



Measurement of Λ_c^+ production in Pb–Pb
collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$ with the
ALICE experiment at the LHC

Thesis submitted in accordance with the requirements of the University of
Liverpool for the degree of Doctor in Philosophy by

Clara Bartels

May 2024



Abstract

The quark-gluon plasma (QGP) is an extremely hot and dense form of matter in which quarks and gluons exist in a deconfined state. It was present at the very beginning of the universe just after the Big Bang, before the universe cooled down enough for hadrons to exist. Ultra-relativistic heavy-ion collisions such as the lead-lead collisions at the Large Hadron Collider (LHC) at CERN (Geneva, Switzerland) can be used to recreate this form of matter in the laboratory, creating a very short lived small droplet of QGP just after the collision. These collisions are therefore a way to study not only the properties of the QGP, but also the hadronisation mechanisms at play when the QGP droplet cools down and undergoes a phase transition from free quarks and gluons to hadronic matter.

Heavy quarks, such as the charm or bottom quark, are created at the very beginning of the collision, and interact with the QGP as it expands and cools down, eventually undergoing hadronisation. This means they experience the entire evolution of the system, making heavy-flavour hadrons such as the Λ_c^+ baryon (udc) an excellent probe into the QGP. The relative production of different hadrons, in particular the relative production of baryons and mesons, is an especially useful probe of hadronisation mechanisms. As heavy-flavour hadrons have a very short lifetime, they decay before reaching the detector, and their production has to be reconstructed via their decay products.

In this thesis, the production of Λ_c^+ measured as a function of transverse momentum p_T in Pb–Pb collisions at a centre-of-mass energy per nucleon-nucleon collision of $\sqrt{s_{\text{NN}}} = 5.02$ TeV via the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel is presented. The Λ_c^+ production yield is measured in two collisional centrality classes, central and semi-central,

in two and three p_T bins respectively. A previous measurement of charmed D meson production by the ALICE collaboration is used to calculate the Λ_c^+/D^0 production ratio, and a measurement of Λ_c^+ production in proton-proton (pp) collisions is used to determine the nuclear modification factor R_{AA} . The experimental results of this thesis work are merged with other ALICE measurements performed using a previous analysis done via the $\Lambda_c^+ \rightarrow \text{pK}_S^0$ decay channel, and compared to theoretical models which implement charm quark transport and hadronisation in different ways.

It is found that the Λ_c^+/D^0 ratio is enhanced in central Pb–Pb collisions with respect to pp collisions at intermediate p_T , while it approaches that of pp collisions in the high- p_T region. This indicates that hadronisation differs in Pb–Pb collisions compared to pp collisions. This measurement significantly improves precision at high p_T when added to previous results from the $\Lambda_c^+ \rightarrow \text{pK}_S^0$ channel, and this work gives the most precise measurement of Λ_c^+ production in Pb–Pb collisions to date.

Acknowledgements

Finishing my PhD would not have been possible without the invaluable support of many people. Firstly I would like to thank my supervisors. Marielle Chartier for her support and understanding at all times. Jaime Norman for his constant support and patient help at every step of my analysis. I would not have been able to finish my PhD without him. And Roy Lemmon and Jon Dainton for their valuable input when writing my thesis.

I would also like to thank the wonderful colleagues at Liverpool University, many of whom have become friends over the years, as well as the fantastic people I met at CERN. My thanks also goes to my friends back in Germany, who have helped me through the past four years despite the distance.

Thanks also goes to my family, my parents and my sister, who have supported me my whole life and made this possible in the first place. Special thanks also go to my aunt, whose academic career has proven a big inspiration for me.

Lastly I would like to thank Duncan, who has been there every step of the way and supported me through every struggle along the way.

Contents

Abstract	i
Acknowledgements	iii
Contents	vii
List of Figures	xv
List of Tables	xvi
1 Introduction	1
1.1 Quantum Chromodynamics	2
1.2 Quark-gluon plasma	6
1.3 Heavy-ion Collisions	10
1.3.1 Stages of a heavy-ion collision evolution	10
1.3.2 Hadronisation mechanisms	11
1.3.3 Properties of heavy-ion collisions	13
1.4 Measurements in Heavy-ion collisions	16
2 Heavy-flavour physics	22
2.1 Heavy-flavour production in pp collisions	22
2.2 Heavy-flavour production in Pb–Pb collisions	23
2.2.1 Interactions in the QGP	24
2.3 Previous measurements	27
2.3.1 Small systems	28
2.3.2 Pb–Pb collisions	31
2.4 Aim of this thesis	33

3	Experimental setup	35
3.1	The LHC	35
3.2	The ALICE detector	37
3.2.1	The Inner Tracking System	40
3.2.2	The Time Projection Chamber	42
3.2.3	The Time of Flight detector	44
3.2.4	The V0 detector	45
3.3	Trigger system	47
3.4	Track and vertex reconstruction	48
3.5	Particle identification	49
4	Data Analysis	52
4.1	Overview	52
4.2	Event Selection	53
4.2.1	Monte Carlo simulation	54
4.3	Candidate selection	54
4.3.1	Candidate reconstruction and track selection	54
4.3.2	Particle identification	55
4.3.3	Λ_c^+ candidate selection	56
4.3.4	Initial kinematic and topological selection	57
4.4	Multivariate Selection	59
4.4.1	Boosted Decision Trees	60
4.4.2	BDT training	62
4.4.3	Raw yield	70
4.5	Corrections	72
4.5.1	Efficiency	72
4.5.2	Feed-down	75
4.6	Systematic Uncertainties	79
4.6.1	Raw Yield Extraction	80
4.6.2	BDT cut value	82
4.6.3	MC p_T -shape	82
4.6.4	Tracking Efficiency	85
4.6.5	Pre-selection PID efficiency	87
4.6.6	Feed-down subtraction	88
4.6.7	Summary	90

5	Results	93
5.1	Λ_c^+ production yield	94
5.2	Λ_c^+/D^0 ratio	97
5.3	Theoretical Models	98
5.4	Comparison to models	102
6	Conclusion and Outlook	105
A	Machine Learning performance plots	108
	References	123

List of Figures

1.1	The particles of the standard model [1].	3
1.2	Feynman diagrams for the interaction vertices of quarks and gluons. The two rightmost are possible due to the self-interaction of gluons, and do not have an equivalent in QED.	4
1.3	A summary of measurements of the strong coupling constant α_s as a function of momentum transfer Q [2].	5
1.4	A schematic of the phase diagram of QCD as a function of temperature and baryon chemical potential (see text) [3].	7
1.5	The pressure p , energy density ϵ , and entropy density s normalised to the 4 th (3 rd) power of temperature T , obtained through Lattice QCD calculations [4]. The darker lines are the predictions from the Hadron Resonance Gas (HRG) model [5]. There is a strong increase in the cross-over transition region (yellow band), but even at $T \approx 400$ MeV the energy density is still roughly 20% below the ideal gas limit [3].	9
1.6	The stages of evolution of a heavy-ion collision at ALICE [3] (see text). b denotes the impact parameter (see section 1.3.3).	12
1.7	A schematic view of a heavy-ion collision showing the impact parameter b , which determines the centrality of the collision, as well as the participating and spectating nucleons [6].	15
1.8	The measured normalised charged multiplicity as a function of the number of participating nucleons for different collision systems, scaled for different collision energies and compared to PYTHIA predictions (see text) [3].	16
1.9	The production yield per unit of rapidity in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV for different hadron species measured by ALICE at mid-rapidity and over the full p_T range, compared to four different statistical hadronisation models. Figure taken from [3].	17

1.10	The nuclear modification factors for charged particles measured by ALICE for p–Pb collisions and central (0-5% centrality) and peripheral (70-80% centrality) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV as a function of p_{T} [7].	19
1.11	The Λ/K_s^0 ratio measured by ALICE in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV as a function of p_{T} for different centrality classes, and compared to measurements done in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV [8].	20
2.1	The p_{T} -differentiated production cross section of Λ_c^+ baryons (left) and D^0 mesons (right) measured by ALICE in pp collisions at $\sqrt{s} = 5.02$ TeV [9, 10]. They are compared to different pQCD calculations [11, 12] (see text).	28
2.2	Left: the Λ_c^+/D^0 baryon-to-meson ratio measured by ALICE in pp collisions at $\sqrt{s} = 5.02$ TeV, compared to several models. Right: The D^+/D^0 and D_s^+/D^0 ratio measured in pp collisions at $\sqrt{s} = 7$ TeV, also compared to several models [13].	29
2.3	The nuclear modification factor R_{pPb} of Λ_c^+ baryons measured by ALICE in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV compared to that measured for D mesons (left) and to two models (right) [9]	30
2.4	The nuclear modification factor R_{AA} for prompt and non-prompt D meson production measured by ALICE in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV [14].	32
2.5	The p_{T} -integrated production yields per units of rapidity for different charm-hadron species measured by ALICE in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ GeV, compared to predictions from a statistical hadronisation model [3, 15].	33
3.1	The CERN accelerator complex [16].	36
3.2	The ALICE detector setup for Run 2 [17]. The beam-pipe is surrounded by the central barrel, which includes the Inner Tracking System (ITS) that is shown in an insert. On the right side of the image, the forward parts of the detector system are located, such as the dipole magnet and muon spectrometer.	38
3.3	The impact parameter resolution of ALICE for Pb–Pb collisions in Run 2, compared to that projected for Run 3. [18] The open and closed markers show the longitudinal and transverse directions, respectively.	40

3.4	Schematic of the Inner Tracking System used in ALICE in Run 1 and Run 2 [19]. Shown are the six silicon detector layers, consisting of two layers each of SPD, SDD, and SSD.	41
3.5	Specific ionising energy loss dE/dx of tracks measured in the TPC as a function of momentum in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV [20]. The black lines represent the Bethe-Bloch parametrisations for various charged particles. Figure taken from the ALICE figure repository. . .	43
3.6	Relativistic particle velocity β vs momentum measured by the TOF detector in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV (Run 2) [21]. The signatures of the individual particles are labelled and clearly distinguishable. Figure taken from the ALICE figure repository.	46
3.7	The distribution of the sum of amplitudes in the V0 detectors (VOA and VOC) in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ GeV (black), with a fit to the Glauber model (red). The different centrality classes are shown in the shaded areas. [22]	47
4.1	A schematic of the reconstructed $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay, with some reconstruction variables labelled. PV and SV denote the primary and secondary decay vertices, respectively, DCA is the distance of closest approach, and θ is the pointing angle (defined in the text).	58
4.2	The signal and background distribution of four variables (cosine of the pointing angle $\cos(\theta_p)$, σ_{vertex} , distance of closest approach (dca), and dist_{12}), with the chosen pre-selection cut value marked in black. . . .	60
4.3	An example for a decision tree. At every node, the variable (X,Y,Z,W) and cut value (0, 5, 10) with the best signal to background separation are chosen to divide the dataset (A,B,C, D, ...).	63
4.4	The signal and background distribution of the variables used for the training, shown in the $12 < p_T < 24$ GeV/ c bin for central collisions. The individual variables, from left to right and top to bottom, are the decay length, the decay length in the x-y plane, the normalised decay length in the x-y plane, the cosine of the pointing angle, the cosine of the pointing angle in the x-y plane, the impact parameter in the x-y plane, the distance of closest approach, the σ of the vertex reconstruction, the distance between the 1 st and 2 nd tracks, the distance between the 2 nd and 3 rd tracks, the maximum difference between measured and calculated impact parameter, and the impact parameter for the three tracks (prong 0,1, and 2).	64

4.5	The learning curve of the BDT model, shown in the $12 < p_{\text{T}} < 24 \text{ GeV}/c$ bin for both centrality classes, which shows a good convergence between training and testing datasets.	65
4.6	The ROC curve of the BDT model, shown in the $12 < p_{\text{T}} < 24 \text{ GeV}/c$ bin for both centrality classes. A comparison is made between the testing and training data sets, and showing a good agreement and a reasonably high AUC value.	66
4.7	The individual impact of each training variable on the model, shown in the $12 < p_{\text{T}} < 24 \text{ GeV}/c$ bin for both centrality classes. The x-axis denotes the impact on the model, where a higher value means a more signal-like impact, and a lower one indicates background-like. The colour shows the variable value of the individual instance, with red showing a high value and blue a low one. They are ordered from most impactful (top) to least impactful (bottom).	67
4.8	The BDT model output, shown in the $12 < p_{\text{T}} < 24 \text{ GeV}/c$ bin for both centrality classes. It shows the separation between signal-like and background-like events. There is good agreement between training and testing datasets.	68
4.9	The expected significance of the Λ_c^+ invariant mass peak versus the BDT cut in the $12 < p_{\text{T}} < 24 \text{ GeV}/c$ bin for both centrality classes of this analysis. Note that the sharp rise at very high BDT cut value does not represent a real increase in significance but is instead a consequence of the extremely low statistics in this region.	70
4.10	Invariant mass distributions of Λ_c^+ candidates in the 0-10% centrality interval. Shown are the $6 < p_{\text{T}} < 8 \text{ GeV}/c$, $8 < p_{\text{T}} < 12 \text{ GeV}/c$, and $12 < p_{\text{T}} < 24 \text{ GeV}/c$ intervals. Fits to the Λ_c^+ signal (blue) and background (red) are also shown.	73
4.11	Invariant mass distributions of Λ_c^+ candidates in the 30-50% centrality interval. Shown are the $4 < p_{\text{T}} < 6 \text{ GeV}/c$, $6 < p_{\text{T}} < 8 \text{ GeV}/c$, $8 < p_{\text{T}} < 12 \text{ GeV}/c$, and $12 < p_{\text{T}} < 24 \text{ GeV}/c$ intervals. Fits to the Λ_c^+ signal (blue) and background (red) are also shown.	74
4.12	The mean (left) and width (right) of the Gaussian fits to the Λ_c^+ signal peak in Monte Carlo simulation compared to the fit in data when leaving the width as a free parameter.	75
4.13	The invariant mass distribution for Λ_c^+ candidates in central Pb–Pb collisions in the $24 < p_{\text{T}} < 36 \text{ GeV}/c$ bin. Due to a lack of statistics the extension of the analysis to this momentum bin was not successful, as the significance of 2 is well below the minimum required 3.	76

4.14	The invariant mass fit results for this analysis as a function of p_T compared to results from a previous ALICE analysis for the $\Lambda_c^+ \rightarrow pK_S^0$ channel, with both collision centrality classes shown. On the left is the significance of the fit, with the imposed minimum quality of Significance > 3 shown as a dotted line. On the right are the signal to background ratios for each bin. It can be seen that while this analysis is less performant in the low momentum bins, it offers an improvement at higher momentum.	77
4.15	The detector acceptance corrected efficiency for prompt and feed-down Λ_c^+ baryons for central (left) and semi-central (right) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.	78
4.16	Fraction of prompt Λ_c^+ baryons calculated as a function of p_T for central (left) and semi-central (right) Pb–Pb collisions as a function of p_T . The central value is shown as the red line, and the systematic uncertainties are represented by the black dotted lines.	79
4.17	The results of the multi-trial variation of the fitting parameters of the raw yield extraction. They show, starting from the top left: The distribution of the fit width σ , in cases where it is fixed or left free, with the value from the main analysis shown in red; The distribution of the mean of the fits; the raw yields from the trials without (top) and with (bottom) bin counting; the distribution of the raw yield from all the trials; and the χ^2/NDF (top) and significance (bottom) of the trials. These plots are for Pb–Pb collisions in the 30-50% centrality bin and $12 < p_T < 24$ GeV/ c bin, and yield an RMS of 12%.	83
4.18	The invariant mass distributions for central Pb–Pb collisions in the $12 < p_T < 24$ GeV/ c momentum bin resulting from varying the BDT cut around the central value of 0.7.	84
4.19	The corrected Λ_c^+ yields resulting from the different BDT cut values applied, normalised with the yield found with the central cut value. The shaded area represents the systematic uncertainty derived from this.	85
4.20	The corrected Λ_c^+ yields resulting from the different p_T MC p_T -distribution shapes applied, compared to that from the TAMUmax model used in the analysis. The shaded area represents the systematic uncertainty derived from this.	86
4.21	The systematic uncertainty on the matching efficiency between TPC and ITS, shown for pions and kaons combined, as well as for protons, for central and semi-central collisions.	88

4.22	The relative single-track quality selection efficiency shown for pions for the additional TPC selections in experimental data over that in MC data, for central collisions.	89
4.23	The combined Λ_c^+ systematic tracking uncertainty for central (left) and semi-central (right) collisions. The individual contributions from matching efficiency (red) and the TPC tracks (green) are also shown.	90
4.24	The corrected yields resulting from the different PID preselections, normalised by the yield from the main analysis. The shaded area represents the systematic uncertainty derived from this.	91
4.25	The uncertainty on the prompt R_{AA} fraction resulting from the variations of the $R_{\text{AAFeed-Down}}/R_{\text{AAprompt}}$ hypothesis for both central (left) and semi-central (right) collisions, as described in the text.	92
5.1	The p_{T} -differential production yield of Λ_c^+ baryons measured in the $\Lambda_c^+ \rightarrow \text{pK}^-\pi^+$ channel in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV for central (left) and semi-central (right) events. The bars and boxes represent the statistical and systematic uncertainties, respectively.	94
5.2	The p_{T} -differential production yield of Λ_c^+ baryons measured in the $\Lambda_c^+ \rightarrow \text{pK}^-\pi^+$ channel in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV compared to that measured in the $\Lambda_c^+ \rightarrow \text{pK}_\text{S}^0$ channel [23]. The two measurements are in good agreement. The bars and boxes represent the statistical and systematic uncertainties, respectively.	95
5.3	The p_{T} -differential production yield of Λ_c^+ baryons measured in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV from the combined results of the $\Lambda_c^+ \rightarrow \text{pK}^-\pi^+$ and $\Lambda_c^+ \rightarrow \text{pK}_\text{S}^0$ channels. The bars and boxes represent the statistical and systematic uncertainties, respectively.	96
5.4	The relative systematic (top) and statistical (bottom) uncertainties for central (left) and semi-central (right) events for the two different decay channels, $\Lambda_c^+ \rightarrow \text{pK}^-\pi^+$ (red) and $\Lambda_c^+ \rightarrow \text{pK}_\text{S}^0$ (blue), with the total uncertainties found by merging the two results shown in black.	98
5.5	The Λ_c^+/D^0 ratio measured in the $\Lambda_c^+ \rightarrow \text{pK}^-\pi^+$ channel in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV for central (left) and semi-central (right) events. The measurement of D^0 production are taken from [24]. The bars and boxes represent the statistical and systematic uncertainties, respectively.	99

5.6	The Λ_c^+/D^0 ratio measured in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel (red) compared to that from the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel (blue), both measured at $\sqrt{s_{NN}} = 5.02$ TeV. The black data points are the average of the two decay channels, and the bars and boxes represent the statistical and systematic uncertainties, respectively.	100
5.7	The merged Λ_c^+/D^0 ratio measured in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for two centralities compared to that measured in pp collisions also at $\sqrt{s_{NN}} = 5.02$ TeV [25]. The bars and boxes represent the statistical and systematic uncertainties, respectively.	101
5.8	The merged Λ_c^+/D^0 ratio in Pb–Pb (central and semi-central) and pp collisions measured at $\sqrt{s_{NN}} = 5.02$ TeV compared to three theoretical models: SHMc [15], Catania [26], and TAMU [27] (see text).	103
5.9	Top: The merged nuclear modification factor R_{AA} measured at $\sqrt{s_{NN}} = 5.02$ TeV for central and semi-central events compared to three theoretical models: SHMc [15], Catania [26], and TAMU [27] (see text). The bottom figure shows the same figure from the previously published analysis [23], and the increase in precision at high p_T provided by this work can be clearly observed.	104
6.1	The integrated luminosity of the Pb–Pb data taking run at ALICE in October 2023 [28].	106
A.1	The learning curve showing the RMSE of the training and testing data sets for central events.	109
A.2	The learning curve showing the RMSE of the training and testing data sets for semi-central events.	110
A.3	The ROC curve for the training and testing data set for central events.	111
A.4	The ROC curve for the training and testing data set for semi-central events.	112
A.5	The impact of the different training variables shown for central events.	113
A.6	The impact of the different training variables shown for semi-central events.	114
A.7	The model output distribution for central events, showing the separation between signal and background events for training and testing data sets.	115
A.8	The model output distribution for semi-central events, showing the separation between signal and background events for training and testing data sets.	116
A.9	The significance optimisation plots for central collisions.	117

A.10	The significance optimisation plots for semi-central collisions.	118
A.11	The invariant mass plots for central collisions with a fit to the Λ_c^+ mass in the three p_T bins.	119
A.12	The invariant mass plots for semi-central collisions with a fit to the Λ_c^+ mass in the three p_T bins.	120
A.13	The invariant mass MC plots for central collisions with a fit to the Λ_c^+ mass in the three p_T bins.	121
A.14	The invariant mass MC plots for semi-central collisions with a fit to the Λ_c^+ mass in the four p_T bins.	122

List of Tables

3.1	Properties of the six ITS layers. Numbers taken from [29].	42
4.1	The different decay modes of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel and their respective branching ratios and energies [2].	53
4.2	The pre-selection cuts applied on reconstructed tracks (bottom rows) and Λ_c^+ candidates (top rows).	59
4.3	The hyperparameters of the XGBoost model used in this analysis. . .	62
4.4	Number of signal and background training events used in central collisions.	63
4.5	The number of signal and background training events used in semi-central collisions.	65
4.6	BDT cut chosen in this analysis for each centrality and momentum bin	70
4.7	Summary of the systematic uncertainties found from the various sources discussed in this section.	92

Chapter 1

Introduction

The ALICE experiment (A Large Ion Collider Experiment) at the LHC (Large Hadron Collider) at CERN in Geneva (Switzerland) was built to study a deconfined state of matter called quark-gluon plasma (QGP), that is speculated to have existed at the very beginning of the universe. It can be recreated in high-energy heavy-ion collisions, such as lead-lead (Pb–Pb) collisions at LHC. Decades of experimental study since the 1980s has lead to remarkable progress in determining QGP properties. Despite this, there are still many aspects of the QGP that are unknown or not understood. These include a precise characterisation of its constituents and how they interact, what phenomena emerge from its presence in a hadronic collision, and by what mechanisms hadrons form as the QGP cools down and undergoes a phase transition back to ordinary hadronic matter. Heavy-flavour hadrons provide a good probe into some of these questions, as their high mass means they are created only at the very beginning of the collision and experience all stages of the QGP evolution. In order to gain information about the QGP properties, it is therefore important to study the production of these hadrons. In this thesis, the production yield of Λ_c^+ baryons is measured in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, which provides a useful probe of QGP properties and hadronisation mechanisms.

In this chapter, the QGP and its properties are discussed, and heavy-ion collisions are described. In chapter 2, heavy-flavour production, which this thesis focuses on, is introduced. The role of heavy-flavour constituents in investigating the properties

of the QGP, as well as heavy-flavour production measurements from small systems such as proton-proton (pp) collisions are discussed. Chapter 3 gives an overview of the experimental setup used for this analysis, i.e. of the Large Hadron Collider and the ALICE detector. In chapter 4, the analysis procedure is described in detail. The results are discussed in chapter 5, and they are combined with the results of a previous measurement by the ALICE collaboration done in a different decay channel. These results are compared to theoretical models which implement different hadronisation mechanisms. Finally, chapter 6 discusses these results and offers an outlook on some future measurements.

In the following, error bars generally denote statistical uncertainties, while error boxes show systematic uncertainties.

1.1 Quantum Chromodynamics

The standard model of particle physics was developed in the second half of the 20th century and it is currently our best understanding of all elemental particles and their fundamental interactions. It describes three types of fundamental particles and their interactions: Quarks and leptons are fermions with a spin of $\frac{1}{2}$, while gauge bosons have a spin of 1.

There are three families of quarks, which have one positively charged ($+\frac{2}{3}$) and one negatively charged ($-\frac{1}{3}$) quark each. These are up (u), down (d), charm (c), strange (s), top (t) and bottom (b) quarks. The lighter flavour quarks (u,d) are the most stable, while the heavier ones quickly decay into the lighter quarks. In addition to their electromagnetic charge, quarks also carry a so-called “colour charge”, which allows them to interact via the strong force that is mediated by gluons. This will be elaborated on in section 1.1. Leptons also have three families. They consist of the negatively charged electron (e), muon (μ), and tau (τ), as well as the corresponding neutrinos (ν_e, ν_μ, ν_τ). While e , μ , and τ have a negative charge of 1, the neutrinos are electrically neutral and interact only via the weak force.

In addition to these 12 fermions, there are also gauge bosons in the standard model, which act as mediators for the three fundamental forces: gluons mediate the

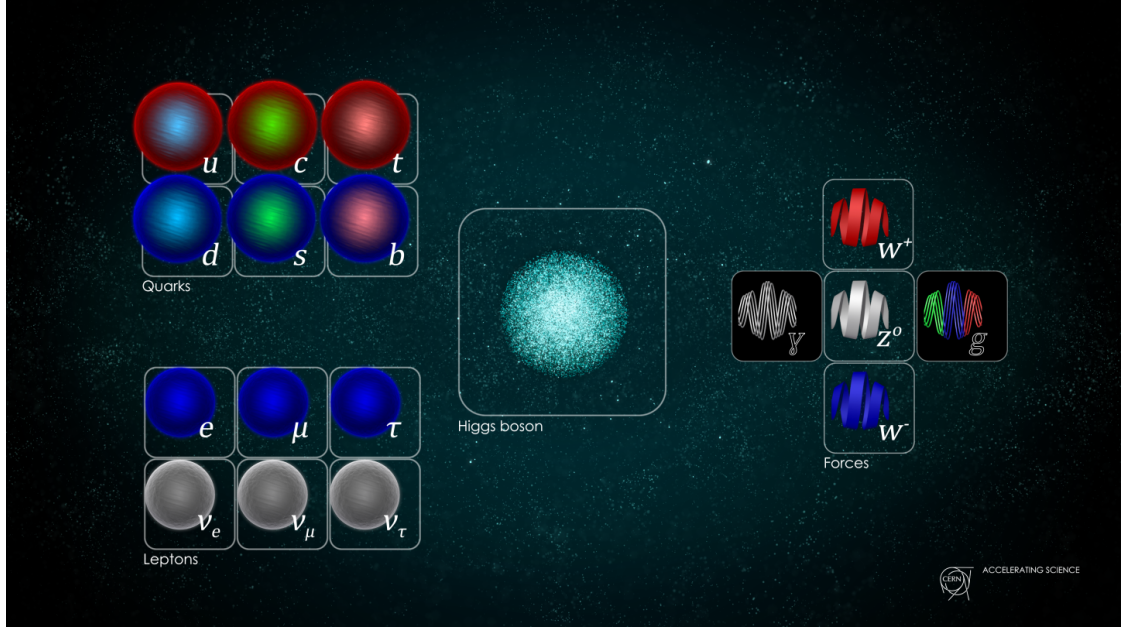


Figure 1.1: The particles of the standard model [1].

strong force, photons the electromagnetic force, and Z^0 and $W^\pm$ bosons mediate the weak force. The last particle in the standard model to be discovered was the Higgs boson [30, 31], which has a spin of 0. It is an excitation in the Higgs field, which gives mass to the elemental particles. Figure 1.1 shows the 18 particles that make up the standard model. Note that while the gluon is shown as one particle, there are actually eight different types of gluons which correspond to different colour states.

Quantum Chromodynamics (QCD) is the theory of interaction between quarks, which is mediated by gluons. It is a non-abelian gauge field theory that describes the interactions of the strong force [32]. A crucial difference to electromagnetic interactions is that gluons, the force-carriers of the strong interaction, are self-interacting. This means that unlike photons, which do not carry electromagnetic charge, gluons do carry a colour charge, and can therefore interact with other gluons. The Lagrangian of the QCD is given by

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{\mu\nu}^\alpha G_{\alpha}^{\mu\nu} + \sum_f \bar{\Psi}_f (i\gamma^\mu D_\mu - m_f) \Psi_f \quad (1.1)$$

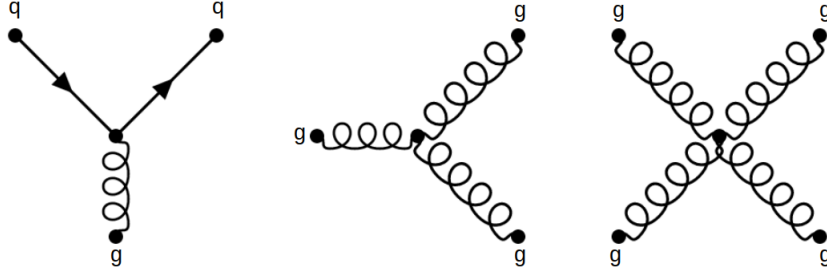


Figure 1.2: Feynman diagrams for the interaction vertices of quarks and gluons. The two rightmost are possible due to the self-interaction of gluons, and do not have an equivalent in QED.

where f runs over the six quark flavours, Ψ_f is a quark field of flavour f , and m_f is the mass of a quark flavour f . $G_{\mu\nu}^\alpha$ is the gluonic field tensor, which unlike its QED equivalent has a third term accounting for the self-interaction of gluons:

$$G_{\mu\nu}^\alpha = \delta_\mu A_\nu^\alpha - \delta_\nu A_\mu^\alpha - gf^{\alpha\beta\gamma} A_\mu^\beta A_\nu^\gamma \quad (1.2)$$

The index α runs over the 8 colour degrees of freedom of the gluon field. Figure 1.2 shows the Feynman diagram of possible interaction vertices in QCD, which include the self-interacting terms that are not present in QED. As a consequence of this self-interaction, the screening effect, which also occurs in QED, is counteracted by an anti-screening effect stemming from gluon loop interactions. The interplay between these two effects determines the strength of the coupling constant α_s , which determines the strength of the interaction between quarks and gluons. Unlike its QED equivalent, it does not increase but decreases with the momentum transfer Q of the interaction as:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f) \cdot \ln(Q^2/\Lambda_{\text{QCD}}^2)} \quad (1.3)$$

where n_f is the number of quark flavours, and Λ_{QCD} is the scale parameter of QCD, which depends on the number of quark flavours used in its calculation. The dependence of α_s on Q is shown in figure 1.3.

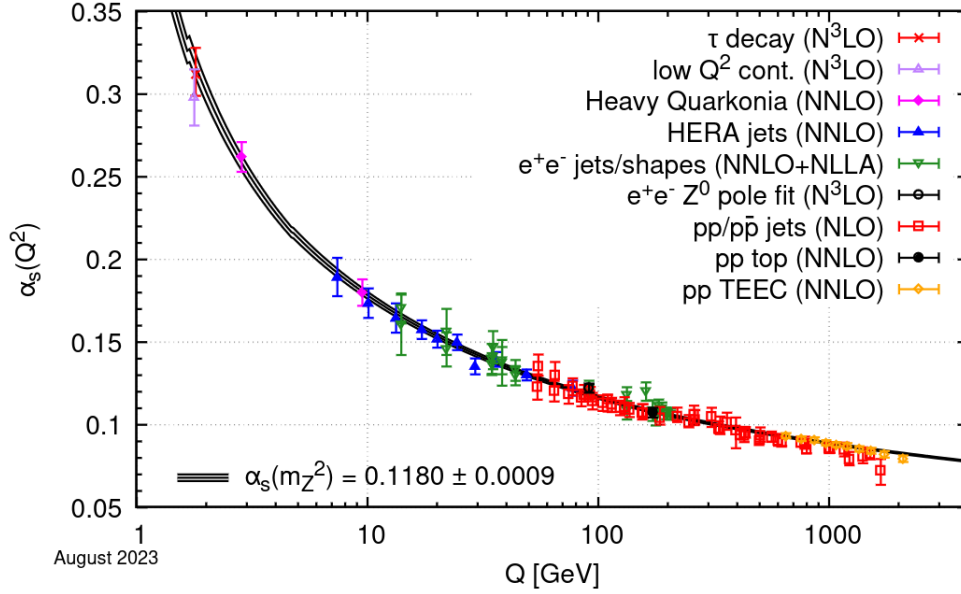


Figure 1.3: A summary of measurements of the strong coupling constant α_s as a function of momentum transfer Q [2].

For momentum transfers larger than the QCD scale ($\Lambda_{\text{QCD}} \gtrsim 200 \text{ MeV}$), α_s becomes small and the quarks are weakly bound, a behaviour which is called asymptotic freedom. Due to the small value of α_s , this region can be investigated with a perturbative calculation (pQCD). In it, the process is expanded to higher orders of α_s , but only lower orders contribute significantly, so the higher orders can be neglected [33]. At lower momentum transfers, α_s becomes larger, and the perturbative approach breaks down as higher orders dominate. In this region, a phenomenon called confinement is observed, which means quarks are strongly bound in hadrons and cannot be separated [34]. Another feature observed in this region is spontaneous breaking of chiral symmetry [35]. In this non-perturbative regime, theoretical QCD calculations can be done using Lattice QCD [36, 37] or Effective Field Theories [38].

Many aspects of QCD, such as the measurement of α_s shown in figure 1.3, can be studied using pp collisions. Another way to gain more information about QCD is to study extended systems that are subject to the strong interaction. In the study

of electromagnetic interactions, this question is addressed in the field of condensed matter physics. In that case, phenomena such as superconductivity emerge, that do not rise directly from the rules of the microscopic interactions [39]. A parallel to condensed matter physics can be found in QCD studies. Heating up an extended system of quarks and gluons to a very large energy density produces a quark-gluon plasma (QGP). For such a system, confinement no longer applies, and chiral symmetry is approximately restored [40]. Knowledge of the dynamics of its fundamental constituents (quarks and gluons) sheds little light on an understanding of the bulk properties of QCD matter, and completely new phenomena may emerge. The study of the QGP therefore provides a unique way of further investigating QCD behaviour.

1.2 Quark-gluon plasma

The QGP is a special state of nuclear matter that occurs at very high energy density, and in which quarks and gluons are no longer confined inside hadrons but exist in a free state. It is thought to have existed at the very beginning of the universe, up to about 10^{-6} s after the Big Bang. Figure 1.4 shows the phase diagram of QCD, with the temperature T plotted against the baryochemical potential μ_B , which is proportional to the net baryon number. The QGP present at the beginning of the universe corresponds to a low net baryon number and a high temperature. A complimentary situation has also been theorised, with a QGP occurring at relatively low temperatures and high baryon density. This is speculated to be the case in the core of neutron stars [41]. In the laboratory, a situation with low net baryon number and high temperature can be created in high-energy heavy-ion collisions, where a huge amount of energy is deposited in a small volume, leading to a QGP being created.

The first attempt to describe nuclear matter in extreme conditions was made as early as the 1930s, with Oppenheimer and Volkoff calculating its properties at far higher than normal nuclear density [42]. Similar calculations were carried out by Zeldovich in the 1960s [43] and Baym and Chin in the 1970s [44]. The theoretical models of the time were not sufficient to describe nuclear matter at these extremes,

and conflicting theories existed. The fundamental constituents of nuclear matter, quarks and gluons, were not known until the proposal of quarks by Gell-Mann and Zweig in 1964 [45, 46], and their subsequent discovery in inelastic scattering experiments in 1969 [47, 48]. A breakthrough came with the discovery of asymptotic freedom in 1973 by Gross, Wilzek, and Politzer [49, 33]. It led to the postulation of a dense quark soup at very high temperatures by Collins and Perry [50], as well as the existence of a second-order phase transition between the hadronic and quark–gluon phases of matter, as described by Cabibbo and Parisi [40]. A host of possible signatures of the QGP were introduced, such as an excess of multi-strange baryons compared to proton-proton collisions as a consequence of the increased production of strange quarks [51], and a suppression of J/Ψ due to its binding energy being lower than the QGP temperature [52]. Other observables included an enhancement of the di-lepton spectrum [53] and the emission of high-energy photons.[54, 55].

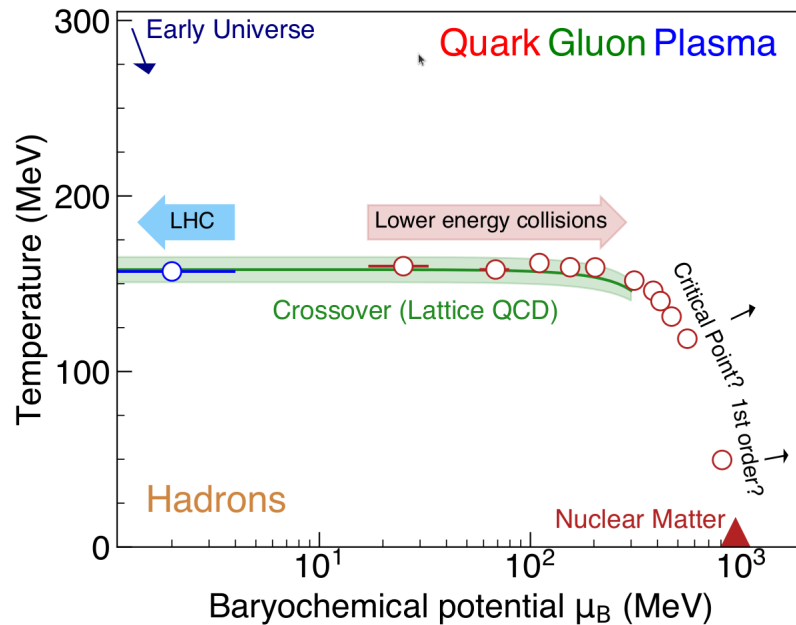


Figure 1.4: A schematic of the phase diagram of QCD as a function of temperature and baryon chemical potential (see text) [3].

With the start of AGS at BNL [56] and CERN’s SPS [57], both in 1986, nu-

clear collisions were available at an energy of $\sqrt{s_{\text{NN}}} = 5 \text{ GeV}$ and $\sqrt{s_{\text{NN}}} = 20 \text{ GeV}$, respectively, and during the 1990s, many of these QGP signatures were observed [58, 59, 60, 61, 57]. The co-existence of so many of these signatures was seen as evidence for the existence of the new state of matter [57]. The start of the first heavy-ion collider RHIC in 2000 meant an increase of the collision energy to 130 GeV [62]. In 2002, RHIC observed for the first time the suppression of high- p_{T} particles in Au–Au collisions, which was seen as evidence for the energy loss of hard partons in the QGP, and initiated further studies of jet modification in the medium [63]. Another important discovery was that the QGP exhibits behaviour as a nearly ideal fluid [64], as found from measurements of the nuclear modification factor R_{AA} (see section 2.3) at high transverse momentum p_{T} , as well as in measurements of anisotropic flow at low p_{T} . When the LHC started operation in 2008, its heavy-ion programme offered a further significant increase in the available collision energy, with Pb–Pb collisions initially running at 2.76 TeV and increasing to 5.02 TeV in 2015. The ALICE experiment was built with the main purpose of measuring those collisions (see chapter 3). The increase in not only energy but also hadronic multiplicity offered ideal conditions to study the chemical composition of the medium, as well as its behaviour, with unprecedented precision.

A useful tool for theoretical descriptions of the QGP is lattice QCD, which can be used to describe static problems in the non-perturbative region of QCD. In this approach, quarks are designated as the lattice sites, with gluons as the connections between those sites. With sufficient computational power, the QCD equation of state can be calculated using realistic masses for the dominant quark flavours (up, down, and strange) and an extrapolation to the continuum limit [65]. This can then be used to extract the thermodynamic properties of the QGP, such as pressure or energy density, as well as their temperature dependence. Calculations of lattice QCD show that the system undergoes a smooth transition from confined to deconfined state at a critical temperature $T_c \approx 155 \text{ MeV}$ [4]. It is important to note that this calculation was done for a system with zero net baryon density, which corresponds to the situation in the early universe before hadrons existed. Due to the continuous nature of the thermodynamic variables, a cross-over state occurs, where confined and

de-confined matter are present at the same time. This corresponds to the so-called pseudo-critical temperature T_{pc} . The transition to deconfined matter coincides with a strong increase in energy density ϵ , indicating the release of many new degrees of freedom in the system, correlating to quarks and gluons instead of confined hadrons. This can be seen in figure 1.5, which shows the pressure, energy density and entropy density, normalised to the 4th power of the temperature. The results are extrapolated to higher temperatures, and as shown in the figure, even at $T \approx 400$ MeV, the energy density of a non-interacting ideal-gas limit is not reached. Below this limit, the QGP exists as a strongly interacting system of quarks and gluons, behaving as an almost perfect liquid [66].

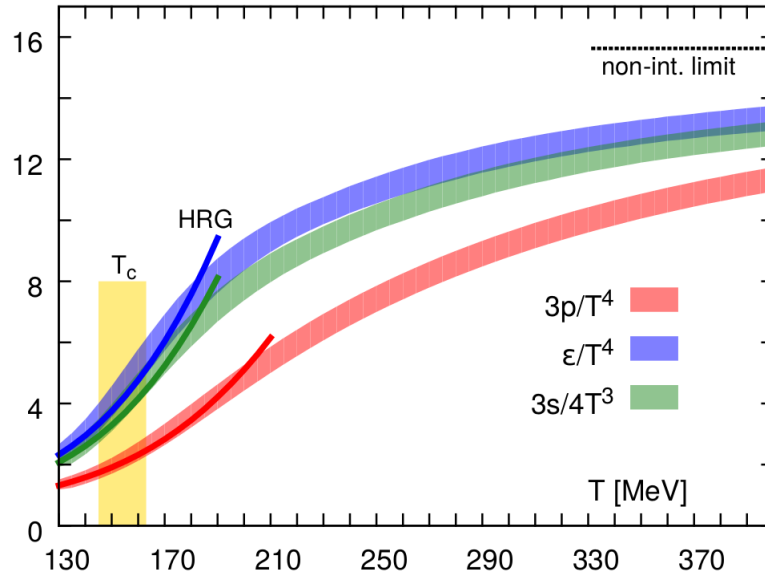


Figure 1.5: The pressure p , energy density ϵ , and entropy density s normalised to the 4th (3rd) power of temperature T , obtained through Lattice QCD calculations [4]. The darker lines are the predictions from the Hadron Resonance Gas (HRG) model [5]. There is a strong increase in the cross-over transition region (yellow band), but even at $T \approx 400$ MeV the energy density is still roughly 20% below the ideal gas limit [3].

1.3 Heavy-ion Collisions

Ultra-relativistic heavy-ion collisions are a powerful tool for studying the quark-gluon plasma, as they offer the chance to create and observe the system in the laboratory. The main challenge is the short lifetime of the QGP (in the order of 10^{-15}s), which rapidly cools and hadronises. It therefore needs to be studied indirectly by measuring the particles resulting from the collision.

1.3.1 Stages of a heavy-ion collision evolution

Figure 1.6 shows the evolution of the QGP created in Pb–Pb collisions at ALICE. It consists of several stages, from the initial collision to the particles reaching the detector:

- **Initial state** In the initial state, the energy and entropy distribution is governed by the partons participating in the collision, leading to a “lumpy” distribution. In this weakly-coupled pre-equilibrium phase, soft partons are created in small- Q^2 interactions. Large- Q^2 interactions produce high-momentum quarks and gluons, which lead to the production of jets in the QGP.
- **Collision** The collision happens over a very short time frame of less than $1\text{ fm}/c$. The crossing time, which is the time it takes the nucleons to pass each other completely, is $\tau_{\text{cross}} = 2R/\gamma$ for Pb–Pb collisions at the LHC, where R is the nuclei radius and the Lorentz factor γ is estimated at 1500. This means that the collision happens much faster than the time scale of the strong interaction. A crucial property of the collision is how centrally the two nuclei collide, and therefore how many nucleons are participating in the collision. This is quantified by something called the centrality of the collision, which is described in section 1.3.3.
- **QGP formation** After about $\tau \approx 1\text{ fm}/c$, the system reaches thermal equilibrium, and a QGP is created. The quarks and gluons in it are strongly coupled and form a collective medium. Due to the pressure gradients stemming from

the original parton distribution, the system rapidly expands. At this point, the system behaves as a nearly-perfect liquid, and the expansion occurs collectively. This means that it can be described using relativistic hydrodynamics.

- **Hadronisation** As the system expands, it also cools. Once the medium temperature falls below the critical temperature T_c , the QGP ceases to exist and reverts back into hadrons. At this point a hadron gas is formed, which is a system that has different degrees of freedom but still behaves collectively. This hadronisation can happen via several different processes, and is described in more detail in 1.3.2. The hadrons created still interact via both elastic and inelastic processes, so that the hadronic species abundances are not yet fixed.
- **Freeze-out** The chemical freeze-out occurs when the temperature of the system has gone down sufficiently for the inelastic collisions of the hadrons to cease, at which point the particle abundances become fixed. Elastic interactions continue, changing the momentum distribution of the hadrons until the so-called kinematic freeze-out. This occurs at roughly $\tau \approx 10 \text{ fm}/c$, when the hadron momentum distributions become fixed.

After freeze-out, the particles then fly through the vacuum before reaching the innermost layer of the detector. The measurement of these particles offers valuable insights into the development and the properties of the QGP.

1.3.2 Hadronisation mechanisms

Of particular interest in the evolution of the QGP is the hadronisation process. As mentioned above, hadronisation describes the formation of hadrons from the deconfined quark-gluon plasma, and is a fundamental process of QCD. Due to the small momentum transfers and consequently large coupling constant involved, it cannot be described using perturbative QCD, which makes it challenging to study. Instead, various phenomenological models of different hadronisation mechanisms are used. In heavy-ion collisions, there are expected to be two mechanisms by which hadrons are formed from the QGP: fragmentation and coalescence.

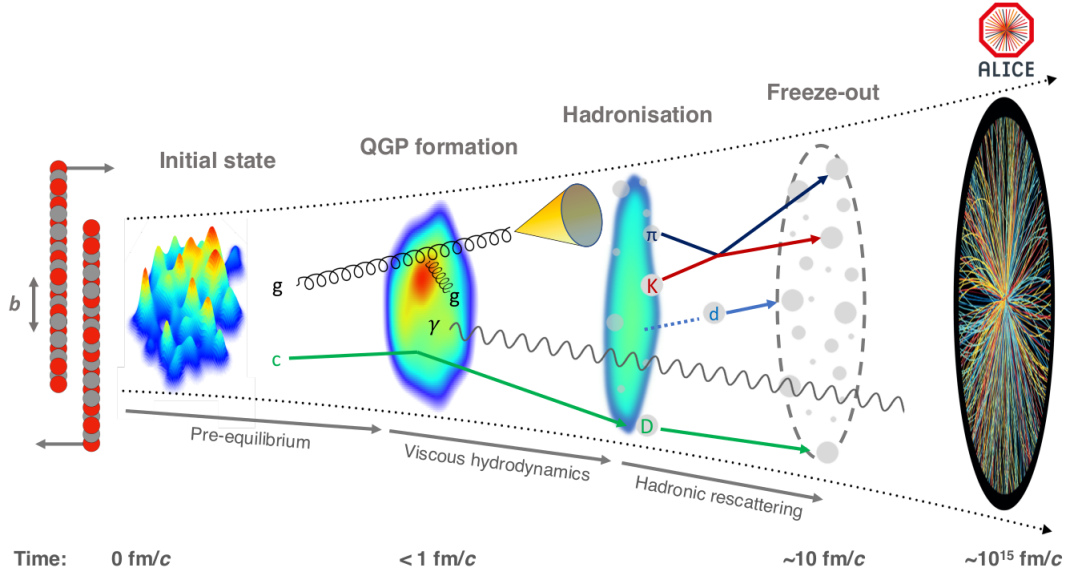


Figure 1.6: The stages of evolution of a heavy-ion collision at ALICE [3] (see text). b denotes the impact parameter (see section 1.3.3).

Fragmentation occurs when a highly energetic quark creates one or more quarks from the vacuum with which it then forms into a hadron. This process is expected to dominate in the high momentum range ($p_T > 8 \text{ GeV}$), where there are more quarks, and the resulting hadron has a momentum somewhere below that of the fragmenting parton.

Provided deconfined quarks are close enough to each other in the medium, they can also combine into hadrons directly, without the need to create additional quarks. This process is called recombination, or coalescence. Since the quarks have to be close enough to each other in both space and momentum, this process is expected to dominate at low and intermediate momentum. The momentum of the hadron created is essentially equal to the sum of the momenta of the quarks that combined into it.

Most hadronisation models that implement coalescence use the Wigner formalism. It describes the expected number of hadrons produced with a momentum P as

[67]

$$\frac{d^2 N_H}{d^2 P} = g_H \int \prod_{i=1}^n \frac{d^3 p_i}{(2\pi)^3 E_i} p_i d\sigma_i f_{q_i}(x_i, p_i) f_H(x_1, \dots, x_n, p_1, \dots, p_n) \delta^{(2)} \left(p_T - \sum_{i=1}^n p_{T,i} \right) \quad (1.4)$$

Here, g_H is a statistical factor to form a colourless hadron, $d\sigma_i$ is an element of a space-like hypersurface, and f_{q_i} and f_H are the Wigner functions of the partons and hadron respectively, which represent the space and momentum distribution of quarks. In the case of the Catania model [26], a Gaussian distribution is assumed for charmed hadrons, but different models implement this in different ways.

1.3.3 Properties of heavy-ion collisions

There are several parameters that characterise a heavy-ion collision. The analysis results are expected to vary depending on these collision parameters, making it important to measure them. Two kinematic properties that will be mentioned in this analysis, centrality and rapidity, are described here.

Rapidity

In collider experiments, the geometry of the collision is usually defined so that the z-axis is parallel to the incoming beams, allowing the position and direction of the particles to be classified using the polar angle ϕ transverse to it and the azimuthal angle θ parallel to it. However, instead of characterising the position using θ , it can be useful to use a quantity that is invariant to Lorentz boosts along the beam axis. This is especially true in heavy-ion collisions, where not all partons participate in the collision, meaning that the centre of mass of the collision moves along the beam axis and is essentially unknown. To this end, the rapidity y is defined:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z c}{E - p_z c} \right) \quad (1.5)$$

where E is the energy and $p_z = p \cos \theta$ is the momentum along the z -axis. For a particle moving directly transverse to the beam axis ($p = p_T$), the rapidity is 0, while for a particle travelling along the beam axis it approaches $\pm\infty$, depending on whether it travels forwards or backwards. Under a longitudinal Lorentz boost, this rapidity changes only by an additive constant, meaning that the difference in rapidity between two positions is invariant under the Lorentz boost.

For highly relativistic energies, the rapidity simplifies to

$$y \approx \eta = -\ln \left(\frac{\theta}{2} \right) \quad (1.6)$$

η is called the pseudo-rapidity. It depends only on the azimuthal angle θ , and is therefore straightforward to determine. Usually, a pseudo-rapidity range is defined for a given analysis based on detector acceptance, denoting the angles along the beam axis for which particles are considered.

Centrality

Lead nuclei are extended systems and are significantly larger than protons. This means that the collisions do not always happen head-on, but have a wide range of how centrally the two lead ions collide, as shown in figure 1.7. The amount of energy deposited in the collision greatly depends on how central the collision is, and how many nucleons partake in it, making it of great interest when investigating the properties of the QGP created in it. To quantify this, a variable called centrality is introduced, which is defined as

$$c = \frac{b}{R_1 + R_2} \quad (1.7)$$

where R_1 and R_2 are the radii of the colliding nuclei and b is the impact parameter, which represents the distance between the two centres of the nuclei at collision. This means a perfectly head-on collision is assigned a centrality of 0, while a completely peripheral collision has a centrality of 1.

The impact parameter cannot be measured directly, but instead can be found experimentally via two different methods: One is measuring the multiplicity of charged

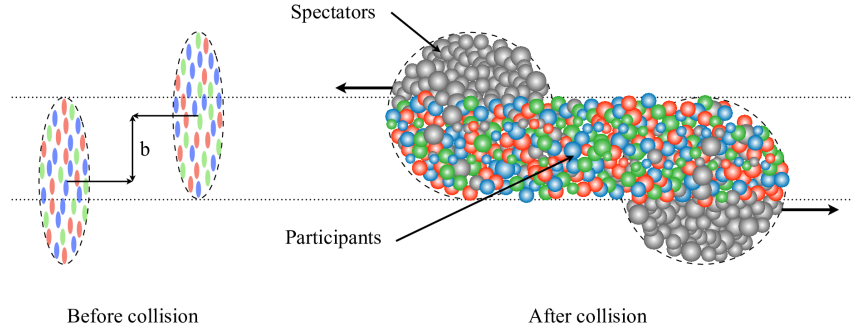


Figure 1.7: A schematic view of a heavy-ion collision showing the impact parameter b , which determines the centrality of the collision, as well as the participating and spectating nucleons [6].

particles N_{ch} , which depends monotonically on the impact parameter. The other is measuring the energy carried by particles close to the beam pipe, which corresponds to the number of spectator nucleons N_{spec} , i.e. nucleons not participating in the collisions. Figure 1.8 shows the charged particle multiplicity normalised by the number of participating nucleon pairs as a function of N_{part} , as calculated by a Glauber Model [68], which is described in more detail in section 3.2.4. It shows measurements from both Pb–Pb and Xe–Xe collisions done at ALICE, as well as results from RHIC measured in Au–Au and Cu–Cu collisions, which are scaled to account for the different collision energies. They are compared to predictions from PYTHIA 8.3 calculations [69] using two different selections, one on N_{ch} and one on N_{part} . At very central collisions there is a rise for the N_{ch} -selected calculations attributed to fluctuations of the charged-particle multiplicity at a fixed number of particle sources. This selection bias can be overcome by using the energy deposited by spectator nucleons for the centrality calculation (see section 3.2.4).

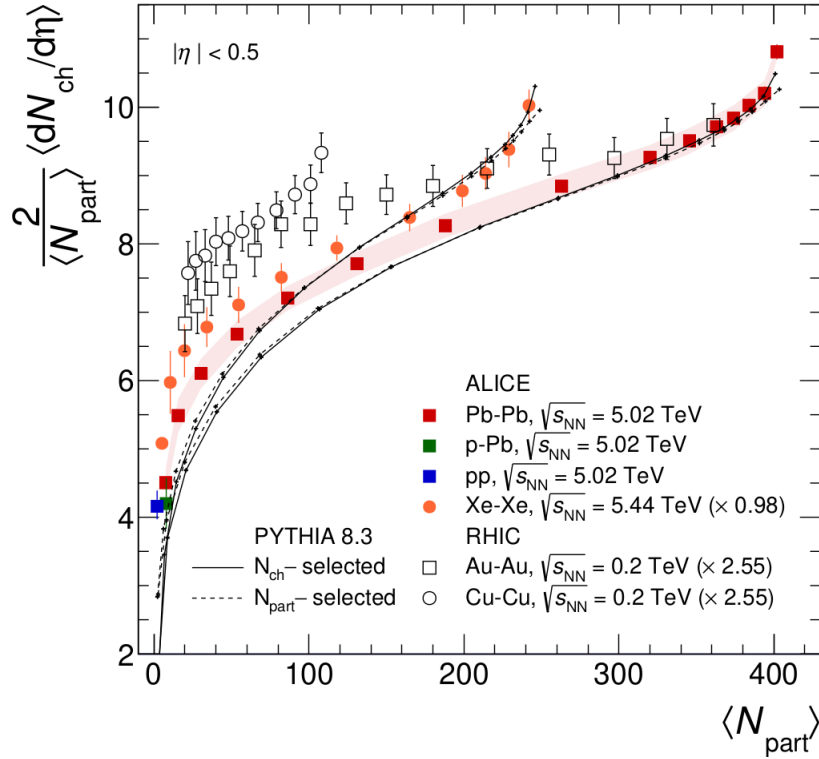


Figure 1.8: The measured normalised charged multiplicity as a function of the number of participating nucleons for different collision systems, scaled for different collision energies and compared to PYTHIA predictions (see text) [3].

1.4 Measurements in Heavy-ion collisions

An important observable that can be used to study the properties of the QGP is the particle multiplicity with respect to rapidity $\frac{dN}{dy}$, or the particle multiplicity measured differentially as a function of transverse momentum $\frac{d^2N}{dydp_T}$. The presence of a QGP is expected to have different effects on the number and momentum range of different hadrons produced. These can then be compared to the distribution of other particle species. Figure 1.9 shows the production yield of different light-flavour hadrons at mid-rapidity, integrated over the entire p_T -range, measured in central Pb-Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76 \text{ TeV}$. They are compared to four different theoretical

calculation that implement particle production via statistical hadronisation models (SHMs). SHMs assume a fully thermalised system before the QGP transitions into a hadron gas, and implement hadronisation in a purely statistical manner. These models describe the measurement results well, indicating that the total production yield of light flavour hadrons in Pb–Pb collisions is well understood.

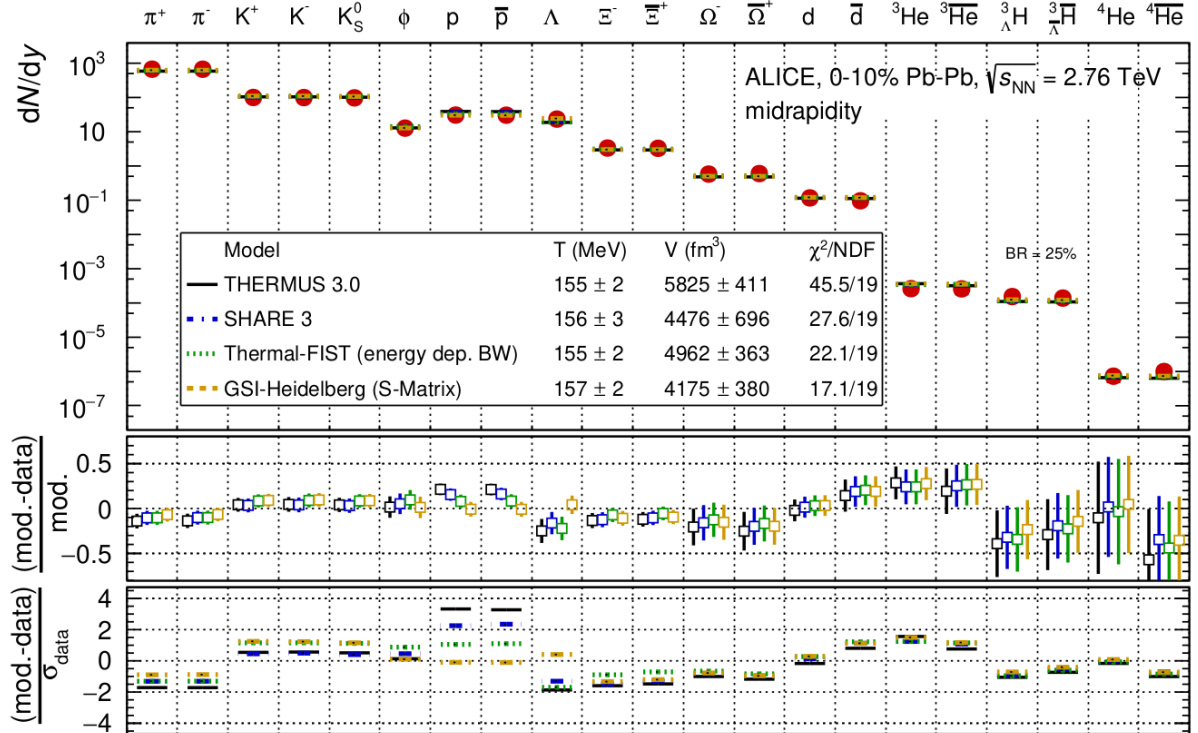


Figure 1.9: The production yield per unit of rapidity in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV for different hadron species measured by ALICE at mid-rapidity and over the full p_{T} range, compared to four different statistical hadronisation models. Figure taken from [3].

Another useful observable is the nuclear modification factor R_{AA} , which compares the particle production in heavy-ion collisions to those in proton-proton collisions. It is defined as

$$R_{\text{AA}}(p_{\text{T}}) = \frac{1}{\langle T_{\text{AA}} \rangle} \frac{dN_{\text{AA}}(p_{\text{T}})/dp_{\text{T}}}{d\sigma_{\text{pp}}(p_{\text{T}})/dp_{\text{T}}} \quad (1.8)$$

where $dN_{AA}(p_T)/dp_T$ is the p_T -differential production yield of a particle in AA collisions, and $d\sigma_{pp}(p_T)/dp_T$ is the p_T -differential cross section in pp collisions. It is corrected by the average nuclear overlap function $\langle T_{AA} \rangle$, which is found by dividing the average number of nucleon pairs participating in the collision N_{coll} by the inelastic nucleon-nucleon cross section. Due to this scaling, a nuclear modification factor of $R_{AA} = 1$ would indicate that a heavy-ion collision can be viewed as a superposition of multiple nucleon-nucleon collisions, where no QGP is formed. On the other hand, a deviation from 1 implies a fundamental difference in these collisions. It is expected that in the presence of a QGP, the R_{AA} of produced hadrons is significantly below unity at high p_T , due to the in-medium energy loss of the quarks and gluons. Another reason for modifications to hadron production can be so-called cold nuclear matter effects, which occur not due to the formation of a QGP but the presence of nuclei in the collision. To investigate these effects and separate them from QGP-related modifications, measurements in intermediate collision systems, such as p-Pb collisions, are crucial.

Figure 1.10 shows the nuclear modification factor for charged particles measured by ALICE in p-Pb collisions and central and peripheral Pb-Pb collisions. For intermediate and high p_T , the p-Pb measurement shows agreement with unity, which suggests that there is no in medium energy loss in this system. For Pb-Pb collisions on the other hand, the R_{AA} is significantly below unity, and a hierarchy from peripheral to central collisions can also be seen, with central collisions showing a R_{AA} substantially below that of peripheral ones. This suggests a significant energy loss of high-momentum quarks within the medium, and a strong dependence of this energy loss on the size of the QGP droplet that is produced.

Additionally, measurements of the relative production of baryons and mesons can give insights into how the presence of the QGP affects the hadronisation process. In this so-called baryon-to-meson ratio, the production yield of a baryon is compared to that of a meson with similar quark content. The presence of coalescence, for example, is expected to lead to an enhancement in the baryon-to-meson ratio at intermediate p_T . This is due to the aforementioned fact that the momentum of a hadron produced in a coalescence process is similar to the sum of the momenta of the

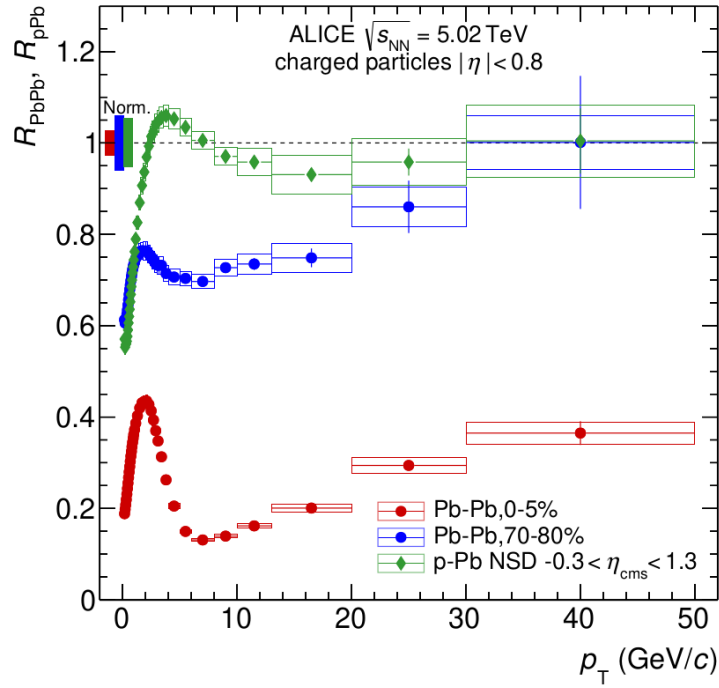


Figure 1.10: The nuclear modification factors for charged particles measured by ALICE for p–Pb collisions and central (0–5% centrality) and peripheral (70–80% centrality) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV as a function of p_{T} [7].

quarks that combine into it. Since baryons consist of three quarks and mesons of two, a baryon would have a higher momentum than a meson produced from quarks with the same momentum. In reverse, this means that a baryon and a meson of the same momentum are combined from quarks with different momentum. Since the number of produced quarks decreases with increasing p_{T} , one can therefore expect more baryons relative to mesons to be produced if coalescence occurs in the hadronisation process.

Figure 1.11 shows the Λ/K_{s}^0 ratio measured by ALICE in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV as a function of p_{T} , for different centrality classes, which are compared to a measurement done in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV. It shows that for very peripheral collisions, the Pb–Pb measurement is consistent with that from pp collisions, while for more central collisions a peak emerges at intermediate p_{T} . It is

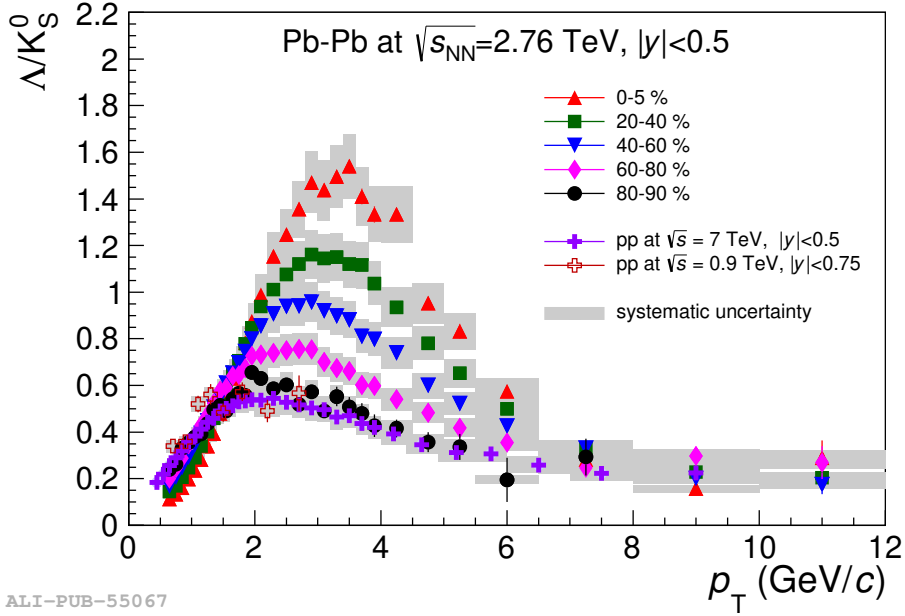


Figure 1.11: The Λ/K_s^0 ratio measured by ALICE in Pb-Pb collisions at $\sqrt{s_{NN}}=2.76$ TeV as a function of p_T for different centrality classes, and compared to measurements done in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV [8].

important to note however that when integrating the Λ and K_s^0 yields over the entire p_T range, they are consistent among all centrality classes with the pp measurements. This suggests that the presence of the medium leads to a redistribution of momenta rather than an increase in the total baryon production [8]. Another possible reason for this redistribution are hydrodynamical flow effects, which push the heavier baryons to higher momentum than the lighter mesons. Calculations implementing this effect are only applicable at low momenta of $p_T < 2$ GeV/c, above which they deviate from the measurements, but still present the need to disentangle them from the effects of the QGP medium.

Both Λ and K_s^0 are light flavour hadrons, meaning they consist of only up, down, and strange quarks. Their lighter mass means that they are produced at all stages of the collision, which means there are larger uncertainties when using these hadrons to study particle production. Heavy-flavour quarks on the other hand are produced

only at the beginning of the collision, making them a more precise probe into QGP properties and hadronisation mechanisms, as described in the next chapter.

Chapter 2

Heavy-flavour physics

Heavy-flavour quarks, such as the c and b quarks, are created in hard scattering processes at the beginning of heavy-ion collisions, when the energy is high enough to produce them. Therefore, their abundance is fixed at the start of the collision and they go on to experience the full evolution of the system, interacting with the quark-gluon plasma in elastic and inelastic collisions. This makes open heavy-flavour hadrons ideal probes for investigating the quark-gluon plasma.

This chapter details the theoretical framework used to describe the production of heavy-flavour quarks in pp and $Pb-Pb$ collisions and motivates the use of open heavy-flavour hadrons as probes into the properties of the quark-gluon plasma in its deconfined state as well as the hadronisation process. It also gives a brief outline of previous heavy-flavour measurements at ALICE, both in $Pb-Pb$ collisions and in small systems such as pp and $p-Pb$ collisions.

2.1 Heavy-flavour production in pp collisions

Protons consist of quarks and gluons, and a pp collision must therefore be described as many individual parton-parton interactions. The production cross section of a heavy-flavour hadron h in a pp collision can be expressed as [70]

$$d\sigma_{pp \rightarrow h+X}^{hard} = f_i^a(x_i, Q^2) \otimes f_j^b(x_j, Q^2) \otimes \sigma_{ij \rightarrow q+X}^{hard}(x_i, x_j, Q^2) \otimes D_{q \rightarrow h}(z, Q^2) \quad (2.1)$$

where $f_{i,j}^{a,b}(x_{i,j}, Q^2)$ is the parton distribution function (PDF) describing the probability of finding a parton species i carrying the fraction x_i of the momentum of the proton, and $\sigma_{ij \rightarrow q+X}^{hard}(x_i, x_j, Q^2)$ is the scattering cross section of two partons i and j forming the parton species q . Since heavy-flavour hadrons are produced in hard scattering processes with a large momentum transfer, the coupling constant is small, and the later can be described with pQCD methods. The last term, $D_{q \rightarrow h}(z, Q^2)$, represents the likelihood of a parton q fragmenting into the hadron h , called fragmentation function. These fragmentation functions were long believed to be universal, and once measured in one collision system are used as input into theoretical calculations for other collision systems [71]. As will be discussed later in this chapter, this assumption has since been challenged.

2.2 Heavy-flavour production in Pb–Pb collisions

In ultra-relativistic heavy-ion collisions such as Pb–Pb collisions at the LHC, the ambient temperature of the QCD matter produced is lower than the mass of the heavy-flavour quarks: $m_Q \ll T_{QGP}$. This means that heavy quarks are only created in the initial hard scattering between constituents quarks or gluons within the nuclei, but not in the quark-gluon plasma itself. While they lose energy travelling through the medium, their multiplicity is fixed.

Similar to equation 2.1, the production cross section of a heavy-flavour hadron h in a heavy ion collision can be expressed as

$$d\sigma_{AA \rightarrow h+X}^{hard} = A \cdot A \cdot f_i^N(x_i, Q^2) \otimes f_j^N(x_j, Q^2) \otimes d\sigma_{ij \rightarrow Q\bar{Q}+X}^{hard}(x_i, x_j, Q^2) \otimes R_N^{QGP}(E, m_Q, T, \alpha, \dots) \otimes D_{q \rightarrow h}^N(z, Q^2) \quad (2.2)$$

Here, A is the mass number of the colliding nuclei, and $\otimes$ represents a convolution over each term. $d\sigma_{if \rightarrow Q\bar{Q}+X}^{hard}(x_i, x_j, Q^2)$ is the partonic production cross section of the heavy quark within the hadron, and $f_{i,j}^N(x_{i,j}, Q^2)$ are the parton distribution functions describing the probability to find a parton species i carrying the fraction

x_i of the momentum of the nucleon. They are given by

$$f_i^N(x_i, Q^2) = R_i^N(x_i, Q^2)f_i(x_i, Q^2) \quad (2.3)$$

$f_i(x_i, Q^2)$ is the PDF of a particle species i , and $R_i^N(x_i, Q^2)$ is the nuclear modification applied to it to take into account the parton distribution in a nucleus compared to that of a free nucleon.

The last two terms of equation 2.2 are taking into account the effects of the medium. $R_N^{QGP}(E, m_Q, T, \alpha, \dots)$ denotes the influence of energy loss in the medium, as well as collective flow, on the heavy quark momentum. $D_{q \rightarrow h}^N(z, Q^2)$ represents a fragmentation function in the medium, which is modified by other hadronisation mechanisms present in the medium, such as coalescence, which is described in the previous chapter in 1.3.2.

2.2.1 Interactions in the QGP

Studying the interactions of heavy quarks in the QGP can give valuable insights into the properties of the QGP. This is done using transport models to simulate the interactions of heavy quarks with the medium partons. The predicted particle cross sections given by these models can then be compared to experimental data. The models used to compare to the experimental results presented in this thesis are described in chapter 5, but the theoretical basis for them is outlined in this section.

Heavy flavour quarks interact with the medium in two different ways. At the low momentum range, this interaction mainly happens through elastic collisions, while at higher momentum energy is lost via gluon radiation. The total energy loss can be described by $\Delta E = \Delta E_{\text{coll}} + \Delta E_{\text{rad}}$.

The movement of a heavy flavour particle through QCD matter can be described by a Boltzmann equation [72]. The phase space evolution of the heavy quark phase space distribution function f_Q is given by:

$$\left[\frac{\partial}{\partial t} + \frac{\vec{p}}{E_p} \frac{\partial}{\partial \vec{x}} + \vec{F} \frac{\partial}{\partial \vec{p}} \right] f_Q(t, \vec{x}, \vec{p}) = C[f_Q] \quad (2.4)$$

Here, $E_p = \sqrt{m_Q^2 + \vec{p}^2}$ is the quark energy, $\vec{F}$ is a force induced by an external field, and $C[f_Q]$ is the collisional integral. For a static medium in equilibrium, the distribution function follows the Boltzmann distribution: $f_Q = d_Q e^{-E_p/T}$. d_Q here denotes the heavy quark degeneracy, and T is the temperature of the medium. The first two terms on the left of eq. 2.4 govern the space-time evolution of f_Q , while the third term represents the effect of the total mean field that affects the quark, especially in the early stages of the collision. The collisional integral $C[f_Q]$ on the right side of the equation takes into account the scattering amplitude between the heavy quark and the medium partons. For heavy quarks such as the charm quark, the relation $m_Q \gg T$ holds. Additionally, the momentum transfer is typically small, which suggests that the quarks experience multiple incoherent collisions that are typical of a Brownian motion. The Boltzmann equation can therefore be transformed into a Fokker-Planck equation:

$$\frac{\partial}{\partial t} f_Q(t, \vec{p}) = \gamma \frac{\partial}{\partial p_i} [p_i f_Q(t, p)] + D_p \Delta_{\vec{p}} f_Q(t, p) \quad (2.5)$$

Where γ is the drag coefficient, which describes the energy loss on the medium per unit of time in the absence of fluctuation. D_p is the momentum diffusion coefficient, which represents the momentum broadening per unit of time. In the thermal equilibrium, the two coefficients are linked with the Einstein relation:

$$D_p = m_Q \gamma T \quad (2.6)$$

Here, m_Q is the mass of the heavy quark and T is the medium temperature. The drag coefficient is also related to the spatial diffusion constant D_s :

$$D_s = \frac{T}{m_Q \gamma} \quad (2.7)$$

The Fokker-Planck equation can be solved stochastically using a Langevin model [73]. For a non-relativistic limit, the changes in position and momentum are given

by

$$\Delta \vec{x} = \frac{\vec{p}}{m_Q} \Delta t \quad (2.8)$$

$$\Delta \vec{p} = (-\gamma \vec{p} + \zeta) \Delta t \quad (2.9)$$

where ζ is given by a Gaussian probability distribution.

Gluon radiation is a $2 \rightarrow 3$ process, in which a gluon is radiated. The relative amplitude of the different scattering processes is given by the Casimir factor C_R used in these calculations:

$$\begin{aligned} C_R &= N_c \quad \text{for } g \rightarrow gg \\ C_R &= (N_c^2 - 1)/2N_c \quad \text{for } g \rightarrow qg \\ C_R &= 1/2 \quad \text{for } g \rightarrow q\bar{q} \end{aligned} \quad (2.10)$$

Here, $N_c = 3$ is the number of quark colours. The number of gluons radiated off gluons is therefore $9/4$ times higher than that radiated off quarks. The probability of gluon radiation is calculated in [74]. Here, a soft scattering approximation is done using the BDMPS formalism. In it, the characteristic gluon energy is given by

$$\omega_c = \frac{1}{2} \hat{q} L^2 \quad (2.11)$$

Here, L is the path length in the medium, and $\hat{q}$ is the transport coefficient. For gluons with energy $\omega < \omega_c$ and a kinematic constraint $R = \omega_c L \rightarrow \infty$, the energy distribution of the radiated gluons can be estimated as

$$\omega \frac{dI}{d\omega} \approx \frac{2\alpha_s C_R}{\pi} \sqrt{\frac{\omega_c}{2\omega}} \quad (2.12)$$

with α_s the strong coupling constant. The average parton energy loss is found by integrating this energy distribution:

$$\langle \Delta E \rangle = \int_0^\infty d\omega \omega \frac{dI}{d\omega} = \frac{\alpha_s C_R}{2} \omega_c \quad (2.13)$$

With equation 2.11 this gives a L^2 -dependence of the average energy loss.

Gluon radiation is suppressed at angles smaller than the ratio of the quark mass to its energy. This is known as the “dead-cone” effect [75] and has recently been observed directly by the ALICE collaboration [76]. It means that the suppression of energy loss via radiation decreases for heavier quarks. This, together with the different Casimir factors, leads to a flavour-dependent hierarchy of radiative energy loss:

$$\Delta E_{\text{rad}}(g) > \Delta E_{\text{rad}}(q) > \Delta E_{\text{rad}}(c) > \Delta E_{\text{rad}}(b) \quad (2.14)$$

Here, q denotes the light quarks u, d and s .

By measuring the relative abundances of different hadrons in Pb–Pb collisions compared to those in pp or p–Pb collisions, further information can be gained about the in-medium energy loss of quarks and gluons in the QGP.

2.3 Previous measurements

Heavy-flavour hadrons have a very short lifetime, which presents a difficulty in studying them as only their decay products reach the detector. There are different strategies to utilise these to study heavy-flavour production, depending on the goal of the analysis and the data available. The most conceptually straightforward way is to choose a known decay channel of a particle and fully reconstruct it. This involves topological and kinematic cuts on the decay product and the secondary decay vertex, and is the approach used in this thesis, further described in chapter 4. Another approach is to look for leptons stemming from heavy-flavour decays. This can be done either by looking for displaced leptons from secondary decay vertices, or by measuring the lepton yield and subtracting known sources of lepton production. A third option is to reconstruct jets. By reconstructing tracks and particles inside them, heavy-flavour jets can be distinguished from light-flavour jets [70].

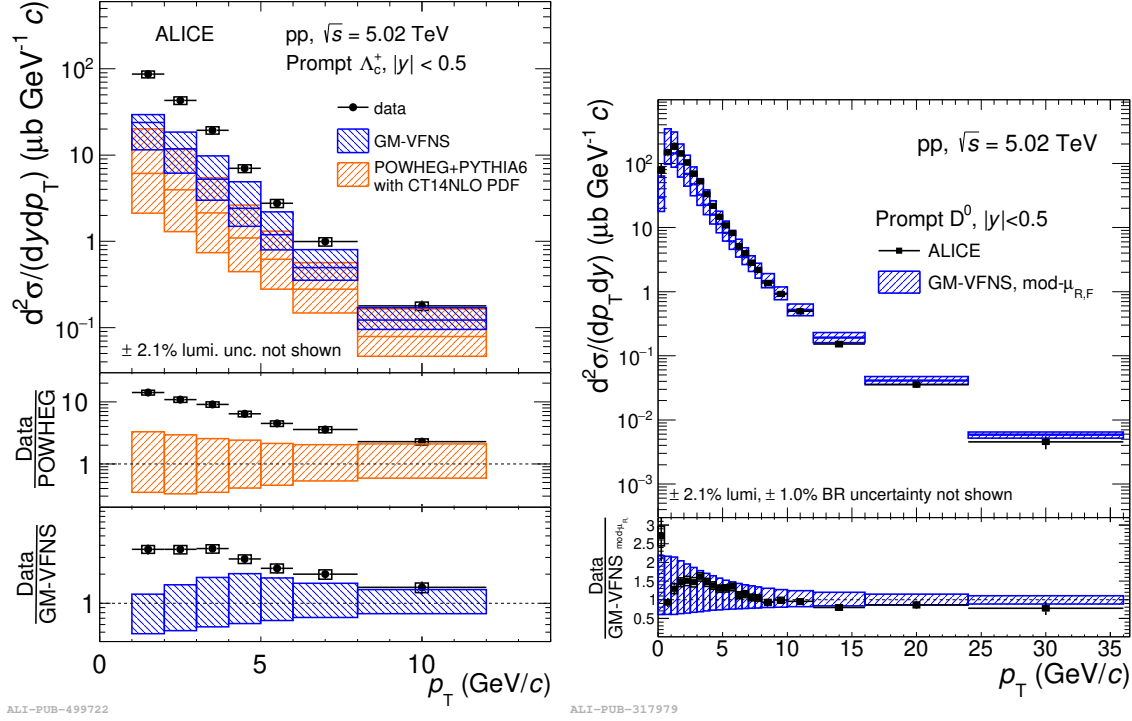


Figure 2.1: The p_T -differentiated production cross section of Λ_c^+ baryons (left) and D^0 mesons (right) measured by ALICE in pp collisions at $\sqrt{s} = 5.02$ TeV [9, 10]. They are compared to different pQCD calculations [11, 12] (see text).

2.3.1 Small systems

While the QGP is only expected to occur in heavy-ion collisions, measurements of heavy-flavour production are also carried out in small systems such as proton-proton and proton-lead collisions. Aside from providing a reference frame for heavy-ion measurements, important information can also be gained on the QCD processes in these systems. This includes studying the parton distribution function of protons, as well as gaining insight into hadronisation processes and the fragmentation function.

The production of Λ_c^+ baryons was measured by the ALICE collaboration in proton-proton collisions at collision energies from $\sqrt{s} = 5$ to 13 TeV [9, 25, 77], and is shown in the left panel of figure 2.1 for $\sqrt{s} = 5.02$ TeV. The p_T -differential cross

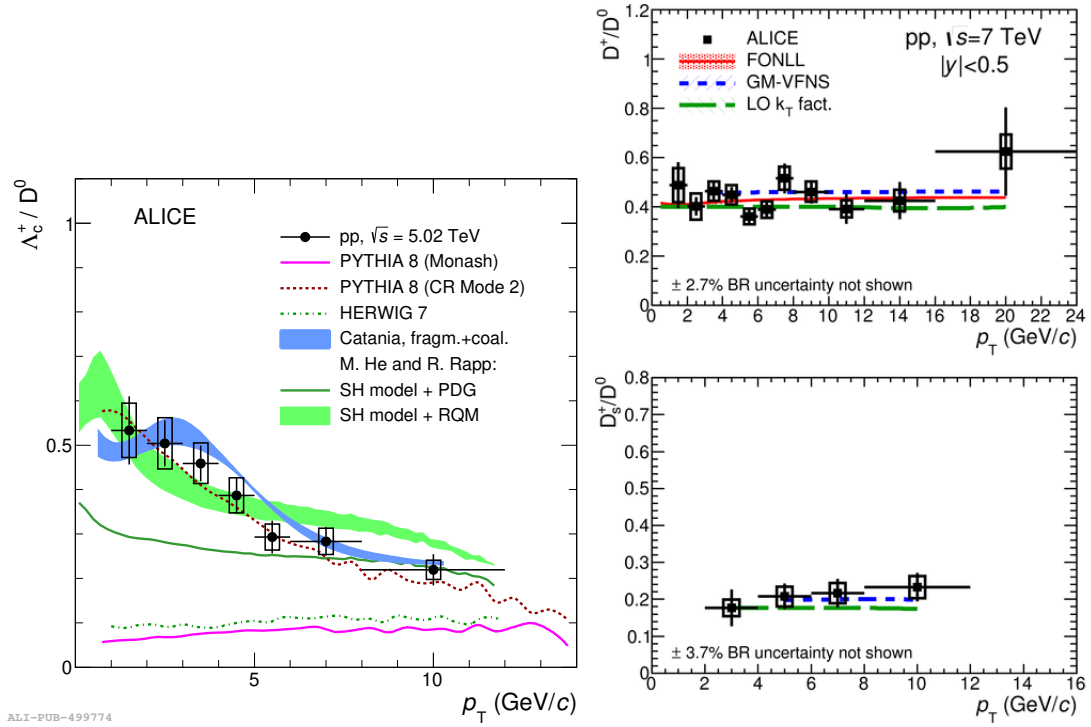


Figure 2.2: Left: the Λ_c^+/D^0 baryon-to-meson ratio measured by ALICE in pp collisions at $\sqrt{s}=5.02$ TeV, compared to several models. Right: The D^+/D^0 and D_s^+/D^0 ratio measured in pp collisions at $\sqrt{s} = 7$ TeV, also compared to several models [13].

section is compared to several different model predictions. The left side shows a comparison to predictions from analytical pQCD calculations using GM-VFNS [12] and next-to-leading order pQCD calculations using the POWHEG event generator [11]. Both underestimate the Λ_c^+ prediction, with the GM-VNFS calculation being closer to data. On the other hand measurements of the D^0 production yield also measured in pp collisions at $\sqrt{s} = 5$ to 13 TeV show a good agreement with predictions from the GM-VNFS calculation, as seen in figure 2.1 for $\sqrt{s_{NN}} = 5.02$ GeV (right panel). This indicates that there is something wrong in the pQCD models used for these predictions, which use pQCD quark production calculations convoluted with parton distribution functions and fragmentation functions.

Fragmentation functions can be probed by measuring the production ratios of different hadrons. To this end, the Λ_c^+/D^0 ratio was also measured at $\sqrt{s} = 5$ to 13 TeV [78], and is shown in figure 2.2 for $\sqrt{s} = 5.02$ TeV (left). It was found that models using a fragmentation function tuned on e^+e^- measurements severely under-predict the measurement. In addition, they predict a flat p_T dependence, which is also not seen in the data. This hints at a fundamental difference in the fragmentation process between e^+e^- and pp collisions. On the other hand, models that include coalescence in pp collisions, colour reconnection beyond the leading order, or additional undiscovered charmed baryon states [26, 79, 15], describe the data well. A different picture is seen when measuring the production ratios of different D mesons. The right panel of figure 2.2 shows the ratios of D^+ over D^0 mesons (top), and D_s^+ mesons over D^0 mesons (bottom), both measured at $\sqrt{s} = 7$ TeV. As seen in figure 2.1 (right), the models using fragmentation functions tuned on measurements from e^+e^- collisions describe the D-meson measurement within uncertainties, and the ratios do not appear to exhibit a dependence on momentum.

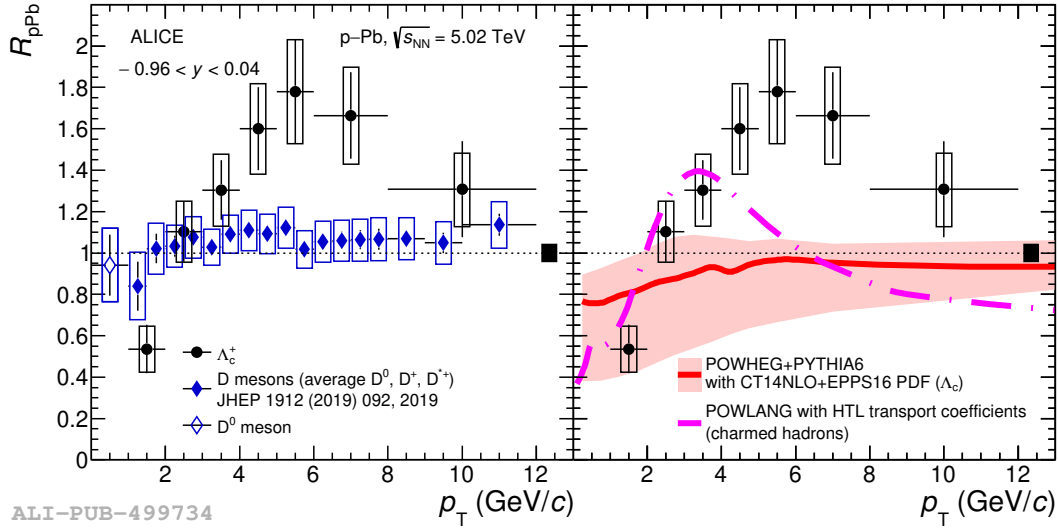


Figure 2.3: The nuclear modification factor R_{pPb} of Λ_c^+ baryons measured by ALICE in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared to that measured for D mesons (left) and to two models (right) [9]

Similar measurements were also done in p–Pb collisions [9]. Figure 2.3 shows the modification factor R_{pPb} for Λ_c^+ baryons also measured at $\sqrt{s} = 5$ to 13 TeV. It hints at an enhancement at intermediate p_T for p–Pb collisions in relation to pp collisions at intermediate momentum, which is not found for D mesons. The right panel shows a comparison to two different models. One is the POWHEG+PYTHIA model [11], which uses a pQCD calculation and a modification to the parton distribution function in the nucleus. The other is the POWLANG model [80], which implements the formation of a small QGP. While the latter qualitatively follows the p_T dependency seen in the data, neither model describes the measurement in a quantitative way, making it clear that enhancement of the R_{pPb} is not yet fully understood.

2.3.2 Pb–Pb collisions

Several different heavy-flavour measurements have been done in Pb–Pb collisions at ALICE. Figure 2.4 shows the nuclear modification factor as a function of p_T for prompt and non-prompt mesons, measured in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. While prompt mesons are produced directly through the hadronisation of a charm quark, non-prompt D mesons are decay products of beauty hadrons. By utilising the displacement of the secondary decay vertex, the meson yields from the two different sources can be separated. At intermediate p_T , a clear hint emerges that non-prompt mesons are less suppressed compared to prompt mesons. This indicates a greater suppression of charm quarks compared to the heavier bottom quarks. This is in line with the dead cone effect discussed in section 2.2.1, which predicts a hierarchy of radiative energy loss dependent on quark flavour masses.

A comparison of the yields found for different charmed hadron species (non-strange [24] and strange D mesons [81], Λ_c^+ baryons [23], and J/ψ mesons [82]) is shown in figure 2.5. They were measured at mid-rapidity in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. For D^0 and J/ψ mesons the measurements were done down to $p_T = 0$, while for the other D mesons and Λ_c^+ baryons the yield at low p_T was extrapolated using the differential production yield ratios relative to D^0 measurements. The measurements are compared to a statistical hadronisation model [15],

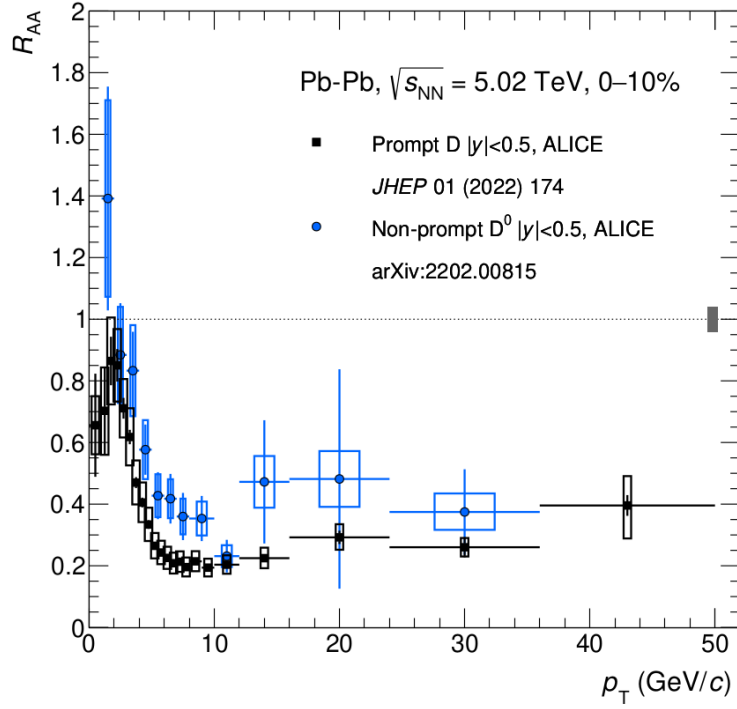


Figure 2.4: The nuclear modification factor R_{AA} for prompt and non-prompt D meson production measured by ALICE in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV [14].

which is described in section 5.3. While the measurements agree with the model within uncertainties for the mesons, there is a larger discrepancy for the Λ_c^+ yield. This is mitigated when additional excited baryon states beyond those in the PDG are added to the model.

The Λ_c^+ measurements shown in figure 2.5 are done in the $\Lambda_c^+ \rightarrow \text{p}K_{\text{S}}^0$ channel. They will be merged with the results of this analysis to increase the overall precision. The merging process and the results of this are described in chapter 5.

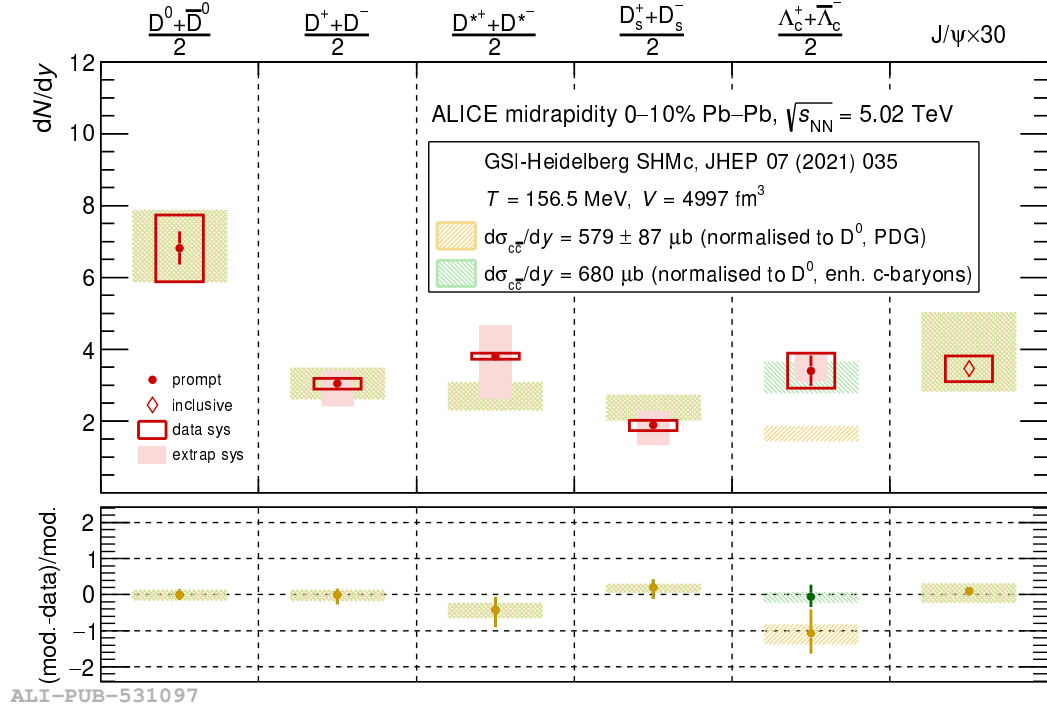


Figure 2.5: The p_T -integrated production yields per units of rapidity for different charm-hadron species measured by ALICE in central Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ GeV, compared to predictions from a statistical hadronisation model [3, 15].

2.4 Aim of this thesis

As described in this chapter, heavy flavour measurements in Pb-Pb collisions are fundamental to understanding the formation and hadronisation of QGP in heavy-ion collisions. In this thesis, the production of Λ_c^+ baryons is measured for the first time in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. It is merged with a previous measurement done in the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel, making this combined measurement the most precise measurement of Λ_c^+ production in Pb-Pb collisions to date. Using previous measurements in pp collisions as well as measurements of D^0 production in Pb-Pb collisions, the nuclear modification factor R_{AA} and the Λ_c^+/D^0 ratio are also measured. The $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel presents a serious challenge due to its three-body final state and the short lifetime of the Λ_c^+ ,

with the decay length being of the same order as the detector resolution. This means this analysis is at the very limit of what can be measured with this data, and is more precise in the high momentum region, where there is less background present.

Chapter 3

Experimental setup

The data analysed in this work was taken with the ALICE experiment at CERN. ALICE (A Large Ion Collider Experiment) is one of the four main experiments at the Large Hadron Collider (LHC), situated at Point 2. It is purpose-built to measure collisions of heavy ions. At present, these are lead ions (Pb–Pb), but there is a possibility of colliding other ions as well such as Xe–Xe. ALICE also takes data of proton-proton (pp) and proton-lead (p–Pb) collisions. In this chapter, the ALICE experimental setup is described. Section 3.1 describes the LHC and the additional pre-accelerators, which accelerate the lead ions and protons to ultra-relativistic energies. Section 3.2 describes the general setup of the ALICE detector and an overview of the different sub-detectors used in this analysis for triggering and particle tracking and identification. Finally, sections 3.3, 3.4, and 3.5 cover trigger, tracking, and particle identification procedures respectively.

3.1 The LHC

The Large Hadron Collider (LHC) is the largest particle accelerator in the world [83]. It was built at the European Organisation of Nuclear Research (CERN), located near Geneva, and began operation in 2008. It is a circular collider with a circumference of 27 km, located about 100 m underground. It was built mainly for

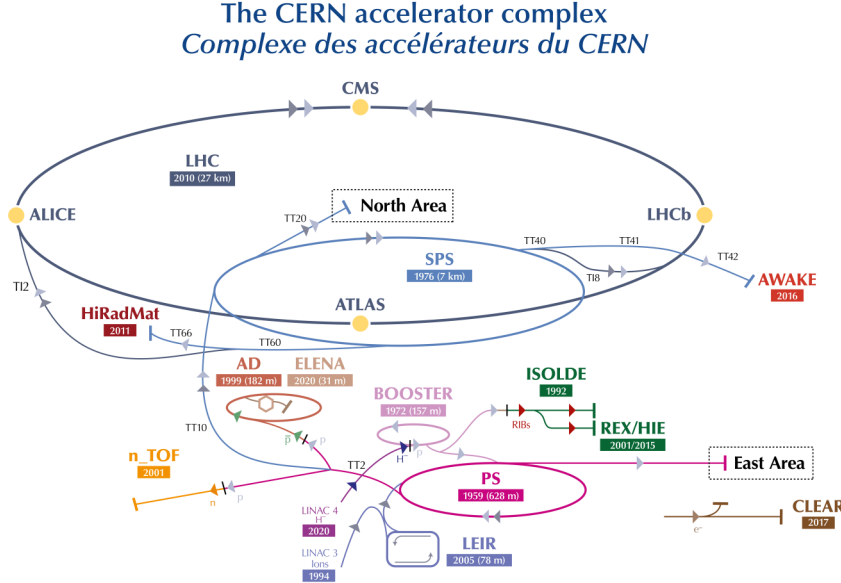


Figure 3.1: The CERN accelerator complex [16].

proton-proton collisions but also collides heavy ions. For proton-proton collisions, the planned maximum centre-of-mass energy was $\sqrt{s} = 14 \text{ TeV}$, with a peak luminosity of $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. For lead nuclei collisions, this was $\sqrt{s_{\text{NN}}} = 5 \text{ TeV}$ and $\mathcal{L} = 10^{27} \text{ cm}^{-2} \text{ s}^{-1}$. In order to ensure comparability between the collision systems, the proton-proton data is also taken at the same centre of mass energy per nucleon-nucleon collision as the lead-lead collisions. In Run 2, which took place between 2015 and 2018 and the data of which is analysed in this thesis, a maximum centre-of-mass energy of $\sqrt{s} = 14 \text{ TeV}$ was reached for proton-proton collisions, and $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$ for lead-lead collisions. Following the long shutdown period 2 (LS2, 2018-2022), the LHC restarted operation in 2023 for Run 3, so far reaching a peak centre-of-mass energy of $\sqrt{s} = 13.6 \text{ TeV}$ for proton-proton collisions and $\sqrt{s_{\text{NN}}} = 5.36 \text{ TeV}$ in lead-lead collisions. While this is slightly lower than it was during Run 2, the luminosity is much higher in Run 3, targeted to be about $\mathcal{L} = 210^{34} \text{ cm}^{-2} \text{ s}^{-1}$ for pp collisions. This vastly increases the amount of data that can be taken.

Several pre-existing accelerators from earlier experiments are used as a system of pre-accelerators. Figure 3.1 shows the accelerator complex as well as the various experiments that are operating at CERN. The protons used in the LHC are obtained by stripping the electrons off hydrogen atoms from a hydrogen bottle. The first acceleration stage is done in the linear accelerator LINAC2, up to an energy of 50 MeV. They are then injected into the circular Proton Synchrotron Booster (PSB), where they are accelerated to an energy of 1.4 GeV. From there, they are sent to another circular collider, the Proton Synchrotron (PS), which accelerates them to an energy of about 25 GeV. The last pre-accelerator is the Super Proton Synchrotron (SPS), which brings the protons up to an energy of 450 GeV. They are then injected into the LHC as two beams going in opposite directions, where they are accelerated to their maximum collision energies before they are collided at the interaction points (IP). For lead ions, the acceleration procedure is slightly different. They are first created by heating a lead sample up to 800 °C to form a lead vapour, which is then ionised. The first acceleration step is performed by a linear accelerator called LINAC3, which brings the lead ions to an energy of 4.2 MeV. They are then sent to the Low Energy Ion Ring (LEIR), where they are brought to about 72 MeV. From here, they are injected into the PS and follow the same acceleration path as protons. They reach the LHC with an energy of 177 GeV. In the LHC, both beams are accelerated to 2.51 TeV each.

The four main experiments operating around the LHC ring are two general-purpose detectors (ATLAS [84] and CMS [85]), one detector built for b-quark physics (LHCb [86]), and the ALICE experiment [19] which is described in this chapter. The Pb–Pb data used in this work was taken during Run 2 of the LHC in 2018.

3.2 The ALICE detector

The ALICE detector is a general-purpose detector designed specifically for the high-multiplicity environment of heavy-ion collisions. It was built to study the colour-deconfined state of matter that existed immediately before the Big Bang, the so-called quark-gluon plasma (QGP). It is also capable of measuring proton-proton (pp)

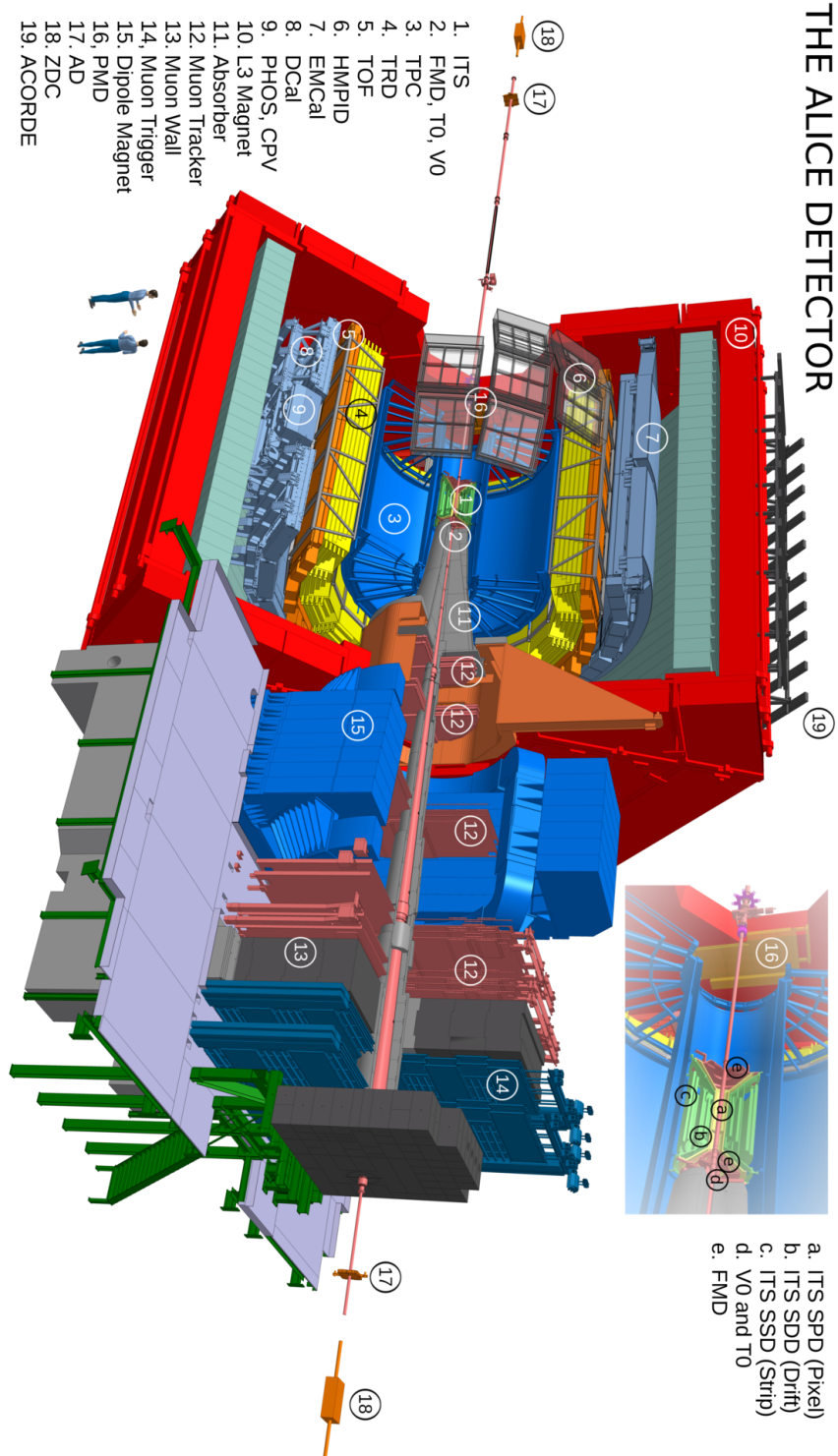


Figure 3.2: The ALICE detector setup for Run 2 [17]. The beam-pipe is surrounded by the central barrel, which includes the Inner Tracking System (ITS) that is shown in an insert. On the right side of the image, the forward parts of the detector system are located, such as the dipole magnet and muon spectrometer.

and proton-lead (p-Pb) collisions, which can be used as baseline collision systems when investigating the properties of the QGP.

The detector, which is shown in figure 3.2, measures $26 \times 16 \times 26 \text{ m}^3$ and weighs 10 000 t. It is formed of two main components: the central barrel, which is arranged around the beam pipe at the collision point, and the forward muon spectrometer, which is situated further along the beam-pipe with respect to the collision point. The central barrel covers a pseudorapidity range of $-0.9 < \eta < 0.9$ and the full azimuthal range. It is placed within a large solenoid magnet, which was previously used by the L3 experiment at LEP¹, and creates a uniform magnetic field along the beam axis of up to 0.5 T. Going from the centre outwards, the central barrel is composed of the Inner Tracking System (ITS) [89], the Time Projection Chamber (TPC) [90], the Transition Radiation Detector (TRD) [91], the Time Of Flight detector (TOF) [92], the Ring Imaging Cherenkov detector (HMPID) [93], and three electromagnetic calorimeters (PHOS [94], EMCal [95], and DCal [96]). The forward muon spectrometer covers the pseudorapidity range of $-4 < \eta < -2.5$, i.e. the forward direction, and the full azimuthal range. Its purpose is the detection and tracking of muons. To achieve this, it has several absorber layers, cathode plate chambers, and resistive plate chambers, as well as a dipole magnet. In addition to the central barrel and forward muon spectrometer there are several other smaller detectors located at forward and backward rapidity, whose main task is global event characterisation and triggering. These are the Zero Degree Calorimeter (ZDC) [97], Photon Multiplicity Detector (PMD) [98], Forward Multiplicity Detector (FMD) [99], T0, V0A and V0C. There is also an array of scintillators (ACORDE) [100], which is placed on top of the L3 magnet in order to trigger on muons coming from cosmic rays. This is used for alignment purposes and cosmic ray measurements. In this section, the main sub-detectors used in this analysis will be described in more detail.

The coordinate system used by the ALICE experiment is a right-handed Cartesian coordinate system. Its origin point lies on the interaction point 2 (IP2), which is where the beam collision happens. The z axis is oriented along the beam line,

¹The L3 experiment [87] operated at the Large Electron-Positron (LEP) collider [88], which until the year 2000 ran in the tunnel that is now occupied by the LHC.

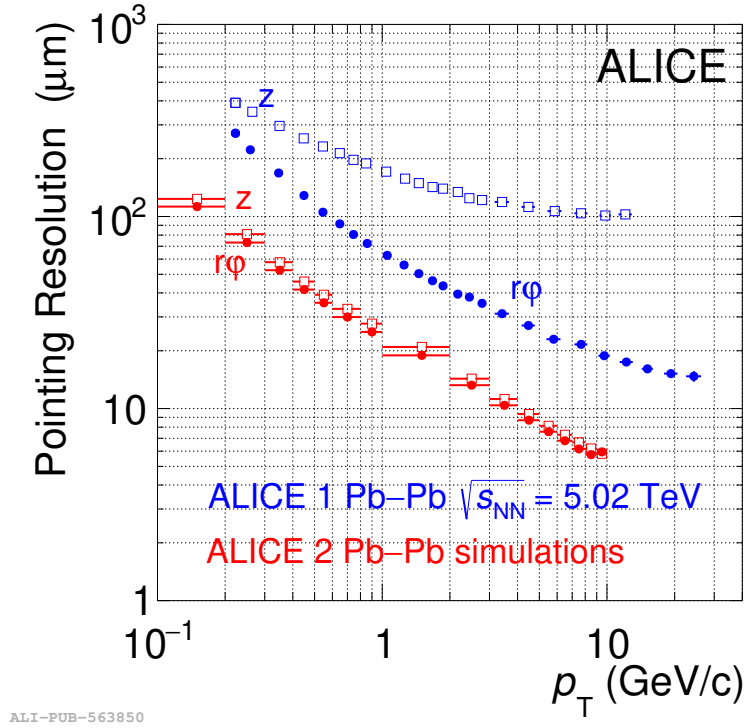


Figure 3.3: The impact parameter resolution of ALICE for Pb–Pb collisions in Run 2, compared to that projected for Run 3. [18] The open and closed markers show the longitudinal and transverse directions, respectively.

and points away from the forward muon spectrometer. The x axis points towards the centre of the LHC, and the y coordinate points upwards. A spherical coordinate system is also used. Here, the polar angle ϕ is defined in the $y - z$ plane, and the azimuthal angle θ is defined in the $x - y$ plane.

Figure 3.3 shows the detector resolution of the track impact parameter for Pb–Pb collisions in the longitudinal (along the beam axis) and transverse direction-. It also includes a simulation of the performance expected of the updated detector in Run 3.

3.2.1 The Inner Tracking System

The ITS is the detector placed closest to the beam pipe with an inner radius of 3.9 cm. Its main task is the reconstruction of charged particle tracks as well as primary and

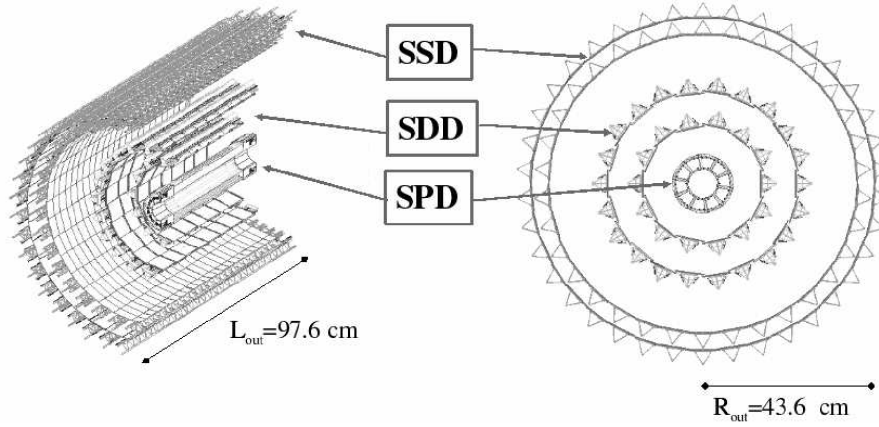


Figure 3.4: Schematic of the Inner Tracking System used in ALICE in Run 1 and Run 2 [19]. Shown are the six silicon detector layers, consisting of two layers each of SPD, SDD, and SSD.

secondary vertices. It consists of six layers of silicon detectors, divided into three subsections of different detector technologies. The inner two layers are called the Silicon Pixel Detector (SPD). The SPD is composed of hybrid silicon pixels, arranged in two-dimensional sensor matrices formed of 256×160 pixels with dimensions of $50 \mu\text{m}$ in the $r\phi$ direction and $425 \mu\text{m}$ in z . Its granularity enables a good spatial resolution, and it covers a pseudorapidity range of $|\eta| < 1.98$. Combined with the forward multiplicity detectors (FMD) which have a pseudorapidity range of $-3.4 < \eta < 5.1$, this provides full coverage for measuring charged-particle multiplicity.

The middle two layers of the ITS make up the silicon drift detector (SDD). It has 260 modules, with an active area of $72.7 \times 75.3 \text{ mm}$ each. The modules are divided into two drift volumes by a high-voltage central cathode, with electrons moving in opposite directions due to an electric field. The silicon strip detector (SSD) consists of the outer two layers of the ITS. These are two-sided modules that each contain 768 parallel strips with a pitch of $95 \mu\text{m}$. Both the SDD and the SSD cover a pseudorapidity region of $|\eta| < 0.9$. The technical details of the six ITS layers are summarised in table 3.1.

Layer	Detector	$r(\text{cm})$	$z(\text{cm})$	Material Budget ($\%X_0$)	Spatial Resolution $r\phi \times z(\mu\text{m}^2)$
1	SPD	3.9	14.1	1.14	12×100
2	SPD	7.6	14.1	1.14	12×100
3	SDD	15.0	22.2	1.13	35×25
4	SDD	23.9	29.7	1.26	25×25
5	SSD	38.0	43.1	0.83	20×830
6	SSD	43.0	48.9	0.86	20×830

Table 3.1: Properties of the six ITS layers. Numbers taken from [29].

For Run 3, a new Inner Tracking System has been installed in ALICE to replace the one used in Run 1 and Run 2, called ITS2 [101], which greatly improves tracking and vertexing capabilities.

3.2.2 The Time Projection Chamber

TPC is the main tracking detector in the central barrel. It covers the full azimuthal range, with a radial coverage of $85\text{ cm} < r < 247\text{ cm}$, and a length of about 5 m, it covers a pseudorapidity range of $-0.9 < \eta < 0.9$. It has an active volume of 90 m^3 , which is filled with a changing mixture of gas. In Run 1, this was a mixture of Ne-CO₂-N (90-50-5), which was changed to Ar-CO₂ (88-12) for Run 2. The gas mixture is chosen to optimise drift speed, have a low radiation length, keep diffusion low, and be chemically stable. Charged particles hitting the gas volume ionise the gas. The electrons created are accelerated towards one of the two end plates by an electric field with a potential gradient along the z -axis of about 400 V cm^{-1} . The drift speed of $27\text{ cm }\mu\text{s}^{-1}$ results in a maximum drift time of $100\text{ }\mu\text{s}$. The two end plates are made up of 18 trapezoidal sectors each, where multi-wire proportional chambers (MWPCs) are mounted. They are separated into an inner and outer readout chamber (IROC and OROC). The MWPCs consist of three layers of wire and a readout pad. The first layer is the gating grid, which can be opened or closed by alternating its current. In its closed state, it prevents ions from escaping back into the drift volume, which can cause a distortion in the electrical field. The closed gate also stops the electrons

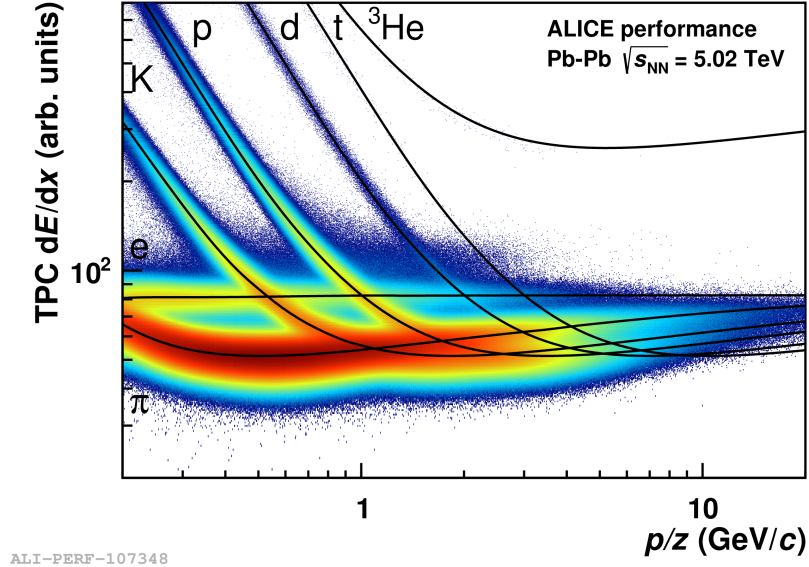


Figure 3.5: Specific ionising energy loss dE/dx of tracks measured in the TPC as a function of momentum in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [20]. The black lines represent the Bethe-Bloch parametrisations for various charged particles. Figure taken from the ALICE figure repository.

from drifting into the amplification region. It is opened in intervals matching the maximum drift time of approximately $100 \mu\text{s}$. The next two wire layers are used to amplify the signal into an electron avalanche, which is then read out by the cathode pads. There are 159 rows of cathode pads, for a total of 557568 pads [102]. These vary in size, from $4 \times 7.5 \text{ mm}^2$ in the IROC and $10 - 15 \times 6 \text{ mm}^2$ in the OROC.

The TPC can detect tracks of charged particles, as well as identify them. The particle identification is done by utilising the ionisation energy loss per unit length, dE/dx . This is described by the Bethe-Bloch formula. In the ALICE collaboration, a parametrisation of this formula is used that was first proposed by the ALEPH collaboration [103]:

$$f(\beta\gamma) = \frac{P_1}{\beta^{P_4}} \left(P_2 - \beta^{P_4} - \ln \left(P_3 + \frac{1}{(\beta\gamma)^{P_5}} \right) \right) \quad (3.1)$$

Here, β and γ are the relativistic velocity and the Lorentz factor, and $P_1 - P_5$ are fit parameters. Figure 3.5 shows the specific energy loss of tracks in the TPC measured in Pb–Pb collisions as $\sqrt{s_{\text{NN}}} = 2.76$ TeV against the momentum (on a logarithmic scale). The contributions of the individual particles are labelled and clearly visible, allowing to distinguish between different particle species. Depending on the region however, particles are not always clearly identifiable, and often more than one particle species hypothesis has to be considered.

3.2.3 The Time of Flight detector

The TOF detector is an array of Multi-gap Resistive Plate Chambers (MRPCs) built on a cylindrical shell with an inner radius of 370 cm and outer radius of 399 cm. It covers the full azimuthal range and a pseudorapidity range of $-0.9 < \eta < 0.9$, and it has an active surface area of about 141 m². Its main task is particle identification at intermediate momentum range, which is done by measuring the time it takes charged particles to reach the TOF from the interaction point. The detector is divided into 18 sectors along the azimuthal directions, which are each divided into 5 modules along the z -direction. Each of these modules contains between 15 and 19 MRPCs, for a total of 1593 MRPC strips. Each strip consists of two cathodes and a central anode, with the space between them filled by a stack of fine glass plates separated by a 250 μm gap. A strong, uniform electric field is applied across the whole volume. Each strip has an active area of 120×7.4 cm², and is filled with a gas mixture of 90% C₂H₂F₄, 5% i-C₄H₁₀, and 5% SF₆. A charged particle hitting the detector volume ionises the gas, and the electric field ensures the electrons are accelerated enough to create an electron avalanche. The glass plates stop the avalanches, which induces a signal on pick-up electrodes.

Particle identification with the TOF is done by measuring a particle's time of flight along a known length, and using this to calculate the particle mass with the following equation:

$$m = \frac{p}{c} \sqrt{\frac{c^2 t^2}{L^2} - 1} \quad (3.2)$$

Here, p is the momentum of the particle, t is the measured flight time, and L is the trajectory length. The difference in flight time between two particles, t_1 and t_2 , of the same momentum and trajectory length but different masses, m_1 and m_2 , is given by

$$t_1 - t_2 = \frac{L}{2c} \left(\frac{m_1^2 c^2 - m_2^2 c^2}{p^2} - 1 \right) \quad (3.3)$$

The PID capability to differentiate between two particles, expressed in the number of standard deviations $n\sigma$, is limited by the time resolution δt :

$$n\sigma = \frac{t_1 - t_2}{\delta t} \quad (3.4)$$

The TOF has an intrinsic time resolution of 80 ps. In order to measure the time of flight, a start time is needed. This is given by the T0 detector, which has a time resolution of ~ 50 ps for pp collisions and ~ 25 ps for Pb-Pb collisions [104]. The efficiency of the T0 detector is 50% for pp collisions and 100% for central Pb-Pb collisions. In cases where the T0 signal is absent, the start time can be estimated based on particle arrival times in the TOF. The total time resolution in cases where the T0 signal is used is given by

$$\sigma_{tot}^2 = \sigma_{TOF}^2 + \sigma_{T0}^2 \quad (3.5)$$

Figure 3.6 shows the particle velocity measured by the TOF plotted against momentum. It provides good particle identification information for momenta of up to 2.5 GeV/c for pions and kaons, and 4 GeV/c for protons [19]. Combined with the TPC, which offers particle identification at high as well as low momenta, particle species can be separated over the whole momentum range used in this analysis.

3.2.4 The V0 detector

The V0 detector is used for triggering and to determine event multiplicity. It consists of two modules of scintillator arrays placed along the beamline at opposite ends of the detector. Both cover the full azimuthal range and a pseudorapidity of $2.8 < \eta < 5.1$

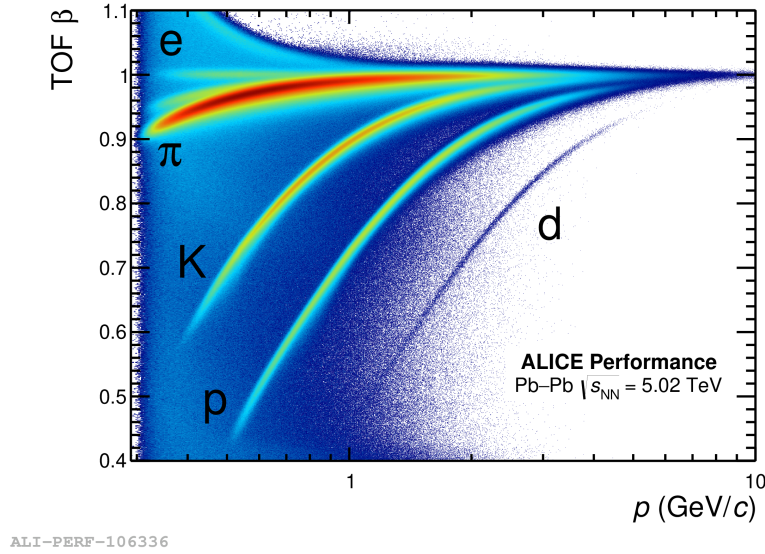


Figure 3.6: Relativistic particle velocity β vs momentum measured by the TOF detector in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV (Run 2) [21]. The signatures of the individual particles are labelled and clearly distinguishable. Figure taken from the ALICE figure repository.

(V0A) and $-3.7 < \eta < -1.7$ (V0C), respectively. They serve as a minimum bias (MB) trigger for the central barrel detectors, requiring a signal in both V0A and V0C for Pb–Pb collisions.

The V0 detector is also used to estimate the centrality of the collision. As described in section 1.3.3, the centrality cannot be measured directly, but needs to be inferred from the average charged-particle multiplicity N_{ch} . For this purpose, a Glauber Monte Carlo model is used, together with a model for particle production that is based on a negative binomial distribution (NBD). From this, the V0 amplitude distribution can be simulated, i.e. the experimental multiplicity distribution. Figure 3.7 shows this distribution fitted to experimental data, with the centrality regions corresponding to each amplitude shown underneath the curve. From this model, the V0 amplitude measured for an event can be used to estimate the centrality of the collision.

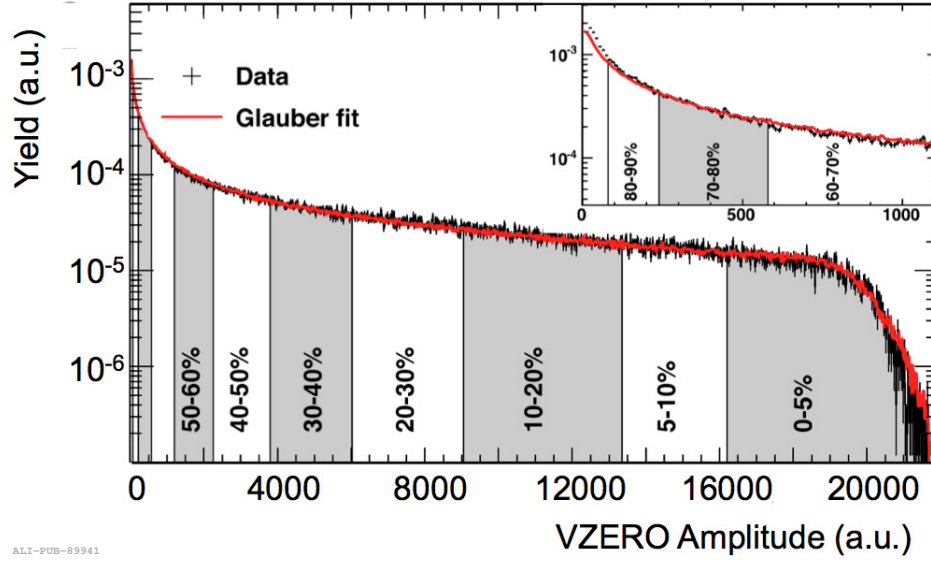


Figure 3.7: The distribution of the sum of amplitudes in the V0 detectors (VOA and VOC) in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ GeV (black), with a fit to the Glauber model (red). The different centrality classes are shown in the shaded areas. [22]

3.3 Trigger system

During Run 1 and 2, the trigger decision in ALICE is made by the Central Trigger Processor (CTP) according to signals from the individual detectors as well as the LHC bunch filling scheme. It is evaluated every clock cycle (25 ns). The trigger system is based on three hardware-based stages: Level0 (L0), Level1 (L1), and Level2 (L2). This is done to accommodate the different readout times of the individual detectors. The L0 trigger decision happens 0.9 μ s after a collision, and includes information from the V0, T0, EMCal, PHOS, and muon trigger detectors. Events triggered in this level are passed on to the next stage. The L1 trigger decision happens 6.5 μ s after the collision, which is defined by the latency of the computation required by TRD and EMCal detectors, as well as the signal propagation time from the ZDC detectors, which are 113 m away from the interaction point. The last trigger stage happens 100 μ s after the collision and accounts for the drift time of the TPC. If an event passes the L2 trigger, the data is sent to the Data Acquisition System

(DAQ) and the High Level Trigger (HLT), where event building and online data compression are performed.

3.4 Track and vertex reconstruction

The track and vertex reconstructions are done offline, using the information from the central barrel detectors. In a first step, the raw data is converted into clusters using information such as position, signal strength, and signal times. Each cluster corresponds to a single particle producing hits as it interacts with a detector element. Clusters in the two innermost layers of the ITS (the SPD) are used to produce track fragments called tracklets. These are used to reconstruct the primary collision vertex, which is done by determining the position where the largest number of tracklets converge. The full tracks are constructed using a three stage approach, more specifically an inward-outward-inward propagation system.

This scheme starts at the outer radius of the TPC. Two clusters from the outer two pad rows are used to build track “seeds”, first by combining them with the vertex position and then by using three clusters without the vertex position. Then, the tracks are propagated inward by combining them with the next cluster along the trajectory that fulfils a proximity cut, and updating the track accordingly. This continues until the inner edge of the TPC is reached. In cases where two or more tracks share clusters, those with a worse reconstruction are rejected by imposing a maximum fraction of shared clusters. This is done to avoid fitting the same track twice. Tracks that have fewer than 20 out of the possible 159 maximum clusters in the TPC are also rejected. After a TPC track has been accepted, a preliminary particle identification is done using the ionisation energy loss, and this information is used in the next steps. The TPC tracks are then propagated to the outer edge of the ITS and used as track seeds for the ITS tracks. In the ITS the procedure from the TPC track reconstruction is repeated, however, from layer to layer each cluster is used to propagate the track, provided it passes a proximity cut. This means that each track seed produces a “tree” of tracks, out of which the one with the highest reconstruction quality is saved. The tracks are extrapolated to the interaction vertex.

In the next stage, the tracks are re-fitted outwards using the same clusters as previously. When they reach the outer edge of the TPC, they are propagated to the TRD and TOF, where time-of-flight information is assigned. However, the propagation of the tracks is not used to update track information, but only used for PID purposes. In the third and final stage, the tracks are once again fitted inwards, using all available cluster information.

3.5 Particle identification

There are several detectors in ALICE providing PID information, and they can either be used by themselves or combined to provide complementary information. The standard way of doing PID is the $n\sigma$ method. Here, the selection is made on the difference between the expected signal of a particle $\hat{S}$ and the measured detector signal S . $n\sigma$ is given by:

$$n\sigma = \frac{S - \hat{S}(H_i)}{\sigma_i} \quad (3.6)$$

where H_i is a specific particle hypothesis and σ_i is the detector resolution. Combining the information from the TPC and TOF detectors makes it possible to separate particle species. However there is still some overlap, and rectangular cuts are not necessarily optimal. An alternative PID method based on Bayesian probabilities has been developed, as described in [105]. The idea is to use probabilities, which has the added advantage of including detectors with non-Gaussian responses easily. A conditional probability can be defined that a particle of species H_i will produce a signal S in a given detector α . For a Gaussian detector response this can be expressed as:

$$P(S_\alpha|H_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}n\sigma_i^2} = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(S_\alpha - \hat{S}_\alpha(H_i))^2}{2\sigma_i^2}}. \quad (3.7)$$

For detectors with a non-Gaussian response, an alternative parametrisation can be used. The advantage of this is that the probabilities from different detectors can be combined:

$$P(\vec{S}|H_i) = \prod_{\alpha=TPC,TOF,\dots} P_\alpha(S_\alpha|H_i) \quad (3.8)$$

Here, $P(\vec{S}|H_i)$ is the conditional probability that a set of detector signals $\vec{S}$ is seen for a given particle species H_i . To find the probability of a particle being of species H_i when a set of signals $\vec{S}$ is observed, Bayes theorem [106] is used:

$$P(H_i|\vec{S}) = \frac{P(\vec{S}|H_i)C(H_i)}{\sum_{k=e,\mu,\pi,\dots} P(\vec{S}|H_k)C(H_k)} \quad (3.9)$$

Here, $C(H_i)$ is an a priori probability of measuring the particle species H_i , also called a “prior”, while the conditional probability $P(H_i|\vec{S})$ is called a posterior probability.

The priors are an estimation of the true particle yield. They are calculated as a function of momentum and dependent on the centrality class. This is done with an iterative procedure on data taken from the respective collision system. The priors are arbitrarily normalised to the abundance of pions, and they are set to one in the first step. The Bayesian probability $P_n(H_i|\vec{S})$ is then calculated using these flat priors, and used as weights for the p_T spectra for the $n+1$ step. This is repeated until the priors converge. They are found to be consistent with true particle abundances to within roughly 10%.

For a given track, the Bayesian probability is calculated for each particle species, and a selection can be made on it. This can be done by assigning the identity of the particle species with the highest probability to the track, or by accepting each particle species that has a probability higher than a certain threshold. The latter can lead to a track being assigned more than one particle species. The aim is to maximise the purity, i.e. the ratio between the number of correctly identified particles and the total selected. Purity is not the only concern however, with another consideration being the efficiency, which is the ratio between the number of particles assigned a certain species and the total number of particles of that species. That means that the more cuts are applied to filter out background, the more signal events are also lost. A balance is therefore required between having a good particle identification and keeping enough events to carry out a useful analysis.

Further considerations on the question of how to select and cut on data with respect to not only achieving a good signal-to-background ratio, but also keeping

the size of the dataset to a manageable level, will be given in the following chapter, where the analysis procedure is described.

Chapter 4

Data Analysis

In this chapter, the measurement of the p_T -differential production yield of promptly produced Λ_c^+ baryons is described. It is performed in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel in Pb–Pb collisions at a centre-of-mass energy per nucleon-nucleon collision of $\sqrt{s_{NN}} = 5.02$ TeV. Section 4.1 gives a brief overview of the analysis and the decay channel investigated. The initial event selection and the Monte Carlo (MC) simulation used are described in section 4.2. Section 4.3 details how candidate reconstruction and track selection is applied to the data, followed by a loose selection criteria on the decay topology and a Bayesian probability approach to the initial particle identification cut. Then, a machine learning model is trained on data and MC to optimally select the Λ_c^+ signal (4.4). Corrections for the Λ_c^+ reconstruction efficiency and feed-down from beauty decays are then applied to the extracted raw yield (4.5). Lastly, the systematic uncertainties on the results of this analysis are calculated (4.6).

4.1 Overview

The Λ_c^+ baryon has a quark content of udc and is the lowest mass charmed baryon with a mass of 2286.46 ± 0.14 MeV/c [2]. Due to its short lifetime of $201.5 \times 10^{-15} \pm 2.7$ s ($c\tau = 60.4$ μ m), it decays before it reaches the ALICE detector, and is measured

Decay mode	Branching ratio (Γ_i/Γ)	p(MeV/c)
$pK\pi$ non-resonant	$(3.5 \pm 0.4)\%$	823
$p\bar{K}^*(892)^0$	$(1.96 \pm 0.27)\%$	685
$\Delta(1232)^{++}K^-$	$(1.08 \pm 0.35)\%$	710
$\Lambda(1520)\pi^+$	$(2.2 \pm 0.5)\%$	628

Table 4.1: The different decay modes of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel and their respective branching ratios and energies [2].

via a reconstruction of its decay products. This analysis uses the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel, which has a branching ratio of $6.35 \pm 0.33\%$ [2]. The decay includes three resonant decay modes as well as one non-resonant one, which are shown in table 4.1.

The analysis was performed selecting two different centrality classes: central (0-10% centrality) and semi-central (30-50% centrality) events. The Λ_c^+ production yield was successfully extracted in the p_T regions $6 < p_T < 8$ GeV/ c (semi-central events only), $8 < p_T < 12$ GeV/ c , and $12 < p_T < 24$ GeV/ c . Initially, lower momentum regions ($6 < p_T < 8$ GeV/ c for central and $4 < p_T < 6$ GeV/ c for semi-central events) were also investigated but since this analysis is at the limit of what is possible with this data and the tracking and vertexing performances of the ALICE detector in Run 2, signal extraction was ultimately unsuccessful.

4.2 Event Selection

The data used in this analysis was taken by the ALICE detector in the 2018 Pb–Pb run with a collision energy of $\sqrt{s_{NN}} = 5.02$ TeV. A minimum-bias trigger is used to select events that have at least one hit in each of the two arrays of the V0 detector (V0A and V0C). Two additional online trigger selections were added based on the V0-signal amplitude, with the aim of increasing the number of central and semi-central events. In addition to that, the primary vertex reconstructed from the full tracks was required to be no more than 10 cm from the centre of the detector along the z plane, and to be within 5 cm of the one reconstructed using only the SPD detector. The centrality information needed to perform the analysis in two distinct

centrality bins is based on the response of the V0M detector, as described in section 3.2.4.

4.2.1 Monte Carlo simulation

In this analysis, two Monte Carlo simulation datasets are used, one for training the machine learning model, and a second one to calculate reconstruction efficiencies. The first simulation is enriched with $\Lambda_c^+ \rightarrow pK^-\pi^+$ decays in order to gain enough signal events for training. This is done by simulating Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ GeV using the HIJING 1.36 event generator [107]. Then, the PYTHIA 8 generator [108] with Monash 2013 tune [109] was used to inject $c\bar{c}$ and $b\bar{b}$ pairs. The Λ_c^+ baryons created are forced to decay via the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel to simulate enough signal events for training the machine learning model and correcting the Λ_c^+ yield for the reconstruction efficiencies. To simulate the resulting particles' traversal through the ALICE detector, the GEANT3 [110] package is used, reproducing the detector's conditions at the time of data taking. This includes alignment, active channels, noise, and gain, which are tuned to match the data-taking conditions for each individual data-taking run.

4.3 Candidate selection

4.3.1 Candidate reconstruction and track selection

The Λ_c^+ candidates are reconstructed from the decay products of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel, i.e. a proton, a negatively charged kaon, and a positively charged pion. The charge conjugate (an anti- Λ_c^+ decaying into an antiproton, a negatively charged kaon, and a positively charged pion) is also reconstructed, meaning that both Λ_c^+ and Λ_c^- baryons are used. This is corrected for later by adding a factor of $\frac{1}{2}$ in the calculation of the yield.

To create the Λ_c^+ decay candidates, three tracks are combined, two positive ones and a negative one for the Λ_c^+ baryon, and two negative ones and a positive one for

the Λ_c^- baryon. From these tracks, a secondary vertex is reconstructed to define the point where the Λ_c^+ candidate decayed. This is done by first taking two oppositely charged tracks and finding the point of minimum distances between their reconstructed trajectory, and then by taking all three tracks and finding the minimum of the total distances.

Pre-selection track quality cuts are made on the individual tracks to select those used for candidate reconstruction. These include several requirements. A track needs to have a minimum number of clusters in the ITS and TPC detectors, pass the refit stage, and be within a pseudo-rapidity range that ensures minimal degradation of the TPC response. The exact selection parameters are listed in table 4.2, and identical to those in a previous analysis of a similar decay.

4.3.2 Particle identification

In addition to the track quality selection, a pre-selection on the particle identification is also made to ensure that they are the correct species. For this, two different PID strategies were explored, which are also described in section 3.5:

- **$n\sigma$ approach:** For a given track, its detector signal is compared to the expected signal for a specific particle species. The difference is expressed in multiples of the detector resolution, and a selection can then be applied to only allow tracks where this difference is within a certain threshold, usually around 3σ . This approach means that it is possible for a track to be selected as more than one particle species.
- **Bayesian probability:** A Bayesian probability is calculated for each track for the different particle species (see section 3.5). For a track to be considered as a certain particle, a minimum probability can then be imposed. Alternatively, each track can also be assigned only the particle species for which it has the highest probability. This approach is described in detail in [105].

Originally, the $n\sigma$ method was used for PID. However, the large data volume and subsequent file sizes proved a significant hurdle to the analysis procedure, which

many of the computational steps taking in the order of days to run. In order to mitigate this a Bayesian PID strategy was trialled, which was previously used in an analysis of the same decay in pp collisions, where it was found to significantly reduce background. When applying this and running the analysis again, it was found that the background could be greatly reduced using the Bayesian method, to about 20% the size of what it was for the previous method. This allowed for a much more manageable dataset size in terms of CPU time needed for the analysis, while keeping the loss of signal events to an acceptable limit. Comparing the signal that was extracted with each method showed that they were consistent, and so ultimately the Bayesian approach was chosen to accommodate the computational limitations.

4.3.3 Λ_c^+ candidate selection

In order to achieve optimal signal extraction, cuts are applied on various topological and kinematic variables. In this analysis, triplets are constructed of three particle tracks and the decay is reconstructed as shown in figure 4.1. The displaced (or secondary) vertex of this decay makes it possible to exploit additional properties to separate true Λ_c^+ decays from background. This background includes any protons, kaons, and pions produced either directly in the hadronisation stage, or stemming from decays of other particles. The variables used for signal and background separation are defined as follows:

- **Distance of closest approach (DCA):** The distance of closest approach measures how closely each particle track passes by the reconstructed secondary decay vertex. A selection is made on the maximum from between the three values is used.
- **Quality of the reconstructed secondary vertex:** The quality of the reconstructed secondary vertex is defined as $\sigma_{\text{vert}} = \sqrt{\sigma_0^2 + \sigma_1^2 + \sigma_2^2}$, where σ_i are the distances between the vertex and the individual tracks.
- **Decay length:** The decay length measures the distance between the primary (PV) and secondary (SV) vertices, i.e. the distance the Λ_c^+ baryon travelled

before decaying. The primary vertex is taken as the centre of the collision, and the secondary vertex is reconstructed from the three particles' trajectories.

- **Decay length using track pairs:** An alternative way to determine the decay length is the distance between the primary vertex and the point of closest approach between two of the three tracks. This is done for oppositely charged tracks, and is referred to as “ dist_{12} ” and “ dist_{23} ”. This additional method is used in order to mitigate the uncertainty on the previous method that comes from the reconstruction of the secondary vertex.
- **Normalised decay length in x-y plane:** Since the resolution in the plane perpendicular to the beam line is better than in the z-direction, an additional calculation of the decay length is done by projecting it into the x-y plane.
- **Cosine of pointing angle:** The pointing angle is defined as the angle between the flight direction of the reconstructed Λ_c^+ candidate, and the continuation of the connecting line between primary and secondary vertices, i.e. the expected flight direction. If the vertices and candidate decay product momenta are reconstructed precisely, a value close to 1 is expected.
- **impact parameter:** This is the minimum distance of a track to the primary vertex, which is given for each track individually. As the decay is secondary, it is not expected to be 0.
- **max norm d0d0exp:** The maximum difference between the impact parameter measured and that calculated.

4.3.4 Initial kinematic and topological selection

In addition to the track selection, cuts prior to the multivariate selection are also made on the kinematic and topological variables described in section 4.3.3. These so-called pre-selection cuts are defined by studying the signal and background distributions of Λ_c^+ decays in Monte Carlo simulations, where signal and background

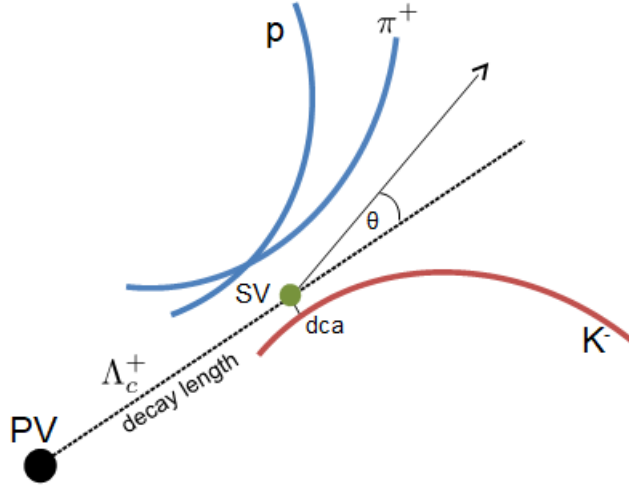


Figure 4.1: A schematic of the reconstructed $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay, with some reconstruction variables labelled. PV and SV denote the primary and secondary decay vertices, respectively, DCA is the distance of closest approach, and θ is the pointing angle (defined in the text).

Λ_c^+ candidates are known. An example of these distributions for some topological variables is shown in figure 4.2, with the proposed cut value shown ¹. A decision of where to apply a cut is made based on where the best separation between signal and background can be made while still reducing the size of the dataset to a manageable level. Due to this constraint, the cuts vary from looser ones for the semi-central events, to tighter ones for the lower-momentum central ones where there are significantly more Λ_c^+ candidates. The chosen values for all these pre-selection cuts are shown in table 4.2, which also includes the pre-selection criteria on the tracks that were described in section 4.3.1.

¹Note that these figures are from an early stage of the analysis, and the cut values are outdated. They are merely shown as an example of how the preselection cuts are motivated.

centrality	0-10%		30-50%
$p_T(\text{GeV}/c)$	0 – 8	> 8	
σ_{vert}	< 0.3	< 0.045	< 0.06
$\cos\Theta_{\text{pointing}}$	> 0.95	> 0.25	> 0.
dca cut	< 0.03	< 0.0375	< 0.05
dist prim-sec (decay length)	> 0.00625	> 0.00625	> 0.
dist ₁₂	> 0.03	> 0.125	> 0.
PID	Bayesian maximum probability		
min nTPC clusters	50		
TPCrefit	true		
ITSrefit	true		
$ \eta $	< 0.8		
p_T	> 0.6 GeV/ c		

Table 4.2: The pre-selection cuts applied on reconstructed tracks (bottom rows) and Λ_c^+ candidates (top rows).

4.4 Multivariate Selection

The main challenge of this analysis is the large combinatorial background, and the need to separate the signal events from this. Due to the complicated nature of the decay having three final state particles, filtering out the background presents a significant difficulty.

For the pre-selection cuts, the decision on where to cut was made by plotting the signal and background distributions and manually deciding on an optimal cut value. Since there are a lot of different variables to be cut on, testing and choosing cut values for all of them for signal extraction is somewhat impractical. Another downside to these so-called rectangular cut method is the possibility of some variables being correlated with others, further complicating this approach. Instead, a machine learning model is used to optimally select Λ_c^+ signal candidates and reject Λ_c^+ background candidates. This is done by first training an algorithm to optimise this selection on a dataset with known classification - i.e. a Monte Carlo simulation dataset where signal and background events are known - before applying it to the dataset with unknown classification. The specific type of algorithm used are Boosted Decision Trees

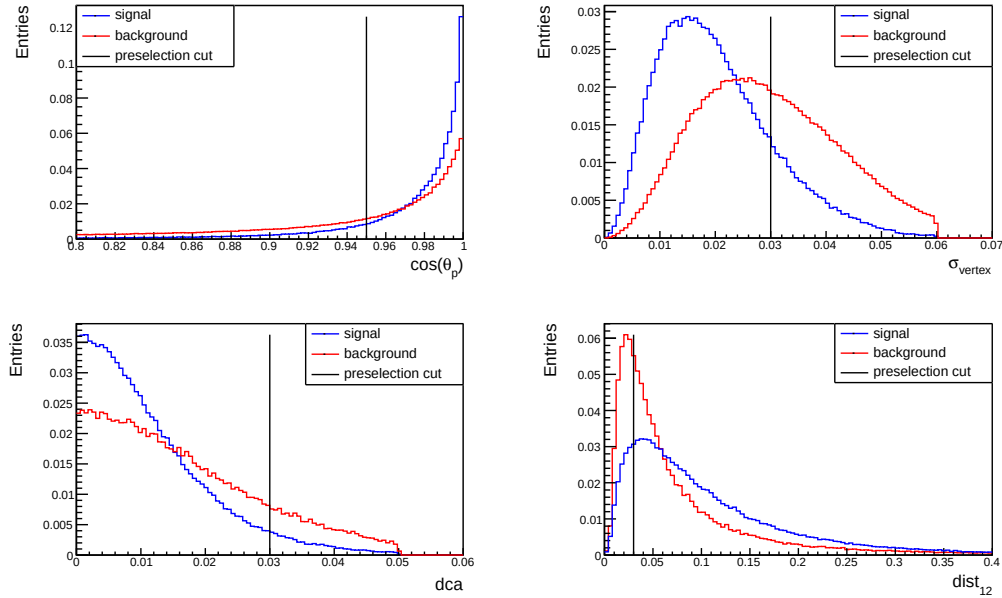


Figure 4.2: The signal and background distribution of four variables ($\cos(\theta_p)$, σ_{vertex} , distance of closest approach (dca), and dist_{12}), with the chosen pre-selection cut value marked in black.

(BDT), which are implemented in the MLHEP software package [111] developed for the ALICE collaboration. This package operates with Pandas Dataframes and the training and testing are done using the xGBoost algorithm [112].

4.4.1 Boosted Decision Trees

Decision trees work by splitting up the data in a series of nodes. At each node, it checks which variable and cut position yields the best separation between signal and background events. The data is then split into the events that pass the cut and those that do not. In both directions, this is then followed by another node, applying another cut to split the sets. This continues until certain criteria, for example a minimum number of steps, are met. As a result, the decision space is split up into many different branches that are each given a score between 0 and 1, denoting how background- or signal-like these events are. Each candidate is propagated along

the nodes depending on which of the cuts it passes and fails, until it reaches the end of the branch. Since each branch has been classified during the training, the candidate can then be assigned a BDT score on which a selection can be made, which only candidates with a value above a certain threshold being kept. A schematic representation of a Decision tree is shown in figure 4.3.

Decision trees are prone to overfitting, which happens when a model is very good at predictions for the specific dataset it was trained on, but performs poorly on others. To increase the stability of the model, several decision trees can be combined into a “forest”. Small decision trees, called “weak learners”, are grown with a maximum step size of somewhere around 2-5. Boosting is an iterative combination, where each subsequent tree “learns” from the mistakes of the previous one. This is done by giving the events that were misclassified at the previous step a greater relative weight in the next tree. Further trees are grown until a certain number is reached. The response of the model is then based on a combination of all of the different trees. To further increase stability, for each tree a random subset of the training data is selected, as well as a random subset of the training variables that will be used to grow that tree.

The specific algorithm used in this analysis is called XGBoost [112], which is an open-source package for tree boosting. It was originally released in 2014 and quickly became popular in data science applications due to its performance and speed. There are a variety of tuneable parameters in an XGBoost model:

- **max depth:** The maximum depth of each individual tree, after which it is terminated. If a point is reached before then where no minimum loss reduction² is seen between steps, no further steps are taken in that node. The tree depth is kept small in order to prevent overfitting.
- **learning rate:** After each boosting step, the weights of new variables are shrunk by this value to make the boosting process more conservative.

²A loss function is defined for the model in order to quantify how well it performs. In this case it tracks how well the model distinguishes between signal and data, and a lower value means better performance.

XGBoost parameters	
max-depth	3
learning-rate	0.1
number of trees	850
gamma	0
min-child-weight	3
subsample	0.8
colsample-bytree	0.8

Table 4.3: The hyperparameters of the XGBoost model used in this analysis.

- **number of trees:** The number of individual trees used in the model which is usually of the order of 100 or 1000.
- **gamma:** The minimum loss reduction required for a new step to be added to a tree. The higher this number is, the more conservative the model is.
- **min child weight:** The minimum sum of instance weight needed in a child. If a node has a sum less than this, no further steps are taken.
- **subsample:** The fraction of the training instances that is randomly chosen at the start of each boosting interaction.
- **colsample bytree:** The fraction of variables that are randomly chosen to train each tree.

The full list of parameters chosen is shown in table 4.3. They were taken from a previous ALICE analysis into the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel in Pb–Pb collisions, which uses the same Machine Learning framework as this work [23].

4.4.2 BDT training

The model is trained on both Monte Carlo simulation data and experimental data. The signal Λ_c^+ candidates used come from simulated Λ_c^+ decays. As the Monte Carlo data set is not sufficiently large, the background Λ_c^+ candidates used are candidates from experimental data in the “sideband” region of the invariant mass distribution.

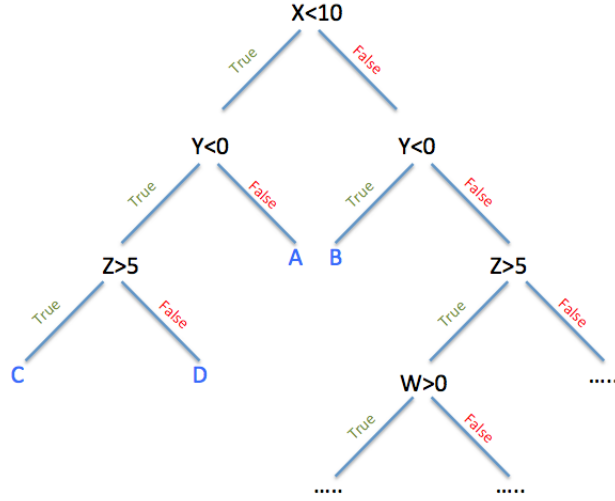


Figure 4.3: An example for a decision tree. At every node, the variable (X,Y,Z,W) and cut value (0, 5, 10) with the best signal to background separation are chosen to divide the dataset (A,B,C, D, ...).

$p_T(\text{GeV}/c)$	# of training events (signal)	# of training events (background)
6-8	9950	499985
8-12	21385	500255
12-24	23475	195415

Table 4.4: Number of signal and background training events used in central collisions.

This means that they are from a region in the dataset where there are no signal events, in this case $m_{\text{cand}} < 2.23 \text{ GeV}/c^2$ and $m_{\text{cand}} > 2.33 \text{ GeV}/c^2$. They are split into a training and a testing set. The testing set provides a statistically independent dataset to test the model and ensures it does not interpret statistical fluctuations in the training set as features in the data (so-called “overfitting”). The number of events used for testing and training is shown in tables 4.4 and 4.5. Individual models are trained for each p_T bin of this analysis. Note that while the ideal number of events used for the training to converge well were found in the region of 500000, the Monte Carlo production data was not sufficiently large for this training set size,

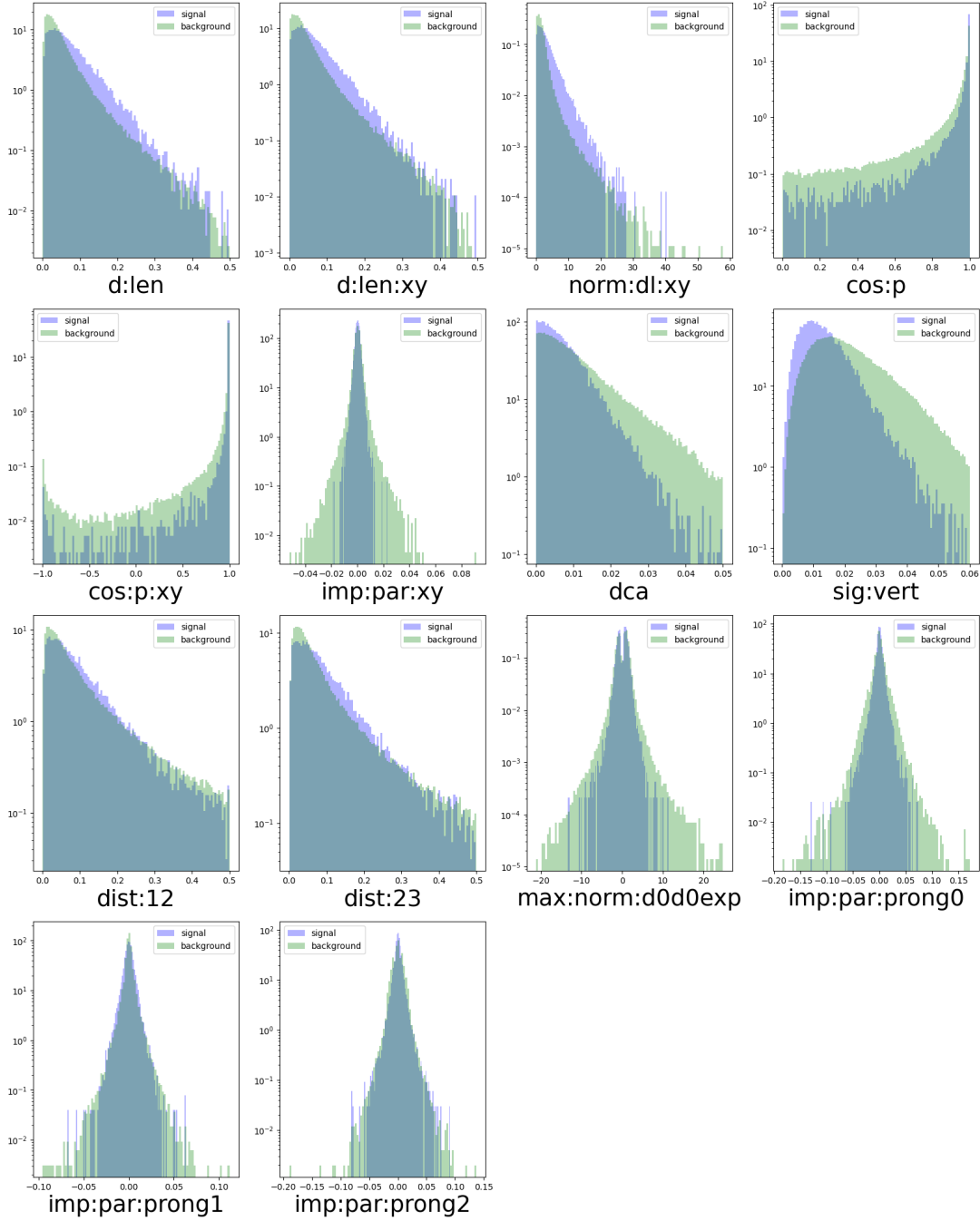


Figure 4.4: The signal and background distribution of the variables used for the training, shown in the $12 < p_T < 24 \text{ GeV}/c$ bin for central collisions. The individual variables, from left to right and top to bottom, are the decay length, the decay length in the x-y plane, the normalised decay length in the x-y plane, the cosine of the pointing angle, the cosine of the pointing angle in the x-y plane, the impact parameter in the x-y plane, the distance of closest approach, the σ of the vertex reconstruction, the distance between the 1st and 2nd tracks, the distance between the 2nd and 3rd tracks, the maximum difference between measured and calculated impact parameter, and the impact parameter for the three tracks (prong 0,1, and 2).

$p_T(\text{GeV}/c)$	# of training events (signal)	# of training events (background)
4-6	37750	500540
6-8	25780	495485
8 -12	32990	500855
12-24	33205	195170

Table 4.5: The number of signal and background training events used in semi-central collisions.

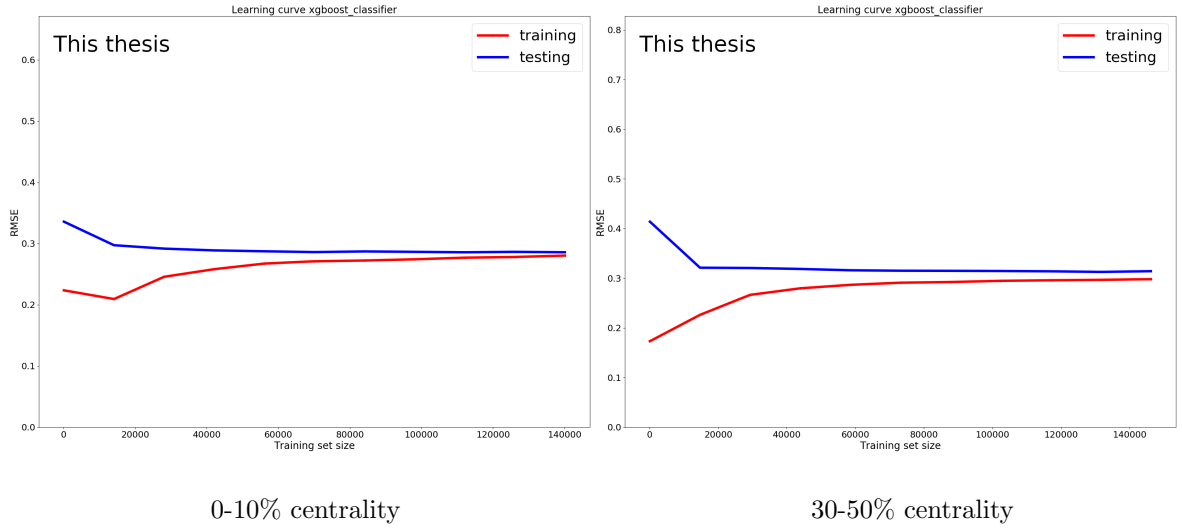


Figure 4.5: The learning curve of the BDT model, shown in the $12 < p_T < 24 \text{ GeV}/c$ bin for both centrality classes, which shows a good convergence between training and testing datasets.

which is why the numbers shown are lower. Similarly in the high momentum bins the number of background events also fell below this threshold.

The BDT uses input variables relating to the decay reconstruction. The particle identification method is omitted here, as it is optimised using the Bayesian probability method. Figure 4.4 shows the distribution of signal and background events for each training variable in for $12 < p_T < 24 \text{ GeV}/c$ in central collisions. Separation between signal and background is seen for each variable.

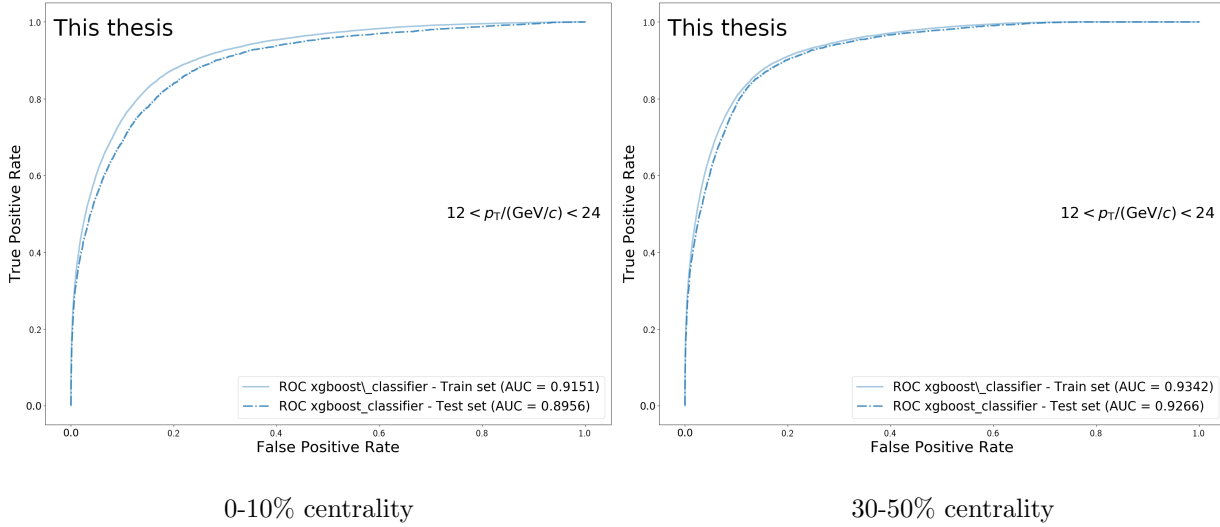


Figure 4.6: The ROC curve of the BDT model, shown in the $12 < p_T < 24 \text{ GeV}/c$ bin for both centrality classes. A comparison is made between the testing and training data sets, and showing a good agreement and a reasonably high AUC value.

A learning curve is drawn to check that the training and testing converge. It plots the RMSE (root mean squared error) for testing and training events as a function of the training set size, and can be used to determine if a larger sample size (no convergence) or more input variables (large RMSE) are needed. The learning curves for this analysis are shown for the $12 < p_T < 24 \text{ GeV}/c$ bin in figure 4.5, and the training and testing curves converge to a reasonable level.

Another indication of model performance is the receiver operating characteristic (ROC) curve, which plots the true positive against the false positive rate. A true positive here means a signal event correctly identified as such, while a false positive is a background event wrongly selected as signal. The ROC quality is normally measured by the area under the curve (AUC), which for an ideal model would be 1, as it would have a 100% true positive rate and a 0% false positive one. For this analysis, it is plotted for both the training and the testing data set, in order to check that the model does not overfit. As shown in figure 4.6 for the $12 < p_T < 24 \text{ GeV}/c$

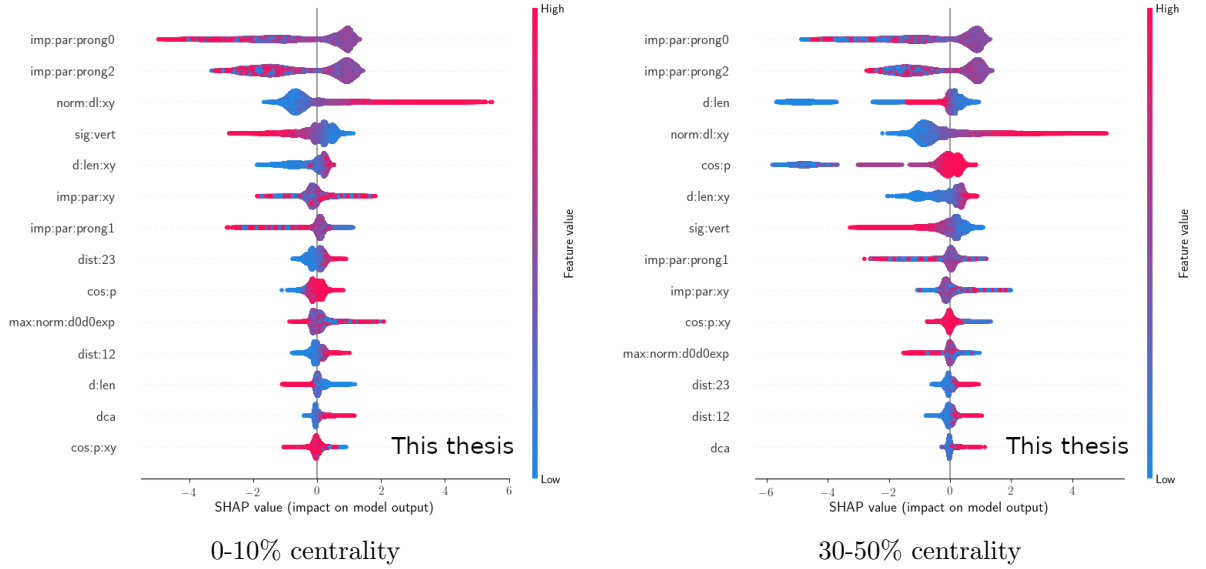


Figure 4.7: The individual impact of each training variable on the model, shown in the $12 < p_T < 24 \text{ GeV}/c$ bin for both centrality classes. The x-axis denotes the impact on the model, where a higher value means a more signal-like impact, and a lower one indicates background-like. The colour shows the variable value of the individual instance, with red showing a high value and blue a low one. They are ordered from most impactful (top) to least impactful (bottom).

bin for both central and semi-central collisions, they are in good agreement at a high AUC value of around 0.9, compared to values between 0.7 and 0.8 for a similar analysis. This shows that the model is rejecting a high amount of background while maintaining good signal efficiency for both training and testing data sets.

Figure 4.7 shows so-called SHAP (SHapley Additive exPlanations) plots, which are used to assess the influence of the individual training variables on the model output. The input variables are ordered by how much each variable contributes to the overall model output, but they also provide additional information in comparison to simple linear plots. Each instance is plotted as a dot, which is coloured depending on the value of the variable. It does not give an absolute value but an approximation of high (red) and low (blue). In addition to that, its position on the x-axis indicates

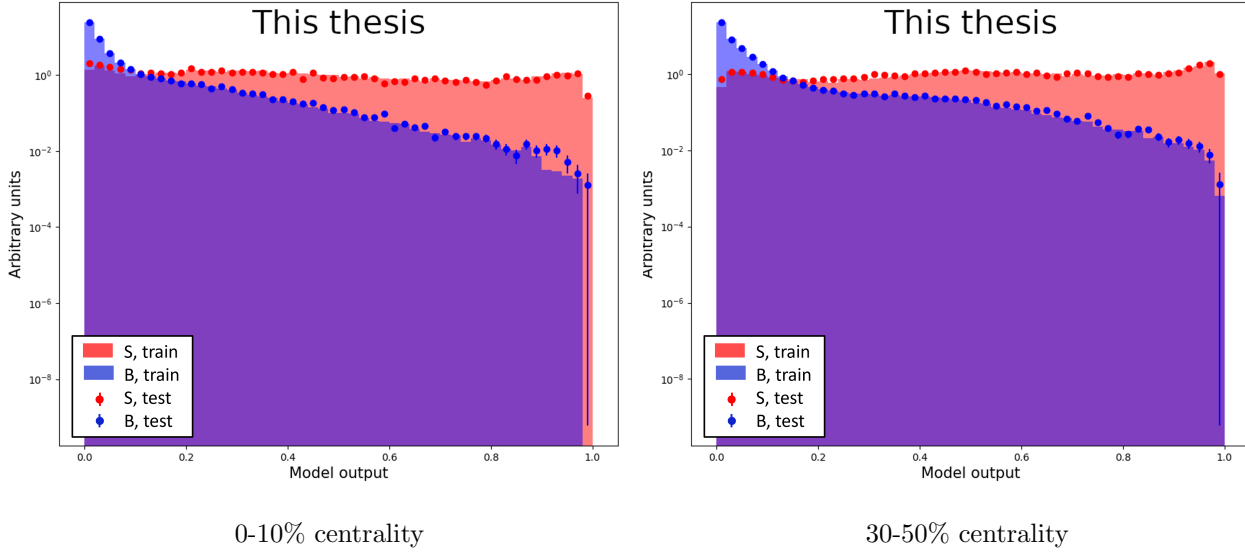


Figure 4.8: The BDT model output, shown in the $12 < p_T < 24 \text{ GeV}/c$ bin for both centrality classes. It shows the separation between signal-like and background-like events. There is good agreement between training and testing datasets.

how much and in which direction it contributes to the model output. These plots give insight into what each variable contributes to the model, as well as check for possible inconsistencies or errors in case the outcome is not as expected. As for example a low σ_{vert} is expected for a good candidate reconstruction, it is expected that a high value (red colour) has a negative impact on the BDT score towards the left of the plot), whereas a low value (blue) is found on the right side. By looking at the shape of each individual variable's cluster its distribution and range can also be estimated. In the case of σ_{vert} both plots show the expected colour distribution, with more events in the centre of the distribution than at the extreme points. From the SHAP plots in this analysis, it can be concluded that in all momentum bins the impact parameter, decay length, σ of the vertex reconstruction, and the pointing angle generally have the highest impact on the BDT model.

Finally, the BDT output is shown in figure 4.8. It shows a good agreement

between training and testing samples. The model output assigns a value to each candidate that represents whether the candidate is more likely to be signal or background. This ranges from 0, for very background-like, to 1, for very signal-like. The distributions look as expected, with the number of background events decreasing for tighter cuts on the model output, while signal events increase.

To use the trained BDT model, a cut then needs to be made on the model output to select only candidates with an output number over a certain value. This cut is chosen by optimising the expected statistical significance of the extracted signal, which minimises the expected statistical uncertainty of the measurement. For a pre-defined range of cut values, the expected number of reconstructed signal and background per event is calculated in order to estimate the expected significance. The expected signal S is calculated using the following equation:

$$S = 2 \cdot \left(\frac{d\sigma}{dp_T} \right)_{\Lambda_c^+ \rightarrow pK_S^0} \cdot \Delta p_T \cdot \text{BR} \cdot (\text{Acc} \times \epsilon) \cdot T_{AA} / f_{\text{prompt}} \quad (4.1)$$

Here, the assumptions are that $\text{BR} = 6.23\%$ and $f_{\text{prompt}} = 0.9$ for the entire p_T range. T_{AA} is the nuclear overlap function, which represents the number of nucleon-nucleon collisions at a given impact parameter. $\left(\frac{d\sigma}{dp_T} \right)_{\Lambda_c^+ \rightarrow pK_S^0}$ is taken from a recent measurement by ALICE of the Λ_c^+ production yield via the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel [23], instead of using a FONLL calculation as was previously done. The background B meanwhile is estimated from a fit to the background, scaled for the number of events used vs. the number of total events. Combining signal and background scaled to the expected number of events gives the estimated significance:

$$\text{Significance} = \frac{S}{\sqrt{S+B}} \quad (4.2)$$

This is then plotted against the BDT cut value, which allows to find the cut value with the highest expected significance. The significance optimisation is shown in figure 4.9 for the $12 < p_T < 24 \text{ GeV}/c$ bin. Table 4.6 lists the BDT cut values that were chosen for the individual momentum bins.

$p_T(\text{GeV}/c)$	4-6	6-8	8-12	12-24
BDT cut (central)	-	0.45	0.5	0.7
BDT cut (semi-central)	0.5	0.5	0.75	0.75

Table 4.6: BDT cut chosen in this analysis for each centrality and momentum bin

For better readability, the BDT performance figures in this section only show the results in one of the p_T bins, namely the $12 < p_T < 24 \text{ GeV}/c$ one. The performance figures for the other bins are shown in appendix A.

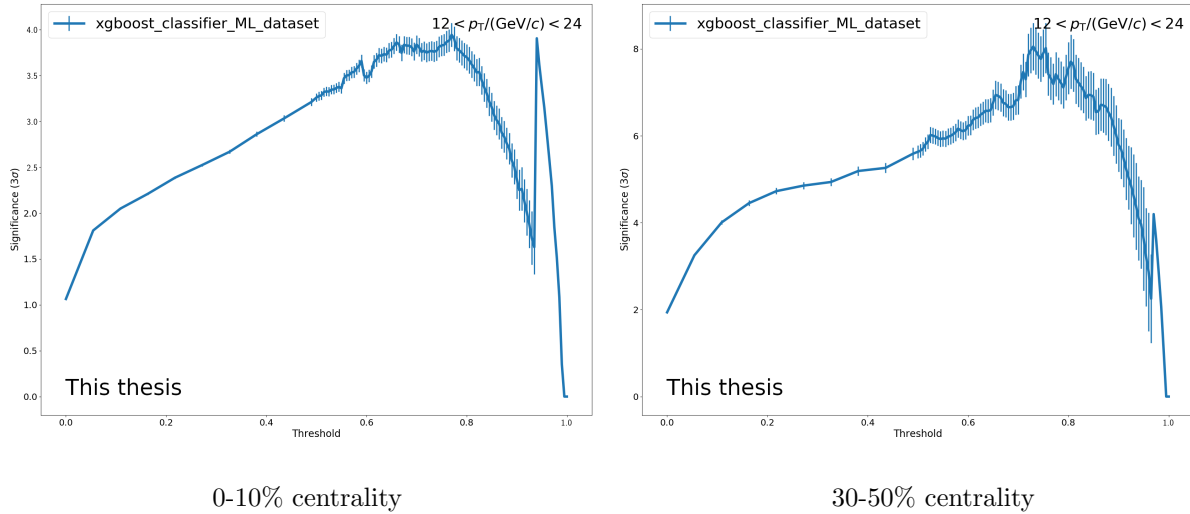


Figure 4.9: The expected significance of the Λ_c^+ invariant mass peak versus the BDT cut in the $12 < p_T < 24 \text{ GeV}/c$ bin for both centrality classes of this analysis. Note that the sharp rise at very high BDT cut value does not represent a real increase in significance but is instead a consequence of the extremely low statistics in this region.

4.4.3 Raw yield

After the BDT cut is applied, the invariant mass distribution of the events that pass it is plotted, and a fit to the distribution is made to extract the Λ_c^+ signal.

The fit function used for the signal peak is a Gaussian, while for the background, a second order polynomial is used. Figures 4.10 and 4.11 show these fits for central and semi-central collisions, respectively. The figures also show the fit parameters as well as the number of signal and background events under the peak. Due to the small signal and large background, to improve stability, the σ value (the width of the fit, corresponding to the mean of the Gaussian fit function) of the data fit is fixed to that from the fit of the Monte Carlo simulation. To ensure that this does not influence the results, the mean and width of the fit are compared to the case where the σ is not fixed, which is shown in figure 4.12. They show that while there is some discrepancy, the mean and width agree within the expected range.

In order to ensure that the signal is not caused by a statistical fluctuation, a significance of at least 3 is required for the signal to be accepted. Since the mass of the Λ_c^+ is known and the expected peak position is therefore also known, this value is lower than the more commonly used 5. This was observed for all momentum bins except for $6 < p_T < 8 \text{ GeV}/c$ in central collisions and $4 < p_T < 6 \text{ GeV}/c$ in semi-central collisions, and it was therefore decided not to include them in the final results. An attempt was also made to extend the analysis to higher momentum, as the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel is expected to perform better in this region than the $\Lambda_c^+ \rightarrow pK_S^0$ channel. For this, a bin of $24 < p_T < 36 \text{ GeV}/c$ was defined. Since the number of background candidates is not enough to train a reliable BDT model in this region, the model from the $12 < p_T < 24 \text{ GeV}/c$ bin was used for training, and applied to the data in the $24 < p_T < 36 \text{ GeV}/c$ region. The resulting invariant mass plot is shown in figure 4.13 for central collisions. Unfortunately, the statistics available are too low, and thus no signal could be extracted in this p_T region.

To gauge the performance of the signal extraction, it is useful to compare the precision and kinematic reach of this analysis to the same analysis performed in the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel by ALICE in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02 \text{ TeV}$ [23]. Figure 4.14 shows two of the fit parameters, the significance and the signal-to-background ratio, for the two different analyses. It shows that the signal-to-background ratio is higher in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel analysed in this thesis work, which indicates that while fewer signal events are found than in the $\Lambda_c^+ \rightarrow pK_S^0$

channel, there are also fewer background events. It can also be seen that while the $\Lambda_c^+ \rightarrow \text{pK}_S^0$ channel generally offers higher significance at low momentum, at higher momentum this analysis shows a clear improvement over it.

4.5 Corrections

The p_T -differential yield of the prompt production of Λ_c^+ baryons is calculated as

$$\left. \frac{dN\Lambda_c^+}{dp_T} \right|_{|y|<0.5} = \frac{f_{\text{prompt}} \times \frac{1}{2} N_{|y|<y_{\text{fid}}}^{\Lambda_c^{\pm \text{raw}}}}{\Delta p_T \times c_{\Delta y} \times (A \times \epsilon)_{\text{prompt}} \times \text{BR} \times N_{\text{ev}}} \quad (4.3)$$

Here, $N_{|y|<y_{\text{fid}}}^{\Lambda_c^{\pm \text{raw}}}$ is the raw yield extracted for a given p_T interval Δp_T . The $\frac{1}{2}$ -factor takes into account that both Λ_c^+ and Λ_c^- particles were measured and that what is reported is only the prompt Λ_c^+ production yield. BR denotes the branching ratio of the decay studied, and N_{ev} is the number of events. The factor $c_{\Delta y}$ is included to normalise the corrected yield to one unit of rapidity. The fraction of the raw yield that comes from promptly produced Λ_c^+ , f_{prompt} , and the detector acceptance and efficiency $(A \times \epsilon)_{\text{prompt}}$ are described below.

4.5.1 Efficiency

The measured raw yield is corrected for the detector acceptance A and the efficiency ϵ of prompt Λ_c^+ . The efficiency is also required to determine the factor f_{prompt} (see below). The efficiencies are calculated using Monte Carlo simulations of Λ_c^+ -enhanced Pb–Pb collisions, where the final state particles are transported through the simulated ALICE detector. This allows for individual efficiencies to be calculated for each step of the track and candidate selections, as well as the BDT selection. The efficiencies are highly dependent on the p_T distribution of the simulated Λ_c^+ candidates. The p_T shape is re-weighted to match that of the TAMU model [27], which has been shown to accurately describe Λ_c^+ production in the p_T region of this analysis, and which is described in section 5.3. The Monte Carlo simulation used in

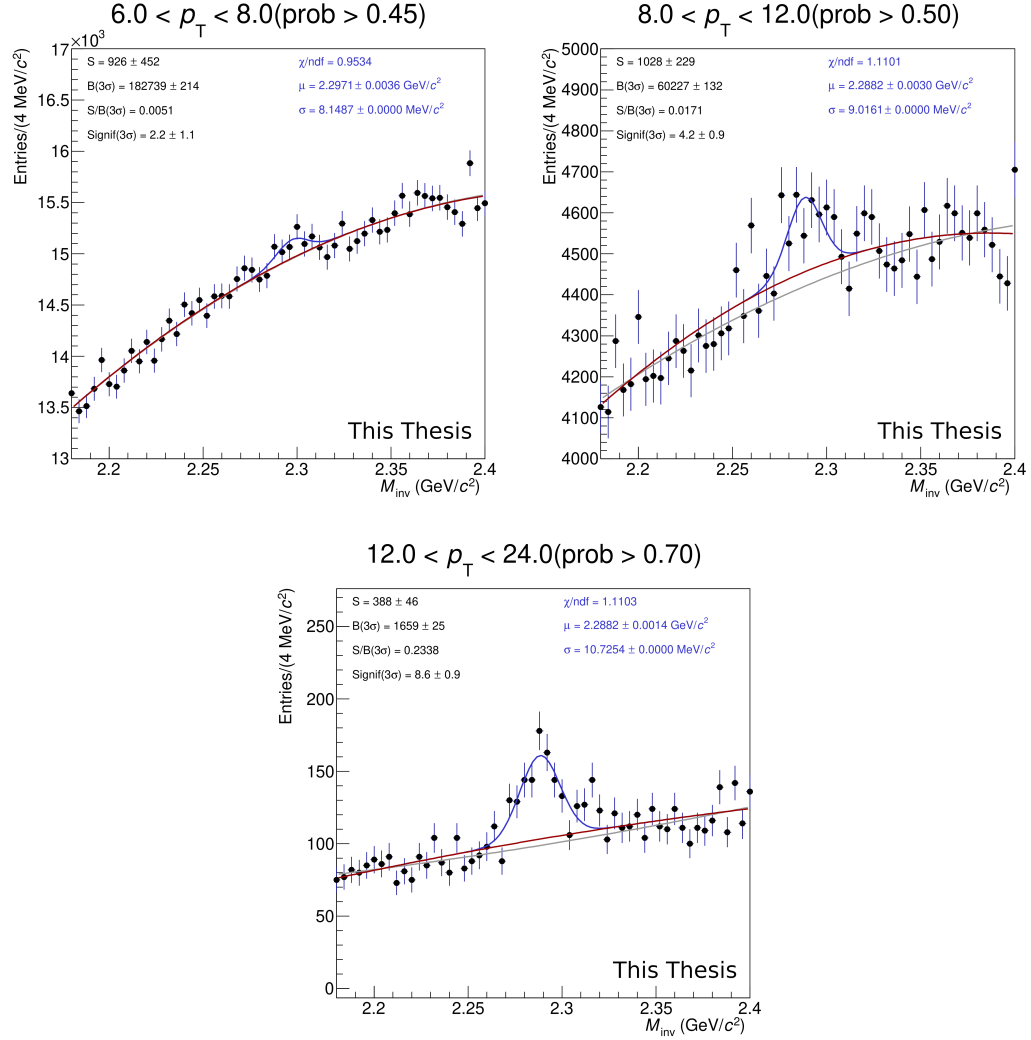


Figure 4.10: Invariant mass distributions of Λ_c^+ candidates in the 0-10% centrality interval. Shown are the $6 < p_T < 8 \text{ GeV}/c$, $8 < p_T < 12 \text{ GeV}/c$, and $12 < p_T < 24 \text{ GeV}/c$ intervals. Fits to the Λ_c^+ signal (blue) and background (red) are also shown.

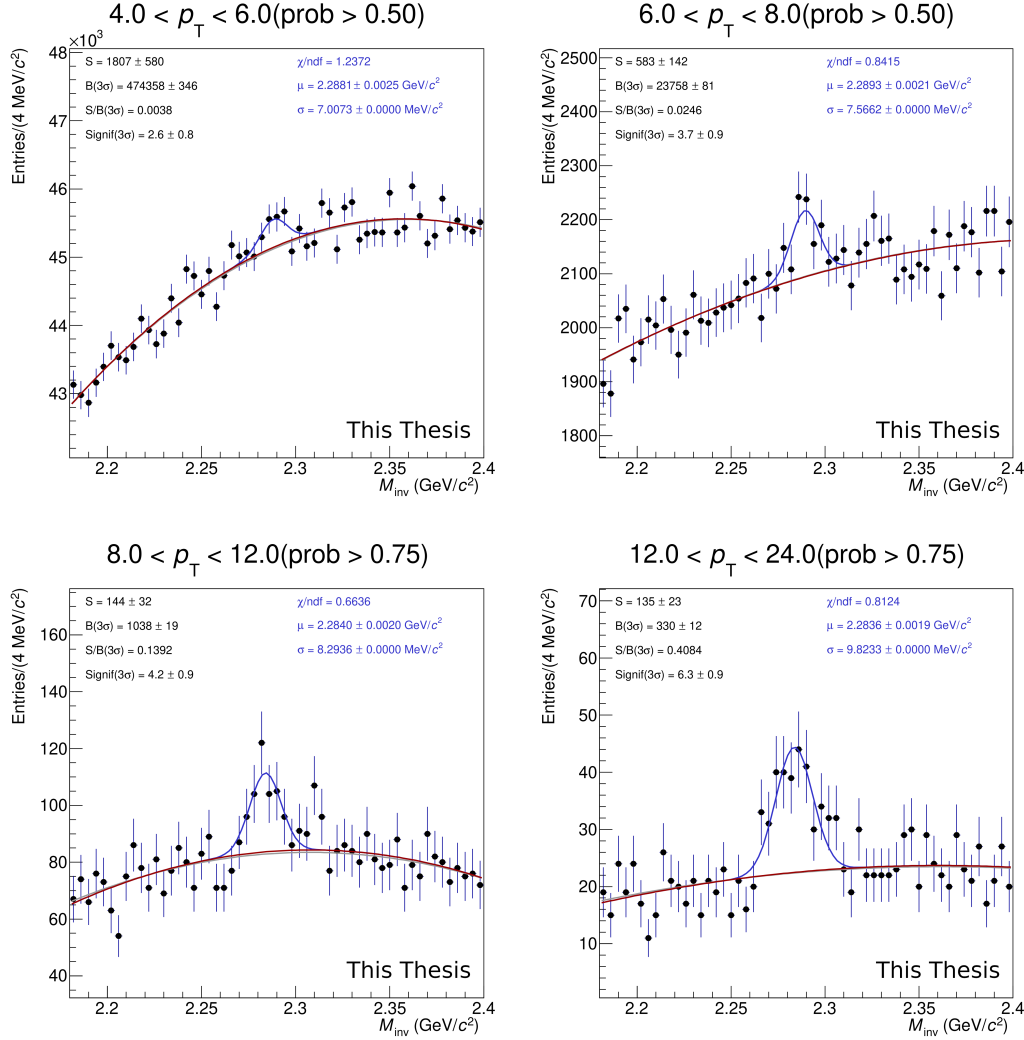


Figure 4.11: Invariant mass distributions of Λ_c^+ candidates in the 30-50% centrality interval. Shown are the $4 < p_T < 6 \text{ GeV}/c$, $6 < p_T < 8 \text{ GeV}/c$, $8 < p_T < 12 \text{ GeV}/c$, and $12 < p_T < 24 \text{ GeV}/c$ intervals. Fits to the Λ_c^+ signal (blue) and background (red) are also shown.

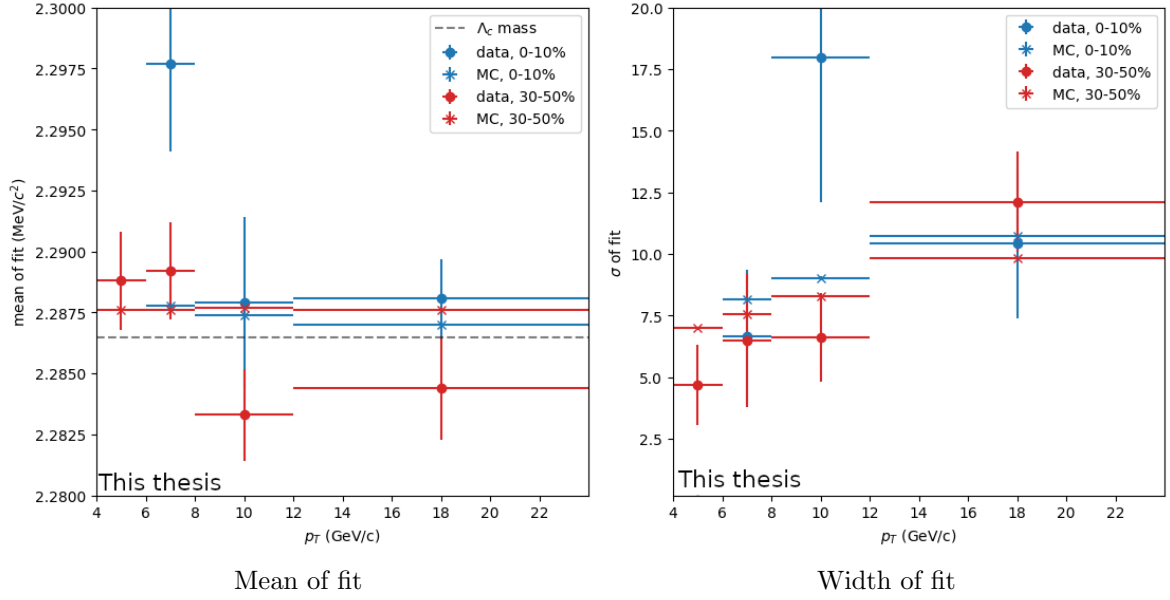


Figure 4.12: The mean (left) and width (right) of the Gaussian fits to the Λ_c^+ signal peak in Monte Carlo simulation compared to the fit in data when leaving the width as a free parameter.

the efficiency calculation is different to the one used for training the BDT model, and is therefore independent of it. Additionally, an acceptance factor A is calculated, which is defined as the number of Λ_c^+ candidates within the rapidity range of this analysis ($-0.5 < y < 0.5$), divided by the total number of Λ_c^+ candidates within the acceptance of the detector. The resulting total efficiencies for both prompt and feed-down Λ_c^+ candidates are shown in figure 4.15. They range from $\sim 0.5\%$ to $\sim 3\%$ for prompt Λ_c^+ and $\sim 2\%$ to $\sim 3.5\%$ for non-prompt Λ_c^+ .

4.5.2 Feed-down

The Λ_c baryon can also be produced from decays of Λ_b or other hadrons containing a b quark. Since this analysis measures only prompt Λ_c baryons, a feed-down subtraction needs to be applied to determine the fraction of the raw yield that is expected to originate from feed-down Λ_c^+ . The feed-down Λ_c^+ production cross section is de-

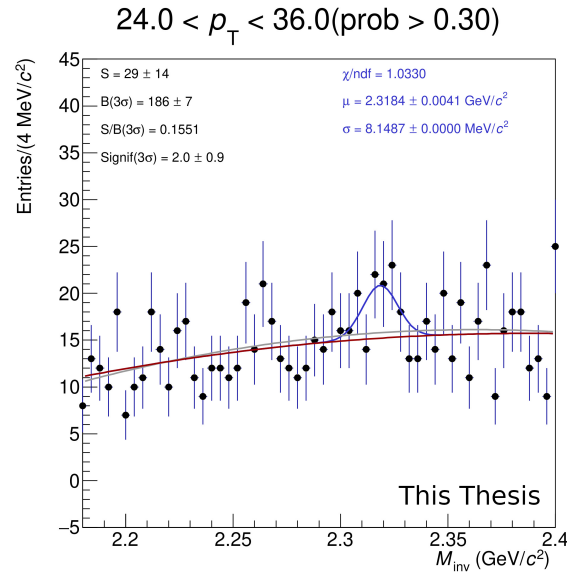


Figure 4.13: The invariant mass distribution for Λ_c^+ candidates in central Pb–Pb collisions in the $24 < p_T < 36 \text{ GeV}/c$ bin. Due to a lack of statistics the extension of the analysis to this momentum bin was not successful, as the significance of 2 is well below the minimum required 3.

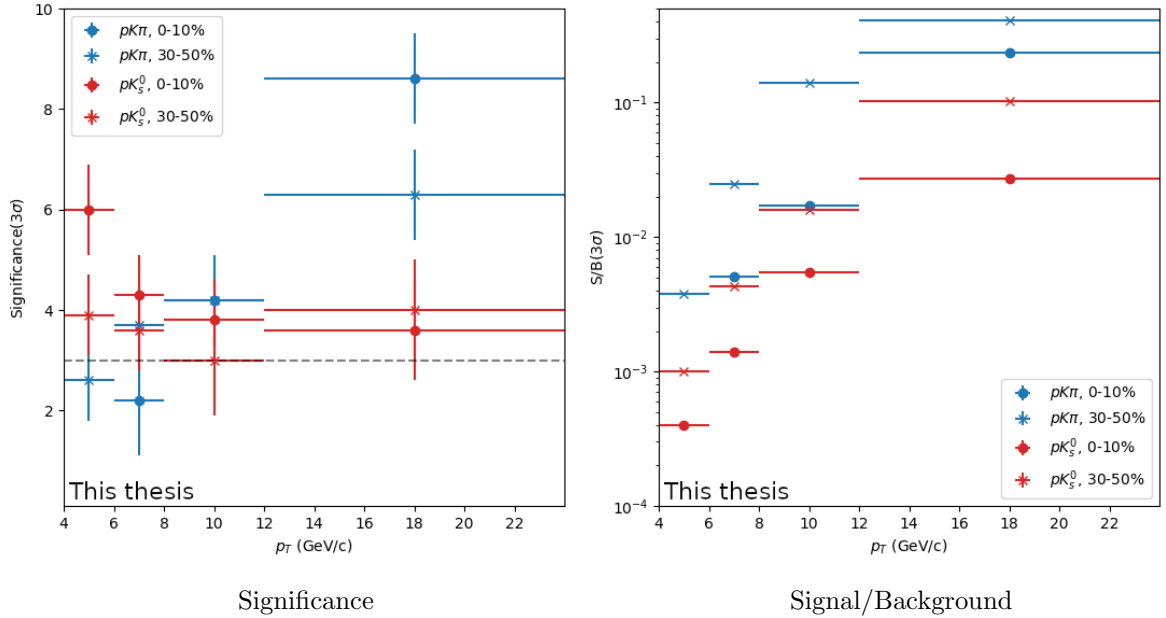


Figure 4.14: The invariant mass fit results for this analysis as a function of p_T compared to results from a previous ALICE analysis for the $\Lambda_c^+ \rightarrow pK_S^0$ channel, with both collision centrality classes shown. On the left is the significance of the fit, with the imposed minimum quality of $\text{Significance} > 3$ shown as a dotted line. On the right are the signal to background ratios for each bin. It can be seen that while this analysis is less performant in the low momentum bins, it offers an improvement at higher momentum.

terminated using several different calculations. The first step is obtaining the beauty production cross section, which is taken from FONLL calculations [113, 114]. Since these are done for pp collisions, the production cross section is scaled by the nuclear overlap function $\langle T_{AA} \rangle$, which can be interpreted as the number of nucleon-nucleon collisions at a given impact parameter. This way, the expected production is scaled for the number of expected binary nucleon-nucleon collisions in the centrality class of interest. LHCb measurements of the $\Lambda_b/(B^0 + B^+)$ ratio at $\sqrt{s} = 13 \text{ TeV}$ [115] are used to estimate the fraction of b quarks that fragment into Λ_b . The decay kinematics of $\Lambda_b \rightarrow \Lambda_c + X$ are taken from PYTHIA8 calculations. In order to account

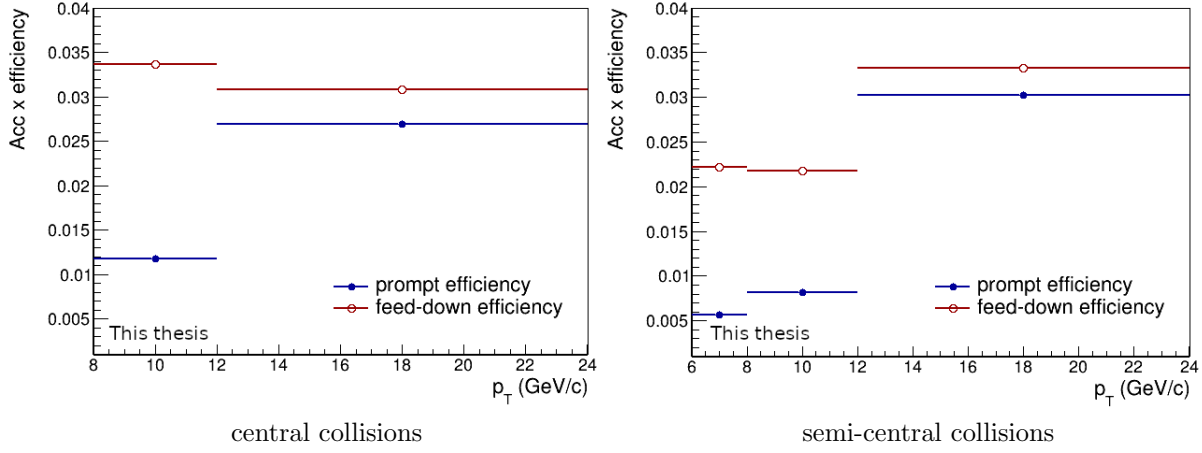


Figure 4.15: The detector acceptance corrected efficiency for prompt and feed-down Λ_c^+ baryons for central (left) and semi-central (right) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.

for the fact that the $b\bar{b}$ production is not measured in Pb–Pb collisions, a theoretical estimation from the Catania model is used of $R_{AA}^{\text{non-prompt}}/R_{AA}^{\text{prompt}} = 2$ [116]. Using these input values, the prompt fraction is calculated as

$$\begin{aligned}
 f_{\text{prompt}} &= 1 - \left(N_{\Lambda_c^+}^{\text{feed-down raw}} / N_{\Lambda_c^+}^{\text{raw}} \right) \\
 &= 1 - \langle T_{AA} \rangle \times \left(\frac{d^2\sigma}{dy dp_T} \right)_{\text{feed-down}}^{\text{FONLL}} \times R_{AA}^{\text{feed-down}} \\
 &\quad \times \frac{(\text{Acc} \times \epsilon)_{\text{feed-down}} \times \Delta y \Delta p_T \times \text{BR} \times N_{\text{evt}}}{N_{\Lambda_c^+}^{\text{raw}} / 2}
 \end{aligned} \tag{4.4}$$

The $1/2$ factor of the raw yield is to account for both particles and anti-particles having been measured, while the FONLL calculation only contains the particles. Figure 4.16 shows the feed-down fraction, calculated separately for central and semi-central Pb–Pb collisions. For central collisions, a prompt fraction of around 60% for lower p_T and around 80% for higher p_T was found, while for semi-central collisions the prompt fraction is slightly higher, between 70% and 90%. The figure also shows

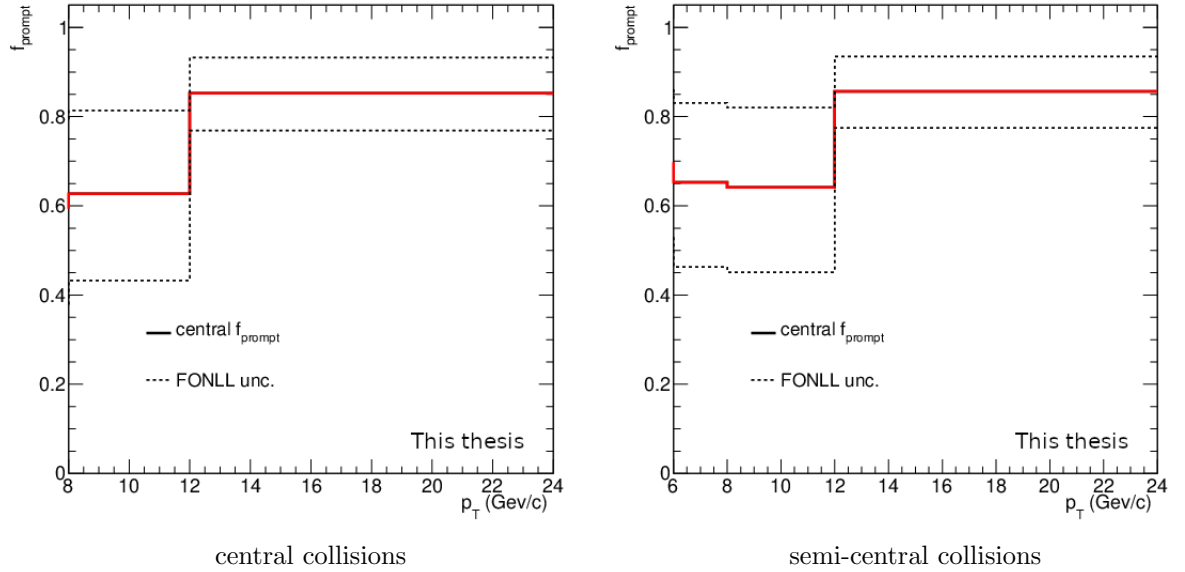


Figure 4.16: Fraction of prompt Λ_c^+ baryons calculated as a function of p_T for central (left) and semi-central (right) Pb–Pb collisions as a function of p_T . The central value is shown as the red line, and the systematic uncertainties are represented by the black dotted lines.

the systematic uncertainties on these calculations, which are discussed in 4.6.6.

4.6 Systematic Uncertainties

In this section, the systematic uncertainties estimated in this analysis are discussed. A wide range of sources of uncertainty are evaluated, from those coming from steps or choices in the Λ_c^+ reconstruction, selection, and signal extraction, to global uncertainties like that on the branching ratio. The sources of systematic uncertainties include:

- **Yield extraction:** the uncertainties on the raw yield extraction, which are found by repeating the signal fit while changing the fit parameters.

- **BDT cut:** The uncertainties coming from the BDT cut selection. The cut value is varied to estimate the influence on the result.
- **MC p_T -shape:** This uncertainty originates from the fact that the p_T distribution of MC data used for the efficiency calculation is not necessarily that of the true shape of the experimental data. To mitigate this, different p_T -distribution shapes are applied and the impact on the result is compared between them.
- **Tracking efficiency:** Uncertainty stemming from the difference in the track reconstruction efficiency determined in simulation to that in data.
- **Pre-selection PID efficiency:** Uncertainty from the Bayesian PID selection. It is estimated by changing the Bayesian probability cut and comparing efficiencies.
- **Feed-down subtraction:** Uncertainty on the prompt fraction, which is investigated by varying the hypothesis parameters.

4.6.1 Raw Yield Extraction

As described above, the raw yield is determined via a fit to the invariant mass spectrum. In order to determine this systematic uncertainty, the fitting procedure is repeated many times while varying the details of the fitting method. A total of 450 trials was run for each p_T and centrality bin. The fit parameters that were varied include:

- **Background fit function:** In the main analysis, a second-order polynomial is fit to the background. To ensure that the choice of function does not bias the result, different functions are used. In this case, polynomials of the second, third, and fourth order were considered.
- **Signal width:** In the analysis, the Gaussian width of the peak was set to the value obtained from the MC fit. In the multi-trial approach, this value was varied by $\pm 15\%$, while leaving it as a free parameter was also considered.

- **Fit range:** The minimum of the fit range for the Λ_c^+ mass was varied between (2.140, 2.170, 2.180, 2.190, 2.200) GeV/c, and the maximum range between (2.380, 2.390, 2.400, 2.410, 2.420) GeV/c.
- **Bin counting:** By default, the raw yield is extracted from the fit to the invariant mass by integrating the area under the signal peak and subtracting the background by integrating the same area of the background fit. Alternatively, the number of signal events can be found by counting the entries in each individual bin that lies within the signal peak and adding them up, which is called bin counting. This is more prone to statistical fluctuation, which is why the former method is used in the main analysis. It is still important, however, to check that they do not deviate from one another in a significant way.

A minimum significance of 2.7 was set for the results of these fits, and they were required to have $\chi^2/\text{NDF} < 2$. An example of the outcome of such a series of refits is shown in figure 4.17 for semi-central events in the $12 < p_T < 24$ GeV/c bin. The systematic uncertainty stemming from the yield extraction was found by looking at the RMS of the multi-trial distribution.

4.6.2 BDT cut value

The cut on the Boosted Decision Tree response is determined through the significance optimisation procedure described in 4.4.1. There is a systematic uncertainty on this method, as there is some discrepancy between the data and MC. The detector is not modelled perfectly, and might have a different response in the simulation than it does in reality. In addition to that, the results from the pK_S^0 channel that go into calculating the BDT efficiency also have an uncertainty on them.

In order to evaluate this systematic uncertainty, the analysis was repeated several times with variations to the BDT cut value. It was changed in steps of 0.05 from -0.15 to +0.15 of the original value in the main analysis. It was ensured that the resulting variation in efficiency is at least 20%, to cover an acceptably large region. An example of the variations in the invariant mass plots found as a result of those changes is shown in figure 4.18 for the $12 < p_T < 24 \text{ GeV}/c$ bin for central events. As in the previous section, fits with a significance lower than 2.7 were ignored. Those that passed the threshold were then used to calculate the corrected yield for each of the variations, which was compared to the main analysis, as shown in figure 4.19. The systematic error was taken as the RMS of the relative variations compared to the main value.

4.6.3 MC p_T -shape

Another source of systematic uncertainty investigated is the one stemming from the p_T distribution of the generated Λ_c^+ signal in the Monte Carlo simulation. It is found by varying the shape of the p_T distribution of the generated Λ_c^+ signal according to predictions from different models than that used in the main analysis. Specifically, this is done by using predictions of the Λ_c^+ to D^0 ratio and multiplying it by the D^0 p_T -distribution shape, thereby obtaining different hypotheses for the Λ_c^+ p_T -distribution shape. Then, the p_T distribution in the Monte Carlo simulation used for the efficiency is re-weighted to match that found from the model. To estimate the systematic uncertainty, this step is repeated for different models, and the corrected yield is measured using these different efficiencies. Figure 4.20 shows the variation

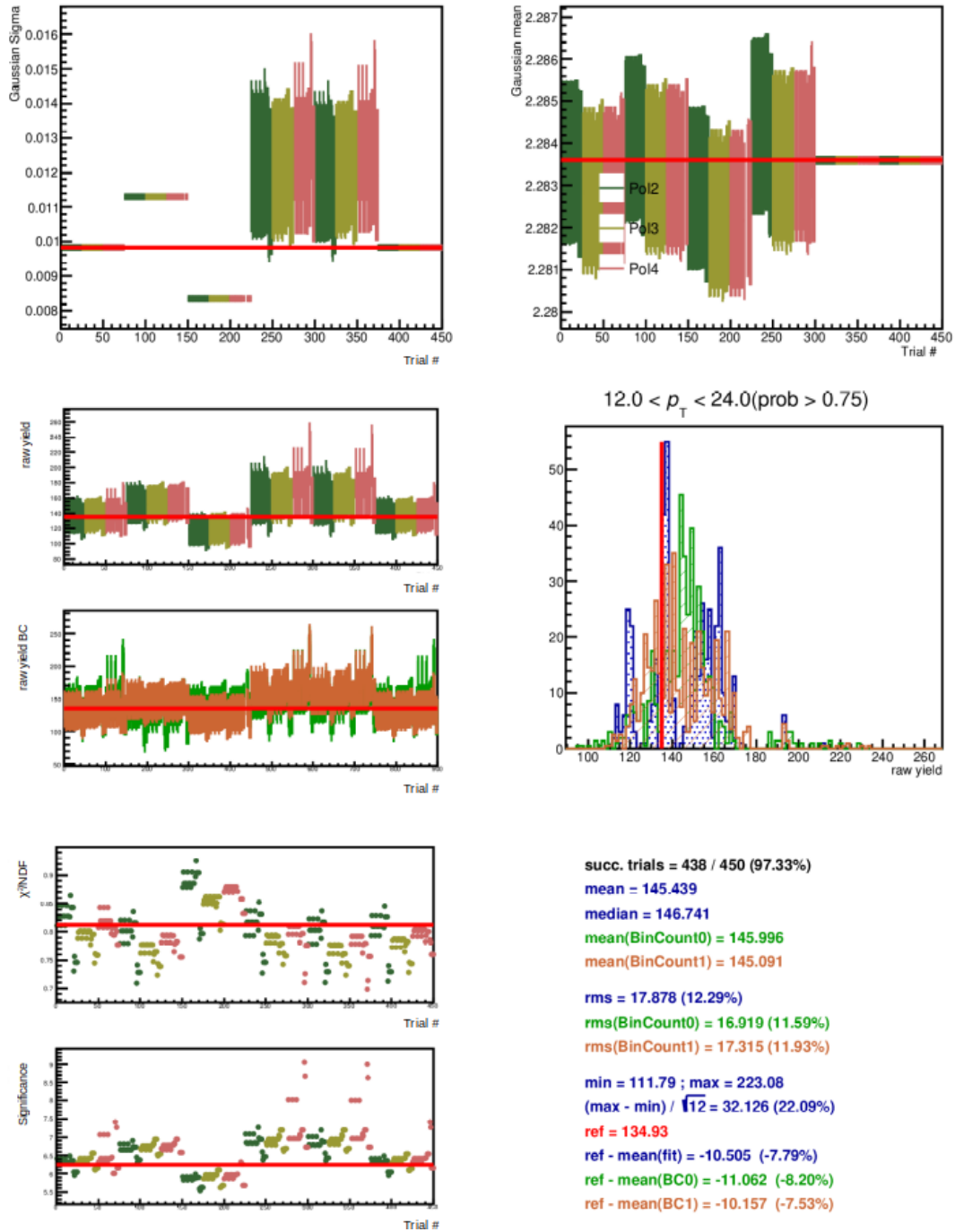


Figure 4.17: The results of the multi-trial variation of the fitting parameters of the raw yield extraction. They show, starting from the top left: The distribution of the fit width σ , in cases where it is fixed or left free, with the value from the main analysis shown in red; The distribution of the mean of the fits; the raw yields from the trials without (top) and with (bottom) bin counting; the distribution of the raw yield from all the trials; and the χ^2/NDF (top) and significance (bottom) of the trials. These plots are for Pb–Pb collisions in the 30-50% centrality bin and $12 < p_T < 24 \text{ GeV}/c$ bin, and yield an RMS of 12%.

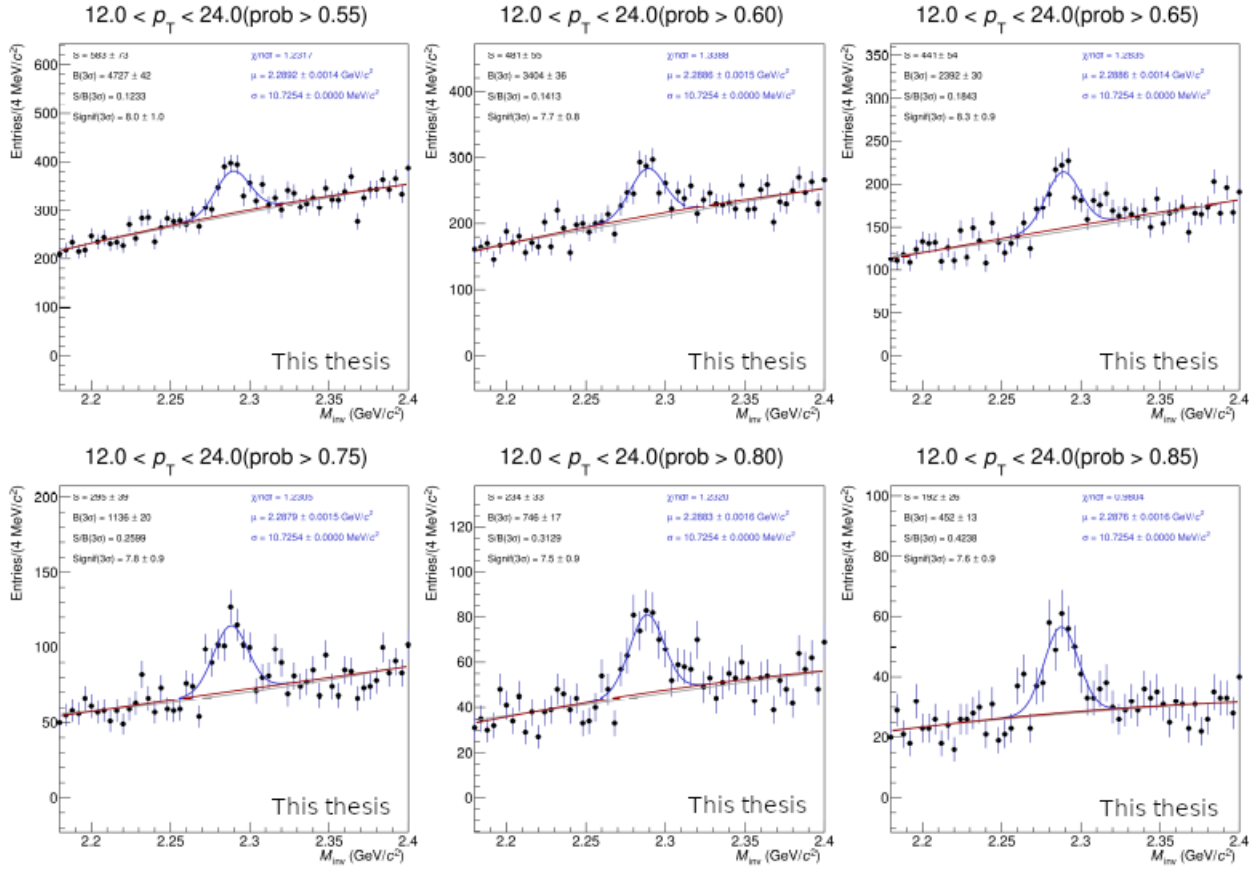


Figure 4.18: The invariant mass distributions for central Pb-Pb collisions in the $12 < p_T < 24 \text{ GeV}/c$ momentum bin resulting from varying the BDT cut around the central value of 0.7.

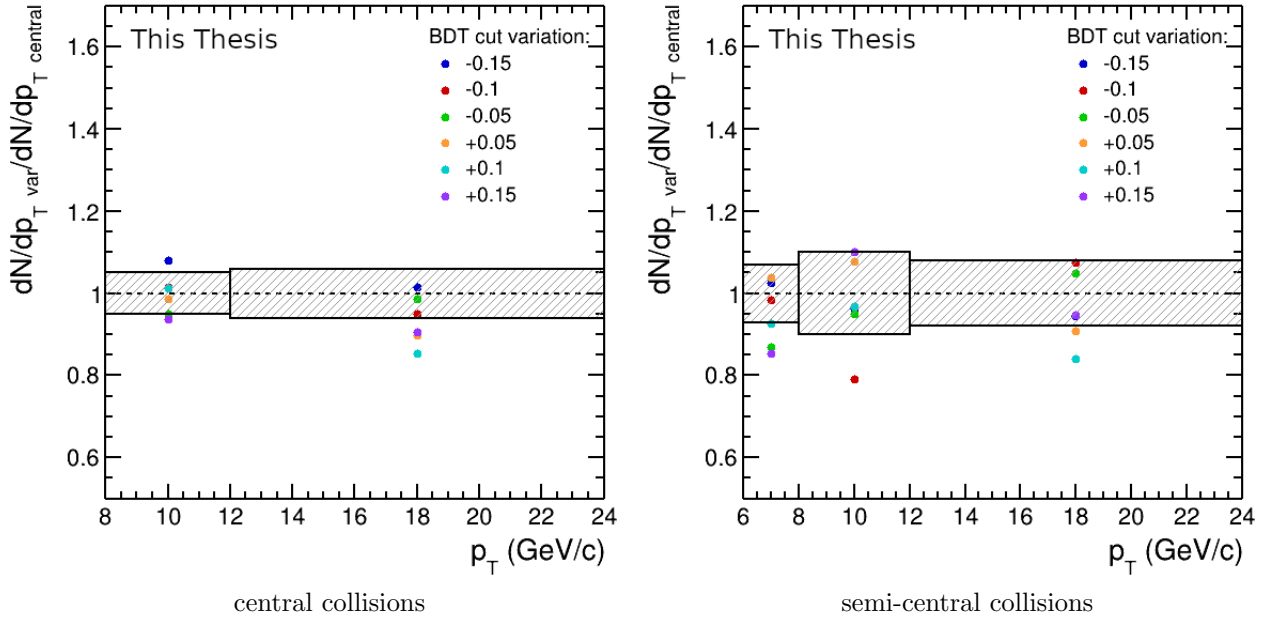


Figure 4.19: The corrected Λ_c^+ yields resulting from the different BDT cut values applied, normalised with the yield found with the central cut value. The shaded area represents the systematic uncertainty derived from this.

from the different models when compared to that used in the main analysis. The models used for the variations are TAMUmin [27], Catania [26], and SHM [15], which are described in section 5.3. The uncertainty is found by calculating the RMS of the relative changes.

4.6.4 Tracking Efficiency

There are two main sources of uncertainty on the track-reconstruction efficiency. One is the track-quality selection in the TPC, which comes from possible imperfections in the simulation of the track reconstruction in the TPC. The other stems from the simulation of the propagation of tracks from the TPC inwards to the ITS.

The matching efficiency is defined as the fraction of the number of tracks that pass

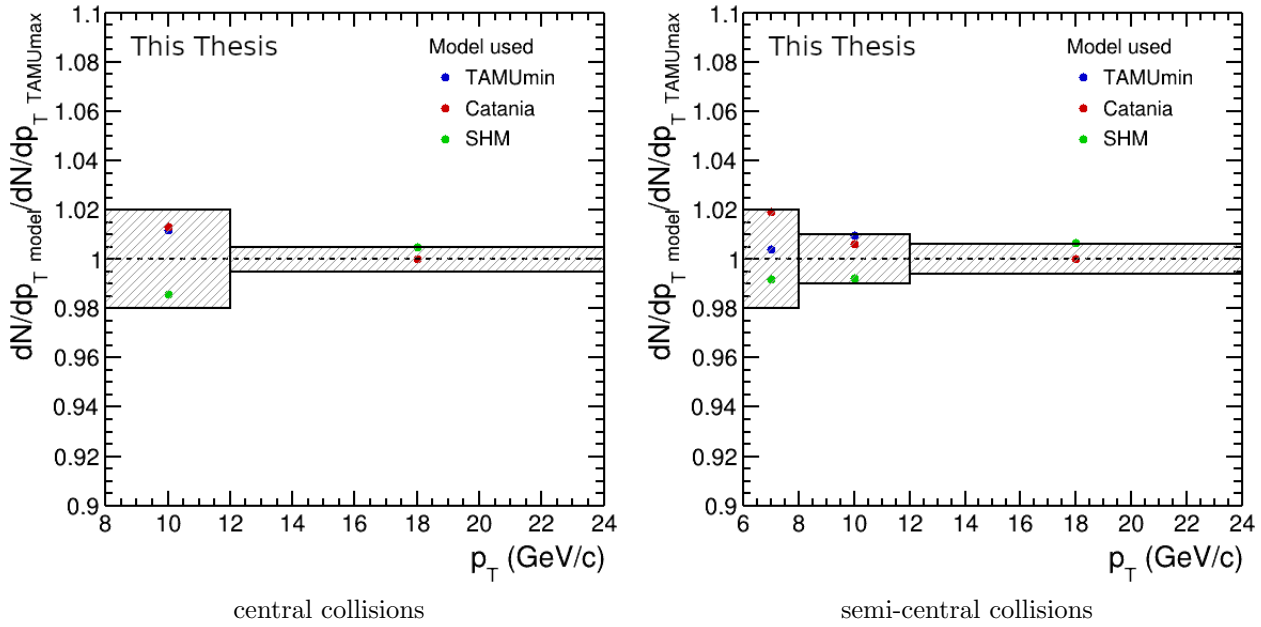


Figure 4.20: The corrected Λ_c^+ yields resulting from the different p_T MC p_T -distribution shapes applied, compared to that from the TAMUmax model used in the analysis. The shaded area represents the systematic uncertainty derived from this.

the ITS-TPC refit stage compared to those in the TPC:

$$\epsilon = \frac{\# \text{tracks}_{\text{ITSTPC}}}{\# \text{tracks}_{\text{TPC}}} \quad (4.5)$$

The systematic uncertainty on this is estimated by looking at the difference of this efficiency found in experimental data to that from the Monte Carlo simulation data:

$$\text{syst} = \frac{\epsilon_{\text{exp}} - \epsilon_{\text{MC}}}{\epsilon_{\text{exp}}} \quad (4.6)$$

It is shown in figure 4.21 for central and semi-central collisions.

The single track-quality selection uncertainty is obtained individually for pions and protons. For this, relatively pure samples of both particles are taken. For pions,

those are taken by selecting K_S^0 decay products (K_S^0 decays into $\pi^+\pi^-$), and for protons from Λ_c^+ decays. Five different variations of the track-quality selection are tested, which include:

- increasing the required number of TPC crossed rows to 120 instead of 70;
- increasing the minimum ratio of crossed rows over findable clusters in the TPC to 0.9, up from 0.8;
- varying the number of PID clusters to be larger than 40 or 60.

For each variation, the change in efficiency with respect to the central efficiency is determined for both data and MC. The ratio of these variations is then taken, and is shown for pions in figure 4.22, with the plots for other particles shown in appendix A. The uncertainty was taken to be the maximum relative change of all the variations. The single track uncertainties are propagated to the Λ_c^+ candidates by simulating $\Lambda_c^+ \rightarrow pK^-\pi^+$ decays with PYTHIA. The total uncertainty is found by summing the single-track contributions in quadrature:

$$\sigma_{\text{tot}} = \sqrt{(\sigma_p^{\text{TPC}})^2 + (\sigma_p^{\text{match}})^2} + \sqrt{(\sigma_K^{\text{TPC}})^2 + (\sigma_K^{\text{match}})^2} + \sqrt{(\sigma_\pi^{\text{TPC}})^2 + (\sigma_\pi^{\text{match}})^2} \quad (4.7)$$

It is shown in figure 4.23.

4.6.5 Pre-selection PID efficiency

Before the BDT model is trained, particles are identified following a Bayesian probability approach, as described in section 4.3.2. In it, each track is assigned the particle species with the highest probability. In order to determine the systematic uncertainty from this PID strategy, the selection criteria are varied. Instead of being assigned a particle species based on the maximum Bayesian probability, a track is assigned to each species for which it passes a minimum probability threshold. This is done for several different threshold values, namely 0.5, 0.6, and 0.7. Then, the analysis is run again for each of those different PID pre-selections, and the resulting corrected yield is compared to that of the main analysis. To avoid the influence

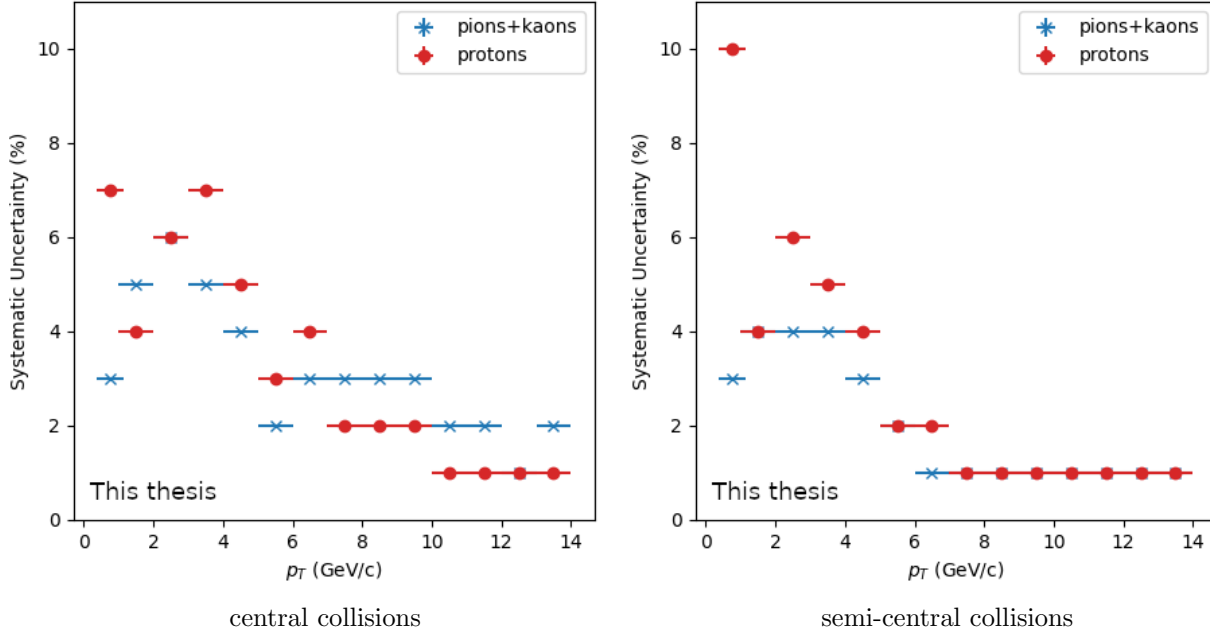


Figure 4.21: The systematic uncertainty on the matching efficiency between TPC and ITS, shown for pions and kaons combined, as well as for protons, for central and semi-central collisions.

of the BDT model on these variations, no new BDT model is trained, but instead the one from the main analysis is applied, and the same cut on the BDT response is used. Figure 4.24 shows the variation in the corrected yield when applying the different pre-selection PID cuts, with the uncertainties calculated as the RMS of this distribution.

4.6.6 Feed-down subtraction

There are several sources of systematic uncertainties from the feed-down subtraction that are taken into account. Firstly, the FONLL calculation of the beauty-quark cross section has theoretical uncertainties associated with it. They are determined by varying the b-quark mass, and the renormalisation and factorisation scales in the

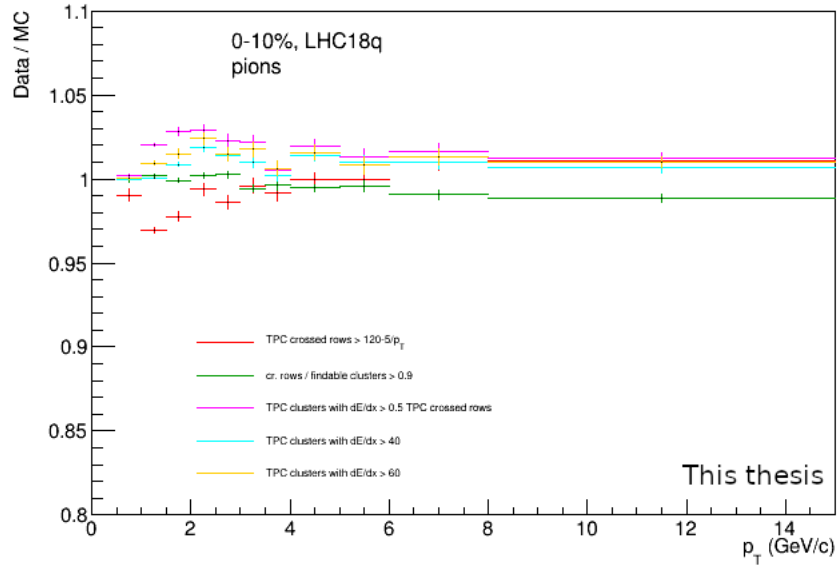


Figure 4.22: The relative single-track quality selection efficiency shown for pions for the additional TPC selections in experimental data over that in MC data, for central collisions.

calculation. Secondly, the parameters from the b-quark fragmentation into Λ_b from the measurement done by LHCb [115] were varied within its uncertainty via a fit to the Λ_b over B meson ratio. Thirdly, an uncertainty coming from the estimation of $R_{AA}^{\text{non-prompt}}/R_{AA}^{\text{prompt}} = 2$ was considered. This was determined by varying the hypothesis of $R_{AA}^{\text{non-prompt}}/R_{AA}^{\text{prompt}}$ between 1 and 3. The resulting variation of the prompt R_{AA} is shown in figure 4.25, where it can be seen that the variation is quite high for low- p_T bins but smaller for high- p_T bins. The total uncertainty coming from these considerations ranges from about 10% to 30%, which is relatively high, and is shown in section 4.5.2 in figure 4.16. It is important to note that unlike the other uncertainties discussed here, this uncertainty source is asymmetric.

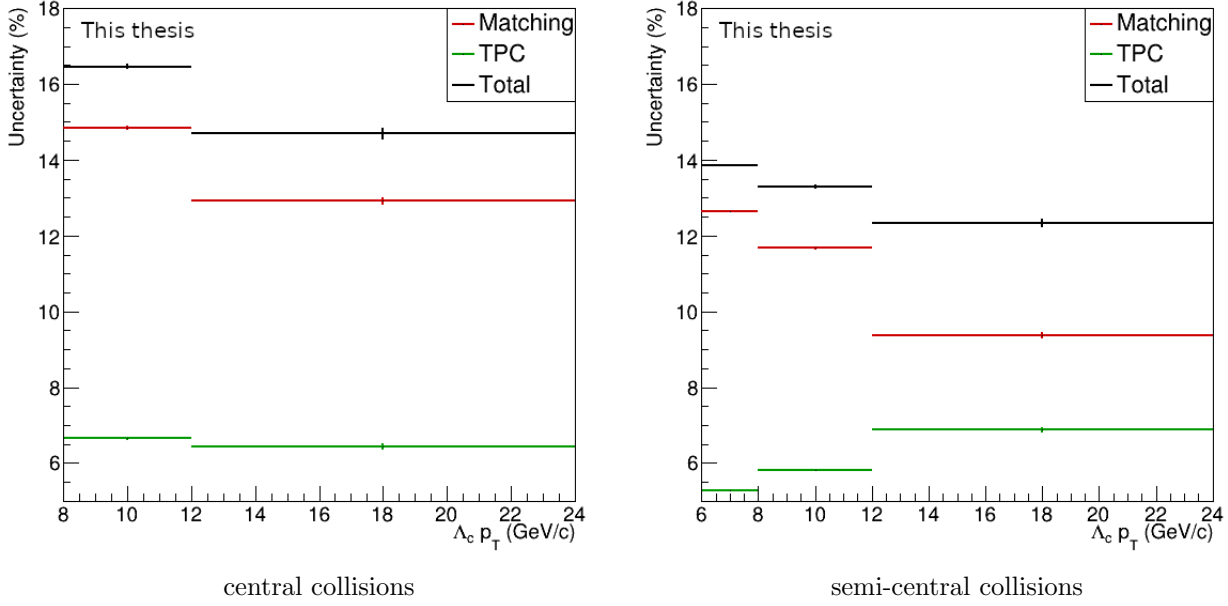


Figure 4.23: The combined Λ_c^+ systematic tracking uncertainty for central (left) and semi-central (right) collisions. The individual contributions from matching efficiency (red) and the TPC tracks (green) are also shown.

4.6.7 Summary

The systematic uncertainties are summarised in table 4.7. Aside from the uncertainties described in this section, it also includes the one stemming from the uncertainty on the branching ratio [2]. The data was taken using a minimum bias trigger, meaning the uncertainty from this source is expected to be very small. The individual uncertainties are assumed to be uncorrelated, and the final uncertainty is found by summing them in quadrature.

The largest contributions to the total uncertainty come from the feed-down subtraction and the yield extraction. Another major source is the tracking uncertainty. The later is expected to be significantly reduced in future measurements, due to the upgrade of the ALICE detector for Run 3. Additionally, the uncertainties decrease at higher transverse momentum, which indicates the possibility of extending this

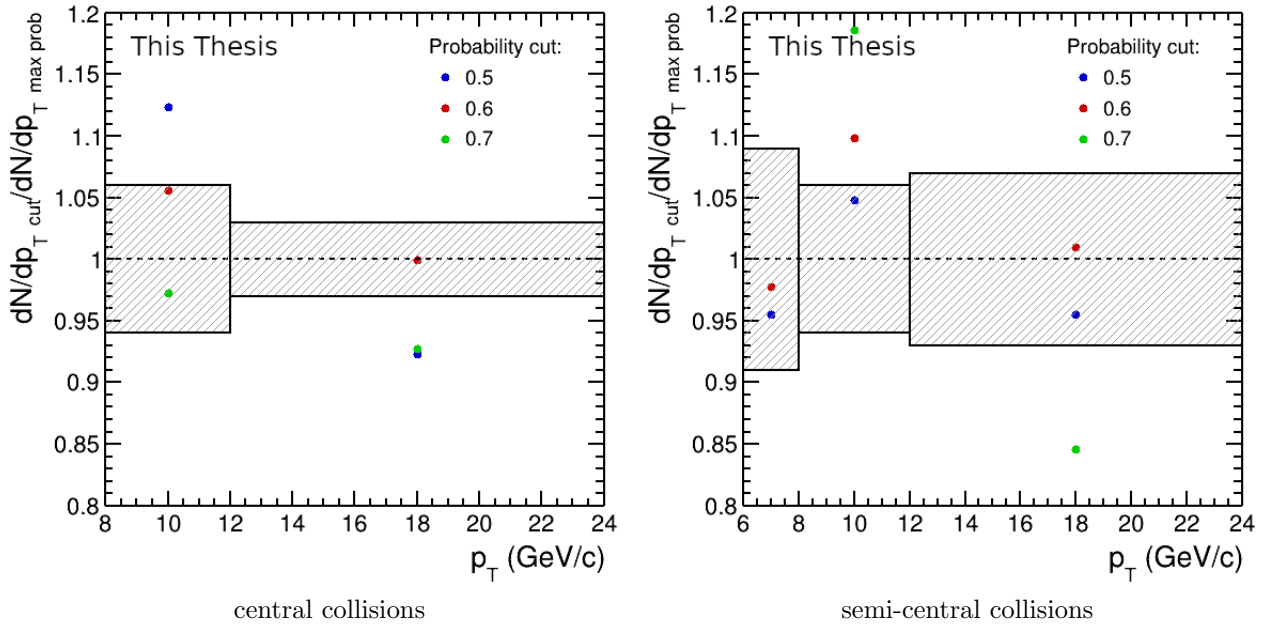


Figure 4.24: The corrected yields resulting from the different PID preselections, normalised by the yield from the main analysis. The shaded area represents the systematic uncertainty derived from this.

measurement in that region once more data becomes available.

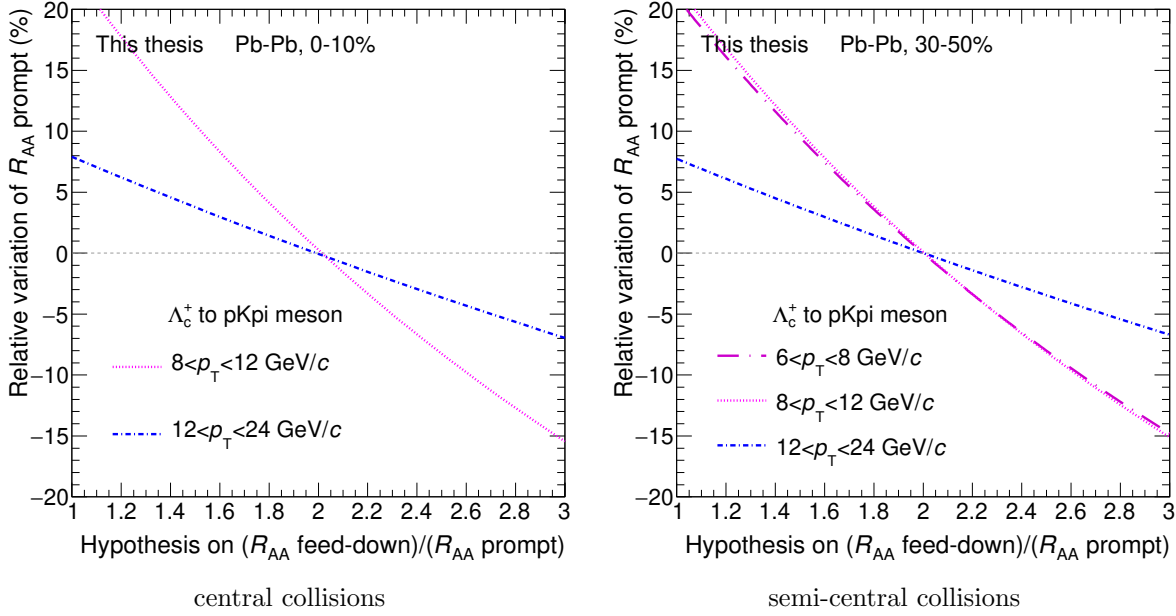


Figure 4.25: The uncertainty on the prompt R_{AA} fraction resulting from the variations of the $R_{AA}^{\text{Feed-Down}}/R_{AA}^{\text{prompt}}$ hypothesis for both central (left) and semi-central (right) collisions, as described in the text.

centrality	0-10%		30-50%		
$p_T(\text{GeV}/c)$	8-12	12-24	6-8	8-12	12-24
Yield extraction	29%	13%	10%	10%	12%
BDT cut	5%	6%	7%	10%	8%
MC p_T shape	2%	0.5%	2%	1%	0.6%
Tracking	16%	15%	14%	13%	12%
PID	6%	3%	9%	6%	7%
Feed-down	+30%	+9%	+27%	+28%	+9%
	-31%	-10%	-29%	-29%	-10%
Branching ratio	5.5%		5.5%		

Table 4.7: Summary of the systematic uncertainties found from the various sources discussed in this section.

Chapter 5

Results

In this chapter, the results of the data analysis described in the previous chapter are shown. The fully corrected production yield of Λ_c^+ baryons measured in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel is presented in Pb–Pb collisions in the centrality regions of 0-10% and 30-50% as a function of p_T in the range of $6 < p_T < 24 \text{ GeV}/c$, and is shown in section 5.1. The measurements reported in this analysis are made in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel. They are compared to previous ALICE measurements made in the $\Lambda_c^+ \rightarrow pK_S^0$ channel, and the results are merged to find the combined production yield. Using measurements of the D^0 production yields performed with ALICE [24], the ratio of these two yields, called the Λ_c^+/D^0 ratio, is also shown, and presented in section 5.2. The results are also compared to earlier measurements made in pp collisions. Section 5.3 describes three theoretical models, and 5.4 shows the comparison between the results and these models. It also shows the nuclear modification factor R_{AA} , which is found by comparing the production yields measured in Pb–Pb collisions to those found in pp measurements, in comparison to the same models.

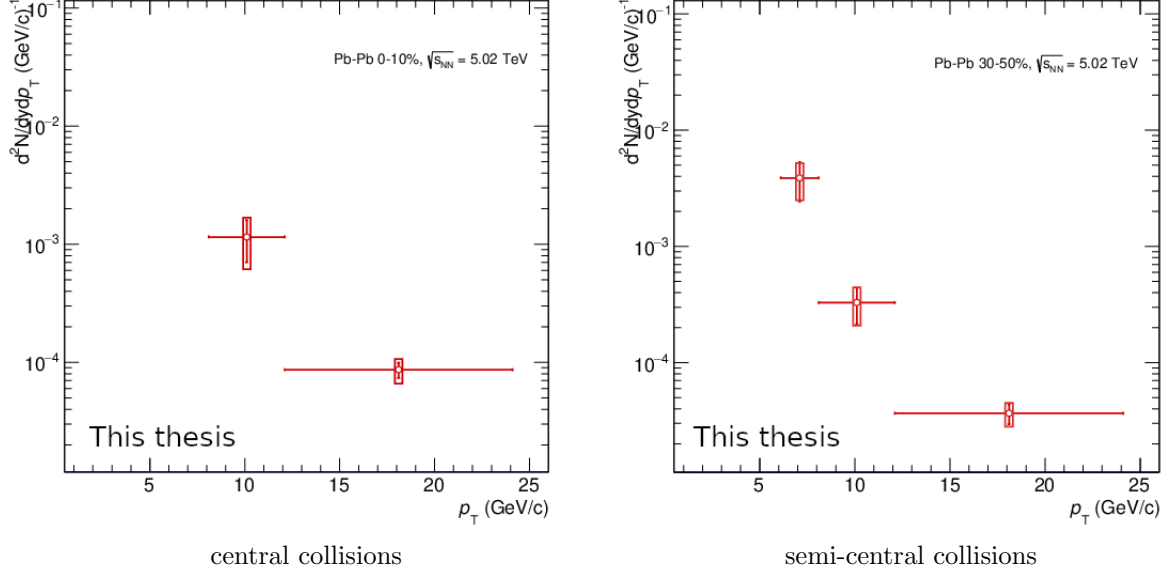


Figure 5.1: The p_T -differential production yield of Λ_c^+ baryons measured in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for central (left) and semi-central (right) events. The bars and boxes represent the statistical and systematic uncertainties, respectively.

5.1 Λ_c^+ production yield

The p_T -differential production yield of Λ_c^+ baryons has been measured in Pb-Pb collisions with a centre-of-mass energy of $\sqrt{s_{NN}} = 5.02$ TeV in the rapidity range of $|\eta| < 0.5$. The yield was reconstructed via the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel using a Boosted Decision Tree machine learning model. This was done in two centrality regions: central collisions at 0 – 10% centrality and semi-central events at 30-50% centrality. The production yield was measured in two p_T bins in central events and three p_T bins in semi-central events. The measured p_T -differentiated production yield is shown in figure 5.1.

As mentioned before, a previous measurement was done in ALICE to measure the Λ_c^+ production yield in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the $\Lambda_c^+ \rightarrow$

pK_S^0 channel [23]. Figure 5.2 shows the comparison between the production yield measurements from the two decay channels, and the two analyses agree with one another within uncertainties.

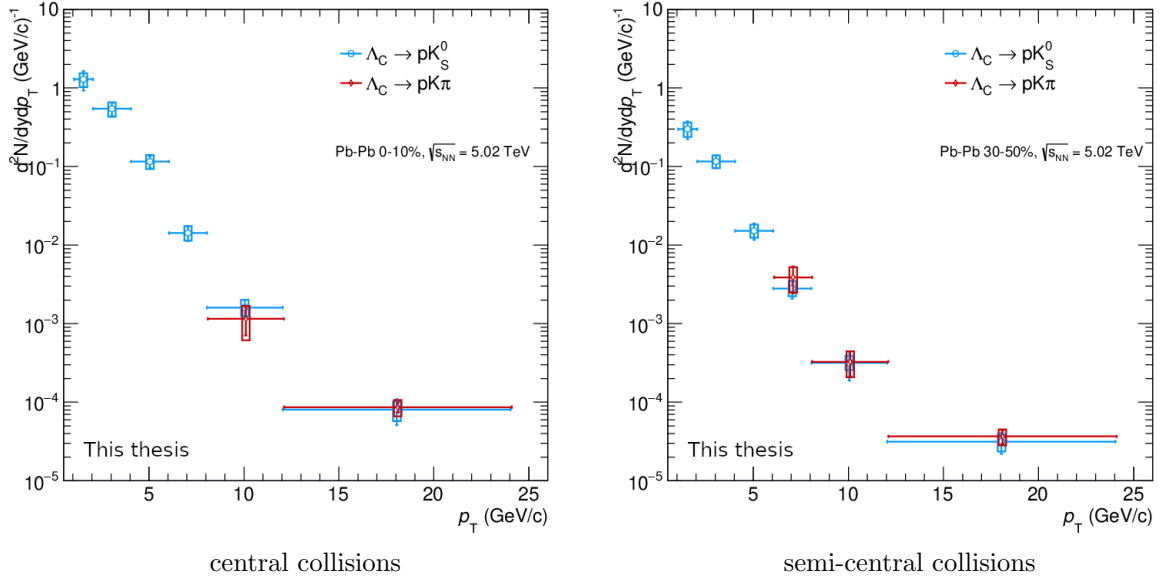


Figure 5.2: The p_T -differential production yield of Λ_c^+ baryons measured in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared to that measured in the $\Lambda_c^+ \rightarrow pK_S^0$ channel [23]. The two measurements are in good agreement. The bars and boxes represent the statistical and systematic uncertainties, respectively.

The results from the two decay channels are combined in order to increase the precision of the overall measurement. To do this, the results are weighted by their uncertainties. Here, it is important to distinguish between correlated and uncorrelated uncertainties between the two measurements. The uncorrelated uncertainties are those specific to the analysis procedure, which in this case are those on the raw yield extraction, the BDT cut efficiency, the particle identification strategy, and that of the branching ratio ¹. Since the analyses are carried out in different decay

¹In reality the uncertainties on the branching ratio are partially correlated, as one is used in the calculation of the other (see [2]). As the branching ratio uncertainty has only a minimal influence here however they are treated as uncorrelated for the purpose of this analysis.

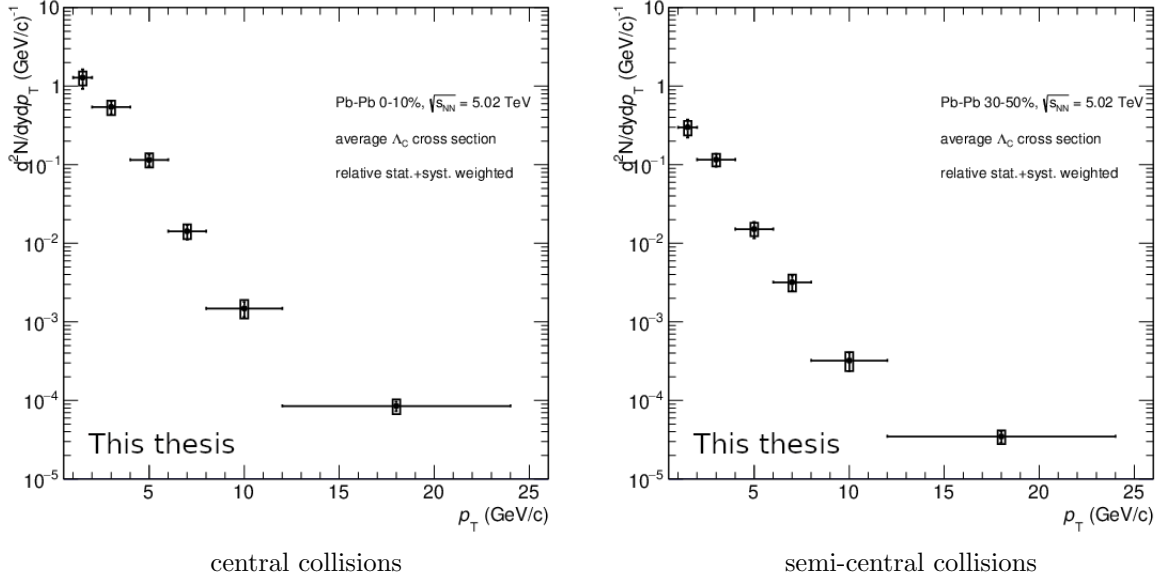


Figure 5.3: The p_T -differential production yield of Λ_c^+ baryons measured in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV from the combined results of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow pK_S^0$ channels. The bars and boxes represent the statistical and systematic uncertainties, respectively.

channels, the statistical uncertainties are also assumed to be uncorrelated. The correlated efficiencies are those that are the same for both analyses, namely the tracking uncertainty and the Monte Carlo p_T -shape uncertainty. In order to merge the measurements in the two decay channels, their results are averaged and weighted by the quadratic sum of their relative statistical and uncorrelated systematic uncertainties: $1/(w_i^{\text{uncorr}})^2$. Figure 5.3 shows the combined production yield obtained from combining the two decay channels. To get a better idea of the improvement on the Λ_c^+ production measurement in Pb–Pb collisions, the relative uncertainties of this analysis are compared to those from the $\Lambda_c^+ \rightarrow pK_S^0$ measurement. Figures 5.4 show both the systematic and statistical uncertainties as a function for p_T for the two different decay channels. The uncertainties resulting from the merging of the two analyses are also plotted, and show that adding the $\Lambda_c^+ \rightarrow pK^-\pi^+$ measurement improves the

statistical uncertainties by a large enough margin to outweigh the increase in the systematic uncertainty at lower p_T , while also yielding a clear improvement in the systematic uncertainties at high p_T .

In the following, all measurements of Λ_c^+ production in Pb–Pb collisions shown are derived from the combination of these two decay channels.

5.2 Λ_c^+/D^0 ratio

As described in chapter 1, the relative production of charm baryons and mesons can give an insight into hadronisation mechanisms, and how the presence of a quark-gluon-plasma might affect it. While this analysis measures the production yield of the Λ_c^+ baryon, the lightest charmed baryon, in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel, a similar analysis was done for the production of the D^0 meson, the lightest charmed meson, in the $D^0 \rightarrow K^-\pi^+$ decay channel [24]. By taking the ratio of these two measurements as a function of transverse momentum, the Λ_c^+/D^0 ratio is found, which is shown in figure 5.5. As done for the production yield, the results for the Λ_c^+/D^0 ratio measured for the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel are also compared to those found for the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel. Figure 5.6 shows the comparison of the two measurements, as well as the merged results for the Λ_c^+/D^0 ratio in Pb–Pb collisions.

In order to study how hadronisation mechanisms are influenced by the collision system, the production yields measured in Pb–Pb collisions are compared to those from measurements done in pp and p–Pb collisions [25]. Figure 5.7 shows a comparison between the Λ_c^+/D^0 ratio in central and semi-central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and that in pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV. It indicates a sequential enhancement, going from pp collisions to semi-central to central Pb–Pb collisions, at intermediate p_T . This suggests that either the hadronisation mechanism is modified in Pb–Pb collisions compared to pp collisions, for example via coalescence, or that there are other mechanisms in the presence of the QGP that redistribute the momentum of baryons and mesons. One example of a possible effect could be flow, which may affect baryons and mesons differently.

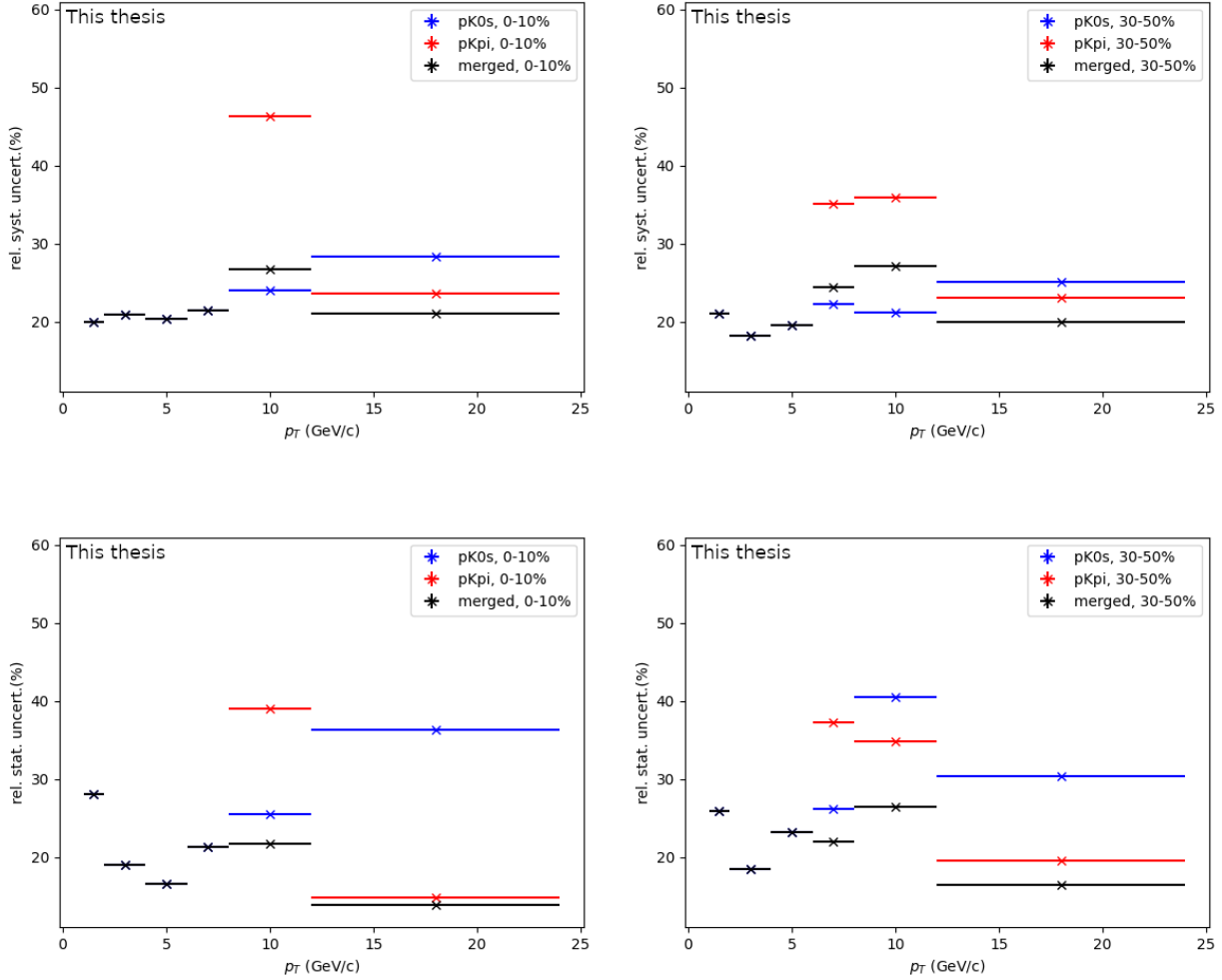


Figure 5.4: The relative systematic (top) and statistical (bottom) uncertainties for central (left) and semi-central (right) events for the two different decay channels, $\Lambda_c^+ \rightarrow pK^-\pi^+$ (red) and $\Lambda_c^+ \rightarrow pK_S^0$ (blue), with the total uncertainties found by merging the two results shown in black.

5.3 Theoretical Models

The results of this analysis are compared to three different theoretical predictions. These are the GSI-Heisenberg statistical hadronisation model (SHMc) [15], the Cata-

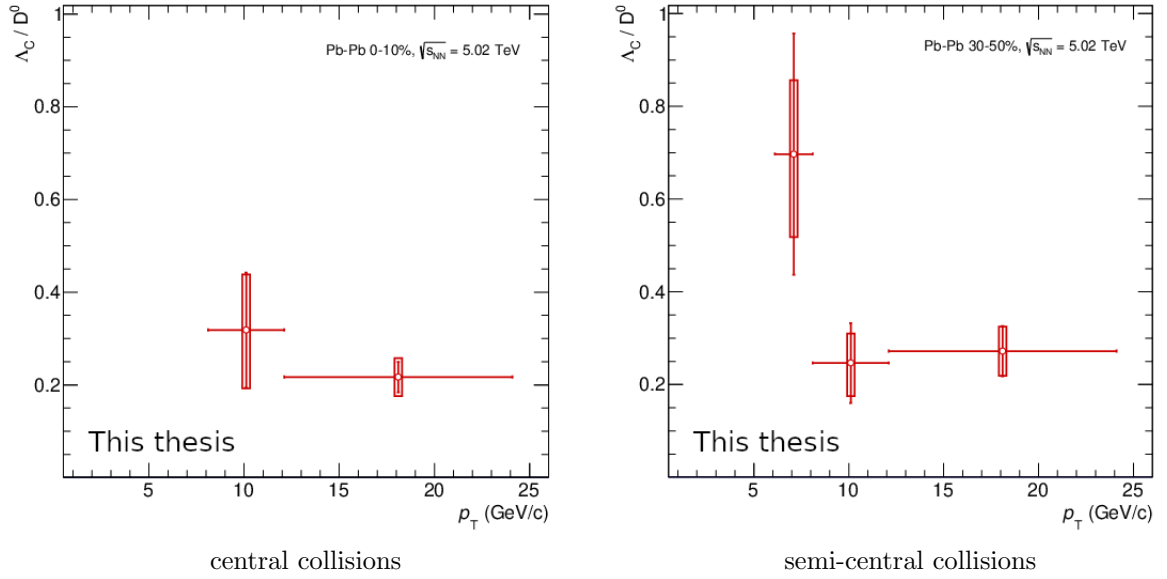


Figure 5.5: The Λ_c^+ / D^0 ratio measured in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for central (left) and semi-central (right) events. The measurement of D^0 production are taken from [24]. The bars and boxes represent the statistical and systematic uncertainties, respectively.

nia model [26], and the TAMU model [27].

In the GSI-Heisenberg SHMc model [15], the charm-hadron p_T spectra are modelled with a combination of a blast-wave model and a core-corona approach. In a blast-wave model, the spectra are obtained using a simple freeze-out² surface with a specific freeze-out temperature, which in this case is determined from light hadron analyses to be 156.5 MeV [117]. A typical limitation of a blast-wave model is that it does not take into account the modification of the particle spectra due to the feed-down from resonance decays. To overcome this, in this model the resonance decays are modelled using the FastReso method [118]. While the blast-wave model can be used to describe the soft momentum spectrum for charm hadrons, the high-

²The freeze-out temperature is the point in the hadronisation of the QGP at which the hadron species become fixed. See section 1.3.1.

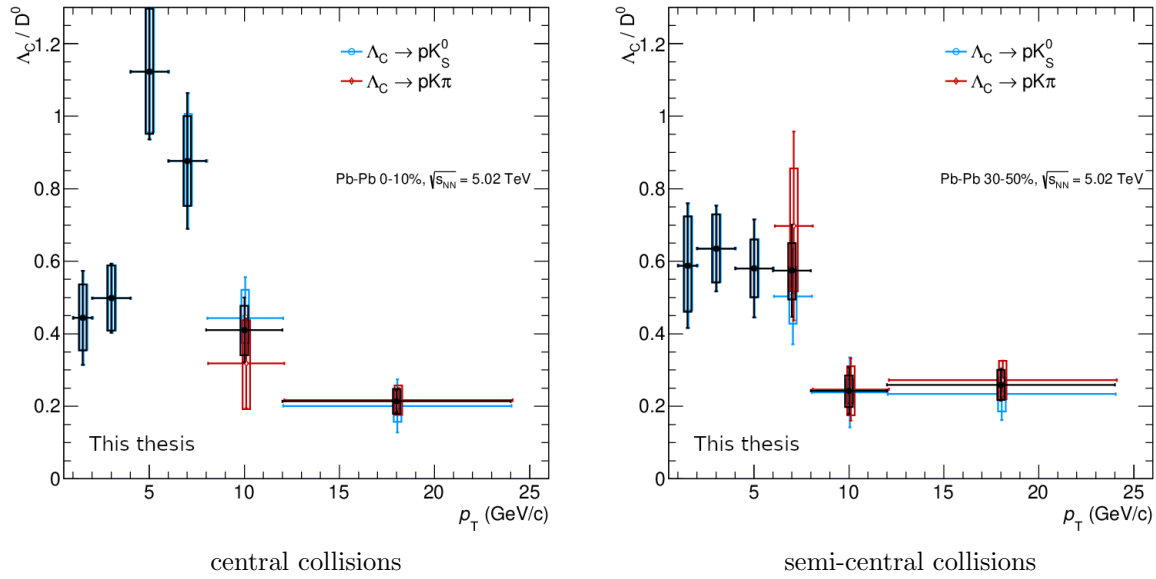


Figure 5.6: The Λ_c^+/D^0 ratio measured in the $\Lambda_c^+ \rightarrow pK^- \pi^+$ decay channel (red) compared to that from the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel (blue), both measured at $\sqrt{s_{NN}} = 5.02$ TeV. The black data points are the average of the two decay channels, and the bars and boxes represent the statistical and systematic uncertainties, respectively.

energy tail is calculated using a core-corona approach. In it, the corona region of the nucleon–nucleon collision is defined as the region where the overlap density is small (set to 10% of the maximum density in this case), where nucleons undergo only one or fewer collisions on average. From this, it is assumed that no QGP forms, and the collision here is treated as a pp collision. In the core part, on the other hand, it is assumed that there are many collisions and the charm quarks are fully thermalised.

In the Catania model [26], hadronisation is modelled to happen via both fragmentation and coalescence. For light partons, the thermal distribution in the quark-gluon plasma is modelled with a temperature similar to the phase transition temperature $T_c = 160$ MeV. It also includes radial flow, which is constrained by the total multiplicity and total transverse energy. The heavy quark transport in the quark-gluon

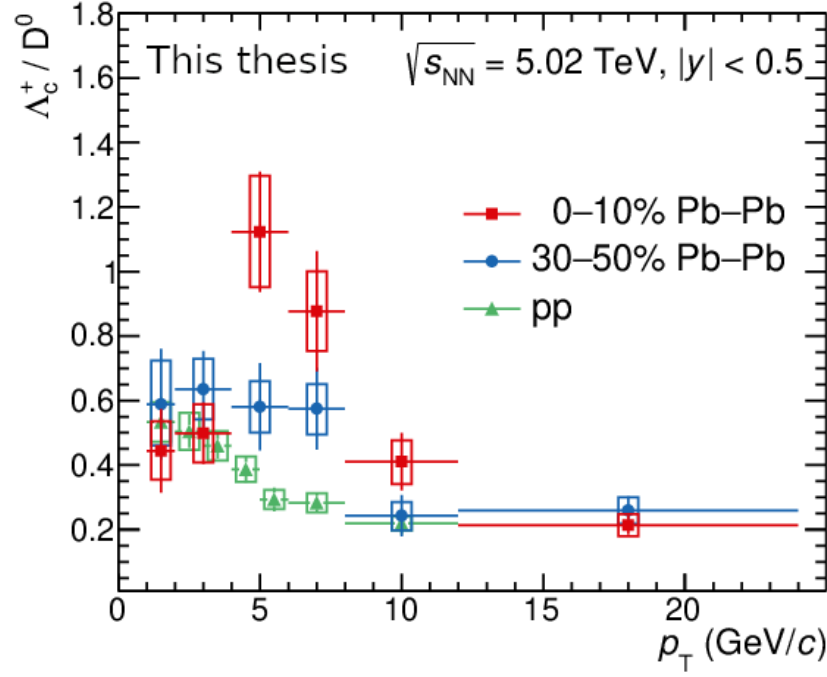


Figure 5.7: The merged Λ_c^+/D^0 ratio measured in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for two centralities compared to that measured in pp collisions also at $\sqrt{s_{NN}} = 5.02$ TeV [25]. The bars and boxes represent the statistical and systematic uncertainties, respectively.

plasma is found by solving the relativistic Boltzmann equation to obtain the transverse momentum distribution. In this model, charmed quarks at low momentum $p_T \simeq 0$ GeV are set to hadronise via coalescence, while quarks of finite momentum that do not are made to fragment. This ensures that all quarks of any momentum undergo hadronisation. The model also takes into account contributions from the decay of heavy-hadron resonances. It finds that the p_T dependence of the resulting baryon-to-meson ratio appears to be sensitive to the quark mass.

The TAMU model [27] is a 4-momentum-conserving recombination model for heavy-flavour mesons and baryons, where coalescence is implemented with a resonance recombination mode. In this model, an event-by-event implementation of heavy-quark distributions in an expanding medium is obtained using a Langevin simulation. It recovers the thermal and chemical equilibrium limits and accounts

for space-momentum correlations (SMCs) between heavy quarks and the partons of the hydrodynamically expanding QGP. The SMCs of high-momentum quarks with high-flow partons in the medium considerably increase the momentum range of coalescence, and lead to an enhancement in the baryon-to-meson ratio at intermediate p_T .

Both the Catania and TAMU models are also available for pp collisions, where the calculations are modified to take the different collision environment into account.

5.4 Comparison to models

Figure 5.8 shows the merged Λ_c^+/D^0 ratio measured in central and semi-central Pb–Pb collisions as well as in pp collisions at, $\sqrt{s_{NN}} = 5.02$ TeV, compared to the models discussed in the previous section. The Catania model is not available for semi-central Pb–Pb collisions, and thus is not shown there, while the SHMc model is only available for Pb–Pb and not for pp collisions.

The SHMc model describes the data in semi-central collisions at intermediate and high momentum, but somewhat underpredicts it at low momentum. In central collisions, it significantly underestimates the Λ_c^+/D^0 ratio at intermediate momentum of $4 < p_T < 8$ GeV/c. The Catania model describes the pp data well. However, in central Pb–Pb collisions, it overpredicts the Λ_c^+/D^0 ratio at low momentum and underpredicts it at intermediate momentum. The TAMU model accurately reproduces the Λ_c^+/D^0 ratio in Pb–Pb collisions in both shape and magnitude. It also predicts the results in pp collisions well. The results suggest that hadronisation happens via both fragmentation and coalescence, and that there is a modified hadronisation mechanism for all hadronic systems compared to e^+e^- collisions.

Figure 5.9 shows the experimentally measured nuclear modification factor R_{AA} (see section 1.4), compared to the same three models. The data indicates an enhancement at intermediate p_T , with the Λ_c^+ production suppressed both at low and high p_T in Pb–Pb collisions compared to pp. This is not replicated in the Catania model, which shows a downward trend for R_{AA} with increasing momentum. The SHMc model does replicate the slight peak at intermediate p_T , but overall under-

predicts the data. The TAMU model describes the data better, although it shows a shift towards lower p_T compared to the data in both central and semi-central collisions. Overall this shows a clear need for further study to fully understand hadron production in both pp and p-Pb collisions.

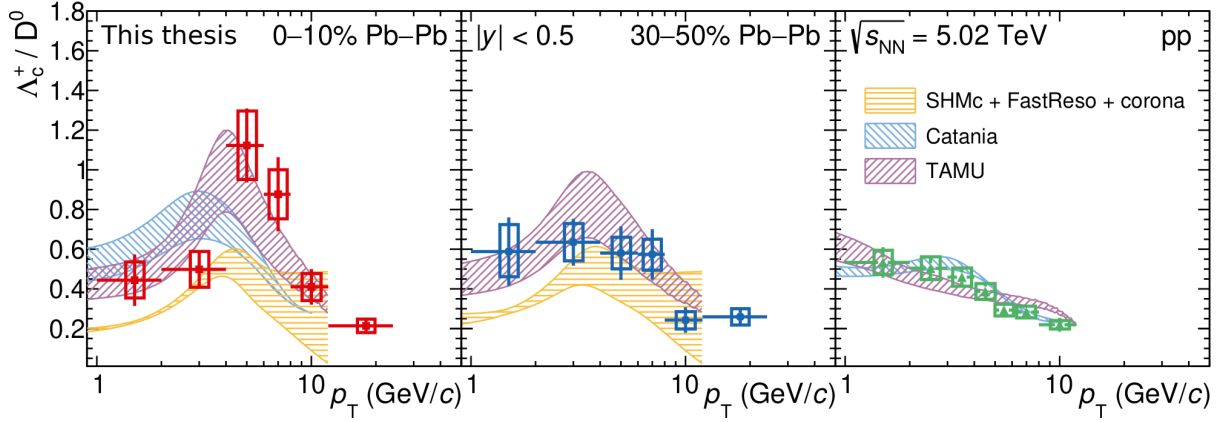


Figure 5.8: The merged Λ_c^+/D^0 ratio in Pb-Pb (central and semi-central) and pp collisions measured at $\sqrt{s_{NN}} = 5.02$ TeV compared to three theoretical models: SHMc [15], Catania [26], and TAMU [27] (see text).

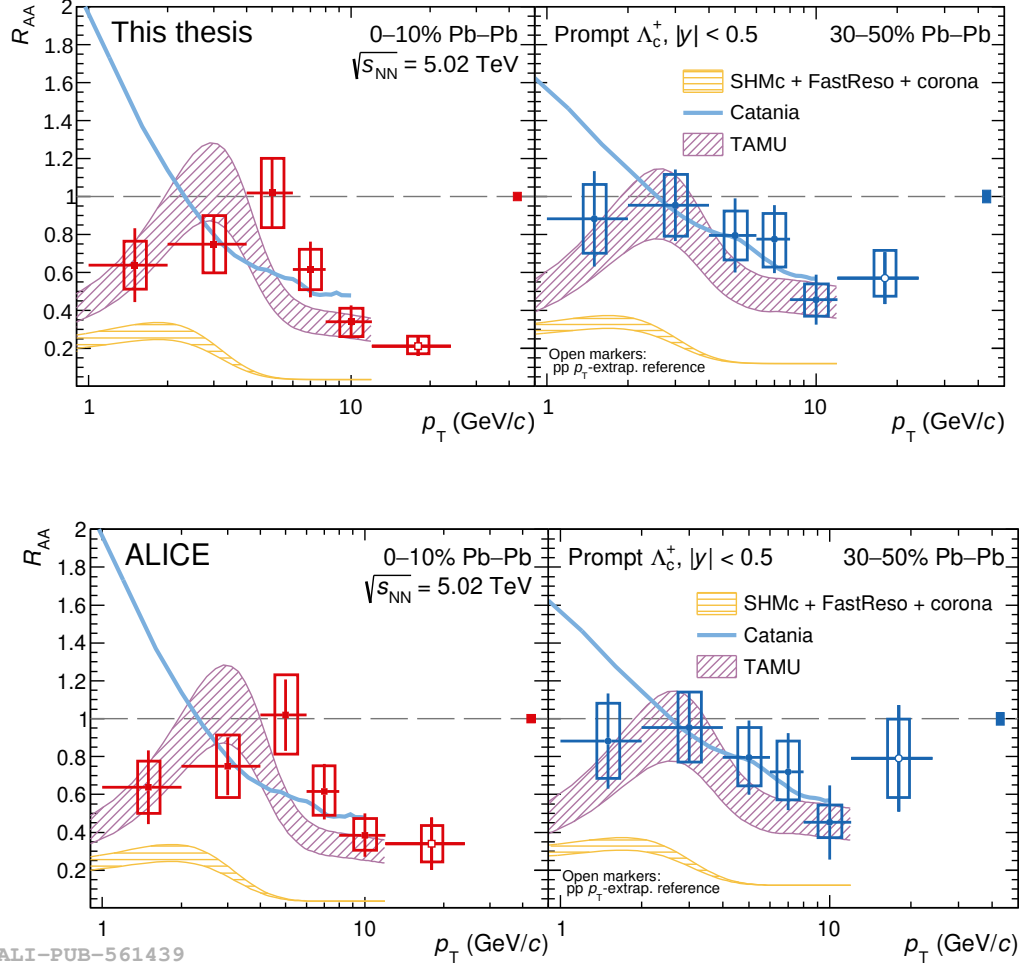


Figure 5.9: Top: The merged nuclear modification factor R_{AA} measured at $\sqrt{s_{NN}} = 5.02$ TeV for central and semi-central events compared to three theoretical models: SHMc [15], Catania [26], and TAMU [27] (see text). The bottom figure shows the same figure from the previously published analysis [23], and the increase in precision at high p_T provided by this work can be clearly observed.

Chapter 6

Conclusion and Outlook

In this thesis, the measurements of the production of the Λ_c^+ baryon in Pb–Pb collisions at a centre-of-mass energy per nucleon–nucleon collision of $\sqrt{s_{\text{NN}}} = 5.02$ TeV with ALICE was presented. The analysis was carried out in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel and separated into two different collisional centrality classes. The p_T -differential production yield was measured in the $8 < p_T < 24$ GeV/ c momentum range in central collisions and $6 < p_T < 24$ GeV/ c in semi-central collisions. The results were compared to those measured in the $\Lambda_c^+ \rightarrow pK_S^0$ channel and merged, weighted by the relative statistical and uncorrelated systematic errors. Using measurements of the D^0 production cross section, the Λ_c^+/D^0 ratio was also measured. The nuclear modification factor R_{AA} was obtained using measurements in pp collisions [25]. Both the Λ_c^+/D^0 ratio and R_{AA} were compared to three theoretical models, namely the Heidelberg SHMc model, the Catania model, and the TAMU model. The results are described well by models that implement hadronisation in the QGP via both fragmentation and coalescence.

The measurement was also attempted in lower momentum bins. However due to the challenging nature of this analysis, including the huge combinatorial background, the three-body decay, and the short lifetime of the Λ_c^+ baryon, a measurement down to lower p_T was beyond the scope of the performance of ALICE for the Run 2 data. It is the first time that Λ_c^+ production has been measured in Pb–Pb collisions using the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay channel, and the results represent an important baseline for

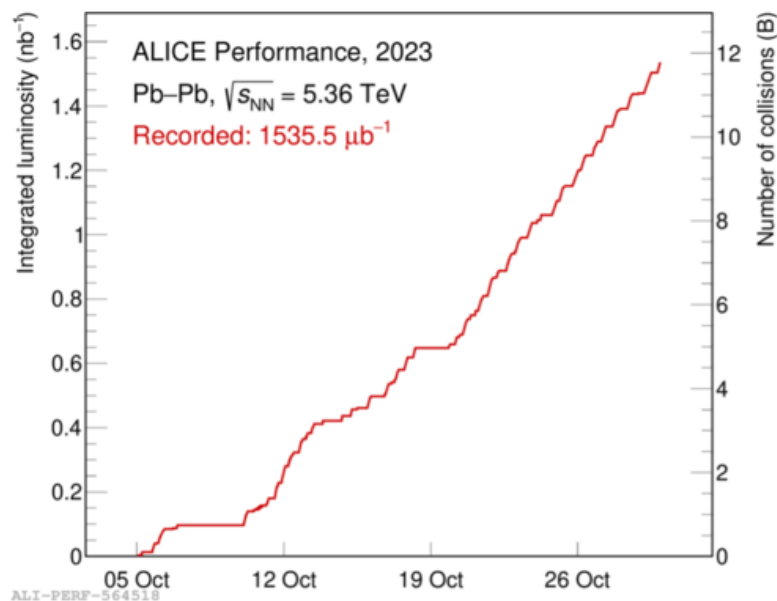


Figure 6.1: The integrated luminosity of the Pb–Pb data taking run at ALICE in October 2023 [28].

future measurements. Furthermore, the theoretical models used here do not include a prediction for momenta above $24 \text{ GeV}/c$, which is where this analysis proves the highest precision. Having access to higher p_T momentum regions and comparing them to these results would provide further insight into the validity of these models.

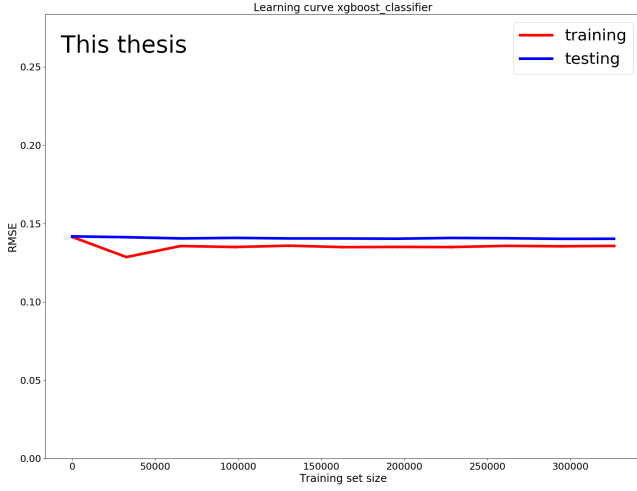
The ALICE detector has since been upgraded for Runs 3 and 4 of the LHC [18], with the first Pb–Pb data taken at the end of 2023. The maximum collision energy reached was $\sqrt{s_{NN}} = 5.36 \text{ TeV}$, the highest energy ever recorded. More importantly, the improved detector layout and readout system led to a dramatic increase in integrated luminosity, as shown in figure 6.1. Over the span of the five-week run, ALICE recorded almost 12 billion Pb–Pb events, which is 40 times more than the total of the previous data taking runs from 2010 to 2018 [28]. The analysis of Run 3 data is expected to constrain the Λ_c^+ measurements further, and deliver a more detailed study of hadronisation mechanisms.

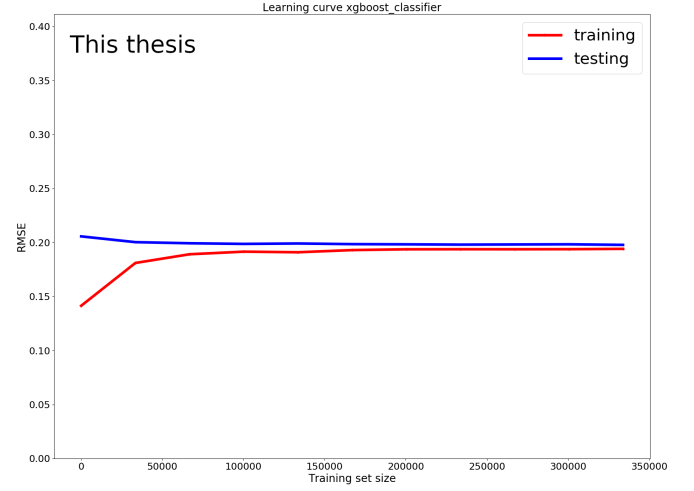
Additional measurements are also planned. This includes studying the flow of Λ_c^+ baryons within the expanding QGP via measuring the elliptic flow v_2 coefficient,

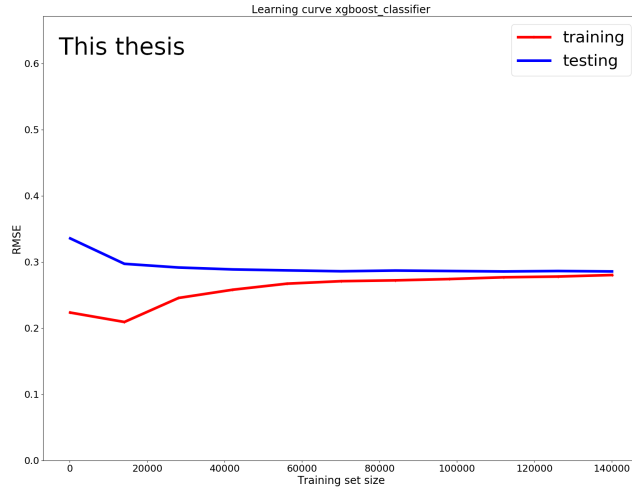
which is expected to provide important information about how they interact in the medium, as well as further constraining hadronisation measurements. Measurements of other charmed baryon productions are also planned, such as the Ξ_c and Σ_c baryons. Both have already been measured in pp collisions, where they exhibit a similar enhancement in relation to charmed mesons when compared to measurements in e^+e^- collisions as was seen for Λ_c^+ baryons (see 2.3.1) [119, 120]. Other measurements may also be possible, such as that of Ω_c , which has also been previously measured in pp collisions [121]. Overall, Run 3 is expected to further increase the theoretical understanding of hadronisation and other QCD processes by a significant amount.

Appendix A

Machine Learning performance plots

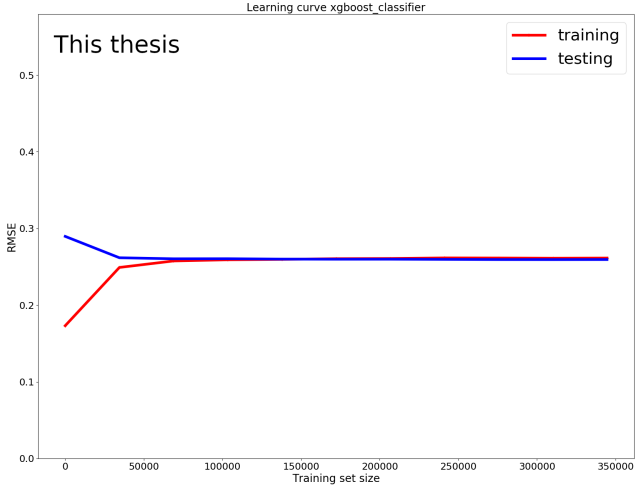


$$6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$$


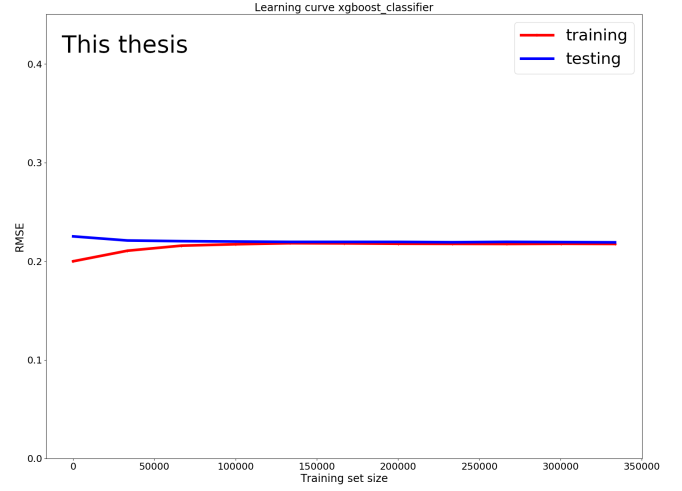
$$8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$$


$$12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$$

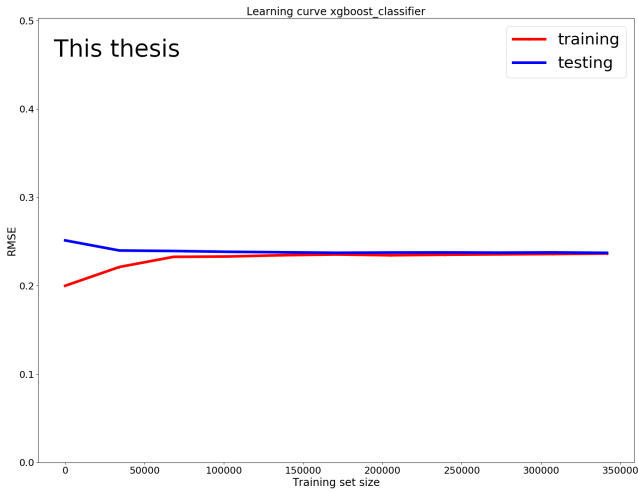
Figure A.1: The learning curve showing the RMSE of the training and testing data sets for central events.



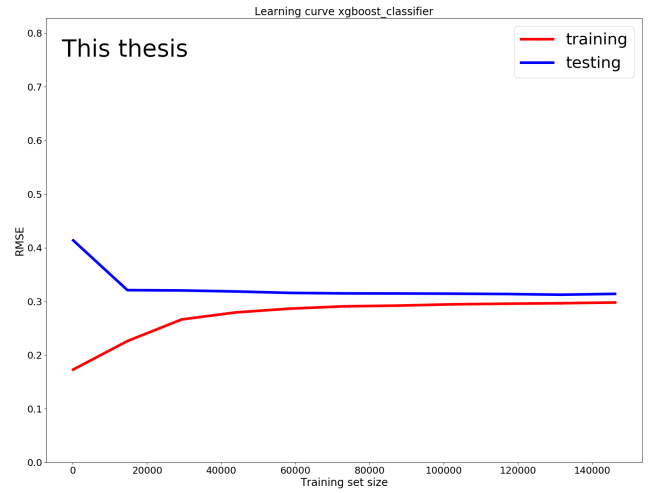
$$4 \text{ GeV}/c < p_T < 6 \text{ GeV}/c$$



$$6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$$

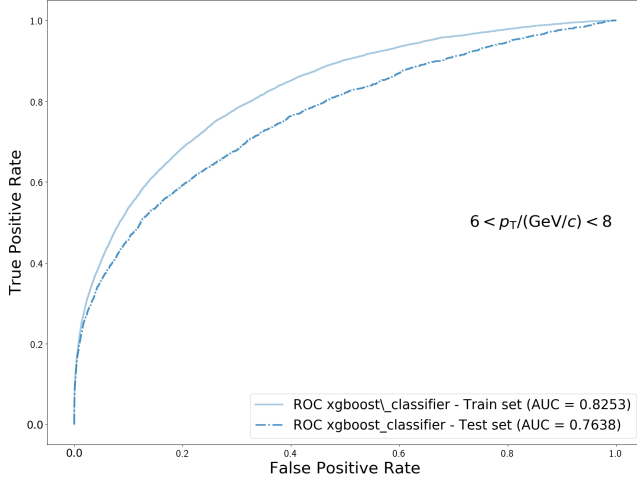


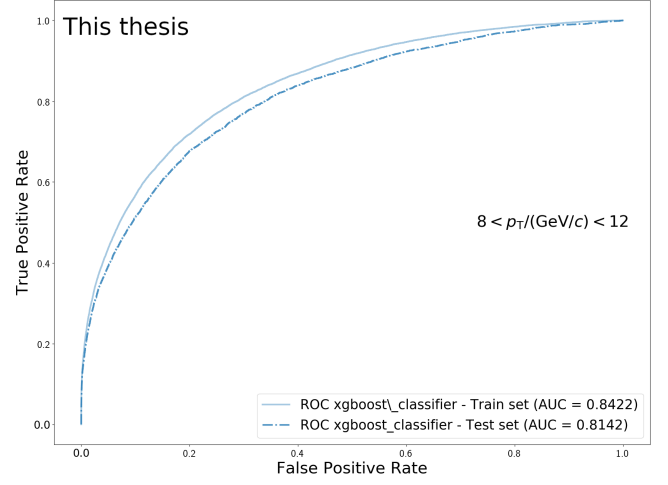
$$8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$$

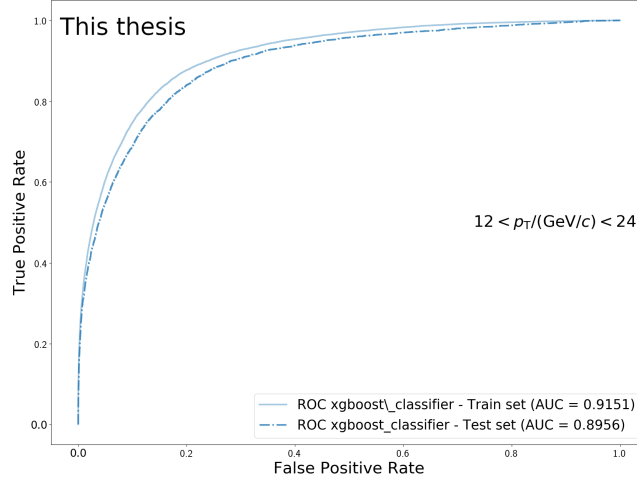


$$12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$$

Figure A.2: The learning curve showing the RMSE of the training and testing data sets for semi-central events.

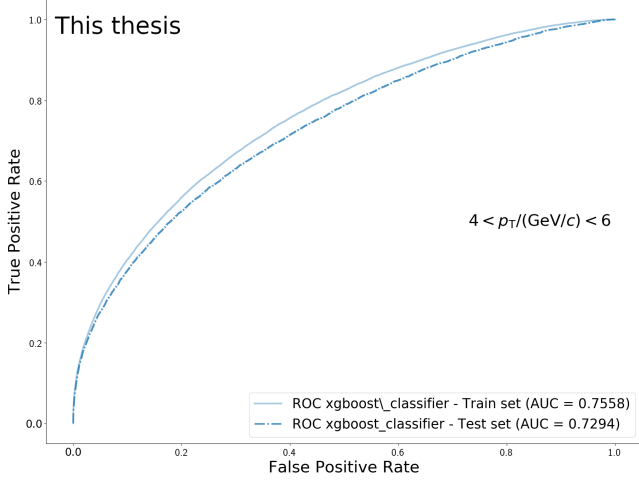


$$6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$$


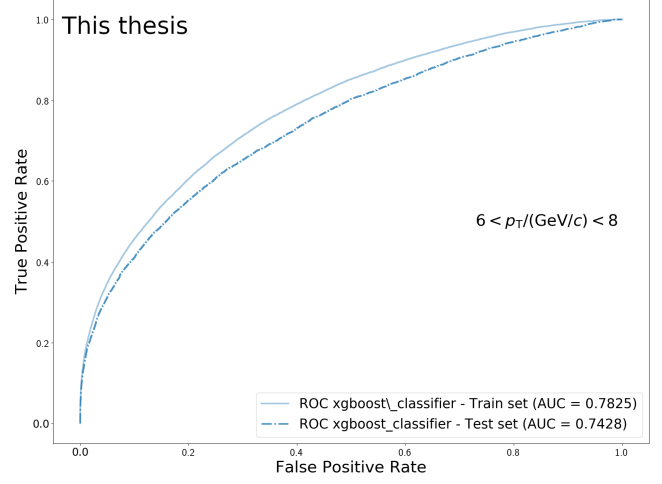
$$8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$$


$$12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$$

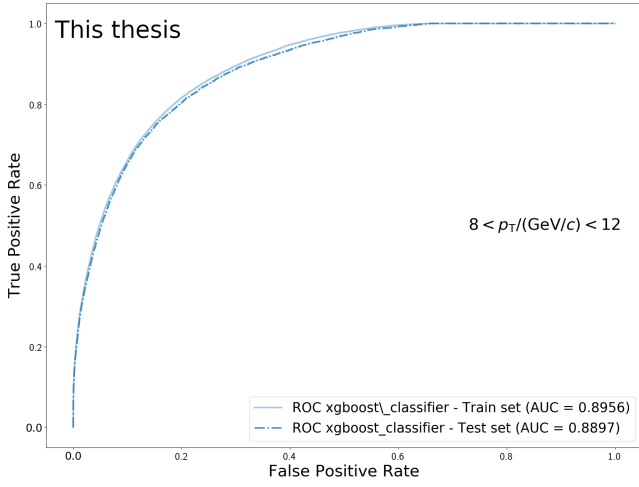
Figure A.3: The ROC curve for the training and testing data set for central events.



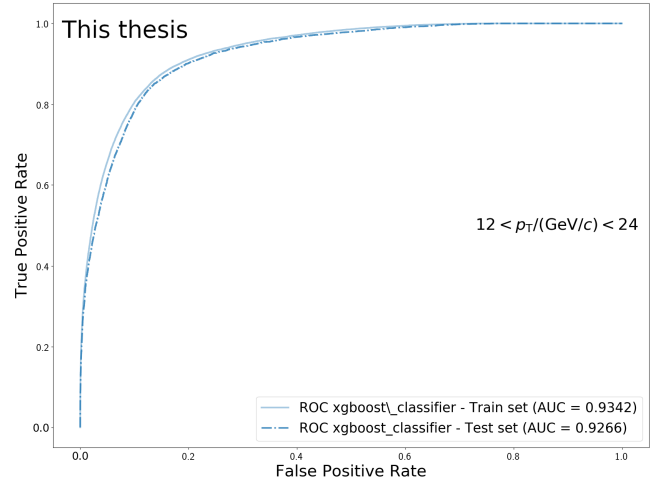
$$4 \text{ GeV}/c < p_T < 6 \text{ GeV}/c$$



$$6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$$



$$8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$$



$$12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$$

Figure A.4: The ROC curve for the training and testing data set for semi-central events.

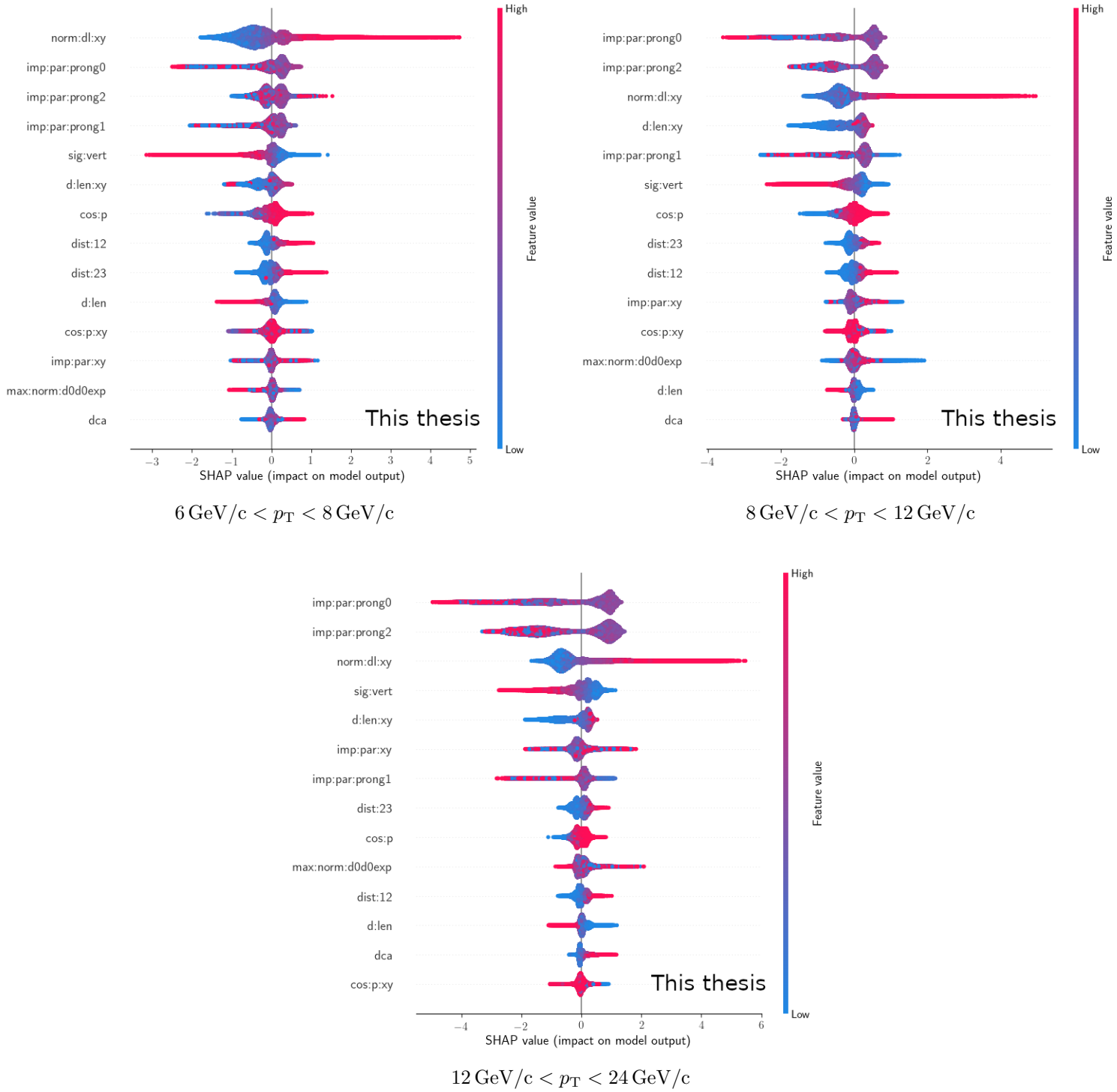


Figure A.5: The impact of the different training variables shown for central events.

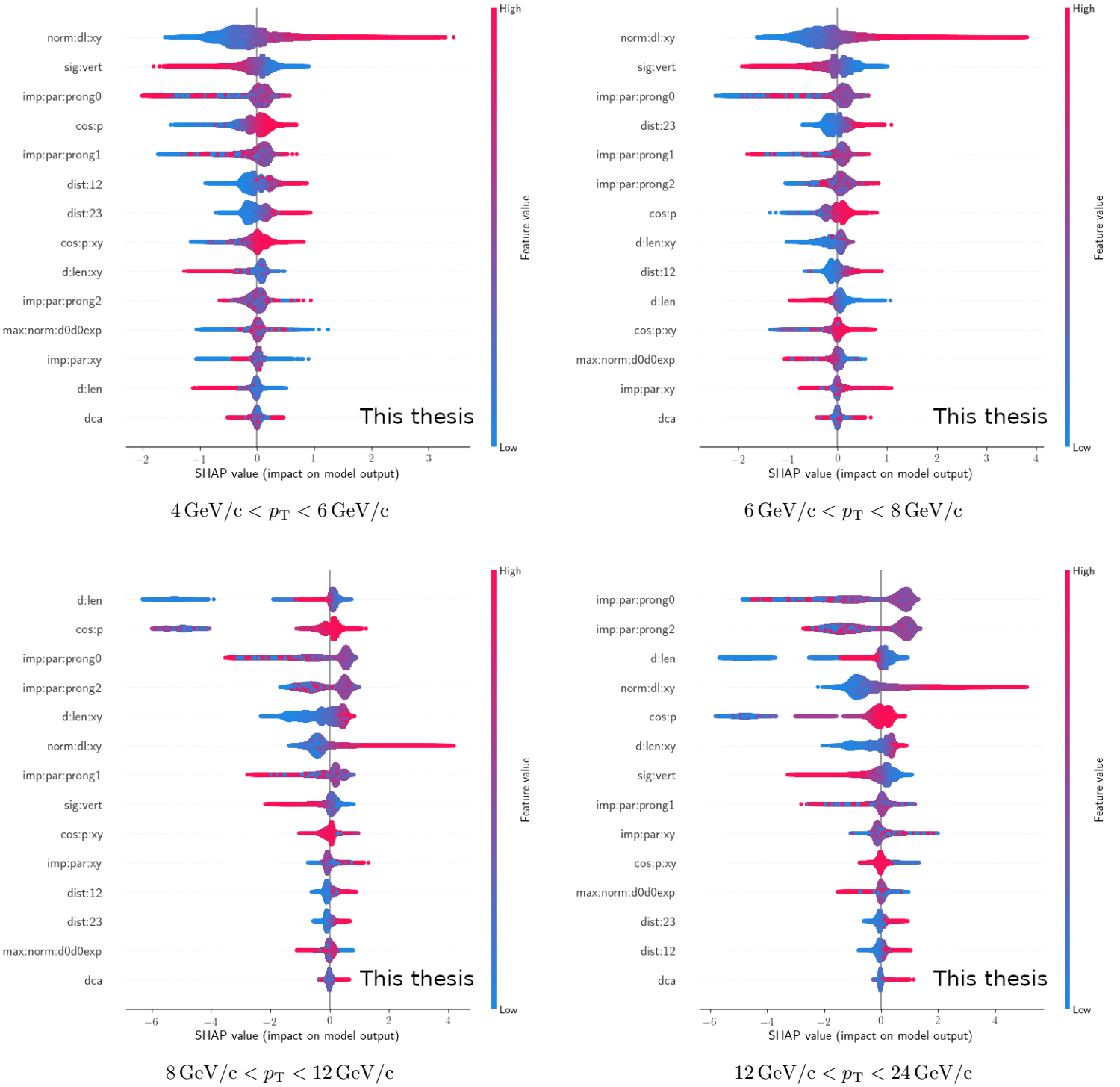


Figure A.6: The impact of the different training variables shown for semi-central events.

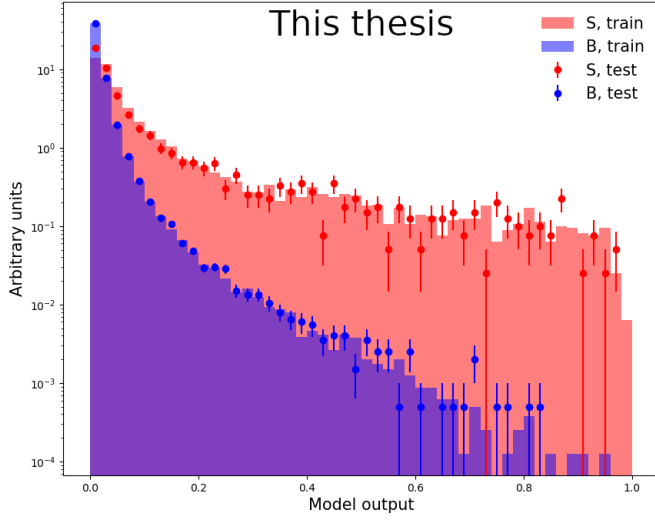
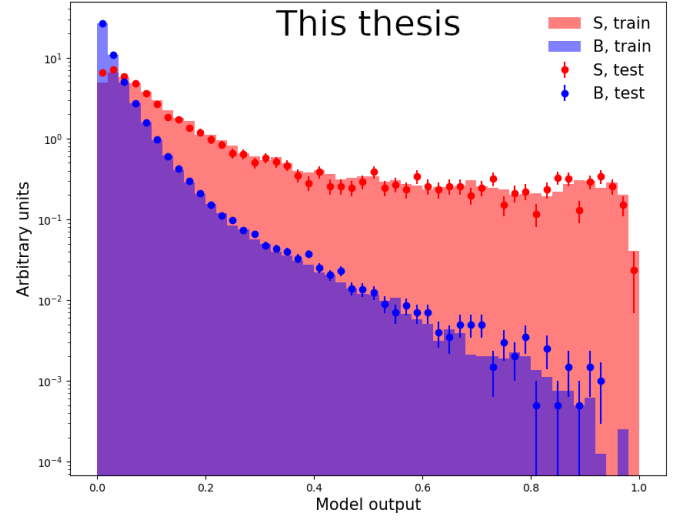
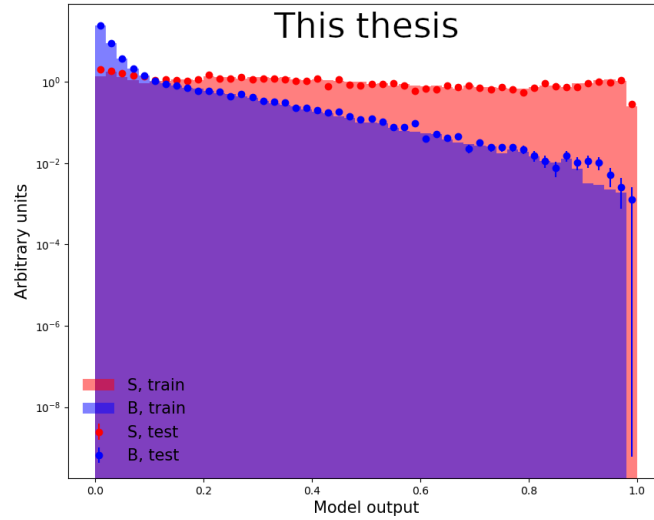
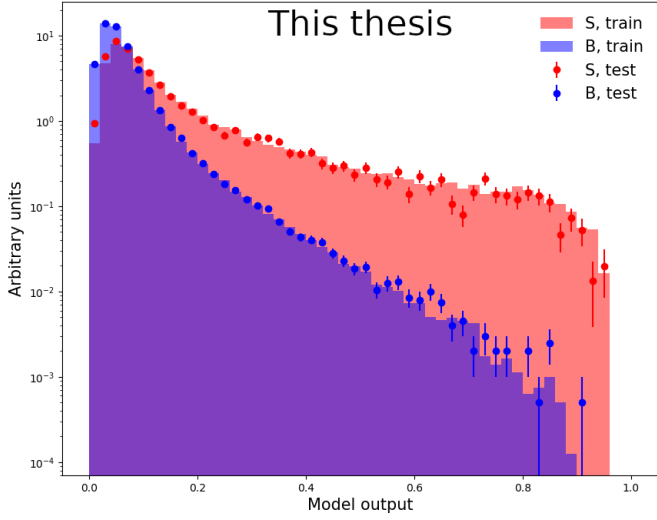
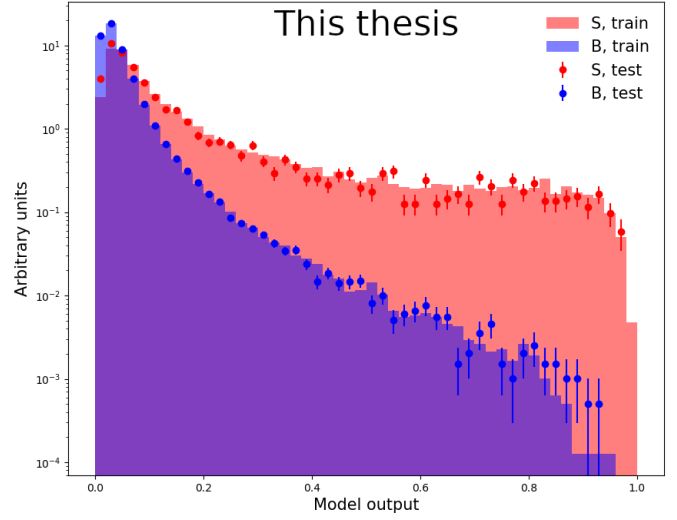

 $6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$

 $8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$

 $12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$

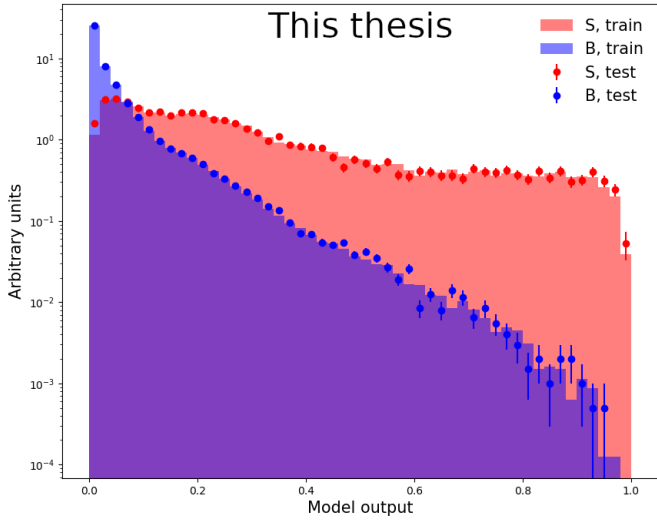
Figure A.7: The model output distribution for central events, showing the separation between signal and background events for training and testing data sets.



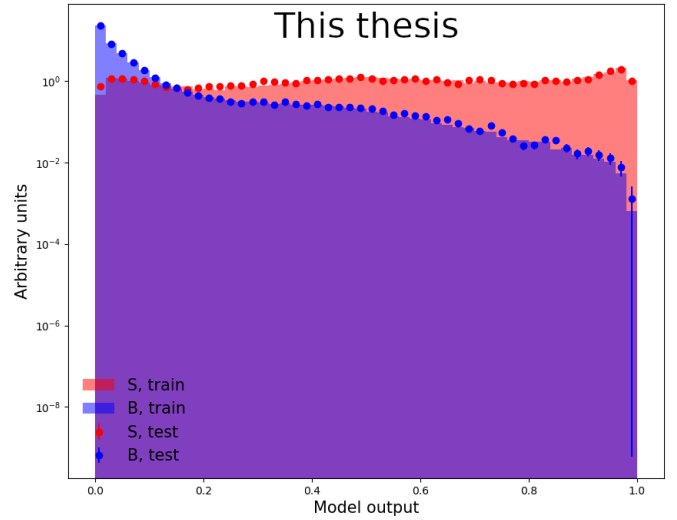
$4 \text{ GeV}/c < p_T < 6 \text{ GeV}/c$



$6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$

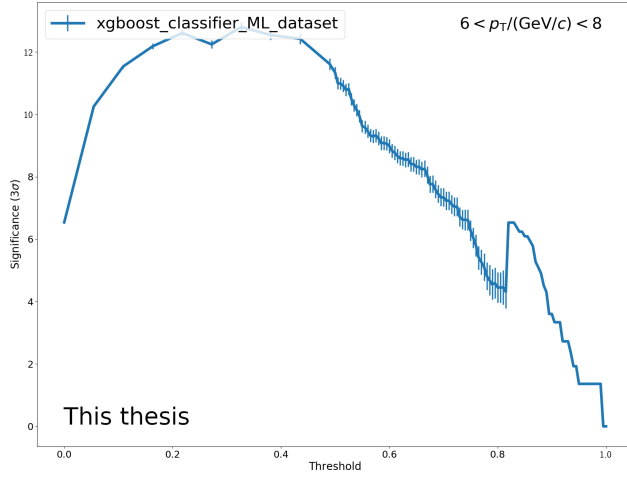


$8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$

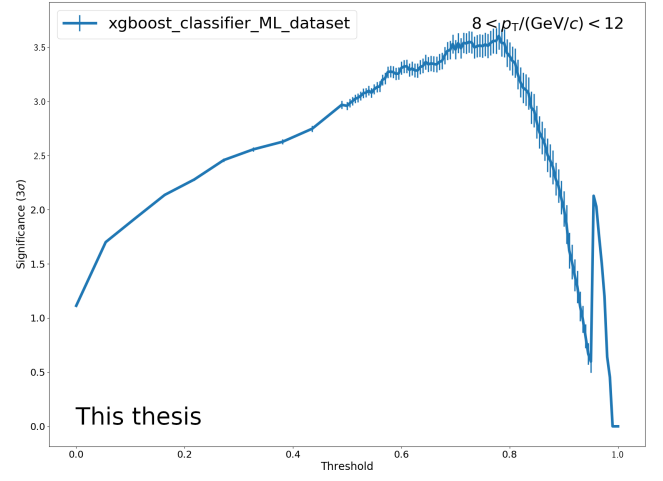


$12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$

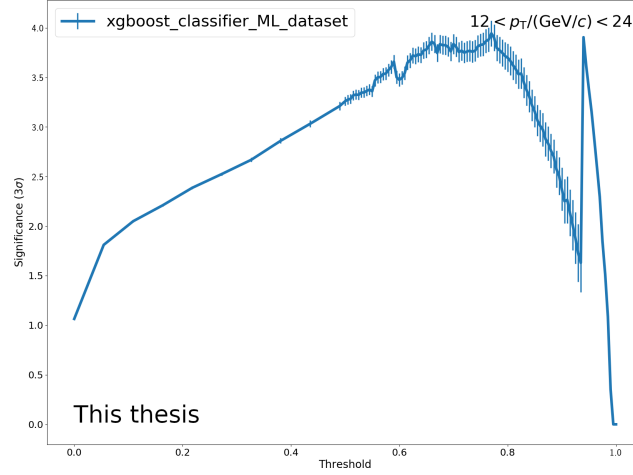
Figure A.8: The model output distribution for semi-central events, showing the separation between signal and background events for training and testing data sets.



$$6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$$

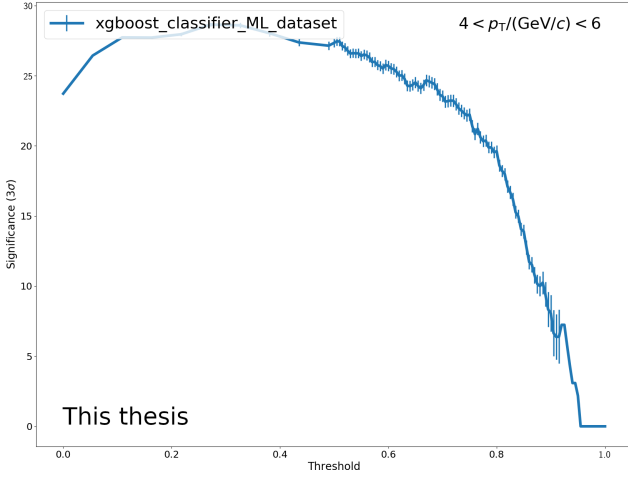


$$8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$$

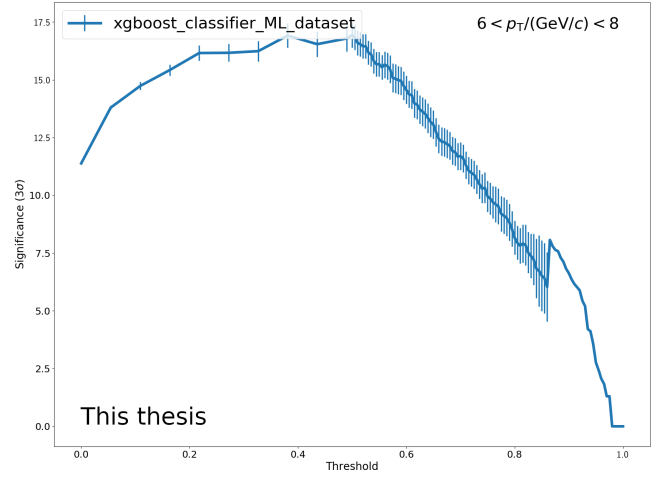


$$12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$$

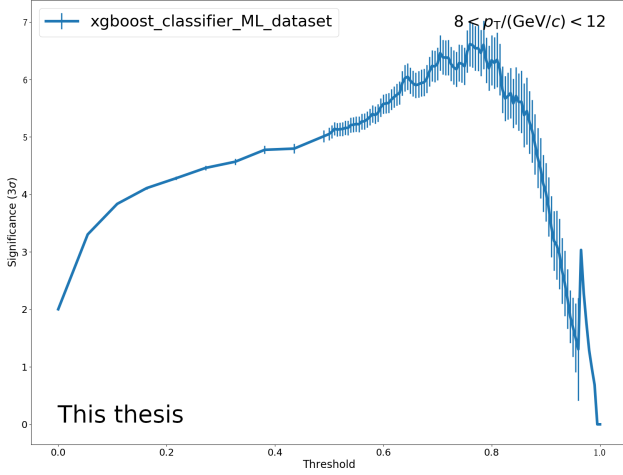
Figure A.9: The significance optimisation plots for central collisions.



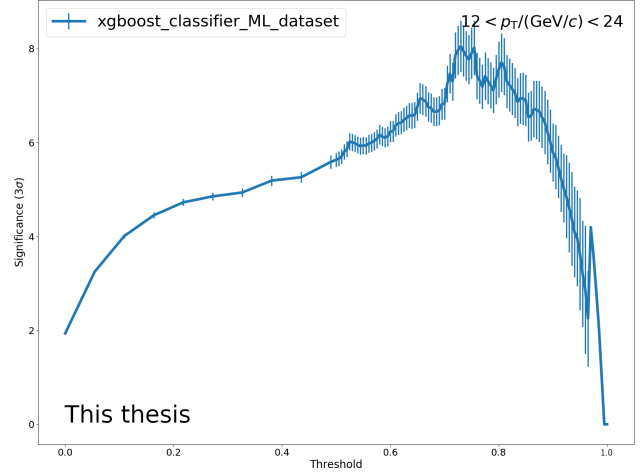
$$4 \text{ GeV}/c < p_T < 6 \text{ GeV}/c$$



$$6 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$$



$$8 \text{ GeV}/c < p_T < 12 \text{ GeV}/c$$



$$12 \text{ GeV}/c < p_T < 24 \text{ GeV}/c$$

Figure A.10: The significance optimisation plots for semi-central collisions.

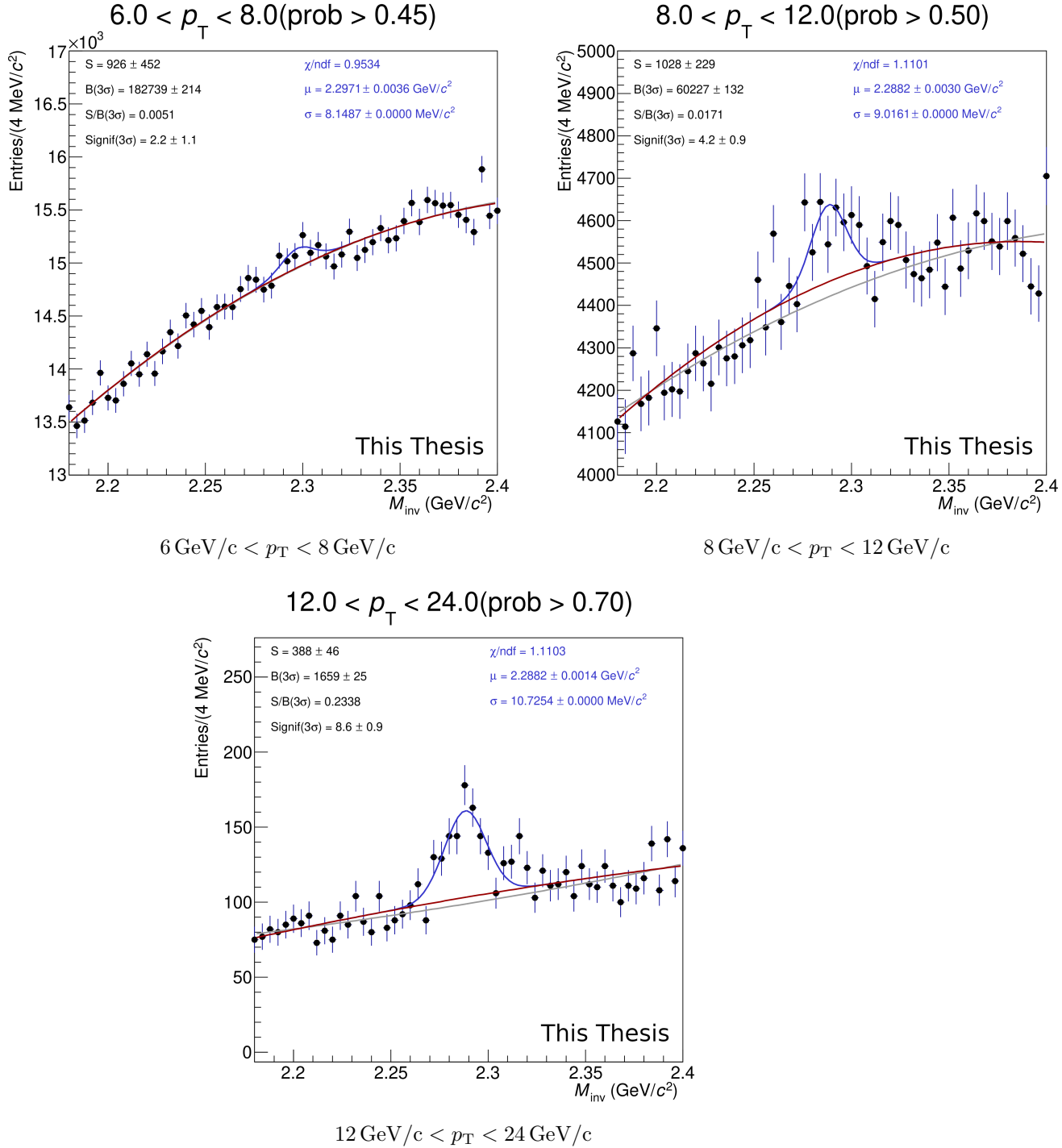


Figure A.11: The invariant mass plots for central collisions with a fit to the Λ_c^+ mass in the three p_T bins.

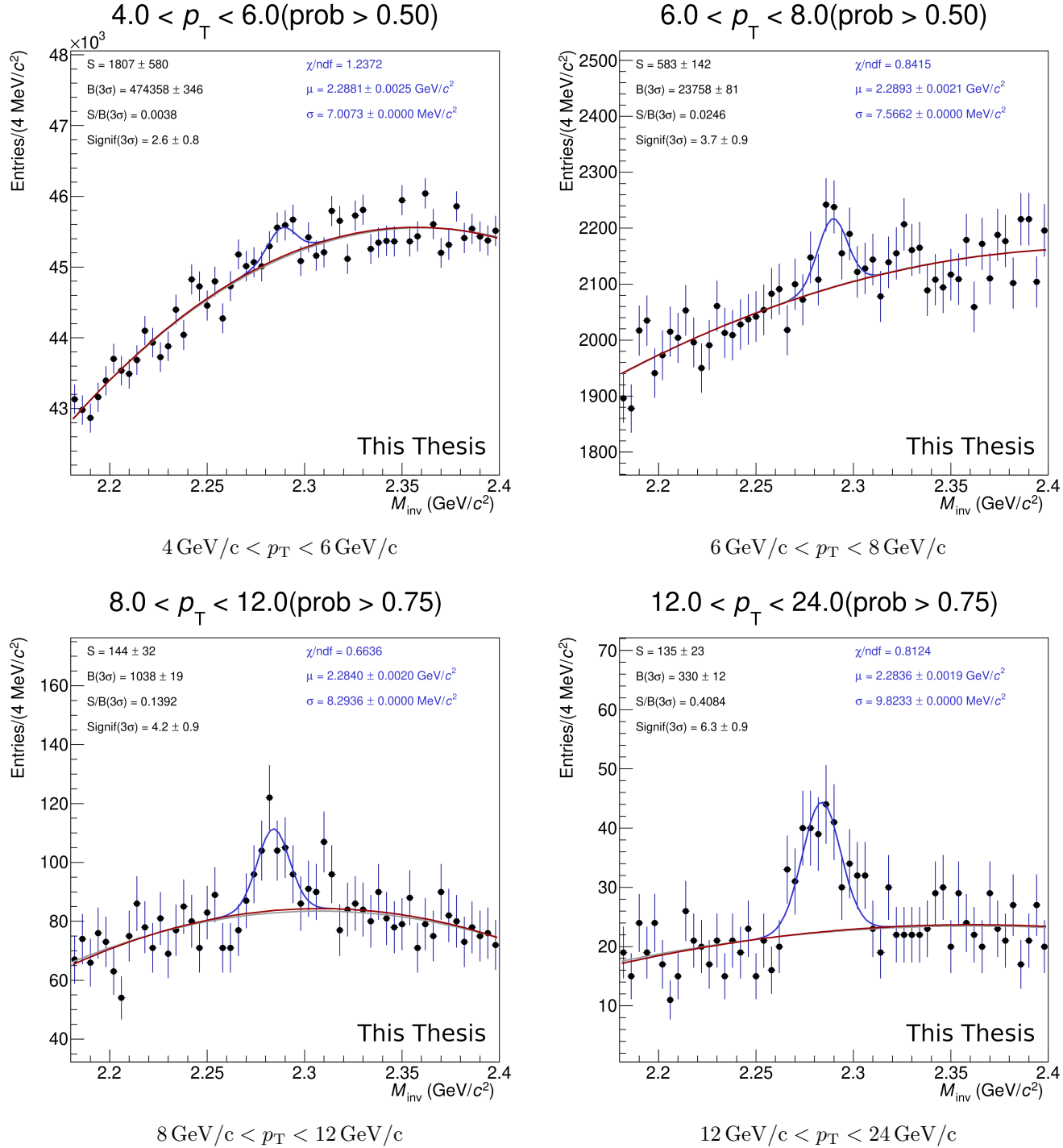


Figure A.12: The invariant mass plots for semi-central collisions with a fit to the Λ_c^+ mass in the three p_T bins.

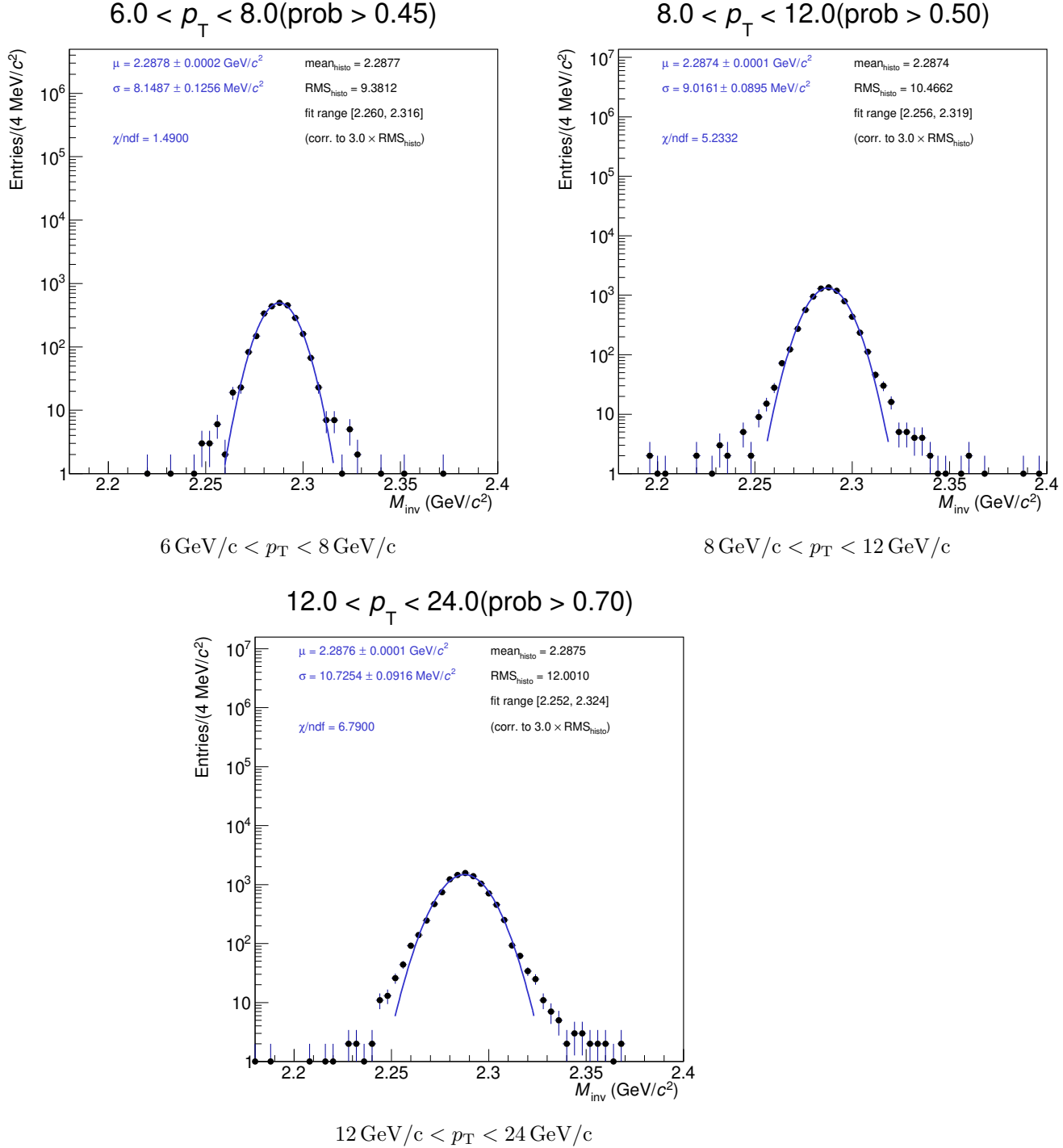


Figure A.13: The invariant mass MC plots for central collisions with a fit to the Λ_c^+ mass in the three p_T bins.

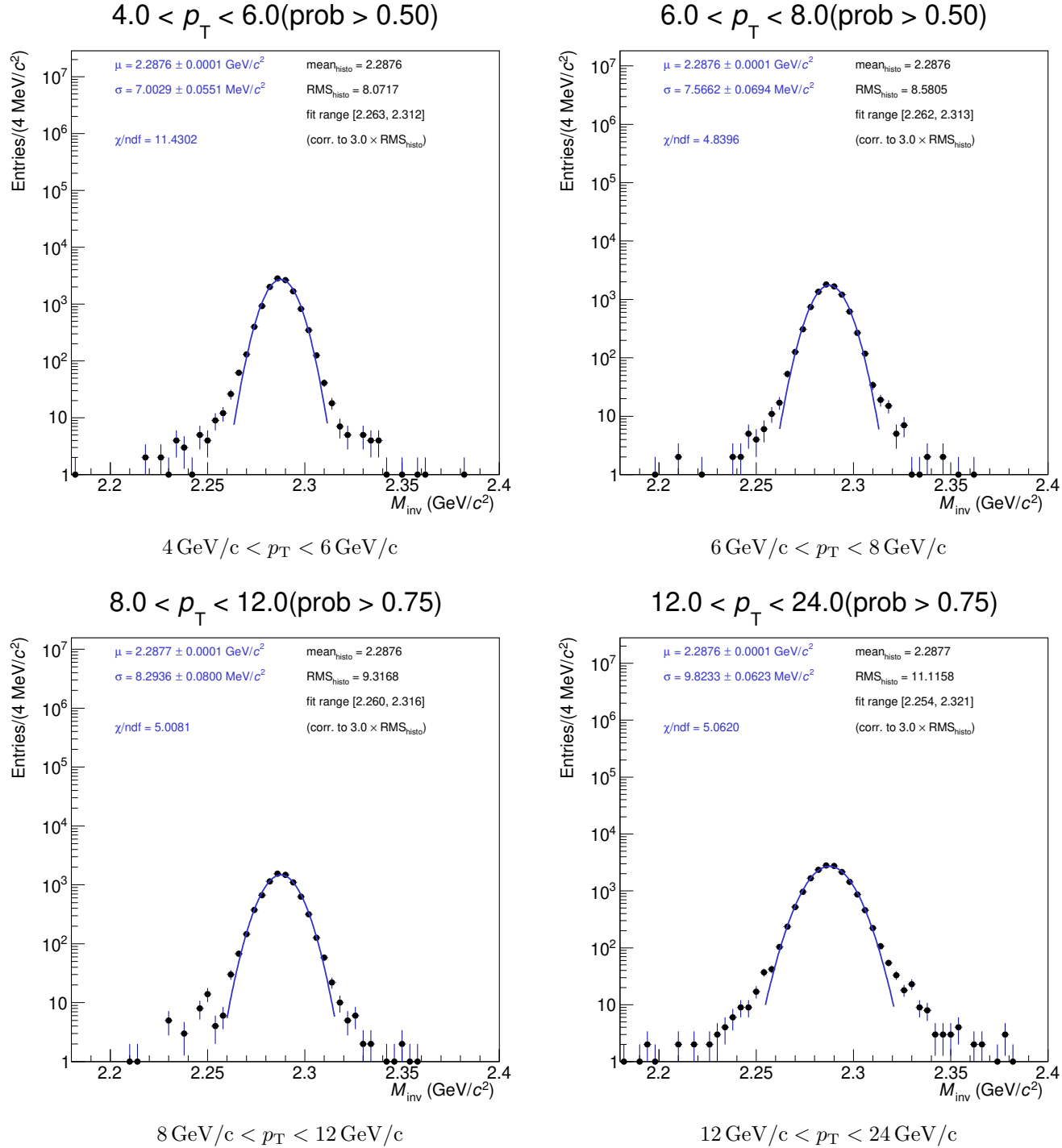


Figure A.14: The invariant mass MC plots for semi-central collisions with a fit to the Λ_c^+ mass in the four p_T bins.

Bibliography

- [1] CERN. <https://home.web.cern.ch/science/physics/standard-model>.
- [2] C. P. et al (Particle Data Group). Review of Particle Physics. *Chin. Phys. C*, 40(10):100001, 2016.
- [3] A. Collaboration. The ALICE experiment – A journey through QCD. 2022. arXiv:2211.04384 [nucl-ex].
- [4] A. B. et al. Equation of state in 2+1-flavor QCD. *Physical Review D*, 90(9), November 2014. 10.1103/physrevd.90.094503.
- [5] F. Karsch et al. Hadron resonance mass spectrum and lattice QCD thermodynamics. *The European Physical Journal C*, 29(4):549–556, August 2003.
- [6] The LHCb Collaboration. Centrality determination in heavy-ion collisions with the LHCb detector. *JINST*, 17(05):P05009, 2022. arXiv:2111.01607.
- [7] The ALICE Collaboration. Transverse momentum spectra and nuclear modification factors of charged particles in pp, p-Pb and Pb-Pb collisions at the LHC. *Journal of High Energy Physics*, 2018(11), November 2018.
- [8] The ALICE Collaboration. K_s^0 and Λ production in Pb–Pb collisions at $\sqrt{s_{NN}}=2.76$ TeV. *Physical Review Letters*, 111(22), November 2013. 10.1103/physrevlett.111.222301.

- [9] The ALICE Collaboration. Λ_c^+ production in pp and p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ eV. *Physical Review C*, 104(5), November 2021. 10.1103/physrevc.104.054905.
- [10] The ALICE Collaboration. Measurement of D^0 , D^+ , D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE. *The European Physical Journal C*, 79(5), May 2019.
- [11] S. Frixione et al. A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *Journal of High Energy Physics*, 2007(09):126–126, September 2007. 10.1088/1126-6708/2007/09/126.
- [12] B. Kniehl et al. $\Lambda^\pm$ production in pp collisions with a new fragmentation function. *Physical Review D*, 101(11), June 2020.
- [13] The ALICE Collaboration. Measurement of D-meson production at mid-rapidity in pp collisions at $\sqrt{s}=7$ TeV. *The European Physical Journal C*, 77(8), August 2017. 10.1140/epjc/s10052-017-5090-4.
- [14] The ALICE Collaboration. Measurement of beauty production via non-prompt D^0 mesons in Pb-Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. *Journal of High Energy Physics*, 2022(12), December 2022. 10.1007/jhep12(2022)126.
- [15] A. A. et al. The multiple-charm hierarchy in the statistical hadronization model. *JHEP*, 7:035, 2021. arXiv:2104.12754 [hep-ph].
- [16] CERN. CERN’s acellerator complex. <https://cds.cern.ch/images/CERN-GRAPHICS-2019-002-1>, Dec 2019. Accessed on 2021-10-08.
- [17] A. Tauro. 3D ALICE Schematic RUN2 . <https://alice-figure.web.cern.ch/node/11219>, May 2017. Accessed on 2023-09-11.
- [18] The ALICE Collaboration. ALICE upgrades during the LHC Long Shutdown 2. 2023. arXiv:2302.01238 [physics.ins-det].

- [19] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, August 2008. <https://dx.doi.org/10.1088/1748-0221/3/08/S08002>.
- [20] The ALICE Collaboration. The ALICE figure repository, 2016. <https://alice-figure.web.cern.ch/>.
- [21] F. Carnesecchi. Performance of the ALICE Time-Of-Flight detector at the LHC. *Journal of Instrumentation*, 14(06):C06023–C06023, jun 2019. <http://dx.doi.org/10.1088/1748-0221/14/06/C06023>.
- [22] The ALICE Collaboration. Centrality determination of Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV with ALICE. *Physical Review C*, 88(4), October 2013. <https://doi.org/10.1103/PhysRevC.88.044909>.
- [23] The ALICE Collaboration. Constraining hadronization mechanisms with Λ_c^+/D^0 production ratios in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. *Phys. Lett. B*, 839, 2023. arXiv:2112.08156 [nucl-ex].
- [24] The ALICE Collaboration. Prompt D^0 , D^+ , and D^{*+} production in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. 2022. arXiv:2110.09420 [nucl-ex].
- [25] The ALICE Collaboration. First measurement of Λ_c^+ production down to $p_{\text{T}} = 0$ in pp and p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. *Phys. Rev. C*, 107, 2023. arXiv:2211.14032 [nucl-ex].
- [26] S. P. et al. Charmed hadrons from coalescence plus fragmentation in relativistic nucleus–nucleus collisions at RHIC and LHC. *Eur. Phys. J. C*, 78:348, 2018. arXiv:1712.00730 [hep-ph].
- [27] M. He and R. Rapp. Hadronization and charm-hadron ratios in heavy-ion collisions. *Phys. Rev. Lett.*, 124, 2020. arXiv:1905.09216 [nucl-th].
- [28] The ALICE Collaboration. ALICE bags about twelve billion heavy-ion collisions, 2023. https://alice-collaboration.web.cern.ch/2023_Dec_heavy_ion.

-
- [29] G. Contin. Performance of the present ALICE Inner Tracking System and studies for the upgrade. *Journal of Instrumentation*, 7(06):C06007, June 2012. <https://dx.doi.org/10.1088/1748-0221/7/06/C06007>.
- [30] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, September 2012.
- [31] The CMS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, September 2012. 10.1016/j.physletb.2012.08.020.
- [32] F. Wilczek. Quantum field theory. *Rev. Mod. Phys.*, 71:S85–S95, March 1999. <https://link.aps.org/doi/10.1103/RevModPhys.71.S85>.
- [33] H. D. Politzer. Reliable Perturbative Results for Strong Interactions? *Phys. Rev. Lett.*, 30:1346–1349, June 1973. <https://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [34] J. Greensite. *An Introduction to the Confinement Problem*, volume 821. Springer, 2011.
- [35] V. Koch. Aspects of Chiral Symmetry. *International Journal of Modern Physics E*, 06(02):203–249, 1997. <https://doi.org/10.1142/S0218301397000147>.
- [36] S. Muroya et al. Lattice QCD at Finite Density: An Introductory Review. *Progress of Theoretical Physics*, 110(4):615–668, October 2003. 10.1143/ptp.110.615.
- [37] C. Ratti. Lattice QCD and heavy ion collisions: a review of recent progress. *Reports on Progress in Physics*, 81(8):084301, July 2018. 10.1088/1361-6633/aabb97.

- [38] C. Burgess. An Introduction to Effective Field Theory. *Annual Review of Nuclear and Particle Science*, 57(1):329–362, November 2007. <https://doi.org/10.1146>
- [39] T. O’Connor. Emergent Properties. *American Philosophical Quarterly*, 31(2):91–104, 1994.
- [40] N. Cabibbo and G. Parisi. Exponential hadronic spectrum and quark liberation. *Physics Letters B*, 59(1):67–69, 1975. <https://www.sciencedirect.com/science/article/pii/0370269375901586>.
- [41] E. Annala et al. Evidence for quark-matter cores in massive neutron stars. *Nature Physics*, 16(9):907–910, June 2020. <https://doi.org/10.1038/2Fs41567-020-0914-9>.
- [42] J. R. Oppenheimer and G. M. Volkoff. On Massive Neutron Cores. *Phys. Rev.*, 55:374–381, February 1939. 10.1103/PhysRev.55.37474.
- [43] Y. B. Zeldovich. The equation of state at ultrahigh densities and its relativistic limitations. *Zh. Eksp. Teor. Fiz.*, 41:1609–1615, 1961.
- [44] G. Baym and S. Chin. Can a neutron star be a giant MIT bag? *Physics Letters B*, 62(2):241–244, 1976.
- [45] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214–215, 1964. 10.1016/S0031-9163(64)92001-3.
- [46] G. Zweig. An SU_3 model for strong interaction symmetry and its breaking; Version 2. 1964. 10.17181/CERN-TH-412”.
- [47] E. D. Bloom et al. High-Energy Inelastic $e - p$ Scattering at 6° and 10° . *Phys. Rev. Lett.*, 23:930–934, October 1969. 10.1103/PhysRevLett.23.930.
- [48] M. Breidenbach et al. Observed Behavior of Highly Inelastic Electron-Proton Scattering. *Phys. Rev. Lett.*, 23:935–939, October 1969. 10.1103/PhysRevLett.23.935.

- [49] D. J. Gross and F. Wilczek. Ultraviolet Behavior of Non-Abelian Gauge Theories. *Phys. Rev. Lett.*, 30:1343–1346, June 1973. 10.1103/PhysRevLett.30.1343.
- [50] J. C. Collins and M. J. Perry. Superdense Matter: Neutrons or Asymptotically Free Quarks? *Phys. Rev. Lett.*, 34:1353–1356, May 1975. 10.1103/PhysRevLett.34.1353.
- [51] J. Rafelski and B. Müller. Strangeness Production in the Quark-Gluon Plasma. *Phys. Rev. Lett.*, 48:1066–1069, April 1982. 10.1103/PhysRevLett.48.1066.
- [52] T. Matsui and H. Satz. J/ψ suppression by quark-gluon plasma formation. *Physics Letters B*, 178(4):416–422, 1986.
- [53] G. E. Brown and M. Rho. Scaling effective Lagrangians in a dense medium. *Phys. Rev. Lett.*, 66:2720–2723, May 1991. 10.1103/PhysRevLett.66.2720.
- [54] J. Kapusta et al. High-energy photons from quark-gluon plasma versus hot hadronic gas. *Phys. Rev. D*, 44:2774–2788, November 1991. 10.1103/PhysRevD.44.2774.
- [55] P. Ruuskanen. Electromagnetic probes of quark-gluon plasma in relativistic heavy-ion collisions. *Nuclear Physics A*, 544(1):169–182, 1992.
- [56] C. Ogilvie. Review of nuclear reactions at the AGS. *Nuclear Physics A*, 698(1–4):3–12, February 2002. 10.1016/S0375-9474(01)01343-4.
- [57] U. Heinz and M. Jacob. Evidence for a New State of Matter: An Assessment of the Results from the CERN Lead Beam Programme, 2000.
- [58] The NA50 Collaboration. Evidence for deconfinement of quarks and gluons from the J/ψ suppression pattern measured in Pb–Pb collisions at the CERN-SPS. *Physics Letters B*, 477(1):28–36, 2000. <https://www.sciencedirect.com/science/article/pii/S0370269300002379>.

- [59] The WA97 Collaboration. Strangeness enhancement at mid-rapidity in Pb–Pb collisions at 158 at GeV/c. *Physics Letters B*, 449(3):401–406, 1999. <https://www.sciencedirect.com/science/article/pii/S0370269399001409>.
- [60] CERES/NA45 Collaboration. Low-mass e^+e^- pair production in 158 A GeV Pb–Au collisions at the CERN SPS, its dependence on multiplicity and transverse momentum. *Physics Letters B*, 422(1-4):405–412, March 1998. <https://doi.org/10.1016%2Fs0370-2693%2898%2900083-5>.
- [61] The NA49 Collaboration . Energy dependence of pion and kaon production in central Pb–Pb collisions. *Physical Review C*, 66(5), November 2002.
- [62] M. Harrison et al. RHIC project overview. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 499(2):235–244, 2003. The Relativistic Heavy Ion Collider Project: RHIC and its Detectors.
- [63] The PHENIX Collaboration. Suppression of Hadrons with Large Transverse Momentum in Central Au–Au collisions at $\sqrt{s_{\text{NN}}} = 130$ GeV. *Physical Review Letters*, 88(2), December 2001. 10.1103/physrevlett.88.022301.
- [64] E. Shuryak. Why does the quark–gluon plasma at RHIC behave as a nearly ideal fluid? *Progress in Particle and Nuclear Physics*, 53(1):273–303, July 2004. 10.1016/j.ppnp.2004.02.025.
- [65] S. Borsányi et al. Full result for the QCD equation of state with 2+1 flavors. *Physics Letters B*, 730:99–104, 2014. <https://doi.org/10.1016%2Fj.physletb.2014.01.007>.
- [66] P. F. Kolb and U. Heinz. Hydrodynamic description of ultrarelativistic heavy-ion collisions. 2003. arXiv:nucl-th/0305084.
- [67] V. Minissale et al. Coalescence plus fragmentation approach for the hadronization mechanism of heavy hadrons from AA to pp collisions. *EPJ Web Conf.*, 276:02013, 2023. 10.1051/epjconf/202327602013.

- [68] M. L. Miller et al. Glauber Modeling in High-Energy Nuclear Collisions. *Annual Review of Nuclear and Particle Science*, 57(1):205–243, November 2007. <https://doi.org/10.1146%2Fannurev.nucl.57.090506.123020>.
- [69] C. Bierlich et al. Diffractive and non-diffractive wounded nucleons and final states in pA collisions. *Journal of High Energy Physics*, 2016(10), October 2016. 10.1007/jhep10(2016)139.
- [70] Andronic, A. et al. Heavy-flavour and quarkonium production in the LHC era: from proton–proton to heavy-ion collisions. *The European Physical Journal C*, 76(3), February 2016.
- [71] V. A. . Z. O. Lisovyi, M. Combined analysis of charm-quark fragmentation-fraction measurements. *Eur. Phys. J. C*, 76, 2016. <https://doi.org/10.1140/epjc/s10052-016-4246-y>.
- [72] R. R. H. Prino. Open Heavy Flavour QCD Matter and in Nuclear Collisions. *J. Physics G*, 43, 2016.
- [73] D. T. G. D. Moore. How Much do Heavy Quarks Thermalize in a Heavy Ion Collision? *Phys.Rev. C*71, 064904, 2005. <https://doi.org/10.1103/PhysRevC.71.064904>.
- [74] U. A. W. C. A. Salgado. Calculating Quenching Weights. *Phys.Rev. D*68, 014008, 2003. <https://doi.org/10.1103/PhysRevD.68.014008>.
- [75] Y. L. Dokshitzer et al. On specific QCD properties of heavy quark fragmentation (‘dead cone’). *Journal of Physics G: Nuclear and Particle Physics*, 17(10):1602, October 1991. 10.1088/0954-3899/17/10/023.
- [76] The ALICE Collaboration. Direct observation of the dead-cone effect in quantum chromodynamics . *Nature*, 605:440–446, 2022.
- [77] G. Luparello. ALICE measurements of heavy-flavour production in pp and p–Pb collisions at the LHC. *Nuclear and Particle Physics Proceedings*, 289-290:277–

- 280, 2017. 8th International Conference on Hard and Electromagnetic Probes of High Energy Nuclear Collisions.
- [78] The ALICE Collaboration. Λ_c^+ production and baryon-to-meson ratios in pp and p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ eV}$ at the LHC. *Physical Review Letters*, 127(20), November 2021. 10.1103/physrevlett.127.202301.
- [79] S. P. Christiansen, J.R. String formation beyond leading colour. *J. High Energ. Phys.*, 2015. [https://doi.org/10.1007/JHEP08\(2015\)003](https://doi.org/10.1007/JHEP08(2015)003).
- [80] A. Beraudo et al. Heavy-flavour production in high-energy d–Au and p–Pb collisions. *Journal of High Energy Physics*, 2016(3), March 2016.
- [81] The ALICE Collaboration. Measurement of prompt D_s^+ -meson production and azimuthal anisotropy in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$. *Physics Letters B*, 827:136986, April 2022. 10.1016/j.physletb.2022.136986.
- [82] The ALICE Collaboration. Measurements of inclusive J/ψ production at midrapidity and forward rapidity in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$, 2023.
- [83] L. Evans and P. Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001, August 2008. <https://dx.doi.org/10.1088/1748-0221/3/08/S08001>.
- [84] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, August 2008. <https://dx.doi.org/10.1088/1748-0221/3/08/S08003>.
- [85] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, August 2008. <https://dx.doi.org/10.1088/1748-0221/3/08/S08004>.
- [86] The LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, August 2008. <https://dx.doi.org/10.1088/1748-0221/3/08/S08005>.

-
- [87] LEP Experiments Committee. Technical proposal: L3. Technical report, 1983. <https://cds.cern.ch/record/297266>.
 - [88] S. Myers. *The LEP Collider, from design to approval and commissioning*. John Adams' memorial lecture. CERN, Geneva, 1991. Delivered at CERN, 26 Nov 1990.
 - [89] The ALICE Collaboration. *ALICE Inner Tracking System (ITS): Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 1999. <https://cds.cern.ch/record/391175>.
 - [90] THE ALICE Collaboration. *ALICE time projection chamber: Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 2000. <https://cds.cern.ch/record/451098>.
 - [91] P. Cortese. *ALICE transition-radiation detector: Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 2001. <https://cds.cern.ch/record/519145>.
 - [92] *ALICE Time-Of-Flight system (TOF): Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 2000. <https://cds.cern.ch/record/430132>.
 - [93] F. Piuz et al. *ALICE high-momentum particle identification: Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 1998. <https://cds.cern.ch/record/381431>.
 - [94] V. I. M. et al. *ALICE Photon Spectrometer (PHOS): Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 1999. <https://cds.cern.ch/record/381432>.
 - [95] The ALICE Collaboration. *ALICE Electromagnetic Calorimeter Technical Design Report*. Technical report, 2008. <https://cds.cern.ch/record/1121574>.

- [96] J. Allen and C. B. et al. ALICE DCal: An Addendum to the EMCal Technical Design Report Di-Jet and Hadron-Jet correlation measurements in ALICE. Technical report, 2010. <https://cds.cern.ch/record/1272952>.
- [97] M. Gallio et al. *ALICE Zero-Degree Calorimeter (ZDC): Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 1999. <https://cds.cern.ch/record/381433>.
- [98] *ALICE Photon Multiplicity Detector (PMD): Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 1999. <https://cds.cern.ch/record/451099>.
- [99] The ALICE Collaboration. *ALICE forward detectors: FMD, TO and VO: Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 2004. <https://cds.cern.ch/record/781854>.
- [100] A. Fernández et al. ACORDE a cosmic ray detector for ALICE. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 572(1):102–103, 2007. Frontier Detectors for Frontier Physics.
- [101] The ALICE Collaboration. Technical Design Report for the Upgrade of the ALICE Inner Tracking System. Technical report, 2014. <https://cds.cern.ch/record/1625842>.
- [102] Y. A. e. a. J. Alme. The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 622(1):316–367, October 2010. <https://doi.org/10.1016%2Fj.nima.2010.04.042>.
- [103] W. B. L. Rolandi, W. Riegler. *Particle Detection with Drift Chambers*. Springer, 2008.

-
- [104] The ALICE Collaboration. Determination of the event collision time with the ALICE detector at the LHC. *The European Physical Journal Plus*, 132(2), feb 2017. <http://dx.doi.org/10.1140/epjp/i2017-11279-1>.
- [105] T. A. Collaboration. Particle identification in ALICE: a Bayesian approach. *Eur. Phys. J. Plus*, 131(168), 2016. <https://doi.org/10.1140/epjp/i2016-16168-5>.
- [106] P. Gregory. *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support*. Cambridge University Press, 2005.
- [107] X.-N. Wang and M. Gyulassy. hijing: A Monte Carlo model for multiple jet production in pp, pA, and AA collisions. *Phys. Rev. D*, 44:3501–3516, December 1991.
- [108] T. Sjöstrand et al. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, jun 2015.
- [109] P. Skands et al. Tuning PYTHIA 8.1: the Monash 2013 tune. *The European Physical Journal C*, 74(8), aug 2014.
- [110] R. Brun et al. *GEANT: Detector Description and Simulation Tool; Oct 1994*. CERN Program Library. CERN, Geneva, 1993. Long Writeup W5013.
- [111] Machine learning package for high-energy physics. <https://github.com/alishw/MachineLearningHEP>.
- [112] T. Chen and C. Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pp. 785–794, New York, NY, USA, 2016. Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>.

- [113] M. Cacciari et al. The p_T spectrum in heavy-flavour hadroproduction. *Journal of High Energy Physics*, 1998(05):007–007, May 1998. <https://doi.org/10.1088%2F1126-6708%2F1998%2F05%2F007>.
- [114] M. Cacciari et al. Theoretical predictions for charm and bottom production at the LHC. *Journal of High Energy Physics*, 2012(10), October 2012. <https://doi.org/10.1007%2Fjhep10%282012%29137>.
- [115] The LHCb Collaboration. Measurement of b hadron fractions in 13 TeV pp collisions collisions. *Physical Review D*, 100(3), August 2019. <https://doi.org/10.1103%2Fphysrevd.100.031102>.
- [116] S. K. Das et al. Propagation of heavy baryons in heavy-ion collisions. *Physical Review D*, 94(11), December 2016. <https://doi.org/10.1103%2Fphysrevd.94.114039>.
- [117] A. Bazavov et al. Chiral crossover in QCD at zero and non-zero chemical potentials. *Physics Letters B*, 795:15–21, August 2019.
- [118] E. G. A. Mazeliauskas, S. Floerchinger and D. Teaney. FastReso – program for computing irreducible components of the particle distribution from direct resonance decays. <https://github.com/amazeliauskas/FastReso>, Dec 2019.
- [119] The ALICE Collaboration. Measurement of the cross sections of Ξ_c^0 and Ξ_c^+ baryons and branching-fraction ratio $\text{BR}(\Xi_c^0 \rightarrow \Xi^- e^+ \nu_e)/\text{BR}(\Xi_c^0 \Xi_c^0 \rightarrow \Xi \pi^+)$ in pp collisions at 13 TeV. *Physical Review Letters*, 127(27), December 2021. 10.1103/physrevlett.127.272001.
- [120] The ALICE Collaboration. Measurement of prompt D^0 , Λ_c^+ , and $\Sigma_c^{0,++}(2455)$ production in pp collisions at $\sqrt{s} = 13$ TeV. *Physical Review Letters*, 128(1), January 2022. 10.1103/physrevlett.128.012001.
- [121] The ALICE Collaboration. First measurement of Ω_c^0 production in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 846, 2023. <https://doi.org/10.1016/j.physletb.2022.137625>.