

Longitudinal Fluctuations in Relativistic Heavy Ion Collisions

by
Raquel Quishpe

A dissertation submitted to the Department of Physics,
College of Natural Sciences and Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Physics

Chair of Committee: Anthony Timmins

Committee Member: Rene Bellwied

Committee Member: Lisa Koerner

Committee Member: Greg Morrison

Committee Member: Ioannis Pavlidis

University of Houston
May 2022



Copyright 2022, Raquel Quishpe

DEDICATION

I dedicate this dissertation work to my parents, Lucia and Luis Quishpe. It is thanks to their hard work throughout my life that I was able to have the opportunity to pursue my dreams.

ACKNOWLEDGMENTS

During my PhD I had the opportunity to meet new people and meet again some old mates to whom I am grateful for sharing this journey with me. First of all, I would like to thank Anthony Timmins, my PhD advisor, who introduced me to the ALICE Collaboration, and supported me by giving me the opportunity to work on different projects throughout my career. Among those projects is the Hyperloop Train System, lead by Jan Fiete Grosse-Oetringhaus, to whom I am extremely grateful for making me part of this project, thanks to this I was able to enjoy the dynamics of working in a big experiment like CERN and earned a lot of new software skills. I would like to also thank the Hyperloop team for the great discussion and team work, Armenuhi Abramyan, Raluca Cruceru and Costin Grigoras. Moreover, during the time I was at CERN, I became part of the ALICE Service Work Board as a Junior representative, I would like to thank Kai Schweda and the other Service Work Board members for letting me contribute in the decisions made and for guiding me throughout the process.

Furthermore, I would like to thank Rene Bellwied and Claudia Ratti, for their guidance and warm welcome to their research groups as well as to their lives. Thanks to the other members of my PhD dissertation committee as well, Lisa Koerner, Greg Morrison and Ioannis Pavlidis, for their constructive critic that shaped this dissertation work.

I want to also thank my Physics PhD colleagues, specially, as other fellows used to call us, the “dream team”, Vidushi Adlakha, Geethal Amila, Yuli Chang and Rauf Giwa, thanks for becoming like a family while we were all far away from home. In addition, I would like to thank my Ecuadorian friends who joined me in the US, Joaquin Grefa, Pablo Lopez, Evelyn Yanez and Wilson Yanez, thanks for bringing a piece of home to me. Finally, I would like to thank my parents, Luis and Lucia Quishpe, and sister, Mercy Quishpe, for all the constant love and support. Also, a special thank to my fiancé, Loïc Betemps, for his constant support during these last years of my PhD.

ABSTRACT

Collisions of relativistic heavy-ions aim to create a unique state of matter, known as the Quark Gluon Plasma (QGP), where quarks and gluons are in a deconfined state. So far, one of the most suitable probes to study the properties of the QGP are the flow coefficients ν_n , which quantify the anisotropic particle distributions in the final state, and that can be obtained via Fourier decomposition. These flow coefficients have been extensively studied. Moreover, it has been seen that the pseudorapidity of the final particle distribution in a heavy ion collision is sensitive to event-by-event longitudinal fluctuations, and the shape can be decomposed using Legendre polynomials, analogous to the Fourier decomposition in anisotropic flow. A significant asymmetry is observed, and can be quantified with the first order coefficient, a_1 , from the longitudinal decomposition. The presence of a longitudinal asymmetry in nuclei-nuclei collisions does not agree with the boost invariance assumed in various models that describe the initial state of a relativistic heavy ion collision. The description of the asymmetry can reveal nontrivial information about the structure of the fireball in the longitudinal direction and hence can be used as a probe of the initial conditions of the evolution of the system. In this thesis, the longitudinal decomposition for Xe-Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV and Pb-Pb collisions $\sqrt{s_{NN}} = 5.02$ TeV from the LHC Run 2 data measured with the ALICE detector is presented and compared to models.

Furthermore, during the Long Shutdown 2 (LS2), the ALICE detector went through several upgrades that allows it to collect in the LHC Run 3 and Run 4 about two order of magnitude more data compared to the LHC Run 2 and Run 1. The data collection rate increase required the development of a new software in ALICE, the Online-Offline Computing System (O²), along which a new organized analysis system was developed, the Hyperloop train system. This thesis will also present the concept and successful development of the Hyperloop train system that has been actively used since June 2020 in ALICE.

Contents

DEDICATION	iii
ACKNOWLEDGMENTS	iv
ABSTRACT	v
LIST OF FIGURES	xi
LIST OF TABLES	xii
1 INTRODUCTION	1
1.1 The Standard Model	2
1.2 The Strong Force and Quantum Chromodynamics	3
1.2.1 QCD at small distances	3
1.2.2 QCD at large distances	4
1.3 Lattice QCD	5
1.4 Creating a QGP using Relativistic Heavy Ion Collisions	6
1.5 The QCD phase diagram	9
1.6 Probes of the Quark Gluon Plasma (QGP)	9
1.6.1 Anisotropic Flow	12
1.7 Outline	14
2 LONGITUDINAL FLUCTUATIONS IN A HEAVY-ION COLLISION	17
2.1 Kinematic variables	17
2.1.1 Geometry	17
2.1.2 Impact parameter and centrality	17
2.2 Overview of theoretical models in HIC	19
2.2.1 Color Glass Condensate (CGC)	19
2.2.2 Monte Carlo Glauber	20
2.2.3 IP-Glasma	21
2.2.4 Heavy Ion Jet Interaction Generator (HIJING)	21
2.2.5 A Multi-Phase Transport Model (AMPT)	22
2.3 Previous results on longitudinal fluctuations	22
3 HIGH ENERGY PHYSICS EXPERIMENTS	29
3.1 Large Hadron Collider (LHC)	29
3.1.1 LHC Timeline	29
3.2 A Large Ion Collider Experiment (ALICE)	30
3.2.1 Time Projection Chamber (TPC)	31
3.2.2 Inner Tracking System (ITS)	34
3.2.3 VZERO (V0)	36
4 ALICE TRAIN SYSTEM	37
4.1 ALICE Computing Tools	37
4.2 The LEGO Train System	37

4.2.1	LEGO Web Interface	38
4.3	From the LEGO trains to Hyperloop trains	40
4.4	The Hyperloop Train System	41
4.4.1	Bookkeeping and comparison	47
4.4.2	Automatic Composition	49
4.4.3	Analysis Facility Staging	50
4.5	Hyperloop - current status	50
5	ANALYSIS METHOD	51
5.1	DATA PROCESSING AND EVENT SELECTION	52
5.1.1	Multiplicity and Centrality	54
5.1.2	Background and pileup rejection	55
5.1.3	Detector non-uniformity	57
5.1.4	Systematic uncertainties	57
5.1.5	Non-flow contributions	60
6	RESULTS	61
6.1	Background method comparison	61
6.2	Systematic checks	62
6.3	Non-flow contributions	74
6.4	Comparison to Monte Carlo data	79
7	CONCLUSION AND OUTLOOK	86
A	Overwatch	89
B	Bibliography	90
	BIBLIOGRAPHY	91

List of Figures

1	The Standard Model. Figure taken from [standardmodelref].	2
2	Normalized pressure, energy density, and entropy density as a function of the temperature. The gray points are from the HotQCD collaboration [lattice2014]. The colored points are from the Wuppertal–Budapest Collaboration [BORSANYI201499]. Figure taken from [ratti2018].	5
3	Evolution of a relativistic heavy ion collision. Figure taken from [hicevolutionfig]. .	7
4	Schematic QCD phase diagram for nuclear matter. The green band represents the result for the crossover transition, in the μ_B region accessible to lattice QCD calculations [qcdphase2020]. The open points show experimental results for the determination of the chemical freeze-out parameters [qcdphase2018][qcdphase2006][qcdphase2021]. The location of atomic nuclei is also shown (nuclear matter), as well as conjectured regions for the presence of a first order phase transition and of a critical point. Figure taken from [hicevolutionfig].	10
5	Schematic 3D-view of a peripheral heavy ion collision. (Left) Incoming heavy-ion nuclei accelerated along the z -axis. (Right) Almond shape like overlap region after the collision.	13
6	Decomposition of one initial condition into its first 4 harmonic deformations. The Φ_n are the participant plane angles for each deformation relative to the impact-parameter direction. Figure taken from [qiu2013eventbyevent].	14
7	Flow coefficients for A–A collisions measured with the ALICE detector.	15
8	Schematic cartoon of the geometry of the ALICE detector.	18
9	The impact parameter (b) in a heavy-ion collision. Left, xz view of the incoming nuclei. Right, xy , transverse view of the colliding nuclei.	18
10	A cartoon example of the correlation of the final state observable N_{ch} with Glauber calculated quantities (b , N_{part}). Figure taken from [mcglauber2007].	19
11	Gauge field configurations in the form of “flux tubes” of longitudinal chromo-electric and chromo-magnetic fields screened on transverse distance. Figure taken from [Iancu:2012xa].	20
12	Initial energy density (arbitrary units) in the transverse plane for heavy-ion collision events produced with different models. Figures taken from [kln2012].	21
13	Schematic longitudinal particle distribution at early stages of the collision.	24
14	LRC strength measured with STAR on 2009. Forward-backward correlation strength for 0-10% most central Au–Au collisions and p–p from data. HIJING and PSM predictions are shown as well. Figure taken from [lrc2009].	25
15	First order longitudinal coefficients from HIJING and AMPT events calculated by the ATLAS Collaboration.	27
16	Magnitude from short range correlations, SRC (left), and long range correlations, a_1 (right), as a function of the final multiplicity N_{ch} . Figure taken from [jia2017]. . . .	28
17	Diagram of the LHC. Figure taken from [Stark2020].	30
18	LHC luminosity timeline. Figure adapted from [Stark2020].	30
19	Diagram of the ALICE detector. Figure taken from [alicefig2014]	31
20	3D view of the TPC field cage and service support wheels, without readout chambers. Figure taken from [alicetech2008].	32

21	TPC dE/dx measurements in the full 2011 Pb–Pb statistics. Figure taken from [aliceloi2014].	33
22	Layout of the ITS. It is composed of three detectors, each of them consists of two layers: the Silicon Pixel Detector (SPD) closely surrounding the beam pipe, the Silicon Drift Detector (SDD), and the Silicon Strip Detector (SSD). Figure taken from [alicetech2008].	34
23	Layout of the ITS for Run 3. It consists of 3 Inner Barrel layers (IB) and 4 Outer Barrel layers (OB). Figure taken from [Musa:1431539].	35
24	Sketches of VZERO-A and VZERO-C arrays showing their segmentation. Scintillator thicknesses are 2.5 and 2 cm respectively. The scintillator segments on both sides of the dashed lines are connected to the same photomultipliers. Figure taken from [vzero2013].	36
25	Procedure of a LEGO train run. After a successful train test, the analysis jobs run. Then the results are merged into a single output file.	39
26	Performance of CF Pb–Pb web page (by 2019).	40
27	Example of Hyperloop wagon configuration. O^2 parameters are supported and displayed in a user-interactive interface. From top to bottom the sample parameters type displayed are array of strings, integer, matrix of double numbers with columns and rows labels, matrix of double numbers, histogram binning with fixed width, histogram binning with variable width.	41
28	Procedure of a Hyperloop train run. Wagon tests are immediate and automatic after the user enables a train. A train run can be composed by a train operator, or by the system. Figure taken from [quishpe2021hyperloop].	43
29	Test statistics of a wagon for an active dataset. (Top left) Proportional set size memory vs. test ID. (Top right) CPU/Wall time and Throughput vs. test ID. (Bottom left) Output size vs. test ID.	44
30	Test graphs from train run test. Table-bytes-per-table vs. time. Graphs can also be displayed per device.	45
31	Grid statistics. Shows overall jobs performance. (Center left) Number of jobs vs. Files per job. (Center right) Number of jobs vs. CPU time per job. (Bottom left) Number of jobs vs. wall time per job.	46
32	Wagon comparison. Comparison of wagon configurations corresponding to two different timestamps. One timestamp is on cyan, and the other green. The common parameters are displayed on gray.	47
33	Train run comparison. Comparison of two different train run configurations. One train run is on cyan, and the other green. The common parameters are displayed on gray. Direct links to the specific timestamp dataset and wagons are available.	48
34	Wagon history. Detailed change log of a wagon configuration, including the user name and time at which a change was made. On the right, by clicking on two check boxes, a user can compare two different timestamps configuration from the same wagon.	49
35	Distribution of the sum of amplitudes in the V0 scintillators for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The distribution is fitted with the NBD-Glauber fit shown as a line. The inset shows a zoom of the most peripheral region. Figure taken from [centALICE:2018tvk].	55

36	Distribution of the sum of amplitudes in the V0 scintillators for Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. The distribution is fitted with the NBD-Glauber fit shown as a line. The inset shows a zoom of the most peripheral region. Figure taken from [centALICE:2018tvk].	56
37	Primary vertex irregularity in Xe–Xe data	58
38	First order longitudinal coefficients, a_1 in Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. Comparison of different backgrounds. (Red) a_1 calculated using a random uniform background. (Blue) a_1 calculated using random pseudorapidity distribution based background.	61
39	First order longitudinal coefficients, a_1 in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. Period LHC18r. Comparison of different backgrounds. (Red) a_1 calculated using a random uniform background. (Blue) a_1 calculated using random pseudorapidity distribution based background.	62
40	Comparison to different systematic uncertainties in Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV in a pseudorapidity window $ \eta < 0.8$	63
41	Comparison to different systematic uncertainties in Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV in a pseudorapidity window $ \eta < 0.6$	64
42	Comparison to different systematic uncertainties in Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV in a pseudorapidity window $ \eta < 0.4$	65
43	Comparison to different systematic uncertainties in Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV in a pseudorapidity window $ \eta < 0.2$	66
44	Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $ \eta < 0.8$	68
45	Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $ \eta < 0.6$	69
46	Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $ \eta < 0.4$	70
47	Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $ \eta < 0.2$	71
48	First longitudinal coefficient a_1 for Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. Includes systematic (box) and statistical (line) errors.	72
49	First longitudinal coefficient a_1 for Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, period LHC18r. Includes systematic (box) and statistical errors (line).	72
50	a_1 vs. $\langle dN_{ch}/d\eta \rangle$ in different pseudorapidity windos for Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV and Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. Includes systematic (box) and statistical (line) errors.	73
51	First longitudinal coefficient a_1 for Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV, period LHC17n. Comparison among charged, positive and negative tracks. Includes systematic (box) and statistical errors (line).	75
52	First longitudinal coefficient a_1 for Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, period LHC18r. Comparison among charged, positive and negative tracks. Includes systematic (box) and statistical errors (line).	76
53	Comparison of different momenta ranges [GeV/c] in Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV, period LHC17n.	77

54	Comparison of different momenta ranges [GeV/c] in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, period LHC18r.	78
55	a_1 from Xe–Xe MC HIJING data in different pseudorapidity windows. Comparison of generated vs. reconstructed charged tracks.	79
56	(a)(b)(c) a_1 from Xe–Xe MC Hijing data in different pseudorapidity windows for charged, positive and negative tracks. (d) Comparison among charged tracks in different pseudorapidity windows.	80
57	a_1 from Xe–Xe raw data and MC Hijing in different pseudorapidity windows. Comparison of data vs. MC charged tracks.	81
58	a_1 from Pb–Pb MC Hijing data in different pseudorapidity windows. Comparison of generated vs. reconstructed charged tracks.	82
59	(a)(b)(c) a_1 from Pb–Pb MC Hijing data in different pseudorapidity windows for charged, positive and negative tracks. (d) Comparison among charged tracks in different pseudorapidity windows.	83
60	a_1 from Pb–Pb raw data and MC Hijing in different pseudorapidity windows. Comparison of data vs. MC charged tracks.	84
61	a_1 from Pb–Pb and Xe–Xe raw data and MC HIJING in different pseudorapidity windows. Statistical (line) and systematic (box) errors included.	85
62	Overwatch architecture. Figure taken from [raymond’ehlers’2018’1309376].	89
63	Overwatch WebApp. Figure taken from [raymond’ehlers’2018’1309376].	90

List of Tables

1	Usage statistics of the LEGO train system. Data for 2012 and 2014 taken from [zimmermann2015alice].	39
2	Run 2 raw and MC data used for this thesis.	53
3	Centrality classes for Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV and Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. Data taken from [centpb2016][centxe2019][centALICE:2018tvk].	55
4	Summary of systematic checks applied for Xe–Xe raw data and MC in this thesis. .	59
5	Summary of systematic checks applied for Pb–Pb raw data and MC in this thesis. .	59
6	Summary of systematic error contribution for Xe–Xe collisions in different pseudo-rapidity windows.	67
7	Summary of systematic error contribution for Pb–Pb collisions in different pseudo-rapidity windows.	67

1 Introduction

Describing what the Universe is made of and how it works has been one of the main queries in our history. Throughout the years, most of the main basic components of matter and their interactions have been successfully described in a theoretical model, the Standard Model. Among those basic components we have quarks and gluons, which cannot be found by themselves in nature under normal circumstances. Quarks and gluons are held together by the strong force in the form of hadrons, and it has been observed that the only way to see them by themselves is in a very hot and dense state of matter known as the Quark Gluon Plasma (QGP). For the QGP to happen, extreme temperatures are required, about 100 000 times the temperature of the core of the Sun. In nature, this extreme temperature has only been seen a few microseconds after the Big Bang, nearly 14 billion years ago. To reproduce the conditions needed for the QGP, relativistic heavy ion collisions have been created in large experiments located at CERN (European Organization for Nuclear Research) and BNL (Brookhaven National Laboratory).

Several analyses have been performed in order to investigate the properties of the QGP via relativistic viscous hydrodynamics. Anisotropic flow and longitudinal fluctuations in heavy-ion collisions have been seen as signatures sensitive to the formation of the QGP. Furthermore, theoretical models integrating these concepts have been developed aiming to constrain thermodynamic properties of the initial state of the collision, such as the shear viscosity and density fluctuations.

Throughout this chapter, basic concepts are described such as the Standard Model, the theory that describes the strong interaction (Quantum Chromodynamics, QCD), the Quark Gluon Plasma (QGP) and the evolution of a relativistic heavy-ion collision. Finally, an overview of the aim of this thesis is presented.

1.1 The Standard Model

Matter interacts through four forces: electromagnetic, strong, weak and gravity. With the exception of the gravitational force, the Standard Model (SM) is a theory that successfully describes most of existing particles and their interactions. In the SM, the particles are divided in two types, leptons and quarks, each of them containing six particles and their corresponding antiparticles. Quarks make up hadrons, like protons and neutrons, and are identified based on their flavor: up, down, charm, strange, top and bottom. Likewise, the leptons are the electron and electron neutrino, muon and muon neutrino, and tau and tau neutrino. Particle interactions can take place thanks to particle mediators, known as Gauge bosons, i.e. the photons for the electromagnetic force, the $W^\pm$ and Z bosons for the weak force, gluons for the strong force, and the Higgs boson as the responsible for giving mass the particles mentioned above. Figure 1 summarizes the properties of quarks, leptons and Gauge bosons in the SM.

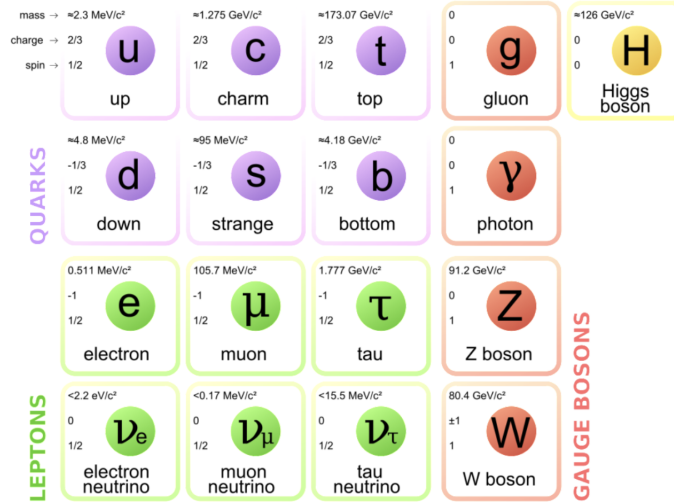


Figure 1: The Standard Model. Figure taken from [standardmodelref].

1.2 The Strong Force and Quantum Chromodynamics

The strong force is responsible for holding the atomic nuclei together and is described through a non-Abelian gauge field theory known as Quantum Chromodynamics (QCD). As its name states, the feature that it describes is the color charge (an analog to the electric charge from Quantum Electrodynamics, QED). The three color charges in QCD are red, green and blue. Each quark can only carry one color charge, and gluons can carry two color charges. Since gluons possess color charge, they can interact among themselves, which is not the case for photons in QED.

At small distances the non-Abelian QCD potential is assumed to be of the Coulomb type, whereas at large distances it must increase indefinitely so that quarks are confined inside a hadron. Hence, the potential for the interaction between two quarks can be simplified as,

$$V(r) = -\frac{4\alpha_s}{3r} + kr \quad (1)$$

where r is the distance between the two quarks, α_s is the running coupling constant, and k is a string constant whose value is $\approx 0.85 \text{ GeV fm}^{-1}$ [perkins'2000]. The first term of equation 1 dominates at small values of r , which arises from single-gluon exchange. The second term, is associated to confinement at large values of r .

1.2.1 QCD at small distances

The running coupling constant in equation 1 depends on the four momentum transfer (q^2), and from first-order perturbative QCD calculation it can be approximated as,

$$\alpha_s(q^2) = \frac{4\pi}{\beta_1 \ln(q^2/\Lambda^2)} \quad , \quad \Lambda = q_0 e^{-2\pi/\beta_1 \alpha_s(q_0^2)} \quad (2)$$

$\alpha_s(q_0^2)$ is the boundary of the solution, with q_0 being the initial four momentum transfer; and $\beta_1 = (11N_c - 2n_f)/3$ where n_f is the number of flavors and N_c the number of colors, in QCD 6 and 3 respectively, giving $\beta_1 = 5$. In the literature, Λ is known as the QCD scale constant, an

infrared cutoff, and it has a value of 0.2 GeV [**lambdaqcd2000**]. A high-momentum transfer, i.e. $q^2 \gg \Lambda_{QCD}$, involves a short-range interaction, and is referred as hard processes. As α_s tends to zero, so does the potential in equation 1, and hence it allows the use of perturbation methods to calculate physical quantities like cross sections at high energies, i.e. the perturbative regime corresponds to $\alpha_s \ll 1$. Furthermore, at very small distances and high values of q^2 , α_s decreases, vanishing asymptotically, and in the limit of very large q^2 , quarks can be considered to be “free”, also known as asymptotic freedom.

Asymptotic freedom causes interactions between particles to become asymptotically weaker as the energy increases, i.e. small distances. This leads to quarks and gluons no longer being confined inside hadrons, but to be almost free in a new hot and dense state of matter known as the Quark Gloun Plasma (QGP). The QGP can only be achieved under extreme conditions. In nature, the QGP has only been present a few microseconds during the creation of the Universe at the Big Bang, about 14 billion years ago. Nowadays, it can be reproduced via ultra-relativistic heavy ion collisions; however, understanding how the system evolves, is still yet to be fully explained.

1.2.2 QCD at large distances

At large distances ($\approx$ more than 1 fm), the second term in equation 1 dominates and becomes linearly proportional to the distance, which supports the idea that gluons carry color charge, and can thus interact with themselves. Moreover, at large distances, α_s increases making it impossible to detach quarks from hadrons, also known as confinement. As a result of color confinement, quarks and gluons are never isolated under ordinary circumstances; they are found in nature in the form of hadrons. Hadrons are made of two or more quarks, where the total color charge is neutral, they are divided in mesons and baryons. Mesons are made of an even number of quarks, usually quark-antiquark pairs where the neutral color charge comes from having a color and the corresponding anticolor charge, e.g. pions. Baryons are made of an odd number of quarks, usually three quarks, where in order to have a neutral color charge, it requires a combination of red, blue and green

color charge resulting into a color neutral charge, similar to a light beam. Examples of baryons are protons and neutrons.

1.3 Lattice QCD

In the non-perturbative regime, QCD needs to be solved numerically. Simulations have been performed by placing quarks and gluons on a discretized lattice and simulating their interactions in thermal equilibrium. This is known as Lattice QCD. Computational numerical methods are used in order to predict thermodynamic properties of hadronic matter from a quantitative perspective, like temperature and energy density. Figure 2 shows the extrapolation to the continuum of lattice simulations of (2+1)-flavor QCD from the HotQCD and Wuppertal–Budapest (WB) Collaborations, as well as the temperature dependence of other thermodynamical quantities such as pressure (p) and entropy density (s) [lattice2014]. The figure also shows the Stefan-Boltzmann (SB) limit for the pressure and the scaled entropy.

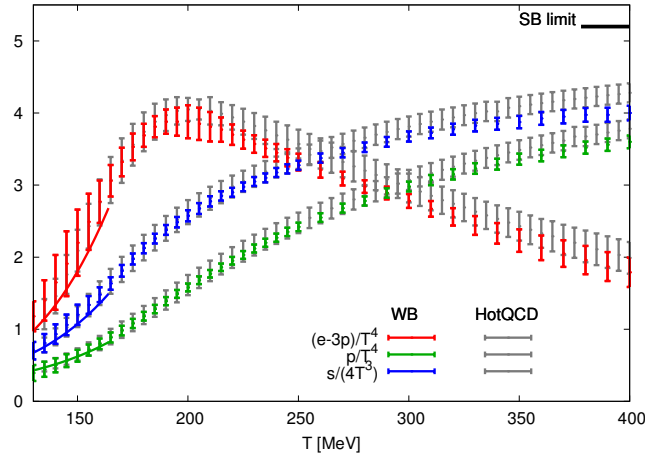


Figure 2: Normalized pressure, energy density, and entropy density as a function of the temperature. The gray points are from the HotQCD collaboration [lattice2014]. The colored points are from the Wuppertal–Budapest Collaboration [BORSANYI201499]. Figure taken from [ratti2018].

At sufficiently high temperatures, thermodynamics may be described in terms of a weakly interacting quark-gluon gas, and at infinite temperature all thermodynamic quantities will converge to the ideal gas limit. One of the most relevant quantities to describe in a system is its Equation of State (EoS). In the case of zero net-baryonic density ($\mu_B = 0$), which is the case in the LHC, the EoS of QCD is known with high precision, and the different approaches of the two collaborations HotQCD and WB showing good agreement. However, the extrapolation shown in Figure 2 shows that at the highest temperature value, $T = 400$ MeV, the entropy, energy density and pressure are still 13%, 18% and 27%, respectively, below the ideal gas limit, which should have been reached at very high T due to the asymptotic freedom of QCD. The latter suggests a possible large non-perturbative contribution in the gluonic sector of QCD present in the bulk thermodynamic observables, which so far were suppressed as the calculations done are only in the quark sector. Moreover, the simulations predict that the phase transition from hadronic degrees of freedom to the QGP is a broad analytical crossover, which happens gradually in the temperature range $T \simeq 145 - 165$ MeV[ratti2018]. Due to the nature of the transition, it is not possible to unambiguously define a transition temperature, and different observables yield slightly different values.

1.4 Creating a QGP using Relativistic Heavy Ion Collisions

In order to reproduce a QGP, it is necessary to arrive to extremely high energy densities and temperature, about 100 000 times the temperature of the core of the Sun. These extreme conditions can be fulfilled by colliding two heavy atomic nuclei accelerated to energies of the order of 10 GeV and upwards, i.e. they travel at a speed close to the speed of light. High energy colliders such as Relativistic Heavy-Ion Collider (RHIC) at Brookhaven National Laboratory (BNL) and the Large Hadron Collider (LHC) at CERN (European Organization for Nuclear Research), have accelerated gold and lead nuclei respectively.

As explained in [PhysRevD.27.140], since the heavy atomic nuclei travel at a speed close to the speed of light, the bunches of particles traveling are highly Lorentz contracted, and can be

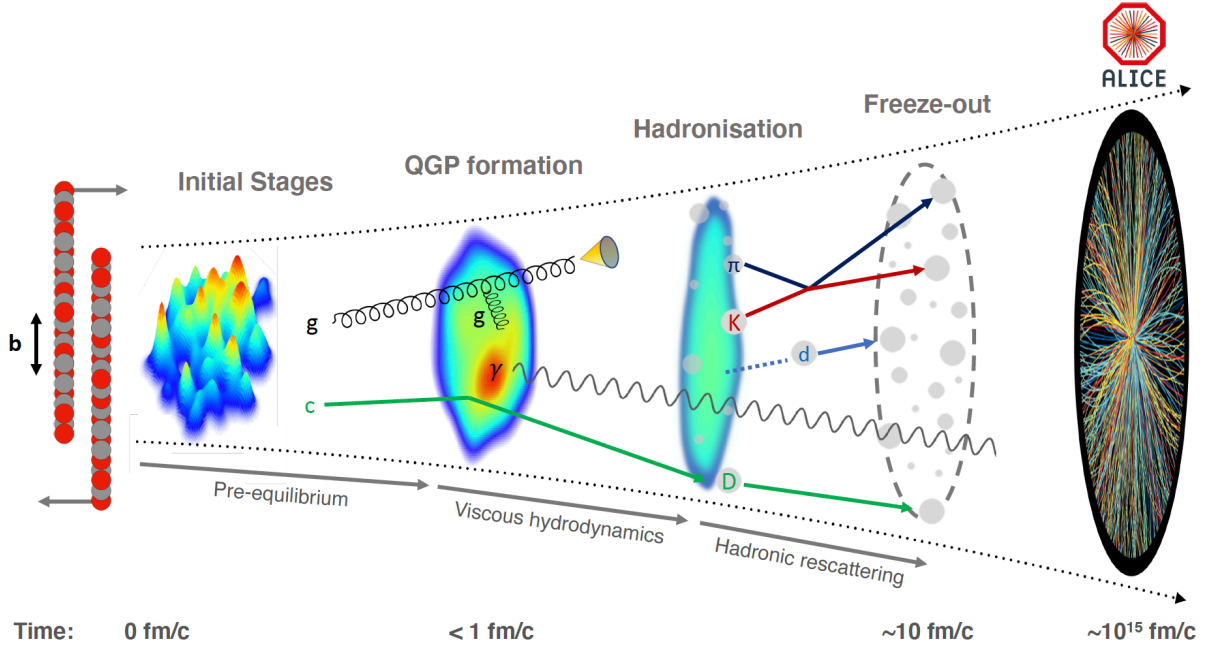


Figure 3: Evolution of a relativistic heavy ion collision. Figure taken from [hicevolutionfig].

considered as flat pancakes. Hence, it can be assumed that the expansion takes place mainly in the longitudinal direction. The latter allows to define time in terms of the proper time $\tau \equiv \sqrt{t^2 - z^2}$. Figure 3 displays the evolution of a relativistic heavy ion collision, which can be divided in the following phases [Iancu:2012xa]:

- Before the collision takes place, the Lorentz contracted pancakes are mostly composed of gluons, which carry a tiny fraction of longitudinal momenta of their parent nucleons ($x \ll 1$), and with a density that increases rapidly as $1/x$. Due to the uncertainty principle, gluons making up this high-density system carry relatively large transverse momenta. Following QCD asymptotic freedom, the Gauge coupling in these gluon interactions is relatively weak.
- At the beginning of the collision, the incoming nuclei hit each other, and the interactions take place. Hard processes, i.e. transfer of large momenta, take place and are responsible of the production of hard particles like jets or heavy quarks.

- In the pre-equilibrium phase, most of the multiplicity that will be present in the final state is generated, i.e. most of the hadrons that are seen in the detectors at the last stage are produced via the fragmentation and hadronisation of the initial state gluons liberated at this stage. Right after being liberated, they form a relatively dense medium, e.g. Pb+Pb collisions at the LHC have approximately an average energy density of $\varepsilon > 15 \text{ GeV/fm}^3$ - 15 times higher than the density of a nucleon.
- If the resulting partons did not interact with each other, or if the interactions are negligible, the partons will rapidly separate and independently evolve towards the final state hadrons via fragmentation and hadronisation. However, at RHIC and the LHC, heavy ion collisions have shown collective phenomena, which demonstrates that the partons actually did interact significantly before they evolved towards the final state hadrons. Hence, the dense partonic matter produced in the intermediate stages of a heavy ion collision may be a strongly coupled fluid, as in order to compensate the rapid expansion it required strong interactions to take place in such a short thermalization time ($\tau \sim 1 \text{ fm/c}$).
- The outcome of this thermalization is a hot and dense state, the QGP. However, experimentally, this state does not last long enough, as the partonic matter keeps expanding, cooling down and eventually hadronising. The temperature is space and time dependent, so the equilibrium is reached only locally. Hadronisation takes place when the local temperatures become of the order of the critical temperature T_c for deconfinement, which is $T_c \simeq 150/180 \text{ MeV}$ from QCD studies. At the LHC, Pb+Pb collisions this happens around $\tau \sim 10 \text{ fm/c}$.
- As the system continues to expand and cool down, hadrons continue to interact with each other. The point where the inelastic processes stop and the chemical (hadronic) composition of the system does not change anymore is known as the chemical freeze-out. This occurs almost at the same time as the system transitions to a phase of a hot hadron gas.
- At $\tau \sim 20 \text{ fm/c}$, the collision rate becomes smaller than the expansion rate, i.e. the density has decreased so much that the hadrons no longer interact with each other. The transition

from a fluid state to a system of free particles is known as the kinetic freeze-out. The free particles fly into the vacuum and are eventually detected by experiments.

1.5 The QCD phase diagram

As mentioned before, during the intermediate stages of a heavy ion collision, the created hot and dense matter undergoes a phase transition from a strongly coupled fluid to a hadronic gas. This process can be represented by a phase diagram, as in Figure 4. This phase diagram is depicted in terms of the so called baryon chemical potential (μ_B) in the x -axis and the temperature in the y -axis. The baryon chemical potential, μ_B , refers to the difference of baryons and antibaryons present, i.e. matter minus anti-matter. For instance, $\mu_B = 0$ means that there is the same amount of baryons as antibaryons, and $\mu_B > 0$ means that baryons dominate over anti-baryons. Heavy ion collisions at high energies, like the ones produced in RHIC or the LHC, correspond to a low μ_B , this happens because even if there is an excess of baryons, this number becomes negligible due to the excessive number of newly created particles which have both, matter and antimatter. On the other hand, collisions at lower energies, correspond to larger values of μ_B as the original number of nucleons brought into the collision becomes significant.

At a high μ_B and low temperature, a first order phase transition may happen between the QGP and the phase of confined hadron gas, while a cross-over is believed to mark a boundary between the hadron gas and the QGP at low μ_B and high temperature. Furthermore, the hadron gas undergoes a phase transition to the QGP when sufficient energy density and temperature are reached. Lattice QCD estimates this value to be around $\varepsilon \approx 0.18 - 0.5 \text{ GeV/fm}^3$ and $T_c \approx 145 - 165 \text{ MeV}$ [lattice2014]. Characterizing this phase transition remains as a constant research topic in the field.

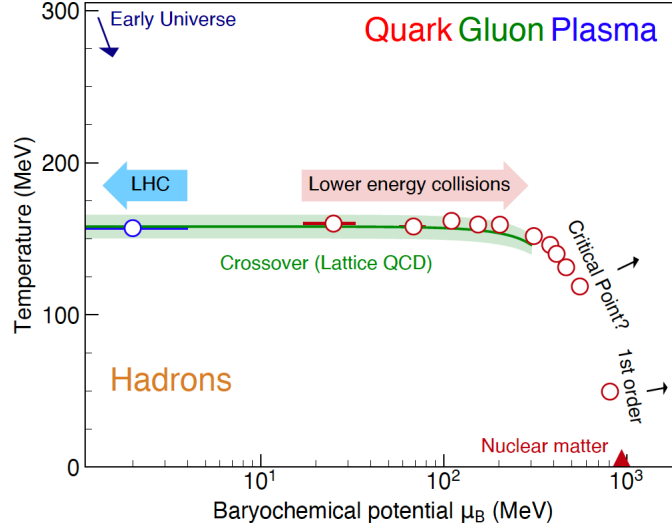


Figure 4: Schematic QCD phase diagram for nuclear matter. The green band represents the result for the crossover transition, in the μ_B region accessible to lattice QCD calculations [qcdphase2020]. The open points show experimental results for the determination of the chemical freeze-out parameters [qcdphase2018][qcdphase2006][qcdphase2021]. The location of atomic nuclei is also shown (nuclear matter), as well as conjectured regions for the presence of a first order phase transition and of a critical point. Figure taken from [hicevolutionfig].

1.6 Probes of the Quark Gluon Plasma (QGP)

It has been experimentally observed that near the critical temperature, the QGP behaves more like a nearly perfect fluid than like a free gas of quarks and gluons. Relativistic hydrodynamics has been used to model the QGP evolution and has helped in the understanding of particle spectra, flow and correlations.

As the system created in a heavy-ion collision expands, approximations can be done considering a local thermal equilibrium. Hydrodynamics provides an effective macroscopic description of this approximation, and can be derived from the underlying microscopic (kinetic) description through an expansion in gradients of the local thermodynamic variables, where the zeroth order represents an ideal fluid. Navier–Stokes theory accounts for terms that are linear in gradients which describe momentum and baryon number transport across fluid cells due to bulk and shear viscosity as well

as heat conduction. However, Relativistic Navier–Stokes theory is acausal [**HISCOCK1983466**], i.e. superluminal signal propagation causes short-wavelength perturbations to grow exponentially, leading to physical and numerical instabilities. As a consequence, it is necessary to include second-order gradient terms, which can give physical transport quantities, such as viscosities and the heat conductivity. One of the most widely used theoretical frameworks that deals with this regularization is the Israel–Stewart theory [**doi:10.1146/annurev-nucl-102212-170540**].

For a system with $\mu_B \sim 0$, like the collisions that take place at the LHC and top RHIC energies, the heat conduction effects can be neglected. This leaves the bulk(ζ) and shear(η) viscosities, and the relaxation times τ_ζ and τ_η for the bulk and shear viscous pressures as the main transport coefficients controlling the collective dynamics. These are then combined with the entropy density s and the temperature T , resulting in the specific viscosities η/s and ζ/s . These specific viscosities can be constrained by experimental data from heavy ion-collisions, specially η/s . The shear viscosity washes out small structures in the initial state anisotropies, reducing the observed final momentum anisotropy in terms of flow coefficients (see Section 1.6.1). In theoretical approaches such as the anti-de Sitter/conformal field theory (AdS/CFT), the strong-coupling limits for the specific shear and bulk viscosities have been found to be $\eta/s \approx 1/(4\pi)$ [**shear2001**] and $\zeta/\eta \approx 2(1/3 - c_s^2)$ [**BUCHEL2008286**].

Collective phenomena have been observed for heavy ion collisions at RHIC and the LHC, which are related to the formation of a strongly interacting QGP and can be described using relativistic hydrodynamics. In the final particle distribution, this collective behavior can be manifested from significant particle correlations. The analysis of the final particle distribution is crucial in order to understand the transport properties of the QGP. The geometry of each collision depends on the so called impact parameter(b) which refers to the distance between the centers of the two nuclei in the transverse plane, and the event-by-event fluctuations. The observables sensitive to geometry provide a powerful tool to study the QGP. So far, the experimental data has corroborated that the

shear viscosity value $\eta/s \approx 0.08 - 0.20$ is close to the boundary mentioned before. Furthermore, it has been seen that the shear viscosity can depend on the temperature of the QGP, which stands as in RHIC smaller values of η/s have been measured, compared to those at the LHC[**shear2011**]. On the other hand, the bulk viscosity ζ becomes non-negligible around the phase transition region, where $\zeta/s \approx 0.3$ or larger around T_c [**bulk2015**]. After local equilibrium, the hydrodynamic approach is no longer valid; instead, the evolution of the system requires simulations that describe the hadron cascade around T_c from the QGP to the hadron gas.

The probes of the QGP can be divided into two categories, hard and soft probes. Hard probes refer to high momentum transfer, like modification of the jets in the medium. High- p_T partons are created at very early stages of the collisions during hard scatterings, and then fragment into di-jets. The hard scattered partons and the subsequent parton shower interact strongly with the QGP and its constituents via elastic and radiative processes before hadronization into jets of observable particles. Energy is lost from the jet cone as jets interact with the QGP. Hence, the comparison to jet and high- p_T observables in vacuum (proton-proton collisions) can probe the modification due to medium effects.

On the other hand, soft probes refer to low momentum transfer. Due to the geometry of the colliding system, one of the soft probes that has been extensively analyzed over the years has been anisotropic flow. Thanks to the extensive investigation, an unprecedented level of precision has been achieved in the study of anisotropic flow and its connection to the properties of the QGP and the initial state of the collision. In addition to that, longitudinal fluctuations can also contribute to further constraint properties of the initial state of the collision, this will be described in Section 2.

1.6.1 Anisotropic Flow

The relativistic heavy ions colliding can be considered as highly Lorentz contracted pancake. This fact has lead to an extensive analysis of the transverse plane of the collision, i.e. xy - plane. The collision can be characterized using the impact parameter b , where $b \approx 0$ represents a central collision, $b > 0$ a semi-central collision, and $b \gg 0$ a peripheral collision. In a central collision, almost all the nucleon inside a nucleus participate in the collision, while in a peripheral collision, only a small fraction of nucleons do. As a result, different types of heavy-ion collisions give a different overlapping transverse area. A central collision will give a circular geometry, while a peripheral collision is more like an almond shape, like in Figure 5.

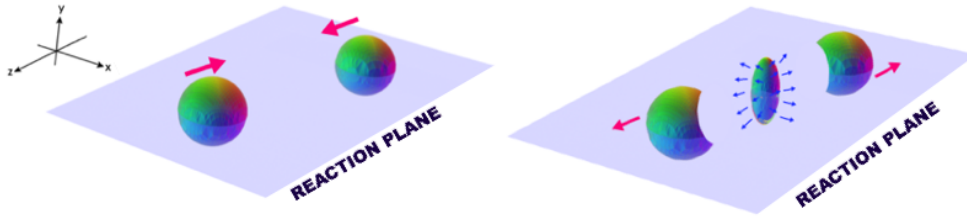


Figure 5: Schematic 3D-view of a peripheral heavy ion collision. (Left) Incoming heavy-ion nuclei accelerated along the z -axis. (Right) Almond shape like overlap region after the collision.

In reality, the almond shape is not a perfect ellipse in the xy -plane. However, it has been seen that the geometry of the distribution can be decomposed via Fourier decomposition. Analyses of this decomposition have determined that only the $\cos$ term of the Fourier decomposition is significant,

$$\frac{dN}{d\Phi} \propto 1 + \sum_{n=1}^{\infty} 2\nu_n \cos(n(\phi - \Psi_n)) \quad (3)$$

where Φ_n is the azimuthal angle of the particles, Ψ_n is the symmetry plane angle of the n -th harmonic, and ν_n are the so known flow coefficients corresponding to a harmonic n . Figure 6 illustrates a cartoon of the first 4 harmonic deformations. Each of these flow coefficients ν_n is sensitive to

the centrality of the collision due to their relation to initial anisotropies. For instance, a peripheral collision is characterized by a pronounced elliptic shape, and hence the second order coefficient ν_2 is significant, also known as elliptic flow. For a central collision, higher order coefficients contribute if the overlap geometry is completely driven by the fluctuations in nucleon positions inside the nuclei. Furthermore, it has been seen that the flow coefficients are directly proportional to hydrodynamic properties of the initial state as $\nu_n \approx \kappa_n \varepsilon_n$, where ε_n is the corresponding eccentricity coefficient governing the shape of the initial state and κ_n describes the response to the medium. In particular κ_n is sensitive to the shear viscosity over entropy density ratio (η/s) and the lifetime of the system.

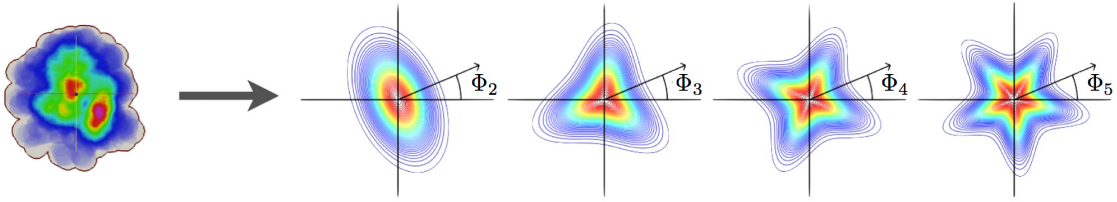
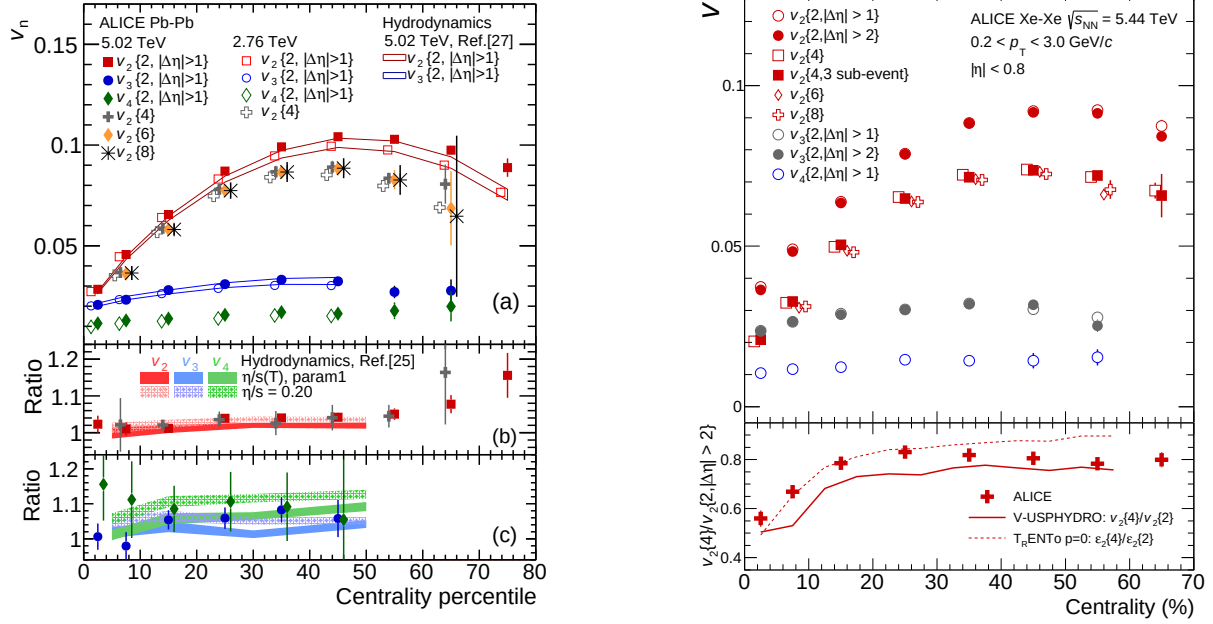


Figure 6: Decomposition of one initial condition into its first 4 harmonic deformations. The Φ_n are the participant plane angles for each deformation relative to the impact-parameter direction. Figure taken from [qiu2013eventbyevent].

At the LHC, the experiment that focuses on the studies of heavy-ion collisions is ALICE. Extensive analyses of different order flow coefficients have taken place. Figure 7 shows results of ν_n for two-particle and multi-particle correlations measured with the ALICE detector for Pb–Pb collisions at 5.02 TeV and 2.76 TeV, and for Xe–Xe collisions at 5.44 TeV. For Pb–Pb collisions, the contribution of elliptic flow, ν_2 , is larger than the one in Xe–Xe collisions. And ν_3 is larger in Xe–Xe collisions than in Pb–Pb. The differences in the magnitude of the flow coefficients for each symmetric system is related to the nature of the collision, i.e. the size of the incoming nuclei. Predictions using hydrodynamical models, like MC Glauber and TRENTo, using different values of the shear viscosity over the entropy ratio (η/s) coincide with experimental measurements (within $\sim 5\%$) [xexeflow2018]. An overall analysis of anisotropic flow for pp, p–pb, Pb–Pb and Xe–Xe

collisions, can be found at [allflow2019]. For the purpose of this thesis, only symmetric A–A collision systems are summarized.



(a) (a) Anisotropic flow ν_n integrated over the p_T range $0.2 < p_T < 5.0$ GeV/c, as a function of event centrality from Pb–Pb collisions. The ratios of different ν_n between Pb–Pb collisions at 5.02 TeV and 2.76 TeV, are presented in (b) and (c). Figure taken from [pbpbflow2016].

(b) Top panel: Charged particle ν_n integrated over the transverse momentum range $0.2 < p_T < 3.0$ GeV/c as a function of centrality from Xe–Xe collisions. Bottom panel: Ratios of $\nu_2\{4\}/\nu_2\{2\}$ compared to some theoretical predictions. Figure taken from [xexeflow2018]

Figure 7: Flow coefficients for A–A collisions measured with the ALICE detector.

1.7 Outline

Throughout the years, extensive analyses aiming to describe the properties of the QGP as well as the initial state have been done. Among the signatures of the QGP, anisotropic flow has been successfully studied via Fourier decomposition, where the coefficients coming from this decomposition, ν_n , described different shapes of the transverse plane of the collision, and these measurements have been linked with thermodynamic properties of the initial state of the colliding system, such

as the shear viscosity over entropy density ratio (η/s) and its connection to the eccentricity described by ν_n . Further constraints to the initial state of a heavy-ion collision may be done, with the study of other QGP signatures. This thesis aims to describe longitudinal fluctuations (z -axis) of a heavy-ion collision by using an analog method to that of anisotropic flow, i.e. decomposing the longitudinal dependent distribution by using Legendre polynomials using the data collected by the ALICE detector. The longitudinal coefficients, a_n , describing this decomposition may provide further constraints on the properties of the system such as an expected asymmetry of the QGP. The latter suggests a break of the boost-invariance at the very beginning of the collision, which contradicts the assumptions done in some of the theoretical models described above.

Section 2 will describe the previous theoretical and experimental results in the study of longitudinal fluctuations, and describe the geometry of a heavy-ion collision. Section 3 will describe the ALICE detector, which is the detector used for the data analyzed throughout this thesis. The data to be analyzed corresponds to the LHC Run 2 period, i.e. data collected in 2015 and 2018. A comparison between the ALICE detector during Run 2 versus the upgrades done for Run 3 (2020) is highlighted. Section 5 shows the analysis method using Legendre polynomials to decompose event-by-event pseudorapidity distributions. Section 5.1 describes the data processing and event selection. Section 6 shows the results of the first longitudinal coefficients, a_n , measured with the ALICE detector. Finally, Section 7 will discuss the results and show the conclusions of this thesis.

Furthermore, there are still several open questions in this field of study, which will not be discussed in detail in this thesis, e.g. the study of heavy-flavour particles, quarkonium states, real and virtual photons, jets and their correlations with other probes [[aliceloi2014](#)]. So far, the analysis of these phenomena has been limited due to the low statistics available in data collected by the high energy detectors. As a result, at the LHC at CERN, since 2018, during the Long Shutdown 2 (LS2), the ALICE detector went through upgrades that allows it to collect two times more data during LHC Run 3 than in Run 1 and Run 2. The high data collection rate, as well as the upgrades

done to the ALICE detector, required the implementation of a new analysis framework, and along with it, the development of a new organized analysis framework for Run 3, known as the Hyperloop train system. These hardware and software upgrades require the contribution of all the members of the ALICE Collaboration, in particular, during my PhD, I have extensively contributed in the software upgrades. Appendix A briefly describes my contribution to the software development of the online data quality monitoring, Overwatch, during 2018. Section 4 provides an extensive explanation to my contribution in the software development of the Hyperloop train system (from 2019 to present).

2 Longitudinal Fluctuations in a Heavy-Ion Collision

2.1 Kinematic variables

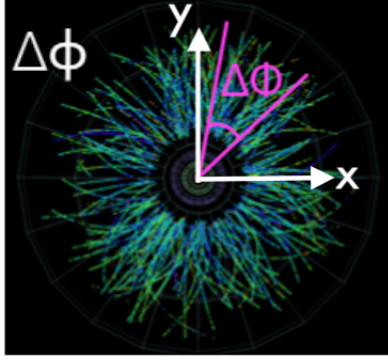
Throughout this thesis, the kinematic variables that will be used are based on the geometry of the heavy-ion colliders, and therefore, it is important to first give an overview of the kinematic terms that are used, such as pseudorapidity, impact parameter, and centrality.

2.1.1 Geometry

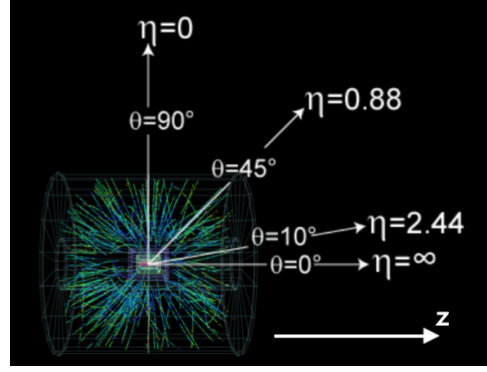
The overall geometry of a high energy detector can be considered as that of a cylinder, where in coordinate system, the z – *axis* is taken to be along the incoming beam axis, and the transverse plane xy perpendicular to it. One of the quantities that is invariant for different observers when the particles are traveling at a speed close to the speed of light, i.e. Lorentz invariant, is the total momentum of the interacting particles. However, this quantity is difficult to directly measure as the energies and momentum fractions of the respective constituents that interact are not known a priori. Instead, the quantities that are used in hadron colliders are the rapidity and pseudorapidity. Rapidity (y) depends on the energy and the total momentum, which is mass dependent. On the other hand, the pseudorapidity only depends on the polar angle θ , and is defined as $\eta = -\ln[\tan(\theta/2)]$. The pseudorapidity, η , is related to the angle with respect to the axis of the colliding beams, i.e. $\eta = 0$ for particle trajectories that are perpendicular to the beam, and positive or negative values for those at an angle to the beam (Figure 8b).

2.1.2 Impact parameter and centrality

In high energy physics, the impact parameter (b) is defined as the distance of closest approach that would occur for the colliding particles if they just followed their initial straight-line trajectories. In particular, in heavy-ion collisions, the bunches of incoming particles are considered to be highly Lorentz contracted pancakes since they travel at a speed close to the speed of light. Hence, the impact parameter refers to the distance between the centers of the two nuclei in the transverse



(a) Diagram of transverse area.



(b) Diagram of pseudorapidity

Figure 8: Schematic cartoon of the geometry of the ALICE detector.

plane. Figure 9 illustrates the impact parameter of a heavy-ion collision.

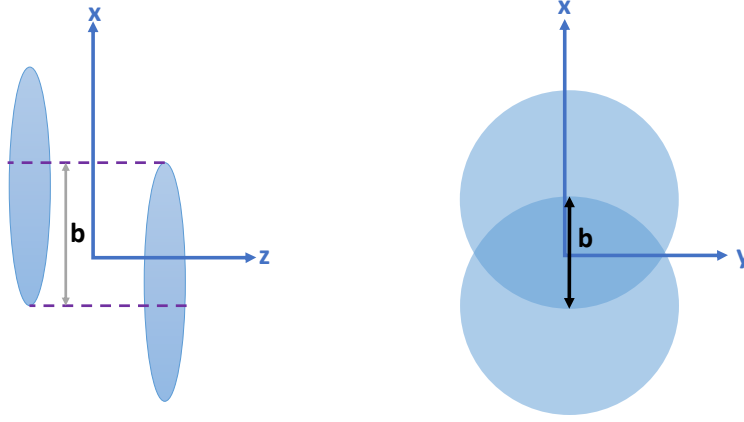


Figure 9: The impact parameter (b) in a heavy-ion collision. Left, xz view of the incoming nuclei. Right, xy , transverse view of the colliding nuclei.

However, in reality b is not directly measurable, hence, in high energy experiments, it is assumed that the impact parameter is monotonically related to particle multiplicity, both at mid and forward rapidity. So for $b = 0$ all the particles in the target and projectile participate in the collision, a.k.a. central collision. For $b > 0$, some but not all the particles from incoming nuclei participate in the collision, a.k.a semi-central collision. For $b \gg 0$, only a small fraction of particles from the incoming nuclei participate in the collision, a.k.a peripheral collision. Hence, the percentage of overlap of the colliding nuclei is discussed in terms of centrality classes, as depicted in Figure 10.

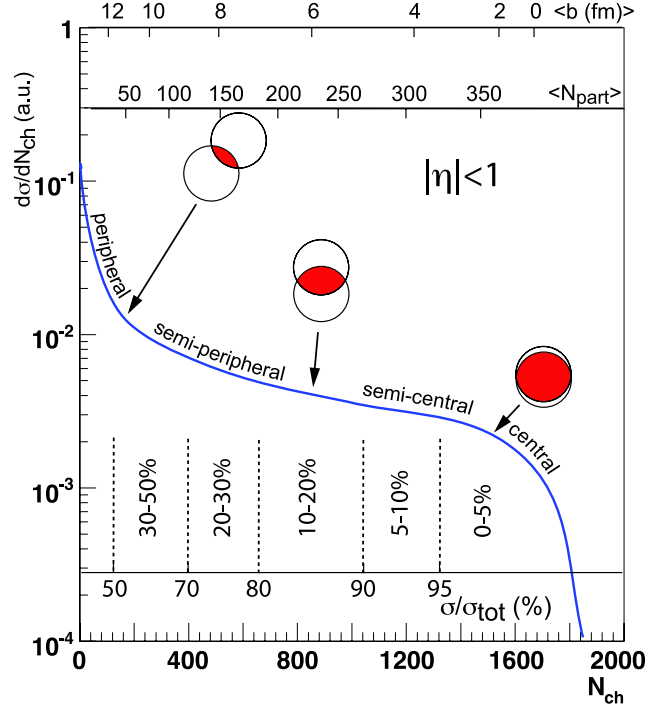


Figure 10: A cartoon example of the correlation of the final state observable N_{ch} with Glauber calculated quantities (b , N_{part}). Figure taken from [mcglauber2007].

2.2 Overview of theoretical models in HIC

Throughout this thesis, a number of theoretical models will be discussed, and this subsection gives a brief overview of them.

2.2.1 Color Glass Condensate (CGC)

The effective field theory that describes QCD at high energy limit is the Color Glass Condensate (CGC) [cgc2010]. The early stages of a hadron-hadron heavy-ion collision can be described with this model. At high energies, the hadron is dilute, and it is characterized by high gluon occupation numbers. Because of its high density, the CGC is weakly coupled, and it can be studied within

perturbative QCD. For hadron-hadron collisions, like the ones that take place at RHIC and the LHC, in order to model the bulk of particle production, it is necessary to consider many-body correlations associated with gluon saturation in both nuclei, together with the multiple scattering between the two saturated gluon distributions.

In nuclear collisions, CGC dynamics produces a field of configurations at early times known as Glasma, i.e. immediately after the collision, the fields contained within nuclei are now color screened on transverse distance between the flat pancakes in form of strong longitudinal chromo-electric and chromo-magnetic fields, this is depicted on Figure 11. Each of the flux tubes can produce particles along the longitudinal direction. As a result, Glasma fields generate long-range rapidity correlations, which experimentally manifest in the observation of the so-called “ridge” present in near-side long-range two-particle correlations of the final particle distribution.

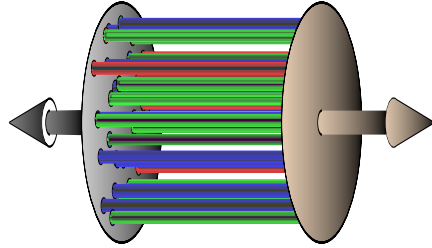


Figure 11: Gauge field configurations in the form of “flux tubes” of longitudinal chromo-electric and chromo-magnetic fields screened on transverse distance. Figure taken from [Iancu:2012xa].

2.2.2 Monte Carlo Glauber

The Monte Carlo (MC) Glauber model is based on a simple geometrical picture of a nucleus-nucleus collision. It is used to describe the initial state of the collision. Nucleons are assumed to travel on straight line trajectories, independent of whether they collide with other nucleons or not. After a nucleon is struck by another nucleon, there is definitely a highly excited baryonic object [mcglauber2007]. In the MC framework the nucleons of the two nuclei are distributed in

space according to the nucleon density profile, parametrized with a Woods-Saxon function. Then an impact parameter b of the nucleus-nucleus collision is chosen randomly. Finally the resulting initial density profile is composed of nucleons that suffered at least of one collision.

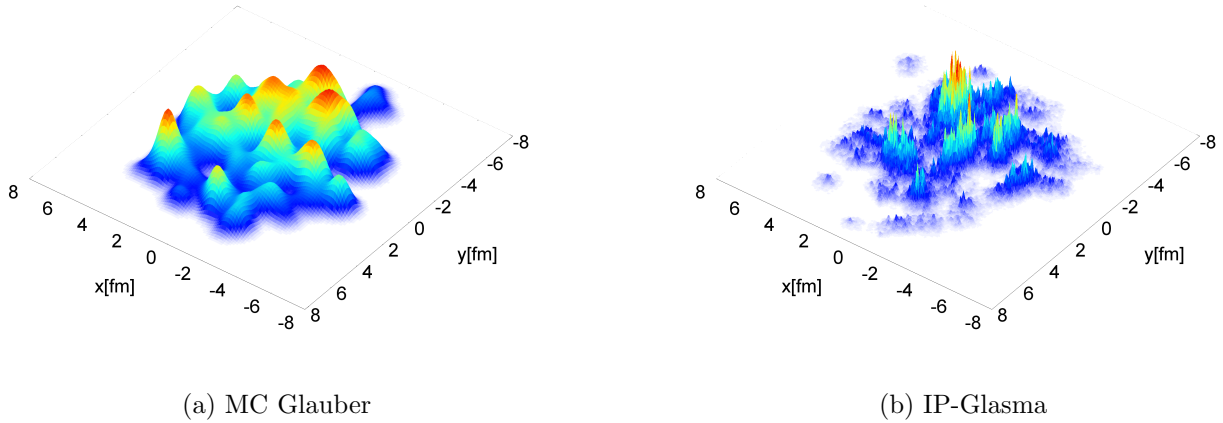


Figure 12: Initial energy density (arbitrary units) in the transverse plane for heavy-ion collision events produced with different models. Figures taken from [kln2012].

2.2.3 IP-Glasma

Odd flow harmonics are entirely driven by fluctuations, hence, it is crucial to have a realistic description of quantum fluctuations in multi-particle production in order to properly describe the final state dynamics. Hence, an impact parameter dependent saturation model has been developed, the so called IP-Glasma, which is based on the CGC framework. This model determines fluctuating configuration of color charges in two incoming highly energetic nuclei. The Glasma distribution is computed using classical Yang-Mills equations and matched event-by-event to viscous relativistic hydrodynamics in order to compute harmonics of the flow distribution[kln2012].

2.2.4 Heavy Ion Jet Interaction Generator (HIJING)

HIJING is a Monte Carlo program that studies jets and the associated particle production in high energy p-p, p-A and A-A collisions [hijing1994][hijing2011]. It combines a QCD-inspired model

of jet production with the Lund model for jet fragmentation. Hard or semi-hard parton scatterings with transverse momenta of a few GeV are expected to dominate high energy ion collisions. HIJING provides underlying events with realistic fluctuations coming from jets, and incorporates mechanisms such as multiple minijet production, soft excitation, nuclear shadowing of parton distribution functions and jet interactions in dense hadronic matter. Furthermore, it produces initial conditions for parton and hadron cascade models such as the AMPT model.

2.2.5 A Multi-Phase Transport Model (AMPT)

The AMPT model includes initial partonic and final hadronic interactions and the transition between them[**ampt2005**]. It aims to provide a kinetic description of all essential stages of heavy ion collisions over a wide range of center-of-mass energies from $\sqrt{s_{\text{NN}}}$ of a few GeV up to 5500 GeV. It uses HIJING for generating the initial conditions[**hijing2011**], Zhang’s Parton Cascade (ZPC) for modeling partonic scatterings[**zpc1998**], the Lund string fragmentation model or a quark coalescence model for hadronization, and A Relativistic Transport (ART) model for treating hadronic scatterings.

There are two versions of the model, the default AMPT, and the AMPT model with string melting. In the default AMPT model partons are recombined with their parent strings when they stop interacting, and the resulting strings are converted to hadrons using the Lund string fragmentation model. In the AMPT with string melting, a quark coalescence model is used instead to combine partons into hadrons. The latter has been more successful for describing anisotropic flow measurements from heavy-ion collisions. Therefore, HIJING and AMPT models allows us to understand how longitudinal multiplicity fluctuations in the early time can be affected by the final-state interactions.

2.3 Previous results on longitudinal fluctuations

The pseudorapidity of the final particle distribution is sensitive to event-by-event (EbyE) longitudinal fluctuations. So far, most of the theoretical models that explain the evolution of a heavy-ion collision, in particular the initial state, rely on the approximation of boost invariance along the longitudinal direction since the incoming nuclei are highly Lorentz contracted. However, it has been seen that the density fluctuations influence the early time entropy production well before the onset of the collective flow. The latter manifests as long-range correlations (LRC) in pseudorapidity which originate at early stages of the collision before the longitudinal expansion separates the particles by large distances, and hence can be used as a probe of the initial conditions of the evolution of the system.

The main consequence from the longitudinal fluctuations has been the asymmetric density of the fireball in one side of the rapidity (or pseudorapidity) window than in the other. There may be different sources for this asymmetry, like the forward-backward difference in number of collisions, or the difference in the number of flux-tubes described in the CGC/Glasma picture [bzdak2013]. In CGC the correlations of the particles created at early stages of the collisions can spread over large rapidity intervals, unlike the particles produced at later stages. Hence, the description of the asymmetry can reveal nontrivial information about the structure of the fireball in the longitudinal direction. Furthermore, it can test various models of particle production in hadronic collisions.

The forward-backward asymmetry can be easily interpreted as the difference between the number of participants coming from the target and the projectile. Figure 13a shows a cartoon of the incoming bunches of particles in the longitudinal directions, and Figure 13b shows the difference of the number of participants in the collision. In this example, 15 nucleon participants come from the target (N_{part}^f), and 10 from the projectile (N_{part}^b), so if we measure the asymmetry as,

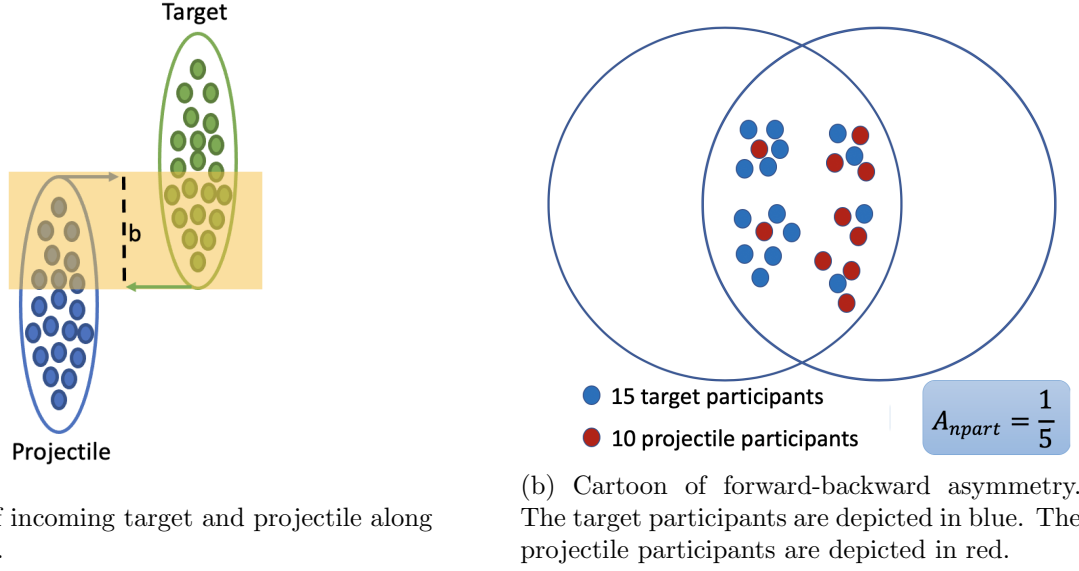


Figure 13: Schematic longitudinal particle distribution at early stages of the collision.

$$A_{npart} = \frac{N_{part}^f - N_{part}^b}{N_{part}^f + N_{part}^n} \quad (4)$$

we obtain that the asymmetry factor is of $\frac{1}{5}$. In a p–A collision, a forward-backward asymmetry is expected as the number of initial participants is different. However, in an A–A collision, excluding effects due to the detector, in principle there should not be a significant forward-backward asymmetry since the system is symmetric, which disagrees with what has been seen in data collected by RHIC and the LHC.

Different methods have been applied in order to characterize the forward-backward asymmetry. For instance, using data from RHIC measured by STAR for Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV, the LRC strength was determined using the dependence of the average charged particle multiplicity in the backward hemisphere, $\langle N_b \rangle$, on the event multiplicity in the forward hemisphere, N_f , such that $\langle N_b \rangle = a + bN_f$, where a is a constant, and b measures the correlation strength [lrc2009]:

$$b = \frac{\langle N_f N_b \rangle - \langle N_f \rangle \langle N_b \rangle}{\langle N_f^2 \rangle - \langle N_f \rangle^2} \quad (5)$$

In 2009, the results were compared to models like the Parton String Model (PSM) and HIJING version 1.383 (see Figure 14). The predictions from PSM and HIJING are lower than the data. In particular, the strong fall of the correlation strength, b , with respect to $\Delta\eta$ for HIJING, provides some constraints on the contribution of the impact parameter fluctuations to the correlation strength. Hence the need to include the effects of forward-backward asymmetry in the models.

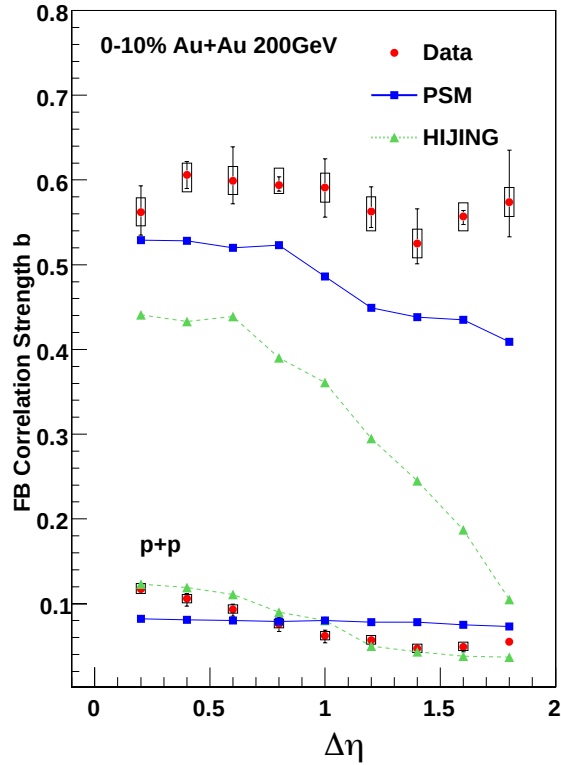


Figure 14: LRC strength measured with STAR on 2009. Forward-backward correlation strength for 0-10% most central Au–Au collisions and p–p from data. HIJING and PSM predictions are shown as well. Figure taken from [lrc2009].

Later on, a systematic study of the factorial moments of the multiplicity distribution for central Au–Au collisions data from STAR [bzdak2012], revealed a significant boost invariance on the

EbyE basis, which may substantially affect the hydrodynamic expansion of the system. From this result, Bzdak and Teaney proposed the decomposition of longitudinal fluctuations of the fireball using orthogonal Chebyshev polynomials [bzdak2013], analog to the Fourier decomposition in the transverse plane for anisotropic flow. In their work, the decomposition was implemented for the rapidity y , but can be extended to any longitudinal variable, for instance, the pseudorapidity η . In their study, it was found that by using Chebyshev polynomials, each polynomial represents a different component of the fluctuating rapidity density of the fireball, and hence, the coefficients extracted from the expansion reveal nontrivial information about the fireball structure.

In more recent works, the idea of decomposing the longitudinal fluctuations with orthogonal polynomials, has been extended to use Legendre polynomials. In 2016, the ATLAS Collaboration proposed a single-particle method to obtain the shape components directly from each event, and a two-particle correlation method which give the ensemble RMS-average of the shape components [jia2016]. The analysis uses pseudorapidity, η , as the longitudinal variable. It defines a forward-backward correlation as:

$$C(\eta_1, \eta_2) = \frac{\langle N(\eta_1)N(\eta_2) \rangle - \langle N(\eta_1) \rangle \langle N(\eta_2) \rangle}{\langle N(\eta_1) \rangle \langle N(\eta_2) \rangle} \quad (6)$$

where $N(\eta) = dN/d\eta$ is the multiplicity density distribution in pseudorapidity per event, and the average is over the event ensemble, i.e. different centrality. Considering a single particle distribution, Equation 6 can be defined in terms of the normalized observed multiplicity density distribution $R(\eta) = N(\eta)/\langle N(\eta) \rangle$:

$$C(\eta_1, \eta_2) = \langle R(\eta_1)R(\eta_2) \rangle - \frac{\delta(\eta_1 - \eta_2)}{\langle N(\eta_1) \rangle} \quad (7)$$

hence, in the absence of EbyE fluctuations, $R(\eta) = 1$ and $C = 1$. Furthermore, the observed multiplicity density distribution, $R(\eta)$ can be defined in terms of Legendre polynomials in the rapidity range $[-Y, Y]$,

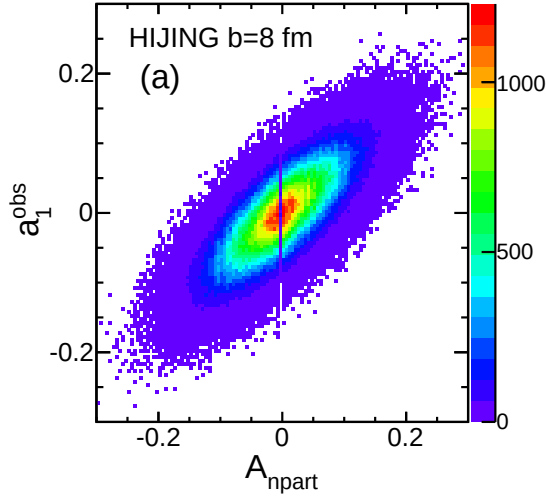
$$R(\eta) = 1 + \sum_n^{\infty} a_n^{obs} T_n(\eta) \quad , \quad T_n(\eta) = \sqrt{n + \frac{1}{2}} P_n(\eta/Y) \quad (8)$$

where P_n represents different orders of the Legendre polynomials. The longitudinal coefficients, a_n are then extracted from the correlation as,

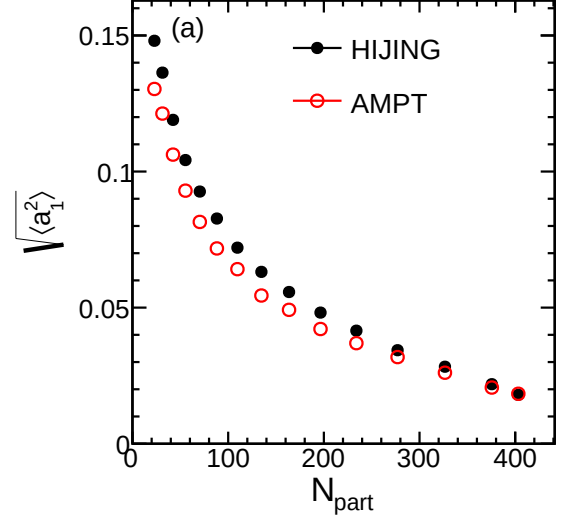
$$\langle a_n a_m \rangle = \frac{1}{Y^2} \int [C(\eta_1, \eta_2) - 1] \frac{T_n(\eta_1) T_m(\eta_2) + T_n(\eta_2) T_m(\eta_1)}{2} d\eta_1 d\eta_2. \quad (9)$$

The first order longitudinal coefficients a_1 for Pb–Pb collisions at LHC energy of $\sqrt{s_{NN}} = 2.76$ TeV were determined theoretically for the HIJING and AMPT models, with an analysis that was carried out in a pseudorapidity range $|\eta| < Y = 6$, and fixed impact parameter $b = 8$ fm. From HIJING events, a strong correlation between a_1^{obs} and A_{npart} (calculated with Equation 4) is observed as shown in Figure 15a. This strong correlation suggests that the FB asymmetry in the multiplicity distribution is indeed driven by the number of participant nucleons in the two colliding nuclei. Furthermore, in Figure 15b it can be seen how the magnitude of a_1 depends on the centrality, i.e. the strength of the signal increases for more peripheral collisions.

In a later work, the longitudinal coefficients, a_n , were extracted for Pb–Pb collisions at energy of $\sqrt{s_{NN}} = 2.76$ TeV, p–pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and p–p collisions at $\sqrt{s} = 13$ TeV, collected by the ATLAS detector [jia2017]. In order to extract longitudinal fluctuations coming from the initial state of a heavy-ion collision only, it is necessary to get rid of longitudinal fluctuations that are generated during the final state of the collision, which are present in short range correlations (SRC). Effects from SRC come from phenomena generated at the final state of the collision, such as resonance decays, jet fragmentation and Bose-Einstein correlations. In Figure 16, for $|\eta| < 2.4$, the result showed a significant difference from longitudinal fluctuations calculated from short range correlations (SRC) for the different colliding system, whereas a_1 was found to be consistent for the colliding systems in long range correlations (LRC).



(a) Event by event correlation between a_1^{obs} and A_{npart} from HIJING events with $b = 8 fm$. Figure taken from [jia2016]



(b) Centrality dependence of a_1 for HIJING and AMPT events. Figure taken from [jia2016].

Figure 15: First order longitudinal coefficients from HIJING and AMPT events calculated by the ATLAS Collaboration.

One of the main purposes of the ALICE experiment is the study of the QGP, hence, the study of the properties of the initial state of a heavy-ion collision. The data collected by the ALICE detector may provide further details of the analysis of longitudinal fluctuations using decomposition with Legendre polynomials. The goal of this thesis is to study for the first time event by event longitudinal fluctuations with ALICE collected data applying the Legendre polynomials decomposition.

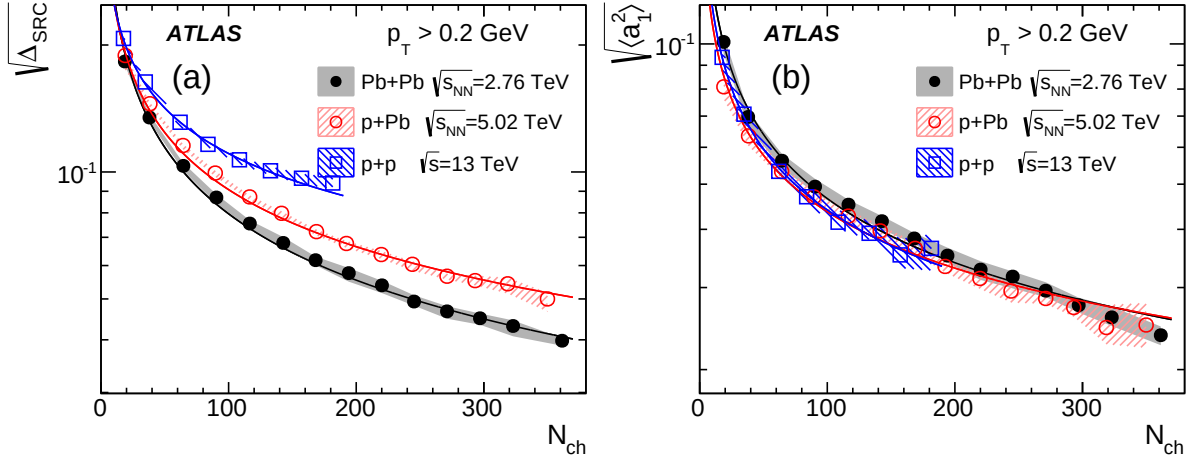


Figure 16: Magnitude from short range correlations, SRC (left), and long range correlations, a_1 (right), as a function of the final multiplicity N_{ch} . Figure taken from [jia2017].

3 High Energy Physics Experiments

As mentioned before, among the heavy ion colliders that have the design needed to reproduce a QGP are the Relativistic Heavy-Ion Collider (RHIC) at Brookhaven National Laboratory (BNL) and the Large Hadron Collider (LHC) at CERN (European Organization for Nuclear Research). This thesis uses data from one of the experiments at the LHC, ALICE.

3.1 Large Hadron Collider (LHC)

The LHC at CERN is a 27 km super-conducting particle accelerator located approximately 100 m underground (Figure 17). It sits on the border between Switzerland and France, near the city of Geneva. It was built between 1998 and 2008, steering the first beam on 10 September 2008. The LHC collides beams of protons and/or ions in opposite directions at four locations along the super-conducting ring, where four main experiments are located: ALICE (A Large Ion Collider Experiment), ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid experiment) and LHCb (LHC beauty). Each of the experiments focuses on the studies of different phenomena. CMS and ATLAS are the biggest experiments in the LHC; they use general-purpose detectors to

investigate the largest range of physics possible, e.g. the study of the Higgs boson and physics beyond the Standard Model. On the other hand, ALICE and LHCb have detectors specialised for focussing on specific phenomena, e.g. the LHCb investigates differences between matter and antimatter by studying the beauty quark, and ALICE studies the physics of strongly interacting matter at extreme energy densities where the QGP takes place.

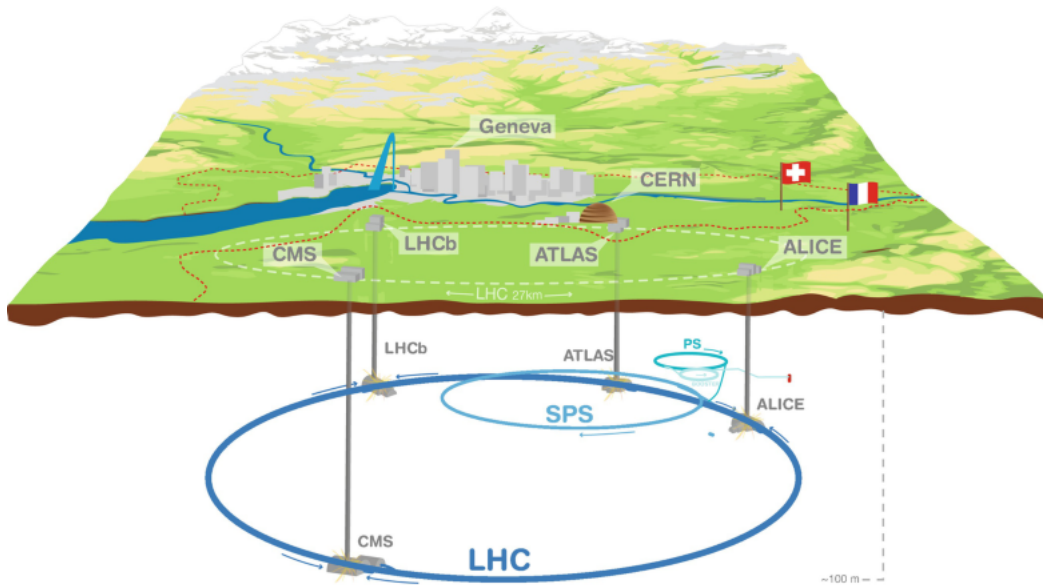


Figure 17: Diagram of the LHC. Figure taken from [Stark2020].

3.1.1 LHC Timeline

The LHC performance is measured in terms of its luminosity, which is proportional to the number of collisions that occur per unit of time. The higher the luminosity, the more data the experiments can collect to allow them to analyze rare physics phenomena. The evolution of the LHC high luminosity is depicted on Figure 18. Throughout the years, different collisions have taken place at different energies at the LHC. The data that is analyzed in this thesis corresponds to data taken during Run 2. Furthermore, the description of the ALICE detector that follows in the next subsection includes upgrades that have been done during the Long Shutdown 2 (LS2) for Run 3 and 4.

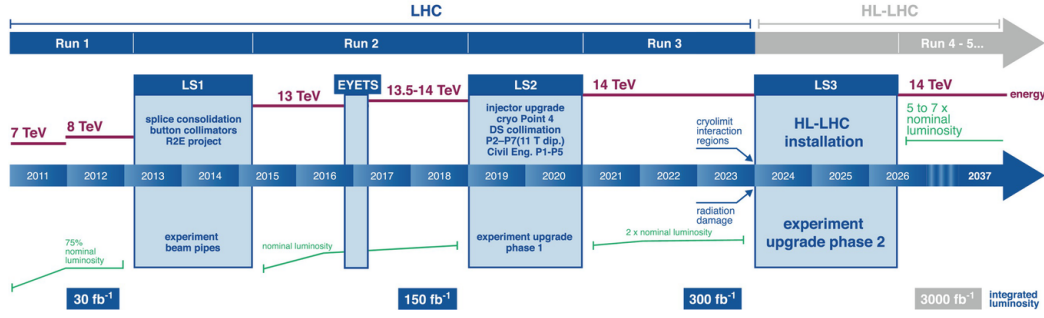


Figure 18: LHC luminosity timeline. Figure adapted from [Stark2020].

3.2 A Large Ion Collider Experiment (ALICE)

ALICE is a general-purpose heavy-ion experiment designed to study the physics of strongly interacting matter and the QGP in nucleus–nucleus (A–A), proton-nucleus (p–A), and proton-proton (p–p) collisions at the LHC. It has a redundant and robust tracking system and several systems for particle identification. The ALICE detector consists of a central barrel, a muon spectrometer at forward rapidity and a set of detectors for triggering and event characterization [aliceloi2014] depicted on Figure 19. An overview of the main tracking and particle identification detectors is given in the next subsections; a complete description of all ALICE detectors can be found in [alicetech2008].

3.2.1 Time Projection Chamber (TPC)

The TPC is the main and biggest ALICE detector; its main purpose is the tracking of charged-particle momentum measurements with good two-track separation, particle identification, and vertex determination. The phase space covered by the TPC in pseudo-rapidity is $|\eta| < 0.9$ for tracks with full radial track length (matches in Inner Tracking System (ITS), Transition Radiation Detector (TRD), and Time-Of-Flight(TOF) detectors); for reduced track length (at reduced momentum resolution), an acceptance up to about $|\eta| = 1.5$ is accessible. The TPC covers the full azimuth (with the exception of the dead zones). A large transverse momenta, p_T , range is covered from low p_T of about 0.1 GeV/c up to 100 GeV/c with good momentum resolution. During LHC Run

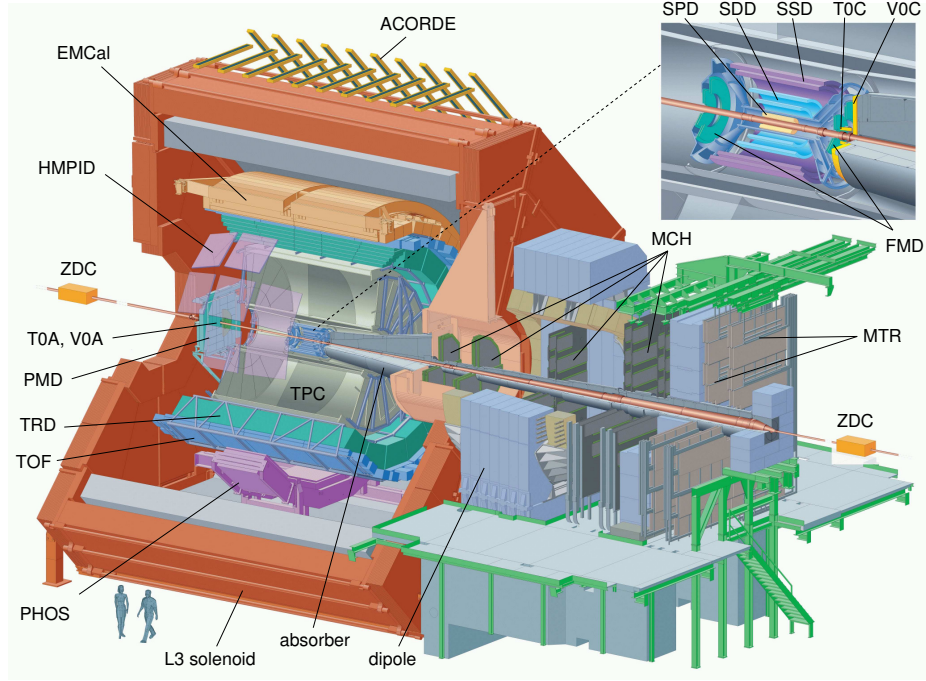


Figure 19: Diagram of the ALICE detector. Figure taken from [alicefig2014]

1 and 2 the detector consisted of a large cylindrical field cage, covering $87 < r < 247$ cm and $250 < z < 250$ cm in radial and longitudinal directions, respectively. A central high voltage electrode at $z = 0$ divided the active drift volume of 90 m^3 into two halves providing an electric drift field of 400 V/cm . The volume was filled a mixture of Ne, CO_2 , N_2 gas whose composition varied with the performance aiming to keep multiple scattering and space charge effect at a minimum. For instance, the maximum electron drift time for a Ne- CO_2 (90-10) gas mixture is about $100 \mu\text{s}$. The primary electrons were transported over a distance of up to 2.5 m on either side of the central electrode to the end plates.

Multi-Wire Proportional Chambers (MWPC) with cathode pad readout were mounted into 18 trapezoidal sectors at each end plate [alicetech2008]. In the radial direction these end plates are split up into two chambers. The readout chambers employ a commonly used scheme of wire grids, including a plane of anode wires, cathode wires, and a gating grid. In the absence of a trigger, the gating grid wires are biased by a voltage that prevents electrons to enter the amplification region

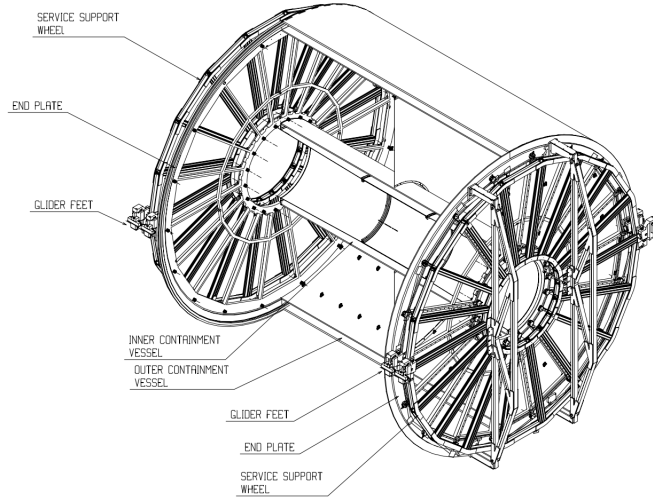


Figure 20: 3D view of the TPC field cage and service support wheels, without readout chambers. Figure taken from [alicetech2008].

and ions from preceding events to penetrate into the drift volume. The maximum drift time of ions from the amplification region to the gating grid is about $180 \mu s$. In order to provide efficient ion gating, the gating grid must be closed for $180 \mu s$ after the end of the electron drift. In addition to that, a $100 \mu s$ past protection is applied to avoid event pile-up. Hence the read out rate is limited to a maximum of 3.5 kHz.

The TPC provides particle identification (PID) via the measurement of the specific energy loss by ionization dE/dx . The solid lines in Figure 21 are a parametrization of the detector response based on the Bethe-Bloch formula [Blum:1105920].

For Run 3 and 4, the goal is to operate the TPC in an ungated mode while keeping the ion space charge induced drift field distortions at a tolerable level. Furthermore, the excellent tracking and PID capabilities present in the detector during Run 1 and 2 are being kept. As a result, the MWPC based readout chambers used in the past were replaced by Gas Electron Multipliers (GEM) detectors which feature intrinsic ion blocking without additional gating and exhibit excellent rate capabilities. Thanks to this upgrade, it is possible to have continuous operation where Pb-Pb

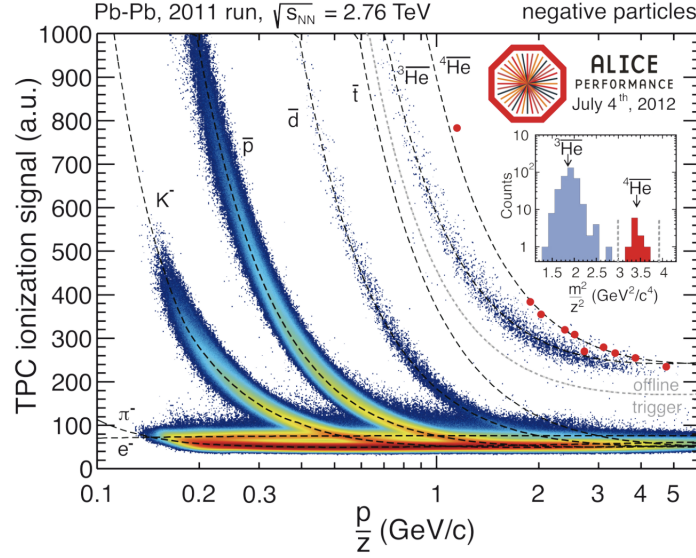


Figure 21: TPC dE/dx measurements in the full 2011 Pb-Pb statistics. Figure taken from [aliceloi2014].

collisions at a rate of 50kHz can be recorded. The latter implies an increase by about a factor 100 in recorded luminosity, leading to about $2.5 \cdot 10^9$ central Pb-Pb events in a typical year of heavy-ion running[aliceloi2014].

3.2.2 Inner Tracking System (ITS)

The ITS consists of six layers of silicon detectors placed coaxially to the beam pipe, shown in Figure 22. Its main tasks are to localize the primary vertex with a resolution better than $100 \mu\text{s}$ in order to reconstruct secondary vertices from the decays of hyperons and D and B mesons, to track and identify particles with momentum below 200 MeV/c, to improve the momentum and angle resolution of particles reconstructed by the TPC, and to reconstruct particles that traverse the dead regions of the TPC [alicetech2008].

The Silicon Pixel Detector (SPD) is the inner most silicon detector, it has two layers located at 3.9 cm and 7.6 cm each from the beam axis. Its goal is to determine the primary and secondary

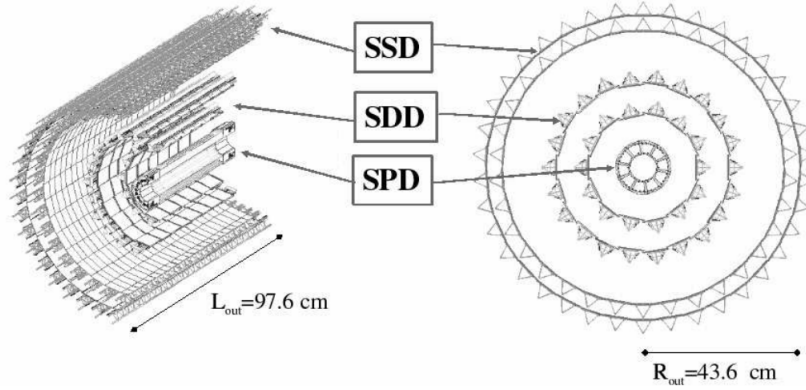


Figure 22: Layout of the ITS. It is composed of three detectors, each of them consists of two layers: the Silicon Pixel Detector (SPD) closely surrounding the beam pipe, the Silicon Drift Detector (SDD), and the Silicon Strip Detector (SSD). Figure taken from[alicetech2008].

vertex as well as the measurement of impact parameter of secondary tracks. The first layer has a more extended pseudo-rapidity coverage ($|\eta| < 1.98$) which, together with the Forward Multiplicity Detectors (FMD), provides continuous coverage for the measurement of charged particle multiplicity. The Silicon Drift Detector (SDD) has two layers located at 15 cm and 23.9 cm from the beam axis. The SDD is able to provide a two-dimensional spatial information in moderate particle density without the necessity of separation into small pixel cells. One dimension is obtained from the signal readout along the anode strip, and second dimension is obtained from the drift time of the electrons towards the anode. As a result it has multi-particle tracking capability and provides energy loss measurements for particle identification. The Silicon Strip Detector (SSD) has two layers located at 38.0 cm and 43.0 cm from the beam axis. It connects the ITS tracks to the tracks in the TPC and also provides dE/dx information.

The ITS Run 2 (ITS 2) design is limited by its poor readout rate capabilities, as a result, for Run 3, the ITS consists of seven layers based on Monolithic Active Pixel Sensors (MAPS), where the ITS 2 is the first large-area silicon tracker (Figure 23). The pseudorapidity coverage is of $|\eta| < 1.22$ for tracks from 90% luminous region.

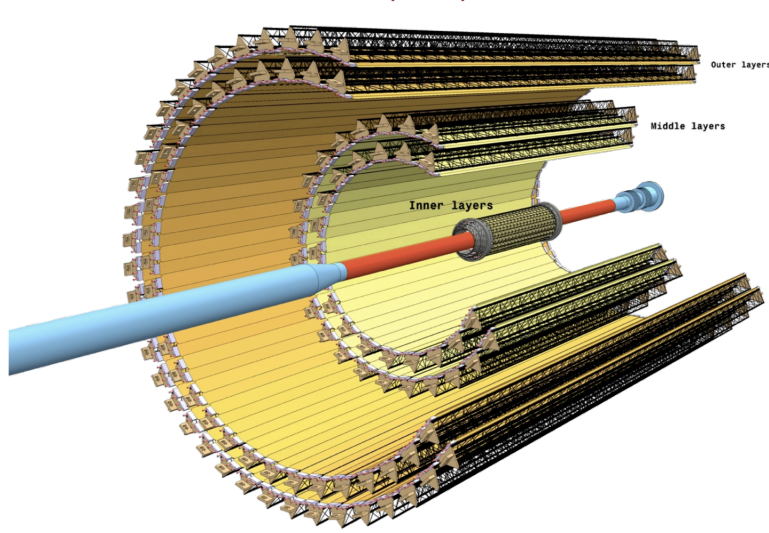


Figure 23: Layout of the ITS for Run 3. It consists of 3 Inner Barrel layers (IB) and 4 Outer Barrel layers (OB). Figure taken from [Musa:1431539].

3.2.3 VZERO (V0)

The V0 detector is composed of two plastic scintillator arrays, V0A and V0C situated at 90 cm and 340 cm on either side of the collision vertex ($z = 0$) with pseudorapidity coverage of $-3.7 < \eta < -1.7$ and $2.8 < \eta < 5.1$ respectively; a detailed description can be found at [vzero2013]. It has a time resolution better than 1 ns. It provides a highly efficient interaction trigger, a selective centrality trigger for Pb–Pb collisions and it is also used for rejection of beam-gas interactions.

Details about other ALICE detectors and the upgrades they have gone through from Run 2 to Run 3 can be found at [alicetech2008] and [aliceloi2014].

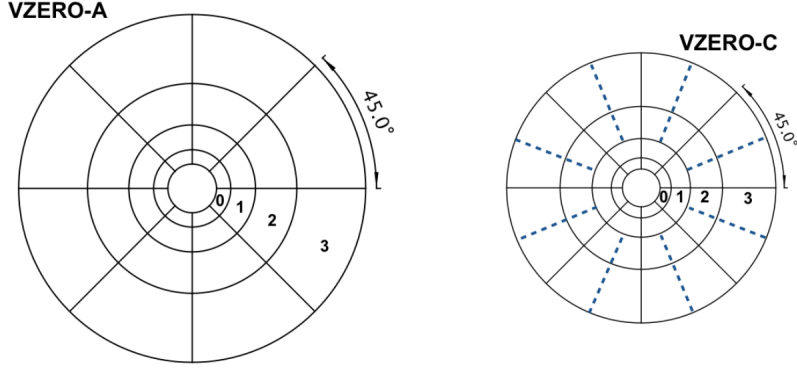


Figure 24: Sketches of VZERO-A and VZERO-C arrays showing their segmentation. Scintillator thicknesses are 2.5 and 2 cm respectively. The scintillator segments on both sides of the dashed lines are connected to the same photomultipliers. Figure taken from [vzero2013].

4 ALICE Train System

4.1 ALICE Computing Tools

During LHC Run 1, the ALICE experiment collected 7.2 PB of raw data. In order to handle this huge amount of data, ALICE has implemented a global distributed computing system, ALICE Grid. The ongoing major upgrade for LHC Run 3 will enable the experiment to cope with an increase in data rate collection to be of 50 kHz for Pb–Pb collisions during Run 3, around two orders of magnitude more than in Run 2. Due to this increased data rate collection, the Offline-Online computing system (O^2) has been developed in order to have an efficient trigger and read-out system. In addition to that, one of the upgrades done for the trigger was the implementation of an online monitoring, the ALICE overwatch project (more information to be found in Appendix A).

ALICE analyses mostly deal with large datasets using the distributed Grid infrastructure. In LHC running periods 1 and 2, ALICE developed a system of analysis trains, the so-called “LEGO trains”, that allowed the user to configure analysis tasks, a.k.a wagons, that are expected to be run on the same data. The LEGO train system builds upon existing tools: the ALICE analysis framework as well as the Grid submission and monitoring infrastructure. This centralized system

improved the resource utilization and provided a friendly user interface (UI). Currently, 90% of ALICE analyses use the train system. However, the new ALICE computing system, O², is not compatible with the LEGO trains. Analysis trains will also be the main workhorse for analysis in Run 3, hence, a new infrastructure, Hyperloop, is being developed based on the successful concept of the LEGO trains.

4.2 The LEGO Train System

The concept of the LEGO train framework consists of combining multiple analysis tasks which should analyze the same dataset and running them within the same analysis job. By doing this, the data is read only once to be used for multiple analysis tasks, instead of reading data for every analysis task and for each user like in the case of using the Grid only. As a result, less time is spent loading the data and the CPU efficiency increases. Furthermore, by having a web-based GUI, the users do not have to deal with the complexity of the Grid since they only need to define their code on the webpage and wait for the analysis results to be available where the jobs have been properly merged.

In the LEGO train system, the trains are divided per physics working groups, PWGs, (i.e. jet analysis, particle correlations, etc) and per data type and collision system (p-p, Pb-Pb or p-pb raw data or the corresponding Monte Carlo). The process of a train run is described in Figure 25. In order to join a train, a user has to add their code for an analysis task as a wagon inside of the corresponding train. After that, the user needs to enable that wagon in the dataset required. Once a wagon is enabled, a train operator (the expert who has the permissions to test and run trains) combines several users wagons in a train run and tests the train run. The train operators are assigned per PWG. If the train run test is successful, the operator proceeds to submit the train run to the Grid, and the users will receive their results merged without the need of monitoring the evolution of their analysis jobs submission. A more detailed explanation of the LEGO train system can be found in [zimmermann2015alice] and [Zimmermann:2314849].

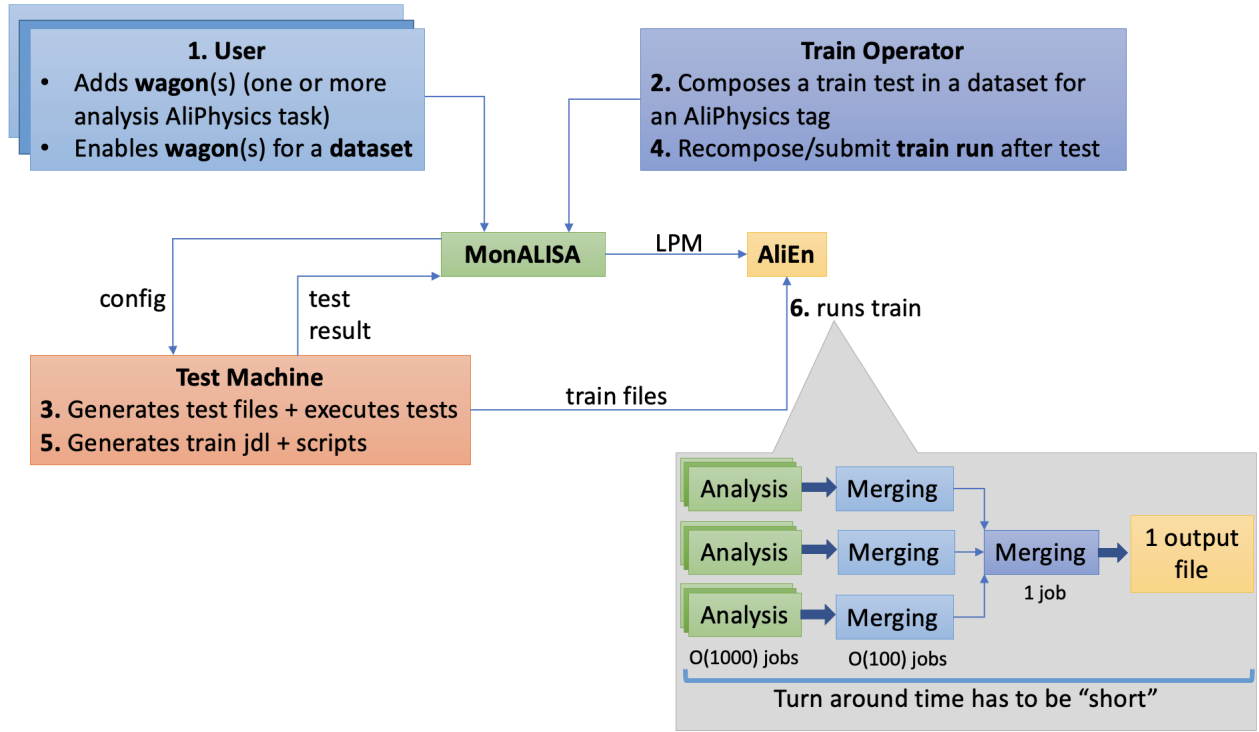


Figure 25: Procedure of a LEGO train run. After a successful train test, the analysis jobs run. Then the results are merged into a single output file.

4.2.1 LEGO Web Interface

The LEGO train web interface has a database developed using PostgreSQL, an open source object-relational database system. The backend is developed with MonALISA (Monitoring Agents using a Large Integrated Services Architecture), which is the ALICE Grid monitoring tool set. The frontend is basic HTML, with static web pages that are the same for all users and train operators.

The LEGO train system has had a huge impact within the ALICE community, such that over the last 9 years, the number of trains and users has increased significantly, the progress can be seen in Table 1. In 2020, 16000 train runs and 172 million Grid jobs were submitted.

As explained before, the trains are divided per PWG and data type, currently resulting in

	2012	2014	2019
Users	60	188	>470
Trains	42	79	113
Train runs	1537	7447	16524
Number of jobs	12M	34M	192M
Train wagons/run	14.9	8.9	17.4

Table 1: Usage statistics of the LEGO train system. Data for 2012 and 2014 taken from [zimmermann2015alice].

about 87 active trains. The big amount of trains, plus the increasing number of users, decreased the efficiency loading in the static web pages, as the data contained got eventually too large. For instance, for one of the most common trains used, the Correlations and Flow (CF) Pb–Pb train web page (Figure 26) required about 26 seconds to load by the end of 2019.

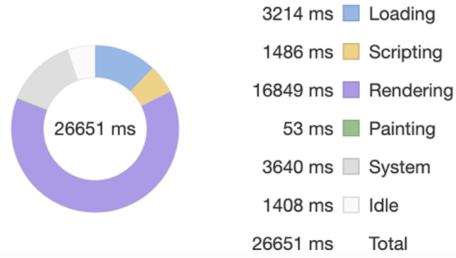


Figure 26: Performance of CF Pb–Pb web page (by 2019).

Furthermore, since trains are divided per PWG, train operators are divided per PWG as well. On average, there are 5 train operators per PWG who take care of the configuration of datasets available per train, and of testing, composition and submission of train runs. In addition to that, the large number of active trains has led to some train runs using the same dataset not being run together as they belonged to a different PWG trains.

4.3 From the LEGO trains to Hyperloop trains

During the Long Shutdown 2 (LS2), the ALICE detector went through upgrades allowing to collect in Run 3 roughly 2 orders of magnitude more data than in Run 1 and 2. This led to the implementation of a new analysis framework, the Online-Offline Computing System (O²)[**Buncic:2015ari**], which is not compatible with the LEGO train system. Given the new analysis framework, as well as other needed upgrades (e.g. large web page time load detailed in 4.2.1), a new train system has been developed for Run 3 since mid 2019, the Hyperloop train system. In the long term, the LEGO trains will be phased out and replaced by the Hyperloop trains. To allow users to utilize Run 2 data with the new software O², the Run 2 data is being converted to the Run 3 data format, and can be handled with the Hyperloop train system. In addition to that, Hyperloop train runs will not be allowed to be submitted to the ALICE Grid only, but they can also be submitted to ALICE Analysis Facilities where different sets of data are stored.

4.4 The Hyperloop Train System

The Hyperloop train system has been developed based on the solid concept of the LEGO trains. The web interface database has been developed using PostgreSQL with additional fields storing validity range (i.e. timestamp range) that allow bookkeeping for wagons and datasets (4.4.1), which was not part of the LEGO train system. The latter allows wagons and train run configuration comparison, which can be used to efficiently understand what has changed in the configuration between two train runs. The backend has been developed in MonALISA. The frontend has been developed using a modern and interactive library, React.js [**reactjs**].

In order to improve the web page load, as well as to facilitate the navigation for users and train operators, instead of having static web pages, the Hyperloop UI offers different personalized web pages for users and train operators. Trains are no longer divided per PWG, instead the concept of an “analysis” is introduced. An analysis is created using a task manager, JIRA, where analyzers

The screenshot displays a user interface for configuring Hyperloop wagons. It features several sections, each with a title, a lightbulb icon, a trash icon, and a configuration area:

- array-string:** A list of three string inputs labeled I1, I2, and I3. I1 and I2 have a minus button, while I3 has both minus and plus buttons.
- integer:** A single integer input field containing the value 2.
- labeled-matrix:** A matrix with 3 rows (labeled r 1, r 2, r 3) and 4 columns (labeled c 1, c 2, c 3, c 4). The values are:

r 1	1.1000	1.2000	1.2999	1.3999
r 2	2.0999	2.2000	2.2999	2.4000
r 3	3.0999	3.2000	3.2999	3.4000
- matrix-double:** A 3x3 matrix of double values:

1.1	1.2	1.3
2.1	2.2	2.3
3.1	3.2	3.3
- fixed-hist-binning:** A configuration panel with a dropdown set to 'Fixed Width', 'Bins' set to 14, 'Min' set to 0, and 'Max' set to 1400.
- variable-hist-binning:** A configuration panel with a dropdown set to 'Variable Width' and a list of values: 0,10,20,30,40,50.

Figure 27: Example of Hyperloop wagon configuration. O^2 parameters are supported and displayed in a user-interactive interface. From top to bottom the sample parameters type displayed are array of strings, integer, matrix of double numbers with columns and rows labels, matrix of double numbers, histogram binning with fixed width, histogram binning with variable width.

that belong to the analysis are defined, most likely by the group Project Leader or a PWG Convener. Analyzers can add, edit and enable any wagon that is part of the analyses they belong to. The usage of JIRA allows an easier management of analysis approval in ALICE.

Users have reading access to all web pages within the UI. The web page where users have writing rights is the *My Analyses* web page. The *My Analyses* web page displays all analyses the analyzer belongs to. It is a personalized web page which saves the user configuration, e.g. analyses order. Wagons are created per analysis from an O^2 workflow task. An O^2 workflow task is available in a daily release of the O^2 software, a.k.a. package tag. Any analyzer can edit the configuration of a wagon within the analysis. In order to edit a wagon configuration, the LEGO trains required a long list of code as an input from the user, however, Hyperloop supports all different types of O^2 parameters (e.g. integer, string, arrays, matrices, labeled matrices, etc) and displays them in a

user-friendly way like in Figure 27. If a workflow task produces derived data, this can be activated in the wagon configuration, and will be saved to the output if the train run in which the wagon is included is run as a derived data production. Furthermore, similar to the LEGO trains, subwagons can be added to a wagon. A subwagon is used when a user requires to run the same configuration of a wagon, but changing only one or a few of the parameters value, e.g. this feature is commonly used when running systematic checks.

Another way of creating a wagon, is by cloning it from a different wagon. Users have access to all analyses available in the system in the *All Analyses* web page. Users can access any analysis in a read only view, and clone a wagon from a different analysis to any analysis where the user belongs to. The cloned wagon will have exactly the same configuration as the original one, including subwagons and dependencies.

Active datasets available in the system can be added per analysis in *My Analyses*. All analyses activated for the same dataset can be combined and run together. The process of a train run submission is explained in Figure 28. For the LEGO trains, a user could only enable the wagon for a dataset with the latest available package tag, and if a different package tag or configuration was required, the user needed to send an email to the train operators with a specific request. For Hyperloop, a user can enable a wagon for a dataset with the following types of package tags:

- **Fixed:** specific package tag
- **Newer:** any package tag not older than the package tag the wagon was enabled with
- **Future pull request:** the workflow task with the latest updates is not available in a released package tag, but is part of a pull request. It will be tested as soon as the pull request is merged and released in the O² software

For the LEGO trains, in order to test an enabled wagon, it had to be composed in a train run by a train operator (Figure 25). LEGO train operators compose train runs on a schedule

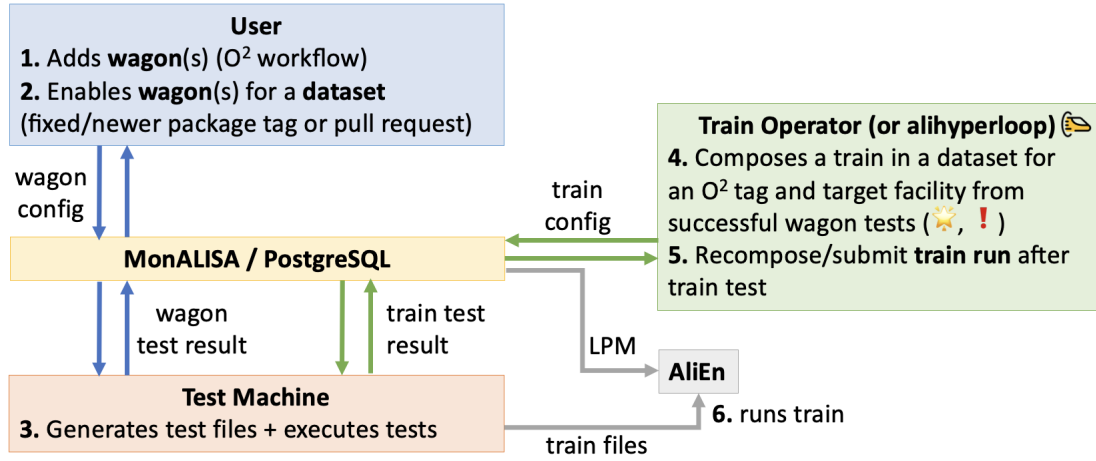


Figure 28: Procedure of a Hyperloop train run. Wagon tests are immediate and automatic after the user enables a train. A train run can be composed by a train operator, or by the system. Figure taken from [quishpe2021hyperloop].

defined by their PWG, e.g. Mondays, Wednesday and Fridays, the other days would be done only upon special request. In Hyperloop, as soon as a wagon is enabled for a dataset, a wagon test will take place. The user gets an immediate feedback of the test results from the system. Furthermore, interactive graphs have been implemented, showing the test statistics of a wagon (Figure 29).

For train operators, there are web pages to add and modify runlists, datasets and train runs. Common users have read only access to those pages. In the LEGO trains, runlists were defined by ALICE Data Preparation Group (DPG) in a Twiki page. If a train operator needed to add or modify a dataset runlist, the train operator needed to access that Twiki page, and manually copy and paste the runlist, and update manually the runlist along with the DPG updates. In Hyperloop, the *DPG runlists* web page has been designed with the purpose of allowing ALICE DPG to define the raw data and MC runlists in the UI. If a runlist is modified, Hyperloop automatically updates the datasets using that runlist through the whole system.

All datasets available in the system are displayed in the *Datasets* web page. Datasets are staged



Figure 29: Test statistics of a wagon for an active dataset. (Top left) Proportional set size memory vs. test ID. (Top right) CPU/Wall time and Throughput vs. test ID. (Bottom left) Output size vs. test ID.

in different analysis facilities (4.4.3). Furthermore, a schedule for automatic composition along with other parameters can be defined in the dataset configuration(4.4.2). There are 4 types of datasets:

- **Data:** Run 3 raw data defined from a runlist available in the DPG runlists web page
- **MC:** Run 3 Monte Carlo data defined from a runlist available in the DPG runlists web page
- **RUN2:** Converted Run 2 data
- **HY:** Derived data produced in a Hyperloop train run

All datasets with enabled wagons and final test result of success or warning are displayed in the *Train Submission* web page. Train operators can compose a train run per dataset based on the package tag, target facility maximum memory, wagons common dependencies and wagons configuration compatibility. An algorithm has been developed returning the ideal wagon combination for

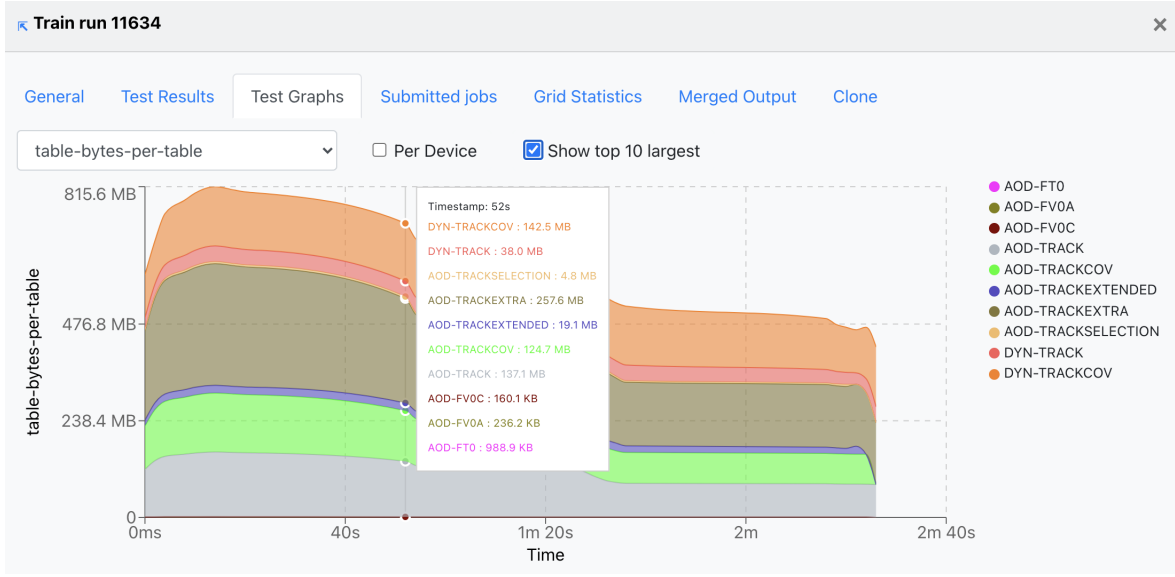


Figure 30: Test graphs from train run test. Table-bytes-per-table vs. time. Graphs can also be displayed per device.

a train run based on the parameters mentioned before. A train run can be composed with a flag for automatic submission, i.e. if the train run test is successful, the train run will be automatically submitted. Similar to the LEGO trains, the train run final result is merged.

The *Train Runs* web page displays all available train runs in the system. The train run result includes interactive graphs for the overall performance (Figure 30), Grid statistics (Figure 31), as well as direct links to the dataset and wagons specific configuration that participated in the train run. In addition to the general settings, tests and jobs results, when a train run is done, it can be cloned. The cloned train run will be created with the current dataset configuration, but with the wagons configuration that belong to the timestamp of the original train. If a train run has been submitted but the jobs are stuck, a summarized error log can be found in the *Trains with issues* web page. Furthermore an overall summary of the current system status can be found in the *Dashboard* web page, e.g. number of wagons waiting to be tested, AliEn jobs ran in the last week, number of finished trains during the last week, etc.

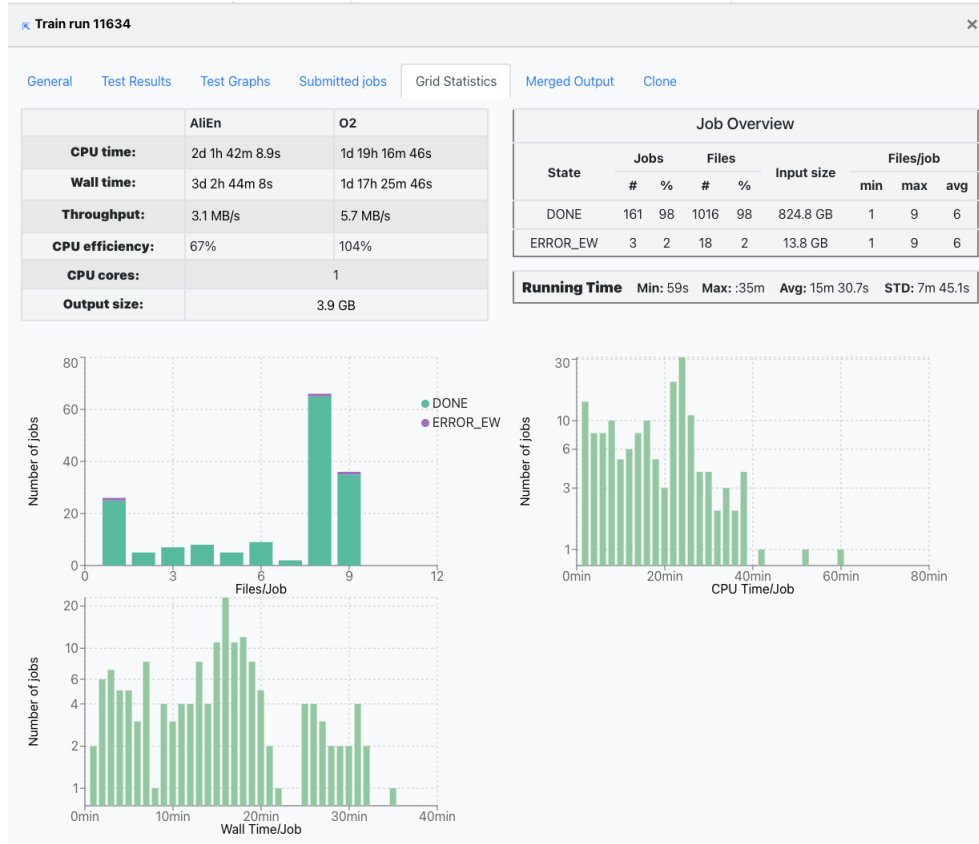


Figure 31: Grid statistics. Shows overall jobs performance. (Center left) Number of jobs vs. Files per job. (Center right) Number of jobs vs. CPU time per job. (Bottom left) Number of jobs vs. wall time per job.

4.4.1 Bookkeeping and comparison

The validity ranges added to some fields for the database, allows to save wagons, datasets and runlists configurations per timestamp. Thanks to this, a detailed change-log of the configuration can be accessed for wagons, datasets and runlists, where the time and user name who modified the configuration is specified (Figure 34). Direct links to a read only view of the configuration at a specific are also available, and it is possible to compare two wagon configurations from different timestamps (Figure 32).

The fact that wagons and datasets have a specific configuration per timestamp allows to have

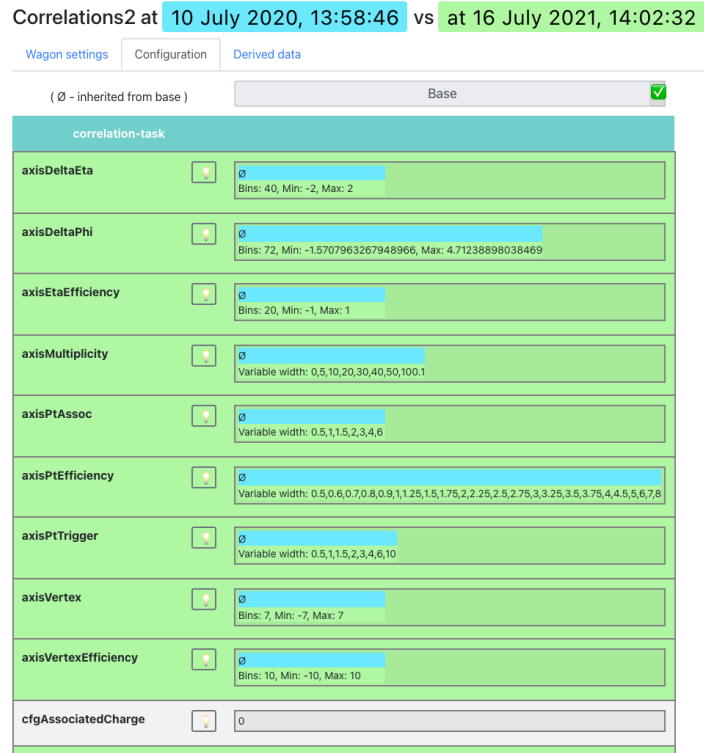


Figure 32: Wagon comparison. Comparison of wagon configurations corresponding to two different timestamps. One timestamp is on cyan, and the other green. The common parameters are displayed on gray.

train run comparison as well. In the *Train Runs* web page, two train runs can be selected and compared (Figure 33). Direct links to the participating wagons and datasets are available for the specific timestamp. The train run comparison is a tool that will allow to improve the performance of the system.

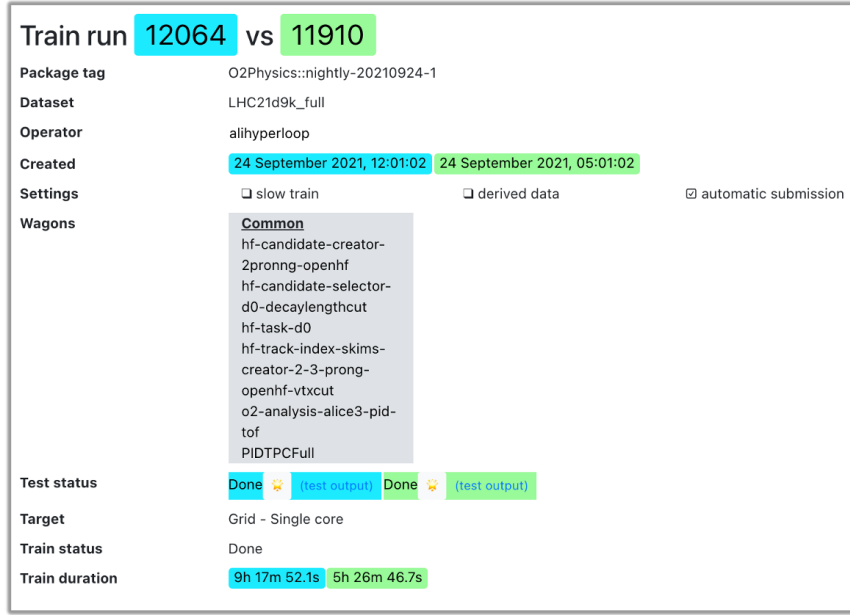


Figure 33: Train run comparison. Comparison of two different train run configurations. One train run is on cyan, and the other green. The common parameters are displayed on gray. Direct links to the specific timestamp dataset and wagons are available.

4.4.2 Automatic Composition

Automatic train composition can be defined in a dataset configuration, with the options “full” and/or “scheduled”. If “full” then a train run is composed if the memory consumption is full, i.e. the train run memory consumption is at least 75% that of the target facility memory. If “scheduled”, a train run is composed at the scheduled time regardless of the train run being “full”. This feature takes into account parameters such as the maximal number of trains per analysis per week and maximal CPU time. In addition to that, the same parameters as in the manual train submission are required, i.e. package tag, target facility maximum memory, wagons dependencies compatibility and wagons configuration compatibility.

4.4.3 Analysis Facility Staging

A large dataset can contain a subset of smaller datasets, also known as linked datasets. In order to not waste computing resources when running an analysis with a large dataset, running successfully over a smaller dataset is required. If the train run over a small dataset succeeds, any user from the analyses involved can request to run over a larger dataset upon approval of the PWGs conveners from the analyses.

4.5 Hyperloop - current status

The Hyperloop train framework has been successfully implemented based on the solid concept of the LEGO trains. It has been used in the ALICE Analysis Challenge, which began in June 2020. It is also used to validate the Run 3 Monte Carlo. So far, 38 analyses have been created for the Hyperloop framework, where 58 analyzers are actively using it. Continuous optimization is being done along with the O² software upgrades. Analyzers from different PWGs have been testing the framework using converted Run 2 data. Their results have shown a factor 3-10 event throughput improvement from AliPhysics (LEGO trains) to O² (Hyperloop trains), further optimization is foreseen. To date, Run 2 converted data and ALICE 3 MC are available in Hyperloop, and more than 1700 train runs have been successfully submitted to the ALICE Grid and Analysis Facilities.

Thanks to all the automatic actions and improvements in the UI, i.e. unified trains, automatic wagon tests, automatic train composition, automatic submission and algorithms computing the wagons compatibility, there is no longer need to have a huge number of train operators for the UI maintenance as in the LEGO trains, instead, there will be global centralized submission with central train operators.

5 Analysis Details

As mentioned in Section 2, different methods have been used in order to investigate the effect of longitudinal fluctuations as a probe of the QGP. This thesis focuses on decomposing EbyE pseudorapidity distributions using Legendre polynomials for single events measured with the ALICE detector. Based on the methods described in [bzdak2013] and [jia2016], longitudinal fluctuations can be extracted for a single particle distribution $R(\eta)$ by replacing its definition in Equation 8, for a range $[-Y, Y]$

$$R(\eta) = 1 + \sum_n^{\infty} a_n T_n(\eta) = 1 + \sum_n^{\infty} a_n \sqrt{n + \frac{1}{2}} P_n\left(\frac{\eta}{Y}\right) \quad (10)$$

$$R(\eta) - 1 = \sum_n^{\infty} a_n \sqrt{n + \frac{1}{2}} P_n\left(\frac{\eta}{Y}\right) \quad * \int P_m\left(\frac{\eta}{Y}\right) d\left(\frac{\eta}{Y}\right)$$

$$\int (R(\eta) - 1) P_m\left(\frac{\eta}{Y}\right) d\left(\frac{\eta}{Y}\right) = \sum_n^{\infty} a_n \sqrt{n + \frac{1}{2}} \int P_n\left(\frac{\eta}{Y}\right) P_m\left(\frac{\eta}{Y}\right) d\left(\frac{\eta}{Y}\right)$$

applying Legendre polynomials orthogonality, i.e.

$$\int_{-1}^1 P_n(x) P_m(x) dx = \frac{2}{2n+1} \delta_{nm}$$

$$\int (R(\eta) - 1) P_m\left(\frac{\eta}{Y}\right) d\left(\frac{\eta}{Y}\right) = \sum_n^{\infty} a_n \sqrt{n + \frac{1}{2}} \left(\frac{2}{2n+1}\right) \delta_{nm}$$

$$\frac{1}{Y} \int (R(\eta) - 1) P_n\left(\frac{\eta}{Y}\right) d\eta = a_n \left(n + \frac{1}{2}\right)^{-1/2}$$

$$a_n = \frac{1}{Y} \sqrt{n + \frac{1}{2}} \int (R(\eta) - 1) P_n\left(\frac{\eta}{Y}\right) d\eta \quad (11)$$

As a result, $R(\eta)$ from the observable gives as a_n^{obs} . In addition to that, two possible background

subtractions (a_n^{ran}) have been analyzed,

- *Uniform background*: random events generated from a uniform random distribution within the pseudorapidity window sampled over the average events multiplicity $\langle N(\eta) \rangle$
- *Pseudorapidity based background*: random events generated from an EbyE pseudorapidity distribution within the pseudorapidity window, and sampled over the average events multiplicity $\langle N(\eta) \rangle$

hence, the non-statistical component from the longitudinal fluctuations can be obtained after averaging over many events as:

$$\langle a_n^2 \rangle = \langle (a_n^{obs})^2 \rangle - \langle (a_n^{ran})^2 \rangle \quad (12)$$

5.1 Data Processing and Event Selection

This thesis analyses data from Run 2 period in the LHC collected by the ALICE detector. As mentioned in Section 3.2, the ALICE experiment is composed of different detectors. The data used for the analysis along this thesis comes from tracks collected by the TPC and ITS detectors. In general, the data collected by the ALICE detectors is stored in the ALICE Grid in a format known as Event Summary Data (ESD). ESDs contain all the information from the raw data, e.g. primary vertex, tracking, secondary vertices, etc. Since ESDs contain a large amount of information, they occupy a large amount of memory. If all users from the ALICE collaboration use ESDs for their analysis, this will imply an extremely high consumption of computer resources. In order to avoid the waste of computer resources, as well as time consumption, the concept of Analysis Object Data, AOD, was introduced in ALICE. An AOD contains only relevant and lighter analysis-related information of an ESD, which is more manageable in terms of memory. Another computing resource introduced in ALICE are NanoAODs, which are AODs, but with less information. The latter is not used in this thesis.

During the LHC data taking periods, the ALICE detectors record the passage of particles made by the collisions in a particular beam crossing, known as an event. In an event, often one collision dominates the signature in the detector. The uninterrupted collection of events is known as a run. The runs are grouped into periods, which refer to the lapse of time that the collisions are ongoing throughout the LHC. Furthermore, MC production for each period has been created by the ALICE Data Preparation Group (DPG). ESDs and AODs are available for MC data as well.

The collision systems analyzed throughout this thesis are Xe–Xe at $\sqrt{s_{\text{NN}}} = 5.44$ TeV and Pb–Pb at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. During 2017, Xe–Xe collisions took place in the LHC for calibration purposes; there are 2 runs (280234 and 280235) taken in the period LHC17n. Since the longitudinal coefficients computed in this thesis are calculated for the first time in ALICE, using Xe–Xe collisions data represents a basic test environment, as the data stored for the 2 runs is about ~ 1.5 million events, which occupies around 6.461 TB, and hence, can be fully tested in a local computer before accessing the ALICE Grid as explained in Section 4.1 and Section 4. Furthermore, a Xe–Xe collision is a symmetric colliding system, like the Pb–Pb collision. Moreover, a Xe–Xe collision shows a significant contribution of triangular flow compared to that of a Pb–Pb collision where the elliptic flow dominates (Figure 7 [xexeflow2018]). The overall raw data and MC productions used for this analysis is exposed in Table 2.

System	Type	$\sqrt{s_{\text{NN}}}$ [TeV]	Period (or tag)	Runs
Xe–Xe	Raw	5.44	LHC17n	(280234 - 280235)
Xe–Xe	MC	5.44	LHC17j2 (HIJING)	
Pb–Pb	Raw	5.02	LHC18r	(295581 – 296623)
Pb–Pb	MC	5.02	LHC20e3a (HIJING minimum bias)	

Table 2: Run 2 raw and MC data used for this thesis.

The track reconstruction and particle identification are done by the TPC detector, and the ITS is used to improve the spatial and momentum resolution, as explained in Section 3.2. The latter helps to reject background from secondary tracks, which may originate from weak decays,

conversion, secondary hadronic interactions in the detector material, and pileup. Pileup events occur when more than one interaction has occurred within the same bunch crossing. The 2018 Pb–Pb data collected by the TPC contains pileup. Furthermore, the QGP probe analyzed in this thesis corresponds to soft probes, i.e. low momentum range, hence, the tracks considered for the analysis are those where $0.2 < p_T < 2.0$ GeV/c.

5.1.1 Multiplicity and Centrality

The number of particles that participated in a heavy-ion collision can be described by the impact parameter b , which is the distance between the centers of the two nuclei in the transverse plane (Section 2.1.1). Since b is not directly measurable, an experimental proxy of centrality is used to characterize the amount of nuclear overlap in the collisions. One of the typical features related to the collision centrality are the number of nucleons participating in the collision, N_{part} , and the number of binary nucleon-nucleon collisions, N_{coll} , among the participant nucleons. Participants refer to the nucleons which undergo at least one nucleon-nucleon collision, while those who do not are known as spectators.

The centrality is defined as the percentile of the hadronic cross section corresponding to a particle multiplicity, or an energy deposited. By coupling a MC Glauber simulation to a model of particle production, it is possible to obtain the produced particle multiplicity which can be compared to the one measured experimentally. Centrality classes can be defined based on the simulated multiplicity distribution. In ALICE, to reproduce the experimental multiplicity, the MC Glauber simulation is coupled to a negative binomial distribution (NBD) [**centALICE:2018tvk**]. Figures 35 and 36 show the distribution of V0 amplitudes fitted by a Glauber-NBD fit for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and Xe–Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV respectively. The multiplicity of the two colliding systems differs as the type of colliding nuclei is different, i.e. ^{208}Pb and ^{129}Xe .

For this thesis, the centrality classes that are used are: 0-5%, 5-10%, 10-20%, 20-30%, 30-40%,

40-50%, 50-60%, 60-70%. The multiplicity and centrality for Pb–Pb and Xe–Xe for these centrality classes is summarized in Table 3 [centpb2016][centxe2019][centALICE:2018tvk].

%	$\langle dN_{ch}/d\eta \rangle$ Pb–Pb	$\langle dN_{ch}/d\eta \rangle$ Xe–Xe
0-5	1942.5 ± 53.5	1167 ± 26
5-10	1585.5 ± 46	939 ± 24
10-20	1180 ± 31	706 ± 17
20-30	786 ± 20	478 ± 11
30-40	512 ± 15	315 ± 8
40-50	318 ± 12	198 ± 5
50-60	183 ± 8	118 ± 3
60-70	96.3 ± 5.8	64.7 ± 2

Table 3: Centrality classes for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and Xe–Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV. Data taken from [centpb2016][centxe2019][centALICE:2018tvk].

5.1.2 Background and pileup rejection

For the Xe–Xe collisions that took place in 2017, the interaction probability per bunch-crossing was sufficiently small that the chance of two hadronic interactions occurring within the integration time of the involved detectors, so-called pileup events, was negligible. However, that was not the case for the Pb–Pb collisions that took place in 2018. There are two main categories of pileup, same-bunch-crossing pileup and out-of-bunch (OOB) pileup. Same-bunch-crossing pileup means that two or more collisions occurred in the same bunch crossing. This is seen by all detectors, and can be removed at the event selection level with cuts based on multiple reconstructed vertices. On the other hand, out-of bunch pileup means one (or more) collisions occurred in bunch crossings different from the one which triggered the acquisition. This affects differently the detectors depending on their readout time, and can also be removed at the event selection level, but with cuts based on multiple reconstructed vertices and on correlations between number of tracks/tracklets/amplitudes in different detectors.

The presence of pileup in the collected data has some negative consequences. For instance, the degradation of tracking performance for events with extremely high occupancy in TPC due

to strong pileup of Pb–Pb collisions in TPC, as well as the degradation of electron ID via TPC dE/dx in Pb–Pb events with pileup in TPC. As a result, it is necessary to remove pileup from the Pb–Pb data. ALICE DPG has developed tools that take care of Pb–Pb pileup, available in *AliPhysics* (ALICE Analysis Repository). Furthermore, in order to have a realistic signal for model comparison, the Monte Carlo events produced for Pb–Pb include pileup that matches the TPC performance. For this thesis, libraries were developed in the *AliPhysics* framework using *AliRoot*, also the DPG methods were implemented to remove OOB pileup from Pb–Pb raw data and MC respectively [**dpgreport**]. These libraries are available in the *AliPhysics* repository, <https://github.com/alisw/AliPhysics/tree/master/PWGCF/EBYE/LongAsymmetry>.

The OOB removal methods mentioned are cuts based on correlations between number of tracks/tracklets/amplitudes in the SDD+SSD vs. TPC clusters. Multi-dimensional pile-up correction is carried out having the SDD+SSD vs. TPC cluster correlation fitted with a N order polynomial, and the terms *looser cut*, *even more looser cut*, and *very loose cut* refer to the zero order been varied from -8000., -12000. and -16000. respectively. These parameters are pileup cuts based on V0 past-future information, online vs. offline VOM correlation, and online vs. offline SPD Fast-OR signal correlation.

5.1.3 Detector non-uniformity

The analysis of longitudinal fluctuations is done based on the final particle pseudorapidity distribution. The ALICE detector acceptance changes depending on the primary vertex (PV) position along the beam axis (PVz). For the centrality range of interest of this thesis, 0 – 70%, non uniform track acceptance was observed in Xe–Xe data for $|PVz| > 8$ cm, see Figure 37. Hence, only tracks with $|PVz| < 8$ cm are considered for the analysis in both, Xe–Xe and Pb–Pb collisions.

5.1.4 Systematic uncertainties

The track finding in ALICE can be performed with various detectors. Systematic uncertainties can arise due to the quality of the track i.e. how close its momentum corresponds to the actual particle momenta, and whether that track is from the actual collision or background. By varying the track quality, as will be done, one can gauge the impact of detector resolution on the final physics observables that will be presented in the next chapter. For this thesis, the analysis is done with the so-called *Global tracks*, i.e. TPC+ITS tracks, and *Hybrid tracks*, i.e. tracks passing standard TPC+ITS cuts, with SPD hit request and golden χ^2 cut (maximum χ^2 for vertex constrained tracks). In ALICE data format, these tracks are called in terms of FilterBits (FB), where for Global tracks FB=96, and Hybrid tracks FB=768.

As mentioned in Section 3.2.1, the TPC has two readout chambers, where there is a total of 570132 pads that form the cathode plane of the multi-wire proportional chambers. In the radial direction, there are 159 of these pads. The charge signals in a pad row associated with a track is known as a cluster. As a result, due to the number of pads, a particle flying away from the collision point in the radial direction creates a track with a maximum of 159 clusters. Hence, the more clusters a track has, the better the track can be reconstructed [**Leermakers:2045797**].

Even though the track that a particle that travels through the TPC end plates leaves can be reconstructed from the clusters, due to limited detector efficiency, not all pad rows the track crosses will give a signal, i.e. the number of pad rows the track passes will be higher than the number of clusters. A pad row that does not give a signal is added to the number of clusters to obtain the number of crossed rows in the pad rows on both neighboring sides that do give a signal [**Leermakers:2045797**]. This parameter is referred as the number of TPC crossed rows. Therefore, in this thesis, the parameters that have been varied based on previous published Xe–Xe and Pb–Pb data analysis in ALICE, are the minimum number of TPC clusters N_{cl} and the minimum number of TPC crossed rows.

A summary of the systematic uncertainties that are used in this thesis analysis is available in Tables 4 and 5 for Xe–Xe and Pb–Pb data respectively.

Systematic	$\chi^2/NDF <$	TPC n crossed rows $>$	TPC NCls $>$	Tracks
Default	4	70	70	Global
1	3	.	.	.
2	.	100	.	.
3	.	120	.	.
4	.	.	80	.
5	.	.	.	Hybrid

Table 4: Summary of systematic checks applied for Xe–Xe raw data and MC in this thesis.

Systematic	$\chi^2/NDF <$	TPC n crossed rows $>$	TPC NCls $>$	Tracks	OOB pileup cut
Default	2.5	70	70	Global	remove all OOB pileup
1	4	.	.	.	.
2	.	100	.	.	.
3	.	120	.	.	.
4	.	.	80	.	.
5	.	.	.	Hybrid	.
6	.	.	.	.	no cut
7	.	.	.	.	looser cut
8	.	.	.	.	even more looser cut
9	.	.	.	.	very loose cut

Table 5: Summary of systematic checks applied for Pb–Pb raw data and MC in this thesis.

5.1.5 Non-flow contributions

Short range correlations may contaminate the signal coming from the shape fluctuations. The source of those effects may be due to resonance decays and local charge conservation [bozek2012]. These problems can be suppressed by studying the contribution of positive and negative charged tracks separately. In soft probes, the local charge-conservation mechanism is applied in the statistical hadronization that occurs after the hydrodynamic evolution, and thus, tracks coming from

resonances can be considered to have the same sign charge. In addition to that, the impact of resonance can be studied by varying the momenta range, $0.2 < p_T < 0.6$ [GeV/c] and $0.6 < p_T < 2.0$ [GeV/c], where tracks coming from resonances will correspond to the upper p_T range. Overall, due to local charge conservation effects, it is expected to have a weaker contribution to the magnitude of a_1 when the tracks are coming from resonances.

6 Results

6.1 Background method comparison

After applying the method mentioned in Section 5, i.e. Equations 11 and 12, the first order coefficients are calculated for Pb–Pb and Xe–Xe collisions. Two different background methods have been studied. Figures 38 and 39 display the comparison between the two different types of background subtraction methods for the two collision systems with the default parameters. It can be seen that regardless of the pseudorapidity window ($|\eta| < 0.8, 0.6, 0.4$) and the collision system, both of the background methods agree. Since the background based on the pseudorapidity distribution comes from the observed EbyE pseudorapidity distribution, and hence is realistic, for the rest of the analysis, all longitudinal coefficients will use this background subtraction only.

6.2 Systematic checks

As mentioned in Section 5.1.4, in order to analyze the systematic uncertainty, it is necessary to vary the track quality with different cuts from the detectors. These systematic checks were calculated for Pb–Pb and Xe–Xe collisions. For both collision systems, $|\eta| < 0.2$ has a bigger difference from the systematic checks vs. the default configuration as there is low statistics in this pseudorapidity window. Figures 40, 41, 42, and 43 show the comparison of the different systematic uncertainties in Xe–Xe collision with respect to its default configuration for charged tracks in different pseudorapidity windows, $|\eta| < 0.8, 0.6, 0.4, 0.2$.

Table 6 summarizes the contributions from the systematic check vs. the default configuration in Xe–Xe collisions, the percentages from the smallest and biggest ratio is given. From the sources of systematic errors, the one that has the biggest difference, is when using hybrid tracks vs. global tracks, which can be seen in Figures 40e, 41e, 42e and 43e. This difference is due to the acceptance of the tracks as hybrid tracks have an extra SPD hit request. Figure 48 displays the final result of a_1 for charged tracks in different pseudorapidity windows including systematic and statistical errors. As the pseudorapidity decreases, the magnitude of the asymmetry decreases as well. The latter is expected as the number of participants can be approximated to be directly proportional to η by comparing Equations 4 and 8.

$ \eta <$	0.8	0.6	0.4	0.2
$\chi^2/NDF < 3$ (%)	0.42 - 8.69	1.21 - 10.52	2.58 - 12.02	2.35 - 17.54
TPC n crossed rows > 100 (%)	0.07 - 1.45	0.02 - 4.63	0.26 - 3.91	0.33 - 6.63
TPC n crossed rows > 120 (%)	0.03 - 2.62	0.10 - 2.41	0.09 - 3.85	0.94 - 9.13
TPC NCl's > 80 (%)	0.07 - 1.44	0.37 - 4.79	0.25 - 4.04	0.24 - 6.47
Hybrid tracks (%)	0.76 - 18.15	0.16 - 11.09	0.02 - 16.44	0.20 - 17.45

Table 6: Summary of systematic error contribution for Xe–Xe collisions in different pseudorapidity windows.

For Pb–Pb collisions, Figures 44, 45, 46 and 47 show the comparison of the different systematic uncertainties in Pb–Pb collision with respect to its default configuration for charged tracks in different pseudorapidity windows. Figures 44f, 45f, 46f and 47f display a bigger difference due to the tracks acceptance. Figure 49 shows the final result of a_1 for charged tracks in different pseudorapidity windows including systematic and statistical errors. Table 7 summarizes the contributions from the systematic check vs. the default configuration in Pb–Pb collisions, the percentages from the smallest and biggest ratio is given.

$ \eta <$	0.8	0.6	0.4	0.2
$\chi^2/NDF < 4$ (%)	0.09 - 0.42	0.02 - 0.29	0.22 - 1.22	0.05 - 2.45
TPC n crossed rows > 100 (%)	0.03 - 0.61	0.13 - 0.78	0.28 - 1.15	0.43 - 2.01
TPC n crossed rows > 120 (%)	0.15 - 1.39	0.01 - 1.90	0.03 - 3.34	1.40 - 2.82
TPC NClS > 80 (%)	0.02 - 0.35	0.07 - 0.62	0.17 - 0.55	0.16 - 1.93
Hybrid tracks (%)	0.44 - 21.79	0.09 - 11.76	0.35 - 23.10	1.43 - 30.86
No pileup cuts (%)	0.24 - 0.95	0.001 - 0.83	0.002 - 0.88	0.06 - 3.31
Looser cut (%)	0.04 - 0.51	0.01 - 0.94	0.02 - 0.79	0.02 - 1.82
Even more looser cut (%)	0.07 - 0.89	0.02 - 0.54	0.08 - 1.20	0.09 - 2.24
Very loose cut (%)	0.02 - 1.01	0.06 - 0.93	0.08 - 0.82	0.43 - 1.74

Table 7: Summary of systematic error contribution for Pb–Pb collisions in different pseudorapidity windows.

Histograms

Analysis: Hyperloop Framework Test Analysis

Workflow: o2-analysis/tutorial-histograms

Dependencies:

Compare [selected](#) Unselect All

	Wagon is created by jgrosseo	30 June 2020, 20:43:59 ☒
	Subwagon base is created by jgrosseo	
	Wagon (configuration) is created by jgrosseo	30 June 2020, 20:44:37 ☐
	Configuration pt-histogram/pTCut of base subwagon is created by jgrosseo	
	Configuration etaphi-histogram/phiCut of base subwagon is created by jgrosseo	
	Wagon is updated by jgrosseo	03 September 2020, 09:55:59 ☒
	Wagon is updated by jgrosseo	04 September 2020, 15:22:41 ☐
	Wagon (configuration) is updated by jgrosseo	08 September 2020, 09:53:36 ☐
	Configuration pt-histogram/pTCut of base subwagon is updated by jgrosseo	
	Wagon is updated by jgrosseo	08 September 2020, 09:53:36 ☐
	Wagon is updated by jgrosseo	08 September 2020, 09:58:15 ☐
	Wagon is updated by jgrosseo	08 September 2020, 09:58:21 ☐
	Wagon (configuration) is removed by jgrosseo	11 May 2021, 11:35:15 ☐
	Configuration pt-histogram/pTCut of base subwagon is removed by jgrosseo	
	Configuration etaphi-histogram/phiCut of base subwagon is removed by jgrosseo	

Figure 34: Wagon history. Detailed change log of a wagon configuration, including the user name and time at which a change was made. On the right, by clicking on two check boxes, a user can compare two different timestamps configuration from the same wagon.

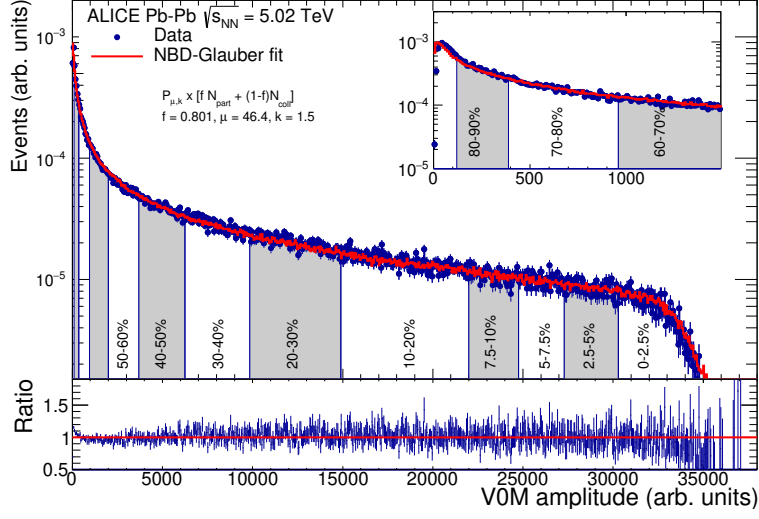


Figure 35: Distribution of the sum of amplitudes in the V0 scintillators for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The distribution is fitted with the NBD-Glauber fit shown as a line. The inset shows a zoom of the most peripheral region. Figure taken from [centALICE:2018tvk].

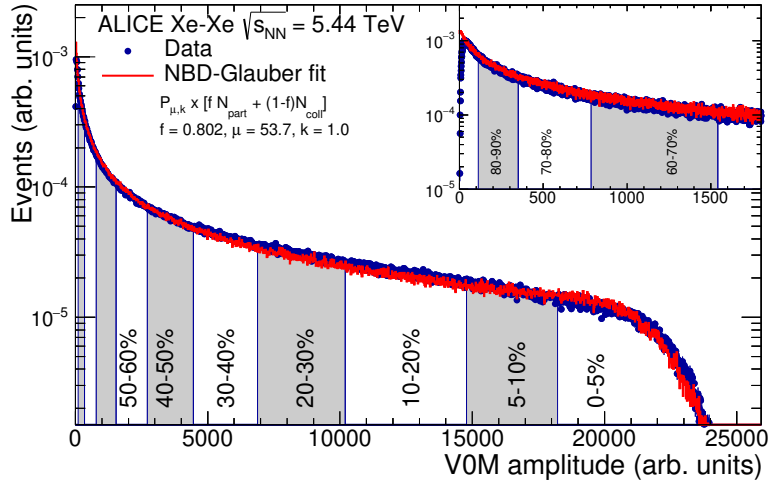


Figure 36: Distribution of the sum of amplitudes in the V0 scintillators for Xe–Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV. The distribution is fitted with the NBD-Glauber fit shown as a line. The inset shows a zoom of the most peripheral region. Figure taken from [centALICE:2018tvk].

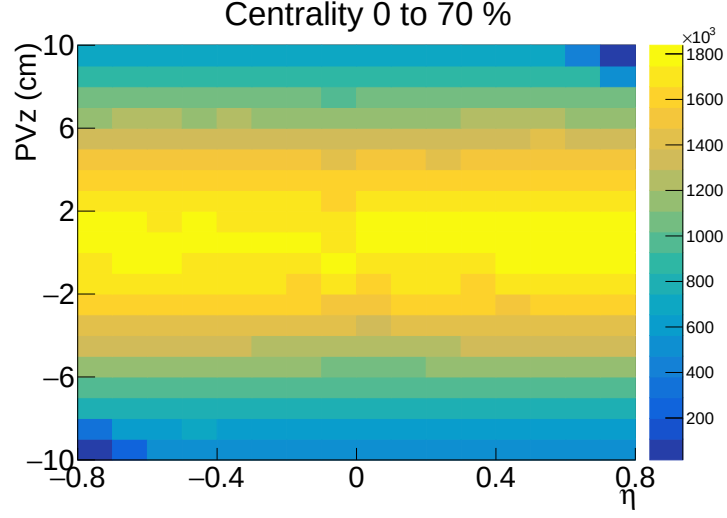


Figure 37: Primary vertex irregularity in Xe-Xe data

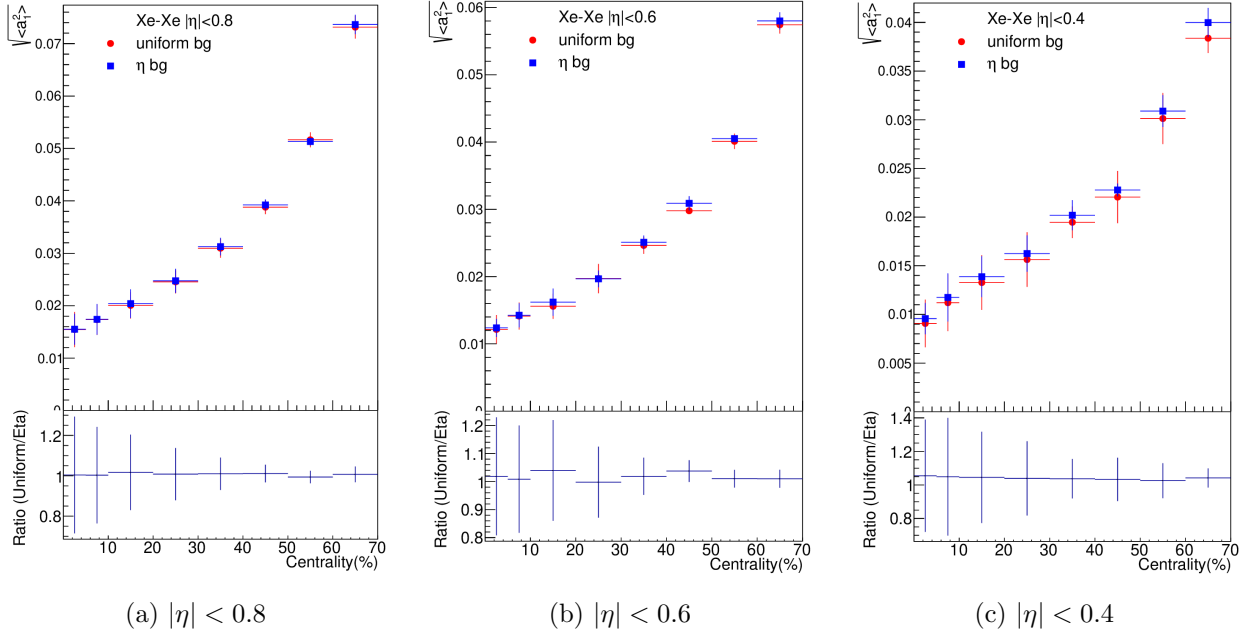


Figure 38: First order longitudinal coefficients, a_1 in Xe-Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. Comparison of different backgrounds. (Red) a_1 calculated using a random uniform background. (Blue) a_1 calculated using random pseudorapidity distribution based background.

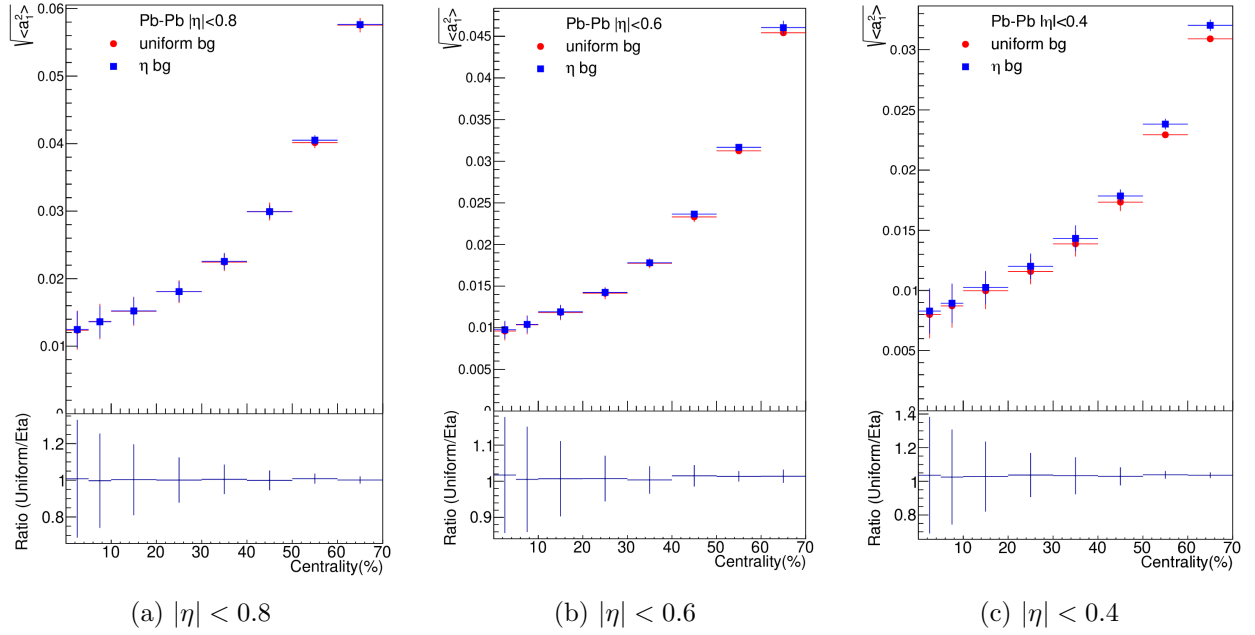


Figure 39: First order longitudinal coefficients, a_1 in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Period LHC18r. Comparison of different backgrounds. (Red) a_1 calculated using a random uniform background. (Blue) a_1 calculated using random pseudorapidity distribution based background.

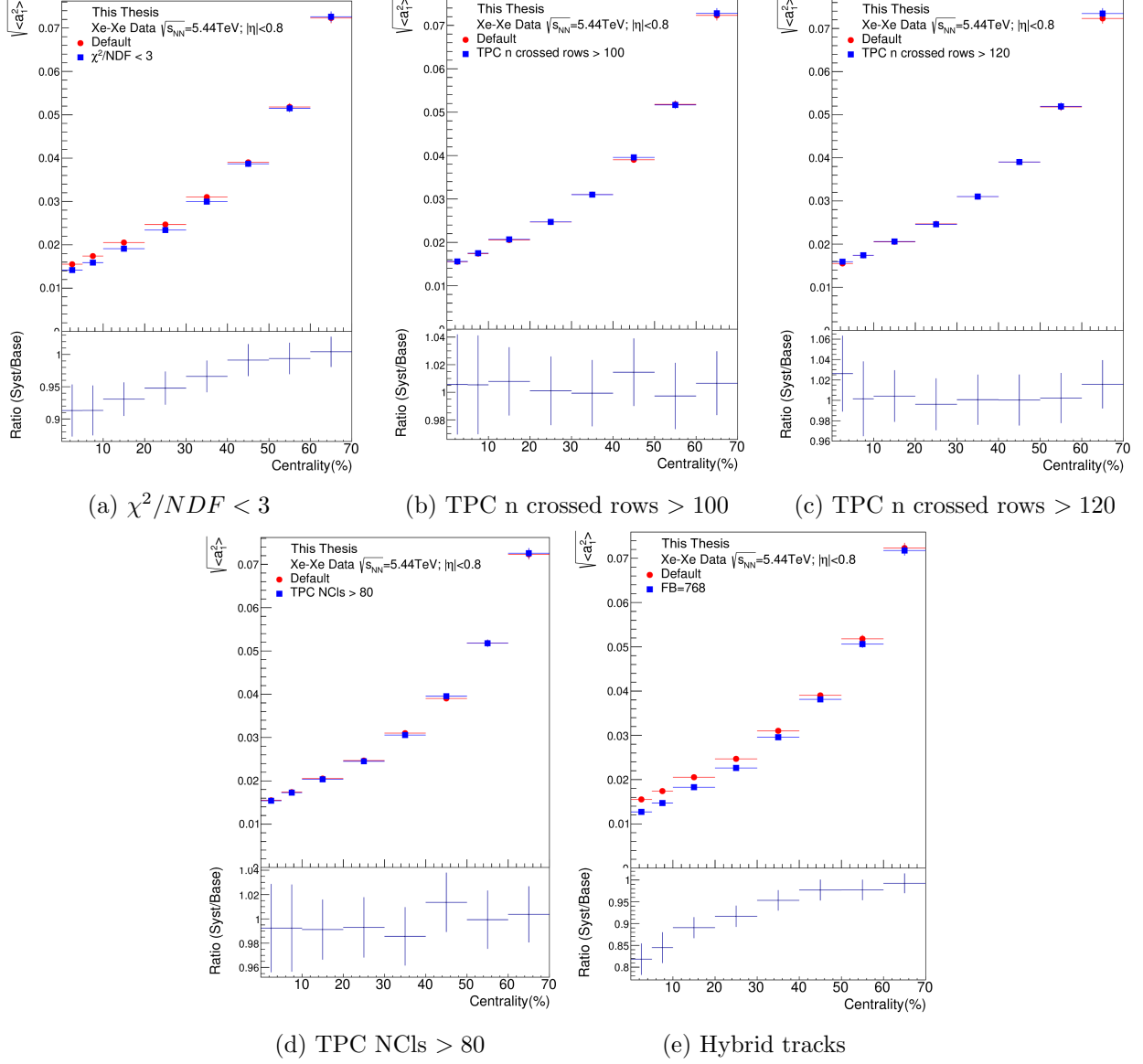


Figure 40: Comparison to different systematic uncertainties in Xe-Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV in a pseudorapidity window $|\eta| < 0.8$.

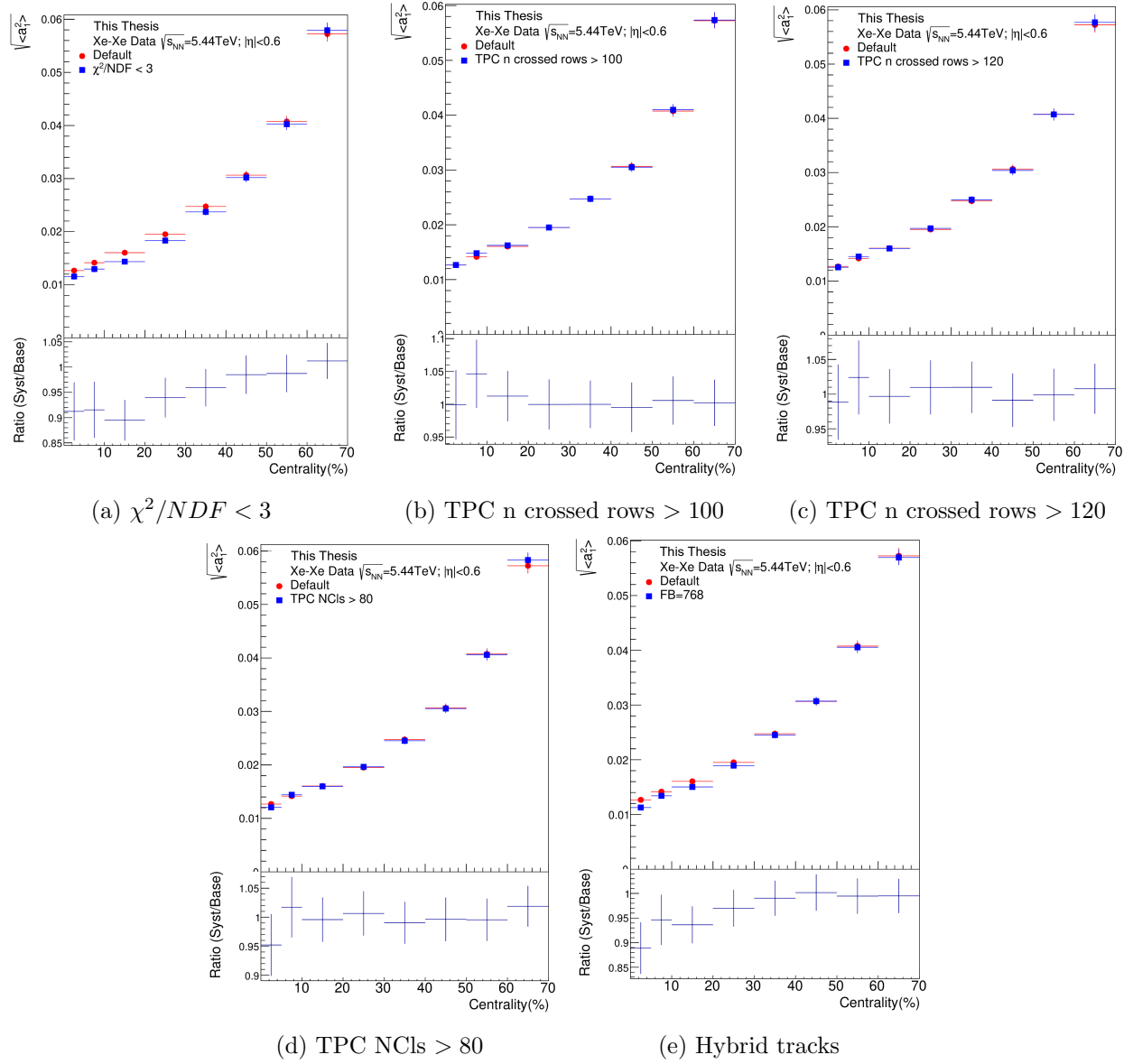


Figure 41: Comparison to different systematic uncertainties in Xe-Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV in a pseudorapidity window $|\eta| < 0.6$.

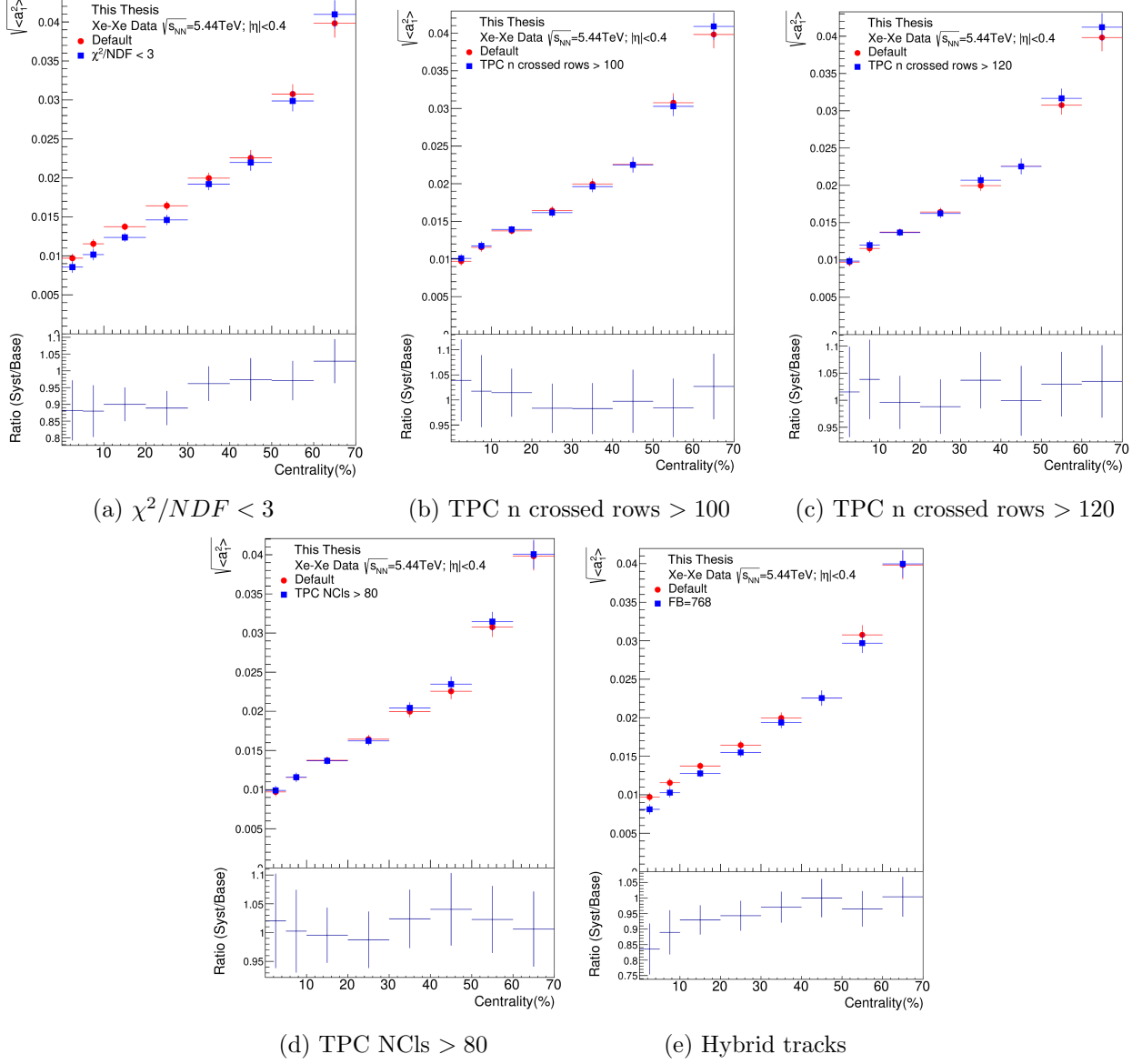


Figure 42: Comparison to different systematic uncertainties in Xe-Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV in a pseudorapidity window $|\eta| < 0.4$.

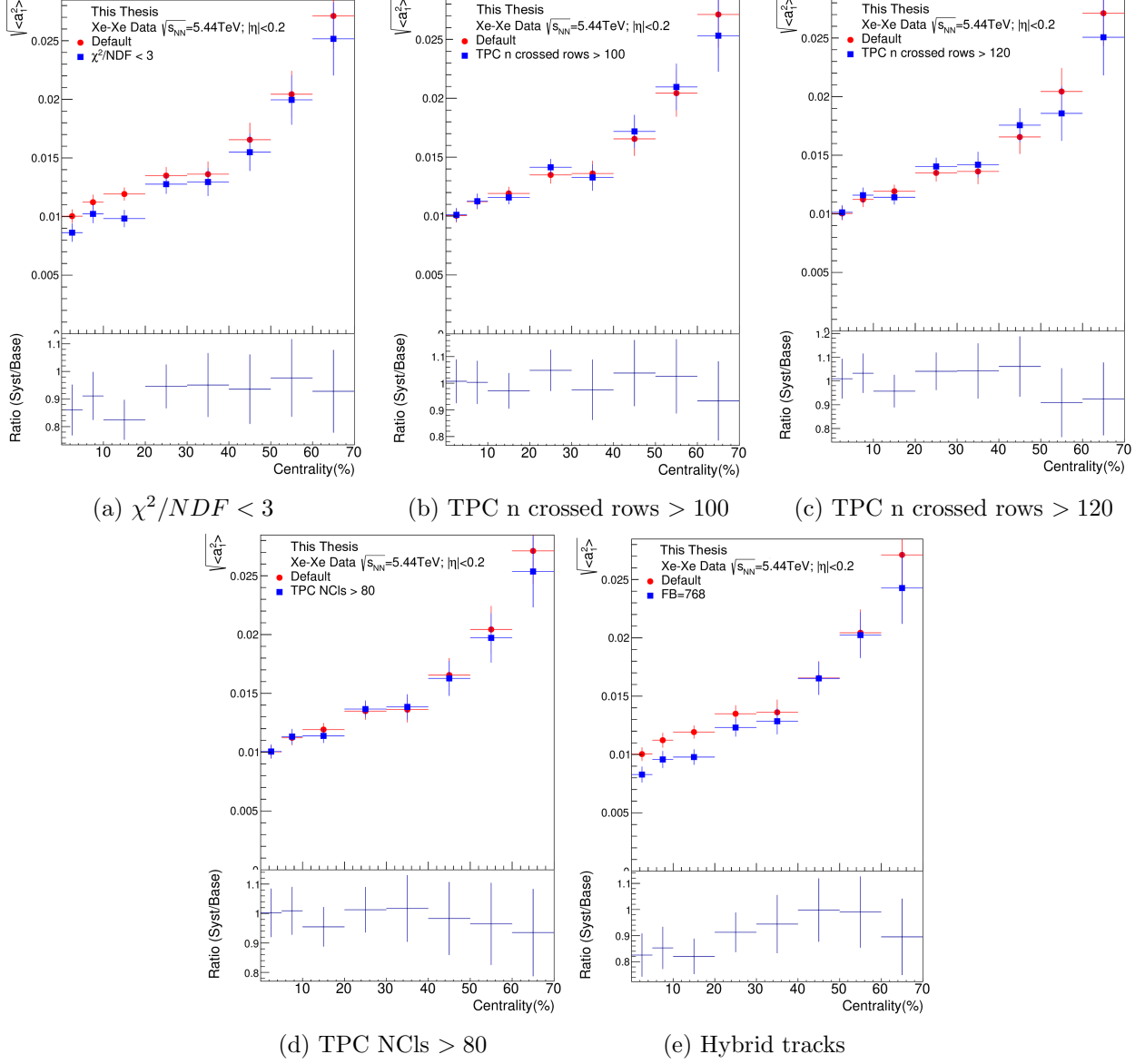
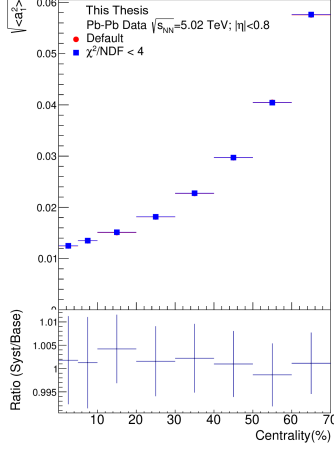
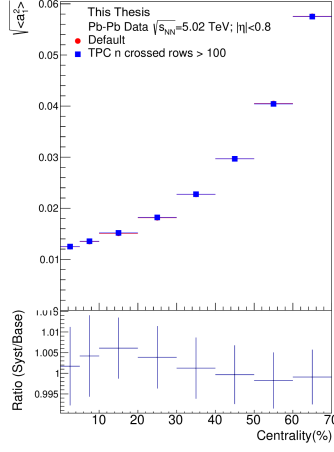


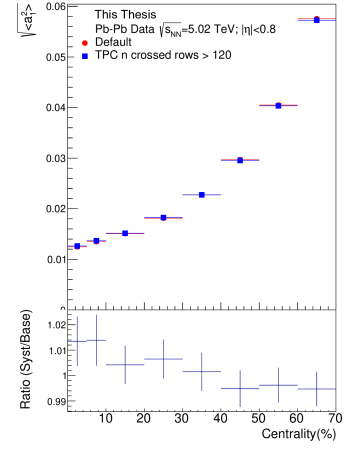
Figure 43: Comparison to different systematic uncertainties in Xe-Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV in a pseudorapidity window $|\eta| < 0.2$.



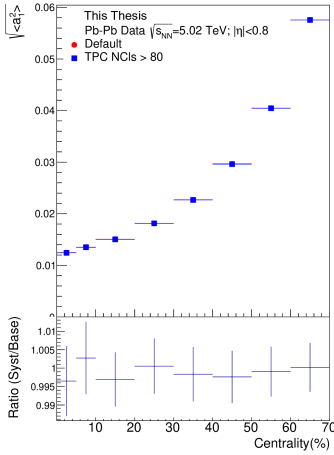
(a) $\chi^2/NDF < 4$



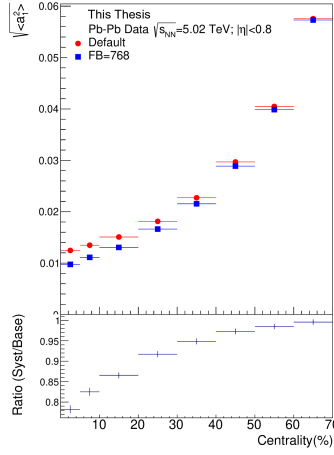
(b) TPC n crossed rows > 100



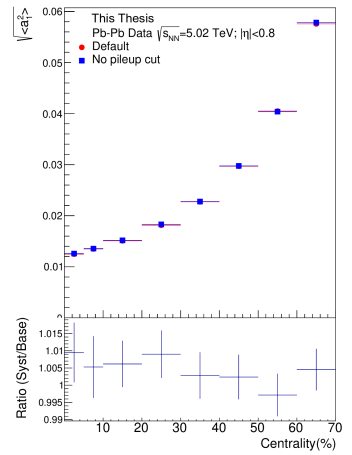
(c) TPC n crossed rows > 120



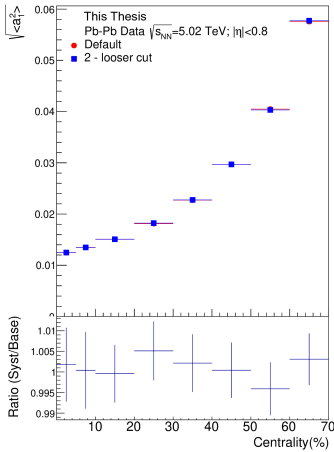
(d) TPC NCIs > 80



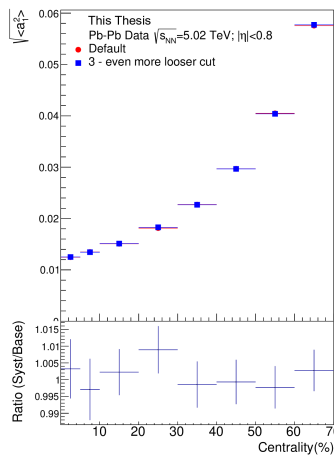
(e) Hybrid tracks



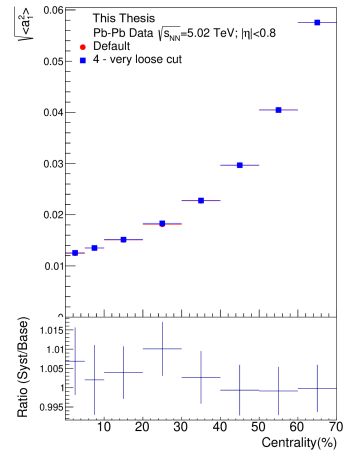
(f) No pileup cuts



(g) SDD+SSD clusters vs
TPC clusters - Looser cut

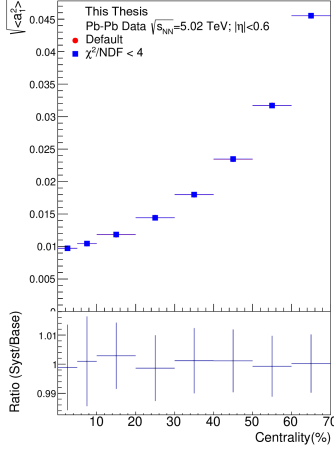


(h) SDD+SSD clusters vs
TPC clusters - Even more
looser cut

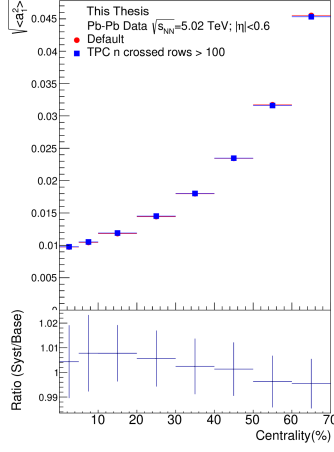


(i) SDD+SSD clusters vs
TPC clusters - Very loose cut

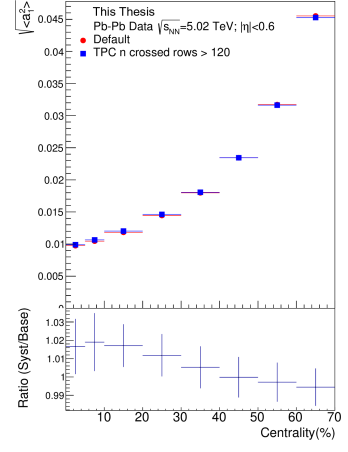
Figure 44: Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $|\eta| < 0.8$.



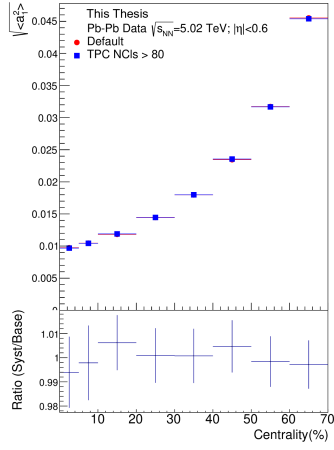
(a) $\chi^2/NDF < 4$



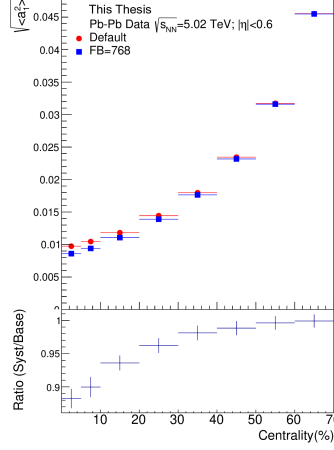
(b) TPC n crossed rows > 100



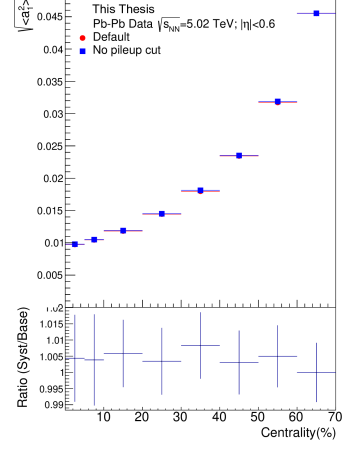
(c) TPC n crossed rows > 120



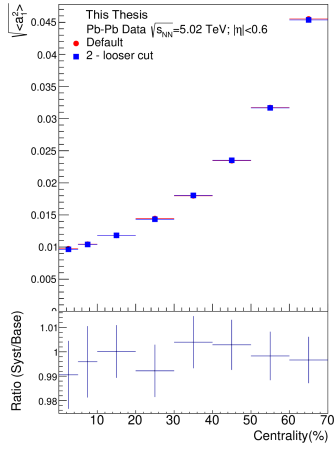
(d) TPC NCLs > 80



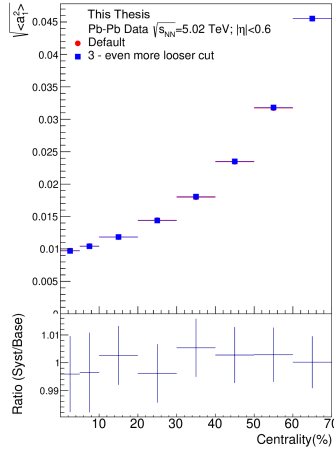
(e) Hybrid tracks



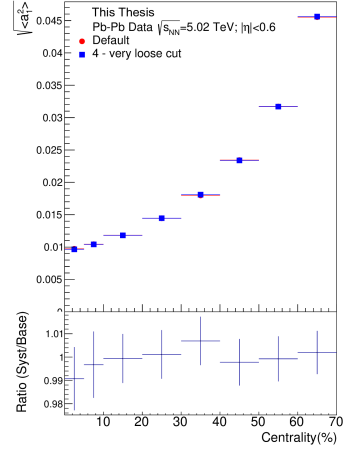
(f) No pileup cuts



(g) SDD+SSD clusters vs TPC clusters - Looser cut

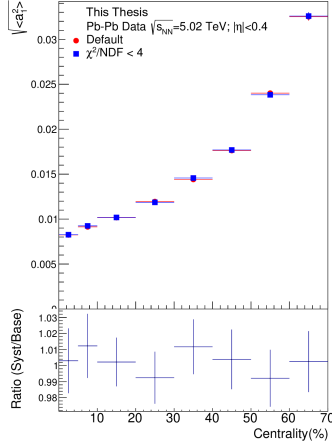


(h) SDD+SSD clusters vs TPC clusters - Even more looser cut

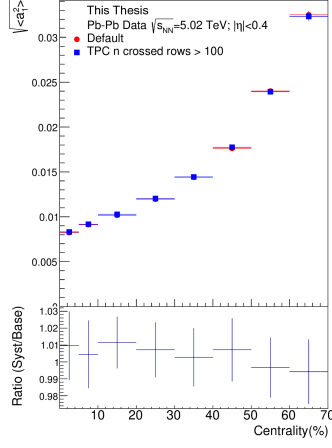


(i) SDD+SSD clusters vs TPC clusters - Very loose cut

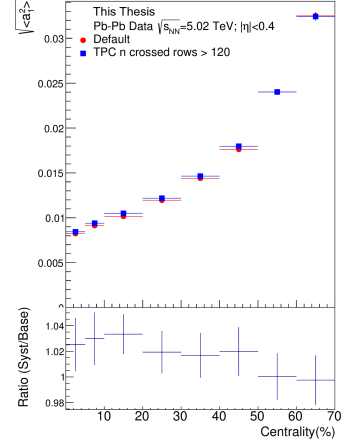
Figure 45: Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $|\eta| < 0.6$.



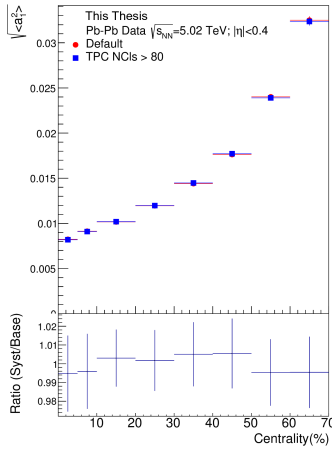
(a) $\chi^2/NDF < 4$



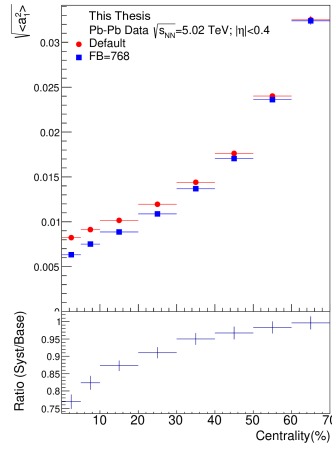
(b) TPC n crossed rows > 100



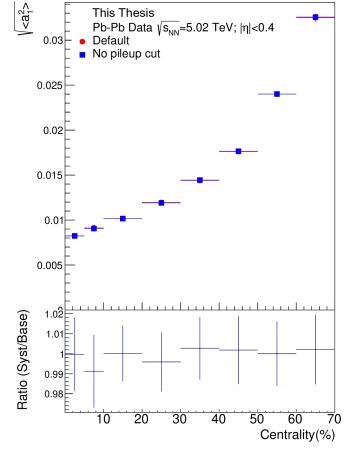
(c) TPC n crossed rows > 120



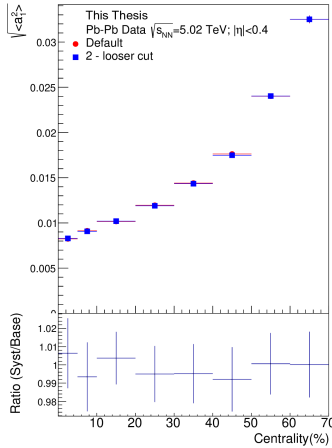
(d) TPC NCIs > 80



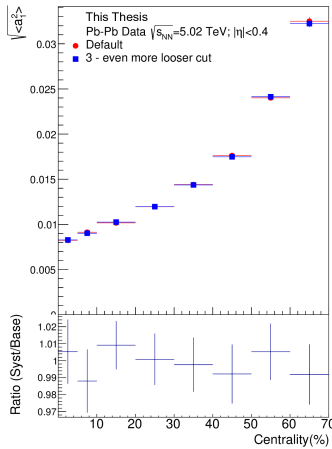
(e) Hybrid tracks



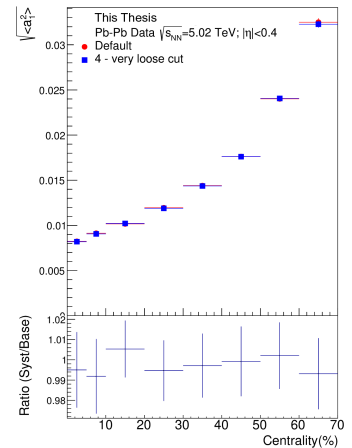
(f) No pileup cuts



(g) SDD+SSD clusters vs TPC clusters - Looser cut

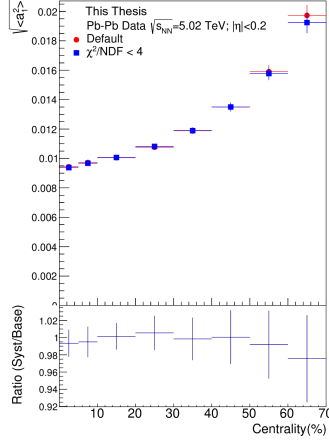


(h) SDD+SSD clusters vs TPC clusters - Even more looser cut

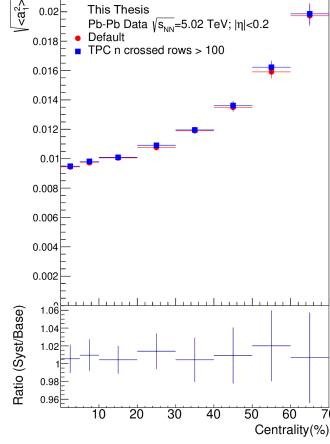


(i) SDD+SSD clusters vs TPC clusters - Very loose cut

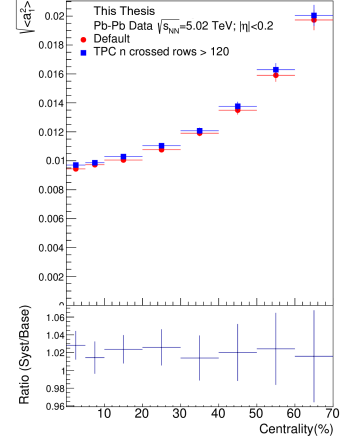
Figure 46: Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $|\eta| < 0.4$.



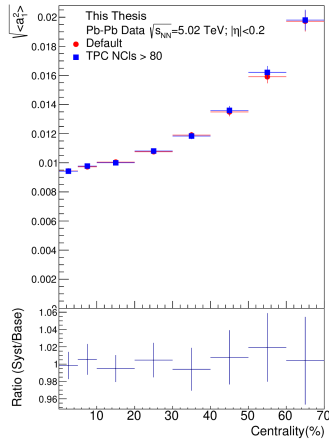
(a) $\chi^2/NDF < 4$



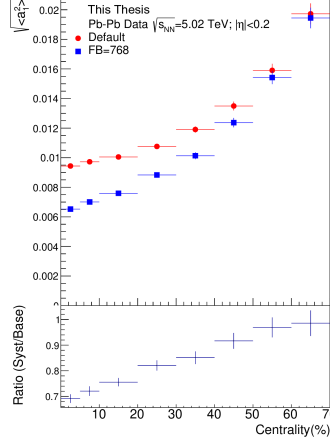
(b) TPC n crossed rows > 100



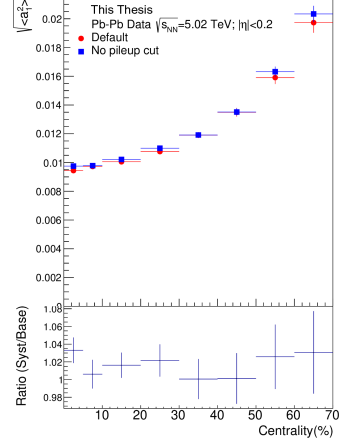
(c) TPC n crossed rows > 120



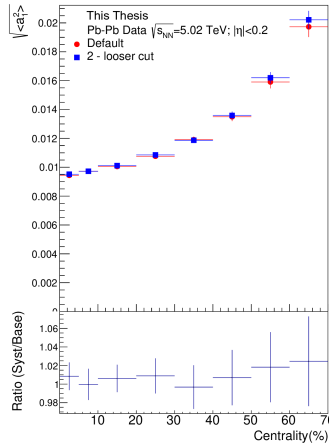
(d) TPC NCIs > 80



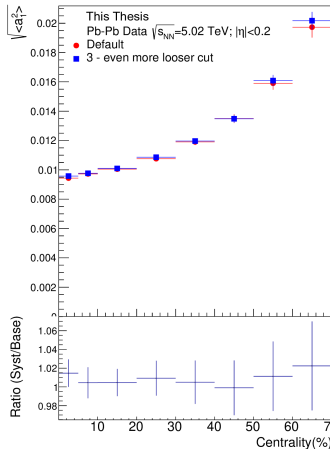
(e) Hybrid tracks



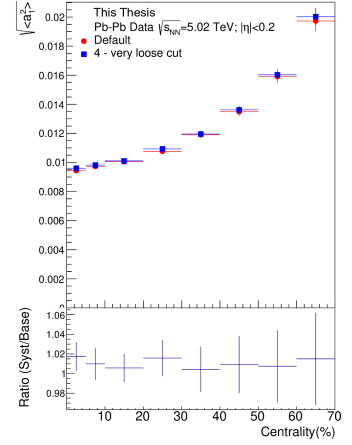
(f) No pileup cuts



(g) SDD+SSD clusters vs TPC clusters - Looser cut



(h) SDD+SSD clusters vs TPC clusters - Even more looser cut



(i) SDD+SSD clusters vs TPC clusters - Very loose cut

Figure 47: Comparison to different systematic uncertainties in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, period LHC18r, in a pseudorapidity window $|\eta| < 0.2$.

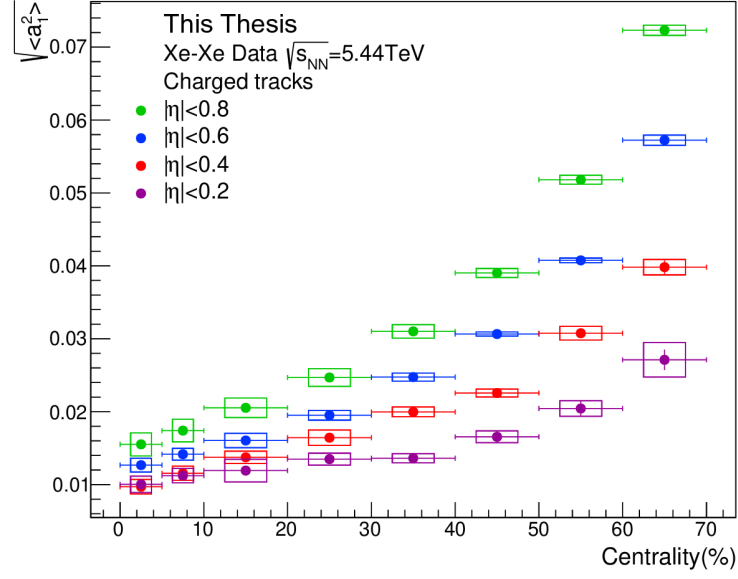


Figure 48: First longitudinal coefficient a_1 for Xe–Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV. Includes systematic (box) and statistical (line) errors.

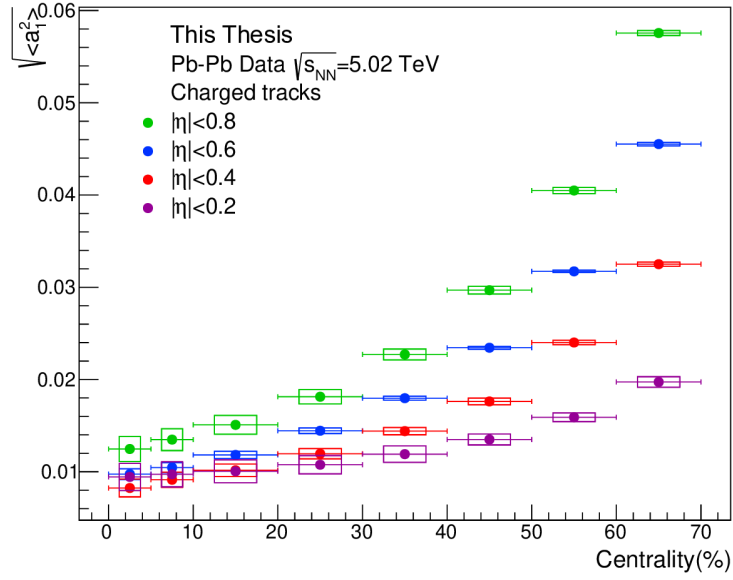


Figure 49: First longitudinal coefficient a_1 for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, period LHC18r. Includes systematic (box) and statistical errors (line).

Figure 50 displays a_1 vs. $\langle dN_{ch}/d\eta \rangle$ for different pseudorapidity windows, $|\eta| < 0.8, 0.6, 0.4, 0.2$. The values of $\langle dN_{ch}/d\eta \rangle$ are taken from Table 3. As mentioned in Section 2.3, Figure 16, ATLAS showed that at a given multiplicity N_{ch} , $\langle a_1^2 \rangle \sim 1/N_{ch}$, which is numerically similar for different collision systems [jia2017]. Figure 50 agrees with this scaling.

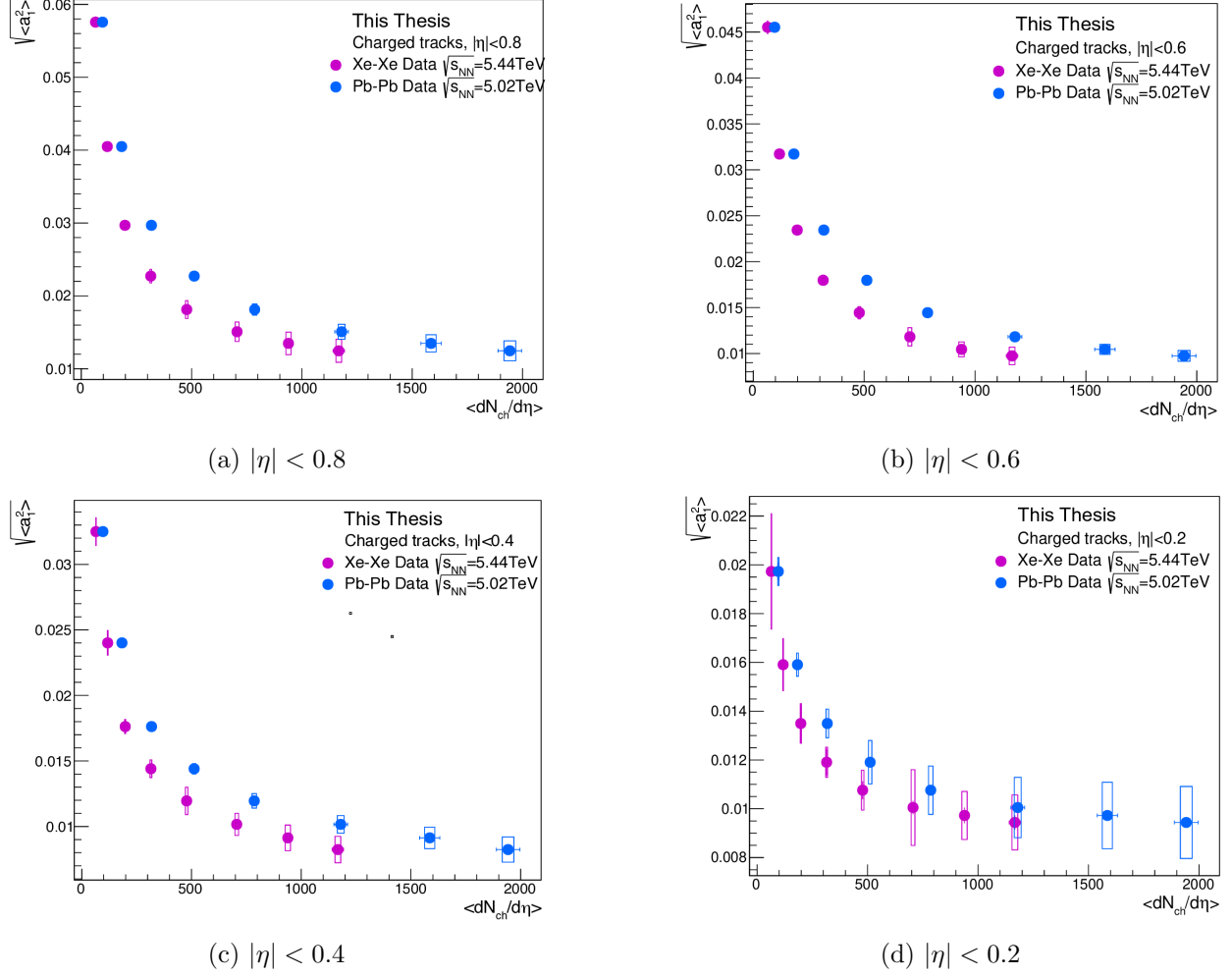


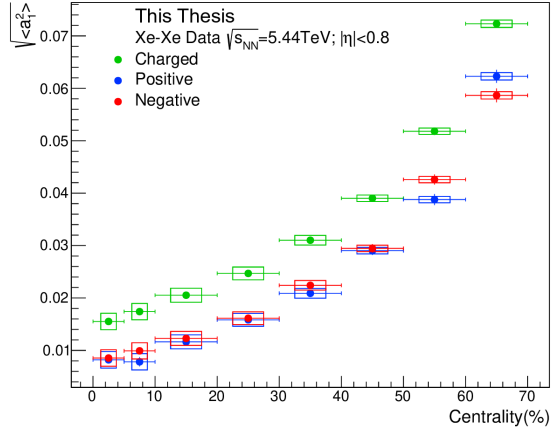
Figure 50: a_1 vs. $\langle dN_{ch}/d\eta \rangle$ in different pseudorapidity windows for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and Xe–Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV. Includes systematic (box) and statistical (line) errors.

6.3 Non-flow contributions

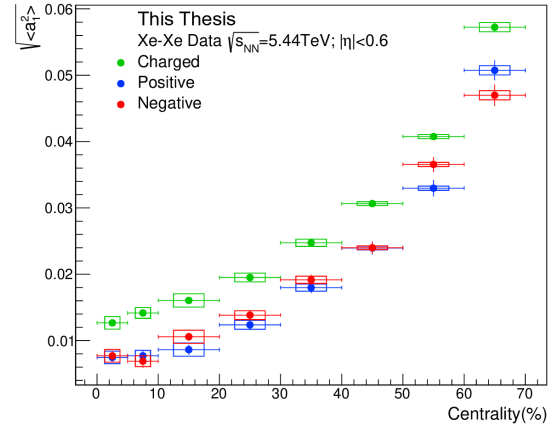
In order to analyze the effects due to non-flow, for instance, short correlations coming from resonances, the analysis has been implemented for single charged tracks, i.e. positive tracks only, or negative tracks only, and been compared to the charged tracks signal. Figure 51 shows the a_1 for charged, positive and negative tracks in different pseudorapidity windows for Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. In Figures 51c and 51d, the fact that the negative and positive tracks result is bigger than the one from charged tracks as it approaches peripheral collisions is justified due to low statistics from the Xe–Xe data (~ 1.5 million events).

On the other hand, Figure 52 shows a_1 for charged, positive and negative tracks in different pseudorapidity windows for Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. In Figure 52d, a_1 coming from positive and negative charged tracks are higher in magnitude due to low statistics.

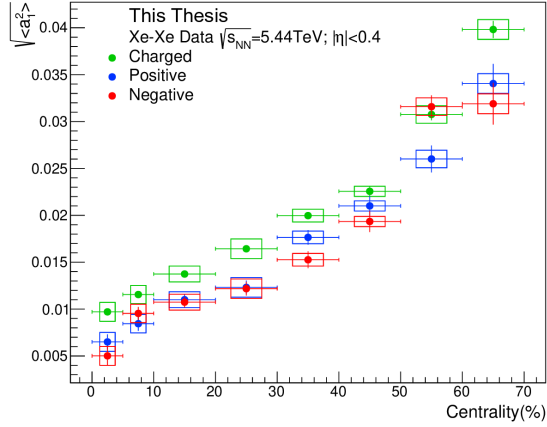
In addition to single charged hadrons, different p_T ranges are shown in Figures 53 and 54. Low momenta $0.2 < p_T < 0.6$ GeV/c and high momenta $0.6 < p_T < 2.0$ GeV/c. It can be seen that in central collisions, the magnitude of a_1 from low momenta charged tracks is bigger than that of higher momenta charged tracks.



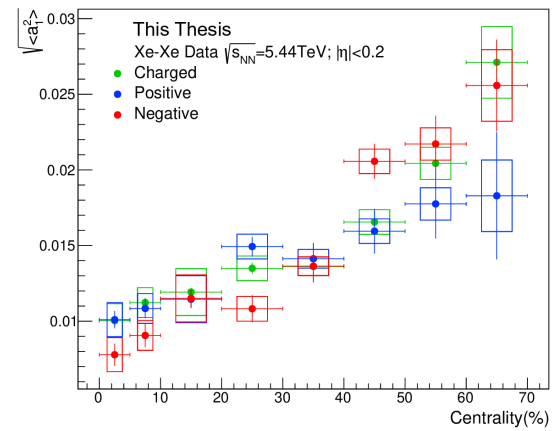
(a) $|\eta| < 0.8$



(b) $|\eta| < 0.6$

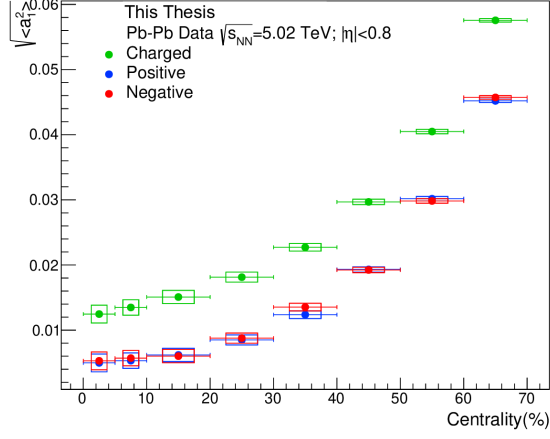


(c) $|\eta| < 0.4$

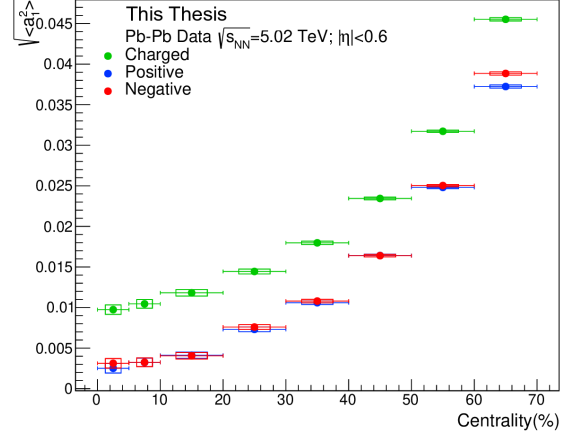


(d) $|\eta| < 0.2$

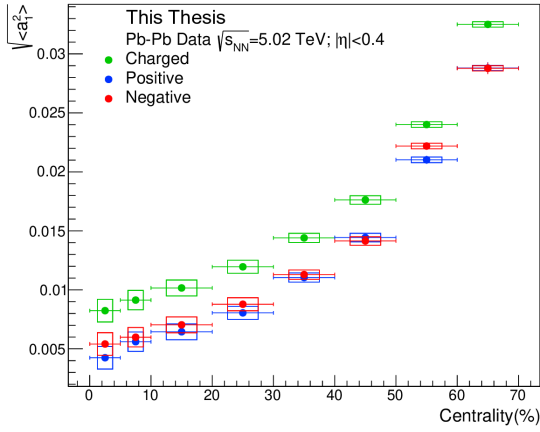
Figure 51: First longitudinal coefficient a_1 for Xe–Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV, period LHC17n. Comparison among charged, positive and negative tracks. Includes systematic (box) and statistical errors (line).



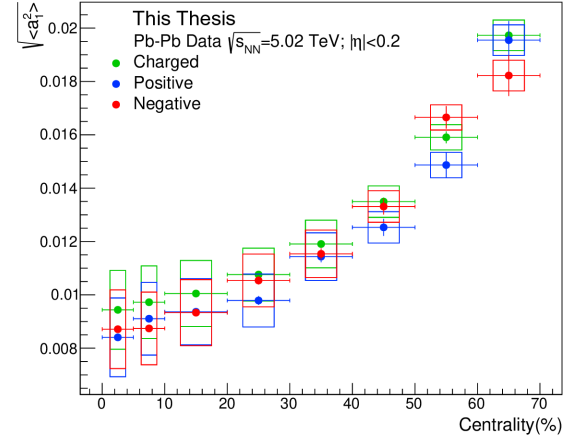
(a) $|\eta| < 0.8$



(b) $|\eta| < 0.6$

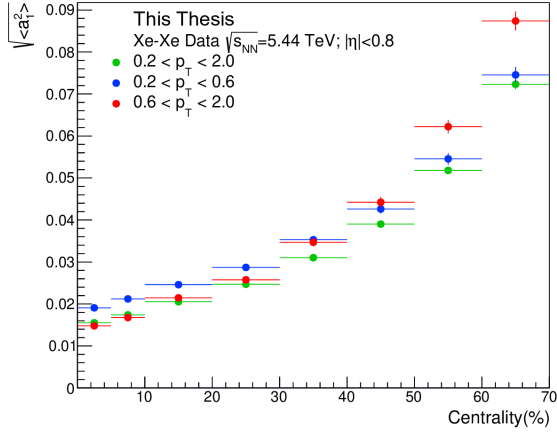


(c) $|\eta| < 0.4$

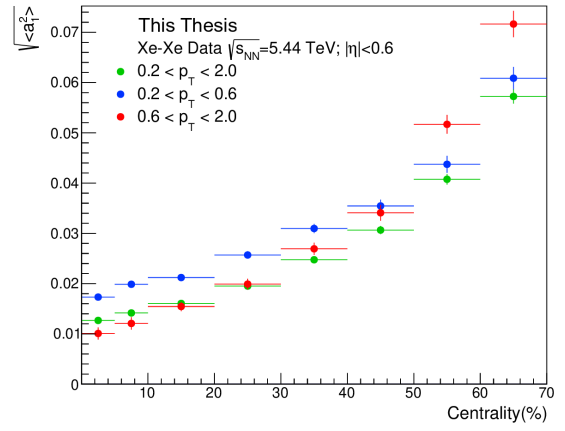


(d) $|\eta| < 0.2$

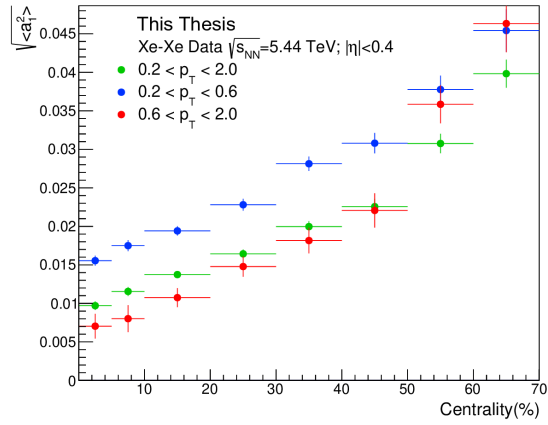
Figure 52: First longitudinal coefficient a_1 for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, period LHC18r. Comparison among charged, positive and negative tracks. Includes systematic (box) and statistical errors (line).



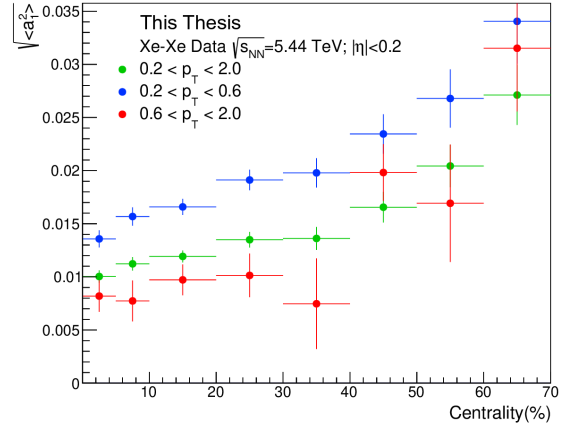
(a) $|\eta| < 0.8$



(b) $|\eta| < 0.6$

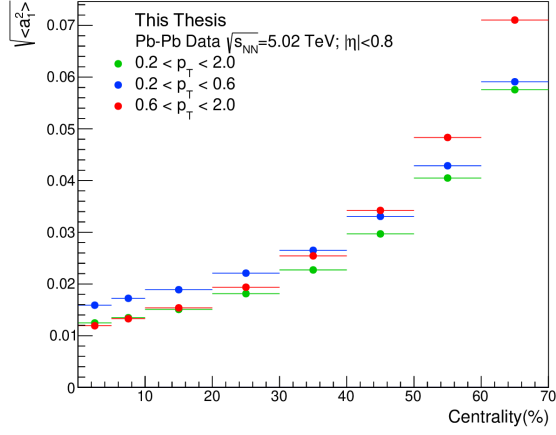


(c) $|\eta| < 0.4$

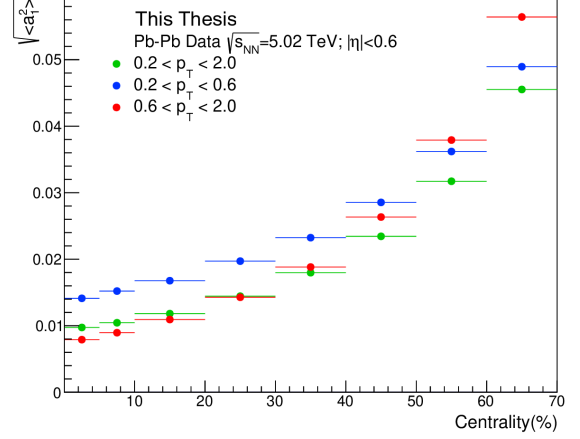


(d) $|\eta| < 0.2$

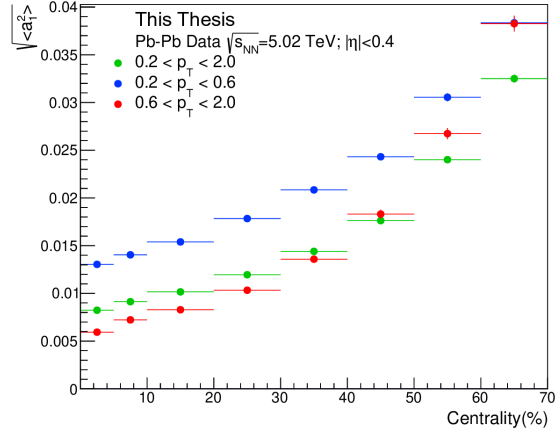
Figure 53: Comparison of different momenta ranges [GeV/c] in Xe–Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV, period LHC17n.



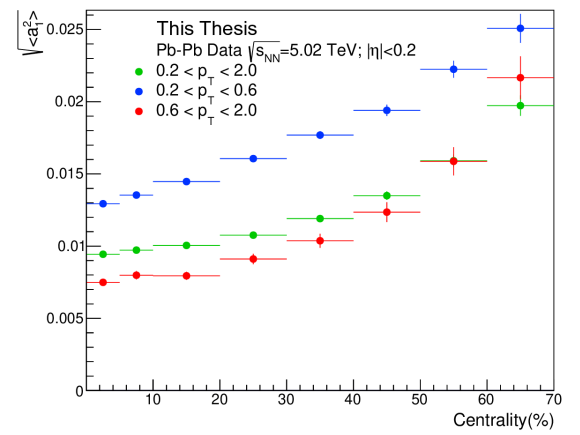
(a) $|\eta| < 0.8$



(b) $|\eta| < 0.6$



(c) $|\eta| < 0.4$



(d) $|\eta| < 0.2$

Figure 54: Comparison of different momenta ranges [GeV/c] in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, period LHC18r.

6.4 Comparison to Monte Carlo data

In order to validate the analysis method, MC closure tests are performed based on HIJING events. As it was seen in the previous plots, there is not enough statistics for $|\eta| < 0.2$, therefore, this pseudorapidity window will not be considered in the comparison to MC data.

Figure 55 shows the reconstructed vs. generated charged tracks for Xe–Xe MC HIJING data at $\sqrt{s_{NN}} = 5.44$ TeV. This MC data belongs to period LHC17j2. A difference of about 20% is expected as more generated events are created in order to have MC data. Figure 56 displays the comparison of charged, positive and negative tracks for Xe–Xe HIJING data. The a_1 coefficients obtained from Xe–Xe MC HIJING data is compared with Xe–Xe raw data from Figure 48 and presented in Figure 57.

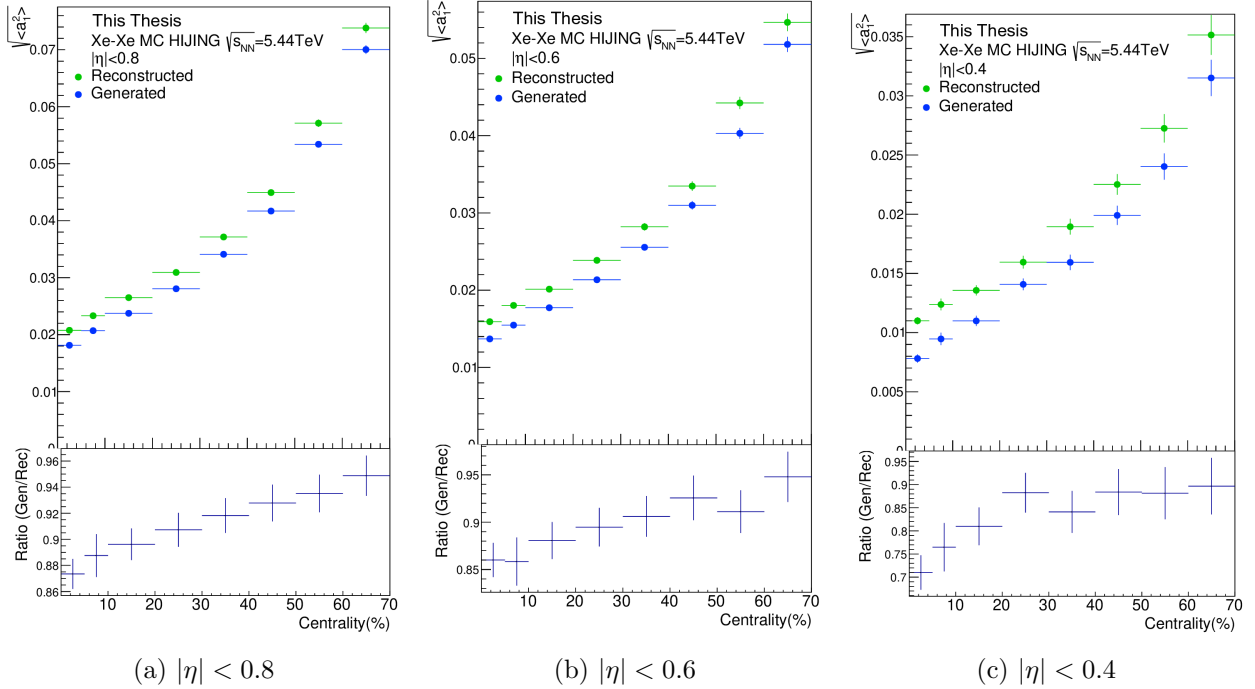
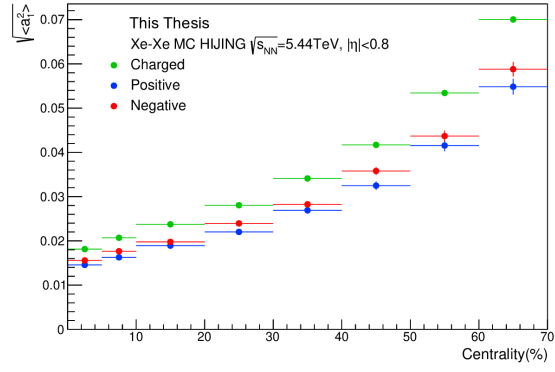
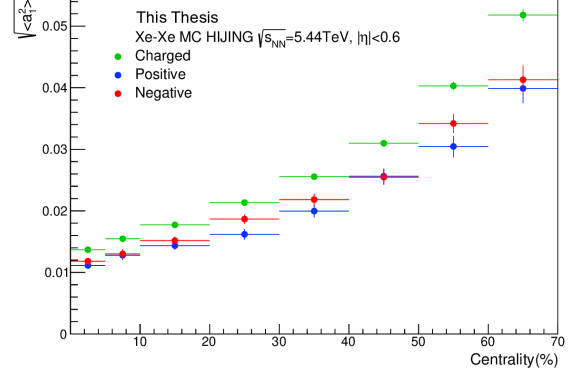


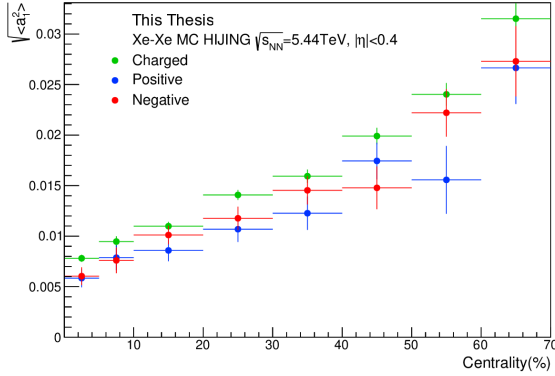
Figure 55: a_1 from Xe–Xe MC HIJING data in different pseudorapidity windows. Comparison of generated vs. reconstructed charged tracks.



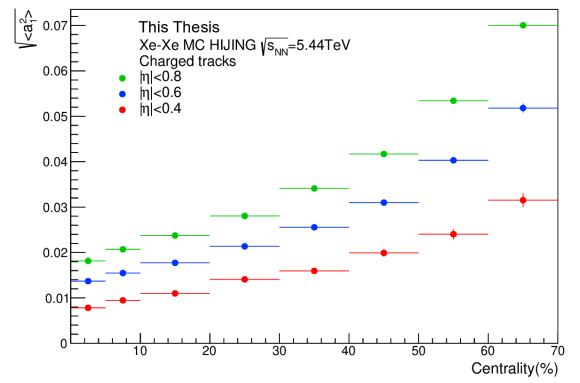
(a) $|\eta| < 0.8$



(b) $|\eta| < 0.6$



(c) $|\eta| < 0.4$



(d) Charged tracks

Figure 56: (a)(b)(c) a_1 from Xe-Xe MC Hijing data in different pseudorapidity windows for charged, positive and negative tracks. (d) Comparison among charged tracks in different pseudorapidity windows.

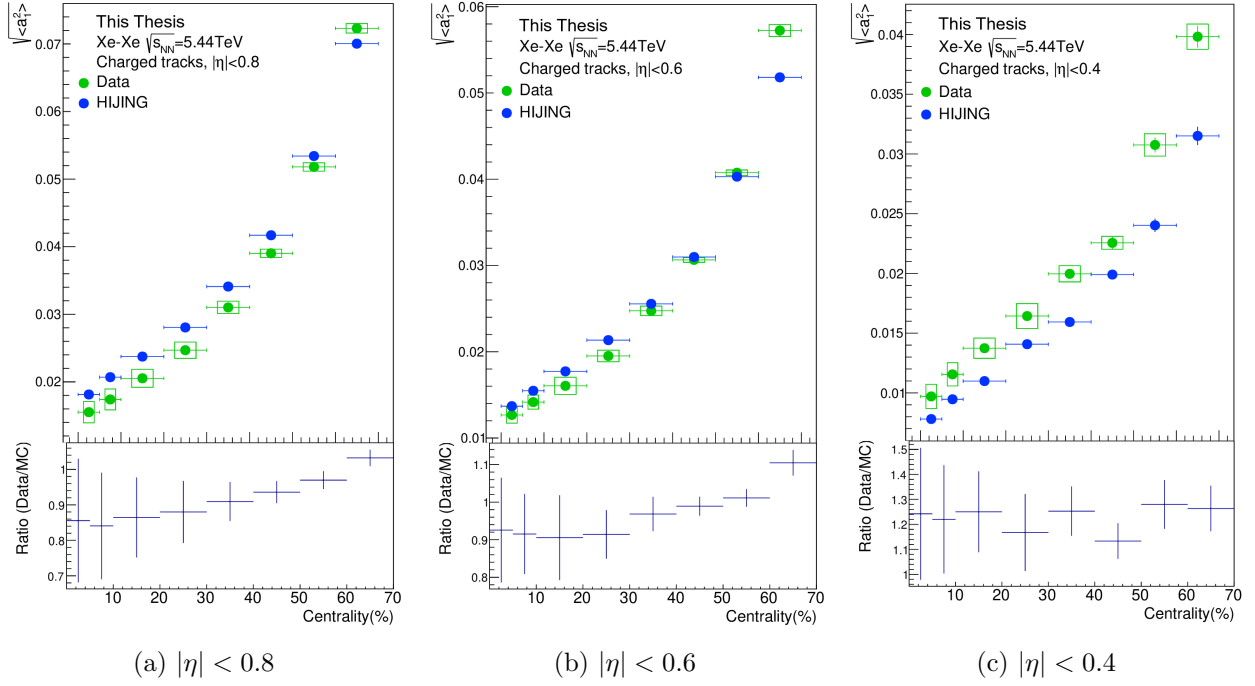


Figure 57: a_1 from Xe-Xe raw data and MC Hijing in different pseudorapidity windows. Comparison of data vs. MC charged tracks.

Likewise, Figure 58 displays the reconstructed vs. generated charged tracks for Pb–Pb MC HIJING data at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. This MC data belongs to period LHC20e3a. Figure 59 displays the comparison of charged, positive and negative tracks for Pb–Pb HIJING data. The a_1 coefficients obtained from Pb–Pb MC HIJING data is compared with Pb–Pb raw data from Figure 49 and presented in Figure 60.

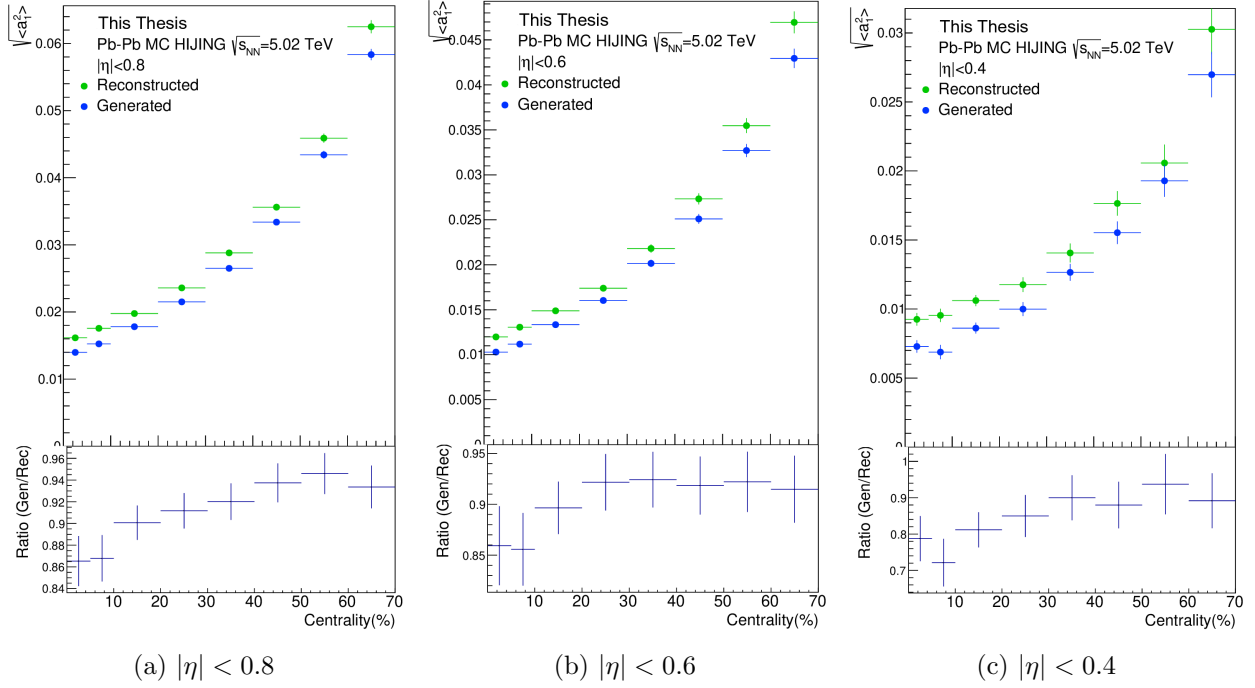
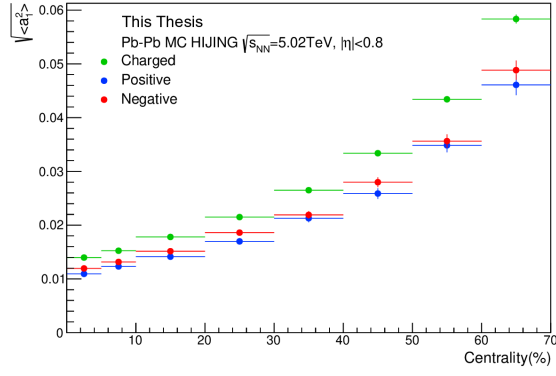
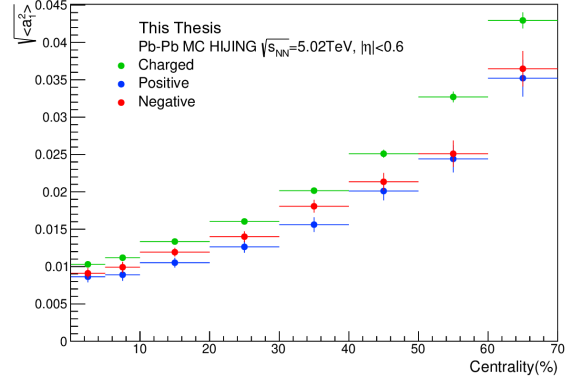


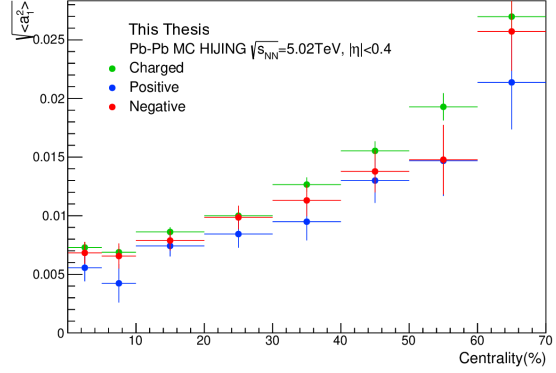
Figure 58: a_1 from Pb–Pb MC Hijing data in different pseudorapidity windows. Comparison of generated vs. reconstructed charged tracks.



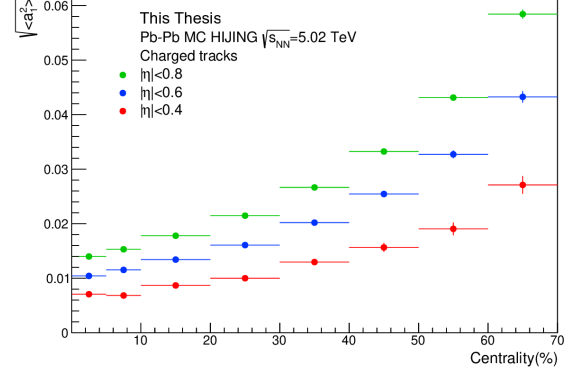
(a) $|\eta| < 0.8$



(b) $|\eta| < 0.6$



(c) $|\eta| < 0.4$



(d) Charged tracks

Figure 59: (a)(b)(c) a_1 from Pb–Pb MC Hijing data in different pseudorapidity windows for charged, positive and negative tracks. (d) Comparison among charged tracks in different pseudorapidity windows.

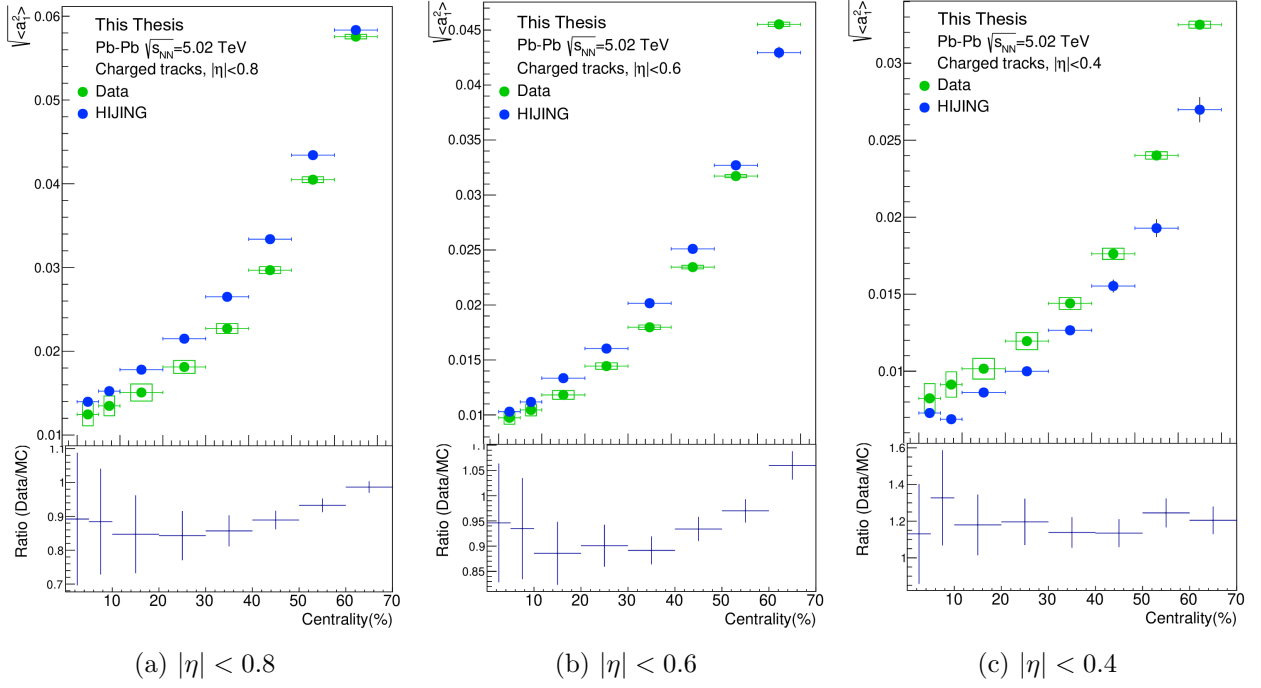


Figure 60: a_1 from Pb-Pb raw data and MC Hijing in different pseudorapidity windows. Comparison of data vs. MC charged tracks.

Finally, a summary of the analysis of the first order longitudinal coefficient (raw and MC data), a_1 is presented in Figure 61 including statistical and systematic errors. The values of $\langle dN_{ch}/d\eta \rangle$ are taken from Table 3. The asymmetry scaling of a_1 ($\langle a_1^2 \rangle \sim 1/N_{ch}$) agrees with the results obtained by ATLAS (Figure 16), regardless of the pseudorapidity window.

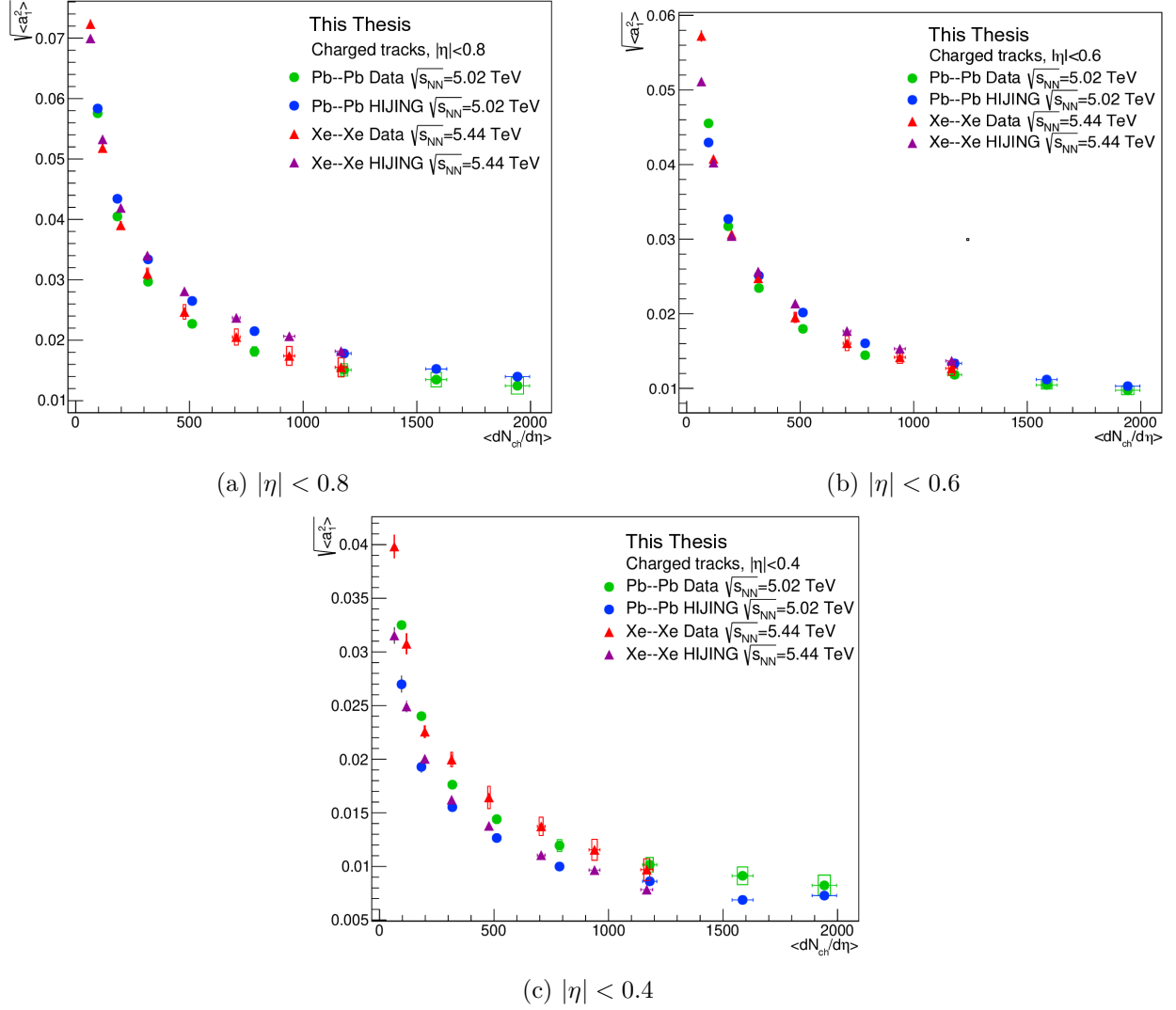


Figure 61: a_1 from Pb-Pb and Xe-Xe raw data and MC HIJING in different pseudorapidity windows. Statistical (line) and systematic (box) errors included.

7 Conclusion and Outlook

During LHC Run 1 and 2, ALICE developed an organized analysis framework, the LEGO train system, with the goal of increasing the CPU efficiency of the analysis jobs and thus to get more user analysis done with the same amount of computing resources hiding the complexity of the Grid to the users. The latter is done collecting multiple analysis tasks which should analyze the same dataset and by running them within the same analysis job. During LS2, ALICE went through major upgrades for the LHC Run 3 and 4. These upgrades implied the creation of a new computing system, ALICE O² software, which required the development of a new organized analysis framework. As a result, a new organized analysis framework has been successfully developed in ALICE, which will be used for LHC Run 3 and Run 4, the Hyperloop train system. Hyperloop has been developed based on the solid concept of the LEGO train system. Various improvements such as global train submission, change log of wagons and datasets, immediate automatic wagon testing, automatic composition, etc, have been implemented during the last two years. A first official version of Hyperloop was deployed as of June 2020, and has been used by ALICE users ever since. Furthermore, Hyperloop has evolved over the last months along with the development of the ALICE O² software. So far, over 1700 Hyperloop train runs have been successfully submitted to the ALICE Grid and Analysis Facilities. Run 2 converted data and Run 3 MC data are already available in the Hyperloop framework. Thanks to the new design of the train system, i.e. the development of algorithms that allow several automatic actions that no longer requires the intervention of a train operator, a global 24/7 centralized train submission is envisaged for LHC Run 3 and 4. As of January 2022, global centralized train submission will begin with the participation of institutes through ALICE Service Work: 1 USA, 1-2 Europe, 1 Asia. Further final development is foreseen for the Hyperloop framework during early 2022, such as Web App testing, with the goal of having a final version ready for Run 3 and Run 4.

Several analyses have been carried out in ALICE using the organized analysis framework. For

instance, the analysis of the QGP and the initial state of a heavy-ion collision at the LHC. Initial state models generate energy or density profiles prior to the time of the QGP formation in the overlap region of a heavy-ion collision. Hence, an accurate description of the initial state helps constrain the calculations of key QGP medium properties. As the QGP behaves like an almost ideal fluid, it has been seen that relativistic viscous hydrodynamic calculations using various initial state models can often describe the multiplicity distributions, particle spectra, and integrated flow measurements simultaneously. Over the past decades, the measurement of initial state azimuthal anisotropies have been extensively studied. In particular, the description of observed azimuthal anisotropies via Fourier decomposition, with the so known flow coefficients, ν_n . A comparison of the observed flow coefficients with initial state models (e.g. MC Glauber and TRENTo) have allowed to constrain QGP transport properties such as the shear and bulk viscosity to entropy ratio, η/s and ζ/s .

So far, various models that describe the initial state of a relativistic heavy-ion collision assume boost invariance, i.e. invariance along the longitudinal direction. However, by attempting to quantify the longitudinal fluctuations, a significant forward-backward asymmetry of the nucleon participants has been observed. This shows a breaking of boost invariance. A similar analysis technique as that of Fourier decomposition in azimuthal anisotropies has been proposed in previous analyses of longitudinal fluctuations. The study of a longitudinal observable like the rapidity or pseudorapidity via decomposition using orthogonal polynomials, like the Legendre polynomials, allows to quantify the forward-backward asymmetry with the first order coefficient of the decomposition, a_1 . Over the past decades, other experiments, like ATLAS, have extracted a_1 from two-particle correlations. This thesis has calculated the longitudinal coefficient a_1 from a single event pseudorapidity distribution for the first time in the ALICE collaboration using LHC Run 2 data. The a_1 coefficients were obtained for raw data and HIJING MC events in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV and Xe–Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV. The results obtained agree with previous calculations from other experiments, i.e. the magnitude of a_1 increase for peripheral collisions and there is a

universal scaling that at a given multiplicity, N_{ch} , follows $\langle a_1^2 \rangle \sim 1/N_{ch}$. Furthermore, non-flow contributions, such as the effect of resonances, were analyzed using two different approaches. The first one by studying the positive and negative charged tracks separately, and the second one by studying different ranges of transverse momenta, p_T , where the upper range of p_T would represent the tracks coming from resonances. A weak a_1 was observed for tracks coming from resonances, which was expected from local charge conservation effects.

The method developed throughout this thesis, now available in the ALICE analysis framework, *AliPhysics*, can be used in future research works to extract higher order longitudinal coefficients for any single event pseudorapidity distribution. The physics behind higher order coefficients is yet to be understood. Although the HIJING and AMPT models provide a reasonable description of the a_1 , other models, such as those that are CGC based, need to be tested to see if a better agreement is reached. This will help disentangle whether the asymmetries are described better in the context of gluon fluctuations, as opposed to nucleon fluctuations used for HIJING and AMPT.

A Overwatch

The ALICE Overwatch project (Online Visualization of Emerging tRends and Web Accessible de-Tector Conditions using the HLT) has been designed with the purpose of providing real-time online data monitoring and quality assurance using time-stamped data from the ALICE High Level Trigger (HLT) and Data Quality Monitoring (DQM). The Overwatch framework requires ROOT 6, and has been developed using Flask, a microframework for Python.

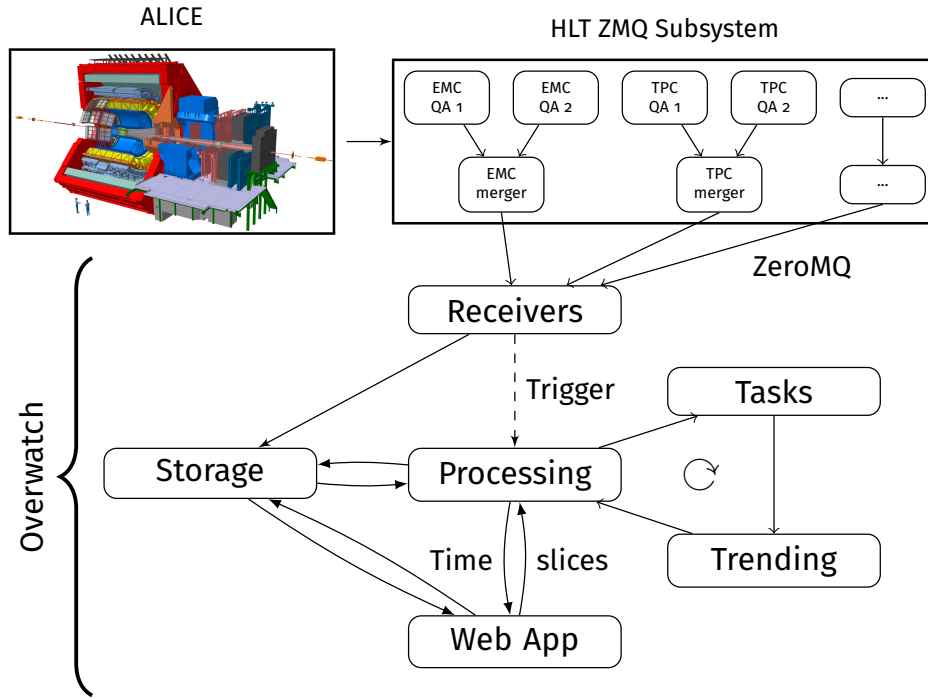


Figure 62: Overwatch architecture. Figure taken from [raymond'ehlers'2018'1309376].

The Overwatch architecture is shown on Figure 62. The incoming data is handled by the receivers, which then make data available to be processed by the processing module. The processing module aims to provide a user-friendly and viewable form, which has been developed using a module of ROOT, PyROOT. Finally, the output of the processing is then visualized via the WebApp, Figure 63.

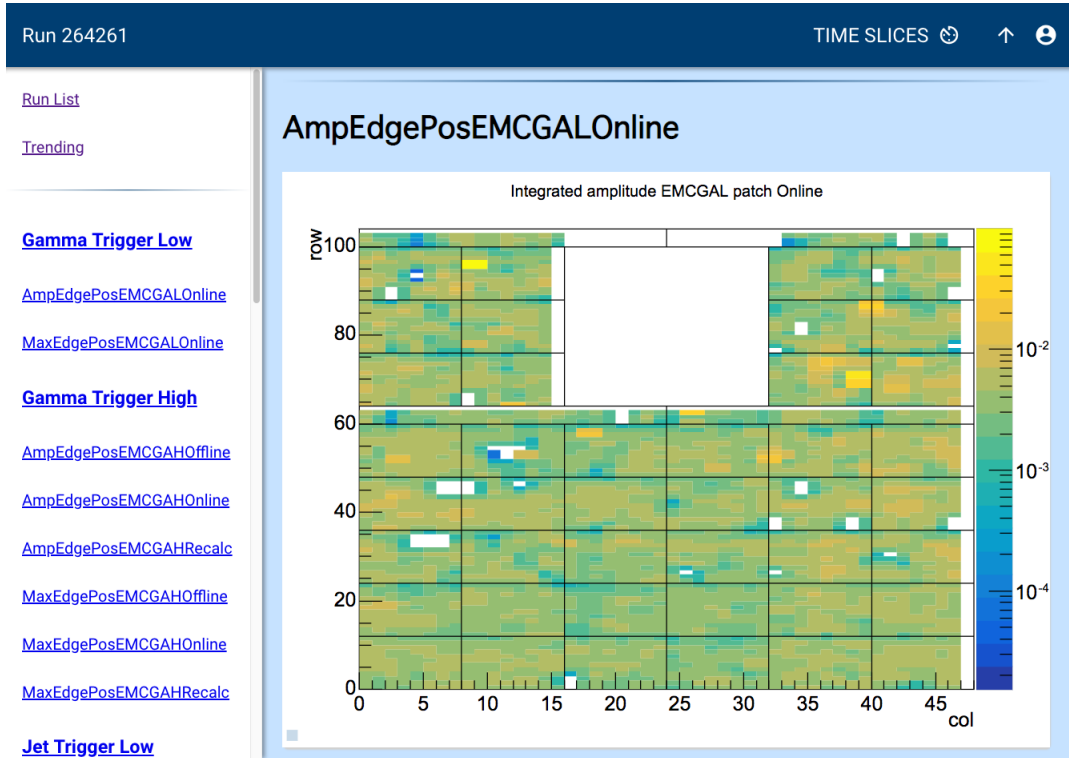


Figure 63: Overwatch WebApp. Figure taken from [raymond'ehlers'2018'1309376].

Overwatch has unique capabilities within ALICE, as it uses time-stamped, persistently stored information, providing real-time and post-mortem data exploration. Furthermore, it has been used as a testing ground for O² Quality Control (QC) developments, specially for the trending and alarms subsystems. A more extensive documentation of Overwatch is found in [raymond'ehlers'2018'1309376].