Analysis of D^0 and D^{*+}-meson production in pp and p–Pb collisions with ALICE at the LHC

Jeremy Wilkinson

Physikalisches Institut
Universität Heidelberg
2016

Dissertation
submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

put forward by
Jeremy John Wilkinson, MSci
born in York, United Kingdom
Oral examination: 28/07/2016

Analysis of D^0 and D^{*+}-meson production in pp and p–Pb collisions with ALICE at the LHC

Abstract

This thesis presents measurements of D-meson production in the central barrel of the AL-ICE detector in pp and p–Pb collisions. The reconstruction of D^0 mesons in the hadronic channel $D^0 \to K^- \pi^+$ was studied in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ using a Bayesian particle identification (PID) method, in order to test the validity of this new approach. Comparisons were made between these results and those obtained with established PID methods. Consistency was found between the different approaches, as well as an increase of the signal-to-background ratio and a similar or greater statistical significance for most of the implementations of the Bayesian approach. Further measurements of $D^{*+} \to D^0 \pi^+$ were made as a function of charged-particle multiplicity in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$. The aim was to test the role of multi-parton interactions (MPI) and possible collective phenomena in small systems at LHC energies. The results for D^{*+} mesons were consistent with the other D-meson species studied by ALICE (D^0 and D^+). The measurements against mid-rapidity multiplicity showed consistency with previous results from pp collisions; however, a slower increase of the relative D-meson yield was found as a function of multiplicity at large rapidity for p–Pb collisions than pp collisions. The results for both multiplicity estimators were reproduced by phenomenological models, both with and without viscous hydrodynamics.

Zusammenfassung

In dieser Dissertation werden Messungen zur Produktion von D-Mesonen im zentralen Rapiditätsbereich des ALICE-Detektors präsentiert. Dazu wurde die Rekonstruktion von D^0-Mesonen im hadronischen Zerfallskanal $D^0 \to K^- \pi^+$ in pp-Kollisionen bei $\sqrt{s} = 7\,\text{TeV}$ unter Verwendung einer bayesschen Teilchenidentifikationsmethode (PID-Methode) studiert. Durch den Vergleich mit Ergebnissen der bewährten PID-Methode wurde die Gültigkeit der Resultate des bayesschen Ansatzes zunächst verifiziert, und anschließend ein verbessertes Signal-zu-Untergrund-Verhältnis sowie ähnliche oder erhöhte statistische Signifikanz ermittelt. Weitere Messungen von $D^{*+} \to D^0 \pi^+$ wurden als Funktion von Multiplizität von geladenen Teilchen in p–Pb-Kollisionen bei $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$ durchgeführt, mit dem Ziel den Einfluss von Multi-Parton-Wechselwirkungen (MPI) und mögliche kollektive Phänomene in kleinen Kollisionssystemen bei LHC-Energien zu überprüfen. Die Ergebnisse für D^{*+}-Mesonen entsprechen denen der ebenfalls von ALICE inspizierten D^0- und D^+-Mesonen. Die Messungen gegen Multiplizität bei mittlerer Rapidität zeigten Übereinstimmung mit früheren Messungen in pp-Kollisionen, im Gegensatz zu einem langsameren Aufstieg des relativen D-Meson-Ertrags als Funktion der Multiplizität bei höherer Rapidität in p–Pb-Kollisionen gegenüber pp-Kollisionen. Phänomenologische Modelle – sowohl mit als auch ohne viskose Hydrodynamik – beschreiben die Ergebnisse beider Rapiditätsbereiche.

Contents

Contents

List of Figures

3   A Large Ion Collider Experiment at the LHC

4   Data analysis

5   A Bayesian approach to PID in ALICE

6 D^{*+}-meson production as a function of multiplicity

List of Tables

List of Tables

1 | Introduction

The production of charm and beauty quarks in high-energy proton and heavy-ion collisions is among the key observables at the Large Hadron Collider. Due to their large masses, these heavy quarks can only be produced in the initial stages of a Pb–Pb collision, before the thermalisation time of the Quark–Gluon Plasma (QGP), and not during thermal processes at later stages. This means that they are able to probe the full evolution of the system. One important measurement to make in this system is the nuclear modification factor R_{AA}, which compares the p_T-dependent yields of particles in Pb–Pb collisions with those from a superposition of binary pp collisions. This quantity is sensitive to the energy loss experienced by particles as they propagate through the medium. Measurements of the R_{AA} for various species therefore allow possible flavour dependences to be measured, particularly the so-called 'dead-cone effect' [1], which predicts a mass ordering of energy loss, leading heavier-flavour quarks to lose less energy through interactions within the QGP than lighter quarks. Also of interest is the collective flow of particles due to interactions between the constituents of the QGP, which may be characterised through hydrodynamic calculations within a locally thermalised medium. As they are present throughout the lifetime of the system, heavy quarks are expected to undergo more interactions within the medium, and therefore to be more sensitive to in-medium flow than lighter quarks. This is studied through the anisotropic flow, most commonly characterised by the elliptic flow parameter v_2, which allows the extent of collective motion in the expansion of the QGP to be studied. Measurements of both the nuclear modification and flow of heavy quarks in Pb–Pb collisions thus allow the transport coefficients of the QGP to be examined in greater detail.

In the light-flavour sector, measurements of baryon-to-meson production ratios have shown an enhancement of protons with respect to pions, and more significantly the strange Λ baryons with respect to K_S^0 mesons, at intermediate momenta in central Au–Au and Pb–Pb collisions compared to pp collisions [2, 3], potentially due to the effects of quark coalescence on hadronisation [4]. Measurements of the charmed Λ_c^+ baryon in Pb–Pb collisions would allow these measurements to be extended into the heavy-flavour sector, thus determining whether similar processes affect the hadronisation of heavier quarks in the medium. However, due to its short decay length and

limitations in the tracking resolution of current particle detectors, the Λ_c^+ baryon remains elusive in heavy-ion collisions at the LHC. Charmed D mesons at ultra-low momenta are similarly challenging to measure due to large levels of combinatorial background, thus limiting the momentum range at which nuclear effects on charm production can be precisely studied.

Neither challenge can be adequately overcome with improved topological selections alone, thus necessitating the development and implementation of a new particle identification approach using Bayes' Theorem [5] in order to improve the precision of these measurements. However, before any such method can be used more widely in the ALICE framework, it must be carefully established that the detector response is properly modelled and corrected for in Monte Carlo simulations, and thus that the use of this method does not lead to significant biases in the final results of an analysis with respect to previously used methods. In the course of this work, this validation was performed for a full analysis of the reconstruction of $D^0 \to K^- \pi^+$ decays (and charge conjugates) in proton–proton collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ in ALICE using this novel Bayesian PID approach to combine the signals from multiple detectors [6]. The performance, in terms of the signal-to-background ratio, statistical significance, and reconstruction efficiency, was tested for a variety of different implementations of the Bayesian PID approach, and the systematic uncertainties of this approach as compared to previously established methods were evaluated. In addition, the ability of the PID framework to properly correct for altered descriptions of the expected detector response through different choices of Bayesian priors was tested.

Measurements of charm production in proton–proton and proton–nucleus collisions are also undertaken in ALICE to better understand and disentangle effects arising within the QGP from those occurring either due to increased numbers of independent binary collisions or the presence of nuclear matter in the initial state. However, recent measurements of high-multiplicity events in these small systems have led to a sea change in the study of ultrarelativistic hadronic collisions. Due to a much smaller collision region with respect to Pb–Pb collisions, the pp and p–Pb collision systems were originally thought to be too small and short-lived to sustain a thermalised medium. However, high-multiplicity proton–lead collisions have been shown to exhibit collective behaviour, as evidenced by a 'double-ridge' structure in the study of angular two-particle correlations [7–9], similar to that observed in central Pb–Pb collisions. This raises the possibility of studying even small collision systems in terms of relativistic hydrodynamics.

Due to only being produced in initial hard scatterings, increased heavy-quark yields per event are expected to correspond to significant upward fluctuations of the number of parton–parton scatterings in a collision, known as Multi-Parton Interactions (MPI) [10]. These effects have been observed in the simultaneous production of multiple pairs of heavy quarks in proton–antiproton and proton–proton collisions at

Tevatron and the LHC [11–15]. As the overall charged-particle multiplicity is dominated by light-flavour particles, which may be produced in soft processes at late times in the evolution of the system, it is particularly interesting to study their relation to heavy-flavour production at high multiplicities in order to examine the extent of multi-parton interactions on high-multiplicity events. Charm and beauty production were studied in ALICE as a function of the charged-particle multiplicity in proton–proton collisions at $\sqrt{s} = 7\,\mathrm{TeV}$, where a faster scaling was seen for D- and B-meson production as compared to the overall charged-particle multiplicity, and the results were seen to be consistent with predictions that included relativistic hydrodynamics [16].

In order to compare the effects of multi-parton interactions with those due to the presence of a nucleus in the initial state, similar measurements of D-meson yields as a function of charged-particle multiplicity were also made in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ [17]. This was performed in this work for the reconstruction of D^{*+} mesons in the channel $\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+$ (and charge conjugates) as a function of multiplicity. The charged-particle multiplicity was estimated in two regions: at mid-rapidity (the same region as the D mesons) and at large backward rapidity (the Pb-going direction), in order to judge nuclear effects on the multiplicity in intervals both correlated with and separate from the region in which the D-meson yields are studied. The multiplicity estimation and correction procedure was implemented and carefully tested in order to properly estimate the charged-particle yield from the detector signals, and to remove systematic effects on the multiplicity occurring due to the detector response and acceptance. Finally, the relative D^{*+}-meson yields obtained in this work are compared with those of D^0 and D^+ mesons measured by ALICE [17], as well as with the equivalent results from proton–proton collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ [16] and with model calculations under the EPOS framework, both with and without the inclusion of viscous hydrodynamics in the final state [18, 19].

This thesis is structured as follows. Section 2 gives an overview of the theoretical and experimental context behind Quantum Chromodynamics (QCD) and measurements of heavy-flavour particles in proton–proton, lead–lead and proton–lead collisions, as well as motivating the study of high-multiplicity events in small systems. Section 3 then outlines the experimental apparatus of the ALICE detector, including the subdetectors that are relevant to the analyses presented in this work. The tracking and PID methodology of ALICE is also described. Section 4 deals with the data analysis techniques that are used for the presented work, including the selection of D-meson candidates from the geometrical properties of their decays, and the signal extraction and correction procedure.

Section 5 introduces the principles behind the Bayesian PID approach, and its application to the study of $\mathrm{D}^0 \to \mathrm{K}^-\pi^+$, in order to appraise its validity for use in other measurements such as $\Lambda_c^+ \to \mathrm{pK}^-\pi^+$. This is done by means of comparison with measurements using previously published PID methods. The performance of the

method in terms of the signal quality and statistical significance are also discussed, as well as the effect of the choice of Bayesian priors on the results of the analysis. The work presented in this section was also published by the ALICE Collaboration in [6].

Section 6 discusses the multiplicity-dependent measurement of D^{*+} mesons in p–Pb collisions. The signal extraction procedure and relevant systematic uncertainties for D^{*+} mesons in this analysis are explained in further detail, as are the methods used to estimate and correct the charged-particle multiplicity in both of the studied rapidity regions. Results are presented for D^{*+} mesons, and compared with equivalent results for the other D-meson species measured by ALICE (D^0 and D^+). They are then compared with previous measurements taken by ALICE in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ [16], and with the results given by phenomenological models. The work presented in this section was also published by the ALICE Collaboration in [17].

2 | Theoretical and experimental background

This section outlines some of the relevant theoretical and experimental background behind the measurements presented in this work, and attempts to bring the results into the context of existing measurements of charm production in hadronic collisions. Section 2.1 outlines the strong nuclear force, which mediates the interactions between quarks and gluons, and the theory behind the phase transition to a deconfined state in ultrarelativistic heavy-ion collisions, in terms of Quantum Chromodynamics (QCD). Common variables in heavy-ion collisions are defined in Section 2.2. Section 2.3 introduces the concept of heavy flavours, and brings their relevance in high-energy collisions into focus. Sections 2.4 and 2.5 then concentrate on existing measurements of heavy-flavour physics in Pb–Pb and p–Pb collisions at the LHC. Finally, Section 2.6 deals with the significance of high-multiplicity events in small systems, and their relation to multi-parton interactions (MPI) and possible collective effects.

2.1 Quantum Chromodynamics (QCD)

The four fundamental forces that govern the interactions between particles are the electromagnetic force, the weak nuclear force, the strong nuclear force, and gravitation. The exchange particles of these four forces, the quantities they couple to, and their effective ranges are outlined in Tab. 2.1. The first three of these are described by the Standard Model (SM) of particle physics.

The strong nuclear force is described by the theory of Quantum Chromodynamics (QCD), and acts on colour-charged particles, i.e. quarks and gluons. The quarks are spin-$\frac{1}{2}$ fermions that have three possible colour charges: r, g or b ($\bar{\text{r}}$, $\bar{\text{g}}$ or $\bar{\text{b}}$ in the case of antiquarks). Quarks constitute hadronic matter, in bound states known as baryons ($|\text{qqq}\rangle$, $|\bar{\text{q}}\bar{\text{q}}\bar{\text{q}}\rangle$) and mesons ($|\text{q}\bar{\text{q}}\rangle$). Recent experimental evidence has also pointed to the existence of more exotic forms of hadronic matter, namely tetraquarks ($|\text{qq}\bar{\text{q}}\bar{\text{q}}\rangle$) [21] and pentaquarks ($|\text{q}\bar{\text{q}}\text{qqq}\rangle$) [22]. Gluons are spin-1 gauge bosons that mediate the strong force between quarks within a hadron. Unlike the photon. which is charge-neutral, gluons possess a non-zero value of the quantum number they couple to (colour), and so are able to undergo self-coupling processes. They exist in eight

	Electromagnetic	Weak	Strong	Gravitation
Exchange particle	γ	$W^{\pm}$, Z^0	g	graviton (hypothetical)
Couples to	charge	weak isospin	colour	mass
Relative strength	1	10^{-11}	100	10^{-24}
Range	∞	$0.01\,\text{fm}$	$1\,\text{fm}$	∞

Table 2.1: The fundamental forces of nature. Each force is listed with its exchange particle, coupling strength relative to the EM interaction, and effective range. Values from [20].

possible linearly independent colour–anticolour configurations due to the SU(3) nature of QCD. When bound into hadrons, quarks and gluons are collectively known as 'partons'.

Flavour	Charge (e)	Bare mass $\left(\text{MeV}/c^2\right)$	Constituent mass $\left(\text{MeV}/c^2\right)$
up	$+^2/_3$	$2.3^{+0.7}_{-0.5}$	220–360
down	$-^1/_3$	$4.8^{+0.5}_{-0.3}$	220–360
strange	$-^1/_3$	95 ± 5	419–540
charm	$+^2/_3$	1275 ± 25	1550–1710
beauty	$-^1/_3$	4180 ± 30	~ 4000
top	$+^2/_3$	$173210 \pm 510 \pm 710$	–

Table 2.2: The six quark flavours with their electrical charges with respect to the elemental charge e, bare masses and constituent masses. Bare masses from [23], constituent masses from [24, 25]. The top quark has not been observed to hadronise, and so its constituent mass is not listed.

The six quark flavours are listed in Tab. 2.2 along with their masses. Large differences can be seen between the bare (Higgs) masses and the confined (or constituent) masses. This difference arises as the quark masses stem from a combination of their coupling to the Higgs field in the electroweak regime (bare) and spontaneous symmetry breaking in the QCD vacuum (constituent). These two effects are compared in Fig. 2.1, where it is further highlighted that the light-quark (u, d, s) masses are dominated by QCD symmetry breaking, and the heavy-quark (c, b, t) masses are almost entirely due to Higgs coupling.

Up and down quarks are abundant in the Universe, being the valence constituents of the protons and neutrons that make up ordinary nuclear matter. It is speculated that strange quarks may also be present in nature at the core of neutron stars [26]. The heavier quarks (charm, beauty and top) are not seen as valence quarks in ordinary matter, but may be produced in quark–antiquark pairs through electroweak (strong)
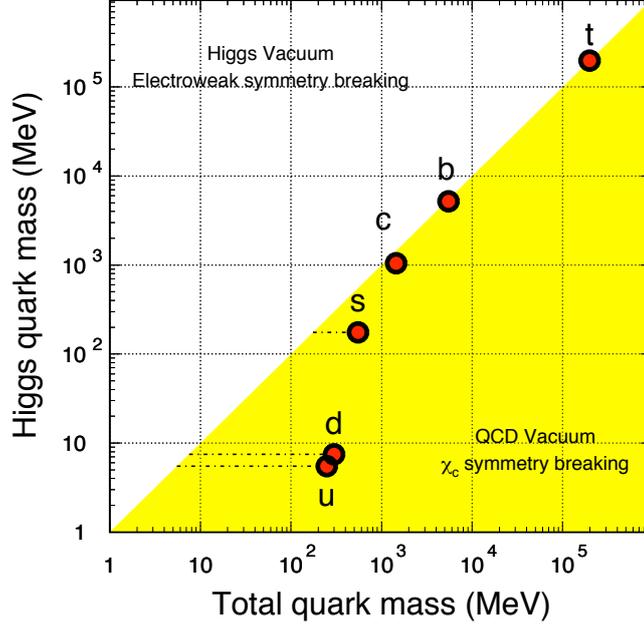
Figure 2.1: Comparative plot of quark masses in the QCD and Higgs vacua. For the light quarks (u, d, s) a large proportion of the mass arises due to chiral symmetry breaking in the QCD vacuum. Figure from [25].

interactions in e^+e^- (hadron) colliders at ultrarelativistic energies. The production of single top quarks as a result of electroweak interactions has also been observed in proton–antiproton and proton–proton collisions at Tevatron and the LHC [27–30]. Virtual quark–antiquark pairs may also be created within a hadron by the splitting of a gluon with sufficiently high momentum, with a decreasing probability for increasing quark mass; these virtual quarks are known as 'sea quarks'. The existence of quarks was confirmed by deep inelastic scatterings (DIS) of electrons on protons [31, 32].

Quarks have never been experimentally observed as free particles, and instead are seen bound into hadrons. This phenomenon occurs due to the properties of the QCD potential between a quark–antiquark pair:

$$V_{\mathrm{QCD}}(r) = -\frac{4}{3}\frac{\alpha_{\mathrm{s}}(r)}{r} + \sigma \cdot r. \tag{2.1}$$

Here, V_{QCD} is the QCD potential between the quark and antiquark, r is the distance between them, α_{s} is the QCD gauge coupling parameter and σ is the QCD string tension. The first term is sometimes known as a 'Coulomb-like' term, whose $1/r$ dependence arises from the exchange of a single, massless spin-1 boson.

Though it is often termed as a 'coupling constant', α_{s} varies with the distance between two partons. This variation is termed as a 'running coupling', illustrated
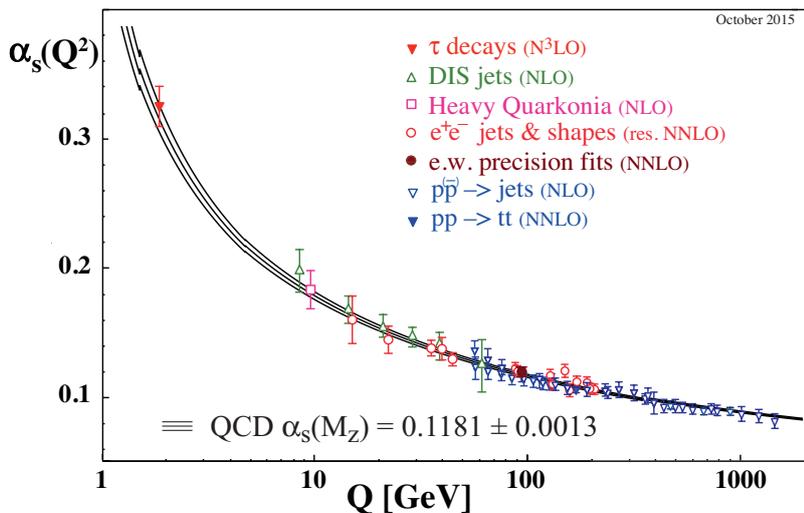
Figure 2.2: Running of the QCD coupling strength $\alpha_{\rm s}$ as a function of the four-momentum transfer Q^2. Measurements (shown as points) are compared with a parameterisation of $\alpha_{\rm s}$ based on measurements at the scale of the Z-boson mass. The order to which perturbation theory is used is indicated in parentheses (NLO = next-to-leading order; NNLO = next-to-next-to-leading order; res. NNLO = NNLO with resummed next-to-leading logarithms; N^3LO = next-to-next-to-next-to-leading order). Figure from [23].

in Fig. 2.2 as a function of the four-momentum transfer, Q^2. Large Q^2 values correspond to higher collision energies and smaller separations between partons, and vice-versa for low Q^2.

The coupling strength may be determined using perturbative calculations of Quantum Chromodynamics (pQCD). The running coupling strength $\alpha_{\rm s}$ can be written at the leading-logarithm approximation in terms of the four-momentum transfer of the gluon as follows [33]:

$$\alpha_{\rm s}(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln(Q^2/\Lambda_{\rm QCD}^2)},\tag{2.2}$$

where n_f is the number of active flavours and $\Lambda_{\rm QCD}$ is the QCD scale parameter, which corresponds to the energy scale below which perturbative coupling becomes divergent. Here, $\Lambda_{\rm QCD}$ is a free parameter, and must be determined experimentally; it is of the order of $\sim$200 MeV [23]. This running of the coupling strength has been verified at a variety of energy scales using results from relativistic colliders, as shown in Fig. 2.2 [23].

At low temperatures and large quark–antiquark separation distances, the second term of Eq. (2.1) dominates, causing the potential to increase linearly as a function of r. An infinite input of energy would therefore be required for quarks to escape this

potential; this phenomenon is known as 'confinement'. At extreme distances between a pair of quarks, it becomes more energetically favourable for the colour string to break and produce a quark–antiquark pair from the vacuum (a process known as 'fragmentation') than for the separation to increase any further. Conversely, as $r \to 0$ (i.e. large Q^2), the string term is less relevant, and the coupling of the Coulomb term weakens due to the running of α_s at high Q^2. Here, the force effectively vanishes, causing the quarks to act as quasi-free particles. This effect is known as 'asymptotic freedom'.

The considerations above apply at lower energy densities. It is, however, possible for this confinement to be overcome under extremes of temperature and/or pressure. This is illustrated by the QCD phase diagram, which is shown in Fig. 2.3 as a function of net baryonic density ρ (the total baryon number per unit volume) and temperature [34]. The white region in the lower-left corner shows ordinary hadronic matter under confinement conditions; the point at $(\rho = 1, T \approx 0)$ corresponds to nuclear matter under standard conditions in the present Universe, and the red and blue arrows indicate the regions of the phase diagram that are probed by high-energy collider experiments.
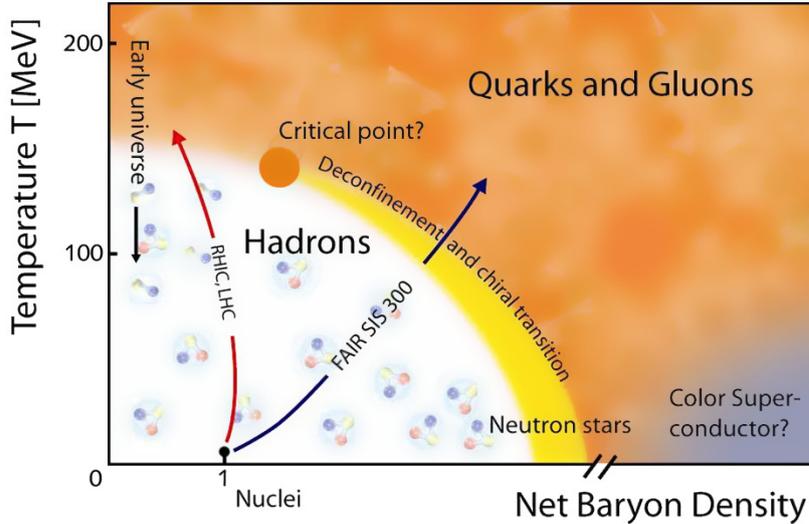
Figure 2.3: The phase diagram for QCD matter, as a function of net baryonic density and temperature. The arrows indicate regions of the phase diagram that are probed by collider experiments. Figure (edited) from [34].

At high baryonic densities and low temperatures, a colour-superconducting state is postulated to occur. Under these conditions, possibly similar to those at the core of a neutron star, the pressure is such that the separation between hadrons is smaller than their radius ($\sim 1\,\mathrm{fm}$), causing them to effectively overlap. At this point, each quark can no longer be meaningfully attributed to any given hadron [35].

The other possible extreme is a high temperature and low baryonic density. Hagedorn originally proposed, under a pion gas approximation, that there would be an upper temperature limit for hadronic matter, T_0, with a value of ~120–160 MeV [36]. The concept of this temperature limit originates from the number density of particles, ρ, in a hadron gas. This increases exponentially as a function of mass, according to

$$\rho(m) \propto e^{\frac{m}{T_0}}. \tag{2.3}$$

According to thermodynamic constraints, where each mass level between m and $m + \mathrm{d}m$ is populated by $e^{-m/T}$ particles, the total number of particles as a function of T then becomes

$$N(m) \propto \int_0^\infty \rho(m) e^{-\frac{m}{T}} \mathrm{d}m = \int_0^\infty e^{\frac{m}{T_0} - \frac{m}{T}} \mathrm{d}m, \tag{2.4}$$

which diverges for temperatures greater than T_0. At this point, the introduction of further energy into the system would produce additional particles, and not affect the kinetic energy (and so temperature) of the system. This was said by Hagedorn to show that this was a limiting physical temperature. It was later proposed by Cabibbo and Parisi [35] that, far from showing a temperature limit, the singularity instead signifies a second-order phase transition to a deconfined phase above the so-called 'critical temperature', T_c. As the temperature increases, large numbers of free colour charges are produced and fill the vacuum between quarks, effectively screening the strong interaction and reducing its range to such short scales that the strong force is no longer effective at the size of the hadron. This leads to a screening effect of the Coulomb term. In addition, the QCD string tension σ decreases as a function of temperature according to $\sigma/\sigma_0 = \sqrt{1 - (T/T_c)^2}$ [37]. For temperatures above T_c, the Coulomb term is therefore modified to a short-ranged Yukawa potential, and the string tension vanishes. This process is termed as Debye screening, analogous to the charge screening observed for the electromagnetic interaction in an electromagnetic plasma. The quark–antiquark potential is therefore reparameterised such that [38]

$$V_{\mathrm{QCD}}(r) = -\frac{4}{3}\frac{\alpha_{\mathrm{s}}}{r} e^{-\frac{r}{r_{\mathrm{D}}}}, \tag{2.5}$$

where r_{D} is the Debye radius. At separations larger than r_{D}, a pair of quarks can be treated as being unbound, and the presence of free colour charges shortens this length. Deconfinement thus occurs when r_{D} reaches values $\ll 1$ fm. This state of matter, a large, hot system in which quarks and gluons are effectively unbound, is known as a Quark–Gluon Plasma (QGP). It is postulated that these conditions match those of the early Universe, a few microseconds after the start of the Big Bang.

The presence of a phase transition above the critical temperature has been confirmed by calculations of QCD on the lattice (a non-perturbative approach, applicable

to low energy scales, wherein space-time is discretised into a grid with quarks at the nodes, connected by gluon fields). Thermodynamic quantities may be computed under this approach, for example the energy density ε, pressure p and entropy density s/T^3:

$$\varepsilon = \frac{T^2}{V}\left(\frac{\partial Z}{\partial V}\right)_V, \tag{2.6}$$

$$p = T\left(\frac{\partial \ln Z}{\partial V}\right)_V, \text{ and} \tag{2.7}$$

$$\frac{s}{T^3} = \frac{\varepsilon + p}{T^4}, \tag{2.8}$$

where V is the volume, related to the cube of the lattice spacing a, and Z is the partition function [39]. For an ideal gas (i.e. no interactions between constituent particles) of fermions and bosons, ε is given by

$$\varepsilon_{\text{fermion}} = n_{\text{dof}}\frac{7\pi^2}{240}T^4, \text{ and} \tag{2.9}$$

$$\varepsilon_{\text{boson}} = n_{\text{dof}}\frac{\pi^2}{30}T^4, \tag{2.10}$$

where n_{dof} is the relevant number of degrees of freedom. The factor of 7/8 between the two ϵ values accounts for the differences between Fermi-Dirac statistics for fermions and Bose-Einstein statistics for bosons. In a QGP system, the fermions are quarks, whose n_{dof} is given by $n_{\text{spin}} \times n_{\text{colour}} \times n_{\text{charge}} \times n_{\text{flavour}} = 2 \times 3 \times 2 \times n_{\text{flavour}}$, where the active number of flavours varies depending on the calculation framework. For $\varepsilon_{\text{boson}}$, i.e. the gluon energy density, n_{dof} is given by $n_{\text{spin}} \times n_{\text{colour}} = 2 \times 8$. By contrast, when the partons are confined into hadrons, as in a pion gas, the number of degrees of freedom is three (due to pions having no spin and three possible isospin states: π^+, π^0 and π^-).

In the ideal gas limit, the system is expected to obey the Stefan–Boltzmann law, whereby the quantities $\frac{\varepsilon}{T^4}$, $\frac{3p}{T^4}$ and $\frac{3s}{4T^3}$ become constant as a function of temperature. Lattice calculations of these quantities are shown in Fig. 2.4 as a function of temperature, for the '2 + 1' quark flavour scheme (two light quark flavours of equal mass, and one heavier flavour) [40]. The calculations are shown by shaded bands; the darker solid curves give further predictions under the approximation of a hadronic resonance gas (HRG). The vertical band at $T = 154 \pm 9\,\text{MeV}$ corresponds to the crossover region around the critical temperature T_{c}, and the horizontal line at $95\pi^2/60$ is the Stefan–Boltzmann limit for the energy density for an ideal gas. Although the calculated energy density asymptotically approaches this limit, it does not reach it within the shown temperature scale due to residual interactions between quarks and gluons in
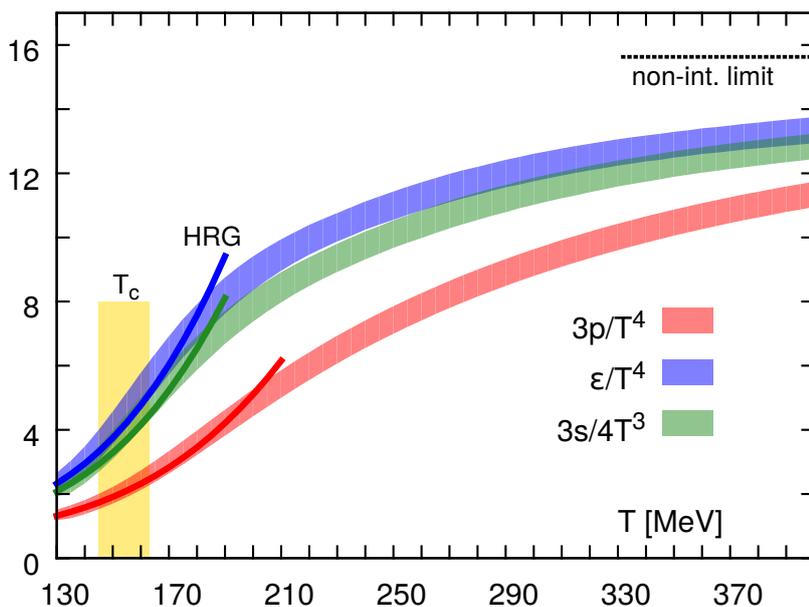
Figure 2.4: Calculations of $3p/T^4$, ε/T^4 and $3s/4T^3$ on the lattice as a function of temperature, shown alongside calculations under a hadron resonance gas (HRG) approximation. The critical temperature $T_{\mathrm{c}} = 154 \pm 9\,\mathrm{MeV}$ is shown as a vertical shaded band, and the Stefan–Boltzmann limit for the energy density of an ideal gas ($95\pi^2/60$) as a horizontal dashed line. Figure from [40].

the system. The central predictions under the two approximations are consistent for $T < 150\,\mathrm{MeV}$, but the calculations begin to diverge for temperatures above $170\,\mathrm{MeV}$. The onset of the steep increase in the energy density at the marked temperature is attributed to the number of degrees of freedom increasing from the hadronic to the partonic level due to the phase transition to a QGP.

The evolution of a QGP produced in an ultrarelativistic heavy-ion collision is illustrated in Fig. 2.5. Here, the collision between the pair of beams occurs at $(z = 0, t = 0)$, where z is the beam axis (see Section 2.2) and t is the time. The light cone of the collision, defined by lines of constant proper time ($\tau = \sqrt{t^2 - z^2}$), is indicated by diagonal black lines. Some important time and temperature scales are also indicated: the thermalisation time τ_0, the critical temperature T_{c}, the chemical freeze-out temperature T_{ch} and the kinetic freeze-out temperature T_{fo}.

The system thermalises quickly, reaching local thermal equilibrium at a proper time $\tau_0 \approx 1\,\mathrm{fm}/c$ after the collision. The medium has a very small shear-viscosity-to-entropy-density ratio, and may be described according to viscous relativistic hydrodynamics.

As the QGP cools, it undergoes a series of sequential (or potentially simultaneous) freeze-outs. Below T_{c}, the system regresses to a gas of hadrons. Once it reaches the chemical freeze-out temperature T_{ch}, inelastic collisions between the constituents cease
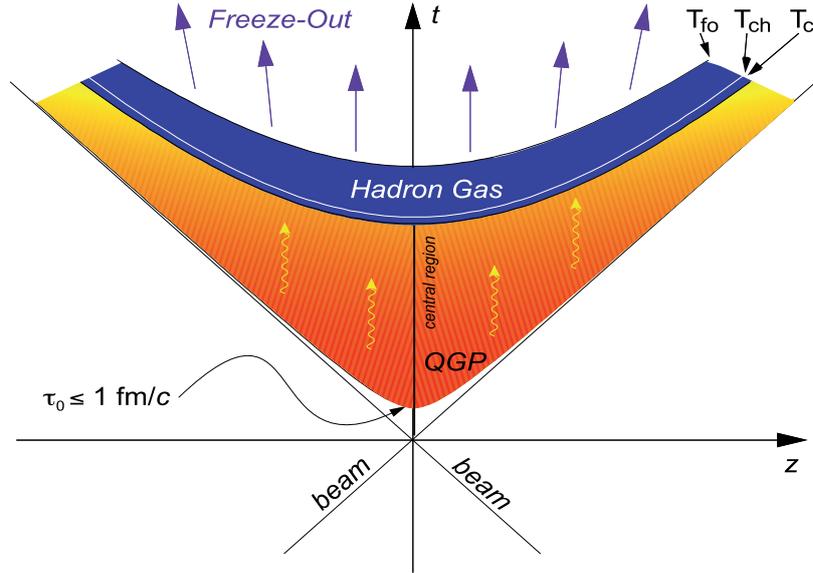
Figure 2.5: Evolution of the QGP in a heavy-ion collision. The horizontal axis z refers to the beam direction; the vertical axis t is the time. The diagonal black lines represent the light cone of the collision. The points at which the system reaches critical temperature (T_c), chemical freeze-out (T_{ch}) and kinetic freeze-out (T_{fo}) are also marked. Figure from [41].

and the hadronic composition becomes fixed. Finally, once the system reaches T_{fo}, it is said to be in a state of kinetic freeze-out, where no further elastic interactions between the constituent hadrons occur and the momenta of the hadrons become fixed. The overall lifetime of the system is strongly dependent on the energy density at the start of the collision, with an average QGP lifetime in Au–Au collisions at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ of $\tau_{QGP}^{RHIC} = 6.2$, and in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$ of $\tau_{QGP}^{LHC} = 14\,\mathrm{fm}/c$ for a critical temperature $T_c \approx 180\,\mathrm{MeV}$ [42].

Key observables in this system include thermal photons, to which the QGP is transparent as they do not participate in the strong interaction, and so their spectra can be used as a gauge of the temperature; light-flavour hadrons at low momenta, which are mainly produced at later stages of the evolution and as such are able to describe the temperature evolution and hydrodynamics of the QGP; and so-called 'hard probes' such as jets and heavy quarks, which are only produced in the initial hard scatterings of the heavy-ion collision, and so give better insight into the early stages of the collision system.

2.2   Hadronic colliders

In high-energy particle and nuclear physics experiments, particles are collided at ultrarelativistic energies in order to produce interactions between them. This section briefly outlines some of the common terminology used in such experiments.

In a collider experiment, a beam of charged particles is accelerated close to the speed of light using radio-frequency electric fields for acceleration and magnetic fields for the steering and focussing of the beams. At these velocities, relativistic effects come into play for the kinematics of the system. The energy available for particle production in the colliding system is defined by the centre-of-mass energy, $\sqrt{s}$. In a fixed-target experiment, for example a proton beam colliding with a hydrogen target, the centre-of-mass energy according to special relativity is $\sqrt{s} = \sqrt{2m_{\mathrm{p}}^2 c^4 + 2m_{\mathrm{p}} c^2 E_{\mathrm{beam}}}$, where m_{p} is the mass of the proton and E_{beam} is the total energy of the accelerated particle in the lab frame, according to $E = \sqrt{p^2 c^2 + m^2 c^4}$. By contrast, a symmetric collision system with head-on collisions between two beams of particles of equal energy will have a centre-of-mass energy of $2E_{\mathrm{beam}}$, meaning a much larger available energy than in the fixed-target case.

The coordinate system of the collision is usually defined in a cylindrical fashion, using r, ϕ and z as coordinates. The colliding particles travel along the z axis and nominally meet in the central point of the detector, $z = 0$. The radial distance outwards from the z-axis is denoted by r, and ϕ is the azimuthal angle about the z-axis.

Particle production rates in colliders are generally studied differentially, as a function of p_{T} and y. The p_{T} of a particle is its transverse momentum, i.e. the component of its momentum perpendicular to the beam axis, and y is the particle's rapidity. The rapidity serves as a measure of the proportion of the particle's momentum in the direction of the beam, and is additive under Lorentz transformations. It is calculated as

$$y = \frac{1}{2} \ln \frac{E + p_z c}{E - p_z c}, \tag{2.11}$$

where p_z is the longitudinal (z-)component of the particle's momentum and E is its total energy. A particle with $y = 0$ in the centre-of-mass system, or close to this value, is termed as being at 'mid-rapidity', and large positive or negative y is defined as 'forward' or 'backward' rapidity. In a symmetric collision system, the definition of the positive and negative direction is arbitrary; in an asymmetric collision such as p–Pb, it depends on the rapidity shift of the centre-of-mass system with respect to the lab system, with particles travelling in the same direction as the centre-of-mass system being termed as 'forward'. An analogous quantity known as the 'pseudorapidity', η, may also be defined in terms of the zenith angle θ with respect to the beam axis, in the lab frame:

$$\eta_{\text{lab}} = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right). \tag{2.12}$$

The advantage of using η is that it is a purely geometrical quantity, independent of the particle kinematics, and so it is commonly used to define the positions of detectors within an experiment. The naming convention for small and large η follows that for y.

In heavy-ion collisions, rather than referring to the total energy of each nucleus, the collision energy is defined as the energy per colliding nucleon–nucleon pair, $\sqrt{s_{\text{NN}}}$. The maximum attainable $\sqrt{s_{\text{NN}}}$ for heavy-ion collisions is usually smaller than the maximum $\sqrt{s}$ for proton–proton collisions, as the ability to accelerate and steer these heavier particles by electric and magnetic fields is affected by their smaller charge-to-mass ratio with respect to protons. Proton–proton collisions may also be taken at a matching $\sqrt{s}$ in order to compare production rates in nucleus–nucleus collisions with the assumption of a simple superposition of pp collisions.

Heavy-ion collisions may also be characterised by their centrality. The centrality is characterised by the impact parameter b, defined as the distance of closest approach between the centres of the two nuclei. The centrality is usually either classified into percentiles, with 0% signifying the most central (head-on) collisions, and 100% indicating the most peripheral (glancing). This may be estimated based on percentiles of the distribution of the charged-particle multiplicity, with the highest produced multiplicities corresponding to the most central collisions. It may also be classified in terms of N_{coll}, the number of binary collisions in the system, which is estimated by a Glauber model approximation [43, 44].

2.3   Heavy-flavour particles

Heavy-flavour hadrons, i.e. those containing charm and beauty quarks, provide a unique probe into the properties of the medium formed in heavy-ion collisions. The top quark is also considered to belong to the heavy-flavour quarks, but it decays before it can hadronise due to the large difference between the top and beauty masses; it therefore is not mentioned in the following.

Heavy quarks are largely produced in hard scatterings according to the processes outlined as Feynman diagrams in Fig. 2.6. These processes, from left to right, are known as gluon fusion, gluon splitting, and flavour excitation. Gluon fusion is a standard leading-order (LO) pair production process, in which a pair of gluons, one from each of the colliding nucleons, fuse to produce a pair of quarks. In gluon splitting (next-to-leading order, NLO), a gluon from one nucleon scatters off a gluon from the other nucleon and then splits into a $Q\bar{Q}$ pair. Finally, in flavour excitation (NLO),

an off-shell quark originating from a virtual gluon splitting process in one nucleon undergoes a hard scattering with a gluon from the other nucleon, and is brought onto the mass shell as a result.
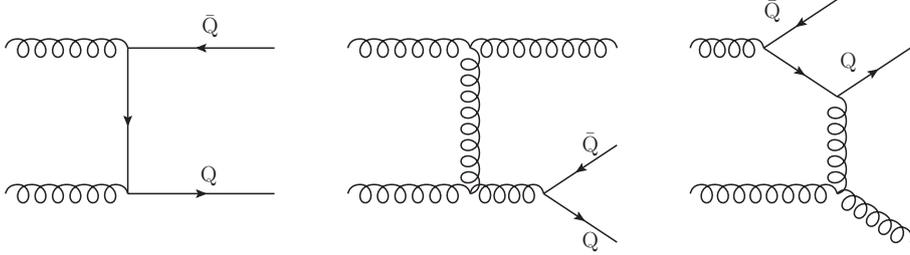
Figure 2.6: Feynman diagrams of leading-order and next-to-leading-order processes for heavy-quark (Q) production in hadronic collisions. The processes illustrated are: (left) gluon fusion, (middle) gluon splitting and (right) flavour excitation.

Charm, being the lightest of the heavy quarks, is the most commonly produced in ultrarelativistic hadronic collisions. Due to the strong force obeying flavour conservation, charm quarks produced in gluon scatterings are always produced in $c\bar{c}$ pairs, which may either hadronise into flavour-neutral charmonia (bound $|c\bar{c}\rangle$ states) or into open-charmed baryons and mesons (particles with a non-zero charm quantum number, most commonly D mesons and Λ_c^+ baryons). This hadronisation occurs shortly after the quark is produced, through a process known as fragmentation. This process is soft, meaning that it cannot be modelled perturbatively, and so its parameters must be determined from experimental results. It is possible to perform this parameterisation from simple quantum mechanical first principles, for example using the Peterson function [45]:

$$D_{\mathrm{Q}}^{\mathrm{H}}(z) \propto \frac{1}{z[1 - (1/z) - \varepsilon_{\mathrm{Q}}/(1-z)]^2}, \qquad (2.13)$$

where 'H' and 'Q' refer respectively to the hadron and the heavy quark; z is the ratio of the hadron's total energy to that of the heavy quark ($z = E_{\mathrm{H}}/E_{\mathrm{Q}}$), and ε_Q is the fragmentation parameter, typically defined as the ratio of the light-quark mass to the heavy-quark mass in a meson. Typical values of this parameter are $\sim$0.006 for beauty quarks, or $\sim$0.06 for charm quarks [46]. The Peterson function is shown in Fig. 2.7 for these values.

The fragmentation function is characterised by a peak at $z \approx 1 - \sqrt{\varepsilon_{\mathrm{Q}}}$, which indicates the most likely value of z for a produced hadron. The kinematic explanation for the difference between the two functions is the difference in mass between the charm and beauty quark: as the beauty quark is much heavier than charm, the momentum of the final hadron is far more dependent on that of the beauty quark in a B meson than on the charm quark in a D meson. This varying level of influence
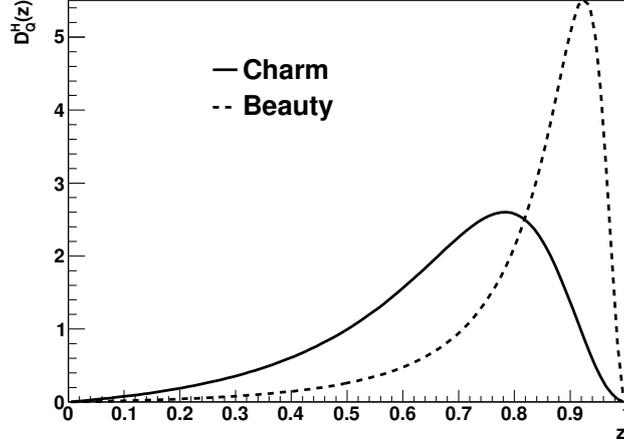
Figure 2.7: The Peterson fragmentation functions for charm and beauty to all hadrons, with the fragmentation parameters $\varepsilon_c = 0.06$ and $\varepsilon_b = 0.006$. Both curves are normalised to unity. Figure from [47].

also broadens the distribution for charm quarks with respect to beauty quarks. As it cannot be calculated perturbatively, the distribution must be tuned to experimental parameters such as x_Q, the energy of a weakly decaying hadron containing the heavy quark Q, normalised to the beam energy. This determination was done at the Z-resonance and gave peak values in the fragmentation for $\langle x_b \rangle = 0.702 \pm 0.008$ and $\langle x_c \rangle = 0.484 \pm 0.008$ [48]. The fragmentation function may also be parameterised by other approaches, including the Kartvelishvili–Likhoded–Petrov function [49], the Bowler function [50], and the Collins–Spiller function [51].

The fragmentation fractions (the proportions of charm quarks fragmenting to different hadron species) are outlined in Fig. 2.8. Approximately 1% of charm quarks hadronise into charmonia, and ∼99% to open-charmed particles. The yields of heavy quarkonia (bound $|c\bar{c}\rangle$ and $|b\bar{b}\rangle$ states) are of particular importance in heavy-ion collisions, as the suppression of charmonium production at high temperatures has long been considered to be a characteristic signature of the formation of a QGP [52]. While the J/ψ meson is relatively strongly bound, and is expected to remain bound at temperatures up to ∼$2T_c$, the excited states have weaker binding energies and so are more likely to sequentially dissociate at much lower temperatures (∼1.1–$1.2T_c$) [53,54]. In addition, calculations under the Statistical Model approach [55,56] have shown that J/ψ hadronisation proceeds via statistical processes; a precise determination of the number of $c\bar{c}$ pairs produced in the collisions is required in order to make reliable predictions in this framework. As the proportion of $c\bar{c}$ pairs that hadronise into charmonia is roughly 1%, studies of open charm production therefore serve as vital baselines for these processes, as they provide a much larger statistical basis for precise

measurements of the production cross section of c$\bar{\text{c}}$ pairs.
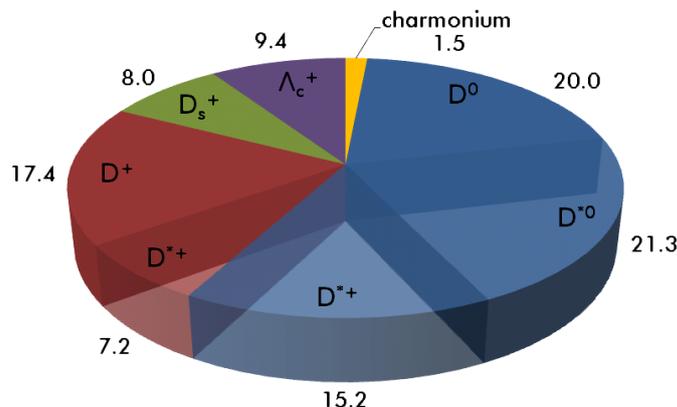
Figure 2.8: Fragmentation fractions of charm quarks to mesons and baryons. The blue region refers to D^0 mesons in the final state; the red region refers to D^+ mesons in the final state. The solid segments are prompt mesons, and the transparent segments represent the feed-down from the labelled resonant states (D^{*+} and D^{*0}). Figure from [57].

Open heavy-flavour production can be studied in various ways. In ALICE, this is done either through the semileptonic decays of D and B mesons to electrons (muons) at mid-(forward) rapidity, or through the kinematic reconstruction of the hadronic decays of D mesons at mid-rapidity. Each system has its own advantages: analyses of heavy-flavour electrons allow the contribution from beauty hadrons to be measured separately from the inclusive spectrum by means of a template fit to the impact parameter distribution [58], and analyses of heavy-flavour decay muons at large rapidities allow smaller regions in Bjorken-x to be probed [59].

The main advantage of fully reconstructing the hadronic decays is that it allows the full kinematic information of the original particle to be reconstructed. In ALICE, this is done for the D^0, D^+, D^{*+} and D_s^+ mesons, and Λ_c^+ baryons (and their respective charge conjugates). The analysed decay channels, as well as their mean lifetimes and branching fractions, are listed in Tab. 2.3, and the Feynman diagrams of the $D^0 \to K^-\pi^+$ and $D^{*+} \to D^0\pi^+$ decay processes are shown in Fig. 2.9. The $\overline{D}^0 \to K^-\pi^+$ channel also exists, but is doubly Cabibbo-suppressed and so its branching fraction is negligible with respect to $D^0 \to K^-\pi^+$.

Apart from $D^{*+} \to D^0\pi^+$, which is a strong decay that occurs almost directly after its production ($c\tau \approx 2.4\,\text{pm}$; $\tau \approx 8 \times 10^{-21}$ s, calculated from its decay width $\Gamma = 83.4 \pm 1.8\,\text{keV}$ [23] using $\tau = \hbar/\Gamma$), the remaining processes are weak decays that occur some time after the initial fragmentation. The finite decay length of these processes can be exploited with the use of a high-resolution tracking detector, as the decay particles (also referred to as 'daughters') originate from a secondary vertex,

Species	Decay channel	$\mathcal{B}$	$c\tau$ (μm)
D^0	$\mathrm{D}^0 \to \mathrm{K}^-\pi^+$	$3.93 \pm 0.04\%$	122.9
	$\mathrm{D}^0 \to \mathrm{K}^+\pi^-$	$(1.399 \pm 0.027) \times 10^{-4}$	
D^+	$\mathrm{D}^+ \to \mathrm{K}^-\pi^+\pi^+$	$9.46 \pm 0.24\%$	311.8
D^{*+}	$\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+$	$67.7 \pm 0.5\%$	$\sim 2.4 \times 10^{-6}$
	$(\to \mathrm{K}^-\pi^+\pi^+)$	$(2.66 \pm 0.03\%)$	
D_s^+	$\mathrm{D}_\mathrm{s}^+ \to \phi\pi^+ \to \mathrm{K}^-\mathrm{K}^+\pi^+$	$2.28 \pm 0.12\%$	149.9
Λ_c^+	$\Lambda_\mathrm{c}^+ \to \mathrm{p}\mathrm{K}^-\pi^+$	$6.84^{+0.32}_{-0.40}\%$	$59.9\,\mu$m
	$\Lambda_\mathrm{c}^+ \to \mathrm{p}\mathrm{K}_\mathrm{S}^0$	$3.21 \pm 0.30\%$	

Table 2.3: Open heavy-flavour decay channels measured by the ALICE Collaboration, listed with their branching ratios $\mathcal{B}$ and mean lifetimes $c\tau$ (in μm). All respective charge conjugates are also measured. The D^{*+} decay length is computed from its decay width $\Gamma = 83.4 \pm 1.8\,$keV. Values from [23].
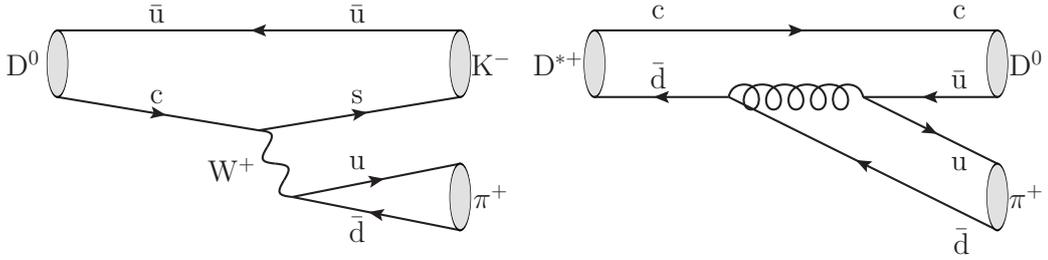
Figure 2.9: Feynman diagrams of the $\mathrm{D}^0 \to \mathrm{K}^-\pi^+$ (left) and $\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+$ (right) decays.

which has some displacement from the primary interaction vertex. The tracking and reconstruction of secondary vertices is discussed in Section 3.3, and the reconstruction procedure for D mesons in ALICE is covered in further detail in Section 4.

While measurements of D-meson production are well established in ALICE, the study of Λ_c^+ baryons remains a challenge at present. In pp collisions, the Λ_c^+ baryon accounts for a significant proportion of charm-quark hadronisation ($\sim 9.4\%$), and its measurement in Pb–Pb collisions may shed light on additional hadronisation mechanisms within the QGP. Previous measurements of light-flavour baryon-to-meson ratios (p/π and $\Lambda/\mathrm{K}_\mathrm{S}^0$) have been made in Au–Au and Pb–Pb collisions at RHIC and the LHC [2,3], and have shown an enhancement above unity at intermediate p_T, with values larger than those measured in pp collisions. This enhancement was also observed to be more significant for the strange baryon-to-meson ratio than the non-strange one. One possible explanation for this difference between pp and heavy-ion collisions is the effect of quark coalescence on hadronisation within the QGP [4]. Studies

of the production of Λ_c^+ baryons in Pb–Pb collisions would allow these baryon-to-meson ratio measurements to be extended into the heavy-flavour sector, in order to determine whether such effects also play a significant role for heavy-quark hadronisation. However, while the Λ_c^+ baryon has been observed in proton–proton collisions at $\sqrt{s} = 7\,\text{TeV}$ by the LHCb Collaboration [60], it remains elusive in heavy-ion collisions at the LHC due to its short decay length and the limited spatial resolutions of tracking detectors. This greatly motivates the need for improvements in particle identification (PID) techniques so that the effects of combinatorial background can be reduced in channels such as $\Lambda_c^+ \to \text{p}\text{K}^-\pi^+$, where selections on the decay geometry alone are not sufficient. A trial measurement of Λ_c^+ production in this channel using a novel Bayesian PID approach was presented by ALICE in [6], and is discussed further in Section 5.5.

The rate of production of a particle species is characterised by its cross section, measured in barns ($1\,\text{b} = 10^{-28}\,\text{m}^2$). The production cross section of D mesons can be computed from measurements of their yields in proton–proton collisions, as follows:

$$\left.\frac{\mathrm{d}\sigma^{\mathrm{D}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{|y|<0.5} = \frac{1}{2}\frac{1}{\Delta y_{\mathrm{fid}}\Delta p_{\mathrm{T}}}\frac{f_{\mathrm{prompt}}(p_{\mathrm{T}})\cdot N_{\mathrm{raw}}^{\mathrm{D}}(p_{\mathrm{T}})\big|_{|y|<y_{\mathrm{fid}}}}{(\mathrm{Acc}\times\varepsilon)_{\mathrm{prompt}}(p_{\mathrm{T}})\cdot\mathcal{B}\cdot\mathcal{L}_{\mathrm{int}}}. \tag{2.14}$$

Here, σ^{D} is the production cross section, p_{T} is the transverse momentum of the D meson, f_{prompt} is the fraction of D mesons that are produced from a prompt $c\bar{c}$ pair as opposed to feed-down from the decay of a B meson, $N_{\mathrm{raw}}^{\mathrm{D}}$ is the raw yield of D mesons, $\mathrm{Acc}\times\varepsilon$ is the acceptance-times-efficiency factor determined in Monte Carlo simulations, $\mathcal{B}$ is the branching ratio for the measured decay channel, and $\mathcal{L}_{\mathrm{int}}$ is the integrated luminosity for inelastic pp collisions in the dataset under consideration. y_{fid} is referred to as the fiducial acceptance in rapidity, which varies polynomially from 0.5 at low p_{T} to 0.8 for $p_{\mathrm{T}} > 5\,\text{GeV}/c$. The factor of $\Delta y_{\mathrm{fid}}(= 2y_{\mathrm{fid}})$ in the denominator corrects the cross section to $|y| < 0.5$ over the whole p_{T} range; this equation thus holds under the assumption that the rapidity distribution of $\frac{\mathrm{d}\sigma}{\mathrm{d}y}$ is flat at mid-rapidity.

The large masses of heavy quarks ($m_{\mathrm{Q}} \gg \Lambda_{\mathrm{QCD}}$, see Section 2.1) mean that they can be studied under perturbative frameworks, and thus they allow direct comparisons to be made between data and perturbative QCD (pQCD) models. One such framework is the Fixed Order plus Next-to-Leading Logarithm approach (FONLL) [61,62]. This approach attempts to reduce the uncertainties of next-to-leading-order calculations at high p_{T} by combining a matched calculation of the LO and NLO Feynman diagrams of heavy-quark production processes (see Fig. 2.6) with a resummation of the next-to-leading logarithms, computed with the nucleon parton distribution functions according to the following simplified formula [63]:

$$\sigma_Q^{\text{FONLL}}(p_T, m) = \text{FO} + (\text{RS} - \text{FOM0})G(p_T, m), \qquad \textbf{(2.15)}$$

where m is the heavy-quark mass, σ_Q^{FONLL} is the production cross section of the heavy quark, FO is the NLO calculation including the quark mass computed in the MNR framework [64], FOM0 is the massless limit of this NLO calculation, and RS is the resummed calculation at the massless limit. The factor of $G(p_T, m)$ is a suppression term, $G = p_T^2/(p_T^2 + 5^2 m^2)$, which ensures that the massless limit is only considered at high p_T ($p_T > 5m$).
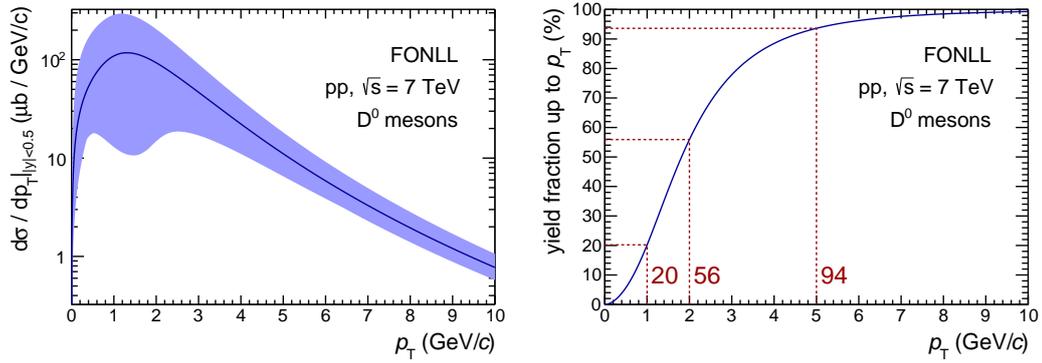
Figure 2.10: FONLL calculations [61,62] of the p_T-differential D^0-meson production cross section at mid-rapidity, as a function of p_T. Left: p_T spectrum; right: cumulative fraction of the total cross section as a function of p_T. The calculations are shown as a curve, and their uncertainties as a shaded band. The red numbers and dashed lines mark percentiles of the cumulative distribution, integrated from $p_T = 0\,\text{GeV}/c$. Figures from [65].

Calculations of the D^0 production cross section under this approach are shown in Fig. 2.10 as a p_T spectrum and as a cumulative integral. The left-hand plot shows that a large proportion of the cross section lies at low p_T, and that the theoretical uncertainties also grow in this region. This highlights the importance of precise measurements of charm production at low p_T, as if the theoretical uncertainties are large, the systematic uncertainties on any extrapolation of the data into this region will increase accordingly. It can be seen from the right-hand panel of Fig. 2.10 that approximately 20% of the charm production cross section sits in the range $0 < p_T < 1\,\text{GeV}/c$; a further 36% is predicted to be within $1 < p_T < 2\,\text{GeV}/c$. Current ALICE measurements of D-meson production at mid-rapidity extend to $p_T = 1\,\text{GeV}/c$; the standard topological selections become less effective at momenta lower than this due to the spatial resolution of the Inner Tracking System. Further studies are ongoing in the ALICE Collaboration to improve the p_T reach of D^0 measurements down to $p_T = 0\,\text{GeV}/c$ by relaxing the topological selections in order to measure the D-meson yield in this region [65].

Figure 2.11 shows the measured total c$\bar{\text{c}}$ production cross section in pp and proton–nucleus collisions as a function of the centre-of-mass energy $\sqrt{s}$. The p–A results are scaled by the number of nucleons in the nucleus. ALICE measurements of the open-charm production cross section in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ [66] and $\sqrt{s} = 2.76\,\text{TeV}$ [67] for $1 < p_\text{T} < 24\,\text{GeV}/c$ were extrapolated to the full phase space to determine the total c$\bar{\text{c}}$ production cross section [47, 66].
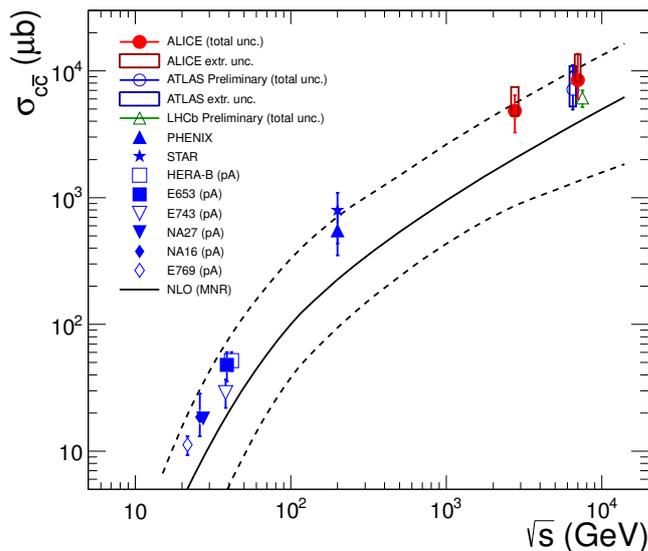
Figure 2.11: Total charm–anticharm production cross section in pp collisions as a function of the collision energy $\sqrt{s}$. Results from a variety of experiments [46, 60, 66–70] are compared with MNR pQCD calculations and their uncertainty bands [64]. Figure from [67].

The results at $\sqrt{s} = 7\,\text{TeV}$ were found to be consistent with measurements by the ATLAS and LHCb collaborations at the same $\sqrt{s}$ [60,70]. The results at $\sqrt{s} = 2.76\,\text{TeV}$ and $\sqrt{s} = 7\,\text{TeV}$ [66,67] also showed consistency with the upper band predicted by pQCD calculations made in the MNR framework [64], which is the next-to-leading-order precursor of FONLL used in Eq. (2.15). The MNR and FONLL frameworks are equivalent in their calculations of the total charm cross section, with the FONLL framework giving smaller uncertainties than MNR at higher p_T due to the resummations at the massless limit; at low p_T, which is more relevant for the total cross section, the two frameworks are approximately equivalent [47].

The agreement with pQCD calculations was also confirmed by comparison with other experiments at a wide range of collision energies, also shown in Fig. 2.11. All of the experimental measurements are found in the upper uncertainty band of the model calculations, potentially implying that the cross section calculated under the NLO approach is artificially suppressed. The FONLL framework was shown to

similarly underestimate the cross section of D-meson production in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ when compared to measurements by the LHCb Collaboration in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ [71]. One possible explanation for this is that the charm quark mass employed by the central calculations ($m_\mathrm{c} = 1.5\,\mathrm{GeV}/c^2$) is too large, which would lead to a reduced gluon splitting cross section at the leading order. This has been confirmed by more recent analyses of deep inelastic scattering at HERA using a minimal-subtraction-scheme ($\overline{\mathrm{MS}}$) approach for the running of the charm quark mass [72,73], giving a charm-quark mass of $\sim 1.26\,\mathrm{GeV}/c^2$ [74].

A further interesting quantity to measure for charm in pp collisions, and a particular motivation for measuring D^{*+} production, is the P_v, the ratio of charged $\left(\left|\mathrm{c\bar{d}}\right\rangle \text{ and } \left|\mathrm{\bar{c}d}\right\rangle\right)$ D mesons that are produced in a vector state (spin $= 1$) to the sum of those produced in a vector or pseudoscalar state (spin $= 0$). P_v is expressed as the ratio of the D^{*+} cross section to the sum of D^{*+} and the cross section of D^+ mesons not originating from D^{*+} decays:

$$P_\mathrm{v} = \frac{\sigma_\mathrm{tot}^{\mathrm{D}^{*+}}}{\sigma_\mathrm{tot}^{\mathrm{D}^{*+}} + \sigma_\mathrm{tot}^{\mathrm{D}^+} - \sigma_\mathrm{tot}^{\mathrm{D}^{*+}} \cdot (1 - \mathcal{B}_{\mathrm{D}^{*+} \to \mathrm{D}^0 \pi^+})} = \frac{\sigma_\mathrm{tot}^{\mathrm{D}^{*+}}}{\sigma_\mathrm{tot}^{\mathrm{D}^+} + \sigma_\mathrm{tot}^{\mathrm{D}^{*+}} \cdot \mathcal{B}_{\mathrm{D}^{*+} \to \mathrm{D}^0 \pi^+}}, \tag{2.16}$$

where σ_tot^i is the total production cross section for the particle i, and $\mathcal{B}_{\mathrm{D}^{*+} \to \mathrm{D}^0 \pi^+} = 67.7 \pm 0.5\%$ [23] is the branching ratio for a D^{*+} meson to decay with a D^0 meson in the final state. The subtraction of $\mathcal{B}_{\mathrm{D}^{*+} \to \mathrm{D}^0 \pi^+}$ from unity accounts for all of the decay channels containing a D^+ meson in the final state. Naïve spin counting would lead to a value of $P_\mathrm{v} = 3/(3+1) = 0.75$.

Global measurements of P_v can be seen in Fig. 2.12 for a variety of collider experiments at different energies [66,67,70,75–80]. All of the measurements are consistent with one another within the experimental uncertainties. The global mean (plotted as a yellow vertical band) was computed as a weighted average, with the inverse of the uncertainties as weights, and found to be $P_\mathrm{v}^\mathrm{avg} = 0.60 \pm 0.01$. This observation differs from the spin-counting hypothesis, which depends on the assumption of an infinitely large heavy-quark mass in heavy-quark effective theory; this would lead to a negligible mass difference between the regular and excited D-meson states. In reality the difference between the D^{*+} and D^+ meson masses is $\sim 7.5\%$. As the mass of the beauty quark is much larger than that of charm, future measurements of B mesons may be able to better test this theory, as the difference between the regular and resonance B-meson states is on the order of 9‰.

In addition, the charm P_v allows tests to be made of the temperature-dependent statistical model of hadronisation [55,56] and combinations of the Lund fragmentation functions with Clebsch–Gordan coupling [81]. The statistical model predicts a
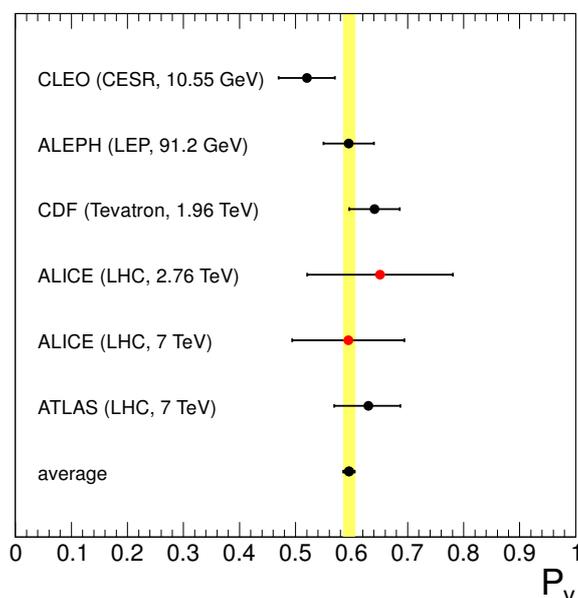
Figure 2.12: P_v measurements for D^{*+} and D^+ mesons at a variety of collider experiments [66, 67, 70, 75–80]. The global weighted average, $P_\mathrm{v}^\mathrm{avg} = 0.60 \pm 0.01$, is plotted as a yellow shaded band. Figure from [67].

P_v value of 0.58 ± 0.13 for a temperature of 164 MeV, based on the calculation of the ratio of the D^{*+} yield to the D^+ yield as $3 \cdot (m_{D^{*+}}/m_{D^+})^2 \cdot \exp\left(-(m_{D^{*+}} - m_{D^+})/T\right)$. The Lund function makes its estimate based on an exponential in terms of the difference between the D^{*+} and D^+ masses, which serves as a correction to a spin counting approach applied using Clebsch–Gordan coefficients to account for the coupling between the quark–antiquark pair and the final bound state. Under this framework, the predicted P_v is 0.63. Both of these models differ strongly from the spin counting hypothesis, and are consistent with the measured value for P_v.

2.4   Open charm production in Pb–Pb collisions at the LHC

The heavy quarks have masses much larger than the crossover temperature of the QGP ($\approx 154\,\mathrm{MeV}$, see Section 2.1), meaning that they can only be produced during initial hard scatterings in heavy-ion collisions, and not thermally produced at later times within the QGP. In addition, as the heavy quarks obtain the majority of their mass through their coupling to the Higgs field rather than confinement (see Tab. 2.2 and Fig. 2.1), they retain their masses under the partial restoration of chiral symmetry that occurs in the QGP. Conversely, the lighter quarks lose the majority of their masses under these conditions. Therefore, kinematic studies of heavy-flavour hadrons provide

insight into the full evolution of the QGP. They can therefore be used to gauge the process of energy loss within the medium, as they interact strongly with the rest of the medium as they traverse it. As flavour is conserved in strong interactions and the lifetime of the QGP is short, the number of charm quarks remains constant over the lifetime of the QGP, and thus they experience the full evolution of the system. This makes them vital probes of the conditions formed in high-energy Pb–Pb collisions.

The nuclear modification factor, R_{AA}, of D mesons is used as one avenue for studying the properties of the QGP. The R_{AA} is defined as the ratio of the measured yield in Pb–Pb collisions to that measured in pp collisions at the same $\sqrt{s}$, scaled by the number of binary collisions:

$$R_{AA} = \frac{1}{\langle T_{AA} \rangle} \cdot \frac{\mathrm{d}N_{AA}/\mathrm{d}p_T}{\mathrm{d}\sigma_{pp}/\mathrm{d}p_T}, \tag{2.17}$$

where T_{AA} is the nuclear overlap function, which is determined via Glauber model calculations and is proportional to the number of binary nucleon–nucleon collisions (N_{coll}) [43,44]; N_{AA} is the particle yield in Pb–Pb collisions; and σ_{pp} is the production cross section in pp collisions. An R_{AA} equal to unity implies no modification of the spectrum; an R_{AA} below unity implies a suppression with respect to pp collisions due to energy loss in the medium through collisional (from elastic scatterings with other quarks and gluons) or radiative (gluon bremsstrahlung) processes. Typically in ALICE, the pp reference data are taken at $\sqrt{s} = 7\,\mathrm{TeV}$ and scaled to $\sqrt{s} = 2.76\,\mathrm{TeV}$ using FONLL pQCD calculations, as the integrated luminosity of pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ is higher than that at $\sqrt{s} = 2.76\,\mathrm{TeV}$, leading to reduced statistical uncertainties. This procedure was tested for multiple different models for the scaling of D-meson cross sections at $\sqrt{s} = 7\,\mathrm{TeV}$ to existing data from the CDF collaboration at $\sqrt{s} = 1.96\,\mathrm{TeV}$ [47, 82], whereby it was found that FONLL gave the smallest systematic uncertainties overall. Further checks were made in [67] for the rescaling of pp data at $\sqrt{s} = 7\,\mathrm{TeV}$ to $\sqrt{s} = 2.76\,\mathrm{TeV}$, where the rescaled results were found to be within 5–10% of the 2.76 TeV data.

The D-meson R_{AA} measured by ALICE is shown in Fig. 2.13 as a function of p_T and centrality. The left-hand panel shows the average for D^0, D^+ and D^{*+} mesons at $5 < p_T < 8\,\mathrm{GeV}/c$ as a function of N_{part}, compared with ALICE results for charged pions [83]. Here it can be seen that the degree of suppression increases with increasing centrality (and therefore energy density) for both particle species. A possible mass ordering is also seen, in that the central points for D mesons systematically show a slightly lesser suppression than those for pions. This behaviour is predicted by the so-called 'dead cone effect' [1], which lends mass and colour-charge dependences to energy loss in the medium, as the available phase space for gluon radiation is limited for heavier particles at intermediate momenta due to the conservation of angular momentum [84–87]. Due to this effect, one expects a sequential modification of the

energy loss, namely $\Delta E_{\rm b} < \Delta E_{\rm c} < \Delta E_{\rm u,d,s}$; however, the full extent of this effect cannot be adequately distinguished with the current experimental uncertainties.
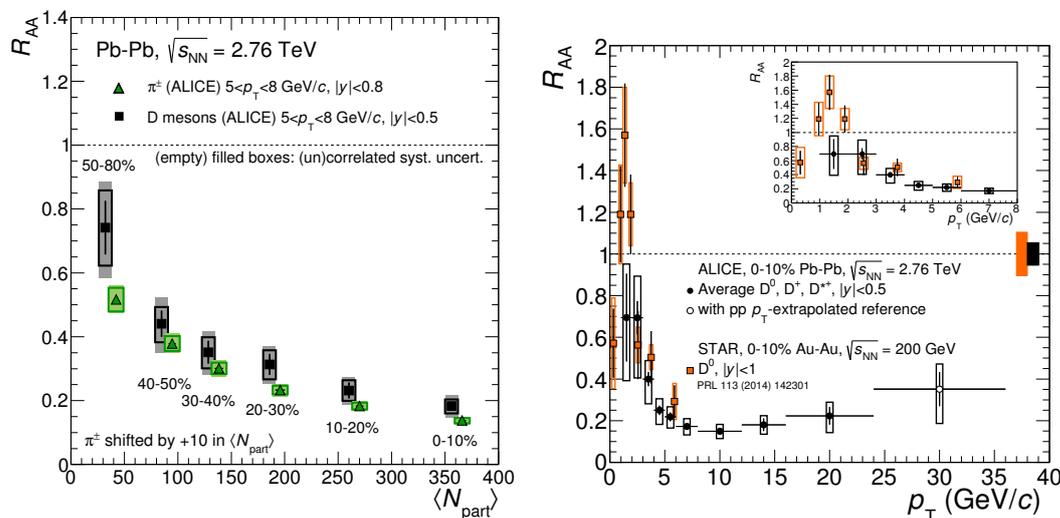
Figure 2.13: D-meson $R_{\rm AA}$ measured by the ALICE Collaboration in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 2.76\,{\rm TeV}$. Left: D mesons with $5 < p_{\rm T} < 8\,{\rm GeV}/c$ as a function of $N_{\rm part}$, compared with measurements for charged pions [83, 88]; right: as a function of $p_{\rm T}$ in the 10% most central collisions, compared with STAR measurements for $\rm D^0$ mesons in central Au–Au collisions at $\sqrt{s_{\rm NN}} = 200\,{\rm GeV}$ [89, 90]. Figures from [88, 90].

It is also useful to compare measurements from different colliders. Due to the lower collision energy ($\sqrt{s_{\rm NN}} = 200\,{\rm GeV}$), Au–Au collisions at RHIC contain a smaller energy density than Pb–Pb collisions at the LHC, and so the conditions of the produced medium differ. In particular, the mean lifetime of the QGP is shorter for Au–Au collisions at RHIC than for Pb–Pb collisions at the LHC, with measurements by the ALICE Collaboration confirming that the radius of the medium is 10–35% larger at the LHC than RHIC, and the mean lifetime is ∼40% larger [91]. The right-hand panel of Fig. 2.13 shows the $p_{\rm T}$ dependence of the D-meson $R_{\rm AA}$ for the 10% most central Pb–Pb collisions, this time plotted against STAR measurements of the $\rm D^0$-meson $R_{\rm AA}$ in Au–Au collisions at $\sqrt{s_{\rm NN}} = 200\,{\rm GeV}$ [89]. The inlay shows a zoomed-in view of the low-$p_{\rm T}$ region, $0 < p_{\rm T} < 8\,{\rm GeV}/c$. The strongest suppression of the D-meson yield is seen at intermediate $p_{\rm T}$ ($5 < p_{\rm T} < 15\,{\rm GeV}/c$), with an increasing $R_{\rm AA}$ at high $p_{\rm T}$. A peak in the $R_{\rm AA}$ can be seen at low $p_{\rm T}$ ($1 < p_{\rm T} < 2\,{\rm GeV}/c$) in the STAR measurements, and similar is implied by the ALICE points being equal for $1 < p_{\rm T} < 2\,{\rm GeV}/c$ and $2 < p_{\rm T} < 3\,{\rm GeV}/c$. This peak is explained by the effects of hydrodynamic flow and energy loss: particles at high-$p_{\rm T}$ lose a significant proportion of energy in the medium, and are therefore shifted to lower momenta; particles at very low $p_{\rm T}$ gain momentum through flow and are shifted to higher $p_{\rm T}$. The combination of these effects leads to a relative enhancement of the spectrum at intermediate $p_{\rm T}$.

Although the suppression is similar between the two systems at intermediate p_{T} ($2 < p_{\mathrm{T}} < 6\,\mathrm{GeV}/c$) within the uncertainties, a clear difference between the two experiments is seen at $1 < p_{\mathrm{T}} < 2\,\mathrm{GeV}/c$ despite large systematic uncertainties. Here, the ALICE data are suppressed by a factor of ~ 1.4, while the STAR data are enhanced by a factor of ~ 1.6. Going to ultra-low p_{T}, the STAR data again show a suppression. Future ALICE measurements of the D-meson R_{AA} with increased statistics and improved tracking will be vital in order to reduce the systematic uncertainties and extend the p_{T} reach to $0\,\mathrm{GeV}/c$ in order to better study these effects at higher energy densities.

A further observable in nucleus–nucleus collisions is particle flow, which gives an indication of collective expansion in the medium, offering further insight into QCD interactions in the expanding fireball. As heavy quarks are produced in the earliest stages, they participate in the entirety of the expansion, and also have a longer expected thermalisation time than lighter quarks [92]. This property can be used to examine the extent to which heavy quarks thermalise within the QGP, giving more information about the transport coefficients of the medium and the potential effects of quark coalescence on heavy-quark hadronisation [93, 94].

Flow arises due to a combination of pressure gradients in the medium and partonic energy loss within the QGP. It is most commonly studied in the context of anisotropic flow, which arises due to the spatial inhomogeneity of the collision system in non-central collisions. This is illustrated in the left-hand panel of Fig. 2.14, where the collision region (overlap between the two nuclei) is shown as an orange spheroid.

The impact parameter b between the two nuclei is the distance between their centres, and the grey grid is the 'reaction plane', defined as the plane that contains both the beam axis and the impact parameter. The spatial anisotropy of this system in the initial state translates to an anisotropy in momentum for produced particles in the final state, due to the larger pressure gradient induced in the reaction plane than outside it. The right-hand panel shows a simulation of the expansion of the system for an Au–Au collision at $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$, with an impact parameter $b = 7\,\mathrm{fm}$. It can be seen that although the system has a high level of spatial eccentricity shortly after the collision ($3.2\,\mathrm{fm}/c$), this quickly quenches due to the momentum anisotropy (within $8\,\mathrm{fm}/c$).

It is possible to make a Fourier expansion of particle distributions as a function of the azimuthal angle ϕ about the beam axis z, with respect to the reaction plane Ψ_{RP} [98]:

$$\frac{\mathrm{d}N}{\mathrm{d}\,(\phi - \Psi_{\mathrm{RP}})} = \frac{N_0}{2\pi}\left(1 + 2v_1\cos\left(\phi - \Psi_{\mathrm{RP}}\right) + 2v_2\cos\left(2(\phi - \Psi_{\mathrm{RP}})\right) + ...\right), \qquad \textbf{(2.18)}$$

where N_0 is a normalisation parameter and the Fourier coefficients v_{n} are the flow
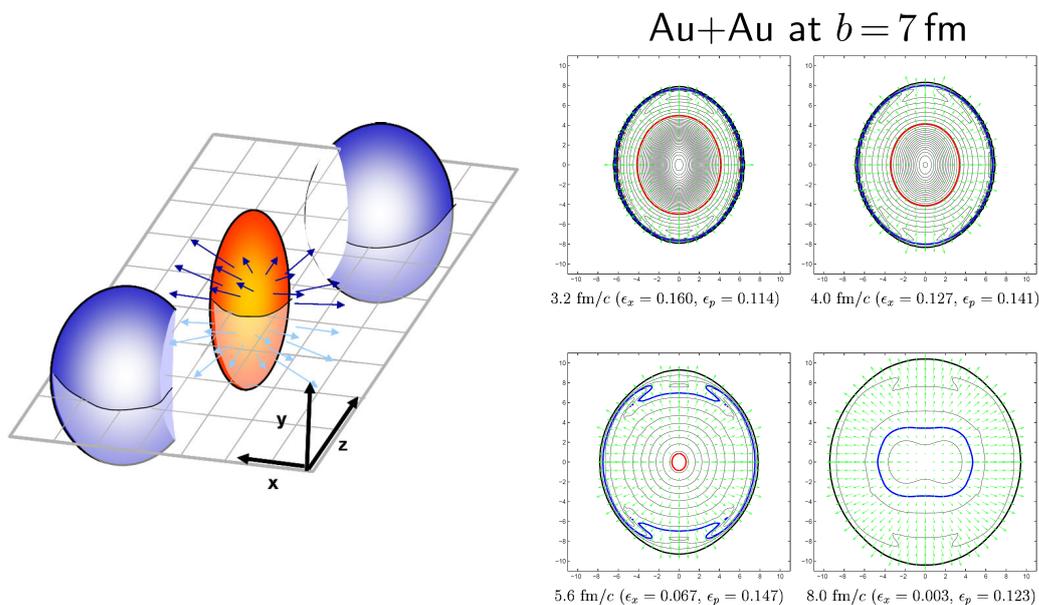
Figure 2.14: Left: schematic diagram of a non-central heavy-ion collision, with two nuclei travelling in the $\pm z$ directions. The blue regions represent the spectator (i.e. non-colliding) nucleons of each nucleus, and the orange spheroid represents the collision region. The grey grid indicates the reaction plane. Right: simulation of elliptic flow in Au–Au collisions with an impact parameter $b = 7\,\text{fm}$ [95]. Figures from [96, 97].

parameters for each harmonic n. v_1 is known as the 'directed flow' parameter, which characterises isotropic flow, and v_2 is referred to as the 'elliptic flow' parameter, which covers the first-order anisotropy of the expansion due to the overall shape of the collision region. Further harmonics can also be studied (v_3, 'triangular flow', etc.); these occur mainly due to fluctuations of nucleons in the initial state.

Each v_n harmonic can be computed for a particle species by fitting the ϕ distribution of the nth harmonic with the following:

$$\frac{\mathrm{d}N}{\mathrm{d}\Delta\phi_n} = \frac{N_0}{2\pi}\left(1 + 2v_n\cos\left(n\Delta\phi_n\right)\right); \quad \Delta\phi_n = \phi - \Psi_n; \qquad \textbf{(2.19)}$$

where Ψ_n is the event plane angle for each harmonic. This is known as the event-plane method. These functions can then be integrated in $\Delta\phi$ for each n in order to determine the particle yields in-plane and out-of-plane, and thus derive a formula relating these quantities to v_n [98]:

$$N_{\text{in-plane}} = \frac{N_0}{2\pi} \cdot 2n \cdot \int\limits_{0}^{\pi/2n} \left(1 + 2v_n \cos\left(n\Delta\phi\right)\right) \mathrm{d}\Delta\phi = \frac{N_0}{2} + \frac{2N_0 v_n}{\pi}$$

$$N_{\text{out-of-plane}} = \frac{N_0}{2\pi} \cdot 2n \cdot \int\limits_{\pi/2n}^{\pi/n} \left(1 + 2v_n \cos\left(n\Delta\phi\right)\right) \mathrm{d}\Delta\phi = \frac{N_0}{2} + \frac{2N_0 v_n}{\pi}$$

$$\rightarrow \frac{N_{\text{in-plane}} - N_{\text{out-of-plane}}}{N_{\text{in-plane}} + N_{\text{out-of-plane}}} = \frac{4v_n}{\pi} \implies v_n = \frac{\pi}{4}\frac{N_{\text{in-plane}} - N_{\text{out-of-plane}}}{N_{\text{in-plane}} - N_{\text{out-of-plane}}}. \qquad \textbf{(2.20)}$$

This method is used in ALICE to measure the elliptic flow v_2 of particles including D mesons [99]. For this measurement, 'in-plane' refers to particles produced within the ranges $-\frac{\pi}{4} < \phi < \frac{\pi}{4}$ and $\frac{3\pi}{4} < \phi < \frac{5\pi}{4}$ with respect to the event plane, and 'out-of-plane' refers to particles within $\frac{\pi}{4} < \phi < \frac{3\pi}{4}$ and $\frac{5\pi}{4} < \phi < \frac{7\pi}{4}$.
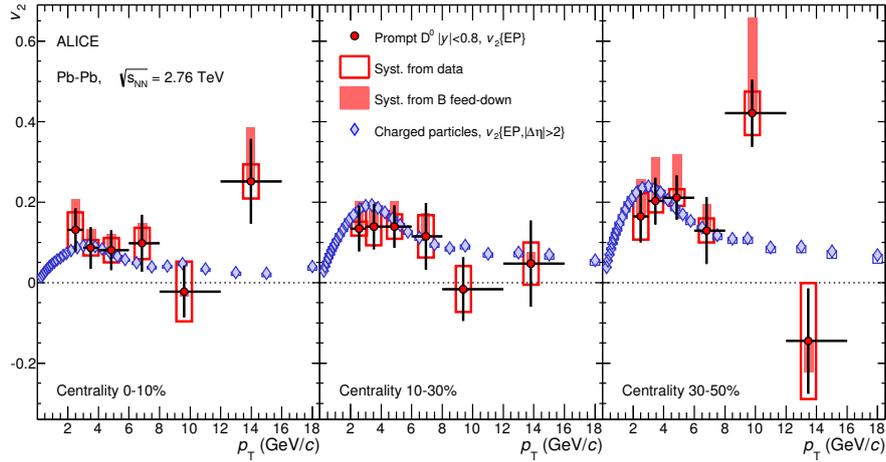
Figure 2.15: Elliptic flow v_2 of D^0 mesons in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76\,\text{TeV}$, measured by the ALICE Collaboration, in three centrality classes as a function of p_T. The results are plotted alongside similar measurements for charged particles. Figure from [99].

Measurements of D^0-meson elliptic flow are shown in Fig. 2.15 as a function of p_T in three centrality classes. The results are compared with similar measurements for light-flavour charged particles in the same centrality classes. The results for the two particle types are comparable within uncertainties, indicating that heavy-flavour and light-flavour hadrons participate in the collective expansion of the medium to a similar degree.

In order to conclusively determine the underlying processes in charm production, it is necessary to compare these results with theoretical models. This is done in Fig. 2.16
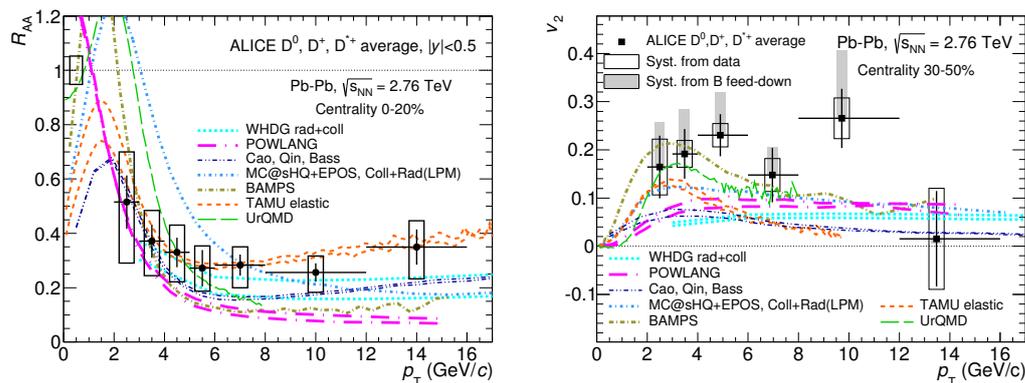
Figure 2.16: ALICE results for D^0-, D^+- and D^{*+}-meson (left) R_{AA} and (right) v_2 as a function of p_T, compared with theoretical calculations [100–111]. Figures from [99].

for the R_{AA} and v_2 of D mesons. Here, a variety of different models that combine both initial- and final-state nuclear effects are shown for each measurement. It can be seen that although many of the models are able to reproduce one of the observables within the experimental and theoretical uncertainties, it remains a challenge for both observables to be predicted simultaneously.

2.5 Open charm production in p–Pb collisions at the LHC

While heavy-flavour measurements in Pb–Pb collisions, and their comparison with pp results, allow some information to be obtained about energy loss in the medium, they alone do not tell the full story. In order to obtain a fuller picture of the processes occurring in the QGP, it is vital to also make measurements in proton–nucleus (p–Pb) collisions.

Due to the smaller system size, the characteristic timescale of the system in p–Pb collisions is much shorter than that in Pb–Pb collisions, making it more likely to dissipate before thermalisation occurs. This means that comparisons between p–Pb and Pb–Pb collisions allow the conditions occurring in the final state due to QGP formation to be disentangled from those happening simply due to the presence of a nucleus in the collision. These effects are commonly termed collectively either as 'nuclear' or 'Cold Nuclear Matter' (CNM) effects, although in reality they are a combination of CNM effects and so-called 'initial-state effects'. Initial-state effects include for example the modification of nuclear parton distribution functions with respect to simple superpositions of proton PDFs, whereas CNM effects are typically those occurring due to the interaction of particles interacting with remnants of the colliding nuclei, such as final-state energy loss.

The nuclear PDFs may be expressed as [112]

$$f_i^A(x, Q^2) = R_i^A(x, Q^2) f_i(x, Q^2), \tag{2.21}$$

where f_i^A is the bound nucleon PDF for a parton species i, f_i is the free nucleon PDF, determined by deep inelastic scattering, R_i^A is the partonic nuclear modification factor, and x and Q^2 respectively refer to the Bjorken-x factor (the momentum fraction carried by the parton) and the four-momentum transfer. Depending on x, the factor R_i^A in heavy nuclei can show an enhancement, due to effects known as antishadowing ($0.05 < x < 0.3$) [113] or Fermi motion ($x \approx 1$) [114], or a suppression due to the EMC effect ($0.3 < x < 0.7$) [115] or nuclear shadowing ($x < 0.05$) [116]. The production of $c\bar{c}$ pairs at LHC energies corresponds to a Bjorken-x of the order of 10^{-4} [117, 118], making nuclear shadowing the dominant process in this regime. This is expected to lead to a decreased interaction cross section for hard scattering processes, and thus heavy-flavour production, making open heavy-flavour production in p–Pb collisions at LHC energies an effective probe of nuclear shadowing.

Initial- and final-state radiation are also of particular interest as nuclear effects. The distribution of partonic transverse momenta, k_T, may be widened due to undergoing soft collisions in the initial stages, or partons may radiate gluons in the initial state and thus reduce the energy of the collision system between partons. This effect is known as k_T broadening. Here, the effect of multiple scatterings between partons is of relevance. The production mechanism of high-k_T partons is determined by the coherence length l_c [119],

$$l_c = \frac{\sqrt{s}}{m_N k_T}, \tag{2.22}$$

where m_N is the mass of the nucleon. For the short coherence length limit, $l_c \ll r_A$ (where r_A is the nuclear radius), the interactions with each nucleon are considered to occur incoherently, leading to a so-called Cronin enhancement [120, 121] due to a large number of initial- and final-state interactions. For the other coherence length limit, $l_c \gg r_A$, the nucleus may be considered to have a single collective parton distribution, with which partons interact coherently, causing gluon radiation processes to be slowed [119]. As heavy-quark production processes are largely associated with gluon radiation, this coherence effect would lead to a suppression of their production in this regime.

A typical observable used for proton–nucleus collisions is the nuclear modification factor R_{pA}:

$$R_{pA} = \frac{1}{A} \cdot \frac{d\sigma^{pA}/dp_T}{d\sigma^{pp}/dp_T}, \tag{2.23}$$

where A is the mass number of the heavy nucleus, and σ^{pA} and σ^{pp} are the production

cross sections in proton–nucleus and proton–proton collisions, respectively. As an observable it is comparable with the Pb–Pb nuclear modification factor R_{AA}; however, while R_{AA} is a per-event quantity, R_{pA} is a comparison of overall cross sections and so is scaled by the number of nucleons in the nucleus rather than the number of binary collisions (see Eq. (2.17)).

The Cronin effect is seen in Fig. 2.17 for calculations of the R_{pAu} factor of pions for p–Au collisions at (left) $\sqrt{s_{NN}} = 200\,\text{GeV}$ and (right) $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ [119].
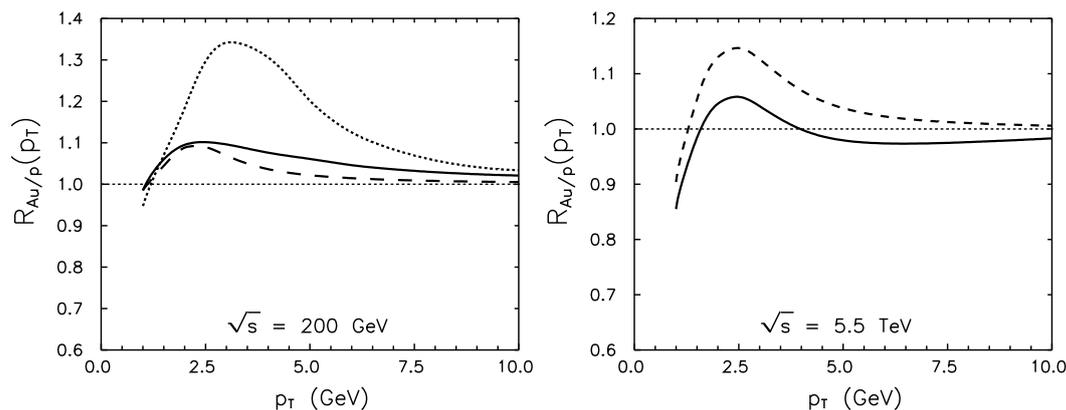
Figure 2.17: Calculations of the pion R_{pAu} at (left) $\sqrt{s_{NN}} = 200\,\text{GeV}$ and (right) $\sqrt{s_{NN}} = 5.02\,\text{TeV}$. The different curve styles correspond to different calculation parameters, discussed in the text. Figures from [119].

In the calculations for RHIC energies ($\sqrt{s_{NN}} = 200\,\text{GeV}$), the dotted line corresponds to the limit of $l_c \ll r_A$, where antishadowing is seen, and the dotted curve corresponds to the limit of $l_c \gg r_A$, where shadowing is expected to dominate. The solid curve is an interpolation between the two limits. In each case, an enhancement of the nuclear modification factor by about 10% is predicted for $1 < p_T < 3\,\text{GeV}/c$, with a broader enhancement seen for shorter coherence lengths. By contrast, the LHC calculations ($\sqrt{s_{NN}} = 2.76\,\text{TeV}$) are shown without (with) the inclusion of gluon shadowing effects as a dashed (solid) curve. If gluon shadowing is neglected, a similarly large enhancement is predicted as for RHIC in the range $1 < p_T < 4\,\text{GeV}/c$; however, when shadowing is included, the enhancement is only of the order of 5% and in a much narrower p_T region. Additionally, a suppression of the nuclear modification factor is seen for lower p_T in both regimes, as well as a small suppression at higher p_T for the case with shadowing. Measurements of the charged-particle R_{pPb} at ALICE have shown behaviour that appears rather similar to these predictions for the central points; however, the 'Cronin-like' enhancement is only barely significant with the current experimental uncertainties [122].

Measurements of D-meson production in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ may be undertaken in order to ascertain the extent of these nuclear effects on the pro-

duction of open heavy-flavour particles. The measurement of the R_{pPb} factor for the average of D^0, D^+ and D^{*+} mesons by ALICE [123] is shown in Fig. 2.18 as a function of p_{T}, and in comparison with theoretical models that include nuclear effects [61, 112, 124–126] (left) and results from Pb–Pb collisions in ALICE, in the 0–10% and 30–50% centrality classes (right) [90].
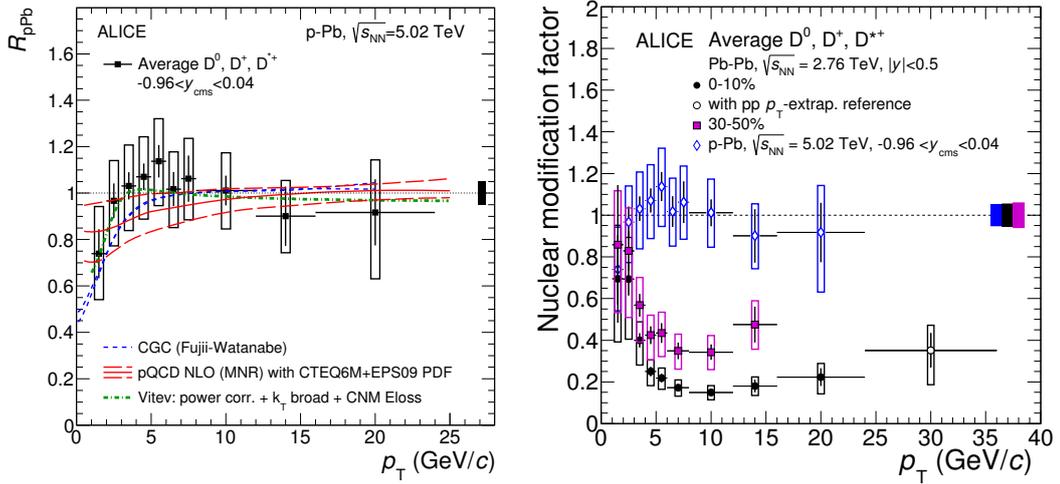
Figure 2.18: R_{pPb} results for mid-rapidity D-meson production measured by the ALICE Collaboration in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$. Left: compared with models including nuclear effects [61, 112, 124–126]; right: compared with R_{AA} results from Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$ in two centrality classes [90]. Figures from [90, 123].

The R_{pPb} of D mesons shows consistency with unity at high p_{T}, and falls off slightly for $1 < p_{\mathrm{T}} < 2\,\mathrm{GeV}/c$. The left-hand panel of Fig. 2.18 shows comparisons between these results and models including CNM and initial-state effects [61, 112, 124–126]; the results were found to be consistent with all of these models, within the uncertainties. A small increase, attributable to the effects of a hydrodynamically expanding medium, was predicted at $p_{\mathrm{T}} \approx 5\,\mathrm{GeV}/c$ in [125] (shown as a green dotted curve). Such an increase would be consistent with a Cronin enhancement of charmed mesons, which would be expected at lower collision energies in this p_{T} region and is often interpreted as a result of multiple scatterings within the nucleus [120, 121]. However, the R_{pPb} results are consistent with unity within the experimental uncertainties, meaning that the measurements are currently not sensitive to any nuclear effects in p–Pb collisions. This was also confirmed by measurements of electrons from semileptonic heavy-flavour hadron decays (i.e. a combination of beauty and charm), which showed consistency with unity within the measurement uncertainties over a wide p_{T} range [127].

From the right-hand panel of Fig. 2.18, it can be seen that D-meson production in both central and semi-central Pb–Pb collisions is also strongly suppressed with respect

to p–Pb collisions, especially at higher p_T. Additionally, the EPS09 PDFs [112] that are used to modify the nuclear PDFs for p–Pb collisions, and successfully reproduce the R_{pPb} result, predict only a small initial-state effect in Pb–Pb collisions [128], further implying that the suppression in Pb–Pb collisions is chiefly due to final-state energy loss in the medium rather than initial-state effects. Future measurements with larger data samples and therefore smaller experimental uncertainties may be able to better determine the behaviour of the spectrum at low p_T and conclusively state the presence or absence of a Cronin effect for charm production in p–Pb collisions.

2.6 High-multiplicity events in small systems

High-multiplicity events in ultrarelativistic collisions have attracted much attention of late, particularly in small collision systems (pp and p–Pb collisions). Although the small system size with respect to Pb–Pb collisions was thought to preclude the formation of a large and long-lived medium, events with high multiplicities have displayed some effects akin to those seen in Pb–Pb collisions.
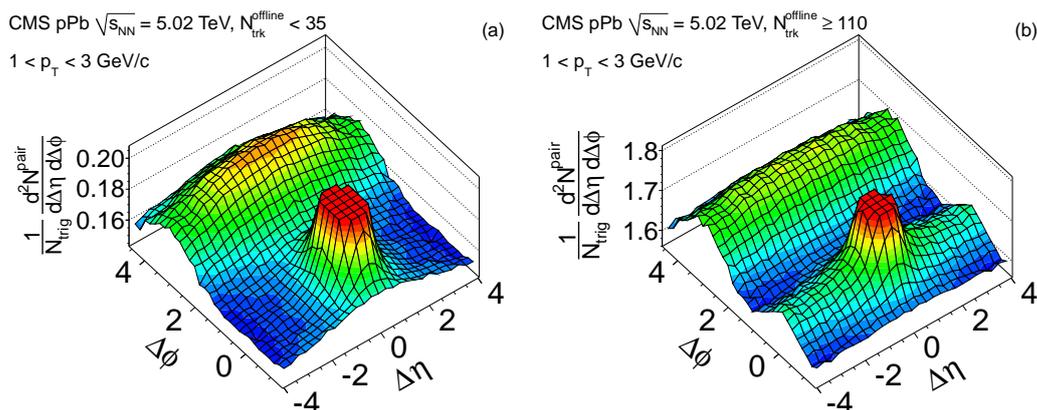
Figure 2.19: Plots of angular two-particle correlations in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ at (left) low and (right) high multiplicity, measured by the CMS Collaboration. Figure from [7].

One of the first examples of this in p–Pb collisions at the LHC was the observation of a 'double ridge' structure in measurements of two-particle correlations in high-multiplicity p–Pb collisions by the CMS Collaboration [7], shown in Fig. 2.19. These correlations plot the differences in azimuthal angle ($\Delta\phi$) and pseudorapidity ($\Delta\eta$, related to the zenith angle θ by $\eta = -\ln(\tan(\theta/2))$, see Section 2.2) between pairs of particles produced by the collision, and serve as an indicator of the extent of collectivity between the produced particles. In nucleus–nucleus collisions, a pair of broad peaks (or 'ridges') in $\Delta\eta$, in opposing regions of $\Delta\phi$, implies a significant

collective effect on particle motion in the system, and is seen as being indicative of hydrodynamic flow in the QGP; its observation in proton–lead collisions has led to the idea that even small systems can show collectivity if the particle density is high enough. This measurement was later confirmed by the ALICE [8] and ATLAS [9] collaborations, which both saw similar double-ridge structures in high-multiplicity p–Pb collisions. In addition, sequential suppression of ψ' mesons with respect to J/ψ has been observed in small-impact-parameter d–Au [129] and p–Pb collisions [130]. These effects have raised the question of whether something resembling a QGP is in fact seen in small systems at high particle multiplicities.

High-multiplicity events are the consequence of multiple effects. In pp collisions, they occur due to a combination of a smaller impact parameter between the two protons (the separation between their centres), and upward fluctuations of the gluon PDFs, both of which lead to a larger number of potential gluon–gluon scatterings in each collision. These increased numbers of scattering events are termed as 'multi-parton interactions', or MPI [10]. MPI contributions have also been observed in the multiple production of c$\bar{\text{c}}$ and b$\bar{\text{b}}$ pairs in proton–antiproton and proton–proton collisions at both Tevatron and the LHC [11–15].

In p–Pb collisions, high-multiplicity events also have a contribution from MPI, but this is accompanied by the larger number of binary nucleon–nucleon collisions that may occur in this system ($\langle N_{\text{coll}} \rangle = 6.9$, determined through the ratio between the proton–nucleon and p–Pb interaction cross sections [131]). Multiplicity-dependent measurements in pp and p–Pb collisions therefore probe the interplay between these two effects.

Heavy-flavour production as a function of multiplicity

Heavy-flavour production rates as a function of charged-particle multiplicity in pp and p–Pb collisions are of particular interest. The generation of the overall charged-particle multiplicity is largely dominated by light particles that are produced in soft production processes, whereas heavy quarks are only produced due to initial hard scatterings. Therefore, comparisons between the yields of light- and heavy-flavour particles probe the correlations between the two production mechanisms.

Studies of open heavy-flavour yields as a function of charged-particle multiplicity were made in pp collisions by the ALICE Collaboration, through measurements of D mesons and non-prompt J/ψ mesons (i.e. those from the decays of B mesons) [16]. In ALICE, the charged-particle multiplicity is estimated in two regions, both at mid-rapidity through the number of track segments counted by the inner tracker of the detector, and at forward and backward rapidity by counting the signal from charged particles in the V0 hodoscopes in these regions (see Section 6.1 for further information about multiplicity determination). The mid-rapidity multiplicity estimator includes

charged particles from the same hard scattering that produced the D meson, as well as the D meson's own decay products, potentially introducing autocorrelations into the measurement. The measurement of multiplicity at large rapidities was therefore also undertaken in order to judge the extent of this bias (if any), as the estimator is decoupled from the rapidity region in which D mesons are studied.

The multiplicity dependence of heavy-flavour production was studied through the use of relative yields, defined as the ratio of the per-event yield in each multiplicity class to the multiplicity-integrated yield per event (see Section 6). This quantity is shown for both particle types for $2 < p_T < 4\,\mathrm{GeV}/c$ as a function of mid-rapidity charged-particle multiplicity in the left-hand panel of Fig. 2.20. Here, the charged-particle multiplicity is given as the pseudorapidity density of charged particles within the measured rapidity region $(\mathrm{d}N_{ch}/\mathrm{d}\eta)$ relative to the overall mean $(\langle \mathrm{d}N_{ch}/\mathrm{d}\eta \rangle)$. Here, $\langle \mathrm{d}N_{ch}/\mathrm{d}\eta \rangle_{|\eta|<1.0} = 6.01 \pm 0.01\,\text{(stat.)}^{+0.20}_{-0.12}\,\text{(syst.)}$ [132]. The right-hand panel shows the relative D^0-meson yield as a function of forward and backward multiplicity, $-3.7 < \eta < -1.7$ and $2.8 < \eta < 5.1$, where the detector signal N_{V0} was simply normalised to the average in the data sample rather than being converted to an N_{ch} value.
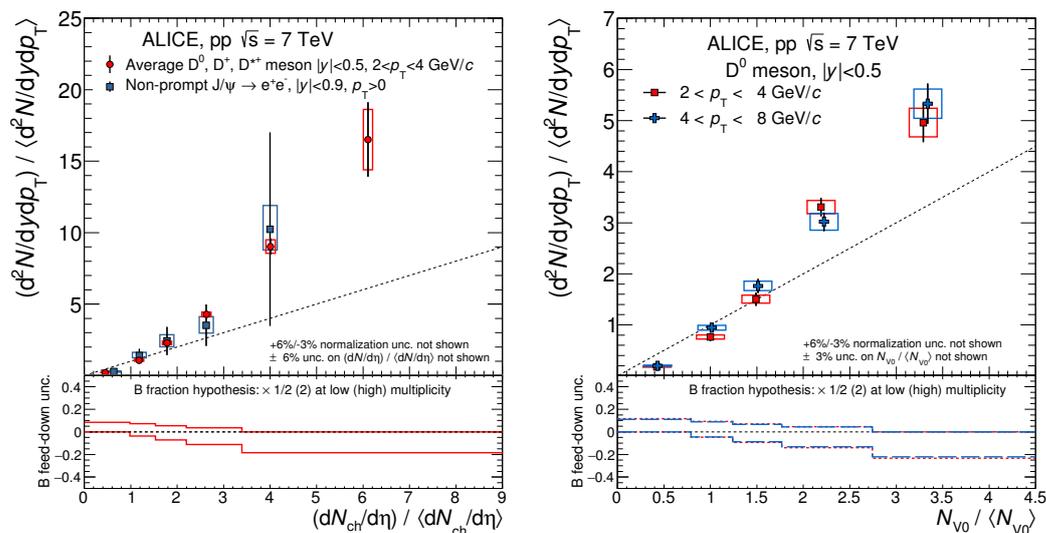
Figure 2.20: D-meson and non-prompt J/ψ production as a function of charged-particle multiplicity in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$. Left: Average of D^0, D^+ and D^{*+} mesons compared with non-prompt J/ψ as a function of multiplicity at mid-rapidity. Right: D^0 mesons at mid-rapidity in two p_T intervals, as a function of multiplicity at forward and backward multiplicity. Figures from [16].

For both multiplicity estimators, an upward curve was found for the relative D-meson yield as a function of multiplicity, implying a faster scaling of hard scatterings than soft processes at high multiplicities. The trend was roughly similar in both

rapidity regions, implying that there is no directional dependence of charged-particle production overall, and that any autocorrelations in the mid-rapidity region are small. In addition, no significant p_T dependence was seen in the results. The relative non-prompt J/ψ yield (i.e. the relative beauty yield) was found to scale similarly to the relative D-meson yield, implying that the scaling of hard processes is independent of the flavour of the produced heavy quarks.

In order to understand the underlying processes, it is important to compare these results with model calculations. One relevant model is the PYTHIA 8 event generator [133]. Hadronic collisions in this model begin with an initial two-to-two hard process, predominantly the scattering of a charm sea quark off another quark (for example $c + u \rightarrow c + u$) or gluon fusion (see Fig. 2.6). In this model the sea-quark scattering is more significant, contributing 9% of the total D-meson production as compared to 2% for gluon fusion [16]. This initial scattering is followed by consecutive hard interactions in the colliding protons (MPI). Gluons may also split into a charm–anticharm pair (as in Fig. 2.6), and the sources of these gluons can either be the preceding hard processes, or initial- and final-state radiation (ISR/FSR). When the model calculations are integrated over all multiplicity, initial- and final-state radiation contributes about 62% of the overall D-meson production cross section, MPI contributes ∼21%, and the initial hard scattering corresponds to ∼11% overall [16].

The PYTHIA calculations of the relative D-meson yield are illustrated in Fig. 2.21. The top-left and top-right panels show the contributions from each process for charm and beauty quarks, respectively. In the top-left plot, it can be seen that MPI and initial/final-state radiation both lead to an increasing gradient in the relative D-meson yield at high multiplicity. By contrast, the contribution from the first hard scattering saturates at relatively low multiplicity. This is explained by the fact that increased amounts of ISR/FSR and MPI are able to contribute both to the production of charm quarks and to the total multiplicity at all multiplicities, while the initial scattering is only significant for charm production at low multiplicities. The steeper onset seen for MPI in beauty production (top-right panel) corresponds to the larger beauty mass, which allows for a higher level of event activity for MPI and ISR/FSR processes. The lower panels show the corresponding predictions for charm quarks, for (left) all contributions in PYTHIA and (right) only the initial hard scattering process. The predictions are given in five p_T intervals. A p_T ordering is seen for the sum of all processes, whereby the relative yield increases more steeply at higher p_T than at lower p_T. When only the initial hard scattering is considered, the D-meson yield peaks at low multiplicity before decreasing at higher multiplicity for low p_T values, and shows a linear increase followed by a saturation for higher p_T ($8 < p_T < 12\,\mathrm{GeV}/c$). The contributions of MPI to the total charged-particle multiplicity were also studied in the framework, finding that at low multiplicities MPI only played a small role, whilst events with multiplicities of five times the average were found to have over sixteen
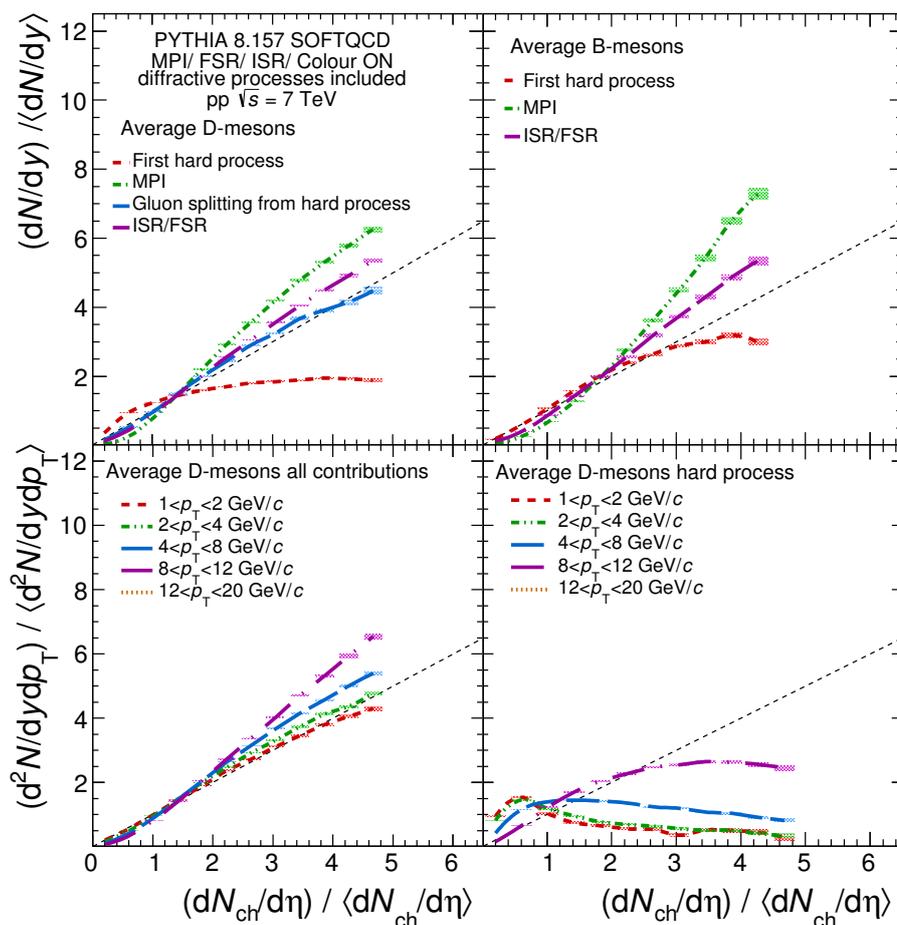
Figure 2.21: Relative D- and B-meson yields as a function of the relative charged-particle multiplicity at mid-rapidity in pp collisions at $\sqrt{s} = 7\,\text{TeV}$, calculated under the PYTHIA 8 framework [133]. The top panels show the contributions from the initial hard process, multi-parton interactions, gluon splitting and initial- and final-state radiation. The lower panels show the relative D-meson yield for five p_T intervals: on the left, for the sum of all processes, on the right, for only the first hard scattering. The dashed grey identity line is shown to guide the eye. Figure from [16].

parton–parton interactions [16].

Figure 2.22 shows comparisons between the D-meson results for the mid-rapidity multiplicity estimator with model calculations. Included among these are the EPOS 3 event generator [18, 19], percolation model calculations [134, 135], and calculations with PYTHIA 8.157 [133], with the full set of effects discussed above. The results are shown in four p_T intervals from 1 to $12\,\text{GeV}/c$.

EPOS 3 employs a 'parton-based Gribov–Regge' model [18], in which individual parton–parton scatterings are represented by parton ladders, or Pomerons. Each ladder is characterised by a hard pQCD process, with the inclusion of initial- and
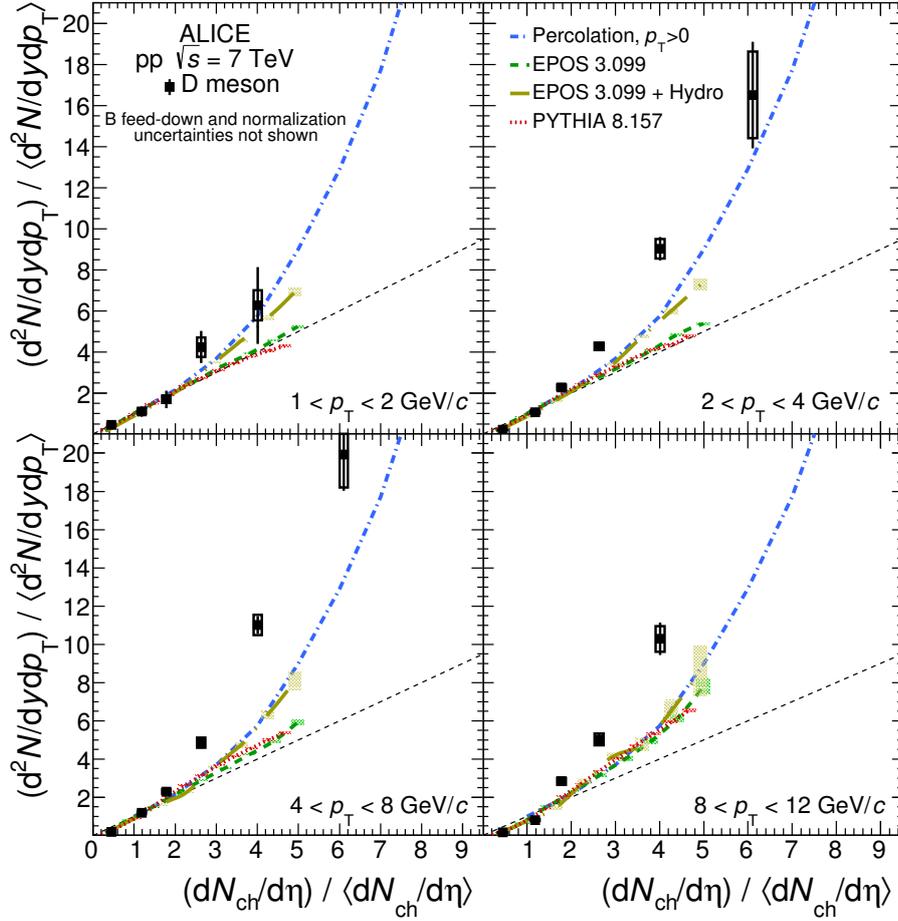
Figure 2.22: D-meson production as a function of charged-particle multiplicity in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ in four p_T intervals. The results for the average of D^0, D^+ and D^{*+} mesons are compared with theoretical calculations [18, 19, 133–135]. Figure from [16].

final-state radiation. String fragmentation processes are used to account for hadronisation, and non-linear effects are considered according to a saturation scale Q_s, which depends on the number of participants and the energy of the Pomeron. Under these considerations, the number of multi-parton interactions is directly proportional to both the number of hard scatterings and the total multiplicity. A consistent approach is applied in this model for pp, p–Pb and Pb–Pb collisions. In addition, viscous hydrodynamics may be applied to the core of the collision to account for the evolution of the system [19].

The percolation model [134, 135] is based on colour sources of finite size being exchanged between the colliding nucleons and interacting with one another. The effective number of these colour sources is reduced by coherence between them. At

high energy densities, the total yield of charged particles from soft processes is expected to be reduced, as the effective size of each source scales with the inverse of its mass, leading to a saturation of lighter particles. Due to their smaller effective size, heavier quarks are less affected by this. The percolation calculations shown here are integrated over all p_T.

The results shown in Fig. 2.22 show large differences between the predictions of the different models. For all of them, the relative D-meson yield at high p_T ($8 < p_T < 12\,\mathrm{GeV}/c$) shows a faster-than-linear increase; for $p_T < 8\,\mathrm{GeV}/c$ the increase is roughly linear for PYTHIA and EPOS without hydrodynamics, which would imply equal scaling for soft and hard production processes. By contrast, the p_T-integrated percolation model calculations are roughly consistent with the EPOS calculations with hydrodynamics, both of which show a greater-than-linear increase of the relative D-meson production rate with respect to multiplicity for all p_T. While all of the models underestimate the relative D-meson yields at high multiplicity for $2 < p_T < 12\,\mathrm{GeV}/c$, the percolation and hydrodynamic models describe the data better than PYTHIA or pure EPOS, potentially indicating that flow effects, and therefore collectivity, are relevant for charged particles at high multiplicities in this system. Similar measurements in p–Pb collisions allow these effects to be quantified in a slightly larger system, as well as judging whether the presence of a nucleus in the initial state has an effect on the possible collectivity, see Section 6.

3 | A Large Ion Collider Experiment at the LHC

Measurements of ultrarelativistic hadronic collisions require sophisticated experimental apparatus, not only to create a high-energy collision in the first place, but also to reconstruct and analyse the particles that are produced. For the work presented in this thesis, this took the form of the ALICE detector system [136], which is one of the major four experiments of the CERN Large Hadron Collider. This section gives an overview of the ALICE detector systems employed for measurements of open heavy-flavour production in this work. Section 3.1 introduces the Large Hadron Collider (LHC). Section 3.2 describes the ALICE detector, and the relevant subdetectors that were used for the measurements presented in this thesis. Sections 3.3 and 3.4 then outline some of the general reconstruction techniques used in ALICE, namely the tracking procedure and the methods used for particle identification (PID).

3.1 The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) [137] is to date the world's largest particle accelerator. It is based at CERN (Conseil Européen pour la Recherche Nucléaire), which straddles the border between France and Switzerland. The accelerator itself occupies the former tunnel of the Large Electron–Positron Collider (LEP), and is a ring of 27 km in circumference, approximately 100 m underground.

The LHC (shown with the rest of the CERN accelerator complex in Fig. 3.1) is host to a variety of particle and nuclear physics experiments, the four largest being ATLAS, CMS, LHCb and ALICE. ATLAS (A Toroidal LHC Apparatus) [138] and CMS (Compact Muon Solenoid) [139] are two general-purpose detectors, with one of their central aims being the search for and characterisation of the Higgs boson, discovered in 2012 [140,141], as well as signs of new physics. LHCb (LHC Beauty) [142] is a series of forward detectors used for the study of b-physics. ALICE (A Large Ion Collider Experiment) [136] is the dedicated heavy-ion experiment of the LHC, whose aim is to study the nature of the Quark–Gluon Plasma (QGP) formed in ultrarelativistic nucleus–nucleus collisions.

Figure 3.1: Schematic of the CERN accelerator complex. The locations of the four main LHC experiments (ALICE, ATLAS, CMS and LHCb) are shown around the LHC ring. Image (edited) from [143].

Protons are accelerated by a linear accelerator (LINAC 2) before being passed into the Proton Synchrotron Booster, and sequentially accelerated via the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) before being injected into the LHC at energies of 450 GeV. Lead nuclei follow a similar series, but are accelerated via the LINAC 3 and Low-Energy Ion Ring (LEIR) before being passed into the PS. Their injection energy at the LHC is 177 GeV per nucleon. Once injected, the beams are further accelerated around the LHC ring before being brought to collision in the centre of each detector. During RUN 1 of the LHC (2009–2013) the maximum possible collision energy was $\sqrt{s} = 8$ TeV for pp collisions, $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for p–Pb collisions and $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV for Pb–Pb collisions; during the first Long Shutdown (2013–2015) improvements were made to the accelerator system to allow energies of $\sqrt{s} = 13$ TeV to be reached for pp collisions, $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for Pb–Pb collisions, and $\sqrt{s_{\mathrm{NN}}} = 8$ TeV for p–Pb collisions. The LHC machine is described in further detail in [137].

3.2 A Large Ion Collider Experiment (ALICE)

The main function of ALICE [136, 144] is to study the Quark–Gluon Plasma (QGP) that is expected to be formed in the high temperatures and energy densities present in central Pb–Pb collisions at LHC energies (see Section 2.1). Its main strengths lie in the tracking and identification of particles at mid-rapidity in high-multiplicity nuclear

collisions.

A schematic of the ALICE detector system is shown in Fig. 3.2. ALICE consists of the Central Barrel at mid-rapidity, and the Muon Arm at forward rapidity. The central-barrel detectors sit inside the L3 magnet, which is a large solenoid providing a magnetic field of $0.5\,\mathrm{T}$ along the beam axis.
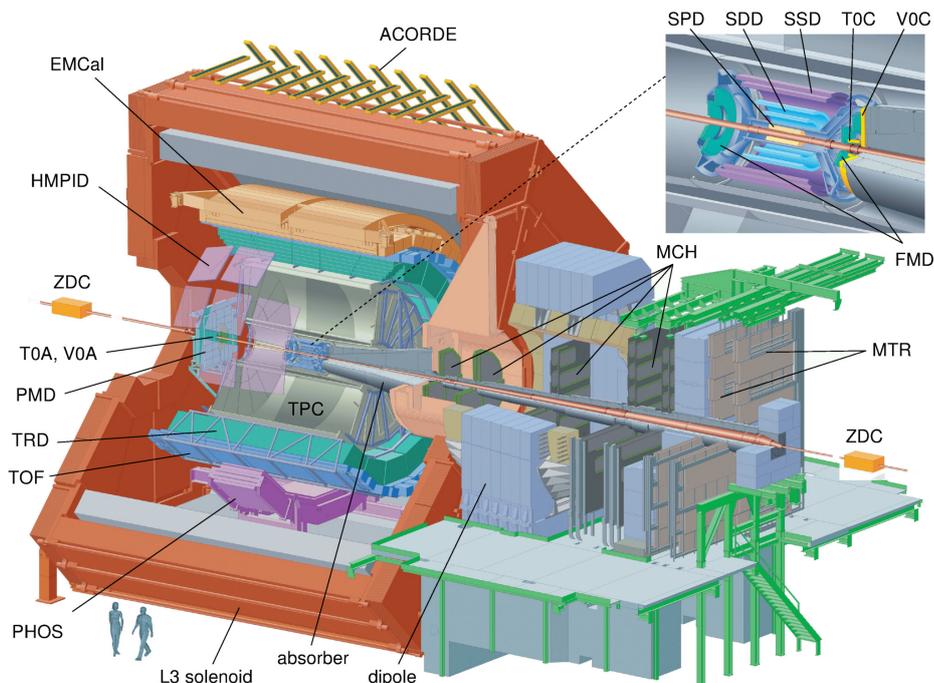
Figure 3.2: The ALICE Central Barrel and Forward Muon Arm, with detector subsystems labelled. The inlay at the top-right shows a zoomed-in view of the beam line (orange) and the Inner Tracking System. Figure from [144].

The following subsections present the detectors in the ALICE Central Barrel that are relevant for the analyses in this work. A list of the detectors and their acceptances is given in Tab. 3.1.

The Inner Tracking System (ITS)

The Inner Tracking System, or ITS [145], is the closest detector to the interaction point. It is formed of six layers of silicon detectors, two each of Silicon Pixel Detectors (SPD), Silicon Drift Detectors (SDD) and Silicon Strip Detectors (SSD), at radii between 3.9 and $43\,\mathrm{cm}$ from the beam line. Each layer has full azimuthal coverage, and pseudorapidity coverage ranging between $|\eta| < 0.9$ for the SDD layers and $|\eta| < 2.0$ for the innermost SPD layer.

The SPD accounts for the innermost two layers of the ITS, with a spatial resolution of $12\,\mu\mathrm{m}$ in z and $100\mu\mathrm{m}$ in $r\phi$. The third and fourth layers of the ITS comprise the

Detector	Acceptance	Position (cm)	Technology	Main purpose
SPD	$\lvert\eta\rvert < 2.0$	$r = 3.9$	Si pixel	tracking, vertex
	$\lvert\eta\rvert < 1.4$	$r = 7.6$	Si pixel	tracking, vertex
SDD	$\lvert\eta\rvert < 0.9$	$r = 15.0$	Si drift	tracking, PID
	$\lvert\eta\rvert < 0.9$	$r = 23.9$	Si drift	tracking, PID
SSD	$\lvert\eta\rvert < 1.0$	$r = 38$	Si strip	tracking, PID
	$\lvert\eta\rvert < 1.0$	$r = 43$	Si strip	tracking, PID
TPC	$\lvert\eta\rvert < 0.9$	$85 < r < 247$	Ne/Ar drift+MWPC	tracking, PID
TOF	$\lvert\eta\rvert < 0.9$	$370 < r < 399$	MRPC	PID
V0	$2.8 < \eta < 5.1$	$z = 329$	scintillator	charged particles
	$-3.7 < \eta < -1.7$	$z = -88$	scintillator	charged particles
T0	$4.6 < \eta < 4.9$	$z = 370$	quartz	time, vertex
	$-3.3 < \eta < -3.0$	$z = -70$	quartz	time, vertex

Table 3.1: List of detectors used in the analyses presented in this work, and their respective radial positions and polar acceptances in pseudorapidity. All of the listed detectors have full azimuthal coverage. Values from [144].

SDD, which has a resolution of $25\,\mu\mathrm{m} \times 35\,\mu\mathrm{m}$ in $(z \times r\phi)$. The final two layers form the SSD, whose spatial resolution is $(830 \times 20)\mu\mathrm{m}^2$ $(z \times r\phi)$.

The primary function of the ITS is to precisely determine the location of the interaction point (primary vertex) and the point at which particles decay (secondary vertex), in order that the decay topology in each event can be better estimated. The primary vertex is defined as the point in space with the greatest level of convergence of 'tracklets' reconstructed by the SPD. Tracklets in this context are defined as the lines joining pairs of hits in the two layers of the detector. These can also be used to estimate the multiplicity of an event, under the assumption that the total number of primary charged particles (N_{ch}) is proportional to the number of tracklets counted by the detector for a collision. Typically this is performed at mid-rapidity, $\lvert\eta\rvert < 1.0$, to remove possible edge effects on the estimation. The multiplicity determination and correction for multiplicity-dependent analyses is discussed in further detail in Section 6.1.

A visual representation of track reconstruction in the ITS is shown in Fig. 3.3, which shows a cross section of the event display for the first pp collision recorded by ALICE at $\sqrt{s} = 900\,\mathrm{GeV}$ in 2009 [146]. The hits in each layer of the ITS are shown as dots, with red lines drawn to represent the reconstructed tracks. The tracking algorithm that is used to determine tracks in the detector is further discussed in Section 3.3.

As well as contributing clusters to the charged-particle tracking algorithm, the SDD and SSD provide analogue outputs that allow them to give PID information via specific energy loss $(\mathrm{d}E/\mathrm{d}x)$ measurements in the silicon of the detectors. In order
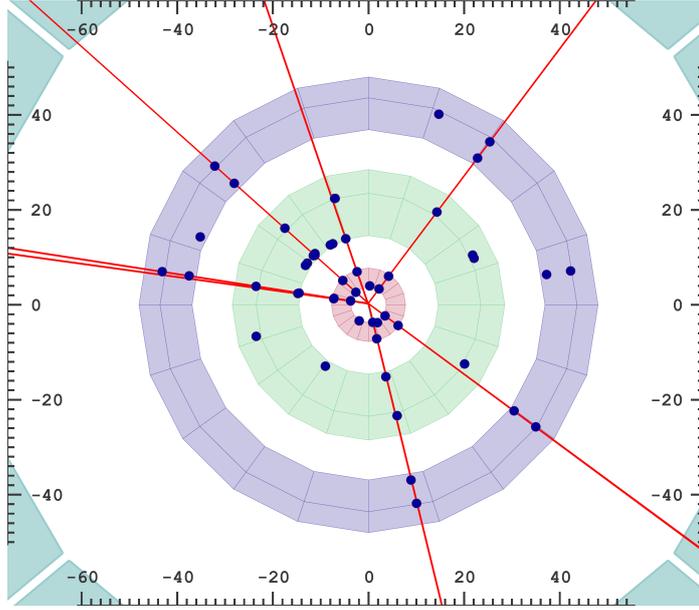
Figure 3.3: Cross-sectional view of the ITS in the event display for the first proton–proton collision reconstructed by ALICE at $\sqrt{s} = 900\,\text{GeV}$ in 2009. The dark blue dots indicate hits in the layers of the ITS, and the red lines indicate reconstructed tracks. The SPD, SDD and SSD are shown shaded in pink, green and blue, respectively. Image from [147], also published in [146].

for PID measurements to be made, the detector response must be parameterised according to each detector species. The method used to parameterise the energy loss distribution was first proposed by the ALEPH Collaboration [148], using a modified form of the Bethe–Bloch formula [149, 150],

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

where β and γ are the relativistic velocity and Lorentz factor, and P_i are fit parameters [144].

The energy loss of charged particles in material is a statistical process, and so the energy loss signal in a detector is described by a distribution about the Bethe–Bloch expectation. This may be described by a Landau distribution [151], which is characterised by a sharp onset in the probability density for energy deposits lower than the mean, and a shallower decline for energy deposits above the mean (the 'Landau tail'). In order to reduce the effects of this upper tail and more simply approximate this distribution with a Gaussian function, a truncated mean is typically taken of the signals, whereby larger energy deposits out of a selection of hits for a given particle are discarded. In the ITS, the outer four detector layers are considered, with a mean taken between the lowest two dE/dx signals if four hits are measured, or a weighted

average of the lowest two if three are measured (with weight 1 for the lowest, and 0.5 for the second-lowest) [144]. Taken in aggregate over the entire data sample, this leads to an approximately Gaussian response about the expectation for each species. The PID performance of the ITS is shown in Fig. 3.4, which plots the measured dE/dx as a function of momentum. A separation of the measured signals can be seen between kaons and pions up to $\sim 1\,\mathrm{GeV}/c$, and between protons and pions up to $\sim 2\,\mathrm{GeV}/c$.

ALI-PUB-72353
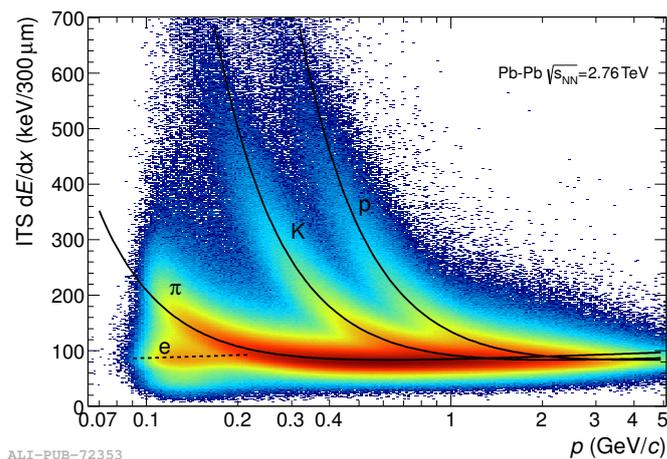
Figure 3.4: Distribution of energy loss in the ITS as a function of the particle momentum for electrons, pions, kaons and protons. The momentum and dE/dx were both measured using only the ITS. Figure from [144].

The Time Projection Chamber (TPC)

Surrounding the ITS in the central barrel is a large gas Time Projection Chamber (TPC) [152]. The TPC serves as the main tracker in the central barrel of ALICE. It covers the full azimuth at mid-rapidity ($|\eta| < 0.9$), and lies at a radial distance between 85–247 cm from the beam axis, with a length of 5 m along the beam direction. The ALICE TPC is the largest detector of its kind to date, with a total drift volume of $88\,\mathrm{m}^3$. It was filled with a Ne–CO_2-based gas mixture in RUN 1, with a ratio of roughly 90% Ne to 10% CO_2.

The TPC employs a cylindrical field cage that is used to maintain a constant electrostatic field within the detector. The field is aligned to the direction of the beam axis. Atoms of the noble gas in the detector are ionised by interactions with the passing charged particles, and the electrons that are freed then drift to the end plates of the detector due to the applied electric field. The end plates contain a series of readout pads consisting of Multi-Wire Proportional Chambers (MWPCs), which collect and amplify (with a gain factor of 7,000–8,000 [152]) the electrons released from ionised gas atoms in the detector.
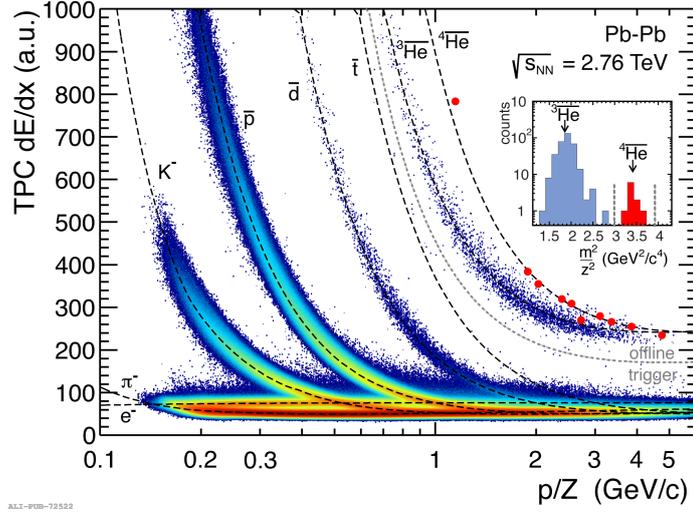
Figure 3.5: TPC dE/dx signal as a function of rigidity for electrons, pions, kaons, antiprotons and light antinuclei in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$. The black curves represent the Bethe-Bloch parameterisation for each particle species. The in-laid panel shows the mass hypothesis distributions measured by TOF for light antin-uclei (^{3}He and ^{4}He). Figure from [144].

In addition to its tracking capabilities, the TPC provides PID information in the form of measurements of momentum, charge, and specific energy loss (dE/dx) of particles passing through the gas. As with the ITS, a truncated mean is used for the dE/dx measurement in the TPC dE/dx in order to reduce the Landau tail of the energy loss distribution [151]. The truncated mean in the TPC discards the 5% smallest and 25% largest signals for a given track, leaving a near-Gaussian response with a resolution of ∼5.2% in pp collisions or ∼6.5% in the 0–5% most-central Pb–Pb collisions [144]. The energy loss distribution in the TPC can be seen in Fig. 3.5 for electrons, pions, kaons, protons and light nuclei in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76\,\text{TeV}$. The detector allows charged-particle ratios to be measured reliably up to high momenta, with a good level of separation in the relativistic rise region up to transverse momenta of $p_T = 20\,\text{GeV}/c$ when measured on a statistical basis [144].

The Time of Flight Detector (TOF)

The Time of Flight detector (TOF) [153] is essential for hadron identification at inter-mediate momenta, offering clear separation between pions and kaons with momenta up to $2.5\,\text{GeV}/c$, and pions and protons up to $4\,\text{GeV}/c$. It covers the full azimuth, with a pseudorapidity coverage $|\eta| < 0.9$, at a radial distance of 370–$399\,\text{cm}$ from the beam line. It is formed of an array of Multi-gap Resistive Plate Chambers (MRPCs). As a charged particle passes through these chambers, it induces an avalanche of electrons

between the layers of the detector, the timing of which is read out by the detector with high precision.

TOF is complemented at forward and backward rapidity by the T0 detector [154], a pair of arrays of Cherenkov counters, which is used for determining the start-time of a collision. The T0C and T0A detectors respectively sit at $-3.28 < \eta < -2.97$ and $4.61 < \eta < 4.92$, on either side of the IP, and each contains 12 counters consisting of a quartz radiator and a photomultiplier tube. The difference between the reconstructed times of the hits in the TOF and T0 detectors serves as a measurement of the flight times of charged particles between the collision point and TOF.

Once the intrinsic detector resolution, contributions from calibration and electronics, start-time uncertainty, and the tracking and momentum resolution have been taken into account, the overall TOF time resolution for charged particles is $\sim 80\,\mathrm{ps}$ [144].

The relationship between the particle momentum and its velocity allows an estimation of the mass to be made. The velocity β is inversely proportional to t_{TOF}, and is related to the rest mass m_0 by

$$m_0 = \frac{p}{\gamma \beta c},$$
(3.2)

where p is the particle momentum, γ is the Lorentz factor, and c is the speed of light.
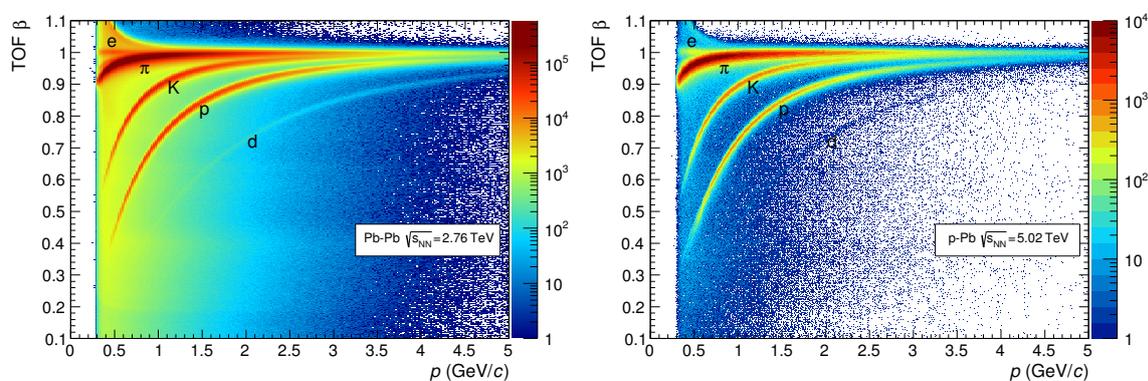
Figure 3.6: Overview of β measurements for particles reaching the TOF detector, with expectations drawn for electrons, pions, kaons, protons and deuterons. Left: Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$. Right: p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$. Figures from [144].

The performance of the TOF detector is shown in Fig. 3.6 for Pb–Pb and p–Pb collisions. In each panel, the measured β of each particle is plotted against its momentum p (as measured in the TPC), and compared with parameterisations of the expected signals for various particle species. A direct comparison of the two collision systems for particles with $0.95 < p < 1.05\,\mathrm{GeV}/c$ is also given in Fig. 3.7. The signals

for Pb–Pb collisions sit on top of a high level of background, which is strongly reduced in p–Pb collisions. This implies that the background occurs due to the mismatching of tracks between the TPC and TOF in high-multiplicity events, rather than the resolution of the detector itself. The fraction of mismatched events between the TPC and TOF increases with the occupancy of the TOF, with the most central Pb–Pb collisions ($\sim 10^4$ TOF hits) giving a pad occupancy of 6.7% and a mismatch fraction of 6.5% [144]. This mismatch fraction is accounted for in the tuning of Monte Carlo simulations to the detector response when efficiency corrections are made, see Section 5.3.
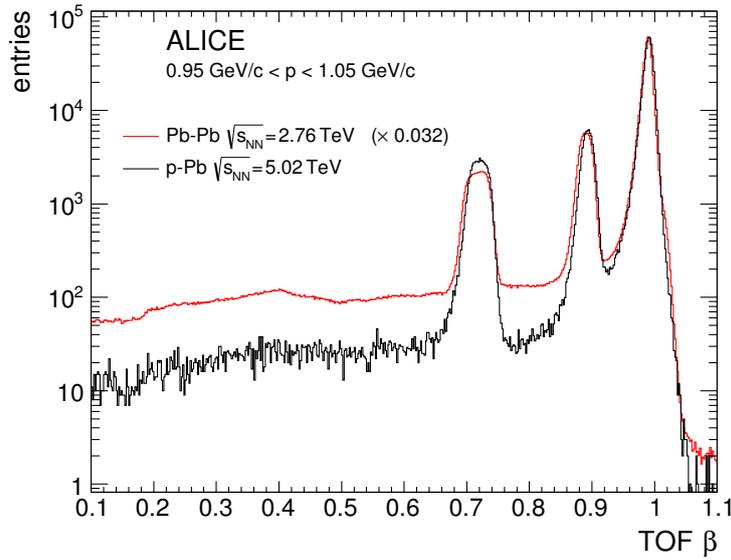
Figure 3.7: Comparison of β values measured by TOF for minimum-bias p–Pb and Pb–Pb collisions, for $0.95 < p < 1.05\,\mathrm{GeV}/c$. The Pb–Pb curve is normalised to the peak at $\beta = 0.99$ (the expectation for pions in this momentum interval) for p–Pb collisions. A larger background (due to mismatches from high-multiplicity events) is seen for Pb–Pb collisions than for p–Pb collisions. Figure from [144].

The V0 Detector

The V0 detector consists of the V0A and V0C hodoscopes on either side of the interaction point. In the standard ALICE coordinate system, V0A covers $2.8 < \eta < 5.1$, 329 cm away from the IP, and V0C covers $-3.1 < \eta < -1.7$, 88 cm from the IP and in front of the hadronic absorber of the muon arm. In p–Pb collisions, the signs of the quoted pseudorapidity intervals are swapped when the Pb nucleus is travelling towards V0A. Each of the two detectors consists of a circular array of 32 scintillators. A common minimum-bias trigger requirement in ALICE is that coincident hits are recorded in both of these detectors.

The amplitude measured by V0 corresponds to the charged-particle multiplicity produced in a collision, the determination and correction of which is outlined in Section 6.1. In heavy-ion collisions this estimate may also be used as a basis for centrality determination, as illustrated in Fig. 3.8 for the amplitude measured by the V0A counter in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$ [131]. In this case, the multiplicity in the Pb-going direction is measured. The centrality is based on percentiles in this distribution, where the highest multiplicities are assigned as the most central events (0%), and the lowest multiplicities as the most peripheral (100%). These percentiles are marked in the figure by vertical lines. The distribution is also overlaid with a fit under the NBD-Glauber approximation, which is based on a Glauber Monte Carlo calculation for the number of participant nucleons [155–157] combined with a negative binomial distribution (NBD) of the particle multiplicity in individual nucleon–nucleon collisions. The NBD approximation is justified by its successful description of the charged-particle multiplicity in minimum-bias pp and p$\bar{\text{p}}$ collisions in broad regions of rapidity [146, 158]. Similar methodology is also used in ALICE for the estimation of the collision centrality in Pb–Pb collisions [159].
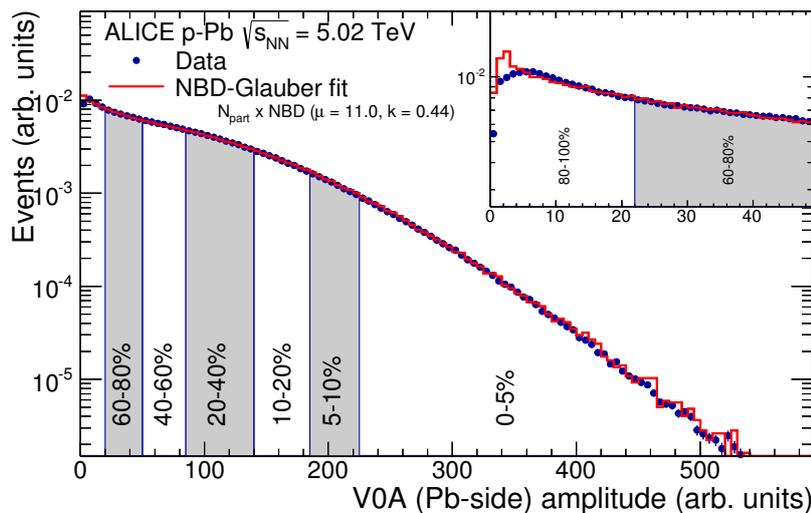
Figure 3.8: Distribution of the sum of amplitudes in the V0A hodoscopes in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$. The boundaries between centrality classes are indicated by vertical lines. An NBD-Glauber fit [146, 155–158] to the distribution is shown as a red line. The inlay shows a zoomed-in view of the 0–50 amplitude region, corresponding to the 40% most peripheral events. Figure from [131].

3.3 Charged-particle tracking in ALICE

The tracking algorithm for the ALICE central barrel [160, 161] relies heavily on the TPC and ITS described in the previous sections. First, the preliminary location of

the interaction vertex must be determined. This measurement uses the SPD (the innermost two layers of the ITS), and defines the primary vertex as the point of convergence of the largest number of SPD tracklets (pairs of hits spanning the two SPD layers, which are extrapolated to the centre of the detector). This assumes that the majority of tracklets stem from primary particles. Tracks that do not converge on this point are then sequentially re-analysed in order to find secondary vertices, as discussed later in this section.

The actual track reconstruction follows a three-stage process (inward–outward–inward). The TPC provides the starting point for the algorithm. Each particle produces hits at up to 159 space points in the detector (one for each pad-row). These are reconstructed from two-dimensional clusters according to pad-row–time planes. Tracks are then found and fitted to these space points using the Kalman filter algorithm, which allows track recognition and reconstruction to be performed simultaneously. The track is seeded using two separate constraints, either with the interaction vertex and two of the TPC clusters at large radius, or with a triplet of clusters excluding the vertex. The seeds are extrapolated towards the inner radius of the TPC in steps, whereby the closest cluster to the track is chosen at each step, according to a cut on the proximity. Tracks are rejected if they have fewer than 20 total clusters, or if over half of the expected clusters for a track are not present. These tracks are then used as the seeds for tracking in the ITS, which continues the cluster finding process until the primary vertex is reached.

The second iteration follows the same process outwards from the primary vertex through the ITS and TPC, again using the Kalman filter, after which it is extended to the TOF and Transition Radiation Detector (TRD, $r = 290\,\mathrm{cm}$), for which it is assigned a flight time and an attempted match to a TRD tracklet, respectively. Clusters are assigned for both detectors. While these outer detectors are not used for the determination of a particle's kinematics as of RUN 1 [144], they provide information for particle identification purposes and improve the momentum resolution at high p_{T} due to offering a longer lever arm for tracking.

Finally, the third iteration proceeds inwards from the outer edge of the TPC. Here, the energy loss signals in the TPC and outer layers of the ITS are used to re-fit the track parameters, namely its position, direction and curvature. The algorithm makes cuts on the distance of closest approach between the tracks and the primary vertex, in order to remove contributions from secondary tracks (those originating from interactions with detector material and particle decays). After this step, the fully reconstructed tracks are extrapolated to the nominal beam line in order to make a second, more precise determination of the primary vertex. The vertex resolution achieved by this procedure is dependent on the track multiplicity, with resolutions of $97\,\mu\mathrm{m}$ achieved for low-multiplicity events ($\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta = 5$) and $32\,\mu\mathrm{m}$ for higher-multiplicity events ($\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta = 25$) [144]. The momentum resolution for tracks

reconstructed by this algorithm is shown in Fig. 3.9, where it is represented by the inverse-p_T resolution [144]:

$$\frac{\sigma_{p_T}}{p_T} = p_T \cdot \sigma_{1/p_T}.$$ (3.3)

This quantity is shown in Fig. 3.9 as a function of $1/p_T$. This is given for TPC standalone tracks and tracks from combined TPC–ITS tracking, both with and without the vertex constraint applied at the start of the seeding. It can be seen that the vertex constraint leads to a greatly improved p_T resolution for TPC-standalone tracking, while the TPC–ITS tracking algorithm performs just as well without the constraint as with it.
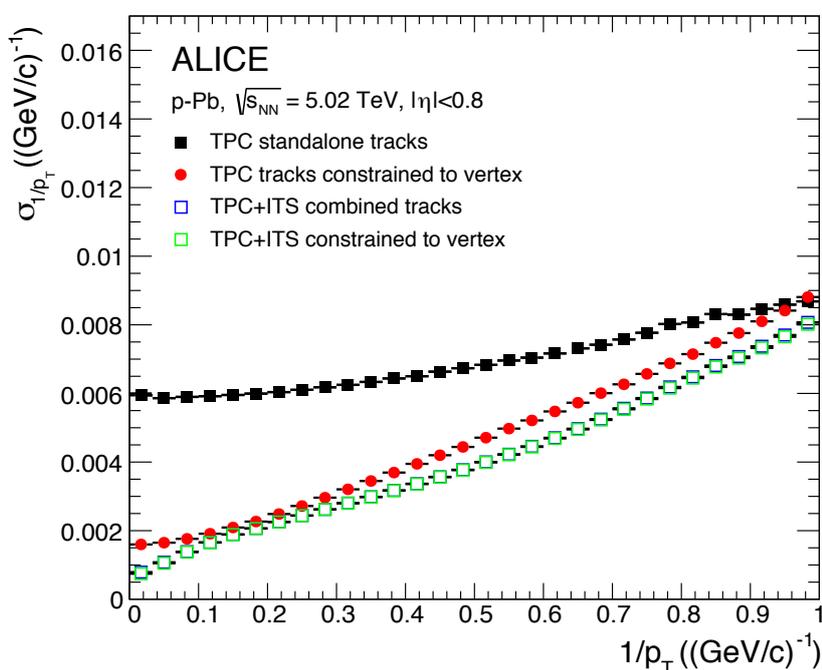
Figure 3.9: Inverse p_T resolution (see Eq. (3.3)) for primary track reconstruction in ALICE. Figure from [144].

Particles with short lifetimes with respect to the scale of the detector, such as D mesons (see Section 4.2), cannot be directly reconstructed by this procedure as they decay long before they would reach any active detector material. These must therefore be reconstructed from the secondary vertices at which they decay. A schematic of two such decays, $K_S^0 \rightarrow \pi^- \pi^+$ and $\Xi^- \rightarrow \Lambda^0 \pi^- \rightarrow p \pi^- \pi^-$, is shown in Fig. 3.10. Here, the two primary particles decay at some point between the first and second layers of the ITS, but do not induce a hit in the SPD as they are neutral. However, their charged decay daughters may be reconstructed in the outer layers as follows. Tracks reconstructed from the Kalman fits described above that lie outside a maximal requirement of distance of closest approach from the primary vertex when extrapolated
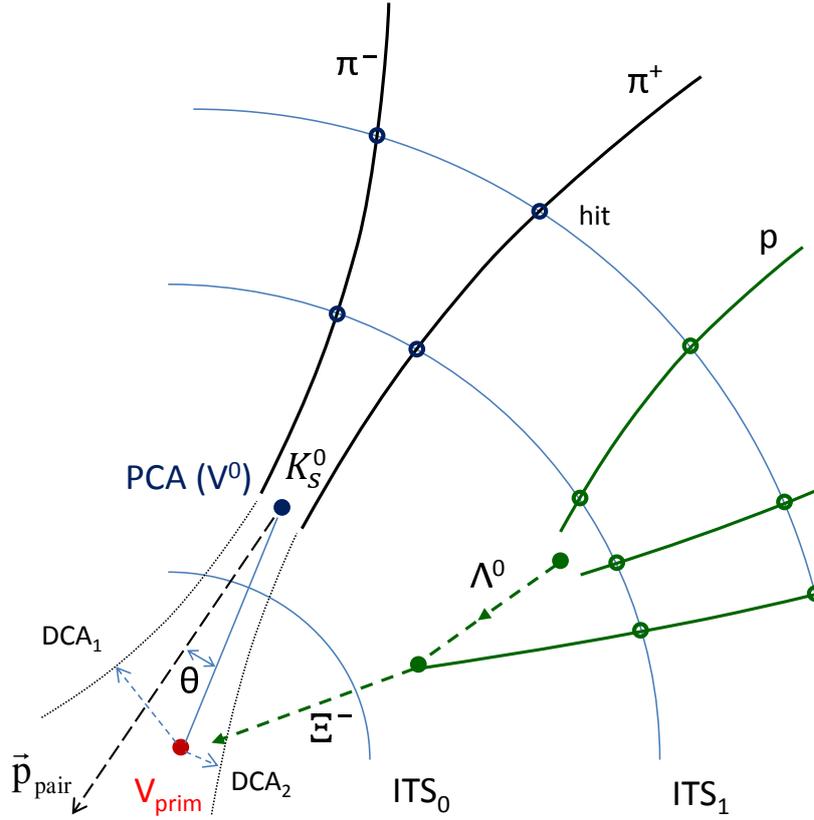
Figure 3.10: Schematic of secondary track reconstruction for K_S^0 and Ξ^- decays. Reconstructed tracks, extrapolated to candidate secondary vertices, are shown as solid lines, and extrapolations towards the primary vertex as dashed lines. The light blue quarter-circles represent layers of the ITS (not to scale), and the red point is the reconstructed primary vertex. Figure from [144].

to the beam line (1 mm for Pb–Pb collisions, or 0.5 mm for pp collisions [144]) are matched in pairs of opposite charges, known as neutral vertex candidates (or V^0 candidates). The point of closest approach (PCA) between the tracks is computed for each such pair, and is subjected to a series of quality cuts. The PCA must lie closer to the primary vertex than the innermost detector hit of either track, the DCA between the two tracks must be smaller than 1.5 cm, and $\cos\theta$ (where θ is the angle between the total momentum vector of the two tracks and the flight line between the candidate and the primary vertex) must be greater than 0.9. Ξ^- decays, as shown in the schematic, are reconstructed by pairing V^0 candidates with reconstructed invariant masses close to that of Λ with further secondary tracks, with additional constraints on the Ξ^- decay length and the PCA between the Λ and the secondary decay track. Further selection cuts may be made on the decay topologies of other particles in the offline data analysis in order to reconstruct them. This is of particular importance for short-lived particles such as D mesons, whose reconstruction procedure is covered

in Section 4.2.

3.4   PID in ALICE

ALICE uses a wide variety of detectors and techniques to identify particles in the Central Barrel, as outlined in Section 3.2. The primary technique used is the specific energy deposit, dE/dx, of charged particles as they pass through active detector material. In the SDD and SSD, this is the silicon of the detector; in the TPC and TRD, the detector gas. In each case, the measured dE/dx can be compared with the momentum, and used to derive the expected particle species based on the Bethe-Bloch formula [162]:

$$-\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{m_e c^2} \cdot \frac{nz^2}{\beta^2} \cdot \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \cdot \left[\ln\left(\frac{2m_e c^2 \beta^2}{I \cdot (1-\beta^2)} \right) - \beta^2 \right], \tag{3.4}$$

where e is the elementary charge, m_e is the electron rest mass, n is the number density of electrons in the target material, z is the charge of the particle, and I is the average ionisation potential of the target atom. An example of the PID signal for the TPC is shown in Fig. 3.5, where a clear separation can be seen between pions and kaons for $p < 0.5\,\mathrm{GeV}/c$, and between protons and pions for $p < 1\,\mathrm{GeV}/c$.

Another important estimate of the particle species can be obtained based on comparing the momentum of a charged particle with its $\beta\gamma$, computed according to its flight time from the interaction point to the TOF detector. The flight time t_{TOF} is defined for each charged-particle track as the difference between the collision time measured by T0 and the reconstructed time of the TOF hit. From this, the mass of the particle can be determined according to Eq. (3.2). A plot illustrating the separation power of the TOF detector is shown in Fig. 3.11. Here, the measured signals for particles with $1.5 < p_T < 1.6\,\mathrm{GeV}/c$ produced in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$ are shown alongside model fits for pions, kaons and protons. The t_{TOF} measurements are shown as the difference from the expectation for kaons in this momentum range, divided by the expected resolution for kaons.

The overall separation power of the TPC, TOF, High-Momentum PID detector (HMPID) and ITS for pions, kaons and protons is shown in Fig. 3.12 as a function of p_T. The TPC is the most effective detector for distinguishing between pions and kaons at low p_T ($< 1\,\mathrm{GeV}/c$) and at higher p_T ($> 3\mathrm{GeV}/c$), with the TOF detector performing the best at intermediate p_T ($0.5 < p_T < 3\,\mathrm{GeV}/c$). For kaons and protons, dE/dx measurements again prove to be the most effective for $p_T < 1.5\,\mathrm{GeV}/c$, with TOF providing excellent separation for $0.5 < p_T < 5\,\mathrm{GeV}/c$. In each case, certain p_T ranges are best served by a single detector, whereas others may benefit from the PID information from a variety of different methods.
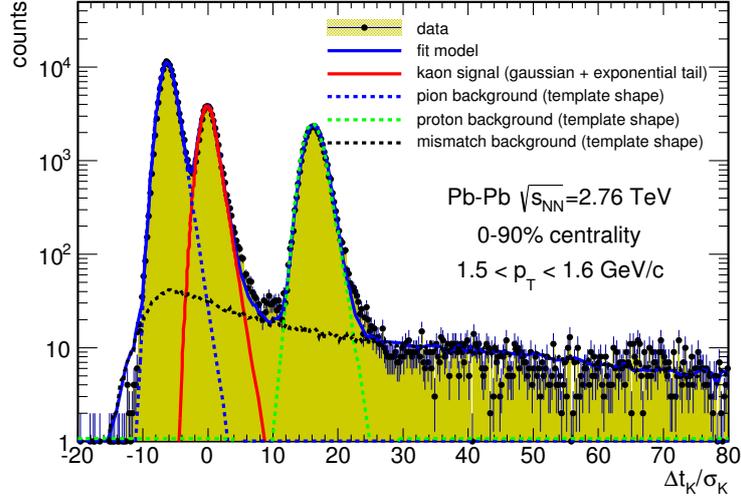
Figure 3.11: Comparison of t_{TOF} measurements for particles at $1.5 < p_{\mathrm{T}} < 1.6\,\mathrm{GeV}/c$ in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$, compared with the expectation for kaons, in units of the expected resolution for kaons. The peaks for pions, kaons and protons are fitted with Gaussian functions. Figure from [144].

While each individual detector is able to provide reliable PID information on its own, PID is at its most powerful when the signals from multiple detectors can be combined directly. An example of this is shown in Fig. 3.13 as a comparison of the TPC and TOF responses and their respective expectations for pions, kaons and protons for $2.5 < p_{\mathrm{T}} < 3.0\,\mathrm{GeV}/c$. A clear separation in $\mathrm{d}E/\mathrm{d}x$ can be seen between pions and kaons, and in t_{TOF} between kaons and protons. The t_{TOF} bands for kaons and pions are merged, as are the $\mathrm{d}E/\mathrm{d}x$ bands for protons and kaons. In each case, however, the point-to-point separation between the expected signals for each pair of species is larger if the signals are combined than if each detector (axis) is considered alone.

The simplest approach to performing PID in a given detector is through a selection on its raw signal, S, e.g. t_{TOF} or $\mathrm{d}E/\mathrm{d}x$. Most commonly, this is done using some discriminating variable ξ,

$$\xi = f(S, R), \tag{3.5}$$

where R is the expected detector response for a specific particle species. R is usually dependent on the properties of each particle track, for example the charge (Z), the momentum (p), or the track length (L). While the response function can be Gaussian for certain detectors, it more commonly represents a more complex parameterisation that takes various subtle detector effects into account.

Assuming a Gaussian response, e.g. the $\mathrm{d}E/\mathrm{d}x$ response of the TPC after the
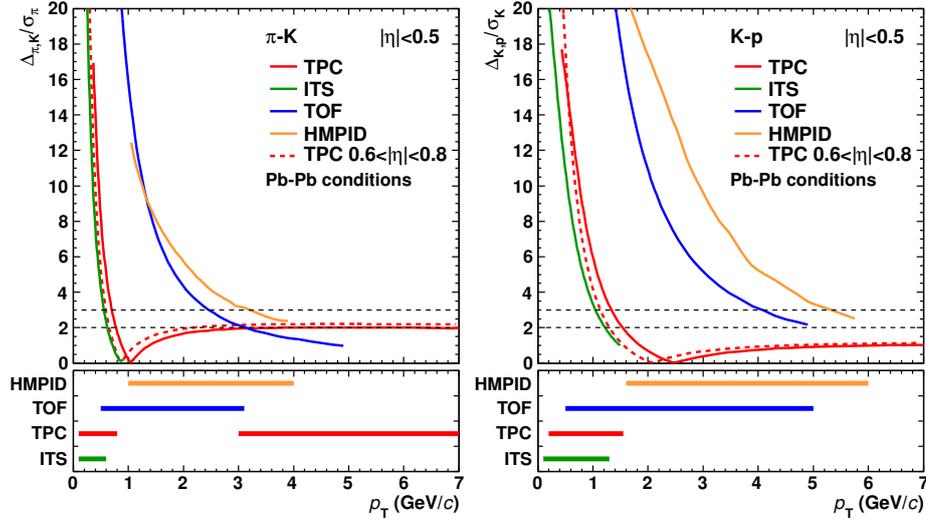
Figure 3.12: Hadron PID separation power in the central barrel of ALICE as a function of p_T. Left (right): separation between pions and kaons (kaons and protons). The value on the y-axis is the peak-to-peak distance divided by the pion (kaon) resolution, averaged over $|\eta| < 0.5$. The lower panels show the ranges for which the separation power is better than 2σ for each detector. Figure from [144].
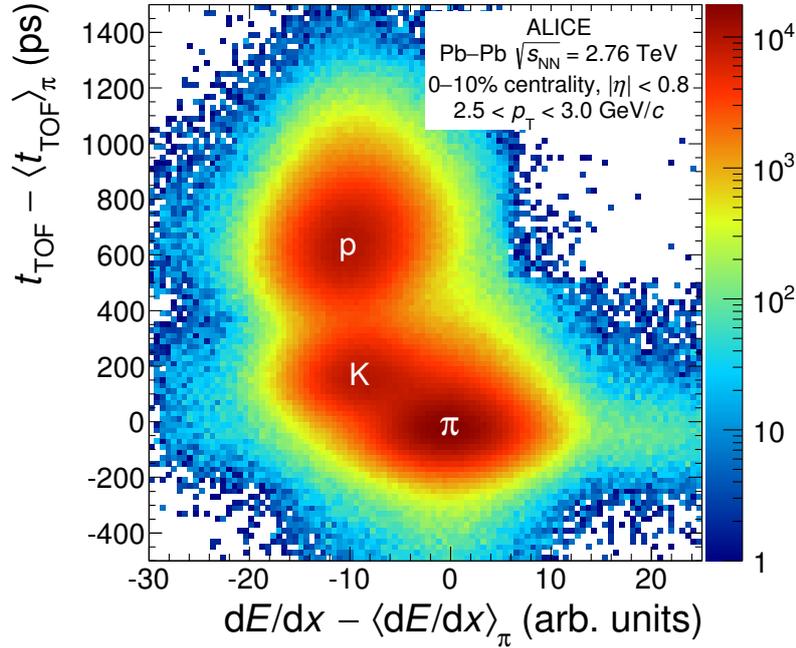
Figure 3.13: Example of combined TPC–TOF particle identification. The TPC response (x-axis) and TOF response (y-axis) are shown for charged pions, kaons and protons with respect to the expectation for pions. Figure from [6].

subtraction of the Landau tail through the use of the truncated mean, R can be described for the species hypothesis H_i by the expected mean signal $\hat{S}(H_i)$ and the expected detector resolution σ. In ALICE, H_i typically refers to electrons, muons, pions, kaons or protons, but can also include light nuclei such as deuterons, tritons, ^{3}He nuclei and ^{4}He nuclei. In most analyses ξ is represented by the n_σ variable ('number of sigmas'), which is unitless and describes the distance of a signal from the expectation in units of the resolution. For a species i and a detector α, n_σ may be expressed as

$$n_{\sigma_\alpha^i} = \frac{S_\alpha - \hat{S}(H_i)_\alpha}{\sigma_\alpha^i}. \tag{3.6}$$

Here, the resolution is denoted as σ_α^i due to it being dependent on both the detector and the species measured. For simplicity, this is taken as read in the rest of this thesis, and so is referred to simply as σ. It is possible to make simple selections based on the n_σ variable; typically this will involve a cut of $|n_\sigma| < 2$ or $|n_\sigma| < 3$, in which all particles lying within the given region around the expectation are accepted as being compatible with that hypothesis. In regions where the detector responses are well separated for different species, this leads to a clear identity being assigned to each track; in other cases, a particle may be selected as being compatible with multiple mass hypotheses. For cases where the response is perfectly Gaussian, it is possible to combine the n_σ values as a quadratic sum to account for multiple detectors. However, if the parameterisations are more complex, a more sophisticated approach is needed, such as the Bayesian approach outlined in Section 5.

4 | Data analysis

The methods used to measure open charm production in this thesis depend on kinematically reconstructing the hadronic decays of D mesons. This is performed at mid-rapidity ($|\eta| < 0.8$) using the ALICE detector, corresponding to $|y| < 0.5$ in pp collisions and $-0.96 < y_{\mathrm{cms}} < 0.04$ in p–Pb collisions. While ALICE also performs measurements of open heavy-flavour hadrons in semileptonic decay channels, both to electrons at mid-rapidity (e.g. [127]) and to muons at forward rapidity (e.g. [59]), the hadronic reconstructions carry the advantage that they allow the full kinematics of the initial state to be constructed (the semileptonic decays involve a neutrino, which is undetectable).

This section deals with the analysis strategies that are used for the reconstruction of D mesons in the ALICE central barrel. Section 4.1 outlines the datasets that were used for the analyses in this work. Section 4.2 covers the topological selections that are used to reconstruct a D^0 or D^{*+}-meson candidate from its decay tracks. Section 4.3 briefly discusses the PID selections that are used to identify the daughter tracks from the decays. Finally, Section 4.4 summarises the signal extraction procedures that are used to obtain D-meson yields from the invariant mass spectra.

4.1 Datasets

The analyses presented in this work make use of data collected by the ALICE detector during the pp and p–Pb campaigns during Run 1 of the LHC (2009–2013).

The pp data used in Section 5 were from the 2010 pp running periods. Two proton beams were collided, each with a beam energy $E_{\mathrm{beam}} = 3.5\,\mathrm{TeV}$, leading to a collision energy in the centre-of-mass system of $7\,\mathrm{TeV}$. The dataset corresponds to $\sim 3 \times 10^8$ events and an integrated luminosity of $5\,\mathrm{nb}^{-1}$. The data were collected with a minimum-bias trigger, requiring coincident hits in both of the V0 detectors, and a primary vertex within $|z_{\mathrm{vtx}}| < 10\,\mathrm{cm}$ around the centre point of the detector coordinate system. The triggering efficiency is $\sim 87\%$ for inelastic pp collisions overall, and 100% for events containing a D meson [66].

The p–Pb data used in Section 6 were collected in 2013, again using a minimum-bias trigger requiring coincident hits in both the V0A and V0C detectors and a pri-

mary vertex within $|z_{\text{vtx}}| < 10\,\text{cm}$. The dataset corresponds to $\sim 1 \times 10^8$ events, and an integrated luminosity of $48.6 \pm 1.6\,\mu\text{b}^{-1}$. The triggering efficiency in this dataset is 96.4% for p–Pb events overall, and 100% for events containing a D meson [17]. Due to the asymmetry of the system and differing charge-to-mass ratio between the proton and neutron, the beam energy was different for the two species, $E_{\text{beam}}^{\text{p}} = 4\,\text{TeV}$ and $E_{\text{beam}}^{\text{Pb}} = 1.58\,\text{TeV}$ per nucleon, leading to a collision energy in the centre-of-momentum system per binary nucleon pair of $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$. The asymmetry in the beam energies leads to a shift in the rapidity of the system with respect to the laboratory frame, $\Delta y = 0.465$, in the proton-going direction. As mentioned in Section 2.2, this shift causes a change to the definition of the coordinate system of the collision: particles moving in the direction of the proton are termed as having positive, or 'forward', rapidity, and particles in the Pb-going direction are termed as having negative, or 'backward', rapidity.

ALICE performs proton–lead collisions in two configurations, termed as 'p–Pb' and 'Pb–p'. The p–Pb system has the Pb nucleus travelling anticlockwise around the LHC ring, in the direction of the V0A with respect to the centre of ALICE (positive z-direction), and Pb–p has the Pb nucleus travelling clockwise, towards the V0C detector and muon arm. The minimum-bias p–Pb dataset only involved collisions in the p–Pb configuration, and so the V0A detector is termed as being at 'backward' rapidity in the following. This data sample was recorded over two data-taking periods, referred to later in this thesis as 'LHC13b' and 'LHC13c'.

4.2 Topological selections

After the initial hadronisation, D^0 mesons have a non-zero lifetime before weakly decaying. The mean lifetime of a D^0 meson is $c\tau = 122.9\,\mu\text{m}$, and the 'golden channel' in which they are measured in ALICE is $\text{D}^0 \to \text{K}^- \pi^+$ (and respective charge conjugates, branching fraction $\mathcal{B} = 3.93 \pm 0.04\%$) [23]. The kinematics of the decay can be used in order to reconstruct candidates, and the tracking and vertexing information provided by the TPC and ITS are crucial for this.

A schematic diagram of the decay of a D^0 meson can be seen in Fig. 4.1. The D^0 meson is produced through hadronisation at the primary vertex, and then flies some distance before decaying at the secondary vertex. The blue and red arrows respectively represent the trajectories of the pion and kaon that are produced by the decay. Their trajectories cross at the secondary vertex, and are extrapolated beyond this in order to determine their impact parameters d_0 (the smallest distance between the secondary particle tracks and the primary vertex), which are indicated by the pale green arrows. The dashed line represents the flight line of the D^0 meson between the primary and secondary vertices, and its expected trajectory in the absence of a
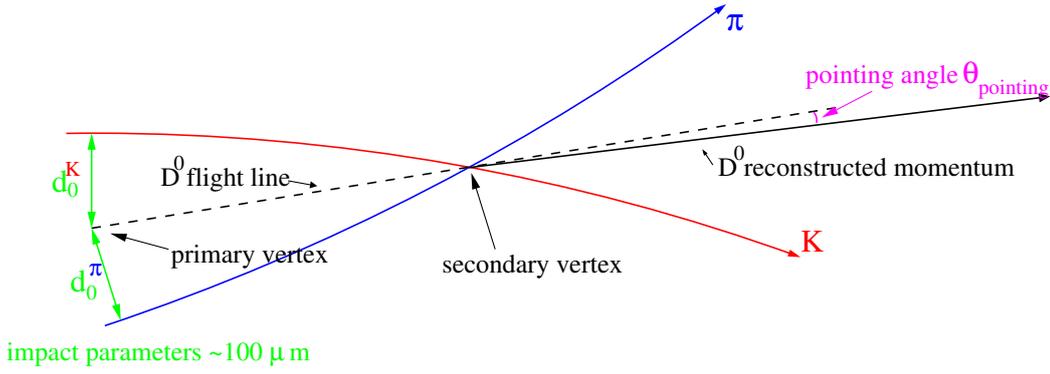
Figure 4.1: Schematic diagram of the kinematics of a $D^0 \to K^- \pi^+$ decay. Figure from [163].

decay. The distance between the two vertices is referred to as the decay length. The solid black arrow shows the total momentum of the reconstructed kaon–pion pair, and the angle between this momentum and the D-meson flight line is referred to as the 'pointing angle' θ_{point}.

These kinematic properties can be exploited in order to strongly reduce the level of combinatorial background that would be seen when simply reconstructing all kaon-pion pairs. Typically, selections are made on the following geometrical properties:

– DCA, the distance of closest approach between the two daughter tracks around the secondary vertex;

– the decay length (distance between the primary and secondary vertices);

– L_{xy}, the projection of the decay length onto the transverse plane, normalised by its resolution;

– $\cos(\theta_{\mathrm{point}})$, the cosine of the pointing angle;

– $\cos(\theta^*)$, the cosine of the angle between the kaon flight line in the D^0 rest frame and the boost direction; and

– $\left| d_0^{\mathrm{K}} \right|$, $\left| d_0^{\pi} \right|$, and $d_0^{\mathrm{K}} \times d_0^{\pi}$, the impact parameters of the two daughter tracks measured in the plane transverse to the beam direction, and their product.

The x–y plane is used for the decay length measurement as the azimuthal resolution in the ITS is finer than the longitudinal resolution. Further kinematic constraints are also applied to the D-meson candidate itself, namely:

– ΔM_{D^0}, the difference between the reconstructed invariant mass of the D^0 meson and the world average ($1864.84 \pm 0.05\,\mathrm{MeV}/c^2$ [23]); and

- $p_{\mathrm{T}}^{\mathrm{K}}$ and p_{T}^{π}, the transverse momenta of the secondary particle tracks.

For the D^{*+} meson, the channel that is measured is the strong decay $\mathrm{D}^{*+} \to \mathrm{D}^0 \pi^+$ (and charge conjugates, $\mathcal{B} = 67.7 \pm 0.5\%$ [23]). The lifetime of this resonance is extremely short ($c\tau \approx 2.4\,\mathrm{pm}$, from $\tau = \hbar/\Gamma$ with the decay width $\Gamma = 83.4 \pm 1.8\,\mathrm{keV}$ [23]), meaning that the decay essentially occurs at the production vertex. In addition, the phase space of this decay is limited due to the small mass difference between the D^{*+} meson and the $\mathrm{D}^0 \pi^+$ final state (see Section 4.4), limiting the momentum of the charged pion. The decay is therefore studied by pairing a reconstructed D^0 candidate with a soft (i.e. low-momentum) pion at the primary vertex. The usual selections apply for the D^0 candidate as above, and further selections are made on:

- $\Delta M_{\mathrm{D}^{*+}} = M_{\mathrm{K}\pi\pi} - M_{\mathrm{D}^0}$, the difference in invariant mass between the reconstructed triplet and the D^0 meson candidate;

- $p_{\mathrm{T}}^{\pi_{\mathrm{s}}}$, the transverse momentum of the soft pion; and

- $\theta_{\pi_{\mathrm{s}}}$, the angle between the soft pion's momentum and the decay plane of the D^0 meson (defined by the momentum vectors of the D^0 decay daughters).

The secondary vertices were reconstructed with tracks in a pseudorapidity range $|\eta| < 0.8$, each with transverse momenta $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}/c$. In addition, they were required to have at least 70 (out of a maximum possible 159) hits in the TPC, and at least one hit (out of two) in the SPD. The soft pion candidate from the D^{*+}-meson decay was more loosely selected, with the requirement of $p_{\mathrm{T}} > 0.1\,\mathrm{GeV}/c$ and at least three (out of a possible six) hits in the ITS. No PID selection criteria were applied to this pion.

The topological and kinematic selections are chosen based on information from Monte Carlo simulations, with the aim of maximising the signal and minimising the background. This is done using a charm-enriched Monte Carlo sample to simulate the signal, and a minimum-bias sample for the estimate of the background. Efficiencies are also determined for the selection cuts from Monte Carlo simulations, see Section 4.4. As the kinematics of the decay tracks change with the momentum of the particle, the selections are varied as a function of p_{T}. The specific cut values that were used for the analyses in this thesis were chosen to match those already used in the analyses published in [66, 123], and are given in the analysis-specific sections (Tab. 5.4 for D^0 mesons in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$, and Tab. 6.3 for D^{*+} mesons in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$). The systematic uncertainties assigned to variations in the selection cuts, based on the reliability of their description in the Monte Carlo simulations for efficiency determination, are $\sim$10% for D^0 mesons in pp collisions [66], and $\sim$8% for D^{*+} mesons in p–Pb collisions [123].

4.3 D-meson PID

In addition to the kinematic selections mentioned above, the standard analysis of D mesons in ALICE relies heavily on information from the TPC and TOF detectors for the identification of the decay daughters, as outlined in Sections 3.2 and 3.4.

Typically, compatibility cuts are applied to the signal of either the TPC or TOF detector. As it has greater separation power for a larger range of momenta, the TOF is primarily used where available; however, if a track has no valid TOF hit, the signal in the TPC can be used alone. The standard analysis of D mesons makes use of n_σ compatibility selections (see Eq. (3.6)) for the PID, i.e. particles are accepted as being compatible with the pion or kaon hypothesis if they lie within a certain multiple of σ from the expectation for each detector.

For the analysis of D^0 mesons in pp collisions presented in Section 5 using n_σ PID (consistent with the analysis in [66]), pions and kaons were selected based on a 3σ compatibility cut around the mean for both detectors individually. Where information from the TOF detector was available, this was used; otherwise, the TPC dE/dx was used. For the analysis of $D^{*+} \to D^0 \pi^+$ in p–Pb collisions (see Section 6), PID was used only for reconstructed D^{*+} candidates with $p_T < 8\,\text{GeV}/c$. No PID requirements were applied to the soft pion, only the topological selections referred to in Section 4.2. A 2σ compatibility selection was made on both the kaon and pion decay tracks from the D^0 meson in the TPC, and a 3σ selection in the TOF signal if available.

The PID selection efficiency in both cases is roughly 97%. The systematic uncertainty on the n_σ PID method used for D^0 mesons measured in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ and for D^{*+} mesons in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$, calculated by comparing the efficiency-corrected yield obtained with n_σ PID with that obtained without PID is of the order of 5% [66, 123].

4.4 Signal extraction and correction

D-meson candidates that pass the selection cuts are used to fill invariant mass distributions for each analysed p_T interval. In multiplicity-dependent analyses, they are generated per combination of p_T and multiplicity class, as well as for the multiplicity-integrated case. For D^0 mesons, the invariant mass of the reconstructed kaon–pion pair ($M_{K\pi}$) is computed and used for this; for D^{*+}, the quantity $\Delta M = M_{K\pi\pi} - M_{D^0}$ is used. The motivation for this is that the difference between the D^{*+} and D^0 masses is small ($145.4257 \pm 0.0017\,\text{MeV}/c^2$), and is only slightly larger than the pion mass ($139.5702 \pm 0.0004\,\text{MeV}/c^2$), thus limiting the phase space of the produced D^0 and π^+ decay daughters and meaning that the majority of their kinetic energy comes from the D^{*+} meson's own kinetic energy, rather than from the mass difference between

the initial and final state. This restricts the kinematic phase space of the decay, thus giving a greatly improved mass resolution for this difference as compared to simply measuring the invariant mass of the reconstructed triplet. The D^0-meson spectrum also contains a contribution from the decay $\overline{D}^0 \to K^- \pi^+$, which manifests as an incorrect assignment of the D-meson invariant mass; however, as this decay is doubly Cabibbo-suppressed, its contribution to the yield is negligible [66].

D-meson signals are then extracted from the histograms by fitting them with pre-defined functions. For both D^0 and D^{*+} mesons, the signal peak is fitted with a Gaussian function. The signal also sits on top of some amount of combinatorial background, which must be subtracted to obtain a final yield. For the D^0 meson, this is modelled by an exponential function; for D^{*+}, this is an exponential multiplied by a power-law function, given by

$$B(\Delta M) = a\sqrt{\Delta M - M_\pi} \cdot e^{b(\Delta M - M_\pi)}, \qquad (4.1)$$

where ΔM is the mass difference between the reconstructed triplet and the D^0 candidate, and a and b are free fit parameters. The background is first determined by considering only points within side-bands of the invariant mass distributions, i.e. regions $> 4\sigma$ away from the signal peak, where σ is the signal width. This approximation is then narrowed by recomputing this fit over the full invariant mass distribution to give the final background estimate. This function is integrated in a 3σ region around the mass peak to give the background that is subtracted from the yield. As part of the systematic uncertainty determination, the choice of background may also be varied (for D^0, a linear, power-law or polynomial function; for D^{*+}, a power-law function).

Example invariant mass histograms for D^0 mesons with $2 < p_T < 3 \, \text{GeV}/c$ in pp collisions at $\sqrt{s} = 7 \, \text{TeV}$, and for D^{*+} mesons with $4 < p_T < 8 \, \text{GeV}/c$ in p–Pb collisions at $\sqrt{s_{NN}} = 5.02 \, \text{TeV}$, are shown in Fig. 4.2 along with their fits. The yields are extracted directly from the integral to the total fit, with the integral of the background fit subtracted. The alternative method of bin counting is used to determine the systematic uncertainties: here, the total of signal and background are taken from the contents of the bins of the invariant mass histogram within some distance around the mean, and the background is subtracted based on the fit function within the signal region. The yield extraction uncertainty for D^0 mesons in pp collisions at $\sqrt{s} = 7 \, \text{TeV}$ was found to be $\sim 20\%$ at low p_T and $\sim 10\%$ at high p_T [66]. For the minimum-bias analysis of D^{*+} mesons in p–Pb collisions presented in [123], the yield extraction uncertainty was found to be between $\sim 2\%$ at intermediate p_T and $\sim 10\%$ at low p_T; the determination and size of the yield extraction uncertainties on the multiplicity-dependent analysis of D^{*+} mesons presented in this work is discussed in more detail in Section 6.2.

In order to give a physical quantity such as a cross section, the raw yields ex-
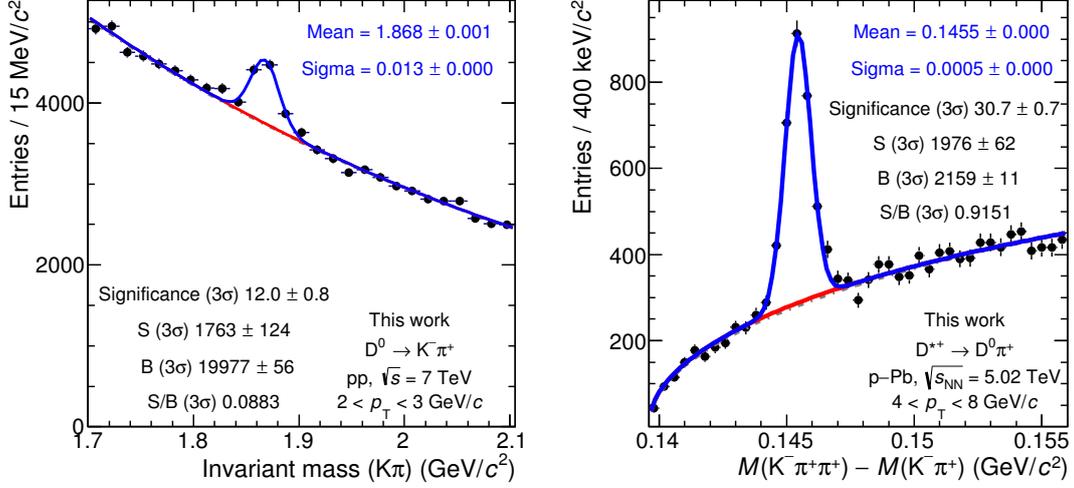
Figure 4.2: Invariant mass histograms for D-meson reconstruction. Left: D^0 mesons with $2 < p_T < 3\,\text{GeV}/c$ in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. Right: D^{*+} mesons with $4 < p_T < 8\,\text{GeV}/c$ in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$. The background functions are given by red curves; the sum of signal + background by blue curves. Each plot shows the signal, background statistical significance and signal-to-background ratio given by the fit in the signal region.
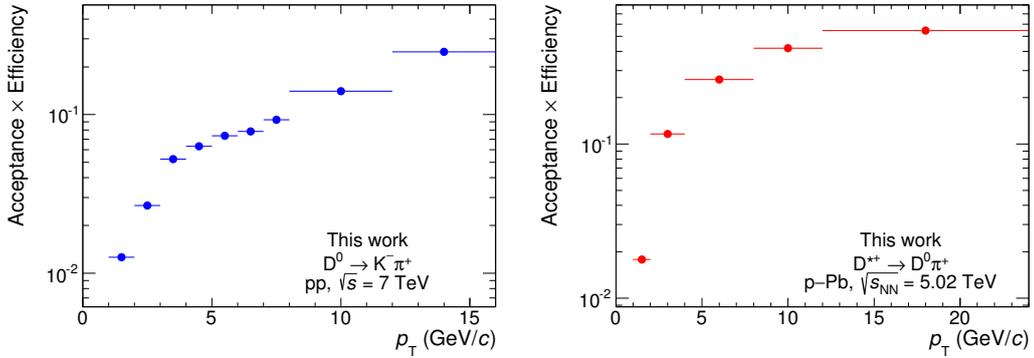
Figure 4.3: Reconstruction and selection efficiencies for (left) D^0 mesons in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ and (right) D^{*+} mesons in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$, as a function of p_T.

tracted by this procedure must be corrected for their reconstruction and selection efficiencies, ε. These are determined in charm-enriched Monte Carlo simulations based on the PYTHIA 6.4.21 event generator [164] using the Perugia-0 tune [165], with the GEANT3 package [166] used to simulate the detector response, geometry and material budget. Further minimum-bias simulations also allow for the determination of acceptance factors α based on the geometrical acceptance of the detectors. The reconstruction and selection efficiencies for prompt D^0 mesons in pp collisions and

D^{*+} mesons in p–Pb collisions used in this thesis are shown in Fig. 4.3.

The efficiencies vary as a function of p_T and multiplicity, although the differences as a function of multiplicity are small as the detector response is not significantly multiplicity-dependent in p–Pb collisions, and the selection cuts are only varied with p_T. For both D^0 mesons in pp collisions and D^{*+} mesons in p–Pb collisions, the efficiency is smallest at low p_T and increases at higher p_T, before reaching a plateau. The low efficiency at low p_T arises due to the strength of the topological cuts in this region; as the cuts loosen at higher p_T, the efficiencies rise, before saturating due to the detector acceptance.

5 | A Bayesian approach to PID in ALICE

Particle identification (PID) is one of the major strengths of ALICE, and in particular is essential for the reconstruction of heavy-flavour particle decays, as outlined in Section 4. As explained in Section 3.4, the ALICE detector profits from having access to multiple complementary PID technologies. Often, compatibility cuts are applied separately to the individual detectors; however, PID is at its most effective when these signals can be combined with one another in a simple manner. This can be done by applying a probabilistic, rather than frequentist, approach to the PID responses of the detectors. In addition, in regions where the separation between the expected signals for two species is smaller than the detector's separation power, it is possible for tracks to be assigned more than one possible identity using the n_σ compatibility approach; in such cases it may be useful to glean the likelihood for each possible species to judge which hypothesis provides the best match to the signal.

The Bayesian PID technique is one such approach, allowing probability estimates to be made for each species for a given detector signal. These estimates can then be combined as a simple product if many detectors are used, and the resulting probability can be used in selections.

A Bayesian PID approach is already used in the ALICE Transition Radiation Detector (TRD) in order to combine the signals induced in different layers using their likelihoods [167], and was proposed during the early stages of the design of ALICE for combining PID signals from different detectors [117,118]. However, it was decided that a comprehensive understanding of the detector responses was necessary before this method could be fully implemented and used in analyses. In order to benchmark this new method, a detailed study of the Bayesian PID approach was performed in multiple decay channels in order to test its performance, and crucially, to ensure that its use does not bias the final results of an analysis through incorrect evaluations of the PID efficiency. In order to do this, it was tested in ALICE for high-purity samples of pions, kaons and protons from two-pronged V^0 particle decays, single-particle spectra, and a full analysis of the decays of D^0 mesons in the channel $D^0 \to K^- \pi^+$ [6].

This section outlines the basis behind the Bayesian method in the context of the ALICE PID framework (Section 5.1), with a focussed discussion of the determination

of prior probabilities in ALICE (Section 5.2). Section 5.3 discusses the PID efficiency determination for Bayesian PID, and Section 5.4 compares a variety of possible strategies that are applied in order to translate the Bayesian probabilities into a decision on the particle species. Section 5.5 motivates the use of Bayesian PID in ALICE, with reference to a trial measurement of $\Lambda_c^+ \to pK^-\pi^+$ performed by ALICE in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ [6]. Finally, Section 5.6 presents the case study performed for this thesis, in a full analysis of D^0-meson production in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. The validation focussed on the major PID detectors in ALICE that had full azimuthal coverage during RUN 1 of the LHC, i.e. TPC and TOF; however, the method can also be applied to the other PID detectors in ALICE (ITS, TRD, the Electromagnetic Calorimeter EMCal, and the High-Momentum PID Detector HMPID). The work presented in this section was also published in [6].

5.1   Bayesian particle identification

The Bayesian PID approach uses the conditional probability that a signal S will be seen in the detector for some hypothesised species H_i, $P(S|H_i)$. Assuming a Gaussian response, for example in the TPC (after the subtraction of the Landau tail) and TOF detectors, this probability corresponds to

$$P(S|H_i) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\text{n}_\sigma^2} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(S-\hat{S}(H_i))^2}{2\sigma^2}}, \tag{5.1}$$

where n_σ is the difference between the measured signal S and the expected signal $\hat{S}$ for the species H_i, in units of the detector resolution σ, as in Eq. (3.6). The probability distribution for a given detector is therefore intrinsically defined by the detector response itself. For detectors with non-Gaussian responses, the probability is instead given by a suitable alternative parameterisation.

The use of probabilities is advantageous at it converts the detector responses to entirely unitless quantities that may simply be combined multiplicatively. This greatly simplifies the problem of combining the detector responses for different detector technologies. The combined probability for a particle that belongs to the species H_i to give a set of signals $\vec{S}$ for a series of detectors is thus

$$P(\vec{S}|H_i) = \prod_{\alpha=\text{TPC,TOF,...}} P_\alpha(S_\alpha|H_i), \tag{5.2}$$

where $\vec{S} = (S_\text{TPC}, S_\text{TOF},...)$, and P_α and S_α respectively represent the probability and signal for a detector α.

While $P(\vec{S}|H_i)$ represents the probability of measuring the combination of signals $\vec{S}$ given the knowledge of the particle species, the more interesting quantity for PID

is the probability of a particle being of some species H_i given a set of observed signals (i.e. $P(H_i|\vec{S})$). The relationship between these two conditional probabilities is given by Bayes' theorem [5],

$$P(H_i|\vec{S}) = \frac{P(\vec{S}|H_i)C(H_i)}{\sum_{k=e,\mu,\pi,\ldots}P(\vec{S}|H_k)C(H_k)}, \qquad (5.3)$$

where $C(H_i)$ represents the 'a priori' (or 'prior') probability of measuring the particle species H_i according to its relative abundance, and $P(H_i|\vec{S})$ is referred to as the 'a posteriori' (or 'posterior') probability. The summation on the denominator of this equation combines the probabilities of all possible species, normalising them such that $\sum_{H_i} P(H_i|\vec{S}) = 1$. Selections based on the probability $P(H_i|\vec{S})$ in this case effectively correspond to selecting on the purity of the species H_i in the data sample, i.e. the proportion of correctly identified H_i particles over all particles identified as H_i.

H_i can be any long-lived charged particle; in ALICE this typically covers electrons, pions, kaons, protons, and light nuclei such as deuterons, tritons and ^{4}He nuclei. Although muons may also be considered, they are usually neglected in the priors for the central-barrel detectors at mid-rapidity, as their mass is very similar to pions, making the two species almost indistinguishable over a wide p_T range. This would lead to a reduction in the pion efficiency if muons were included [6].

5.2 Bayesian priors in ALICE

The priors $C(H_i)$ in Eq. (5.3) represent an approximate measurement of the true abundances of each particle species. Due to possible differences in the particle ratios between collision systems, they are determined separately for pp collisions at $\sqrt{s} = 7\,\text{TeV}$, Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76\,\text{TeV}$, and p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$. They take the form of a p_T distribution of the yield of each particle species in all collision systems, but are also computed in centrality classes for Pb–Pb collisions and V0A multiplicity classes for p–Pb collisions, as the abundances and detector response may also vary with these parameters. In order to normalise the distributions, they are expressed as relative abundances with respect to pions (the most abundant species) in the sample.

In each system, the priors are computed via an iterative method, itself based upon a Bayesian approach in the TPC. Tracks were selected with a minimal set of single-track cuts (a minimum of 80 out of a possible maximum 159 clusters in the TPC, and maximum distances of closest approach of 3 cm from the primary vertex in both the z-direction and x–y plane) For each step, n, the iteration proceeds as follows. In the first step ($n = 0$), all of the distributions are set to unity for all p_T (this case is hereafter referred to as 'flat priors'). Bayesian posterior probabilities are

then computed for each track and species hypothesis according to Eq. (5.3), giving $P_n(H_i|S)$. The identified yields, which are expressed as p_T spectra, $Y(H_i, p_T)$, are then filled for the step $n+1$ as a sum over all measurements in each p_T interval, with the single-track probabilities used as weights:

$$Y_{n+1}(H_i, p_T) = \sum_S P_n(H_i|S). \tag{5.4}$$

The results for each species are then normalised against the pion yield for that step, to give the priors to be used in the following step:

$$C_{n+1}(H_i, p_T) = \frac{Y_{n+1}(H_i, p_T)}{Y_{n+1}(H_\pi, p_T)}. \tag{5.5}$$

This process is then repeated, with the posterior probabilities from one step fed back as prior probabilities for the next, until the priors converge (i.e. until they no longer change with successive iterations). This is demonstrated in Fig. 5.1, where the kaon priors (the K/π ratios) for p–Pb collisions in the 10–20% V0A multiplicity class are shown for various steps of the iteration, in the left panel as the priors at each step, and in the right panel as the ratio of each prior distribution to that of the preceding step, C_n/C_{n-1}. It was found that the prior extraction stabilised after approximately five iterations.
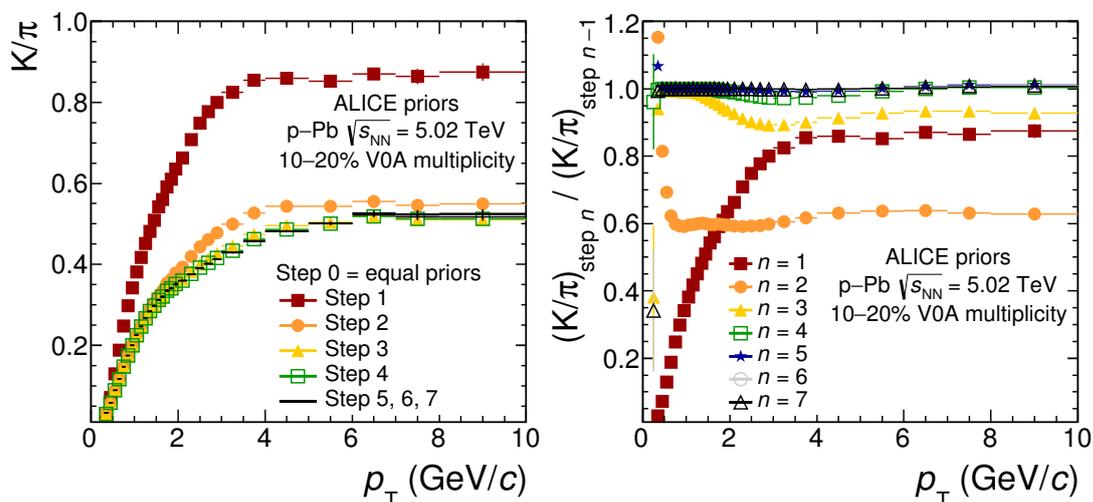
Figure 5.1: An example of the iterative prior extraction procedure for data from p–Pb collisions (for the 10–20% V0A multiplicity class). The extracted K/π ratio of the priors is shown as a function of p_T at each step of the iteration (left) and as a ratio of the value between each successive step (right). Step 0 refers to the initial ratio, which is set to unity for all p_T. Figure from [6].

These priors were determined for global tracks, i.e. tracks reconstructed in both

the ITS and the TPC. Priors extracted via this method in each collision system form the basis of the Bayesian PID approach in ALICE, and thus are referred to as the 'standard priors'. In the following, these standard priors are referred to as $C_{\text{TPC}}(H_i)$. These priors are valid for all TPC tracks fulfilling minimal track selection criteria; however, as the geometrical acceptance and material budget varies for each detector, corrections must be made to the priors in order for them to be useful for other central-barrel detectors. This is achieved through the use of so-called 'propagation factors' (F_α), which vary for each detector and are dependent on the particle species, p_{T} and charge. The factors are based on the ratio between the efficiency of matching the TPC to another detector for pions and that of another species, based on parameterisations of the detector response in Monte Carlo simulations. As with the priors themselves, the propagation factors are calculated relative to the pion distribution. As an example, the F_{TOF} distributions for positive and negative kaons and protons are shown for $p_{\text{T}} > 0.75\,\text{GeV}/c$ in Fig. 5.2. The factors are determined separately for positive and negative particles, due to differing levels of absorption in the detector material. This leads to a smaller propagation factor overall for negative particles than positive particles, especially at low p_{T}.
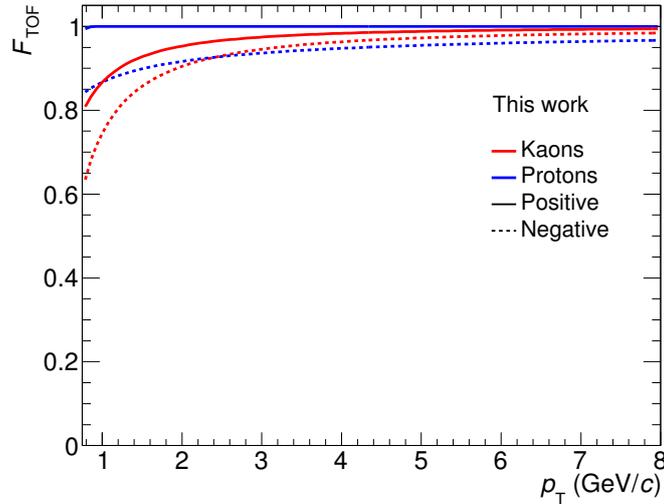
Figure 5.2: Propagation factor (F_{TOF}) distributions used to propagate the K$/\pi$ and p$/\pi$ priors from TPC to TOF, as a function of p_{T}. The solid lines are the parameterisations for positive particles; the dashed lines, for negative particles. The paramaterisations are determined for $p_{\text{T}} > 0.75\,\text{GeV}/c$.

For a detector α and species i, the propagated prior distribution is defined as

$$C_\alpha(H_i, p_{\text{T}}) = F_\alpha(H_i, p_{\text{T}}) \times C_{\text{TPC}}(H_i, p_{\text{T}}). \qquad (5.6)$$

This propagation is currently available in ALICE for the TRD, TOF, EMCal and

HMPID detectors; separate sets of priors were also determined in ALICE for tracks based only on ITS hits ('ITS standalone') for use at low momenta.

As the priors represent the relative abundance of each particle species, they can be checked for consistency with the data by directly comparing them with the true abundances of pions, kaons and protons measured by ALICE in pp, Pb–Pb and p–Pb collisions [168–170]. This comparison is shown for the K/π and p/π ratios for pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ and Pb–Pb collisions at $\sqrt{s_\mathrm{NN}} = 2.76\,\mathrm{TeV}$ in Fig. 5.3, and for p–Pb collisions at $\sqrt{s_\mathrm{NN}} = 5.02\,\mathrm{TeV}$ in Fig. 5.4. In each figure, the upper panels show the particle ratios for the measured spectra (filled symbols) and the extracted Bayesian priors (open symbols) for various centrality and multiplicity classes for Pb–Pb and p–Pb collisions, respectively. The lower panels show the ratios of the spectra to the priors. In each case, the extracted priors were within roughly 10% of the measured spectra over the majority of the p_T range for all collision systems [6].

Further checks were made in order to ensure that the observed differences between the priors and the measured abundances would not have an adverse effect on later analyses. In [6], this was performed for the analysis of single-particle spectra using flat priors (equal to unity for all species); in this work (see Section 5.6) it was also performed for the analysis of $\mathrm{D}^0 \to \mathrm{K}^- \pi^+$ using modified priors with a 10% surplus and deficit of kaons, altered uniformly over all p_T, in order to match the size of the kaon–pion discrepancy seen in Fig. 5.3. In both cases, no significant deviation was found between the efficiency-corrected yields from the Bayesian PID method using the standard and altered priors.

5.3 PID efficiency, contamination, and misidentification

Ideally, the PID selections would correctly identify all reconstructed particles in the sample, thus allowing the direct extraction of yields from the invariant mass spectra. However, due to the limited detector resolution, as well as the inclusion or exclusion of particles based on selection criteria, corrections must be made to the raw yields in order to extract a physical quantity (production cross section, p_T spectrum, etc.). These corrections are based on estimates of the efficiency of the selection criteria. The PID efficiency is defined in terms of each individual species i; for simplicity, this section deals with the most common case of three hadron species (pions, kaons, protons), but in practice this can take the form of any of the species listed in Section 5.1. The PID efficiency can be expressed by the efficiency matrix ϵ^PID,

$$\epsilon^\mathrm{PID} = \begin{pmatrix} \epsilon_{\pi\pi} & \epsilon_{\pi\mathrm{K}} & \epsilon_{\pi\mathrm{p}} \\ \epsilon_{\mathrm{K}\pi} & \epsilon_{\mathrm{KK}} & \epsilon_{\mathrm{Kp}} \\ \epsilon_{\mathrm{p}\pi} & \epsilon_{\mathrm{pK}} & \epsilon_{\mathrm{pp}} \end{pmatrix}, \tag{5.7}$$
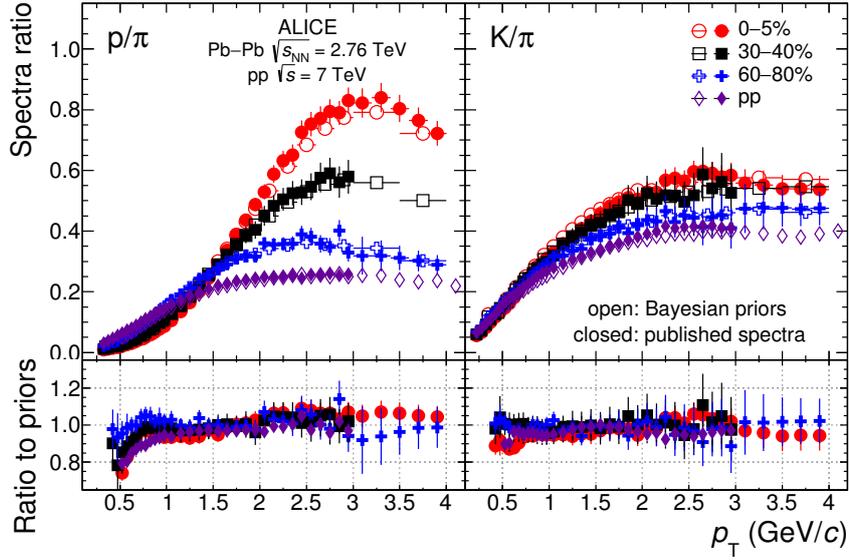
Figure 5.3: The proton/pion ratio (left) and kaon/pion ratio (right), as measured by ALICE [168, 169] using TPC and TOF PID (filled symbols), compared with the TPC priors (open symbols) for Pb–Pb and pp collisions. For Pb–Pb, the results are reported for different centrality classes. These particle ratios are calculated for $|\Delta y| < 0.5$. The double ratios (the measured abundances divided by the Bayesian priors) are shown in the lower panels. Figure from [6].
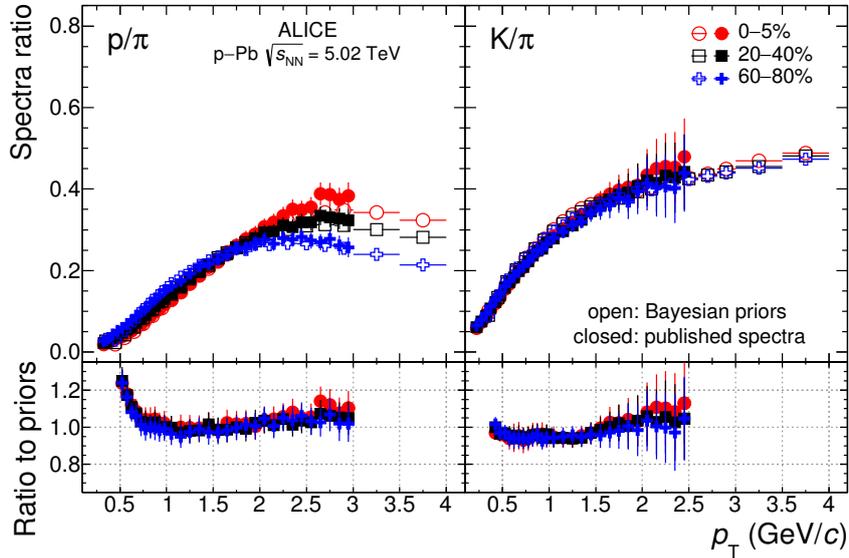
Figure 5.4: The proton/pion ratio (left) and kaon/pion ratio (right), as measured by ALICE [170] using TPC and TOF PID (filled symbols), compared with the TPC priors as described in the text (open symbols) for p–Pb collisions for different V0A multiplicity classes. Particle ratios are calculated for $|\Delta y| < 0.5$. The double ratios (the measured abundances divided by the Bayesian priors) are shown in the lower panels. Figure from [6].

where the diagonal elements (ϵ_{ii}) represent the PID efficiencies of the species i, and the off-diagonal terms ($\epsilon_{ij}, i \neq j$) give the probability of misidentifying a particle i as some other species j.

Abundance vectors for pions, kaons and protons may be defined as

$$\vec{A}_{\text{meas}} = \begin{pmatrix} \pi_{\text{meas}} \\ \text{K}_{\text{meas}} \\ \text{p}_{\text{meas}} \end{pmatrix} \text{ and } \vec{A}_{\text{true}} = \begin{pmatrix} \pi_{\text{true}} \\ \text{K}_{\text{true}} \\ \text{p}_{\text{true}} \end{pmatrix}, \tag{5.8}$$

where the elements of $\vec{A}_{\text{meas}}$ ($\vec{A}_{\text{true}}$) represent the measured (true) abundances of each species. This then allows the redefinition of the PID efficiencies as

$$\epsilon_{ii} = \frac{N_{i \text{ identified as } i}}{A^i_{\text{true}}}, \tag{5.9}$$

leading to a re-casting of the abundance vectors in terms of the transpose of ϵ^{PID}:

$$\begin{pmatrix} \pi_{\text{meas}} \\ \text{K}_{\text{meas}} \\ \text{p}_{\text{meas}} \end{pmatrix} = \begin{pmatrix} \epsilon_{\pi\pi} & \epsilon_{\pi\text{K}} & \epsilon_{\pi\text{p}} \\ \epsilon_{\text{K}\pi} & \epsilon_{\text{KK}} & \epsilon_{\text{Kp}} \\ \epsilon_{\text{p}\pi} & \epsilon_{\text{pK}} & \epsilon_{\text{pp}} \end{pmatrix}^{\text{T}} \cdot \begin{pmatrix} \pi_{\text{true}} \\ \text{K}_{\text{true}} \\ \text{p}_{\text{true}} \end{pmatrix}. \tag{5.10}$$

The true abundances of each species may thus be extracted from the measured abundances by inverting the PID matrix:

$$\vec{A}_{\text{true}} = \left(\epsilon^{\text{PID,T}} \right)^{-1} \times \vec{A}_{\text{meas}}. \tag{5.11}$$

The elements of ϵ^{PID} are altered by the strength of the PID cut, as well as having dependences on the particle p_{T} and rapidity, the collision centrality, and the event multiplicity.

The level of contamination of the PID sample must also be taken into account. The contamination c_{ji} is not the same as the misidentification probability ϵ_{ij} (the off-diagonal elements of ϵ^{PID}), and is instead defined as

$$c_{ji} = \frac{N_{i \text{ identified as } j}}{A^j_{\text{meas}}}, i \neq j. \tag{5.12}$$

It is, however, possible to connect the misidentification probability to the contamination through the following formula:

$$c_{ji} = \frac{\epsilon_{ij} A^i_{\text{true}}}{\epsilon_{jj} A^j_{\text{true}} + \sum_{j \neq k} \epsilon_{jk} A^k_{\text{true}}}. \tag{5.13}$$

In analyses, there is an interplay between the efficiency and contamination that arises due to a PID strategy. For example, a more inclusive cut, e.g. increasing the width of an n_σ PID selection, will naturally accept a larger proportion of particles of the correct species. However, broadening the selection criteria will also cause more particles of other species to be selected, especially where the separations between the expected signals for various species become smaller than the detectors' separation power. A high level of contamination can also be described as a low 'purity' in the sample, which manifests as a higher level of background in the final result, and so a lower signal-to-background ratio.

Detector	Discrepancy	Refinement through TuneOn-Data
TPC	Offset of dE/dx response	Splines extracted from data are sampled, giving an identical Gaussian response for Monte Carlo and data.
	Differing η dependence of dE/dx	Simulated signal sampled from η-dependent splines extracted from data.
	Multiplicity-dependent dE/dx signal	Simulated signal sampled from multiplicity-dependent splines extracted from data.
TOF	Non-Gaussian tails in TOF signals	Tails are added randomly to Monte Carlo signals, giving a response similar to that in data on average.
	Pb–Pb matching efficiency overestimated in MC due to geometric misalignment	Number of matched hits decreased randomly to reproduce data.
	Estimate of mismatched TPC–TOF tracks too low due to underestimated secondary particle production in Pb–Pb collisions	Mismatched hits added to Monte Carlo to reproduce pattern in data

Table 5.1: List of procedures applied to the TPC and TOF responses in Monte Carlo simulations by the 'TuneOnData' option in AliRoot.

While it is theoretically possible to estimate these quantities based on pure data, it is more reliable to base them on Monte Carlo simulations, so as to have a definition of the true species of each particle. For PID in general, and particularly for the Bayesian PID approach, care must be taken to ensure that the behaviour of the detectors in the

simulations matches the true conditions in each data taking period. This is handled in the ALICE analysis framework AliRoot by the so-called 'TuneOnData' option. This corrects a number of discrepancies that may be seen between data and Monte Carlo for the TPC and TOF, as outlined in Tab. 5.1. In the TPC, the corrections are based on splines (comparisons between the measured and expected $\mathrm{d}E/\mathrm{d}x$) extracted from the data set. Of particular importance are the multiplicity corrections, which are applied to offset the effect of the high multiplicities experienced in central Pb–Pb collisions due to the occupancy in the detector, and the η dependence correction, which accounts for the larger level of diffusion undergone by an electron ionised in the centre of the detector than one released closer to the end plates of the detector. For the TOF detector, the major tuning corrections are the addition of non-Gaussian tails to the detector response, to better emulate the true signal distribution, and the addition and subtraction of hits from the sample in Pb–Pb collisions to simulate the increased number of mismatched tracks and reduced matching efficiency seen in data for high-multiplicity Pb–Pb events.

5.4   Bayesian PID strategies

There are multiple possible ways of using the species probabilities that are computed for each track in the detector as PID selection criteria. In order to comprehensively test the capabilities of the Bayesian PID approach, multiple such strategies were implemented and compared with one another for the $\mathrm{D}^0 \to \mathrm{K}^- \pi^+$ analysis. Each method is assumed to be equally valid a priori, but comes with its own benefits and drawbacks, discussed below.

The simplest of the three strategies is the **maximum-probability** method. For this case, each particle track was identified based on the highest species probability, effectively making a selection on the 'most likely' particle type for a given detector response. This is the most intuitive case, and results in a single species being associated to each track; however, it does not take account of possible edge cases where the expected detector response functions for multiple species overlap and the probabilities for multiple species are close to one another.

The second approach is based on a **fixed probability threshold**, whereby a particle is accepted as being compatible with a given hypothesis if its probability is above some predetermined value. Due to the normalisation of the probabilities in Eq. (5.3), this effectively corresponds to a maximum-probability approach with the additional imposition of a purity requirement if the threshold is above 50%. For thresholds below 50%, on the other hand, it allows for cases where a particle is compatible with multiple species, as long as they conform to some purity requirement. In each of these two methods, one or more explicit identities are given to each particle

track by the PID procedure, and the PID efficiency ε_{ii} corresponds to the proportion of particles of a given species that are identified correctly.

The final approach tested in this thesis is a **weighted Bayesian** method. This case approaches the problem of identifying particles slightly differently to the other two methods, in that it makes no attempt to assign a specific identity per track. Instead, all particles are accepted and used to fill the final result distributions (usually invariant mass spectra), with some weight W_i. Here W_i is simply defined as the Bayesian posterior probability for the track to belong to the species i. For primary particles, this method is very similar in nature to the prior extraction procedure outlined in Section 5.2. For particles decaying into multiple tracks, W_i is defined as the overall probability that the primary particle belonged to the species under consideration, which may be defined as the joint probability of the relevant decay products in the channel under consideration. Under the assumption that the species of the two daughter tracks are independent of one another, this can be estimated by the product of the single-track probabilities. For example, for $D^0 \to K^-\pi^+$ (and charge conjugates), the weights for D^0 and $\overline{D}^0$ are defined as

$$
\begin{aligned}
W_{D^0} &= P(K^- \cap \pi^+) = P_{K^-} \times P_{\pi^+}, \text{ and} \\
W_{\overline{D}^0} &= P(K^+ \cap \pi^-) = P_{K^+} \times P_{\pi^-},
\end{aligned}
\tag{5.14}
$$

where P_i in each case corresponds to the single-track probability for the species i. As this selection is purely probabilistic, and performs no specific acceptance or rejection of particles, it can be considered to be a 'pure Bayesian' PID approach. In this case, the PID efficiency does not confer information about the actual detection efficiency, but rather the average weight that is applied to true particle candidates in the Monte Carlo sample. However, in terms of correcting the raw yields, these two quantities serve the same purpose.

In every case, there is a tradeoff between the purity and the efficiency of a measurement. The most intuitive example of this is the fixed-threshold method: increasing the threshold increases the purity requirement for particles to be accepted, and so by necessity the efficiency will decrease accordingly as more particles are rejected. The key ingredient of the validation is, therefore, to ensure that varying this purity through the use of different cuts does not significantly alter the final results of the analysis. This will be discussed further in Section 5.6.

5.5 Motivation: the case of $\Lambda_c \to pK\pi$

Analysing the three-pronged decay of $\Lambda_c^+ \to pK^-\pi^+$ (and respective charge conjugates) is of particular interest in the charm sector, as it would allow measurements of the baryon-to-meson ratio in heavy-ion collisions to be extended into the charm

sector, see Section 2.3. The three-pronged decay $\Lambda_c^+ \to pK^-\pi^+$ proceeds through a combination of multiple resonant channels [23], which are listed in Tab. 5.2.

Λ_c^+ decay mode	$\mathcal{B}$ ($\Gamma_i/\Gamma_{\text{total}}$)	Γ_i/Γ_1
$pK^-\pi^+$ (inclusive)	$(6.84^{+0.32}_{-0.4})\%$	1
$pK^-\pi^+$ (non-resonant)	$(3.8\pm0.4)\%$	0.56 ± 0.06
$\Lambda(1520)\pi^+$	$(2.4\pm0.6)\%$	0.35 ± 0.08
$p\overline{K}^*(892)^0$	$(2.13\pm0.30)\%$	0.31 ± 0.04
$\Delta(1232)^{++}K^-$	$(1.18\pm0.27)\%$	0.17 ± 0.04

Table 5.2: List of Λ_c^+-baryon decay channels with $pK^-\pi^+$ in the final state considered in ALICE. 'Γ_{total}' in the second column refers to the total decay width of the Λ_c^+ baryon, and 'Γ_1' in the third column refers to the inclusive $pK^-\pi^+$ mode listed on the first row. Values from [23].

Due to the short decay length of Λ_c^+ baryons ($c\tau = 59.9\,\mu$m) and the limited spatial resolution of the ITS, a purely topological analysis of Λ_c^+ baryons remains a challenge with the ALICE detector system as of RUN 1. As such, while it has been observed in pp collisions at $\sqrt{s} = 7$ TeV by LHCb [60], the production of Λ_c^+ baryons has not yet been studied in heavy-ion collisions at the LHC. The high combinatorial background makes it difficult to extract a statistically significant raw yield, particularly at low p_{T}, meaning that a robust PID strategy is vital to a successful analysis.

A trial analysis of $\Lambda_c^+ \to pK^-\pi^+$ was performed by the ALICE Collaboration in pp collisions at $\sqrt{s} = 7$ TeV [6], on a sample of $\sim3\times10^8$ minimum-bias events. In order to compare the possible PID strategies, three were applied: a standard n_σ cut, maximum-probability Bayesian PID, and an alternative 'minimum-σ' approach.

The n_σ PID method was based on the TOF response if available, and the TPC alone otherwise. A 3σ compatibility cut was applied on the TOF response for all p_{T} for pions, kaons and protons. Where the TPC response was used, a variable selection between 2 and 3σ was applied for protons and kaons depending on p_{T}, and a 3σ selection was applied for pions for all p_{T}.

The Bayesian PID approach was applied on a single-track basis for each of the decay daughters, under the maximum-probability criterion listed in Section 5.4. For this approach, the TPC and TOF responses were combined, or where the TOF signal was unavailable, the TPC was used alone.

The minimum-σ PID method is based upon choosing the particle species whose expected signal in each detector is the closest to the measured signal, i.e. where the estimated n_σ value for all detectors is the smallest. In this way, it serves as a kind of middle ground between the n_σ and Bayesian approaches: while it takes no input in the form of priors, and is performed in terms of the detector resolution σ, it makes

an exclusive selection of the estimated most likely species for each particle. In this sense, it attempts to choose the best hypothesis for each candidate without explicitly estimating the probabilities.

Figure 5.5 shows the invariant mass spectra of $\Lambda_c^+ \to pK^-\pi^+$ in two p_T intervals: on the left, in a wide p_T interval $(2 < p_T < 6\,\mathrm{GeV}/c)$; on the right, in a narrower p_T interval $(3 < p_T < 4\,\mathrm{GeV}/c)$. The results from n_σ PID, minimum-σ PID and maximum-probability Bayesian PID are plotted on the same axes in each case. For signal extraction purposes, the signal peaks are fitted with Gaussian functions and the backgrounds with exponential functions. In addition, the values of the statistical significance Σ obtained from these fits, defined as $\Sigma = S/\sqrt{S+B}$, where S is the signal and B is the background, are listed in Tab. 5.3.
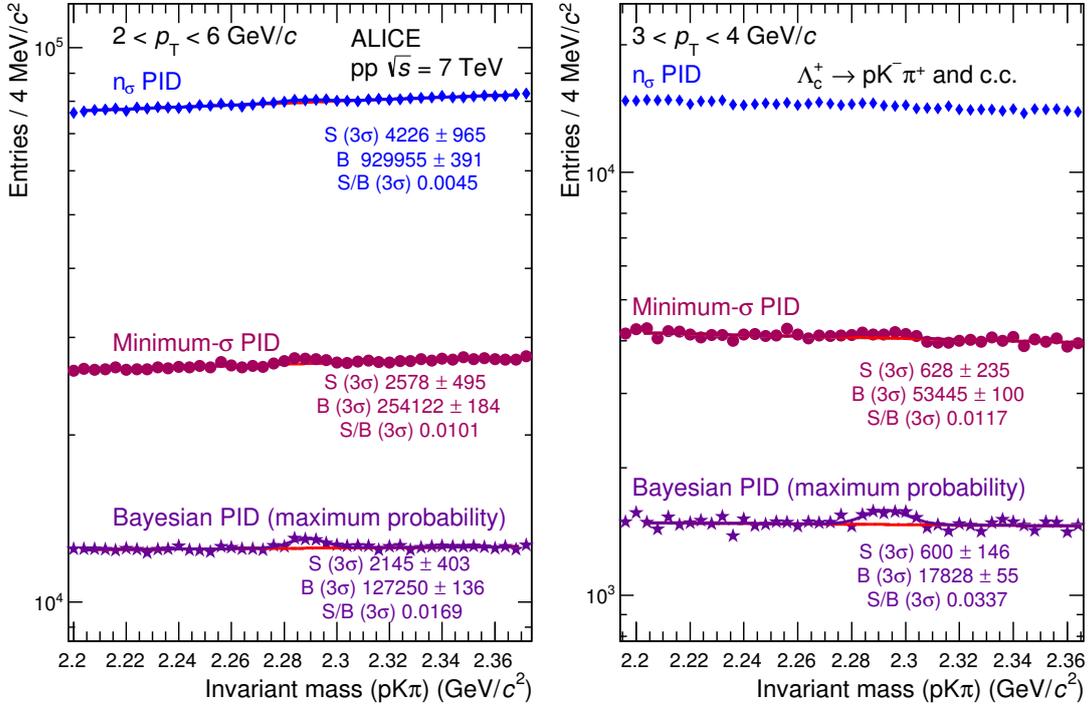
Figure 5.5: Invariant mass spectra of $\Lambda_c^+ \to pK^-\pi^+$ measured by ALICE using n_σ PID, minimum-σ PID and maximum-probability Bayesian PID for (left) $2 < p_T < 6\,\mathrm{GeV}/c$ and (right) $3 < p_T < 4\,\mathrm{GeV}/c$. Due to the low statistical significance, it was not possible to extract a stable signal for n_σ PID for $3 < p_T < 4\,\mathrm{GeV}/c$, therefore this fit and its results are not shown. Figure from [6].

For both p_T intervals, a strong reduction in the background of the n_σ PID method can be seen for minimum-σ PID, and it is decreased further when Bayesian PID is applied. For $2 < p_T < 6\,\mathrm{GeV}/c$, the background is greatly reduced and the signal-to-background ratio more than doubles for minimum-σ PID as opposed to n_σ PID; for maximum-probability Bayesian PID the increase in signal-to-background ratio is a

PID method	$2 < p_\mathrm{T} < 6\,\mathrm{GeV}/c$	$3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$
n_σ	4.4 ± 1.1	–
Minimum-σ	5.1 ± 1.0	2.7 ± 1.0
Bayesian (max. prob.)	6.0 ± 1.1	4.4 ± 1.1

Table 5.3: Statistical significance of Λ_c^+ baryons, for two p_T intervals and three PID methods.

factor of 3.75 compared to n_σ PID. A sequential increase of the statistical significance is also seen between the methods, with the highest statistical significance (and thus smallest statistical uncertainties) seen for Bayesian PID. Comparing the raw signal counts that are shown in Fig. 5.5, it can also be inferred that the minimum-σ and Bayesian methods have a roughly similar efficiency, even though the Bayesian PID method gives a markedly higher signal-to-background ratio and statistical significance.

For $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$, no stable fit could be obtained for the n_σ PID method due to the small signal-to-background ratio and significance, and so this fit and its results are not plotted. For the minimum-σ method, a very small signal can be seen on top of the background; however, the statistical significance (< 3) is such that the statistical uncertainties on this measurement would be prohibitively large. For the Bayesian PID method, the significance in this interval again sees an improvement over minimum-σ PID, with a factor-3 increase in the signal-to-background ratio over the minimum-σ method.

From these results it was concluded that the most promising PID method for a full analysis of $\Lambda_\mathrm{c}^+ \to \mathrm{pK}^-\pi^+$ would be a Bayesian approach, as it would allow a great increase in the precision of the measurement, both in terms of the attainable p_T resolution and the overall statistical uncertainties. However, the method must first be validated for an established analysis, to ensure that it can reliably reproduce previously measured results. This is covered in the remainder of this section.

5.6 Validation of Bayesian PID for $\mathbf{D^0 \to K^-\pi^+}$

In order to ensure that the use of a Bayesian PID method does not cause a significant bias in the final results of analyses, it must be tested against the standard PID approach for an established analysis, and the final corrected yields compared to check for consistency. In [6] this was performed for a variety of two-pronged V^0 decays, which serve as high-purity sources of pions, kaons and protons, as well as for the overall corrected p_T spectra of pions, kaons and protons. This thesis focuses on the final, and most stringent, validation channel presented in [6], namely a complete analysis of D^0-meson production in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$, in the decay channel

$D^0 \to K^- \pi^+$ (and respective charge conjugates).

This section presents the results of the analysis using a variety of Bayesian PID methods, and compares them with analyses both without PID and using the standard n_σ PID method. The aim was not only to judge whether any bias was introduced into the analysis by the Bayesian method, but also to compare the performance of multiple different Bayesian PID strategies in terms of their signal quality. The signal-to-background ratio, statistical significance and PID efficiency of each approach are given, as well as a comparison of the corrected yield from each analysis. Finally, a further study on the dependence of the final result on the choice of priors is illustrated.

This analysis was performed on $\sim 3 \times 10^8$ pp events at $\sqrt{s} = 7\,\text{TeV}$, collected in 2010 during the Run 1 data taking campaign, as described in Section 4.1. As the goal was to reproduce the previous D^0-meson results, the kinematic and topological selections outlined in Section 4.2 were chosen to match those used for the original pp analysis published in [66]. The cuts themselves are given in Tab. 5.4.

p_T (GeV/c)	$p_T^{K,\pi}$ (GeV/c)	$\cos\theta^*$	$d_0^K \times d_0^\pi$ (cm^2)	$\cos\theta_{\text{pointing}}$
1–2	> 0.4	< 0.8	< -0.00025	> 0.8
2–3	> 0.7	< 0.8	< -0.00015	> 0.85
3–4	> 0.7	< 0.8	< -0.00008	> 0.85
4–5	> 0.7	< 0.8	< -0.00008	> 0.85
5–6	> 0.7	< 0.8	< -0.00008	> 0.85
6–8	> 0.7	< 0.8	< -0.00008	> 0.85
8–12	> 0.7	< 0.9	< -0.00005	> 0.85
12–16	> 0.7	< 1.0	$-$	> 0.85

All p_T classes	
$|\Delta M|$ (GeV/c^2)	< 0.4
dca (cm)	< 0.03
$\left|d_0^{K,\pi}\right|$ (cm)	< 0.1
$\cos\theta^*$	< 0.8
$|d_0/\delta d_0|$	> 0.5
decl/δdecl	$> 1.$
decl (cm)	$> \text{Min}(p_T \times 0.0066 + 0.01, 0.06)$

Table 5.4: D^0 selection cuts used for the analysis presented in this section. A dash represents a cut that was not applied in a given p_T interval. The definitions of these cuts are given in Section 4.2.

All of the analyses presented in this section used the same set of topological se-

lections, but differed in the PID strategy used. The n_σ PID method was based on the TPC and TOF responses, with each considered separately. Both TPC and TOF used a 3σ compatibility cut for the decay tracks. Tracks without a valid TOF signal used the TPC information alone for PID; tracks where the two detectors' signals were incompatible with one another were considered to be unidentified, but kept as being compatible with both the pion and kaon hypotheses. The final decision of the D^0 candidate's identity was based on the compatibility of the final state with each of the $K^\mp\pi^\pm$ hypotheses. For cases where the candidate was accepted as potentially belonging to both mass hypotheses, the actual assigned identity depended on the results of the topological selections.

In order to make a comprehensive check of the Bayesian method's performance, the analysis was repeated using each of the three Bayesian strategies outlined in Section 5.4. For the fixed-threshold method, a wide variety of purity thresholds was chosen: 40%, 50%, 70% and 80% were used. The two tracks were required to be oppositely charged. The maximum-probability and fixed-threshold methods were applied on a single-track basis to the decay daughters, and the identified particles were used to reconstruct D mesons as follows:

- – if both tracks were identified as possible kaons, the candidate was accepted as being compatible with both the D^0 and $\overline{D}^0$ hypotheses;

- – if one track was identified as a kaon and the other as a pion, the candidate was accepted as a D^0 if the negative track was a kaon and the positive track was a pion, and vice-versa for $\overline{D}^0$;

- – if neither track was identified as being compatible with the kaon hypothesis, the candidate was rejected;

- – if one or both of the tracks were not compatible with either the kaon or the pion hypothesis, the candidate was rejected.

For the weighted PID method, no explicit assignment was made for the decay tracks; instead, each candidate was accepted as being compatible with both the D^0 and $\overline{D}^0$ hypotheses. Instead, the weights for the candidate to be a D^0 and a $\overline{D}^0$ were computed according to Eq. (5.14), and these were used to fill the invariant mass histogram for its respective p_T interval. A further analysis without any PID selection was also performed, in order to test the systematic uncertainties of the Bayesian approaches against that of n_σ PID itself. All of these analyses were run on the same sample of events in order to ensure the maximum possible statistical correlation.

Yield extraction

The yield was extracted from distributions of the invariant mass of the reconstructed Kπ pair. If a candidate was accepted by the topological and PID selection steps, it was used to fill the invariant mass histogram for the corresponding p_T interval. The raw yield was extracted from the histogram based on a fit to the invariant mass distribution, as described in Section 4.4. The signal was fitted with a Gaussian function, and the background was fitted for the side-bands (regions $> 5\sigma$ away from the signal peak) with an exponential function and extrapolated to the full range. While the mean of the Gaussian function is usually close to the true mean D^0 meson mass $(1.864\,\mathrm{GeV}/c^2$ [23]), the mass is typically left as a free parameter in the fit. In addition, the width of the signal peak, σ, is p_T-dependent and is also usually left as a free parameter. Both of these are determined per p_T interval.

For the analyses presented in this section, the mass and width were left free only for the n_σ PID method, and the extracted values were applied as fixed parameters for the remaining PID methods. This was done in order to minimise the effect of any potential yield-extraction-specific differences between the various methods when comparing them, as well as to ensure that fluctuations in the signal width did not affect the validation. One caveat here is the requirement that the signal samples be statistically correlated, so that fluctuations in the data would not lead to differing natural signal widths for each method. The statistical correlation was tested by making a count of the number of candidates that were accepted by each Bayesian PID method, and comparing this with the number of candidates that were accepted by both Bayesian and n_σ PID. In every p_T interval, and for every tested choice of Bayesian PID method (maximum-probability and all four probability thresholds), over 99% of the candidates selected by Bayesian PID were also accepted by n_σ PID. This was also true irrespective of whether the count was made over the entire invariant mass range or only within 3σ around the expected signal peak. This result implies that the candidates selected with Bayesian PID are a fully correlated subset of those selected by n_σ PID. This justifies the fixing of the yield extraction parameters, and confirms that the statistical uncertainties can be assumed to be fully correlated between the n_σ and Bayesian PID methods. In the following comparisons, the systematic uncertainties due to yield extraction are therefore not taken into account, and only the statistical uncertainties are considered.

Figure 5.6 shows a comparison of the invariant mass distributions resulting from the analyses without PID, using the n_σ PID method, and with maximum-probability Bayesian PID, for three D-meson p_T intervals between 1 and $4\,\mathrm{GeV}/c$. The full list of raw D^0 yields for each p_T interval and PID method is given in the appendix, in Tab. A.1. The level of background is strongly reduced for the n_σ PID case compared with the analysis without PID, with a further reduction when the analysis is performed

with maximum-probability Bayesian PID. The p_T interval $1 < p_T < 2\,\text{GeV}/c$ highlights the need for robust PID in this analysis: without PID, the signal-to-background ratio is too small to extract a stable signal in this interval, limiting the p_T reach of the analysis in a crucial region ($1 < p_T < 2\,\text{GeV}/c$ is estimated to contain some 36% of the D^0 production cross section at $\sqrt{s} = 7\,\text{TeV}$, see Section 2.3). Due to this instability, the analysis without PID is neglected in the following comparisons for this p_T interval.

In each p_T interval, the decrease in background is accompanied by an increase in the signal-to-background ratio; however, the signal itself also decreases due to the reduced efficiency of the PID methods. This is discussed in more detail later in this section.
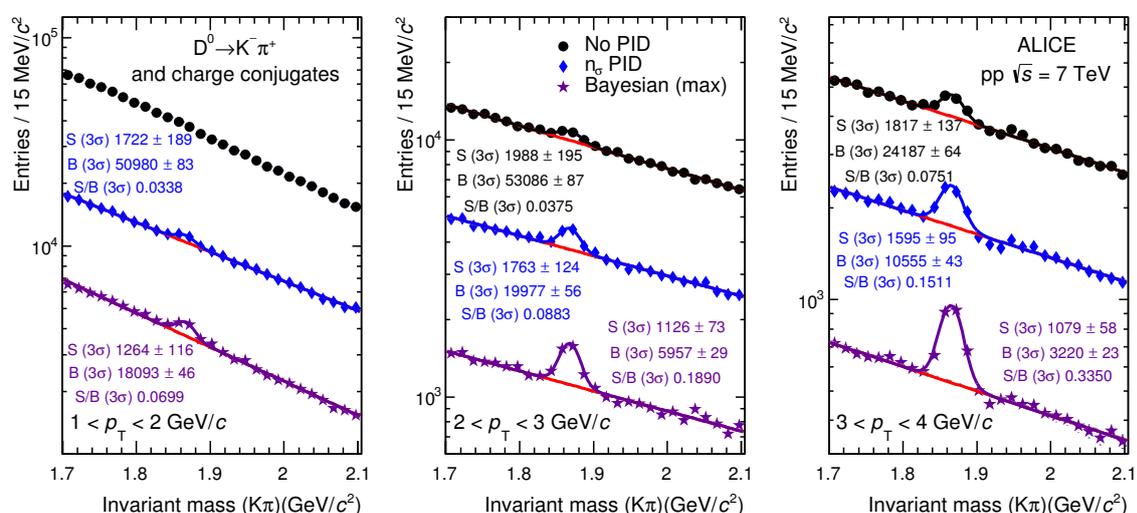
Figure 5.6: A comparison of the invariant mass distributions in three p_T intervals for D^0 candidates in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ obtained without PID, with n_σ PID, and with Bayesian PID using the maximum-probability condition. Due to the low statistical significance, it was not possible to extract a stable signal without PID for $1 < p_T < 2\,\text{GeV}/c$, therefore this fit and its results are not shown. Figure from [6].

Signal-to-background ratio and significance

The signal-to-background ratio and statistical significance Σ can be seen as estimates of the quality of the signal. A high signal-to-background ratio implies a high signal purity, and a high statistical significance is seen when the statistical uncertainties are small. Σ is defined as the inverse of the relative statistical uncertainty,

$$\Sigma = \frac{S}{\sqrt{S+B}}, \tag{5.15}$$

where S and B are the signal and background, respectively. S and B are both estimated within a range of 3σ around the signal peak, where σ is the width of the Gaussian function, as discussed in the previous section.

Figure 5.7 illustrates the signal-to-background ratio and significance for all of the Bayesian PID methods tested. It should be noted that the perceived 'jump' in the statistical significance for all of the PID methods in the p_T intervals $8 < p_T < 12\,\mathrm{GeV}/c$ and $12 < p_T < 16\,\mathrm{GeV}/c$ above the trend seen for $p_T < 8\,\mathrm{GeV}/c$ is not because of some effect inherent to the analysis, but is simply due to integrating the signal over a wider p_T range; this is done in order to improve the precision in a region where the sample size is limited and so the statistical uncertainties would otherwise be prohibitively large. A similar increase in the significance would be seen if e.g. the p_T intervals in the range $4 < p_T < 8\,\mathrm{GeV}/c$ were combined together.

In every p_T interval, it was found that the analysis without PID gave the lowest signal-to-background ratio. Applying n_σ PID gave a significant increase in this ratio for all p_T, and a further increase by a factor of at least two was seen for all of the Bayesian PID methods. In addition, while the performance of each of the Bayesian methods was similar, it was observed that for the fixed-threshold method, the signal-to-background ratio increases with an increasing threshold, confirming the idea that increasing the probability requirement also increases the obtained signal purity.
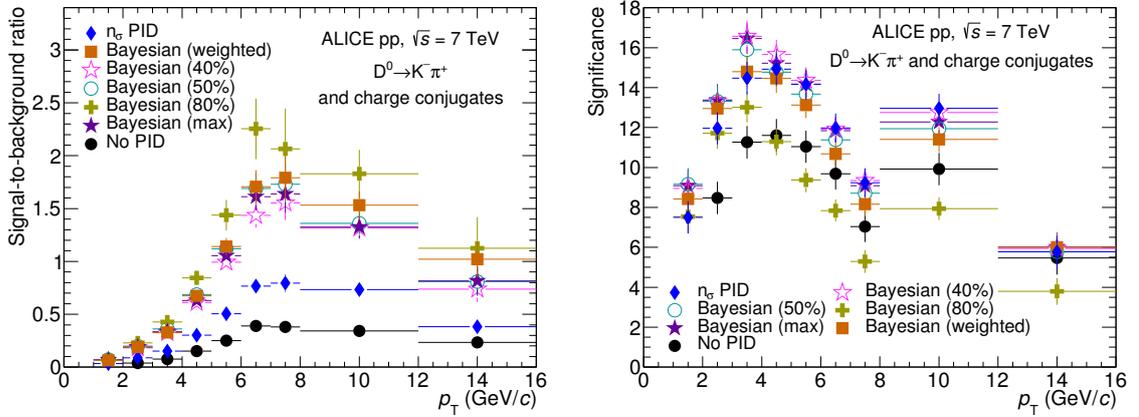
Figure 5.7: (Left) Signal-to-background ratio and (right) statistical significance of D^0 mesons in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ as a function of p_T for various methods of particle identification. The increase in statistical significance at $8 < p_T < 12\,\mathrm{GeV}/c$ is an effect of the width of the p_T interval increasing from 1 to $4\,\mathrm{GeV}/c$. Figures from [6].

The improvement in the statistical significance is less clear-cut. At low p_T ($1 < p_T < 4\,\mathrm{GeV}/c$), the significance is by far the lowest for the analysis without PID. All of the Bayesian methods, with the exception of the 80% threshold, gave a higher significance than the n_σ PID method in this region, with the 40%, 50% and maximum-

probability methods almost indistinguishable from one another. At higher p_T ($p_T > 4\,\mathrm{GeV}/c$), the advantage in terms of the significance for Bayesian PID over n_σ PID begins to wane, with the majority of the methods showing a roughly similar or slightly lower statistical significance compared to n_σ PID. In this region, the significance for the 80% threshold becomes significantly lower than that of the analysis without PID. This was also observed for the 70% probability threshold (not shown in these figures due to space concerns): the statistical significance for this strategy was greater than that seen for the stricter 80% threshold at all p_T, but was lower than n_σ PID overall. The implication of this is that while a strong Bayesian PID cut greatly improves the signal purity, extremely tight PID selections can restrict the statistics of the measurement by excluding more signal candidates, meaning an increase in statistical uncertainties.

In summary, while Bayesian PID confers a benefit in terms of the statistical uncertainty at low p_T, it seems to lose this effect at high p_T. This is likely due to the reduction of the separation power for all detectors at higher p_T. In addition, stricter threshold cuts such as 80% are more likely to introduce larger statistical uncertainties into the measurement due to the lowered statistical significance, indicating that they are undesirable for precision measurements of rare probes for which statistics are limited.

PID efficiencies

The PID efficiency of each method is defined as the proportion of true D^0 mesons that are correctly identified by the PID conditions applied. This was determined for each PID method using Monte Carlo simulations based on the PYTHIA 6.4.21 event generator [164] with the Perugia-0 tune [165], and the GEANT3 package [166] to simulate the detector conditions. While the efficiencies are also determined with regards to the topological selections and detector acceptance, the cuts are identical for all of the analyses presented in this section and only the PID selection differs. The overall reconstruction and selection efficiency for D^0 mesons using the n_σ PID method with all topological selections can be seen in Fig. 4.3.

The efficiencies of the different Bayesian PID methods used, and n_σ PID, are compared in Fig. 5.8. As the analysis without PID does not reject any particles, the efficiency in this case is trivially 100%, and so it is not shown here.

The efficiency of the n_σ PID method was found to be the highest (approximately 97%), which is expected due to it being a wide, inclusive compatibility cut. For Bayesian PID, the efficiency of each fixed-threshold method depends on the choice of parameters: a higher probability threshold leads to a lower efficiency. This behaviour is expected due to the fact that the tighter selections are more exclusive, cutting out more of the true signal in regions where the estimated purity is likely to be lower.
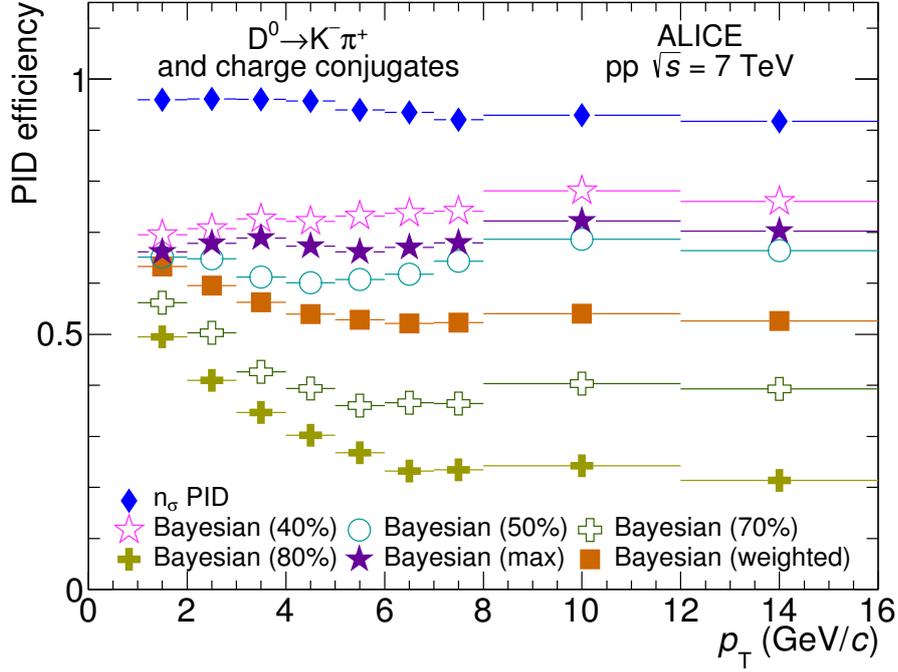
Figure 5.8: A comparison of the PID efficiencies for $D^0 \to K^- \pi^+$ in pp collisions at $\sqrt{s} = 7$ TeV obtained using several PID strategies, as a function of p_T. Figure from [6].

In addition, it can be seen that the efficiency falls steeply towards higher p_T for the stricter Bayesian PID strategies (70% and 80%). Similarly, for the weighted case, the average weight for D^0 mesons was found to be between 50 and 65%, depending on p_T. Due to the reduction of the separation power between kaons and pions for the TPC and TOF at higher p_T (see Fig. 3.12), the probability of a detector signal belonging to a given species falls; this effect is exhibited in the slight decrease of the average weight (the product of the single-track probabilities) when going to higher p_T. However, the efficiency is roughly flat as a function of p_T for the lower thresholds (40% and 50%) and maximum-probability Bayesian PID method, whose efficiency lay between these two thresholds. This effect is not unexpected since, as already mentioned in Section 5.4, the 50% threshold effectively applies a maximum-probability cut with an additional purity requirement.

Comparison of corrected yields

The corrected yield, when all selection efficiencies are taken into account, is an estimate of the true number of D^0 mesons in the sample. Under ideal detector conditions, this would be identical to the raw yield; however, as the detector and selection efficiencies are not equal to unity, these must be corrected for. This is particularly

important when trying to compare the results between different PID methods: due to the vastly different PID efficiencies, as seen in Fig. 5.8, the raw yields are not directly comparable with one another, but the corrected yields should be consistent if the modelling of the detector response in Monte Carlo simulations is correct.

The corrected yields are obtained by dividing the raw yields by their respective efficiencies. If a ratio of the corrected yield to the result using n_σ PID is taken, this effectively gives an estimate of the systematic uncertainty of each PID method with respect to the standard analysis. A similar approach was used in [66] to estimate the systematic uncertainty of the n_σ PID method itself, by comparing it with the analysis without PID. The ratio is defined as

$$R_\kappa = \frac{Y_\kappa^{\mathrm{corr}}}{Y_{\mathrm{n}_\sigma}^{\mathrm{corr}}} = \frac{Y_\kappa^{\mathrm{raw}} \big/ \epsilon_\kappa^{\mathrm{PID}}}{Y_{\mathrm{n}_\sigma}^{\mathrm{raw}} \big/ \epsilon_{\mathrm{n}_\sigma}^{\mathrm{PID}}}, \tag{5.16}$$

where the index κ denotes the PID method under study, and Y_κ^{raw} and Y_κ^{corr} respectively denote the raw and corrected yields for that method. When obtaining a full production cross section, as in Eq. (2.14), the efficiency of all selection criteria and the detector acceptance factor α are used; however, for the comparisons presented here, only the PID strategy is changed while keeping all other conditions the same, and so the acceptance factor and topological selection efficiencies cancel out in the final ratio. This means that only the PID efficiency is corrected for in R_κ. Taken as an aggregate over the full p_{T} range, R_κ gives an idea of the level of bias (if any) in the final result with respect to n_σ PID. The statistical uncertainties on R_κ were estimated under the assumption of full correlation between Bayesian and n_σ PID, as justified earlier in this section. The uncertainties were propagated according to

$$\sigma_{R_\kappa} = \left| \frac{\sigma_{Y_\kappa^{\mathrm{corr}}}}{Y_{\mathrm{n}_\sigma}^{\mathrm{corr}}} - \frac{\sigma_{Y_{\mathrm{n}_\sigma}^{\mathrm{corr}}}}{Y_\kappa^{\mathrm{corr}}} \right|, \tag{5.17}$$

where each instance of σ_x is the statistical uncertainty on the variable x. Due to this, the uncertainties mostly cancel out in the final ratio.

Figure 5.9 shows the corrected yield ratios for Bayesian PID to n_σ PID for (left) the various fixed probability thresholds and (right) the maximum-probability and weighted Bayesian PID methods, and the analysis without PID. Both plots also show the 5% systematic uncertainty inherent to the n_σ PID approach, as quoted for the original ALICE analysis of D^0 mesons in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ [66].

The results were largely consistent with the n_σ PID method, with some larger deviations observed at higher p_{T} due to limited statistics in this region. For the fixed probability thresholds, the average PID systematic with respect to n_σ PID was found to be 9.3%, 8.3%, 6.2% and 4.7% for the 40%, 50%, 70% and 80% thresholds,
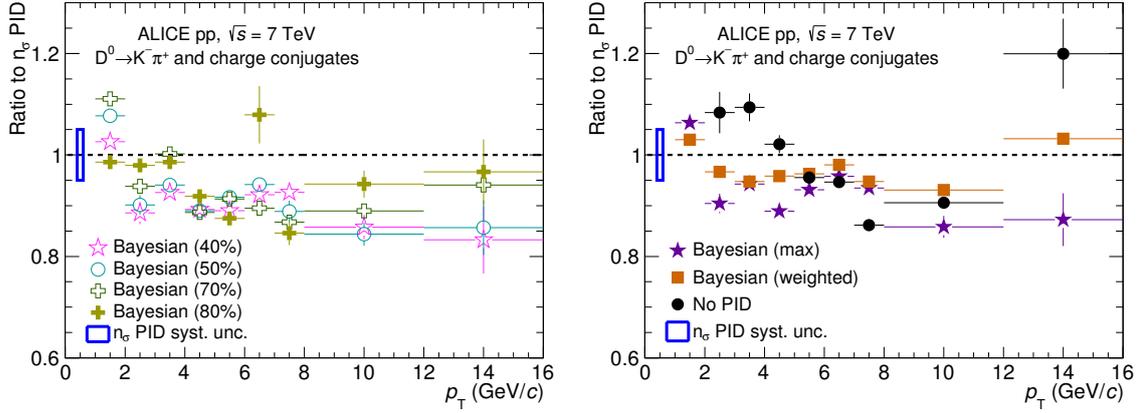
Figure 5.9: Ratios of corrected yields obtained using Bayesian PID to that obtained using n_σ PID, for (left) the 40%, 50%, 70% and 80% Bayesian probability thresholds, and (right) maximum-probability and weighted Bayesian PID, and no PID. The 5% systematic uncertainty on the n_σ PID method [66] is shown as a blue box at $0.5\,\mathrm{GeV}/c$. Figure from [6].

respectively. This ordering implies that a higher-purity cut is more easily corrected for in the Monte Carlo simulations. For maximum-probability Bayesian PID, the average difference is 7.1%; for the weighted Bayesian method this is 2.7%. Finally, the average difference for the analysis without PID was 4.7%, confirming the quoted systematic uncertainty of 5% on the n_σ PID method [66].

On the whole, a small systematic loss of signal was seen compared to the n_σ method. The full list of average and maximum deviations from the n_σ PID method can be found in Tab. A.2, where it is highlighted that the largest deficits tend to occur at higher values of p_T; by contrast, the method performs better on the whole at lower p_T, where statistics are more abundant and the detector responses are better separated. While these deviations point to an imperfect reconstruction of the true detector response, and thus an overestimate of the true PID efficiency, the corrected yields on the whole lead to a similarly sized PID systematic uncertainty that is not significantly larger than that of the n_σ PID method itself. It is therefore concluded that the use of Bayesian PID is valid for this and other similar analyses, and that the stricter probability thresholds and weighted methods appear to give the most consistent response as compared to n_σ PID.

Effect of choice of priors

As the Bayesian PID approach involves taking an additional input in the form of prior probability distributions, which have an effect on the expected particle abundance distributions, it is also important as part of the validation to ensure that the effect

of the choice of the priors is minimal. To this end, further tests were made to ensure the stability of the corrected yield against the choice of priors. In order to test priors that differed from the standard priors but were still relatively close to reality, the modified sets of priors had the relative abundance of kaons scaled up or down by 10% uniformly in all p_T intervals. This deviation is similar to the actual difference observed between the priors and the measured kaon–pion ratios in Fig. 5.3, thus testing whether deviations of the size seen in the prior extraction procedure have a significant effect on the final result. The ratios were computed similarly to those in Eq. (5.16), but using the same PID method in the numerator and denominator:

$$R_\kappa^\mathrm{prior} = \frac{Y_\kappa^\mathrm{corr,alt}}{Y_\kappa^\mathrm{corr,std}}, \tag{5.18}$$

where the superscripts 'alt' and 'std' refer to the analysis with altered priors and with standard priors, respectively. As with the calculation of R_κ, the uncertainties were assumed to be fully correlated, and so were propagated according to Eq. (5.17). Figure 5.10 shows this ratio for (left) the analysis with fixed probability thresholds, and (right) for maximum-probability and weighted Bayesian PID. The list of average and maximum deviations from the results using standard priors is given in Tab. A.3.
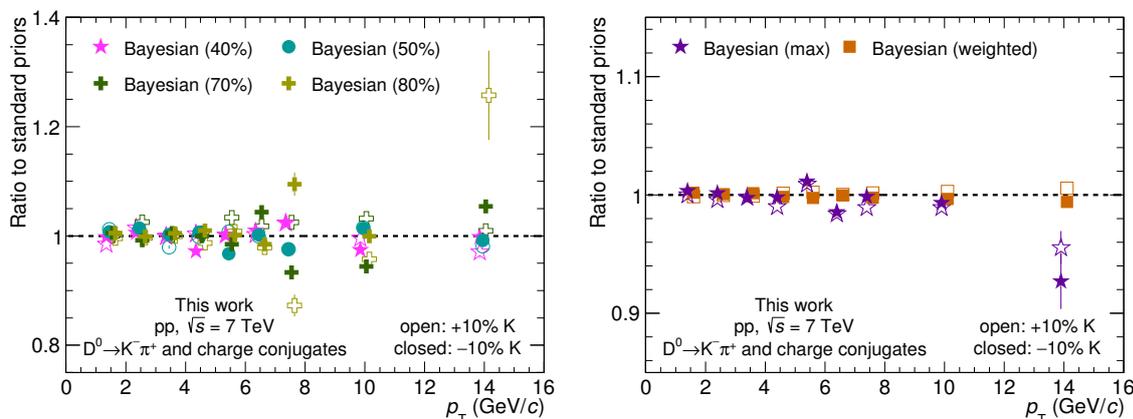
Figure 5.10: Ratios of corrected yields using modified priors to those obtained using the standard priors for (left) fixed-threshold Bayesian PID and (right) maximum-probability and weighted Bayesian PID. The open points show the priors modified with 10% extra kaons; the filled points, with 10% fewer kaons. The points corresponding to each method are shifted slightly in each p_T interval for visibility.

With the exception of the 80% probability threshold, the difference introduced into the corrected yield of each method by changing the choice of priors was found to be much smaller than the aforementioned PID systematics overall, and for all cases apart from the 70% and 80% thresholds with added kaons, a difference of less than 1% was seen on average. However, it was also seen that the deviation was

affected by the amount of available statistics for the most exclusive cuts. The 70% and 80% thresholds, as well as the maximum-probability method, showed significant deviations from unity in either the $7 < p_{\mathrm{T}} < 8\,\mathrm{GeV}/c$ interval (the highest p_{T} interval with a width of $1\,\mathrm{GeV}/c$) or the $12 < p_{\mathrm{T}} < 16\,\mathrm{GeV}/c$ interval (where the number of events was smallest). The weighted Bayesian method, and the more inclusive 40% and 50% thresholds, each remained stable without significant deviations over the full p_{T} range.

While at first glance the large differences in the low-statistics regions may appear to be due to statistical significances, this is not necessarily the case: for $12 < p_{\mathrm{T}} < 16\,\mathrm{GeV}/c$, the deviations of the maximum-probability method are at their largest, despite the significance being practically identical to those of the 40% and 50% thresholds. Nor is it necessarily due to the exclusivity of the cut, as the efficiency of the maximum-probability method is among the highest in this p_{T} region. One possible explanation is that the finite Monte Carlo statistics mean that the efficiency is slightly over- (under)estimated for more (less) inclusive cuts when the priors are altered. However, as the differences introduced were on the whole rather small, it can be concluded that the effect of the choice of priors is insignificant for the majority of Bayesian PID strategies. In addition, in [6], further checks were made on the analysis of single-particle spectra using flat priors for all species; these were also found to not introduce significant changes in the final result.

5.7 Conclusions & outlook

It was verified that the Bayesian PID approach presents a valid choice for the analysis of $\mathrm{D}^0 \to \mathrm{K}^- \pi^+$, paving the way for its use in other analyses in ALICE. Furthermore, the relative performances of various Bayesian PID strategies were tested. It was found that strongly exclusive PID cuts, such as an 80% threshold on the Bayesian probabilities of individual tracks, greatly increase the purity of the result, exhibited by higher signal-to-background ratios; however, this is counteracted by a reduction in statistical significance (and therefore an increase in the statistical uncertainties). More inclusive cases, on the other hand, achieve both a greater purity than the standard n_σ PID method and a greater (similar) statistical significance compared to the n_σ PID method at low (high) p_{T}. The specific choice of PID method for future analyses under the Bayesian PID framework is strongly contingent on the requirements of each analysis. Cross section and R_{AA} measurements, where statistics are at a premium, require the smallest possible statistical uncertainties, and so one of the less stringent methods would be most suitable; conversely, for e.g. identified two-particle correlations, where higher purities are desired, a stricter cut may be applied. In general, the tests of the Bayesian PID method have shown that, despite a small overestimate of

the PID efficiency in Monte Carlo, it is able to satisfactorily correct for a wide range of cut requirements.

The weighted Bayesian method presents a particularly interesting case: its basis is entirely probabilistic, and although it makes no explicit selection of candidates, it gives among the highest purity in the final yield, as well as the greatest stability against the choice of priors and in terms of the comparison of the corrected yield with that from the n_σ PID method.

The Bayesian PID probabilities are currently being used as one of the inputs for a Boosted Decision Tree (BDT) analysis of $\Lambda_c^+ \rightarrow pK^-\pi^+$ using the Toolkit for Multivariate Analysis (TMVA) [171], which is being performed in parallel with analyses of the two-pronged decay channel $\Lambda_c^+ \rightarrow pK_S^0$ in pp and p–Pb collisions in ALICE. In addition, now that the Bayesian PID framework has been established and validated in the ALICE TPC and TOF, it can be applied to analyses of other particle species using the wide variety of other PID detectors in the ALICE central barrel.

6 │ D^{*+}-meson production as a function of multiplicity

As mentioned in Section 2.6, the study of particle production as a function of charged-particle multiplicity in small systems allows the role of multi-parton interactions, in addition to possible collective effects, to be better examined. In particular, heavy-flavour production as a function of multiplicity allows the correlations between soft and hard particle production processes to be more closely examined. This section deals with the production of D^{*+} mesons, and their reconstruction in the channel D$^{*+} \to$ D$^0 \pi^+$, as a function of multiplicity in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. The multiplicity is estimated in two intervals: mid-rapidity ($|\eta| < 1.0$), and backward rapidity (the Pb-going direction, $-2.8 < \eta < -5.1$), and the determination and correction of the multiplicity in both regions is outlined in Section 6.1. Section 6.2 describes the reconstruction of D^{*+} mesons in ALICE, including the selection efficiency and the determination of the systematic uncertainties on the yield extraction procedure. Section 6.3 then gives the results of the analysis, as well as comparisons with D^0 and D$^+$ mesons in p–Pb collisions. The results are also compared with equivalent measurements in pp collisions at $\sqrt{s} = 7$ TeV [16], and with the results given by model calculations, using both multiplicity estimators.

The work presented in this section was also published as part of [17].

6.1 Multiplicity determination and correction

The charged-particle multiplicity in p–Pb collisions is estimated in two rapidity regions: mid-rapidity ($|\eta| < 1.0$) and backward rapidity ($-5.1 < \eta < -2.8$, in the Pb-going direction). In both cases, η refers to η_{lab}, the pseudorapidity in the lab system with respect to the nominal origin of the ALICE coordinate system ($z = 0$, $r = 0$). The mid-rapidity multiplicity is measured using the number of tracklets, $N_{\mathrm{tracklets}}$, reconstructed in an event by the SPD. A 'tracklet' in this context is defined as the line joining a pair of hits in each of the layers of the SPD and compatible with the primary vertex of the collision. The backward-rapidity estimator makes use of the V0A hodoscope, whose signal is proportional to the charged-particle multiplicity in this rapidity region. For pp and Pb–Pb collisions, multiplicity and centrality esti-

mators usually involve both of the V0 counters; however, the asymmetry of p–Pb collisions justifies the use of only the Pb-going direction. This section deals with the methods that were used to classify and correct the charged-particle multiplicity in both rapidity regions in this work.

Multiplicity corrections

As the geometry and evolution of the collision system do not depend on its location within the detector, the distribution of multiplicity should ideally be independent of the position of the primary vertex. However, due to detector effects, the actual multiplicity that is measured does not follow this expectation. In addition, as the data sample was taken over the course of multiple runs, LHC13b and LHC13c (see Section 4.1), time dependences may appear due to parts of the detector being inactive at certain times. The distributions must therefore be corrected for, in order to remove any possible time or position dependences that may affect the true multiplicity distribution.

The vertex position is most commonly referred to in terms of the distance from the origin along the beam axis, z_{vtx}. In the ALICE coordinate system, the Pb nucleus travels in the positive z direction in the analysed data sample, and the proton travels in the negative z direction. The z_{vtx} distribution of the raw measured $N_{\mathrm{tracklets}}^{\mathrm{raw}}$ ($N_{\mathrm{V0A}}^{\mathrm{raw}}$) in the LHC13b p–Pb data taking period is shown in the left- (right-)hand panel of Fig. 6.1, as a heatmap overlaid with the mean at each z_{vtx} point (black line).

For the multiplicity at mid-rapidity, the mean rises with increasing z_{vtx} for $|\eta| < 6\,\mathrm{cm}$, and falls off steeply for $|\eta| > 6\,\mathrm{cm}$. This non-uniformity comes about as a result of the limited geometrical detector acceptance, as tracklets are less likely to fall within the acceptance of the detector if they are produced closer to the edges. In addition, fluctuations can be seen in the distribution due to a varying number of dead channels in the SPD as a function of z. For the backward-rapidity multiplicity measurement, the geometrical acceptance plays a large role, as the solid angle between the interaction vertex and the detector coverage becomes smaller with increasing distance from the detector. This manifests as a gradual linear decrease of the N_{V0A} value with increasing z_{vtx}.

The profiles of the multiplicity distributions are shown for both of the analysed datasets and both multiplicity estimators in Fig. 6.2, in order to check the time dependence over both data taking periods (LHC13b and LHC13c). In the left-hand panel, the difference between the two data taking periods can be seen, as the average number of tracklets reconstructed by the SPD is slightly smaller in the later period (LHC13c) than in the earlier one (LHC13b), particularly for negative values of z_{vtx}, due to an increased number of dead SPD channels in the later data taking period. By comparison, there is no dependence of the V0A response on the data taking period
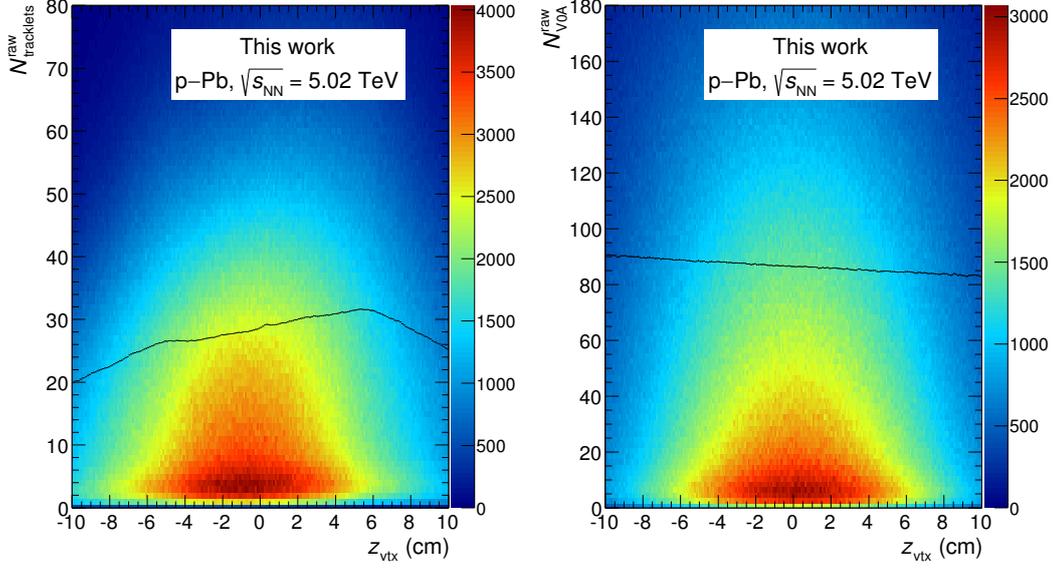
Figure 6.1: Raw multiplicity distributions as a function of z_{vtx} for the LHC13b p–Pb data taking period. Left: $N_{tracklets}$ measured within $|\eta| < 1$. Right: N_{V0A} measured within $-5.1 < \eta < -2.8$. The frequencies of each multiplicity value within the data sample are represented by a heatmap. The black lines represents the overall mean $N_{tracklets}$ and N_{V0A} as a function of z_{vtx}.

in question.
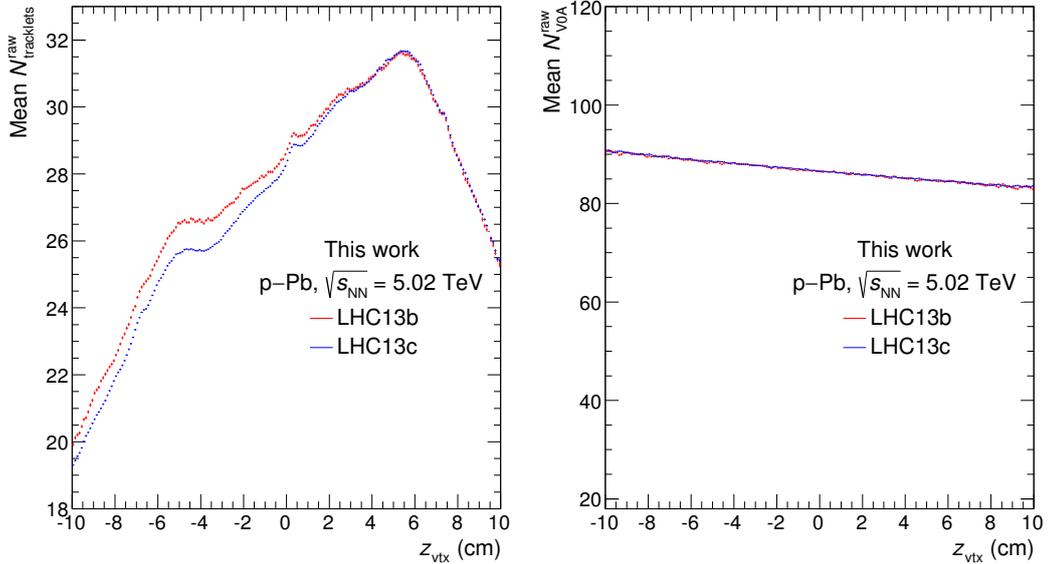
Figure 6.2: Raw multiplicity profiles as a function of z_{vtx} for both p–Pb data taking periods. Left: $N_{tracklets}$ measured within $|\eta| < 1$. Right: N_{V0A} measured within $-5.1 < \eta < -2.8$.

In order to correct for these effects, it is necessary to normalise the multiplicity distributions as a function of z_{vtx}. Although the shape of the aberrations differs between the two rapidity regions, the method used to correct the distributions is the same for both. Analogous methods were also used for the z_{vtx} correction in the multiplicity-dependent analysis of D mesons in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ in ALICE [16].

The profile distributions shown in Fig. 6.2 for each of the data taking periods formed the basis of the correction of the multiplicity distributions. Each profile was used to correct events in its corresponding data taking period. A Poissonian correction method was applied on an event-by-event basis to normalise the distributions to some reference value, chosen as the overall minimum value of the mean multiplicity as a function of z_{vtx} for each estimator in the analysed sample ($N_{\text{tracklets}} = 19.44$, $N_{\text{V0A}} = 82.7$). These values effectively become the mean of each distribution; however, as the final multiplicity intervals are all normalised to the mean, the specific choice of reference value is somewhat arbitrary for the final result. The correction was performed according to

$$
N_{\text{tracklets}}^{\text{corr}} =
\begin{cases}
N_{\text{tracklets}}^{\text{raw}} - \mathcal{P}(N_{\text{tracklets}}^{\text{prof}} - N_{\text{tracklets}}^{\text{ref}}), & N_{\text{tracklets}}^{\text{prof}} > N_{\text{tracklets}}^{\text{ref}}, \\
N_{\text{tracklets}}^{\text{raw}} + \mathcal{P}(N_{\text{tracklets}}^{\text{ref}} - N_{\text{tracklets}}^{\text{prof}}), & N_{\text{tracklets}}^{\text{prof}} < N_{\text{tracklets}}^{\text{ref}},
\end{cases}
\tag{6.1}
$$

where $\mathcal{P}(\mu)$ is a random variable following a Poisson distribution with mean μ, $N_{\text{tracklets}}^{\text{prof}}$ is the mean value of the multiplicity at the same z_{vtx} position, from the multiplicity profile for that data taking period, and $N_{\text{tracklets}}^{\text{ref}}$ is the reference multiplicity. The formula used for N_{V0A} is the same as in Eq. (6.1), but with $N_{\text{tracklets}}$ replaced by N_{V0A} in every case. The Poisson distribution was used here in order to ensure an integer number of counts for each multiplicity estimator.

The corrected distributions are shown in Fig. 6.3 for the two-dimensional distributions of $N_{\text{tracklets}}^{\text{corr}}$ and $N_{\text{V0A}}^{\text{corr}}$ against z_{vtx}, and in Fig. 6.4 for the profiles per data-taking period. The profile of the $\langle N_{\text{tracklets}} \rangle$ distribution retains some of its z_{vtx} shape; however, the differences between the maximum and minimum of the profile are of the order of 1%, much smaller than in the uncorrected case ($\sim 50\%$). Both of the data-taking periods are found to be consistent with one another. For the distribution of $\langle N_{\text{V0A}} \rangle$, the profiles of both data-taking periods are found to be equal and flat as a function of z_{vtx}, with only some fluctuations seen in the uncertainties of the two periods.

Conversion between $N_{\text{tracklets}}$ and N_{ch}

The $N_{\text{tracklets}}$ estimator simply counts the number of tracks seen by the detector, and thus is subject to the detector efficiency and other experimental effects. The
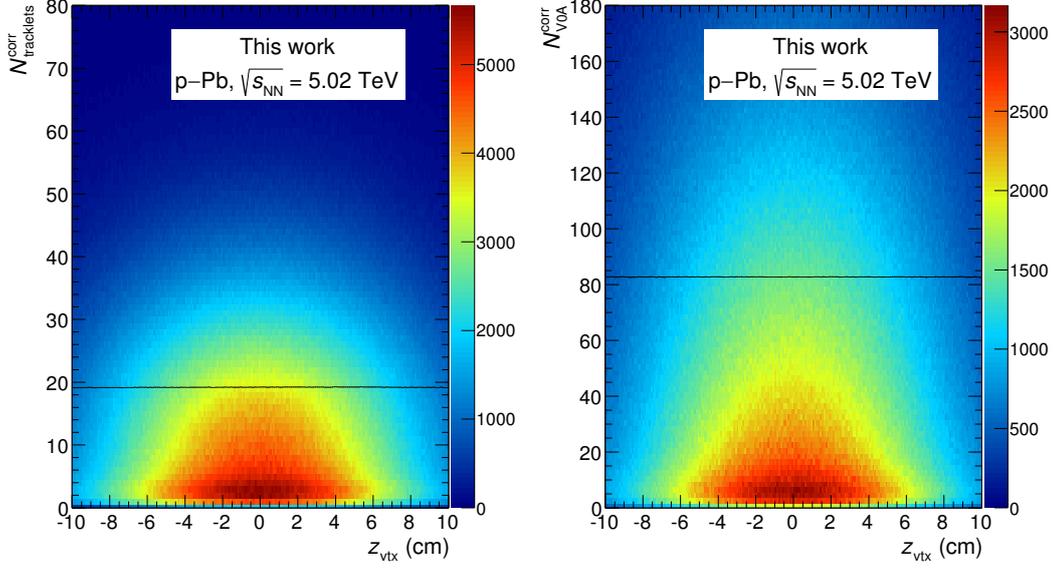
Figure 6.3: Corrected multiplicity distributions as a function of z_{vtx} for the LHC13b p–Pb data taking period. Left: $N_{\text{tracklets}}$ measured within $|\eta| < 1$. Right: N_{V0A} measured within $-5.1 < \eta < -2.8$. The frequencies of each multiplicity value within the data sample are represented by a heatmap. The black lines represent the overall mean $N_{\text{tracklets}}$ and N_{V0A} as a function of z_{vtx}.
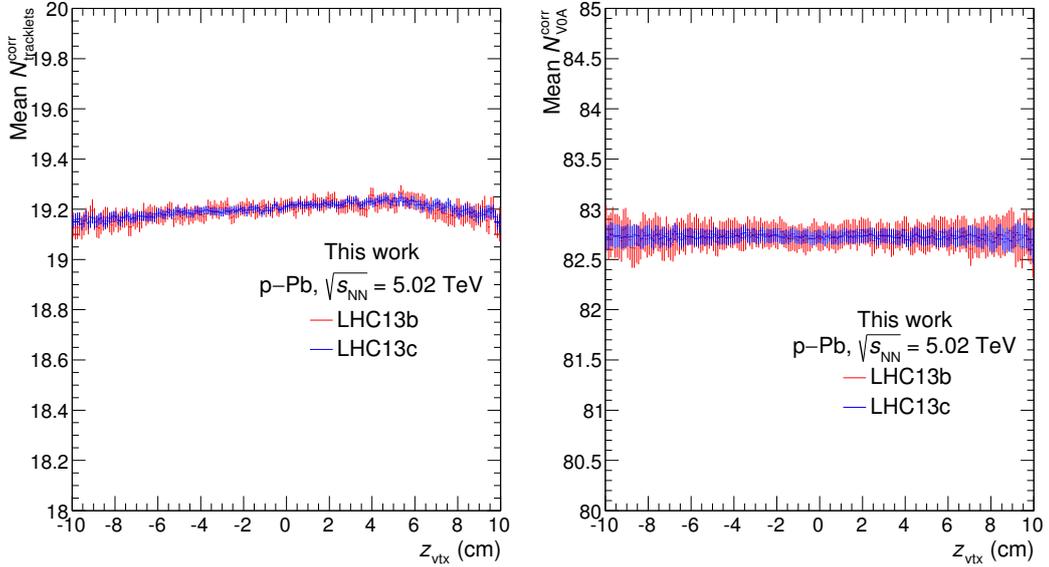
Figure 6.4: Zoomed-in view of the corrected multiplicity profiles as a function of z_{vtx} for both p–Pb data taking periods. Left: $N_{\text{tracklets}}$ measured within $|\eta| < 1$. Right: N_{V0A} measured within $-5.1 < \eta < -2.8$.

actual physical quantity of interest is N_{ch}, the number of primary charged particles produced in the interaction. For the purposes of this analysis, the charged-particle multiplicity density measured by ALICE at mid-rapidity, $\langle \mathrm{d}N_{\text{ch}}/\mathrm{d}\eta \rangle_{|\eta|<1.0} = 17.64 \pm 0.01(\text{stat.}) \pm 0.68(\text{syst.})$, measured by ALICE for inelastic p–Pb collision events at $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$ with at least one charged particle within $|\eta| < 1.0$ [172], is used as a reference value. The requirement of at least one charged particle is an LHC convention that allows results from different experiments to be directly compared with minimal model dependence. This necessitates the conversion of the measured detector response $N_{\text{tracklets}}$ to a physical N_{ch} value, in order for the two quantities to be comparable. This conversion was performed using minimum-bias Monte Carlo simulations. A comparison was made between the number of generated physical primary particles that were generated in the simulation, and the corrected number of simulated tracklets reconstructed in the SPD. The correlation between these two quantities is shown in Fig. 6.5. The coloured region is a heat map of points, with the horizontal axis corresponding to the number of reconstructed tracks in the simulation, and the vertical axis corresponding to the total number of charged particles generated in the simulation. The black points are the profile, i.e. the mean N_{ch} value seen for each $N_{\text{tracklets}}$ point, and the red line is a linear fit to the distribution.
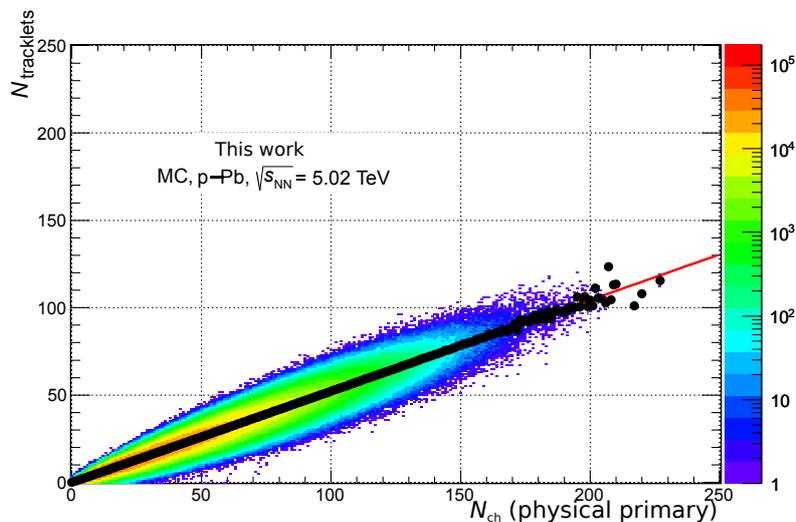
Figure 6.5: Comparison between $N_{\text{tracklets}}$ and N_{ch} in minimum-bias Monte Carlo simulations. The N_{ch} value corresponds to all physical primary charged particles generated in the simulation. The profile (mean at every $N_{\text{tracklets}}$ value) is shown as black points; a linear fit to the distribution is shown as a red line.

The linear fit to the distribution yielded an estimate of $0.52 N_{\text{tracklets}}/N_{\text{ch}}$, with an approximately Gaussian spread of points about the central line. This was cross-checked by making a Gaussian fit to the distribution at $N_{\text{ch}}^{\text{MC}} = 2\langle \mathrm{d}N_{\text{ch}}/\mathrm{d}\eta \rangle_{|\eta|<1.0}^{\text{data}} =$

35.28 [172]. The mean of this Gaussian was found to be 18.27, which was consistent with the proportionality factor of the overall distribution ($18.27/35.28 = 0.52$). Under the assumption of a roughly flat distribution of N_{ch} as a function of η at mid-rapidity, the measured $N_{tracklets}$ values were therefore converted to $dN_{ch}/d\eta$ by dividing by 0.52 (to convert to N_{ch}) and then by a further factor of two (the width of the pseudorapidity range, $\Delta\eta = 2$. Linear fits were also made within the individual multiplicity classes in order to judge the local deviations of the scaling factor from the global value. Further checks were made of the proportionality for the charm-enriched Monte Carlo data set used for the charm-enriched Monte Carlo sample used to correct for the reconstruction efficiencies; the scaling of N_{ch} against $N_{tracklets}$ was found to be consistent. The spread of the Gaussian and deviations for the linear fits for individual multiplicity classes, in combination with the experimental uncertainties on the previously measured $\langle dN_{ch}/d\eta\rangle$, led to a 5% global systematic uncertainty on the measured $dN_{ch}/d\eta$ values, and a 6.3% systematic uncertainty on the final relative $dN_{ch}/d\eta/\langle dN_{ch}/d\eta\rangle$ values globally.

For the analysis using the multiplicity at backward rapidity, no explicit conversion to N_{ch} was made. Instead, the average corrected V0A signal for the analysed event sample, $\langle N_{V0A}\rangle = 82.7$, was used as a reference value. The uncertainty on the relative backward multiplicity measurements, based on the variance of the multiplicity about the mean ratio within each class, was 5% globally.

Multiplicity intervals for analysis

The corrected multiplicity ranges and their averages are shown in Tab. 6.1 for the $dN_{ch}/d\eta$ analysis, and in Tab. 6.2 for the N_{V0A} analysis. The number of events analysed in each class is also given for both estimators. The intervals are expressed in terms of the aforementioned mean charged-particle multiplicity at mid-rapidity, $\langle dN_{ch}/d\eta\rangle_{|\eta|<1.0} = 17.64 \pm 0.01$ (stat.) ± 0.68 (syst.). As can be seen, the highest multiplicity interval for the mid-rapidity multiplicity estimator probes extreme values of N_{ch}, reaching a subsample of $< 1\%$ of the total number of events. The highest-multiplicity events with a valid D*+-meson candidate within the range of the invariant mass spectrum had $N_{tracklets} = 131$ ($dN_{ch}/d\eta = 126$) and $N_{V0A} = 603$ for each estimator, i.e. seven times the mean multiplicity in each case.

6.2 D*+-meson reconstruction

The reconstruction of D*+ mesons for this analysis followed the procedure outlined in Section 4. D*+ mesons were reconstructed at mid-rapidity ($-0.96 < y_{cms} < 0.04$) in five p_T intervals in the range $1 < p_T < 24\,\mathrm{GeV}/c$, for six mid-rapidity multiplicity

$N_{\mathrm{tracklets}}$ range	$\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle_i$	$\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle_i \big/ \langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$	$N_{\mathrm{events}}/10^6$
$[1, 21]$	9.8	0.56	58.6
$[22, 28]$	24.0	1.36	12.8
$[29, 34]$	30.3	1.72	8.0
$[35, 43]$	37.3	2.11	7.6
$[44, 69]$	50.3	2.85	6.4
$[70, 199]$	75.3	4.27	0.5
Integrated	17.6	–	96.6

Table 6.1: Mid-rapidity multiplicity intervals used for the D^{*+}-meson analysis in Section 6. The d$N_{\mathrm{ch}}/\mathrm{d}\eta$ column represents the average multiplicity in each interval (5.0% global uncertainty not quoted). The third column is the average in each multiplicity interval divided by the global average ($\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle_{|\eta|<1.0} = 17.64$, 6.3% global uncertainty not quoted). The final column is the number of analysed events in each multiplicity interval, which is used for the normalisation of the final result.

N_{V0A} range	$\langle N_{\mathrm{V0A}} \rangle_i$	$\langle N_{\mathrm{V0A}} \rangle_i \big/ \langle N_{\mathrm{V0A}} \rangle$	$N_{\mathrm{events}}/10^6$
$[0, 90]$	40.0	0.48	36.9
$[91, 132]$	109.5	1.32	12.3
$[133, 172]$	150.0	1.81	10.0
$[173, 226]$	195.5	2.36	20.5
$[227, 798]$	272.5	3.29	16.8
Integrated	82.7	–	96.4

Table 6.2: Backward-rapidity multiplicity intervals used for the D^{*+}-meson analysis in Section 6. The N_{V0A} column represents the average multiplicity in each interval (5.0% global uncertainty not quoted). The second column is the average in each multiplicity interval divided by the global average ($\langle N_{\mathrm{V0A}} \rangle = 82.7$), 6.3% global uncertainty not quoted). The final column is the number of analysed events in each multiplicity interval, which is used for the normalisation of the final result.

classes and five backward-rapidity multiplicity classes. The topological selections used to reconstruct D^{*+} mesons were mostly the same as used for the R_{pPb} analysis published by ALICE in [123], apart from the distance of closest approach (DCA) cut at low p_{T} ($1 < p_{\mathrm{T}} < 2\,\mathrm{GeV}/c$), which was loosened in order to improve the statistical significance in the small subsamples of events of the individual multiplicity classes. All of the selection criteria were kept constant as a function of multiplicity in all p_{T} intervals. The full list of reconstruction and selection cuts is given in Tab. 6.3; the definitions of the variables are given in Section 4.2. As the Bayesian PID method

described in Section 5 was not yet established in ALICE at the time of this analysis, an n_σ PID method was used instead. As mentioned in Section 4.3, the PID for D*⁺ mesons relied on 2- and 3σ compatibility selections for the decay daughters of the D^0 meson in the TPC and TOF, respectively; no PID selection was imposed on the soft pion candidate.

p_T (GeV/c)	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–12	12–16	16–24
$\left\|\Delta M_{D^0}\right\|$ (MeV/c^2) <	32	32	32	32	36	36	36	50	94	94
dca (cm) <	0.025	0.03	0.03	0.04	0.05	0.05	0.06	0.08	0.1	0.1
$\cos\theta^*$ <	0.8	0.85	0.9	0.9	1	1	1	1	1	1
p_T^K (GeV/c) >	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.4	0.4	0.3
p_T^π (GeV/c) >	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.4	0.4	03
$\|d_0^K\|$ (cm) <	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2
$\|d_0^\pi\|$ (cm) <	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2
$d_{0,K} \times d_{0,\pi}$ (cm^2) <	−0.3	−0.25	−0.13	−0.04	0.015	0.05	0.06	0.1	0.5	0.5
$\cos(\theta_{point})$ >	0.93	0.9	0.88	0.88	0.85	0.7	0.7	0.6	0.6	–
$\cos(\theta_{point,xy})$ >	0.97	–	–	–	–	–	–	–	–	–
L_{xy} (cm) >	3.5	3.5	3.5	3.5	2.5	1.5	1.	0	0	0

Table 6.3: Summary table of the reconstruction and selection cuts used for D*⁺ mesons in p–Pb collisions. The cut titles are defined in Section 4.2.

The reconstruction and selection efficiencies are shown as a function of midrapidity multiplicity for each studied p_T interval in Fig. 6.6. The left-hand panel shows the efficiency for prompt D*⁺ mesons; the right-hand panel, the efficiency for D*⁺ mesons originating from the decays of B mesons ('feed-down'). These were determined using heavy-flavour-enriched Monte Carlo simulations based on the PYTHIA 6.4.21 event generator [164] with the Perugia-0 tune [165], with particles transported through the detector with GEANT3 [166]. The underlying event for each PYTHIA-generated c$\bar{c}$ or b$\bar{b}$ pair was generated using the HIJING v1.36 [173] event generator, which simulates multiple jet production and nuclear shadowing in nuclear collisions. The efficiency rises as a function of p_T due to the selection cuts loosening at high p_T, as mentioned above; however, the efficiencies are roughly constant as a function of multiplicity due to the consistency of the detector response and cuts against the multiplicity. The efficiency for feed-down D*⁺ mesons is found to be larger than that for prompt D*⁺ mesons, due to the decay length cut being more efficient for the longer-lived B meson than for D mesons. Finally, the multiplicity shapes of both the prompt and feed-down efficiencies are largely consistent, meaning that any multiplicity dependence of D*⁺-meson production cannot be attributed simply to an alteration in the feed-down contribution.
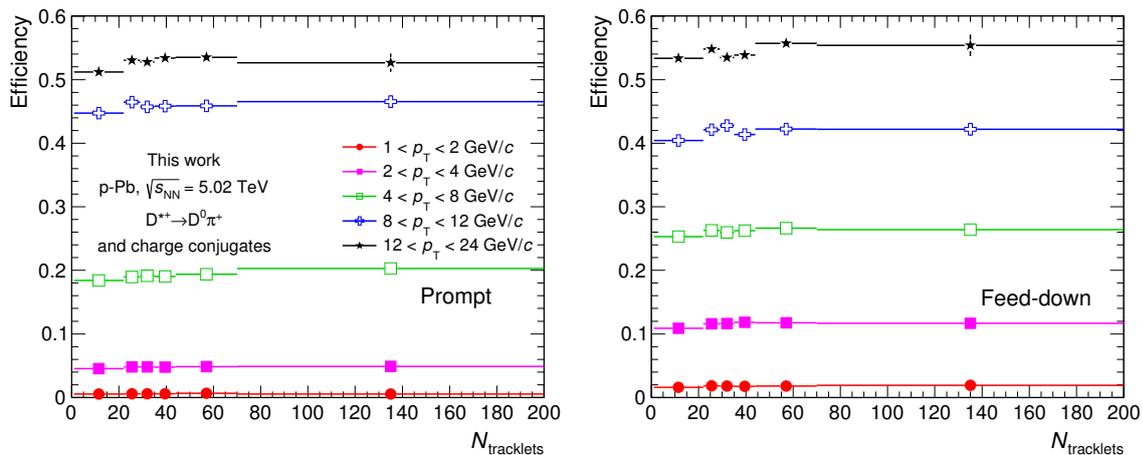
Figure 6.6: Reconstruction and selection efficiency of D^{*+} mesons in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, as a function of mid-rapidity multiplicity. Left: prompt D^{*+} mesons; right: D^{*+} mesons from the decays of B mesons.

Yield extraction

As described in Section 4.4, the extraction of the D^{*+}-meson yield is performed using a fit to the invariant mass distribution. In this case, the mass difference between the full reconstructed $K^-\pi^+\pi^+$ triplet and the reconstructed $K^-\pi^+$ doublet of the D^0 decay daughter, $\Delta M = M_{K\pi\pi} - M_{K\pi}$, was used for the invariant mass analysis in order to improve the signal resolution (see Section 4.4). The signal is modelled using a Gaussian function, and the background is estimated by a power law multiplied by an exponential, as given in Eq. (4.1). The width per mass bin in the invariant mass histograms was varied as a function of p_T between $400\,\text{keV}/c^2$ and $1\,\text{MeV}/c^2$, to reduce statistical fluctuations in regions with few counts. The overall signal width was found to be consistent with those extracted from the Monte Carlo simulations that were used to determine the efficiencies.

In order to reduce systematic effects based on measuring the signal in different multiplicity classes, the Gaussian fit parameters for the signal (the mean and standard deviation) were determined for the multiplicity-integrated case in each p_T interval, and then fixed to these values for each multiplicity interval. The use of consistent parameters in each interval is justified by the detector response (typified by the efficiencies shown in Fig. 6.6) being largely independent of the multiplicity, and by the individual multiplicity classes being statistically correlated subsamples of the multiplicity-integrated dataset. In the majority of p_T and multiplicity classes, a stable signal was obtained using this method.

The invariant mass spectra are shown with their fits in Fig. 6.7 for (top) all five p_T intervals for all multiplicity, and (bottom) three N_{V0A} intervals (low, intermediate
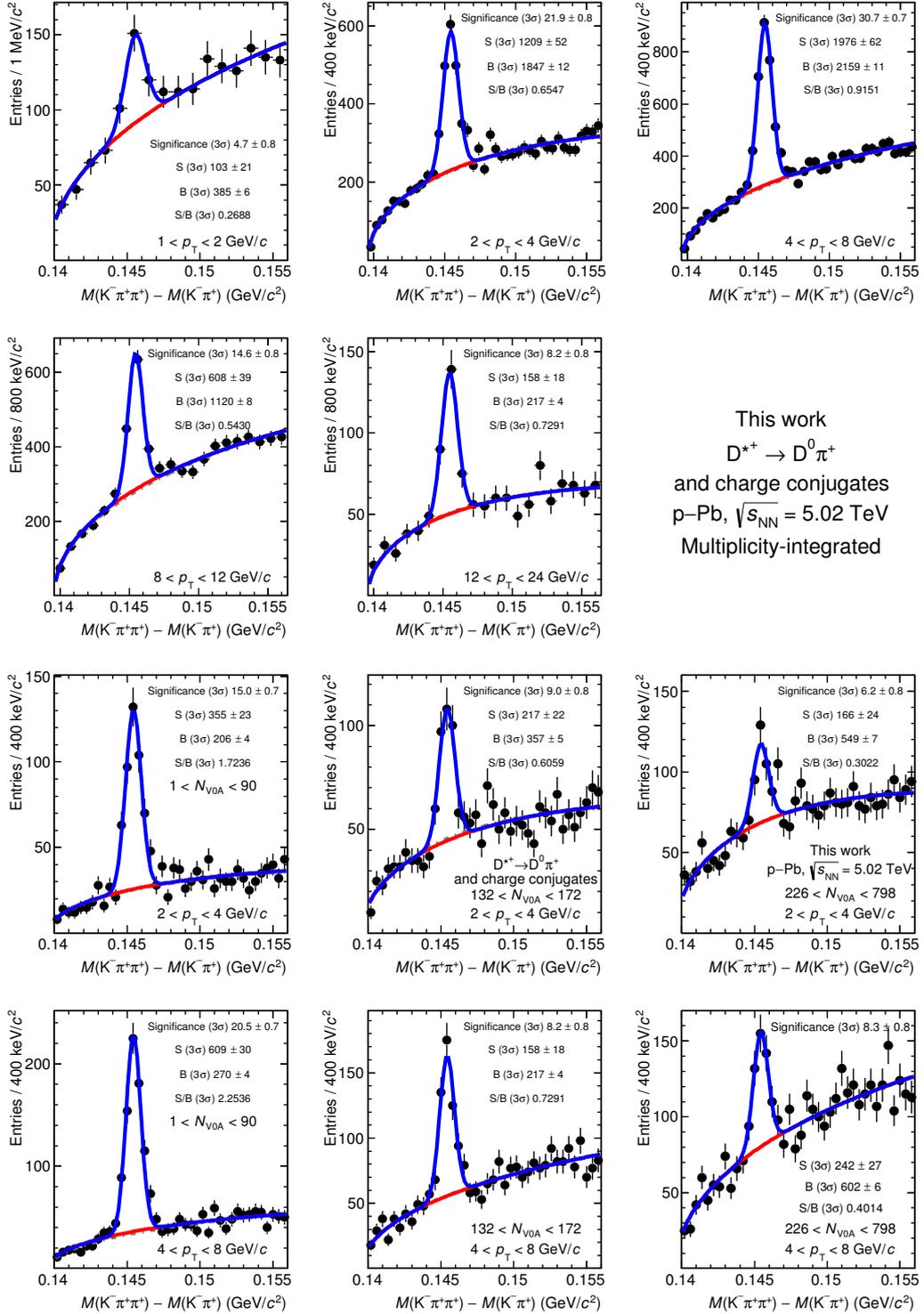
Figure 6.7: Invariant mass distributions of the mass difference for $D^{*+} \to D^0\pi^+$ in two p_T intervals and three multiplicity intervals. The top five panels show the multiplicity-integrated invariant mass spectra for five p_T intervals. The lower six panels show the distributions for two p_T intervals (rows) and three multiplicity intervals (columns).

and high) for two p_T intervals. The raw yields for all p_T and multiplicity intervals, along with their statistical uncertainties, are outlined in the appendix, in Tabs. B.1 and B.2.

Calculation of self-normalised yields

The observable that is considered in this analysis is the so-called 'self-normalised yield', defined as

$$
\frac{\left(\mathrm{d}^2 N^{\mathrm{D}}\big/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}\right)^j}{\left\langle \mathrm{d}^2 N^{\mathrm{D}}\big/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}\right\rangle} = \left(\frac{1}{N^j_{\mathrm{events}}}\frac{N^j_{\mathrm{raw\ D}}}{\varepsilon^j_{\mathrm{prompt\ D}}}\right)\bigg/\left(\frac{1}{N_{\mathrm{MB\ trigger}}/\varepsilon_{\mathrm{MB\ trigger}}}\frac{\langle N_{\mathrm{raw\ D}}\rangle}{\langle\varepsilon_{\mathrm{prompt\ D}}\rangle}\right),
$$

(6.2)

where $N_{\mathrm{raw\ D}}$ is the raw yield of D mesons, j refers to the multiplicity interval under consideration, $\varepsilon_{\mathrm{prompt\ D}}$ is the reconstruction and selection efficiency, and N^j_{events} is the number of p–Pb collision events analysed for each multiplicity class. This formula is used under the assumption that the fraction of D mesons stemming from feed-down from B-meson decays is independent of multiplicity, and thus cancels out in the ratio; however, an additional uncertainty is assigned based on the feed-down fraction determined from the efficiencies.

The variables in the denominator of Eq. (6.2) correspond to the multiplicity-integrated case. Here, $N_{\mathrm{MB\ trigger}}$ refers to the total number of inelastic p–Pb events that were triggered on in the minimum-bias sample. $\varepsilon_{\mathrm{MB\ trigger}}$ is the trigger efficiency $(96.4\pm3.1\%)$ [174]. PYTHIA Monte Carlo simulations confirmed that the minimum-bias trigger efficiency for D mesons is 100%, meaning that the number of D mesons in the triggered event sample is unaffected by the correction of the number of events for the trigger efficiency.

Systematic uncertainties

The efficiency corrections for variations in the selection cuts and PID selections in the minimum-bias analysis of D^{*+} mesons in p–Pb collisions lead to systematic uncertainties of ~8% and 5%, respectively [123]. For the relative yields as a function of multiplicity, however, these contributions cancel out in the final ratio due to the consistency of the detector response at all multiplicities, and so they are not considered as sources of systematic uncertainty for this analysis [16, 17]. The feed-down contribution from the decays of B mesons was also assumed to remain constant as a function of multiplicity, so no feed-down was explicitly subtracted from the D-meson yields. This is justified by previous measurements of beauty production as a function of multiplicity in pp collisions, as well as by PYTHIA simulations [16] and the fact

of the feed-down efficiency shape being roughly constant as a function of multiplicity (Fig. 6.6). A feed-down uncertainty was, however, applied based on the uncertainty of these efficiencies. This was estimated using the same charm-enriched Monte Carlo event set used to compute the selection efficiencies. The uncertainty based on the feed-down subtraction for D^{*+} mesons was found to be between $\sim$5–10%, depending on multiplicity.

The uncertainty on the multiplicity determination was judged by taking a series of linear fits in the individual multiplicity classes and comparing the extracted $N_{\text{ch}}/N_{\text{tracklets}}$ scaling factors with the global one used for the multiplicity conversion, see Section 6.1. This led to a 5% systematic uncertainty being assigned to the relative multiplicity for all multiplicity classes.

The most important contribution to the systematic uncertainties in this analysis comes from the yield extraction procedure. The measured yields can fluctuate based on the method used to estimate the signal, and the deviation of the final result due to this is estimated and assigned as a systematic uncertainty. For the purposes of this estimate, the 'central' result was taken to be that defined by the criteria outlined in the previous section: an automatic determination of the mean and width of the Gaussian in the multiplicity-integrated case; the fixing of these values in each p_{T} interval for the multiplicity-binned analysis; an invariant mass resolution per histogram bin of between $400\,\text{keV}/c^2$ and $1\,\text{MeV}/c^2$, depending on p_{T}; and a background defined by the function given in Eq. (4.1).

The following variations of the fit procedure were used to estimate the systematic uncertainty, with the number of variations given in parentheses:

– adjusting the mass bin width in the invariant mass distributions (4);

– applying a power-law or exponential $\times$ power-law function for the background (2);

– changing the lower and upper mass limits of the fit to the background distribution (5 $\times$ 5); and

– leaving the mass and width of the Gaussian function free in the individual multiplicity bins instead of fixing it (2).

The total number of combinations of variations was therefore 400. Additional checks were also made using bin counting within five σ ranges between 3σ and 5σ around the mean for the fit with the central parameters, using two methods for the background estimation (the fit to the side bands and to the full invariant mass range), adding a further ten variants of the fitting procedure. The estimation was performed via a multi-trial approach, whereby each possible combination of variations was tested. The same variations were applied to both the multiplicity-integrated case and each

multiplicity bin, in order to ensure that both the numerator and denominator of the relative yield were fitted in a consistent way. In order to rule out unrealistic fluctuations, the following conditions had to be met for a trial to be accepted:

- $\frac{2}{3}\sigma_{\text{int}} < \sigma_{\text{mult}} < \frac{3}{2}\sigma_{\text{int}}$, where σ_{int} and σ_{mult} are the Gaussian width for the multiplicity-integrated and multiplicity-binned cases, respectively;

- $\Sigma > 3.0$, where Σ is the statistical significance;

- Fit probability > 0.05, defined as the probability for the estimated χ^2 and number of degrees of freedom based on the incomplete gamma function.

For accepted trials, a ratio was taken between the self-normalised yield and the central value. The distribution of the ratio of the raw yield in a given multiplicity bin to the multiplicity-integrated raw yield was taken for each variation of the fitting procedure:

$$\Delta = \frac{Y_{\text{bin}}^{\text{var}}}{Y_{\text{int}}^{\text{var}}}, \tag{6.3}$$

where the superscript 'var' refers to each individual variation of the fit parameters, and the subscripts 'bin' and 'int' refer respectively to the results in individual multiplicity bins and multiplicity-integrated.
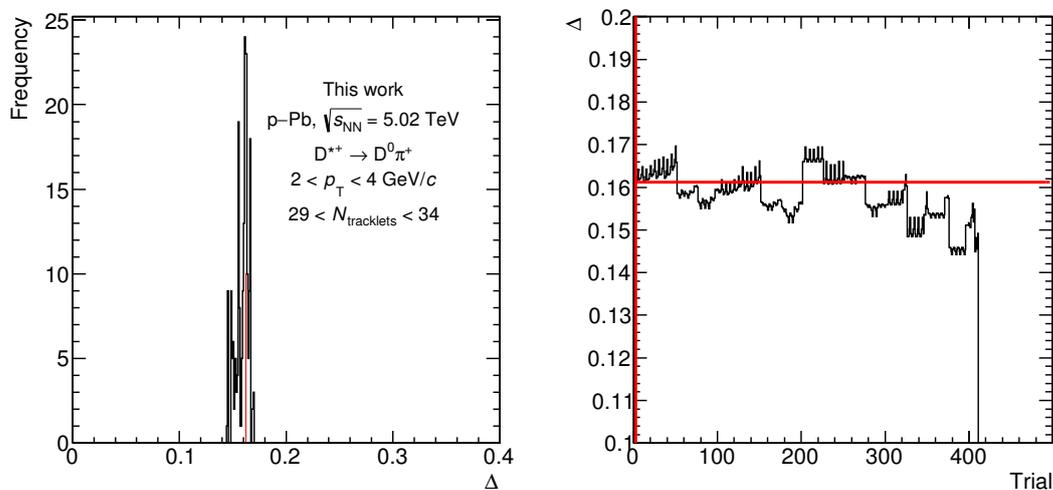
Figure 6.8: Multi-trial fit used to determine the systematic uncertainty on the yield extraction of D^{*+} mesons in multiplicity bins. Left: frequency distribution of Δ (see Eq. (6.3)) for successful trials. Right: Δ as a function of the trial number for all trials. The red lines correspond to the central values, i.e. the default fitting parameters. These plots are for $2 < p_T < 4\,\text{GeV}/c$ and $29 < N_{\text{tracklets}} < 34$.

The frequency distribution of Δ, an example of which is shown in Fig. 6.8, was then approximated with a Gaussian function, and the expected systematic uncertainty was

extracted based on the width of this distribution divided by the mean, and taking into account the distance of this mean from the value extracted with the central parameters. p_T/multiplicity intervals where the fit with the central parameters fell outside the bulk of the trials, or where the Gaussian approximation did not hold for the set of accepted trials, were excluded from the final results as having unstable fits. This typically occurred at intermediate and high multiplicity for the lowest p_T interval, and at the highest multiplicity interval for intermediate and high p_T, due to the limited D^{*+}-meson statistics in these regions leading to either low signal-to-background ratios or large statistical fluctuations in the mass peak. The number of accepted trials for the valid intervals varied between ~ 50 and the full 410. The overall systematic uncertainties varied within the range 4–9% as a function of p_T, being larger at low and high p_T and lower at intermediate p_T, but were found to be constant as a function of multiplicity for each estimator. The full list of yield extraction uncertainties can be found in Tabs. B.3 and B.4.

6.3   Results

Comparison of D^{*+} with D^0 and D^+ mesons

The results for D^{*+}-meson production are shown in Fig. 6.9 as a function of multiplicity for the $N_\mathrm{tracklets}$ multiplicity estimator, and in Fig. 6.10 as a function of the N_V0A estimator. In both figures, the results for D^{*+} mesons are plotted alongside those measured by ALICE for D^0 and D^+ mesons in the channels $D^0 \to K^- \pi^+$ and $D^+ \to K^- \pi^+ \pi^+$ (and respective charge conjugates).

Each p_T interval is shown as a single panel for all three species, with the individual species slightly offset horizontally for better visibility. Diagonal identity lines ($y = x$) are plotted to guide the eye; they correspond to the case where hard particle production processes scale 1:1 with soft particle production processes. The systematic uncertainties due to yield extraction are plotted as boxes, the statistical uncertainties as lines; the systematic uncertainties due to feed-down from B mesons is shown in the lower panels, and the global uncertainties on the event normalisation and multiplicity correction are not plotted.

All of the D-meson species are consistent with one another within the experimental uncertainties, with minimal dependence on p_T seen for both estimators. In addition, the D^{*+} meson has statistical and systematic uncertainties comparable with those found for D^0 and D^+ mesons, and smaller feed-down uncertainties overall. As the results are similar for D^0, D^+ and D^{*+}, they were combined as an average over all species to better study common features between the three. This was done as a weighted average between the three species, and is shown in Fig. 6.11, which plots all five p_T intervals together for each estimator. The averages were calculated with the
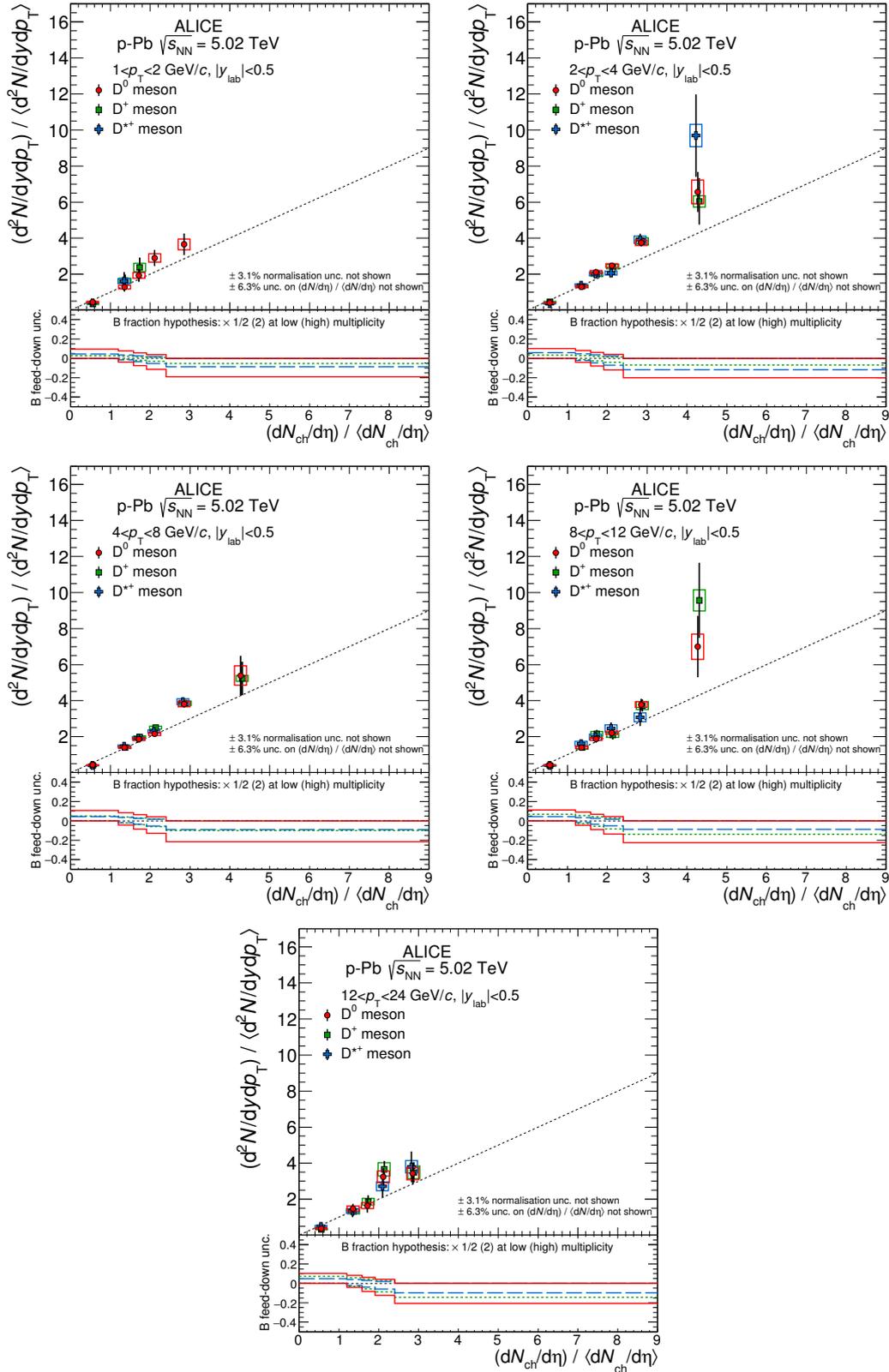
Figure 6.9: Results for D*^{*+}-, D⁰- and D⁺-meson production as a function of mid-rapidity multiplicity in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$, for five p_T intervals between 1 and 24 GeV/c. The lower panels show the systematic uncertainties due to the feed-down from B-meson decays. Figures from [17].
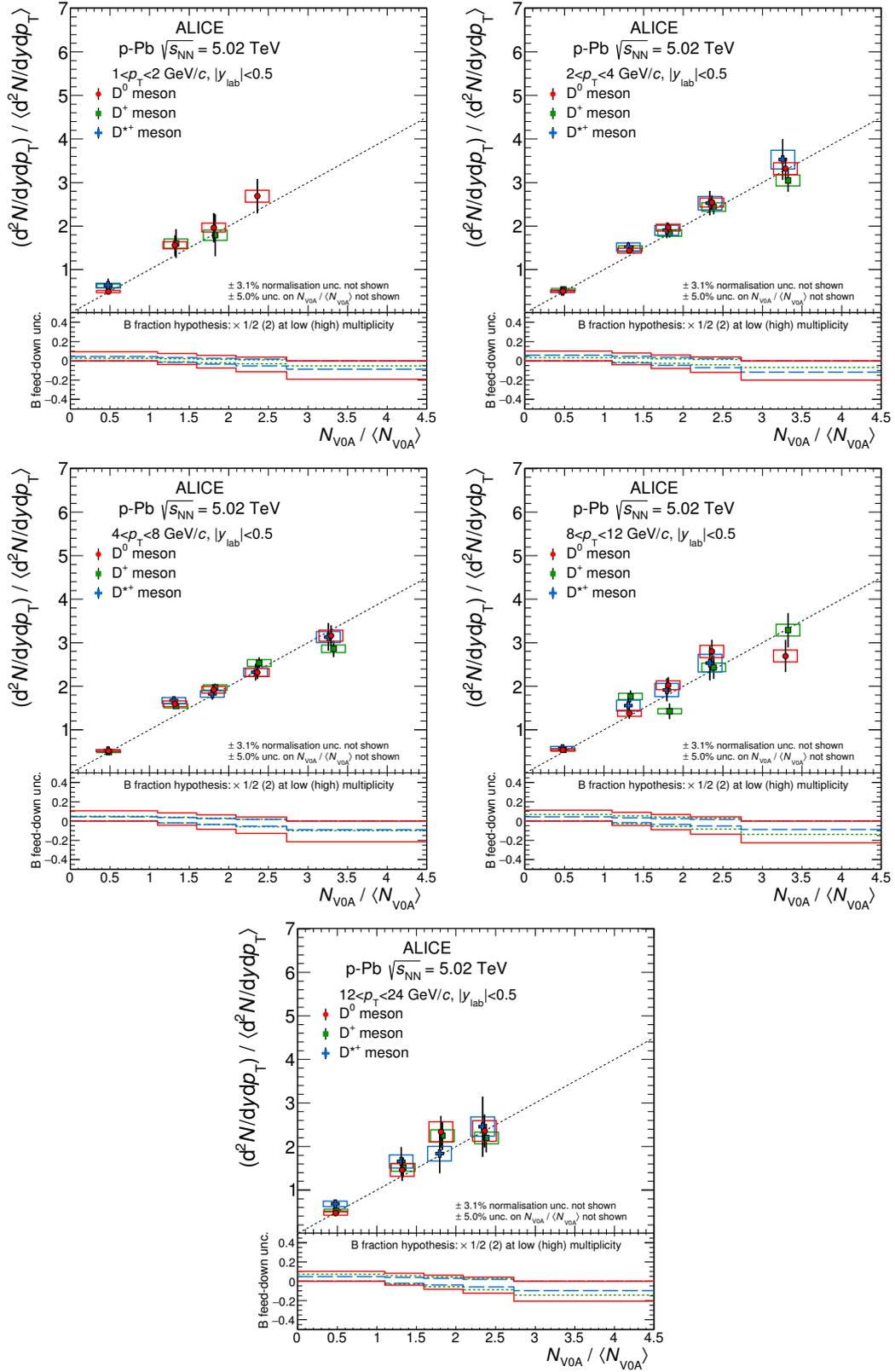
Figure 6.10: Results for D*+-, D0- and D+-meson production as a function of backward-rapidity multiplicity in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$, for five p_T intervals between 1 and 24 GeV/c. The lower panels show the systematic uncertainties due to the feed-down from B-meson decays. Figures from [17].

inverse square of the relative statistical uncertainties as a weight for each D-meson species. The full lists of the average results of all D mesons for each p_T and multiplicity class for both estimators are given in Tabs. B.5 and B.6. As with Figs. 6.9 and 6.10, the uncertainties on the B-meson feed-down fraction are shown separately, and the individual p_T intervals are slightly offset horizontally for visibility. Both the relative yields and the feed-down uncertainties were found to be approximately independent of p_T.
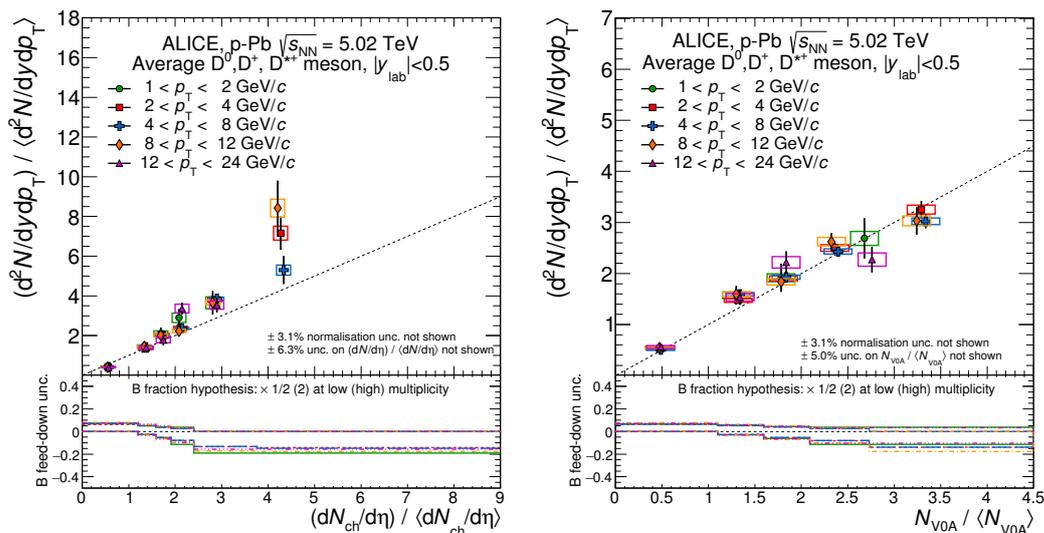

Figure 6.11: Average relative yield for D^0, D^+ and D^{*+} mesons as a function of charged-particle multiplicity at (left) mid- and (right) backward rapidity in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$. Results are shown for five p_T intervals within the range $1 < p_T < 24\,\text{GeV}/c$. The lower panels show the systematic uncertainties due to the feed-down from B-meson decays. Figures from [17].

For the mid-rapidity multiplicity estimator, there appears to be a faster-than-linear increase of the D-meson yields; when the multiplicity is estimated at backward rapidity, the increase is more linear. In order to better interpret the differences between the two rapidity regions, it is necessary to make comparisons with equivalent measurements in proton–proton collisions and theoretical models.

Comparison with pp measurements

Previous measurements of D-meson production as a function of multiplicity in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ were reported by ALICE in [16], where multi-parton interactions were shown to be a major factor in charm production at high multiplicities (see Section 2.6). Comparisons between the multiplicity-dependent results from pp and p–Pb collisions allow the contributions from MPI to be compared with from those occurring due to the presence of a nucleus in the collision system.
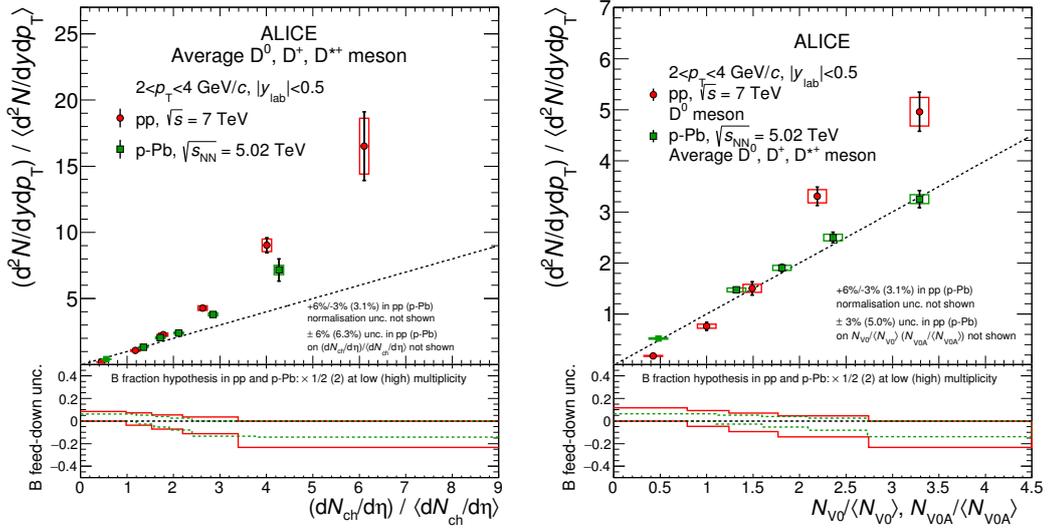
Figure 6.12: Relative D-meson yields in $|y_{\text{lab}}| < 0.5$ as a function of (left) relative charged-particle multiplicity at mid-rapidity $|\eta| < 1.0$, and (right) at backward rapidity $-5.1 < \eta < -2.8$ (including $1.7 < \eta < 3.7$ for pp collisions), for pp collisions [16] and p–Pb collisions [17]. Results are shown for $2 < p_{\text{T}} < 4\,\text{GeV}/c$. The relative yields are given in the upper panels with statistical (systematic) uncertainties drawn as bars (boxes). The uncertainty of the B feed-down fraction is drawn separately in the lower panels. A diagonal $x = y$ line (grey) is shown to guide the eye. Figure from [17]

Figure 6.12 shows the comparisons between the p–Pb and pp results for the average of all D mesons, in both of the rapidity ranges for which the charged-particle multiplicity was studied. The pp measurement with the mid-rapidity multiplicity estimator used an identical definition of multiplicity to that of the p–Pb measurement (based on $N_{\text{tracklets}}$ within $|\eta| < 1.0$ in the SPD). While the p–Pb measurement only uses the V0A counter at backward rapidity, the pp measurement employed both of the V0 detectors ($-5.1 < \eta < -2.8$ and $1.7 < \eta < 3.7$) due to the symmetry of the collision system. Only the D^0 yield was measured with this estimator for pp collisions.

For the $N_{\text{tracklets}}$ multiplicity estimator, the results were found to be consistent between pp and p–Pb collisions for all of the studied p_{T} and multiplicity intervals, implying that the multiplicity scaling is roughly similar for MPI and N_{coll} in this region. However, for the V0 estimator, the p–Pb results are lower than those from pp collisions at high multiplicity. It is possible that the shape of the collision region accounts for the differences between the results for the two systems: the presence of the Pb nucleus may lead to a higher rate of soft particle production in the Pb-going direction than at mid-rapidity, as the Pb-going multiplicity is expected to scale with the number of colliding nucleons, potentially explaining the observation that the rise of the D-meson yield relative to the backward multiplicity is slower than that relative to the mid-rapidity multiplicity estimator.

Comparison with models

The results for the average of all three D-meson species are compared with theoretical models in Figs. 6.13 and 6.14 for mid-rapidity and backward-rapidity multiplicity, respectively. In each case, the results (shown as black points) are compared in four p_T intervals within $1 < p_T < 12\,\mathrm{GeV}/c$ with calculations performed in the EPOS3 model [18,19], which was outlined in Section 2.6. EPOS uses a parton-based Gribov–Regge approach [18] to multiple scatterings, in which individual parton–parton scatterings are treated as parton ladders composed of a hard pQCD process, with the inclusion of initial- and final-state radiation. A saturation scale is used to parameterise non-linear parton evolution effects. Two versions of the calculations are used, one with a pure EPOS calculation, and one including a (3+1)-dimensional hydrodynamic evolution of the collision system [19], which introduces viscous flow to the final state.

In both rapidity intervals, the results were found to be consistent with the models within uncertainties. For the backward-rapidity multiplicity estimator, the two theory curves are in strong agreement with one another, diverging only at very high multiplicity for the highest p_T interval considered, and successfully predict the linear behaviour shown by the relative D-meson yields in the multiplicity range considered. On the other hand, there is a clear difference between the predictions with and without hydrodynamics when the multiplicity is computed at mid-rapidity. Here, when hydro is not considered, the theory curve (dark green) is consistent with the diagonal line, while the calculations with hydro show an upward curve as a function of multiplicity. This effect is more pronounced in the higher p_T intervals than in $1 < p_T < 2\,\mathrm{GeV}/c$.

The predictions including viscous hydrodynamics describe the data better than those without hydro. A similar behaviour was predicted by this model for the multiplicity dependence of charm production in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$, see Section 2.6 and Fig. 2.22. The EPOS calculations compute scatterings in pp and p–Pb collisions in a consistent fashion. In this model, the curvature is attributed to a reduction of the overall charged-particle yield in this region at high multiplicity due to hydrodynamic effects, rather than an enhancement of the charm yield. This occurs due to a proportion of the thermal energy being used to generate collective motion in the form of flow instead of particle production. As the charm yield is only dependent on hard rather than soft scatterings, it remains unaffected by this process [175]. As the result with hydro is more consistent with the data, these results therefore may hint at possible flow effects for charged particles in high-multiplicity p–Pb collisions at mid-rapidity, which would be in line with the double-ridge effect observed in p–Pb collisions at high multiplicities (see Fig. 2.19). By contrast, the effects due to flow in this model are lessened for charged particles at large Pb-going rapidities, explaining the linearity of this result in the measured multiplicity region.
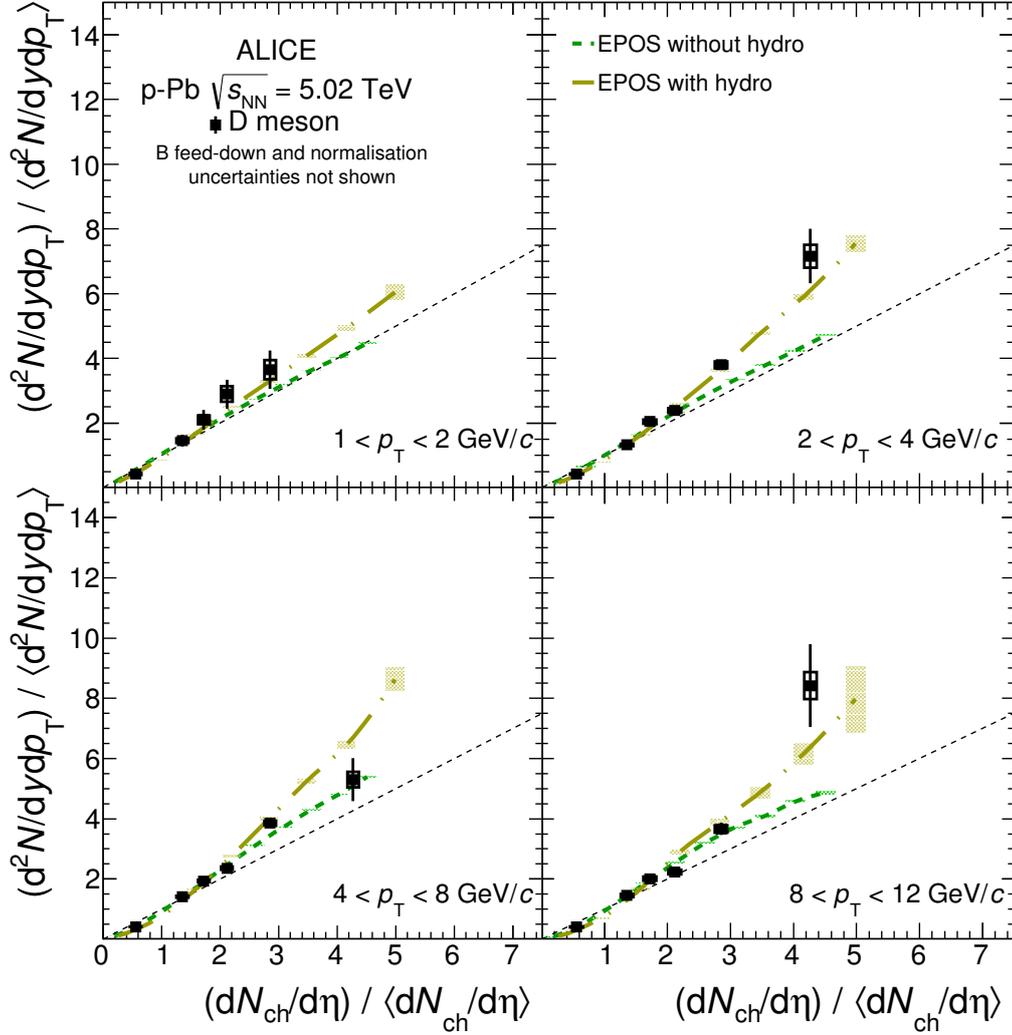
Figure 6.13: Average relative D-meson yield as a function of relative charged-particle multiplicity at mid-rapidity in four p_T intervals, compared with theoretical calculations. The systematic uncertainties on the data due to renormalisation ($\pm 3.1\%$), the $\left(\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta\right)\big/\left\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta\right\rangle$ values ($\pm 6.3\%$), the feed-down contribution are not shown. The black points represent the D-meson measurements; the coloured lines and shaded boxes respectively show the EPOS3.116 calculation curves and their statistical uncertainties, with and without hydrodynamics [18, 19]. A diagonal $x = y$ line (grey) is shown to guide the eye. Figure from [17].
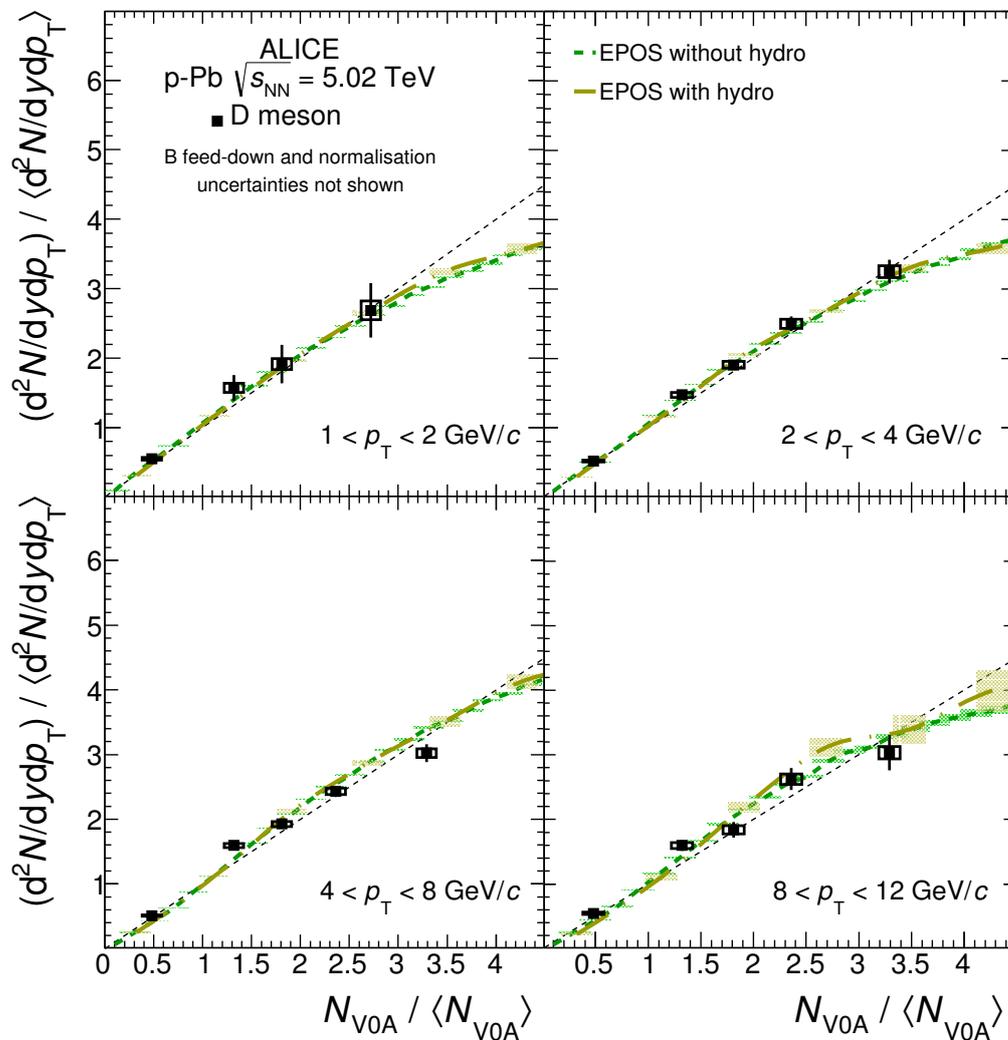
Figure 6.14: Average relative D-meson yield as a function of relative charged-particle multiplicity at mid-rapidity in four p_T intervals, compared with theoretical calculations. The systematic uncertainties on the data due to renormalisation ($\pm 3.1\%$), the $(N_{V0A})/\langle N_{V0A}\rangle$ values ($\pm 5.0\%$), the feed-down contribution are not shown. The black points represent the D-meson measurements; the coloured lines and shaded boxes respectively show the EPOS3.116 calculation curves and their statistical uncertainties, with and without hydrodynamics [18, 19]. A diagonal $x = y$ line (grey) is shown to guide the eye. Figure from [17].

6.4 Conclusions & outlook

The production of D^{*+} mesons at mid-rapidity was measured in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV as a function of the charged-particle multiplicity in two rapidity classes, mid-rapidity ($|\eta| < 1.0$) and backward rapidity (Pb-going direction, $-5.1 < \eta < -2.8$) with the ALICE detector. The decay of D^{*+} was measured in the channel D$^{*+} \to$ D$^0 \pi^+$, with the further decay D$^0 \to$ K$^- \pi^+$. The relative D^{*+} yields as a function of multiplicity for both multiplicity estimators were found to be consistent within uncertainties with those for D^0 and D$^+$ mesons measured in the same collision system in the channels D$^0 \to$ K$^- \pi^+$ and D$^+ \to$ K$^- \pi^+ \pi^+$. An average was taken between the three measured species to reduce the overall uncertainties; the resulting relative D-meson yields were found not to have any significant dependence on p_T for either multiplicity estimator. For the mid-rapidity estimator, a faster-than-linear increase was seen of the relative D-meson yield with respect to the charged-particle multiplicity, implying a faster scaling of hard scattering processes than soft particle production processes at high multiplicity densities $(dN_{ch}/d\eta \big/ \langle dN_{ch}/d\eta \rangle \approx 4.5)$. A similar scaling has also been seen for the multiplicity dependence of J/ψ-meson production at mid-rapidity in p–Pb collisions in ALICE [176]. Conversely, for the estimator at backward rapidity, the relative D-meson yield was found to scale roughly equally with the overall multiplicity up to high multiplicities in all p_T intervals.

Comparisons were then made between the results in p–Pb collisions and those already reported by the ALICE Collaboration in proton–proton collisions at $\sqrt{s} = 7$ TeV [16]. The results for the mid-rapidity estimator showed consistency between both collision systems, implying that the multiplicity scaling is roughly equivalent between pp and p–Pb collisions in this region. This implies that at mid-rapidity, the contributions from MPI and the increased number of binary nucleon–nucleon collisions in p–Pb collisions scale similarly to the MPI contributions in pp collisions at high multiplicities. When the multiplicity in the Pb-going direction was considered, however, a shallower increase for the relative D-meson yield could be seen for p–Pb collisions than pp collisions. This implies that, for this estimator, the rate of scaling of hard processes with respect to soft processes is slower in p–Pb collisions than in pp collisions, likely due to the presence of a nucleus in the collision.

Model calculations under the EPOS3.116 framework [18, 19], both with and without the inclusion of viscous hydrodynamics, were compared with the D-meson results in p–Pb collisions for both multiplicity estimators. For the backward-rapidity multiplicity estimator, both versions of the model calculations were consistent with one another, and successfully reproduced the roughly equal increase of D-meson and charged-particle yields for this estimator. For the mid-rapidity multiplicity, the model without hydrodynamics again predicted a roughly linear increase; the calculations

with hydrodynamics showed an upward curvature consistent with the D-meson measurements. It is therefore concluded that predictions involving viscous hydro better describe the behaviour of D-meson production as a function of multiplicity. Under the assumptions made in the model, the upward curvature may be attributed to a suppression of the charged-particle yield at high multiplicity, due to some proportion of the energy in the system being used for the generation of flow rather than for charged-particle production; as charm quarks are produced in the initial hard scatterings, their production rate would be unaffected by this. This result may imply an element of flow being present for light-flavour particles in high-multiplicity p–Pb events at mid-rapidity, whereas the hydrodynamic flow effects predicted by the model at backward rapidity are less significant.

Further p–Pb collisions are planned to be performed during RUN 2 of the LHC at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ and $\sqrt{s_{NN}} = 8\,\text{TeV}$, extending both the available statistics at the existing collision energy and pushing measurements to an even higher energy density. As the most significant contribution to the systematic uncertainties came from the yield extraction procedure, the increased statistics would allow both the statistical and systematic uncertainties of the multiplicity-dependent measurement to be reduced, as well as improving the granularity in terms of the number of p_T and multiplicity classes that can be stably measured for each meson. Triggering on high-multiplicity events may also allow an extension of the results to higher multiplicities.

7 | Conclusion and outlook

Results were presented for measurements of D-meson production at mid-rapidity in proton–proton and proton–lead collisions, using the ALICE detector.

A Bayesian particle identification technique was implemented and validated for the study of D^0 mesons in pp collisions at $\sqrt{s} = 7\,\text{TeV}$, in order to establish this method and determine its validity for use in other analyses in ALICE. This was motivated specifically by the case of $\Lambda_c^+ \to pK^-\pi^+$, whose signal sits atop a high level of combinatorial background, making previously established 'n_σ' PID methods ineffective for this analysis. It was previously shown that a Bayesian approach to combining the signals of different detectors (TPC and TOF) makes such an analysis viable in the current setup of the ALICE detector [6]. The D^0 analysis showed a heightened signal-to-background ratio (and thus signal purity) over n_σ PID for all p_T, and an increased (similar) statistical significance at low (high) p_T for all but the strictest Bayesian PID selections. Due to the exclusivity of the Bayesian PID selections, a lower reconstruction efficiency was seen, with tighter probability selections leading to further decreased efficiencies. Despite this, when the corrected yields for each method were compared, the results were found to be largely consistent with one another, giving a PID systematic uncertainty within 3–10% with respect to the n_σ method, depending on the strategy chosen. These deviations occur due to imperfections in the knowledge of the detector response, particularly in the tails; however, these systematic uncertainties still compare favourably with that of the n_σ PID method, which is 5% [66]. The dependence of the result on the choice of priors was also tested by varying the prior distributions of kaons, and was found to be negligible (within roughly 1% overall for the majority of the tested methods). The Bayesian PID approach is thus considered to be a valid one for the analysis of $D^0 \to K^-\pi^+$. This analysis was performed alongside other validations using high-purity samples of pions, kaons and protons from two-pronged particle decays, and single-particle spectra [6], in which the method was fully validated for analyses using the TPC and TOF detectors. Methods already exist in the ALICE PID framework to extend this approach to the other central-barrel detectors in ALICE. In particular, the single-track Bayesian PID probabilities are currently being used as an input parameter for a boosted decision tree analysis of $\Lambda_c^+ \to pK^-\pi^+$ using the Toolkit for Multivariate Analysis (TMVA) [171]

in pp and p–Pb collisions in ALICE [6], alongside complementary measurements in the channel $\Lambda_c^+ \to pK_S^0$.

The production of D^{*+} mesons was also studied in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ as a function of charged-particle multiplicity. The motivation was to estimate the role of multi-parton interactions and binary collision scaling in charm production by comparing results between pp and p–Pb collisions, as well as to examine possible collective effects in high-multiplicity p–Pb collisions. D^{*+} mesons were measured at mid-rapidity and the multiplicity was estimated at both mid-rapidity and backward rapidity. The multiplicity in both regions was corrected on an event-by-event basis by a data-driven statistical method to remove detector acceptance-induced dependences on the z-position of the vertex, and a Monte Carlo-based method was used to convert the number of tracklets in the SPD to the true pseudorapidity density of charged particles. A faster-than-linear increase of the relative D^{*+} yield was seen for the mid-rapidity estimator, along with a roughly linear increase for the estimator at backward rapidity. The results for D^{*+} mesons were found to be consistent with ALICE measurements of D^0 and D^+ mesons at all p_T and multiplicity, within uncertainties. Comparisons between pp results [16] and p–Pb results [17] for the average of all three D-meson species showed consistency for the mid-rapidity multiplicity estimator, implying that the effects of N_{coll} and MPI scaling in p–Pb collisions match those of MPI in pp collisions in this rapidity region. However, the estimator at backward (and in pp collisions, forward) rapidity showed a steeper increase of the relative D-meson yield for pp collisions than p–Pb collisions as a function of multiplicity. This behaviour was confirmed by EPOS3 phenomenological models, both with and without the application of viscous hydrodynamics [18, 19]. In this scheme, the faster-than-linear increase seen with the mid-rapidity multiplicity estimator was better reproduced by the version that included hydrodynamics; this is interpreted in this model as a reduction of the overall charged-particle yield at high multiplicity due to flow, rather than an enhancement of the overall D-meson yield. When the multiplicity is estimated at backward rapidity (the Pb-going direction), these effects disappear and the data are reproduced well by both versions of the model, implying a lesser effect of hydrodynamics on charged particles in this region.

RUN 2 of the LHC brings with it higher collision energies and luminosities in pp ($\sqrt{s} = 13\,\text{TeV}$), p–Pb ($\sqrt{s_{NN}} = 8\,\text{TeV}$) and Pb–Pb ($\sqrt{s_{NN}} = 5.02\,\text{TeV}$) collisions, allowing new energy densities to be reached as well as providing improved statistics for rare probes such as open heavy-flavour hadron production. In addition, reference measurements in pp collisions at $\sqrt{s} = 5.02\,\text{TeV}$ will make this the first energy studied in all three collision systems (pp, p–Pb and Pb–Pb), allowing more direct comparisons to be made between the three. The p–Pb data taking campaign in 2016 at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ will increase the available statistics in this system and thus allow the experimental uncertainties on the measurements to be reduced. This will allow a

more precise determination of the R_{pPb} of D mesons, potentially making the measurement sensitive to the small nuclear effects present in this system, as well as providing an opportunity for multiplicity-dependent measurements to probe higher multiplicities through improved triggers. Further p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 8\,\mathrm{TeV}$ will allow access new high energy densities in this small system. The ALICE Transition Radiation Detector, which was partially installed in RUN 1 and completed ahead of RUN 2, will also allow for improved triggering and particle identification for electrons from the semileptonic decays of heavy-flavour particles to electrons at mid-rapidity.

During the second Long Shutdown (2019–2020), a significant upgrade programme will be undertaken within both the ALICE detector and the LHC as a whole. Of particular interest for heavy-flavour measurements are the upgrades to the ALICE ITS and TPC. The ITS will be refitted with seven layers of Monolithic Active Pixel Sensors (MAPS), replacing the current six-layer SPD/SDD/SSD setup [177]. While this will remove the ability to make PID measurements with this detector, it will improve the spatial resolution by a factor of 3 and extend its tracking capabilities to ultra-low p_{T}, as well as allowing for continuous read-out from the detector. In addition, the high-luminosity upgrades to the LHC will lead to collision rates reaching up to $50\,\mathrm{kHz}$ for Pb–Pb collisions. This necessitates a continuous read-out mode for the TPC, which is not possible with its current technology. During the second Long Shutdown it will be upgraded to use Gas Electron Multiplier (GEM) foils for amplification, allowing an increase by a factor of 100 of the measurable event rate. These improvements will greatly improve the precision of measurements in rare and difficult-to-access channels such as $\Lambda_{\mathrm{c}}^{+} \to \mathrm{pK}^{-}\pi^{+}$, and D mesons at ultra-low p_{T}, allowing both access to new observables and vast improvements in the current precision of the open charm cross section. Finally, the upgrades to the ITS will facilitate the measurement of the hadronic decays of B mesons, allowing ALICE measurements of reconstructed open heavy flavours to move into the beauty sector [178].

Appendices

A | Validation of Bayesian PID for $D^0 \to K\pi$

p_T (GeV/c)	no PID	n_σ	40%	50%	70%	80%
1–2	2436 ± 350	1722 ± 189	1281 ± 119	1259 ± 115	1121 ± 105	877 ± 97
2–3	1988 ± 195	1763 ± 124	1149 ± 74	1072 ± 69	866 ± 60	737 ± 54
3–4	1817 ± 137	1595 ± 95	1117 ± 60	956 ± 53	710 ± 44	568 ± 39
4–5	1027 ± 77	962 ± 57	647 ± 37	538 ± 33	351 ± 25	279 ± 22
5–6	608 ± 49	598 ± 38	414 ± 25	354 ± 23	209 ± 16	149 ± 14
6–7	336 ± 31	331 ± 25	241 ± 18	206 ± 16	116 ± 12	89 ± 10
7–8	180 ± 23	193 ± 19	144 ± 14	120 ± 12	66 ± 9	42 ± 7
8–12	387 ± 36	397 ± 28	286 ± 20	247 ± 18	153 ± 13	97 ± 10
12–16	158 ± 27	121 ± 20	83 ± 13	75 ± 12	49 ± 8	27 ± 6

p_T (GeV/c)	max. prob.	weighted
1–2	1264 ± 116	1171 ± 116
2–3	1126 ± 73	1057 ± 70
3–4	1079 ± 58	886 ± 53
4–5	601 ± 35	520 ± 32
5–6	392 ± 24	324 ± 22
6–7	228 ± 17	181 ± 15
7–8	133 ± 13	104 ± 11
8–12	265 ± 19	215 ± 16
12–16	81 ± 12	72 ± 11

Table A.1: Raw D^0-meson yields in pp collisions at $\sqrt{s} = 7$ TeV with each applied Bayesian PID method.

Method	Average deviation	Maximum deviation	At p_T interval
40% threshold	-9.3%	-16.8%	12–16 GeV/c
50% threshold	-8.3%	-15.6%	8–12 GeV/c
70% threshold	-6.2%	-13.3%	7–8 GeV/c
80% threshold	-4.7%	-15.4%	7–8 GeV/c
Max. probability	-7.2%	-14.2%	8–12 GeV/c
Weighted	-2.7%	-6.9%	8–12 GeV/c
no PID	$+4.7\%$	$+35.8\%$	1–2 GeV/c

Table A.2: List of average and maximum deviations for corrected D^0-meson yields from n_σ PID for each PID method. The final two columns give the maximum deviation from the result with n_σ PID, and the p_T interval in which this maximum is seen.

Method	Average deviation	Maximum deviation	At p_T interval
40% threshold +10% kaons	-0.05%	2.93%	12–16 GeV/c
40% threshold -10% kaons	-0.17%	2.78%	4–5 GeV/c
50% threshold +10% kaons	-0.22%	2.43%	7–8 GeV/c
50% threshold -10% kaons	-0.19%	3.23%	5–6 GeV/c
70% threshold +10% kaons	$+1.65\%$	3.37%	5–6 GeV/c
70% threshold -10% kaons	-0.43%	6.70%	7–8 GeV/c
80% threshold +10% kaons	$+0.58\%$	25.72%	12–16 GeV/c
80% threshold -10% kaons	$+8.99\%$	70.94%	12–16 GeV/c
Max. probability +10% kaons	-0.98%	4.45%	12–16 GeV/c
Max. probability -10% kaons	-0.94%	7.30%	12–16 GeV/c
Weighted +10% kaons	$+0.13\%$	0.56%	12–16 GeV/c
Weighted -10% kaons	-0.13%	0.60%	12–16 GeV/c

Table A.3: List of average and maximum deviations for analyses with modified priors for each PID method. Percentages are given with respect to the corrected yields obtained for each method using the standard priors. The final two columns give the maximum deviation from the result with standard priors, and the p_T interval in which this maximum is seen.

B | D^{*+} as a function of multiplicity

p_{T} (GeV/c)	Multiplicity interval						
	1–200	1–21	22–28	29–34	35–43	44–69	70–200
1–2	103 ± 21	22± 7	22 ± 7	–	–	–	–
2–4	1209 ± 52	240± 19	216 ± 19	194 ± 19	194 ± 22	298±30	58 ± 14
4–8	1976 ± 62	446± 24	369 ± 24	300 ± 23	342 ± 27	499±35	–
8–12	608 ± 39	141± 15	125 ± 15	96 ± 14	114 ± 17	119±22	–
12–24	158 ± 18	44± 8	27 ± 7	–	33 ± 8	38±10	–

Table B.1: Summary table of the raw yields per p_{T} and N_{trk} multiplicity interval for D^{*+} mesons. All yields are presented with their statistical uncertainties. Excluded intervals are marked with a dash.

p_{T} (GeV/c)	Multiplicity interval					
	1–798	1–90	91–131	132–172	173–225	226–798
1–2	103 ± 21	35 ± 8	–	–	–	–
2–4	1209 ± 52	355± 23	273 ± 23	217 ± 22	193 ± 24	166 ± 24
4–8	1976 ± 62	609± 30	493 ± 29	341 ± 26	291 ± 26	242 ± 27
8–12	608 ± 39	208± 19	142 ± 18	111 ± 17	98 ± 17	–
12–24	158 ± 18	64 ± 11	39 ± 9	27 ± 7	25 ± 8	–

Table B.2: Summary table of the raw yields per p_{T} and V0A multiplicity interval for D^{*+} mesons. All yields are presented with their statistical uncertainties. Excluded intervals are marked with a dash.

	Multiplicity interval					
p_T (GeV/c)	1–21	22–28	29–34	35–43	44–69	70–200
1–2	7%	7%	–	–	–	–
2–4	5%	5%	5%	5%	5%	5%
4–8	4%	4%	4%	4%	4%	–
8–12	8%	8%	8%	8%	8%	–
12–24	9%	9%	–	9%	9%	–

Table B.3: Summary table of the yield extraction uncertainties per p_T and $N_\mathrm{tracklets}$ multiplicity interval for D^{*+} mesons. Excluded intervals are marked with a dash.

	Multiplicity interval				
p_T (GeV/c)	1–90	91–131	132–172	173–225	226–798
1–2	7%	–	–	–	–
2–4	6%	6%	6%	6%	6%
4–8	4%	4%	4%	4%	4%
8–12	8%	8%	8%	8%	–
12–24	9%	9%	9%	9%	–

Table B.4: Summary table of the yield extraction uncertainties per p_T and V0A multiplicity interval for D^{*+} mesons. Excluded intervals are marked with a dash.

	$(\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta)/\langle\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta\rangle$		
	0.56 ± 0.04	1.35 ± 0.09	1.72 ± 0.11
p_{T} (GeV/c)	$\mathrm{d}^2N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}/\langle\mathrm{d}^2N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}\rangle$		
1–2	$0.42\pm0.04\pm0.03^{+0.03}_{-0.00}$	$1.46\pm0.18\pm0.08^{+0.07}_{-0.04}$	$2.10\pm0.30\pm0.13^{+0.09}_{-0.11}$
2–4	$0.43\pm0.01\pm0.01^{+0.03}_{-0.00}$	$1.33\pm0.05\pm0.04^{+0.07}_{-0.03}$	$2.05\pm0.08\pm0.06^{+0.08}_{-0.11}$
4–8	$0.42\pm0.01\pm0.01^{+0.03}_{-0.00}$	$1.41\pm0.04\pm0.03^{+0.07}_{-0.04}$	$1.93\pm0.06\pm0.04^{+0.08}_{-0.10}$
8–12	$0.41\pm0.02\pm0.01^{+0.03}_{-0.00}$	$1.45\pm0.08\pm0.05^{+0.09}_{-0.05}$	$2.01\pm0.13\pm0.07^{+0.10}_{-0.13}$
12–24	$0.40\pm0.04\pm0.02^{+0.03}_{-0.00}$	$1.39\pm0.15\pm0.08^{+0.09}_{-0.04}$	$1.77\pm0.26\pm0.13^{+0.09}_{-0.12}$

	$(\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta)/\langle\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta\rangle$		
	2.11 ± 0.13	2.85 ± 0.18	4.27 ± 0.27
p_{T} (GeV/c)	$\mathrm{d}^2N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}/\langle\mathrm{d}^2N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}\rangle$		
1–2	$2.90\pm0.44\pm0.24^{+0.11}_{-0.33}$	$3.65\pm0.60\pm0.30^{+0.00}_{-0.70}$	–
2–4	$2.40\pm0.09\pm0.07^{+0.06}_{-0.19}$	$3.80\pm0.13\pm0.11^{+0.00}_{-0.51}$	$7.16\pm0.84\pm0.35^{+0.00}_{-1.02}$
4–8	$2.36\pm0.08\pm0.05^{+0.06}_{-0.18}$	$3.86\pm0.11\pm0.08^{+0.00}_{-0.51}$	$5.30\pm0.71\pm0.25^{+0.00}_{-0.79}$
8–12	$2.23\pm0.15\pm0.08^{+0.07}_{-0.22}$	$3.67\pm0.21\pm0.12^{+0.00}_{-0.63}$	$8.42\pm1.38\pm0.45^{+0.00}_{-1.49}$
12–24	$3.37\pm0.30\pm0.22^{+0.11}_{-0.32}$	$3.53\pm0.37\pm0.21^{+0.00}_{-0.55}$	–

Table B.5: Average of relative D^0, D^+ and D^{*+} meson yields in several multiplicity and p_{T} intervals for p–Pb collisions at $\sqrt{s_{\mathrm{NN}}}=5.02$ TeV, as a function of the relative charged-particle multiplicity at mid-rapidity. Values are given with their uncertainties, in the following order: statistical, systematic and feed-down contribution. The global normalisation uncertainty of 3.1% is not shown. Excluded intervals are shown as a dash. Table from [17].

	$N_{\mathrm{V0A}}/\langle N_{\mathrm{V0A}}\rangle$		
	0.48 ± 0.02	1.32 ± 0.07	1.81 ± 0.09
p_{T} (GeV/c)	$\mathrm{d}^2 N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}/\langle\mathrm{d}^2 N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}\rangle$		
1–2	$0.55\pm0.05\pm0.02^{+0.04}_{-0.00}$	$1.57\pm0.19\pm0.06^{+0.09}_{-0.05}$	$1.92\pm0.28\pm0.08^{+0.09}_{-0.12}$
2–4	$0.52\pm0.01\pm0.01^{+0.03}_{-0.00}$	$1.47\pm0.05\pm0.04^{+0.08}_{-0.04}$	$1.91\pm0.07\pm0.05^{+0.08}_{-0.10}$
4–8	$0.51\pm0.01\pm0.01^{+0.03}_{-0.00}$	$1.59\pm0.04\pm0.03^{+0.08}_{-0.04}$	$1.93\pm0.06\pm0.04^{+0.08}_{-0.10}$
8–12	$0.55\pm0.02\pm0.02^{+0.04}_{-0.00}$	$1.60\pm0.08\pm0.05^{+0.10}_{-0.05}$	$1.84\pm0.12\pm0.06^{+0.09}_{-0.13}$
12–24	$0.56\pm0.04\pm0.03^{+0.04}_{-0.00}$	$1.53\pm0.15\pm0.07^{+0.09}_{-0.05}$	$2.22\pm0.22\pm0.11^{+0.11}_{-0.14}$

	$N_{\mathrm{V0A}}/\langle N_{\mathrm{V0A}}\rangle$	
	2.36 ± 0.12	3.29 ± 0.16
	(2.72 ± 0.14)	
p_{T} (GeV/c)	$\mathrm{d}^2 N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}/\langle\mathrm{d}^2 N/\mathrm{d}y\mathrm{d}p_{\mathrm{T}}\rangle$	
1–2	$(2.69\pm0.39\pm0.14^{+0.10}_{-0.31})$	
2–4	$2.50\pm0.11\pm0.07^{+0.07}_{-0.20}$	$3.25\pm0.17\pm0.09^{+0.00}_{-0.45}$
4–8	$2.43\pm0.09\pm0.05^{+0.06}_{-0.19}$	$3.02\pm0.14\pm0.07^{+0.00}_{-0.42}$
8–12	$2.62\pm0.17\pm0.08^{+0.09}_{-0.27}$	$3.03\pm0.27\pm0.10^{+0.00}_{-0.53}$
12–24	$(2.27\pm0.25\pm0.13^{+0.08}_{-0.24})$	

Table B.6: Average of relative D^0, D^+ and D^{*+} meson yields in several multiplicity and p_{T} intervals for p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, as a function of the relative charged-particle multiplicity at mid-rapidity. Values are given with their uncertainties, in the following order: statistical, systematic and feed-down contribution. The global normalisation uncertainty of 3.1% is not shown. For $1 < p_{\mathrm{T}} < 2$ GeV/c and $12 < p_{\mathrm{T}} < 24$ GeV/c, the final two multiplicity intervals are merged for D^0 and D^+ mesons, and the results are shown in parentheses. Table from [17].

References

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

[2] **STAR** Collaboration, "Measurements of identified particles at intermediate transverse momentum in the STAR experiment from Au–Au collisions at $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$," `arXiv:nucl-ex/0601042 [nucl-ex]`.

[3] **ALICE** Collaboration, "K_S^0 and Λ production in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$," *Phys. Rev. Lett.* **111** (Nov, 2013) 222301.

[4] R. C. Hwa and L. Zhu, "Multi-minijet Contribution to Hadronic Spectra and Correlations in Pb-Pb Collisions at 2.76 TeV and beyond," `arXiv:1202.2091 [nucl-th]`.

[5] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences*. Cambridge University Press, 2005.

[6] **ALICE** Collaboration, "Particle identification in ALICE: a Bayesian approach," `arXiv:1602.01392 [physics.data-an]`.

[7] **CMS** Collaboration, "Observation of long-range near-side angular correlations in proton–lead collisions at the LHC," *Phys. Lett.* **B718** (2013) 795–814, `arXiv:1210.5482 [nucl-ex]`.

[8] **ALICE** Collaboration, "Long-range angular correlations on the near and away side in p–Pb collisions at $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$," *Phys. Lett.* **B719** (2013) 29–41, `arXiv:1212.2001 [nucl-ex]`.

[9] **ATLAS** Collaboration, "Observation of associated near-side and away-side long-range correlations in $\sqrt{s_{NN}} = 5.02\,\mathrm{TeV}$ proton–lead collisions with the ATLAS detector," *Phys. Rev. Lett.* **110** no. 18, (2013) 182302, `arXiv:1212.5198 [hep-ex]`.

[10] M. Strikman, "Transverse nucleon structure and multiparton interactions," *Acta Phys. Polon.* **B42** (2011) 2607–2630, `arXiv:1112.3834 [hep-ph]`.

[11] **CDF** Collaboration, "Double parton scattering in $\bar{p}p$ collisions at $\sqrt{s} = 1.8\mathrm{TeV}$," *Phys. Rev.* **D56** (1997) 3811–3832.

[12] **CMS** Collaboration, "Measurement of prompt J/ψ pair production in pp collisions at $\sqrt{s} = 7$ Tev," *JHEP* **09** (2014) 094, arXiv:1406.0484 [hep-ex].

[13] **D0** Collaboration, "Observation and studies of double J/ψ production at the Tevatron," *Phys. Rev.* **D90** no. 11, (2014) 111101, arXiv:1406.2380 [hep-ex].

[14] **D0** Collaboration, "Evidence for simultaneous production of J/ψ and Υ mesons," *Phys. Rev. Lett.* **116** no. 8, (2016) 082002, arXiv:1511.02428 [hep-ex].

[15] **LHCb** Collaboration, "Production of associated Υ and open charm hadrons in pp collisions at $\sqrt{s} = 7$ and 8 TeV via double parton scattering," arXiv:1510.05949 [hep-ex].

[16] **ALICE** Collaboration, "Measurement of charm and beauty production at central rapidity versus charged-particle multiplicity in proton–proton collisions at $\sqrt{s} = 7$ TeV," *JHEP* **09** (2015) 148, arXiv:1505.00664 [nucl-ex].

[17] **ALICE** Collaboration, "Measurement of D-meson production versus multiplicity in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV," arXiv:1602.07240 [nucl-ex].

[18] H. Drescher *et al.*, "Parton based Gribov-Regge theory," *Phys.Rept.* **350** (2001) 93–289, arXiv:hep-ph/0007198 [hep-ph].

[19] K. Werner *et al.*, "Analysing radial flow features in p–Pb and pp collisions at several TeV by studying identified particle production in EPOS3," *Phys.Rev.* **C89** (2014) 064903, arXiv:1312.1233 [nucl-th].

[20] S. Hughes, "Lecture notes: Electricity & magnetism." http://web.mit.edu/sahughes/www/8.022/lec01.pdf.

[21] **Belle** Collaboration, "Observation of a narrow charmoniumlike state in exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}$J/$\psi$ decays," *Phys. Rev. Lett.* **91** (Dec, 2003) 262001.

[22] **LHCb** Collaboration, "Observation of J/ψp resonances consistent with pentaquark states in $\Lambda_b^0 \to$ J/ψK$^- p$ decays," *Phys. Rev. Lett.* **115** (Aug, 2015) 072001.

[23] **Particle Data Group** Collaboration, K. A. Olive *et al.*, "Review of Particle Physics," *Chin. Phys.* **C38** (2015) 090001 (and 2015 update).

[24] V. Borka Jovanovic *et al.*, "Constituent quark masses obtained from hadron masses with contributions of Fermi-Breit and Glozman-Riska hyperfine interactions," *Phys. Rev.* **D82** (2010) 117501, `arXiv:1011.1749 [hep-ph]`.

[25] X. Zhu *et al.*, "D$\overline{\text{D}}$ correlations as a sensitive probe for thermalization in high-energy nuclear collisions," *Phys. Lett.* **B647** (2007) 366–370, `arXiv:hep-ph/0604178 [hep-ph]`.

[26] F. Weber, "Strange quark matter and compact stars," *Prog. Part. Nucl. Phys.* **54** (2005) 193–288, `arXiv:astro-ph/0407155 [astro-ph]`.

[27] **CDF** Collaboration, "First observation of electroweak single top quark production," *Phys. Rev. Lett.* **103** (2009) 092002, `arXiv:0903.0885 [hep-ex]`.

[28] **D0** Collaboration, "Observation of single top quark production," *Phys. Rev. Lett.* **103** (2009) 092001, `arXiv:0903.0850 [hep-ex]`.

[29] **ATLAS** Collaboration, "Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS detector," *Phys. Rev.* **D90** no. 11, (2014) 112006, `arXiv:1406.7844 [hep-ex]`.

[30] **CMS** Collaboration, "Measurement of the t-channel single-top-quark production cross section and of the $|V_{\text{tb}}|$ CKM matrix element in pp collisions at $\sqrt{s} = 8\,\text{TeV}$," *JHEP* **06** (2014) 090, `arXiv:1403.7366 [hep-ex]`.

[31] E. D. Bloom *et al.*, "High-energy inelastic e–p scattering at 6° and 10°," *Phys. Rev. Lett.* **23** (Oct, 1969) 930–934.

[32] M. Breidenbach *et al.*, "Observed behavior of highly inelastic electron–proton scattering," *Phys. Rev. Lett.* **23** (Oct, 1969) 935–939.

[33] G. Sterman *et al.*, "Handbook of perturbative QCD," *Rev. Mod. Phys.* **67** (Jan, 1995) 157–248.

[34] "Das CBM Experiment an FAIR." `http://astro.uni-wuppertal.de/html/Projekte/CBM.htm`.

[35] N. Cabibbo and G. Parisi, "Exponential hadronic spectrum and quark liberation," *Physics Letters B* **59** no. 1, (1975) 67–69.

[36] R. Hagedorn, "Hadronic matter near the boiling point," *Nuovo Cim.* **A56** (1968) 1027–1057.

[37] P. Bicudo, "QCD string tension curve, the ferromagnetic magnetization, and the quark–antiquark confining potential at finite temperature," *Phys. Rev. D* **82** (Aug, 2010) 034507.

[38] C. Wong, *Introduction to High-energy Heavy-ion Collisions.* Introduction to High-energy Heavy-ion Collisions. World Scientific, 1994.

[39] A. Vuorinen, *The pressure of QCD at finite temperature and quark number density.* PhD thesis, Helsinki U., 2003. `arXiv:hep-ph/0402242 [hep-ph]`. `https://inspirehep.net/record/1265574/files/thepress.pdf`.

[40] **HotQCD** Collaboration, "Equation of state in (2+1)-flavor QCD," *Phys. Rev. D90* (2014) 094503, `arXiv:1407.6387 [hep-lat]`.

[41] M. Kliemant, R. Sahoo, T. Schuster, and R. Stock, "Global Properties of Nucleus–Nucleus Collisions," *Lect. Notes Phys.* **785** (2010) 23–103, `arXiv:0809.2482 [nucl-ex]`.

[42] C. Markert, R. Bellwied, and I. Vitev, "Formation and decay of hadronic resonances in the QGP," *Physics Letters B* **669** no. 1, (2008) 92 – 97.

[43] R. J. Glauber and G. Matthiae, "High-energy scattering of protons by nuclei," *Nucl. Phys.* **B21** (1970) 135–157.

[44] M. L. Miller *et al.*, "Glauber modeling in high energy nuclear collisions," *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 205–243, `arXiv:nucl-ex/0701025 [nucl-ex]`.

[45] 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.

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

[47] J. Wilkinson, "Analysis and extrapolation of D-meson cross sections at the LHC," Master's thesis, Universität Heidelberg & Imperial College London, 2011.

[48] S. Schael *et al.*, "Precision electroweak measurements on the *Z* resonance," *Phys. Rept.* **427** (2006) 257–454, `arXiv:hep-ex/0509008 [hep-ex]`.

[49] V. Kartvelishvili, A. Likhoded, and V. Petrov, "On the fragmentation functions of heavy quarks into hadrons," *Physics Letters B* **78** no. 5, (1978) 615 – 617.

[50] M. G. Bowler, "e$^+$e$^-$ production of heavy quarks in the string model," *Zeitschrift für Physik C Particles and Fields* **11** no. 2, (1981) 169–174. http://dx.doi.org/10.1007/BF01574001.

[51] P. D. B. Collins and T. P. Spiller, "The fragmentation of heavy quarks," *J. Phys.* **G11** (1985) 1289.

[52] T. Matsui and H. Satz, "J/ψ Suppression by Quark–Gluon Plasma Formation," *Phys. Lett.* **B178** (1986) 416–422.

[53] H. Satz, "Colour deconfinement and quarkonium binding," *J. Phys.* **G32** (2006) R25, arXiv:hep-ph/0512217 [hep-ph].

[54] H. Satz, "Probing the States of Matter in QCD," *Int. J. Mod. Phys.* **A28** (2013) 1330043, arXiv:1310.1209 [hep-ph].

[55] A. Andronic, P. Braun-Munzinger, K. Redlich, and J. Stachel, "Evidence for charmonium generation at the phase boundary in ultra-relativistic nuclear collisions," *Phys. Lett.* **B652** (2007) 259–261, arXiv:nucl-th/0701079 [NUCL-TH].

[56] A. Andronic, F. Beutler, P. Braun-Munzinger, K. Redlich, and J. Stachel, "Statistical hadronization of heavy flavor quarks in elementary collisions: Successes and failures," *Phys. Lett.* **B678** (2009) 350–354, arXiv:0904.1368 [hep-ph].

[57] K. Schweda, "Prompt production of D mesons with ALICE at the LHC," arXiv:1402.1370 [nucl-ex]. Habilitation thesis, Universität Heidelberg.

[58] **ALICE** Collaboration, "Measurement of electrons from beauty hadron decays in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Lett.* **B721** (2013) 13–23, arXiv:1208.1902 [hep-ex].

[59] **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, arXiv:1201.3791 [hep-ex].

[60] **LHCb** Collaboration, "Prompt charm production in pp collisions at $\sqrt{s} = 7$ TeV," *Nucl. Phys.* **B871** (2013) 1–20, arXiv:1302.2864 [hep-ex].

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

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

[63] J. Baines *et al.*, "Heavy quarks (Working Group 3): Summary Report for the HERA-LHC Workshop Proceedings," `arXiv:hep-ph/0601164 [hep-ph]`.

[64] 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.

[65] Christian Möhler, "D-meson production at ultra-low transverse momentum in proton–proton collisions with ALICE at the LHC," Master's thesis, Universität Heidelberg, 2014.

[66] **ALICE** Collaboration, "Measurement of charm production at central rapidity in proton–proton collisions at $\sqrt{s} = 7\,\text{TeV}$," *JHEP* **1201** (2012) 128, `arXiv:1111.1553 [hep-ex]`.

[67] **ALICE** Collaboration, "Measurement of charm production at central rapidity in proton–proton collisions at $\sqrt{s} = 2.76$ TeV," *JHEP* **07** (2012) 191, `arXiv:1205.4007 [hep-ex]`.

[68] **STAR** Collaboration, "Measurements of D^0 and D^* production in pp collisions at $\sqrt{s} = 200\,\text{GeV}$," *Phys. Rev.* **D86** (2012) 072013, `arXiv:1204.4244 [nucl-ex]`.

[69] **PHENIX** Collaboration, A. Adare *et al.*, "Heavy quark production in pp and energy loss and flow of heavy quarks in Au–Au collisions at $\sqrt{s_{\text{NN}}} = 200$ GeV," *Phys. Rev.* **C84** (2011) 044905, `arXiv:1005.1627 [nucl-ex]`.

[70] **ATLAS** Collaboration, "Comparison of $D^{(*)}$ meson production cross sections with FONLL and GM-VFNS predictions,". `https://cds.cern.ch/record/1378479`.

[71] **LHCb** Collaboration, "Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13\,\text{TeV}$," *JHEP* **03** (2016) 159, `arXiv:1510.01707 [hep-ex]`.

[72] S. Alekhin and S. Moch, "Heavy-quark deep-inelastic scattering with a running mass," *Phys. Lett.* **B699** (2011) 345–353, `arXiv:1011.5790 [hep-ph]`.

[73] S. Alekhin and S. O. Moch, "Running Heavy-Quark Masses in DIS," in *19th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2011) Newport News, Virginia, April 11-15, 2011.* 2011. `arXiv:1107.0469 [hep-ph]`. `https://inspirehep.net/record/916873/files/arXiv:1107.0469.pdf`.

[74] **H1 & ZEUS** Collaboration, "Combination and QCD Analysis of Charm Production Cross Section Measurements in Deep-Inelastic ep Scattering at HERA," *Eur.Phys.J.* **C73** (2013) 2311, `arXiv:1211.1182 [hep-ex]`.

[75] **CDF II** Collaboration, "Measurement of prompt charm meson production cross sections in p$\overline{\text{p}}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys. Rev. Lett.* **91** (Dec, 2003) 241804.

[76] **CLEO** Collaboration, "Charm production in nonresonant e$^+$e$^-$ annihilations at $\sqrt{s} = 10.55$ GeV," *Phys. Rev. D* **37** (Apr, 1988) 1719–1743.

[77] **CLEO** Collaboration, "Erratum: Charm production in nonresonant e$^+$e$^-$ annihilations at $\sqrt{s} = 10.55$ GeV," *Phys. Rev. D* **39** (Mar, 1989) 1471–1472.

[78] **ALEPH** Collaboration, "Study of charm production in Z decays," *Eur. Phys. J.* **C16** (2000) 597–611, `arXiv:hep-ex/9909032 [hep-ex]`.

[79] **CDF** Collaboration, "Measurement of the Direct Charm Meson Production Cross Section at CDF," 2013. `http://www-cdf.fnal.gov/physics/new/bottom/030403.blessed-dxsec/`. CDF Note 6623.

[80] **STAR** Collaboration, M. Calderon de la Barca Sanchez *et al.*, "Open charm production from d + Au collisions in STAR," *Eur. Phys. J.* **C43** (2005) 187–192.

[81] C. D. Buchanan and S. B. Chun, "Simple predictive model for flavor production in hadronization," *Phys. Rev. Lett.* **59** (Nov, 1987) 1997–2000.

[82] R. Averbeck *et al.*, "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," `arXiv:1107.3243 [hep-ph]`.

[83] **ALICE** Collaboration, "Production of charged pions, kaons and protons at large transverse momenta in pp and Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV," *Phys. Lett.* **B736** (2014) 196–207, `arXiv:1401.1250 [nucl-ex]`.

[84] S. Wicks, W. Horowitz, M. Djordjevic, and M. Gyulassy, "Heavy quark jet quenching with collisional plus radiative energy loss and path length fluctuations," *Nucl. Phys.* **A783** (2007) 493–496, `arXiv:nucl-th/0701063 [nucl-th]`.

[85] N. Armesto, C. A. Salgado, and U. A. Wiedemann, "Medium induced gluon radiation off massive quarks fills the dead cone," *Phys. Rev.* **D69** (2004) 114003, `arXiv:hep-ph/0312106 [hep-ph]`.

References

[86] M. Djordjevic and M. Gyulassy, "Heavy quark radiative energy loss in QCD matter," *Nucl. Phys.* **A733** (2004) 265–298, `arXiv:nucl-th/0310076 [nucl-th]`.

[87] B.-W. Zhang, E. Wang, and X.-N. Wang, "Heavy quark energy loss in nuclear medium," *Phys. Rev. Lett.* **93** (2004) 072301, `arXiv:nucl-th/0309040 [nucl-th]`.

[88] **ALICE** Collaboration, "Centrality dependence of high-p_T D meson suppression in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *JHEP* **11** (2015) 205, `arXiv:1506.06604 [nucl-ex]`.

[89] **STAR** Collaboration, "Observation of D^0 meson nuclear modifications in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV," *Phys. Rev. Lett.* **113** no. 14, (2014) 142301, `arXiv:1404.6185 [nucl-ex]`.

[90] **ALICE** Collaboration, "Transverse momentum dependence of D-meson production in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *JHEP* **03** (2016) 081, `arXiv:1509.06888 [nucl-ex]`.

[91] **ALICE** Collaboration, "Two-pion Bose-Einstein correlations in central Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *Phys. Lett.* **B696** (2011) 328–337, `arXiv:1012.4035 [nucl-ex]`.

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

[93] S. Batsouli, S. Kelly, M. Gyulassy, and J. L. Nagle, "Does the charm flow at RHIC?," *Phys. Lett.* **B557** (2003) 26–32, `arXiv:nucl-th/0212068 [nucl-th]`.

[94] V. Greco, C. M. Ko, and R. Rapp, "Quark coalescence for charmed mesons in ultrarelativistic heavy ion collisions," *Phys. Lett.* **B595** (2004) 202–208, `arXiv:nucl-th/0312100 [nucl-th]`.

[95] P. F. Kolb and U. W. Heinz, "Hydrodynamic description of ultrarelativistic heavy ion collisions," `arXiv:nucl-th/0305084 [nucl-th]`.

[96] Z. Matthews, "Quantum Diaries: Flow." `http://www.quantumdiaries.org/2011/11/07/flow/`.

[97] U. Heinz, "RHIC serves the perfect fluid – hydrodynamic flow of the Quark–Gluon Plasma," 2005.

134

`http://pyweb.swan.ac.uk/xqcd/talks/heinz1.pdf`. Extreme QCD Conference, Swansea, Aug. 2–5.

[98] R. Grajcarek, *Anisotropic flow of the charmed D^{*+} meson in non-central Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV*. PhD thesis, Universität Heidelberg, 2013.

[99] **ALICE** Collaboration, "Azimuthal anisotropy of D meson production in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV," *Phys. Rev.* **C90** no. 3, (2014) 034904, `arXiv:1405.2001 [nucl-ex]`.

[100] S. Wicks, W. Horowitz, M. Djordjevic, and M. Gyulassy, "Elastic, inelastic, and path length fluctuations in jet tomography," *Nuclear Physics A* **784** no. 1–4, (2007) 426 – 442.

[101] W. A. Horowitz and M. Gyulassy, "The Surprising Transparency of the sQGP at LHC," *Nucl. Phys.* **A872** (2011) 265–285, `arXiv:1104.4958 [hep-ph]`.

[102] W. A. Horowitz, "Testing pQCD and AdS/CFT Energy Loss at RHIC and LHC," *AIP Conf. Proc.* **1441** (2012) 889–891, `arXiv:1108.5876 [hep-ph]`.

[103] W. M. Alberico *et al.*, "Heavy-flavour spectra in high energy nucleus–nucleus collisions," *Eur. Phys. J.* **C71** (2011) 1666, `arXiv:1101.6008 [hep-ph]`.

[104] M. Monteno *et al.*, "Heavy-flavor dynamics in nucleus–nucleus collisions: from RHIC to LHC," *J. Phys.* **G38** (2011) 124144, `arXiv:1107.0256 [hep-ph]`.

[105] S. Cao, G.-Y. Qin, and S. A. Bass, "Heavy-quark dynamics and hadronization in ultrarelativistic heavy-ion collisions: Collisional versus radiative energy loss," *Phys. Rev.* **C88** (2013) 044907, `arXiv:1308.0617 [nucl-th]`.

[106] J. Uphoff, O. Fochler, Z. Xu, and C. Greiner, "Elliptic flow and energy loss of heavy quarks in ultrarelativistic heavy ion collisions," *Phys. Rev. C* **84** (Aug, 2011) 024908.

[107] O. Fochler, J. Uphoff, Z. Xu, and C. Greiner, "Jet quenching and elliptic flow at the RHIC and the LHC within a pQCD-based partonic transport model," *Journal of Physics G: Nuclear and Particle Physics* **38** no. 12, (2011) 124152. `http://stacks.iop.org/0954-3899/38/i=12/a=124152`.

[108] J. Uphoff, O. Fochler, Z. Xu, and C. Greiner, "Open Heavy Flavor in Pb+Pb Collisions at $\sqrt{s} = 2.76$ TeV within a Transport Model," *Phys. Lett.* **B717** (2012) 430–435, `arXiv:1205.4945 [hep-ph]`.

[109] T. Lang *et al.*, "Heavy quark transport in heavy ion collisions at energies available at the BNL Relativistic Heavy Ion Collider and at the CERN Large Hadron Collider within the UrQMD hybrid model," *Phys. Rev.* **C93** no. 1, (2016) 014901, `arXiv:1211.6912 [hep-ph]`.

[110] T. Lang, H. van Hees, J. Steinheimer, Y.-P. Yan, and M. Bleicher, "Heavy quark transport at RHIC and LHC," *Journal of Physics: Conference Series* **426** no. 1, (2013) 012032. `http://stacks.iop.org/1742-6596/426/i=1/a=012032`.

[111] M. Nahrgang, J. Aichelin, P. B. Gossiaux, and K. Werner, "Influence of hadronic bound states above T_c on heavy-quark observables in Pb + Pb collisions at the CERN Large Hadron Collider," *Phys. Rev. C* **89** (Jan, 2014) 014905.

[112] 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, `arXiv:0902.4154 [hep-ph]`.

[113] I. Schmidt, "Physics of nuclear antishadowing," in *11th International Conference on Elastic and Diffractive Scattering: Towards High Energy Frontiers: The 20th Anniversary of the Blois Workshops, 17th Rencontre de Blois (EDS 05) Chateau de Blois, Blois, France, May 15-20, 2005.* 2005. `arXiv:hep-ph/0511173 [hep-ph]`.

[114] A. Szczurek and A. Budzanowski, "Fermi motion and nuclear modification factor," *Mod. Phys. Lett.* **A19** (2004) 1669–1680, `arXiv:nucl-th/0311025 [nucl-th]`.

[115] **EMC** Collaboration, "The ratio of the nucleon structure functions f_2^n for iron and deuterium," *Physics Letters B* **123** no. 3, (1983) 275 – 278.

[116] M. Arneodo, "Nuclear effects in structure functions," *Physics Reports* **240** no. 5, (1994) 301 – 393.

[117] **ALICE** Collaboration, F. Carminati *et al.*, "ALICE: Physics performance report, volume I," *J. Phys.* **G30** (2004) 1517–1763.

[118] **ALICE** Collaboration, B. Alessandro *et al.*, "ALICE: Physics performance report, volume II," *J. Phys.* **G32** (2006) 1295–2040.

[119] B. Z. Kopeliovich, J. Nemchik, A. Schafer, and A. V. Tarasov, "Cronin effect in hadron production off nuclei," *Phys. Rev. Lett.* **88** (2002) 232303, `arXiv:hep-ph/0201010 [hep-ph]`.

[120] J. W. Cronin, H. J. Frisch, M. J. Shochet, J. P. Boymond, P. A. Piroué, and R. L. Sumner, "Production of hadrons at large transverse momentum at 200, 300, and 400 gev," *Phys. Rev. D* **11** (Jun, 1975) 3105–3123.

[121] D. Antreasyan, J. W. Cronin, H. J. Frisch, M. J. Shochet, L. Kluberg, P. A. Piroué, and R. L. Sumner, "Production of hadrons at large transverse momentum in 200, 300, and 400 GeV pp and p–nucleus collisions," *Phys. Rev. D* **19** (Feb, 1979) 764–778.

[122] **ALICE** Collaboration, "Transverse momentum dependence of inclusive primary charged-particle production in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV," *Eur. Phys. J.* **C74** no. 9, (2014) 3054, `arXiv:1405.2737 [nucl-ex]`.

[123] **ALICE** Collaboration, "Measurement of prompt D-meson production in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV," *Phys.Rev.Lett.* **113** no. 23, (2014) 232301, `arXiv:1405.3452 [nucl-ex]`.

[124] R. Sharma, I. Vitev, and B.-W. Zhang, "Light-cone wave function approach to open heavy flavor dynamics in QCD matter," *Phys. Rev. C* **80** (Nov, 2009) 054902.

[125] A. M. Sickles, "Possible evidence for radial flow of heavy mesons in d + Au collisions," *Physics Letters B* **731** (2014) 51–56.

[126] Z.-B. Kang *et al.*, "Multiple scattering effects on heavy meson production in p+A collisions at backward rapidity," *Phys. Lett. B* **740** (2015) 23–29.

[127] **ALICE** Collaboration, "Measurement of electrons from heavy-flavour hadron decays in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV," *Phys. Lett.* **B754** (2016) 81–93, `arXiv:1509.07491 [nucl-ex]`.

[128] **ALICE** Collaboration, "Suppression of high transverse momentum D mesons in central Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *JHEP* **09** (2012) 112, `arXiv:1203.2160 [nucl-ex]`.

[129] **PHENIX** Collaboration, "Nuclear modification of ψ', χ_c, and J/ψ production in d + Au collisions at $\sqrt{s_{NN}} = 200$ GeV," *Phys. Rev. Lett.* **111** (Nov, 2013) 202301.

[130] **ALICE** Collaboration, "Suppression of $\psi(2S)$ production in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV," *JHEP* **12** (2014) 073, `arXiv:1405.3796 [nucl-ex]`.

[131] **ALICE** Collaboration, "Centrality dependence of particle production in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV," *Phys. Rev.* **C91** no. 6, (2015) 064905, `arXiv:1412.6828 [nucl-ex]`.

[132] **ALICE** Collaboration, "Charged-particle multiplicity measurement in proton–proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC," *Eur. Phys. J.* **C68** (2010) 345–354, `arXiv:1004.3514 [hep-ex]`.

[133] Sjöstrand, Torbjorn and Mrenna, Stephen and Skands, Peter Z., "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.* **178** (2008) 852–867, `arXiv:0710.3820 [hep-ph]`.

[134] E. G. Ferreiro and C. Pajares, "High multiplicity pp events and J/ψ production at LHC," *Phys. Rev.* **C86** (2012) 034903, `arXiv:1203.5936 [hep-ph]`.

[135] E. G. Ferreiro and C. Pajares, "Open charm production in high multiplicity proton–proton events at the LHC," `arXiv:1501.03381 [hep-ph]`.

[136] **ALICE** Collaboration, "The ALICE experiment at the CERN LHC," *Journal of Instrumentation* **3** no. 08, (2008) S08002. `http://stacks.iop.org/1748-0221/3/i=08/a=S08002`.

[137] L. Evans and P. Bryant, "Lhc machine," *Journal of Instrumentation* **3** no. 08, (2008) S08001. `http://stacks.iop.org/1748-0221/3/i=08/a=S08001`.

[138] **ATLAS** Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider," *Journal of Instrumentation* **3** no. 08, (2008) S08003. `http://stacks.iop.org/1748-0221/3/i=08/a=S08003`.

[139] **CMS** Collaboration, "The CMS experiment at the CERN LHC," *Journal of Instrumentation* **3** no. 08, (2008) S08004. `http://stacks.iop.org/1748-0221/3/i=08/a=S08004`.

[140] **ATLAS** Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett.* **B716** (2012) 1–29, `arXiv:1207.7214 [hep-ex]`.

[141] **CMS** Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett.* **B716** (2012) 30–61, `arXiv:1207.7235 [hep-ex]`.

[142] **LHCb** Collaboration, "The LHCb Detector at the LHC," *Journal of Instrumentation* **3** no. 08, (2008) S08005. `http://stacks.iop.org/1748-0221/3/i=08/a=S08005`.

[143] C. De Melis, "The CERN accelerator complex. Complexe des accélérateurs du CERN,". `http://cds.cern.ch/record/2119882`.

[144] **ALICE** Collaboration, "Performance of the ALICE Experiment at the CERN LHC," *Int. J. Mod. Phys.* **A29** (2014) 1430044, `arXiv:1402.4476 [nucl-ex]`.

[145] **ALICE** Collaboration, *ALICE Inner Tracking System (ITS): Technical Design Report.* ALICE Technical Design Report. CERN, 1999. `https://cds.cern.ch/record/391175`.

[146] **ALICE** Collaboration, "First proton–proton collisions at the LHC as observed with the ALICE detector: measurement of the charged particle pseudorapidity density at $\sqrt{s} = 900\,\mathrm{GeV}$," *Eur. Phys. J.* **C65** (2010) 111–125, `arXiv:0911.5430 [hep-ex]`.

[147] **ALICE** Collaboration, "LHC restart - Monday 23 November 2009," 2009. `http://aliceinfo.cern.ch/Public/en/Chapter1/restartd4.html`.

[148] W. Blum, W. Riegler, and L. Rolandi, *Particle detection with drift chambers; 2nd ed.* Springer, Berlin, 2008. `https://cds.cern.ch/record/1105920`.

[149] H. Bethe, "Bremsformel für Elektronen relativistischer Geschwindigkeit," *Zeitschrift für Physik* **76** no. 5, (1932) 293–299.

[150] F. Bloch, "Zur Bremsung rasch bewegter Teilchen beim Durchgang durch Materie," *Annalen der Physik* **408** no. 3, (1933) 285–320.

[151] L. Landau, "On the energy loss of fast particles by ionization," *J. Phys.(USSR)* **8** (1944) 201–205.

[152] J. Alme *et al.*, "The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **622** no. 1, (2010) 316 – 367.

[153] **ALICE** Collaboration, "ALICE Time-Of-Flight system (TOF): Technical Design Report,". `https://cds.cern.ch/record/430132`.

[154] M. Bondila *et al.*, "ALICE T0 detector," *IEEE Trans. Nucl. Sci.* **52** (2005) 1705–1711.

[155] M. Rybczynski and W. Broniowski, "Wounded nucleon model with realistic nucleon–nucleon collision profile and observables in relativistic heavy-ion collisions," *Phys. Rev.* **C84** (2011) 064913, `arXiv:1110.2609 [nucl-th]`.

[156] W. Broniowski, P. Bozek, and M. Rybczynski, "Fluctuating initial conditions in heavy-ion collisions from the Glauber approach," *Phys. Rev.* **C76** (2007) 054905, `arXiv:0706.4266 [nucl-th]`.

References

[157] W. Broniowski and M. Rybczynski, "Two-body nucleon–nucleon correlations in Glauber models of relativistic heavy-ion collisions," *Phys. Rev.* **C81** (2010) 064909, arXiv:1003.1088 [nucl-th].

[158] **ALICE** Collaboration, "Charged-particle multiplicity measurement in proton–proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC," *Eur. Phys. J.* **C68** (2010) 89–108, arXiv:1004.3034 [hep-ex].

[159] **ALICE** Collaboration, "Centrality determination of Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV with ALICE," *Phys. Rev.* **C88** no. 4, (2013) 044909, arXiv:1301.4361 [nucl-ex].

[160] M. Ivanov *et al.*, "Track reconstruction in high density environment," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **566** no. 1, (2006) 70 – 74. Proceedings of the 1st Workshop on Tracking in High Multiplicity Environments.

[161] Y. Belikov, M. Ivanov, K. Safarik, and J. Bracinik, "TPC tracking and particle identification in high density environment," *eConf* **C0303241** (2003) TULT011, arXiv:physics/0306108 [physics].

[162] P. Sigmund, *Particle Penetration and Radiation Effects: General Aspects and Stopping of Swift Point Charges.* Springer Series in Solid-State Sciences. Springer Berlin Heidelberg, 2006.

[163] A. Dainese, *Charm production and in-medium QCD energy loss in nucleus nucleus collisions with ALICE: A Performance study.* PhD thesis, Padua U., 2003. arXiv:nucl-ex/0311004 [nucl-ex].

[164] Torbjörn Sjöstrand and Stephen Mrenna and Peter Skands, "Pythia 6.4 physics and manual," *Journal of High Energy Physics* **2006** no. 05, (2006) 026. http://stacks.iop.org/1126-6708/2006/i=05/a=026.

[165] P. Z. Skands, "The Perugia Tunes," in *Proceedings, 1st International Workshop on Multiple Partonic Interactions at the LHC (MPI08),* pp. 284–297. 2009. arXiv:0905.3418 [hep-ph].

[166] R. Brun *et al.*, *GEANT: Detector Description and Simulation Tool; Oct 1994.* CERN Program Library. CERN, 1993. https://cds.cern.ch/record/1082634.

[167] Y. Pachmayer for the **ALICE** Collaboration, "Particle identification with the ALICE Transition Radiation Detector," *Nucl. Instrum. Meth.* **A766** (2014) 292–295, `arXiv:1402.3508 [physics.ins-det]`.

[168] **ALICE** Collaboration, "Measurement of pion, kaon and proton production in proton–proton collisions at $\sqrt{s} = 7\,\mathrm{TeV}$," *Eur. Phys. J.* **C75** (2015) 226, `arXiv:1504.00024 [nucl-ex]`.

[169] **ALICE** Collaboration, "Pion, kaon, and proton production in central Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$," *Phys. Rev. Lett.* **109** (2012) 252301, `arXiv:1208.1974 [hep-ex]`.

[170] **ALICE** Collaboration, "Multiplicity Dependence of pion, kaon, proton and lambda production in p–Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$," *Phys. Lett.* **B728** (2014) 25–38, `arXiv:1307.6796 [nucl-ex]`.

[171] A. Hoecker *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," `arXiv:physics/0703039 [physics]`.

[172] **ALICE** Collaboration, "Pseudorapidity density of charged particles in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$," *Phys. Rev. Lett.* **110** (2013) 032301, `arXiv:1210.3615 [nucl-ex]`.

[173] X.-N. Wang and M. Gyulassy, "HIJING: A Monte Carlo model for multiple jet production in pp, pA, and AA collisions," *Phys. Rev. D* **44** (Dec, 1991) 3501–3516.

[174] **ALICE** Collaboration, "Measurement of visible cross sections in proton–lead collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV in van der Meer scans with the ALICE detector," *JINST* **9** no. 11, (2014) P11003, `arXiv:1405.1849 [nucl-ex]`.

[175] K. Werner *et al.*, "Charm production in high multiplicity pp events," `arXiv:1602.03414 [nucl-th]`.

[176] M. Winn, *Inclusive J/ψ production at mid-rapidity in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$.* PhD thesis, Universität Heidelberg, 2016.

[177] F. Reidt, *Studies for the ALICE Inner Tracking System Upgrade.* PhD thesis, Universität Heidelberg, 2016.

[178] J. Stiller, *Full kinematic reconstruction of charged B mesons with the upgraded Inner Tracking System of the ALICE Experiment.* PhD thesis, Universität Heidelberg, 2016.

[179] D. Binosi and L. Theußl, "JaxoDraw: A graphical user interface for drawing Feynman diagrams," *Computer Physics Communications* **161** no. 1–2, (2004) 76 – 86.

Acknowledgements

I'm also grateful to the internal review committees for both papers: Michael Weber, Peter Hristov, Alexander Kalweit, Gianluca Usai, Torsten Dahms, and Leonardo Milano, and to the ALICE Editorial Board, institution groups and individual members who lent their comments and suggestions during the preparation of both papers. The input we received during the internal review stages was comprehensive, always appreciated, and greatly helped to further strengthen the quality of both publications.

The Feynman diagrams in Figs. 2.6 and 2.9 were drawn using JaxoDraw [179].

Finally, my warmest thanks go to my Mum Denise, my Dad Alan, my sister Charlotte, and the rest of my family (both human, feline and canine) for their love and support through the good and the bad throughout my life and particularly these last few years. You have been a bedrock and an inspiration for everything I've done, and no matter how far away I went, and no matter how long for, I always knew there was a home I could go back to with good music, great food, and wonderful company. Without you, none of this would have been possible.