

Searching for new physics in $c \rightarrow u$ transitions

Author: Marko Stamenkovic
marko.stamenkovic@epfl.ch

Supervisor: Dr. Luca Pescatore
Director: Prof. Dr. Tatsuya Nakada

June 23, 2017

Abstract

Rare decays are processes suppressed in the Standard Model (SM) which happen only at loop level. As new particles could contribute in the loops at the same level as the Standard Model, these decays are very sensitive to new physics. In particular, we consider the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ which is a $c \rightarrow u\ell^+\ell^-$ flavour changing neutral current forbidden at a tree level in the Standard Model. A search for this decay is performed with data collected by the LHCb in 2016, at a center of mass energy of $\sqrt{s} = 13$ TeV corresponding to 1.66 fb^{-1} of integrated luminosity. Two strategies are investigated: promptly produced Λ_c^+ candidates or coming from the semi-leptonic $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-\bar{\nu}_\mu$ decay. Two normalisation channels are also investigated: the $\Lambda_c^+ \rightarrow pK^-\pi^+$ or $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decay. The optimal configuration is found to be the prompt production mode combined with the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ normalisation channel. Using the CL_S method, an expected upper limit is set at

$$\mathcal{B}(\Lambda_c^+ \rightarrow p\mu^+\mu^-) < 9.67 \times 10^{-8}.$$



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Contents

1	Introduction	3
2	Theory	3
2.1	Rare decays	4
2.2	Charm decays	4
3	Motivation	5
4	The LHCb experiment	6
5	Strategy	7
5.1	Available datasets	10
5.2	Definition of signal and sideband regions	10
5.3	Fit model	10
6	Selection	12
6.1	Candidate preselection	12
6.2	Proton Particle identification	13
6.3	Multivariate analysis	15
6.3.1	The boosted decision tree classifier	16
6.3.2	The multilayer perceptron classifier	16
6.3.3	Prompt mode	17
6.3.4	SL mode	19
6.3.5	Optimisation of the parameters	20
6.3.6	Choice of the MVA classifier	21
6.3.7	Optimisation of the MVA output requirement	23
6.4	Hardware and software triggers requirements	24
6.5	Muon Particle Identification	26
6.6	Decision between the prompt and SL modes	26
6.7	Calibration of particle identification	28
6.8	Efficiency of the full selection	29
6.9	Expected number of events	30
6.10	Backgrounds	31
7	Upper limit determination	34
8	Prospect on systematic uncertainties	35
9	Blind fit	36
10	Discussion of the results	37
11	Conclusion	39
12	Appendix	41
12.1	The $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$ contamination in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ sample	41
12.2	Comparison of the signal, sideband and simulation distributions	41
12.3	MVA training correlation matrices	42
12.4	Trigger	43
12.5	PID calibration analysis : PIDCalib, MCResampling and PIDGen	44
	References	48

1 Introduction

The Standard Model of particles physics (SM) is a quantum field theory classifying all observed elementary particles and describing three of the four known fundamental forces in the universe : the electromagnetic, weak and strong interactions. The theory was mainly developed throughout the latter half of the 20th century in a common effort involving many scientists around the world. The current formulation was finalised in the middle of the 1970s with the experimental confirmation of the existence of quarks. Several experimental verifications such as the confirmation of the top quark [1, 2], the tau neutrino ν_τ observation [3] and the Higgs boson discovery [4, 5] assessed the general confidence in the SM and made it one of the most successful theory in modern physics. In addition, important properties of the weak currents, W and Z^0 bosons were predicted with great accuracy. One example is the predicted and measured widths of the W and Z^0 bosons [6] which can be found in Tab. 1.

Table 1: Predicted and measured widths of the W and Z^0 bosons [6].

Width	Predicted [GeV]	Measured [GeV]
Γ_{Z^0}	2.4960 ± 0.0002	2.4952 ± 0.0023
Γ_W	2.0915 ± 0.0005	2.085 ± 0.042

Nevertheless, some phenomena remain unexplained and the SM falls short of being a complete theory of fundamental interactions. It does not fully explain the baryon asymmetry in the universe nor does it explain a theory of gravity. Models as the accelerating universe expansion [7, 8] possibly described by dark energy cannot be incorporated. Furthermore, experimentally established effects such as dark matter or neutrino oscillations [9] and their non-zero masses do not find an explanation within the SM scope. These reasons motivate the search for new physics.

2 Theory

In the SM, flavour changing weak decays, in the quark or lepton sector, are decays involving a transition from an initial flavour state to a different one, e.g. $b \rightarrow c$ transitions. For the quark sector, the strength of flavour changing weak decays is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [10, 11] which has 4 free parameters, 3 mixing angles suggesting that the the mass eigenstates participating in the weak interactions are rotated with respect to the flavour eigenstates and 1 complex CP -violating phase. The unitarity of the matrix imposed by the universality of weak interactions implies couplings of high magnitude between quarks of the same family while transitions between different generations of quarks are expected to be highly suppressed, also referred to *Cabibbo suppressed*.

The present analysis is focused on Flavour Changing Neutral Currents (FCNC) which are a class of processes involving a transition from one flavour state to another while conserving the charge. The decay of interest is the rare decay $\Lambda_c^+ \rightarrow p \mu^+ \mu^-$ which is a $c \rightarrow u \ell^+ \ell^-$ FCNC process, where $\ell = e, \mu$.

2.1 Rare decays

Due to the Glashow-Iliopoulos-Maiani (GIM) mechanism [12], the FCNC processes are forbidden at a tree level but can occur through loop diagrams such as penguin or W box diagrams. Most of the attention on FCNC has been focused on rare processes involving $s \rightarrow d\ell^+\ell^-$ and $b \rightarrow s\ell^+\ell^-$ transitions. The analogous FCNC process $c \rightarrow u\ell^+\ell^-$ in the charm sector has gained less attention and therefore there are plenty of opportunities to analyse this decay. The Feynman diagrams of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay are displayed in Fig. 1. These decays are called *rare decays* because the branching fractions are expected to be small, in the case of the $c \rightarrow u\ell^+\ell^-$ transition, typically $\sim 10^{-9}$ [13, 14, 16] or lower. Additional contributions to the loops from particles beyond the SM are not necessarily suppressed which makes these decays very sensitive to new physics. Due to the virtuality, particles with higher mass scales could also contribute to the loops. As penguin decays are well understood from a theoretical point of view, those processes are particularly interesting because they allow precise comparisons with measurements.

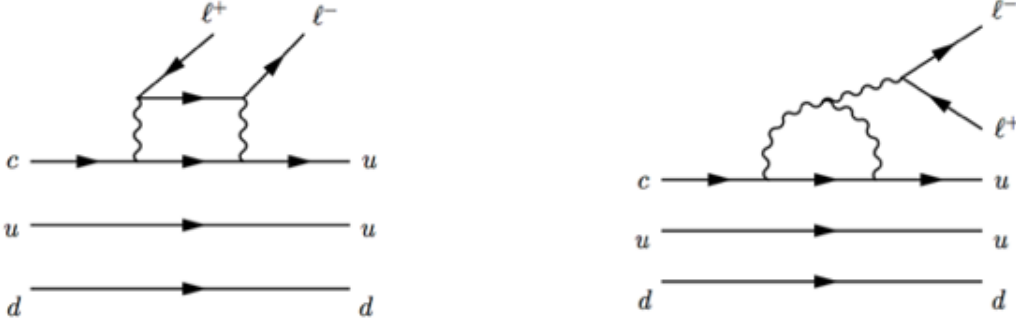


Figure 1: Feynman diagrams for the $\Lambda_c^+ \rightarrow p\ell^+\ell^-$ decay, which is a $c \rightarrow u\ell^+\ell^-$ transition, W -box (left) and penguin (right) diagrams.

2.2 Charm decays

Semilpetonic charm hadron decays induced by $c \rightarrow u\ell^+\ell^-$ transitions are flavour changing processes occurring at low energies, at scales $\mu \ll M_W$ where μ represents the mass energy scale considered in the Lagrangian of the interaction and M_W is the mass of the W^\pm boson. It is then convenient to pass from a full theory $\mathcal{L}_{(\text{full EW} \times \text{QCD})}$ of the electroweak interaction to an effective field theory \mathcal{L}_{eff} by removing the high-energy degrees of freedom, i.e. by integrating out the W^\pm boson and all the particles with a higher mass $m \geq M_W$. The lagrangian density of the effective field theory \mathcal{L}_{eff} is then given by

$$\mathcal{L}_{(\text{full EW} \times \text{QCD})} \rightarrow \mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{QED} \times \text{QCD}} \left(\begin{array}{c} \text{quarks} \\ \& \text{leptons} \end{array} \right) + \sum_n C_n(\mu) Q_n \quad (1)$$

where $\mathcal{L}_{\text{QED} \times \text{QCD}}$ is the lagrangian density of the Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD), $C_n(\mu)$ are the Wilson coefficients describing the short distance effects and Q_n are the local interaction terms

for the long distance interactions. This formulation of the effective field theory can be seen as a modern version of the Fermi theory for weak interactions and the informations on the electroweak scale physics are encoded in the values of the Wilson coefficients. The model is predictive because all the C_n coefficients can be computed. The ability to observe or constrain new physics is closely related to the accuracy of determining the SM contribution. Therefore, a precise evaluation of the Wilson coefficients within the SM is necessary. Further theoretical descriptions of the FCNC processes and the computation of the Wilson coefficients for the rare charm decays can be found in [15], [16] and [17].

3 Motivation

Rare decays and the semileptonic $c \rightarrow u\ell^+\ell^-$ processes offer several observables that can be used to study effects of physics beyond the SM. The most direct one appears in decay rates and would be measured as an excess in the expected branching fractions. Further informations can be found in observables such as the forward-backward asymmetries that characterise the angular distributions of final state particles. Complementary informations can also be found in the observable computed from the ratio between similar decays. It is also interesting to mention that lepton flavour or lepton universality breaking modes would be a direct evidence of new physics and can be tested with the same set of observables.

The present analysis is focused on the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ and aims at measuring its branching fraction. Previous measurement have been conducted: the BaBar collaboration found an excess of $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ candidates with a deviation of 2.6σ [18] from the background-only null hypothesis in 2011. They set an upper limit on the branching fraction $\mathcal{B}(\Lambda_c^+ \rightarrow p\mu^+\mu^-) < 44 \times 10^{-6}$ at 90% confidence level (CL). Furthermore, a blinded analysis performed at the LHCb set an expected limit of $\mathcal{B}(\Lambda_c^+ \rightarrow p\mu^+\mu^-) < 6.71 \times 10^{-8}$ at 95% CL [19]. This previous analysis was performed in the data collected by the LHCb detector at a center of mass energy $\sqrt{s} = 7$ and 8 TeV, corresponding to an integrated luminosity of 3 fb^{-1} . Concerning the present analysis, the integrated luminosity is 1.66 fb^{-1} and the center of mass energy $\sqrt{s} = 13\text{ TeV}$. We expect to reach a sensitivity similar to the one achieved by the LHCb previous measurement. The integrated luminosities \mathcal{L} of the BaBar and LHCb experiments are reported in Tab. 2.

Table 2: Integrated luminosities for the BaBar and LHCb experiments.

Experiment	\mathcal{L}
BaBar 2011 (e^+e^- -collision)	384 fb^{-1}
LHCb 2011	1 fb^{-1}
LHCb 2012	2 fb^{-1}
LHCb 2015	0.279 fb^{-1}
LHCb 2016	1.66 fb^{-1}

4 The LHCb experiment

The LHCb experiment [21] is a heavy flavour physics experiment at the LHC at CERN in Geneva, Switzerland. Its main purpose is to search for indirect evidence of new physics in CP violation and rare decays of beauty and charm hadrons, by looking for the effects of new particles in processes that are predicted in the SM.

Quark mixing in the SM is described by the CKM matrix, which has a single source of CP violation. Since the level of CP violation in weak interactions cannot explain the matter-antimatter asymmetry in the universe, new sources of CP violation beyond the SM are needed. Along with CP violating phenomena, dark matter as well as extensions of the SM are studied by looking for effects of new particles in predicted processes, flavour changing neutral currents are one example. To examine such possibilities, CP violation and rare decays of hadrons containing b and c quarks must be studied with large data samples, using many different decay modes.

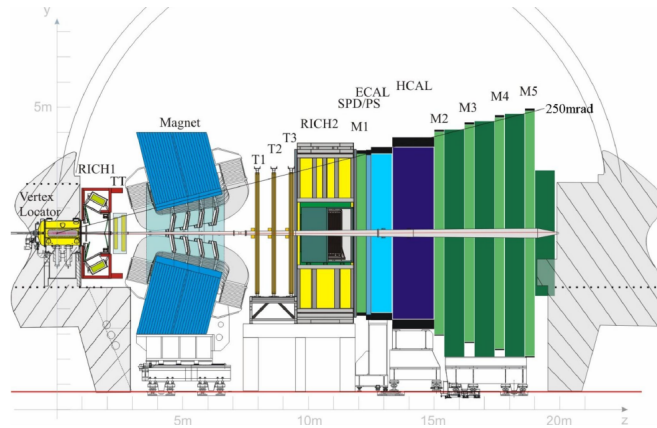


Figure 2: View of the LHCb detector.

The LHCb detector is a single-arm spectrometer [21, 22] covering the pseudorapidity range $2 < \eta < 5$. Figure 2 shows the different sub-detectors, which can be subdivided into five main components.

Tracking system

The high precision tracking system consists of the VERtex LOcator (VELO) [23], situated around the interaction region and containing 42 silicon modules that provide a measurement for the r and ϕ coordinates, a large area silicon-strip detector (TT) [24] located upstream of a dipole magnet with a bending power of 4 Tm, and three stations (T1-T3) [25] of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of the momentum p of charged particles with a relative uncertainty of 0.5% at low momentum up to 1% at 200 GeV/ c . The impact parameter (IP), defined as the minimal distance of a track to a primary vertex (PV), is measured with a resolution of $(15 + 29/p_T)$, where the transverse momentum p_T is in GeV/ c [26].

RICH (Rich Imaging Cherenkov detectors)

Two RICH detectors [27] enable charged particle identification by deducing velocities from the radius of the Cherenkov ring. The RICH system is designed to identify charged hadrons (π , K and p) and to distinguish final states of otherwise identical topology.

Calorimeter system

The calorimeter system is composed of a Scintillating Pad Detector (SPD), a Preshower (PS), a electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL) [28]. This detector is used to provide identifications of electrons, photons and hadrons, as well as their energies and positions. Information from calorimeters also contribute to the Level zero trigger.

Muon detection system

Five stations (M1-M5) of rectangular shape, placed along the beam axis, compose the muon detection system and, together with the calorimeters, contribute to the Level zero trigger of the experiment. Stations M2 to M5 are placed downstream the calorimeters and are interleaved with 80 cm thick iron absorbers. Their information is used to identify and trace penetrating muons both in the online and offline analysis. Station M1 is instead located in front of the calorimeters and is used in the first level trigger [29].

Trigger

The online event selection is performed by a trigger. This consists of a hardware followed by a software stage both taking informations from the calorimeters and muon detection system and, in the case of the software stage, VELO informations. Between the two stages, near real-time alignment and calibration are effectuated and the trigger receives the updated constants. The same alignment and calibration are applied to the offline reconstruction, ensuring a high-quality particles identification (PID) information between the trigger and the offline software. The present analysis exploits candidates reconstructed directly in the trigger.

At the hardware trigger stage, events are required to have a muon with high transverse momentum p_T or a hadron, photon or electron with high transverse energy in the calorimeters. For hadrons, the transverse energy threshold is 3.5 GeV. The software trigger requires a two-, three-, or four-track secondary vertex with a significant displacement from any primary pp interaction vertex. At least one charged particle must have a transverse momentum $p_T > 1.6$ GeV/ c and be inconsistent with originating from a primary vertex. A multivariate algorithm [30] is used for the identification of secondary vertices consistent with the decay of a b or c hadron.

5 Strategy

The purpose of this analysis is to study the sensitivity to the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ in data collected by the LHCb detector in 2016. A blinded analysis is performed to avoid any artificial biases. We explore two production modes, prompt Λ_c^+ production directly related to the primary pp -collision or Λ_c^+ candidates pro-

duced in the semi-leptonic (SL) decay¹ $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$, and two normalisation channels, $\Lambda_c^+ \rightarrow pK^-\pi^+$ or $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decay. The study is organised in two parts. We first determine the best production modes along with the normalisation channels to ensure a quality measurement of the investigated decay. Then, we analyse the various possible background contaminations as well as the sensitivity we can expect within the limits of our selection process. From now on, we refer to the semi-leptonic production of Λ_c^+ candidates as the SL mode.

Production mode

We analyse the prompt production mode in opposition with the Λ_c^+ baryons produced via the SL decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$. The cross section of the c -quark is greater than the cross section of the b -quark and therefore the advantage of the prompt production mode over the SL one is the larger number of events that might enable a more precise measurement. Indeed, for the SL mode, the number of Λ_c^+ candidates is bounded by the number of Λ_b^0 baryons produced in the pp -collision. On the other hand, the SL mode has the advantage to have 4 final particles which are detected as well as two vertices and *de facto* less combinatorial possibilities in the reconstruction of the Λ_c^+ candidate. We therefore expect less combinatorial background in the SL data sample.

Normalisation channel

Two normalisation channels are considered in the analysis: the $\Lambda_c^+ \rightarrow pK^-\pi^+$ and the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decays. One purpose of the study is to determine the most suitable channel for a precise measurement of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$. Table 3 details the different branching ratios of the considered decays in both normalisation channels. The high branching fraction of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay implies a large number of events and increases the precision of the normalised measurement. The direct disadvantage is that the analysed rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ is semi-leptonic while the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel is fully hadronic. Therefore, applying the same selection on both channels is impossible. In particular, the trigger requirements might greatly reduce the number of candidates we can possibly measure, on one channel or the other. On the other hand, not applying the same selection necessarily increases the systematic uncertainties we can expect. For this reason, we also study the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decay as the measured particles are one proton p and two muons $\mu^+\mu^-$. Normalising by this channel is possible through the exact same selection as the one used for the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ but the available number of candidates is greatly reduced by the low branching fraction of the full decay chain $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$. Taking into account the production mode, one main advantage of the SL sample is that the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel also benefits from the extra muon μ^- produced in the primary decay. Therefore, muon-specific requirements might be applied on this fully hadronic channel.

The first step of the analysis is to determine the selection requirements common to all the decays. We determine this part of the selection using the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel because of its high number of reconstructed Λ_c^+ candidates, the presence of a proton in the final state and the low visibility of the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel. The requirement on the proton is determined first. This is done by optimising the significance as a function of the selection on the

¹The data sample of the prompt production mode also contains Λ_c^+ candidates produced via the SL decay Λ_b^0 . Nevertheless, the contribution is assumed to be negligible.

Table 3: Normalisation channels and their branching fractions [6].

Decay	BF
$\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$	0.62 ± 0.27
$\Lambda_c^+ \rightarrow p K^- \pi^+$	$(5.0 \pm 1.3) \times 10^{-2}$
$\Lambda_c^+ \rightarrow p \phi$	$(1.04 \pm 0.21) \times 10^{-3}$
$\phi \rightarrow \mu^+ \mu^-$	$(2.87 \pm 0.19) \times 10^{-4}$
$\Lambda_c^+ \rightarrow p \phi (\rightarrow \mu^+ \mu^-)$	$(2.98 \pm 0.21) \times 10^{-7}$

proton PID variable. Then, to ensure a high combinatorial background rejection, we perform a multivariate data analysis (MVA), performed using the Toolkit for MultiVariate data Analysis (TMVA) [41]. Two classifiers are investigated: the boosted decision tree (BDT) and the multilayer perceptron (MLP). By fitting the sideband of the $\Lambda_c^+ \rightarrow p \mu^+ \mu^-$ data sample, we optimise the Punzi figure of merit [42] in order to determine the optimal requirement to apply on the MVA classifier output variable.

The following step of the analysis is to determine the most suitable production mode and normalisation channel. A study of the trigger requirements gives us further informations on the normalisation channel to use. In addition, we establish the best muon PID requirement by optimising the significance of the $\Lambda_c^+ \rightarrow p \phi (\rightarrow \mu^+ \mu^-)$ decay. By measuring the $\Lambda_c^+ \rightarrow p \phi (\rightarrow \mu^+ \mu^-)$ decay twice, once per production modes (prompt or SL), we are able to determine the optimal conditions to analyse the rare channel $\Lambda_c^+ \rightarrow p \mu^+ \mu^-$.

Once the optimal configuration is determined, we examine the efficiency of the selection on the investigated rare decay and the normalisation channel. We also look for different contamination channels misidentified under the $\Lambda_c^+ \rightarrow p \mu^+ \mu^-$ mass hypothesis and predict the number of contamination events we can expect in our data sample. Finally, we test the sensitivity of our model with the CL_S method [46].

Table 4 details the mass and lifetime τ of the different particles involved in our analysis. The hadrons p , K^- , π^+ and leptons μ^\pm are directly detected as tracks in the detector and allow us to reconstruct the charm decay of the Λ_c^+ baryon.

Table 4: Values of relevant parameters for the Λ_c^+ , ϕ , p , K^- , π^+ and μ^\pm . [6]

Particle	Mass [MeV/ c^2]	τ [s]
Λ_c^+	2286.46 ± 0.14	$(2.00 \pm 0.06) \times 10^{-13}$
ϕ	1019.445 ± 0.020	$(1.55 \pm 0.01) \times 10^{-22}$
p	938.272	Detected particle
K^-	493.611 ± 0.013	Detected particle
π^-	139.57018 ± 0.00035	Detected particle
μ^\pm	105.66	Detected particle

5.1 Available datasets

The rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ is searched in a data sample collected by the LHCb experiment in 2016 at a center-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 1.66 fb^{-1} . Simulation samples were produced for the 2016 signal sample using PYTHIA 8 [31] at the generation stage. Decays of hadronic particles are described using EVTGEN [32] and the interactions of the generated particles with the detector as well as its response are simulated using the GEANT4 toolkit [33] as described in Ref. [34]. The data and signal simulated samples are also available for the normalisation channel $\Lambda_c^+ \rightarrow pK^-\pi^+$. The prompt and SL data samples are available separately. Table 5 shows the number of simulated events for the $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay for the prompt and SL modes. It is important to mention that the simulated sample for the SL mode is the one from 2012 and therefore does not contain the full set of trigger lines of the 2016 exploitation.

Table 5: Number of simulated events for the decay $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow p\mu^+\mu^-$, prompt and SL modes.

Decay	Prompt	SL
$\Lambda_c^+ \rightarrow pK^-\pi^+$	4,015,311	1,043,784
$\Lambda_c^+ \rightarrow p\mu^+\mu^-$	4,023,148	2,016,804

5.2 Definition of signal and sideband regions

In the invariant mass region of the reconstructed Λ_c^+ candidate, we define a window of $\pm 30 \text{ MeV}/c^2$. This window around the nominal Λ_c^+ candidate mass $m_{\Lambda_c^+}$ represents the blind we apply to the analysis. Furthermore, contamination arising from resonant modes are avoided by imposing a veto on the dimuon invariant mass of the candidate for the ϕ , ω and η mesons mass regions. Table 6 shows the various definitions of the blinded region of interest, the vetoes we apply to the rare $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay and the requirement defined to select the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ candidates.

5.3 Fit model

All the signal shapes in the present analysis are determined using an unbinned maximum likelihood fit to the invariant mass distribution of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ candidates in the range $[2250, 2330] \text{ MeV}/c^2$. This range is chosen to fully contain the reconstructed $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ candidates as well as rare candidates from $\Lambda_c^+ \rightarrow p\mu^+\mu^-$.

The signal shapes are described by a distribution referred to as *Hypatia* [35]. The Hypatia probability density function is an extension of the widely used Crystal Ball (CB) [36, 37, 38] function and defined as :

Table 6: Blinded region for the invariant mass $m(p\mu^+\mu^-)$, dimuon invariant mass vetoes on resonant channels as well as control sample signal definition.

Invariant mass $m(p\mu^+\mu^-)$
Low sideband $< 2260 \text{ MeV}/c^2$
$ m(p\mu^+\mu^-) - m_{\Lambda_c^+} < 30 \text{ MeV}/c^2$
High sideband $> 2320 \text{ MeV}/c^2$
Invariant mass $m(\mu^+\mu^-)$
Vetoes on $\Lambda_c^+ \rightarrow p\mu^+\mu^-$
$ m(\mu^+\mu^-) - m_\phi > 40 \text{ MeV}/c^2$
$ m(\mu^+\mu^-) - m_\omega > 40 \text{ MeV}/c^2$
$ m(\mu^+\mu^-) - m_\eta > 40 \text{ MeV}/c^2$
Selection for $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$
$ m(\mu^+\mu^-) - m_\phi < 35 \text{ MeV}/c^2$

$$I(m, \mu, \sigma, \lambda, \zeta, \beta) \propto \begin{cases} G \propto ((m - \mu)^2 + A_\lambda(\zeta)\sigma^2)^{\frac{1}{2}\lambda - \frac{1}{4}} e^{\beta(m - \mu)} K_{\lambda - \frac{1}{2}}(\zeta \sqrt{1 + \left(\frac{(m - \mu)^2}{A_\lambda(\zeta)\sigma^2}\right)}) & , \text{ if } \frac{m - \mu}{\sigma} > -a \\ \frac{G(\mu - a\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{(1 - m/(n \frac{G(\mu - a\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{G'(\mu - a\sigma, \mu, \sigma, \lambda, \zeta, \beta)} - a\sigma))^n} & , \text{ otherwise} \end{cases} \quad (2)$$

where $G(m, \mu, \sigma, \lambda, \zeta, \beta)$ is the generalised hyperbolic distribution described in Ref. [39], G' is its derivative, m is a free variable and corresponds to the measured mass, μ is the mean value, σ the resolution, a is called the transition point, n is the power-law exponent, β controls the tail of the density function, K_λ is the λ th-order Bessel function of third kind and A_λ is a parameter defined as

$$A_\lambda^2 = \frac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)} \quad (3)$$

where ζ is a physical parameter in the range $[0.1, 10]$.

When it comes to fitting a mass peak where energy losses are to be depicted, the Hypatia function preserves the properties of the CB probability density function. In addition, the main advantage over the CB model is that it considers the per-event uncertainties that occur in the process of measuring particles in counting experiments. For these reasons, the Hypatia model is preferable from a practical point of view.

For all the fits, the low and high mass region in the invariant mass $m(p\mu^+\mu^-)$ are mainly combinatorial background, defined as random combination of tracks, and we chose to describe them using a first order Chebyshev polynomial.

For some specific fits, special requirements, *e.g.* fixed parameters or data-driven corrections, are described explicitly when they are used.

6 Selection

6.1 Candidate preselection

A preselection is applied in order to reconstruct the Λ_c^+ candidates. Full list of requirements are shown in Tab. 7. The preselection is built in a way to keep the requirements as similar as possible for the three investigated channels $\Lambda_c^+ \rightarrow p\mu^+\mu^-$, $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$. The candidate selection aims to maximise the yield of $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ signals.

Table 7: Candidate preselection common to all the examined samples.

Particles	Prompt	SL
Λ_c^+	Max DOCA < 0.15 $\cos \theta_D > 0.9999$ $\chi^2/\text{ndf} < 5$ $\chi_{\text{IP}}^2 < 25$	$\sum_{i=1}^3 p_{T_i} > 1800 \text{ MeV}/c$ $\cos \theta_D > 0.99$ $\chi^2/\text{ndf} < 6$ FD $\chi^2 > 100$
Λ_b^0	- -	$\cos \theta_D > 0.999$ $\chi^2/\text{ndf} < 6$
Extra μ	-	DLL $_{\mu\pi} > 0$ $p_T > 800 \text{ MeV}/c$ $p > 3000 \text{ MeV}/c$
p	DLL $_{p\pi} > -5$	DLL $_{p\pi} > 4$ $\chi_{\text{IP}}^2 > 9$ DLL $_{p\pi}$ -DLL $_{K\pi} > 0$ $p > 2000 \text{ MeV}/c^2$
K	DLL $_{K\pi} > -5$	DLL $_{K\pi} > 4$
μ	IsMuon	IsMuon
All daughters	$p_T > 300 \text{ MeV}/c^2$ $p > 2000 \text{ MeV}/c^2$ GhostProb < 0.1 Track $\chi^2/\text{ndf} < 5$ $\chi_{\text{IP}}^2 > 5$	$p_T > 250 \text{ MeV}/c^2$ - GhostProb < 0.5 Track $\chi^2/\text{ndf} < 4$ $\chi_{\text{IP}}^2 > 4$

For the prompt mode, the Λ_c^+ candidate is reconstructed from three charged tracks and is required to have a decay vertex of high quality. The track is considered consistent with the primary vertex PV if χ_{IP}^2 lies below 25, where χ_{IP}^2 is defined as the difference between the vertex-fit χ^2 of a PV reconstructed with and without the track under consideration. The cosine of the angle between the Λ_c^+ candidate momentum p and the primary vertex direction is required to be larger than 0.9999 corresponding to a pointing angle of $\theta_D < 0.8^\circ$. The quality of the Λ_c^+ vertex is assured by a maximum distance between the tracks of the daughter particles, also referred to as the distance of closest approach (MaxDOCA), below 0.15 as well as a vertex-fit requirement $\chi_{\text{vtx}}^2/\text{ndf} < 5$, where ndf is the number of degrees of freedom. All the daughters are also required to have a momentum p larger than 2000 MeV/c, a transverse momentum p_T larger than 300 MeV/c, a track fit χ^2/ndf lower than 5, a χ_{IP}^2 exceeding 5 and a track GhostProb requirement lower than 0.1, where GhostProb is the probability of a track to be identified as a ghost track. In addition, the tracks are required to be identified as proton p or kaon K particles by constraining the difference

of the log-likelihood under different mass hypotheses $DLL_{p\pi}$ and $DLL_{K\pi}$ on the hadronic mode while the muon-specific flag **IsMuon** is required on the dimuon one. The latter being the most constraining, one can expect a difference in the efficiencies.

For the SL mode, the **Max DOCA** and χ^2_{IP} requirements are removed for the Λ_c^+ candidate. Instead, the sum of the momenta for the three decay products is required above a threshold $\sum_{i=1}^3 p_{T_i} > 1800 \text{ MeV}/c$. As the Λ_c^+ decay is not the primary decay, a requirement on the quality of the flight distance-fit $\text{FD}\chi^2 > 100$ is imposed. For the Λ_b^0 candidate, a good decay vertex is obtained through the requirements $\cos\theta_D > 0.999$ and $\chi^2/\text{ndf} < 6$. The requirements on the transverse momentum p_T are the same for all the daughters with the exception of the muon and the extra muon from the primary decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$. Concerning the χ^2_{IP} requirement, the proton selection is different than the one for all the other decay products.

Figure 3 shows the invariant mass distribution $m(pK^-\pi^+)$ fitted for the prompt and SL modes. One can notice that the muon generated in the semi-leptonic $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$ decay gives an extra handle to efficiently reject back-ground during the candidate preselection and significantly improves the purity of the sample.

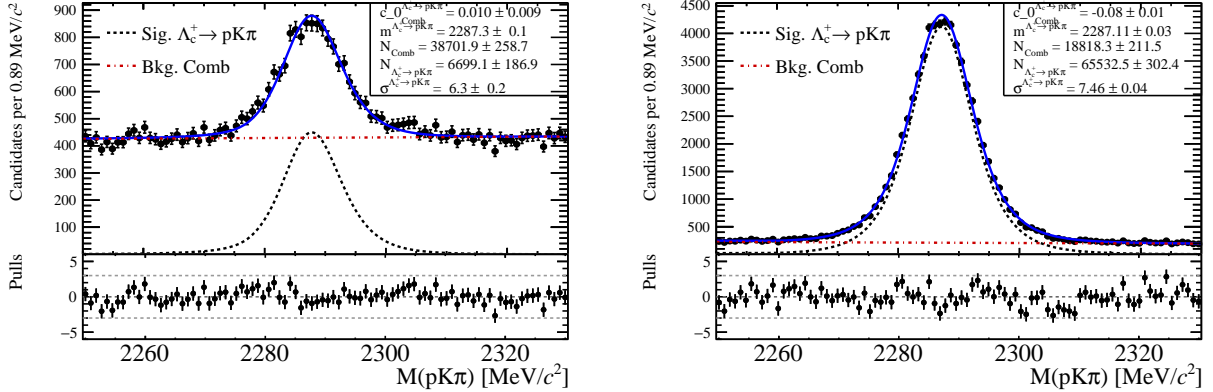


Figure 3: Fit to the $\Lambda_c^+ \rightarrow pK^-\pi^+$ data after the preselection, prompt (left) and SL (right), signal (black) and combinatorial background (red).

6.2 Proton Particle identification

The first requirement of the selection to be optimised is the PID of the proton. The reason is that the PID variables are difficult to simulate² due to the complexity of the measurement process and the lack of a sufficient model. Therefore, in a blinded analysis, one does not have access to the number of signal events that pass the selection and it becomes difficult to find an appropriate figure of merit to maximise. Since for the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel there is a clear distinction between the signal and combinatorial background, we use it to determine the proton PID requirement. The requirement is based on the **ProbNN** variable defined as the output of a MVA classifier combining informations from the RICH detectors, the different tracking systems, the calorimeters and the muon stations.

²The calibration of the simulated PID variables is discussed in details in Sec. 6.7.

The variable represents the probability for a measured particle to be identified as a given particle: K , π , p , μ or e . A significance parameter $\frac{N_{Sig}}{\sqrt{N_{Sig}+N_{Bkg}}}$ is optimised to find the best selection to apply on the proton particle identification **ProbNNp** variable, where N_{Sig} is the number of signal events and N_{Bkg} the number of background events extracted from the fit to the data in a window of $[\mu - 2\sigma, \mu + 2\sigma]$ where μ is the mean of the probability density function and σ the corresponding standard deviation. Since every investigated channel contains a proton, the determined selection is applied on all the available data samples.

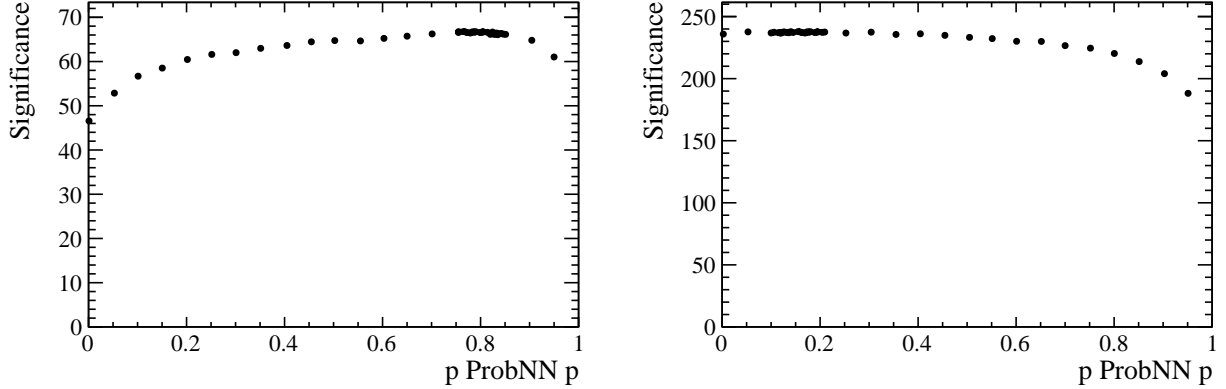


Figure 4: Significance as a function of the **ProbNNp** requirement for the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel, prompt (left) and SL (right) modes.

Figure 4 shows the significance as a function of the lower cut on the proton **ProbNNp** variable. Table 8 lists the optimal selection maximising the significance.

Table 8: Proton **ProbNNp** selection for prompt and SL modes.

Prompt	SL
$p \text{ ProbNNp} > 0.805$	$p \text{ ProbNNp} > 0.155$

A visualisation of the background rejection due to the proton **ProbNNp** requirement can be seen on Fig. 5 where the dashed black line corresponds to the data sample after the preselection, the blue field represents the sideband regions as defined in Tab. 6 and the full red line describes the candidates removed by the selection. The high threshold set for the prompt mode tends to remove a significant number of combinatorial candidates. Concerning the SL mode, the high purity of the signal limits the discrimination power of the optimal requirement. Furthermore, one can notice the rejected candidates peaking around the Λ_c^+ mass. An explanation comes from the intrinsic efficiency of the selection on the signal sample, removing part of real candidates. For the prompt mode, one can notice an excess of removed candidates inducing a slightly asymmetrical behaviour on the left side of the region corresponding to the nominal mass of the Λ_c^+ candidate. A shift to the left side in the reconstructed mass is explainable by a loss of energy in the reconstruction. For this reason, we look for similar decays where the final state particles are reconstructed under the mass hypothesis of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel. One explanation is given by

the $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$ decay where the π^0 is not reconstructed. The branching fraction is $\mathcal{B}(\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0) = (4.9 \pm 0.4) \times 10^{-2}$ which is of the same order as the branching fraction of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel. Most of the events are expected to be suppressed due to the preselection but due to the high statistics of the mode, some of them might still be present. For this reason, a strong requirement on the proton **ProbNNp** variable might indeed remove $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$ candidates. A brief analysis leading to this conclusion is reported in App. 12.1.

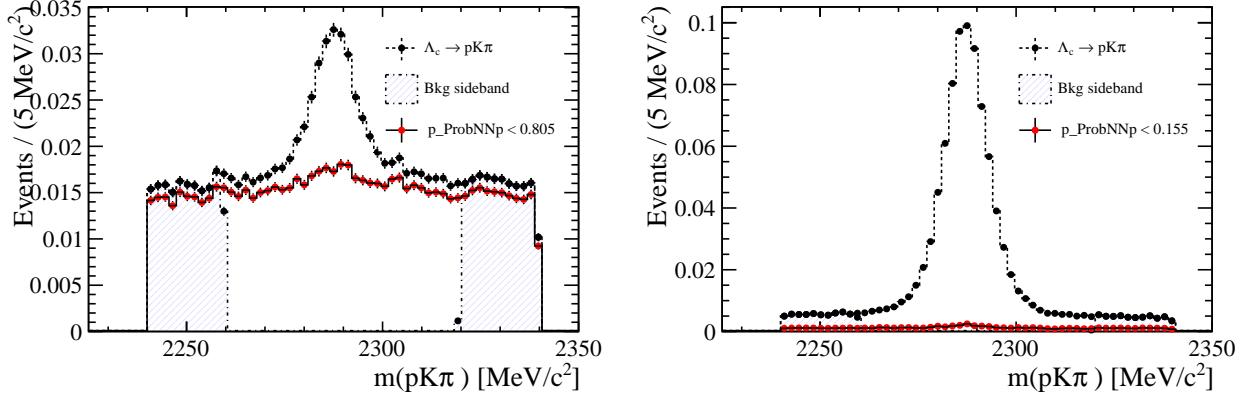


Figure 5: Mass distribution $m(pK^-\pi^+)$ for the $\Lambda_c^+ \rightarrow pK^-\pi^+$ data sample : after the preselection (black dashed), removed by the optimal proton PID requirement (red) and data sideband (blue field) for the prompt (left) and SL (right) modes.

Figure 6 shows the fit to the data after the proton selection. For the prompt mode, one can notice the high rejection of the combinatorial background events.

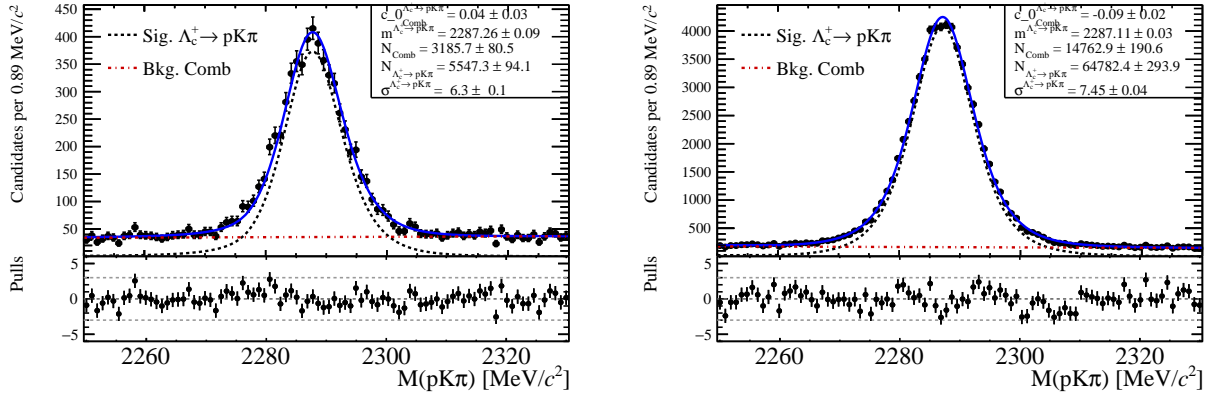


Figure 6: Fit to the $\Lambda_c^+ \rightarrow pK^-\pi^+$ data with proton p **ProbNNp** selection applied for the prompt (left) and SL (right) modes.

6.3 Multivariate analysis

To ensure the rejection of the combinatorial background, the MVA combines informations from low correlated variables exploiting them at best. The purpose is to classify a given event as signal or background. A solution is to train a classifier on given signal and background samples where the nature of the event is known

a priori. Two different classifiers are considered: the BDT and the MLP. An optimisation is performed on three different parameters for each methods and the best resulting configurations are then compared. The signal sample is provided by the simulation while the background is extracted from the sideband of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ invariant mass. The two samples are split in two halves which are used separately for the training and testing of the classifiers. We first present the two different classifiers, the different input variables, the resulting outputs after an optimisation of relevant parameters as well as their respective trainings and testings, separately for the prompt and SL modes, then we determine the most suitable classifier to apply to both modes by comparing the *receiver operating characteristics* (ROC) curves and finally we optimise the selection with respect to the Punzi figure of merit [42].

6.3.1 The boosted decision tree classifier

A *decision tree* [41] is a binary structure where each nodes corresponds to a variable with a requirement that can be fulfilled or not. Repeated decisions are taken on a single variable at a time until a stop-criterium is satisfied. This way, the data sample is split into many regions and depending on the majority of training events that end up in the final *leaf* of a node, the tree is able to classify events as signal or background. Boosting algorithms extend the concept by sequentially applying a decision tree to a reweighted version of the training sample. Each tree is computed from the same training ensemble and events misclassified during the training of a given tree are attributed a higher weight in the training of the following tree. The decision from several trees are then averaged into a single classifier leading to a clear enhancement of performances in most of the cases. The ranking of the input variables is ensured by counting how often the variable is used to split the nodes and the number of events present in the corresponding *leaf*.

6.3.2 The multilayer perceptron classifier

The MLP classifier [41] is a set of neurons constituted by a combination of variables where each neuron has an individual response at a given set of input signals. The neurons are organised into layers and, by applying an external signal, the network is put into a final state representing a classification of the event as signal or background. During the training phase, the *back propagation* [41] algorithm determines the neurons as well as their respective weights in order to achieve the given final states. The variable ranking is achieved by computing the importance I_i of the input variable i following the formula

$$I_i = \bar{x}_i^2 \sum_{j=1}^{n_h} (w_{ij}^{(1)})^2 \quad (4)$$

where \bar{x}_i is the sample mean of the input variable i , $w_{ij}^{(1)}$ is the weight of the path from the variable i in the input layer to the variable j in the first hidden layer and n_h is the number of variables in the first hidden layer.

6.3.3 Prompt mode

Input variables

Figure 7 displays the distributions of the 12 different variables used to train the two MVA classifiers. The discriminating variables are determined by comparing the data signal, data sideband and the simulated distributions of the different variables. A brief description of the comparison method can be found in App. 12.2. The variables are mainly focusing on the Λ_c^+ vertex and quality variables such as the vertex-fit $\chi_{\text{vtx}}^2/\text{ndf}$, the lifetime τ , transverse momentum p_T , the $\cos\theta_D$, where θ_D is the pointing angle, the χ^2 of the kinematic fit referred to as the Decay Tree Fitter (DTF) [43], the impact parameter (IP) as well as the maximum distance of closest approach (MaxDOCA). The only discriminating variable specific to the final-state particles is the maximum of the three tracks **GhostProb**. Because of the high number of Λ_c^+ candidates and the fact that no PID variable are used as an input to the classifier, the MVA is trained on the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel but is applied on all the data samples. Therefore, we use only one variable specific to the daughters in order to avoid biases on final states requirements. A further information is provided by the isolation variables A_{p_T} corresponding to the imbalance of transverse momentum p_T of nearby tracks compared to the one of the candidate:

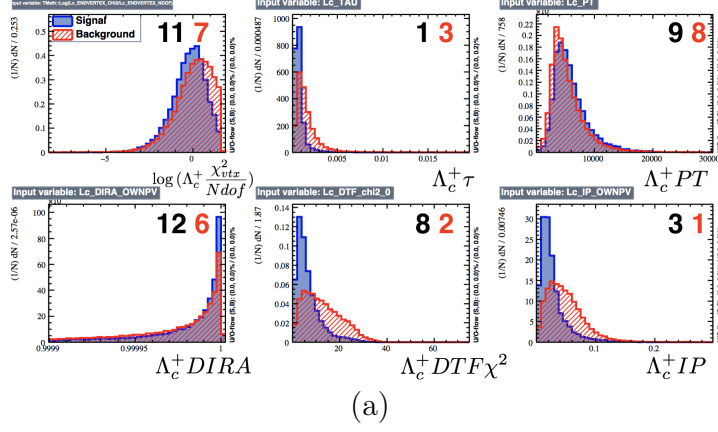
$$A_{p_T} = \frac{p_T(\Lambda_c^+) - (\sum \vec{p})_T}{p_T(\Lambda_c^+) + (\sum \vec{p})_T} \quad (5)$$

where $(\sum \vec{p})_T$ is the transverse component of the vector sum momenta of all charged particles within a cone around the candidate, excluding the signal track of the particle under consideration (p, K, π or μ). The cone radius is defined in the plane of pseudo-rapidity and azimuthal angle (η, ϕ) as $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Another isolation variable computed is the $\Lambda_c^+ \chi_{\text{vtx}}^2$ **mass 2 tracks**. The variable is computed by matching every possible pair of tracks, excepted for the signal tracks of the decay product, and fitting the vertex of the Λ_c^+ candidate again. For the vertex-fit with the smallest χ^2 , the mass of the Λ_c^+ particle is computed for each event.

Table 9: Ranking of the MVA variables for BDT and MLP, prompt.

Ranking	BDT	MLP
1	$\Lambda_c^+ \tau$	$\Lambda_c^+ \text{IP}$
2	$p \ A_{p_T}$	$\Lambda_c^+ \text{DTF} \chi^2$
3	$\Lambda_c^+ \text{IP}$	$\Lambda_c^+ \tau$
4	$\pi^+ \ A_{p_T}$	$\Lambda_c^+ \chi_{\text{vtx}}^2$ mass 2 tracks
5	$\Lambda_c^+ \chi_{\text{vtx}}^2$ mass 2 tracks	Max GhostProb
6	Max GhostProb	$\Lambda_c^+ \cos\theta_D$
7	$K^- \ A_{p_T}$	$\Lambda_c^+ \log \chi_{\text{vtx}}^2/\text{ndf}$
8	$\Lambda_c^+ \text{DTF} \chi^2$	$\Lambda_c^+ \text{PT}$
9	$\Lambda_c^+ \text{PT}$	$\Lambda_c^+ \log \text{MaxDOCA}$
10	$\Lambda_c^+ \log \text{MaxDOCA}$	$p \ A_{p_T}$
11	$\Lambda_c^+ \log \chi_{\text{vtx}}^2/\text{ndf}$	$\pi^+ \ A_{p_T}$
12	$\Lambda_c^+ \cos\theta_D$	$K^- \ A_{p_T}$

MVA Variables: Ranking BDT MLP



MVA Variables: Ranking BDT MLP

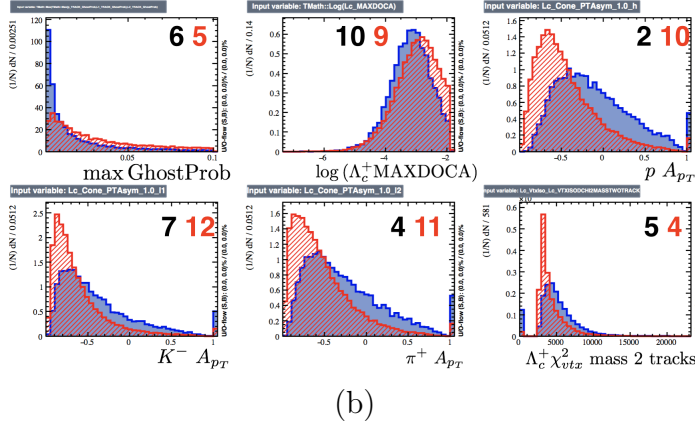


Figure 7: Set of variables used as input to the MVA classifiers for the prompt mode, signal (blue) and background (red), variables ranking for BDT and MLP.

In Fig. 7 and Tab. 9, a comparison of the considered variables is drawn for the simulation and background obtained from the sideband of the data sample as well as the ranking of the variables within the two analysed MVA methods: BDT and MLP. From the correlation matrices (see App. 12, Fig. 24), one can deduce three sets of correlated variables :

- 1) the three transverse momenta asymmetries A_{pT}
- 2) Λ_c^+ MaxDOCA and $\Lambda_c^+ \log \chi_{vtx}^2 / ndf$
- 3) Λ_c^+ DTF χ^2 and Λ_c^+ IP.

The ranking of the BDT classifier can potentially be understood by looking at the correlated variables. If two correlated variables are spanning the same tree, a strong selection on the first node reduces significantly the discrimination power of the following node, leading to a smaller number of events in the final *leaf* and *de facto* a lower ranking. This hypothesis might partly explain the positions of the asymmetrical variables A_{pT} (2, 4 and 7) as well as the Λ_c^+ IP and Λ_c^+ DTF χ^2 variables (3 and 8).

In the case of the MLP classifier, the neurons are stimulated by a combination of input variables and therefore one correlated variable can hardly affect the ranking of the others. This consideration partly explains the ranking of the Λ_c^+

IP and Λ_c^+ DTF χ^2 (1 and 2) as they seem to be among the most discriminating variables in our input set.

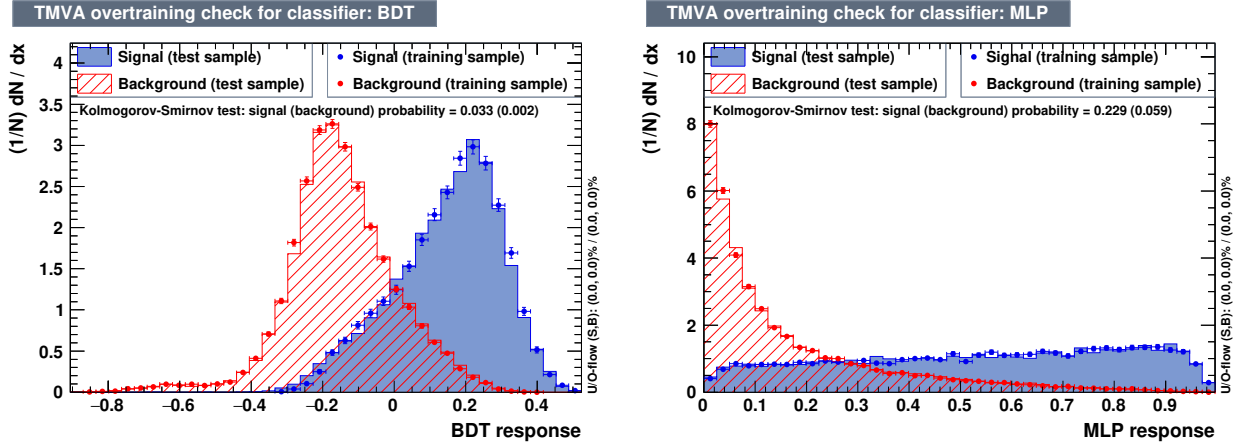


Figure 8: Overtraining test for the MVA classifiers BDT (left) and MLP (right), prompt.

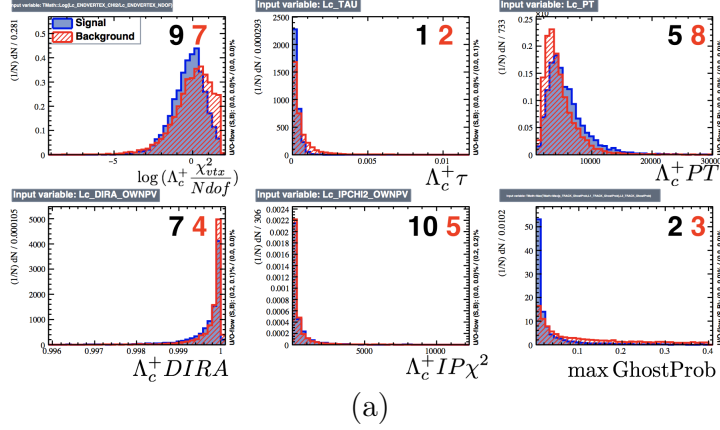
Once the input variables defined, the MVA methods are trained by classifying given events in the final state, signal or background. By knowing the true nature of the event, weights are determined within the system depending on its ability to correctly identify signal from background events. Due to the limited size of the simulation sample after the preselection ($\sim 30'000$ events), one has to pay a particular attention to the overtraining of the classifiers. Overtraining occurs when a machine learning problem has too few degrees of freedom because most of the parameters of the algorithm were adjusted on a small number of data points. Its manifestation usually translates in an increased performance of the classifier over an objectively achievable one. In other words, the classifier performs well only on the sample it was trained on and cannot be used to classify different samples. Heuristically, the BDT classifier tends to be more sensitive to the overtraining effects. It is partly explained by the large number of nodes involved in the technique. For this reason, a testing sample is used to assess the performance of the classifier by comparing the results with the training sample. Figure 8 displays the training and testing comparison for the output variable of the two classifiers: BDT and MLP. The main conclusion of the figure is that the overtraining is not observed and the classifiers can be applied on the various samples.

6.3.4 SL mode

The same procedure is applied to the SL mode. The main difference lies in the fact that the muon μ produced in the semi-leptonic decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$ enables to use extra informations to reject combinatorial background. Therefore, we remove the most correlated variables with respect to the prompt mode and add three new variables : $\Lambda_b^0 \chi_{\text{vtx}}^2/\text{ndf}$, μ track **GhostProb** and χ^2/ndf .

The variables and their respective rankings are shown in Fig. 9 and Tab. 10. The Λ_c^+ DTF χ^2 and Λ_c^+ IP variables were replaced by the Λ_c^+ χ_{IP}^2 . The comparison between the different rankings of the prompt and SL mode suggests that the variable $\Lambda_c^+ \tau$ and **GhostProb** variables offer the greatest discrimination power between signal and background events.

MVA Variables: Ranking **BDT** **MLP**



MVA Variables: Ranking **BDT** **MLP**

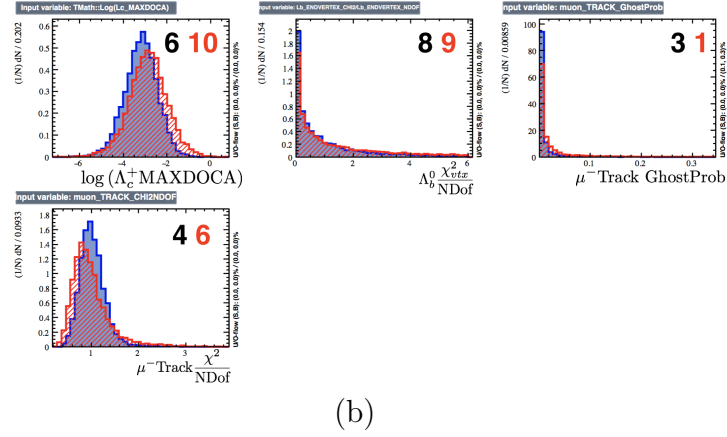


Figure 9: Set of variables used as input to the MVA classifiers : SL mode, variables ranking : BDT and MLP.

In addition, Fig. 10 compares the distributions of the MVA classifiers output variables for the training and testing samples. The classifiers are also not overtrained.

6.3.5 Optimisation of the parameters

The best configuration of the two classifiers presented in Figs. 8 and 10 is found through a three dimensional optimisation of their parameters. The parameters to optimise are advised by the TMVA manual [41]. For the BDT, we simultaneously optimise the maximal depth of the decision tree (**MaxDepth**), the minimal percentage of training events required in a *leaf* node (**MinNodeSize**) and the learning rate β for the adaptative boosting algorithm (**AdaBoost**). For the MLP classifier, the number of training cycles (**Epochs**), the number of layers in the hidden layer architecture (**HiddenLayers**) and the overtraining testing rate (**TestRate**). The optimisation is performed by computing the classifiers for each configurations and comparing the resulting ROC curves. The values of the best parameters for the BDT and MLP classifiers for the prompt and SL modes can be found in Tab. 11.

Table 10: Ranking of the MVA variables for BDT and MLP, SL production mode.

Ranking	BDT	MLP
1	$\Lambda_c^+ \tau$	$\mu^- \text{Track GhostProb}$
2	Max GhostProb	$\Lambda_c^+ \tau$
3	$\mu^- \text{Track GhostProb}$	Max GhostProb
4	$\mu^- \text{Track } \chi^2/\text{ndf}$	$\Lambda_c^+ \chi_{\text{IP}}^2$
5	$\Lambda_c^+ \text{PT}$	Max GhostProb
6	$\Lambda_c^+ \text{MaxDOCA}$	$\mu^- \text{Track } \chi^2/\text{ndf}$
7	$\Lambda_c^+ \cos \theta_D$	$\log \Lambda_c^+ \chi_{\text{vtx}}^2/\text{ndf}$
8	$\Lambda_b^0 \chi^2/\text{ndf}$	$\Lambda_c^+ \text{PT}$
9	$\log \Lambda_c^+ \chi_{\text{vtx}}^2/\text{ndf}$	$\Lambda_b^0 \chi^2/\text{ndf}$
10	$\Lambda_c^+ \chi_{\text{IP}}^2$	$\Lambda_c^+ \text{MaxDOCA}$

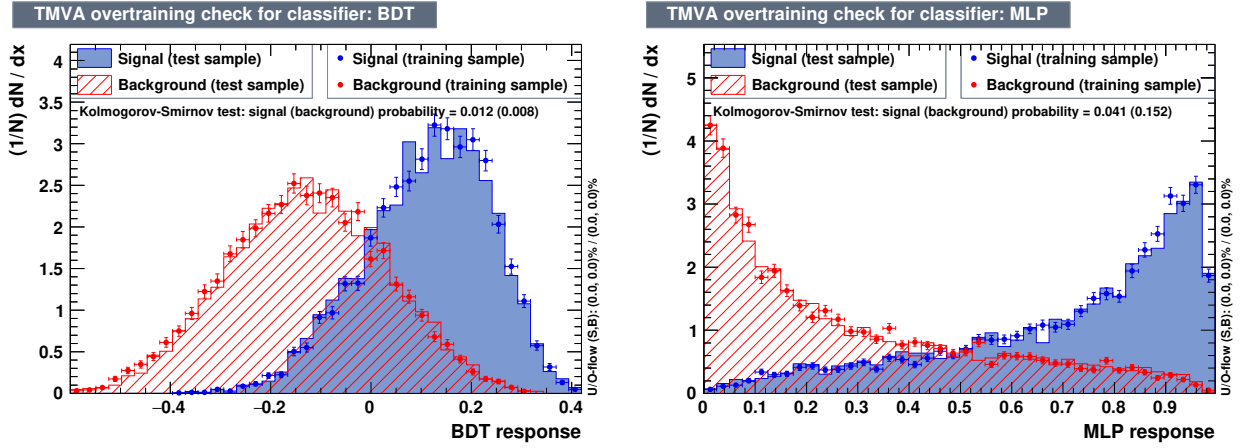


Figure 10: Overtraining test for the MVA classifiers BDT (left) and MLP (right), SL production mode.

6.3.6 Choice of the MVA classifier

After training the classifiers for both production modes, their ROC curves are compared. The discrimination power of a variable can be determined by computing the rejection of the background as a function of the signal efficiency. The ideal scenario corresponds to a unity background rejection independently of the signal efficiency. In this case, the distribution is a step function and the integral of the distribution is unitary. A qualitative comparison of two variables can be achieved by integrating the respective ROC curves.

Figure 11 compares the different ROC curves for the BDT and MLP classifiers for the prompt and the SL mode. Table 12 summarises the values of the ROC curve integral on the domain $[0,1]$. The BDT classifier performs better in both cases and we chose to apply it on all the signal and data samples available.

In order to avoid biases, the linear correlation between the MVA output variable BDT and the invariant mass distribution must be verified. It is therefore paramount to confirm that the classifier is not correlated to the invariant mass distribution in order to preserve the distribution of the signal candidates we expect to see in the rare channel $\Lambda_c^+ \rightarrow p\mu^+\mu^-$. In addition, if too many informations are provided to the classifier, it may be able to infer the invariant

Table 11: Optimal parameters for the MVA the BDT and MLP classifiers, for prompt and SL modes.

Classifier	Parameters	Prompt	SL
BDT	Maximal depth	4	5
	Minimal node size	0.08	0.07
	Adaboost β	0.2	0.2
MLP	Epochs	500	500
	Hidden layers	9	11
	Testing rate	5	5

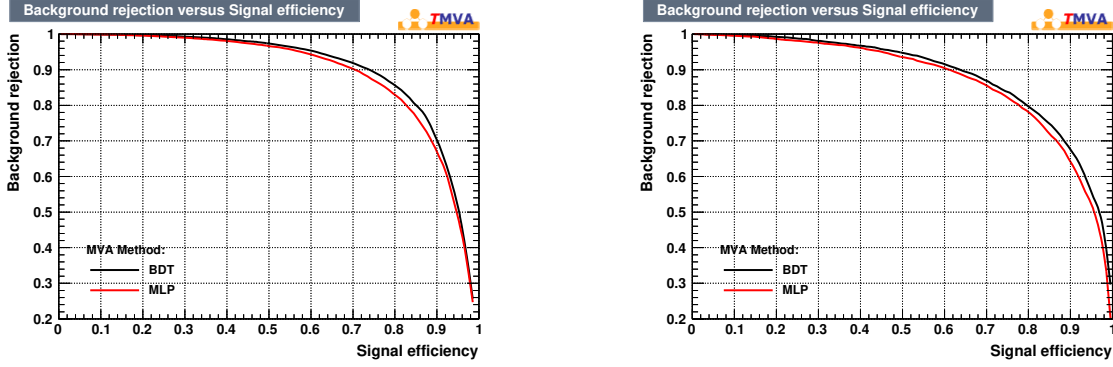


Figure 11: ROC curves for the different MVA classifiers, prompt (left) and SL (right) modes.

mass. Therefore, the MVA output might create artificial peaks. A simple test is given by the linear correlation between the MVA classifier and the invariant mass. This evaluation can be found in Fig. 12. The black dots represent the correlation profile between the mean value of the BDT output variable and the invariant mass distribution of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel while the grey dots represent the distribution of events as a function of the two variables. The flatness of the correlation profile indicates a low dependency between the two variables of interest.

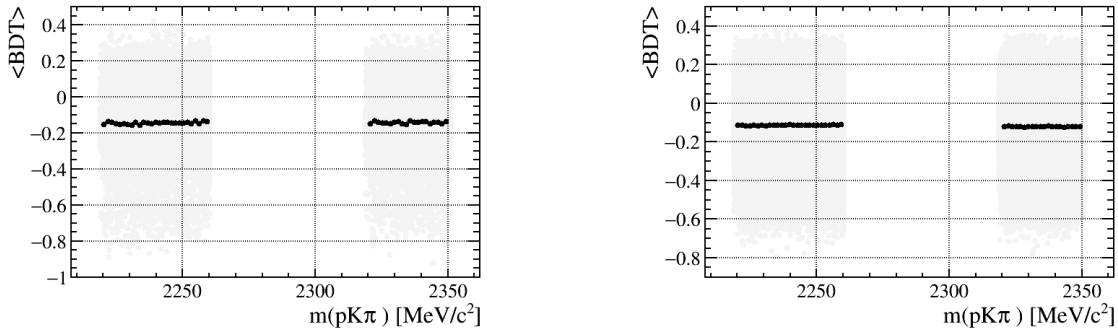


Figure 12: BDT output variable as a function of the invariant mass reconstruction for the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel, distribution (grey) and correlation profile (black), prompt (left) and SL (right).

Table 12: Integral of the ROC curves on the domain $[0,1]$ for the BDT and MLP classifiers, prompt and SL mode.

Classifiers	Prompt	SL
BDT	0.903	0.881
MLP	0.893	0.864

6.3.7 Optimisation of the MVA output requirement

So far, the main requirements were optimised using the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel because of its high statistics, the clear distinction between the signal and the background candidates and the low visibility of the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decay. The further requirements are found using the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ data sample.

Since the signal is not observed, we cannot optimise the selection on the BDT classifier using the significance as previously done for the proton **ProbNNp** requirement. Therefore, we use the Punzi figure of merit [42] which is optimal for optimisations aiming at a discovery.

The figure of merit is defined as

$$\frac{\varepsilon(t)}{a/2 + \sqrt{B(t)}} \quad (6)$$

where $\varepsilon(t)$ represents the efficiency of the selection t on the signal simulation sample, $a = 5$ is a constant and $B(t)$ is the measured number of combinatorial background candidates obtained from the data sample extrapolating in the signal region of interest after finding the probability density function with a fit to the sideband data. The main advantage of this figure of merit over the significance is that no prior knowledge on the cross section is required since the expression is linear in the number of signal events. Consequently, the criterium is unambiguously preferable from a practical point of view.

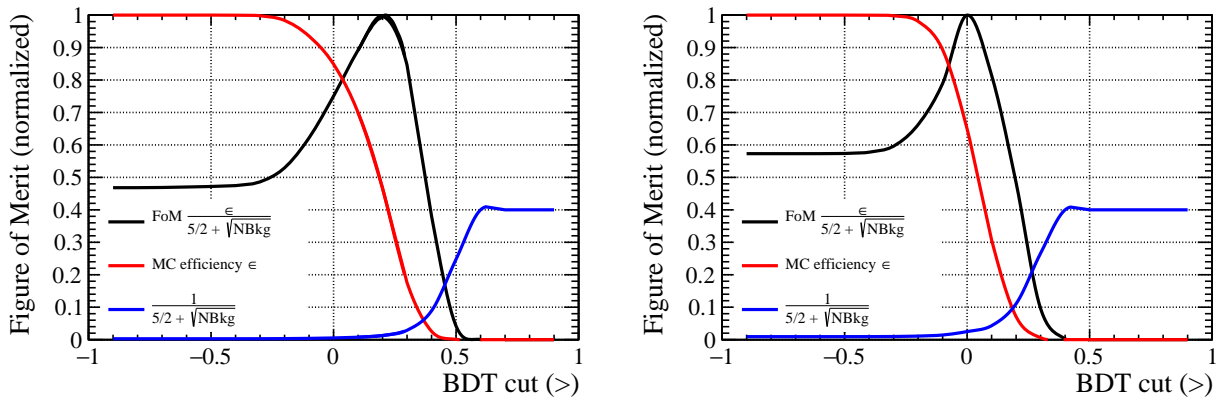


Figure 13: The Punzi figure of merit (solid black line), signal simulation efficiency (red) and denominator dependent on the extrapolated number of background events N_{Bkg} (blue) for the prompt (left) and SL (right) modes.

Figure 13 shows the figure of merit as a function of the BDT output variable. By varying the selection on the BDT classifier, we compute the efficiency on the

signal simulation sample (red line) as well as the denominator of Eq. 6 (blue line) by fitting the sideband region in the data sample and extrapolating the expected number of background events in the region of interest. The solid black line represents the figure of merit as the multiplication of the two former. The optimal requirements are presented in Tab. 13.

Table 13: Selection on the BDT classifier variable

Prompt	SL
BDT > 0.21	BDT > 0.0

6.4 Hardware and software triggers requirements

Analysing the hardware and software triggers specifications is required to determine the production mode with respect to the normalisation channel. To reduce systematic uncertainties of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$, one must guarantee a selection as similar as possible on the normalisation channel. Since the $\Lambda_c^+ \rightarrow pK^-\pi^+$ decay is a fully hadronic decay, a muon-specific trigger selection would lower the available signal yield of the control sample and therefore reduce significantly the precision of the measurement. On the other hand, the SL mode offers the advantage that a muon μ is also present in the primary decay of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel, and triggers the muon-specific requirements which would have a good efficiency on both channels.

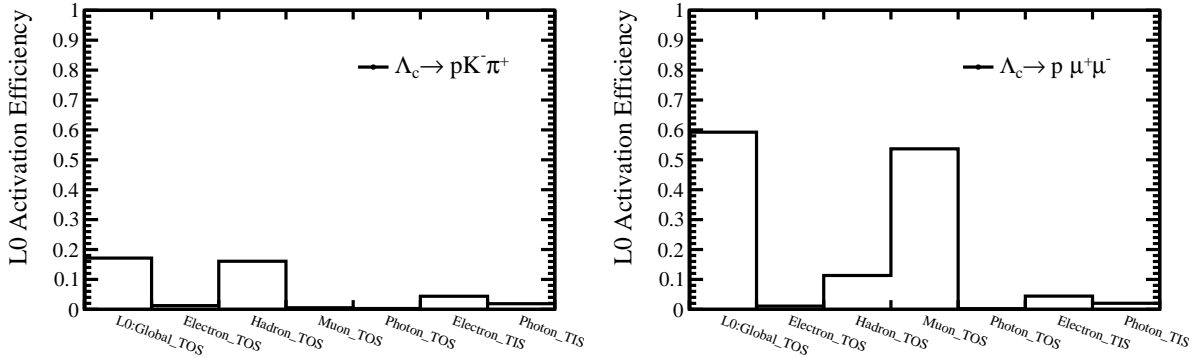


Figure 14: Hardware (L0) triggers for $\Lambda_c^+ \rightarrow pK^-\pi^+$ (left) and $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ (right) simulation samples for prompt mode.

For the prompt mode, Fig. 14 displays the fraction of events that activated the various hardware (L0) trigger for the $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ channels. As one can expect, less than $\sim 1\%$ of events activated the hardware and software muon-specific triggers on the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel. Using the hadron-specific hardware trigger would be a viable solution as both decays contain a proton but then we would remove $\sim 90\%$ of events on the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay, which we want to analyse. This is too constraining to measure a rare decay. For this reason, we exclude the possibility of using the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel as a normalisation channel if the prompt production mode is chosen. Concerning

the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ normalisation channel, the $\sim 60\%$ activation of the hardware trigger enables the measurement of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ within the prompt data sample.

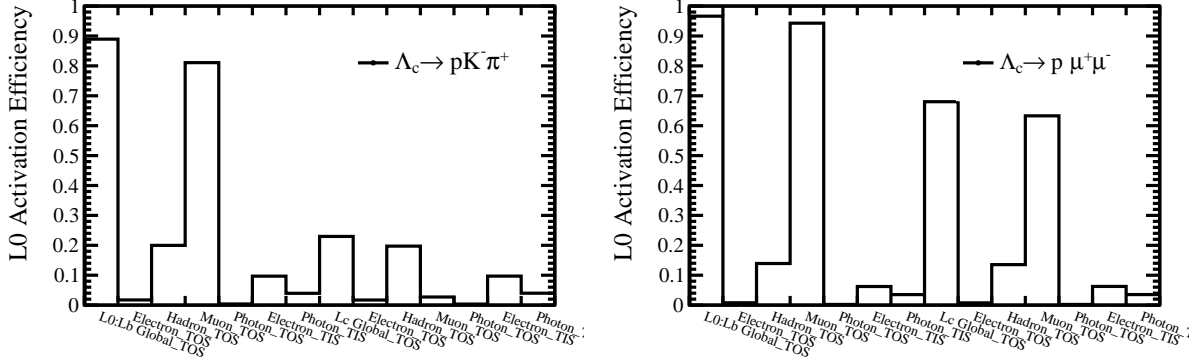


Figure 15: Hardware (L0) triggers for $\Lambda_c^+ \rightarrow pK^-\pi^+$ (left) and $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ (right) simulation samples for the SL mode.

Figure 15 displays the same requirements for the SL mode but with the considerations on the Λ_c^+ and Λ_b^0 baryons. The $\sim 80\%$ activation of the Λ_b^0 muon-specific trigger on the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel enables the possibility to use the hadronic mode as a normalisation for the rare decay.

Table 14: Full trigger requirements for the analysis of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$, prompt mode associated to the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel, SL mode associated to the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel.

Normalisation channel	Trigger requirements
$\Lambda_c^+ \rightarrow pK^-\pi^+$ (SL)	L0: Λ_b^0 Muon Decision Hlt1: Λ_c^+ Track All L0 <i>or</i> Track Muon Hlt2: Λ_c^+ Phys
$\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ (Prompt)	L0: Λ_c^+ Muon Decision Hlt1: Λ_c^+ Track Muon <i>or</i> MVA <i>or</i> MVA Muon <i>or</i> Two Track MVA Hlt2: Λ_c^+ Phys

Table 14 details the final trigger requirements applied on the simulation and data samples in order to analyse the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$. For the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ normalisation channel, we examine the software trigger HLT1 and its muon-specific requirements using the prompt mode simulation sample. We combine the requirements with an efficiency greater than 20% with a *or* logical statement. For the $\Lambda_c^+ \rightarrow pK^-\pi^+$ normalisation channel, it is important to note that the HLT1 trigger for the SL mode is a specific trigger line of the 2012 simulation, for this reason, an extended analysis is required on the 2016 samples. Concerning the HLT2 software trigger, the specific triggers (without any prior requirement applied, see App. 12.4 Tab. 25) are unusually low. For this reason, we choose to use a preliminary requirement Λ_c^+ Phys which corresponds to a logical *or* statement of all the trigger requirements. Nevertheless, a more

detailed analysis is needed to determine the specific requirements pertinent for the analysis.

6.5 Muon Particle Identification

The selection is finalised with a requirement on the muons μ^\pm **ProbNNmu** variables. Due to physical considerations, we assume that a selection optimising the significance of the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel, which is well visible after the MVA requirement, should maximise the chance of observing the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$. For the prompt mode, two muons are present in the products of the decay and the requirements are optimised simultaneously. For the SL mode, the optimisation is performed on the two products at the same time but a further requirement on the muon produced in the primary decay $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-\bar{\nu}_\mu$ is considered.

SL mode: $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-\bar{\nu}_\mu$ muon optimisation

The optimal requirement on the extra muon from the primary semi-leptonic $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-\bar{\nu}_\mu$ decay is determined by optimising the significance on the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel. The same procedure as for the proton optimisation is applied and the final requirement added to the selection is $\mu^- \text{ProbNNmu} > 0.8$.

Prompt and SL mode comparison

The requirements on the two muons μ^\pm produced in the Λ_c^+ decay are also optimised using the significance criteria on the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel. The optimisation is processed by setting an upper limit on the μ **ProbNNmu** requirement and computing the resulting significance. The optimal requirement for the prompt and SL modes can be found in Tab. 15 along with the resulting parameters of interest.

Table 15: Results of the fit to the invariant mass distribution of the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decay, prompt and SL modes.

$\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$	Prompt	SL
Muon particle identification	$\mu \text{ ProbNNmu} > 0.225$	$\mu \text{ ProbNNmu} > 0.005$
Mean value μ	$(2288.4 \pm 0.8) \text{ MeV}/c^2$	$(2279 \pm 3) \text{ MeV}/c^2$
Standard deviation σ	$(9 \pm 1) \text{ MeV}/c^2$	$9 \text{ MeV}/c^2$ (fixed)
Number of signal events N_{Sig}	127 ± 13	30 ± 7
Number of background events N_{Bkg}	44 ± 12	35 ± 10
Significance	9.7 ± 0.9	3.7 ± 0.7

6.6 Decision between the prompt and SL modes

The optimal production mode is determined by comparing the different significances with the full selection applied. The fit is performed by first computing the parameters of the fit model with the help of the simulated sample for the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay. The result is showed in Fig. 16 while Fig. 17 displays the fit to the data for the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel for the prompt and SL modes. The fit model is Hypatia as described in Sec. 5.3 and the different parameters are fixed from a prior fit to the simulation data sample. The signal and background

yields are extracted from the resulting probability density function. The prompt mode contains a larger number of signal events than the SL one. This can be explained by the larger cross section of the c -quark with respect to the b -quark $\sigma(c\bar{c}) \sim 20 \times \sigma(b\bar{b})$ and the supplementary decay of the Λ_b^0 candidate in the SL mode. While trying to fit the SL mode invariant mass, the standard deviation, when let free to vary, was twice as big as the standard deviation of the fit for the prompt mode. The standard deviation is closely related to the resolution of the detector. As the experiment is the same and the kinematics of the Λ_c^+ candidates are expected to be similar for the two production modes, there is no obvious reason to consider an excessively different standard deviation. For this reason, in the SL mode, we fix the standard deviation $\sigma_{\text{SL}}^{\text{DATA}}$ from the fit to the simulation sample and correct it with the ratio of standard deviations obtained in the fit to the simulation and data sample in the prompt mode. This data-driven correction is given by

$$\sigma_{\text{SL}}^{\text{DATA}} = \sigma_{\text{SL}}^{\text{MC}} \times \frac{\sigma_{\text{Prompt}}^{\text{DATA}}}{\sigma_{\text{Prompt}}^{\text{MC}}} \quad (7)$$

where the label MC represents the Monte-Carlo generated simulation samples.

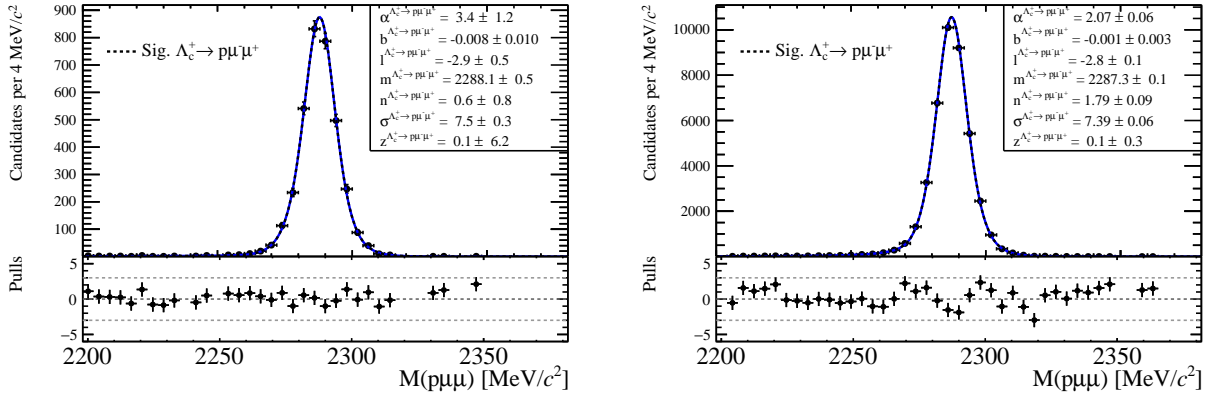


Figure 16: Fit to the simulated sample for the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ channel, prompt (left) and SL (right) modes.

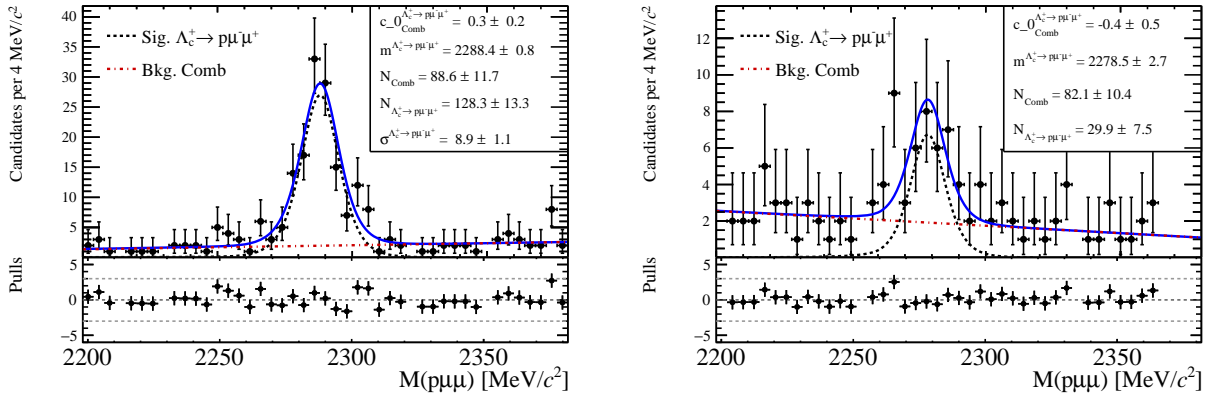


Figure 17: Fit to the data for the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel, prompt (left) and SL (right) modes.

To determine which production mode is the most suitable for the analysis of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ we compare the significance. The branching fraction of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ is expected to be at least two orders of magnitude lower than the one of the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel. A high number of $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ signal events should therefore maximise our chance to measure the rare decay or set a low value on the expected upper limit. Furthermore, the sensitivity to a signal of the rare mode also depends on the level of background and a higher significance of the peak is an indication of a better signal to background ratio. Since the significance of the peak for the prompt mode is approximatively three times larger, the analysis of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ rare channel should be performed with Λ_c^+ candidates promptly produced in the pp -collisions. Following the analysis of the trigger requirements, we deduce that the normalisation channel to use is the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decay.

Table 16: Full selection, the software trigger (Hlt1) requirements are separated by a *or* logical statement.

$\Lambda_c^+ \rightarrow p\mu^+\mu^-$	Prompt	SL
$p \text{ ProbNNp} >$	0.805	0.155
BDT $>$	0.21	0.0
Hardware trigger (L0)	Λ_c^+ Muon Decision	
Software trigger (Hlt1)	Λ_c^+ Track Muon	
(separated by a <i>or</i> logical statement)	MVA	
	Muon MVA	
	Two Track MVA	
Software trigger (Hlt2)	Λ_c^+ Phys	
Extra $\mu \text{ ProbNNmu} >$	-	0.8
$\mu \text{ ProbNNmu} >$	0.225	0.005

The full selection for the prompt and SL modes on the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel is summarised in Tab. 16.

6.7 Calibration of particle identification

Simulations of processes happening inside the LHCb detector are non-trivial and when it comes to the PID, the contributions of the different components of the detector to the PID variables, e.g. **ProbNN**, are not sufficiently well-modeled. Modeling the interaction of a traversing particle in a detector requires informations from the kinematics of the particle, the occupancy of the detector that may differ from one event to another and experimental conditions such as the temperature, the gas pressure and alignments. In order to solve this issue, a data-driven tool was developed to help analysts compute the efficiencies related to PID selection requirements. The **PIDCalib** package is a set of tools designed to compute the efficiencies specific to one PID requirement or to re-sample a complete simulation sample. It provides generalised calibration samples and uses tag-and-probe and **sWeights** methods to compute the efficiencies [44]. The present analysis uses the calibration sample in order to compute the various efficiencies of the **ProbNN** variables involved in the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ selection. A discussion of the different resampling tools can be found in App. 12.5.

The kinematics of the signal tracks and the occupancy of the detector are usually different than the ones of the calibration sample. For this reason, the calibration sample is split into bins with sufficient amount of data and the efficiency separately in each bin. The final efficiency of the signal is obtained with a weighted average over the kinematics of the signal. The particle identification efficiency averaged over the simulated signal sample is given by

$$\varepsilon_{\text{PID}} = \frac{\sum_i^{N_{\text{toys}}} \sum_j^{N_{\text{bins}}} n_{i,j} \varepsilon_{i,j}}{\sum_i^{N_{\text{toys}}} \sum_j^{N_{\text{bins}}} n_{i,j}} \quad (8)$$

where i is the index of the kinematics/occupancy variable bin and j is the index of the bin of the toys used to propagate the uncertainties.

Throughout the analysis, we use the **PIDCalib** package to compute the efficiency on two occasions: the efficiency of the PID requirement in the final selection and the probability of mis-identification of similar decay products as those of the signal decay. The kinematic variables used to parametrise the efficiency are the momentum p and the pseudo-rapidity $\eta = \text{arctanh}\left(\frac{p_z}{|p|}\right)$, where p_z is the component of the momentum along the beam axis. The default binning proposed by the **PIDCalib** package is used for both computations and can be found in Tab. 17.

Table 17: Main binning used to obtain the efficiencies of the **ProbNN** variables for the muons and the proton tracks as well as for the computation of the mis-identification probability of the kaon and pions as a muon.

Particles	Variables	Binning	Unit
p, K, π	p	[3,9.3,15.6,19,24.4,29.8,35.2,40.6,51.4,56.8,62.2,67.6,73,78.4,83.8,89.2,94.6,100]	GeV/ c
	η	[1.5,2.375,3.25,4.125,5]	
μ	p	[3,6,8,10,12,14.5,17.5,21.5,27,32,40,60,70,100]	GeV/ c
	η	[1.5,2.375,3.25,4.125,5]	

6.8 Efficiency of the full selection

Computing the different efficiencies is required to guarantee the possibility to normalise the decay we are interested in, namely the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ rare decay or to compute the expected number of events arising from the different contamination channels. The normalisation implies that the ratio of efficiencies are taken into account and therefore, it is not required to know the absolute value of the efficiencies to give a good estimate. The systematic uncertainties due to technical considerations of the measurement are expected to mostly cancel and the ratio delivers the meaningful information. The efficiency $\varepsilon(t)$ of a requirement t is extracted from the simulated sample and given by

$$\varepsilon(t) = \frac{N(t+s)}{N(s)} \quad (9)$$

where N is the number of events that passed the requirement and s is the requirement applied prior to the requirement of interest. The final efficiency is simply computed by multiplying the efficiencies of all the requirements. In this analysis, the normalisation channel $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ is extracted from the simulated sample of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay by applying a window $|m(\mu^+\mu^-) - m_\phi| < 35 \text{ MeV}/c^2$, where m_ϕ is the mass of the ϕ meson (see Tab. 6). We define three mass regions: the low mass region, the ϕ -mass region and the high mass region. The low (high) mass region represents the dimuon invariant mass reconstruction where the momenta of the two muons are the lowest (highest). The efficiency of each requirement on the full rare decay sample can be recovered by computing the weighted average of the efficiencies in the three different regions.

Table 18: Efficiencies of the different selections computed from the simulated sample, each requirement is applied in addition to the previous requirement.

Requirement	$\Lambda_c^+ \rightarrow p\mu^+\mu^-$		
	low mass region	ϕ -mass region	high mass region
Preselection	$(9.77 \pm 0.06) \times 10^{-3}$	$(1.10 \pm 0.02) \times 10^{-2}$	$(9.49 \pm 0.10) \times 10^{-3}$
BDT	0.421 ± 0.005	0.41 ± 0.01	0.361 ± 0.007
L0	0.516 ± 0.008	0.56 ± 0.02	0.58 ± 0.02
Hlt1	0.79 ± 0.02	0.76 ± 0.04	0.78 ± 0.03
Hlt2	0.63 ± 0.02	0.66 ± 0.04	0.69 ± 0.03
p ProbNNp	0.78330 ± 0.00001	0.78283 ± 0.00002	0.77600 ± 0.00002
$\mu^+\mu^-$ ProbNNmumu	0.81583 ± 0.00004	0.8116 ± 0.0001	0.83891 ± 0.00008
p and $\mu^+\mu^-$ PID	0.64856 ± 0.00003	0.64414 ± 0.00010	0.65909 ± 0.00006
Full selection	$(6.9 \pm 0.1) \times 10^{-4}$	$(8.2 \pm 0.4) \times 10^{-4}$	$(7.1 \pm 0.2) \times 10^{-4}$

For the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ normalisation channel, the efficiencies are reported in Tab 18. Different behaviours are to be noted. The preselection tends to accept more events in the invariant mass reconstruction of the ϕ meson and this consideration remains true throughout the full selection. The BDT variable is always descending and indicates that the requirement removes more events with higher momenta of the muons. The efficiencies of the hardware trigger stage (L0) are ascendant. This can be explained by the fact that particles with high momenta have a higher probability of going above the hardware trigger threshold. The efficiencies of the PID requirements are computed with the help of the PIDCalib package. The efficiencies of the final selection in the three mass regions are computed by multiplying all the others.

Table 19 shows the efficiency of the full selection for the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$. The events arising from resonant decays of light unflavoured mesons are excluded by imposing a veto on the mass of the ϕ , ρ/ω and η particles.

6.9 Expected number of events

The normalisation channel $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ and the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ are analysed from the same data sample. The events are measured using the same detector and therefore, the efficiencies related to the technical considerations are expected to cancel by taking the ratio. This way, it is possible to compute the number of expected signal events in the data sample from the yield

Table 19: Efficiency of the selection on the rare channel $\Lambda_c^+ \rightarrow p\mu^+\mu^-$

Candidate selection	$\Lambda_c^+ \rightarrow p\mu^+\mu^-$
Selection only (no PID)	$(1.08 \pm 0.02) \times 10^{-3}$
$ m(\mu^+\mu^-) - m_\phi > 40 \text{ MeV}/c^2$	0.89 ± 0.02
$ m(\mu^+\mu^-) - m_\omega > 40 \text{ MeV}/c^2$	0.91 ± 0.02
$ m(\mu^+\mu^-) - m_\eta > 40 \text{ MeV}/c^2$	0.93 ± 0.02
All vetoes	0.73 ± 0.02
p and $\mu^+\mu^-$ PID	0.65059 ± 0.00003
Full selection + Vetoes	$(5.13 \pm 0.09) \times 10^{-4}$

of the normalisation channel. Logically, this consideration extends to different decays reconstructed in the data sample, e.g. the contaminations channels reconstructed as Λ_c^+ candidates. The number of expected N_X events for a decay X is given by

$$N(X) = N(\Lambda_c^+ \rightarrow p\phi) \frac{\mathcal{B}(X)}{\mathcal{B}(\Lambda_c^+ \rightarrow p\phi)} \frac{\varepsilon(X)}{\varepsilon(\Lambda_c^+ \rightarrow p\phi)} \quad (10)$$

where $N_{\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)}$ is the yield of $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ candidates extracted from a fit to the data sample, \mathcal{B} the branching fraction and ε the efficiency.

As a brief reminder, the main parameters of the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ normalisation channel are reported in Tab. 20.

Table 20: Branching fraction, efficiency and number of Λ_c^+ candidates for the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ normalisation channel.

Parameters	\mathcal{B}	ε	N
$\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$	$(2.97 \pm 0.63) \times 10^{-7}$	$(8.2 \pm 0.4) \times 10^{-4}$	127 ± 13

6.10 Backgrounds

In the present analysis, two sources of background are considered:

- combinatorial background: candidates reconstruction by random combinations of tracks;
- candidates reconstructed from tracks of similar decays where the final particles are misidentified under signal hypothesis.

The MVA classifiers as described in Sec. 6.3 are optimised to reject combinatorial background. The remaining candidates are described in the fit by a Chebyshev polynomial of first order.

The background resulting from misidentified tracks is analysed in similar channels with a proton in the decay products or with three hadrons possibly misidentified as the decay of interest $\Lambda_c^+ \rightarrow p\mu^+\mu^-$. Eight different modes are considered and their respective branching fractions are presented in Tab. 21. Modes without muons in the final state are expected to be highly suppressed due to the preselection requirement `IsMuon`. The considered modes containing muons

$D^+ \rightarrow K^+ \mu^- \mu^+$ and $D^+ \rightarrow \pi^+ \mu^- \mu^+$ are FCNC processes and their contribution is expected to be at the same level as the rare decay of interest $\Lambda_c^+ \rightarrow p \mu^+ \mu^-$. Due to the constraining requirement $p \text{ ProbNNp} > 0.805$, we expect the contribution of these channels to be negligible. The other requirements either contain a proton in the final state or have a high branching fraction and have to be taken into account.

Table 21: Possible contamination channels where the final particles can be misidentified as those of the signal decay [6].

Mode	Branching Fraction \mathcal{B}
$\Lambda_c^+ \rightarrow p \pi^+ \pi^-$	$(3.5 \pm 2.0) \times 10^{-3}$
$\Lambda_c^+ \rightarrow p K^+ K^-$	$(7.7 \pm 3.5) \times 10^{-4}$
$\Lambda_c^+ \rightarrow p K^+ \pi^-$	$(5.0 \pm 1.3) \times 10^{-2}$
$D^+ \rightarrow K^+ K^- \pi^+$	$(9.51 \pm 0.34) \times 10^{-3}$
$D^+ \rightarrow K^+ \pi^- \pi^+$	$(5.19 \pm 0.26) \times 10^{-4}$
$D^+ \rightarrow K^+ K^+ K^-$	$(8.5 \pm 2.0) \times 10^{-5}$
$D^+ \rightarrow K^+ \mu^- \mu^+$	$< 4.3 \times 10^{-6}$ CL= 90%
$D^+ \rightarrow \pi^+ \mu^- \mu^+$	$< 7.3 \times 10^{-8}$ CL= 90%

As the observable analysed is the invariant mass reconstruction of the Λ_c^+ candidates in the region [2200,2380] MeV/ c^2 , we need to estimate the expected number of background events in the region of interest. Because too few number of events pass the selection in the contamination simulated sample, we use the `TGenPhaseSpace` method from the ROOT framework [45] to generate the different background channels and reconstruct the candidates under the mass hypothesis of the $\Lambda_c^+ \rightarrow p \mu^+ \mu^-$ decay.

Let the energy³ of a particle i be $E_i^2 = m_i^2 + p_i^2$, where m_i is its mass and \vec{p}_i its momentum. In the Lorentz notation, the four-vector $P_i' = (E_i, \vec{p}_i)$ describes the kinematic of the particle in the laboratory frame and the four-vector P_i represents the same particle in the rest frame. The relation between the two frames is enabled by the Lorentz invariance of the norm of the four-vectors and can be written as $P_i^2 = P_i'^2$.

For example, the invariant mass reconstruction $m(p\pi^+\pi^-)$ of the 3-body decay $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ under the mass hypothesis of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay is given by

$$m(p\pi^+\pi^-)^2 = (E_p + E_{\pi^+} + E_{\pi^-})^2 - (\vec{p}_p + \vec{p}_{\pi^+} + \vec{p}_{\pi^-})^2 \quad (11)$$

where the mass of the pion m_π was replaced by the mass of the muon m_μ in the energy

$$E_\pi = \sqrt{m_\mu^2 + p_\pi^2}. \quad (12)$$

As the mass of the muon is lower than the mass of the pion $m_\mu < m_\pi$, following Eq. 11 and 12, the invariant mass reconstruction is lowered. The different masses of the particles produced in the decay of interest $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ as well as in the different background channels under investigation can be found in Tab. 4.

Figure 18 displays the different simulated backgrounds samples reconstructed under the mass hypothesis of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay. A high difference between the mass of the candidates and the particles they are misidentified for

³We use the natural system of units, where the speed of light $c = 1$.

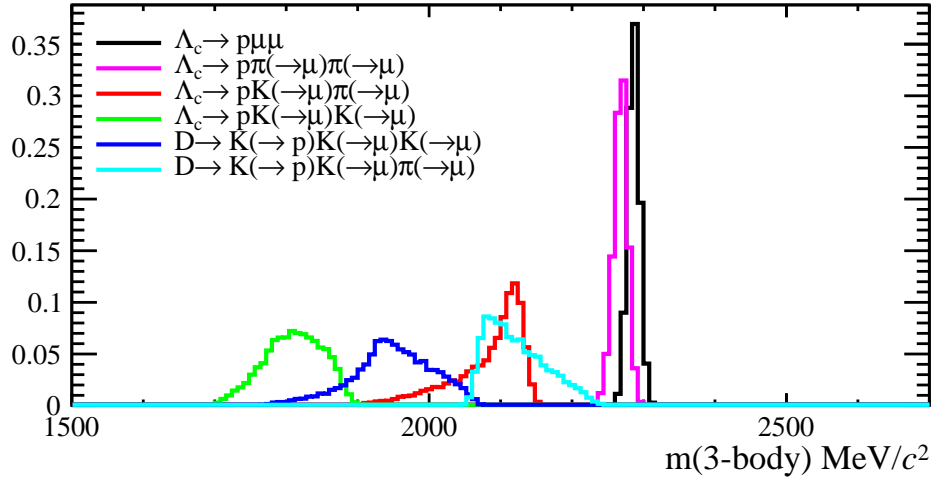


Figure 18: Simulation of different background samples reconstructed under the mass hypothesis of $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay.

implies a greater shift in the invariant mass reconstruction. The asymmetrical behaviour of the $D^+ \rightarrow K^+K^-\pi^+$ and $\Lambda_c^+ \rightarrow pK^-\pi^+$ is due to the imbalance in the momenta induced by the different masses of the decay products with respect to the signal decay where two final state particles are muons. Decays as the $\Lambda_c^+ \rightarrow p\pi^+\pi^-$, $\Lambda_c^+ \rightarrow pK^+K^-$ and $D^+ \rightarrow K^+K^+K^-$ where the two decay products reconstructed as muons are the same exhibit a shape approximatively symmetric once the invariant mass is reconstructed. From the phase-space analysis, we conclude that the two channels susceptible to contaminate the region of interest [2200,2380] MeV/ c^2 are the $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ and $D^+ \rightarrow K^+K^-\pi^+$ decays.

Following Eq. 10, it is possible to compute the expected number of events for the two contamination channels. The selection without the PID requirements is applied to the simulated contamination samples. Since the selection is designed to measure a semi-leptonic decay and the two modes are fully hadronic, we expect them to be highly suppressed. For the $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ decay, no events from the simulated sample pass the selection. For this reason, we choose to set an upper limit for the expected number of events arising from the contamination channel. By assuming the number of events follows a Poisson distribution, we set a limit at 90 % Confidence Level (CL) [6]. Concerning the $D^+ \rightarrow K^+K^-\pi^+$ decay, a few number of events satisfy the selection and we are able to estimate the contamination.

To compute the efficiency of the full selection, the PID requirements and their acceptance to the mis-identifications need to be taken into account. The mis-identification probability corresponds to the efficiency of the proton and muon PID requirement on true misidentified particle tracks. Using the `PIDCalib` package, we compute the efficiencies of the PID selection requirements on the decay products of the contamination channels $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ and $D^+ \rightarrow K^+K^-\pi^+$. Due to the low number of events satisfying the selection, we choose to determine the efficiencies in the worst case scenario. Independently from the kinematic variables p and η from which the mis-identification probability is parametrised, we extract its maximal value from the calibration sample for the two polarities of the magnets and then take the average value. The efficiency of the full

selection is given by multiplying the efficiencies of the selection and the efficiency of the mis-identification. The results are detailed in Tabs. 22 and 23. From the background study, because of their presence in the invariant mass of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay, one can conclude that the only possible contaminations arise from the $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ and $D^+ \rightarrow K^+K^-\pi^+$ decays. Thanks to the full selection and especially the PID requirements, the number of events susceptible to impact the region of interest for the analysis can be neglected for the $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ channel. Concerning the $D^+ \rightarrow K^+K^-\pi^+$ decay, the contamination is present in the left data sideband and not under the invariant mass of interest, we choose to count it in the combinatorial share.

Table 22: Identification probability for candidates from the $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ and $D^+ \rightarrow K^+K^-\pi^+$ decays to be reconstructed as $\Lambda_c^+ \rightarrow p\mu^+\mu^-$.

Efficiency ε	Magnet Up	Magnet Down	Mean value
$p(\rightarrow p)$	0.9602 ± 0.0001	0.9563 ± 0.0001	0.9583 ± 0.0002
$\pi(\rightarrow \mu)$	$(2.17 \pm 0.02) \times 10^{-2}$	$(1.88 \pm 0.02) \times 10^{-2}$	$(2.03 \pm 0.04) \times 10^{-2}$
$K(\rightarrow p)$	0.7141 ± 0.0008	0.7141 ± 0.0008	0.714 ± 0.002
$K(\rightarrow \mu)$	$(1.3 \pm 0.1) \times 10^{-2}$	$(1.3 \pm 0.3) \times 10^{-2}$	$(1.3 \pm 0.4) \times 10^{-2}$
$p\pi^+\pi^-(\rightarrow p\mu^+\mu^-)$	-	-	$(3.9 \pm 0.1) \times 10^{-4}$
$K^+K^-\pi^+(\rightarrow p\mu^+\mu^-)$	-	-	$(1.9 \pm 0.6) \times 10^{-4}$

Table 23: Branching fraction, efficiency of the full selection, efficiency of the mis-identification and upper limit on the expected number of events for the $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ and $D^+ \rightarrow K^+K^-\pi^+$ contamination modes.

Parameter	$\Lambda_c^+ \rightarrow p\pi^+\pi^-$	$D^+ \rightarrow K^+K^-\pi^+$
Branching fraction \mathcal{B}	$(4.3 \pm 0.4) \times 10^{-3}$	$(9.5 \pm 0.3) \times 10^{-3}$
Efficiency selection	$(4 \pm 3) \times 10^{-7}$	$(2 \pm 1) \times 10^{-6}$
Efficiency mis-identification	$(3.9 \pm 0.1) \times 10^{-4}$	$(1.9 \pm 0.6) \times 10^{-4}$
Expected N	< 0.382 at 90 % CL	
	2 ± 1	

7 Upper limit determination

The expected upper limit on the branching fraction of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ rare decay is determined using the CL_S technique [46]. In a counting experiment where the measured number of events are well-described by a Poisson law, two hypotheses are tested simultaneously: the *Signal + Background* ($S+B$) and *Background-only* (B). The purpose is to quantify the agreement between the data sample and each of the hypothesis. The ratio of the likelihoods, \mathcal{Q} , is given by

$$\mathcal{Q} = \frac{\mathcal{L}_{S+B}(\mu = 1)}{\mathcal{L}_B(\mu = 0)} \quad (13)$$

where μ is the signal strength while \mathcal{L}_{S+B} and \mathcal{L}_B are the likelihood functions of respectively the $S+B$ and B hypotheses. The test-statistic is defined as $-2 \ln \mathcal{Q}$

and the probability distributions of the test-statistic are defined as

$$CL_{S+B} = P_{S+B}(\ln \mathcal{Q} \leq \ln \mathcal{Q}_{obs}) \quad (14)$$

and

$$CL_B = P_B(\ln \mathcal{Q} \leq \ln \mathcal{Q}_{obs}) \quad (15)$$

where $\ln \mathcal{Q}_{obs}$ is the test-statistic observed in the data, P is the probability distribution and CL can be interpreted as the agreement of a given hypothesis with the data sample. For example, the CL_{S+B} probability represents the probability to measure a result less compatible with signal when the signal hypothesis is true. The CL_S is defined as

$$CL_S = \frac{CL_{S+B}}{CL_B}. \quad (16)$$

From Eq. 16, it is interesting to note that the CL_S observable is not a "*a true frequentist confidence or probability*" [46] but exhibits some interesting properties in the sensitivity. In the case of a clear distinction between signal and background, the observable is a good approximation of the $S + B$ probability $CL_S \sim CL_{S+B}$. The CL_{S+B} probability is a statement about the signal *and* the background and in the case of a vanishing signal, an increased uncertainty on the background tends to improve the sensitivity. In that case, the CL_S observable would indicate a decreased sensitivity. From a practical point of view, the CL_S method represents an approximate confidence level in the signal-only hypothesis.

In order to determine an upper limit, one treats the observable CL_S as a probability and sets a threshold α , *e.g.* $CL_S < 0.05$ which corresponds to an exclusion limit with a 95% CL. The expected upper limit is achieved by replacing the $\ln \mathcal{Q}_{obs}$ in Eqs. 14 and 15 by the test-statistic of the B hypothesis $\ln \mathcal{Q}_B$. This method is also known as the *Asimov data test* [47].

The implementation of the CL_S method is provided by the RooStats [48] framework.

8 Prospect on systematic uncertainties

Systematic effects arise due to the uncertainties on the branching fraction of the control sample, the efficiency ratio, the yield ratio, the choice of a specific fitting model for the invariant mass peak, the choice of fixing parameters in the fit of the rare decay and the variation of ratios of efficiencies due to the finite size of the PIDCalib package. The present analysis presents a preliminary result on the branching fraction of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay. The main sources of systematic uncertainties are reviewed but a proper estimation will be done in a future analysis.

Branching fraction of the control sample

A systematic uncertainty of 21.5% is due to the branching fraction of the normalisation channel $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$. This uncertainty is expected to be the main source of systematic uncertainties in the analysis.

Finite size of the simulated samples

Due to the finite size of the simulated samples, a systematic uncertainty has to be accounted for the computation of the expected upper limit through the CL_S

method. The statistical uncertainties on the efficiency estimation is considered as a source of systematic uncertainties in the determination of the upper limit of the branching fraction. From Tab. 18, these systematic uncertainties are expected to be negligible ($\sim 1.5\%$).

Efficiency ratio

Our limited ability to determine the ratio of the efficiencies for the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ and the normalisation channel $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ is expected to induce systematic uncertainties. The geometrical acceptance of the detector and the response of the sub-detectors, which depend on the angular distributions of the final state particles might differ between the two decays and therefore not completely cancel in the ratio as it is assumed in the present analysis. An estimate could be determined by recalculating the acceptance assuming a flat $m(\mu^+\mu^-)$ distribution.

The shape of the invariant mass distribution

An assumption is done on the model used to describe the shape of the invariant mass distribution and, since this choice is arbitrary, this induces systematic uncertainties on the measurement of the branching fraction. The same applies when fit parameters are fixed. An estimate of those systematic uncertainties can be achieved by toys studies or by constraining the parameters of the fit with a Gaussian distribution.

PID corrections

The tools of the PIDCalib package used to compute the efficiency of the PID requirements are expected to induce statistical and systematic uncertainties. The statistical uncertainties arise from the finite size of the calibration sample while the the systematic ones are inherent to the data-driven technique involved. In order to estimate the systematic uncertainties, one should vary the binning of the kinematic variables p and η and compare the value of the efficiencies. Another source of systematic uncertainty might be attributed to the choice of the phase space model. An estimation could be achieved by comparing the efficiency correction in the extreme bins of the $m(\mu^+\mu^-)$ distributions or by changing the binning.

9 Blind fit

A blind fit to the invariant mass distribution, $m(p\mu^+\mu^-)$, is performed by excluding candidates in the signal mass region $[2200, 2380] \text{ MeV}/c^2$. The result is displayed in Fig. 19.

The expected upper limit is computed in order to test the sensitivity of the analysis. The signal strength parameter μ corresponds to the number of signal events, which are directly converted into the branching fraction of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ using the Eq. 10. The CL_S method is computed using the asymptotic formula for the profile likelihood ratio [47]. In the present case, the systematic uncertainties for the expected upper limit are not taken into account, in order to compute them, toys studies have to be performed. The CL_S distribution is reported in Fig. 20.

The upper limit for the BaBar experiment and the expected upper limits for the LHCb Run I and Run II are reported in Tab. 24. The higher statistics in the data collected by the LHCb during the Run I and Run II leads to a

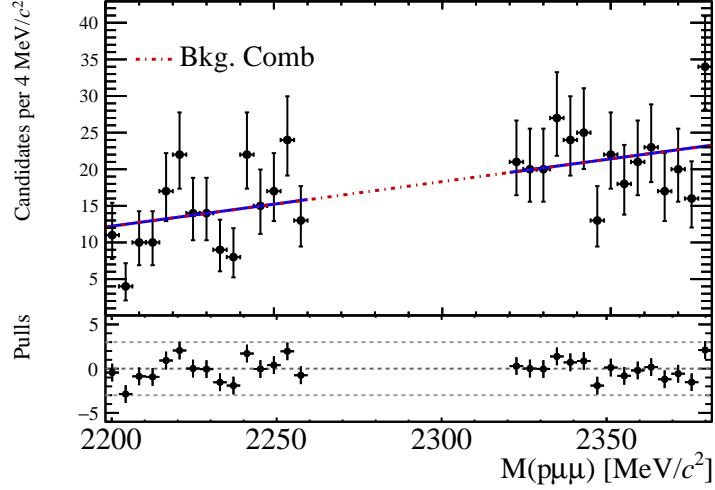


Figure 19: Blinded fit to the invariant mass reconstruction, $m(p\mu^+\mu^-)$, for the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ channel, for the data sample.

significant improvement in the sensitivity for this decay with respect to the BaBar experiment.

Table 24: Upper limits for the branching fraction of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ measured by the BaBar, LHCb 2011 and 2012 (Run I) and LHCb 2016 (Run II) experiments.

Experiment	\mathcal{L} [fb^{-1}]	$\mathcal{B}(\Lambda_c^+ \rightarrow p\mu^+\mu^-)$
BaBar (e^+e^- -collision)	384	$< 44 \times 10^{-6}$ 90 % CL
LHCb Run I (expected)	3	$< 6.71 \times 10^{-8}$ 95 % CL
LHCb Run II (expected)	1.66	$< 9.67 \times 10^{-8}$ 95 % CL

10 Discussion of the results

The BaBar analysis

Due to the general higher sensitivity of the LHCb experiment, an improvement over the BaBar results is expected. An approximative estimation of the improvement expected by the LHCb over the BaBar experiment can be obtained by comparing their results on very similar decays. The $D^+ \rightarrow \pi^+\mu^+\mu^-$ decay is a $c \rightarrow u\ell^+\ell^-$ FCNC process presented in the same analysis as the one for the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay for BaBar [18] and in the analysis in the data collected by the LHCb in 2011 at an integrated luminosity of 1 fb^{-1} [49] corresponding to less than 1/3 of the full data sample available for the 2011 and 2012 exploitation. Both collaborations set an observed upper limit for the branching fraction of the decay and the ratio between their respective results is of the order 10^{-2} . By comparison, we expect to reach a sensitivity at least of the order 10^{-7} on the observed upper limits of the LHCb Run I and Run II analyses.

Furthermore, the BaBar experiment measured an excess of 2.6σ of $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ candidates yielding $11.1 \pm 5.0 \pm 2.5$ events. If the measurement corresponds to true signal events, taking into account the sensitivity, the LHCb

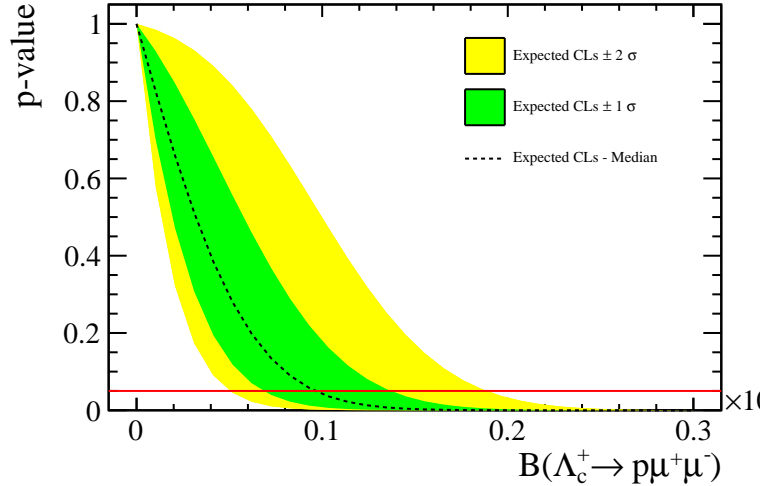


Figure 20: Expected upper limit for the branching fraction of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ computed with the CL_S method.

experiment should measure several hundreds of signal events. As the branching fraction of the $c \rightarrow u\ell^+\ell^-$ transitions are expected to be $\sim 10^{-9}$ [13, 14, 16], this would be a clear evidence of new physics in FCNC processes. An analysis in the full data sample collected by the LHCb during Run II, which is expected to reach 5 fb^{-1} of integrated luminosity, would achieve a precise measurement of the phenomenon. Nevertheless, such a consideration would raise several questions and one of the first to be addressed would be to understand why the LHCb $D^+ \rightarrow \pi^+\mu^+\mu^-$ analysis did not see the hint of an excess.

The LHCb Run I analysis

The LHCb Run I analysis is performed in data collected by the LHCb detector in 2011 and 2012 at a center-of-mass energy of 7 and 8 TeV, corresponding to an integrated luminosity of 3 fb^{-1} . The preselection *Tau23MuTau2PMuMuLine* is used in the first stage where relatively loose topological requirements are imposed and the *IsMuon* flag is required to be true. The normalisation channel is the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ decay. The MVA classifier is the BDT and the PID requirements *p ProbNNp* and *μ ProbNNmu* are used to select proton and muon tracks. The optimisation criterium is the expected upper limit at 95% CL and the optimisation is performed simultaneously on the BDT and the *ProbNN* variables.

The final results of the expected upper limit indicate that the LHCb Run I analysis has an increased sensitivity with respect to the one that can be obtained with 2016 alone. The integrated luminosity of the Run I analysis is indeed twice as big as the one available for the Run II in 2016 but this difference is compensated by the gain in the center-of-mass energy, respectively 7 and 8 TeV for 2011 and 2012 and 13 TeV for 2016. This consideration results in an increased cross-section $\sigma(c\bar{c})$ for the Run II data sample. Determining if the result is coherent with respect to the number of $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ candidates hypothetically produced in the collisions for Run I and Run II is a complex task. Since the normalisation channel is the same in both analyses, comparing the total number of $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ candidates produced would provide a quality information

on the analysis. The total number of events is given by

$$N^{\text{tot}}(\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)) = \mathcal{L} \times \sigma(c\bar{c}) \times f_{c \rightarrow \Lambda_c^+} \times 2 \times \mathcal{B}(\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)) \quad (17)$$

where \mathcal{L} is the integrated luminosity, $\sigma(c\bar{c})$ is the cross-section of the $c\bar{c}$ -production and $f_{c \rightarrow \Lambda_c^+}$ the fragmentation fraction of c -quarks into Λ_c^+ candidates. Comparing the total number of events produced should give a preliminary information on the sensitivity of both analysis. Nevertheless, the increased center-of-mass energy of the 2016 exploitation might affect the level of background, which is also determining the sensitivity. As the value of the $\sigma(c\bar{c})$ cross-section is not available for the data sample measured in 2016, this computation is out of reach for this analysis but should be taken into account in a future search for the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ rare decay.

However, the comparison of the final results for both analyses indicates that some improvements could be achieved. Throughout the analysis, the BDT and the **ProbNN** variables for the proton and the muons were considered uncorrelated. This is an approximation done to simplify the optimisation but a three dimensional optimisation should increase the performance of the selection and lead to a better discrimination of the background. In addition, optimisation of the expected upper limit as a figure of merit might lead to a significant improvement of the result as this is the criterium we would like to observe. It would also be a practical method to avoid the approximations due to the fact that the optimisation of the proton **ProbNNp** variable is performed using the significance criterium on the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel. To some extent, the optimisation of the muon **ProbNNmu** requirement on the $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ channel is also an approximation. The kinematic of the rare decay is a 3-body decay. Therefore, the phase-space distribution might affect the distribution of the **ProbNN** variables and an optimisation on decays with a different kinematic might lower the sensitivity of the analysis.

Nevertheless, the present analysis gives an interesting prospect on the necessity to perform a complete analysis of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ in the full data sample that will be collected by the LHCb in the Run I and Run II exploitation. If the unblinding of the Run I analysis suggests that some signal events of $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ candidates are present in the data sample, then an analysis in the full data sample provided by the LHCb at the end of the second run would be sufficient to perform a measurement of the branching fraction. In the other case, the measurement would remain out of reach even using the complete data sample.

11 Conclusion

A search for the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ has been conducted using data collected by the LHCb experiment in pp -collisions, corresponding to a 1.66 fb^{-1} integrated luminosity at a center-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$. Excluding the resonant contributions and assuming a phase-space model, an expected upper limit is set on the branching fraction. This result is expected to represent a factor of 100 improvement compared to the previous measurement performed by the BaBar experiment. The 95% expected upper limit on the branching fraction is

$$\mathcal{B}(\Lambda_c^+ \rightarrow p\mu^+\mu^-) < 9.67 \times 10^{-8}. \quad (18)$$

During the study, two production modes were analysed: Λ_c^+ candidates promptly produced in the pp -collisions or produced in the semi-leptonic decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$. Alongside with the production mode, two normalisation channels were investigated: the $\Lambda_c^+ \rightarrow pK^-\pi^+$ or $\Lambda_c^+ \rightarrow p\phi(\rightarrow p\mu^+\mu^-)$ decays. The prompt production mode was determined to give an optimal configuration to analyse the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ and from the trigger analysis, the only normalisation channel compatible with the prompt mode is the $\Lambda_c^+ \rightarrow p\phi(\rightarrow p\mu^+\mu^-)$.

A further improvement of the expected upper limit could be achieved by developing several aspects:

The multivariate data analysis

The TMVA package used to compute the MVA classifiers is a complete set of tools widely used by physicist in their analyses. Nevertheless, the toolkit has not been updated since 2013 and when it comes to optimisation, the complexity related to finding the best input variables or the most efficient classifier which is not overtraining increases dramatically. The `scikit-learn` project is a machine-learning library written in Python compatible with analyses performed at the LHC. It offers a set of interesting tools such as the *cross-validation*, *grid search optimisation* and *recursive features elimination* that helps extract the variables exhibiting the highest information gain and efficiently optimise the different parameters involved in the training and testing of a classifier. For this reason, computing the MVA classifiers using the `scikit-learn` library might enhance the performances of the MVA classifier in the rejection of combinatorial background.

The PID variables resampling

A limitation in the analysis comes from the fact that the PID variables are not well modeled and that the simulation of these variables is not accurate enough to be used in the various optimisations. This is the main reason why the proton PID requirement was optimised using the $\Lambda_c^+ \rightarrow pK^-\pi^+$ data sample. Better results would be achieved by directly optimising a figure of merit on the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$. Similarly to the BDT requirement optimisation, the Punzi figure of merit would give better results if one could access the PID variable in the simulation sample. Such a possibility is enabled by the `PIDGen` tool from the `PIDCalib` package. The tool allows the calibration of PID distributions directly from a data-driven calibration sample. A corrected PID variable enables an easier computation of the efficiency and could be used in optimisation processes, therefore leading to significant improvements in the analysis. The method has also the advantage of preserving the correlation among the variables, meaning that the correction does not cause a loss of information. The tools are still in development and the muon specific calibration samples were not available for the present analysis but this method is worth being investigated in future searches for rare decays at LHCb. A preliminary analysis of the resampling tools is presented in App. 12.5.

12 Appendix

12.1 The $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$ contamination in the $\Lambda_c^+ \rightarrow pK^-\pi^+$ sample

Figure 21 displays the invariant mass of the mis-reconstructed candidates of the $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$, $\Lambda_c^+ \rightarrow p\pi^+\pi^-$ and $\Lambda_c^+ \rightarrow pK^+K^-$ decays under the mass hypothesis of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel. Due to the difference in the mass of the K and π mesons, the two decays with the same particles in the decay products are not contributing in the mass region of the Λ_c^+ candidate. On the other hand, the reconstruction of the $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$ decay exhibits a contamination in the region of interest. This is entirely due to the non-reconstruction of the true π^0 pion. This consideration possibly explains the slightly asymmetrical behaviours seen in Fig. 6 for the prompt mode.

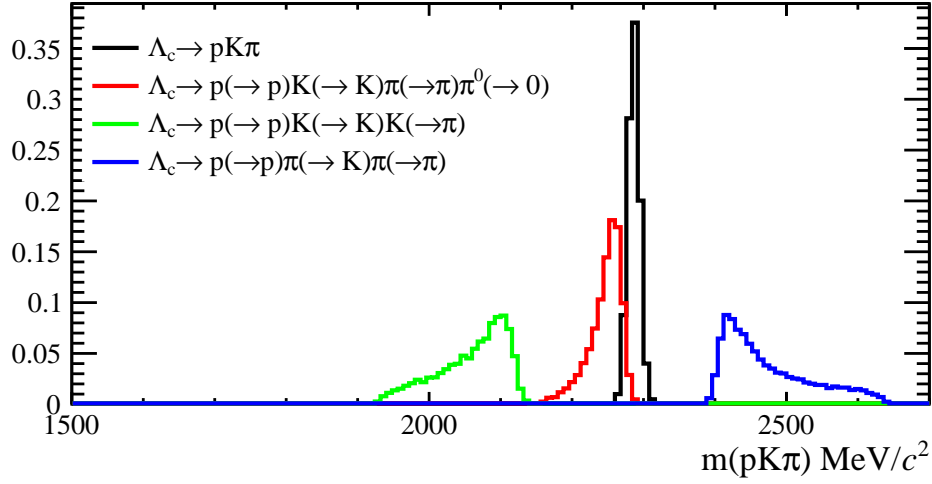


Figure 21: Misreconstruction of the $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$ decay under the $\Lambda_c^+ \rightarrow pK^-\pi^+$ mass hypothesis.

In order to complete a more rigorous analysis, one should have access to the simulated sample of the $\Lambda_c^+ \rightarrow pK^-\pi^+\pi^0$ decay. The number of events expected to contaminate the decay of interest could be deduced following Eq. 10 but rewritten with respect to some normalisation channel. As the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel is only used to compute the proton PID requirement and train the MVA classifiers, we choose to not investigate such details.

12.2 Comparison of the signal, sideband and simulation distributions

In order to select the most discriminating variables for the inputs of the MVA classifiers, we compare the distributions of the variables of the $\Lambda_c^+ \rightarrow pK^-\pi^+$ channel. We extract the signal from the region $[2260, 2320]$ MeV/ c^2 corresponding to the blinded region in the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ sample. The combinatorial background is extracted from the data sideband defined as the region below and above the region of interest. To control if the distribution of the variable is correct in the signal region, we compare it to the distribution of the simulated sample.

Nevertheless, the region of interest contains signal and background events. A separation is applied by computing the signal to background ratio within the region of interest and removing the corresponding ratio of background events in the region of interest. A fit to the data is performed on the sample right after the preselection and the yields of the signal and background are extracted to compute the ratio. Figure 22 shows the distributions of the proton momentum p

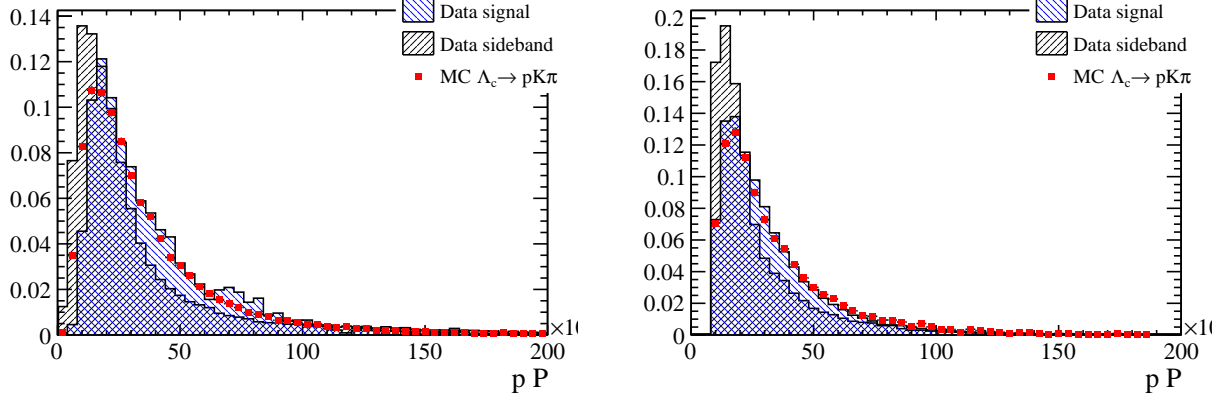


Figure 22: Comparison of the momentum p of the proton for the data signal defined in the region $[2260, 2320]$ MeV/c^2 , the data sideband and the simulated sample for the prompt (left) and the SL (right) modes.

for the data signal, data sideband and the simulated sample after the data-driven correction, for the prompt and the SL modes. A slight difference appears between the data signal and the simulated distribution of the variable. This is due to the fact that the fit to the data performed to extract the yields of the signal and the background is performed on a reduced data sample in order to accelerate the process. Therefore the signal to background ratio is only an approximation of the true ratio in the region of interest and causes some deviations between the distributions. Nevertheless, one can assert that the simulated sample reproduces the data signal correctly. This method is used to analyse the distributions of a set of variables heuristically determined to discriminate the signal from the background. Then, the variables exhibiting the best separation are used as input variable for the MVA classifier. Finally, from the results of the MVA analysis, the most correlated variables with the smallest information gain are removed and the classifiers are computed again.

Figure 23 shows the distribution of the isolation variable A_{p_T} defined in a cone of radius $R = 1$ excluding the track of the proton (see Eq. 5).

12.3 MVA training correlation matrices

The correlation matrices of the input variables for the MVA classifiers are presented in Figs. 24 and 25. The correlation between the respective variables might affect the rankings of the various variables for the BDT classifier.

For the prompt mode, the most correlated variables are the asymmetric variables A_{p_T} for the three decay products, the maximum distance of closest approach $\Lambda_c^+ \log \text{MaxDOCA}$ with the $\Lambda_c^+ \log \chi^2_{\text{vtx}}/\text{ndf}$ and the $\Lambda_c^+ \text{DTF } \chi^2$ with the $\Lambda_c^+ \text{IP}$ variable.

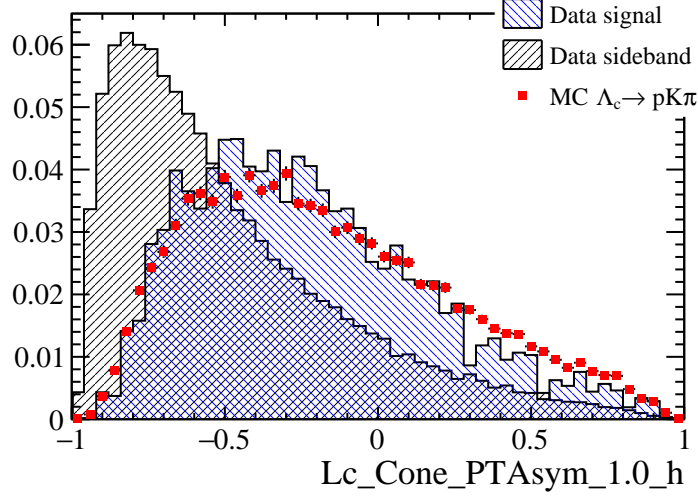


Figure 23: Distribution of the isolation variable A_{p_T} excluding the proton track and a cone of radius $R = 1$ for the data signal, the data sideband and the simulated sample for the prompt mode.

Concerning the SL mode, the Λ_c^+ $\log \text{MaxDOCA}$ and Λ_c^+ $\log \chi_{\text{vtx}}^2/\text{ndf}$ also exhibit a higher correlation. The variables specific to the tracks of the extra muon, $\mu \text{ Track } \chi_{\text{vtx}}^2/\text{ndf}$ and GhostProb are correlated.

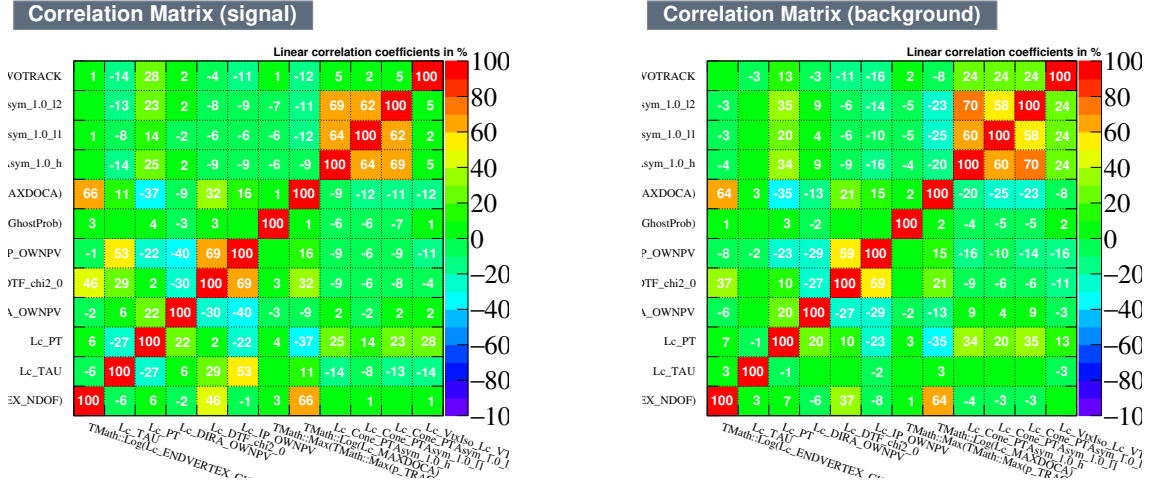


Figure 24: Correlation matrices for the prompt mode, signal (left) and combinatorial background (right).

12.4 Trigger

Table 25 shows the different trigger lines and their efficiency of activation for the prompt mode. The values are computed without any prior requirements applied. The Hlt2 Phys statement is a *or* logical statement between all the lines. A detailed analysis of the main lines shows that none of the investigated variables is able to explain the Phys activation rate. This consideration points to two main conclusions, either the Hlt2 software variables are computed wrongly or there is another requirement not taken into account in the present analysis.

Table 25: Software triggers requirements for the prompt mode, $\Lambda_c^+ \rightarrow pK^-\pi^+$ and $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ channels.

Trigger	Requirement	$\Lambda_c^+ \rightarrow pK^-\pi^+$	$\Lambda_c^+ \rightarrow p\mu^+\mu^-$
Hlt1 TOS	Global	0.354 ± 0.002	0.660 ± 0.002
	Phys	0.354 ± 0.002	0.660 ± 0.002
	Tack All Muon	$(2.4 \pm 0.2) \times 10^{-3}$	0.249 ± 0.002
	Track MVA	0.144 ± 0.002	0.295 ± 0.002
	Track Muon MVA	$(1.7 \pm 0.2) \times 10^{-3}$	0.198 ± 0.002
	Two Track MVA	0.224 ± 0.002	0.479 ± 0.002
	Single Muon High p_T	$(2.7 \pm 0.8) \times 10^{-4}$	$(2.30 \pm 0.07) \times 10^{-2}$
Hlt2 TOS	Global	0.510 ± 0.005	0.663 ± 0.002
	Phys	0.120 ± 0.003	0.384 ± 0.002
	Topo 2 Body	$(1.7 \pm 0.4) \times 10^{-3}$	$(7.8 \pm 0.4) \times 10^{-3}$
	Topo 3 Body	$(2.4 \pm 0.5) \times 10^{-3}$	$(1.53 \pm 0.06) \times 10^{-2}$
	Topo 4 Body	No Events	$(5 \pm 3) \times 10^{-5}$
	Topo E 2 Body	$(4 \pm 0.4) \times 10^{-3}$	$(7.8 \pm 0.4) \times 10^{-3}$
	Topo E 3 Body	$(2 \pm 1) \times 10^{-4}$	$(1.1 \pm 0.5) \times 10^{-4}$
	Topo E 4 Body	No Events	No Events
	Topo Mu 2 Body	$(3.6 \pm 0.6) \times 10^{-3}$	$(3.13 \pm 0.08) \times 10^{-2}$
	Topo Mu 3 Body	$(1.8 \pm 0.4) \times 10^{-3}$	$(1.63 \pm 0.06) \times 10^{-2}$
	Topo Mu 4 Body	No Events	$(7 \pm 4) \times 10^{-5}$
	Topo Mu E 2 Body	$(2 \pm 1) \times 10^{-4}$	$(9 \pm 1) \times 10^{-4}$
	Topo Mu E 3 Body	No Events	$(3.6 \pm 0.9) \times 10^{-4}$
	Topo Mu E 4 Body	No Events	No Events
	Rare Charm Lc2PMuMu	$(1.0 \pm 0.3) \times 10^{-3}$	$(8.3 \pm 0.1) \times 10^{-2}$
	Rare Charm Lc2PMuMu SS	No Events	$(1.1 \pm 0.5) \times 10^{-4}$
	Rare Charm Lc2PMue	$(1 \pm 1) \times 10^{-4}$	$(5.0 \pm 0.3) \times 10^{-3}$
	Rare Charm Lc2pee	$(1 \pm 1) \times 10^{-4}$	$(2 \pm 2) \times 10^{-5}$
	Charm Hadron D2HHH	No Events	No Events
	DiMuon Detached	No Events	No Events
	Charm Semi Leptonic 3 body D2KMuMu	No Events	No Events
	Topo Mu 2 Body BBDT	No Events	No Events

The **PIDGen** method is based on an unbinned PID resampling using a kernel density estimation. It offers two set of tools, the PID resampling similar to the one of the **MCResampling** and a PID transformation where the PID variables are calibrated in such a way that they are distributed as in the data. The approach of the PID variable transformation aims to keep the correlation of the PID variable with the other variables. Therefore, it would be possible to use it as any other variables, *e.g.* in computations of MVA classifiers. Nevertheless, the tool is in an early development stage and the main features are not updated for the calibration of simulated samples of the 2016 exploitation of the LHCb. The only feature available which is useful for the present analysis is the PID resampling based on a random generation for the p **ProbNNp** variable.

The results of the efficiencies computed from the calibration samples labeled **PIDCalib**, **MCResampling** and **PIDGen** methods are reported in Tab. 26. The efficiency computed with the **PIDGen** method is the less in agreement with the others. A detailed analysis is required to understand the origin of the difference. Although, it is important to mention that the preselection contains a PID requirement $DLL_{p\pi} > -5$ that might alter the quality of the PID resampling. Concerning the **MCResampling**, the results are not in complete agreement with the standard method. The main reason comes from the limited number of events

Table 26: Efficiencies of the different PID requirements under three different corrections, PIDCalib, MCResampling and PIDGen.

Requirement	$\Lambda_c^+ \rightarrow p\mu^+\mu^-$		
	low mass region	ϕ -mass region	high mass region
PIDCalib p ProbNNp	0.78330 ± 0.00001	0.78283 ± 0.00002	0.77600 ± 0.00002
MCResampling p ProbNNp	0.80 ± 0.01	0.81 ± 0.03	0.82 ± 0.02
PIDGen p ProbNNp	0.617 ± 0.009	0.60 ± 0.03	0.61 ± 0.02
PIDCalib $\mu^+\mu^-$ ProbNNmumu	0.81583 ± 0.00004	0.8116 ± 0.0001	0.83891 ± 0.00008
MCResampling μ ProbNNmu	0.78 ± 0.01	0.79 ± 0.03	0.80 ± 0.02

available to perform the correction. This limit has two origins: the limited size of the calibration samples that induces a systematic uncertainty on the correction and the limited number of events in the simulated sample subject to higher statistical fluctuations when the PID variable is randomly generated.

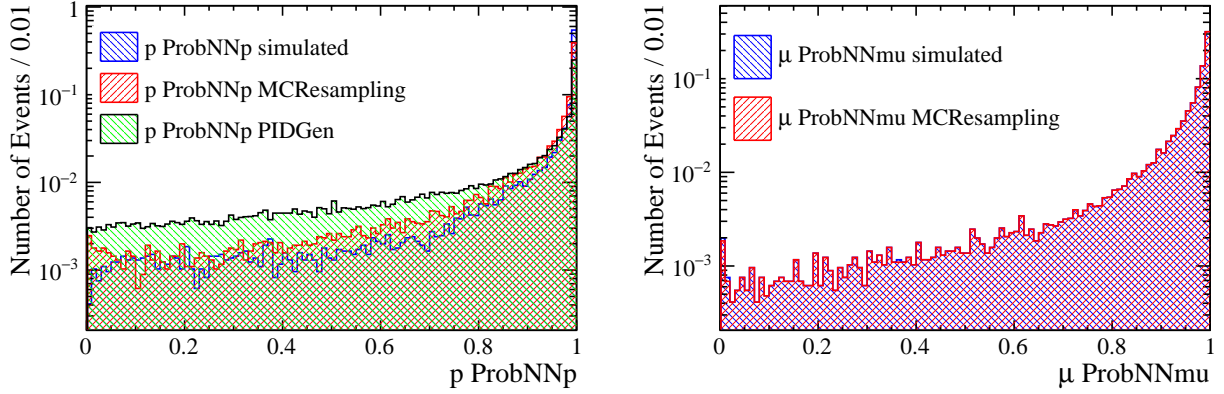


Figure 26: Comparison of the simulated PID variable, the MCResampling and PIDGen corrections for the proton (left) and the muon (right).

Figure 26 displays the simulated PID variables (blue), the MCResampling (red) and PIDGen (green) corrected variables for the proton and muon PID variables. The PIDGen correction is systematically above the other two which might indicate that either the calibration sample cannot be used for the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ due to the fact that the phase-space model is not the same as the one of the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ or that the correction was computed wrongly. Concerning the MCResampling, the proton PID variable is subject to significant corrections in opposition to the muon specific variable where the corrected distribution is almost identical to the simulated variable. The LHCb detector is particularly sensitive to muons and it is possible that the simulation of the muon PID variables is sufficiently well modeled to lead to such results, especially because the variable is less dependent on the RICH informations. Nevertheless, further analyses of the PID variables on different modes are required in order to test the viability of the simulated muon PID variables.

In the present analysis, it was decided to use the normal calibration over the MCResampling method for several reasons. The PIDCalib calibration method is widely used in analyses performed at LHCb in comparison to the MCResampling method and therefore, we expect the major issues to be publicly reported and

addressed. The **MCResampling** method is not sustained for Run II data and custom modifications were needed in order to compute the corrected variables. Nevertheless, it is interesting to mention that the values of the efficiency of the PID requirements on the normalisation channel $\Lambda_c^+ \rightarrow p\phi(\rightarrow \mu^+\mu^-)$ and the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ are relatively close and therefore cancel in the computation of the expected branching fraction. A computation of the expected upper limit for the branching fraction of the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay using the efficiencies obtained with the **MCResampling** gives

$$\mathcal{B}(\Lambda_c^+ \rightarrow p\mu^+\mu^-) < 9.87 \times 10^{-8} \quad (19)$$

which is close to the final result of the analysis $\mathcal{B}(\Lambda_c^+ \rightarrow p\mu^+\mu^-) < 9.67 \times 10^{-8}$ computed with the standard method. Finally, a further search for the rare decay $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ should investigate the possibilities offered by the **PIDGen** tools in order to exploit the advantages of using the PID variables as the other variables, *e.g.* as an input variable for the MVA classifiers.

References

- [1] The D0 collaboration, S. Abachi *et al*, "*Search for high mass top quark production in $p\bar{p}$ -collisions at $\sqrt{s} = 1.8$ TeV*", Phys.Rev.Lett.74:2422-2426 (1995), [arXiv:hep-ex/9411001](#)
- [2] The CDF collaboration, "*Observation of top quark production in $p\bar{p}$ collisions*", Phys.Rev.Lett.74:2626-2631,1995, [arXiv:hep-ex/9503002](#)
- [3] The DONUT collaboration, "*Observation of tau neutrinos interactions*", Phys.Lett.B504:218-224 (2001), [arXiv:hep-ex/0012035](#)
- [4] The ATLAS collaboration, G. Aad *et al*, "*Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at LHC*", Phys.Lett. B716 (2012) 1-29, [arXiv:1207.7214](#)
- [5] The CMS collaboration, S. Chatrchyan *et al*, "*Observation of a new boson of 125 GeV with the CMS experiment at the LHC*", Phys. Lett. B 716 (2012) 30, [arXiv:1207.7235](#)
- [6] The Particle Data Group, C. Patrigniani *et al*, "*Review of particle physics*", Chin.Phys. C40 (2016)
- [7] A. G. Riess *et al*, "*Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*", Astron.J.116:1009-1038, 1998, [arXiv:astro-ph/9805201](#)
- [8] S. Perlmutter *et al*, "*Measurements of Omega and Lambda from 42 High-Redshift Supernovae*", Astrophys.J.517:565-586 (1999), [arXiv:astro-ph/9812133](#)
- [9] The SNO collaboration, B. Aharmim *et al*, "*Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory*", Phys. Rev. C 88, 025501 (2013)
- [10] N. Cabibbo, "*Unitary symmetry and leptonic decays*", Phys. Rev. Lett. 10 (1963) 531-533
- [11] M. Kobayashi and T. Maskawa, "*CP Violation in the Renormalisable theory of weak interactions*", Prog. Theor. Phys. 49 (1973) 652-657
- [12] S. L. Glashow, J. Iliopoulos and L. Maiani, "*Weak interactions with Lepton-Hadron symmetry*", Phys. Rev. D2 (1970), 1285
- [13] S. Fajfer *et al* "*Updated constraints on new physics in rare charm decays*", Phys.Rev.D76:074010 (2007), [arXiv:0706.1133](#)
- [14] S. Fajfer *et al* "*On $D \rightarrow X_u \ell^+ \ell^-$ within the Standard Model and Frameworks like the Littlest Higgs Model with T Parity*", Phys.Rev.D83: 114006 (2011), [arXiv:1101.6053](#)
- [15] S. de Boer and G. Hiller "*Flavor & new physics opportunities with rare charm decays into leptons*", Phys. Rev.D93,074001 (2016), [arXiv:1510.00311](#)
- [16] S. Fajfer *et al* "*Rare charm meson decays $D \rightarrow Pl^+ l^-$ and $c \rightarrow ul^+ l^-$ in SM and MSSM*", Phys.Rev.D64:114009 (2001), [arXiv:hep-ph/0106333](#)
- [17] G. Burdman *et al* "*Rare Charm Decays in the Standard Model and Beyond*", Phys.Rev.D66:014009 (2002), [arXiv:hep-ph/0112235](#)

- [18] The BaBar Collaboration, "*Searches for Rare or Forbidden Semileptonic Charm Decays*", BABAR-PUB-11/008, SLAC-PUB-14482 (2011) , [arXiv:1107.4465](#)
- [19] M. Chrzaszcz *et al*, "*Search for the $\Lambda_c^+ \rightarrow p\mu^+\mu^-$ decay*", LHCb-ANA-2016-015 (2016)
- [20] The LHCb collaboration, R.Aaij *et al*, "*LHCb Detector Performance*", CERN-PH-EP-2014-290, LHCb-DP-2014-002 (2014), [arXiv:1412.6352](#)
- [21] The LHCb collaboration, R.Aaij *et al*, "*LHCb Detector Performance*", CERN-PH-EP-2014-290, LHCb-DP-2014-002 (2014), [arXiv:1412.6352](#)
- [22] The LHCb collaboration, A. A .Alves Jr. *et al*, "*The LHCb Detector at LHC*", JINST, **3** (2008) S08005. [arXiv:1211.3055](#)
- [23] The LHCb collaboration, R.Aaij *et al*, "*Performance of the LHCb Vertex Locator*", JINST **9** (2014) P09007, [arXiv:1405.7808](#)
- [24] The LHCb Collaboration, "*LHCb inner tracker: Technical Design Report*", CERN-LHCC-2002-029. LHCb-TDR-011.
- [25] The LHCb Outer Tracker Group, R.Arink *et al*, "*Performance of the LHCb Outer Tracker*", JINST **9** (2014) P01002, [arXiv:1311.3893](#)
- [26] The LHCb collaboration, R.Aaij *et al*, "*Measurement of the J/ψ pair production cross-section in pp collisions at $\sqrt{s} = 13\text{TeV}$* ", CERN-EP-2016-307,LHCb-PAPER-2016-057 (2017), [arXiv:1612.07451](#)
- [27] The LHCb RICH collaboration, M.Adolfi *et al*, "*Performance of the LHCb RICH detector at the LHC*", CERN-PH-EP-2014-290, [arXiv:1211.6759](#)
- [28] The LHCb collaboration, R.Aaij *et al*, "*Performance of the LHCb Calorimeters*", LHCb-DP-2013-004, in preparation.
- [29] The LHCb collaboration, A. A. Alves *et al*, "*Performance of the LHCb muon system*", JINST **8** (2013) P02022, [arXiv:1211.1346](#)
- [30] V. V. Gligorov and M. Williams, "*Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*", JINST **8** (2013) P02013 , [arXiv:1210.6861](#)
- [31] T. Sjöstrand *et al*, "*A brief introduction to PYTHIA 8.1*", Comput. Phys. Commun. **178** (2008) 852, [arXiv:0710.3820](#)
- [32] D. J. Lange, "*The EvtGen particle decay simulation package*", Nucl. Instrum. Meth. **A462** (2001) 152.
- [33] Geant4 collaboration, J. Allison *et al*, "*Geant4 developments and applications*", IEEETrans. Nucl. Sci. **53** (2006) 270
- [34] M. Clemencic *et al*, "*The LHCb simulation application, Gauss: Design, evolution and experience*", J. Phys. Cong. Ser. **331** (2011) 032023
- [35] D. Martinez Santos and F. Dupertuis, "*Mass distributions marginalized over per-event errors*", NIM A, 764, 150 (2014), [arXiv:1312.5000](#)
- [36] M. Oreglia, "*A study of the reactions $\psi'\gamma\gamma\psi$* ", Ph.D. thesis (1980).
- [37] J. Gaiser, "*Charmonium spectroscopy from radiative decays of the J/ψ and ψ'* ", Ph.D. thesis (1982).

- [38] T. Skwarnicki, "*A study of the radiative cascade transitions between Upsilon-prime and Upsilon resonances*", Ph.D. thesis, Institute of Nuclear Physics, Krakow, DESY-F31-86-02 (1986).
- [39] A. J. McNeil *et al* "*Quantitative Risk Management Concepts*", Techniques and Tools, Princeton University Press
- [40] N.L. Johnson "*Systems of Frequency Curves Generated by Methods of Translation*", Oxford university press on behalf of Biometrika Trust (1949), [jstor.org](http://www.jstor.org)
- [41] A. Hoecker *et al*, "*TMVA - Toolkit for Multivariate Data Analysis*", PoS ACAT:040 (2007), [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039)
- [42] G. Punzi, "*Sensitivity of searches for new signals and its optimization*", in *Statistical Problems in Particle Physics, Astrophysics and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79 (2003). [arXiv:physics/0308063](https://arxiv.org/abs/physics/0308063)
- [43] W. D. Hulsbergen, "*Decay chain fitting with a Kalman filter*", Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0503191](https://arxiv.org/abs/physics/0503191)
- [44] L. Anderlini *et al*, "*The PIDCalib package*", Tech. Rep. LHCb-PUB-2016-021, CERN (2016)
- [45] Rene Brun and Fons Rademakers, "*ROOT - An object oriented data analysis framework*", Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86. See also <http://root.cern.ch/>.
- [46] A. L. Read, "*Presentation of search results: the CLs technique*", Journal of Physics G: Nuclear and Particle Physics **28** (2002)
- [47] G. Cowan *et al*, "*Asymptotic formulae for likelihood-based tests of new physics*", Eur.Phys.J.C71:1554 (2011), [arXiv:1007.1727](https://arxiv.org/abs/1007.1727)
- [48] L. Moneta *et al*, "*The RooStats project*", [arXiv:1009.1003](https://arxiv.org/abs/1009.1003)
- [49] The LHCb collaboration, R.Aaij *et al*, "*Search for $D_{(s)}^+ \rightarrow \pi^+ \mu^+ \mu^-$ and $D_{(s)}^+ \rightarrow \pi^- \mu^+ \mu^-$ decays*", Physics Letters B 724 (2013) 203-212, [arXiv:1304.6365](https://arxiv.org/abs/1304.6365)

Acknowledgements

I would like to express my gratitude to Dr. Luca Pescatore for the constant support, the advices but also the pertinent questions he asked in order to make me progress and gain the most out of this master thesis. None of this work would have been possible without him. I would also like to thank Dr. Maurizio Martinelli for the suggestions and solutions he proposed throughout the analysis. Finally, I would like to thank Prof. Dr. Nakada, Prof. Dr. Schneider, Prof. Dr. Bay, Dr. Blanc and all the people involved in the LHCb group at EPFL for the environment they built in order to address the current and future challenges of modern physics. If *science stands on the shoulder of giants*, the LHCb group at EPFL fully assumes its role of forming the scientists of tomorrow.