Department of Physics and Astronomy
University of Heidelberg

Master Thesis in Physics
submitted by

Carolina Reetz

born in Frankenthal/Pfalz (Germany)

August 2022

Measurement of Ξ_c^+ in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ALICE detector

This Master Thesis has been carried out by Carolina Reetz at the
Physikalisches Institut of the University of Heidelberg
under the supervision of
Prof. Dr. Silvia Masciocchi

Abstract

High-energy hadronic collisions are used to study the fundamental nature of strongly interacting matter. Nucleus-nucleus and proton–proton (pp) collisions at high-energy particle colliders allow investigating strongly interacting matter under extreme conditions, like high temperatures, and characterising the fundamental theory underlying nuclear physics, Quantum Chromodynamics (QCD). Measurements of heavy-flavour hadron production in pp collisions provide an important test for perturbative QCD and serve as a reference for production measurements in heavy-ion collisions, where heavy-flavour hadrons act as a sensitive probe for the hot nuclear matter state predicted at high temperatures, the quark-gluon plasma (QGP). The production cross section ratio of charmed baryons and mesons is sensitive to the fragmentation functions, which describe the probability of a charm quark to hadronise into a specific charmed hadron species, and which have been assumed to be universal across different collision systems. Measurements of charmed baryon and meson production are hence important to study charm hadronisation mechanisms in different collision systems.

This thesis presents the latest p_{T}-differential production cross section measurement of the charm-strange Ξ_c^+ baryon in the transverse momentum range $3 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$ at midrapidity in pp collisions at the centre-of-mass energy $\sqrt{s} = 13\,\mathrm{TeV}$ recorded by the ALICE detector at the LHC. The short-lived particle is reconstructed via its weak decay to a Ξ^- baryon and two pions, employing the KFParticle software package for the full reconstruction of particle decays, developed for the CBM experiment. Reconstructed particle candidates are selected based on their decay topology with a multivariate analysis approach using the machine learning tool XGBoost.

The Ξ_c^+/D^0 production cross section ratio is measured at midrapity and compared with several model predictions, which consider different charm quark hadronisation mechanisms. It is found, that the ratio is significantly enhanced compared to results from $\mathrm{e}^+\mathrm{e}^-$ and $\mathrm{e}^-\mathrm{p}$ collisions, suggesting that the fragmentation of charm into hadrons is modified for baryons and mesons in different systems. The measurement provides important constraints to model predictions, in particular being sensitive to charm-strange baryon production.

Zusammenfassung

Hochenergetische Hadronenkollisionen dienen der Untersuchung der grundlegenden Natur von stark wechselwirkender Materie. Kern-Kern und Proton-Proton (pp) Kollisionen an Hochenergie-Teilchenbeschleunigern ermöglichen die Untersuchung stark wechselwirkender Materie unter extremen Bedingungen, wie hohe Temperaturen, und die Charakterisierung der grundlegenden Theorie der Kernphysik, der Quantenchromodynamik (QCD). Messungen der Produktion von Heavy-Flavour Baryonen in pp Kollisionen sind ein wichtiger Test für perturbative Quantenchromodynamik (QCD) und dienen als Referenz für Messungen in Schwerionenkollisionen, in denen Heavy-Flavour Hadronen als empfindliche Sonde für den bei hohen Temperaturen vorhergesagten Zustand der Kernmaterie, das Quark-Gluon-Plasma (QGP), dienen. Das Verhältnis der Produktions-Wirkungsquerschnitte von Baryonen und Mesonen mit Charm-Quarks bietet indirekt Zugang zu den Fragmentierungsfunktionen, die die Wahrscheinlichkeit beschreiben, dass ein Charm-Quark in eine bestimmte Hadronenspezies mit Charm hadronisiert, und von denen angenommen wurde, dass sie unabhängig vom Kollisionssystem sind. Messungen der Produktion von Baryonen und Mesonen mit Charm sind daher wichtig, um die Hadronisierungsmechanismen von Charm-Quarks in verschiedenen Kollisionssystemen zu untersuchen.

In dieser Arbeit wird die Messung des p_{T}-differentiellen Produktions-Wirkungsquerschnitts des Ξ_c^+-Baryons im Transversalimpulsbereich $3 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$ bei mittlerer Rapidität in pp-Kollisionen bei der Schwerpunktsenergie $\sqrt{s} = 13\,\mathrm{TeV}$ mit dem ALICE Detektor am LHC präsentiert. Das kurzlebige Teilchen wird über seinen schwachen Zerfall in ein Ξ^--Baryon und zwei Pionen rekonstruiert, wobei das für das CBM-Experiment entwickelte Softwarepaket KFParticle verwendet wird. Die rekonstruierten Teilchenkandidaten werden anhand ihrer Zerfallstopologie mit einer multivariaten Analyse selektiert, wozu das Machine Learning Paket XGBoost angewandt wird.

Das Wirkungsquerschnitts-Verhältnis Ξ_c^+/D^0 wird gemessen und mit verschiedenen theoretischen Vorhersagen verglichen, die jeweils verschiedene Hadronisierungsmechanismen für Charm-Quarks berücksichtigen. Es zeigt sich, dass das Verhältnis im Vergleich zu den Ergebnissen aus $\mathrm{e}^+\mathrm{e}^-$- und $\mathrm{e}^-\mathrm{p}$-Kollisionen signifikant höher ist, was darauf hindeutet, dass sich die Fragmentierung von Charm in Hadronen für Baryonen und Mesonen in verschiedenen Systemen unterscheidet. Die Messung liefert wichtige Anhaltspunkte für theoretische Modelle und bietet insbesondere die Möglichkeit, die Produktion von Baryonen mit Charm- und Strange-Quarks zu untersuchen.

Contents

1. Introduction

1.1. Quantum Chromodynamics

The standard model of particle physics [1] characterises all known elementary particles and the electromagnetic, weak, and strong forces, which are the interactions between them. The fundamental point-like spin-$\frac{1}{2}$ particles, the fermions, include quarks and leptons and appear in three different generations with increasing masses but the same fundamental interactions. The forces between the elementary particles are each described by a quantum field theory (QFT) where interactions between particles are mediated via the exchange of spin-1 gauge bosons.

This short introduction to the underlying theory relevant for this work is mainly summarised from Ref. [2].

Quantum Chromodynamics (QCD) is the relativistic quantum field theory describing the strong interaction and it is associated with an invariance under SU(3) local phase transformations. Analogously to Quantum Electrodynamics (QED), which describes the electromagnetic interaction and has one conserved electric charge, the conserved charge associated with QCD is *colour*, taking three states labelled as red, green, and blue. The colour charge is carried by the quarks with flavours up, down, strange, charm, beauty, and top. The symmetry under local gauge transformations requires eight gauge fields, which correspond to eight massless gauge bosons, the gluons, mediating the strong interaction and connecting quark states of different colours. Due to colour conservation at the QCD interaction vertex, gluons therefore must carry colour and anticolour charge themselves. This leads to the fact that gluons can self-interact, which is the reason for the very different behaviour of QCD compared to QED.

The effective strength of the strong interaction between colour charges results from the sum of all possible processes. These include higher-order corrections to the bare QCD interaction vertex. All these corrections are absorbed in the definition of an effective strong coupling strength, $\alpha_{\mathrm{s}}(Q^2)$, depending on the momentum transfer, Q^2. For example, a mediating gluon can emit a quark-antiquark pair which annihilates back to a gluon, which is described as a fermionic loop in the Feynman diagram of the interaction. As a consequence, the initial charges are surrounded by a cloud of virtual $\mathrm{q\bar{q}}$ pairs. In QCD, additional bosonic loop diagrams occur owing to the gluon self-interaction. Whereas the virtual colour neutral $\mathrm{q\bar{q}}$ clouds have the effect of screening the interacting colour charges, the virtual gluon cloud carries colour and leads to an *anti-screening* effect, which can be interpreted as the opposite of screening. The dominance of the anti-screening mechanism in QCD leads to the fact that the effective coupling becomes small for large momentum transfers Q^2, and diverges at small Q^2. This gives rise to the concepts of *colour confinement* and *asymptotic freedom*. Colour confinement implies that coloured objects cannot propagate as free particles but are only observed confined in colourless bound states, called hadrons, at small values of Q^2 (increasing distance). At large values of Q^2, the coupling becomes small and the quarks and gluons can travel distances exceeding the size of a hadron, and hence can be treated as quasi-free particles rather than being strongly bound within the hadron. This concept is known as asymptotic freedom. The evolution of α_s with the energy scale, Q, is experimentally well established, Ref. [3] gives a summary of the experimental values as function of Q. In high-energy regimes ($Q^2 \gtrsim 1\,\mathrm{GeV}^2/c^2$), which are accessible with high-energy collider experiments, α_s is sufficiently small for perturbative QCD (pQCD) to be applicable to calculate interactions between coloured objects, using higher orders of α_s. For low-energy interactions, α_s is of $\mathcal{O}(1)$ and therefore becomes too large for a perturbative approach. In this non-perturbative low-energy regime, which for example applies to the discussion of the later stages of the hadronisation process, the computational technique of lattice QCD (lQCD) [4] is used, where calculations are performed on a discrete space-time grid.

1.2. Nuclear matter under extreme conditions

As a result of the variation of α_s with the energy scale, QCD predicts distinct phases of nuclear matter with different dominant degrees of freedom, represented in a QCD phase diagram. A sketch of the current understanding of this phase diagram is shown in Figure 1.1. In the case of ordinary nuclear matter at finite temperature $T \approx 0$ and $\mu_B \approx 1\,\mathrm{GeV}$, where μ_B denotes the baryo-chemical potential (excess of matter over antimatter), confinement states that quarks and gluons are bound into colour-neutral hadrons. For extremely high temperatures and/or densities, quarks and gluons are expected to move freely over distances larger than the size of a nucleon and form a deconfined state of matter, which is called *quark-gluon plasma (QGP)* [6]. In the limit of zero baryo-chemical potential, a smooth crossover transition to a QGP state is predicted by lQCD [7] at the critical temperature $T_\mathrm{c} = (156.5 \pm 1.5)\,\mathrm{MeV}$ [8]. Microseconds after its creation in the Big Bang, the universe is believed to have been a hot QGP before

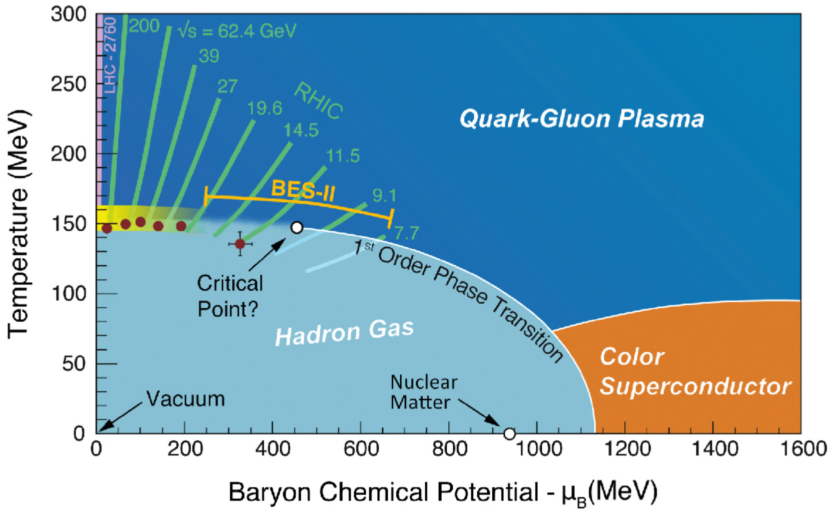


Figure 1.1.: Sketch of the QCD phase diagram as function of temperature, T, and baryo-chemical potential, μ_B. Figure taken from [5].

cooling down below the critical temperature T_c [9]. At finite temperatures ($T \approx 0$) and increasing baryo-chemical potential, on the other hand, a first-order phase transition of nuclear matter to a deconfined state with potential colour-superconducting properties is predicted [10], which is believed to exist in the core of neutron stars.

Since its prediction [11, 12], various experimental evidence for a QGP state of strongly interacting matter has been collected [13]. The early universe is not directly experimentally accessible, but the QGP can be investigated in high-energy nucleus-nucleus collisions. At collider experiments, where high temperatures and energy densities are reached, the QCD phase diagram is probed in the region of $\mu_B \approx 0$. At lower energies, a high amount of the nucleons are stopped in the collision region and the baryon density is high. This allows to study the phase diagram in the region of non-vanishing baryo-chemical potential. At the Relativistic Heavy Ion Collider (RHIC) a beam-energy scan, where the energy of the colliding beams is lowered systematically, together with measurements of the STAR (Solenoidal Tracker at RHIC) detector in fixed target mode, are used to probe the QCD phase diagram at different ”starting points” to search for evidence of a critical point and a first-order phase transition at non-zero baryo-chemical potential [14].

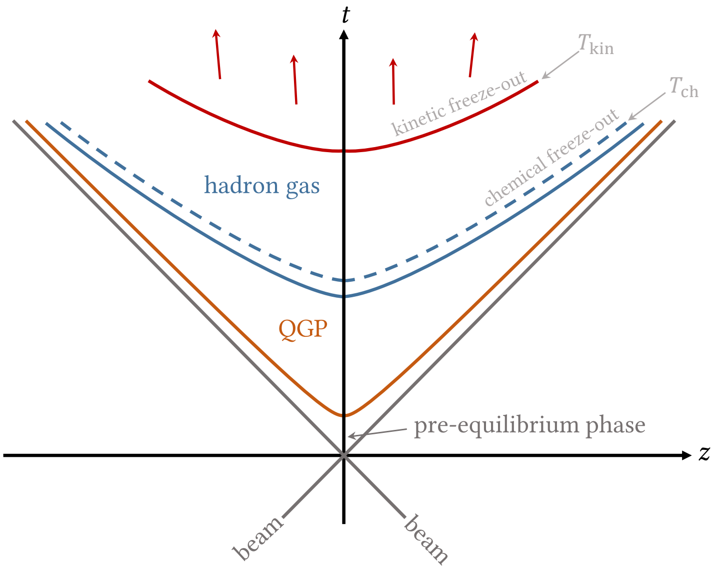


Figure 1.2.: Sketch of the evolution of a heavy-ion collision as space-time diagram. The evolving system after the initial collision time of the two beams is depicted in the light-cone at times $t > 0$.

1.3. High-energy nuclear collisions

When two ultra-relativistic heavy ions collide, the system undergoes a complex evolution [5], depicted in Figure 1.2 as space-time diagram. For a head-on collision of two highly Lorentz-contracted heavy ions, the energy density is highest at the moment that the nuclei collide ($\tau = 0$) [5]. Subsequently, the participating partons can undergo hard processes with large momentum transfer. At this *pre-equilibrium* stage, the system is far from equilibrium and the interaction rate between the partons is high, leading to a rapid approach to local thermal equilibrium at an expected thermalisation time of $\tau_0 \lesssim 1\,\text{fm}/c$ [15]. In case the energy density at τ_0 exceeds the energy density of a hadron, the produced quarks and gluons cannot be described as confined anymore, but a local thermalised equilibrium *QGP phase* forms, which cools down while expanding. It was shown that the macroscopic evolution of the medium can be described by relativistic fluid dynamics [16]. The QGP phase lasts up to the point where the temperature of the medium reaches the critical value T_c and undergoes the crossover transition discussed in section 1.2. During this *hadronisation stage*, the quarks and gluons get confined into hadrons at the critical temperature, T_c. The produced *hadron gas* expands further while the constituents undergo inelastic collisions. When the *chemical freeze-out* temperature, T_{ch}, is reached, inelastic scattering stops and the hadron abundancies (yields) are fixed.

Once the temperature drops below the *kinetic freeze-out* limit, T_{kin}, also elastic interactions between the particles cease and the hadrons stream away freely to the detector.
The described evolution cannot be directly observed in measurements, but indirect probes have to be used to extract information about the different stages and properties of the QGP. Furthermore, it is crucial to gain an extensive understanding of the processes in smaller systems, like proton–proton (pp) collisions, which are studied in this thesis, and which serve as a reference for measurements in heavy-ion collisions. For a comprehensive investigation, both systems have to be understood in detail.

1.4. Charm production in high-energy collisions

Due to their large mass ($m_{\text{c}} \simeq 1.3\,\text{GeV}/c^2$ [3]), charm quarks can only be produced in the initial hard scattering of a high-energy collision with large momentum transfer $Q^2 > 4m_{\text{c}}^2$. In these regions, the coupling strength, α_{s}, is small enough to apply pQCD calculations for the computation of charm production.
When two high-energy hadrons collide, the incident partons scatter and interact via colour fields (*initial hard partonic scattering*). The interacting partons can either be the valence quarks, the gluons, or the seaquarks (-antiquarks) of the incoming hadrons. During one pp collision, multiple partonic interactions (MPIs) can occur. Within the hard scattering processes, charm quark-antiquark pairs ($\text{c}\bar{\text{c}}$) are for example produced in pQCD processes like gluon fusion gg $\rightarrow$ $\text{c}\bar{\text{c}}$ or annihilation of light quarks and antiquarks $\text{q}\bar{\text{q}}\rightarrow$ $\text{c}\bar{\text{c}}$ [17]. The Leading Order diagrams of the two processes are depicted in Figure 1.3. Heavy quarks or antiquarks are also produced in *parton showers*, as depicted in Figure 1.3. The participating initial or final state partons from the hard scattering processes can emit initial-state (ISR) or final-state radiation (FSR) in the form of gluons or photons, which possibly split into $\text{c}\bar{\text{c}}$ pairs [17]. Furthermore, heavy quarks (antiquarks) can also come from the parton sea of the collided protons, when the companion antiquark (quark) underwent a hard partonic interaction.
The lower limit of Q^2 for charm production to happen corresponds to an upper bound for the production time of about $\tau \sim 1/Q \lesssim 0.1\,\text{fm}/c$, which is smaller than the expected formation time of a QGP, τ_0, implying that charm production is mostly unaffected by a possible medium.
The heavy quarks produced in the initial hard processes get confined into colour-neutral hadrons in the non-perturbative process of *hadronisation*.

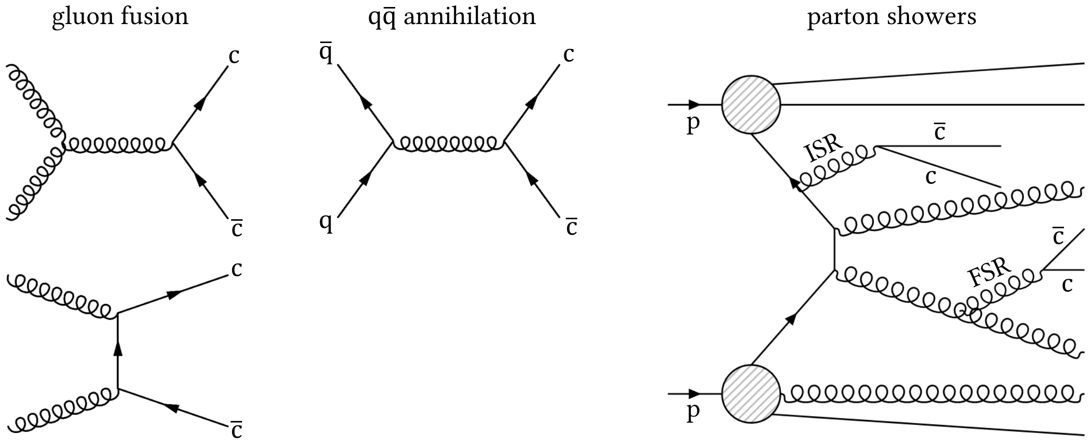


Figure 1.3.: Left: Leading Order QCD diagrams of gluon fusion and quark-antiquark (qq̄) annihilation. Right: Parton showers from initial-state (ISR) and final-state radiation (FSR) emitted by partons in the hard scattering of a pp collision.

1.5. Open heavy flavour hadron production in proton–proton collisions

1.5.1. Factorisation approach

In pp collisions, the production cross section of open heavy flavour hadrons, which are particles containing at least a heavy quark (in this case charm) and other lighter quarks, is computed exploiting the QCD *factorisation theorem* [18] as follows:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}}^{pp\rightarrow H_cX} = \sum_{i,j=q,\bar{q}g} f_i(x_1,\mu_f^2)f_j(x_2,\mu_f^2)\frac{\mathrm{d}\sigma^{ij\rightarrow c\bar{c}}}{\mathrm{d}p_{\mathrm{T}}}D_{c\rightarrow H_c}(z_c = p_{H_c}/p_c,\mu_f^2). \tag{1.1}$$

In the formula, c refers to the charm quark, H_c to the open heavy-flavour hadrons, and p_{T} to their transverse momentum. Within this approach, their production cross sections are described as a convolution of the *parton distribution functions (PDFs)* of the colliding protons, $f_i(x_1,\mu_f^2)f_j(x_2,\mu_f^2)$, the parton hard scattering cross section of cc̄ pair production, $\mathrm{d}\sigma^{ij\rightarrow c\bar{c}}/\mathrm{d}p_{\mathrm{T}}$, and the *fragmentation functions (FFs)*, $D_{c\rightarrow H_c}(z_c,\mu_f^2)$. The PDFs give the probability to find a parton of flavour i carrying a fraction x_i of the total momentum in the incoming proton (p), where μ_f^2 is the factorisation scale. The probability of a charm quark c to hadronise into the hadron H_c (*fragmentation fraction*) carrying the momentum fraction z_c is included in the FF. While the hard scattering charm production cross section can be computed as power expansion in terms of α_s with pQCD, as discussed earlier, for both, the PDFs and FFs, the underlying processes are non-perturbative and hence these functions have to be tuned on measurements. While the PDFs are parameterised from deep inelastic scattering experiments, like $\mathrm{e^-p \rightarrow e^-X}$, the FFs are typically taken from measurements in $\mathrm{e^+e^-}$ collisions and are assumed to be universal among different collision systems.

1.5.2. Ratios of charm hadron production cross sections

Models based on pQCD calculations exploiting the factorisation theorem with FFs tuned on $\mathrm{e^+e^-}$ data generally describe the measurements of D- and B-meson (containing a charm and beauty quark respectively) production cross sections down to low transverse momenta in pp collisions at several centre-of-mass energies at the Large Hadron Collider (LHC) [19–21]. However, it was shown that these model calculations are not able to capture measurements of Λ_c^+-baryon production in pp collisions at midrapidity at the centre-of-mass energies $\sqrt{s} = 5.02\,\mathrm{TeV}$ and 7 TeV reported by the ALICE Collaboration [22–24]. Similar observations were made with the measurement of Ξ_c^0-baryon production (a particle containing a strange (s) quark in addition to the heavy charm quark) in pp collisions at the centre-of-mass energies $\sqrt{s} = 7\,\mathrm{TeV}$ [25] and $5.02\,\mathrm{TeV}$ [26].

Since the PDFs and the hard scattering are independent of the final measured hadron species, only the fragmentation function remains when computing ratios of production cross sections of different heavy flavour hadron species. Therefore, measurements of hadron-to-hadron production cross section ratios are sensitive to fragmentation fractions and heavy flavour hadronisation mechanisms.

The reported charm-baryon measurements from the ALICE Collaboration show higher Λ_c^+/D^0 and Ξ_c^0/D^0 production cross section ratios compared to previous measurements of Λ_c^+-baryon production in $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collisions (see for example Refs. [27, 28]), showing an enhanced baryon over meson production in pp collisions. Baryon enhancement is also observed in heavy-ion collisions, which is shown for example by the measurement of Λ_c^+/D^0 in lead–lead (Pb–Pb) collisions at the centre-of-mass energy per nucleon-nucleon pair $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ measured by the ALICE Collaboration [29]. These results suggest that the fragmentation fractions of charm quarks are non-universal across different collision systems, in contrary to expectations, and demand for more differential studies of charm hadronisation across various collision systems.

In order to understand the observed enhancement of charm-baryon production in hadronic collisions compared to $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ measurements, different theoretical approaches are proposed.

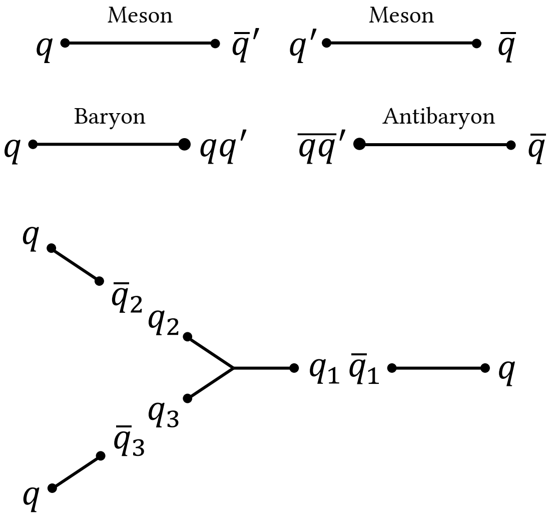


Figure 1.4.: Hadronisation in the picture of string fragmentation. Top: Meson production via string breaking quark-antiquark (q′q̄′) creation. Middle: Baryon production via string breaking diquark pair production. Bottom: Hadronisation by q′q̄′ production in a junction string topology.

1.5.3. Hadronisation models

PYTHIA event generator with Lund string fragmentation

State-of-the-art Monte Carlo (MC) event generators like PYTHIA [30] are used to model charm hadron yields implementing different charm hadronisation processes, where the fragmentation functions are tuned on e^+e^- and e^-p measurements. PYTHIA is modelling the hadronisation process with the Lund *string fragmentation* approach [31]. In the string fragmentation picture, colour fields between partons resemble *strings*, which connect a colour object at one of their ends with an anticolour charge at the other end. Since the QCD potential between a quark and an antiquark grows linearly with the distance between them, a high-tension string, i.e. an increasing potential, forms between the two connected partons as they move away from each other. At some point, it is energetically more favourable to break the string via the production of a quark-antiquark (qq̄) pair from the vacuum, rather than extending it further. This leads to the production of mesons (containing a quark and an antiquark), or further fragmenting of the newly created individual strings. The top sketch in Figure 1.4 shows a simplified picture of this meson production mechanism via qq̄ pair creation and string breaking. The equivalent production process for baryons (containing three valence quarks) is depicted in the middle figure and works by string breaking via the creation of a diquark and an anti-diquark.

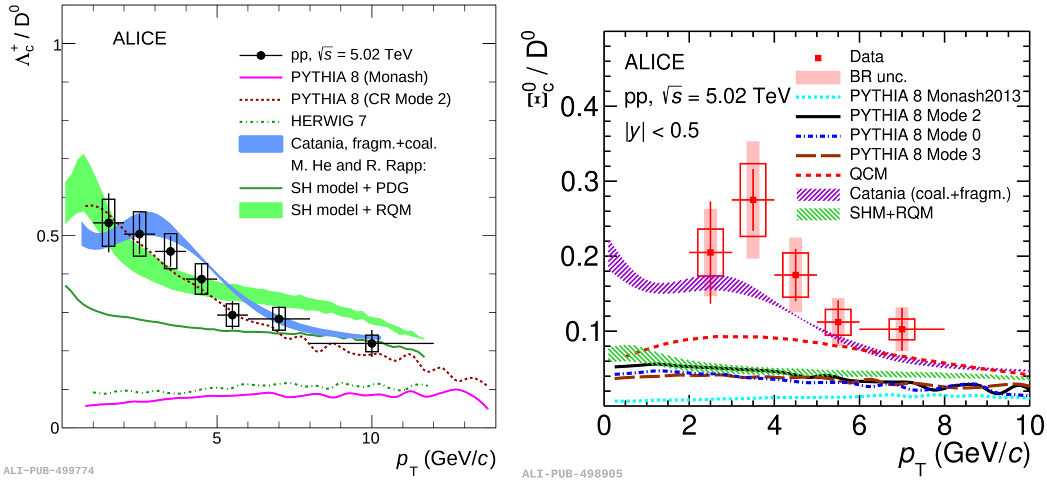


Figure 1.5.: Baryon-to-meson ratios measured in pp collisions at $\sqrt{s} = 5.02\,\mathrm{TeV}$ at midrapidity by the ALICE Collaboration as function of the transverse momentum. Left: The Λ_c^+/D^0 ratio [23, 24] compared to theoretical predictions. Right: The Ξ_c^0/D^0 ratio [26] compared to different model predictions.

An important question to be addressed by the model is which partons are connected via strings in the first place. In the leading colour (LC) approximation, a parton is colour-connected to one other parton belonging to the same MPI system. So-called *colour reconnection (CR)* models introduce new colour connection topologies, allowing partons to connect beyond LC across various hard scattering processes or with beam remnants, minimising the string lengths between them [30, 32]. Beyond that, even more recent CR models [33] allow the partons to colour-connect to two other partons and form *junctions*, in which the string pieces from each of the three quarks (or antiquarks) meet in a Y-shaped topology. The hadronisation process of such a topology is depicted in the bottom sketch of Figure 1.4. The three strings are breaking via the production of $\mathrm{q\bar{q}}$ and diquark pairs forming mesons and baryons (antibaryons), as discussed above. The produced quarks which are nearest to the junction, one from each of the string pieces, form a baryon $\mathrm{q_1q_2q_3}$. This additional baryon production mechanism leads to a baryon enhancement in models with CR beyond LC.

The PYTHIA event generator comes in different tunes, all implementing different sets of parameters tuned on different measurements. Most typically, PYTHIA 8 is used with the Monash tune [34].

Figure 1.5 shows the baryon-to-meson ratios Λ_c^+/D^0 [23, 24] and Ξ_c^0/D^0 [26] measured in pp collisions at the centre-of-mass energy $\sqrt{s} = 5.02\,\mathrm{TeV}$ at midrapidity by ALICE, compared to different model predictions. Predictions from PYTHIA 8 with the Monash tune, with parameters tuned on e^+e^- and e^-p, are found to significantly underestimate the ratios. Models with CR beyond LC, like PYTHIA 8 with Mode 2 or Mode 3 are reproducing the Λ_c^+/D^0 ratio much better but are not able to capture the Ξ_c^+/D^0 ratio.

Quark (re-)combination mechanism

The observed baryon enhancement in high-energy pp collisions compared to e^+e^- and e^-p measurements, together with other measurements in pp collisions showing a similar phenomenology to heavy-ion collisions, point to a possible formation of a hot QCD medium also in pp collisions, even though the energy density is expected to be insufficiently high for a QGP to be created.

While fragmentation happens in vacuum, an alternative mechanism for hadronisation of partons in the presence of a parton-rich environment, called *quark (re-)combination mechanism (QCM)* or *coalescence*, is considered, which was initially proposed as hadronisation mechanism in heavy-ion collisions. In this model, heavy quarks in a deconfined medium coalesce with other light quarks. They pick up either a comoving antiquark or two comoving quarks from the medium, which are close in phase space, to form a meson or baryon.

Due to the possible formation of a deconfined medium in pp collisions, the measured Ξ_c^0/D^0 is compared to the QCM prediction to test the contribution of coalescence to charm hadronisation in pp collisions. The right panel of Figure 1.5 shows that the QCM model [35] does not describe the ratio. This sets constraints on the model parameters and possibly suggests that a pure coalescence approach is not sufficient.

Catania coalescence model

The Catania model [36] implements hadronisation of charm quarks via coalescence together with vacuum fragmentation and assumes the presence of a thermalised medium of light quarks (u, d, s). The momentum spectrum of hadrons, which are formed by coalescence of quarks, are computed based on the Wigner function describing the spatial and momentum distribution of quarks in a hadron, where the width of the Wigner function is related to the root mean square charge radius of the hadron taken from the quark model [36]. The Catania model adopts a Gaussian shape for the Wigner distribution function in space and momentum, which is normalised to guarantee that the total probability for coalescence in the limit $p \to 0$ is 1. Furthermore, a statistical factor is taken into account in the computation of the hadron spectra, giving the probability that two (three) random quarks have the right quantum numbers to match the quantum number of the considered hadron.

Within the Catania model, the probability of a charm quark to hadronise via coalescence or fragmentation depends on its transverse momentum. The coalescence probability is high for low quark momenta and quickly decreases with increasing transverse momentum, meaning that all charm quarks hadronise via coalescence at low transverse momenta whereas in the high momentum region, fragmentation is the dominant contribution.

The Catania model is compared to the measured Λ_c^+/D^0 and Ξ_c^+/D^0 ratios in Figure 1.5. In the case of the Λ_c^+/D^0 ratio, the model provides a good description of the data, and it is the model that is closest to the measured Ξ_c^0/D^0 ratio over the full transverse momentum interval. This indicates the possibility that charm quark coalescence takes place in pp collisions.

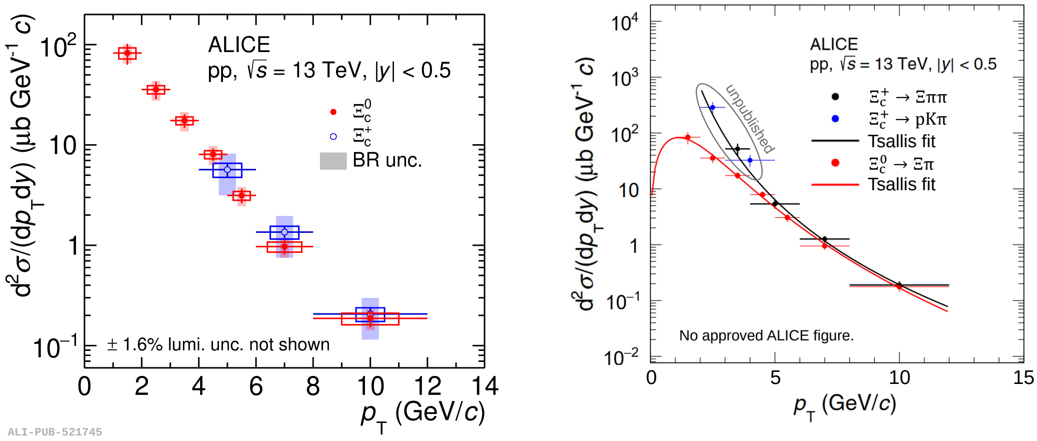
Statistical hadronisation model

Within the statistical hadronisation model (SHM) [37] particle yields in heavy-ion collisions are computed based on statistical weights governed by the mass of the possible hadron states at the hadronisation temperature.

The model is compared to measurements in pp collisions and is partly found to describe the data successfully. The Λ_c^+/D^0 ratio in the left panel of Figure 1.5 is compared to a SHM [38] using the baryon states listed by the Particle Data Group [3], which is found to underpredict the data, especially at low transverse momenta. If additional excited charm baryon states, which are not yet observed but predicted by the relativistic quark model (RQM) [39], are taken into account, the model is able to describe the Λ_c^+/D^0 ratio. However, even with an extended list of charm baryon states, the model underestimates the Ξ_c^+/D^0 ratio by the same amount as PYTHIA 8 with CR tunes, which might indicate a yet incomplete list of charm baryon resonances in the RQM.

1.6. Analysis motivation

The discussed results demand for more differential measurements in the heavy flavour baryon sector, both in pp collisions and in larger systems. The measurement of open charm hadron production in pp collisions is a powerful tool for the test of pQCD calculations describing the production of charm quarks in the hard scattering. In heavy-ion collisions, charm quarks are an excellent probe for the created QGP due to their early



(a)

(b)

Figure 1.6.: (a) Cross section of prompt Ξ_c^0 (full red markers) and Ξ_c^+ (open blue markers) baryons as function of p_T in pp collisions at $\sqrt{s} = 13$ TeV. The statistical and systematic uncertainties are represented by the error bars and empty boxes, respectively. The systematic uncertainties due to the branching ratio (BR) are shown as shaded boxes. Figure taken from [41]. (b) Published cross section measurement of prompt Ξ_c^0 baryons [41] (red markers) as function of p_T in pp collisions at $\sqrt{s} = 13$ TeV compared to partly unpublished measurements of the Ξ_c^+ baryon cross section in two different decay channels (black and blue markers). The measurements are fitted with a Tsallis function [42].

production time and flavour conservation by the strong interaction. Hence, the measurement of charm hadron production in these systems allows studying and characterising the properties of the QGP.

This work focuses on the measurement of the production of the charm-strange Ξ_c^+ baryon in pp collisions, which is expected to be even more enhanced with respect to Λ_c^+ production due to strangeness enhancement observed in pp collisions [40]. The result provides important constraints for several charm quark hadronisation models in pp collisions, in particular being sensitive to charm-strange baryon production. Moreover, it will serve as a reference and preparation for the planned measurement in Pb–Pb collisions to study the influence of a QGP, which is yet not feasible with the data currently available.

The Ξ_c^+ baryon (usc) is a short-lived particle ($c\tau(\Xi_c^+) = 136.6$ μm [3]) which is not directly detectable in the detector due to its short life time. It is fully reconstructed via its hadronic decay to a Ξ baryon and two pions (π) in pp collisions at $\sqrt{s} = 13$ TeV. The production cross section of Ξ_c^+ baryons with a transverse momentum (p_T) between 4 and 12 GeV/c was already measured in pp collisions at $\sqrt{s} = 13$ TeV in ALICE via the same decay [41]. The result is shown in Figure 1.6a. It is compared to the measurement of its isospin partner, the Ξ_c^0 baryon, which was measured in ALICE via its hadronic decay to a Ξ baryon and a π in the same system at the same collision energy [41]. The results between 4 and $12\,\mathrm{GeV}/c$ are similar, which is expected and well understood by isospin symmetry.

However, a first attempt to extend the measurement of the Ξ_c^+ cross section to transverse momenta between 3 and $4\,\mathrm{GeV}/c$ yields a result higher by a factor of ~ 3 compared to its isospin partner. A similar result was observed for the measurement of the Ξ_c^+ baryon via its decay to a proton (p), a Kaon (K), and a π in the same collision system at transverse momenta between 2 and $3\,\mathrm{GeV}/c$. Both measurements are shown in Figure 1.6b and are compared to the published Ξ_c^0 cross section. Since isospin symmetry suggests the same production cross section for both baryons, these observations are unexpected. Possibly unobserved higher resonance states that feed differently into Ξ_c^+ and Ξ_c^0 baryons could lead to deviations between the corresponding production cross sections.

In a traditional analysis approach, different features describing the topology of the decay or providing particle identification (PID) information are used to differentiate between signal and background candidates. For the reduction of the large combinatorial background, selection criteria based on these features are applied to the reconstructed candidates. These criteria are usually tuned manually to maximise the statistical significance of the extracted signal. This possibly introduces the risk to enhance the yield via statistical fluctuations. Both measurements of the Ξ_c^+ cross section in Figure 1.6b are performed with standard reconstruction and rectangular selection-based analysis techniques. Since the risks described above are specifically pronounced for low p_T intervals, where the combinatorial background is large and the signal more difficult to extract, the two standard analyses are likely biased in the low-p_T region. This might explain the unexpected high results for the measured cross section.

This work aims to contribute with a cross section measurement of the Ξ_c^+ baryon at $p_\mathrm{T} < 4\,\mathrm{GeV}/c$ to verify this assumption. The analysis developed in this work will exploit new and more refined analysis techniques for the candidate reconstruction and selection. The decay chain reconstruction is performed using the *KFParticle package* [43]. This software is based on the Kalman filter (KF) method and has been developed for the complete reconstruction of short-lived particles by the Condensed Baryonic Matter (CBM) Collaboration. To reduce the risk of a potential bias in the candidate selection, supervised Machine Learning (ML) techniques in the form of Boosted Decision Trees (BDTs) are used to classify the reconstructed candidates as signal or background, employing the XGBoost library [44].

2. The ALICE detector at the LHC

The Large Hadron Collider (LHC) at the European Organisation for Nuclear Research (CERN) nearby Geneva is a superconducting hadron accelerator and the world's most powerful particle collider. It is installed in the $26.7\,\text{km}$ tunnel originally built for the LEP (Large Electron-Positron Collider) machine, along which protons (or lead nuclei) are accelerated in opposite directions in two beam pipes. It is designed for a maximum centre-of-mass energy per nucleon–nucleon pair of $\sqrt{s} = 14\,\text{TeV}$ for pp collisions and $\sqrt{s_{\text{NN}}} = 5.5\,\text{TeV}$ for Pb–Pb collisions [45].
The beams are crossing at so-called interaction points where they are brought to collision. There are four main experiments of the accelerator complex, which are each located at one of the interaction points. Apart from the two high luminosity experiments ATLAS [46] and CMS [47], designed for the search of the Higgs boson and beyond Standard Model physics, as well as the LHCb [48] detector for the study of CP violation, searches, and flavour physics, the LHC has one experiment dedicated to Pb–Pb ion operation, which is called A Large Ion Collider Experiment (ALICE). In the past, also Xenon nuclei were collided, which proves versatility and can be seen as an important test for future plans to collide smaller nuclei like Oxygen.
ALICE is designed as a general-purpose detector for relativistic heavy-ion collisions at the CERN LHC measuring a wide range of physics observables. Its physics programme focuses on the investigation of a small part of the phase diagram of strongly interacting matter and the physics of the QGP, both at extreme values of energy density and temperature [49]. This requires the ability to measure particles down to low transverse momenta with high precision (for $100\,\text{MeV}/c$ pions a relative momentum resolution of $2\,\%$ is achieved [49]). The detector granularity is optimised to cope with the high charged particle multiplicity, which is reached in high-energy heavy-ion collisions (up to $\mathrm{d}N/\mathrm{d}\eta = 4000$ [49]).
Figure 2.1 shows a schematic overview of the ALICE detector layout with its subdetector systems, as it was installed during the second data-taking run at the LHC (Run 2), between 2015 and 2018. The ALICE detector consists of 18 different subdetector systems arranged in a central barrel, which covers the central rapidity region ($|\eta| < 0.9$), and a muon spectrometer at forward rapidity ($-4.0 \leqslant \eta \leqslant -2.5$) [45]. The central barrel detectors are embedded in the large solenoid L3 magnet (reused from the L3 experiment at LEP), which provides a magnetic field with a nominal strength of $B = 0.5\,\text{T}$ parallel

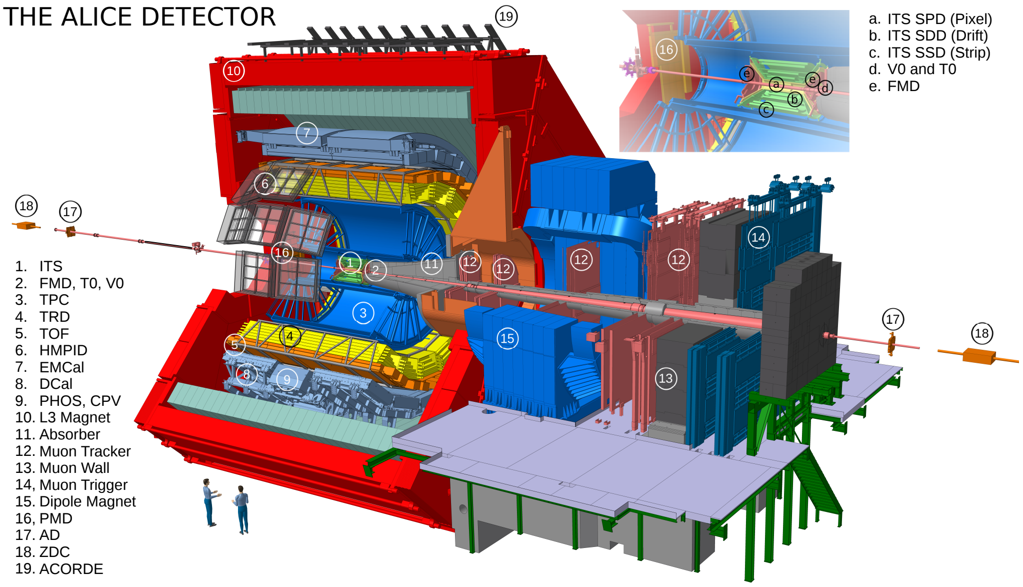


Figure 2.1.: Schematic view of the ALICE detector with its subdetector systems as installed during Run 2 [50].

to the beam axis [49]. The innermost detectors of the central barrel are the two tracking detectors, the Inner Tracking System (ITS) and the Time Projection Chamber (TPC). They are surrounded by the Transition Radiation Detector (TRD) and the Time-Of-Flight (TOF) detector, which provides PID in the intermediate momentum range [45]. These most central subdetectors cover the full azimuth around the beamline, in contrast to the three outermost detector parts, the High Momentum Particle Identification Detector (HMPID) designed to extend the PID capability for high momentum particles, and the two electromagnetic calorimeters, the Photon Spectrometer (PHOS) and the Electromagnetic Calorimeter (EMCal) with the Di-Jet Calorimeter (DCal). Several other detector systems are used for triggering and event characterisation. The VZERO detector consists of two arrays of scintillator counters, VZEROA ($2.8 < \eta < 5.1$ [49]) and VZEROC ($-3.7 < \eta < -1.7$ [49]), located asymmetrically on either side of the interaction point [45]. They are used for the triggering of minimum-bias (MB) events in the central barrel and to reject interactions between the beam particles and the residual gas in the beam pipe. The MB trigger requires a coincidence in the two VZERO detector components.

The most important detectors used for tracking, the ITS and the TPC, and for PID, the TPC and the TOF, which are employed in this analysis, are described in more detail in the following sections.

A Cartesian coordinate system [51] is defined for ALICE, with the point of origin in the centre of the detector denoted as the nominal interaction point (IP), the z-axis parallel

to the mean beam direction, and the x-direction oriented towards the LHC centre. It follows the LHC rules, which are also used by the other main LHC experiments.
The kinematics of a particle of known mass inside the detector is fully described by three independent variables, its momentum in xy-direction, referred to as transverse momentum p_{T}, the azimuth angle ϕ and the polar angle θ. Typically, θ is replaced by the rapidity y, which is Lorentz-invariant along the z-axis.

2.1. Inner Tracking System

The ITS consists of six layers of silicon detectors arranged cylindrically around the beam pipe. It is located closest to the interaction point where a track density of up to $50\,\mathrm{tracks/cm^2}$ in heavy-ion collisions is expected [49]. This influenced the choice of Silicon Pixel Detectors (SPD) for the two innermost layers and Silicon Drift Detectors (SDD) for the following two layers. The two outermost layers use Silicon Strip Detectors (SSD). Its main tasks are charged particle tracking and the reconstruction of the primary vertex (PV), which is the measured collision point. The high spatial resolution ($12\,\mu\mathrm{m}$ in $r\phi$ [49]) of the two SPD layers, together with the SSD layers, allows to measure the impact parameter of secondary tracks (distance of closest approach between the trajectory and the PV) and to reconstruct the secondary vertices (decay points) of short-lived hadrons with high precision. Both are fundamental quantities for the reconstruction of weakly decaying charm hadrons. Above that, the four layers are crucial for the matching of tracks from the TPC (described in the next section) to the ITS during the tracking procedure, improving the angle and momentum resolution of high momentum particles [49]. Together with the two middle SDD layers, the SSD provides a measurement of the specific ionisation energy loss, which is needed for the PID of low momentum particles in the ITS. Due to the sufficient PID capability of the TPC and the TOF detector, the PID information from the ITS is not used in this work.

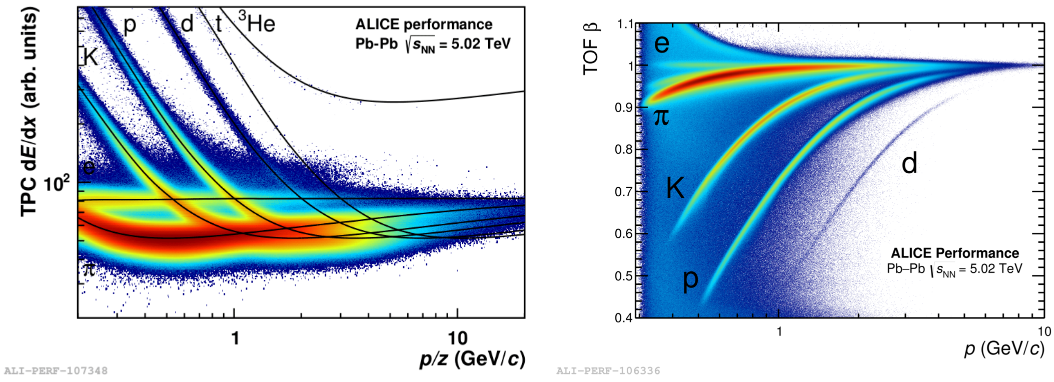
2.2. Time Projection Chamber

The TPC is a very unique instrument for tracking and charged particle identification measurements in the high multiplicity environment given inside the ALICE detector. It is optimised for a large momentum range of about $100\,\mathrm{MeV}/c$ to $100\,\mathrm{GeV}/c$ (assuming a nominal magnetic field of $0.5\,\mathrm{T}$) [49], with good momentum resolution. The large cylindrical detector with an active volume of about $90\,\mathrm{m^3}$ [52] covers the full azimuth around the beamline. It has an inner radius of $85\,\mathrm{cm}$, an outer radius of $250\,\mathrm{cm}$, and a length of $500\,\mathrm{cm}$ along the z-direction [49]. The chamber is divided into two parts by a central high voltage electrode. The endplates on each outer end of the cylinder are divided into 18 trapezoidal sectors in the azimuthal direction and 2 regions in the radial direction (due to the radial dependence of the track density), resulting in 72 Readout Chambers (ROCs) overall. The interior of the TPC is filled with a $\mathrm{NeCO_2N_2}$ or $\mathrm{ArCO_2N_2}$ gas mixture with a small radiation length and low multiple scattering rates. During Run 2, the gas was changed from the Neon mixture to the Argon mixture and back again due to large observed space-charge distortions.

Charged particles traversing the gas-filled chamber ionise the gas molecules on their way and the freed electrons drift towards the endplates, under the influence of a precise axial electric field of $400\,\mathrm{V/cm}$ in the beam direction, with a maximal drift time of $t_\mathrm{D} \sim 90 - 108\,\mathrm{\mu s}$, depending on the gas mixture [49]. At the endplates, the electron signal produces an avalanche in multi-wire proportional chambers (MWPCs), which induces a mirror charge that is read out via clustering by several readout pads. The readout chambers contain 159 rows of pads along the radial extension and about 560 000 pads overall, which means that a passing track can maximally induce 159 clusters in the TPC. Not in every pad row that a particle crosses a cluster is detected. Therefore, the *number of crossed rows* is defined as the sum of the number of clusters and the number of pad rows without signal, but with clusters in both adjacent rows. The *number of findable clusters* is the number of pad rows, which have possible clusters based on the geometry of the track and the chambers.

The track position in $r\phi$ can be determined by the pad position of the deposited charge, whereas the position in the zy-plane must be calculated using the total drift time and the drift velocity of the electrons arising from the ionisation. This is how the information for a full 3D reconstruction of the tracks is provided.

A measurement of the particle's momentum simultaneously to the specific energy loss per unit path length via ionisation ($\mathrm{d}E/\mathrm{d}x$) in the TPC provides PID information for charged particles over a wide momentum range. The specific energy loss via ionisation is proportional to the number of released charges and can be described by the Bethe-Bloch formula [3], which depends on the particle species and its momentum, as well as the properties of the traversed medium. The PID information can be extracted by comparing the measured energy loss in the TPC with the expected $\mathrm{d}E/\mathrm{d}x$ for a specific particle species and momentum. The expected energy loss is based on a parameterisation



(a) (b)

Figure 2.2.: (a) Measured specific energy loss in the TPC in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02\,\text{TeV}$ as function of the rigidity. The black lines represent the expected energy loss calculated with the mass hypothesis of the indicated particle species. (b) Particle velocity (β), measured by the TOF detector, as function of momentum in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02\,\text{TeV}$.

of the Bethe-Bloch equation [53]

$$f(\beta\gamma) = \frac{P_1}{\beta^{P_4}}\left[P_2 - \beta^{P_4} - \ln\left(P_3 + \frac{1}{(\beta\gamma)^{P_5}}\right)\right], \tag{2.1}$$

where P_{1-5} are parameters from fits to measured data, β is the particle velocity, and γ is the Lorentz factor.

Figure 2.2a shows the measured specific energy loss in the TPC in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02\,\text{TeV}$ as function of the rigidity (momentum divided by charge) of positively charged tracks. The decisive quantity for PID is typically the resolution σ of the dE/dx measurement of a track [52]. A specific particle species i can be identified by the deviation of the measured dE/d$x_{\rm TPC}$ from the expected energy loss, $\langle \mathrm{d}E/\mathrm{d}x\rangle_i$, in terms of the measurement resolution, $\sigma_{\mathrm{d}E/\mathrm{d}x}$, as follows:

$$n^i_\sigma = \frac{\frac{\mathrm{d}E}{\mathrm{d}x}_{\rm TPC} - \left\langle\frac{\mathrm{d}E}{\mathrm{d}x}\right\rangle_i}{\sigma_{\mathrm{d}E/\mathrm{d}x}}. \tag{2.2}$$

The TPC has excellent PID capabilities on a track-by-track basis for light particles below 1 GeV/c, and via statistical methods for higher momentum light particles [53]. In order to cover also tracks up to a few GeV/c, the PID information has to be complemented by additional measurements provided by other subdetector systems [45].

2.3. Time-of-Flight detector

The Time-Of-Flight (TOF) detector is a large array of Multi-gap Resistive-Plate Chambers (MRPC), providing PID information for the intermediate momentum range. It covers the full azimuth with 18 sectors in ϕ and 5 segments in the z-direction, at 370-399 cm radius from the beamline. Each module consists of a group of 10-gap double-stack MRPCs, which are placed inside a gas-filled box transversely to the beam direction. Each MRPC is made up of two stacks of equally spaced resistive glass plates, creating 10 small gas-filled gaps. A uniform electric field is provided over the whole sensitive volume [49].

Traversing charged particles will ionise the gas between the glass plates and start a gas avalanche process, which produces a signal on the pick-up electrodes at the outer surfaces. The total signal is taken as the sum of the signals from all gas-filled gaps, resulting in an intrinsic time resolution of about $40\,\mathrm{ps}$ [49]. The PID information is based on the time-of-flight $\tau = t_{\mathrm{TOF}} - t_0$ of the particles from the interaction point to the TOF detector, with the arrival time t_{TOF}. The start time t_0 of the corresponding event is determined by the T0 detector, which is composed of two arrays of Cherenkov counters placed on each side of the nominal interaction point at forward rapidity. To provide increased resolution at higher multiplicities, t_0 can also be estimated by the particle arrival times in the TOF detector using a combinatorial algorithm based on a χ^2 minimisation between all possible mass hypotheses [53].

The particle velocity can be expressed by $\beta = L/(c\tau)$, with L being the length along the particle trajectory. The TOF PID performance is shown in Figure 2.2b, where the measured velocity as function of the measured particle momentum for Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ is plotted. Thus, the TOF detector provides separation for pions and kaons below $2.5\,\mathrm{GeV}/c$ and up to $4\,\mathrm{GeV}/c$ for protons and kaons, with a significance better than 3σ [53].

2.4. Track and vertex reconstruction

A collision with usually one main interaction point, referred to as primary vertex (PV), and a variety of tracks measured in the detector, is called *event*. For the reconstruction of an event, the tracking in the central barrel plays a major role. In the previous section, the TPC detector was used as an example to explain the effect of charged particles travelling through the detector material. Each particle carrying an electric charge induces signals measuring the position in space where it has passed the detector. The resulting task is to allocate these space points to individual tracks and reconstruct the related kinematics. Due to the high charged-particle multiplicity density, the track finding and vertex reconstruction become a challenging task [53].

The tracking process in the central barrel [53] involves several reconstruction steps. The first step towards the reconstruction of a track is the clusterisation process, which is performed separately for each detector. During this process, the detector signal is converted into clusters characterised by their position and the related errors.

The tracking in the central barrel starts with the estimation of a preliminary interaction vertex using clusters in the two layers of the SPD. Pairs of clusters in the SPD are used to reconstruct *tracklets*, and the interaction vertex is defined as the point where most of the tracklets converge.

The actual track reconstruction follows an inward-outward-inward scheme [53] starting with the track finding step in the outermost pad-rows of the TPC by building track seeds with several clusters and the position of the estimated interaction vertex as a constraint. The built seeds are propagated inwards to the inner TPC radius and updated at each step by assigning clusters that fulfil certain proximity cuts using the Kalman Filter algorithm [54]. Only those tracks with at least 20 out of 159 possible clusters in the TPC are accepted for reconstruction. For the propagated tracks, a preliminary PID based on the ionisation loss in the TPC is determined.

The reconstructed TPC tracks arising from the described track finding step in the TPC are then propagated to the outer layer of the SSD. They become the seeds for the track finding in the ITS, which follows in principle the scheme used for the TPC but taking into account the higher track density. For each TPC track, a decision tree is built and filled with all prolongation candidates fitted to exactly this track sorted by the reduced χ^2 of this fit. After applying an algorithm preventing two tracks from sharing too many clusters between each other, the track with the highest quality according to its χ^2 from each decision tree is added to the reconstructed event. Figure 2.3 shows the so-called ITS prolongation efficiency, which is the fraction of TPC tracks that are prolonged to the ITS. For the requirement to find at least two clusters in the ITS (black markers), the matching efficiency is about $95\,\%$ in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$. The requirement of at least one cluster in the SPD (red markers) results in a matching efficiency of about $85\,\%$. With those clusters not used for the ITS-TPC tracks, a standalone ITS reconstruction is done due to the decreasing acceptance and reconstruction efficiency for low momenta in the TPC.

In the second step, the reconstructed tracks are first extrapolated to their point of clos-

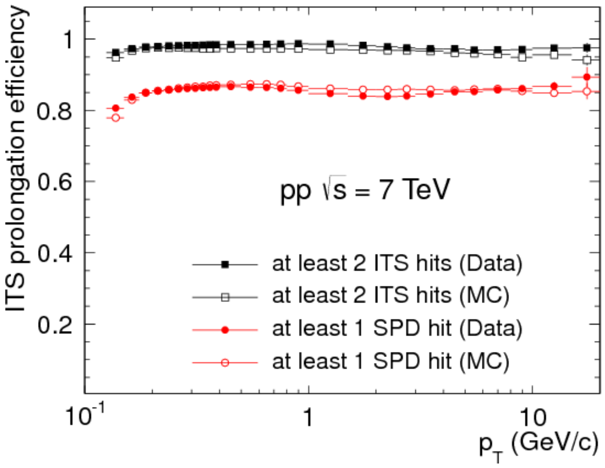


Figure 2.3.: ITS prolongation efficiency as function of transverse momentum for pp collisions in real data (full markers) and simulated MC data (open markers) with different requirements of ITS layer contributions [53].

est approach to the preliminary interaction vertex and then propagated in the outwards direction by the use of the Kalman Filter algorithm and of the clusters found in the track finding stage. At each updating step, several pieces of information are updated until the filtering reaches the TRD where the tracks are matched to tracklets in the six layers of the subdetector. In a similar way, the tracks are matched afterwards to clusters in the TOF and signals in the EMCal, PHOS, and HMPID.

In the last step, the reconstructed tracks are again propagated from the outer TPC radius inwards to the interaction vertex and the track state vector parameters together with its covariance matrix are determined. The last stage of the event reconstruction procedure is the final determination of the PV by a precise vertex fit.

2.5. ALICE detector upgrade

During the LHC long shutdown 2 (LS2), major upgrades [55] were installed in the ALICE detector to prepare it for the upcoming larger interaction rates in Pb–Pb collisions in Run 3. The main novelties are the upgrade of the TPC readout system enabling continuous data-taking, the replacement of the ITS of Run 1 and 2 together with a smaller-diameter beam pipe to allow for an improved tracking and vertex resolution, the new Muon-Forward-Tracker (MFT) [56], the Fast-Interaction-Trigger (FIT) [57] replacing the former trigger detectors, and a new online-offline (O^2) software framework.

One of the key upgrades in the central barrel is the installation of the new ITS [58], which fully replaces the former ITS and consists of seven layers based on Monolithic Active Pixel Sensors (MAPS). Due to the newly installed beam pipe with an inner radius of 18.2 mm, the inner layers are located closer to the beam axis, which is a major ingredient for the improved resolution on the track impact parameter measurement. Furthermore, the tracking and momentum resolution will be improved significantly by a reduction of the overall material budget. The design aims at a high readout rate of 50 kHz in Pb–Pb (compared to previous few kHz) and 1 MHz in pp collisions. These improvements will allow for more precise physics measurements, especially at low transverse momenta.

The second major central barrel upgrade is related to the readout system of the large TPC [59]. The MWPCs, which included a gating grid to prevent ion backflow into the chamber gas, were limited to a readout rate of a few kHz. To be able to cope with the increased interaction rate, the MWPCs are replaced by Gas Electron Multiplier (GEM) chambers for signal amplification. The new chambers allow for a continuous readout mode resulting in the possibility of a full reconstruction of all collisions. However, the accumulation of space charge in the drift volume due to the continuous readout leads to drift field distortions. Hence, the TPC upgrade to a continuous readout mode demands for complex and innovative calibration and correction procedures, as well as ways for efficient data compression, which can be tackled by the new software framework.

The prospects of Run 3 data-taking related to the work in this thesis are discussed in chapter 7.

3. Analysis strategy

3.1. Decay chain reconstruction with the KFParticle package

Figure 3.1 schematically shows the weak decay of a prompt Ξ_c^+ baryon, which is reconstructed in this analysis and where prompt means that the Ξ_c^+ is created in the PV of the collision.

The usual method of reconstructing short-lived particles is by determining the secondary vertex of their decay as the point at the distance of closest approach (DCA) between the daughter particle trajectories. The mother particle is then directly reconstructed by extrapolating the daughter particles' parameters to that vertex, where their momentum and energy are summed up. That is, traditional vertexing packages put focus on the reconstruction of the production and decay vertices.

The KFParticle package, on the other hand, provides a method for the estimation of the decayed particle's parameters and the associated covariance matrix in addition to the reconstruction of the production and decay vertices [60]. It is a software package that has been developed for the complete reconstruction of short-lived particles. The underlying algorithm is based on the Kalman filter method [54].

The Kalman Filter algorithm is a mathematical procedure for the iterative estimation of the state of a dynamical system based on a set of measurements with inaccuracies. It takes a set of n random measurements m_k with $k = 1...n$ and provides an optimal estimate for an unknown state vector $\mathbf{r}$ with its covariance matrix $\mathbf{C}$. In the general case, $\mathbf{r}$ can change between two measurements. The covariance matrix contains all covariances of the state vector with the variances of the single state vector components on its diagonal. The algorithm [61] starts with an initial approximation of $\mathbf{r}_0$ and $\mathbf{C}_0$ (initialisation step). With the knowledge of the impact of one measurement on the change of the state vector, a prediction of the evolution of $\mathbf{r}$ and $\mathbf{C}$ is made based on this first approximation (prediction step). For each measurement m_k, the state vector is updated, $\mathbf{r}_k$, to give the optimum estimation based on the first k measurements (filter step). The estimate $\mathbf{r}_n$ after the last update is then calculated via a geometrical fit with all measurements and gives the optimal estimation of the state vector based on the total set of measurements.

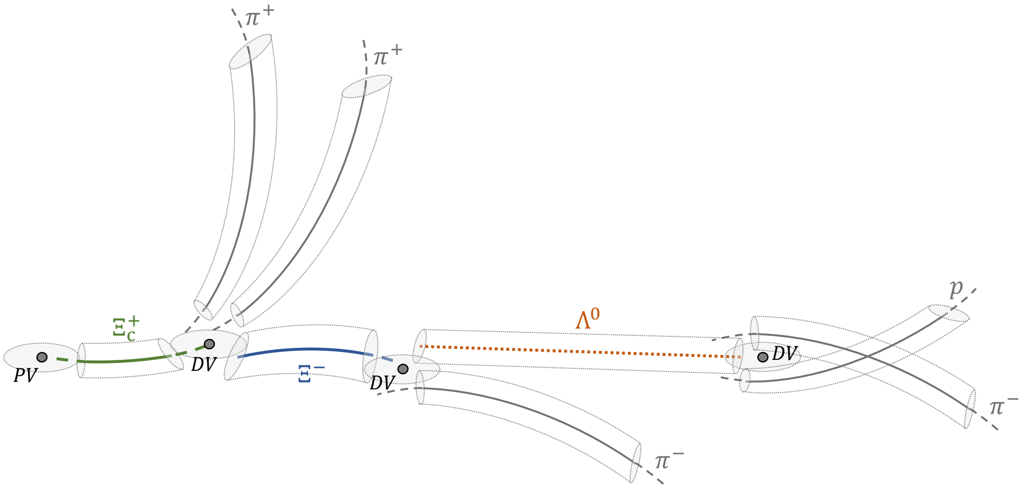


Figure 3.1.: Sketch of the hadronic decay of a prompt Ξ_c^+ baryon. The primary vertex (PV) is the collision point where the Ξ_c^+ baryon is produced. The localisation errors of the PV and the decay vertices (DV), the reconstructed particles, and the measured tracks are indicated in grey (not to scale).

Following this general algorithm of the Kalman filter, the first step in the reconstruction of a decayed mother particle is the approximation of the position of its decay point and its momentum and energy. This estimation is based on the kinematics of the daughter tracks. The KF software uses the following geometry-independent parameterisation of the particle as the state vector [61]:

$$\mathbf{r} = (x, y, z, p_x, p_y, p_z, E)^{\mathrm{T}} \tag{3.1}$$

with the position $(x, y, z)^{\mathrm{T}}$ along the trajectory, the particle momentum $(p_x, p_y, p_z)^{\mathrm{T}}$ and its energy E.

After the first approximation of the decay vertex, one of the daughter particles is transported to this initial vertex approximation. At this initial position, the daughter particle's momentum and its covariance matrix are calculated. The consequential estimate of the daughter m_1 is then used for the filter step of the Kalman filter to provide an estimation of the mother particle's state vector and covariance matrix. This filter step is repeated for the next daughter track m_2 until all n daughters of the decayed particle are treated.

After the reconstruction of the mother at its decay vertex, the parameter $s = \frac{l}{p}$ is added to the state vector $\mathbf{r} = (x, y, z, p_x, p_y, p_z, E, s)^{\mathrm{T}}$, where l is the length of the particle trajectory in the laboratory coordinate system and p is the momentum [60]. In the last optional step, all parameters of the particle are transported to its production vertex, which

is used as a measurement for the filtering of the particle's parameters at this vertex.
The final optimal estimation of the state vector and its covariance matrix contains all information needed to obtain a full description of the particle, both at the production and the decay vertex [61]. Hence, KFParticle allows the determination of the particle track in its full extent. Some information, that is not explicitly included in the state vector, can be easily calculated. The particle momentum p, the invariant mass M, the length of flight in the laboratory system L, and the proper time $c\tau$ are determined as follows [61]:

$$\begin{aligned} p &= \sqrt{p_x^2 + p_y^2 + p_z^2} \\ M &= \sqrt{E^2 - p^2} \\ L &= sp \\ c\tau &= sM. \end{aligned} \tag{3.2}$$

After assigning the daughter tracks to a reconstructed secondary vertex, they can be removed from the PV fit, which corresponds to a *subtraction* of a measurement from the vertex [60]. Analogously, it is possible to add the reconstructed mother particle as a measurement to the PV fit if the vertex corresponds to its production vertex. This procedure usually helps to improve the vertex precision. KFParticle fully considers the uncertainties of a track. This information is exploited when the tracks are used to reconstruct a decay. As a consequence, the daughter tracks are updated within their own uncertainty bands.

3.1.1. Constrained fits and variables

In many cases, it is possible to improve the estimator of the state vector obtained in the fitting procedure by applying certain assumptions on the features of the particle it represents in the form of constraints. These constraints are treated as one-dimensional measurements (in case of the particle mass, without error) by the Kalman filter in the secondary vertex fit [60].
If the invariant mass M of the reconstructed particle (see Equation 3.2) is known, the parameters of the particle can be refitted with the measured value M^2 and the measurement matrix $H_{M^2} = (0, 0, 0, -p_x, -p_y, -p_z, E, 0)$ [61]. This penalises the state vector parameters in the way that all daughter particles are required to form the invariant mass M, which updates the mass and momentum information of the reconstructed particle [62]. This is called the *mass constraint*. It is particularly important for long decay chains

like $\Xi^- \rightarrow \Lambda\pi^- \rightarrow p\pi^-\pi^-$ where the mass distribution of the reconstructed Λ introduces an uncertainty to the mass of the reconstructed Ξ^- [62].

In the case that an assumption can be made on the reconstructed particle's production vertex, the particle is transported to its production point and a *topological constraint* can be set to the state vector parameters. The constraint is used to align the mother particle to point to its production vertex (or any other vertex) [60]. If the particle is required to point back to the PV, the according vertex uncertainties can also be taken into account in the topological constraint.

After the fitting routine, different quantities describing the vertex fit quality can be extracted and subsequently used for the selection of specific reconstructed particle candidates.

The $\chi^2_{\mathrm{geo}}/\mathrm{NDF}$ refers to the geometrical fitting procedure for the reconstruction of a decayed particle by its daughter particles. It describes whether the trajectories of the daughter particles of a decay intersect within their uncertainties [62], and therefore expresses the quality of the vertex. NDF is the number of degrees of freedom. For a good vertex fit, the probability is high that the daughter trajectories intersect within their errors.

In the case of a reconstructed particle that has been assigned to a production vertex using a topological constraint, the $\chi^2_{\mathrm{topo}}/\mathrm{NDF}$ quantifies the probability of the hypothesis that the particle truly emerges from the assigned vertex. It characterises whether a particle is produced in the region of the production vertex taking into account the localisation error of the vertex and the uncertainties on the particle trajectory [62]. A large χ^2_{topo} indicates a tension between the considered particle and vertex (within their uncertainties), and a disfavour of the assumption that the given vertex is the production point of the particle.

3.2. Candidate selection with Machine Learning techniques

In the search for rare signals, it is crucial to reduce the number of background candidates to be able to extract these signals with good statistical significance. Especially 3-prong decays require careful consideration of possibilities to reduce the large combinatorial background.

Some limitations of a traditional analysis approach were already described above. Furthermore, many candidates do not show all characteristics of signal or background and are therefore rejected in a standard analysis based on ”rectangular” selection criteria if they fail only one particular selection criterion. In this way, it is possible that one criterion is decisive in the rejection of a particle candidate on its own, regardless of other selection criteria.

In order to use all available information, it is necessary to additionally exploit correlations in feature space. To add safety to the procedure, the selection can be optimised blindly. Therefore, in this analysis, Boosted Decision Trees (BDT) are used to classify signal and background. The algorithm is trained on a labelled set of signal and background candidates and learns the differences between the samples in feature space. For this binary classification task, the library *XGBoost* [44], which is a gradient boosting algorithm, is used.

3.2.1. Supervised learning approach for a binary classification task

Machine Learning techniques are widely used in data analysis of high-energy physics experiments to solve difficult classification or regression problems [63, 64]. In the case that a class label needs to be predicted, one speaks of a *classification* task, whereas in a *regression* problem, a quantity is to be predicted.

A supervised ML algorithm [65] takes a given dataset as input and uses its characteristics to carry out a classification or regression task. In the case of a classification problem, the dataset D consists of N input *instances* $\mathbf{x}_i = \{x_1, x_2, ..., x_n\}$ with n associated *features*, and a set of class labels $y_i \in \{0, 1\}$, which are the final output to be predicted. A labelled training sample $D = \{\mathbf{x}_i, y_i\}^N$ with N instances therefore contains a set of known $(\mathbf{x}, y)$-values.

Assuming there exists a function $f(\mathbf{x}) = y$ mapping $\mathbf{x}$ to y, the learning goal is to find an approximation $\hat{f}(\mathbf{x})$ of this function given D. $\hat{f}(\mathbf{x})$ is found by solving an optimisation task where a specified loss function $L(y, \hat{f}(\mathbf{x}))$, characterising the quality of the prediction for an object $\mathbf{x}_i$, is minimised [66]. Thus, the algorithm uses the training data to learn the correlations between the input variables and their label, to fit a model. This approximation can then be used to classify data with unlabelled input instances.

In addition to the training set, a similar test set is used to assess the quality of the trained model. Especially BDTs are sensitive to overfitting, which means that the trained model is too complex and unable to generalise. Such models are fitting the training data so accurately that even fine details are captured by it. This comes with the caveat that the model is not able to generally classify similar data. To control overfitting, the trained model is tested on a labelled test set similar to the training set. If the quality of the training and test samples differ significantly, the model is likely overfitting the training

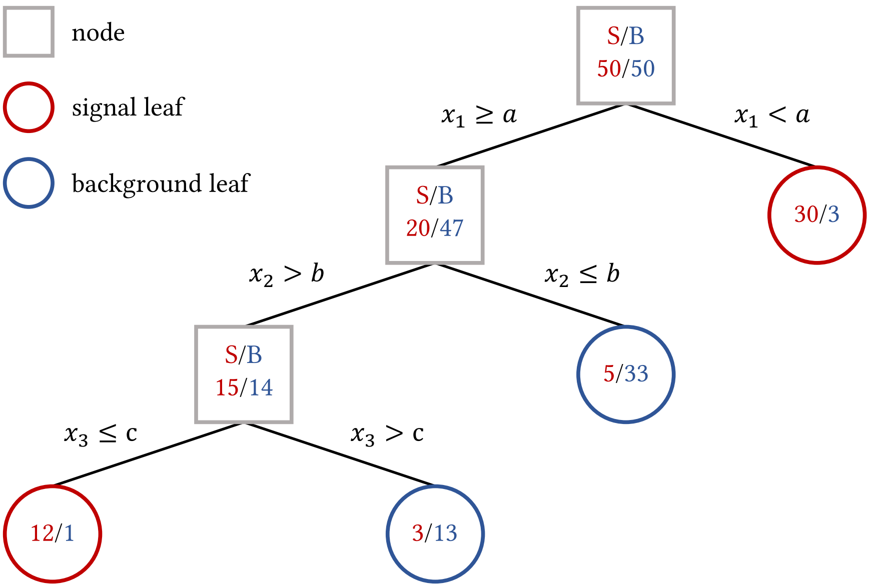


Figure 3.2.: Schematic of a decision tree for binary classification of signal (S, in red) and background (B, in blue) candidates. The initial sample at the first node is split into branches. The terminal nodes are shown as circles and are either signal leaves in the case that the signal candidates are dominant or background leaves if the opposite is the case.

data. In such cases, the model complexity needs to be reduced. This can be achieved in many different ways, for example by reducing the number of features describing the input instances, also discussed in section 4.2.

To control the quality of the model, the loss function L is combined with a regularisation term Ω, which penalises the complexity of the model. Together they form a so-called *regularised objective*, which needs to be minimised in the optimisation problem [67]. Thereby the resulting model will tend to be simple but predictive.

3.2.2. Boosted decision trees and XGBoost

Boosted Decision Trees are particularly robust ML algorithms [68] and are strong in dealing with missing values in the data or imbalanced datasets in general [66]. The ability to deal with these characteristics often makes them a favoured tool in high-energy physics.

A schematic of a single decision tree is shown in Figure 3.2. Suppose one has a labelled set of candidates, which need to be classified as either signal (red) or background (blue). Each candidate has several features, which can be used to distinguish between the two classes. Starting from an *initial node*, which contains all labelled candidates, one of the features is used to split the set into two parts, called *branches*, depending on the value of this feature. The splitting procedure is repeated for each of the branches, every time choosing the variable and splitting value with the best separation for the considered branch. At each stage, the branch which yields the highest increase in the quality of the separation is chosen to be split. At the end of the splitting procedure, several final branches remain, which become the so-called *leaves*. The leaves are either classified as signal if the signal candidates are dominating the branch or as background, if the opposite is the case [68].

Single decision trees are well established and have been used for quite some time [69]. Nevertheless, they are known to be unstable and therefore referred to as *weak learners* [66]. However, great improvement can be achieved by combining a large number of weak learners to form a decision tree ensemble. This procedure, which results in a powerful multivariate algorithm, is called *boosting* [66]. It is not limited to decision trees but can be applied to any classifier.

A generic boosting algorithm for a training sample $\mathbb{T}_k$ is implemented as follows [70]:

```
Initialise $\mathbb{T}_1$
for k in 1...$N_\text{tree}$
        train classifier $T_k$ on $\mathbb{T}_k$
        assign weight $\alpha_k$ to $T_k$
        modify $T_k$ into $T_{k+1}$
```

N_tree denotes the number of decision trees and T_k is the classifier at the k-th iteration. Assume one is starting with a sample of unweighted candidates from which a decision tree is built as described above. Whenever a candidate is misclassified by this tree, meaning a signal candidate which ends up on a background leaf or the other way around, this candidate's weight will be increased, it will be *boosted* [68]. Subsequently, a new tree T_{k+1} is created with the updated weights and the procedure is repeated. At each iteration k, the next tree T_k is added that most improves the model by minimising the regularised objective function [67]

$$\mathcal{L}^k = \sum_{i=1}^{N} L(y_i, \hat{y}_i^{k-1} + T_k(\mathbf{x}_i)) + \Omega(T_k), \tag{3.3}$$

where $\hat{y}_i^k$ is the prediction at the k-th iteration and N denotes the number of instances. In this way, the ensemble of decision trees is created by iteratively adding weak learners to the ensemble. The final boosted output will be a weighted average of all weak learners [70]

$$\hat{y}_i = \sum_{k=1}^{N_{\text{tree}}} \alpha_k T_k(\mathbf{x}_i), \tag{3.4}$$

where $\mathbf{x}_i$ is the i-th input instance with its associated discriminating features.
Each candidate will be followed through all of the trees in the ensemble and receives a score. Each time it lands on a signal leaf, it will be given a score 1 and if it is assigned to a background leaf, it will get the score -1. The final score is the renormalised sum of all the scores including possible weights [68]. A high score will be connected to a high probability of the candidate being signal and if the score is low, the candidate is most likely background.
In the case of *gradient boosting*, the Tailor expansion of the loss function L up to the second order is used to optimise the objective function in Equation 3.3 [71]. *XGBoost*, which stands for "Extreme Gradient Boosting", [44] is an optimised gradient boosting library that is used for the binary classification task in this work. It is based on a decision tree ensemble model consisting of several classification trees, or CARTs, as described above.
Before the model is trained in XGBoost a list of parameters, which control the structure and the functioning of the algorithm, have to be set. They either depend on the single trees in the ensemble or on the model itself. The list of these *hyperparameters* used in this work to describe the model is the following:

- **Maximum depth**: Number of nodes along the longest path from the initial node to the last leaf. It indicates the maximum depth of a tree in the ensemble.

- **Learning rate**: Magnitude of change between sequential trees. It quantifies how much the weights are adjusted for each boosting iteration and therefore controls the step size of the gradient boosting.

- **Number of estimators**: Number of decision trees included in the ensemble model.

- **Minimum child weight**: Lower bound on the number of candidates in a node that is required to undergo further splitting. If the construction of a tree results in a node where the sum of the instance weights is less than the value of this parameter, a leaf is formed and the building process of the tree stops at this point. This parameter therefore controls the node splitting.

- **Subsample ratio of training instances**: Ratio of the subsamples of the instances used for training, which are the rows in the training data. XGBoost can randomly sample the training data set before growing the tree at each boosting iteration, also called *bagging*. The default value is 1, which means that there would be no subsampling of rows. Decreasing the value will lead to stronger subsampling.

- **Subsample ratio of columns**: Ratio of the subsamples of the attributes used for building each tree. XGBoost can randomly sample the columns of the training data before growing the tree at each boosting iteration. As for subsampling of the rows, the default value is 1.

The hyperparameters are directly affecting the performance of the classifier [66]. Therefore, the tuning of these parameters improves the performance of the given ML algorithm compared to the default parameters. Since some of the hyperparameters are depending on each other, they should not be tuned independently. For example, the boosting specific learning rate, which controls the step size of the gradient boosting, is correlated to the number of estimators. A large number of trees in the ensemble can lead to overfitting, which can be prevented by adjusting the learning rate [66]. In this work, the hyperparameters are optimised with a Bayesian approach, which is implemented in the Heavy-Ion Physics Environment for Machine Learning (hipe4ML) [72], which provides helper functions for the python ML toolkit scikit-learn. The approach is described in subsection 4.2.2.

4. Data analysis

In this work, the Ξ_c^+ baryon ($m = 2467.71 \pm 0.23\,\mathrm{MeV}/c^2$) is measured via its hadronic decay to two positively charged π and a Ξ^- baryon, which further decays into a Λ and a π^-. The neutral Λ baryon further decays into a proton and a π^-. The results in this work are presented for prompt Ξ_c^+, which are produced in the primary collision. The contribution from Ξ_c^+ coming from the decay of a beauty hadron (feed-down) is removed. Less than $10\,\%$ of the prompt Ξ_c^+ in the decay channel $\Xi_c^+ \to \Xi^-\pi^+\pi^+$ are decaying via a resonance $\Xi_c^+ \to \Xi(1530)^0\pi^+ \to \Xi^-\pi^+\pi^+$ [3]. This fraction of resonantly decaying particles has to be taken into account in the training process of the BDT model for the binary classification of signal and background, as well as in the efficiency calculation since they can be affected differently by some selection criteria.

The analysis is performed in the transverse momentum range $3 < p_\mathrm{T}(\Xi_c^+) < 12\,\mathrm{GeV}/c$, divided into four intervals, and therefore provides the first measurement of the Ξ_c^+ baryon down to $p_\mathrm{T}(\Xi_c^+) = 3\,\mathrm{GeV}/c$ in ALICE. The results are compared to a measurement of the prompt Ξ_c^0 production cross section and a previously measured prompt Ξ_c^+ cross section (for $p_\mathrm{T}(\Xi_c^+) > 4\,\mathrm{GeV}/c$) obtained with standard analysis techniques [41].

4.1. Candidate reconstruction

4.1.1. Monte Carlo sample and event selection

The minimum-bias triggered data sample used in this analysis consists of pp collisions at the centre-of-mass energy $\sqrt{s} = 13\,\mathrm{TeV}$ recorded by ALICE in the years 2016, 2017 and 2018 during LHC Run 2. A total number of 1.9 billion events are analysed, which corresponds to an integrated luminosity of $\mathcal{L}_\mathrm{int} = 32.08 \pm 0.51\,\mathrm{nb}^{-1}$. This value is obtained from three van der Meer scans (one per year) and reported in [73] together with the determination procedure.

The analysis also requires the use of signal generated in Monte Carlo (MC) simulations for the ML training and to determine the reconstruction efficiency. About 60 million MC events from the event generator PYTHIA 8 [74] with the Monash tune [34] are used in this analysis. Two-thirds of the events are produced by injecting a $\mathrm{c\bar{c}}$ pair, whereas a $\mathrm{b\bar{b}}$ pair was injected for the remaining events. Each event by definition has to contain a Ξ_c^+ baryon, which is forced to decay via the decay channel of interest, $\Xi_c^+ \to \Xi^-\pi^+\pi^+$.

The simulated final state particles are transported through the detector material by the detector response simulation code GEANT3 [75]. The digitised result closely resembles the real data produced in the detector. Two independent samples of the dedicated MC production are used for the training of the classifying ML model and the reconstruction efficiency and acceptance correction during the computation of the production cross section.

To select relevant events for a physics analysis, different selection criteria are required to be satisfied by the MB triggered events in the offline analysis of the recorded data before the particle reconstruction.

The recorded events can include multiple collisions, so-called *pileup*, either occurring in the same crossing of the LHC bunches (same-bunch-crossing pileups), or in bunch crossings different from the one that triggered the data acquisition (out-of-bunch pileups) [76]. Pileup events of the first nature can be identified based on the fact that multiple vertices, separated by a few cm along the beam direction, are reconstructed from measured tracks since the collisions occur close in time. In the case of out-of-bunch pileup events, belonging tracks can be identified by correlating the information in different detector systems, since the effect of these events is different on each subdetector due to different readout times. Pileup events with multiple collisions are not of interest for this physics analysis and are therefore excluded in the event selection by rejecting events with multiple vertices (less than $1\,\%$ [73]).

To avoid edge effects, the position of the PV along the beam direction is required to be reconstructed within the range of $\pm\,10\,\mathrm{cm}$ from the nominal interaction point.

4.1.2. Decay reconstruction and preselection

The reconstruction of the Ξ_c^+ baryon is conducted using the KFParticle package and it starts with the selection of unlike sign tracks of protons and pions, which are combined to reconstruct the secondary vertex of the Λ baryon decay. Particles like the Λ baryon, which are neutral and carry strangeness, have a displaced decay vertex from the PV and are not tracked in the detector. Their oppositely charged decay products leave a typical V-shaped signature in the detector, which is why this type of decay is referred to as V^0 decay.

The particle identification (PID) of the two daughter tracks is ensured by the $\mathrm{d}E/\mathrm{d}x$ measurement in the TPC. For the V^0 reconstruction, protons and pions are selected with the criteria $|n\sigma_{\mathrm{TPC}}|(\mathrm{p}) < 3$ and $|n\sigma_{\mathrm{TPC}}|(\pi) < 3$. Furthermore, the number of clusters in the TPC from which the tracks are reconstructed is required to be larger than

70, whereas the number of clusters used for the determination of the specific energy loss must be at least 50. The ratio of the number of crossed rows over the number of findable clusters in the TPC is required to be larger than 0.8. Finally, for the pseudorapidity of the two daughter tracks the criterion $|\eta| < 0.8$ must apply, to avoid any edge effects due to the limited pseudorapidity coverage ($|\eta| < 0.9$) of the central barrel detectors. In order to select good quality Λ candidates, the candidate mass should not deviate by more than $0.01\,\mathrm{GeV}/c$ from the mass given by the Particle Data Group (PDG) ($1115.683 \pm 0.006\,\mathrm{MeV}/c^2$) [3].

The next step is the reconstruction of the Ξ^- decay vertex by combining the selected Λ candidates with tracks of secondary π^-. For these pion tracks, the same track selection criteria as for the V^0 daughter particles are applied. The TPC PID of the π^- track is complemented by the information from the TOF detector, in case it is available, by applying the criterion $|n\sigma_{\mathrm{TOF}}|(\pi) < 5$. Additionally, the π^- is required to have a transverse momentum larger than $0.15\,\mathrm{GeV}/c$, which is the minimum transverse momentum of particles that are only reconstructed by the TPC. Similar to the V^0 candidates, the mass of the reconstructed Ξ^- is required to lie within a certain range from the mass given by the PDG ($1321.71 \pm 0.07\,\mathrm{MeV}/c^2$) [3] to reject background from outside the peak region. The corresponding selection criterion depends on the transverse momentum of the reconstructed Ξ_c^+ candidates. Below $6\,\mathrm{GeV}/c$, the mass should not deviate more than $\pm\,0.004\,\mathrm{GeV}/c^2$ from the central value PDG mass, whereas for $6 < p_{\mathrm{T}}(\Xi_c^+) < 8\,\mathrm{GeV}/c$ ($8 < p_{\mathrm{T}}(\Xi_c^+) < 12\,\mathrm{GeV}/c$) the mass window is chosen to be $\pm\,0.005\,\mathrm{GeV}/c^2$ ($\pm\,0.006\,\mathrm{GeV}/c^2$) around the PDG value.

In the last step of the reconstruction, the selected Ξ^- candidates are combined with two positively charged pion tracks to reconstruct the Ξ_c^+ candidates. The pion tracks fulfil the same selection criteria as before, apart from their transverse momentum, which is required to be larger than $0.4\,\mathrm{GeV}/c$ to remove part of the large combinatorial background at low transverse momenta. In addition, the tracks are required to have left at least 3 hits in the ITS to ensure that they are primary.

During the reconstruction, the mass constraint is applied to the Λ and Ξ^- candidates, as well as the topological constraint, which is fitting the Ξ^- and the Ξ_c^+ to the PV.

To reduce the combinatorial background at an early stage of this analysis and to exploit the full performance of the ML algorithm, the reconstructed Ξ_c^+ candidates are preselected based on some of their topological and kinematic decay features. The corresponding loose selection criteria are listed in Table 4.1, where the pointing angle (PA) is defined as the angle between the momentum vector of the reconstructed particle and

Table 4.1.: Preselection criteria applied in this analysis.

Decay feature	**Criterion**
$\lvert\eta(\Xi_c^+)\rvert$	< 0.8
$p_{\mathrm{T}}(\pi^+ \leftarrow \Xi_c^+)$	> 0.4
$\mathrm{PA}(\Lambda \rightarrow \Xi)$	< 0.5
$\chi^2_{\mathrm{topo}}(\Xi \rightarrow \mathrm{PV})$	$> 0.$
$\chi^2_{\mathrm{topo}}(\Xi_c^+ \rightarrow \mathrm{PV})$	$> 0. and < 50.$
$\chi^2_{\mathrm{geo}}(\Xi_c^+)$	$> 0. and < 50.$

the line connecting its assigned production vertex with its decay vertex. The value of the PA is expected to be small if the reconstructed particle points back to its production vertex. The pseudorapidity selection reported in Table 4.1 corresponds to the geometrical detector acceptance.

4.2. Binary classification with XGBoost

In this analysis, the gradient boosting machine XGBoost is used to separate reconstructed signal candidates from the background. Especially in the case of a rare signal like the Ξ_c^+ with large combinatorial background, it is necessary to achieve a good separation between candidates of both types. The previously described preselection of the reconstructed candidates already results in a rough separation of signal and background candidates. However, the classification can be significantly improved by the use of a BDT model.

The BDT model is trained on a set of candidates after selecting several input features for the classification and an optimised set of model hyperparameters. Subsequently, the model is tested on an independent set of candidates to validate its performance.

The analysis is conducted separately in four p_{T} intervals of the reconstructed Ξ_c^+ and a different model is trained, tested, and validated in each interval. In the following, p_{T} refers to the transverse momentum of the reconstructed Ξ_c^+ candidates.

4.2.1. Input sample

The input sample for the binary classification task consists of a set of reconstructed prompt Ξ_c^+ candidates, which are described by their geometrical, kinematic, PID, and topological decay features. The set contains true signal candidates from MC simulations (including resonant decay as well as direct decay), and combinatorial background,

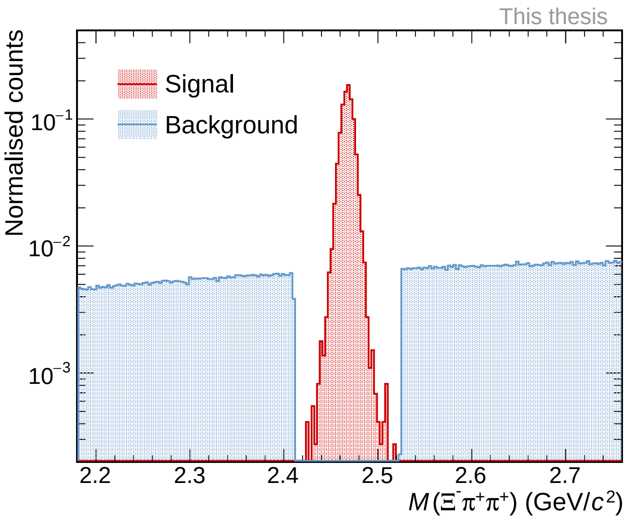


Figure 4.1.: Signal (red) and background (blue) invariant mass spectrum of the training instances in the interval $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$, normalised to the number of candidates.

which is taken from real data with the requirement that the candidate invariant mass lies outside the mass region of the Ξ_c^+ baryon ($2.411 - 1.525\,\mathrm{GeV}/c^2$). The input candidate invariant mass spectrum for the model trained in the p_{T} interval between 3 and $4\,\mathrm{GeV}/c$ is shown in Figure 4.1. Since the number of simulated events is limited, the full available MC sample is used for the classification task, whereas only a random fraction of $20\,\%$ of the real events is needed to provide enough background candidates for the model training and testing process. The proportion of true signal decaying directly and background for the models trained in the p_{T} interval between 3 and $6\,\mathrm{GeV}/c$ is $1:2$. This choice is made to maximise the number of candidates used for training, which improves the model performance while keeping the proportion as balanced as possible at the same time. For the p_{T} range above $6\,\mathrm{GeV}/c$, where the number of background candidates in real data rapidly decreases, all available candidates from the used data fraction are taken into account. The exact numbers of the different types of candidates used for training and testing of the BDT model are listed in Table 4.2.

To create independent training and test sets, the input sample is randomly split into two parts: $60\,\%$ for training and $40\,\%$ for testing. A performance study has been made to test the effect of different split values on the model outcome. Due to the limited number of candidates available for the model training and testing, and to address the fact that the model performance improves with an increasing number of training instances, the configuration of $60:40$ was chosen.

Table 4.2.: Number of signal and background candidates used for the BDT model training and testing in the analysed p_{T} intervals.

p_{T} (GeV/c)	(3, 4)	(4, 6)	(6, 8)	(8, 12)
Prompt, direct decay	14790	33381	21854	14976
Prompt, resonant decay	3326	9440	7331	5416
Background	29580	66762	34231	5951

4.2.2. Hyperparameter optimisation

Before the actual training process, the model hyperparameters have to be chosen. It is desirable to use a set of hyperparameters returning the best model performance. This optimisation task can be tackled with different methods. Apart from random or grid searches, which are extremely expensive, there is the much more efficient approach of *bayesian optimisation*, which is applied in this analysis. This iterative procedure uses the information from previous evaluations to create a mapping (called surrogate) of a specific set of hyperparameters to a probability of a score on the objective function. This probability model is updated after each evaluation of the objective function with a new set of hyperparameters. Therefore, the bayesian method tends to find better hyperparameters by reasoning about the best parameter set based on previous trials.

In principle, the algorithm can evaluate different sets of parameters until the model performs optimally on the used training data, which would lead to a non-generalisable model. To solve this overfitting problem, the k-fold cross-validation method can be used [77]. The approach involves randomly dividing the training data into k non-overlapping groups (folds) and evaluating each parameter set on $k-1$ folds. The remaining fold is treated as a validation set. This fitting procedure is repeated k times, each time permuting the folds used for optimisation and validation. The final cross-validation estimate is taken as the average over the resulting k permutations.

This analysis uses the bayesian optimisation method with k-fold cross-validation implemented in hipe4ML [72] with the parameter ranges listed in Table 4.3. The parameter range is defined in a way that the optimisation algorithm does not always converge towards the lower or upper edge of the given interval. In general, the choice of parameter ranges for optimisation needs to be handled carefully considering memory consumption, the risk of overfitting, performance, and conservatism of the resulting model. For example, particularly deep trees are expensive to evaluate, consume lots of memory, and introduce the risk of overfitting, at the same time they lead to increased model performance. Overfitting can be controlled by subsampling the training data, and by making

Table 4.3.: Parameter ranges for the hyperparameter optimisation and optimised set of model parameters for the analysed p_T intervals.

p_T (GeV/c)	(3, 4)	(4, 6)	(6, 8)	(8, 12)	range
Maximum depth	3	2	3	2	$(1, 3)$
Learning rate	0.02	0.08	0.06	0.04	$(0.01, 0.1)$
Number of estimators	363	240	241	363	$(100, 1000)$
Minimum child weight	2.7	6.8	5.9	2.3	$(1, 10)$
Subsample ratio of rows	0.93	0.91	0.86	0.87	$(0.8, 1.)$
Subsample ratio of columns	0.91	0.90	0.91	0.89	$(0.8, 1.)$

the algorithm more conservative, which can be achieved by choosing large minimum child weights or a small learning rate.

In this analysis, the training set is divided into five folds of approximately equal size for cross-validating the evaluation on the different parameter sets. Finally, the best set of hyperparameters is chosen after ten optimisation steps. Table 4.3 lists the optimal parameters found for the BDT models, which are applied in this analysis to extract the final results of this work.

4.2.3. Input feature selection

The BDT model performs the classification task based on a set of features. The choice of these training features should be treated carefully, taking into account different considerations.

Naively one would expect that a large number of different features used for separation results in the best classification possible. Though, there is a difference between good performance on the training data and a generally good model performance on different data sets. Too complex models are likely to be overfitting the training data. Therefore, it is always a trade-off between model performance and complexity.

Furthermore, correlations between input features have to be taken into account. If two training features are highly correlated, they may contain similar information and it should be considered to include only one of them to keep the model simpler. Features which are correlated for signal candidates and anticorrelated for background, or vice versa, are likely to have high discriminating power and can therefore be exploited in the classification. On the other hand, correlations between training features and the observable, which is the Ξ_c^+ invariant mass, for background candidates need to be avoided. They potentially lead to a modification in the background shape of the invariant mass

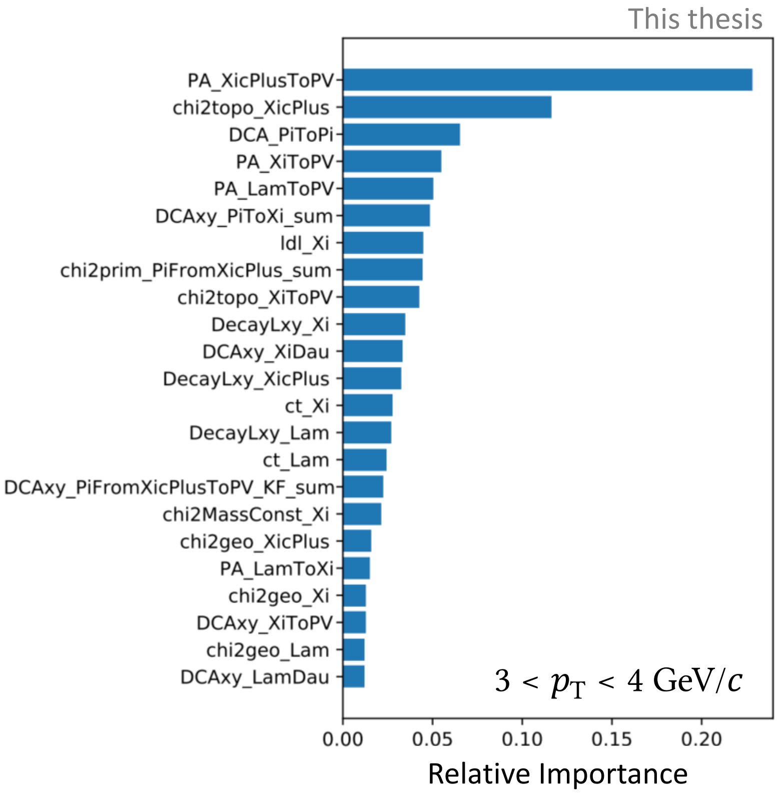


Figure 4.2.: Relative feature importance for the BDT model trained with candidates in the interval $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$ with all available decay features included in the training process.

spectrum, artificially enhancing or reducing the extracted signal.

To reach a not too complex model with high performance, a first model is trained in each p_{T} interval using all available decay features as input. Depending on how often the features are used in the building process of the BDT, a feature importance is assigned to them, which is the average impact on the model output. From this ranking, the features with the highest separation power and the largest gain in model performance are chosen. Figure 4.2 shows the feature importance ranking for the model trained in the interval $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$ including all available decay features in the training process. A definition of all listed variables can be found in Appendix A.
Additionally to this first model training, physical arguments have been considered during the choice of input features to ensure a good quality of the reconstructed vertices. The two most important features, the PA and the χ^2_{topo} of the Ξ_c^+ pointing back to the PV, characterise if the reconstructed candidate points back to the PV, its production ver-

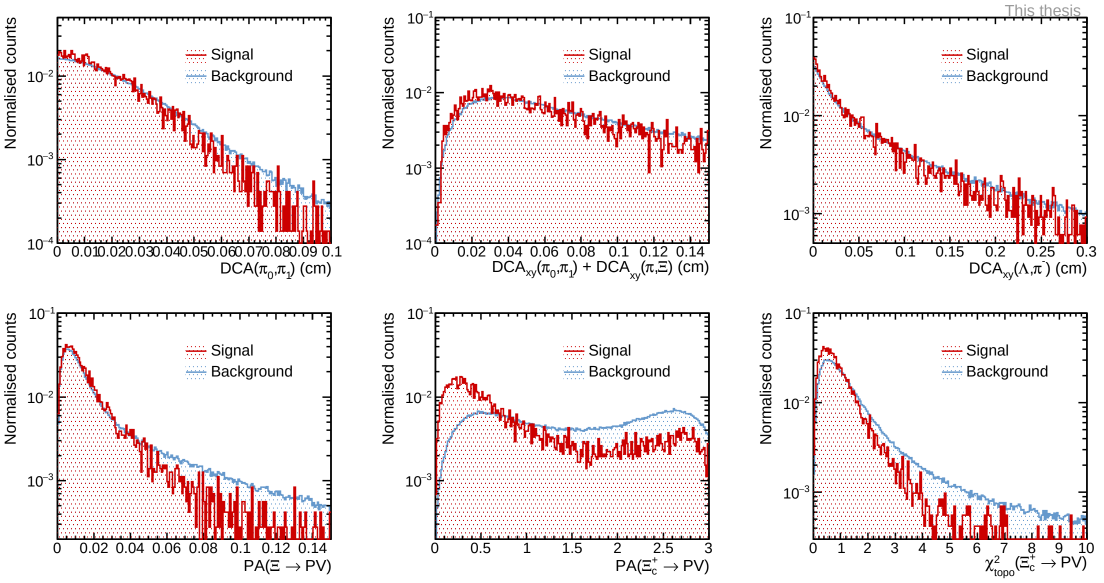


Figure 4.3.: Signal (red) and background (blue) distributions of the selected training features in the range $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$, normalised to the number of candidates. From left to right, top to bottom: DCA between the two pions coming from the Ξ_c^+ in three dimensions, $(\mathrm{DCA_{xy}}(\pi_0,\,\pi_1) + \mathrm{DCA_{xy}}(\pi_0,\,\Xi^-) + \mathrm{DCA_{xy}}(\pi_1,\,\Xi^-))$, DCA between Ξ^- daughters in xy-direction, PA of the Ξ^- to the PV, PA of the Ξ_c^+ to the PV, and χ^2_topo of the Ξ_c^+ to the PV.

tex. The values are small for true signal. Since the Ξ_c^+ is a very short-lived particle, the Ξ^- candidate should also point back to the PV, where the Ξ_c^+ is produced, for signal candidates. Therefore, the PA of the Ξ^- pointing back to the PV, which should be small for true Ξ_c^+, is also included in the model. Furthermore, to ensure a good quality of the Ξ_c^+, the DCA between the two primary pion tracks in three dimensions, as well as the sum of the DCA between the two pions, and the pions and the Ξ^- in xy-direction, $(\mathrm{DCA_{xy}}(\pi_0,\,\pi_1) + \mathrm{DCA_{xy}}(\pi_0,\,\Xi^-) + \mathrm{DCA_{xy}}(\pi_1,\,\Xi^-))$, are added to the model. Both have smaller values for signal candidates. Finally, the DCA between the Ξ^- daughters in xy-direction is taken as an input feature to guarantee a good Ξ^- vertex quality. Taking all these arguments into account and looking at the feature importance ranking, six different training features were selected. The distributions for signal and background candidates used in the training in $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$ are shown in Figure 4.3. A significant difference between the two distributions is a sign of high discriminating power. Especially the PA of the Ξ_c^+ to the PV is distributed differently for signal and background. True Ξ_c^+ candidates tend to lower values, whereas the background is distributed nearly uniformly. The χ^2_topo of the Ξ_c^+ to the PV also shows some deviation between signal and

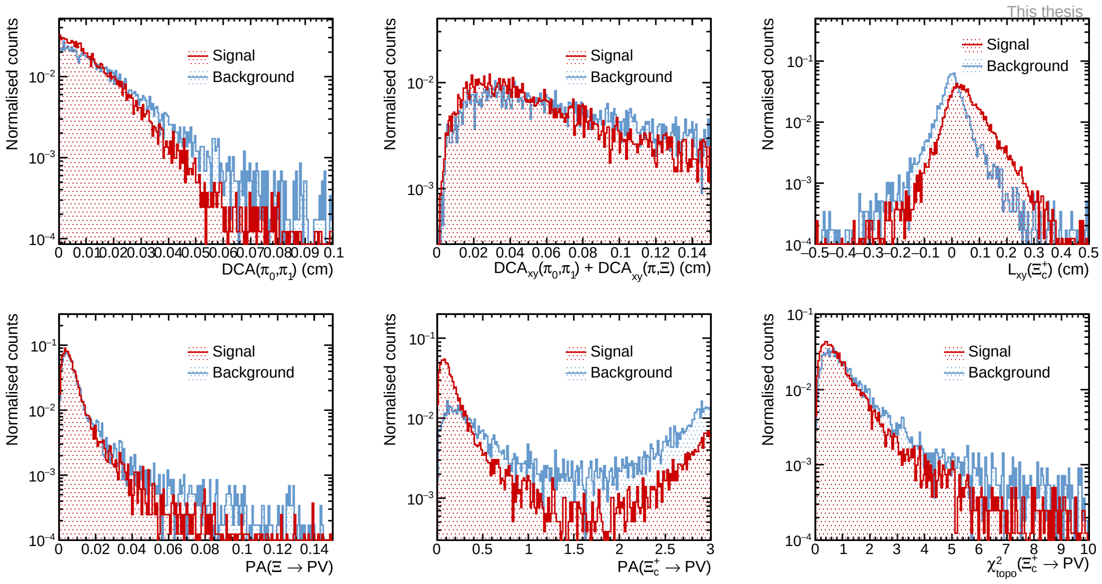
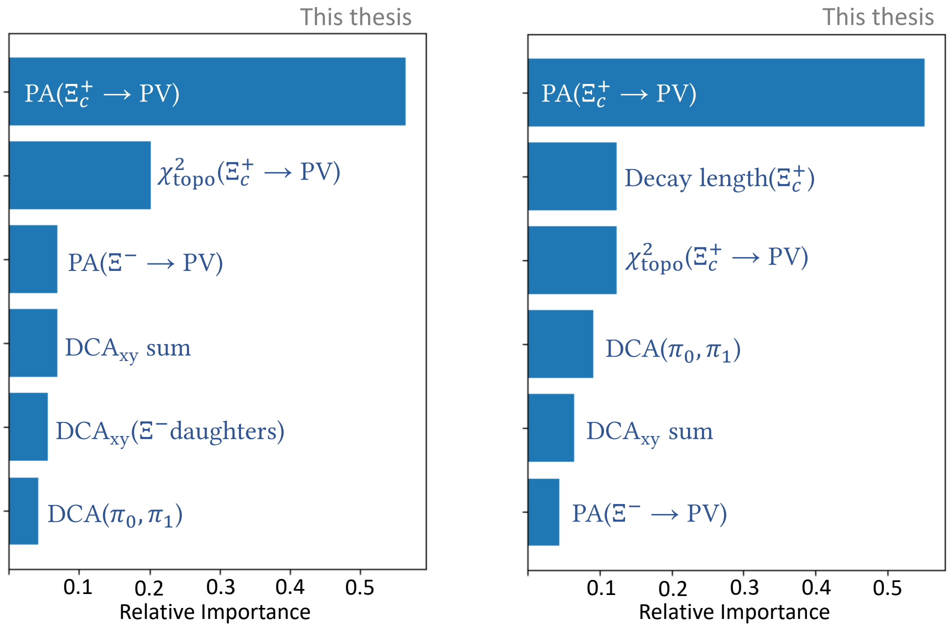


Figure 4.4.: Signal (red) and background (blue) distributions of the selected training features in the range $8 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$, normalised to the number of candidates. From left to right, top to bottom: DCA between the two pions coming from the Ξ_c^+ in three dimensions, $(\mathrm{DCA}_{\mathrm{xy}}(\pi_0,\,\pi_1) + \mathrm{DCA}_{\mathrm{xy}}(\pi_0,\,\Xi^-) + \mathrm{DCA}_{\mathrm{xy}}(\pi_1,\,\Xi^-))$, decay length of Ξ_c^+ in xy-direction, PA of the Ξ^- to the PV, PA of the Ξ_c^+ to the PV, and χ^2_{topo} of the Ξ_c^+ to the PV.

background, since the χ^2_{topo} of the topological constraint is small in the case that the particle points back to its assigned production vertex. These results hint at the fact that the two variables are the most decisive classification features in this analysis.

For high momenta, where the particles are Lorentz-boosted, the Ξ_c^+ decay length becomes an important feature to discriminate signal and background since the decay vertex is more displaced from the PV. For the models trained in the range $4 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$ the decay length of the Ξ_c^+ is therefore added to the training. It replaces the DCA between the Ξ^- daughters in xy-direction, which was found to be the least important of the selected features in this p_{T} range. The feature distributions for signal and background in the p_{T} interval between 8 and $12\,\mathrm{GeV}/c$ are shown in Figure 4.4. The highest deviation between signal and background is observed in the PA of Ξ_c^+ to the PV and the Ξ_c^+ decay length in xy-direction. The decay length in the xy-plane is calculated by the KFParticle package considering the direction of the particle momentum. The background is distributed uniformly around 0 since primary π^- are selected for the reconstruction of the decay vertex. For signal, on the other hand, the values are shifted to positive values due to the displaced topology in the case of true signal candidates. However, due to a limited



(a) $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$. **(b)** $8 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$.

Figure 4.5.: Relative feature importance ranking of the selected training features for the BDT models in the low and high p_{T} range.

vertex resolution, also negative values of the decay length in the xy-plane are observed. The relative importance of the selected classification features can be seen in Figure 4.5 for the low and high p_{T} models. For all trained models, the PA of Ξ_c^+ to the PV is most decisive, followed by the χ^2_{topo} of Ξ_c^+ to the PV for the lower p_{T} ranges and the Ξ_c^+ decay length at high transverse momenta. These results suggest that the decay topology of the Ξ_c^+ can indeed be fully exploited, even at low transverse momenta.

Figure 4.6 shows the correlation matrices for all selected classification criteria in the signal and the background sample for the range $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$. No correlation between any input feature and the Ξ_c^+ invariant mass is observed, which is an important criterion for the selection of training features. Several correlations, which are visible in the background sample but less pronounced in the signal, are likely to be exploited by the BDT model for the classification.

4.2.4. Sample weights

As mentioned earlier, $10\,\%$ of the Ξ_c^+ baryons in the analysed decay channel are decaying via a resonance, Ξ^*. In order to study the differences between the two types of signal

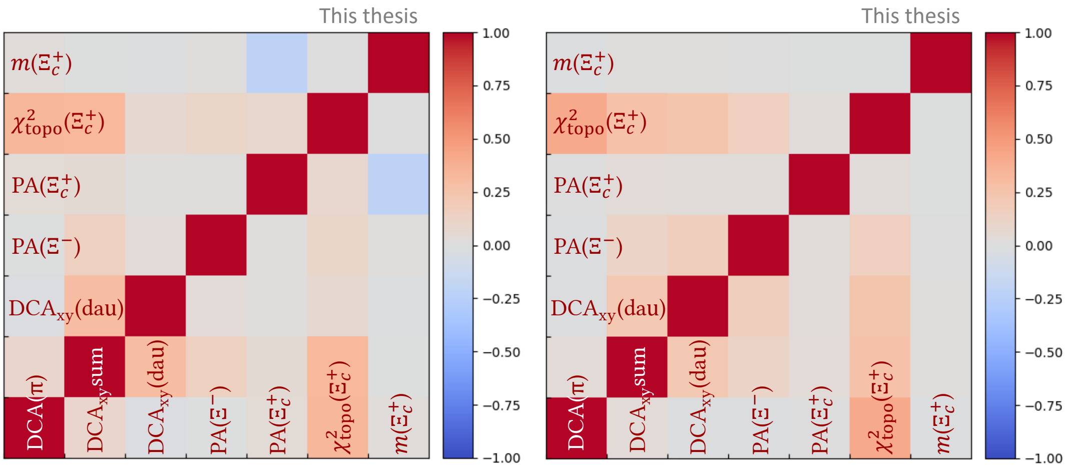


Figure 4.6.: Correlation matrix of the training features for signal (left) and background (right) in the range $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$. Correlations are indicated in red, and anticorrelations in blue.

candidates, the selected training features and the reconstruction and selection efficiency were investigated separately. A small deviation was observed in the DCA between the two primary pions. The distribution for true signal candidates coming from a direct decay and from a resonant decay, as well as for background taken from data, is shown in Figure 4.7 in the p_{T} range between 3 and $4\,\mathrm{GeV}/c$. The differences are explicable by the fact that in the case of the resonant decay, one of the two pions is originating from the resonance Ξ^* and not directly from the Ξ_c^+. The topology, as well as the kinematics (like the momentum distribution of the two π^+), will therefore differ compared to the direct decay vertex. This possibly results in a different reconstruction and selection efficiency for both candidate types.

Following these observations, the directly and resonantly decaying signal candidates have to be weighted in the training process according to their natural abundances. In order to exploit the full statistics, all available signal candidates decaying via a resonance are included in the input sample. Therefore, the ratio between directly and resonantly decaying particles varies across the different p_{T} ranges (see Table 4.2). Nevertheless, XGBoost allows assigning weights to the training instances during the fitting procedure. This functionality is used to weight resonantly and directly decaying candidates according to the ratio $1 : 10$ during the training, ensuring that the algorithm does not over-learn the characteristics of the resonant decays.

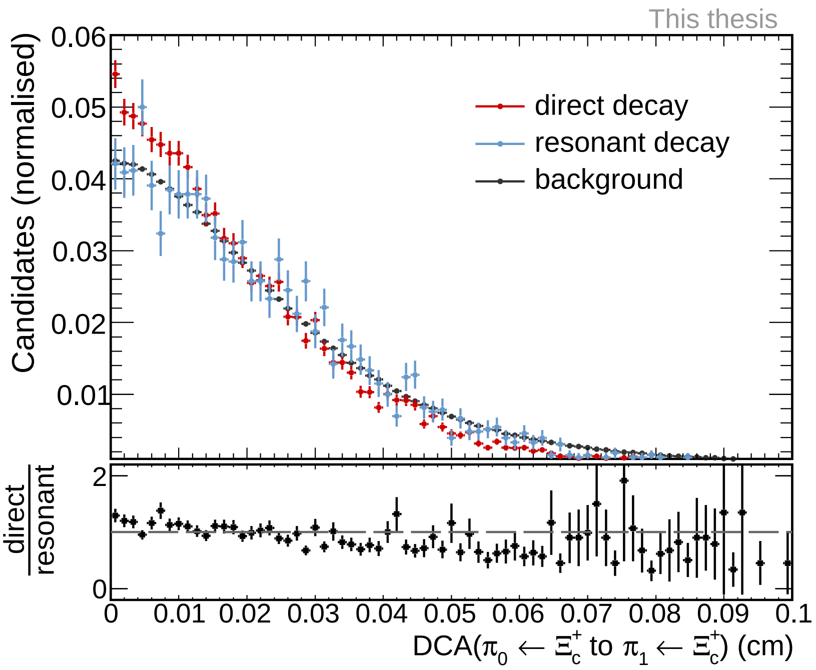


Figure 4.7.: The DCA between the two pions (π_0, π_1) from Ξ_c^+ for true signal candidates decaying via a resonance (blue) or decaying directly (red), as well as for background (black) in the range $3 < p_T < 4\,\text{GeV}/c$. The ratio between the red and blue distribution is indicated below the plot.

4.2.5. Model performance and output

After defining the input sample, selecting the training features, and setting the hyperparameters, the BDT model is trained on the training set. Subsequently, it is tested on the independent test set to evaluate the model performance. In the optimal case, the model is decisive and generalisable, meaning that it is neither overtrained nor undertrained and the deviation between the training and test set is small. Furthermore, a perfect model is at the same time $100\,\%$ efficient and $100\,\%$ pure. Efficiency refers to the ability to identify the instances of the signal class correctly, and purity connects to the number of instances that are falsely assigned to this class. These characteristics, which are combined in the model performance, can be evaluated in several ways.

The learning curves, shown in Figure 4.8 for the model trained in the p_T interval between 3 and $4\,\text{GeV}/c$, are defined as the root-mean-square error (RMSE) of the training set (red line) and the test set (blue line), which is the deviation of the model prediction from the observation. Therefore, the *experience* of the model is plotted on the x-axis as the number of candidates used for training, and the *learning* is plotted on the y-axis as the RMSE. For a training sample with only a few candidates, the fitting problem is trivial, thus the

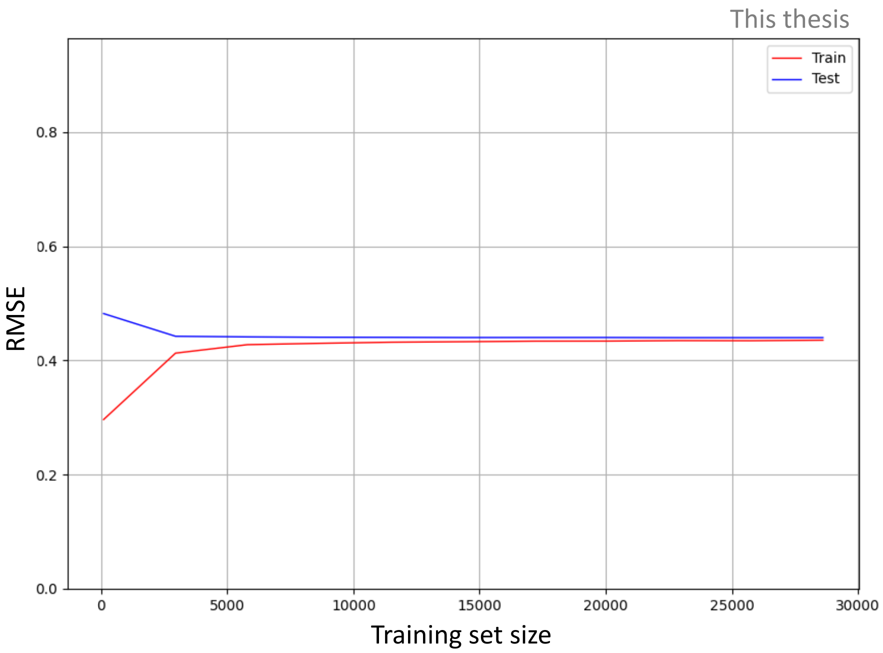


Figure 4.8.: Learning curves: Root-mean-square error (RMSE) of the training set (red) and the test set (blue) as function of the training set size for the model in the range $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$.

RMSE is small (red line), and the model captures the training data nearly perfectly. The test set, on the other hand, is poorly described by the model. When taking into account more training instances, the error on the training data increases, since more points are added to the fit. At the same time, the error on the test set decreases, which shows that the model becomes more general and is able to describe the test data. For a training set with roughly 5000 instances, the two curves start to converge and stabilise at a common value. Thus, with larger input samples, overfitting can be avoided and the model performance can be stabilised. Figure 4.8 therefore verifies that the model in the range $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$ is neither undertrained nor overtrained.

Another method to assess the model performance and to control for efficiency and purity is the Receiver Operating Characteristic (ROC). It represents the dependence of the signal *efficiency* on the *error rate* of a model by plotting the true positive rate against the false positive rate of the signal class for different classification values. The true positive rate is defined as the fraction of correctly classified instances out of all instances of the signal class (efficiency), and the false positive rate is the fraction of wrongly classified instances out of all instances of the background class ($1 -$ purity). The ROC curves for

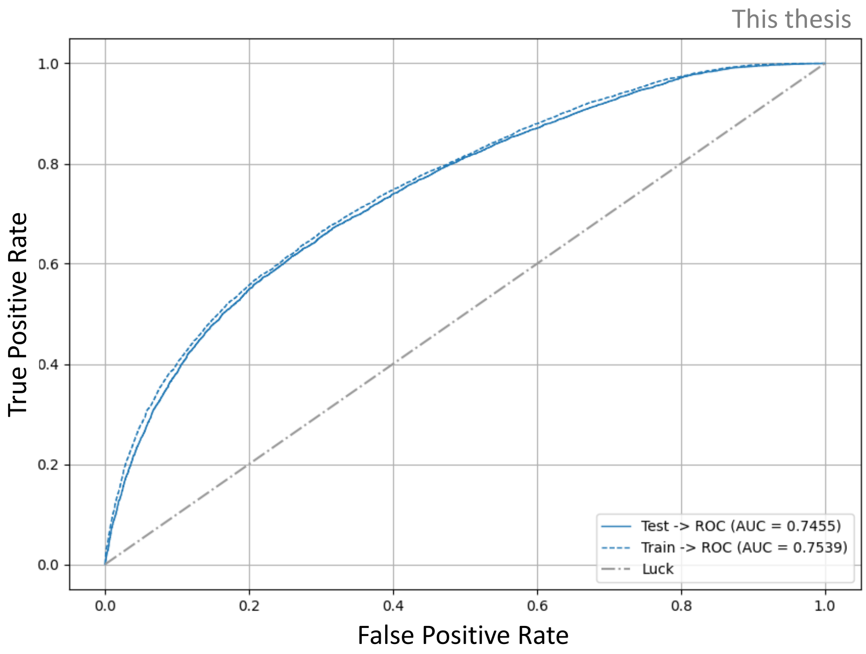
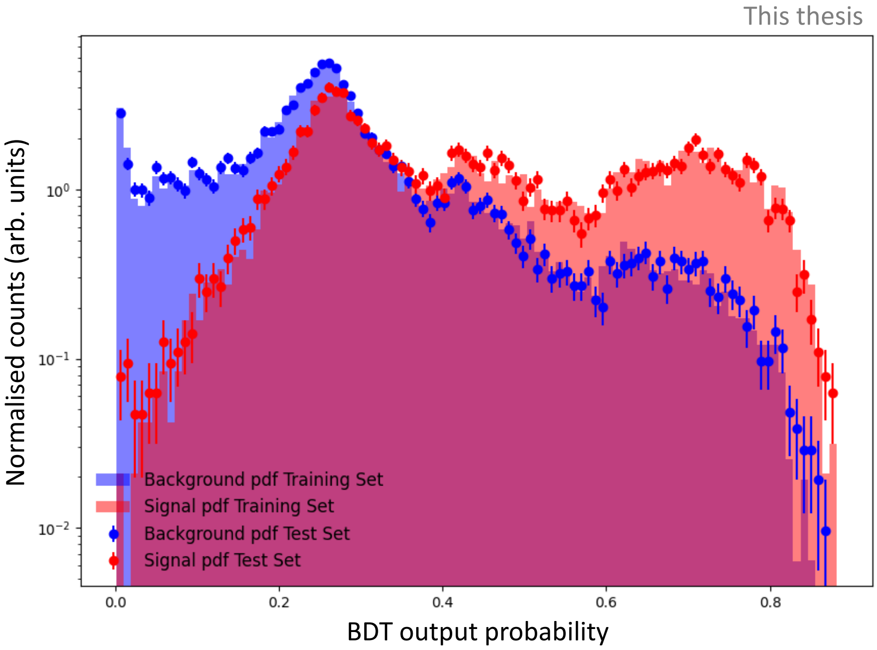


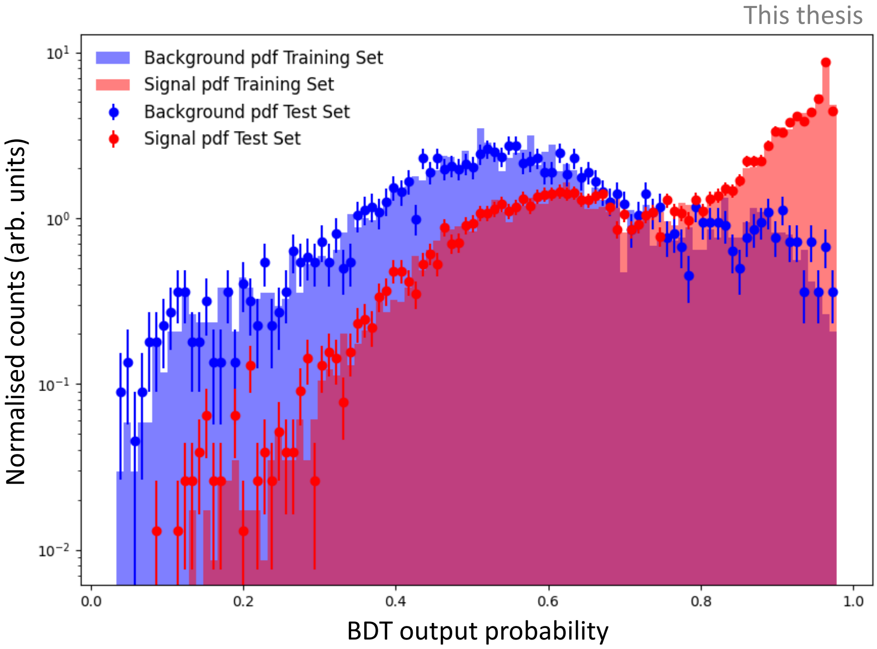
Figure 4.9.: Receiver Operating Characteristic (ROC) curves of the training set (blue dashed line) and the test set (blue solid line) for the model in the range $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$.

the model trained in the range $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$ are depicted in Figure 4.9. The blue dashed line shows the ROC curve of the training set and the solid line describes the test set. A large deviation between the two curves is a sign of overfitting, which cannot be observed in this case. The grey dashed line on the diagonal indicates the case if the success rate and error rate are equal and therefore describes a random classifier. A model with the ROC curve lying below this line is misinterpreting the data. The Area Under the ROC Curve (AUC) can be interpreted as the probability that the model classifies a true candidate of one class correctly. The perfect classification would result in a point in the top left corner of the plot. In reality, a $100\,\%$ pure model is likely to have low efficiency, meaning that all classified signal are indeed true signal candidates but with very small statistics. A model with $100\,\%$ signal efficiency, on the other hand, probably identifies all signal candidates correctly but will not reject much background. For a well-performing classifier, the AUC should therefore be maximised, and purity and efficiency need to be balanced.

For the remaining p_T intervals a similar or better performance was observed.



(a) $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$.



(b) $8 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$.

Figure 4.10.: Model output probability for signal (red) and background (blue) candidates in the training set (bars) and the test set (full markers) for low and high p_{T} models.

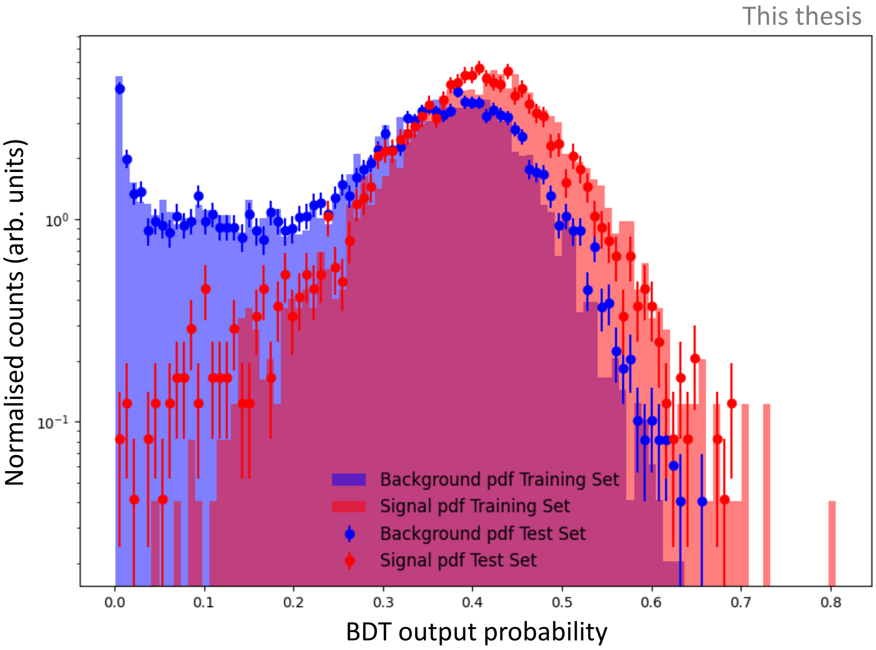


Figure 4.11.: Model output probability for signal (red) and background (blue) candidates in the training set (bars) and the test set (full markers) for the model in $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$, excluding PA($\Xi_c^+ \to$ PV) as classification feature. The model is not applied in this analysis.

The resulting BDT model classifies the candidates of the training set and the test set with a certain probability based on the training features. The model output is a single response variable for each candidate, which describes the probability of the candidate to be signal according to the BDT. This response variable, the BDT probability, is shown in Figure 4.10 for signal (red) and background (blue) candidates of the training set (bars) and the test set (full markers) in the p_{T} range between 3 and $4\,\mathrm{GeV}/c$ and between 8 and $12\,\mathrm{GeV}/c$. Background and signal candidates are distributed at low and high probabilities respectively. For an ideal classifier, the signal would be peaking at 1 and background at 0 probability, which would allow for perfect separation. In general, the distribution of the test sample follows the training set distribution, which is a sign of a good model performance without overfitting or underfitting. Due to low statistics at high transverse momenta, the two samples are deviating from each other at low probabilities in Figure 4.10b but still lie within their statistic uncertainties.

For Lorentz-boosted particles at high transverse momenta (Figure 4.10b), the signal output is peaking at maximal probability. Here, the topological features of the more displaced decay topology allow for an even better separation between signal and background candidates.

The enhanced structure in the signal distribution compared to background at values larger than $60\,\%$ in Figure 4.10a was investigated by removing single classification features from the training. It was found to be originating from the separation based on the PA of the Ξ_c^+ to the PV. The output probability of the according model, not including the PA as input feature, is shown in Figure 4.11 for the range $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$. The signal structure at high output probabilities is not observed in this case.

4.2.6. Working point determination

The final trained BDT models described in the previous sections are applied to the full analysed data sample. Subsequently, the response variable is used as a selection criterion to exclude all candidates below a certain threshold value to reject as many background candidates as possible. It is desirable to extract signal candidates with a high significance. However, a tuning of the selection on the BDT output probability to maximise the signal significance introduces the risk to amplify statistical fluctuations. This possible bias can be avoided by a blind optimisation of the selection criterion on the BDT output probability, the so-called *working point (WP)*.

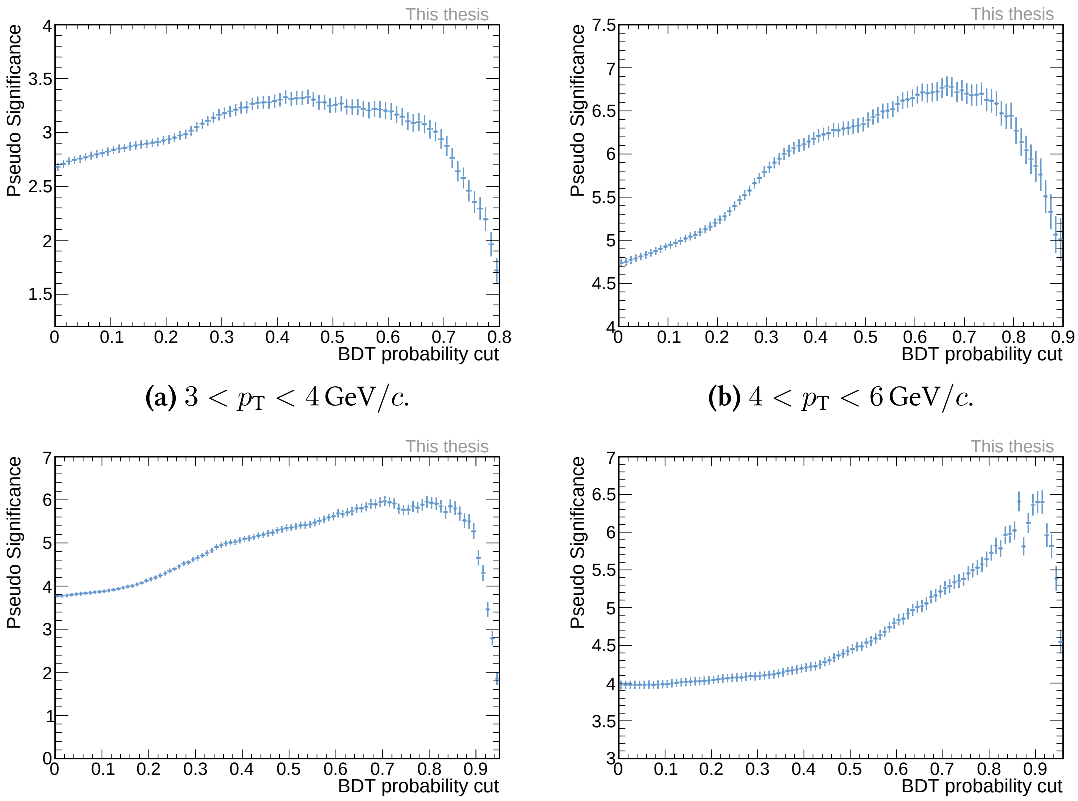
The WP is determined by calculating a so-called pseudo significance, S, which is an estimate of the significance expected in data, based on an expected number of signal, s, and background candidates, b:

$$S = \frac{s}{\sqrt{s+b}}. \tag{4.1}$$

The signal is estimated based on the measured production cross section of the isospin partner Ξ_c^0 [41]. The expected raw yield is calculated by rearranging Equation 6.1 using the Ξ_c^0 measurement as cross section prediction, correcting for the efficiency of this analysis (see section 4.4) as function of the selection on the BDT probability, and normalising by the rapidity and p_{T} bin width. A realistic number of background candidates in the signal region is estimated from part of the data sample. The signal region is defined from true MC signal as a 3σ range around the mean of the signal peak. Excluding the signal

Table 4.4.: Determined working point (WP) values for the analysed p_{T} intervals together with the chosen selection criterion on the BDT output probability.

p_{T} (GeV/c)	(3, 4)	(4, 6)	(6, 8)	(8, 12)
Working point	0.45	0.66	0.70	0.86
Selection criterion	0.67	0.66	0.70	0.85



(a) $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$.

(b) $4 < p_\mathrm{T} < 6\,\mathrm{GeV}/c$.

(c) $6 < p_\mathrm{T} < 8\,\mathrm{GeV}/c$.

(d) $8 < p_\mathrm{T} < 12\,\mathrm{GeV}/c$.

Figure 4.12.: Pseudo significance as function of the selection criterion on the BDT output probability.

region, the invariant mass spectrum in data is fitted with a second-order polynomial, which is extrapolated to the signal region. The integral below the fit function in the range of the signal, scaled up to the full analysed data sample, is taken as an estimate for the number of background candidates. This procedure is repeated as function of the selection on the BDT probability.

The resulting pseudo significance as function of the selection on the BDT output probability is shown in Figure 4.12 for each analysed p_T interval. The WP is defined as the selection that maximises the pseudo significance. The according values for all p_T intervals are presented in Table 4.4.

At low transverse momenta (Figure 4.12a) the pseudo significance exhibits a plateau region for intermediate values, suggesting a stable significance over a wide range of selections. Due to the high background at low transverse momenta, it is necessary to reduce it significantly to be able to extract signal (compare Figure 4.10a). Therefore, a selection criterion on the right edge of the pseudo significance plateau was chosen for the analysis

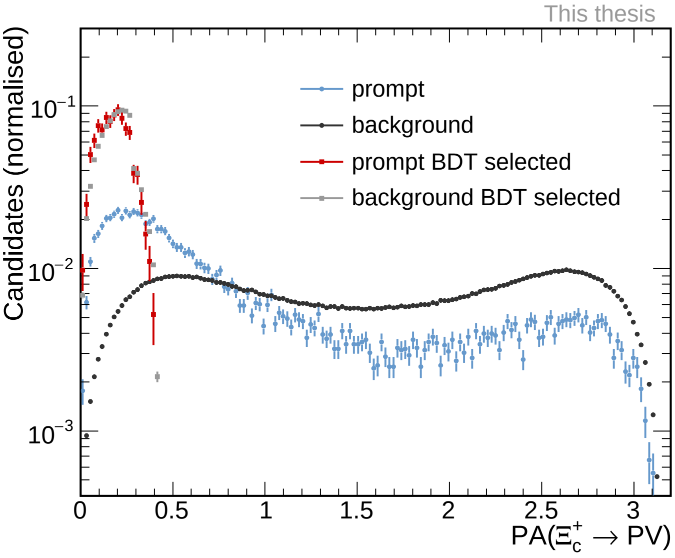


Figure 4.13.: PA of Ξ_c^+ to the PV for preselected signal (blue) and background candidates (black) in the range $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$, compared to signal (red) and background candidates (grey) with the additional selection criterion BDT probability $>$ 0.67, normalised to the number of candidates.

(see Table 4.4). For intermediate momenta (Figure 4.12b and 4.12c), the plateau region is less pronounced. The pseudo significance rather shows a peaking structure towards higher values, where the model is able to reject a lot of background while preserving enough signal to be extracted. The selection criteria in this p_T range are therefore chosen according to the determined WP, as indicated in Table 4.4. Figure 4.12d shows that the pseudo significance is peaking sharply at high values for large transverse momenta. In comparison with Figure 4.10b, it becomes clear that the model can select signal with high efficiency even for tight selections. To avoid the statistical fluctuation at the determined WP, which is visible in Figure 4.12d, a slightly deviating value was chosen as the selection criterion in this p_T interval.

The model performance is verified by the feature distributions for reconstructed signal and background before and after the applied BDT selection. Figure 4.13 shows the PA($\Xi_c^+ \to$ PV) distribution in the range $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$ for candidates before and after the BDT selection indicated in Table 4.4. The uniform background structure is strongly suppressed by the BDT selection and only candidates with small values are selected. As a result, also the signal tail at higher values is rejected but due to the low number of candidates in this region, enough signal efficiency is preserved. Similar observations have been made for the remaining classification features after the applied

BDT selection.

4.3. Raw yield extraction

The number of signal candidates is extracted from maximum likelihood fits to the invariant mass spectrum of all candidates after applying the BDT selection indicated in Table 4.4. The results for the different p_{T} intervals are reported in Figure 4.14 respectively. The blue line describes the total fit function, which contains a Gaussian fit to the signal peak and an exponential shape for the background spectrum, which is represented by the red line. The signal region is defined as 3σ interval around the mean of the signal peak ($\mu_{\mathrm{MC}} - 3\sigma_{\mathrm{MC}},\ \mu_{\mathrm{MC}} + 3\sigma_{\mathrm{MC}}$), where mean μ_{MC} and width σ_{MC} are both extracted from a Gaussian fit to true signal candidates from MC. The fitting procedure starts with an estimation of the background by a fit to the mass sidebands excluding a $4\sigma_{\mathrm{MC}}$ range around the peak mean. The signal region is modelled with a Gaussian function subsequently, with the width fixed to σ_{MC} and the initial mean taken as μ_{MC}. After a simultaneous fit of the signal and the background over the full range, the integrals below the background function as well as under the total fit are calculated and the signal counts are extracted as the difference between both.

The mean, μ, and width, σ of the signal peak together with the values of the raw yield, s, the number of background, b, the signal-to-background ratio, s/b, and the signal significance, S, within 3σ are reported. A significance smaller than 3 is interpreted to not significantly describe a signal peak structure, whereas a higher significance suggests a signal peak structure on top of the background spectrum. Throughout all analysed p_{T} intervals a significant signal peak was found in the candidate spectrum. For low transverse momenta, where the combinatorial background is high, the signal-to-background ratio is small and the signal can only be extracted with low significance. The largest signal-to-background ratio is observed at high momenta and the best signal significance is achieved at intermediate p_{T}. The p_{T}-differential background subtracted residuals of the invariant mass fits are presented in Figure 4.15.

The extracted values are compared to the values obtained in the previous analysis with standard reconstruction and selection techniques [41], and presented in Figure 4.16. In this analysis, the raw yield (left) is systematically lower compared to the published results. Though, only in the transverse momentum range between 6 and $8\,\mathrm{GeV}/c$ the values do not lie within their statistic uncertainties. The significance, on the other hand, is higher throughout all p_{T} intervals, not overlapping with the previous values within their

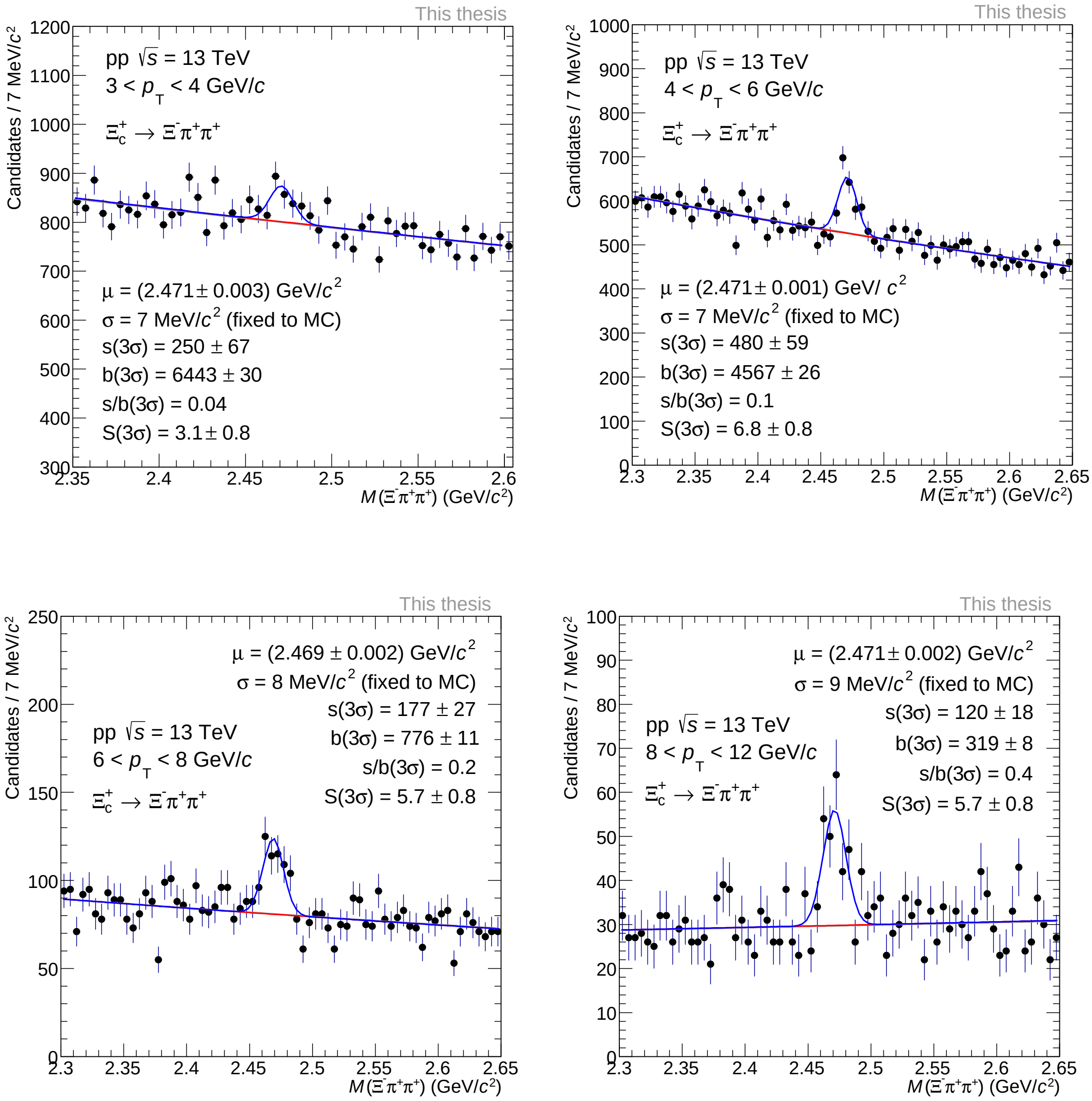


Figure 4.14.: Invariant mass spectrum of Ξ_c^+ candidates and charge conjugates in $3 < p_{\rm T} < 4\,{\rm GeV}/c$, $4 < p_{\rm T} < 6\,{\rm GeV}/c$, $6 < p_{\rm T} < 8\,{\rm GeV}/c$, and $8 < p_{\rm T} < 12\,{\rm GeV}/c$ respectively. The background is fitted with an exponential function (red line) and the signal peak with a Gaussian. The values of the mean (μ) and the width (σ) of the peak are indicated on the plot. The total fit function is shown by the blue line. The number of extracted signal (s) and background candidates (b) are reported together with the signal-to-background ratio (s/b) in the signal region ($\mu \pm 3\sigma$) and the signal significance (S).

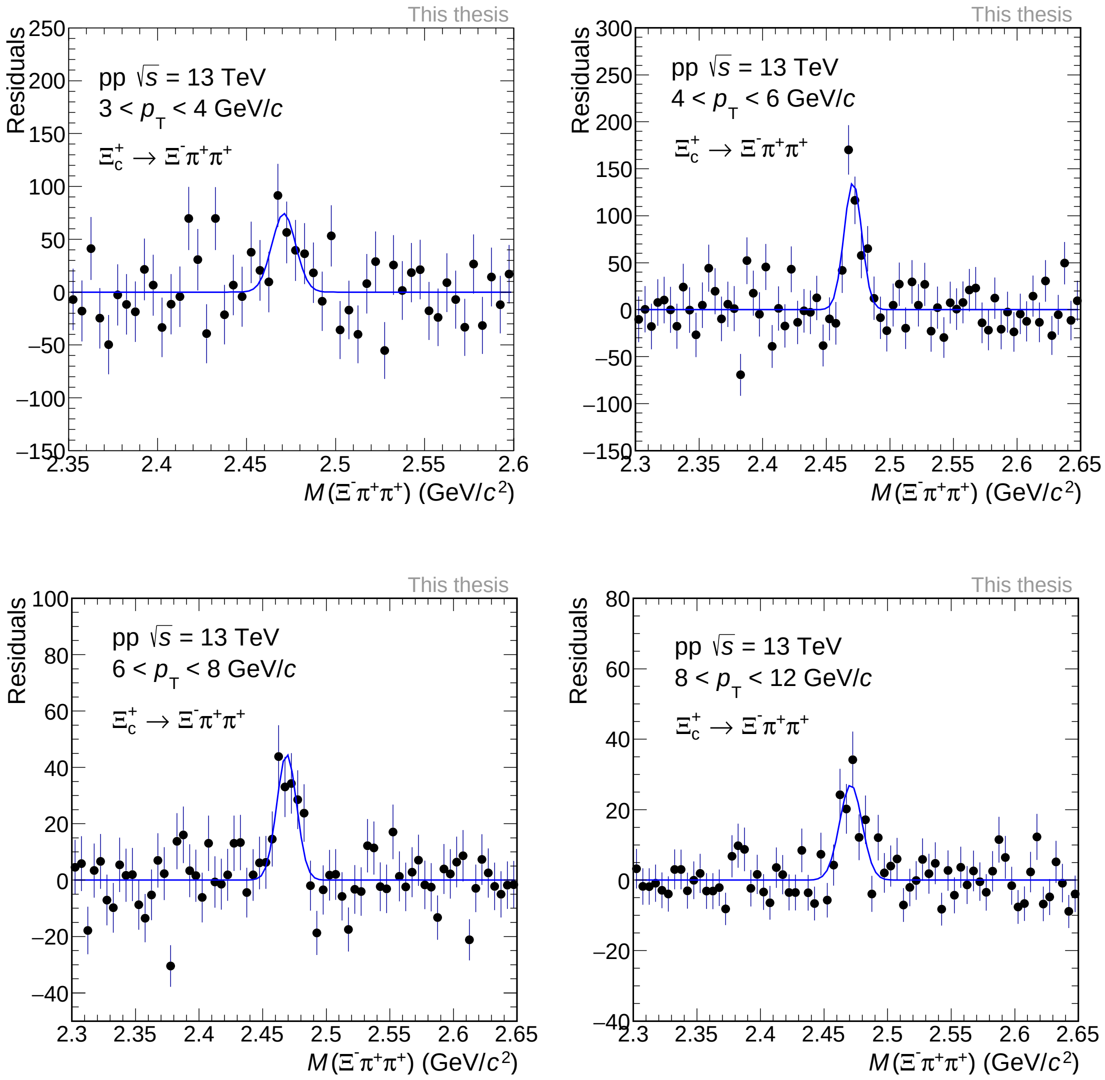


Figure 4.15.: Background subtracted residuals of the invariant mass fit of Ξ_c^+ candidates and charge conjugates in $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$, $4 < p_{\mathrm{T}} < 6\,\mathrm{GeV}/c$, $6 < p_{\mathrm{T}} < 8\,\mathrm{GeV}/c$, and $8 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$ respectively.

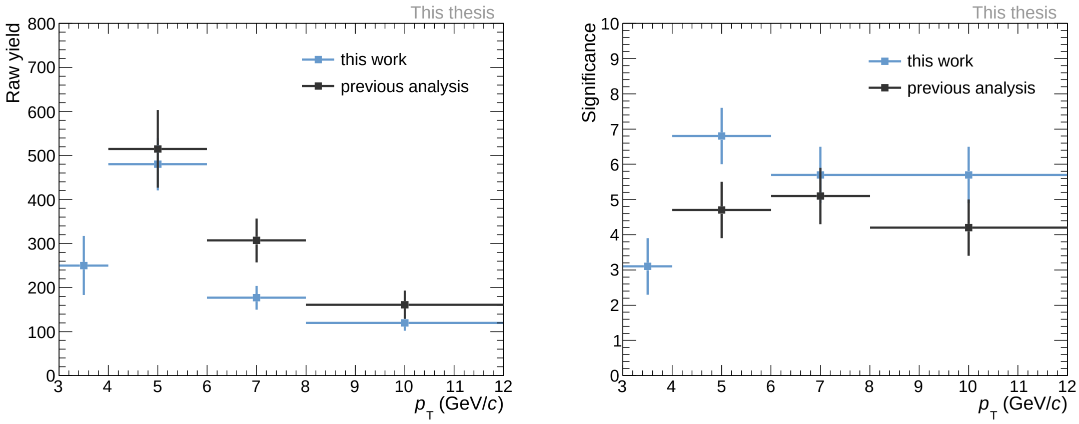


Figure 4.16.: Raw yield (left) and signal significance (right) values as function of p_{T} for this work (blue) compared to results of the previously conducted standard analysis [41] (black) in the same decay channel and collision system.

uncertainties in the p_{T} range between 4 and $6\,\mathrm{GeV}/c$. The increased significance suggests a higher background rejection in this work compared to the published one, which is verified by Figure 4.17. The number of background candidates (left) and the signal-to-background ratio (right) are presented respectively. A lower number of background and an increased signal-to-background ratio are observed throughout all p_{T} intervals compared to the standard analysis. The reported comparison verifies the fact that it is possible to achieve a better signal-background separation with a multivariate analysis approach.

4.4. Efficiency correction

The extracted raw yield needs to be corrected for the limited detector acceptance, as well as the reconstruction and selection efficiency of the analysis in order to compute the production cross section. The total acceptance and signal efficiency is defined as the ratio between the number of reconstructed signal candidates after the BDT selection and the number of generated prompt Ξ_c^+ in MC. It can be split into a reconstruction, preselection, and BDT efficiency. An independent MC sample is used for the efficiency calculation in order to avoid correlations between the BDT training and the BDT selection efficiency determination. Due to the slightly different decay topology of the resonant decay compared to the direct decay of a Ξ_c^+ to a Ξ^- and two π^+, the efficiency is determined separately for the two candidate types. The total acceptance and efficiency

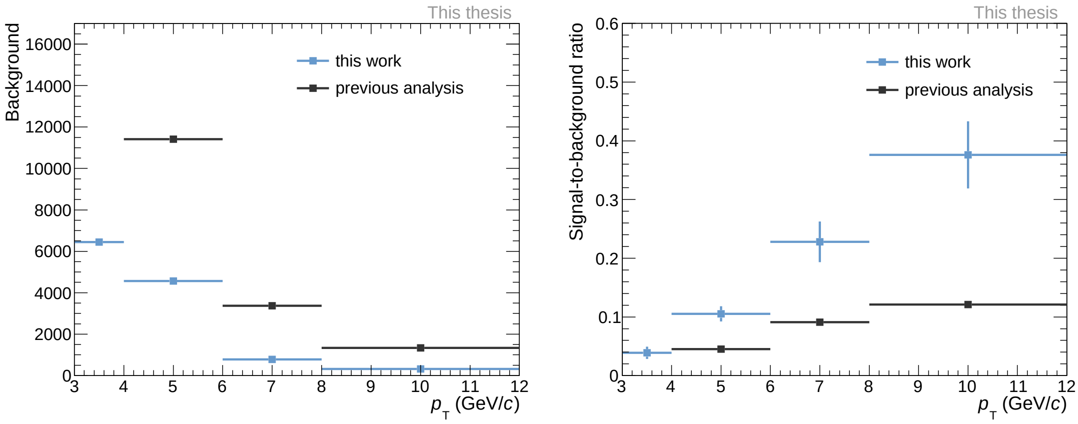
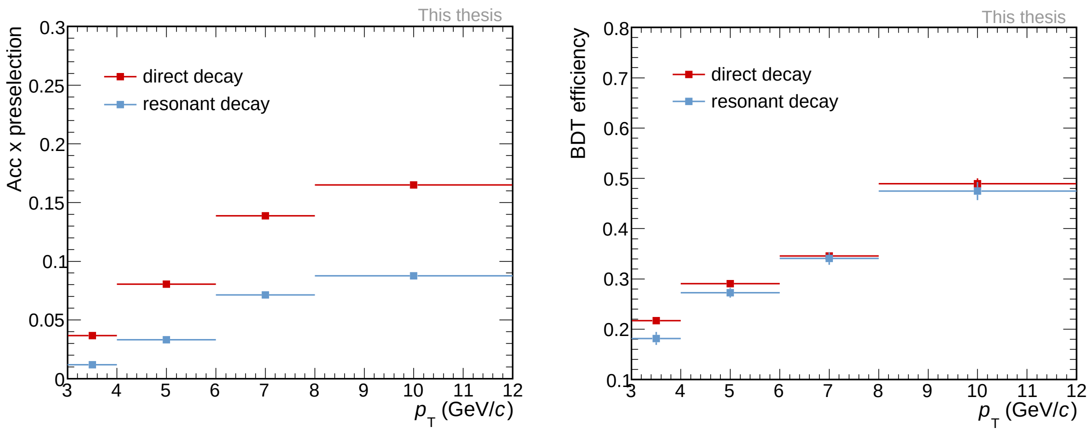


Figure 4.17.: Number of background candidates (left) and signal-to-background-ratio (right) values as function of p_T for this work (blue) compared to results of the previously conducted standard analysis [41] (black) in the same decay channel and collision system.

is then defined as the weighted average of the total efficiency of resonantly and directly decaying particles, following the ratio $1 : 10$.

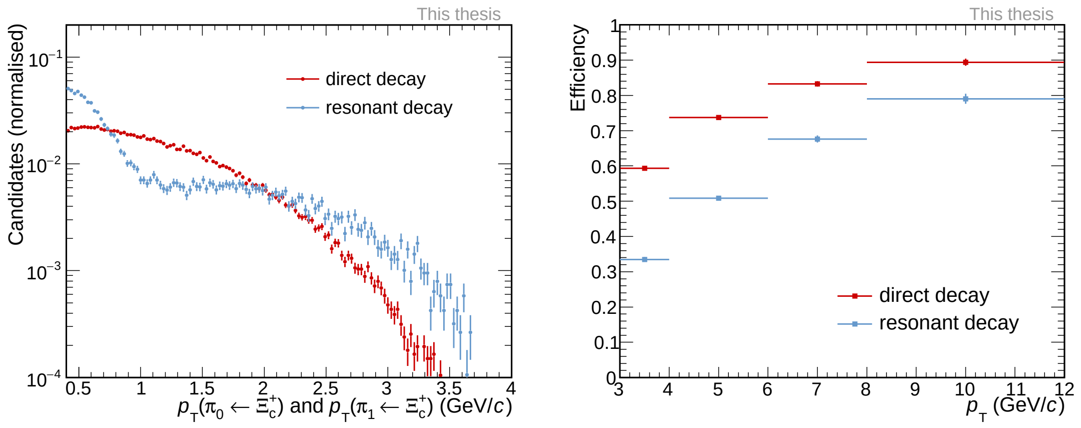
The acceptance and preselection efficiency is calculated as the number of signal candidates after the track selection and the preselection is applied, over the number of generated prompt particles. It includes the detector acceptance and the reconstruction efficiency. The values as function of p_T are presented in Figure 4.18a for directly and resonantly decaying candidates respectively. The preselection efficiency for candidates decaying via a resonance is systematically lower compared to the direct decay over the whole p_T range. The discrepancy is understood by a study of the p_T spectrum of the two π^+ coming from the Ξ_c^+. The MC p_T spectrum of the two π^+ is shown in Figure 4.19a for directly and resonantly decaying true signal candidates respectively. It becomes clear that the distribution of the candidates from a resonant decay contains a hard pion spectrum belonging to the primary pion, and a softer part for the pion coming from the resonance. The preselection criterion on the pion p_T to be larger than $0.4\,\mathrm{GeV}/c$ therefore results in a higher rejection of resonantly decaying candidates, due to the softer pion p_T spectrum, which reduces the preselection efficiency for the resonant decay. This is verified by Figure 4.19b, which shows the efficiency of the pion p_T selection criterion for directly and resonantly decaying signal candidates respectively. In this case, the efficiency is defined as the number of preselected signal candidates without a criterion on the p_T of the two pions, over the number of candidates with such a selection applied.



(a) Preselection efficiency.

(b) BDT selection efficiency.

Figure 4.18.: Efficiency for directly decaying (red) and resonantly decaying (blue) reconstructed signal candidates as function of p_{T}.



(a)

(b)

Figure 4.19.: (a) p_{T} spectrum of the two π^+ in the range $4 < p_{\mathrm{T}} < 6\,\mathrm{GeV}/c$ normalised to the number of candidates. (b) Signal efficiency of the preselection criterion $p_{\mathrm{T}}(\pi_0, \pi_1) < 0.4\,\mathrm{GeV}/c$ as function of p_{T} for directly decaying candidates (red) and candidates decaying via a resonance (blue).

The BDT selection efficiency as function of p_{T} is shown in Figure 4.18b. It is defined as the number of reconstructed signal candidates after the BDT selection is applied, over the number before the selection. The values are comparable for the direct and resonant decay since the BDT model is trained with both particle types and can classify them with similar efficiency.

The various efficiencies are combined into a total efficiency, which is calculated as the ratio between the number of selected signal candidates after the BDT selection and the number of generated prompt particles. The total efficiency therefore includes the detector acceptance as well as the reconstruction efficiency. The resonant and direct decay are treated separately and are combined subsequently into a weighted average, which is taken as the overall total efficiency to compute the production cross section. The result is presented in Figure 4.20. As expected, the total efficiency of resonantly decaying candidates is lower compared to the direct decay due to the difference in the preselection efficiency. Comparing the result with the values of the previous analysis, the weighted average of this analysis is found to be lower, which is understood by the fact that more candidates are rejected by the applied selections, as presented in Figure 4.16 and 4.17. At low momenta, however, the multivariate analysis results in a strongly improved signal efficiency by a factor 2.5, which allows for the signal extraction with relatively high significance even in this low momentum range.

4.5. Feed-down subtraction

The extracted raw yield, $N_{\mathrm{raw}}^{\Xi_c^+ + \Xi_c^-}$, includes contributions from beauty hadron decays (mostly Ξ_b^0 and $\mathrm{B_s}$ [30]), which are subtracted in order to obtain the p_{T}-differential production cross section of prompt Ξ_c^+ baryons. Therefore, the measured raw yield, $N_{\mathrm{raw}}^{\Xi_c^+ + \Xi_c^-}$, is corrected by the raw yield fraction of prompt Ξ_c^+

$$f_{\mathrm{prompt}} = 1 - \frac{N_{\mathrm{feed-down}}^{\Xi_c^+ + \Xi_c^-}}{N_{\mathrm{raw}}^{\Xi_c^+ + \Xi_c^-}}. \tag{4.2}$$

The yield of feed-down $\Xi_c^{\pm}$, $N_{\mathrm{feed-down}}^{\Xi_c^+ + \Xi_c^-}$, is estimated from the cross section of Λ_c^+ baryons originating from Λ_b^0 decays, $(\mathrm{d}^2\sigma/\mathrm{d}p_{\mathrm{T}}\mathrm{d}y)_{\mathrm{FD,\,FONLL}}^{\Lambda_c^+}$, using the beauty-quark production cross section from fixed-order next-to-leading-log (FONLL) calculations [78, 79]. The fraction of beauty quarks fragmenting into beauty hadrons is taken from the LHCb measurement of beauty fragmentation fractions in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ [80], and

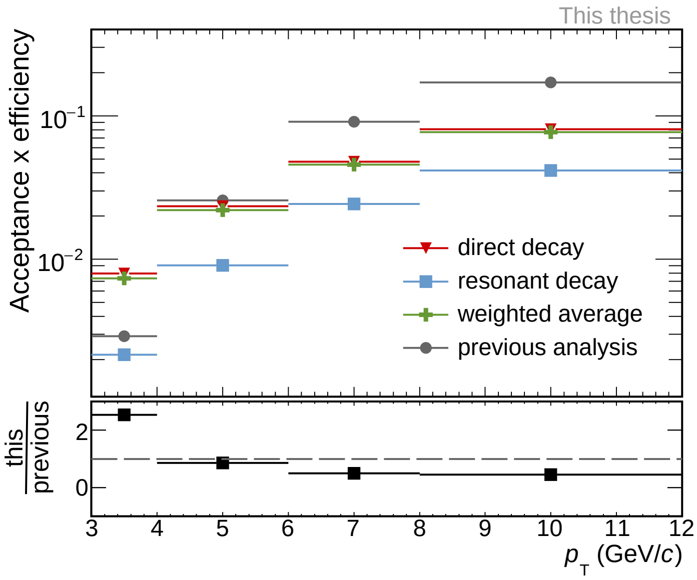


Figure 4.20.: Total efficiency for directly decaying (red) and resonantly decaying (blue) reconstructed prompt signal candidates as function of p_T, together with the overall total efficiency of this work (green) and the results of the previous analysis (grey) [41]. The ratio of the two latter distributions is presented beneath the plots.

the decay kinematics of beauty hadrons decaying into a final state with a Λ_c^+ is taken from PYTHIA 8.

The cross section of Λ_c^+ from Λ_b^0 decays is scaled by the fraction of Ξ_b decaying in a final state with a Ξ_c^+ over the fraction of Λ_b^0 decaying in a final state with a Λ_c^+ baryon, which are predicted by PYTHIA 8 [30] to be 50.5 % and 82 % respectively. The assumptions are made that the p_T shapes of the feed-down Λ_c^+ and prompt Ξ_c^+ cross sections are similar, that the cross section ratio Ξ_c^0/Λ_c^+ is similar for inclusive and feed-down particles, and that the isospin partners Ξ_c^0 and Ξ_c^+ have same yields. Under these assumptions, the predicted feed-down Λ_c^+ cross section, $(\mathrm{d}^2\sigma/\mathrm{d}p_T\mathrm{d}y)^{\Lambda_c^+}_{\mathrm{FD,\,FONLL}}$, is scaled by the ratio of the measured p_T-differential cross sections of prompt Ξ_c^0 [41], $(\mathrm{d}^2\sigma/\mathrm{d}p_T\mathrm{d}y)^{\Xi_c^0}_{\mathrm{prompt.}}$, and prompt Λ_c^+ [81], $(\mathrm{d}^2\sigma/\mathrm{d}p_T\mathrm{d}y)^{\Lambda_c^+}_{\mathrm{prompt}}$, to obtain an estimation for the p_T-differential feed-down Ξ_c^+ cross section

$$\left(\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_T\mathrm{d}y}\right)^{\Xi_c^+}_{\mathrm{FD}} = \frac{(\mathrm{d}^2\sigma/\mathrm{d}p_T\mathrm{d}y)^{\Xi_c^0}_{\mathrm{prompt}}}{(\mathrm{d}^2\sigma/\mathrm{d}p_T\mathrm{d}y)^{\Lambda_c^+}_{\mathrm{prompt}}} \cdot \left(\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_T\mathrm{d}y}\right)^{\Lambda_c^+}_{\mathrm{FD,\,FONLL}} . \tag{4.3}$$

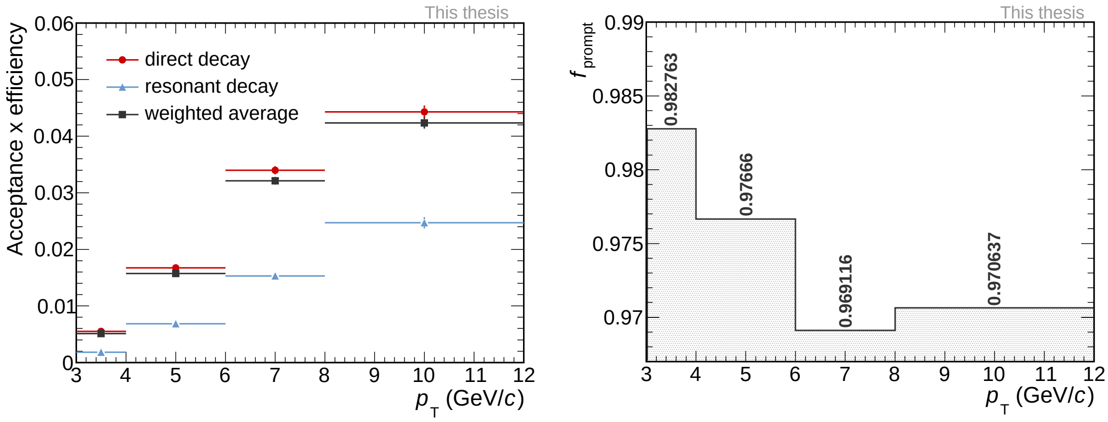


Figure 4.21.: Left: Efficiency for directly decaying (red) and resonantly decaying (blue) reconstructed feed-down candidates as function of p_{T}, together with the total feed-down efficiency (black). Right: Fraction of prompt Ξ_c^+ in this analysis as function of p_{T}.

From this estimation, the feed-down $\Xi_c^{\pm}$ yield, $N_{\mathrm{feed-down}}^{\Xi_c^+ + \Xi_c^-}$, is computed with different corrections

$$N_{\mathrm{feed-down}}^{\Xi_c^+ + \Xi_c^-}(p_{\mathrm{T}}) = 2 \cdot \left(\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_{\mathrm{T}}\mathrm{d}y} \right)_{\mathrm{FD}}^{\Xi_c^+} (p_{\mathrm{T}}) \cdot \Delta p_{\mathrm{T}} \cdot \Delta y \cdot \mathcal{L}_{\mathrm{int}} \cdot \mathrm{BR} \cdot (\mathrm{Acc} \times \varepsilon)_{\mathrm{feed-down}}(p_{\mathrm{T}}). \quad (4.4)$$

Where the factors 2, Δp_{T}, and Δy account for the contribution from antiparticles, and the p_{T} and rapidity bin width. The branching ratio, BR, of $\Xi_c^+ \to \Xi^- \pi^- \pi^-$ and the integrated luminosity, $\mathcal{L}_{\mathrm{int}}$, also enter as correction factors. Finally, the detector acceptance and reconstruction efficiency of feed-down Ξ_c^+, $(\mathrm{Acc} \times \varepsilon)_{\mathrm{feed-down}}(p_{\mathrm{T}})$, needs to be taken into account. It is computed in the same way as the reconstruction efficiency of prompt Ξ_c^+ reported in section 4.4. The observed efficiency values for feed-down Ξ_c^+ are presented in the left panel of Figure 4.21.

Finally, the estimated feed-down $\Xi_c^{\pm}$ yield is used to calculate f_{prompt} according to Equation 4.2. The result is reported in the right panel of Figure 4.21 and the values range between 97 % and 98 %.

5. Systematic uncertainties

Different steps in the analysis procedure introduce systematic uncertainties, which have to be included in the final production cross section. The main sources of systematic uncertainties in this analysis are the BDT probability selection, the raw yield extraction procedure, the track quality selection, the limited ITS-TPC matching efficiency, the deviating MC p_{T} shape, and the estimation of the fraction of prompt Ξ_c^+ candidates. A summary of the estimated values is reported in Table 5.1. The various contributions of systematic uncertainties are assumed to be uncorrelated and are therefore summed in quadrature to be combined into a total systematic uncertainty on the cross section measurement. The global uncertainties due to the BR and the luminosity are not included in the total systematic uncertainty but are stated separately.

Table 5.1.: Applied relative systematic uncertainties for the analysed p_{T} intervals.

p_{T} (GeV/c)	(3, 4)	(4, 6)	(6, 8)	(8, 12)
BDT selection (%)	3	2	4	4
Yield extraction (%)	9	8	6	7
Track quality selection(%)	5	5	5	5
ITS-TPC matching efficiency (%)	3	3	4	4
MC p_{T} shape (%)	1	2	-	-
f_{prompt} (%)	$+2$ -2	$+2$ -3	$+3$ -4	$+3$ -3
Branching ratio (%)	44	44	44	44
Luminosity (%)	1.6	1.6	1.6	1.6

5.1. BDT probability selection

Potential differences between the training features in MC compared to real data might introduce a systematic uncertainty due to the BDT probability selection, which corresponds to a set of selections on the topological input features. The discrepancies can lead to a biased BDT selection efficiency, which affects the final efficiency correction. The size of this contribution is estimated by a variation of the BDT probability selection criterion within a certain range around the central value efficiency.

The raw yield is extracted for a number of BDT probability selections, corresponding

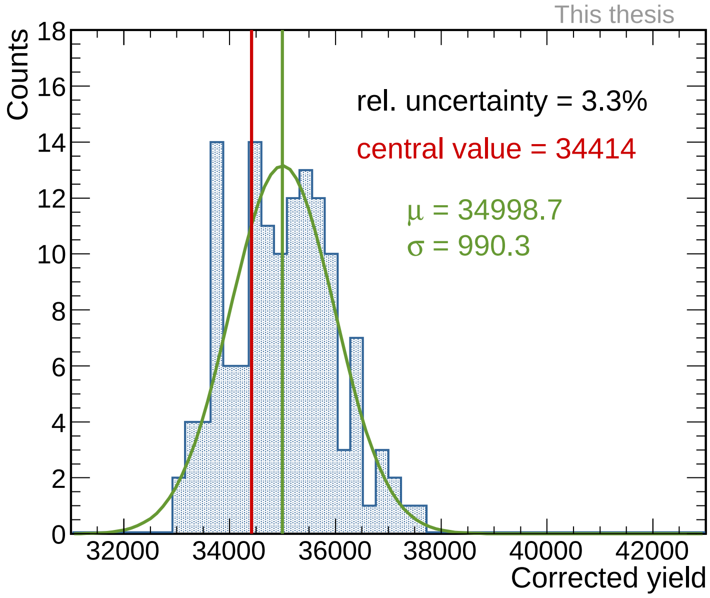
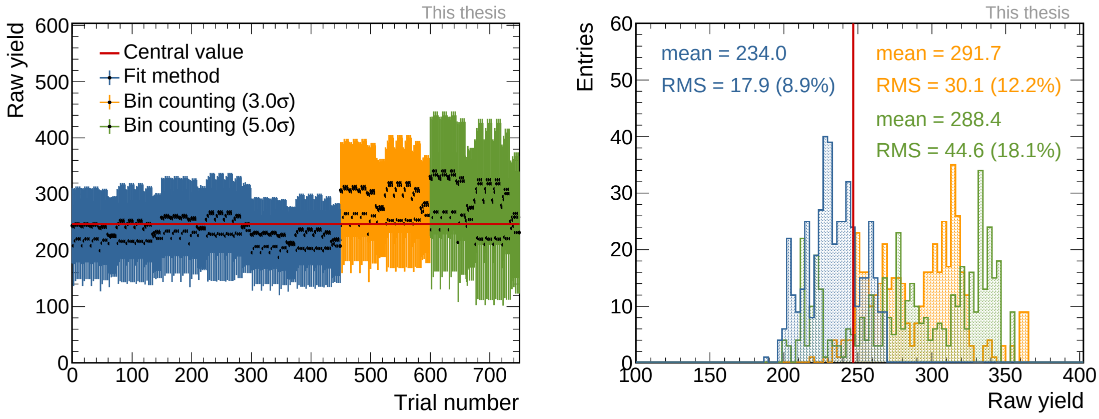


Figure 5.1.: Corrected yield distribution (blue) for a BDT selection range corresponding to ±30% around the central value efficiency in the interval $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$. The central value is reported by the red line, and the Gaussian fit with the peak mean is indicated by the green line.

to an efficiency variation of $\pm\,30\,\%$ ($\pm\,40\,\%$ for $8 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$) around the central value. In extreme cases, the signal significance might be lower than the value of 3. These cases are not included in the estimation of the systematic uncertainty. The extracted yield is corrected by the total efficiency, which is computed separately taking into account the varying BDT efficiency. The corresponding efficiency corrected yield distribution is shown in Figure 5.1 as an example for the p_{T} range between 3 and $4\,\mathrm{GeV}/c$. It is fitted with a Gaussian function (green line) to extract the peak mean, μ, and width, σ, of the distribution (reported in green). The systematic uncertainty due to the BDT selection is obtained by adding in quadrature the peak width and the distance between the Gaussian mean (reported in green) and the central value (reported in red). In this p_{T} range, a relative systematic uncertainty of $3.3\,\%$ was estimated. Similar values were observed for the remaining p_{T} intervals (see Table 5.1).

5.2. Yield extraction

The previously described fitting procedure for the extraction of the raw yield from the candidate invariant mass spectrum (section 4.3) represents another source of systematic



(a) Raw yield as function of the trial number.

(b) Raw yield distribution.

Figure 5.2.: Multitrial approach of the raw yield extraction with the fit method (blue), and the bin counting method in a 3σ range (orange), and a 5σ range (green) around the signal peak, in the interval $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$. The central value is indicated by the red line.

uncertainties. The choice of the initial fit parameters, the signal and background function shape, as well as the mass histogram binning and fit range, might systematically influence the yield extraction. To estimate the size of this uncertainty, the invariant mass fit is repeated several times in each p_{T} interval, varying the settings just mentioned.

Across the trials, the background fit function type is changed between an exponential shape and a first- and second-order polynomial. The width of the Gaussian signal fit is fixed in all cases for low momenta, either to the central value, which is taken from the MC signal fit, or increased (decreased) by $10\,\%$ compared to this value. For transverse momenta larger than $4\,\mathrm{GeV}/c$, the sigma of the signal fit is additionally let free for one-fourth of the trials. The upper and lower limits of the fit range are varied between five different values within the intervals $(2.21,\ 2.39)$ and $(2.56,\ 2.64)$ respectively. Finally, different binnings of the invariant mass histogram are considered.

For each trial, the raw yield is extracted from the fit. The result as function of the trial number is reported in Figure 5.2a (blue markers) for all fits with a χ^2_{red} between 0.5 and 2.5. The values are distributed around the central value, which is indicated by the red line. From the distribution of all extracted raw yields, the systematic uncertainty is determined. An example of this distribution is shown in Figure 5.2b for the interval $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$ (blue distribution). The systematic uncertainty is taken as the root-mean-square (RMS) of the distribution, summed in quadrature with the distance between the mean of the histogram (reported in blue) and the central value (red line).

Table 5.2.: Variations of TPC track quality selections according to [82].

Track parameter	1	2	3	4	5	6
Crossed rows	> 65	> 70	> 70	> 75	> 80	> 85
Crossed rows / findable	> 0.75	> 0.75	> 0.80	> 0.85	> 0.90	> 0.90
dE/dx clusters	> 40	> 45	> 50	> 55	> 60	> 60

In this case, a relative systematic uncertainty of $8.9\,\%$ was estimated. Both, the mean and the RMS of the histogram are reported on the plot in blue. For all p_{T} intervals, the central value was found to lie well within the distribution.

As an additional check, the results from the described fit method are compared to a different signal extraction technique, called bin-counting. Instead of a fit to the whole invariant mass spectrum, only the background is estimated by a fitting procedure, varying the function shape, fit range and histogram binning as above. Thereby two cases of fit functions for the background estimation are considered: firstly, the initial fit to the mass sidebands excluding the signal region, and secondly the background component taken from the final signal and background fit to the whole spectrum range. Subsequently, the background is subtracted from the spectrum and the bin counts in a 3σ (5σ) range around the previously defined peak mean are summed up to make up the raw yield.

The extracted results of the bin-counting in the p_{T} range between 3 and 4 GeV$/c$ for a 3σ and 5σ range are reported in Figure 5.2 by the orange and green markers respectively. The bin-counting method yields systematically higher results in both cases compared to the fit method, due to fluctuations in the signal region. This is understood by the residual distribution of the initial fit to the invariant mass spectrum shown in section 4.3, where a tendency of positive fluctuations in the signal region is observed. In the remaining p_{T} intervals, the bin-counting method is reproducing the results of the fit method well.

5.3. Track selection and ITS-TPC matching efficiency

The track quality selections on the decay daughters reported in subsection 4.1.2 introduce further systematic uncertainties, as well as the efficiency of the ITS-TPC track matching described in section 2.4, due to possible discrepancies between MC and real data. Since the tracking in this work is similar to the published analysis [41], the systematic uncertainties are inherited.

In the published analysis, the uncertainty on the track quality selection was estimated by varying the selection criteria related to the TPC tracking as shown in Table 5.2. The uncertainty associated with the tracking was estimated to be 5 % for the whole analysed p_{T} range, obtained by summing 3 % due to the tracking of the Ξ^- decay daughters, and 1 % for each π^+ track from the decay of the Ξ_c^+ [82]. The same value is adopted for the interval $3 < p_{\mathrm{T}} < 4\,\mathrm{GeV}/c$ in this analysis.

The total systematic uncertainty related to the ITS-TPC matching efficiency was estimated by the published analysis as the arithmetic sum of the systematic uncertainty on each track, which was computed centrally for each data period. The uncertainty is solely taken into account for the π^+ tracks coming from the Ξ_c^+, since the ITS is only used for these tracks. A value of 4 % (3 %) was obtained for the high (low) momentum range [82]. Since, the p_{T} dependence was found to be small, a value of 3 % is considered for the candidates with a p_{T} between 3 and $4\,\mathrm{GeV}/c$ in this analysis.

5.4. Monte Carlo p_{T}-shape

The shape of the Ξ_c^+ p_{T} spectrum in the MC sample, which is used in this analysis for the efficiency calculation, can be modified compared to real data. Hence, the potential discrepancy between generated and measured candidates imposes an additional systematic uncertainty during the efficiency calculation, which has to be taken into account. Since the used MC sample is similar to the published analysis, the values are inherited as before.

The estimation procedure for this source of uncertainty follows the description in [41]. The ratio between the generated candidate p_{T} distribution and the measured spectrum is used to re-weight each candidate. The p_{T} dependent weights are fitted three times with a varying exponential function, resulting in a central, a minimum, and a maximum value for the candidate weights extracted from each of the fit functions. The generated candidates are then weighted with the according values, and for each of the re-weighted candidate spectra, the total analysis efficiency is computed. Finally, the ratio of the efficiencies with respect to the default spectrum is taken as the systematic uncertainty. It was observed to be negligible for candidates with $p_{\mathrm{T}} > 6\,\mathrm{GeV}/c$, and to amount to 2 % (1 %) in the interval $4 < p_{\mathrm{T}} < 6\,\mathrm{GeV}/c$ (below $4\,\mathrm{GeV}/c$). The observed value of 1 % is inherited for the candidates with a p_{T} between 3 and $4\,\mathrm{GeV}/c$ in this analysis.

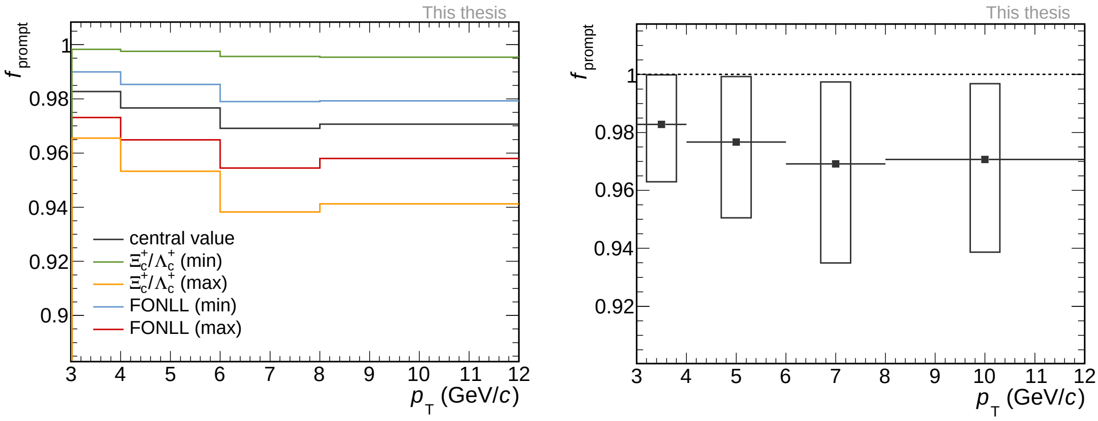


Figure 5.3.: Left: Fraction of prompt Ξ_c^+. The central value is reported in black, and the upper (lower) value limits due to the minimum (maximum) FONLL prediction and the down-scaled (up-scaled) Ξ_c^0/Λ_c^+ ratio in blue and green (red and orange) respectively. Right: Central value estimation of the fraction of prompt Ξ_c^+ with the assigned systematic uncertainties presented as boxes.

5.5. Estimation of prompt fraction

For the estimation of f_{prompt}, the uncertainty on the feed-down Λ_c^+ cross section FONLL prediction [78, 79] and various assumptions on the cross section ratio Ξ_c^0/Λ_c^+ are considered. The non-prompt Λ_c^+ uncertainty is estimated by varying the beauty quark mass, as well as the factorisation and renormalisation scales in the FONLL calculation, as described in [79]. The two contributions are summed in quadrature together with the uncertainty on the parton distribution functions to obtain the overall upper and lower systematic uncertainty on the FONLL prediction.

For the estimation of f_{prompt}, the non-prompt Λ_c^+ cross section is scaled by the ratio of prompt Ξ_c^0 over prompt Λ_c^+ cross sections, relying on the assumption that the ratio Ξ_c^0/Λ_c^+ is the same for prompt and feed-down baryons. In order to account for possible differences between the Ξ_c^0/Λ_c^+ and Ξ_b^-/Λ_b^0 ratios, the cross section ratio is scaled up by a conservative factor of 2. The lower uncertainty on the ratio is obtained by scaling it down by a factor of 0.05 to capture the Ξ_b^-/Λ_b^0 value measured at forward rapidity by the LHCb Collaboration [83].

Using the defined upper and lower limits on the FONLL prediction and the Ξ_c^0/Λ_c^+ cross section ratio, varying values for f_{prompt} are computed according to section 4.5 and the result is presented in the left panel of Figure 5.3. The central f_{prompt} value (section 4.5) is reported in black, while the lower (upper) limits due to the varied FONLL calculation and the different assumptions on Ξ_c^0/Λ_c^+ are shown in red and orange (blue and green) respectively. The final upper and lower bounds on the value of f_{prompt} are each taken as the quadratic sum of the two contributions. They are indicated as boxes in Figure 5.3 together with the central value, and they range between $-2.1\,\%$ and $-3.6\,\%$, and $+1.8\,\%$ and $+3.9\,\%$.

5.6. Branching ratio and luminosity

The branching ratio of the analysed decay and the integrated luminosity carry uncertainties, which need to be taken into account since they both enter the cross section computation.

The integrated luminosity is determined in van der Meer scans [73] ($\mathcal{L}_{\text{int}} = 32.08\,\text{nb}^{-1}$), and the relative uncertainty is reported to be $1.6\,\%$ for pp data recorded between 2016 and 2018 [73]. The branching fraction of the analysed decay is calculated from the individual decay components. The branching fraction of $\Xi_c^+ \to \Xi^-\pi^+\pi^+$ is taken from a recent Belle measurement of the absolute branching fractions of the Ξ_c^+ baryon [84], it is reported to be $\mathcal{B}(\Xi_c^+ \to \Xi^-\pi^+\pi^+) = (2.86 \pm 1.21^{\text{stat.}} \pm 0.38^{\text{syst.}})\,\%$. The fractions of $\Xi^- \to \Lambda\pi^-$ $(99.887 \pm 0.035)\,\%$ and $\Lambda \to \text{p}\pi^-$ $(63.9 \pm 0.5)\,\%$ are taken from the Particle Data Group [3]. The branching fraction of the analysed decay chain, therefore, results in $1.8\,\%$ with a relative uncertainty of $44.4\,\%$.

6. Results and discussion

6.1. p_T-differential cross section

In this analysis, the production cross section of prompt Ξ_c^+ baryons measured in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ with the ALICE detector is presented. The measurement is performed at midrapidity ($|y| < 0.5$) in the transverse momentum range $3 < p_\mathrm{T} < 12\,\mathrm{GeV}/c$. The Ξ_c^+ baryon is reconstructed via the hadronic decay channel $\Xi_c^+ \to \pi^+\pi^+$ with the Ξ^- decaying in $\Xi^- \to (\pi^-\Lambda) \to \pi^-\mathrm{p}\pi^-$, including the charge conjugate modes. The p_T-differential production cross section is obtained according to

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_\mathrm{T}\mathrm{d}y} = \frac{1}{2} \cdot \frac{f_\mathrm{prompt}(p_\mathrm{T}) \cdot N_\mathrm{raw}^{\Xi_c^+ + \Xi_c^-}(p_\mathrm{T})}{(\mathrm{Acc} \times \varepsilon)_\mathrm{prompt}(p_\mathrm{T}) \cdot \mathrm{BR} \cdot \mathcal{L}_\mathrm{int} \cdot \Delta y \cdot \Delta p_\mathrm{T}}. \tag{6.1}$$

The p_T dependent fraction of prompt Ξ_c^+, $f_\mathrm{prompt}(p_\mathrm{T})$, is calculated from an estimated number of Ξ_c^+ from beauty baryon decays, and is found to be 0.97 on average. The $\Xi_c^\pm$ raw yield, $N_\mathrm{raw}^{\Xi_c^+ + \Xi_c^-}$, in a given p_T interval with width Δp_T is extracted from fits to the candidate invariant mass spectrum in this p_T interval. To account for a limited detector acceptance (Acc) and the reconstruction and selection efficiency (ε), the selected Ξ_c^+ are corrected by the p_T dependent product of the geometrical acceptance and the efficiency of prompt candidates in this analysis, $(\mathrm{Acc} \times \varepsilon)_\mathrm{prompt}(p_\mathrm{T})$. The result is normalised by the rapidity interval $\Delta y = 1.6$ of the measurement under the assumption that the Ξ_c^+ rapidity distribution is uniform in the range $|y| < 0.8$, as well as by the width of the according momentum interval, Δp_T. The branching ratio, BR, and the integrated luminosity, $\mathcal{L}_\mathrm{int}$, have to be taken into account as normalisation factors, where the total branching ratio is computed as the product of the individual decay components to be $\mathrm{BR} = (1.83 \pm 0.08)\,\%$, and the integrated luminosity of the analysed sample was determined to be $\mathcal{L}_\mathrm{int} = (32.08 \pm 0.52)\,\mathrm{nb}^{-1}$. Finally, the normalisation factor $\frac{1}{2}$ accounts for the measured antiparticles, since the cross section is computed as the average of Ξ_c^+ and Ξ_c^- and the extracted raw yield contains both particles.

The resulting p_T-differential production cross section is depicted in Figure 6.1. The statistical and systematic uncertainties are represented by vertical error bars and boxes respectively. The shaded boxes show the global uncertainty due to the branching ratio. These conventions apply to all results presented in this chapter.

The reported result (red) is compared to published ALICE measurements of the Ξ_c cross

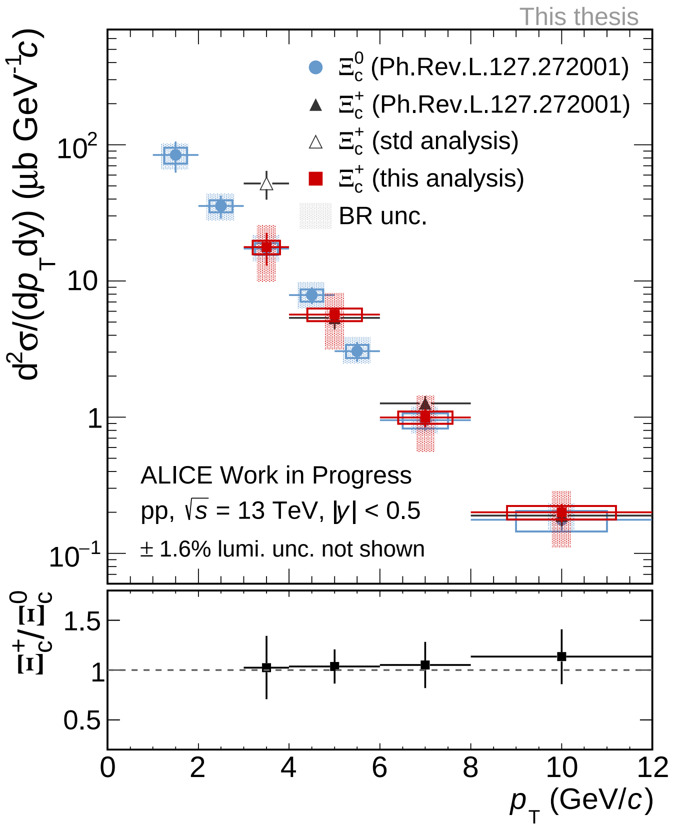


Figure 6.1.: The p_{T}-differential production cross section of prompt Ξ_c^+ and Ξ_c^0 baryons in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$. The result of this analysis is reported by the red markers, with statistical and systematic uncertainties shown as vertical lines and boxes respectively. The ratio Ξ_c^+/Ξ_c^0 of this analysis over the published measurement [41] (blue markers) is presented below the plot.

section [41]. The Ξ_c^0 production cross section (blue) was obtained from a combined measurement of $\Xi_c^0 \rightarrow \Xi^- e^+ \nu_e$ and $\Xi_c^0 \rightarrow \Xi^- \pi^+$ in the transverse momentum range $1 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$. The Ξ_c^+ measurement (black full markers) was performed in the same system and decay channel as this work in the interval $4 < p_{\mathrm{T}} < 12\,\mathrm{GeV}/c$, using standard reconstruction and analysis techniques. In addition, the unpublished analysis attempt, to extend the Ξ_c^+ measurement to lower momenta (black open marker) is shown in the figure.

Thanks to the improved reconstruction and selection in this analysis, it was possible to cover a wider p_{T} range compared to the published analysis, extending the measurement to $p_{\mathrm{T}} = 3\,\mathrm{GeV}/c$. Due to the increasing Ξ_c production at low p_{T} compared to higher momenta, this is an important step on the way to a more precise measurement of the to-

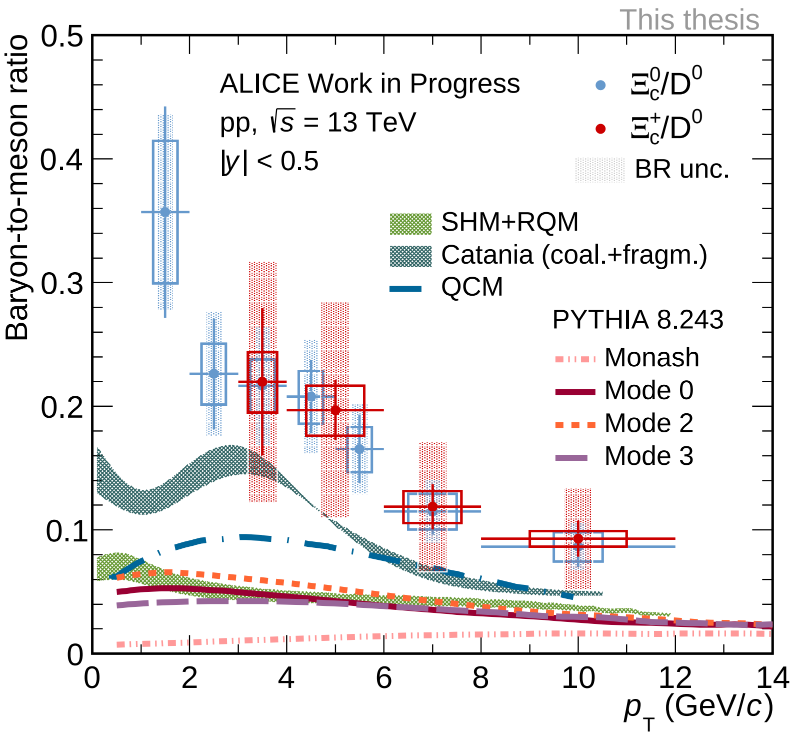


Figure 6.2.: Prompt cross section ratios Ξ_c^+/D^0 (red) and Ξ_c^0/D^0 (blue) [41] in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$, compared to model predictions [30, 32, 35, 36, 38].

tal inclusive Ξ_c^+ production cross section. With the here presented measurement, the by a factor ~ 3 higher result compared to the Ξ_c^0 baryon in the interval $3 < p_\mathrm{T} < 4\,\mathrm{GeV}/c$ obtained in the previous analysis attempt cannot be confirmed. The result rather follows the trend of the Ξ_c^0 measurement over the whole transverse momentum range.
The published measurements agree with the here presented result within their statistical and systematic uncertainties. This finding confirms the expected result suggested by isospin symmetry and a similar feeding from higher resonance states for the two baryons.

6.2. Baryon-to-meson ratio

To study the charm haronisation process, the Ξ_c^+/D^0 cross section ratio is presented, which is sensitive to the charm fragmentation function. The ratio is computed by dividing the measured production cross sections of the Ξ_c^+ baryon and the D^0 meson [81]. Figure 6.2 shows a comparison of the Ξ_c^+/D^0 result obtained in this analysis (red) and the published Ξ_c^0/D^0 ratio (blue) [41]. The systematic uncertainties related to the ITS-TPC matching efficiency, as well as the uncertainty due to the FONLL prediction during the estimation of f_{prompt} are propagated as correlated. All other sources of systematic uncertainties are assumed to be uncorrelated in the ratio. The two results are in good agreement within their statistical and systematic uncertainties and a similar p_T-dependence is observed. The measured ratio shows a decreasing trend for $p_\mathrm{T} > 3\,\mathrm{GeV}/c$, the maximum value is $\Xi_c^+/\mathrm{D}^0 \approx 0.22$, and the smallest value is about 0.1. The result suggests that the fragmentation of charm into baryons and mesons differs, and is non-universal.

The presented result is compared to model predictions depicted in Figure 6.2 by different bands and lines indicated in the figure. The PYTHIA 8.2 event generator with the Monash tune [30], implementing charm hadronisation via vacuum fragmentation with the fragmentation parameters tuned on $\mathrm{e}^+\mathrm{e}^-$ data, significantly underpredicts the measurement by a factor of 22 in the low-p_T region and by a factor of about 5 at high p_T. In addition, different tunes including CR beyond the LC approximation [32] are considered (Mode 0, 2, and 3 in the figure), where the Mode 3 tune also takes into account the formation of junction topologies, increasing the baryon production. All three models predict a nearly uniform enhancement of the baryon-to-meson ratio in the low-p_T region, but still underestimate the data by a factor of about $4 - 6$ (for $p_\mathrm{T} < 4\,\mathrm{GeV}/c$). These results provide further evidence for the modification of fragmentation functions in pp collisions compared to $\mathrm{e}^+\mathrm{e}^-$ and $\mathrm{e}^-\mathrm{p}$ collisions.

The measurement is further compared to a SHM [38], where hadron yields in heavy-ion collisions are computed via statistical weights based on the mass of the particles. The presented model takes into account an extended charm baryon spectrum including additional yet unobserved excited baryon states, which are predicted by the relativistic quark model (RQM) [39] and lQCD [85]. As already discussed in section 1.5.3, this model describes the measured Λ_c^+/D^0 ratio in pp collisions [23]. However, Figure 6.2 shows that it underestimates the presented Ξ_c^+/D^0 ratio by the same amount as the PYTHIA tunes with CR beyond LC.

Furthermore, models including hadronisation via quark (re-)combination or coalescence are considered to describe the measured baryon-to-meson ratio. The quark (re-)combination mechanism (QCM) [35] requires the presence of a parton rich environment from which the charm quark can pick up a comoving light antiquark (two comoving light quarks) to form a charmed meson (baryon). Figure 6.2 shows that this model predicts a more enhanced baryon-to-meson ratio than the ones discussed so far, but it is still not able to describe the measured Ξ_c^+/D^0 ratio, especially at mid and low p_T. Finally, the result is compared to the Catania coalescence model [36], which applies a coalescence approach together with fragmentation, originally developed for heavy-ion collisions. The model assumes the formation of a hot QCD matter at finite temperature. The model band in Figure 6.2 is closest to the measured result over the whole p_{T} interval.

Overall, all considered models are unable to describe the presented results, except the Catania model, which comes closest to the data.

7. Conclusion and outlook

The p_T-differential production cross section of the prompt charm-strange $\Xi_c^+ \to \Xi^-\pi^+\pi^+$ baryon was measured at midrapidity in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ in the transverse momentum range $3 < p_\mathrm{T} < 12\,\mathrm{GeV}/c$ with the ALICE detector. Applying a multivariate analysis technique, it was possible to extend the published Ξ_c^+ measurement in the same decay channel and system at low p_T to $p_\mathrm{T} = 3\,\mathrm{GeV}/c$. In the lowest p_T interval, the signal was extracted with an efficiency 2.5 times higher than in the previous analysis attempt.

The challenging measurement of charm baryon production at low transverse momenta, where the production rates are high and the detector acceptance decreases, is an important step to understanding charm production and hadronisation and reaching a measurement in larger systems like Pb–Pb collisions. Therefore, it was attempted to extract a measurement even at $p_\mathrm{T} < 3\,\mathrm{GeV}/c$. The analysis was conducted with the same strategy presented in the previous chapters and it was possible to train a model with comparable performance with respect to the higher p_T intervals. A detailed study of the different preselection criteria and model input variables was done to reach the best classification performance possible while preserving enough efficiency to be able to extract signal from the large combinatorial background. Figure 7.1 shows the fitted invariant mass spectrum of the selected candidates in the range $2 < p_\mathrm{T} < 3\,\mathrm{GeV}/c$. Only candidates with a BDT probability larger than 0.2 were selected. While the model output suggested a relatively good classification performance, it was not possible to extract signal with high significance from the spectrum. Most probably the signal in Figure 7.1 is an enhanced statistical fluctuation and the result is therefore not feasible.

In general, the presented Ξ_c^+ measurement is in good agreement with the published result of the isospin partner Ξ_c^0 in the full measured p_T range and it provides important constraints for various model predictions. Furthermore, a measurement of the p_T-integrated Ξ_c^+ cross section will be extracted by extrapolating the presented measurement in the range $p_\mathrm{T} < 3\mathrm{GeV}/c$ and $p_\mathrm{T} > 12\,\mathrm{GeV}/c$. Together with other charm hadron measurements it will be used to measure charm fragmentation fractions and the total $\mathrm{c\bar{c}}$ cross section in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ at midrapidity in ALICE. Up to now, the inclusive Ξ_c^+ production was considered in these measurements taking into account twice the measured Ξ_c^0 cross section, assuming isospin symmetry.

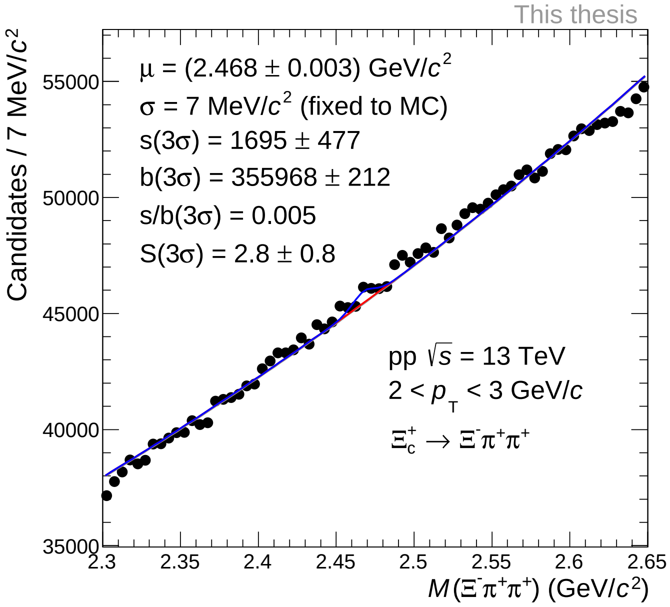


Figure 7.1.: Fit of the invariant mass spectrum of Ξ_c^+ candidates and charge conjugates in the range $2 < p_\mathrm{T} < 3\,\mathrm{GeV}/c$.

The Ξ_c^+/D^0 production cross section ratio was measured in the range $3 < p_\mathrm{T} < 12\,\mathrm{GeV}/c$, for the first time in ALICE providing a result at $p_\mathrm{T} < 4\,\mathrm{GeV}/c$. The ratio agrees with the published Ξ_c^0/D^0 ratio. An enhanced baryon-to-meson ratio is observed compared to measurements in $\mathrm{e}^+\mathrm{e}^-$ and $\mathrm{e}^-\mathrm{p}$ collisions, which is a common observation of several other measurements in the charm baryon sector in high-energy hadronic collisions. Several tunes of the MC event generator PYTHIA are not able to describe the enhanced ratio. These results provide a strong indication for a non-universal fragmentation of charm into baryons and mesons in different collision systems.

The presented result is best described by the theoretical predictions made with the Catania model, suggesting that charm quarks hadronise via coalescence at low transverse momenta even in pp collisions. This finding suggests the unexpected formation of hot QCD matter in pp collisions. The result in pp collisions provides an important reference for future measurements in Pb–Pb collisions.

The data from the upcoming Run 3 at the LHC will provide an improved tracking resolution and higher statistics due to detector upgrades and larger data-taking rates. The improved precision will make a measurement at even lower transverse momenta possible. Furthermore, the planned Ξ_c^+ measurement in Pb–Pb collisions will become feasible. The vertexing capabilities of the KFParticle package will serve as a powerful reconstruction tool in these analyses. Possibly they will be even more important for the extraction of extremely rare signals with large combinatorial background, like multi-charm baryons with long decay chains, i.e. $\Xi_{cc}^{++} \rightarrow \Xi_c^{+}\pi^{+}$ and $\Xi_{cc}^{+} \rightarrow \Xi_c^{0}\pi^{+}$.

List of Acronyms

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC ApparatuS
AUC	Area Under the ROC Curve
BDT	Boosted Decision Tree
BR	branching ratio
CERN	European Organisation for Nuclear Research
CMS	Compact Muon Solenoid
CR	colour reconnection
DCA	distance of closest approach
DCal	Di-Jet Calorimeter
EMCal	Electromagnetic Calorimeter
FF	fragmentation function
FIT	Fast-Interaction-Trigger
FONLL	fixed-order next-to-leading-log
GEM	Gas Electron Multiplier
hipe4ML	Heavy-Ion Physics Environment for Machine Learning
HMPID	High Momentum Particle Identification Detector
IP	interaction point
ITS	Inner Tracking System
LC	leading colour
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty experiment
lQCD	lattice QCD
LS2	long shutdown 2
MAPS	Monolithic Active Pixel Sensors
MB	minimum-bias
MC	Monte Carlo
MFT	Muon-Forward-Tracker
ML	Machine Learning
MPI	multiple partonic interaction
MWPC	multi-wire proportional chamber

PA	pointing angle
Pb–Pb	lead–lead
PDF	parton distribution function
PHOS	Photon Spectrometer
PID	particle identification
pp	proton–proton
pQCD	perturbative QCD
PV	primary vertex
QCD	Quantum Chromodynamics
QCM	quark (re-)combination mechanism
QED	Quantum Electrodynamics
QFT	quantum field theory
QGP	quark-gluon plasma
RHIC	Relativistic Heavy Ion Collider
RMS	root-mean-square
RMSE	root-mean-square error
ROC	Receiver Operating Characteristic
ROC	Readout Chamber
SDD	Silicon Drift Detectors
SHM	statistical hadronisation model
SPD	Silicon Pixel Detectors
SSD	Silicon Strip Detectors
TOF	Time-Of-Flight
TPC	Time Projection Chamber
TRD	Transition Radiation Detector
WP	working point

Bibliography

[1] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, *The standard model of particle physics.* Rev. Mod. Phys. **71** (2 Mar. 1999), S96–S111. DOI: `10.1103/RevModPhys.71.S96`.

[2] M. Thomson, *Modern Particle Physics.* Cambridge University Press, New York, 2013. ISBN: 978-1-107-03426-6.

[3] **Particle Data Group** Collaboration, P. A. Zyla *et al.*, *Review of Particle Physics.* PTEP **2020**.8 (2020). and 2021 update, 083C01. DOI: `10.1093/ptep/ptaa104`.

[4] Kenneth G. Wilson, *Confinement of quarks.* Phys. Rev. D **10** (8 Oct. 1974), 2445–2459. DOI: `10.1103/PhysRevD.10.2445`.

[5] W. Busza, K. Rajagopal, and W. van der Schee, *Heavy Ion Collisions: The Big Picture and the Big Questions.* Annu. Rev. Nucl. Part. S **68**.1 (2018), 339–376. DOI: `10.1146/annurev-nucl-101917-020852`.

[6] E. V. Shuryak, *Quark-gluon plasma and hadronic production of leptons, photons and psions.* Phys. Lett. B **78**.1 (1978), 150–153. ISSN: 0370-2693. DOI: `10.1016/0370-2693(78)90370-2`.

[7] Y. Aoki *et al.*, *The order of the quantum chromodynamics transition predicted by the standard model of particle physics.* Nature **443**.7112 (Oct. 2006), 675–678. DOI: `10.1038/nature05120`.

[8] P. Steinbrecher, *The QCD crossover at zero and non-zero baryon densities from Lattice QCD.* Nucl. Phys. A **982** (2019). The 27th International Conference on Ultra-relativistic Nucleus-Nucleus Collisions: Quark Matter 2018, 847–850. ISSN: 0375-9474. DOI: `10.1016/j.nuclphysa.2018.08.025`.

[9] D. Boyanovsky, H. J. de Vega, and D. J. Schwarz, *Phase Transitions in the Early and Present Universe.* Annu. Rev. Nucl. Part. S **56**.1 (2006), 441–500. DOI: `10.1146/annurev.nucl.56.080805.140539`.

[10] E. Laermann and O. Philipsen, *Lattice QCD at finite temperature.* Annu. Rev. Nucl. Part. S **53**.1 (2003), 163–198. DOI: `10.1146/annurev.nucl.53.041002.110609`.

[11] N. Cabibbo and G. Parisi, *Exponential hadronic spectrum and quark liberation.* Phys. Lett. B **59**.1 (1975), 67–69. ISSN: 0370-2693. DOI: `10.1016/0370-2693(75)90158-6`.

[12] J. C. Collins and M. J. Perry, *Superdense Matter: Neutrons or Asymptotically Free Quarks?* Phys. Rev. Lett. **34** (21 May 1975), 1353–1356. DOI: 10.1103/PhysRevLett.34.1353.

[13] A. Andronic, *An overview of the experimental study of quark-gluon matter in high-energy nucleus-nucleus collisions.* Int. J. Mod. Phys. A **29**.22 (Aug. 2014), 1430047. DOI: 10.1142/s0217751x14300476.

[14] **STAR** Collaboration, K. Lokesh, *STAR Results from the RHIC Beam Energy Scan-I.* Nucl. Phys. A **904-905** (2013). The Quark Matter 2012, 256c–263c. ISSN: 0375-9474. DOI: 10.1016/j.nuclphysa.2013.01.070.

[15] A. Kurkela *et al.*, *Effective kinetic description of event-by-event pre-equilibrium dynamics in high-energy heavy-ion collisions.* Phys. Rev. C **99** (3 Mar. 2019), 034910. DOI: 10.1103/PhysRevC.99.034910.

[16] S. Z. Belen/kji and L. D. Landau, *Hydrodynamic theory of multiple production of particles.* Nuovo Cim **3** (1956), 15–31. DOI: 10.1007/BF02745507.

[17] S. G. Weber, *Multiplicity dependent J/ψ production in proton-proton collisions at the LHC.* PhD thesis. Technische Universität Darmstadt, 2018.

[18] J. C. Collins, D. E. Soper, and G. Sterman, *Heavy particle production in high-energy hadron collisions.* Nucl. Phys. B **263**.1 (1986), 37–60. ISSN: 0550-3213. DOI: 10.1016/0550-3213(86)90026-X.

[19] **ALICE** Collaboration, S. Acharya, D. Adamová, and S. P. Adhya, *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**.388 (2019). DOI: 10.1140/epjc/s10052-019-6873-6.

[20] **ALICE** Collaboration, S. Acharya *et al.*, *Measurement of beauty and charm production in pp collisions at $\sqrt{s} = 5.02\,TeV$ via non-prompt and prompt D mesons.* JHEP **2021**.220 (2021). DOI: 10.1007/JHEP05(2021)220.

[21] **CMS** Collaboration, V. Khachatryan *et al.*, *Measurement of the B^+ Production Cross Section in pp Collisions at $\sqrt{s} = 7\,TeV$.* Phys. Rev. Lett. **106** (11 Mar. 2011), 112001. DOI: 10.1103/PhysRevLett.106.112001.

[22] **ALICE** Collaboration, S. Acharya *et al.*, *Λ_c^+ production in pp collisions at $\sqrt{s} = 7\,TeV$ and in p-Pb collisions at $\sqrt{s_{NN}} = 5.02\,TeV$.* JHEP **04** (2018), 108. DOI: 10.1007/JHEP04(2018)108.

[23] **ALICE** Collaboration, S. Acharya *et al.*, Λ_c^+ *Production and Baryon-to-Meson Ratios in pp and p-Pb Collisions at* $\sqrt{s_{\mathrm{NN}}} = 5.02\ \text{ TeV}$ *at the LHC*. Phys. Rev. Lett. **127** (20 Nov. 2021), 202301. DOI: 10.1103/PhysRevLett.127.202301.

[24] **ALICE** Collaboration, S. Acharya *et al.*, Λ_c^+ *production in* pp *and in* $p-Pb$ *collisions at* $\sqrt{s_N N} = 5.02 TeV$. Phys. Rev. C **104**.5 (Nov. 2021). DOI: 10.1103/physrevc.104.054905.

[25] **ALICE** Collaboration, S. Acharya *et al.*, *First measurement of* Ξ_c^0 *production in pp collisions at* $\sqrt{s} = 7\,TeV$. Phys. Lett. B **781** (2018), 8–19. DOI: 10.1016/j.physletb.2018.03.061.

[26] **ALICE** Collaboration, S. Acharya *et al.*, *Measurement of the production cross section of prompt* Ξ_c^0 *baryons at midrapidity in pp collisions at* $\sqrt{s} = 5.02\,TeV$. JHEP **2021**.159 (2021). DOI: 10.1007/JHEP10(2021)159.

[27] **CLEO** Collaboration, P. Avery *et al.*, *Inclusive production of the charmed baryon* Λ_c^+ *from* e^+e^- *annihilations at* $\sqrt{s} = 10.55\,GeV$. Phys. Rev. D **43** (11 June 1991), 3599–3610. DOI: 10.1103/PhysRevD.43.3599.

[28] **ZEUS** Collaboration, H. Abramowicz *et al.*, *Measurement of* D^+ *and* Λ_c^+ *production in deep inelastic scattering at HERA*. JHEP **2010**.9 (2010). DOI: 10.1007/JHEP11(2010)009.

[29] **ALICE** Collaboration, S. Acharya *et al.*, “Λ_c^+ production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02\,TeV$”. 2021. Submitted to Phys. Rev. Lett. B.

[30] T. Sjöstrand *et al.*, *An introduction to PYTHIA 8.2*. Computer Physics Communications **191** (2015), 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024.

[31] B. Andersson *et al.*, *Parton Fragmentation and String Dynamics*. Phys. Rep. **97** (1983), 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[32] J. R. Christiansen and P. Z. Skands, *String formation beyond leading colour*. JHEP 2015 (3 2015). DOI: 10.1007/JHEP08(2015)003.

[33] T. Sjöstrand and P. Z. Skands, *Baryon number violation and string topologies*. Nucl. Phys. B **659**.1 (2003), 243–298. DOI: 10.1016/S0550-3213(03)00193-7.

[34] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 tune*. Eur. Phys. J. C **74**.8 (2014). DOI: 10.1140/epjc/s10052-014-3024-y.

[35] J. Song, H. Li, and F. Shao, *New features of low p_T charm quark hadronization in pp collisions at $\sqrt{s} = 7\,TeV$.* Eur. Phys. J. C **78**.344 (2018). DOI: 10.1140/epjc/s10052-018-5817-x.

[36] 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. DOI: 10.1016/j.physletb.2021.136622.

[37] A. Andronic *et al.*, *Decoding the phase structure of QCD via particle production at high energy.* Nature **561** (2018), 321–330. DOI: 10.1038/s41586-018-0491-6.

[38] M. He and R. Rapp, *Charm-baryon production in proton-proton collisions.* Phys. Lett. B **795** (2019), 117–121. DOI: 10.1016/j.physletb.2019.06.004.

[39] 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** (1 July 2011), 014025. DOI: 10.1103/PhysRevD.84.014025.

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

[41] **ALICE** Collaboration, S. Acharya *et al.*, *Measurement of the Cross Sections of Ξ_c^0 and Ξ_c^+ Baryons and of the Branching-Fraction Ratio* $\mathrm{BR}(\Xi_c^0 \to \Xi^- e^+ \nu_e)/\mathrm{BR}(\Xi_c^0 \to \Xi^- \pi^+)$ *in pp Collisions at* $\sqrt{s} = 13$ TeV. Phys. Rev. Lett. **127** (27 2021), 272001. DOI: 10.1103/PhysRevLett.127.272001.

[42] C. Tsallis, *What are the Numbers that Experiments Povide.* Química Nova **17** (1994), 468–471.

[43] I. Kisel, I. Kulakov, and M. Zyzak, *Standalone first level event selection package for the CBM experiment.* In: *2012 18th IEEE-NPSS Real Time Conference.* 2012, 1–6. DOI: 10.1109/RTC.2012.6418385.

[44] xgboost developers, *XGBoost Documentation.* 2021. URL: https://xgboost.readthedocs.io/en/stable/index.html# (visited on 05/03/2022).

[45] **ALICE** Collaboration, F. Carminati *et al.*, *ALICE: Physics Performance Report, Volume I.* J. Phys. **G30** (2004), 1517–1763. DOI: 10.1088/0954-3899/30/11/001.

[46] **ATLAS** Collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider.* JINST **3**.08 (2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003.

[47] **CMS** Collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC.* JINST **3**.08 (2008), S08004–S08004. DOI: `10.1088/1748-0221/3/08/s08004`.

[48] **LHCb** Collaboration, A. Augusto Alves *et al.*, *The LHCb Detector at the LHC.* JINST **3**.08 (2008), S08005–S08005. DOI: `10.1088/1748-0221/3/08/s08005`.

[49] **ALICE** Collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC.* JINST **3** (2008), S08002. DOI: `10.1088/1748-0221/3/08/S08002`.

[50] A. Tauro, *ALICE Schematics.* General Photo.

[51] L. Betev *et al.*, *Definition of the ALICE Coordinate System and Basic Rules for Subdetector Components Numbering.* ALICE-INT-2003-038 (2003).

[52] **ALICE TPC** Collaboration, J. Alme *et al.*, *The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events.* Nucl. Instr. and Meth. **A 622** (2010), 316–367. DOI: `10.1016/j.nima.2010.04.042`.

[53] **ALICE** Collaboration, B. Abelev *et al.*, *Performance of the ALICE Experiment at the CERN LHC.* Int. J. Mod. Phys. **A29** (2014), 1430044. DOI: `10.1142/S0217751X14300440`.

[54] R. Kalman, *A New Approach to Linear Filtering and Prediction Problems.* ASME J. Basic Eng. **82**.1 (Series D 1960), 35–45. DOI: `10.1115/1.3662552`.

[55] **ALICE** Collaboration, B. Abelev *et al.*, *Upgrade of the ALICE Experiment: Letter Of Intent.* J. Phys. G: Nucl. Part. Phys. **41**.8 (July 2014), 087001. DOI: `10.1088/0954-3899/41/8/087001`.

[56] **ALICE** Collaboration, *Technical Design Report for the Muon Forward Tracker.* Tech. rep. Jan. 2015.

[57] **ALICE** Collaboration, P. Antonioli, A. Kluge, and W. Riegler, *Upgrade of the ALICE Readout and Trigger System.* Tech. rep. Sept. 2013.

[58] **ALICE** Collaboration, B. Abelev *et al.*, *Technical Design Report for the Upgrade of the ALICE Inner Tracking System.* J. Phys. G: Nucl. Part. Phys. **41**.8 (July 2014), 087002. DOI: `10.1088/0954-3899/41/8/087002`.

[59] **ALICE** Collaboration, *Upgrade of the ALICE Time Projection Chamber.* Tech. rep. Oct. 2013.

[60] S. Gorbunov, *On-line reconstruction algorithms for the CBM and ALICE experiments.* PhD thesis. Faculty of Computer Science and Mathematics Johann Wolfgang Goethe University Frankfurt, 2013.

[61] S. Gorbunov and I. Kisel, *Reconstruction of decayed particles based on the Kalman filter.* CBM-SOFT-note-2007-003 (2007).

[62] M. Zyzak, *Online selection of short-lived particles on many-core computer architectures in the CBM experiment at FAIR.* PhD thesis. Faculty of Computer Science and Mathematics Johann Wolfgang Goethe Uniersity Frankfurt am Main, 2015.

[63] B. Denby, *Neural networks and cellular automata in experimental high energy physics.* Computer Physics Communications **49**.3 (1988), 429–448. DOI: 10.1016/0010-4655(88)90004-5.

[64] K. Albertsson *et al.*, *Machine Learning in High Energy Physics Community White Paper.* J. Phys. Conf. Ser. **1085**.2 (2018), 022008. DOI: 10.1088/1742-6596/1085/2/022008. eprint: 1807.02876.

[65] J. Friedmann, *Greedy Function Approximation: A Gradient Boosting Machine.* Ann. Statist. **29** (2000). DOI: 10.1214/aos/1013203451.

[66] A. S. Cornell *et al.*, *Boosted decision trees in the era of new physics: a smuon analysis case study.* JHEP **2022**.4 (2022). DOI: 10.1007/jhep04(2022)015.

[67] T. Chen and C. Guestrin, *XGBoost: A scalable tree boosting system.* In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* KDD '16. 2016, 785–794. DOI: 10.1145/2939672.2939785.

[68] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification.* Nuclear Instruments and Methods in Physics **543** (2-3 2005), 577–584. DOI: 10.1016/j.nima.2004.12.018..

[69] L. Breimann *et al.*, *Classification and Regression Trees.* Monterey, CA: Wadsworth and Brooks, 1984.

[70] Y. Coadou, *Boosted Decision Trees and Applications.* EPJ Web of Confereces **55** (2013), 02004–. DOI: 10.1051/epjconf/20135502004.

[71] xgboost developers, *XGBoost Tutorials. Introduction to Boosted Trees.* 2021. URL: https://xgboost.readthedocs.io/en/stable/tutorials/model.html (visited on 05/03/2022).

[72] L. Barioglio *et al.*, *Minimal heavy-ion physics environment for Machine Learning (hipe4ML).* URL: https://doi.org/10.5281/zenodo.5070131 (visited on 05/03/2022).

[73] **ALICE** Collaboration, *ALICE 2016-2017-2018 luminosity determination for pp collisions at $\sqrt{s} = 13\,TeV$.* Tech. rep. July 2021.

[74] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1.* Computer Physics Communications **178**.11 (2008), 852–867. DOI: 10.1016/j.cpc.2008.01.036.

[75] R. Brun *et al.*, *GEANT Detector Description and Simulation Tool* (1994). DOI: 10.17181/CERN.MUHF.DMJ1.

[76] M. Arslandok *et al.*, *Track Reconstruction in a High-Density Environment with ALICE.* Particles **5**.1 (2022), 84–95. DOI: 10.3390/particles5010008.

[77] G. James *et al.*, *An Introduction to Statistical Learning. with Applications in R.* Ed. by G. Gasella, S. Fienberg, and I.Olkin. Second. Springer Texts in Statistics. New York: Springer, 2021. Chap. 5. ISBN: 978-1-0716-1417-4.

[78] M. Cacciari, M. Greco, and P. Nason, *The p_{T} spectrum in heavy-flavour hadroproduction.* J. High Energy Phys. **1998**.05 (May 1998), 007–007. DOI: 10.1088/1126-6708/1998/05/007.

[79] M. Cacciari *et al.*, *Theoretical predictions for charm and bottom production at the LHC.* J. High Energy Phys. **2021**.137 (2021). DOI: 10.1007/JHEP10(2012)137.

[80] **LHCb** Collaboration, R. Aaij *et al.*, *Measurement of b hadron fractions in 13 TeV pp collisions.* Phys. Rev. D **100** (3 Aug. 2019), 031102. DOI: 10.1103/PhysRevD.100.031102.

[81] **ALICE** Collaboration, S. Acharya *et al.*, *Measurement of Prompt D^0, Λ_c^+, and $\Sigma_c^{0,++}(2455)$ Production in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV.* Phys. Rev. Lett. **128** (1 Jan. 2022), 012001. DOI: 10.1103/PhysRevLett.128.012001.

[82] G. Luparello, "Ξ_c^+ reconstruction via hadronic decay channel in pp collisions at $\sqrt{s} = 13\,$TeV". Analysis Note. 2020.

[83] **LHCb** Collaboration, R. Aaij *et al.*, *Measurement of the mass and production rate of Ξ_b^- baryons.* Phys. Rev. D **99** (5 Mar. 2019), 052006. DOI: 10.1103/PhysRevD.99.052006.

[84] **Belle** Collaboration, Y. B. Li *et al.*, *First measurements of absolute branching fractions of the Ξ_c^+ baryon at Belle.* Phys. Rev. D **100** (3 Aug. 2019), 031101. DOI: 10.1103/PhysRevD.100.031101.

[85] R. A. Briceno, H. Lin, and D. R. Bolton, *Charmed-baryon spectroscopy from lattice QCD with* $N_f = 2 + 1 + 1$ *flavors*. Phys. Rev. D **86** (9 Nov. 2012), 094504. DOI: 10.1103/PhysRevD.86.094504.

A. Appendix

Table A.1.: List of features considered in this analysis.

Variable name	Description
chi2geo_Lam	χ^2 of geometrical fit of Λ daughter tracks
chi2geo_Xi	χ^2 of geometrical fit of Ξ^- daughters
chi2geo_XicPlus	χ^2 of geometrical fit of Ξ_c^+ daughters
chi2MassConst_Xi	χ^2 of mass constraint fit of Ξ^- daughters
chi2prim_PiFromXicPlus_sum	χ^2 of primary vertex fit of π^+ tracks from Ξ_c^+
chi2topo_XicPlus	χ^2 of topological constraint fit of Ξ_c^+ to the primary vertex
chi2topo_XiToPV	χ^2 of topological constraint fit of Ξ^- to the primary vertex
ct_Lam	$c\tau(\Lambda)$
ct_Xi	$c\tau(\Xi^-)$
DCA_PiToPi	DCA between the two π^+ from Ξ_c^+
DCAxy_LamDau	DCA between Λ daughter tracks in xy-direction
DCAxy_PiFromXicPlusToPV_KF_sum	Sum of DCA between the two π^+ and the primary vertex in xy-direction
DCAxy_PiToXi_sum	Sum of DCA between the two π^+ and Ξ^- in xy-direction
DCAxy_XiDau	DCA between Ξ^- daughter particles in xy-direction
DCAxy_XiToPV	DCA between Ξ^- and primary vertex in xy-direction
DecayLxy_Lam	Decay length of Λ in xy-direction
DecayLxy_Xi	Decay length of Ξ^- in xy-direction
DecayLxy_XicPlus	Decay length of Ξ_c^+ in xy-direction
ldl_Xi	Distance between Ξ^- production and decay vertex l normalised by the associated uncertainty Δl

PA_LamToPV	Angle between Λ momentum vector and line connecting Λ decay vertex with primary vertex
PA_LamToXi	Angle between Λ momentum vector and line connecting Λ decay vertex with Ξ^- vertex
PA_XicPlusToPV	Angle between Ξ_c^+ momentum vector and line connecting Ξ_c^+ decay vertex with primary vertex
PA_XiToPV	Angle between Ξ^- momentum vector and line connecting Ξ^- decay vertex and primary vertex

Declaration of Authorship

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 16. August 2022