University of Glasgow

Particle Physics Experimental

Ph.D Thesis

Author:
Stephen Ogilvy

Supervisors:
Dr. Lars Eklund
Dr. Paul Soler

December 19, 2014

Abstract

LHCb is the dedicated heavy flavour experiment on the LHC accelerator ring. An analysis measuring the relative branching fractions of $\Lambda_c^+ \to phh'$ decays, where $hh' \in \{K^-\pi^+, K^-K^+, \pi^-\pi^+, \pi^-K^+\}$, is presented using $1024.8 \pm 35.9\,\mathrm{pb}^{-1}$ of pp collisions gathered during 2011. This includes a search for the unobserved decay mode $\Lambda_c^+ \to p\pi^-K^+$. LHCb's software-based detector alignment utilises a track χ^2 minimisation approach, which is susceptible to so-called "weak modes" of misalignments. A study of the effects on physics analyses of weak mode misalignment in the LHCb's vertex locator (VELO) is presented, with a strategy for their constraint. Charged hadron discrimination in LHCb is provided by two ring imaging Cherenkov detectors (RICH detectors). A data-driven study on the performance of the aerogel Cherenkov radiator is presented. Particle identification information is used to construct variables describing the likelihood of hadronic particle hypotheses of reconstructed tracks, on which cuts are placed in typical physics analyses. The preparation of new data samples allowing for the data-driven efficiency correction of particle identification cuts on proton tracks is described.

Contents

1 $\Lambda_c^+ \to phh'$ relative branching fractions **2**
1.1 Introduction . . . 2
1.2 Data and processing . . . 3
1.2.1 LHCb data set . . . 3
1.2.2 Monte Carlo samples . . . 3
1.3 Selection . . . 4
1.3.1 Stripping and trigger of prompt data . . . 5
1.3.2 Stripping and trigger of semileptonic data . . . 7
1.3.3 Offline selection . . . 9
1.3.4 Mass spectra . . . 18
1.4 Prompt backgrounds . . . 22
1.4.1 $D_s^+ \to hhh / D^+ \to hhh$ mis-reconstructions . . . 22
1.4.2 Mis-ID in Λ_c decays . . . 32
1.4.3 Reassigning mass hypotheses in data . . . 34
1.4.4 Summary . . . 34
1.5 Semileptonic backgrounds . . . 36
1.5.1 Mis-ID in Λ_c decays . . . 36
1.5.2 Reassigning mass hypotheses in data . . . 37
1.5.3 Summary . . . 37

References **39**

Chapter 1

$\Lambda_c^+ \to phh'$ relative branching fractions

1.1 Introduction

The Λ_c is a charmed baryon of mass 2286.46 ± 0.14 MeV [1]. The decays of charmed baryons can be used to study colour suppression of W^+ exchange diagrams. Table 1.1 reproduces the current knowledge of the relative branching ratios of $\Lambda_c^+ \to phh'$ decays, where $hh' \in \{K^-\pi^+, K^-K^+, \pi^-\pi^+, \pi^-K^+\}$, as collected by the Particle Data Group (PDG) [1]. In 2011, LHCb collected significantly larger samples of the three hitherto observed Λ_c decays than those gathered in previous measurements.

The candidate led an analysis to measure the branching fractions of Λ_c to a proton and two charged hadrons, without hyperon mediation. The aim of the analysis was to measure the branching ratios of the three suppressed modes with respect to the Cabibbo-favored (CF) mode $\Lambda_c^+ \to pK^-\pi^+$ with the best statistical precision to date. This goal entails the first observation of the doubly Cabibbo-suppressed (DCS) decay mode $\Lambda_c^+ \to p\pi^-K^+$. This search was conducted blindly by remaining ignorant of all $\Lambda_c^+ \to p\pi^-K^+$ candidates reconstructed with a mass within ± 25 MeV of the world-average Λ_c^+ mass until the analysis procedure was finalised.

The need to finalise the analysis procedure prior to unblinding the DCS mode requires the analysis procedure to be developed utilising control modes. The singly-Cabibbo suppressed (SCS) modes $\Lambda_c^+ \to pK^-K^+$ and $\Lambda_c^+ \to p\pi^-\pi^+$ are good candidates for control modes, however, as indicated in Table 1.1, the relative branching fractions of these modes are known with poor precision. We worked to obtain the requisite confidence in our analysis methods by simultaneously examining samples of $\Lambda_c^+ \to phh'$ decays where the Λ_c is produced at the primary interaction vertex (PV) and where it is produced in semileptonic Λ_b^0 decays. The two samples are statistically independent and have different production, triggering, and selection mechanisms. By obtaining consistent relative branching ratios for the SCS modes with respect to the CF mode in these two independent samples we could be sufficiently confident that our methods will yield accurate results with the DCS mode.

In this thesis, Chapter 1 outlines the datasets and simulated samples used, the selection of candidates and the investigation of peaking backgrounds. Chapter **??** then outlines the efficiency corrections and yield extractions in the analyses. Chapter **??** then provides the

Mode	BF	Refs	Note
$\Lambda_c^+ \to pK^-\pi^+$	$(5.0 \pm 1.3) \times 10^{-2}$	[2–4]	Derived from $\mathcal{B}(\overline{B} \to \Lambda_c^+ X)$
$\to p\overline{K}^*(892)^0$	$(1.6 \pm 0.5) \times 10^{-2}$		Inclusive $\overline{K}^*(892)^0$
$\to \Delta(1232)^{++}K^-$	$(8.6 \pm 3.0) \times 10^{-3}$		
$\to \Lambda(1520)^0\pi^+$	$(1.8 \pm 0.6) \times 10^{-2}$		Inclusive $\Lambda(1520)^0$
$\to pK^-\pi^+$ nonresonant	$(2.8 \pm 0.8) \times 10^{-2}$		
$\Lambda_c^+ \to p\pi^-\pi^+$	$(3.5 \pm 2.0) \times 10^{-3}$	[5]	
$\to pf_0(980)$	$(2.8 \pm 1.9) \times 10^{-3}$		Inclusive $f_0(980)$
$\Lambda_c^+ \to pK^-K^+$	$(7.7 \pm 3.5) \times 10^{-4}$	[6, 7]	
$\to p\phi$	$(8.2 \pm 2.7) \times 10^{-4}$		Inclusive ϕ
$\to pK^-K^+$ non-ϕ	$(3.5 \pm 1.7) \times 10^{-4}$		
$\Lambda_c^+ \to p\pi^-K^+$	$< 2.3 \times 10^{-4}$	[8]	CL = 90%

Table 1.1: World-average branching fractions for $\Lambda_c^+ \to phh'$ decays as recorded in the 2012 PDG Review of Particle Properties [1]. Where noted, the PDG has adjusted the branching fractions for resonant decays to include all final states of the indicated resonances, not just the phh' final state of interest.

sources of systematic uncertainty in the analysis. We then give the methods utilised for extracting confidence intervals for the yield of $\Lambda_c^+ \to p\pi^-K^+$ and gives the final branching fraction results in the analysis.

1.2 Data and processing

1.2.1 LHCb data set

The measurements described here are based on $1024.8 \pm 35.9\,\text{pb}^{-1}$ of $\sqrt{s} = 7\,\text{TeV}$ $p-p$ collisions observed with the full LHCb detector [9]. The data was collected throughout 2011 under varying beam, detector, and trigger conditions. We use all of the available data that has been marked OK by data quality checks. For the prompt analysis, we find $588.2 \pm 20.6\,\text{pb}^{-1}$ with the magnet polarity down (MagDown) and $435.7 \pm 15.2\,\text{pb}^{-1}$ for the magnet polarity up (MagUp). For the SL analysis, we find $589.4 \pm 20.6\,\text{pb}^{-1}$ for the MagDown polarity and $435.5 \pm 15.2\,\text{pb}^{-1}$ for the MagUp polarity.

1.2.2 Monte Carlo samples

We use Monte Carlo (MC) events from the LHCb full event and detector simulation for signal and background studies and to estimate reconstruction and selection efficiencies. To replicate the selections applied to the real data we use modified versions of the stripping selections which have all PID selection criteria removed, such that the PID selection may be replicated without reliance on the poorly modelled PID discriminants in the simulation.

All other selection criteria remain the same as those in the version of the stripping run on real data.

The simulation utilised in the analyses employ one of two generator-level cuts, applied at the EvtGen level. These are:

`DaughtersInLHCb` - This requires all charged daughters in the generated decay to be produced in the range $0.01\,\text{rad} < \theta_{\text{charged}} < 0.4\,\text{rad}$ of the z-axis and all neutral daughters to be produced in the range $0.005\,\text{rad} < \theta_{\text{neutral}} < 0.4\,\text{rad}$ of the z-axis.

`LHCbAcceptance` - This requires the head particle/mother of the specified signal decay to be produced in the range $0 < \theta_{signal} < 0.4\,\text{rad}$ of the z-axis.

All generated MC samples utilise the `DaughtersInLHCb` cut, with the exception of the semileptonic CF $\Lambda_c^+ \to pK^-\pi^+$ which utilises the `LHCbAcceptance` cut.

The prompt $\Lambda_c^+ \to phh'$ MC samples were generated without a resonance structure. All decays in the simulation for both analyses proceed via. the `PHSP` EvtGen decay model [10]. The generated decays accordingly populate the phase space uniformly. A fraction of the prompt MC events contain only Λ_c^+ produced in b-hadron decays. In order to ensure that the efficiencies computed from the prompt MC samples are well defined, we process these samples with a filter that retains only events that contain a promptly produced signal decay. The filter traces the ancestry of the generated signal Λ_c^+. If any of its ancestors has a mean lifetime longer than 0.1 fs then it is classified as not-prompt.

The semileptonic simulation used for the CF mode $\Lambda_c^+ \to pK^-\pi^+$ include an admixture of pure phase-space and pseudo-resonance structure of the decay generated according to the following fragment of EvtGen code:

```
Decay MyLambda_c+
  0.02800         p+       K-       pi+          PHSP;
  0.01065         p+       Myanti-K*0            PHSP;
  0.00860         Delta++  K-                    PHSP;
  0.00414         MyLambda(1520)0  pi+           PHSP;
Enddecay
```

The semileptonic samples of the SCS decays $\Lambda_c^+ \to pK^-K^+$ and $\Lambda_c^+ \to p\pi^-\pi^+$ also include an admixture of pure phase-space and pseudo-resonance structure. In the sample used for $\Lambda_c^+ \to pK^-K^+$ 48.4% of the signal Λ_c^+ undergo a pure phase space decay $\Lambda_c^+ \to pK^-K^+$ and the remaining 51.6% decay through an intermediate ϕ resonance, $\Lambda_c^+ \to p\phi(K^-K^+)$. In the sample used for $\Lambda_c^+ \to p\pi^-\pi^+$, 55.6% of the signal Λ_c^+ undergo a pure phase space decay $\Lambda_c^+ \to p\pi^-\pi^+$ and the remaining 44.4% decay through an intermediate $f_0(980)$ resonance, $\Lambda_c^+ \to pf_0(980)(\pi^-\pi^+)$.

1.3 Selection

Herein we outline the selection of candidate $\Lambda_c^+ \to phh'$ decays in the prompt and semileptonic analyses. Selection takes place at several levels: at the stripping, trigger and offline

selections. The central gathering of interesting decays is known as the "stripping" selection, and its purpose is to gather decays of interest in a general fashion, such that several different measurements may be performed using the same decay mode. In the central processing of data, the stripping lines have a finite bandwidth. They must therefore make an initial selection of the decays for general purpose which preserves as much signal as possible while keeping the retention of candidates within acceptable limits. A rigid selection of trigger requirements is enforced in this analysis such that the efficiency of the trigger selections can be reliably measured. Finally, a further selection is made to adapt the selection in the stripping for the specific needs of this analysis, called the "offline" selection. This includes a multi-variate selection for use in selecting the rare decays of $\Lambda_c^+ \to p\pi^- K^+$.

1.3.1 Stripping and trigger of prompt data

Stripping

The stripping selection reconstructs the general pattern $\Lambda_c^+ \to phh'$. Table 1.2 lists the selection criteria applied at the stripping level. The main goals of the selection are to preserve signal while vetoing combinatoric backgrounds and backgrounds from other charm decays. The multiplicity of LHCb events is high due to the hadronic production environment. Consequently, cuts on the impact parameter (IP) of reconstructed heavy-flavour particles are often a useful discriminant in rejecting the large combinatoric backgrounds from unrelated tracks. Due to the short Λ_c lifetime (approximately 200 fs) relative to the D and B mesons, this approach provides less discriminatory power over the rejection of combinatorics.

We therefore utilise a suite of tight cuts on the momenta (p) and transverse momenta (p_{T}) of the candidate daughters. The requirement is made that all candidate daughters have a significant impact parameter with respect to the primary interaction. A cut is specifically placed upon the minimum daughter $IP\chi^2$, which is the increase in the associated primary vertex χ^2 when the daughter track is included in the primary vertex fit. PID selection is also employed at this stage to reduce the retention of the stripping line while vetoing a large number of candidates which are combinations of unrelated tracks. Requirements are made that the reconstructed tracks are of a high quality, specifically that the maximum normalised track χ^2 be below a given threshold and that the "Clone distance", a quantity based on the Kullback-Liebler distance between tracks [11], must exceed a given value to remove clone tracks which do not share common hits.

All permutations of pairs of the daughter tracks should have a reconstructed distance of closest approach ($DoCA$) no greater than a maximum threshold. Requirements are placed on the Λ_c vertex quality and on the distance between the Λ_c decay vertex and the primary vertex. We require that the Λ_c candidate points back to the primary interaction by placing a minimum requirement on the cosine of the angle between the reconstructed Λ_c momentum and the displacement vector between the primary vertex and the Λ_c decay vertex, a quantity referred to as the $DIRA$. We require that the reconstructed candidate

has a computed lifetime below 1.2 ps, which preserves most Λ_c signal while vetoing significant amounts of meson decays which are misreconstructed. Finally we require the candidate mass to be within 90 MeV/c^2 of the nominal PDG Λ_c mass.

	Variable name	Stripping
$h^{\pm}$	Track $\chi^2/N_{\rm ndf}$	< 5
	Clone distance	> 5000
	IP χ^2	$> 4, 8$
	$p_{\rm T}$	$> 400, 1200\,{\rm MeV}/c$
	$\lvert\vec{p}\rvert$	$> 3.2\,{\rm GeV}/c$
p	$\log\left(\mathcal{L}_p/\mathcal{L}_\pi\right)$	> 5
K	$\log\left(\mathcal{L}_K/\mathcal{L}_\pi\right)$	> 5
π	$\log\left(\mathcal{L}_K/\mathcal{L}_\pi\right)$	< 0
$h_i h_j$	$DoCA$	$< 0.10\,{\rm mm}$
Λ_c^+	Vertex $\chi^2/N_{\rm ndf}$	< 20
	VD χ^2	> 16
	$DIRA$	> 0.9999
	Proper time τ	$\in [0, 1.2]\,{\rm ps}$
	$\lvert m_{phh'} - m_{\Lambda_c^+{\rm PDG}}\rvert$	$< 90\,{\rm MeV}/c^2$

Table 1.2: Selection criteria for prompt $\Lambda_c^+ \to phh'$ candidates in the stripping selection. h denotes the pion, kaon, or proton product of a Λ_c^+ decay and is used to indicate cuts that are applied to all three. Where two values for the daughter h $p_{\rm T}$ lower limit are indicated, all three daughters are required to satisfy the looser cut and at least one of the daughters must satisfy the tighter cut. $h_i h_j$ denotes every pairing of the daughters.

Trigger

While a dedicated HLT2 trigger line operated in 2011 for the $\Lambda_c^+ \to pK^-\pi^+$ mode, there were no such lines for the prompt SCS and DCS Λ_c decays for 2011. In order to have a consistent trigger chain for the four modes we adopt the trigger requirement that the event was triggered independently of the Λ_c candidate, or Triggered Independently of Signal (TIS). The efficiencies of TIS chains are typically very low, but a TIS chain provides the benefit of trigger efficiency cancellation between the $\Lambda_c^+ \to phh'$ modes. The specific requirements are that any L0 trigger was fired independently of the Λ_c candidate, and that the event was triggered by at least one of a collection of physics analysis trigger algorithms in the HLT independently of the Λ_c candidate.

1.3.2 Stripping and trigger of semileptonic data

Stripping

Each of the stripping lines in the semileptonic analysis reconstruct candidates according to the pattern $\Lambda_b^0 \to \Lambda_c^+ \mu^\pm$, $\Lambda_c^+ \to phh'$. Table 1.3 lists the selection criteria applied at the stripping level. The selection requirements on the $\Lambda_c^+ \to phh'$ component of the reconstruction are broadly similar to that used in the prompt analysis, but with looser momentum requirements placed upon the daughters. The relaxation of these cuts is possible due to the additional discrimination acquired from information on the Λ_b^0 decay. The minumum requirement on the Λ_c daughter $IP\chi^2$ is raised, as the daughters of the semileptonically-produced Λ_c decays have higher impact parameters than for daughters of those Λ_c which are produced promplty at the primary interaction. Similar selection criteria are applied to the reconstructed Λ_c candidates as are applied to those in the prompt analysis.

Muon candidates are required to pass a series of track quality and kinematic requirements. The tracks are also required to have associated hits in all muon stations. The Λ_c candidates and muon candidates together form Λ_b^0 candidates which are required to form a common vertex which is significantly displaced in z from the primary vertex. The Λ_b^0 candidate is also required to point back to the primary interaction. As the neutrino in the decay is not reconstructed, the mass of the reconstructed Λ_c μ system is allowed to be significantly lower than the Λ_b^0 nominal mass, with a permitted $m_{\Lambda_c\mu}$ range of 2.5 – 6.0 GeV/c.

Trigger

For the semileptonic decays we use a trigger requirement that focuses on the muon in the Λ_b^0 decay and thus reduces the relative bias on the Λ_c^+ decay phase space. The muon in the Λ_b^0 decay is required to have triggered the L0 muon trigger for that event. The Λ_b^0 candidate is required to have triggered at least one of a series of algorithms in the HLT, designed to select decays of heavy-favour hadrons to final states including muons by exploiting the decay topology.

	Variable name	Stripping
$h^{\pm}$	Track χ^2/N_{ndf}	< 4
	Clone distance	> 5000
	IP χ^2	> 9
	p_{T}	> 300
	$\lvert\vec{p}\rvert$	$> 2\,\text{GeV}/c$
p	$\log(\mathcal{L}_p/\mathcal{L}_\pi)$	> 4
	$\log(\mathcal{L}_p/\mathcal{L}_K)$	$> 10^{-10}$
K	$\log(\mathcal{L}_K/\mathcal{L}_\pi)$	> 4
π	$\log(\mathcal{L}_K/\mathcal{L}_\pi)$	< 10
π for $\Lambda_c^+ \to p\pi^-\pi^+$	$\log(\mathcal{L}_K/\mathcal{L}_\pi)$	< 0
$h_i h_j$	$DoCA$ χ^2	< 20
Λ_c^+	Vertex χ^2/N_{ndf}	< 6
	$p_{\text{T}p} + p_{\text{T}h} + p_{\text{T}h'}$	$> 1800\,\text{MeV}/c$
	VD χ^2	> 100
	$DIRA$	> 0.99
	$\lvert m_{phh'} - m_{\Lambda_c^+\text{PDG}}\rvert$	$< 80\,\text{MeV}/c^2$
μ	Track χ^2/N_{ndf}	< 4
	IP χ^2	> 9
	p_{T}	$> 800\,\text{MeV}/c$
	$\lvert\vec{p}\rvert$	$> 3\,\text{GeV}/c$
	$\log(\mathcal{L}_\mu/\mathcal{L}_\pi)$	> 0
Λ_b^0	Vertex χ^2/N_{ndf}	< 6
	$z_{\Lambda_c^+} - z_{\Lambda_b^0}$	$> -9999\,\text{mm}$
	$DIRA$	> 0.999
	$m_{\Lambda_c^+\mu}$	$\in (2.5, 6)\,\text{GeV}/c^2$

Table 1.3: Stripping selection criteria for the semileptonic $\Lambda_b^0 \to \Lambda_c^+\mu^{\pm}$, $\Lambda_c^+ \to phh'$ candidates in the analysis. h denotes the pion, kaon, or proton product of a Λ_c^+ decay and is used to indicate cuts that are applied to all three. $h_i h_j$ denotes every pairing of the daughters.

1.3.3 Offline selection

The offline selection is designed to taylor the output of the general-purpose datasets from the stripping selection output and adopt the selection for the specific purposes of our analysis. All offline selections are trained using 10% of the $\Lambda_c^+ \to pK^-\pi^+$ data from each respective analysis. The ${}_s\mathcal{P}lots$ method [12] is used to extract the signal and background distributions from the data. This data is then excluded from the remainder of the analysis, with an appropriate scaling factor included in the final branching fraction calculations. It is desirable to utilise the same offline cuts in the selection of all four modes for the cancellation of systematic uncertainties. The selections therefore are all trained to maximise sensitivity to the unobserved doubly-Cabibbo suppressed mode. To this end an additional global signal weighting of

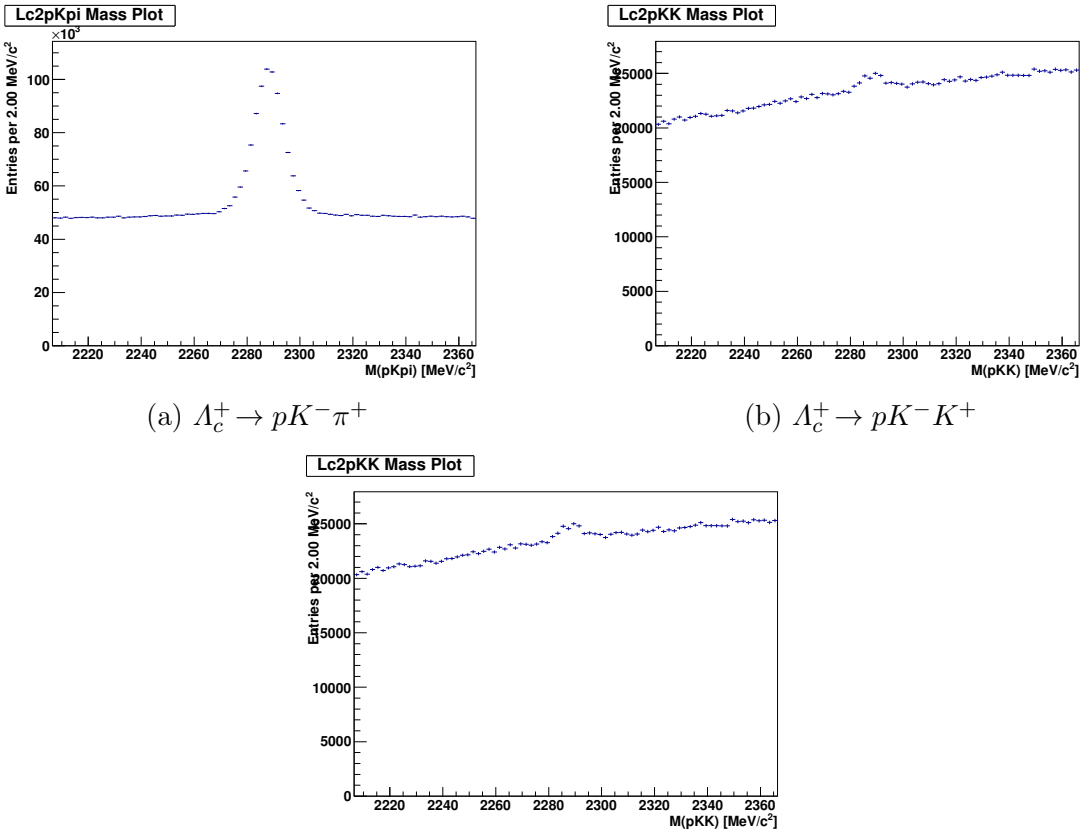
$$\frac{|\mathrm{V_{us}}|^2|\mathrm{V_{cd}}|^2}{|\mathrm{V_{cs}}|^2|\mathrm{V_{ud}}|^2} \approx 0.003, \tag{1.1}$$

which is a crude indication of the expected relative rate of $\Lambda_c^+ \to p\pi^-K^+$ over $\Lambda_c^+ \to pK^-\pi^+$, was applied in the trainings. Examination of the sidebands of the Cabibbo-favoured and doubly-Cabibbo suppressed data indicate that the combinatoric background in the doubly-Cabibbo suppressed mode is approximately 33 % lower than that in the Cabibbo favoured mode, as shown in Figure 1.3. This is taken into account in the selection training by applying a global combinatoric background weight of 0.67.

The output from the stripping lines contain a high combinatoric component, which must be further reduced in order to accurately fit the mass distributions. The mass distributions at stripping level for the modes which are not blinded are shown in Figure 1.1 for the promptly selected candidates and Figure 1.2 for the semileptonically selected candidates. This necessitates the use of cuts on the PID varibles to veto combinatorics from wrongly identified particles. Selection training has shown that PID DLL variables possess the greatest discrimination between signal and combinatoric background (compared to track/vertex quality, event kinematic info etc.). These are simultaneously optimised using the CROP tool, and performed in both the prompt and the semileptonic selections. The optimisations use the training figure of merit $\mathrm{S}/\sqrt{\mathrm{S}+\mathrm{B}}$, where S and B are the sums of the signal and background weights respectively.

The efficiencies of some stages of the selection must be evaluated on an event-by-event basis. The stripping selection efficiency, for example, must be calculated by assigning a per-candidate local efficiency depending on the candidate's location in the phase space parameterising the resonant $\Lambda_c^+ \to phh'$ decay (see Section **??**). The ${}_s\mathcal{P}lots$ method is used for background subtraction in all yield extractions. In order to preserve the event weight normalisation, each selected Λ_c candidate must have daughter kinematics which correspond to a valid region of the decay phase space, such that a valid stripping efficiency may be assigned to the event.

To ensure each Λ_c candidate has valid daughter kinematics the DecayTreeFitter tool [13] is employed. In contrast to typical pattern reconstructions, which begin with the final daughters and work backwards to form parent particles, this tool simultaneously fits the entire decay chain using information on the momenta and vertex positions, and is able to



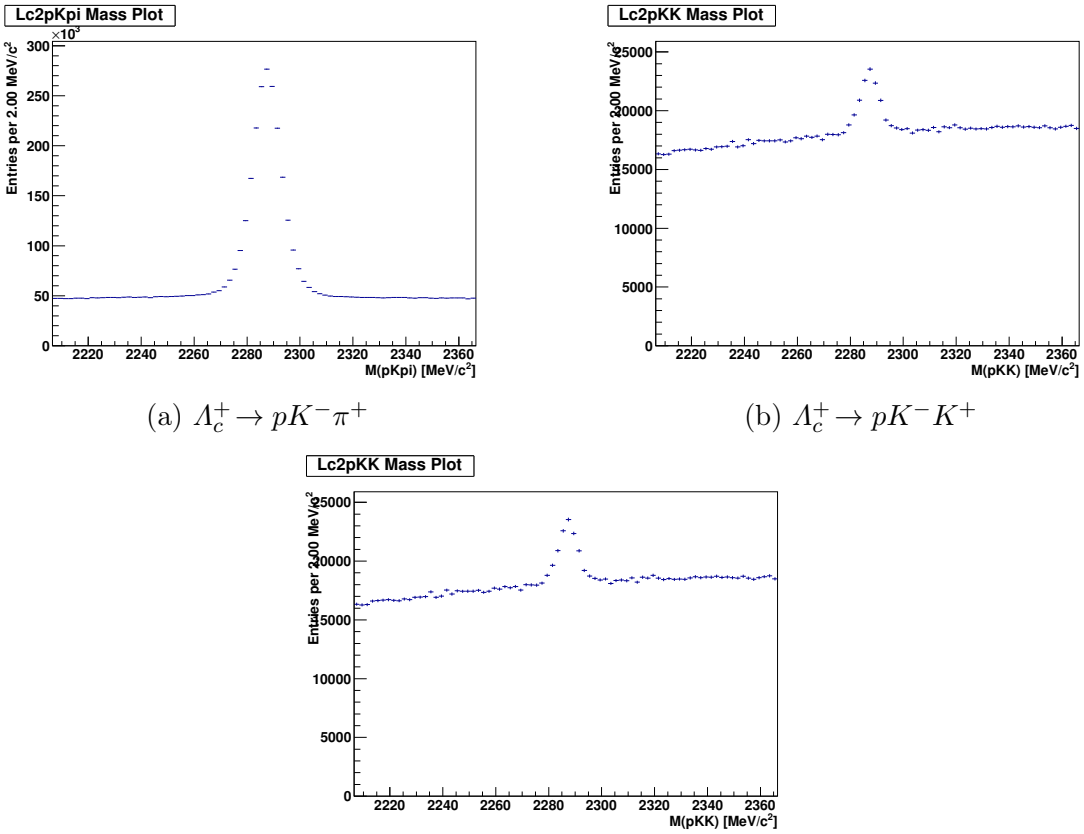
(a) $\Lambda_c^+ \to pK^-\pi^+$ (b) $\Lambda_c^+ \to pK^-K^+$

(c) $\Lambda_c^+ \to p\pi^-\pi^+$

Figure 1.1: The mass distributions of the raw prompt stripping output, before the application of trigger cuts.

utilise constraints using nominal particle masses. Our implementation uses a constraint on the Λ_c mass and no primary-vertex constraint. Convergence of the DecayTreeFitter algorithm is made necessary for an event to pass the selection. This has a minimal effect on the signal region, and primarily vetos events in the high mass sideband. The effect of this convergence criterion is demonstrated in the promptly selected $\Lambda_c^+ \to pK^-\pi^+$ in Figure 1.4. This criterion is applied to both the prompt and semileptonic selections.

In order to ascertain the efficiencies of the PID DLL cuts used in the selections it is necessary to utilise the data-driven `PIDCalib` technique. This technique uses calibration tracks from decay modes which may be cleanly reconstructed without the use of the PID discriminants. These tracks are binned in p and η to derive local PID efficiencies, which are then used to ascribe event-by-event PID efficiencies to the signal decays in this analysis (this procedure is outlined more fully in Section **??**). There exist regions of the $p - \eta$ phase space for which $\Lambda_c^+ \to phh'$ signal tracks exist, but for which no calibration data exists. This is due to the lower p_{T} of protons in $\Lambda^0 \to p\pi^-$ decays than those protons in $\Lambda_c^+ \to phh'$ decays. In such regions it is impossible to attribute a PID efficiency to the signal data. It



(a) $\Lambda_c^+ \to pK^-\pi^+$

(b) $\Lambda_c^+ \to pK^-K^+$

(c) $\Lambda_c^+ \to p\pi^-\pi^+$

Figure 1.2: The mass distributions of the raw semileptonic stripping output, before the application of trigger cuts.

is therefore necessary to include a variety of fiducial cuts on the kinematics of the daughter tracks in the $\Lambda_c^+ \to phh'$ signal decays to exclude Λ_c candidates with daughter particles falling in such regions. This ensures a reliable PID efficiency is ascribed to each event of the signal data. The vetoed regions are given in Table 1.4, and illustrated in Figure 1.5. The vetoes exclude between 15 – 30 % of signal candidates varying on a per-mode basis.

Daughter Particle	Vetoed Regions
π	$\eta < 2.0$, $\eta > 4.5$, $p < 5GeV/c^2$, $p > 100GeV/c^2$
K	$\eta < 2.0$, $\eta > 4.5$, $p < 5GeV/c^2$, $p > 100GeV/c^2$
p	$\eta < 2.0$, $\eta > 4.5$, $p < 15.6GeV/c^2$, $p > 100GeV/c^2$, $(2.0 < \eta < 2.625$ and $p > 29.3GeV/c^2)$, $(2.625 < \eta < 3.25$ and $p > 44.83GeV/c^2)$

Table 1.4: Kinematically vetoed regions of the p and η phase space such that the data-driven `PIDCalib` method may be properly utilised.

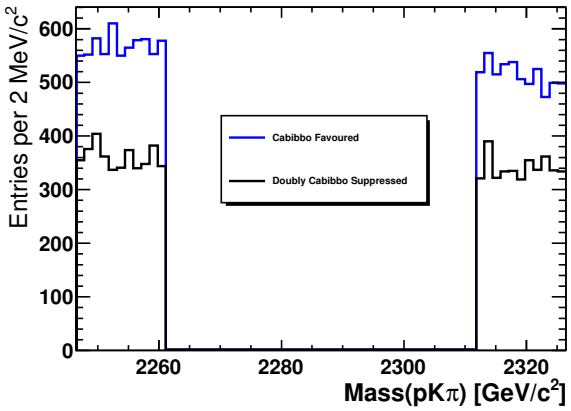
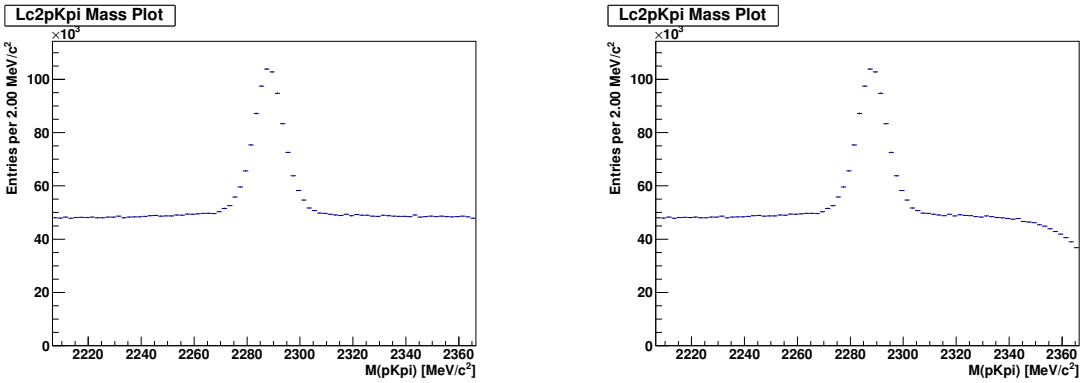


Figure 1.3: The mass sidebands from data of the promptly selected $\Lambda_c^+ \to pK^-\pi^+$ and $\Lambda_c^+ \to p\pi^-K^+$, after the final selection. Integrating these distributions and taking the ratio of the doubly-Cabibbo suppressed over the Cabibbo favoured gives combinatoric background ratio of 0.67.



(a) (b)

Figure 1.4: Mass distributions of promptly selected $\Lambda_c^+ \to pK^-\pi^+$ before (a) and after (b) the application of the DTF configuration and convergence criterion.

Offline selection of prompt candidates

The Λ_c mean lifetime is 200fs [1], which is much lower than that of charmed mesons (for example, the D^+ mean lifetime is 1040ps). Due to the resultant low separation of the Λ_c decay vertex with respect to the primary vertex, combinatoric backgrounds resulting from combinations of unrelated tracks are large in the case of promptly produced Λ_c. In order to suppress these backgrounds, a multivariate selection is used in tandem with the offline PID DLL cuts in the selection of $\Lambda_c^+ \to p\pi^-K^+$ candidates.

Multivariate selections combine information from numerous input variables describing a process and combine it to form one variable. When attempting to discriminate between

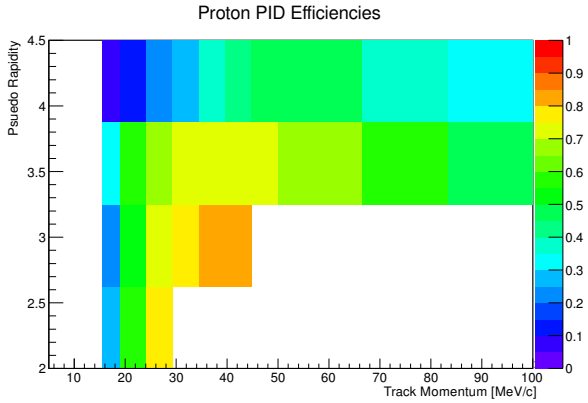


Figure 1.5: Kinematically vetoed regions of the p and η phase space for promptly selected protons. White bins indicate vetoed regions, coloured bins give the average efficiency of the PID DLL cuts on proton tracks in that region, as evaluated using the $\Lambda^0 \to p\pi^-$ PID calibration samples.

multiple species in a sample, a simple selection placed on the multivariate can possess much greater discriminatory power than simple selections placed upon the input variables. A decision tree (BDT, [14]) consists of a branching tree of nodes. Starting from one node, a simple selection requirement is made on one or a combination of the input variables with a binary output. Each output leads to a separate node whereby another requirement is made. Ultimately at the final nodes signals are assigned a value corresponding to their likelihood of being each species. A boosted decision tree (BDT [15]) consists of a series of trees in a "forest", whereby the architecture of each tree is informed by the outcomes of the previous trees. Several algorithms exist to dictate the development of the forest, or the "boosting". One such algorithm, ADABOOST [16] , assigns greater weights for a given tree in the training phase to events which were misclassified in previous trees.

A boosted decision tree was trained using the ADABOOST adaptive boosting algorithm on the $\Lambda_c^+ \to pK^-\pi^+$ data. The ${}_s\mathcal{P}lots$ technique is utilised to disentangle the signal and background input variable distributions. The training and testing of the BDT was performed in TMVA [17]. Initially the TMVA default values for BDT architecture were utilised. The input variables were chosen such that the selection possesses minimum sensitivity to the kinematics of the daughter particles in the Λ_c decay to ensure the BDT's selection of events has a uniform efficiency across the decay modes. These variables are:

- Reconstructed Λ_c p_{T}.
- Λ_c maximum $DoCA$, the distance of closest approach between any of the possible pairs of particles in the Λ_c decay.
- Λ_c $DIRA$, the direction angle of the reconstructed Λ_c momentum w.r.t. the direction of flight to the best primary vertex.
- Λ_c end vertex χ^2.

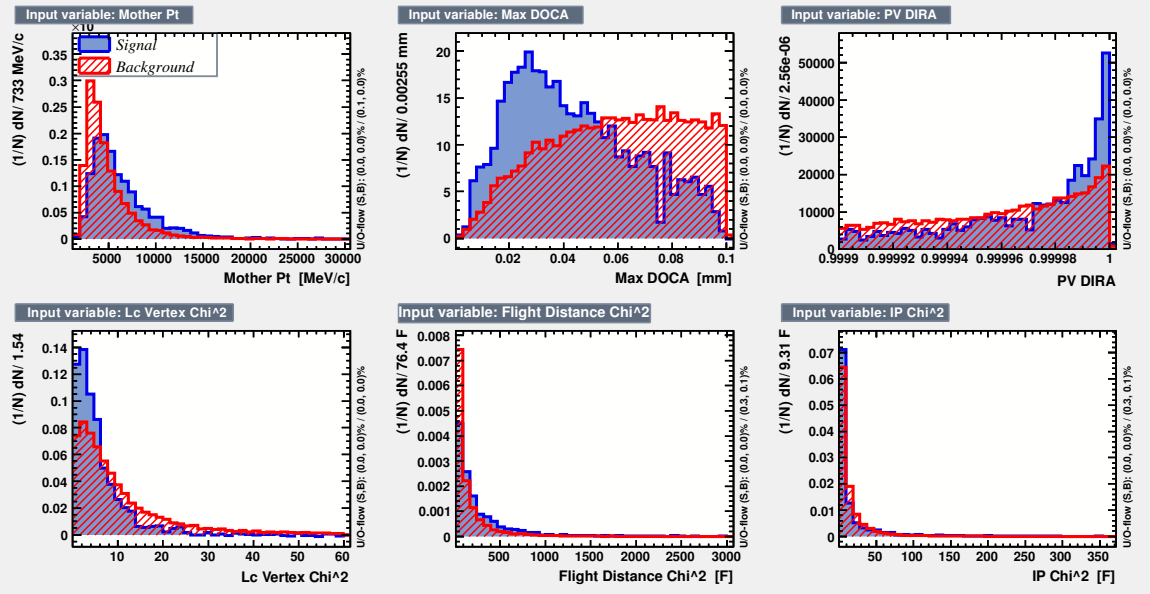


Figure 1.6: BDT input variables distributons for signal and combinatoric background, from the promptly selected $\Lambda_c^+ \to pK^-\pi^+$ training sample.

- Λ_c flight distance χ^2 with respect to the primary vertex.
- Λ_c $IP\chi^2$, the impact paramter χ^2 w.r.t. the best primary vertex, or the increase to the χ^2 of the associated primary vertex when the reconstructed Λ_c momentum is included in the primary vertex fit.

The distributions of these variables for signal and background for the prompt selection training sample, before the offline PID cuts, are given in Figure 1.6. These are acquired using the ${}_s\mathcal{P}lots$ technique to extract the signal and combinatoric distributions.

In the selection of the Cabibbo favoured and singly-Cabibbo suppressed modes, we do not apply any MVA offline selection but do apply offline PID cuts. In the selection of $\Lambda_c^+ \to p\pi^-K^+$ candidates, the offline PID cuts and the BDT itself are optimised in tandem recursively to attain the most powerful discrimination attainable. This procedure involves the following steps:

1. A BDT is trained on the training sample without any offline PID cuts applied.
2. The BDT response cut is then optimised in CROP with the cuts on the PID DLL variables, using $S/\sqrt{S+B}$ as a figure of merit. As the PID cuts will exclude some background events which the BDT is trained to exclude, the BDT will not be optimally trained for a selection in conjunction with the PID DLL cuts.
3. A new BDT is then trained on a data sample which has had the optimum PID DLL cuts as found in step 2 applied as a preselection. The new BDT response cut is then optimised in CROP in tandem with the cuts on the PID DLL variables to find new optimum PID cuts for this tree architecture.

4. This process is repeated, with new BDTs being trained until the optimum PID response as found by CROP converges.

At this point we can at the very least say that a local maximum in the N-dimensional cutspace has been found. A cross check is carried out to look for higher maxima of signal significance. A set of arbitrary PID cuts are applied to the training sample initially, and the simultaneous PID/BDT optimisation is reperformed. The arbitrary PID cuts used are:

- proton DLL(p - π) > 10.
- proton DLL(p - K) > 10.
- kaon DLL(K - π) > 8.

The same optimum PID and BDT cuts are attained, demonstrating that the maximum found is likely a global maximum.

At this stage an optimisation of the tree architecture takes place. The optimisation process takes into account both signal and background discrimination and the compatibility of the BDT response between the training and testing phases (as evaluated with a Kolmogorov Smirnov test [18] by TMVA). The former should be maximised for obvious reasons. The latter gives an indication of the level of overtraining, where the training phase becomes sensitive to the statistical fluctuations in the training sample, leading to a sub-optimal MVA architecture. The variables defining the tree architecture which were optimised are:

$\mathrm{N_{trees}}$ - The number of trees in the forest.

Boost type - The algorithm which determines the boosting procedure.

Adaptive Boost β - The learning rate for the adaptive boost algorithm.

Maximum Tree Depth - The maximum depth of the trees, or depth of nodes per tree.

Steps in Node Optimisation - At each node in the tree the data is split according to a binary decision based on the input variable of interest at the node. This variable is the number of steps in the variable range used to find the optimum splitting point at each node.

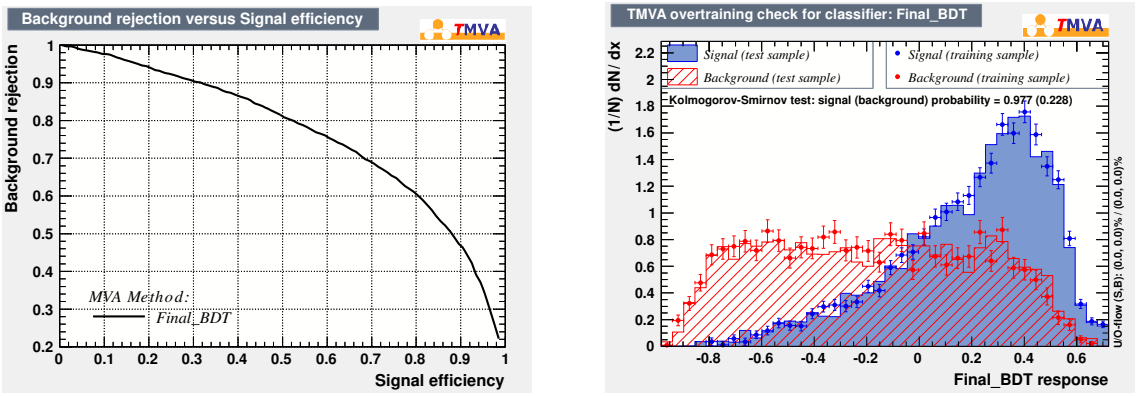
The optimum values of these parameters are given in Table 1.5.

The final BDT classifier output, overtraining check and ROC curve are shown in Figure 1.7. The final PID DLL and BDT response cuts are given in Table 1.6. Assuming the naive expected ratio of doubly-Cabibbo suppressed to Cabibbo favoured decays, the maximum $\mathrm{S}/\sqrt{\mathrm{S}+\mathrm{B}}$ significance of the promptly selected $\Lambda_c^+ \to p\pi^- K^+$ may be estimated. After the final PID cuts, this is done by summing the signal and background weights and varying the BDT response to find the optimum cut.

The tighter offline PID cuts used in the selection of all prompt modes (along with the semileptonic PID cuts) are given in Table 1.6.

Table 1.5: The BDT architecture variables which were optimised, showing the default and optimum values.

Architecture Variable	Default Value	Tested Values	Optimum Value
Boosting Type	Adaptive	Adaptive, Gradient, Bagging	Adaptive
Adaptive Boost β	0.5	$0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 1.0$	0.3
N_{trees}	400	$10, 20, 40, 50, 80, 100, 150, 200, 300, 400$	50
Maximum tree depth	3	$2, 3, 4, 5$	3
Steps in Node Optimisation	20	$10, 12, 14, 16, 18, 20, 22, 24, 26$	20



(a) (b)

Figure 1.7: The receiver operating characteristic (ROC) curve (a) and the classifier output distributions for the training/testing phases (b) for the final BDT. The Kolmogorov-Smirnov tests for compatibility in the training and testing phases indicate that the training and testing sample distributions agree in both signal and background.

Offline selection of semileptonic candidates

The higher separation from the primary vertex in semileptonically produced Λ_c over those produced promptly, due to the high mean lifetime of the Λ_b^0 (approximately 1500 fs), results in a much lower combinatoric background in the semileptonic data. Investigation of multivariate analysis and of further rectangular cuts on kinematic properties of the Λ_b^0 and Λ_c^+ decays over those in the stripping demonstrated no significant gains in discrimination between signal and combinatoric backgrounds. Tighter cuts on the PID DLL variables compared to those in the stripping were shown to offer significant discrimination. As such PID DLL cuts were optimised in CROP for maximum sensitivity to the $\Lambda_c^+ \to p\pi^-K^+$ mode, with the optimum cuts shown in Table 1.6. It should be noted that the tighter PID cuts in the semileptonic $\Lambda_c^+ \to p\pi^-\pi^+$ come from the stripping line for this mode, and are very sub-optimal for the other $\Lambda_c^+ \to phh'$ modes. As such these are the only differences in offline selection between the semileptonic modes.

Selection	Particle	PID Cuts
Prompt $\Lambda_c^+ \to phh'$	p	$\mathrm{DLL}_{(\mathrm{p}-\pi)} > 20, \mathrm{DLL}_{(\mathrm{p}-\mathrm{K})} > 12$
	K	$\mathrm{DLL}_{(\mathrm{K}-\pi)} > 10, \mathrm{DLL}_{(\mathrm{K}-\mathrm{p})} > -8$
	π	$\mathrm{DLL}_{(\mathrm{K}-\pi)} < 0$
SL $\Lambda_c^+ \to p\pi^-\pi^+$	p	$\mathrm{DLL}_{(\mathrm{p}-\pi)} > 20, \mathrm{DLL}_{(\mathrm{p}-\mathrm{K})} > 9$
	π	$\mathrm{DLL}_{(\mathrm{K}-\pi)} < 0$
Other SL $\Lambda_c^+ \to phh'$	p	$\mathrm{DLL}_{(\mathrm{p}-\pi)} > 20, \mathrm{DLL}_{(\mathrm{p}-\mathrm{K})} > 9$
	K	$\mathrm{DLL}_{(\mathrm{K}-\pi)} > 10$
	π	$\mathrm{DLL}_{(\mathrm{K}-\pi)} < 10$

Table 1.6: The final PID cuts used in each of the selections in the analysis.

1.3.4 Mass spectra

The mass spectra at each stage of selection are given herein to demonstrate the effects of the selection at each stage. These are given in Figure 1.8 – Figure 1.13. The mass range is centred on the nominal Λ_c mass $\pm$ 80 MeV/c^2. The terms in the legend are as follows:

Raw - The mass spectra of all stripped candidates, i.e. the raw output of the stripping selections.

Trigger - The mass spectra of the stripped candidates which pass the trigger requirements.

Trigger and PID - The mass spectra of the stripped candidates which pass the trigger requirements and the tight offline PID cuts.

Full Selection - The mass spectra of the stripped candidates passing the trigger, offline PID, kinematic vetoes and DTF convergence requirement.

In the prompt, the TIS trigger chain vetoes very few candidates. This is expected due to the lack of dedicated Hlt2 lines for the Cabibbo-suppressed modes and the inefficiency of the Cabibbo-favoured dedicated Hlt2 line - what prompt $\Lambda_c^+ \to phh'$ LHCb recorded in 2011 was almost entirely triggered independently of the Λ_c decay. In the semileptonic, the effect of the TOS chain is to veto a large amount of background while vetoing a relatively lower signal fraction - this is to be expected given the signal-enhancing cuts in the topological trigger. In both analyses the background is reduced considerably with the application of the PID cuts. The kinematic vetoes can be seen to veto signal and background indiscriminately, while the DTF convergence requirement results in candidates at high masses being vetoed in those channels with kaons in the final state.

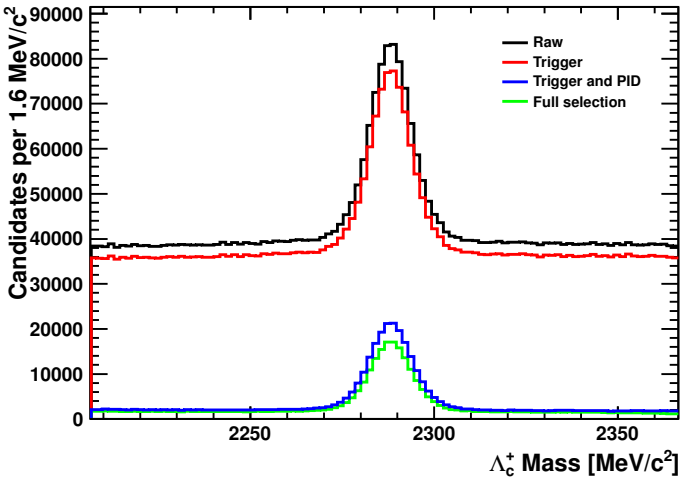


Figure 1.8: The mass distributions of the prompt $\Lambda_c^+ \to pK^-\pi^+$ for each stage of selection.

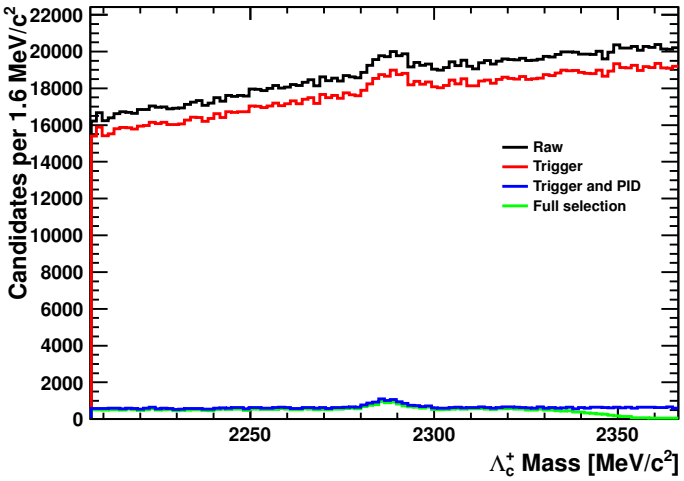


Figure 1.9: The mass distributions of the prompt $\Lambda_c^+ \to pK^-K^+$ for each stage of selection.

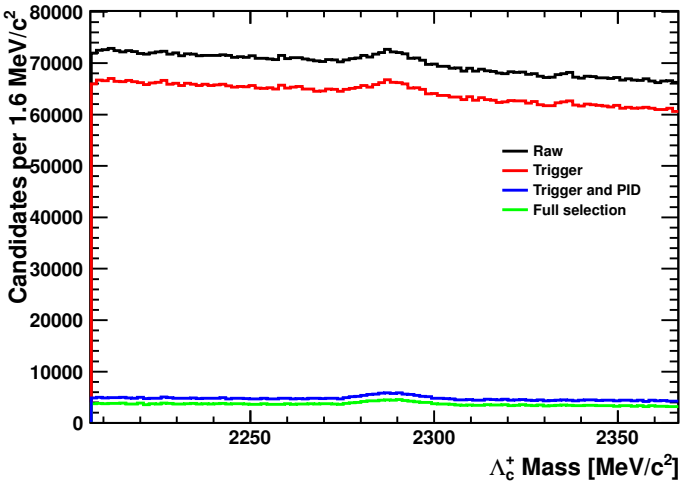


Figure 1.10: The mass distributions of the prompt $\Lambda_c^+ \to p\pi^-\pi^+$ for each stage of selection.

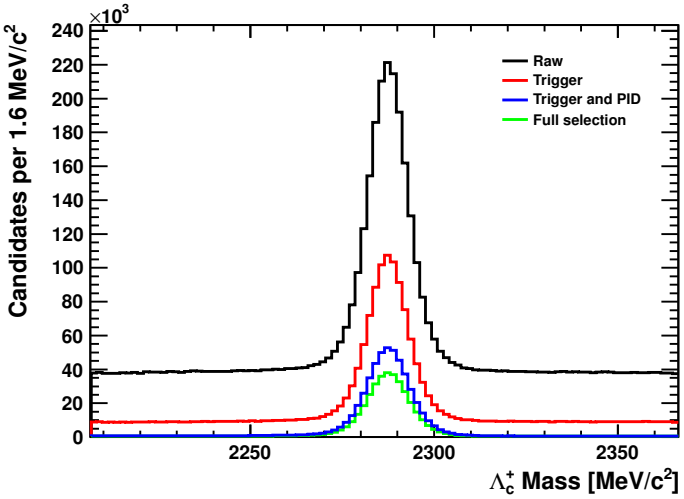


Figure 1.11: The mass distributions of the SL $\Lambda_c^+ \to pK^-\pi^+$ for each stage of selection.

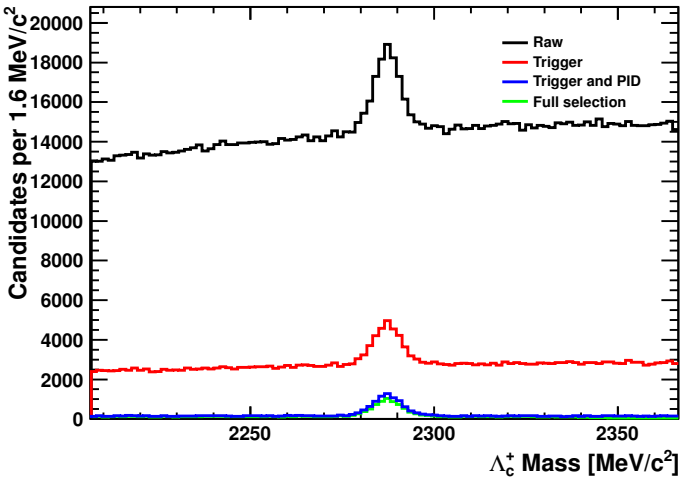


Figure 1.12: The mass distributions of the SL $\Lambda_c^+ \to pK^-K^+$ for each stage of selection.

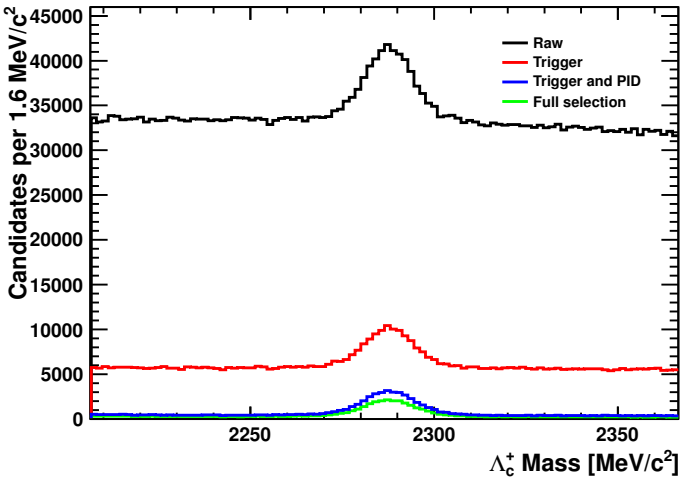


Figure 1.13: The mass distributions of the SL $\Lambda_c^+ \to p\pi^-\pi^+$ for each stage of selection.

1.4 Prompt backgrounds

Herein we discuss the potential backgrounds in the prompt analysis. There are two broad categories of backgrounds that our selections may be prone to: reflections from D^+ decays and Λ_c decays. There are a number of charmed meson decays which, if mis-reconstructed in some fashion, may result in peaking backgrounds in the |lc candidate mass window around the nominal. A brief summation of these are as follows (in all cases h refers to any permutation of charged kaons and pions):

$D^+ \to hhh / D_s^+ \to hhh$ **reflections** - all the D^+/D_s^+ daughters are reconstructed, but at least one of the daughters are misidentified as another charged hadron.

$D^+ \to hhhh^0 / D_s^+ \to hhhh^0$ **partial reconstruction** - the neutral daughter hadron, for example a π^0 or $K^0_{\rm S}$, is not reconstructed and at least one of the charged daughters is misidentified as another charged hadron.

$D^0 \to hhhh$ **partial reconstruction** - one charged daughter is not reconstructed, and at least one of the charged daughters is misidentified as another charged hadron.

The misidentification has the effect of smearing the candidate mass distribution and shifting it from the nominal Λ_c mass in non-trivial ways depending on the decay kinematics. To parametrise these distributions we apply our selection to simulated D decays while forcing the required mass hypotheses on the set of D daughters such that candidates manifest with a final state corresponding to one of the $\Lambda_c^+ \to phh'$ modes under investigation.

1.4.1 $D_s^+ \to hhh / D^+ \to hhh$ mis-reconstructions

Method

There are a large number of D^+/D_s^+ decays which, with a single final state kaon or pion mis-identified as a proton, can lead to their manifestation in the same final state as one of the $\Lambda_c^+ \to phh'$ modes. To account for these $D^+ \to hhh / D_s^+ \to hhh$ reflections, we consider the mis-identifications which can arise which will likely result in candidates in the region of the Λ_c mass.

The possible reflections are constrained by one of the same-sign daughters being mis-identified as a proton, resulting in only four generic cases whereby single mis-identification reflections are relevant. As the number of decays which can result in such a mis-identification are numerous, we investigate first the modes which have the highest branching fractions corresponding to these general cases of mis-identification.

These decays are outlined in Table 1.7. The most abundantly produced D^+ mode in which a pion mis-identification as proton is possible is $D^+ \to \pi^+\pi^+K^-$, with a measured branching fraction of (9.13 ± 0.19) % [19]. D^+ decays with a same-sign kaon in the final state require at least one Cabibbo-suppressed transition, and their branching fractions are correspondingly lower. The most abundant of these is the decay $D^+ \to \pi^+K^-K^+$, with a branching fraction of $(9.54 \pm 0.26) \times 10^{-3}$ [19]. The highest branching fraction of a D_s^+

with a same-sign pion or kaon in the final state is in both cases the decay $D_s^+ \to \pi^+ K^- K^+$, with a branching fraction of (5.49 ± 0.27) % [20], [21].

$D^+ \to hhh$ Mode	D^+ decay $\mathcal{B}$ [%]	mis-ID	$\Lambda_c^+ \to phh'$ Mode
$D^+ \to \pi^+\pi^+K^-$	$(9.13 \pm 0.19) \times 10^{-2}$	$\pi^+ \to p$	$\Lambda_c^+ \to pK^-\pi^+$
$D^+ \to K^+K^-\pi^+$	$(9.54 \pm 0.26) \times 10^{-3}$	$K^+ \to p$	$\Lambda_c^+ \to pK^-\pi^+$
$D_s^+ \to K^+K^-\pi^+$	$(5.49 \pm 0.27) \times 10^{-2}$	$\pi^+ \to p$	$\Lambda_c^+ \to pK^-K^+$
$D_s^+ \to K^+K^-\pi^+$		$K^+ \to p$	$\Lambda_c^+ \to pK^-\pi^+$

Table 1.7: The decays which correspond to a $D^+ \to hhh/D_s^+ \to hhh$ reflection where a positive charged kaon or pion is mid-identified as a proton. We provide the branching fractions and the mis-ID corresponding to the $\Lambda_c^+ \to phh'$ modes in which the reflections manifest. The decays given possess some of the highest branching fractions of the D^+/D_s^+ decays which can result in a $\Lambda_c^+ \to phh'$ final state under a single misidentification.

Efficiency and yield estimations

The numbers of simulated events passing the various stages of our selection are detailed in Table 1.8. For all modes, zero events pass the full selection. A finite number of candidates pass just the trigger and stripping, and a finite number of candidates pass just the stripping and PID. As the trigger chain in the prompt analysis is TIS, the subset of events passing the TIS requirement should not be dependent on the PID response. We can therefore factorise the individual efficiencies to better estimate the overall efficiency of our selection on the backgrounds. The full efficiency of the selection is therefore given as:

$$\epsilon_{total} = \frac{\epsilon_{(\text{strip,PID})}\epsilon_{(\text{strip,trig})}}{\epsilon_{strip}} \tag{1.2}$$

where ϵ_{strip} is the efficiency of the stripping selection alone, $\epsilon_{(\text{strip,PID})}$ is the efficiency of the stripping and PID selections and $\epsilon_{(\text{strip,trig})}$ is the efficiency of the stripping and trigger selections (as the stripping selection is applied to the baseline already it is necessary to include it in both the expressions for the PID and trigger efficiencies and factor it out). The efficiencies for the modes are shown in Table 1.9. Where non-zero numbers of events pass orthogonal selection, we provide an efficiency as per Equation 1.2. If no events survive any orthogonal selections, we provide upper limits. These efficiencies are rough calculations - the PID efficiency is derived using the PID response in the simulation, which is known to be poorly modelled. The numbers of events surviving the selections are very low, such that the binomial errors on the efficiencies are high. Nonetheless, they serve as a reasonable approximation when estimating the orders of magnitude of the efficiencies of our prompt selection on the D decays. The efficiencies for the full prompt selection are estimated to be of order $\mathcal{O}(10^{-9})$ – $\mathcal{O}(10^{-8})$ for the various D decays considered.

The prompt charm cross sections at $\sqrt{s}$ 7 TeV at LHCb have been measured in [22]. Integrated over the LHCb acceptance, those relevant to this analysis are:

- $\sigma(\Lambda_c) = 230 \pm 77\ \mu$b
- $\sigma(D^+) = 676 \pm 137\ \mu$b
- $\sigma(D_s^+) = 194 \pm 38\ \mu$b
- $\sigma(D^0) = 1488 \pm 182\ \mu$b

Taking these, with the estimated efficiencies of our selection on the above D decays, we can estimate the number of charged D reflections surviving the prompt Λ_c selection in the 2011 dataset. The typical expected D candidates passing our selection are estimated as:

$$N_{\text{pass}} = \mathcal{L}\sigma\mathcal{B}\epsilon_D \tag{1.3}$$

where $\mathcal{L}$ is the integrated luminosity of the 2011 dataset, σ is the production cross section in the LHCb acceptance of the parent D, $\mathcal{B}$ is the branching fraction of the D decay leading to the reflection and ϵ_D is the estimated efficiency of the prompt Λ_c selection on the D decay. These are summarised in Table 1.10. The estimations for N_{pass} are of the order of 1000 for the decays considered.

This somewhat low yield is to be expected given our selection's design. Tight PID cuts are used on every final state daughter in the prompt selection. In addition to suppressing combinations of unrelated tracks (of which there are many, due to the short Λ_c mean lifetime and low PV displacement), these serve to eliminate the majority of mis-identified reflections. The maximum lifetime cut on the reconstructed Λ_c candidate of 1.2 ps removes a significant fraction (approximately 35 %) of genuine D decays, which have a mean lifetime of around 1 ps compared to 0.2 ps for the Λ_c. The minimum momentum requirement on the proton of 15.6 GeV/c and the requirement that at least one daughter have a transverse momentum above 1.2 GeV/c suppress the kaon/pion tracks which come from D decays (the proton momenta in genuine Λ_c decays are typically greater than those of the D decay daughters).

There are many components of our selection which will apply equally well to excluding other D decays. The lifetime cut will uniformly veto the same fraction of all D decays. The number of kaons and pions in the final state of each D decay will affect the momenta of the final state particles, although this is not expected to be a large effect. As such, to first order the efficacy of the momentum and PID requirements in vetoing other D decays which have not been directly examined should be comparible to the modes which were investigated. All other D decays which can result in final state reflections in the Λ_c reconstruction have lower branching fractions than the decays investigated herein, resulting in lower expected contamination in the 2011 Λ_c data. In the case of reflections arising from two or three mis-identifications, we may assume that the tight PID cuts make it significantly less likely that a D decay with multiple incorrectly identified daughter tracks will pass our selection than in the case of a single mis-identification.

The number of reflection candidates expected to pass the selection for the highest D^+ and D_s^+ branching fractions should not be a significant factor in our fitting procedure if, as expected, they are spread across the full Λ_c mass window. We also provide the

ratio of the number of reflection candidates expected to pass full selection in 2011 to the number of candidates in the corresponding $\Lambda_c^+ \to phh'$ signal mode in which the reflection manifests. Given that 3 of the 4 reflections possible from the high branching ratio D^+/D_s^+ decays considered manifest in the Cabibbo-favoured mode these reflection over signal ratios are generally at the per-cent level, and negligible in this analysis given other limiting systematics.

$D^+ \to hhh$ Mode	N_{total}	mis-ID	N_{strip}	$N_{strip,PID}$	$N_{strip,trig}$	$N_{strip,trig,pid}$
$D^+ \to \pi^+\pi^+K^-$	5128979	$\pi^+ \to p$	37492	40	125	0
$D^+ \to K^+K^-\pi^+$	10150948	$K^+ \to p$	15141	41	329	0
$D_s^+ \to K^+K^-\pi^+$	1007498	$\pi^+ \to p$	332	0	2	0
		$K^+ \to p$	41	0	7	0

Table 1.8: The simulated data used in the single mis-identification background studies. Shown are the numbers of events within 90 MeV/c^2 of the Λ_c mass passing the prompt Λ_c selection at each stage of selection.

$D^+ \to hhh$ Mode	mis-ID	ϵ_{strip}	$\epsilon_{(PID,strip)}$	$\epsilon_{(trig,strip)}$	ϵ_{total}
$D^+ \to \pi^+\pi^+K^-$	$\pi^+ \to p$	$(7.309 \pm 0.037) \times 10^{-3}$	$(7.8 \pm 1.2) \times 10^{-6}$	$(2.4 \pm 0.2) \times 10^{-5}$	$(2.6 \pm 0.5) \times 10^{-8}$
$D^+ \to K^+K^-\pi^+$	$K^+ \to p$	$(1.491 \pm 0.001) \times 10^{-3}$	$(4.0 \pm 0.6) \times 10^{-6}$	$(3.2 \pm 0.2) \times 10^{-5}$	$(8.8 \pm 1.5) \times 10^{-8}$
$D_s^+ \to K^+K^-\pi^+$	$\pi^+ \to p$	$(3.295 \pm 0.181) \times 10^{-4}$	$< 1.3 \times 10^{-7}$ @ 95 % CL	$(2.0 \pm 1.4) \times 10^{-6}$	$< 1.3 \times 10^{-7}$ @ 95 % CL
	$K^+ \to p$	$(4.069 \pm 0.635) \times 10^{-5}$	$< 1.3 \times 10^{-7}$ @ 95 % CL	$(7.0 \pm 2.6) \times 10^{-6}$	$< 1.3 \times 10^{-7}$ @ 95 % CL

Table 1.9: The simulated data used in the single mis-identification background studies. Shown are the efficiencies for the various stages of the selection and the calculated efficiency of the full selection.

$D^+ \to hhh$ Mode	$\sigma(mother)/\sigma(\Lambda_c)$	Decay $\mathcal{B}$	mis-ID	N_{final}	$\Lambda_c^+ \to phh'$ signal fraction [%]
$D^+ \to \pi^+\pi^+K^-$	2.939 ± 1.150	$(9.13 \pm 0.19) \times 10^{-2}$	$\pi^+ \to p$	$(1.6 \pm 0.5) \times 10^3$	1.2
$D^+ \to K^+K^-\pi^+$		$(9.54 \pm 0.26) \times 10^{-3}$	$K^+ \to p$	$(0.6 \pm 0.2) \times 10^3$	0.5
$D_s^+ \to K^+K^-\pi^+$	0.084 ± 0.327	$(5.49 \pm 0.27) \times 10^{-2}$	$\pi^+ \to p$	$< 1.4 \times 10^3$ @ 95 % CL	51
			$K^+ \to p$	$< 1.4 \times 10^3$ @ 95 % CL	1.1

Table 1.10: The estimated number of candidates from each charged D decay mode expected to pass the prompt Λ_c selection in the 2011 dataset. Also given are the ratio of this number of candidates to the number of candidates in the corresponding $\Lambda_c^+ \to phh'$ signal mode in which the reflection manifests.

Simulated mass distributions

We use the simulated data to investigate the shapes of the mass distributions of misreconstructed candidates. If the distributions do not peak within the Λ_c mass window, and have shapes which can be well modelled by a first order polynomial, it may be safely assumed that even if the number of D decays which pass our selection in this mass region is much higher in reality than our calculations say, this contribution will be accounted for by the fit model for the combinatoric background component.

As no simulated D decays pass our full selection, we investigate the shape of the mass distribution of the mis-reconstructed D candidates at various stages of the selection. We investigate the mass distribution after the stripping selection only, the mass distribution after the stripping and PID selections, and the mass distribution after the trigger and stripping selections.

An example of the D^+ reflection mass distribution, $D^+ \to K^-\pi^+\pi^+$, π^+ mis-ID as p, is given in Figure 1.14 for the full candidate mass distribution, with equivalent plots of these quantites in just the Λ_c^+ stripping mass window given in Figure 1.15. As can be seen, the candidates after stripping have mass distributions which either peak in the Λ_c mass region (albeit with very broad peaks) or which are not flat in this region. As can be seen in the distributions for candidates after stripping and PID and the distributions after stripping and trigger these structures become flatter as the selection is applied. It is therefore expected that the distributions of charged D reflections passing the full Λ_c selection will be well modelled by a first order polynomial. Our combinatoric background is also well described by a first order polynomial, and so any sum of these will simply be a new first order polynomial, for which our fit model will account.

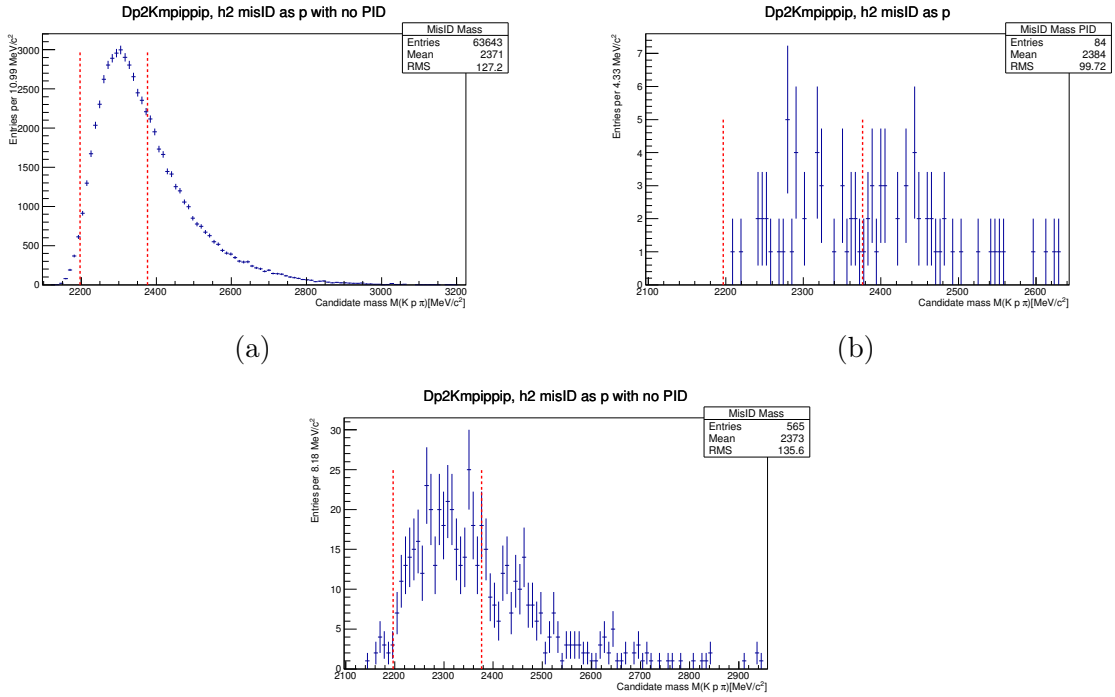


Figure 1.14: The mass distributions of simulated $D^+ \to K^-\pi^+\pi^+$ candidates with the proton mass hypothesis forced upon one of the daughter pions. Shaded red lines show the Λ_c mass region from the stripping. (a) has applied the stripping selection alone, (b) the stripping and PID, (c) the stripping and trigger.

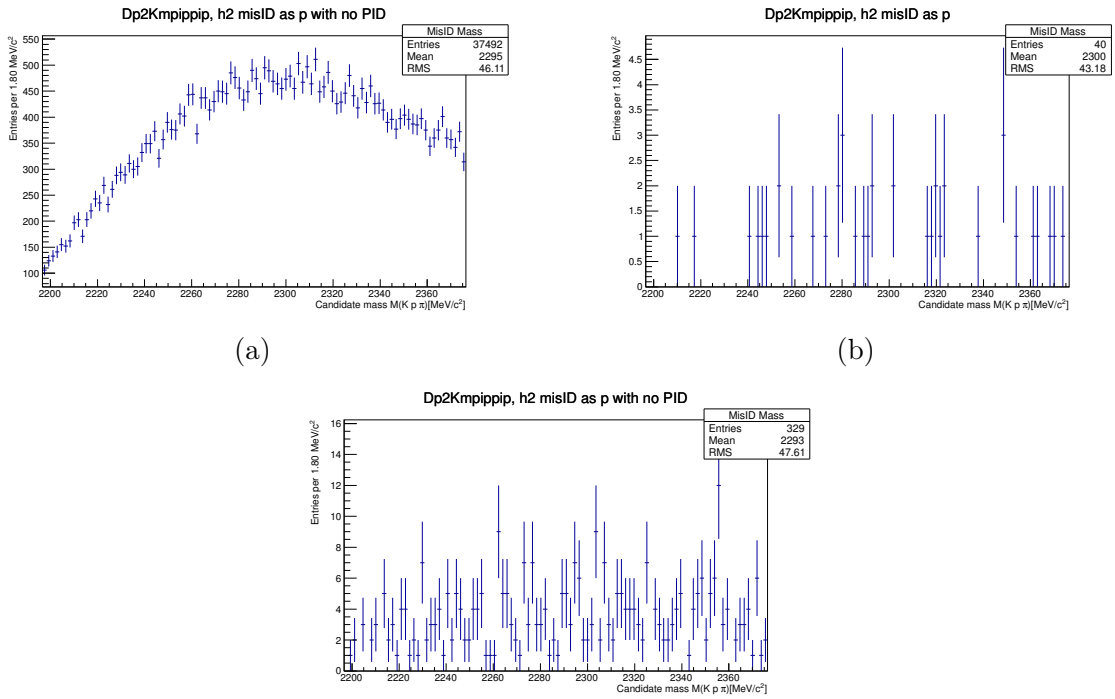


Figure 1.15: The mass distributions of simulated $D^+ \to K^-\pi^+\pi^+$ candidates with the proton mass hypothesis forced upon one of the daughter pions. The mass region shown is that applied in the $\Lambda_c^+ \to phh'$ stripping. (a) has applied the stripping selection alone, (b) the stripping and PID, (c) the stripping and trigger.

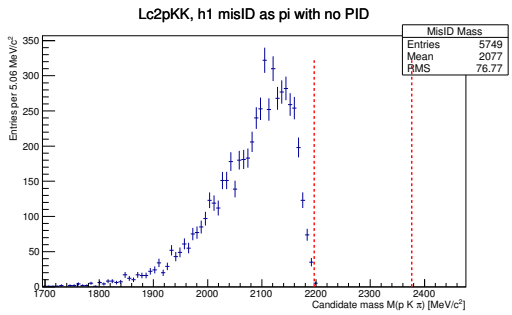
1.4.2 Mis-ID in Λ_c decays

In the case of cross-feed between the different $\Lambda_c^+ \to phh'$ modes, single mis-identification of any daughter hadron is expected to result in a reconstructed mass distribution peaking away from the Λ_c mass. Double and triple mis-identification in the singly Cabibbo-suppressed modes is also expected to take the reconstructed candidate outside the mass range of interest. For the Cabbibo favoured and doubly-Cabibbo suppressed modes, when the pion is mis-identified as a kaon and the kaon is mis-identified as a pion, the mass distribution peaks close to the Λ_c mass, although the distribution will be highly smeared from the double mis-identification.

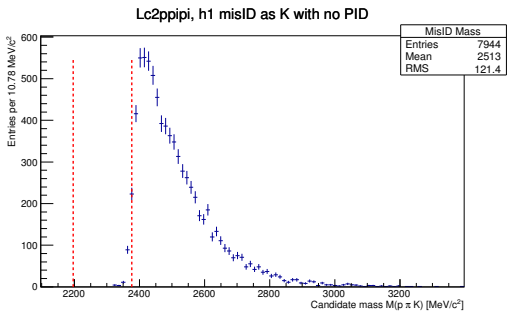
As the $\Lambda_c^+ \to phh'$ real data has PID cuts applied at the stripping level, we use simulated $\Lambda_c^+ \to phh'$ decays for these studies. The samples have undergone the stripping selection with PID cuts removed as described in Section 1.2.2. As in the case of the studies of D reflections, we force the wrong mass hypothesis on one or more of the truth matched signal decay daughters and recompute the invariant mass of the parent candidate.

Single mis-ID

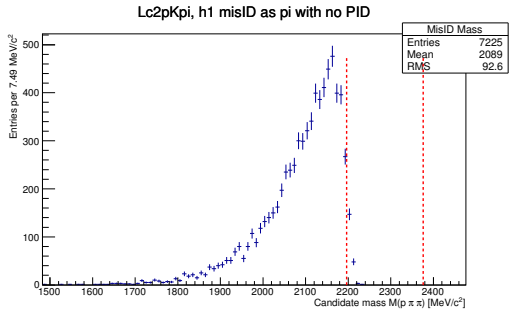
The case of a single daughter mis-identification, as expected, results in mass distributions outside of the signal region. These cases are shown in Figure 1.16, where no PID or trigger requirements are placed on the samples. As the mass distributions all peak outwith the range of interest we conclude that such cross feed between the $\Lambda_c^+ \to phh'$ modes is not a concern.



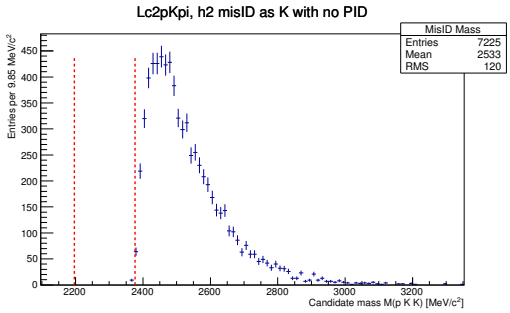
(a) $\Lambda_c^+ \to pK^-K^+$, $K^- \to \pi^-$



(b) $\Lambda_c^+ \to p\pi^-\pi^+\pi^- \to K^-$



(c) $\Lambda_c^+ \to pK^-\pi^+$, $K^- \to \pi^-$



(d) $\Lambda_c^+ \to pK^-\pi^+$, $\pi^+ \to K^+$

Figure 1.16: The mass distributions of simulated $\Lambda_c^+ \to phh'$, with the wrong mass hypothesis forced on one of the daughter mesons. The mass window indicated in red is that applied in the $\Lambda_c^+ \to phh'$ stripping. The selection applied is a version of the stripping with all PID cuts removed, and no trigger requirements have been made.

1.4.3 Reassigning mass hypotheses in data

As a more direct check we reassign the mass hypotheses of the daughter particles in the prompt $\Lambda_c^+ \to phh'$ data, to those mass hypotheses which correspond to the final states of all possible D reflections. We show an example of these checks for $\Lambda_c^+ \to pK^-\pi^+$ with the proton mass hypothesis changed to a kaon in Figure 1.17. No sharply peaking structures at either the D^+ or D_s^+ mass are found in any of the $\Lambda_c^+ \to phh'$ data. We take this as evidence that our final datasets correspond to only candidates which are genuine Λ_c signal and candidates which are formed from unrelated tracks, and that reflection misreconstructions of D decays are sufficiently suppressed in our selections.

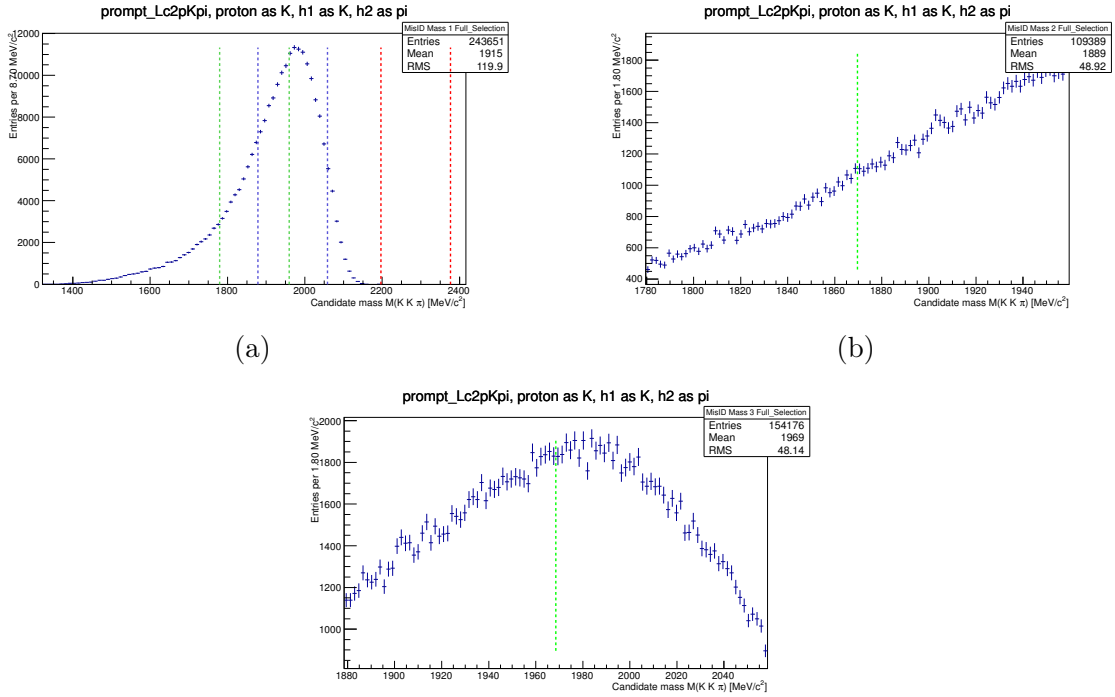


Figure 1.17: The candidate mass distribution of prompt $\Lambda_c^+ \to pK^-\pi^+$ data after final selection, where the p mass hypothesis is changed to that of the K^+. Shown in (a) is the whole mass distribution, with the D^+ mass region denoted in green with a zoom of this region in (b) and the D_s^+ mass region denoted in blue with a zoom of this region in (c). This probes for the presence of reflections from mis-identified $D^+/D_s^+ \to K^+K^-\pi^+$.

1.4.4 Summary

The likely peaking backgrounds from misidentified $D^+ \to hhh/D_s^+ \to hhh$ and $\Lambda_c^+ \to phh'$ decays have been investigated. We do not also explicitly investigate the case of partially reconstructed $D^+ \to hhhh^0/D_s^+ \to hhhh^0$- the missing neutral particle's momentum

will result in mass distributions which are lower compared to those resulting from mis-identification of fully reconstructed charged D decays. The variation in the missing neutral's momentum will also result in a higher smearing of the mass distributions. The branching fractions of four-body D decays are generally significantly lower than their three-body counterparts, leading to further suppression of possible candidates resulting from this mis-reconstruction. These factors indicate that any cross-feed from $D^+ \to hhhh^0/D_s^+ \to hhhh^0$ decays will be less significant than in the case of $D^+ \to hhh/D_s^+ \to hhh$ decays.

Our investigations using simulated data of the mass shapes indicate that distributions of any prominent reflections should be approximately linear across the Λ_c signal region. Our estimations using simulation of reflection yields passing the prompt selection place estimates on the number of D^+ reflections passing the selection. These numbers are of the order 10^3, although the errors on these estimates are of order 35 %. We are only able to place conservative upper limits on D_s^+ reflections using simulation of 10^5. As such we investigate the mass shapes of our real data with candidate mass hypotheses reassigned. We find no evidence of D reflection backgrounds contaminating our final data sets.

We do not explicitly investigate the corresponding cases of misidentified $D^0 \to hhhh$ decays where one charged daughter is not reconstructed. The features of the prompt selection which make it so robust against charged D reflections also make it robust against partial reconstruction of neutral D mesons. The only difference between the two is that the reconstructed charged daughter kinematics will be lower in the case of neutral D decays due to the missing charged daughter's missing momentum. Our selection should therefore have a lower efficiency in the selection of these candidates compared to the investigated charged D decays due to the minimum daughter momentum requirements. While the measured cross section for D^0 is higher than that of D^+ or D_s^+, we still estimate that the number of reflection candidates from these sources is negligible.

We investigate cross-feed between the $\Lambda_c^+ \to phh'$ modes. In single mis-ID cases we expect this to be negligible due to the shifting of the reconstructed candidate mass outside the signal mass window. *Comment on double-mis ID wil be here when we have the DCS distributions*.

We conclude that the only relevant background in the prompt analysis will be non-peaking combinations of unrelated tracks. This informs our choice of fit model, outlined in Section **??**.

1.5 Semileptonic backgrounds

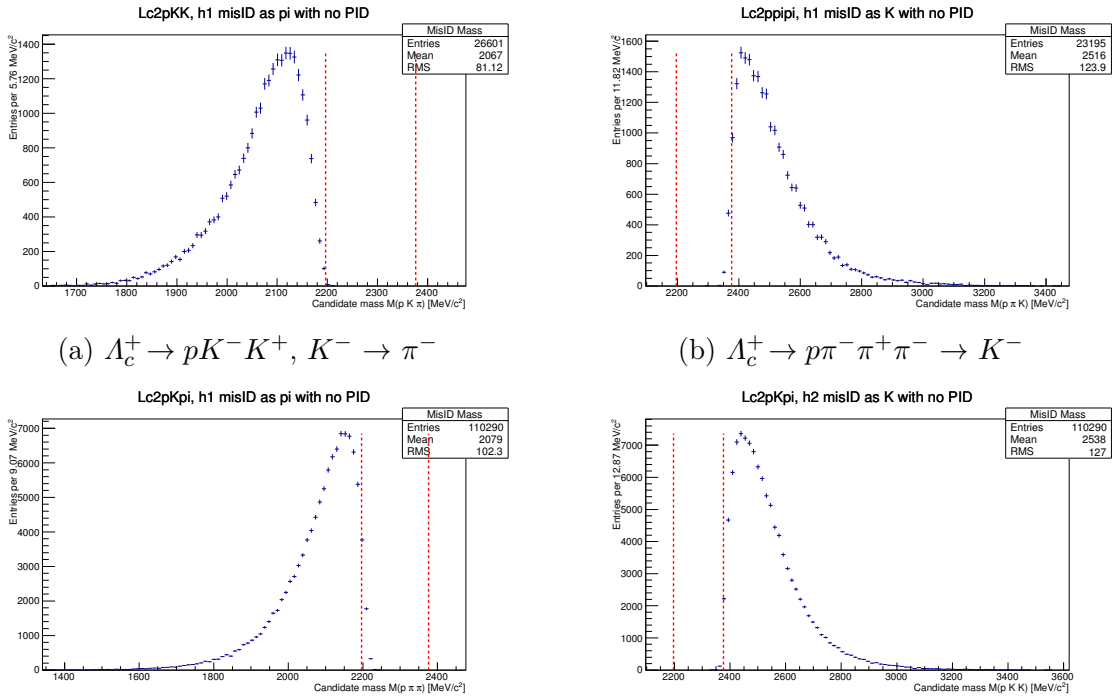
Herein we discuss the potential backgrounds in the main semileptonic TOS analysis.

1.5.1 Mis-ID in Λ_c decays

As in the case of the prompt analysis, we investigate cross-feed between the $\Lambda_c^+ \to phh'$ decay modes.

Single mis-ID

The case of a single daughter mis-identification, as in the prompt analysis, results in mass distributions outside of the signal region. These cases are shown in Figure 1.18, where no PID or trigger requirements are placed on the samples. As the mass distributions all peak outwith the range of interest we conclude that such cross feed between the $\Lambda_c^+ \to phh'$ modes is not a concern.



(a) $\Lambda_c^+ \to pK^-K^+$, $K^- \to \pi^-$

(b) $\Lambda_c^+ \to p\pi^-\pi^+\pi^- \to K^-$

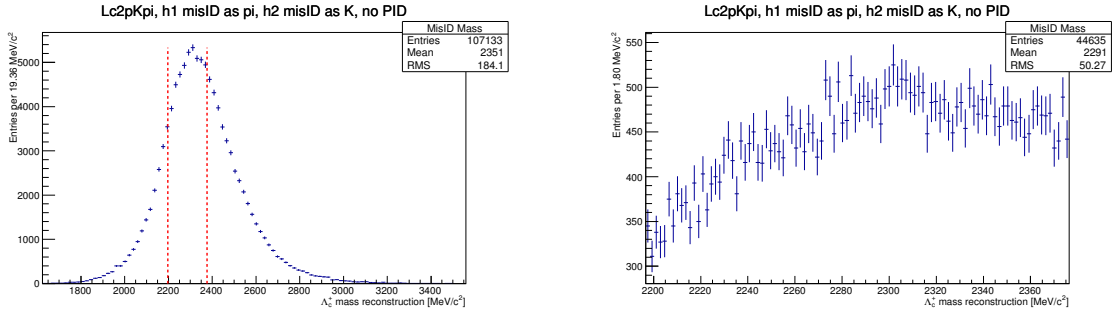
(c) $\Lambda_c^+ \to pK^-\pi^+$, $K^- \to \pi^-$

(d) $\Lambda_c^+ \to pK^-\pi^+$, $\pi^+ \to K^+$

Figure 1.18: The mass distributions of simulated $\Lambda_b^0 \to \Lambda_c^+\mu^-\overline{\nu}_\mu$, $\Lambda_c^+ \to phh'$, with the wrong mass hypothesis forced on one of the daughter mesons. The mass region indicated in red is that applied in the $\Lambda_b^0 \to \Lambda_c^+\mu^-\overline{\nu}_\mu$, $\Lambda_c^+ \to phh'$ stripping. The selection applied is a version of the stripping with all PID cuts removed, and no trigger cuts have been applied.

Double mis-ID

As in the prompt analysis, we investigate the case of double mis-ID in the Cabibbo favoured mode, where both daughter mesons are incorrectly identified. The full mass distribution of these candidates and a close-up of the Λ_c mass window in the stripping selection is shown in Figure 1.19, where no PID or trigger requirements have been placed on the reconstructed candidates.



(a) (b)

Figure 1.19: The mass distributions of simulated $\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu, \Lambda_c^+ \to pK^-\pi^+$, with the wrong mass hypothesis forced on both of the daughter mesons. The mass region indicated in red in (a) is that applied in the $\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu, \Lambda_c^+ \to phh'$ stripping, (b) shows just this region. The selection applied is a version of the stripping with all PID cuts removed, and no trigger cuts have been applied.

1.5.2 Reassigning mass hypotheses in data

As a more direct check we reassign the mass hypotheses of the daughter particles on the semileptonic $\Lambda_c^+ \to phh'$ data to check for sharply peaking structures, which would be indicative of D reflections surviving the selection. No structures are observed. We show an example of these checks for $\Lambda_c^+ \to pK^-\pi^+$ with the proton mass hypothesis changed to a kaon in Figure 1.20.

1.5.3 Summary

It was considered unnecessary to investigate in simulation the case whereby a $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu, D^+ \to hhh$ decay is mis-reconstructed as the charge conjugate of a $\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu, \Lambda_c^+ \to phh'$ decay. The shape of the candidate mass spectra resulting from such a mis-ID should be similar to the equivalent reflections in the prompt analysis. Despite the higher TOS trigger efficiencies of D mesons and the lower PID cuts no evidence of reflections from $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu, D^+ \to hhh$ decays was found in checks with the real data. As such we assume that our selection suppresses reflections from other charm decays such

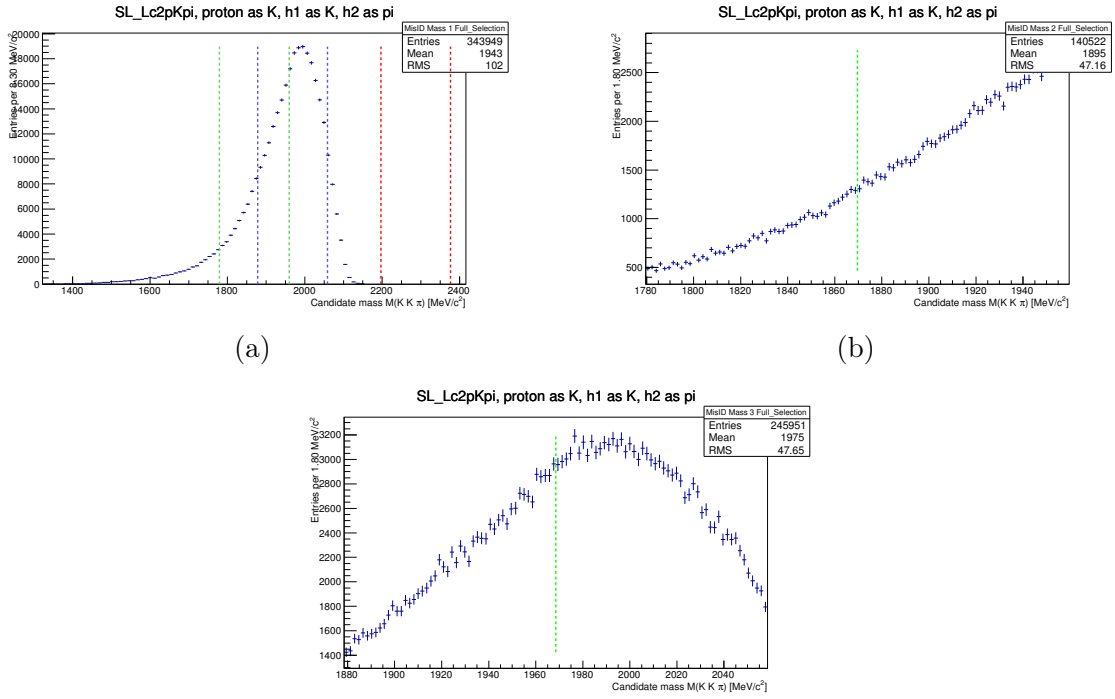


Figure 1.20: The mass distribution of SL $\Lambda_c^+ \to pK^-\pi^+$ data after final selection, where the p has been mis-identified as a K^+. Shown top is the whole mass distribution, with the D^+ mass region denoted in green (bottom left plot) and the D_s^+ mass region denoted in blue (bottom right plot). This probes for the presence of reflections from mis-identified $D^+/D_s^+ \to K^+K^-\pi^+$.

that they are negligible. Upon unblinding the DCS mode we will investigate the mass distributions of double mis-ID candidates to check for contamination.

The conclusions we make regarding the SL backgrounds are broadly similar to those we state for the prompt in Section 1.4.4. We conclude that we have through examination of simulated and real data shown that our selections sufficiently suppress the backgrounds under consideration to a level whereby they may be considered negligible in the semileptonic analysis.

Bibliography

[1] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.

[2] E791 Collaboration, E. Aitala *et al.*, *Multidimensional resonance analysis of* $\Lambda_c^+ \to pK^-\pi^+$, Phys. Lett. **B471** (2000) 449, arXiv:hep-ex/9912003.

[3] ACCMOR Collaboration, A. Bozek *et al.*, *A Study of* Λ_c^+ *decays into* $pK^-\pi^+$, $pK^-\pi^+\pi^0$ *and* $pK^-\pi^+\pi^0\pi^0$, Phys. Lett. **B312** (1993) 247.

[4] M. Basile *et al.*, *A Measurement of two resonant contributions in the* Λ_c^+ *branching ratios*, Nuovo Cim. **A62** (1981) 14.

[5] ACCMOR Collaboration, S. Barlag *et al.*, *Measurement of frequencies of various decay modes of charmed particles* D^0, D^+, D_s^+ *and* Λ_c^+ *including the observation of new channels*, Z. Phys. **C48** (1990) 29.

[6] Belle Collaboration, K. Abe *et al.*, *Observation of Cabibbo suppressed and W exchange* Λ_c^+ *baryon decays*, Phys. Lett. **B524** (2002) 33, arXiv:hep-ex/0111032.

[7] CLEO Collaboration, J. Alexander *et al.*, *Observation of the Cabibbo suppressed charmed baryon decay* $\Lambda_c^+ \to p\phi$, Phys. Rev. **D53** (1996) 1013, arXiv:hep-ex/9508005.

[8] FOCUS Collaboration, J. Link *et al.*, *Search for* $\Lambda_c^+ \to pK^+\pi^-$ *and* $D_s^+ \to K^+K^+\pi^-$ *using genetic programming event selection*, Phys. Lett. **B624** (2005) 166, arXiv:hep-ex/0507103.

[9] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[10] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[11] M. Needham, *Clone track identification using the Kullback-Liebler distance*, .

[12] M. Pivk and F. L. Diberder, *A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (2005), no. 1âĂŞ2 356 .

[13] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[14] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*, Wadsworth and Brooks, Monterey, CA, 1984. new edition [**?**]?

[15] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classifcation and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[16] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119.

[17] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[18] J. Massey, Frank J., *The Kolmogorov-Smirnov Test for Goodness of Fit*, Journal of the American Statistical Association **46** (1951), no. 253 pp. 68.

[19] CLEO Collaboration, S. Dobbs *et al.*, *Measurement of absolute hadronic branching fractions of D mesons and e+ e- -> D anti-D cross-sections at the psi(3770)*, Phys. Rev. **D76** (2007) 112001, arXiv:0709.3783.

[20] CLEO Collaboration, P. Onyisi *et al.*, *Improved Measurement of Absolute Hadronic Branching Fractions of the D_s^+ Meson*, Phys. Rev. **D88** (2013), no. 3 032009, arXiv:1306.5363.

[21] Belle Collaboration, A. Zupanc *et al.*, *Measurements of branching fractions of leptonic and hadronic D_s^+ meson decays and extraction of the D_s^+ meson decay constant*, JHEP **1309** (2013) 139, arXiv:1307.6240.

[22] LHCb collaboration, R. Aaij *et al.*, *Prompt charm production in pp collisions at $\sqrt{s}$ = 7 TeV*, Nucl. Phys. **B871** (2013) 1, arXiv:1302.2864.