Λ_c^+ baryon production measurements with the ALICE experiment at the LHC

Jaime Norman

University of Liverpool

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

December 2017

Abstract

Quantum chromodynamics, the quantum field theory that describes the strong interaction, demonstrates a property known as asymptotic freedom which weakens the strong coupling constant α_s at high energies or short distances. The measurement of particles containing heavy quarks, i.e. charm and beauty, in high-energy particle collisions is a stringent test of the theory of quantum chromodynamics in the regime where α_s is small. In addition, asymptotic freedom leads to a phase transition of nuclear matter at high temperatures or energy densities to a phase known as the Quark-Gluon Plasma, where quarks and gluons are deconfined, and this state of matter can be studied in relativistic heavy-ion collisions. Particles containing heavy quarks, i.e. charm and beauty, have been proposed as probes of the properties of the Quark Gluon Plasma, where the measurement of mesons and baryons can offer insight into the transport properties of the medium and mechanisms related to the formation of hadrons during the transition back to 'confined' quark states. Proton-proton and proton-lead collisions offer a crucial benchmark for these measurements, and can also reveal important insights into particle production and interaction mechanisms.

The goal of this thesis is to investigate the production of the charmed baryon Λ_c^+ in high-energy particle collisions with the ALICE detector at the Large Hadron Collider. The measurements presented will test predictions utilising perturbative (small α_s) and non-perturbative (large α_s) methods, will test possible cold-nuclear-matter modifications of the Λ_c^+ yield in proton-lead collisions, and will set the stage for future measurements in lead-lead collisions. The measurements are carried out by reconstructing the hadronic decay channel $\Lambda_c^+ \to \mathrm{pK}^-\pi^+$, making selections on its decay topology, extracting the signal via an invariant mass analysis, and finally correcting for its selection and reconstruction efficiency. A multivariate technique (Boosted Decision Trees) has been developed and

is utilised in order to improve the signal extraction by optimally combining discriminating variables related to the Λ_c^+ decay topology. This technique has also been investigated as a possible approach to measuring the Λ_c^+ baryon in lead-lead collisions in the future, after the upgrade of the ALICE Inner Tracking System will make this measurement possible.

The transverse momentum dependence of the Λ_c^+ baryon production cross section has been measured in proton-proton collisions at a centre-of-mass energy of 7 TeV and proton-lead collisions at a centre-of-mass energy per nucleon-nucleon collision of 5.02 TeV, in the transverse momentum range $2 < p_{\mathrm{T}} < 12$ GeV/c, and is shown to be under-predicted by theoretical calculations. The baryon-to-meson ratio Λ_c^+/D^0 has been measured to be consistent in proton-proton collisions and proton-lead collisions and under-predicted by theoretical calculations. The nuclear modification factor R_{pPb} is measured to be consistent with unity and in agreement with the D meson R_{pPb}, indicating no significant modification of the Λ_c^+ yield in proton-lead collisions with respect to proton-proton collisions within the experimental uncertainties. Finally, Boosted Decision Trees have been shown to significantly improve the statistical precision with which the measurement of the Λ_c^+ baryon can be made in lead-lead collisions with the ALICE detector in the future.

Author's Contribution

Throughout my PhD I have been working as part of the ALICE collaboration. The work presented in this thesis extends the $\Lambda_c^+ \to pK^-\pi^+$ analysis in pp and p–Pb collisions to include a multivariate selection (using Boosted Decision Trees) of Λ_c^+ candidates. This was carried out extending the framework developed by collaborators, and while I used much of the pre-existing framework all analysis work in this thesis, including the Λ_c^+ candidate selection, signal extraction, corrections and systematic uncertainties was performed by me. I also played a large role in the analysis performed in pp and p–Pb collisions using the standard, cut-based method, and developed the strategy to merge the Λ_c^+ analyses measured using different decay channels, though this is not presented in detail here. The preliminary results shown in this thesis were presented by myself at the 'Strangeness in Quark Matter' conference in 2017 [1].

An extension of this work also involved a study to evaluate the statistical improvement that will be gained from using a multivariate approach to extract the Λ_c^+ signal in Pb–Pb collisions, after the upgraded ALICE Inner Tracking System is installed. This was carried out as an extension of and comparison to work performed for the Inner Tracking System Technical Design Report, and this work constitutes a chapter of this thesis.

In addition to the work presented here, I helped review the analysis of the D meson production cross section measurements at $\sqrt{s} = 5$ TeV as part of the Analysis Review Committee, which have been presented as preliminary results at conferences. For my service task, I worked on a laser soldering technique that was being explored to bond the silicon chips of the Inner Tracking System upgrade to a circuit-board. I also had the opportunity to give overview talks on open heavy-flavour measurements at ALICE - at the 'High p_T physics in the RHIC-LHC era' workshop at Brookhaven National Laboratory, and at the '55th International Winter Meeting on Nuclear Physics' in Bormio, Italy [2].

All the work presented in this thesis is the result of my own work, except where explicit reference is made to the work of others.

Acknowledgements

Firstly, big thanks to my primary supervisor Marielle Chartier. From when I begun my undergraduate project to finishing my PhD, Marielle has supported and encouraged me throughout. Thanks also to my secondary supervisors, John Dainton for patiently talking through many ideas in this work, and Roy Lemmon for his feedback and help in shaping this thesis. Many thanks to Rossella Romita for first introducing me to the ALICE experiment and this analysis, and to Marcel Araujo Silva Figueredo, with whom the analyses in this thesis were finalised. It was a pleasure to work closely with both of you during the first couple of years of my PhD, and I learned a lot from both of you.

During my time in ALICE I have been lucky to work in collaboration with many people, with whom I shared many interesting discussions. Thanks to those who I have worked with directly on the Λ_c^+ analysis, in particular to Chiara Zampolli and Andrea Alici for their help at the start of this analysis. Thanks to to Davide Caffarri and Zaida Conesa del Valle for the helpful discussions as well as the opportunity to present heavy-flavour results from ALICE at interesting conferences, and thanks to Andre Mischke, Jesus Guillermo Contreras Nuno, Francesco Prino and Martin Voelkl for their help in reviewing the analyses at different stages. Thanks to Andrea Dainese and Cristina Terrevoli for their useful discussions in shaping the studies presented in chapter 6 of this work. I would also like to thank Petra Riedler, Luciano Musa and Antoine Junique for the opportunity to work on the upgrade of the Inner Tracking System for my service task.

Liverpool has become a second home over the past seven years. Thanks to everyone I met during my undergraduate degree and my PhD, in particular Kārlis Dreimanis, John Anders and David Morris for all the jokes and the drinks. Huge thanks to Jonathan Tinsley for proofreading sections of this thesis. Thanks also to all my non-physicist friends for giving me a welcome escape from physics when needed and for pretending to be interested in this work.

Thanks to my family; my sister Meryn, and my parents Maggie and Bob - I would not be writing this if it wasn't for my parents consistent encouragement and support. And lastly to the 'other' Alice, Alice Colquhoun, for sharing all the stress, the travelling, and enjoyment of the last few years. She has been much more than a $16 \times 16 \times 26$ m^3, $10,000$ tonne particle physics experiment to me, and for that I thank her.

For Peter Norman,

"science is either physics

or stamp collecting."

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In our current understanding of nature and the universe at the most fundamental scale, there exists four fundamental forces, which act in different ways. Gravity, which acts upon anything with mass, the electromagnetic force which acts on electrically charged particles, the weak force, which is the cause of some radioactive decays, and the strong force, which binds together quarks into particles known as hadrons, as well as confining protons and neutrons within the atomic nucleus. The latter three forces are all described in terms of the Standard Model of particle physics, which also describes the particles which couple to these forces. Gravity has so far eluded integration with the Standard Model.

The Standard Model was born out of the two great revolutions of physics in the 20th century. The first is quantum mechanics, which describes nature at small scales. Its initial formulation explains the structure of atoms, composed of a nucleus of protons and neutrons, orbited by electrons, and thus explains much of atomic physics. The second is special relativity, which describes the propagation of light and matter up to high speeds. The relativistic wave equation for a free spin-1/2 particle, expressed in natural units as $(i\partial\!\!\!/ - m)\psi = 0$, is consistent with both theories, and among other things, predicted the existence of antimatter, accounts for the fine structure of hydrogen, and justifies that particles have quantised spin. It doesn't however account for particle creation and annihilation. The development of Quantum Field Theory (QFT) led to a full theory consistent with both quantum mechanics and special relativity, while accounting for observed violations in particle 'number conservation' - conceptually, QFT describes fields pervading all space and particles existing as excitations of these fields. The Standard Model is a manifestation of quantum field theories which rose from a combined effort of theoretical development, and experimental observations which provided remarkable confirmation of many predictions from what we now understand as the theoretical basis for describing the known universe.

1.1 The Standard Model

In the Standard Model there exist 17 fundamental particles. There are 12 spin-1/2 particles known as 'fermions', and 5 integer spin particles known as 'bosons'.

Within the fermions there exists 6 types of 'quarks', which interact via the strong force and electroweak force (and can be said to have 'colour charge', as well as 'electric charge' and 'weak isospin'). More will be said about quarks in the next section. There are 6 more types of fermions known as 'leptons', which do not carry colour charge, including the electron, the muon and the tau which have electric charge, and accompanying each of these are 3 types of neutrinos (electron neutrino, muon neutrino and tau neutrino), which are electrically neutral and interact only via the weak force. The property that distinguishes between the different types of fermions is called 'flavour', and flavour-changing processes can occur under the weak interaction.

The bosons in the Standard Model have spin-1 (known as 'gauge bosons') or spin-0 ('scalar bosons'). The gauge bosons in the Standard Model are defined as the force carriers, which mediate the strong, weak and electromagnetic interactions. The photon mediates the electromagnetic interaction, and is massless. The W and Z bosons (W^+, W^-, Z) are massive, and mediate the weak interaction. The gluon mediates the strong interaction between quarks. The colour charge of gluons gives different types of gluons arising from the different possible combinations of colour and anticolour charge. The only spin-0 boson within the Standard Model is the Higgs Boson - a massive particle which is responsible for giving other massive particles their (bare) mass.

The interactions within the Standard Model are formulated within the principle of local gauge invariance, with the gauge symmetry described by the group product $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$.

1.2 Quantum Chromodynamics

The rapid progression of particle physics and detector technology in the 1950's led to a large number of new particle discoveries, known then and now as hadrons. Properties such as charge, isospin and 'strangeness' were assigned to them, but an overall theory was lacking. To impose some order on the plethora of unexplained particles, Gell-Mann and Ne'eman proposed a flavour symmetry under the special unitary group SU(3) such that the baryons and mesons correspond to irreducible representations of

the group [3]. Later in 1964, Gell-Mann and Zweig proposed that hadrons were composite particles formed from up, down and strange (u, d, s) spin-1/2 particles with fractional electric charges [4, 5]. This was initially more as a useful mathematical construct, however electron-proton scattering experiments where the electron rebounded inelastically from protons, known as deep-inelastic scattering, were proposed to indicate whether the proton was composed of more fundamental particles [6, 7]. These experiments indeed suggested the proton contains smaller, point-like objects, then named 'partons', and this model was further verified with observed phenomena in neutrino-nucleon, electron-positron and proton-proton collisions, from which conclusive evidence was obtained that these partons were charged particles with spin-1/2. With the prediction of the charm quark [8] and its subsequent discovery [9, 10], the quark model was confirmed. The discovery of the beauty quark (sometimes called 'bottom' quark) in 1977 [11] and the top quark in 1995 [12] completed the six types of quarks that are within the Standard Model. Table 1.1 gives the mass and electric charge of each of these quarks.

Generation	Name	Mass (GeV/c^2)	Charge/e
1	up (u)	$2.2^{+0.6}_{-0.4} \times 10^{-3}$	$+2/3$
	down (d)	$4.7^{+0.5}_{-0.4} \times 10^{-3}$	$-1/3$
2	strange (s)	$96^{+8}_{-4} \times 10^{-3}$	$-1/3$
	charm (c)	1.27 ± 0.03	$+2/3$
3	beauty (b)	$4.18^{+0.04}_{-0.03}$	$-1/3$
	top (t)	$173.21 \pm 0.51 \pm 0.71$	$+2/3$

Table 1.1: The properties of the three quark generations, including their mass and electric charge [13].

Quarks have a property called 'colour charge', with baryons composed of a red, green and blue quark, and mesons composed of a colour-anticolour pair, forming colour singlet states with a net colour charge of 0. The property of colour was first proposed to explain how quarks could exist in some observed hadrons (e.g. Δ^{++} with quark content uuu [14]) with otherwise identical quantum states without an additional quantum number to prevent the violation of the Pauli exclusion principle. Colour was verified from e^+e^- data where the ratio $R = e^+e^- \rightarrow \text{hadrons} / e^+e^- \rightarrow \mu^+\mu^- \propto N_C \sum_f Q_f^2$ with Q_f the electric charge of a quark with flavour f confirmed the number of colour charges $N_C = 3$ [13].

The unification of the electromagnetic and weak nuclear forces within the framework of gauge field theories led to their application for the strong interaction. The

quantum field theory of the strong interaction is known as Quantum Chromodynamics (QCD), and is a non-Abelian gauge theory based on the symmetry group SU(3). The non-Abelian nature of QCD leads to the gluon carrying colour charge, meaning gluons can self-interact, generating 3- and 4-gluon interaction vertices. This property of gluon self-interactions represents a major difference to Quantum Electrodynamics (QED), the (Abelian) quantum field theory describing the electromagnetic interaction, where the photon does not carry electric charge. This results is some of the fundamental phenomena arising from QCD, as will be explained in the next section.

1.2.1 Confinement

The strength and behaviour of the coupling strength of quantum field theories at different scales can be understood through the 'vacuum polarisation'. The vacuum, according to QFT, is not a vacuum in the sense of being completely empty, but instead contains virtual particles. In QED if a charge is inserted in the vacuum, it polarises the electron-positron pairs, effectively 'screening' the charge, and cancelling out some of the charge seen at some distance r. As r increases, or if the charge is probed at smaller energies, there are more virtual particles screening the charge such that the effective charge is reduced, and conversely at smaller r or if probed at higher energies the screening effect reduces, such that the effective charge increases. The strength of the force is determined by the coupling constant α and its change with energy or distance (or, as often expressed, with momentum transfer squared Q^2) is known as 'running coupling'.

In QCD the colour charge is screened in a similar way by quark-antiquark (colour-anticolour) pairs, decreasing the overall colour charge. However, since gluons are self-interacting, this effect also occurs with gluons where gluon loops instead contribute to the overall colour charge, leading to an *anti-screening* effect. The interplay of these screening and anti-screening effects determines how the *strong* coupling α_s scales with energy or distance. One can define the β function which describes the running of the coupling α_s at a given scale μ:

$$\beta(\alpha_s) = \mu^2 \frac{d\alpha_s(\mu^2)}{d\mu^2}. \tag{1.1}$$

To first order, the β-function for an SU(N) gauge theory with n_f flavours of quarks has been calculated as $\beta_1(\alpha_s) = \frac{\alpha_s^2}{\pi} b_0 = \frac{\alpha_s^2}{\pi}\left(-\frac{11N}{6} + \frac{n_f}{3}\right)$[15] which for 6 quark flavours is negative. This means that the anti-screening effect is greater than the screening effect, and demonstrates the property of QCD called *asymptotic freedom* - that at

large energies $\alpha_s(\mu)$ will approach zero (and conversely, that $\alpha_s(\mu)$ diverges at small energies). This behaviour of α_s is confirmed by experiments where its value is evaluated at different scales using various experimental methods, as shown in figure 1.1. At the mass of the Z boson, $\alpha_s(M_Z) = 0.1181 \pm 0.0011$ [13].

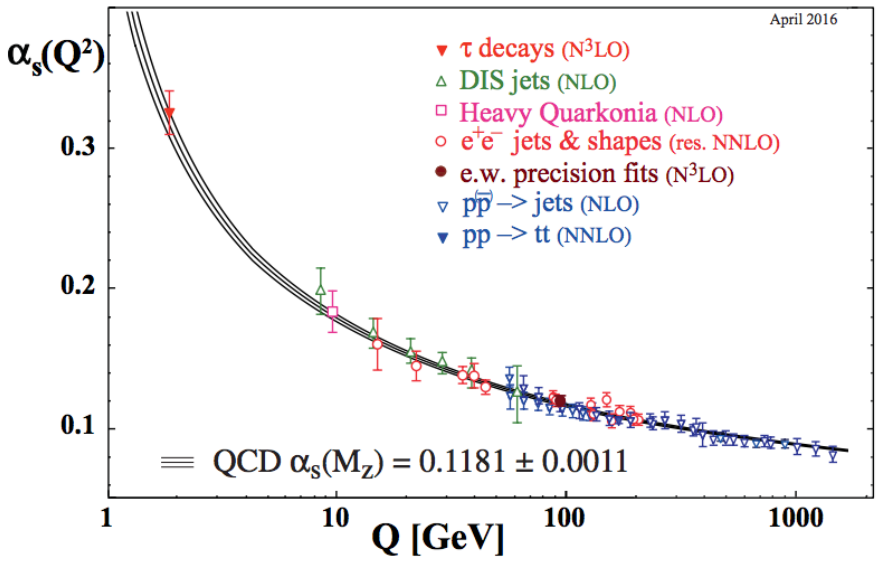

Figure 1.1: Summary of measurements of $\alpha_s(Q)$ as a function of the respective energy scale Q. The brackets indicate the order to which perturbation theory is used [13].

Predictive calculations can be performed by expanding the relevant observable in a perturbative series in powers of α_s. For large energies (or small distances), the higher-order terms get smaller and smaller such that they can be neglected, and the observable can be calculated to a given order of α_s. This idea forms the basis of perturbative QCD (pQCD), and is the method used for calculating observables (e.g. production cross sections) at high-energy colliders. However at small energies (or large distances) where $\alpha_s \sim 1$, a perturbative expansion is no longer possible, and one must rely on non-perturbative methods for predictions.

Asymptotic freedom was discovered by Gross, Wilczek and Politzer in 1973 [16–19], an achievement for which they won the Nobel Prize in 2004 [20], and is one of the most important properties of QCD. As well as describing the property of confinement at low energies it also implies that bulk QCD matter at high energy densities will undergo a phase transition from confined hadrons to a state where the strong force weakens such that quarks and gluons are no longer confined.

1.2.2 Deconfinement

The concept of a critical point of hadronic matter was first postulated in the mid-1960's by Hagedorn [21], before the formulation of QCD and the discovery of quarks, where it was proposed that there would be an upper temperature limit for hadronic matter. Hadronic interactions provide abundant resonance production, and the density of particle species as a function of their mass $\rho(m)$ was found to follow the functional form

$$\rho(m) \propto e^{\frac{m}{T_0}} \tag{1.2}$$

where T_0 is a parameter which when fit to the data comes out to be around 160 MeV. Hagedorn interpreted T_0 as the limiting temperature of hadronic matter, due to the fact that the partition function for a density of this form diverges, effectively introducing an (unphysical) infinite number of degrees of freedom, for temperatures $T > T_0$.

It was later suggested by Cabibbo [22] that this limiting temperature represents the existence of a phase boundary, and temperatures larger than a critical temperature T_C may represent a phase where quarks are not confined within hadrons. A simplistic model illustrating this phase transition, called the Bag Model [23], relates pressure and temperature for these phases. In this model, hadrons are modelled as a 'bag' of quarks with a constant, positive potential energy B which acts as an external pressure. With this consideration, an ideal gas can be thought of as a gas of bags, and 'quark matter' can be thought of as a single, big bag whose pressure is reduced by B. The pressure as a function of temperature for a gas of massless pions P_π, and an ideal deconfined state of quarks and gluons with two flavours and three colours P_{qg}, can be written in the Stefan-Boltzmann[1] form [24]:

$$P_\pi = g\frac{\pi^2}{90}T^4 \tag{1.3}$$

$$P_{qg} = g\frac{\pi^2}{90}T^4 - B. \tag{1.4}$$

Here g is the effective number of degrees of freedom of the considered particles, with $g = 3$ for pions and $g = 37$ for quarks and gluons[2], and B represents the

[1]The Stefan-Boltzmann form for the pressure is valid for a relativistic, non-interacting gas.

[2]For pions, there are 3 pion species leading to the degree of freedom $g = 3$. For a quark-gluon gas there are 2×8 degrees of freedom for the gluon accounting for the 2 spin states and 8 types of gluon, and $\frac{7}{9}(2 \times 2 \times 3 \times 2)$ degrees of freedom for the quark taking into account the 2 spin states, 2 flavours, 3 colours and 2 particle-antiparticle combinations, and the factor $\frac{7}{8}$ taking into account

bag pressure, taking into account the difference between the pressure in a vacuum and the ground state for quarks and gluons in a medium. Taking the crossover of the two pressures, and the bag pressure determined empirically from hadron spectroscopy as $B^{1/4} \simeq 0.2$ GeV [25, 26], one obtains

$$T_C \simeq 0.72 B^{1/4} \simeq 150 \text{ MeV} \tag{1.5}$$

The change in the active degrees of freedom of nuclear matter at high temperatures is clear from the temperature dependence of the energy density ϵ. Shown in figure 1.2 is ϵ/T^4 as a function of T, calculated on the lattice[3] for two massless quarks, three massless quarks and two massless quarks with one massive quark (s) with its real mass [27]. An abrupt change in the energy density is seen around the critical temperature ($T/T_C = 1$), representing the transition to a deconfined phase where quarks which were confined within hadrons are liberated and begin contributing to the bulk thermodynamic properties. Above the critical temperature the energy density depends on the number of quarks flavours up to high temperatures, such that the active light quark degrees of freedom could be verified experimentally. As $T \rightarrow \infty$ the energy density approaches the ideal (Stefan-Boltzmann) limit for a gas of massless quarks and gluons. In the end, lattice QCD predicts a critical temperature of hadronic matter of around 170 MeV [24] - very close to that predicted from the simple bag model. Calculations on the lattice also reveal a critical energy density of $\epsilon_0 \sim 0.5 - 1$ GeV/fm^3 [28].

So far, it has been outlined that, theoretically, a state of matter exists at sufficiently high temperatures or densities where quarks and gluons are deconfined. This state of matter is known as the Quark-Gluon Plasma (QGP). Figure 1.3 shows the phase diagram of nuclear matter as we know it today, where the vertical axis is the temperature of the system and the horizontal axis shows the baryo-chemical potential[4] μ_B. At low μ_B, lattice QCD results [29] indicate that a rapid crossover occurs as the temperature changes, whereas models based on effective field theories of QCD which reproduce other aspects of the phase diagram [30–33] suggest that a first-order

the difference between Fermi-Dirac and Bose-Einstein statistics. This leads to the degree of freedom $g = 37$.

[3]To study the nature of the equation of state (phase diagram) under QCD, calculations can be done within the framework of Lattice QCD. This is a non-perturbative approach which can be applied at small energy scales or long distances, formulated on a grid of space-time points where quarks are defined at lattice nodes, which are connected by gluon fields.

[4]The baryo-chemical potential is a measure of how the free energy of the system changes when the baryon number is changed, and represents the net baryon density of the system. It is also closely related to the overall density of the system.

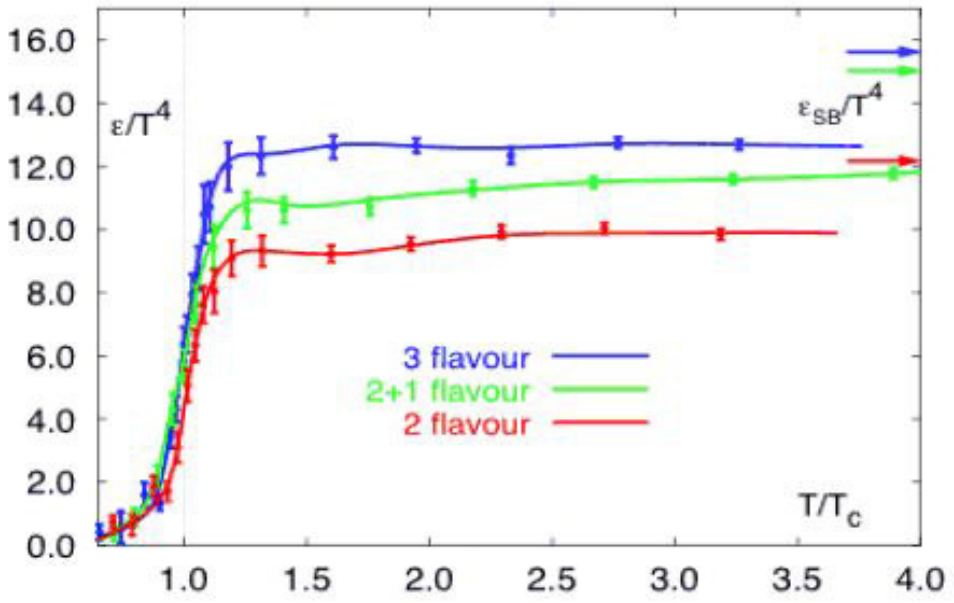

Figure 1.2: The energy density ϵ divided by T^4, equivalent to the number of degrees of freedom, computed on the lattice. The arrows on the right axis show the limit for a perfect massless gas [27].

phase transition occurs for low T and large μ_B. This implies that at some point in the phase diagram there exists a critical point, the search for which is an active field of study (see e.g. [34]). Going further to larger μ_B and smaller T a 'colour superconductor' phase is theorised to exist, also in a region where the QCD coupling is weak [35].

In practice, the temperatures and densities required to achieve this state are so large that they do not exist in nature at present (except for perhaps in the core of neutron stars, in the colour superconducting phase indicated in figure 1.3 [35]). However, the universe was once more compact, and therefore hotter and denser than the present day, and up until $\sim 10^{-6}$ sec after the big bang the temperature of the universe was high enough such that the whole universe was a Quark-Gluon Plasma. The expansion of the universe caused cooling until quarks could form hadrons. The more stable hadrons (protons and neutrons) went on to form nuclei in the process of nucleosynthesis, and the density anisotropies of the universe allowed stars, galaxies and the universe we see today to form.

To recreate these conditions in the lab, heavy ions are collided at highly relativistic energies. In such (central) collisions of nuclei a large amount of energy is liberated in a volume equalling the size of the atomic nuclei ($R \propto A^{1/3}$ where A is the mass number

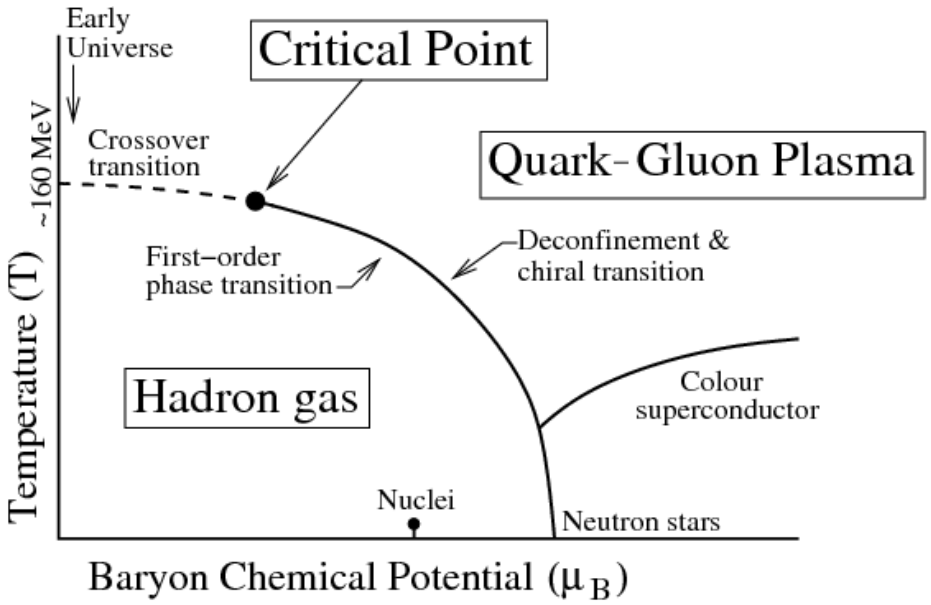

Figure 1.3: The QCD phase diagram (temperature as a function of baryo-chemical potential μ_B) of nuclear matter [36].

of the colliding nuclei), producing many particles, and the huge energy density in this small region of space should create sufficiently high temperatures to cross the phase boundary indicated in figure 1.3. The first conclusive evidence of a deconfined state of matter came from the Relativistic Heavy Ion Collider (RHIC), which primarily collides gold (Au-Au) beams at a maximum energy of $\sqrt{s_{\rm NN}} = 200$ GeV. The ALICE detector has been built at the Large Hadron Collider (LHC) with the primary aim of studying lead–lead (Pb–Pb) collisions at a maximum energy per nucleon-nucleon collision ~ 30 times that at RHIC, $\sqrt{s_{\rm NN}} = 5.5$ GeV. A description of the ALICE apparatus is given in chapter 3. The remainder of this chapter will give a brief outline of heavy-ion collisions, examples of observables used to study properties of the Quark-Gluon Plasma and a selection of important results in the light quark sector relevant for the work of this thesis.

Measurements in proton-proton (pp) collisions are an essential aspect of studying the Quark-Gluon Plasma in heavy-ion collisions. They can be used as a reference, from which any difference in heavy-ion collisions can be inferred to be down to physical differences due to the presence of a hot, interacting medium. pp collisions can also be used to address several topics within QCD where ALICE can compliment the other LHC detectors. Finally, the LHC has run with proton-lead (p–Pb) collisions, which are crucial to separate effects due to the Quark-Gluon Plasma from those that may be due to other effects occurring with the presence of a lead nucleus (known as 'cold

nuclear matter' effects), as well as offering insights into other aspects of QCD. These collision systems will be mentioned when necessary, since the measurements presented in this thesis are performed in pp and p–Pb collisions.

1.3 Quark-Gluon Plasma in heavy-ion collisions

Figure 1.4 shows a 2d representation of the light-cone, indicating two nuclei approaching each other and the measurable phases of the collision, in one space (horizontal axis) and one time (vertical) dimension. Nuclei with 'rest' radius R are Lorentz contracted in laboratory frame, and the crossing time of the nuclei in the laboratory frame can be estimated as $\tau_{cross} = 2R/\gamma$ - for lead ions accelerated at the LHC at $\sqrt{s_{\rm NN}} = 2.76$ TeV , $\gamma \sim 1500$ and $\tau_{cross} = 10^{-3}$ fm/c. The crossing time is therefore much shorter than the time scale of the strong interaction[5] $\tau_{strong} \sim 1/\Lambda_{QCD} \sim 1$ fm/c. The formation of new partons occurs in a timescale of the order of the strong interaction $\tau_{form} \sim \tau_{strong} \sim 1$ fm/c, which continue to form and interact through secondary interactions before reaching thermal equilibrium.

After reaching thermal equilibrium, large pressure gradients reflecting the initial collisional impact geometry causes the partonic system to rapidly expand. This expansion occurs collectively due to the low viscosity of the medium, and is generally described using relativistic hydrodynamics.

After rapid expansion and cooling the medium reaches a critical temperature T_C where the energy density is small enough for quarks to hadronise, and the medium becomes an interacting gas of hadrons. At some point after hadronisation inelastic interactions between the medium's constituents cease and the hadron species abundance becomes fixed. This point is known as 'chemical freeze-out'. Elastic interactions between hadrons then continue for a brief time in a 'hadronic phase' until both inelastic and elastic processes cease and the momenta of the hadrons become fixed. This point is known as 'kinetic freeze-out'.

There are a number of methods that can be used to study each phase of the system's evolution. The very early stages of the medium can be probed by electromagnetic radiation, for example photons and dileptons (e^+e^- or $\mu^+\mu^-$ pairs), which are formed either from the interaction of quarks and gluons or by quark-antiquark annihilation - since they interact electromagnetically and not strongly, once formed

[5] Λ_{QCD} is known as the QCD scale coming from the expression for the strong coupling $\alpha_s(\mu) = \frac{1}{b_0 \ln \frac{\mu^2}{\Lambda^2_{QCD}}}$ (to first order), where $\Lambda_{QCD} \approx 200$ MeV. This energy scale also defines the time scale of the strong interaction.

they escape the medium with no further modification. Studying the hydrodynamic evolution of the system through the collision geometry and the final state momentum distribution can give information about bulk properties, such as the viscosity and the extent to which the medium thermalises. The total hadron yields can be used to study chemical freeze-out, and the relative hadron yields can give insight into hadronisation. Since the transition to QGP only occurs for light quarks (u, d, s), heavy quarkonia measurements (J/ψ, $\Psi(2S)$ and Υ) can be used to test 'suppression'[6] and 'regeneration'[7] mechanisms. Finally, partons formed in hard-scattering (large-momentum-transfer) processes at the start of the collision can lose energy traversing the medium, and quantifying this energy loss can give information about the transport properties of the medium.

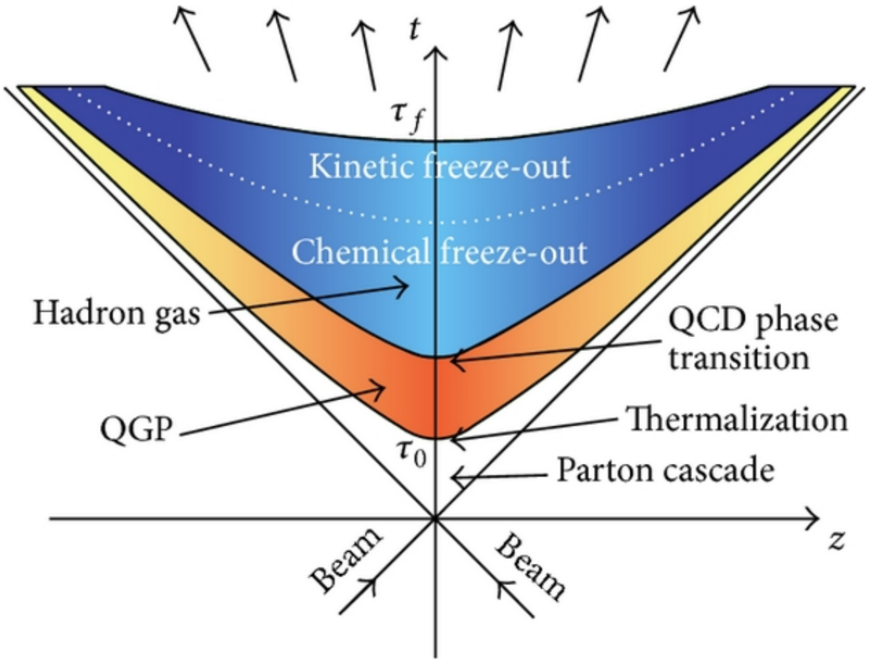

Figure 1.4: The space-time evolution of a heavy-ion collision, as a 2d representation of the light-cone. The nuclei collide at $(t, z) = 0$, and the created medium expands in time. The light-cone is shown as the diagonal straight lines, and each stage of the evolution of the collision, as described in the text, is indicated on the figure. Taken from [37].

[6]Quarkonia can exist bound in the medium up to temperatures similar to their binding energies at which point they can 'melt', suppressing the final quarkonia yield with respect to pp collisions.

[7]The high charm cross section at the LHC means many $c\bar{c}$ pairs are created in central Pb–Pb collisions. Quarkonia formed from hadronisation via 'coalescence' of two independently-produced c quarks can lead to a regeneration effect, reducing the level of suppression.

1.3.1 Kinematics of the collision

The following section will briefly outline some important measures used when studying colliding proton or heavy-ion beams.

1.3.1.1 Rapidity

The direction along the beam line (along the z-axis) is called longitudinal and the direction perpendicular to the beam line (in the x-y plane) is called transverse. The 3-momentum p is usually decomposed into the longitudinal (p_{z}) and transverse (p_{T}) momentum components, with p_{T} being a vector quantity which is invariant under a Lorentz boost along the longitudinal direction. The rapidity y of a particle with mass m and energy E is defined in natural units as

$$y = \frac{1}{2} \ln \left(\frac{E + p_{\mathrm{z}}}{E - p_{\mathrm{z}}} \right) \tag{1.6}$$

which changes by an additive constant under longitudinal Lorentz boosts. The shape of the rapidity distribution remains unchanged under a longitudinal Lorentz boost, so when moving from the centre-of-mass frame (CMS) to the laboratory frame (LS) the rapidity distribution is the same with the y-scale shifted by the rapidity in the centre-of-mass frame y_{cms}. For a particle emitted at an angle θ relative to the beam axis, the rapidity can be written as $y = \frac{1}{2} \ln \left[\frac{\sqrt{m^2+p^2}+p\cos\theta}{\sqrt{m^2+p^2}-p\cos\theta} \right]$ which at very high energies simplifies to

$$y = -\ln\left(\tan\theta/2\right) = \eta. \tag{1.7}$$

η is known as the pseudorapidity, which depends only on θ, and is the same for a particle of any mass, momentum or energy.

1.3.1.2 Centrality

Since heavy nuclei contain many nucleons with finite radius r_0, nuclei are extended objects with a radius $R \simeq r_0 A^{1/3}$. The geometry of a heavy-ion collision thus plays an important role in the description of its physics. The main quantity used to determine the collision geometry is the impact parameter b, defined as the transverse distance between the centres of two colliding nuclei, and can carry values from 0 to $R_1 + R_2$, where R_1 and R_2 are the diameters of the two nuclei. The 'centrality' of a collision is often used, and is given as a percentage of the impact parameter with respect to the added radii $b/(R_1 + R_2)$, where 0% centrality is a completely head-on

collision, and 100% centrality is a completely peripheral collision at the edge of where the nuclei spatially overlap. A given centrality corresponds to other geometrical quantities useful for centrality determination, including the number of participant nucleons N_{part}, defined as the number of nucleons in one nucleus which undergo one or more binary collisions with nucleons from the other nucleus, and the number of binary nucleon-nucleon collisions N_{coll}, defined as the total number of nucleon-nucleon collisions in an event.

The Glauber model [38] is used to estimate N_{coll} and N_{part} for a given centrality. In the Glauber model a nuclear density profile (usually a Woods-Saxon distribution [39]) is used to derive $N_{\text{coll}}(b)$ and $N_{\text{part}}(b)$ for a given b, done either numerically using a *Monte Carlo* (MC) approach, or analytically with the *optical limit approximation* approach, as detailed in [38]. Since there is a monotonic relationship between b and the particle multiplicity, centrality classes can be defined in terms of measured multiplicity intervals.

In Pb–Pb collisions the centrality of the collision is strongly correlated with the number of participant nucleons. However in p–Pb collisions the correlation is much weaker due to fluctuations in the distribution of the nucleons in the lead nucleus. In addition, hard scatterings can contribute significantly to the overall particle multiplicities. This leads to significant multiplicity fluctuations for a given centrality which are of a similar magnitude of the total range of multiplicities, and care has to be taken when choosing a centrality estimator to understand any possible biases present in the centrality estimator. The measurement in p–Pb collisions in this thesis is performed in minimum-bias, centrality-integrated p–Pb collisions, but for details related to centrality measurements in p–Pb collisions see [40].

1.3.2 Multiplicities and energy densities

The multiplicity of produced particles is an important property of the collisions related to the collision geometry, as mentioned in the previous section, as well as the initial parton densities and the energy density produced in the collision. In p–Pb collisions, or deuteron-gold (d–Au) at RHIC, the multiplicity can also be related to the initial parton distribution in the nucleus and can be used to understand the initial state. To compare bulk particle production in different collision systems the primary charged-particle density per unit of pseudorapidity $\mathrm{d}N_{\text{ch}}/\mathrm{d}\eta$ is scaled by the number of nucleon pairs, $\frac{2}{\langle N_{part}\rangle}\langle \mathrm{d}N_{\text{ch}}/\mathrm{d}\eta\rangle$, and is shown in figure 1.5 as a function of $\sqrt{s_{\text{NN}}}$. The results for p–Pb(d–Au) collisions are found to be consistent with the scaling of inelastic pp collisions and are described by the function $s_{NN}^{0.103}$, whereas the value for central

nucleus-nucleus (AA) collisions scales more rapidly with energy and is described by the function $s_{NN}^{0.155}$. In p–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV, $\frac{2}{\langle N_{part}\rangle}\langle {\rm d}N_{\rm ch}/{\rm d}\eta\rangle = 2.3\pm0.3$ [41], which is over four times smaller than that for Pb–Pb collisions at the same energy, where $\frac{2}{\langle N_{part}\rangle}\langle {\rm d}N_{\rm ch}/{\rm d}\eta\rangle = 10.1 \pm 0.3$ [42].

The initial energy density ϵ_0 reached in heavy-ion collisions can be estimated using the Bjorken model [43] from the measured transverse energy $E_{\rm T}$

$$\epsilon_0(\tau_0) = \frac{1}{\pi R^2}\frac{1}{\tau_0}\frac{{\rm d}E_{\rm T}}{{\rm d}y} \tag{1.8}$$

where the initial fireball is a cylinder of length ${\rm d}z = \tau_0 {\rm d}y$ and transverse radius $R \sim A^{1/3}$. For a central Pb-Pb collision the overlap area $\pi R^2 \sim 150$ fm^2, and the formation time occurs roughly on the timescale of the strong interaction $\tau_0 \sim 1$ fm/c. The measured ${\rm d}E_{\rm T}/{\rm d}\eta$ reaches 1700 GeV in the most central (0-5%) Pb-Pb collisions at 2.76 TeV, with ${\rm d}E_{\rm T}/{\rm d}y$ determined to be around 10% larger, giving the total energy density $\epsilon_0 \approx 12$ GeV/fm^3 [44]. It is noted that this is much larger than the critical energy density predicted from lattice QCD calculations of $\sim 0.5-1$ GeV/fm^3. At RHIC this was measured as ~ 5 GeV/fm^3 in Au–Au collisions at $\sqrt{s_{\rm NN}} = 200$ GeV

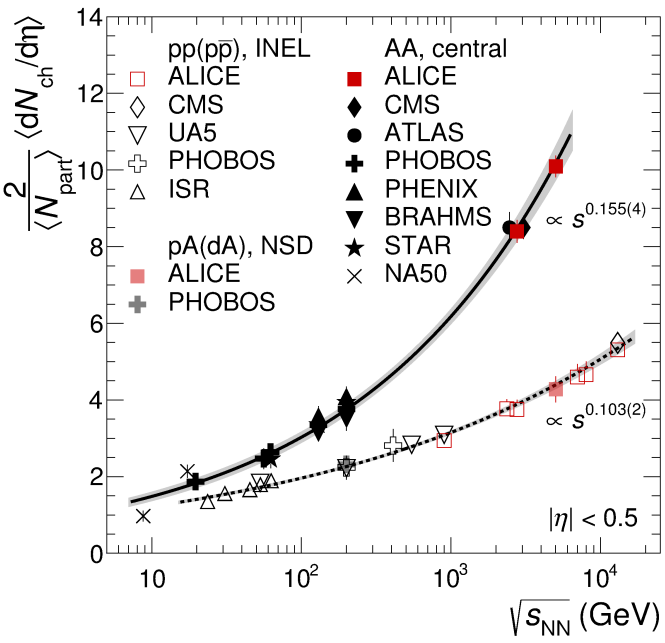

Figure 1.5: The charged-particle pseudorapidity density at mid-rapidity normalised to the number of participating nucleons, for proton-proton, proton-nucleus and central nucleus-nucleus collisions, as a function of the centre-of-mass energy per nucleon-nucleon collision [42].

[45, 46] and at top SPS energies as ~ 3 GeV/fm^3 [47], indicating the critical point was also surpassed in both of these cases.

1.3.3 Thermal properties of the Quark-Gluon Plasma

Integrated hadron yields give information about the extent to which the system reaches thermal equilibrium. A thermalised deconfined system undergoing a phase change should drive the system towards chemical equilibrium[8]. It has been observed that the hadron yield can be described by a statistical hadronisation model (SHM) [48] dependent only on temperature T, volume V and baryo-chemical potential μ_B, from Au–Au collisions at around $\sqrt{s_{\mathrm{NN}}} \sim 2$ GeV to Pb–Pb collisions at the LHC [49] - thus T and μ_B at chemical freeze-out can be extracted from the measured hadron yields. A global fit to the integrated hadron yields in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV, where T and μ_B are the free parameters of the fit, is shown in figure 1.6. Temperature values of around $T = 160$ MeV are extracted from these data.

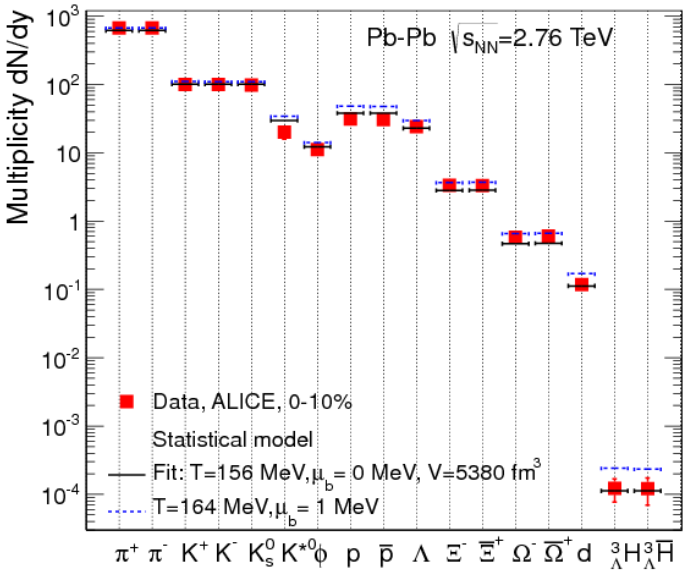

Figure 1.6: Hadron yields in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV measured by ALICE and fit with the statistical hadronisation model. The fit yields a temperature $T = 156$ MeV. Results of the model when fixing $T = 164$ MeV are shown to demonstrate the sensitivity of the fit [49].

[8] Analogous to thermal equilibrium, where the temperature of a system does not change temporally or spatially, chemical equilibrium implies that the relative abundance of particle type is in equilibrium and does not change.

The temperature extracted from the same fitting procedure for lower energy collisions decreases to around 50 MeV for Au–Au collisions at $\sqrt{s_{\mathrm{NN}}} = 2.24$ GeV, while the baryo-chemical potential decreases smoothly as a function of energy. In this way the phase diagram in figure 1.3 can be roughly mapped out [50].

1.3.4 Nuclear modification factor

One way to study the hottest and densest phases of high-energy heavy-ion collisions is to study particles emitted with high transverse momentum. High-p_{T} partons (that is, quarks or gluons) traversing the medium are sensitive to a range of effects, which are expected to modify their distribution with respect to pp collisions. In particular, energy loss due to induced gluon radiation or elastic collisions is known as 'jet quenching', and its measurement can be used to study the thermodynamic or transport properties of the QGP. The nuclear modification factor R_{AB} of some observable Φ_{AB} (where A and B are the respective objects colliding with an impact parameter b), measured as a function of, e.g, collision energy $\sqrt{s_{\mathrm{NN}}}$, transverse momentum p_{T}, or rapidity y, is defined by comparing to the same observable measured in pp collisions Φ_{pp}, where R_{AB} is given as

$$R_{\mathrm{AB}}(\sqrt{s_{\mathrm{NN}}}, p_{\mathrm{T}}, y; b) = \frac{\text{“hot/cold QCD medium”}}{\text{“QCD vacuum”}} \propto \frac{\Phi_{AB}(\sqrt{s_{\mathrm{NN}}}, p_{\mathrm{T}}, y; b)}{\Phi_{pp}(\sqrt{s_{\mathrm{NN}}}, p_{\mathrm{T}}, y)} \tag{1.9}$$

which can be more specifically defined, in p–Pb or AA collisions, here as a function of p_{T}, as

$$R_{\mathrm{AA}} = \frac{\mathrm{d}N_{\mathrm{AA}}/\mathrm{d}p_{\mathrm{T}}}{\langle N_{\mathrm{coll}}\rangle\, \mathrm{d}N_{\mathrm{pp}}/\mathrm{d}p_{\mathrm{T}}} = \frac{\mathrm{d}N_{\mathrm{AA}}/\mathrm{d}p_{\mathrm{T}}}{\langle T_{\mathrm{AA}}\rangle\, \mathrm{d}\sigma_{\mathrm{pp}}^{\mathrm{INEL}}/\mathrm{d}p_{\mathrm{T}}} \tag{1.10}$$

$$R_{\mathrm{pPb}} = \frac{\mathrm{d}N_{\mathrm{pPb}}/\mathrm{d}p_{\mathrm{T}}}{\langle N_{\mathrm{coll}}\rangle\, \mathrm{d}N_{\mathrm{pp}}/\mathrm{d}p_{\mathrm{T}}} = \frac{\mathrm{d}N_{\mathrm{pPb}}/\mathrm{d}p_{\mathrm{T}}}{\langle T_{\mathrm{pPb}}\rangle\, \mathrm{d}\sigma_{\mathrm{pp}}^{\mathrm{INEL}}/\mathrm{d}p_{\mathrm{T}}} \tag{1.11}$$

where $N_{\mathrm{AA/pPb/pp}}$ is the observable's yield in AA, p–Pb or pp collisions, respectively, $\sigma_{\mathrm{pp}}^{\mathrm{INEL}}$ is the observable's inelastic cross section in pp collisions, $\langle N_{\mathrm{coll}}\rangle$ is the average number of binary nucleon-nucleon collisions for the considered centrality interval, and $\langle T_{\mathrm{AA}}\rangle = N_{\mathrm{coll}}/\sigma_{\mathrm{pp}}$ is the nuclear overlap function, which can be thought of as the effective integrated luminosity per collision. A nuclear modification factor less than unity indicates a suppression, and a nuclear modification factor greater than unity indicates an enhancement, with respect to the same number of independent pp collisions.

A measured high-p_{T} particle can be interpreted as the leading particle within a jet. Measurements of the centrality dependence of the nuclear modification of identified particles (p, K, π) in Pb–Pb collisions [51], as shown in figure 1.7 shows a clear decrease in the R_{AA} with more central collisions. A similar suppression at high p_{T} for each hadron species is seen, suggesting that there is no dependence between the energy loss in the medium and the particle species. At intermediate p_{T}, there is an interplay between hydrodynamics and jet quenching, which leads to a mass dependent R_{AA}. The modified partonic density in the nucleus with respect to a free proton, as well as other non-perturbative effects, is responsible for the suppression of the R_{AA} at low p_{T}. Finally, baryons and mesons may hadronise differently, manifesting in a baryon 'enhancement' at intermediate p_{T} (explained in section 1.3.5). The direct measurement of jets, while more difficult experimentally, are the closest objects to the initial parton that can be properly defined. The measurement of both charged [52] and calorimetric [53] jets in Pb–Pb collisions at the LHC shows a similar level of suppression to charged particles.

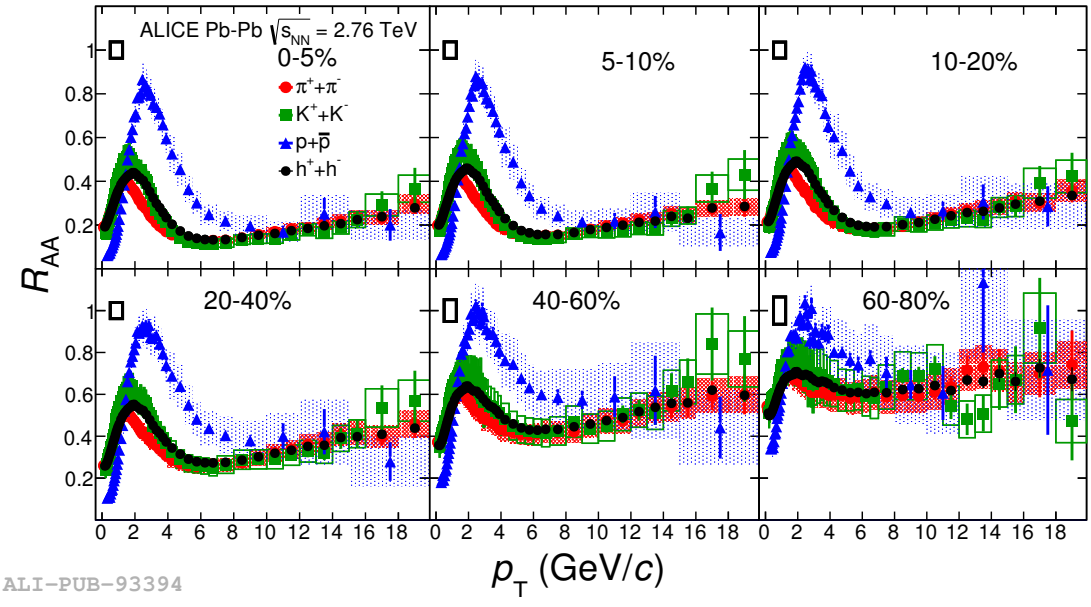

Figure 1.7: The nuclear modification factor R_{AA} in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV as a function of p_{T} for different particle species. Results for different collision centralities are shown [51].

The nuclear modification factor of charged hadrons in p–Pb collisions [54] is shown in figure 1.8 in comparison to the nuclear modification factor of charged hadrons in Pb–Pb collisions [55]. For $p_{\mathrm{T}} > 2$ GeV/c the R_{pPb} is consistent with unity, showing that the large suppression seen at high p_{T} in Pb–Pb collisions is due to jet quenching

effects. Note that the R_{AA} of photons [56] as well as $\mathrm{W}^{\pm}$ [57] and Z^{0} [58] bosons is consistent with unity, since these particles do not interact strongly and therefore are not affected by the strongly-interacting medium.

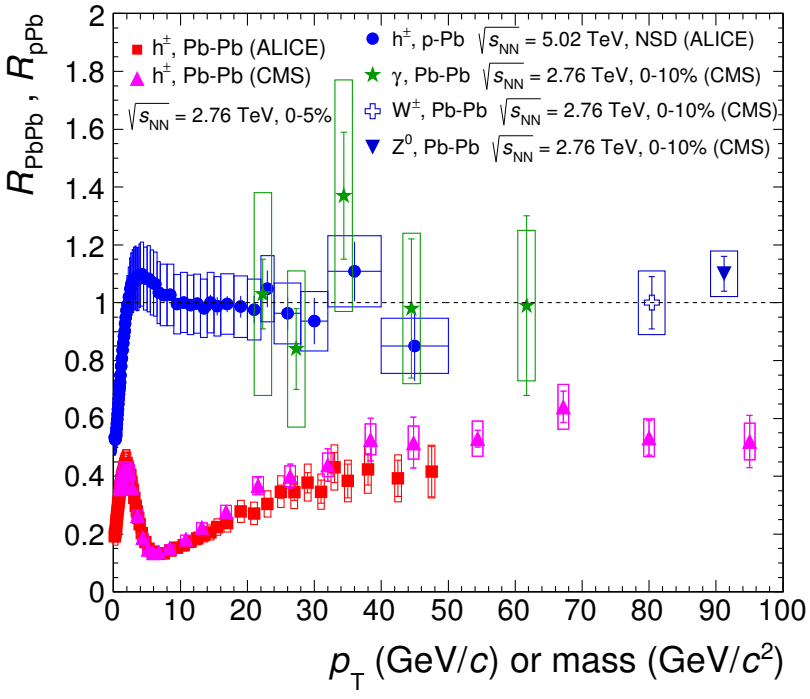

Figure 1.8: The nuclear modification factor as a function of p_{T} of charged hadrons in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV [55] and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [54], and the nuclear modification factor of γ [56], $\mathrm{W}^{\pm}$ [57] and Z^{0} [58] in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV.

1.3.5 Baryon-to-Meson Anomaly

It has been observed in heavy-ion collisions at RHIC that the $\Lambda/\mathrm{K}^{0}_{\mathrm{S}}$ and p/π ratios[9] at intermediate p_{T} ($\sim 2-6$ GeV/c) show a marked enhancement in central heavy-ion collisions with respect to peripheral heavy-ion and pp collisions [59, 60]. The same observation has been made at ALICE [61]. It has been suggested that a new mode of hadronisation can occur after deconfinement, where two or three quarks with low and similar momentum can combine to form a baryon or meson. In this scenario of hadronisation, given the name 'recombination' or 'coalescence' [62], baryons appear

[9]When writing ratios of particles (e.g. p/π, $\Lambda/\mathrm{K}^{0}_{\mathrm{S}}$, $\Lambda_{\mathrm{c}}/\mathrm{D}^{0}$), this refers to the ratios of the number of particles produced, i.e. the particle 'yields' either from their integrated yields, or differentially as a function of e.g. p_{T} or event multiplicity. This is often calculated from production cross section measurements.

at higher momentum since their total momentum is the sum of the three constituent quarks, rather than two for mesons. As an alternative or additional explanation, at lower p_{T} the hydrodynamical evolution of the system could also contribute to baryon enhancement, where radial flow would push baryons (which are heavier than mesons) to higher momentum.

Figure 1.9 shows a comparison between central and peripheral heavy-ion collisions, comparing Au–Au results from the STAR experiment at RHIC, and Pb-Pb collisions from ALICE [61]. The shape of both results is seen to be very similar, with the peak of the ratio shifted slightly to higher values and higher p_{T} for the ALICE results. The data is compared to models including a hydrodynamical calculation [63], a recombination calculation [64], and the EPOS event generator [65], which takes into account the interaction between jets and the hydrodynamically expanding medium. At low p_{T} the hydrodynamical model describes the data, but deviates after $p_{\mathrm{T}} \gtrsim 2\mathrm{GeV}/c$, while the recombination model qualitatively reproduces the shape but overestimates the baryon enhancement. The EPOS model describes the data well. Recent calculations including a combination of recombination and flow also describe the data well [66].

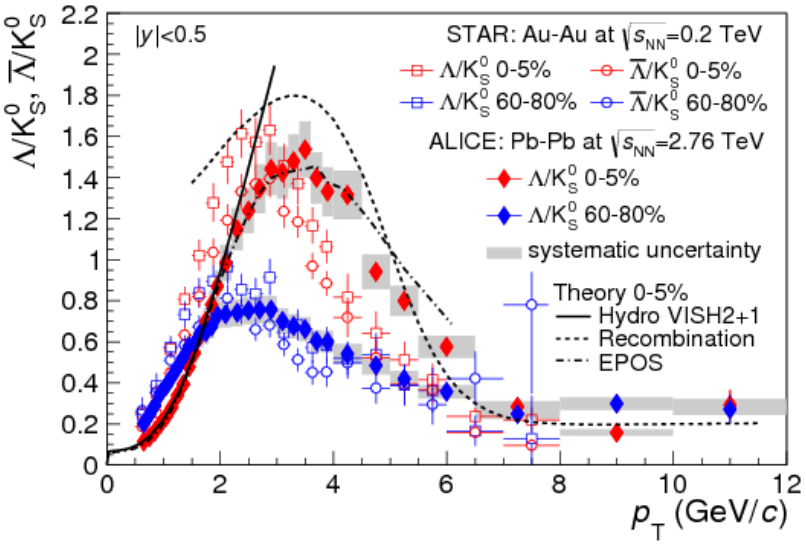

Figure 1.9: The $\Lambda/\mathrm{K}^0_{\mathrm{S}}$ ratio as a function of p_{T} [61] in central and peripheral Au–Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ MeV, and in central and peripheral Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV, in comparison to models [63–65]

To understand the interplay between hydrodynamics - where the different hadron masses would be responsible for the different flow patterns, and hadronisation (recombination) mechanisms - where the difference in the baryon-to-meson ratio is down

to the number of quarks constituting the hadron - resonances such as the $K^*(892)$ and $\phi(1020)$, which have masses similar to the proton, can also contribute important information. A flat $p/\phi(1020)$ ratio has been measured in Pb–Pb collisions [67, 68], suggesting that the difference in baryon-to-meson ratio could be down to hydrodynamics, and challenging the presumed role of recombination in the enhanced $\Lambda/\mathrm{K}^0_\mathrm{S}$ and p/π ratios.

In asymmetric collision systems (e.g. d–Au at RHIC, p–Pb at ALICE) a deficit of particle production is seen at low p_T, and an enhancement is seen at intermediate p_T. This effect (known as the 'Cronin Effect') has traditionally been explained as multiple soft scatterings before the hard scattering and subsequent fragmentation of the scattered parton, modifying the p_T distribution of the observed final-state hadron. This should not have a particle species dependence, since the particle species should only depend on the hadronisation of the scattered parton. However an enhancement of the proton yield in d–Au collisions relative to pp collisions, larger than that of protons or kaons and increasing with more central collisions, has been measured at RHIC (figure 1.10, left) [69]. A similar effect can be seen in the $\Lambda/\mathrm{K}^0_\mathrm{S}$ ratio as well as the p/π ratio measured by ALICE in p–Pb collisions at $\sqrt{s_\mathrm{NN}} = 5.02$ TeV [70]. The $\Lambda/\mathrm{K}^0_\mathrm{S}$ ratio can be seen in figure 1.10 (right) where the centrality classes 0-5 % and 60-80 % are shown for p–Pb collisions and the 0-5 % and 80-90% centrality classes are shown for Pb–Pb collisions for comparison.

The qualitative similarity between p–Pb and Pb–Pb collisions is likely driven by a common mechanism. Hadronisation through recombination of both soft and hard partons has been suggested to account for the proton enhancement, with such models describing these data well [71]. On the other hand, as will be discussed in the following pages, recent measurements of central p–Pb collisions and high multiplicity pp collisions at the LHC indicate that collective expansion may also be present in these systems, which may affect the p_T spectra in p–Pb collisions in a similar way to Pb–Pb collisions.

1.3.6 Collectivity

Collectivity in heavy-ion collisions has been studied since fixed target experiments in the late 1970's, and it relates to the physics governing the expansion of the system. 'Collective flow', which is defined as the correlation between the position of matter and its direction of flow, signals the presence of multiple interactions between the constituents of the medium created in the collision, and a greater flow generally signals that the matter comes closer to thermal equilibrium. Collective flow can be

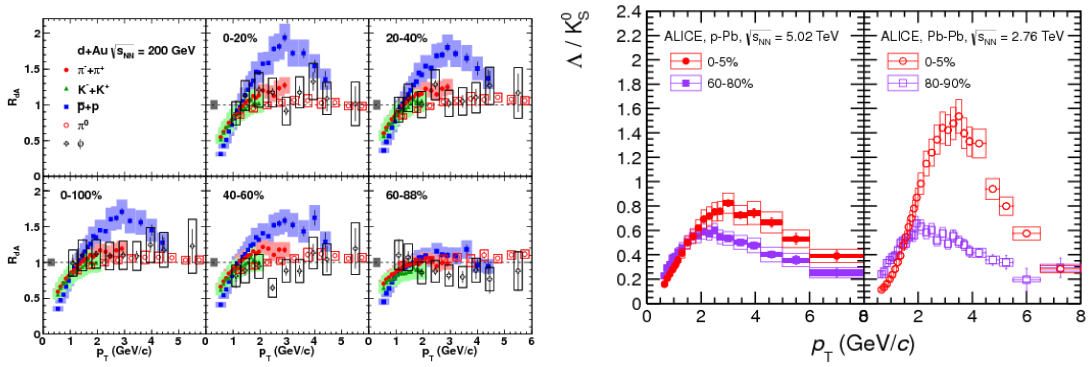

Figure 1.10: Left: R_{dAu} of pions, kaons, protons, π^0 and ρ particles as a function of p_{T} in different centrality intervals in d–Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV [69]. Right: The $\Lambda/\mathrm{K}^0_{\mathrm{S}}$ ratio in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV in the 0-5 % and 60-80 % centrality classes, and in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV in the 0-5 % and 80-90% centrality classes [70].

studied by expanding the azimuthal momentum, or the momentum distribution in the plane perpendicular to the beam direction ($r\phi$), in a Fourier series:

$$\frac{dN}{d\phi} = \frac{N}{2\pi}[1 + 2v_1\cos(\phi) + 2v_2\cos(2\phi) + ...] \tag{1.12}$$

where ϕ is the azimuthal angle between the particle and reaction plane, and v_n are the Fourier coefficients of the nth harmonic. A semi-central heavy-ion collision creates pressure gradients along the horizontal axis much larger than that in the vertical axis. This is depicted pictorially in figure 1.11. If the mean free path among the produced particles is larger than the system size (left), the particles escape with their initial momenta and their azimuthal distribution does not depend on their azimuthal angle. On the other hand, if the mean free path is much smaller than the size of the system, 'collectivity' can build up, and flow is enhanced along the horizontal axis. The amplitude of this enhancement is described by the magnitude of the second Fourier harmonic v_2, known as elliptic flow [72], which is sensitive to bulk properties of the system such as the viscosity and speed of sound, and can be used to probe the degree to which the system thermalises.

Elliptic flow has been measured at the LHC and RHIC [73–75], and attempts to derive the shear viscosity over entropy η/s [76] have found values close to the lowest theoretical bound obtained from calculations based on the strong coupling limit of conformal field theory, with the conjectured correspondence to anti-de Sitter space (AdS/CFT) [77]. In addition, higher order coefficients of the Fourier series in

equation 1.12 (v_3-v_6) have been measured to be non-zero [74, 78], and are ascribed to fluctuations of the positions of nucleons in the overlap region.

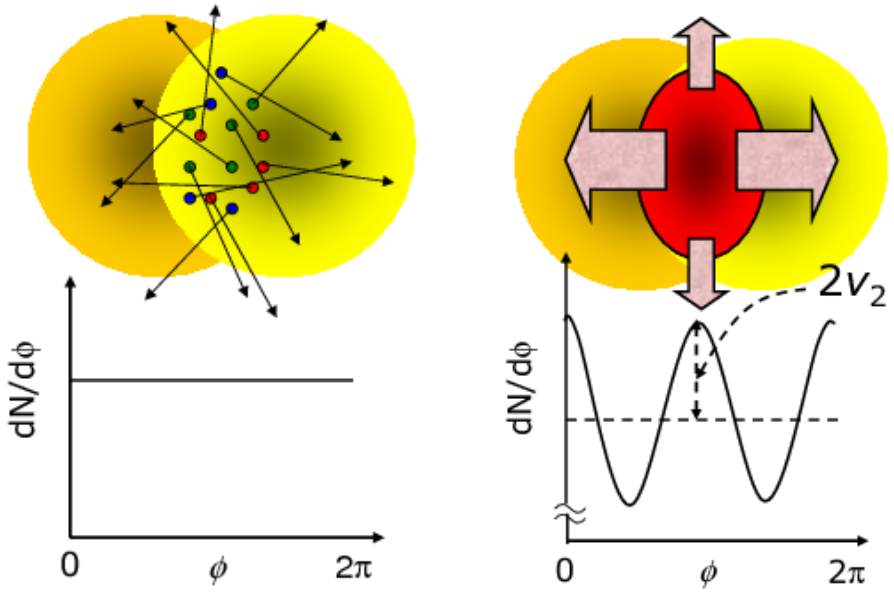

Figure 1.11: Normalised angular distribution $\mathrm{d}N/\mathrm{d}\phi$ of a semi-central heavy-ion collision. Figure taken from [79].

1.3.6.1 Collectivity in smaller systems

While there may be more than one parton-parton interaction in a proton-proton collision (in fact the 'hard' cross section exceeds the total pp cross section at LHC energies below a p_{T} scale of a few GeV [80]), they are usually treated as uncorrelated collisions with no interference between each interaction. In proton-lead collisions multiple nucleon-nucleon collisions occur but it was initially thought that the system would not live long enough to thermalise and undergo collective expansion. However, recent unexpected behaviour in the angular two-particle correlations have called for a rethink into whether some form of collectivity occurs in high multiplicity p–Pb and pp collisions.

The study of two-particle correlations by measuring the relative angles $\Delta\phi$ and $\Delta\eta$ between pairs of particles - a 'trigger' particle in a $p_{\mathrm{T,trig}}$ interval and an 'associated' particle in a $p_{\mathrm{T,assoc}}$ interval - show a 'near-side' jet peak at $\Delta\phi \sim 0, \Delta\eta \sim 0$, where both trigger and associated particle originate from the same parton, and a broader 'away-side' peak at $\Delta\phi \sim 0, \Delta\eta \sim 0$ extending over a wide η range, where the associated particle originates from the recoiling jet. In nucleus-nucleus collisions a

near-side 'ridge' extending over a large η range is explained to be due to modified jet-related correlations due to the collective expansion of the system. Recently these 'long-range correlations' have also been measured in pp and p–Pb collisions. Examples can be seen in figure 1.12. The left figure shows the two-particle angular correlations as a function of relative azimuthal angle $\Delta\phi$ and $\Delta\eta$ in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV measured by ALICE [81] where the peripheral collision distribution has been subtracted from the central collision distribution. The right figure shows the two-particle angular correlations in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS [82] in the highest multiplicity ($N_{ch} \geq 120$) events. The anisotropic flow coefficients can be extracted from these distributions and are found to be non-zero and quantitatively similar in magnitude for both p–Pb and pp collisions, at different energies, suggesting the ridge arises from a similar physical mechanism in both collision systems. The origin of this phenomena is still unknown, and attempts to describe it include initial-state effects (gluon saturation, colour connections) and final-state effects (multiple parton interactions, collective effects arising from the high-density system), see [83] for a recent review on these phenomena.

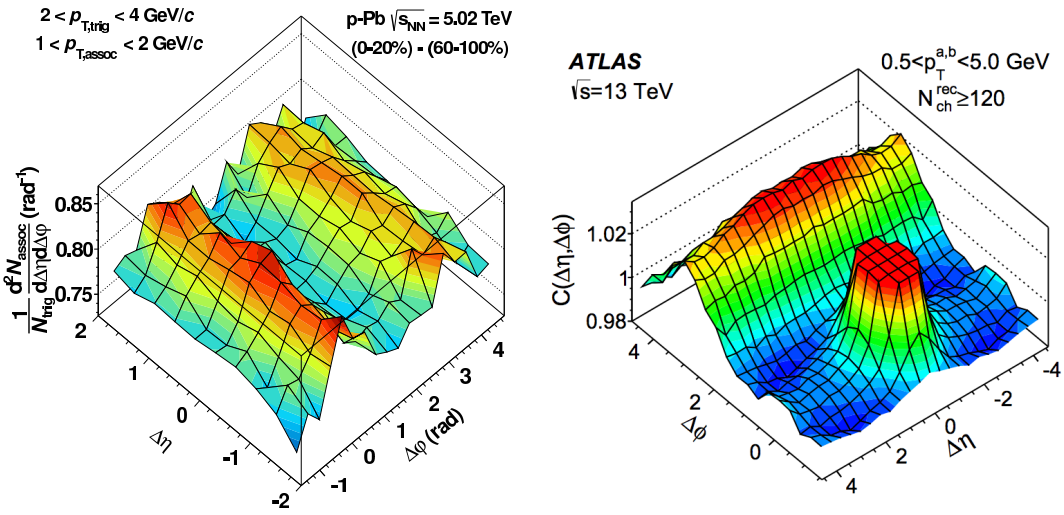

Figure 1.12: Left: The difference in the two-particle correlation function of high-multiplicity and low-multiplicity p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV as measured by ALICE [81]. Right: The two-particle correlation function in high-multiplicity pp collisions at $\sqrt{s} = 13$ TeV as measured by ATLAS [82]. In both cases $\Delta\phi$ and $\Delta\eta$ are the difference in ϕ and η, respectively, between a trigger particle and an associated particle.

1.4 Summary

In this chapter, it has been outlined that the asymptotically free nature of QCD, and the bulk properties of nuclear matter at high temperatures and energy densities leads to a point where quarks and gluons become deconfined. It has been illustrated that toy models and lattice QCD predict a phase transition temperature of around 160 MeV, and that by recreating these temperatures in the lab it is possible to study this state of matter. Heavy-ion collisions have been introduced as a way to achieve these goals, and some observables used to study the Quark-Gluon Plasma have been introduced. A selection of results from heavy-ion collisions have been presented, including measurements of the energy density and multiplicities, the thermal properties of the medium, ways to study the collective behaviour of the system, and possible tools to study hadronisation mechanisms. These discussed results have all been from the 'light' sector (that is, hadrons containing, and limited to, u, d and s quarks). The study of hadrons containing heavier quarks (i.e. charm c and beauty b) also offer a unique probe of the Quark-Gluon Plasma, as well as important insights into QCD, and the next chapter will focus on the motivations, theoretical description, and measurements of these particles.

Chapter 2

Heavy-flavour production - from pp to heavy-ion collisions

Open heavy-flavour (charm and beauty) production in pp collisions can be used to understand important aspects of QCD. In particular it is an important testing ground for state-of-the-art perturbative QCD predictions, and also offers a reference for measurements in different collision systems. Open heavy-flavour production in p–Pb collisions can be modified by various initial- and final-state effects, such as modifications to the initial state of the lead nucleus with respect to free protons, and multiple scattering that may occur in the final state (known as cold nuclear matter effects). Finally, heavy-flavour measurements in Pb–Pb collisions offer a unique way to probe the strongly-interacting matter created in such collisions, to determine properties of the medium and to probe hadronisation mechanisms. This chapter will present some theoretical details of heavy-flavour production in each of these collision systems, as well as some experimental results. A summary of the theory of heavy-flavour production is given in section 2.1, with a description of heavy-flavour production in pp collisions (section 2.1.1), p–Pb collisions (section 2.1.2) and Pb–Pb collisions (section 2.1.3). This will be followed by an overview of experimental results in section 2.2, detailing measurements in pp and p–Pb collisions (section 2.2.1) and Pb–Pb collisions (section 2.2.2). This chapter aims to put the measurements presented in this work in the wider context of the study of heavy-flavour production in different collision systems.

2.1 Theoretical description

2.1.1 Heavy-flavour production in pp collisions

In proton-proton collisions, the proton must be considered as a composite structure of quarks and gluons, which in high-energy collisions are given the name 'partons'. An inelastic proton-proton collision is therefore a superposition of one or more inelastic *parton-parton* interactions. For high-energy parton interactions the QCD coupling α_S is small, and as such the cross section for the production of a parton c in a *parton-parton* collision $ab \to c + X$ can be computed using perturbative techniques. Under the assumption that the parton-parton interaction occurs on much shorter timescales than any interactions that can occur before the collision among partons within the proton, or after the collision with the outgoing parton, the production of a high-p_{T} or high-mass hadron h in a *proton-proton* collision $AB \to h$ can be calculated using the QCD "factorisation theorem" [84], where

$$d\sigma^{hard}_{AB\to h} = f_{b/B}(x_1, Q^2) \otimes f_{a/A}(x_2, Q^2) \otimes d\sigma^{hard}_{ab\to c}(x_1, x_2, Q^2) \otimes D_{c\to h}(z, Q^2). \quad (2.1)$$

The assumption here is that any interaction between the parton belonging to the colliding proton, or any interaction during the evolution of the outgoing parton into their hadronic final state occur at longer distances, and these interactions can be factorised into the perturbative parton-parton scattering cross section $d\sigma^{hard}_{ab\to c}$, and two non-perturbative terms:

- A parton distribution function (PDF) $f_{a/A}(x, Q^2)$, which parametrises the non-perturbative dynamics of the proton. It represents the probability of finding a parton of flavour a with a momentum fraction $x = p_{parton}/p_{proton}$ of the proton A, at a given momentum transfer Q^2.
- A fragmentation function (FF) $D_{c\to h}(z, Q^2)$, which describes the probability that the outgoing parton c fragments into a hadron h with fractional momentum z of the initial parton.

The parton distribution functions of a proton and the fragmentation functions of a quark cannot be calculated perturbatively (since they are related to phenomena occurring at low energies/large time scales), and therefore rely on experimental input.

2.1.1.1 Heavy-quark production

At a partonic level, heavy-quark production can proceed via the interaction of two quarks, two gluons, or a quark and a gluon. At leading order (LO) there are two processes which contribute to heavy quark production; $q + \bar{q} \rightarrow Q + \bar{Q}$ and $g + g \rightarrow Q + \bar{Q}$. The Feynman diagrams for these processes are shown in figure 2.1. At Next-to-Leading Order (NLO) virtual one-loop corrections[1] are added to the previous processes, plus $2 \rightarrow 3$ processes: 1) $q + \bar{q} \rightarrow Q + \bar{Q} + g$, 2) $g + g \rightarrow Q + \bar{Q} + g$, 3) $g + q \rightarrow q + Q + \bar{Q}$, and 4) $g + \bar{q} \rightarrow \bar{q} + Q + \bar{Q}$. Examples of some of these diagrams are shown in figure 2.2.

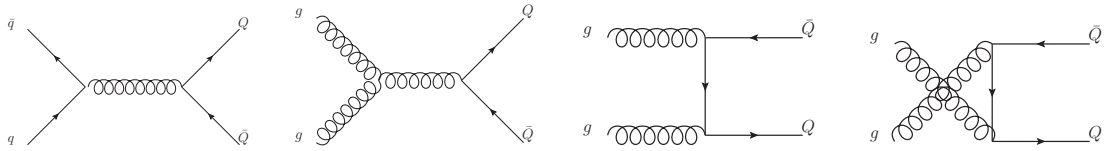

Figure 2.1: LO diagrams contributing to heavy-quark production.

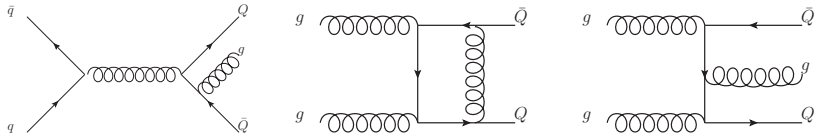

Figure 2.2: Examples of NLO diagrams contributing to heavy-quark production.

pQCD calculations introduce two scale factors - the *renormalisation scale* μ_r introduced when renormalising the strong coupling, and the *factorisation scale* μ_f which appears when factorising the long-distance and short-distance effects, to obtain the factorised cross section given in equation 2.1. These scales are usually chosen to be of the order of the hard process being calculated, and for the production of a heavy quark this can be taken as the mass of the quark. pQCD calculations are reliable at scales larger than the QCD scale $\mu_f, \mu_r > \Lambda_{QCD}$, and since the masses of heavy quarks, $m_c \simeq 1.5$ GeV/c^2 and $m_b \simeq 4.5$ GeV/c^2 are larger than Λ_{QCD}, pQCD is applicable to the production of heavy quarks down to low transverse momentum. There are however some considerations when calculating cross sections in different p_{T} regions, and several schemes have been developed for calculating heavy quark production.

The simplest scheme for calculating heavy-flavour production is the Fixed Flavour Number Scheme (FFNS), where the heavy quark is not considered as an active par-

[1]one-loop corrections account for the emission and absorption of a gluon by the incoming or outgoing partons. Its 'virtuality' refers to the fact that interacting gluons do not have to obey the energy-momentum relation. Since this allows particles to have different masses, virtual particles are also often referred to as being 'off the mass shell'.

ton in the proton. A sum over all possible subprocesses $a + b \rightarrow Q + X$ is performed in equation 2.1 for all considered active quarks a, b ($u, \bar{u}, d, \bar{d}, s, \bar{s}, g$ for charm and beauty production. The quarks $c, \bar{c}$ can be added for calculating beauty production). The range of applicability of FFNS is roughly $0 \leq p_{\mathrm{T}} \lesssim 5 \times m_Q$. Differential implementations of FFNS are often referred to as NLO MNR [85].

For large p_{T} relative to the heavy-quark mass a perturbative expansion ceases to be applicable, which can be dealt with by treating the outgoing quark like a massless parton, absorbing terms which would spoil the perturbative expansion into the PDF and FF, where the scale dependence is governed by the DGLAP equations (see section 2.1.1.2). This scheme is known as the Zero-Mass Variable-Flavour-Number Scheme (ZM-VFNS) and is applicable at large transverse momenta [86].

FFNS and ZM-VFNS are applicable in complimentary p_{T} regions. A scheme attempting to unify these two frameworks has been created, named the General-Mass Variable-Flavour-Number Scheme (GM-VFNS) [87, 88], which is valid from small to large transverse momenta.

The Fixed-Order plus Next-to-Leading-Logarithms calculation (FONLL) [89, 90] has also been developed, as a variant of GM-VFNS based on matching NLO massive (low p_{T}) and massless (high p_{T}) calculations. The final cross section is a convolution of different contributions and is calculated as

$$\mathrm{d}\sigma_{FONLL} = \mathrm{d}\sigma_{FO} + (\mathrm{d}\sigma_{RS} - \mathrm{d}\sigma_{FOM0}) \times G(m_Q, p_{\mathrm{T}}). \tag{2.2}$$

Here $\mathrm{d}\sigma_{FO}$ is the NLO calculation, performed with massive quarks (at fixed-order). The cross section calculated in the massless limit (resummed next-to-leading-log, or RS) is represented by $\mathrm{d}\sigma_{RS}$. Since there are terms shared by both approaches the fixed order terms at the massless limit are subtracted, given by the term $\mathrm{d}\sigma_{FOM0}$. $G(m_Q, p_{\mathrm{T}})$ is a matching function which is forced to approach 1 as $p_{\mathrm{T}} \rightarrow \infty$, whose functional form is arbitrary and given by $G(m_Q, p_{\mathrm{T}}) = p_{\mathrm{T}}^2/(p_{\mathrm{T}}^2 + a^2 m_Q^2)$, where the constant a is chosen to suppress the FO calculation at $m_{\mathrm{T}} > 5m_Q$.

In addition to the analytic calculations described above which only provide an accurate description of inclusively produced heavy hadrons, heavy-flavour production can be studied through Monte Carlo (MC) event generators which allow for a more complete description of the final state, as well as allowing for transport through simulated detectors. An example of a widely used MC generator is PYTHIA [91, 92], which includes hard interactions and soft interactions, parton distributions, initial- and final-state parton showers, multi-parton interactions, jet and quark fragmentation and hadronisation, and hadron decay. Hard processes in PYTHIA are performed at

at leading-order + leading-log accuracy - more recently NLO generators have been developed, including MC@NLO [93] and POWHEG [94, 95]. Generators such as HIJING [96] extend PYTHIA to include effects present in heavy-ion collisions such as nuclear shadowing (see section 2.1.2), multiple minijet production and jet quenching. EPOS [65], based on Gribov-Regge theory [97], and DIPSY, based on a coloured dipole picture (see [98] and references within), are examples of recent MC generators that have been developed with the aim of simulating and gaining deeper insight into the physics of pp and heavy-ion collisions, in particular to understand soft and semi-hard processes where perturbative QCD ceases to be applicable.

2.1.1.2 Parton distribution functions

Parton distribution functions are determined experimentally at a given energy scale, and the scale dependence of the PDFs is given by the DGLAP evolution [99–102] in terms of the calculable splitting functions, $P_{aa\prime}$, as

$$\frac{\partial f_{a/A}}{\partial \ln Q^2} = \sum_{a\prime=q,g} P_{aa\prime} \otimes f_{a\prime/A} \tag{2.3}$$

where the splitting functions can be expanded as a perturbative series in the strong coupling $\alpha_S(Q^2)$. In this way parton distribution functions determined from experimental data at lower energies can be scaled for higher energy regimes, such as those at the LHC. Many PDF sets exist (see e.g. [103–106]), calculated using a combination of hard-scattering input data, including fixed target and colliding beam deep-inelastic scattering (DIS) data to constrain the quark and gluon PDFs, Drell-Yan data to constrain high-x sea quarks, Tevatron jet data to constrain high-x gluon distributions, and W and Z data which provides access to different quark contributions [107].

2.1.1.3 Heavy-quark fragmentation

Fragmentation functions of heavy quarks have been studied in e^+e^- and electron-hadron collisions, where the clean environment allows for the determination of the state of the bare quark before hadronisation. Through fits to data the fragmentation functional form can be determined through some parametrisation. For example, the Peterson fragmentation function [108] is given by

$$\frac{dN}{dz} = \frac{1}{z[1-(1/z)-\epsilon_P/(1-z)]^2}. \tag{2.4}$$

Here ϵ_P is the only free parameter, and is expected to vary roughly as the ratio of the light quark mass (or sum of the two light quark masses for baryons) to the

heavy quark mass. To this effect, a harder fragmentation function, or smaller value of ϵ_P, is expected for lighter D mesons, meaning that a c quark fragmenting to a D meson will not lose much of its initial momentum, and the Λ_c^+ baryon is expected to have a softer fragmentation function, losing more of the initial quark momentum in hadronisation than a D meson would. Additionally b hadrons are expected to have a harder fragmentation function than charmed hadrons.

The probability that a quark will fragment into a given hadron h, given the name 'fragmentation fraction', is included in the fragmentation function. Measurements of the fragmentation fractions of charm quarks to hadrons (D^0, D^+, D^{*+}, D_s^+, Λ_c^+) have been made in e^+e^- collisions at B-factories, from Z decays, in photoproduction, in deep inelastic scattering and in pp collisions. A combined analysis of charmed hadron production from different experiments has been performed [109], and a summary of the fragmentation fractions $f(c \to h)$ from this analysis is shown in figure 2.3. The measurements are in good agreement within uncertainties, supporting the hypothesis that quark fragmentation is independent of production process (fragmentation universality). The violation of fragmentation universality has however been hinted at in the beauty sector [13], where differences between fragmentation fractions of b quarks to b baryons have been seen in $Z \to b\bar{b}$ decays in e^+e^- collisions, and and hadron collisions ($p\overline{p}$ collisions at the Tevatron, or pp collisions at the LHC), where a p_T dependence in the fragmentation fractions has also been observed in $p\overline{p}$ collisions by CDF [110] and pp collisions by LHCb [111], though the available precision does not allow a firm conclusion to be drawn.

Hadronisation of heavy quarks (as well as light quarks) is modelled in MC generators using phenomenological, non-perturbative models. The most widely used hadronisation model is the Lund string model [112, 113], where the potential between two colour-connected quarks moving apart is described by a colour string which breaks through creation of quark-antiquark pairs. New $q\bar{q}$ pairs are produced when it becomes more energetically favourable for the string to break into smaller strings, modelled through quantum tunnelling with the probability

$$\frac{\mathrm{d}P}{\mathrm{d}p_T} \propto \kappa \exp\left(-\frac{\pi m_T^2}{\kappa}\right) \tag{2.5}$$

where the transverse mass of the quarks is the quadratic sum of the quark mass m and transverse momentum p_T, $m_T = \sqrt{m^2 + p_T^2}$, and κ is the longitudinal energy density of the string, or the 'string tension', which is estimated from experimental data and lattice calculations to be $\kappa \approx 1$ GeV/fm for a $q\bar{q}$ string in vacuum [114].

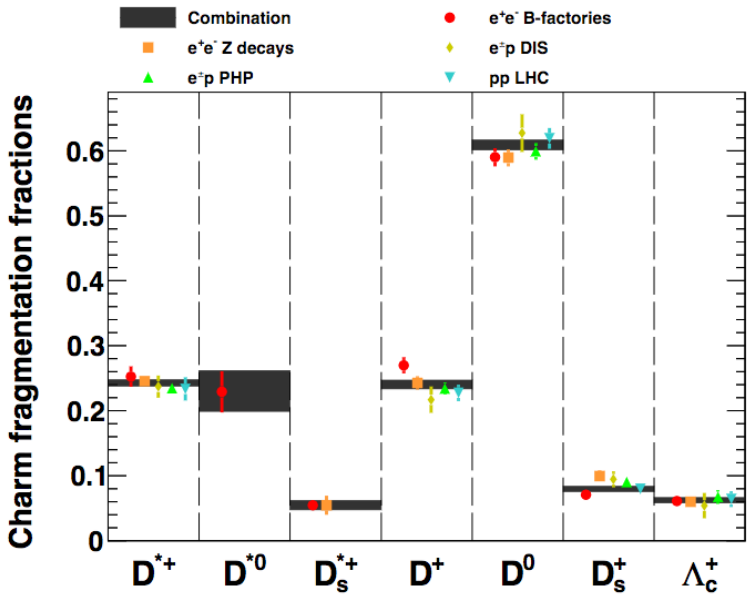

Figure 2.3: The charm-quark fragmentation fractions measured in different experiments. The global combination is shown with the shaded bands [109].

Fragmentation parameters can be determined from e^+e^- colliders, for example in the clean $e^+e^- \to Z \to q\bar{q}$ reaction, and directly applied to hadron colliders; any discrepancy between the results has to be due to phenomena not active at lepton colliders. The increased density of quarks and gluons is one effect in proton-proton collisions that could affect hadronisation. Colour reconnection models [115, 116] provide an extension on the string model where colour correlations can influence hadronisation.

Build on the Lund string model for hadronisation, colour reconnection occurs when colour strings begin to overlap. In a scenario without colour reconnection, a scattered parton at mid-rapidity may be colour connected with part of the remaining proton. With colour reconnection however, partons from two separate parton scatterings may become colour-connected, coming from a preference for the colour strings to be as short as possible. Since this gives a large transverse boost, this can increase the $\langle p_{\mathrm{T}} \rangle$. A sketch of this scenario is shown in figure 2.4. In addition, colour reconnection allows for 'baryon junctions', as depicted in figure 2.5. Since this topology is closely related to baryons, the result can be a baryon enhancement. Reconnections occur given some constrains, including SU(3) colour rules from QCD governing which reconnections are allowed, ensuring that the strings are causally connected, and taking into account the string length to decide whether a possible reconnection is favoured.

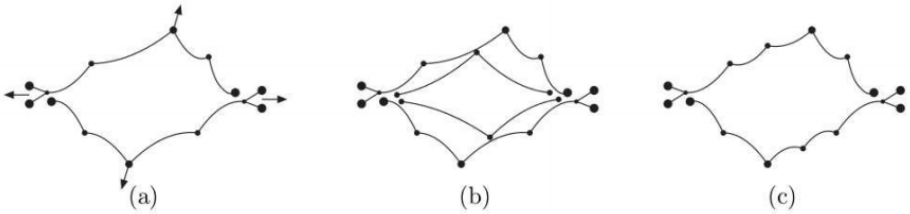

Figure 2.4: Sketch of how colour reconnection can occur depicting, from left to right, a) The outgoing partons of the first hard scattering color connected to the beam remnants, b) the second hard scattering, and c) color reconnected strings [117].

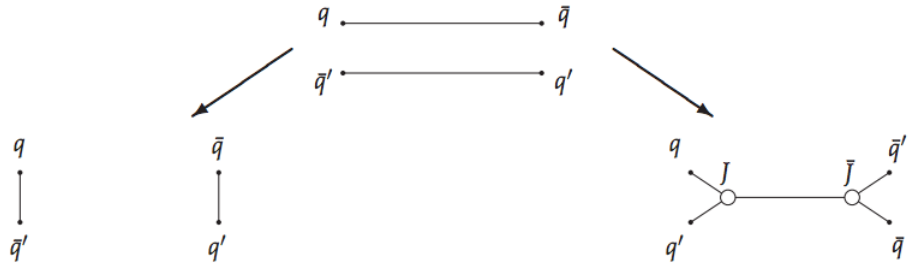

Figure 2.5: Sketch of possible colour reconnections of two $q\bar{q}$ pairs (top), including colour reconnection with the other quark(antiquark) (left), and the formation of a baryon junction (right) [116].

Rope hadronisation [118] is implemented in the MC event generator DIPSY [98], which involves interaction between strings, such that when several parton pairs are next to each other they can coherently interact to form a colour rope. The effects of the interaction is determined by SU(3) colour rules, and can end up with the strings forming a 'rope' which is hadronised with a higher effective string tension, and thus with more energy available for fragmentation. This results in more strange quarks and diquarks produced, leading to a higher number of baryons and more strangeness among the resulting hadrons.

Figure 2.6 shows the Λ/K ratio for MC generators PYTHIA8 which includes colour reconnection, and DIPSY which includes hadronisation via ropes [116]. Both models show an increase in the Λ/K ratio. It is noted that enhanced baryon production, as well as the distinctive 'peak' structure which moves towards higher p_{T} with larger multiplicities that is seen for colour reconnection is normally attributed to radial flow in Pb–Pb and p–Pb collisions [61, 70], and is an intriguing hint at a possible connection between flow (which there is evidence for in high-multiplicity pp collisions) and colour reconnection. This feature is also seen in the colour reconnection model

for the proton/pion ratio [117].

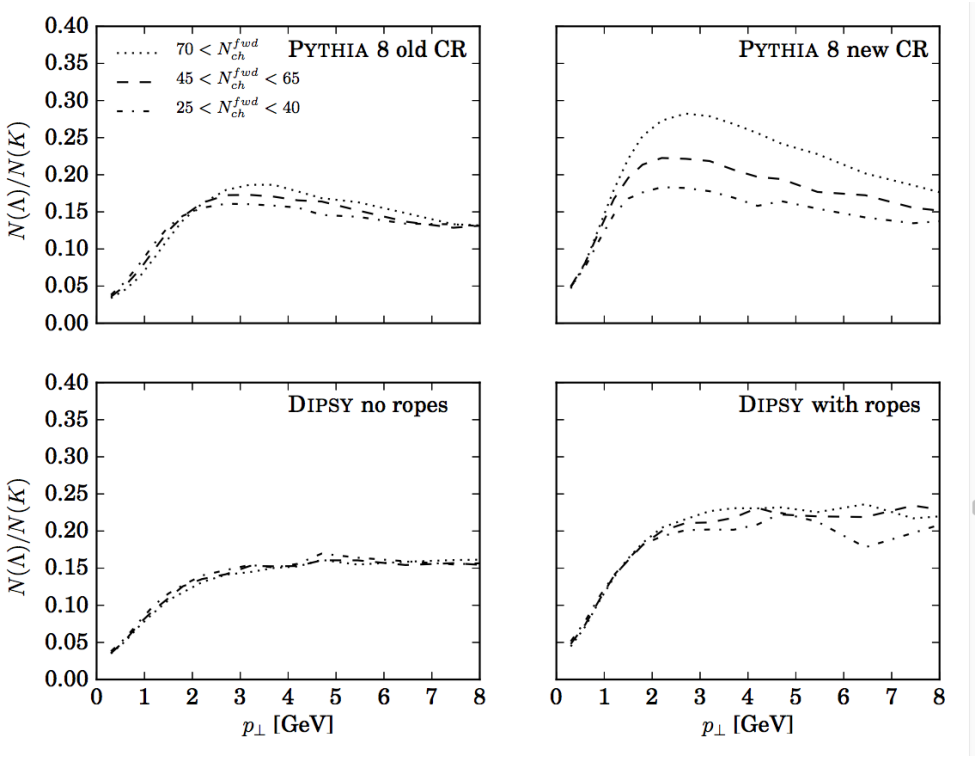

Figure 2.6: The Λ/K ratio in pp collisions from MC generators PYTHIA8 and DIPSY as a function of p_{T} in three forward-rapidity multiplicity intervals. On the right the colour reconnection models are shown, and on the left the previous hadronisation models are shown [116].

2.1.2 Heavy-flavour production in p–Pb collisions

In order to fully characterise effects present in Pb–Pb collisions, an understanding of all phenomena which could contribute to observables is needed. Effects due to the presence of the nuclei (such as modifications to the nuclear parton distribution function, or multiple scatterings of partons before and/or after the hard scattering) are known as cold nuclear matter (CNM) effects. These effects can be studied in proton-nucleus (p–A) collisions, or deuteron-nucleus (d–A) collisions, where the energy density is smaller than in AA collisions at similar collision energies. Recent results from the LHC and RHIC (see chapter 1) have called into question the idea that the medium created in these collisions truly is 'cold', which could increase the energy loss of partons traversing the medium, and collective behaviour of the expanding system could impart 'flow' to the heavy particles produced, modifying the heavy-flavour yield.

The production cross section of a heavy quark pair in (minimum-bias) p–A collisions assuming factorisation as in equation 2.1 is given by

$$d\sigma^{hard}_{pA\to Q\bar{Q}+X} = A \cdot f_i^N(x_1, Q^2) \otimes f_j^N(x_2, Q^2) \otimes d\sigma^{hard}_{ij\to Q\bar{Q}+X}(x_1, x_2, Q^2) \tag{2.6}$$

where A is the mass number of the nucleus and f^N is the PDF for the bound proton in the nucleus for partons of species $i(j)$, related to that of a free proton PDF by

$$f_i^N(x_i, Q^2) = R_i^N(x_i, Q^2) f_i(x_i, Q^2) \tag{2.7}$$

where $R_i^N(x, Q^2)$ is the nuclear modification to the free proton PDF f_i for a parton of species i. This factor stems from the fact that the structure of high-energy nuclei is different from a superposition of the non-interacting component nucleons, and as such the PDF of the nucleus is modified. A value less than 1 at a given fractional momentum x and momentum transfer Q^2 indicates a depletion of partons with respect to a free nucleon, while a value greater than 1 indicates an enhancement. An illustration of $R_i^A(x, Q^2)$ is shown in figure 2.7. The features in different x regions can be explained as follows:

- **Shadowing:** The timescales to produce a heavy quark pair $Q\bar{Q}$ are typically of the order of $\tau_c \sim 1/m_{Q\bar{Q}} \lesssim 0.1$ fm/c in the $Q\bar{Q}$ rest frame. However in the rest frame of the nucleus, this time $\tau_c \sim E_{Q\bar{Q}}/m^2_{Q\bar{Q}}$ can be larger than the nucleus. This leads to shadowing effects due to destructive interferences from multiple scattering on different nucleons [119].

- **Anti-shadowing:** At intermediate x, constructive interferences due to multiple scatterings dominate [119].

- **EMC effect:** the suppression of the nuclear PDF at large x ($x > 0.2$) is thought to originate from the fact that quarks in bound nucleons move through a larger confinement volume, and as such carry less momentum than quarks in free nucleons, though a complete theoretical description of the EMC effect is still elusive [120].

- **Fermi motion:** The enhancement of the nuclear PDF at high x is due to the fact that nucleons in a nucleus are moving with larger transverse momentum compared to a free proton.

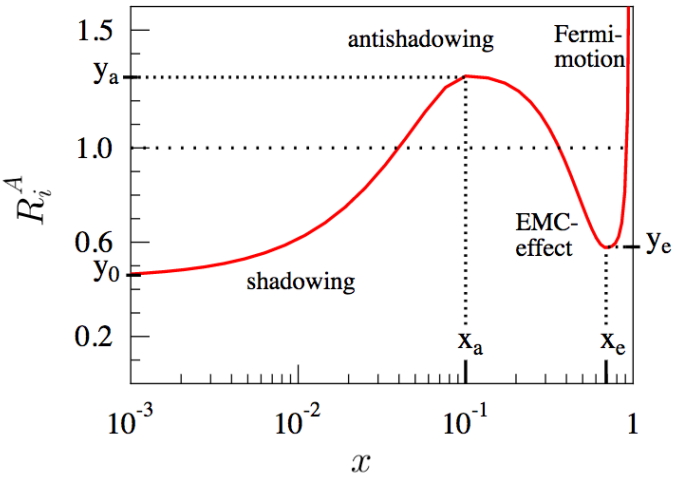

Figure 2.7: Sketch of the nuclear modification to the free proton parton density function [121]. See text for explanation of the different regions indicated on the plot.

Calculations have been performed to determine parametrisations of the nPDFs at NLO in pQCD [121], examples of which can be seen in figure 2.8, including contributions from valence and sea quarks, and gluons. The nuclear PDFs are determined through fits to experimental data, including deep inelastic lepton-nucleus scattering (DIS), Drell-Yan (DY) dilepton production and inclusive pion production measured at RHIC which improves in particular the gluon densities. Note that there are large uncertainties at low x, particularly for the gluon distributions, highlighting the need to constrain the PDFs at low x.

Heavy-flavour production can also be affected by multiple, softer, elastic scatterings of partons in the nucleus before or after the initial hard scattering. A parton, after each scattering, will acquire additional transverse momentum k_{T}, which grows with the number of elastic collisions. This extra momentum given to the hard parton in its final state can lead to an enhancement of the R_{pPb} at $p_{\mathrm{T}} \gtrsim 2$ GeV/c, expected to disappear as $p_{\mathrm{T}} \to \infty$, and is usually know as k_{T} broadening, or the Cronin Effect [122].

Finally, recent results indicating collectivity in high-multiplicity p–Pb collisions (see chapter 1) suggests that flow could be imparted on to heavy-flavour hadrons. A first attempt [123] to describe the expanding medium with a Blast-Wave fit [124, 125] which describes heavy-flavour decay electron data from RHIC well, predicts a small enhancement of D mesons of up to 30% at around $2 < p_{\mathrm{T}} < 4$ GeV/c for the 5% highest multiplicity p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. Assuming that the presence of

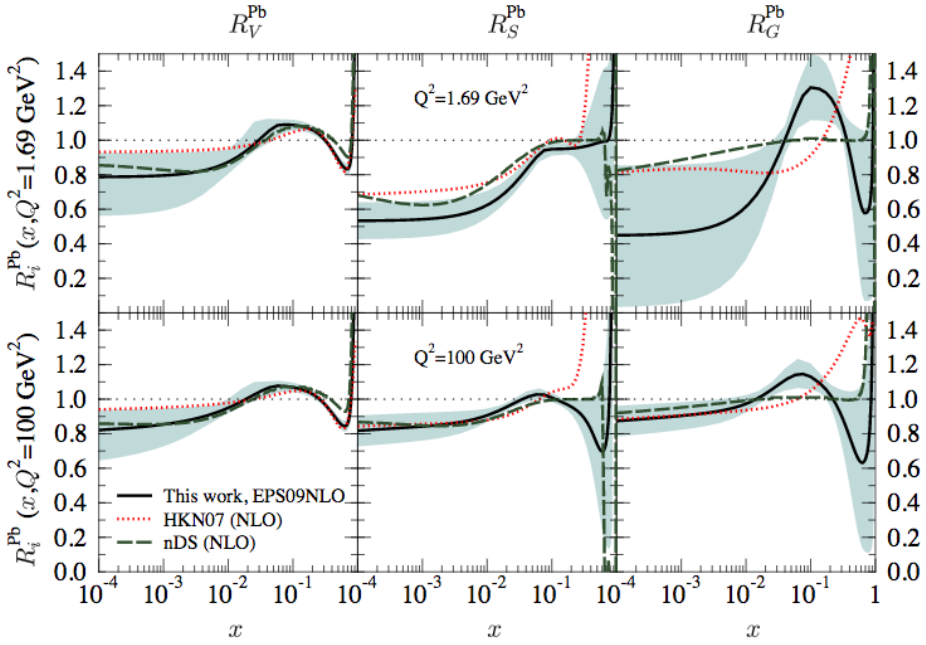

Figure 2.8: Comparison of the valence (left) and sea quark (middle), and gluon (right) modifications R_i^N at $Q^2 = 1.69\text{GeV}/c^2$ (top) and $Q^2 = 100\text{GeV}/c^2$ (bottom) from different analyses at NLO [121].

collectivity also indicates a 'hot' and deconfined medium is created, medium-induced energy loss may affect the heavy particle p_{T} spectra. Some calculations [126, 127] modifying the propagation and hadronisation of heavy quarks due to the presence of a deconfined medium indicate a moderate ($< 20 - 30\%$) suppression of D mesons at higher p_{T} caused by energy loss of the heavy quark.

2.1.3 Heavy-flavour production in Pb–Pb collisions

The study of heavy quarks offer a valuable probe of the medium produced in high energy heavy-ion collisions. Since the mass of charm and beauty quarks are much greater than the temperature of the system $m_Q >> T$, $Q\bar{Q}$ pairs are produced almost exclusively from the initial hard scattering processes on short time scales, meaning the heavy quark number after the initial hard scattering is essentially fixed. Heavy-quark production is calculable perturbatively, and any modification to their differential yield, e.g. their transverse momentum spectra, can be attributed to the modification of their kinematics by the medium. In addition, heavy quarks are not expected to become fully thermally equilibrated with the medium, and thus keep some memory of the interaction of the QGP. Heavy quarks are therefore very useful probes of the

QGP and of the underlying QCD force.

The production cross section of a hadron h in (minimum-bias) nucleus-nucleus collisions (with mass number A) can be expressed as

$$d\sigma^{hard}_{AA\to h+X} = A\cdot A\cdot f_i^N(x_i,Q^2)\otimes f_j^N(x_j,Q^2)\otimes d\sigma^{hard}_{ij\to Q\bar{Q}+X}(x_i,x_j,Q^2) \\ \otimes R_N^{\mathrm{QGP}}(E,m_Q,T,\alpha,...)\otimes D^N_{c\to h}(z,Q^2). \tag{2.8}$$

Here the production cross section $d\sigma^{hard}_{ij\to Q\bar{Q}+X}(x_i,x_j,Q^2)$ is calculated as in section 2.1.1.1, and the nuclear PDFs $f^N_{i,j}(x_i,j,Q^2)$ are calculated as in section 2.1.2. The two extra terms represent the effects of the quark-gluon plasma - R_N^{QGP} represents the modification of the heavy quark momentum due to energy loss in the medium and/or collective flow, and $D^N_{c\to h}(z,Q^2)$ represents the *medium-modified* fragmentation function due to hadronisation mechanisms which come into play at freeze-out, such as quark coalescence. The following subsections describe these mechanisms in more detail.

2.1.3.1 Heavy quark energy loss in the QGP

The energy loss ΔE of a high energy parton passing through the QGP as represented by R_N^{QGP} in equation 2.8 depends on the properties of the particle (energy E, mass m_Q) as well as the medium (temperature T, particle-medium interaction coupling α, thickness L). Energy loss in the medium primarily occurs via either radiative or collisional energy loss, where the total energy loss can be expressed as $\Delta E = \Delta E_{coll} + \Delta E_{rad}$. Collisional energy loss dominates at low particle momentum, and occurs through elastic scatterings with the medium constituents. Radiative energy loss occurs through inelastic scatterings within the medium through the emission of a gluon or photon, and instead dominates at higher momentum.

In-medium energy loss can be modelled using a Brownian motion framework[2] using the Fokker-Planck equation[3], which describes the time evolution of the heavy-quark momentum distribution under the influence of *drag* and *diffusion* forces. A simplified version of the Fokker-Planck equation [128] describes the time evolution of the heavy quark probability density f_Q, and is given as

$$\frac{\partial}{\partial t}f_Q(t,p) = \gamma\frac{\partial}{\partial p_i}\left[p_i f_Q(t,p)\right] + D_p\Delta_{\vec{p}}f_Q(t,p) \tag{2.9}$$

[2]In general, Brownian motion describes the random motion of a particle in a liquid or gas arising from collisions with fast-moving particles constituting the liquid or gas.

[3]The Fokker-Planck equation approximates the Boltzmann equation when the momentum transfer to the heavy quark is small relative to the total momentum of the heavy quark.

This framework introduces some variables used to describe the medium - the *drag coefficient* γ represents the fractional momentum loss per unit of time in the absence of fluctuations, and the *momentum diffusion coefficient* D_p represents the increase in the variance of the momentum distribution per unit time. The drag and diffusion coefficients are linked through the Einstein relation $D_p = m_Q \gamma T$ where m_Q is the mass of the heavy quark and T is the temperature of the medium. The *spatial diffusion coefficient* D_s which describes the broadening of the spatial distribution with time is also defined and is related to the drag and momentum coefficient through $D_s = T^2/D_p = T/(m_Q\gamma)$. The Fokker-Planck equation can be solved stochastically with a Monte Carlo simulation using the Langevin equation [129], where in the non-relativistic limit the change in position $\mathrm{d}\vec{x}$ and change in momentum $\mathrm{d}\vec{p}$ can be given by

$$\mathrm{d}\vec{p} = -(\gamma(p)\vec{p} + \xi)\mathrm{d}t \tag{2.10}$$

$$\mathrm{d}\vec{x} = \frac{\vec{p}}{m_Q}\mathrm{d}t \tag{2.11}$$

The change in momentum is given by a differential equation with a 'deterministic' part and a 'stochastic' part. The deterministic part is realised by the drag term γ, and the stochastic part is realised by the term ξ which is sampled from a Gaussian probability distribution with width proportional to the momentum diffusion coefficient D_p [129].

For radiative energy loss processes ($2 \to 3$ processes), the difference in the quark and gluon couplings has to be taken into account. The relative amplitude of the 3 possible vertices are determined by the colour, or 'Casimir' factors C_R, and are $C_R = N_C$ for $g \to gg$, $C_R = (N_C^2 - 1)/2N_C$ for $g \to qg$ and $C_R = 1/2$ for $g \to q\bar{q}$, where $N_C = 3$ is the number of colours in QCD. This means that the average number of gluons radiated by a gluon is 9/4 times higher than that radiated by a quark [130]. Radiation off a heavy quark is also suppressed at angles smaller than the ratio of the quark mass m_Q to its energy E. This effect is known as the 'dead cone effect' [131], and it leads to a suppression of radiative energy loss which increases for heavier quarks. This leads to a hierarchy of flavour-dependent radiative energy loss, $\Delta E(g) > \Delta E(q) > \Delta E(c) > \Delta E(b)$ [130].

Interaction rates are determined through the calculation of the scattering matrix (or cross section) of a given radiative or collisional process, from which the drag and diffusion coefficients can be calculated. Various approaches have been developed to calculate the heavy-quark interactions within the medium, including pQCD

calculations of radiative and collisional energy loss, as well as non-perturbative approaches employing lattice-QCD calculated heavy-quark potentials and/or calculations of transport coefficients through a T-matrix approach, and the calculation of transport coefficients using the AdS/CFT conjecture. For full reviews of these types of calculations, see [128, 132].

Phenomenologically, the nuclear modification factor R_{AA} can be used to experimentally test different scenarios of heavy-quark energy loss, and help constrain the transport properties of the medium. The elliptic flow v_2 can be used to test the level to which heavy quarks thermalise in the medium. In addition, the v_2 can be used to test the path length dependence of in-medium energy loss - since semi-central collisions lead to different path lengths for heavy quarks in the azimuthal plane ϕ, the amount of energy loss in-plane and out-of-plane will also help constrain energy loss models.

2.1.3.2 Heavy-quark hadronisation in the QGP

As a reminder, in heavy-ion collisions a hadronisation mechanism other than fragmentation (where a quark or diquark will appear from the vacuum to form a meson or baryon) has been proposed. Coalescence, or recombination, describes the formation of hadrons through the coalescence of quarks close to each other in position and momentum space [62]. The measurement of the Λ_{c}^{+} baryon in heavy-ion collisions has been proposed as a probe of hadronisation [133]. In particular, if quark coalescence plays a significant role in the hadronisation of charm quarks, an enhancement of the baryon-to-meson ratio $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ is expected [134, 135]. The measurement in the charm sector would further constrain open questions from the strange $\Lambda/\mathrm{K}_{\mathrm{S}}^{0}$ ratio and interpretations based on the interplay between hadronisation mechanisms and flow (see section 1.3.5). Furthermore, it has been predicted that binary, coloured bound states (gg, qg, qq) may exist within the QGP [136], and coalescence also occuring between a charm quark and a 'diquark' state ud would create a further enhancement of the baryon-to-meson ratio [134]. This idea has been developed further in [135], where the effect of coalescence on the beauty baryon-to-meson ratio $\Lambda_{\mathrm{b}}^{0}/\mathrm{B}$ was also studied though this is not described here. Λ_{c}^{+} formation via coalescence could also affect the heavy-flavour decay electron spectrum, since the branching ratios for Λ_{c}^{+} to electrons is much smaller than that for D mesons, and a Λ_{c}^{+} enhancement had been proposed as a possible explanation for the large heavy-flavour decay electron suppression in Au–Au collisions seen at RHIC [137, 138].

The transverse momentum spectrum of a meson or baryon $\frac{\mathrm{d}N_M}{\mathrm{d}\boldsymbol{p}_M}$ or $\frac{\mathrm{d}N_B}{\mathrm{d}\boldsymbol{p}_B}$, formed through coalescence can be calculated from the individual quark transverse momentum distributions $\frac{\mathrm{d}N_i}{\mathrm{d}\boldsymbol{p}_i}$, derived from the 'GKL' implementation explained in [62] and given in [135] for mesons and baryons as

$$\frac{\mathrm{d}N_M}{\mathrm{d}\boldsymbol{p}_M} = g_M \frac{(2\sqrt{\pi}\sigma)^2}{V} \int \mathrm{d}\boldsymbol{p}_1 \mathrm{d}\boldsymbol{p}_2 \frac{\mathrm{d}N_1}{\mathrm{d}\boldsymbol{p}_1} \frac{\mathrm{d}N_2}{\mathrm{d}\boldsymbol{p}_2} \times \exp(-\boldsymbol{k}^2\sigma^2)\delta(\boldsymbol{p}_M - \boldsymbol{p}_1 - \boldsymbol{p}_3) \tag{2.12}$$

$$\begin{aligned}\frac{\mathrm{d}N_B}{\mathrm{d}\boldsymbol{p}_B} = g_B \frac{(2\sqrt{\pi})^6(\sigma_1\sigma_2)^3}{V^2} \int \mathrm{d}\boldsymbol{p}_1 \mathrm{d}\boldsymbol{p}_2 \mathrm{d}\boldsymbol{p}_3 \frac{\mathrm{d}N_1}{\mathrm{d}\boldsymbol{p}_1} \frac{\mathrm{d}N_2}{\mathrm{d}\boldsymbol{p}_2} \frac{\mathrm{d}N_3}{\mathrm{d}\boldsymbol{p}_3} \\ \times \exp(-\boldsymbol{k}_1^2\sigma_1^2 - \boldsymbol{k}_2^2\sigma_2^2)\delta(\boldsymbol{p}_B - \boldsymbol{p}_1 - \boldsymbol{p}_2 - \boldsymbol{p}_3)\end{aligned} \tag{2.13}$$

where it is assumed heavy quarks with transverse momentum $\mathbf{p_1}$ and light quarks with transverse momenta $\mathbf{p_2}$ (and $\mathbf{p_3}$) are produced uniformly distributed in a cylinder of volume V, $g_{M,B}$ is a statistical factor accounting for the degeneracy of the coloured quarks, σ is the 'width parameter' related to the radius of the produced meson or baryon, and $\mathbf{k}$ is the relative transverse momentum between the heavy quark and light antiquark. The production of a baryon from a heavy quark and a bound diquark can be expressed similarly to equation 2.12.

In Ref. [135] the baryon-to-meson ratio $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ (and $\Lambda_\mathrm{b}^0/\mathrm{B}$) is determined at RHIC energies (Au–Au collisions, $\sqrt{s_{\mathrm{NN}}} = 200$ GeV) from a purely thermal model where the integrated $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ ratio comes out as $\left(\frac{\Lambda_\mathrm{c}}{\mathrm{D}^0}\right) \simeq 0.24-0.27$, and from a coalescence model considering coalescence of a charm quark with two quarks, and a charm quark with a [ud] diquark state. In the coalescence model, using a thermal initial momentum distribution for light quarks and diquarks, taking the heavy quark distribution from pp collisions of the same energy, including heavy quark energy loss, and including fragmentation and resonance contributions the final baryon-to-meson ratios are

$$\frac{\Lambda_\mathrm{c}^+}{\mathrm{D}^0} = 0.83 \quad \text{(three-quark)} \quad , \quad \frac{\Lambda_\mathrm{c}^+}{\mathrm{D}^0} = 1.34 \quad \text{(diquark)}. \tag{2.14}$$

For the three-quark model this leads to $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ ratio around 6.4 times greater than the predictions from PYTHIA [91], and around 3.1 times larger than the purely thermal model. A further enhancement in $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ can be seen for the diquark model, where the $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ ratio is around 10.3 times larger than the predictions from PYTHIA and around 5.0 times larger than the purely thermal model. Figure 2.9 shows the Λ_c^+ and D^0 p_T spectrum, and the $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ ratio at RHIC energies at mid-rapidity as a function of p_T. A large difference can be seen for both the diquark and three-quark

model up to ~ 6 GeV/c, with a difference in the diquark and three-quark model up to $\sim 3-4$ GeV/c. This highlights the motivation for a Λ_c^+ measurement down to low p_T.

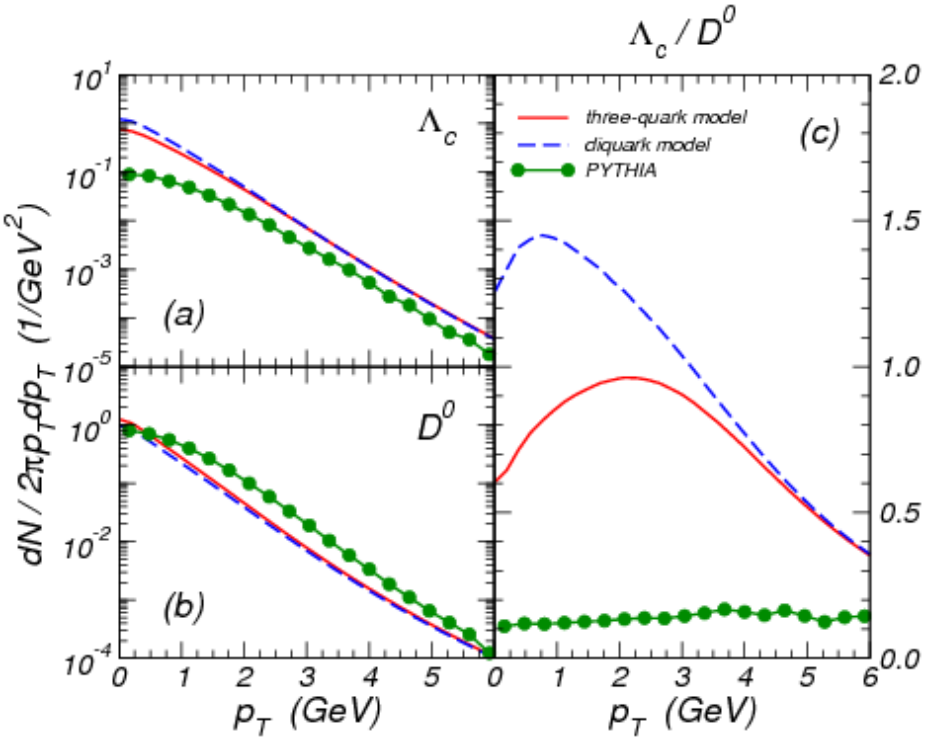

Figure 2.9: Spectra of (a) Λ_c^+ and (b) D^0, and (c) the ratio Λ_c^+/D^0 at $|y| \geq 0.5$ for central Au-Au collisions at $\sqrt{s_{NN}} = 200$ GeV, including the three-quark and diquark coalescence models, and results from PYTHIA [135]. Note that 'Λ_c/D^0' in the figure refers to the ratio of *particles*, as refered to as Λ_c^+/D^0 in the text.

The diquark enhancement can also be thought of perhaps more simply in terms of the probabilities of quark collisions in the medium. The Λ_c is formed from a three-body collision among the c, u and d quark constituents if there are no diquark states in the medium. However, if $[ud]$ diquark states exist this becomes a two-body collision, which is statistically more likely to occur.

It is noted that further studies into Λ_c^+ propagation in the hadronic phase after chemical freeze-out have determined that the Λ_c^+ spectrum should be relatively unaffected by this phase [139–141], thus allowing one to ignore this part of the system's evolution when interpreting the modification of the Λ_c^+ yield.

2.2 Experimental Results

Experimentally there are different techniques to study open heavy-flavour production. A full exclusive reconstruction of the heavy-flavour particles is possible, for example the reconstruction of $B^0 \rightarrow J/\psi\phi$, $D^0 \rightarrow K^-\pi^+$, or as described in detail in this thesis, $\Lambda_c^+ \rightarrow pK^-\pi^+$. Selection of (semi)-inclusive decays is also possible, an example is where a J/ψ a few hundred μm from the primary vertex is reconstructed and supposed to come from a b-decay. Another technique is detection of leptons (e, μ) from inclusive heavy-flavour decays, by subtracting known background sources (electrons from so-called Dalitz decays, e.g. $\pi^0 \rightarrow e^+e^-\gamma$, photon conversions $\gamma \rightarrow e^+e^-$, light hadron decays, J/ψ, $W^{\pm}$, etc.) from the inclusive electron spectra. Finally, one can reconstruct c- and b-jets, where one can either search for a reconstructed heavy-flavour decay within a jet, or use properties of the jet (e.g. displaced vertices) to distinguish between light-flavour jets (i.e. u, d, s, and in the case of b-jets, c). This section will outline some of the many recent open heavy-flavour measurements - here I mainly concentrate on measurements of fully reconstructed heavy-flavour decays (focusing mainly on exclusive reconstruction of open charm decays), but for a full recent review of measurements in the RHIC and LHC era see [132, 142].

2.2.1 Experimental results in pp and p–Pb collisions

The p_{T}-differential cross section of D^0, D^+, D^{*+} and D_s^+ mesons [143–145], heavy-flavour decay muons [146, 147], heavy-flavour decay electrons [148, 149] and beauty-decay electrons [150, 151] have been measured by ALICE in pp collisions at $\sqrt{s} =$ 2.76 TeV and $\sqrt{s} = 7$ TeV and are found to be compatible with pQCD predictions including calculations from FONLL [152], GM-VFNS [88, 153] and the k_T-factorisation scheme [154]. Preliminary cross section measurements of D mesons have also been performed at $\sqrt{s} = 5.02$ TeV, $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV. Examples of the D^0, D^+ and D^{*+} cross sections at different energies are shown in figure 2.10 in comparison to theoretical predictions. It is pointed out that the D meson cross sections at $\sqrt{s} = 5.02$ TeV were used to validate the rescaled 7 TeV cross section used as the reference for the preliminary R_{AA} measurement in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [155, 156]. While relying on theoretical input to rescale the 7 TeV reference to 5.02 TeV, the 7 TeV result offers smaller uncertainties than the 5.02 TeV measurement.

Charm production has also been studied in pp collisions at RHIC [157–159] and in $p\overline{p}$ collisions at Tevatron [160], and at lower energy fixed-target experiments [161].

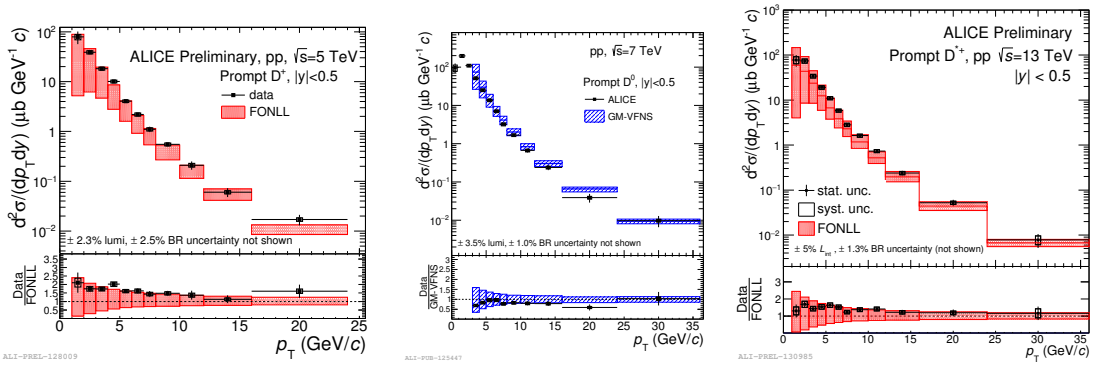

Figure 2.10: p_{T}-differential cross section measurements of prompt D^{+} at $\sqrt{s}$ = 5.02 TeV (left), D^{0} at $\sqrt{s}$ = 7 TeV (centre) and D^{*+} at $\sqrt{s}$ = 13 TeV (right), compared to predictions from FONLL [152] and GM-VFNS [88, 153].

The scaling of the total inclusive charm production cross section in pp and p–A collisions as a function of collision energy $\sqrt{s}$ is shown in figure 2.11. The total cross section is described within uncertainties by NLO pQCD calculations [162] with the measurements falling on the upper edge of the theoretical uncertainty.

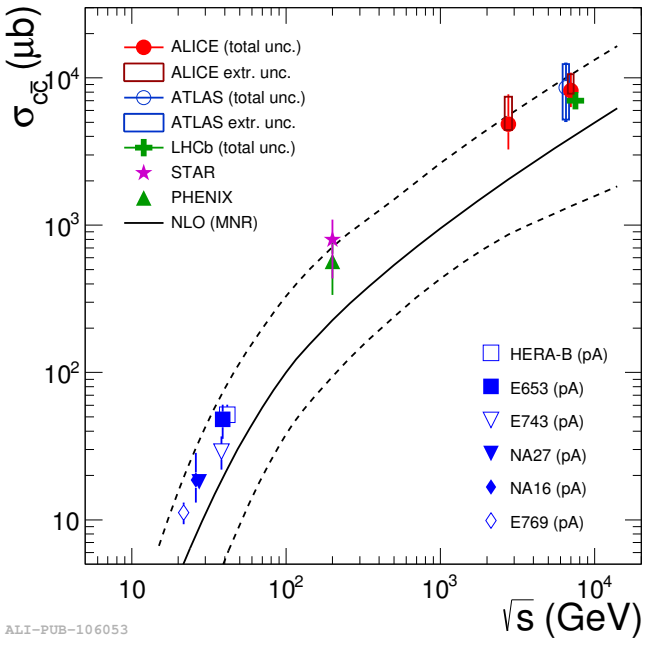

Figure 2.11: Total inclusive charm production cross section in nucleon–nucleon collisions as a function of $\sqrt{s}$ [144]. Data are from pA collisions for $\sqrt{s} < 100$ GeV and from pp collisions for $\sqrt{s} > 100$ GeV. Data from pA collisions were scaled by $1/A$. Results from NLO pQCD calculations (MNR [162]) and their uncertainties are shown as solid and dashed lines.

The nuclear modification factors R_{pPb} of D mesons (including $\mathrm{D}^0, \mathrm{D}^+$, D^{*+} and D_s^+) have been measured at ALICE at mid-rapidity [144], and are found to be compatible with each other. Figure 2.12 shows the averaged $\mathrm{D}^0, \mathrm{D}^+$ and D^{*+} R_{pPb}. Within uncertainties, the measured D-meson R_{pPb} is compatible with unity. This confirms the large suppression of D mesons in Pb–Pb collisions is indeed a final-state effect due to the interaction with the hot, deconfined medium (see next section). The D-meson R_{pPb} is compared to models - in the left panel models are shown which include only cold nuclear matter effects, including NLO pQCD predictions [162] with EPS09 parameterisation of the nuclear parton distribution function (nPDF) [121], a leading-order pQCD calculation including shadowing, k_T broadening and cold nuclear matter energy loss [163], a model based on the Colour Glass Condensate [164], and a higher-twist calculation based on incoherent multiple scattering [165]. All models are consistent with the data except that based on incoherent multiple scattering, which is disfavoured by the data at low p_{T}. The right panel shows a comparison with models which assume a QGP is formed - the DUKE model [127] includes both collisional and radiative processes, and the POWLANG model [126] includes only collisional processes, considering two different transport coefficients. While the data are consistent with the models within uncertainties, the shape of the R_{pPb} slightly disfavours these models. The comparison with models shows the motivation for a precise measurement at low p_{T}, since this is where the models deviate from each other. The D^0 meson R_{pPb}

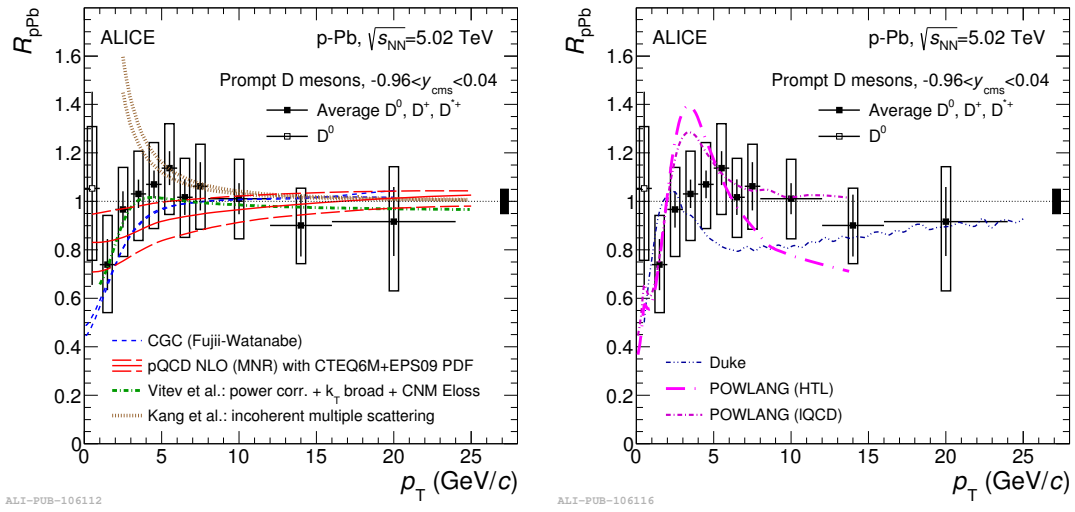

Figure 2.12: The nuclear modification factor R_{pPb} of D mesons in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV as a function of p_{T}, averaged over all D meson species [144], compared with models that include cold nuclear matter effects [121, 162–165] (left) and models that include a small QGP formation [126, 127] (right).

has also been measured at LHCb in the rapidity region $2.5 < y < 4$ [166], allowing different Bjorken-x regions of the nuclear PDF to be probed, and is found to be in good agreement with a NLO prediction computed using the EPS09 nPDF [121].

The multiplicity dependence of the nuclear modification factor has also been studied for D mesons by measuring the 'centrality-dependent' nuclear modification factor $Q_{\rm pPb}$ given by

$$Q_{\rm pPb} = \frac{1}{\langle T_{\rm pPb}^{cent} \rangle} \frac{{\rm d}N_{\rm pPb}^{cent}/{\rm d}p_{\rm T}}{{\rm d}\sigma_{\rm pp}/{\rm d}p_{\rm T}} \tag{2.15}$$

where ${\rm d}N_{\rm pPb}^{cent}/{\rm d}p_{\rm T}$ is the D meson yield in p–Pb collisions in a given centrality class, ${\rm d}\sigma_{\rm pp}/{\rm d}p_{\rm T}$ is the D meson cross section in pp collisions and $\langle T_{\rm pPb}^{cent} \rangle$ is the average nuclear overlap function in a given centrality class, equal to the number of binary nucleon–nucleon collisions divided by the inelastic nucleon–nucleon cross section. Figure 2.13 shows the $Q_{\rm pPb}$ measured in p–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV [167] where the centrality is estimated [40] using the energy deposited in the Zero Degree Neutron Calorimeter (ZNA). The $Q_{\rm pPb}$ in all centrality classes are seen to be consistent with each other and with unity, suggesting there is no centrality dependence in the nuclear modification of D mesons within the experimental uncertainties.

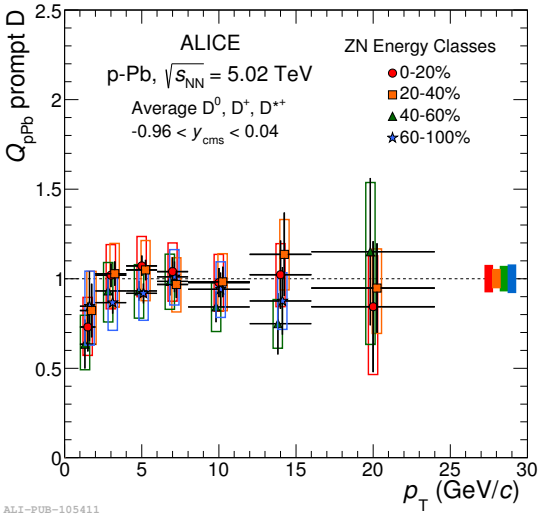

Figure 2.13: The averaged D meson centrality dependent nuclear modification factor $Q_{\rm pPb}$ in p–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV in the centrality classes 0-20%, 20-40%, 40-60% and 60-100%, using the ZNA centrality estimator [167].

While there are a wealth of charmed meson and heavy-flavour lepton production results from hadron-hadron colliders, there are relatively few charmed baryon production measurements. The $\Lambda_{\rm c}^+$ baryon $p_{\rm T}$ and y differential cross section has been

measured by LHCb at $\sqrt{s} = 7$ TeV. Shown in figure 2.14 (left) is the p_{T}-differential cross section of the Λ_{c}^{+} from $2-8$ GeV/c in the rapidity range $2 < y < 4.5$ [168]. The LHCb measurement is compared to predictions from GM-VFNS, and the prediction is seen to be in good agreement with the measured data, though the data is at the upper edge of the theoretical uncertainty. Previous measurements of Λ_{c}^{+} production in hadron-hadron collisions have been at much lower centre-of-mass energies [169, 170].

It is also noted here that Λ_{c}^{+} production has been studied in ep collisions at HERA [171] and in $e^{+}e^{-}$ collisions at energies around the Υ resonance - including at CLEO at $\sqrt{s} = 10.55$ GeV [172, 173], at ARGUS at $\sqrt{s} = 10.2$ GeV [174], at Belle at $\sqrt{s} = 10.6$ GeV [175] and at BaBar at $\sqrt{s} = 10.54$ GeV [176]. These measurements were based on the differential momentum spectrum $\mathrm{d}\sigma/\mathrm{d}x_p$ where x_p is the ratio of the hadron momentum to the maximum momentum available to this hadron. Figure 2.14 (right) shows the differential Λ_{c}^{+} production rate (where prompt and non-prompt Λ_{c}^{+} are not separated) normalised by the number of $e^{+}e^{-} \rightarrow \mathrm{q\overline{q}}$ events measured by BaBar, CLEO and BELLE [176]. All measurements are seen to agree within uncertainties. The fragmentation function used in GM-VFNS is evolved from fits to $e^{+}e^{-}$ collider data.

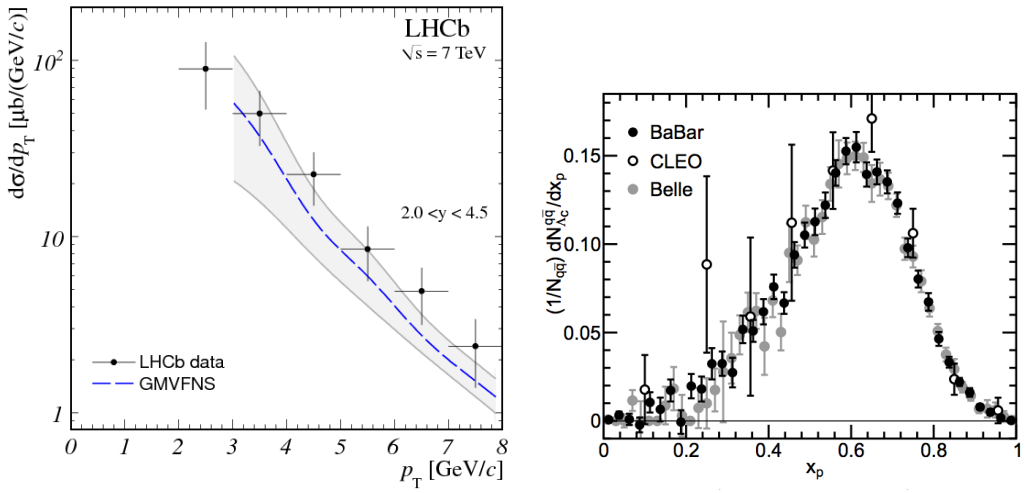

Figure 2.14: Left: The p_{T}-differential cross section of Λ_{c}^{+}, compared with GM-VFNS calculations [168]. Right: the x_p-differential Λ_{c}^{+} production rate per $e^{+}e^{-} \rightarrow \mathrm{q\overline{q}}$ event as measured by BaBar, CLEO and BELLE [176].

ALICE has also measured the azimuthal correlations of heavy-flavour particles with charged particles, and studied D meson production as a function of the event multiplicity in pp and p–Pb collisions. For further information on heavy-flavour

results from ALICE in pp and p–Pb collisions, from Run 1 at the LHC, see [2] and references within.

2.2.2 Experimental results in Pb–Pb collisions

The nuclear modification factor $R_{\rm AA}$ of $\rm D^0, D^+, D^{*+}$ and $\rm D_s^+$ has been measured by ALICE in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV [177, 178]. The left panel of figure 2.15 shows the $R_{\rm AA}$ for D mesons (averaged) and charged pions measured by ALICE, and non-prompt J/ψ coming from B meson decays measured by CMS, as a function of the number of participant nucleons N_{part} [178]. The pion $R_{\rm AA}$ is consistent with the D meson $R_{\rm AA}$, and the J/ψ $R_{\rm AA}$ is larger than both the D meson and pion $R_{\rm AA}$. While the quark energy-loss hierarchy described in section 2.1.3.1 is clearly seen for charm and beauty it is not immediately apparent for pions and D mesons. However an interpretation based on the colour-charge and quark-mass energy-loss dependence is not trivial, since it should take into account the different fragmentation functions (harder for charm quarks compared to light quarks) and the fact that at low $p_{\rm T}$ soft production processes could modify the pion yield. The right panel of figure 2.15 shows the $R_{\rm AA}$ of D mesons as a function of $p_{\rm T}$ measured by ALICE [177] compared to theoretical predictions, where the heavy quark interaction with the medium is calculated from pQCD inspired models which include radiative and/or collisional energy loss (BAMPS, POWLANG, MC@sHQ and Cao, Qin, Bass), lattice QCD models (TAMU), or a quasi-particle model (PHSD) (see [142] and references within). All models shown use a Fokker-Planck approach or similar to simulate the heavy quark transport (see section 2.1.3.1), and include an underlying hydrodynamic expansion of the system.

The measurement of the elliptic flow (v_2) of heavy-flavour particles also offers further insight into the transport properties of the medium. At low $p_{\rm T}$ it measures the extent to which heavy quarks thermalise in the medium, at low-to-intermediate $p_{\rm T}$ it is expected to be sensitive to the heavy quark hadronisation mechanism, and at high $p_{\rm T}$ the v_2 is also sensitive to the path-length dependence of in-medium energy loss, complimenting the measurement of heavy-flavour $R_{\rm AA}$. The v_2 of D mesons, heavy-flavour decay electrons and heavy-flavour decay muons [181] has been measured at $\sqrt{s_{\rm NN}} = 2.76$ TeV, and is measured to be of the same order (or slightly smaller) than the v_2 of charged particles, suggesting charm quarks thermalise significantly in the medium.

Predictions of the $R_{\rm AA}$ of D mesons and heavy-flavour decay leptons which include the underlying hydrodynamical expansion of the system (see above) can also make

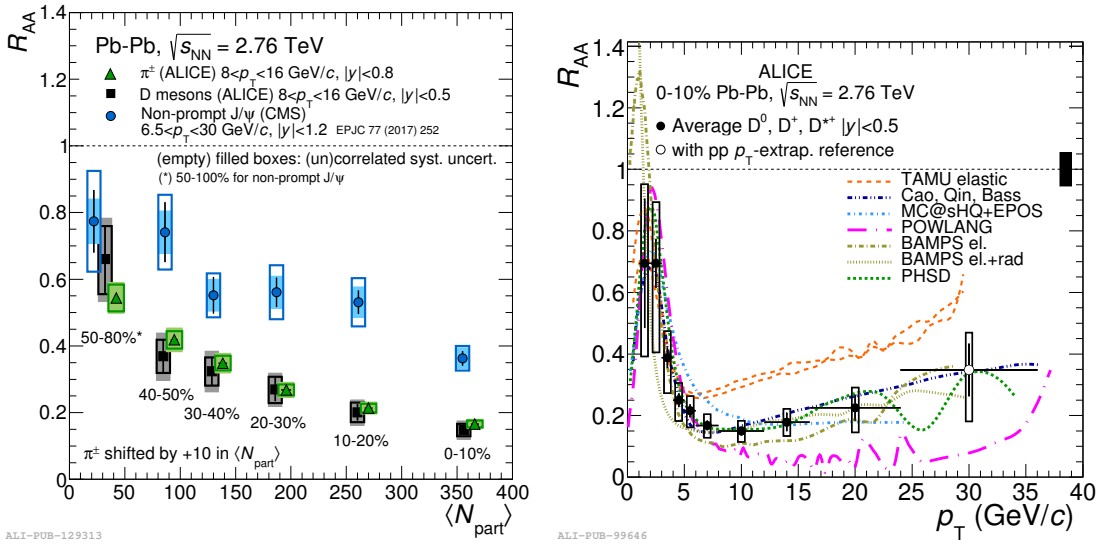

Figure 2.15: Left: R_{AA} as a function of N_{part} for D mesons [178], pions [179] and non-prompt J/ψ [180]. Right: D meson R_{AA} as a function of p_{T} in comparison to models (see [142] and references within). Both measurements are made in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV.

predictions for the elliptic flow v_2 of heavy-flavour, though it is still a challenge for models to describe both observables (see [142] and references within). Some data-driven approaches to extract transport coefficients from the measured R_{AA} and v_2 of heavy quarks have also been attempted. For example, a Bayesian analysis has been performed [182] to obtain a posterior probability distribution for the temperature dependence of the charm quark spatial diffusion coefficient using D meson R_{AA} and v_2 measurements from ALICE and STAR, where a constrained value compatible with lattice QCD calculations is obtained.

The first measurement of the Λ_{c}^{+} baryon in nucleus-nucleus collisions was presented at Quark Matter 2017, measured in Au–Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV/c by the STAR experiment [183] and is shown in figure 2.16. The measurement was made in $3 < p_{\mathrm{T}} < 6$ GeV/c in a 10-60% centrality interval. The measured $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^0$ ratio is about unity - significantly higher than the PYTHIA [91] prediction, and higher than the statistical hadronisation model (SHM) [184–186]. The data is compatible with models which include diquark and three-quark coalescence [134, 135], though it should be noted that these predictions are for the 0-5% centrality class. The data is also compared to a prediction for the baryon-over-total-charm-meson $\Lambda_{\mathrm{c}}/\mathrm{D}$ from a model which includes charm quark coalescence and interaction in the hadronic phase [139], where one might expect a factor of ∼1.5 increase when considering the $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^0$

ratio [183]. The suggestion from these results is that Λ_c^+ baryon production is consistent with models involving hadronisation via coalescence, including coalescence from diquarks. However, for a proper interpretation of these results a reference measurement is crucial, either in pp collisions or in peripheral Au–Au collisions at the same centre-of-mass energy.

2.3 Aim of this thesis

This chapter has described the motivation for measuring heavy-flavour (in particular charm) production in pp, p–Pb and Pb–Pb collisions. A brief description of some theoretical concepts in each collision system, and some highlights of experimental results have been shown. From this discussion, some questions arise:

- **Is the production of heavy baryons in pp collisions at mid-rapidity and at LHC energies reproduced by pQCD calculations?** D meson production (and overall charm production) is described well by pQCD predictions over a wide transverse momentum range, and Λ_c^+ production is also described

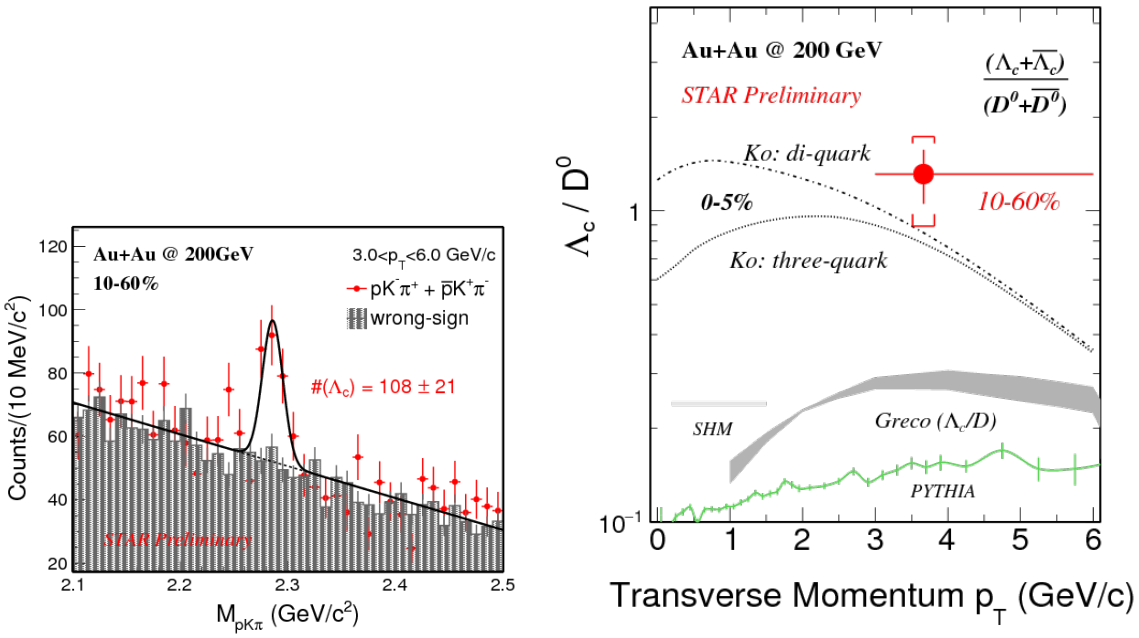

Figure 2.16: Left: The invariant mass distribution of the Λ_c^+ baryon in Au–Au collisions at $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV/c in a 10-60% centrality interval measured by STAR [183]. Right: The corresponding baryon-to-meson ratio Λ_c^+/D^0 [183] in comparison to theoretical models [134, 135, 139, 184–186].

well at forward rapidity. A measurement of charmed baryon production in pp collisions at mid-rapidity will contribute towards an understanding of hadron production mechanisms.

- **Is the process of *hadronisation* of heavy quarks described by models in high-energy pp collisions?** It has been shown that hadronisation models in pp collisions (rope hadronisation, colour reconnection) lead to modified predictions of the baryon-to-meson ratio in the strange sector, consistent with experimental data, and also reproduce some features of high-multiplicity pp collisions. A measurement of the charmed baryon Λ_c^+ and thus a measurement of the baryon-to-meson ratio in the charm sector will help to constrain phenomenological models of QCD in the non-perturbative regime.

- **Is the Λ_c^+ yield and the baryon-to-meson ratio in the charm sector modified by *cold nuclear matter effects* in p–Pb collisions?** Initial- and final-state effects in p–Pb collisions may modify the Λ_c^+ yield relative to pp collisions - results from D mesons suggest a small level of modification in the charm sector, though an enhancement of baryons has been measured in high-multiplicity d–Au and p–Pb collisions relative to pp collisions.

- **Does *coalescence* of heavy quarks with light quarks contribute to hadronisation in the charm sector, in a hot, deconfined medium created in heavy-ion collisions?** It has been shown in this chapter that the charmed baryon-to-meson ratio in Pb–Pb collisions will be sensitive to in-medium hadronisation (coalescence) mechanisms, and may be able to reveal predicted light diquark states. While the current ALICE experimental setup does not allow for a Λ_c^+ baryon measurement in Pb–Pb collisions, it is useful to evaluate the predicted reach of this measurement for Run 3 of the LHC, after ALICE's Inner Tracking System has been upgraded.

- **Experimentally, can the signal extraction procedure to measure the Λ_c^+, through the measurement of its decay products and selection on its topological and kinematic properties, be improved?** The small production cross section of the Λ_c^+ relative to the strange Λ baryon and charmed mesons, the 'messy' environment of pp collisions and, to a larger extent, heavy-ion collisions, and the three-prong decay of the Λ_c^+ baryon's most easily reconstructed hadronic decay modes $\Lambda_c^+ \rightarrow \mathrm{pK}^-\pi^+$ and $\Lambda_c^+ \rightarrow \mathrm{pK}_s^0$ all together means the signal extraction of the Λ_c^+ remains challenging, even in pp collisions. Its

short lifetime ($c\tau \sim 60$ μm) also means a very precise detector resolution is needed to resolve the secondary decay vertex of the Λ_c^+ baryon. Currently in ALICE strange and charmed hadrons are measured by making 'rectangular' cuts on the decay topology, but more optimal way to distinguish true Λ_c^+ baryons from the large combinatorial background coming from uncorrelated sources may aid in improving the precision of the Λ_c^+ measurement.

This thesis aims to contribute towards answering these questions. Chapter 4 describes the data analysis procedure to obtain a measurement of the p_{T}-differential cross section of the Λ_c^+ baryon in pp collisions at $\sqrt{s} = 7$ TeV and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 7$ TeV, with data samples that the ALICE experiment collected in 2010 and 2013, respectively. Chapter 5 then shows the results derived from these measurements. Finally, chapter 6 will detail a study performed to evaluate the performance of a multivariate approach aimed at reducing the combinatorial background in the Λ_c^+ analysis after the upgrade of the ALICE Inner Tracking System (see the following chapter for a description of the current Inner Tracking System) in Pb–Pb collisions. Before this, the next chapter will describe the current setup of the ALICE experiment.

Chapter 3

ALICE Experimental Setup

A Large Ion Collider Experiment (ALICE) [187] is one of the 4 main experiments at the Large Hadron Collider (LHC) [188], the accelerator built to collide charged particles at unprecedented energies (TeV scale). ALICE has been designed and optimised for the high-multiplicity environment of Pb–Pb collisions, in order to study the deconfined state of matter known as the Quark-Gluon Plasma (see chapter 1). This chapter will give a description of the current ALICE experimental setup. Section 3.1 will introduce the main features of the LHC. An overview of the ALICE experimental apparatus will be given in section 3.2, along with the main detectors used in the analysis presented in this thesis. Section 3.3 will outline in more detail the features of these detectors, including their tracking and particle identification capabilities. A short description of the trigger and data taking conditions is given in section 3.4.

3.1 The Large Hadron Collider

The LHC is a large particle collider built at the European Organization for Nuclear Research (CERN), near Geneva, Switzerland. Located 100 m underground it spans 27 km in circumference, and makes use of pre-existing accelerator systems to act as injectors. It has been designed to collide proton beams, with a maximum centre-of-mass energy of 14 TeV at a luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, as well as Lead (Pb) ions with a maximum energy of 2.8 TeV per nucleon at a peak luminosity of 10^{27} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

A schematic of the layout of the LHC can be seen in figure 3.1. The main experiments are built at 4 interaction points - ATLAS [189] and CMS [190], 2 general purpose detectors designed to confirm or disprove the existence of the Higgs boson (discovered in 2012 [191, 192]) and to search for physics beyond the Standard Model; LHCb [193], specialising in b-physics to measure the parameters of CP violation; and ALICE. Points 3 and 7 each contain two collimation systems, used to focus the beam

and control beam losses. Point 4 contains two RF systems, one per beam, which accelerate the beam up to its operating energy. Point 6 contains the beam dump, which is where the beam is extracted from the machine.

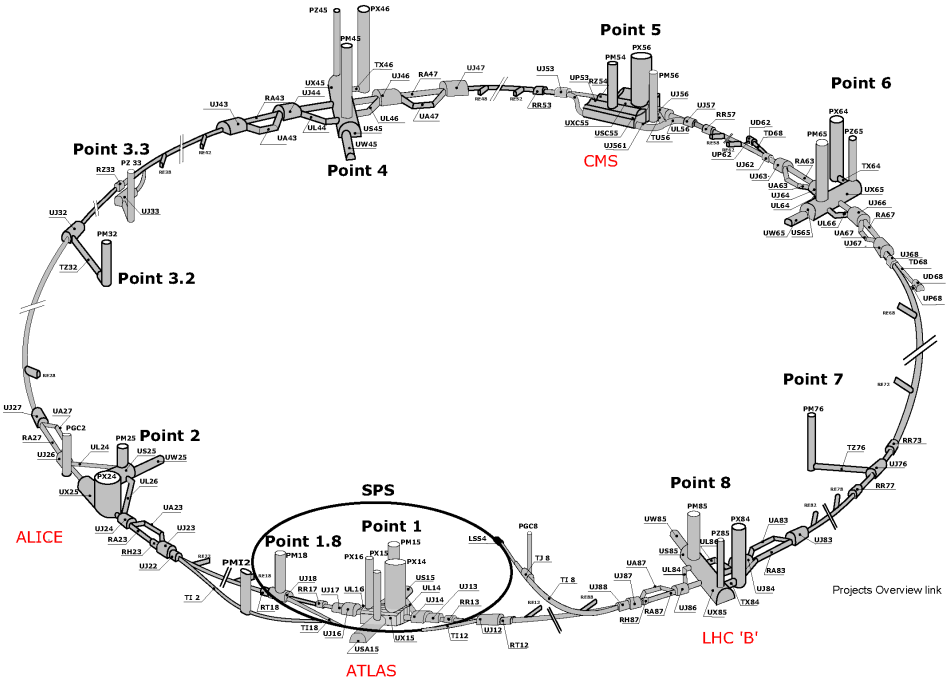

Figure 3.1: Schematic of the layout of the LHC, including each of the main experiments and all other points of entry [194].

The LHC first collided proton beams in 2009 at a centre-of-mass energy of 0.9 TeV. During Run 1 (November 2009 - February 2013) it has provided pp collisions at $\sqrt{s} = 0.9, 2.76, 7$ and 8 TeV, increasing in energy from 2010-2013. During this time there were also 3 dedicated heavy-ion programs, providing Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV per nucleon-nucleon collision in 2010 and 2011, and in January 2013 there was a p-Pb run at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV per nucleon-nucleon collision. The integrated luminosity delivered to all experiments can be seen for all heavy-ion data taking periods in Run 1 in figure 3.2 (left). After a shutdown period the LHC restarted in April 2015 for Run 2, delivering pp collisions at $\sqrt{s} = 13$ TeV. The heavy-ion runs of Run 2 have included Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV at the end of 2015, and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV and $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV at the end of 2016. The integrated luminosity of pp collisions in 2016 can be seen in figure 3.2 (right). Note that since a maximal luminosity is not a priority at LHCb or ALICE they run with ‘luminosity levelling’, where the beams are offset from one another and slowly

moved towards each other, so that the luminosity remains constant, hence the reduced delivered integrated luminosity.

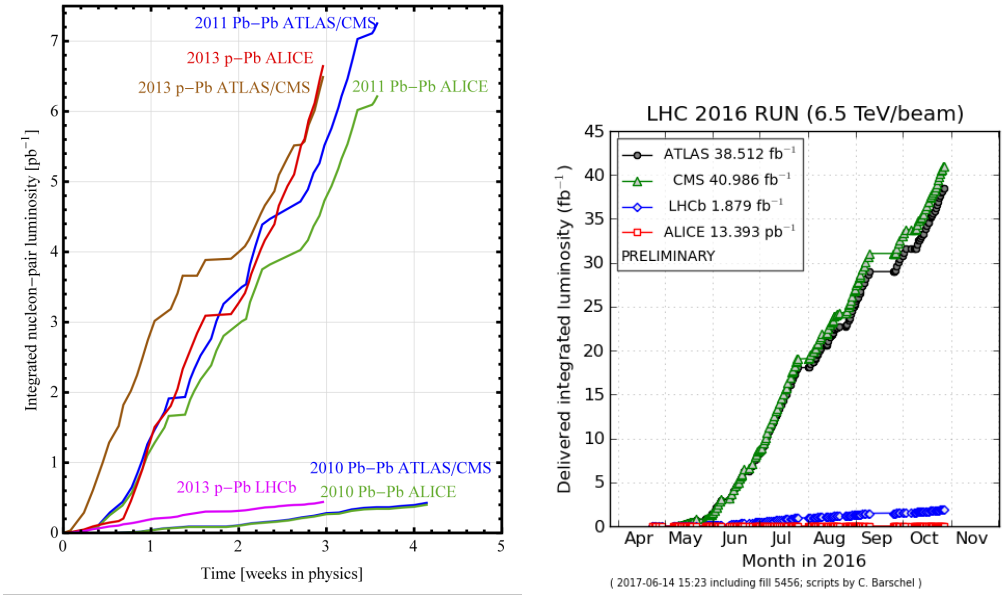

Figure 3.2: Left: Delivered integrated luminosity per nucleon-nucleon pair at the LHC for collisions involving ions, during Run 1 [195]. Right: Delivered integrated luminosity for proton-proton collisions in 2016 [196].

3.2 The ALICE detector

The ALICE detector was first conceptualised at the end of 1990, and was officially approved by the CERN research board in 1997 [197]. After an extensive design and R&D effort through most of the 1990's, ALICE was constructed and began data taking at the start of the LHC's operation in 2008. Extrapolation from the running conditions found at heavy-ion experiments at the time, namely fixed-target experiments at the CERN SPS and the Brookhaven AGS, required a factor of 300 increase in centre-of-mass energy and a factor of 7 in beam mass. For this reason, the detector design was general purpose, that is, able to measure physics signals of possible interest even if the reason was established later on, and flexible, allowing potential modifications during the design and construction process [198].

ALICE has been designed with many physics considerations in mind. One of the main challenges with heavy-ion collisions is the extreme number of particles produced in a central Pb–Pb event. Original estimates from the charged-particle multiplicity density at mid-rapidity in a central Pb–Pb event ranged from $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta = 2000$ up

to $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta = 8000$, although it has now been measured to be $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \approx 1600$ [199] in the most central Pb–Pb events at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV, a factor of ~ 400 larger than the average proton-proton multiplicity of $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \approx 4$ [200] at similar energies. A large reach for particle momentum measurements was also needed, from tens of MeV/c to over 100 GeV/c, achieved with very low material thickness to reduce multiple scattering at low p_{T} and a large tracking distance for good resolution at high p_{T}. Particle Identification (PID), essential for many physics signals, is performed employing many PID techniques.

A schematic of ALICE is shown in figure 3.3. The ALICE detector consists of a cylindrical central barrel and a forward muon spectrometer. The Inner Tracking System (ITS) is a six-layer silicon detector closest to the beam pipe, and is used for particle tracking and vertex reconstruction. The Time Projection Chamber (TPC) is the main tracking detector also used for PID. The Time-Of-Flight detector (TOF) is used to identify particles via their time of flight. These detectors are described in more detail below. Electron identification is improved by the Transition Radiation Detector (TRD) and photons are detected with the Photon Spectrometer (PHOS), and both electrons and photons are detected by the Electromagnetic Calorimeter (EMCal), which is also used for jet measurements. High-momentum particles can also be identified with the HMPID detector. These central detectors are all embedded in a solenoid magnet with $B = 0.5$ T. The muon spectrometer is located at negative rapidity ($-4.0 < \eta < -2.4$), and consists of layers of absorbers, muon trigger (MTR) and muon tracker (MCH) systems, contained within a dipole magnet with a field integral of 3 Tm, to identify muons from heavy-flavour decays and quarkonia. In addition, the Forward Multiplicity Detector (FMD), Zero Degree Calorimeters (ZDC), T0 and V0 detectors are used to measure global event characteristics, and to trigger events of interest. The Photon Multiplicity Detector (PMD) measures the multiplicity and spatial distribution of photons in the forward region. Finally, the ALICE COsmic Ray DEtector (ACORDE) is used to trigger on cosmic rays for alignment and cosmic-ray physics purposes. For a more detailed description of each detector subsystem see [187] and references within.

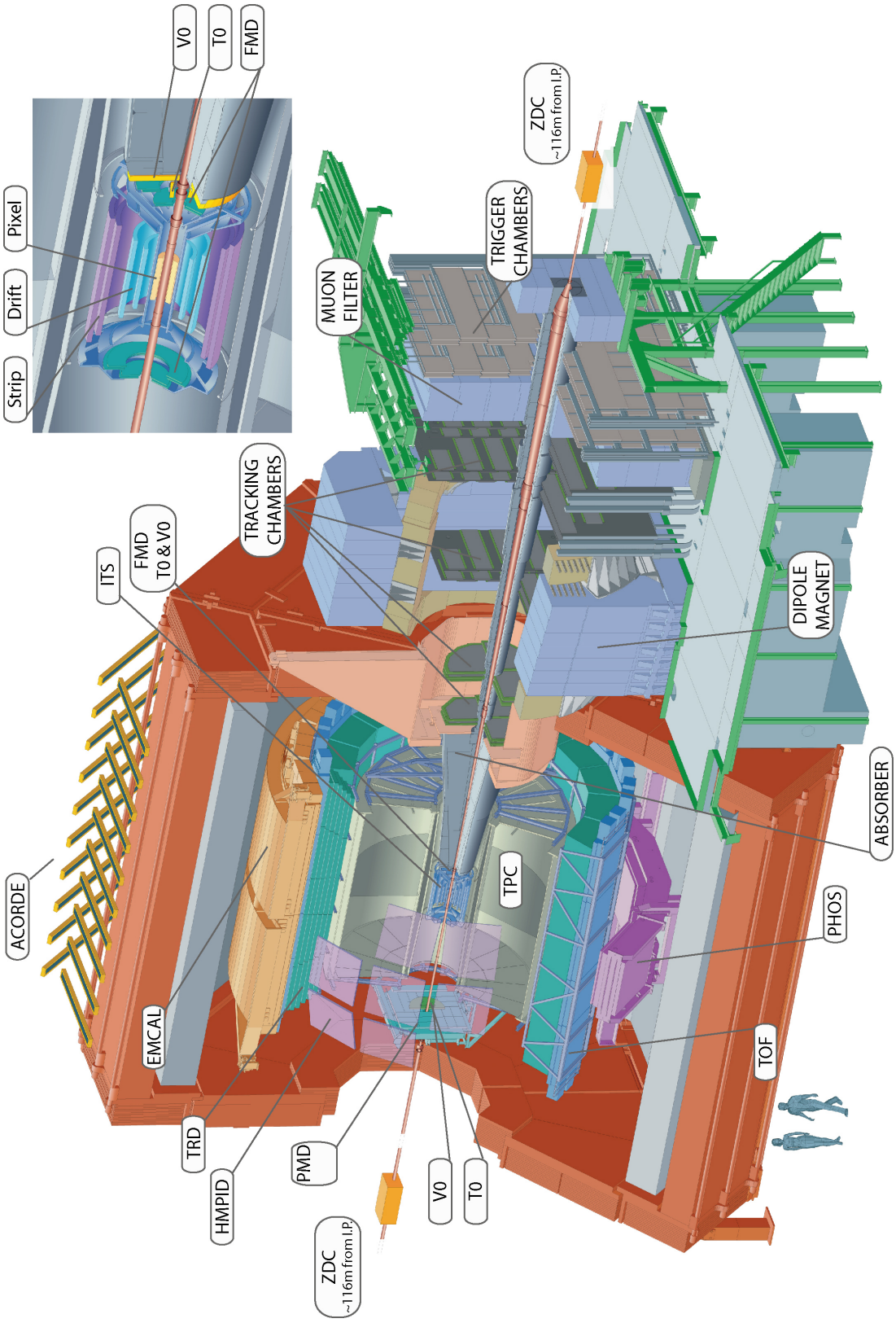

Figure 3.3: Schematic of the ALICE detector. Each individual sub-detector is labelled [201].

3.2.1 The Inner Tracking System

The ITS is a six-layer silicon detector, built primarily for tracking particles and reconstructing particle decays with a very high spatial accuracy. It is also able to contribute to particle identification of low-momentum particles, and improve the momentum resolution of the TPC. It consists of three types of detector, with each sub-detector composed of two layers - The Silicon Pixel Detector (SPD), the Silicon Drift Detector (SDD) and the Silicon Strip Detector (SSD), as shown schematically in figure 3.4.

The ITS, covering the full azimuthal range, has been built with many design considerations in mind. Due to the very high track density in heavy-ion collisions the ITS has high granularity and a good two-track separation. The very small impact parameter resolution[1] needed to reconstruct short-lived particles has led to the ITS detectors having a spatial resolution of a few tens of μm. The rapidity coverage ($|\eta| < 0.9$ for all layers, except the inner SPD layer covering $|\eta| < 1.98$) allows for a good efficiency in detecting high mass, low transverse-momentum particles. Since the momentum and impact parameter resolution for particles with small transverse momenta are dominated by multiple scattering in the detector, the material in the active volume has been kept as small as possible, such that the detectors active thickness amounts to 0.4% of the detectors average radiation length X_0, and the additional material is kept to a comparable thickness. The outer two layers (SDD, SSD) also offer PID capabilities for low-momentum particles. Table 3.1 shows the parameters for each of the three detector types.

The Silicon Pixel Detector (SPD) constitutes the innermost 2 layers of the ITS. It is based on hybrid silicon pixels, consisting of a two-dimensional matrix of reverse biased silicon diodes. The sensor matrix of the silicon chips includes 256×160 pixels (or cells), measuring 50 μm ($r\phi$) $\times$ 425 μm (z), which allows for a spatial precision of 12(100) μm to be achieved in the $r\phi(z)$ direction. The readout is binary, meaning the signal is only read out if the pre-amplified and shaped signal is above a certain set threshold. 1200 chips make up the full SPD for a total of 9.8×10^6 cells. The extended pseudo-rapidity coverage ($|\eta| < 1.98$) of the inner SPD layer provides, together with the Forward Multiplicity Detectors (FMD), continuous coverage for the measurement of charged-particle multiplicity.

The Silicon Drift Detector (SDD) makes up the two middle layers of the ITS. The SDD provides very good multitrack capability, as well as a $\mathrm{d}E/\mathrm{d}x$ measurement for

[1]The impact parameter is defined as the minimum distance of a track to the interaction vertex. This quantity defines the spatial resolution of the detector

Parameter	Silicon Pixel	Silicon Drift	Silicon Strip
r position of inner (outer) layer (cm)	3.9(7.6)	15.0(23.9)	38.0(43.0)
Spatial precision $r\phi$ (μm)	12	35	20
Spatial precision z (μm)	100	25	830
Two track resolution $r\phi$ (μm)	100	200	300
Two track resolution z (μm)	850	600	2400
Cell size (μm^2)	50×425	202×294	95×40000
Active area per module (mm^2)	12.8×69.6	72.5×75.3	73×40
Readout channels per module	40960	2×256	2×768
Total number of modules	240	260	1698
Total number of readout channels (k)	98	133	2608
Total number of cells (M)	9.84	23.0	2.6

Table 3.1: Parameters of the various detector types used in the Inner Tracking System [187]. A module represents a single sensor element.

particle identification. The active sensors consist of a series of parallel, implanted p^+ strips which are connected to a voltage divider providing the voltage needed to fully deplete the volume of the detector, and the edge of the detectors consist of a series of n^+ pads. An ionising particle traversing the detector creates electron-hole pairs, and the holes are collected by the p^+ strips, while the electrons drift towards the n^+ pads where they are collected. The coordinate perpendicular to the drift direction is then given by the centroid of the collected charge as a function of the anode position, and the coordinate parallel to the drift direction is given by the centroid of the signal as

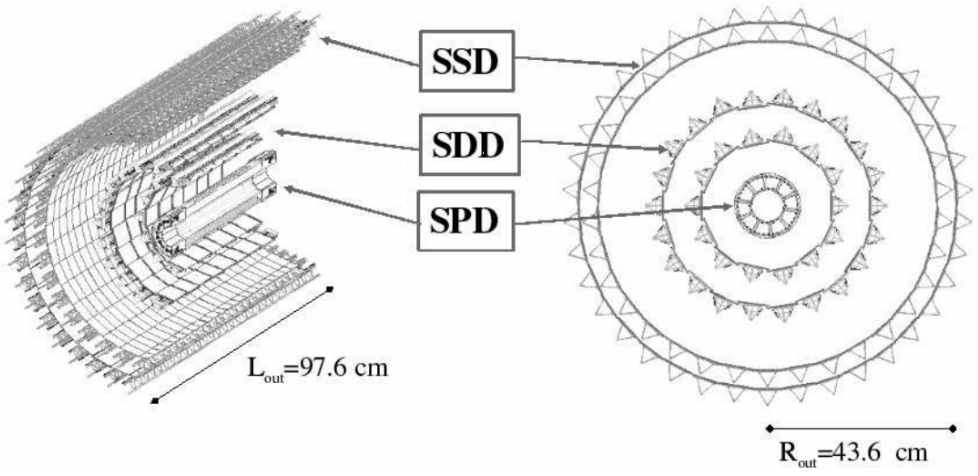

Figure 3.4: Schematic of the ITS layout. The SSD (inner two layers), SDD (middle two layers) and SSD (outer two layers) are labelled [187].

a function of time, achieving a spatial precision of 35(25) μm along the drift(anode) direction, corresponding to the $r\phi(z)$ coordinates.

The Silicon Strip Detector (SSD) make up the outer two layers of the ITS. This detector provides good position resolution plus $\mathrm{d}E/\mathrm{d}x$ information, and is crucial for the matching of tracks from the TPC to the ITS. The sensors are each 75×42 mm^2 and are made of n-type silicon. On both sides silicon strips are implanted, 40 mm long and intersecting with an angle $\alpha = 35$ mrad. One strip side is p^+ doped and the other is n^+ doped, and are biased to fully deplete the sensor. The readout electronics digitise a total of 2.6 million analogue channels. A spatial resolution of around 20(830) μm can be achieved in the $r\phi(z)$ direction.

3.2.2 Time Projection Chamber

The Time Projection Chamber is the main tracking detector in the central barrel of ALICE. It covers the full azimuthal range and a pseudorapidity coverage of $|\eta| < 0.9$ for the full radial track length, with the active volume spanning a radial position of $848 < r < 2466$ mm. It has been optimised to track particles in the high-multiplicity environment of Pb–Pb collisions. It is also used for PID over a wide momentum range.

The detector consists of a large cylindrical field cage, filled with 90 m^3 of a gas mixture[2] of primarily Ar or Ne. Charged particles propagating through the TPC ionise the gas. The ionised electrons then drift at a speed of 2.7 cm/μs under the large voltage (a potential gradient of 400 V/cm in the z direction creating a total potential of 100 kV) towards the readout on one of the two end plates. Multi-wire proportional chambers (MWPCs) make up the end plates, where the ionised electrons cause an electron avalanche which reaches the cathode pad readout. The position of the charged deposition on the cathode gives the two-dimensional track position in $r\phi$, and the time taken for the electrons to drift to the end plates gives the track position in z. In this way, a position resolution of 1100(800) μm in $r\phi$ and 1250(1100) μm in z is attainable for tracks at the inner (outer) radii.

PID is performed by the TPC via a measurement of the energy loss per unit length $\mathrm{d}E/\mathrm{d}x$, or specific energy loss, which is directly related to the number of electrons ionised by the charged particle propagating through the TPC. The specific energy loss of a given particle is described by the Bethe-Bloch formula [13] and in ALICE a

[2]During Run 1 the TPC was filled with a Ne/CO_2 (90-10) or $Ne/CO_2/N_2$ (90-10-5) mixture. For Run 2 this mixture was changed to Ar/CO_2 (90-10)

parametrisation of this formula, initially proposed by the ALEPH collaboration [202], is used [203] which can be written as

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

where P_{1-5} are parameters determined from fits to measured data. Figure 3.5 shows the specific energy loss as measured in Pb–Pb collisions as a function of the track momentum, along with the Bethe-Bloch characterisations for different charged particles, where a separation between particle species (e, π, K, p as well as deuteron) can clearly be seen. At low momenta the separation between species is clear and tracks can be identified on a track-by-track basis, and at high momenta tracks can be identified on a statistical basis. A relativistic rise offers further separation between species at high momentum.

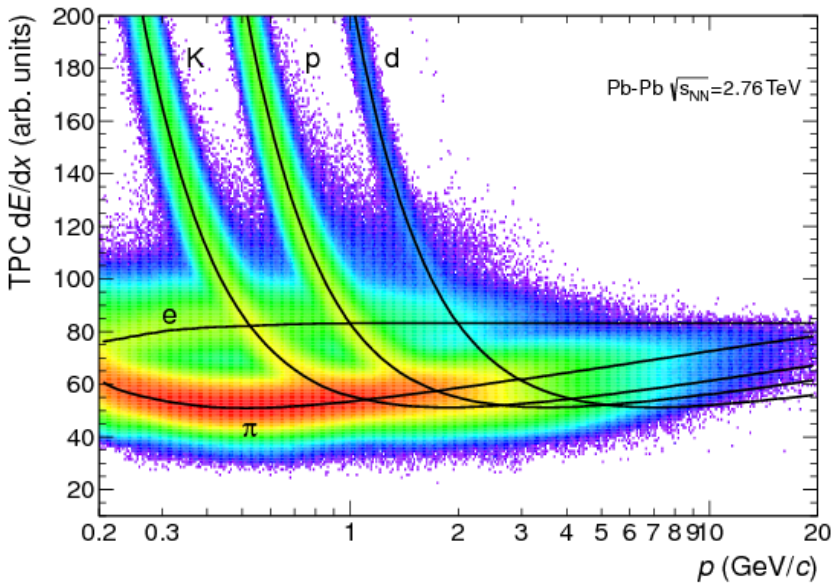

Figure 3.5: Measured $\mathrm{d}E/\mathrm{d}x$ as a function of particle momentum in the TPC, in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV. The black lines indicate the Bethe-Bloch $\mathrm{d}E/\mathrm{d}x$ parameterisation for various charged particles [203].

3.2.3 Time-Of-Flight Detector

The Time-Of-Flight detector is a large array of Multi-gap Resistive Plate Chambers (MRPCs), covering the full azimuthal range with a rapidity coverage of $|\eta| < 0.9$. It provides PID in the intermediate momentum range, at higher momentum than

the TPC, via the measurement of the time of flight of charged particles from the interaction point to the TOF.

The MRPCs, which form the basic module of the TOF, consist of stacks of resistive glass plates, placed inside gas-tight modules, with a high voltage applied to the external surfaces of the stack. A charged particle ionises the gas within the TOF modules, and the high electric field causes an electron ionisation avalanche. The avalanche is stopped when reaching a gas plate, inducing a signal on pick-up electrodes. There are between 15 and 19 MRPC strips in a TOF module; 5 modules make up a sector, and the full TOF detector contains 18 sectors, giving a total of 1638 MRPC strips.

The start time for the TOF detector is provided by the T0 detector, which achieves a time resolution of $20 - 25$ ps for Pb–Pb collisions and 40 ps for pp collisions. The efficiency of the T0 detector is 100% for the 60% most central Pb–Pb collisions, and around 50% for pp collisions. The start time can also be estimated by the TOF detector, based on particle arrival times at the detector, which is useful for when the T0 signal is absent. The intrinsic time resolution of the TOF detector is around 80 ps, and the total time resolution is a combination of the time resolution of the TOF detector and the event time resolution measured with the T0 detector such that $\sigma^2_{tot} = \sigma^2_{TOF} + \sigma^2_{T0}$. The total time resolution of pion tracks is shown in figure 3.6. The resolution improves with increasing number of tracks used to define the start time of the collision, and at the limit of high track multiplicity the time resolution approaches the intrinsic TOF detector resolution.

The PID capabilities of the TOF is illustrated in figure 3.7, where the particles' velocity as a fraction of the speed of light β is plotted as a function of momentum. Kaons and pions can be distinguished up to around 2.5 GeV/c, and protons and pions can be distinguished up to around 4 GeV/c.

3.2.4 V0 detector

The V0 detector system provides triggering in pp, p–Pb and Pb–Pb collisions, and is also used to judge the multiplicity of the event in p–Pb and Pb–Pb collisions allowing the event to be classified by the centrality, either at trigger level or analysis level. The V0 detector system consists of two disks of segmented plastic scintillator tiles installed on both sides of the interaction point, named V0A and V0C, and they cover the pseudorapidity regions $2.8 < \eta < 5.1$ (V0A) and $-3.7 < \eta < -1.7$ (V0C). Each disk consists of 32 counters distributed in 4 rings. Scintillation light produced by particles traversing the detector's volume is transported to photomultiplier tubes. The signal produced is proportional to the number of particles hitting the detector,

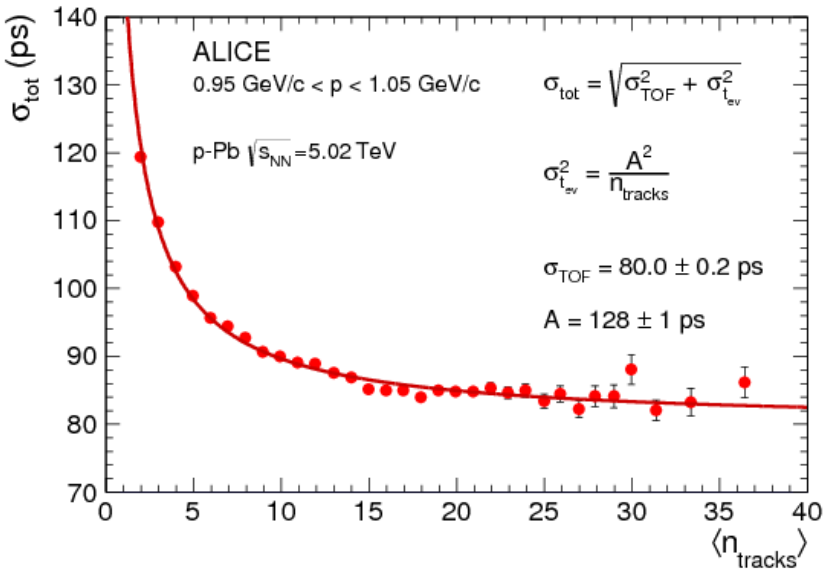

Figure 3.6: The time resolution of of the TOF detector for pions as a function of the number of tracks used to define the start time of the collision, in p–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV. Shown is a small slice in pion momentum [203].

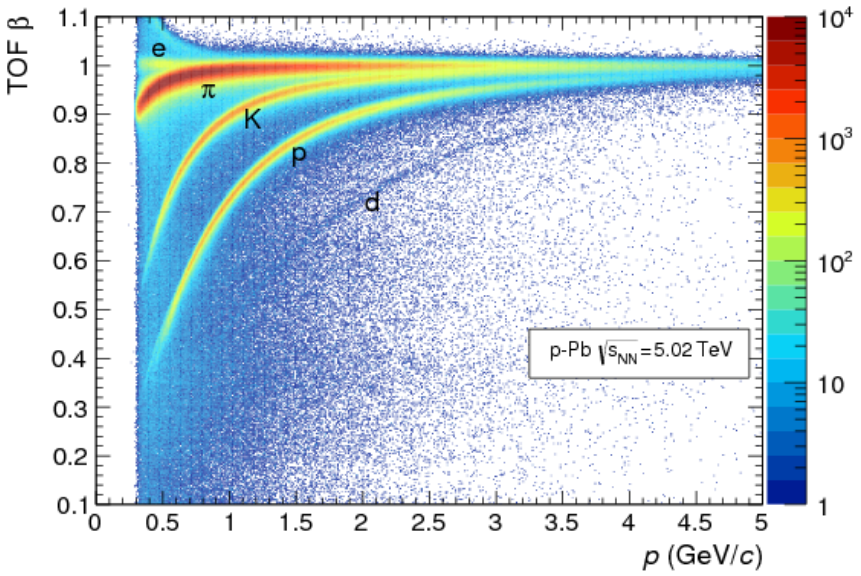

Figure 3.7: Charged particle velocity $\beta(=\frac{v}{c})$ measured in the TOF as a function of the particle momentum, in p–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV [203].

which is proportional to the number of primary particles created in the collision, allowing the a centrality determination using a Glauber model [38].

3.3 Detector performance

3.3.1 Tracking

Particle track finding [203] is performed by the ITS and TPC. Prior to this procedure 'clusterisation' is performed, where the detector data are converted into 'clusters' defined by information such as positions, signal amplitudes and signal times with their associated errors. The interaction vertex is then determined using 'tracklets' - lines defined by pairs of clusters in each SPD layer - where the vertex is found and defined when the maximum number of tracklets converges onto a single point. The tracks are then reconstructed using a three-stage approach.

The first stage begins at the TPC. Tracks are searched for starting at large radius at the edge of the TPC, where the track is built inwards. Track 'seeds' are built first with two TPC clusters and the interaction vertex, then with three clusters and no vertex, and are propagated inwards, at each step updating with the nearest cluster provided it fulfils a proximity cut. Tracks are accepted providing they have a minimum number of clusters and that they miss no more than 50% of expected clusters for the position of the track. Reconstruction of the same track is avoided by rejecting tracks which share a fraction of common clusters, where the worse of the two tracks is rejected. After the track has been propagated to the inner TPC edge a preliminary particle identification is performed based on the specific energy loss, and this information is used in the following steps. The tracks are then propagated to the outermost ITS layers, becoming seeds for the ITS track finding, and the tracks are propagated inwards based on clusters in the ITS layers. At this point any 'path' layer-to-layer is saved provided it passes a proximity cut, such that each TPC track produces a tree of track hypotheses in the ITS, at which point the highest quality candidate (based on the χ^2 of each path) from each hypothesis tree is saved, and each track is extrapolated to the point of closest approach to the interaction vertex. The second stage involves propagation outwards. The tracks are re-fitted using a Kalman filter technique [204] in the outward direction using the same clusters from the first stage. The tracks are extended and matched to clusters in the TRD, and the TOF where a time-of-flight is assigned. At the final stage all tracks are propagated inwards from the TPC to the ITS, where the tracks are re-fitted.

The impact parameter resolution is shown as a function of p_{T} in figure 3.8, for all charged particles in 3 collision systems - pp, p–Pb and Pb–Pb collisions. Due to the larger multiplicities in heavier systems the interaction vertex can be determined more precisely, and as such the impact parameter resolution is improved. Topological

selection on charmed weakly-decaying particles (D mesons, Λ_c^+ baryon) also becomes more effective at higher p_T due both to the improved impact parameter resolution, plus the longer decay length of the particle in the lab frame.

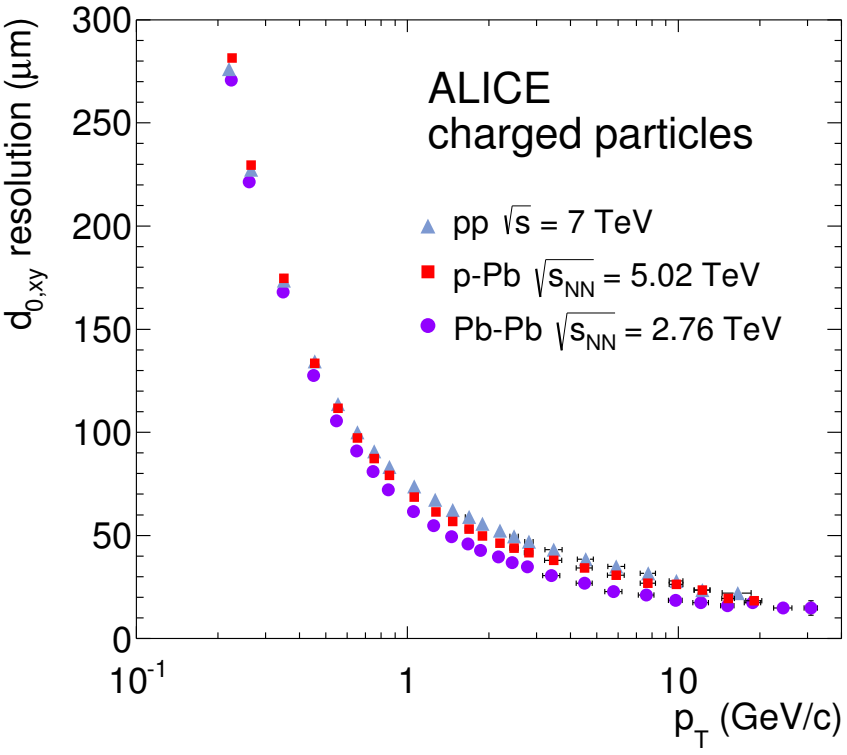

Figure 3.8: The impact parameter resolution as a function of p_T in pp collisions at $\sqrt{s}$ = 7 TeV, in p–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV and in Pb–Pb collisions at $\sqrt{s_{NN}}$ = 2.76 TeV [203].

3.3.2 Particle identification

PID in ALICE can either be performed using a single detector, or using a combination of information from more than one detector to provide optimal identification. Traditionally this can be done using an 'nσ' approach, where a selection is made on the difference between the measured signal from the detector S and the expected response from the detector $\hat{S}$ given a particle species hypothesis H_i, divided by the detector resolution σ_i where $n\sigma = S - \hat{S}(H_i)/\sigma_i$. This can be done for multiple detectors to cover different kinematic regions and/or particle type. Figure 3.9 shows the separation of the expected pion, kaon and proton signals in the TPC and TOF, where a separation can be seen for each particle species, though there is some overlap.

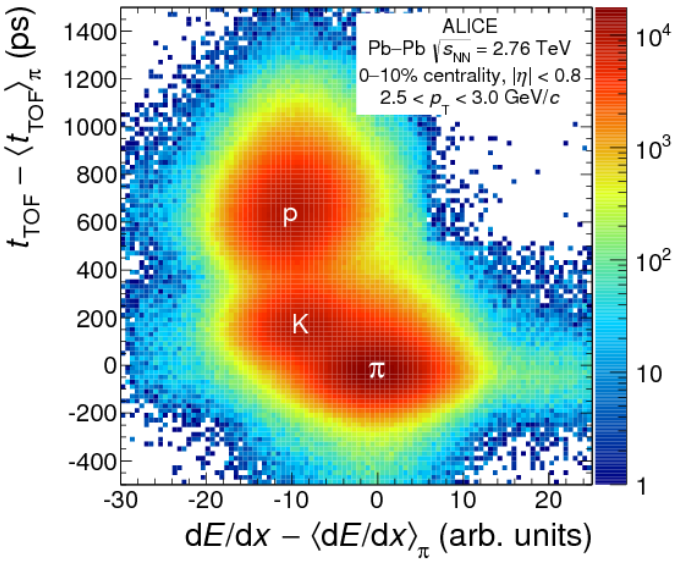

Figure 3.9: Combined particle identification information in the TPC and TOF from Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV. The PID signals are expressed in terms of the deviation from the expected response for pions in each detector [205].

However, a rectangular nσ selection on these signals may not be optimal, and within ALICE a Bayesian PID method has been developed [205] in order to provide a more optimal PID approach by expressing and combining the detector signals in terms of probabilities. For a given detector α a conditional probability that a particle of species H_i will produce a signal S can be defined, $P(S|H_i)$. If the detector response is Gaussian, this is defined as

$$P(S|H_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{1}{2}\mathrm{n}\sigma^2} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(S-\hat{S}(H_i))^2}{2\sigma^2}}. \tag{3.2}$$

$P(S|H_i)$ may be determined by comparing the detector response (i.e the dE/dx in the TPC, or the time-of-flight in the TOF) to its expected response. The probabilities from different detectors can then be combined:

$$P(\vec{S}|H_i) = \prod_{\alpha=TPC,TOF,\ldots} P_\alpha(S_\alpha|H_i). \tag{3.3}$$

This can be interpreted as the conditional probability that the set of detector signals $\vec{S}$ will be seen for a given particle species H_i. The probability that the particle is of species H_i, given the detector signal $\vec{S}$, is then expressed through Bayes' theorem

[206]:

$$P(H_i|\vec{S}) = \frac{P(\vec{S}|H_i)C(H_i)}{\sum_{k=e,\mu,\pi,...} P(\vec{S}|H_k)C(H_k)}. \tag{3.4}$$

Here, $C(H_i)$ is an a priori probability of measuring the particle species H_i, and is also known as the 'prior'. The conditional probability $P(H_i|\vec{S})$ is known as the posterior probability.

The priors are evaluated using an iterative procedure on a sample of data, and are computed as a function of transverse momentum. They are also computed dependent on the centrality in Pb–Pb collisions and dependent on the multiplicity class in p–Pb collisions. The priors are defined as a ratio with respect to pions, e.g. K/π or p/π, and all set to 1 at the beginning of the procedure. The posterior probabilities $P_n(H_i|S)$ are determined using these priors and are used as weights to fill p_{T} spectra starting from the inclusive measured p_{T} spectra, for all signals induced by particles in the sample. The priors used in the next iteration step are determined from the ratio of these spectra with that obtained for pions, and these steps are repeated until the priors converge. The priors are found to describe the measured particle abundances within $\sim 10\%$ over a wide momentum range for all collision systems [205].

Once the Bayesian probability has been calculated for each particle species for a given track, a selection based on this information can be applied, which can include

- *Fixed threshold:* The track is accepted as belonging to a particle species if the probability is larger than a given threshold;
- *Maximum probability:* The track is given the identity of the particle species with the highest probability;
- *Weighted:* All tracks reaching the PID step are accepted, with weights corresponding to the Bayesian probabilities associated to the each track.

Each of these PID strategies (as well as the nσ strategy) aim to maximise the purity (defined as the ratio between the number of correctly identified particles with the total number of particles selected) at a given efficiency of the PID strategy (defined as the proportion of particles of a given species which are identified correctly). Figure 3.10 shows a comparison of the nσ and Bayesian PID methods in the signal extraction of the $\Lambda_{\mathrm{c}}^{+} \rightarrow \mathrm{pK}^{-}\pi^{+}$ and its charge conjugate in $\sim 3\times10^{8}$ pp collisions at $\sqrt{s} = 7$ TeV with ALICE. The $\Lambda_{\mathrm{c}}^{+} \rightarrow \mathrm{pK}^{-}\pi^{+}$ candidate selection is based on topological selections and PID. For the nσ approach the TPC and TOF were used, and between a 2 and 3σ

selection was made depending on the detector, p_{T} and particle species, optimised for the Λ_{c} decay. The minimum-σ approach identifies each track to belong to the particle species which results in the minimum nσ, combining TPC and TOF by adding their respective nσ in quadrature. The Bayesian approach adopts the maximum probability criterion. In the $2 < p_{\mathrm{T}} < 6$ GeV/c region the statistical significance improves from 4.4 ± 1.0 with the minimum-σ approach to 6.0 ± 1.1 for the Bayesian approach. Additionally the signal-to-background ratio increases by a factor of around three with the background reducing by a factor of around seven. Despite the Bayesian approach yielding a lower selection efficiency, it can be concluded that the Bayesian approach is the optimal PID method in extracting the signal of the Λ_{c}^{+} baryon.

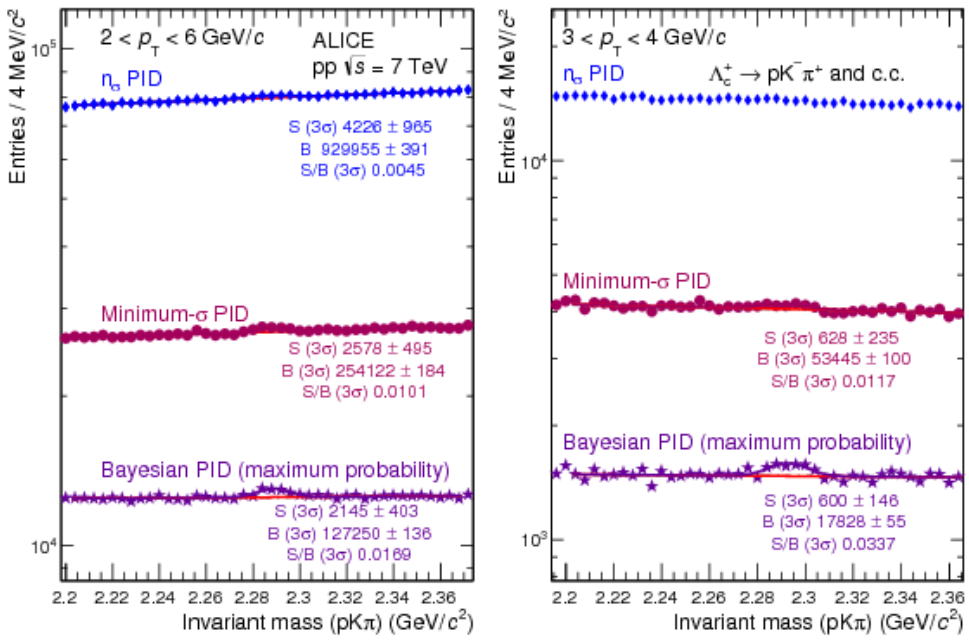

Figure 3.10: The invariant mass spectra of pKπ triplets in the Λ_{c} invariant mass window, in the p_{T} ranges $2 < p_{\mathrm{T}} < 6$ GeV/c (left) and $3 < p_{\mathrm{T}} < 4$ GeV/c (right), where 4 different PID methods are compared; nσ, minimum-σ and Bayesian. The signal S, background B and signal-to-background ratio S/B is shown for each method [205].

3.4 Trigger and data taking

The trigger decision is generated by the Central Trigger Processor (CTP) of ALICE, based on detector signals and information about the LHC filling scheme. The CTP evaluates the trigger inputs at every machine clock cycle ($\sim$ 25 ns). The difference

in the readout time of each detector means the trigger system is designed with three levels. The Level 0 trigger decision (L0) can take information from the V0, T0, EMCal, PHOS, and MTR detectors and is evaluated $\sim 0.9\mu$s after the collision. The accepted events from the L0 decision are further evaluated by the Level 1 trigger (L1) around 6.5μs after the collision. The propagation time of the signal from the ZDC and the computation time of the TRD and EMCal defines this latency. The L0 and L1 decisions are sent to the detectors triggering the buffering of the data in the detector front-end electronics if the events are accepted. The Level 2 trigger decision (L2) is taken after around 100 μs and takes input from the TPC, where the drift time in the TPC corresponds to the timing of the L2 trigger decision. If the L2 trigger decision is passed the data is sent in parallel to the DAQ and the High Level Trigger (HLT). The HLT performs a more detailed analysis of the data, and reduces the data volume before storage.

Minimum-bias (MB) events are events selected by trigger conditions defined such that as little bias as possible is induced from the trigger conditions. In pp collisions in the data taking period of 2010 a minimum interaction trigger was used requiring a logical OR between a hit in the SPD and in either of the two V0 detectors (V0A or V0C). In p–Pb collisions in the data taking period of 2013 the minimum interaction trigger instead required coincident signals in both scintillator arrays of the V0 system. The interaction rate during these periods ranged from a few kHz to a few tens of kHz - enough to reach close to the maximum detector readout rate but leaving the mean number of interactions per bunch crossing low to minimise the number of pile-up events in the same bunch crossing.

During the 2011 Pb–Pb running period, the interaction rate provided by the LHC reached $3 - 4$ kHz, and ALICE ran with minimum bias, centrality and rare triggers at the same time. The centrality triggers in Pb–Pb collisions are based on summed signals from the V0 detector, where central (0-10%) and semi-central (0-50%) events are selected. Thresholds are applied separately to the total signal in the V0A and V0C detectors, and a coincidence of the two detectors is required.

A variety of rare triggers are defined depending on the physics to be studied, using different detectors. The high-multiplicity trigger is based on the hit multiplicity in the outer layer of the SPD. The PHOS and EMCal provide triggers at L0 based on a minimum energy deposit within a window of calorimetric cells. The EMCal also provides triggers on photons/electrons, and on jets at L1. The TRD can be used at L1 to trigger on jets, single electrons, double electrons (Quarkonia trigger), and electrons in coincidence with the EMCal. The muon triggers are implemented at L0, and consist

of single-muon and dimuon triggers. Triggers requiring a particle at mid-rapidity and no particles within the intermediate pseudorapidity region covered by the V0 detector, and a similar trigger condition requiring a single muon at forward pseudorapidity but nothing at mid-rapidity was defined to select ultra-peripheral collisions[3]. Finally a cosmic ray trigger using the ACORDE detector collects high-multiplicity muon cosmic ray events. For more information about the running conditions of ALICE during Run 1 see [203].

[3]Ultra-peripheral collisions are defined as peripheral collisions with an impact parameter $b \geq R_1 + R_2$ where R_1 and R_2 are the radii of the two colliding particles, respectively. In these cases there are no hadronic interactions; instead electromagnetic interactions can occur.

Chapter 4

Data Analysis

This chapter describes the $\Lambda_c^+ \to pK^-\pi^+$ data analysis performed in pp collisions at $\sqrt{s} = 7$ TeV and p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The aim of this analysis is to extract the p_T-differential Λ_c^+ baryon production cross section in pp and p–Pb collisions. This is achieved through event, track and Λ_c^+ candidate pre-selection (sections 4.2-4.3), multivariate selection (section 4.4), extraction of the Λ_c^+ signal (section 4.5), and calculating efficiency and feed-down corrections (section 4.6). The systematic uncertainties are also evaluated and presented in section 4.7. Finally, for the measurement of the nuclear modification factor R_{pPb}, the pp measurement is scaled to the same energy per nucleon-nucleon collision as the p–Pb measurement, $\sqrt{s} = 5.02$ TeV, and this procedure is detailed in section 4.8.

4.1 Introduction

The lowest-mass charmed baryon is the Λ_c^+ with a quark content udc and a rest mass of 2286.46 $\pm$ 0.14 MeV/c^2 [13]. The Λ_c^+ measurement presented in this thesis uses the $\Lambda_c^+ \to pK^-\pi^+$ decay channel with a branching ratio of $6.35 \pm 0.33\%$, which includes 3 types of decay through a coupled resonance as well as one non-resonant decay. Table 4.1 lists the branching ratios of the individual decay channels considered within this analysis. The $\Lambda_c^+ \to pK^-\pi^+$ branching ratio as given in [13] is an average of the model-independent measurements by the Belle experiment at the KEKB e^+e^- collider [207], and the BESIII experiment at the BEPC II e^+e^- collider [208].

As indicated in table 4.1, the Λ_c^+ can decay to the $pK^-\pi^+$ three-body final state via an intermediate resonant two-body final state. This can be depicted pictorially using a Dalitz plot, which shows the splitting of total energy in the Λ_c^+ rest frame given to the three decay products. The Λ_c^+ Dalitz plot is shown in figure 4.1 - the intermediate resonances $\Lambda(1520)$, $K^*(892)^0$ and $\Delta(1232)^{++}$ are indicated as lines drawn

Λ_c^+ decay mode	branching ratio($\Gamma_i/\Gamma_{\text{total}}$)	Γ_i/Γ_1	q (MeV/c)
$pK^-\pi^+$	$(6.35 \pm 0.33)\%$	1	823
$p\overline{K}^*(892)^0$	$(1.98 \pm 0.28)\%$	0.31 ± 0.04	685
$\Delta(1232)^{++}K^-$	$(1.09 \pm 0.25)\%$	0.17 ± 0.04	710
$\Lambda(1520)\pi^+$	$(2.2 \pm 0.5)\%$	0.35 ± 0.08	627
$pK^-\pi^+$nonresonant	$(3.5 \pm 0.5)\%$	0.55 ± 0.06	823

Table 4.1: Properties of the resonant and non-resonant $\Lambda_c^+ \to pK^-\pi^+$ decay modes, including branching ratios as fractions of the decay widths Γ, and the energy released in the decay q. These numbers are taken from [13].

parallel to the axes connecting the two daughter particles. The branching fractions of each resonant decay has been measured by studying the resonant substructure of the $\Lambda_c^+ \to pK^-\pi^+$ decay channel [209–211].

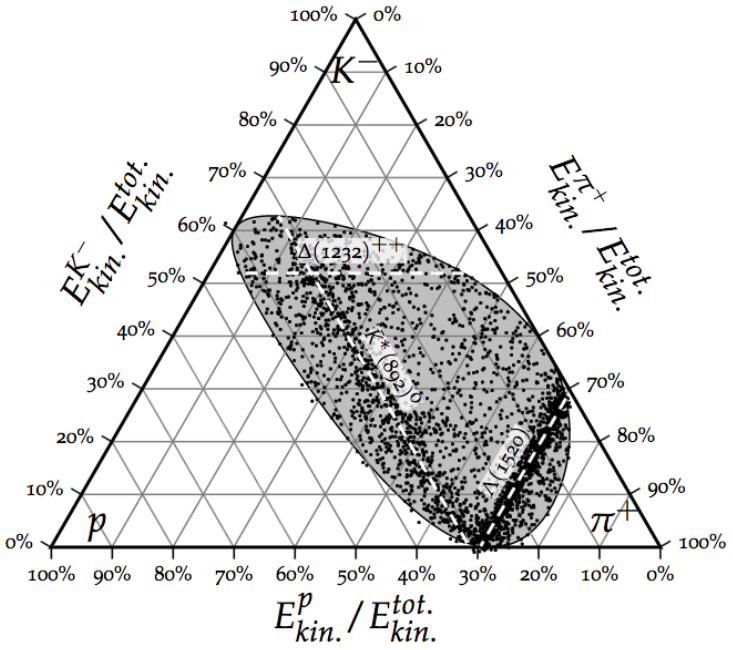

Figure 4.1: Dalitz plot of the $\Lambda_c^+ \to pK^-\pi^+$ decay [212]. See text for explanation.

A Λ_c^+ baryon that is produced in pp collisions from the fragmentation of a c-quark, produced in a hard scattering process, into a charmed hadron (either a Λ_c or a higher mass resonance) is referred to a 'prompt Λ_c'. There is also a fraction of measured Λ_c^+ that come from the decay of a beauty hadron, the majority of which come from the decay $\Lambda_b^0 \to \Lambda_c + X$. These are referred to from now on as 'non-prompt Λ_c'.

The $p_{\rm T}$-differential cross section of the prompt production of $\Lambda_{\rm c}^{+}$ baryons, in both pp and p-Pb collisions, is determined as

$$\left.\frac{{\rm d}\sigma^{\Lambda_{\rm c}{}^{+}}}{{\rm d}p_{\rm T}}\right|_{|y_{\rm lab}|<0.5} = \frac{1}{2}\frac{1}{\Delta y \Delta p_{\rm T}}\frac{f_{\rm prompt}(p_{\rm T})\cdot N_{\rm raw}^{\Lambda_{\rm c}^{\pm}}(p_{\rm T})\Big|_{|y_{\rm lab}|<y_{\rm fid}(p_{\rm T})}}{(Acc\times\epsilon)_{\rm prompt}(p_{\rm T})\cdot {\rm BR}\cdot\mathcal{L}_{\rm int}} \tag{4.1}$$

The raw yields $N_{\rm raw}^{\Lambda_{\rm c}^{\pm}}(p_{\rm T})$ of the decay $\Lambda_{\rm c}^{+}\rightarrow {\rm pK}^{-}\pi^{+}$ and its charged conjugated decay $\Lambda_{\rm c}^{-}\rightarrow \bar{\rm p}{\rm K}^{+}\pi^{-}$ are extracted via a fit to the invariant mass distribution of the identified decay products of the $\Lambda_{\rm c}{}^{\pm}$ candidates, within the fiducial acceptance of the detector $y_{\rm fid}(p_{\rm T})$[1]. To correct for the contribution of $\Lambda_{\rm c}^{+}$ particles coming from a b-baryon decay, the raw yield is multiplied by the fraction of prompt $\Lambda_{\rm c}^{+}$, $f_{\rm prompt}(p_{\rm T})$. The $(Acc\times\epsilon)$ correction accounts for the detector acceptance, and the reconstruction and selection efficiencies (including PID, event selection and multivariate selection). The cross section is normalised by the branching ratio BR and the integrated luminosity $\mathcal{L}_{int}=N_{pp(pPb),MB}/\sigma_{MB}$, where $N_{pp(pPb),MB}$ is the total number of pp(pPb) events passing the minimum bias trigger, and σ_{MB} is the minimum bias trigger cross section. The factor $\frac{1}{\Delta y\Delta p_{\rm T}}$ takes into account the width of the $p_{\rm T}$ bins, and the y interval, and the factor 1/2 gives the $\Lambda_{\rm c}^{+}$ cross section, given that $\sigma^{\Lambda_{\rm c}^{+}}=(\sigma^{\Lambda_{\rm c}^{+}}+\sigma^{\Lambda_{\rm c}^{-}})/2$.

In p–Pb collisions, because of the different energies per nucleon of the proton and lead beams, the nucleon-nucleon centre-of-mass frame is shifted in rapidity, moving with a rapidity $|\Delta y_{\rm NN}|=0.465$. The measurement of the $\Lambda_{\rm c}^{+}$ cross section in $|y_{\rm lab}|<0.5$ thus leads to a shift in the centre-of-mass rapidity coverage of $-0.96<y_{\rm cms}<0.04$.

As described in section 4.5 and as mentioned above, the measurement involves extracting the signal of both $\Lambda_{\rm c}^{+}$ and $\Lambda_{\rm c}^{-}$ through an invariant mass analysis of both ${\rm pK}^{-}\pi^{+}$ and $\bar{\rm p}{\rm K}^{+}\pi^{-}$ triplets. For this reason and for brevity, from here on in '$\Lambda_{\rm c}$' refers to '$\Lambda_{\rm c}^{+}+\Lambda_{\rm c}^{-}$', unless stated otherwise.

4.2 Event Selection

4.2.1 pp collisions

The pp data sample analysed in this work was collected with the ALICE detector in 2010. The beam energies were 3.5 TeV per beam, leading to a collision energy of $\sqrt{s}=7$ TeV. The data sample was collected with a minimum bias trigger, which

[1]The acceptance of $\Lambda_{\rm c}$ candidates drops steeply for $y_{\rm lab}>0.5$ at low $p_{\rm T}$ and for $y_{\rm lab}>0.8$ for $p_{\rm T}>5$ GeV/c, so a fiducial acceptance cut is applied, defined as $|y_{\rm fid}(p_{\rm T})|<-\frac{0.2}{15}p_{\rm T}^{2}+\frac{1.9}{15}p_{\rm T}+0.5$ for $p_{\rm T}\leq 5$ GeV/c and $|y_{\rm fid}(p_{\rm T})|<0.8$ for $p_{\rm T}>5$ GeV/c.

required signals in the SPD layers or in either of the V0 arrays (V0A or V0C). A selection was also made on the position of the primary vertex of the interaction requiring it to be within ±10 cm from the centre of the detector in the z plane, and a maximum difference of ±5 mm between the primary vertex with that reconstructed with just the SPD. In total about 310 million events were analysed after event selection, corresponding to an integrated luminosity of about $\mathcal{L}_{int} = 5$ nb^{-1}.

4.2.2 p–Pb collisions

The p-Pb data sample analysed in this work was collected with the ALICE detector at the end of 2013. The beam energies were 4 TeV for protons and 1.58 TeV per nucleon for the lead nuclei, leading to a collision energy per nucleon-nucleon collision of $\sqrt{s_{\rm NN}} = 5.02$ TeV. The data sample was collected with a minimum bias trigger, which required coincident signals in both scintillator arrays of the V0 detector. The same conditions as for the pp sample on the position of the primary vertex were also required. In total about 100 million events were analysed after event selection, corresponding to an integrated luminosity of about $\mathcal{L}_{int} = 50$ μb^{-1}.

4.3 Track and Λ_c candidate selection

The $\Lambda_c^+ \to {\rm pK}^-\pi^+$ and corresponding charged conjugate decays are reconstructed from the detection and identification of the 3 final state particles: a proton, a negatively charged kaon and a positively charged pion (${\rm pK}^-\pi^+$), or the charged conjugate ($\bar{\rm p}{\rm K}^+\pi^-$). The candidates are constructed by combining a negative and two positive tracks for Λ_c^+, while Λ_c^- candidates require a positive and two negative tracks. The secondary vertex is reconstructed, first for a given pair of opposite sign tracks from the minimum of the distance between the two helices representing the tracks, then from the 3-track triplet from a minimisation of the total distance between the three tracks.

To select Λ_c candidates, track quality cuts are made on each individual track. These include a minimum number of clusters in the ITS and TPC, a requirement that the tracks pass the refit stage[2], and that the tracks propagate in a pseudorapidity range where there is minimal degradation of the TPC response. Loose kinematic and topological selections on single-track variables and variables related to the reconstructed candidate are also made, rejecting background in regions where there

[2]The refit stage refers to the 3rd tracking step as described in section 3.3.1 where the tracks are refit towards the reconstructed vertex.

	p-Pb	pp
Track selection	n TPC clusters ≥ 70 TPCrefit = true n. ITS clusters ≥ 4 ITSrefit = true $\lvert\eta\rvert < 0.8$ $p_{\mathrm{T}} > 0.4$ GeV/c	n TPC clusters ≥ 70 TPCrefit = true n. ITS clusters ≥ 4 ITSrefit = true $\lvert\eta\rvert < 0.8$ $p_{\mathrm{T}} > 0.3$ GeV/c
PID proton	$\lvert n\sigma_{TPC\|\|TOF}\rvert < 3$	$p_{\mathrm{T}} < 1.0$ GeV/c: $\lvert n\sigma_{TPC}\rvert < 3$ $p_{\mathrm{T}} > 1.0$ GeV/c: $\lvert n\sigma_{TPC\|\|TOF}\rvert < 3$
PID kaon	$\lvert n\sigma_{TPC\|\|TOF}\rvert < 3$	$p_{\mathrm{T}} < 0.55$ GeV/c: $\lvert n\sigma_{TPC}\rvert < 3$ $p_{\mathrm{T}} > 0.55$ GeV/c: $\lvert n\sigma_{TPC\|\|TOF}\rvert < 3$
PID pion	$\lvert n\sigma_{TPC}\rvert < 3$	$\lvert n\sigma_{TPC}\rvert < 3$
Λ_{c} candidate pre-selection	$\sigma_{vertex} < 0.06$ cm $DCA < 0.05$ cm $\cos\theta_{pointing} > 0$	$\sigma_{vertex} < 0.06$ cm $DCA < 0.05$ cm $\cos\theta_{pointing} > -1$

Table 4.2: The track and candidate pre-selection criteria applied in the p–Pb (left) and pp (right) analyses.

is minimal signal (referred to as pre-selection from here on in). The pre-selection includes a minimum p_{T} cut of the individual tracks, a maximum distance of closest approach of each track from the reconstructed secondary vertex, and in the case of the p–Pb analysis a cut on the cosine of the pointing angle (see section 4.4.2). In addition, loose PID cuts are made on the tracks, based on the difference between the predicted detector response ($\mathrm{d}E/\mathrm{d}x$ in the TPC, or time-of-flight in the TOF) and the measured value, quantified by the number of sigma difference between these two values, nσ. The track and candidate selections are summarised in table 4.2, for the pp and p–Pb analyses.

4.4 Multivariate Selection

Most physics analyses seek to extract a physics 'signal' with the highest possible precision, often achieved by reducing the 'background' by making selections on certain features of the signal in question - these features exploit physics knowledge of the signal, and can include topological and kinematic properties of the signal, or detector responses. Making tighter selections can reduce the background but often comes with the consequence of also reducing the signal. The question of how to optimise the signal/background levels is important, and multivariate techniques are often used

in high-energy physics to perform this task. Machine learning algorithms aim to use samples of known signal/background make-up (often simulated) to respond to samples of unknown make-up, in order to optimise the signal/background separation in these samples.

Due to the large combinatorial background in the analysis presented in this thesis, a multivariate approach has been chosen as an attempt to improve the signal extraction. In addition, this analysis also serves as an interesting benchmark for the application of multivariate techniques within ALICE. The Toolkit for MultiVariate Analysis (TMVA) [213] is a ROOT-integrated machine learning environment which includes a host of multivariate algorithms, and is used for the multivariate selection stage of this analysis.

4.4.1 Boosted Decision Trees

Boosted Decision Trees (BDTs) are one such machine learning algorithm. A series of binary yes/no decisions to fulfil a certain cut condition are used to split the data at each step. The data sample before each cut decision forms a 'node', and the variable and cut position which gives the best signal/background separation is chosen at each node. Events passing the cut condition go one way, and those failing go the other, continuing until certain criteria are fulfilled (e.g number of steps taken, or minimum number of events in a node). The variable phase space is thus split into many regions, and candidates occupying each region are classified as being "signal-like" or "background-like". An example decision tree is visualised in figure 4.2.

Statistical fluctuations in the input training data can cause an instability of the decision trees. Boosting is a way to increase the stability of decision trees with respect to these fluctuations, while also enhancing the performance of the method. Small decision trees are grown (typically with a maximum number of steps, also referred to as 'tree depth', of 2-5), and each subsequent tree is grown with events that were misclassified in the previous tree given a larger weight. This continues until the number of trees grown reaches a specified value. These trees form a 'forest', and the decision tree response is based on a combined response of all trees in the forest.

Many boosting algorithms exist for this task. One such algorithm is Adaptive Boosting, or AdaBoost [214]. Here, every decision tree is grown where events misclassified in the previous tree are given a boost weight α, derived from the misclassification rate of the previous tree err, where

$$\alpha = \frac{1 - err}{err}. \tag{4.2}$$

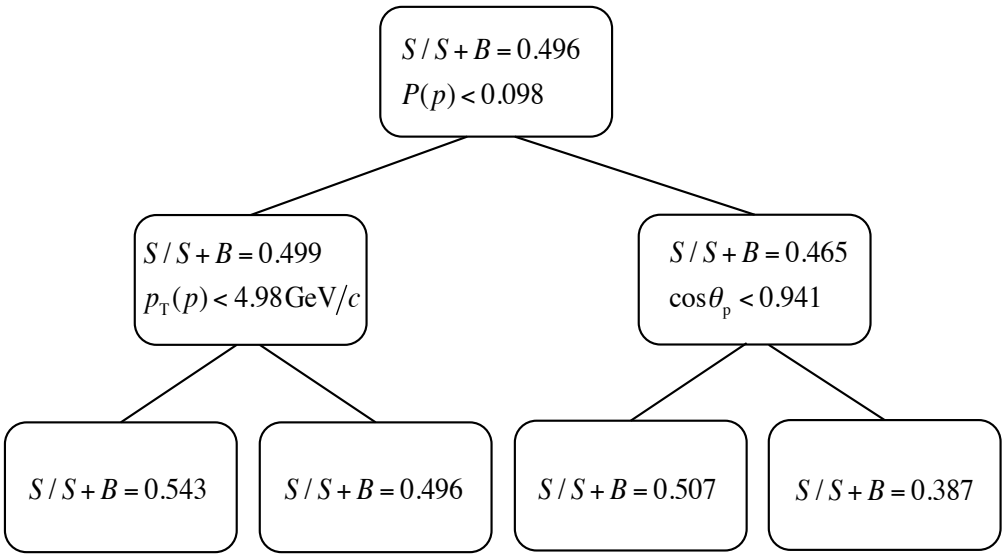

Figure 4.2: The 150th decision tree constructed for the p_{T} bin $6 < p_{\mathrm{T}} < 8$ GeV/c of the p–Pb analysis. Each variable and cut depicted are those which give the best separation between signal and background when cut on, where the 'signal-like' sub-sample goes left and the 'background-like' sub-sample goes right.

The weights of the entire event sample are then renormalised so that the sum of all weights remains constant. The final boosted event response $y_{\mathrm{Boost}}(\mathbf{x})$ is then given by

$$y_{\mathrm{Boost}}(\mathbf{x}) = \frac{1}{N} \cdot \sum_{i}^{N} \ln \alpha_i \cdot h_i(\mathbf{x}) \tag{4.3}$$

where $\mathbf{x}$ represents the vector of input variables, $h(\mathbf{x})$ is the response of an individual tree and N is the number of trees in the forest. Large values of $y_{\mathrm{Boost}}(\mathbf{x})$ indicate signal-like events and small values indicate background-like events. The performance of AdaBoost can also be enhanced by enforcing a "slow learning", where the boost weight is modified by an exponent, $\alpha \to \alpha^{\beta}$ with $\beta < 1$, which allows for a larger number of boost steps.

4.4.2 Discriminating Variables

A sketch of the $\Lambda_{\mathrm{c}}^{+} \to \mathrm{pK}^{-}\pi^{+}$ decay is shown in figure 4.3. Weakly decaying charmed hadrons (D-mesons, Λ_{c} baryon) have proper decay lengths $c\tau$ greater than or of the order of the impact parameter resolution of the ALICE tracking detectors. The reconstruction of weakly-decaying particles in ALICE is therefore based on exploiting the

displacement of the particles from the primary vertex. The Λ_c baryons proper decay length ($c.\tau \sim 60$ μm) means resolving its decay vertex is at the limit of the current detector, although selection can still be made on variables relating to the displaced vertex. The variables used as BDT input include kinematic variables (variables relating to the kinematics of the final state), topological variables (variables relating to the Λ_c decay topology), and PID variables (variables relating to particle identification information). A total of 12 variables are used in this analysis:

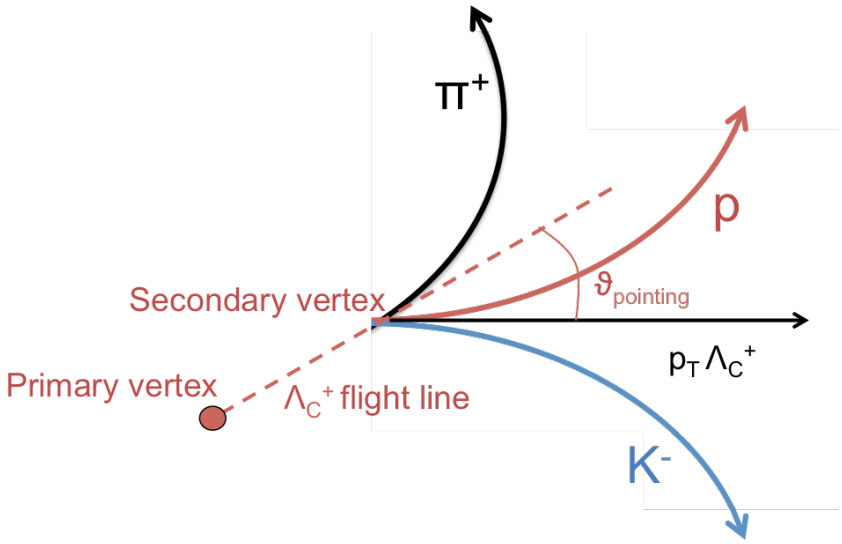

Figure 4.3: A sketch of the $\Lambda_c^+ \to pK^-\pi^+$ decay, representing the primary and secondary vertices.

4.4.2.1 Kinematic

The transverse momentum of the Λ_c candidate decay products, the proton, kaon and pion, are used as BDT input variables. These are referred to from now on as $p_T(p)$, $p_T(K)$ and $p_T(\pi)$. While $p_T(p)$ and $p_T(K)$ not offer very high discrimination power on their own, $p_T(\pi)$ can offer reasonable discrimination, particularly at high Λ_c p_T.

4.4.2.2 Topological

The topological variables included as BDT input variables include:

- **Cosine of the pointing angle:** The pointing angle $\theta_{pointing}$ is defined as the angle between the momentum vector of the reconstructed particle (in this case the Λ_c), and the decay path from the primary vertex to the secondary vertex.

A value of $\cos\theta_{pointing} \sim 1$ would be expected for perfect vertex reconstruction, however in practice the finite detector resolution means the $\cos\theta_{pointing}$ distribution of the Λ_c candidates is smeared out to lower values. This variable offers good discrimination, especially at higher p_T where the detector resolution is better.

- **Decay length:** The decay length L is constructed as the distance between the reconstructed primary and secondary vertex. A cut on the decay length separates the signal from background sources originating from the primary vertex.

- **Decay length using track pairs:** The decay length is also calculated between the primary vertex and the point of the distance of closest approach between 2 of the 3 tracks. This is calculated for oppositely charged tracks and is referred to as $dist_{12}$.

- **Quality of reconstructed vertex:** The quality of the reconstructed vertex σ_{vert} is defined as $\sigma_{vert} = \sqrt{\sigma_0^2 + \sigma_1^2 + \sigma_2^2}$, where $\sigma_{0(1,2)}$ is the distance between the secondary vertex and track 0(1,2).

- **Distance of closest approach:** The distance between each track and the reconstructed secondary vertex is known as the distance of closest approach, or DCA and the maximum DCA of the 3 tracks is included as an input variable to the BDT.

- **Normalised decay length in XY plane:** The resolution in the x-y plane (the plane perpendicular to the beam line) is known to be better than that in the z plane, and so projecting a variable in this plane can add discriminating power. The decay length in the XY plane normalised by its uncertainty L_{XY} is included as an input variable to the BDT.

A comparison of the simulated signal and background distributions of these variables are shown in figure 4.4. The corresponding distributions for candidates taken from the sidebands of the invariant mass distribution in data is also shown, which should be reproduced by the background sample. The signal and background samples used for the BDT training is discussed in section 4.4.3, but it is useful to point out here that a separation can be seen in the signal and background distributions, and that the simulated background is well described by the distribution taken from the sidebands of the data sample.

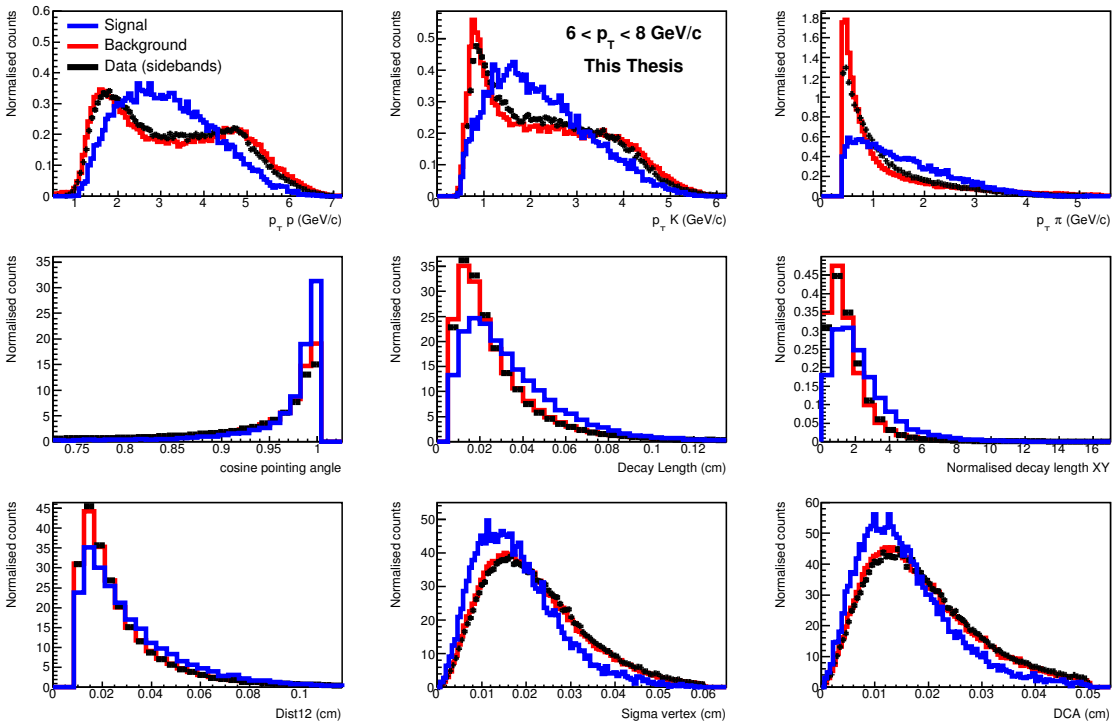

Figure 4.4: A comparison of the signal, background and data sidebands distributions of the discriminating variables used in the p–Pb analysis, in the p_{T} interval $6 < p_{\mathrm{T}} < 8$ GeV/c.

4.4.2.3 PID

Loose pre-selection PID is applied using a cut on the number of sigma a track is away from the predicted PID hypothesis, however the purity of identified particles from this loose PID is relatively low. The Bayesian PID framework [205], a description of which is given in section 3.3.2 allows for a more effective combination of the PID capabilities of the ALICE detectors, and uses a Bayesian approach to identify particles with a

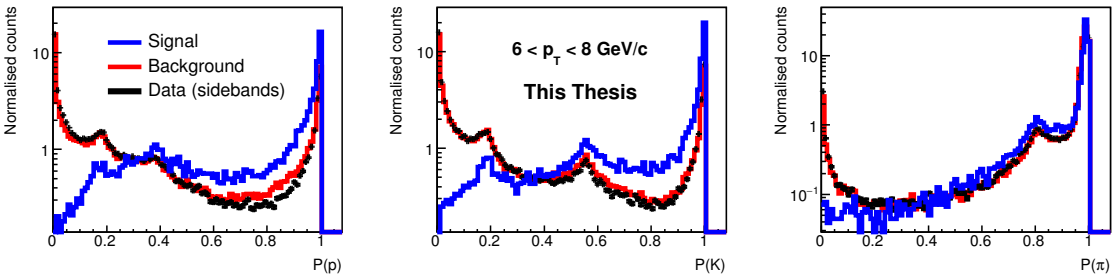

Figure 4.5: A comparison of the signal, background and data sidebands distributions of the PID variables used in the p–Pb analysis, in the p_{T} interval $6 < p_{\mathrm{T}} < 8$ GeV/c.

much lower contamination rate. The posterior probabilities for each track to be the identity it has been assigned, $P(p)$, $P(K)$ and $P(\pi)$ for protons, kaons and pions respectively, offer very good discrimination power. The 3 corresponding variables are included as input variables to the BDT. Figure 4.5 shows an example of these variables in one p_{T} bin.

Each variable's discriminating power can be visualised by plotting the signal efficiency as a function of the background rejection, known as a Receiver-Operator-Characteristic (ROC) curve. An example ROC curve describing cuts on each single kinematic variable used in this analysis in shown in figure 4.6. The area under the curve can be used as a measure of the discrimination power, with a purely diagonal line with area 0.5 representing the 'worst' scenario where the variable offers no discrimination power, and a curve with area 1 representing the 'perfect' case where full background rejection can be achieved without loss of signal. The curves in the figure can be seen to be close to the diagonal line, meaning selection on each single variable offers poor discrimination power. On its own, the cosine of the pointing angle can be seen to give the best discrimination, while $dist_{12}$ offers the poorest discrimination.

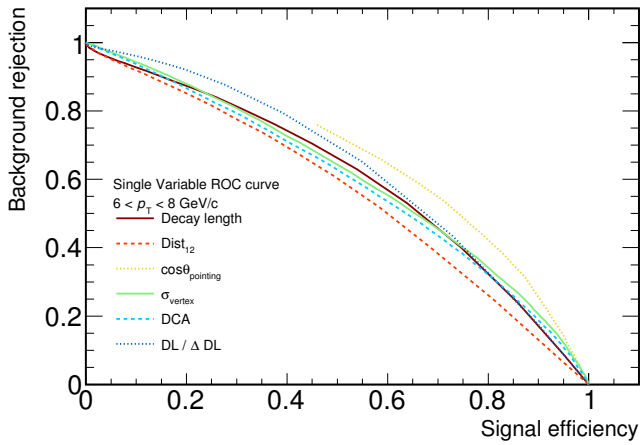

Figure 4.6: The ROC curve (signal efficiency vs. background rejection) of the kinematic variables used in the p–Pb analysis, in the p_{T} interval $6 < p_{\mathrm{T}} < 8$ GeV/c.

4.4.3 BDT training

The BDT algorithm is trained using signal and background samples chosen to represent the true signal and background make-up in data. The BDT parameters were chosen using an iterative process, where each parameter was changed one by one,

and the performance was evaluated using the ROC curve for each particular parameter set. These parameters include the boosting algorithm, the total number of trees trained, the maximum depth of each tree, the minimum node size before terminating the branch as a percentage of the total sample size, and the total number of cuts to use when scanning across each variable to determine the best cut at each node. There is also the option to use bagged boost, where the classifier is repeatedly trained using a randomised sub-sample of training events, and the final classifier is an average of these individual classifiers. A maximum depth of 2 and total number of trees of 500 gave the optimal signal efficiency. Each of the other options showed minimal improvement with respect to the default values. A summary of the parameters chosen is shown in table 4.3. The settings were chosen to be the same for the pp and p–Pb analyses.

BDT parameter	Choice
N. trees	500
Max. depth	2
Min. node size	2.5%
Use bagged boost	no
N. cuts	50
AdaBoost β	0.5

Table 4.3: The BDT parameters chosen for this analysis, using the adaptive boost algorithm.

Simulated data is used to generate samples that accurately represent the kinematics of the signal and background of this analysis. Monte Carlo (MC) generators aim to simulate the physics of particle collisions, including the hard scattering processes and following parton showers as well as the underlying event, hadronisation, and unstable particle decays. The particles created from event generation are then transported through the simulated ALICE detector using GEANT3 [215], which includes a detailed description of the detector layout, detector response and the conditions of the luminous region. The MC simulations are tuned to include the conditions specific to each run in data. Different training samples were used for the pp and p–Pb analyses.

The signal sample for the training stage in the p–Pb analysis was taken from simulated p–Pb events, using the HIJING event generator [96]. Around 4 million p–Pb events were generated. Λ_c baryons were forced to decay to a $pK\pi$ final state, and the event was required to contain at least 1 prompt Λ_c generated in the event with $|y| < 1.5$ and with the Λ_c daughters within $|y| < 1.5$. This leads to around

$100,000$ Λ_c reconstructed in the p_T range of this analysis (∼11,000 - 20,000 in each p_T bin).

The signal sample for the training stage in the pp analysis was taken from simulated pp events, using the PYTHIA 6.4.21 event generator [91] with Perugia-0 tuning [216]. Around 4 million pp events were generated. The events were required to contain at least one $c\bar{c}$ pair. This leads to around $40,000$ Λ_c reconstructed in the p_T range of this analysis (∼5,000 - 10,000 in each p_T bin).

The background samples for the training stage in both analyses were created using the sidebands of the invariant mass distribution in the respective data samples for each collision system. The sidebands are defined as $M_{\Lambda_c} - 0.18 < M < M_{\Lambda_c} - 0.13$ GeV/c^2 and $M_{\Lambda_c} + 0.13 < M < M_{\Lambda_c} + 0.18$ GeV/c^2 where $M_{\Lambda_c} = 2.2864$ GeV/c^2. 90,000 Λ_c candidates were used, randomly selected over all runs in the respective data samples. This number was chosen to be larger than the signal samples so that statistical fluctuations are not significant, and constant as a function of p_T.

The training stage was performed using one dataset to train the BDT, and another as an independent test sample. This split is done to ensure a statistically independent evaluation of the BDT algorithm is performed with the test sample, which is used to test the level of overtraining, which can occur when statistical fluctuations are interpreted as features of a variable. The background sample was split in half, with 45,000 candidates used for both the training and test sample. In order to maximise the number of signal candidates, a small fraction of the total signal sample was taken as a test sample (1000 signal events), and the rest were used as the training sample. The BDT training is performed individually for each p_T interval. Shown in figure 4.7 are the BDT responses in pp and p–Pb collisions in the range $6 < p_T < 8$ GeV/c. It can be seen that the training and test sample agree reasonably well, indicating minimal overtraining.

The ROC curve for the p–Pb analysis is shown in figure 4.8 in the p_T interval $6 < p_T < 8$ GeV/c. For comparison, the ROC curve corresponding to optimised rectangular cuts is also shown - it can be seen that the BDT performs better than the optimised rectangular cuts, with a higher background rejection for a given signal efficiency. This suggests the use of BDTs in reducing the background in this analysis is promising. The next step is to apply the trained BDTs on the pp and p–Pb data samples.

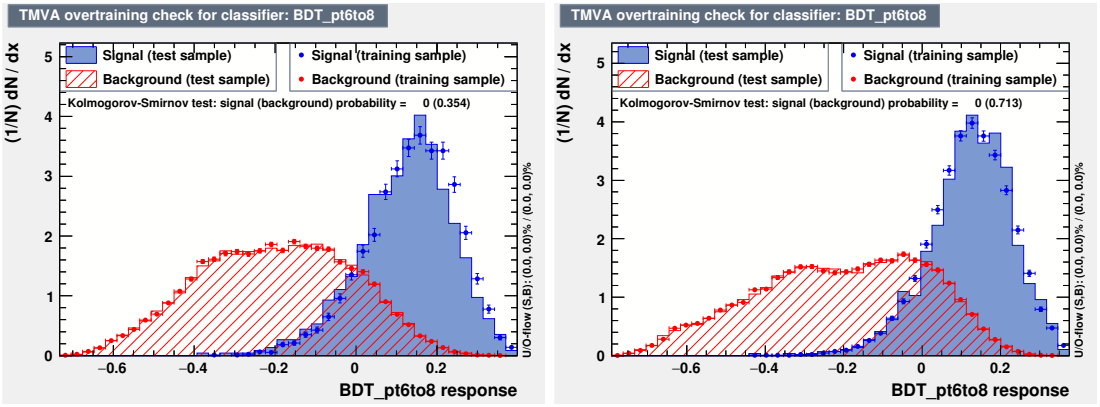

Figure 4.7: The BDT response evaluated with the training and test samples in the pp analysis (left) and in the p–Pb analysis (right) in the p_{T} interval $6 < p_{\mathrm{T}} < 8$ GeV/c.

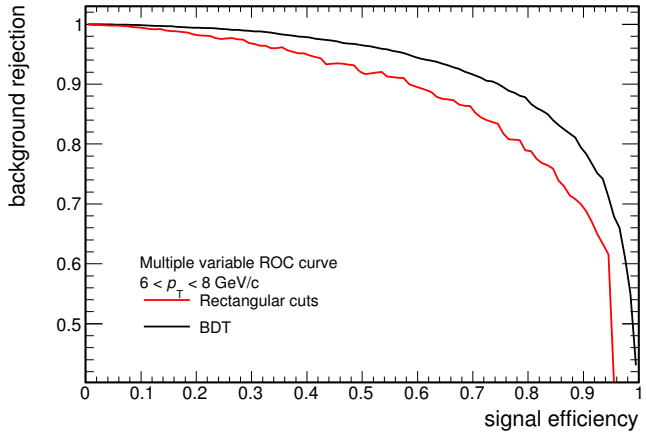

Figure 4.8: The ROC curve (signal efficiency vs. background rejection) for the BDT analysis and for optimised rectangular cuts, in p–Pb collisions in the p_{T} interval $6 < p_{\mathrm{T}} < 8$ GeV/c.

4.4.4 BDT application

The BDT weights learned during the training phase are applied to the datasets described in section 4.2. The BDT response is evaluated for each 3-prong Λ_{c} candidate. Figure 4.9 shows the BDT response distributions in data, compared to the response for simulated signal and background test samples. The BDT response in data matches the background response well. In this case, as the signal-to-background ratio is very low, the signal in data is negligible and it is expected that the data would follow the background. The signal sample is that used for the efficiency correction described in section 4.6.1, and a cut on the signal response determines the BDT cut efficiency. While the signal response distribution in data cannot be known, the similarity be-

tween the background and the data response indicates that the simulated events provide a good description of data, and the transformation that the BDT algorithm performs on the simulation matches that performed on data, so a cut on the signal response should accurately describe the amount of signal rejected by the cut when determining the BDT cut efficiency. Any discrepancy between the 'real' and 'simulated' signal response will manifest itself in a systematic difference in the BDT cut efficiency, and the systematic uncertainty associated with the BDT cut is estimated by varying the BDT response cut and checking the effect on the final result. This is described in section 4.7.2.

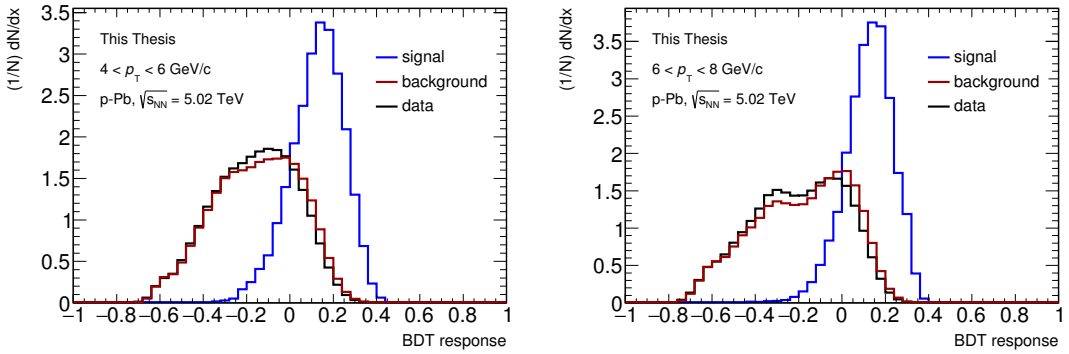

Figure 4.9: Examples of the BDT response for signal and background Λ_c candidates from simulated p–Pb collisions, compared to Λ_c candidates from p–Pb data, in the invariant mass windows $4 < p_T < 6$ GeV/c (left) and $6 < p_T < 8$ GeV/c (right).

The optimal cut on the BDT response is determined by scanning the response, and fitting the Λ_c invariant mass peak at each cut. For this scan, a second order polynomial is used to fit the background, and a Gaussian is used to fit the signal, as described in the next section. The signal and background entries are evaluated by integrating the signal and background regions within $\pm 3\sigma$ of the peak width, and the best cut is chosen by maximising the significance (defined as $S/\sqrt{S+B}$) while ensuring the mean and width of the peak agree with the MC values. This was performed for each p_T interval. Figure 4.10 shows the significance of the signal extraction in an example p_T interval for both the pp and p–Pb analyses, where a distinctive 'peak' can be seen at a BDT response close to the chosen optimal cutting point. It is noted here that the significance as expressed above is roughly equal to the inverse of the statistical uncertainty of the signal - for example, a significance of 4 would denote a statistical uncertainty of $\sim 25\%$.

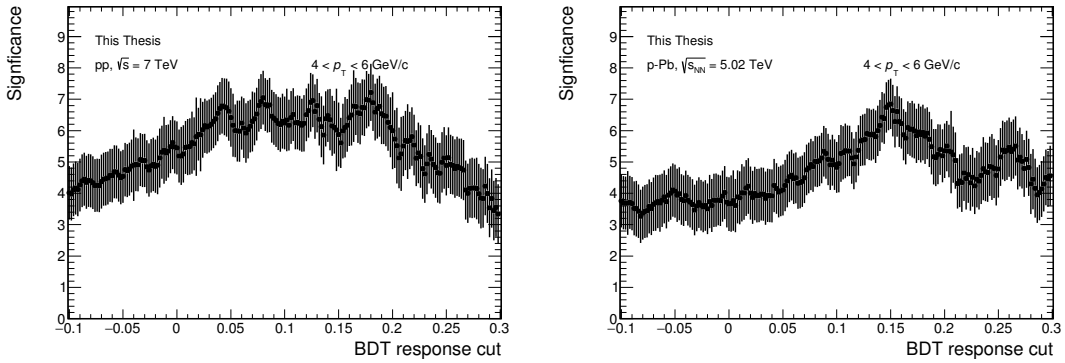

Figure 4.10: The BDT cut optimisation procedure: significance of the signal extraction as a function of BDT response in the p_{T} interval $4 < p_{\mathrm{T}} < 6$ GeV/c in the pp analysis (left) and the p–Pb analysis (right).

4.5 Signal extraction

The Λ_{c} signal (raw yield) is extracted via a fit to the invariant mass spectra of the proton-kaon-pion triplets which pass the PID, pre-selection and BDT cuts. The invariant mass of the Λ_{c} candidates is calculated as

$$M(pK\pi) = \sqrt{(E_p + E_K + E_\pi)^2 - (\boldsymbol{p}_p + \boldsymbol{p}_K + \boldsymbol{p}_\pi)^2}. \tag{4.4}$$

The fitting procedure is performed in two steps and is done with the Minuit [217] package. The first step fits the background distribution in the sideband region of the invariant mass spectra. In the second step the signal is fitted with a Gaussian function and the full signal-plus-background distribution is then re-fit using a combined fit function. For the background a second-order polynomial function is used and for the signal a Gaussian function is used, such that the combined fit function has the form

$$f(M) = P_1 + P_2 M + P_3 M^2 + \frac{P_4}{\sqrt{2\pi}\sigma} e^{-\frac{(M+\mu)}{2\sigma^2}} \tag{4.5}$$

where P_1, P_2 and P_3 are coefficients of the polynomial background function, P_4 is a normalisation coefficient of the Gaussian function, and σ and μ are the width and mean of the Gaussian signal function, respectively. The signal is determined as the integral of the Gaussian between $\pm 3\sigma$, and the background is determined as the integral of the background function in the same mass range.

Since the width of the peak has been verified to agree with expectation from simulation, and the relatively small significance ($\sim 4 - 7$) can lead to instability of the fitting procedure, the width of the Gaussian function used to model the signal is

fixed to that obtained from a fit to the invariant mass of the signal in Monte Carlo simulation. It is verified that this does not worsen the χ^2/ndf of the fit, and that the final result of the fit does not systematically affect the raw yield by a significant amount. The systematic effect of this choice is accounted for in the estimation of the systematic uncertainty of the raw yield extraction (see section 4.7.1).

Figure 4.11 shows the Λ_c candidate invariant mass spectra in the p-Pb analysis, in 4 p_T intervals. The mean and width of the mass peak is shown, as well as the signal and significance calculated as $S/\sqrt{S+B}$, and their respective errors. Figure 4.12 shows the invariant mass spectra in the pp analysis, in 4 p_T intervals. The width of the peak is larger in the pp analysis, due to the worse p_T resolution of single tracks, and for this reason the bin width in the invariant mass spectra is chosen to be larger than that in the p–Pb analysis (6 MeV/c^2 instead of 4 MeV/c^2). As well as this, the default invariant mass window is also widened slightly ($2.19 - 2.38$ GeV/c^2 instead of $2.20 - 2.37$ GeV/c^2) to allow for a slightly larger background fit window. Table 4.4 summarises the cut on the BDT response and the corresponding extracted raw yield for each p_T interval, in each collision system.

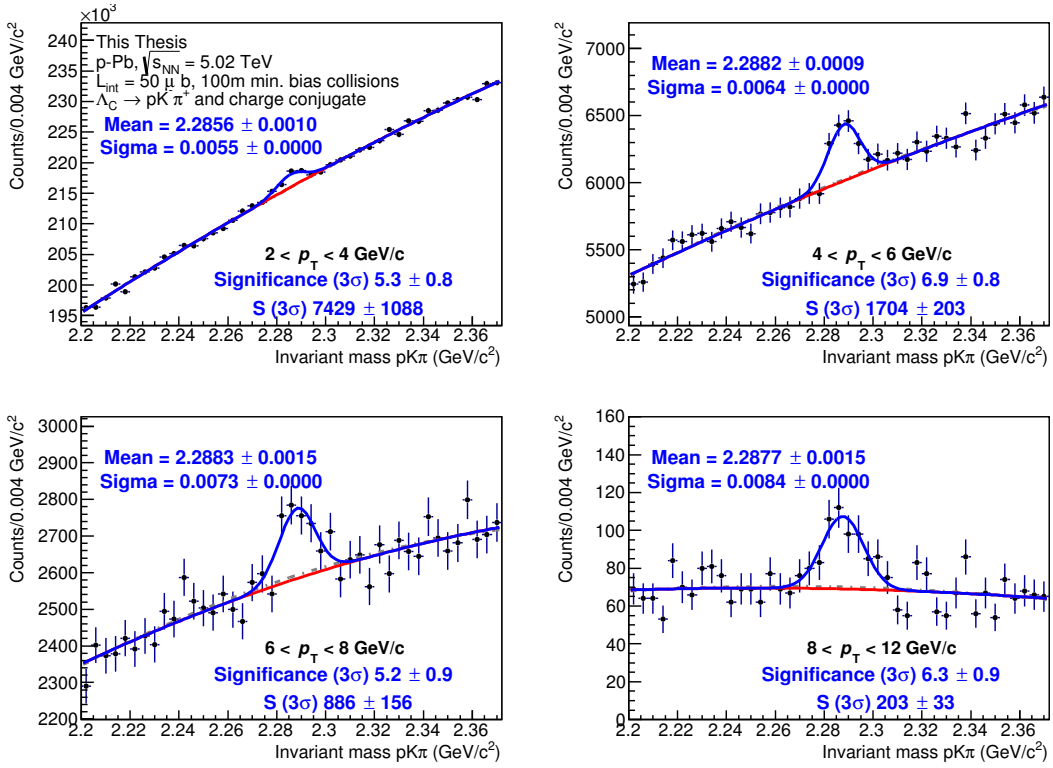

Figure 4.11: The Λ_c candidate invariant mass spectra in 4 p_T intervals, and the fit to the signal and background, in the p-Pb analysis.

The width and mean of the Gaussian peaks used to model the signal when leaving both the width and mean as free parameters of the fit are shown in figure 4.13 and figure 4.14 for the pp and p–Pb analyses, respectively, in comparison to the width and mean obtained in MC. The values obtained from the fit agree with the MC expectation within $\sim 2\sigma$.

	p–Pb		pp	
p_{T} interval (GeV/c)	BDT cut	raw yield	BDT cut	raw yield
$2 < p_{\mathrm{T}} < 4$	0.02	7429 ± 1088	0.07	4130 ± 612
$4 < p_{\mathrm{T}} < 6$	0.15	1704 ± 203	0.08	1988 ± 244
$6 < p_{\mathrm{T}} < 8$	0.03	886 ± 156	0.15	298 ± 71
$8 < p_{\mathrm{T}} < 12$	0.17	203 ± 33	0.16	148 ± 40

Table 4.4: BDT response cut and corresponding raw yield extracted in each p_{T} interval in the pp and p–Pb analyses.

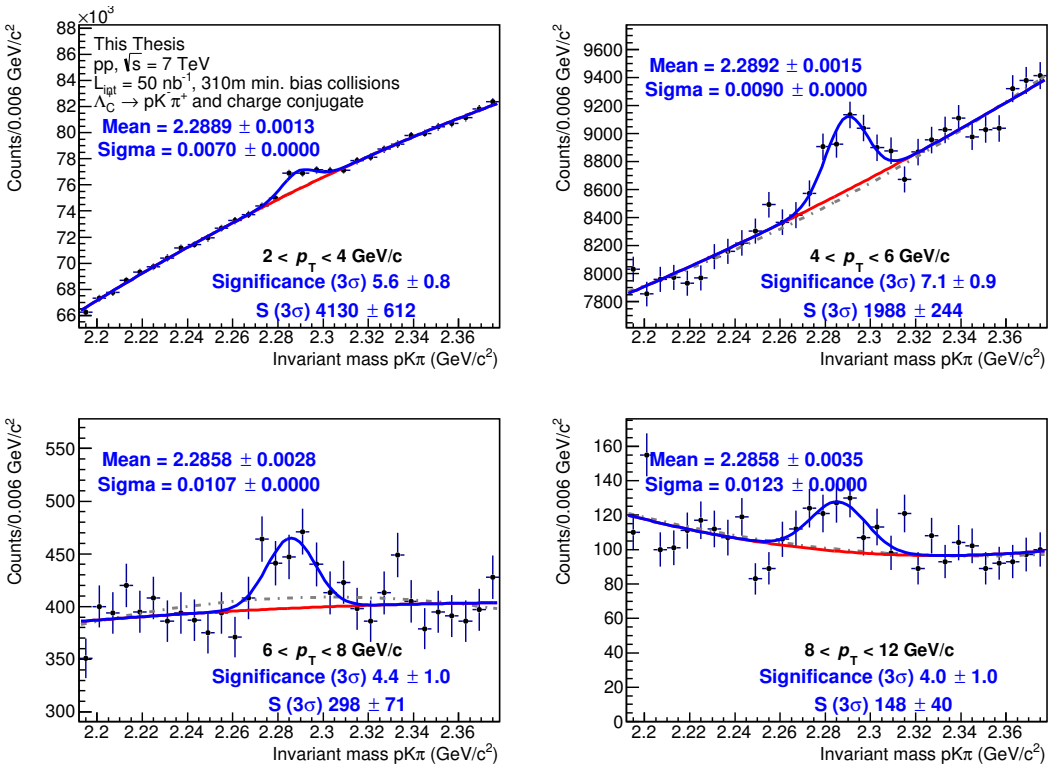

Figure 4.12: The Λ_{c} candidate invariant mass spectra in 4 p_{T} intervals, and the fit to the signal and background, in the pp analysis.

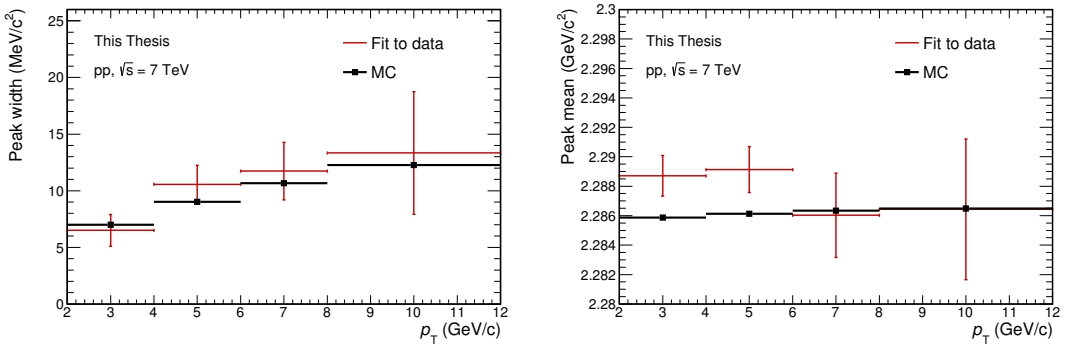

Figure 4.13: The mean and width of the Gaussian function used to model the Λ_c signal in the pp analysis, compared to expectation from MC.

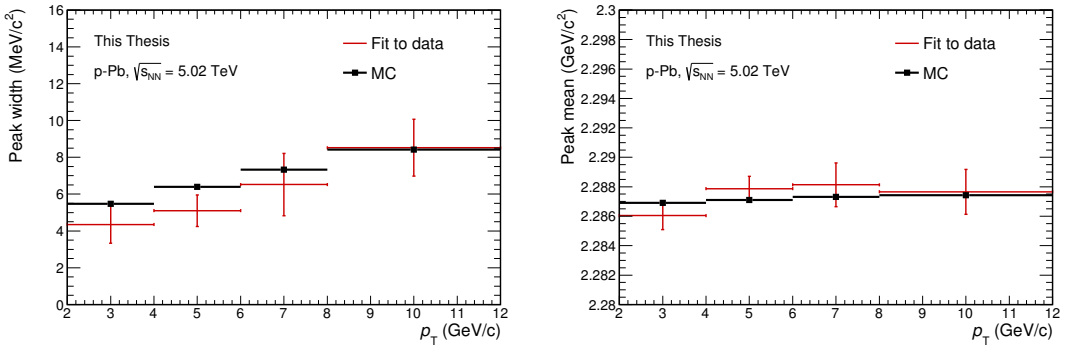

Figure 4.14: The mean and width of the Gaussian function used to model the Λ_c signal in the p–Pb analysis, compared to expectation from MC.

4.6 Corrections

To obtain the p_T-differential production cross section of the prompt Λ_c^+ baryon in pp and p–Pb collisions in the lab rapidity region $|y_{lab}| < 0.5$, the raw yield needs to be corrected with various factors, which are described in the following sections.

4.6.1 Efficiencies

For the efficiency correction in the pp analysis, pp events are simulated using PYTHIA [91, 92] with Perugia-0 tune [216], and each event is required to contain either a $c\bar{c}$ or $b\bar{b}$ pair. For the efficiency correction in the p–Pb analysis, p–Pb events are simulated using the HIJING event generator [96], and every event is enriched with a $c\bar{c}$ or $b\bar{b}$ pair using PYTHIA with Perugia-0 tune. In both collision systems, each Λ_c baryon generated is forced to decay to either $\Lambda_c^+ \to pK^-\pi^+$ or $\Lambda_c^+ \to pK_s^0$. The particles created from event generation are then transported through the simulated ALICE

detector using GEANT3 [215].

The total efficiency is determined as the number of Λ_c signal candidates which pass the reconstruction, pre-selection, PID and BDT cuts, divided by the total number of generated Λ_c within the fiducial acceptance region $|y| < y_{\rm fid}(p_{\rm T})$, and is calculated in two steps. The first step determines the 'pre-selection' efficiency ϵ_{pre} which takes into account all steps of reconstruction and selection prior to the BDT step. The second step determines the BDT cut efficiency ϵ_{BDT}, which is the fraction of Λ_c signal which pass the BDT selection. The BDT response of each Λ_c signal candidate passing the pre-selection cuts is determined, and the efficiency is calculated using the number of candidates with a BDT response greater than that applied in data, $\epsilon_{BDT} = N_{\rm cut}/N_{\rm nocut}$, for each $p_{\rm T}$ interval. The total efficiency is then calculated as $\epsilon_{total} = \epsilon_{pre} \times \epsilon_{BDT}$, referred to as ϵ from now on.

The Λ_c baryon decays to a pKπ final state either non-resonantly or via an intermediate resonant state. Due to the different q values of each resonant decay the kinematics of each decay will vary, and thus the efficiency for each decay may vary due to the acceptance region of the detector and the kinematic/topological selection. The total efficiency ϵ is therefore taken as the weighted average of all total efficiencies ϵ_i (where i represents each resonant/non-resonant decay), with weights w_i equal to their branching ratios as listed in table 4.1, given by

$$\epsilon = \frac{\sum_{i=1}^{4} w_i \cdot \epsilon_i}{\sum_{i=1}^{4} w_i}. \tag{4.6}$$

The acceptance factor *Acc* takes into account the fiducial acceptance of the detector, defined as the number of Λ_c within $|y| < 0.5$ divided by the number of Λ_c with all daughters within the fiducial acceptance of the detector $y_{\rm fid}$ (see section 4.1). This is calculated using a Monte Carlo simulation, where Λ_c particles are generated within PYTHIA and the geometry of the decay is used to calculate *Acc*. Since the acceptance factor is purely geometrical, the same factor can be used in both pp and p–Pb collisions, for both prompt and non-prompt Λ_c decays.

GEANT3 versions older than v1-14-6 are known to contain an incorrect description of the interaction cross section of protons and antiprotons with the detector material, where GEANT3 overestimates the elastic scattering cross section by around 25% [218]. Since the pp simulation used to determine the efficiency was produced using an older version of GEANT3 (v1-11-17), a correction is applied for the pp efficiency calculation. The Monte Carlo simulation package FLUKA [219, 220] gives a more accurate description of the (anti)proton interaction cross sections, so a correction

factor $C(p_{\mathrm{T}}) = \epsilon_{Geant3}(p_{\mathrm{T}})/\epsilon_{FLUKA}(p_{\mathrm{T}})$, which is a function of the (anti)proton p_{T} is calculated, and the Λ_{c} efficiency is obtained dividing the $\Lambda_{\mathrm{c}}^{+}/\Lambda_{\mathrm{c}}^{-}$ efficiencies by these factors, which are shown in the left panel of figure 4.15. This leads to around a 4%(17%) higher efficiency for $\Lambda_{\mathrm{c}}^{+}(\Lambda_{\mathrm{c}}^{-})$ in the Λ_{c} p_{T} interval $2 < p_{\mathrm{T}} < 3$ GeV/c, reducing to around 0.7%(6%) in the interval $8 < p_{\mathrm{T}} < 12$ GeV/c. The corrected efficiency is then taken as the average of the corrected efficiencies for Λ_{c}^{+} and Λ_{c}^{-}. The final $Acc \times \epsilon$ with and without the correction factor are shown in the right panel of figure 4.15. This leads to an 11% difference at low p_{T}, reducing to a 4% difference at high p_{T}.

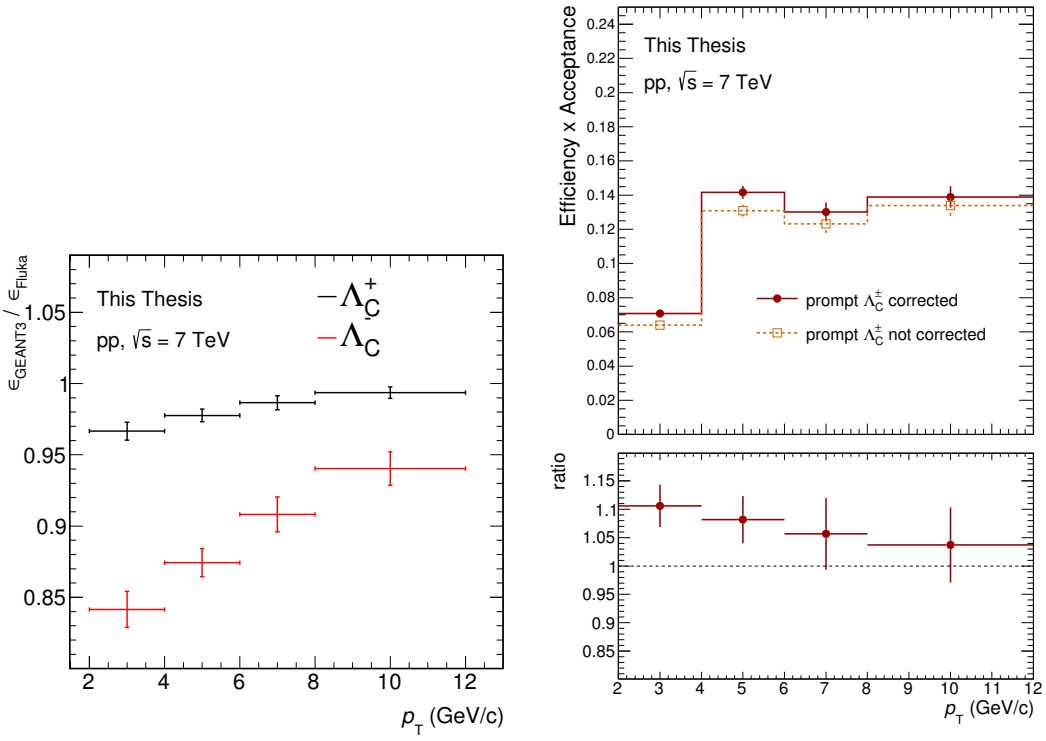

Figure 4.15: Left: Correction factors for Λ_{c}^{+} and Λ_{c}^{-} to take into account the overestimation of the elastic scattering cross section of protons and antiprotons in GEANT3. Right: A comparison of the corrected and non-corrected total efficiency × acceptance.

The pp pre-selection efficiency, BDT cut efficiency, acceptance factor and total efficiency × acceptance are shown in figure 4.16. In pp collisions the total efficiency x acceptance goes from around 0.07 for prompt and non-prompt Λ_{c} at low p_{T}, up to around 0.14 for prompt Λ_{c} and 0.21 for non-prompt Λ_{c} at high p_{T}.

The p–Pb pre-selection efficiency, BDT cut efficiency, acceptance factor and total efficiency × acceptance are shown in figure 4.17. In p–Pb collisions the total efficiency

x acceptance is around 0.05 for prompt and 0.07 for non-prompt Λ_c at low p_T, increasing to around 0.15 for prompt and 0.24 for non-prompt Λ_c at high p_T. The final efficiency in this figure is weighted to reproduce the correct multiplicity distribution, as explained in the next section.

In p–Pb collisions the difference between the prompt and non-prompt efficiencies is due to the pre-selection (as explained in section 4.3) and the cut on the BDT response. The similarity between the prompt and non-prompt Λ_c efficiencies in pp collisions is due both to the loose pre-selection on the Λ_c decay and the loose cut on the BDT response. A tighter cut on the BDT response at higher p_T selects a larger fraction of non-prompt Λ_c, hence the larger efficiency at high p_T. The tighter pre-selection in p–Pb collisions already removes a larger fraction of prompt Λ_c, while the BDT cut roughly removes an equal fraction of prompt and non-prompt Λ_c, which is reflected in the p–Pb efficiencies.

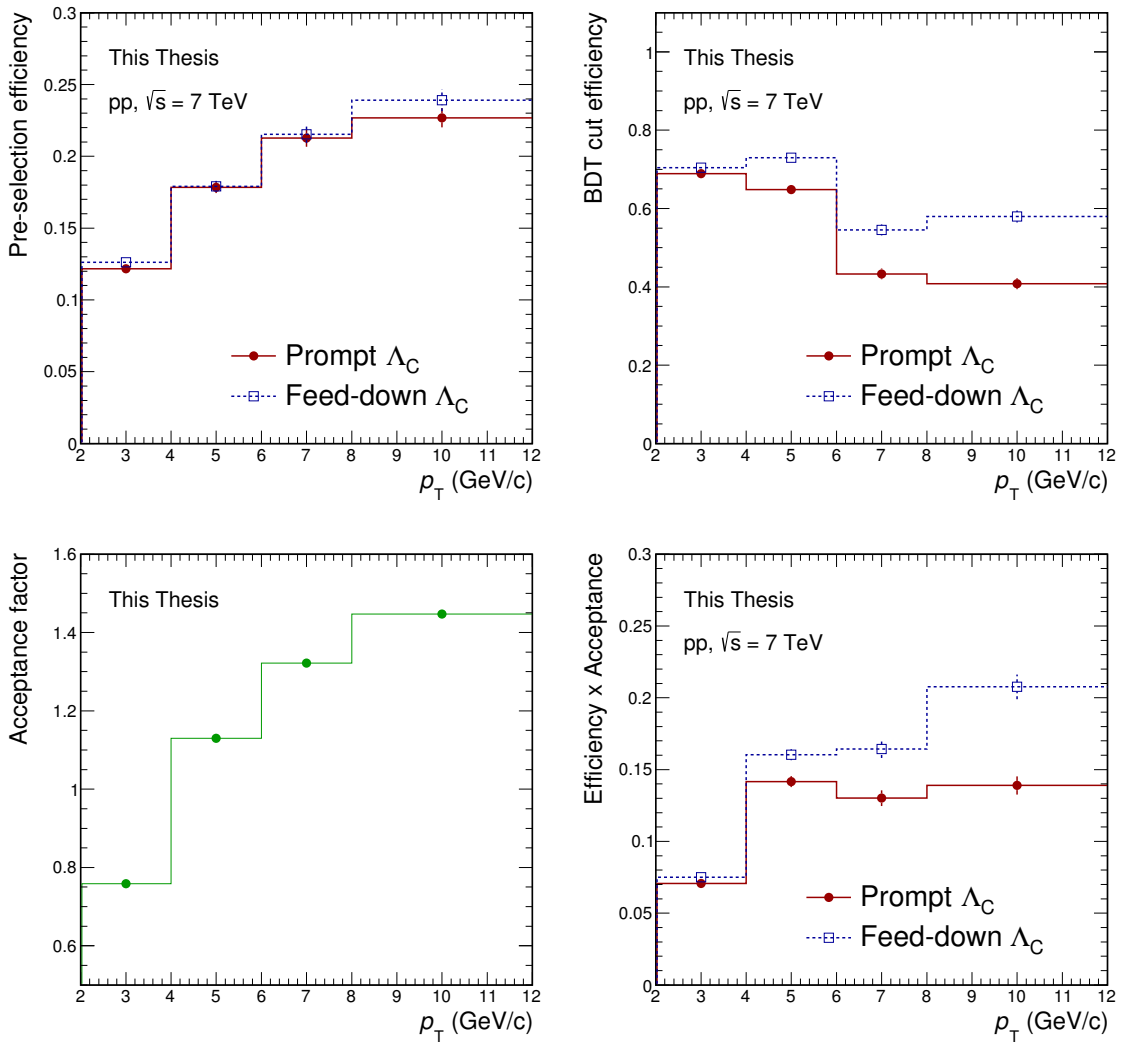

Figure 4.16: Clockwise from top left: pre-selection efficiency, BDT cut efficiency, total efficiency x acceptance and acceptance factor, for the pp analysis. The prompt and non-prompt Λ_c efficiencies are shown separately. The total efficiency x acceptance includes the proton/antiproton efficiency correction described in the text.

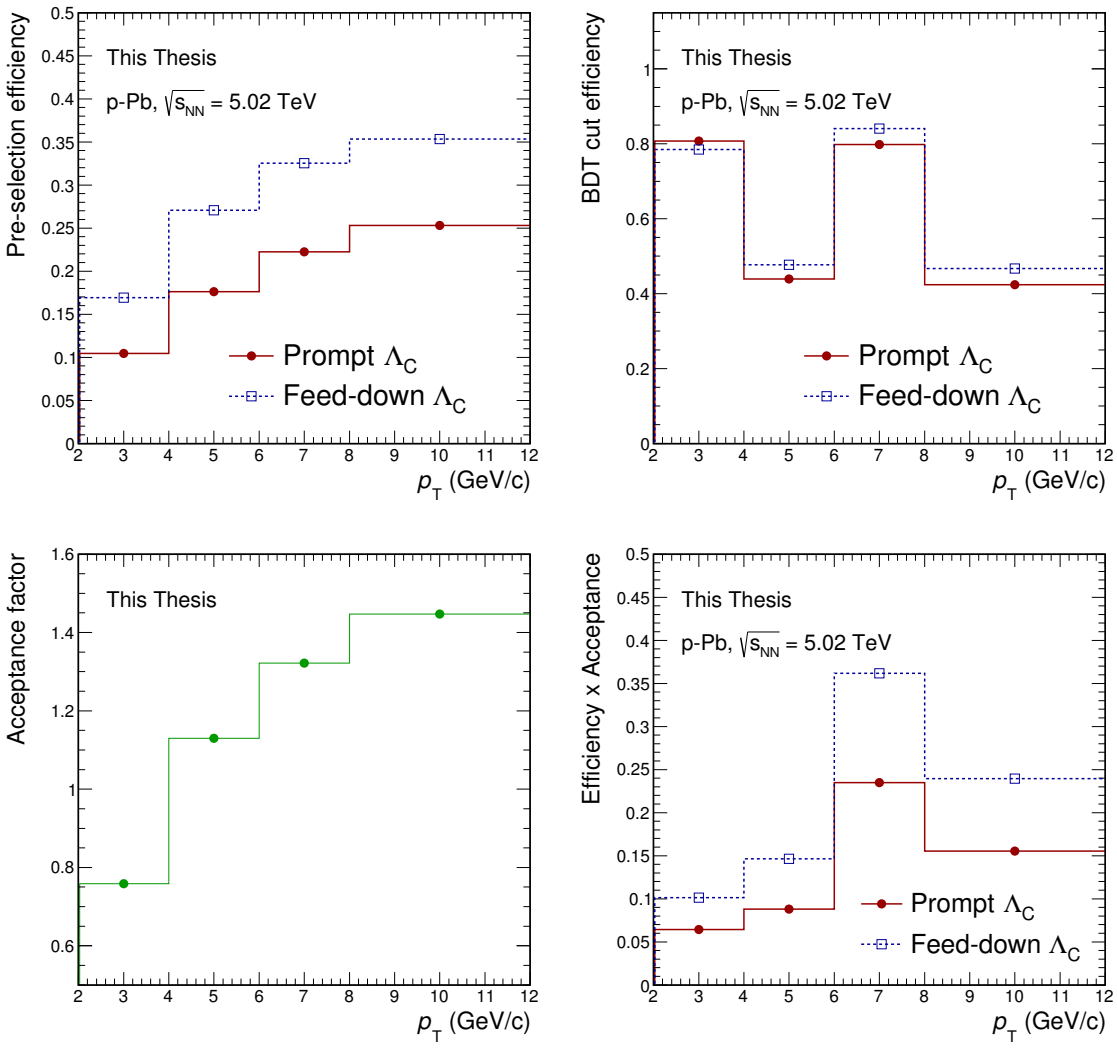

Figure 4.17: Clockwise from top left: pre-selection efficiency, BDT cut efficiency, total efficiency x acceptance and acceptance factor, for the p-Pb analysis. The prompt and non-prompt Λ_c efficiencies are shown separately.

4.6.1.1 p–Pb multiplicity distribution correction

The multiplicity in the p–Pb MC simulations used in this analysis does not describe that in data, which can affect the efficiency correction. In particular, the resolution on the reconstruction of the primary vertex is improved for events with a higher multiplicity, which leads to a multiplicity dependence on the efficiency when making selections on variables related to the primary vertex.

To determine the effect of the multiplicity bias on the reconstruction and selection efficiency, the MC sample is weighted to reproduce the multiplicity distribution in data. 4 weights have been tried:

1. The ratio data/MC of the number of tracklets distribution for events passing the event selection criteria
2. The ratio data/MC of the number of tracklets distribution for events passing the physics selection criteria
3. The ratio data/MC of the number of tracklets distribution for events with at least one D^0 candidate
4. The ratio data/MC of the number of tracklets distribution for events with at least one D^0 candidate in the D^0 mass window

Weights (3) and (4) have been evaluated for D^0 candidates, and using them for this analysis assumes that the $N_{tracklets}$ distribution looks similar for events with D^0 and Λ_c candidates. This is a fair assumption since selecting on the mass region for weight (4) does not change the weight much, suggesting that the multiplicity distribution for 'hard' events containing a D^0 candidate should be very similar to 'hard' events containing a Λ_c candidate. The efficiency using each weight is shown in figure 4.18 - a maximum deviation of around 3% is seen for the efficiency when weighting. Weight (4) is used for the final efficiencies in p–Pb collisions, and the total efficiency x acceptance shown in figure 4.17 takes into account these weights.

4.6.2 Feed-down correction

As described above, the Λ_c baryons reconstructed in this analysis originate not only from prompt Λ_c production, but also from decays of non-prompt Λ_b^0 baryons which then weakly decay to $\Lambda_c + X$. This is corrected for by calculating the fraction of the reconstructed Λ_c yield which is 'prompt', called f_{prompt}. The method of feed-down

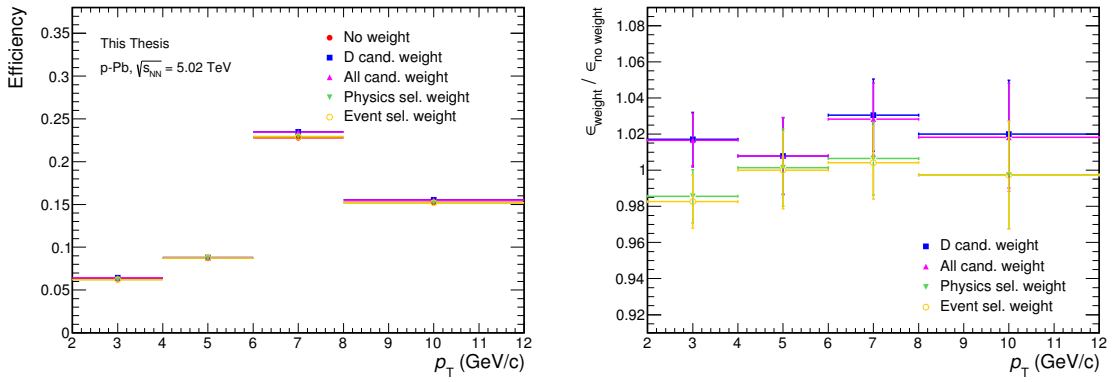

Figure 4.18: The efficiency weighted according to the 4 weights described in the text for the p–Pb analysis, where each weight considers different multiplicity distributions. The efficiency is shown on the left, and the ratio with the unweighted efficiency is shown on the right

correction used in this analysis uses the measured raw yield $N_{\rm raw}^{\Lambda_c}$, and the feed-down raw yield $N^{\Lambda_c,{\rm feed-down}}$ which was estimated from the production cross section of Λ_b^0 calculated within FONLL [152] where the fragmentation fraction of beauty quarks into Λ_c baryons $f(b \to \Lambda_c) = 0.073$ was taken as the combined average from e^+e^- collisions at LEP [221] and the $\Lambda_b^0 \to \Lambda_c + {\rm X}$ decay kinematics were modelled using the EvtGen [222] package. f_{prompt} was then calculated applying the correction factors for acceptance and reconstruction efficiency $\epsilon_{\Lambda_b^0 \to \Lambda_c}$ of non-prompt Λ_c, the branching ratio BR of $\Lambda_b^0 \to \Lambda_c + {\rm X}$ and the integrated luminosity $\mathcal{L}_{\rm int}$:

$$f_{\rm prompt} = 1 - \frac{N^{\Lambda_c,{\rm feed-down}}}{N_{\rm raw}^{\Lambda_c}} = 1 - \frac{1}{N_{\rm raw}^{\Lambda_c}} \cdot \mathcal{L}_{\rm int} \cdot {\rm BR} \cdot ({\rm Acc} \times \epsilon_{\Lambda_b^0 \to \Lambda_c}) \int_{\Delta y \Delta p_T} \frac{{\rm d}\sigma_{\rm FONLL}^{\Lambda_b^0}}{{\rm dydp_T}}. \quad (4.7)$$

This is referred to as the 'N_b' method. $f_{\rm prompt}$ calculated with this method is around 0.98 at low $p_{\rm T}$, and around 0.92 at high $p_{\rm T}$, corresponding to a non-prompt fraction of the raw yield of $\sim 2 - 8\%$. As explained in section 4.7.5, two additional methods are considered in order to evaluate the systematic uncertainty on f_{prompt}, considering in addition predictions of charm production, and the measured Λ_b^0 production cross section.

4.7 Systematic Uncertainties

This section describes the evaluation of the systematic uncertainties in these analyses. The uncertainties on the cross section measurements arise from uncertainties on the different ingredients used for these measurements (equation 2.1) - namely the raw yield, the prompt fraction, the efficiency correction (which is a combination of the uncertainty on the tracking efficiency, the uncertainty on the pre-selection and the uncertainty on the BDT selection), and the global uncertainty associated to the branching ratio of $\Lambda_c^+ \to pK^-\pi^+$ and the integrated luminosity. Each uncertainty source is assumed to be uncorrelated with all other sources, so the final uncertainty is determined summing all systematic uncertainties in quadrature. The total relative systematic uncertainties for each source are summarised at the end of this section.

4.7.1 Yield Extraction

When choosing a fitting procedure to extract the raw yield, there is some uncertainty inherent in the fit. The systematic uncertainty due to the yield extraction has been estimated by varying the fit procedure. The fits to the invariant mass distribution were performed many times, varying the following parameters from their default values:

- The invariant mass bin width. In p–Pb collisions, 5 values are chosen of 2, 3, 4 (default), 5 and 6 GeV/c^2. In pp collisions, 5 values are chosen of 4, 5, 6 (default), 7 and 8 GeV/c^2.
- The lower limit of the invariant mass for the fit. In p–Pb collisions, 5 values are used between 2.18 and 2.22 GeV/c^2 in regular steps. In pp collisions since the default fit window is slightly wider 5 values between 2.17 and 2.23 GeV/c^2 are used.
- The upper limit of the invariant mass for the fit. Similarly to changing the lower limit, 5 values are used between 2.35 and 2.39 GeV/c^2. In pp collisions 5 values are used between 2.36 and 2.40 GeV/c^2.
- The background fit function. Up to 4 different fitting functions are used, including an exponential, linear, and second and third-order polynomial functions.
- The Gaussian width and mean of the signal peak. 4 variations were used here - fixing the Gaussian width to the expectation from MC and leaving the mean

free, leaving the mean and width free, fixing the Gaussian width and mean to the expectation from MC and fixing the Gaussian mean to the expectation value while leaving the width free.

This gives in total $5 \times 5 \times 5 \times 4 \times 4 = 2000$ trials. The trials were discarded and considered not good fits if the $\chi^2/\text{ndf} > 1.5$, if the width of the Gaussian peak $\sigma > 2.0 \times \sigma(MC)$ or $\sigma < 0.5 \times \sigma(MC)$ and if the significance < 3.0. In addition, since the linear and exponential fit are not good approximations of the background in the p_{T} intervals 2-4 and 4-6 GeV/c, the trials with these fit functions were skipped for these p_{T} intervals, giving a total of 1000 trials. An example of the raw yield as a function of the trial is shown in figure 4.19, where the trials corresponding to the Gaussian width and mean options, and the trials corresponding to the choice of background fit function are shown.

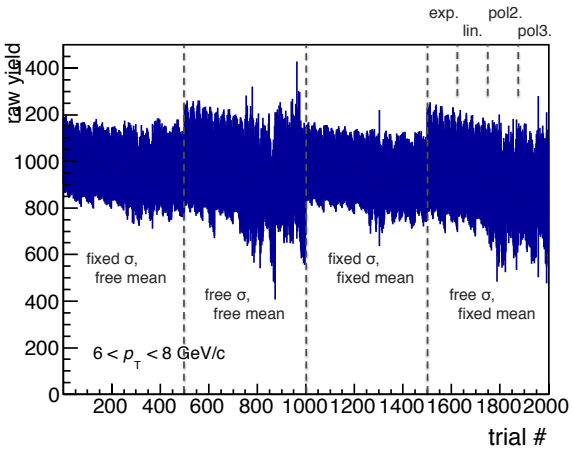

Figure 4.19: Raw yield as a function of the trial number for the p–Pb analysis in the range $6 < p_{\mathrm{T}} < 8$ GeV/c. The trials corresponding to the choice of Gaussian width and mean (fixing to MC expectation or leaving free), and the trials corresponding to the choice of background fit function (exponential, linear, 2nd order polynomial and 3rd order polynomial) are indicated on the figure.

For the p–Pb analysis, the raw yield as a function of the trial for each p_{T} bin is shown in figure 4.20, and the spread of the raw yield is shown in figure 4.21. The corresponding plots for the pp analysis are shown in figure 4.22 and figure 4.23. The spread in the raw yield distribution used for the assessment of the systematic uncertainty is the RMS, shown on the figures. The RMS in the case of a flat distribution ($\max - \min/\sqrt{12}$) is also shown in the figures. The RMS normalised by the mean is assigned as the systematic uncertainty where the distribution is symmetric and

gaussian. In the cases where the distribution is not symmetric, the max $-$ min$/\sqrt{12}$ is used to guide the choice of systematic uncertainty.

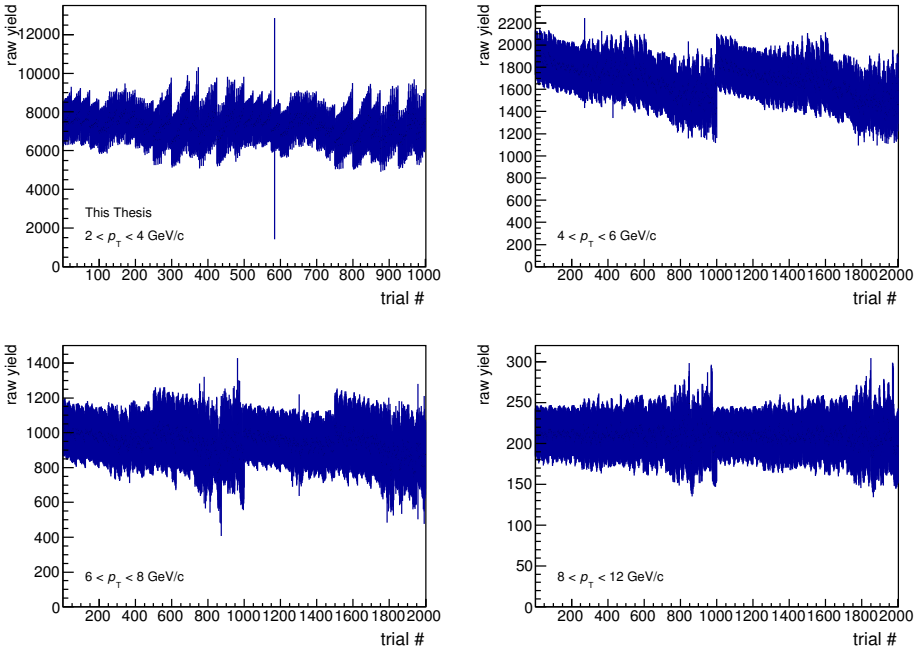

Figure 4.20: Raw yield as a function of the trial number for each p_{T} interval in the p–Pb analysis.

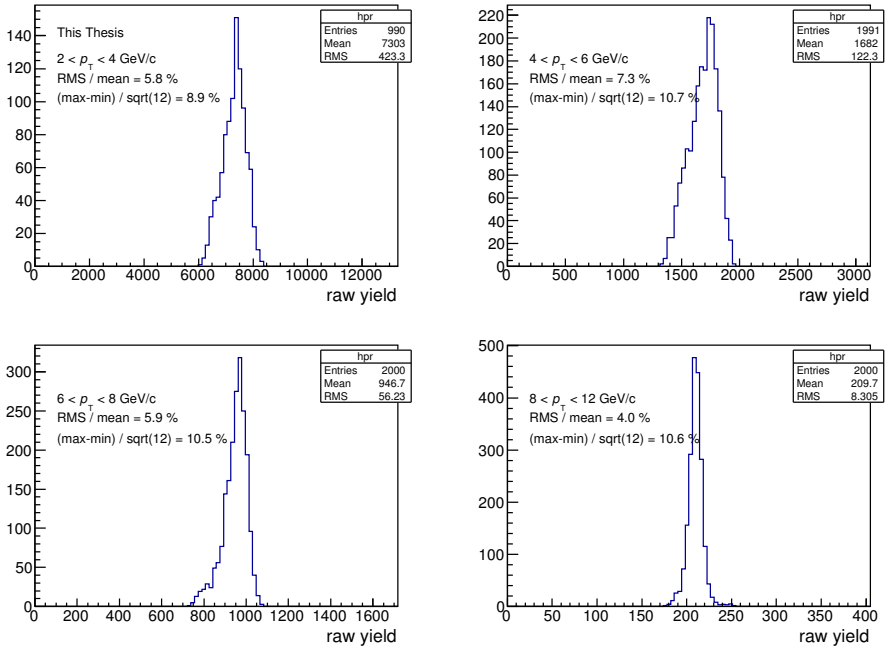

Figure 4.21: Projection of the raw yield spread for each p_{T} interval in the p–Pb analysis.

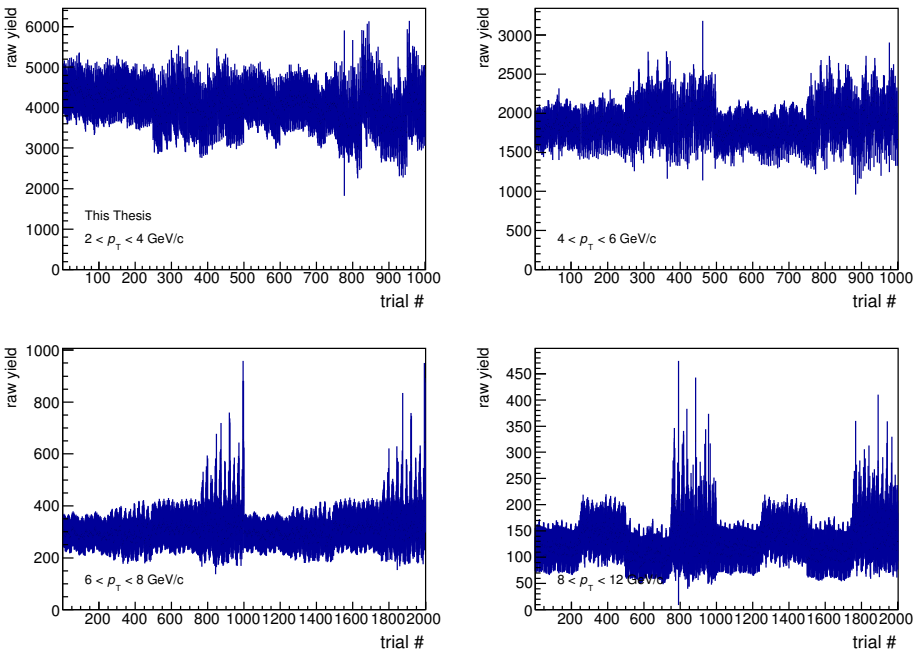

Figure 4.22: Raw yield as a function of the trial number for each p_{T} interval in the pp analysis.

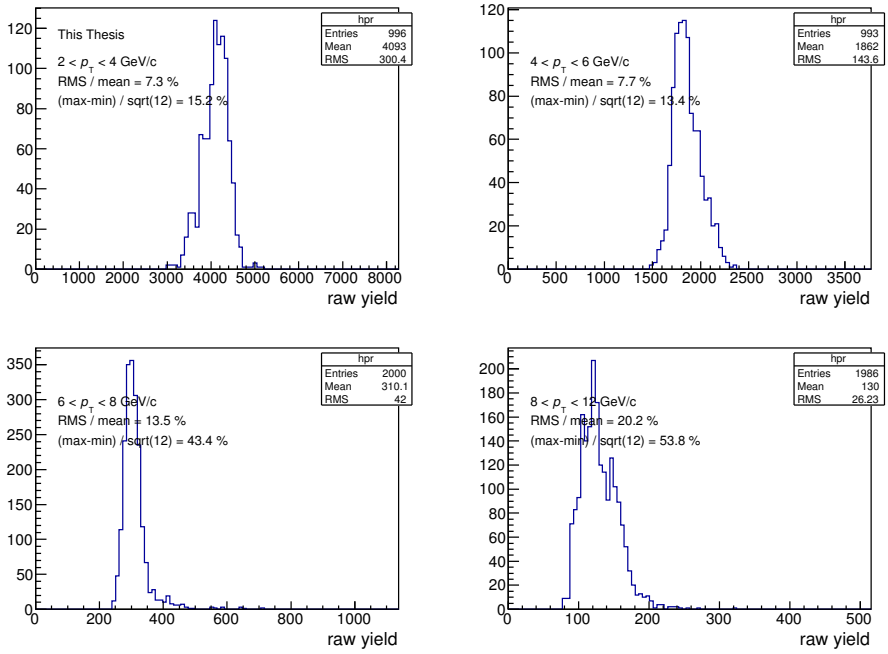

Figure 4.23: Projection of the raw yield spread for each p_{T} interval in the pp analysis.

4.7.2 BDT cut Efficiency

A systematic uncertainty can arise from discrepancies between data and simulation when using simulation to correct for some selection. In the case of this analysis, the BDT cut efficiency relies on the variables used as input to the BDT to be accurate - If a detector response is not modelled correctly in simulation, this may cause a change in one or more variable's distribution which can lead to a change in the BDT response, meaning the BDT cut efficiency will either be overestimated or underestimated. The effect of this discrepancy on the BDT cut efficiency has been evaluated by changing the cut on the BDT response and calculating the variation of the measured cross section in each p_{T} interval.

The analysis was repeated many times, with varying cuts on the BDT response, 'scanning' the BDT response in a pre-defined range. The gap between successive cuts on the BDT response was chosen to be 0.002, and the range of the BDT response cuts was taken as ± 0.05 from the central value in each p_{T} interval (giving a maximum total of 50 separate BDT cutting points per p_{T} interval). Extra quality cuts were made to decide whether to discard each cut, including a minimum BDT cut efficiency of 0.3, and a minimum signal extraction significance of 3. In p–Pb collisions, the maximum variation in this method causes a $^{+7}_{-16}\%$ difference in efficiency in the p_{T} interval $2 < p_{\mathrm{T}} < 4$ GeV/c, and a $^{+38}_{-32}\%$ difference in efficiency in the p_{T} interval $8 < p_{\mathrm{T}} < 12$ GeV/c. In pp collisions, the variation causes a $^{+15}_{-22}\%$ difference in the p_{T} interval $2 < p_{\mathrm{T}} < 4$ GeV/c, and a $^{+47}_{-27}\%$ difference in the p_{T} interval $8 < p_{\mathrm{T}} < 12$ GeV/c.

Figure 4.24 shows the cross section in each p_{T} interval as a function of the BDT response for the p–Pb analysis and 4.25 shows the spread of the cross section in the cut range considered. Figures 4.26 and 4.27 show the same for the pp analysis. The systematic uncertainty for each p_{T} interval is calculated using the same measures of the spread as for the raw yield extraction, the RMS and the $\max - \min/\sqrt{12}$ normalised by the mean of the distribution. In p–Pb collisions the two measures of the spread stay below 10% of the mean of the distribution. In pp collisions the two measures of the spread stay below 13%.

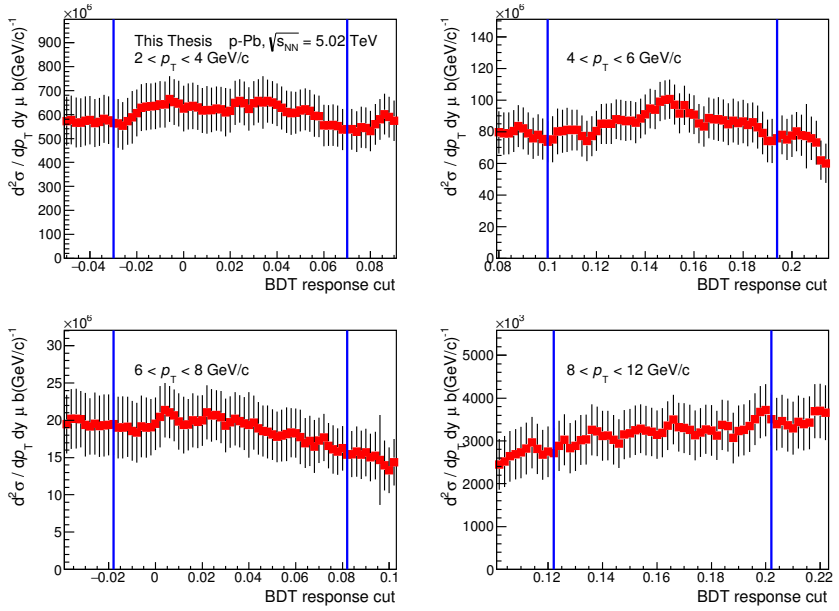

Figure 4.24: The cross section in each p_{T} interval as a function of the BDT response cut in the p–Pb analysis. The vertical lines indicate the minimum and maximum BDT cut used to evaluate the BDT cut systematic uncertainty.

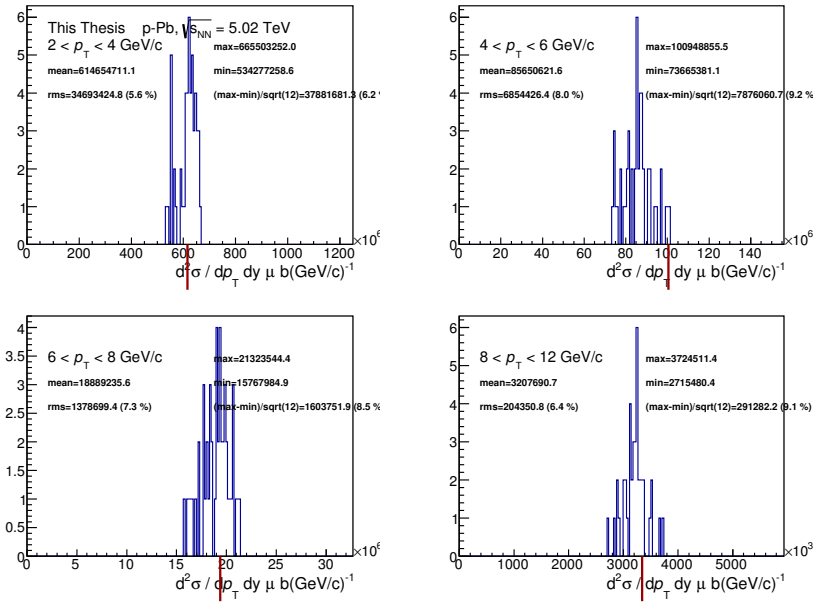

Figure 4.25: The spread of the cross section in each p_{T} interval for all BDT cuts in the considered BDT response range, in the p–Pb analysis.

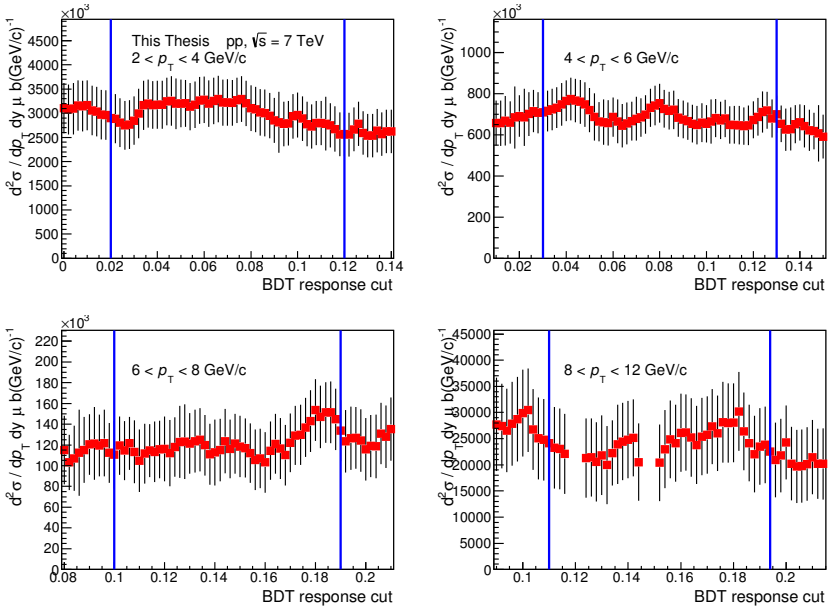

Figure 4.26: The cross section in each p_{T} interval as a function of the BDT response cut in the pp analysis. The vertical lines indicate the minimum and maximum BDT cut used to evaluate the BDT cut systematic uncertainty.

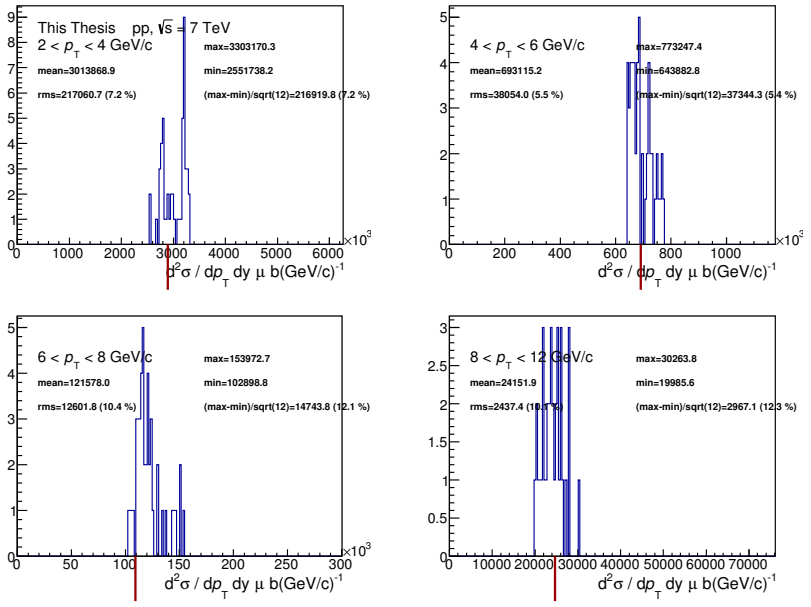

Figure 4.27: The spread of the cross section in each p_{T} interval for all BDT cuts in the considered BDT response range, in the pp analysis.

There are caveats in evaluating the uncertainty due to the mismatch between data and simulation in this way. There are uncertainties inherent in the method of extracting the raw yield through a fit to the invariant mass distributions, and these uncertainties may be the cause of the variation in the cross section. In addition the values of the raw yield extracted at each cut value are fully correlated meaning the same statistical fluctuations in the invariant mass distribution are present at each cut chosen. In order to further investigate the extent to which a data-simulation mismatch could affect the efficiency correction, the sensitivity of the input variable distributions on the final efficiency was studied. Here, parameters within the simulation are changed according to a best estimate of how they might be incorrectly modelled in simulation. This causes the distributions of the input variables (and thus the BDT response distribution) to change, and the resulting variation of the efficiency can give an indication of how sensitive the efficiency is to these 'incorrect' parameters of the simulation. These studies have only been performed in the p–Pb analysis, but for the purpose of this thesis it is safe to assume that the conclusions will be the similar for the pp analysis, since the variation in the final cross section with the BDT cut is of the same order.

4.7.2.1 PID variables

The PID variables, namely the probabilities for each track to be correctly assigned to a species of particle, play a large role in the multivariate selection, since they offer a large discrimination power with respect to the other input variables. The simulation is seen to do a good job at describing the data background (see figure 4.5), though any difference between the simulated variables and the true distribution in data could lead to a bias in the BDT cut efficiency. To estimate the systematic uncertainty associated with the inclusion of these PID variables in the multivariate selection, the ratio of the priors for kaons over the priors for pions used in the calculation of the Bayesian probability are modified from the default values. The priors are changed by a maximum of $\pm 30\%$, a number conservatively chosen based on the maximum mismatch between the default priors determined through an iterative procedure and the measured particle abundances [205].

Figure 4.28 shows the ratio of the BDT cut efficiencies for each prior modification with the efficiency using default priors, where the K/π priors are changed by $\pm 20\%$ and $\pm 30\%$. The efficiency changes by a maximum of $2-3\%$ at low p_{T}, reducing to $1-2\%$ at high p_{T}.

4.7.2.2 Topological variables

The input variables including the decay length, the decay length significance, σ_{vert}, DCA, $dist_{12}$ and the cosine of the pointing angle are all affected by the impact parameter resolution of the reconstructed tracks - namely the distance of closest approach of the track to the vertex from which they originated. It is known that the impact parameter resolution in MC simulations does not exactly match that observed in data [203]. To estimate the effect this has on the BDT cut efficiency, the impact parameter resolution in Monte Carlo is changed to match that in data. This is performed following the 'hybrid' method developed for studies related to the ALICE ITS upgrade [223, 224]. For this study the charged track impact parameter resolution in $r\phi$ in p–Pb simulations is changed to match that in data by rescaling the difference between the reconstructed track position and the 'true' MC track position, such that the position of each track y is scaled according to

$$y = y_{MC} + \sigma^{\text{data}}_{d_{0,r\phi}}(p_{\text{T}})/\sigma^{\text{MC}}_{d_{0,r\phi}}(p_{\text{T}})\dot{(}y - y_{MC}). \tag{4.8}$$

The variables associated to the reconstructed decay are then recalculated, and the Λ_{c} reconstruction and selection efficiency, including the BDT selection, is calculated.

Figure 4.29 (left) shows the ratio of the impact parameter resolution in MC with that in data, as a function of the track p_{T}. The description in simulation is worse at high p_{T}. Figure 4.29 (right) shows ratio of the efficiencies comparing the two impact parameter resolution cases. It can be seen that the change in efficiency is minimal ($< 1.5\%$ at low p_{T}, and less than a percent at intermediate/high p_{T}).

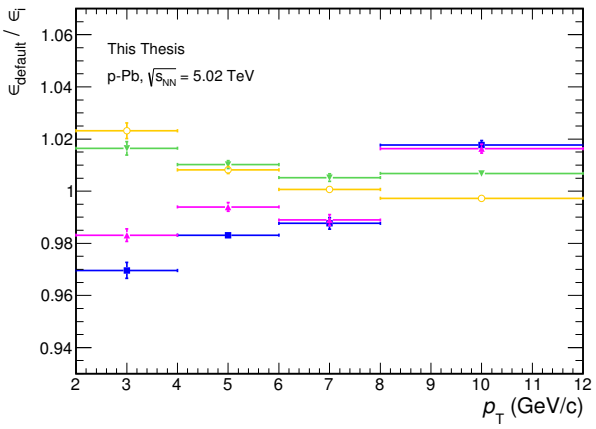

Figure 4.28: Ratio of the efficiencies for each modification of the K/π priors with respect to the default priors, in the p–Pb analysis.

The studies performed on the sensitivity of the BDT cut efficiency on the impact parameter resolution and choice or priors shows that the modification of these parameters has minimal effect on the BDT cut efficiency. These studies also justify the assigned BDT cut systematic uncertainty in p–Pb collisions. If the uncertainty assigned is seen approximately as the (quadratic) sum of the uncertainty coming from each input variable (kinematic, topological and PID), the assigned systematic of 8% at low p_{T} and 6% at high p_{T} (as summarised in section 4.7.9) can be seen as the 'conservative' sum of the topological ($\sim 1\%$) plus PID ($\sim 2-3\%$) plus an uncertainty on the the p_{T} of the daughter tracks, presumably of the same order.

4.7.3 Tracking efficiency

The systematic uncertainty associated with the tracking efficiency was estimated for the measurement of the p_{T}-differential cross section of D mesons measured via hadronic decay channels in pp and p–Pb collisions [144]. The uncertainty was estimated by varying track selection criteria. This includes variations to single-track cuts, including the number of TPC crossed rows, the number of TPC clusters associated with a track, the number of TPC clusters associated to a track which are used for PID, and the number of TPC crossed rows relative to the maximum number of TPC clusters that can be geometricaly assigned to the track. The resulting uncertainty was estimated to be 4% per track for pp collisions and 3% per track for p–Pb collisions, independent of track p_{T}. Since the track selection and thus the

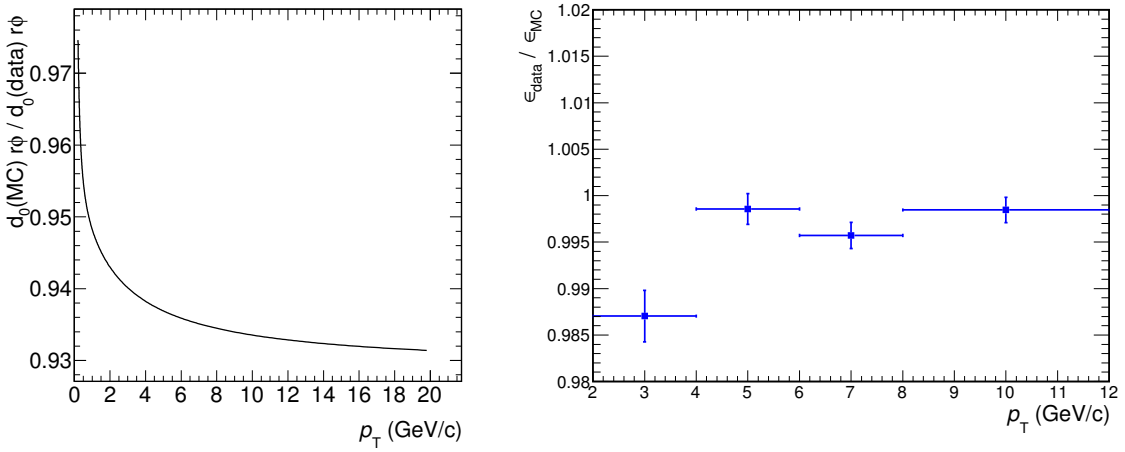

Figure 4.29: Left: The ratio of the impact parameter resolution in $r\phi$ of charged tracks from simulation with that in data, in p–Pb collisions. Right: The ratio of the total efficiency with the impact parameter resolution from simulation, to the total efficiency with the impact parameter resolution in data.

tracking efficiency is the same for this analysis, the same per-track uncertainty was assigned. The $\Lambda_c^+ \to pK^-\pi^+$ decay involves three tracks, so a p_T-independent systematic uncertainty of 12% is assigned in pp collisions and 9% is assigned in p–Pb collisions.

4.7.4 Monte Carlo p_T shape

Λ_c baryons are generated in the MC simulation with a p_T shape as given in PYTHIA. A discrepancy between the generated shape and the true Λ_c p_T shape will lead to a systematic shift of the calculated efficiencies.

To estimate the p_T shape systematic uncertainty, the efficiency is calculated applying weights determined from different p_T shape hypotheses - namely the p_T shape in FONLL [152], divided by the generated p_T shape. For the p–Pb analysis 3 generated p_T shapes are checked - the Λ_c p_T shape in simulated p–Pb collisions, the D meson p_T shape in simulated p–Pb collisions, and the D meson p_T shape in simulated pp collisions. For the pp analysis 2 generated p_T shapes are checked - the Λ_c p_T shape in simulated pp collisions, and the D meson p_T shape in simulated pp collisions. The uncertainty in the p_T shape can be assumed to fall within these three cases. A flat weight was also used as an extreme case, not to define an uncertainty but to check the sensitivity of the p_T shape on the efficiency.

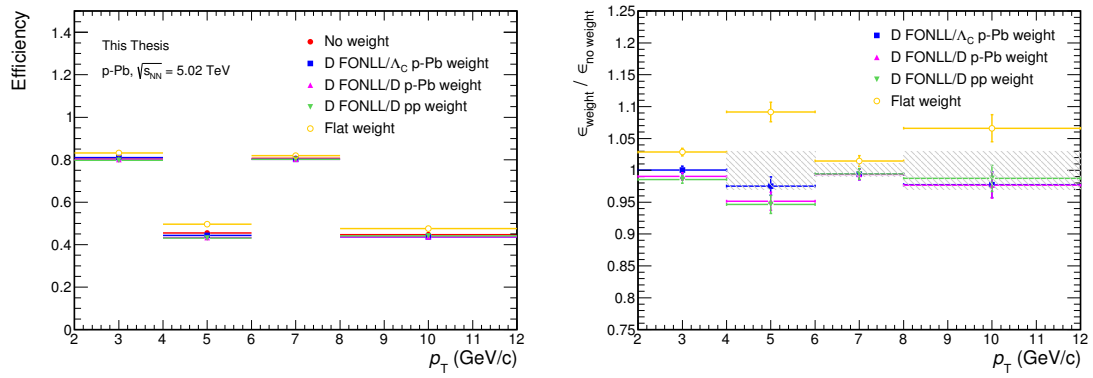

Figure 4.30: Generated p_T shape efficiency, considering different efficiency weights (left), and the ratio of the unweighted efficiency with the weighted efficiencies (right), for the p–Pb analysis.

The BDT cut efficiency for prompt Λ_c for each weighted option and the ratio with the unweighted efficiency in the p–Pb analysis is shown in figure 4.30. A systematic shift is seen in the BDT cut efficiencies, which ranges between $0 - 3\%$ when using the Λ_c p_T shape in PYTHIA. The same for the pp analysis is shown in figure 4.31,

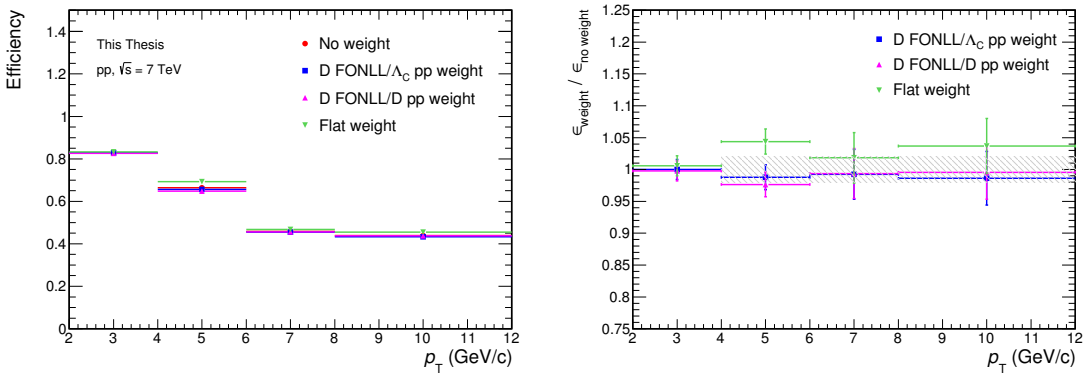

Figure 4.31: Generated p_{T} shape efficiency, considering different efficiency weights (left), and the ratio of the unweighted efficiency with the weighted efficiencies (right), for the pp analysis.

where a similar systematic shift is seen. The weighted pre-selection efficiency has been verified to be consistent within uncertainties with the unweighted efficiency. Taking into account that the prompt efficiency variation will be dominant in the variation of the total cross section, the systematic shift in the BDT cut efficiency using the 'D FONLL / Λ_{c} p–Pb(pp) weight' is used to define the p_{T} shape systematic uncertainty in p–Pb(pp) collisions.

4.7.5 Feed-down correction

The uncertainty on f_{prompt}, the fraction of the raw yield coming from beauty feed-down, is estimated from the uncertainties defined in the FONLL prediction of b production as well as changing the method used to obtain the $\Lambda_{\mathrm{b}}^{0} \rightarrow \Lambda_{\mathrm{c}} + \mathrm{X}$ cross section, taking into account existing measurements of Λ_{b}^{0} production at the LHC. The uncertainty in the $\Lambda_{\mathrm{b}}^{0} \rightarrow \Lambda_{\mathrm{c}} + \mathrm{X}$ cross section is determined by varying the parameters in FONLL including the b-quark mass, the PDF parametrisation, and the renormalisation and factorisation scales within their uncertainties, and taking into account the uncertainty in the fragmentation fraction $f(b \rightarrow \Lambda_{\mathrm{b}}^{0})$ and the uncertainty in the branching ratio $BR(\Lambda_{\mathrm{b}}^{0} \rightarrow \Lambda_{\mathrm{c}} + \mathrm{X})$. Two additional methods to estimate f_{prompt} were also considered:

(i) In the f_c method, the relative fraction f_{prompt} is entirely based on calculated cross sections for prompt and non-prompt production of Λ_{c} within the FONLL

framework:

$$f_{\mathrm{prompt}} = 1 \Bigg/ \left(1 + \frac{\epsilon_{\Lambda_{\mathrm{b}}^0 \to \Lambda_{\mathrm{c}}} \cdot \mathrm{BR}_b}{\epsilon_{\Lambda_{\mathrm{c}}} \cdot \mathrm{BR}_{\mathrm{c}}} \cdot \frac{\int_{\Delta y \Delta p_{\mathrm{T}}} \frac{\mathrm{d}\sigma^{\Lambda_{\mathrm{b}}^0}_{\mathrm{FONLL}}}{\mathrm{dydp_T}}}{\int_{\Delta y \Delta p_{\mathrm{T}}} \frac{\mathrm{d}\sigma^{\Lambda_{\mathrm{c}}}_{\mathrm{FONLL}}}{\mathrm{dydp_T}}} \right) \tag{4.9}$$

This method has been used for the measurement of the D meson cross sections within ALICE (see e.g. [143]) to evaluate the uncertainty, given that beauty production may not be described well by FONLL at the LHC.

(ii) Since Λ_{b}^0 production has been measured by LHC experiments at 7 TeV, it has also been considered that, assuming the Λ_{b}^0 prediction used to determine the feed-down fraction is consistent with measurements, one could also purely rely on the N_b method. The most recent measurement of D mesons at ALICE [145] followed a similar line of reasoning, since the production of B mesons is well described by FONLL at the LHC. The Λ_{b}^0 cross section has been measured by LHCb [225] to be up to a factor 2 above the FONLL prediction (taking into account the fragmentation fraction ($b \to \Lambda_{\mathrm{b}}^0 = 0.21$) measured by LHCb at forward rapidity [111]) in the same p_{T} region as this analysis, and has been measured by CMS [226] to be around a factor 1.6 above the FONLL prediction in the highest p_{T} region of this analysis ($10 < p_{\mathrm{T}} < 13$ GeV/c). The Λ_{b}^0 production cross section from FONLL has therefore been rescaled by a factor of 2 to include the deviation seen in measurements, and the feed-down fraction is determined following the N_b method. This method is named '$N_b \times 2$'.

The feed-down fraction determined from the N_b, f_c and $N_b \times 2$ methods is shown in figure 4.32. Since the theoretical uncertainties are larger than the uncertainties from the measured raw yield, the more data-driven method N_b (the method not relying on the theoretical charm prediction) has smaller uncertainties, and is chosen as the central f_{prompt}, and the method $N_b \times 2$ considering the deviation of the measured Λ_{b}^0 cross section from the theoretical expectation is also used to evaluate the uncertainty. The uncertainty on f_{prompt} is chosen as the envelope of the methods N_b and $N_b \times 2$. The f_c method is considered a check and is not used to evaluate the uncertainty.

4.7.6 Pre-selection PID Efficiency

As explained in section 4.3 loose pre-selection PID is applied prior to the BDT selection. This pre-selection PID efficiency is high (90%) relative to the BDT selection, so

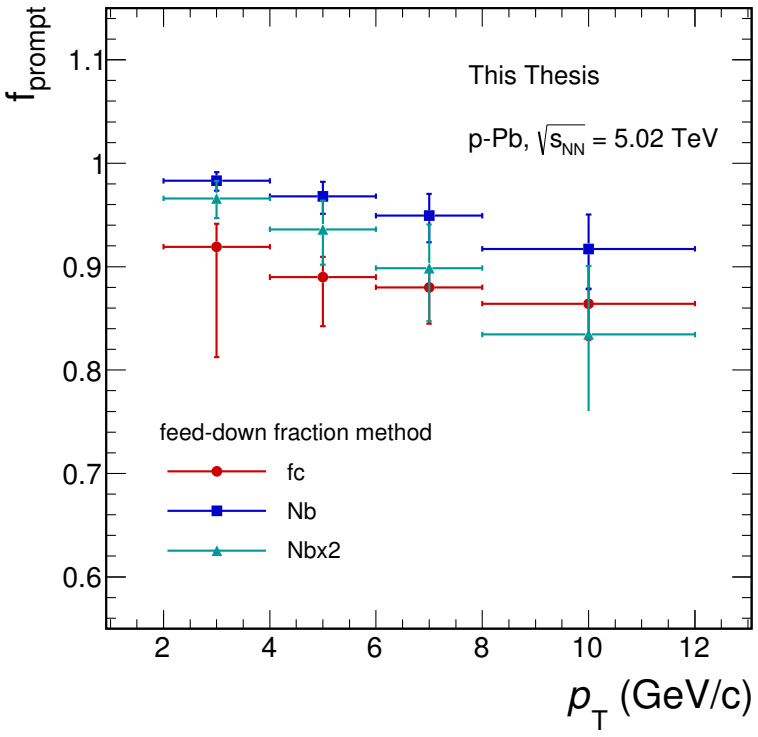

Figure 4.32: The fraction of prompt Λ_c in the p–Pb analysis, showing the 3 methods N_b, f_c and $N_b \times 2$ as described in the text.

the uncertainty is expected to be small. The systematic uncertainty due to this selection is estimated by performing the analysis using a different PID strategy, namely a 3σ cut in just the TPC. This selection was found to have a negligible effect on the final cross section, so no uncertainty is assigned due to the pre-selection PID.

4.7.7 Positive/Negative Λ_c

To verify that the measurement of the Λ_c^+ cross section through halving the measurement of both Λ_c^+ and Λ_c^- does not introduce a bias in the cross section due to unequal numbers of Λ_c^+ and Λ_c^- being produced, and to validate the positive/negative efficiencies, the signal is extracted for both Λ_c^+ and Λ_c^- and the cross section is evaluated for both cases. Since dividing the data sample into positive and negative contributions further reduces the statistics the signal is extracted merging the analysis p_T intervals into two: $2 < p_T < 6$ GeV/c and $6 < p_T < 12$ GeV/c. The separate Λ_c^+ and Λ_c^- cross sections are compared with the charge-integrated Λ_c cross section in the same p_T intervals. For the BDT stage, a BDT is trained per p_T bin and per charge case (Λ_c^+, Λ_c^- or charge-integrated $\Lambda_c^\pm$), and applied on the corresponding data sample. The BDT

optimal cutting point is determined for the charge-integrated case as in section 4.5, and the same cutting point is used for the charge-separated cases. The corresponding fits to the invariant mass distributions are shown in figure 4.34. The corrected yield for each case is then determined. Figure 4.35 shows the ratio of $2\times$ the corrected yield for the charge-separated cases with the charge-integrated case. The ratio is consistent with unity within uncertainties. Since no asymmetry in the production of Λ_c^+ and Λ_c^- is expected, at least within the experimental accuracy achievable here, this is considered a cross check and no systematic uncertainty is assigned.

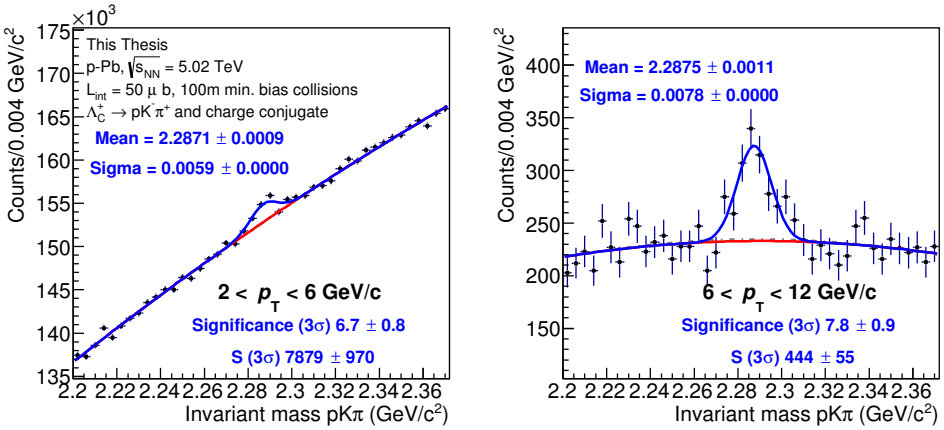

Figure 4.33: The invariant mass spectra for Λ_c candidates in the 2 p_T intervals $2 < p_T < 6$ GeV/c and $6 < p_T < 12$ GeV/c, and the fit to the signal and background, in p-Pb collisions.

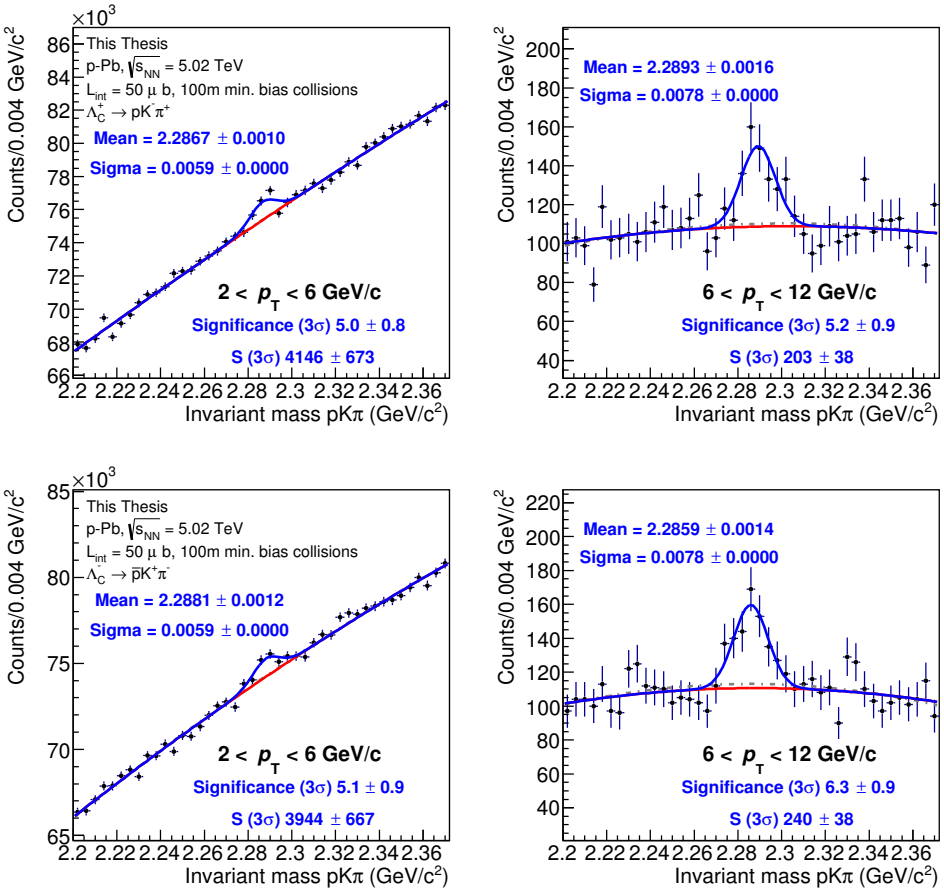

Figure 4.34: The invariant mass spectra for Λ_c candidates in the 2 p_T intervals $2 < p_T < 6$ GeV/c and $6 < p_T < 12$ GeV/c, and the fit to the signal and background, in p-Pb collisions, separating the $\Lambda_c^+ \to pK^-\pi^+$ candidates (top panels) and $\Lambda_c^- \to \bar{p}K^+\pi^-$ candidates (bottom panels).

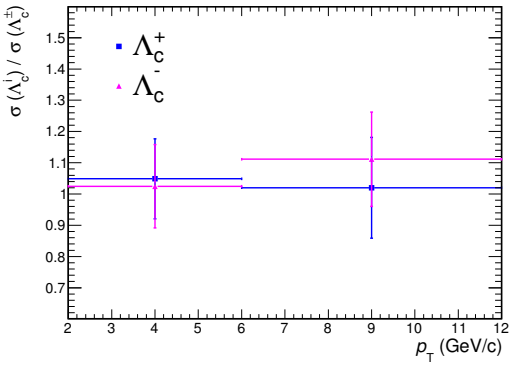

Figure 4.35: The ratio between the corrected yield of Λ_c^+ and Λ_c^-, over the charge integrated $\Lambda_c^\pm$ corrected yield

4.7.8 Normalisation

Uncertainties on the branching ratio and the integrated luminosity $\mathcal{L}_{\rm int}$ contribute to the overall Λ_c cross section uncertainty, which originate from the uncertainty on the measured $\Lambda_c^+ \rightarrow {\rm pK}^-\pi^+$ branching ratio, and the interaction cross section of pp or p–Pb collisions for the trigger condition of the particular analysed dataset. These uncertainty are considered as global uncertainties and apply to all $p_{\rm T}$ intervals.

The branching fraction of $\Lambda_c^+ \rightarrow {\rm pK}^-\pi^+$ is $(6.35 \pm 0.33)\%$ [13] - a fractional uncertainty of 5.0%.

The pp interaction cross section at $\sqrt{s} = 7$ TeV for the trigger condition in this analysis (a hit in either side of the V0 detector) is 62.2 ± 2.2 mb [227], giving a fractional uncertainty of 3.5% [227].

The p–Pb interaction cross section at $\sqrt{s_{\rm NN}} = 5.02$ TeV for the trigger condition in this analysis (a hit in both sides of the V0 detector V0A and V0C) is 2.09 ± 0.07 b [228], a fractional uncertainty of 3.5%. An additional uncorrelated uncertainty of 1% applies to the luminosity measurement in p–Pb collisions resulting from the differences in the luminosity obtained using the V0 and T0 detector, bringing the total p–Pb interaction cross section uncertainty in p–Pb collisions to 3.7%.

4.7.9 Uncertainties summary

This section has described the estimation of the systematic uncertainty sources of the $p_{\rm T}$-differential cross section of the Λ_c baryon in each measured $p_{\rm T}$ interval. A summary of the systematic uncertainty sources is shown in table 4.5. Figure 4.36 also shows a summary of the systematic uncertainties in the pp and p–Pb analyses. In the

	pp analysis, $\sqrt{s} = 7$ TeV			
p_{T} interval (GeV/c)	2-4	4-6	6-8	8-12
Yield extraction	10%	10%	15%	25%
Cut efficiency	8%	8%	10%	10%
Pre-selection PID	negligible			
$\Lambda_{\mathrm{c}}^{+}, \Lambda_{\mathrm{c}}^{-}$	negligible			
Tracking	12%	12%	12%	12%
MC p_{T} shape	0%	2%	2%	2%
Feed-down	$^{+0.5}_{-2}$%	$^{+0.6}_{-3}$%	$^{+1}_{-7}$%	$^{+2}_{-10}$%
Branching fraction	5.0%			
Luminosity	3.5%			
	p-Pb analysis, $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV			
Yield extraction	6%	8%	8%	4%
Cut efficiency	8%	8%	8%	6%
PID variables	~2-4%			
topological variables	~1%			
Pre-selection PID	negligible			
$\Lambda_{\mathrm{c}}^{+}, \Lambda_{\mathrm{c}}^{-}$	negligible			
Tracking	9%	9%	9%	9%
MC p_{T} shape	0%	3%	1%	3%
Feed-down	$^{+0.8}_{-4}$%	$^{+1}_{-7}$%	$^{+2}_{-11}$%	$^{+4}_{-18}$%
Branching fraction	5.0%			
Luminosity	3.7%			

Table 4.5: Summary of the total systematic uncertainties assigned for each systematic uncertainty source, expressed as a percentage of the cross section in each p_{T} interval. The uncertainty on the branching fraction and luminosity are global uncertainties assigned to all p_{T} bins.

pp analysis, the tracking and yield extraction systematic uncertainties are dominant at low p_{T}, and the yield extraction systematic uncertainty is dominant at high p_{T} due to the small statistical significance of the signal extraction. In the p–Pb analysis, the tracking, yield extraction and BDT cut efficiency systematic uncertainties are the largest uncertainty sources and are of roughly the same order. While at low p_{T} the relative statistical and systematic uncertainties are roughly the same in both the pp and p–Pb analysis, at high p_{T} the uncertainty in the pp analysis is larger due to the smaller statistical significance of the signal extraction, and the corresponding larger statistical uncertainty and yield extraction uncertainty.

The relative statistical uncertainties on the p_{T}-differential cross section measured

in this work are shown in figure 4.36 for the pp and p–Pb analyses, in comparison to the statistical uncertainty from the analyses performed with a standard, rectangular cut-based method [229, 230]. In general, an improvement in the statistical uncertainty is obtained in this work from the use of BDTs. In the pp analysis, a smaller statistical uncertainty is obtained in all p_{T} intervals (though the p_{T} range $2 < p_{\mathrm{T}} < 4$ GeV/c in the BDT analysis is merged to 1 bin rather than 2 in the standard analysis). In addition, the $6 < p_{\mathrm{T}} < 12$ GeV/c range becomes accessible in the BDT analysis, where the signal can be extracted in two additional p_{T} bins. In the p–Pb analysis, a smaller statistical uncertainty is obtained from the BDT analysis in all p_{T} intervals except in the range $8 < p_{\mathrm{T}} < 12$ GeV/c where the uncertainty in the standard analysis is slightly smaller.

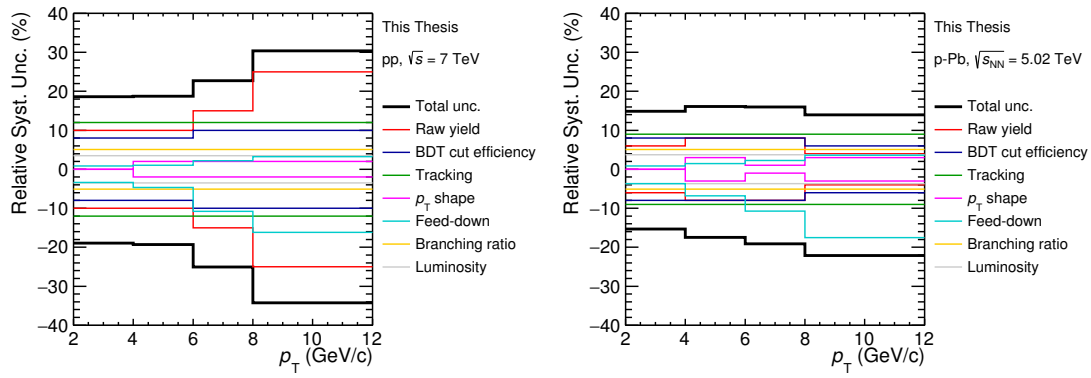

Figure 4.36: Summary of the systematic uncertainty sources contributing to the overall p_{T}-differential cross section systematic uncertainty for the pp analysis (left) and the p–Pb analysis (right).

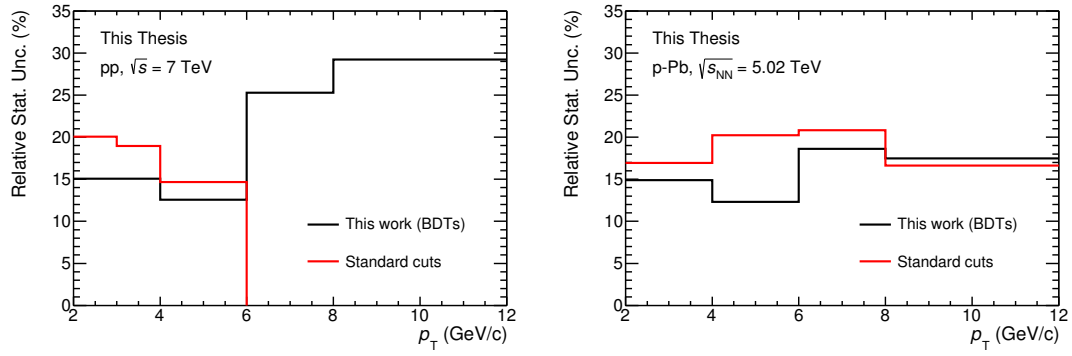

Figure 4.37: The statistical uncertainty of the p_{T}-differential cross section from this work, compared to the statistical uncertainty of the analysis performed using a standard, rectangular cut-based method, for the pp analysis (left) and the p–Pb analysis (right).

correlation with pp(p–Pb) Λ_c			
	p–Pb(pp) Λ_c	D^0	p_T
Yield extraction	U	U	U
Cut efficiency	U	U	C
Tracking	U	C	C
MC p_T shape	U	U	C
Feed-down	C	C	C
Branching fraction	C	U	C
Luminosity	U	C	C

Table 4.6: Summary of the correlation between the pp and p–Pb Λ_c cross section systematic uncertainty sources and different quantities. 'U' determines the uncertainty source to be uncorrelated, and 'C' determines the uncertainty source to be correlated. The 'p–Pb(pp) Λ_c' column states the correlation with the Λ_c analysis in the different collision system (relevant for the R_{pPb} measurement), the 'D^0' column states the correlation with the D^0 analysis in the same collision system (relevant for the Λ_c/D^0 ratio), and the 'p_T' column states the correlation among the different p_T intervals (relevant for the integrated p_T cross section).

In chapter 5 the p_T-differential cross section measurement of the Λ_c baryon in the two collision systems are compared to other measured quantities. The ratio of the Λ_c cross section in p–Pb collisions relative to pp collisions quantifies the nuclear modification in p–Pb collisions, and the baryon-to-meson ratio Λ_c/D^0 gives insight into the relative production of baryons and mesons. The Λ_c cross section is also integrated over p_T to obtain a more inclusive measurement less dependent on p_T. The systematic uncertainty sources between the Λ_c baryon in different collision systems, the D^0 meson, and the correlation between each p_T interval are either assumed to be fully correlated, or fully uncorrelated, and the correlation is summarised in table 4.6. How the uncertainties are propagated depending on the correlation is described in appendix A.1.

4.8 pp reference scaling for the R_{pPb} measurement

The nuclear modification factor in p–Pb collisions, R_{pPb}, describes the modification of a given observable (in this case the Λ_c yield) in p–Pb collisions, relative to pp collisions, scaled by the number of binary nucleon-nucleon collisions. In other words, it measures how a given observable differs from the same number of *independent* nucleon-nucleon collisions - any change can be put down to either initial- or final-state effects from the presence of the nucleus or presence of multiple interactions in

the final state. In minimum bias collisions $R_{\rm pPb}$ can be expressed as

$$R_{\rm pPb} = \frac{1}{A}\frac{{\rm d}\sigma^{{\rm prompt}\Lambda_{\rm c}}_{\rm pPb}/{\rm d}p_{\rm T}}{{\rm d}\sigma^{{\rm prompt}\Lambda_{\rm c}}_{\rm pp}/{\rm d}p_{\rm T}} \tag{4.10}$$

where ${\rm d}\sigma^{{\rm prompt}\Lambda_{\rm c}}_{\rm pPb(pp)}/{\rm d}p_{\rm T}$ is the $p_{\rm T}$-differential production cross section of prompt $\Lambda_{\rm c}$ in p–Pb(pp) collisions, and $A = 208$ is the mass number of lead. The measurements of prompt $\Lambda_{\rm c}$ production presented in this work are at different collision energies per nucleon-nucleon collision (5.02 TeV and 7 TeV in pp and p–Pb collisions respectively). For this reason the pp results should be scaled to the same energy as the p–Pb results to compute the $R_{\rm pPb}$. This procedure was performed in previous D meson measurements [144, 231] as described in detail in [232] using the ratio of FONLL perturbative QCD calculations at $\sqrt{s} = 5.02$ TeV and $\sqrt{s_{\rm NN}} = 7$ TeV [152] to scale the pp reference, such that

$$R_{\rm pPb} = \frac{1}{A}\frac{{\rm d}\sigma^{\rm 5TeV}_{\rm pPb}/{\rm d}p_{\rm T}}{\frac{{\rm d}\sigma^{\rm 5TeV}_{\rm FONLL}/{\rm d}p_{\rm T}}{{\rm d}\sigma^{\rm 7TeV}_{\rm FONLL}/{\rm d}p_{\rm T}}{\rm d}\sigma^{\rm 7TeV}_{\rm pp}/{\rm d}p_{\rm T}}. \tag{4.11}$$

The uncertainty on the scaling is then calculated by evaluating the scaling ratio and changing the parameters of the calculation, then taking the envelope of all variations as the uncertainties. The parameters within FONLL that are changed include:

- The bare c-quark mass. The central mass is set to $m_c = 1.5$ GeV/c^2. The uncertainty on the mass is taken to be ± 0.2 GeV/c^2 giving the minimum and maximum masses as 1.3 and 1.7 GeV/c^2, respectively.
- The proton PDF. The PDF set in the FONLL calculation is the CTEQ6.6 PDFs [104]. The minimum and maximum uncertainty bounds of the PDF set form the uncertainty for the PDF in FONLL.
- The factorisation scales μ_F and μ_R. The central values are taken as $\mu_R = \mu_F = \mu_0 = \sqrt{m^2 + p_{\rm T}^2}$. The uncertainty on the scales are evaluated such that $\mu_0/2 < \mu_R, \mu_F < 2\mu_0$ with $1/2 < \mu_R/\mu_F < 2$.

The ratio of the bare c-quark cross sections at 5 TeV and 7 TeV, and the uncertainty is shown in figure 4.38. The uncertainty is asymmetric, and is largest at low $p_{\rm T}$.

The c-quark 'fragments' into a $\Lambda_{\rm c}$ baryon which carries a fraction of the initial c-quark momentum z. The probability of the $\Lambda_{\rm c}$ fragmenting with a given z is described by the fragmentation function (see section 2.1.1). A fragmentation function is said

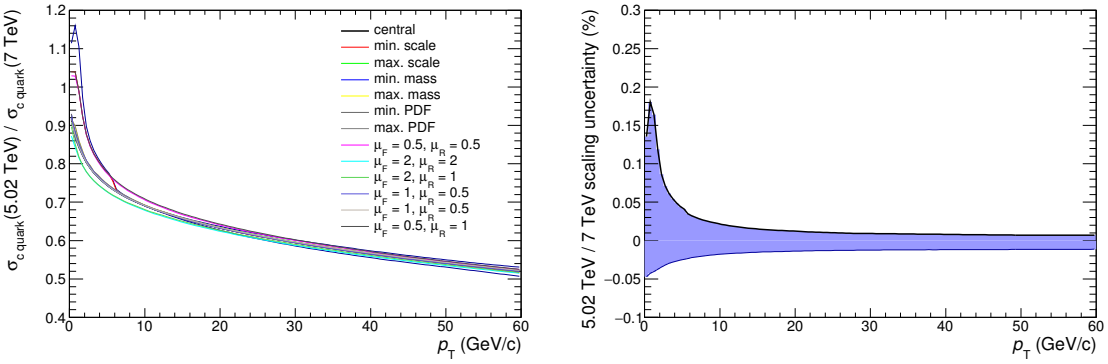

Figure 4.38: Left: The p_{T}-differential bare c-quark cross section ratio of 5 TeV and 7 TeV in FONLL, considering the changes of the quark mass, the PDF and the scales within their uncertainties as defined in FONLL. Right: The upper and lower uncertainty on this ratio.

to be 'hard' when the hadron keeps a large fraction of the momentum of the initial c-quark, and 'soft' when the hadron loses momentum relative to the initial c-quark. However the Λ_{c} fragmentation function is not well known/measured, so production cross sections are not readily available in FONLL. To calculate a scaling factor for the Λ_{c} baryon, different fragmentation functions were tested to check the sensitivity of the scaling factor on the choice of fragmentation function.

In FONLL, the BFCY [233] parametrisation describes the fragmentation of a c quark to pseudoscalar ($\mathrm{D}^{+}, \mathrm{D}^{0}$) or vector ($\mathrm{D}^{*+}$) mesons. This fragmentation function is tested, and the scaling factor changes by less than 1% between pseudoscalar and vector mesons. The Peterson parametrisation [108] of the fragmentation function is also tested (see equation 2.4), which has a single free parameter ϵ_P which is expected to vary roughly as the ratio of the light (di)quark mass to the heavy quark mass in the hadron. In this case a small value of ϵ_P describes a 'harder' fragmentation function, while a large value describes a 'softer' fragmentation function. For the Peterson parametrisation the ratio of the light quark and heavy quark masses gives $\epsilon_P = 0.005$, giving a fragmentation function very close to that from the BFCY parametrisation. This is also close to the value used for Λ_{c} in GM-VFNS [234]. A value of $\epsilon_P = 0.05$ is the default value used for charm fragmentation in PYTHIA when the Peterson fragmentation is enabled. As an additional check, the $\Lambda_{\mathrm{c}}/\mathrm{D}^{0}$ ratio was also calculated given the D^{0} shape from FONLL and assuming different fragmentation functions from $\epsilon_P = 0$ (bare c-quark cross section) to $\epsilon_P = 0.1$, and the ratio with all these variations was seen to be consistent with the measured $\Lambda_{\mathrm{c}}/\mathrm{D}^{0}$ ratio in pp collisions.

Based on these trials, the central point and upper/lower bounds of the uncertainty for the scaling factor are defined as follows:

- The central value is the ratio of the FONLL prediction for D^+.
- The upper uncertainty is the ratio of the FONLL prediction for the bare c-quark ('hardest' case).
- The lower uncertainty is the ratio of the bare c-quark cross section + the fragmentation function with Peterson $\epsilon = 0.1$ ('softest' case).

The uncertainty on the scaling factor is the envelope of these points plus the associated uncertainties from FONLL. Figure 4.39 shows the scaling factor for these three cases, with their associated uncertainty. The uncertainty relative to the central (D^+) point goes from $+12, -5\%$ in the p_T bin 2-4 GeV/c, to $+5, -3\%$ in the p_T bin 8-12 GeV/c.

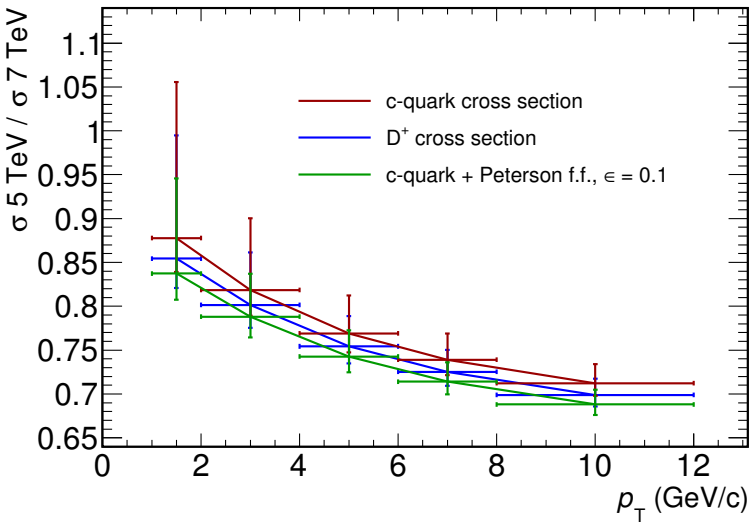

Figure 4.39: The scaling factor for the pp reference in the $R_{\rm pPb}$ measurement as a function of p_T, for the central, lower and upper uncertainty cases described in the text. The vertical bars describe the envelope of the changes in the quark mass, the PDF and the scales in FONLL [235].

Chapter 5

Results

This chapter presents the results of the data analysis detailed in chapter 4. The production cross section of the Λ_c^+ baryon has been measured as a function of p_T in the range $2 < p_T < 12$ GeV/c in two collision systems: in pp collisions at $\sqrt{s} = 7$ TeV, and in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. These results are shown in section 5.1, along with comparisons to theoretical calculations for the pp results. A comparison between results obtained with 2 data analysis methods to select Λ_c candidates, a traditional cut-based approach and the multivariate approach presented in this thesis, is also shown in this section. The baryon-to-meson ratio, Λ_c^+/D^0, has been measured and compared to predictions from MC generators, and is shown in section 5.2. The Λ_c^+ nuclear modification factor R_{pPb} has been measured and is shown in section 5.3. The results are then discussed in section 5.4. Finally, the $\Lambda_c^+ \to pK^-\pi^+$ analyses (including the p–Pb BDT analysis and the pp and p–Pb analyses using rectangular cuts) are merged with independent analyses of the $\Lambda_c^+ \to pK_s^0$ and $\Lambda_c^+ \to \Lambda e^+\nu_e$ decay channels, as explained in section 5.5. As a reminder, the p_T-differential cross section of the Λ_c^+ baryon is measured rather than the particle-plus-antiparticle ($\Lambda_c^+ + \Lambda_c^-$), and the results presented in this chapter refer to the positively charged baryon, Λ_c^+ only. Λ_c still refers to both Λ_c^+ and Λ_c^-.

5.1 Λ_c^+ production cross section in pp and p–Pb collisions

The p_T-differential production cross section of the Λ_c^+ baryon has been measured via its hadronic decay channel $\Lambda_c^+ \to pK^-\pi^+$. It has been measured as a function of transverse momentum in pp collisions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV in the centre-of-mass rapidity range $|y| < 0.5$, and in p–Pb collisions at a centre-of-

mass per nucleon-nucleon collision of $\sqrt{s_{\rm NN}} = 5.02$ TeV in the centre-of-mass rapidity range $-0.96 < y < 0.04$. Figure 5.1 shows the $\Lambda_{\rm c}^+$ baryon $p_{\rm T}$-differential production cross section measured in pp collisions, and figure 5.2 shows the $\Lambda_{\rm c}^+$ $p_{\rm T}$-differential production cross section measured in p–Pb collisions. Both production cross section measurements have been performed in the $p_{\rm T}$ range $2 < p_{\rm T} < 12$ GeV/c.

The production cross section measurements presented in this work rely on selection of the $\Lambda_{\rm c}$ candidates using Boosted Decision Trees, to remove background candidates where the measured proton, kaon and pion candidates come from uncorrelated sources. The measurement has also been performed using a traditional cut-based approach in both pp and p–Pb collisions [229, 230], where 'rectangular' selections are made on variables which are similar to those used as 'input' to the Boosted Decision Tree in this work[1]. A comparison with this approach is shown in figure 5.3, for pp and p–Pb collision. The ratio of the two methods is also shown in the bottom panels of this figure, where only the statistical uncertainties are shown and the uncertainties are considered uncorrelated. Where the $p_{\rm T}$ binning is different, the cross section for the merged bin and the corresponding statistical uncertainties, which are uncorrelated with $p_{\rm T}$, has been calculated. The two methods are seen to be in good agreement, with the pp cross sections consistent within $\sim 15\%$ and the p–Pb cross sections consistent within $\sim 30\%$. The degree of correlation between the two methods is difficult to assess - both the statistical and systematic uncertainties have some degree of correlation - but it can be qualitatively estimated that the measurements are consistent within statistical and systematic uncertainties to within 2σ.

A perturbative QCD calculation for the $\Lambda_{\rm c}^+$ cross section with GM-VFNS [87, 88] at NLO (which provides a good description of the D-meson $p_{\rm T}$-differential cross section at the same energy [145]) has been performed and is compared to the measured $\Lambda_{\rm c}^+$ baryon $p_{\rm T}$-differential cross section in figure 5.4. The calculation has been performed with the CTEQ 6.6 [104] parametrisation of the parton densities, using the 'global' fragmentation functions from [236], and assuming $m_c = 1.5$ GeV/c^2. The renormalisation and factorisation scales μ_R and μ_R correspond to $\mu_i = x_i m_{\rm T}$ as justified in section 2.1.1.1 where $m_{\rm T}$ is the transverse mass of the charm quark and $x_i = 1$ for the central prediction, and x_i is varied such that $x_i \in 0.5, 1, 2$ to determine the uncertainty in the calculation. With this scale choice the prediction is only valid above $p_{\rm T} \gtrsim 3$ GeV/c, and due to the binning of the measured pp cross section the

[1]The traditional cut-based analysis makes selections on all variables used in this analysis, except the normalised decay length in the XY plane.

prediction is shown for $p_{\mathrm{T}} > 4$ GeV/c. GM-VFNS significantly underpredicts the measured Λ_{c}^{+} cross section, by a factor of around 3 in the $4 < p_{\mathrm{T}} < 6$ GeV/c interval, and by a factor of around 2 in the $6 < p_{\mathrm{T}} < 12$ GeV/c interval.

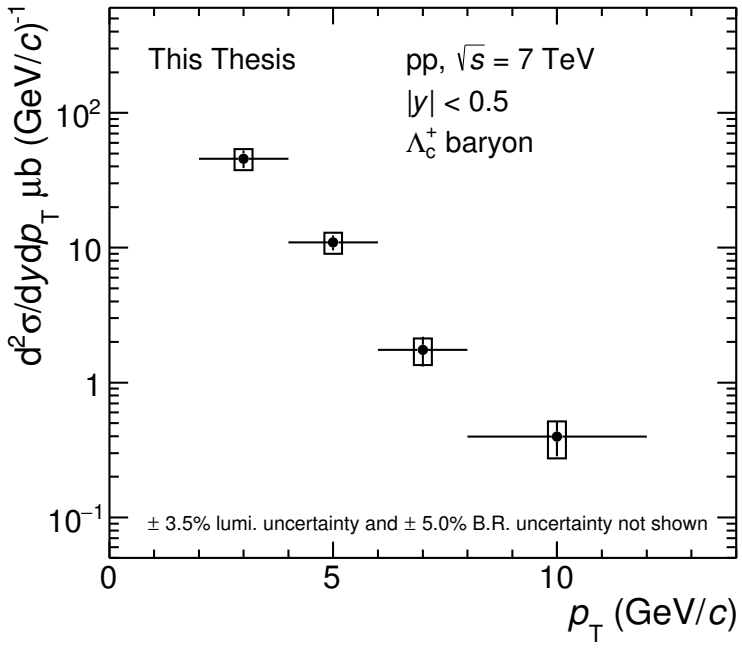

Figure 5.1: p_{T}-differential cross section of the Λ_{c}^{+} baryon measured in the $\Lambda_{\mathrm{c}}^{+} \to \mathrm{pK}^{-}\pi^{+}$ channel in pp collisions at $\sqrt{s} = 7$ TeV. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

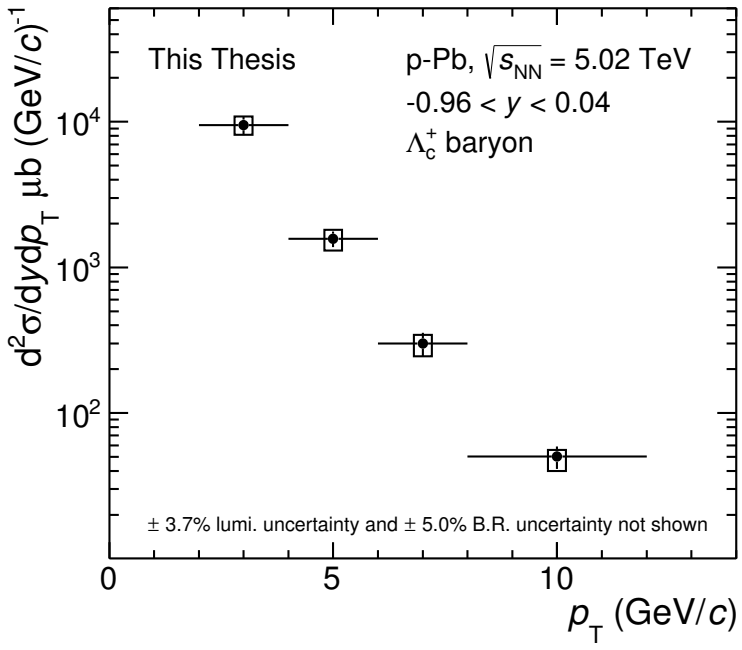

Figure 5.2: p_{T}-differential cross section of the Λ_{c}^{+} baryon measured in the $\Lambda_{\mathrm{c}}^{+} \to \mathrm{pK}^{-}\pi^{+}$ channel in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

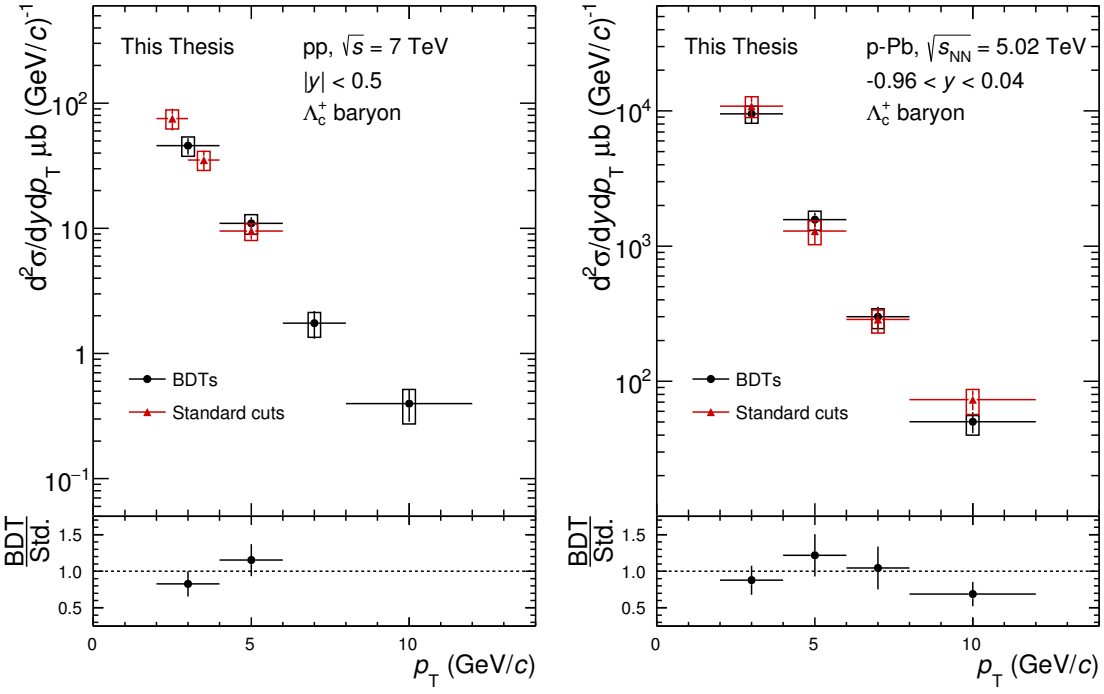

Figure 5.3: p_T-differential cross section of the Λ_c^+ baryon measured in the $\Lambda_c^+ \to pK^-\pi^+$ channel in pp collisions at $\sqrt{s} = 7$ TeV (left) and in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV (right) in this work (i.e. using Boosted Decision Trees as part of the analysis method), compared to the analysis performed using a standard, rectangular cut-based method. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes. The bottom panels show the ratio between the two methods, where only statistical uncertainties are shown and the uncertainties are considered as uncorrelated.

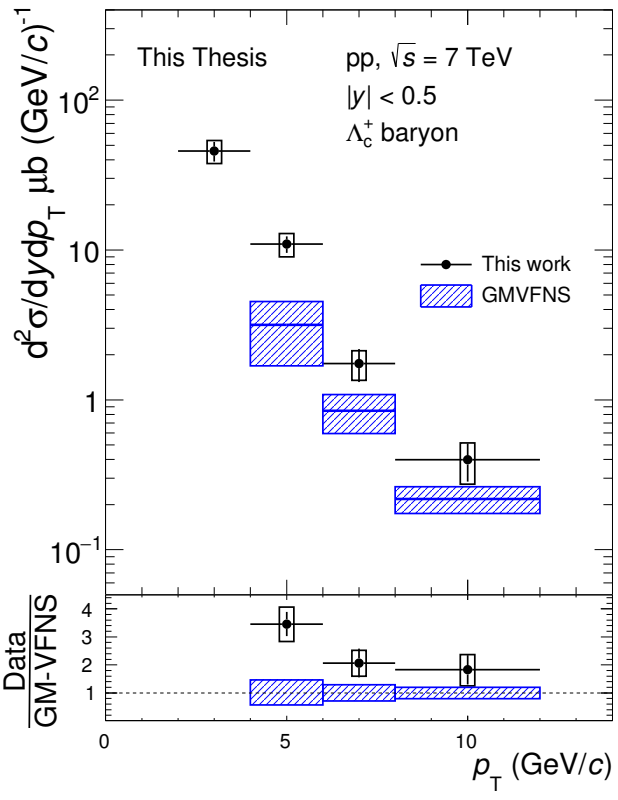

Figure 5.4: p_{T}-differential cross section of the Λ_{c}^{+} baryon measured in the $\Lambda_{\mathrm{c}}^{+} \rightarrow \mathrm{pK}^{-}\pi^{+}$ channel in pp collisions compared to the GM-VFNS calculation. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes. The bottom panel shows the ratio of the measured data with the GM-VFNS calculation [87, 88] (GM-VFNS from I. Schienbein).

5.2 Λ_c^+/D^0 ratio in pp and p–Pb collisions

The D^0, D^+, D^{*+} and D_s^+ meson production cross sections are all measured by ALICE in pp, p–Pb and Pb–Pb collisions, via the hadronic decay channels $D^0 \to K^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$, $D^{*+} \to D^0\pi^+$ and $D_s^+ \to \phi\pi^+ \to K^-K^+\pi^+$. All the aforementioned D meson production cross sections are measured in pp collisions at $\sqrt{s} = 7$ TeV and in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, using an analysis method based on the selection of decay topologies displaced from the interaction vertex, similar to the Λ_c baryon. The D^0 has also been measured down to $p_T = 0$ using an analysis based on the estimation and subtraction of the uncorrelated combinatorial background without reconstruction of the D^0 decay vertex [144, 145].

Using the Λ_c^+ production cross section measurements presented in this work, and the D^0 production cross section measurements in pp [145] and p–Pb [144] collisions, the baryon-to-meson ratio Λ_c^+/D^0 has been measured in pp and p–Pb collisions in the p_T range $2 < p_T < 12$ GeV/c. It is noted here that the baryon-to-meson ratio Λ_c^+/D^0 is equivalent to the antiparticle-plus-particle baryon-to-meson ratio $(\Lambda_c^+ + \Lambda_c^-)/(D^0 + \bar{D}^0)$, since the D meson measurements in [144, 145] and the Λ_c^+ measurement presented in this work reconstruct both particle and antiparticle, and then determine the cross section by dividing the particle-plus-antiparticle cross section by a factor of two. The Λ_c^+/D^0 ratio in both collision systems is shown in figure 5.5. The correlation between the systematic uncertainties is given in table 4.6. The Λ_c^+/D^0 ratio in both collision systems is seen to agree within statistical and systematic uncertainties. A hint of a decrease with p_T is also seen in both collision systems, although the data are also consistent with a flat Λ_c^+/D^0 ratio within uncertainties.

Figure 5.5 also shows comparisons of the Λ_c^+/D^0 ratio to predictions from MC generators. These predictions include PYTHIA8 [92] with the Monash tune [237] and a tune which includes colour reconnection [116], and DIPSY [98] with rope hadronisation[2] [118]. The measured Λ_c^+/D^0 ratio is higher than all predictions from MC. While the PYTHIA8 Monash tune and DIPSY with rope hadronisation have a relatively constant ratio around ~ 0.1, the PYTHIA prediction with colour reconnection raises the Λ_c^+/D^0 ratio by a factor of $\sim 2-3$, bringing the prediction closer to the data. The falling Λ_c^+/D^0 ratio with p_T is also approximately reproduced with the colour reconnection model.

The Λ_c^+/D^0 ratio in pp and p–Pb collisions integrated over the measured p_T range $2 < p_T < 12$ GeV/c is shown in figure 5.6, as a function of the centre-of-mass

[2]The colour reconnection and rope hadronisation models are discussed in chapter 2.

rapidity y. To obtain the Λ_c cross section integrated in the measured p_T interval the raw yield extraction systematic uncertainty is taken as uncorrelated with p_T, and all other systematic uncertainties are conservatively assumed to be correlated with p_T, as shown in table 4.6. The statistical uncertainties are uncorrelated with p_T. The Λ_c/D^0 ratio is obtained assuming the same correlation for the D^0 cross section with p_T. The Λ_c^+/D^0 ratio in each collision system is

$$(\Lambda_c^+/D^0)_{pp} = 0.5501 \pm 0.0679 \text{ (stat.)} \, ^{+0.0945}_{-0.0952} \text{ (syst.)} \tag{5.1}$$

$$(\Lambda_c^+/D^0)_{p-Pb} = 0.6371 \pm 0.08122 \text{ (stat.)} \, ^{+0.1669}_{-0.1044} \text{ (syst.)} \tag{5.2}$$

The measured Λ_c^+/D^0 ratios in pp and p–Pb collisions are consistent with each other within uncertainties.

The Λ_c^+/D^0 ratio as presented in this work is compared with the Λ_c^+/D^0 as measured by LHCb in pp collisions at $\sqrt{s} = 7$ TeV, shown in figure 5.7 as a function of p_T and in figure 5.8 as a function of the centre-of-mass rapidity in the common p_T range of each analysis $2 < p_T < 8$ GeV/c. The LHCb measurement as a function of p_T is systematically lower than the measurement presented in this work. However the trend as a function of y seen by LHCb, where the Λ_c^+/D^0 decreases towards mid-rapidity, is not consistent with the measurement presented in this work.

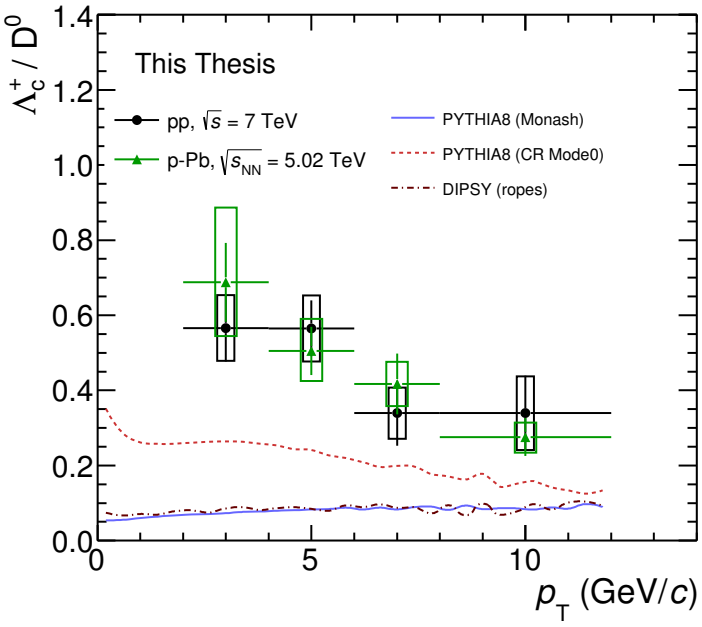

Figure 5.5: The Λ_c^+/D^0 ratio as a function of p_T measured in pp collisions at $\sqrt{s}$ = 7 TeV and in p–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes. The measurements are compared to predictions [238] from pp MC generators [92, 98, 116, 118, 237].

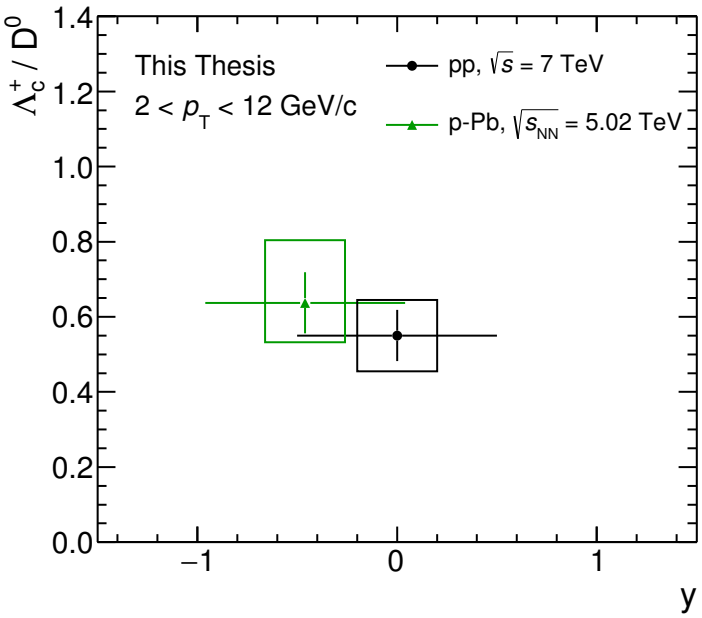

Figure 5.6: The Λ_c^+/D^0 ratio integrated over the measured p_T range $2 < p_T < 12$ GeV/c, as a function of the centre-of-mass rapidity y, in pp collisions at $\sqrt{s}$ = 7 TeV and in p–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

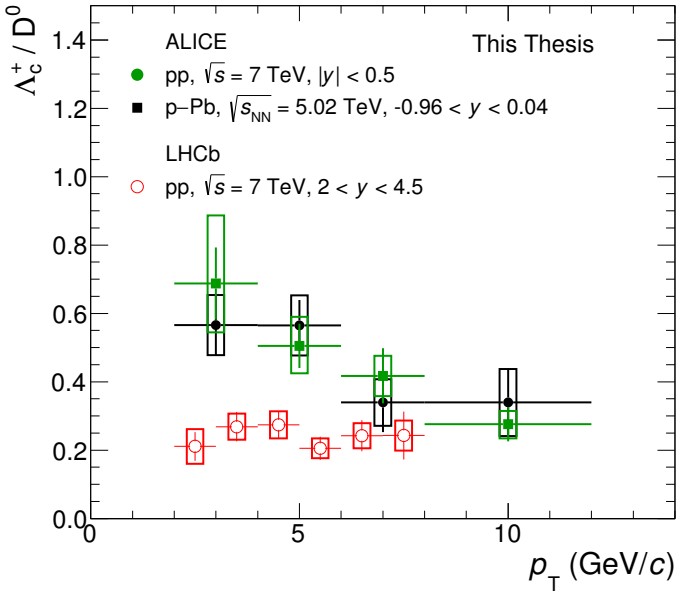

Figure 5.7: The Λ_c^+/D^0 ratio in pp collisions $\sqrt{s} = 7$ TeV and in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV as presented in this work compared with that measured by LHCb in pp collisions at $\sqrt{s} = 7$ TeV, as a function of p_T. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

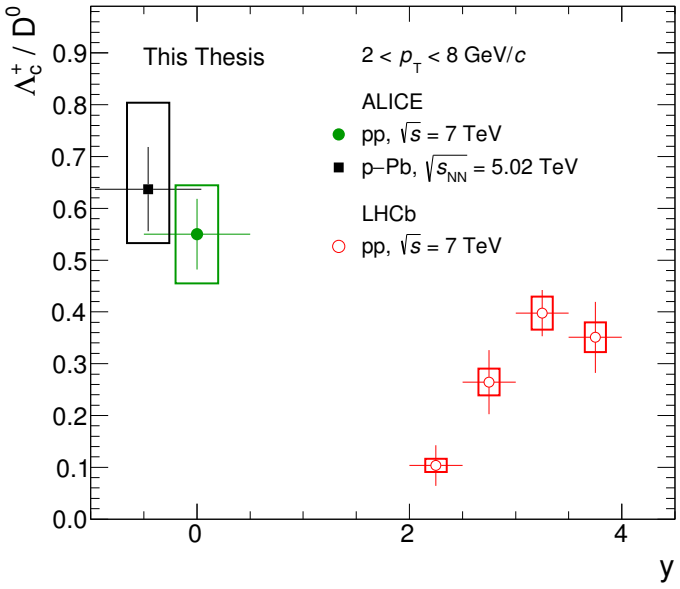

Figure 5.8: The Λ_c^+/D^0 ratio in pp collisions $\sqrt{s} = 7$ TeV and in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV as presented in this work compared with that measured by LHCb in pp collisions at $\sqrt{s} = 7$ TeV, as a function of the centre-of-mass rapidity y in the p_T range $2 < p_T < 8$ GeV/c. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

5.3 The nuclear modification of the Λ_c^+ yield

Figure 5.9 shows the p_T-differential cross section of the Λ_c^+ baryon in p–Pb collisions in comparison with the p_T-differential cross section in pp collisions, scaled to the centre-of-mass energy $\sqrt{s} = 5.02$ TeV as described in section 4.8, and scaled by the mass number $A = 208$. The two measurements are seen to be compatible within uncertainties. The nuclear modification factor in p–Pb collisions R_{pPb} is then computed by dividing the Λ_c^+ cross section in p–Pb collisions by the pp reference:

$$R_{pPb} = \frac{1}{A} \frac{d\sigma_{pPb}^{prompt\Lambda_c}/dp_T}{d\sigma_{pp}^{prompt\Lambda_c}/dp_T} \tag{5.3}$$

The Λ_c^+ R_{pPb} is shown in figure 5.10. The measured R_{pPb} is compatible with unity within the statistical and systematic uncertainties.

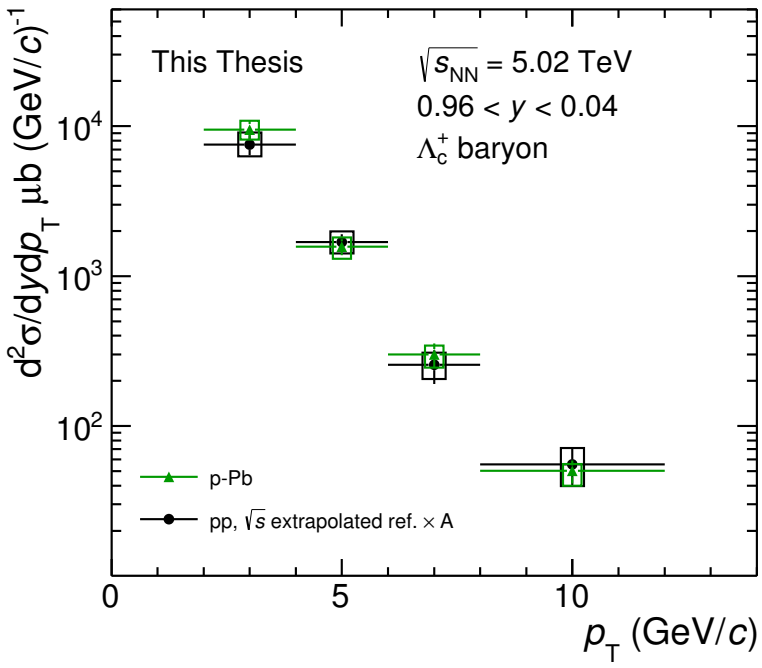

Figure 5.9: p_T-differential cross section of the Λ_c^+ baryon measured in the $\Lambda_c^+ \to pK^-\pi^+$ channel in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and in pp collisions at $\sqrt{s} = 7$ TeV, scaled by the Pb mass number $A = 208$ and scaled to match the p–Pb $\sqrt{s_{NN}}$. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

The R_{pPb} has also been measured for D^0,D^+,D^{*+} and D_s^+ mesons [144]. The average D meson (D^0,D^+ and D^{*+}) R_{pPb} is shown in figure 5.11 in comparison with the Λ_c^+ R_{pPb}. Both are seen to be compatible with each other.

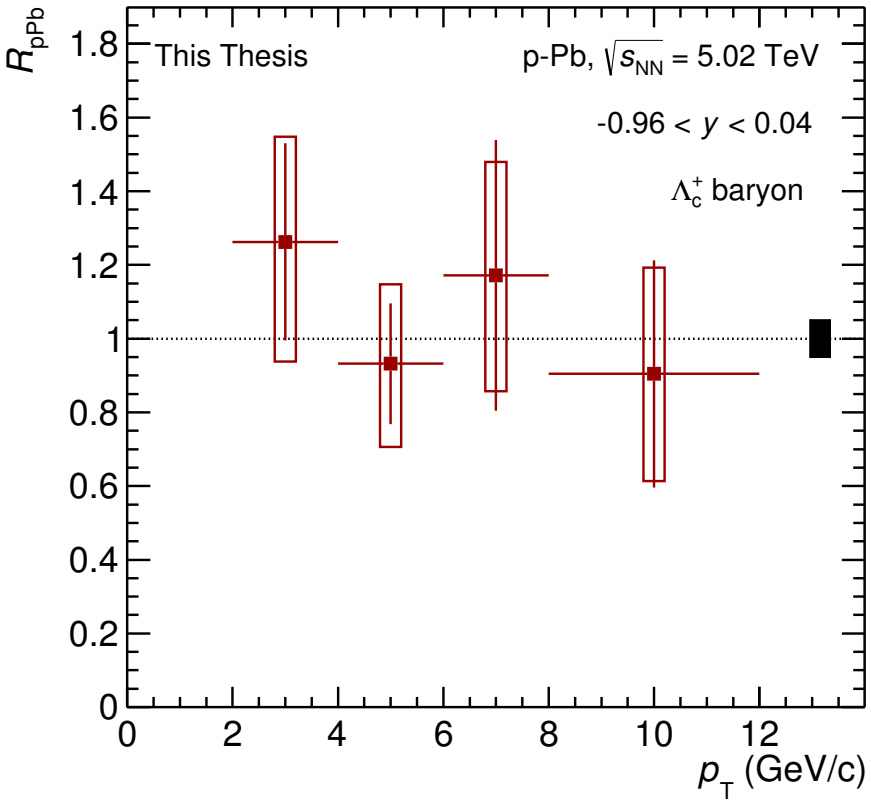

Figure 5.10: The nuclear modification factor R_{pPb} of the Λ_c^+ baryon at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes, and the normalisation uncertainty (coming from the uncertainty on the luminosity in the p–Pb measurement and the pp reference) is shown as the box on the right of the plot.

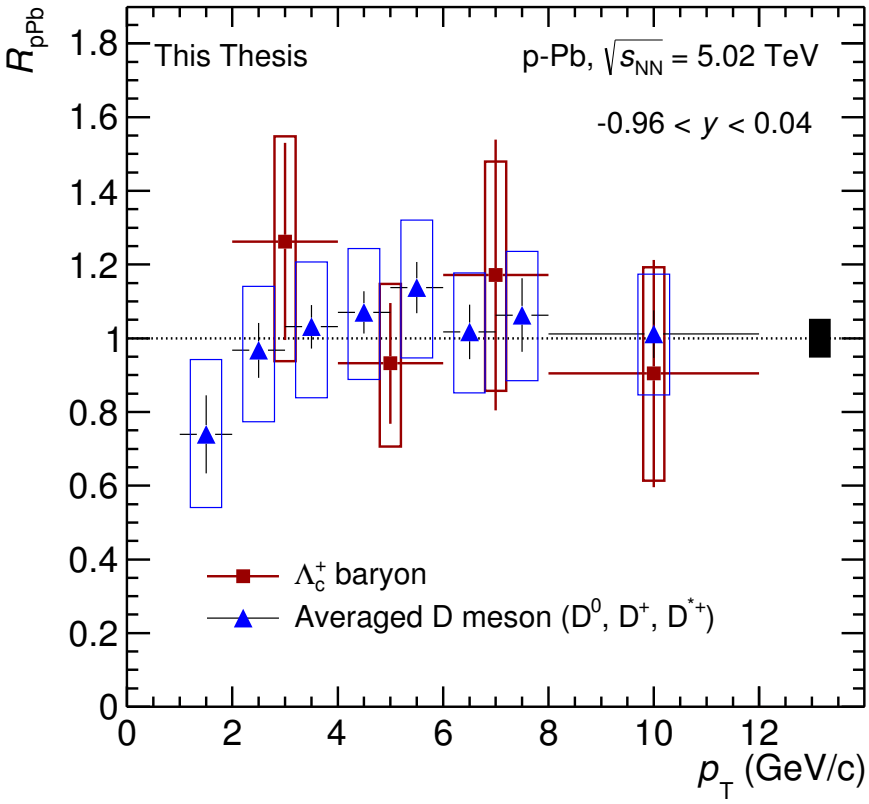

Figure 5.11: Comparison of the nuclear modification factor $R_{\rm pPb}$ of the Λ_c^+ baryon with the averaged $R_{\rm pPb}$ of D^0, D^+ and D^{*+} [144]. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes, and the normalisation uncertainty (coming from the uncertainty on the luminosity in the p–Pb measurement and the pp reference) is shown as the box on the right of the plot.

5.4 Discussion

This chapter has presented the measurements of the p_{T}-differential cross section of the Λ_{c}^{+} baryon in pp collisions at $\sqrt{s} = 7$ TeV and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, the baryon-to-meson ratio $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ in both collision systems, and the nuclear modification factor R_{pPb} of the Λ_{c}^{+} baryon.

The measurement in pp collisions is larger than the pQCD calculation from GM-VFNS by a factor of 2-3. GM-VFNS describes the D^{0}, D^{+}, D^{*+} and $\mathrm{D}_{\mathrm{s}}^{+}$ meson production cross sections within uncertainties [145], and describes Λ_{c}^{+} production at forward rapidities [168]. The measured baryon-to-meson ratio $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ in pp collisions is also significantly larger than predictions from PYTHIA and DIPSY, although colour reconnection models implemented in PYTHIA bring the predicted ratio closer to the measured data. The measured baryon-to-meson ratio $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ in p–Pb collisions is compatible with that measured in pp collisions.

For further insight, a comparison of the measured fragmentation fraction $f(c \to \Lambda_{\mathrm{c}}^{+})$ to the integrated baryon-to-meson ratio in the p_{T} region of the measurement can be made. The fragmentation fraction $f(c \to \Lambda_{\mathrm{c}}^{+})$ is measured from a combined analysis of different measurements (see [109] and references within, summarised in figure 2.3) to be 0.0623 ± 0.0041, and the fragmentation fraction $f(c \to \mathrm{D}^{0})$ is measured to be 0.6086 ± 0.0076, leading to the ratio $f(c \to \Lambda_{\mathrm{c}}^{+})/f(c \to \mathrm{D}^{0}) = 0.109 \pm 0.007$[3]. The measured $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratio in pp collisions in the range $2 < p_{\mathrm{T}} < 12$ GeV/c is larger than this ratio by a factor of 5, and is in excess of the ratio by 3.8σ considering the statistical and systematic uncertainties of both measurements. However care should be taken in a direct comparison of these two values, since the measured $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratio from this work is in a given p_{T} range, and thus does not include the full p_{T} phase space. The fraction of Λ_{c}^{+} produced in this p_{T} region with respect to the full p_{T} range is estimated from PYTHIA to be around 50%. An extrapolation of the Λ_{c}^{+} cross section down to $p_{\mathrm{T}} = 0$ is outside the scope of this work, though the large uncertainties and relatively few p_{T} bins of this measurement would suggest an extrapolation would itself have large uncertainties.

Altogether, the data suggest a larger Λ_{c}^{+} production cross section than theoretical predictions which include a pQCD treatment of charm-quark production and hadronisation based on phenomenological models. The measurements are also larger than previous measurements in different collision systems at different centre-of-mass

[3]Note that the branching ratios for this calculation have been updated to the latest values in the PDG [13].

energies, albeit in a different kinematic window, which suggests an enhanced $\Lambda_{\rm c}^+$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV and in p–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV with respect to previous measurements. The $\Lambda_{\rm c}^+/{\rm D}^0$ ratio is however, at least qualitatively, better reproduced by models which modify the hadronisation mechanism in terms of colour reconnection, allowing a baryon enhancement. This suggests that the difference between the $\Lambda_{\rm c}^+$ production measurement and theory is likely to be due to a modified hadronisation mechanism enhancing baryon production. This is backed up by the fact that the production of D mesons is well described by pQCD predictions including GM-VFNS - the same theoretical scheme which underpredicts $\Lambda_{\rm c}^+$ production. The agreement in pp and p–Pb collisions of the $\Lambda_{\rm c}^+/{\rm D}^0$ suggests the mechanisms responsible for $\Lambda_{\rm c}^+$ production and hadronisation are common in both collision systems.

In addition, since the LHCb measurement of the $\Lambda_{\rm c}^+/{\rm D}^0$ ratio in pp collisions at the same centre-of-mass energy $\sqrt{s} = 7$ TeV and in the rapidity region $2 < y < 4.5$ is below that presented in this work, this work suggests that there is a rapidity dependence in the relative production of baryons and mesons. The fact that the LHCb production cross section is described by GM-VFNS also suggests this rapidity dependence in $\Lambda_{\rm c}^+$ production that is not described by GM-VFNS.

The measured $\Lambda_{\rm c}^+$ production cross section in p–Pb collisions at $\sqrt{s_{\rm NN}} = 5$ TeV agrees with the same measurement in pp collisions scaled to the same energy, and scaled by the lead mass number. The $\Lambda_{\rm c}^+$ nuclear modification factor $R_{\rm pPb}$ is thus compatible with unity, and consistent with the $R_{\rm pPb}$ of D mesons in the same $p_{\rm T}$ range. Within uncertainties this suggests that there is minimal modification in the $\Lambda_{\rm c}^+$ yield due to cold nuclear matter effects, within the experimental uncertainties.

5.5 Combined analysis of $\Lambda_{\rm c}^+ \to {\rm pK}^-\pi^+$, $\Lambda_{\rm c}^+ \to {\rm pK}_{\rm s}^0$ and $\Lambda_{\rm c}^+ \to \Lambda{\rm e}^+\nu_{\rm e}$

The $\Lambda_{\rm c}^+$ $p_{\rm T}$-differential cross section has also been measured by others in the decay channels $\Lambda_{\rm c}^+ \to {\rm pK}_{\rm s}^0$ and $\Lambda_{\rm c}^+ \to \Lambda{\rm e}^+\nu_{\rm e}$ [1]. The $\Lambda_{\rm c}^+ \to {\rm pK}_{\rm s}^0$ analysis uses a similar technique to the analysis presented in this work, where ${\rm K}_{\rm S}^0$ mesons ($\to \pi^+\pi^-$) and protons are reconstructed and identified, selection on the decay topology is made and the signal is extracted via an invariant mass analysis, correcting for the reconstruction and selection efficiency and acceptance and subtracting $\Lambda_{\rm c}^+$ contributions from $\Lambda_{\rm b}^0$. For the $\Lambda_{\rm c}^+ \to \Lambda{\rm e}^+\nu_{\rm e}$ analysis, Λ and e^+ pairs and their charge conjugates are reconstructed and identified, and the uncorrelated background is subtracted by

subtracting wrong-sign pairs ($\bar{\Lambda}e^+$ and Λe^-), after correcting for contributions of Λ_b^0 semileptonic decays to wrong-sign pairs, and of $\Xi_c^{0,+}$ decays to right-sign pairs, and correcting for the missing momentum of the neutrino, similar to the approach performed in e.g. [239, 240]. The measurements are then averaged. The BDT analysis in p–Pb collisions performed in this work contributes to the averaged p–Pb cross section. The BDT analysis in pp collisions performed in this work does not contribute to the averaged pp cross section for the publication, but instead provides a useful cross check for this measurement.

The decay-channel averaged p_T-differential cross sections of the Λ_c^+ baryon in pp collisions and p–Pb collisions is shown in figure 5.12. The measurements in different decay channels and using different analysis techniques agree within uncertainties. The (merged) Λ_c^+/D^0 ratio and nuclear modification factor R_{pPb} are shown in figure 5.13 and 5.14, respectively. With the combination of different measurements, a higher precision result is achieved. Predictions for the R_{pPb} are also shown for the merged analysis, which include POWHEG+PYTHIA6 pQCD production [241] with CT10NLO [242] plus the EPS09 parametrisation of the nuclear PDF [121], and the POWLANG model for charmed hadrons [126] where heavy-quark transport through a deconfined medium is calculated. The measurement is consistent with both models within the experimental uncertainties.

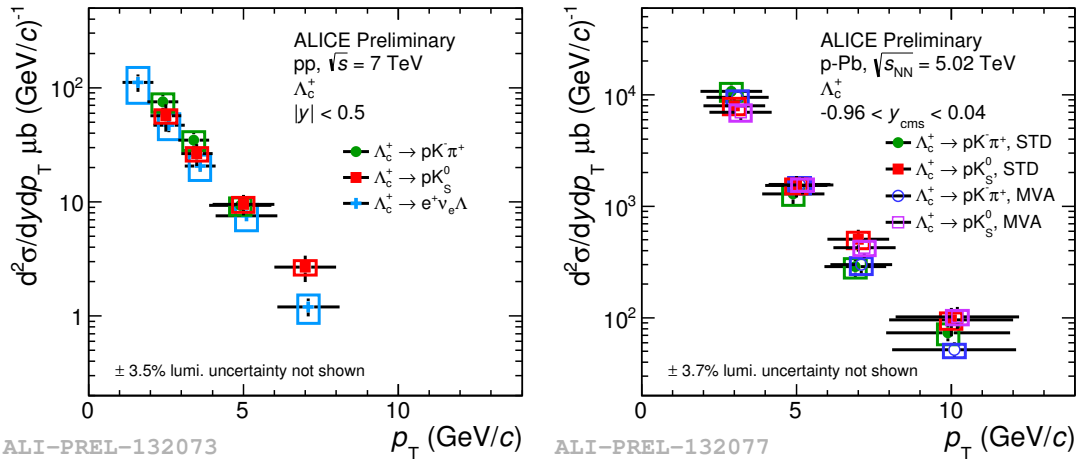

Figure 5.12: p_T-differential cross sections of the Λ_c^+ baryon measured in the decay channels $\Lambda_c^+ \to pK^-\pi^+$, $\Lambda_c^+ \to pK_s^0$ and $\Lambda_c^+ \to \Lambda e^+\nu_e$ in pp and p–Pb collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s_{NN}} = 5.02$ TeV, respectively [1]. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

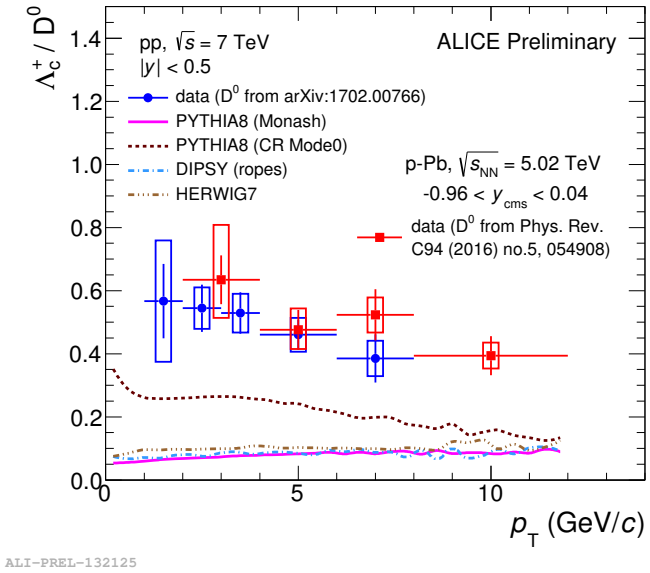

Figure 5.13: Λ_c^+/D^0 ratio from the merged analyses as described in the text, in pp collisions at $\sqrt{s} = 7$ TeV and in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [1]. The results are compared to predictions from MC generators [92, 98, 116, 118, 237]. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

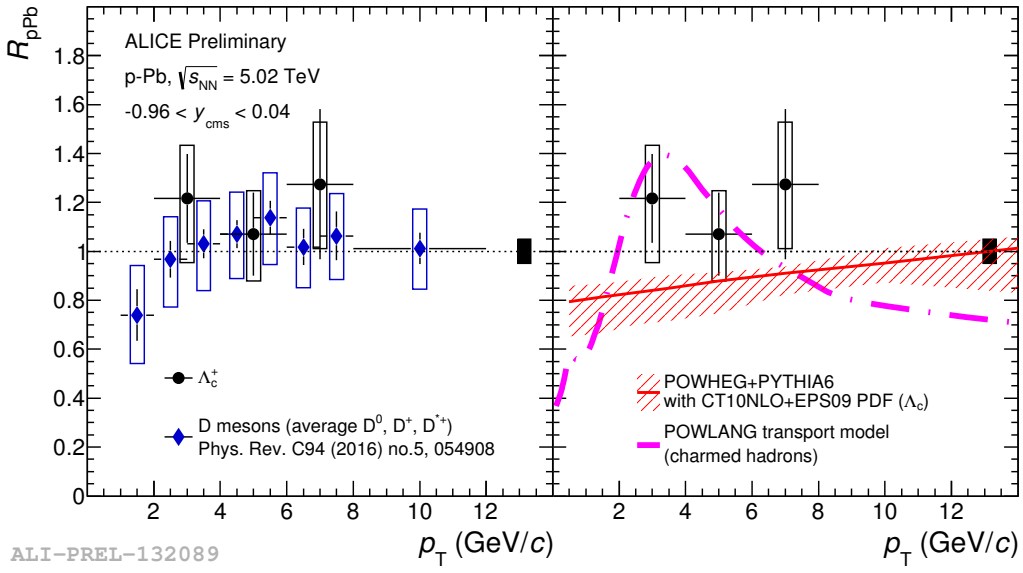

Figure 5.14: R_{pPb} from the merged analyses as described in the text, in pp collisions at $\sqrt{s} = 7$ TeV and p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [1]. These results are compared to predictions including POWHEG+PYTHIA6 [241] plus the CT10NLO+EPS09 [121, 242] parameterisation of the nuclear PDF, and a POWLANG transport model [126]. The statistical uncertainties are indicated by the vertical lines and the systematic uncertainties are indicated by the boxes.

Chapter 6

Towards a measurement of the Λ_{c}^{+} baryon in Pb–Pb collisions

The Inner Tracking System currently installed at ALICE (see chapter 3) has allowed for measurements of charmed mesons (D^0,D^+,D^{*+} and $\mathrm{D}_{\mathrm{s}}^+$) over a wide transverse momentum range, in pp, p–Pb and Pb–Pb collisions, made possible due to the relatively long lifetimes of the mesons - the proper lifetime $c\tau \sim 310$ μm for D^+, and ~ 120 μm for D^0. To reduce the combinatorial background, selections can be made on variables related to the displaced vertex of the D meson, and the current impact parameter resolution of the ITS (~ 150 μm for ~ 0.4 GeV/c tracks, down to ~ 20 μm for ~ 10 GeV/c tracks) allows for large background reduction. Measurements at low transverse momentum however are challenging with currently available datasets - measurements below 1 GeV/c are possible in pp and p–Pb with large statistical uncertainties, and are not possible in Pb–Pb. Furthermore, charmed baryon measurements (Λ_{c}) have not been possible yet in Pb–Pb - its proper lifetime $c\tau \sim 60$ μm is lower than the impact parameter resolution of the present ITS in the transverse momentum range of the majority of Λ_{c} daughters, meaning topological selections are not enough for the large combinatorial background reduction needed for signal extraction. In addition, the present ITS detector also has limited read-out rate capabilities. The ITS can run at a maximum rate of 1 kHz - much less that the maximum planned interaction rate of 50 kHz in Run 3 [223]. For rare heavy-flavour probes such as the Λ_{c} baryon and B mesons a dataset with much larger integrated luminosity is vital for precise measurements.

A complete upgrade strategy for the ITS has thus been developed. A seven-layer silicon pixel detector will be build which will include a reduced material budget, smaller pixel segmentation, and a shorter read-out time. The full ALICE upgrade

strategy Letter of Intent (LoI) was approved in March 2013 [243], and the Technical Design Report (TDR) for the ITS upgrade was published in March 2014 [223].

The physics performance of the ITS upgrade has been studied for the Λ_c, using a full detailed simulation and reconstruction of the new ITS, where Λ_c candidates are reconstructed and particle identification and topological selections are applied, in order to assess the expected reach and precision of an R_{AA} and Λ_c/D^0 measurement. A multivariate treatment of the selection may further improve the background rejection, and so far has not yet been attempted but is motivated to understand whether the measurement may be improved. One might hope that due to the improvement in the detector resolution and thus the 'power' of the discriminating variables that a greater level of improvement in the background rejection and thus the statistical significance of the yield extraction may be possible, with respect to the improvement achieved in the Run 1 analysis presented in chapter 4.

This chapter presents a feasibility study on the use of multivariate analysis techniques (specifically Boosted Decision Trees) in the Λ_c analysis of Run 3, and compares to a standard rectangular cut approach. An overview of the ITS upgrade specifications are detailed in section 6.1. Some physics performance studies carried out to estimate the physics reach of the ITS upgrade are shown in section 6.2. An extension of these studies to include a multivariate selection is detailed in section 6.3.

6.1 Upgrade specifications

The ITS upgrade is based on 7 concentric layers of silicon pixels, grouped as an inner barrel (layers 0, 1 and 2) and an outer barrel (layers 3, 4, 5 and 6), azimuthally segmented in units called staves. The staves consist of a space frame which acts as a support structure, a cold plate containing the cooling pipes and an integrated circuit containing the flexible printed circuit onto which the pixel chips are bonded. The outer barrel staves are further segmented in azimuth into halves, named half staves, and longitudinally into modules. Table 6.1 lists the ITS upgrade's main geometrical parameters.

The ITS upgrade features many improvements with respect to the current ITS. The first detection layer will be closer to the beam pipe (with a beam pipe inner radius of 19 mm compared to the current 29 mm). The material budget will also be significantly reduced - the new pixel chip technology allows for a factor of 7 improvement in the thickness of the chip (50 μm instead of 350 μm), and a highly optimised scheme for the power distribution, which along with improvements to the cooling,

	Inner Barrel			Outer barrel			
	Inner layers			Middle layers		Outer layers	
	layer 0	layer 1	layer 2	layer 3	layer 4	layer 5	layer 6
Radial position (min.) (mm)	22.4	30.1	37.8	194.4	243.9	342.3	391.8
Radial position (max.) (mm)	26.7	34.6	42.1	197.7	247.0	345.4	394.9
Length (sensitive area) (mm)	271	271	271	843	843	1475	1475
Pseudorapidity coverage (deg)	±2.5	±2.3	±2.0	±1.5	±1.4	±1.4	±1.3
Active area (cm^2)	421	562	702	10483	13104	32105	36691
Pixel chip dimensions (mm^2)				15×30			
N. pixel chips	108	144	180	2688	3360	8232	9408
N. staves	12	16	20	24	30	42	48
Pixel size (μm^2)	30×30			30×30			

Table 6.1: Geometric parameters of the ITS upgrade. Table modified from [223] according to recent developments [244].

mechanics and other detector elements will help achieve a detector with a radiation length of around $0.3\%X_0$ per layer for the inner three layers. The maximum readout time will also be substantially improved, with the upgraded readout architecture allowing the detector to take data at 100 kHz for Pb–Pb collisions, and 400 kHz for pp collisions [223].

At the heart of the upgrade campaign is the upgraded pixel chip. Extensive R&D efforts have lead to the development of the ALPIDE pixel chip [244] to be developed. The ALPIDE pixel chip is based on Monolithic Active Pixel Sensors (MAPS), where the silicon sensor and readout chip are integrated in one silicon wafer. This allows for a significant improvement in spatial resolution, material thickness and signal-over-noise for a given power consumption over the current ITS setup, where the silicon pixel sensors are bump-bonded onto read-out electronics. 0.18 μm CMOS technology will be used, as depicted in figure 6.1. A charged particle traversing the chip's active volume will liberate electron/hole pairs. The electrons drift around the chip until they enter the depleted volume and are collected by the n-well collection diode. The chip features a deep p-well, which shields the other n-well containing the PMOS transistor from the active volume, preventing them from collecting signal charge. This allows for full CMOS implementation without deterioration of the charge, which means more complex logic can be implemented in the pixel, which in turn minimises the digital circuitry at the edge of the chip, which would add to the insensitive area of the chip. Another benefit to this layout is that the readout can be improved from the typical 'rolling shutter readout', where each pixel is read out one-by-one, to a more power

efficient scheme. This has lead to a power consumption of less than 40 mW cm^2. Test beam results have shown excellent performance of the chip prototypes, with a detection efficiency above 99% a fake hit rate better than 10^{-5} and a spacial resolution of 5 μm over a large range of operational settings [244].

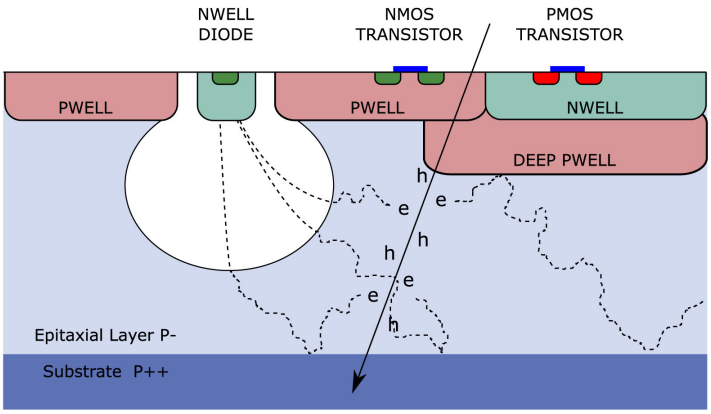

Figure 6.1: Schematic view of the pALPIDE pixel structure [244]

6.2 Upgrade physics performance

Full Monte Carlo simulations have been developed to study the performance of the ITS upgrade, and to optimise its layout [223]. Simulated Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} =$ 5.5 TeV were generated using the HIJING [96] event generator tuned to a particle multiplicity expected from $\sqrt{s_{\mathrm{NN}}}$ scaling laws where $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \propto s_{NN}^{0.15}$ [199]. A full GEANT simulation of the detector has been carried out, which includes transport through the detector, tracking and vertex reconstruction. The detector response is a crucial input to the full Monte Carlo simulations, which in this case is the response of the pixel chip. Characteristics of the pixel chip including the average pixel noise distribution, the fake pixel rate and the charge spread function[1] have been extracted from test beam measurements, and the pixel response simulation takes into account these characteristics to generate charge contributions for each pixel which are digitised and formed into clusters. The track finding follows the same algorithm as described in section 3.3.1, and tracks are propagated to the TPC. The ITS stand-alone tracking performance is also evaluated. Figure 6.2 shows the impact parameter resolution and tracking efficiency as a function of transverse momentum for the ITS upgrade,

[1]The average pixel noise distribution describes the analogue noise charge in the pixels, the fake pixel rate is the number of 'fake' pixels which fire, and the charge spread function describes the distribution of cluster charge as a function of the distance from the track.

and a comparison with the current ITS. A significant improvement in the impact parameter resolution of around a factor 3(5) will be achievable in the $r\phi(z)$ plane at $p_T = 400$ MeV/c compared to the current ITS. The tracking efficiency will stay above $\sim 80\%$ for particles down to $p_T \sim 100$ MeV.

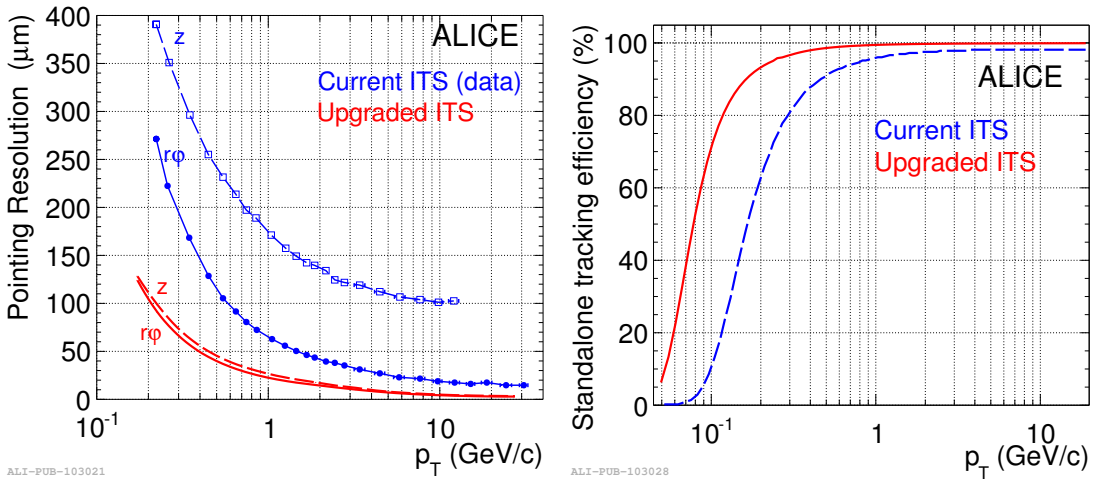

Figure 6.2: Left: the impact parameter resolution for primary charged pions in $r\phi$ and z for the ITS upgrade, compared with the corresponding resolution for the current ITS. Right: the ITS stand-alone tracking efficiency for primary charged pions for the ITS upgrade, compared with the current ITS [223].

The improved resolution will be crucial for many measurements down to low p_T, including precision measurements of charm and beauty mesons (D, B) to quantify in-medium energy loss and thermalisation of heavy quarks, and low mass dielectrons to address questions related to the temperature and lifetime of the system, and the generation of hadron masses. In addition, the baryon-to-meson ratio (Λ_c/D, Λ_b^0/B) aims to address open questions concerning hadronisation of heavy quarks in the medium, including the role of coalescence as a hadronisation mechanism and the possible presence of diquarks in the medium (see chapter 2).

Performance studies to assess the reach of these observables have been performed for the ITS TDR [223]. For these studies MC simulations of central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV with impact parameter $0 < b < 5$ fm corresponding to the centrality class $0 - 10\%$ have been used. HIJING [96] events were enhanced with heavy-quark pairs generated using PYTHIA6 for the heavy-flavour performance studies. For the study of heavy-flavour decays (D mesons, Λ_c baryon) tracks and vertices were reconstructed, charm decay candidates were built, and selections were made on the candidates to reduce combinatorial background, including PID and topological

selections. The signal and background is scaled to 1.6×10^{10} events corresponding to the target integrated luminosity for the upgrade physics programme is 10 nb^{-1} for Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.5$ TeV[2], and the signal is scaled to the expected $\mathrm{d}N/\mathrm{d}y \times \mathrm{BR} = 1.4$, and an estimation of the $R_{\rm AA}$ equal to the $R_{\rm AA}$ of the strange Λ baryon.

Figure 6.3 shows the expected signal-over-background and significance ($S/\sqrt{S+B}$) of the $\Lambda_{\rm c}$ for the signal and background scaling described above, where the significance is shown for the centrality class $0-20\%$ assuming that the significance / $\sqrt{N_{events}}$ is the same in both $0-10\%$ and $0-20\%$. In this centrality class the measurement will be able to reach down to 2 GeV/c with a significance of ~ 8 (which corresponds to a relative statistical uncertainty of 12%). Figure 6.4 shows the nuclear modification factor and $\Lambda_{\rm c}/{\rm D}^0$ ratio, with corresponding statistical uncertainties. The statistical uncertainties for the pp reference are assumed to be negligible compared with the statistical uncertainties in Pb–Pb. The systematic uncertainties in this plot were estimated from the experience of the D meson measurements, and from the expected uncertainties from the $\rm D^0$. It is worth pointing out here that the systematic uncertainties are dominated by the uncertainty on the feed-down, which was determined from the theoretical uncertainty on prompt $\Lambda_{\rm c}$ and non-prompt $\Lambda_{\rm b}^0 \rightarrow \Lambda_{\rm c} + {\rm X}$ production. It may be possible to reduce this uncertainty since $\Lambda_{\rm c}$ and $\Lambda_{\rm b}^0$ production has been measured at the LHC with smaller uncertainties than the uncertainties from theory.

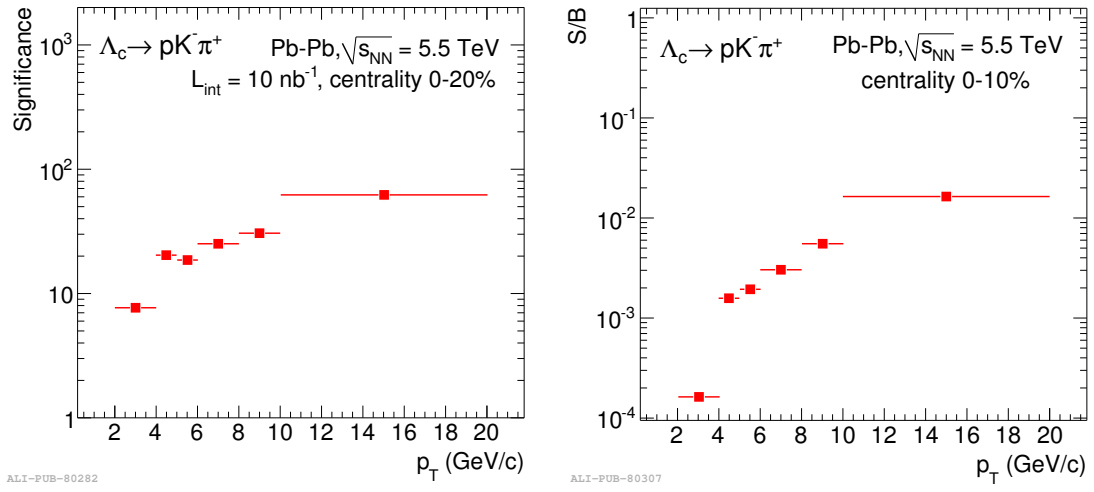

Figure 6.3: Expected significance (left) and S/B (right) of the $\Lambda_{\rm c}^+$ baryon as a function of $p_{\rm T}$, for 1.6×10^{10} Pb–Pb events at $\sqrt{s_{\rm NN}} = 5.5$ TeV, in the centrality class 0-20 % [223].

[2]to be taken over $3\frac{1}{2}$ months of operation with Pb beams in the LHC

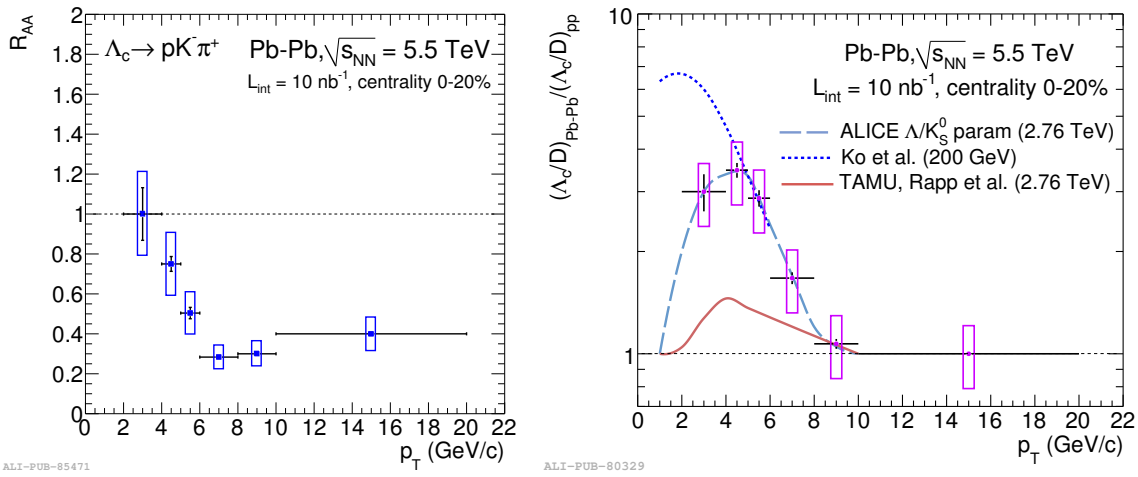

Figure 6.4: Expected Λ_c measurements after the ITS upgrade. Left: The nuclear modification factor R_{AA} for the Λ_c in central Pb–Pb collisions. Right: The baryon-to-meson double ratio Λ_c/D^0, compared to model calculations [135, 245]. These figures are for the $0-20\%$ centrality class assuming $\mathcal{L}_{int} = 10$ nb^{-1}. The expected statistical uncertainties are shown as vertical lines, and the expected systematic uncertainties are shown as boxes [223].

These studies demonstrate that given ALICE achieves its target integrated luminosity, a measurement of the Λ_c baryon will be achievable with good precision down to 2 GeV/c, where different hadronisation scenarios will be distinguishable. This data would be taken over a number of years during Run 3 (typically the heavy-ion run at the LHC is for 1 month per year). Under this assumption ~ 3 nb^{-1} of integrated luminosity would be taken in the first year of Run 3, which would correspond to a significance of around 4 in the p_T interval $2 < p_T < 4$ GeV/c. Furthermore, to further resolve the shape of the R_{AA} and Λ_c/D^0, finer binning may be desirable, especially where a 'steep' gradient may be expected in the mid-p_T ($4 \lesssim p_T \lesssim 10$ GeV/c) region. In both of these cases, it would clearly be beneficial to achieve the most optimal background rejection, to achieve greater statistical precision.

6.3 Analysis

To assess the performance of BDTs compared to the studies already performed, PID and cuts are made which match those in the TDR study. For the PID $n\sigma$ cuts were made in both the TPC and TOF. In the TOF a $\pm 3\sigma$ cut was made for pions, kaons and protons. In the TPC p_T-dependent $n\sigma$ cuts were made for kaons and protons, where for kaons a $\pm 3\sigma$ cut was made for $p_T < 0.5$ GeV/c, tightening to $\pm 2\sigma$ above

that, and for protons a $\pm 3\sigma$ cut was made up to $p_{\mathrm{T}} < 0.8$ GeV/c, tightening to $\pm 2\sigma$ above that, which reflects the p_{T} regions where pions begin to contaminate the kaon/proton $\mathrm{d}E/\mathrm{d}x$ curve. For pions a p_{T}-independent cut of $\pm 3\sigma$ was made in the TPC. For the TDR study p_{T}-dependent cuts were tuned and applied on variables including the p_{T} of the decay products, $\cos\theta_{pointing}$, the decay length, $Dist_{12}$, σ_{vert} and the DCA of the decay products. Figure 6.5 shows the distribution of these variables in the p_{T} interval $6 < p_{\mathrm{T}} < 8$ GeV/c. Note that on tracks with $p_{\mathrm{T}} > 2$ GeV/c a cut on the impact parameter $|d_0| > 25$ μm is made, which explains the 'jump' in the $p_{\mathrm{T}}^{p,K,\pi}$ signal and background distributions at 2 GeV/c. Table 6.2 details the cuts that were made in each p_{T} interval. After these selection criteria were applied, the signal and background were scaled to account for differences between the MC simulation and the expected dataset. The signal and background were scaled to $\mathcal{L}_{int} = 10$ nb^{-1} corresponding to 1.6×10^{10} events in the centrality class 0-20%. The signal yield was also scaled to the expected $\mathrm{d}N/\mathrm{d}y = 1.4 \times BR$, and scaled to the R_{AA} for strange Λ baryons, where $R_{\mathrm{AA}} =:$ 1 for $p_{\mathrm{T}} < 4$ GeV/c, 0.7 for $4 < p_{\mathrm{T}} < 5$ GeV/c, 0.5 for $5 < p_{\mathrm{T}} < 6$ GeV/c and 0.3 for $p_{\mathrm{T}} > 6$ GeV/c.

The analysis with rectangular cuts was redone as a reference for the BDT analysis, and also to verify that this reference is valid with respect to the studies that have already been carried out. The significance and S/B qualitatively agree with the that achieved in figure 6.3.

The BDT analysis is done as a direct comparison to the standard analysis. For this reason, all selection *prior* to the BDT (PID, track selection) is done in the same way. The variables used as input to the BDT are the same as those described in table 6.2, and no selection is made on these variables prior to the BDT. For the BDT stage, the signal sample corresponds to $6000 - 12,000$ prompt Λ_{c} candidates depending on the p_{T} interval, and for the background sample around $300,000$ background Λ_{c} candidates from the HIJING event are used (ensuring no candidates are used from the injected charm signal), sampled from 6 runs. The signal and background samples are split in half, with one half used to train the BDT (training sample), and the other half used to test the BDT (test sample). The BDT parameters were varied one-by-one to determine the best configuration in the same procedure as the analysis on data (see section 4.4.3), and the optimal parameters were determined to be the same as those in table 4.3, with the exception of the maximum depth of the tree which was increased from 2 to 3. Figure 6.6 shows the BDT response in 2 p_{T} intervals.

The ROC curve (as a reminder - the signal efficiency vs. background rejection) corresponding to these trained BDTs is shown in figure 6.7. The signal efficiency and

	Λ_c p_T interval (GeV/c)						
variable	2-3	3-4	4-5	5-6	6-8	8-10	10-20
p_T proton (GeV/c)	0.5	0.9	1.0	1.0	1.1	1.2	1.4
p_T kaon (GeV/c)	0.9	0.8	1.0	1.1	1.3	1.7	2.3
p_T pion (GeV/c)	0.75	0.8	0.8	0.9	1.0	0.8	1.1
Decay length (μm)	110	120	140	130	140	160	170
DCA (μm)	50	30	35	30	25	20	18
σ_{vertex} (μm)	60	50	40	40	30	30	20
$Dist_{12}$ (μm)	100	150	150	150	180	180	200
$\cos\theta_{pointing}$	0.97	0.99	0.99	0.99	0.99	0.99	0.99

Table 6.2: Optimised p_T-dependent cuts on variables used in this study [246].

background rejection corresponding to the cuts shown in table 6.2 is also shown in this figure. It can be seen that the ROC curve of the BDT analysis is higher than the standard cut point, meaning better performance can be achieved using BDTs. To assess the improvement that one could achieve, two lines of attack are followed:

- The cut on the BDT response is chosen to **improve the background rejec-**

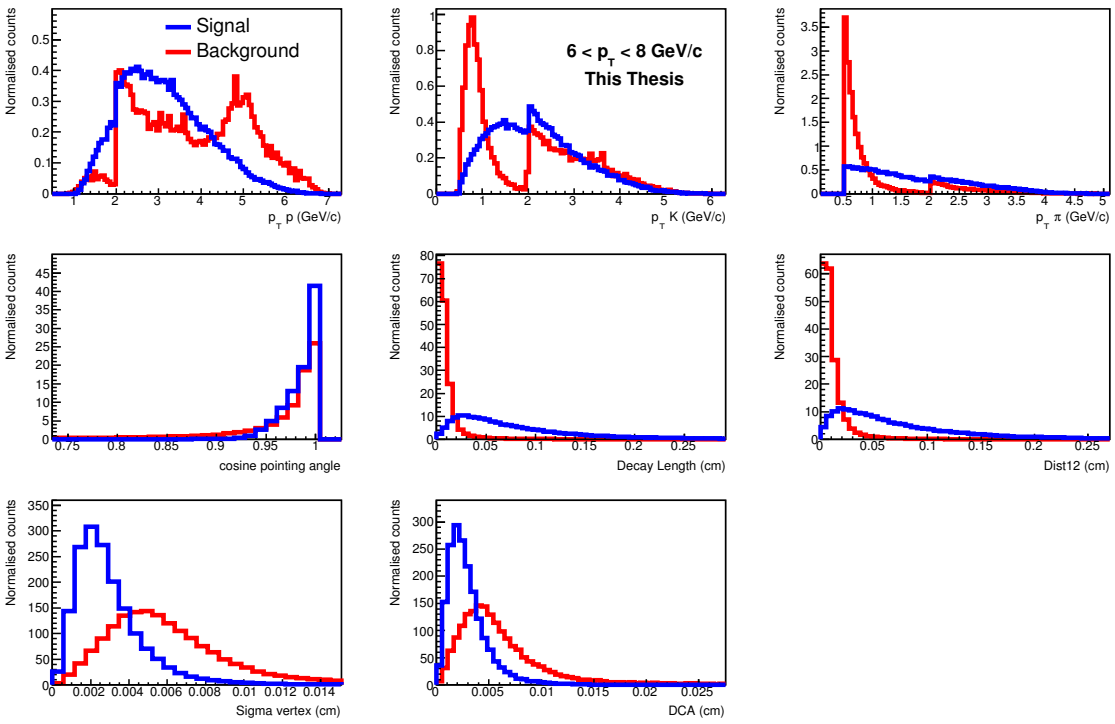

Figure 6.5: The signal and background distributions for the input variables used in the BDT, in the p_T interval $6 < p_T < 8$ GeV/c.

tion by cutting at the **same signal efficiency**.

- The cut on the BDT response is chosen to **improve the signal efficiency** by cutting at the **same background rejection**.

These two scenarios are illustrated in figure 6.7, where the signal efficiency and background rejection corresponding to the rectangular cuts is marked by the star, and the two arrows depict the scenario where the background rejection is improved, and the scenario where the signal efficiency is improved. The points where the arrows intersect the ROC curve for the trained BDT represent the new and improved signal efficiency/background rejection.

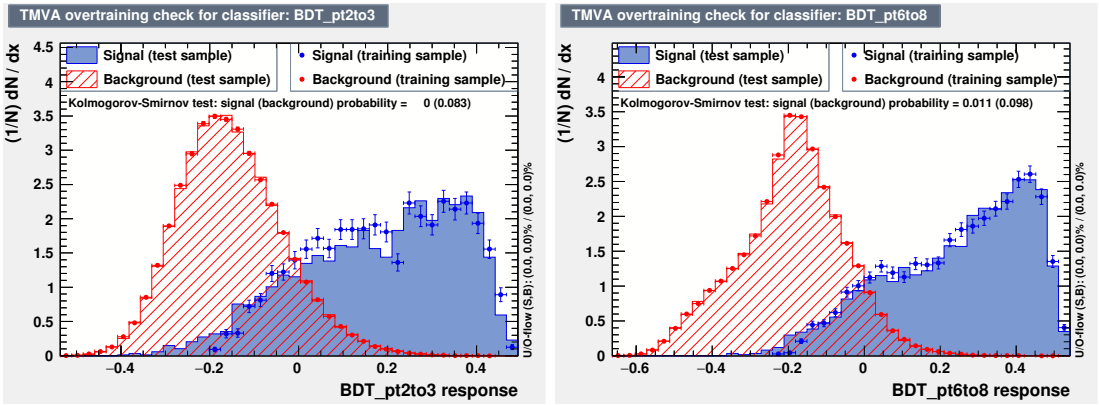

Figure 6.6: The BDT response for the signal (blue) and background (red), for the p_{T} intervals $2 < p_{\mathrm{T}} < 3$ GeV/c (left) and $6 < p_{\mathrm{T}} < 8$ GeV/c (right). The training sample is shown as markers, and the test sample is shown as the filled area.

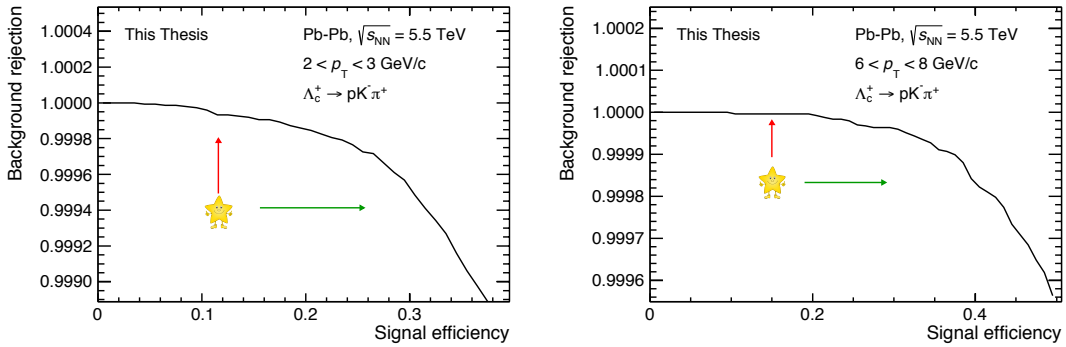

Figure 6.7: The ROC curve in the p_{T} intervals $2 < p_{\mathrm{T}} < 3$ GeV/c (left) and $6 < p_{\mathrm{T}} < 8$ GeV/c (right). The signal efficiency and background rejection achieved for the rectangular cuts in the same p_{T} interval is shown by the star. The two scenarios to judge the improvement using BDTs with respect to the standard cuts as described in the text are shown by the arrows.

The significance can simply be scaled from the significance achieved with the rectangular cuts, to match the improvement in the signal efficiency or background rejection, evaluating these values from the test sample. Alternatively, the trained BDTs can be applied on the full dataset. The latter option has the benefit of increasing the statistics of the background sample - since the background rejection is already very high with only one background candidate in $\sim$ 100,000 surviving the cut on the BDT response, the number of background candidates used to evaluate the new background rejection is very low, and thus the statistical uncertainty is large. Both options have been tried and agree well within statistical uncertainties.

The significance and signal-over-background are shown in figure 6.8 for the case where the background rejection has been improved with respect to the rectangular cut-based method, and where the trained BDTs are applied on the full dataset. The statistical uncertainties are determined for the signal and background from the number of signal and background candidates remaining after the selection, and the uncertainties for the significance and signal-over-background are obtained as detailed in appendix A.2. A significant improvement in the signal-over-background up to a factor $30-40$ can be obtained with BDTs, cutting on the BDT response at the same signal efficiency as the rectangular cut set. In particular at low p_{T}, using the rectangular cuts a significance of around 6 is achieved, which would correspond to a statistical uncertainty of 17%. This is improved by a factor of about 5 by using BDTs increasing the significance to around 30, which corresponds to a statistical uncertainty of 3%.

Figure 6.9 shows the significance and signal-over-background for the alternative case where the signal efficiency has been improved with respect to the rectangular cut-based method. Here a significant improvement is also seen, where keeping more signal for the same level of background rejection allows for an improvement in the significance of a factor of around 2-3 in most p_{T} bins. While this case results in a smaller statistical improvement than the case where the background rejection is improved, the increased signal efficiency may help to reduce other sources of uncertainty - for example, it can be expected that the systematic uncertainty due to the cut on the BDT response would be reduced given the higher signal efficiency.

The improvement in the signal extraction for the case where the background rejection is improved can be seen in figure 6.10, where the invariant mass distributions from the topological analysis is compared to the invariant mass distributions obtained from the BDT analysis. Here the signal and background distributions have been fit with a Gaussian and second order polynomial, respectively, the signal and background distributions are recalculated from the fit, scaling according to the expected number

of events, the expected Λ_c cross section and expected R_{AA} (see above), and the distributions have been re-fit using the same procedure as in data (see section 4.5).

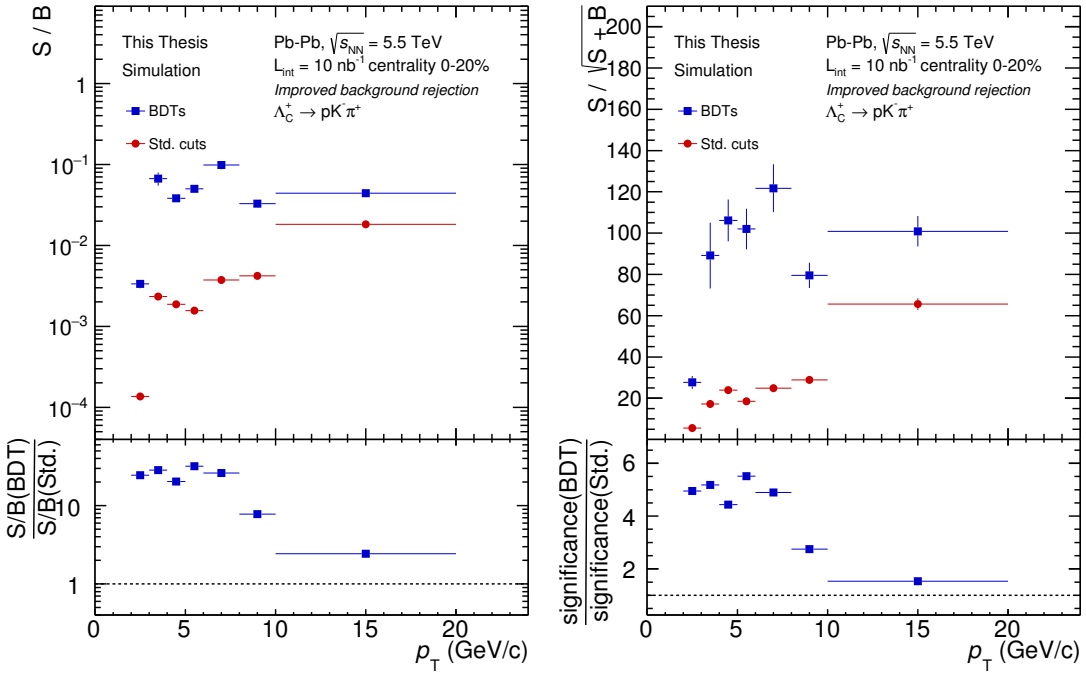

Figure 6.8: Left: The signal-over-background achieved using rectangular cuts, and using BDTs where the background rejection is improved with respect to the rectangular cut method. Right: the corresponding significance.

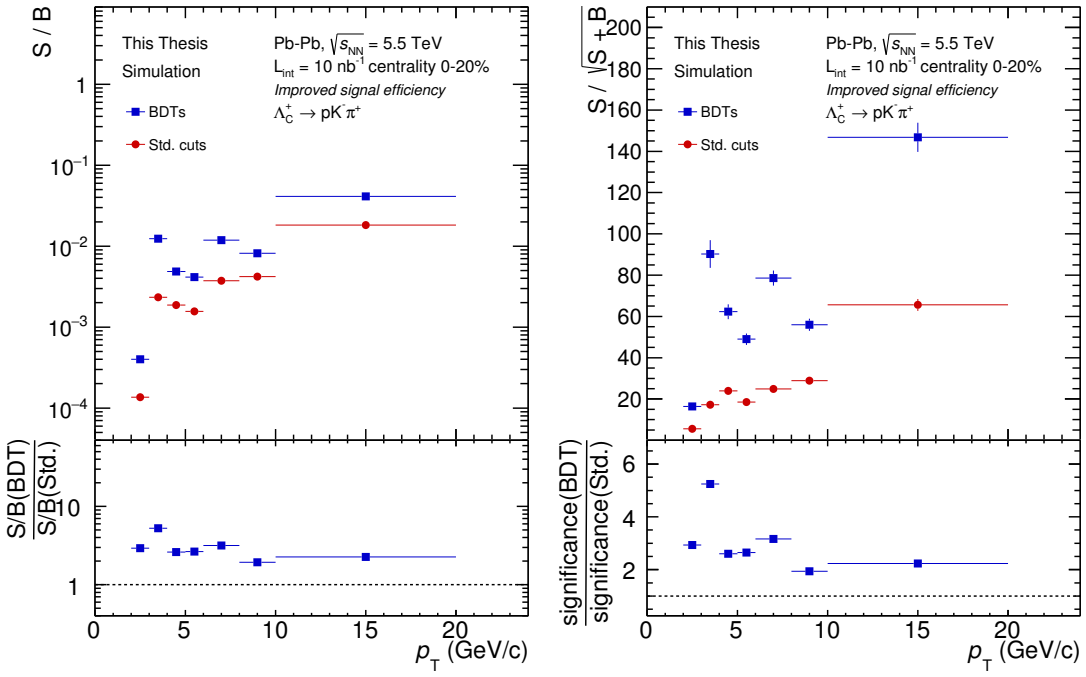

Figure 6.9: Left: The signal-over-background achieved using rectangular cuts, and using BDTs where the background rejection is improved with respect to the rectangular cut method. Right: the corresponding significance.

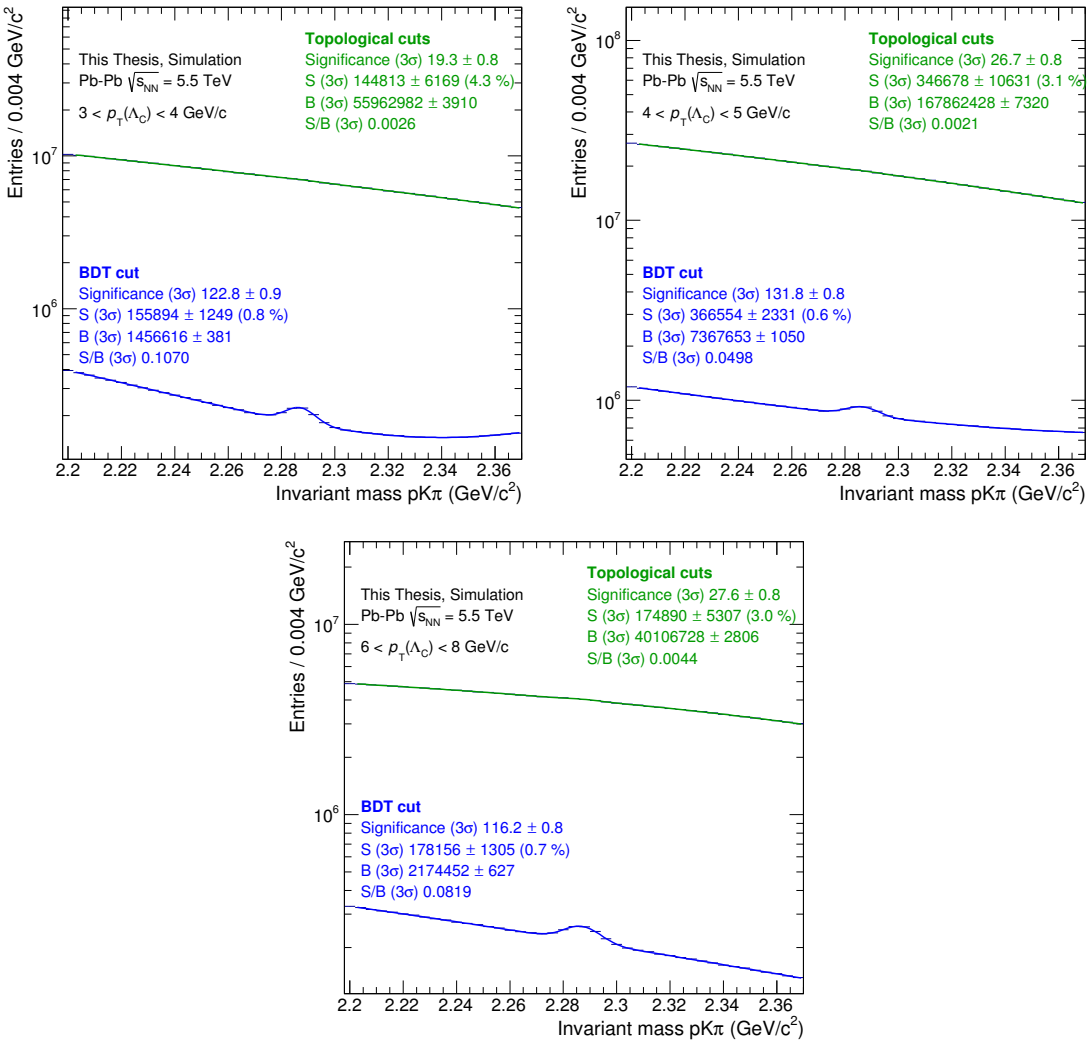

Figure 6.10: Example invariant mass distributions of prompt Λ_c signal plus background Λ_c candidates, comparing optimised 'rectangular' topological selections with selection using Boosted Decision Trees. The signal and background is scaled to the expected number of events, and the signal is scaled to the estimated Λ_c cross section and R_{AA} in minimum-bias Pb–Pb collisions. The width of the peak is fixed to that in simulation, and the BDT cut is tuned to reproduce the same signal efficiency as obtained with topological cuts.

So far it has been demonstrated that a considerable improvement in the Λ_c selection can be achieved over the rectangular cuts optimised for previous studies. The level of improvement quoted is the level of improvement relative to these cut sets. This improvement may therefore be overestimated if the cut set that has been chosen is not optimal. Some work has therefore been done to try to improve the rectangular cuts, using cut optimisation techniques that are also provided by TMVA.

The simulated annealing algorithm [247, 248] aims to solve minimisation problems with several minima. It is inspired by the process of annealing in condensed matter physics, where atoms in a slowly cooling metal will move towards the state of lowest energy. This principle is converted into an algorithm which achieves slow convergence of an optimisation problem with multiple solutions, where the solutions are cut sets with maximised significance. This algorithm achieves good performance in this analysis.

A comparison of the ROC curves for the BDT and for the cuts optimised using simulated annealing is shown in figures 6.11 and 6.12. These figures also show the relative improvement in the background efficiency, calculated as 1 - (background rejection), as well as the improvement in the significance, under the scenario that the background rejection is improved for the same signal efficiency. The improvement in the significance is calculated assuming that

$$S << B \quad \Rightarrow \quad \text{significance} = \frac{S}{\sqrt{B}} \tag{6.1}$$

so the ratio of the significances:

$$\frac{\text{significance(BDT)}}{\text{significance(cuts)}} = \sqrt{\frac{\epsilon_B^{cuts}}{\epsilon_B^{BDT}}} \tag{6.2}$$

where ϵ_B^{cuts} is the background efficiency with the optimised rectangular cuts and ϵ_B^{BDT} is the background efficiency with the BDT.

A significant improvement with respect to the rectangular cuts can still be achieved, with at least a factor of 10 improvement in the background rejection obtainable in most p_{T} intervals. This factor increases at mid-p_{T} with a factor of ~ 30 achievable in the $5 < p_{\mathrm{T}} < 6$ GeV/c. This improvement in the background rejection corresponds to an improvement in the significance of around a factor 3 for a background rejection factor of 10, and around a factor 5 for a background rejection factor of around 30. In general, the improvement factor is larger at smaller signal efficiencies, though in the high and low p_{T} intervals the improvement seems to peak at a given signal efficiency. This cut optimisation algorithm appears to choose cuts which give either similar or

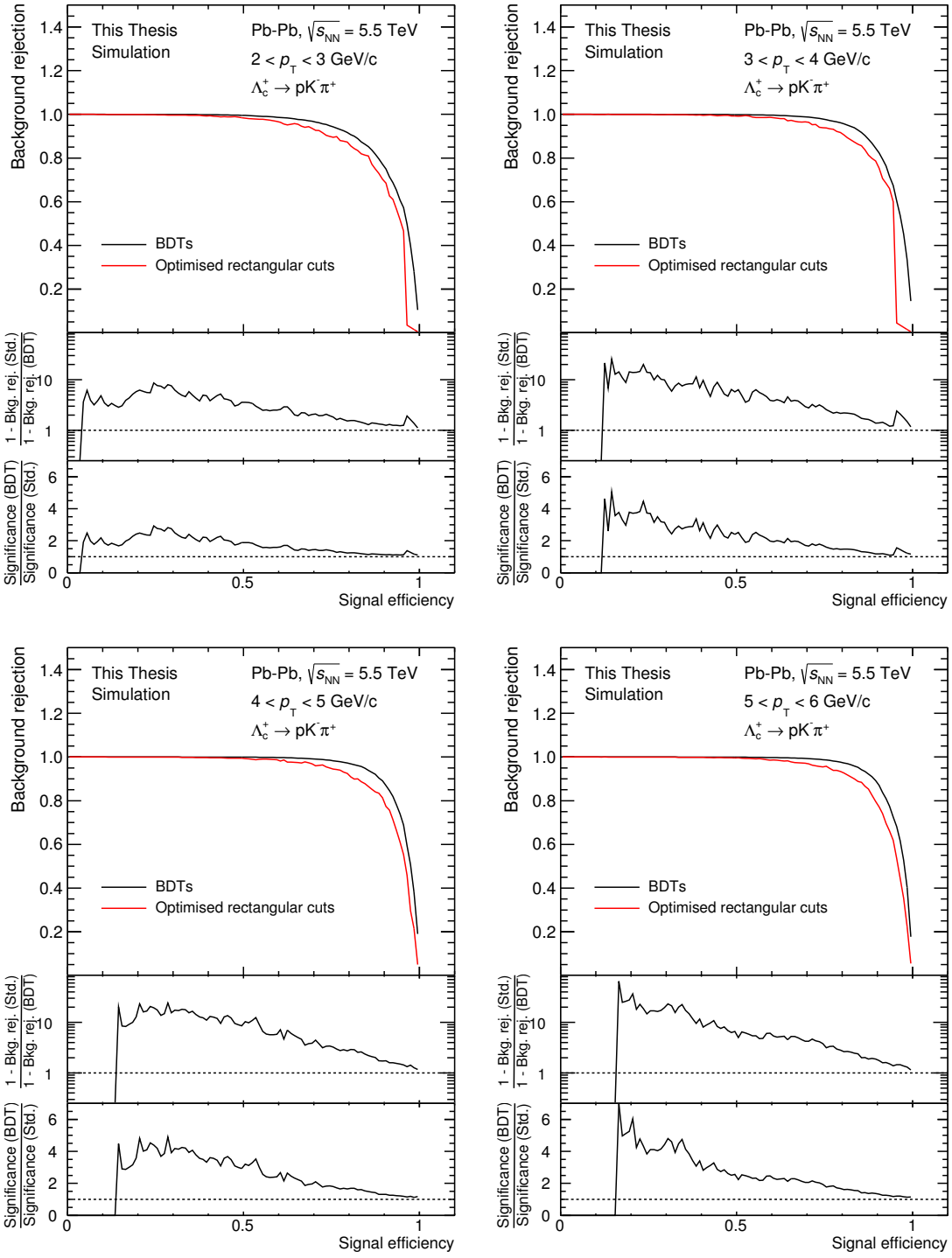

Figure 6.11: The ROC curves of the BDT method compared to the ROC curve for optimised rectangular cuts, in different p_{T} intervals from $2 < p_{\mathrm{T}} < 6$ GeV/c. The middle panels show the ratio of the background efficiency (defined as 1 - background rejection) as a function of signal efficiency. The bottom panels show the corresponding improvement in significance that this improvement in background rejection would bring.

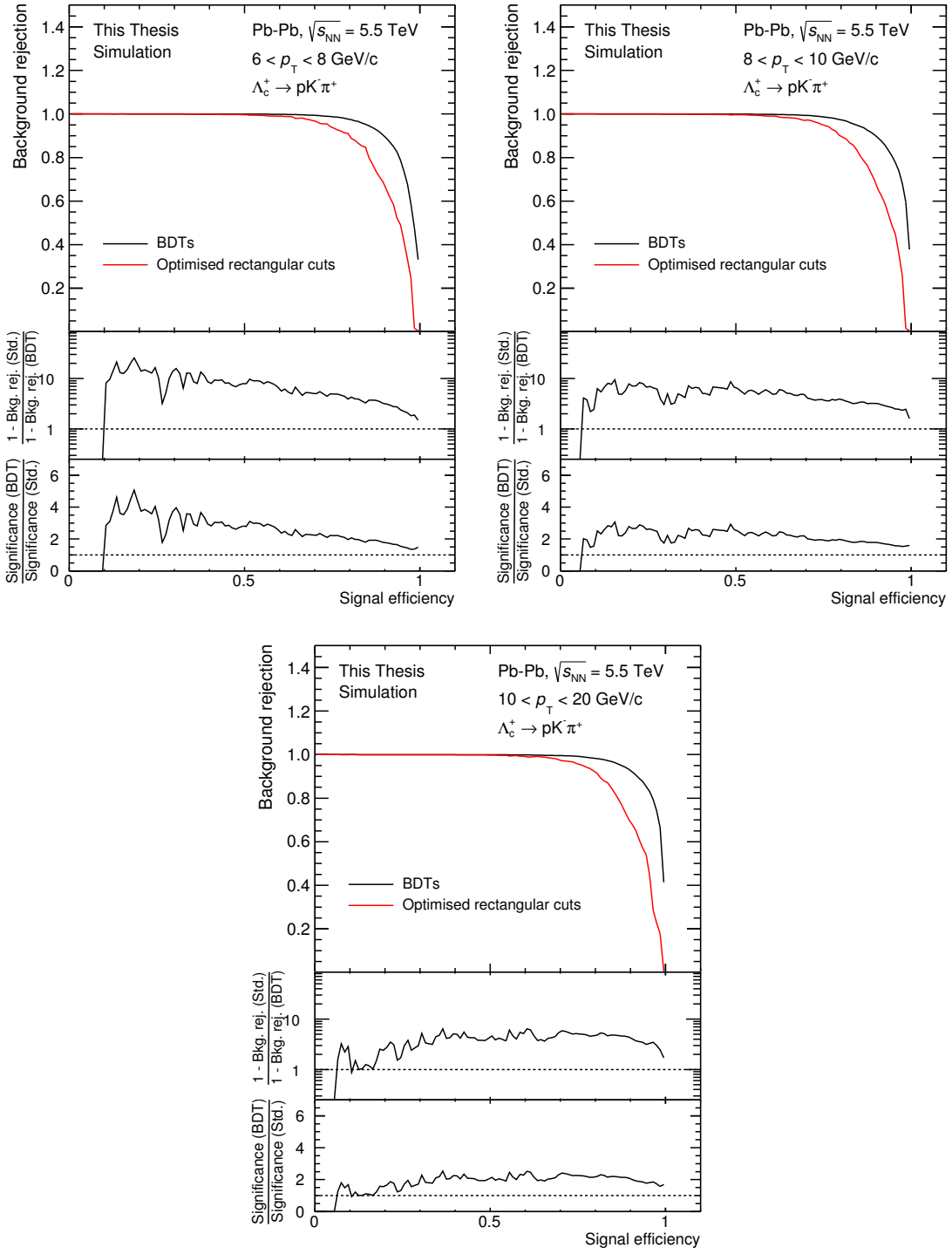

Figure 6.12: The ROC curves of the BDT method compared to the ROC curve for optimised rectangular cuts, in different p_T intervals from $6 < p_T < 20$ GeV/c. The middle panels show the ratio of the background efficiency (defined as 1 - background rejection) as a function of signal efficiency. The bottom panels show the corresponding improvement in significance that this improvement in background rejection would bring.

slightly better performance than the optimised cuts given in table 6.2 for a given signal efficiency cutting point. It is clear however that the BDT can still significantly outperform a standard rectangular cut-based method. It is also interesting to note that the BDT seems give the largest improvement in the mid-p_{T} region, with the largest factor of improvement achieved in the p_{T} range $4 < p_{\mathrm{T}} < 8$ GeV/c.

Chapter 7

Conclusions and Perspectives

This thesis has presented a measurement of the p_{T} dependence of the Λ_{c}^{+} baryon differential cross section in pp collisions at a centre-of-mass energy $\sqrt{s} = 7$ TeV and in p–Pb collisions at a centre-of-mass energy per nucleon-nucleon collision $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV in the range $2 < p_{\mathrm{T}} < 12$ GeV/c. A novel multivariate approach was developed, which combined discriminating variables related to the Λ_{c}^{+} decay in an optimal way. This approach reduced the statistical uncertainty of the measurement, and extended the p_{T} reach in pp collisions. A study to assess how multivariate analysis techniques could be used in the measurement of the Λ_{c}^{+} baryon in Pb–Pb collisions has also been presented, which found that a significant improvement in the signal extraction and resulting statistical uncertainties can be achieved with such a technique.

This work has found that in pp collisions at $\sqrt{s} = 7$ TeV the measured p_{T}-differential cross section of the Λ_{c}^{+} baryon is underestimated by theoretical predictions. The ratio of the p_{T}-differential cross section of the Λ_{c}^{+} baryon to that of the D^{0} meson, $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$, is also underestimated by theoretical predictions, though a model including colour reconnection [115, 116] brings the prediction for the $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratio closer to the data. The $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratio in pp collisions at $\sqrt{s} = 7$ TeV is measured to be consistent with that in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV within experimental uncertainties, suggesting common hadronisation mechanisms are at play in both collision systems. Finally, the nuclear modification factor R_{pPb} is measured to be consistent with unity and consistent with the averaged D meson R_{pPb}, indicating no significant modification of the Λ_{c}^{+} yield in p–Pb collisions with respect to pp collisions.

For the publication which is currently in preparation, the measurement in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV presented in this work is combined with the measurement in the decay channel $\Lambda_{\mathrm{c}}^{+} \rightarrow \mathrm{pK}_{\mathrm{s}}^{0}$. The measurements include a rectangular cut-based approach and a multivariate approach in both decay channels, and the

combination of the measurements makes a reduction in the uncertainty possible. The measurement in pp collisions at $\sqrt{s} = 7$ TeV in the publication is a combination of the decay channels $\Lambda_c^+ \to pK^-\pi^+$, $\Lambda_c^+ \to pK_s^0$ and $\Lambda_c^+ \to \Lambda e^+\nu_e$, and the analysis in pp collisions presented in this work provides a cross check for this measurement.

These results set the stage for future measurements of the Λ_c^+ baryon, which will extend the work presented in this thesis. Run 2 at the LHC begun in 2015 and ALICE has collected pp collision data at 5 TeV and 13 TeV. A measurement in pp collisions at 5 TeV would provide a reference for a p–Pb and possible Pb–Pb measurement which would remove the systematic uncertainty associated with the energy scaling procedure presented in this work, and while the current collected data sample at 5 TeV is modest ($\sim$ 100 million minimum bias collisions), a long 5 TeV run is foreseen for the end of 2017, which will make a higher precision measurement at this energy possible. The p–Pb data at $\sqrt{s_{NN}} = 5.02$ TeV collected in Run 2 is larger than the Run 1 dataset by a factor of ~ 6 ($\sim$ 600 million minimum bias collisions), and will allow to reduce the statistical uncertainty of the measurement presented in this thesis, which could further constrain models including initial- and final-state effects in p–Pb collisions. The larger dataset may also allow for a more differential investigation into Λ_c^+ production as a function of the event multiplicity, which may help to constrain effects that modify the Λ_c^+ yield in high multiplicity p–Pb collisions. The upgrade of ALICE's Inner Tracking System will make a measurement of the Λ_c^+ baryon possible in Pb–Pb collisions during Run 3 at the LHC, and will also allow for precision measurements of the Λ_c^+ baryon to be made in pp and p–Pb collisions.

The use of multivariate analysis methods has a growing importance within the high-energy physics community. This thesis has demonstrated that a multivariate approach can aid in reducing the uncertainty in the measurement of the production of the Λ_c^+ baryon, and these techniques could be extended to the measurement of other physics signals. The measurement of charmed mesons could be improved - for example, the D_s^+ meson has a proper decay length slightly larger than the Λ_c^+ baryon, and the measurement in Pb–Pb collisions currently suffers from large uncertainties [249]. More exotic heavy-flavour particles (B mesons, Λ_b^0 and $\Xi_c^{0,+}$ baryons), providing a probe to further constrain energy loss and hadronisation models in Pb–Pb collisions, may also be reachable via their hadronic decay channels, and their decay topology may allow for topological and kinematic properties of the decay to be distinguished from background with the use of a multivariate algorithm. The application of deep learning techniques also offers the capability of recognising complex correlations in data. Recent applications of deep learning at the LHC include jet

flavour [250, 251] and substructure [252–254] classification, exotic particle (Higgs, Supersymmetric particles) searches [255], as well as studies into using deep learning for event reconstruction [256], which show promising results. The use of these techniques with ALICE in the future may help offer new insights into the properties of the Quark-Gluon Plasma.

Appendices

Appendix A

Uncertainties

A.1 Correlation between uncertainty sources

The final p_{T}-differential cross section measurement of the Λ_{c} baryon in a given collision system is compared to other measured quantities. The ratio of the Λ_{c} cross section in p–Pb collisions relative to pp collisions quantifies the nuclear modification in p–Pb collisions, and the baryon-to-meson ration $\Lambda_{\mathrm{c}}/\mathrm{D}^0$ gives insight into the relative production of baryons and mesons. The Λ_{c} cross section is also integrated over p_{T} to obtain a more inclusive measurement less dependent on p_{T}. The propagation of the statistical and systematic uncertainties in each p_{T} interval is described in this section.

For additive quantities $y = A + B$ (integrated cross section), the absolute uncorrelated statistical and systematic uncertainties sum quadratically:

$$\Delta y = \sqrt{\Delta A^2 + \Delta B^2}. \tag{A.1}$$

The systematic uncertainty sources that are assumed to be fully correlated sum linearly:

$$\Delta y = \Delta A + \Delta B. \tag{A.2}$$

In the ratio of two quantities $y = A/B$ ($\Lambda_{\mathrm{c}}/\mathrm{D}^0$, R_{pPb} measurements) the relative uncorrelated statistical and systematic uncertainties sum quadratically:

$$\frac{\Delta y}{y} = \sqrt{\left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2}. \tag{A.3}$$

The systematic uncertainty sources that are assumed to be fully correlated partially cancel depending on the relative uncertainty of A and B:

$$\frac{\Delta y}{y} = |\frac{\Delta A}{A} - \frac{\Delta B}{B}|. \tag{A.4}$$

A.2 Significance and signal-over-background

After the selection criteria applied on Λ_c candidates, a given number of signal S and background B candidates remain with statistical uncertainties $\Delta S = \frac{1}{\sqrt{S}}$ and $\Delta B = \frac{1}{\sqrt{B}}$. The upgrade studies use the significance $\frac{S}{\sqrt{S+B}}$ and signal-to-background ratio S/B as a 'figure of merit' of the selection criteria applied. The uncertainty of a function $f(x, y, ...)$ which is a function of one or more variables $x, y, ...$ is given by

$$\Delta f^2 = \left(\frac{\mathrm{d}f}{\mathrm{d}x}\Delta x\right)^2 + \left(\frac{\mathrm{d}f}{\mathrm{d}y}\Delta y\right)^2 + ... \tag{A.5}$$

So it follows that the uncertainty on the S/B is given as

$$\Delta(S/B)^2 = \left(\frac{1}{B}\Delta S\right)^2 + \left(-\frac{S}{B^2}\Delta B\right)^2 \tag{A.6}$$

and it follows that the relative uncertainty is the quadratic sum of the relative uncertainties of S and B

$$\frac{\Delta(S/B)}{(S/B)} = \sqrt{\left(\frac{\Delta S}{S}\right)^2 + \left(\frac{\Delta B}{B}\right)^2}. \tag{A.7}$$

The uncertainty on the significance is determined assuming $S << B$, such that the significance can be approximated as $\frac{S}{\sqrt{B}}$. The uncertainty is then given by

$$\Delta sgnf. = \sqrt{\frac{\Delta S^2}{B} + \frac{S^2\Delta B^2}{4B^3}}. \tag{A.8}$$

Appendix B

Cross sections

B.1 p_{T}-differential cross section in pp and p–Pb collisions

p_{T} range (GeV/c)	cross section (μb(GeV/c)$^{-1}$)					
$2 < p_{\mathrm{T}} < 4$	45.7	$\pm$6.9	$\pm$8.0	$^{+0.2}_{-1.0}$	$\pm$2.3	$\pm$1.6
$4 < p_{\mathrm{T}} < 6$	10.9	$\pm$1.4	$\pm$1.9	$^{+0.1}_{-0.3}$	$\pm$0.6	$\pm$0.4
$6 < p_{\mathrm{T}} < 8$	1.75	$\pm$0.44	$\pm$0.38	$^{+0.020}_{-0.12}$	$\pm$0.09	$\pm$0.06
$8 < p_{\mathrm{T}} < 12$	0.398	$\pm$0.113	$\pm$0.118	$^{+0.006}_{-0.040}$	$\pm$0.020	$\pm$0.014

Table B.1: Λ_{c}^{+} baryon p_{T}-differential cross section in pp collisions at $\sqrt{s} = 7$ TeV. The uncertainties given are, from left to right; statistical, systematic, feed-down, branching ratio and luminosity contributions.

p_{T} range (GeV/c)	cross section (μb(GeV/c)$^{-1}$)					
$2 < p_{\mathrm{T}} < 4$	9510	$\pm$1420	$\pm$1340	$^{+81}_{-348}$	$\pm$485	$\pm$352
$4 < p_{\mathrm{T}} < 6$	1570	$\pm$193	$\pm$242	$^{+23}_{-107}$	$\pm$80	$\pm$58
$6 < p_{\mathrm{T}} < 8$	300	$\pm$56	$\pm$43	$^{+7}_{-32}$	$\pm$15	$\pm$11
$8 < p_{\mathrm{T}} < 12$	50.1	$\pm$8.8	$\pm$5.3	$^{+1.8}_{-8.8}$	$\pm$2.6	$\pm$1.9

Table B.2: Λ_{c}^{+} baryon p_{T}-differential cross section in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. The uncertainties given are, from left to right; statistical, systematic, feed-down, branching ratio and luminosity contributions.

Bibliography

[1] ALICE collaboration, J. Norman, *Charmed meson and baryon measurements in pp and p-Pb collisions with ALICE at the LHC*, in *17th International Conference on Strangeness in Quark Matter*, 2017. 1711.02393.

[2] ALICE collaboration, J. Norman, *Open heavy-flavour measurements in pp and p–Pb collisions with ALICE at the LHC*, in *55th International Winter Meeting on Nuclear Physics (Bormio 2017) Bormio, Italy*, 2017. 1706.05223.

[3] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry*, .

[4] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, *Phys. Lett.* **8** (1964) 214–215.

[5] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 2*, . Version 1 is CERN preprint 8182/TH.401, Jan. 17, 1964.

[6] E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, G. Miller, L. W. Mo et al., *High-Energy Inelastic $e-p$ Scattering at 6 and 10* deg, *Phys. Rev. Lett.* **23** (Oct, 1969) 930–934.

[7] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. DeStaebler et al., *Observed behavior of highly inelastic electron-proton scattering*, *Phys. Rev. Lett.* **23** (Oct, 1969) 935–939.

[8] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak interactions with lepton-hadron symmetry*, *Phys. Rev. D* **2** (Oct, 1970) 1285–1292.

[9] J. J. Aubert, U. Becker, P. J. Biggs, J. Burger, M. Chen, G. Everhart et al., *Experimental observation of a heavy particle j*, *Phys. Rev. Lett.* **33** (Dec, 1974) 1404–1406.

[10] J. E. Augustin, A. M. Boyarski, M. Breidenbach, F. Bulos, J. T. Dakin, G. J. Feldman et al., *Discovery of a narrow resonance in e^+e^- annihilation*, *Phys. Rev. Lett.* **33** (Dec, 1974) 1406–1408.

[11] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, *Phys. Rev. Lett.* **39** (1977) 252–255.

[12] CDF collaboration, F. Abe et al., *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626–2631, [hep-ex/9503002].

[13] Particle Data Group collaboration, C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.

[14] K. A. Brueckner, *Meson-Nucleon Scattering and Nucleon Isobars*, *Phys. Rev.* **86** (1952) 106–109.

[15] D. J. Gross, *The discovery of asymptotic freedom and the emergence of qcd*, *Proceedings of the National Academy of Sciences of the United States of America* **102** (2005) 9099–9108, [http://www.pnas.org/content/102/26/9099.full.pdf].

[16] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346.

[17] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. i*, *Phys. Rev. D* **8** (Nov, 1973) 3633–3652.

[18] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. ii*, *Phys. Rev. D* **9** (Feb, 1974) 980–993.

[19] H. D. Politzer, *Reliable perturbative results for strong interactions?*, *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349.

[20] Nobel Media AB 2014, "The Nobel Prize in Physics 2004." http://www.nobelprize.org/nobel_prizes/physics/laureates/2004/. [Online; accessed 15-March-2017].

[21] R. Hagedorn, *Statistical thermodynamics of strong interactions at high-energies*, *Nuovo Cim. Suppl.* **3** (1965) 147–186.

[22] N. Cabibbo and G. Parisi, *Exponential Hadronic Spectrum and Quark Liberation*, *Phys. Lett.* **B59** (1975) 67–69.

[23] A. Chodos, R. L. Jaffe, K. Johnson, C. B. Thorn and V. F. Weisskopf, *New extended model of hadrons*, *Phys. Rev. D* **9** (Jun, 1974) 3471–3495.

[24] H. Satz, *The Thermodynamics of Quarks and Gluons*, *Lect. Notes Phys.* **785** (2010) 1–21, [0803.1611].

[25] T. DeGrand, R. L. Jaffe, K. Johnson and J. Kiskis, *Masses and other parameters of the light hadrons*, *Phys. Rev. D* **12** (Oct, 1975) 2060–2076.

[26] H. B. Nielsen and M. Ninomiya, *Color Ferromagnetic Vacuum States in QCD and Two Loop Energy Densities Bounds on MIT Bag Constant in Relation to QCD* Λ, *Nucl. Phys.* **B169** (1980) 309–328.

[27] F. Karsch, *Lattice results on QCD thermodynamics*, *Nucl. Phys.* **A698** (2002) 199–208, [hep-ph/0103314].

[28] F. Karsch, *Lattice QCD at high temperature and density*, *Lect. Notes Phys.* **583** (2002) 209–249, [hep-lat/0106019].

[29] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz and K. K. Szabo, *The Order of the quantum chromodynamics transition predicted by the standard model of particle physics*, *Nature* **443** (2006) 675–678, [hep-lat/0611014].

[30] Y. Nambu and G. Jona-Lasinio, *Dynamical model of elementary particles based on an analogy with superconductivity. ii*, *Phys. Rev.* **124** (Oct, 1961) 246–254.

[31] B.-J. Schaefer and J. Wambach, *The Phase diagram of the quark meson model*, *Nucl. Phys.* **A757** (2005) 479–492, [nucl-th/0403039].

[32] C. Ratti, M. A. Thaler and W. Weise, *Phases of QCD: Lattice thermodynamics and a field theoretical model*, *Phys. Rev.* **D73** (2006) 014019, [hep-ph/0506234].

[33] B.-J. Schaefer, J. M. Pawlowski and J. Wambach, *Phase structure of the polyakov-quark-meson model*, *Phys. Rev. D* **76** (Oct, 2007) 074023.

[34] X. Luo and N. Xu, *Search for the QCD Critical Point with Fluctuations of Conserved Quantities in Relativistic Heavy-Ion Collisions at RHIC : An Overview*, 1701.02105.

[35] M. G. Alford, A. Schmitt, K. Rajagopal and T. Schfer, *Color superconductivity in dense quark matter*, *Rev. Mod. Phys.* **80** (2008) 1455–1515, [0709.4635].

[36] R. S. Bhalerao, *Relativistic heavy-ion collisions*, . Comments: Updated version of the lectures given at the First Asia-Europe-Pacific School of High-Energy Physics, Fukuoka, Japan, 14-27 October 2012. Published as a CERN Yellow Report (CERN-2014-001) and KEK report (KEK-Proceedings-2013-8), K. Kawagoe and M. Mulders (eds.), 2014, p. 219. Total 21 pages.

[37] A. Tawfik and A. G. Shalaby, *Balance Function in High-Energy Collisions*, *Adv. High Energy Phys.* **2015** (2015) 186812.

[38] M. L. Miller, K. Reygers, S. J. Sanders and P. Steinberg, *Glauber modeling in high energy nuclear collisions*, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 205–243, [nucl-ex/0701025].

[39] R. D. Woods and D. S. Saxon, *Diffuse surface optical model for nucleon-nuclei scattering*, *Phys. Rev.* **95** (Jul, 1954) 577–578.

[40] ALICE collaboration, J. Adam et al., *Centrality dependence of particle production in p-Pb collisions at* $\sqrt{s_{\rm NN}}$ *= 5.02 TeV*, *Phys. Rev.* **C91** (2015) 064905, [1412.6828].

[41] ALICE collaboration, B. Abelev et al., *Pseudorapidity density of charged particles in p + Pb collisions at* $\sqrt{s_{NN}} = 5.02$ *TeV*, *Phys. Rev. Lett.* **110** (2013) 032301, [1210.3615].

[42] ALICE collaboration, J. Adam et al., *Centrality dependence of the charged-particle multiplicity density at midrapidity in Pb-Pb collisions at* $\sqrt{s_{\rm NN}}$ *= 5.02 TeV*, *Phys. Rev. Lett.* **116** (2016) 222302, [1512.06104].

[43] J. D. Bjorken, *Highly Relativistic Nucleus-Nucleus Collisions: The Central Rapidity Region*, *Phys. Rev.* **D27** (1983) 140–151.

[44] ALICE collaboration, J. Adam et al., *Measurement of transverse energy at midrapidity in Pb-Pb collisions at* $\sqrt{s_{\rm NN}} = 2.76$ *TeV*, *Phys. Rev.* **C94** (2016) 034903, [1603.04775].

[45] PHENIX collaboration, S. S. Adler et al., *Systematic studies of the centrality and* $\sqrt{s_{\rm NN}}$ *dependence of the* $dE_T/d\eta$ *and* $dN_{ch}/d\eta$ *in heavy ion collisions at mid-rapidity*, *Phys. Rev.* **C71** (2005) 034908, [nucl-ex/0409015]. [Erratum: Phys. Rev.C71,049901(2005)].

[46] STAR collaboration, J. Adams et al., *Measurements of transverse energy distributions in Au + Au collisions at $\sqrt{s_{NN}}$ = 200-GeV*, *Phys. Rev.* **C70** (2004) 054907, [nucl-ex/0407003].

[47] NA49 collaboration, T. Alber et al., *Transverse energy production in Pb-208 + Pb collisions at 158-GeV per nucleon*, *Phys. Rev. Lett.* **75** (1995) 3814–3817.

[48] P. Braun-Munzinger, K. Redlich and J. Stachel, *Particle production in heavy ion collisions*, nucl-th/0304013.

[49] J. Stachel, A. Andronic, P. Braun-Munzinger and K. Redlich, *Confronting LHC data with the statistical hadronization model*, *J. Phys. Conf. Ser.* **509** (2014) 012019, [1311.4662].

[50] J. Cleymans, H. Oeschler, K. Redlich and S. Wheaton, *Comparison of chemical freeze-out criteria in heavy-ion collisions*, *Phys. Rev.* **C73** (2006) 034905, [hep-ph/0511094].

[51] ALICE collaboration, J. Adam et al., *Centrality dependence of the nuclear modification factor of charged pions, kaons, and protons in Pb-Pb collisions at $\sqrt{s_{\rm NN}}$ = 2.76 TeV*, *Phys. Rev.* **C93** (2016) 034913, [1506.07287].

[52] ALICE collaboration, B. Abelev et al., *Measurement of charged jet suppression in Pb-Pb collisions at $\sqrt{s_{NN}}$ = 2.76 TeV*, *JHEP* **03** (2014) 013, [1311.0633].

[53] ATLAS collaboration, G. Aad et al., *Measurement of the jet radius and transverse momentum dependence of inclusive jet suppression in lead-lead collisions at $\sqrt{s_{NN}}$= 2.76 TeV with the ATLAS detector*, *Phys. Lett.* **B719** (2013) 220–241, [1208.1967].

[54] ALICE collaboration, B. Abelev et al., *Transverse momentum distribution and nuclear modification factor of charged particles in p-Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV*, *Phys. Rev. Lett.* **110** (2013) 082302, [1210.4520].

[55] ALICE collaboration, B. Abelev et al., *Centrality Dependence of Charged Particle Production at Large Transverse Momentum in Pb–Pb Collisions at $\sqrt{s_{\rm NN}}$ = 2.76 TeV*, *Phys. Lett.* **B720** (2013) 52–62, [1208.2711].

[56] CMS collaboration, S. Chatrchyan et al., *Measurement of isolated photon production in pp and PbPb collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *Phys. Lett.* **B710** (2012) 256–277, [1201.3093].

[57] CMS collaboration, S. Chatrchyan et al., *Study of W boson production in PbPb and pp collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *Phys. Lett.* **B715** (2012) 66–87, [1205.6334].

[58] CMS collaboration, S. Chatrchyan et al., *Study of Z production in PbPb and pp collisions at* $\sqrt{s_{\rm NN}} = 2.76$ *TeV in the dimuon and dielectron decay channels*, *JHEP* **03** (2015) 022, [1410.4825].

[59] PHENIX collaboration, S. S. Adler, S. Afanasiev, C. Aidala, N. N. Ajitanand, Y. Akiba, J. Alexander et al., *Scaling properties of proton and antiproton production in* $\sqrt{s_{\rm NN}} = 200$ *gev au+au collisions*, *Phys. Rev. Lett.* **91** (Oct, 2003) 172301.

[60] STAR collaboration, J. Adams, M. M. Aggarwal, Z. Ahammed, J. Amonett, B. D. Anderson, M. Anderson et al., *Measurements of identified particles at intermediate transverse momentum in the STAR experiment from Au+Au collisions at* $\sqrt{s_{\rm NN}} = 200$ *GeV*, Tech. Rep. nucl-ex/0601042, Jan, 2006.

[61] ALICE collaboration, B. Abelev et al., K^0_S *and* Λ *production in Pb-Pb collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *Phys. Rev. Lett.* **111** (2013) 222301, [1307.5530].

[62] R. J. Fries, V. Greco and P. Sorensen, *Coalescence Models For Hadron Formation From Quark Gluon Plasma*, *Ann. Rev. Nucl. Part. Sci.* **58** (2008) 177–205, [0807.4939].

[63] H. Song and U. W. Heinz, *Multiplicity scaling in ideal and viscous hydrodynamics*, *Phys. Rev.* **C78** (2008) 024902, [0805.1756].

[64] H. Song, S. A. Bass and U. Heinz, *Elliptic flow in 200 A GeV Au+Au collisions and 2.76 A TeV Pb+Pb collisions: insights from viscous hydrodynamics + hadron cascade hybrid model*, *Phys. Rev.* **C83** (2011) 054912, [1103.2380]. [Erratum: Phys. Rev.C87,no.1,019902(2013)].

[65] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko and K. Werner, *EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*, *Phys. Rev.* **C92** (2015) 034906, [1306.0121].

[66] V. Minissale, F. Scardina and V. Greco, *Hadrons from coalescence plus fragmentation in AA collisions at energies available at the BNL Relativistic Heavy Ion Collider to the CERN Large Hadron Collider*, *Phys. Rev.* **C92** (2015) 054904, [1502.06213].

[67] ALICE collaboration, B. B. Abelev et al., $K^*(892)^0$ *and* (1020) *production in Pb-Pb collisions at* $\sqrt{sNN}$ *= 2.76 TeV*, *Phys. Rev.* **C91** (2015) 024609, [1404.0495].

[68] ALICE collaboration, J. Adam et al., $K^*(892)^0$ *and* $\phi(1020)$ *meson production at high transverse momentum in pp and Pb-Pb collisions at* $\sqrt{s_{\rm NN}}$ *= 2.76 TeV*, 1702.00555.

[69] PHENIX collaboration, A. Adare, S. Afanasiev, C. Aidala, N. N. Ajitanand, Y. Akiba, H. Al-Bataineh et al., *Spectra and ratios of identified particles in au+au and d+au collisions at* $\sqrt{s_{NN}} = 200$ *gev*, *Phys. Rev. C* **88** (Aug, 2013) 024906.

[70] ALICE collaboration, B. B. Abelev et al., *Multiplicity Dependence of Pion, Kaon, Proton and Lambda Production in p-Pb Collisions at* $\sqrt{s_{NN}}$ *= 5.02 TeV*, *Phys. Lett.* **B728** (2014) 25–38, [1307.6796].

[71] R. C. Hwa and C. B. Yang, *Final state interaction as the origin of the Cronin effect*, *Phys. Rev. Lett.* **93** (2004) 082302, [nucl-th/0403001].

[72] R. Snellings, *Elliptic Flow: A Brief Review*, *New J. Phys.* **13** (2011) 055008, [1102.3010].

[73] STAR collaboration, K. H. Ackermann et al., *Elliptic flow in Au + Au collisions at* $\sqrt{s_{NN}}$ *= 130 GeV*, *Phys. Rev. Lett.* **86** (2001) 402–407, [nucl-ex/0009011].

[74] ALICE collaboration, J. Adam et al., *Anisotropic flow of charged particles in Pb-Pb collisions at* $\sqrt{s_{\rm NN}} = 5.02$ *TeV*, *Phys. Rev. Lett.* **116** (2016) 132302, [1602.01119].

[75] ALICE collaboration, K. Aamodt et al., *Elliptic flow of charged particles in Pb-Pb collisions at 2.76 TeV*, *Phys. Rev. Lett.* **105** (2010) 252302, [1011.3914].

[76] D. A. Teaney, *Viscous Hydrodynamics and the Quark Gluon Plasma*, in *Quark-gluon plasma 4* (R. C. Hwa and X.-N. Wang, eds.), pp. 207–266. 2010. 0905.2433. DOI.

[77] P. Kovtun, D. T. Son and A. O. Starinets, *Viscosity in strongly interacting quantum field theories from black hole physics*, *Phys. Rev. Lett.* **94** (2005) 111601, [hep-th/0405231].

[78] ATLAS collaboration, G. Aad et al., *Measurement of the azimuthal anisotropy for charged particle production in $\sqrt{s_{NN}} = 2.76$ TeV lead-lead collisions with the ATLAS detector*, *Phys. Rev.* **C86** (2012) 014907, [1203.3087].

[79] T. Hirano, N. van der Kolk and A. Bilandzic, *Hydrodynamics and Flow*, *Lect. Notes Phys.* **785** (2010) 139–178, [0808.2684].

[80] M. Bahr, J. M. Butterworth and M. H. Seymour, *The Underlying Event and the Total Cross Section from Tevatron to the LHC*, *JHEP* **01** (2009) 065, [0806.2949].

[81] ALICE collaboration, B. Abelev et al., *Long-range angular correlations on the near and away side in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, *Phys. Lett.* **B719** (2013) 29–41, [1212.2001].

[82] ATLAS collaboration, G. Aad et al., *Observation of Long-Range Elliptic Azimuthal Anisotropies in $\sqrt{s}$ =13 and 2.76 TeV pp Collisions with the ATLAS Detector*, *Phys. Rev. Lett.* **116** (2016) 172301, [1509.04776].

[83] K. Dusling, W. Li and B. Schenke, *Novel collective phenomena in high-energy protonproton and protonnucleus collisions*, *Int. J. Mod. Phys.* **E25** (2016) 1630002, [1509.07939].

[84] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, [hep-ph/0409313].

[85] M. L. Mangano, P. Nason and G. Ridolfi, *Heavy quark correlations in hadron collisions at next-to-leading order*, *Nucl. Phys.* **B373** (1992) 295–345.

[86] F. Aversa, P. Chiappetta, M. Greco and J. P. Guillet, *QCD Corrections to Parton-Parton Scattering Processes*, *Nucl. Phys.* **B327** (1989) 105.

[87] B. A. Kniehl, G. Kramer, I. Schienbein and H. Spiesberger, *Inclusive D*+- production in p anti-p collisions with massive charm quarks*, *Phys. Rev.* **D71** (2005) 014018, [hep-ph/0410289].

[88] B. A. Kniehl, G. Kramer, I. Schienbein and H. Spiesberger, *Collinear subtractions in hadroproduction of heavy quarks*, *Eur. Phys. J.* **C41** (2005) 199–212, [hep-ph/0502194].

[89] M. Cacciari, M. Greco and P. Nason, *The P(T) spectrum in heavy flavor hadroproduction*, *JHEP* **05** (1998) 007, [hep-ph/9803400].

[90] M. Cacciari, S. Frixione and P. Nason, *The P(T) spectrum in heavy flavor photoproduction*, *JHEP* **03** (2001) 006, [hep-ph/0102134].

[91] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, [hep-ph/0603175].

[92] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [0710.3820].

[93] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029, [hep-ph/0204244].

[94] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146].

[95] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [1002.2581].

[96] X.-N. Wang and M. Gyulassy, *HIJING : A Monte Carlo model for multiple jet production in pp, pA, and AA collisions*, *Phys. Rev.* **D44** (Dec, 1991) 3501–3516.

[97] H. J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog and K. Werner, *Parton based Gribov-Regge theory*, *Phys. Rept.* **350** (2001) 93–289, [hep-ph/0007198].

[98] C. Flensburg, G. Gustafson and L. Lonnblad, *Inclusive and Exclusive Observables from Dipoles in High Energy Collisions*, *JHEP* **08** (2011) 103, [1103.4321].

[99] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450. [Yad. Fiz.15,781(1972)].

[100] L. N. Lipatov, *The parton model and perturbation theory*, *Sov. J. Nucl. Phys.* **20** (1975) 94–102. [Yad. Fiz.20,181(1974)].

[101] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B126** (1977) 298–318.

[102] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641–653. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[103] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J.* **C63** (2009) 189–285, [0901.0002].

[104] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump et al., *Implications of CTEQ global analysis for collider observables*, *Phys. Rev.* **D78** (2008) 013004, [0802.0007].

[105] NNPDF collaboration, R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, A. Piccione et al., *A Determination of parton distributions with faithful uncertainty estimation*, *Nucl. Phys.* **B809** (2009) 1–63, [0808.1231]. [Erratum: Nucl. Phys.B816,293(2009)].

[106] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo et al., *A first unbiased global NLO determination of parton distributions and their uncertainties*, *Nucl. Phys.* **B838** (2010) 136–206, [1002.4407].

[107] R. Placakyte, *Parton Distribution Functions*, in *Proceedings, 31st International Conference on Physics in collisions (PIC 2011): Vancouver, Canada, August 28-September 1, 2011*, 2011. 1111.5452.

[108] C. Peterson, D. Schlatter, I. Schmitt and P. M. Zerwas, *Scaling violations in inclusive e^+e^- annihilation spectra*, *Phys. Rev. D* **27** (Jan, 1983) 105–111.

[109] M. Lisovyi, A. Verbytskyi and O. Zenaiev, *Combined analysis of charm-quark fragmentation-fraction measurements*, *Eur. Phys. J.* **C76** (2016) 397, [1509.01061].

[110] CDF collaboration, T. Aaltonen et al., *Measurement of Ratios of Fragmentation Fractions for Bottom Hadrons in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96-TeV*, *Phys. Rev.* **D77** (2008) 072003, [0801.4375].

[111] LHCb collaboration, R. Aaij et al., *Measurement of b-hadron production fractions in* 7 TeV *pp collisions*, *Phys. Rev.* **D85** (2012) 032008, [1111.2357].

[112] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 31–145.

[113] B. Andersson, *The Lund model*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **7** (1997) 1–471.

[114] C. Bierlich, G. Gustafson, L. Lnnblad and A. Tarasov, *Effects of Overlapping Strings in pp Collisions*, *JHEP* **03** (2015) 148, [1412.6259].

[115] J. R. Christiansen and P. Z. Skands, *String Formation Beyond Leading Colour*, *JHEP* **08** (2015) 003, [1505.01681].

[116] C. Bierlich and J. R. Christiansen, *Effects of color reconnection on hadron flavor observables*, *Phys. Rev.* **D92** (2015) 094010, [1507.02091].

[117] A. Ortiz Velasquez, P. Christiansen, E. Cuautle Flores, I. Maldonado Cervantes and G. Pai, *Color Reconnection and Flowlike Patterns in pp Collisions*, *Phys. Rev. Lett.* **111** (2013) 042001, [1303.6326].

[118] T. Biro, H. Nielsen and J. Knoll, *Colour rope model for extreme relativistic heavy ion collisions*, *Nuclear Physics B* **245** (1984) 449 – 468.

[119] N. Armesto, *Nuclear shadowing*, *J. Phys.* **G32** (2006) R367–R394, [hep-ph/0604108].

[120] D. Higinbotham, G. A. Miller, O. Hen and K. Rith, "The EMC effect still puzzles after 30 years." "http://www.cerncourier.com/cws/article/cern/53091". [Online; accessed 17-March-2017].

[121] K. J. Eskola, H. Paukkunen and C. A. Salgado, *EPS09: A New Generation of NLO and LO Nuclear Parton Distribution Functions*, *JHEP* **04** (2009) 065, [0902.4154].

[122] A. Accardi, *Cronin effect in proton nucleus collisions: A Survey of theoretical models*, hep-ph/0212148.

[123] A. M. Sickles, *Possible Evidence for Radial Flow of Heavy Mesons in d+Au Collisions*, *Phys. Lett.* **B731** (2014) 51–56, [1309.6924].

[124] P. F. Kolb, P. Huovinen, U. W. Heinz and H. Heiselberg, *Elliptic flow at SPS and RHIC: From kinetic transport to hydrodynamics*, *Phys. Lett.* **B500** (2001) 232–240, [hep-ph/0012137].

[125] E. Schnedermann, J. Sollfrank and U. W. Heinz, *Thermal phenomenology of hadrons from 200-A/GeV S+S collisions*, *Phys. Rev.* **C48** (1993) 2462–2475, [nucl-th/9307020].

[126] A. Beraudo, A. De Pace, M. Monteno, M. Nardi and F. Prino, *Heavy-flavour production in high-energy d-Au and p-Pb collisions*, *JHEP* **03** (2016) 123, [1512.05186].

[127] Y. Xu, S. Cao, G.-Y. Qin, W. Ke, M. Nahrgang, J. Auvinen et al., *Heavy-flavor dynamics in relativistic p-Pb collisions at* $\sqrt{s_{NN}} = 5.02$ *TeV*, *Nucl. Part. Phys. Proc.* **276-278** (2016) 225–228, [1510.07520].

[128] R. Rapp and H. van Hees, *Heavy Quarks in the Quark-Gluon Plasma*, in *Quark-gluon plasma 4*, pp. 111–206, 2010. 0903.1096. DOI.

[129] G. D. Moore and D. Teaney, *How much do heavy quarks thermalize in a heavy ion collision?*, *Phys. Rev.* **C71** (2005) 064904, [hep-ph/0412346].

[130] D. d. Enterria and B. Betz, *High-pT Hadron Suppression and Jet Quenching*, pp. 285–339. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. 10.1007/978-3-642-02286-9.

[131] Y. L. Dokshitzer and D. E. Kharzeev, *Heavy quark colorimetry of QCD matter*, *Phys. Lett.* **B519** (2001) 199–206, [hep-ph/0106202].

[132] F. Prino and R. Rapp, *Open Heavy Flavor in QCD Matter and in Nuclear Collisions*, *J. Phys.* **G43** (2016) 093002, [1603.00529].

[133] K. S. Sateesh, *An Experimental signal for diquarks in quark gluon plasma*, *Phys. Rev.* **D45** (1992) 866–868.

[134] S. H. Lee, K. Ohnishi, S. Yasui, I.-K. Yoo and C.-M. Ko, *Lambda(c) enhancement from strongly coupled quark-gluon plasma*, *Phys. Rev. Lett.* **100** (2008) 222301, [0709.3637].

[135] Y. Oh, C. M. Ko, S. H. Lee and S. Yasui, *Heavy baryon/meson ratios in relativistic heavy ion collisions*, *Phys. Rev.* **C79** (2009) 044905, [0901.1382].

[136] E. V. Shuryak and I. Zahed, *Toward a theory of binary bound states in the quark-gluon plasma*, *Phys. Rev. D* **70** (Sep, 2004) 054507.

[137] P. R. Sorensen and X. Dong, *Suppression of non-photonic electrons from enhancement of charm baryons in heavy ion collisions*, *Phys. Rev.* **C74** (2006) 024902, [nucl-th/0512042].

[138] G. Martinez-Garcia, S. Gadrat and P. Crochet, *Consequences of a Lambda(c)/D enhancement effect on the non-photonic electron nuclear modification factor in central heavy ion collisions at RHIC energy*, *Phys. Lett.* **B663** (2008) 55–60, [0710.2152]. [Erratum: Phys. Lett.B666,533(2008)].

[139] S. Ghosh, S. K. Das, V. Greco, S. Sarkar and J.-e. Alam, *Diffusion of Λ_c in hot hadronic medium and its impact on Λ_c/D ratio*, *Phys. Rev.* **D90** (2014) 054018, [1407.5069].

[140] L. Tolos, J. M. Torres-Rincon and S. K. Das, *Transport coefficients of heavy baryons*, *Phys. Rev.* **D94** (2016) 034018, [1601.03743].

[141] S. K. Das, J. M. Torres-Rincon, L. Tolos, V. Minissale, F. Scardina and V. Greco, *Propagation of heavy baryons in heavy-ion collisions*, *Phys. Rev.* **D94** (2016) 114039, [1604.05666].

[142] A. Andronic et al., *Heavy-flavour and quarkonium production in the LHC era: from protonproton to heavy-ion collisions*, *Eur. Phys. J.* **C76** (2016) 107, [1506.03981].

[143] ALICE collaboration, B. Abelev et al., *Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **01** (2012) 128, [1111.1553].

[144] ALICE collaboration, J. Adam et al., *D-meson production in p-Pb collisions at $\sqrt{s_{\rm NN}}$ =5.02 TeV and in pp collisions at $\sqrt{s}$ =7 TeV*, *Phys. Rev.* **C94** (2016) 054908, [1605.07569].

[145] ALICE collaboration, S. Acharya et al., *Measurement of D-meson production at mid-rapidity in pp collisions at $\sqrt{\mathbf{s}} = \mathbf{7}$ TeV*, 1702.00766.

[146] ALICE collaboration, B. Abelev et al., *Production of muons from heavy flavour decays at forward rapidity in pp and Pb-Pb collisions at* $\sqrt{s_{NN}}$ *= 2.76 TeV*, *Phys. Rev. Lett.* **109** (2012) 112301, [1205.6443].

[147] ALICE collaboration, B. Abelev et al., *Heavy flavour decay muon production at forward rapidity in proton-proton collisions at* $\sqrt{s}$ *= 7 TeV*, *Phys. Lett.* **B708** (2012) 265–275, [1201.3791].

[148] ALICE collaboration, B. Abelev et al., *Measurement of electrons from semileptonic heavy-flavour hadron decays in pp collisions at* $\sqrt{s}$ *= 7 TeV*, *Phys. Rev.* **D86** (2012) 112007, [1205.5423].

[149] ALICE collaboration, B. B. Abelev et al., *Measurement of electrons from semileptonic heavy-flavor hadron decays in pp collisions at* $\sqrt{s} = 2.76$ *TeV*, *Phys. Rev.* **D91** (2015) 012001, [1405.4117].

[150] ALICE collaboration, B. Abelev et al., *Measurement of electrons from beauty hadron decays in pp collisions at* $\sqrt{s} = 7$ *TeV*, *Phys. Lett.* **B721** (2013) 13–23, [1208.1902]. [Erratum: Phys. Lett.B763,507(2016)].

[151] ALICE collaboration, B. B. Abelev et al., *Beauty production in pp collisions at* $\sqrt{s}$ *= 2.76 TeV measured via semi-electronic decays*, *Phys. Lett.* **B738** (2014) 97–108, [1405.4144].

[152] M. Cacciari, S. Frixione, N. Houdeau, M. L. Mangano, P. Nason and G. Ridolfi, *Theoretical predictions for charm and bottom production at the LHC*, *JHEP* **10** (2012) 137, [1205.6344].

[153] B. A. Kniehl, G. Kramer, I. Schienbein and H. Spiesberger, *Inclusive Charmed-Meson Production at the CERN LHC*, *Eur. Phys. J.* **C72** (2012) 2082, [1202.0439].

[154] R. Maciula and A. Szczurek, *Open charm production at the LHC - k_t-factorization approach*, *Phys. Rev.* **D87** (2013) 094022, [1301.3033].

[155] A. M. Barbano, *D-meson nuclear modification factor and elliptic flow measurements in Pb–Pb collisions at* $\sqrt{s_{\rm NN}}$*=5.02 TeV with ALICE at the LHC*, in *Quark Matter 2017 proceedings, forthcoming*, 2017.

[156] Terrevoli, C, D^0 *measurements in pp and Pb–Pb collisions at* $\sqrt{s_{\rm NN}}$*=5.02 TeV with ALICE at the LHC*, in *Quark Matter 2017 proceedings, forthcoming*, 2017.

[157] PHENIX collaboration, A. Adare et al., *Measurement of high-p(T) single electrons from heavy-flavor decays in p+p collisions at $\sqrt{s}$ = 200-GeV*, *Phys. Rev. Lett.* **97** (2006) 252002, [hep-ex/0609010].

[158] STAR collaboration, H. Agakishiev et al., *High p_T non-photonic electron production in $p+p$ collisions at $\sqrt{s} = 200$ GeV*, *Phys. Rev.* **D83** (2011) 052006, [1102.2611].

[159] STAR collaboration, L. Adamczyk et al., *Measurements of D^0 and D^* Production in $p+p$ Collisions at $\sqrt{s} = 200$ GeV*, *Phys. Rev.* **D86** (2012) 072013, [1204.4244].

[160] CDF collaboration, D. Acosta et al., *Measurement of prompt charm meson production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. Lett.* **91** (2003) 241804, [hep-ex/0307080].

[161] C. Lourenco and H. K. Wohri, *Heavy flavour hadro-production from fixed-target to collider energies*, *Phys. Rept.* **433** (2006) 127–180, [hep-ph/0609101].

[162] M. L. Mangano, P. Nason and G. Ridolfi, *Heavy-quark correlations in hadron collisions at next-to-leading order*, *Nuclear Physics B* **373** (1992) 295 – 345.

[163] R. Sharma, I. Vitev and B.-W. Zhang, *Light-cone wave function approach to open heavy flavor dynamics in QCD matter*, *Phys. Rev.* **C80** (2009) 054902, [0904.0032].

[164] H. Fujii and K. Watanabe, *Heavy quark pair production in high energy pA collisions: Open heavy flavors*, *Nucl. Phys.* **A920** (2013) 78–93, [1308.1258].

[165] Z.-B. Kang, I. Vitev, E. Wang, H. Xing and C. Zhang, *Multiple scattering effects on heavy meson production in p+A collisions at backward rapidity*, *Phys. Lett.* **B740** (2015) 23–29, [1409.2494].

[166] LHCb collaboration, *Study of cold nuclear matter effects using prompt D^0 meson production in p*Pb *collisions at LHCb*, .

[167] ALICE collaboration, J. Adam et al., *Measurement of D-meson production versus multiplicity in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, *JHEP* **08** (2016) 078, [1602.07240].

[168] LHCb collaboration, R. Aaij et al., *Prompt charm production in pp collisions at sqrt(s)=7 TeV*, *Nucl. Phys.* **B871** (2013) 1–20, [1302.2864].

[169] E. E. Kluge, *Charmed particle production at the CERN intersecting storage rings*, .

[170] LEBC-EHS collaboration, M. Aguilar-Bentez et al., Λ_c *production characteristics in proton-proton interactions at 400 GeV/c*, *Phys. Lett. B* **199** (Jul, 1987) 462–468. 18 p.

[171] ZEUS collaboration, H. Abramowicz et al., *Measurement of charm fragmentation fractions in photoproduction at HERA*, *JHEP* **09** (2013) 058, [1306.4862].

[172] CLEO collaboration, D. Bortoletto et al., *Charm Production in Nonresonant e+ e- Annihilations at s**(1/2) = 10.55-GeV*, *Phys. Rev.* **D37** (1988) 1719. [Erratum: Phys. Rev.D39,1471(1989)].

[173] CLEO collaboration, P. Avery et al., *Inclusive production of the charmed baryon Lambda(c) from e+ e- annihilations at s**(1/2) = 10.55-GeV*, *Phys. Rev.* **D43** (1991) 3599–3610.

[174] ARGUS collaboration, H. Albrecht et al., *Observation of the Charmed Baryon* $\Lambda(c)$ *in* e^+e^- *Annihilation at 10-GeV*, *Phys. Lett.* **B207** (1988) 109–114.

[175] Belle collaboration, R. Seuster et al., *Charm hadrons from fragmentation and B decays in e+ e- annihilation at s**(1/2) = 10.6-GeV*, *Phys. Rev.* **D73** (2006) 032002, [hep-ex/0506068].

[176] BaBar collaboration, B. Aubert et al., *Inclusive Lambda(c)+ Production in e+ e- Annihilations at s**(1/2) = 10.54-GeV and in Upsilon(4S) Decays*, *Phys. Rev.* **D75** (2007) 012003, [hep-ex/0609004].

[177] ALICE collaboration, J. Adam et al., *Transverse momentum dependence of D-meson production in Pb-Pb collisions at* $\sqrt{s_{\rm NN}}$ *= 2.76 TeV*, *JHEP* **03** (2016) 081, [1509.06888].

[178] ALICE collaboration, J. Adam et al., *Centrality dependence of high-*p_T *D meson suppression in Pb-Pb collisions at* $\sqrt{s_{\rm NN}} = 2.76$ *TeV*, *JHEP* **11** (2015) 205, [1506.06604].

[179] ALICE collaboration, B. B. Abelev et al., *Production of charged pions, kaons and protons at large transverse momenta in pp and PbPb collisions at* $\sqrt{s_{\rm NN}}$ *=2.76 TeV*, *Phys. Lett.* **B736** (2014) 196–207, [1401.1250].

[180] CMS collaboration, V. Khachatryan et al., *Suppression and azimuthal anisotropy of prompt and nonprompt* J/ψ *production in PbPb collisions at* $\sqrt{s_{_{NN}}} = 2.76$ TeV, *Eur. Phys. J.* **C77** (2017) 252, [1610.00613].

[181] ALICE collaboration, J. Adam et al., *Elliptic flow of muons from heavy-flavour hadron decays at forward rapidity in PbPb collisions at* $\sqrt{s_{\rm NN}} = 2.76$ *TeV*, *Phys. Lett.* **B753** (2016) 41–56, [1507.03134].

[182] Y. Xu, M. Nahrgang, J. E. Bernhard, S. Cao and S. A. Bass, *A data-driven analysis of the heavy quark transport coefficient*, 1704.07800.

[183] STAR collaboration, G. Xie, Λ_c *Production in Au+Au Collisions at* $\sqrt{s_{NN}}$ *= 200 GeV measured by the STAR experiment*, in *26th International Conference on Ultrarelativistic Nucleus-Nucleus Collisions (Quark Matter 2017) Chicago,Illinois, USA, February 6-11, 2017*, 2017. 1704.04353.

[184] I. Kuznetsova and J. Rafelski, *Heavy flavor hadrons in statistical hadronization of strangeness-rich QGP*, *Eur. Phys. J.* **C51** (2007) 113–133, [hep-ph/0607203].

[185] A. Andronic, P. Braun-Munzinger, K. Redlich and J. Stachel, *Charmonium and open charm production in nuclear collisions at SPS/FAIR energies and the possible influence of a hot hadronic medium*, *Phys. Lett.* **B659** (2008) 149–155, [0708.1488].

[186] A. Andronic, P. Braun-Munzinger, K. Redlich and J. Stachel, *Heavy quark(onium) at LHC: the statistical hadronization case*, *J. Phys.* **G37** (2010) 094014, [1002.4441].

[187] ALICE collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.

[188] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[189] ATLAS collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[190] CMS collaboration, S. Chatrchyan et al., *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[191] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [1207.7214].

[192] CMS collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [1207.7235].

[193] LHCb collaboration, A. A. Alves, Jr. et al., *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.

[194] "LHC Underground Layout." "http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/pagesources/LHC-Underground-Layout.png". [Online; accessed 13-May-2016].

[195] W. Fischer and J. M. Jowett, *Ion Colliders*, *Rev. Accel. Sci. Technol.* **7** (Jan, 2015) 49–76. mult. p.

[196] CERN, "Luminosity plots pp 2011." "https://lpc.web.cern.ch/lumiplots_2011_pp.htm". [Online; accessed 28-April-2017].

[197] "Minutes of the 130th meeting of the CERN research board held on 14 February 1997 CERN/DG/RB 97-252, M130." http://committees.web.cern.ch/committees/RB/RBMinutes/RBMinutes130.html". [Online; accessed 18-October-2017].

[198] C. Fabjan and J. Schukraft, *The Story of ALICE: Building the dedicated heavy ion detector at LHC*, in *'The Large Hadron Collider: A marvel technology', EPFL-Press Lausanne, Switzerland, 2009 (Editor: L. Evans), chapter 5.4*, 2011. 1101.1257.

[199] ALICE collaboration, K. Aamodt et al., *Charged-particle multiplicity density at mid-rapidity in central Pb-Pb collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *Phys. Rev. Lett.* **105** (2010) 252301, [1011.3916].

[200] ALICE collaboration, K. Aamodt et al., *Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC*, *Eur. Phys. J.* **C68** (2010) 89–108, [1004.3034].

[201] Wikipedia, "Schematics of the ALICE subdetectors." https://commons.wikimedia.org/wiki/File:2012-Aug-02-ALICE_3D_v0_with_Text_(1)_2.jpg. [Online; accessed 23-July-2017].

[202] W. Blum, W. Riegler and L. Rolandi, *Particle detection with drift chambers*. Springer Science & Business Media, 2008.

[203] ALICE collaboration, B. B. Abelev et al., *Performance of the ALICE Experiment at the CERN LHC*, *Int. J. Mod. Phys.* **A29** (2014) 1430044, [1402.4476].

[204] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth.* **A262** (1987) 444–450.

[205] ALICE collaboration, J. Adam et al., *Particle identification in ALICE: a Bayesian approach*, *Eur. Phys. J. Plus* **131** (2016) 168, [1602.01392].

[206] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences*. Cambridge University Press, New York, NY, USA, 2005.

[207] Belle collaboration, A. Zupanc, C. Bartel, N. Gabyshev, I. Adachi, H. Aihara, D. M. Asner et al., *Measurement of the branching fraction $\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)$*, *Phys. Rev. Lett.* **113** (Jul, 2014) 042002.

[208] BESIII collaboration, M. Ablikim et al., *Measurements of absolute hadronic branching fractions of Λ_c^+ baryon*, *Phys. Rev. Lett.* **116** (2016) 052001, [1511.08380].

[209] E791 collaboration, E. M. Aitala et al., *Multidimensional resonance analysis of $\Lambda_c^+ \to pK^-\pi^+$*, *Phys. Lett.* **B471** (2000) 449–459, [hep-ex/9912003].

[210] ACCMOR collaboration, A. Bozek et al., *A Study of Λ_c^+ decays into $pK^-\pi^+$, $pK^-\pi^+\pi^0$ and $pK^-\pi^+\pi^0\pi^0$*, *Phys. Lett.* **B312** (1993) 247–252.

[211] M. Basile et al., *A measurement of two resonant contributions in the Λ_C^+ branching ratios*, *Il Nuovo Cimento A (1965-1970)* **62** (1981) 14–20.

[212] M. Mager, P. Braun-Munzinger and L. Musa, *Studies on the upgrade of the ALICE central tracker*, Ph.D. thesis, Darmstadt, Tech. U., Nov, 2012.

[213] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, *PoS* **ACAT** (2007) 040, [physics/0703039].

[214] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, *Journal of Computer and System Sciences* **55** (1997) 119 – 139.

[215] R. Brun, F. Bruyant, M. Maire, A. C. McPherson and P. Zanarini, *GEANT3*, .

[216] P. Z. Skands, *The Perugia Tunes*, in *Proceedings, 1st International Workshop on Multiple Partonic Interactions at the LHC (MPI08): Perugia, Italy, October 27-31, 2008*, pp. 284–297, 2009. 0905.3418.

[217] F. James and M. Roos, *MINUIT: a system for function minimization and analysis of the parameter errors and corrections*, *Comput. Phys. Commun.* **10** (Jul, 1975) 343–367. 38 p.

[218] ALICE collaboration, K. Aamodt et al., *Midrapidity antiproton-to-proton ratio in pp collisions at* $\sqrt{s} = 0.9$ *and* 7 *TeV measured by the ALICE experiment*, *Phys. Rev. Lett.* **105** (2010) 072002, [1006.5432].

[219] A. Ferrari, P. R. Sala, A. Fasso and J. Ranft, *FLUKA: A multi-particle transport code (Program version 2005)*, .

[220] T. T. Böhlen, F. Cerutti, M. P. W. Chin, A. Fassò, A. Ferrari, P. G. Ortega et al., *The FLUKA Code: Developments and Challenges for High Energy and Medical Applications*, *Nuclear Data Sheets* **120** (June, 2014) 211–214.

[221] L. Gladilin, *Fragmentation fractions of c and b quarks into charmed hadrons at LEP*, *Eur. Phys. J.* **C75** (2015) 19, [1404.3888].

[222] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[223] B. A. et al and T. A. Collaboration, *Technical Design Report for the Upgrade of the ALICE Inner Tracking System*, *Journal of Physics G: Nuclear and Particle Physics* **41** (2014) 087002.

[224] L. Musa, *Upgrade of the Inner Tracking System Conceptual Design Report. Conceptual Design Report for the Upgrade of the ALICE ITS*, Tech. Rep. CERN-LHCC-2012-013. LHCC-P-005, CERN, Geneva, Aug, 2012.

[225] LHCb collaboration, R. Aaij et al., *Study of the production of Λ_b^0 and $\overline{B}^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \to J/\psi p K^-$ branching fraction*, *Chin. Phys.* **C40** (2016) 011001, [1509.00292].

[226] CMS collaboration, S. Chatrchyan et al., *Measurement of the Λ_b cross section and the $\bar{\Lambda}_b$ to Λ_b ratio with $J/\Psi\Lambda$ decays in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys. Lett.* **B714** (2012) 136–157, [1205.0594].

[227] ALICE collaboration, B. Abelev et al., *Measurement of inelastic, single- and double-diffraction cross sections in proton–proton collisions at the LHC with ALICE*, *Eur. Phys. J.* **C73** (2013) 2456, [1208.4968].

[228] ALICE collaboration, B. B. Abelev et al., *Measurement of visible cross sections in proton-lead collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV in van der Meer scans with the ALICE detector*, *JINST* **9** (2014) P11003, [1405.1849].

[229] M. Figueredo, R. Romita, J. Norman, K. Schweda and J. Wilkinson, $\Lambda_{\rm C}$ *to $pK\pi$ in p-p collisions with Std cuts*, . Unpublished ALICE internal note.

[230] M. Figueredo and J. Norman, $\Lambda_{\rm C}$ *to $pK\pi$ in p-Pb collisions with Std cuts*, . Unpublished ALICE internal note.

[231] ALICE collaboration, B. B. Abelev et al., *Measurement of prompt D-meson production in $p-Pb$ collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, *Phys. Rev. Lett.* **113** (2014) 232301, [1405.3452].

[232] R. Averbeck, N. Bastid, Z. C. del Valle, P. Crochet, A. Dainese and X. Zhang, *Reference Heavy Flavour Cross Sections in pp Collisions at $\sqrt{s} = 2.76$ TeV, using a pQCD-Driven $\sqrt{s}$-Scaling of ALICE Measurements at $\sqrt{s} = 7$ TeV*, 1107.3243.

[233] E. Braaten, K.-m. Cheung, S. Fleming and T. C. Yuan, *Perturbative QCD fragmentation functions as a model for heavy quark fragmentation*, *Phys. Rev.* **D51** (1995) 4819–4829, [hep-ph/9409316].

[234] B. A. Kniehl and G. Kramer, *Charmed-hadron fragmentation functions from CERN LEP1 revisited*, *Phys. Rev.* **D74** (2006) 037502, [hep-ph/0607306].

[235] F. Prino. private communication.

[236] T. Kneesch, B. A. Kniehl, G. Kramer and I. Schienbein, *Charmed-meson fragmentation functions with finite-mass corrections*, *Nucl. Phys.* **B799** (2008) 34–59, [0712.0481].

[237] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, *Eur. Phys. J.* **C74** (2014) 3024, [1404.5630].

[238] Y. Watanabe. private communication.

[239] H. Albrecht, H. Cronstrm, H. Ehrlichmann, T. Hamacher, R. Hofmann, T. Kirchhoff et al., *Observation of semileptonic decay*, *Physics Letters B* **303** (1993) 368 – 376.

[240] CLEO collaboration, J. P. Alexander, C. Bebek, K. Berkelman, K. Bloom, T. E. Browder, D. G. Cassel et al., *First observation of the decay $\Xi_c^+ \to \Xi^0 e^+ \nu_e$ and an estimate of the Ξ_c^+ Ξ_c^0 lifetime ratio*, *Phys. Rev. Lett.* **74** (Apr, 1995) 3113–3117.

[241] S. Frixione, P. Nason and G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126, [0707.3088].

[242] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin et al., *New parton distributions for collider physics*, *Phys. Rev.* **D82** (2010) 074024, [1007.2241].

[243] ALICE collaboration, B. A. et al., *Upgrade of the ALICE Experiment: Letter Of Intent*, *Journal of Physics G: Nuclear and Particle Physics* **41** (2014) 087001.

[244] M. Mager, *ALPIDE, the Monolithic Active Pixel Sensor for the ALICE ITS upgrade*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **824** (2016) 434 – 438. Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors.

[245] M. He, R. J. Fries and R. Rapp, $\mathbf{D_s}$*-Meson as Quantitative Probe of Diffusion and Hadronization in Nuclear Collisions*, *Phys. Rev. Lett.* **110** (2013) 112301, [1204.4442].

[246] C. Terrevoli. private communication.

[247] P. J. M. Laarhoven and E. H. L. Aarts, eds., *Simulated Annealing: Theory and Applications*. Kluwer Academic Publishers, Norwell, MA, USA, 1987.

[248] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *Equation of state calculations by fast computing machines*, *The Journal of Chemical Physics* **21** (1953) 1087–1092, [http://dx.doi.org/10.1063/1.1699114].

[249] ALICE collaboration, J. Adam et al., *Measurement of D_s^+ production and nuclear modification factor in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV*, *JHEP* **03** (2016) 082, [1509.07287].

[250] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban and D. Whiteson, *Jet Flavor Classification in High-Energy Physics with Deep Neural Networks*, *Phys. Rev.* **D94** (2016) 112002, [1607.08633].

[251] P. T. Komiske, E. M. Metodiev and M. D. Schwartz, *Deep learning in color: towards automated quark/gluon jet discrimination*, *JHEP* **01** (2017) 110, [1612.01551].

[252] P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, *Jet Substructure Classification in High-Energy Physics with Deep Neural Networks*, *Phys. Rev.* **D93** (2016) 094034, [1603.09349].

[253] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images deep learning edition*, *JHEP* **07** (2016) 069, [1511.05190].

[254] G. Louppe, K. Cho, C. Becot and K. Cranmer, *QCD-Aware Recursive Neural Networks for Jet Physics*, 1702.00748.

[255] P. Baldi, P. Sadowski and D. Whiteson, *Searching for Exotic Particles in High-Energy Physics with Deep Learning*, *Nature Commun.* **5** (2014) 4308, [1402.4735].

[256] A. Farbin, *Event Reconstruction with Deep Learning*, *PoS* **ICHEP2016** (2016) 180.