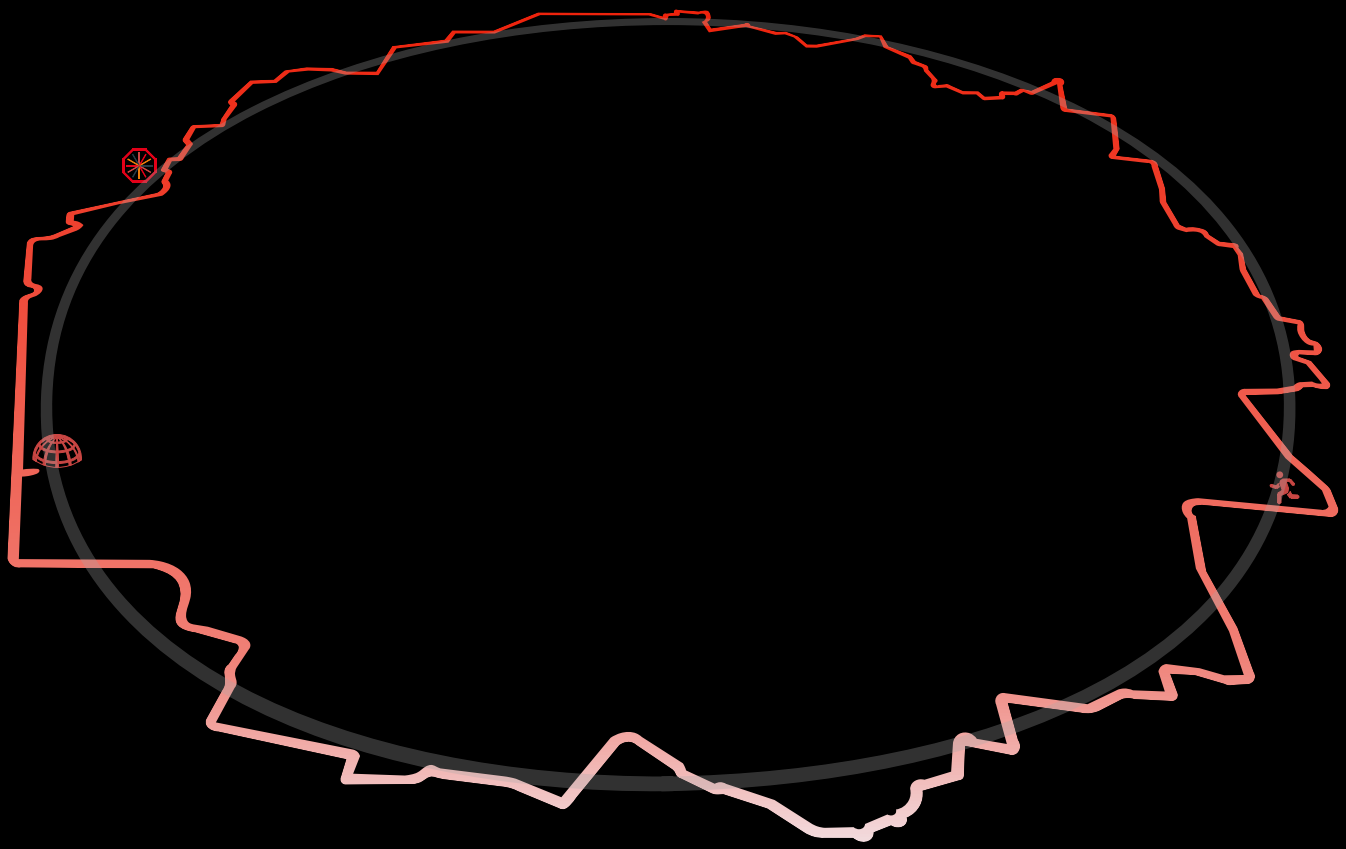
Hadronisation of Heavy Quarks

Production measurements of heavy-flavour hadrons from small to large collision systems



Luuk Vermunt

Hadronisation of Heavy Quarks

production measurements of heavy-flavour hadrons from small to large collision systems

Luuk Vermunt

ISBN: 978-94-6423-773-3

A catalogue record is available from the Utrecht University Library.

Cover design by Luuk Vermunt. *Running the Large Hadron Collider: 34.5 km, 2h55m.*
Printed by ProefschriftMaken

Hadronisation of Heavy Quarks

production measurements of heavy-flavour hadrons from small to large collision systems

Hadronisatie van Zware Quarks

productie metingen van zware hadronen in kleine tot grote botsingssystemen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 16 mei 2022 des middags te 12:15 uur

door

Lucas Anne Vermunt

geboren op 4 september 1993
te Nijmegen

Promotor:	Prof.dr. R. Snellings
Copromotoren:	Dr. A. Mischke[†]
	Dr. A. Grelli

This research was partially funded by the Netherlands Organisation for Scientific Research.

"I've got a million things that I need to do, but they're all secondary"

Alex Turner

Contents

CHAPTER 1

General introduction

1

1.1 Introduction

In any field of science, a good collaboration between theorists and experimentalists is crucial to push further the boundaries of our knowledge. Where theorists mostly try to think about how the world *could* be, the job of an experimentalist is to figure out which of these ideas is actually correct. Even though the "glory" is, arguably, most of the time received for the theoretical work, the key to success is a joint effort between both type of scientists. A good example is the field of elementary particle physics, which investigates the theory that all matter around us is made of smaller, elementary particles. In the 1950s and 60s, when beams of ultra-relativistic particles became experimentally available, and dozens of — at that point seen as — elementary particles where discovered, theorists started to develop an overarching theory trying to describe all particles and interactions. In the early 70s, these developments lead to the *Standard Model* (SM) of particle physics [1], later affirmed by many experimental discoveries, with the Higgs boson in 2012 [2, 3] as final piece in the jigsaw. A puzzle with holes in it though, and quite significant ones [4]. For decades, theorists and experimentalists have been working together on new theories to explain the deficiencies of the SM, but without luck[1]. As a good friend of mine in microbiology wrote: "*Particle physicists must now all be experts in designing specific experiments that can test if a theory is significantly false*". But humility demands that we accept the scientific riddles that we cannot solve here and now. On the other hand, how awful would it be if we were able to understand everything right away, with nothing left to discover? Anyhow, if the future will give us a contradictory observable explained by new theories, or a postulated theory beyond the SM which is validated by measurements, progress will follow from a collaboration between both type of physicists.

In the dissertation you are currently reading, I will present my own small contribution to our understanding of the universe, conducted in these four years of research within the ALICE experiment [7]. The work focuses on a better understanding of one of the fundamental interactions incorporated in the SM: the strong interaction, described by quantum chromodynamics (QCD). In the world of high-energy physics, where major efforts are devoted to precision measurements of the Standard Model and searches for theories beyond it, nuclear physics was — and perhaps still is — seen as an underdog. Fortunately, this is gradually changing after recent and exiting measurements that demonstrate the importance of achieving a better understanding of QCD for particle physics in general [8–13]. My work on this topic can be divided in three independent experimental analyses and one performance simulation study, each presented in a separate chapter in this manuscript. To appreciate better the results, I will first provide some of the theoretical background on heavy-ion physics in this chapter and specifically on heavy-flavour physics in the next. In this latter chapter I will end with the aim of this work, making the connection with a few of the big open questions in the field. Chapter 3 and 4 are dedicated to the experimental setup and the used analysis strategies and methods. Finally, in Chapters 5 to

[1] Very recently, two exiting measurements from the LHCb [5] and Muon g-2 [6] Collaborations were published, inconsistent with SM predictions, which might hint to new particles or forces not accounted for in the Standard Model. The results are not statistical significant yet though, thus so far we can only hope it is the beginning of a breakthrough.

8, my own analyses are presented: D-meson production in pp collisions, charm-hadron ratios as function of event multiplicity in pp collisions, Λ_c^+-baryon production in Pb–Pb collisions, and a prospect study for the measurement of B_s^0 mesons in the near future, respectively.

Knowing the scientific community a bit, I can safely assume that most of you do not have the time to read a full dissertation like this. Of course I would still advise you to do so, but if time is scarce, I would recommend you to read at least the first (introduction) and last (discussion) section of the analysis chapters, where, respectively, the importance of the measurement is introduced and my personal interpretation and view on the results are presented. Having a glimpse at the figures in the result section does no harm either. I hope you enjoy reading this dissertation.

1.2 Standard Model and quantum chromodynamics

The Standard Model of particle physics, which is a relativistic quantum field theory based on a local gauge symmetry of the type SU(3) × SU(2) × U(1), is able to classify all the known elementary particles and describes the strong, weak, and electromagnetic interactions. The fourth known interaction, gravitation, cannot be described with this theory[2]. Each force acts on specific charges, and is "carried" by different particles called gauge bosons. The electromagnetic force acts between particles with electric charge and is carried by photons (γ); the weak interaction is carried by $W^{\pm}$ and Z^0 bosons and affects particles with flavour charge; gluons (g) carry the strong interactions between particles with colour charge. The SM includes twelve elementary particles (and twelve antiparticles) with half-integer spin, known as fermions, which are classified in two groups: the quarks (experience all three types of interaction) and the leptons (experience the electromagnetic and/or the weak interaction). A full description of the SM can be found in Ref. [14]. In this thesis we will focus on the strong force, which is described by *quantum chromodynamics*.

Historically the roots of QCD lie with nuclear physics and the description of ordinary matter like protons and neutrons. Nowadays, it is used (and required) to describe what happens at high-energy hadron colliders. It is seen as a precise and elegant theory, and appears as an expanded version of quantum electrodynamics (QED), the relativistic quantum field theory of electrodynamics, which is a generation older. Despite the beauty of the theory, its mathematical description [15] is outside the scope of this dissertation, and we will therefore rely on an "illustrative" language.

Where QED has just one kind of charge (electric charge), in QCD the conserved charge, called *colour charge*, can take three values: red, green, and blue. As for electric charge in QED, the colour charges are conserved in all physical processes, and the interaction is exchanged by massless gauge bosons, called gluons, in a similar way as photons do in QED. The colour charge is carried by quarks, which are elementary particles, for which we know six different

[2]This is one of the major flaws of the Standard Model, together with the strong CP problem, neutrino oscillations, matter–antimatter asymmetry, and the nature of dark matter and dark energy [4].

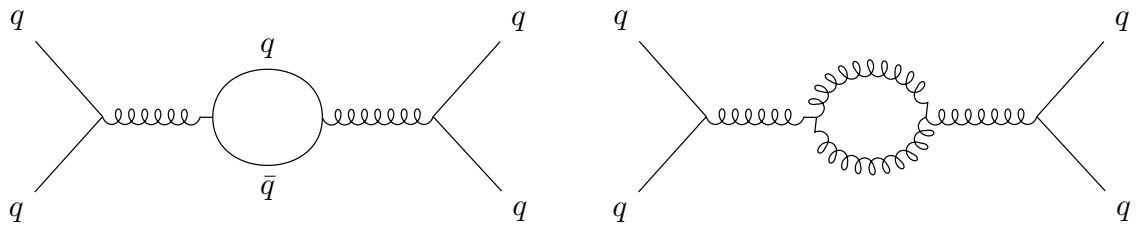


Figure 1.1: First-order Feynman diagrams defining what is called screening (left) and anti-screening (right) in the QCD interaction.

kind of flavours: up, down, strange, charm, beauty, and top, often denoted by their first letter. Ordinary matter consists mostly of u and d quarks, where the others are much heavier and decay through the weak interaction. Quarks are spin-1/2 point particles, like electrons.

Besides the she similarities between QED and QCD, there are also crucial differences. Where photons are charge neutral, gluons must carry unbalanced colour charges, since they are allowed to change one colour into another, and colour charge itself is a conserved quantity. This leads to the most distinct feature of QCD with respect to the other forces, gluons are coloured, hence they can self-interact. These self-interactions results in what is called *anti-screening* in colour interaction.

To understand anti-screening, and its consequences, a brief detour to charge screening in QED is helpful. Interactions among particles are effectively determined by summing all possible processes that can occur. For example, the process of a photon in transit should be considered together with that photon turning into an electron–positron pair which annihilate back to a photon at a later moment. Since electrons emit virtual photons all the time, they will be surrounded by a cloud of virtual electron–positron pairs, where the positive charges of the positrons partially "block" the negative charge of the electron. This process is called screening, and reduces the interaction strength of the electromagnetic force. Likewise, the QCD vacuum consists of virtual $q\bar{q}$ pairs, screening the initial colour charges (left panel Fig. 1.1). However, due to the gluon self-interaction, a pair of gluons can be formed as well (right panel), and because those carry colour charge, the effective charge becomes larger, hence increasing the interaction strength. This mechanisms is called anti-screening, and is actually the dominant one for QCD. This follows from the calculation (to lowest non-trivial order) of the so-called beta function, which for the SU(3) symmetry group of QCD is negative as long as there are no more than sixteen flavours of quark. Since there are only six known quark flavours, anti-screening dominates and the effective strong *coupling constant* (α_s) thus becomes small at short distances, i.e. it reduces for increased momentum exchange (Q^2) between the two quarks, which is confirmed by many experimental results as presented in Fig. 1.2. The running of the coupling constant defines two of the unique phenomenons in QCD, *colour confinement* and *asymptotic freedom*. For decreasing Q^2 (increasing distances), α_s diverges, meaning that quarks are strongly bound in hadrons and cannot be separated. For large values of Q^2 (decreasing distances), α_s decreases to the point that the hadron constituents can be viewed as quasi-free particles.

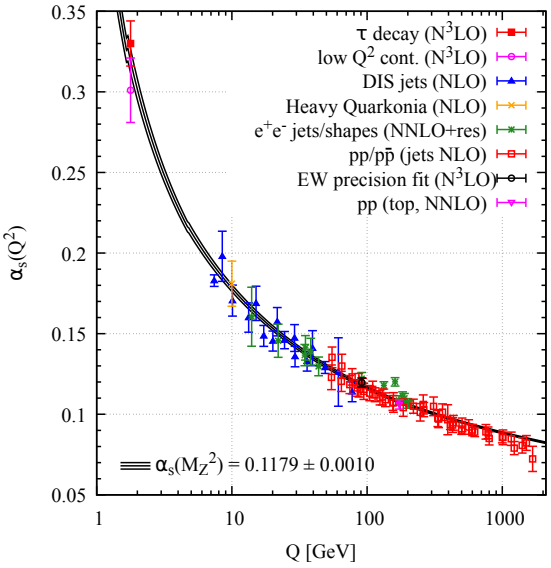


Figure 1.2: Summary of the experimental values of the strong coupling constant, α_s, as function of the energy scale Q, compared to predictions from QCD. Figure taken from Ref. [16].

It is not easy to relate the underlying concepts of QCD to observed phenomena in nature, it quickly becomes too impractical to solve the equations analytically. For high-energy QCD interactions ($Q^2 \gtrsim 1$ GeV$^2/c^2$) one can rely on *perturbative QCD* (pQCD), making use of terms involving powers of α_s to calculate the transition elements of the scattering matrix. For low-energy interactions, α_s becomes too large for a perturbative approach, so effective theories and/or numerical methods need to be used instead. *Lattice QCD* (lQCD) [17] is one of those approaches, using non-perturbative calculations on a discrete space–time grid. It has been successfully applied to many properties of hadrons, as well as used to determine fundamental parameters of the SM. Phenomenology is a third way to "solve" QCD, in particular by studying certain regimes in the phase diagram of nuclear matter that can be both probed by experiments and (partly) constrained by perturbative or lattice QCD calculations. Now we are in the fortunate situation that such regimes are very interesting, making it one of the main objectives of high-energy nuclear physics. This will be explained in more detail in the next sections.

1.3 High-energy nuclear physics

As a result of the large variation in the strong coupling constant, the properties of nuclear matter can be very different: it is expected to exhibit different states, especially at extremely high temperatures and/or densities. In the early universe, microseconds after the Big Bang, or in the core of neutron stars — both phenomena in nature that we are eager to understand better — such conditions should be met, i.e. nuclear matter will behave differently than in normal day life, making it of interest to study the phase diagram of nuclear matter. These circumstances can be investigated in the laboratory via high-energy collisions between heavy nuclei[3], while lQCD calculations can constrain phenomenological models in this regime, providing us with both theoretical and experimental handles to tackle the problem.

[3]The reason why we call high-energy nuclear physics often heavy-ion physics.

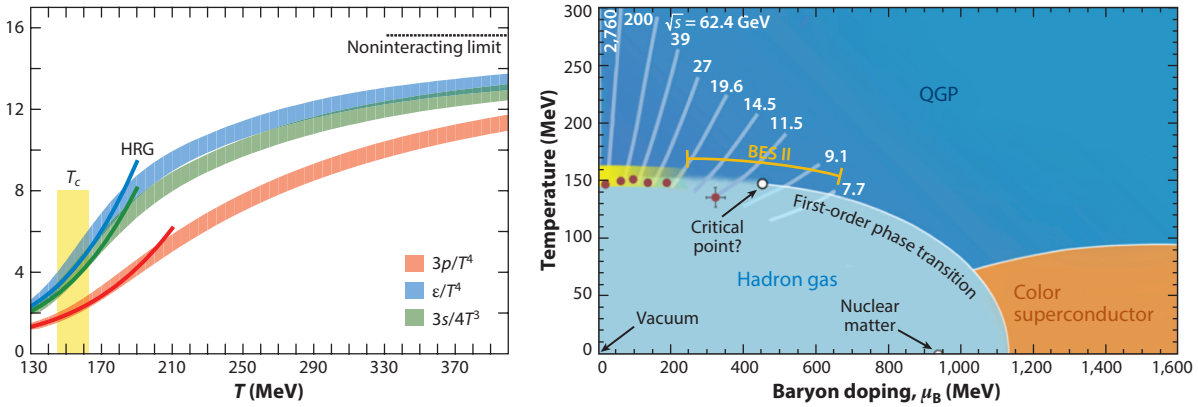


Figure 1.3: Left: Predictions of lattice QCD for the pressure p, energy density ϵ, and entropy density s (normalised to T^4 to make them dimensionless) of hot QCD matter as a function of temperature T (coloured bands). A continuous crossover around $T_c \approx 155$ MeV from a hadron resonance gas (coloured lines) to a quark–gluon plasma is shown. At extremely high temperature, the non-interacting limit (Stefan–Boltzmann values) is asymptotically approached. Right: Sketch of the current understanding of the QCD phase diagram as a function of temperature T and the chemical potential for baryon number μ_B (the excess of quarks over antiquarks). The regions of the phase diagram traversed by the expanding and cooling droplets of QGP formed in heavy-ion collisions with different collision energies are also shown. Figures taken from Refs. [18, 24].

The thermodynamic behaviour of QCD as a function of temperature is actually one of those topics that can be calculated directly with lattice QCD [18]. In the left panel of Fig. 1.3, lQCD predictions for the pressure p, energy density ϵ, and entropy density s of QCD matter in thermal equilibrium at temperature T are shown. The rise with increasing temperature in ϵ/T^4 and s/T^4 indicates the crossover from a hadron resonance gas (HRG) to a new phase. This predicted phase, where colour is deconfined, is what we call the *quark–gluon plasma* (QGP). A sketch of the current understanding of the QCD phase diagram is shown in the right panel of Fig. 1.3. In the limit of $\mu_B = 0$ (baryo-chemical potential, the excess of quarks over antiquarks), the transition is expected around $T_c \approx 155$ MeV, corresponding to an energy density of about $\epsilon \approx 0.7$ GeV/fm^3 [19]. The early universe is believed to have crossed the QGP phase a few microseconds after the Big Bang, before it cooled down to below the critical temperature T_c. When instead the baryo-chemical potential is increased (for $T \approx 0$, as is the case for ordinary matter), a transition to a phase with colour superconducting properties is predicted, hypothesised to exist in the core of neutron stars [20]. Global symmetries of the QCD Lagrangian control the nature of the phase transition, which is believed to be a crossover (as a consequence of chiral symmetry [21, 22]) in the low-μ_B region, while at higher μ_B it is expected to be a first order transition [23]. For this reason, a critical point in the phase diagram is predicted in Fig. 1.3 and looked for by experiments.

1.4 Heavy-ion collisions

Since the 1980s, several heavy-ion experiments took place at particle accelerators at international laboratories, starting from fixed-target experiments operating at centre-of-mass energy

per nucleon-pair ($\sqrt{s_{\rm NN}}$) of the order of 10 GeV (AGS at BNL and SPS at CERN), to collider experiments with top energies of 200 GeV (RHIC at BNL) and 5.5 TeV (LHC at CERN). Each experiment probing a different "starting point" in the QCD phase diagram, with fixed-target experiments in the $\mu_{\rm B}$ interval of 200–500 MeV and collider experiments at their top energies in the region $\mu_{\rm B} \approx 0$. The main reason for this difference is because at low energies the amount of nucleons that are stopped in the collision region is high, increasing the net baryon density. At RHIC, there is the possibility to do a beam-energy scan (BES), which changes the initial T and $\mu_{\rm B}$, with the purpose to identify the order of the phase transition and locate its corresponding critical point (if it exists).

The collision between two ultra-relativistic heavy nuclei is rather complex, so lets first quickly summarise what happens in a typical proton–proton (pp) collision, which is considered as a relatively simple system and the guide for heavy-ion collisions. Colliding protons, where the beam energy is divided over the three valence quarks and the parton sea, interact via colour fields between partons from the projectile and the target. The scattered partons travel outwards, emitting QCD radiation (gluons) that will form a parton shower, until hadronisation takes place. This showering process can be explained in view of string fragmentation models[4], where the colour fields resemble strings. When the connected partons move away from each other, it gets at a certain time energetically more favourable for a new quark–antiquark pair to appear, rather than extending the string further, and therefore splitting the string into two. This continues until the mass of a string piece gets small enough, when the breaking stops and it is identified as a hadron. This is called *string fragmentation* in vacuum, and will be discussed in more detail in Chapter 2. Instead of the scattered partons, a bunch of colour-neutral hadrons will be observed in particle detectors, clustered together in what are called jets.

Despite these few, (semi-)hard scattering processes in pp collisions, the effective energy density is still low given the mean free path of the scattered partons is large compared to the system size. In heavy-ion collisions instead, where the colliding nuclear system is larger and the parton density higher, the situation is much more complex. The different consecutive phases during the space–time evolution of such collision, for the case where a QGP is created or not, are sketched in Fig. 1.4, and the current understanding will be discussed below. Much more can be written on each stage, but the focus here will lie on providing the foundations to understand this work. Interested readers can have a look at Ref. [25].

- **Initial state.** Prior to the collision, the incident nuclei are, due to their ultra-relativistic velocities, two Lorentz-contracted disks in the centre-of-mass frame (which is the same as the laboratory frame at colliders). These disks are mainly composed out of gluons that carry only tiny fractions (x) of the longitudinal momenta of their parent nucleons but relatively large transverse momenta (e.g. $k_\perp \simeq 2$ GeV/c for $x = 10^{-4}$), and whose densities rapidly increase with $1/x$. These high-density gluon systems include a large fraction of the colliding partons and therefore determine the dynamics of the early stages in heavy-ion

[4]This is just one of the interpretations, there are different ideas on the market. More details on these other implementations will be given at a later moment.

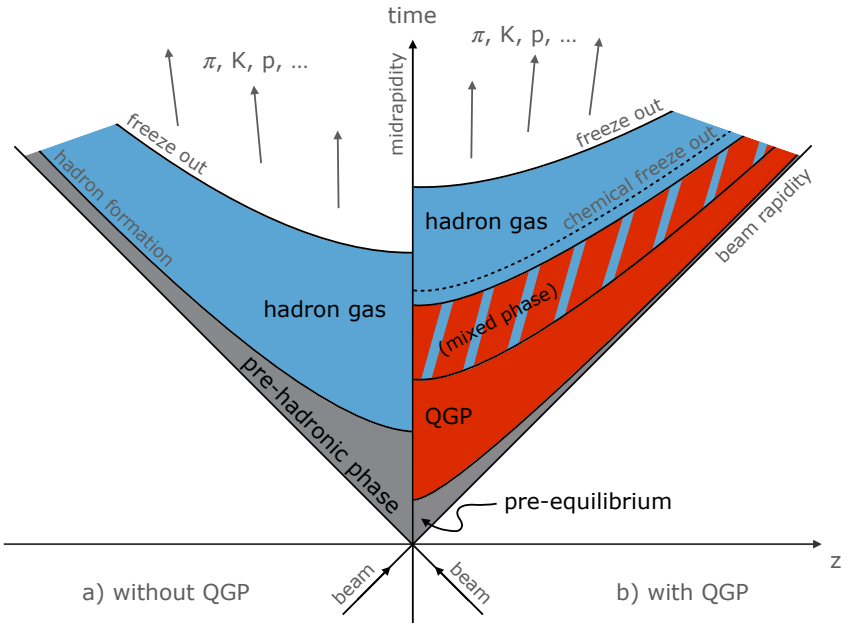


Figure 1.4: A sketch of the evolution of a heavy-ion collision. The left side of the coordinate system ($z < 0$) represents the case where no QGP is created, i.e. when the medium does not reach the critical energy density. The right side ($z > 0$) shows the case where a QGP is created. The shown "mixed phase" only occurs in case of a first-order phase transition, which is currently not believed to take place for most type of ultra-relativistic heavy-ion collisions. See text for more details on the different phases.

collisions. A detailed understanding of these small-x gluons in the nuclei wave function is mandatory to properly characterise the produced QCD medium. For example, bulk observables, used to extract the medium transport coefficients, are as much sensitive to features of the initial stage than to QGP properties itself. See Refs. [26, 27] for a review of the challenges on this gluonic form of matter.

- **Pre-equilibrium.** At $\tau = 0$ the two nuclei hit, which is the moment where the maximum energy density occurs, but this is merely a consequence of Lorentz contraction, and the system is clearly very far from equilibrium. The phase that follows is called pre-equilibrium in which the interactions start developing. The processes involving relatively large transferred momenta, the "hard processes", occur faster ($\Delta\tau \sim 1/Q$, the time during which fluctuations with virtuality Q can survive according to the Heisenberg's uncertainty principle). The products of these first interactions, like jets, direct photons, dilepton pairs, and heavy quarks, are generally the most important probes of the latter phases. Semi-hard processes, with $Q \sim 1$ GeV, and responsible for most of the multiplicity in the final state, occur at $\tau \sim 0.2$ fm/c. At this stage, the initial-state (hard) gluons are liberated and start to radiate softer gluons, which will quickly overwhelm the number of primary hard gluons. A thermal reservoir of soft gluons is formed, which, at first, carries only a small fraction of the total energy, but extracts quickly more energy from the hard gluons via multiple elastic scatterings [28]. The more interactions, the higher the parton density and temperature of the system will get. The dynamics of this non-equilibrium state of partonic matter, see Refs. [27, 29] for more details, is thought to last for about 1 fm/c (although no direct

constraints from data are available), and till recently there were no practical frameworks to connect numerical calculations of it to ones for the next phases [30, 31].

- **QGP.** In the case the interactions between the produced partons would be negligible, they would rapidly separate and evolve independently towards the final-state hadrons (left side of Fig. 1.4). If they instead would interact quite strongly, as is the case for heavy-ion collisions at RHIC and the LHC, they would rapidly approach thermal equilibrium leading to the quark–gluon plasma phase. Given that the partonic matter keeps expanding and cooling down, this thermal equilibrium is only reached locally, i.e. the temperature is space and time dependent. It was shown that, macroscopically, the expansion of the system can be examined with *relativistic fluid dynamics* [32, 33], predicted to have very low specific shear viscosity η/s [34]. This hydrodynamic phase at the LHC lasts for about 10 fm/c, and quantifying the transport coefficients experimentally became an important aspect of heavy-ion physics [35]. The QGP has some more distinct experimental features that will be explained in more detail in Sec. 1.5.

- **Hadronisation.** When the local temperature falls below the critical temperature T_c the crossover transition occurs and hadronisation will take place — the quarks and gluons get confined inside colour-singlet hadrons — in a process called the *chemical freeze out*. There may be additional hadronisation mechanisms possible with respect to hadronisation in vacuum, which is the main topic of this dissertation (for all details, see Chapter 2). The formed hadronic system, a so-called *hot hadron gas*, is still relatively dense and preserves local thermal equilibrium while further expanding, until eventually the system dilutes and the hadrons stop interacting with each other. This is what is called the (kinetic or thermal) *freeze out*, and after this the particles will freely stream until they reach the detector.

Besides the different phases in a heavy-ion collision, the initial collision geometry is also an aspect that needs to be taken into account because of the finite size of the atomic nuclei. These "geometrical" quantities, which are typically expressed as the impact parameter (b), the number of participating nucleons (N_part), and the number of binary nucleon–nucleon collisions (N_coll), can be described with *Glauber model* calculations [36]. The collision between two nuclei is described as an incoherent superposition of binary nucleon–nucleon interactions, considering protons and neutrons point-like and indistinguishable. Furthermore, the model assumes that at sufficiently high energies, these nucleons will not be deflected when the nuclei pass through each other. With these assumptions, called the *optical limit*, a simple analytic description based on just the nuclear density function and the nucleon–nucleon inelastic cross section is possible. Within this framework it can be shown that the cross section for hard processes is proportional to N_coll, where the one for soft processes was believed to be proportional to N_part [36].

The optical limit has the disadvantage that it does not locate nucleons at specific spatial coordinates, but uses continuous density functions instead. On average this is a proper approximation, but when looking event-by-event a more realistic description is required. This emerged with the Glauber Monte Carlo approach used nowadays by realistic particle production codes. In the

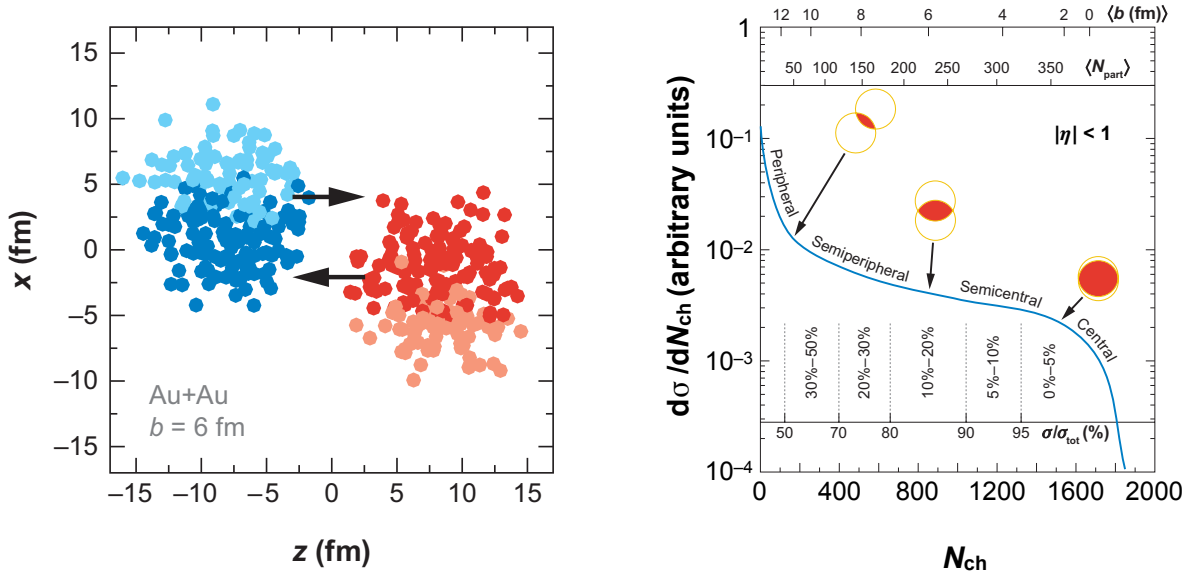


Figure 1.5: Left: Glauber Monte Carlo event in the transverse plane for a Au–Au collision with impact parameter $b = 6$ fm. The nucleons are drawn as circles, where the darker ones are participating. Right: Illustration of the correlation between the final-state observable N_{ch} (number of charged-primary particles) with calculated quantities from the Glauber model (b and N_{part}). Figures taken from Ref. [36].

Monte Carlo ansatz, both the positions of the nucleons and the impact parameter of the collision are randomly generated according to the respective nuclear density and $\mathrm{d}\sigma/\mathrm{d}b \propto 2\pi b$ distribution, respectively. In the left panel of Fig. 1.5 an illustration is shown for a Au–Au collision with impact parameter $b = 6$ fm. In this approach, the N_{part} and N_{coll} values are obtained for each generated event.

Neither N_{part} nor N_{coll} can be directly measured experimentally, however, mean values can be extracted for classes of measured events by mapping a measured distribution to the corresponding distribution obtained from Glauber calculations. This is performed by defining *centrality classes*, defined as the percentile of the hadronic cross section, in both the measured and calculated distributions and connecting the mean values. The mapping procedure is experiment and collision system specific. The ALICE experiment uses for example the VZERO amplitude distribution [37], but $\mathrm{d}N_{\mathrm{evt}}/\mathrm{d}N_{\mathrm{ch}}$, as used for the final-state observable in the illustration in the right panel of Fig. 1.5, is possible as well. This procedure will be discussed in more detail in Sec. 3.3.

1.5 Experimental observables

In this chapter, we slowly moved away from pure theory and started discussing the rationale behind experiments more often. This is in line with how the field is currently evolving: it is a science driven by experiments[5]. This last section of this chapter is therefore dedicated to

[5] I might be biased here as an experimental physicist.

some of the main experimental probes and signatures of the QGP and the current theoretical understanding of these observations.

The experimental probes of the QGP can essentially be divided in two classes: measurements of particle yields and measurements of particle correlations. Yields are commonly measured with respect to kinematic variables like the transverse momentum (p_{T}) or pseudorapidity (η) of an identified particle[6], or with respect to event characteristics like the charged hadron multiplicity (N_{ch}). Bare yields provide information on the cross sections of the underlying production process [38], where ratios of yields can be more exploratory into specific physical mechanisms. For example, ratios of yields between nucleus–nucleus and proton–proton collisions can probe in-medium energy loss mechanisms [39], where ratios of yields between different hadron species are sensitive to hadronisation processes [40, 41]. Alternatively, particle correlations can be studied, which are usually calculated between kinematic variables of specific identified particles and/or between all produced particles in an event. For example, correlations in the azimuthal angle provide information on anisotropic flow [42], while momentum correlations are instead sensitive to the spatial evolution of the system [43] or in-medium energy loss mechanisms [44]. In the next few subsections, the topics that are relevant for this work will be discussed in more detail.

1.5.1 Particle multiplicity

The most basic quantity, the total amount of particles produced in a collision, provides already important information about the properties of the created medium, e.g. the collision geometry and the energy density produced. The first option we have seen above when discussing the Glauber models, the latter can be shown with the Bjorken hydrodynamic picture for nucleus–nucleus collisions [45]. The Bjorken energy density can be estimated with

$$\epsilon_{\mathrm{BJ}} = \frac{\langle m_{\mathrm{T}} \rangle}{\tau_0 \cdot A} \left. \frac{\mathrm{d}N(\tau_0)}{\mathrm{d}y} \right|_{y=0}, \tag{1.1}$$

where A is the transverse overlap area of the collision region, τ_0 the thermalisation time for the QGP (usually estimated as 1 fm/c), m_{T} the transverse mass defined as $m_{\mathrm{T}} = (m^2 + p_{\mathrm{T}}^2)^{1/2}$, and $\mathrm{d}N/\mathrm{d}y$ the particle multiplicity per unit of rapidity[7]. Of course, given the complex sequence of stages that precedes particle production, this is a simplification, but it provides a good estimate of the energy density of the medium via the measured charged-particle pseudorapidity density at midrapidity.

A collection of charged-particle multiplicity measurements in different collision systems is presented in the left panel of Fig. 1.6 [46]. From the single measurements, the energy density can be estimated via Eq. 1.1, which is found to be ~ 14 GeV/fm^3 at the LHC [47] and about a factor 3 smaller at RHIC [48], both well above the critical value obtained in Sec. 1.3. The averaged

[6] The transverse momentum is defined as $p_{\mathrm{T}} = (p_{\mathrm{x}}^2 + p_{\mathrm{y}}^2)^{1/2}$ and the pseudorapidity as $\eta = -\ln(\tan(\theta/2))$, where θ is the polar angle. The pseudorapidity is an approximation of the rapidity $y = 1/2 \ln((E - p_{\mathrm{z}})/(E + p_{\mathrm{z}}))$ when $p_{\mathrm{T}} \gg m$.

[7] In experiments we do not measure neutral particles, but we can compensate for this using $\mathrm{d}N/\mathrm{d}y \approx \frac{3}{2}\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}y$.

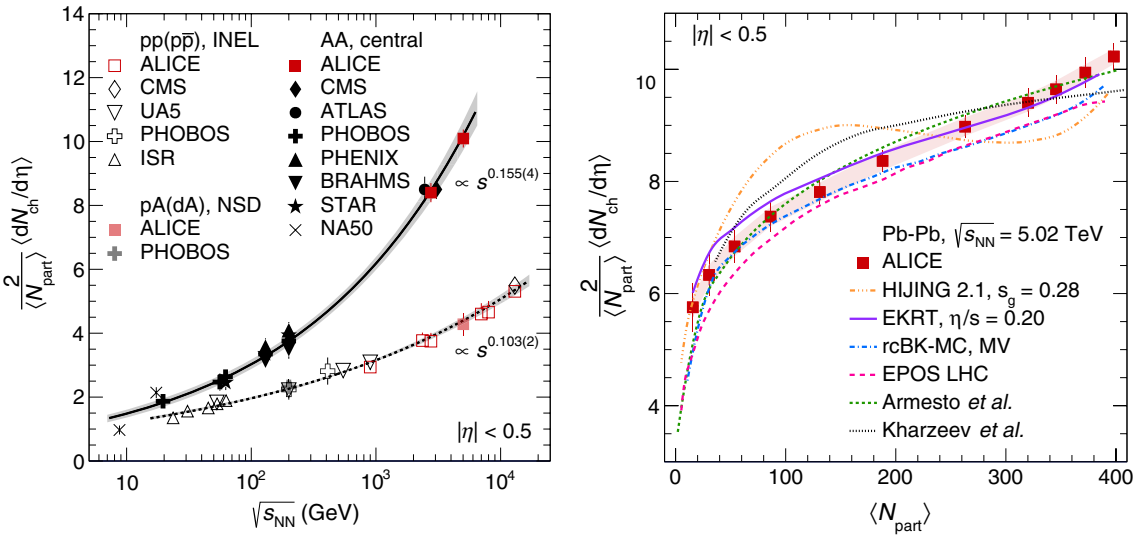


Figure 1.6: Left: A collection of measurements of the charged-particle pseudorapidity density ($\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$) for different collisions as function of centre-of-mass energy ($\sqrt{s_{\mathrm{NN}}}$). Right: The $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ as function of N_{part} for Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV compared to different model predictions. The division of $2/\langle N_{\mathrm{part}} \rangle$ in both figures is introduced to compare different colliding systems. See Ref. [46] and references therein.

charged-particle pseudorapidity density, $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$, shows a smooth trend as function of $\sqrt{s_{\mathrm{NN}}}$ for different collision systems. The rise of $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ is observed to be steeper for nucleus–nucleus collisions with respect to proton–proton and proton–nucleus collisions. The fact that the measurements in p–A and d–A systems can be parametrised with the same power-law function as for pp systems, imply that the observed difference for nucleus–nucleus collisions is not only related to the multiple interactions between the participating nucleons since the proton in p–A collisions also encounters multiple nucleons.

The right panel of Fig. 1.6 shows the centrality dependence, via N_{part} as discussed above, of $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ for Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV compared to model predictions [46]. The charged-particle pseudorapidity density rises steeply for peripheral collisions, but for $N_{\mathrm{part}} > 50$ the slope is decreased. Model calculations that try to describe this observation fall in two broad categories: models combining pQCD processes with soft interactions or models with different implementations to tame the gluon density in the initial stage. The latter ones do in particular a good job describing the measurement, suggesting there might be a gluon saturation domain that limits the number of produced partons and, in turn, of particles [26, 27]. Another implication of these measurements, which is not yet fully understood, is that the energy and centrality dependence are surprisingly independent of each other [46–48]. This may indicate that the naively expected increase with energy resulting from the increase of hard relative to soft processes does not play a leading role in determining the particle multiplicity.

1.5.2 Identified hadron spectroscopy

Besides the overall number of particles produced, measurements of identified hadron yields provide important information on the chemical composition of the aforementioned hadron gas.

By assuming that the system is in thermal and chemical equilibrium at the chemical freeze out, we can see the system as a *grand-canonical ensemble*, and make use of the fact we can describe this in the language of statistical mechanics. In this way, the number density of each state (n_i) is given by a derivative of the partition function

$$n_i = N/V = -\frac{T}{V}\frac{\partial \ln Z_i}{\partial \mu} = \frac{g_i}{2\pi^2}\int_0^\infty \frac{p^2 \mathrm{d}p}{\mathrm{e}^{(E_i-\mu_i)/T} \pm 1}. \tag{1.2}$$

Here, g_i an μ_i are the degrees of freedom and the chemical potential of species i, respectively, which ensure quantum number conservation. The main point here is, if this holds, the critical temperature of the chemical freeze out and volume of the system can be determined from a single fit to the measured abundances of different states. Extra phenomenological factors, implementing an ad hoc suppression of hadrons with strange (γ_s) or charm (γ_c) valence quarks, were added in some versions to better describe the data. These suppression factors account for an incomplete thermalisation of the strange and charm quarks and are typically needed in peripheral A–A, p–A, and pp collisions.

Figure 1.7 shows an example of such a fit to p_T-integrated yields measurements by the ALICE Collaboration in central Pb–Pb collisions for three different *statistical hadronisation* models [49]. The models describe remarkably well the data, that span over nine orders of magnitude. The extracted freeze-out temperature is found to be (156 ± 2) MeV, where the volume is of the order of 5000 fm^3. The fact that the fits work so well support the picture that there is a chemical equilibrium over a large volume, implying the system is thermalised.

The primary hadron abundances will be frozen after the chemical freeze out, but the particles will still elastically scatter among themselves. Only after the kinematic freeze out, when the system becomes too dilute for interactions to occur, the momentum spectra get fixed. In case thermal equilibrium is assumed, the low p_T spectra can be described by a Maxwell–Boltzmann distribution

$$\frac{1}{p_\mathrm{T}}\frac{\mathrm{d}N}{\mathrm{d}p_\mathrm{T}} \propto \mathrm{e}^{-\sqrt{m^2+p_\mathrm{T}^2}/T}. \tag{1.3}$$

The parameter T represents here the kinematic freeze-out temperature, and should be similar for different particle species. In nucleus–nucleus collisions, however, a mass dependency in the slope is observed due to a collective motion in the transverse plane caused by the internal pressure in the QGP [50]. This is called *radial flow*, and its term $\frac{m}{2}\langle\beta_T\rangle^2$, where $\langle\beta_\mathrm{T}\rangle$ is the average radial-expansion velocity of the medium, should be superimposed to the thermal random motion. The fact that the system is expanding, usually described by hydrodynamic models, introduces an additional complication. Blast-wave models [38] were developed to estimate the kinematic freeze-out temperature (and $\langle\beta_\mathrm{T}\rangle$) via fits to the measured p_T-differential spectra distributions of different hadron species. For Pb–Pb collisions at $\sqrt{s_\mathrm{NN}} = 5.02$ TeV, $\langle\beta_\mathrm{T}\rangle$ was found to increase with centrality, reaching ~ 0.66 for the 5% most central collisions, while T_kin decreases from (161 ± 6) to (90 ± 3) MeV [50][8]. Note that these blast-wave models do not replace full

[8]For peripheral collisions $T_\mathrm{chem} \sim T_\mathrm{kin}$, which means that the hadronic gas phase is very short, although the estimation of this kinematic freeze-out temperature may be affected by event-selection and geometry biases [51].

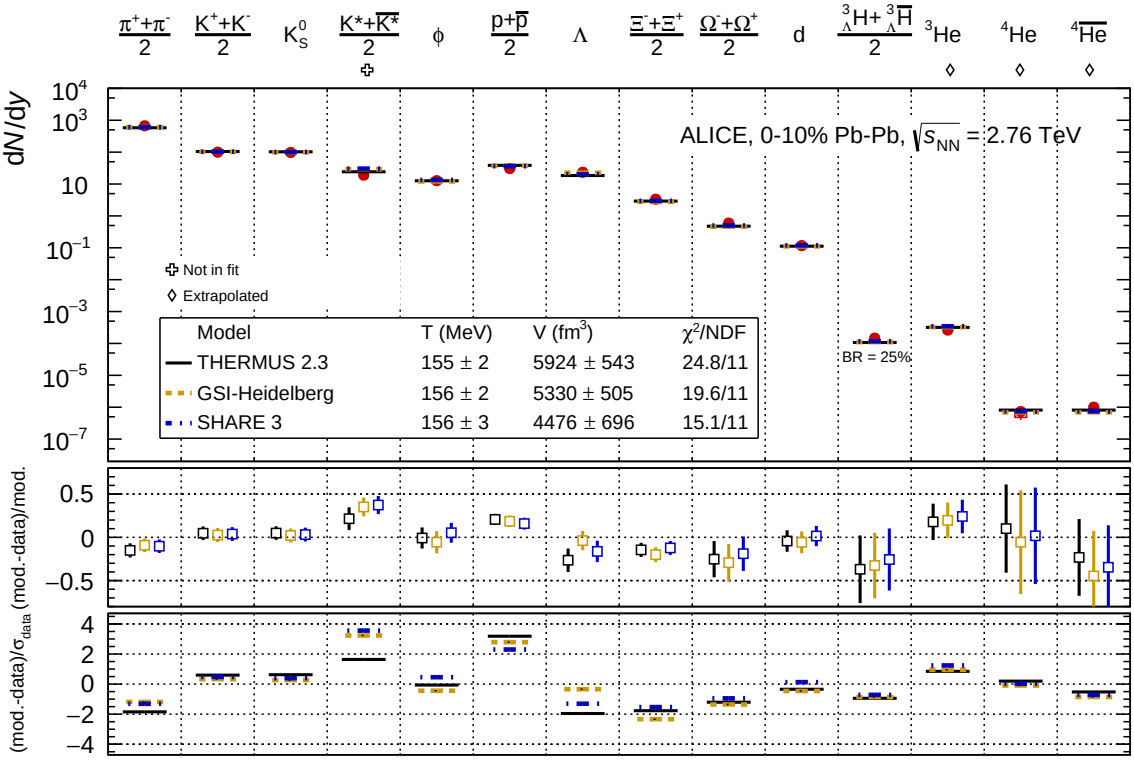


Figure 1.7: Three different thermal model fits to light-flavour hadron yields in central (0–10%) Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV as measured by the ALICE Collaboration [49]. The upper panel shows the fit results together with the data, whereas the lower two panels show the difference between model and data normalised to the model value and to the experimental uncertainties, respectively.

hydrodynamical calculations, but simply allow for a fast comparison of the hydrodynamical parameters of different systems via the measured $p_{\rm T}$-spectra.

So far we have seen that identified hadron spectroscopy provides a tool for studying the collision dynamics beyond that of global event characterisation. Additionally, the relative yields between different hadron species may give insights into specific physical mechanisms. For example, at RHIC, the study of $p_{\rm T}$-differential baryon-to-meson ratios provided new puzzles in the early 2000's [52, 53]. Generally, the amount of suppression experienced due to interactions with the produced colour deconfined medium seemed to be different for baryons and mesons. Theoretical attempts to describe these measurements lead to the idea[9] that recombination (or coalescence) of medium constituents is the dominant hadronisation mechanism at low $p_{\rm T}$ [55, 56]. This will be discussed in much more details in Chapter 2.

Another interesting ratio is the one between (multi-)strange and non-strange hadrons, which can be used to observe the predicted strangeness enhancement in the QGP [57]. An increased production of hadrons containing strange quarks was proposed as a signature of deconfinement, because of the partial restoration of the chiral symmetry when the phase transition occurred.

[9]To be precise, the quark recombination model was first invoked to explain another aspect of RHIC data: the observed flavour pattern in the elliptic flow [54]. A few months later it was used as well to describe the baryon-to-meson anomaly.

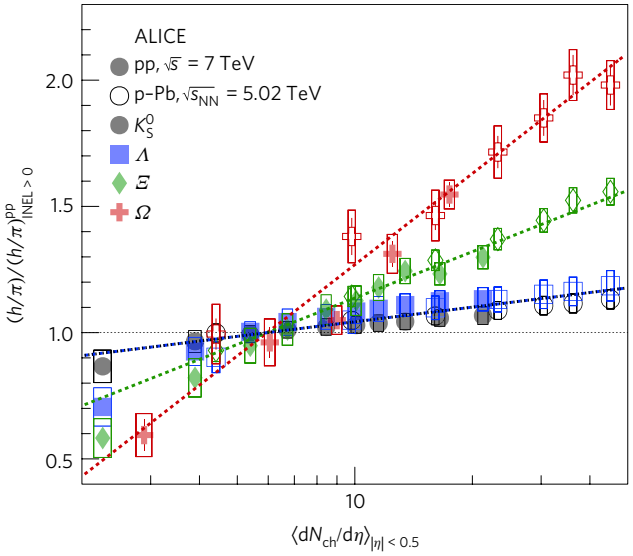


Figure 1.8: The p_T-integrated ratios of (multi-)strange hadrons to pions to the values measured in the inclusive pp sample as function of multiplicity as measured by the ALICE Collaboration [8]. The observed enhancement with multiplicity increases with strangeness content.

Due to the restoration, the mass of the strange quark decreases from the constituent to the current value of about 100 MeV/c^2 [16], hence an abundant production of strange quark–antiquark pairs is expected because of the mass being smaller than the temperature. This idea was confirmed in the 1990s by the NA35, WA97, and NA57 Collaborations [58–60], and observed as well in heavy-ion collisions at RHIC [61] and LHC [62]. Recently, and unexpectedly, this observation was "extended" to high-multiplicity pp collisions [8]. A smooth increase in the ratio of the p_T-integrated yields of (multi-)strange hadrons to those of charged pions was observed from low-multiplicity pp collisions to central Pb–Pb collisions. The observed enhancement increases with strangeness content rather than with mass or baryon number, as shown in Fig. 1.8, and could not be reproduced by any of the commonly used theoretical models. It may point towards a common underlying physics mechanism which gradually compensates the strangeness suppression, i.e. small collision systems exhibiting characteristic features understood to be connected to the formation of a QGP.

1.5.3 In-medium energy loss

Just like a charged particle that loses energy while traversing normal matter, a particle with colour charge should lose energy as it passes through the coloured QGP medium. If one can describe these in-medium energy loss mechanisms theoretically, then it should be possible to extract properties of the medium, like the colour charge density, from experimental measurements. Hard probes, i.e. processes with large momentum exchange that occur before the QGP is formed, are in this sense interesting. Such processes, whether they occur in elementary or nucleus–nucleus collisions, will lead to productions of jets, a collimated spray of hadrons from the fragmentation of a high-p_T parton. In case a QGP is formed these jets will be "quenched", as predicted by Bjorken in the early 1980s [63]. Jet quenching was first observed at RHIC [39],

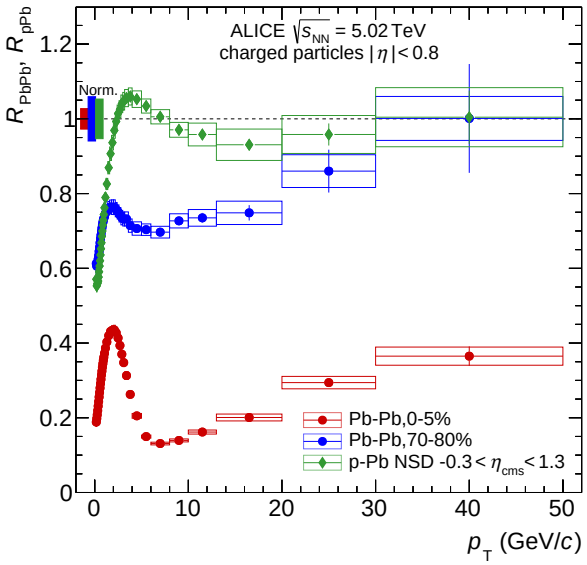


Figure 1.9: The nuclear modification factors for charged particles as measured by ALICE in central (0–5%) and peripheral (70–80%) Pb–Pb collisions and in minimum-bias p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [64].

quantified using the so-called nuclear modification factor

$$R_{\mathrm{AA}}(p_{\mathrm{T}}, y) \equiv \frac{1}{\langle T_{\mathrm{AA}} \rangle} \frac{\mathrm{d}^2 N_{\mathrm{AA}}/\mathrm{d}p_{\mathrm{T}}\mathrm{d}y}{\mathrm{d}^2 \sigma_{\mathrm{pp}}/\mathrm{d}p_{\mathrm{T}}\mathrm{d}y} = \frac{1}{\langle N_{\mathrm{coll}} \rangle} \frac{\mathrm{d}^2 N_{\mathrm{AA}}/\mathrm{d}p_{\mathrm{T}}\mathrm{d}y}{\mathrm{d}^2 N_{\mathrm{pp}}/\mathrm{d}p_{\mathrm{T}}\mathrm{d}y}, \tag{1.4}$$

where $\langle T_{\mathrm{AA}} \rangle$ is the nuclear overlap function, defined as the average number of nucleon–nucleon collisions $\langle N_{\mathrm{coll}} \rangle$, divided by the inelastic nucleon–nucleon cross section [36]. A R_{AA} equal to unity is expected in absence of medium effects, while a difference from unity implies modifications of the p_{T} distribution of the produced hadrons. This can be due to in-medium energy loss via inelastic (radiative energy loss) and elastic (collisional energy loss) scattering processes and the dynamics of the hadronisation process in presence of a QGP, but also other processes not related to the presence of the QGP, so-called *cold-nuclear matter* (CNM) effects, can cause a deviation from unity. These CNM effects can be assessed by studying the nuclear modification of hadrons in p–A collisions.

In Fig. 1.9, the nuclear modification factor of charged particles as a function of p_{T} as measured by the ALICE Collaboration [64] in central (0–5%) and peripheral (70–80%) Pb–Pb collisions and in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV is shown. For high-p_{T} particles ($p_{\mathrm{T}} \gtrsim 6$ GeV/c), the nuclear modification factor is compatible with unity for p–Pb collisions, indicating no visible effects in the R_{AA} related to the initial state are expected. The maximum at intermediate p_{T} ($2 < p_{\mathrm{T}} < 6$ GeV/c) is related to the so-called Cronin effect [65]. The nuclear modification factor in central Pb–Pb collisions is significantly suppressed, where for peripheral collisions, the suppression is smaller and approaches unity for the highest p_{T} intervals. The measurement of the R_{AA} is therefore a sensitive probe of medium-induced effects, and comparisons to theoretical calculations can be used to extract transport properties of the medium [66].

1.6 Discussion

In this chapter, I gave an overview on the theoretical background of high-energy nuclear physics, the research field devoted to the theory of the strong interaction, quantum chromodynamics. Despite many of the young researchers seeing this field as the study of the quark–gluon plasma, this is only one of the many phases in which QCD can express itself. For a complete understanding of QCD — or more delimited, to understand how QCD behaves in a heavy-ion collision — a much broader and multi-disciplinary knowledge is required.

I do not deem myself an expert on all the theoretical aspects that were discussed above, and for that reason I tried to refer to more detailed reviews in the text. Actually, I think there are only very few people that can grasp everything we know at this moment about QCD. On the experimental side, there are also many more observations that were not and will not be discussed in this dissertation. Given that they are nevertheless interesting, let me list here three topics together with a recent publication: anisotropic flow [67], the melting of quarkonia [68], and direct photon production [69].

The field of ultra-relativistic heavy-ion physics has seen an unprecedented level of progress in the last few decades, however, much remains to be discovered. In Refs. [25, 70] a nice overview of the big questions is presented. I finish this chapter repeating the few questions where this work tries to contribute:

(i) **Is there a QGP formed in small systems?** This question highlights a "clash" between the approaches of the heavy-ion and high-energy physics (HIP and HEP) communities on how to understand soft multi-particle production in hadronic collisions. Where heavy-ion physics pursues a collective hydrodynamical picture, the standard assumption in HEP is that the interactions between physical degrees of freedom are negligible. Both views can explain many of the observed characteristics for Pb–Pb and pp collisions, respectively. However, the scope of heavy-ion physics goes beyond hot nuclear matter, and smaller systems like pp — where thermalisation was initially not assumed — are actively studied as well. Surprisingly, many similarities with traditional heavy-ion physics were highlighted [8–13], forcing the authors of one of the most actively used general-purpose event generators in the HEP community to conclude that there is "*more physics at play than traditionally assumed*" [71]. If this means a small droplet of QGP is produced in pp collisions (the "HIP view") or if the similarities with a hot, deconfined medium can instead be explained by introducing short-range colour correlations (the "HEP view"), is currently under debate.

(ii) **What is the level of interplay among the quark flavours in the hadronisation phase?** It is clear by now that hadronic modelling without taking into account the parton phase-space density, like vacuum fragmentation as it was developed studying e^+e^- collision data, is no longer valid in presence of any reservoir of partons. I wrote here "any", since (just like for the first question) it is an important point for both HIP and HEP. The QGP phase is an extreme case, but also the outgoing parton showers in pp events can undergo secondary interactions. There have been many theoretical attempts introducing modified

hadronisation mechanisms to describe the measurements, ranging from statistical hadronisation, to coalescence mechanisms or reconnections between colour charges, suggesting significant differences with the picture of a freely-streaming parton while fragmenting. It is to be seen which of these ideas are appropriate, and to which extent.

(iii) **What is the level of chemical and/or kinetic thermalisation achieved by the hadron populations?** In the last couple of decades there has been an unreasonable success of the statistical hadronisation model, which is able to describe dozens of light-flavour hadron production ratios for many collision systems and collision energies with basically one free parameter. We do not know yet *why* or *how* it works this well, although for light-flavour hadrons it can be imagined that the light quarks "evolve into" or were "born into" thermal equilibrium when assuming a hot-QCD medium. This means, however, that the underlying dynamics of the hadronisation can no longer be retrieved for such particles. To accomplish a true understanding of what is happening in the hadronisation phase, it is important to study particle species which are partly or not at all thermalised, i.e. disagree with predictions from the statistical hadronisation model.

At the end of the next chapter, after the required background on open heavy-flavour physics has been discussed, we will connect the aim of this dissertation with these big open questions in the field.

CHAPTER 2

Open heavy-flavour physics

Open heavy-flavour hadrons, i.e. particles containing at least a charm or beauty quark and other lighter quarks, are among the most important tools for the study of QCD in high-energy hadronic collisions. In proton–proton collisions, heavy-flavour production provides important tests of perturbative QCD calculations down to low transverse momenta. Meanwhile, heavy-flavour production in nucleus–nucleus collisions represents a sensitive probe of the hot and dense QCD matter. The large masses of heavy quarks, $m_{\mathrm{c}} \simeq 1.3$ GeV/c^2 and $m_{\mathrm{b}} \simeq 4.2$ GeV/c^2 [16], together with flavour conservation in strong interactions, is essentially what makes heavy-flavour physics appealing for the study of quantum chromodynamics. The consequences of these two points will be discussed in detail in this chapter, where the main features are, in addition, also summarised in the "Important implications" boxes.

The sections in this chapter will follow the various stages during the evolution of a collision, starting from the production of heavy quarks, going through the interactions with the cold- and hot-nuclear medium, and finishing with the hadronisation process. We will conclude this chapter by connecting open heavy-flavour physics, especially the physics behind the analyses performed for this dissertation, with a few of the big open questions in the field, as listed in Chapter 1.

2.1 Open heavy-flavour production

The strong interaction, which conserves flavour charge, requires heavy quarks (Q) to be produced in quark–antiquark pairs. The leading-order (LO) production processes ($\mathcal{O}(\alpha_{\mathrm{s}}^2)$) are quark–antiquark annihilation, $\mathrm{q\overline{q}} \rightarrow \mathrm{Q\overline{Q}}$, and gluon fusion, $\mathrm{gg} \rightarrow \mathrm{Q\overline{Q}}$ [72]. At next-to-leading order (NLO), which corresponds to $\mathcal{O}(\alpha_{\mathrm{s}}^3)$, flavour-excitation and gluon-splitting processes contribute to the production as well, both represented by $\mathrm{gq} \rightarrow \mathrm{Q\overline{Q}q}$ or $\mathrm{gg} \rightarrow \mathrm{Q\overline{Q}g}$. The minimum virtuality of these scatterings is $Q^2 \gtrsim 4m_{\mathrm{b,c}}^2$, hence essentially restricted to the hard scatterings in the initial stages of the collision following the argument made in Sec. 1.4.

When considering nucleus–nucleus collisions, the production time of charm quarks ($\tau \sim 1/Q \lesssim 1/2m_{\mathrm{c}} \approx 0.1$ fm/c) is smaller than the formation time of a possible QCD medium, so the production should be unaffected by the properties of the medium. The NLO gluon-splitting production process, where the quark and antiquark are produced with a small relative opening angle, is an exception to this picture. In this case, the production time is increased by a boost factor $E_{\mathrm{g}}/(2m_{\mathrm{c}}) \sim E_{\mathrm{c}}/m_{\mathrm{c}}$, and therefore the charm production time will be of the order of $E_{\mathrm{c}}/(2m_{\mathrm{c}}^2)$. This means that charm quarks with $E_{\mathrm{c}} \gtrsim 15$ GeV from gluon splitting will have similar formation times as the QGP, and their production can therefore be affected by the medium properties. The fraction of charm quarks produced via gluon splitting, estimated with pQCD calculations [73] and with Monte Carlo simulations [74], was found to be moderate though (10–20%). For beauty quarks, large fractions of about 50% are predicted [75], but the heavier quark mass pushes the relevance of this effect to higher energies. Another exception would be secondary (thermal) production from processes like $\mathrm{gg} \rightarrow \mathrm{c\overline{c}}$ in the QGP, which, however, can be considered negligible unless initial QGP temperatures much larger than currently experimentally reachable are assumed [76]. It is therefore reasonable to assume that heavy-flavour hadrons up

to $p_{\mathrm{T}} \sim 50$ GeV/c probe the full medium evolution.

Important implications

Due to their large masses, heavy-flavour production is restricted to the initial stages of the collision, well before the production of a possible QGP.

Because heavy quarks are produced in hard scatterings with large momentum transfer Q, for which the strong coupling constant is significantly smaller than unity (see Fig. 1.2), the elementary cross section of the production processes $\sigma_{\mathrm{ab}\to\mathrm{Q}\overline{\mathrm{Q}}}$ can be calculated down to low p_{T} in the framework of perturbative QCD. These calculations are "universal", i.e. representing some particular kinematic configurations, and can therefore be tried for light partons as well. There are, however, two main advantages of working with heavy quarks. First of all, the contribution of gluon emission processes from the final-state parton need to be properly included in the calculation, which for light flavours will lead to a singularity in case of collinear emission. For heavy quarks, instead, the mass will act as a cut-off for these collinear singularities, which makes the calculation possible, even at low p_{T} [72]. Another issue emerges with the NLO gluon-splitting processes, which will give rise to potentially large logarithms, $N(\mathrm{Q}\overline{\mathrm{Q}}) \sim \alpha_{\mathrm{s}}/(6\pi)\log(p_{\mathrm{T}}^2/m^2)$. For light quarks, where $m \sim 0$, this calculation will not be possible since the perturbation series is no longer an expansion in a small parameter, leading again to a collinear divergence. Because of the large values of m_{Q}, these logarithmic terms at NLO do not prevent a perturbative approach, although it might bias NLO calculations at large p_{T} [72].

Important implications

The heavy-quark mass acts as a cut-off for initial- and final-state collinear singularities, making perturbative calculations of the production processes possible down to low p_{T}.

From the elementary heavy quark–antiquark cross section calculation, the production of open heavy-flavour hadrons (h_{Q}) in pp collisions can be obtained exploiting the (collinear) *factorisation theorem* [77]

$$\mathrm{d}\sigma_{h_{\mathrm{Q}}+\mathrm{X}} \simeq \sum_{a,b} \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \int_0^1 \mathrm{d}z\, f_{\mathrm{a}}(x_{\mathrm{a}};\mu_{\mathrm{F}}^2) \otimes f_{\mathrm{b}}(x_{\mathrm{b}};\mu_{\mathrm{F}}^2) \otimes \mathrm{d}\tilde{\sigma}_{\mathrm{ab}\to\mathrm{Q}\overline{\mathrm{Q}}}(x_a, x_b;\mu_{\mathrm{F}}^2,\mu_{\mathrm{R}}^2) \otimes D_{\mathrm{Q}\to h_{\mathrm{Q}}}(z). \tag{2.1}$$

Here, the parton distribution functions (PDFs) are depicted by $f_{\mathrm{a}}(x_{\mathrm{a}};\mu_{\mathrm{F}}^2)$, which describe the probability of finding a parton of flavour a carrying a fraction x_{a} of the momentum in the colliding proton. The second term is the partonic cross section of the $\mathrm{Q}\overline{\mathrm{Q}}$ production process. The dividing line between these two ingredients is drawn at an arbitrary[1] scale, μ_{F}, called the *factorisation scale*. The partonic cross section also depends on the renormalisation scale μ_{R} at which the strong coupling constant is evaluated. Both scales are usually determined

[1] This should be read as "user-defined". We have the right to choose μ_{F} as we wish because the final result will be independent of μ_{F} if the calculation were done exactly.

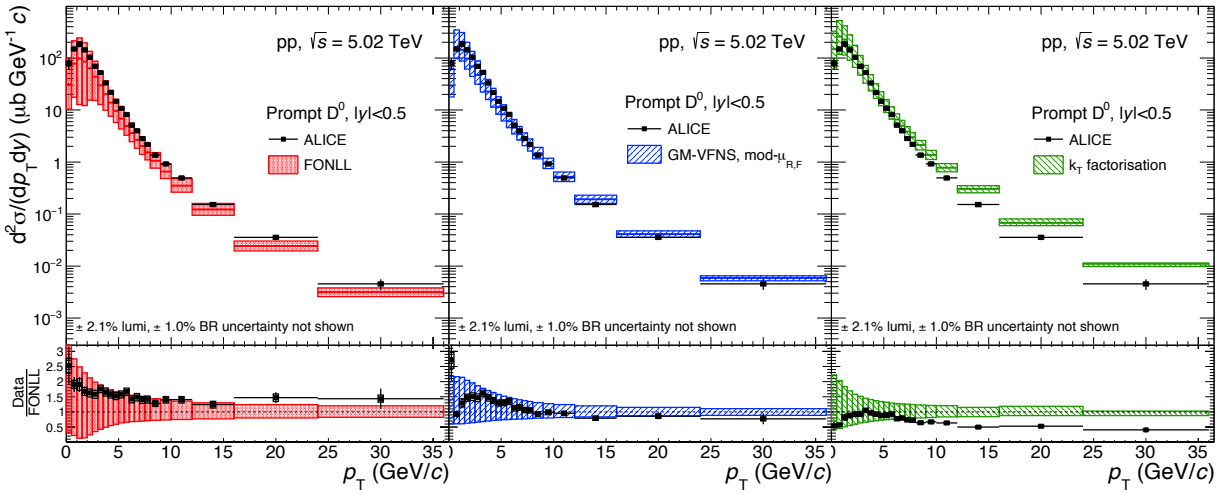


Figure 2.1: The p_{T}-differential production cross section for prompt D^0 in $\sqrt{s} = 5.02$ TeV pp collisions as measured by the ALICE Collaboration [82]. Comparisons to predictions from three pQCD calculations (from left to right: FONLL [83], GM-VFNS [84], and k_{T}-factorisation [85]) are shown, which describe the measurements within their theoretical uncertainties. As shown in Ref. [86], such pQCD calculations provide a good description of the measurements of beauty mesons as well.

by the hard scale of the process $\mu_{\mathrm{R}} \sim \mu_{\mathrm{F}} \sim \sqrt{m_{\mathrm{Q}}^2 + p_{\mathrm{T,Q}}^2}$, and varied around this nominal value in order to estimate theoretical uncertainties. The last ingredient, $D_{\mathrm{Q}\to h_{\mathrm{Q}}}(z)$, is the fragmentation function (FF) that describes the probability of a heavy quark Q hadronising in the hadron h_{Q} with momentum $p_h = zp_{\mathrm{Q}}$. The fragmentation fractions are usually taken to be scale-independent, although scale-dependent ones can be considered as well [78]. Only the elementary cross section can be calculated with perturbative QCD, the PDFs and FFs describe both non-perturbative processes, and are typically parametrised from measurements. The parton distribution functions are based on deep inelastic scattering experiments (at a certain, arbitrary, scale μ_0) and evolved to any other μ_{F} scale using the so-called DGLAP equations [79–81]. The fragmentation functions are usually taken from measurements in $\mathrm{e}^+\mathrm{e}^-$ collisions assuming universality among different collision systems, although, as we will see later, recent measurements are questioning this assumption.

The state-of-the-art pQCD calculations exploiting the collinear factorisation theorem, valid in the entire kinematic range ($p_{\mathrm{T}} \ll m_{\mathrm{Q}}$, $p_{\mathrm{T}} \simeq m_{\mathrm{Q}}$, $p_{\mathrm{T}} \gg m_{\mathrm{Q}}$), are performed with the FONLL (Fixed-Order-Next-to-Leading-Log) [87, 88] or GM-VFNS (General Mass Variable Flavour Number Scheme) [89, 90] frameworks. FONLL calculations adopt a NLO approach for low p_{T} matched to next-to-leading-log (NLL) calculations for high p_{T}. A similar approach is used in the GM-VFNS framework, although the logarithmic correction terms are treated differently. Both predictions are compared to the p_{T}-differential cross section for D^0 mesons in pp collisions at $\sqrt{s} = 5.02$ TeV [82] in Fig. 2.1. They are compatible with the measurements within their theoretical uncertainties — which are dominant over the experimental ones — which come mainly from the choice of the scales in the perturbative calculation. In case of production ratios, e.g. between different energies, rapidity regions, or hadron species, the theoretical uncertainty due to the choice of the factorisation and renormalisation scales become subdominant [91], making the

measurement more sensitive to the PDFs and/or FFs.

An effective and efficient alternative to collinear factorisation is the so-called k_{T}-factorisation approach [92, 93], which performs the calculations at LO taking into account the transverse momenta (k_{T}) of both partons that enter the hard process. This explicit treatment of the k_{T} of the incident partons implies that the sum of the $\mathrm{Q\overline{Q}}$ transverse momenta no longer cancels, making it feasible to also study kinematic correlations. The initial k_{T}, together with gluon-emission processes to effectively account for higher-order corrections, are encoded in so-called unintegrated gluon distribution functions, for which (in contrary to collinear PDFs) different theoretical approaches differ considerably [94]. A comparison of k_{T}-factorisation predictions with the measured production of D^0 mesons at $\sqrt{s} = 5.02$ TeV is shown in Fig. 2.1 as well. It describes the data at low and intermediate p_{T} within the theoretical uncertainties, that are also estimated by varying the relevant scales.

Besides the (mostly) analytic pQCD calculations discussed above, general-purpose Monte-Carlo generators like PYTHIA [95] and HERWIG [96] can also be used for predictions of heavy-flavour production. They allow for a more complete description of the final state, including initial- and final-state parton showers, hadronisation processes, hadron decays, and detector response, although with the disadvantage that they have only LO+LL accuracy. This accuracy, together with the fact that much tuning on data is involved and the predictions do not come (naturally) with a theoretical uncertainty, makes them less realistic references to compare to. Monte Carlo generators with NLO accuracy in the hard scattering like POWHEG [97] and MC@NLO [98] are available as well, which can be matched to the parton shower of PYTHIA and HERWIG for a complete modelling. Comparisons of POWHEG+PYTHIA calculations to measured D-meson spectra (and collinear factorisation calculations) can be found in Ref. [99]. Overall, the predictions agree within the theoretical uncertainty band, which are again dominated by the variations of the factorisation and renormalisation scales, although the central calculations tend to underestimate the data.

All statements on data-theory comparisons above were, intentionally, only made for heavy-flavour mesons. This is because the first measurements of heavy-flavour baryon production at the LHC [100–109] were puzzling[2]: all theoretical predictions that describe well the meson sector significantly underestimate the baryon-production measurements. An example is presented in Fig. 2.2, where the prompt Λ_{c}^{+}-baryon production cross section in pp collisions at $\sqrt{s} = 5.02$ TeV as measured by the ALICE Collaboration [102, 103] is compared to GM-VFNS [84], PYTHIA 8 [95], and POWHEG [97] predictions, which all significantly underestimate the data. This attracted much theoretical attention in the last years [85, 110–114], and is very relevant for this dissertation as well. However, before we can discuss this in more detail, we first need to go through the next few sections. We will come back to this topic in Sec. 2.4 when discussing hadronisation.

[2]In literature there are two types of Λ_{c}^{+} "puzzles": the data-versus-model comparisons and the ALICE-versus-LHCb at $\sqrt{s} = 7$ TeV disagreement. We will not discuss the latter one here, but the general consensus in the field is that the measurement of LHCb needs a revision. One argument for this is the "unphysical-looking" $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^0$ ratio versus rapidity, see for example Fig. 10 of Ref. [100].

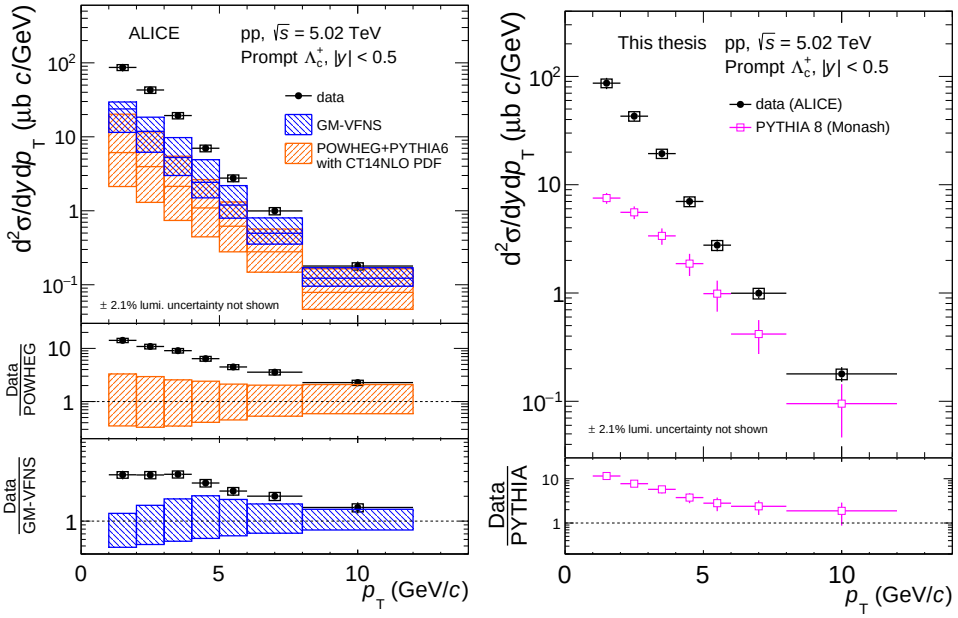


Figure 2.2: The p_{T}-differential production cross section for prompt Λ_{c}^{+} in $\sqrt{s} = 5.02$ TeV pp collisions as measured by the ALICE Collaboration [102, 103]. Comparisons to predictions from GM-VFNS [84], POWHEG [97], and PYTHIA 8 [95] are shown, which all significantly underestimate the measurements at low p_{T}.

2.2 Cold-nuclear matter effects

It is important to have a quantitative understanding of the cold-nuclear matter (CNM) effects, induced by the presence of nuclei in the initial state, before to characterise the properties of the medium produced in nucleus–nucleus collisions. This is typically performed by studying the production in proton–nucleus collisions. Open heavy-flavour production is affected by two CNM effects: (i) a modification of the effective partonic luminosity due to the different dynamics of partons in nuclei and (ii) multiple scatterings of partons in the nucleus before (and after) the hard scattering.

The first effect is dominant for heavy flavours at LHC energies, and is a consequence of the larger parton density in nuclei, which modifies the parton distribution function. The structure of a high-energy nuclei is simply much more complex than a superposition of non-interacting component nucleons, as we have seen in Sec. 1.4. The PDF for a bound proton in the nucleus with A nucleons (the so-called nuclear PDF, or nPDF, $f^{\mathrm{A}}(x;\mu_{\mathrm{F}}^2)$) can be related to that of a free proton ($f^{\mathrm{p}}(x;\mu_{\mathrm{F}}^2)$) by the ratio

$$R(x;\mu_{\mathrm{F}}^2) = \frac{f^{\mathrm{A}}(x;\mu_{\mathrm{F}}^2)}{A \cdot f^{\mathrm{p}}(x;\mu_{\mathrm{F}}^2)}. \tag{2.2}$$

Different mechanisms are thought to be responsible for the modification in different regions of x, and schematically four regions can be identified. First of all, a depletion ($R < 1$) is observed at small $x \lesssim 10^{-2}$, called *shadowing* [115], which can be understood as due to phase-space saturation (i.e. a high gluon fusion probability) in the high-density gluon systems in the

initial state. A possible enhancement ($R > 1$) known as *anti-shadowing* is seen at intermediate values $10^{-2} \lesssim x \lesssim 10^{-1}$, usually discussed as coming from the application of conservation rules for momentum [116]. Then there is the *EMC effect* (European Muon Collaboration effect), a depletion that takes place at large $x \gtrsim 10^{-1}$ [117], for which the nature is not yet uncovered. Finally, an enhancement of the nPDF at $x \gtrsim 0.8$ is found because of *Fermi motion*, meaning the nucleons in the nucleus are moving with larger transverse momentum compared to a free proton. There is no overarching theoretical calculation for the observed pattern over the entire x range, so the nuclear PDFs are calculated based on phenomenological parametrisations via global fits to several sets of experimental data [118, 119], as for the proton PDFs. Heavy-flavour production at midrapidity at the LHC will result in Bjorken $x \sim 10^{-4}$ for charm and $\sim 10^{-3}$ for beauty [86], and therefore mostly influenced by a modification of the nuclear PDFs in the shadowing region.

The second relevant CNM effect for heavy-flavour production are the multiple elastic scatterings of the partons with the constituents of the target nucleus that may occur before and after the hard scattering. The parton will acquire additional k_{T} after each scattering, responsible for a shift of the transverse momentum spectrum, usually known as k_{T}-broadening, or the Cronin effect [120]. On the other hand, the parton will also lose energy when travelling through cold-nuclear matter[3], which has the same qualitative behaviour as final-state energy loss in the QGP, although much smaller in magnitude [121].

For several years, the paradigm was that proton–nucleus collisions were only useful to probe the above CNM effects coming from the initial state. However, recent experimental data [122–125] opened the question if the large set of particles produced in p–A collisions form a "medium" with some degree of collectivity. In this case, final-state effects will impact the measurement of heavy-flavour production in such collisions as well, making a quantitative interpretation of the genuine CNM effects more complex. Besides in-medium energy loss mechanisms, that we will discuss in more detail in the next section, acquired radial and elliptic flow may influence the p_{T}-spectrum shape of open heavy-flavour hadrons. This was investigated in Refs. [126–128], where it was concluded that these flow effects, in a small and short-living medium, indeed leave signatures in the final hadronic observables of heavy-flavour hadrons, visible via an enhancement in the nuclear modification factor R_{pPb} at intermediate p_{T}.

Finally, besides the possible interplay with final-state effects, it is also not given that the CNM effects can easily be extrapolated from p–A collisions to the two-nuclei system in A–A collisions. Fortunately, phenomenological works have shown that such an extrapolation is possible for the relevant CNM effects for open heavy-flavour production. Regarding the nuclear PDF effects, as long as the factorisation theorem holds for the production cross section, the nuclear modification factors also factorise, i.e. $R_{\mathrm{AA}}^{\mathrm{CNM}}(y) = R_{\mathrm{pA}}(+y) \cdot R_{\mathrm{pA}}(-y)$ [86]. Also the cold-nuclear matter effects due to multiple scatterings in A–A collisions at midrapidity are, because of the symmetry in the $\mathrm{Q\overline{Q}}$ production, roughly double the nuclear modification as expected in p–A collisions [129].

[3]Either via radiative or collisional energy-loss mechanisms, discussed in more detail in the next section.

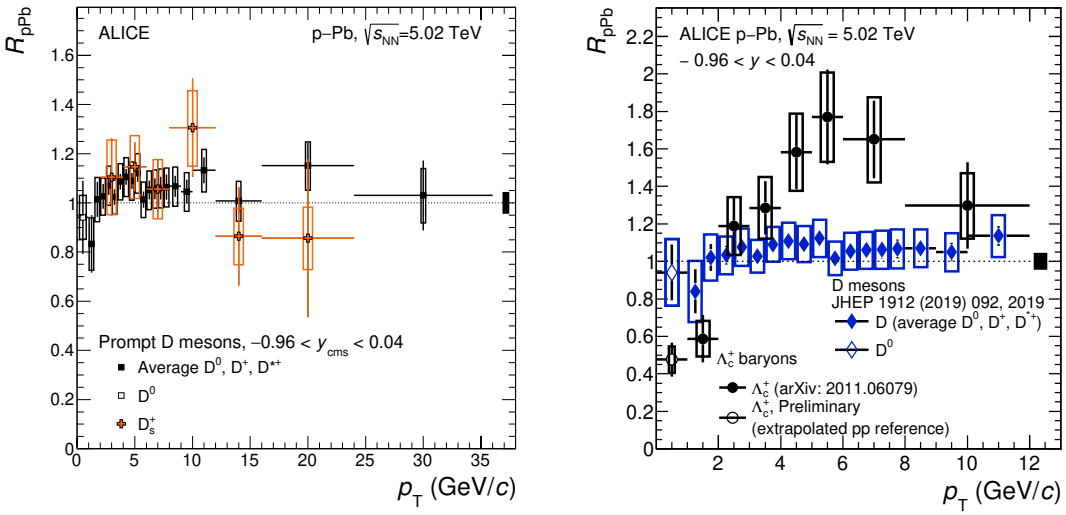


Figure 2.3: The nuclear modification factor R_{pPb} of prompt D mesons (an average of the non-strange D meson species together with the $\mathrm{D_s^+}$) and Λ_c^+ baryons in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV as measured by the ALICE Collaboration [102, 130].

In Fig. 2.3, the nuclear modification factor in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV is shown for non-strange D, $\mathrm{D_s^+}$, and Λ_c^+ hadrons [102, 130]. As discussed above, it is difficult to get a complete picture of the cold-nuclear matter effects due to the interplay among the different effects, and, possibly, the inclusion of final-state effects as well. What one can conclude though, is that for the D mesons, the R_{pPb} is consistent with unity, meaning that the effects are either negligible, or cancel each other. A suppression of D mesons in Pb–Pb collisions will therefore be a final-state effect, coming from interactions with a hot QCD medium. For the Λ_c^+ baryons, an enhancement (depletion) in the R_{pPb} at intermediate (low) p_T is observed, although still with relatively large uncertainties. This implies that it should be taken into account for a study of Λ_c^+ production in Pb–Pb collisions. The R_{pPb} for Λ_c^+ cannot be quantitatively described by theoretical calculations on the market, so a full understanding is still lacking. It is, however, interesting to note the similarities with the light-flavour sector [124, 131, 132], where an enhanced R_{pPb} was also observed for baryons with respect to mesons (and unity), believed to be induced from radial flow.

2.3 Hot-nuclear matter effects

For open heavy-flavour production to be a sensitive probe of the hot-nuclear matter effects induced by a QGP, a crucial requirement is that the number of heavy quarks remains constant during the evolution of the medium. As discussed in Sec. 2.1, thermal production of heavy quarks is negligible at the current achievable collision energies [76], and therefore heavy-flavour production is restricted to the initial stages of the collision. As shown in Ref. [133], also the $\mathrm{c\overline{c}}$ annihilation rate in the QGP can be safely neglected for the RHIC and LHC scenarios, fulfilling the prerequisite of a stable number of heavy quarks. Therefore, charm and beauty quarks, in contrary to light quarks and gluons, preserve their "identity" while traversing the medium, and can be tagged from low to high p_T through measurements of heavy-flavour hadrons.

Important implications

Heavy quarks retain their flavour and mass identity while traversing the QGP and can therefore be tagged throughout all p_{T} ranges by measurements of heavy-flavour hadrons.

During their propagation through the medium, heavy quarks lose part of their energy interacting with the medium constituents, occurring either via elastic (collisional energy loss) and inelastic (radiative energy loss) scattering processes. The dominant source for ultra-relativistic heavy quarks ($p_{\mathrm{Q}} \gg m_{\mathrm{Q}}$) is generally considered to be gluon radiation. For relativistic quarks ($p_{\mathrm{Q}} \lesssim 10m_{\mathrm{Q}}$), the importance of elastic scattering processes is believed to increase with decreasing p_{Q} [134, 135]. When studying bulk thermalisation of heavy quarks, radiative energy loss can therefore be neglected as most of the heavy quarks are produced with $p_{\mathrm{Q}} \lesssim m_{\mathrm{Q}}$. Considering only collisional energy loss, the thermal relaxation time of heavy quarks was found to be much larger than for light quarks (by a factor $\sim m_{\mathrm{Q}}/T$), and comparable to (or longer than) the medium lifetime [136]. This — together with the fact that heavy-flavour production is restricted to the initial stages of the collision — thus means that the momentum spectra of heavy-flavour hadrons may encode a "memory" of their interaction history, and the observed modifications with respect to a baseline measurement can serve as a gauge of the heavy-flavour coupling strength to the medium.

Important implications

The momentum spectra of heavy-flavour hadrons encode a "memory" of the heavy-quark interaction history because of their long thermal relaxation times.

The radiative in-medium energy loss mechanism originates from q(g)Q $\rightarrow$ q(g)Qg processes, where q (g) is a medium light quark (gluon), analogous to the electromagnetic bremsstrahlung process. The total radiative energy loss depends on the properties of the medium, and in particular on the medium path length L. This dependency originates from the long formation times of these inelastic scatterings, which will contribute coherently. The total energy loss $\Delta E = p_{\mathrm{Q}}^{\mathrm{in}} - p_{\mathrm{Q}}^{\mathrm{out}}$ can therefore only be evaluated at the end of the in-medium path length. This dependency on L is quadratic, as can be shown within the BDMPS-Z formalism [137]

$$\langle \Delta E_{\mathrm{rad}} \rangle \propto \frac{\alpha_{\mathrm{s}} C_R}{2} w_c = \frac{\alpha_{\mathrm{s}} C_R}{4} \hat{q} L^2. \tag{2.3}$$

Here, w_c is the characteristic energy of the emitted gluons, C_R are the Casimir factors for the QCD vertices, and $\hat{q}$ the transport coefficient characterising the medium-induced k_{T}^2 transferred to the projectile per unit path length. In addition, a quark-mass dependent correction is expected to be introduced through a phenomenon known as the *dead-cone effect* [134], whereby gluon-bremsstrahlung from an emitter (with mass m and energy E) is suppressed for angles smaller than m/E, relative to the direction of the emitter. The dead-cone effect is very relevant for heavy quarks, reducing the emission in the collinear region, while for the light-flavour sector radiative energy loss is the dominant mechanism even in the low-p_{T} regime. This leads to a flavour

hierarchy in the radiative energy loss for relativistic energies: $\Delta E(\mathrm{g}) > \Delta E(\mathrm{q}) > \Delta E(\mathrm{Q})$.

The collisional energy loss mechanisms occur via q(g)Q → q(g)Q scatterings, and are getting important for merely relativistic quarks. Let us first, however, briefly consider collisional energy loss in the ultra-relativistic limit ($p_{\mathrm{Q}} \gg m_{\mathrm{Q}}$), where similarities can be drawn with the "simpler" situation of a fast muon crossing a QED $\mathrm{e^+e^-}$ plasma. The two partonic processes that need to be considered in this case are Coulomb collisions (with electrons and positrons) and Compton collisions (with photons). A detailed review is presented in Ref. [138], where it is shown that the total energy loss due to elastic collisions will have a linear dependency on the in-medium path length and a logarithmic dependency on the initial parton energy.

For a general description of the space–time evolution of heavy quarks undergoing multiple elastic scatterings in a QGP — which is important since these scatterings transfer the radial and anisotropic flows of the QGP to the heavy quarks — one can rely on transport theories via the Boltzmann equation [139]. Since $m_{\mathrm{Q}} \gg T$, the typical momentum exchange of heavy quarks with the medium, $\vec{q}^2 \simeq T^2$, is small compared to its thermal momentum, $p_{\mathrm{Q}}^{\mathrm{th}} \simeq \sqrt{2m_{\mathrm{Q}}T}$, enabling a Brownian motion treatment[4]. In this limit, the Boltzmann equation can be approximated by the Fokker–Planck equation [140], which can be solved stochastically with a Monte-Carlo simulation using the macroscopic Langevin equation [136]

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = -\eta_D(p)\vec{p} + \vec{\xi}. \tag{2.4}$$

Here, $\eta_D(p)$ is a momentum drag (friction) coefficient and $\vec{\xi}$ delivers random (thermal) momentum kicks. The drag coefficient is related to the momentum (spatial) diffusion coefficient D_p (D_s) through the Einstein fluctuation–dissipation relation $D_p = m_{\mathrm{Q}}\eta_D T$ ($D_s = T/(m_{\mathrm{Q}}\eta_D)$). The spatial diffusion coefficient is directly related to the thermal relaxation time of heavy quarks $\tau_{\mathrm{Q}} = m_{\mathrm{Q}}/TD_s$, where a larger value of D_s implies a looser coupling with the medium (less frequent rescatterings), and therefore a longer thermalisation time. This thus implies that heavy-flavour production can be a sensitive probe for the transport coefficients of the QGP.

Important implications

Because the typical thermal momentum of a heavy quark is much larger than the typical momentum transfer from the medium, a diffusion treatment can be adopted.

Constraining the transport properties of the medium is typically performed via comparisons of model predictions to measurements of the nuclear modification factor R_{AA} and the elliptic flow v_2[5]. Various theoretical approaches exist to calculate the collisional and/or radiative heavy-quark interactions with the medium, from which the drag and diffusion coefficients follow (see

[4] The heavy quark receives many relatively small momentum kicks from the medium, equivalent to the (natural) Brownian motion phenomenon.

[5] This is the dominant harmonic coefficient in mid-central collisions, characterised by the magnitude of the second Fourier coefficients $v_n = \langle \cos n(\phi - \Psi_n) \rangle$, where ϕ is the particle-momentum azimuthal angle, the brackets denote the average over all measured particles, and Ψ_n is the symmetry-plane angle relative to the $\mathrm{n^{th}}$ harmonic.

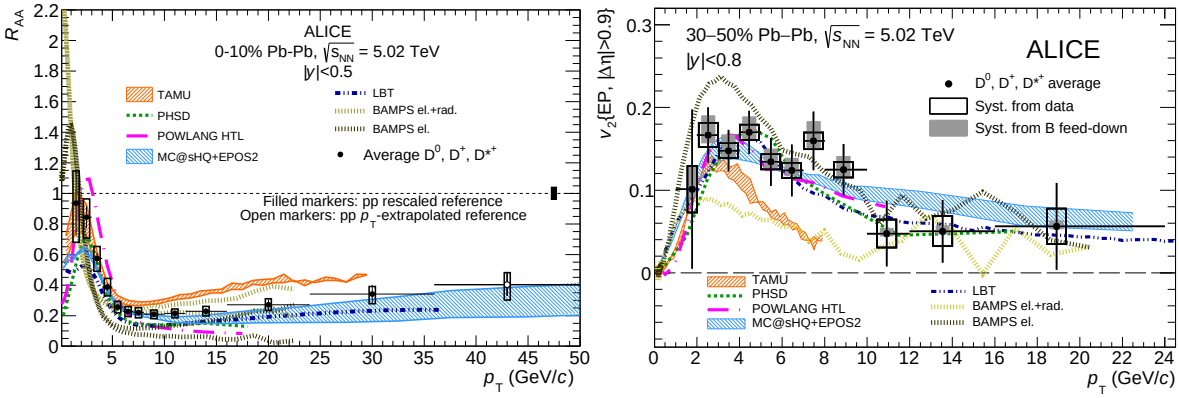


Figure 2.4: The average $R_{\rm AA}$ (left) and v_2 (right) of $\rm D^0$, $\rm D^+$, and $\rm D^{*+}$ mesons in, respectively, central and mid-central Pb–Pb collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV as measured by the ALICE Collaboration, compared with models that have predictions for both observables [141, 144].

Refs. [139, 140] for detailed reviews on the different type of calculations). When predictions for the $R_{\rm AA}$ and v_2 observables of these models are compared to measurements, constraints on the transport coefficients can be put after, e.g., calculating the agreement with data via the χ^2/ndf [141, 142] or calibrating the model parameters using a Bayesian model-to-data analysis [143]. Measurements of heavy-flavour hadrons open therefore a unique access to the transport coefficients of the quark–gluon plasma.

In Fig. 2.4, the nuclear modification factor $R_{\rm AA}$ and elliptic flow v_2 of non-strange D mesons in, respectively, central and mid-central Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV as measured by the ALICE Collaboration are shown [141, 144]. The average $R_{\rm AA}$ of the three non-strange D-meson species shows a clear nuclear modification for $p_{\rm T} > 3$ GeV/c, is compatible with that of charged particles for $p_{\rm T} > 6$ GeV/c, and hints to less suppression for the strange $\rm D_s^+$ meson (see Ref. [144] for these latter two observations). The elliptic flow for the non-strange D mesons is larger than zero for $2 < p_{\rm T} < 10$ GeV/c, similar to the v_2 of charged particles in the common $p_{\rm T}$ interval (1–16 GeV/c), and compatible with that of the strange $\rm D_s^+$ meson [141]. Transport model predictions (see references in Ref. [144]) that compute both observables are compared to the measurements in Fig. 2.4, but most of them fail in describing both simultaneously. Such measurements provide therefore important constraints for the understanding of the interaction of heavy quarks with the high-density QCD medium.

2.4 Heavy-flavour hadronisation

For heavy-flavour hadronisation we consider typically two type of mechanisms, fragmentation and in-medium hadronisation via recombination (coalescence), which were discussed already briefly in Sec. 2.1 and 1.5.2, respectively. Here we will discuss them in more detail, starting with the "traditional" vacuum fragmentation in the next paragraphs, and discussing a few of the relevant[6] ideas for hadronisation via recombination after.

[6]Recombination/coalescence is a term that has a very wide usage for processes at very different time scales in high-energy (nuclear) physics, hence the specification.

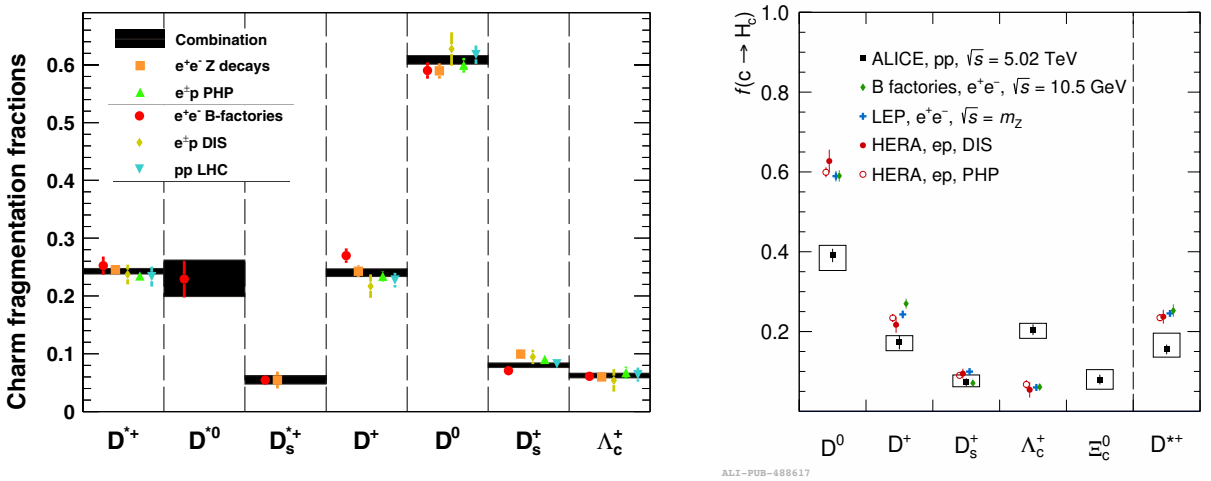


Figure 2.5: Left: The charm fragmentation fractions $f(\mathrm{c} \to h_\mathrm{c})$ as measured in $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collisions by different experiments [146]. Fractions for pp collisions are shown as well, but these are directly based on the Λ_c^+ production measurement of LHCb [101], which is being questioned by the field. Right: The charm fragmentation fractions, including for the first time the contribution of the heavier charm baryon Ξ_c^0 [105], in pp collisions at $\sqrt{s} = 5.02$ TeV as measured by ALICE [147].

The clean environment in $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collisions allows for a determination of the bare quark state before hadronisation, making it possible to study the fragmentation functions of heavy quarks. Via fits of these measurements, with for example the Peterson fragmentation function [145], the fragmentation functional form can be parametrised. The Peterson fragmentation function has only one free parameter, which is expected to ensure a harder fragmentation function (the quark will loose less of its initial momentum) for a lighter charm-hadron specie. The same holds for b hadrons, although in general a harder fragmentation function is expected for beauty than for charm hadrons. The fragmentation functions include the probability of a quark fragmenting into a given hadron h, called the *fragmentation fraction*. A summary of the fragmentation fractions for charm hadrons, from a combined analysis performed using different experiments, is shown in Fig. 2.5 [146]. The measurements in $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ are compatible within uncertainties, on which the fragmentation universality assumption was built. However, as shown before in Fig. 2.2, this universality seems not to hold in fully hadronic collision systems like pp [100–109], which is taken into account for the charm-sector by ALICE [147] in the right panel of Fig. 2.5. A clear difference with respect to the charm fragmentation fractions measured in $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collisions is observed.

In the state-of-the-art Monte Carlo event generators, the hadronisation of a parton (both for light or heavy flavours) is modelled using phenomenological, non-perturbative implementations of the fragmentation process, such as via the formation of strings (PYTHIA [95]), clusters (HERWIG [96]), or ropes (DIPSY [148]). The Lund string model [149] from PYTHIA is the most widely used one, where two partons interact via colour fields resembled by strings. When the connected partons move away from each other, it will, at a certain time, get energetically more favourable for a new $\mathrm{q\bar{q}}$ pair to appear, rather than extending the string further, and therefore breaking the string into two. This is modelled via quantum tunnelling, with the

probability depending on the transverse mass and the longitudinal energy density (or string tension).

The fragmentation parameters in these models are tuned[7] on e^+e^- and e^-p measurements, used directly in calculations for hadronic collisions as well. Discrepancies between such predictions and pp measurements are therefore likely to come from phenomena not present at lepton colliders, where colour correlations and the possibility for multiple parton interactions (MPI) due to the increased density of quarks and gluons in the initial state are most likely. The first measurements of Λ production at the LHC [152, 153] showed such a discrepancy, leading to the development of phenomenological extensions of the string model for pp collisions via colour reconnections [95, 110] and rope hadronisation [154] (a mechanism developed originally for heavy-ion collisions [155]).

The main question such models need to address is between which partons the confining potentials arise. Traditionally, the leading-colour (LC) approximation was used, where a quark (gluon) is connected to one (two) other partons within the same MPI scattering system. For example, a scattered quark at midrapidity may be colour connected with a parton in the remnant of the proton, but not with a quark from a different MPI. Colour reconnection (CR) models [95] improve this picture by allowing a fraction of partons to "forget" their LC connections and build new ones based on a preference of shorter strings. More realistic CR models beyond the leading-colour approximation were developed recently [110] introducing SU(3) colour-algebra selection rules to the string-length minimisation strategy. This colour reconnection scenario allows for "junctions" to be formed, with a topology closely related to baryons, and resulting therefore in a baryon enhancement.

The colour rope hadronisation model [154] introduces interactions between strings (treated as flux tubes with fixed radius in this model), based on the idea that those next to each other in geometrical space can act coherently. Following SU(3) colour rules, it can either form a "final-state swing" (reconnects colour dipoles) or a "colour rope". The latter one will hadronise with a higher effective string tension, i.e. there will be more energy available for fragmentation. This results in more strange quarks and diquarks produced in string breakups, and thus more baryons and strangeness among the final hadrons.

Figure 2.6 shows a comparison of the predictions from PYTHIA 8, with and without CR beyond leading colour approximation, and DIPSY, with and without colour rope hadronisation, for several hadron-specie ratios as function of event activity [156]. Both extended models show an increase in baryon production (p/π and Λ/K ratios), where an extra strangeness enhancement is only observed for DIPSY with rope hadronisation (K/π[8] and Ξ/Λ). While these phenomenological extensions of vacuum fragmentation have been compared to light-flavour measurements

[7]Event generators come with different tunes: a set of (non-physical) flags, modes, and parameters tuned in such a way that the event generator reproduces, to a certain extent, a variety of different measurements. Many different tunes exist, but in ALICE we typically use PYTHIA 6 with the Perugia 2011 tune [150] or PYTHIA 8 with the Monash 2013 tune [151].

[8]The PYTHIA model with CR beyond the LC approximation lies below its baseline for the K/π ratio due to phase-space constraints for low invariant-mass strings [156].

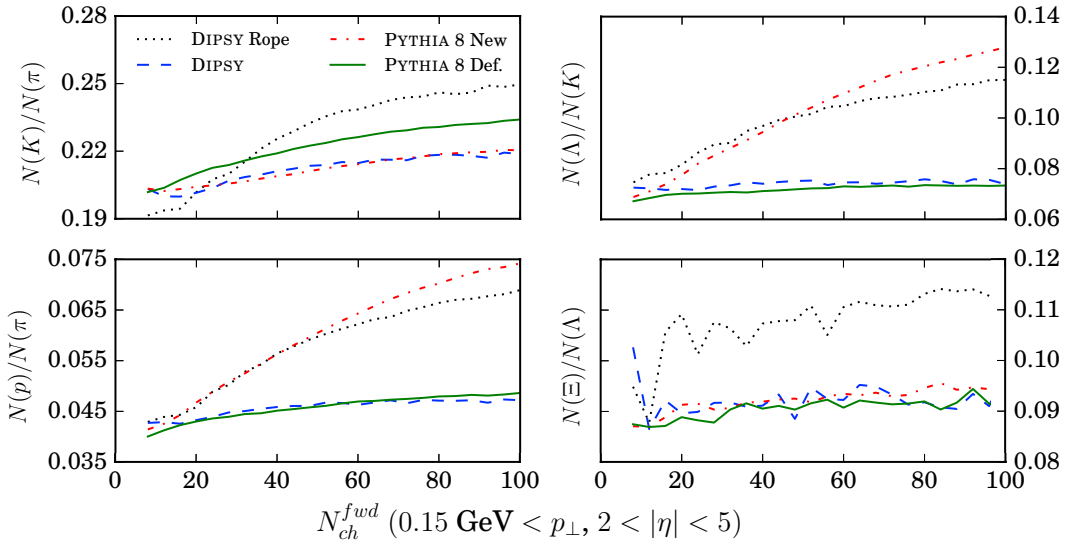


Figure 2.6: Ratios of identified hadrons in the central region ($|\eta| < 1$) as function of event activity (multiplicity at forward rapidities) for pp collisions at $\sqrt{s} = 13$ TeV as predicted by the PYTHIA 8 and DIPSY event generators with and without CR beyond leading colour approximation and rope hadronisation, respectively. Figure adjusted from Ref. [156].

extensively [8, 102, 157], and in general provide for a better description of the hadron-specie ratios, comparisons to measurements in the heavy-flavour sector are sparse. Nevertheless, heavy quarks provide interesting constraints for these novel hadronisation mechanisms, because they come from a finite source (the hard scattering) and are therefore always a string end point. One of the model authors stated that the heavy-flavour sector may actually be more constraining than the light-flavour sector if we understand the measurements in the correct way (P. Skands, private communication, May 2021).

Important implications

Since heavy quarks are coming from a finite source, they are (among) the most constraining probes for hadronisation mechanisms.

It is clear by now that vacuum fragmentation is no longer valid in presence of any reservoir of partons. An alternative mechanisms for parton hadronisation, with respect to the two model implementations discussed above, is called *recombination* (or coalescence). The word "recombination" has a wide usage in the history of high-energy physics: from the mechanism to form nuclei built from nucleons at the kinetic freeze out, to explanations for the "leading particle effect" by combining partons from the hard scattering with valence quarks in the beam remnants [158]. Here, we are however considering recombination between constituents of the colour-deconfined medium, i.e. a heavy quark combining with one or two thermal partons in the QGP to form a hadron. The main difference with respect to hadronisation via fragmentation is that in this case $p_h = p_{\rm Q} + p_{\rm q,1}(+ p_{\rm q,2})$, instead of $p_h < p_{\rm Q}$.

Details of the different implementations of recombination are beyond the scope of this chapter, but the fundamental requirement is that the partons have to be close in phase space in order to

have a relevant overlap of their wave functions. The details of the dynamical process are then discarded in favour of exploiting an adiabatic approximation, which considers a projection of the initial state onto the final clustered state. Two implementations of recombination considered often in recent literature regarding heavy-flavour hadronisation are the quark coalescence model via the Wigner-function formalism [159] considered in the Catania model [160], and the Resonance Recombination Model (RRM) [161] used by the TAMU model [162]. One of the main differences between both implementations is that the quark coalescence model is an instantaneous projection, which means interactions with the surrounding medium are not included, and therefore four-momentum is not conserved. The RRM is not instantaneous and proceeds via the formation of resonant states, governed by different time scales for different resonant states. In this case four-momentum is conserved in the coalescence process.

It is expected that the recombination mechanism gives an important contribution at low p_{T}, while fragmentation will be dominant at high p_{T} [135, 160]. Measurements of the production of strange-charm mesons ($\mathrm{D_s^+}$) and charm baryons (Λ_c^+) in heavy-ion collisions represent a sensitive probe of the recombination mechanism, especially their production ratios to the non-strange D mesons compared to this ratio in pp collisions. For $\mathrm{D_s^+}$ mesons, the $\mathrm{D_s^+/D^0}$ ratio is expected to be enhanced due to the strangeness enhancement in the QGP, "acquired" by the production of $\mathrm{D_s^+}$ mesons via recombination. This picture is supported by production measurements in Pb–Pb collisions [163]. Also the $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ ratio is predicted to be enhanced in nucleus–nucleus collisions in case quark coalescence plays a significant role for charm quarks [164, 165]. Similarly as for the light-flavour baryons, recombination will push charm baryons to higher momenta with respect to charm mesons because of the additional constituent quark. As alternative (or additional) explanations, the radial flow of the system and the possibility that binary, coloured bound states of light partons exist in the QGP [166] could also contribute to a baryon enhancement.

Let us now go back to the Λ_c^+ puzzle, the enhancement observed in pp collisions with respect to measurements in $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ systems. Before one can conclude on the type of Λ_c^+ hadronisation mechanisms in A–A collisions, a proper understanding of the reference pp system should be obtained. The baryon-to-meson ratio $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ in pp collisions at $\sqrt{s} = 5.02$ TeV [102, 103] compared to model predictions [96, 110–112, 151] is shown in the left panel of Fig. 2.7. Predictions from PYTHIA 8 with the Monash tune and HERWIG 7, both with fragmentation parameters tuned on $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collision data, significantly underestimate the measurement in pp. Two of the "extended" hadronisation models discussed in this section, PYTHIA 8 with CR beyond the LC (CR Mode 2 in the legend) and Catania including fragmentation and recombination, do a much better job reproducing the data. A third model, based on statistical hadronisation (see Sec. 1.5.2) with an augmented set of baryon stated predicted by the Relativistic Quark Model (RQM) [167] also catches well the measurement. This, including the simple non-universality of the fragmentation fractions argument, provides us with four possible explanations for the Λ_c^+ puzzle in pp collisions. In the right panel of Fig. 2.7, the first[9] measurement in nucleus–nucleus

[9]Which was at the start of this work also the only available measurement in nucleus–nucleus collisions. In the meantime, more measurements of Λ_c^+ production in heavy-ion collisions appeared, but I keep those for the dedicated Chapter 7.

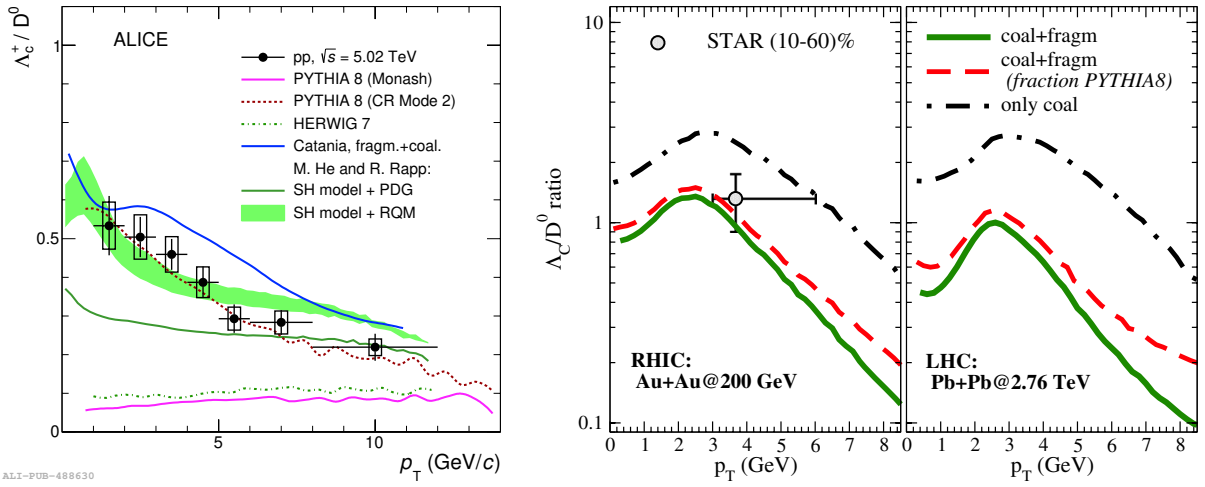


Figure 2.7: Left: The Λ_c^+/D^0 ratio measured in pp collisions at $\sqrt{s} = 5.02$ TeV as measured by the ALICE Collaboration [102, 103] compared to different theoretical calculations [96, 110–112, 151]. Right: The first measurement of Λ_c^+ production in heavy-ion collisions as measured by the STAR Collaboration in $\sqrt{s_{NN}} = 200$ GeV Au–Au collisions [168] compared to theoretical predictions including hadronisation via recombination, including predictions for Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV [160].

collisions by the STAR Collaboration [168] is shown, compared to theoretical calculations including recombination. The measurement was found to be significantly higher than the SHM and PYTHIA 6 prediction [168], and compatible with calculations including hadronisation via recombination [160]. However, more precise and more differential results, as well as reference measurements at the same collision energy, are of crucial importance for a proper interpretation of these results. This was one of the main objectives of this work, and will be discussed in Chapter 7.

2.5 Discussion

This chapter was dedicated to open heavy-flavour physics, a sub-field in high-energy nuclear physics, and the subject of this dissertation. The motivations for measuring open heavy-flavour hadrons in pp, p–Pb, and Pb–Pb collisions have been described via a (brief) description of the theoretical concepts and a discussion of some of the intriguing experimental results. I did not limit the discussion to what was known at the start of this doctoral dissertation (October 2017), since the field proceeds on many fronts at the same time, and recent work like, e.g., the experimental and theoretical efforts on Λ_c^+ production in pp collisions of the last two years, is very relevant for the discussion on Λ_c^+ production in Pb–Pb collisions that will be presented in Chapter 7. Regarding the four measurements presented in the next chapters, I instead limited the discussion to what was known at the start of this research. For a complete overview, I refer to the relevant chapters: Chapter 5 for charm-meson production (D^{*+}) in pp collisions; Chapter 6 for heavy-flavour production (D_s^+ and Λ_c^+) in pp multiplicity event classes; Chapter 7 for charm-baryon production (Λ_c^+) in Pb–Pb collisions; and Chapter 8 for beauty-meson production (B_s^0) in Pb–Pb collisions.

To finish this chapter, let me make the connection between the aim of this thesis (presented via a series of questions that were asked at the start of each analysis) with the three relevant big open questions in the field of heavy-ion physics (as stated in the previous chapter):

- **Is the centre-of-mass energy dependence of charm production reproduced by pQCD calculations at LHC energies, i.e. are the conclusions made on the nuclear modification of charm quarks based on a pp-rescaled reference correct?** A previous publication by ALICE [169] has shown that D-meson production in pp collisions at $\sqrt{s} = 7$ TeV is described well by pQCD predictions. This is important, since heavy-flavour production measurements in p–Pb and Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV [144, 170] require a reference measurement in pp collisions at the same centre-of-mass energies (which did not exist), and therefore relied so far on a $p_{\rm T}$-dependent $\sqrt{s}$-scaling factor from FONLL calculations [87, 88]. Any conclusion on the effects of heavy-quark energy loss and the dynamics of heavy-quark hadronisation in heavy-ion collisions — especially via the nuclear modification factor — thus depends on the models implementation of the $\sqrt{s}$-dependence in D-meson production. To have a model-independent measurement, it is important to provide the measurement in pp collisions at the same centre-of-mass energy per nucleon–nucleon collision as for the Pb–Pb collision. Furthermore, as pointed out in Ref. [91], the production cross section ratios at different LHC energies will make the theoretical uncertainty due to the choice of the factorisation and renormalisation scale subdominant with respect to the uncertainty on the PDFs, providing a sensitivity to "another aspect" of the pQCD calculation.

- **Is the process of hadronisation of heavy quarks in pp collisions properly described by model predictions?** This question is approached in three different ways in this dissertation, probing the difference in the production between pseudoscalar ($\rm D^0$) and vector ($\rm D^{*+}$) mesons, between non-strange ($\rm D^0$) and strange ($\rm D_s^+$) mesons, and between mesons ($\rm D^0$) and baryons (Λ_c^+)[10]. The importance of the latter two comparisons have been discussed in detail in Sec. 2.4, and such measurements will help to constrain the phenomenological models of QCD in the non-perturbative regime of hadronisation. In addition, with the first comparison the fragmentation functions for vector mesons, for which only one dedicated parametrisation is on the market [171], can be constrained. This can, for example, be seen in Fig. 2.8 for the $\rm D^+/D^0$ (both non-strange pseudoscalar mesons) and $\rm D^{*+}/D^0$ ratios in pp collisions at $\sqrt{s} = 7$ TeV [169], where a much larger spread in the model predictions is observed for the $\rm D^{*+}/D^0$ ratio, mainly due to the different choices for the $\rm D^{*+}$ fragmentation function.

- **Do we observe a multiplicity dependency in the strange-to-non-strange or baryon-to-meson production ratios for charm hadrons in pp collisions, and, if this is the case, can this be explained with initial state or "standard" QCD**

[10] Let me already mention here that this work is performed in close collaboration with other ALICE members, where each person was responsible for the analysis of one of the hadrons. I therefore do not "claim" all these hadrons as mine, but, given my large efforts on the general aspects and finalisation of these analyses, I will show and discuss the final results of those hadrons in the relevant chapters as well.

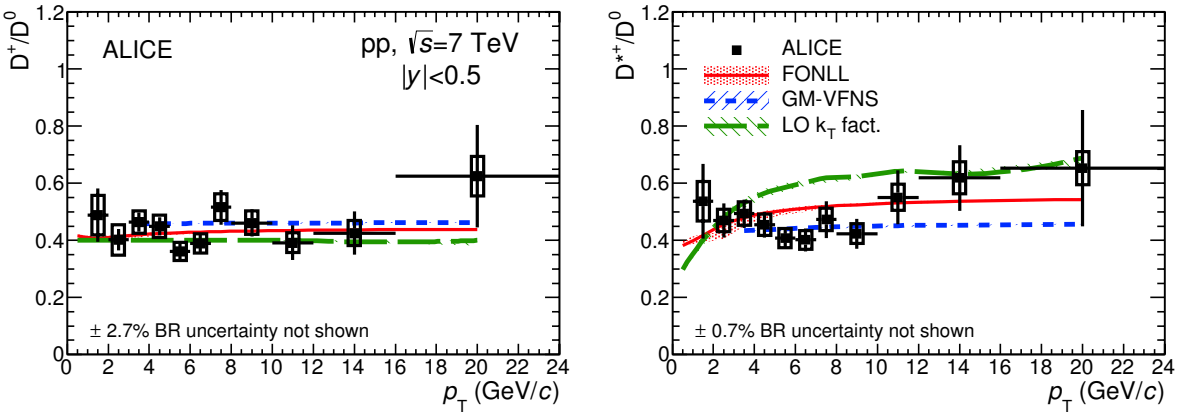


Figure 2.8: Ratios of D-meson production cross sections as a function of p_{T}, in particular the ratio of two pseudoscalar mesons $\mathrm{D}^+/\mathrm{D}^0$ (left) and the vector-to-pseudoscalar ratio $\mathrm{D}^{*+}/\mathrm{D}^0$ (right), in pp collisions at $\sqrt{s} = 7$ TeV compared to model predictions [169].

effects, or do we need to assume QGP-like mechanisms in pp collisions? This question is motivated by observations in the light-flavour sector [8, 157, 172], where enhancements of baryons and (multi-)strange hadrons were observed in high-multiplicity pp collisions, hinting to the presence of a collective, equilibrated system already in pp systems. The study of heavy-flavour production as a function of multiplicity can provide important further insights, since, as discussed in Sec. 2.4, the heavy quarks are coming from a finite source. They are therefore a more genuine probe of the underlying hadronisation mechanisms, in contrary to light flavours, which can easily be "popped out of the vacuum". In addition, heavy quarks are less likely to thermalise in case a small droplet of QGP in pp collisions would be produced, and therefore may carry more information on possible medium-induced interactions. On the other hand, the study of the baryon-to-meson ratio in ultra-low pp multiplicities might approach the $\Lambda_{\mathrm{c}}^+/\mathrm{D}^0$ ratio measured in $\mathrm{e}^+\mathrm{e}^-$ collisions and with that contribute to our understanding of the Λ_{c}^+ puzzle.

- **In the presence of a hot, deconfined medium created in Pb–Pb collisions, does recombination of charm quarks with light medium constituents contribute to the hadronisation of charm quarks?** As discussed extensively in Sec. 2.4, the presence of the colour-deconfined quark–gluon plasma will alter the heavy-quark hadronisation mechanisms with respect to vacuum fragmentation. The quark recombination (or coalescence) picture is nowadays the main approach to heavy-flavour hadronisation in many heavy-ion models, although different theories exploiting, e.g., statistical hadronisation with heavier-state resonances are in agreement with the sparse experimental data as well. The charm baryon-to-meson ratio, $\Lambda_{\mathrm{c}}^+/\mathrm{D}^0$, is shown to be sensitive to in-medium hadronisation mechanisms [164, 165], and a measurement of Λ_{c}^+ production in Pb–Pb collisions is therefore of great importance to shed light on the hadronisation and/or thermalisation phase for the charm sector.

- **Experimentally, will it be feasible to extend the in-medium heavy-flavour energy loss and hadronisation picture to the beauty sector in LHC Run 3 and**

4? From a theoretical point of view, beauty hadrons are a cleaner probe than charm hadrons for the strongly-coupled QGP (in terms of the implementation of microscopic interactions and transport because of their larger mass, and as a measure of the coupling strength without saturation because of possible thermalisation). The current experimental setup of the ALICE experiment does not allow for measurements of beauty hadrons in Pb–Pb collisions, and our experimental knowledge on the behaviour of heavy quarks in a colour-deconfined medium is so far mainly constrained by the charm sector. It is useful to evaluate the predicted performance of beauty-hadron measurements, like the B_s^0 meson, for Run 3 and 4 of the LHC, after the upgrade of the Inner Tracking System of ALICE.

CHAPTER 3

Experimental setup

In this chapter, a brief introduction to the accelerator complex at the European Organisation for Nuclear Research (CERN), and especially the Large Hadron Collider (LHC), will be presented. Afterwards, a detailed description of the ALICE (A Large Ion Collider Experiment) detector will be given, with the focus on the central barrel detectors that are utilised in this work. The chapter ends discussing some of the general procedures performed in the ALICE offline framework, like centrality determination and track reconstruction.

3.1 CERN accelerator complex

The Large Hadron Collider is the most powerful particle collider in the world, built at CERN, nearby Geneva, Switzerland. The LHC is a two-ring superconducting hadron accelerator, installed in a tunnel spanning a circumference of 26.7 km, situated 50 to 175 meters underground [173]. It is designed to accelerate protons as well as lead nuclei (Pb^{82+}) with a maximum centre-of-mass energy (per nucleon–nucleon pair) of $\sqrt{s} = 14$ TeV and $\sqrt{s_{NN}} = 5.5$ TeV, respectively. It benefits from the previous accelerators built at CERN in the last decades. In Fig. 3.1, a schematic view of the CERN accelerator complex is shown, presenting the preceding stages before the LHC, as well as the apparatus used by other experimental collaborations. The accelerating process for protons starts with them being extracted from a hydrogen tank, after which they are accelerated via stages in a linear accelerator (LINAC 2), the Proton Synchrotron Booster (BOOSTER), Proton Synchrotron (PS), and finally the Super Proton Synchrotron (SPS), before being injected in the LHC accelerator ring. The accelerating process of ions mainly differs for the initial two stages, where after extraction the atoms are accelerated and ionised in a linear machine (LINAC 3), after which they are accumulated in a dedicated ion ring (LEIR), before being injected in the PS.

There are four main experiments at the LHC, located at so-called interaction points where the

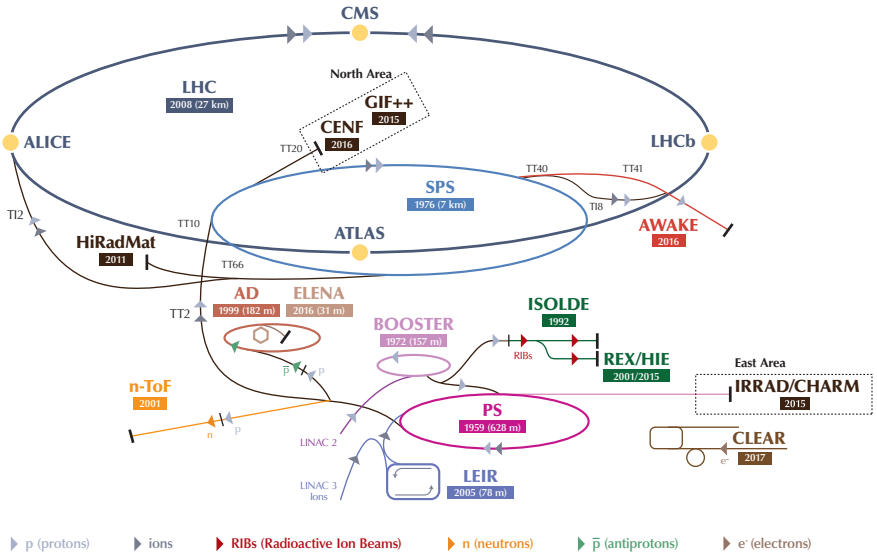


Figure 3.1: Schematic view of the layout of the CERN accelerator complex, including the main experiments [174].

beam pipes cross. The ATLAS (A Toroidal LHC Apparatus) [175] and CMS (Compact Muon Solenoid) [176] experiments are two general-purpose detectors designed to measure the Higgs boson and to search for physics beyond the Standard Model. The LHCb experiment [177] is specialised in the study of CP violation in beauty hadron-decays, while the ALICE experiment is dedicated to the study of heavy-ion collisions[1]. There are four more interaction points at the LHC which are used to collimate, accelerate, or dump the beams.

In 2009, the LHC collided its first proton beams, indicating the start of Run 1 (from November 2019 to February 2013). Proton–proton collisions were recorded at a centre-of-mass energy of $\sqrt{s} = 0.9$, 2.76, 7, and 8 TeV. During this run, three dedicated heavy-ion programs took place, providing Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV per nucleon–nucleon collision. In April 2015, the LHC restarted for Run 2, in which pp collisions at $\sqrt{s} = 5.02$ and 13 TeV, p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ and 8.16 TeV, and two Pb–Pb runs at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV were delivered. This work will be fully based on the data recorded by ALICE in LHC Run 2, so in the sections below we will discuss the ALICE detector with the setup as has been used during this period.

3.2 The ALICE detector

The ALICE detector is designed as a general-purpose detector for ultra-relativistic heavy-ion collisions [7], i.e. designed to address the physics of the quark–gluon plasma and the phase diagram of QCD matter. This requires the ability to track particles down to low momenta (about 100 MeV for pions) in environments with large charged-particle multiplicities at midrapidity (up to $\mathcal{O}(10^3)$ for central Pb–Pb collisions). Therefore, detectors with high granularity and low material budget have been developed. Another fundamental feature of ALICE is particle identification (PID), essential for many physics signals.

A schematic view of the ALICE detector, and its subdetectors, is shown in Fig. 3.2. It can be divided in two main parts: the *central barrel*, covering the central rapidity region ($|\eta| < 0.9$), and the *muon spectrometer* at forward rapidities ($-4 < \eta < -2.5$). The central barrel is embedded in the L3 Magnet (from the L3 experiment at LEP), which provides a maximum $B = 0.5$ T magnetic field parallel to the beam axis. A detailed overview of the ALICE experiment and its performance can be found in Refs. [7, 178–180]. In the next few sections the detectors directly employed in this dissertation are described.

3.2.1 Inner Tracking System

The Inner Tracking System (ITS) [182], depicted in the left panel of Fig. 3.3, is a six-layer silicon detector located closest to the interaction point. It is built primarily for charged particle tracking and the determination of the primary vertex (the collision point). For this work, it is also fundamental for the reconstruction of secondary vertices (point of decay for short-lived hadrons)

[1] Nowadays, all four experiments have a heavy-ion program.

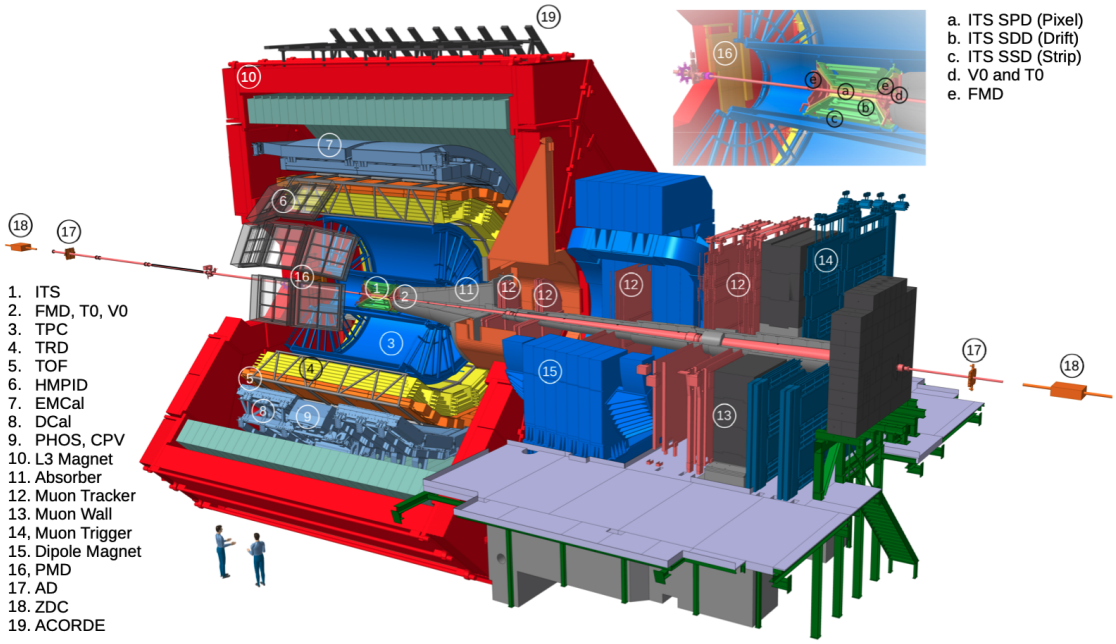


Figure 3.2: Schematic view of the layout of the ALICE detector during LHC Run 2 [181].

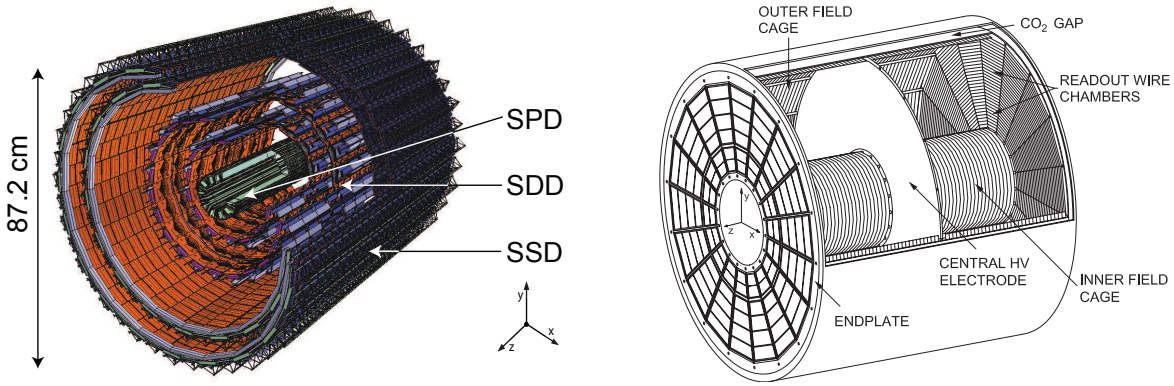


Figure 3.3: Schematic view of the ITS (left) and TPC (right) detectors [182, 183]. See text for more details.

and the measurement of the impact parameter of tracks (defined as the distance of closest approach between the track trajectory and the primary vertex). Three different technologies are used for the inner, middle, and outer layers.

The two innermost layers are equipped with Silicon Pixel Detectors (SPD), and are due to their position and resolution fundamental for the reconstruction of weakly-decaying strange, charm, and beauty hadrons. Thanks to the fast response, the SPD layers can be used as a trigger as well. Finally, the SPD layers provide a midrapidity multiplicity estimator via so-called tracklets, defined as track segments built by associating pairs of hits in both layers. The Silicon Drift Detectors (SDD), which are the third and fourth layers, provide a very good multitrack capability, as well as the possibility for particle identification via $\mathrm{d}E/\mathrm{d}x$ measurements. This latter feature is however not used for this dissertation because of the excellent PID capabilities of other detectors. The two outermost layers, equipped with double-sided Silicon Strip Detectors

(SSD) are fundamental for the prolongation of tracks from the main tracking detector of ALICE (the Time Projection Chamber, discussed in the next section) to the ITS. Also the SSD layers provide $\mathrm{d}E/\mathrm{d}x$ information. The main differences between the three technologies are summarised in Tab. 3.1.

Table 3.1: Parameters of the various type of detector layers in the ITS detector [7].

Parameter		Silicon Pixel	Silicon Drift	Silicon Strip
r position of inner(outer) layer	(cm)	3.9(7.6)	15.0(23.9)	38.0(43.0)
Material budget of inner(outer) layer	($\%X_0$)	1.14(1.14)	1.13(1.26)	0.83(0.86)
Spatial resolution $r\phi$	(µm)	12	35	20
Spatial resolution z	(µm)	100	25	830
Two track resolution $r\phi$	(µm)	100	200	300
Two track resolution z	(µm)	850	600	2400
Cell size	($\mu\mathrm{m}^2$)	50×425	202×294	95×40000

3.2.2 Time Projection Chamber

The Time Projection Chamber (TPC) [183], illustrated in the right panel of Fig. 3.3, is the main tracking detector of the central barrel. It is a cylindrical gas detector, placed symmetrically around the interaction point from about 85 to 250 cm, and covering the full azimuthal angle. With an overall length of 500 cm along the beam direction it covers a pseudorapidity range of $|\eta| < 0.9$ for full radial tracks (tracks matched to hits in the more inward and outward detectors). The TPC is optimised to track particles in high-multiplicity environments in a large transverse momenta range from 100 MeV/c up to 100 GeV/c, with good momentum resolution. In addition, the TPC detector has excellent PID capabilities over a wide momenta range.

The detector is made of a large cylindrical field cage, filled with a 90 m^3 gas mixture of primarily Neon with $\mathrm{CO}_2/\mathrm{N}_2$ or Argon with $\mathrm{CO}_2/\mathrm{N}_2$ for the first two years of Run 2[2]. Charged particles that traverse the TPC will ionise the gas, and the freed electrons will drift towards the outer ends of the cylinder under influence of a highly uniform electrostatic field of 400 V/cm in the beam direction. On the end plates, multi-wire proportional chambers with cathode pad readout are mounted, divided in 18 trapezoidal sectors. The signal amplification of the freed electrons is provided through the avalanche effect occurring in these readout chambers. The two-dimensional track position in $r\phi$ can be constructed from the position of the deposited charge on the cathode, where the z-direction is calculated from the drift time of the electrons.

Particle identification can be performed in the TPC via the measurement of the specific energy loss $\mathrm{d}E/\mathrm{d}x$, which is directly proportional to the number of freed electrons. The specific energy loss depends on the particle specie and momentum, and can be described by the Bethe-Bloch formula [16]. The identification of a particle is based on the comparison of the measured $\mathrm{d}E/\mathrm{d}x$ with the expectation for a specific hadron specie with momentum p. The $\mathrm{d}E/\mathrm{d}x$ information for a given track is extracted from the 65% lowest-amplitude clusters assigned to the track.

[2]The gas was changed back to the Neon mixture in 2017, because larges phase-charge distortions were observed with the Argon mixture.

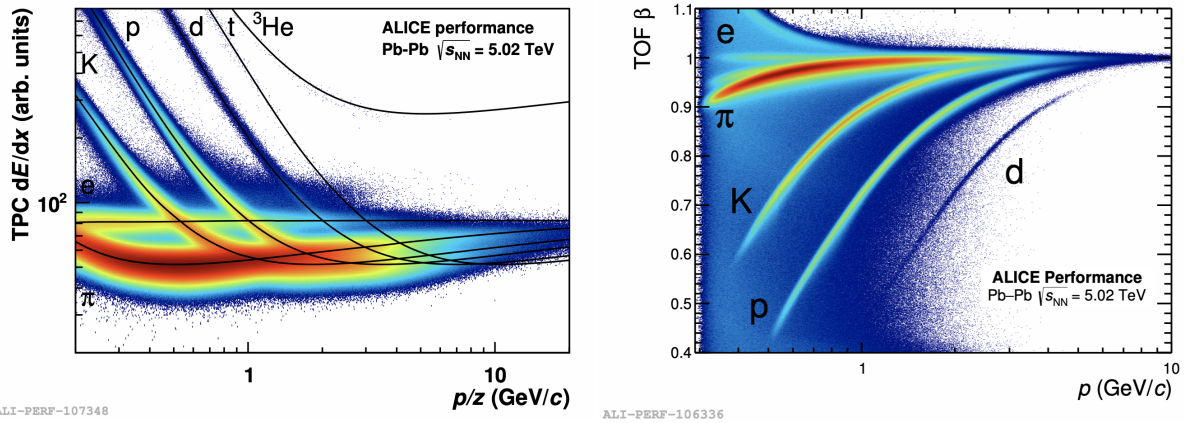


Figure 3.4: Typical particle identification performance in ALICE for Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV with low interaction ratio. The left panel shows the specific energy loss $\mathrm{d}E/\mathrm{d}x$ in the TPC as function of the magnetic rigidity for positive tracks, with superimposed the Bethe-Bloch calculations for specific particle species. The right panel shows the velocity β as measured with the TOF detector as function of momentum. The contribution of wrongly associated tracks can be seen outside the bands of the different particle species.

For the expected $\mathrm{d}E/\mathrm{d}x$, ALICE relies on a parametrisation of the Bette-Bloch formula with five parameters determined from fits to measured data [180]. In the left panel of Fig. 3.4, the measured specific energy loss in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV is shown as function of magnetic rigidity (p/z, with momentum p and charge number z) for seven particle species. The particle identification of pions and kaons (protons) can be performed on a track-by-track basis for $p \lesssim 0.7$ (1.5) GeV/c, where the largest separation is achieved. The PID information is typically expressed in numbers of σ from the expected $\mathrm{d}E/\mathrm{d}x$ resolution for a track with a given momentum. These excellent PID capabilities, in particular in the TPC detector, are one of the distinctive features of the ALICE detector, and of great importance of the performed analyses in this dissertation.

3.2.3 Time-Of-Flight detector

The main purpose of the Time-Of-Flight (TOF) detector [184] is the particle identification in the intermediate momentum range. It is a large array of Multigap Resistive Plate Chambers (MRPCs), positioned at 370–399 cm from the beam axis, covering the pseudorapidity range $|\eta| < 0.9$ and the full azimuthal angle. Any ionisation produced by a traversing charged particle in a chamber will lead, via the gas avalanche process, to observed signals on the pick-up electrodes. Each TOF module is a 10-gap double-stack MPRC strip, and has a time resolution of about 40 ps.

The PID is provided by the measurement of the time-of-flight of charged particles from the interaction point to the TOF. The start time t_0 is determined with the T0 detector discussed in the next section, or, alternatively, estimated using the arrival times in the TOF detector. In the latter case, a combinatorial algorithm based on a χ^2 minimisation between all mass hypotheses is used. The particle identification is based on the measured velocity (β) distribution in the TOF as function of the measured momentum by the TPC. It depends therefore on the matching efficiency

of TPC tracks to TOF hits, which reduces rapidly to zero for tracks with $p_{\mathrm{T}} < 0.7$ GeV/c due to energy loss, while at high p_{T} it is of the order of 70%, reflecting mainly the geometrical acceptance and inactive modules. The overall TOF resolution for tracks with $p = 1$ GeV/c is around 60–80 ps in Pb–Pb collisions, while it is 100 ps in proton–proton events because of the larger uncertainty on the determination of t_0. The PID capabilities of the TOF detector are shown in the right panel of Fig. 3.4, where the velocity β is plotted as function of momentum. Up to $p \simeq 4$ GeV/c, the TOF provides a separation between protons and kaons with more than 3σ, while kaons and pions can be distinguished up to around 2.5 GeV/c.

3.2.4 V0 and T0 detectors

The V0 detector [185] consists out of two arrays of 32 scintillator counters (called V0A and V0C), installed on both sides of the ALICE interaction point. Both V0 detectors are segmented in 4 radial rings and 8 azimuthal sectors, and located 340 cm (V0A) and 90 cm (V0C) away from the nominal primary vertex. They cover the pseudorapidity intervals $2.8 < \eta < 5.1$ and $-3.7 < \eta < -1.7$, respectively, and the full azimuthal angle. The V0 detector has several functions: triggering, rejection of beam–gas collisions, luminosity determination, and is used to define event characteristics like centrality or multiplicity. The latter one can be achieved since the signal produced, via scintillation light of traversing particles, is proportional to the number of particles hitting the detector. The triggers provided by the V0 arrays are exploiting AND logic between the V0A and V0C signal, used to define minimum-bias, high-multiplicity, mid-central, and central event triggers.

The T0 detector is used to determine the start time for the TOF detector and is composed of two arrays of Cherenkov counters, placed on about 375 cm (T0A) and 73 cm (T0C) on either side of the nominal interaction point, covering the full azimuth and the pseudorapidity intervals $4.61 < \eta < 4.92$ and $-3.28 < \eta < -2.97$, respectively. This detector achieves a 20–25 ps time resolution for Pb–Pb collisions and 40 ps for pp collisions.

3.2.5 Trigger system

The trigger decision, to store or to not store a collision, is generated by the Central Trigger Processor (CTP) of ALICE. It receives trigger inputs from the trigger detectors and, in case the trigger requirements are fulfilled, provides trigger signals to readout detectors. It evaluates the trigger inputs every ~ 25 ns, but the different readout times of each detector means it is divided in three levels. Fast trigger detectors — like the SPD, V0, and T0 — send Level-0 (L0) trigger inputs for which the L0 trigger decision is made about 0.9 µs after the collision. The accepted events are further processed by the Level-1 (L1) trigger algorithm, which expects trigger inputs from "slower" trigger detectors after around 6.5 µs. In case no trigger input is received by the CTP, the readout detectors stop registering the event. A third-level trigger (L2) is foreseen after 100 µs, which is approximate the maximum drift time of electrons in the TPC. Additional requirements can be set in the L2 trigger as well, to, e.g., avoid the presence of two sequential central Pb–Pb collisions. In case also the L2 trigger requirements are fulfilled, the event is send

to the Data Acquisition System (DAQ). The DAQ manages the data flow from the electronics up to the archiving on tape, and includes the High-Level Trigger (HLT) system, which rejects events that are not interesting after a fast track reconstruction and an online analysis. Events selected by the HLT are then registered on tape.

The triggers used in this work are the minimum-bias, two type of high-multiplicity, and two type of centrality triggers. Minimum-bias (MB) events are defined such that the events have as little bias as possible from the trigger condition, and requires signals in both V0A and V0C in coincidence with the proton bunch arrival time. To enrich the data statistics for pp collisions in the highest multiplicity regions, high-multiplicity triggers based on a minimum selection of fired chips in the SPD (HMSPD) and of V0 signal amplitudes (HMV0) were used. The centrality triggers are based on the V0 signal amplitudes, and adopted to enhance the statistics of central and mid-central events for Pb–Pb collisions.

3.3 Centrality determination

The centrality, as introduced in Sec. 1.4, is an experimental quantity to characterise a heavy-ion collision. It quantifies the overlap between the colliding nuclei using a percentage scale (where 0% is a head-on collision, and 100% means no overlap), and therefore relates to the — experimentally inaccessible — impact parameter b. Theoretically, the centrality c_{th} of an A–A collision with impact parameter b is defined by integrating the impact parameter distribution $\mathrm{d}\sigma/\mathrm{d}b$ like

$$c_{\mathrm{th}} = \frac{\int_0^b (\mathrm{d}\sigma/\mathrm{d}b')\mathrm{d}b'}{\int_0^\infty (\mathrm{d}\sigma/\mathrm{d}b')\mathrm{d}b'} = \frac{1}{\sigma_{\mathrm{tot}}} \int_0^b \frac{\mathrm{d}\sigma}{\mathrm{d}b'}\mathrm{d}b', \tag{3.1}$$

where σ_{tot} is the total hadronic cross section. As explained in Sec. 1.4, this is mapped with the experimental definition of centrality, providing access to theoretical quantities like b, N_{part}, and N_{coll} for experimental data. The centrality c_{exp} in Pb–Pb collisions in ALICE is defined as the percentile of the hadronic cross section corresponding to a particle multiplicity above a given threshold [37]

$$c \approx \frac{1}{\sigma_{\mathrm{tot}}} \int_{N_{\mathrm{ch}}^{\mathrm{thr}}}^\infty \frac{\mathrm{d}\sigma}{\mathrm{d}N'_{\mathrm{ch}}}\mathrm{d}N'_{\mathrm{ch}}. \tag{3.2}$$

Here, $\mathrm{d}\sigma$ is replaced by the number of observed events $\mathrm{d}n$ with a given multiplicity and σ_{tot} by the total number of observed events, corrected for the trigger efficiency and determined in a region where the non-hadronic interactions are negligible. The used multiplicity estimator for Pb–Pb collisions in ALICE is the total signal measured in the V0 detectors, which is a proxy of N_{ch}. The V0 signal is normalised at an anchor point, which corresponds to 90% of the total hadronic cross section derived from a fit of the distribution from Glauber Monte Carlo simulations [36] combined with a simple model for particle production (NBD) [37, 186] to the measured distribution. The events with multiplicities lower than those from the anchor point are not considered for the centrality determination since they will be contaminated by electromagnetic background events. In order to relate the centrality classes to the collision geometry and decompose the particle production into the contributions due to soft and hard

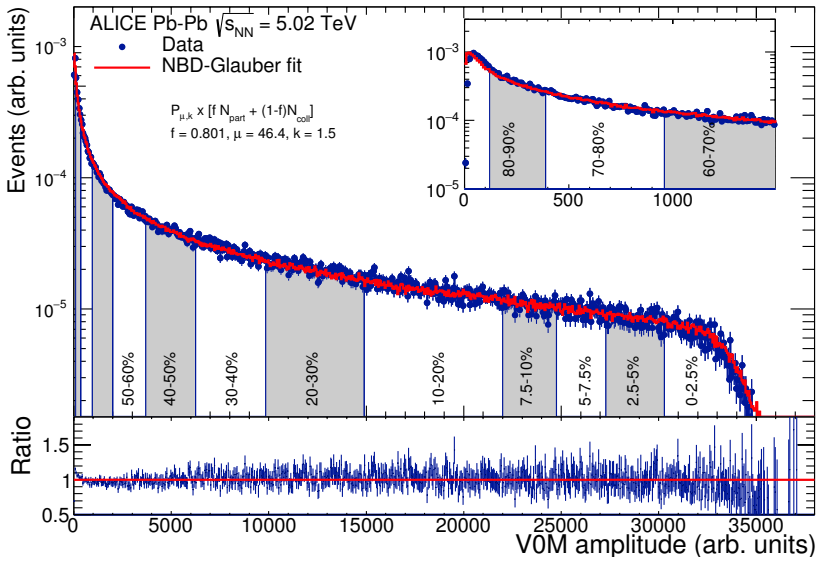


Figure 3.5: Distribution of the total V0 signal, i.e. the sum of the amplitudes in both the V0A and V0C detector, for Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV [37]. The distribution is fitted with the NBD-Glauber fit.

interactions, the same NBD-Glauber model fit can be used. Figure 3.5 shows the distribution of the total signal amplitude in the V0 detectors fitted with such NBD-Glauber function for Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV [37]. The resolution on the determined centrality limits is smaller than 0.1% for central collisions to about 2% for mid-central collisions.

3.4 Track and vertex reconstruction

The procedures for the track reconstruction and vertex finding are performed offline starting from the clusterisation step, which converts the raw detector data into "clusters", separately for each detector. Clusters are characterised by variables like position, signal amplitude, and signal times. From the information of the SPD clusters, a preliminary interaction vertex is determined as next step. This is the space point that minimises the distance among the tracklets, the track segments reconstructed by associating pairs of clusters in the two SPD layers. In case no single point is found, for example in low multiplicity pp events, the algorithm performs an one-dimensional minimisation along the beam axis instead. On the other hand, because of interaction pile up in pp collisions, multiple vertices can be expected as well, which are determined by repeating the algorithm several times after discarding the clusters that already contributed.

The next step is the track reconstruction, which is performed in three stages, following an inward–outward–inward scheme. The first inward stage starts by finding tracks in the TPC, looking first at large radii at the edge of the TPC. A track "seed" starts from two such TPC clusters and the preliminary interaction vertex[3], after which they are updated with every nearest cluster that is found by the Kalman Filter algorithm [187] until the TPC inner radius is reached. The number of TPC clusters that is used to built a track is a quality criteria of the track. The

[3]The constraint from this preliminary interaction point is removed when more than 2 clusters are found.

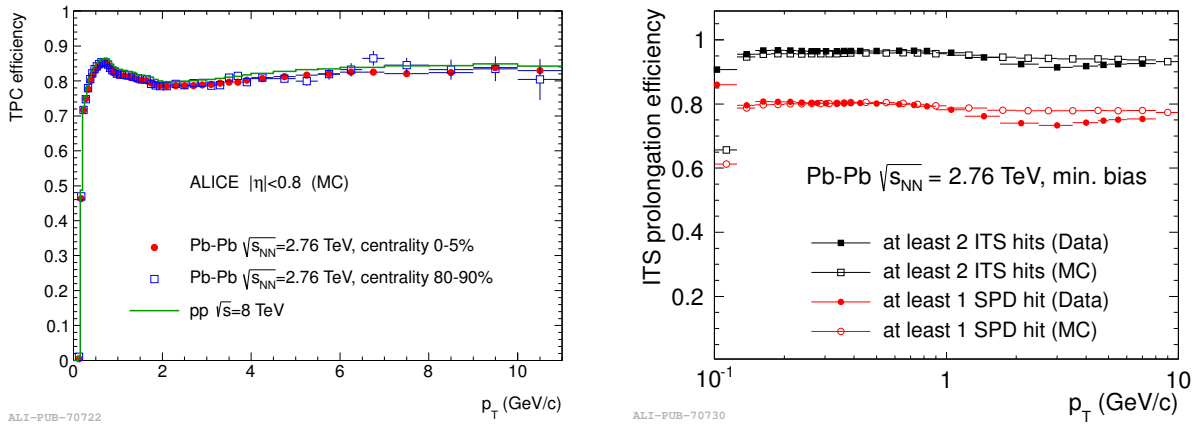


Figure 3.6: Left: The TPC tracking efficiency for primary particles as function of p_{T} for simulated pp and Pb–Pb collisions at $\sqrt{s} = 8$ TeV and $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV, respectively [180]. The central and peripheral centrality classes indicate that the efficiency does not depend on the detector occupancy. Right: The ITS-TPC matching efficiency as function of p_{T} for data and simulations for Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV [180].

reconstruction efficiency for TPC tracks in simulated pp and Pb–Pb collisions is shown in the left panel of Fig. 3.6 [180]. From the distributions it is clearly visible that the efficiency is independent of the detector occupancy. The efficiency drops for tracks with $p_{\mathrm{T}} < 0.5$ GeV/c caused by energy loss in the detector material.

The TPC tracks are matched to clusters in the outermost SSD layer, which become the seed for the track finding in the ITS. A similar procedure is adapted, where the seed is propagated inwards, attaching at each ITS layer all the ITS clusters within a proximity cut. This means that each TPC track produces a "tree" of track hypotheses in the ITS. The highest quality candidate (based on the χ^2) from each hypothesis tree is in the end added to the reconstructed TPC track. The right panel of Fig. 3.6 shows the fraction of TPC tracks that have a prolongation in the ITS, the so-called matching efficiency, for real and simulated Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} =$ 2.76 TeV [180]. In case at least two ITS (one SPD) clusters are required, the ITS-TPC matching efficiency is about 95% (75%) for central Pb–Pb collisions.

The obtained tracks are then propagated outwards to match to clusters in the TRD and TOF, and the signals in the EMCal, PHOS, and HMPID detectors. Finally, another inward step is performed, refitting the found tracks using all the previously found clusters, which is then extrapolated to the distance of closest approach to the interaction vertex. After the full tracking procedure, the final position of the primary vertex is calculated in case two or more tracks were reconstructed, which improves the resolution by about a factor 2.5 with respect to the preliminary primary vertex. The resolution (and the related variables like the impact parameter) improves with increasing charged-particle multiplicity by roughly the square root of the number of contributing tracks, from around 200 µm for low-multiplicity pp collisions ($\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \approx 2$), up to 50 µm for high-multiplicity pp collisions ($\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \approx 40$) and 10 µm for Pb–Pb collisions. The resolution on the impact parameter improves as well with increasing p_{T}, from about 60 µm at $p_{\mathrm{T}} = 1$ GeV/c to 20 µm at 10 GeV/c for charged particles in Pb–Pb collisions, which makes the topological selection of high-p_{T} weakly-decaying particles more effective [180].

3.5 The ALICE offline framework

The offline framework of ALICE is based on the ROOT software [188], designed specifically for high-energy physics experiments to deal with big data processing, statistical analyses, and data visualisation and storage. The ALICE offline framework is used for simulations, alignment and calibration of detectors, reconstruction of events, and for performing the actual analyses on the collected experimental data. Simulations are generated with Monte Carlo event generators, where the PYTHIA [95, 189] generators for pp collisions and the HIJING [190] generator for p–Pb and Pb–Pb collisions are typically used for the analyses in this dissertation. A detailed description of the detector is included in the simulations, so detector responses to the passage of particles are also simulated via transport packages like GEANT [191, 192]. The simulations are anchored to actual data-taking runs, to adopt the same detector and data-taking conditions.

The reconstructed data (both real and simulated) are stored in so-called ESD (Event Summary Data) ROOT files, containing all the physical information required for analyses and quality assurance for each subdetector. To reduce the output size, and therefore the required computational time when analysing such file, AOD (Analysis Object Data) ROOT files are produced, containing only the information needed for analyses. All computational tasks (e.g. generating simulations, reconstruction of data, or performing analyses) are performed on the worldwide LHC computing grid, an infrastructure connecting several computing centres from more than 40 countries.

CHAPTER 4

Analysis methods

In this dissertation, the D^{*+}, D_s^+, and B_s^0 mesons and the Λ_c^+ baryon (and their charge conjugates[1]) are investigated, all composed out of a heavy quark and one or two light quarks. These hadrons have a mean proper decay length of at most a few hundreds of micrometers and therefore cannot be revealed directly since they decay before reaching the first detector layer. They are detected, instead, via an exclusive reconstruction of the following hadronic decay channels: $D^{*+} \to D^0\pi^+ \to K^-\pi^+\pi^+$, $D_s^+ \to \phi\pi^+ \to K^+K^-\pi^+$, $\Lambda_c^+ \to pK_S^0 \to p\pi^+\pi^-$, and $B_s^0 \to D_s^-\pi^+ \to \phi\pi^-\pi^+ \to K^+K^-\pi^-\pi^+$. In order to suppress the large combinatorial background, selections on the properties of these decay channels are required. Besides exploiting the particle identification capabilities of ALICE, an important selection opportunity comes from the fact that with the spatial resolution of the ALICE detector it is possible to reconstruct the secondary decay vertex. For most of the here considered heavy-flavour hadrons, the precision on this secondary vertex is high enough to exploit the separation with the primary interaction vertex as selection criteria.

In this chapter, the data samples as well as the analysis strategies and methods that were used for the measurements presented in this dissertation will be introduced. Even though four separate analyses with four different heavy-flavour hadron species on four different data samples were performed, several parts are in common, hence this dedicated chapter. For all details, or analysis-specific methods, I refer to the next four chapters.

4.1 Data samples

The data analysed in this manuscript were collected by the ALICE detector in LHC Run 2, during the 2016, 2017, and 2018 pp runs at $\sqrt{s} = 5.02$ TeV and $\sqrt{s} = 13$ TeV and the 2018 Pb–Pb run at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. All data samples were collected with the minimum-bias trigger, although enhanced with the two high-multiplicity triggers and the two centrality triggers for pp collisions at $\sqrt{s} = 13$ TeV and Pb–Pb at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, respectively. The characteristics of the data samples are summarised in Tab. 4.1.

Table 4.1: Summary of the characteristics of the used data samples for the analyses presented in this dissertation. The numbers for the Pb–Pb data samples are after selecting a centrality interval of 0–10% and 30–50%.

System	$\sqrt{s_{\mathrm{NN}}}$	Trigger	N_{events}	$\mathcal{L}_{\mathrm{int}}$ ($\mu\mathrm{b}^{-1}$)	Year
pp	5.02	MB	990×10^6	19300 ± 400	2017
	13	MB	1800×10^6	32000 ± 1600	2016–2018
	13	HMV0	1000×10^6	n.a.	2016–2018
	13	HMSPD	300×10^6	n.a.	2018
Pb–Pb	5.02	MB or central	100×10^6	130.5 ± 0.5	2018
	5.02	MB or mid-central	85×10^6	55.5 ± 0.2	2018

In addition, dedicated Monte Carlo simulations with a detailed description of the ALICE apparatus geometry, detector response (via the GEANT3 transport package [191]), and anchored to

[1] In the remaining of the text, the charge conjugates are always considered, unless explicitly mentioned otherwise.

each data-taking period to ensure the same data-taking conditions, were used by each analysis to, e.g., optimise the selection criteria and compute the acceptance and efficiency corrections. For the $\sqrt{s} = 5.02$ TeV pp analysis, the PYTHIA 6.4.25 event generator [189] with the Perugia-11 tune [150] was adopted, where for $\sqrt{s} = 13$ TeV pp analysis, the more recent PYTHIA 8.243 event generator [95] with the Monash 2013 tune [151] was used. In order to achieve a good statistical precision, each simulated pp event contained at least one $\mathrm{c\overline{c}}$ or $\mathrm{b\overline{b}}$ pair and the charm hadrons were forced to decay in their decay channel of interest. The decays of the beauty hadrons were not forced, to not introduce biases in the estimation of the feed-down charm-hadron selection and reconstruction efficiencies. Different simulations were used for the training of machine-learning algorithms (see Sec. 4.2.4), where only the events were stored that had at least once the charm-hadron decay of interest produced within $|y| < 1$.

In case of heavy-ion collisions, the HIJING 1.36 generator [190] was used to simulate the underlying Pb–Pb collision, while the charm-hadron signals were injected via PYTHIA 8.243 pp events as described above. The number of injected PYTHIA events increased with increasing collision centrality. Since the $\mathrm{B_s^0}$ analysis is not a real data analysis, but a performance study for upgraded ITS setups, the simulations have some unusual features. These will be discussed in more detail in the dedicated Chapter 8.

4.1.1 Offline event selection

In addition to the online trigger selections discussed in Sec. 3.2.5, an offline event selection is performed to reject background coming from interactions between one of the beams and the residual gas present in the beam vacuum tube. This is performed exploiting the timing information of the V0 detectors and the correlation between the number of hits and tracklets in the SPD detector. Furthermore, in order to have a uniform acceptance in pseudorapidity, collision vertices were required to be within ± 10 cm (in the z-direction) of the nominal interaction point. This guarantees the pseudorapidity coverage of $|\eta| < 0.8$ in each ITS layer, as well as a symmetric geometrical acceptance. Finally, to reduce the superposition of more than one collision within the colliding bunches (so-called in-bunch *pile-up* events), events with multiple reconstructed primary vertices were rejected. In particular, pile-up events are rejected in case another vertex is found with at least 5 contributors and at least 0.8 cm away from the main primary vertex. This is especially important for pp collisions, while the probability for in-bunch pile-up events for Pb–Pb collisions is negligible because of the lower interaction rate. Out-of-bunch pile-up events, where one or more collisions occur in bunch crossings different from the one which triggered the data acquisition, are rejected by the request of at least a hit in one of the two innermost layers of the ITS, which, as we will see later, is required for at least one of the decay daughters in these analyses.

These offline event selections reduce the primary-vertex reconstruction efficiency, which will be taken into account when calculating the reconstruction efficiency of the heavy-flavour hadrons. However, for a correct normalisation, the number of events where no vertex could be reconstructed within the $|z| < 10$ cm interval has to be taken into account as well. The proper

normalised number of events is obtained via

$$\begin{aligned} N_{\mathrm{ev}}^{\mathrm{norm}} &= N^{\mathrm{reco\,vtx}}\big|_{|z_{\mathrm{vtx}}|\leq 10\,\mathrm{cm}} + N^{\mathrm{no\,vtx}}\big|_{|z_{\mathrm{vtx}}|\leq 10\,\mathrm{cm}} \\ &= N^{\mathrm{reco\,vtx}}\big|_{|z_{\mathrm{vtx}}|\leq 10\,\mathrm{cm}} + N^{\mathrm{no\,vtx}} - N^{\mathrm{no\,vtx}}\big|_{|z_{\mathrm{vtx}}|>10\,\mathrm{cm}} \\ &= N^{\mathrm{reco\,vtx}}\big|_{|z_{\mathrm{vtx}}|\leq 10\,\mathrm{cm}} + N^{\mathrm{no\,vtx}} - N^{\mathrm{no\,vtx}} \cdot N^{\mathrm{reco\,vtx}}\big|_{|z_{\mathrm{vtx}}|>10\,\mathrm{cm}} / N^{\mathrm{reco\,vtx}}, \end{aligned} \tag{4.1}$$

where $N^{\mathrm{reco\,vtx}}$ and $N^{\mathrm{no\,vtx}}$ are the number of events with or without a primary vertex reconstructed, respectively. Overall, this correction has a 10% effect for the pp data samples, where for Pb–Pb it is a negligible contribution. The final (corrected) number of events after the offline event selections, together with the corresponding integrated luminosity ($\mathcal{L}_{\mathrm{int}}$), are reported in Tab. 4.1. For the Pb–Pb data samples, the central (0–10%) and mid-central (30–50%) centrality selections are applied. Because of the trigger turn-on curves, it is not trivial to calculate an integrated luminosity for the HMV0 and HMSPD data samples, hence no value is provided in Tab. 4.1. Considering only the data periods for which the triggers are fully efficient in the multiplicity region of interest (see Chapter 6), an integrated luminosity of about 8.5 pb^{-1} and 0.8 pb^{-1} can be estimated for the HMV0 and HMSPD triggered data samples, respectively.

4.2 Reconstruction and selection of heavy-flavour hadrons

The D^{*+}, $\mathrm{D}_{\mathrm{s}}^{+}$, and $\mathrm{B}_{\mathrm{s}}^{0}$ mesons and the Λ_{c}^{+} baryon are measured via an exclusive reconstruction of the following hadronic decay channels: $\mathrm{D}^{*+} \to \mathrm{D}^{0}\pi^{+} \to \mathrm{K}^{-}\pi^{+}\pi^{+}$, $\mathrm{D}_{\mathrm{s}}^{+} \to \phi\pi^{+} \to \mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}$, $\Lambda_{\mathrm{c}}^{+} \to \mathrm{pK}_{\mathrm{S}}^{0} \to \mathrm{p}\pi^{+}\pi^{-}$, and $\mathrm{B}_{\mathrm{s}}^{0} \to \mathrm{D}_{\mathrm{s}}^{-}\pi^{+} \to \phi\pi^{-}\pi^{+} \to \mathrm{K}^{+}\mathrm{K}^{-}\pi^{-}\pi^{+}$. The main properties of these hadrons and decay channels are summarised in Tab. 4.2. The meson candidates are built by combining pairs or triplets of tracks with the proper charge-sign combination, starting with the "intermediate" hadron states D^{0} and $\mathrm{D}_{\mathrm{s}}^{-}$ for the D^{*+} and $\mathrm{B}_{\mathrm{s}}^{0}$ mesons, respectively. For these combination of tracks, the decay vertex is reconstructed as the space point that minimises the "distance" $\mathcal{D}$ between the considered tracks

$$\mathcal{D} = \sqrt{\sum_{i=1}^{N_{\mathrm{prongs}}} \left[\left(\frac{x_i - x_0}{\sigma(x_i)} \right)^2 + \left(\frac{y_i - y_0}{\sigma(y_i)} \right)^2 + \left(\frac{z_i - z_0}{\sigma(z_i)} \right)^2 \right]}, \tag{4.2}$$

where N_{prongs} is the number of decay tracks, and (x_i, y_i, z_i) the coordinates with their uncertainties σ of track i at the distance of closest approach with respect to the reconstructed decay vertex (x_0, y_0, z_0). For the $\mathrm{B}_{\mathrm{s}}^{0}$ meson, this process is repeated by pairing the $\mathrm{D}_{\mathrm{s}}^{-}$ candidate with a positive-charge track (bachelor) and reconstructing another decay vertex. Also the D^{0} candidate is combined with a positive-charge track to form the D^{*+} candidate, only in this case no additional decay vertex is reconstructed because of the very small $c\tau$ of the D^{*+} decay, hence the decay-vertex of the D^{*+} cannot be resolved from the primary vertex. For the $\Lambda_{\mathrm{c}}^{+} \to \mathrm{pK}_{\mathrm{S}}^{0}$, the reconstruction starts from a $\mathrm{K}_{\mathrm{S}}^{0} \to \pi^{+}\pi^{-}$ candidate, a so-called V^{0}-decay, which is reconstructed as a pair of opposite-sign charged tracks forming a neutral decay vertex displaced from the primary vertex. The V^{0}-candidate is paired with a proton-candidate track. Because of the relatively large uncertainty on the $\mathrm{K}_{\mathrm{S}}^{0}$ track close to the interaction vertex and, in addition,

Table 4.2: Summary of the main properties of the analysed hadrons and their studied decay channels [16]. The $c\tau$ of the intermediate hadron states in the exploited decay channels are 122.9 µm for the D^0 meson, 46.4 fm for the ϕ meson, and 2.68 cm for the K^0_S. The PDG reports the resonance width (Γ) for the D^{*+} and ϕ meson, where here the $\tau = 1/\Gamma$ conversion is used.

Hadron	Mass (MeV/c^2)	$c\tau$ (µm)	Decay channel	$\Gamma_i/\Gamma_{\text{total}}$ (%)
D^{*+} $(c\overline{d})$	2010.26 ± 0.05	$2.37 \cdot 10^{-6}$	$D^0[\rightarrow K^-\pi^+]\pi^+$	2.67 ± 0.02
D_s^+ $(c\overline{s})$	1968.34 ± 0.07	151.2	$\phi(1020)[\rightarrow K^+K^-]\pi^+$	2.24 ± 0.08
Λ_c^+ (cdu)	2286.46 ± 0.14	60.7	$K^0_S[\rightarrow \pi^+\pi^-]p$	1.10 ± 0.06
B_s^0 $(s\overline{b})$	5366.88 ± 0.14	454.2	$D_s^-\pi^+$	$(6.72 \pm 0.01) \cdot 10^{-3}$

the comparable track impact-parameter resolution of the ITS and the mean decay length of the Λ_c^+, this analysis cannot fully benefit of the reconstruction and selection of secondary vertex topologies, and a decay-vertex reconstruction was therefore not exploited.

The momenta of the reconstructed heavy-flavour hadrons follow from momentum conservation applied on the point along each track closest to the decay vertex (primary vertex in case of the D^{*+} and Λ_c^+). The invariant mass of each candidate is calculated as

$$M^2_{\text{inv}} = \left(\sum_{i=1}^{N_{\text{prongs}}} E_i\right)^2 - \left\|\sum_{i=1}^{N_{\text{prongs}}} \vec{p_i}\right\|^2 = \left(\sum_{i=1}^{N_{\text{prongs}}} \sqrt{m_i^2 + \|\vec{p_i}\|^2}\right)^2 - \left\|\sum_{i=1}^{N_{\text{prongs}}} \vec{p_i}\right\|^2, \qquad (4.3)$$

where m_i is the used mass hypothesis for the decay tracks. For most of the here considered decay chains, the ambiguity in assigning the mass is resolved by the charge of the bachelor tracks. Only for the two tracks having the same charge sign as the D_s^+ candidate, both the pion- and kaon-mass hypothesis are considered. The contribution of signal D_s^+ candidates with the wrong decay-particle mass assignment (so-called *reflections*) present in the invariant-mass peak region are, after the selections discussed below, negligible.

The considered decay channels in this dissertation all have relatively low branching ratios ($\Gamma_i/\Gamma_{\text{total}}$), while the reconstruction procedure will cause a large number of combinatorial background candidates. It is therefore of crucial importance to increase the signal-to-background ratio and the statistical significance of the measurements. Rather strict selection criteria on the track-quality, decay topology and kinematics, and particle identification of the decay tracks are therefore required. All three categories will be discussed in more detail in the subsections below.

4.2.1 Selections on track quality

Track-quality selections were applied to the decay tracks of the heavy-flavour candidates. In particular (except for the soft-pion track for the D^{*+} and the V^0-decay tracks for the Λ_c^+), only tracks with $|\eta| < 0.8$, $p_{\text{T}} > 0.3$ GeV/c (0.5 for Pb–Pb), at least 70 out of 159 associated crossed TPC pad rows, a ratio of crossed rows to findable clusters in the TPC larger than 0.8, a $\chi^2/\text{ndf} < 2$ (2.5 for Pb–Pb) in the TPC, and a minimum number of two hits (out of six) in the ITS, with at least one in the SPD, were considered. This last requirement is a good compromise between the impact-parameter and decay-vertex resolution, and the candidate-reconstruction

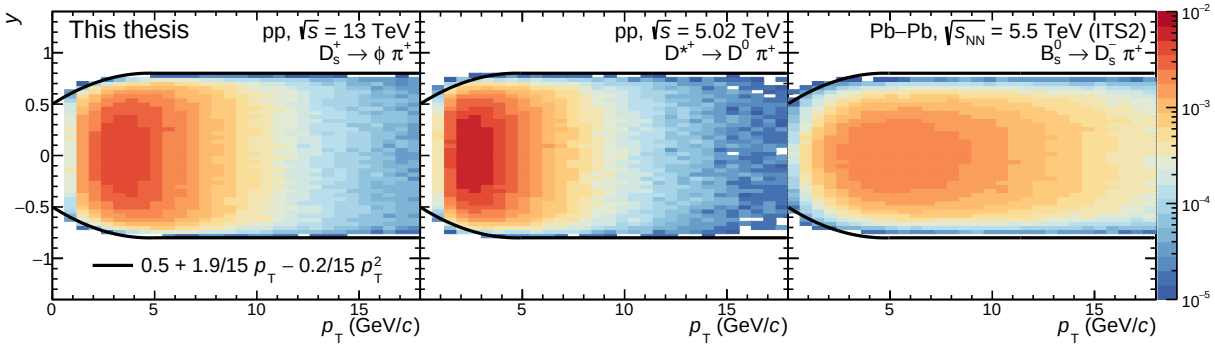


Figure 4.1: The rapidity as function of transverse momentum of reconstructed $\mathrm{D_s^+}$ (left), $\mathrm{D^{*+}}$ (middle), and $\mathrm{B_s^0}$ (right) mesons in different collision systems. The solid lines define the fiducial acceptance region $y_{\mathrm{fid}}(p_{\mathrm{T}})$.

efficiency. For the soft-pion candidate of the $\mathrm{D^{*+}}$ decay, tracks with $|\eta| < 0.9$, $p_{\mathrm{T}} > 80$ MeV/c, and two points in the ITS, with at least one in the SPD were considered instead. In addition, the requirement of a successful refit in the TPC from the track reconstruction stage was removed, because its momentum is in general too low to cross the full TPC (at least 200 MeV/c). For the $\mathrm{K_S^0}$ daughters in the Λ_{c}^+ decay, the ITS-refit and ITS-cluster requirements were released, while $p_{\mathrm{T}} > 100$ MeV/c and the same TPC track-quality selections as described above were used. Finally, in Pb–Pb collisions, a selection on the impact parameter of tracks with $p_{\mathrm{T}} < 2$ GeV/c of $|d_0| > 25$ µm (15 µm for the upgraded ITS geometry) is applied in order to select only displaced tracks in this kinematic region to reduce the amount of combinatorial background. Moreover, at least 50 clusters available for PID in the TPC were required for the 2018 Pb–Pb data sample, to reduce the effects of the not sufficiently simulated track reconstruction at the TPC sector edges.

As a consequence of these track selection criteria, the detector acceptance for the heavy-flavour hadrons varies as a function of rapidity. For this reason, a fiducial acceptance selection was applied on the rapidity of the candidates, $|y| < y_{\mathrm{fid}}(p_{\mathrm{T}})$, where the factor $y_{\mathrm{fid}}(p_{\mathrm{T}})$ was defined as a second-order polynomial function, increasing from 0.5 to 0.8 in the transverse momentum range $0 < p_{\mathrm{T}} < 5$ GeV/c, and a constant term, $y_{\mathrm{fid}} = 0.8$, for $p_{\mathrm{T}} > 5$ GeV/c. This is illustrated in Fig. 4.1 for $\mathrm{D_s^+}$, $\mathrm{D^{*+}}$, and $\mathrm{B_s^0}$ mesons for different collision systems. The different trends for the different mesons are because of the looser track-quality selections used for the soft pion track in the $\mathrm{D^{*+}}$ analysis, and the upgraded ITS geometry exploited for the $\mathrm{B_s^0}$ analysis (see Chapter 8 for more details). Given that these heavy-flavour analyses are typically performed for multiple hadrons at the same time (and the upgraded ITS analyses are still performance studies), the same fiducial acceptance selection is used, defined by the second-order polynomial function

$$y_{\mathrm{fid}}(p_{\mathrm{T}}) = 0.5 + \frac{1.9}{15} p_{\mathrm{T}} - \frac{0.2}{15} p_{\mathrm{T}}^2. \tag{4.4}$$

4.2.2 Selections on decay topology and kinematics

The decay channels studied in this dissertation are all characterised by a weakly-decaying heavy-flavour hadron in the chain, and therefore have proper decay lengths $c\tau$ greater than ($\mathrm{D^0}$, $\mathrm{D_s^+}$,

B_s^0) or of the order of (Λ_c^+) the impact parameter resolution of the ALICE detector. Combinatorial background candidates are, instead, predominantly built from tracks originating from the primary vertex. The main criteria to discriminate heavy-flavour signal from the combinatorial background are therefore based on the displaced decay topologies. In addition, kinematic variables related to the decay chain are also exploited, which, especially for the $\Lambda_c^+ \rightarrow pK_S^0$ decay channel, will help to improve the selection of signal. All important selection variables for the four decay chains will be briefly discussed below, showing as well examples of the distributions for prompt and feed-down signal[2] and combinatorial background. The distributions for prompt and feed-down signal are taken from simulations, where the combinatorial background distributions are extracted from candidates in the sidebands (excluding the signal region) of the invariant-mass distribution in data.

- **Decay length** (L, L_{xy})**:** The decay length is defined as the distance between the primary and decay vertices in three-dimensional space (L) or in the transverse plane (L_{xy}). The projection in the transverse plane is considered since the resolution in the x and y coordinates is better than in the z coordinate. The effect from the curvature in their trajectories due to the magnetic field can safely be neglected for heavy-flavour hadrons because of the proper decay lengths of only several hundreds of micrometers.

  Typical distributions for the decay length are shown in the left (for D_s^+ in pp collisions) and middle panel (for B_s^0 in Pb–Pb collisions with the upgraded ITS setup) of Fig. 4.2, where the signal distributions are "shifted" towards higher values. Because for feed-down D_s^+ candidates, the decay length includes as well the distance travelled by prior beauty hadrons, it is on average more displaced than prompt D_s^+ candidates. The non-zero displacement for combinatorial background is only due to the detector resolution, which will be improved significantly with the upgraded ITS detector in LHC Run 3-4, hence the improved separation power between signal and background B_s^0 candidates. The decay length is enlarged for high-momentum heavy-flavour hadrons because of the Lorentz boost.

  Selections on the decay length are used by the D^{*+} (via the D^0 decay length), the D_s^+, and the B_s^0 (via both the B_s^0 and D_s^+ decay lengths) analyses. As mentioned before, the decay vertex for the $\Lambda_c^+ \rightarrow pK_S^0$ decay is not reconstructed and can therefore not be exploited. The decay length (or $c\tau$) of the K_S^0 candidate is used, but has only limited separation power as we will see later, because the proper decay length of the Λ_c^+ is negligible with respect to the one of the K_S^0.

- **Normalised decay length in the transverse plane** ($L_{xy}/\sigma(L_{xy})$)**:** This is defined as the projected decay length in the transverse plane L_{xy} divided by its uncertainty $\sigma(L_{xy})$. It is considered since it provides a better background rejection than the nominal decay length variables, especially at low p_T. The right panel of Fig. 4.2 shows the normalised

[2]Prompt signal candidates originate directly from the hadronisation of a charm (beauty in case of B_s^0) quark or from a strongly-decaying charm (beauty) excited hadron, while feed-down signal candidates are coming from beauty-hadron decays, and therefore only relevant for the charm hadrons.

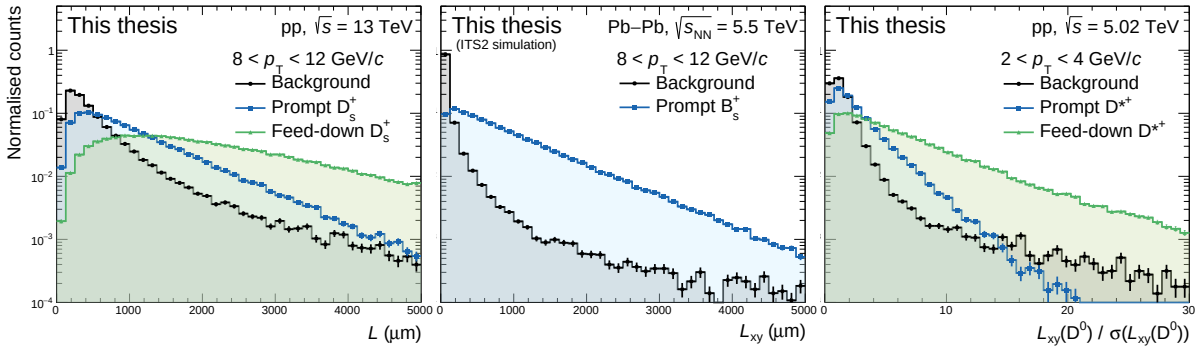


Figure 4.2: Left: Decay-length distribution of prompt D_s^+, feed-down D_s^+, and combinatorial-background candidates in pp collisions at $\sqrt{s} = 13$ TeV for $8 < p_T < 12$ GeV/c. Middle: Distribution of the decay-length projected in the transverse plane for B_s^0 and combinatorial-background candidates for $8 < p_T < 12$ GeV/c in simulated Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV with the upgraded ITS setup for LHC Run 3. Right: Normalised decay-length distribution in the transverse plane for prompt, feed-down, and combinatorial-background D^0 candidates coming from D^{*+} decays in pp collisions at $\sqrt{s} = 5.02$ TeV for $2 < p_T < 4$ GeV/c.

4

decay-length distribution for the D^0 daughters of low-p_T D^{*+} candidates in pp collisions. This variable is considered in all three meson analyses.

- **Cosine of pointing angle** ($\cos\theta_p$, $\cos\theta_p^{xy}$)**:** The pointing angle θ_p is the angle between the line connecting the primary and decay vertices and the momentum vector of the reconstructed candidate. To better exploit the x- and y-coordinate resolutions, a projection in the transverse plane (θ_p^{xy}) is used as well. The cosine of the pointing angle should be exactly 1 for prompt and about 1 for feed-down signal candidates in case of a perfect vertex reconstruction. Resolution effects, however, smear out the distributions to lower values. For combinatorial background, instead, the pointing angle does not have to point to the primary vertex, hence the $\cos\theta_p$ distribution will be flatter. The distributions for high-p_T B_s^0 and low-p_T D_s^+ mesons are shown in the left and middle panel of Fig. 4.3, showing the expected ordering around $\cos\theta_p \sim 1$ and the improved separation power with the ITS2 setup. All four analyses exploit this selection variable, where for the D^{*+} and Λ_c^+ decay channels, the decay vertex of the D^0 and K_S^0 mesons are used, respectively.

- **Cosine theta star** ($\cos\theta^*$)**:** The θ^* angle is the angle between the flight line of the reconstructed particle, i.e. the line connecting the primary and decay vertices in the lab frame, and one of the decay tracks in the rest frame of the reconstructed heavy-flavour hadron. As shown in the right panel of Fig. 4.3 for D^0 candidates from D^{*+} decays, the distribution for combinatorial background peaks around $\cos\theta^* \sim 1$, while the signal distributions are more flat. This selection variable is only used in the D^{*+} analysis.

- **Impact parameter in the transverse plane** (d_0^{xy})**:** The impact parameter of the reconstructed heavy-flavour hadron in the transverse plane is defined as the distance of closest approach between the reconstructed flight line (or measured track in case one considers a decay daughter) and the primary vertex position, projected in the transverse plane. Again, the detector resolution will give rise to a smearing effect, making the impact parameter non-zero for prompt heavy-flavour hadrons. The distribution will get narrower

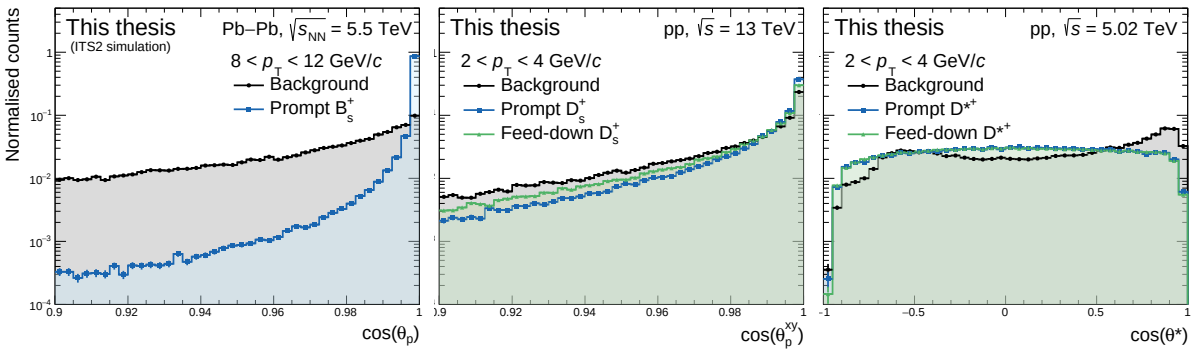


Figure 4.3: Left: Cosine of pointing angle for B_s^0 and combinatorial-background candidates for $8 < p_T < 12$ GeV/c in simulated Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV with the upgraded ITS setup for LHC Run 3. Middle: Cosine of pointing angle distribution in the transverse plane of prompt D_s^+, feed-down D_s^+, and combinatorial-background candidates in pp collisions at $\sqrt{s} = 13$ TeV for $2 < p_T < 4$ GeV/c. Right: Cosine θ^* distribution for prompt, feed-down, and combinatorial-background D^0 candidates coming from D^{*+} decays in pp collisions at $\sqrt{s} = 5.02$ TeV for $2 < p_T < 4$ GeV/c.

for higher p_T and, in general, for LHC Run 3 because of the improved tracking detectors. Besides the possibility to reject background, as shown in the left panel of Fig. 4.4 for D_s^+ candidates, it can be used to suppress the contribution of feed-down D_s^+ mesons as well, which has a much wider distribution.

All four analyses are applying selections on the d_0^{xy}-related variables. While for the heavy-flavour mesons the most powerful is the impact parameter for the reconstructed candidate, the impact parameter of the proton track is the only accessible (and useful) one for the $\Lambda_c^+ \to pK_S^0$ analysis.

- **Product of impact parameters in the transverse plane** ($d_{0,i}^{xy} \cdot d_{0,j}^{xy}$)**:** In addition to the normal impact parameter, the D^{*+} (via the D^0 meson) and B_s^0 analyses apply as well a selection on the product of the impact parameters of their decay tracks (between the K^- and π^+ daughter tracks for the D^0 and the D_s^- and π^+ tracks for the B_s^0 meson). Because these are neutral hadrons with a two-prong decay, the multiplication of the impact parameters will give rise to an asymmetry for signal, while it is on average symmetric for combinatorial background. This is visible in the middle panel of Fig. 4.4 for D^0 candidates from D^{*+} decays in pp collisions.

- **Maximum normalised difference between the measured and the expected impact parameters in the transverse plane** ($\max\{(d_{0,i}^{reco} - d_{0,i}^{exp})/\sigma_\Delta\}$)**:** This selection is, like for the normal impact parameters, an effective way to reject feed-down and combinatorial background candidates. The selection is applied for decay daughter i to the maximum value of $(d_{0,i}^{reco} - d_{0,i}^{exp})/\sigma_\Delta$, where $d_{0,i}^{reco}$ is the measured impact parameter of track i and $d_{0,i}^{exp}$ and σ_Δ are defined as

$$\begin{aligned} d_{0,i}^{exp} &= L_{xy} \cdot \sin\theta_i^{xy}, \\ \sigma_\Delta &= \sqrt{\sigma^2_{d_{0,i}^{xy}(meas)} + \sigma^2_{d_{0,i}^{xy}(exp)}}. \end{aligned} \tag{4.5}$$

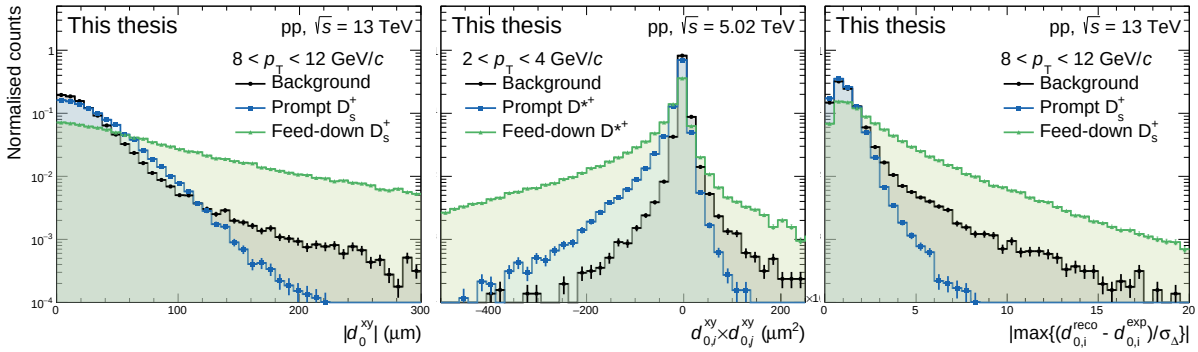


Figure 4.4: Left: The absolute value of the impact-parameter distribution in the transverse plane of prompt $\mathrm{D_s^+}$, feed-down $\mathrm{D_s^+}$, and combinatorial-background candidates in pp collisions at $\sqrt{s} = 13$ TeV for $8 < p_{\mathrm{T}} < 12$ GeV/c. Middle: Product of the measured d_0^{xy} of the $\mathrm{D^0}$ (from $\mathrm{D^{*+}}$ decays) decay daughters in pp collisions at $\sqrt{s} = 5.02$ TeV for $2 < p_{\mathrm{T}} < 4$ GeV/c. Right: The absolute value of the max $\{(d_{0,i}^{\mathrm{reco}} - d_{0,i}^{\mathrm{exp}})/\sigma_\Delta\}$ distribution for $\mathrm{D_s^+}$ candidates in pp collisions at $\sqrt{s} = 13$ TeV.

4

Here θ_i^{xy} is the angle between the reconstructed momentum of the heavy-flavour hadron and that of the i-th daughter track, in the transverse plane. Its uncertainty is neglected, so

$$\sigma_{d_{0,i}^{\mathrm{xy}}(\mathrm{exp})} = \sigma(L_{\mathrm{xy}}) \cdot \sin\theta_i^{\mathrm{xy}}. \tag{4.6}$$

The right panel of Fig. 4.4 shows an example of the distribution for high-p_{T} $\mathrm{D_s^+}$ mesons in pp collisions.

- **Track dispersion at the decay vertex** (σ_{vtx})**:** This variable takes the square root of the sum of the squared distances of closest approach between the decay tracks and the reconstructed decay vertex

$$\sigma_{\mathrm{vtx}} = \sqrt{\sum_{i=1}^{N_{\mathrm{prongs}}} d_i^2}. \tag{4.7}$$

Again, resolution effects on the tracking and vertexing give rise to a non-zero value for true heavy-flavour hadrons, while it is pushed on average to higher values for combinatorial background, as shown in the left panel of Fig. 4.5 for $\mathrm{B_s^0}$ mesons. This variable is especially useful for three-prong decays, and hence only used for the $\mathrm{D_s^+}$ and $\mathrm{B_s^0}$ (via the $\mathrm{D_s^+}$) analyses.

- **Difference between invariant mass intermediate state and PDG world average** (ΔM)**:** A very effective way to reduce the combinatorial background is to select on the difference of the invariant mass of one of the intermediate hadron states with respect to the PDG world average [16]. Such selection is applied for all four decay channels, especially on the $\mathrm{D^0}$, $\phi(1020)$, $\mathrm{K_S^0}$, and $\mathrm{D_s^+}$ mesons. In the middle panel of Fig. 4.5, the $|\Delta M(\mathrm{KK})| = |M_{\mathrm{inv}}(\mathrm{KK}) - M_{\mathrm{PDG}}(\phi(1020))|$ is shown, where a clear peak in the signal distributions is visible, while the combinatorial background is uniformly distributed. A selection on this particular variable is what makes the contribution of reflections (see Sec. 4.2), coming from the two possible mass hypotheses for the $\mathrm{D_s^+}$ decay topology, negligible.

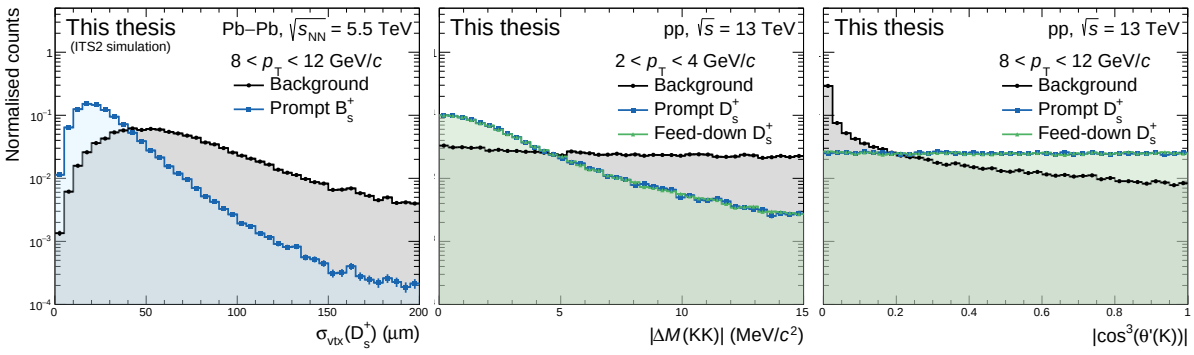


Figure 4.5: Left: The track dispersion at the decay vertex of the D_s^+ mesons (from B_s^0-meson decays) for $8 < p_T < 12$ GeV/c in simulated Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV with the upgraded ITS setup for LHC Run 3. Middle: Difference between the invariant mass of the KK-pair in D_s^+-meson decays and the PDG value of the $\phi(1020)$ mass in pp collisions at $\sqrt{s} = 13$ TeV for $2 < p_T < 4$ GeV/c. Right: The $|\cos^3\theta'(K)|$ distribution for D_s^+-candidates decays with $8 < p_T < 12$ GeV/c in pp collisions at $\sqrt{s} = 13$ TeV.

- **Cosine cubed of one of the D_s^+-decay angles** ($\cos^3\theta'(K)$)**:** This is a specific variable for the $D_s^+ \to \phi\pi^+ \to K^+K^-\pi^+$ decay, for which the spin of the ϕ meson is aligned to its direction of motion relative to the D_s^+ meson. As a consequence, it can be shown that the cube of the cosine of the angle between one of the kaons and the pion in the KK rest frame is expected to be flat for signal D_s^+ within $[-1, +1]$, while it peaks at zero for combinatorial background. This is illustrated in the right panel of Fig. 4.5 for D_s^+-meson candidates in pp collisions. This variable is used as well for the B_s^0 analysis.

Most of the selection variables are related to the displaced decay topology of the weakly-decaying heavy-flavour hadrons, which, as explained above, cannot be exploited for the $\Lambda_c^+ \to pK_S^0$ decay channel given the poor resolution on its decay vertex. The measured impact parameter of the proton-candidate track is therefore the only variable sensitive to the displaced decay topology. The distribution for prompt Λ_c^+, feed-down Λ_c^+, and combinatorial-background candidates is illustrated in the left panel of Fig. 4.6, showing a reduced background-rejection potential compared to the impact parameter distributions for the other here presented heavy-flavour hadrons. Other variables exploited in the Λ_c^+ analysis are related to achieve an as pure as possible sample of K_S^0 mesons, via, e.g., a selection on the decay length (middle panel of Fig. 4.6), on the cosine of the pointing angle (right panel) of the K_S^0 meson, or in the Armenteros-Podolanski space [193]. Because the large proper decay length of the K_S^0 meson, these variables are hardly sensitive to the additional displaced decay topology of the Λ_c^+ baryon. This makes that the particle identification selections, discussed in the next section, are fundamental for the $\Lambda_c^+ \to pK_S^0$ analysis.

4.2.3 Selections on particle-identification information

As introduced in Chapter 3, the specific energy loss dE/dx measured in the TPC and the time-of-flight measured in the TOF detector can be used to apply particle identification (PID) selections on the decay tracks. A track is considered to be compatible with the pion, kaon, or

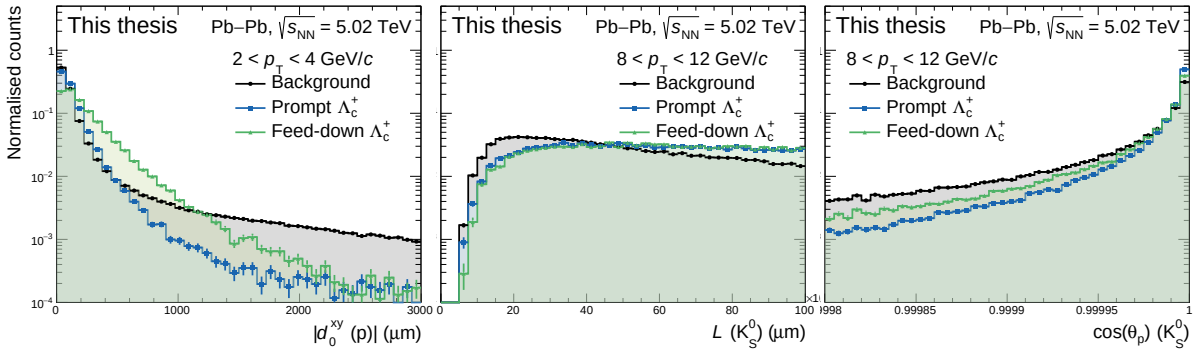


Figure 4.6: Measured impact-parameter distribution in the transverse plane of proton-candidate tracks (left), decay length distribution of K_S^0 candidates (middle), and cosine of the pointing angle distribution of K_S^0 candidates (right) coming from prompt Λ_c^+, feed-down Λ_c^+, and combinatorial-background candidates in central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.

proton mass hypothesis depending on the difference between the measured and expected signal S in the TPC or TOF. This is expressed in units of resolution σ as

$$N_\sigma(\pi, \mathrm{K}, \mathrm{p}) = \left| \frac{S_{\mathrm{meas}} - S_{\mathrm{exp}}(\pi, \mathrm{K}, \mathrm{p})}{\sigma(\pi, \mathrm{K}, \mathrm{p})} \right| . \tag{4.8}$$

Different PID strategies are adopted for the different analyses, briefly summarised below.

- **PID strategy D^{*+}:** The particle-identification selections are applied to the D^0 decay tracks, while PID is not considered for the soft-pion track. For both D^0 decay tracks, a $N_\sigma^{\mathrm{TPC}} < 3$ and $N_\sigma^{\mathrm{TOF}} < 3$ compatibility selection with the pion or kaon mass was asked. Tracks without a TOF signal were identified using only the TPC information. Based on this PID information and the charge sign of the decay tracks, D^0 candidates were accepted or rejected, according to their compatibility with the $\mathrm{K}^-\pi^+$ (in case of a positively-charged soft pion) or the $\mathrm{K}^+\pi^-$ (in case of a negatively-charged soft pion) final state.

- **PID strategy D_s^+:** The particle-identification for the D_s^+ analysis is two-fold. First of all, a general "prefiltering" selection is applied, after which the selected candidates, including some of the PID variables, are passed to a machine-learning algorithm.

  For the prefilter selection, a conservative $N_\sigma < 3$ compatibility selection was applied to the difference between the measured and expected signals for the particle TPC $\mathrm{d}E/\mathrm{d}x$ and TOF time-of-flight using an OR logic between the two detector responses. Tracks without hits in the TOF detector were identified using only the TPC information, and vice versa. Tracks without any PID information were treated as non-identified, but still used in the analysis and considered to be compatible with both a pion and a kaon hypotheses. In addition, since the decay particle in the $\mathrm{D}_\mathrm{s}^+ \to \phi\pi^+ \to \mathrm{K}^+\mathrm{K}^-\pi^+$ decay with opposite charge sign with respect to the D_s^+ meson has to be a kaon, a triplet was rejected if the opposite-sign track was not compatible with the kaon mass hypothesis.

  The PID variables given to the machine-learning algorithm are a combination of the TPC and TOF information. In this way, the number of variables used in the machine-learning

training is kept low and tracks without PID information in one of the detectors are properly accounted for. In the end, six PID variables are obtained, two for each track, which contain the information on the compatibility of the track with the pion and kaon hypotheses according to

$$N_{\sigma}^{\mathrm{comb}}(\pi,\mathrm{K}) = \begin{cases} N_{\sigma}^{\mathrm{TPC}}(\pi,\mathrm{K}) & \text{only TPC info;} \\ N_{\sigma}^{\mathrm{TOF}}(\pi,\mathrm{K}) & \text{only TOF info;} \\ \frac{1}{\sqrt{2}}\sqrt{(N_{\sigma}^{\mathrm{TPC}}(\pi,\mathrm{K}))^2 + (N_{\sigma}^{\mathrm{TOF}}(\pi,\mathrm{K}))^2} & \text{TPC and TOF info.} \end{cases} \tag{4.9}$$

- **PID strategy Λ_{c}^{+}:** For the proton track, a conservative $N_{\sigma} < 3$ compatibility selection was applied to both the TPC and TOF PID information as prefiltering selection. Tracks without hits in the TOF detector were identified using only the TPC information. No particle-identification selection was applied for the $\mathrm{K}_{\mathrm{S}}^{0}$ decay tracks. Afterwards, one PID variable is provided to the machine-learning algorithm, in particular $N_{\sigma}^{\mathrm{comb}}(\mathrm{p})$ similarly as defined in Eq. 4.9 for pions and kaons.

- **PID strategy $\mathrm{B}_{\mathrm{s}}^{0}$:** Only a prefiltering PID selection is applied for the $\mathrm{B}_{\mathrm{s}}^{0}$ analysis, following the conservative $N_{\sigma} < 3$ compatibility strategy as discussed above for the $\mathrm{D}_{\mathrm{s}}^{+}$ meson, where the additional bachelor track has to be compatible with the pion mass hypothesis. No additional PID variables are considered in the machine-learning algorithm.

4.2.4 Multivariate selection

Machine-learning methods, where an algorithm "learns" to perform a task for which it is not explicitly programmed for, are well suited for classification and regression problems. Applications in high-energy physics get therefore more and more common in recent years, see for example Refs. [194, 195] for an overview. In this dissertation, machine learning (ML) techniques based on the Boosted Decision Tree (BDT) algorithm are used to improve the candidate selection for the $\mathrm{D}_{\mathrm{s}}^{+}$, Λ_{c}^{+}, and $\mathrm{B}_{\mathrm{s}}^{0}$ analyses. Since this is one of the most basic applications of ML in high-energy physics, and fairly well established by now, we will be brief in this section.

The *internal parameters* of the BDT are set by a procedure called *training* on a data sample where the algorithm knows the difference between the *class* typologies. Ideally, the model learns some general patterns which are related to the different classes, so it can be employed to label unclassified *instances*. The trained model will then provide an *output score* based on the instance's *features*, which is a numerical value related to the probability of an instance belonging to the different classes. The final classification, i.e. candidate selection, is based on a threshold value of the output score. To achieve the best performance, the *external parameters* of the BDT should be optimised by the user for the specific problem. For more details on these concepts, see Ref. [196].

The selection of $\mathrm{D}_{\mathrm{s}}^{+}$, Λ_{c}^{+}, and $\mathrm{B}_{\mathrm{s}}^{0}$ candidates were performed using multivariate techniques based on the Boosted Decision Tree algorithm provided by the XGBoost package [197]. The BDT

training was performed for each p_{T} interval considering two type of classes: prompt signal candidates from dedicated simulations (as introduced in Sec. 4.1) and combinatorial background candidates from the sidebands of the invariant-mass distribution in data. The input data samples were split randomly into a training (80%) and a testing (20%) data set, with an upper limit between 50–200k candidates to avoid *overtraining*. The selection features considered in the optimisation included both topological and particle identification variables, chosen via an iterative process where the variables below an *importance* threshold were removed. The external parameters of the BDT, so-called hyperparameters, were optimised using a Bayesian optimisation [198], performed independently for each p_{T} interval. The testing data sample was, in the end, used as a statistically independent evaluation of the performance of the BDT algorithm, for example to test the level of overtraining.

Typical distributions for the BDT output for the training and testing data samples are shown in Fig. 4.7 for $\mathrm{D_s^+}$ mesons in pp collisions and Fig. 4.8 for Λ_{c}^+ baryons in Pb–Pb collisions. The left panels show an example of the BDT responses, for which the distribution to the respective signal class is close to unity, while the background distributions are typically more shifted towards zero. The right panels show the Receiver-Operator-Characteristic (ROC) curves, which plots the background rejection as function of the signal selection efficiency. The possible values of the "Area under the ROC Curve" (AUC) are comprehended between 0.5 and 1, where the first corresponds to a random classification and the latter to a perfect discrimination of the two class samples. The AUC gives a global estimation of the model performance, i.e. not related to the threshold value that will be chosen. The presented BDT for the $\mathrm{D_s^+}$ analysis clearly shows an improved performance with respect to the Λ_{c}^+ one, which can be attributed due the increased separation potential for $\mathrm{D_s^+}$ mesons because of the reconstructed decay vertex, as well as the less severe environment in pp collisions[3]. The compatibility for the BDT output response and the ROC curves between the training and testing data samples demonstrate that the BDTs were not overtrained.

The choice of the selections on the BDT output response was performed by estimating the expected significance ($S_{\mathrm{exp}}/\sqrt{S_{\mathrm{exp}} + B_{\mathrm{exp}}}$) for different selection criteria based on theoretical predictions. The expected signal, normalised per event, was estimated using FONLL predictions [83] as

$$S_{\mathrm{exp}} = A \cdot \left(\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}} \right)_{\mathrm{prompt}}^{\mathrm{FONLL}} \cdot \epsilon_{\mathrm{prompt}} / f_{\mathrm{prompt}}, \tag{4.10}$$

where A is a factor ensuring the proper normalisation for each p_{T} interval[4], $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}$ the FONLL production cross section for prompt hadrons, $\epsilon_{\mathrm{prompt}}$ the reconstruction and selection efficiency, and f_{prompt} the fraction of prompt hadrons. The f_{prompt} fraction can be estimated via the FONLL cross sections for prompt and feed-down charm hadrons and their respective Monte

[3] For practical reasons, the harsher environment in Pb–Pb collisions also requires the need to use tighter prefiltering selections, hence there is less "room" for the BDT to differentiate between signal and background candidates.

[4] Including terms like, e.g., the branching ratio, acceptance factor, and widths of the p_{T} and y intervals. This factor is independent on the used selection on the BDT output response.

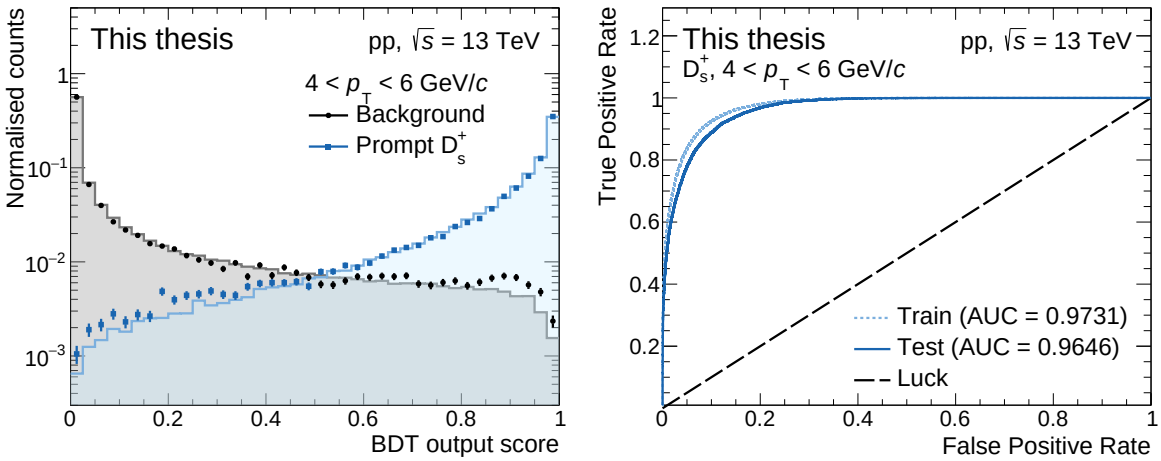


Figure 4.7: Left: BDT output response for prompt D_s^+ and combinatorial background candidates in pp collisions at $\sqrt{s} = 13$ TeV for $4 < p_T < 6$ GeV/c. The histograms present the responses in the training data sample, while the markers are from the independent testing data sample. Right: The corresponding ROC curves and AUC values for the training and testing data sample. See text for more details.

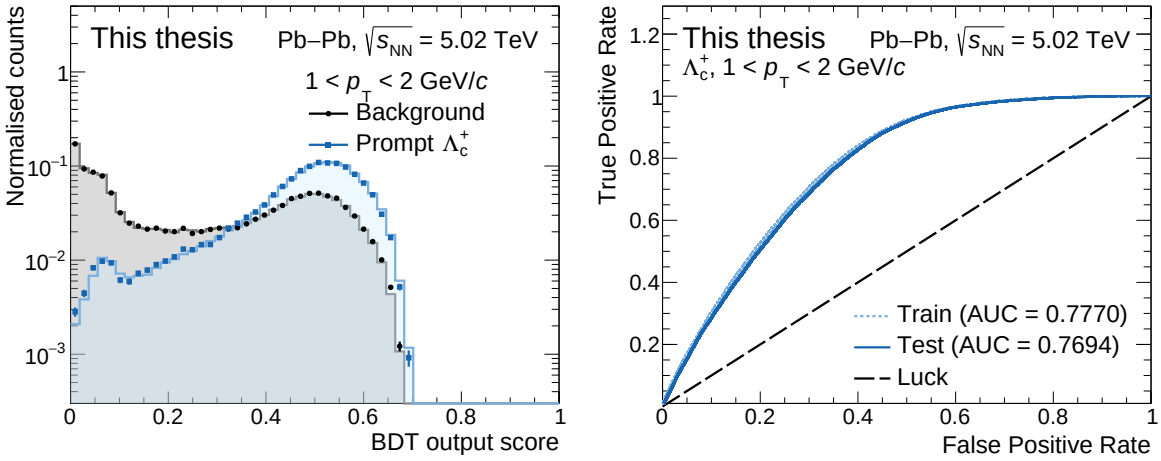


Figure 4.8: Left: BDT output response for prompt Λ_c^+ and combinatorial background candidates in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for $1 < p_T < 2$ GeV/c. The histograms present the responses in the training data sample, while the markers are from the independent testing data sample. Right: The corresponding ROC curves and AUC values for the training and testing data sample. See text for more details.

Carlo selection efficiencies. The expected background under the signal peak, B_{exp}, was instead evaluated from a small fraction of the data by fitting the sidebands of the invariant-mass distribution. The optimal selection on the BDT output response was selected as the value that guarantees a high significance and a sufficiently high selection efficiency. It is noted here that the significance, as defined above, is roughly equal to the inverse of the statistical uncertainty of the signal, and therefore a higher significance ensures a more precise measurement.

Selections based on properly optimised and trained BDTs are likely to have an improved performance, i.e. a larger statistical significance and/or signal-to-background ratio with a higher selection efficiency, with respect to rectangular selections. This is partly due to the non-linearity in the selections on the underlying variables, but it is also simply easier to optimise the selections

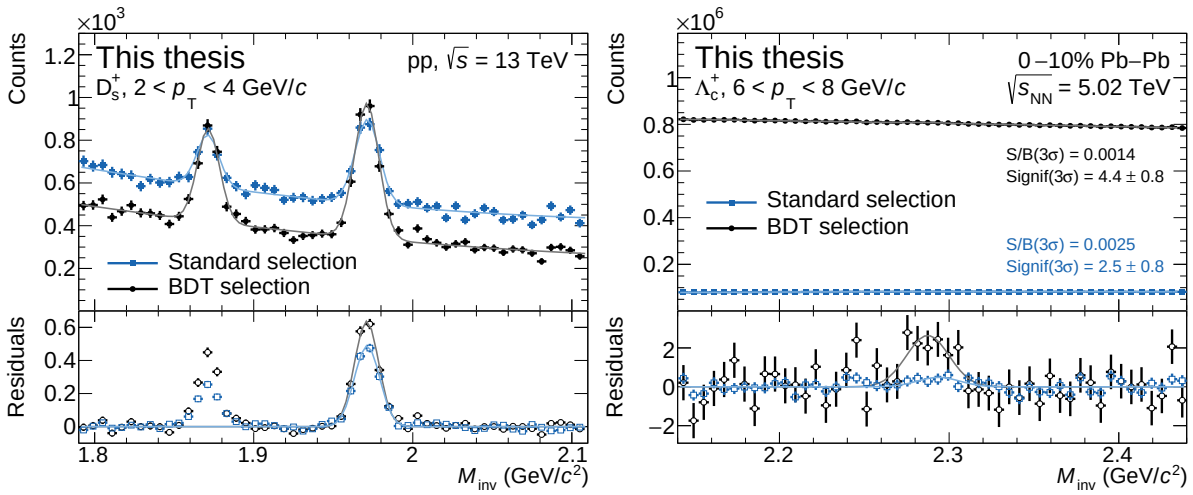


Figure 4.9: A comparison of the invariant-mass distributions for $\mathrm{D_s^+} \to \phi\pi^+ \to \mathrm{K^+K^-}\pi^+$ in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_\mathrm{T} < 4$ GeV/c (left) and $\Lambda_\mathrm{c}^+ \to \mathrm{pK_S^0}$ in 0–10% Pb–Pb collisions at $\sqrt{s_\mathrm{NN}} = 5.02$ TeV for $6 < p_\mathrm{T} < 8$ GeV/c (right) with optimised rectangular selections (from Refs. [82, 199]) and the optimised BDT selection used in this dissertation. The lines represent a fit — composed out of a second-order polynomial function and one or two Gaussian functions — to the data. The bottom panels show the residual distribution after subtracting the background fit function.

given one has to find the optimal selection on one "variable". Exploiting machine-learning algorithms for the selection comes with the downside that possible correlations between the training variables and the candidates invariant mass could lead to modifications of the invariant-mass distribution, possibly leading to fit instabilities from complex background shapes or even introducing "fake signal". For the BDTs used in this dissertation, this was studied on simulations of purely combinatorial background, leading to the removal of the $\cos\theta^*$ variable for the $\mathrm{D_s^+}$ and Λ_c^+ analyses.

Comparisons of optimised rectangular selections (from previous analyses [82, 199]) and optimised selections on the BDT output response variable (as introduced above) via their final invariant-mass distributions are shown in Fig. 4.9 for the $\mathrm{D_s^+} \to \phi\pi^+ \to \mathrm{K^+K^-}\pi^+$ analysis in pp collisions at $\sqrt{s} = 13$ TeV and $\Lambda_\mathrm{c}^+ \to \mathrm{pK_S^0}$ in Pb–Pb collisions at $\sqrt{s_\mathrm{NN}} = 5.02$ TeV. In both cases, an improvement in the significance of about a factor 1.7 is found for the BDT selections with respect to the rectangular selections. The signal-to-background (S/B) ratio also improves by a factor 2.3 for the $\mathrm{D_s^+}$ analysis, whereas the optimal BDT selection for the Λ_c^+ analysis leads to a lower S/B and a much higher background level. The Λ_c^+ selection efficiency is instead a factor four higher for the BDT analysis. This difference can be attributed to the removal of the $\mathrm{sgn}\{\vec{p} \cdot \vec{d}_0^{\,\mathrm{xy}}(\mathrm{p})\} \cdot d_0^{\mathrm{xy}}(\mathrm{p})$ variable from the selections, which was observed to introduce instabilities in the analysis.

4.3 Signal extraction and corrections

The final observables for the analyses presented in this dissertation are — or are directly related to — the p_T-differential yields of (prompt) heavy-flavour hadrons within $|y| < 0.5$, which are

evaluated as

$$\left.\frac{\mathrm{d}N^{h_{\mathrm{c,b}}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{|y|<0.5} = \frac{f_{\mathrm{prompt}} \cdot \frac{1}{2}\, N_{\mathrm{raw}}^{h_{\mathrm{c,b}}}\Big|_{|y|<y_{\mathrm{fid}}}}{\Delta p_{\mathrm{T}} \cdot \alpha_y \cdot (A \times \epsilon)_{\mathrm{prompt}} \cdot \mathrm{BR} \cdot N_{\mathrm{ev}}^{\mathrm{norm}}}. \tag{4.11}$$

The raw yield values, $N_{\mathrm{raw}}^{h_{\mathrm{c,b}}}$, extracted in a given p_{T} interval (of width Δp_{T}), are divided by a factor two and multiplied by the prompt fraction f_{prompt} to obtain the charge-averaged yields of prompt heavy-flavour hadrons. Furthermore, they were divided by $\alpha_y \cdot (A \times \epsilon)$, enclosing the rapidity coverage and the acceptance-times-efficiency, the BR of the decay channel, and by the number of analysed events $N_{\mathrm{evt}}^{\mathrm{norm}}$. In case of additional selections on a multiplicity estimator, the number of analysed events needs to be further corrected for the trigger efficiency $N_{\mathrm{ev}}^{\mathrm{norm}}/\epsilon_{\mathrm{trig}}$. The p_{T}-differential yield is directly related to the cross section $\mathrm{d}\sigma^{h_{\mathrm{c,b}}}/\mathrm{d}p_{\mathrm{T}}$ by replacing $N_{\mathrm{ev}}^{\mathrm{norm}}$ with the integrated luminosity $\mathcal{L}_{\mathrm{int}} = N_{\mathrm{ev}}^{\mathrm{norm}}/\sigma_{\mathrm{MB}}$, where σ_{MB} is the cross section for the MB trigger condition that is estimated via van der Meer scans [200].

The raw yields of the heavy-flavour hadrons were extracted in each p_{T} interval via binned maximum-likelihood fits to the candidate invariant-mass distributions. The fitting function consisted of a Gaussian term to estimate the signal and an exponential, polynomial, or threshold function (depending on hadron specie and p_{T}) to estimate the background. The width of the Gaussian was in certain cases fixed to the values obtained from simulations to improve the fit stability. The contribution of signal candidates with the wrong decay-particle mass assignment were assumed to be negligible after the selections. For the $M(\mathrm{KK}\pi)$ distribution, i.e. the $\mathrm{D_s^+}$ invariant-mass distribution, an additional Gaussian was used to describe the signal of the decay $\mathrm{D^+ \to K^+K^-\pi^+}$, present on the left side of the $\mathrm{D_s^+}$-meson signal as visible in Fig. 4.9. More examples of the invariant-mass distributions, together with the result of the fits, will be shown in each dedicated chapter.

The $\alpha_y \cdot (A \times \epsilon)$ corrections were obtained from simulations as introduced above, including a detailed description of the detector geometry and its response as well as the LHC beam conditions. The efficiency is defined as

$$\epsilon = \frac{N_{\mathrm{reco+sel}}\big|_{|y|<y_{\mathrm{fid}}}}{N_{\mathrm{gen}}^{\mathrm{dau\,in\,acc}}\big|_{|y|<y_{\mathrm{fid}}}}, \tag{4.12}$$

where $N_{\mathrm{reco+sel}}$ is the number of reconstructed (signal) $h_{\mathrm{c,b}}$ after all the selections and $N_{\mathrm{gen}}^{\mathrm{dau\,in\,acc}}$ the number of generated heavy-flavour hadrons (in the decay channel of interest) with their daughters fulfilling the $|\eta| < 0.9$ and $p_{\mathrm{T}} > 0.1$ GeV/c acceptance requirements. In both cases, only candidates within the fiducial acceptance volume, as defined in Eq. 4.4, were included. Furthermore, only events with a generated z-vertex within ± 10 cm were considered, while for $N_{\mathrm{reco+sel}}$ also the event selection criteria (as defined in Sec. 4.1.1) are required. A separation is made between prompt and feed-down charm hadrons, since, in general, different efficiencies are expected due to the geometrical selections applied on the decay-vertex topology, which for feed-down charm hadrons are more displaced due to the larger h_{b} lifetime ($c\tau \approx 500$ µm). In addition, the efficiency is expected to vary with p_{T} and multiplicity because of the track-reconstruction efficiency and primary-vertex resolution.

The acceptance term times the enclosed rapidity coverage, $\alpha_y \cdot A$, is defined as

$$A' = \alpha_y \cdot A = \frac{N_{\rm gen}^{\rm dau\ in\ acc}\big|_{|y|<y_{\rm fid}}}{N_{\rm gen}\big|_{|y|<0.5}}, \tag{4.13}$$

which can be calculated directly from the used simulations, or estimated using a PYTHIA toy Monte Carlo approach. The latter has the advantage that, besides a "direct" control of the statistical uncertainty, different rapidity shapes of the heavy-flavour hadrons can be used. It was verified, however, that the rapidity shape of $h_{\rm c,b}$ in PYTHIA is nearly flat for $|y| < 0.8$, confirmed by FONLL calculations [83] and experimental measurements [170]. Therefore, $\alpha_y \simeq y_{\rm fid}/0.5$ and thus varies from 1 to 1.6 as function of $p_{\rm T}$, meaning A' can be larger than unity. In case the acceptance and efficiency correction terms are calculated from the same simulations, as done for the analyses in this dissertation, the partial correlation between $N_{\rm reco+sel}$ and $N_{\rm gen}$ is approximated as $\rho = \sqrt{\epsilon \cdot (0.5/y_{\rm fid}(p_{\rm T})) \cdot (A'/1.6)}$. To use more realistic momentum and/or multiplicity distributions for the $(A' \cdot \epsilon)$ determination, reweighting techniques ($N = \sum_i w_i N_i$) were used on both the generated and reconstructed number of candidates.

Finally, for the prompt fraction correction, typically two methods are used within ALICE: a theory-driven and data-driven estimation. In the theory-driven method, the $f_{\rm prompt}$ fractions were evaluated in each $p_{\rm T}$ interval by subtracting the calculated raw yield of feed-down charm hadrons ($N_{\rm raw}^{\rm feed-down}$) from the measured inclusive raw yield ($N_{\rm raw}$)

$$\begin{aligned} f_{\rm prompt} &= 1 - \frac{N_{\rm raw}^{\rm feed-down}}{N_{\rm raw}} \\ &= 1 - \left(\frac{{\rm d}^2\sigma}{{\rm d}p_{\rm T}{\rm d}y}\right)^{\rm feed-down}_{\rm FONLL,\,PYTHIA} \cdot \frac{(A' \times \epsilon)_{\rm feed-down} \cdot \Delta p_{\rm T} \cdot {\rm BR} \cdot \mathcal{L}_{\rm int}}{\frac{1}{2} \cdot N_{\rm raw}}. \end{aligned} \tag{4.14}$$

The beauty-quark production cross section was estimated with FONLL calculations [83], the fraction of beauty quarks that fragment into B mesons was estimated from $\rm e^+e^-$ measurements [201] and those into $\Lambda_{\rm b}^0$ from the $\Lambda_{\rm b}^0/({\rm B}^0 + {\rm B}^+)$ ratio measured by LHCb in pp collisions at $\sqrt{s} = 13$ TeV [108], and the beauty-hadron decay kinematics modelled from PYTHIA 8 simulations (or the EvtGen package [202] for the older $\rm D^{*+}$ analysis). In case of analyses in Pb–Pb collisions, the nuclear modification factor of charm hadrons from beauty-hadron decays was also taken into account assuming $R_{\rm AA}^{\rm feed-down} = a \cdot R_{\rm AA}^{\rm prompt}$, where a is a constant, as well as the nuclear overlap function $T_{\rm AA}$. A systematic uncertainty on the estimation of $f_{\rm prompt}$ was evaluated by varying the FONLL parameters, as well as a variation on the $R_{\rm AA}^{\rm feed-down}$ hypothesis. The $f_{\rm prompt}$ values depend on the topological selections and the hadron specie, but are typically around 0.9. More details for the specific analyses are given in the relevant chapters.

The precision of the data-driven method, which extracts the $f_{\rm prompt}$ fraction from an unbinned log-likelihood fit to the $d_0^{\rm xy}$ distribution of heavy-flavour candidates in a 2σ interval in the invariant-mass signal region, is too poor given the fine $p_{\rm T}$ intervals and the limited data sample sizes. Therefore, all charm-hadron analyses presented in this dissertation are exploiting the theory-driven estimation. A comparison of both methods is given in Fig. 4.10 for $\rm D^0$ mesons in

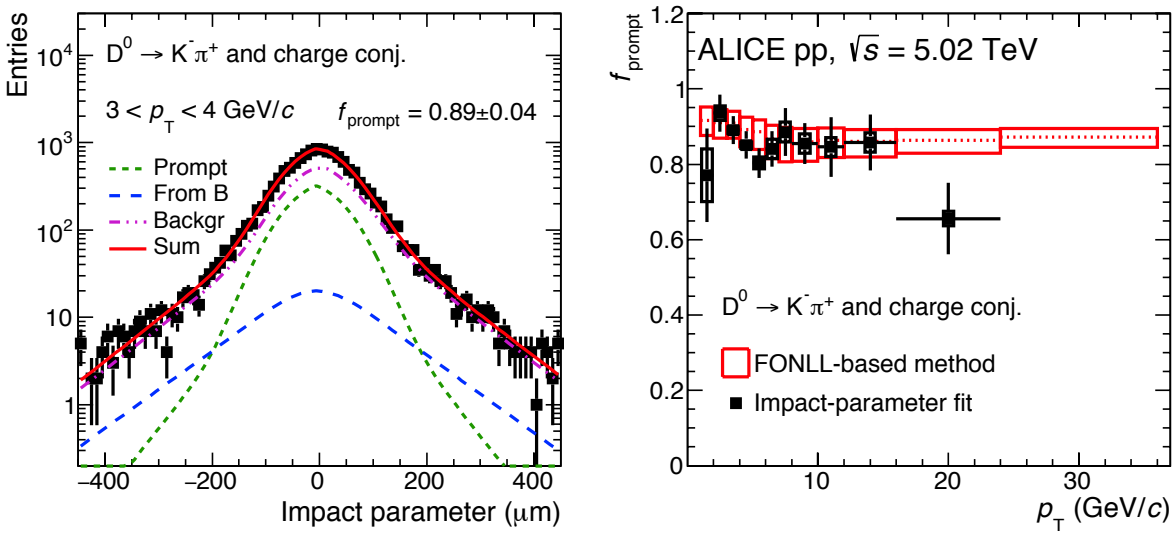


Figure 4.10: Left: Example of a fit to the impact-parameter distribution of D^0 candidates to estimate the f_{prompt} fraction in pp collisions at $\sqrt{s} = 5.02$ TeV [82]. The curves show the fit functions describing the prompt, feed-down, and background candidates, as well as their sum. Right: Comparison of the extracted f_{prompt} fractions from the data-driven and FONLL-based method. The estimation of the systematic uncertainty on the data-driven method (shown as boxes) are explained in detail in Ref. [170].

pp collisions at $\sqrt{s} = 5.02$ TeV. For this comparison, wider p_T intervals were adopted compared to the actual analysis [82], to improve the quality of the impact-parameter fit. Both methods are found to be compatible within uncertainties.

CHAPTER 5

D-meson production in pp collisions at $\sqrt{s} = 5.02$ TeV

Abstract – The production measurement of prompt D^{*+} mesons in proton–proton collisions at $\sqrt{s} = 5.02$ TeV with the ALICE detector is reported. This analysis is performed in close collaboration with the analysers of the D^0, D^+, and D_s^+ mesons in the same data sample. The D^{*+} mesons were reconstructed at midrapidity ($|y| < 0.5$) via their hadronic decay channel $D^{*+} \to D^0\pi^+ \to K^-\pi^+\pi^+$ (and its charge conjugate) and measured in the transverse momentum interval $1 < p_T < 36$ GeV/c. Thanks to the higher integrated luminosity, an analysis in finer p_T intervals with respect to the previous measurements at $\sqrt{s} = 7$ TeV was performed, allowing for a more detailed description of the p_T shape. The measured p_T-differential production cross sections are compared to the results at $\sqrt{s} = 7$ TeV and to six different theoretical calculations. The D^{*+}/D^0 production ratio and the visible and p_T-extrapolated D^{*+} production yields are also reported. This measurement will, in addition, allow for a more accurate determination of the nuclear modification factor in p–Pb and Pb–Pb collisions performed at the same nucleon–nucleon centre-of-mass energy.

In collaboration with C. Terrevoli (D^0), F. Prino (D^0), F. Grosa (D_s^+), and R. Bala (D^+)

Published in Eur. Phys. J. C ***79*** *(2019) 388*

5.1 Introduction

Measurements of heavy-flavour production in proton–proton collisions at LHC energies are a sensitive test of quantum chromodynamics calculations with the factorisation approach, as discussed extensively in Chapter 2. In such schemes, the p_{T}-differential production cross sections of heavy-flavour hadrons are calculated as a convolution of (i) the parton distribution functions of the incoming protons, (ii) the partonic scattering cross section, calculated as a perturbative series in powers of the strong coupling constant α_{s}, and (iii) the fragmentation function, which parametrises the evolution of a heavy quark into a given heavy-flavour hadron species.

In this chapter, we consider the collinear factorisation theorem, implemented in terms of the squared momentum transfer Q^2 [77], and factorisation exploiting the partonic transverse momentum k_{T} [92, 93]. Typical collinear factorisation calculations for LHC energies are available in the general-mass variable-flavour-number scheme, GM-VFNS [84, 89, 203, 204], and in the fixed-order plus next-to-leading-logarithms approach, FONLL [83, 87]. Both schemes have next-to-leading order accuracy, with all-order resummation of next-to-leading logarithms. Within the k_{T}-factorisation framework, heavy-flavour production cross-section calculations exist only at leading order (LO) approximation in α_{s} [85, 92]. Previous publications [169, 205] have shown that D-meson production in pp collisions is well described by these calculations at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 2.76$ TeV. This description at different collision energies is important for heavy-flavour production measurements in p–Pb and Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV as well, because those so far relied on a p_{T}-dependent $\sqrt{s}$-scaling factor extracted from FONLL calculations [141, 144]. For a model-independent study of the effects induced by cold and hot strongly-interacting matter, it is therefore fundamental to provide the measurement in pp collisions at the same centre-of-mass energy per nucleon–nucleon collision as in the larger systems.

The ratios of the p_{T}-differential cross sections of open charm hadron species can help to constrain the phenomenological models of QCD in the non-perturbative regime that aim to describe heavy-quark hadronisation. The fragmentation functions of charm quarks to vector (D^{*+}) and pseudoscalar (D^0) can be constrained via the measured $\mathrm{D}^{*+}/\mathrm{D}^0$ ratio, while the $\mathrm{D}_{\mathrm{s}}^+/\mathrm{D}^0$ ratio is sensitive to the difference in hadronisation of strange and non-strange mesons. The relative abundances of open charm hadron species also test the statistical hadronisation scenario of charm quarks into hadrons [206], where a thermal description of charm-hadron production is assumed.

In this chapter, the p_{T}-differential production cross section of the D mesons, measured in $|y| < 0.5$ in pp collisions at $\sqrt{s} = 5.02$ TeV, are reported together with their ratios. It is organised as follows. In Section 5.2 the D^{*+}-meson selection procedure is explained. Sections 5.3 and 5.4 describe the D^{*+} raw-yield extraction and acceptance-times-efficiency and prompt fraction corrections, respectively. In Section 5.5, the evaluation of the systematic uncertainties for the D^{*+} analysis is discussed. Finally, in Section 5.6 and 5.7, the results will be presented and discussed, including as well some of the interesting features of the D^0, D^+, and $\mathrm{D}_{\mathrm{s}}^+$ production measurements, which were performed in close collaboration (and published together) with the

here presented D^{*+} measurement.

5.2 Selection criteria

The D^{*+} mesons were reconstructed via the $\mathrm{D}^{*+} \to \mathrm{D}^0\pi^+ \to \mathrm{K}^-\pi^+\pi^+$ decay channel, with branching ratio $2.67 \pm 0.02\%$ [16]. The pion from the D^{*+} decay is referred to as "soft pion" because of the limited phase space available in the decay. Since the D^{*+} decays strongly, its decay length is of the order of a few picometers, hence it is not possible to distinguish the decay vertex from the primary vertex. In order to increase the statistical significance, the topological selections are therefore mainly based on the D^0-decay topology (separated from the primary vertex by a few hundred micrometers), as explained in Section 4.2.2. The selection criteria were tuned to have a large statistical significance ($S/\sqrt{S+B}$), while preserving a relatively high selection efficiency and signal-to-background ratio. The topological selections are based on the $\sqrt{s} = 7$ TeV analysis [169], optimising in particular the distance of closest approach between the D^0 decay tracks and the reconstructed secondary vertex, the D^0 $\cos\theta_\mathrm{p}$, and the $d_0^{\mathrm{xy}}(K) \cdot d_0^{\mathrm{xy}}(\pi)$ via a multidimensional approach[1]. A summary of some of the threshold values are reported in Tab. 5.1. The values depend on p_T, since at low momentum the contribution from the combinatorial background is larger, while at high p_T the statistical significance of the D^{*+} signal is already relatively high before applying topological selections. In addition to the topological and kinematical selections, track-quality and particle-identification selections are applied as described in Sec. 4.2.

Table 5.1: Threshold values for the topological and kinematical selections applied to the D^{*+} (or D^0) candidates in pp collisions at $\sqrt{s} = 5.02$ TeV for four p_T intervals. At high p_T, fewer selections are needed to extract a D^{*+} signal with good statistical significance.

p_T (GeV/c)			1–2	4–5	9–10	24–36
$\lvert M_\mathrm{inv}(\mathrm{K}\pi) - M_\mathrm{PDG}(\mathrm{D}^0)\rvert$	(MeV/c^2)	$<$	32	32	50	74
DCA	(µm)	$<$	380	420	1000	-
$d_0^{\mathrm{xy}}(K) \cdot d_0^{\mathrm{xy}}(\pi)$	(µm^2)	$<$	$-1.35 \cdot 10^4$	$-0.28 \cdot 10^4$	-	-
$\cos\theta_\mathrm{p}$		$>$	0.8	0.8	0.7	0.2
$\cos\theta_\mathrm{p}^{\mathrm{xy}}$		$>$	0.88	-	-	-
$\cos\theta^*(\pi)$		$<$	0.9	0.9	-	-
$p_\mathrm{T}(\mathrm{K})$	(GeV/c)	$>$	0.6	1.0	0.6	0.3
$p_\mathrm{T}(\pi)$	(GeV/c)	$>$	0.6	1.0	0.6	0.3
$L_{\mathrm{xy}}/\sigma(L_{\mathrm{xy}})$		$>$	2.7	-	-	-

5.3 Raw-yield extraction

The D^{*+} raw yields were obtained from binned maximum likelihood fits to the invariant-mass difference distribution between D^{*+} and D^0 candidates ($\Delta M = M(\mathrm{K}\pi\pi) - M(\mathrm{K}\pi)$) in nineteen

[1]This procedure could be improved by calculating the expected significance, based on theoretical calculations of D^{*+} production and a fraction of the statistics in data, instead of optimising on the full data sample. In this way, one avoids biases induced by statistical fluctuations under the D^{*+} peak. This is the procedure employed by the machine-learning analyses presented in this dissertation.

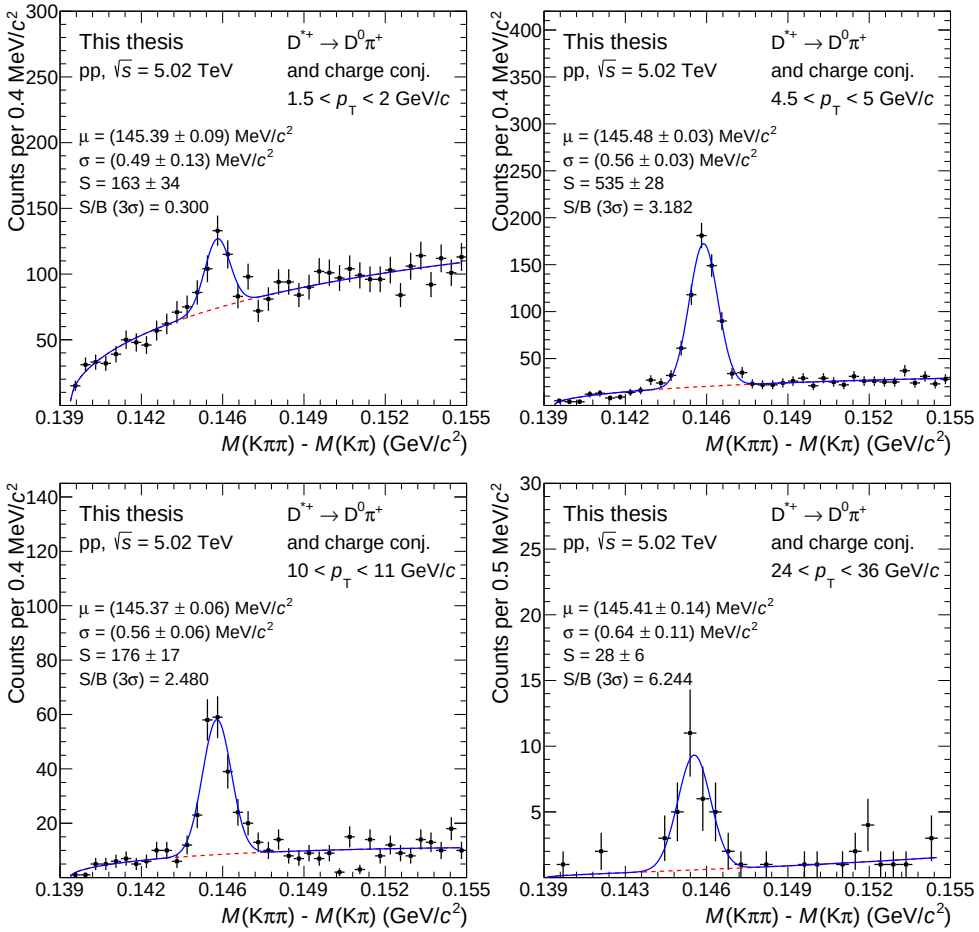


Figure 5.1: Invariant-mass difference distribution of D^{*+} candidates and charge conjugates in $1.5 < p_{\mathrm{T}} < 2$ GeV/c, $4.5 < p_{\mathrm{T}} < 5$ GeV/c, $10 < p_{\mathrm{T}} < 11$ GeV/c, and $24 < p_{\mathrm{T}} < 36$ GeV/c, respectively. The blue solid lines show the total fit function, while the red dashed lines are the combinatorial background terms. The values of the mean (μ) and the width (σ) of the signal peak are reported, together with the signal counts (S) and the signal-to-background ratio (S/B) in the mass interval $(\mu - 3\sigma, \mu + 3\sigma)$.

p_{T} intervals between $1 < p_{\mathrm{T}} < 36$ GeV/c. Because of the small mass difference between both mesons, the D^{*+} signal will be visible as a narrow peak at $\Delta M \approx 145.43$ MeV/c^2 [16] close to the threshold (m_π) and thus in a rather low combinatorial background region. The resolution in ΔM is dominated by the momentum resultion of the soft pion. In the fit function, the signal was modelled with a Gaussian, while the background was described by the kinematic threshold function $a\sqrt{\Delta M - m_\pi} \cdot \mathrm{e}^{b(\Delta M - m_\pi)}$. The non-Gaussian component in the D^{*+} signal distribution, mainly driven by the soft-pion track reconstruction[2], was studied in simulations by using the sum of two Gaussian functions and found to have an effect on the raw-yield extraction of 1–2%. Examples of the mass-difference fits at low, intermediate, and high p_{T} are reported in Fig. 5.1. The blue solid lines show the total fit functions and the red dashed lines are the combinatorial-background terms. The values of the raw yield, signal-to-background ratio, and significance (within 3σ) vary between 23 ± 6 to 735 ± 39, 0.30 to 6.24, and 6.1 ± 1.1 to 21.8 ± 0.7,

[2]The non-Gaussian tails in the distribution can be attributed to the uncertainties on the low-momentum soft-pion tracks, as well as to events where the π^+ decays in flight to a μ^+ and clusters from both the π^+ and μ^+ track segments were used in the track reconstruction [207].

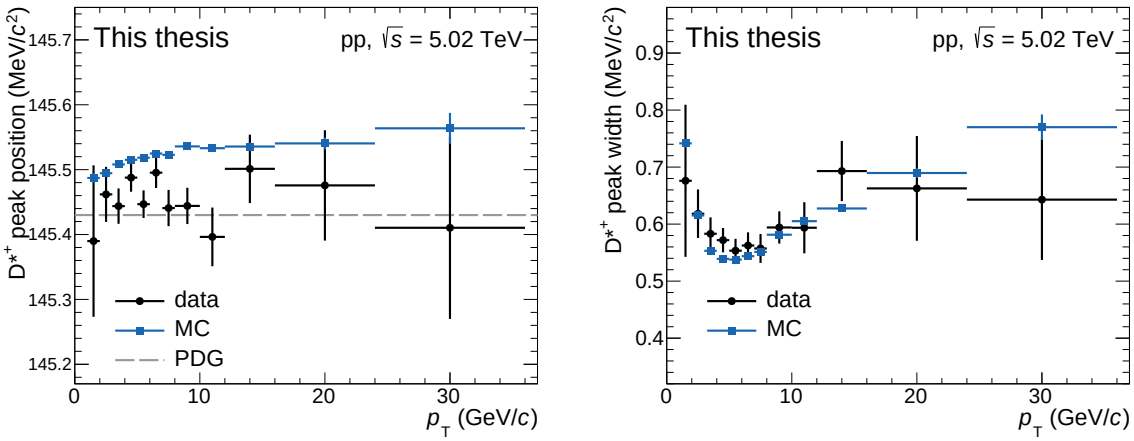


Figure 5.2: The Gaussian mean (left) and width (right) of the fitted D^{*+}-meson peak in data and MC simulations. The world-average value from the PDG [16] is reported as well. To reduce statistical fluctuations, broader p_T intervals are adopted with respect to the actual analysis.

respectively. The best significance is achieved at intermediate p_T, while the signal-to-background ratio is largest at high p_T.

In Fig. 5.2, the Gaussian mean and width of the fitted D^{*+} peaks are compared to the values obtained from Monte Carlo simulations. For this comparison, the fits were performed in wider p_T intervals to reduce the statistical fluctuations, but the same conclusions hold for the analysis in finer p_T intervals. While the extracted widths in data are compatible with simulations, the peak position is typically lower than simulations by about two standard deviations. Similar shifts in the peak position are observed for the other D mesons, and are believed to come from residual misalignment of the detectors. These discrepancies are taken into account in the evaluation of the source of systematic uncertainty related to the raw-yield extraction.

There is no contamination from other (resonant) decay channels, contributing to the D^{*+} signal, that should be corrected for. In addition, the ambiguity in assigning the decay-particle mass for the D^0 daughters is resolved by the charge of the soft-pion track. Also the contribution from neutral-meson decays to two particles (especially those from D^0 mesons not originating from $D^{*+} \to D^0\pi^+$) combined with a pion-candidate track is subdominant and not contaminating the peak region in the ΔM distributions in a significant way [208]. The extracted raw yields from the invariant-mass fits can therefore be directly used in Eq. 4.11.

5.4 Corrections

The acceptance and efficiency corrections were determined using dedicated ALICE Monte Carlo simulations (see Sec. 4.1) as described in detail in Sec. 4.3. Figure 5.3 shows the $\alpha_y \cdot A$ (left panel) and $A \cdot \epsilon$ for prompt and feed-down D^{*+} mesons within the fiducial acceptance region (right panel) as function of p_T. The acceptance term is presented as estimated directly from the used ALICE simulations, as well as when using a PYTHIA toy Monte Carlo approach with two different generated y-shapes. The different methods are compatible within 0.5%. The acceptance-times-efficiency increases with increasing p_T from about 0.2% at low p_T to 50% at

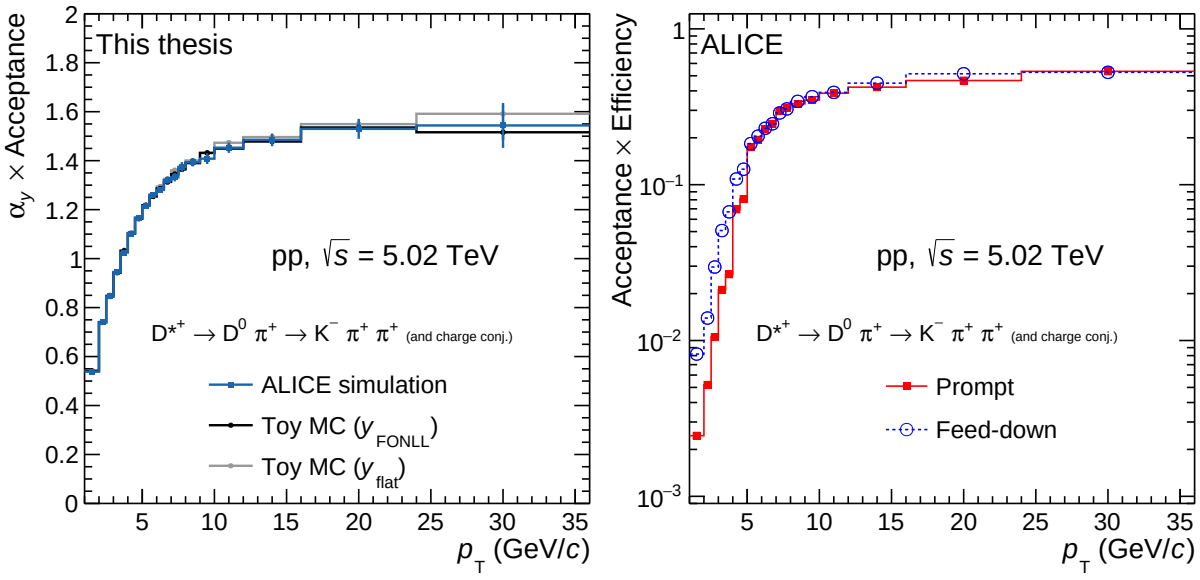


Figure 5.3: Left: The acceptance term for D^{*+} mesons in pp collisions at $\sqrt{s}$ = 5.02 TeV estimated with three different procedures. Right: Acceptance-times-efficiency for D^{*+} mesons as function of p_T. The efficiencies are shown for prompt (solid lines) and feed-down (dotted lines) D^{*+} mesons.

high p_T due to the looser topological selections (and slightly for the p_T dependence in the track- and vertex-reconstruction). The displacement from the primary vertex is, on average, larger for beauty hadrons due to their long lifetimes ($c\tau \approx 500$ µm [16]) resulting in a more efficient selection of feed-down compared to prompt D^{*+} mesons. The simulations used the GEANT3 transport package [191], but it was verified for this analysis that when using GEANT4 [192] the result was compatible within 2%.

The fraction of prompt D^{*+} was evaluated with the theory-driven method as discussed in Sec. 4.3, which uses the cross section of feed-down D^{*+} mesons from FONLL calculations [83] folded with the $h_b \to D + X$ decay kinematics from the EvtGen package [202]. A systematic uncertainty was estimated by varying the FONLL parameters, as prescribed in Ref. [83]. In particular, the renormalisation and factorisation scales were varied independently in the range $0.5 < \mu_R(\mu_F)/\mu_0 < 2$ (with μ_0 the default value in central FONLL predictions), as well as the beauty-quark mass between 4.5 and 5 GeV/c^2. The uncertainties on the CTEQ6.6 PDFs [209] used in the FONLL calculation were taken into account as well. The final prompt fractions in each p_T interval, including their systematic uncertainties, are shown in the right panel of Fig. 5.4. The f_{prompt} fractions depend on the applied topological selections and vary between 0.92 for $p_T < 5$ GeV/c to about 0.95 at intermediate p_T. The data-driven method (see Sec. 4.3) was tried as well, but was found to be unsuccessful for many p_T intervals for this D^{*+} analysis because of the limited number of selected candidates. The left panel of Fig. 5.4 shows an example of the fit to the impact-parameter distribution of D^{*+} mesons in the transverse-momentum interval $3 < p_T < 4$ GeV/c, while in the right panel the extracted prompt fractions with the data-driven method are compared to those calculated with the FONLL-based approach. For the p_T intervals in which the impact-parameter fit converged, the extracted prompt fractions are compatible within uncertainties, as is also the case for the D^0 and D_s^+ analyses [82].

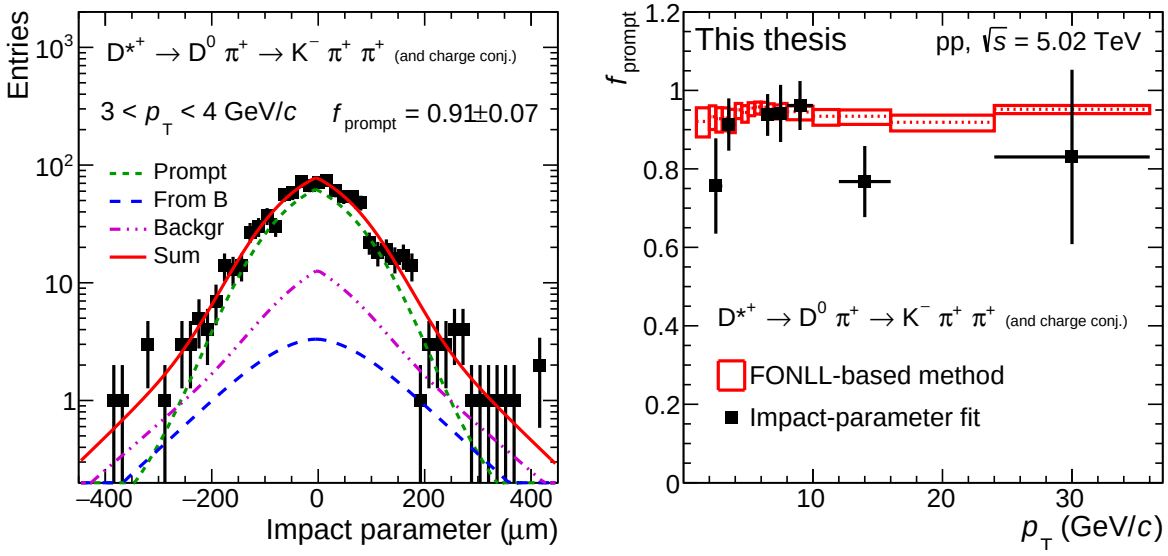


Figure 5.4: Left: Fit to the impact-parameter distribution of D^{*+} candidates in $3 < p_T < 4$ GeV/c. The curves show the fit functions describing the prompt, feed-down, and background contributions, as well as their sum. Right: The estimated prompt fractions of D^{*+} mesons from the FONLL-based approach (dashed line) and their uncertainties (red boxes) compared to the prompt fractions as estimated with the data-driven method in coarser p_T intervals (squared markers).

5.5 Systematic uncertainties

Systematic uncertainties on the D^{*+}-meson cross section were estimated considering the following sources: the extraction of the raw yield from the invariant-mass distributions (δ_1); the track-reconstruction efficiency (δ_2); the D^{*+}-meson selection efficiency (δ_3); the PID selection efficiency (δ_4); the p_T shape of generated D^{*+} mesons in the simulation (δ_5); and the subtraction of the feed-down from beauty-hadron decays (δ_6). In addition, the uncertainties on the branching ratios (δ_7) and on the integrated luminosity (δ_8) were considered. In the subsections below, source δ_1–δ_5 will be discussed in more detail, while the systematic uncertainty assigned for the feed-down subtraction was already discussed in Sec. 5.4.

5.5.1 Raw-yield extraction

The systematic uncertainties on the raw-yield extraction were evaluated by repeating the fits to the invariant-mass spectra several hundred times varying the fit interval and the functional form of the background fit function, the so-called multi-trial approach. In particular, two lower (140.0 and 140.5 MeV/c^2) and six upper (between 155 and 165 MeV/c^2) limits for the invariant-mass fit interval were used. Next to the threshold function, a power-law function ($a(\Delta M - m_\pi)^b$) was also considered as background fit function. A shift of the invariant-mass distributions of 0.1, 0.2, or 0.3 MeV/c^2 was applied as well before rebinning to the final 0.4 MeV/c^2 bin widths. The fits were performed with the width and/or mean of the Gaussian fixed to the values extracted from simulations, as well as leaving both as free parameters. Only fits with a $\chi^2/\text{ndf} < 3$ and a width of the signal peak within $0.5 < \sigma/\sigma_{MC} < 2$ were considered.

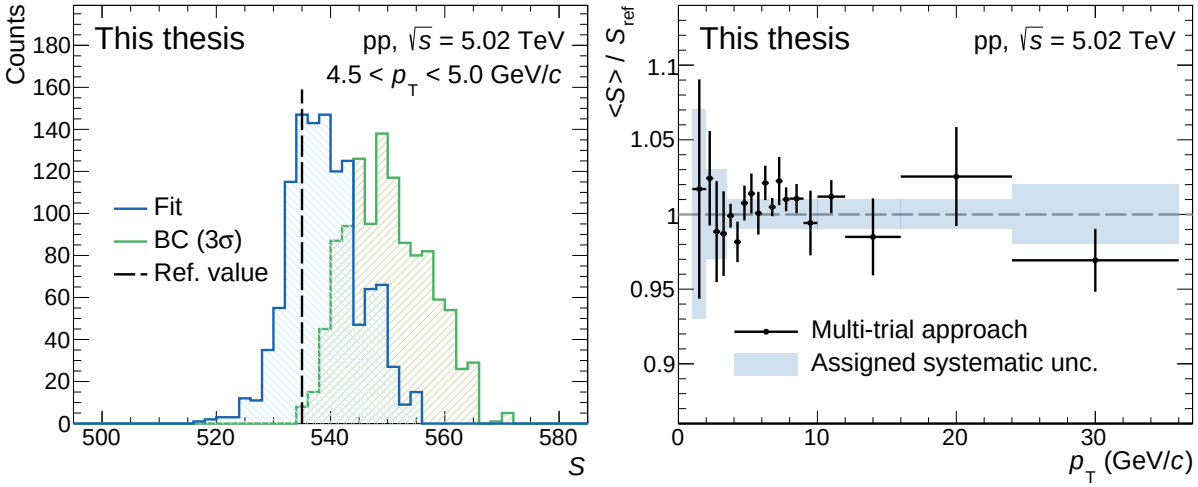


Figure 5.5: Left: Raw-yield distributions obtained with the multi-trial approach via the fit and bin-counting (BC) procedures for D^{*+} mesons in $4.5 < p_T < 5$ GeV/c in pp collisions at $\sqrt{s} = 5.02$ TeV. Right: Summary of the multi-trial approach in each D^{*+}-meson p_T interval, showing the mean of the raw-yield distribution divided by the reference raw-yield values from the central fits (the error bars represent the RMS of the distribution) and the final assigned relative systematic uncertainties.

The same strategy was performed using a bin-counting method, in which the signal yield was obtained by integrating the invariant-mass histogram in a $\pm 3\sigma$ region around the D^{*+} peak after subtracting the background, estimated from a fit to the side-bands only. This test was performed to have a signal-shape independent estimation of the raw yield, and is therefore sensitive to the non-Gaussian tails in the D^{*+} signal shape. The difference between the fit and bin-counting method is of the order of 2% for all p_T intervals, which is in agreement with a 1–2% effect estimated when using the sum of two Gaussian functions for the signal fit function as discussed above.

The left panel of Fig. 5.5 shows an example of the multi-trial distributions for $4.5 < p_T < 5$ GeV/c. The systematic uncertainty was evaluated by considering the RMS of the distributions of the signal yields obtained from all these variations as well as the shift from the central fit to the mean of the distribution, presented in the right panel of Fig. 5.5. The final assigned relative systematic uncertainties range between 1% and 7%. As additional stability check for p_T intervals with low statistics, this multi-trial approach was repeated for different final invariant-mass bin widths (from 0.1 to 0.6 MeV/c^2). No significant deviations from the central histogram binning were found.

5.5.2 Track-reconstruction efficiency

The systematic uncertainty on the track-reconstruction efficiency includes contributions of the track-finding procedure in the TPC, the prolongation to the ITS detector, and possible imperfections in describing the track-quality selections in simulation. It was estimated by varying the track-quality selection criteria and by comparing the probability to prolong the tracks from the TPC to the ITS hits (matching efficiency) in data and simulation.

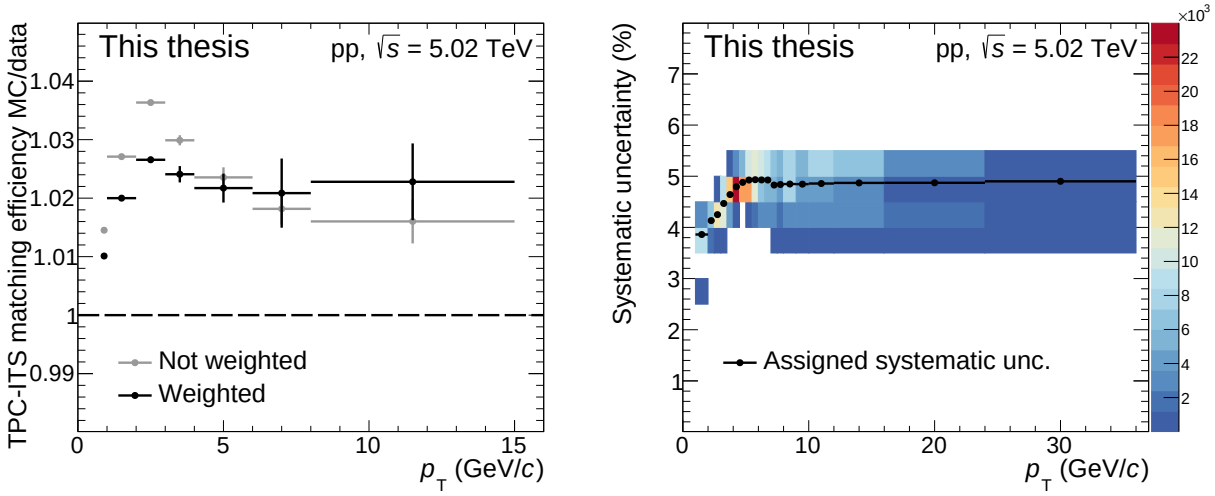


Figure 5.6: Left: The relative difference between the TPC-ITS matching efficiency in data and simulation before and after the weighting procedure for primary and secondary particles. Right: The assigned systematic uncertainty due to the track-reconstruction efficiency. See text for more details.

To evaluate a possible systematic uncertainty due to the track-quality selections on the decay daughters from the D^0, the D^{*+} analysis was repeated with three alternative TPC track-quality criteria. In particular, a number of TPC crossed rows larger than $120-5/(p_\mathrm{T}\ [\mathrm{GeV}/c])$, a number of TPC clusters of at least 0.65 times the number of TPC crossed rows, or a ratio of crossed rows to findable clusters in the TPC of 0.9, were required. For the soft-pion candidate track, for which no TPC track-quality requirements were required (see Sec. 4.2.1), three alternative ITS track-quality criteria were used instead. Instead of at least one hit in the SPD, variations requiring a hit in the first or both SPD layers were considered, as well as a minimum of three points in the full ITS. The variations of the D^{*+} cross section with respect to the central analysis were observed to be 1% for the D^0 track-quality variations (0.5% per track) and 1% for the variations applied to the soft-pion track, both without a p_T-dependence.

To obtain the matching efficiency, defined as the fraction of TPC tracks that have a prolongation in the ITS (see Sec. 3.4), the abundances of primary and secondary particles in data (see Ref. [210] for the definitions) were estimated via template fits to the track impact-parameter d_0^{xy} distributions, where the relative abundances in the simulation were weighted to match those in data. This is motivated by the observation that the ITS-TPC matching efficiency is much larger for primary produced particles. After the weighting procedure, the systematic uncertainty due to the matching efficiency was defined as the relative difference of the efficiencies in data and simulation, as shown in left panel of Fig. 5.6. These values were assigned as a p_T-dependent per-track uncertainty for the D^0 decay daughters, while for the soft-pion track no uncertainty due to the matching efficiency was considered.

The total assigned systematic uncertainty due to the tracking efficiency follows from a quadratic sum of the different per-track contributions, propagated to the D^{*+} mesons via the kinematics of the decay tracks. The final assigned values range from 3.5% to 5%, as shown in the right panel of Fig. 5.6.

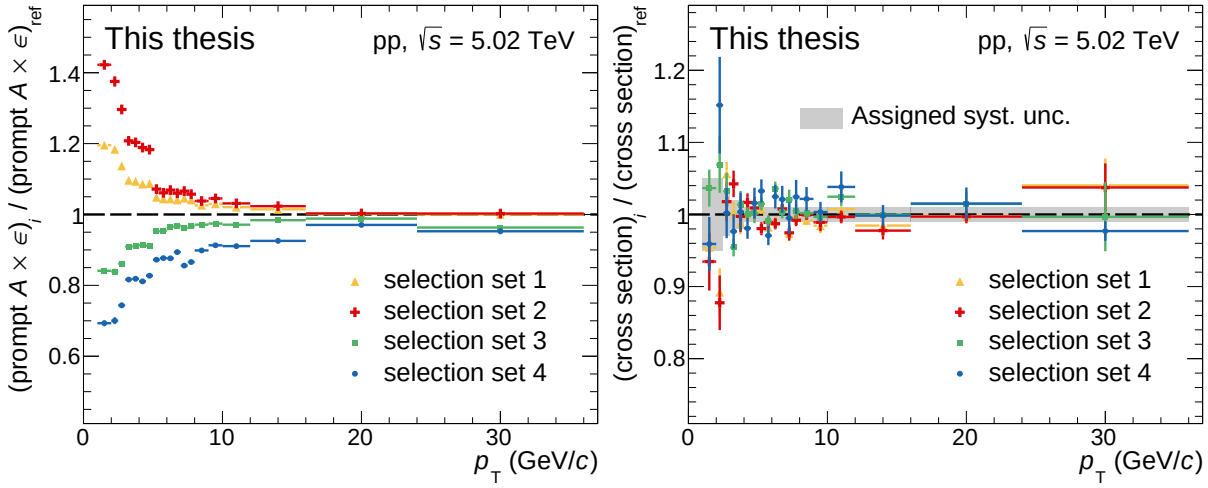


Figure 5.7: Ratio of the prompt acceptance-times-efficiency (left) and cross section (right) for D^{*+} mesons in pp collisions at $\sqrt{s}$ = 5.02 TeV between the four sets with varied topological selections and the central analysis. The final assigned relative systematic uncertainties are presented by the grey band.

5.5.3 Topological, kinematical, and PID selection efficiency

The systematic uncertainty on the selection efficiency on the D^{*+} decay topology originates from imperfections in the simulation of the decay kinematics and topology and of the resolutions and alignments of detectors in the simulation. To evaluate a systematic uncertainty, the analysis was repeated with different sets of selection criteria, resulting in a significant modification of the efficiencies, raw yield, and background values. The most effective D^{*+} selection variables, in particular the distance of closest approach of the D^0 decay tracks and the reconstructed secondary vertex, the D^0 $\cos\theta_{\text{p}}$ and $L_{\text{xy}}/\sigma(L_{\text{xy}})$, and the $d_0^{\text{xy}}(K) \cdot d_0^{\text{xy}}(\pi)$, were varied simultaneously with respect to the chosen central values resulting in two looser and two tighter sets[3]. In the left panel of Fig. 5.7, the difference in prompt acceptance-times-efficiency for the different sets with respect to the central selections are shown. At high p_{T}, the differences are rather small due to the loose topological selections.

The analysis was repeated for each set in the same way as discussed before, only the width of the Gaussian function were fixed to the value of the central analysis in order to limit fluctuations due to the fit performance. The systematic uncertainty was assigned considering the RMS with respect to unity of the cross-section ratio of each set with the central one, as shown in the right panel of Fig. 5.7. The relative systematic uncertainties range from 1% to 5%, depending on p_{T}.

Similarly, a difference in the $N_\sigma^{\text{TPC, TOF}}(\pi, \text{K})$ PID distributions in data and simulation can bias the measurement. To estimate the systematic uncertainty on the PID selection efficiency, the analysis was repeated without PID selections. The signal extraction was feasible in all p_{T} intervals, except for $1 < p_{\text{T}} < 2$ GeV/c, for which a tighter $N_\sigma^{\text{TPC}} < 2$ and $N_\sigma^{\text{TOF}} < 2$ compatibility selection with the pion or kaon mass was asked instead. The resulting cross sections were found

[3]The variations were limited to four sets because of practical reasons. As we will see later in Chapter 6 and 7, a much more thorough evaluation of this source of systematic uncertainty is performed for the analyses using BDTs.

Table 5.2: Summary of the relative systematic uncertainties on the D^{*+} measurement in different p_{T} intervals.

p_{T} (GeV/c)		1.0–2.0	2–2.5	4.5–5.0	6.0–6.5	10–12	24–36
Signal yield	(δ_1)	7%	3%	1%	1%	1%	2%
Tracking efficiency	(δ_2)	4%	4%	5%	5%	5%	5%
Selection efficiency	(δ_3)	5%	3%	1%	1%	1%	1%
PID efficiency	(δ_4)	0	0	0	0	0	0
p_{T} shape in MC	(δ_5)	5%	1%	0	0	0	0
Feed-down	(δ_6)	$^{+3.8}_{-4.3}$%	$^{+3.0}_{-3.5}$%	$^{+1.9}_{-2.6}$%	$^{+1.3}_{-1.9}$%	$^{+1.8}_{-2.5}$%	$^{+1.0}_{-1.2}$%
Branching ratio	(δ_7)	1.3%					
Luminosity	(δ_8)	2.1%					

to be compatible with those obtained with the conservative $N_\sigma < 3$ compatibility PID selection and therefore no systematic uncertainty was assigned. This was confirmed with a per-track study performed for the $\mathrm{D_s^+}$ analysis in the same data sample [211], for which an analysis without PID selection was not feasible. Relatively pure samples of pions and kaons in the TPC were selected from $\mathrm{V^0}$ ($\mathrm{K^0_S}$ and Λ) decays and tracks satisfying $N_\sigma^{\mathrm{TOF}}(\mathrm{K}) < 0.2$, respectively. The relative difference between the PID selection efficiencies for the $N_\sigma < 3$ compatibility selection in data and simulation was assigned as a p_{T}-dependent per-track systematic uncertainty, propagated to the D mesons via the kinematics of the decay-daughter tracks. The final systematic effect is below 1%, confirming the decision of assigning no uncertainty due to the PID selection efficiency.

5.5.4 Generated MC p_{T} shape

The systematic uncertainty due to the generated D^{*+}-meson p_{T} shape, which could introduce a bias in case of a non-realistic distribution, was estimated by using FONLL [83] as an alternative generator with respect to PYTHIA to simulate the D^{*+}-meson p_{T} distribution. The weights are defined as $(\mathrm{d}N/\mathrm{d}p_{\mathrm{T}})_{\mathrm{FONLL}}/(\mathrm{d}N/\mathrm{d}p_{\mathrm{T}})_{\mathrm{PYTHIA6}}$ and applied to both the generated and reconstructed D^{*+} mesons in the $(A' \cdot \epsilon)$ determination. The systematic uncertainty is defined as the ratio between the prompt efficiencies for both p_{T}-shape hypotheses, and was found to be 5% for $1 < p_{\mathrm{T}} < 2$ GeV/c, 1% for $2 < p_{\mathrm{T}} < 2.5$ GeV/c, and negligible at higher p_{T}. The efficiency varies more steeply in the low-p_{T} region, hence it is more sensitive to the generated p_{T} shape.

5.5.5 Summary

The contributions of these different sources of uncertainties (δ_1–δ_6) were summed in quadrature to obtain the total systematic uncertainty in each p_{T} interval, which varies from 5.4% to 11.3% for the D^{*+} meson. This is of the same order as the total systematic uncertainty for the D^0 and D^+ mesons, while for the $\mathrm{D_s^+}$ meson it is in general 2% larger. The p_{T}-differential cross sections have an additional global normalisation uncertainty due to the uncertainties on the integrated luminosity [200] and on the branching ratios of the considered D-meson decays [16]. A summary of the systematic uncertainties for the D^{*+} meson is reported in Table 5.2 for different p_{T} intervals.

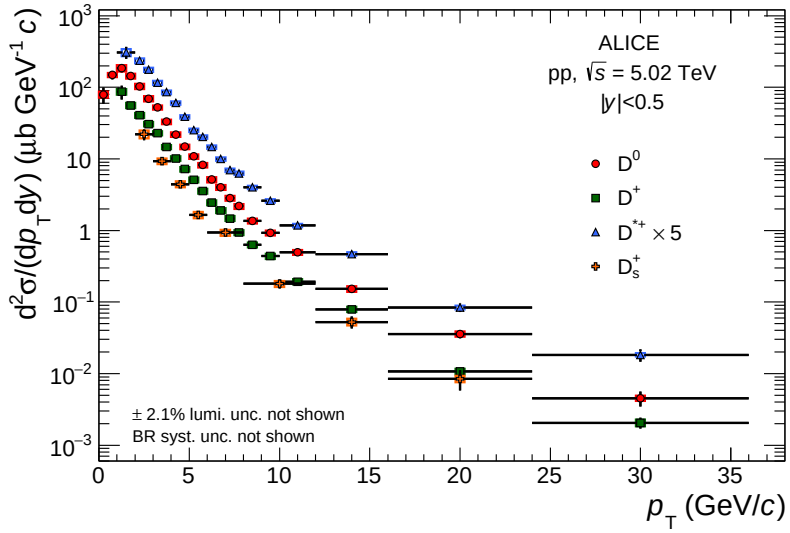


Figure 5.8: The p_{T}-differential production cross section of prompt D^0, D^+, D^{*+}, and $\mathrm{D}_{\mathrm{s}}^+$ mesons in pp collisions at $\sqrt{s} = 5.02$ TeV. For the D^0 meson, the result in the $0 < p_{\mathrm{T}} < 1$ GeV/c interval is obtained from the analysis without decay-vertex reconstruction. The D^{*+}-meson cross section is scaled by a factor 5 for better visibility.

The systematic uncertainty on the raw-yield extraction (δ_1) is considered uncorrelated among the different p_{T} intervals given the substantial variations in the S/B as function of p_{T}, while the other contributions (δ_{2-8}), are mostly correlated among the different p_{T} intervals. In case of charm-hadron ratios, e.g. $\mathrm{D}^{*+}/\mathrm{D}^0$, the total systematic uncertainties are computed treating the contributions due to the tracking efficiency, the PID efficiency, the feed-down subtraction, and luminosity as correlated among the different D-meson species ($\delta_{2,4,6,8}$), while all other sources were propagated as uncorrelated. Finally, for the ratio of production cross sections at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 5.02$ TeV, the contribution originating from the feed-down subtraction (δ_6) was treated as correlated, while all the other systematic uncertainties on the cross sections were propagated as uncorrelated between the measurements at the two different energies, except for the uncertainty on the BR (δ_7) which cancels out in such ratio.

5.6 Results

The p_{T}-differential cross section for prompt D^{*+}-meson production in pp collisions at $\sqrt{s} =$ 5.02 TeV in $|y| < 0.5$ is depicted in Fig. 5.8. The vertical error bars represent the statistical uncertainties and the systematic uncertainties are depicted as boxes around the data points, as for all the figures in the result sections. Thanks to the higher integrated luminosity of pp collisions at $\sqrt{s} = 5.02$ TeV, the here presented measurement of the D^{*+} meson has a wider p_{T} coverage and is, in addition, more differential in p_{T} with respect to a previous ALICE measurement at $\sqrt{s} = 7$ TeV [169]. Compared to this measurement at $\sqrt{s} = 7$ TeV, the total statistical and systematic uncertainties are reduced by a factor of about 1.4 and 1.1–1.6[4], respectively.

[4]This reduction of about a factor 1.6 is mainly driven by the systematic uncertainty due to the selection efficiency (δ_3) for which a conservative choice at high p_{T} was made for the analysis at $\sqrt{s} = 7$ TeV.

Figure 5.8 shows also the p_{T}-differential cross section for prompt D^0, D^+, and D_s^+ mesons, for which the analyses were performed in close collaboration with the here presented D^{*+} production measurement. The D mesons were reconstructed at midrapidity via their hadronic decay channels $\mathrm{D}^0 \to \mathrm{K}^-\pi^+$, $\mathrm{D}^+ \to \mathrm{K}^+\mathrm{K}^-\pi^+$, and $\mathrm{D}_\mathrm{s}^+ \to \phi\pi^+ \to \mathrm{K}^+\mathrm{K}^-\pi^+$, exploiting a similar displaced decay-vertex topology. Instead, an alternative analysis technique based on particle identification and on the estimation and subtraction of the combinatorial background was used to measure the production of D^0 mesons down to $p_{\mathrm{T}} = 0$ [82]. The prompt D^0-meson p_{T}-differential cross section is compatible with the one measured by the CMS collaboration at the same centre-of-mass energy in $|y| < 1$ and $2 < p_{\mathrm{T}} < 100$ GeV/c [212], while for the D^{*+}, D^+, and D_s^+ mesons no other $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}$ measurements at midrapidity and $\sqrt{s} = 5.02$ TeV exist.

In Fig. 5.9, the measured prompt D^{*+}-meson p_{T}-differential cross section is compared with results of pQCD calculations with different schemes, in particular FONLL predictions [83], two calculations using the GM-VFNS framework with different prescriptions to regulate the divergences at low p_{T} (dubbed as mod-$\mu_{\mathrm{R,F}}$ [84] and SACOT-m_{T} [204]), and a calculation based on k_{T} factorisation [85]. In the GM-VFNS(mod-$\mu_{\mathrm{R,F}}$) model, the value of charm mass is set to 1.3 GeV/c^2, while for the other predictions the mass is set to 1.5 GeV/c^2. Furthermore, the four frameworks utilise different sets of PDFs (CTEQ6.6 [209] for FONLL, CTEQ14 [213] for GM-VFNS(mod-$\mu_{\mathrm{R,F}}$), NNPDF3.1 [214] for GM-VFNS(SACOT-m_{T}), and MMHT2014 [215] for the k_{T} factorisation calculation) and different fragmentation functions. More details on the conceptional differences between these schemes were previously discussed in Section 2.1. The theoretical uncertainties are estimated by varying the factorisation and renormalisation scales (μ_{F}, μ_{R}) and the charm-quark mass in FONLL and the k_{T}-factorisation calculations, while for the GM-VFNS predictions only μ_{R} and $\mu_{\mathrm{R,F}}$ are varied for GM-VFNS(mod-$\mu_{\mathrm{R,F}}$) and GM-VFNS(SACOT-m_{T}), respectively.

The measured D^{*+} cross section, as well as the cross section for the other non-strange D mesons [82], is described within uncertainties by FONLL and the two GM-VFNS calculations. The data lie systematically on the upper edge of the uncertainty band of the FONLL predictions, while for the two calculations in the GM-VFNS framework, the central values of the predictions tend to underestimate the data at low and intermediate p_{T}. The k_{T}-factorisation predictions describe the data at low and intermediate p_{T}, but overshoot them for $p_{\mathrm{T}} \gtrsim 10$ GeV/c. The D_s^+-meson production tends to be underestimated by the three pQCD calculations (predictions from the FONLL framework are not included because the D_s^+-meson fragmentation functions are not included) in the measured p_{T} range [82].

The p_{T}-differential production cross section of prompt D^{*+} mesons is also compared in Fig. 5.9 to the predictions obtained from the PYTHIA 8 event generator with the default Monash 2013 tune [151], with and without colour-reconnection beyond leading colour approximation [110] (see Sec. 2.4). The theoretical uncertainty shown are purely statistical due to the limited size of the simulations. At high p_{T}, both PYTHIA calculations are compatible with the measurement, while at low p_{T} they undershoot the data by a factor 1.5–2. This discrepancy is likely due to a mistake that entered in the Monash tuning process, where the wrong D^{*+}/D ratio was used because the

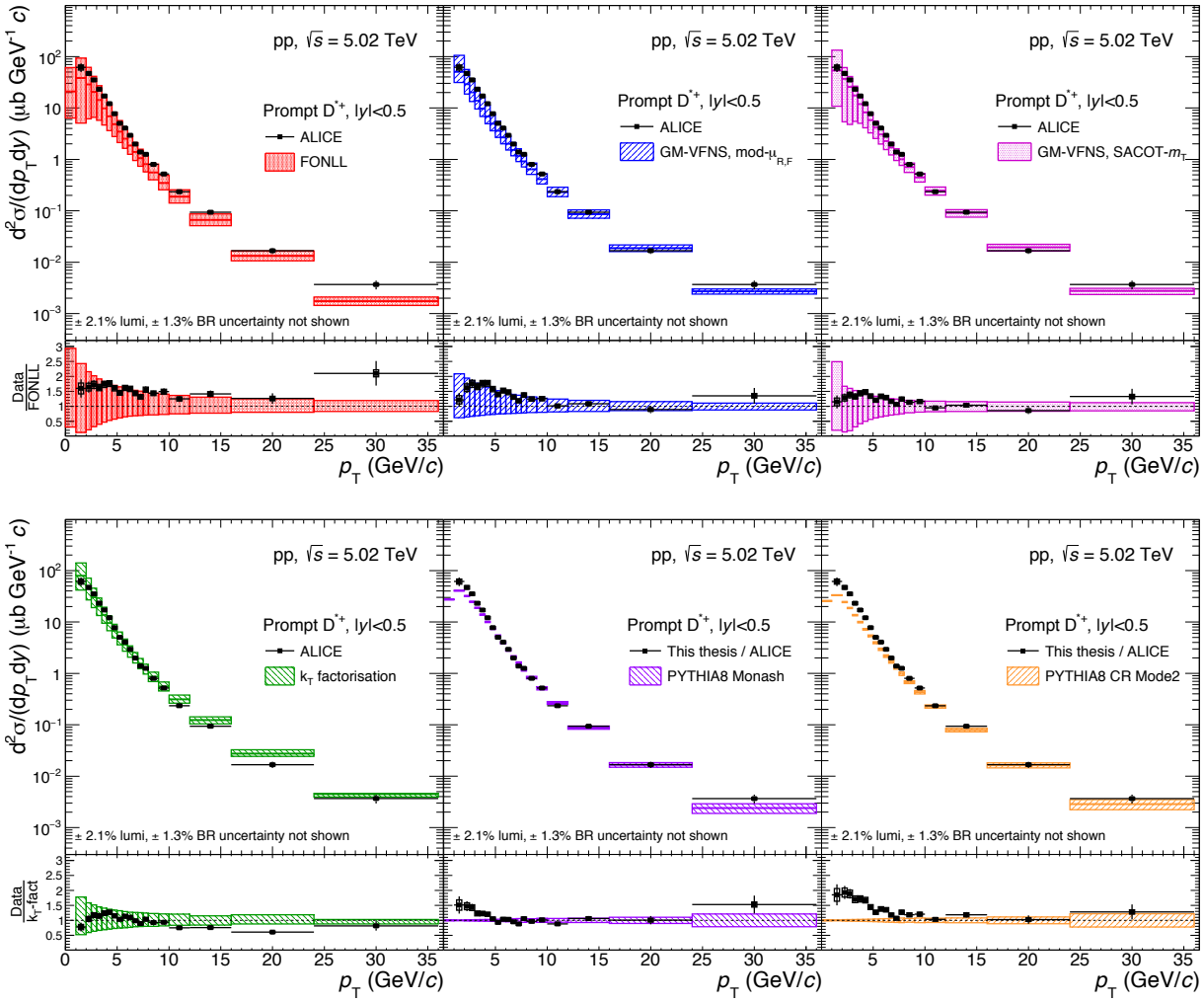


Figure 5.9: The p_{T}-differential production cross section of prompt D^{*+} mesons in pp collisions at $\sqrt{s} = 5.02$ TeV compared to pQCD calculations (FONLL [83], GM-VFNS(mod-$\mu_{\mathrm{R,F}}$) [84], GM-VFNS(SACOT-m_{T}) [204], and k_{T} factorisation [85]) and to predictions from MC event generators (PYTHIA 8 with the Monash tune [151] with and without colour-reconnection beyond leading colour approximation [110]). The ratios between the data and the theoretical predictions are shown in the lower panels.

rates were taken from separate, inconsistent sources (P. Skands, private communication, May 2021).

To study the evolution of prompt D-meson production with the centre-of-mass energy of the collision, the ratios of the production cross sections in pp collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 5.02$ TeV were computed and compared to FONLL predictions. This is an important comparison for the heavy-flavour production measurements in p–Pb and Pb–Pb collisions performed during LHC Run 1 [144, 170], since the required pp reference for these measurements was measured at a different centre-of-mass energy and relied on a p_{T}-dependent $\sqrt{s}$-scaling factor extracted from FONLL calculations. In Fig. 5.10, the results for the D^{*+} and D^{0} mesons are shown, which agree with FONLL predictions within one standard deviation. In the FONLL predictions, the uncertainties originating from scale variations and from the PDFs cancel out to a large extent [91].

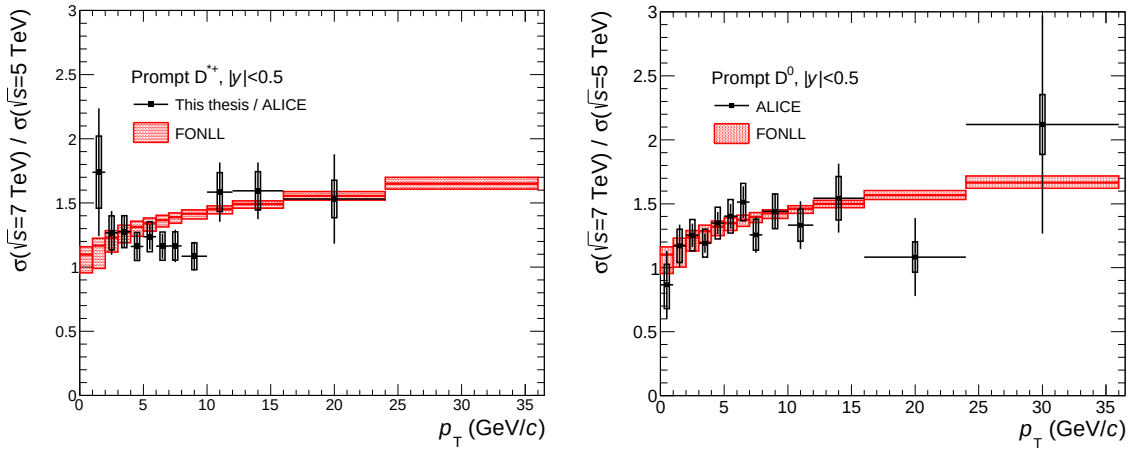


Figure 5.10: Ratios of D^{*+} (left) and D^0 (right) production cross sections in pp collisions at $\sqrt{s} = 7$ TeV [169] and $\sqrt{s} = 5.02$ TeV compared to FONLL pQCD calculations [83].

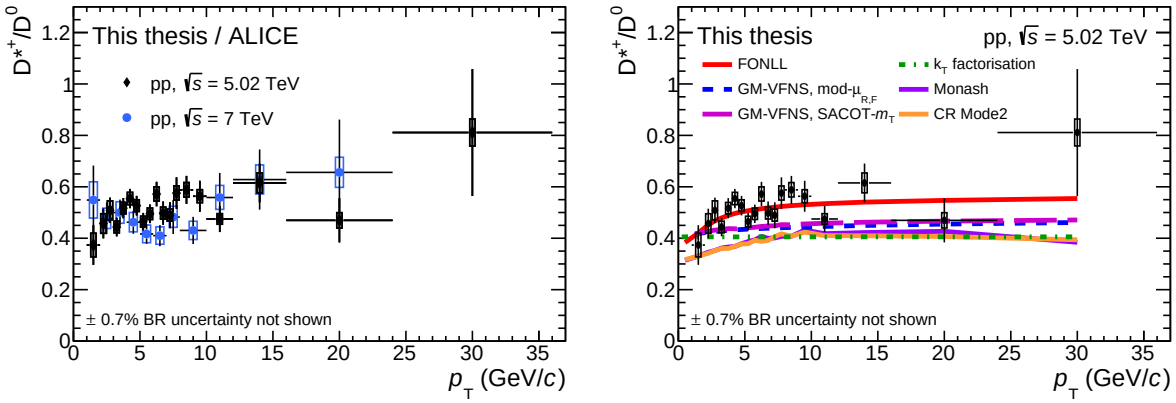


Figure 5.11: Left: The $\mathrm{D}^{*+}/\mathrm{D}^0$ production ratio in pp collisions at $\sqrt{s} = 5.02$ TeV and $\sqrt{s} = 7$ TeV [169]. Right: The $\mathrm{D}^{*+}/\mathrm{D}^0$ at $\sqrt{s} = 5.02$ TeV compared to pQCD calculations (FONLL [83], GM-VFNS(mod-$\mu_{\mathrm{R,F}}$) [84], GM-VFNS(SACOT-m_{T}) [204], and k_{T} factorisation [85]) and to predictions from MC event generators (PYTHIA 8 with the Monash tune [151] with and without colour-reconnection beyond leading colour approximation [110]).

The ratios of the p_{T}-differential cross section of prompt D mesons, in particular the measured $\mathrm{D}^{*+}/\mathrm{D}^0$, are presented in Fig. 5.11 for pp collisions at $\sqrt{s} = 5.02$ TeV and $\sqrt{s} = 7$ TeV. The production ratios at the two different energies are compatible within uncertainties. The $\mathrm{D}^{*+}/\mathrm{D}^0$ ratio at $\sqrt{s} = 5.02$ TeV is compared as well to the different pQCD calculations and to predictions from the PYTHIA 8 event generator. The FONLL predictions, with fragmentation fractions taken from $\mathrm{e}^+\mathrm{e}^-$ collision data [201], describe the data, while the other models are slightly undershooting the ratio.

Finally, the visible D^{*+}-meson production cross section was evaluated by integrating the p_{T}-differential cross sections over the narrower p_{T} intervals in $1 < p_{\mathrm{T}} < 36$ GeV/c. Starting from this visible cross section, the production cross section per unit of rapidity $\mathrm{d}\sigma/\mathrm{d}y$ in $|y| < 0.5$ was computed based on an extrapolation procedure to account for the $p_{\mathrm{T}} < 1$ GeV/c and $p_{\mathrm{T}} > 36$ GeV/c intervals. The extrapolation factor was computed using the FONLL central parameters (and verified with those from the GM-VFNS calculation) to evaluate the ratio between the total

Table 5.3: Visible and p_{T}-extrapolated production cross section of prompt D^{*+} mesons in $|y| < 0.5$ in pp collisions at $\sqrt{s} = 5.02$ TeV.

Kinematic range	Extr. factor	$\mathrm{d}\sigma/\mathrm{d}y|_{|y|<0.5}$ (µb)
$1 < p_{\mathrm{T}} < 36$ GeV/c	n.a.	$143 \pm 12\,(\mathrm{stat}) \pm 11\,(\mathrm{syst}) \pm 3\,(\mathrm{lumi}) \pm 2\,(\mathrm{BR})$
$p_{\mathrm{T}} > 0$	$1.24^{+0.34}_{-0.08}$	$178 \pm 15\,(\mathrm{stat}) \pm 14\,(\mathrm{syst}) \pm 4\,(\mathrm{lumi}) \pm 2\,(\mathrm{BR})\,^{+48}_{-12}\,(\mathrm{extr})$

production cross section in $|y| < 0.5$ and that in the experimentally covered p_{T} range. The systematic uncertainty on the extrapolation factor was estimated considering sources due to the PDFs uncertainties and due to the variations of the charm-quark mass and renormalisation and factorisation scales. The computed visible and extrapolated prompt D^{*+}-meson $\mathrm{d}\sigma/\mathrm{d}y$ in $|y| < 0.5$ are presented in Tab. 5.3.

The ratios of the p_{T}-integrated yields of the different D-meson species were computed from the cross sections integrated over the common p_{T} range. The measured $\sigma(\mathrm{D}^{*+})/\sigma(\mathrm{D}^0)$ in $1 < p_{\mathrm{T}} < 36$ GeV/c is $0.43 \pm 0.04\,(\mathrm{stat}) \pm 0.03\,(\mathrm{syst}) \pm 0.003\,(\mathrm{BR})$. This is compatible within uncertainties with the ALICE results in pp at $\sqrt{s} = 7$ TeV [169] and Pb–Pb at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [144], as well as with the pp measurements of the LHCb Collaboration at forward rapidity ($2.0 < y < 4.5$) at three different collision energies [101, 216, 217] and the results in $\mathrm{e}^+\mathrm{e}^-$ collision systems measured at LEP [218].

5.7 Discussion

In this chapter, the analysis and the main physics results of the prompt D^{*+} production measurement in pp collisions at $\sqrt{s} = 5.02$ TeV were discussed. Thanks to the higher integrated luminosity of this data sample, the measurement is more precise, more differential, and has a wider p_{T} coverage with respect to previous ALICE measurements [169, 205]. Measurements of D mesons (in pp, but also in the larger systems) are, however, no longer front-page news, and much have been said over the years. In this section, I will contribute as well and discuss (my personal) interpretation of these results.

I will start with the first sentence of this chapter: "Heavy-flavour production measurements in pp are a sensitive test of QCD calculations with the factorisation approach". As we have seen in Fig. 5.9, the p_{T}-differential cross section of prompt D mesons are described within the theoretical uncertainties by pQCD models, indicating our understanding of charm meson production is quite good. The p_{T} shape of the cross sections is best reproduced by the central FONLL predictions, although it underestimates the magnitude by about 50%. This may indicate a preference of the data for a charm-quark mass smaller than the default of 1.5 GeV/c^2. Nevertheless, the p_{T} shape of central FONLL calculations is still a good starting point for more-advanced charm-hadron models. For example, it is the shape of the p_{T} distribution that affects the R_{AA} and v_2 observables. A recent collaboration between many authors of heavy-flavour diffusion and energy-loss models therefore also used FONLL calculations as generic input to compare their frameworks [219].

In Fig. 5.9, we can also note that the current experimental uncertainties are much smaller than the theoretical ones. Naively, one would thus expect a large constraining power to the model calculations. The theoretical uncertainties arise, however, mainly from variations in the choice of the scales for the perturbative calculations (dominant with respect to uncertainties on the PDFs and FFs), hence not much can be improved here[5]. As discussed in the text, it is therefore more interesting to look at cross-section ratios, for which the uncertainties originating from scale variations mostly cancel.

In Fig. 5.10, the ratios of the production cross section in pp collisions at $\sqrt{s}$ = 7 TeV and $\sqrt{s}$ = 5.02 TeV were thus computed and compared to FONLL calculations. Here, however, the experimental and theoretical uncertainties are of similar magnitude, hence there is again limited power to constrain the calculation. The observation that FONLL reproduces well the p_{T} shape of this $\sigma(7\,\mathrm{TeV})/\sigma(5.02\,\mathrm{TeV})$ ratio, tells us that the conclusions made on the nuclear modification of charm quarks based on earlier R_{AA} (and R_{pPb}) measurements of D mesons — which used an $\sqrt{s}$-scaled pp reference at $\sqrt{s}$ = 7 TeV [144, 170] — remain true. Of course, any future nuclear modification measurement at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV should use this new, and more precise, measurement at the same collision energy as pp baseline. This was indeed the case [130, 220].

As discussed, the ratios of production cross sections between different charm-hadron species probe mainly the hadronisation part of the production process. With the presented $\mathrm{D}^{*+}/\mathrm{D}^0$ ratio in Fig. 5.11, the fragmentation functions for vector mesons can be studied. At the end of Chapter 2, we noted that for a long time only one dedicated parametrisation for vector D mesons was on the market [171], and therefore the spread among the theoretical predictions for the $\mathrm{D}^{*+}/\mathrm{D}^0$ ratio at $\sqrt{s} = 7$ TeV was much larger than, for example, for the $\mathrm{D}^+/\mathrm{D}^0$ ratio. There has, however, been quite some theoretical efforts regarding the D^{*+} fragmentation function in the last years [221, 222]. These studies were triggered by methodological and theoretical improvements of the theoretical methods and observed large discrepancies between data and predictions for D^{*+} in jet production [223] (and in general by the new D^{*+} measurements at the LHC). These efforts are not yet incorporated in the model predictions shown in this chapter[6], and might improve the model-data comparison of the $\mathrm{D}^{*+}/\mathrm{D}^0$ ratio shown in Fig. 5.11. Let me end by noting that the $\mathrm{D}^{*+}/\mathrm{D}^+$ ratio, even though I did not study it, might be more natural to study because of the same quark content ($\mathrm{c\bar{d}}$).

There is one other cross-section ratio that can help to constrain the pQCD model calculations, namely the ratio in different rapidity regions exploiting the LHCb measurement at forward rapidities [216], as presented for D^0 and D^{*+} mesons in Fig. 5.12. The ratios of the — more precise — D^0 production cross sections between mid- and forward rapidities show a hint of a different slope between data and FONLL calculations. This observation might help to improve

[5] I ignore here the "obvious" improvements to include analytic solutions of higher-order corrections in the theoretical calculations, which is not a trivial task due to the complexity of QCD.

[6] The D^{*+} fragmentation function for the k_{T}-factorisation predictions actually changed between Refs. [94] and [85], but I consider this a downgrade. Instead of using the dedicated FF for D^{*+} of Ref. [171], they are now using the Peterson fragmentation function [145] for all charm hadrons, hence the constant prediction with p_{T}.

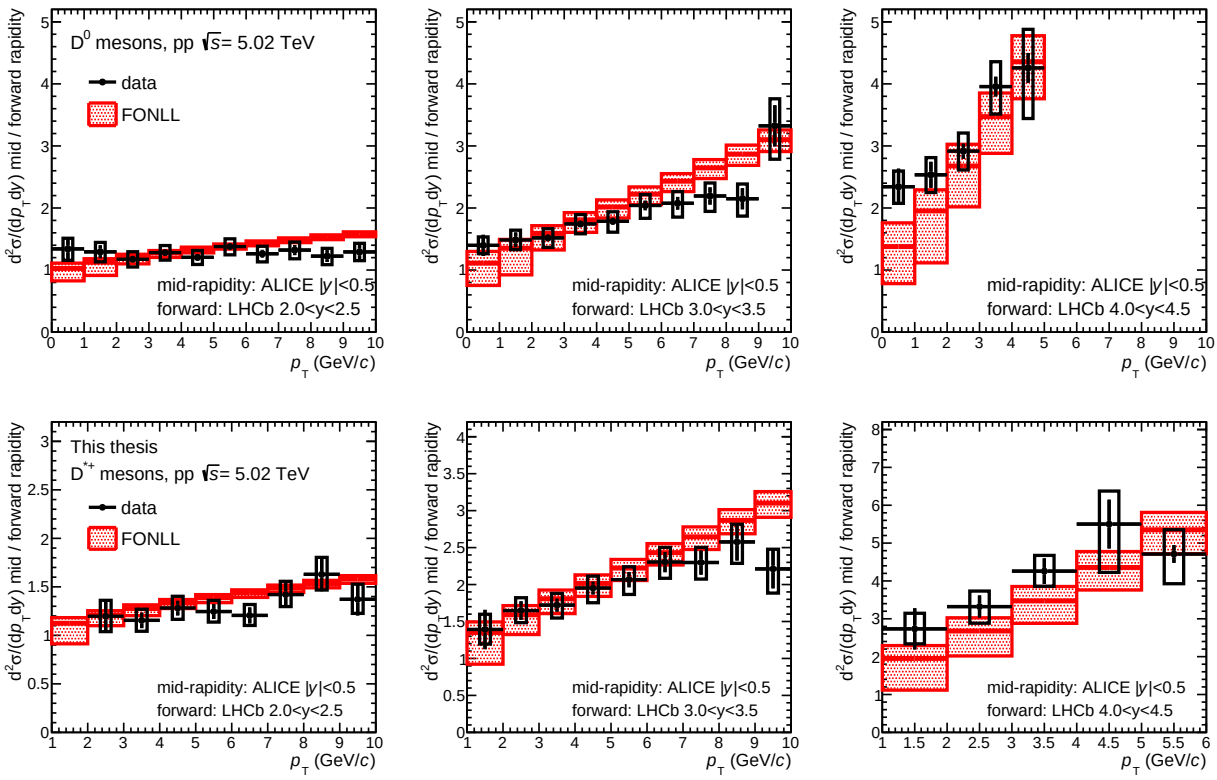


Figure 5.12: Ratios of D^0 (top) and D^{*+} (bottom) production cross section per unit of rapidity at mid-rapidity ($|y| < 0.5$) to those measured by the LHCb Collaboration [216] in three forward rapidity regions, as a function of p_T. Predictions from FONLL calculations [83] are compared to the data.

the rapidity description of D meson productions in FONLL calculations, via, for example, constraints on the gluon PDFs as advertised in Ref. [91].

To summarise, I think that charm-meson production measurements have been very useful in the past decade to improve our understanding of the charm sector at the LHC. However, we are currently at the level of "optimal experimental precision". New spectra measurements of the D^0, D^+, and D^{*+} at the (three times larger) $\sqrt{s} = 13$ TeV pp data sample will likely not provide new insights, nor constrain the theoretical models further. It is time to move to more differential studies (e.g. production versus multiplicity or production in jets) or more exotic hadrons (e.g. heavy-flavour baryons or $B_{s/c}^{0/+}$ mesons). In the next chapter, we will already see an example of such analysis.

CHAPTER 6

Charm-hadron production versus multiplicity in pp collisions at $\sqrt{s} = 13$ TeV

Abstract – The production of prompt D^0, D_s^+, and Λ_c^+ hadrons, and their ratios, $\mathrm{D}_\mathrm{s}^+/\mathrm{D}^0$ and $\Lambda_\mathrm{c}^+/\mathrm{D}^0$, are measured in pp collisions at $\sqrt{s} = 13$ TeV at midrapidity ($|y| < 0.5$) with the ALICE detector. The measurements are performed as a function of the candidate transverse momentum, in intervals of charged-particle multiplicity, measured with two multiplicity estimators covering different pseudorapidity regions. While the strange to non-strange $\mathrm{D}_\mathrm{s}^+/\mathrm{D}^0$ ratios indicate no significant multiplicity dependence, the baryon-to-meson $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ ratios show a multiplicity-dependent enhancement, with a significance of 5.3σ for $1 < p_\mathrm{T} < 12$ GeV/c, comparing the highest multiplicity interval with respect to the lowest one. On the contrary, the p_T-integrated ratios, extrapolated down to $p_\mathrm{T} = 0$, do not show a significant dependence on multiplicity within the (relatively large) uncertainties. The measured charm-hadron yield ratios are compared to PYTHIA predictions and to calculations based on a statistical hadronisation model, as well as to the measured ratios in p–Pb and Pb–Pb collisions.

In collaboration with A. Gal, B. Volkel, C. Terrevoli, and L. Dello Stritto
Manuscript submitted to Phys. Lett. B (arXiv:2111.11948 [nucl-ex])

6.1 Introduction

Heavy-flavour hadrons are produced in high-energy particle collisions through the hadronisation of the corresponding heavy-flavour quarks, which in turn typically originate in early hard scattering processes. As discussed extensively in Chapter 2, the most common theoretical approach to describe this production is based on the QCD factorisation theorem [77]. Regarding the hadronisation of heavy quarks, these theoretical calculations rely on the assumption that fragmentation functions, which are typically measured in e^+e^- or e^-p collisions [171], are universal across all collision systems and energies. Systematic measurements of the relative production of heavy-flavour hadrons, performed in different collision systems, provide an excellent experimental benchmark to test this assumption.

As we have seen for the results presented in Chapter 5, perturbative calculations at next-to-leading order, with next-to-leading-log resummation [83, 87, 90, 203], can successfully describe the production of strange and non-strange charm mesons and their ratios in pp collisions as a function of transverse momentum and rapidity. In contrast, theoretical calculations based on the factorisation theorem provide a poor description of heavy-flavour baryon production in hadronic collisions. In particular, measurements of the Λ_c^+ production cross section in pp and p–Pb collisions at $\sqrt{s} = 7$ TeV [100] and $\sqrt{s_{NN}} = 5.02$ TeV [102, 103] have shown a larger p_T-differential cross section compared to calculations where the charm fragmentation is tuned on e^+e^- and e^-p measurements [90, 203]. This contradiction, especially well visible when looking at the baryon-to-meson Λ_c^+/D^0 ratio, is what we called the Λ_c^+ puzzle in Chapter 2.

The measurements of the Λ_c^+/D^0 and D_s^+/D^0 ratios also play an important role in the study of quark–gluon plasma. In heavy-ion collisions, measurements of baryon-to-meson and of strange to non-strange hadron production ratios are considered to be powerful observables to study the hadronisation mechanisms in the QGP [158]. The most recent measurement of the Λ_c^+/D^0 ratio in Pb–Pb collisions will be discussed extensively in Chapter 7, but it is consistent with the hypothesis of an enhancement of the Λ_c^+/D^0 ratio with respect to pp collisions in the intermediate p_T region. A hint of enhancement of the D_s^+/D^0 ratio in central Pb–Pb collisions with respect to pp collisions was also observed at $\sqrt{s_{NN}} = 5.02$ TeV in the intermediate p_T region $2 < p_T < 8$ GeV/c, as expected in presence of a sizeable contribution of coalescence processes and increased strangeness production in the medium [163].

These observations in pp and Pb–Pb collisions triggered additional studies on charm-hadron production as function of event multiplicity. By comparing the D_s^+/D^0 and Λ_c^+/D^0 ratios in high-multiplicity pp collisions with the values measured in Pb–Pb collisions, it can be investigated if modified hadronisation mechanisms, like recombination processes, already start to play a role in pp collisions. Furthermore, low-multiplicity pp events might shed light on the differences observed in the Λ_c^+/D^0 ratio between pp and e^+e^- collision systems.

Also in the light-flavour sector there is interest in the evolution of identified hadrons with multiplicity. The ALICE and CMS Collaborations observed an enhancement of baryon-to-meson ratios (p/π and K/π) at intermediate p_T in high-multiplicity pp and p–Pb collisions, similar to

what is observed in heavy-ion collisions [132, 224]. The ALICE Collaboration performed recently also several measurements related to strangeness production versus multiplicity, considering primary ($\mathrm{K^0_S}$, Λ) and multi-strange (Ξ, Ω, ϕ) hadrons [8, 157, 225]. The mass and multiplicity dependencies are reminiscent of the patterns seen in p–Pb and Pb–Pb collisions at the LHC. Furthermore, it has been demonstrated that the evolution of the baryon-to-meson ratios as a function of $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ exhibits a universal pattern for all collision systems. This behaviour might indicate a common mechanism at work that depends solely on final-state multiplicity density. Another observation is that the enhancement in the particle ratios increases with strangeness content rather than with mass or baryon number of the hadron. The evolution of this strangeness enhancement steadily increases from low multiplicity pp to high multiplicity p–Pb and reaches the values observed in Pb–Pb collisions. This may point towards a common underlying physics mechanism which gradually compensates the strangeness suppression in fragmentation.

In this chapter, the first measurement of the production yields of prompt $\mathrm{D^0}$, $\mathrm{D_s^+}$, and Λ_c^+ and corresponding ratios, $\mathrm{D_s^+/D^0}$ and $\Lambda_\mathrm{c}^+/\mathrm{D^0}$, in pp collisions at $\sqrt{s} = 13$ TeV, as a function of the charged-particle pseudorapidity density $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$, are presented. The study was performed considering events selected according to the charged-particle density at mid- and forward rapidities, in order to investigate the effects of possible biases originating from the determination of the multiplicity in the same pseudorapidity region as in which the charm hadrons are reconstructed. This chapter is organised as follows. In Section 6.2, the event classification and general correction strategies are introduced. Sections 6.3 to 6.6 present the analysis regarding $\mathrm{D_s^+}$-meson production in midrapidity multiplicity event classes[1], discussing the selection procedure, raw-yield extraction, corrections, and evaluation of the systematic uncertainties, respectively. Finally, in Section 6.7 and 6.8, the results will be presented and discussed in detail.

6.2 Event classification

Events are classified based on their event multiplicity, calculated using two estimators. At midrapidity ($|\eta| < 1$), it was estimated via the number of tracklets N_{trkl}, defined as track segments built by associating pairs of hits in the two SPD layers. The event multiplicity in the forward rapidity region was estimated based on the percentile distribution p_{V0M} of the V0M amplitude, which is the sum of signal amplitudes in the V0A and V0C scintillators. Typical self-normalised distributions for both multiplicity estimators are shown in Fig. 6.1.

The pp collision data at $\sqrt{s} = 13$ TeV used for this analysis was collected in the years 2016, 2017, and 2018. Three trigger setups were employed. The minimum-bias (MB) trigger required signals in both V0A and V0C in coincidence with the proton bunch arrival time. To enrich the data sample in the highest multiplicity regions, high-multiplicity triggers based on a minimum selection of fired chips in the SPD (HMSPD) or of V0 amplitudes (HMV0) were used, which were fully efficient for $N_{\mathrm{trkl}} > 65$ and $p_{\mathrm{V0M}} < 0.1\%$, respectively. Offline selection criteria

[1]Since I am responsible for the $\mathrm{D_s^+}$ analysis with the midrapidity multiplicity estimator, I limit the analysis-related part of the discussion in this dissertation to this analysis, even though I made large efforts for the other $\mathrm{D_s^+}$ and $\Lambda_\mathrm{c}^+ \to \mathrm{pK^0_S}$ analyses as well.

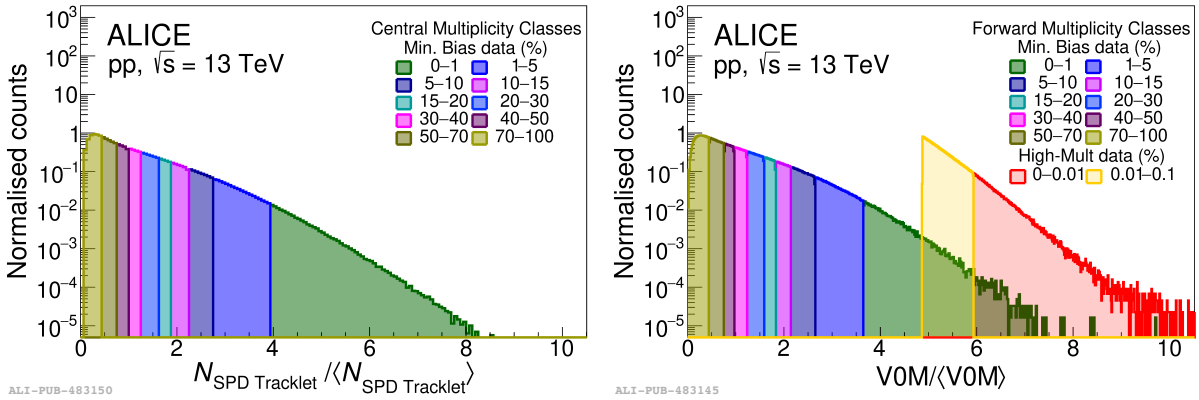


Figure 6.1: Left: The distribution of the total number of SPD tracklets in an event (within $|\eta| < 2$) scaled by its average value in pp collisions at $\sqrt{s} = 13$ TeV. Right: The distribution of the V0M amplitude scaled by its average value. The percentile values of the multiplicity classes are fractions of the visible cross section $\Delta\sigma/\sigma_{\mathrm{MB}}$. Figures taken from Ref. [226].

were applied as introduced in Sec. 4.1.1. In addition, events were required to have at least one charged particle within the pseudorapidity region $|\eta| < 1$ (the so-called INEL$_{>0}$ class). This class of events minimises diffractive corrections and therefore helps the comparison to theoretical calculations. It corresponds to about 75% of the total inelastic cross section [226, 227]. After the aforementioned selections, the data sample consists of 1.8 billion MB triggered events as well as of 0.3 billion and one billion events collected with the HMSPD and HMV0 triggers, respectively.

The events were assigned to three or four multiplicity intervals based on the corresponding observables N_{trkl} and p_{V0M}, as presented in Table 6.1. The last N_{trkl} and p_{V0M} intervals contain data collected with the HMSPD and HMV0 triggers, respectively. Table 6.1 shows also the mean charged-particle multiplicity density ($\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$) of charged-primary particles N_{ch}, since N_{trkl} and p_{V0M} are not physical primary observables. The conversion procedure will be introduced in the following subsection. Finally, the trigger correction ϵ_{INEL} is provided, which is required to account for those events which fulfil the INEL$_{>0}$ requirement but were not selected by the trigger, as specified in Ref. [226]. Where needed, a weighted average is performed to merge the finer multiplicity intervals used. The quoted uncertainty is defined as the envelope of the weighted average of the analysed periods. Note that the signal detection efficiency for charm

Table 6.1: Summary of the multiplicity event classes at midrapidity (N_{trkl}) and forward rapidity (p_{V0M} [%]). The average charged-particle densities $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ at midrapidity are shown, together with the value corresponding to the INEL$_{>0}$ event class. The trigger efficiency ϵ_{INEL} is also reported for each multiplicity interval, as estimated in Ref. [226].

Mult. estimator	Mult. interval	$\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$	ϵ_{INEL}
N_{trkl}	$[1, 9]$	3.10 ± 0.02	0.862 ± 0.015
	$[10, 29]$	10.54 ± 0.01	0.997 ± 0.002
	$[30, 59]$	22.56 ± 0.07	1 (negl. unc.)
	$[60, 99]$	37.83 ± 0.06	1 (negl. unc.)
p_{V0M} [%]	$[30, 100]$	4.41 ± 0.05	0.897 ± 0.013
	$[0.1, 30]$	13.81 ± 0.14	0.997 ± 0.001
	$[0, 0.1]$	31.53 ± 0.38	1 (negl. unc.)
INEL$_{>0}$		6.93 ± 0.09	0.920 ± 0.003

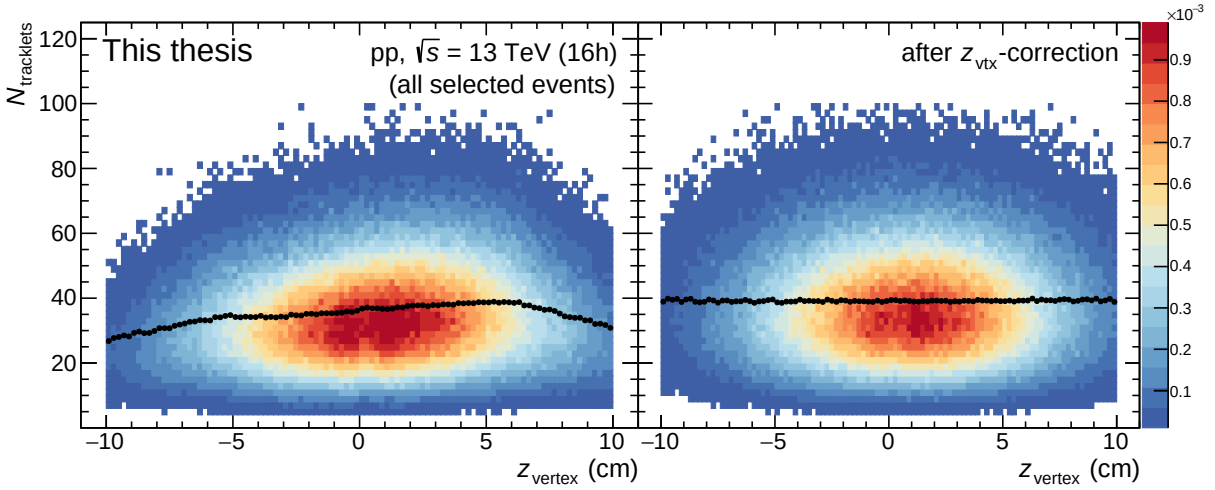


Figure 6.2: The two-dimensional N_{trkl} against z_{vtx} distribution before (left) and after (right) the z_{vtx}-correction procedure, for events belonging to the 16h period. The N_{trkl} profiles are indicated by the solid black markers.

hadrons is not modified by the INEL$_{>0}$ requirement, because of the SPD requirements for the decay tracks.

6.2.1 Corrections to multiplicity estimators

The acceptance of the SPD in pseudorapidity changes with longitudinal vertex position z_{vtx} and, in addition, the acceptance-times-efficiency changes with time due to variations of the inactive channels. Therefore, a data-driven correction procedure was applied on an event-by-event basis to the number of tracklets depending on the z_{vtx} position and data taking period. In particular, the raw profiles of the two-dimensional distribution of N_{trkl} against z_{vtx} are used to correct the measured number of tracklets for each event. The aim is to equalise the average number of tracklets among the data-taking periods in the three years and to correct for the z_{vtx}-dependency, according to $N_{\mathrm{trkl}}^{\mathrm{corr}} = N_{\mathrm{trkl}} \cdot \langle N_{\mathrm{trkl}}^{\mathrm{ref}} \rangle / \langle N_{\mathrm{trkl}}^{\mathrm{period}}(z) \rangle$. Here, $\langle N_{\mathrm{trkl}}^{\mathrm{ref}} \rangle$ is the maximum value in the average number of tracklets in a chosen reference period. In this case, one of the first periods in 2016 was chosen as reference to have the highest SPD acceptance. The $\langle N_{\mathrm{trkl}}^{\mathrm{period}}(z) \rangle$ term represents the mean number of tracklets for events with its vertex at a given value of z for the data-taking period under consideration. The Poisson probability distribution is employed to get the corrected number of tracklets for each event. In Fig. 6.2, the two-dimensional N_{trkl} against z_{vtx} distribution is shown before and after the correction procedure. For the forward multiplicity estimator, p_{V0M}, the percentiles are calculated run-by-run and therefore self-normalise the effect of ageing. In addition, no sizeable z_{vtx}-dependence is observed in the V0M amplitudes, hence no z_{vtx} correction was applied for this multiplicity estimator.

A second correction is applied to the midrapidity multiplicity estimator in order to reduce the effect of auto-correlations that could arise from the measurement of the charged-particle multiplicity density in the same rapidity region as the reconstruction of the charm hadrons. The tracklets coming from a D-meson or Λ_{c}^{+}-baryon decay track with a hit in the first and second ITS layer were subtracted from the total number of tracklets. For the $\Lambda_{\mathrm{c}}^{+} \rightarrow \mathrm{pK}_{\mathrm{S}}^{0}$ analysis, only

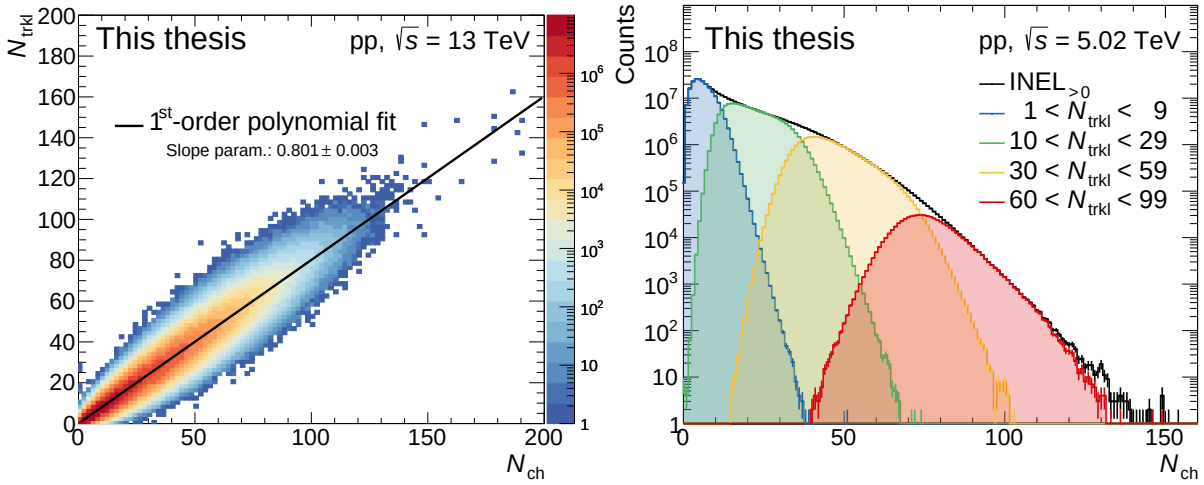


Figure 6.3: Left: The correlation between the physical primary N_{ch} with the detector-dependent N_{trkl} distribution within $|\eta < 1|$ (after the z_{vtx}-correction). The distribution is fitted with a first-order polynomial to show the almost linear correlation between both variables. Right: The distributions of the physical primary N_{ch} sliced in the different N_{trkl} multiplicity intervals.

the proton-candidate track is considered for this subtraction procedure. It is performed on a candidate basis, hence a maximum of three tracklets can be subtracted. Especially for the low-multiplicity event classes, this correction can have a sizeable effect. A possible remaining bias could be induced by the charged particles produced in the fragmentation of the charm quarks or by decays of excited charm states that are not subtracted from the number of tracklets.

6

6.2.2 Mean charged-particle multiplicity density

The N_{trkl} and p_{V0M} multiplicity estimators are not physical primary observables. They are detector-dependent and, for the midrapidity estimator, can change over time when the SPD ages. Therefore, a conversion is needed to a "physical value" that can be used outside the ALICE experiment as well. For this we use the mean charged-particle multiplicity density $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ of charged-primary particles N_{ch}, whose definition is given in Ref. [210]. For the p_{V0M} percentiles the conversion (and values) reported in Ref. [226] were used, while the conversion of the specific N_{trkl} intervals used in this analysis was performed separately.

The conversion is performed via the correlation between the physical primary N_{ch} and the detector-dependent N_{trkl} distributions within $|\eta| < 1$ for selected events. Monte Carlo simulations have shown that the number of tracklets is proportional to $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ within 2% [228], as also visible in the left panel of Fig. 6.3. The reported values of $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ in Tab. 6.1 are estimated as the mean of the N_{ch} distributions after applying the N_{trkl} interval selections, as presented in the right panel of Fig. 6.3. The uncertainty on the quoted $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ values is obtained as the RMS of the $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ distribution calculated for each period separately. Note that the distributions are reported for $|\eta| < 1$, while $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ is given per unit of pseudorapidity, hence the factor two difference between the values in Fig. 6.3 and Tab. 6.1.

There is a clear overlap in the N_{ch} distribution between the different event classes, hence a

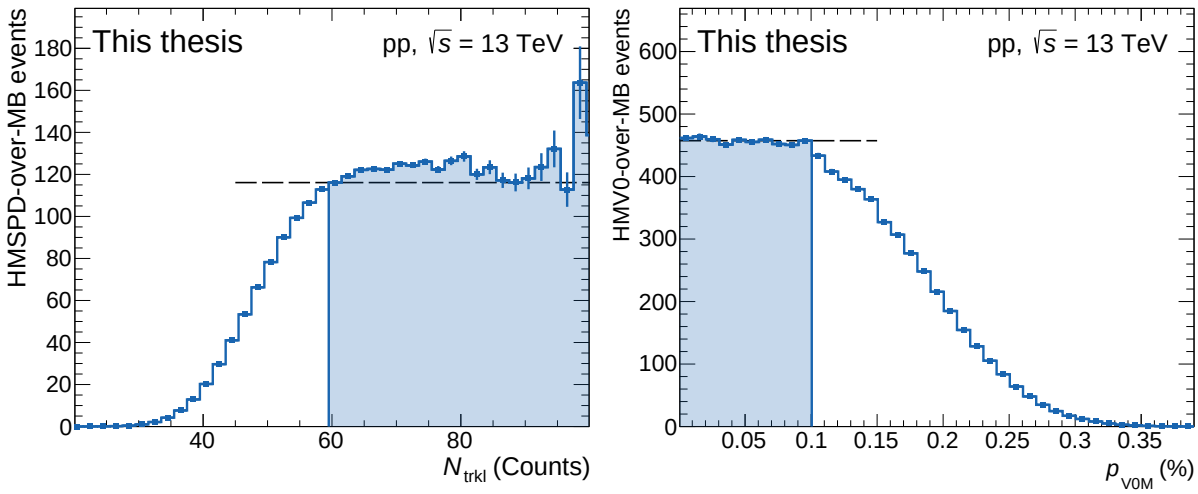


Figure 6.4: The trigger turn-on curves, defined as the ratio of HM-over-MB events as function of the relevant multiplicity estimator, for the HMSPD (left) and HMV0 (right) trigger. The filled area indicates the multiplicity interval for which the HM trigger is used. The dashed line is shown to highlight if the trigger was fully efficient in the considered range. The enhancement factor between the minimum-bias and HM triggered events, i.e. the ratio not saturating at one, is mainly due to the larger live-time of the HM triggers.

sharp selection on the reconstructed multiplicity does not correspond to a sharp selection on the true multiplicity estimator. It is therefore not trivial to estimate a minimum and maximum for the $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ distributions. Several strategies were considered, where in the end a first-order global polynomial fit to the profiles — weighted with the RMS of the distribution in each bin, instead of using the statistical uncertainties[2] — was used. The final minimum and maximum values of $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ in the four midrapidity event classes are: [0.7–5.6], [6.2–17.7], [18.3–35.8], and [36.4–60.0].

6

6.2.3 High-multiplicity trigger inefficiency

The high-multiplicity triggers HMSPD and HMV0 are fully efficient for $N_{\mathrm{trkl}} > 65$ and $p_{\mathrm{V0M}} < 0.1\%$, respectively, as can be seen via their trigger turn-on curves presented in Fig. 6.4. The trigger turn-on curves are defined as the ratio of the N_{trkl} or p_{V0M} distributions in minimum-bias and HM triggered events. No trigger correction is therefore needed for the HMV0 trigger in the studied p_{V0M} interval, while for the N_{trkl} event class we need to take into account the trigger inefficiency of the HMSPD trigger in the range $60 < N_{\mathrm{trkl}} < 65$.

The number of events as well as the charm-hadron candidate invariant-mass distributions were corrected for the HM trigger inefficiency by means of a data-driven re-weighting procedure. The N_{trkl}-dependent event weights were defined from the ratio of the measured distributions of the number of tracklets in the HM and minimum-bias trigger samples (i.e. the inverse of the ratios in Fig. 6.4). The effect of this correction on the raw yield per event was of about 2–4%, depending on the hadron species. Since the minimum-bias and high-multiplicity SPD triggers have different prescales, their ratio does not saturate at one. To extract the weights, the

[2] This was chosen to give more "weight" to the high-multiplicity region, where the statistics is limited.

ratio was thus first normalised to one in the high-multiplicity region. Different normalisation techniques were studied, which all gave the same corrected results within a few per mile.

6.3 Selection criteria

The $\mathrm{D_s^+}$ mesons were reconstructed via the $\mathrm{D_s^+} \to \phi\pi^+ \to \mathrm{K^+K^-}\pi^+$ decay channel, with branching ratio $2.24 \pm 0.08\%$ [16]. The optimisation of the selections on the topological and particle identification variables, as introduced in Sec. 4.2, is performed using a machine-learning approach (see Sec. 4.2.4). In particular, the Boosted Decision Trees (BDT) technique, using the toolkit from XGBoost [197], was employed. The training sample was assembled considering the background from the sidebands of the candidate invariant-mass distribution in data, and the prompt $\mathrm{D_s^+}$ signal candidates from MC simulations. Independent BDTs were trained for each p_{T} interval for the multiplicity-integrated event class. The normalised decay length ($L_{\mathrm{xy}}/\sigma(L_{\mathrm{xy}})$, the invariant-mass of the ϕ meson, the $\cos^3\theta'(\mathrm{K})$, and the $N_{\sigma}^{\mathrm{comb}}(\pi, \mathrm{K})$ variables for the opposite charge-sign decay particle are among the most important training variables in all p_{T} intervals. The optimal selections on the BDT outputs were tuned to maximise the expected statistical significance estimated by FONLL predictions [83, 87], as introduced in Sec. 4.2.4. The f_{prompt} parameter in Eq. 4.10 was set to 0.9 for all BDT selections, which is a typical value obtained for the measurement of $\mathrm{D_s^+}$ mesons in pp collisions. The BDT threshold values for the different p_{T} intervals vary between 0.96 and 0.98. In order to minimise sources of variations other than those relative to the multiplicity selections, the same BDTs and same selections were used in all the multiplicity intervals.

6.4 Raw-yield extraction

The $\mathrm{D_s^+}$ raw yields were obtained from binned maximum likelihood fits to the invariant-mass distribution of $\mathrm{D_s^+}$ candidates and charge conjugates, $M(\mathrm{KK}\pi)$, in six p_{T} intervals between $1 < p_{\mathrm{T}} < 24$ GeV/c for each multiplicity interval. A Gaussian function was used to describe the signal peak, while a second-order polynomial (for the lowest two p_{T} intervals) or exponential function was chosen to model the background shape. An additional Gaussian function was introduced to describe the signal of the decay $\mathrm{D^+} \to \mathrm{K^+K^-}\pi^+$ ($\mathrm{BR} = (9.68 \pm 0.18) \cdot 10^{-3}$ [16]), present on the left side of the $\mathrm{D_s^+}$ signal peak. To improve the fit stability, the width of the $\mathrm{D^+}$ peak in the $M(\mathrm{KK}\pi)$ invariant-mass spectrum was fixed to the width of the $\mathrm{D_s^+}$ signal peak extracted from data, multiplied by the ratio between the $\mathrm{D^+}$ and $\mathrm{D_s^+}$ peak widths obtained from simulations. Typical values of this ratio in simulations were found to be about 0.91. The widths of the Gaussian peaks used to describe the $\mathrm{D_s^+}$ signal component were left unconstrained in the fit procedure for the multiplicity-integrated sample. For the fits to the invariant-mass spectra for the multiplicity-dependent analyses, the widths were then fixed to the one extracted in data for the multiplicity-integrated case.

A few examples of the invariant-mass fits at different p_{T} and multiplicity intervals at midrapidity are reported in Fig. 6.5. The blue solid lines show the total fit functions, while the red dashed

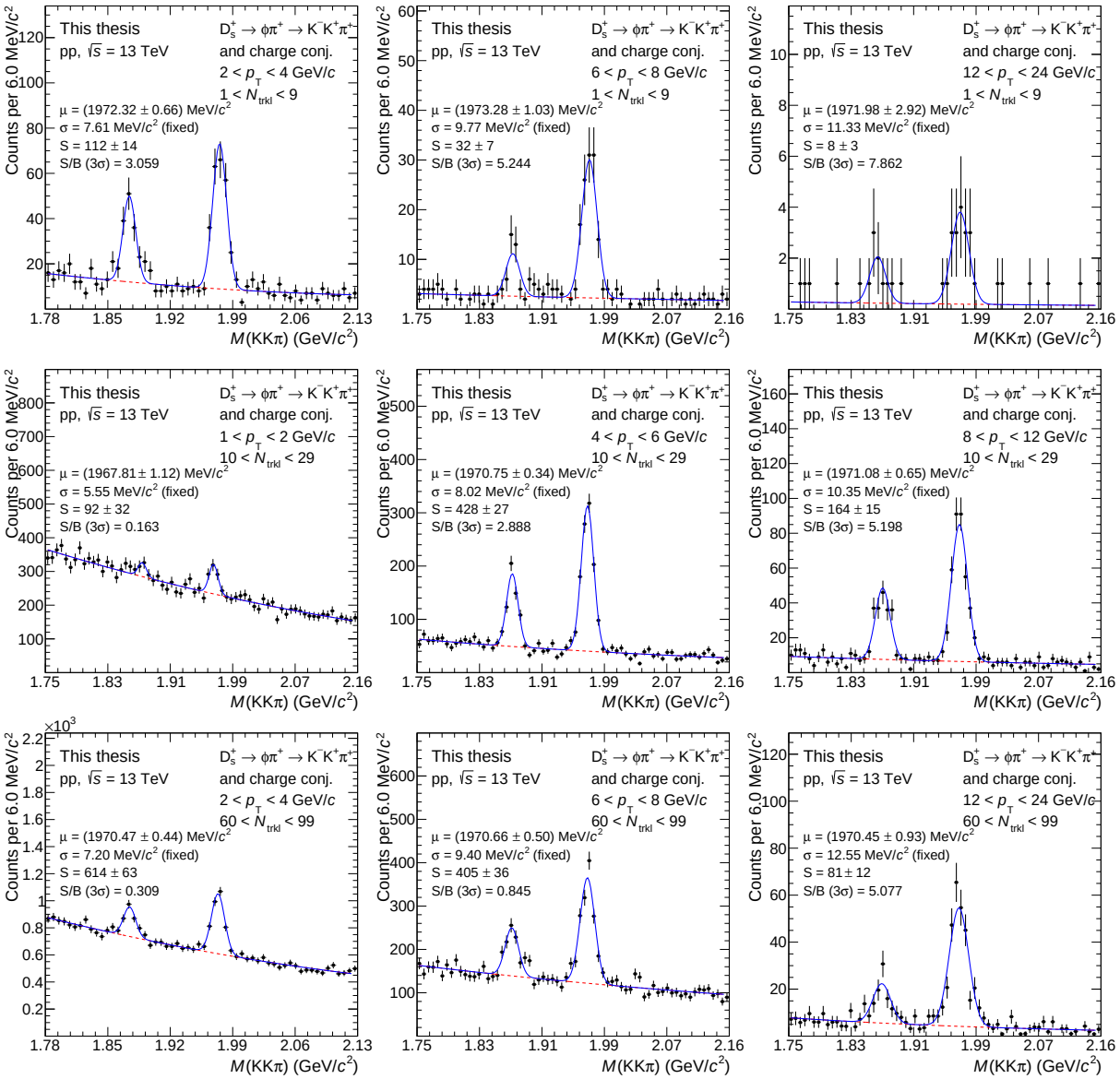


Figure 6.5: Invariant-mass distribution of D_s^+ candidates and charge conjugates in low (left), intermediate (middle), and high (right) p_T intervals, for low (top), intermediate (middle), and high (right) multiplicity intervals at midrapidity. The blue solid lines show the total fit function, while the red dashed lines are the combinatorial background terms. The values of the mean (μ) and the width (σ) of the signal peak are reported, together with the signal counts (S) and the signal-to-background ratio (S/B) in the mass interval $(\mu - 3\sigma, \mu + 3\sigma)$. The invariant-mass distributions in the highest multiplicity interval are corrected for the HMSPD trigger inefficiency, as explained in Sec. 6.2.3.

lines are the combinatorial-background terms. The values of the extracted raw yield, signal-to-background ratio, and significance (within 3σ) are presented in Fig. 6.6. A significance higher than three for the $1 < p_T < 2$ GeV/c interval could only be achieved in the INEL$_{>0}$ and $10 < N_{trkl} < 29$ event classes, hence this p_T interval was removed for the other multiplicity intervals. The Gaussian width of the fitted D_s^+ peaks vary between 5.5 to 12.55 MeV/c^2, increasing with p_T as expected from the p_T-resolution of the decay tracks[3]. The extracted Gaussian means

[3]The HMSPD trigger was not active in all periods, hence the Gaussian width of the $60 < N_{trkl} < 99$ multiplicity interval slightly differs with respect to the other multiplicity intervals.

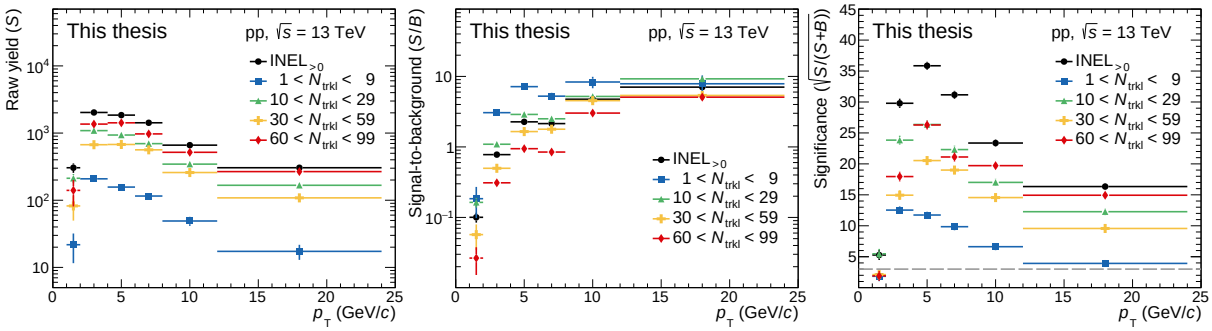


Figure 6.6: Summary of the performance of the D_s^+ invariant-mass fits in the different multiplicity intervals at midrapidity for pp collisions at $\sqrt{s} = 13$ TeV. The extracted values for the raw yield (left), the signal-to-background ratio (middle), and significance ($S/\sqrt{S+B}$) (right) are shown. The dashed line indicates a significance of three, which is ALICE's limit to make an observation.

vary around 1971 MeV/c^2, slightly above the PDG average of 1968.34 ± 0.07 MeV/c^2 [16]. The peak position is compatible with the values extracted from simulations within two standard deviations. This difference will be taken into account while evaluating the source of systematic uncertainty due to the raw-yield extraction.

The D_s^+ signal peak can be contaminated by the contribution of reflected signal, i.e. D_s^+ candidates with the wrong-mass hypothesis assigned to the pion and kaon tracks with the same charge sign. The invariant-mass distribution for reflected signal is centred around the D_s^+ mass, although with a much broader RMS. In case the final selected number of reflected candidates is substantial, this could bias the signal extraction. However, the particle identification and the $M_{\rm inv}(\rm KK) - M_{\rm PDG}(\phi(1020))$ selections reduce the amount of reflections significantly. After these selections, the reflected signal was observed to have a relative contribution of less than 0.1% to the D_s^+ extracted yield (within 5σ), as found in dedicated studies for the pp analysis at $\sqrt{s} = 5.02$ TeV [82, 211]. Since the same type of selections are applied for this analysis at $\sqrt{s} = 13$ TeV, no additional contribution for the K–π reflected signal was added to the invariant-mass fits.

There might be contamination from other D_s^+ decay channels, via other intermediate resonances, as well. Given the similar mass of the $\phi(1020)$, the $K^{*0}(892)$, and the $f_0(980)$ mesons, the $D_s^+ \to \overline{K}^{*0}(892)K^+ \to K^-\pi^+K^+$ (BR $= 2.58 \pm 0.08\%$) and $D_s^+ \to f_0(980)\pi^+ \to K^+K^-\pi^+$ (BR $= 1.14 \pm 0.31\%$) decay channels are expected to give the largest contribution [16]. The resonant decay channels via the $f_0(1370)$, $f_0(1710)$, and $K^{*0}(1430)$ mesons will be effectively rejected by the $M_{\rm inv}(\rm KK) - M_{\rm PDG}(\phi(1020))$ selection, and, in addition, have a branching ratio of one or two order of magnitude smaller. The contamination in the extracted D_s^+ signal of these two resonant decay channels was studied in detail for the pp measurement at $\sqrt{s} = 5.02$ TeV [82, 211]. The contamination of the $D_s^+ \to \overline{K}^{*0}(892)K^+ \to K^-\pi^+K^+$ decay was found to be less than 1% based on studies on the $A \cdot \epsilon$ factor. It is more difficult to estimate the contamination from the $D_s^+ \to f_0(980)\pi^+ \to K^+K^-\pi^+$ decay, because of the large uncertainty on the f_0 mass and width according to the PDG [16]. Based on a preliminary measurement of the ALICE Collaboration, that measured $M(f_0) \approx 974$ MeV/c^2 and $\Gamma(f_0) \approx 38$ MeV/c^2, a relatively small contamination

of 2.9% was found. Similarly as for the $\sqrt{s} = 5.02$ TeV measurement, no correction was applied for this analysis at $\sqrt{s} = 13$ TeV, since the same type of selections are applied.

6.5 Corrections

The acceptance-times-efficiency corrections were determined following the general procedure described in Sec. 4.3. To account for the multiplicity dependence of the efficiency, which is driven by the primary-vertex resolution improving with increasing multiplicity, the generated events were weighted on the basis of their charged-particle multiplicity in order to match the multiplicity distribution observed in data. The multiplicity weights are defined for each year as the ratio of the $N_{\rm trkl}$ distributions in data and MC, for selected events with a charm-hadron that has an invariant mass within a 20 MeV/c^2 interval around the PDG nominal value. For the multiplicity estimator at midrapidity, the $N_{\rm trkl}$ selection is applied in the simulations when calculating the efficiency, while this is not the case for the forward multiplicity estimator due to an uncalibrated $p_{\rm V0M}$ distribution in the simulations. To account for this, the multiplicity weights were instead extracted as function of $N_{\rm trkl}$, with the $p_{\rm V0M}$ multiplicity selection applied only for data. No weighting procedure is applied to merge the three years, since the relative number of events between the three years is similar for data and simulation. For the estimation of the acceptance factor, the ALICE simulations are directly used instead of the toy MC simulation procedure, because the underlying rapidity-shape of the $\rm D_s^+$ mesons is modified with respect to the predicted FONLL y-shape (especially for the lowest midrapidity multiplicity event class).

The acceptance-times-efficiency for prompt $\rm D_s^+$ mesons in the fiducial acceptance region for the different event classes, as well as their ratio with the $\rm INEL_{>0}$ event class, are presented in Fig. 6.7. The acceptance-times-efficiency increases from about 0.003 (0.006) at low $p_{\rm T}$ to about 0.20 at high $p_{\rm T}$ for prompt (feed-down) $\rm D_s^+$ mesons. The efficiency for feed-down $\rm D_s^+$ mesons is higher than that of prompt $\rm D_s^+$ mesons at low $p_{\rm T}$ because of the, on average, larger displacement from the primary vertex for beauty-hadron decays. The largest difference with respect to the efficiency computed in the $\rm INEL_{>0}$ class is observed at low $p_{\rm T}$ in $1 < p_{\rm T} < 2$ GeV/c, where the difference reaches 35%, while it decreases to about 5% in $12 < p_{\rm T} < 24$ GeV/c. Note that the HMSPD trigger was only active in 2018, hence in the ratio of the $60 < N_{\rm trkl} < 99$ interval to $\rm INEL_{>0}$, ageing effects of the detector are visible as well.

The fraction of prompt $\rm D_s^+$ was evaluated with the theory-driven method as introduced in Sec. 4.3. In particular, this procedure uses (i) the beauty-quark production cross section from FONLL calculations [83, 87], (ii) the $(A \times \epsilon)$ for feed-down $\rm D_s^+$ mesons, (iii) beauty-quark fragmentation fractions determined from LHCb data for $\rm b \to \Lambda_b^0$ [108] and from $\rm e^+e^-$ measurements for $\rm b \to B$ [201], and (iv) the decay kinematics from PYTHIA 8 simulations [95]. In contrary to the published analyses at $\sqrt{s} = 5.02$ TeV and $\sqrt{s} = 7$ TeV [82, 169], the $h_{\rm b} \to \rm D_s^+ + X$ kinematics is described by PYTHIA 8 simulations, instead of the EvtGen package [202], since in EvtGen the $\rm B_s^0 \to D_s^+ + X$ contribution is not taken into account[4]. A systematic uncertainty

[4]Unlike other B and D meson species, the $\rm B_s^0$ meson decays with similar abundances in $\rm D_s^+$ and $\rm D_s^-$.

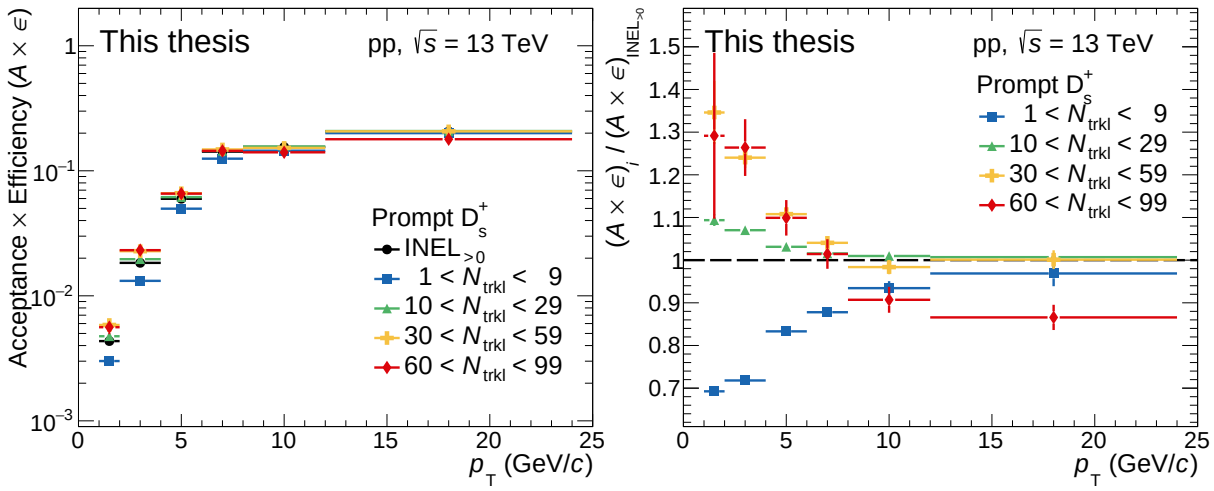


Figure 6.7: Prompt acceptance-times-efficiency as function of p_T for the different multiplicity event classes at midrapidity (left) and their ratio with the INEL$_{>0}$ event class (right). Note that the HMSPD trigger was only active in 2018, hence the ratio to INEL$_{>0}$ for the $60 < N_{trkl} < 99$ interval show ageing effects of the detector as well.

was estimated by varying the FONLL parameters [83], including as well the uncertainties on the CTEQ6.6 PDFs [209]. The f_{prompt} fraction was assumed to be independent of the event multiplicity and therefore computed for the minimum-bias event class. This assumption is justified by the expected weak dependence of the feed-down fraction with multiplicity [228], predicted also by PYTHIA 8 simulations, and the relatively small variations of the feed-down efficiencies observed in the different multiplicity intervals. The values of f_{prompt} range from 0.81 at low p_T to 0.87 in $6 < p_T < 8$ GeV/c.

6

6.6 Systematic uncertainties

Systematic uncertainties on the D_s^+-meson corrected yields in different multiplicity intervals were estimated considering the following sources, starting with the ones similar for the minimum-bias analyses (like presented in Chapter 5): the extraction of the raw yield from the invariant-mass distributions (δ_1); the track-reconstruction efficiency (δ_2); the D_s^+-meson selection efficiency (δ_3); the PID selection efficiency (δ_4); the p_T shape of generated D_s^+ mesons in the simulation (δ_5); the subtraction of the feed-down from beauty-hadron decays (δ_6); and the uncertainties on the branching ratios (δ_7). Since the corrected yield per event is reported, the uncertainty on the integrated luminosity was not considered as systematic source. Instead the uncertainty on the ϵ_{INEL} values (see Tab. 6.1) were assigned (δ_8).

In addition to these general minimum-bias sources of systematic uncertainty, new sources related to the multiplicity dependence of the analysis are addressed: the multiplicity and z_{vtx} shape of the simulated events ($\delta_{9,10}$); the HM trigger inefficiency correction (δ_{11}); the assumed multiplicity independence of f_{prompt} (δ_{12}); the N_{trkl} interval limits in the simulation (δ_{13}); and the statistical uncertainty on the acceptance-times-efficiency values (δ_{14})[5].

[5]For minimum-bias analyses, the statistical uncertainty on the acceptance-times-efficiency values is negligible with respect to other sources, hence it was not considered in Chapter 5.

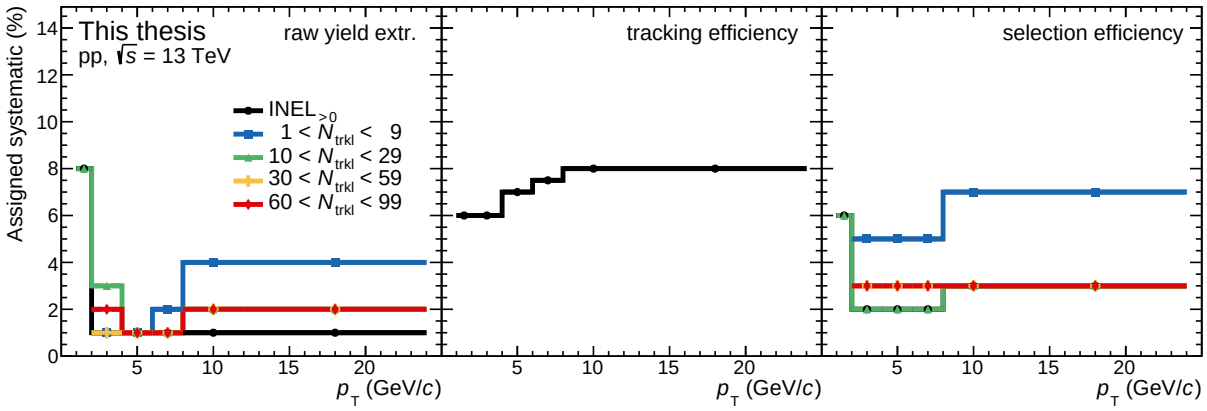


Figure 6.8: Summary of the assigned systematic uncertainties for the raw-yield extraction (left), the tracking efficiency (middle), and the selection efficiency (right). See text for more details.

In the subsections below, source δ_1–δ_6 will be briefly discussed, while source δ_9–δ_{14} will be discussed in more detail. In Sec. 5.5, more details on the procedures of the systematic sources which are in common with the minimum-bias analyses can be found.

6.6.1 Sources in common with minimum-bias analyses

The systematic uncertainty on the raw-yield extraction was evaluated in each combination of the studied p_{T} and multiplicity intervals by repeating the fit to the invariant-mass spectra varying the fit range and the background fit function via the so-called multi-trial approach. The systematic uncertainty, assigned based on the RMS of the distribution and shift of the mean with respect to the central value, ranges between 1% and 8% depending on the p_{T} and multiplicity interval, as presented in the left panel of Fig. 6.8.

Systematic effects on the track-reconstruction efficiency were estimated by varying the TPC track-quality selection criteria and by comparing the prolongation probability of the TPC tracks to the ITS hits in data and simulation. The final assigned systematic uncertainty follows from the quadratic sum of both per-track contributions, propagated to the $\mathrm{D_s^+}$ mesons via the kinematics of the decay tracks. It ranges from 6% to 8%, increasing with increasing p_{T} (see middle panel of Fig. 6.8). This systematic source is assumed to be multiplicity independent, hence the values for the $\mathrm{INEL_{>0}}$ event class are assigned as well to the different multiplicity intervals.

The systematic uncertainty on the selection efficiency originates from imperfections in the simulation of the decay kinematics and topology and from the resolutions and alignments of detectors in the simulation. It was estimated by comparing the corrected yields obtained by repeating the analysis with different selections on the BDT output, resulting in a significant modification of the efficiencies, raw yield, and background values. This procedure, illustrated in Fig. 6.9, covers both the PID and topological selection efficiency. The variations are performed separately for the different multiplicity and p_{T} intervals. The assigned systematic uncertainties are largest at low p_{T} — where the selection criteria are strict — reaching 6% for the $\mathrm{D_s^+}$ analysis (see right panel of Fig. 6.8). Possible systematic effects due to the loose PID selection applied prior to the

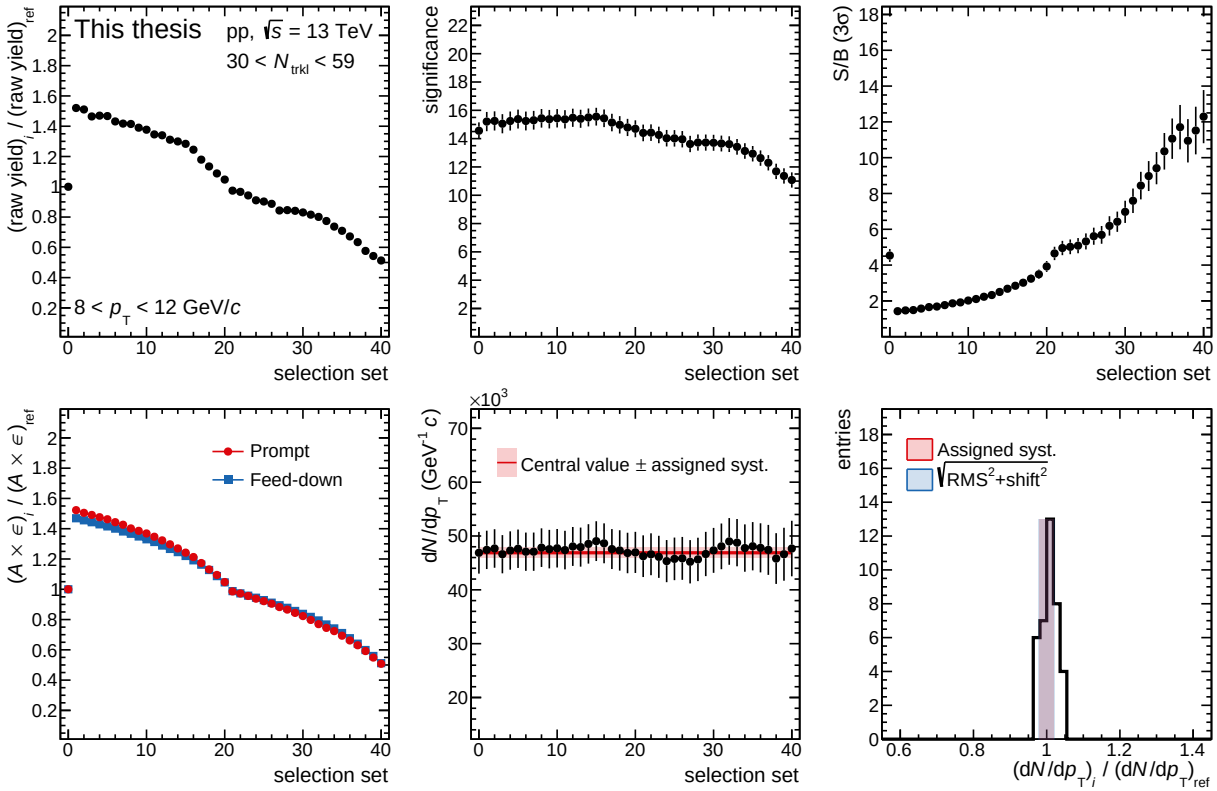


Figure 6.9: Summary on the estimation of the systematic uncertainty on the selection efficiency for $\mathrm{D_s^+}$ mesons in pp collisions at $\sqrt{s} = 13$ TeV for $8 < p_{\mathrm{T}} < 12$ GeV/c in the $30 < N_{\mathrm{trkl}} < 59$ multiplicity interval. The variation of the extracted raw yield (top-left), the statistical significance (top-middle), the signal-to-background ratio (top-right), the variation of the acceptance-times-efficiency (bottom-left), and the corrected yield per event (bottom-middle) as function of the selection set are presented. The ratio of the $\mathrm{d}N/\mathrm{d}p_{\mathrm{T}}$ per event for each selection set with the central one is reported in the bottom-right panel.

machine-learning one, were investigated comparing pion and kaon PID selection efficiencies in data and simulation. The differences were found to be negligible.

The possible differences between the real and simulated $\mathrm{D_s^+}$-meson p_{T} spectra result in a further source of systematic uncertainty, which was evaluated by reweighting the p_{T} shape from the PYTHIA 8 generator [95] to match the one from FONLL calculations [83, 87]. The difference between both predictions (i.e. the weights) are shown in the left panel of Fig. 6.10. This contribution, found to be multiplicity independent, ranges from 1% to 6% for $p_{\mathrm{T}} < 6$ GeV/c while it is negligible at higher p_{T}. These values, and the multiplicity independence, are confirmed by a reweighting procedure exploiting multiplicity-dependent weights, defined as the ratios between Tsallis-Lévy fits [229] to the measured production yields (as will be shown later in Sec. 6.7) and the generated p_{T} shapes from PYTHIA.

The systematic uncertainty on the prompt fraction was estimated by varying the FONLL parameters (in particular the b-quark mass and the renormalisation and factorisation scales) in the calculation of the feed-down $\mathrm{D_s^+}$ meson cross section. The assigned uncertainty values for the $\mathrm{D_s^+}$ meson range from $^{+12.4}_{-12.3}$% at low p_{T} to $^{+4.5}_{-3.6}$% at high p_{T}, as presented in the right panel of Fig. 6.10.

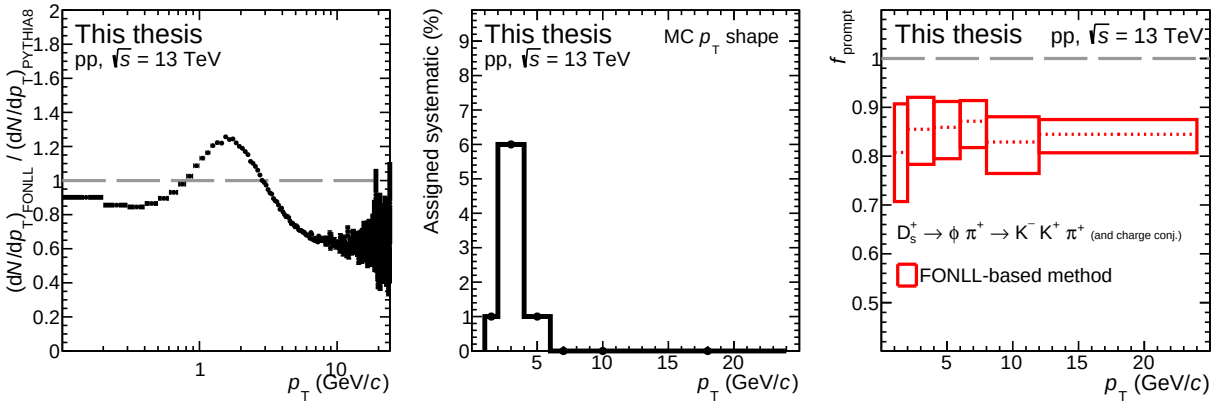


Figure 6.10: Left: Ratio between the normalised $\mathrm{d}N/\mathrm{d}p_{\mathrm{T}}$ spectra of FONLL [83, 87] calculations and PYTHIA 8 [95] predictions for prompt $\mathrm{D_s^+}$ mesons at $\sqrt{s}$ = 13 TeV TeV. Middle: Summary of the assigned systematic uncertainty due to the generated p_{T} shape. Right: The estimated prompt fraction of $\mathrm{D_s^+}$ mesons from the FONLL-based approach (dashed line) and the assigned systematic uncertainty (red boxes).

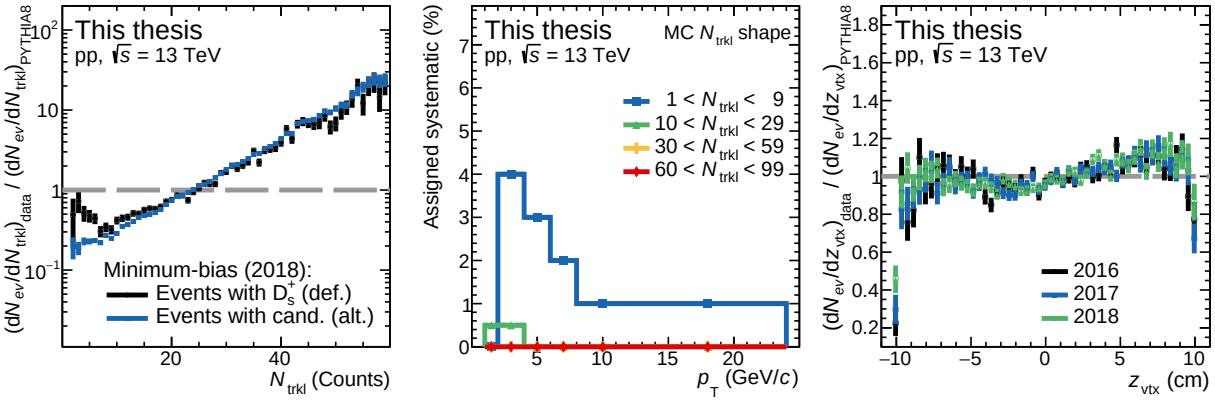


Figure 6.11: Left: Ratio between the normalised N_{trkl} distributions in data and simulations for both the default and systematic-alternative case for the midrapidity multiplicity estimator. Middle: Summary of the assigned systematic uncertainty due to the generated multiplicity shape in the different event classes. Right: Comparison of the z_{vtx} weights for the three years for the $\mathrm{INEL}_{>0}$ event class.

6.6.2 Sources related to the multiplicity dependence

6.6.2.1 Generated multiplicity and z_{vtx} shape

Similarly as the generated $\mathrm{D_s^+}$-meson p_{T} shape, the simulated event multiplicity could introduce a bias in case of a non-realistic distribution. As discussed in Sec. 6.5, the N_{trkl} distributions in the simulations are tuned to reproduce the measured distributions in data to account for the multiplicity dependence of the efficiencies. The systematic uncertainty due to this procedure was evaluated by changing the weight functions used to reproduce the measured charged-particle multiplicity in the simulation. For the central value of the efficiencies, the weights were obtained from the N_{trkl} distributions of events with a selected $\mathrm{D_s^+}$ candidate within a 20 MeV/c^2 invariant-mass window, while for the systematic variation this invariant-mass requirement was removed. The comparison of both weight distributions are shown in the left panel of Fig. 6.11. This alternative weighting distribution leads to a maximum deviation of about 4% on the $\mathrm{d}N/\mathrm{d}p_{\mathrm{T}}$

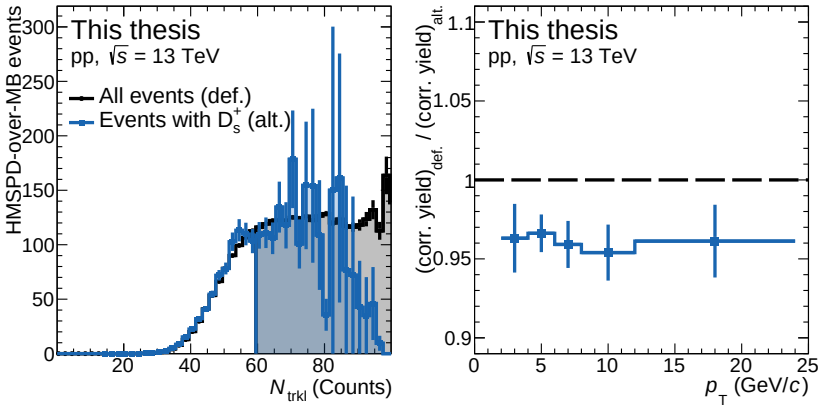


Figure 6.12: Left: The HMSPD trigger turn-on curve, defined as the ratio of HM-over-MB events as function of $N_{\rm trkl}$, for all events (default) and events with a selected $\rm D_s^+$ candidate within a 20 MeV/c^2 invariant-mass window (alternative). The filled area indicates the multiplicity interval for which the HM trigger is used. Right: Ratio between the $\mathrm{d}N/\mathrm{d}p_{\rm T}$ distributions obtained with the alternative and default HMSPD trigger inefficiency weighting procedure.

observed at low $p_{\rm T}$ and low multiplicity (middle panel of Fig. 6.11), which is assigned as systematic uncertainty.

A slight dependence of the efficiencies on the primary vertex position along the beam axis, $z_{\rm vtx}$, is also observed. Hence, a further data-driven reweighing procedure was performed to investigate possible biases induced by differences between the $z_{\rm vtx}$ distribution in simulations and in data. The weights are defined as the ratio of the $z_{\rm vtx}$ distributions for events with a selected $\rm D_s^+$ candidate within a 20 MeV/c^2 invariant-mass window. The $z_{\rm vtx}$ weights for the three different years are shown in the right panel of Fig. 6.11. A $p_{\rm T}$-dependent systematic uncertainty is estimated, resulting in a small contribution of about 0.5% for $p_{\rm T} < 4$ GeV/c, and negligible elsewhere.

6.6.2.2 Trigger inefficiency correction

As explained in Sec. 6.2.3, to correct for the HMSPD trigger inefficiency in the $60 < N_{\rm trkl} < 99$ multiplicity interval, a data-driven reweighting procedure was used to correct the number of events and $\rm D_s^+$-candidate invariant-mass distributions. The central weights are defined as the inverse of the ratio (normalised to one in the high-multiplicity region) between the measured $N_{\rm trkl}$ distributions in the HM and minimum-bias trigger samples for selected events. As a systematic variation, the weights were calculated adding the requirement of a selected $\rm D_s^+$-candidate within a 20 MeV/c^2 invariant-mass window. The trigger turn-on curve for these events, as well as the difference between the final $\mathrm{d}N/\mathrm{d}p_{\rm T}$, are presented in Fig. 6.12. The difference on the corrected yield is about 4%, which is thus assigned as systematic uncertainty.

To validate the procedure used to correct for the trigger inefficiency, the analysis in $60 < N_{\rm trkl} < 99$ is performed with the minimum-bias triggered data as well. The $\mathrm{d}N/\mathrm{d}p_{\rm T}$ distribution is compatible within uncertainties with the central analysis using the HMSPD trigger including correction for the trigger inefficiency. The minimum-bias sample is, however, affected by much larger uncertainties.

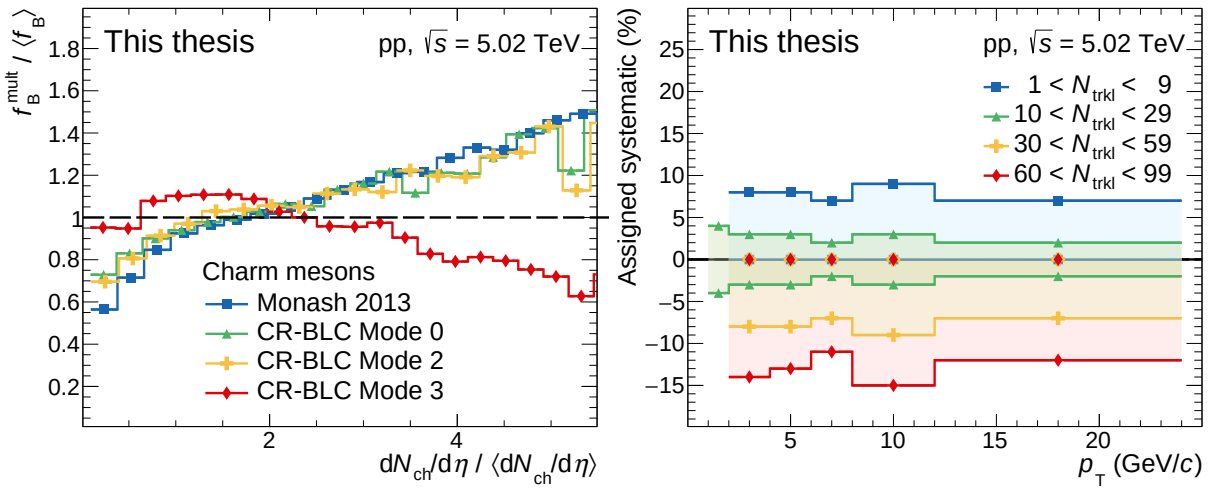


Figure 6.13: Left: Ratio of the $f_{\mathrm{B}}^{\mathrm{mult}}/\langle f_{\mathrm{B}} \rangle$ distributions for charm-hadrons as function of $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta/\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ for various PYTHIA 8 tunes (no uncertainties are shown). Right: Summary of the assigned systematic uncertainty due to the assumption of a multiplicity-independent prompt fraction in the different event classes.

6.6.2.3 Feed-down subtraction

For the calculation of the prompt fraction with the FONLL-based approach, we made the assumption that the relative contribution of beauty hadron decays to the charm-hadron yields in the different multiplicity intervals is constant and equal to the multiplicity-integrated class. There could, however, be a dependency with the charged-particle multiplicity density. A systematic uncertainty on this assumption, in addition to the FONLL variations, is thus assigned for the feed-down subtraction.

The dependency of the prompt fraction with event multiplicity was studied with different PYTHIA 8 tunes, in particular with Monash [151] and the colour-reconnection beyond leading colour approximation modes [110]. In the left panel of Fig. 6.13, the relative contribution of feed-down charm hadrons in multiplicity intervals with respect to the multiplicity-integrated class is presented[6]. In particular, the $f_{\mathrm{B}}^{\mathrm{mult}}/\langle f_{\mathrm{B}} \rangle$ distributions, with $f_{\mathrm{B}} = 1 - f_{\mathrm{prompt}}$, were studied as function of $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta/\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$, which show a global increasing trend from about 0.7 to 1.5 from the lowest to the highest multiplicity event class. From these distributions, multiplicity-dependent variations were defined and applied to the estimated f_{B} fractions in data. The obtained range of f_{B} values are converted to an interval in f_{prompt}, which is assigned as systematic uncertainty. The resulting values depend on the p_{T} interval and the multiplicity class, and are summarised in the right panel of Fig. 6.13.

The uncertainties coming from the FONLL-based approach for the estimation of the f_{prompt} fraction are taken into account in this procedure, hence the systematic variations are applied on a lower, central, and upper limit of the f_{B} fractions in each p_{T} interval. Only in case of the $(\mathrm{d}N/\mathrm{d}p_{\mathrm{T}})_{\mathrm{mult}}/(\mathrm{d}N/\mathrm{d}p_{\mathrm{T}})_{\mathrm{INEL}>0}$ ratio, the uncertainties coming from the FONLL variation are not taken into account. It is thus assumed that the values of the pQCD scales and quark mass

[6]The study was performed separately for each charm-hadron specie and found to be compatible within uncertainties. We therefore report the "charm-hadron integrated" case to have a better statistical precision.

do not depend on multiplicity. The assigned systematic uncertainty in this case is therefore smaller, varying for intermediate p_{T} between $^{+5}_{-0}$% at low and $^{+0}_{-8}$% at high multiplicity.

6.6.2.4 Other sources of systematic uncertainties

Systematic effects due to the dependence of the efficiency on the N_{trkl} interval limits were studied as well, since the decay tracks of the reconstructed candidates were removed from the multiplicity estimation in data, but not in the computation of the efficiency in MC, for practical reasons. A systematic uncertainty was evaluated by comparing the efficiency computed in a N_{trkl} interval shifted by one or two units (for two- or three-body decays, respectively) with the one in the default intervals. A systematic uncertainty was observed for the $1 < N_{\mathrm{trkl}} < 9$ interval, ranging from 3% to 7% depending on p_{T}. No systematic uncertainty is assigned for the other multiplicity classes

Because of the limited statistics of the simulated samples, especially after the multiplicity interval selections, the relative statistical uncertainty on the selection efficiency is assigned as systematic uncertainty. It depends on the multiplicity intervals, reaching 4% for the highest and 0.5–2% for the other multiplicity intervals.

6.6.3 Summary

The contributions of these different sources of uncertainties (δ_1–δ_{14}) were summed in quadrature to obtain the total systematic uncertainty in each p_{T} interval, which varies from $^{+10\%}_{-10\%}{}^{+17\%}_{-15\%}$ in the lowest to $^{+11\%}_{-15\%}{}^{+15\%}_{-19\%}$ in the highest multiplicity interval. The systematic source due to the assumed multiplicity independence of the prompt fraction is dominant, and (due to its dependence on f_{prompt}) the reason why the total uncertainties differ between the D^0, $\mathrm{D}_{\mathrm{s}}^+$, and Λ_{c}^+ analyses.

In the next section, we will see three different type of observables, for which the strategy for the propagation of the statistical (δ_{stat}) and systematic uncertainty sources is now briefly explained. In case of the ratio of the $\mathrm{d}N/\mathrm{d}p_{\mathrm{T}}$ (per event) between the multiplicity and the $\mathrm{INEL}_{>0}$ event classes, the multiplicity-dependent systematic sources were considered as uncorrelated ($\delta_{8,\,9,\,11-13}$) and the contributions of the tracking and PID efficiency, the shape of MC p_{T} spectra and z_{vtx} distribution, the beauty feed-down, and the branching ratio as correlated ($\delta_{2,\,4-7,\,10}$). The statistical uncertainties and the systematic uncertainties related to the selection efficiency and to the raw-yield extraction were considered partially correlated with respect to the measurement performed in the $\mathrm{INEL}_{>0}$ class ($\delta_{\mathrm{stat},\,1,\,3,\,14}$). When considering charm-hadron ratios, the total systematic uncertainties are computed treating the contributions due to the raw-yield extraction, the selection efficiency, the shape of the MC p_{T} spectra, the z_{vtx} distribution, the statistical uncertainties, and the branching ratio as uncorrelated ($\delta_{\mathrm{stat},\,1,\,3,\,5,\,7,\,10,\,14}$). The systematic effect deriving from the variation of the multiplicity-interval limits was propagated as partially correlated (δ_{13}), while the other systematic uncertainties were assumed to be fully correlated. Finally, in case of double charm-hadron ratios between different multiplicity event classes, the

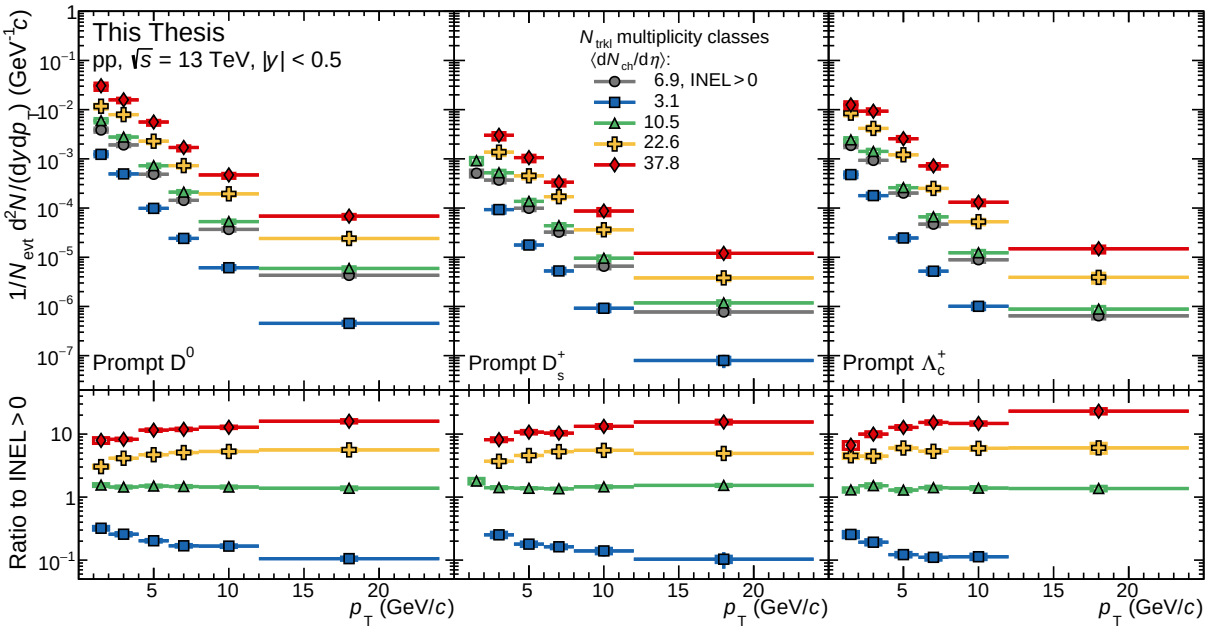


Figure 6.14: Transverse-momentum spectra of D^0, D_s^+, and Λ_c^+ hadrons measured in pp collisions at $\sqrt{s} = 13$ TeV for different multiplicity event classes at midrapidity. The corresponding ratios to $\mathrm{INEL}_{>0}$ are shown in the bottom panels.

raw-yield extraction, the selection efficiency, the shape of the MC p_T spectra, the z_vtx distribution, and the statistical uncertainties are considered as uncorrelated ($\delta_{\mathrm{stat},\,1,\,3,\,5,\,10,\,14}$), the multiplicity-interval limits as partially correlated (δ_{13}), while the other sources cancel out in the double ratio.

6.7 Results

The p_T-differential spectra of D_s^+ mesons, measured in $|y| < 0.5$, are shown in Fig. 6.14 for minimum bias ($\mathrm{INEL}_{>0}$) and the four multiplicity classes selected using the N_trkl estimator at midrapidity. The bottom panels present the ratios to the $\mathrm{INEL}_{>0}$ class. The p_T-differential spectra of D^0 and Λ_c^+ hadrons are presented as well, where the p_T-differential corrected yield of Λ_c^+ was obtained averaging the results from the two decay channels $\Lambda_\mathrm{c}^+ \rightarrow \mathrm{pK}^-\pi^+$ and $\Lambda_\mathrm{c}^+ \rightarrow \mathrm{pK}_\mathrm{S}^0$ to achieve a more precise measurement of the Λ_c^+ production. The correlation between the statistical and systematic uncertainties is taken into account following the strategy explained in Ref. [102]. The multiplicity-dependent systematic sources were, in addition, considered as correlated between the two decay channels. The measured p_T spectra of the three particles increase from the lowest to the highest multiplicity class, and their ratios to $\mathrm{INEL}_{>0}$ increase (decrease) with increasing p_T for the highest (lowest) multiplicity class, suggesting a plateau towards $p_\mathrm{T} > 10$ GeV/c. Such a plateau, i.e. the slopes of the particle spectra becoming independent of the multiplicity class, is indeed expected from pQCD calculations [230].

In order to investigate potential differences in the $\langle \mathrm{d}N_\mathrm{ch}/\mathrm{d}\eta \rangle$ dependence of the D^0-meson production with respect to the D_s^+ meson and Λ_c^+ baryon, the $\mathrm{D}_\mathrm{s}^+/\mathrm{D}^0$ and $\Lambda_\mathrm{c}^+/\mathrm{D}^0$ yield ratios are compared in different event multiplicity classes in Fig. 6.15, considering both the forward and

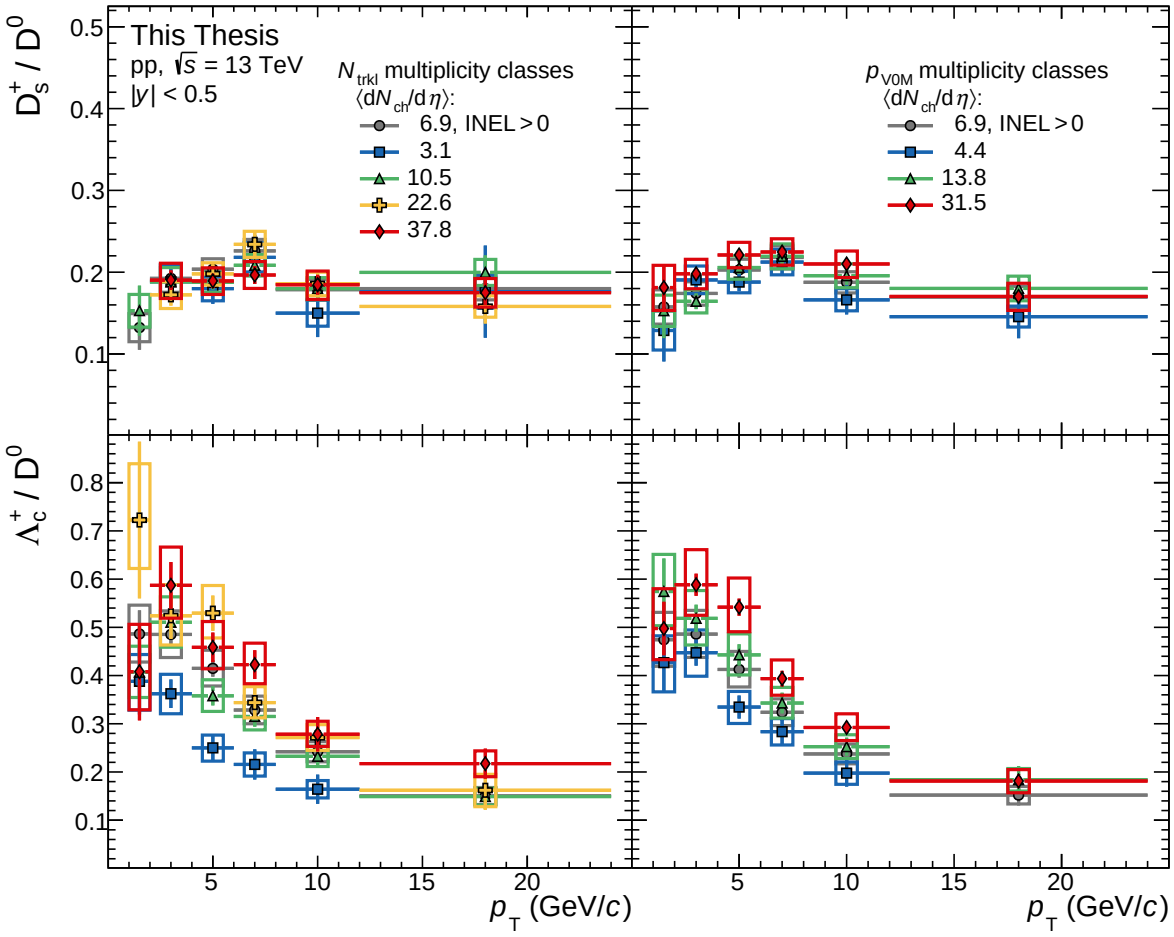


Figure 6.15: The $\mathrm{D_s^+/D^0}$ (top) and $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ (bottom) ratios measured in pp collisions at $\sqrt{s} = 13$ TeV for different multiplicity event classes at mid- (left) and forward (right) rapidity.

midrapidity estimators. Within the current experimental uncertainties, the $\mathrm{D_s^+/D^0}$ ratios are independent of p_T in the measured p_T range and compatible with the average of the p_T-integrated measurements performed at $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ colliders: 0.17 ± 0.03 [146, 201]. A dependence of these ratios with multiplicity, as seen for the ratio of (multi-)strange hadrons to $\pi^\pm$ [8, 225], is not observed within the uncertainties.

The $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ ratios show, on the contrary, an evident dependence on multiplicity, and a hierarchy is observed going from the lowest to the highest multiplicity intervals, for both the N_trkl and p_V0M estimators. In the lowest multiplicity class the ratio is still higher than the average of corresponding measurements performed in $\mathrm{e^+e^-}$ collisions at LEP, which was found to be $0.113 \pm 0.013(\mathrm{stat}) \pm 0.006(\mathrm{syst})$ [102, 201]. In order to estimate a significance level for the increase with multiplicity, the $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ ratio between the highest multiplicity (HM) over the lowest multiplicity (LM) event class at midrapidity was computed. The probability to observe the measured $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ double-ratio (DR) larger than one — reported for both the $\mathrm{D_s^+/D^0}$ and $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ in Fig. 6.16 — corresponds to a significance of 5.3σ in $1 < p_\mathrm{T} < 12$ GeV/c. Here, we considered as null hypothesis DR = 1, we took into account the statistical and systematic uncertainties, and, with the aim of investigating the least favourable case, the measured values

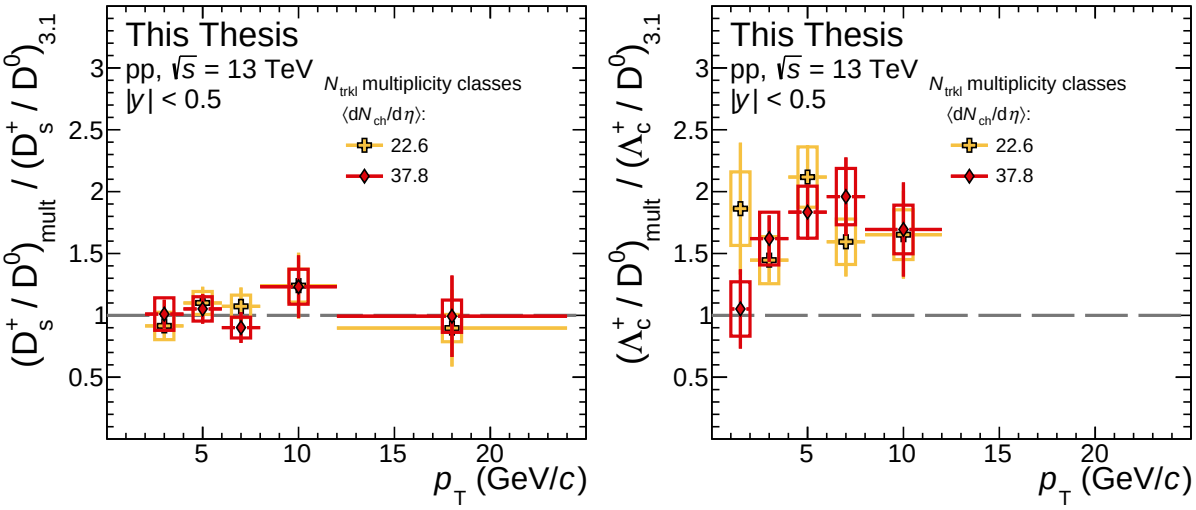


Figure 6.16: The $\mathrm{D_s^+/D^0}$ (left) and $\Lambda_\mathrm{c}^+/\mathrm{D^0}$ (right) double-ratios between the lowest and two highest multiplicity event classes at midrapidity for pp collisions at $\sqrt{s} = 13$ TeV.

in all p_T intervals were shifted down by one standard deviation, by considering the sources of systematic uncertainties correlated with p_T that do not cancel out in the double ratio ($\delta_{3,5}$).

The $\mathrm{D_s^+/D^0}$ ratios are shown in Fig. 6.17 as a function of $\langle \mathrm{d}N_\mathrm{ch}/\mathrm{d}\eta \rangle$ in five different p_T intervals from 2 to 24 GeV/c. The measured ratios in the different midrapidity multiplicity event classes in pp collisions at $\sqrt{s}$ = 13 TeV are compared to the ratio in (forward) multiplicity event classes in p–Pb [130] and Pb–Pb [144, 163] collisions at $\sqrt{s_\mathrm{NN}} = 5.02$ TeV. As already observed in Ref. [144], there is a hint of an enhancement of the $\mathrm{D_s^+/D^0}$ ratio in Pb–Pb collisions with respect to minimum-bias pp and p–Pb collisions. With the current uncertainties it is, however, not possible to say if this enhancement is a continuous trend from small to large systems as for light-flavour strange hadrons [8, 157, 225], or if there is a discontinuity between pp/p–Pb and Pb–Pb multiplicities[7].

6.7.1 Theory comparison

The measured charm-hadron ratios for the lowest and highest multiplicity class at midrapidity are compared to model predictions from MC generators in Fig. 6.18. The simulations with the PYTHIA event generator were performed with the Monash 2013 tune [151], with and without colour-reconnection beyond leading colour approximation (CR-BLC) [110] (see Sec. 2.4). For the latter, three modes are suggested by the authors, applying different constraints on the allowed reconnections among colour sources, in particular concerning the causality connection among strings involved in a reconnection, and time dilation caused by relative boosts of the strings. The simulations are shown in intervals of primary particle multiplicities selected at midrapidity, evaluated by studying the correlation between N_trkl intervals and N_ch values in simulation. The

[7] Keep in mind that a trend in a given p_T interval could also be due to a modification of the p_T shape, e.g. due to radial flow.

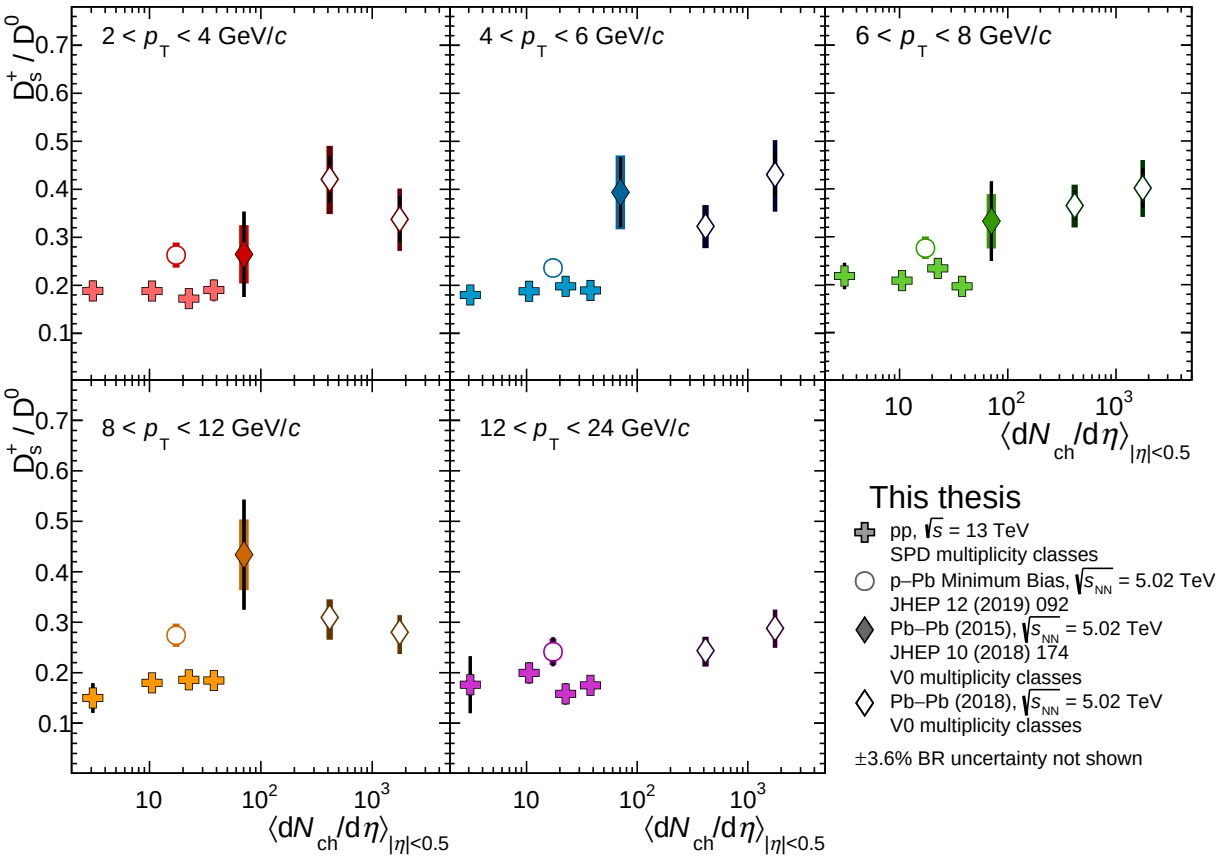


Figure 6.17: The D_s^+/D^0 ratios, as a function of primary charged particles per unit of pseudorapidity, in pp at $\sqrt{s}$ = 13 TeV, p–Pb at $\sqrt{s}$ = 5.02 TeV [130], and Pb–Pb at $\sqrt{s_{NN}}$ = 5.02 TeV [144, 163] in five different p_T intervals from 2 to 24 GeV/c.

6

estimated intervals are $1 \leq N_{ch} \leq 12$ and $N_{ch} > 75$ for the lowest and highest multiplicity event classes, respectively.

The measured D_s^+/D^0 ratio at low and high multiplicity are compatible with PYTHIA predictions with the Monash and CR-BLC tunes. As we already have seen when introducing the Λ_c^+ puzzle, the Monash tune does not reproduce the Λ_c^+/D^0 ratio in minimum-bias pp collisions, neither for multiplicity event classes. In addition, it does not predict a multiplicity dependence. Meanwhile, the CR-BLC tunes describe the Λ_c^+/D^0 decreasing trend with p_T, and are closer to the overall magnitude, as also observed for minimum-bias pp collisions at $\sqrt{s}$ = 5.02 and 13 TeV [102, 231]. The CR-BLC tunes show a clear dependence with multiplicity, qualitatively reproducing the trend observed in data.

The measurements in Fig. 6.18 are also compared with a canonical ensemble statistical hadronisation (CE-SH) model prediction [232], where the authors generalise their grand-canonical ensemble statistical hadronisation (GCE-SH) investigation of charm-hadron production [111]. In the usual implementation of a grand-canonical ensemble, quantum charges like electric charge, baryon number, and charm number are conserved on average and the hadronisation temperature is usually assumed to be a universal parameter for heavy-ion collisions, hence the hadron yield

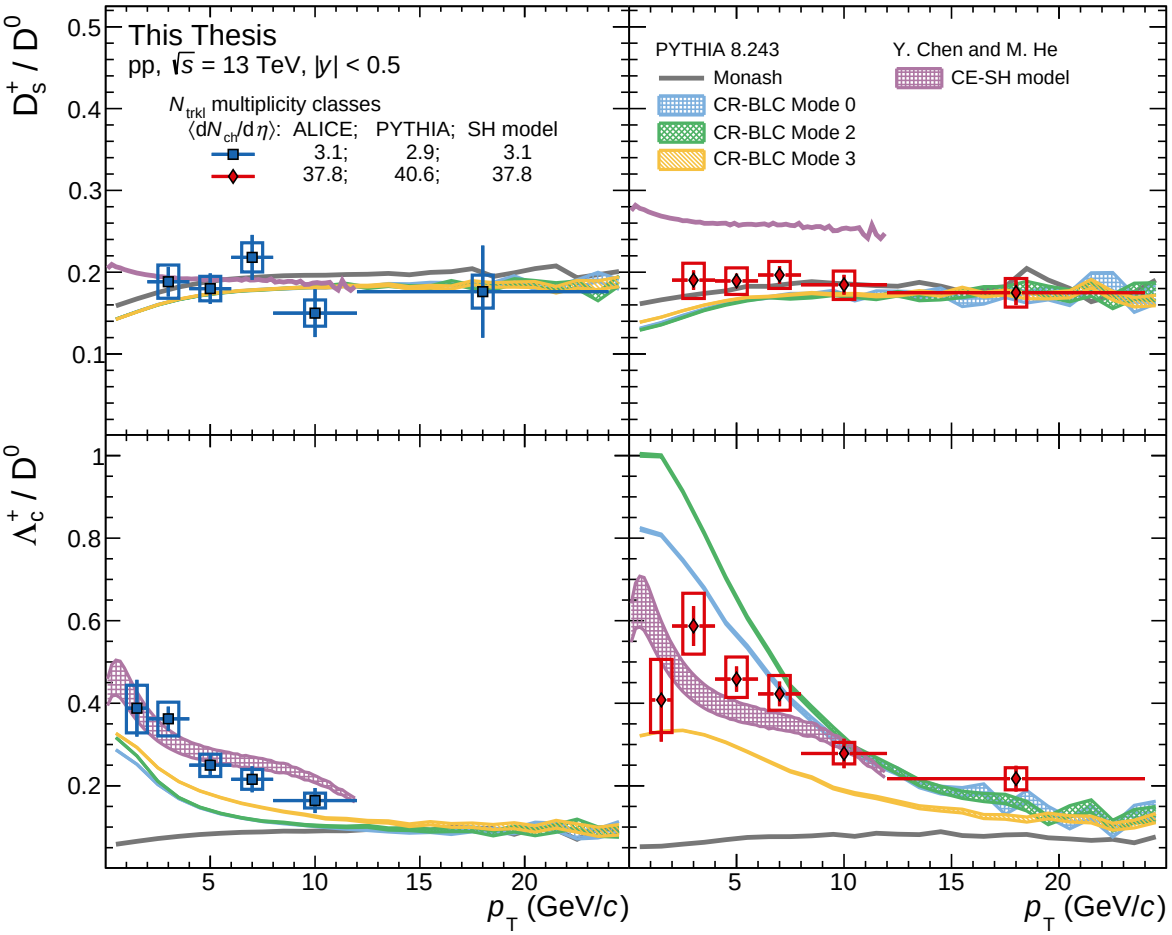


Figure 6.18: The D_s^+/D^0 (top) and Λ_c^+/D^0 (bottom) ratios measured in pp collisions at $\sqrt{s} = 13$ TeV for the lowest (left) and highest (right) multiplicity event classes at midrapidity. The measurements are compared to PYTHIA predictions with the Monash 2013 tune [151] and the CR-BLC modes [110] and the CE-SH model [232], estimated in similar multiplicity classes. The uncertainty bands for the PYTHIA 8 predictions are the statistical uncertainties on the simulations, while for the CE-SH model they refer to the variation of the branching ratios for the augmented charm-baryon states.

ratios would be the same in all collision centrality intervals. The canonical ensemble treatment, however, requires an exact conservation of quantum charges, which becomes increasingly important for small collision systems like pp and e^+e^- [233]. Such canonical corrections introduce a system size (i.e. $\langle dN_{ch}/d\eta \rangle$) dependence for the predicted charm hadron production in the CE-SH model, similar as what was seen in the strangeness sector [234]. The version of their GCE-SH model based on measured charm-baryon spectrum reported by the PDG [16] was observed to strongly underestimate the minimum-bias Λ_c^+/D^0 measurements [102]. For this reason, for the Λ_c^+/D^0 case, the underlying charm-baryon spectrum is augmented to include additional excited baryon states predicted by the Relativistic Quark Model (RQM) [167]. For the D_s^+/D^0 predictions, only the PDG case is shown, since RQM does not modify the D-meson yields with respect to the PDG set. The model calculations describe the Λ_c^+/D^0 ratios reproducing the multiplicity dependence. The D_s^+/D^0 prediction is compatible with the measurement for the low multiplicity class, while it does overestimate the data in thhe highest multiplicity interval.

6.7.2 Transverse momentum-integrated production yields and ratios

The $p_{\rm T}$-integrated yields of $\rm D^0$, $\rm D_s^+$, and Λ_c^+ were computed by integrating the $p_{\rm T}$-differential spectra in their measured range and extrapolating them down to $p_{\rm T} = 0$ in each multiplicity interval. The PYTHIA predictions with CR-BLC Mode 2 were used for the extrapolations in each multiplicity interval. The extrapolation factor was computed as the ratio of the PYTHIA full-$p_{\rm T}$ spectrum integral to the integral in the visible $p_{\rm T}$ range, which was then multiplied to the visible $p_{\rm T}$-integrated charm-hadron yield. To estimate an uncertainty on the extrapolation, the procedure was repeated considering also CR-BLC Mode 0 and Mode 3, and two different functions to fit the spectra (Tsallis-Lévy [229] and a power-law function). The fits were performed considering the statistical and $p_{\rm T}$-uncorrelated sources of systematic uncertainties, and repeated shifting up and down the data by one sigma of the $p_{\rm T}$-correlated systematic uncertainties. The envelope of all trials was assigned as extrapolation uncertainty. The extrapolation factors and the $p_{\rm T}$-integrated yields for all thee charm hadrons in the $\rm INEL_{>0}$, the lowest, and the highest midrapidity multiplicity event classes are reported in Tab. 6.2.

Table 6.2: The $p_{\rm T}$-extrapolated production yields per event ($\times 10^3$) of prompt $\rm D^0$, $\rm D_s^+$, and Λ_c^+ in the $\rm INEL_{>0}$, the lowest, and the highest midrapidity multiplicity event classes in $|y| < 0.5$ in pp collisions at $\sqrt{s} = 13$ TeV.

$\langle {\rm d}N_{\rm ch}/{\rm d}\eta \rangle$	Hadron	Extr. factor	$1/N_{\rm ev} \cdot {\rm d}N/{\rm d}y\vert_{\vert y\vert<0.5}$ ($\times 10^3$)
$\rm INEL_{>0}$	$\rm D^0$	$1.27^{+0.03}_{-0.03}$	$11.67 \pm 0.26\,(\rm stat)\,^{+0.91}_{-0.96}\,(\rm syst)\,^{+0.28}_{-0.32}\,(\rm extr)$
	$\rm D_s^+$	$1.24^{+0.02}_{-0.08}$	$1.92 \pm 0.13\,(\rm stat)\,^{+0.22}_{-0.23}\,(\rm syst)\,^{+0.02}_{-0.13}\,(\rm extr)$
	Λ_c^+	$1.34^{+0.02}_{-0.09}$	$5.72 \pm 0.24\,(\rm stat)\,^{+0.52}_{-0.53}\,(\rm syst)\,^{+0.08}_{-0.40}\,(\rm extr)$
3.1	$\rm D^0$	$1.45^{+0.01}_{-0.13}$	$3.61 \pm 0.14\,(\rm stat)\,^{+0.42}_{-0.34}\,(\rm syst)\,^{+0.02}_{-0.33}\,(\rm extr)$
	$\rm D_s^+$	$2.53^{+0.24}_{-0.03}$	$0.60 \pm 0.04\,(\rm stat)\,^{+0.10}_{-0.09}\,(\rm syst)\,^{+0.06}_{-0.08}\,(\rm extr)$
	Λ_c^+	$1.63^{+0.00}_{-0.34}$	$1.47 \pm 0.13\,(\rm stat)\,^{+0.18}_{-0.18}\,(\rm syst)\,^{+0.00}_{-0.31}\,(\rm extr)$
37.8	$\rm D^0$	$1.14^{+0.11}_{-0.00}$	$91.04 \pm 3.25\,(\rm stat)\,^{+\ 7.93}_{-13.58}\,(\rm syst)\,^{+8.72}_{-0.13}\,(\rm extr)$
	$\rm D_s^+$	$1.49^{+0.49}_{-0.07}$	$13.83 \pm 0.53\,(\rm stat)\,^{+\ 1.66}_{-\ 2.54}\,(\rm syst)\,^{+2.45}_{-0.65}\,(\rm extr)$
	Λ_c^+	$1.18^{+0.05}_{-0.05}$	$45.22 \pm 3.83\,(\rm stat)\,^{+\ 4.97}_{-\ 5.31}\,(\rm syst)\,^{+1.76}_{-1.93}\,(\rm extr)$

The $p_{\rm T}$-integrated $\rm D_s^+/D^0$ and $\Lambda_c^+/\rm D^0$ yield ratios as a function of $\langle {\rm d}N_{\rm ch}/{\rm d}\eta \rangle$ are shown in Fig. 6.19. The results are compared to the $p_{\rm T}$-integrated PYTHIA predictions. The $\rm D_s^+/D^0$ ratios are constant with multiplicity, as expected from the agreement between the $p_{\rm T}$-differential ratios in the different multiplicity event classes. The PYTHIA calculations — which predict an increase in the $\rm D_s^+/D^0$ ratio at extremely low charged-particle multiplicity densities, i.e. special events with just one or zero produced charged-primary particles at midrapidity — agree with the data. For the $\Lambda_c^+/\rm D^0$ ratios, on the contrary, an increase of the $\Lambda_c^+/\rm D^0$ ratio with $\langle {\rm d}N_{\rm ch}/{\rm d}\eta \rangle$ was expected from the $p_{\rm T}$-differential ratios[8]. With the current uncertainties, however, the $\Lambda_c^+/\rm D^0$ ratios (excluding the $\rm INEL_{>0}$ class) are compatible with a constant with a significance of about 1σ. The data disfavour the Monash prediction in the whole multiplicity range, and tend to be below (above) the CR-BLC Mode 2 (3) for the two highest multiplicity intervals.

[8] At least naively, from the 5.3σ significance observation that the $\Lambda_c^+/\rm D^0$ ratio in the highest multiplicity event class is higher than the lowest class in the visible $p_{\rm T}$ interval.

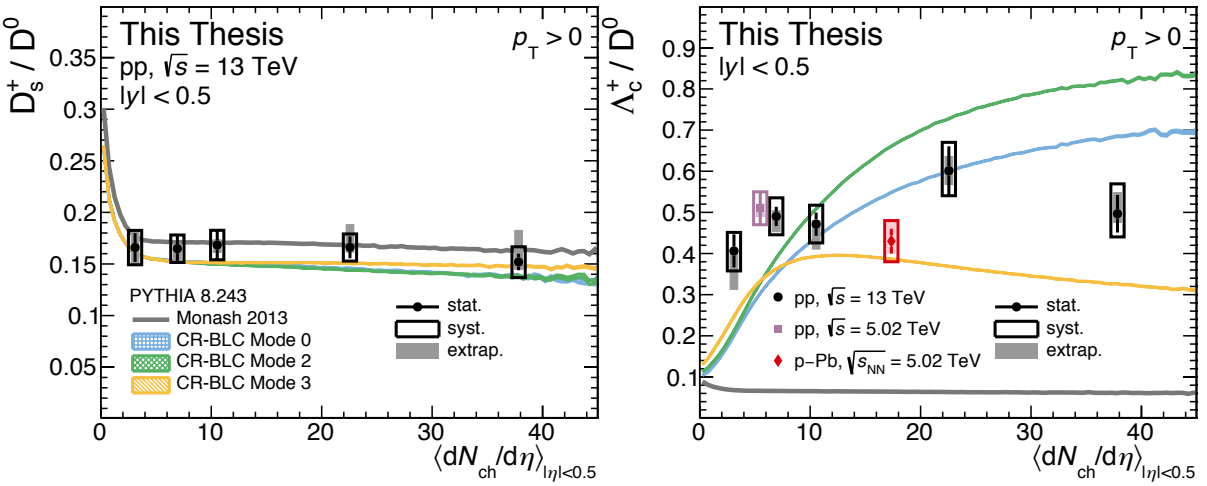


Figure 6.19: The p_{T}-integrated $\mathrm{D_s^+/D^0}$ (left) and $\Lambda_{\mathrm{c}}^+/\mathrm{D^0}$ (right) ratios as function of $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ for pp collisions at $\sqrt{s}$ = 13 TeV compared to PYTHIA predictions [110, 151]. The $\Lambda_{\mathrm{c}}^+/\mathrm{D^0}$ measurements performed in pp and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV from Ref. [102, 103] are also shown.

6.8 Discussion

In the previous section, the main physics results of these charm-hadron production versus multiplicity measurements were presented. I tried to report them as an unbiased reader, i.e. stick to observations as one would do in an official ALICE document. In this section, I will go a bit deeper and discuss (my personal) physical interpretation of these results. Given the lack of a visible multiplicity dependence in the $\mathrm{D_s^+/D^0}$ ratio with the current uncertainties — which is in line with $\mathrm{K^+}/\pi^+$ results [157] — I will focus the discussion mostly on the Λ_{c}^+.

The hardening of the p_{T} spectra with increasing $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$, as we saw in Fig. 6.14, is also observed for the light-flavour hadrons [225, 230]. The effect was shown to be more pronounced for protons than for kaons and pions, and similar for the comparison between strange mesons and baryons. The hardening of the spectra with increasing multiplicity could be attributed to an increasing number of high-p_{T} jets [235], but also the mass dependence of spectral shape modifications — observed in Pb–Pb collisions [50], and usually associated with the hydrodynamical evolution of the system — could play a role. In the latter case, it would thus be a hint of the production of a droplet of QGP in pp collisions. Differentiating between both cases requires a detailed model study to, for example, the slope parameter of the spectra shapes at intermediate p_{T}. While for light-flavour hadrons such study was inconclusive [225, 230], it might be interesting for charm hadrons[9], which are less likely to thermalise because of their large masses. Currently, the experimental precision of these charm-hadron production measurements is too small for such study.

As discussed in Chapter 1, the relative yields between different hadron species may actually give more insights. The conclusions on the reported $\Lambda_{\mathrm{c}}^+/\mathrm{D^0}$ ratios as function of p_{T} (Figs. 6.15

[9]Given the small differences in mass among the here studied charm hadrons (200–400 MeV/c^2), I would do such study with the $\mathrm{D^0}$ meson, to achieve the best precision.

and 6.16) and the extrapolated p_{T}-integrated $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratios as function of multiplicity (Fig. 6.19) might, at first sight, seem contradictory though. On the one hand we observe a significant enhancement with multiplicity of the baryon-to-meson ratios in the visible p_{T} range suggesting a modification of the hadronisation mechanisms with multiplicity, while the ratios extrapolated to $p_{\mathrm{T}} > 0$ show no dependency with $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$. One can argue that the relatively large uncertainties on the p_{T}-extrapolated ratios "spoil" the significance calculation, i.e. the data is in agreement with a rising trend with multiplicity as well. Another explanation would be that there is indeed no enhancement of charm baryon production with respect to charm mesons, and that the increase observed in the p_{T}-differential ratios is due to a different redistribution of momentum for baryons and mesons. If this would have been a comparison between Pb–Pb collisions to pp, I would argue that such modification of the p_{T} shape is expected because of radial flow (which is the caveat I had in mind when discussing the trends in small p_{T} intervals of Fig. 6.17). In small systems like pp, it is difficult to imagine that radial flow builds up in a sufficient way to justify the observation, so the explanation should likely be sought somewhere else. This is an active topic of discussion in the field though, since many (especially light-flavour) results in pp multiplicity classes are in compliance with this hydrodynamic origin picture.

To investigate further on the underlying physical mechanisms that can explain a remarkable observation, comparisons to model predictions are needed. Unfortunately, there are currently no theoretical predictions for both the $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratios as function of p_{T} and $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ on the market, besides our own PYTHIA 8 simulations, for which the CR-BLC modes [110] qualitatively describe the enhancement for the p_{T}-differential ratios, but fail to describe the constant trend for the extrapolated ratios. This is not that worrisome though, since the tuning for these CR-BLC modes was performed basically on light-flavour results only, and the authors do not expect it to work out-of-the-box for heavy-flavour hadrons (P. Skands, private communication, May 2021). The measurements presented in this chapter thus provide an interesting possibility for them to further constrain their specific theoretical concept. More in general, while the factorisation approach is satisfactory to describe the inclusive production of heavy-flavour mesons in minimum-bias pp events, we currently lack a proper understanding of the interplay between the soft and hard components of the event. The multiplicity dependence of charm-hadron production is thus a topic in which further theoretical developments are required.

That all being said, let me make the necessary side remark that the significance calculations and comparisons to theory in this chapter are performed using the analyses exploiting the midrapidity multiplicity estimator. It is, however, known that measuring the p_{T} spectra of charged particles in the same kinematic region as where the multiplicity selection is performed introduces a trivial difference with respect to non-coinciding rapidity regions. It is therefore common practice in the light-flavour sector to use the p_{V0M} multiplicity estimator. The authors of Ref. [236], which performed a simulation study for heavy-flavour probes regarding this fragmentation bias, also highly advise to do the same to disentangle between autocorrelation and true correlation effects between the hard probe and the underlying event. For future work, this should certainly be taken into account in more detail, but for now I can only comment that the conclusions of this

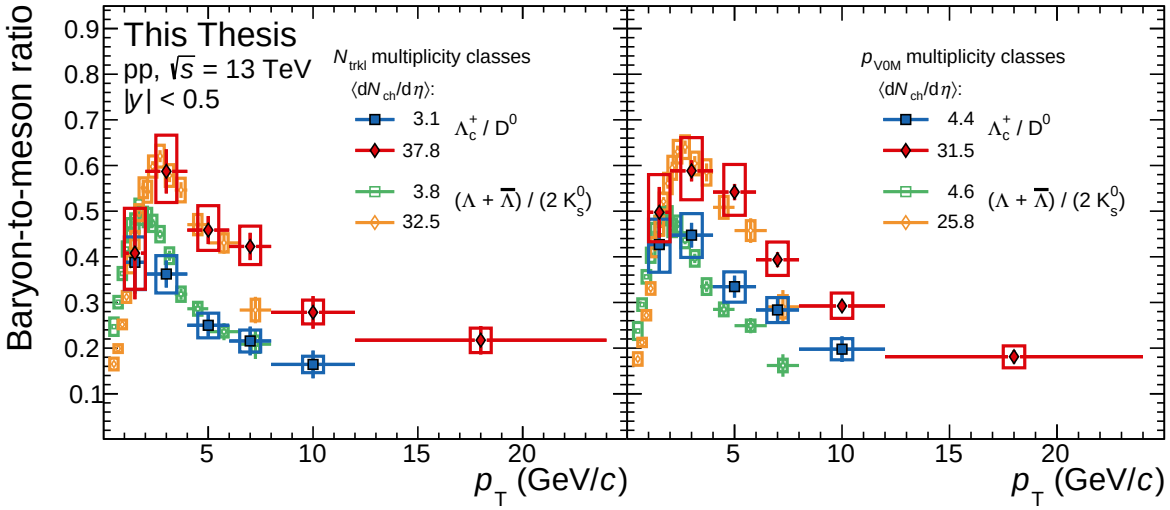


Figure 6.20: The baryon-to-meson ratios in the light-flavour, based on measurements from Ref. [225] and charm sector measured in pp collisions at $\sqrt{s}$ = 13 TeV for similar low- and high-multiplicity event classes at mid- (left) and forward (right) rapidity.

chapter do not change significantly when using the $p_{\rm V0M}$ distributions[10].

Finally, it is interesting to note that the light-flavour equivalent, the $\Lambda/{\rm K}^0_{\rm S}$ baryon-to-meson ratio, shows a remarkably similar enhancement with $\langle {\rm d}N_{\rm ch}/{\rm d}\eta \rangle$ for the $p_{\rm T}$-differential ratios (see Fig. 6.20), while the $p_{\rm T}$-integrated ratios show no significant dependence [8, 225]. These similarities between the light- and heavy-flavour baryon-to-meson ratios were observed as well in minimum-bias pp and p–Pb collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV both in terms of shape and magnitude [102], and also qualitatively similar trends were observed for Pb–Pb collisions as we will see in the next chapter. These similarities, if not accidental, are surprising given that in the vacuum fragmentation scenario, the light-flavour hadron production has a significant contribution from gluon fragmentation, whereas heavy-flavour hadrons are primarily produced through the fragmentation of a charm quark produced in the initial hard scattering. In addition, at low $p_{\rm T}$, light-flavour hadrons originate mainly from small-momentum soft scattering processes. These observations hint at a potential common mechanism for light- and charm-baryon formation in hadronic collisions at LHC energies.

In summary, I think these charm hadron measurement as function of multiplicity are rather interesting, shed further light on the $\Lambda_{\rm c}^+$ puzzle, and, in general, will teach us more about charm hadronisation and the possible hydrodynamic picture in pp collisions. The measurements reported in this chapter are, however, limited by the large systematic uncertainties. In addition, to extract a good signal, coarse intervals in multiplicity and transverse momentum had to be used, despite this being the largest data sample collected by the ALICE Collaboration so far. Further measurements with Run 3 data are in my opinion fundamental to investigate on the shape of the $p_{\rm T}$-integrated baryon-to-meson ratios with multiplicity, to extend the multiplicity reach to lower and higher multiplicity intervals, and to provide more precise measurements.

[10]Note that this "re-analysis" was done with less care than the results using the midrapidity multiplicity estimator reported in this chapter. Especially for the $p_{\rm T}$-integrated ratios and comparisons with models.

CHAPTER 7

Λ_c^+-baryon production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV

Abstract – The p_T-differential production yield of prompt Λ_c^+ baryons was measured in central (0–10%) and mid-central (30–50%) Pb–Pb collisions at the centre-of-mass energy per nucleon–nucleon pair $\sqrt{s_{NN}} = 5.02$ TeV, with the ALICE detector at the LHC. The Λ_c^+ baryons and their charge conjugates were reconstructed via the decay channel $\Lambda_c^+ \rightarrow pK_S^0 \rightarrow p\pi^+\pi^-$ at midrapidity ($|y| < 0.5$) in the transverse momentum interval $1 < p_T < 24$ GeV/c. The Λ_c^+/D^0 production ratio, the Λ_c^+ nuclear modification factor R_{AA}, and the Λ_c^+ p_T-integrated production yield are reported. The results are more precise and more differential in transverse momentum and centrality than previous measurements. The Λ_c^+/D^0 ratio in central Pb–Pb collisions in the interval $4 < p_T < 8$ GeV/c was found to be larger than the one measured in pp collisions, while the p_T-integrated Λ_c^+/D^0 ratio is compatible within uncertainties with the ratio in pp collisions. Theoretical calculations that model the charm-quark transport in the quark–gluon plasma and include hadronisation via both coalescence and fragmentation mechanisms describe the p_T-differential Λ_c^+/D^0 ratio and R_{AA} within uncertainties.

Manuscript submitted to Phys. Rev. Lett. (arXiv:2112.08156 [nucl-ex])

7.1 Introduction

Measurements of open heavy-flavour production in heavy-ion collisions provide a unique method to investigate the mechanisms of in-medium energy loss and the hadronisation of heavy quarks from the quark–gluon plasma [86]. The hadronisation of heavy quarks into open heavy-flavour hadrons is expected to be influenced by the presence of a deconfined medium: theoretical calculations that include modified hadronisation via quark coalescence or via a resonance recombination approach [160, 162, 164, 165, 237] (see Chapter 2 for more details) predict a significant enhancement of the Λ_c^+/D^0 ratio in heavy-ion collisions at low and intermediate transverse momentum compared to the expected ratio in pp collisions. In addition, the possibility that binary, coloured bound states of light partons exist in the QGP [166] could further contribute to the baryon enhancement [164, 165]. The study of such a potential enhancement of p_T-integrated yield requires a good understanding of Λ_c^+ production in smaller collision systems, which actually showed surprising features at LHC energies [102, 103]. This is what we called the Λ_c^+ puzzle throughout this dissertation, and boils down to the observation of a significant enhancement of the Λ_c^+/D^0 yield ratio in pp collisions at midrapidity with respect to measurements in e^+e^- and e^-p collisions.

In Chapter 6, the measurement of the Λ_c^+/D^0 (and D_s^+/D^0) ratio in intervals of charged-particle multiplicity in pp collisions at $\sqrt{s}$ = 13 TeV was discussed [238], motivated by the observation of the Λ_c^+ puzzle in minimum-bias pp collisions. As we have seen, this measurement revealed a significant increase of the Λ_c^+/D^0 ratio in the measured $1 < p_T < 12$ GeV/c range from the lowest to the highest multiplicity interval, although the extrapolated p_T-integrated Λ_c^+/D^0 ratios seem to depend less on the multiplicity. The low-multiplicity pp data suggest that in a hadronic collision, even at relatively small multiplicities, charm-quark hadronisation proceeds differently than in the e^+e^--collision environment. Whether the p_T-differential Λ_c^+/D^0 ratio keeps evolving with multiplicity up to the typical multiplicities of Pb–Pb collisions, and whether an overall p_T-integrated enhancement of Λ_c^+ production relative to the D^0 one is present at higher multiplicities, is an open question, fundamental to understand charm-quark hadronisation.

The Λ_c^+ production in nucleus–nucleus collisions was measured for the first time at the LHC by ALICE in Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV in the 0–80% centrality interval for $6 < p_T < 12$ GeV/c [199]. The Λ_c^+/D^0 ratio was found to be close to unity, larger than the corresponding ratio measured in pp collisions, and well described by calculations including hadronisation vvia coalescence mechanisms [160, 162]. The Λ_c^+/D^0 ratio measured in the interval $3 < p_T < 6$ GeV/c by the STAR Collaboration in Au–Au collisions at $\sqrt{s_{NN}}$ = 200 GeV [239], is significantly larger than the value obtained with PYTHIA simulations of pp collisions [110, 151] and in agreement with model calculations including hadronisation via coalescence [160, 162, 165, 240–242]. The CMS measurement in Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV [104], performed in the interval $10 < p_T < 20$ GeV/c, is consistent with the pp result, suggesting that modified hadronisation due to the presence of a deconfined medium has no significant effect in this p_T range.

In the light-flavour sector, baryon-to-meson ratios like the p/π [50] and Λ/K_S^0 [243] in Pb–Pb collisions are studied extensively as well. The p_T-dependent ratios show distinctive peaks with

respect to pp collisions at intermediate p_T (~ 3 GeV/c), being more pronounced than observed for meson-to-meson ratios, while the p_T-integrated ratios show no significant enhancement. This increase is attributed to the hadron-mass dependent p_T shift due to the presence of radial flow in heavy-ion collisions, which affects heavier particles more than lighter ones, and to hadronisation via coalescence, which will push baryons to higher momenta with respect to mesons because of the additional constituent quark (even in the absence of flow). The production of Λ_c^+ and D^0 hadrons are affected as well by radial flow effects, thus providing an additional push of the Λ_c^+/D^0 ratio to higher momenta. The study of Λ_c^+-baryon production, compared to the one of non-strange D and D_s^+ mesons, could help to better disentangle the role of coalescence and radial flow because of the smaller mass differences than for light-flavour hadrons. On a similar note, the effect of hadronic diffusion on the Λ_c^+ and D^0 yields is not clearly addressed in the current literature or with inconsistent conclusions [244, 245].

In this chapter, the measurements of the prompt Λ_c^+ p_T-differential yields and Λ_c^+/D^0 yield ratios performed using the 2018 Pb–Pb dataset are reported. These measurements extend the study of the hadronisation properties of charm baryons to the extreme conditions realised in central Pb–Pb collisions. The measurement is presented in central (0–10%) and mid-central (30–50%) collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the range $1 < p_T < 24$ GeV/c. The values of the Λ_c^+/D^0 ratios are reported as a function of p_T and compared to theoretical predictions. The p_T-integrated Λ_c^+ production yield, extrapolated to $p_T = 0$, and the nuclear modification factor R_{AA} are also presented. This chapter is organised in a similar way as the previous ones, discussing first the selection procedure, raw-yield extraction, corrections, and evaluation of the systematic uncertainties, after which the results are presented and discussed in detail.

7.2 Selection criteria

The Λ_c^+ baryon and its charge conjugate were reconstructed by exploiting the topology of the hadronic decay channel $\Lambda_c^+ \to pK_S^0 \to p\pi^+\pi^-$ (BR = $1.10 \pm 0.06\%$) [16]. The Λ_c^+-candidate selection was performed using multivariate techniques based on the Boosted Decision Tree (BDT) algorithm provided by the XGBoost package [197], as introduced in Sec. 4.2. The BDT training was performed considering as signal candidates prompt Λ_c^+ decays simulated with PYTHIA 8.243 [95], embedded into a Pb–Pb event simulated with HIJING 1.36 [190] and transported through the ALICE detector geometry using GEANT3 [191]. Background candidates were taken from the sidebands of the invariant-mass distribution in data. The selection features considered in the optimisation included both topological and particle-identification variables. The $N_\sigma^{comb}(p)$ PID variable is the most important training variable in all p_T intervals. The optimal selections on the BDT outputs were tuned to maximise the expected statistical significance estimated by FONLL predictions [83, 87], as introduced in Sec. 4.2.4. The f_{prompt} parameter was set to 0.9 for all BDT selections and the R_{AA} parameter was assumed to be equal to one[1]. The threshold values for the different p_T intervals vary between 0.31 and 0.92.

[1]Both these assumptions could be improved by calculating the values "on-the-fly" for each BDT selection. The values chosen now, however, do not bias the measurement in any way.

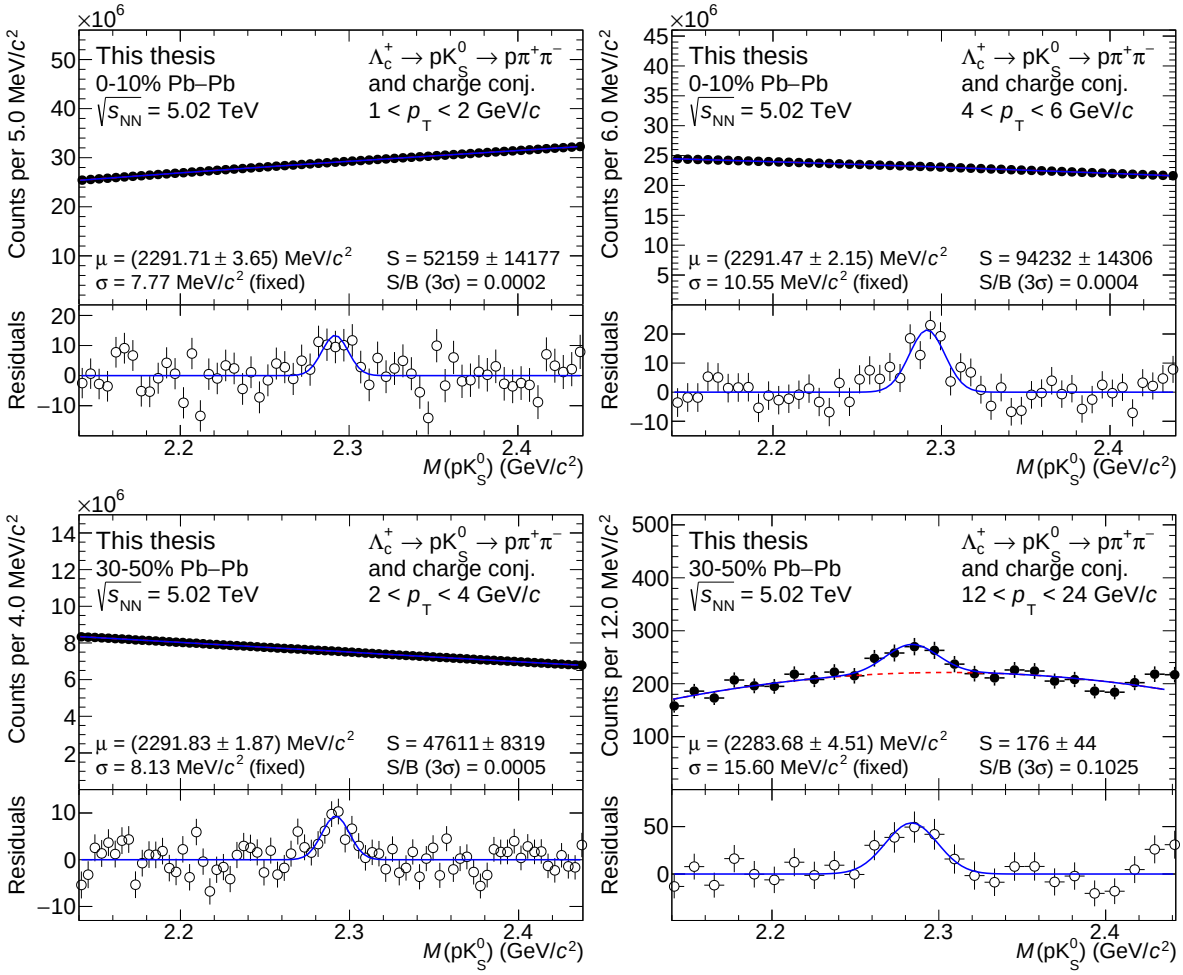


Figure 7.1: Invariant-mass distribution of Λ_c^+ candidates and charge conjugates in low (left) and intermediate or high (right) p_T intervals for 0–10% (top) and 30–50% (bottom) Pb–Pb collisions at midrapidity. The blue solid lines show the total fit function, while the red dashed lines are the combinatorial background terms. The values of the mean (μ) and the width (σ) of the signal peak are reported, together with the signal counts (S) and the signal-to-background ratio (S/B) in the mass interval $(\mu - 3\sigma, \mu + 3\sigma)$. The bottom panels show the residuals of the invariant-mass fit, to better visualise the signal peak.

7.3 Raw-yield extraction

The yields of Λ_c^+ baryons were extracted in six p_T intervals in $1 < p_T < 24$ GeV/c for both centrality classes via binned maximum-likelihood fits to the candidate invariant-mass distributions. The fitting function consisted of a Gaussian term to estimate the signal and a second-, third-, or fourth-order polynomial function (depending on the p_T interval) to estimate the background. The width of the Gaussian was fixed in each p_T interval to the values obtained from simulations to improve the fit stability. The choice of the order of the polynomial function used for the central fit was based on a detailed study of the residuals of the invariant-mass fit. Differences in the performance between free and fixed Gaussian widths, as well as the different background fit functions, are taken into account in the evaluation of the raw-yield extraction systematic source.

Examples of the invariant-mass fits for central and mid-central Pb–Pb collisions are reported in Fig. 7.1. The blue solid lines show the total fit functions and the red dashed lines are the

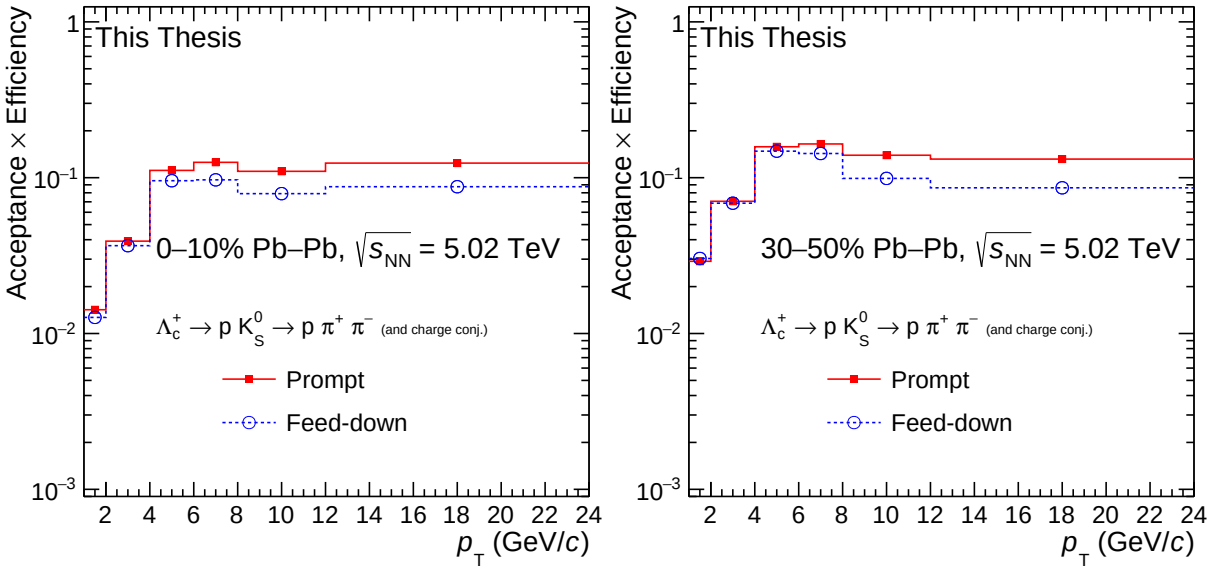


Figure 7.2: The acceptance-times-efficiency for Λ_c^+ baryons as function of p_T in 0–10% (left) and 30–50% (right) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The efficiencies are shown for prompt (solid lines) and feed-down (dotted lines) Λ_c^+ baryons.

combinatorial-background terms. The bottom panels report the residuals. The values of the raw yield, signal-to-background ratio, and significance (within 3σ) vary between 176 ± 44 to $(142 \pm 25) \cdot 10^3$, 0.0001 to 0.1, and 3.0 ± 1.1 to 6.0 ± 0.9, respectively. The raw-yield extraction is challenging, especially at low p_T, with signal-to-background ratios below one per mille. The Gaussian width of the fitted Λ_c^+ peaks, fixed to the value extracted from simulations, vary between 7.5 to 18.8 MeV/c^2, increasing with p_T as expected from the p_T-resolution of the decay tracks. The extracted Gaussian means vary around 2290 MeV/c^2, slightly above the PDG average of 2286.46 ± 0.14 MeV/c^2 [16].

There is no contamination from other (resonant) decay channels, neither from reflected candidates (coming from the possible ambiguity in assigning the decay-particle mass), hence the extracted raw yields from the invariant-mass fits are directly used in Eq. 4.11. Since the statistical significance of the Λ_c^+ signals are rather low, the invariant-mass background shapes were compared with simulations. The shapes of background Λ_c^+ candidates from an ALICE general-purpose Monte Carlo simulation (i.e. mimics data as well as possible) after the central selections on the BDT output scores were found to be compatible with the shapes observed in Fig. 7.1, and the invariant-mass fits did not converge to any Λ_c^+ signal. This indicates that the raw yields are not affected by an "artificial" peak created by the ML model, as is sometimes the fear for low-significance signals.

7.4 Corrections

Figure 7.2 shows the acceptance-times-efficiency $A \cdot \epsilon$ for prompt and feed-down Λ_c^+ baryons within the fiducial acceptance as function of p_T for both centrality classes. The acceptance term is estimated directly from the used ALICE simulations. The generated p_T spectrum used to

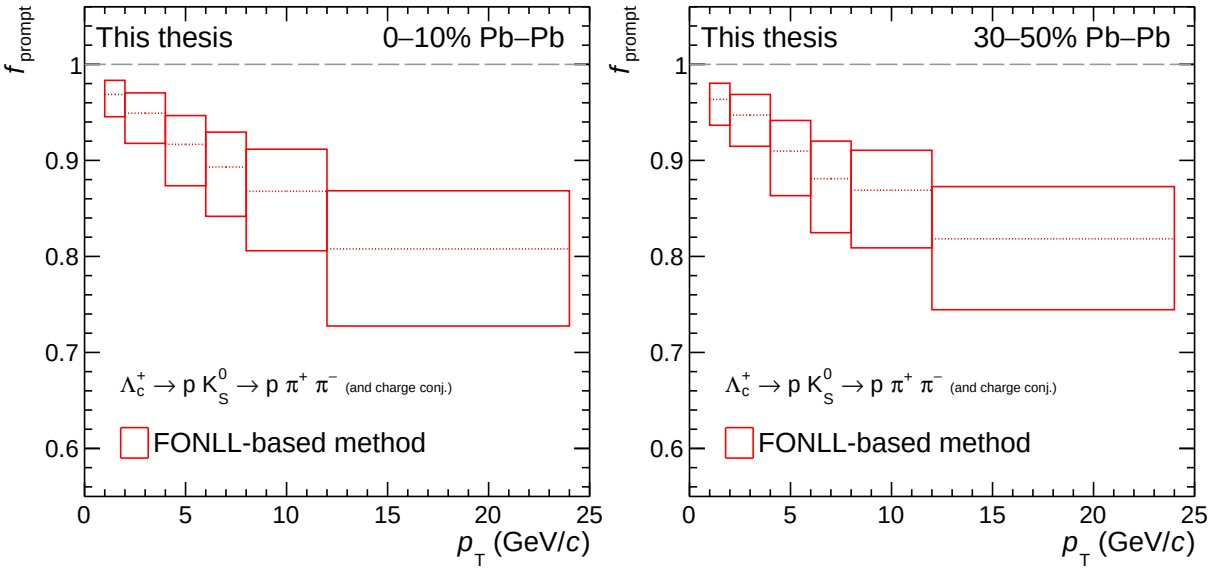


Figure 7.3: The estimated prompt fraction for Λ_c^+ baryons as function of p_T in 0–10% (left) and 30–50% (right) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV using the FONLL-based approach.

calculate the efficiencies was reweighted to reproduce the shape obtained from the D^0 measurement [220] multiplied by Λ_c^+/D^0 calculations from the TAMU model [162] in 0–10% and 30–50% Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The $(A \times \epsilon)$ for prompt Λ_c^+ baryons varies from 1 (3)% at low p_T to about 12 (16)% at high p_T for central (mid-central) collisions. The difference with respect to the $A \cdot \epsilon$ for feed-down Λ_c^+ is induced by the impact-parameter in the transverse plane of the proton-candidate track, which is among the BDT training variables and sensitive to the on-average further displacement from the primary vertex for beauty hadrons.

The fraction of prompt Λ_c^+ baryons was evaluated with the theory-driven method as introduced in Sec. 4.3. In particular, the beauty-quark production cross section was estimated with FONLL calculations [83, 87], the fraction of beauty quarks that fragment into Λ_b^0 was estimated from the $\Lambda_b^0/(B^0 + B^+)$ ratio measured by LHCb in pp collisions at $\sqrt{s} = 13$ TeV [108][2], and the $\Lambda_b^0 \to \Lambda_c^+ + X$ decay kinematics taken from PYTHIA 8. Furthermore, the branching ratio $f(\Lambda_b^0 \to \Lambda_c^+ + X) = 33\%$ is used for the normalisation [16] and the central hypothesis of the ratio $R_{AA}^{non-prompt}/R_{AA}^{prompt}$ was considered to be equal to 2 for Λ_c^+ as predicted by the "Catania" theoretical calculation [237] since no measurements of beauty-baryon production in nucleus–nucleus collisions are available.

The resulting f_{prompt} fraction was found to be about 97% at low p_T and about 81% at high p_T, as presented in Fig. 7.3 for both centrality intervals. The shown systematic uncertainty is due to the variation of the FONLL parameters (as prescribed in Ref. [83]) and variations of the parametrisation of the Λ_b^0 fragmentation fraction (see Ref. [102]). A systematic uncertainty is also assigned for the $R_{AA}^{non-prompt}/R_{AA}^{prompt}$ hypothesis, which will be discussed later.

[2]See Ref. [102] for more details on how the LHCb measurement is used. The parametrisation assumes $f_u + f_d + f_s + f_{\Lambda_b^0} = 1$ (the fraction of b quarks that hadronise in $\overline{B}^0$, B^-, $\overline{B}_s^0$, and Λ_b^0, respectively), $f_u = f_d$, and no rapidity dependence in $f_{\Lambda_b^0}$.

7.4.1 Proton–proton reference

The f_{prompt} calculation, as well as the actual nuclear modification factor of Λ_c^+ baryons, requires a pp reference in the same p_T intervals. The p_T-differential Λ_c^+ production cross section in pp collisions at $\sqrt{s} = 5.02$ TeV from Ref. [102] is used as baseline. This measurement is performed in seven p_T intervals between $1 < p_T < 12$ GeV/c, and was thus rebinned to the same p_T intervals as the Λ_c^+ measurement in Pb–Pb considering the statistical uncertainties and the raw-yield extraction systematic uncertainty as uncorrelated with p_T ($\delta_{\text{stat},1,14}$).

The p_T-differential cross section of prompt Λ_c^+ in pp collisions was extrapolated to $12 < p_T < 24$ GeV/c exploiting the Λ_c^+ and D^0 measurements at $\sqrt{s} = 5.02$ and 13 TeV [82, 231], assuming no $\sqrt{s}$-dependence for the Λ_c^+/D^0 ratio. This assumption is true within statistical and systematic uncertainties in the full measured p_T range. Two different procedures were tried as well to validate the central method: using (i) the FONLL theoretical prediction for D^0 mesons [83] times an ad-hoc parametrisation of the Λ_c^+/D^0 ratio in pp at $\sqrt{s} = 5.02$ TeV [82] and (ii) using the Λ_c^+ measurement in $12 < p_T < 24$ GeV/c in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [102] scaled by $A = 208$ and assuming $R_{\text{pPb}} = 1$. All three methods were found to be compatible within the uncertainties.

7.5 Systematic uncertainties

The systematic uncertainties on the Λ_c^+ corrected yields include contributions from: the extraction of the raw yield from the invariant-mass distributions (δ_1); the track-reconstruction efficiency (δ_2); the Λ_c^+-baryon selection efficiency (δ_3); the PID selection efficiency (δ_4); the p_T shape of generated Λ_c^+ baryons in the simulation (δ_5); and the subtraction of the feed-down from beauty-hadron decays (δ_6). In addition, the uncertainties on the branching ratios (δ_7), the centrality interval limits (δ_{13}), and the statistical uncertainty on the acceptance-times-efficiency values (δ_{14}) were considered. In the subsections below, the different sources will be discussed in more detail.

7.5.1 Raw-yield extraction

The systematic uncertainty on the raw-yield extraction was evaluated in each p_T interval for both centrality classes by repeating the fit to the invariant-mass spectra varying the fit range and the background fit function via the so-called multi-trial approach. In particular, five lower (between 2.12 and 2.16 GeV/c^2) and five upper (between 2.42 and 2.46 GeV/c^2) limits for the invariant-mass fit interval were used. A second-, third-, and fourth-order polynomial function were considered as background fit function. A shift of the invariant-mass distributions of 1 to $(r-1)$ MeV/c^2 (with integer steps) was applied as well before rebinning to the final r MeV/c^2 bin widths. The fits were performed with the width and/or position of the signal peak fixed to the values extracted from simulations, as well as leaving both as free parameters. Only fits with a $\chi^2/\text{ndf} < 3$, a width of the signal peak within $0.5 < \sigma/\sigma_{\text{MC}} < 2$, and a significance higher than 2.8 were considered. In case these requirements reject most of the free-width trials,

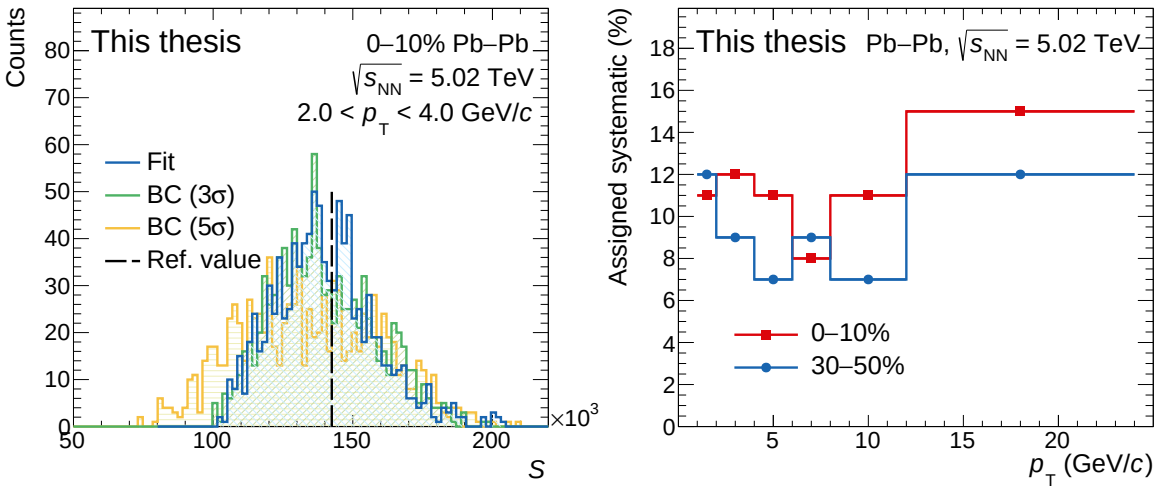


Figure 7.4: Left: Raw-yield distributions obtained with the multi-trial approach via the fit and bin-counting (BC) procedures for Λ_c^+ baryons in $2 < p_T < 4$ GeV/c in 0–10% Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Right: Summary of the assigned systematic uncertainties for the raw-yield extraction in both centrality classes.

variations of ±15% the MC width were applied as well. The same strategy was performed using a bin-counting method, in which the signal yield was obtained by integrating the invariant-mass distribution in a $\pm 3\sigma$ and $\pm 5\sigma$ region around the Λ_c^+ peak after subtracting the background, estimated from a fit to the side-bands only. The left panel of Fig. 7.4 shows an example of the multi-trial distributions for $2 < p_T < 4$ GeV/c for the central centrality class.

The systematic uncertainties, assigned based on the RMS of the distribution and shift of the mean with respect to the value of the central fit, range between 7% and 15% depending on the p_T, as summarised in the right panel of Fig. 7.4.

7.5.2 Track-reconstruction efficiency

Systematic effects on the track-reconstruction efficiency were estimated by varying the TPC track-quality selection criteria and by comparing the prolongation probability of the TPC tracks to the ITS hits in data and simulation (matching efficiency). This latter source is evaluated centrally in ALICE via the procedure explained in Sec. 5.5 and was for this analysis only taken into account for the proton-candidate track, since for the K_S^0 decay daughters no ITS requirements are asked. For proton tracks with $p_T < 5$ GeV/c, the final systematic uncertainty due to the matching efficiency varies between 5% and 10%, while at higher p_T it reduces to 1%.

A possible systematic uncertainty due to the track-quality selections on the decay daughters is estimated individually for protons and pions with a per-track study. Relatively pure samples of protons (from Λ decays) and pions (from K_S^0 decays) were selected based on the Armenteros-Podolanski value [193]. The same track-quality and PID selections as the central Λ_c^+ analysis were required. Five variations of the track-quality selection were tested to evaluate a systematic uncertainty. In particular, a number of TPC crossed rows larger than $120 - 5/(p_T$ [GeV/c]), a ratio of crossed rows to findable clusters in the TPC of 0.9, and a number of clusters available

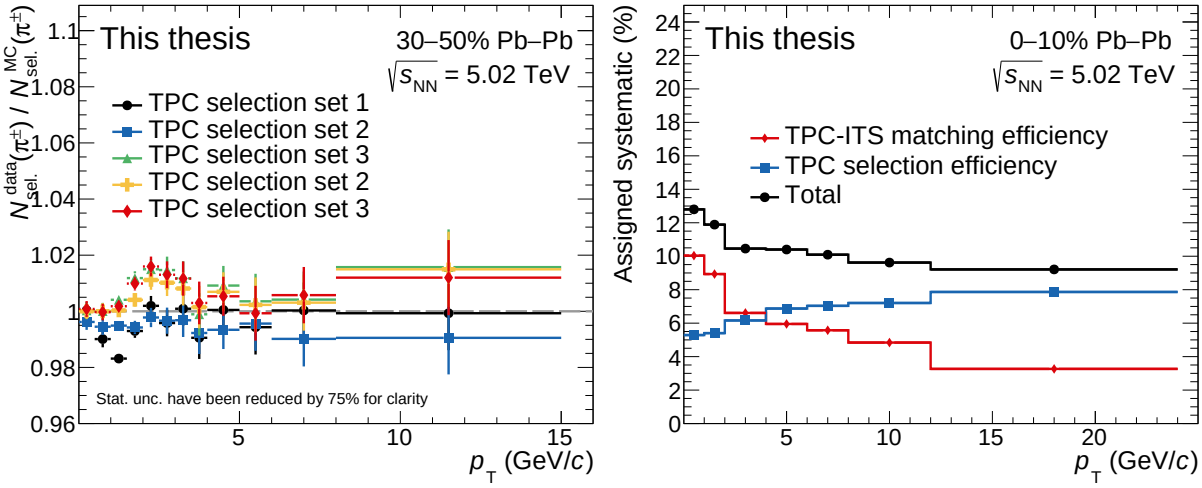


Figure 7.5: Left: The relative difference between data and simulation in the selected number of pion tracks for the central and alternative TPC track-quality selections in mid-central (30–50%) Pb–Pb collisions. Right: The total assigned systematic uncertainty due to the track-reconstruction efficiency for central (0–10%) Pb–Pb collisions, including as well both independent sources.

for particle identification in the TPC of at least 50% of the number of TPC crossed rows were required. The number of clusters for PID was also varied to be larger than 40 or 60, to estimate the effect of the minimum number of 50 clusters that was added to reduce the effects of the not sufficiently simulated track reconstruction at the TPC sector edges.

The assigned systematic uncertainty related to the track-quality selections is defined as the maximum deviation from one in the data-to-simulation ratio of $N_{\mathrm{sel}}^{\mathrm{data}}(\mathrm{p},\pi)/N_{\mathrm{sel}}^{\mathrm{MC}}(\mathrm{p},\pi)$, where $N_{\mathrm{sel}}(\mathrm{p},\pi)$ is the number of selected protons and pions, respectively. The low statistics for protons in the used simulations introduced significant fluctuations in this double ratio, so the maximum deviation was instead taken from a first-order polynomial fit. For pions it was found to vary between 1% and 2.5% per track, while for protons a larger uncertainty of 2% to 6% was found. The left panel of Fig. 7.5 shows an example for pions in mid-central Pb–Pb collisions.

The final assigned systematic uncertainty follows from the quadratic sum of the different per-track contributions, propagated to the Λ_c^+ baryons via the kinematics of the decay tracks[3]. The total uncertainty due to the track-reconstruction efficiency varies between 8% and 12%, where the matching efficiency is dominant at low p_{T} and the TPC track-quality efficiency at high p_{T}, as illustrated in the right panel of Fig. 7.5.

7.5.3 Topological, kinematical, and PID selection efficiency

The systematic uncertainty on the selection efficiency originates from imperfections in the simulation of the decay kinematics and topology and from the resolutions and alignments of detectors in the simulation. It was estimated by comparing the corrected yields obtained by repeating the analysis with different selections on the BDT output score, resulting in a significant modification

[3]Because of practical reasons, the selection on the BDT output score is not included when propagating. We thus assume that the Λ_c^+ kinematics are approximately the same after the prefilter and the final selections.

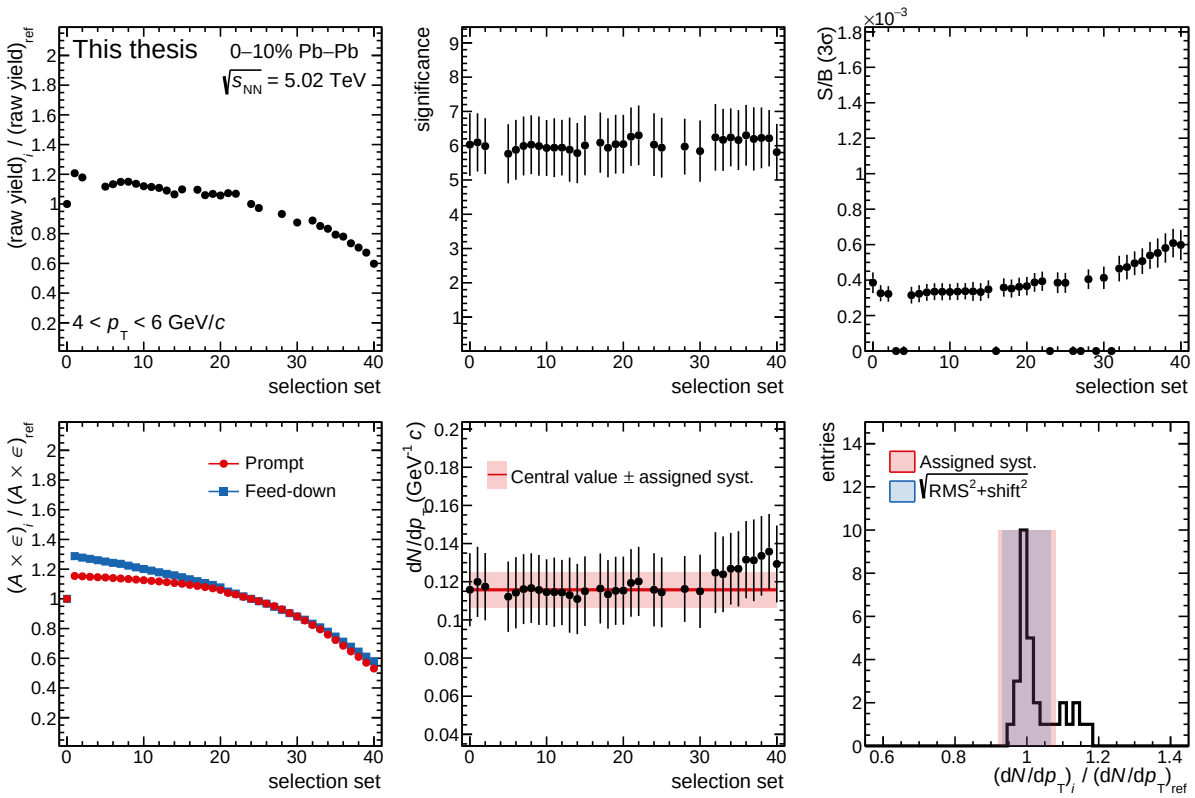


Figure 7.6: Summary on the estimation of the systematic uncertainty on the selection efficiency for Λ_c^+ baryons in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for $4 < p_T < 6$ GeV/c in the central centrality interval. The variation of the extracted raw yield (top-left), the statistical significance (top-middle), the signal-to-background ratio (top-right), the variation of the acceptance-times-efficiency (bottom-left), and the corrected yield per event (bottom-middle) as function of the selection set are presented. The ratio of the dN/dp_T for each selection set with the central one is reported in the bottom-right panel.

of the efficiencies, raw yield, and background values. This procedure, illustrated in Fig. 7.6 for $4 < p_T < 6$ GeV/c in the central centrality class, covers both the PID and topological selection efficiency. The variations are performed separately for the different centrality and p_T intervals. The systematic uncertainty, normally assigned based on the RMS of the distribution and shift of the mean with respect to the central value, is affected by the low statistical significance of the (optimised) central selections. To not include these statistical fluctuations in the assigned systematic uncertainty, it was therefore chosen to assign a conservative p_T-independent value of 8% (7%) for central (mid-central) Pb–Pb collisions.

Possible systematic effects due to the loose PID selection, applied prior to the machine-learning one, were investigated comparing the proton and pion PID selection efficiencies in data and simulation. The differences were found to be negligible, and therefore no extra systematic uncertainty was assigned.

7.5.4 Generated MC p_T shape

As discussed above, the simulated p_T distribution from PYTHIA 8 is reweighted to a more realistic distribution. In particular, it is tuned to reproduce the shape obtained from the D^0

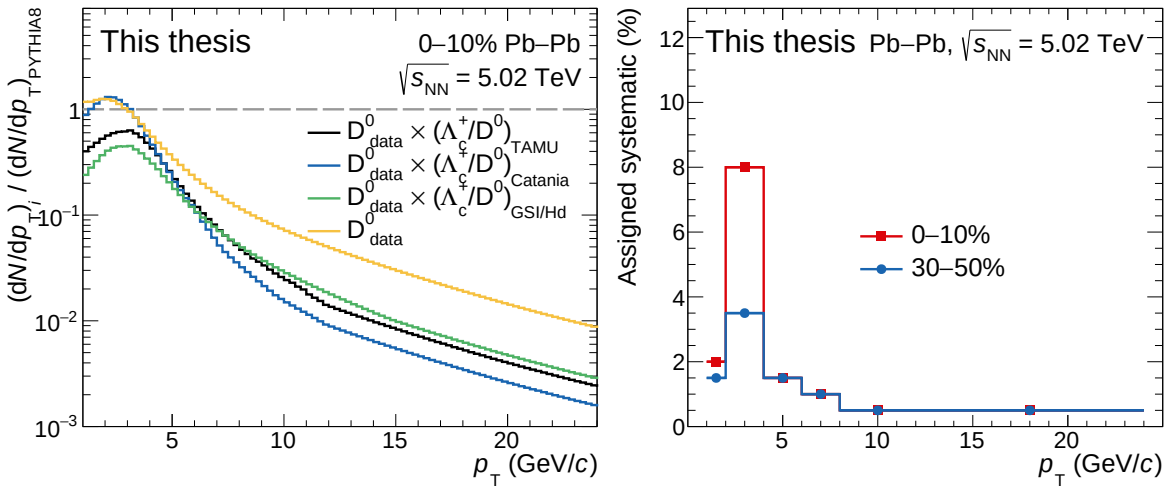


Figure 7.7: Left: Ratio between the normalised $\mathrm{d}N/\mathrm{d}p_\mathrm{T}$ distributions for prompt Λ_c^+ baryons from dedicated theoretical predictions [160, 162, 246] or measurements [220] and PYTHIA 8 [95] simulations. Right: Summary of the assigned systematic uncertainty due to the generated p_T shape.

measurement [220] multiplied by Λ_c^+/D^0 calculations from the TAMU model [162] in 0–10% and 30–50% Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV. To estimate a possible systematic uncertainty, the p_T weights were varied considering different Λ_c^+/D^0 theoretical predictions (from the Catania [160] and GSI–Heidelberg SH [246] models) as well as to reweight them to just the measured D^0 p_T distribution in Pb–Pb collisions [220]. The difference between the different weights are shown in the left panel of Fig. 7.7. The assigned systematic uncertainties are based on the ratio between the prompt efficiencies for the central and alternative p_T-shape hypotheses, defined as the RMS of the differences with respect to one. This contribution ranges from between 0.5% to 8% for central and between 0.5% to 3.5% for mid-central Pb–Pb collisions.

7.5.5 Feed-down subtraction

The systematic uncertainty on the prompt fraction was estimated by varying the FONLL parameters (in particular the b-quark mass and the renormalisation and factorisation scales) in the calculation of the feed-down Λ_c^+-baryon cross section. In addition, the function describing the parametrisation of the Λ_b^0 fragmentation fraction from the $\Lambda_b^0/(\mathrm{B}^0 + \mathrm{B}^+)$ production measurement by LHCb in pp collisions at $\sqrt{s}$ = 13 TeV [108] was varied. The three free fit parameters of Ref. [108] were varied independently within their uncertainties, after which the envelope defined the lower and upper uncertainty bound. For $p_\mathrm{T}(\Lambda_b^0) < 5$ GeV/c, which is outside the kinematic range of the LHCb measurement, the lower uncertainty bound was taken to be equal to the value at $p_\mathrm{T}(\Lambda_b^0) = 5$ GeV/c, while the upper uncertainty bound was estimated from the p_T-dependent fit. In order to account for a possible $\sqrt{s}$-dependence of the fragmentation fractions, an additional reduction of the lower bound was considered based on the spread of the LHCb measurements at different collision energies. The assigned systematic uncertainty values for the Λ_c^+ baryons range from $^{+1.5}_{-2.4}$% at low p_T to $^{+7.5}_{-9.9}$% at high p_T, as already presented in Fig. 7.3.

For the part related to the hypothesis of $R_\mathrm{AA}^\mathrm{non-prompt}$, an interval $1/3 < R_\mathrm{AA}^\mathrm{non-prompt}/R_\mathrm{AA}^\mathrm{prompt} < 3$

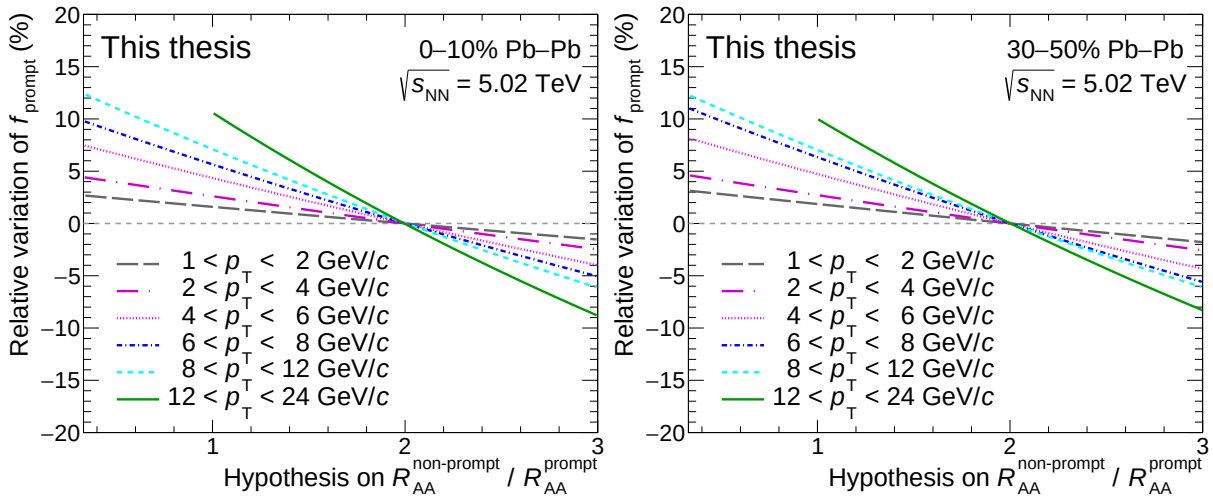


Figure 7.8: Relative variation of the Λ_c^+ f_{prompt} fractions coming from the variations on the $R_{AA}^{non-prompt}$ hypothesis for central (left) and mid-central (right) Pb–Pb collisions. The assigned systematic uncertainties are the minimal and maximal values for each shown p_T interval (i.e. at 1/3(1) and 3, respectively).

was considered, wider with respect to that used for non-strange D mesons [220] to cover possible yet unmeasured differences between the modification of charm- and beauty-baryon production in the QCD medium. For $p_T > 12$ GeV/c, the same variations as for the non-strange D mesons ($1 < R_{AA}^{non-prompt}/R_{AA}^{prompt} < 3$) were used instead, because the hadronisation of charm quarks at high p_T is expected to be dominated by the fragmentation mechanism. The variations for the non-strange D mesons are defined by comparing the R_{AA} measurements and theoretical predictions of prompt D mesons [144] with that of non-prompt J/ψ decays [247, 248]. The assigned uncertainty values for the Λ_c^+ baryons coming from the $R_{AA}^{non-prompt}$ hypothesis are summarised in Fig. 7.8 and range from $^{+2.7}_{-1.5}$% at low p_T to $^{+10.5}_{-8.8}$% at high p_T.

7.5.6 Other sources of systematic uncertainties

Systematic effects due to the uncertainty on the fraction of the hadronic cross section used in the Glauber fit (used to determine the centrality classes, as explained in Sec. 1.4) were studied as well. It was estimated from the variation of the D^0-meson dN/dp_T when the limits of the centrality classes are shifted by 1.1% (e.g. shifted from 30–50% to 30.3–50.6% and 29.7–49.5%). The resulting uncertainty is common for all charm hadron species and p_T intervals, and was found to be less than 0.1% for central and about 2% for mid-central Pb–Pb collisions.

Because of the limited statistics of the simulated samples, the relative statistical uncertainty on the selection efficiency is assigned as systematic uncertainty. It varies from 1% at low p_T to 4% at high p_T.

7.5.7 Summary

The contributions of these different sources of uncertainties ($\delta_{1-6,14}$) were summed in quadrature to obtain the total systematic uncertainty in each p_T interval, which varies from 17%–23% and

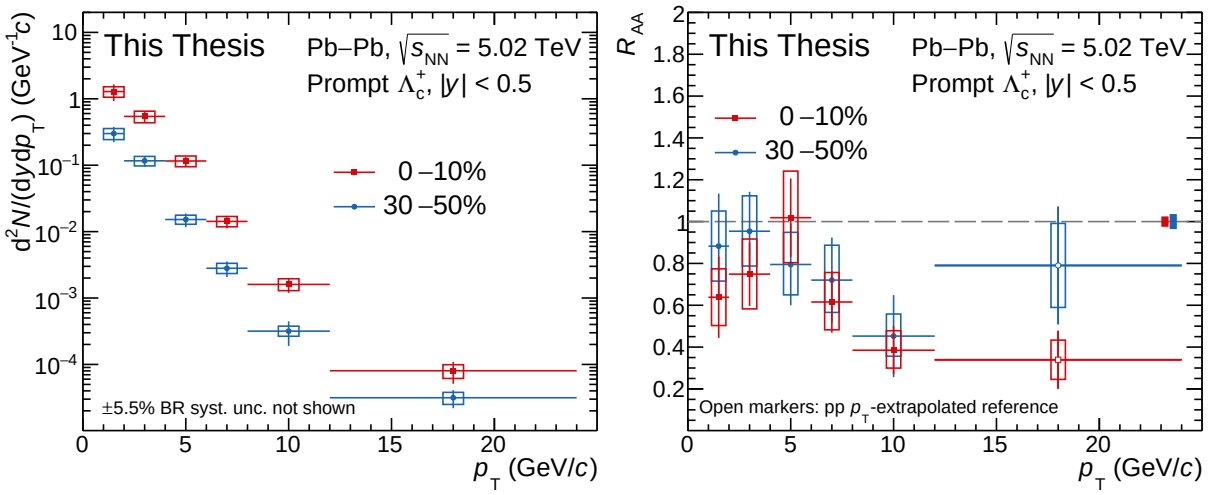


Figure 7.9: Left: The p_T-differential production yields of prompt Λ_c^+ in central (0–10%) and mid-central (30–50%) Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV. Right: The nuclear modification factor of prompt Λ_c^+ in central (0–10%) and mid-central (30–50%) Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV. The normalisation uncertainties are shown as a shaded box around unity.

15%–20% for the central and mid-central measurements. The sources due to the uncertainty on the branching ratio (δ_7; of 5.5% [16]) and centrality interval limits (δ_{13}) are considered in an additional global normalisation uncertainty. The uncertainty of the pp normalisation uncertainty [102] and of the average nuclear overlap function [37] is included in the global normalisation uncertainty when considering the nuclear modification factor R_{AA}.

For the Λ_c^+/D^0 ratio, the Λ_c^+ and D^0 uncertainties were considered as uncorrelated except for the tracking (δ_2) and the feed-down (δ_6) contribution, which partially cancel in the ratio, and the source due to the centrality intervals (δ_{13}), which fully cancels out. For the R_{AA}, the pp and Pb–Pb uncertainties were added in quadrature except for the branching ratio uncertainty (δ_7; due to the different considered decay modes) and the feed-down contribution (δ_6), which partially cancel out. When comparing Λ_c^+ (or Λ_c^+/D^0) production in Pb–Pb to pp collisions, all uncertainties except the branching ratio uncertainty (δ_7) were considered to be uncorrelated between the different collision systems. Furthermore, we consider the statistical and raw-yield extraction uncertainties ($\delta_{stat,\,1,\,14}$) as uncorrelated when rebinning or integrating over p_T.

7.6 Results

The p_T-differential production yields of prompt Λ_c^+ baryons, measured in the rapidity interval $|y| < 0.5$, are shown in Fig. 7.9 (left panel). Thanks to centrality triggers adopted by the ALICE Collaboration for the Pb–Pb 2018 data sample, the here presented measurement of the Λ_c^+ dN/dp_T has a wider p_T coverage, and is, in addition, more differential in p_T and centrality with respect to previous measurements in A–A collisions [104, 199, 239]. In the right panel of Fig. 7.9, the nuclear modification factor, R_{AA}, is presented. For $p_T < 6$ GeV/c, the R_{AA} is found to be compatible with unity, within uncertainties, for both central and mid-central collisions, while a suppression due to the charm-quark energy loss in the QGP is observed for higher p_T.

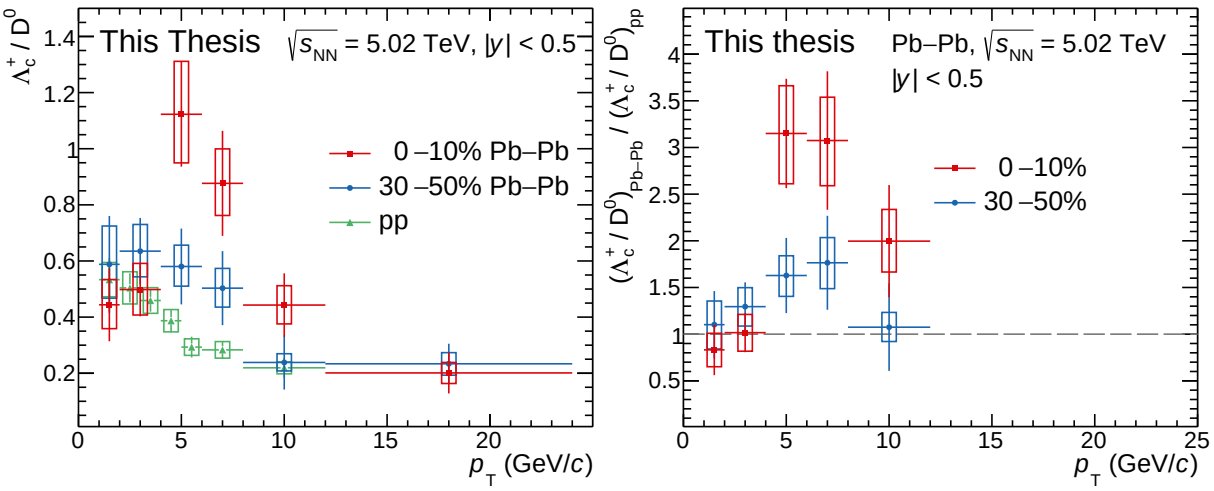


Figure 7.10: Left: The Λ_c^+/D^0 ratio in central and mid-central Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV compared to the results obtained from pp collisions [102, 103]. Right: The Λ_c^+/D^0 double-ratios between the two Pb–Pb centrality classes and pp collisions.

In the left panel of Fig. 7.10, the ratio between the production yields of Λ_c^+ baryons and D^0 mesons [220] measured in the same centrality intervals are presented and compared to the pp measurement at the same collision energy [102, 103]. The ratios increase from pp to mid-central and central Pb–Pb collisions for $4 < p_T < 8$ GeV/c with a significance of 2.0σ and 3.7σ, respectively. This significance level is estimated via the double Λ_c^+/D^0 ratio between Pb–Pb and pp collisions, rebinning the pp measurements to the Λ_c^+ p_T intervals in Pb–Pb, as presented in the right panel of Fig. 7.10. The probability to observe the measured double-ratio to be larger than one in $4 < p_T < 8$ GeV/c was tested by taking into account the statistical and systematic uncertainties. With the aim of investigating the least favourable case, the measured values in the considered p_T intervals were shifted down by one standard deviation considering the sources of systematic uncertainties correlated with p_T that do not cancel out in the double ratio ($\delta_{2,3,4,5}$). The same calculation was performed for the full measured $1 < p_T < 12$ GeV/c range, leading to a significance level of 2.9σ and 4.4σ for mid-central and central collisions, respectively.

The Λ_c^+/D^0 ratios are shown in Fig. 7.11 as a function of $\langle dN_{ch}/d\eta \rangle$ in five different p_T intervals from 2 to 24 GeV/c. The measured ratios in the different centrality classes for Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV are compared to the midrapidity multiplicity event classes in pp collisions at $\sqrt{s}$ = 13 TeV [238] and to minimum-bias pp and p–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV [102]. There is a clear sign of a continuous rising trend with $\langle dN_{ch}/d\eta \rangle$, that extends to Pb–Pb multiplicities, in the Λ_c^+/D^0 ratios at intermediate p_T[4].

7.6.1 Theory comparison

Figure 7.12 compares the p_T-differential Λ_c^+/D^0 ratios measured in central and mid-central collisions with different theoretical predictions: Catania [160], TAMU [162], and the GSI–Heidelberg statistical hadronisation model (SHMc) [246]. The predictions of Catania and TAMU for pp

[4] A trend in a given p_T interval could also be due to a modification of the p_T shape, e.g. due to radial flow. It is better to look at the trend with $\langle dN_{ch}/d\eta \rangle$ for the p_T-integrated Λ_c^+/D^0 ratios, which we will discuss later.

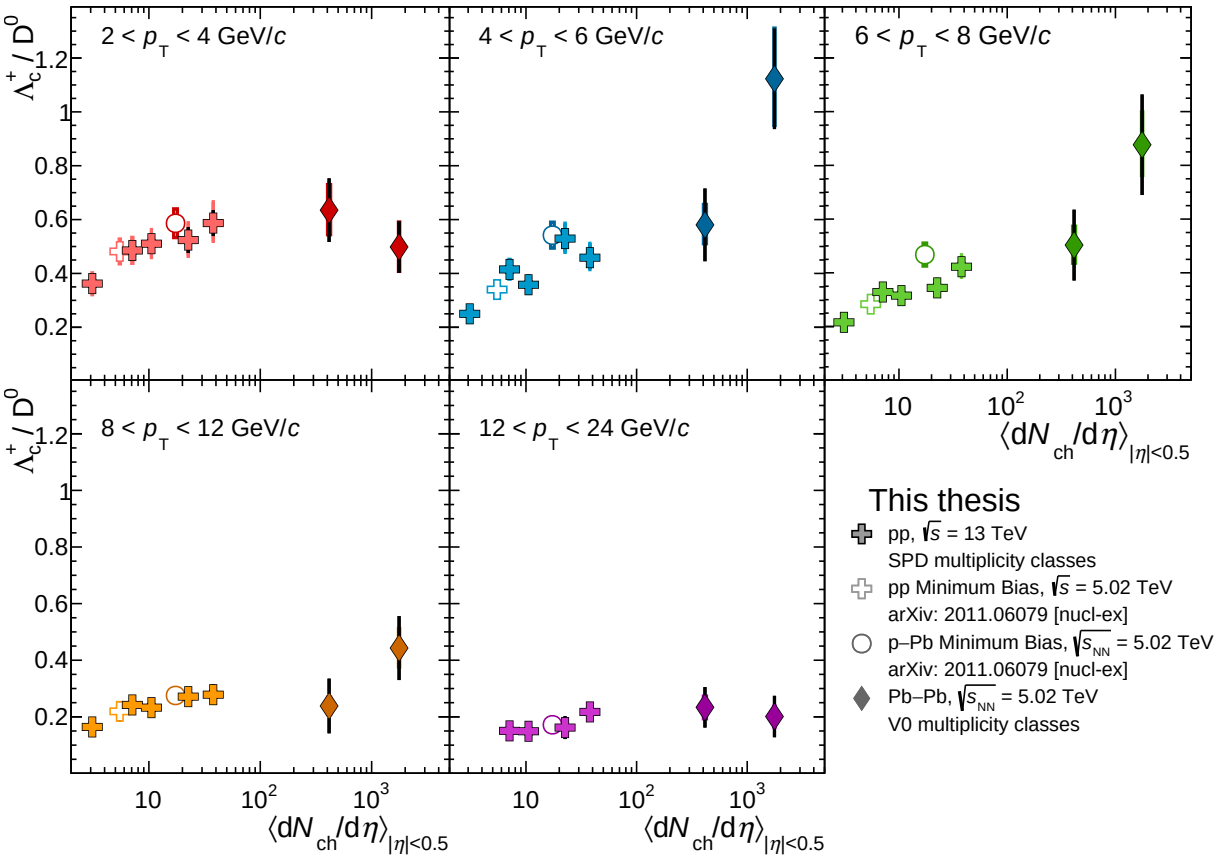


Figure 7.11: The Λ_c^+/D^0 ratios, as a function of primary charged particles per unit of pseudorapidity in pp at $\sqrt{s} = 5.02$ TeV [102] and $\sqrt{s} = 13$ TeV [238], p–Pb at $\sqrt{s} = 5.02$ TeV [102], and Pb–Pb at $\sqrt{s_{NN}} = 5.02$ TeV in five different p_T intervals from 2 to 24 GeV/c.

collisions [111, 112] are also compared with the existing measurement in pp collisions [102, 103]. The Catania model assumes that a QGP is formed in both pp and Pb–Pb collisions. In Pb–Pb collisions heavy-quark transport is implemented via the Boltzmann equation, and in both pp and Pb–Pb collisions hadronisation occurs either via coalescence, implemented through the Wigner formalism, or via fragmentation in case the quarks do not undergo coalescence. The TAMU model describes charm-quark transport in an expanding medium with the Langevin equation and hadronisation proceeds primarily via coalescence, implemented with a resonance recombination model [161]. Left-over charm quarks not undergoing coalescence are hadronised via fragmentation. In pp collisions, the charm-hadron abundances are instead determined with a statistical hadronisation approach. In both collision systems, the underlying charm-baryon spectrum includes yet-unobserved excited states [16] predicted by the relativistic quark model [167] and lattice QCD [111]. Finally, for the SHMc predictions the charm-hadron p_T spectra are modelled within a core-corona approach. The core contribution represents the central region of the colliding nuclei where charm quarks achieve local thermal equilibrium in a hydrodynamically expanding QGP. The charm-hadron spectra in the corona contribution are, instead, parameterised from measurements in pp collisions. The p_T-spectra modification due to resonance decays is computed using the FastReso package [249]. More details on the different models can

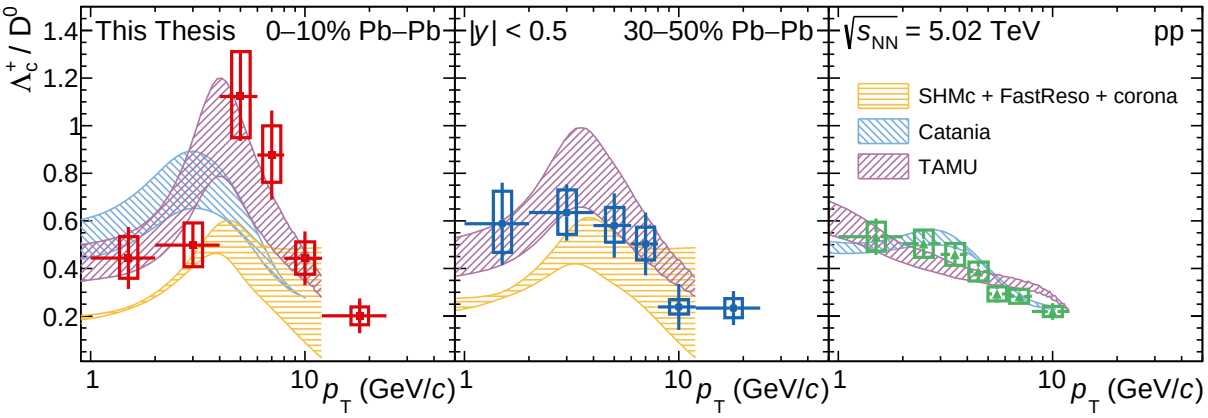


Figure 7.12: The Λ_c^+/D^0 yield ratio as a function of p_T in 0–10% (left) and 30–50% (middle) Pb–Pb and pp (right) collisions at $\sqrt{s_{NN}}$ = 5.02 TeV compared with predictions of different theoretical calculations [111, 112, 160, 162, 246, 249].

be found in the relevant references, since they are beyond the scope of this dissertation.

The SHMc describes the Λ_c^+/D^0 ratio in mid-central collisions, but underpredicts the ratio in $4 < p_T < 8$ GeV/c in central collisions. The prediction of the Catania model in central collisions overestimate (underestimate) the Λ_c^+/D^0 ratio at low (intermediate) p_T. The TAMU predictions reproduce the magnitude and shape of the Λ_c^+/D^0 ratios. While both coalescence model calculations are able to describe the Λ_c^+/D^0 ratio in Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV in the 10–80% centrality interval [239], the TAMU model reproduces significantly better the data in central Pb–Pb collisions.

The same models provide also a prediction for the nuclear modification factor for prompt Λ_c^+ in both centrality intervals at $\sqrt{s_{NN}} = 5.02$ TeV, as presented in Fig. 7.13. For the description of the charm p_T spectra in pp collisions, the Catania [112][5] and TAMU [111] models use as starting point both FONLL calculations [83, 87], while the SHMc [246] relies on a parametrisation of the Λ_c^+ production measurement in pp [102] via a three-parameter fit. The TAMU prediction reproduces the R_{AA} in both centrality classes at all p_T, while the Catania model agrees relatively well for $p_T > 2$ GeV/c. The SHMc significantly underpredicts the measured nuclear modification factors. As already noted by the authors in Ref. [246], this suggests that the corona description is somewhat schematic and could be further optimised.

7.6.2 Transverse momentum-integrated production yields and ratios

The Λ_c^+ production yield for $p_T > 0$ was estimated by summing up the p_T-differential yields, extrapolating the Λ_c^+/D^0 ratio down to $p_T = 0$, and exploiting the measurement of the D^0 yield in the interval $0 < p_T < 1$ GeV/c [220]. The p_T-differential Λ_c^+/D^0 ratio was interpolated using the shape predicted by TAMU [162], Catania [160] (not available for 30–50%), SHMc [246], and blast-wave [38] calculations. The shape from TAMU was chosen for the central value based on

[5]Note that the authors of the Catania model updated their predictions for the Λ_c^+/D^0 ratio (including a theoretical uncertainty band), while for the R_{AA} a prediction of an older version of their model is shown.

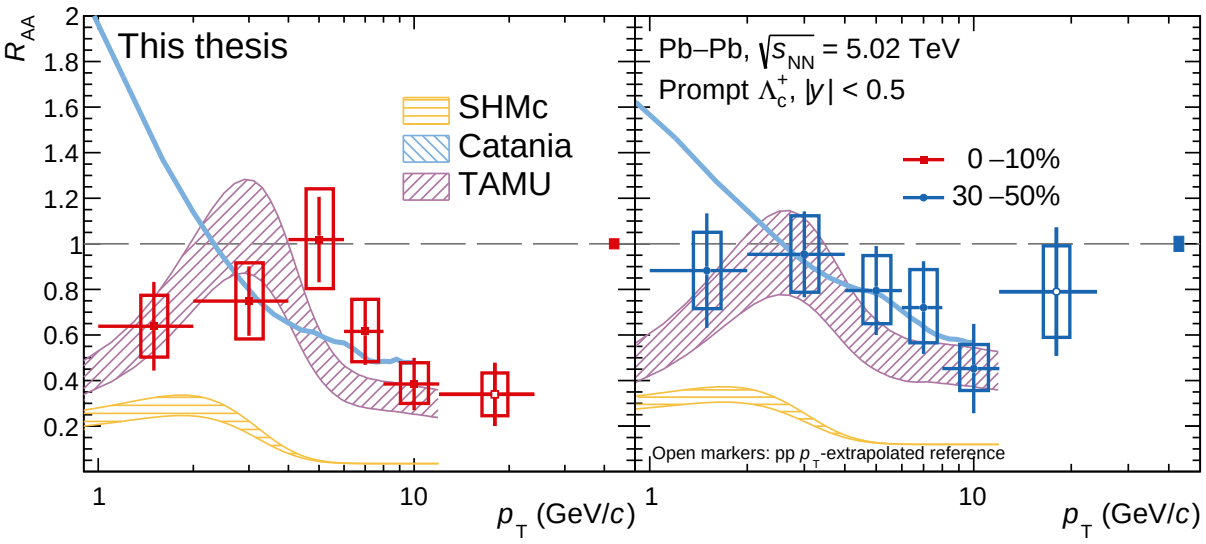


Figure 7.13: The p_T-differential nuclear modification factor of prompt Λ_c^+ in 0–10% (left) and 30–50% (right) Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV compared to the predictions of different theoretical calculations [160, 162, 246]. The normalisation uncertainties are shown as a shaded box around unity.

the χ^2/ndf of the fits. The fits, for which the normalisation is left as free parameter, are shown in Fig. 7.14 for both the central and mid-central centrality intervals. The extrapolated Λ_c^+/D^0 ratio for $p_T < 1$ GeV/c was then multiplied by the dN/dp_T of D^0 mesons in $0 < p_T < 1$ GeV/c. The obtained extrapolated Λ_c^+ yields amount to approximate 20% and 19% of the total production yields — calculated as the sum of the extrapolated and measured yields — for 0–10% and 30–50%, respectively. The difference between the obtained yields with the other fit functions was considered in the systematic uncertainty due to the extrapolation. The total uncertainty on the measured dN/dp_T of D^0 mesons in $0 < p_T < 1$ GeV/c enters the extrapolation uncertainty as well.

The results for the prompt Λ_c^+ production yields per unit of rapidity in $|y| < 0.5$ are

$$dN/dy\,|_{0–10\%} = 3.28 \pm 0.42\,(\text{stat}) \pm 0.44\,(\text{syst}) \pm 0.16\,(\text{BR})\,^{+0.46}_{-0.29}\,(\text{extr});$$
$$dN/dy\,|_{30–50\%} = 0.70 \pm 0.09\,(\text{stat}) \pm 0.09\,(\text{syst}) \pm 0.04\,(\text{BR})\,^{+0.07}_{-0.05}\,(\text{extr})\,.$$

The SHMc [246] predicts lower values, $dN/dy = 1.55 \pm 0.23$ and $dN/dy = 0.316 \pm 0.036$, respectively. The difference between the measured and predicted dN/dy, presented in Fig. 7.15, is 2.6σ and 2.8σ considering the total uncertainties on the measurement for central and mid-central Pb–Pb collisions, respectively.

The measured p_T-integrated Λ_c^+/D^0 ratios, obtained dividing the p_T-integrated Λ_c^+ and D^0 yields [220], are $0.48^{+0.13}_{-0.12}$ for 0–10% and $0.55^{+0.14}_{-0.13}$ for 30–50% centrality classes, taking into account the correlation between the measured and extrapolated uncertainties, hence only a total uncertainty is reported. The extrapolation uncertainty is divided into a correlated part (total extrapolation uncertainty considering only the TAMU shape) and uncorrelated part (total extrapolation uncertainty of all shapes minus the one of the TAMU model) with respect to the

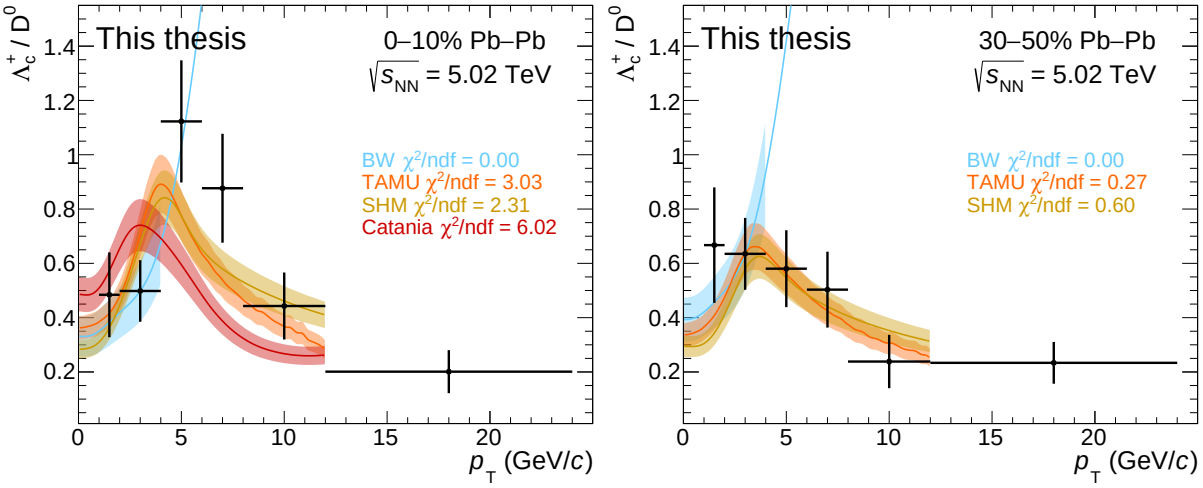


Figure 7.14: Fits of the Λ_c^+/D^0 ratio in central (left) and mid-central (right) Pb–Pb collisions using the p_T shape of TAMU [162], Catania [160], SHMc [246], and blast-wave [38] calculations. The normalisation is left free, and the blast-wave calculations are restricted to $p_T < 4$ GeV/c (the central line is continued for $p_T > 4$ GeV/c just for illustration). The shown uncertainty bands are obtained considering the statistical and p_T-uncorrelated systematic uncertainties on the measured data points (black uncertainty bars) for the fits. The contribution of the p_T-correlated systematic uncertainties (not shown) is taken into account by repeating the fit shifting up and down the black data points with 1σ.

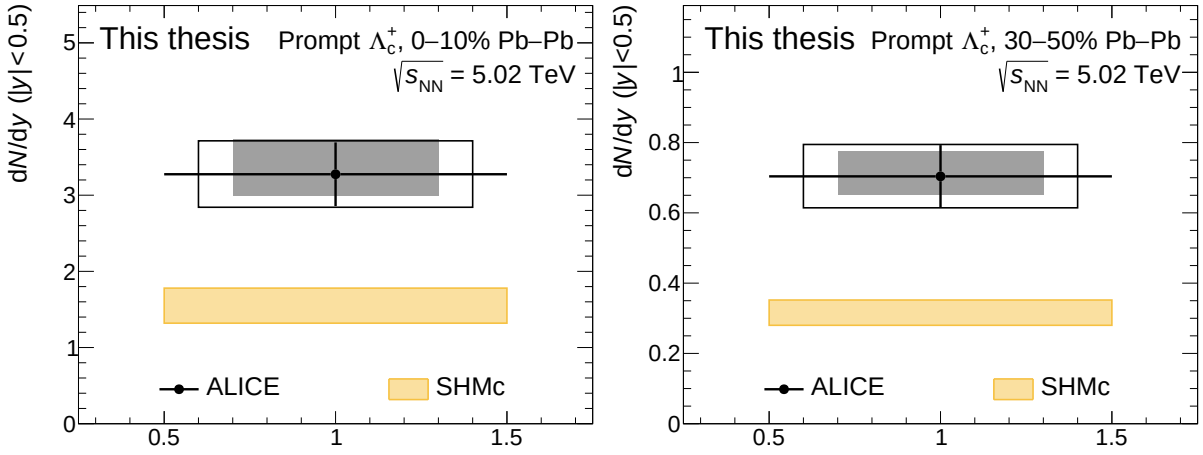


Figure 7.15: Transverse momentum-integrated production yield for prompt Λ_c^+ in 0–10% (left) and 30–50% (right) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in $|y| < 0.5$ compared to predictions from the SHMc [246]. The statistical (bars), systematic (boxes), and extrapolation (shaded area) uncertainties are shown.

measured uncertainties. The uncorrelated part is summed in quadrature with the measured uncertainties, while the correlated part is added linearly. The different sources of systematic uncertainty are propagated as usual for the Λ_c^+/D^0 ratio. The p_T-integrated Λ_c^+/D^0 ratios in both centrality intervals in Pb–Pb are compatible with the results in minimum-bias pp and p–Pb collisions [102] within one standard deviation of the combined uncertainties, as presented in Fig. 7.16.

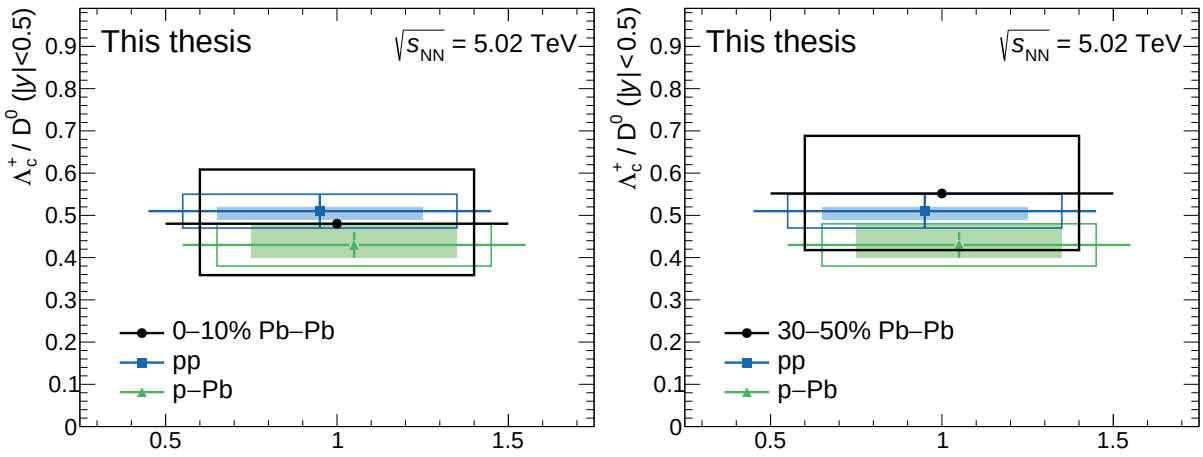


Figure 7.16: Transverse momentum-integrated Λ_c^+/D^0 yield ratios in 0–10% (left) and 30–50% (right) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in $|y| < 0.5$ compared to the results in minimum-bias pp and p–Pb collisions [102]. For Pb–Pb, only the total uncertainties are shown (box), while for pp and p–Pb a separation is made between statistical (bars), systematic (boxes), and extrapolation (shaded area) uncertainties.

7.7 Discussion

In this chapter, the measurement of prompt Λ_c^+ baryons in central and mid-central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV was presented. The measurement has a wider p_T coverage and is more differential in p_T and centrality with respect to previous measurements in A–A collisions [104, 199, 239], providing an important further glance to charm-quark hadronisation in heavy-ion collisions. As for the previous chapters, I will use this section to provide (my view on the) physics interpretation of these interesting results. There is a close connection with the discussion of Chapter 6.

Let me first quickly recap the theoretical interest of this measurement. The Λ_c^+/D^0 ratio has been proposed several times as an observable that should be enhanced (with respect to an only-fragmentation picture) by theoretical models that include modified hadronisation by the presence of a deconfined medium via quark coalescence or resonance recombination. The authors of Refs. [160, 162, 164, 165, 237] expect this to be an overall, p_T-integrated, enhancement, i.e. no depletion of the ratio is expected at low p_T to cancel a possible enhancement at intermediate p_T[6]. In these models, the enhancement follows from arguments like a higher coalescence probability for baryons [160, 162, 165, 237], the existence of diquark states [164, 165], or simply due to the larger values for the baryo-chemical potential [250]. A further enhancement is expected in Ref. [162] by an augmented set of resonance charm baryon states, but its impact should be similar for pp and Pb–Pb systems. In addition, a redistribution over p_T, leading to an enhancement at intermediate p_T and depletion at low p_T, can be expected for the Λ_c^+/D^0 ratio following observations for the light-flavour baryon-to-meson ratios [50, 243]. This increase is attributed to the presence of radial flow in heavy-ion collisions, which affects heavier particles

[6]Note that for the TAMU model in Fig. 7.12 this does not seem to be true, but one should keep in mind that their pp and Pb–Pb models are intrinsically different. Based on Fig. 5 in Ref. [162], it seems to be the case that also the TAMU Pb–Pb model expects a p_T-integrated enhancement for the Λ_c^+/D^0 ratio.

more than lighter ones[7], and to hadronisation via coalescence, which will push baryons to higher momenta with respect to mesons because of the additional constituent quark. These — by now "trivial" — effects affect the heavy-flavour baryon-to-meson ratios as well, and I suppose that they are taken into account in the Λ_c^+/D^0 model predictions mentioned above, although this is not highlighted in the manuscripts. The effect of hadronic diffusion on the Λ_c^+ and D^0 yields is also not clearly addressed in the current literature or with inconsistent conclusions [244, 245].

As predicted for intermediate p_T, the Λ_c^+/D^0 ratios indeed increase from pp to mid-central and central Pb–Pb collisions for $4 < p_T < 8$ GeV/c with a significance of 2.0σ and 3.7σ, respectively. The measurements are also reasonably described by theoretical calculations that include both coalescence and fragmentation processes when describing the hadronisation of heavy flavours in the QGP. On the contrary, however, in Fig. 7.16 (and Fig. 6.19), we have seen that the p_T-integrated Λ_c^+/D^0 ratios do not show an enhancement with increasing multiplicity from low-multiplicity pp to central Pb–Pb collisions. Following the line of thoughts of the heavy-flavour coalescence models above, this may suggest that the charm hadronisation and Λ_c^+ production do not differ significantly in Pb–Pb and pp collisions, and that the observed baryon enhancement at intermediate p_T is predominantly caused by a different modification of the hadron spectra due to interactions in the hadronic phase. An alternative explanation is that charm hadronisation is indeed modified in Pb–Pb collisions, but the different kinematic distributions of light and charm quarks at hadronisation time in pp and Pb–Pb collisions modulate the amplitude and p_T dependence of the Λ_c^+/D^0 ratio.

These last two sentences are kept rather general to cover multiple possibilities that can explain these two observations. In the first case, one can think about the effect that processes like the collective hydrodynamic expansion of the system and hadronic rescattering between the chemical and thermal freeze out can have. An example would be radial flow, developed in either the partonic or hadronic phase, which "pushes" the heavier Λ_c^+ baryons towards higher $\langle p_T \rangle$ with respect to D^0 mesons. This would change the p_T-differential Λ_c^+/D^0, but leave the p_T-integrated values unaltered. I explicitly specified partonic and hadronic radial flow, since I believe the radial flow from the partonic phase cannot cause a "radial-flow bump" in the Λ_c^+/D^0 ratio in case of transport models like Catania, because the momentum spectrum of hadrons at the hadronisation temperature is fully determined from the quark's kinematic distributions (see for example Eq. 1 of Ref. [160]), i.e. the mass of the hadron does not enter these type of calculations and the difference in the sum of the single quark's masses for Λ_c^+ and D^0 is negligible. Any modification of the underlying kinematic distributions of quarks caused by radial flow (or any other nuclear-medium effect) should thus cancel[8] in case of charm-hadron

[7] The expected effect of radial flow on the p_T-shape of the Λ_c^+/D^0 ratio can be illustrated with a blast-wave calculation [38]. This has been shown in Fig. 7.14, since it was one of the shapes considered in the Λ_c^+ yield extrapolation procedure.

[8] The modified kinematic distributions may of course alter the charm-meson and -baryon coalescence probabilities in an independent way, and thereby modify the p_T-shape of the Λ_c^+/D^0 ratio. However, this effect is typically not meant when saying "radial flow modifies the p_T-shape of the Λ_c^+/D^0 ratio from pp to Pb–Pb collisions". It is equivalent to stating that the prediction of a coalescence model is sensitive to the hydrodynamic scenario used, which is a rather redundant statement.

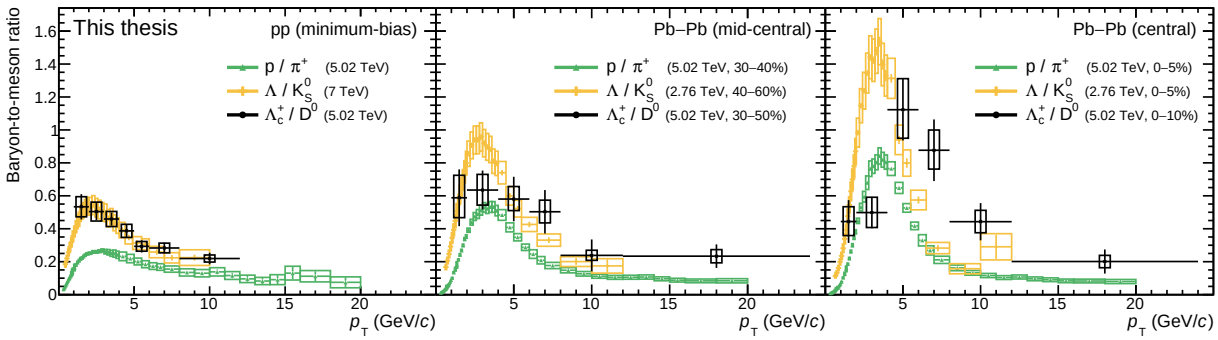


Figure 7.17: The baryon-to-meson ratios in the light-flavour sector, based on measurements from Ref. [50, 243], and charm sector [102] measured in minimum-bias pp (left), mid-central Pb–Pb (middle), and central Pb–Pb (right) collisions.

ratios. Therefore, only radial flow developed in the hadronic phase could cause a "radial flow bump" at intermediate p_T. In case of hydrodynamic or statistical hadronisation models, which typically use the Cooper–Frye formula [251], the mass of the hadron enters in the hadronisation calculation and radial flow from the partonic phase can certainly cause a p_T modification in the Λ_c^+/D^0 ratio. This is also clearly visible in the Λ_c^+/D^0 prediction of the SHMc model [246].

The second part of the above sentence refers to unconventional implementations of modified charm hadronisation in presence of a deconfined medium in Pb–Pb collisions. I wrote "unconventional" here, since the implementations of quark coalescence and resonance recombination we saw throughout this chapter, all predict an overall enhancement of the Λ_c^+/D^0 ratio (i.e. an enhancement also in the p_T-integrated ratio) and should thus actually be "ruled out" by the data. However, one can imagine some modified implementation of the same coalescence/recombination idea, with a different dependency on, for example, the relative velocity between light and charm quarks for the meson and baryon recombination probability equations. This idea, which comes with a disclaimer[9], should be able to explain the observations also in case such modified hadronisation mechanisms are active in pp collisions, since the average light-quark velocity is higher in Pb–Pb than in pp systems.

Let us now go back to the point I ended the previous discussion with: the similarities between the heavy- and light-flavour sector concerning the baryon-to-meson ratios. We have seen that there are remarkable agreements between the Λ_c^+/D^0 and Λ/K_S^0 ratios in minimum-bias pp and p–Pb events [102], as well as for multiplicity event classes in pp [238], both in terms of shape and magnitude. Also in Pb–Pb collisions, the Λ_c^+/D^0 ratio as function of p_T follows a trend qualitatively similar to that of the p/π [50] and Λ/K_S^0 [243] ratios. As one can see in Fig. 7.17, the agreement between the Λ_c^+/D^0 and Λ/K_S^0 ratios is smaller for Pb–Pb systems though, but all three type of ratios still present a qualitatively similar enhancement at intermediate p_T and a shift of the maximum value towards higher p_T values with increasing multiplicity from pp

[9]Let me play the card of an experimentalist here: theoretically I cannot motivate this idea of a different dependence on the kinematic distributions of light and charm quarks. Actually, the resonance recombination implementation of the TAMU model [162] has a $v_{rel}(\vec{p_q}, \vec{p_Q})$ dependency (following from the Boltzmann equation which they use as starting point) that predicts an overall Λ_c^+/D^0 enhancement. This section should thus be read as an experimentalist trying to start a theory discussion.

to Pb–Pb collisions. The same effect we saw in Fig. 7.11, where the Λ_c^+/D^0 ratio smoothly increased as function of $\langle dN_{ch}/d\eta \rangle$ in small p_T intervals from small to large systems. Similar behaviour, including the accompanying decreasing trend with multiplicity for low p_T, is also seen in the Λ/K_S^0 ratio [225, 243]. This depicts itself in a basically constant p_T-integrated Λ/K_S^0 ratio from pp to Pb–Pb multiplicities. All these similarities between the light- and heavy-flavour baryon-to-meson ratios — for which the mass difference is not too different — might hint to a similar origin for the observations reported in this chapter for the Λ_c^+/D^0 ratio. The discussion on how to interpret these observations in the light-flavour sector is focusing on a radial boost from a hydrodynamic origin picture. As discussed above, we might have to consider this as explanation for these heavy-flavour results as well, i.e. there is no p_T-integrated enhancement in the Λ_c^+/D^0 ratio.

Another possibility is that we just cannot appreciate the enhancement in the p_T-integrated Λ_c^+/D^0 ratios because of the large uncertainties. More precise measurements in the future are fundamental. This measurement is therefore one of ALICE's main goals for LHC Run 3 and 4, where we expect an increase of the integrated luminosity by about a factor 50 and an improved tracking resolution by a factor 3 to 6 (see Chapter 8 for more details). In this sense, I would also like to highlight a future $\Sigma_c^{0,+,++}/\Lambda_c^+$ measurement, which should actually be a better signal for the existence of diquarks in the QGP [252].

I will end this chapter with the charm family portrait in heavy-ion collisions, via the nuclear modification factor, shown in Fig. 7.18. I do not believe that the R_{AA} is the interesting observable for the Λ_c^+ baryon (more physics can be obtained for the Λ_c^+/D^0 ratio and D^0 R_{AA}), but given the emphasis on this variable in the field of heavy-ion physics, the new comparison among the charm hadrons earns a place in this manuscript.

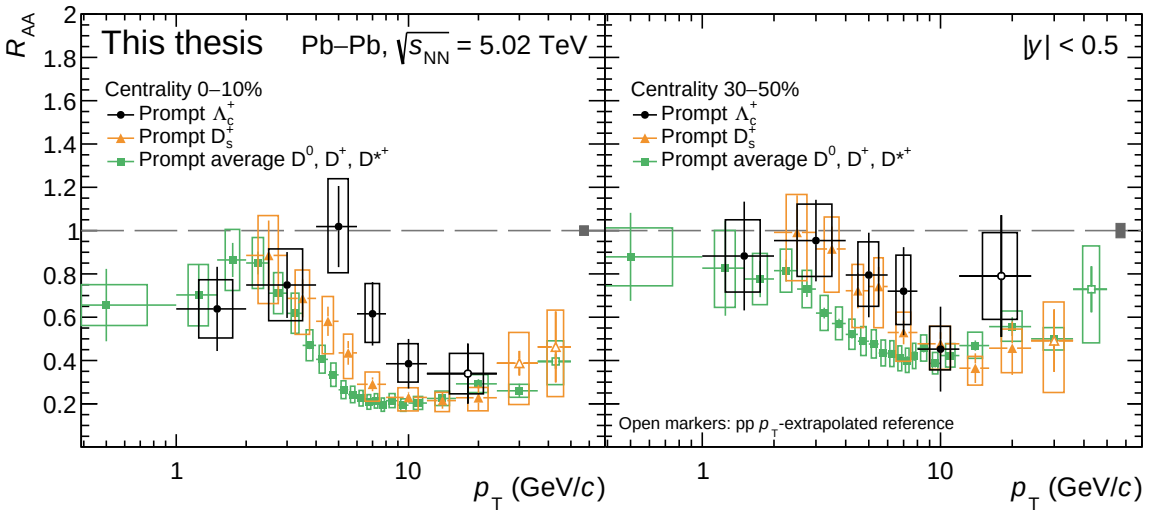


Figure 7.18: Right: The nuclear modification factor of prompt Λ_c^+ baryons, D_s^+ mesons [163], and the average of the non-strange D mesons [220] in central (left) and mid-central (right) Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV. The normalisation uncertainties are shown as a shaded box around unity.

CHAPTER 8

Beauty-hadron production with the ALICE Upgrade

Abstract – With the installation of a new inner barrel for the ALICE Inner Tracking System for LHC Run 4 (2027–2030), several measurements related to beauty hadronisation will be rendered possible. This upgrade project foresees the construction of three new fully cylindrical ultra-thin layers of monolithic active pixel sensors, which is expected to largely improve the pointing resolution of the tracking, in particular at low momenta, with respect to the ALICE Inner Tracking System installed during the second long LHC shutdown. With the expected performance of this upgrade project, the measurement of the $\mathrm{B_s^0}$ meson in Pb–Pb collisions will get feasible and complement, and extend to lower p_T, the existing measurement of the CMS Collaboration and their projection for the High Luminosity LHC era. Together with measurements of Λ_b^0 baryons and non-prompt $\mathrm{D_s^+}$ mesons, which can be achieved with good precision, this upgrade will make it possible to extend the in-medium heavy-flavour energy loss and hadronisation picture to the beauty sector.

Manuscript for publication in preparation

8.1 Introduction

Over the years, many experimental observations from experiments at RHIC and LHC showed evidence that heavy quarks (especially charm) interact strongly with the QGP. These interactions will lead to the thermalisation of low-momentum heavy quarks, which then take part in the expansion and hadronisation of the medium. As discussed extensively in Chapter 2, heavy-flavour hadrons will provide information on all stages of the evolution of the system and, in addition, uniquely probe the quark-mass dependence of the different mechanisms believed to take place in the QGP. While data are becoming more and more precise in the charm sector (see for example the previous three chapters), measurements of beauty hadrons in heavy-ion collisions are still scarce. From a theoretical point of view, however, beauty hadrons represent a cleaner probe than charm hadrons to study the QGP [219]. The LHC Run 3 and 4 will open a new precision era for these heavy-flavour measurements which will likely contribute heavily to our understanding of heavy-flavour physics.

In the second long LHC shutdown (2019–2021), the machine and the tracking detectors of the different LHC experiments got major upgrades, which will result in higher accumulated luminosity as well as an improved precision. These upgrades will already make it possible to further quantify the properties of the QGP with heavy-flavour probes, and the expected physics performances of the four different LHC experiments is discussed extensively in Ref. [253]. The major — for heavy-flavour probes relevant — upgrade of the ALICE experiment includes the installation of a new Inner Tracking System (ITS2) based on monolithic active pixel sensors (MAPS) [254]. The ITS2 will enable the measurement of several heavy-flavour hadrons with unprecedented precision, as discussed in the technical design report [254]. The rapid development of the MAPS technology, however, now allows for a significant further performance increase. Another upgrade of the inner barrel of the ALICE Inner Tracking System (ITS3) is therefore intended during the third long LHC shutdown, which will bring another set of crucial observables like charm-strange baryons and beauty-strange mesons in heavy-ion collisions into reach.

This chapter is organised as follows. First, we discuss briefly the final and envisaged detector layout of, respectively, the ITS2 and ITS3 detector and the type of simulations used to access the resulting performances. Then we discuss in detail the analysis performed to extract the expected physics performances of the $\mathrm{B_s^0}$ meson, projected on the R_{AA} observable, with both ITS setups in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.5$ TeV. The chapter ends with a general discussion, comparing as well to the expected performance on the measurement of non-prompt $\mathrm{D_s^+}$ mesons, another interesting way to access beauty hadronisation in presence of a QGP. For a more detailed discussion of the other core motivations for the installation of the ITS3 detector, e.g. the measurement of c-deuteron or the new concept of strangeness tracking, we refer to Ref. [255].

8

8.2 Upgrade specifications

Details on the ALICE Inner Tracking System installed during the first two LHC runs (ITS1) were given in Sec. 3.2.1. It consisted out of six cylindrical layers of silicon detectors located at

radii between 39 and 430 mm. The inner radius was limited by the radius of the beampipe, while the outer radius was determined by the requirement to match tracks to those in the TPC. Overall the boundaries on the number, position, and segmentation of the different layers were set by the technological limitations and available funds at that point. Besides the insufficient impact parameter resolution to perform certain more exotic analyses, a crucial limitation of the ITS1 detector was its read-out rate capabilities that could run at a maximum of 1 kHz. In addition, it was basically impossible to access the ITS1 detector for maintenance during the LHC shutdowns, leading to many inactive channels over the years.

The installation of the ITS2 will ensure an improved performance in many respects. First of all, the first detection layer will be located closer to the beampipe (from 23 to 18 mm), which will also have a smaller radius (18.2 to 16 mm) and be less thick (800 to 500 µm). In addition, the ITS2 has an extra seventh layer, has a significantly reduced material budget (from 1.14 to $0.35\%X_0$ for the first layer), a finer granularity, and a maximum readout above the expected 50 kHz Pb–Pb interaction rate for LHC Run 3 and 4. These upgrades will improve the ability of the ALICE detector to distinguish secondary decay vertices from the primary vertex and track very low transverse-momentum particles, which allows for the measurements of rare probes (like heavy-flavour hadrons) with outstanding precision down to $p_{\mathrm{T}} = 0$. This all was possible after significant R&D over many years that lead to the so-called ALPIDE chip, in which the sensor collecting the charge signal and the readout circuit are hosted in the same substrate of silicon. Each ALPIDE chip contains more than half a million pixels in an area of 15×30 mm^2. The ALPIDE chips are tiled up in staves to cover the full surface of the ITS. Also a new design for the ITS stave was developed, based on an ultralight carbon structure embedded with water cooling pipes (which operate at room temperature). In May 2021 the ITS2 installation was successfully completed, and it is — at the moment of writing — waiting to take its first real data (a pilot run of pp data at 900 GeV scheduled for October 2021).

Even though no real data is yet taken with the ITS2, ALICE is already thinking about an ITS3 detector, following from the wish to reduce the material thickness of the inner layers even further. The main technological advance that will be exploited in the ITS3 is called "stitching", that allows to construct a new generation of MAPS with an area up to the wafer size (of about 28×9.4 cm). One can thus achieve a fully cylindrical sensor of MAPS, that due to the small sensor thickness (30–40 µm) and the flexible nature of silicon can be bend to radii of 1.8 cm. The innermost detection layer can thus naturally surround the beampipe. In addition, the current design should be well within the range for which air cooling is applicable. It should thus be possible to build a cylindrical layer of silicon-only sensors without the need of using staves, significantly reducing the material thickness by about a factor six. The ITS3 will therefore further improve the tracking precision and efficiency at low transverse momentum. It is planned to be installed in the ALICE experiment during LS3 (2025–2027). For all details, we refer to the expression of interest [256] and the letter of intent [257].

The installation of ITS3 in LS3 will imply that it takes data starting from at least LHC Run 4. However, for an apples-to-apples comparison of the ITS2 and ITS3 detectors, the studies

presented in this chapter will be based on an expected luminosity of 10 nb^{-1} that is assumed exclusively for each detector.

8.2.1 Simulations with upgraded ITS setups

The expected performance of measurements with the ITS2 and ITS3 geometry are obtained via simulation studies. The studies were performed on central Pb–Pb events at $\sqrt{s_{\mathrm{NN}}} = 5.5$ TeV exploiting the HIJING 1.36 generator [190] for the underlying Pb–Pb event, PYTHIA 8.243 [95] to inject $\mathrm{c\overline{c}}$ (or $\mathrm{b\overline{b}}$) pp events on top, as well as further injection of single particles like the $\mathrm{B_s^0}$ meson to enhance the data sample, and GEANT3 [191] for the detector response. The simulations contain a detailed description of the ALICE apparatus with the ITS2 setup, and can thus be directly used to assess the expected performance for this case. To estimate the performance of the ITS3 layout, a fast simulation scheme called the *hybrid* method was used instead. Here, the full-simulation with the ITS2 setup have been corrected to mimic the ITS3 momentum and impact parameter resolutions. This is achieved by applying a scaling factor, evaluated as the ratio of the ITS2 and ITS3 resolutions, to the residuals of the track impact parameters and to the track transverse momentum with respect to their true MC values, e.g.

$$p_{\mathrm{T,\,reco.}}^{\mathrm{ITS3}} = p_{\mathrm{T,\,true}}^{\mathrm{MC}} + \frac{(p_{\mathrm{T,\,rec.}}^{\mathrm{ITS2}} - p_{\mathrm{T,\,true}}^{\mathrm{MC}})}{f_{\mathrm{scaling}}}. \tag{8.1}$$

It was demonstrated that this method reproduces the performance of a full simulation [254]. The scaling factor f_{scaling} was studied using a Fast Monte Carlo Tool (FMCT), which accounts for multiple scattering, secondary interactions, and detector occupancy, but ignores the energy loss of particles traversing the detector. The impact parameter and transverse momentum resolutions of both ITS setups, estimated with the FCMT and full simulations, are presented in Ref. [257].

8.3 Physics performance of $\mathrm{B_s^0}$ mesons

The production measurement of $\mathrm{B_s^0}$ mesons will shed additional light on the general physics cases of beauty-hadron production, but more importantly, it is an effective probe to study beauty-quark hadronisation mechanisms. As discussed extensively in Chapter 2, these hadronisation mechanisms belong to the non-perturbative domain of QCD and a first-principle description is still missing. In presence of any reservoir of partons, like in the QGP, vacuum fragmentation is no longer valid, and the hadronisation mechanism is expected to be modified. This will lead to a modification of the relative abundance of the various heavy-flavour hadron species produced. The most striking effects predicted are an enhancement of the baryon-to-meson ratio (see Chapter 7) and an increase of strange heavy-flavour hadrons, like $\mathrm{D_s^+}$ and $\mathrm{B_s^0}$, due to the strangeness-rich QGP [258, 259]. Several measurements of the production of $\mathrm{D_s^+}$ mesons in Pb–Pb collisions exist [144, 163, 260], while the beauty counterpart was only measured by the CMS Collaboration [261]. However, the precision of this latter measurement was not large enough to make any strong conclusions.

To estimate the expected performance of these upgrade projects, simulation studies for the measurement of the $\mathrm{B_s^0}$-meson production with the ITS2 and ITS3 geometry were carried out. The analysed decay channel, $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ ($\mathrm{BR} = (3.00 \pm 0.23) \cdot 10^{-3}$) with $\mathrm{D_s^-} \to \phi\pi^- \to \mathrm{K^+K^-}\pi^-$ ($\mathrm{BR} = (2.24 \pm 0.08) \cdot 10^{-2}$) [16], strongly benefits from the improved track spatial resolution of the Inner Tracking System upgrades because of the small mean proper decay length and large combinatorial background. As the decay branching ratio is very small, this analysis requires large integrated luminosities. Another interesting possibility would therefore be to study the $\mathrm{B_s^0}$ mesons via non-prompt $\mathrm{D_s^+}$ mesons, which will be briefly mentioned in the discussion of this chapter. With this method it is, however, not possible to distinguish between $\mathrm{D_s^+}$ mesons coming from different beauty-hadron species and also the p_T of the beauty-hadron cannot be reconstructed.

8.3.1 Selection criteria

The optimisation for the $\mathrm{B_s^0}$ analysis is performed using a machine learning approach as introduced in Sec. 4.2.4, replacing the traditionally used linear selections. Since machine learning models can apply more complex and refined selections, accounting for non-linear relations between selection variables, it is expected to improve the $\mathrm{B_s^0}$ signal extraction, especially at low p_T where the combinatorial background is high. In addition, using the Boosted Decision Trees (BDT) technique has a practical advantage, since it is rather quick to find an optimised selection on the BDT output score for both ITS setups compared to optimising two sets of rectangular selections.

The reconstruction and selection of $\mathrm{B_s^0}$ candidates rely on the displaced decay-vertex topology of both the $\mathrm{B_s^0}$ and the $\mathrm{D_s^-}$ mesons in the decay chain. Typical selection variables that are provided to the BDTs are, for example, the cosine of the pointing angle for the $\mathrm{B_s^0}$ meson, the decay length and invariant mass of the $\mathrm{D_s^-}$ meson, and the product of impact parameters in the transverse plane of the $\mathrm{D_s^-}$ meson and pion-candidate track. Particle identification is only required at the prefiltering stage, i.e. not part of the training variables of the BDT, to factorise out the effect of PID when comparing the performance of both ITS setups. See Sec. 4.2 for a more detailed discussion on the $\mathrm{B_s^0}$-meson reconstruction and selection.

The hyperparameters of the BDTs are optimised using a Bayesian optimisation method [198], performed separately for each p_T interval. For each p_T interval, it was chosen to use a double BDT structure for the $\mathrm{B_s^0}$ selection, in sequence, which lead to a factor 2–4 improved performance with respect to a selection on the output of a single BDT. This was understood due to the very loose (basically none) $\mathrm{B_s^0}$ topological selections that were used at prefiltering level. A single BDT was found to target mainly the "obvious differences" between background and signal $\mathrm{B_s^0}$ candidates, making the optimal selection on the BDT output score converge at very high values. Instead of using these optimised selections, the first BDT is used as an extra prefiltering selection, with a loose selection on the ML probability (0.3–0.7 for ITS2, 0.3–0.5 for ITS3). The $\mathrm{B_s^0}$ candidates that pass these selections are then used to train the second BDT (with the same training variables, but a different Bayesian optimisation of the hyperparameters), for which

the most optimal selection was chosen. This second selection was optimised to obtain a high statistical significance, chosen after a scan of the BDT probability output (> 0.91 for ITS2 and > 0.98 for ITS3).

8.3.2 Expected precision

To estimate the expected precision on the measurement of the $\mathrm{B_s^0}$ mesons, one needs to estimate the expected number of signal and background candidates within a 3σ region around the $\mathrm{B_s^0}$ mass peak, being σ the Gaussian width of the peak. The expected signal is calculated according to the following equation

$$S = 2 \cdot \Delta p_{\mathrm{T}} \cdot \mathrm{BR} \cdot N_{\mathrm{ev}}^{\mathrm{exp}} \cdot \alpha_y \cdot (A \times \epsilon) \cdot \langle T_{\mathrm{AA}}^{0\text{–}10\%} \rangle \cdot \\ (\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}})_{|y|<0.5}^{\mathrm{b,\,FONLL}} \cdot f(\mathrm{b} \to \mathrm{B}) \cdot F(\mathrm{B/B_s^0})^{\mathrm{LHCb}} \cdot R_{\mathrm{AA}}^{\mathrm{B_s^0,\,TAMU}}, \tag{8.2}$$

where Δp_{T} and α_y enclose the transverse momentum and rapidity coverage, BR the branching ratio of the full decay chain, and $N_{\mathrm{ev}}^{\mathrm{exp}}$ the expected number of events corresponding to the expected luminosity of $\mathcal{L}_{\mathrm{int}} = 10\ \mathrm{nb}^{-1}$ (7.7 billion events). The acceptance-times-efficiency for $\mathrm{B_s^0}$ mesons after selections are calculated from the respective MC productions with the ITS2 or ITS3 setup. The last five terms are needed to properly calculate the expected number of $\mathrm{B_s^0}$ candidates in $N_{\mathrm{ev}}^{\mathrm{exp}}$ Pb–Pb collisions. The $\langle T_{\mathrm{AA}}^{0\text{–}10\%} \rangle$ is the average nuclear overlap function for the 0–10% centrality class ($23.07\ \mu\mathrm{b}^{-1}$), which is multiplied by the beauty-quark production in pp collisions at $\sqrt{s} = 5.5$ TeV in $|y| < 0.5$ predicted by FONLL calculations [83], $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}$. This is multiplied again by the fragmentation fraction $f(b \to \mathrm{B}) = 40.7\%$ from $\mathrm{e^+e^-}$ collision data [262], and to the fraction $F(\mathrm{B/B_s^0})^{\mathrm{LHCb}} = 2 \cdot f_{\mathrm{s}}/(f_{\mathrm{d}} + f_{\mathrm{u}}) = 2 \cdot (0.122 \pm 0.006)$ from a recent measurement of the LHCb Collaboration in pp collisions at $\sqrt{s} = 13$ TeV [108]. Finally, $R_{\mathrm{AA}}^{\mathrm{B_s^0,\,TAMU}}$ is the nuclear modification factor of $\mathrm{B_s^0}$ mesons predicted by the TAMU model for the 0–10% centrality class [259]. Since we consider the signal within 3σ of the $\mathrm{B_s^0}$ resolution of the invariant mass peak, S is multiplied by the factor 0.9973 (not indicated in Eq. 8.2).

The expected number of background within 3σ after the application of the selection criteria is estimated considering the following sources

$$B = \frac{B_{\mathrm{SB\ data}}^{\mathrm{current}}(3\sigma)}{B_{\mathrm{MC}}^{\mathrm{current}}(3\sigma)} \cdot B_{\mathrm{MC}}^{\mathrm{upgrade}} + B_{\mathrm{FONLL+MC}}^{(\mathrm{pr.\ D_s^-})+\pi} + \sum_{h_{\mathrm{b}}} B_{\mathrm{FONLL+MC}}^{(\mathrm{fd.\ D_s^-})+\pi} \cdot \frac{F_{\mathrm{b}\to h_{\mathrm{b}}\to \mathrm{D_s^-}+\mathrm{X}}}{F_{\mathrm{b}\to \mathrm{D_s^-}+\mathrm{X}}}. \tag{8.3}$$

Here, $B_{\mathrm{MC}}^{\mathrm{upgrade}}$ is the number of combinatorial-background candidates candidates extracted from an exponential fit to the invariant-mass distribution coming from the HIJING underlying event, scaled to account for MC–data differences by $B_{\mathrm{MC}}^{\mathrm{current}}(3\sigma)$ and $B_{\mathrm{SB\ data}}^{\mathrm{current}}(3\sigma)$, the corresponding number of background candidates for the LHC Run 2 MC simulations and real data for Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, respectively. Two additional background sources are considered, coming from true $\mathrm{D_s^-}$ mesons (injected by PYTHIA) combined with background pions (from HIJING). This source has been calculated separately for prompt and feed-down $\mathrm{D_s^-}$ mesons, for which the expected number of $\mathrm{D_s^-}$ signal has been estimated with a similar FONLL+TAMU

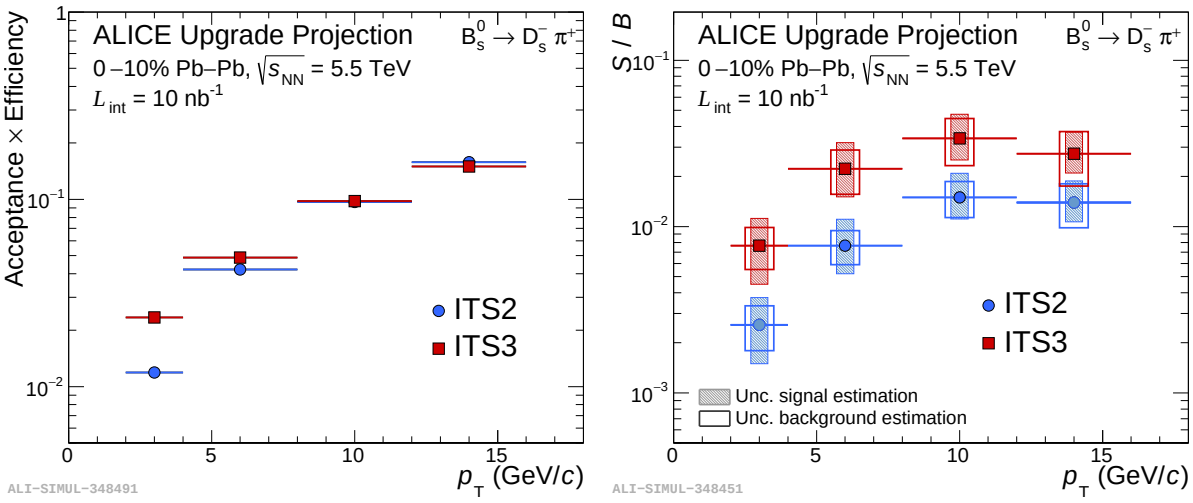


Figure 8.1: Left: The B_s^0 acceptance-times-efficiency for the ITS2 and ITS3 setup, both optimised to have a high expected statistical significance. Right: The estimated signal-over-background ratio for both ITS setups.

strategy as used to estimate the B_s^0 signal. The background coming from feed-down D_s^- mesons have been separated for B^0, B^+, B_s^0, and Λ_b^0 hadrons (indicated by the sum over h_b), and properly scaled to the expected numbers by the F factors. Exploiting a weighting strategy, the transverse momentum of the D_s^- mesons is then propagated to the expected p_T spectrum of the background B_s^0 candidates. In the end, these additional background sources of true D_s^- mesons with background pion-candidate tracks account for 1% to 5% of the total background, increasing with p_T.

The different ingredients to estimate the expected significance ($S/\sqrt{S+B}$) corresponding to the selections reported, are shown in Fig. 8.1, with $A \times \epsilon$ on the left and the expected S/B on the right. A clear improvement for the ITS3 setup is visible, leading to a larger selection efficiency at low p_T, and a larger background rejection in general. Figure 8.2 shows the expected significance for a luminosity of $\mathcal{L}_{int} = 10$ nb^{-1} of 0–10% central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV. The improvement of the statistical significance for the B_s^0 mesons is of about 1.4 up to a factor 2.5 at low p_T. This improvement provides an additional p_T interval in reach for the ITS3 setup, both at low and high transverse momentum.

8.3.3 Systematic uncertainties

This analysis has two types of systematic uncertainties, one on the estimation of the expected signal and background (the boxes shown in Figs 8.1 and 8.2), and the expected "real" systematic uncertainties if such a B_s^0 measurement would be performed on real data. Both types will now be discussed separately.

The largest uncertainty on the calculation of the expected number of B_s^0 signal is coming from the theoretical input, especially the FONLL predictions. The uncertainties on the used experimental

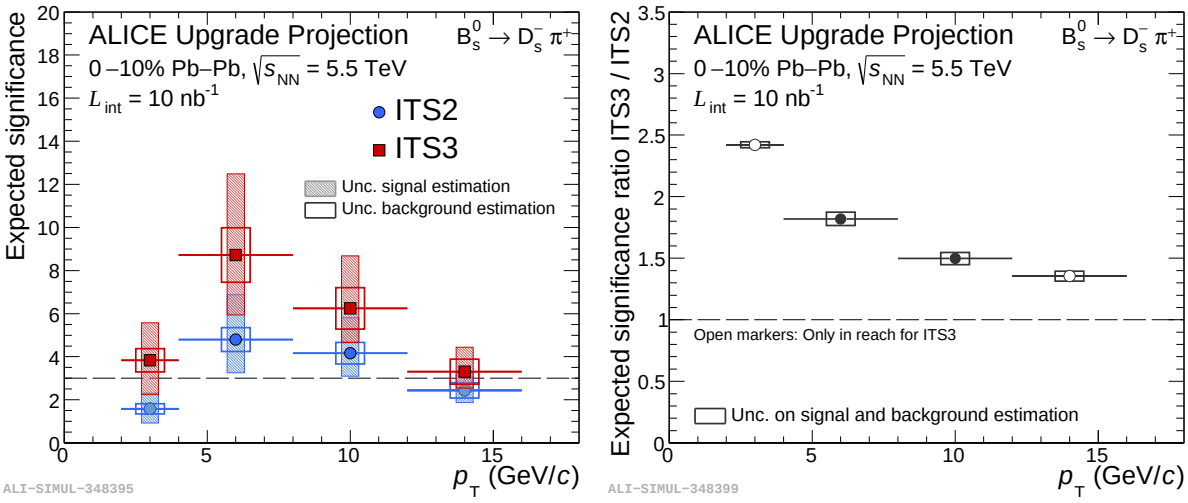


Figure 8.2: Left: The expected significance of the $B_s^0 \to D_s^- \pi^+$ decay for a luminosity of $\mathcal{L}_{int} = 10$ nb^{-1} of 0–10% central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV, for the ITS2 and ITS3 setup in blue and red, respectively. The dashed line indicates a significance of three, which is ALICE's limit to make an observation. Right: The ratio of the expected B_s^0 significances of the ITS3 and ITS2 setups.

measurements of the LHCb Collaboration and other collider experiments (for the used branching ratio and fragmentation fractions), together with the statistical uncertainty on the acceptance-times-efficiency calculation estimated from PYTHIA simulations, are negligible compared to the FONLL theoretical uncertainty. At the moment of writing, only one B_s^0 nuclear modification factor prediction is available, which comes without theoretical uncertainties, hence no systematic has been propagated for this part.

Similar systematic uncertainties exist for the estimation of the background source of true D_s^- with background pion-candidate tracks. However, due to the small contribution of these sources, they are negligible compared to the uncertainty on the estimation of the combinatorial background from HIJING simulations. A large uncertainty on this latter source is assigned due to the low B_s^0 combinatorial background candidates in the simulation. In earlier ALICE upgrade studies [254], a flat 40% systematic uncertainty was assumed, while for this study a parametrisation of the fit parameters as function of BDT output score has been used instead. A fine scan of the BDT probability was performed, going through the fitting procedure for each invariant-mass distribution. It is assumed that the slope parameter does not vary much with the BDT probability, and could therefore be parametrised with a fit. From this fit, and the total number of combinatorial background candidates in the full invariant-mass interval, the other fit parameter of the exponential function was calculated. An upper and lower limit was estimated by using different ranges for fitting the slope parameter and calculating the total number of candidates using bin-counting or by fitting. The corresponding curves for the $4 < p_T < 8$ GeV/c interval with the ITS2 setup are shown in the left panel of Fig. 8.3 as an example. The envelope of these variations was used to quote the uncertainties on the background estimation, which vary between 25% and 40%.

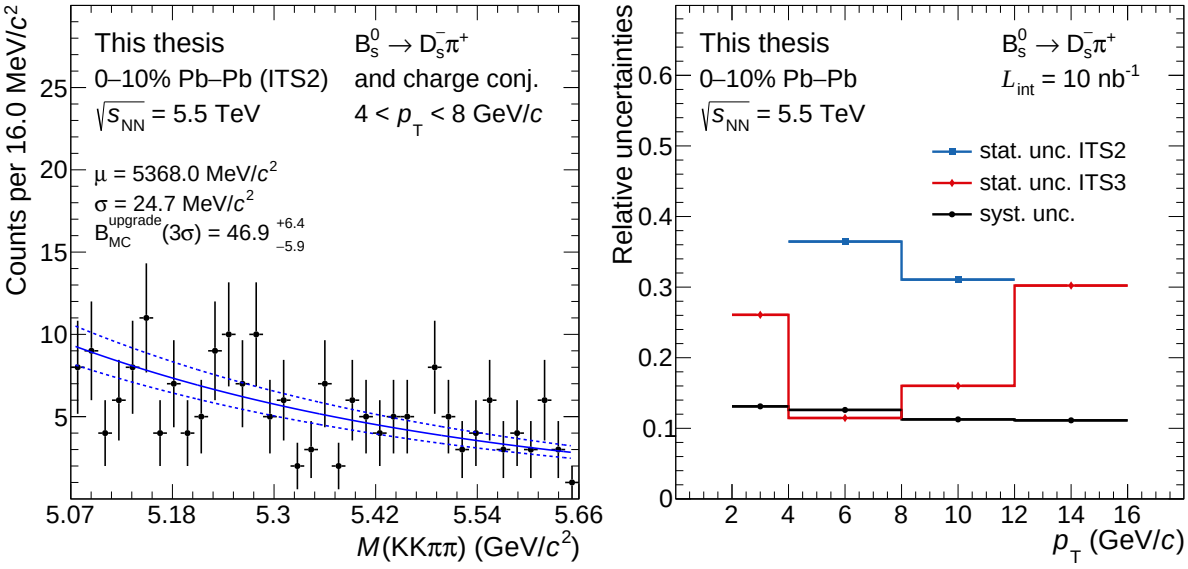


Figure 8.3: Left: Fit to the invariant-mass distribution of combinatorial B_s^0 background candidates for the ITS2 setup with $4 < p_T < 8$ GeV/c. Because of the low number of background candidates, a systematic uncertainty has been estimated, indicated by the dashed blue lines. See text for more details. Right: The expected statistical uncertainty for the ITS2 and ITS3 setups compared to an estimation of the systematic uncertainties for the B_s^0 analysis.

The systematic uncertainties discussed above are not related to the ones that would be quoted on an actual B_s^0 measurement. These have been estimated, or "guessed", based on the D_s^+ Pb–Pb analysis on the 2018 data sample [163]. A systematic uncertainty due to the yield extraction (δ_1), tracking efficiency (δ_2), candidate selection efficiency (δ_3), PID selection efficiency (δ_4), and the generated MC p_T shape (δ_5) are considered. In addition, the uncertainty on the T_{AA} (1.9% [144]) has to be taken into account as well. The uncertainty on the total branching ratio (δ_7; 8.5% [16]) will cancel in the R_{AA} projection, but should be considered otherwise. None of the systematic uncertainties are expected to deviate (significantly) between the ITS2 and ITS3 setup, so equal values have been assigned.

The systematic related to the candidate selection efficiency (8% at low p_T and 4% at high p_T), the PID selection efficiency (0% everywhere), and tracking efficiency (0.5% per track coming from the TPC quality selections and 2% per track from the ITS-TPC matching efficiency), cannot be estimated for the B_s^0 performance analysis directly, and have therefore been guessed based on the D_s^+ Pb–Pb analysis [163]. The raw-yield extraction systematic (6% everywhere) is based as well on the D_s^+ Pb–Pb analysis, but confirmed by two toy MC studies. Finally, the systematic source due to the generated MC p_T shape (1% to 4%) has been estimated directly using the available MC production, weighting the PYTHIA generated p_T spectrum to the one of FONLL calculations [83] and FONLL multiplied with R_{AA} predictions of the TAMU model [259]. The contributions of these different sources of uncertainties were summed in quadrature to obtain the total systematic uncertainty in each p_T interval. The right panel of Fig. 8.3 shows how the total systematic compares to the expected statistical uncertainty of this measurement (estimated as the inverse of the expected significance).

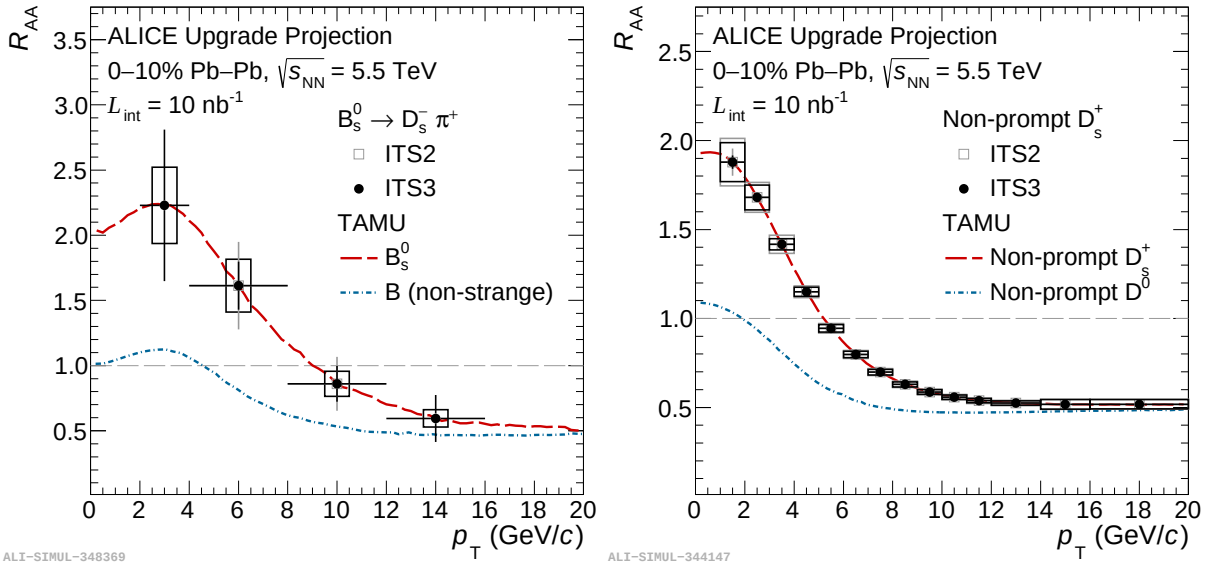


Figure 8.4: Expected nuclear modification factor of B_s^0 mesons (left), compared to the predictions for strange and non-strange B mesons from the TAMU model [259]. The right panel shows the final result performed for a similar study on non-prompt D_s^+ mesons.

8.3.4 Results

Figure 8.4 shows the expected performance projected on the measurement of the R_{AA} of B_s^0 mesons. The uncertainties of the pp reference were not included, considering them small compared to the uncertainties on the measurement in Pb–Pb collisions. The central points were assumed to be equal to the prediction of the TAMU model [259], which is shown with a red-dashed line. The TAMU prediction for non-strange B mesons is shown in blue. At the moment of writing, no ALICE projection exists for the non-strange B-meson R_{AA} for the ITS3 setup, but assuming a better precision due to the larger data sample, one can expect that the precision on the B_s^0 measurement with ITS3 is sufficient to discriminate between the predicted modification of non-strange and strange B mesons by the TAMU model.

8.4 Discussion

In this chapter, the physics performance studies of the $B_s^0 \rightarrow D_s^- \pi^+$ analysis with the upgraded ITS geometries was presented. Where other probes, like the B^+ or Λ_b^0, were already studied in the years prior to the ITS2 installation, the B_s^0 analysis was performed for the first time in ALICE. This for the simple fact that it is a more challenging analysis requiring more events and/or a better resolution to discriminate between signal and background. With the foreseen ITS3 upgrade project, it was worth to estimate the feasibility of beauty-strange hadron production for Pb–Pb collisions in LHC Run 3 and 4.

It will indeed be feasible to measure the B_s^0 meson in Pb–Pb collisions with the upgraded ITS geometries. However, even with this unprecedented precision of the ITS3 detector, only four p_T intervals will be in reach. Nevertheless, with the here presented results, the expected

measurement of the $\mathrm{B_s^0}$ meson will complement and extend to lower p_T the existing measurement of the CMS Collaboration [261] and their projection for the High Luminosity LHC era [263].

The expected performance for the ITS3 setup has, however, a limitation that I did not discuss clearly so far. For a proper comparison, both ITS geometries were assumed to collect $\mathcal{L}_\mathrm{int} = 10\ \mathrm{nb}^{-1}$, a number that is expected for the combined LHC Run 3 and 4 by ALICE. This is a reasonable assumption for these type of performance studies, but one should keep in mind that the ITS3 will, at the earliest, only be installed in LS3, and therefore will collect less events. From past experience we also know that the merging of data samples from different LHC runs is seldom advantageous: the different detector configurations lead to a complicated treatment of systematic uncertainties. So the expected significance for the ITS3 detector might be overestimated.

Personally, I think the ITS3 is a very nice and promising detector, and ALICE should definitely pursue the R&D to work towards a future installation in the ALICE experiment. I also think that it will help to get the beauty sector in heavy-ion collisions in reach with reasonable precision. I do, however, not think that this $\mathrm{B_s^0} \rightarrow \mathrm{D_s^-}\pi^+$ analysis should be one of the main objectives. It might be wiser to focus on the measurement of beauty hadrons with higher fragmentation and branching ratios. For example, the upgrade projection of the B^+ meson that was studied for the TDR of the ITS2 detector [254], has eight p_T intervals in reach within $1 < p_\mathrm{T} < 16$ GeV/c. A good alternative for the $\mathrm{B_s^0}$ meson are non-prompt $\mathrm{D_s^+}$ mesons, for which a large fraction come from $\mathrm{B_s^0}$ decays ($\mathrm{BR}(\mathrm{B_s^0} \rightarrow \mathrm{D_s^+} + \mathrm{X}) = (93 \pm 25)\%$ [16]) and thus sensitive to beauty-strange hadron production in the QGP as well. A similar expected performance study was done for non-prompt $\mathrm{D_s^+}$ mesons, for which the projection on the R_AA observable can be seen in the right panel of Fig. 8.4. It guarantees a significantly higher statistical precision with respect to the $\mathrm{B_s^0}$ analysis, although it has the downside that it is not possible to distinguish $\mathrm{D_s^+}$ mesons coming from different beauty-hadron species and that the p_T of the beauty-hadron parent is not reconstructed.

CHAPTER B

Bibliography

[1] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, "The standard model of particle physics", *Rev. Mod. Phys.* **71** (1999) S96–S111, arXiv:hep-ph/9812285.

[2] **CMS** Collaboration, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[3] **ATLAS** Collaboration, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[4] J. D. Lykken, "Beyond the Standard Model", *Proceedings of the 17th European School of High-energy Physics (ESHEP)* (May, 2010) 101–110, arXiv:1005.1676.

[5] **LHCb** Collaboration, R. Aaij *et al.*, "Test of lepton universality in beauty-quark decays", arXiv:2103.11769 [hep-ex].

[6] **Muon g-2** Collaboration, B. Abi *et al.*, "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm", *Phys. Rev. Lett.* **126** (2021) 141801, arXiv:2104.03281 [hep-ex].

[7] **ALICE** Collaboration, K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002.

[8] **ALICE** Collaboration, J. Adam *et al.*, "Enhanced production of multi-strange hadrons in high-multiplicity proton–proton collisions", *Nature Phys.* **13** (2017) 535–539, arXiv:1606.07424 [nucl-ex].

[9] **PHENIX** Collaboration, C. Aidala *et al.*, "Creation of quark–gluon plasma droplets with three distinct geometries", *Nature Phys.* **15** (2019) 214–220, arXiv:1805.02973 [nucl-ex].

[10] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of long-range near-side two-particle angular correlations in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **116** (2016) 172302, arXiv:1510.03068 [nucl-ex].

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

[12] **ZEUS** Collaboration, I. Abt *et al.*, "Two-particle azimuthal correlations as a probe of collective behaviour in deep inelastic ep scattering at HERA", *JHEP* **04** (2020) 070, arXiv:1912.07431 [hep-ex].

[13] **Belle** Collaboration, A. Abdesselam *et al.*, "Measurement of two-particle correlations in hadronic e^+e^- collisions at Belle", arXiv:2008.04187 [hep-ex].

[14] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.

[15] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the Color Octet Gluon Picture", *Phys. Lett. B* **47** (1973) 365–368.

[16] **Particle Data Group** Collaboration, P. A. Zyla *et al.*, "Review of Particle Physics", *Prog. Theor. Exp. Phys.* **2020** (2020) 083C01.

[17] K. G. Wilson, "Confinement of Quarks", *Phys. Rev. D* **10** (1974) 2445–2459.

[18] **HotQCD** Collaboration, A. Bazavov *et al.*, "Equation of state in (2+1)-flavor QCD", *Phys. Rev. D* **90** (2014) 094503, arXiv:1407.6387 [hep-lat].

[19] S. Borsanyi, "Thermodynamics of the QCD transition from lattice", *Nucl. Phys. A* **904-905** (2013) 270c–277c, arXiv:1210.6901 [hep-lat].

[20] M. G. Alford *et al.*, "Color superconductivity in dense quark matter", *Rev. Mod. Phys.* **80** (2008) 1455–1515, arXiv:0709.4635 [hep-ph].

[21] F. R. Brown *et al.*, "On the existence of a phase transition for QCD with three light quarks", *Phys. Rev. Lett.* **65** (1990) 2491–2494.

[22] Y. Aoki *et al.*, "The order of the quantum chromodynamics transition predicted by the standard model of particle physics", *Nature* **443** (2006) 675–678, arXiv:hep-lat/0611014.

[23] E. Laermann and O. Philipsen, "Lattice QCD at finite temperature", *Ann. Rev. Nucl. Part. Sci.* **53** (2003) 163–198, arXiv:hep-ph/0303042.

[24] A. Aprahamian *et al.*, "Reaching for the Horizon: The 2015 Long Range Plan for Nuclear Science.". Unpublished, Oct, 2015, https://www.osti.gov/biblio/1296778.

[25] W. Busza, K. Rajagopal, and W. van der Schee, "Heavy Ion Collisions: The Big Picture, and the Big Questions", *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 339–376, arXiv:1802.04801 [hep-ph].

[26] J.-P. Blaizot, "High gluon densities in heavy ion collisions", *Rept. Prog. Phys.* **80** (2017) 032301, arXiv:1607.04448 [hep-ph].

[27] F. Gelis, "Color Glass Condensate and Glasma", *Int. J. Mod. Phys. A* **28** (2013) 1330001, arXiv:1211.3327 [hep-ph].

[28] R. Baier *et al.*, "'Bottom up' thermalization in heavy ion collisions", *Phys. Lett. B* **502** (2001) 51–58, arXiv:hep-ph/0009237.

[29] T. Lappi and L. McLerran, "Some features of the glasma", *Nucl. Phys. A* **772** (2006) 200–212, arXiv:hep-ph/0602189.

[30] A. Kurkela *et al.*, "Effective kinetic description of event-by-event pre-equilibrium dynamics in high-energy heavy-ion collisions", *Phys. Rev. C* **99** (2019) 034910, arXiv:1805.00961 [hep-ph].

[31] A. Kurkela *et al.*, "Matching the Nonequilibrium Initial Stage of Heavy Ion Collisions to Hydrodynamics with QCD Kinetic Theory", *Phys. Rev. Lett.* **122** (2019) 122302, arXiv:1805.01604 [hep-ph].

[32] L. D. Landau, "On the multiparticle production in high-energy collisions", *Izv. Akad. Nauk Ser. Fiz.* **17** (1953) 51–64.

[33] S. Z. Belenkij and L. D. Landau, "Hydrodynamic theory of multiple production of particles", *Usp. Fiz. Nauk* **56** (1955) 309.

[34] G. Policastro, D. T. Son, and A. O. Starinets, "The Shear viscosity of strongly coupled $N = 4$ supersymmetric Yang-Mills plasma", *Phys. Rev. Lett.* **87** (2001) 081601, arXiv:hep-th/0104066.

[35] U. Heinz and R. Snellings, "Collective flow and viscosity in relativistic heavy-ion collisions", *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 123–151, arXiv:1301.2826 [nucl-th].

[36] 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.

[37] **ALICE** Collaboration, S. Acharya *et al.*, "Centrality determination in heavy ion collisions", ALICE-PUBLIC-2018-011.

[38] E. Schnedermann, J. Sollfrank, and U. W. Heinz, "Thermal phenomenology of hadrons from 200A GeV S+S collisions", *Phys. Rev. C* **48** (1993) 2462–2475, arXiv:nucl-th/9307020.

[39] **PHENIX** Collaboration, K. Adcox *et al.*, "Suppression of hadrons with large transverse momentum in central Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 130$ GeV", *Phys. Rev. Lett.* **88** (2002) 022301, arXiv:nucl-ex/0109003.

[40] R. J. Fries *et al.*, "Hadron production in heavy ion collisions: Fragmentation and recombination from a dense parton phase", *Phys. Rev. C* **68** (2003) 044902, arXiv:nucl-th/0306027.

[41] V. Topor Pop *et al.*, "Baryon junction loops and the baryon-meson anomaly at high energies", *Phys. Rev. C* **70** (2004) 064906, arXiv:nucl-th/0407095.

[42] J.-Y. Ollitrault, "Anisotropy as a signature of transverse collective flow", *Phys. Rev. D* **46** (1992) 229–245.

[43] M. A. Lisa *et al.*, "Femtoscopy in relativistic heavy ion collisions", *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 357–402, arXiv:nucl-ex/0505014.

[44] L. Vermunt *et al.*, "Influence of final-state radiation on heavy-flavour observables in pp collisions", arXiv:1710.09639 [nucl-th].

[45] J. D. Bjorken, "Highly Relativistic Nucleus-Nucleus Collisions: The Central Rapidity Region", *Phys. Rev. D* **27** (1983) 140–151.

[46] **ALICE** Collaboration, J. Adam *et al.*, "Centrality dependence of the charged-particle multiplicity density at midrapidity in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Rev. Lett.* **116** (2016) 222302, arXiv:1512.06104 [nucl-ex].

[47] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the pseudorapidity and centrality dependence of the transverse energy density in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV", *Phys. Rev. Lett.* **109** (2012) 152303, arXiv:1205.2488 [nucl-ex].

[48] **PHENIX** Collaboration, S. S. Adler *et al.*, "Systematic studies of the centrality and $\sqrt{s_{\mathrm{NN}}}$ dependence of the $\mathrm{d}E_{\mathrm{T}}/\mathrm{d}\eta$ and $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ in heavy ion collisions at mid-rapidity", *Phys. Rev. C* **71** (2005) 034908, arXiv:nucl-ex/0409015. [Erratum: *Phys. Rev. C* **71** (2005) 049901].

[49] **ALICE** Collaboration, S. Acharya *et al.*, "Production of ^{4}He and $^4\overline{\mathrm{He}}$ in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV at the LHC", *Nucl. Phys. A* **971** (2018) 1–20, arXiv:1710.07531 [nucl-ex].

[50] **ALICE** Collaboration, S. Acharya *et al.*, "Production of charged pions, kaons, and (anti-)protons in Pb–Pb and inelastic pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Rev. C* **101** (2020) 044907, arXiv:1910.07678 [nucl-ex].

[51] C. Loizides and A. Morsch, "Absence of jet quenching in peripheral nucleus–nucleus collisions", *Phys. Lett. B* **773** (2017) 408–411, arXiv:1705.08856 [nucl-ex].

[52] **STAR** Collaboration, C. Adler *et al.*, "Midrapidity Λ and $\bar{\Lambda}$ production in Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 130$ GeV", *Phys. Rev. Lett.* **89** (2002) 092301, arXiv:nucl-ex/0203016.

[53] **PHENIX** Collaboration, K. Adcox *et al.*, "Centrality dependence of $\pi^{+/-}$, $\mathrm{K}^{+/-}$, p and $\bar{\mathrm{p}}$ production from $\sqrt{s_{\mathrm{NN}}} = 130$ GeV Au+Au collisions at RHIC", *Phys. Rev. Lett.* **88** (2002) 242301, arXiv:nucl-ex/0112006.

[54] Z.-w. Lin and C. M. Ko, "Flavor ordering of elliptic flows at high transverse momentum", *Phys. Rev. Lett.* **89** (2002) 202302, arXiv:nucl-th/0207014.

[55] V. Greco, C. M. Ko, and P. Levai, "Parton coalescence and anti-proton / pion anomaly at RHIC", *Phys. Rev. Lett.* **90** (2003) 202302, arXiv:nucl-th/0301093.

[56] R. J. Fries *et al.*, "Hadronization in heavy ion collisions: Recombination and fragmentation of partons", *Phys. Rev. Lett.* **90** (2003) 202303, arXiv:nucl-th/0301087.

[57] J. Rafelski and B. Muller, "Strangeness Production in the Quark-Gluon Plasma", *Phys. Rev. Lett.* **48** (1982) 1066. [Erratum: *Phys. Rev. Lett.* **56** (1986), 2334].

[58] **NA35** Collaboration, J. Baechler *et al.*, "Strangeness enhancement in central S+S collisions at 200 GeV/nucleon", *Nucl. Phys. A* **525** (1991) 221C–226C.

[59] **WA97** Collaboration, E. Andersen *et al.*, "Strangeness enhancement at mid-rapidity in Pb–Pb collisions at 158A GeV/c", *Phys. Lett. B* **449** (1999) 401–406.

[60] **NA57** Collaboration, F. Antinori *et al.*, "Study of the production of strange and multi-strange particles in lead–lead interactions at the CERN SPS: The NA57 experiment", *Nucl. Phys. A* **681** (2001) 165–173.

[61] **STAR** Collaboration, G. Agakishiev *et al.*, "Strangeness Enhancement in Cu+Cu and Au+Au Collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV", *Phys. Rev. Lett.* **108** (2012) 072301, arXiv:1107.2955 [nucl-ex].

[62] **ALICE** Collaboration, B. B. Abelev *et al.*, "Multi-strange baryon production at mid-rapidity in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV", *Phys. Lett. B* **728** (2014) 216–227, arXiv:1307.5543 [nucl-ex]. [Erratum: *Phys. Lett. B* **734** (2014), 409–410].

[63] J. D. Bjorken, "Energy Loss of Energetic Partons in Quark-Gluon Plasma: Possible Extinction of High p_{T} Jets in Hadron-Hadron Collisions", FERMILAB-PUB-82-059-THY.

[64] **ALICE** Collaboration, S. Acharya *et al.*, "Transverse momentum spectra and nuclear modification factors of charged particles in pp, p–Pb and Pb–Pb collisions at the LHC", *JHEP* **11** (2018) 013, arXiv:1802.09145 [nucl-ex].

[65] J. W. Cronin *et al.*, "Production of hadrons with large transverse momentum at 200, 300, and 400 GeV", *Phys. Rev. D* **11** (1975) 3105–3123.

[66] M. Xie *et al.*, "Extracting jet transport coefficient via single hadron and dihadron productions in high-energy heavy-ion collisions", *Eur. Phys. J. C* **79** (2019) 589, arXiv:1901.04155 [hep-ph].

[67] **ALICE** Collaboration, S. Acharya *et al.*, "Anisotropic flow of identified particles in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *JHEP* **09** (2018) 006, arXiv:1805.04390 [nucl-ex].

[68] **ALICE** Collaboration, S. Acharya *et al.*, "Centrality and transverse momentum dependence of inclusive J/ψ production at midrapidity in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Lett. B* **805** (2020) 135434, arXiv:1910.14404 [nucl-ex].

[69] **ALICE** Collaboration, J. Adam *et al.*, "Direct photon production in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV", *Phys. Lett. B* **754** (2016) 235–248, arXiv:1509.07324 [nucl-ex].

[70] J. Schukraft, "QM2017: Status and Key open Questions in Ultra-Relativistic Heavy-Ion Physics", *Nucl. Phys. A* **967** (2017) 1–10, arXiv:1705.02646 [hep-ex].

[71] N. Fischer and T. Sjöstrand, "Thermodynamical String Fragmentation", *JHEP* **01** (2017) 140, arXiv:1610.09818 [hep-ph].

[72] M. L. Mangano, "Two lectures on heavy quark production in hadronic collisions", *Proc. Int. Sch. Phys. Fermi* **137** (1998) 95–137, arXiv:hep-ph/9711337.

[73] A. H. Mueller and P. Nason, "Heavy particle content in QCD jets", *Phys. Lett. B* **157** (1985) 226–228.

[74] M. L. Mangano and P. Nason, "Heavy quark multiplicities in gluon jets", *Phys. Lett. B* **285** (1992) 160–166.

[75] A. Banfi, G. P. Salam, and G. Zanderighi, "Accurate QCD predictions for heavy-quark jets at the Tevatron and LHC", *JHEP* **07** (2007) 026, arXiv:0704.2999 [hep-ph].

[76] B.-W. Zhang, C.-M. Ko, and W. Liu, "Thermal charm production in a quark-gluon plasma in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.5$ TeV", *Phys. Rev. C* **77** (2008) 024901, arXiv:0709.1684 [nucl-th].

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

[78] T. Kneesch *et al.*, "Charmed-meson fragmentation functions with finite-mass corrections", *Nucl. Phys. B* **799** (2008) 34–59, arXiv:0712.0481 [hep-ph].

[79] V. N. Gribov and L. N. Lipatov, "Deep inelastic ep-scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[80] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys. B* **126** (1977) 298–318.

[81] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.", *Sov. Phys. JETP* **46** (1977) 641–653.

[82] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of D^0, D^+, D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE", *Eur. Phys. J. C* **79** (2019) 388, arXiv:1901.07979 [nucl-ex].

[83] M. Cacciari *et al.*, "Theoretical predictions for charm and bottom production at the LHC", *JHEP* **10** (2012) 137, arXiv:1205.6344 [hep-ph].

[84] G. Kramer and H. Spiesberger, "Study of heavy meson production in p–Pb collisions at $\sqrt{s} = 5.02$ TeV in the general-mass variable-flavour-number scheme", *Nucl. Phys. B* **925** (2017) 415–430, arXiv:1703.04754 [hep-ph].

[85] R. Maciuła and A. Szczurek, "Production of Λ_c^+ baryons at the LHC within the k_T-factorization approach and independent parton fragmentation picture", *Phys. Rev. D* **98** (2018) 014016, arXiv:1803.05807 [hep-ph].

[86] A. Andronic *et al.*, "Heavy-flavour and quarkonium production in the LHC era: from proton–proton to heavy-ion collisions", *Eur. Phys. J. C* **76** (2016) 107, arXiv:1506.03981 [nucl-ex].

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

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

[89] B. A. Kniehl *et al.*, "Inclusive $\mathrm{D}^{*\pm}$ production in $\mathrm{p\bar{p}}$ collisions with massive charm quarks", *Phys. Rev. D* **71** (2005) 014018, arXiv:hep-ph/0410289.

[90] B. A. Kniehl *et al.*, "Collinear subtractions in hadroproduction of heavy quarks", *Eur. Phys. J. C* **41** (2005) 199–212, arXiv:hep-ph/0502194.

[91] M. Cacciari, M. L. Mangano, and P. Nason, "Gluon PDF constraints from the ratio of forward heavy-quark production at the LHC at $\sqrt{s} = 7$ and 13 TeV", *Eur. Phys. J. C* **75** (2015) 610, arXiv:1507.06197 [hep-ph].

[92] S. Catani, M. Ciafaloni, and F. Hautmann, "Gluon contributions to small x heavy flavour production", *Phys. Lett. B* **242** (1990) 97–102.

[93] J. C. Collins and R. K. Ellis, "Heavy quark production in very high-energy hadron collisions", *Nucl. Phys. B* **360** (1991) 3–30.

[94] R. Maciula and A. Szczurek, "Open charm production at the LHC: k_{T}-factorization approach", *Phys. Rev. D* **87** (2013) 094022, arXiv:1301.3033 [hep-ph].

[95] T. Sjöstrand *et al.*, "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].

[96] J. Bellm *et al.*, "Herwig 7.0/Herwig++ 3.0 release note", *Eur. Phys. J. C* **76** (2016) 196, arXiv:1512.01178 [hep-ph].

[97] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, arXiv:0707.3088 [hep-ph].

[98] S. Frixione, P. Nason, and B. R. Webber, "Matching NLO QCD and parton showers in heavy flavor production", *JHEP* **08** (2003) 007, arXiv:hep-ph/0305252.

[99] M. Klasen *et al.*, "NLO Monte Carlo predictions for heavy-quark production at the LHC: pp collisions in ALICE", *JHEP* **08** (2014) 109, arXiv:1405.3083 [hep-ph].

[100] **ALICE** Collaboration, S. Acharya *et al.*, "Λ_{c}^{+} production in pp collisions at $\sqrt{s} = 7$ TeV and in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *JHEP* **04** (2018) 108, arXiv:1712.09581 [nucl-ex].

[101] **LHCb** Collaboration, R. Aaij *et al.*, "Prompt charm production in pp collisions at $\sqrt{s} = 7$ TeV", *Nucl. Phys. B* **871** (2013) 1–20, arXiv:1302.2864 [hep-ex].

[102] **ALICE** Collaboration, S. Acharya *et al.*, "Λ_{c}^{+} production in pp and in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", arXiv:2011.06079 [nucl-ex].

[103] **ALICE** Collaboration, S. Acharya *et al.*, "Λ_{c}^{+} production and baryon-to-meson ratios in pp and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV at the LHC", arXiv:2011.06078 [nucl-ex].

[104] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Production of Λ_{c}^{+} baryons in proton–proton and lead–lead collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Lett. B* **803** (2020) 135328, arXiv:1906.03322 [hep-ex].

[105] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of the cross sections of Ξ_{c}^{0} and Ξ_{c}^{+} baryons and branching-fraction ratio $\mathrm{BR}(\Xi_{\mathrm{c}}^{0} \to \Xi^{-}\mathrm{e}^{+}\nu_{\mathrm{e}})/\mathrm{BR}(\Xi_{\mathrm{c}}^{0} \to \Xi^{-}\pi^{+})$ in pp collisions at 13 TeV", arXiv:2105.05187 [nucl-ex].

[106] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of b-hadron production fractions in 7 TeV pp collisions", *Phys. Rev. D* **85** (2012) 032008, arXiv:1111.2357 [hep-ex].

[107] **LHCb** Collaboration, R. Aaij *et al.*, "Study of the production of Λ_b^0 and $\overline{B}^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \to J/\psi p K^-$ branching fraction", *Chin. Phys. C* **40** (2016) 011001, arXiv:1509.00292 [hep-ex].

[108] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of b hadron fractions in 13 TeV pp collisions", *Phys. Rev. D* **100** (2019) 031102, arXiv:1902.06794 [hep-ex].

[109] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the Λ_b cross section and the $\bar{\Lambda}_b$ to Λ_b ratio with $J/\psi\Lambda$ decays in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **714** (2012) 136–157, arXiv:1205.0594 [hep-ex].

[110] J. R. Christiansen and P. Z. Skands, "String Formation Beyond Leading Colour", *JHEP* **08** (2015) 003, arXiv:1505.01681 [hep-ph].

[111] M. He and R. Rapp, "Charm-Baryon Production in proton–proton Collisions", *Phys. Lett. B* **795** (2019) 117–121, arXiv:1902.08889 [nucl-th].

[112] V. Minissale, S. Plumari, and V. Greco, "Charm Hadrons in pp collisions at LHC energy within a Coalescence plus Fragmentation approach", *Phys. Lett. B* **821** (2021) 136622, arXiv:2012.12001 [hep-ph].

[113] H.-H. Li *et al.*, "Production of single-charm hadrons by quark combination mechanism in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV", *Phys. Rev. C* **97** (2018) 064915, arXiv:1712.08921 [hep-ph].

[114] J. Song, H.-h. Li, and F.-l. Shao, "New feature of low p_T charm quark hadronization in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **78** (2018) 344, arXiv:1801.09402 [hep-ph].

[115] N. Armesto, "Nuclear shadowing", *J. Phys. G* **32** (2006) R367–R394, arXiv:hep-ph/0604108.

[116] M. Arneodo, "Nuclear effects in structure functions", *Phys. Rept.* **240** (1994) 301–393.

[117] D. F. Geesaman, K. Saito, and A. W. Thomas, "The nuclear EMC effect", *Ann. Rev. Nucl. Part. Sci.* **45** (1995) 337–390.

[118] 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].

[119] K. J. Eskola *et al.*, "EPPS16: Nuclear parton distributions with LHC data", *Eur. Phys. J. C* **77** (2017) 163, arXiv:1612.05741 [hep-ph].

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

[121] I. Vitev, "Non-Abelian energy loss in cold nuclear matter", *Phys. Rev. C* **75** (2007) 064906, arXiv:hep-ph/0703002.

[122] **CMS** Collaboration, S. Chatrchyan *et al.*, "Observation of Long-Range Near-Side Angular Correlations in Proton–Lead Collisions at the LHC", *Phys. Lett. B* **718** (2013) 795–814, arXiv:1210.5482 [nucl-ex].

[123] **CMS** Collaboration, V. Khachatryan *et al.*, "Evidence for Collective Multiparticle Correlations in p–Pb Collisions", *Phys. Rev. Lett.* **115** (2015) 012301, arXiv:1502.05382 [nucl-ex].

[124] **ALICE** Collaboration, B. B. Abelev *et al.*, "Multiplicity Dependence of Pion, Kaon, Proton and Lambda Production in p–Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Lett. B* **728** (2014) 25–38, arXiv:1307.6796 [nucl-ex].

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

[126] A. M. Sickles, "Possible Evidence for Radial Flow of Heavy Mesons in d+Au Collisions", *Phys. Lett. B* **731** (2014) 51–56, arXiv:1309.6924 [nucl-th].

[127] Y. Xu *et al.*, "Heavy-flavor dynamics in relativistic p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Nucl. Part. Phys. Proc.* **276-278** (2016) 225–228, arXiv:1510.07520 [nucl-th].

[128] A. Beraudo *et al.*, "Heavy-flavour production in high-energy d–Au and p–Pb collisions", *JHEP* **03** (2016) 123, arXiv:1512.05186 [hep-ph].

[129] Y.-T. Chien *et al.*, "Jet Quenching from QCD Evolution", *Phys. Rev. D* **93** (2016) 074030, arXiv:1509.02936 [hep-ph].

[130] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of prompt D^0, D^+, D^{*+}, and $\mathrm{D}_{\mathrm{s}}^+$ production in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *JHEP* **12** (2019) 092, arXiv:1906.03425 [nucl-ex].

[131] **STAR** Collaboration, J. Adams *et al.*, "Identified hadron spectra at large transverse momentum in pp and d+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV", *Phys. Lett. B* **637** (2006) 161–169, arXiv:nucl-ex/0601033.

[132] **ALICE** Collaboration, J. Adam *et al.*, "Multiplicity dependence of charged pion, kaon, and (anti)proton production at large transverse momentum in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Lett. B* **760** (2016) 720–735, arXiv:1601.03658 [nucl-ex].

[133] P. Braun-Munzinger, "Quarkonium production in ultra-relativistic nuclear collisions: Suppression versus enhancement", *J. Phys. G* **34** (2007) S471–478, arXiv:nucl-th/0701093.

[134] Y. L. Dokshitzer and D. E. Kharzeev, "Heavy quark colorimetry of QCD matter", *Phys. Lett. B* **519** (2001) 199–206, arXiv:hep-ph/0106202.

[135] 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. C* **88** (2013) 044907, arXiv:1308.0617 [nucl-th].

[136] G. D. Moore and D. Teaney, "How much do heavy quarks thermalize in a heavy ion collision?", *Phys. Rev. C* **71** (2005) 064904, arXiv:hep-ph/0412346.

[137] C. A. Salgado and U. A. Wiedemann, "Calculating quenching weights", *Phys. Rev. D* **68** (2003) 014008, arXiv:hep-ph/0302184.

[138] S. Peigne and A. V. Smilga, "Energy losses in a hot plasma revisited", *Phys. Usp.* **52** (2009) 659–685, arXiv:0810.5702 [hep-ph].

[139] F. Prino and R. Rapp, "Open Heavy Flavor in QCD Matter and in Nuclear Collisions", *J. Phys. G* **43** (2016) 093002, arXiv:1603.00529 [nucl-ex].

[140] R. Rapp and H. van Hees, "Heavy quarks in the quark-gluon plasma", *Quark-gluon plasma 4* (2010) 111–206, arXiv:0903.1096 [hep-ph].

[141] **ALICE** Collaboration, S. Acharya *et al.*, "D-meson azimuthal anisotropy in midcentral Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV", *Phys. Rev. Lett.* **120** (2018) 102301, arXiv:1707.01005 [nucl-ex].

[142] **ALICE** Collaboration, S. Acharya *et al.*, "Transverse-momentum and event-shape dependence of D-meson flow harmonics in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV", *Phys. Lett. B* **813** (2021) 136054, arXiv:2005.11131 [nucl-ex].

[143] W. Ke, Y. Xu, and S. A. Bass, "Linearized Boltzmann-Langevin model for heavy quark transport in hot and dense QCD matter", *Phys. Rev. C* **98** (2018) 064901, arXiv:1806.08848 [nucl-th].

[144] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of $\rm D^0$, $\rm D^+$, $\rm D^{*+}$ and $\rm D_s^+$ production in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV", *JHEP* **10** (2018) 174, arXiv:1804.09083 [nucl-ex].

[145] C. Peterson *et al.*, "Scaling Violations in Inclusive $\rm e^+e^-$ Annihilation Spectra", *Phys. Rev. D* **27** (1983) 105.

[146] M. Lisovyi, A. Verbytskyi, and O. Zenaiev, "Combined analysis of charm-quark fragmentation-fraction measurements", *Eur. Phys. J. C* **76** (2016) 397, arXiv:1509.01061 [hep-ex].

[147] **ALICE** Collaboration, S. Acharya *et al.*, "Charm-quark fragmentation fractions and production cross section at midrapidity in pp collisions at the LHC", arXiv:2105.06335 [nucl-ex].

[148] C. Flensburg, G. Gustafson, and L. Lonnblad, "Inclusive and Exclusive Observables from Dipoles in High Energy Collisions", *JHEP* **08** (2011) 103, arXiv:1103.4321 [hep-ph].

[149] B. Andersson *et al.*, "Parton Fragmentation and String Dynamics", *Phys. Rept.* **97** (1983) 31–145.

[150] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes", *Phys. Rev. D* **82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[151] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur. Phys. J.* **C74** (2014) 3024, arXiv:1404.5630 [hep-ph].

[152] **ALICE** Collaboration, K. Aamodt *et al.*, "Strange particle production in proton–proton collisions at $\sqrt{s} = 0.9$ TeV with ALICE at the LHC", *Eur. Phys. J. C* **71** (2011) 1594, arXiv:1012.3257 [hep-ex].

[153] **CMS** Collaboration, V. Khachatryan *et al.*, "Strange Particle Production in pp Collisions at $\sqrt{s} = 0.9$ and 7 TeV", *JHEP* **05** (2011) 064, arXiv:1102.4282 [hep-ex].

[154] C. Bierlich *et al.*, "Effects of Overlapping Strings in pp Collisions", *JHEP* **03** (2015) 148, arXiv:1412.6259 [hep-ph].

[155] T. S. Biro, H. B. Nielsen, and J. Knoll, "Color Rope Model for Extreme Relativistic Heavy Ion Collisions", *Nucl. Phys. B* **245** (1984) 449–468.

[156] C. Bierlich and J. R. Christiansen, "Effects of color reconnection on hadron flavor observables", *Phys. Rev. D* **92** (2015) 094010, arXiv:1507.02091 [hep-ph].

[157] **ALICE** Collaboration, S. Acharya *et al.*, "Multiplicity dependence of light-flavor hadron production in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. C* **99** (2019) 024906, arXiv:1807.11321 [nucl-ex].

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

[159] C. B. Dover *et al.*, "Relativistic coalescence model for high-energy nuclear collisions", *Phys. Rev. C* **44** (1991) 1636–1654.

[160] S. Plumari *et al.*, "Charmed Hadrons from Coalescence plus Fragmentation in relativistic nucleus-nucleus collisions at RHIC and LHC", *Eur. Phys. J. C* **78** (2018) 348, arXiv:1712.00730 [hep-ph].

[161] L. Ravagli and R. Rapp, "Quark Coalescence based on a Transport Equation", *Phys. Lett. B* **655** (2007) 126–131, arXiv:0705.0021 [hep-ph].

[162] M. He and R. Rapp, "Hadronization and Charm-Hadron Ratios in Heavy-Ion Collisions", *Phys. Rev. Lett.* **124** (2020) 042301, arXiv:1905.09216 [nucl-th].

[163] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of prompt $\mathrm{D_s^+}$-meson production and azimuthal anisotropy in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", arXiv:2110.10006 [nucl-ex].

[164] S. H. Lee *et al.*, "Λ_c^+ enhancement from strongly coupled quark-gluon plasma", *Phys. Rev. Lett.* **100** (2008) 222301, arXiv:0709.3637 [nucl-th].

[165] Y. Oh *et al.*, "Heavy baryon/meson ratios in relativistic heavy ion collisions", *Phys. Rev. C* **79** (2009) 044905, arXiv:0901.1382 [nucl-th].

[166] E. V. Shuryak and I. Zahed, "Towards a theory of binary bound states in the quark-gluon plasma", *Phys. Rev. D* **70** (2004) 054507, arXiv:hep-ph/0403127.

[167] D. Ebert, R. N. Faustov, and V. O. Galkin, "Spectroscopy and Regge trajectories of heavy baryons in the relativistic quark-diquark picture", *Phys. Rev. D* **84** (2011) 014025, arXiv:1105.0583 [hep-ph].

[168] **STAR** Collaboration, G. Xie, "Λ_c^+ Production in Au+Au Collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV measured by the STAR experiment", *Nucl. Phys. A* **967** (2017) 928–931, arXiv:1704.04353 [nucl-ex].

[169] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of D-meson production at mid-rapidity in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **77** (2017) 550, arXiv:1702.00766 [hep-ex].

[170] **ALICE** Collaboration, J. Adam *et al.*, "D-meson production in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV and in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. C* **94** (2016) 054908, arXiv:1605.07569 [nucl-ex].

[171] E. Braaten *et al.*, "Perturbative QCD fragmentation functions as a model for heavy quark fragmentation", *Phys. Rev. D* **51** (1995) 4819–4829, arXiv:hep-ph/9409316.

[172] **CMS** Collaboration, V. Khachatryan *et al.*, "Multiplicity and rapidity dependence of strange hadron production in pp, p–Pb, and Pb–Pb collisions at the LHC", *Phys. Lett. B* **768** (2017) 103–129, arXiv:1605.06699 [nucl-ex].

[173] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001.

[174] E. Mobs, "The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019.". General Photo, Jul, 2019, https://cds.cern.ch/record/2684277.

[175] **ATLAS** Collaboration, G. Aad *et al.*, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003.

[176] **CMS** Collaboration, S. Chatrchyan *et al.*, "The CMS Experiment at the CERN LHC", *JINST* **3** (2008) S08004.

[177] **LHCb** Collaboration, A. A. Alves, Jr. *et al.*, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005.

[178] **ALICE** Collaboration, P. Cortese *et al.*, "ALICE: Physics performance report, volume I", *J. Phys. G* **30** (2004) 1517–1763.

[179] **ALICE** Collaboration, C. W. Fabjan *et al.*, "ALICE: Physics performance report, volume II", *J. Phys. G* **32** (2006) 1295–2040.

[180] **ALICE** Collaboration, B. B. Abelev *et al.*, "Performance of the ALICE Experiment at the CERN LHC", *Int. J. Mod. Phys. A* **29** (2014) 1430044, arXiv:1402.4476 [nucl-ex].

[181] A. Tauro, "ALICE Schematics.". General Photo, May, 2017, https://cds.cern.ch/record/2263642.

[182] **ALICE** Collaboration, K. Aamodt *et al.*, "Alignment of the ALICE Inner Tracking System with cosmic-ray tracks", *JINST* **5** (2010) P03003, arXiv:1001.0502 [physics.ins-det].

[183] J. Alme *et al.*, "The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events", *Nucl. Instrum. Meth. A* **622** (2010) 316–367, arXiv:1001.1950 [physics.ins-det].

[184] A. Akindinov *et al.*, "Performance of the ALICE Time-Of-Flight detector at the LHC", *Eur. Phys. J. Plus* **128** (2013) 44.

[185] **ALICE** Collaboration, E. Abbas *et al.*, "Performance of the ALICE VZERO system", *JINST* **8** (2013) P10016, arXiv:1306.3130 [nucl-ex].

[186] **ALICE** Collaboration, B. Abelev *et al.*, "Centrality determination of Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV with ALICE", *Phys. Rev. C* **88** (2013) 044909, arXiv:1301.4361 [nucl-ex].

[187] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth. A* **262** (1987) 444–450.

[188] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework", *Nucl. Instrum. Meth. A* **389** (1997) 81–86.

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

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

[191] R. Brun *et al.*, "GEANT Detector Description and Simulation Tool", CERN-W5013.

[192] **GEANT4** Collaboration, S. Agostinelli *et al.*, "GEANT4–a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250–303.

[193] J. Podolanski and R. Armenteros, "III. Analysis of V-events", *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **45** no. 360, (1954) 13–30.

[194] K. Albertsson *et al.*, "Machine Learning in High Energy Physics Community White Paper", *J. Phys. Conf. Ser.* **1085** (2018) 022008, arXiv:1807.02876 [physics.comp-ph].

[195] D. Guest, K. Cranmer, and D. Whiteson, "Deep Learning and its Application to LHC Physics", *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 161–181, arXiv:1806.11484 [hep-ex].

[196] T. Mitchell, *Machine Learning*. McGraw-Hill, 1997.

[197] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016) 785–794, arXiv:1603.02754 [cs.LG].

[198] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms", *Proceedings of the Twenty-sixth Annual Conference on Neural Information Processing Systems (NIPS)* (2012) 46–55, arXiv:1206.2944 [stat.ML].

[199] **ALICE** Collaboration, S. Acharya *et al.*, "Λ_c^+ production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV", *Phys. Lett. B* **793** (2019) 212–223, arXiv:1809.10922 [nucl-ex].

[200] **ALICE** Collaboration, S. Acharya *et al.*, "ALICE 2017 luminosity determination for pp collisions at $\sqrt{s} = 5$ TeV", ALICE-PUBLIC-2018-011.

[201] L. Gladilin, "Fragmentation fractions of c and b quarks into charmed hadrons at LEP", *Eur. Phys. J. C* **75** (2015) 19, arXiv:1404.3888 [hep-ex].

[202] D. J. Lange, "The EvtGen particle decay simulation package", *Nucl. Instrum. Meth. A* **462** (2001) 152–155.

[203] B. A. Kniehl *et al.*, "Inclusive Charmed-Meson Production at the CERN LHC", *Eur. Phys. J. C* **72** (2012) 2082, arXiv:1202.0439 [hep-ph].

[204] I. Helenius and H. Paukkunen, "Revisiting the D-meson hadroproduction in general-mass variable flavour number scheme", *JHEP* **05** (2018) 196, arXiv:1804.03557 [hep-ph].

[205] **ALICE** Collaboration, B. Abelev *et al.*, "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].

[206] A. Andronic *et al.*, "Statistical hadronization of heavy flavor quarks in elementary collisions: Successes and failures", *Phys. Lett. B* **678** (2009) 350–354, arXiv:0904.1368 [hep-ph].

[207] **BaBar** Collaboration, J. P. Lees *et al.*, "Measurement of the $D^{*+}(2010)$ natural line width and the $D^{*+}(2010) - D^0$ mass difference", *Phys. Rev. D* **88** (2013) 052003, arXiv:1304.5009 [hep-ex]. [Erratum: *Phys. Rev. D* **88** (2013), 079902].

[208] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of $D^{*\pm}$, $D^{\pm}$ and $D_s^{\pm}$ meson production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *Nucl. Phys. B* **907** (2016) 717–763, arXiv:1512.02913 [hep-ex].

[209] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, arXiv:hep-ph/0201195.

[210] **ALICE** Collaboration, S. Acharya *et al.*, "The ALICE definition of primary particles", ALICE-PUBLIC-2017-005.

[211] F. Grosa, *Strange and non-strange D-meson production in pp, p–Pb, and Pb–Pb collisions with ALICE at the LHC*. PhD thesis, Turin Polytechnic, 2020. Presented 11 Feb 2020.

[212] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Nuclear modification factor of D^0 mesons in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV", *Phys. Lett. B* **782** (2018) 474–496, arXiv:1708.04962 [nucl-ex].

[213] S. Dulat *et al.*, "New parton distribution functions from a global analysis of quantum chromodynamics", *Phys. Rev. D* **93** (2016) 033006, arXiv:1506.07443 [hep-ph].

[214] **NNPDF** Collaboration, R. D. Ball *et al.*, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, arXiv:1706.00428 [hep-ph].

[215] L. A. Harland-Lang *et al.*, "Parton distributions in the LHC era: MMHT 2014 PDFs", *Eur. Phys. J. C* **75** (2015) 204, arXiv:1412.3989 [hep-ph].

[216] **LHCb** Collaboration, R. Aaij *et al.*, "Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 5$ TeV", *JHEP* **06** (2017) 147, arXiv:1610.02230 [hep-ex].

[217] **LHCb** Collaboration, R. Aaij *et al.*, "Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2016) 159, arXiv:1510.01707 [hep-ex]. [Erratum: *JHEP* **09** (2016), 013; Erratum: *JHEP* **05** (2017), 074].

[218] **Belle** Collaboration, R. Seuster *et al.*, "Charm hadrons from fragmentation and B decays in e^+e^- annihilation at $\sqrt{s} = 10.6$ GeV", *Phys. Rev. D* **73** (2006) 032002, arXiv:hep-ex/0506068.

[219] A. Beraudo *et al.*, "Extraction of Heavy-Flavor Transport Coefficients in QCD Matter", *Nucl. Phys. A* **979** (2018) 21–86, arXiv:1803.03824 [nucl-th].

[220] **ALICE** Collaboration, S. Acharya *et al.*, "Prompt D^0, D^+ and D^{*+} production in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV", arXiv:2110.09420 [nucl-ex].

[221] D. P. Anderle *et al.*, "Using hadron-in-jet data in a global analysis of D^* fragmentation functions", *Phys. Rev. D* **96** (2017) 034028, arXiv:1706.09857 [hep-ph].

[222] M. Soleymaninia, H. Khanpour, and S. M. Moosavi Nejad, "First determination of D^{*+}-meson fragmentation functions and their uncertainties at next-to-next-to-leading order", *Phys. Rev. D* **97** (2018) 074014, arXiv:1711.11344 [hep-ph].

[223] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of $D^{*\pm}$ meson production in jets from pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *Phys. Rev. D* **85** (2012) 052005, arXiv:1112.4432 [hep-ex].

[224] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of charged pion, kaon, and proton production in proton–proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **96** (2017) 112003, arXiv:1706.10194 [hep-ex].

[225] **ALICE** Collaboration, S. Acharya *et al.*, "Multiplicity dependence of (multi-)strange hadron production in proton–proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **80** (2020) 167, arXiv:1908.01861 [nucl-ex].

[226] **ALICE** Collaboration, S. Acharya *et al.*, "Pseudorapidity distributions of charged particles as a function of mid- and forward rapidity multiplicities in pp collisions at $\sqrt{s} = 5.02$, 7 and 13 TeV", *Eur. Phys. J. C* **81** (2021) 630, arXiv:2009.09434 [nucl-ex].

[227] **ALICE** Collaboration, J. Adam *et al.*, "Charged-particle multiplicities in proton–proton collisions at $\sqrt{s} = 0.9$ to 8 TeV", *Eur. Phys. J. C* **77** (2017) 33, arXiv:1509.07541 [nucl-ex].

[228] **ALICE** Collaboration, J. Adam *et al.*, "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].

[229] D. Prato and C. Tsallis, "Nonextensive foundation of Levy distributions", *Phys. Rev. E* **60** (1999) 2398.

[230] **ALICE** Collaboration, S. Acharya *et al.*, "Multiplicity dependence of π, K, and p production in pp collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **80** (2020) 693, arXiv:2003.02394 [nucl-ex].

[231] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of prompt D^0, Λ_c^+, and $\Sigma_\mathrm{c}^{0,++}(2455)$ production in pp collisions at $\sqrt{s} = 13$ TeV", arXiv:2106.08278 [hep-ex].

[232] Y. Chen and M. He, "Charged-particle multiplicity dependence of charm-baryon-to-meson ratio in high-energy proton–proton collisions", *Phys. Lett. B* **815** (2021) 136144, arXiv:2011.14328 [hep-ph].

[233] R. Hagedorn and K. Redlich, "Statistical Thermodynamics in Relativistic Particle and Ion Physics: Canonical or Grand Canonical?", *Z. Phys. C* **27** (1985) 541.

[234] V. Vovchenko, B. Dönigus, and H. Stoecker, "Canonical statistical model analysis of pp, p–Pb, and Pb–Pb collisions at energies available at the CERN Large Hadron Collider", *Phys. Rev. C* **100** (2019) 054906, arXiv:1906.03145 [hep-ph].

[235] **ALICE** Collaboration, S. Acharya *et al.*, "Charged-particle production as a function of multiplicity and transverse spherocity in pp collisions at $\sqrt{s} = 5.02$ and 13 TeV", *Eur. Phys. J. C* **79** (2019) 857, arXiv:1905.07208 [nucl-ex].

[236] S. G. Weber *et al.*, "Elucidating the multiplicity dependence of J/ψ production in proton–proton collisions with PYTHIA8", *Eur. Phys. J. C* **79** (2019) 36, arXiv:1811.07744 [nucl-th].

[237] S. K. Das *et al.*, "Propagation of heavy baryons in heavy-ion collisions", *Phys. Rev. D* **94** (2016) 114039, arXiv:1604.05666 [nucl-th].

[238] **ALICE** Collaboration, S. Acharya *et al.*, "Observation of a multiplicity dependence in the p_T-differential charm baryon-to-meson ratios in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:2111.11948 [nucl-ex].

[239] **STAR** Collaboration, J. Adam *et al.*, "First measurement of Λ_c^+ baryon production in Au+Au collisions at $\sqrt{s_\mathrm{NN}} = 200$ GeV", *Phys. Rev. Lett.* **124** (2020) 172301, arXiv:1910.14628 [nucl-ex].

[240] S. Cho *et al.*, "Charmed hadron production in an improved quark coalescence model", *Phys. Rev. C* **101** (2020) 024909, arXiv:1905.09774 [nucl-th].

[241] J. Zhao *et al.*, "Sequential Coalescence with Charm Conservation in High Energy Nuclear Collisions", arXiv:1805.10858 [hep-ph].

[242] S. Cao *et al.*, "Charmed hadron chemistry in relativistic heavy-ion collisions", *Phys. Lett. B* **807** (2020) 135561, arXiv:1911.00456 [nucl-th].

[243] **ALICE** Collaboration, B. B. Abelev *et al.*, "K_S^0 and Λ production in Pb–Pb collisions at $\sqrt{s_\mathrm{NN}} = 2.76$ TeV", *Phys. Rev. Lett.* **111** (2013) 222301, arXiv:1307.5530 [nucl-ex].

[244] S. Ghosh *et al.*, "Diffusion of Λ_c^+ in hot hadronic medium and its impact on $\Lambda_\mathrm{c}^+/\mathrm{D}$ ratio", *Phys. Rev. D* **90** (2014) 054018, arXiv:1407.5069 [nucl-th].

[245] L. Tolos, J. M. Torres-Rincon, and S. K. Das, "Transport coefficients of heavy baryons", *Phys. Rev. D* **94** (2016) 034018, arXiv:1601.03743 [hep-ph].

[246] A. Andronic *et al.*, "The multiple-charm hierarchy in the statistical hadronization model", *JHEP* **07** (2021) 035, arXiv:2104.12754 [hep-ph].

[247] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of prompt and nonprompt charmonium suppression in Pb–Pb collisions at 5.02 TeV", *Eur. Phys. J. C* **78** (2018) 509, arXiv:1712.08959 [nucl-ex].

[248] **ATLAS** Collaboration, M. Aaboud *et al.*, "Prompt and non-prompt J/ψ and ψ(2S) suppression at high transverse momentum in 5.02 TeV Pb–Pb collisions with the ATLAS experiment", *Eur. Phys. J. C* **78** (2018) 762, arXiv:1805.04077 [nucl-ex].

[249] A. Mazeliauskas *et al.*, "Fast resonance decays in nuclear collisions", *Eur. Phys. J. C* **79** (2019) 284, arXiv:1809.11049 [nucl-th].

[250] O. Andreev, "Baryon modes in string breaking from gauge/string duality", *Phys. Lett. B* **804** (2020) 135406, arXiv:1909.12771 [hep-ph].

[251] F. Cooper and G. Frye, "Comment on the Single Particle Distribution in the Hydrodynamic and Statistical Thermodynamic Models of Multiparticle Production", *Phys. Rev. D* **10** (1974) 186.

[252] K. S. Sateesh, "An Experimental signal for diquarks in quark-gluon plasma", *Phys. Rev. D* **45** (1992) 866–868.

[253] Z. Citron *et al.*, "Report from Working Group 5: Future physics opportunities for high-density QCD at the LHC with heavy-ion and proton beams", *CERN Yellow Rep. Monogr.* **7** (2019) 1159–1410, arXiv:1812.06772 [hep-ph].

[254] **ALICE** Collaboration, B. Abelev *et al.*, "Technical Design Report for the Upgrade of the ALICE Inner Tracking System", *J. Phys. G* **41** (2014) 087002.

[255] **ALICE** Collaboration, S. Acharya *et al.*, "Upgrade of the ALICE Inner Tracking System during LS4: study of physics performance", *in preparation* (2021) .

[256] **ALICE** Collaboration, S. Acharya *et al.*, "Expression of Interest for an ALICE ITS Upgrade in LS3", ALICE-PUBLIC-2018-013.

[257] **ALICE** Collaboration, S. Acharya *et al.*, "Letter of Intent for an ALICE ITS Upgrade in LS3", CERN-LHCC-2019-018.

[258] I. Kuznetsova and J. Rafelski, "Charmed hadrons from strangeness-rich QGP", *J. Phys. G* **32** (2006) S499–S504, arXiv:hep-ph/0605307.

[259] M. He, R. J. Fries, and R. Rapp, "Heavy Flavor at the Large Hadron Collider in a Strong Coupling Approach", *Phys. Lett. B* **735** (2014) 445–450, arXiv:1401.3817 [nucl-th].

[260] **ALICE** Collaboration, J. Adam *et al.*, "Measurement of $\mathrm{D_s^+}$ production and nuclear modification factor in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV", *JHEP* **03** (2016) 082, arXiv:1509.07287 [nucl-ex].

[261] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of $\mathrm{B_s^0}$ meson production in pp and Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Lett. B* **796** (2019) 168–190, arXiv:1810.03022 [hep-ex].

[262] **HFLAV** Collaboration, Y. Amhis *et al.*, "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016", *Eur. Phys. J. C* **77** (2017) 895, arXiv:1612.07233 [hep-ex].

[263] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Open heavy flavor and quarkonia in heavy ion collisions at HL-LHC", CMS-PAS-FTR-18-024.

CHAPTER A

Appendices

A.1 Summary

The work presented in this dissertation is dedicated to production measurements of open charm hadrons in proton–proton (pp) and lead–lead (Pb–Pb) collisions with the ALICE detector at the Large Hadron Collider (LHC) at CERN. Studies on the production of open charm hadrons are of paramount importance to investigate charm-quark hadronisation at the LHC, particularly through the evolution of the production ratio between different charm-hadron species. The production and hadronisation of heavy quarks are both described by quantum chromodynamics (QCD), however, while perturbative QCD calculations of heavy-quark production are possible down to low transverse momenta (p_{T}), the typical momentum transfers in hadronisation processes are too small for meaningful QCD calculations. One thus has to resort to phenomenological models to describe hadronisation. Traditional model implementations of hadronisation like vacuum string fragmentation, tuned on — and therefore in good agreement with — $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collision data, can describe well charm-*meson* production in small hadronic collision systems at the LHC. However, this type of hadronic modelling, without taking into account the parton phase-space density, is no longer valid in presence of a reservoir of partons. In large collision systems like Pb–Pb, where a colour-deconfined quark–gluon plasma (QGP) is formed, implementations of alternative hadronisation mechanisms are required, which predict an enhancement of charm-strange hadrons and charm baryons with respect to non-strange charm mesons. It is essential to achieve a good understanding of the hadronisation phase, since it connects what experimentally can be measured (heavy-flavour hadrons) to what theoretically is seen as a sensitive tool to investigate the fundamental aspects of QCD (heavy quarks traversing the QGP).

The need of a better knowledge of hadronisation is not limited to heavy-ion collisions only, also recent measurements in pp collisions, for which the standard assumption in high-energy physics was for many years that the interactions between the physical degrees of freedom are negligible, indicate that there might be more physics at play than initially thought. Many characteristics believed to be a sign of a colour-deconfined medium have been observed in pp systems nowadays, feeding the discussion if a QGP is formed in small collision systems. Regarding heavy-flavour production, measurements of the lightest charm *baryon* (Λ_{c}^+) in pp and p–Pb collisions revealed unexpected features as well, not in line with any of the state-of-the-art model predictions. Typical calculations that were developed to better reproduce the charm baryon-to-meson ratio ($\Lambda_{\mathrm{c}}^+/\mathrm{D}^0$) also rely on alternative hadronisation mechanisms, where some are directly taken from ideas originally developed to describe hadronisation in presence of a colour-deconfined medium.

This manuscript features three independent experimental analyses of charm-hadron production and one performance simulation study of beauty-hadron production across different systems. In particular, the non-strange D^{*+} vector-meson is measured in minimum-bias pp collisions, the production of the charm-strange $\mathrm{D_s^+}$ meson and charm Λ_{c}^+ baryon are studied as function of multiplicity in pp collisions, the Λ_{c}^+ baryon is measured in central and mid-central Pb–Pb collisions, and a prospect for the measurement of the beauty-strange $\mathrm{B_s^0}$ meson in Pb–Pb collisions is presented. The main insights are shortly discussed in the four paragraphs below.

The p_{T}-differential cross section of D^{*+} mesons (as well as the other D mesons) in minimum-bias pp collisions at $\sqrt{s} = 5.02$ TeV is well described by pQCD calculations based on $\mathrm{e}^{+}\mathrm{e}^{-}$ and $\mathrm{e}^{-}\mathrm{p}$ collision data. The experimental uncertainties on the cross section measurements are now significantly smaller than the theoretical ones. However, since the theoretical uncertainties are dominated by the variations of perturbative scales, the constraining power of the bare cross-section measurements is limited. Ratios of the production cross sections between different collision energies, rapidity regions, or hadron species are in that sense more interesting. The measured vector-to-pseudoscalar meson $\mathrm{D}^{*+}/\mathrm{D}^{0}$ ratio shows the importance to incorporate the recently developed fragmentation functions dedicated to D^{*+} mesons into the pQCD models, while the mid-to-forward rapidity-region ratios for D^{0} and D^{*+} mesons hint to a possible improvement required in the rapidity description of D-meson production. These points are details though, the overall good agreement of the pQCD calculations with data indicate that the factorisation scheme together with the vacuum fragmentation picture is reasonable for charm-meson production. The most important feature of this measurement is therefore the collision energy ($\sqrt{s}$), which is similar per nucleon–nucleon collision as for heavy-ion collisions recorded at the LHC. Hence, this measurements can directly, without relying on a model-dependent $\sqrt{s}$-scaling, be used as baseline for charm-meson measurements in heavy-ion collisions.

The study of $\mathrm{D}_{\mathrm{s}}^{+}$-meson and Λ_{c}^{+}-baryon production as function of charged-particle multiplicity density in pp collisions at $\sqrt{s} = 13$ TeV is directly motivated by the surprising observations in Λ_{c}^{+} production in minimum-bias pp and p–Pb collisions, and the predicted enhancement of $\mathrm{D}_{\mathrm{s}}^{+}$ and Λ_{c}^{+} production with respect to non-strange D mesons by modified hadronisation mechanisms in presence of a QGP. While the measured p_{T}-differential production ratios to D^{0} show no multiplicity dependence for $\mathrm{D}_{\mathrm{s}}^{+}$ mesons, a significant enhancement in the visible p_{T} interval is observed for the $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratio. However, extrapolating the production to $p_{\mathrm{T}} > 0$, this enhancement in the $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratio with multiplicity is no longer visible within the uncertainties. This can either be a genuine feature of the data, indicating that the enhancement with multiplicity in the visible p_{T} interval is simply due to a redistribution of momentum acting differently for baryons and mesons, or it can be a consequence of the relatively large experimental uncertainties on the extrapolated ratios. It is interesting that both the observations for the $\mathrm{D}_{\mathrm{s}}^{+}/\mathrm{D}^{0}$ and $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratios are in line with the observations for the light-flavour equivalent ratios: K^{+}/π^{+} and $\Lambda/\mathrm{K}_{\mathrm{S}}^{0}$, respectively. These similarities, if not accidental, hint at a potential common mechanism for light- and charm-hadron formation in hadronic collisions, contrary to expectations from the vacuum-fragmentation scenario. More precise and more differential measurements of charm-hadron production as function of multiplicity are fundamental to investigate further. In addition, theoretical developments are required to properly address the interplay between the soft and hard components of the event as shown by these measurements.

The measurement of Λ_{c}^{+} production in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV has a wider p_{T} coverage and is more differential in p_{T} and centrality with respect to previous measurements in A–A collisions, and thus provide an important further glance to charm-quark hadronisation in

heavy-ion collisions. The p_{T}-differential $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratios show the predicted enhancement by modified hadronisation mechanisms at intermediate p_{T} ($4 < p_{\mathrm{T}} < 8$ GeV/c). However, where these models predict an overall, p_{T}-integrated enhancement of charm baryons over charm mesons, the measured $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ ratios extrapolated to $p_{\mathrm{T}} > 0$ show, for the fist time, that they do not differ significantly between Pb–Pb and pp collisions. In addition, qualitative similarities with the light-flavour $\Lambda/\mathrm{K}_{\mathrm{S}}^{0}$ ratio can again be noted. While for the light-flavour sector, the discussion on how to interpret these observations is focusing on a radial boost from a hydrodynamic origin in all collision systems, the origin in the heavy-flavour sector is not yet known. It might be that the implementations of hadronisation in models needs to be modified to be able to describe the "contradicting" p_{T}-differential and p_{T}-integrated $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ measurements, or that the enhancement seen at intermediate p_{T} is indeed induced only by an effect like radial flow in the hadronic phase. More precise measurements of Λ_{c}^{+} production in Pb–Pb collisions are crucial and one of ALICE's main goals for LHC Run 3 and 4. Further developments on the theory side are likely required as well, for which it might be interesting to make the connection with the Λ_{c}^{+} production as function of event multiplicity in pp collisions.

Finally, a physics performance study for the measurement of beauty-strange $\mathrm{B}_{\mathrm{s}}^{0}$ meson for the ALICE experiment in LHC Run 3 and 4 in Pb–Pb at $\sqrt{s_{\mathrm{NN}}} = 5.5$ TeV is reported. This simulation study is motivated by the installation of a new Inner Tracking System (ITS) based on monolithic active pixel sensors that will enable the measurement of several heavy-flavour hadrons with unprecedented precision. Another upgrade or the ITS is foreseen in the next long LHC shutdown, bringing another set of crucial observables like charm-strange baryons and beauty-strange mesons in heavy-ion collisions into reach. From the theoretical point of view, beauty hadrons represent a much cleaner probe to study the QGP than charm hadrons, and therefore it is important to extend the in-medium heavy-flavour energy loss and hadronisation picture to the beauty sector. With the estimated expected performance of these upgrade projects, the measurement of the $\mathrm{B}_{\mathrm{s}}^{0}$ meson in Pb–Pb collisions will get feasible and complement, and extend to lower p_{T}, the existing measurement of the CMS Collaboration and their projection for the High Luminosity LHC era. However, the measurement remains challenging. Even with the total expected luminosity of $\mathcal{L}_{\mathrm{int}} = 10\ \mathrm{nb}^{-1}$, the estimated statistical precision of the $\mathrm{B}_{\mathrm{s}}^{0}$ signal is limited with just four p_{T} intervals in reach. An interesting alternative to guarantee a significantly higher statistical significance would be to study non-prompt $\mathrm{D}_{\mathrm{s}}^{+}$ mesons.

The analyses presented in this dissertation, especially the two Λ_{c}^{+} measurements, should help to improve our understanding of charm-quark hadronisation. For both analyses a manuscript is close to being published, which will likely attract some theoretical attention. On the experimental side, the story is not finished yet either. More precise and more differential measurements of the charm-baryon sector are foreseen with the data recorded in LHC Run 3 and 4.

A.2 Samenvatting

In dit proefschrift presenteer ik mijn werk gewijd aan productiemetingen van hadronen opgebouwd uit één charm quark en één of meerdere lichtere quarks in proton–proton (pp) en lood-lood (Pb–Pb) botsingen, gemeten met de ALICE detector in de Large Hadron Collider (LHC) op CERN. Het meten van de productie van charm hadronen is van belang voor het onderzoeken van charm-quark hadronisatie. Dit wordt met name gedaan via metingen van de productieverhouding tussen verschillende soorten charm hadronen. Theoretisch gezien worden zowel de productie en hadronisatie van zware quarks (waaronder de charm quark valt) beschreven met kwantumchromodynamica (QCD), maar waar het productieproces met perturbatie QCD berekend kan worden tot lage transversale momenta (p_{T}), is de momentum overdracht in hadronisatieprocessen te klein voor zinvolle QCD berekeningen. Men is daardoor afhankelijk van fenomenologische modellen om hadronisatie te kunnen beschrijven. Traditionele implementaties van hadronisatie, zoals snaarfragmentatie in vacuüm dat is afgesteld op — en daardoor automatisch in overeenkomst met — $\mathrm{e^+e^-}$ en $\mathrm{e^-p}$ data, kunnen de productie van charm *mesonen* in kleine hadronische botsingssystemen goed beschrijven. Dit type modellering houdt echter geen rekening met de faseruimte dichtheid van de quarks en gluonen (partonen) en is daardoor niet meer geldig bij aanwezigheid van een partonenreservoir. Voor grote botsingssystemen zoals Pb–Pb, waarbij een kleurlading-ongebonden quark–gluonplasma (QGP) wordt gevormd, zijn modellen met alternatieve hadronisatiemechanismen vereist. Dit soort modellen voorspellen een toename in de productie van charm-strange mesonen en charm baryonen ten opzichte van charm mesonen zonder strange quark. Het is essentieel om een goed begrip van die hadronisatiefase te verkrijgen, aangezien dit de verbinding maakt tussen wat experimenteel reconstrueerbaar is (hadronen met zware quarks) en wat theoretisch gezien wordt als een gevoelig instrument om de fundamentele aspecten van QCD mee te onderzoeken (zware quarks die een QGP doorkruisen).

De behoefte aan een betere kennis van het hadronisatieproces beperkt zich niet alleen tot botsingen tussen zware ionen, ook recente metingen in pp botsingen, waarvoor het veld van de hoge-energiefysica jarenlang aannam dat de interacties tussen de fysische vrijheidsgraden verwaarloosbaar waren, geven aan dat er mogelijk meer natuurkunde in het spel is dan aanvankelijk werd gedacht. Veel kenmerken waarvan wordt aangenomen dat ze een teken zijn van een kleurlading-ongebonden medium worden tegenwoordig ook waargenomen bij pp botsingen, wat de discussie voedt of in kleine botsingssystemen ook een QGP wordt gevormd. Ook productiemetingen van de lichtste charm baryon (Λ_{c}^+) in pp en p–Pb botsingen lieten onverwachte kenmerken zien, niet in overeenkomst met voorspellingen van de tot dan toe veel gebruikte modellen. Theoretische berekeningen die in staat zijn de charm baryon-tot-meson productieratio ($\Lambda_{\mathrm{c}}^+/\mathrm{D}^0$) beter te reproduceren zijn gebaseerd op alternatieve hadronisatiemechanismen. Sommigen gebruiken daarbij ideeën die oorspronkelijk waren ontwikkeld om het hadronisatieproces in aanwezigheid van een kleurlading-ongebonden QGP te beschrijven.

In dit manuscript worden drie onafhankelijke experimentele analyses van charm-hadron productie en één performance simulatiestudie van beauty-hadron productie in verschillende botsingssystemen gepresenteerd. Allereerst wordt de meting van het niet-strange D^{*+} vectormeson in

minimum-bias pp botsingen besproken, waarna de productie van charm-strange D_s^+ mesonen en charm Λ_c^+ baryonen als functie van de multipliciteit van geladen deeltjes in pp botsingen, de meting van Λ_c^+ baryonen in centrale en half-centrale Pb–Pb botsingen, en een vooruitzicht voor een mogelijke meting van beauty-strange B_s^0 mesonen in Pb–Pb botsingen worden gepresenteerd. In de vier paragrafen hieronder zullen de belangrijkste inzichten kort worden besproken.

De p_T-differentiële werkzame doorsnede van D^{*+} mesonen (evenals de andere D mesonen) in minimum-bias pp botsingen met $\sqrt{s} = 5.02$ TeV wordt goed beschreven door pQCD berekeningen die zijn gebaseerd op metingen in e^+e^- en e^-p botsingssystemen. De experimentele onzekerheden op deze metingen zijn nu aanzienlijk kleiner dan de theoretische. Echter, aangezien de theoretische onzekerheden worden gedomineerd door de variaties van de perturbatie schalen, is er maar een geringe mogelijkheid om de vrijheden van de theoretische modellen te beperken via enkel de werkzame-doorsnedemetingen. Verhoudingen van de productiedoorsnedes tussen verschillende botsingsenergieën, rapiditeitsgebieden, of hadronensoorten zijn in die zin interessanter. De gemeten vector-tot-pseudoscalaire D^{*+}/D^0 ratio toont het belang aan om de recent ontwikkelde D^{*+} fragmentatiefuncties te implementeren in de pQCD modellen, terwijl de midden-tot-voorwaartse rapiditeitsverhoudingen voor D^0 en D^{*+} mesonen een mogelijke verbetering van de rapiditeitsbeschrijving in D-meson productie aanstippen. Dit zijn details echter, de algehele goede overeenstemming tussen de pQCD berekeningen en metingen geven aan dat het factorisatieschema samen met het idee van vacuüm fragmentatie redelijk is voor het beschrijven van charm-mesonen productie. Het belangrijkste kenmerk van deze meting is daarom de botsingsenergie ($\sqrt{s}$), die hetzelfde is per nucleon–nucleonbotsing als voor de zware-ionen botsingssystemen in de LHC. Deze metingen kunnen dus direct, zonder te hoeven vertrouwen op een model-afhankelijke $\sqrt{s}$-schaling, worden gebruikt als referentie voor charm-meson metingen bij botsingen tussen zware ionen.

Het onderzoek naar de D_s^+-meson en Λ_c^+-baryon productie als functie van de multipliciteitsdichtheid van geladen deeltjes in pp botsingen met $\sqrt{s} = 13$ TeV is gemotiveerd door de verrassende waarnemingen in Λ_c^+ productie in minimum-bias pp en p–Pb botsingen, en door de voorspelde toename van D_s^+ en Λ_c^+ productie ten opzichte van niet-strange D mesonen in alternatieve hadronisatiemodellen in aanwezigheid van een QGP. Waar de gemeten p_T-differentiële productieratios tot D^0 geen multipliciteitsafhankelijkheid laten zien voor D_s^+ mesonen, is er een significante toename in de Λ_c^+/D^0 ratio waar te nemen in het gemeten p_T-bereik. Echter, wanneer de productie wordt geëxtrapoleerd naar $p_T > 0$ is deze toename in de Λ_c^+/D^0 ratio met multipliciteit niet langer zichtbaar binnen de onzekerheden. Dit kan ofwel een echt kenmerk zijn van de data, wat betekend dat de toename met multipliciteit in het zichtbare p_T bereik komt door een herverdeling van momentum dat baryonen en mesonen anders beïnvloed, of het kan simpelweg een gevolg zijn van de relatief grote experimentele onzekerheden op de geëxtrapoleerde ratio's. Het is interessant dat zowel de waarnemingen voor de D_s^+/D^0 als Λ_c^+/D^0 ratio's in lijn zijn met observaties voor de equivalente lichte hadronratio's: K^+/π^+ en Λ/K_S^0, respectievelijk. Als deze overeenkomsten niet op toeval berust zijn, duiden ze op een mogelijk gemeenschappelijk mechanisme voor de vorming van lichte en charm hadronen in hadronische botsingen. Dit is in tegenstelling tot wat

men verwacht bij het vacuümfragmentatiescenario. Nauwkeurige en meer differentiële metingen van charm-hadron productie als functie van multipliciteit van geladen deeltjes zijn essentieel als vervolgstap. Daarnaast zijn er theoretische ontwikkelingen nodig om het samenspel tussen de zachte en harde componenten in een botsing, zoals volgt uit deze metingen, goed te kunnen beschrijven.

De meting van de productie van Λ_c^+ baryonen in Pb–Pb botsingen met $\sqrt{s_{NN}} = 5.02$ TeV heeft een groter p_T-bereik en is meer differentieel in p_T en centraliteit ten opzichte van eerdere metingen in A–A botsingssystemen, en biedt dus een belangrijke verdere kijk op charm-quark hadronisatie bij botsingen tussen zware ionen. De gemeten p_T-differentiële Λ_c^+/D^0 ratio's laten een toename zien in het tussenliggende p_T gebied ($4 < p_T < 8$ GeV/c), zoals voorspeld door alternatieve hadronisatiemechanismen in aanwezigheid van een QGP. Waar deze modellen echter een algehele, p_T-geïntegreerde toename van charm baryonen ten opzichte van charm mesonen voorspellen, laten de gemeten Λ_c^+/D^0 ratio's, geëxtrapoleerd naar $p_T > 0$, voor de eerste keer zien dat de Λ_c^+/D^0 waarden niet significant verschillen tussen Pb–Pb en pp botsingen. Daarnaast zijn er opnieuw kwalitatieve overeenkomsten op te merken met de lichte baryon-tot-meson Λ/K_S^0 ratio. Terwijl voor deze laatstgenoemde productieratio de discussie over de interpretatie van de waarnemingen zich concentreert op een radiale boost uit een hydrodynamische oorsprong voor alle botsingssystemen, is de achterliggende oorzaak voor de eigenschappen van de zware baryon-tot-meson ratio's nog niet bekend. Het zou kunnen dat de implementaties van hadronisatie in de fenomenologische modellen aangepast moet worden om de "tegenstrijdige" p_T-differentiële en p_T-geïntegreerde Λ_c^+/D^0 metingen te kunnen beschrijven, of dat de toename van de Λ_c^+/D^0 ratio in het tussenliggende p_T gebied inderdaad wordt veroorzaakt door een effect zoals radiale stroming in de hadronische fase. Nauwkeurige metingen van Λ_c^+ productie in Pb–Pb botsingen zijn cruciaal en een van de hoofddoelen van het ALICE experiment in LHC Run 3 en 4. Verdere ontwikkelingen aan de theoretische kant zijn waarschijnlijk ook nodig, waarbij het interessant zou kunnen zijn om de connectie te maken met de metingen van de Λ_c^+ productie als functie van geladen-deeltjes multipliciteit in pp botsingen.

Ten slotte wordt een vooruitzicht voor een mogelijke meting met het ALICE experiment in LHC Run 3 en 4 van beauty-strange B_s^0 mesonen in Pb–Pb botsingen met $\sqrt{s_{NN}} = 5.5$ TeV gerapporteerd. Deze simulatiestudie is gemotiveerd door de installatie van een nieuwe Inner Tracking System (ITS) op basis van monolithische actieve pixelsensoren die metingen van zware hadronen met ongekende precisie mogelijk zal maken. Een nieuwe upgrade van de ITS is voorzien in de volgende lange LHC-shutdown, waardoor een reeks nieuwe observabelen in zware-ionen botsingen zoals charm-strange baryonen en beauty-strange mesonen binnen bereik komen. Vanuit theoretisch oogpunt zijn beauty hadronen een beter onderzoeksobject om de QGP te bestuderen dan charm hadronen. Het is daarom belangrijk om het beeld van energieverlies in het medium en hadronisatie van zware quarks uit te breiden naar de beauty sector. Uit de bepaalde vooruitzichten van deze ITS upgradeprojecten blijkt dat de meting van de B_s^0 meson in Pb–Pb botsingen haalbaar wordt en daarbij een aanvulling vormt op, en zorgt voor een uitbreiding naar lagere p_T waarden van de bestaande meting van het CMS experiment en hun projectie voor het High

Luminosity LHC tijdperk. De meting blijft echter een uitdaging. Zelfs met de totale verwachtte statistiek van $\mathcal{L}_{\text{int}} = 10\ \text{nb}^{-1}$, is de geschatte statistische precisie van het B_s^0 signaal beperkt met slechts vier p_T intervallen in bereik. Een interessant alternatief om een significant betere statistische significantie te garanderen zou zijn om D_s^+ mesonen van beauty hadronvervallen te bestuderen.

De analyses die in dit proefschrift gepresenteerd worden, en in het bijzonder de twee Λ_c^+ metingen, zouden ons begrip van charm-quark hadronisatie moeten helpen verbeteren. Voor beide analyses staat een manuscript op het punt om gepubliceerd te worden. Dit zal waarschijnlijk de nodige theoretische aandacht trekken. Echter is het verhaal aan de experimentele kant ook nog niet klaar. Nauwkeurigere en meer differentiële metingen van de charm-baryon sector zijn gepland met de data die zal worden vastgelegd in LHC Run 3 en 4.

A.3 Samenvatting voor leken

Deeltjesfysica, of specifieker, zware-ionen fysica. Mijn onderzoeksveld is geen gemakkelijk veld om op een feestje even kort en bondig uit te leggen. Waar sommigen mijn uitleg leken te begrijpen — of in ieder geval deden alsof — waren het vaker vertwijfelde blikken die ik kreeg te zien. Niks ten nadele van mijn gesprekspartner, mijn uitleg liet ook vaak genoeg te wensen over. In dit hoofdstuk hoop ik het goed te maken, en ieder die me niet kon volgen toch wat bij te brengen over mijn onderzoek.

Of mijn uitleg te volgen was of niet, de vervolgvraag was meestal in de trant van: *"Maar wat heeft de maatschappij hieraan?"*. Aangezien ik verwacht dat bij enkelen die dit lezen deze vraag ook zal opkomen, besteed ik er nu alvast wat zinnen aan. Het nut van toegepast onderzoek zal iedereen wel bekend zijn, het vormt de belangrijkste component voor innovatie en is de sleutel tot de ontwikkeling van onze moderne samenleving. Maar fundamenteel onderzoek, wat in wezen alleen gericht is op kennisverrijking zonder dat er direct concrete toepassingen voor hoeven te zijn, wordt vaker in twijfel getrokken. Mijn onderzoek — waar ook nog eens een 27 kilometer lange en 750 gigawattuur per jaar verbruikende[1] deeltjesversneller en detectors zo groot als een klein vijf-verdiepingen appartementencomplex volgestopt met technologie voor nodig zijn — is hier dan weer een extreem voorbeeld van. Ik zie nu dan ook geen praktische toepassingen voor de fenomenen die ik onderzoek; onze drijfveer is enkel het vergroten van de kennis over het universum. Er zijn echter wel "randzaken" van ons onderzoek waar de maatschappij van profiteert. Bijvoorbeeld het wereldwijde web, ontworpen op CERN om informatie-uitwisseling tussen wetenschappers te vergemakkelijken, of protonenbestraling om tumoren te bestrijden, wat gebruik maakt van de deeltjesversnellertechnologie ontwikkeld voor natuurkundig fundamenteel onderzoek. Daarnaast zijn de mogelijkheden voor het gebruik van de ontwikkelde microchips, bedoeld voor het zo precies mogelijk meten van deeltjes, buiten hoge-energetische fysica vrijwel eindeloos. Ons onderzoek verrijkt dus onze kennis, en de manier waarop we het onderzoeken ons alledaagse leven. Nu we dit opgehelderd hebben, is het tijd om naar de natuurkunde te gaan.

Deeltjesfysici zijn eigenlijk een soort historici. Met onze experimenten proberen we terug te gaan naar de eerste momenten na het ontstaan van het universum, enkele microseconden na de oerknal. Dit klinkt wellicht wat onrealistisch, maar als je weet waar en hoe je moet kijken, is bewijs van ons verre verleden overal te vinden. Een simpel voorbeeld is het licht van een ster dat je oog bereikt. Omdat de afstand tussen een willekeurige ster en de aarde zo enorm is, doet het licht (het foton deeltje) er tientallen miljoenen jaren over om ons te bereiken. We kijken dus eigenlijk het verleden in. Met fotonen kun je terugkijken tot ongeveer 380.000 jaar na de oerknal via de zogenoemde kosmische achtergrondstraling. In de periode hiervoor was het universum ondoorzichtig voor licht, alsof je in dichte mist rijdt en het licht van je koplampen niet verder dan een aantal meter reikt. Via metingen van de kosmische achtergrondstraling van een ander elementair deeltje, het neutrino, zou zelfs teruggekeken kunnen worden tot 1 seconde na de oerknal. Wat er in de eerste seconde is gebeurd is echter significant interessanter. Deze

[1] Ter vergelijking, dit is ongeveer 1.6 keer de hoeveelheid die alle woningen in de gemeente Utrecht per jaar verbruiken.

periode heeft het karakter van het universum van vandaag bepaald, en zou zelfs eventueel kunnen voorspellen hoe het ooit zal eindigen. Deeltjesfysici proberen erachter te komen wat er precies in deze eerste seconde is gebeurd.

Volgens onze huidige theorieën is het universum geboren uit een singulariteit, een natuurkundig begrip wat een gebied beschrijft waar alle massa in een extreem klein punt is geperst, oftewel een oneindig massief en oneindig klein punt. Hoe deze singulariteit is gecreëerd weten we (nog) niet, want per definitie bestond er vóór het begin niets. We hebben echter een vrij goed beeld van wat hierna gebeurde. Direct na het ontstaan bevond het universum zich in een extreem warme (quadriljard graden) en kleine (zo groot als een perzik) staat dat zich in minder dan een oogwenk vergrootte met minstens een factor 10^{26} (een 1 met 26 nullen erachter). Dit proces wordt inflatie genoemd. Toen deze fase afliep, verviel "hetgeen" wat de inflatie veroorzaakte — we noemen dit het inflatieveld, met inflaton deeltjes, maar ondanks dat we het een naam hebben gegeven weten we er vrij weinig van af — en overspoelde het universum met materie en straling.

Het universum was op dit moment nog te heet en te dicht om stabiele deeltjes te vormen. Materie bevond zich dus nog niet in de vorm zoals wij dat kennen van hier en nu. Maar hoe dan wel? Laten we het proces voor het gemak omdraaien. Stel je een ijsblokje voor met daaronder een brander waarmee we het oneindig kunnen opwarmen. Het ijs, waarbij de waterstofmoleculen in een kristalstructuur vast zitten, zal gaan smelten: onder invloed van de stijgende temperatuur zullen de moleculen sneller gaan bewegen en "lostrillen". Deze verandering in hoe materie zich representeert noemen we een faseovergang. De volgende overgang, van water naar waterdamp, volgt als de moleculen zo snel beginnen te bewegen dat ze ontsnappen aan de kracht die de vloeistof bij elkaar houdt. Het stopt echter niet bij deze bel waterdamp met vrij bewegende waterstofmoleculen. Als we de temperatuur nog verder opschroeven zullen ook de moleculen gaan "smelten". Oftewel, de onderdelen waaruit het molecuul opgebouwd is (elektronen, protonen, en neutronen) trillen los en gaan vrij van elkaar bewegen. Deze staat van materie noemen we een plasma, en wordt bereikt rond de tienduizend graden Celsius. Dit is nog een relatief lage temperatuur, en plasma's komen dan ook nog vrij vaak voor in de natuur. De zon is bijvoorbeeld in plasmatoestand, maar ook bliksem is een plasma. Bij nog veel hogere temperaturen zal dit smeltproces ook plaatsvinden voor protonen en neutronen, waarna ons ijsblokje veranderd zal zijn in een gigantische mix van quarks (de fundamentele bouwstenen van protonen en neutronen) en gluonen (de deeltjes die ervoor zorgen dat quarks bij normale temperaturen vast blijven zitten in de proton- en neutronstructuur). Deze mix noemen we het quark–gluonplasma, en ons ijsblokje zal zo rond de twee biljoen graden Celsius moeten zijn voordat we deze staat van materie zullen bereiken.

Terug naar de ontwikkeling van ons universum. Direct na de inflatiefase was de temperatuur zo hoog dat alle materie in het universum zich als een quark–gluonplasma representeerde. De eigenschappen van dit plasma hebben dus bepaald hoe alle materie in het heelal zich heeft ontwikkeld. Het universum zelf bleef verder uitzetten en afkoelen, dus het hierboven beschreven opwarmproces van een ijsblokje vond andersom (en in het groot) plaats. Dit ging ongeveer als volgt. Nog steeds binnen de eerste seconde was het universum voldoende afgekoeld waarna

quarks en gluonen samenklonterden in protonen en neutronen, die toen het universum ongeveer 2 minuten oud was atoomkernen begonnen te vormen. Na 380.000 jaar, wanneer het universum nog zo'n 3000 graden is, binden protonen en elektronen met elkaar en worden de eerste neutrale atomen gevormd. Het duurt nog zeker een paar honderd miljoen jaar voordat de eerste sterren zijn gevormd.

Nu vraag je je misschien af wat dit te maken heeft met deeltjesfysica, waar we deeltjes versnellen, op elkaar schieten, en nieuwe deeltjes meten die uit de botsing ontstaan. Hoewel de condities (bijvoorbeeld de bereikte temperatuur en dichtheid) bij zulke botsingsexperimenten anders zijn dan vlak na de oerknal, zijn de natuurwetten hetzelfde. Ze hebben alleen verschillende effecten op verschillende energieschalen. Een botsing tussen twee — tot bijna de lichtsnelheid versnelde — deeltjes bootst dus sommige aspecten van het vroege universum na, en kan ons daardoor iets vertellen over hoe het was. De twee grootste laboratoria waar dit soort experimenten uitgevoerd worden zijn RHIC (Relativistic Heavy Ion Collider) in New York en de LHC (Large Hadron Collider) op CERN in Genève. Hierbij moet je denken aan een paar duizend wetenschappers die samenwerken in één experiment. Het experiment waar ik aan werk, het ALICE experiment aan de LHC, is gewijd aan de studie van de theorie achter de sterke (kern)kracht, de theorie die het bestaan van het quark–gluonplasma postuleert. Dit onderzoek is al enkele tientallen jaren aan de gang, en we leren steeds meer over de eigenschappen van het quark–gluonplasma. Twee in het oog springende eigenschappen zijn als volgt: het plasma heeft een enorm hoge dichtheid (een kubieke centimeter van dit spul zal ongeveer 40 biljoen kilogram wegen) en het gedraagt zich als een bijna perfecte vloeistof (het stroomt met extreem lage relatieve schuifviscositeit[2]).

Intermezzo: Het standaardmodel van de deeltjesfysica

Het is eigenlijk niet mogelijk om mijn onderzoek goed uit te leggen zonder eerst een introductie te geven van de theorie achter deeltjesfysica. Dit intermezzo geeft daarom een (relatief) korte en bondige samenvatting van de wondere wereld van de elementaire deeltjes en krachten. De informatiedichtheid is echter vrij hoog, dus een advies voor de echte leek: bijt jezelf hier niet te erg op vast! Het is niet mijn bedoeling om alle lezers al te verliezen voordat ik aan de beschrijving van mijn eigen onderzoek toegekomen ben. Onderaan dit kader heb ik een schematisch overzicht toegevoegd als naslagwerk van hoe de besproken deeltjes, krachten, en theorieën zich tot elkaar verhouden.

Het standaardmodel van de deeltjesfysica is de theorie die drie van de vier bekende natuurkrachten beschrijft (de elektromagnetische, de zwakke, en de sterke interactie, maar niet de zwaartekracht) en alle tot dusver gemeten elementaire deeltjes classificeert. De theorie is opgesteld in de jaren zeventig, waarna de bevestiging van het bestaan van de top-quark (1995), het tau-neutrino (2000), en het Higgsdeeltje (2012) het standaardmodel verder heeft versterkt.

[2]Viscositeit is de interne wrijving van een vloeistof. Honing (langzaam stromend) heeft bijvoorbeeld een hoge viscositeit, terwijl water (snelstromend) een lagere viscositeitswaarde heeft. Vloeistoffen zonder enige viscositeit (ook wel supervloeistoffen genoemd) hebben vreemde eigenschappen. Als je bijvoorbeeld in zo'n vloeistof roert zal het oneindig blijven draaien.

Het standaardmodel omvat verschillende type elementaire deeltjes die van elkaar onderscheiden kunnen worden door kenmerken zoals *elektrische lading*, *spin* (een eigenschap die je kan vergelijken met de draaiing van een deeltje om zijn as, maar dan de kwantummechanische versie), of *kleurlading* (vergelijkbaar met elektrische lading en volledig losstaand van de visuele perceptie van kleur). Samengevat bestaan er de volgende groepen:

Elementaire fermionen: deeltjes met halftallige spin ($s = 1/2, 3/2, 5/2, ...$).

- *Quarks en antiquarks.* Quarks zijn de basisbouwstenen voor protonen en neutronen, de deeltjes waaruit de kern van een atoom is opgebouwd. Er zijn zes soorten quarks met de "smaken": up, down, strange, charm, beauty, en top. Deze smaken kunnen gepaard worden in drie generaties (up–down, charm–strange, en top–beauty) waarbij de massa significant groter wordt in elke volgende generatie. Alleen materie bestaande uit quarks uit de eerste generatie is stabiel, deeltjes die een quark uit de tweede of derde generatie bevatten vervallen heel snel in deeltjes van de eerste generatie. Voor elke quark bestaat er als antideeltje een antiquark, met tegengestelde lading en smaak (anti-charm bijvoorbeeld). Quarks zijn nooit individueel waargenomen, maar alleen in samengestelde deeltjes die we hadronen noemen. Er zijn twee type hadronen: een meson opgebouwd uit een quark en antiquark; en een baryon die bestaan uit drie quarks of drie antiquarks. Protonen en neutronen zijn voorbeelden van baryonen. Quarks hebben kleurlading (rood, groen, of blauw), maar hadronen zijn "kleurneutraal". Kleurneutraal betekent bijvoorbeeld dat de quarks in het hadron rood–antirood (voor mesonen) of rood–groen–blauw (voor baryonen) gekleurd zijn.

- *Leptonen en antileptonen.* Leptonen komen ook in drie generaties, waarbij elke generatie bestaat uit een geladen lepton en een ongeladen neutrino. De elektron en het elektron-neutrino zijn de bekendste, en vormen de eerste generatie. Hierna volgen respectievelijk de muon en de tau, de zwaardere broertjes van het elektron, samen met de muon-neutrino en tau-neutrino. Elk lepton heeft een antideeltje, dus in totaal zijn er twaalf leptonen bekend. De geladen leptonen, en dan voornamelijk het elektron, kunnen binden met andere deeltjes om samengestelde deeltjes zoals een atoom te vormen. De neutrino's daarentegen interacteren bijna niet en kunnen daardoor maar moeilijk gemeten worden. Op dit moment vliegen er bijvoorbeeld ongeveer 100 miljard neutrino's door het topje van je vinger zonder dat er iets gebeurt.

Elementaire bosonen: deeltjes met integere spin ($s = 0, 1, 2, ...$).

- *IJkbosonen.* Elementaire deeltjes oefenen krachten op elkaar uit via de uitwisseling van ijkbosonen, die fungeren als de "drager" van een van de fundamentele natuurkrachten. Je kunt hierbij denken aan bowlen als analogie uit de klassieke fysica, waar de speler en de kegel twee type elementaire deeltjes zijn, de bowlingbal het

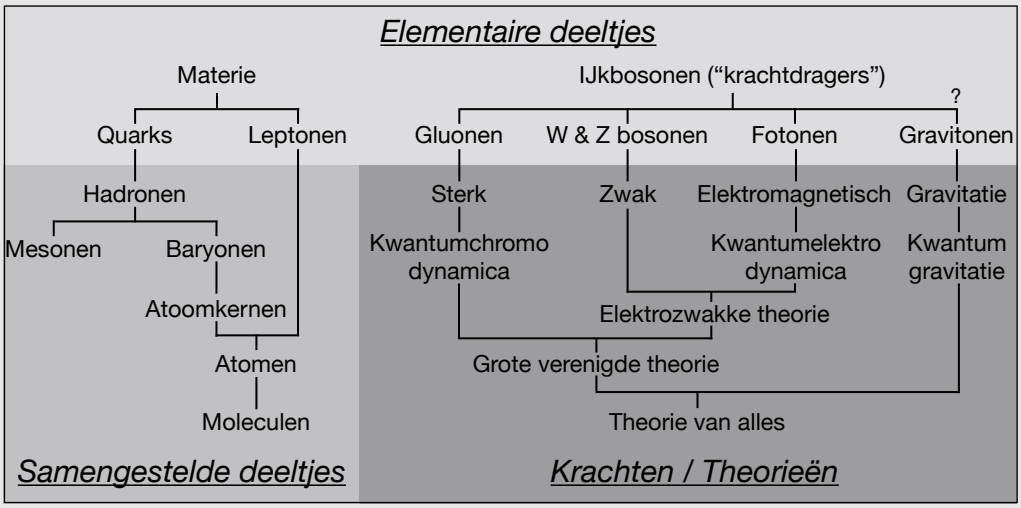
ijkboson is die de kracht van de speler op de kegel overdraagt, en de bowlingbaan het ijkveld representeert. Het standaardmodel omvat drie soorten ijkbosonen: fotonen (voor de elektromagnetische interactie), $W^{\pm}$ en Z^0 bosonen (voor de zwakke interactie), en gluonen (voor de sterke interactie).

- *Scalaire bosonen.* Het standaardmodel omvat maar één scalaire boson, het Higgsdeeltje. Dit deeltje, of beter gezegd het higgsveld, is van fundamenteel belang: het zorgt ervoor dat alle andere deeltjes (behalve het foton en gluon) massa krijgen. Ieder deeltje heeft zijn eigen kwantumveld, maar waar bijvoorbeeld het elektronveld alleen geconcentreerd is op bepaalde plekken (wat we elektronen noemen) en nul is in de rest van de ruimte, heeft het higgsveld de unieke eigenschap dat het overal een hoge waarde heeft. Door deze niet-nulwaarde in vacuüm botst elk deeltje dat met het higgsveld kan interacteren er voortdurend tegenop. Een massaloos deeltje dat met de lichtsnelheid beweegt, maar tussendoor voortdurend tegen het higgsveld aanbotst en daarbij heen en weer geslingerd wordt, zal uiteindelijk (als je alles bij elkaar optelt) langzamer dan het licht lijken te bewegen. En aangezien alleen deeltjes met massa langzamer mogen bewegen dan de lichtsnelheid, zorgt de hoge waarde van het higgsveld in vacuüm er dus voor dat ons massaloze deeltje er nu uitziet en gedraagt alsof het massa heeft. Het Higgsdeeltje is een verdere excitatie van het higgsveld, en zijn ontdekking in 2012 was het langverwachte bewijs van het bestaan van het hier beschreven Higgs mechanisme.

Naast de elementaire deeltjes, kennen we op dit moment vier fundamentele krachten: de *gravitatie* en *elektromagnetische* kracht, werkend over lange afstanden en met effecten die we in het dagelijks leven direct kunnen zien, en de *sterke* en *zwakke* kracht, die krachten uitoefenen op subatomaire schaal (kleiner dan een atoom). De sterke kracht wordt gedragen door het gluon, en is verantwoordelijk voor het binden van quarks in hadronen. De sterke kernkracht, dat protonen en neutronen in atoomkernen bindt, is een "resteffect" van de sterke kracht. De zwakke interactie wordt overgebracht door de $W^{\pm}$ en Z^0 bosonen, en beïnvloedt alle fermionen en het Higgsdeeltje. Het is de enige kracht die smaken van quarks kan veranderen en de reden voor radioactief bètaverval van atomen (een proces waarbij protonen en neutronen in elkaar worden omgezet). De elektromagnetische kracht, gedragen door het foton, is toepasbaar op alle geladen deeltjes, vormt elektrische en magnetische velden, en is de basis voor de elektronische technologie. De gravitatiekracht (oftwel zwaartekracht) wordt niet beschreven door het standaardmodel, maar is, zoals bij iedereen wel bekend, verantwoordelijk voor het bij elkaar houden van grootschalige structuren in het universum. Het hypothetische graviton deeltje wordt vaak gezien als de drager van deze kracht.

Elke interactie kan mathematisch beschreven worden als een veld, via de daarbij behorende kwantumveldentheorie. Kwantumchromodynamica en kwantumelektrodynamica

zijn bijvoorbeeld de theorieën voor respectievelijk de sterke en elektromagnetische interacties. Ook al lijken de vier fundamentele natuurkrachten op alledaagse lage energieën zeer verschillend, geloven veel natuurkundigen dat de krachten gerelateerd zijn en vanaf een zekere energie samenvoegen in één enkele kracht. Dit is al geobserveerd voor de zwakke en elektromagnetische kracht, waarvoor de unificatie-energie relatief laag is. De *elektrozwakke wisselwerking* is de geünificeerde beschrijving van deze twee krachten. Het idee is dat op een gegeven energie kwantumchromodynamica zich hierbij zal voegen (beschreven door de zogenoemde *"grote verenigde theorie"*) en later, op een nog veel hogere energie, ook kwantumgravitatie. Dit zou dan beschreven worden door de (nog niet bestaande) *"theorie van alles"*.



Mijn onderzoek richt zich op kwantumchromodynamica (afgekort met QCD), de theorie die de sterke interactie tussen deeltjes met kleurlading beschrijft. De theorie is in de basis vrij simpel: een quark kan een gluon uitzenden of absorberen; een gluon kan een gluon uitzenden of absorberen; en twee gluonen kunnen direct interageren. De hele theorie van QCD kan bondig worden samengevat met een lagrangiaan: een wiskundige formule waaruit alle dynamica tussen quarks en gluonen volgt. Voor het beschrijven — niet direct hetzelfde als 'begrijpen' — van een willekeurig natuurkundig proces met quarks en gluonen is het voldoende om de lagrangiaan voor de specifieke situatie op te stellen en de bewegingsvergelijkingen[3] op te lossen. Het analytisch oplossen van deze vergelijkingen wordt echter vrij snel onpraktisch. Het kan een heel PhD-traject duren om tot een oplossing te komen voor een vrij simpel probleem. Dit komt omdat de koppelingsconstante, een getal dat de sterkte van een enkele basisinteractie bepaalt, logaritmisch afneemt voor de sterke kracht (een direct effect van het feit dat gluonen met zichzelf kunnen interacteren). Oftewel, de kans dat een proces plaatsvindt neemt toe met afnemende energie of toenemende afstanden tussen de quarks en/of gluonen. Dit zorgt ervoor dat storingstheorie, een veelgebruikt natuurkundig "trucje" waarbij de kansberekening na het meenemen van n

[3]Simpel gezegd zijn deze vergelijkingen de kwantummechanische versie van de bewegingswetten van Newton ($F = m \cdot a$ en actie $\leftrightarrow$ reactie), misschien nog wel bekend van natuurkunde op de middelbare school.

basisinteracties wordt gestopt, niet meer werkt[4]. Dat een QCD-berekening lastig kan zijn wordt bewezen door het feit dat er via de Millenniumprijsproblemen 1 miljoen dollar is uitgeloofd voor een analytisch bewijs van een van de meest karakteristieke eigenschappen van QCD: dat quarks en gluonen niet uit hadronen bevrijd kunnen worden. Dit maakt dat experimentele observaties erg belangrijk zijn voor een beter begrip van kwantumchromodynamica. Zo ook voor het onderzoeken van de eigenschappen van het quark–gluonplasma. Het bestaan van een quark–gluonplasma kan nog worden bewezen via de vergelijkingen van QCD, maar er zijn op dit moment laboratoriumexperimenten nodig om de eigenschappen van deze vorm van materie te onderzoeken.

Dit soort onderzoek vindt plaats bij de RHIC en LHC deeltjesversnellers via botsingen tussen zware ionen (atomen gestript van alle elektronen) die versneld zijn tot bijna de lichtsnelheid. In zulke botsingen worden kleine, kortstondige gebieden geproduceerd waarvan de energiedichtheid vergelijkbaar is met die van het enkele microseconden oude universum, oftewel er wordt een klein beetje quark–gluonplasma gevormd. Deze fase duurt echter extreem kort, na 10^{-23} seconden (honderd triljardste van een seconde) is het plasma alweer verdampt en worden de quarks en gluonen "opgesloten" in hadronen via een proces dat we hadronisatie noemen. We kunnen het quark–gluonplasma dus niet direct waarnemen, maar moeten door middel van het meten van hadronen terug reconstrueren wat er gebeurd is. Sommige type deeltjes zijn hier handiger voor dan anderen of zijn gevoeliger voor andere aspecten van het gevormde medium. Mijn onderzoek gebruikt zware hadronen: deeltjes bestaande uit één charm (of anti-charm) quark en één of meerdere lichtere quarks. Er zijn meerdere redenen waarom zware hadronen goed zijn voor het onderzoeken van het quark–gluonplasma, maar de belangrijkste is dat de productietijd korter en leeftijd van charm quarks langer is dan die van het plasma, en daardoor dus de complete evolutie van het medium meemaken. Dit is niet het geval voor lichtere quarks (worden voornamelijk geproduceerd door het medium zelf) en de top quark (vervalt eerder dan dat het medium hadroniseert). Hadronen met beauty quarks zijn theoretisch gezien een nog beter "meetinstrument", echter zijn deze experimenteel nog niet met voldoende precisie meetbaar.

Het meten van de productie van een zwaar hadron is nog niet voldoende om iets te kunnen zeggen over de eigenschappen van het quark–gluonplasma. Omdat het een indirecte waarneming betreft, is een goed begrip van alle fases die plaatsvinden tijdens een zware-ionen botsing vereist. De initiële distributies van protonen en neutronen in de botsende lood ionen of de specificaties van het charm-quark hadronisatieproces zullen ook invloed hebben op de productie. Om toch iets zinnigs te kunnen zeggen over wat er in de quark–gluonplasma fase is gebeurd, wordt een meting van de productie (of productieverhouding tussen verschillende types hadronen of botsingssystemen) meestal vergeleken met meerdere fenomenologische modellen die de gehele

[4]Het stoppen van de berekening na n basisinteracties is alleen mogelijk als de koppelingsconstante van de interactie altijd veel kleiner dan 1 is, waardoor de kans dat er $(n+1)$ basisinteracties plaatsvinden altijd kleiner is dan de kans op n koppelingen. De kans dat een bepaalde configuratie van interacties tussen deeltjes plaatsvindt is namelijk proportioneel aan de koppelingsconstante tot de macht van het aantal basisinteracties. Bij een waarde veel kleiner dan 1, maakt het voor een zekere n niet meer uit of je de hogere orders wel of niet meeneemt in de berekening. Het effect van het meenemen van deze hogere orders is dan verwaarloosbaar op het eindresultaat van de berekening. Dit is bijvoorbeeld het geval voor berekeningen met kwantumelektrodynamica.

botsing simuleren. Op basis van de overeenkomsten tussen de meting en de modelvoorspellingen worden er conclusies over de onderliggende natuurkunde processen gemaakt. Mijn onderzoek richt zich specifiek op de implementaties van het hadronisatieproces in deze modellen, waarvoor zeer uiteenlopende ideëen bestaan. Het is essentieel om een goed begrip van die hadronisatiefase te verkrijgen, aangezien dit de verbinding maakt tussen wat experimenteel meetbaar is (hadronen met zware quarks) en wat theoretisch gezien wordt als een goed instrument om de eigenschappen van het quark–gluonplasma mee te bepalen (zware quarks die het plasma doorkruisen).

Het is geen simpele taak om de productie van een deeltje te meten, vandaar dat aan één experiment op CERN ook een paar duizend wetenschappers werken. Denk hierbij aan technici die aan de detectoren werken, IT-specialisten die zorgen dat de data opgeslagen en geanalyseerd kan worden, data-analisten (waaronder ik val) die specifieke analyses op de data uitvoeren, theoreten die de metingen pogen te interpreteren, tot managers die alles in goede banen proberen te leiden. De data is dan ook al uitgebreid onder handen genomen voordat ik aan mijn analyses kan beginnen. Het begint allemaal bij de deeltjesversneller zelf, waarbij twee stralen met protonen of lood-ionen in tegengestelde richting worden versneld tot bijna de lichtsnelheid. Als de energie in de versneller is opgelopen tot het gewenste niveau, worden de twee stralen met magneten op elkaar gericht zodat er botsingen kunnen plaatsvinden in het midden van de detector. Uit zulke botsingen ontstaan nieuwe deeltjes, die voor de botsing nog niet bestonden, en alle kanten op vliegen. In een proton–proton (pp) botsing zullen dit, na het hadronisatieproces, ongeveer 10 tot 100 nieuwe deeltjes zijn. Voor een lood–lood (Pb–Pb) botsing moet je denken aan de orde van grootte 1000 tot 10000. Zoals ik al eerder schreef, vliegen deze deeltjes door een detector zo groot als een klein appartementencomplex volgestopt met sensoren. Deze sensoren kunnen *"ja ik zag iets"* zeggen (wat we een hit noemen) of stil blijven. Via reconstructie algoritmes wordt er van alle opgenomen hits een driedimensionaal beeld gemaakt van de botsing, oftewel de banen die de deeltjes aflegden toen ze door de detector vlogen. Een baan die precies in het midden van de detector uitkomt zal waarschijnlijk van een deeltje zijn dat direct in de botsing geproduceerd is, een baan die een meter verderop pas begint zal het product van een verval van een onstabiel deeltje zijn.

De eerste stap van een data-analyse is om uit deze driedimensionale beelden van botsingen — waar we er meestal een paar honderd miljoen van hebben — op zoek te gaan naar een specifiek type deeltje. Sommige analyses hebben geluk, en zoeken naar een soort deeltje dat elke botsing wel een aantal keren geproduceerd wordt. Anderen, zoals mijn analyses naar zware hadronen, zoeken naar een speld in een hooiberg. Een belangrijk onderdeel van zulke analyses is dan ook het scheiden van signaal en achtergrond door het toepassen van selectiecriteria op eigenschappen van de gemeten banen. Voor de meesten analyses in dit manuscript worden deze selectiecriteria geoptimaliseerd met machine learning technieken, maar, met wat meer moeite, kan het ook met de hand. Als het signaal eenmaal uit de data onttrokken is, volgen er nog meerdere correctieprocedures om de meting van de productie invariant te maken (zodat het niet afhankelijk van de selectiecriteria, experimentele opstelling, of grootte van de dataset is). Dit gecorrigeerde signaal van de productie van een type deeltje wordt uiteindelijk gebruikt om metingen mee te doen.

Het bepalen van de onzekerheden op de meting is misschien nog wel belangrijker dan het bepalen van de centrale waarden van de meting zelf, en daar gaat dan ook veel tijd in zitten. Er zijn statistische onzekerheden, die afhankelijk zijn van de grootte van de dataset (meer data geeft een kleinere onzekerheid), en systematische onzekerheden, die volgen uit onzekerheden en/of onvolmaaktheden in de correctieprocedures. De onzekerheden zijn van belang voor het claimen van een ontdekking. In deeltjesfysica is de conventie dat een 5-sigma resultaat — wat betekent dat de kans dat de meting een (statistische) fluctuatie betreft een kans van 1 op 3.5 miljoen is — als een ontdekking gepresenteerd mag worden. Een 3-sigma (99.73% zeker) resultaat noemen we meestal een hint van een observatie, maar wat nog bevestigd moet worden met een toekomstige analyse.

Intermezzo: De bergingswerkers en forensisch onderzoekers van CERN

Het nadeel van een vooraf geschreven uitleg is dat je niet kunt zien of de lezer je nog kan volgen of niet. Ik kan me voorstellen dat bovenstaande paragrafen over — de pracht van — experimenteel onderzoek op CERN voor sommigen niet helemaal zijn binnengekomen. In dit intermezzo stappen we daarom van de werkelijkheid af, en beschrijven we via een metafoor wat we op CERN, en specifiek in mijn experiment, doen.

De Large Hadron Collider op CERN is eigenlijk niets anders dan een snelweg met snel rijdende auto's en vrachtwagens. De auto's representeren de *versnelde protonen*, en de vrachtwagens de veel zwaardere *lood-ionen*. Om het vervolg iets minder luguber te maken gaan we ervan uit dat we te maken hebben met zelfrijdende voertuigen zonder passagiers. Aan de snelweg ligt een archeologische opgraving waar een enorm keramisch beeld uit de oudheid is gevonden. Het beeld staat vol met symbolen en het vermoeden is dat er ergens een verborgen boodschap in is verwerkt die ons veel kan leren over de beschaving van vroeger. De verborgen boodschap symboliseert een *eigenschap van het quark–gluonplasma*, en de vroegere beschaving het *ontstaan van het universum*. Om de boodschap te ontcijferen wordt het beeld met een vrachtwagen naar een instituut gespecialiseerd in symbolen uit de oudheid verplaatst. Op weg naar het instituut gaat het echter helemaal mis en botst de vrachtwagen met het beeld frontaal op een andere vrachtwagen. De snelweg ligt bezaaid met brokstukken en het team van CERN wordt opgeroepen om de situatie te analyseren.

Een uur na het ongeval komen de bergingswerkers (*de technici*) aan, die de positie, grootte, en vorm van elk brokstuk documenteren. Al deze informatie van het plaats van ongeval (*de hits in de detector*) wordt door een team *IT-specialisten* gedigitaliseerd en omgezet in een simplistische simulatievideo van de botsing. Denk hierbij aan een video die gebruikt kan worden om een schuldige aan te wijzen op basis van hoe de brokstukken zijn weggevlogen (*het driedimensionale beeld met banen van deeltjes*), maar die nog niet gedetailleerd genoeg is om een reconstructie te maken van de lading in de vrachtwagens. Dit is de taak van een forensisch onderzoeker (*oftewel ikzelf, een data-analist*), die in de informatie over de duizenden brokstukken op zoek moet gaan naar delen van het keramische beeld met als doel om de verborgen boodschap alsnog te vinden. Sommige stukjes keramiek (*type*

deeltjes) zullen hiervoor interessanter zijn dan anderen. De kans dat een relatief groot restant van het beeld (*een zwaar hadron*) de verborgen boodschap bevat zal namelijk groter zijn dan voor hele kleine stukjes keramiek (*lichte hadronen*). Het loont dus om als forensisch onderzoeker te focussen op een bepaald type broksstuk, in dit geval de grote stukken van het beeld (*een data-analyse naar zware hadronen*). Net zoals de verhouding tussen de productie van lichte en zware hadronen in een lood–lood botsing, zullen er minder grote brokstukken te vinden zijn dan kleine. Hierna houdt de vergelijking echter op, aangezien het aanzienlijk makkelijker is om een groot brokstuk op het plaats van ongeval te vinden ten opzichte van een kleine (terwijl het significant moeilijker is om een zwaar hadron te meten). Het *selectiecriterium nodig om signaal van achtergrond te scheiden* is in dit geval dus simpelweg een filter in de database op de grootte van de brokstukken. Met een beetje geluk worden er een aantal grote stukken keramiek gevonden en bevat één van deze de verborgen boodschap (*een experimentele meting*). Deze brokstukken van het beeld kunnen dan alsnog naar de medewerkers van het instituut (*de theoreten*) gestuurd worden zodat de verborgen boodschap ontcijferd te worden (*meting-versus-model vergelijkingen*).

Er zijn, naast puur de botsing tussen beide vrachtwagens, meer zaken belangrijk om het plaatje compleet te maken. Denk hierbij aan de snelheden van de vrachtwagens en het gedrag van omliggend verkeer vlak voor de botsing (*de initiële distributies van protonen en neutronen in de botsende lood ionen*). Maar ook wat er in het uur na de botsing totdat de bergingswerkers arriveerde is gebeurd (*het hadronisatieproces*). Zijn er stukken keramiek na de botsing verder beschadigd door externe factoren als brand, lekkende brandstof, of omstanders. Dit zijn allemaal factoren die de medewerkers van het instituut moeten weten om de verborgen boodschap zo goed mogelijk te kunnen ontcijferen (*de implementatie van randprocessen in de fenomenologische modellen*). Mijn onderzoek richt zich specifiek op deze tweede externe factor: wat is er gebeurd tussen de botsing zelf en het moment dat de bergingsmedewerkers de brokstukken hebben gedocumenteerd?

Ondanks dat ik dit een aardig gevonden metafoor voor ons onderzoek vind — de inspiratie kwam van een gevaarlijke rotonde naast CERN met de bijnaam de "Large Car Collider" — blijft het natuurlijk een oversimplificatie van wat we echt doen. We kijken bijvoorbeeld niet naar een enkele botsing, maar naar honderden miljoenen tegelijkertijd. En in plaats van verzekerd resultaat (het al gevonden beeld met onbekende symbolen), weten we van tevoren niet of hetgeen we zoeken ook echt interessant zal zijn.

In dit manuscript worden drie onafhankelijke experimentele analyses van charm-hadron productie en één performance simulatiestudie van beauty-hadron productie in verschillende botsingssystemen gepresenteerd. Allereerst wordt de meting van het D^{*+} meson (een meson bestaande uit een charm en anti-down quark in geëxciteerde staat) in pp botsingen besproken. Hierna worden de productie van D_s^+ mesonen (opgebouwd uit een charm en anti-strange quark) en Λ_c^+

baryonen (bestaande uit een charm, up, en down quark) als functie van de multipliciteit van geladen deeltjes in pp botsingen en de meting van Λ_c^+ baryonen in Pb–Pb botsingen gepresenteerd. Dit werk eindigt met een vooruitzicht voor een mogelijke meting van B_s^0 mesonen (bestaande uit een strange en anti-beauty quark) in Pb–Pb botsingen. Deze analyses zijn allen gemotiveerd om de spreiding tussen — en de onzekerheden in — de verschillende modelvoorspellingen te verminderen, waarmee de implementatie van het hadronisatieproces verbeterd kan worden. Waarom specifiek deze metingen zijn gekozen, en de technische details over hoe ze zijn uitgevoerd, liggen helaas buiten de bedoeling van deze samenvatting. In de volgende paragrafen zal ik wel mijn belangrijkste bevindingen kort samenvatten.

De productiemeting van het D^{*+} meson toont aan dat het, met de experimentele precisie van tegenwoordig, van belang wordt om de totale spin van het charm meson mee te nemen in de beschrijving van het hadronisatieproces van charm quarks. Dit is echter een detail die zichtbaar wordt door de steeds betere experimentele precisie, over het algemeen laat deze meting zien dat de productie van charm *mesonen* vrij goed beschreven wordt door modellen. De waargenomen multipliciteitsafhankelijkheid in de D_s^+-meson en Λ_c^+-baryon productiemeting laat echter zien dat er nog wel theoretische ontwikkelingen nodig zijn om het samenspel tusssen de productie van charm *baryonen* en de multipliciteit van geladen deeltjes in de botsing consistent te beschrijven. Daarnaast is het interessant dat deze metingen kwalitatief in lijn zijn met observaties voor hadronen zonder zware quarks, wat in tegenstelling is tot wat men verwacht met het standaard hadronisatiescenario. Ook de productiemeting van het Λ_c^+ baryon in Pb–Pb botsingen laat iets verrassends zien: de door alle modellen van tevoren voorspelde algehele toename van charm baryonen ten opzichte van charm mesonen in Pb–Pb botsingen is in data niet zichtbaar (binnen de experimentele onzekerheden). Dit kan meerdere dingen betekenen, maar wijst er waarschijnlijk op dat het charm hadronisatieproces in aanwezigheid van een quark–gluonplasma nog niet volledig begrepen is. Ten slotte laat de simulatiestudie voor de meting van het B_s^0 meson in Pb–Pb botsingen zien dat het mogelijk wordt om deze meting in de komende jaren met het ALICE experiment uit te voeren. Dit betekent dat we het quark–gluonplasma binnenkort ook kunnen gaan onderzoeken met de — theoretisch gezien betere — beauty quarks.

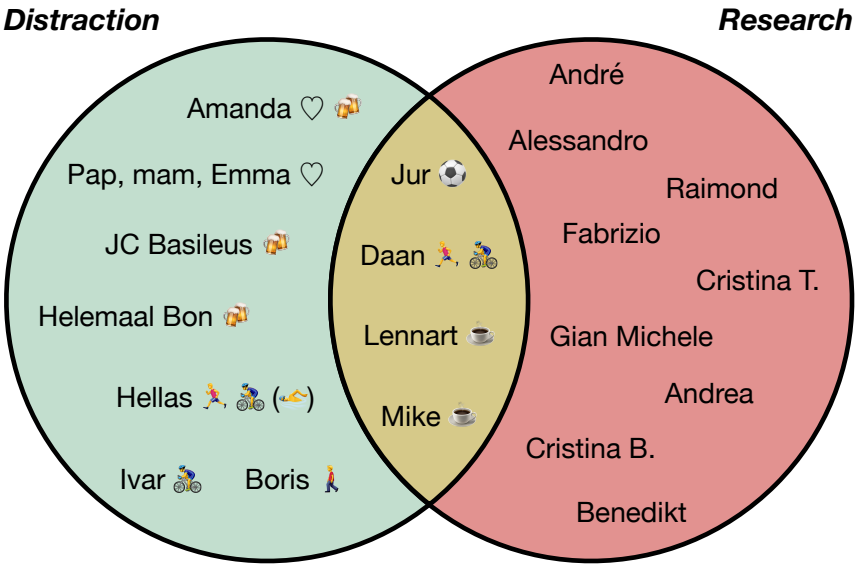
De analyses die in dit proefschrift gepresenteerd worden, en in het bijzonder de twee Λ_c^+ metingen, zijn belangrijke nieuwe resultaten voor ons begrip van het charm hadronisatieproces. Voor beide analyses staat een manuscript op het punt om gepubliceerd te worden en dit zal, zoals in de vorige paragraaf al beschreven, waarschijnlijk de nodige theoretische aandacht trekken. Een betere beschrijving van het hadronisatieproces in de fenomenologische modellen is fundamenteel om zware hadronen te kunnen gebruiken voor het onderzoeken van de eigenschappen van het quark–gluonplasma. Het verhaal aan de experimentele kant is echter ook nog niet klaar. Nauwkeurigere en meer differentiële metingen van de charm-baryon sector zijn gepland met de data die zal worden vastgelegd met de LHC in de komende jaren.

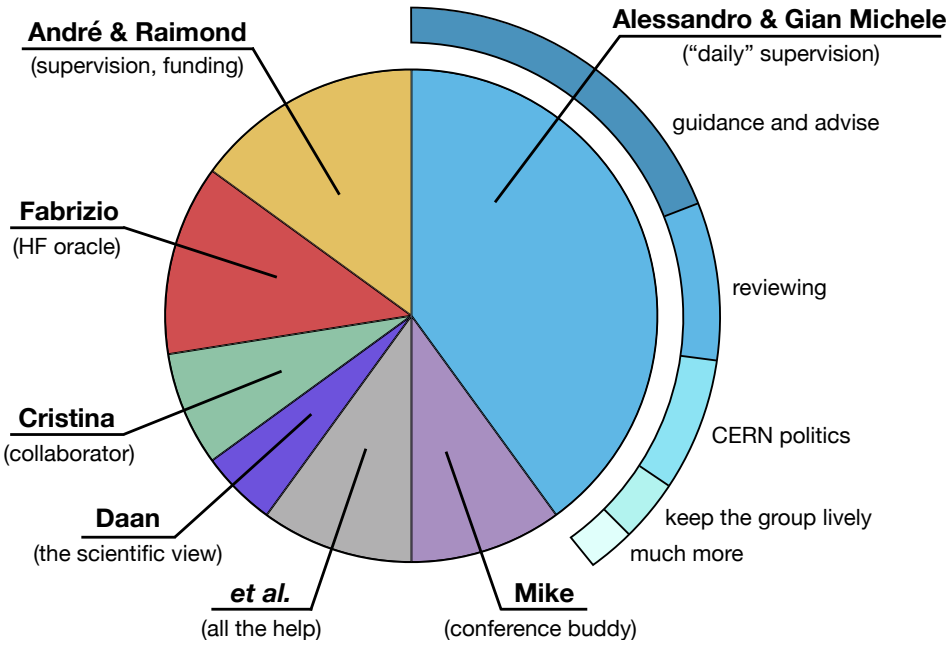
Het veld van de zware-ionen fysica staat niet meer in zijn kinderschoenen, maar is ook zeker nog niet uitgegroeid. Er is nog veel te leren over de eigenschappen van kwantumchromodynamica en het quark–gluonplasma. Op dit moment is de link met het ontstaan van universum, die ik aan het

begin van deze samenvatting legde, ook meer een verkooppraatje dan een echte onderzoeksvraag. Maar wie weet kijken we hier over een aantal jaar alweer anders naar. De potentie is er zeker.

A.4 Acknowledgements

Then suddenly it is over. Four years working as a PhD at Utrecht University finished in this dissertation, with just the acknowledgements left to write. This will probably be the most read chapter in the whole book, so it was a bit unfortunate that it was the most difficult to write. Not because I am not grateful to people — I owe many people a thank you, without them I would not have been able to finish this — it is simply that my non-scientific writing style is a bit underdeveloped after these years. Hence this hybrid approach[5]. In the figures below a scientific breakdown of the acknowledgements is presented. On the next pages, you can read the typical Oscar gratitude speech.





[5]Inspired by a blog post of Matteo Ferla I stumbled upon.

Before anyone else, I would like to thank **André**, who unfortunately no longer has the possibility to read this. André has been my supervisor for my bachelor, master, and PhD project. I thank him for all the support and possibilities he offered me, especially for the one-year stay at CERN. I definitely have to acknowledge André's foresight as well, he knew earlier than me that a career as scientist would suit me. It shows some courage to offer a PhD track to a bachelor student. His assurance helped me to eventually convince myself for a path in science. It was a pleasure to work in your group for all those years and I would have loved to discuss this dissertation with you!

Similarly, my co-promotor **Alessandro** has been part of my journey from the start. You are a person with two sides, but I'm glad I experienced both. I thank you for all the help, support with difficult decisions, and supervision you gave me over the years. It was nice seeing you develop your professor skills, from attending your first ever course during my bachelor thesis to the person supervising non-stop $\mathcal{O}(10)$ students in my last years as PhD. Also your other side — the gentle mockery of others — I enjoyed, it was at least always entertaining when you were around.

Raimond, thanks for being my promotor and stepping in when André was no longer there. You really helped me with all the administrative stuff and with dotting the i's in this dissertation. I also want to thank you for being group leader, and keeping it active by pushing for this merge with the gravitational waves group.

A big thank you to all my colleagues in Utrecht and Nikhef. It's a pity that we could not meet anymore at the coffee machine in these last 1.5 year because of the coronavirus. The two persons with whom I spent the most time in front of this machine (which was because of the "gezelligheid", not because we liked the coffee) are definitely **Mike** and **Lennart**. I am happy that, besides simply sharing the burden[6] of being a student and PhD at SAP, we also became good friends outside office hours. Mike, you were the perfect conference buddy in the last years, lets add some more to our list now we are both postdocs. And somewhere in those years we will manage to do that analysis together. Please stay as excited about science as if it is one of those first times you saw something unexpected with PYTHIA. Similarly, Lennart was the perfect office buddy, the office we shared for at least 20 kilograms of Calvé pindakaas. Thanks for all the times you cooked or baked too much of something delicious that I could finish the day after. I am going to miss your euphoric face when you are talking about your kids.

I very much appreciated the nice and large heavy-flavour group we had in Utrecht. I special shout-out to **Auro**, the cutest scientist of them all with an admirable drive to understand, **Henrique**, the actual group leader who dared to make a live-changing switch, and **Cristina**, for all the help and support when I was young and naive. Marco and Marta also deserve a separate thank you. **Marco** for being one of the very few Dutch people at CERN while I stayed there for a year, and providing me with many possibilities to guide underground tours with you (which I ended up liking a lot). **Marta** I also got to know as a Dutchie at CERN, and I was

[6] Just kidding staff members.

very happy to hear that you joined our group in Utrecht. The lectures I enjoyed TA'ing the most were with you as lecturer. I am also very grateful for the help I received and the fun times I had with the whole Utrecht–Nikhef group these years. Thanks a lot **Annelies**, **Barbara**, **Davide**, **Dimitra**, **Jacopo**, **Monique**, **Naghmeh**, **Panos**, **Paul**, **Shi**, **Syaefudin**, **Thomas**, **Ton**, **Zhanna**, and the others.

That brings me to the second set of "colleagues": people with whom I worked with the most, even though they were employed at universities/laboratories all over the world (at CERN we were used to conference calling long before corona became a thing). A big thank you to the whole ALICE analysis group at CERN for letting me stay for a full year. I learned an incredible amount of new things, had by far the most productive year of my PhD, and — most importantly — had a really great time. In this all, **Gian Michele** had for sure the leading role. Thanks a lot for taking me along in your diehard analysis machine. A big chunk of the work presented in this dissertation was initiated by you. Also **Andrea** and **Benedikt** had a large contribution for which I am grateful, especially with developing our analysis framework. Last but not least, I must thank **Cristina** and **Fabrizio** for bringing our analyses to a good, published end. It was (and is) a pleasure working with you. Fabrizio, you were my "heavy-flavour oracle" throughout these years. I really appreciate all the time you took to answer my questions.

Daan, you are arguably the main reason for the countless of hours I missed during working days because of runs, cycling trips, or (attempts) to swim. Looking at our Strava profiles, it doesn't seem much work could have been performed during the day. Nevertheless we both managed a career in science and sports[7]. We discussed a lot about science as well during our many hours of training. You, unintentionally, taught me how to behave as a duathlete and as a scientist, and for both I thank you. Also, your PhD thesis was the best-looking I have seen so far, and a great source of inspiration.

Jur, I am happy to have a charismatic person as you among my good friends. It was a pity to see you leave our physics studies, but the more exciting that after several years you ended up in the same office building doing a chemistry PhD. I would like to take this opportunity to eternalise our last (of many) games of table football in the office. Winning the first hundred games doesn't mean a thing if you lose the last one... Thanks for agreeing to be my paranymph together with Lennart as well as for your help with the cover. On a similar note I should thank **Jeroen** for pushing me to write a decent "Samenvatting voor leken".

Even though their contribution to the work in this dissertation was minimal, I still owe a thank you to two groups of friends: one in Nijmegen (**Daniel**, **Sjoerd**, **Valerie**, **Wesley**) and the other in Utrecht (**Boris**, **Guus**, **Ivar**, **Jan**, **Jan**, **Jesper**, **Jimmy**, **Marc**, **Thomas**). We didn't see each other as often anymore as I would have liked after I started this PhD trajectory, but when we did it was great as usual. The many beers we shared definitely helped me to take my mind off my work. A special thank you to Boris, for all the walks during the though corona

[7]Okay, for the latter I was mostly known as the "introvert Hellas buddy of the European champion", but well, at least they recognised me at the starting line when you were standing next to me.

times, and to Ivar, who was always up for a nice bike ride. On a similar note, I thank all my triathlon friends for pushing my limits further.

Pap, **mam**, **Emma**, thanks for your unconditional support and love. You were always there for me, giving useful advise, helping with the difficult decisions, and providing a literal helping hand. The foundation of this dissertation was laid by the three of you, far before I knew I was going to pursue a career in academia. Pap, mam, you gave me all the opportunities I could wish for and I know you will encourage me whatever path I take in the future. Emma, my not-so-little little sister, I am really proud on how you approach life. I hope we will see each other more often in the future than these last four years. And when we do, I will keep trying to convince you about the benefits to society of my field of work, even though it will never come close to yours.

Drawing this chapter to a close, there is one person left to thank. The most important of you all, **Amanda**. You have been watching from the sidelines to this physics journey for the last 8.5 years. Years I would not have wanted to spend without you. This dissertation is for sure also your credit. Without your support in the difficult and stressful moments, I don't think I would have been able to finish it at all. It was also in the little things, offering a sympathetic ear when I was stuck or providing me with some snacks when work ended up taking the full evening. There are many other things I could write here to thank you for, but one obviously sticks out. I am extremely grateful that you decided to join me to Geneva, even though it was a big sacrifice on your part. We had a great year, with many adventures. I hope we will keep making new ones in the future.

A.5 List of publications

In the ALICE Collaboration, all scientists contributing to the experiment are on the author list in alphabetic order. I am, therefore, author of *164 papers* in refereed international and national journals. Below I listed only the publications where I was a main contributor.

2021 **Publication:** ALICE Collaboration, "Constraining hadronization mechanisms with Λ_c^+/D^0 production ratios in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV", submitted to Phys. Rev. Lett., `arXiv:2112.08156 [nucl-ex]`.

Publication: ALICE Collaboration, "Observation of a multiplicity dependence in the p_T-differential charm baryon-to-meson ratios in proton–proton collisions at $\sqrt{s} =$ 13 TeV", submitted to Phys. Lett. B, `arXiv:2111.11948 [nucl-ex]`.

Proceeding: L. Vermunt (for the ALICE Collaboration), "Recent ALICE results on charm production and hadronisation", *PoS* Moriond QCD (2021), `arXiv:2105.07154 [nucl-ex]`.

Public note: ALICE Collaboration, "Upgrade of the ALICE Inner Tracking System during LS4: study of physics performance", (in preparation).

2020 **Proceeding:** L. Vermunt (for the ALICE Collaboration), "Open heavy-flavour production from small to large collision systems with ALICE at the LHC", *PoS* ICHEP2020 (2021), 571, `arXiv:2011.06291 [nucl-ex]`.

2019 **Proceeding:** L. Vermunt (for the ALICE Collaboration), "Measurement of Λ_c^+ baryons and D_s^+ mesons in Pb–Pb collisions with ALICE", *PoS* EPS-HEP2019 (2020), 297, `arXiv:1910.11738 [nucl-ex]`.

Publication: ALICE Collaboration, "Measurement of D^0, D^+, D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE", *Eur. Phys. J. C* **79** (2019) 388, `arXiv:1901.07979 [nucl-ex]`.

2018 **Public note:** ALICE Collaboration, "Preliminary Physics Summary: Measurement of D^0, D^+, D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE", ALICE-PUBLIC-2018-006.

2017 **Preprint:** L. Vermunt *et al.*, "Influence of final-state radiation on heavy-flavour observables in pp collisions", `arXiv:1710.09639 [nucl-th]`.