UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Final Report for
Introduction to Research

Study of heavy-flavour baryons with machine learning with ALICE

Supervisor

Dr. Andrea Rossi

Co-supervisor

Dr. Mattia Faggin

Intern

Giovanni Celotto

June 2022 - October 2022

Abstract

A Large Hadron Collider Experiment (ALICE) at the CERN Large Hadron Collider (LHC) studies heavy-ion and proton collisions at ultra-relativistic energies. One of the main goals of the experiment is to study the quark-gluon plasma (QGP), a deconfined state of the strongly-interacting matter reproduced in laboratory via heavy-ion collisions. This state of matter is supposed to have characterized our universe in the first μs after the Big Bang. Heavy-flavour quarks (i.e. charm and beauty) are essential probes to investigate the properties of the QGP, since they are produced in the hard scatterings at the early stages of the collision and they experience the full evolution of the system, losing energy by interacting with the plasma constituents. These effects can be studied by measuring the heavy-flavour hadron production and comparing it with that in proton-proton (pp) collisions, used as a baseline to test our understanding of heavy-flavour physics in nucleus-nucleus collisions. Besides being a reference for heavy-ion collisions, the pp physics has an intrinsic interest, recently aroused by the results related to the relative production of charmed baryons with respect to mesons in pp collisions, showing an enhancement if compared to the predictions driven by measurements in e^-e^+ and e^-p collisions. In this context, the relative production of Λ_c^+ $/D^0$ has proven to be a successful observable [22]. In this report the measurement of the Λ_c^+ production cross section in pp collisions at $\sqrt{s} = 13$ TeV is studied. In Chapter 1 a brief introduction to the strong interaction, the QGP and the heavy-hadron production in pp collisions is presented. In Chapter 2 the acceleration system and the main detectors involved in this analysis are described. In Chapter 3 some preliminary studies on the signal extraction of the $\Lambda_c^+ \to pK^-\pi^+$exploiting the use of cuts on topological variables are given. In Chapter 4 a general description of Machine Learning is outlined and the application of Artificial Intelligence (AI) on the discrimination between signal and background is presented. Finally, a measurement of the Λ_c^+ production cross section in pp collision at $\sqrt{s} = 13$ TeV is given.

Contents

Chapter 1

Hadron physics

1.1 Strong interaction and Quantum Chromodynamics

The ordinary matter is composed by electrons, neutrons and protons. The latter ones are hadrons, and in particular baryons, which means that they have 3 valence quarks: up-up-down (uud) for the proton and up-down-down (udd) for the neutron. These quarks form a bound state, the nucleon, due to the strong interaction mediated by massless gauge bosons, called gluons. In the Standard Model, the strong interaction between quarks and gluons is described by Quantum Chromodynamics (QCD), a non-abelian gauge theory based on the $SU(3)$ local symmetry.
Differently from Quantum Electrodynamics (QED), where photons are electrically neutral, in QCD both quarks and gluons carry a so-called *color* charge. Therefore, gluons can also self interact, as a consequence of the non-abelian nature of QCD.
However, no color-charged particles (neither quarks nor gluons) have ever been observed as free particles in nature. They are always confined in the hadrons, which are singlets under the $SU(3)_C$ group[1] (i.e. they are colorless). The two main types of hadrons are the mesons ($q\overline{q}$), composed by one quark and one antiquark with the corresponding anticolor, and the baryons (qqq), that are bound states of three quarks with different colors. Evidences of tetraquarks and pentaquarks have also been observed, in agreement with QCD predictions [20]. The reason of this phenomenon, usually referred to as *confinement*, relies on the terms appearing in the QCD potential.
From measurements of two-jet angular cross-section in $p\overline{p}$ collisions, the QCD potential at short distances must have a QED-like dependence ($\propto 1/r$). Moreover, from the spins and masses of baryon resonances of the Δ and Λ families, it turned out that a term of the form $V \propto kr$ is required to explain the data at long distances [14]. Thus, a favoured QCD potential is the so-called Cornell potential:

$$V_{\mathrm{QCD}} = -\frac{\alpha_S}{r} + kr \tag{1.1}$$

The second term appearing in Eq. 1.1, which grows linearly with the distance, is responsible for the confinement of quarks inside the hadrons (Section 1.4). The coefficients α_S appearing in the QED-like term is the running strong coupling constant, whose value depends on the energy scale of the process. An example of the running behaviour of the strong coupling constant is shown in Fig. 1.1.

When the energy involved in the process increases, the value of the coupling constant decreases and the regime of *asymptotic freedom* is reached. Vice versa, when the energy decreases (or equivalently the length scale increases), the value of the coupling constant becomes larger, leading to the impossibility of treating QCD as a perturbative quantum field theory. For this reason, non-perturbative approaches must be used. An example of these is given by the *LatticeQCD*, a lattice gauge theory formulated on a grid of points in space and time.

[1] The subscript letter C stands for *color*

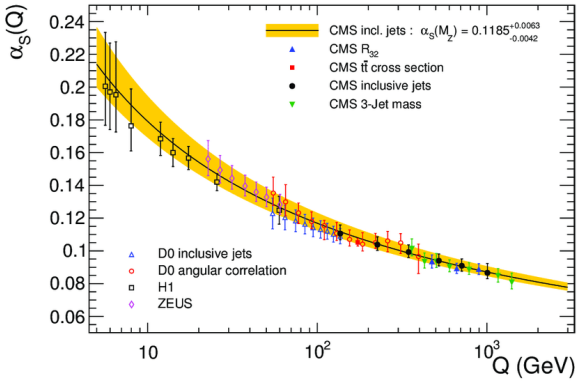


Figure 1.1: Summary of measurements of α_S as a function of the energy scale Q. Picture from [6].

1.2 Quark-Gluon Plasma

Even if quarks and gluons have never been observed as free particles in nature, QCD predicts the existence of the *Quark-Gluon Plasma* (QGP), a state of matter where they are deconfined[2]. This plasma is supposed to have characterized our universe in the first μs after the Big Bang and it can be reproduced in the laboratories using heavy-ion collisions at ultra-relativistic energies. Due to the large energy density required to reach the asymptotic freedom, the QGP can be formed only at extreme conditions of temperature or baryon density number (Fig. 1.2). The existence of a critical temperature close to $T = 155-170$ MeV at zero baryon density is predicted by LatticeQCD. Above this temperature the number of degrees of freedom (dof) of the hadronic system increases, implying the presence of a phase change from a gas of hadrons (mainly pions) to a plasma of quarks and gluons. If we consider the QGP as an ideal fluid composed by elementary particles with mass lower than the temperature of the system, the energy density is expected to grow as the fourth power of the temperature as ruled by the Stefan-Boltzmann law:

$$\epsilon = g\frac{k_B}{\hbar^3 c^3}\frac{\pi^2}{30}T^4 \qquad \text{where} \begin{cases} g = \frac{7}{8}n_{\text{dof}} & \text{for fermions} \\ g = n_{\text{dof}} & \text{for bosons} \end{cases} \tag{1.2}$$

The coefficient g takes into accounts the number of degrees of freedom due to spin, flavours and charges, eventually multiplied by a coefficient 7/8 in the case of fermions [17]. Under these assumptions the ratio between the energy density and the fourth power of the temperature ϵ/T^4 is proportional to the number of degrees of freedom of the system. As one can see in Fig. 1.3, a rise of this ratio is observed in correspondence of the critical temperature T_c. This can be explained as an increase in the number of dof occurring during the phase transition, while moving from $g = 3$ in the colourless gas of pions to $g = 37$ (47.5) in the deconfined plasma of gluons and 2 (3) flavours[3].

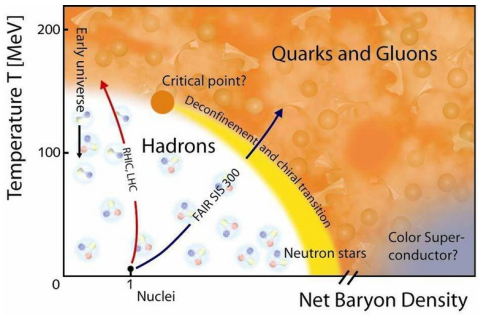


Figure 1.2: Phase diagram of hadronic matter.

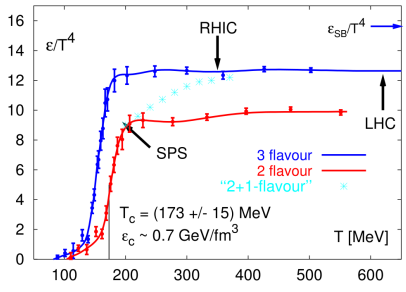


Figure 1.3: Lattice QCD predictions: ϵ/T^4 as a function of the temperature.

[2]This happens due to a modification of the Cornell potential, in which the linear contribution of Eq. 1.1 disappears and the QED-like term is screened by an exponential term[11]

[3]For the gas of pions we have three different isospin states. For the QGP we expect to have eight different gluons, each having two possible spin projections plus two (three) different quark flavours, each having three possible color charge, one corresponding antiparticle, two possible projections of the spin. Overall, in the QGP we have $16+\frac{7}{8}2(3)\cdot 2\cdot 2\cdot 3 = 37(47.5)$ degrees of freedom

1.3 Evolution of the system

In heavy-ion collisions at the LHC, lead (Pb) nuclei are accelerated up to $\sqrt{s_{NN}} = 5.02$ TeV and collide in specific points where energy densities of $\epsilon \geq 1 GeV/fm^3$ are reached, making it possible the formation of the QGP. The typical time in which the plasma is formed is $\tau = 1 fm/c$. After the production of the QGP, a local thermal equilibrium is reached and the system rapidly expands mainly along the longitudinal direction following relativistic hydrodynamic equations. The reason of this anisotropy is that the system is relativistically contracted along that axis[4] (Fig. 1.4).

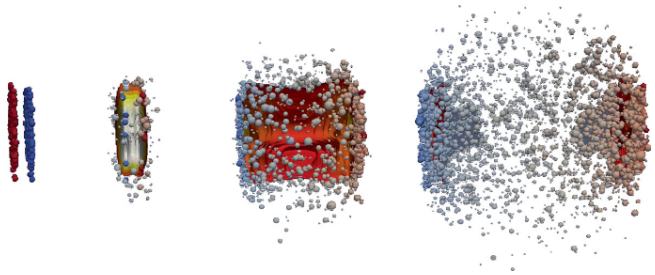


Figure 1.4: Four steps chronologically ordered from left to right of a nucleus-nucleus collision.

As long as the system expands, the temperature decreases down to the phase-transition critical value, when the system starts to convert back into the hadronic state. The hadronization happens at $t \simeq 10 fm/c$. After the hadronization, the system enters in the hadron gas phase when further inelastic scatterings modify the particle species. When the latter ones cease, the abundances of hadron species are frozen (*chemical freeze-out*) and only elastic collisions occur until distances between particles become too large as the system expands (*kinetic freeze-out*) [25]. Final state particles carry information on the QGP and are reconstructed by the detectors. A summary of the evolution of the QGP is depicted in Fig. 1.5.

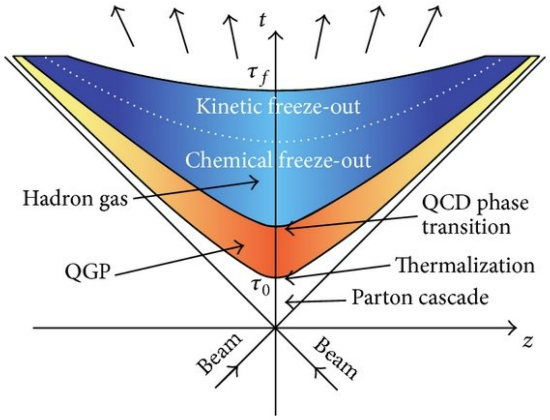


Figure 1.5: The space-time evolution of heavy-ion collision. Picture taken from [18].

Heavy-flavour quarks (i.e. charm and beauty) are essential probes to investigate the properties of the QGP. In fact, due to their large masses, heavy quark pair production ($c\bar{c}$ and $b\bar{b}$) is more likely to occur in the hard scatterings between the initial-state partons, therefore they can experience the full evolution of the system, losing energy by interacting with the plasma constituents [13]. These effects can be studied by measuring the heavy-flavour hadron production. However, in order to fully understand the properties of the QGP in Pb-Pb collisions, a deep comprehension of pp collisions, where no deconfined medium is expected, is also required.

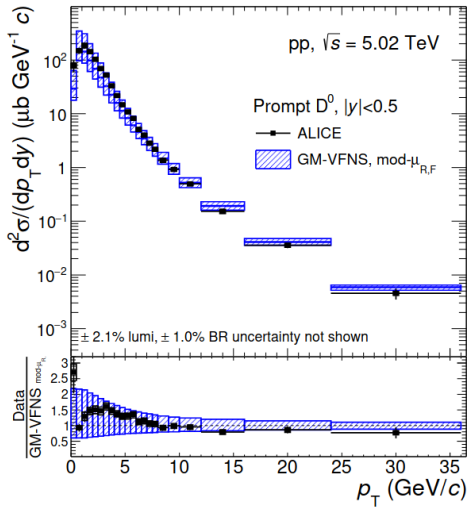
[4] Considering two Pb nuclei colliding at $\sqrt{s_{NN}} = 5.02$ TeV the typical Lorentz factor is $\gamma = \frac{E_N}{M_N c^2} \simeq \frac{2.5\,\text{TeV}}{1\,\text{GeV}} \simeq 2.5 \cdot 10^3$

1.4 Hadron production with ALICE in pp collisions

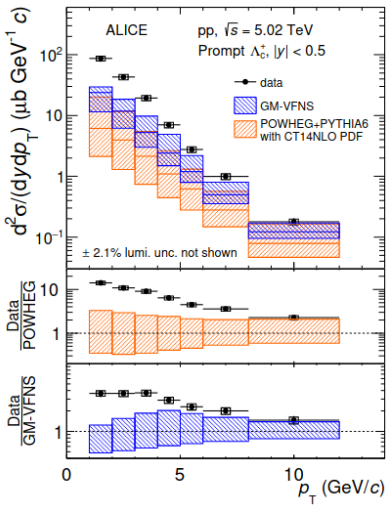
Besides being a reference for the heavy-ion collisions, the pp physics has an intrinsic interest, recently enhanced by the theoretical underestimation of the p_T-differential cross section in heavy-baryon production [22]. In these processes, the value of α_S is small due to the the large Q^2 required for the pair production of heavy quarks, therefore a perturbative approach is appropriate. In perturbative quantum chromodynamics (pQCD) the production cross section of heavy-flavour hadrons (H_Q) in pp collisions $d\sigma_{\mathrm{pp}\to H_Q+X}$ can be expressed factorizing it in three terms [4]:

$$d\sigma_{\mathrm{pp}\to H_Q+X}(\sqrt{s}) = \sum_{i,j=q,\overline{q},g} f_1(x_i,\mu_F^2) f_2(x_j,\mu_F^2) \otimes \sigma^{\mathrm{hard}}_{ij\to Q\overline{Q}}(\alpha_S(\mu_R^2),\mu_F^2,m_Q),x_i x_j s) \otimes D_Q^{H_Q}(z,\mu_F^2) \tag{1.3}$$

Starting with two colliding protons labelled as "1" and "2", the parton distribution functions $f_{1(2)}(x_i,\mu_F^2)$ give the probability for a given parton i inside the proton 1 (2) to carry the momentum fraction $x = p_i/p_p$ of the original proton. The cross section of the hard scattering $\sigma^{\mathrm{hard}}_{ij\to Q\overline{Q}}$ represents the probability of the physifcal process between the partons of the initial state to produce the heavy quarks of the final state ($c\overline{c}$ or $b\overline{b}$). Finally, the fragmentation function $D_Q^{H_Q}(z)$ quantifies the probability for a heavy quark Q to fragment into a hadron H_Q carrying a fraction $z = p_{H_Q}/p_Q$ of the original quark momentum. While the PDF are determined from measurements in deep-inelastic scattering, the fragmentation functions are usually parametrized from measurements of e^+e^- collisions.
The use of the factorization approach has been tested with the comparison between model predictions and experimental measurements. The results from the ALICE experiment has shown a production cross section for the D mesons (Fig. 1.6a) in agreement with models based on the factorization approach, whereas a noticeable underestimation was observed for the Λ_c^+ (Fig. 1.6b) [21, 22]. This may imply that fragmentation functions derived from e^-e^+ collisions cannot be used for pp collisions.



(a) Prompt D^0 p_T-differential production cross sections in pp collisions at $\sqrt{s} = 5.02$ TeV compared to model predictions. Picture from [21].



(b) Prompt Λ_c^+ p_T-differential production cross section in pp collisions at $\sqrt{s} = 5.02$ TeV compared to model predictions. Picture from [22].

Figure 1.6: p_T differential production cross section in pp collisions at $\sqrt{s}$=5.02 TeV for D^0 and Λ_c^+

Not only the production cross section for heavy-baryon is underestimated by theoretical models, but also the predictions on the baryon-to-meson production ratio, which correctly work for e^-e^+ and e^-p collisions, fail in the case of pp collisions. Among these models, one of the most widely used is the Lund string model, which is implemented in Monash, the default tune of PYTHIA event

generator. In this model, a *colour string* linking a pair of $q\bar{q}$ is introduced to explain the production of additional quarks. Assuming to parameterize the potential between a quark and an antiquark with the Cornell potential (Eq. 1.1), in the Lund model only the linear term is taken into account, which is the dominant one at large distances between the pair. Once a $q\bar{q}$ pair is produced in a collision, the increasing distance between the two quarks causes a stretching of the colour string. When the energy stored in the colour string exceeds the amount needed to produce an additional $q\bar{q}$ pair, the string can break into two new quarks, as an effect of the confinement.

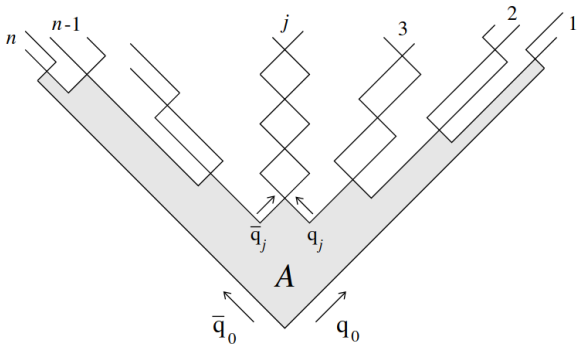
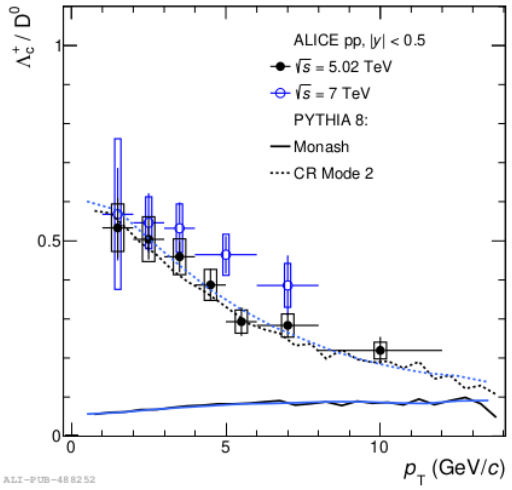
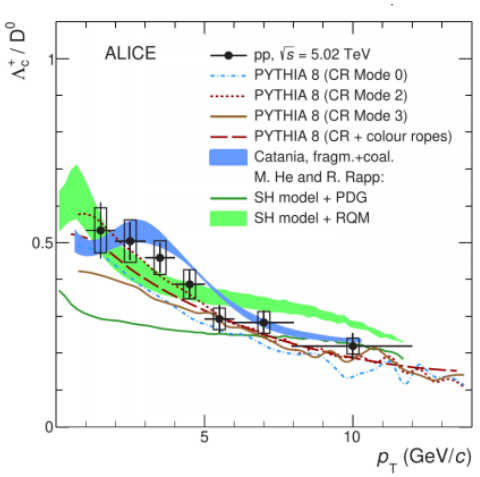


Figure 1.7: Scheme of string fragmentation and color confinement. The grey area indicate where the colour field does not vanish. Picture from [1].

The conventional string fragmentation described in the Lund model had been successfully used for the description of hadron production in e^+e^- and e^-p collisions. However, hadronic collisions at the LHC have shown results not in agreement with the predictions of the model, especially for charmed baryons (Fig. 1.8a). As a consequence, new theoretical models have been developed to explain this discrepancy (Fig. 1.8b) [3, 9, 12, 19].



(a) Λ_c/D^0 ratio in pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV and $\sqrt{s} = 7$ TeV compared with predictions from Pythia predictions with Monash and colour reconnection beyond leading colour approximation (CR).



(b) Λ_c/D^0 ratio in pp collisions at $\sqrt{s} = 5.02$ TeV compared with predictions from Catania model, SHM+RQM model and Herwig.

Figure 1.8: Comparison of Λ_c/D^0 ratio in pp collisions with models. Pictures from [22]

Chapter 2

The ALICE experiment

2.1 The LHC collider

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator. It consists of a 27-kilometer ring of superconducting magnets at a depth ranging from 50 to 175 metres. It was built by the European Organization for Nuclear Research (CERN) between 1998 and 2008 to replace the dismantled Large Electron-Positron Collider (LEP).
The LHC mainly collides proton beams up to $\sqrt{s} = 13.6$ TeV with a peak luminosity of $\mathcal{L} = 10^{38} cm^{-2} s^{-1}$, but it can also accelerate beams of heavy ions: lead-lead (Pb-Pb) collisions and proton-lead (p-Pb) collisions are studied at $\sqrt{s_{NN}} = 5.02$ TeV.
The beams collide in four crossing points where seven experiments, with different structures and purposes, are positioned around.

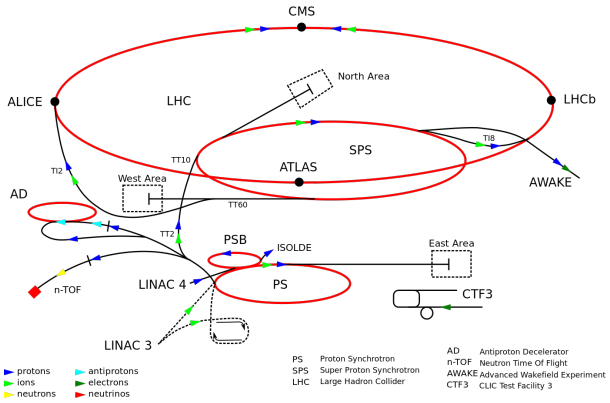


Figure 2.1: LHC accelerator complex.

The proton beams in the LHC are made up of bunches of protons, spaced 25 ns apart, with each one containing more than 10^{11} particles. Before the collisions take place, the two beams inside the accelerator travel at relativistic energies in opposite directions and in separate beam pipes, that are kept at ultrahigh vacuum (10^{-10}mbar). The beams are guided around the accelerators by the presence of a magnetic field of $B = 8.4$T, produced by superconducting electromagnets. The latter are kept at a temperature of $T = -271.3$ °C, provided by a complex distribution system of liquid Helium.

The collider is preceded by a pre-accelerating system, aiming to gradually increase the energy of the protons or lead ions. In particular, this chain involves the *LINear ACcelerator 2* (LINAC2), which takes the protons extracted from hydrogen atoms and accelerates them to ≈ 50 MeV. Then, they are injected in the *Proton Synchrotron Booster* (PSB), where they reach the energy of ≈ 1.4 GeV and sends them to the *Proton Synchrotron* (PS), which makes them travel at the energy ≈ 25 GeV. The last step before entering in the LHC is in the *Super Proton Synchrotron* (SPS), where they are accelerated up to ≈ 450 GeV.

2.2 The ALICE experiment

A Large Ion Collider Experiment (ALICE) at CERN is a general-purpose experiment designed to study the physics of strongly interacting matter and the QGP in nucleus-nucleus collisions at the LHC. In addition to heavy-ion collisions, the ALICE collaboration also studies collisions of protons (pp), which primarily provide reference data for the nucleus-nucleus collisions.
The detector consists of a central barrel, which measures event-by-event hadrons, electrons and photons, and of a forward spectrometer to measure muons. The central part, which covers polar angles from 45° to 135°[1] over the full azimuth, is surrounded by the L3 solenoidal magnet that provides a magnetic field of 0.5 T.

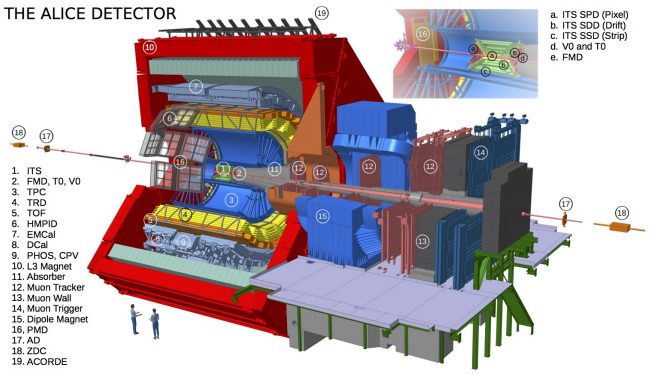


Figure 2.2: ALICE experiment.

The barrel consists of an Inner Tracking System (ITS) of high-resolution silicon detectors, a cylindrical Time-Projection Chamber (TPC) and three particle identification arrays: Time-Of-Flight (TOF) detector, Transition-Radiation Detector (TRD) and a single-arm ring imaging Cherenkov (HMPID). In the outermost region within the L3 magnet, two complementary calorimeters are situated: the Photon Spectrometer (PHOS) and the Electromagnetic Calorimeter (EMCal). The forward muon spectrometer (covering a pseudorapidity range of $-4 < \eta < -2.5$) consists of a complex arrangement of absorbers, a large dipole magnet, and fourteen planes of tracking and triggering chambers. Several smaller detectors (FMD, T0 and V0) for global event characterization and triggering are located at forward angles [2].
In this chapter the main detectors used for the following analysis are briefly described.

2.2.1 ITS

The ITS is the innermost detector surrounding the beampipe. For LHC Run 1 and Run 2, it consisted of six cylindrical layers characterized by different technologies: the first two layers formed the Silicon Pixel Detectors (SPD) and they were located respectively 3.9 and 7.6 cm away from the central axis; the two intermediate layers formed the Silicon Drift Detectors (SDD) and they were 15.0 cm and 23.9 cm far away from the centre; the two outermost layers constituted the Silicon Strip Detectors (SSD), respectively at 38.0 and 43.0 cm (Fig. 2.3a).
The main purposes of the ITS are the following:

- Determination of the primary vertex with a spatial resolution $< 100\mu m$
- Reconstruction of the secondary vertices of charmed or beauty hadrons
- Tracking and identification of charged particles with $p_T < 200$ MeV/c

[1] Using the ALICE coordinate system [5] the angle θ is defined with respect to the axis of collisions. This angle range corresponds to the pseudo-rapidity range $|\eta| < 0.9$

- Improvement of the momentum measurement and angular resolution of the particles reconstructed by the TPC.

For LHC Run 3 a new Monolithic Active Pixel Sensors (MAPS) detector has replaced the detector described above.

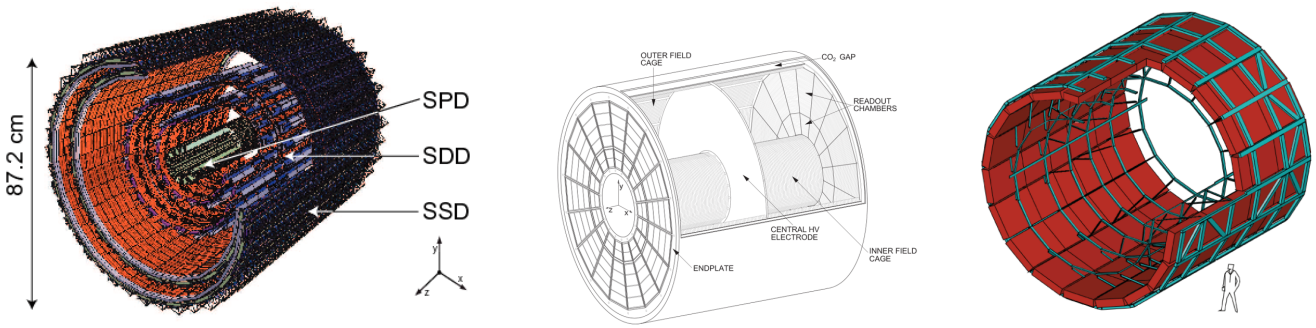
2.2.2 TPC

The TPC is the main detector of the central barrel for charged particles tracking momentum measurements and Particle Identification (PID). From a mechanical point of view it is composed by a cylindrical field cage divided into two drift regions by the presence of a central HV electrode (Fig. 2.3b). The volume of the chamber is $90m^3$ with an inner radius of $R_{\text{int}} = 84.8$ cm and an outer radius of $R_{\text{out}} = 246.6$ cm. The length of the chamber along the z direction (parallel to the beam using the ALICE coordinate system [5]) is 500 cm. As a consequence, the maximum drift path is 2.5 m. In the endplates, the readout chambers are situated. They are divided azimuthally into 18 sectors and further divided into Inner Readout Chambers (IROCs) and Outer Readout Chamber (OROCs) along the radial direction. For Run 2 the detectors were 72 Multiwire Proportional Chambers (MWPCs) that allow a bidimensional reconstruction of the charged particle tracks. The third coordinate is computed from the drift time.

During LHC Run 1 the gas mixtures used was composed of Ne-CO_2-N_2 (90:10:5). During LHC Run 2 the gas mixture of Ar-CO_2 (90:10) was also tested. For the future runs including Run 3, Ne-CO_2-N_2 (90:10:5) is planned to be used due to the smaller space charge distortion it induces with respect to Ar-CO_2 [15]

2.2.3 Time Of Flight (TOF)

The Time-Of-Flight detector (TOF) covers the central region ($|\eta| < 0.9$) with a large cylindrical array ($\approx$170 m^2) at a radial distance of 4.7 m (Fig. 2.3c). It provides charged-particle identification in the intermediate momentum range.
Its main purpose is to measure the time that each particle takes to travel from the vertex to reach it. It has a global time resolution of 80 ps, that makes it possible to provide π/K and K/p separation better than 3σ up to a particle momentum $p \simeq 2.5$ GeV/c and p$\simeq$4 GeV/c respectively.
The smallest unit it is composed of is the double-stack Multigap Resistive Plate Chamber (MRPC) strip. The ALICE TOF array consists of 1593 MRPC strips, subdivided into 18 azimuthal sectors. One strip has a 1207.4 cm^2 active area and it operates in a $C_2H_2F_4$ (90%), i-C_4H_{10} (5%), SF_6 (5%) gas mixture. The MRPC strip is segmented into two rows of 48 pickup pads of $3.5 \cdot 2.5$cm^2 , for a total of about 160000 readout channels. [16]



(a) Inner Tracking System (b) Time Projection Chamber (c) Time Of Flight

Figure 2.3

Chapter 3

Signal of the $\Lambda_c^+ \to pK^-\pi^+$

The baryon Λ_c^+ is composed in terms of valence quarks by an up quark, a down quark and a charm quark (udc). The mass of this baryon is $m = 2286.46 \pm 0.14$ MeV/c^2 and the proper decay length is given by $c\tau = 60.7\mu m$ [24]. The decay channel we are interested in is the $\Lambda_c^+ \to pK^-\pi^+$, whose Branching ratio (BR) is $BR = (6.28 \pm 0.32)\%$. The study of the signal is performed in 6 different transverse momentum (p_T) classes, that are 1-2 GeV/c, 2-4 GeV/c, 4-6 GeV/c, 6-8 GeV/c, 8-12 GeV/c, 12-24 GeV/c. In Chapter 4, the analysis and the measurement of the cross section will proceed only with the p_T bins 2-4 GeV/c, 4-6 GeV/c, 6-8 GeV/c employing machine learning techniques.

3.1 Data

Each Λ_c^+ candidate is reconstructed as a triplet of a proton p, a negatively charged kaon K^- and a positvely charged pion π^+ tracks. The criteria used to select the tracks are shown in Tab.3.1 [8].

Single track variable	**Selection**
$\|\eta\|$	< 0.8
p_T	> 0.3 GeV/c
ITS refit	yes
Track-to-vertex distance in xy plane	< 0.15 cm
Track-to-vertex distance in z direction	< 0.25 cm
Number of TPC crossed rows	> 70
Found/findable clusters in TPC	> 0.8
χ^2 / clusters in TPC	< 4
χ^2 /clusters in ITS	< 36
$\|n_\sigma^{TPC}\|$	< 3
$\|n_\sigma^{TOF}\|$	< 3

Table 3.1: Single track selection criteria used to select candidate p, K$^-$ and π^+ particles.

The above-mentioned tracks are required to be within the pseudorapidity range $|\eta| < 0.8$ and to have a minimum transverse momentum of 300 MeV/c to reduce the combinatorial background when reconstructing triplets at low p_T . The maximum track distance of closest approach (DCA) to the primary vertex must be smaller than $d_{xy} < 0.15$ cm along the transverse plane and smaller than $d_z < 0.25$ cm along the longitudinal direction in order to reduce the number of secondary particles coming from long-lived particle-decays. Besides this, a further reduction of the background is achieved by applying some preliminary PID criteria. The loss of energy of negative charged tracks are required to be compatible within 3σ with a kaon, while the positive charged ones are asked to be compatible with the proton or pion species. This identification is performed using the TPC signal for all the particles and the TOF signal when available. In the latter case, the particle is accepted as a proton, kaon or pion only if identified by both detectors. The Λ_c^+ candidates are finally reconstructed as

triplets among two positive charged particles and a negative charged one selected with the criteria previously mentioned. In the same way, the antiparticles Λ_c^- are reconstructed as triplets of two negative charged particles and a positive charged one. Given the selection criteria and the detector acceptance, in order to avoid border effects all the Λ_c candidates are reconstructed within the fiducial acceptance region, defined by:

$$y_{\rm fid}(p_T) = \begin{cases} 0.5 + \frac{1.9}{15} p_T - \frac{0.2}{15} p_T^2 & \text{for } p_T < 5\text{GeV/c} \\ 0.8 & \text{for } p_T > 5\text{GeV/c} \end{cases} \tag{3.1}$$

3.2 Secondary vertex and topological variables

When a Λ_c^+ is produced in the primary vertex (i.e. the collision vertex between the proton beams), the boost of the baryon makes it travel at a velocity high enough to decay in a point displaced from the primary vertex. The Λ_c^+ reconstruction in the p $K^-\pi^+$ decay channel exploits this displacement between the primary and the secondary vertex. Indeed, for a Λ_c^+ with a momentum equal to its mass, the average decay length in the laboratory frame is $L = \beta\gamma c\tau = \frac{p}{m} c\tau \simeq 61\mu m$, which is comparable with the xy resolution of the ITS (2.2.1). The decay point of the Λ_c^+ baryon, called secondary vertex, is found as the 3D space point (x_0, y_0, z_0) that minimises the distance among the tracks of the decay products (i.e. $pK\pi$), that is minimizing the following quantity:

$$D = \sqrt{\sum_{i=1}^{3} \left(\frac{x_i - x_0}{\sigma_{x_i}}\right)^2 + \left(\frac{y_i - y_0}{\sigma_{y_i}}\right)^2 + \left(\frac{z_i - z_0}{\sigma_{z_i}}\right)^2} \tag{3.2}$$

where (x_i, y_i, z_i) are the coordinates of the i-th track propagated to the point of closest approach among daughter tracks and $(\sigma_{x_i}, \sigma_{y_i}, \sigma_{z_i})$ are the corresponding uncertainties. The position resolution close to the primary vertex given by the innermost layer of the ITS is good enough to allow us to exploit several variables correlated to the displaced topology of the charm-hadron decays, in order to maximize the signal-to-background ratio (S/B). In fact, differently from a real decay of a Λ_c^+ , most of the triplets constituting the background directly come from the primary vertex and therefore they are not characterised by a real displacement. If a displacement is observed in a triplet that constitutes the background, it has to be ascribed to the resolution of the apparatus. A description of the variables exploited in the analysis is given below and a simple cartoon with the $\Lambda_c^+ \to pK^-\pi^+$is in Fig. 3.1.

- **Decay length in the xy plane ($\mathbf{L}_{xy}$).** The decay length is defined as the projection of the distance between the primary vertex and the Λ_c^+ decay point on the xy plane. The reason of the projection is in order to fully exploit the resolution of the ITS along the plane orthogonal to the beam axis, which is significantly better than the one along the z direction.

- **Normalized decay length ($\mathbf{nL}_{xy}$).** This quantity corresponds to the decay length L normalized to the uncertainty estimated from the covariance matrix of the decay product tracks. For the same above-mentioned reason, also in this case only the xy component is considered.

- **Cosine of pointing angle (θ_p).** The pointing angle is defined as the angle between the direction of the reconstructed Λ_c^+ momentum (from the measured momentum of the decay products) and the "flight line" connecting the primary vertex with the Λ_c^+ decay point, determined as explained above in Eq.3.2. This angle is expected to have a null value in the ideal case of infinite resolution of the detectors. Since the cosine of this angle for the signal is expected to be more peaked at 1 than the background, a lower limit will be imposed as a cut to maximize the S/B ratio.

- **Maximum "topomatic" in the transverse plane (d_{res}).** This quantity derives from the transverse decay length (L_{xy}) and the single-track impact parameter d_0 ,defined as the distance of closest approach in the transverse plane between the primary vertex and the reconstructed track propagated at the point of closest approach in xy with respect to it. The maximum "topomatic" d_{res}^{xy} is defined as:

$$d_{\rm res}^{\rm xy} = \max_{i=1}^{3} \left(\frac{d_{0,i}^{\rm xy} - d_{0,i}^{\rm xy}(\exp)}{\sqrt{\sigma^2(d_{0,i}^{xy}) + \sigma^2(d_{0,i}^{xy}(\exp))}} \right) \tag{3.3}$$

where the index i denotes one of the three Λ_c^+ decay products.

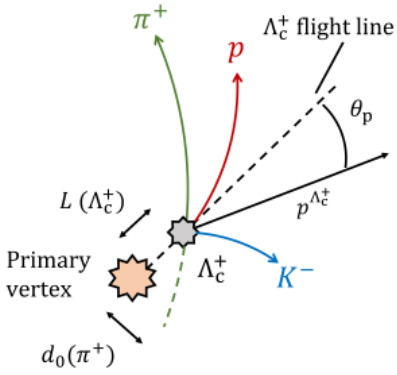


Figure 3.1: Cartoon representing the $\Lambda_c^+ \to pK^-\pi^+$. Picture from [8].

3.2.1 PID selection - a Bayesian approach

One of the key point of this analysis in the identification of the $\Lambda_c^+ \to pK^-\pi^+$candidates is the employment of a Bayesian Particle Identification (PID) method [7]. In this method for a given particle species H_i and a given detector $\alpha = \{\text{TOF, TPC, ...}\}$ with an expected produced signal $\hat{S}_\alpha\left(H_i\right)$ a likelihood $P_\alpha\left(S_\alpha|H_i\right)$ is defined. The meaning of this likelihood is the probability for a particle of species H_i to produce a signal S_α in the detector α (= TPC, TOF, ...). The advantage of using the Bayesian approach from an analytical point of view is that the probabilities of the different detectors can be easily combined making the product of them. Using the Bayes's theorem, one can compute the posterior probability that is the probability that the particle which released a set of signals $\vec{S}$ on the different detectors belongs to the species H_i.

$$P\left(H_i|\vec{S}\right) = \frac{P\left(\vec{S}|H_i\right)C\left(H_i\right)}{\sum\limits_{k=\pi,k,e,\ldots} P\left(\vec{S}|H_k\right)C(H_k)} \tag{3.4}$$

The prior probability $C(H_i)$, which represents our initial knowledge on the particles species crossing the detectors, is given by the relative abundance. The final identification of the particle can rely on different criteria, the one applied in this work is the maximum probability criterion (i.e. a given track is identified with the particle species with the maximum posterior probability).

3.3 Optimization of the S/B ratio

The study of the optimization of the S/B ratio is of deep interest for the best estimate of the baryon production. To do it an algorithm was developed within the analysis framework ROOT.

3.3.1 Description of the algorithm

In the code, the data already selected with the above-mentioned selection criteria (Tab. 3.1 and Bayes PID) are divided into the 6 classes of p_T. At this point, we want to estimate the signal and the background, before applying any cut on topological variables. To do this, first of all the background is fitted using a second-degree polynomial in the sideband regions (Fig. 3.2).

This interpolation is used in order to give some hints for the determination of the background when the full fit of signal and background is performed. The latter involves the sum of a second-degree polynomial and a gaussian function.

$$\underbrace{\frac{A}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}_{\text{Signal}} + \underbrace{p_0 + p_1 \cdot x + p_2 \cdot x^2}_{\text{Background}} \tag{3.5}$$

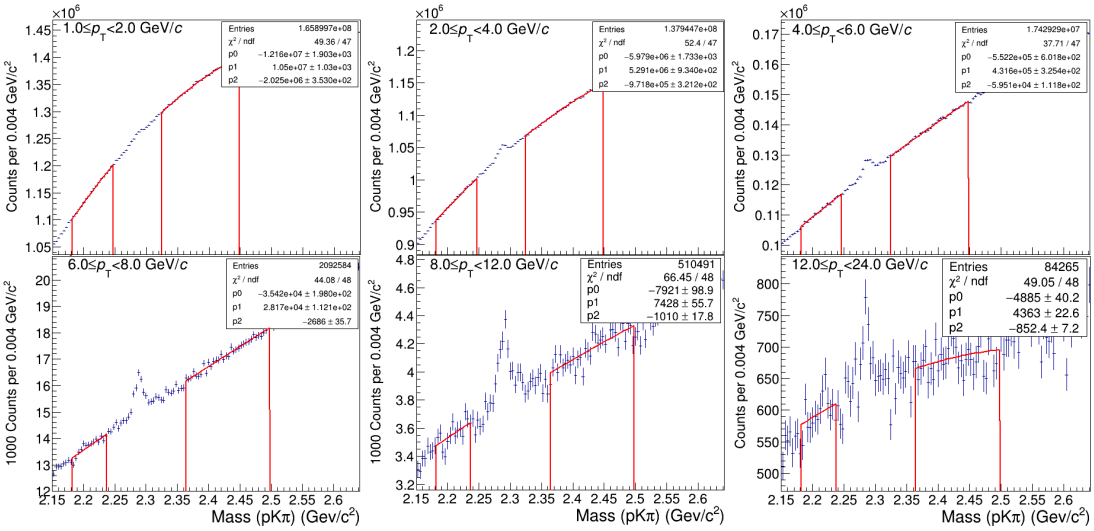


Figure 3.2: Invariant mass distributions of triplets $pK\pi$ for different p_T classes and fit results for the background.

Due to the chosen normalization of the gaussian function, the parameter A gives the number of events under the gaussian, once it is divided by the adopted bin width. In the plots presented in Fig. 3.3, the centroid and the dispersion of the gaussian are reported together with the corresponding errors. The number of events within 3σ is also reported, obtained from the parameter A and exploiting the fact that the area within 3σ under a normalized gaussian is 0.9973. The corresponding error is obtained in the same way. The Background within 3σ is computed integrating the second-degree polynomial appearing in Eq. 3.5. The errors provided is obtained extracting the covariance sub-matrix corresponding to the three parameters from the full covariance matrix of the fit. Finally, the S/B ratio and the statistical significance within 3σ (Signif.) are computed from the definitions in Eq. 3.6. The latter is an indication of the inverse relative uncertainty on the measured signal.

$$S/B(3\sigma) = \frac{S(3\sigma)}{B(3\sigma)} \qquad \text{Signif}(3\sigma) = \frac{S(3\sigma)}{\sqrt{S(3\sigma)+B(3\sigma)}} \tag{3.6}$$

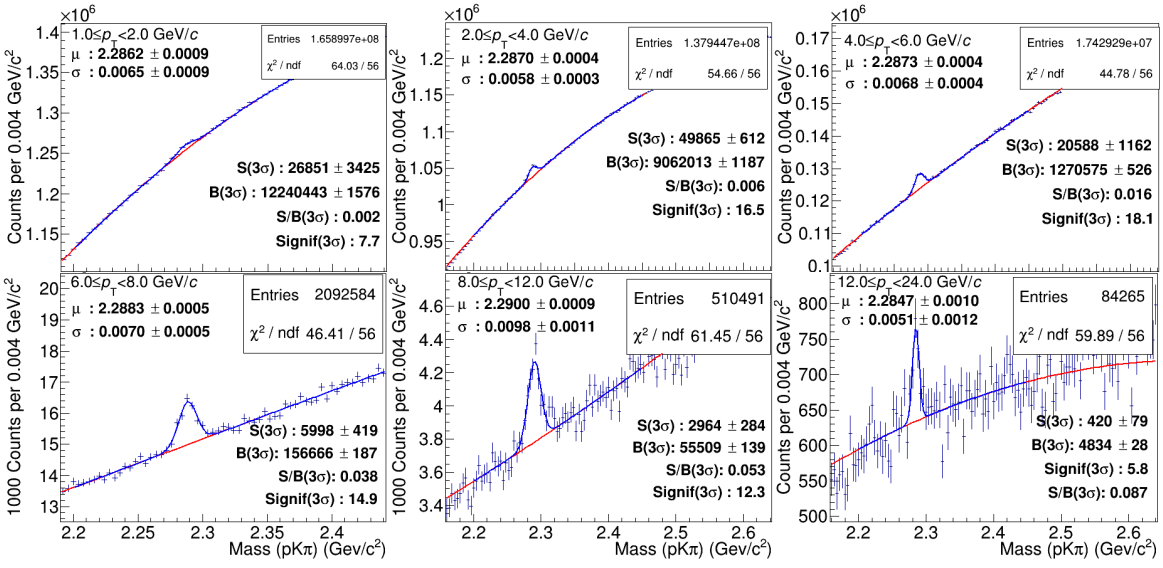
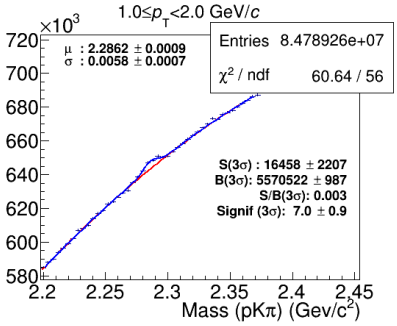


Figure 3.3: Fit of the signal before applying any cut on topological variables.

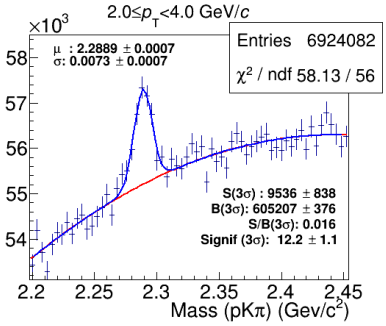
In order to maximize the S/B ratio without reducing the statistical significance, the following procedure was followed. One variable at a time is taken into account and 10 different cuts on that variable are scanned. Every time a cut is performed, the fit procedure described above is done and the S/B ratio and statistical significance are computed. For each case the best compromise between optimizing S/B and the statistical significance was chosen. Once the optimized cut for the first variable is determined, this cut is fixed and a scan over the possible cuts on the second variable starts. When all the 4 variables are optimized, the next p_T class is chosen. The plots presented below show the final optimized situation for all the p_T classes (Fig. 3.4). The table shown below summarizes the cuts applied to each variable to obtain these results. As one can see, improvements in the S/B of a factor 4 were obtained in some cases increasing the statistical significance as well (Fig. 3.4d, 3.4e, 3.4f). Other times a compromise in terms of statistical significance was necessary in order to increase the S/B ratio.

p_T bin (GeV/c)	1 – 2	2 – 4	4 – 6	6 – 8	8 – 12	12 – 24
Cosine(θ_P)	[0.8 - 1]	[0.88 - 1]	[0.84 - 1]	[0.86 - 1]	[0.92 - 1]	[0.94 - 1]
L_{xy} (cm)	[0.00 - 0.05]	[0 - 0.05]	[0.015 - 0.05]	[0.02 - 0.05]	[0.025 - 0.05]	[0.02 - 0.05]
nL_{xy}	[0 - 8]	[2 - 8]	[2 - 8]	[1.5 - 8]	[1 - 8]	[2.5 - 8]
$normImpPar_{xy}$	[0 - 1]	[0 - 1.5]	[0 - 2.5]	[0 - 1.5]	[0 - 4]	[0 - 1]
Initial S/B	0.002	0.006	0.016	0.038	0.053	0.087
Final S/B	0.003	0.016	0.05	0.126	0.183	0.46
Initial Signif	7.7	16.5	18.1	14.9	12.3	5.8
Final Signif	7.0	12.2	16.5	14.9	13.8	6.9

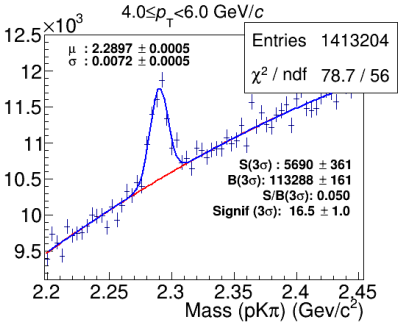
Table 3.2: Cuts applied on the topological variables and summary of S/B and statistical significance.



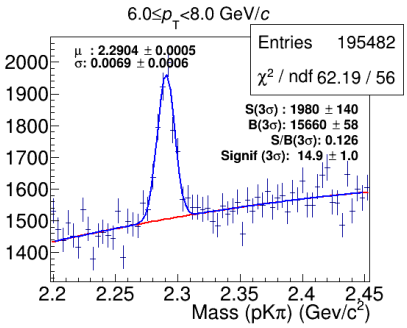
(a) Fit results for p_T class 1-2 GeV/c



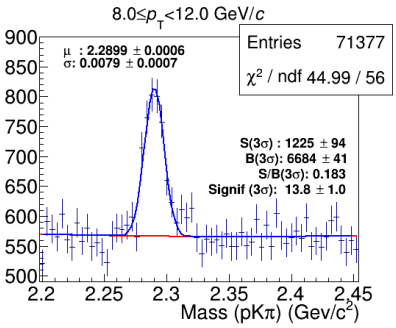
(b) Fit results for p_T class 2-4 GeV/c



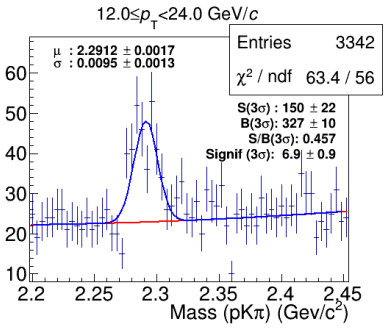
(c) Fit results for p_T class 4-6 GeV/c



(d) Fit results for p_T class 6-8 GeV/c



(e) Fit results for p_T class 8-12 GeV/c



(f) Fit results for p_T class 12-24 GeV/c

Figure 3.4: Optimization results for the 6 p_T bins.

Chapter 4

Application of ML techniques on the dataset

In the previous chapter the problem of discriminating the signal from the background was faced exploiting the event topology and some selection criteria on the decay variables. However some critical issues have been neglected:

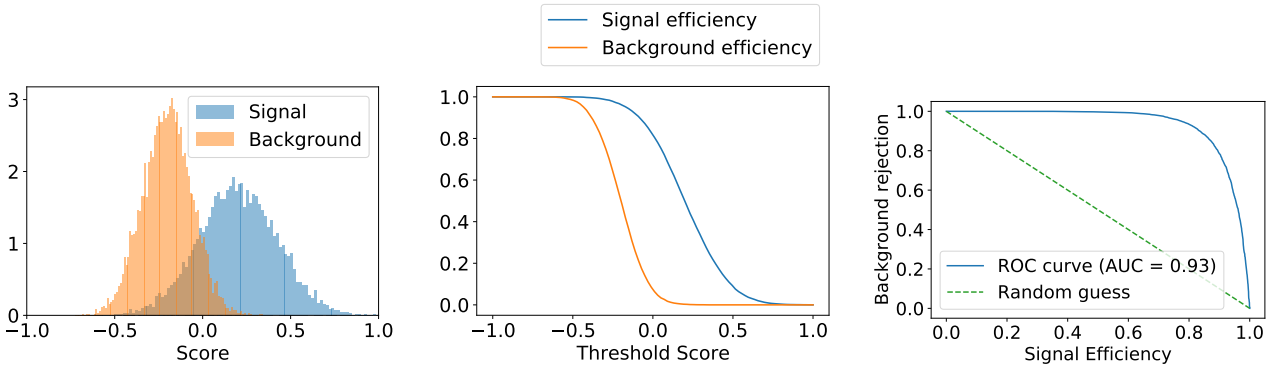
- The optimal cuts were determined according to the fit results, therefore they are extremely sensitive to the statistical fluctuations, which can mask or enhance the real signal (the analysis is not *blind*)
- The optimal cuts that were found correspond to a *local*[1] optimization and not a global one.
- Given the 4 topological variables, a 4-dimensional hyperspace can be considered. The cuts that were done on the topological variables lead to a *connected* subspace of such a hyperspace. A possible better solution can be found considering the union of non-connected subspaces of whatever shape.

4.1 What is Machine Learning?

Machine Learning (ML) is a subfield of artificial intelligence (AI), whose generally aim is to discover patterns in the user data and fit that data into models. The two most common tasks addressed with this tool are *classification* and *regression*. In the first case, the aim is the discrimination among two or more classes, while in the second case a prediction on a given variable is required according to the input data. For our purposes, the distinction between real decay events and background events is an example of classification problem that can be performed using *supervised* ML.
The working principle consists in providing the algorithm a dataset labelled as *class 1* events and another one labelled as *class 2*. The algorithm finds a common pattern among the objects of the same class giving for each event a *score*, a limited value based on the features used to train the model. For this reason, this step is usually referred to as *training*. This value will be later used to decide which class each event of an unknown dataset belongs to, comparing it with a fixed threshold.
In a second step, the *testing*, the algorithm performances are checked with another different known a priori dataset. Finally, the algorithm is used to classify unknown events, assigning a score to them according to what previously learnt. After fixing a given threshold t, all the events with a score higher than t are classified as belonging to the class 2 (signal), the others to the class 1 (background). Of course, the accuracy of this discrimination depends on the shape of the score distributions and their separation. For example, in Fig. 4.1a if $t = -1$ all the events are classified as "signal" and both the signal and background efficiencies are equal to 1 (Fig. 4.1b). The ideal case consists in heaving signal efficiency equal to 1 and backgorund efficiency equal to zero (or equivalently background rejection equal to 1). Since an overlapping region is always present, the rejection of background events is accompanied

[1] In order to find a global optimal set of cuts, a 4-dimensional hyperspace should have been scanned, making this procedure much more time-consuming without solving the other two issues.

by the loss of a fraction of the signal. The receiver operating characteristic (ROC) curve (Fig. 4.1c) represents the background rejection as a function of the signal efficiency. In case of equal output score distributions for the two classes, the ROC curve would correspond to a straight line connecting the points $(1, 0)$ and $(0, 1)$ and the area under it would be 0.5. In the latter case, the model would be completely useless, since each event would be randomly classified as signal or background.
Generally, the better trained the algorithm is, the closer to unity the ROC AUC is. However, this number just gives a general indication of the separation power of the model independently from the threshold adopted and without being sensitive to the relative sizes of the two classes



(a) Class score distribution (b) Signal and background efficiencies (c) ROC curve

Figure 4.1: Examples of score distributions and application of some thresholds in a toy model.

4.1.1 Boosted Decision Trees (BDT)

A *decision tree* (DT) is a flowchart-like structure that takes a set of input features and splits input data recursively based on those features. The data is split in the internal nodes according to the value of one of the input features. The typical structure of a DT is shown in Fig. 4.2, where the signal (S) and background (B) correspond to the two classes of events to be distinguished. The tree develops from a root node where a first test is performed. Such a test corresponds to the comparison of a numeric feature (e.g. the decay length) with a threshold value established to separate the two classes. A single event is then sent to one of the two directions before being eventually tested on a different feature. When a terminal node (a.k.a. leave) is reached, the single event is assigned to a particular class.

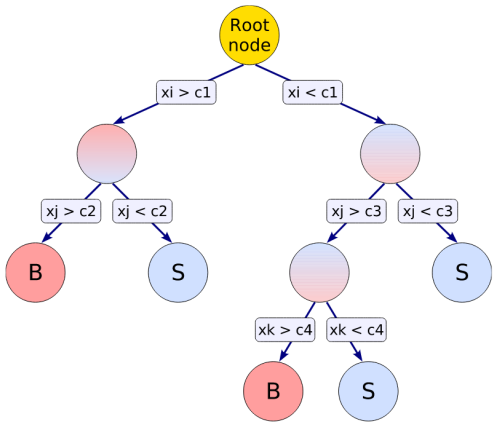


Figure 4.2: Schematic view of a decision tree. Picture from [10].

However, if we used only one DT, the discrimination power will be limited and the algorithm stability may be strongly dependent on a small fluctuation in the training dataset. For this reason,

a possible solution consists in combining the outputs of several DTs into one final classification. Since the creation of several independent trees is computationally demanding and the addition of a redundant DT that does not improve the classification is to be avoided, boosting techniques are usually adopted in this context. In these methods the DTs are built sequentially, trying to add a new DT that compensates the lacks of the previous classifier. This is done re-weighting the event, in such a way that the more often an event is misclassified by the previous classifier, the more important it becomes so that the new classifier will take more care of it. The BDT provides an output score that is related to the outcome of each DTs (called weak learners with respect to it), according to the boosting algorithm. One of the most popular one is called Adaptive Boost (AdaBoost) and this is the algorithm adopted in this work.

4.2 Search for $\Lambda_c^+ \to pK^-\pi^+$signal with AdaBoost

The signal of the decay $\Lambda_c^+ \to pK^-\pi^+$has been searched using Adaboost in the same dataset of pp collisions at $\sqrt{s} = 13$ TeV in the following transverse momentum bins: 2-4 GeV/c, 4-6 GeV/c, 6-8 GeV/c. The $pK\pi$ candidates are reconstructed using the same criteria reported in Table 3.1, but no Bayes PID was employed this time. The optimisation of single-track and topological selections is performed with a BDT model based on the AdaBoost boosting method.

4.2.1 Training and Testing

As previously explained, in the supervised ML the model needs to be trained with data whose class is known in advance. In our case, a set of background events is taken from the $pK\pi$ candidates on the sides of the peak[2]. For what regards the signal, a set of Monte Carlo simulated events is taken. The MC dataset used for the training is composed by $pK\pi$ candidates coming from *prompt* Λ_c^+ decay. This means the Λ_c^+ originated in the primary vertex[3] and did not come from the decay of beauty hadrons. In the latter case, the Λ_c^+ is said to be *non prompt*. It is noteworthy to say that the considered decay $\Lambda_c^+ \to pK^-\pi^+$can occur either in a direct way or via a resonant channel, that is passing through an intermediate state of a short-lived particle [24]. The relative BR of these 4 channels are not correctly simulated in the MC and one should proceed differentiating these 4 situations, since the decay topology and the efficiencies are different. In the following analysis, we are computing one unique efficiency without distinguishing the four decay channels[4]. From the original datasets of background and MC events, half of the events are used for the training and the other half is used for the testing, where the splitting between these two steps is done randomly. As an example, the score distributions and the ROC curves for the training and testing for the p_T bin 6-8 GeV/c are shown below. As one can see from Fig. 4.3b the two curves are quite similar and the AUC is close to 0.9, both are an indication of the fact that the algorithm was well-trained.

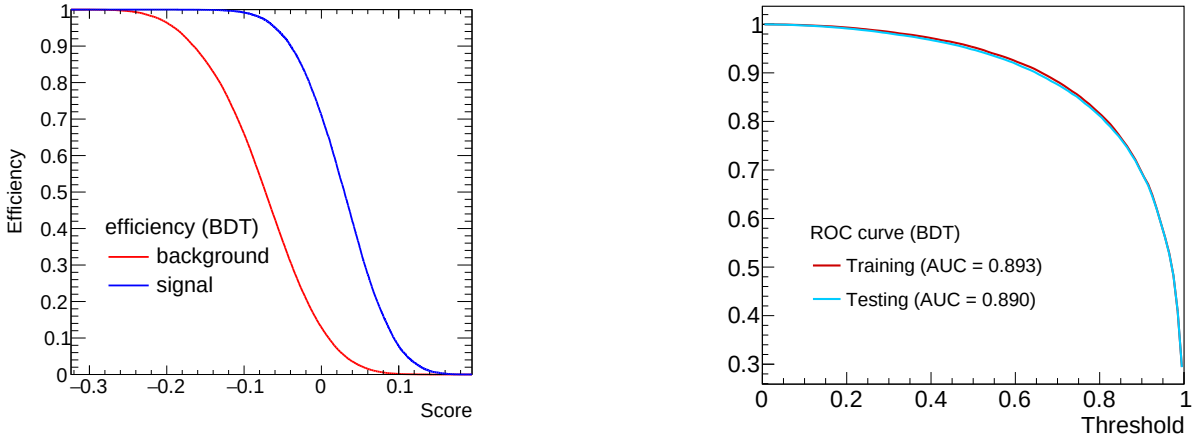
4.2.2 Application

Once the performances of the algorithm have been checked, the model is ready to be used for an unknown dataset, that is the sample of $pK\pi$ candidates reconstructed from the pp collisions. As previously mentioned, a final score is assigned to each $pK\pi$ candidate. First of all, a scan over some possible values of thresholds was done. Some results for the three p_T classes are reported in the appendix. As one can see, when the cut is weak (the first row of the plots) the signal is not well visible, whereas when the threshold on the BDT score increases, the number of background events significantly decreases leading to a higher S/B ratio.

[2]For the background not all the available statistics was used but only 5% to avoid the problem of *overtraining*

[3]Also Λ_c^+ coming from the decay product of charmed hadrons which were in turn prompt hadrons are defined as prompt Λ_c^+

[4]This is a clear limitation of this analysis and it will be outlined in the conclusions after the comparison with other measurements



(a) Signal and background efficiencies

(b) ROC curves and AUC

Figure 4.3: Results of training and testing for 6-8 GeV/c p_T bin.

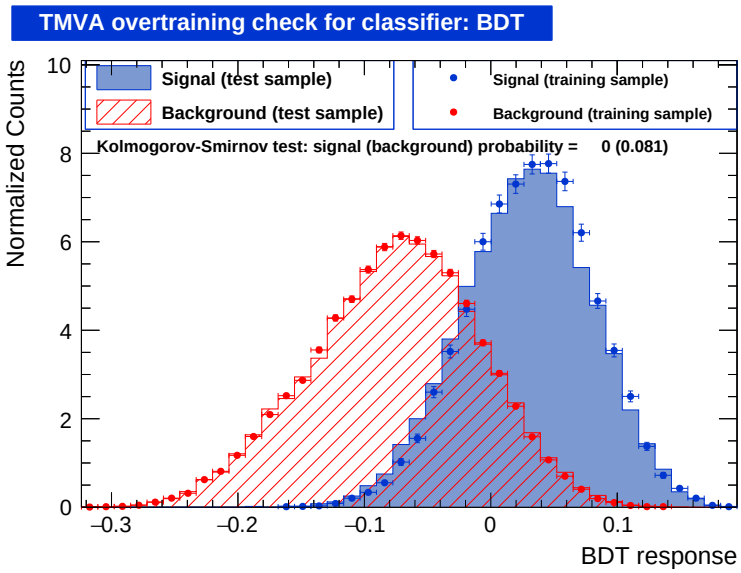


Figure 4.4: BDT score distributions of signal and background for training (histograms) and testing (points) of the 6-8 GeV/c p_T bin.

4.2.3 Determination of the working point

At this point two problems arise:

- The scan is not continuous, therefore an optimal working point[5] based on the S/B ratio and statistical significance will not be accurate
- Even if the first issue was solved, the critical issue we pointed out in the previous chapter related to the sensitivity to the fluctuations of the dataset can be repeated. This problem comes from the fact that we are considering the results of the fit as an indication of the best possible cut.

In order to overcome these issues, we can exploit the fact that knowing the BDT distributions and the *expected* signal and background events in our dataset, a prediction on the working point can be done. What we miss is the evaluation of the expected signal and background events. The latter ones are determined performing a fit on the sides of the invariant mass peak of the events within a given p_T

[5]By working point we mean an optimal cut on the BDT score that maximizes the expected significance

bin. From those parameters the number of events within 3σ is easily evaluated[6]. For what concerns the determination of the signal events, it can be estimated from the following expression:

$$N_{\text{raw}}^{pk\pi} = \left(\frac{d\sigma}{dp_T}^{\text{INPUT}}_{|y|<0.5} \right) (\epsilon \cdot \text{acc})^{\text{PRE-SEL}}_{\text{PROMPT}} \cdot \text{NORM} \tag{4.1}$$

$$\text{where NORM} = \frac{BR \cdot 2 \cdot \Delta p_T \Delta y \cdot \mathcal{L}}{(f_{\text{PROMPT}} = 1)}$$

The first term is the partial differential cross section in the rapidity range $|y| < 0.5$ obtained from another independent decay mode of the Λ_c^+ : $\Lambda_c^+ \to pK_S^0$. The second term is the efficiency of the prompt Λ_c^+ in the preselection phase times the geometrical acceptance. This can be expressed as the following product:

$$\epsilon \cdot \text{acc} = \frac{\text{Reco}^{\text{Analysis}}}{\text{GenAcc}} \frac{\text{GenAcc}}{\text{GenLimAcc}} = \frac{\text{Reco}^{\text{Analysis}}}{\text{GenLimAcc}} \tag{4.2}$$

The term $\text{Reco}^{\text{Analysis}}$ corresponds to the reconstructed prompt Λ_c^+ in the MC sample. This term includes all those Λ_c^+ generated in the MC sample whose decay products satisfy the conditions written in Tab. 3.1 and whose track is within the fiducial acceptance (Eq.3.1). The term GenAcc is the set of potentially reconstructable tracks, which corresponds to the $\text{Reco}^{\text{Analysis}}$ as long as no analysis cuts are applied on the MC data. GenLimAcc is the set of Λ_c^+ which satisfies $|y| < 0.5$, that is the rapidity range in which cross section is expressed.

Finally, the normalization term in Eq. 4.1 takes into account the BR of the decay channel, the integrated luminosity[7], the width of the p_T bin (in our case always $\Delta p_T = 2\text{GeV/c}$), the width of the rapidity Δy at which the differential cross section is measured (in our case $\Delta y = 1$) and a factor 2 due to the corresponding process involving antiparticles $\Lambda_c^- \to \overline{p} K^+ \pi^-$. In principle, we should divide this number by a coefficient called *prompt fraction*, which is the ratio between the prompt and the total number of Λ_c^+ . However, since the production cross section of a charm quark is much larger than the one of a beauty, we can reasonably assume this fraction to be equal to 1. The normalization term is therefore equal in each case considered and its value is NORM $\simeq 7.3 \cdot 10^6$. The summary of the other terms, the final expected signal and background events and the optimal working point at which the cut on the BDT score will be performed are shown in Table 4.1. An example of the predicted significance as a function of the cut on the BDT score showing a maximum in correspondence of the working point WP=0.035 for the p_T class 6-8 GeV/c is shown in Fig. 4.5 (green curve).

p_T (GeV/c)	2–4	4 – 6	6–8
Reco	49956	48507	25687
GenLimAcc	648005	250952	91366
$\epsilon \cdot$ acc	0.077	0.19	0.28
Expected signal	29035	15654	5633
Expected Bkg	23301743	4445156	701478
Working point	0.041	0.038	0.035

Table 4.1: Parameters used for the determination of the working point.

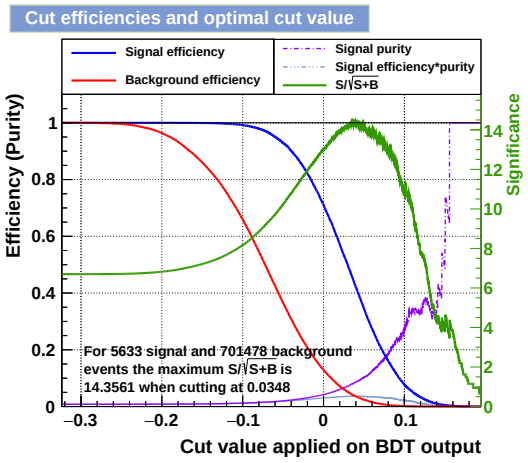


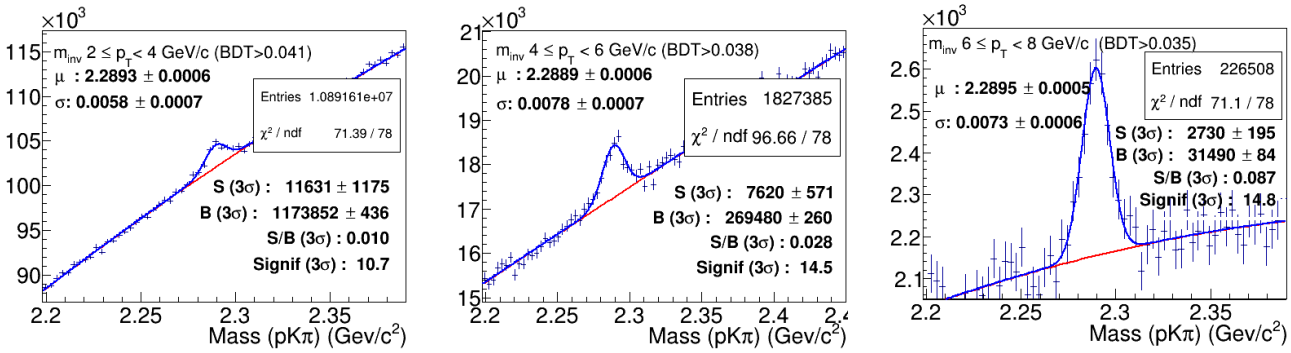
Figure 4.5: Computation of the working point and expected significance for p_T bin 6-8 GeV/c.

[6]To this purpose the mean and the sigma were computed from the MC events where no background is present and the interpolation of the background is performed in points at least 5σ far from the centroid

[7]The integrated luminosity can be in turn expressed as the ratio between the total number of collected events over the minimum bias cross section

4.3 Measurement of the cross section

We can now perform the cut on the BDT score according to the optimal point determined in the previous section and finally perform the fit of the signal as explained in Chapter 3. The plots are shown below (Fig. 4.6).



(a) (b) (c)

Figure 4.6: Fit results after setting the working point.

The signal extracted plays a fundamental role in the determination of the cross section. To compute the cross section, one should take into account that the cut on the BDT score has introduced a further inefficiency, since part of the signal is inevitably lost. Therefore, we need to know how many signal events we have removed after the cut on the BDT score. To do this, the same cut is applied on the MC events, where we know that all the events belongs to the class *signal*. As a consequence, the overall efficiency for both the prompt and nonprompt signals will read:$(\epsilon \cdot \text{acc})^{\text{PRESEL}} \cdot \epsilon_{BDT}$. Done this, one can take the Eq.4.1 and isolate the partial differential cross section:

$$\frac{d\sigma}{dp_T}^{\text{INPUT}}_{|y|<0.5} = \frac{N_{\text{raw}}^{pk\pi} f_{\text{PROMPT}}}{(\epsilon \cdot \text{acc})^{\text{total}}_{\text{PROMPT}} \cdot \text{NORM}} \tag{4.3}$$

The term $N_{\text{raw}}^{pk\pi}$ is the amount of signal we have found after the cut on the working point (i.e. the values of the signal shown in Fig. 4.6), the first term of the denominator represents the total efficiency of the process (selection, geometry and BDT cut). The normalization term is the same found in Eq. 4.1 without f_{PROMPT}, here written explicitly. In this case the prompt fraction is computed from another external input. In particular, given the efficiencies for the prompt and non prompt signal and the differential cross section for non prompt Λ_c^+ , the f_{PROMPT} can be estimated as the complementary to the unity of $f_{\text{NON PROMPT}}$, which in turn is computed from an equation similar to Eq. 4.1. The values of this fraction for the three considered p_T classes range from 88% for the bin 6-8 GeV/c to 93% for the 2-4 GeV/c.
The result of the differential cross section, obtained dividing by the BR of the process $\Lambda_c^+ \to pK^-\pi^+$is shown in Fig. 4.7 compared to the results published by the ALICE collaboration coming from the decay mode $\Lambda_c^+ \to p^+K_S^0$. The two pt-differential cross sections show a difference that can be caused by known limitations of the analysis performed in this work, as discussed in next section. Nevertheless, the difference is not too large[8], and does not give evidence of biases from the usage of machine-learning classification for the determination of the selections applied. Such biases could arise in case the classification relies on signal features poorly described in the Monte Carlo simulation or on spurious correlations among the variables for signal and background.

[8]The relative differences between the measurement of this work and the results from [23] for the three p_T bins considered are 13%, 6% and 10% respectively.

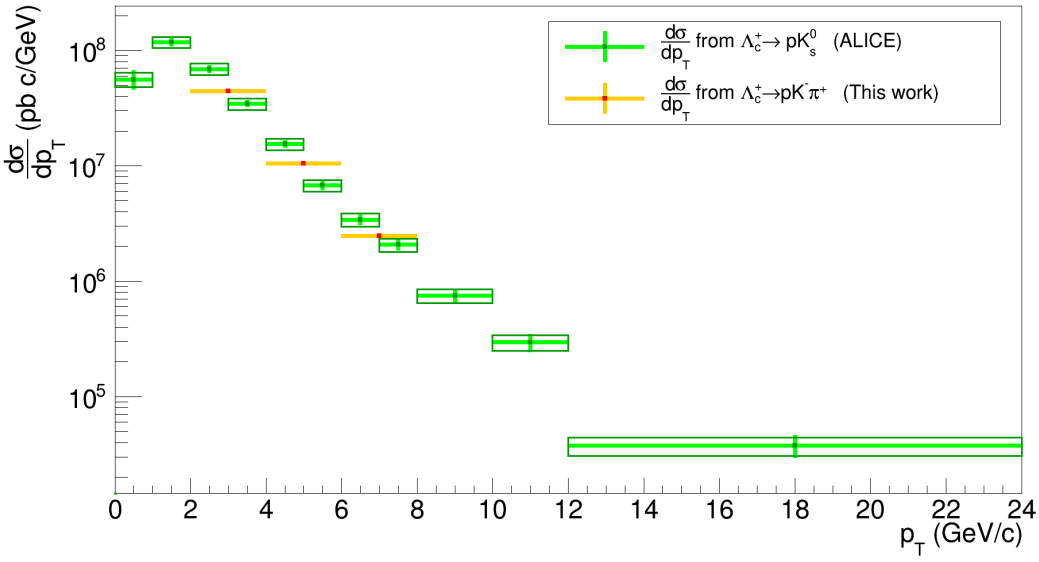


Figure 4.7: Prompt Λ_c^+ p_T-differential production cross section obtained in this work compared to the results published in [23]. Statistical uncertainties are shown as vertical bars, while systematic uncertainties are shown as boxes, and the bin widths are shown as horizontal bars.

4.3.1 Limitations and perspectives

The values we found for the differential cross section are compatible within the uncertainties but slightly smaller than those published by the ALICE collaboration. This is due to some limitations of the performed analysis that are summarized here.

- The 4 decay channels, each characterized by a different efficiency and topology are treated as a unique one. A proper analysis, beyond the aim of this report, should distinguish between these 4 decay modes of the Λ_c^+ and compute a final efficiency as a weighted average of the four efficiencies, where the weight is given by the BR.

- A further limitation consists in having used the same sample for the computation of the efficiency and the training of the model. This is in principle wrong, and one should avoid making use of the training sample during the application phase, especially in case of overtraining, that is when the model learns how to classify the events in the training samples without extrapolating the general patterns from the classes. Nevertheless, given the results of the testing, our case is far from this situation and the influence on the final result is negligible with respect to the first limitation pointed out.

- In the determination of the working point a single cut on the BDT score was taken into account. However, a neighbourhood of that point shows similar values of expected significance (Fig. 4.5). A scan over different values of the working points should have been considered to quantify a contribution of systematic uncertainties on the final measurement.

Conclusions

In this work a measurement of the production p_T-differential cross section of the Λ_c^+ in pp collisions at $\sqrt{s} = 13$ TeV from the decay channel $\Lambda_c^+ \to pK^-\pi^+$using Multivariate Analysis techniques has been presented. Despite of the highlighted limitations, the results are in good agreement with the values published by the ALICE collaboration coming from the independent decay mode $\Lambda_c^+ \to p^+K_S^0$

Appendices

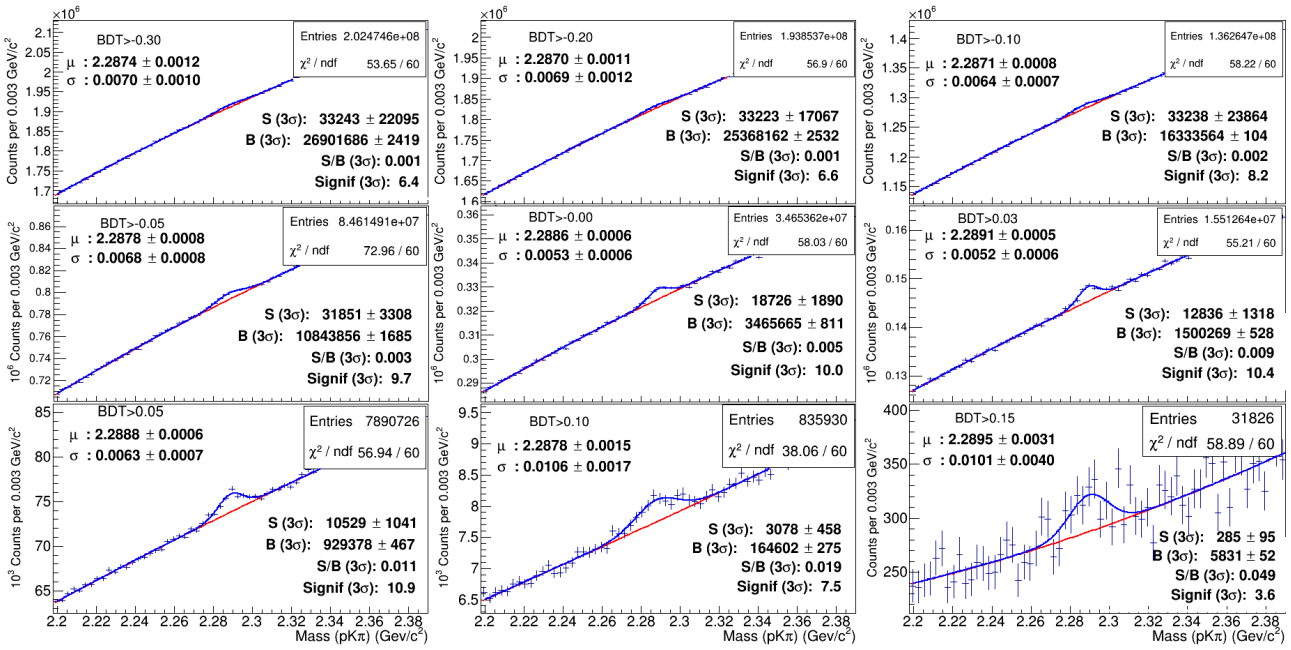


Figure 8: Invariant mass distribution filtered by a scan over 9 possible cuts on the BDT score for the p_T class 2-4 GeV/c.

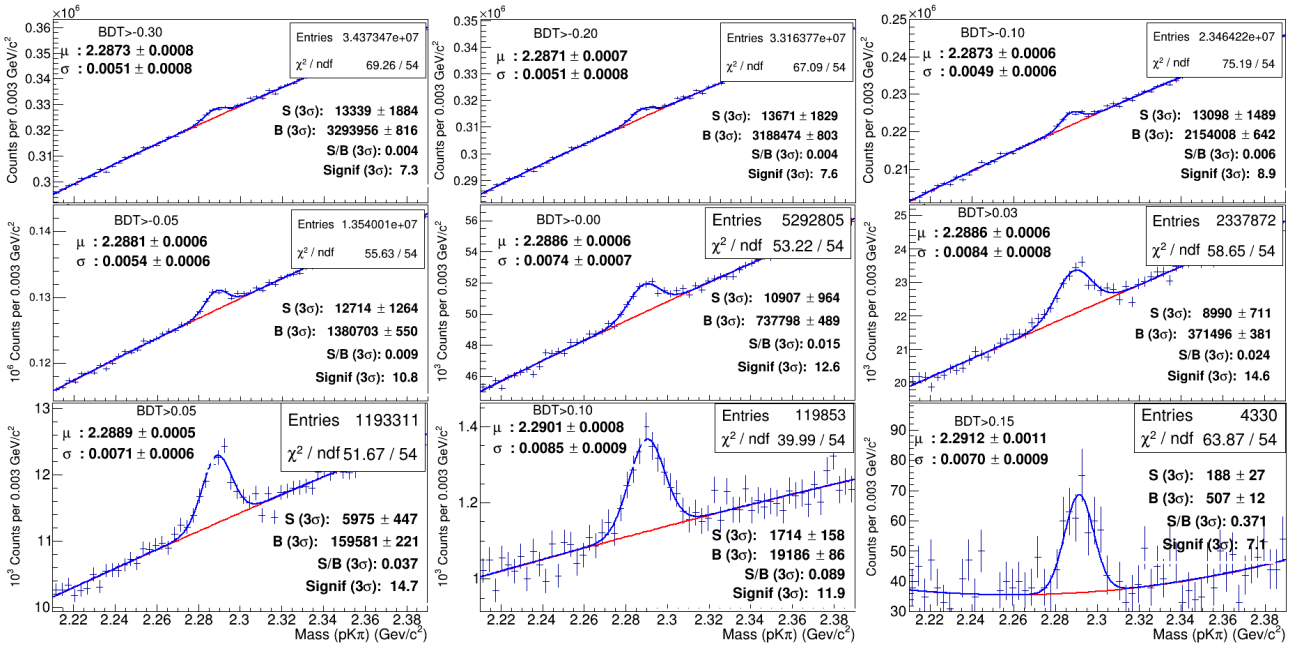


Figure 9: Invariant mass distribution filtered by a scan over 9 possible cuts on the BDT score for the p_T class 4-6 GeV/c.

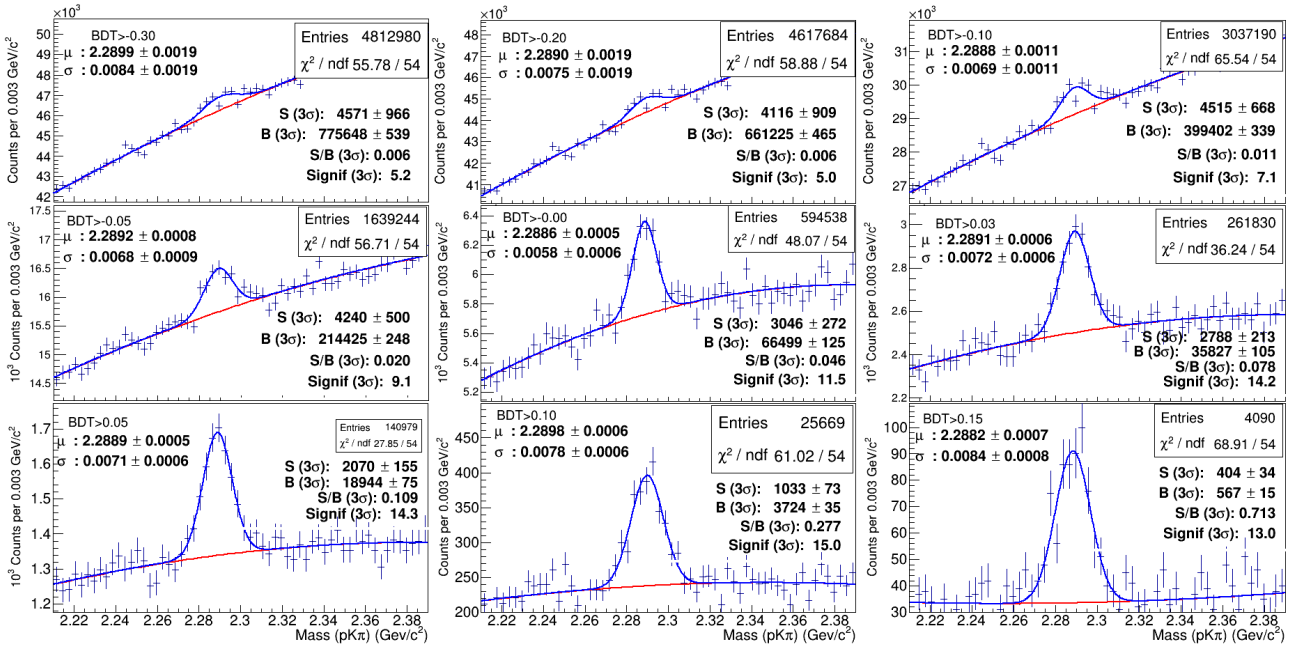


Figure 10: Invariant mass distribution filtered by a scan over 9 possible cuts on the BDT score for the p_T class 6-8 GeV/c.

Bibliography

[1] Bo Andersson, G. Gustafson, Jari Hakkinen, Markus Ringnér, and Peter Sutton. "Is there screwiness at the end of the QCD cascades?". *Journal of High Energy Physics*, page 014, 07 1998.

[2] Iouri Belikov. "ALICE Statistical Wish-list". 2008.

[3] Jesper R. Christiansen and Peter Z. Skands. "String Formation Beyond Leading Colour". *JHEP*, 08:003, 2015.

[4] John C. Collins, Davison E. Soper, and George Sterman. "Factorization of Hard Processes in QCD". 2004.

[5] G. Dellacasa et al. "ALICE: Technical design report of the time projection chamber". 1 2000.

[6] David d'Enterria, Peter Skands, S. Alekhin, Andrea Banfi, Siegfried Bethke, Johannes Bluemlein, K. Chetyrkin, D. d'Enterria, G. Dissertori, X. Tormo, A. Hoang, M. Klasen, T. Klijnsma, S. Kluth, J. Kneur, B. Kniehl, D. Kolodrubetz, J. Kühn, P. Mackenzie, and I. Stewart. "High-precision α_s measurements from LHC to FCC-ee". 12 2015.

[7] J. Adam et al. "Particle identification in ALICE: a Bayesian approach". 2016.

[8] Mattia Faggin. "Measurement of heavy-flavour decay electrons and heavy-flavour baryon production with ALICE experiment at LHC", 2021. Presented 16 Dec 2021.

[9] Min He and Ralf Rapp. "Charm-baryon production in proton-proton collisions". *Physics Letters B*, 795:117–121, 2019.

[10] A. Hoecker *et al.* "TMVA - Toolkit for Multivariate Data Analysis", 2007.

[11] T. Matsui and H. Satz. "J/ψ Suppression by Quark-Gluon Plasma Formation". *Phys. Lett. B*, 178:416–422, 1986.

[12] Vincenzo Minissale, Salvatore Plumari, and Vincenzo Greco. "Charm Hadrons in pp collisions at LHC energy within a Coalescence plus Fragmentation approach". 12 2020.

[13] Jonghan Park. 'Open charm and beauty measurements from small to large systems". *EPJ Web Conf.*, 259:12010. 4 p, 2022.

[14] Donald Hill Perkins. *"Introduction to high energy physics; 4th ed."*. Cambridge Univ. Press, Cambridge, 2000.

[15] Vladimir Peskov, Micro Planicic, and Francisco García. "Technical Design Report for the Upgrade of the ALICE Time Projection Chamber", 03 2014.

[16] Roberto Preghenella. *"The Time-Of-Flight detector of ALICE at LHC: construction, test and commissioning with cosmic rays"*. PhD thesis, alma, Maggio 2009.

[17] Andrea Rossi. "Talk: Hadron Physics and Non Perturbative QCD 2017 - Pollenzo (CN)", 2015. URL: `https://agenda.infn.it/event/13077/contributions/17482/attachments/12668/14270/QGPexpOverview.pdf`.

[18] Shusu Shi. *Event anisotropy v_2 at STAR*. PhD thesis, Hua-Zhong Normal U., 2010.

[19] Jun Song, Hai-hong Li, and Feng-lan Shao. "New feature of low p_T charm quark hadronization in pp collisions at $\sqrt{s} = 7$ TeV". *Eur. Phys. J. C*, 78(4):344, 2018.

[20] Sheldon Stone. "Pentaquarks and Tetraquarks at LHCb", 2015.

[21] S. Acharya *et al.* "Measurement of D^0, D^+ , D^{*+} and D_s^+ production in pp collisions at $\sqrt{s}$ = 5.02 tev with alice". *The European Physical Journal C*, 79(5), may 2019.

[22] S. Acharya *et al.* "Λ_c^+ production and baryon-to-meson ratios in pp and p-Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV at the LHC". *Physical Review Letters*, 127(20), nov 2021.

[23] 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.*, 128:012001, Jan 2022.

[24] R. L. Workman and Others. "Review of Particle Physics". *PTEP*, 2022:083C01, 2022.

[25] Li Yi. *"Study of quark gluon plasma by particle correlations in heavy ion collisions"*. Springer, 2016.