

Measurements of charm production and CP violation with the LHCb detector

A thesis submitted to The University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Science and Engineering

2016

Alex Pearce

Particle Physics Group
School of Physics and Astronomy



The University of Manchester



Contents

Abstract	9
Declaration	11
Copyright	13
Dedication	15
Acknowledgements	17
Preface	19
I Introduction	21
1 Overview	23
2 The Standard Model	25
2.1 CP violation	29
3 The Large Hadron Collider	33
4 The LHCb experiment	35
4.1 The LHCb Detector	37
4.2 Online event reconstruction and selection	47
4.3 Offline data flow and the upgrades for Run 2	50
4.4 Simulation	52

II	Charm production	55
5	Introduction	57
5.1	Analysis overview	58
6	Theory and motivation	63
6.1	Atmospheric neutrino fluxes	67
6.2	Cross-section comparisons with theory	69
7	Datasets and running conditions	73
7.1	Simulated data samples	74
7.2	Crossing angle correction	75
8	Event and candidate selection	77
8.1	Online selection and reconstruction	79
8.2	Offline selection	81
9	Yield extraction	87
9.1	Signal-background discrimination	88
9.2	Prompt-secondary discrimination	89
9.3	Mode-specific details	91
10	Efficiency evaluation	105
10.1	Detector acceptance	106
10.2	Reconstruction	107
10.3	Selection	113
10.4	Particle identification	115
10.5	Signal window	121
10.6	Uncertainties on efficiencies	122
11	Systematic uncertainties	127
11.1	Finite Monte Carlo sample size	127
11.2	Monte Carlo modelling	128
11.3	Monte Carlo truth matching	132
11.4	Particle identification calibration	132

11.5	Luminosity	134
11.6	Tracking efficiency correction	137
11.7	Branching fractions	138
11.8	Fit model	139
11.9	Signal window	139
11.10	Correlations and summary	140
12	Results	143
12.1	Production ratios	143
12.2	Integrated cross-sections	145
12.3	Interpretation and comparison with theory	146
12.4	Measurements at $\sqrt{s} = 5$ TeV	147
13	Summary	157
III	Direct CP violation in charm baryon decays	159
14	Introduction	161
14.1	Analysis overview	163
15	Formalism	165
15.1	Decay phase space	168
16	Datasets	171
16.1	Monte Carlo	173
17	Selection	175
17.1	Stripping and trigger	175
17.2	Offline	177
17.3	Background study	180
17.4	Multiple candidates	183
18	Preliminary fits	193
18.1	Input data and shape parameterisation	193
18.2	Validation	194

19	Kinematic weighting	199
19.1	Weighting with decision trees	200
19.2	Evaluation	201
19.3	Validation	202
19.4	Weight statistics	204
20	Phase space efficiency	221
20.1	Distributions in data	221
20.2	Efficiency modelling	221
21	Measuring A_{CP}^{Raw}	231
21.1	Simultaneous fit	232
21.2	Blinding	233
21.3	Fit results	234
21.4	Validation	238
22	Results	241
22.1	Combining measurements	241
23	Systematic uncertainties	245
23.1	Fit model	245
23.2	Residual background asymmetries	247
23.3	Signal decays from other sources	255
23.4	Summary and combination	256
24	Summary	257
	Appendices	259
A	Charm production yield, efficiency, and systematic tables	261
B	Decision trees	291
	References	297

Word count: 40,000

Abstract

University The University of Manchester

Candidate Alex Pearce

Degree Title Doctor of Philosophy

Thesis Title Measurements of charm production and CP violation with the LHCb detector

Date October 2016

This thesis presents two measurements made using data collected by the LHCb detector, operating at the Large Hadron Collider accelerator at the CERN particle physics laboratory. The first is a measurement of the production rates of promptly produced D^0 , D^+ , D_s^+ , and D^{*+} open charm mesons, using data collected in 2015 at a proton-proton centre-of-mass energy of $\sqrt{s} = 13$ TeV. The second is a search for direct CP violation in two three-body decays of the Λ_c^+ charm baryon, pK^-K^+ and $p\pi^-\pi^+$, using data collected in 2011 at $\sqrt{s} = 7$ TeV and in 2012 at $\sqrt{s} = 8$ TeV. For each measurement, motivation and context are given from the standpoint of improving the theoretical understanding of the Standard Model and searching for signs of physics that cannot be explained by it, and then the various statistical analysis techniques used to extract physical quantities from the data are explained. The systematic limitations of the method are explored and quantified, and then the results are presented.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

1. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
2. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
4. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property University IP Policy, in any relevant Thesis restriction declarations deposited in the University Library, The University Lib-

rary's regulations and in The University's policy on Presentation of Theses.

For Lauren.

Acknowledgements

So many people have done so much for me over the past four-or-so years, and this thesis would never have been finished without them. To anyone who I've had even the most minimal contact with, but I fail to mention: you have truly helped me, and I am very grateful for it. Still, I must be particular.

My supervisors George Lafferty and Sajan Easo helped me immensely when I was starting; Sajan during my Master's project, and George during my PhD. Patrick Spradlin was my third supervisor in everything but name, and did a huge amount to help. Each of them handled my endless questions with patience, and taught me everything I wanted to know.

To everyone I shared an office with at Manchester, Lorenzo, Shanzhen, Giulio, Suzanne, and Kevin, thanks for making being sat in front of a computer all day a bit more bearable. Particular thanks to Igor, for showing me the world outside, and Jon, for telling what were only barely passable as jokes.

Thanks to Chris and Dominik, for sharing the cross-section experience, putting up with the hard times and silly hours, and finishing with an amazing piece of work. I'm really proud of what we managed to do, and I couldn't have done it with anybody else; it's really more your achievement than mine.

I must thank Chris and Marco for allowing me to study at Manchester and for invaluable guidance, and the whole the Manchester group for being nice people to be around. Also the University, the Rutherford Appleton Laboratory, and the Science & Technology Facilities Council for funding this whole business, and to Fred for so deftly pulling the purse strings to allow it to happen.

Vava Gligorov and Patrick Koppenburg have done a lot more for more than they can know, setting me down a path I couldn't have gotten on without them. I'm grateful for their wisdom, guidance, and optimism.

Finally, thanks to everyone within the LHCb collaboration. It's a fantastic environment to work in, and the people I've worked with and spoken to have taught me so much. Coordinators, conveners, reviewers, colleagues, people I only know on mailing lists, and friends, I owe you all.

Preface

The work presented in this thesis was performed in the context of the LHCb experiment, a collaboration of around 800 scientists and engineers. As such, it would not have been possible to perform without the often implicit help of many. The Author worked specifically on the analysis of the data after it was collected by the detector and processed by software common to most other analyses.

The charm production measurements presented in Part II are the result of the work of one Master's student, Christopher Burr, two PhD students, Dominik Müller and the Author, and four post-doctoral researchers, Sajan Easo, Marco Gersabeck, Vava Gligorov, and Patrick Spradlin. The bulk of the analysis work was performed by the three students, with the Author co-ordinating the team as a whole. The Author was primarily responsible for the event and candidate selection, the yield extraction, and some particular efficiency evaluation and systematic studies, although the measurement was truly the work of everyone, with all parts being touched by most analysts at some point. The measurements were published in a peer-reviewed journal in 2016 [1]. Similar measurements using data taken at $\sqrt{s} = 5$ TeV were performed in 2016 by Dominik Müller and the Author [2].

The CP violation search presented in Part III is solely the work of the Author, with supervision from Sajan Easo, George Lafferty, and Patrick Spradlin.

In addition to the work detailed in this thesis, the Author also: implemented a web-based monitoring tool for the VELO sub-detector, the architecture of which was adopted for the central web-based detector monitoring tool used during Run 2; helped found the LHCb Starterkit group, who

aim to teach young researchers software literacy and the basics of the LHCb software suite through three- and four-day workshops; and contributed to a public note on the method of particle identification calibration widely used within the collaboration [3].

Part I

Introduction

1 Overview

The Large Hadron Collider is a circular particle accelerator at CERN in Geneva, Switzerland. It is 27 km in circumference and designed to accelerate two oppositely-circulating beams of protons up to energies of 8 TeV. The beams collide at four points around the ring, at each of which a particle detector records the resulting showers of particles. Two of these detectors, ATLAS and CMS, are designed to study a wide range of phenomena, including searches for particles not seen before the Large Hadron Collider (LHC), such as the Higgs boson, while ALICE is designed to study the quark-gluon plasma, the primary state of matter that is believed to have existed in the early universe. This thesis focuses on the LHCb detector, which is designed to measure the properties of hadrons containing charm and beauty quarks, the two heaviest quarks that can form bound states.

The observed imbalance of matter and antimatter in the universe cannot be explained by established theoretical models, and the physics of heavy flavour, interactions involving charm and beauty quarks, can provide some insight into possible causes. The properties and decays of beauty hadrons have been observed to be different between matter and antimatter [4–8], and similar behaviour is expected in charm hadrons, albeit with a smaller magnitude [9]. The high energy of the LHC beams produces large samples of charm and beauty hadrons [10, 11], allowing for precise measurements to be made of their properties.

This thesis will cover two different analyses. The first is a measurement of the rate of production of mesons containing charm quarks, performed using data taken at a proton-proton centre-of-mass energy of 13 TeV, presented in Part II. The second analysis is a measurement of the relative rate of the

decays of charm and anti-charm baryons, presented in Part III. As both analyses share similar foundations, a general introduction is given in Part I. The Standard Model, the mathematical framework which is able to describe most of what is known about fundamental particles and their interactions, will be presented in Chapter 2. A description of the Large Hadron Collider and its operation will be presented in Chapter 3, before a description of the physics goals and design of the LHCb experiment.

2 The Standard Model

The Standard Model (SM) is a quantum field theory that describes the interactions of three of the four fundamental forces of nature, electromagnetism, the weak force, and the strong force, with a set of fundamental particles. The SM makes many predictions, from the spectral lines of the hydrogen atom to the rate of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay, and so far no experimental results have provided conclusive proof that such predictions are, fundamentally, incorrect. The absence of a description of the fourth fundamental force, gravity, the non-zero masses of neutrinos, and several astronomical observations indicate that, at the very least, the SM is not a complete description of the universe. It does not describe the observed matter-antimatter asymmetry, nor can it account for the apparent presence of dark matter. The purpose of the LHC is to provide experiments with enough data to be able to find minute deviations from the predictions of the SM, which could point to particular extensions of theory. In this chapter, the forces and particles described by the SM are summarised. Particular emphasis is given to the phenomenology of the weak force, which the LHCb detector is optimised to study.

The SM is a powerful predictive model partly due to its use of symmetries, which, through Noether's theorem, give rise to conserved currents. For example, the symmetry of physical laws to rotations in space and translations in time leads to the conservation of total angular momentum and energy, respectively. In a quantum field theory, each of these currents are quantised in units of a quantum charge. Interactions with a field only occur if the particles are charged under that field, such as the familiar electric charge for the electromagnetic field, and all interactions conserve the total charge of the system.

Table 2.1: *Properties of the fundamental particles of the Standard Model [12]. The electric charge is given in units of the electron charge.*

Particle	Electric charge	Spin	Mass (GeV/ c^2)
u, c, t	$2/3$	$1/2$	$2.2 \times 10^{-3}, 1.3, 173$
d, s, b	$-1/3$	$1/2$	$4.7 \times 10^{-3}, 0.1, 4.2$
e^-, μ^-, τ	-1	$1/2$	$0.5 \times 10^{-3}, 0.1, 1.8$
ν_e, ν_μ, ν_τ	0	$1/2$	> 0
γ	0	1	0
g	0	1	0
$W^\pm$	± 1	1	80.4
Z	0	1	91.2
H	0	0	125

Interactions with fields are mediated via the exchange of bosons, particles with integer spin, such as the photon, the force carrier of the electromagnetic field. Fermions are particles with half-integer spins; the quarks, leptons, and neutrinos constitute the fundamental fermions. The known fundamental particles of the SM, along with their measured electric charges and spins, are shown in Table 2.1.

The strong force acts on particles with colour charge, quarks and gluons, via the exchange of gluons. The fact that gluons are charged under the force they mediate significantly complicates the theory of strong interactions. At low energies, the self interaction leads to the phenomenon of confinement, whereby free colour-charged objects cannot be observed, as the attractive force between two colour-charged objects is approximately constant at large separations, and so an infinite amount of energy is required to separate such objects completely. Conversely, at higher energies, these particles become asymptotically free, eventually forming a quark-gluon plasma. The energy scale at which the transition between these two regimes occurs is denoted Λ_{QCD} .

Theoretical studies of the strong force use the framework of quantum chromodynamics (QCD), which can either be used perturbatively, when the energy scale of the process under consideration is much larger than Λ_{QCD} ,

or non-perturbatively, when the opposite is true. It is feasible to make predictions using perturbative QCD, although the computations become more challenging as the order in the perturbative expansions increases. Predictions requiring non-perturbative QCD are generally not possible to perform analytically with today's understanding, although lattice QCD can offer an alternate means of computation for some problems. In most cases, experimental input is necessary to constrain the theory. One such example of a non-perturbative problem is the prediction of pp cross-sections at the LHC, which will be discussed in more detail in Chapter 6.

Through confinement, the six quarks (u , d , c , s , t , and b) cannot be observed as free states. The five lightest quarks bind together to form hadrons, colourless states of two or more quarks bound together with a ‘sea’ of gluons. The top quark decays too quickly to form such states. When quarks and gluons act as outgoing objects in interactions, the observed final state is a shower of colourless particles, where the initial quark or gluon ‘fragments’ into additional coloured objects. These eventually form colourless bound states through the process of hadronisation.

Mesons are hadrons comprising a quark-anti-quark pair, and baryons contain three quarks or three anti-quarks. More exotic states, such as tetra- or penta-quarks, also exist [13–15], but have not been observed outside of particle collider experiments. The behaviour of hadrons can be studied with a particle detector to infer the behaviour of the quarks and gluons. In particular, the LHCb detector is optimised to select and reconstruct the decays of mesons and baryons containing c and b quarks, in order to probe the nature of the weak force. These include the charm mesons D^0 ($c\bar{u}$) and D^+ ($c\bar{d}$), beauty mesons B^0 ($d\bar{b}$) and B^+ ($u\bar{b}$), and the charm and beauty baryons Λ_c^+ (udc) and Λ_b^0 (udb). The decays of these hadrons to sets of comparatively long-lived particles, such as charged pions ($u\bar{d}$) and kaons ($u\bar{s}$), protons (uud), and leptons, allow the properties of the heavy flavour hadrons to be inferred through the interactions of the decay products with the detector.¹

The weak force interacts with particles with a left-handed chirality

¹Throughout this thesis, charge conjugation is implied unless stated otherwise.

and with a non-zero weak isospin T_3 . The chirality can be considered as a Lorentz-invariant version of the helicity (the projection of a particle's spin vector along its momentum vector), and as such encodes the intrinsic 'handedness' of a particle. The symmetry under the parity transformation P , which changes the sign of the spatial coordinates, is maximally violated by the weak force only coupling to particles with a left-handed chirality. The discovery of P violation [16] was soon followed by the discovery of CP violation [17], the breaking of the symmetry of the CP transformation which simultaneously changes the handedness of particles (P) and transforms particles into their anti-particles. (C).

The weak force is particularly interesting to study because it is the only force to have been observed to break the P and CP symmetries. In addition, at high energies the weak field unifies with the electromagnetic field to form the electroweak field. Prior to the process of electroweak symmetry breaking, where the two component fields become physically distinguishable, the electroweak field has four massless degrees of freedom. After the symmetry is broken, two degrees of freedom are observed experimentally as the massive W^+ and W^- bosons, whereas the other two degrees of freedom, W^0 and B^0 , mix to form the observed massive Z boson and the massless photon. The 'spontaneous' breaking of the electroweak unification at low energies is caused by the so-called Higgs mechanism, wherein an additional field, the Higgs field, contributes four additional degrees of freedom to the theory. Three degrees of freedom of the Higgs field mix with three degrees of freedom of the electroweak field to give masses to the $W^\pm$ and Z bosons, whilst the fourth degree of freedom can be interpreted a new boson H , the Higgs particle.

The study of the Higgs particle has only become possible with the LHC, at which the Higgs boson was discovered by the ATLAS and CMS collaborations [18, 19]. Studies continue to measure the properties of the Higgs [20], to see whether it behaves in a way predicted by the SM. Deviations in the behaviour of this particle can indicate the presence of particles not predicted by the SM, observations of which would be considered a major breakthrough in high energy physics. Observations of large CP -violating effects would be

of similar interest, but can relate to different phenomena.

2.1 *CP* violation

As the measurement of *CP* violation in charm and beauty hadron decays is a cornerstone of the LHCb physics programme, and indeed of Part III of this thesis, its mechanism will be describe here in more detail.

The weak interaction eigenstates (d', s', b') are different to the quark mass eigenstates (d, s, b). The weak eigenstates are a superposition of the mass eigenstates, the linear coefficients of which are given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (2.1)$$

where V_{ij} is the coupling of the i to j transition, e.g. $|V_{ud}|^2$ is the relative probability of the transition $d \rightarrow u$. The values of the CKM matrix elements are not predicted by the SM and so must be determined experimentally. There are several methods for determining the coupling constants, depending on the quarks involved, but each generally involves the measurement of one or more decay rates involving a particular quark transition, and require some theoretical input to disentangle the contributions from the strong and weak forces. One example measurement is that of β decay, used as input to compute V_{ud} . Given that the up and down quarks are the most abundant, V_{ud} is particularly well known, with a fractional uncertainty of the order of 10^{-5} . The element V_{ub} is the least well known, with a fractional uncertainty around 2 %, measured using the decays of beauty hadrons [21].

By exploiting the predicated unitarity of the CKM matrix, it can be represented using only three mixing angles and a complex phase δ

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{-i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{-i\delta} & s_{23}c_{13} \\ s_{12}c_{23} - c_{12}c_{23}s_{13}e^{-i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{-i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (2.2)$$

where $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$. In alternative formulations [22], the three angles are parameterised as α , β , and γ and can be visualised as a triangle on the complex plane, as shown in Figure 2.2. One of the key measurements of LHCb is that of the angle γ , which is the least well-measured of the three angles, with the precision on the world average around 9% [23]. This is to be compared with the ultimate precision that could be obtained, given the finite uncertainty on some of the theoretical inputs required to translate the measurements into estimates of the CKM elements, estimated to be up to 4° , depending on the assumptions [24, 25].

The closure of the (α, β, γ) triangle is equivalent to the prediction of the unitarity of the CKM matrix, and hence measuring a value of γ which does not form a triangle would be a sign of physics beyond the SM. The current world averages of the various measurements are presented in the complex plane in Figure 2.2, where some of experimental inputs and their associated uncertainties are shown as bands. The LHCb experiment has performed measurements of most of the inputs shown, with the exception of ε_K which is related to neutral kaon oscillations.

It is the phase in the CKM matrix that permits CP violation. Consider a decay $P \rightarrow f$ and the CP conjugate process $\bar{P} \rightarrow \bar{f}$. The corresponding decay amplitudes $\mathcal{M}_f$ and $\bar{\mathcal{M}}_{\bar{f}}$ can have some phase ϕ , and the CKM mechanism can add an additional phase θ which changes sign under CP , giving

$$\mathcal{M}_f = |a|e^{i(\phi+\theta)}, \quad \bar{\mathcal{M}}_{\bar{f}} = |a|e^{i(\phi-\theta)}, \quad (2.3)$$

where $|a|$ is the magnitude of matrix element. The observable rates of these processes, $|\mathcal{M}_f|^2$ and $|\bar{\mathcal{M}}_{\bar{f}}|^2$, are equal, and so the decay is symmetric under CP . However, for a process that can proceed via two channels, for example, such as in Figure 2.1, the total matrix elements are $\mathcal{M}_f = \mathcal{M}_{f,1} + \mathcal{M}_{f,2}$ and $\bar{\mathcal{M}}_{\bar{f}} = \bar{\mathcal{M}}_{\bar{f},1} + \bar{\mathcal{M}}_{\bar{f},2}$ with

$$\begin{aligned} \mathcal{M}_{f,1} &= |a_1|e^{i(\phi_1+\theta_1)}, & \mathcal{M}_{f,2} &= |a_2|e^{i(\phi_2+\theta_2)}, \\ \bar{\mathcal{M}}_{\bar{f},1} &= |a_1|e^{i(\phi_1-\theta_1)}, & \bar{\mathcal{M}}_{\bar{f},2} &= |a_2|e^{i(\phi_2-\theta_2)}. \end{aligned}$$

It then follows that

$$|\mathcal{M}_f|^2 - |\bar{\mathcal{M}}_{\bar{f}}|^2 = -4|\mathcal{M}_1||\mathcal{M}_2|\sin(\phi_1 - \phi_2)\sin(\theta_1 - \theta_2).$$

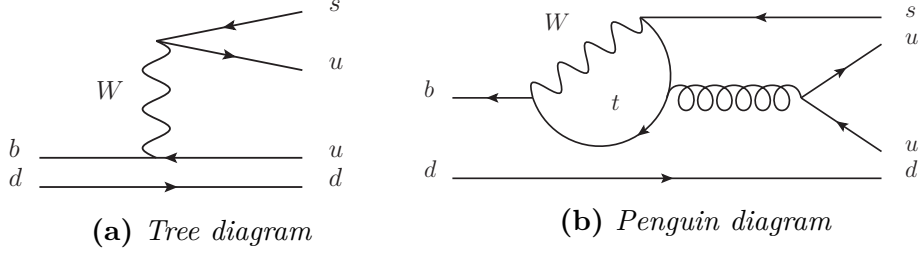
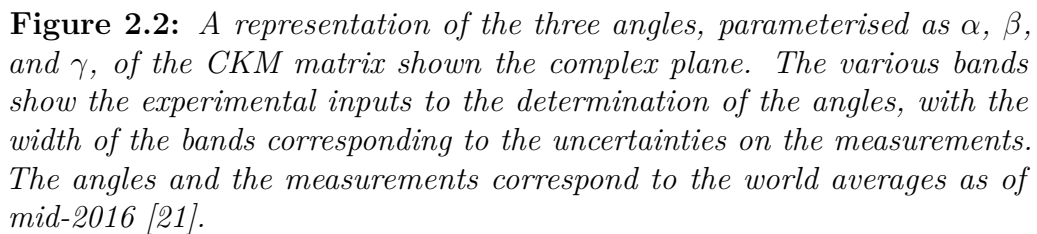


Figure 2.1: Two possible Feynman diagrams for the decay $B^0 \rightarrow K^+ \pi^-$.

This is non-zero, provided that $\theta_1 \neq \theta_2$ and $\phi_1 \neq \phi_2$, and hence introduces a possible CP asymmetry. It is then the interference between channels that permits CP symmetry breaking in weak interactions.

The breaking of the CP symmetry was first observed in kaon decays in 1964 [17], and in 2001 CP violation was first observed in beauty decays [4, 5]. Observing CP violation in charm decays would provide a more complete picture of what is possible in the SM, but it will not solve the problem of the observed baryon asymmetry in the universe, which the SM cannot account for. In the LHC era, it is hoped that signs of physics beyond the Standard Model will point to specific, new theories, that themselves may have explanations for this great mystery.



3 The Large Hadron Collider

As high-energy particle physics has progressed over the last century, the meaning of ‘high energy’ has continuously evolved. Today, high energy refers to the Large Hadron Collider [26], a circular collider which began operating at a proton-proton centre of mass energy of $\sqrt{s} = 7$ TeV in 2010, and in 2015 reached $\sqrt{s} = 13$ TeV. This section will describe the machine, the experiments which it services, and the changes in operation throughout its lifetime so far.

In order to produce collisions, bunches of 1.2×10^{11} protons are obtained through the ionisation of hydrogen gas, and are then accelerated through a chain of progressively larger, more powerful accelerators before being injected into the LHC. During nominal operation, a single LHC ‘fill’ consists of injecting 2808 bunches into the machine, each of which is about 10 cm long, separated from the surrounding bunches by about 7 m, completing a full revolution of the ring at a rate of 11.246 kHz.

The beams are steered around the ring by superconducting dipole magnets, cooled to 1.9 K, and are focused by quadrupole and other, higher-order, magnets. The beams are made to collide at four interaction points (IPs), before which they are focused further, from the usual circulating bunch diameter of 0.2 mm to a width on the order of microns, such that they can collide.

For a collision experiment, the rate N at which a particular process occurs per second is given by

$$N = \mathcal{L}\sigma, \tag{3.1}$$

where $\mathcal{L}$ is the instantaneous luminosity of the accelerator and σ is the

cross-section of the process. At $\sqrt{s} = 13 \text{ TeV}$, the visible proton-proton interaction cross-section is around $\sigma_{\text{vis.}} = 70 \text{ mb}$ [27, 28]. Given that this number is fixed by the beam energy and colliding particle type, the LHC is designed to maximise the luminosity $\mathcal{L}$, which can be expressed as

$$\mathcal{L} = \frac{N_p^2 N_b \nu_{\text{rev.}} \gamma}{4\pi\epsilon\beta^*} F, \quad (3.2)$$

where N_p is the number of particles per bunch; N_b is the number of bunches per beam; $\nu_{\text{rev.}}$ is the bunch revolution frequency; γ is the Lorentz factor; ϵ is the beam emittance during circulation, a measure of the beam size transverse to the beam direction; β^* is a measure of the transverse beam size at a particular IP; and F accounts for non-zero crossing angles between the two beams at an IP. Depending on the experiment, the β^* parameter can be tuned to provide a specific average number ν of collisions per bunch crossing. The design luminosity at the ATLAS and CMS IPs is $1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, which results in a rate of visible proton-proton interactions of 700 MHz.

Running periods

The LHC first began running in 2010 at a proton-proton centre-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$, and this continued until the end of 2011. In 2012, the centre-of-mass energy was increased to $\sqrt{s} = 8 \text{ TeV}$, before the machine was shut down at the end of 2012. This marked the end of Run 1 and the beginning of long shutdown 1 (LS1), during which the accelerator was upgraded in order to prepare the machine for running at $\sqrt{s} = 13 \text{ TeV}$. This new phase of operation, Run 2, began in mid-2015, and is expected to continue until the end of 2018.

4 The LHCb experiment

The LHCb detector [29, 30] is situated at point 8 of the LHC. The general goal of the experiment is to explore the area of heavy flavour physics, that is the interactions of charm and beauty quarks and the hadrons that contain them. In particular, LHCb aims to make, and already has made some of, the most precise measurements of heavy flavour properties to date, as well as to discover decays and particles that were not observed by previous experiments. Heavy flavour has a rich phenomenology: neutral particles can oscillate between their matter and antimatter states, such as $B_{(s)}^0 \leftrightarrow \bar{B}_{(s)}^0$ and $D^0 \leftrightarrow \bar{D}^0$ [31, 32]; bound states of four or five quarks can form [14, 15]; they exhibit distinct matter-antimatter asymmetries [4–8]; and their rare decays can be precisely predicted by the SM [33, 34], as such being very sensitive to contributions from new theories. With this, the LHCb detector must be flexible enough to accommodate a wide physics programme, yet powerful enough to discriminate against the large backgrounds characteristic of a pp collision environment. This chapter will link the properties of heavy flavour decays with the requirements imposed on the detector.

At the LHC, heavy flavour quarks are primarily produced through gluon-gluon fusion, illustrated in Figure 4.1a. With this mechanism, $b\bar{b}$ and $c\bar{c}$ pairs are predominantly produced in the forward region, at low values of θ , as shown in Figure 4.1b, where the polar angle θ is defined as the angle made with the beamline. To exploit this, the LHCb detector is instrumented in the pseudorapidity region $2 < \eta < 5$, where

$$\eta = -\ln \tan \frac{\theta}{2}. \quad (4.1)$$

Particles produced in this region are highly boosted in the laboratory frame,

and so heavy flavour hadrons can fly several millimetres before decaying, given that their lifetimes are of the order of 0.1–1 ps [12]. With a sufficiently sensitive detector, their decay vertices can be spatially distinguished from the primary proton-proton interaction vertices, providing a powerful discriminant between signal and background. Good secondary vertex resolution also allows for good decay time resolution, which is necessary for measuring the fast oscillations of $B_{(s)}^0$ and D^0 mesons, and for measuring any possible decay time asymmetries. In addition, the large displacement of the decay vertices with respect to the primary vertex (PV) causes the particles produced at the secondary vertex to have a high impact parameter (IP), defined in Figure 4.2 as the smallest distance from the particle trajectory to the PV. A sufficient experimental resolution on the IP allows for additional discrimination between random tracks in the event and those from heavy flavour decays. The precise reconstruction of primary and secondary vertices requires a precise tracking system, particularly so around the pp interaction region, and hence the detector design is highly optimised for tracking performance. The tracking system will be described in Section 4.1.

The properties of heavy flavour hadrons, such as their masses and lifetimes, can be inferred by reconstructing the four-momenta of their decay products, and so most analyses at LHCb aim to fully reconstruct the decays of beauty and charm hadrons, such as $B^0 \rightarrow K^+ K^-$, $B_s^0 \rightarrow D_s^\pm K^\mp$ with $D_s^+ \rightarrow K^- K^+ \pi^+$, $\Lambda_b^0 \rightarrow J/\psi p K^-$ with $J/\psi \rightarrow \mu^- \mu^+$, and $D^0 \rightarrow K^\pm \pi^\mp$. It is important to note that certain final states offer higher sensitivities to physics observables than others, and so it is crucial for LHCb to be able to distinguish between different particle species. For example, the decay $D^0 \rightarrow K^+ \pi^-$ is of the order of one thousand times less likely than the decay $D^0 \rightarrow K^- \pi^+$, and so a poor particle identification (PID) performance would result in a search for the $K^+ \pi^-$ final state being dominated by $K^- \pi^+$ ‘background’. In general, misidentification cannot be excluded completely, and so in addition good mass resolution is required to be able to distinguish between different final states in the parent mass spectrum.

An example of the power of, and necessity for, good particle identification and momentum resolution is shown in Figure 4.3. Here, the signal decay is

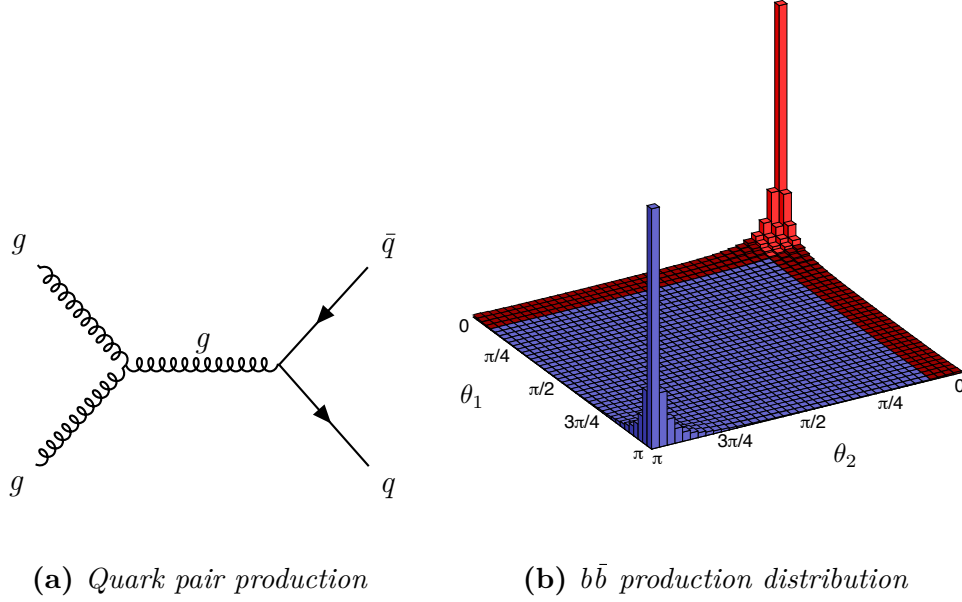


Figure 4.1: Feynman diagram of quark pair production via gluon-gluon fusion (a), and a simulation of the angular distribution of $b\bar{b}$ production at the LHC at $\sqrt{s} = 13$ TeV (b).

$B^0 \rightarrow \pi^+\pi^-$, and the distributions are the two-body invariant mass before PID requirements are made on the tracks (4.3a) and after (4.3b). As can be seen, before the requirements the distribution is dominated by other two-body B^0 , B_s^0 , and Λ_b^0 decays, whereas after the selection the backgrounds are almost entirely removed. The presence of such backgrounds significantly increases the complexity of an analysis and can reduce the precision of the measurement.

In the following, the construction of the detector will be described, as motivated by the physics goals of the collaboration.

4.1 The LHCb Detector

Within the LHCb experiment, the Cartesian co-ordinate system is defined such that the z -axis is aligned along the beam direction, increasing in the

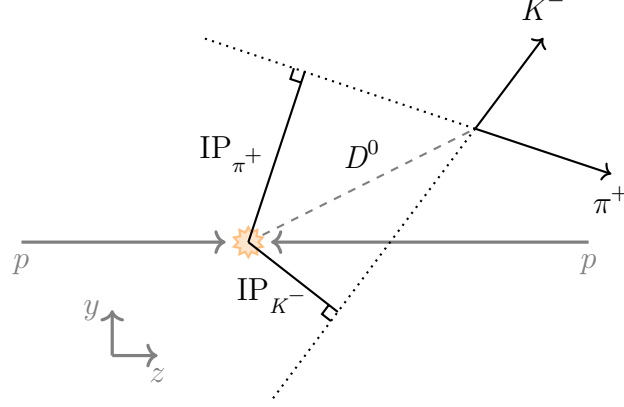


Figure 4.2: Illustration of vertex reconstruction, showing a D^0 meson decaying in flight to a kaon and a pion. The kaon and pion are reconstructed as tracks, and then the D^0 decay vertex is inferred from the point of closest approach of the two tracks. The minimum transverse distance the tracks make when extrapolated back towards the primary proton-proton vertex, the impact parameter (IP), is shown.

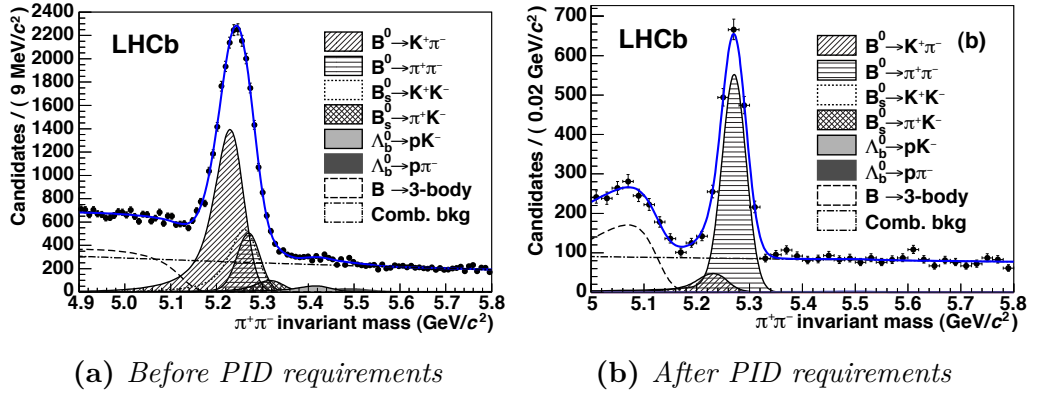


Figure 4.3: Two-body invariant mass spectrum, under the $\pi^- \pi^+$ hypothesis, before (a) and after (b) PID requirements [35]. The contribution from the signal decay is hatched with horizontal lines.

clockwise direction along the LHC ring; the y -axis points vertically upwards; and the x -axis, defined as $x = y \times z$, is then pointing away from the centre of the accelerator. The polar co-ordinate system is defined with the polar angle θ measured with respect to the z -axis, and the azimuthal angle ϕ measured in the xy plane.

A schematic of the detector is given in Figure 4.4. Its forward geometry is evident, with the detector instrumenting the pseudorapidity region $2 < \eta < 5$. A dipole magnet with an integrated field strength of 4 T m bends charged particle trajectories in the xz plane, with its polarity being switched periodically during data-taking. Charged particle trajectories are recorded by the tracking system. This is composed of a vertex detector centred around the proton-proton interaction region, three planes of tracking stations before the magnet, and three planes after. Immediately downstream of the vertex detector (increasing in z) is the first of two ring-imaging Cherenkov (RICH) detectors, each used for particle identification, the second of which is after the tracking stations downstream of the magnet. After this, a calorimetry system is in place to identify neutral particles such as the π^0 and the photon, as well as to measure their energy and that of the charged particles. Finally, a muon identification system is installed after the calorimeters.

The data are selected in real time by a hardware trigger, after which a two-stage software trigger performs an event reconstruction. There are not enough computing resources available to the experiment to record and analyse every proton-proton bunching crossing, and so the trigger is essential to achieve the physics goals.

Each sub-detector will now be described in turn, followed by a description of the trigger system. These descriptions will relate to the detector and its performance during Run 1, and will be followed by Section 4.3 on the offline processing model and the upgrades performed for Run 2.

Tracking

Charged particle trajectories are reconstructed as tracks using hits deposited in the tracking system. This comprises the vertex locator (VELO)

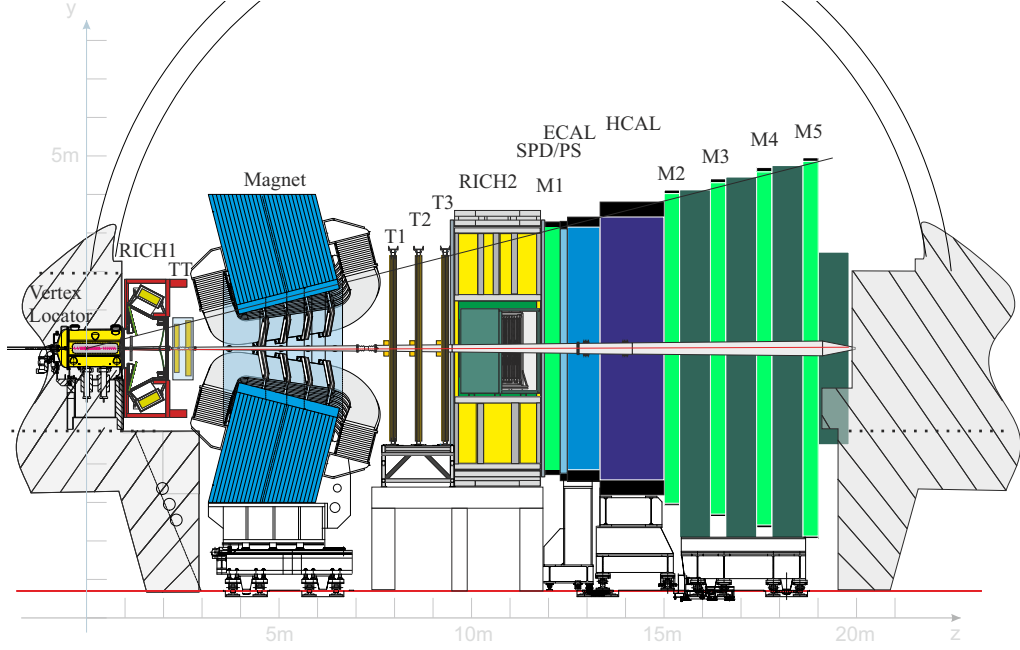


Figure 4.4: A schematic of the LHCb detector. In this Figure, the z -axis, labelled, increases from left to right, the y -axis, also labelled, increases from bottom to top, and the x -axis increases into the page.

surrounding the pp interaction region, the Tracker Turicensis (TT) before the magnet, and the T-stations after the magnet. In Run 1, the tracking system achieved a momentum resolution $\delta p/p$ from 0.5% at $p = 5 \text{ GeV}/c$ to 0.8% at $100 \text{ GeV}/c$, and a track impact parameter resolution varying from around $70 \text{ }\mu\text{m}$ for tracks with a low p_T to $20 \text{ }\mu\text{m}$ at high p_T .

The VELO is an extremely precise silicon strip detector, whose active elements come as close as 8 mm to the LHC beams. Two sets of 24 silicon modules are arranged either side of the beam, as shown in Figure 4.5, each of which measures particle hits in r and ϕ coordinates. Charged particles traversing the active silicon can ionise the material, creating electron-hole pairs that drift towards the electrodes, registering a hit as an electrical signal. The magnetic field strength inside the VELO is approximated to be zero, and so track segments are reconstructed as straight lines using the hits in the sensors. So-called long tracks, the track type used for the majority of

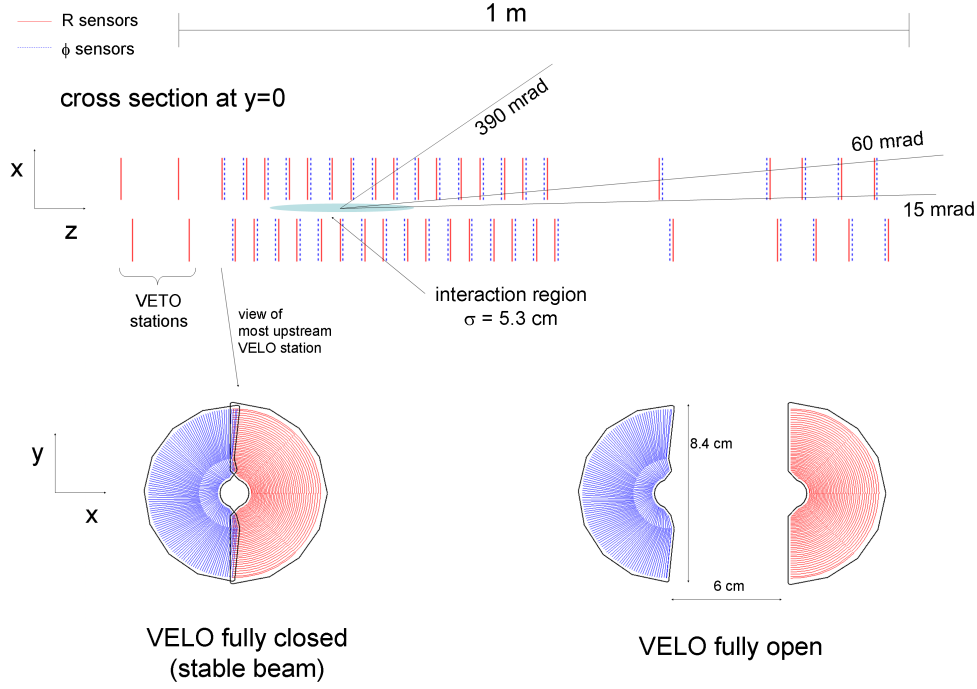


Figure 4.5: A schematic of the VELO sub-detector [29], showing the relative positions of the sensors when the detector is open and closed.

physics analyses, are created by combining VELO segments with hits in the T-stations after the magnet.

The three T-stations, T1–3, each comprise a silicon strip detector close to the beam pipe (the IT) and a drift-tube detector in the outer regions (the OT). The IT has a finer spatial resolution than the OT, and covers a small area around the beam where the particle multiplicities are particularly high. The OT covers an area of approximately 30 m^2 , about 8 times greater than that covered by the IT. Ionising particles liberate electrons from the gas molecules within the tubes, which are detected by anode wires and registered as hits. The distance of the line of ionisation from the wire is determined by measuring the electron drift time, improving the track momentum resolution. The momentum resolution on long tracks is further improved by adding hits detected in the TT, upstream of the T-stations. The TT is a silicon strip

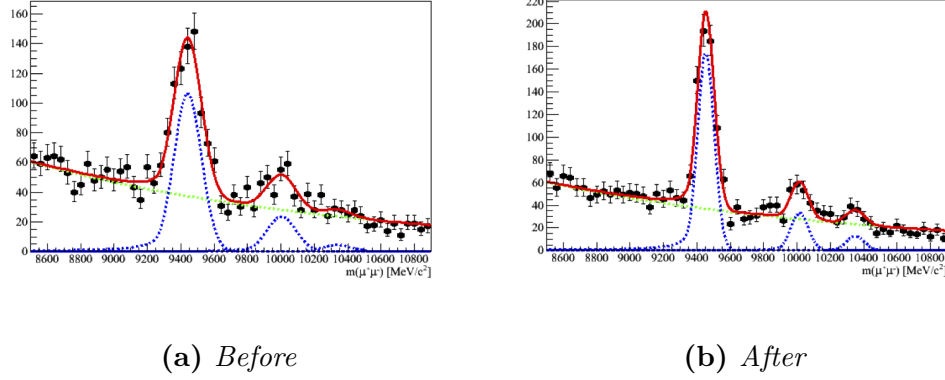


Figure 4.6: *Invariant mass distribution of di-muon candidates in the region of the first three Υ resonances [37]. The left plot (a) shows the data reconstructed with a preliminary alignment, whilst the right plot (b) shows the result of a reconstruction performed with a revised alignment. The di-muon mass resolution improves from 92 MeV/c² to 49 MeV/c².*

detector of the same technology as the IT, with the three stations covering a total area of around 8 m².

The tracking efficiency, defined as the fraction of genuine tracks that are reconstructed given that enough hits were deposited to do so, is measured using a tag-and-probe technique with $J/\psi \rightarrow \mu^+\mu^-$ decays, and described in more detail in Section 10.2. The average tracking efficiency is measured to be above 95 % [36]. The efficient creation of tracks requires the tracker system to be well aligned, but the positions of the various stations can change with time. To compensate for this effect, alignment constants are computed periodically and used in the reconstruction software. The effect of an improved alignment is illustrated in Figure 4.6, where the $\mu^+\mu^-$ mass resolution improves from 92 MeV/c² with an initial alignment to 49 MeV/c² with an improved one [37].

Particle identification

Through the momentum-velocity relation, the mass of a particle can be determined by combining the momentum measurement from the tracking system with a velocity measurement. Velocity measurements are made using

the RICH detectors.

Tracks created by pions, kaons, and protons are identified using the response of two RICH detectors: RICH1 located upstream of the magnet and RICH2 located downstream. The RICHs also provide electron and muon discrimination, albeit at a weaker level. RICH2 has a larger surface area than RICH1 but a smaller angular acceptance, being designed to efficiently identify tracks with momenta in the range 15 to 100 GeV/c, which largely occupy the low- θ region. RICH1 is effective in the momentum range 3 to 60 GeV/c. During Run 1, RICH1 contained both C₄F₁₀ gas and aerogel, of which the latter was removed before the start of Run 2. RICH2 contains CF₄ gas mixed with a small amount of CO₂ to quench scintillation light.

As charged particles travel through one of the radiators, each of refractive index n , the particles emit Cherenkov radiation if they are travelling faster than the phase velocity of light of the radiator. This light is emitted at an angle θ_c to the particle trajectory, and is related to the velocity v of the particle as

$$\cos \theta_c = \frac{1}{n\beta}, \quad (4.2)$$

where $\beta = v/c$. The refractive indices of the radiators are known and are controlled by monitoring the gas pressure and temperature over time. By measuring θ_c the RICH detectors provide velocity measurements of charged tracks.

The Cherenkov light cones are reflected outside of the LHCb detector acceptance by spherical and plane mirrors onto planar arrays of photon detectors. The angle θ_c is related to the radius of the rings, and, for a given momentum, particles of different masses will produced smaller or larger rings. Figure 4.7 illustrates the Cherenkov angles for different particle species as a function of momentum. Separation between mass hypotheses is given by the difference between the angles, denoted on the y -axis. Discrimination can also be provided by the lack of Cherenkov light if a track has a momentum between two momentum thresholds. The Cherenkov angle resolution in RICH1 is around 1.6 mrad, and around 0.66 mrad in RICH2.

In order to assign a particle mass hypothesis to a track, a maximum

likelihood method is used [38]. Initially, all reconstructed tracks are assumed to be pions, and the *predicted* Cherenkov rings assuming this hypothesis set are projected onto the photo-detector planes within both RICH1 and RICH2. The value of the log-likelihood is computed by comparing the predicted number of photons within each photo-detector to that observed. For each track, the mass hypothesis is then changed to each of (e, μ, π, K, p) , the rings assuming the new global hypothesis set are projected, and the log-likelihood is re-computed. The hypothesis giving the largest increase in the log-likelihood is recorded, and the track hypothesis is returned to its original value. After all individual track hypotheses have been trialled, the single hypothesis change that gave the largest increase in the log-likelihood is applied to the respective track. This procedure is repeated until the log-likelihood no longer increases with any change in the hypothesis set. The outputs of this method are the delta log-likelihood (DLL) variables, with one value per track for each of the (e, μ, K, p) hypotheses. Each DLL variable is defined as the difference in the log-likelihood value when the track mass hypothesis is changed from the pion hypothesis to the one of the variable, for example $\text{DLL}_{K-\pi}$. The pion DLL value is then zero by definition.

Calorimetry

Given the momentum measurements from the tracking system and the velocity measurements from the RICH detectors, the particle energy is already determined. For charged hadrons then, precise calorimetry is not necessary. However, the momentum and mass measurements are complex and require a significant amount of time to compute, and so cannot be used in the first stage of the trigger where the proton-proton collision rate must be reduced from 40 MHz down to around 1 MHz. In addition, the RICH detectors are much less sensitive to electron PID than (π, K, p) , and the tracking system cannot detect photons or neutral pions.

To provide a fast positive trigger decision for events containing hadronic decays of heavy flavour, a hadronic calorimeter, the HCAL, is employed. This is a sampling calorimeter, composed of alternating layers of iron and

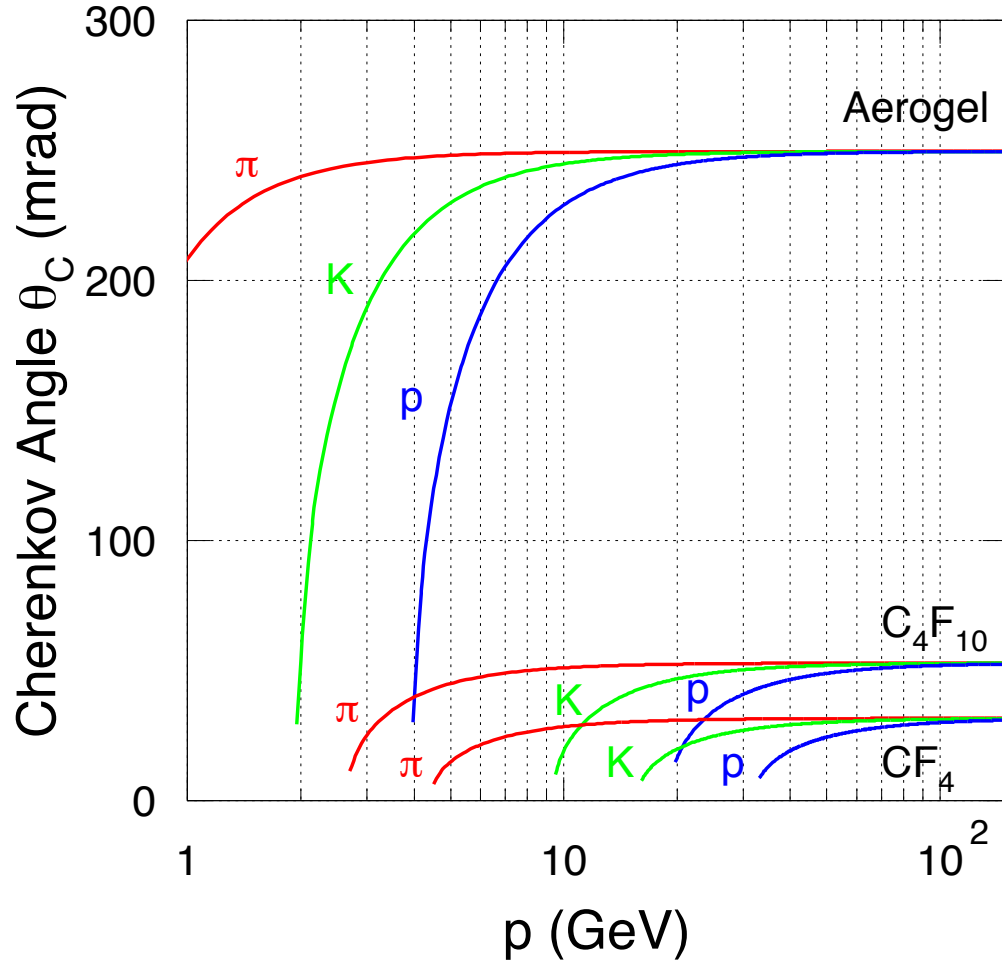


Figure 4.7: Cherenkov angles for different particle species as a function of particle momentum, in the different radiators used in the RICH detectors. The Cherenkov angle resolution in RICH1 (aerogel and C_4F_{10} radiators) is around 1.6 mrad, and around 0.66 mrad in RICH2 (CF_4).

active scintillating material and positioned downstream of RICH2. Hadrons traversing the HCAL deposit energy in these layers, with the scintillation light collected by wavelength-shifting optical fibres and recorded by photomultiplier tubes. This general construction and detection mechanism is common to all LHCb calorimeters. The HCAL is segmented transversely into square cells, 130 mm in width near the beam and 260 mm in width in the outer regions, decreasing in granularity to account for the lower particle multiplicities. In the first-level trigger, the total transverse energy E_T in all clusters of 2×2 cells is computed, and a positive trigger decision is made if the maximum E_T is above some threshold, typically between 3.5 and 3.7 GeV. The HCAL provides the highest rate of positive trigger decisions out of those at the first trigger level, giving a signal efficiency of around 40 % for two-body B decays, and 20 % for four-body D decays, with respect to the offline selection. The resolution on the energy measured by the HCAL has been measured to be $\delta E/E = (69 \pm 5) \%/ \sqrt{E}$ plus a constant term of $(9 \pm 2) \%$ [39].

An electromagnetic calorimeter, the ECAL, is located in front of the HCAL, and measures the energies of electrons and photons. It is segmented in the transverse plane into three regions, increasing in granularity towards the beam. The ECAL is partitioned more finely than the HCAL due to the smaller transverse scale at which electromagnetic showers occur in comparison to hadronic showers. The energy measurements it makes are used in the first trigger level to accept high- E_T electrons and photons, and in the reconstruction of π^0 and η mesons offline. The ability to discriminate between electrons and photons in the trigger decision is provided by the scintillating pad/preshower detector (SPD/PS) which consists of two layers of near-identical scintillator either side of a 15 mm-thick layer of lead (Pb) converter, transversely segmented in the same manner as the ECAL. A positive electron trigger decision requires an E_T measurement of between 2.5 and 3.0 GeV in a cluster of 2×2 cells in the ECAL, along with hits in both the SPD and the PS pads in front of the corresponding ECAL cluster. A positive photon trigger decision is near-identical to this, except that no hits must be present in the corresponding SPD cells. The resolution on

ECAL energy measurements is $\delta E/E = (9.0 \pm 0.5) \%/ \sqrt{E}$, plus a constant term of 0.8 % [39].

Muon reconstruction and identification

A series of five muon stations, M1–5, are used to identify muons. M1 is positioned before the SPD/PS, and M2–5 are after the HCAL, interleaved with 80 cm-thick layers of iron. For almost all regions across the stations, multi-wire proportional chambers are used to collect hits, with the exception being the very inner $20 \times 24 \text{ cm}^2$ region of M1 which uses triple-layered gas electron multiplier (GEM) detectors. Each station is partitioned into four regions of increasing transverse granularity closer to the beam pipe.

A muon momentum of at least 6 GeV/c is required for a full traversal of all five stations. Hits recorded in the muon stations are used to provide p_T measurements used in the first trigger stage, where a track segment must be formed using hits in all stations. The single muon decision requires a minimum segment p_T from 1.48 to 1.76 GeV/c, and the di-muon requires the product of the p_T of the two muons to have a minimum p_T^2 from 1.69 to 2.56 (GeV/c)².

4.2 Online event reconstruction and selection

There are three stages to the LHCb trigger system, run sequentially. Within each stage there is a set of parallel trigger *lines*, and each stage runs only if at least one line in the previous stage gave a positive decision.

The first stage is a hardware trigger, called the level-zero or L0. It comprises a set of decision units as custom circuitry that evaluate decisions at the full 40 MHz of pp collisions provided by the accelerator. The L0 output rate is limited to the 1 MHz rate at which the full detector is able to be read out. The muon stations and the calorimeters can be read at 40 MHz, and it is the information from these sub-detectors that is used for making L0 decisions. The single muon decision requires a high p_T muon, and the

di-muon decision requires two. The hadron, electron, and photon triggers require high E_T signatures in the HCAL or ECAL, as appropriate. On a positive L0 decision, when one or more lines have ‘fired’, the entire detector is read out and the data are sent to the high-level trigger (HLT) computing farm for processing in software. Since 2012, around 20 % of events sent to the farm are deferred to disk, to be processed during the inter-fill periods when the detector is idle [40]. This technique increases the average CPU time available for processing each event, allowing for looser momentum requirements in the reconstruction and hence more efficient triggers.

The first stage of the software trigger, HLT1, performs a simplified track reconstruction in order to confirm the L0 decisions and improve their discriminatory power. Segments in the VELO detector are first reconstructed, and pp vertices are formed. Segments that do not have a significant IP with respect to all PVs or that cannot be matched to hits in the muon stations are discarded. The remaining segments are matched to hits in the T-stations to form long tracks. To save processing time arising due to combinatorics, a search window is defined by a minimum momentum that varied from 3 to 6 GeV/ c and a minimum p_T requirement that varied from 0.5 to 1.25 GeV/ c . The transverse momentum window is not required on tracks that can be matched to hits in the muon stations. With the faster track reconstruction, the mass resolution on $J/\psi \rightarrow \mu^+ \mu^-$ candidates is around 3 % lower than that achieved in the offline reconstruction.

For final states not containing leptons, the most efficient trigger path is through the one-track line, which requires the presence of a single, good quality, high p_T track with a large IP. Typical thresholds were 1.6–1.7 GeV/ c in p_T and 0.1 mm in IP. This line dominates the HLT1 output rate, contributing over 70 % of all triggers. Similar lines exist for muon and di-muon candidates, where the corresponding IP thresholds are not required if the candidate has a very high p_T (or a high $\mu^+ \mu^-$ invariant mass in the case of the di-muon line). The muon triggers have efficiencies upwards of 70 % for B decays containing muons.

Events passing HLT1 are sent to the second software stage, HLT2, at a rate of 80 kHz, where a full event reconstruction is performed, including

the tracks, vertices, and PID information. As more processing time is now available per event, all VELO segments are used in the long-track reconstruction, with looser search window requirements of $p > 3 \text{ GeV}/c$ and $p_T > 0.3 \text{ GeV}/c$, in comparison with HLT1. A relative loss in efficiency of 1 to 2 % per track is seen compared to the offline reconstruction. In HLT2, tracks are combined into vertex candidates, and HLT2 lines are grouped as either *exclusive*, where decays are fully reconstructed, or *inclusive*, where generic signatures are considered.

For beauty decays, the inclusive ‘topological’ lines select two, three, and four-body vertices of charged tracks that are characteristic of those produced by b hadrons. Input tracks are required to have a high IP, and then the resulting vertex is selected based on its distance from the PV, the IP of the vertex momentum vector, the distance of closest approach (DOCA) and the sum of the p_T of the input tracks, and the vertex mass and corrected mass, the latter of which accounts for decay products not explicitly included in the vertex such as long-lived charm hadrons. The DOCA requirements are loose enough to accommodate b hadron decays that include long-lived charm hadrons, where a tertiary vertex would be defined in the offline reconstruction. The vertex quantities are used as input to a boosted decision tree (BDT), and the single output value is used to make the trigger decision [41, 42]. A similar set of topological lines exists that are optimised for b hadron decays that include a high p_T muon, and an inclusive D^{*+} line exists for selecting $D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow hhX$ decays, where hh are two charged hadrons. The topological B triggers are around 75 % efficient on average at selecting fully charged B decays, and the inclusive D^{*+} trigger is between 25 and 90 % efficient. The efficiencies generally increase with increasing heavy flavour p_T , and for the charm triggers are strongly dependent on the multiplicity of the final state, with higher multiplicities having a lower efficiency. The remaining inclusive lines are the single muon and di-muon lines, which are similar to those in HLT1.

Exclusive HLT2 lines fully reconstruct signal decays, such as $B \rightarrow \pi^+ \pi^-$ and $\Lambda_c^+ \rightarrow p K^- \pi^+$. Some of these lines are included to increase the efficiency of the inclusive lines for specific final states, whilst others are the only trigger

path which would give a high efficiency.

Events passing HLT2 are saved to disk. The offline processing then begins with the full event reconstruction, which is more precise than that in the trigger due to the longer processing times permitted. This is exploited in, for example, the reconstruction of long tracks, where two methods are employed offline rather than the one online. The differences between the online and offline reconstruction were removed during LS1, which will be discussed in the following section.

4.3 Offline data flow and the upgrades for Run 2

During Run 1, all data passing HLT2 were saved to disk and were reconstructed offline by a separate software application. The output of this reconstruction comprises tracks and the information associated to them such as the DLL PID responses.

To facilitate easier processing offline for analysts, a central selection is run called the ‘stripping’, which defines stripping lines that reconstruct inclusive and exclusive decays in a similar manner to those defined in HLT2. The lines are grouped into ‘streams’ which group together stripping lines with similar selections, such as those of semileptonic B decays or di-muon decays. In general, any one stripping line contains events which could have been saved due to the decision of any trigger line, and so analysts were able to choose which triggers to include in their analysis dataset offline. This can be a complex procedure, as now there are at least two sets of selections to consider, those in the trigger and those in the stripping, and one must also consider the difference in resolution between the two reconstructions: using the same selection online and offline would result in visible resolution effects around the selection boundaries, which can be hard to model.

To overcome these complications, three parallel efforts were made during LS1. The first of these was a major review of the online and offline reconstruction software, resulting in a large decrease in the execution time

required per event in the offline reconstruction, such that the online reconstruction was made to be identical to that offline. The second effort was the development and deployment of a real-time alignment and calibration of the full detector for each LHC fill. The full output of HLT1 is buffered to a 10 PB disk farm, the detector alignment and calibration constants are computed, and are then applied during the reconstruction performed in HLT2 [37]. The HLT2 processing is then fully asynchronous to the LHC. The third effort was the introduction of the Turbo stream processing model [43]. Given that the additional offline processing was no longer necessary to improve the resolution of the various physics quantities, and that the best possible detector alignment and calibration are already applied, physics analyses could be performed directly on the output of the trigger.

There are several benefits due to the LS1 efforts. The improvement of the online reconstruction allows for parity between the online and offline selections, increasing the trigger efficiency and the possible number of signal candidates available for analysis, and decreasing the total processing time required. This in turn allowed for the inclusion of a new two-track trigger in HLT1, which is able to build and select two-prong vertices, and a set of so-called ‘lifetime unbiased’ HLT1 lines, which create two-body vertices from tracks reconstructed without an IP requirement, allowing for lifetime measurements without the need to model a complex acceptance efficiency as a function of particle lifetime. In addition, these efforts negate the need for an additional, offline reconstruction, saving computing resources, allowing analyses and data quality reports to progress more quickly as the data are made available almost immediately, within hours rather than the days or weeks with the Run 1 processing model. Analyses have already been performed that exploited this model [1, 2, 44].

The Turbo stream contains the signal candidates reconstructed in HLT2, and the data are not centrally processed further offline. The time taken from data-taking to analysable data being available is then considerably reduced with respect to the usual data flow, however the Turbo processing model does have some limitations at present. The most restrictive of these is that the candidates available to analysts are only those explicitly used in the

trigger decision. For example, if a trigger line reconstructs $D^0 \rightarrow K^- \pi^+$ and the candidate passes the selection, analysts can only use the information pertaining to the kaon and pion tracks, the D^0 vertex, and the associated primary vertex. For many analyses, this information is all that is needed; however sometimes additional information about the event is required. Effort is ongoing to resolve these problems, and it is expected that a vast majority of analyses will be using the Turbo stream in Run 3.

4.4 Simulation

The simulation is used to model effects that cannot be modelled using the real data, such as acceptance efficiencies, and to perform detector performance studies during design phases.

Proton-proton collisions are generated by the PYTHIA 8 program [45] using a tuning specific to LHCb [46]. PYTHIA simulates the parton interactions, the fragmentation of the outgoing coloured objects (quarks and gluons), and resulting creation of colourless particles through hadronisation. The decays of the hadrons are simulated by the EVTGEN program, which models phenomena such as branching fractions, neutral meson mixing, CP violation, decay amplitude models, and decay times. Simulated samples for a specific decay chain are generated by running PYTHIA in minimum bias mode, where proton-proton collisions are generated until the particle at the top of the chain is created somewhere in the event. This head particle is then forced to decay by EVTGEN [47] through the decay chain under study. The behaviour of the other particles generated in the event is also controlled by EVTGEN, but is not forced to any particular state.

After EVTGEN, the particles are propagated through a simulation of the entire detector using the GEANT 4 program, where interactions with the active detector elements are recorded as ‘Monte Carlo (MC) hits’. These MC hits are converted into a format mimicking the electrical response of the real detector through an LHCb-specific emulation, and the response is processed through the trigger and reconstruction software in a near-identical manner to real data. The L0 trigger decisions are emulated in software, and the

trigger configuration used for a given data-taking year, such as 2011, is chosen to be the one that was used for the majority of the data-taking.

The simulation is processed in such a way that reconstructed objects can be associated to the MC objects that created them. For charged, stable particles reconstructed via tracks, a particle is associated to an MC particle if at least 70 % of the hits comprising the associated track were deposited by the MC particle in question. The process of assigning MC objects to reconstructed objects is referred to as truth matching. Tracks are classified as ‘ghost’ tracks if there are no MC particles that can be associated to them. Vertices are assigned categories based on the associations of their input tracks. A ‘signal vertex’ is defined as one in which all inputs are associated to MC particles, all inputs have been assigned a particle identity the same as that of their associated MC particle, all inputs have been associated with MC particles which come from the same true MC parent, and the identity of the MC parent matches that assigned to the vertex. Any deviation from these requirements results in the vertex being assigned a particular *background category*, dependent on the nature of the deviations [48]. For example, if at least one track is a ghost, the vertex is classified as a ghost, and if at least one track is associated to an MC particle with a different PID, the vertex is classified as a misidentification. In the case of a vertex that can be associated to an MC particle, the vertex can be further classified as prompt or secondary based on the true lifetime of the MC particle. For consistency of terminology, signal vertices are referred to as having a ‘background category’ of ‘signal’.

In general, separate samples of simulated data are produced on a per-analysis basis, with independent samples generated for different beam energies and dipole magnet polarities (usually only ‘up’ and ‘down’). The generation of MC is done using a globally distributed network of computers, and the resulting data files are then made available to analysts in the same way as for real data, with the additional inclusion of the truth-matching information.

Part II

Charm production

5 Introduction

This part describes a measurement of charm and $c\bar{c}$ production, performed using data taken at $\sqrt{s} = 13\text{ TeV}$ with the LHCb detector. At the LHC, a given charm hadron H_c can either be produced in the proton-proton interaction or in the decays of B hadrons. The latter mechanism is called secondary production, whilst the former is called prompt production. The type of charm hadron produced can either be charmonium, such as the J/ψ meson, or open charm, whose net charm quantum number is non-zero. This measurement is concerned with the prompt production of open charm hadrons. Charm hadrons originating from the strong decay of prompt excited charm states are also considered to be prompt, and here no attempt is made to distinguish between the two sources.

Measurements of $c\bar{c}$ and $b\bar{b}$ cross-sections at the LHC are made to test the predictions of QCD. The production rate of heavy flavour hadrons is expected to vary as a function of hadron p_T and rapidity y , defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \quad (5.1)$$

and so experimental measurements can be compared with both the shape of theoretical predictions as a function of (p_T, y) as well as to the total integrated rate. Measurements can also be used as inputs to constrain predictions. Heavy flavour production rates are also used to estimate sensitivities of experimental searches for rare decays, and also as inputs to atmospheric neutrino experiments, where charm production enters as a background process.

Measuring $c\bar{c}$ and $b\bar{b}$ cross-sections at LHCb is complementary to such efforts made using the general purpose detectors (GPDs), due to the unique

rapidity range of the detector, as well as its ability to cleanly reconstruct low p_T hadrons.

5.1 Analysis overview

The measured cross-section for the inclusive production of an open charm hadron H_c from proton-proton collisions can be expressed as

$$\sigma(H_c) = \frac{N(H_c \rightarrow f)}{\varepsilon(H_c \rightarrow f) \cdot \mathcal{B}(H_c \rightarrow f) \cdot \mathcal{L}_{\text{int.}}}, \quad (5.2)$$

where $N(H_c \rightarrow f)$ is the number of reconstructed and selected prompt charm hadrons H_c in some final state f ; ε is the efficiency of fully reconstructing and selecting the decay $H_c \rightarrow f$, given that it was produced; $\mathcal{B}$ is the branching fraction of the decay; and $\mathcal{L}_{\text{int.}}$ is the integrated luminosity. To compare with the shapes of theoretical predictions, double differential cross-sections can be determined in bins of charm hadron p_T and y , using the width Δp_T of the bin in p_T and the width Δy of the bin in rapidity in which the yield and efficiency measurements were made

$$\left. \frac{d^2\sigma(H_c)}{dp_T dy} \right|_i = \frac{1}{\Delta p_T \Delta y} \frac{N_i(H_c \rightarrow f)}{\varepsilon_i(H_c \rightarrow f) \cdot \mathcal{B}(H_c \rightarrow f) \cdot \mathcal{L}_{\text{int.}}}, \quad (5.3)$$

where i is the index of the (p_T, y) bin.

The analysis presented in this part measures the double differential cross-sections of D^0 , D^+ , D_s^+ , and D^{*+} open charm mesons.¹ The final states chosen are $D^0 \rightarrow K^-\pi^+$, $D^+ \rightarrow K^-\pi^+\pi^+$, $D_s^+ \rightarrow \phi\pi^+$ with $\phi \rightarrow K^-K^+$, and $D^{*+} \rightarrow D^0\pi^+$ with $D^0 \rightarrow K^-\pi^+$, the branching fractions of which are listed in Table 5.1. Charge conjugation is implied, in that the cross-sections are measured as the combined rate of meson and anti-meson production. The measurements are made in the p_T range $0 < p_T < 15 \text{ GeV}/c$ and the rapidity range $2 < y < 4.5$. In y , the measurements are made in five bins of width 0.5, and in p_T the bins are $1 \text{ GeV}/c$ in width in the regions

¹The excited D state is, formally, the $D^*(2010)^+$ meson, but will be referred to throughout as ' D^{*+} '.

Table 5.1: Branching ratios for the different decay modes [12, 49]. The $D_s^+ \rightarrow \phi\pi^+$ branching fraction includes the branching fraction of $\phi \rightarrow K^- K^+$.

Decay mode	$\mathcal{B}$ (%)
$D^0 \rightarrow K^- \pi^+$	3.88 ± 0.05
$D^+ \rightarrow K^- \pi^+ \pi^+$	9.13 ± 0.19
$D_s^+ \rightarrow \phi\pi^+$	2.24 ± 0.13
$D^{*+} \rightarrow D^0 \pi^+$	67.7 ± 0.5

$0 < p_T < 1 \text{ GeV}/c$ and $4 < p_T < 15 \text{ GeV}/c$, and are $0.5 \text{ GeV}/c$ in width in the region $1 < p_T < 4 \text{ GeV}/c$.

The kinematic ranges imposed on the charm hadrons impose a fiducial region on the kinematic phase space, based on the fiducial volume of the detector, where there is instrumentation and where the detection efficiency is relatively high. In addition to such requirements on the parent, the child particles are similarly required to be within the detector's fiducial volume by requiring they occupy a narrow pseudorapidity range. The exact definition of these requirements will be given in Chapter 8.

From the double differential measurements, several derived quantities are computed, namely the ratio of charm production rates measured at $\sqrt{s} = 13 \text{ TeV}$ to those measured by LHCb at 7 TeV [10], ratios of production rates between charm mesons, open charm hadron cross-sections integrated across the (p_T, y) space, and integrated $c\bar{c}$ cross-sections. Each integrated per-meson cross-section is the sum over all (p_T, y) bins of the respective double differential measurements, and the integration is performed only with the fiducial kinematic range of the double differential measurements. The total $c\bar{c}$ cross-sections are computed using fragmentation fractions, each a measurement of the transition probability $f(c \rightarrow H_c)$ that a charm quark will hadronise to a particular open charm hadron. Given the fragmentation fraction and the integrated cross-section, the integrated $c\bar{c}$ cross-section is computed as

$$\sigma(pp \rightarrow c\bar{c}X) = \frac{\sigma(H_c + \bar{H}_c)}{2f(c \rightarrow H_c)}. \quad (5.4)$$

The factor $1/2$ accounts for the inclusion of the charge conjugate (anti-charm)

Table 5.2: *Charm hadron fragmentation fractions [50].*

H_c	$f(c \rightarrow H_c)$
D^0	0.565 ± 0.032
D^+	0.246 ± 0.020
D_s^+	0.224 ± 0.028
D^{*+}	0.080 ± 0.017

states in the measurement, which have been noted explicitly here to avoid ambiguity. For this analysis, the fragmentation fractions are taken from the Particle Data Group (PDG) [50], who compute them as averages of measurements made at e^+e^- colliders operating at a centre-of-mass energy close to the $\Upsilon(4S)$ resonance. These values are given in Table 5.2, where the transition probability for D^0 , $f(c \rightarrow D^0)$, includes contributions from $D^{*+} \rightarrow D^0\pi^+$ decays.

Treatment of uncertainties

Throughout the analysis, the method of Monte Carlo error propagation is used. With this, each quantity entering Equation (5.3) is assumed to be drawn from some probability density function (PDF) with a known set of parameters. Each quantity is then modelled as a vector of ten thousand ‘observations’, with each observation sampled from the corresponding PDF. Computations on scalar quantities then becoming element-wise operations on the sample vectors. Statistics on the resulting sample vector, such as the estimated mean and standard deviation, can then be computed. This technique is used as an analytical error analysis of the measurement is complex: what enters into the cross-section is the inverse of the efficiency, for example, the corresponding probability density function of which may not be obvious, in particular whether the uncertainties can be treated as Gaussian. The use of Monte Carlo error propagation simplifies such analyses.

An example cross-section, measured in the region $12 < p_T < 13 \text{ GeV}/c$, $2.5 < y < 3$, is given in Figure 5.1. The blue curve shows the PDF of the cross-section, as modelled by the MC error propagation technique, when

only the statistical uncertainty is considered, whilst the green curve also includes the contributions from the systematic uncertainties. The green curve is not symmetric, and the mode is shifted with respect to that of the blue.

Unless stated otherwise, each quantity that is measured directly is assumed to follow a normal distribution, whose mean is equal to the measured value and whose width is equal to the measured uncertainty on that value. The computation of that measured uncertainty will be described when the respective quantities are first discussed. When the results of operations using such inputs are presented, such as cross-sections, the central value given is the most probable value of the sample vector representing the result, and the lower and upper uncertainty bounds are given by the 15.9% and 84.1% percentiles either side of the central value.

Part structure

The data used for the measurements were collected in July 2016, under a set of special conditions used for the restart of the LHC after LS1. The conditions of the LHC and of the LHCb detector during this period will be described in Chapter 7, along with a description of the corresponding simulation and a summary of the integrated luminosity measurement. The reconstruction and selection of the collected data will be presented in Chapter 8. The extraction of the prompt charm yields is made on the fully reconstructed and selected data using the method of maximum likelihood fits, as described in Chapter 9. The evaluation of the associated efficiencies is performed using a mixture of simulation and dedicated calibration samples, and is discussed in Chapter 10, after which the systematic uncertainties associated with all procedures used in the measurement are evaluated, for which the description and results are given in Chapter 11. Finally, the cross-sections themselves are computed, and are presented in Chapter 12.

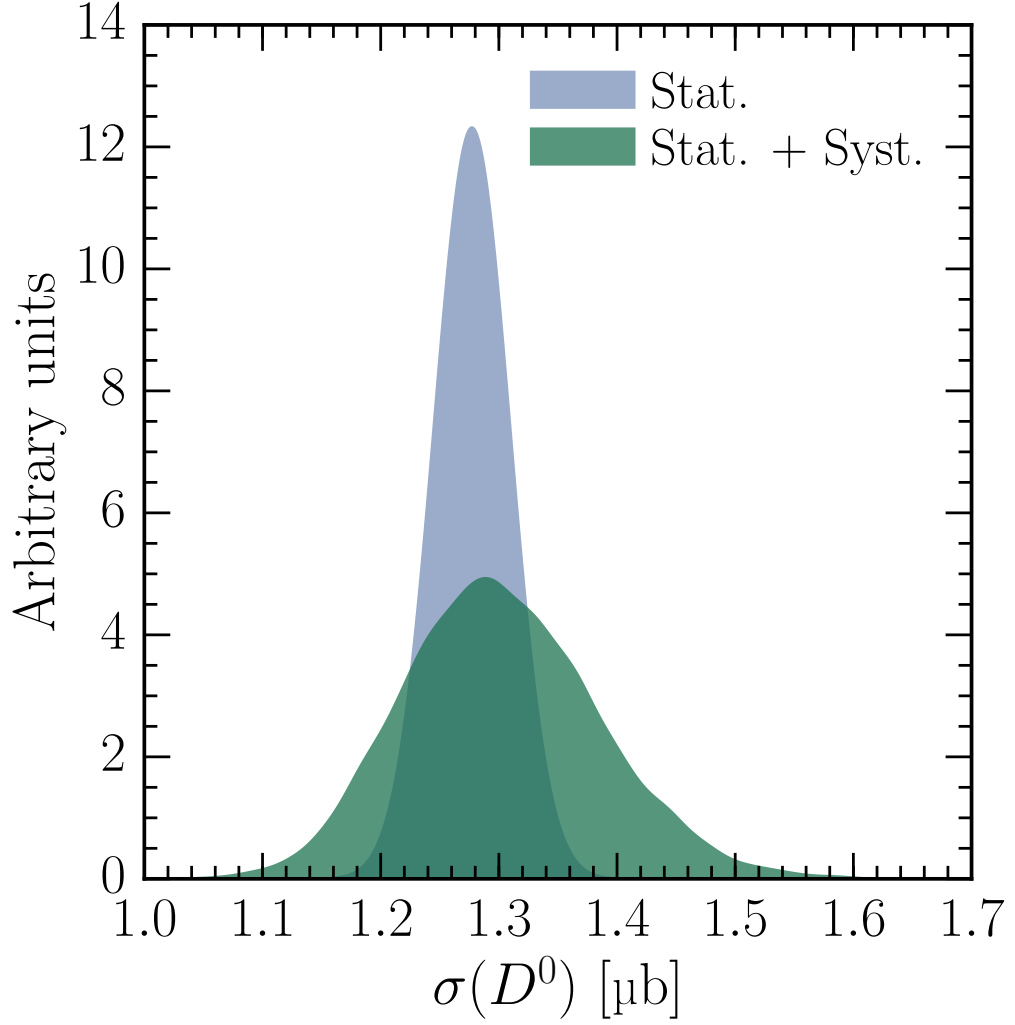


Figure 5.1: *Posterior probability densities for a measurement of the D^0 cross-section in the region $12 < p_T < 13 \text{ GeV}/c$, $2.5 < y < 3$. The blue curve shows the density including contributions only from the statistical uncertainty on the prompt signal count. The green curve shows the density including contributions from all uncertainties that enter the measurement.*

6 Theory and motivation

For an observable cross-section σ of inclusive production of some object P involving two initial-state hadrons h_A and h_B , $h_A h_B \rightarrow PX$, the dominant contribution to a prediction can be expressed as the convolution of several terms

$$\sigma = \sum_{a,b} \sigma_{a,b} \otimes f_{a,A} \otimes f_{b,B}, \quad (6.1)$$

where the sum runs over all parton types a and b in the respective hadrons h_A and h_B ; $\sigma_{a,b}$ is the cross-section of the hard partonic interaction $ab \rightarrow PX$; and f is the parton density function (PDF) for the interacting parton in the respective hadron [51, 52]. The parton-parton cross-sections $\sigma_{a,b}$, for processes such as heavy quark pair production via gluon-gluon fusion shown in Figure 4.1a, can be computed using perturbative QCD. The computations of the PDFs involve low-energy processes below the QCD scale, and so must be evaluated using non-perturbative QCD instead. As the PDFs cannot be computed from first principles, they must be derived by alternate means, and in practice are constrained using measurements of cross-sections by experiments. Additional measurements of cross-sections can constrain the PDFs further, improving their precision.

Parton density functions are parameterised in terms of x , the fraction of the total hadron momentum carried by the interacting parton, and Q^2 , a measure of momentum transfer and inversely proportional to the spatial ‘scale’ at which the hadron can be probed, with a larger Q^2 corresponding to a finer resolvable structure. The PDF $f(x, Q^2)$ for a particular hadron is an intrinsic property, and so, for example, experimental measurements that can constrain the PDF in high-energy proton collisions help not only the

experiments at that collider, but also experiments at other proton colliders. This is possible as the evolution of the PDFs over Q^2 for a given value of x is described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) differential equations [53–55]. Measuring a wide range of x is important as the individual contributions to the total PDF, such as from the gluons and the valence quarks, vary strongly as a function of x , as shown in Figure 6.1. The usage of the DGLAP equations requires an initial PDF from which to evolve to different scales, however, as the dependence on x cannot be computed perturbatively. Such inputs were first provided by deep inelastic scattering (DIS) experiments, which probe the proton structure by colliding leptons and protons. The subsequent evolution of the PDFs have been driven by additional experimental measurements, and by an improved theoretical understanding of the use of the perturbative expansion that must be employed when solving the DGLAP equations.

The treatment of experimental inputs is a matter of debate, as there are somewhat arbitrary choices to be made on what experimental inputs should be used, and on how the various experimental uncertainties should be treated. As a result, different groups of theorists provide so-called ‘PDF sets’, which differ not only in the treatment of the inputs, but also in the treatment of the *theoretical* uncertainties. These include, for example, the choice of order to which the perturbation series are evaluated, and whether the value of the perturbation series scale is a parameter of the fit or if it is fixed. For heavy flavour production, additional uncertainties arise from the choice of *fragmentation function* models [56], which describe how the heavy quarks transition into the observed hadrons, a factor which is not included in the factorisation theorem shown in Equation (6.1).¹ The different PDF sets can be used to compute different cross-section predictions, and so further experimental measurements can help to distinguish between predictions.

¹The fragmentation functions are generally assumed to be process independent, e.g. they are the same for heavy flavour hadron production from $q\bar{q}$ production at both e^-e^+ and pp colliders. They cannot be computed perturbatively, and are extracted from experimental measurements.

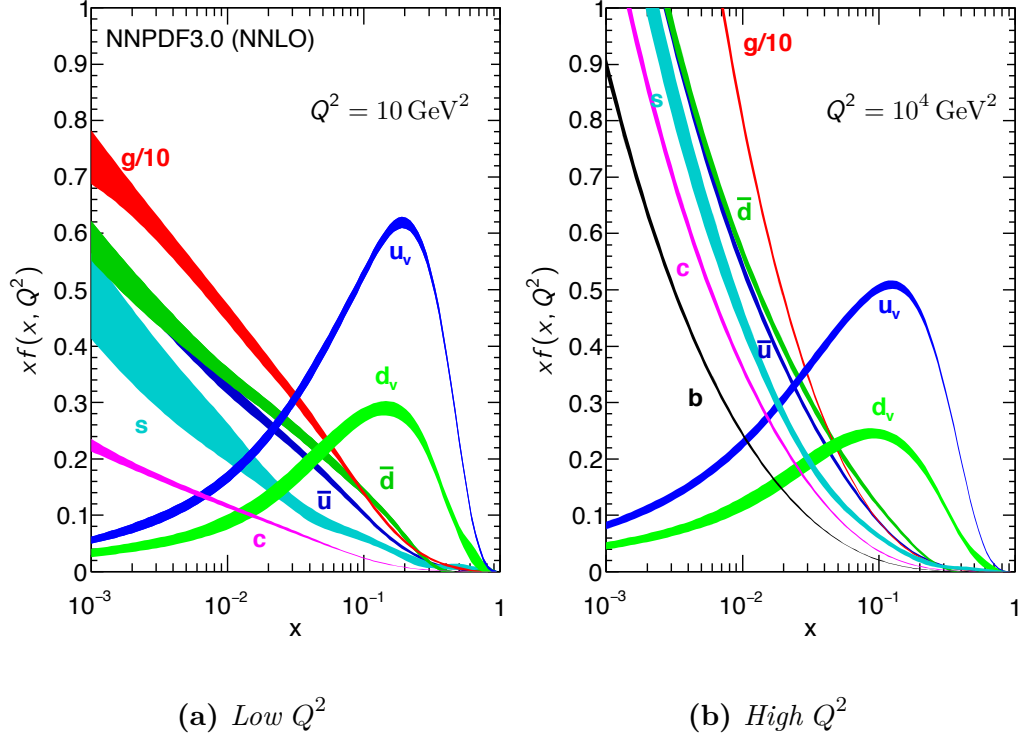


Figure 6.1: Sets of proton parton density functions for low Q^2 (a) and high Q^2 (b), illustrating both x and Q^2 dependence [12]. Each band represents the PDF for a specific parton, as a function of x , shown on the x -axis. The u_v and d_v bands show the contributions from the valance quarks, the g band shows the gluon contribution, and the other bands show the contribution from flavour-specific sea quarks.

Experimental input

Today, inputs from both DIS and hadron-hadron collider experiments serve to constrain proton PDFs over a wide range of x and Q^2 values. The most recent and most powerful DIS measurements have been made by the H1 and ZEUS collaborations at the HERA ep collider [57]. However, these measurements are largely constrained to measure values of x above 10^{-4} , leading to large uncertainties at very low x . This low- x region is particularly interesting to study not only because the current PDF uncertainties there are large, but also because it is where both the gluon and the sea quark contributions to the total PDF are large, as shown in Figure 6.1, and so

low- x measurements probe phenomenologically different processes.

In a hard $2 \rightarrow 2$ process, such as gg fusion or $q\bar{q}$ annihilation producing a heavy quark pair, the extent in x that can be probed with heavy flavour hadron production measurements is given by

$$e^{-y} \frac{\sqrt{p_T^2 + m_q^2}}{\sqrt{s}} < x < e^y \frac{\sqrt{p_T^2 + m_q^2}}{\sqrt{s}}, \quad (6.2)$$

where y and p_T parameterise the hadron kinematics, and m_q is the mass of the heavy quark in the hadron under study [58]. There is then a wide range of x values that can be probed at proton colliders, with the highest sensitivity to low x provided by measurements of charm production at high rapidity. At the energies reached by the LHC, the dominant production mechanism is gluon-gluon fusion, as depicted in Figure 4.1a, and so measurements at the LHC are able to provide particularly strong constraints on the gluon densities. Charm cross-sections have been measured at the LHC in the $|y| < 0.5$ region for $p_T > 1 \text{ GeV}/c$ at $\sqrt{s} = 2.76 \text{ TeV}$ and $\sqrt{s} = 7 \text{ TeV}$ by the ALICE collaboration [59–61], and for pseudorapidity $|\eta| < 2.1$ in the p_T region $3.5 < p_T < 100 \text{ GeV}/c$ at $\sqrt{s} = 7 \text{ TeV}$ by the ATLAS collaboration [62].

Measurements of charm production at $\sqrt{s} = 7 \text{ TeV}$ were made by the LHCb collaboration [10]. By Equation (6.2), measurements of charm hadron production for p_T values of the order of $1 \text{ GeV}/c$ at the lowest rapidities accessible to LHCb can probe x values down to the order of 10^{-6} , with lower values possible for measurements at a higher $\sqrt{s}$, as shown in Figure 6.2. These measurements, along with similar measurements of beauty hadron production at LHCb [11] and the H1 and ZEUS measurements, have been used by the PROSA collaboration [58] to constrain the proton parton density functions; they see a significant reduction in the gluon PDF uncertainty when including the LHCb data, with respect to using the HERA data only, as shown in Figure 6.3. It is expected that additional measurements, at $\sqrt{s} = 13 \text{ TeV}$, can be used to further reduce the gluon PDF uncertainties [63].

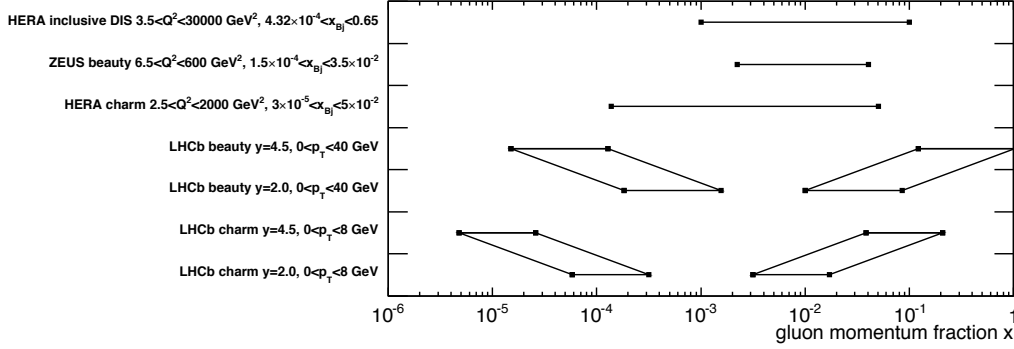


Figure 6.2: Regions of gluon momentum fraction x , on the x -axis, probed by different experiments, indicated on the y -axis [58], where the LHCb regions are for $\sqrt{s} = 7 \text{ TeV}$.

6.1 Atmospheric neutrino fluxes

The constraints imposed on the proton PDFs are also useful in atmospheric neutrino experiments. The IceCube collaboration aims to measure highly energetic neutrinos of astrophysical origin, with energies up to the PeV scale, using the IceCube Neutrino Observatory [64]. This is done to shed light on the mystery of ultra-high-energy cosmic rays (protons), whose source is unknown but which should also produce similarly energetic neutrinos. The so-called conventional neutrino flux constitutes one of the backgrounds for astrophysical neutrino experiments. This is made by cosmic rays interacting in the atmosphere and producing pions and kaons, which then decay to final states containing neutrinos. However, for the high energies the IceCube detector is sensitive to, beginning at around 100 GeV, this background is suppressed. The long lifetime of the mesons increases the probability they will interact with the atmosphere in flight, losing energy through hadronic interactions, hence softening the energy spectrum of the resulting neutrinos. An additional background comes from the production of charm hadrons in the atmosphere, which, with their comparatively short lifetimes, lose very little energy in flight, and so the neutrino spectrum from charm decays is very similar to the cosmic ray energy spectrum. The neutrinos produced in the charm decays can interact with the detector in a similar manner to the signal and exhibit a similar energy spectrum. It is then of key importance

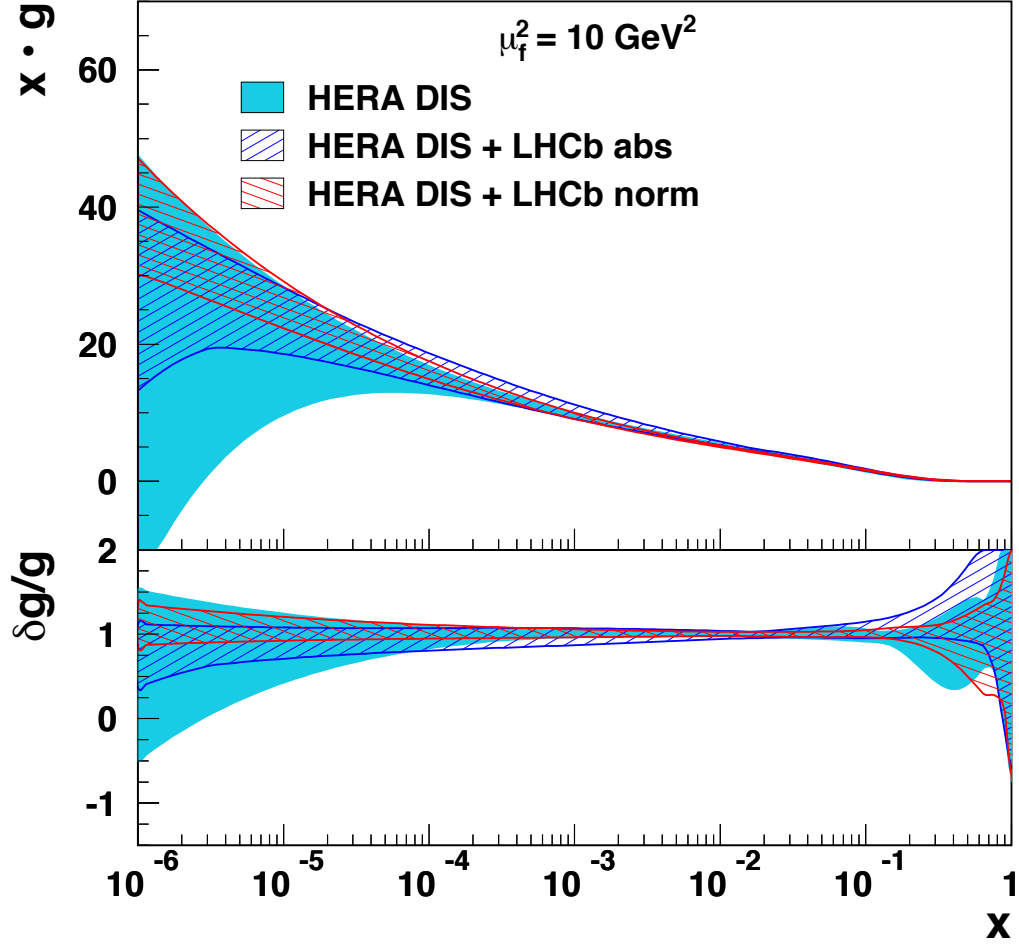


Figure 6.3: *Gluon PDFs after fitting to experimental data [58]. The x -axis shows the proton momentum fraction x , the y -axis on the top plot shows the gluon PDF density, and the y -axis on the bottom shows the relative uncertainty on the model. The solid band shows the fit when only using the HERA measurements as inputs, whilst the hatched bands show the fit when using both the HERA and $\sqrt{s} = 7 \text{ TeV}$ LHCb measurements. The “LHCb norm” measurements are the absolute cross-sections (“LHCb abs”) normalised to a specific (p_T, y) bin, which removes some global experimental uncertainties such as the luminosity.*

that the amount of expected background from these decays is known, else neutrinos from charm decays may be misinterpreted as neutrinos from astrophysical sources. This contamination can be estimated by predicting the proton-proton cross-sections (between the cosmic ray and the nuclei in the atmosphere).

Colliding equal-energy proton beams at a centre-of-mass energy of 13 TeV is equivalent to colliding a proton beam of energy 100 PeV into a fixed target. Hence, measuring charm cross-sections at $\sqrt{s} = 13$ TeV at the LHC can provide constraints on the expected rate of charm production in the atmosphere produced by extremely high-energy cosmic rays. Indeed, the measurements made by LHCb at $\sqrt{s} = 7$ TeV have already been used to that effect, and additional measurements at 13 TeV can help further [63, 65].

6.2 Cross-section comparisons with theory

A useful check of PDFs constrained with past experimental data is to compute cross-section predictions using them, for comparison with new experimental results. This was done in the previous LHCb open charm cross-section analysis, shown in Figure 6.5, where at low p_T the data generally lie above the theory prediction. Additional measurements can help in the understanding of these discrepancies.

For the analysis presented here, a measurement of charm production at $\sqrt{s} = 13$ TeV, three theory groups have provided predictions for the double differential cross-sections. Each set of predictions is computed at next-to-leading order in the strong coupling constant α_s , which is, for inclusive production as shown in Figure 6.4, up to the third power of α_s . The predictions include uncertainties from: the factorisation scale dependence, which is the scale Q^2 at which the perturbative expansion was made; and the renormalisation scale, which is the scale at which the expansion is renormalised to remove ultraviolet divergences. These affect both the evaluation of the partonic cross-sections and the evolution of the PDF sets with the DGLAP equations.

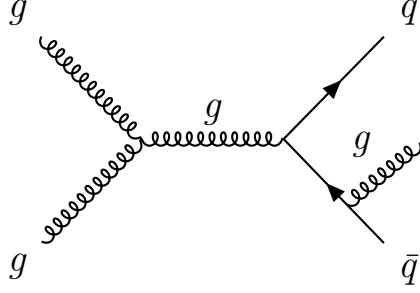


Figure 6.4: *Feynman diagram of next-to-leading order inclusive $q\bar{q}$ pair production via gluon-gluon fusion.*

The POWHEG+NNPDF3.0L predictions [63] use a PDF set that includes the constraints from the $\sqrt{s} = 7$ TeV LHCb measurement [10] (hence the “L” in the PDF set name), and are provided for all (p_T, y) bins for D^0 and D^+ mesons. The generalised-mass variable-flavour-numbering scheme (GMVFNS) predictions [66] are provided for $p_T > 3$ GeV/ c for all mesons under study, and use the CT10 PDF set. These predictions include a factor to account for the $c \rightarrow H_c$ transition probability, obtained by convolving the $c\bar{c} \rightarrow H_c$ cross-sections with the appropriate fragmentation functions. Finally, fixed-order next-to-leading logarithms (FONLL) predictions [67] using the NNPDF3.0 PDF sets are provided for the full (p_T, y) phase space for D^0 , D^+ , and D^{*+} mesons. The FONLL predictions assume unit $c \rightarrow H_c$ transition probability, and so are multiplied by the fragmentation fractions in Table 5.2 for comparison with the measurements. The comparison of these predictions with the data will be given in Chapter 12.

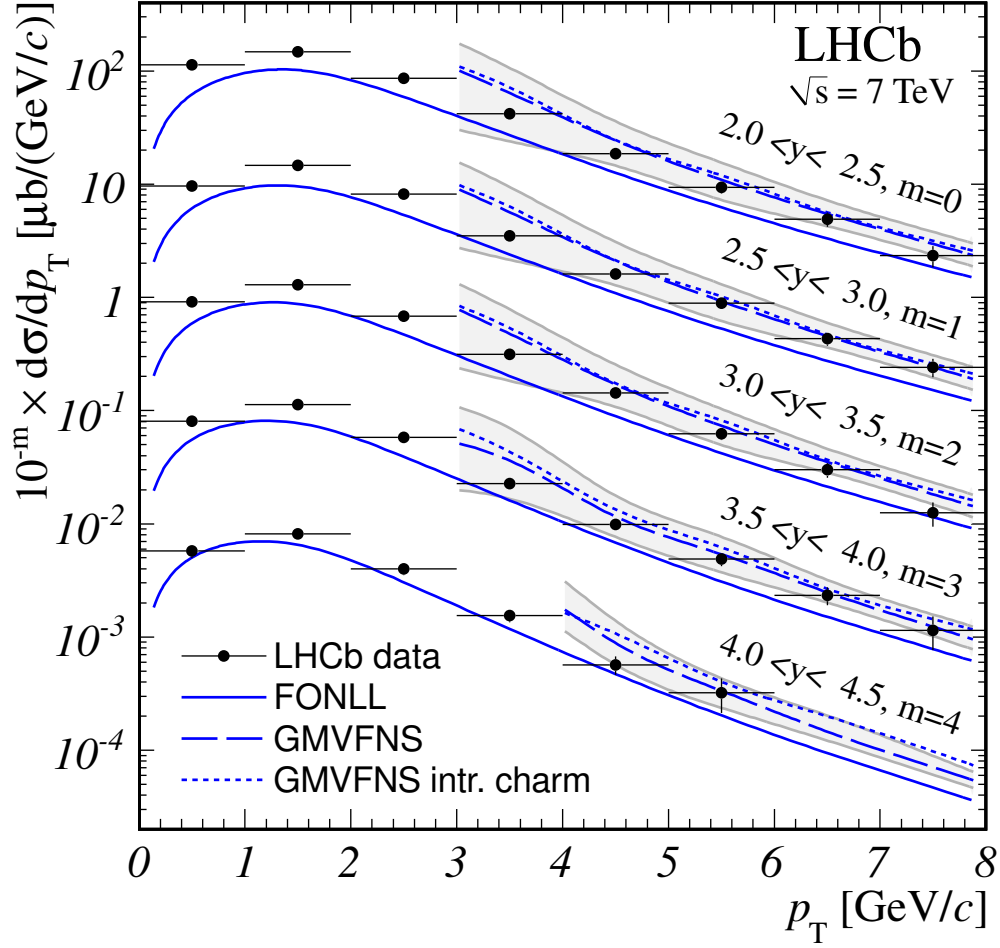


Figure 6.5: Prompt D^0 production rates, measured using the LHCb detector with proton-proton data taken at $\sqrt{s} = 7 \text{ TeV}$ [10], in (p_T, y) bins. The x -axis shows the p_T of the D^0 , and the measurements are offset by a multiplicative factor based on the rapidity bin. The shaded band on the GMVFNS predictions indicate theoretical uncertainties, in contrast to the FONLL predictions where only the central values are shown.

7 Datasets and running conditions

The data used for this measurement were collected at the beginning of Run 2 of the LHC, during a special 15-day ‘early measurements’ period. The number of bunches in the accelerator was gradually increased from 50 at the beginning to 482 bunches at the end of the period, with the number colliding at IP8 increasing from 36 to 397 bunches. In addition, the minimum bunch spacing was set to 50 ns, as in Run 1, rather than the nominal Run 2 spacing of 25 ns. These steps were taken primarily for machine safety, as the LHC began operating at a new centre-of-mass energy of 13 TeV.

The combined dataset used for the measurement corresponds to an integrated luminosity of

$$\mathcal{L}_{\text{int.}} = (4.98 \pm 0.19) \text{ pb}^{-1}. \quad (7.1)$$

The integrated luminosity measurement is performed using a beam-gas imaging technique [68]. For reference, the total integrated luminosity collected in Run 1 at LHCb was 3 fb^{-1} , whilst that for the $\sqrt{s} = 7 \text{ TeV}$ LHCb open charm production measurement was 15 nb^{-1} . All data were taken with the LHCb dipole magnet in the ‘down’ polarity, and were processed via the Turbo stream data flow, as described in Section 4.1. A matching set of simulated MC events is also used in the analysis, which is described in Section 7.1. The selection of events and charm candidates, both in the trigger and offline, will be described in Chapter 8.

7.1 Simulated data samples

This analysis uses simulated samples of $D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow K^- \pi^+$, $D^+ \rightarrow K^- \pi^+ \pi^+$, and $D_s^+ \rightarrow K^- K^+ \pi^+$ decays. The $D^{*+} \rightarrow D^0 \pi^+$ sample is sufficient for studying both $D^{*+} \rightarrow D^0 \pi^+$ and (untagged) $D^0 \rightarrow K^- \pi^+$ efficiencies, as the effect of the soft pion is parameterised in $D^0 p_T$ and y . In addition, the $D_s^+ \rightarrow K^- K^+ \pi^+$ sample is sufficient for studying $D_s^+ \rightarrow \phi \pi^+$ decays, given that the kaon pair is required to be within the same mass window as applied to the data. For each decay mode, samples of 2.5 million events are generated, with an additional 1 million events generated where the signal charm hadron is required to have $p_T > 10 \text{ GeV}/c$. To save computing resources, the signal decays are required to be within the LHCb geometric acceptance after EVTGEN has run. This requires all charged final-state particles to have a positive z component of their three-momentum and to satisfy

$$10 < \theta < 400 \text{ mrad}, \quad (7.2)$$

where θ is the polar angle. As this cut is not 100 % efficient in some (p_T, y) bins, its efficiency must be assessed, as described in Section 10.1, and so dedicated, generator-only datasets are also produced. These data are not processed beyond the EVTGEN step.

In addition to the usual truth-matching procedure discussed in Section 4.4, a prompt/secondary flag is computed for each candidate charm vertex, based on the true lifetime of each particle preceding it in the true decay chain. If the lifetime of any particle in the true ancestry of the charm candidate exceeds a threshold of 0.1 fs, the candidate is flagged as secondary.¹ In what follows, the MC data have been filtered such that only truth-matched, prompt charm hadron candidates remain, unless stated otherwise.

¹The order of magnitude of the lifetimes of the ground-state charm and beauty hadrons is between 0.1 and 1 ps.

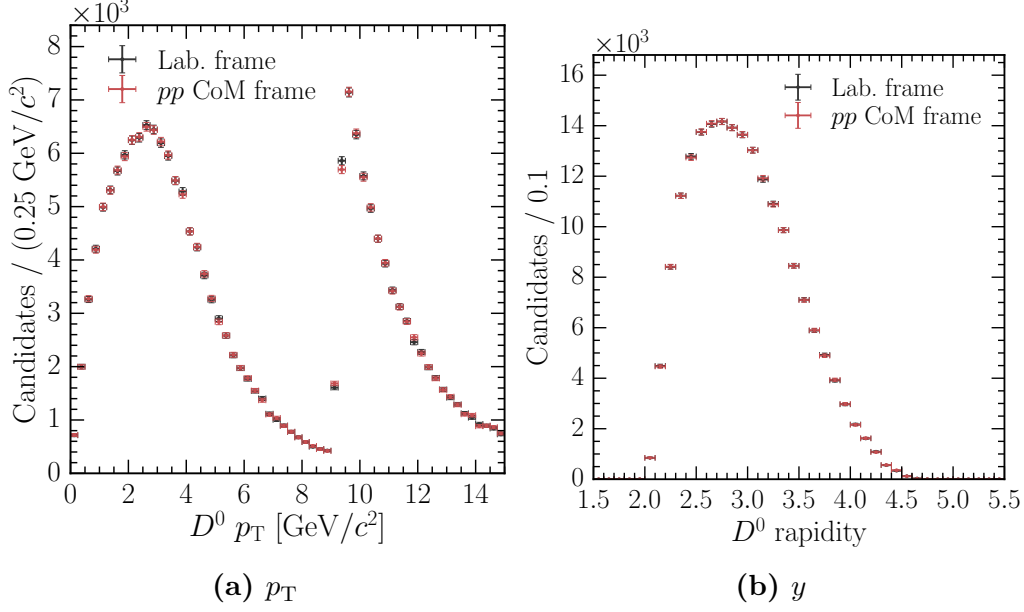


Figure 7.1: Distributions of $D^0 p_T$ (a) and y (b) as measured in the laboratory frame (black) and in the proton-proton centre-of-mass (CoM) frame (red), in the simulated $D^{*+} \rightarrow D^0 \pi^+$, with $D^0 \rightarrow K^- \pi^+$, dataset.

7.2 Crossing angle correction

Due to the non-zero crossing angles of the proton beams, the pp collision frame is boosted with respect to the laboratory frame. As this measurement is made in bins of charm hadron transverse momentum and rapidity measured in the pp collision rest frame, a correction is applied to the p_T and y measured in the laboratory frame. Example p_T and y distributions as measured in the laboratory and pp centre-of-mass frames are given for the simulated $D^{*+} \rightarrow D^0 \pi^+$ dataset in Figure 7.1. Whenever charm hadron p_T or y are mentioned, it is the pp rest frame quantities that are used, unless stated otherwise.

8 Event and candidate selection

The decay modes chosen for this analysis, $D^0 \rightarrow K^- \pi^+$, $D^+ \rightarrow K^- \pi^+ \pi^+$, $D_s^+ \rightarrow \phi \pi^+$ with $\phi \rightarrow K^- K^+$, and $D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow K^- \pi^+$, including their charge conjugates, are the amongst the most probable fully charged, hadronic final states for each charm meson. The advantage of reconstructing such final states is that it allows the invariant mass of the charm hadrons to be computed unambiguously, so that a yield determination can be performed by modelling the peaking signal structure in the mass distribution via fits. These fits will be discussed in detail in Chapter 9, but it is worth noting at this point that such fits are simpler if the data is ‘clean’, that is if the fraction of charm meson signal in the data is high. It is then advantageous to optimise the selection of the data to reduce backgrounds, whilst retaining as much signal as possible to maximise the statistical precision of the measurement.

The selection of fully charged, hadronic decay modes of charm mesons has already been performed by many other analyses at LHCb¹ and their selections share several common features. The first of these features is a requirement on the minimum charm candidate p_T , as there are many soft particles produced directly in the proton-proton collision. The second is to ask that the charm candidate vertex is significantly displaced from the PV, exploiting the long lifetimes of charm hadrons in the laboratory frame. Finally, it is required that the χ_{IP}^2 of the charm meson, a measure of the impact parameter (IP) significance, does not exceed a certain value. This

¹Representative examples include selections of $D^0 \rightarrow K^\mp \pi^\pm$ [32], $D^+ \rightarrow K^- K^+ \pi^+$ [69], and $D_s^+ \rightarrow \phi \pi^+$ [70] decays.

suppresses the background of charm candidates whose three-momentum does not point back to the PV. Each of these three features are powerful discriminators between signal and background.

For this measurement of prompt charm production it is preferable not to make p_T requirements on the charm candidate as the p_T range available for differential measurements would be restricted. Nor is it desirable to make χ_{IP}^2 cuts, as the χ_{IP}^2 distribution is used for prompt-secondary discrimination. The selection used is then informed by previous LHCb studies of fully charged, hadronic charm decays, but with modifications to improve the utility for the measurement in hand. The selection variables that are useful for discriminating between signal and background, in addition to those already discussed, are given and described below. The ‘parent’ is the charm hadron H_c , and the ‘children’ are the charged hadrons in the final state.

The child tracks are firstly selected based on the quality of the track fit, the track fit χ^2 per degree of freedom. For fake (ghost) tracks, the χ^2 is higher on average. Child p and η are required to be in the range in which calibration samples are available for PID efficiency measurements (described in more detail in Section 10.4). The track p_T is required to be above some minimum to reduce the background from random soft tracks originating from the PV, and similarly for the track χ_{IP}^2 , which will be high on average for tracks produced from a displaced, secondary vertex. Finally, PID requirements are made on the tracks to reduce misidentification backgrounds, particularly to suppress the large fraction of pions in the event when a kaon candidate is required.

The combination of tracks used to form an H_c candidate is required to have a maximum pairwise DOCA less than some value. If a vertex can be fitted using the combination, the χ^2 of the fit should be low, indicating a good compatibility with the physical vertex hypothesis. For long-lived charm hadrons, the vertex can be significantly displaced from the PV, and making a minimum requirement on the displacement suppresses much combinatorial background. The momentum vector of the H_c candidate should have the same direction as that of the vector that points from the PV to the decay vertex, and the angle between these two vectors, the direction angle (DIRA),

is required to be less than some value. Finally, the reconstructed mass of the parent vertex must be near the nominal mass of the hadron under study. A suitably large window must be used in order to be able to accurately model the combinatorial background shape in the mass fits.

The numerical requirements made on the discriminatory quantities were not tuned to optimise any specific figure of merit. Instead, the values were chosen based on those used in previous charm analyses, generally being loosened to account for the higher rate available to charm triggers during the early measurements period, and the desire to have as high an efficiency as possible in the low p_T regions of the measurement.

As this analysis uses the Turbo stream, the only selection is that defined using ‘online’ quantities computed in the trigger. The same quantities are used ‘offline’, with no further reconstruction being performed. The description of the selection is split in two: the ‘online’ part, which is the selection defined in the trigger and cannot be tuned once the data have been taken, and the ‘offline’ selection, which is any selection made after the online.

8.1 Online selection and reconstruction

The different purposes and hierarchy of the different stages of the LHCb trigger are described in Section 4.2. For this analysis, at L0 bunch-bunch crossings are accepted randomly by the NOBIAS trigger. The rate at which the NOBIAS trigger accepts events is programmed into the hardware controller manually, with the exact rate depending on the number of colliding bunches at LHCb. The rates were tuned throughout the early measurements period so as to keep the detector ‘dead time’ to a minimum.² The L0 rate is given for each fill in Table 8.1.

In HLT1, the following sequence is performed for events passing L0:

²The detector dead time is the period during which triggers by the L0 are ‘lost’ due to the read-out buffer being full. The buffer is able to hold 16 events, each of which takes 40 LHC clock cycles, of 25 ns each, to be sent to the HLT. Dead time can then occur if several triggers occur too close to one another [71].

1. The full offline charged particle (track) reconstruction is run for hits inside the VELO sub-detector;
2. The full offline PV reconstruction is run, using as input only the VELO tracks reconstructed in the previous step; and
3. The reconstruction of ‘long’ tracks, by first extrapolating VELO tracks into the TT and then, if at least two matching hits are found in the TT, extrapolating through the magnetic field. To decrease the time taken to assess all combinatorial combinations of VELO+TT tracks with hits in the IT and OT, only hits within a search window assuming a p_T of at least 800 MeV/c are considered.

Requiring all child tracks from a vertex to have $p_T > 800$ MeV/c would be inefficient for low p_T charm hadrons, and so the full decay is not reconstructed in HLT1. Instead, the presence of a single, good quality track is required, which is consistent with having originated from a heavy flavour decay by having a large χ^2_{IP} . The full set of selection requirements is given in Table 8.2.

HLT2 performs the full, offline-quality event reconstruction on events passing HLT1. One HLT2 line is defined per charm decay topology, corresponding to: one line for $D^0 \rightarrow K^- \pi^+$; one line for the charged three-body decays, specifically $D^+ \rightarrow K^- \pi^+ \pi^+$ and $D_s^+ \rightarrow K^- K^+ \pi^+$; and one line for $D^{*+} \rightarrow D^0 \pi^+$. In the D^0 , D^+ , and D_s^+ lines, candidates are made by combining long tracks whose particle hypotheses are consistent with the final state under consideration. In the D^{*+} line, D^0 candidates passing the D^0 HLT2 line are combined with pion candidates, referred to here as ‘soft pions’, π_{Soft}^+ . There are three stages to the particle combination: when cuts are applied to the input tracks; when cuts are applied to the four-vector sum of these tracks; and when cuts are applied to the fitted vertex. To save computing resources, the combination does not proceed past a failing step, and combinations where the vertex fit fails to converge are discarded, although the likelihood of a failing vertex fit is suppressed by making DOCA requirements beforehand. An event is saved for offline analysis if at least one HLT2 line ‘fired’, that is if a charm candidate is made that passes all

Table 8.1: List of fills used in the analysis, along with the integrated luminosity $\mathcal{L}_{\text{int.}}$, the L0 NOBIAS rate, and the corresponding effective L0 efficiency (“eff.”) for each fill.

Fill number	$\mathcal{L}_{\text{int.}}(\text{nb}^{-1})$	L0 NOBIAS rate (kHz)	L0 NOBIAS eff. (%)
3976	628 ± 24	250	20.58
3981	1101 ± 43	300	10.84
3988	931 ± 36	300	10.84
3992	1310 ± 51	350	7.84
3996	999 ± 39	350	7.84

Table 8.2: Requirements made on the track that fires the HLT1 trigger line.

Particle	Variable	Cut value
Any track	p_{T}	$> 800 \text{ MeV}/c$
	p	$> 3 \text{ GeV}/c$
	Track χ^2	< 3
	χ_{IP}^2	> 10
	VELO hits	> 9

cuts. The $D^0 \rightarrow K^- \pi^+$, $D_{(s)}^+ \rightarrow h^- h^+ h^+$, and $D^{*+} \rightarrow D^0 \pi^+$ requirements are given in Tables 8.3 to 8.5. Distributions of the invariant masses of the charm candidates after the trigger selection are given in Figures 8.1a, 8.2a, 8.3a, and 8.4a.

8.2 Offline selection

To improve the purity of the samples, the pion PID requirement is tightened to $\text{DLL}_{K-\pi} < 3$. Furthermore, kinematic requirements are made on all D^0 , D^+ , and D_s^+ children such that they are within the fiducial region of phase space available for PID efficiency calibration, specifically $3 < p < 100 \text{ GeV}/c$ and $2 < \eta < 5$. As the three-body samples contain a large fraction of combinatorial background even after these additional cuts, the vertex fit χ^2 and the direction angle cuts are tightened to $\chi^2 < 6$ and $\text{DIRA} < 14 \text{ mrad}$.

To reduce the combinatorial background further in the $D_s^+ \rightarrow K^- K^+ \pi^+$

Table 8.3: Requirements made in the HLT2 $D^0 \rightarrow K^- \pi^+$ selection. The track χ^2 criterion is applied in the reconstruction and listed here for completeness.

Particle	Variable	Cut value
$\pi^\pm, K^\pm$	p_T	$> 250 \text{ MeV}/c$
	p	$p > 2 \text{ GeV}/c$
	Track χ^2	< 3
	χ_{IP}^2	> 16
$\pi^\pm$	$\text{DLL}_{K-\pi}$	< 5
$K^\pm$	$\text{DLL}_{K-\pi}$	> 5
D^0	$m(K^- \pi^+)$	$1784 < m < 1944 \text{ MeV}/c^2$
	K to π DOCA	$< 0.1 \text{ mm}$
	Vertex fit χ^2	< 10
	Direction angle	$< 17 \text{ mrad}$
	Vertex displacement χ^2	> 49

data, a requirement is made on the kaon pair as

$$999.46 < m(K^- K^+) < 1039.46 \text{ MeV}/c^2. \quad (8.1)$$

This corresponds to a $\pm 20 \text{ MeV}/c^2$ window around the nominal mass of the $\phi(1020)$ meson [12], which is referred to here as just ‘ ϕ ’. Around half of all $D_s^+ \rightarrow K^- K^+ \pi^+$ decays proceed via the $\phi \rightarrow K^- K^+$ channel [12], and the narrow width of the ϕ resonance makes this cut efficient at reducing the fraction of combinatorial background in the data, as well as reducing the fraction of $\pi \rightarrow K$ mis-identifications. As the $D_s^+ \rightarrow K^- K^+ \pi^+$ sample is then dominated by resonant $D_s^+ \rightarrow \phi \pi^+$, the D_s^+ sample is referred as ‘ $D_s^+ \rightarrow \phi \pi^+$ ’. The effective reduction of the total number of D_s^+ mesons produced due to the ϕ requirement is treated as part of the branching fraction contribution in Equation (5.3), rather than as part of the selection efficiency. The $D_s^+ \rightarrow \phi \pi^+$ branching fraction used is that given by the CLEO collaboration, where the same $K^- K^+$ mass window definition is used [49], and hence includes the branching fraction of the $\phi \rightarrow K^- K^+$ decay.

Table 8.4: Requirements made in the HLT2 $D_{(s)}^+ \rightarrow h^- h^+ h^+$ selection. The D^+ and D_s^+ candidates are defined according to the three-body mass window after the vertex fit. Cuts of the form $x > x_1, x_2, x_3$ require that all particles satisfy $x > x_1$, at least two satisfy $x > x_2$, and at least one satisfies $x > x_3$. The track χ^2 criterion is applied in the reconstruction and listed here for completeness.

Particle	Variable	Cut value
$\pi^\pm, K^\pm$	p_T	$> 200, 400, 1000 \text{ MeV}/c$
	p	$p > 2 \text{ GeV}/c$
	Track χ^2	< 3
	χ_{IP}^2	$> 4, 10, 50$
$\pi^\pm$	$\text{DLL}_{K-\pi}$	< 5
$K^\pm$	$\text{DLL}_{K-\pi}$	> 5
D^+	$m(h^- h^+ h^+)$	$1789 < m < 1949 \text{ MeV}/c^2$
D_s^+	$m(h^- h^+ h^+)$	$1889 < m < 2049 \text{ MeV}/c^2$
D^+ / D_s^+	Vertex fit χ^2	< 25
	Direction angle	$< 35 \text{ mrad}$
	Vertex displacement χ^2	> 16
	Vertex lifetime τ	$> 0.150 \text{ ps}$

Table 8.5: Requirements made in the HLT2 $D^{*+} \rightarrow D^0 \pi^+$ selection. The input D^0 candidates are those passing the D^0 selection defined in Table 8.3. The track χ^2 criterion is applied in the reconstruction and listed here for completeness.

Particle	Variable	Cut value
$\pi^\pm$	p_T	$> 100 \text{ MeV}/c$
	Track χ^2	< 3
D^{*+}	$m(D^{*+}) - m(D^0)$	$130 < m < 160 \text{ MeV}/c^2$
	Vertex fit χ^2	< 25

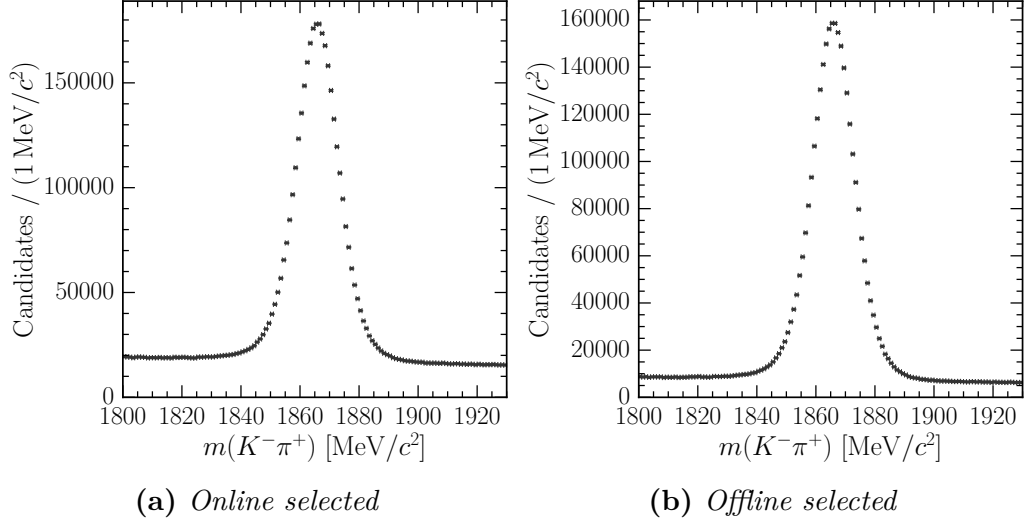


Figure 8.1: Mass distributions of $D^0 \rightarrow K^- \pi^+$ candidates after the online (trigger) selection (a) and after the offline selection (b).

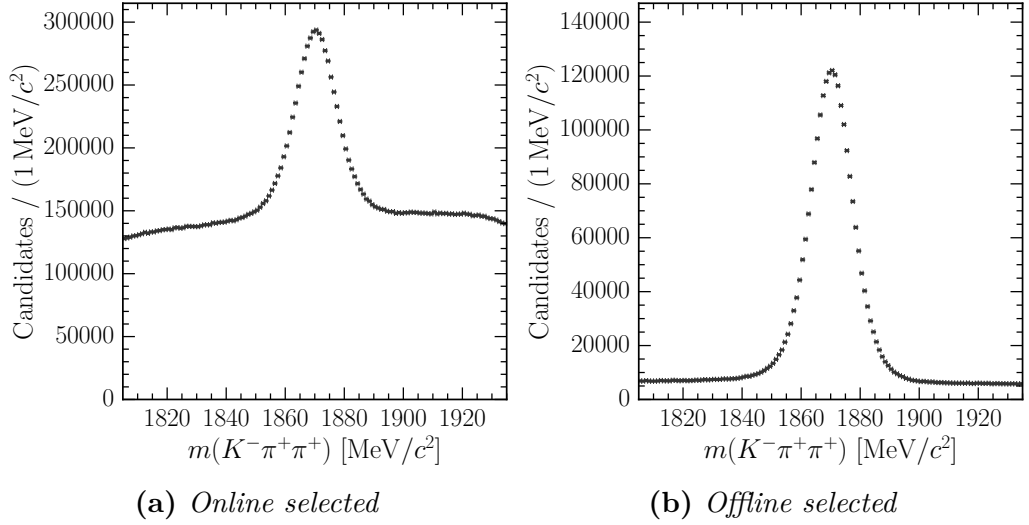


Figure 8.2: Mass distributions of $D^+ \rightarrow K^- \pi^+ \pi^+$ candidates after the online (trigger) selection (a) and after the offline selection (b).

The charm candidate invariant mass distributions after the full offline selection are given in Figures 8.1b, 8.2b, 8.3b, and 8.4b. The number of charm candidates before and after the offline requirements are given in Table 8.6.

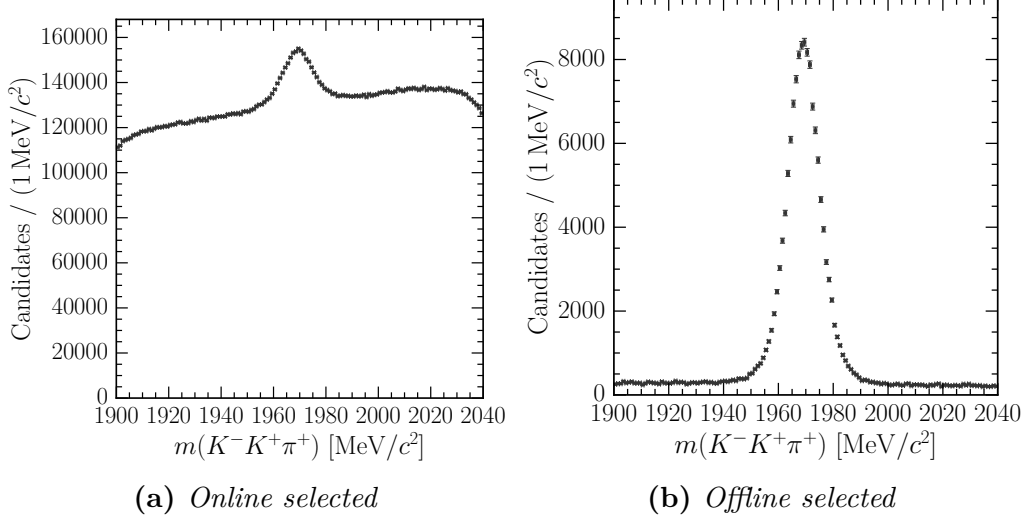


Figure 8.3: Mass distributions of $D_s^+ \rightarrow K^- K^+ \pi^+$ candidates after the online (trigger) selection (a) and after the offline selection (b). The offline selection includes the requirement that the kaon pair are within a $\pm 20 \text{ MeV}/c^2$ window around the nominal $\phi(1020)$ mass.

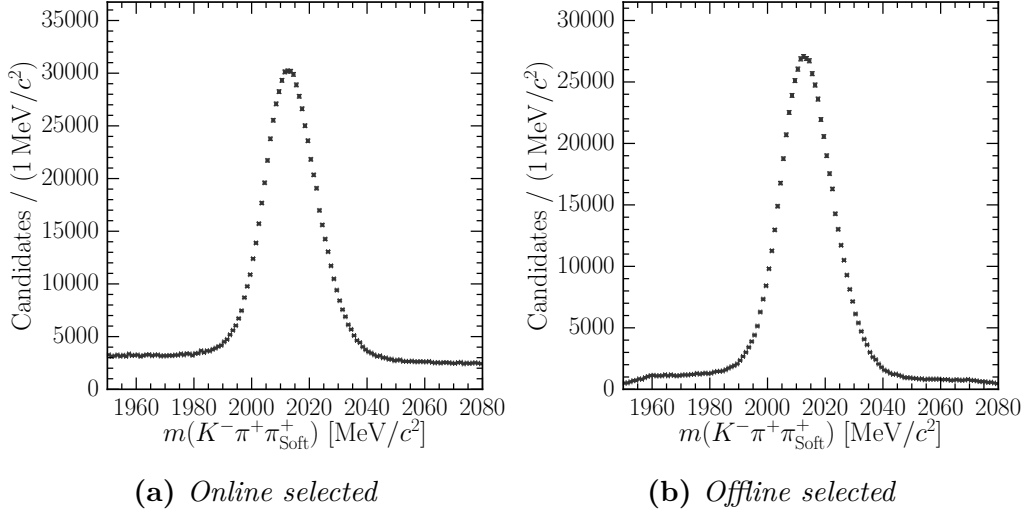


Figure 8.4: Mass distributions of $D^{*+} \rightarrow D^0 \pi^+$ candidates after the online (trigger) selection (a) and after the offline selection (b).

Table 8.6: *Number of candidates before and after the offline selection for each charm candidate under study.*

Decay mode	$N_{\text{Online}} (\times 10^6)$	$N_{\text{Offline}} (\times 10^6)$	$N_{\text{Offline}}/N_{\text{Online}} (\%)$
$D^0 \rightarrow K^- \pi^+$	5.8	3.8	65
$D^+ \rightarrow K^- \pi^+ \pi^+$	25	3.1	12
$D_s^+ \rightarrow K^- K^+ \pi^+$	21	0.16	0.8
$D^{*+} \rightarrow D^0 \pi^+$	1.1	0.72	66

9 Yield extraction

Signal charm mesons can be distinguished from combinatorial background in the invariant mass distribution, as shown in Figure 8.1b for example. However, the contributions from prompt and secondary charm are identical in the invariant mass, and so another distribution must be used to distinguish between the two sources of charm decays.

This analysis uses the logarithm of the χ^2 of the impact parameter, $\ln \chi_{\text{IP}}^2$, to obtain the prompt charm yields. The momentum vector of prompt charm should point back to the PV, and so the $\ln \chi_{\text{IP}}^2$ distribution is expected to be centred around zero. For secondary charm the $\ln \chi_{\text{IP}}^2$ distribution is expected to peak at higher values. This separation is exploited to distinguish between prompt and secondary charm, where the prompt signal yield in each (p_{T}, y) bin is measured by a binned, extended maximum likelihood fit to the $\ln \chi_{\text{IP}}^2$ distribution. The prompt and secondary models in this fit are described analytically as peaking functions. The combinatorial background is described by a non-parametric model taken from the data outside the signal peak in the mass distribution. This background model is normalised in the $\ln \chi_{\text{IP}}^2$ fit to the number of combinatorial background candidates in a signal window defined in the mass distribution, measured in a fit to the charm hadron mass spectrum.

For all cases but the D^{*+} measurement, it is the mass of the charm hadron that is used to discriminate all signal from combinatorial background. For the D^{*+} measurement, the delta mass distribution is used instead, defined as the difference Δm between the reconstructed D^{*+} mass and the reconstructed D^0 mass

$$\Delta m = m(K^- \pi^+ \pi_{\text{Soft}}^+) - m(K^- \pi^+). \quad (9.1)$$

This has a nominal value of $145.4 \text{ MeV}/c^2$ [12]. The narrower signal peak allows for a better signal-to-background ratio.

9.1 Signal-background discrimination

For each one-dimensional mass and Δm fit, one PDF is defined per discriminatory species: signal and combinatorial background. In both the mass and Δm distributions, the prompt and secondary signal shapes are assumed to be identical, and for the Δm fits the combinatorial background is assumed to be indistinguishable from random soft pion backgrounds, where a true D^0 is combined with a random track in the event. The total PDF is constructed as the sum of the per-species PDFs, each f_s , each weighted by the respective yield N_s

$$f(m) = \frac{1}{\sum_s N_s} \sum_s N_s f_s(m), \quad (9.2)$$

$$f(\Delta m) = \frac{1}{\sum_s N_s} \sum_s N_s f_s(\Delta m). \quad (9.3)$$

For every mass fit, the background PDF $f_{\text{Bkg.}}(m)$ is taken to be a first-order polynomial, and the signal PDF $f_{\text{Sig.}}(m)$ is dependent on the charm hadron candidate, with details given in Section 9.3. The combinatorial and random soft pion background in the Δm fits are modelled as a single empirical threshold function of the (un-normalised) form

$$R(x; \Delta m_0, A) = (x - \Delta m_0)^A, \quad (9.4)$$

where Δm_0 is the threshold value, fixed to the charged pion rest mass $m_{\pi^\pm} = 139.57 \text{ MeV}/c^2$ [12]. The signal PDF in Δm is the sum of three normal distributions, sharing a common mean but allowed to have different widths.

One mass or Δm PDF, $f(m)$ or $f(\Delta m)$, is constructed per (p_T, y) bin, and the likelihood is formed as the product of these such that each PDF is fitted simultaneously when the negative log-likelihood (NLL) is minimised.

9.2 Prompt-secondary discrimination

The prompt and secondary signal distributions are modelled by continuous, parametric PDFs. Rather than attempting to parameterise the combinatorial background distribution, a kernel density estimate (KDE) PDF [72], or ‘template’, is created from a histogram of the $\ln \chi_{\text{IP}}^2$ distribution in the lower and upper sidebands of the reconstructed charm hadron mass. In the case of the D^{*+} measurements, the data in the upper sideband of the Δm distribution are used instead. This approach assumes that the mass and $\ln \chi_{\text{IP}}^2$ are uncorrelated in the background sample, such that the $\ln \chi_{\text{IP}}^2$ shape in the signal region is the same as that in the sidebands.

The signal region in each long-lived charm hadron mass distribution is defined as a $40 \text{ MeV}/c^2$ -wide window centred on the nominal rest mass of the given charm hadron, taken to be $1864.84 \text{ MeV}/c^2$, $1869.61 \text{ MeV}/c^2$, and $1968.30 \text{ MeV}/c^2$ for D^0 , D^+ , and D_s^+ [12]. The lower sideband is a $20 \text{ MeV}/c^2$ -wide window centred $50 \text{ MeV}/c^2$ below the centre of the signal window, and the upper sideband is a $20 \text{ MeV}/c^2$ -wide window centred $50 \text{ MeV}/c^2$ above the centre of the signal window. For the D^{*+} measurements, the signal region is defined as a window $6 \text{ MeV}/c^2$ wide, centred on the nominal Δm value of $145.43 \text{ MeV}/c^2$ [12]. The upper sideband, or just ‘sideband’ as a lower sideband region is not defined in Δm , is defined as the region from $4.5 \text{ MeV}/c^2$ to $9 \text{ MeV}/c^2$ above the centre of the signal region. Figures 9.1 and 9.2 show the signal and sideband regions for each mode.

The integral of the $\ln \chi_{\text{IP}}^2$ background template PDF is Gaussian-constrained to the number of background candidates in the mass or Δm signal region. To avoid any complications due to correlations between the background template PDF and the fitted data, only the data in signal region are used in the $\ln \chi_{\text{IP}}^2$ fit. For the D^{*+} measurements, this signal region requirement is made in both the D^0 mass and the Δm distributions.

For each one-dimensional $\ln \chi_{\text{IP}}^2$ fit, one PDF f_s per discriminatory species is assigned: prompt signal, secondary signal, and combinatorial background. The total PDFs are constructed as the sum of per-species

PDFs, each weighted by the respective yield N_s

$$f(\ln \chi_{\text{IP}}^2) = \frac{1}{\sum_s N_s} \sum_s N_s f_s(\ln \chi_{\text{IP}}^2). \quad (9.5)$$

The constraint on the background yield $N_{\text{Bkg.}}$ is applied by multiplying the likelihood by a normal distribution in $N_{\text{Bkg.}}$, whose mean is the number of background candidates measured by the mass fit and whose width is the uncertainty on that value. This penalises the likelihood by reducing its value when the fitted $N_{\text{Bkg.}}$ parameter is far away from the value found in the mass fit.

One $\ln \chi_{\text{IP}}^2$ PDF and background constraint is constructed per (p_T, y) bin, and the likelihood is formed as the product of these such that each PDF is fitted simultaneously when the negative log-likelihood is minimised. The benefit of this construction is that some shape parameters can be shared across bins, reducing the uncertainty on the fit parameters. Which parameters are shared across bins and which are fitted independently is dependent on the type of charm meson. The shared and independent parameters are enumerated in Section 9.3.

The prompt signal model H in the $\ln \chi_{\text{IP}}^2$ distribution is a modified normal distribution, where the width is allowed to be asymmetric with respect to the mean, and the tails are described by exponential functions,

$$H(x; \mu, \sigma, \epsilon, \rho_L, \rho_R) = \begin{cases} \exp\left(\frac{\rho_L^2}{2} + \rho_L \frac{x-\mu}{(1-\epsilon)\sigma}\right) & x < \mu - (\rho_L \sigma (1-\epsilon)), \\ \exp\left(-\left(\frac{x-\mu}{\sqrt{2}\sigma(1-\epsilon)}\right)^2\right) & \mu - (\rho_L \sigma (1-\epsilon)) \leq x < \mu, \\ \exp\left(-\left(\frac{x-\mu}{\sqrt{2}\sigma(1+\epsilon)}\right)^2\right) & \mu \leq x < \mu + (\rho_R \sigma (1+\epsilon)), \\ \exp\left(\frac{\rho_R^2}{2} - \rho_R \frac{x-\mu}{(1+\epsilon)\sigma}\right) & x \geq \mu + (\rho_R \sigma (1+\epsilon)), \end{cases} \quad (9.6)$$

where the parameter μ is the mode of the distribution; σ is the average of the left and right widths; ϵ is the asymmetry between the left and right widths; and $\rho_{L(R)}$ is the exponent for the left (right) tail. The secondary signal distribution is modelled by a normal distribution. These shapes are motivated by the $\ln \chi_{\text{IP}}^2$ distributions for prompt and secondary signal decays observed in the MC samples.

Given the large number of parameters to be fitted, some shape parameters in the prompt and secondary signal $\ln \chi^2_{\text{IP}}$ PDFs are fixed to values obtained from ‘prefits’ to the pure samples of prompt and secondary signal MC described in Section 7.1. The set of parameters that is fixed is dependent on the charm meson species.

9.3 Mode-specific details

As different final states exhibit different features in the mass, Δm , and $\ln \chi^2_{\text{IP}}$ distributions, several aspects of the fits are different between modes. This section gives the list of parameters that are split across (p_T, y) bins, and the list of parameters that are allowed to float after the MC prefits, for each mode.

Initially, no knowledge is assumed on which parameters should be independent (to be ‘split’) across the (p_T, y) bins, nor of whether the $\ln \chi^2_{\text{IP}}$ distribution is well-modelled by the MC. The chosen parameterisation is that which best describes the data. For the mass fits, all parameters are initially fixed such that they are same across all (p_T, y) bins, and are then split one by one when it is observed that a feature varies across the bins, such as the width of the signal peak. All parameters of the $\ln \chi^2_{\text{IP}}$ fits are first kept fixed to values found from the fits to the simulated data, and then individual parameters are floated if the total PDF does not model the data well. Parameters are then split across (p_T, y) bins in the same manner as described below.

Fit details for D^0

For the $D^0 \rightarrow K^- \pi^+$ mode, the signal shape in the mass fit is the sum of a normal and a Crystal Ball distribution [73], sharing a common mode but allowed to have different widths. The Crystal Ball function C is a normal distribution but with a power-law tail on one side, defined (un-normalised)

as

$$C(x; \mu, \sigma, \alpha, n) = \begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma^2}} & \frac{x-\mu}{\sigma} > -\alpha, \\ e^{-\frac{|\alpha|^2}{2}} \left(\frac{n}{|\alpha|}\right)^n \left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma}\right)^{-n} & \frac{x-\mu}{\sigma} \leq -\alpha, \end{cases} \quad (9.7)$$

where α defines where the power-law tail begins, in units of the width of the core Gaussian σ , and n is the power law exponent. The mode and width of the total signal PDF are split across (p_T, y) bins, as is the slope of the background PDF. The mode of the $\ln \chi_{\text{IP}}^2$ prompt signal PDF and the width of the secondary signal PDF are also split across (p_T, y) bins. The mode and width of the prompt signal $\ln \chi_{\text{IP}}^2$ PDF and all parameters of the secondary signal $\ln \chi_{\text{IP}}^2$ PDF are floated in the fit to data, and all other $\ln \chi_{\text{IP}}^2$ shape parameters are fixed to the values from MC.

Prefits to the prompt signal and secondary signal MC samples are given in Figure 9.3.

Fit details for D^+

For the $D^+ \rightarrow K^- \pi^+ \pi^+$ mode, the signal shape in the mass fit is the sum of a normal and a Crystal Ball distribution, sharing a common mode but allowed to have different widths. The mode of the signal mass PDF as well as the width of the Crystal Ball component in the mass PDF are split across (p_T, y) bins, as is the slope of the background mass PDF. The mode of the $\ln \chi_{\text{IP}}^2$ prompt signal and secondary signal distributions are split across (p_T, y) bins. The mode and tail parameters of the signal $\ln \chi_{\text{IP}}^2$ PDF, are floated during the fit to data, with all other shape parameters fixed to the values from MC.

Prefits to the prompt signal and secondary signal data samples in $\ln \chi_{\text{IP}}^2$ are given in Figure 9.4.

Fit details for D_s^+

For the $D_s^+ \rightarrow \phi \pi^+$ mode, the signal shape in the mass fit is the sum of a two normal distributions, sharing a common mean but allowed to have different widths. The mean and width of the signal mass PDF and the slope

of the background mass PDF, along with the mode of the $\ln \chi_{\text{IP}}^2$ prompt signal PDF, are split across (p_{T}, y) bins. The mode of the prompt signal $\ln \chi_{\text{IP}}^2$ PDF and all parameters of the secondary PDF are floated during the fit to data, with all other shape parameters fixed to the values from MC.

Prefits to the prompt signal and secondary signal data samples in $\ln \chi_{\text{IP}}^2$ are given in Figure 9.5.

Fit details for D^{*+}

For the D^{*+} -tagged $D^0 \rightarrow K^- \pi^+$ mode, the width of the widest component of the Δm signal model, described in Section 9.1, is split across (p_{T}, y) bins, along with the mean of the prompt signal $\ln \chi_{\text{IP}}^2$ PDF. The mean of the prompt signal $\ln \chi_{\text{IP}}^2$ PDF and all parameters of the secondary PDF are floated during the fit to data, with all other shape parameters fixed to the values from MC.

Prefits to the prompt signal and secondary signal data samples in D^0 $\ln \chi_{\text{IP}}^2$ are given in Figure 9.6.

Fit results

For all fits, the covariance matrix is checked to be positive definite, and the goodness-of-fit within each (p_{T}, y) bin is checked visually to ensure that the NLL minimisation stopped at a sensible parameter value set. Mass, Δm , and $\ln \chi_{\text{IP}}^2$ fits are given in Figures 9.7 to 9.10, where the data and models shown are the sums of the data and models over all (p_{T}, y) bins. Prompt signal yields per (p_{T}, y) bin are given in Tables A.1 to A.4, and those integrated across all bins in Table 9.1.

The plots show that the models match the data well, but there are some significant discrepancies such as in Figures 9.9b and 9.10b. However, the discrepancies in these integrated plots are exaggerations of those in the individual (p_{T}, y) bin fits: they indicate that there is a consistent bias in the individual bins, but do not necessarily indicate that these are statistically significant. For comparison, Figures 9.11 to 9.14 show the mass (or Δm) and $\ln \chi_{\text{IP}}^2$ fits for each meson in the (p_{T}, y) bin with the highest

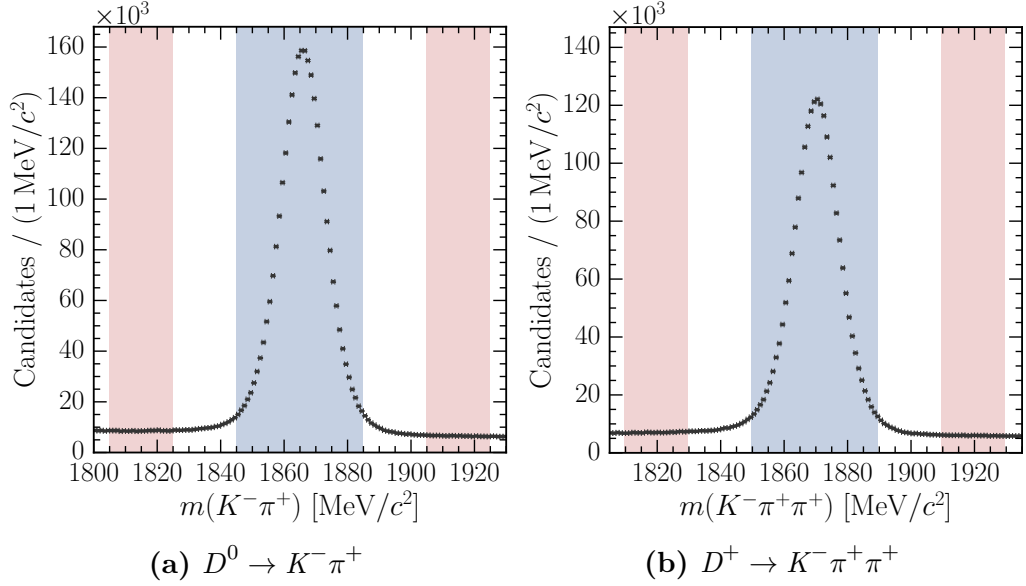


Figure 9.1: Definition of signal window, in blue, and sidebands, in red, for $D^0 \rightarrow K^- \pi^+$ (a) and $D^+ \rightarrow K^- \pi^+ \pi^+$ (b) candidates. The boundaries of the different regions are defined in Section 9.2. The full dataset is shown.

prompt signal yield and in the (p_T, y) bin with the highest combinatorial background yield. These show that, in the individual (p_T, y) bins, the fits model the data significantly better than the integrated plots would suggest, and also that the fit is flexible enough to model both high prompt signal and high combinatorial background datasets. The effect on the cross-section measurements of mis-modelling will be discussed further in Section 11.8.

In some (p_T, y) bins, it is not possible to make a prompt signal yield measurement due to insufficient data in those bins. At this stage in the analysis, a cross-section measurement in a bin is considered unfeasible if there is either no data (zero candidates), or if the prompt signal yield $N_i \pm \sigma_{N_i}$ in that bin as determined in the fit does not satisfy $N_i > 3\sigma_{N_i}$.

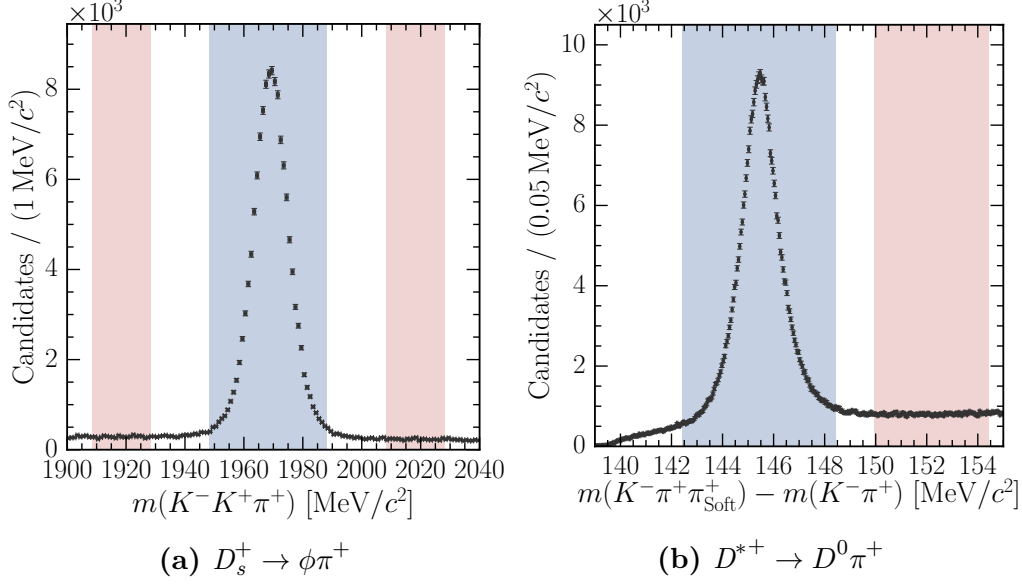


Figure 9.2: Definition of signal window, in blue, and sidebands, in red, for $D_s^+ \rightarrow \phi\pi^+$ (a) and $D^{*+} \rightarrow D^0\pi^+$ (b) candidates. The boundaries of the different regions are defined in Section 9.2. The full dataset is shown.

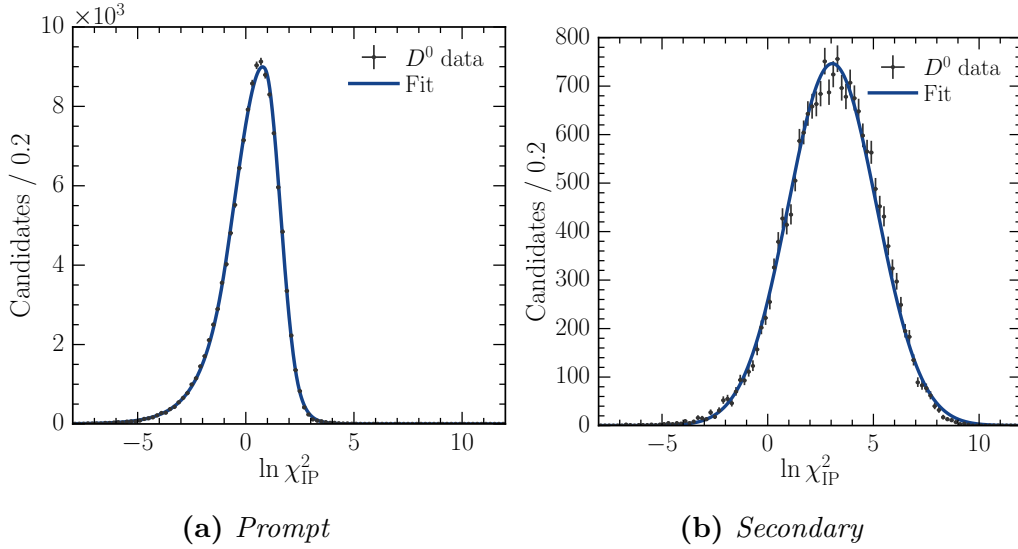


Figure 9.3: Distributions of $\ln \chi^2_{\text{IP}}$ for simulated $D^0 \rightarrow K^-\pi^+$ decays: prompt signal D^0 (a) and secondary signal D^0 (b). The sum of the simultaneous likelihood fits in each (p_{T} , y) bin is overlaid.

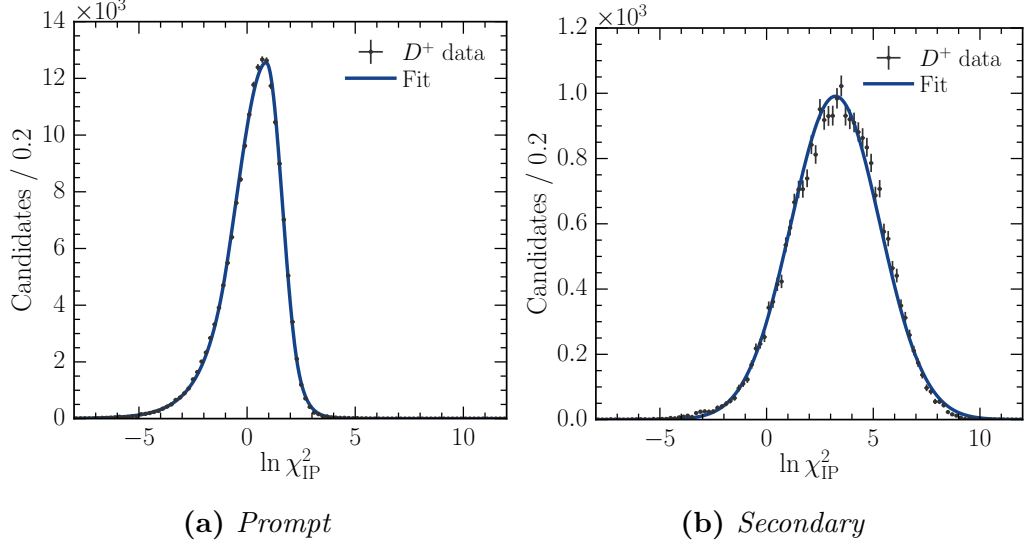


Figure 9.4: Distributions of $\ln \chi_{\text{IP}}^2$ for simulated $D^+ \rightarrow K^- \pi^+ \pi^+$ decays: prompt signal D^+ (a) and secondary signal D^+ (b). The sum of the simultaneous likelihood fits in each (p_{T}, y) bin is overlaid.

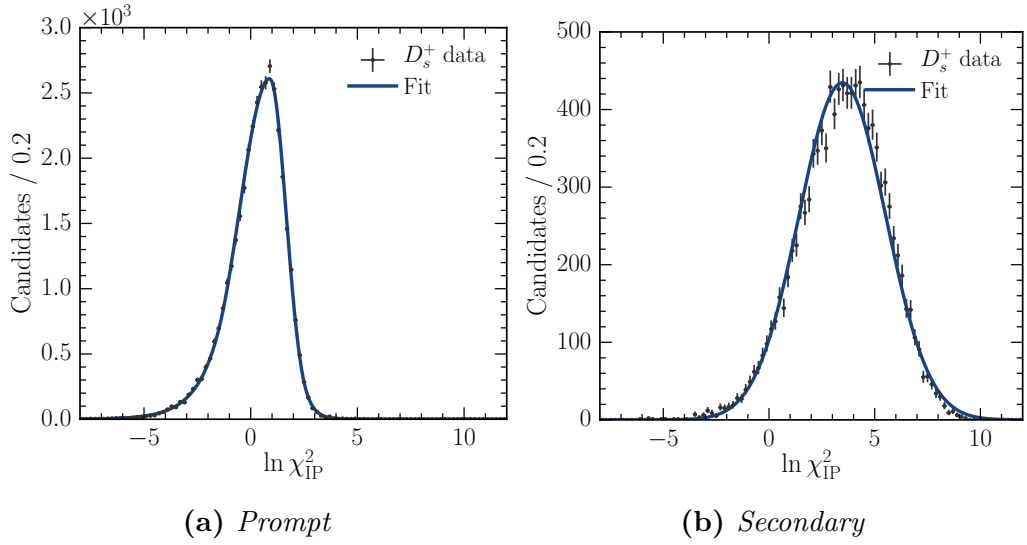


Figure 9.5: Distributions of $\ln \chi_{\text{IP}}^2$ for simulated $D_s^+ \rightarrow \phi \pi^+$ decays: prompt signal D_s^+ (a) and secondary signal D_s^+ (b). The sum of the simultaneous likelihood fits in each (p_{T}, y) bin is overlaid.

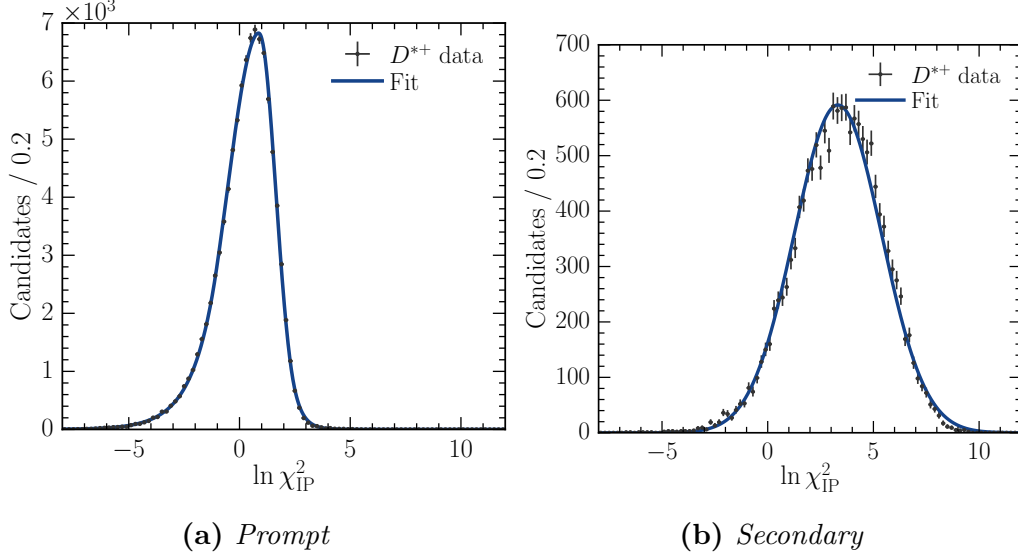


Figure 9.6: Distributions of $D^0 \ln \chi_{\text{IP}}^2$ for simulated $D^{*+} \rightarrow D^0 \pi^+$, with $D^0 \rightarrow K^- \pi^+$, decays: prompt signal D^{*+} (a) and secondary signal D^{*+} (b). The sum of the simultaneous likelihood fits in each (p_{T}, y) bin is overlaid.

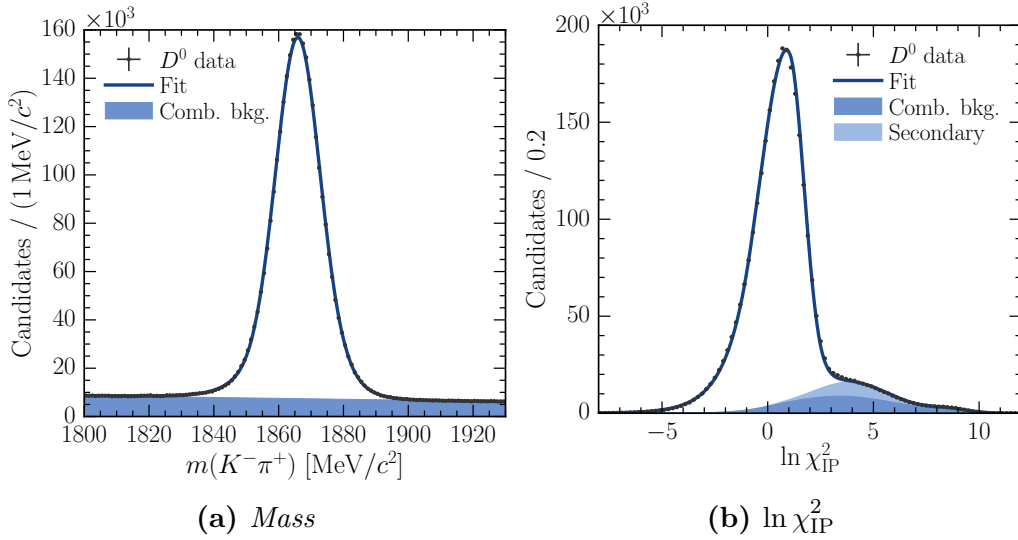


Figure 9.7: Distributions for fully selected $D^0 \rightarrow K^- \pi^+$ candidates: D^0 invariant mass (a) and $D^0 \ln \chi_{\text{IP}}^2$ (b) for a mass window of $\pm 20 \text{ MeV}/c^2$ around the nominal D^0 mass. The sum of the simultaneous likelihood fits in each (p_{T}, y) bin is shown, with components as indicated in the legends.

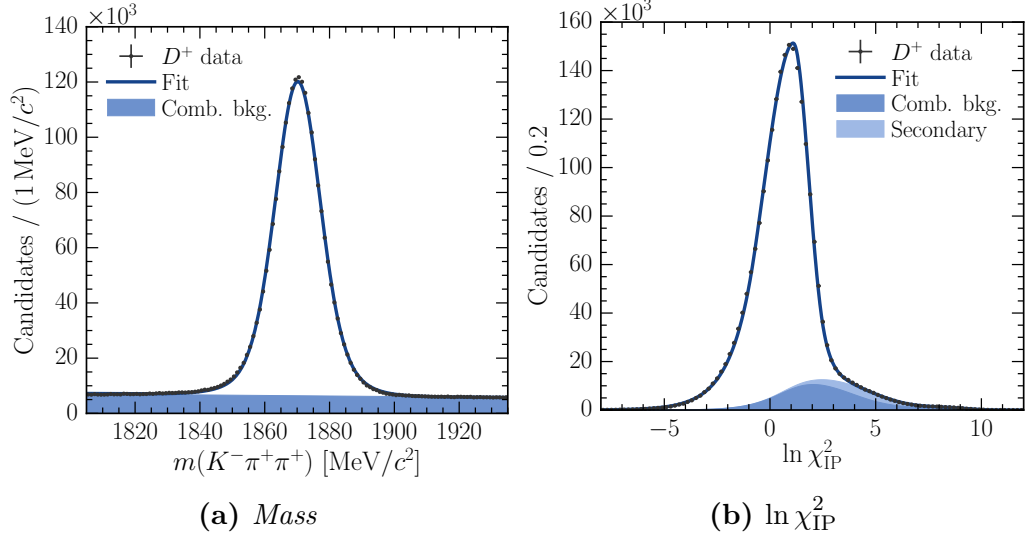


Figure 9.8: Distributions for fully selected $D^+ \rightarrow K^- \pi^+ \pi^+$ candidates: D^+ invariant mass (a) and $D^+ \ln \chi_{\text{IP}}^2$ (b) for a mass window of $\pm 20 \text{ MeV}/c^2$ around the nominal D^+ mass. The sum of the simultaneous likelihood fits in each (p_{T}, y) bin is shown, with components as indicated in the legends.

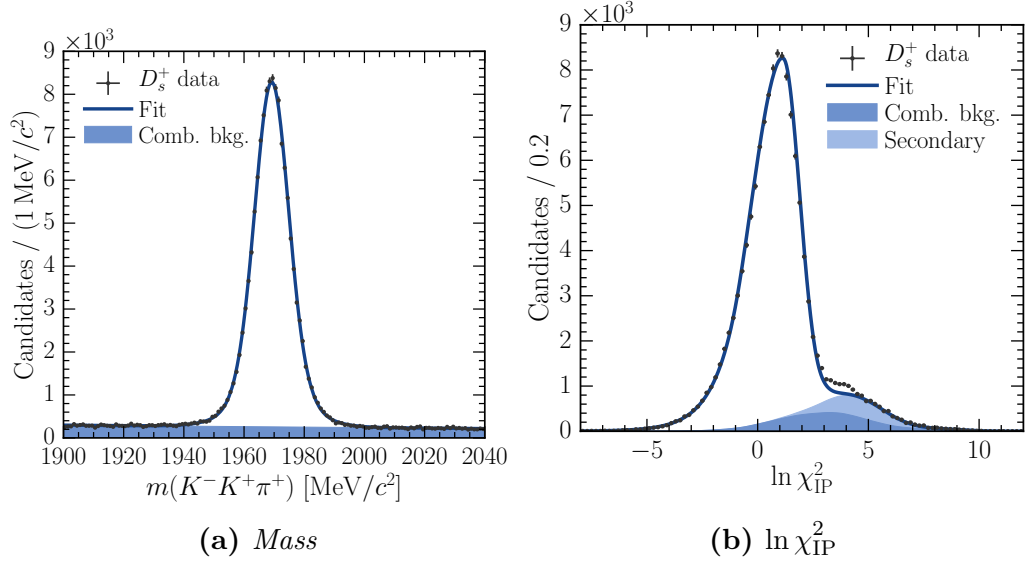


Figure 9.9: Distributions for fully selected $D_s^+ \rightarrow \phi \pi^+$ candidates: D_s^+ invariant mass (a) and $D_s^+ \ln \chi_{\text{IP}}^2$ (b) for a mass window of $\pm 20 \text{ MeV}/c^2$ around the nominal D_s^+ mass. The sum of the simultaneous likelihood fits in each (p_{T}, y) bin is shown, with components as indicated in the legends.

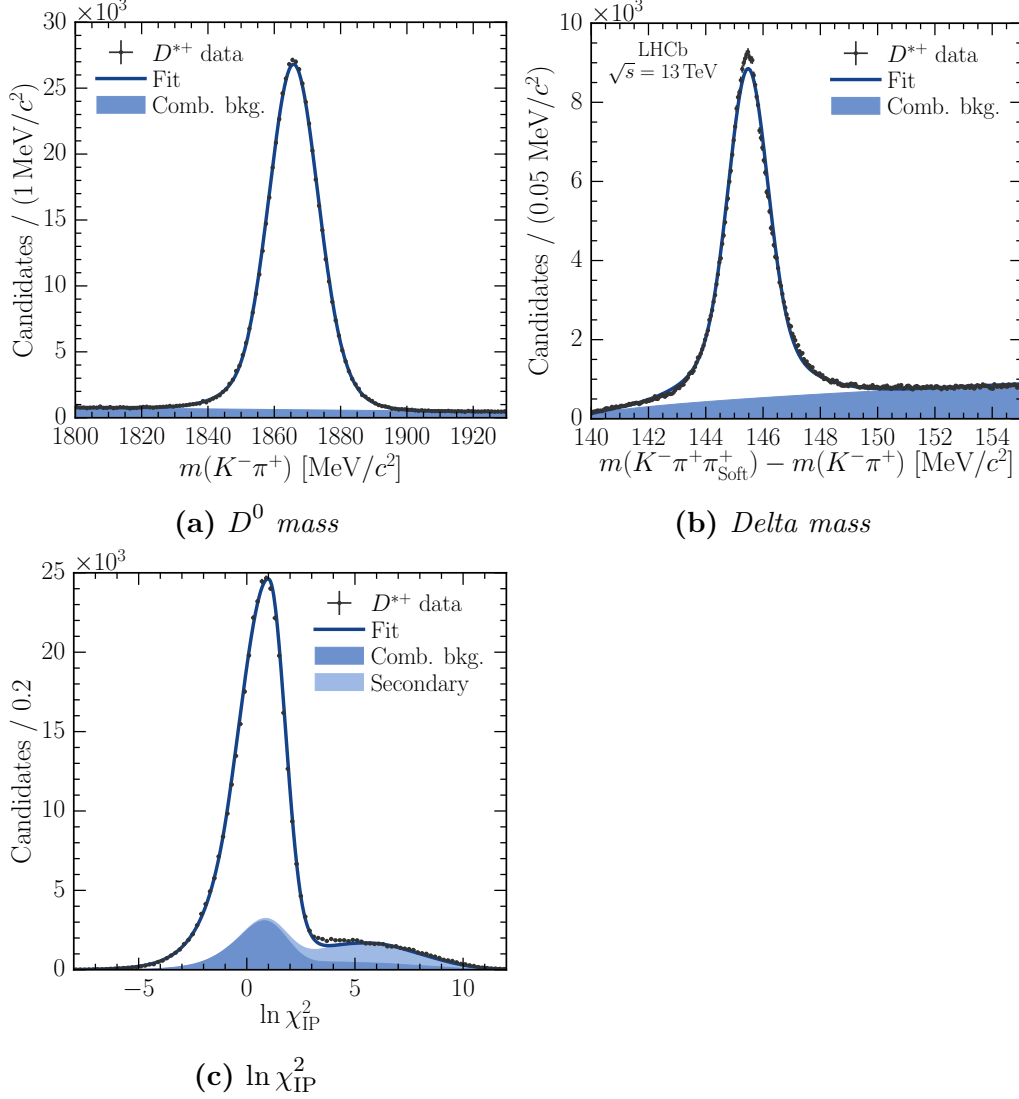


Figure 9.10: Distributions for fully selected D^{*+} candidates, with $D^0 \rightarrow K^-\pi^+$: D^0 invariant mass (a); $\Delta m = m(D^{*+}) - m(D^0)$ (b) for a mass window of ± 20 MeV/ c^2 around the nominal D^0 mass; and $D^0 \ln \chi_{\text{IP}}^2$ (c) with an additional mass window of ± 3 MeV/ c^2 around the nominal Δm value. The sum of the simultaneous likelihood fits in each (p_{T}, y) bin is shown, with components as indicated in the legends.

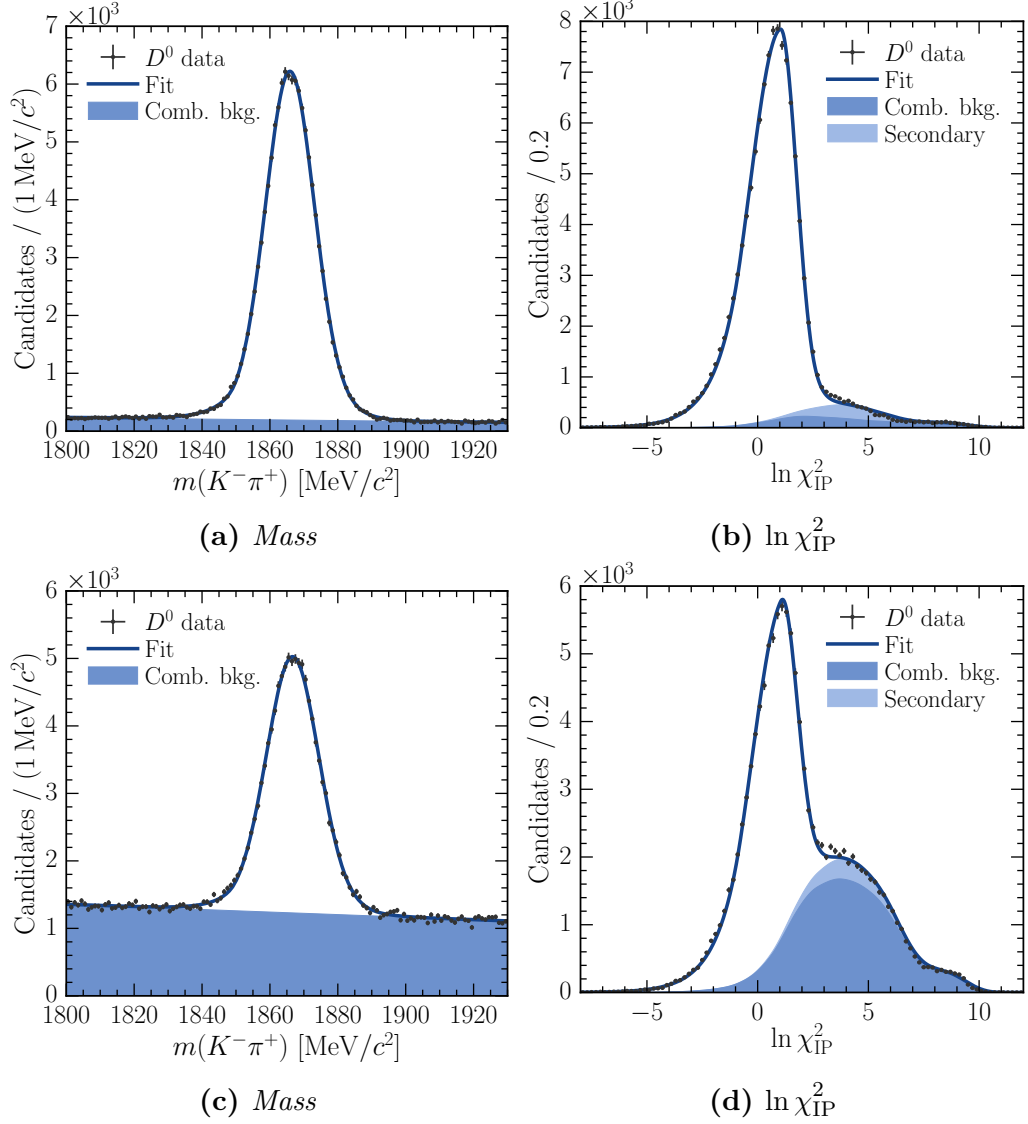


Figure 9.11: Distributions for fully selected $D^0 \rightarrow K^- \pi^+$ candidates: D^0 invariant mass (a and c); and $D^0 \ln \chi_{\text{IP}}^2$ (b and d) for a mass window of $\pm 20 \text{ MeV}/c^2$ around the nominal D^0 mass. The top Figures (a and b) show the data and fits in the region $2 < p_{\text{T}} < 2.5 \text{ GeV}/c$, $3 < y < 3.5$, whilst the bottom Figures (c and d) show the data and fits in the $0 < p_{\text{T}} < 1 \text{ GeV}/c$, $3.5 < y < 4$ region.

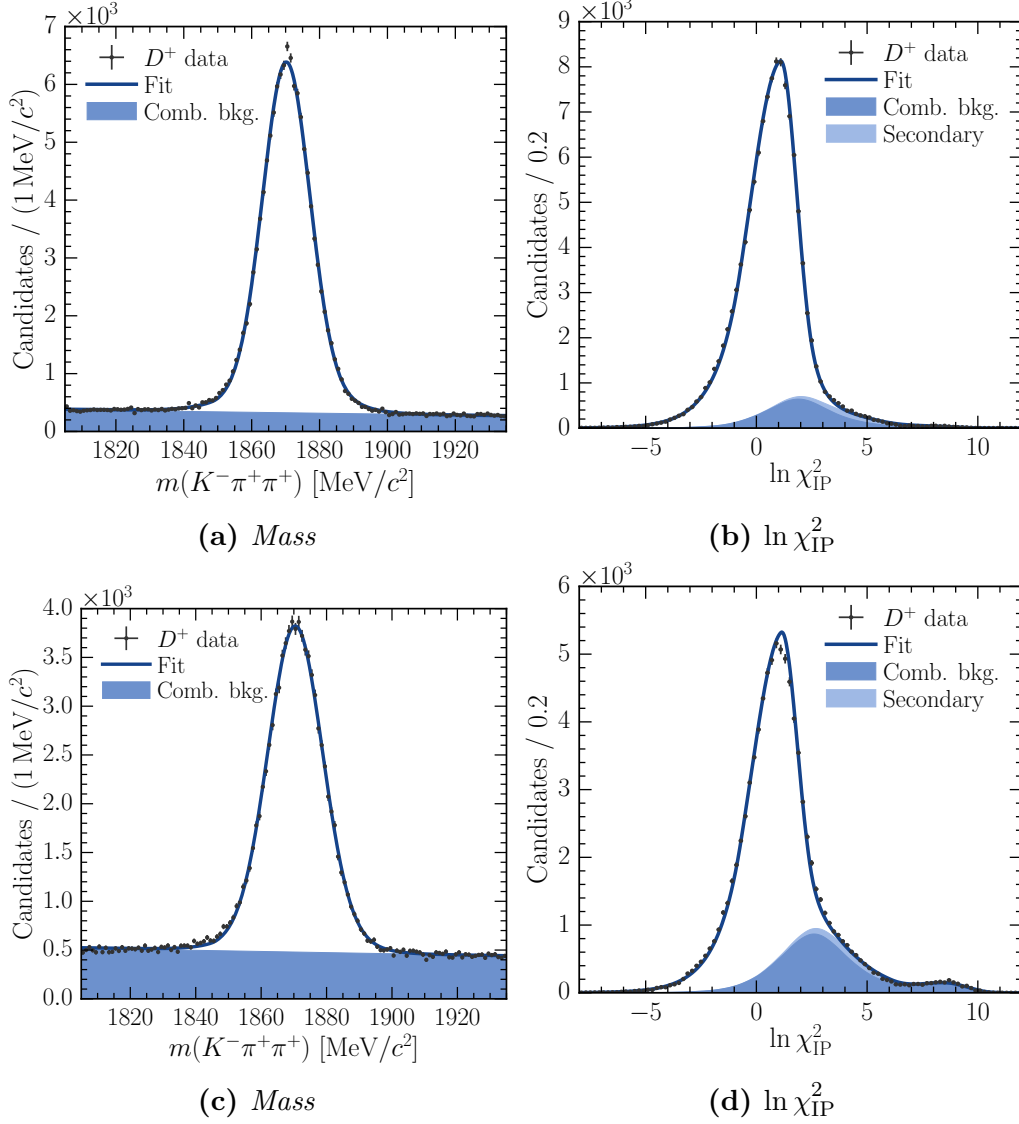


Figure 9.12: Distributions for fully selected $D^+ \rightarrow K^- \pi^+ \pi^+$ candidates: D^+ invariant mass (a and c); and $D^+ \ln \chi_{\text{IP}}^2$ (b and d) for a mass window of $\pm 20 \text{ MeV}/c^2$ around the nominal D^+ mass. The top Figures (a and b) show the data and fits in the region $2.5 < p_{\text{T}} < 3 \text{ GeV}/c$, $3 < y < 3.5$, whilst the bottom Figures (c and d) show the data and fits in the $2 < p_{\text{T}} < 2.5 \text{ GeV}/c$, $3.5 < y < 4$ region.

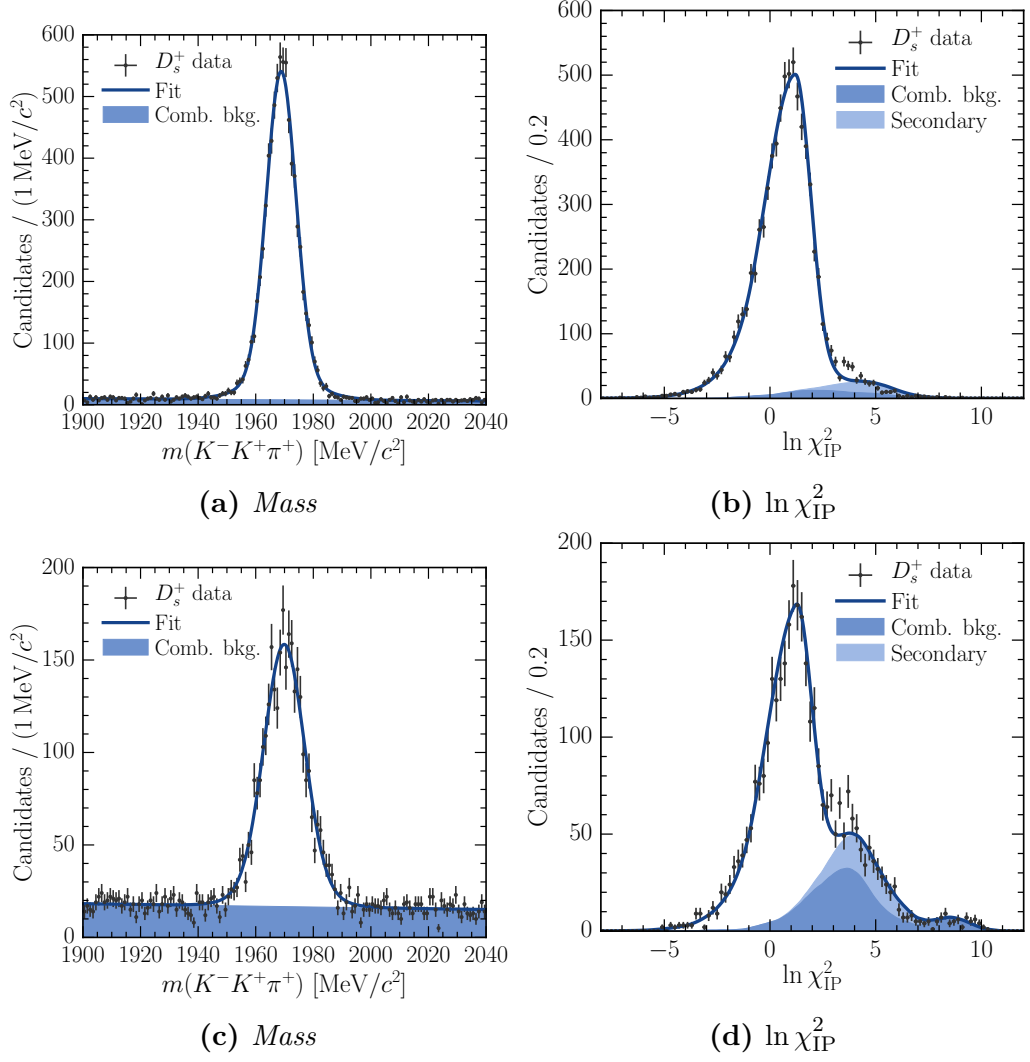


Figure 9.13: Distributions for fully selected $D_s^+ \rightarrow \phi \pi^+$ candidates: D_s^+ invariant mass (a and c); and $D_s^+ \ln \chi_{\text{IP}}^2$ (b and d) for a mass window of $\pm 20 \text{ MeV}/c^2$ around the nominal D_s^+ mass. The top Figures (a and b) show the data and fits in the region $4 < p_{\text{T}} < 5 \text{ GeV}/c$, $2.5 < y < 3$, whilst the bottom Figures (c and d) show the data and fits in the $2 < p_{\text{T}} < 2.5 \text{ GeV}/c$, $3.5 < y < 4$ region.

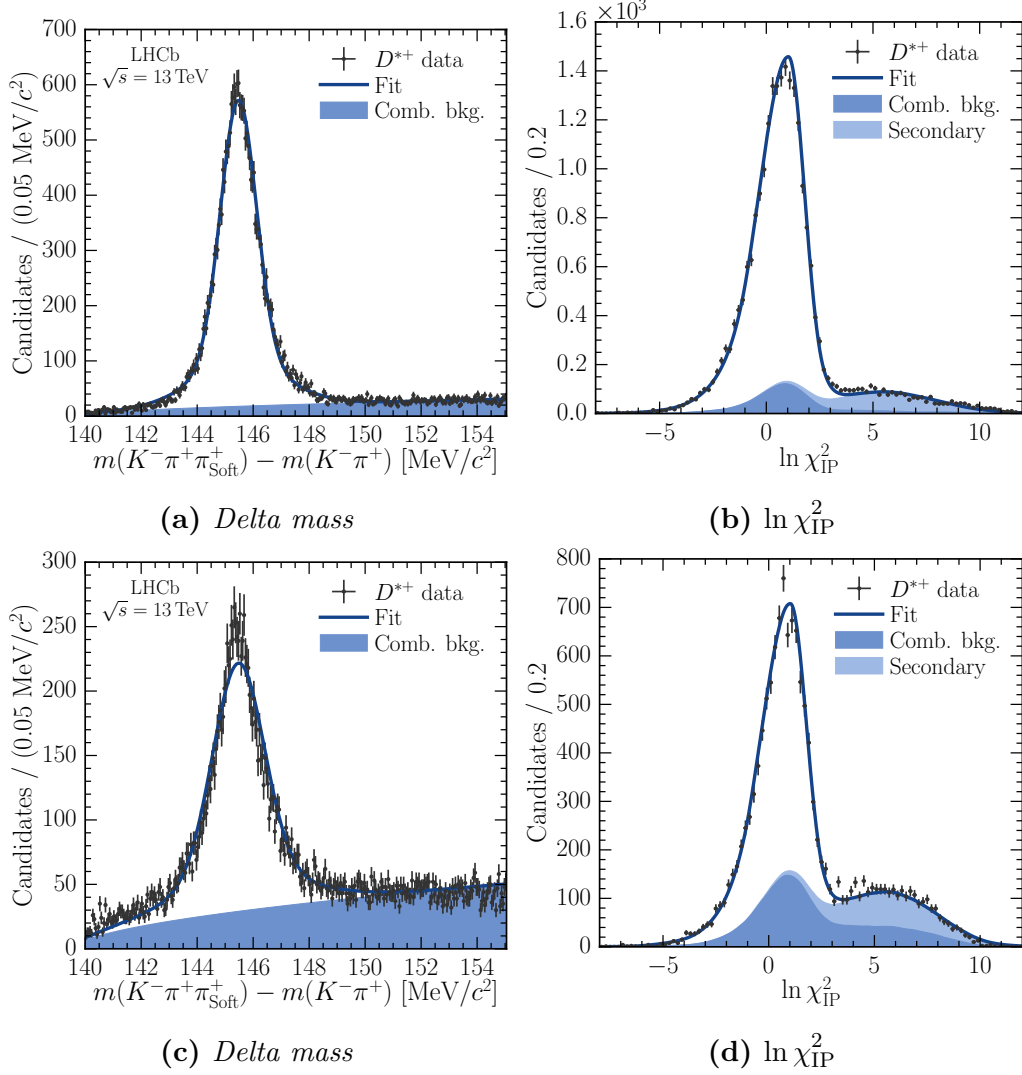


Figure 9.14: Distributions for fully selected $D^{*+} \rightarrow D^0 \pi^+$, with $D^0 \rightarrow K^- \pi^+$, candidates: $\Delta m = m(D^{*+}) - m(D^0)$ (a and c) for a mass window of $\pm 20 \text{ MeV}/c^2$ around the nominal D^0 mass; and $D^0 \ln \chi^2_{\text{IP}}$ (b and d) with an additional mass window of $\pm 3 \text{ MeV}/c^2$ around the nominal D^{*+} - D^0 mass difference. The top Figures (a and b) show the data and fits in the region $4 < p_{\text{T}} < 5 \text{ GeV}/c$, $3 < y < 3.5$, whilst the bottom Figures (c and d) show the data and fits in the $1.5 < p_{\text{T}} < 2 \text{ GeV}/c$, $3.5 < y < 4$ region.

Table 9.1: *Prompt signal yields in the fully selected dataset, summed over all (p_{T}, y) bins in which a measurement is made.*

Hadron	Prompt signal yield
D^0	$(25.77 \pm 0.02) \times 10^5$
D^+	$(19.74 \pm 0.02) \times 10^5$
D_s^+	$(11.32 \pm 0.04) \times 10^4$
D^{*+}	$(30.12 \pm 0.06) \times 10^4$

10 Efficiency evaluation

The efficiency chain defines the fraction of prompt charm mesons that survive the full reconstruction and selection procedure within each (p_T, y) bin. The chain is defined in steps grouped by physical requirements and those imposed by software, with each step being relative to the previous one. This starts with the efficiency of a signal charm hadron to decay within the LHCb acceptance having been produced in a pp collision, $\varepsilon_{\text{Acceptance}}$. The efficiency of such a decay to be fully reconstructed is the reconstruction efficiency $\varepsilon_{\text{Reconstruction}}$, and then the efficiency of the reconstructed decay to be triggered through L0, HLT1, and HLT2 is the trigger efficiency $\varepsilon_{\text{Trigger}}$. There is also the efficiency of triggered candidates to pass through the offline selection $\varepsilon_{\text{Offline}}$, and finally there is the signal window efficiency $\varepsilon_{\text{Sig. win.}}$ due to the requirement imposed before the $\ln \chi^2_{\text{IP}}$ fit. The total efficiency, ε , is the product of these, and is what enters in Equation (5.3).

Two methods are used for efficiency estimation: with simulation, where true information is available before and after all steps; and methods of calibration, where proxy data selected without the use of particular information can be used to assess the effects of using that information in the analysis data. Efficiency evaluation with MC can be simple, but differences between the data and the MC must be accounted for. Calibration techniques use real data, and so can be more robust than MC, but obtaining clean calibration samples can be challenging.

In LHCb, it is known that the PID selection efficiencies are not well-modelled in the MC due to an under-estimation of the detector occupancies. The same is true, albeit to a lesser extent, for the reconstruction efficiencies. Kinematic variables, such as charm hadron and final-state momenta, are

comparatively well-modelled. This motivates a two-step procedure for evaluating the *selection* efficiency, whereby the PID efficiency is computed using calibration techniques, described in Section 10.4, and the efficiency of the remaining requirements is made using MC, detailed in Section 10.3. The reconstruction efficiency will be computed using MC, but corrected with a factor obtained from calibration samples, described in Section 10.1. The extent to which these techniques do not accurately model the efficiencies will be discussed in the context of systematic uncertainties in Section 11.2.

The matching of true MC particles to reconstructed objects, the process of truth-matching as described in Section 7.1, is not 100% efficient, and so a correction factor ρ_{Truth} is computed to account for the deficit.

To summarise: the total efficiency $\varepsilon_i(H_c \rightarrow f)$, for charm candidates in the i th (p_T, y) bin, is the product of the individual efficiencies and correction factors

$$\begin{aligned} \varepsilon_i(H_c \rightarrow f) = & \varepsilon_{\text{Acceptance}} \times \varepsilon_{\text{Reconstruction}} \times \rho_{\text{Truth}} \times \rho_{\text{Tracking}} \\ & \times \varepsilon_{\text{Trigger}} \times \varepsilon_{\text{Offline}} \times \varepsilon_{\text{PID}} \times \varepsilon_{\text{Sig. win.}}. \end{aligned} \quad (10.1)$$

where each efficiency ε is dependent on the charm hadron H_c and the final state f , and is conditional on the previous step.

In this chapter the details of the efficiency computation for each step are given. Example efficiency tables, differential in (p_T, y) , for the two-body $D^0 \rightarrow K^- \pi^+$ and three-body $D^+ \rightarrow K^- \pi^+ \pi^+$ decays are given in Appendix A. The notation for conditional yields will be used throughout, where $N_{B|A}$ denotes the prompt signal yield after process B given that process A preceded B .

10.1 Detector acceptance

The acceptance efficiency, evaluated using simulation, is due to the finite spatial acceptance of the LHCb detector. The acceptance is modelled by a set of cuts requiring that all stable charged particles in the final state have a momentum direction satisfying $10 < \theta < 400 \text{ mrad}$, and that all the particles in the final state point in the positive z direction. These requirements are

applied at the generator level, before the detector simulation, on the true kinematics of the particles.

By counting the number of prompt charm hadrons passing the cut, the acceptance efficiency is defined as

$$\varepsilon_{\text{Acceptance}} = \frac{N_{\text{Accepted}|\text{Generated}}}{N_{\text{Generated}}}. \quad (10.2)$$

The input dataset to this computation is the generator-level set described in Section 7.1. The acceptance efficiencies for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$, in (p_T, y) bins, are given in Tables A.5 and A.6.

10.2 Reconstruction

The reconstruction efficiency $\varepsilon_{\text{Reconstruction}}$ parameterises the fraction of charm meson decays passing the acceptance requirements that are also fully reconstructed as tracks, for the final-state particles, and vertices, for the charm mesons. This folds in several effects: whether the final-state particles have a large enough momentum not to be bent out of the detector acceptance by the magnetic field; whether the final-state particles leave enough hits in the tracking system to be reconstructible (that is, above the minimum threshold at which the tracking system *could* reconstruct a track); whether, given that enough hits were deposited by a final-state particle, a track was actually reconstructed; and whether a vertex can be formed, given that all final-state particles have associated tracks.

The reconstruction efficiency is defined as the ratio of fully reconstructed decays to those passing the acceptance requirements

$$\varepsilon_{\text{Reconstruction}} = \frac{N_{\text{Reconstructed}|\text{Accepted}}}{N_{\text{Accepted}}}. \quad (10.3)$$

This definition is subject to the effects of (p_T, y) bin ‘migration’, whereby mesons are reconstructed as belonging to bins different from those of the true particles. These can be due to resolution as well as any possible biases on the measurements of p_T and y . If the net migration between bins is large, that is if the migrations are not symmetric, then the detector effects should

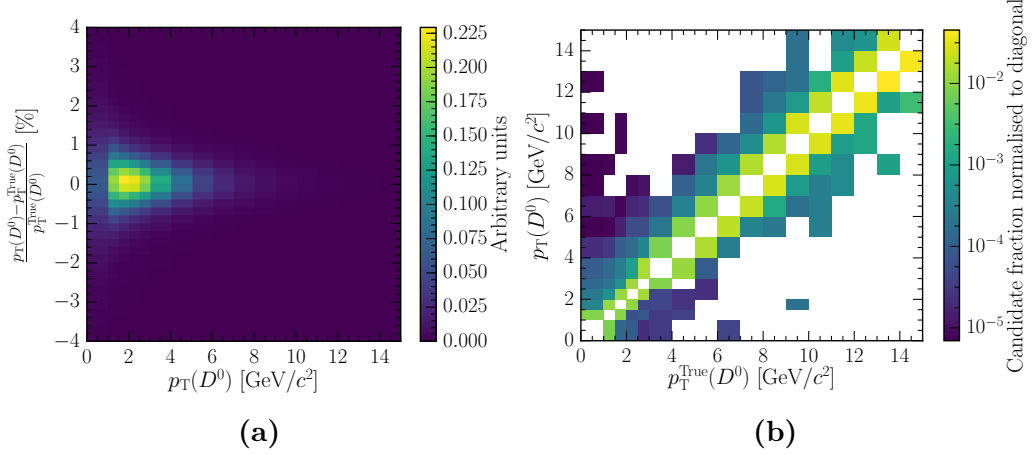


Figure 10.1: *The difference between true and reconstructed D^0 p_T normalised by true D^0 p_T , in bins of reconstructed p_T (a), and a migration matrix (b) showing the fraction of candidates that migrate within a true p_T bin, relative to the diagonal (where there is no migration).*

be corrected for such that the reconstructed meson kinematics are more accurate estimators of their true values. (This process is often referred to as ‘unfolding’ [74].)

Figures 10.1 and 10.2 show the deviations of the reconstructed p_T and y , in bins of those variables, for the $D^0 \rightarrow K^- \pi^+$ MC sample. In both p_T and y the deviation of the reconstructed values from the true values is largely symmetric across the full kinematic range, and the magnitude of the relative deviations is generally below 1 % in p_T and 0.1 % in y . Similarly, the migration matrices show that the fraction of candidates that migrate out of their true (p_T, y) bin is generally well below 1 % in p_T , and below 0.1 % in y , both of which are considered as small migrations. Such small migrations are already accounted for by modelling the reconstruction efficiencies in bins much wider than the resolution.

The reconstruction efficiencies for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$, in (p_T, y) bins, are given in Tables A.7 and A.8.

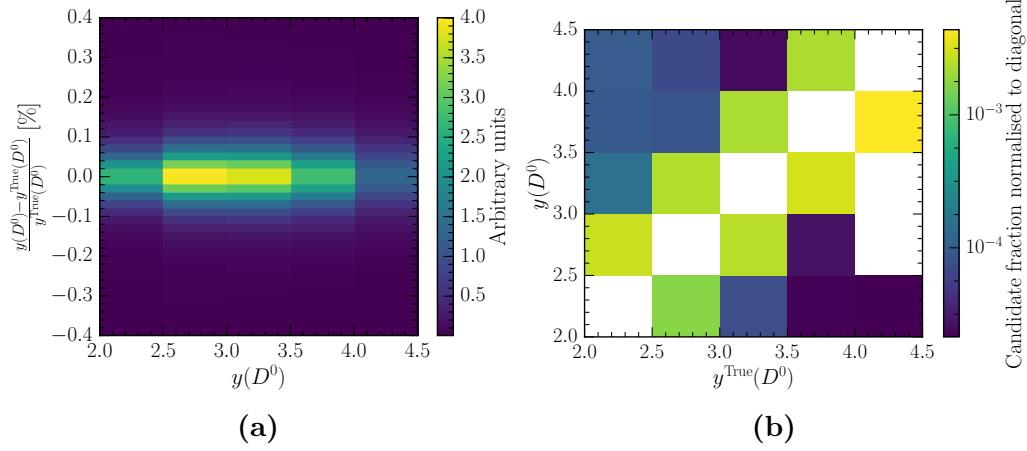


Figure 10.2: The difference between true and reconstructed D^0 rapidity, normalised by true D^0 rapidity, in bins of reconstructed rapidity (a), and a migration matrix (b) showing the fraction of candidates that migrate within a true rapidity bin, relative to the diagonal (where there is no migration).

Truth matching inefficiency

The evaluation of Equation (10.3) requires that the reconstructed objects, tracks and vertices, be correctly associated back to the ‘truth-level’ information, that is the MC objects which were generated and then propagated through the detector simulation. The truth-matching procedure is described in Section 4.4. The background category information is used to filter the MC such that the number of remaining candidates are simply counted to obtain the input to some efficiency computation, such as Equation (10.3). However, the requirement that at least 70 % of a track’s hits are matched to hits created by a single MC particle is not 100 % efficient at accepting signal. The magnitude of the size of the inefficiency can be determined by counting the number of signal decays that fail the truth matching requirement. Figure 10.3 shows the D^0 and D^+ mass distributions for the various background categories. The ‘ghost’ category exhibits a peak identical in nature to that seen in the ‘signal’ category. As not all signal decays that are truly reconstructed will be counted as such when only the ‘signal’ category is considered, associated efficiencies are underestimated.

Given that the number of correctly matched signal decays is N , and the

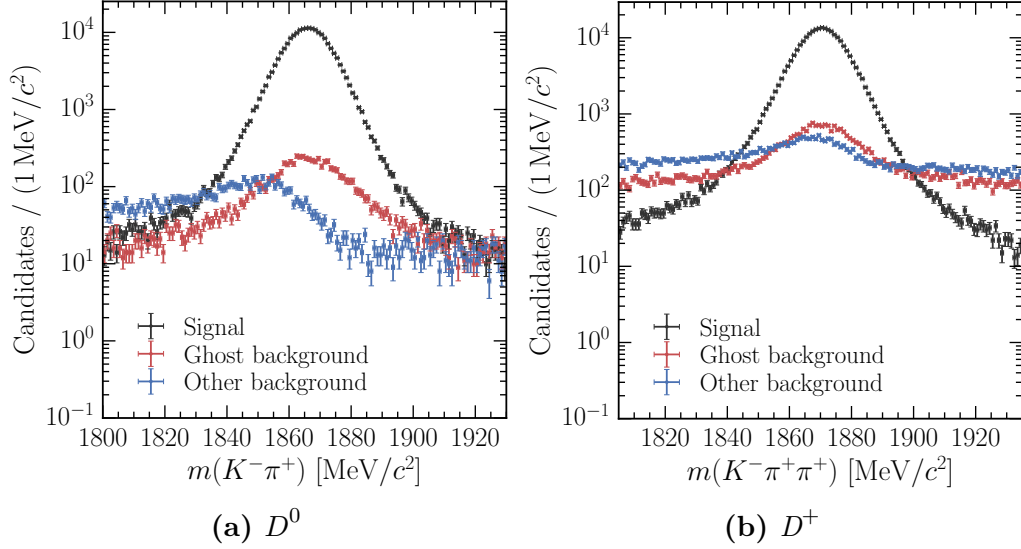


Figure 10.3: Mass distributions of MC D^0 (a) and D^+ (b) candidates for various background categories, as indicated in the legends. The most prominent features are the signal-like peak in the ‘ghost’ distribution, indicative of a truth-matching inefficiency, and the long tail extending below the mass peak in the ‘other’ distribution, which mostly consists of partially reconstructed physics backgrounds such as $D^0 \rightarrow K^- \pi^+ \pi^0$ and $D^+ \rightarrow K^- \pi^+ \pi^+ \pi^0$.

number of incorrectly labelled signal decays is U , the selection efficiency defined in Equation (10.3) is incorrect by a factor

$$\rho_{\text{Truth}} = \frac{U + N}{N}. \quad (10.4)$$

The number of decays passing the signal requirement is used for N , and U is measured with a maximum likelihood fit to the mass or Δm distribution of the decays not passing the requirement. As shown in Figure 10.3, the fraction of incorrectly identified signal decays is small, and so a single fit is performed with the data integrated across all (p_T, y) bins, and a single correction factor is measured. The details of the fit are identical to those for the mass fits used in the charm yield extraction with real data, described in Section 9.1. The result of the fits to the $D^0 \rightarrow K^- \pi^+$ and the $D^+ \rightarrow K^- \pi^+ \pi^+$ simulation are shown in Figure 10.4. Table 10.1 lists the correction factors obtained.

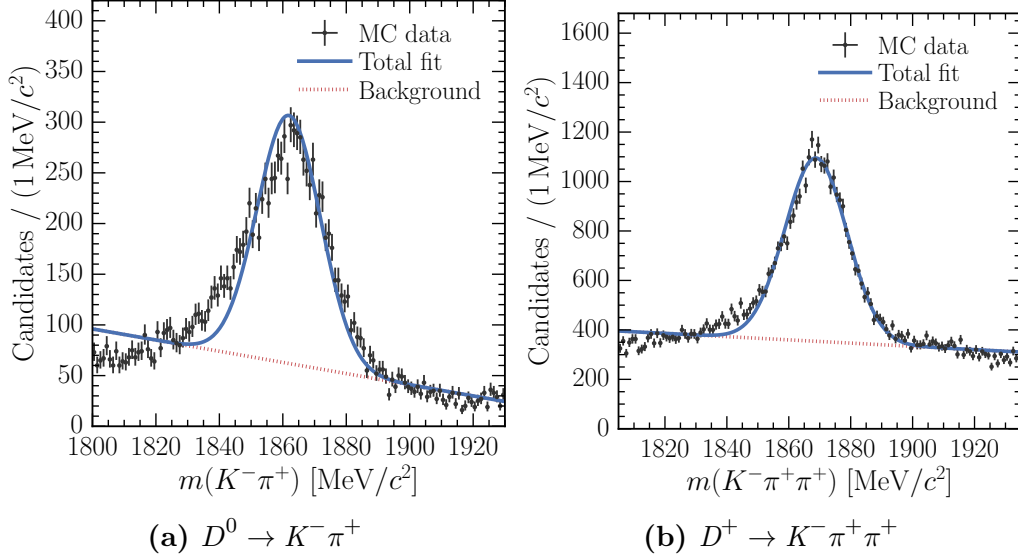


Figure 10.4: Mass distributions of MC D^0 (a) and D^+ (b) candidates not passing the ‘signal’ truth-matching requirement, that is the sum of the “Ghost” and “Other” categories shown in Figure 10.3. Overlaid on each distribution is the result of a maximum likelihood fit, with the total model shown as a blue curve, and the background model as a red dotted line.

Table 10.1: Efficiency $1/\rho_{\text{Truth}}$ of the truth matching requirement applied in the simulation.

Mode	$1/\rho_{\text{Truth}}$ (%)
$D^0 \rightarrow K^- \pi^+$	98.68 ± 0.06
$D^+ \rightarrow K^- \pi^+ \pi^+$	96.53 ± 0.58
$D_s^+ \rightarrow \phi \pi^+$	98.53 ± 0.13
$D^{*+} \rightarrow D^0 \pi^+, D^0 \rightarrow K^- \pi^+$	96.85 ± 0.09

Tracking efficiency correction

The tracking efficiency measures the efficiency for a charged particle to be reconstructed as a track, and is part of the definition of the reconstruction efficiency in Equation (10.3). A data-driven technique is used to compute correction factors based on the observed differences between the data and the MC.

Only long tracks are used in this analysis, which requires that the particles create hits in the VELO and in the T stations. Hits in the TT are not required to form long tracks, although matching hits are added to tracks to improve their momentum resolution. To measure the efficiency of long track creation, $J/\psi \rightarrow \mu^+ \mu^-$ decays are reconstructed using a ‘tag-and-probe’ technique, where one particle is fully reconstructed as a long track and the other track is created using track segments made only using information from the muon stations. The momentum of the latter track, the ‘probe’, can be inferred by treating the magnetic field as an optical system that imparts a momentum ‘kick’ in the x direction. By assuming the muon track segment was made by a particle originating from the PV, the probe segment can be extrapolated back from the muon stations to the yz plane where the kick is modelled, and then the difference in slopes between the PV-plane line and the probe track gives the probe momentum. The invariant mass of the dimuon system can then be computed and the distribution can be used to distinguish between backgrounds and true J/ψ decays. The mass resolution of $J/\psi \rightarrow \mu^+ \mu^-$ decays reconstructed using long+muon tracks is about $200 \text{ MeV}/c^2$, to be compared with a resolution of about $16 \text{ MeV}/c^2$ using long+long tracks. To improve the mass resolution, the muon tracks are combined with hits in the TT, improving the resolution to around $57 \text{ MeV}/c^2$ [36, 75].

The tracking efficiency is evaluated as the ratio of J/ψ mesons reconstructed using the long+muon+TT method, with those that can *also* be reconstructed using the long+long method

$$\varepsilon_{\text{Tracking}} = \frac{N_{\text{long+long|long+muon+TT}}}{N_{\text{long+muon+TT}}}. \quad (10.5)$$

A long track in an event is considered as ‘matched’ to the probe track if there is a 70 % overlap of their hits in the muon stations. The tag-and-probe method is performed using real data and MC, and a ratio is formed

$$\rho_{\text{Tracking}} = \frac{\varepsilon_{\text{Tracking}}^{\text{Data}}}{\varepsilon_{\text{Tracking}}^{\text{MC}}}. \quad (10.6)$$

This ratio is evaluated in bins of probe track momentum and pseudorapidity, as the tracking efficiency is found to vary as a function of these variables, and is shown in Figure 10.5. As it is found to also vary as a function of the total number of reconstructed long tracks in the event, the input tracks in the MC are weighted such that the event multiplicity distributions match those in the data.

The correction factor in Equation (10.6) can be applied to the tracking efficiency found using any decay, as it is only a function of p and η . It is then applied to the reconstruction efficiency in Equation (10.3), where for a single decay the total correction factor is the product of the correction factors for all the tracks comprising the final states. Tracking efficiency correction factors for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$ are given in Tables A.9 and A.10.

The correction factors given do not account for the differences between muon and hadron interactions with the detector. These differences will result in different tracking efficiencies, which may not be the same between data and MC. The effect of these differences will be discussed in the context of systematic uncertainties in Section 11.6.

Combining the truth-matching efficiency and tracking efficiency correction factors, the ‘corrected reconstruction’ efficiency $\varepsilon'_{\text{Reconstruction}}$ can be considered to be

$$\varepsilon'_{\text{Reconstruction}} = \rho_{\text{Truth}} \times \rho_{\text{Tracking}} \times \varepsilon_{\text{Reconstruction}}. \quad (10.7)$$

10.3 Selection

At this stage, the efficiency chain includes the effects of the detector acceptance, the reconstruction, the truth-matching used to evaluate the reconstruction efficiency, and the data-MC differences in the tracking efficiency. What

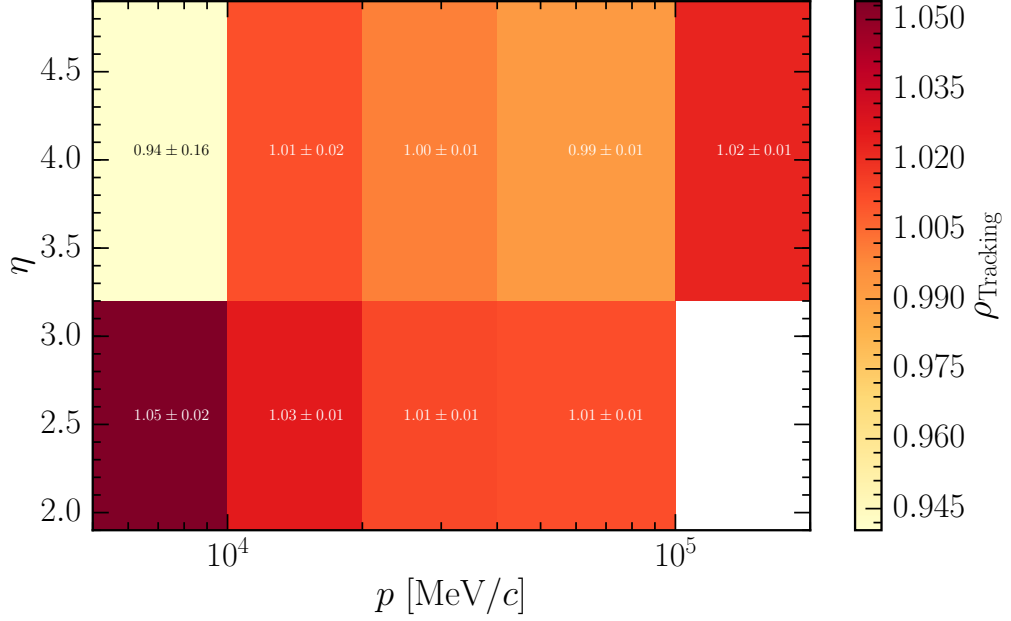


Figure 10.5: Tracking efficiency correction factors ρ_{Tracking} in bins of track momentum and pseudorapidity. The value is labelled in each bin, along with the statistical uncertainty on those values due to the finite size of the calibration datasets. Bins in which a measurement could not be made are empty.

it left is the selection efficiencies due to the online and offline selections, as described in Chapter 8. For simplicity, the ‘online’ and ‘offline’ selection are treated as a single requirement, with the caveat that PID requirements are not included. The PID efficiency can be computed using data-driven methods, and will be described in Section 10.4.

The L0 trigger selects bunch-bunch crossings randomly at a predetermined rate, and so the L0 efficiency is just the ratio of accepted crossings to the total number, or equivalently

$$\varepsilon_{\text{L0}} = \frac{R_{\text{NOBIAS}}}{N_{\text{Coll. bunches}} \nu_{\text{rev}}}, \quad (10.8)$$

where R_{NOBIAS} is the rate at which the L0 trigger was programmed to accept crossings; $N_{\text{Coll. bunches}}$ is the number of bunches in the LHC which collided at the LHCb interaction point; and ν_{rev} is the bunch revolution frequency of 11.246 kHz [26]. Both R_{NOBIAS} and $N_{\text{Coll. bunches}}$ can vary as a function of the

LHC fill, and hence so can $\varepsilon_{\text{NOBIAS}}$. The corresponding values are tabulated in Table 8.1.

The combined HLT1, HLT2, and offline selection efficiency is computed as the ratio of the number of truth-matched decays passing the full selection (but without PID requirements) to the number of truth-matched, reconstructed decays, giving

$$\varepsilon_{\text{Trigger}} \times \varepsilon_{\text{Offline}} = \varepsilon_{\text{Selection}} = \varepsilon_{\text{L0}} \times \frac{N_{\text{Selected|Reconstructed}}}{N_{\text{Reconstructed}}}. \quad (10.9)$$

Selection efficiencies for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$ are given in Tables A.11 and A.12.

10.4 Particle identification

The PID efficiency ε_{PID} is the efficiency of a decay to pass all requirements on the variables that are discriminatory between particle species. The variables are those in the delta log-likelihood (DLL) family, whose computation is described in Section 4.1. This analysis uses the $\text{DLL}_{K-\pi}$ variable to discriminate between kaon and pion tracks, with the exception of the soft pion in the $D^{*+} \rightarrow D^0 \pi^+$ decay on which no PID requirement is made.

It is assumed that the efficiency of a PID cut for a given particle type can be parameterised as a function of three variables: the track momentum p , the track pseudorapidity η , and a measure of the detector occupancy. These particular variables are chosen as the performance of the RICH detectors is known to depend them [76]. As a measure of the detector occupancy, the number N_{SPD} of hits in the SPD detector is used.

A clean sample of $D^0 \rightarrow K^- \pi^+$ decays can be obtained without the use of PID information. The $D^0 \rightarrow K^- \pi^+$ decay is reconstructed, and D^0 mesons are then combined with the remaining tracks in the event to form candidate $D^{*+} \rightarrow D^0 \pi^+$ decays. Kaon (pion) candidates are those with the opposite (same) electric charge as the pion from the D^{*+} . Mis-identifications of other true D^0 decays, such as $D^0 \rightarrow \pi^+ \pi^-$ and $D^0 \rightarrow K^+ K^-$, can be resolved in the D^0 invariant mass spectrum due to the high momentum resolution

of the detector, and true D^{*+} decays can be separated from combinatorial background in the Δm distribution.

The sample of D^{*+} -tagged D^0 decays is referred to here as the calibration sample, and has a 20% overlap with the sample used for the cross-section measurement. This is partly due to the different selection strategy that is necessary when PID information is ignored, and also due to inclusive trigger requirements imposed on the calibration sample. To compute the efficiency of a given PID requirement on either a kaon or a pion track, kaons or pions from the calibration sample are partitioned in p , η , and N_{SPD} as a three-dimensional histogram. The number C_i of signal tracks in the i th bin is computed, and then the number C'_i of signal tracks after the PID requirement is computed, such that the efficiency in the i th bin is

$$\varepsilon_i = \frac{C'_i}{C_i}. \quad (10.10)$$

It is assumed that the $(p, \eta, N_{\text{SPD}})$ partitions are sufficiently small that the PID efficiency within a bin is single-valued. The three-dimensional histogram of efficiencies is an estimate of the efficiency model for the given PID requirement on the given particle.

The track counts C_i and C'_i are obtained with a single two-dimensional fit to the D^0 mass and the Δm distributions. Weights are computed that can be used to statistically remove the background contribution in variable distributions that are uncorrelated with the invariant mass [77]. As p , η , and N_{SPD} are found to be uncorrelated with the fit variables, these weights are used to compute C_i by summing the weights of the tracks that fall in the i th bin. The weights are also used to compute C'_i in the same way, as the DLL variables are found to be uncorrelated with the fit variables. Once the efficiency in each $(p, \eta, N_{\text{SPD}})$ bin is known, the PID efficiency for the same cut on *any* track can be evaluated by looking up which bin the track falls into, and assigning the corresponding efficiency.

For the calibration sample, the average efficiency across the paramet-

erisation space can be expressed as

$$\varepsilon_{\text{PID}} = \frac{\sum_{i \in \text{bins}} C_i \varepsilon_i}{\sum_{i \in \text{bins}} C_i}. \quad (10.11)$$

For some other sample of tracks, referred to here as the reference sample,¹ the contents of each bin can be scaled by a weight w_i based on the fraction of the total reference sample size in the i th bin relative to that of the calibration sample

$$w_i = \frac{R_i}{C_i} \frac{C}{R}, \quad (10.12)$$

where C_i (R_i) is the number of signal calibration (reference) tracks in the i th bin, as before, and C (R) is the total number of signal calibration (reference) tracks in the sample. The average PID efficiency for the weighted reference sample is then

$$\varepsilon_{\text{PID}} = \frac{\sum_{i \in \text{bins}} w_i C_i \varepsilon_i}{\sum_{i \in \text{bins}} w_i C_i}. \quad (10.13)$$

For this cross-section analysis, all the final states have multiple charged tracks, of different species, upon which PID requirements are made. Equation (10.13) gives the average PID efficiency for a final state where a PID requirement is made only on one track, and so the formalism must be extended to the multi-track case. The average efficiency is first modified by substituting Equation (10.12) into Equation (10.13) to give

$$\varepsilon_{\text{PID}} = \frac{\sum_{i \in \text{bins}} R_i \varepsilon_i}{\sum_{i \in \text{bins}} R_i} = \frac{1}{R} \sum_{i \in \text{bins}} R_i \varepsilon_i, \quad (10.14)$$

and then expanding in terms of a_{ij} , the *weight* of the j th reference track in the i th (p, η, N_{SPD}) bin

$$\varepsilon_{\text{PID}} = \frac{1}{R} \sum_{i \in \text{bins}} \sum_j a_{ij} \varepsilon_i. \quad (10.15)$$

¹It is the “reference” sample because it does not have to be equal to the sample of tracks in data one wishes to know the efficiencies for. It can be a sample of simulated data that describes the kinematics of the decay well, or even a different decay mode with similar properties (often called a control channel or a reference channel).

The weight term can account for the possibility of an impure reference sample, where, for example, it can be the same type of background-subtracting weight as used for the creation the efficiency histogram from the calibration sample. With this, the PID efficiency calculation can be thought of as either of two equivalent methods: weighting the calibration sample to look like the reference sample, as in Equation (10.13), or assigning per-track efficiencies to the reference sample based on $(p, \eta, N_{\text{SPD}})$ bins they fall into, as in Equation (10.15). The latter can be generalised to multi-track efficiencies by summing over all *decays* in the reference sample, replacing the per-track efficiency by a per-decay efficiency ε_d

$$\varepsilon_{\text{PID}} = \frac{1}{R} \sum_{d \in \text{decays}} a_d \varepsilon_d, \quad (10.16)$$

where a_d is the weight of the decay, and ε_d is the PID efficiency for that decay. The per-decay efficiency ε_d for the case in which a PID requirement is made on only a single track is ε_t , whilst for a multi-track decay it is

$$\varepsilon_d = \prod_{t \in \text{tracks}} \varepsilon_t. \quad (10.17)$$

In this analysis, the reference sample used for the calibration is the fully selected data, which includes PID cuts applied both in the trigger and offline. The use of a reference sample that includes PID cuts is a valid approach under the assumption that no efficiency bins have a reference sample size of zero [3]. The benefit is that one does not rely on a proxy or MC dataset correctly modelling the track kinematics, and the inter-track correlations.

The statistical component of the uncertainty on the PID efficiency due to the finite size of the calibration sample is computed via a series of pseudo-experiments, which will be described in Section 11.4. As the reference sample is also of a finite size, there is an additional, associated uncertainty on the per-decay PID efficiency. However, the reference sample is identical to the sample entering the fits described in Chapter 9, and so this uncertainty is already accounted for in the uncertainty on the fitted signal yields.

Optimisation of $(p, \eta, N_{\text{SPD}})$ binning scheme

An assumption in the method of PID calibration previously described is that the PID efficiency within a $(p, \eta, N_{\text{SPD}})$ bin is single-valued. This can be guaranteed by employing an infinitely fine binning. In reality, this cannot be used due to the finite size of the calibration sample.

Figure 10.6 illustrates the dependence of the efficiency of the kaon PID requirement on kaon momentum, as computed in equal-probability bins based on the calibration sample populations. The efficiency varies strongly with track momentum, and the equal-probability binning can lead to neighbouring bins where the efficiency is near-constant, as in the 14–25 GeV/ c region where a single bin may be a sufficient model. In addition, the relative softness of the reference sample distribution (black), compared to that of the calibration sample (red), leads to a large fraction of the reference data being covered by a single bin ($p < 14$ GeV/ c). If the efficiency varies strongly within the first bin, the average value of the PID efficiency in that bin is different for the reference and calibration samples, and so estimating the reference efficiency with that obtained from the calibration samples will give a biased result.

In an attempt to minimise potential biases in the evaluation of the PID efficiency on the reference sample, for each one-dimensional projection of the $(p, \eta, N_{\text{SPD}})$ distribution the following iterative procedure is used, assuming each projection is initially partitioned as a single bin:

1. Divide each bin into ten equal-probability sub-bins.
2. For each bin:
 - a) Compute the PID efficiency using the calibration sample in the bin ε and in the ten sub-bins ε_i , and evaluate a χ^2 statistic as

$$\chi^2 = \sum_i^{10} \frac{(\varepsilon - \varepsilon_i)^2}{\delta \varepsilon_i^2}, \quad (10.18)$$
 - b) Compute the p -value of the χ^2 value given 9 degrees of freedom. If $p < 0.001$, the null hypothesis that the PID efficiency within

the bin is constant is rejected, and so the bin should be divided in two. Otherwise, the bin is not split.

3. For each bin that should be divided in two:
 - a) Divide the bin into twenty equal probability sub-bins.
 - b) For each of the 19 new boundaries, compute the χ^2 statistic in the union of the sub-bins below the boundary and in the union of the sub-bins above the boundary using Equation (10.18), and record the sum of the two χ^2 values. The χ^2 computation made under the hypothesis that the efficiency within each union is constant.
 - c) Record the boundary with the smallest χ^2 sum.
 - d) If the two sub-bin unions either side of the recorded boundary do not each contain at least 2000 tracks, the original bin is not split. Otherwise, the original bin is split in two at the recorded boundary.
4. If no additional bins were created in this iteration, stop. Otherwise, repeat from Step 1

The iteration stops when either all bins satisfy the p -value requirement or no further divisions are possible due to the lower sample size limit. The sample size limit of 2000 tracks was chosen to reduce the possibility that the three one-dimensional binning schemes could be combined to form a three-dimensional binning which contained empty bins. Figures 10.7 and 10.8 shows the resulting $(p, \eta, N_{\text{SPD}})$ binning (dark red) for kaon and pion tracks. It can be seen that there are still bins where the finer binning (blue) shows a large variation in the efficiency. The effect of this on the cross-section measurement is considered as a systematic uncertainty in Section 11.4.

The particle identification efficiencies computed using the calibration procedure are given in Tables A.13 and A.14 for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$. The uncertainties given are those due to the finite size of the calibration sample.

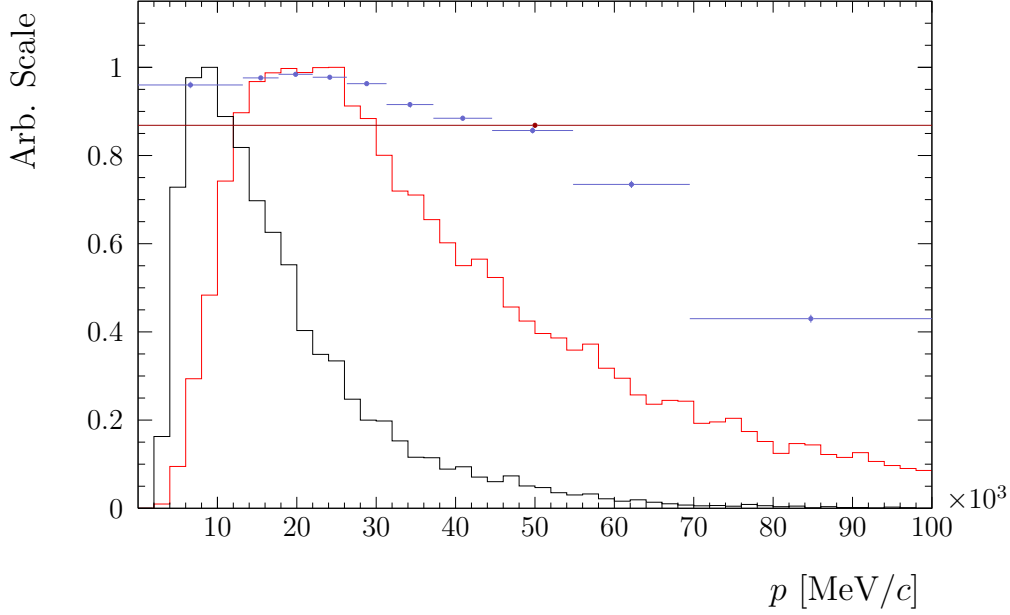


Figure 10.6: *Unoptimised binning scheme (blue) for computing kaon PID efficiencies. The kaon momentum distributions are shown for the calibration sample distribution (red) and the reference sample distribution (black).*

10.5 Signal window

As described in Section 9, the number of prompt signal candidates is measured by fitting the $\ln \chi^2_{\text{IP}}$ distribution of candidates in a signal window defined in the mass distribution. This mass window requirement removes some signal candidates. The efficiency of the requirement is evaluated by computing the fractional integral of the signal component of the model $f_{\text{Sig.}}(m)$ fitted to the mass distribution. In the case of the D^{*+} measurement, there is an additional requirement that the candidates used in the $\ln \chi^2_{\text{IP}}$ fit fall in a signal window defined in the Δm distribution. The efficiency of the Δm signal window requirement is computed in the same way as for the mass signal window requirement, except the signal component of the model fitted to the Δm distribution $g_{\text{Sig}}(\Delta m)$ is used.

To assign an uncertainty to the integrals, the fractional integral is recomputed in a series of 500 pseudo-experiments. Within each one, each shape parameter is assigned a value sampled from a normal distribution

with a mean as the ‘nominal’ value of that parameter, the value found by the fit, and with a width equal to the uncertainty on that nominal value. Correlations between the different shape parameters are taken into account by sampling parameter values from the covariance matrix found in the nominal fit. The variance on the deviation of the integrals, with respect to the integral computed using the nominal shape parameter values, is taken as the square of the uncertainty on the nominal integral value.

As the shape of some mass signal models are (p_T, y) bin dependent, as described in Section 9.3, the signal window efficiency is computed separately in each (p_T, y) bin. Tables A.15 and A.16 give the per-bin signal window efficiencies for each mode.

10.6 Uncertainties on efficiencies

For the efficiencies computed by counting the number N of decays before and the number k after some selection, the estimator $\hat{\varepsilon}$ of the true efficiency is taken to be k/N . The application of the selection on the N decays is considered as a binomial process, with the probability P to observe k counts given the true efficiency ε and the number of initial N counts given by

$$P(k|\varepsilon, N) = \binom{N}{k} \varepsilon^k (1 - \varepsilon)^{N-k}. \quad (10.19)$$

To utilise the MC error propagation discussed in Section 5.1, a PDF of ε given k and N is needed. This can be found by using Bayes’ theorem

$$P(\varepsilon|k, N) \propto P(k|\varepsilon, N)P(\varepsilon|N). \quad (10.20)$$

where $P(\varepsilon|N)$ is the prior probability and $P(k|\varepsilon, N)$ is the posterior probability. For this analysis, a flat prior in ε is chosen, encoding the belief that, given only N , any value of ε between 0 and 1 is reasonable. Substituting such a prior and Equation (10.19) into Equation (10.20), the posterior becomes

$$P(\varepsilon|k, N) \propto \varepsilon^k (1 - \varepsilon)^{N-k}. \quad (10.21)$$

This is the beta distribution $B(\alpha, \beta)$ for $\alpha = k + 1$ and $\beta = N - k - 1$, and is what is used as the generating PDF in the MC error propagation.

For the purposes of visualisation, it is useful to quote an uncertainty on the estimator $\hat{\varepsilon}$. The standard treatments of uncertainty evaluation on an efficiency can yield symmetric uncertainties of zero, or even $\pm 1\sigma$ intervals that include unphysical efficiencies, when extreme values of k and N are used [78]. Instead, asymmetric 68.2% confidence intervals are computed using the Agresti-Coull method [79], and are given as uncertainties on all efficiencies computed as ratios of counts.

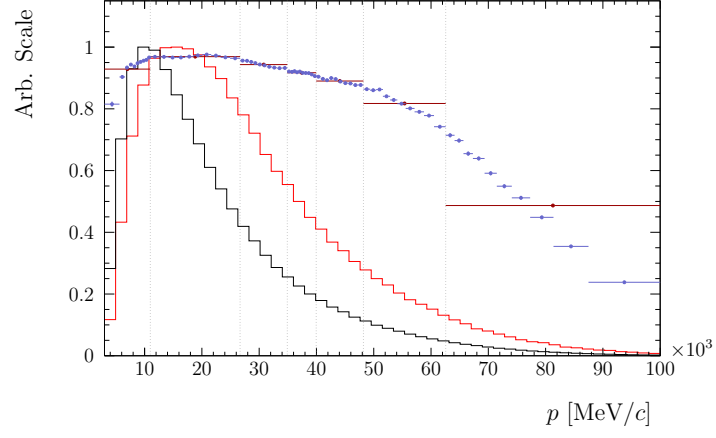
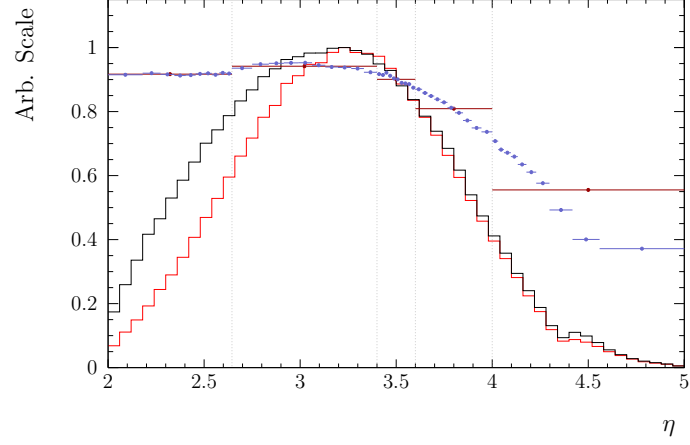
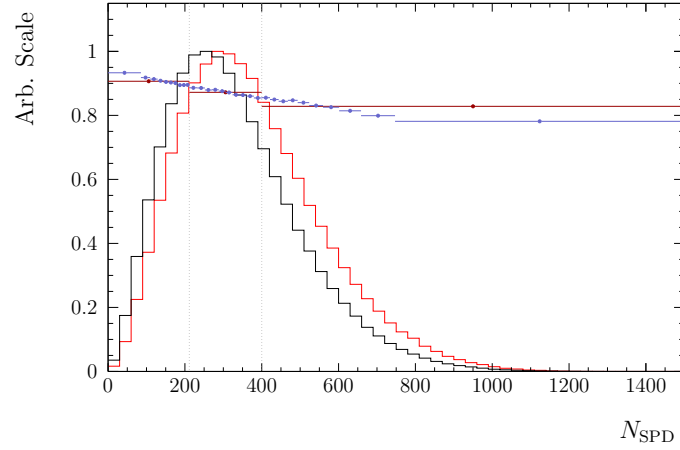
(a) *Track p* (b) *Track η* (c) *N_{SPD}*

Figure 10.7: *Binning scheme (dark red) used for computing kaon PID efficiencies. A finer binning scheme (blue) is given to illustrate the varying efficiency within the nominal binning scheme. The calibration sample distribution (red) and reference sample distribution (black) are also shown. The distributions are kaon momentum (a) and pseudorapidity (b) and the event multiplicity (c).*

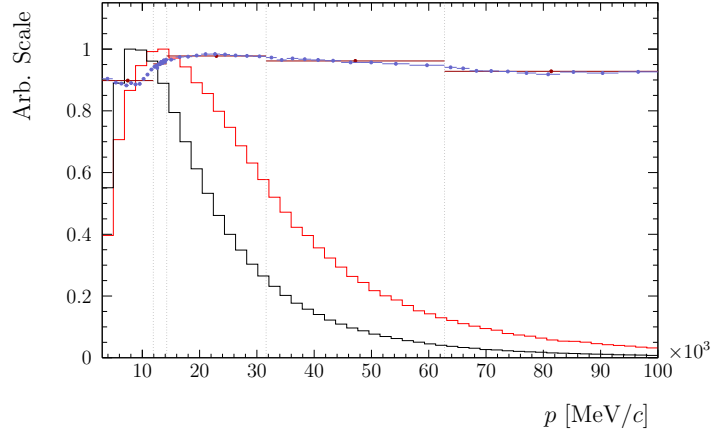
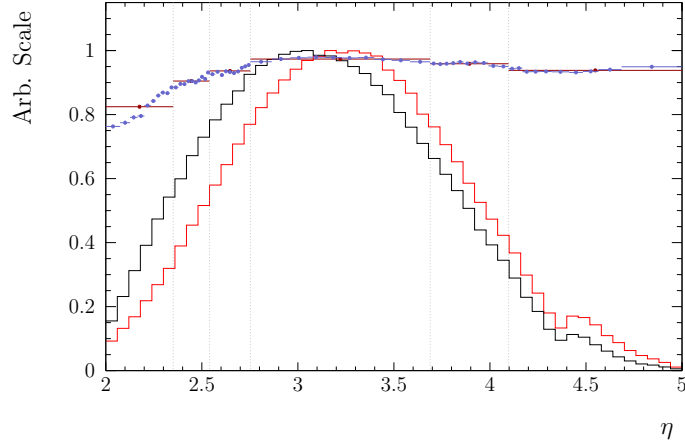
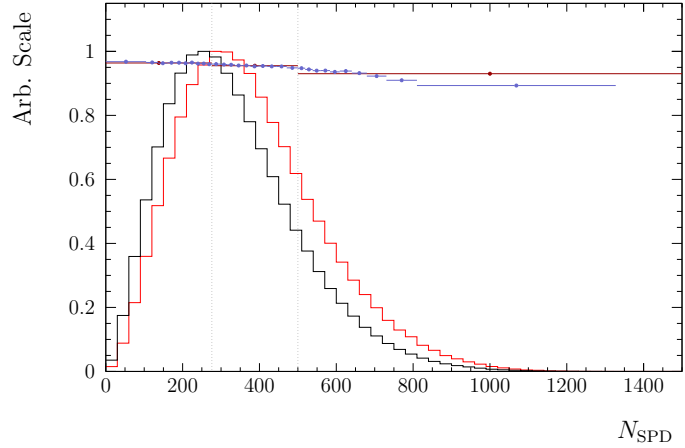
(a) Track p (b) Track η (c) N_{SPD}

Figure 10.8: Binning scheme (dark red) used for computing pion PID efficiencies. A finer binning scheme (blue) is given to illustrate the varying efficiency within the nominal binning scheme. The calibration sample distribution (red) and reference sample distribution (black) are also shown. The distributions are pion momentum (a) and pseudorapidity (b) and the event multiplicity (c).

11 Systematic uncertainties

In the discussion of systematic uncertainties, it is important to consider the correlations of the uncertainties both between (p_T, y) bins and between charm hadrons. Uncertainties that are fully correlated cancel in ratios. In addition, the quantification of the goodness-of-fit of a fit to the cross-section measurements can change significantly depending on whether correlations are accounted for. In what follows, the range of correlations of each uncertainty between bins and between charm hadrons will be given, with a summary presented in Section 11.10.

The systematic uncertainties are treated as nuisance parameters of the cross-section measurement, being propagated as an additive parameter on the respective term they relate to. The nuisance parameters are modelled as probability density functions in the framework of Monte Carlo error propagation, as described in Section 5.1. Typically, the PDF is a normal distribution with a mean of zero and a width corresponding to the size of the systematic variation being assessed, but the exact form will be given alongside the description of the uncertainty.

11.1 Finite Monte Carlo sample size

The Monte Carlo samples used to estimate the acceptance, reconstruction, and selection efficiencies are of a finite size, and so the computed efficiencies carry a statistical uncertainty. This is propagated to the cross-section measurements as a systematic uncertainty. As the set of MC events is unique for each (p_T, y) bin and mode, this systematic uncertainty is uncorrelated between bins and between decay modes. As discussed in Section 10.6, the a

posteriori probability for the estimated efficiency is described by the beta distribution, and so the corresponding nuisance parameter is also modelled by this distribution.

The resulting systematic uncertainties, relative to the central value of the cross-section measurement, are given in Tables A.17 and A.18 for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$.

11.2 Monte Carlo modelling

If the simulation does not correctly model the selection variables, the selection efficiencies that are evaluated with MC, described in Section 10.3, will be incorrect. To assess the accuracy of the Monte Carlo modelling, signal distributions are extracted from the data using the *sPlot* method [77] and are compared to the simulation. With the *sPlot* method, per-candidate signal weights are computed based on the ratio of the signal and background PDFs used in the $\ln \chi^2_{\text{IP}}$ fits described in Section 9.2. The weights are defined in such a way that their sum in a subsample of the data, selected in a way that is uncorrelated with $\ln \chi^2_{\text{IP}}$, is equal to the signal yields in that subsample. For variables that are uncorrelated with $\ln \chi^2_{\text{IP}}$, this then allows signal distributions to be estimated by binning the data as a histogram.

As the definition of the selection efficiency does not include the effects of the PID selection, the distributions of the real data are corrected for the PID efficiency by a per-decay weighting. The signal distributions in data are only available after the selection has been applied, and so the same selection is also applied to the simulation (with the exception of the PID requirements). Comparisons of an example set of variables used in the $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$ selections are shown in Figures 11.1 and 11.2, integrated across all (p_{T}, y) bins for visualisation. The agreement is generally good.

To quantify the effect of the mis-modelling on the selection efficiencies, the following procedure is performed for each selection variable x :

1. Find the cut value y on x which rejects half of the selected simulated sample;

2. Apply the cut value y on x to the real data, counting the number of signal candidates passing the requirement as the sum of signal weights divided by the PID selection efficiencies

$$N_{\text{Passed}} = \sum_i^N I(i, x, y) \frac{w_{i,\text{signal}}}{\varepsilon_{i,\text{PID}}}, \quad (11.1)$$

where $I(i, x, y) = 1$ if the value of x for the i th event passes the requirement y on x , and is zero otherwise;

3. The efficiency-corrected signal yield corresponding to the requirement on x is $2N_{\text{Passed}}$, which should equal the number of signal candidates before the requirement y , if the MC describes the data. The quantity

$$\Delta = 2N_{\text{Passed}} - \sum_N^i \frac{w_{i,\text{signal}}}{\varepsilon_{\text{PID}}}, \quad (11.2)$$

defines the deviation from the assumption that the data is well-modelled by the MC.

This procedure is repeated in every (p_T, y) bin, for a total of $N_{p_T} \cdot N_y \cdot N_{\text{variables}}$ values of Δ .

To account for correlations between variables, which, if neglected, may lead to double-counting of the mis-modelling, the procedure is repeated in ten thousand pseudo-experiments per selection variable. For each trial, each candidate is sampled n times from the real data, with n being sampled from a Poisson distribution of mean 1. The total systematic uncertainty on the efficiency-corrected signal yield due to MC mis-modelling is the mean of the set of ten thousand values of Δ .

The value of Δ for a given variable in a single (p_T, y) bin is usually not statistically significant, but the integration of the values of Δ over all (p_T, y) values can be, and so a difference is evaluated for all (p_T, y) bins. The relative uncertainty on the efficiency-corrected yield N/ε is the difference integrated across all (p_T, y) divided by its uncertainty. This is given for all modes in Table 11.2.

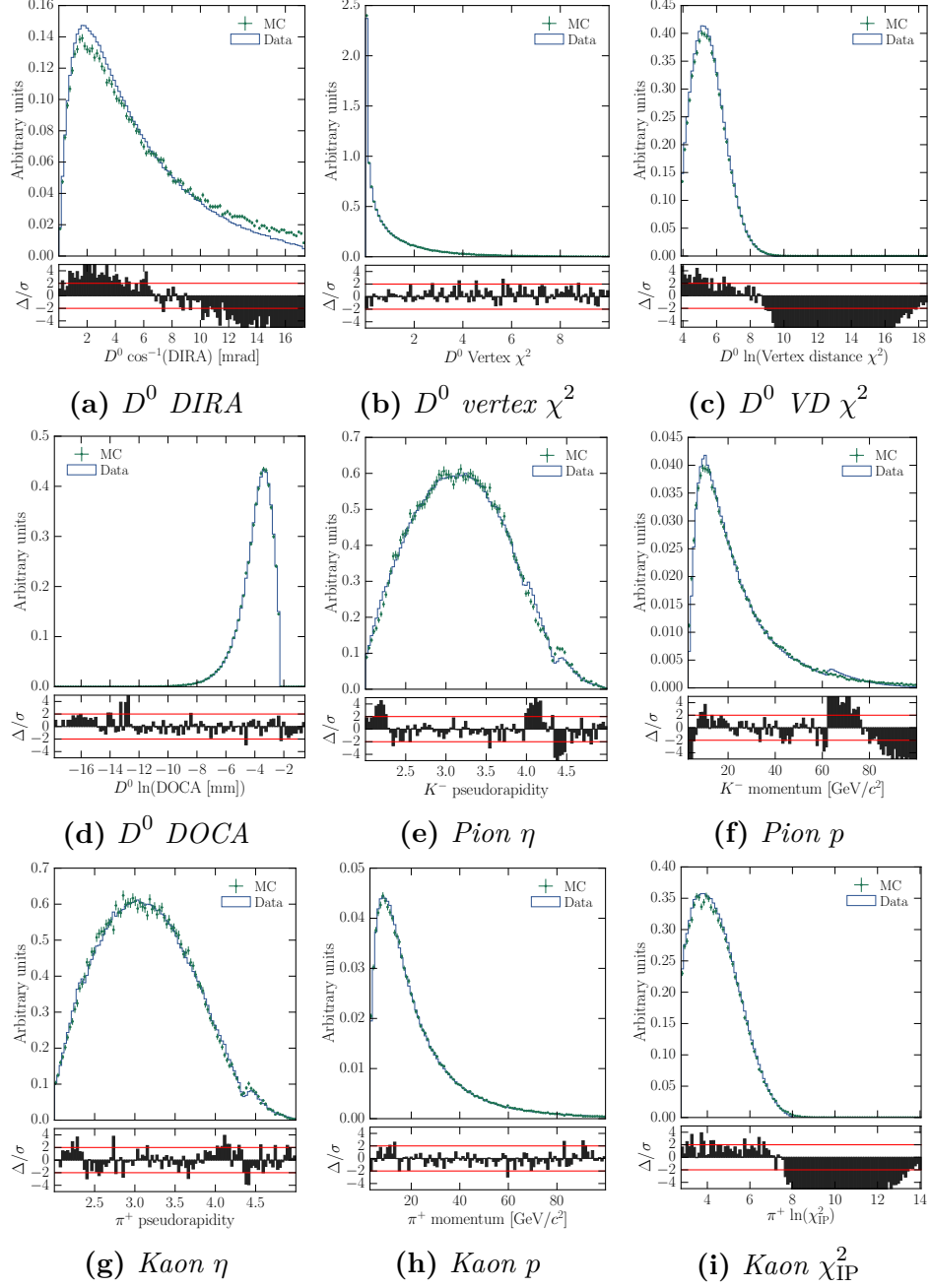


Figure 11.1: Comparison of several variables used in the $D^0 \rightarrow K^- \pi^+$ selection between data (blue line) and MC (green points). Underneath each pair of distributions is a pull plot, showing the significance of the deviation in each histogram bin.

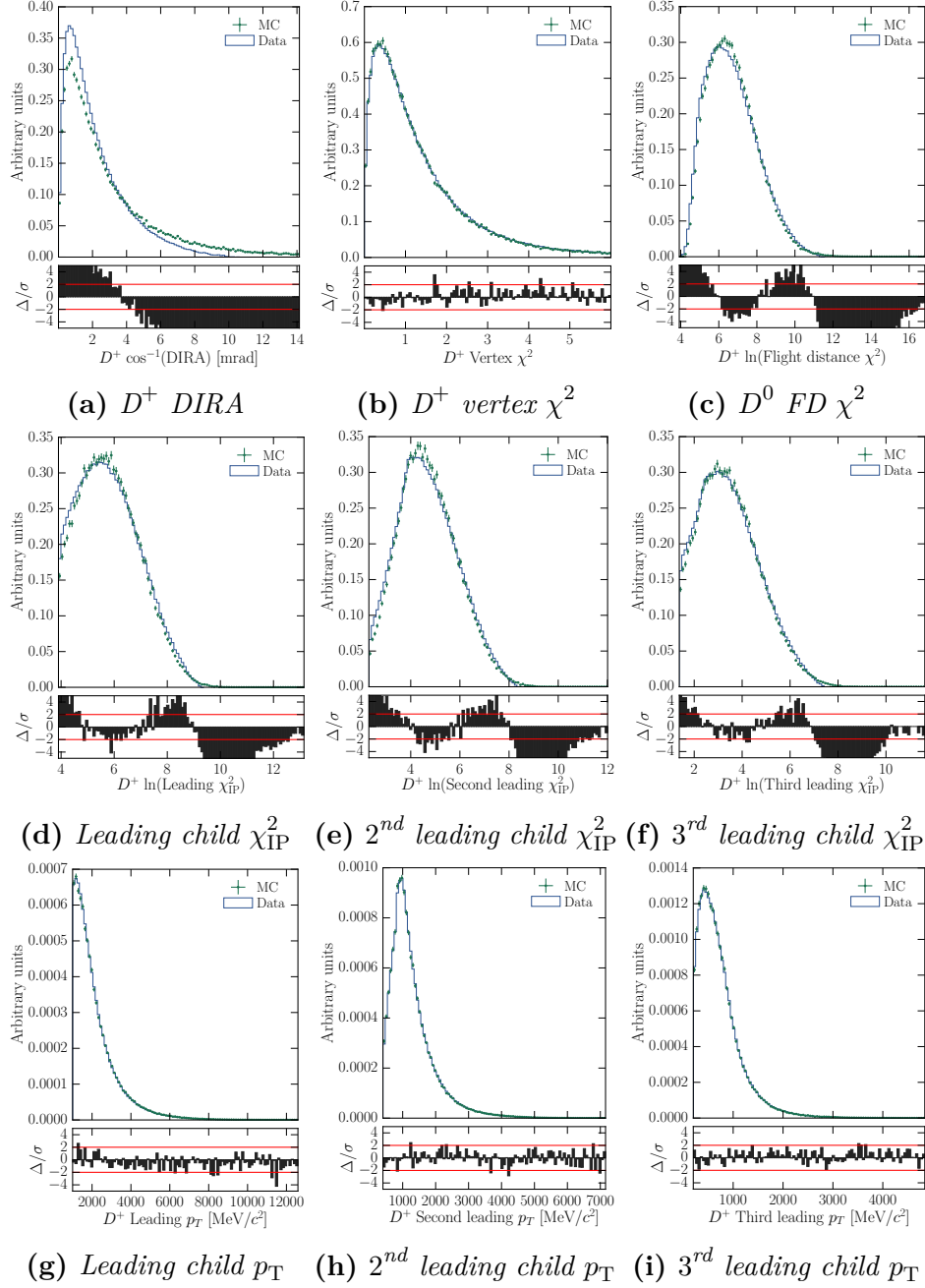


Figure 11.2: Comparison of several variables used in the $D^+ \rightarrow K^- \pi^+ \pi^+$ selection between data (blue line) and MC (green points). Underneath each pair of distributions is a pull plot, showing the significance of the deviation in each histogram bin.

11.3 Monte Carlo truth matching

The truth matching efficiency described in Section 10.2 is computed using an MC sample of finite size, and so carries a statistical uncertainty. As in Section 11.1, this statistical uncertainty is propagated to the cross-section measurements as a systematic uncertainty. The MC sample size is too small to compute the truth matching systematic in (p_T, y) bins, and so a value is computed using the integrated MC. The values for each mode are given in Table 11.3.

11.4 Particle identification calibration

As stated in Section 10.4, a key assumption in the PID calibration is that the PID efficiency ε_{PID} is single-valued in a $(p, \eta, N_{\text{SPD}})$ bin. Generally, this assumption is satisfied by fine binning, but the limited statistics of the calibration sample prohibit this as it would create large statistical uncertainties. There are two systematic uncertainties associated to this effect: the statistical uncertainty on ε_{PID} due to the finite size of the calibration sample; and the breakdown of the assumption that ε_{PID} is single-valued in a $(p, \eta, N_{\text{SPD}})$ bin.

Calibration sample size

The efficiency ε_i in the i th $(p, \eta, N_{\text{SPD}})$ bin of the efficiency histogram has a statistical uncertainty $\delta\varepsilon_i$ due to the finite size of the calibration sample. Each $(p, \eta, N_{\text{SPD}})$ bin contains a disjoint set of calibration tracks, and hence the efficiencies between bins fluctuate independently. There will be a non-zero correlation between the PID efficiencies of tracks in a single decay if those tracks fall in the same efficiency bin, as well as between different decays if there is an overlap in the efficiency bins that are used. To model these correlations, a series of ten thousand pseudo-experiments are performed to propagate the statistical uncertainty on ε_i to a corresponding uncertainty on ε_{PID} .

In each pseudo-experiment, a new set of efficiencies ε'_i is drawn from a normal distribution of mean ε_i and width $\delta\varepsilon_i$. The full calibration procedure is then repeated using the set of ε'_i to obtain an overall efficiency $\varepsilon'_{\text{PID}}$ for that pseudo-experiment. The statistical uncertainty $\delta_C\varepsilon_{\text{PID}}$ on the nominal efficiency is then taken to be the standard deviation of the distribution of $\varepsilon'_{\text{PID}}$.

A common set of efficiency histograms is generated and is used in all pseudo-experiments for computing the uncertainty in each bin for all charm mesons, allowing the correlations to be determined. As different (p_T, y) bins and decay modes have different kinematic distributions for their final-state particles, this correlation is found to be significant within a decay mode only for neighbouring (p_T, y) bins, and negligible between different modes.

The relative uncertainty $\delta_C\varepsilon_{\text{PID}}/\varepsilon_{\text{PID}}$ is given in Tables A.21 and A.22 for $D^0 \rightarrow K^-\pi^+$ and $D^+ \rightarrow K^-\pi^+\pi^+$. The distribution in $\varepsilon'_{\text{PID}}$ is taken as the posterior PDF for this uncertainty in the MC error propagation, translated such that the most probable value is coincident with zero.

Binning scheme

The iterative procedure to find the binning scheme used for the efficiency histograms, described in Section 10.4, is limited by the sample size of the calibration data. To be able to assess the possible shift in ε_{PID} one would see if the calibration sample size was larger, two-dimensional PDFs in track p and η are made using KDEs [72]. The event multiplicity is not modelled since the PID efficiency variation in this feature is seen to be significantly smaller than that in track p and η , and hence ε_{PID} is less sensitive to the binning in that variable.

One two-dimensional KDE is constructed for each particle species both before and after the PID requirements under study. These densities $P_{K/\pi}(p, \eta)$ and $P'_{K/\pi}(p, \eta)$ can be used to obtain the efficiency ε_i in the (p, η) bin of volume V as

$$\varepsilon_i = \frac{C'}{C} \frac{\int_V P'(p, \eta) dp d\eta}{\int_V P(p, \eta) dp d\eta}, \quad (11.3)$$

where C' is the total signal calibration sample size before (after) the requirements. Figure 11.3 shows example one-dimensional KDE projections in kaon η in various p bins in comparison with the signal calibration data. Figure 11.4 shows the one-dimensional projections of the efficiencies in comparison with a fine one-dimensional binned approximation. In both cases, a good agreement is seen between the KDE estimate and the data.

The KDEs are used as generative PDFs to produce pseudo-experimental data with which to perform the PID calibration. Figures 11.5 and 11.6 compare the two-dimensional efficiency histogram from the real calibration data and the nominal binning with that generated from the KDEs, and a good agreement is seen between the two. Unlike with real calibration data, an effectively infinite amount of data can be generated from the KDEs, allowing the effects of the binning scheme to be probed. This is done by computing ε_{PID} with a progressively finer binning, where it is observed that the value of the difference between the ‘KDE efficiency’ and the nominal efficiency converges, as shown in Figure 11.7. The difference $\delta_B \varepsilon_{\text{PID}}$ between the value at convergence and that found in the nominal procedure is taken as the systematic uncertainty on ε_{PID} due to the chosen binning. A normal distribution is assumed for the corresponding nuisance parameter. The same convergence behaviour is seen in all (p_T, y) bins and across all modes, and so the uncertainty is assumed to be fully correlated across them.

The systematic uncertainty due to the choice of $(p, \eta, N_{\text{SPD}})$ binning relative to the measured cross-section value is given in Tables A.21 and A.22 for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$.

11.5 Luminosity

The beam-gas imaging analysis used to measure the integrated luminosity comes with its own set of systematic uncertainties. The dominant components arise from the limited resolution on the measurement of the beam-beam and beam-gas vertices, from the choice of fit model used to parameterise the transverse beam profiles, and the observed variation in measurements of the visible pp cross-section over time [68]. The uncertainty on the luminosity

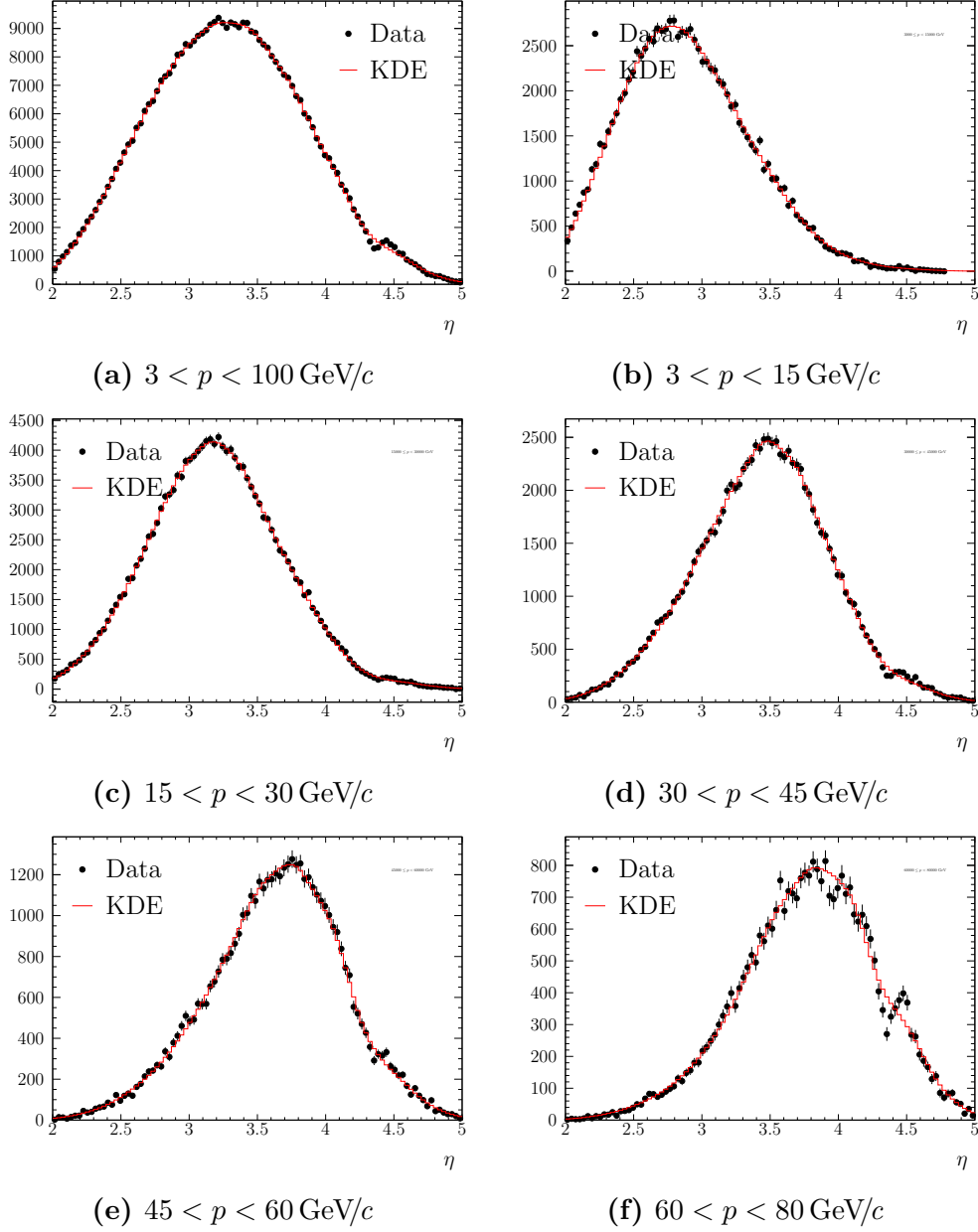


Figure 11.3: Pseudorapidity distribution in the kaon PID signal calibration sample, integrated across all values of kaon momentum (a) and within a series of smaller p bins (b–f).

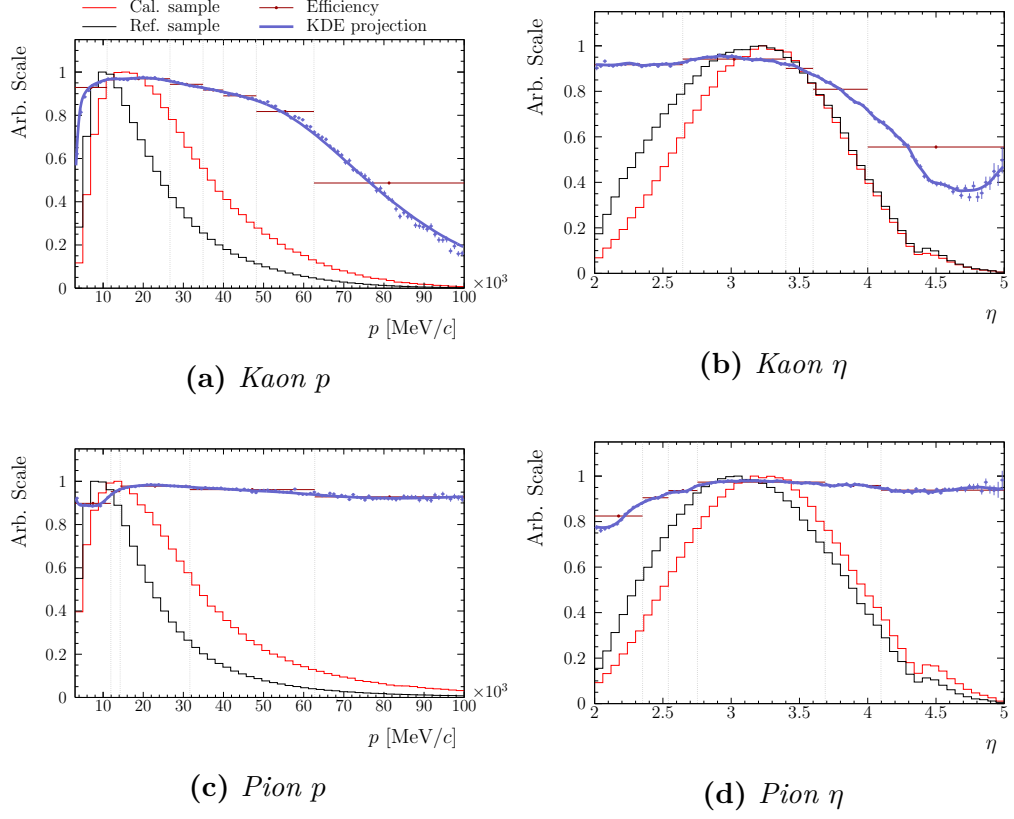


Figure 11.4: As in Figures 10.7 and 10.8, but with an estimate of the efficiency from kernel density estimates overlaid (blue curve).

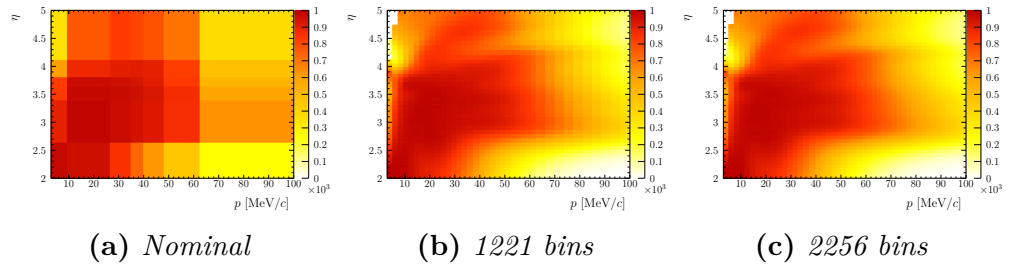


Figure 11.5: Efficiency distributions $\varepsilon_i(p, \eta)$ for kaon tracks in the nominal binning scheme using the real calibration data (a) and higher granularities using data generated from kernel density estimates (b and c).

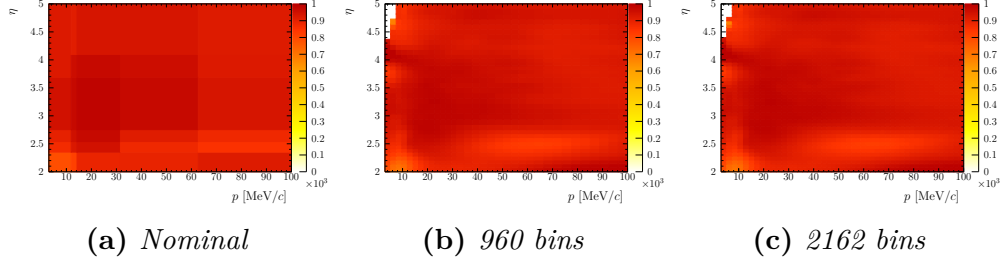


Figure 11.6: Efficiency distributions $\varepsilon_i(p, \eta)$ for pion tracks in the nominal binning scheme using the real calibration data (a) and higher granularities using data generated from kernel density estimates (b and c).

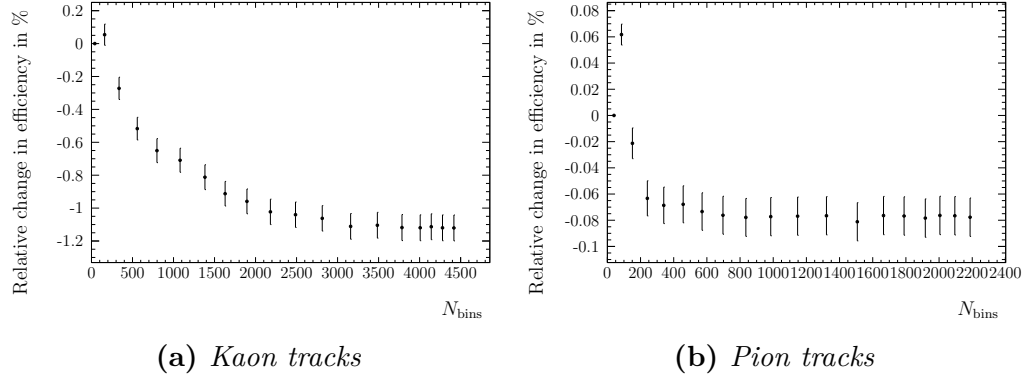


Figure 11.7: Relative change in kaon (a) and pion (b) PID efficiency ε_{PID} , as evaluated using kernel density estimates, when evaluated with a progressively finer histogram binning.

measurement of 3.8 % is propagated to the cross-section measurements, and is fully correlated between all charm mesons and all (p_T, y) bins.

11.6 Tracking efficiency correction

The tracking efficiency corrections are computed as part of a dedicated analysis [36], summarised in Section 10.2, and come with their own set of statistical and systematic uncertainties.

The statistical uncertainties on the correction factors, computed in bins of track p and η , and shown in Figure 10.5. They are propagated to the cross-section measurements with pseudo-experiments in an identical manner

to that used to propagate the statistical uncertainty on the PID efficiency, as presented in Section 11.4. The use of pseudo-experiments encodes the intra- and inter-decay correlations of this uncertainty, and the resulting distribution of uncertainties is used as the PDF of the corresponding parameter on the cross-sections.

The systematic uncertainty on the tracking efficiency correction factors comprises two components: that due to the difference between muon interactions with the detector (used to probe the efficiency) and those of hadrons (used in the cross-section measurements); and that due to the weighting procedure used to equalise the event multiplicity between the real data and the MC. For the former source, a systematic uncertainty of 1.1 % is assigned per pion track and of 1.4 % per kaon track, both of which are dominated by an assumed 10 % uncertainty on the material budget. The latter source contributes 0.4 %. The corresponding nuisance parameters for both sources are modelled as normal distributions.

Unlike the systematic uncertainties on the PID efficiencies, the systematic uncertainties on the correction factors are highly correlated both between (p_T, y) bins and between charm mesons. This is due to the hadronic interaction and weighting uncertainties being dominant and global. The systematic uncertainty on the tracking efficiency correction factor relative to the measured cross-section value is given in Tables A.23 and A.24 for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$.

11.7 Branching fractions

The branching fractions, provided by the Particle Data Group [12] and the CLEO collaboration [49], are listed in Table 5.1. The uncertainties are propagated as a systematic uncertainty on the cross-sections as normal distributions of mean of zero and width equal to the uncertainty. The uncertainties are fully correlated between (p_T, y) bin within a mode, and uncorrelated between all modes, except for that on the $D^0 \rightarrow K^- \pi^+$ branching fraction which enters both the $D^0 \rightarrow K^- \pi^+$ and $D^{*+} \rightarrow D^0 \pi^+$ cross-section measurements, for which a correlation of 95 % is assigned.

11.8 Fit model

The choice of PDFs used to model the data in the fitting procedure described in Chapter 9 is somewhat arbitrary, in that they are not the only functional forms that fit the data well. The magnitude of the difference between the cross-sections computed with other fit models and the nominal model is assigned as a systematic uncertainty on the result. This corresponds to an uncertainty on the true generating PDF for the observed data.

With respect to the nominal signal model used in the $\ln \chi^2_{\text{IP}}$ fit, from Equation (9.6), alternative signal models used in the $\ln \chi^2_{\text{IP}}$ fits are: Equation (9.6) where the Gaussian core is replaced by a ‘double Gaussian’ core, which is the sum of two Gaussian distributions sharing a common mean but having independent widths, and a ‘smeared’ $\ln \chi^2$ distribution, obtained by modifying the usual χ^2 functional form to have a parameterised asymmetry in the widths either side of the mean [80]. Alternative models for the secondary charm component are the PDF used for the prompt signal, and the asymmetric χ^2 function used as an alternative prompt signal model. No alternative models for the background component are tested.

The resulting relative systematic uncertainties are given in Tables A.25 and A.26 for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$.

11.9 Signal window

The signal window applied to the data entering the $\ln \chi^2_{\text{IP}}$ fit, as described in Chapter 9, has an efficiency, as described in Section 10.5. The computation of this efficiency assumes the signal model is known, but the parameters of the model are fitted to the data, and so have an uncertainty, which results in a systematic uncertainty on the signal window efficiency.

The computation of the uncertainty on the signal window efficiency is given in Section 10.5. This is propagated to the cross-sections as the systematic uncertainty, modelled as a normal distribution, and is given in Tables A.27 and A.28 relative to the cross-section values for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$.

11.10 Correlations and summary

The different sources of systematic uncertainty that have been discussed in this chapter can be grouped into three categories: global uncertainties, which are correlated between all bins and all modes; uncertainties which are only correlated between the (p_T, y) bins of a decay mode; and uncertainties which are uncorrelated between bins and modes.

Global uncertainties arise from the luminosity, from the tracking efficiency correction factors, and the $D^0 \rightarrow K^- \pi^+$ branching fraction in between the $D^0 \rightarrow K^- \pi^+$ and $D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow K^- \pi^+$ measurements. Correlated between bins within a decay mode are the uncertainties from the branching fractions, from the arbitrary choice of fit model, and from the PID calibration sample size and binning scheme. The uncertainty due to the finite MC sample size is uncorrelated between bins and between modes.

A summary of the uncertainties entering the cross-section measurements, and their correlations, is given in Table 11.1.

Table 11.1: Systematic uncertainties expressed as fractions of the cross-section measurements, in percent. Uncertainties that are computed bin-by-bin are expressed as ranges giving the minimum to maximum values. Ranges for the correlations between (p_T, y) bins and between modes are also given, expressed in percent.

	D^0	Uncertainties (%)		D^{*+}	Correlations (%)	
		D^+	D_s^+		Bins	Modes
Simulation sample size	2-24	4-55	3-55	2-21	-	-
Simulation modelling	2	1	1	1	-	-
Truth sample size	0.02	0.18	0.11	0.08	-	-
Tracking	3-5	5-17	4-18	5-20	90-100	90-100
PID sample size	0-2	0-1	0-2	0-1	0-100	0-100
PID binning	0-44	0-10	0-20	0-15	100	100
PDF shapes	1-6	1-5	1-2	1-2	-	-
Signal window	0.04-0.4	0.04-1.7	0.06-1.4	0.16-2.84	-	-
Branching fractions	1.2	2.1	5.8	1.5	100	0-95
Luminosity			3.9		100	100

Table 11.2: *Relative uncertainties on the charm cross-sections $\sigma(H_c)$ due to the mis-modelling of the data in the Monte Carlo sample.*

Decay mode	Relative error
$D^0 \rightarrow K^- \pi^+$	1.1 %
$D^+ \rightarrow K^- \pi^+ \pi^+$	1.2 %
$D_s^+ \rightarrow \phi \pi^+$	0.2 %
$D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow K^- \pi^+$	0.91 %

Table 11.3: *Relative uncertainties on the charm cross-sections $\sigma(H_c)$ due to the finite size of the Monte Carlo sample used to assess the truth matching efficiency.*

Decay mode	Relative error
$D^0 \rightarrow K^- \pi^+$	0.02 %
$D^+ \rightarrow K^- \pi^+ \pi^+$	0.18 %
$D_s^+ \rightarrow \phi \pi^+$	0.11 %
$D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow K^- \pi^+$	0.08 %

12 Results

Double differential open charm cross-section measurements are made in bins of charm hadron transverse momentum and rapidity with the relation in Equation (5.3), repeated here for reference,

$$\left. \frac{d^2\sigma(H_c)}{dp_T dy} \right|_i = \frac{1}{\Delta p_T \Delta y} \frac{N_i(H_c \rightarrow f)}{\varepsilon_i(H_c \rightarrow f) \cdot \mathcal{B}(H_c \rightarrow f) \cdot \mathcal{L}_{\text{int.}}}.$$

The prompt signal yields $N_i(H_c \rightarrow f)$ are measured with the maximum likelihood fit described in Chapter 9, and the measurements of the acceptance, reconstruction, and total selection efficiency $\varepsilon_i(H_c \rightarrow f)$ are detailed in Chapter 10. The branching fractions $\mathcal{B}$ are listed in Table 5.1, and the integrated luminosity $\mathcal{L}_{\text{int.}}$ is $(4.98 \pm 0.19) \text{ pb}^{-1}$, as given in Equation (7.1). The estimation of the associated systematic uncertainties on these quantities is described in Chapter 11. With all inputs combined, the measurements are presented in Figures 12.1 and 12.2. They are compared with the three sets of predictions, where available for a given meson and (p_T, y) bin, described in Section 6.2. A comparison of the data with the predictions is given in Section 12.3. Preceding that, measurements of charm production ratios between proton-proton centre-of-mass energies and between charm mesons are presented.

12.1 Production ratios

The predicted ratios of prompt charm production cross-sections between different centre-of-mass energies can be much more precise than absolute predictions, as several common sources of uncertainties cancel [63, 66, 67].

Using the present results obtained at $\sqrt{s} = 13 \text{ TeV}$ and the corresponding results at $\sqrt{s} = 7 \text{ TeV}$ [10], the ratios $R_{13/7}$ are measured for D^0 , D^+ , D_s^+ , and D^{*+} mesons. The $\sqrt{s} = 13 \text{ TeV}$ measurements are re-binned to match the binning used in the $\sqrt{s} = 7 \text{ TeV}$ results and the production ratios are presented for $0 < p_T < 8 \text{ GeV}/c$ and $2 < y < 4.5$. In the calculation of the uncertainties on the ratios, those due to the use of the branching fractions cancel. Correlations of 30 % and 50 % are assumed for the uncertainties on the luminosity measurements and tracking efficiency corrections, respectively, and all other uncertainties are assumed to be uncorrelated. Figure 12.3 shows the measured ratios compared with predictions [63, 66, 67].

Cross-section ratios between different charm mesons are computed. It is generally assumed that the fragmentation fractions $f(c \rightarrow H_c)$, as in Equation (5.4), are universal, in that they do not depend upon the hard process that created the charm quark [50, 81]. Comparing meson cross-section ratios measured at different colliders can test this assumption. Measurements of open charm meson production have been made using e^+e^- collision data by the BaBar, Belle, and CLEO collaborations [82–84], collectively ‘ B factories’, and so comparisons of meson ratios are made with those results. The B factory measurements of D^0 , D^+ , D_s^+ , and D^{*+} production were made using the same set of final states as the analysis described here, and so more precise comparisons can be made by computing the ratios of cross-sections multiplied by the ratio of branching fractions $\sigma(H_c) \times \mathcal{B}(H_c \rightarrow f)$, for example

$$R_{D^+/D^0} = \frac{\sigma(D^+) \mathcal{B}(D^+ \rightarrow K^- \pi^+ \pi^+)}{\sigma(D^0) \mathcal{B}(D^0 \rightarrow K^- \pi^+)}.$$

The uncertainty due to the branching fractions does not enter in this ratio. The meson ratios, differential in p_T and y , are presented in Figures 12.4 and 12.5, along with comparisons to the ratios of the integrated measurements made by the B factories. For most ratios, the Belle and CLEO results are shown for comparison with the LHCb results. However, CLEO did not measure the D_s^+ production cross-section, and so for ratios including D_s^+ the BaBar D_s^+ measurement is combined with the relevant CLEO result.

12.2 Integrated cross-sections

Integrated production cross-sections $\sigma(pp \rightarrow H_c X)$ for each charm meson are computed as the sum of the per-bin measurements, where the uncertainty on the sum takes into account the correlations between bins discussed in Chapter 11. For D_s^+ and D^{*+} mesons, the kinematic region considered is $1 < p_T < 8 \text{ GeV}/c$, $2 < y < 4.5$ due to insufficient data below $p_T = 1 \text{ GeV}/c$, while for D^0 and D^+ the same kinematic region as for the ratio measurements is used. The upper limit is chosen to coincide with that of the LHCb measurements at $\sqrt{s} = 7 \text{ TeV}$.

The D^0 and D^{*+} cross-section results contain bins in which a measurement was not possible, and so the cross-section integration requires a correction factor to account for the missing measurements. A method based on theory calculations is chosen, wherein a multiplicative correction factor is computed as the ratio between the predicted integrated cross-section within the considered kinematic region and the sum of all per-bin cross-section predictions for bins for which a measurement exists. This method uses the POWHEG+NNPDF3.0L predictions [63] for D^0 and the FONLL predictions [67] for D^{*+} . The uncertainty on the extrapolation factor is taken as the difference between factors computed using the upper and lower bounds of the theory predictions, and is propagated to the integrated cross-sections as a systematic uncertainty. Table 12.1 gives the integrated cross-sections for D^0 , D^+ , D_s^+ , and D^{*+} mesons.

The total charm cross-section $\sigma(pp \rightarrow c\bar{c}X)$ is calculated as

$$\sigma(pp \rightarrow c\bar{c}X) = \frac{\sigma(H_c + \bar{H}_c)}{2f(c \rightarrow H_c)}, \quad (12.1)$$

for each decay mode. The term $f(c \rightarrow H_c)$ is the quark to hadron transition probability, and the factor 2 accounts for the inclusion of charge conjugate (anti-charm) states in the measurement. The transition probabilities have been computed using measurements at e^+e^- colliders operating at a centre-of-mass energy close to the $\Upsilon(4S)$ resonance [50], and are listed in Table 5.2. The fragmentation fraction $f(c \rightarrow D^0)$ has an overlapping contribution from $f(c \rightarrow D^{*+})$.

The precision on the $c\bar{c}$ cross-section can be improved by combining the individual measurements. For this, the D_s^+ and D^{*+} inputs are neglected, as the D^0 and D^+ measurements are considerably more precise. The D^0 and D^+ inputs are combined using the method of the best linear unbiased estimator, or BLUE method [85]. The combination of the D^0 and D^+ measurements gives

$$\sigma(pp \rightarrow c\bar{c}X)_{p_T < 8 \text{ GeV}/c, 2.0 < y < 4.5} = 2840 \pm 3 \pm 170 \pm 150 \text{ fb},$$

where the uncertainties are statistical, systematic, and from the uncertainty on the fragmentation fractions.

A comparison of the $c\bar{c}$ cross-sections with predictions in the p_T range $0 < p_T < 8 \text{ GeV}/c$ is given in Figure 12.6. The same Figure also shows a comparison of $\sigma(pp \rightarrow c\bar{c}X)$ for $1 < p_T < 8 \text{ GeV}/c$ computed using each of the four open charm mesons.

Ratios of the integrated $\sigma(H_c) \times \mathcal{B}(H_c \rightarrow f)$ measurements are given in Table 12.2.

12.3 Interpretation and comparison with theory

Figure 12.6b shows that the integrated cross-section measurements agree very well between all four charm mesons in the region $1 < p_T < 8 \text{ GeV}/c$, $2 < y < 4.5$. A similar level of agreement is seen in the measurements including the $0 < p_T < 1 \text{ GeV}/c$ region made with the D^0 and D^+ mesons, as shown in Figure 12.6a. The two absolute predictions for the $c\bar{c}$ cross-section shown on Figure 12.6a agree with the data within their very large uncertainties, however the most precise prediction, the “scaled” number from the POWHEG+NNPDF3.0L group, does not agree with the LHCb measurement. This particular prediction was made by predicting the ratio $R_{13/7}$ and then scaling it with the LHCb measurements at $\sqrt{s} = 7 \text{ TeV}$. The tension between the data and this predictions suggests a systematic mis-modelling that scales with the centre-of-mass energy.

Comparing the double differential cross-section results with the predictions, presented in Figures 12.1 and 12.2, the data consistently lie at or above the upper limit of the $\pm 1\sigma$ uncertainty bands on the predictions. The shape in both p_T and y is well-modelled, but diverges from the models with increasing p_T and decreasing rapidity. The GMVFNS prediction provides the best description of data, although it is unable to make any predictions below $p_T = 3$ GeV/ c due to large theoretical uncertainties. The FONLL prediction set is similarly restricted in the low p_T regions, with very low p_T bins having associated uncertainties spanning over two orders of magnitude. The general underestimation of the cross-sections points to a difficulty in predicting the overall scale of double differential charm cross-sections. The measurements presented here can be used to constrain the parton density functions further, which may help improve the agreement between future measurements and predictions.

The agreement between the ratios of cross-sections at $\sqrt{s} = 13$ and 7 TeV shown in Figure 12.3 is good, although the possibility of interpretation is limited by the large uncertainties on the data, which are driven by the large per-bin uncertainties on the 7 TeV measurement. Such ratios are particularly useful for tuning theoretical predictions, due to aforementioned cancellation of input uncertainties. More precise measurements of cross-section ratios between different centre-of-mass energies are presented in Section 12.4.

The measurements of ratios of cross-sections times branching fractions $\sigma(H_c) \times \mathcal{B}(H_c \rightarrow f)$ presented in Figures 12.4 and 12.5 agree with those made at the B factories, indicating that the assumption of universality of the fragmentation fractions $f(c \rightarrow H_c)$ is valid. There is a weak trend in p_T in some ratios, such in Figures 12.4b and 12.4c, which is consistent with heavier mesons having a harder p_T spectrum [50].

12.4 Measurements at $\sqrt{s} = 5$ TeV

Further measurements of prompt open charm production are made using proton-proton collision data taken in November 2015 at a centre-of-mass energy of $\sqrt{s} = 5$ TeV, with a near-identical trigger scheme and a very

Table 12.1: *Prompt charm production cross-sections integrated over $2 < y < 4.5$ and the p_T ranges given. The computation of the extrapolation factors is described in the text. The first uncertainty on the cross-section is statistical, and the second is systematic and includes the contribution from the extrapolation factor. No extrapolation factor is given for $D_{(s)}^+$ as a measurement is available in every bin of the integrated phase space.*

		Extrapolation factor	Cross-section (tb)
D^0	$0 < p_T < 8 \text{ GeV}/c$	1.0004 ± 0.0009	$3240 \pm 4 \pm 190$
D^+	$0 < p_T < 8 \text{ GeV}/c$	-	$1290 \pm 8 \pm 190$
D^0	$1 < p_T < 8 \text{ GeV}/c$	1.0005 ± 0.0009	$2470 \pm 3 \pm 130$
D^+	$1 < p_T < 8 \text{ GeV}/c$	-	$1000 \pm 3 \pm 110$
D_s^+	$1 < p_T < 8 \text{ GeV}/c$	-	$460 \pm 13 \pm 100$
D^{*+}	$1 < p_T < 8 \text{ GeV}/c$	1.0004 ± 0.0023	$880 \pm 5 \pm 140$

similar analysis technique to the one described here [2]. Such measurements can be compared with those made using proton-lead data taken at a nucleon-nucleon centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$ [86] to compute nuclear modification factors that can be compared to theoretical predictions. In addition, as around 500 times more integrated luminosity was used for the $\sqrt{s} = 5 \text{ TeV}$ measurements than 7 TeV, measurements of cross-section ratios $R_{13/5}$ are much more precise than those of $R_{13/7}$ presented in Section 12.1. An increased precision in ratios between centre-of-mass energies increases the constraining power of the measurements on PDFs. Measurements of $R_{13/5}$ are shown in Figure 12.7, where a good level of agreement is seen with the theoretical predictions.

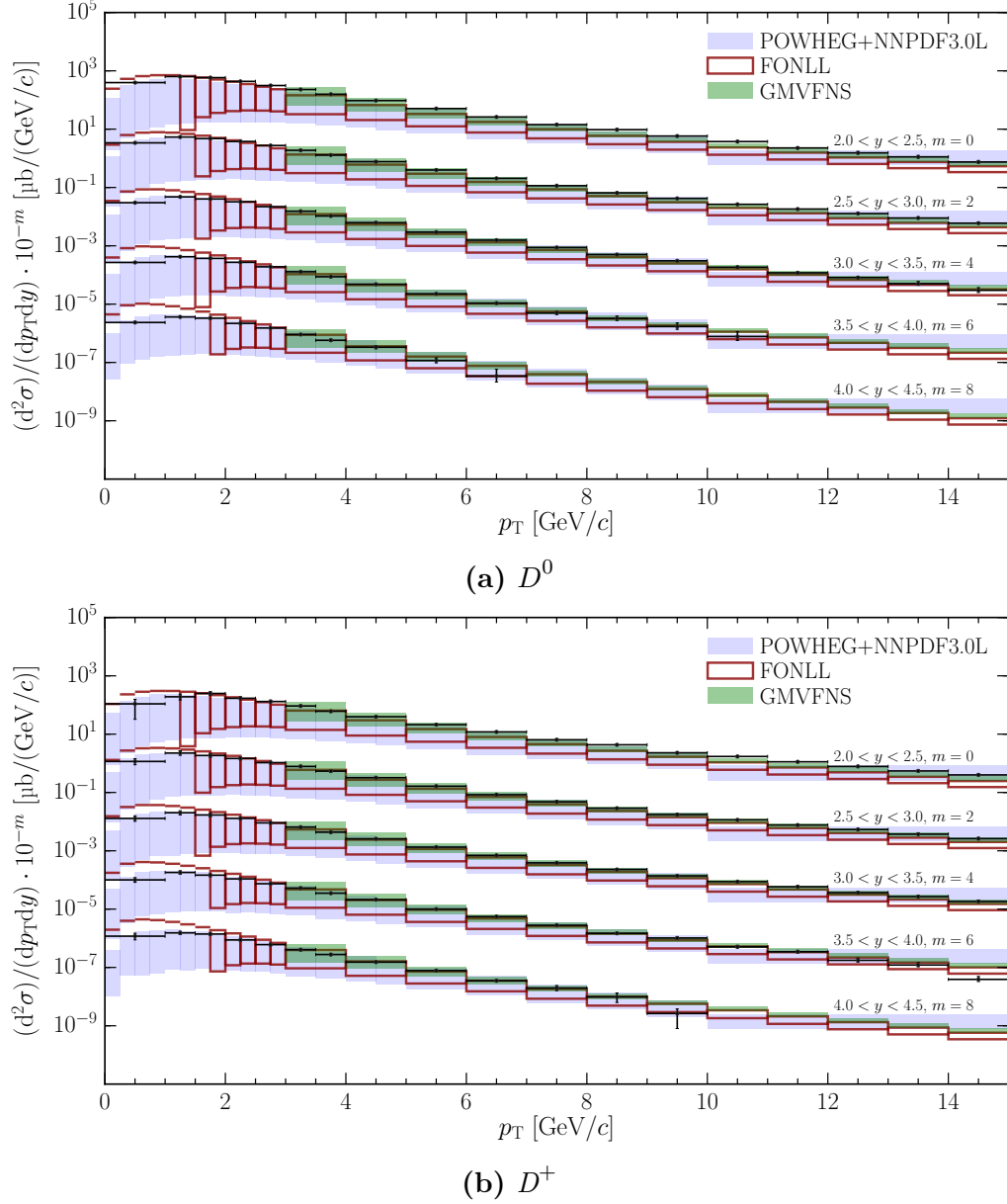


Figure 12.1: Measurements and predictions for the absolute prompt D^0 (a) and D^+ (b) cross-sections at $\sqrt{s} = 13 \text{ TeV}$. Each set of measurements and predictions in a given rapidity bin is offset by a multiplicative factor 10^{-m} , where the factor m is shown on the plots. The boxes indicate the $\pm 1\sigma$ uncertainty band on the theory predictions. In cases where this band spans more than two orders of magnitude only its upper edge is indicated.

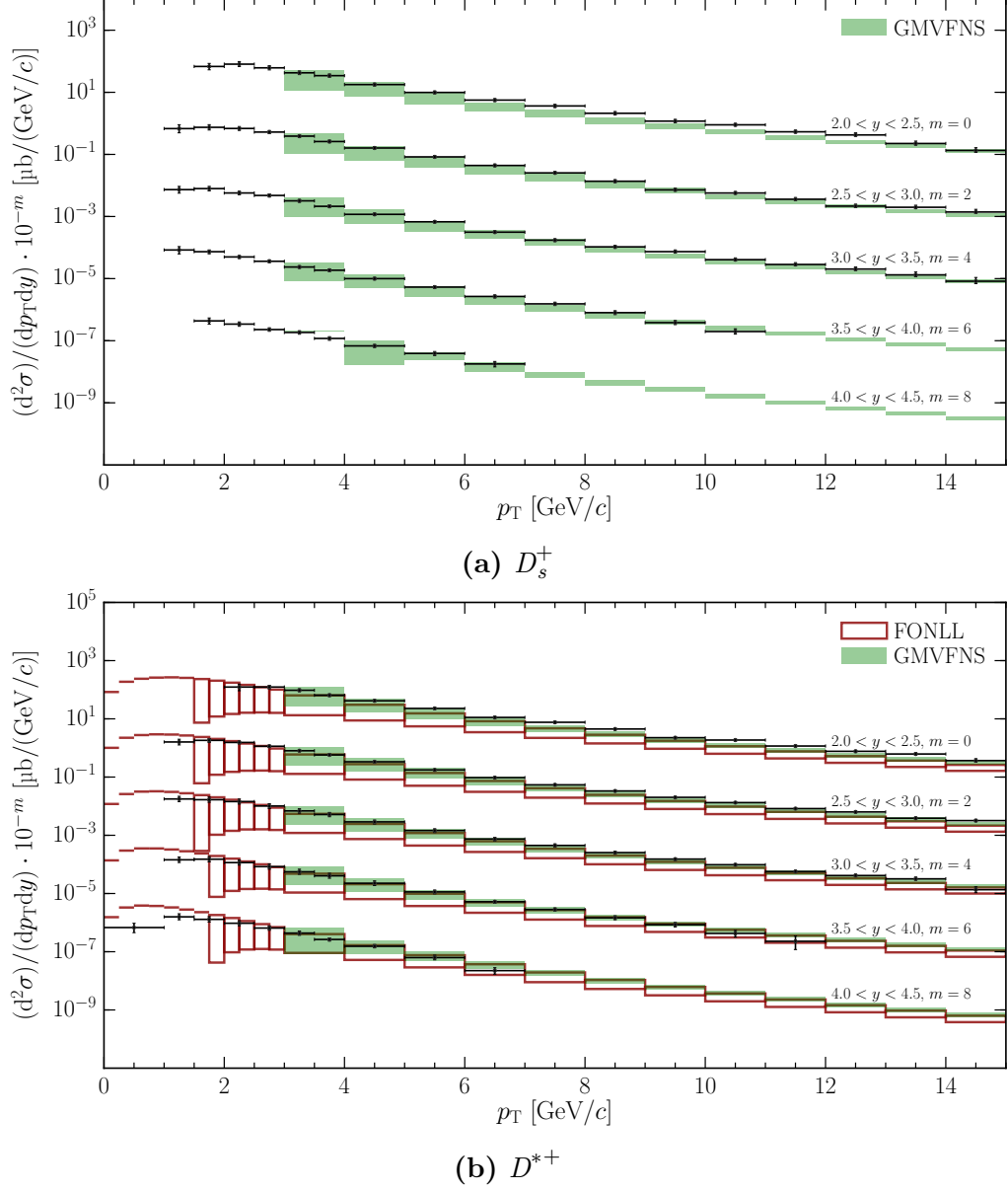


Figure 12.2: Measurements and predictions for the absolute prompt D_s^+ (a) and D^{*+} (b) cross-sections at $\sqrt{s} = 13$ TeV. Each set of measurements and predictions in a given rapidity bin is offset by a multiplicative factor 10^{-m} , where the factor m is shown on the plots. The boxes indicate the $\pm 1\sigma$ uncertainty band on the theory predictions. In cases where this band spans more than two orders of magnitude only its upper edge is indicated.

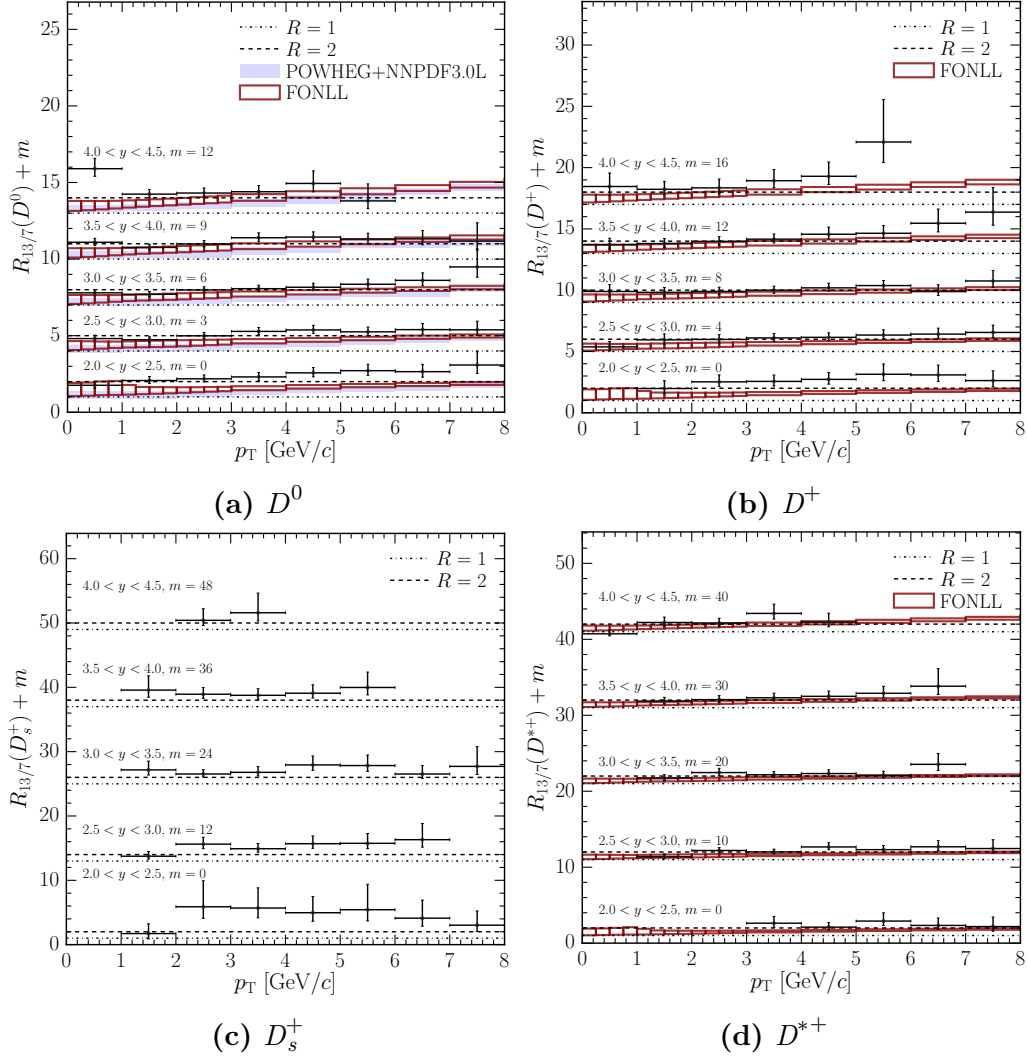


Figure 12.3: Measurements and predictions of the prompt D^0 (a), D^+ (b), D_s^+ (c), and D_s^{*+} (d) cross-section ratios $R_{13/7}$. Each set of measurements and predictions in a given rapidity bin is offset by an additive factor m shown on the plots. The dash-dotted lines indicate the unit ratio for each of the rapidity intervals and the dashed lines indicate a ratio of two. No prediction is available for the D_s^+ ratio.

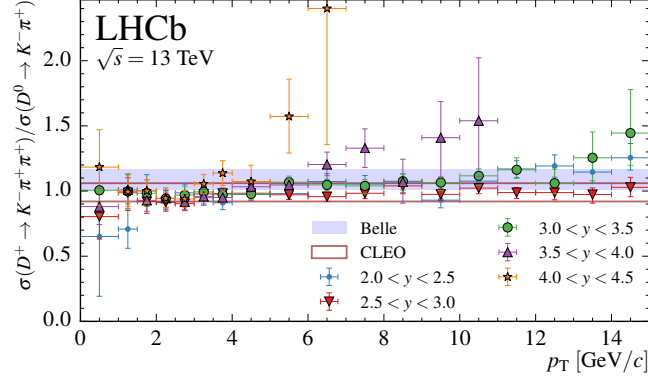
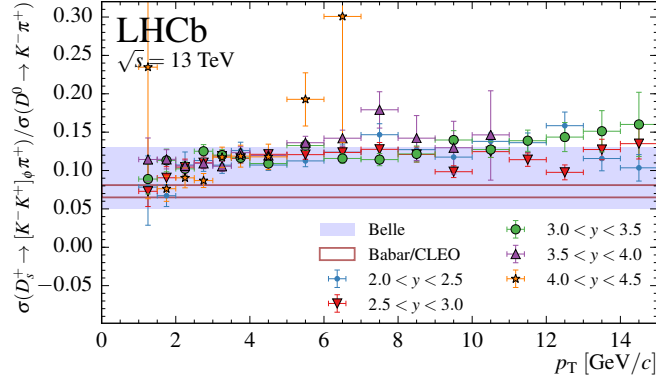
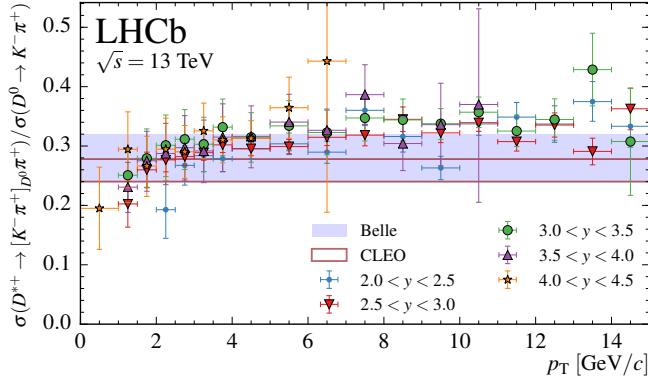
(a) R_{D^+/D^0} (b) $R_{D_s^+/D^0}$ (c) R_{D^{*+}/D^0}

Figure 12.4: Ratios of cross-section times branching fraction measurements of D^+ (a), D_s^+ (b), and D^{*+} (c) mesons with respect to the D^0 measurements. The bands indicate the corresponding ratios computed using measurements from e^+e^- collider experiments [82–84]. The ratios are given as a function of p_T and different symbols indicate different ranges in y . The notation $\sigma(H_c \rightarrow f)$ on each y -axis is shorthand for $\sigma(H_c) \times \mathcal{B}(H_c \rightarrow f)$.

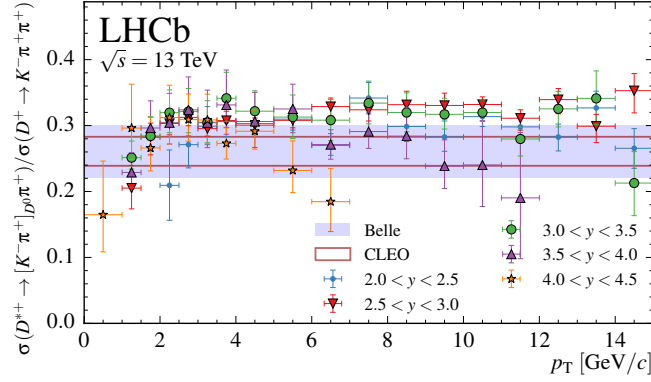
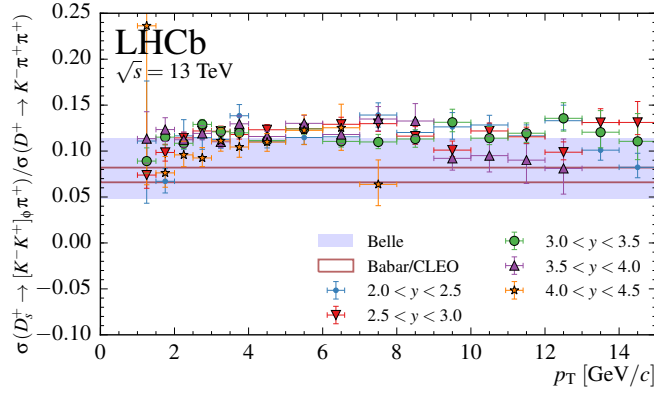
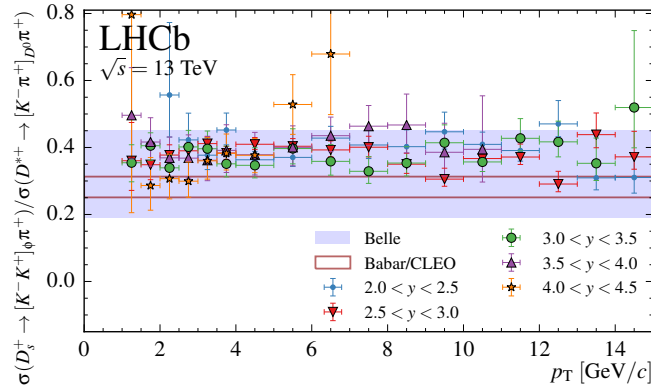
(a) R_{D^{*+}/D^+} (b) $R_{D_s^+/D^+}$ (c) R_{D^{*+}/D_s^+}

Figure 12.5: Ratios of cross-section times branching fraction measurements for D^{*+}/D^+ (a), D_s^+/D^+ (b), and D^{*+}/D_s^+ (c). The bands indicate the corresponding ratios computed using measurements from e^+e^- collider experiments [82–84]. The ratios are given as a function of p_T and different symbols indicate different ranges in y . The notation $\sigma(H_c \rightarrow f)$ on each y -axis is shorthand for $\sigma(H_c) \times \mathcal{B}(H_c \rightarrow f)$.

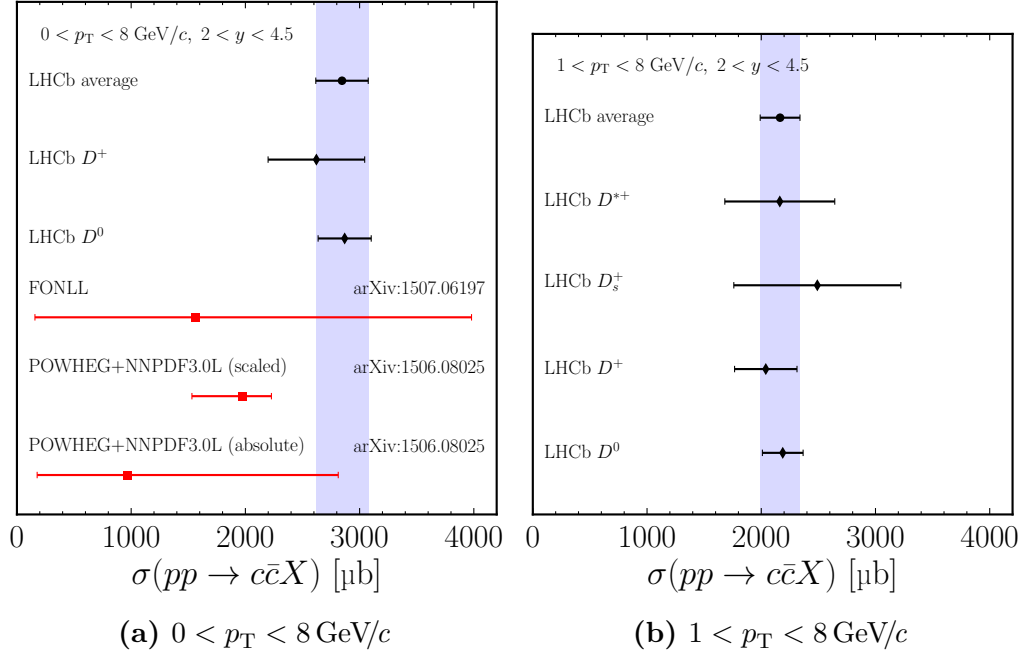


Figure 12.6: Integrated cross-sections (black diamonds), their average (black circle and blue band) and theory predictions (red squares) [63, 67] are shown based on the D^0 and D^+ for $0 < p_T < 8 \text{ GeV}/c$ (a) and for measurements based on all four mesons for $1 < p_T < 8 \text{ GeV}/c$ (b). The “absolute” predictions are based on calculations of the 13 TeV cross-section, while the “scaled” predictions are based on calculations of the 13 to 7 TeV ratio multiplied with the LHCb measurement at 7 TeV [10].

Table 12.2: Ratios of integrated cross-section times branching fraction measurements in the kinematic range $1 < p_{\text{T}} < 8 \text{ GeV}/c$ and $2 < y < 4.5$. The first uncertainty on the ratio is statistical and the second is systematic. The notation $\sigma(H_c \rightarrow f)$ is shorthand for $\sigma(H_c) \times \mathcal{B}(H_c \rightarrow f)$.

Quantity	Measurement
$\sigma(D^+ \rightarrow K^- \pi^+ \pi^+)/\sigma(D^0 \rightarrow K^- \pi^+)$	$0.953^{+0.003+0.060}_{-0.003-0.054}$
$\sigma(D_s^+ \rightarrow \phi \pi^+)/\sigma(D^0 \rightarrow K^- \pi^+)$	$0.106^{+0.003+0.009}_{-0.003-0.010}$
$\sigma(D^{*+} \rightarrow D^0 \pi^+)/\sigma(D^0 \rightarrow K^- \pi^+)$	$0.242^{+0.001+0.027}_{-0.001-0.026}$
$\sigma(D_s^+ \rightarrow \phi \pi^+)/\sigma(D^+ \rightarrow K^- \pi^+ \pi^+)$	$0.112^{+0.004+0.006}_{-0.004-0.009}$
$\sigma(D^{*+} \rightarrow D^0 \pi^+)/\sigma(D^+ \rightarrow K^- \pi^+ \pi^+)$	$0.254^{+0.001+0.016}_{-0.001-0.017}$
$\sigma(D_s^+ \rightarrow \phi \pi^+)/\sigma(D^{*+} \rightarrow D^0 \pi^+)$	$0.444^{+0.013+0.042}_{-0.013-0.052}$

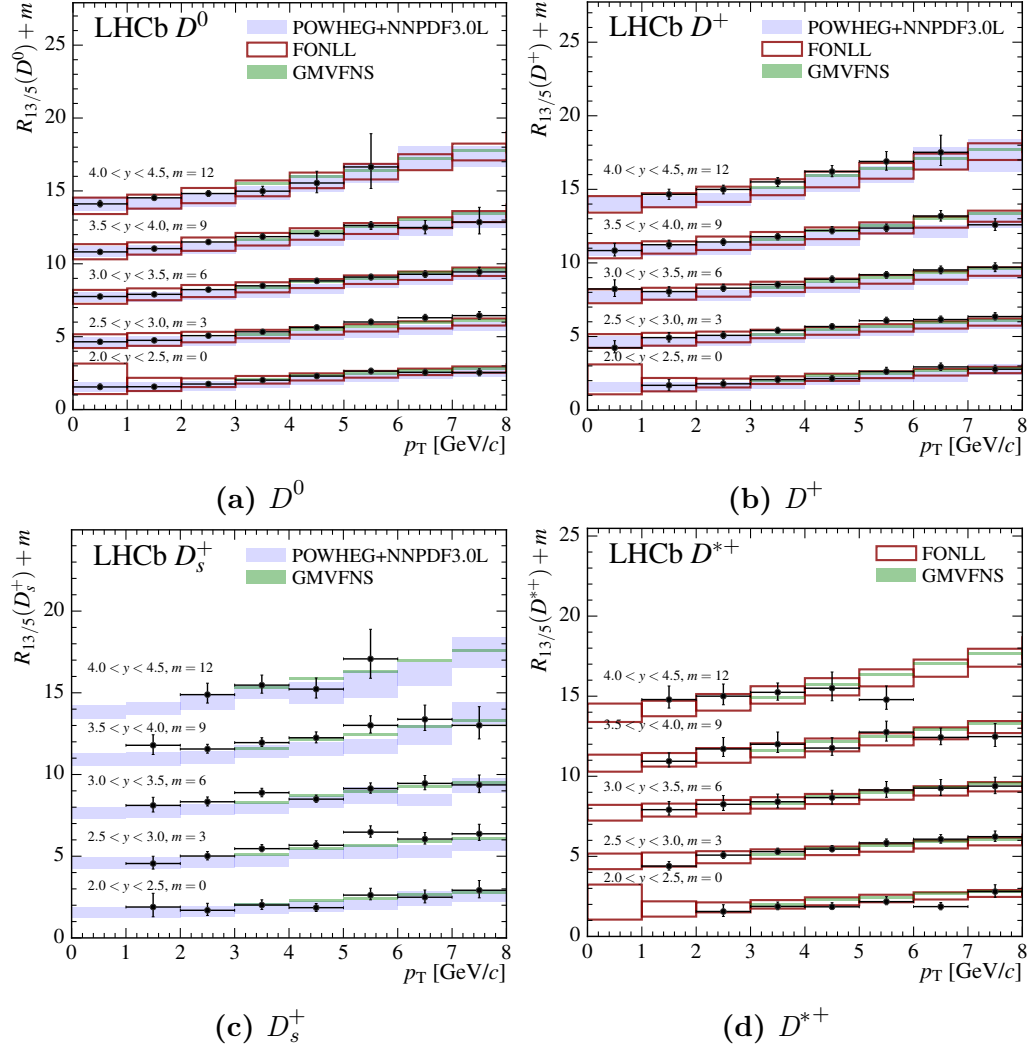


Figure 12.7: Measurements and predictions of the prompt D^0 (a), D^+ (b), D_s^+ (c), and D^{*+} (d) cross-section ratios $R_{13/5}$. Each set of measurements and predictions in a given rapidity bin is offset by an additive factor m shown on the plots.

13 Summary

Measurements of prompt open charm production cross-sections can be used to constrain parton density functions, a key component in QCD predictions of cross-sections at the LHC and beyond. They can also be used to set limits on the background rates in atmospheric neutrino experiments, and are used by the experiments at the LHC to estimate the rates of very rare processes and to plan for future data-taking conditions.

This part has presented the measurements of prompt production cross-sections of the open charm mesons D^0 , D^+ , D_s^+ , and D^{*+} with the LHCb detector at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The measurements are made in bins of charm hadron transverse momentum and rapidity. A significantly higher precision was achieved in comparison to complementary LHCb measurements made at $\sqrt{s} = 7$ TeV, due to a increase in integrated luminosity by a factor 300. The data were analysed immediately after the data-taking thanks to the Turbo stream processing model, unique to the LHCb experiment and new for Run 2 of the LHC. A precise luminosity measurement was made using the beam-gas imaging method, only possible due to the high precision of the LHCb vertex detector.

The data agree with a set of theoretical predictions, each prediction using different inputs and having varying kinematic ranges and uncertainties, although the data generally lie at the upper edge of the uncertainty intervals. This indicates that the (p_T, y) dependence of the cross-sections is well modelled, but that the absolute normalisation is too low. Similar levels of agreement are seen between predictions and measurements of 13 TeV cross-sections relative to those measured at $\sqrt{s} = 7$ TeV and 5 TeV, and for integrated $c\bar{c}$ cross-sections.

It is found that the assumption of universality of charm fragmentation fractions $f(c \rightarrow H_c)$ is valid, as measurements of relative cross-sections between charm mesons agree with similar measurements made at the B factories.

The study of charm physics has the potential to uncover the first hints of physics beyond the SM, although such potential relies on the collection of vast datasets in order to resolve what must be very small effects. With this measurement, it is known that around 3×10^{12} $c\bar{c}$ pairs will be produced within acceptance of the LHCb detector for every 1 fb^{-1} of integrated luminosity, and that the LHCb detector is capable of cleanly and efficiently reconstructing open charm mesons. Despite its name, LHCb is a charm factory.

Part III

Direct CP violation in charm baryon decays

14 Introduction

Direct CP violation has only been observed in the kaon [17, 87] and B meson systems [4–7]. The charm system is the only one in which the SM predicts CP violation but where no CP symmetry breaking has been observed. Particular phenomenological focus has rested on the D^0 meson [9], the lightest charm hadron. The most sensitive experimental probes of direct CP violation in D^0 decays have been measurements of the difference between the rates of direct CP violation in two singly Cabibbo-suppressed (SCS) decays of the D^0 meson, $D^0 \rightarrow K^- K^+$ and $\pi^- \pi^+$.

The rate of direct CP violation in the decay of any particle P to a final state f is parameterised as

$$A_{CP}(P \rightarrow f) = \frac{\Gamma(P \rightarrow f) - \Gamma(\bar{P} \rightarrow \bar{f})}{\Gamma(P \rightarrow f) + \Gamma(\bar{P} \rightarrow \bar{f})}. \quad (14.1)$$

For $D^0 \rightarrow h^- h^+$ decays, this is expected to be of the order of 10^{-3} in the SM [9], and as such is sensitive to contributions from physics beyond the Standard Model.

The $K^- K^+$ and $\pi^- \pi^+$ final states transform into one another under the U -spin transformation, exchanging $d \rightleftharpoons s$ quarks. In some models the sign of $A_{CP}(D^0 \rightarrow K^- K^+)$ is predicted to be opposite to that of $A_{CP}(D^0 \rightarrow \pi^- \pi^+)$ through U -spin symmetry arguments [9], and so the difference between the two is particularly sensitive to direct CP violation

$$\Delta A_{CP} = A_{CP}(D^0 \rightarrow K^- K^+) - A_{CP}(D^0 \rightarrow \pi^- \pi^+). \quad (14.2)$$

After an initial measurement of ΔA_{CP} providing evidence of CP violation in charm decays [88, 89], generating much interest in the theory

community [89], more precise measurements of ΔA_{CP} did not confirm the initial evidence [90, 91]. The current world average is consistent with the hypothesis of no direct CP violation in D^0 decays [92], being¹

$$\Delta A_{CP} = (-0.252 \pm 0.104) \%. \quad (14.3)$$

This is to be compared to theoretical predictions which are uniformly below the percent level, but in which large theoretical uncertainties prevent a consistent number.

Charm baryons, the lightest of which is the Λ_c^+ (udc), are an additional system in which to search for CP violation. The first evidence for CP violation in baryon decays was found only recently [8], using the decays of the lightest beauty baryon, the Λ_b^0 (udb). An analogous measurement of ΔA_{CP} using Λ_c^+ decays, to that made with D^0 decays, is searching for direct CP violation in the decays $\Lambda_c^+ \rightarrow pK^-K^+$ and $\Lambda_c^+ \rightarrow p\pi^-\pi^+$. The dynamics of these SCS decays, which can only be fully parameterised by a five-dimensional phase space [94], are currently poorly understood [12, 95], and so any experimental input can be useful to guide and constrain the theory. The previous generation of heavy flavour collider experiments, such as Belle and BaBar, collected small samples of Λ_c^+ [83], whereas the large charm cross-section at LHCb enables CP violations measurements at the percent level or lower.

The large sample of b hadron decays collected by LHCb can also be used to study CP violation in Λ_c^+ decays. The long Λ_b^0 lifetime, around seven times greater than that of the Λ_c^+ , can be exploited to provide efficient background rejection. The semileptonic decay chain $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-X$ enables the usage of the LHCb muon triggers, described in Section 4.2, requiring a single, high p_T muon whose trajectory is significantly displaced from the PV. Using measurements of the $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-X$ and $\Lambda_c^+ \rightarrow ph^-h^+$ branching fractions [12, 96], and measurements of the Λ_b^0 cross-sections in pp collisions [97], and assuming a total reconstruction and selection efficiency of around 1 %, LHCb should have collected of the order of 10^4 $p\pi^-\pi^+$ decays

¹The world average is not strictly of ΔA_{CP} , but of $\Delta a_{CP}^{\text{dir}}$, which includes effects in *indirect* CP violation, which arise to the possibility of CP violation in D^0 - $\bar{D}^0$ mixing [93].

and 10^3 pK^-K^+ decays during Run 1. The statistical uncertainty on a measurement of ΔA_{CP} will then be dominated by the pK^-K^+ sample size, having a precision on the order of 1 %.

The analysis described in this part measures ΔA_{CP} using $\Lambda_c^+ \rightarrow pK^-K^+$ and $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ decays, reconstructed in association with a high- p_T muon assumed to originate from the decay of a Λ_b^0 . The data were collected with the LHCb detector in 2011 and 2012, corresponding to an integrated luminosity of 3 fb^{-1} . An overview of the analysis will be given in the remainder of this chapter.

14.1 Analysis overview

Rather than measuring the relative difference of decay rates in Equation (14.1), it is simpler to measure the asymmetry in the observed yields N

$$A_{CP}^{\text{Raw}}(f) = \frac{N(f) - N(\bar{f})}{N(f) + N(\bar{f})}, \quad (14.4)$$

where here, and throughout this part, the origin of the final state f from a Λ_c^+ decay is implicit.² This is not necessarily equal to the decay asymmetry in Equation (14.1) as it may contain contributions from associated production and detection asymmetries. The complete formalism and experimental treatment of this will be discussed in Chapter 15.

Any theoretical computation of A_{CP} or ΔA_{CP} for $\Lambda_c^+ \rightarrow ph^-h^+$ decays requires some model for the five-dimensional ph^-h^+ phase space, as it may be that the strength of CP violation varies across it. The selection and reconstruction of the Λ_c^+ candidates may sculpt this distribution, and so the experimental efficiency as a function of the phase space must be computed. The phase space definition is given Section 15.1.

After correcting for any possible background asymmetries with a weighting procedure, and for the phase space efficiencies computed using simulation, the Λ_c^+ and Λ_c^- yields are measured simultaneously for each mode with χ^2 fits. The fits are performed in such a way that the true value of A_{CP}^{Raw}

²It is assumed that charge is conserved, so a final state $\bar{f}$ implies an initial state Λ_c^- .

is hidden to avoid experimental bias, and as such the fits are said to be ‘blinded’.

The description of the data by the fits, any remaining background asymmetries, and Λ_c^+ decays originating from sources other than Λ_b^0 decays are considered in the context of systematic uncertainties. Studies are performed, and deviations from the nominal, measured value of ΔA_{CP} that cannot be corrected for will be assigned as systematic uncertainties on the measurement.

Part structure

The contributions of the various production and detection asymmetries to A_{CP}^{Raw} will be derived and enumerated in Chapter 15. The parameterisation of the five-dimensional phase space will also be given, which is used when constructing the phase space efficiency model. The collision data and the simulation used in the analysis are described in Chapter 16, and the reconstruction and selection of the real data will be presented in Chapter 17. The extraction of the signal yields and of the kinematic distributions of the signal from the fully selected data is presented in Chapter 18, and then those signal distributions are equalised between the two Λ_c^+ decay modes, discussed in Chapter 19, to remove background asymmetries such as the Λ_b^0 production asymmetry. The resulting kinematic weights are combined with reconstruction and selection efficiencies computed as a function of ph^-h^+ phase space, whose method of computation is shown in Chapter 20, after which the combined weights enter the blinded χ^2 fits that are used to measure A_{CP}^{Raw} , described in Chapter 21. The combination of $A_{CP}^{\text{Raw}}(pK^-K^+)$ and $A_{CP}^{\text{Raw}}(p\pi^-\pi^+)$ to measure ΔA_{CP} is given in Chapter 22. Methods for, and the results of, various systematic studies are given in Chapter 23. Finally, a summary is presented in Chapter 24.

15 Formalism

The measurement of A_{CP}^{Raw} , as defined in Equation (14.4), can be contaminated by production and detection asymmetries. By reconstructing $\Lambda_c^+ \rightarrow pK^-K^+$ and $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ decays from $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-X$ decays, A_{CP}^{Raw} can include the effects of the $\Lambda_b^0/\bar{\Lambda}_b^0$ production asymmetry, the μ^-/μ^+ detection asymmetry, and the detection asymmetry of the Λ_c^+ final state $f/\bar{f}$.

The Λ_b^0 production asymmetry is predicted to be non-zero due to the pp collision environment [98], and indeed LHCb has found evidence of such an asymmetry as a function of Λ_b^0 rapidity [97]. The two detection asymmetries are also expected to be non-zero, due to the known differences between the hadron/anti-hadron and muon/anti-muon cross-sections with matter [12]. The final state detection asymmetry can be broken down into two terms: one due to the detection asymmetry of the proton; and another due to the detection asymmetry of the K^-K^+ pair and the $\pi^-\pi^+$ pair. If the meson kinematics are equal within a given mode, for example if the K^- kinematics are identical to those of the K^+ in the pK^-K^+ data, then the final state asymmetry reduces to the proton detection asymmetry.

Both the Λ_b^0 production and the muon detection asymmetries should be independent of the Λ_c^+ final state, given the same Λ_c^+ kinematics. Different acceptance efficiencies between pK^-K^+ and $p\pi^-\pi^+$ can create unequal Λ_c^+ kinematics, but the kinematics can be weighted to equalise them between the modes.

It will now be shown how the background asymmetries entering A_{CP}^{Raw} can be removed in the difference ΔA_{CP} . The number of reconstructed $\Lambda_c^+ \rightarrow f$

decays can be expressed as the product of several effective probabilities

$$N(f) = \mathcal{P}(\Lambda_b^0) \cdot \Gamma(\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X) \cdot \Gamma(\Lambda_c^+ \rightarrow f) \cdot \varepsilon(\mu^-) \cdot \varepsilon(f), \quad (15.1)$$

where $\mathcal{P}(\Lambda_b^0)$ is the probability of producing a Λ_b^0 baryon given a pp collision, and $\varepsilon(\mu^-)$ and $\varepsilon(f)$ are the muon and Λ_c^+ final-state detection efficiencies. A similar expression exists for $N(\bar{f})$, where the particles are replaced by their charge conjugates. It is assumed that the $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X$ decay is CP -symmetric, but that all other factors may not be, for example $\varepsilon(\mu^-) \neq \varepsilon(\mu^+)$.

To express A_{CP}^{Raw} in terms of its component asymmetries, the notation of a general asymmetry parameter X is defined, which describes the asymmetry between two quantities x and $\bar{x}$

$$X = \frac{x - \bar{x}}{x + \bar{x}}. \quad (15.2)$$

This can be rearranged as

$$x = \frac{1}{2}(x + \bar{x})(1 + X), \text{ and} \quad (15.3)$$

$$\bar{x} = \frac{1}{2}(x + \bar{x})(1 - X). \quad (15.4)$$

An asymmetry parameter like that in Equation (15.2) can be defined for each of the relevant terms in Equation (15.1)

$$\begin{aligned} A_P^{\Lambda_b^0}(f) &= \frac{\mathcal{P}(\Lambda_b^0) - \mathcal{P}(\bar{\Lambda}_b^0)}{\mathcal{P}(\Lambda_b^0) + \mathcal{P}(\bar{\Lambda}_b^0)}, \\ A_D^{\mu^-}(f) &= \frac{\varepsilon(\mu^-) - \varepsilon(\mu^+)}{\varepsilon(\mu^-) + \varepsilon(\mu^+)}, \text{ and} \\ A_D^f(f) &= \frac{\varepsilon(f) - \varepsilon(\bar{f})}{\varepsilon(f) + \varepsilon(\bar{f})}, \end{aligned}$$

and the asymmetry in $\Gamma(\Lambda_c^+ \rightarrow f)$ is A_{CP} as in Equation (14.1). Each parameter is, at least implicitly, dependent on the detected Λ_c^+ final state, as each parameter can vary in quantities that may also vary between Λ_c^+ decay modes. Substituting Equation 15.1 into Equation 14.4

$$A_{CP}^{\text{Raw}}(f) = \frac{\mathcal{P}(\Lambda_b^0)\Gamma(f)\varepsilon(\mu^-)\varepsilon(f) - \mathcal{P}(\bar{\Lambda}_b^0)\Gamma(\bar{f})\varepsilon(\mu^+)\varepsilon(\bar{f})}{\mathcal{P}(\Lambda_b^0)\Gamma(f)\varepsilon(\mu^-)\varepsilon(f) + \mathcal{P}(\bar{\Lambda}_b^0)\Gamma(\bar{f})\varepsilon(\mu^+)\varepsilon(\bar{f})},$$

and then substituting each quantity for its equivalent form as in Equations (15.3) and (15.4), all factors of $1/2$ and all factors of the form $(x - \bar{x})$ cancel, leaving

$$A_{CP}^{\text{Raw}}(f) = \frac{Y}{Z}. \quad (15.5)$$

where (dropping the final state parameter temporarily for compactness)

$$\begin{aligned} Y = & A_P^{\Lambda_b^0} A_D^{\mu^-} A_D^f + A_P^{\Lambda_b^0} A_D^{\mu^-} A_{CP} + A_P^{\Lambda_b^0} A_D^f A_{CP} + A_D^{\mu^-} A_D^f A_{CP} \\ & + A_P^{\Lambda_b^0} + A_D^{\mu^-} + A_D^f + A_{CP}, \end{aligned} \quad (15.6)$$

and

$$\begin{aligned} Z = & 1 + A_P^{\Lambda_b^0} A_D^{\mu^-} + A_P^{\Lambda_b^0} A_D^f + A_P^{\Lambda_b^0} A_{CP} + A_D^{\mu^-} A_D^f + A_D^{\mu^-} A_{CP} \\ & + A_D^f A_{CP} + A_P^{\Lambda_b^0} A_D^{\mu^-} A_D^f A_{CP}. \end{aligned} \quad (15.7)$$

Assuming that the individual asymmetries are small, of the order of 1 %, the product of two or more asymmetries is negligible with respect to the leading order, and so

$$A_{CP}^{\text{Raw}}(f) \approx A_{CP}(f) + A_P^{\Lambda_b^0}(f) + A_D^{\mu^-}(f) + A_D^f(f). \quad (15.8)$$

By assuming that these background asymmetries are mode-independent, that is to say $A_D(f) = A_D(g)$ and $A_P(f) = A_P(g)$, the difference between A_{CP}^{Raw} for the pK^-K^+ and $p\pi^-\pi^+$ will only have contributions from A_{CP}

$$\begin{aligned} \Delta A_{CP} &= A_{CP}^{\text{Raw}}(pK^-K^+) - A_{CP}^{\text{Raw}}(p\pi^-\pi^+), \\ &\approx A_{CP}(pK^-K^+) - A_{CP}(p\pi^-\pi^+). \end{aligned} \quad (15.9)$$

The assumption that the production and detection asymmetries in Equation (15.8) are mode independent is not true in general. This can be seen by first making a weaker assumption that the background asymmetries are dependent only on the kinematics of the representative particles. Different final states will in general have different acceptance, reconstruction and selection efficiencies as a function of Λ_c^+ kinematics. The kinematics of the Λ_c^+ are correlated to those of the muon and the Λ_b^0 , and so two samples with

different Λ_c^+ kinematics will likely also have different Λ_b^0 and muon kinematics. Hence, there can still be net production and detection asymmetries in ΔA_{CP} .

The assumption that the background asymmetries depend only on particle kinematics is not unreasonable. The production asymmetry, for example, is a difference in cross-sections, which are usually parameterised by the kinematics of the produced particle (p_T and either η or rapidity). Similarly, a detection asymmetry describes the differences of material interactions between matter and antimatter, and these are dependent on the momentum of the particle in question and, assuming a non-uniform material distribution, its flight path.

As it is not given that the Λ_c^+ kinematics are the same between $\Lambda_c^+ \rightarrow pK^-K^+$ and $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ decays in the data, the data can be weighted to equalise the Λ_c^+ kinematics. The background asymmetries A_P and A_D^μ will then cancel in the difference ΔA_{CP} . The proton kinematics will not necessarily agree after such a weighting, as the energy release in the pK^-K^+ and $p\pi^-\pi^+$ decays is different. These considerations govern the analysis strategy: measure the number of Λ_c^+ candidates in the pK^-K^+ and $p\pi^-\pi^+$ samples after weighting them such that the Λ_c^+ kinematics look alike, such that the Λ_b^0 and μ^- kinematics also agree. Additional weighting may be required to equalise the proton kinematics.

15.1 Decay phase space

The phase space of a decay is the set of variables which fully parameterises all possible dynamics. Although a three-body decay requires 12 parameters to describe the kinematics, under the assumptions of four-momentum conservation and that the masses of the three decay products are known, the number of free parameters is only five. When all particles involved have zero spin, the distribution of the momenta of the decay products is isotropic, such that there are only 2 degrees of freedom. These are usually taken to be two child-pair squared masses, which can be visualised as a Dalitz

plot. In the case of $\Lambda_c^+ \rightarrow ph^-h^+$, the spin $\frac{1}{2}$ of the proton means that the three-body system is no longer rotationally symmetric, and the full five degrees of freedom are required to describe the phase space.

For a meaningful comparison of the measurement of ΔA_{CP} with theoretical predictions, which are not presently available, the efficiency of the Λ_c^+ selection across the five-dimensional phase space must be known. The definition of “selection” includes the effects of the LHCb acceptance and the trigger, stripping, and offline requirements. The efficiency model can either be provided to theorists so that they can apply the same efficiencies to their phase space models, or it can be used to correct the data before the ΔA_{CP} measurement is made.

The five dimensions of the phase space are defined here in a similar way to the $\Lambda_c^+ \rightarrow pK^-\pi^+$ amplitude analysis performed by the E791 collaboration [94]. This defines two child-pair squared masses, of the proton and oppositely-charged child $m^2(ph^-)$ and of the two opposite-sign pseudo-scalars $m^2(h^-h^+)$, and three decay angles. The child-pair squared masses are invariant under Lorentz transformations, but the decay angles are not and so a definition of the frame in which they are computed is required.

The E791 analysis defines a coordinate system in the Λ_c^+ rest frame. The z -axis, also called the quantisation axis or polarisation axis $\hat{\mathbf{P}}_{\Lambda_c^+}$, is perpendicular to the plane of production

$$z = \hat{\mathbf{P}}_{\Lambda_c^+} = \hat{p}_{\text{Beam}} \times \hat{p}_{\Lambda_c^+}, \quad (15.10)$$

where $\hat{p}_{\text{Beam}}$ is the direction of the beam¹ and $\hat{p}_{\Lambda_c^+}$ is the direction of the Λ_c^+ measured in the laboratory frame. The x -axis of the Λ_c^+ rest frame is $\hat{p}_{\Lambda_c^+}$. As the Λ_c^+ candidates in this analysis are not produced directly from the pp collision but in the decays of Λ_b^0 baryons, here the ‘beam direction’ is defined as the direction of the Λ_b^0 momentum vector, which is equal to the direction vector pointing from the pp primary vertex v_{pp} to the $\Lambda_c^+\mu^-$ vertex $v_{\Lambda_c^+\mu^-}$

$$\hat{p}_{\text{Beam}} = \hat{p}_{\Lambda_b^0} = v_{\Lambda_c^+\mu^-} - v_{pp}. \quad (15.11)$$

¹E791 was a fixed-target experiment, colliding a 500 GeV/c pion beam with metal foils.

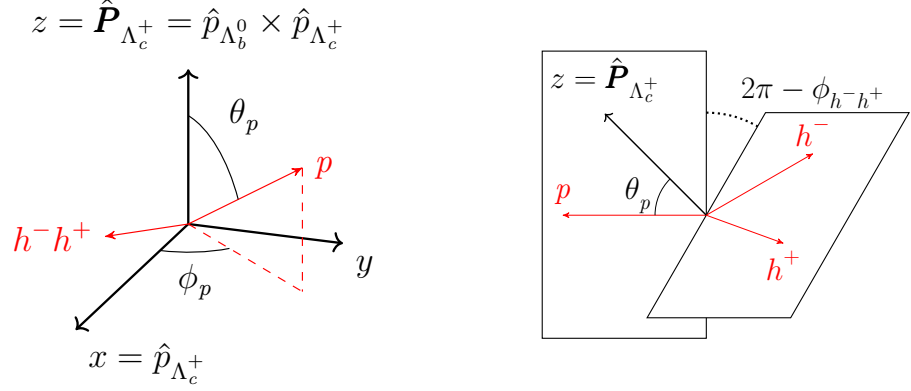


Figure 15.1: Definition of inertial reference frame axes and $\Lambda_c^+ \rightarrow ph^-h^+$ phase space decay angles. Adapted from Figure 1 in the E791 $\Lambda_c^+ \rightarrow pK^-\pi^+$ amplitude analysis paper [94].

With this coordinate system and inertial frame, the three decay angles are defined as:

1. The angle θ_p between the proton momentum vector and the z -axis;
2. The angle ϕ_p between the proton momentum vector and the x -axis;
and
3. The angle $\phi_{h^-h^+}$ between the plane containing the proton momentum vector and the z -axis and the plane containing the two pseudo-scalar meson momentum vectors.

These definitions are illustrated in Figure 15.1

The distributions of the five phase space variables will be presented in Chapter 20, along with the evaluation of the efficiency as a function of the position in phase space.

16 Datasets

This analysis uses all data collected in 2011 and 2012. As the beam conditions changed between the two years, from $\sqrt{s} = 7 \text{ TeV}$ to 8 TeV , the analysis is performed separately on the two sub-samples. The sub-samples are further split into magnet up and magnet down datasets, as the particle detection asymmetries change with the magnet polarity [99]. The individual measurements made using these sub-samples are later combined in an average. The total integrated luminosity is $(2979.0 \pm 40.0) \text{ pb}^{-1}$, the breakdown of which for the different running conditions is given in Table 16.1. This chapter will describe the generalities of the data, such as the processing workflow and application of calibrations, as well as the simulation that will be used.

The data are processed through the offline reconstruction and the stripping, as described in Section 4.3. Two stripping lines select the $\Lambda_c^+ \rightarrow pK^-K^+$ and $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ decays of interest in association with a muon candidate, and a $\Lambda_c^+\mu^-$ vertex is formed and treated as a Λ_b^0 . The candidates passing the stripping selections are required to have been used in specific trigger decisions. As combinatorial background is the dominant component in the data at this stage, as shown in Figure 16.1, candidates passing the stripping and trigger requirements are further filtered offline. Details of the trigger, stripping, and offline selection will be given in Chapter 17.

To improve the Λ_c^+ mass resolution in the datasets, and the consistency between them, two algorithms are run on the data offline. The first of these is the momentum scale calibration. This corrects track momentum by a linear factor α , which varies with running conditions and is found by comparing mass measurements made with $J/\psi \rightarrow \mu^+\mu^-$ and $B^+ \rightarrow J/\psi K^+$ decays to the corresponding world averages [30]. The second algorithm is DECAYTREEFIT-

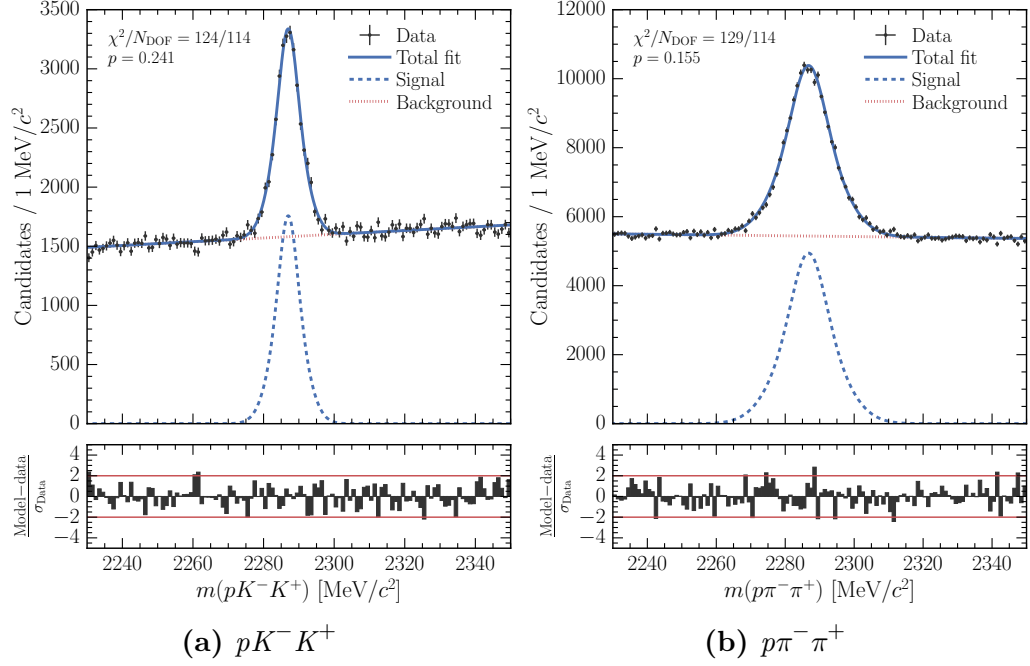


Figure 16.1: Fits to the Λ_c^+ mass spectrum in the 2012 magnet down dataset for pK^-K^+ (a) and $p\pi^-\pi^+$ (b). Only the stripping and trigger selection is applied. The fit that is overlaid is described in Chapter 18.

TER [100]. This fits the entire $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X$ -with- $\Lambda_c^+ \rightarrow ph^-h^+$ decay chain in a single step, rather than the back-to-front reconstruction used in the stripping. The advantage of such a method is that all the information known about the decay can influence the fit parameters, whereas the stripping reconstruction can only propagate information in one direction. Such a global treatment of information is particularly useful when applying constraints in the fit, such as requiring the mass of a vertex combination to be exactly some value, or that the head of the decay originated from the PV. In this analysis, DECAYTREEFITTER is applied to the Λ_b^0 decay with the constraint that the ph^-h^+ vertex has the nominal Λ_c^+ mass of 2286 MeV/c² [12]. This improves the resolution of the Λ_c^+ phase space variables, and restricts the phase space to physical values. Unless stated otherwise, kinematic quantities used in this analysis are those computed by the DECAYTREEFITTER algorithm, with the notable exceptions of the quantities used in the trigger and stripping selections, and the ph^-h^+ invariant mass entering the mass fit.

16.1 Monte Carlo

The only information made available for offline analysis is that passing the stripping. The simulation is used to infer properties of the data before that stage. There are three different sets of Monte Carlo (MC) data that are used in this analysis.

Full detector simulation

The full LHCb Monte Carlo is used to obtain distributions of the phase space variables after the LHCb acceptance, stripping, trigger, and offline requirements. The general procedure of the LHCb MC generation is described in Section 7.1, and the acceptance requirements are identical to those given in Equation (7.2).

The samples contain around 2.5×10^6 generated Λ_b^0 decays per Λ_c^+ final state per magnet polarity, simulated using the most representative data-taking conditions for 2012. These conditions correspond to an average number of pp interactions per bunch crossing of $\nu = 2.5$, a beam energy of 4 TeV, and a trigger configuration representative of the average trigger conditions used in 2012. Each Λ_b^0 is forced to decay to the $\Lambda_c^+ \mu^- \bar{\nu}_\mu$ final state, and the Λ_c^+ is forced to decay to the appropriate $ph^- h^+$ final state with no intermediate resonances. The simulated particles are reconstructed with the same reconstruction software as the real data, and are then passed through a near-identical version of the stripping, with the only difference being the omission of any PID selection, as the related variables are known to be poorly modelled in the MC. The DECAYTREEFITTER algorithm is run on the candidates passing the stripping, but the momentum scale calibration is not applied as the conditions in the MC are stable.

Generator-level simulation

The Monte Carlo samples described in Section 16.1 have the LHCb acceptance requirements imposed upon them, as defined in Equation (7.2). These requirements do not necessarily have a flat efficiency across the $\Lambda_c^+ \rightarrow ph^- h^+$

Table 16.1: *Integrated luminosity for each data sample used in the analysis.*

Year	Polarity	Integrated luminosity (pb^{-1})
2011	Up	422 ± 7
2011	Down	564 ± 10
2012	Up	1001 ± 12
2012	Down	993 ± 12

phase space, and so additional samples are required to study these effects. A sample of 5×10^5 $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- \bar{\nu}_\mu$ decays are generated for each $\Lambda_c^+ \rightarrow ph^-h^+$ mode for each 2012 magnet polarity. These samples are not propagated through the detector simulation, and so only generator-level, or ‘truth level’, information is available.

17 Selection

The Λ_c^+ charm baryon has a lifetime of 0.2 ps, half that of the D^0 and eight times less than that of the B^0 . Indeed, it is one of the shortest-lived weakly decaying heavy flavour hadrons, and as such presents particular challenges when a clean, efficient selection is desired as the secondary decay vertex is not far displaced from the PV. Although the $b\bar{b}$ cross-section is of the order of 100 times less than the $c\bar{c}$ cross-section [10, 11], reconstructing the decay chain $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X$ allows for the exploitation of the long Λ_b^0 lifetime and the highly efficient muon and topological b decay triggers [41, 42]. This analysis selects secondary Λ_c^+ candidates by association with a high- p_T muon displaced from the PV.

17.1 Stripping and trigger

The selections in the stripping lines were not tuned specifically for this analysis, instead serving as general-purpose requirements suitable for a range of studies that use $\Lambda_c^+ \rightarrow ph^- h^+$ decays originating from semileptonic Λ_b^0 decays. As such, they are similar to the selections of other b decays made at LHCb. They exploit the long decay times of b hadrons, such that they fly a significant, measurable distance in the VELO, and the high p_T and large impact parameter of the child vertices and tracks. Additional requirements are similar in nature to those made for charm hadrons, such as track and vertex fit quality and track χ_{IP}^2 , which are described in Chapter 8.

The selection criteria for final state particles entering the stripping lines, namely protons, charged kaons and pions, and muons, is given in Table 17.1. One variable that has not been discussed previously is p_{Ghost} , the ghost

Table 17.1: *Selection of tracks, with associated particle hypotheses, used in the Λ_b^0 and Λ_c^+ stripping selection. Cuts listed under “All” are applied to all tracks.*

Particle	Variable	Cut value
All	p_T	$> 250 \text{ MeV}/c$
	p	$> 2 \text{ GeV}/c$
	Track χ^2/N_{DOF}	< 3
	p_{Ghost}	< 0.5
Muons	p_T	$> 800 \text{ MeV}/c$
	p	$> 3 \text{ GeV}/c$
	χ_{IP}^2	> 4
	$\text{DLL}_{\mu-\pi}$	> 0
Pions	χ_{IP}^2	> 9
	$\text{DLL}_{K-\pi}$	< 4
Kaons	χ_{IP}^2	> 9
	$\text{DLL}_{K-\pi}$	> 4
Protons	χ_{IP}^2	> 9
	$\text{DLL}_{p-\pi}$	> 4
	DLL_{p-K}	$> 1 \times 10^{-10}$

probability, which is the response of an artificial neural network (ANN) trained to discriminate between real and fake (ghost) tracks [101]. The ANN response is transformed in such a way as to be within the range $[0, 1]$, and that the ghost rejection efficiency in simulation varies linearly from zero to 100 % in that range.

The selected protons, pions, and kaons are combined as Λ_c^+ candidate vertices, which are then combined with muon candidates to form Λ_b^0 candidates. The requirements on the combinations and vertices are given in Table 17.2. As in Chapter 8, the selection of composite particles proceeds in two stages: firstly by imposing requirements on the union of the inputs; and secondly on the vertex formed after the fit.

Trigger requirements are made after the Λ_b^0 candidate is created and selected in the stripping lines. They are driven by the powerful discrimination provided by the muon from the Λ_b^0 . The muon candidate is required to have

Table 17.2: *Stripping selection of particle combinations, vertices, and events.*

Object	Variable	Cut value
Event	$N_{\text{Long tracks}}$	< 250
	N_{PV}	> 0
ph^-h^+	$ m - m_{\text{PDG}}(\Lambda_c^+) $	$< 100 \text{ MeV}/c^2$
	$\sum_{\text{Children}} p_{\text{T}}$	$> 1800 \text{ MeV}/c$
	DOCA χ^2	< 20
Λ_c^+ vertex	$ m - m_{\text{PDG}}(\Lambda_c^+) $	$< 80 \text{ MeV}/c^2$
	$\sum_{\text{Children}} p_{\text{T}}$	$> 1800 \text{ MeV}/c$
	Vertex χ^2/N_{DOF}	< 6
	Vertex distance χ^2	> 100
$\Lambda_c^+\mu^-$	m	$< 6200 \text{ MeV}/c^2$
Λ_b^0 vertex	m	$2500 < m < 6000 \text{ MeV}/c^2$
	Vertex χ^2/N_{DOF}	< 6
	DIRA	> 0.999
	$z_{\Lambda_c^+} - z_{\Lambda_b^0}$	$> 0 \text{ mm}$

fired the L0 and HLT1 single muon lines, and at HLT2 the union of the tracks comprising the Λ_b^0 and Λ_c^+ candidates are required to have fired the topological semileptonic b lines, as described in Section 4.2. Distributions of the pK^-K^+ and $p\pi^-\pi^+$ mass in the 2012 magnet down dataset after the stripping and trigger requirements are given in Figure 16.1, as an illustration of the signal purity at this stage.

17.2 Offline

A mass window requirement is imposed on the Λ_c^+ candidates, which is equal to the range used in the mass fits (described in Chapter 18)

$$2230 < m(\Lambda_c^+) < 2350 \text{ GeV}. \quad (17.1)$$

This is the three-body ph^-h^+ invariant mass computed in the vertex fit, rather than that from the decay tree fit. The decay tree fit is required to

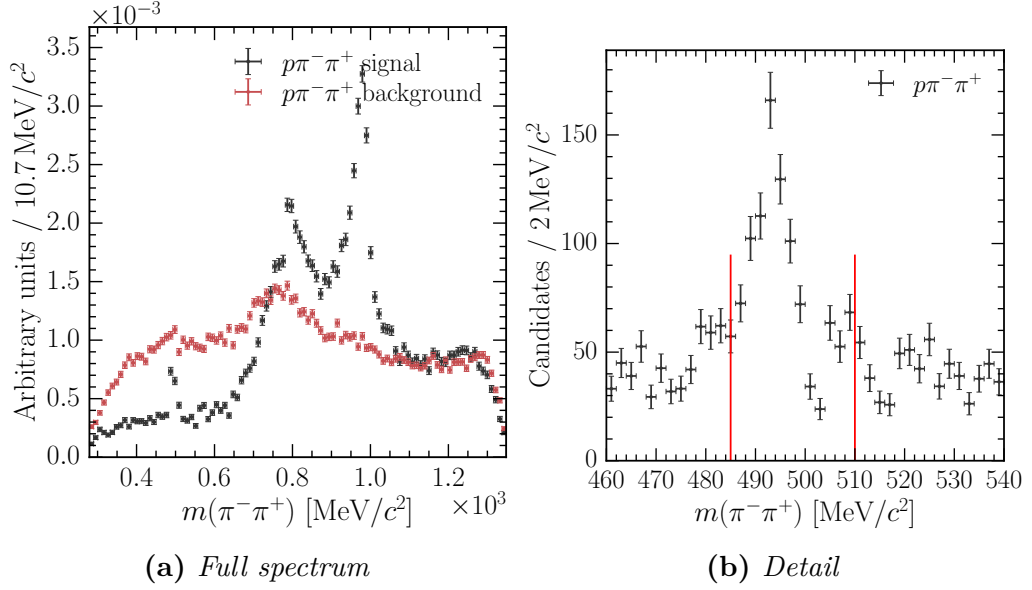


Figure 17.1: Full (a) and zoomed (b) mass spectra of the $\pi^+\pi^-$ pair in the $p\pi^-\pi^+$ decay mode, in the magnet up 2012 data. The background contributions in the black spectra have been statistically subtracted, whereas the red spectrum has the signal component subtracted. The red lines indicate the boundaries of the requirement designed to remove the pK_S^0 component of the sample. Also visible are the $\rho(770)/\omega(782)$ and $f_0(980)$ resonances.

have converged.

The $p\pi^-\pi^+$ mode contains a number of pK_S^0 and $\Lambda^0\pi^+$ candidates, as shown in Figures 17.1 and 17.2. The candidates are removed by applying an exclusion window 40 MeV wide centred near the nominal K_S^0 mass [12]

$$485 < m(\pi^+\pi^-) < 510 \text{ MeV}, \quad (17.2)$$

and an exclusion window 10 MeV wide centred near the nominal Λ^0 mass [12]

$$1100 < m(p\pi^-) < 1120 \text{ MeV}. \quad (17.3)$$

The largest source of combinatorial background is due to $K \rightarrow p$ and $\pi \rightarrow p$ misidentification, as pions and kaons are produced much more numerous than protons. The PID requirements on the proton, kaon, and pion candidates are tightened offline. The ProbNN family of variables is used, which contains one ProbNN variable for each charged particle

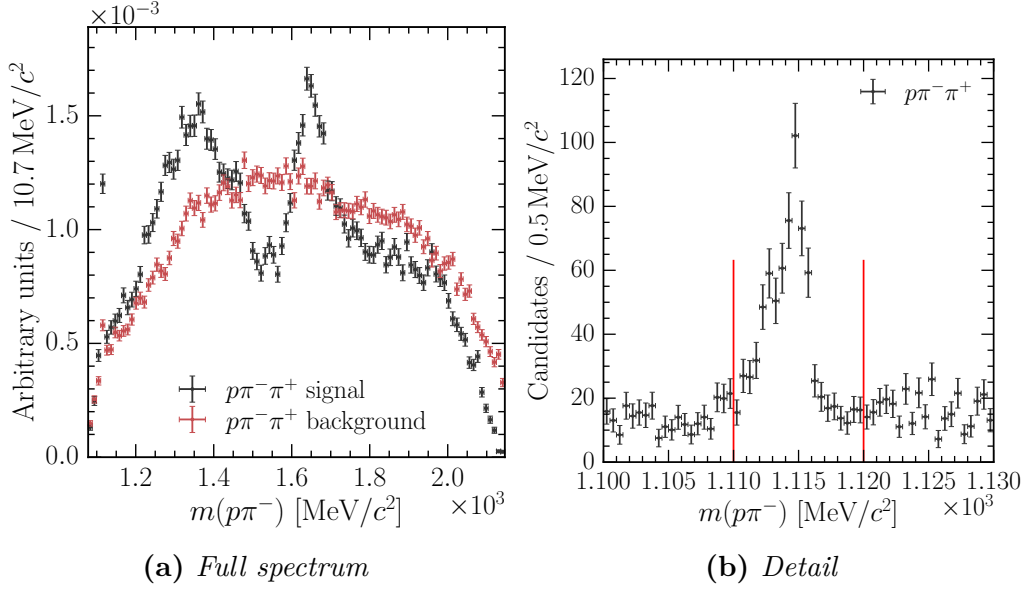


Figure 17.2: Full (a) and zoomed (b) mass spectra of the $p\pi^-$ pair in the $p\pi^-\pi^+$ decay mode, in the magnet up 2012 data. The background contributions in the black spectra have been statistically subtracted, whereas the red spectrum has the signal component subtracted. The red lines indicate the boundaries of the requirement designed to remove the $\Lambda^0\pi^+$ component of the sample. The additional features in the $p\pi^-$ spectrum have not been identified.

hypothesis ($p/\bar{p}$, $K^\pm$, $\pi^\pm$, $\mu^\pm$, $e^\pm$), and is the response of an ANN trained to discriminate each hypothesis from all others. Input variables to the network include the DLL PID variables described in Section 4.1, the track p and p_T , and the total track fit quality and that within each tracking sub-detector. The ANN is trained on a set of truth-matched tracks from a sample of MC generated under 2012 running conditions, and is scaled such that the signal probability is uniform in the range $\text{ProbNN} \in [0, 1]$.

To define a starting point for the ProbNN cut values, a simultaneous optimisation of all PID requirements was performed by maximising the signal significance σ , defined as

$$\sigma = \frac{N_{\text{Sig.}}}{\sqrt{N_{\text{Sig.}} + N_{\text{Bkg.}}}}, \quad (17.4)$$

where $N_{\text{Sig.}}$ and $N_{\text{Bkg.}}$ are the signal and background yields in the sample,

Table 17.3: *Particle identification requirements made on the Λ_c^+ decay products.*

Particle	Variable	Requirement
$p/\bar{p}$	ProbNNp	> 0.3
$K^\pm$	ProbNNk	> 0.3
$\pi^\pm$	ProbNNpi	> 0.3

measured using the fits described in Chapter 18. The optimisation showed that the most powerful discriminants between signal and background were the proton PID information, as expected, and the PID information of the kaon or pion with the same charge as the proton. Applying a selection optimised for the signal significance would then result in the PID requirements being asymmetric in the h^-h^+ meson pair. As it is known that the PID efficiency for a given requirement on a given particle is dependent on the particle kinematics, as discussed in Section 10.4, and that the detection asymmetries also depend on the same kinematics, the same PID cut is applied to both mesons in the Λ_c^+ final state so as not to induce any additional detection asymmetries. The optimal value found for the cut on the meson with opposite sign to the proton was chosen as the value for the cut on the same-sign meson. The PID requirements made offline on the candidate protons, kaons, and pions are given in Table 17.3.

The ph^-h^+ invariant mass distributions after the weak decay vetoes, PID cuts, and DECAYTREEFITTER convergence requirements are given in Figure 17.3 for the 2012 magnet down dataset.

17.3 Background study

Although PID requirements are made on the proton candidate, it is possible that a meson decay will be misidentified as a Λ_c^+ decay. For example, a $D_s^+ \rightarrow K^- K^+ \pi^+$ decay could be misidentified as $\Lambda_c^+ \rightarrow p K^- K^+$ if the pion is assigned the proton mass hypothesis and passes the PID cuts. The presence of any such misidentification (or ‘mis-ID’) can add structures to

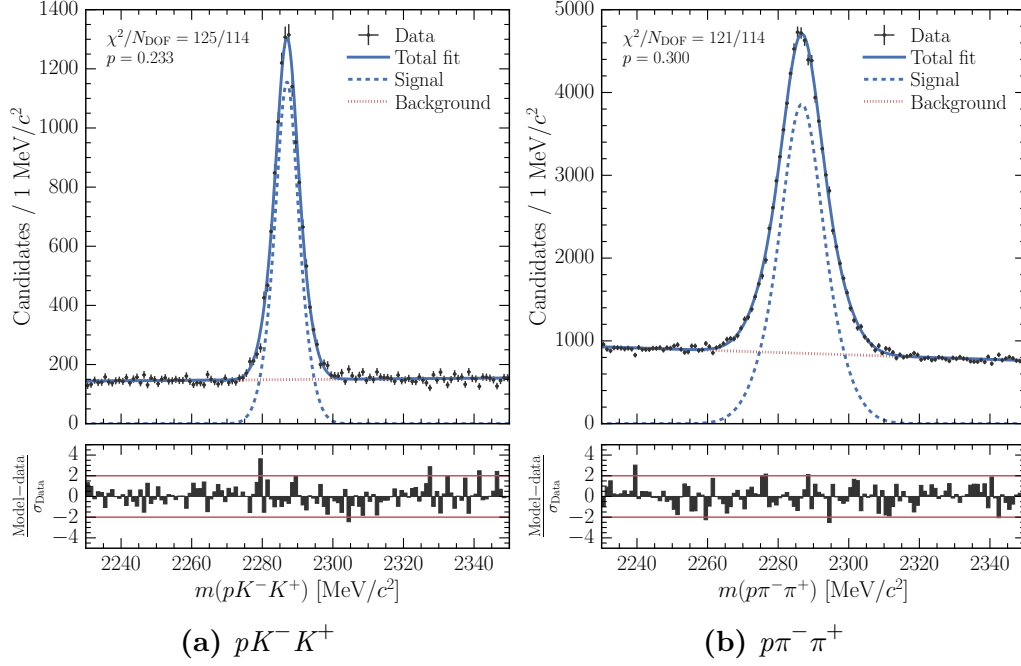


Figure 17.3: Fits to the Λ_c^+ mass spectrum in the 2012 magnet down dataset for $pK^- K^+$ (a) and $p\pi^- \pi^+$ (b). The stripping, trigger, and PID selections are applied, as well as the DECAYTREEFITTER convergence requirements, and the weak decay vetoes for the $p\pi^- \pi^+$ data. The fit that is overlaid is described in Chapter 18.

the Λ_c^+ mass spectrum, making the signal yield extraction more challenging. Such structures may peak near the nominal Λ_c^+ mass and be incorrectly counted as Λ_c^+ signal. Additional backgrounds can arise due to the partial reconstruction of Λ_c^+ decays, which would manifest as a ‘shoulder’ at values of the Λ_c^+ mass below the signal peak. As this feature is not seen in the $pK^- K^+$ or $p\pi^- \pi^+$ mass distributions, shown in Figure 17.3, it is considered insignificant and is not studied further.

One method of checking for the presence of misidentified decays is by inspecting three-body mass distributions for different mass hypothesis sets. For example, by assigning the pion mass to the proton candidate in the $pK^- K^+$ sample and inspecting the invariant mass of the $\pi^+ K^- K^+$ combination. A peak near the D_s^+ mass would indicate a contamination of D_s^+ candidates in the data.

A second way to look for meson backgrounds is to inspect the ‘momentum asymmetry’ parameter of the proton [102, 103]

$$\beta_p = \frac{-p_p + p_{h^+} + p_{h^-}}{p_p + p_{h^+} + p_{h^-}}, \quad (17.5)$$

where p_p is the momentum of the proton candidate and $p_{h^\pm}$ is the momentum of the $h^\pm$ candidate. Along with the mass of the Λ_c^+ , this fully parameterises the mass of the ph^-h^+ system under different hypotheses, and so one can infer the presence of meson backgrounds by inspecting the β_p - $m(ph^-h^+)$ plane. As the relationship between $m(ph^-h^+)$, β_p , and the three-body mass under the background hypothesis is known, contours of the expected mis-ID shape can be drawn on this plane and compared with the data.

Throughout this section, the variables shown are not computed using any information from the DECAYTREEFITTER algorithm described in Chapter 16.

Exchanging mass hypotheses

Figure 17.4 shows the three-body mass distributions in the 2012 magnet down pK^-K^+ data for several mass hypothesis sets, where each set does not contain the proton mass hypothesis. Figure 17.5 shows the same, but where each set does contain the proton mass hypothesis. The two figures show the charm meson and charm baryon background contributions. Figures 17.7 and 17.8 show the same but for the 2012 magnet down $p\pi^-\pi^+$ data.

The distributions from the pK^-K^+ data show contributions from $D^+ \rightarrow K^+K^-\pi^+$, $D_s^+ \rightarrow K^+K^-\pi^+$, and $\Lambda_c^+ \rightarrow pK^-\pi^+$ decays. The $p\pi^-\pi^+$ data show $D^+ \rightarrow K^+\pi^-\pi^+$, $D^+ \rightarrow K^+K^-\pi^+$, $D_s^+ \rightarrow K^+\pi^-\pi^+$, and $D_s^+ \rightarrow K^+K^-\pi^+$ contributions. Each of these backgrounds is removed by vetoing Λ_c^+ candidates whose wrong-mass value falls within 8 MeV of the nominal D^+ , D_s^+ , or Λ_c^+ mass [12].

A double misidentification of pK^-K^+ is present in the pK^-K^+ data, where the hypotheses of the nominal proton and nominal kaon with the same charge as the proton are swapped, visible in Figure 17.5b. This is not removed from the data, as the distribution of candidates that fall in

the signal region of the swapped hypothesis (K^+K^-p) have no irregular structures under the nominal hypothesis (pK^-K^+), as shown in Figure 17.6.

Momentum asymmetry

The momentum asymmetry of the proton is defined in Equation (17.5). Along with the mass of the Λ_c^+ under the nominal three-body mass hypothesis, it fully parameterises alternative three-body hypotheses where the hypothesis of the proton is changed. The $\beta_p-m(\Lambda_c^+)$ plane then provides an additional handle on misidentification backgrounds, in addition to the technique described in Section 17.3.

Figures 17.9 and 17.10 show the $\beta_p-m(\Lambda_c^+)$ plane with 2012 magnet down data. There are distinct contributions from $D^+ \rightarrow K^+K^-\pi^+$ and $D_s^+ \rightarrow K^+K^-\pi^+$ in the pK^-K^+ data, and from $D_s^+ \rightarrow K^+\pi^-\pi^+$ in the $p\pi^-\pi^+$ data. Also shown are the $\beta_p-m(\Lambda_c^+)$ planes after the misidentification vetoes described in Section 17.3 have been applied. No contributions beyond those given in Section 17.3 are seen.

17.4 Multiple candidates

It is possible, but unlikely, that two or more true $\Lambda_c^+ \rightarrow ph^-h^+$ decays are reconstructed in a single LHC bunch crossing. It is more likely that such multiple candidates are caused by unphysical means, such as cloned tracks [104]. In the case where an event has multiple Λ_c^+ candidates, one is selected at random for further analysis and the rest are discarded. The fractions of events containing multiple Λ_c^+ candidates are given in Tables 17.4 and 17.5.

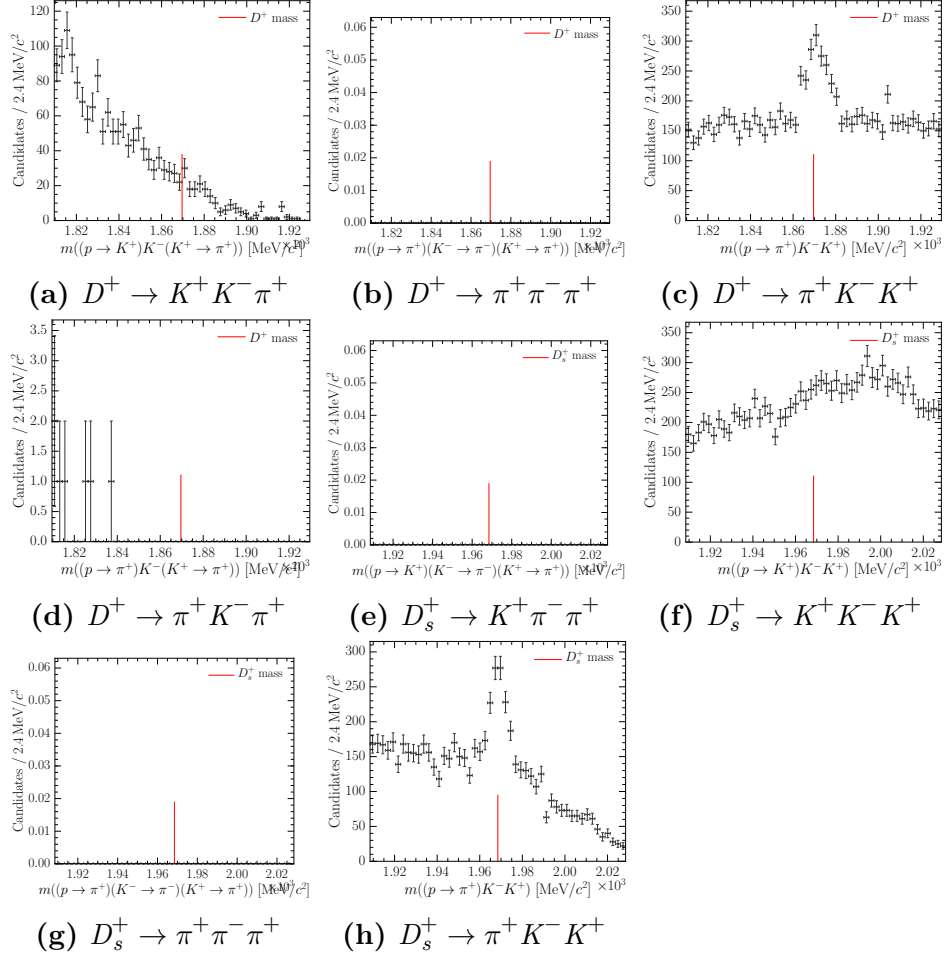


Figure 17.4: Wrong-mass distributions obtained when changing the mass hypotheses of the fully selected Λ_c^+ children in the pK^-K^+ mode, where no child is assigned the proton mass hypothesis. The x-axis on each sub-figure shows the substitutions that have been made. For example, the $(p \rightarrow \pi^+)(K^- \rightarrow \pi^-)K^+$ distribution is where the nominal proton candidate has been assigned the pion mass hypothesis, and the nominal kaon with opposite charge to the proton has been also assigned the pion hypothesis. The vertical red lines indicate the nominal mass of the misidentified particle under study. Only the 2012 magnet down data is shown.

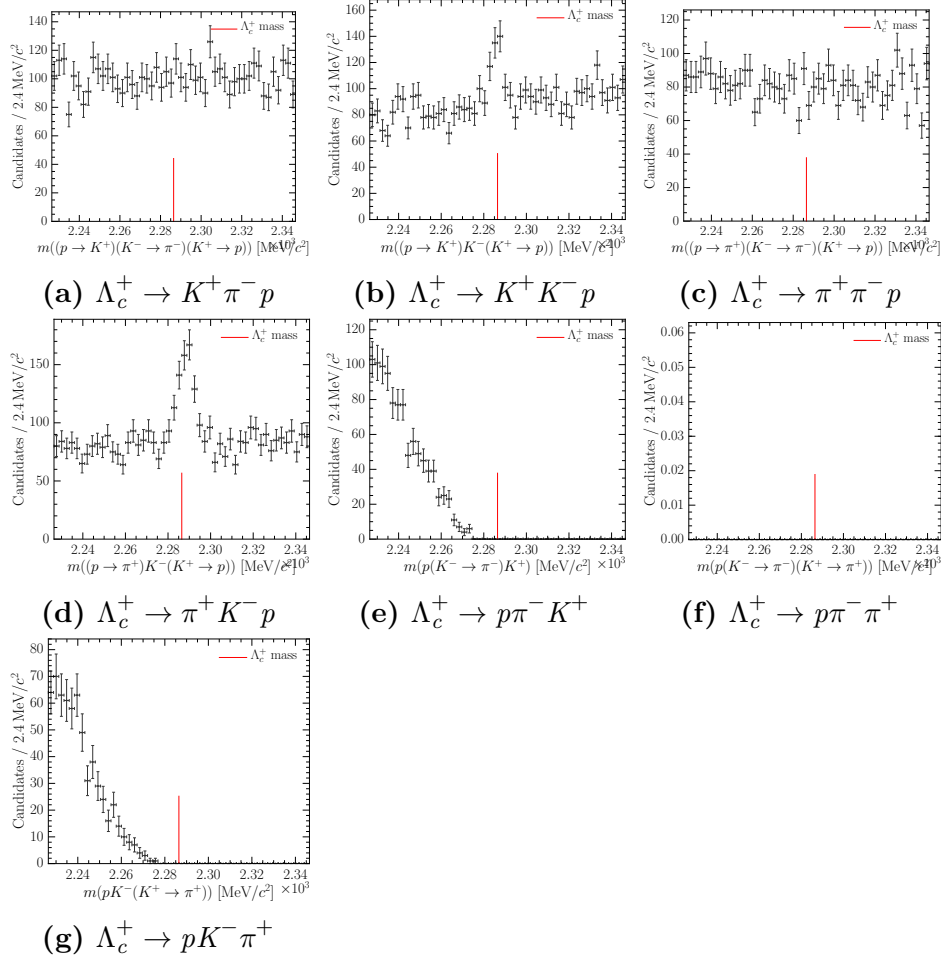


Figure 17.5: Wrong-mass distributions obtained when changing the mass hypotheses of the fully selected Λ_c^+ children in the pK^-K^+ mode, where one child is assigned the proton mass hypothesis. The x-axis on each sub-figure shows the substitutions that have been made. For example, the $pK^-(K^+ \rightarrow \pi^+)$ distribution is where the nominal kaon with same charge as the proton has been also assigned the pion hypothesis. The vertical red lines indicate the nominal mass of the misidentified particle under study. Only the 2012 magnet down data is shown.

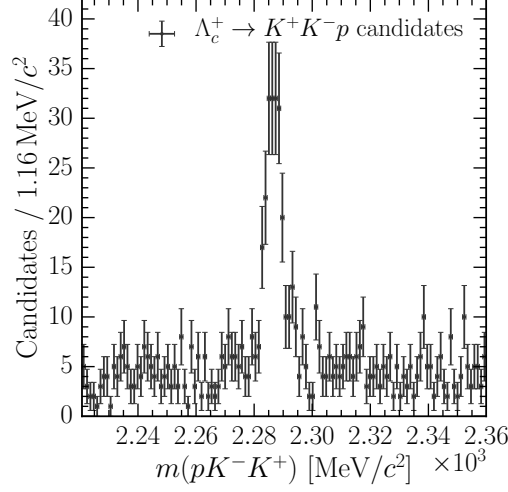


Figure 17.6: Mass spectrum of pK^-K^+ candidates that fall within 8 MeV of the nominal Λ_c^+ mass when the proton and same-sign kaon mass hypotheses are swapped. The mass spectrum of Λ_c^+ candidates with the swapped hypotheses is shown in Figure 17.5b. The lack of structures other than the flat combinatorial-background-like and peaking signal-like components indicate that the presence of this misidentification background does not affect the nominal pK^-K^+ yield extraction.

Table 17.4: Statistics on multiple candidates in the pK^-K^+ data. The total number of events N_{Events} is the number of proton-proton collisions that make up the data sub-sample, and $N_{Candidates}$ is the number of Λ_c^+ candidates. The fraction of events containing at least two candidates is $f_{Multiple}$, and the fraction of those events that fall within the signal window in the Λ_c^+ mass distribution is $f_{Multiple, sig.}$.

Year	Polarity	N_{Events}	$N_{Candidates}$	$f_{Multiple}$ [%]	$f_{Multiple, sig.}$ [%]
2011	Down	11600 ± 100	11800 ± 100	1.39 ± 0.11	0.16 ± 0.06
2011	Up	8300 ± 100	8400 ± 100	1.62 ± 0.14	0.39 ± 0.10
2012	Down	26900 ± 200	27300 ± 200	1.51 ± 0.08	0.23 ± 0.04
2012	Up	26900 ± 200	27400 ± 200	1.65 ± 0.08	0.28 ± 0.05

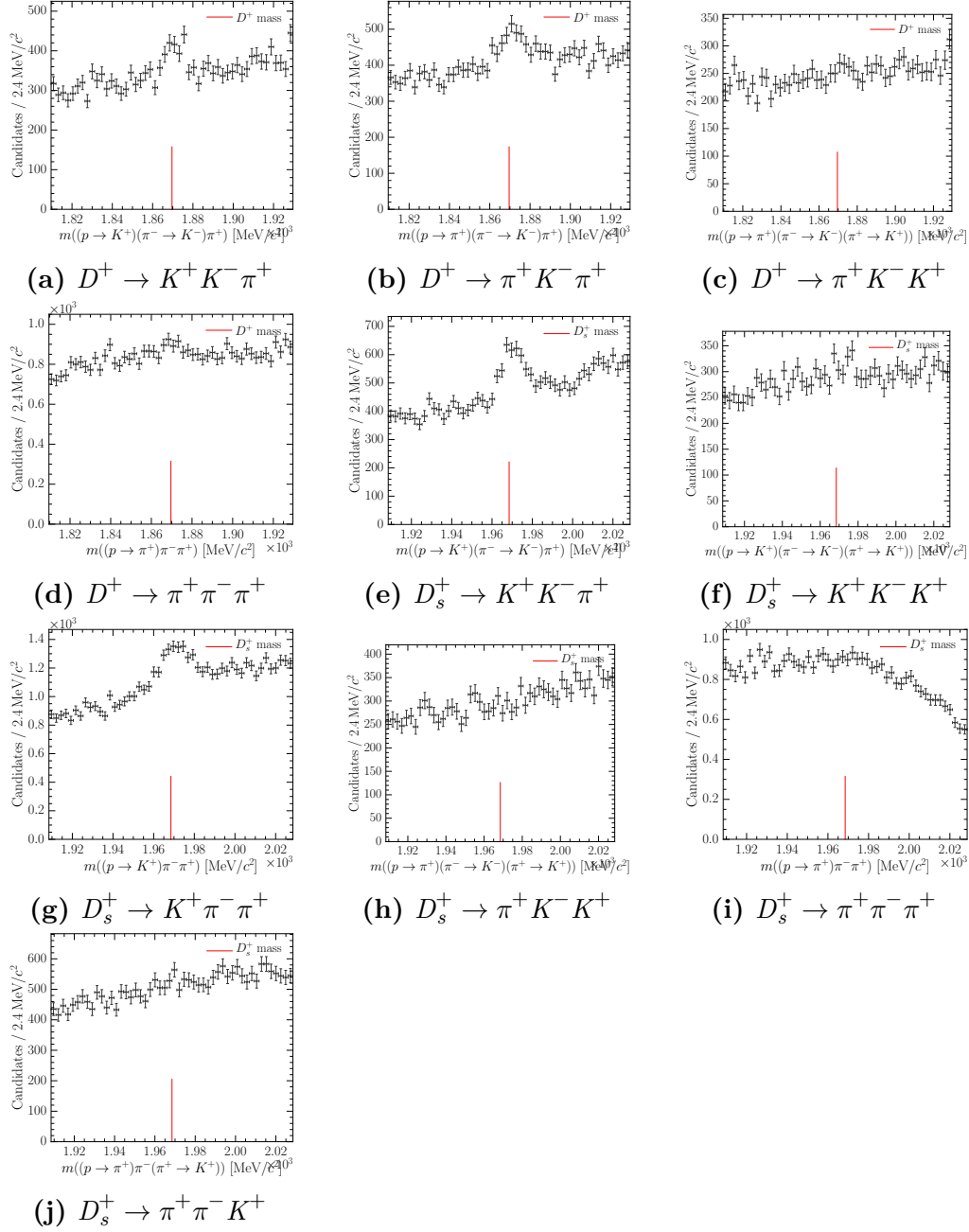


Figure 17.7: Wrong-mass distributions obtained when changing the mass hypotheses of the fully selected Λ_c^+ children in the $p\pi^-\pi^+$ mode, where no child is assigned the proton mass hypothesis. The x-axis on each sub-figure shows the substitutions that have been made. For example, the $(p \rightarrow \pi^+)(\pi^- \rightarrow K^-)\pi^+$ distribution is where the nominal proton candidate has been assigned the pion mass hypothesis, and the nominal pion with opposite charge to the proton has been assigned the kaon hypothesis. The vertical red lines indicate the nominal mass of the misidentified particle under study. Only the 2012 magnet down data is shown.

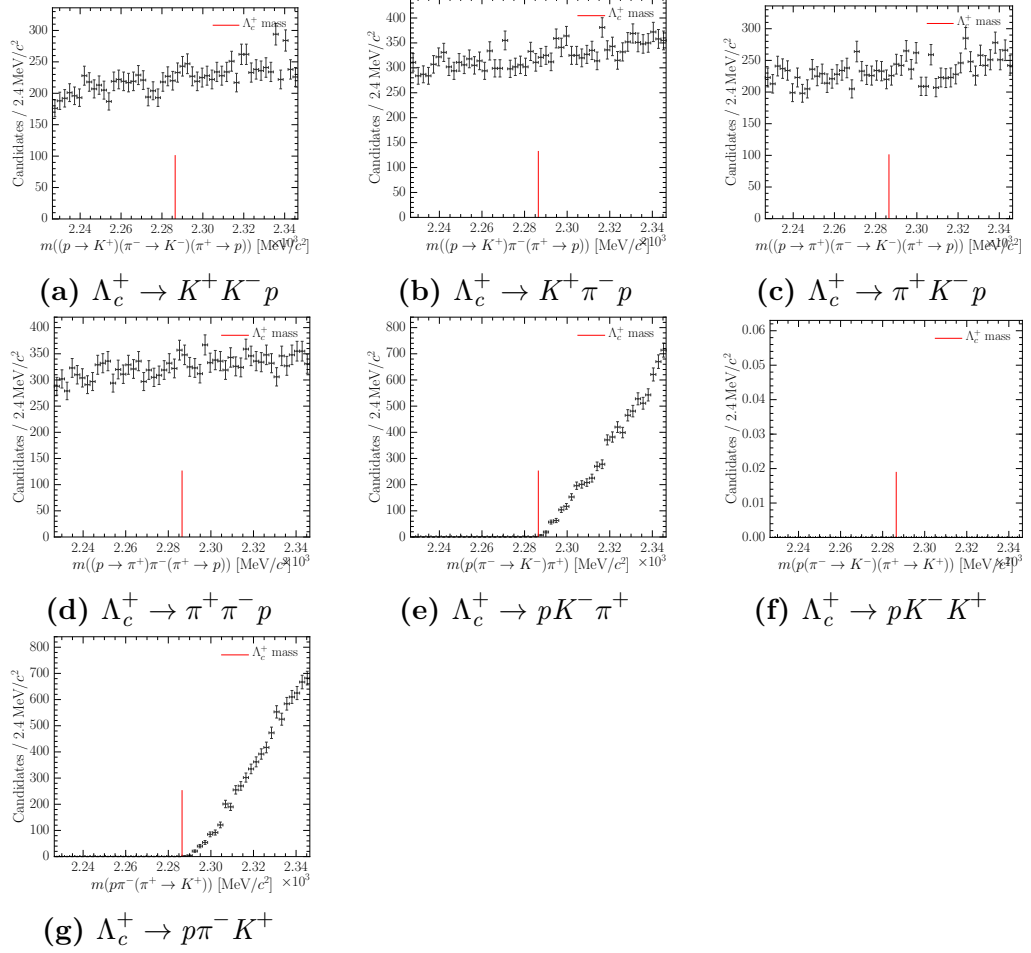


Figure 17.8: Wrong-mass distributions obtained when changing the mass hypotheses of the fully selected Λ_c^+ children in the $p\pi^-\pi^+$ mode, where one child is assigned the proton mass hypothesis. The x-axis on each sub-figure shows the substitutions that have been made. For example, the $p(\pi^- \rightarrow K^-)\pi^+$ distribution is where the nominal pion candidate with opposite charge to the proton candidate has been assigned the kaon mass hypothesis. The vertical red lines indicate the nominal mass of the misidentified particle under study. Only the 2012 magnet down data is shown.

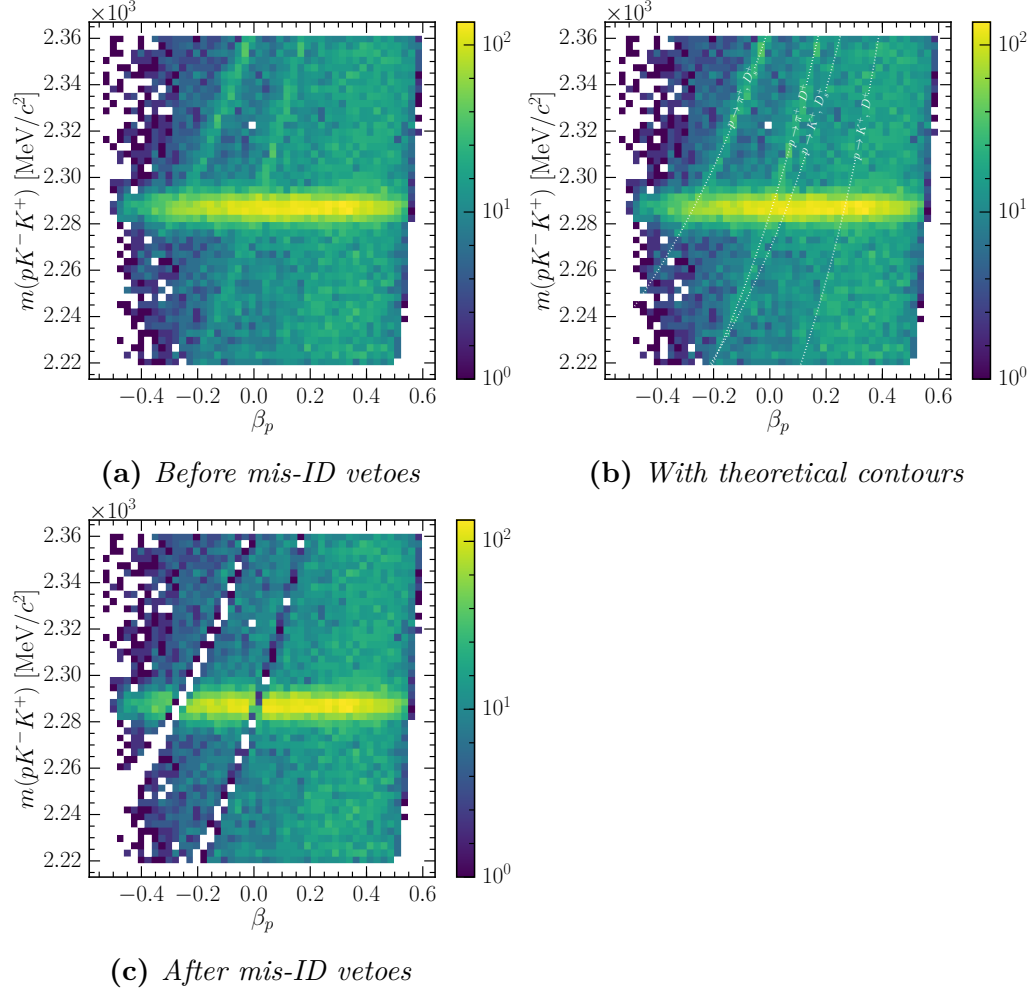


Figure 17.9: Plane of β_p - $m(\Lambda_c^+)$ for the $pK^- K^+$ mode in the 2012 magnet down dataset. The “theoretical contours” shown on Figure 17.9b are the solution of the parametric equation that defines the relationship between $m(\Lambda_c^+)$, β_p , and the mass of the misidentified particle decaying to a particular final state. The plane after the misidentification vetoes, defined in Section 17.3, is shown in Figure 17.9c.

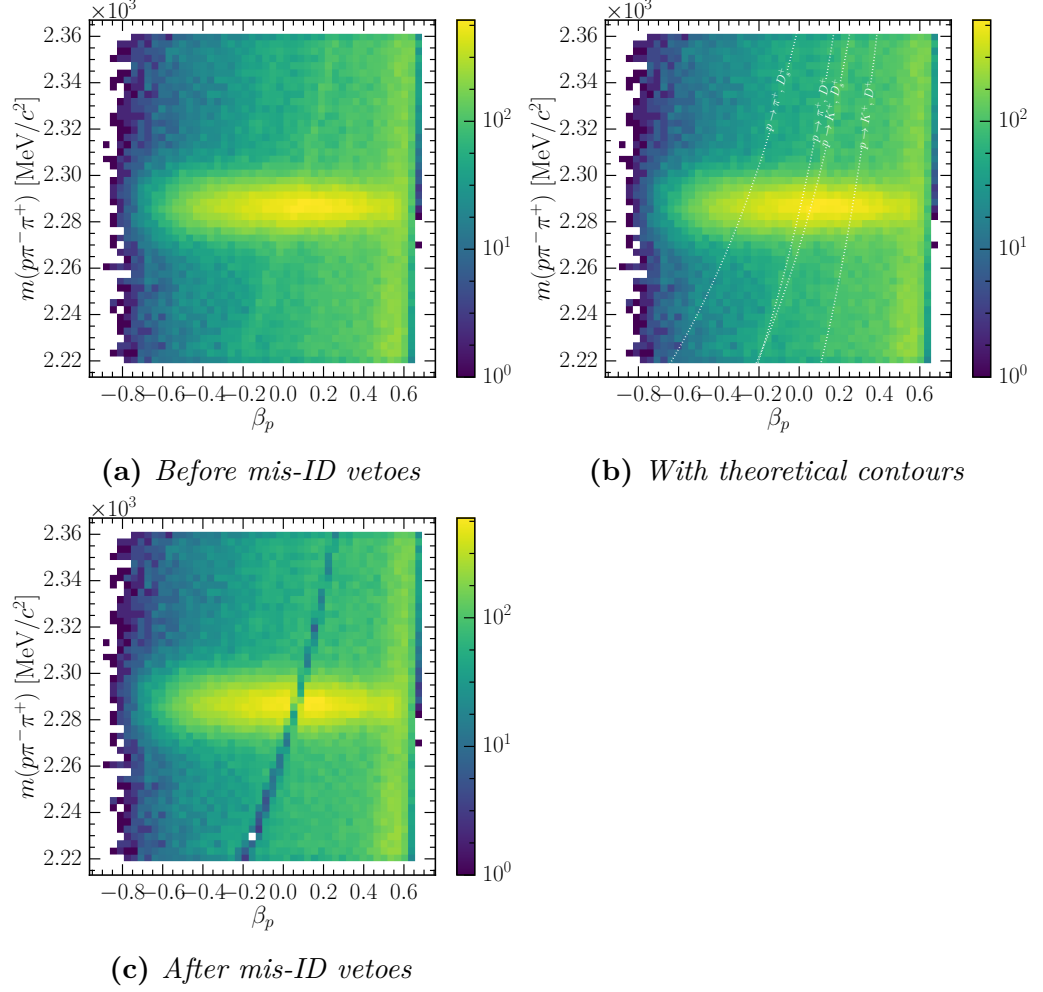


Figure 17.10: Plane of β_p – $m(\Lambda_c^+)$ for the $p\pi^-\pi^+$ mode in the 2012 magnet down dataset. The “theoretical contours” shown on Figure 17.10b are the solution of the parametric equation that defines the relationship between $m(\Lambda_c^+)$, β_p , and the mass of the misidentified particle decaying to a particular final state. The plane after the misidentification vetoes, defined in Section 17.3, is shown in Figure 17.10c.

Table 17.5: *Multiple candidates for the $p\pi^-\pi^+$ data. The total number of events N_{Events} is the number of proton-proton collisions that make up the data sub-sample, and $N_{\text{Candidates}}$ is the number of Λ_c^+ candidates. The fraction of events containing at least two candidates is f_{Multiple} , and the fraction of those events that fall within the signal window in the Λ_c^+ mass distribution is $f_{\text{Multiple, sig.}}$.*

Year	Polarity	N_{Events}	$N_{\text{Candidates}}$	$f_{\text{Multiple}} [\%]$	$f_{\text{Multiple, sig.}} [\%]$
2011	Down	64700 ± 300	65800 ± 300	1.62 ± 0.05	0.23 ± 0.03
2011	Up	46800 ± 200	47600 ± 200	1.80 ± 0.06	0.32 ± 0.04
2012	Down	163100 ± 400	166100 ± 400	1.75 ± 0.03	0.29 ± 0.02
2012	Up	161900 ± 400	165000 ± 400	1.80 ± 0.03	0.32 ± 0.02

18 Preliminary fits

The yield extraction technique used to measure the signal yields entering the computation of A_{CP}^{Raw} , shown in Equation (14.4), is described in Chapter 21. Before then, it is useful to fit to the charge-combined data in order to perform studies that require background subtraction, selection efficiency estimates, and the like. This chapter describes the binned χ^2 fit to the ph^-h^+ invariant mass spectrum that is used for such studies.

18.1 Input data and shape parameterisation

The data entering the fit depends on the study that the information is needed for. For example, the study of the optimal PID requirements described in Section 17.2 requires a mass fit before the offline PID requirements are applied. In all cases, the models and general fit procedure are the same.

The mass spectrum of the three-body ph^-h^+ system is used to discriminate between signal and background. The following will list the functional forms used in the fit to describe these components of the data. In all cases, the normalisation factors are omitted from the definitions, and the variable x represents the Λ_c^+ mass.

For both the pK^-K^+ and $p\pi^-\pi^+$ samples, the signal model f is the linear sum of two normal distributions G sharing a common mean μ but with different widths σ_1 and σ_2 , defined as

$$f(x; \mu, w, \alpha) = \alpha G(x; \mu, \alpha w) + (1 - \alpha) G(x; \mu, \frac{w}{\alpha}). \quad (18.1)$$

The parameter α describes the ‘relative strength’ of each normal distribution and is bounded in the range $[0, 1]$, and the widths σ_1 and σ_2 have been

parameterised as

$$\sigma_1 = \alpha w, \text{ and } \sigma_2 = \frac{w}{\alpha}. \quad (18.2)$$

This enforces the relation $\sigma_1 \leq \sigma_2$, which is useful to know when trying to reason about the fitted values of α , and was observed to increase the stability of the fit compared to having free, independent widths σ_1 and σ_2 . For both modes, the background model g is a first-order polynomial with coefficient a_0

$$g(x) = 1 + a_0 x. \quad (18.3)$$

The signal and background models are combined as a linear sum to form a total function h , where each component is weighted by their respective yields $N_{\text{Sig.}}$ and $N_{\text{Bkg.}}$. The χ^2 fit is performed on the data in the range

$$2230 < m(\Lambda_c^+) < 2350 \text{ GeV}/c^2, \quad (18.4)$$

with bins of width $1 \text{ MeV}/c^2$.

Once a fit has converged and the associated covariance matrix is positive-definite, sWeights [77] are computed and are attached to each candidate. This allows for the statistical subtraction of either the signal or the background component in bins of some variable that is not correlated with the Λ_c^+ mass.

Fits to the fully selected 2012 magnet down dataset for the pK^-K^+ and $p\pi^-\pi^+$ modes are given in Figure 18.1. A good description of the data is seen, and this is representative of the fit quality in all data sub-samples. Tables of signal and background yields measured in the preliminary fit are given for each fully selected subset of the data for pK^-K^+ in Table 18.1 and for $p\pi^-\pi^+$ in Table 18.2. Also given are the number of signal candidates per unit of integrated luminosity $N_{\text{Sig.}}/\mathcal{L}_{\text{int.}}$.

18.2 Validation

To evaluate the bias of the fitter and reliability of the uncertainties, pseudo-experiments are performed by generating MC data from the fitted PDF. One thousand pseudo-experiments are run for each mode and for each data sub-sample. The parameter values of the generating PDFs, from which data

Table 18.1: *Signal and background yields measured in the preliminary fit for each fully selected subset of the pK^-K^+ data.*

Year	Polarity	$N_{\text{Sig.}}$	$N_{\text{Bkg.}}$	$N_{\text{Sig.}}/\mathcal{L}$ [pb]
2011	Down	$3,930 \pm 80$	$6,670 \pm 90$	7.0 ± 0.2
2011	Up	$2,850 \pm 70$	$4,710 \pm 80$	6.8 ± 0.2
2012	Down	$9,250 \pm 120$	$15,160 \pm 140$	9.3 ± 0.2
2012	Up	$8,990 \pm 110$	$15,370 \pm 140$	9.0 ± 0.2

points are drawn, are the same as those found in the fits to the fully selected data, except for the values of yield parameters. Each yield parameter value for each pseudo-experiment is drawn from a Poisson distribution whose mean is set to the value of the yield found in the fits to the data [105].

For each pseudo-experiment, the fitted parameter value and the uncertainty on that fit are recorded, and a pull value p is computed as the difference between the fitted value v' and the true generating value v , divided by the parameter value uncertainty δv

$$p = \frac{v - v'}{\delta v}. \quad (18.5)$$

The true value for each yield parameter is taken to be the mean value of the Poisson distribution from which the generating value is drawn. An unbiased fitter with correct uncertainty coverage should have a distribution of pull values drawn from a unit Gaussian.

The results of the pseudo-experiments are shown in Figure 18.2 and Figure 18.3 for the 2012 magnet down pK^-K^+ and $p\pi^-\pi^+$ data. It is seen that the signal yield parameter is unbiased.

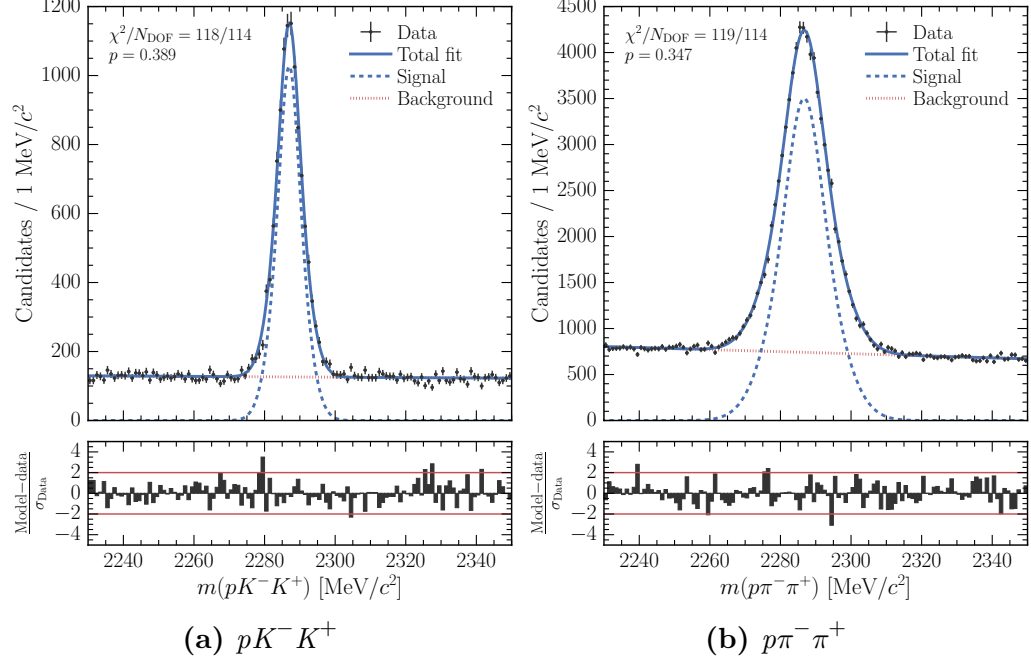


Figure 18.1: Fits to the Λ_c^+ mass spectrum in the 2012 magnet down dataset for pK^-K^+ (a) and $p\pi^-\pi^+$ (b). The full offline selection is applied. The solid blue line is the total fit to the data in black points, and the dotted red and dashed blue lines are the background and signal components, respectively. Below each fit is a pull plot, showing the difference between the total fit model and the data in each bin, normalised by the Poisson uncertainty on the number of entries in that bin.

Table 18.2: Signal and background yields measured in the preliminary fit for each fully selected subset of the $p\pi^-\pi^+$ data.

Year	Polarity	$N_{\text{Sig.}}$	$N_{\text{Bkg.}}$	$N_{\text{Sig.}}/\mathcal{L}$ [pb]
2011	Down	$25,040 \pm 220$	$34,060 \pm 240$	44.4 ± 0.9
2011	Up	$18,370 \pm 190$	$24,360 \pm 210$	43.5 ± 0.9
2012	Down	$59,830 \pm 350$	$88,450 \pm 390$	60.3 ± 0.8
2012	Up	$57,460 \pm 340$	$89,670 \pm 390$	57.4 ± 0.7

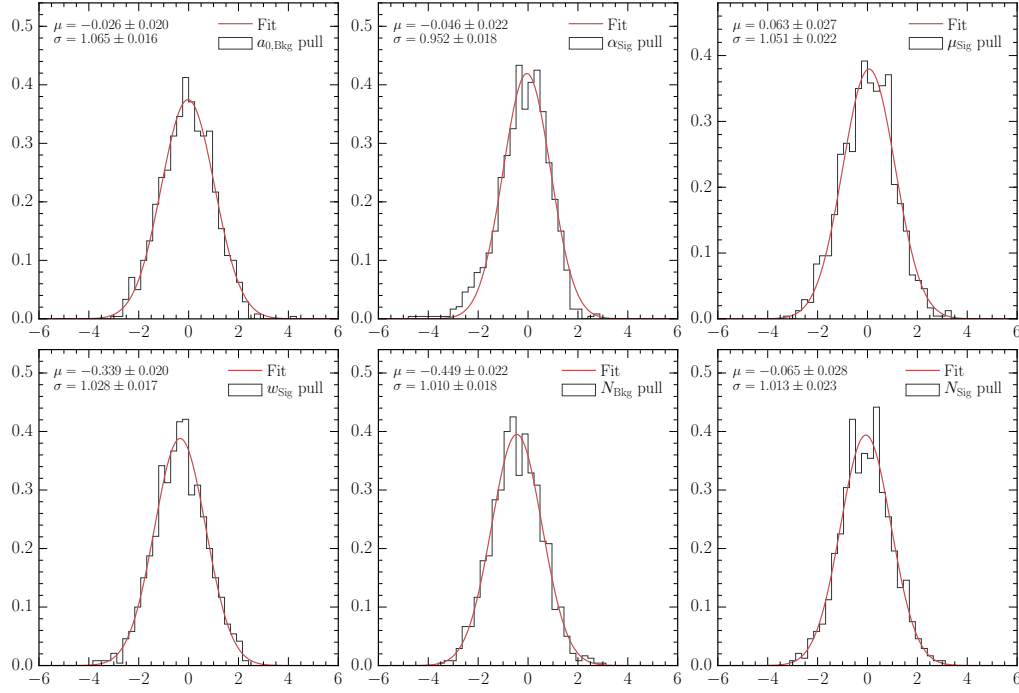


Figure 18.2: Validation of the fits to the Λ_c^+ mass spectrum in the charge-combined 2012 magnet down pK^-K^+ dataset. Each plot shows the distribution of the pull values for a single fit parameter, indicated in the legend, along with a fit of a Gaussian distribution with mean μ and width σ .

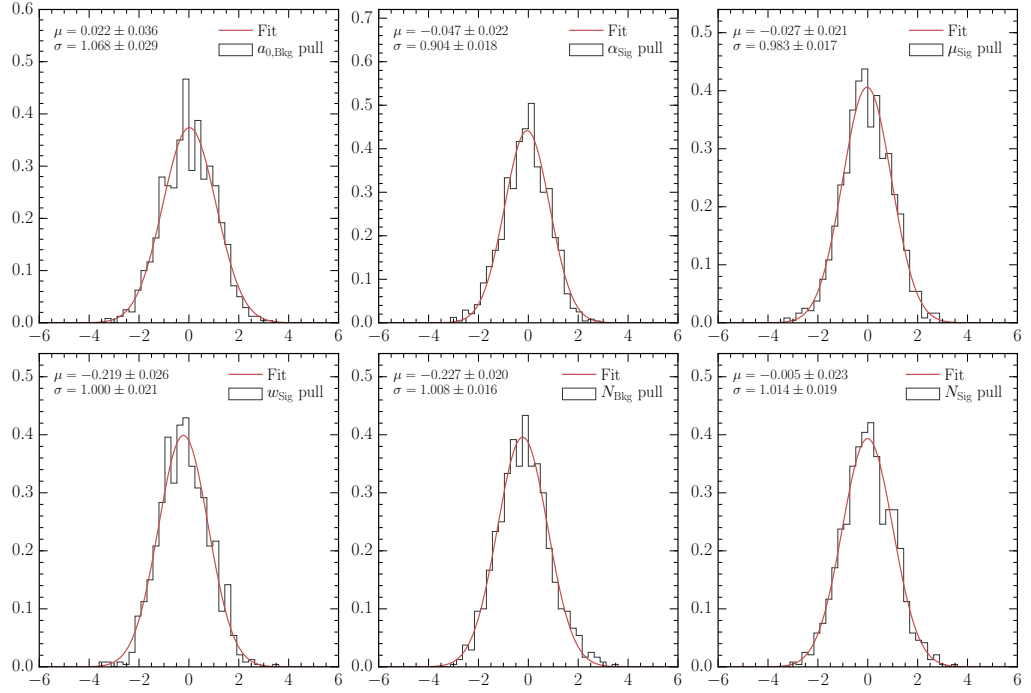


Figure 18.3: Validation of the fits to the Λ_c^+ mass spectrum in the charge-combined 2012 magnet down $p\pi^-\pi^+$ dataset. Each plot shows the distribution of the pull values for a single fit parameter, indicated in the legend, along with a fit of a Gaussian distribution with mean μ and width σ .

19 Kinematic weighting

For the assumptions in Equation (15.9) to hold, the kinematics of the particles contributing to the background asymmetries must be equal between the pK^-K^+ and $p\pi^-\pi^+$ final states. These are the kinematics of the Λ_b^0 , the muon from the Λ_b^0 , and the proton from the Λ_c^+ . Within a particular ph^-h^+ final state, it must also be the case that the h^+ and h^- kinematics are equal in order that $A_D^f(f)$ only contains contributions from the proton detection asymmetry $A_D(p)$. A comparison of these distributions between the 2012 magnet down pK^-K^+ and $p\pi^-\pi^+$ datasets is given in Figures 19.1 to 19.6. The signal distributions are shown, where the contents of each bin is defined as the sum of signal sWeights within that bin, as computed from the result of the mass fit to the charge-combined samples described in Chapter 18. The correlation of the variables with the Λ_c^+ mass is seen to be small, with a correlation coefficient less than 2 % for all variables. Only the muon kinematics across the modes and the h^-h^+ kinematics within the modes agree, and so a weighting is necessary.

A boosted decision tree (BDT) is used to equalise the relevant particle kinematics between the two modes. This is an uncommon technique in physics analyses, with ratios of multi-dimensional histograms being used more often, and so it will be briefly summarised in the following. (An introduction to decisions trees and boosting is presented in Appendix B.) The creation of the specific models using for weighting will then be presented, followed by a discussion on weighting validation.

19.1 Weighting with decision trees

It is common in high-energy physics to want to be able to bring two distributions into agreement, as is the case in this analysis where particle kinematics should be equal between the pK^-K^+ and $p\pi^-\pi^+$ data. Call one possibly multi-dimensional distribution the *target* f_t and the other the *original* distribution f_o . To transform f_o to have the same form as f_t , there is some transformation function W

$$f_t(x) = W(x)f_o(x), \quad (19.1)$$

where x is a possibly multi-dimensional parameter. As the true distributions of f_t and f_o are not known in the data, they are approximated as histograms, partitioning the data in x and counting the number of events that fall in each bin. The transformation function is quantised as w in bins, where in the i th bin

$$f_t(x_i) = w(x_i)f_o(x_i), \quad (19.2)$$

and where $f_{t(o)}(x_i)$ is the count of the target (original) distribution in the i th bin. The histogram of f_o can be transformed to that of f_t by computing the per-bin *weights*

$$w(x_i) = \frac{f_t(x_i)}{f_o(x_i)}. \quad (19.3)$$

This is valid assuming there are no bins where $f_o(x_i) = 0$.

To model the true probability density functions in x as closely as possible, the histogram must have a fine enough binning that all structures are resolved. Specifically, as in the determination of the PID efficiency in Section 10.4, the binning must be fine enough such that difference between the true PDFs of the two distribution within any given bin are small. This must be balanced with the limited statistical precision available in each bin: too fine a binning will result in the original distribution being weighted to match the statistical fluctuations in the target, rather than the physical features. Limited sample sizes become particularly problematic for higher-dimensional weighting, when even 10 equally spaced bins per dimension in a three-dimensional binning requires 1,000 bins. Given physical distributions such as momentum

and pseudorapidity, which have long tails, it can be difficult to invent a binning that avoids sparsely populated or empty bins whilst still accounting for differences in the distributions.

To overcome the limitations of histogram-based weighting, this analysis defers the computation of the per-event weights w_i to a decision tree with gradient boosting [106]. The classifier is trained with the loss function

$$L = \sum_{l \in \text{Leaves}} \frac{(w_{l,o} - w_{l,t})^2}{w_{l,o} + w_{l,t}}, \quad (19.4)$$

where $w_{l,t(o)}$ is the sum of target (original) weights in the l th leaf node of the regression tree. For the initial iteration, and an unweighted set of training events, this is equal to the number of target (original) events in the leaf node. This metric has the tree split the samples into regions where the differences between the sample sizes is maximal, such that the classifier focuses on weighting the entries in those regions. Predictions p_l within a leaf are computed as

$$p_l = \log \frac{w_{l,t}}{w_{l,o}}. \quad (19.5)$$

At the end of the i th iteration, each event from the ‘original’ sample is weighted by $w_i = w_{i-1}e^{p_l}$, where l is the leaf node the event falls in. The target distribution is assigned unit weights $w_i = w_0$. This weighting has the same effect as the weighting in the ADABOOST [107] algorithm, increasing the importance of events in the original distribution that are in regions with a large original/target difference.

The actual weighting of events is identical to that in the histogram weighting: it is the ratio of target to original events in some region. The use of gradient boosting allows for a more intelligent way of finding the ‘bins’ in which the weights are computed, and iteratively improves the weights based on that method.

19.2 Evaluation

The decision tree with gradient boosting is trained with the $p\pi^-\pi^+$ data as the ‘original’ sample and the pK^-K^+ data as the ‘target’. The weighting

procedure is performed separately for each data sub-sample in year and magnet polarity. As it is expected that the Λ_b^0 and muon kinematics are correlated with that of the Λ_c^+ , the Λ_c^+ transverse momentum p_T and pseudorapidity η are included as inputs to the BDT. Also included are the proton p_T and η , as these distributions are seen to disagree after weighting when only the Λ_c^+ kinematics are used. To avoid biases due to fluctuations in the data, two classifiers A and B are trained, and the pK^-K^+ and $p\pi^-\pi^+$ data are each randomly split into two sub-samples 1 and 2: classifier A is trained using sub-sample 1, and the weights for sub-sample 2 are computed using A ; likewise, classifier B is trained using sub-sample 2, and the weights for sub-sample 1 are computed using B . Sub-samples 1 and 2 are then combined for the remainder of the analysis. The input training data are weighted by signal sWeights, as computed from the result of the mass fits described in Chapter 18. Each classifier is boosted in 300 iterations, the maximum number of leaf nodes allowed in the regression tree is four, and the criteria for splitting includes the requirement that any new nodes must contain at least 200 events. The choice of these hyper-parameters will be discussed in Section 19.3, as well as the specific choice of original and target samples.

Figures 19.7 to 19.10 show the Λ_b^0 , muon, Λ_c^+ , and proton kinematics after weighting with the product of signal sWeights and the kinematic weights. Figures 19.11 and 19.12 show the h^-h^+ kinematics, *within* the modes, after the same weighting. There is a considerable improvement in the agreement between the pK^-K^+ and $p\pi^-\pi^+$ data in the Λ_b^0 and proton kinematics. The good agreement in the muon kinematics between the datasets and the h^-h^- kinematics within the datasets does not change. It is noted that the Λ_c^+ kinematic agreement also improves, although this does not affect the measurement of ΔA_{CP} as there is no associated asymmetry.

19.3 Validation

The purpose of the kinematic weighting is to make the particle kinematics the same between the pK^-K^+ and $p\pi^-\pi^+$ modes. Figures 19.7 to 19.10

show good agreement ‘by eye’, but a quantitative assessment is needed. What affects the measurement of ΔA_{CP} is not so much the absolute level of disagreement between the two histograms, for example, but the disagreement in a particular region combined with both the size of the respective asymmetries involved and the density of the data in that region. The effect of the remaining differences on ΔA_{CP} is discussed in the context of systematic uncertainties in Chapter 23.

To provide some quantitative estimate of the agreement, a BDT is employed. If the two samples, the target weighted with signal sWeights and the original weighted with the product of signal sWeights and kinematic weights, are truly identical, then a BDT designed to *discriminate* between them should perform no better than random guessing. The performance of a BDT trained for binary classification can be evaluated by computing the area under the receiver operating characteristic (ROC) curve, or just area under curve (AUC). The ROC curve compares the false positive rate (FPR) of the classifier with the true positive rate (TPR) as a function of the classifier output probability. The TPR is defined as the fraction of true signal events that are correctly classified as such, whilst the FPR is the fraction of background events that are classified as signal. Plotting the FPR increasing along the x -axis and the TPR increasing along the y -axis, the ROC curve for an optimal classifier passes through the top-left corner of the plot, where a large fraction of the signal is accepted and a correspondingly large background fraction is rejected. A classifier that cannot distinguish between signal and background any better than random guessing will have a ROC curve of the form $x = y$. In these two extreme cases, the AUC is 1 and 0.5, respectively.

A BDT, with gradient boosting, is trained to discriminate between the weighted pK^-K^+ and $p\pi^-\pi^+$ data, using three-quarters of the available 2012 magnet down data as input. The Λ_b^0 , muon, and proton p , p_T , η , and azimuthal angle ϕ are used as features. The Λ_c^+ kinematics are not included as there is no asymmetry associated to any residual differences there might be. The maximum number of leaf nodes for any one regression tree is 11, the number of boosting iterations performed is 400, and the

minimum number of samples allowed in a leaf node is 200. The performance of the BDT is evaluated using the remaining quarter of the dataset, and the resulting ROC curve is shown in Figure 19.13. For comparison, it is shown along with a curve obtained when the same classifier is trained using data with no kinematic weighting, as well as a curve when a rudimentary, two-dimensional, 10×10 histogram weighting in $\Lambda_c^+ p_T$ and η is used. It is seen that the inclusion of the kinematic weights significantly reduces the AUC, from 0.61 with no kinematic weights to 0.52 for the BDT weighting. This shows that the BDT weighting does increase the similarity between the kinematic distributions across the modes, and that it can be significantly more powerful than a two-dimensional histogram weighting.

It is the AUC metric evaluated on the training sample that was maximised when choosing the hyper-parameters to use for the weighting. A grid search was performed where the following hyper-parameter values were tested in all combinations: 100, 200, 300, and 400 boosting iterations; a minimum of 50, 100, 200, 300, 500, and 1000 events required to create a new node; and a maximum number of allowed leaf nodes per regression tree of 2, 3, 4, 5, 6, 7, and 8. The hyper-parameters used for the weighting (300 iterations, a minimum of 200 events per node, and a maximum of 4 leaf nodes per tree) were those for which the AUC was smallest.

19.4 Weight statistics

Weighting can increase the statistical contribution of individual entries within a dataset, but it cannot increase the overall statistical power. The number of entries N' in a weighted dataset can be computed as the sum of the per-entry weights w_i , and the variance on this quantity is given as the sum of the squares of the weights

$$N' = \sum_i^N w_i, \quad (\delta N')^2 = \sum_i^N w_i^2 \quad (19.6)$$

For certain values of w_i , it can be that $\delta N' < \delta N$. To prevent this, the number of ‘effective’ entries in the weighted dataset is computed, under the

relation that the relative variance on the effective entries is equal to the relative variance on the weights

$$\frac{N_{\text{Eff.}}}{\delta N_{\text{Eff.}}} = \frac{\sum_i^N w_i}{\sqrt{\sum_i^N w_i^2}}, \quad (19.7)$$

where $N_{\text{Eff.}}$ is assumed to be Poisson-distributed, and so $(\delta N_{\text{Eff.}})^2 = N_{\text{Eff.}}$, giving

$$N_{\text{Eff.}} = \frac{\left(\sum_i^N w_i\right)^2}{\sum_i^N w_i^2}. \quad (19.8)$$

This relation can be enforced by redefining the weights w_i by a factor $W = \sum_i w_i / \sum_i w_i^2$. For unit weights, $N_{\text{Eff.}} = N$ and $\delta N_{\text{Eff.}} = \sqrt{N}$. Otherwise, $N_{\text{Eff.}}$ is the size of the hypothetical dataset that has the same statistical power as the weighted dataset. For this analysis, it provides a handle on how well the two data samples agree in the weighting variables before the weighting, as, in general, a higher level of disagreement requires a larger fraction of candidates to be ‘thrown away’ by the weighting, reducing the statistical power.

Table 19.1 gives several statistics related to the weighting procedure for the $p\pi^-\pi^+$ data. For all data sub-samples, the effective number of candidates is no less than 80% the number of candidates entering the weighting procedure. As there are around five times more $p\pi^-\pi^+$ candidates than pK^-K^+ , the statistical uncertainty of the ΔA_{CP} measurement will still be dominated by the pK^-K^+ sample size, even with an effective 20% reduction in the $p\pi^-\pi^+$ sample size. It is this fact that motivates the choice of the $p\pi^-\pi^+$ mode as the source of the ‘original’ distributions in the weighting and the pK^-K^+ mode as the ‘target’.

Table 19.1: Statistics computed on the weighted $p\pi^-\pi^+$ data for all data sub-samples used in the analysis. The quantities are defined in Section 19.4.

Year	Polarity	N	$\sum w_i$	$\sum w_i^2$	$\sqrt{\sum w_i^2}$	$N_{\text{Eff.}}$	$\delta N_{\text{Eff.}}$	$\frac{N_{\text{Eff.}}}{N}$
2011	Up	42688	39408	44801	211	34663	186	0.812
2011	Down	59044	53627	58867	242	48853	221	0.827
2012	Up	147064	137461	155696	394	121362	348	0.825
2012	Down	148218	138168	153675	392	124227	352	0.838

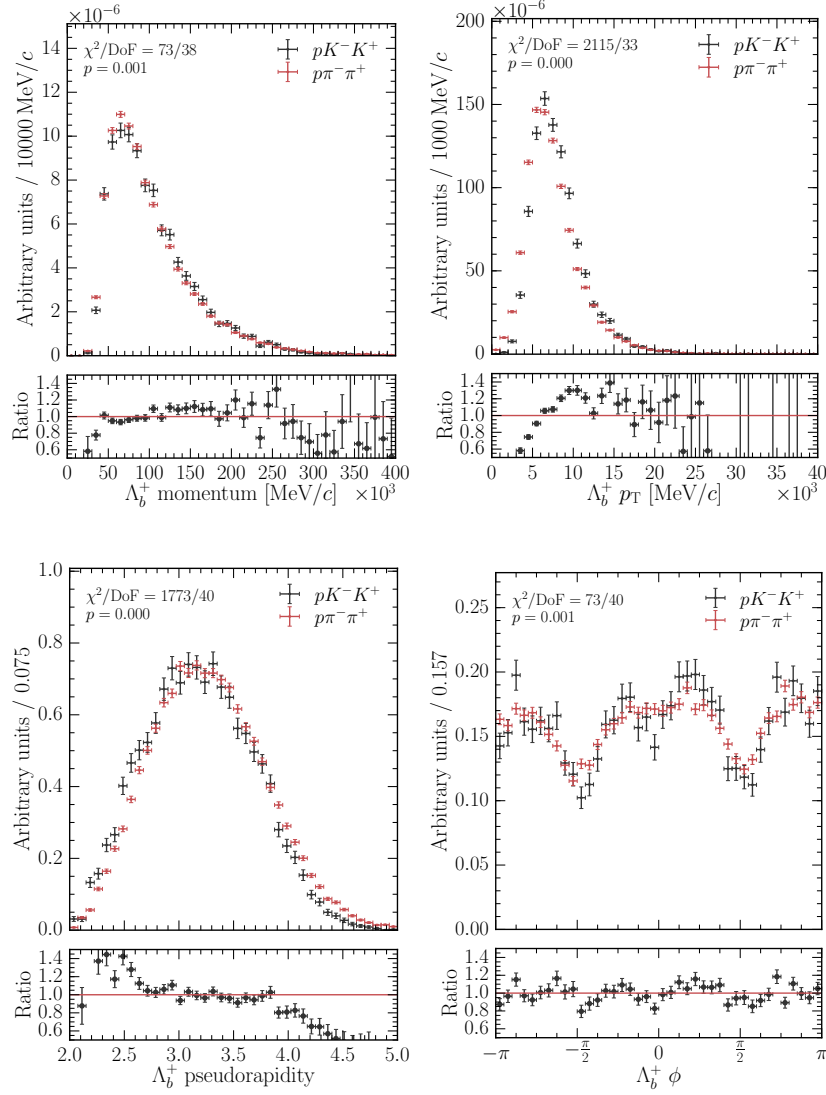


Figure 19.1: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the Λ_b^0 , weighted by signal $sWeights$. The 2012 magnet down data are shown.

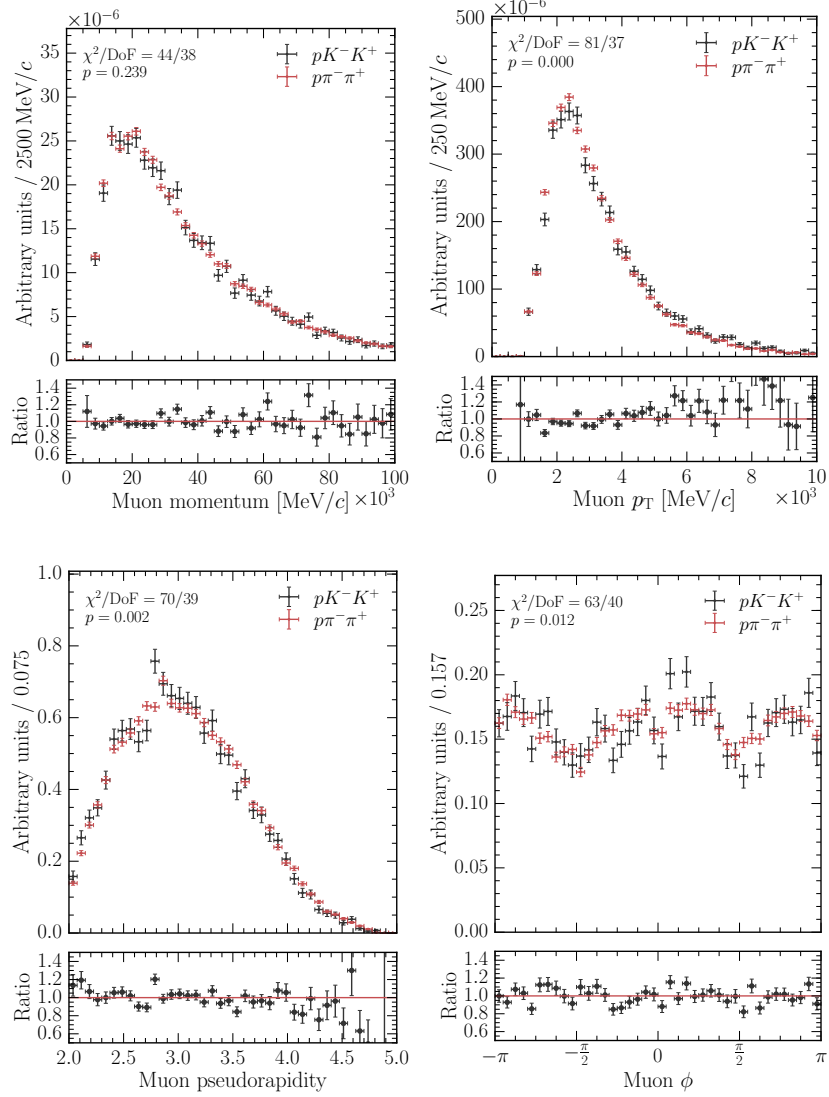


Figure 19.2: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the muon from the Λ_b^0 , weighted by signal s Weights. The 2012 magnet down data are shown.

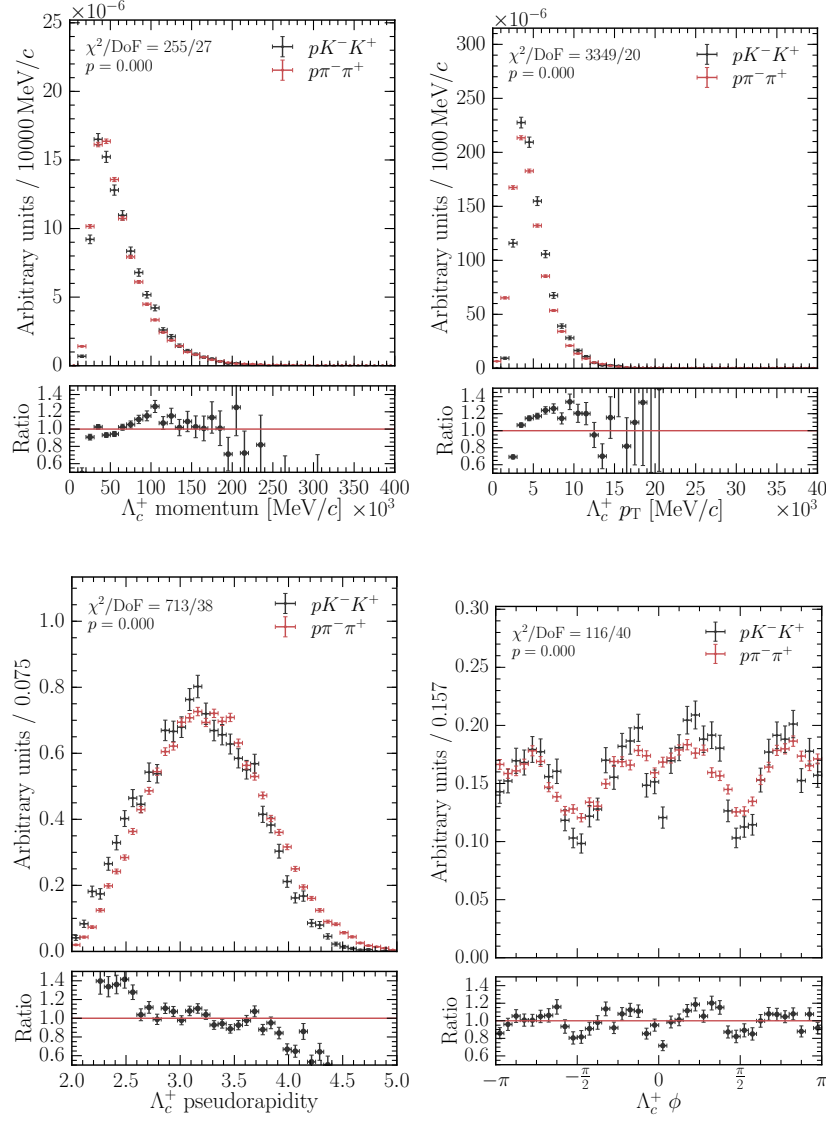


Figure 19.3: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the Λ_c^+ , weighted by signal $sWeights$. The 2012 magnet down data are shown.

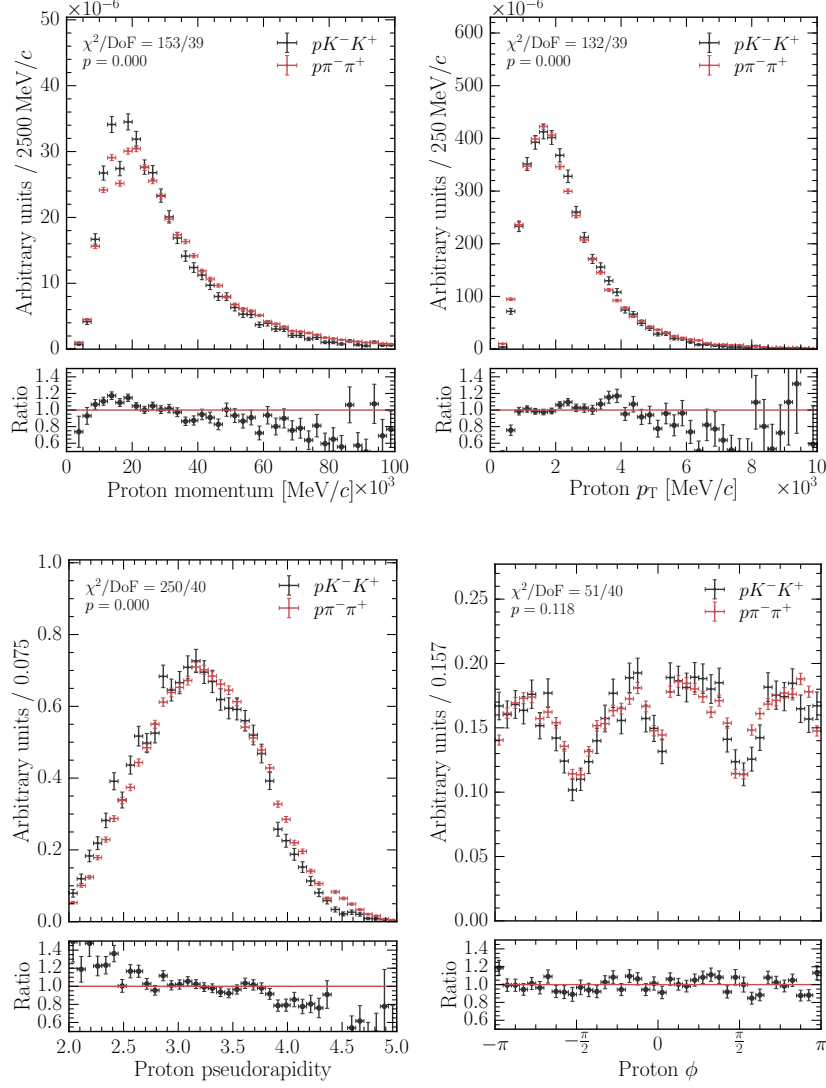


Figure 19.4: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the proton from the Λ_c^+ , weighted by signal $sWeights$. The 2012 magnet down data are shown.

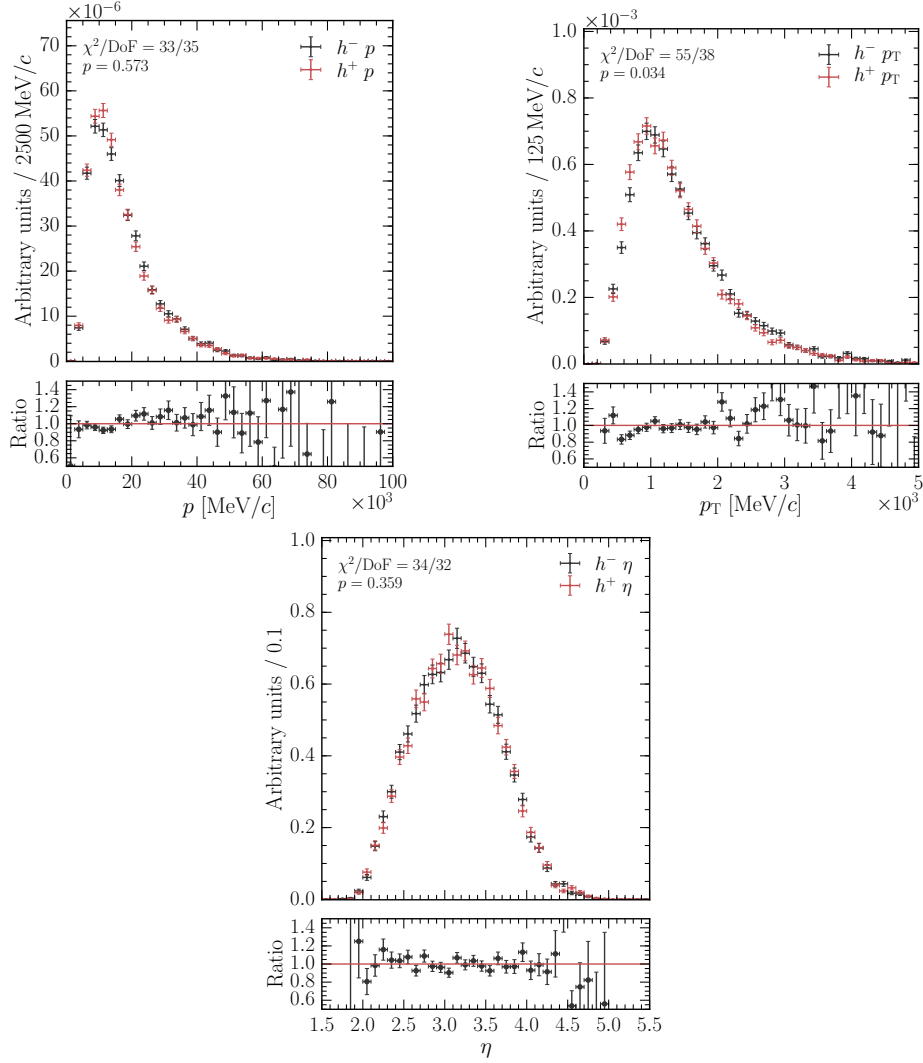


Figure 19.5: Clockwise from the top left: total momentum, transverse momentum, and pseudorapidity of the K^- and K^+ Λ_c^+ children in the $pK^- K^+$ data, weighted by signal $sWeights$. The 2012 magnet down data are shown.

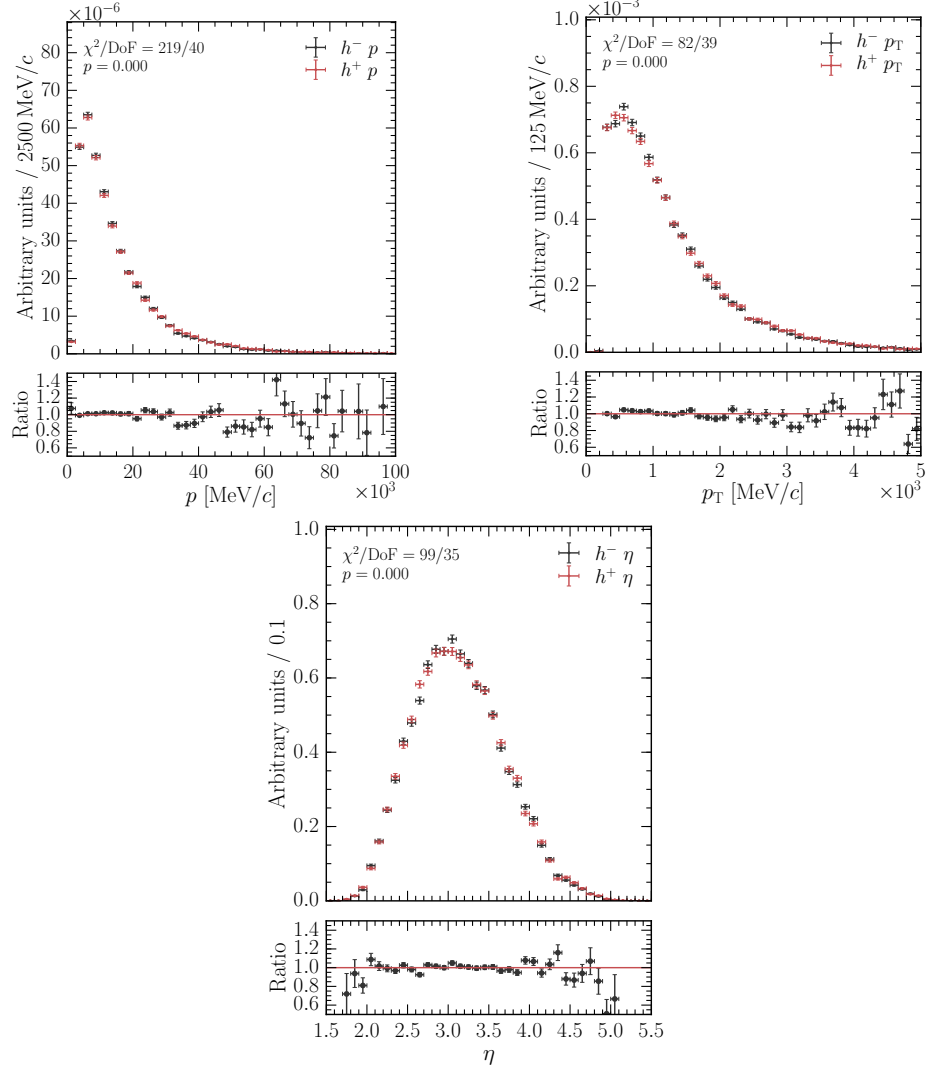


Figure 19.6: Clockwise from the top left: total momentum, transverse momentum, and pseudorapidity of the π^- and π^+ Λ_c^+ children in the $p\pi^- \pi^+$ data, weighted by signal $sWeights$. The 2012 magnet down data are shown.

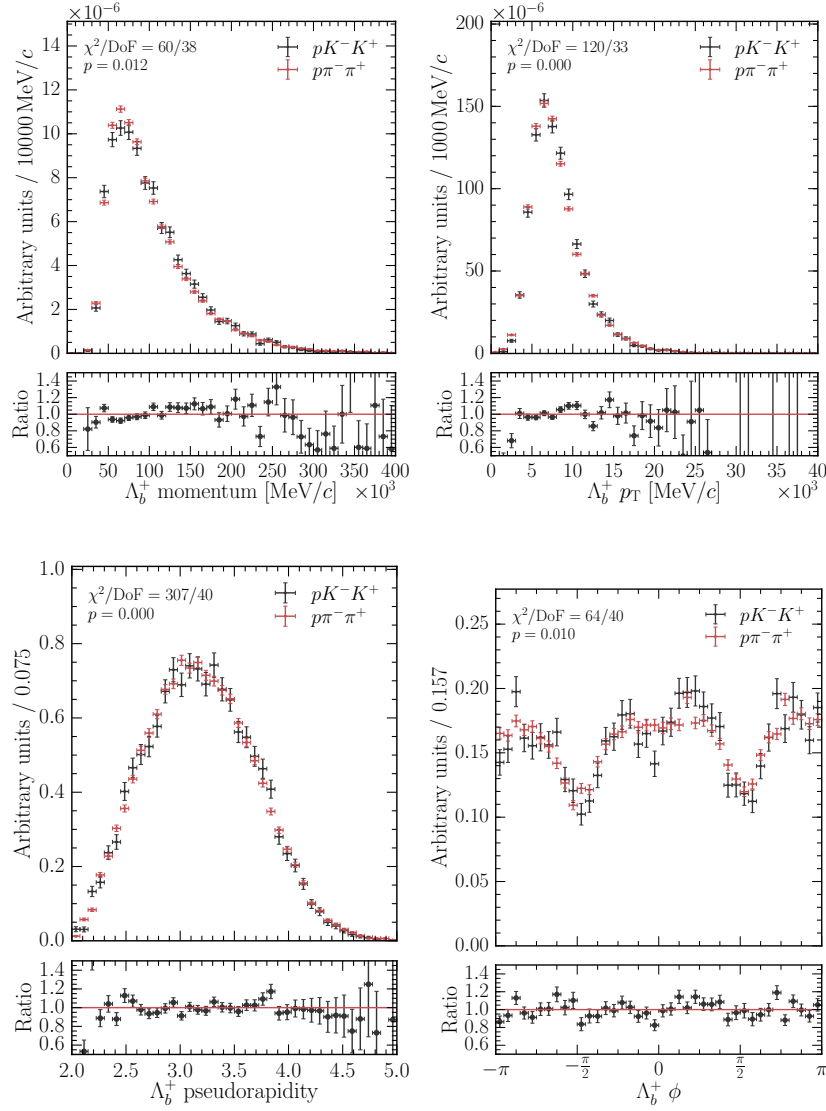


Figure 19.7: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the Λ_b^0 , weighted by the product of signal s Weights and kinematic weights. The 2012 magnet down data are shown.

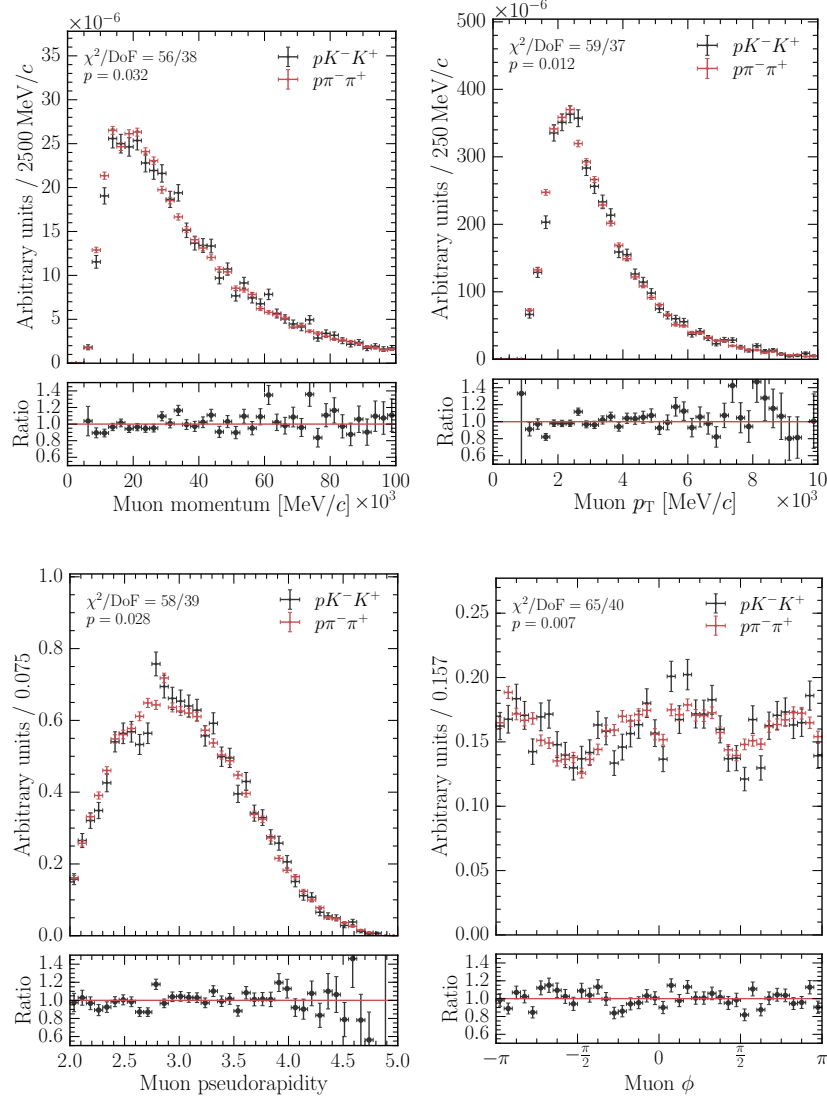


Figure 19.8: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the muon from the Λ_b^0 , weighted by the product of signal $sWeights$ and kinematic weights. The 2012 magnet down data are shown.

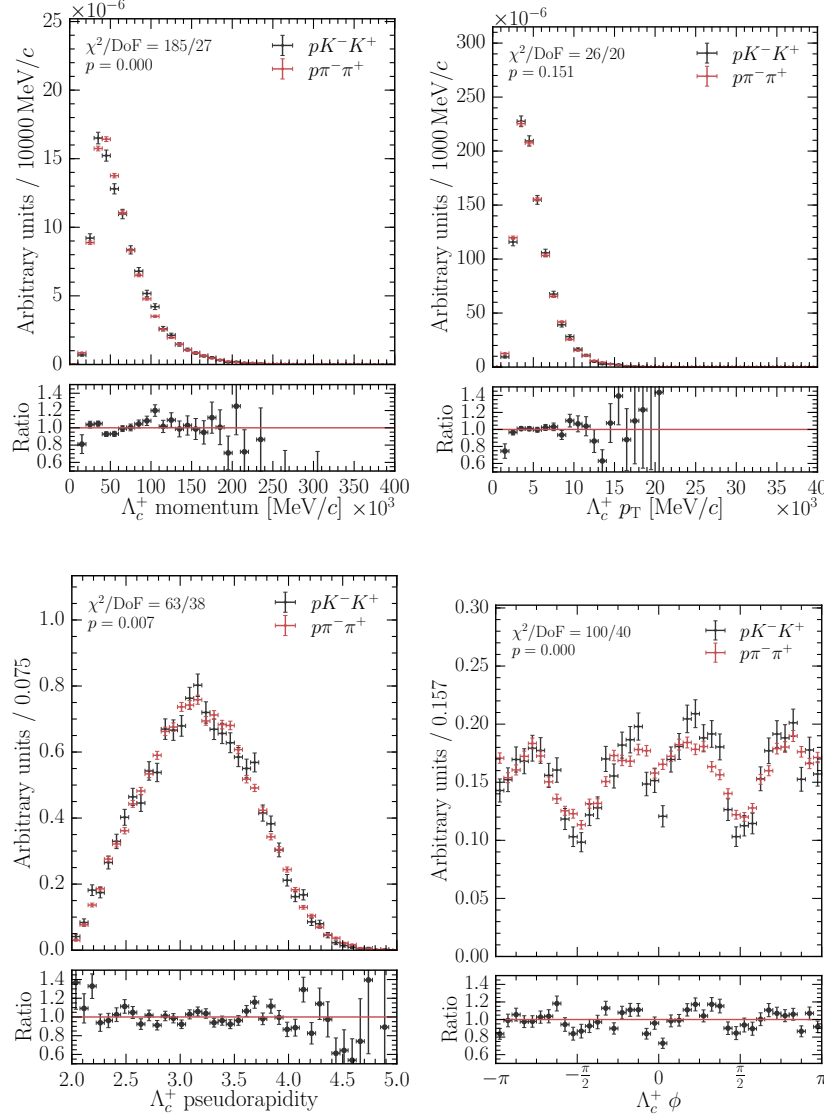


Figure 19.9: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the Λ_c^+ , weighted by the product of signal s Weights and kinematic weights. The 2012 magnet down data are shown.

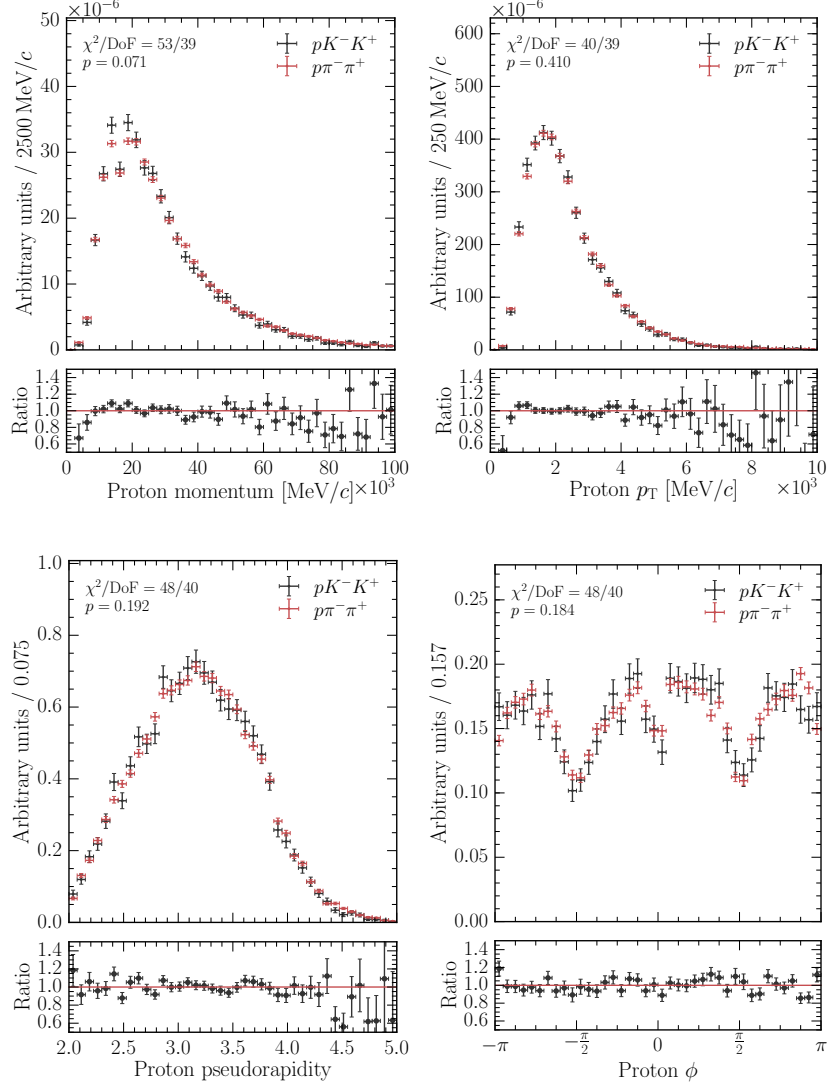


Figure 19.10: Clockwise from the top left: total momentum, transverse momentum, angle ϕ , and pseudorapidity of the proton from the Λ_c^+ , weighted by the product of signal s Weights and kinematic weights. The 2012 magnet down data are shown.

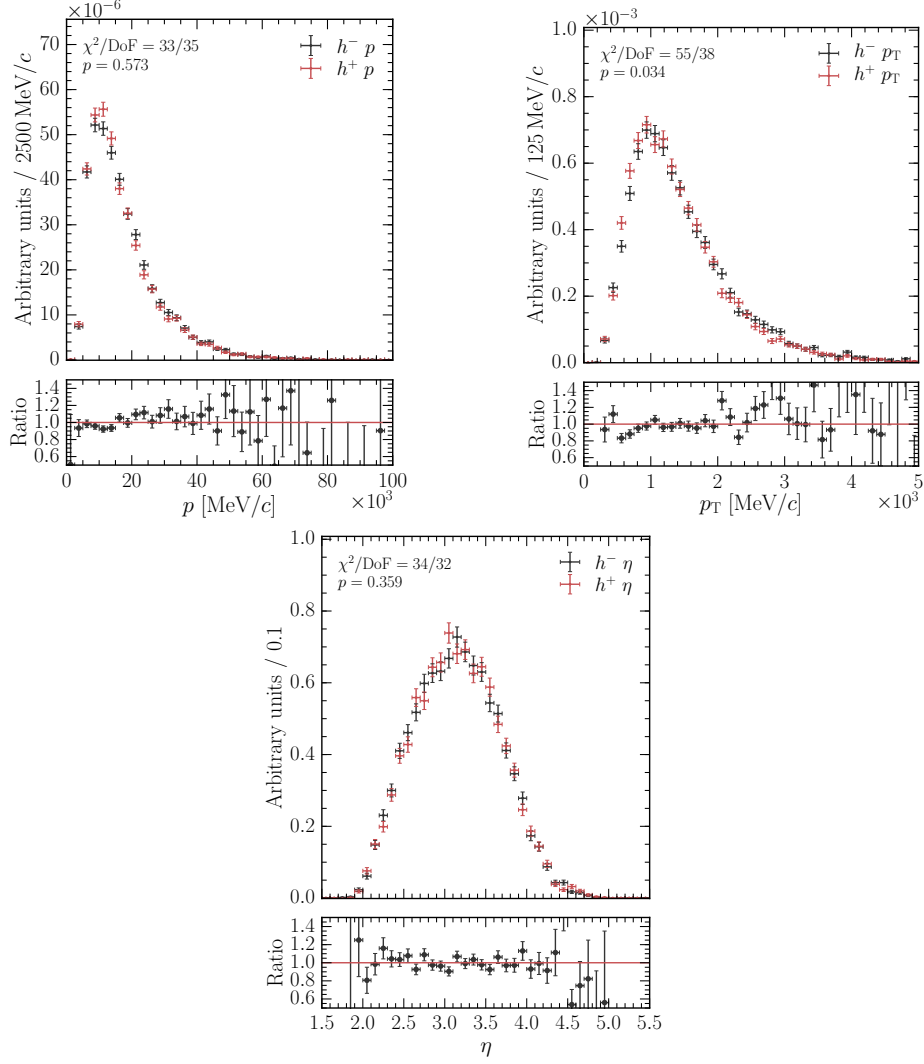


Figure 19.11: Clockwise from the top left: total momentum, transverse momentum, and pseudorapidity of the K^- and K^+ Λ_c^+ children in the $pK^- K^+$ data, weighted by the product of signal s Weights and kinematic weights. The 2012 magnet down data are shown.

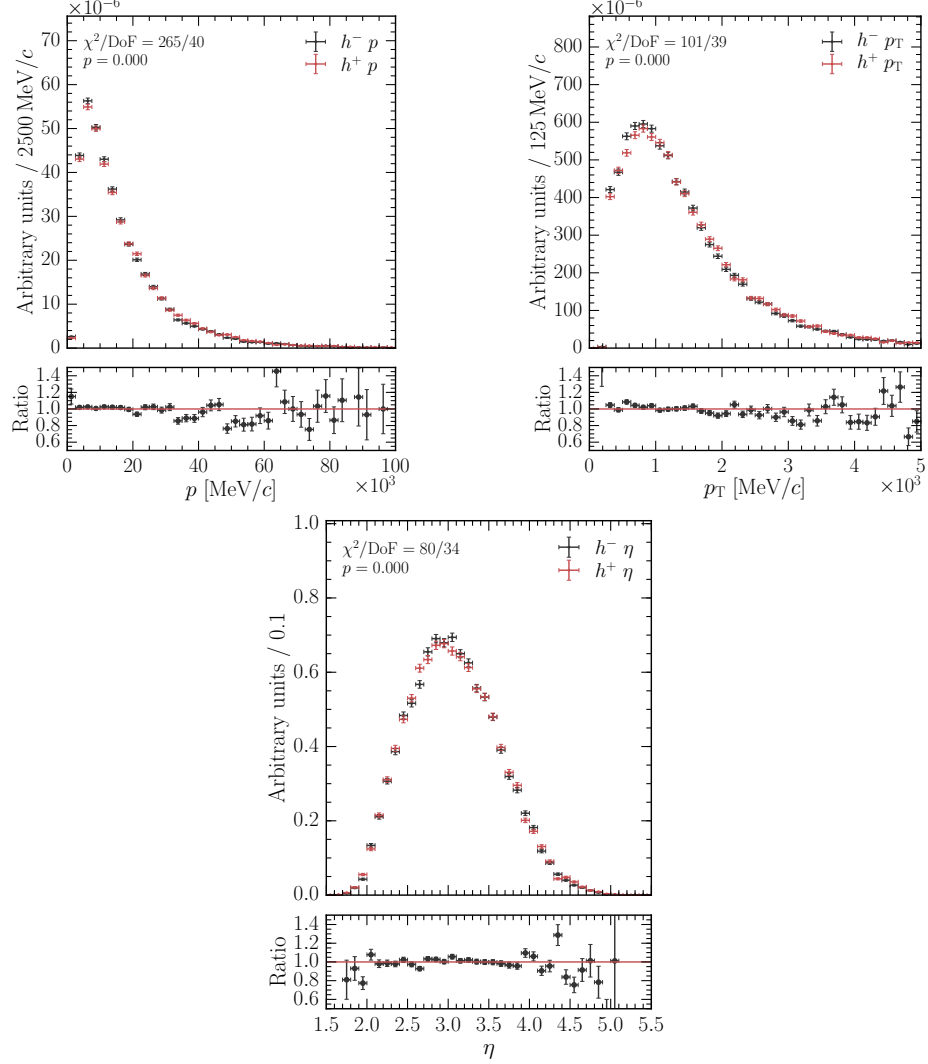


Figure 19.12: Clockwise from the top left: total momentum, transverse momentum, and pseudorapidity of the π^- and π^+ Λ_c^+ children in the $p\pi^-\pi^+$ data, weighted by the product of signal s Weights and kinematic weights. The 2012 magnet down data are shown.

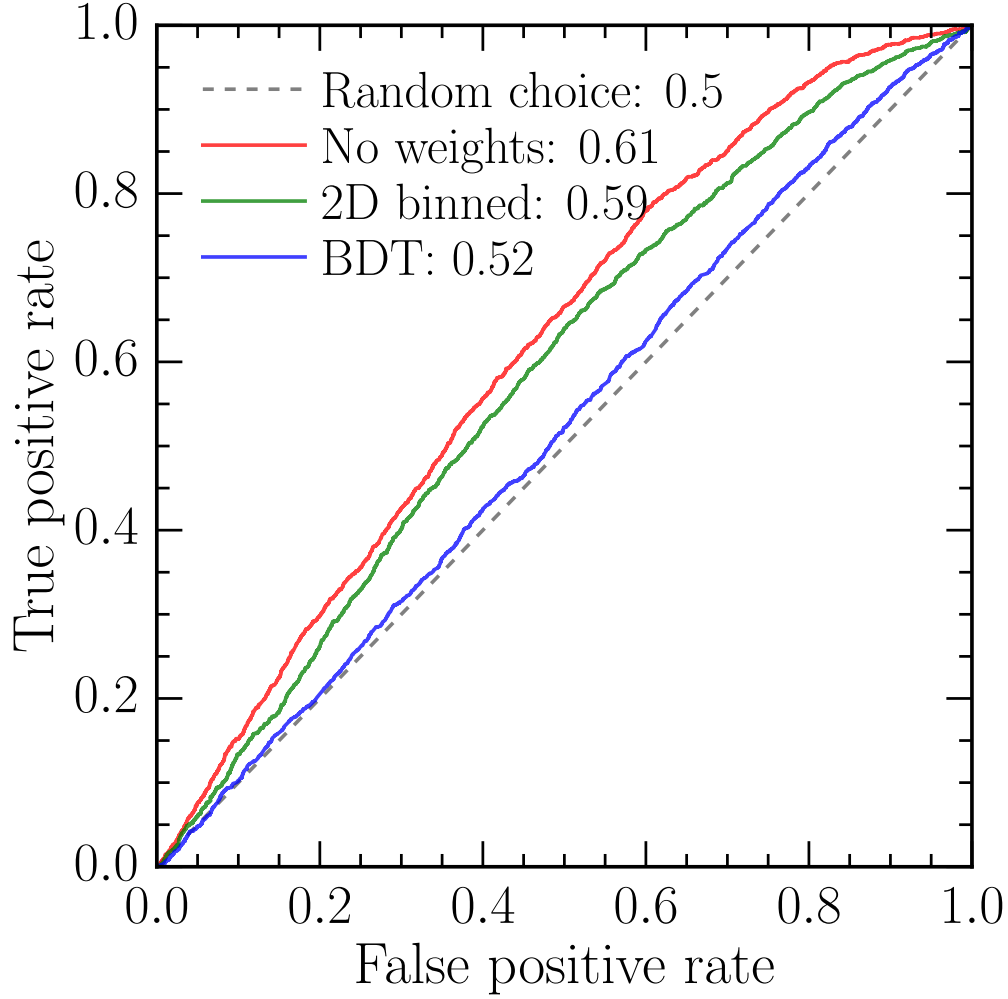


Figure 19.13: ROC curves for different kinematic weighting techniques. The “no weights” data has no kinematic weights applied, only signal $sWeights$. The “2D binned” and “BDT” data uses the product of the respective kinematic weights and the signal $sWeights$ for the $p\pi^-\pi^+$ data, and uses signal $sWeights$ for the pK^-K^+ data. The area under each ROC curve, the AUC, is given in the legend.

20 Phase space efficiency

The definition of the five-dimensional phase space of the $\Lambda_c^+ \rightarrow ph^-h^+$ decays is given in Section 15.1. In this chapter, the phase space variables in the data will be shown, and a description will be given of how a map of the acceptance and selection efficiency across those variables is computed.

20.1 Distributions in data

Figures 20.1 and 20.3 show the one-dimensional projections of $m^2(ph^-)$, $m^2(ph^+)$, $m^2(h^-h^+)$, θ_p , ϕ_p , and $\phi_{h^-h^+}$ in the 2012 magnet down datasets, weighted by signal sWeights. The respective one-dimensional projections from the generator-level MC, described in Section 16.1, are overlaid. Figures 20.2 and 20.4 show the two-dimensional projections of $m^2(ph^+)-m^2(h^-h^+)$, $m^2(ph^-)-m^2(ph^+)$, and $m^2(ph^+)-m^2(h^-h^+)$, and illustrate the complex resonant dynamics in the decays. The child-pair squared mass $m^2(ph^+)$ is not used as part of the phase space parameterisation, but it is shown for completeness. In all cases, the phase space variables are evaluated using the momentum vectors computed using the DECAYTREEFITTER algorithm, where the ph^-h^+ mass is constrained to be the nominal Λ_c^+ mass, as described in Chapter 16.

20.2 Efficiency modelling

As introduced in Section 15.1, for a meaningful comparison with theoretical predictions of ΔA_{CP} , the measurement must account for any sculpting of

the Λ_c^+ phase space due to the selection and the LHCb acceptance. This can be done either by providing an efficiency map for theorists to apply to their models, or by correcting the measurement with it. Such a map can be computed by comparing the phase space distributions of the signal before and after the selection. As the signal distributions before the stripping are not available for analysis, the MC samples described in Section 16.1 are used.

The phase space efficiency can be considered as a five-dimensional function $\varepsilon(\vec{x})$, in the phase space $\vec{x}$, that transforms the original distribution $D(\vec{x})$ to the observed one $D'(\vec{x})$

$$D'(\vec{x}) = \varepsilon(\vec{x})D(\vec{x}). \quad (20.1)$$

If the distributions D and D' were known, the efficiency could be computed as the ratio of the two distributions. Taking these distributions from the MC, an efficiency *map* could be approximated by partitioning the distributions as a five-dimensional histogram and taking the ratio between them. This approach is identical in nature to that defined in the kinematic weighting described in Section 19.1, as an efficiency can be considered as a weighting, and comes with the same drawbacks. In this particular case, it is hard to reduce the dimensionality of the problem, which would make a histogram method more tractable, as it is not clear which dimensions would be near-constant in the efficiency, if any. Indeed, this could be checked by inspecting one-dimensional projections of the efficiency, but this neglects correlations between variables which could significantly alter the total efficiencies. To avoid such difficulties, this analysis employs the gradient boosting decision tree algorithm, as described in Section 19.1, in order to estimate the efficiency function $\varepsilon(\vec{x})$.

The BDT is trained using selected, fully reconstructed truth-matched MC as the ‘target’, and generator-level MC that has not been required to be in the LHCb acceptance as the ‘original’ source. Half the available MC is used for training, whilst the other half is used for evaluation. Figures 20.5 and 20.6 show the one-dimensional distributions of the five phase space variables before and after the selection, as well as one-dimensional projections of the

5D efficiency function learnt by the BDT. For comparison, the ratios between the one-dimensional histograms are also shown, with which the function projections show a good agreement. It can be seen in some cases, such as the $\phi_{h^-h^+}$ projection in both modes, that the BDT does not capture the visible structure of the histogram ratio. This is likely due to the small sample size used in the training, such that the BDT does not ‘see’ such structures as statistically significant, modelling a comparatively flat efficiency.

To correct the real data for the acceptance and selection efficiencies, they are fed through the trained BDT and assigned per-decay weights. These weights are then used in the χ^2 fit to measure A_{CP}^{Raw} .

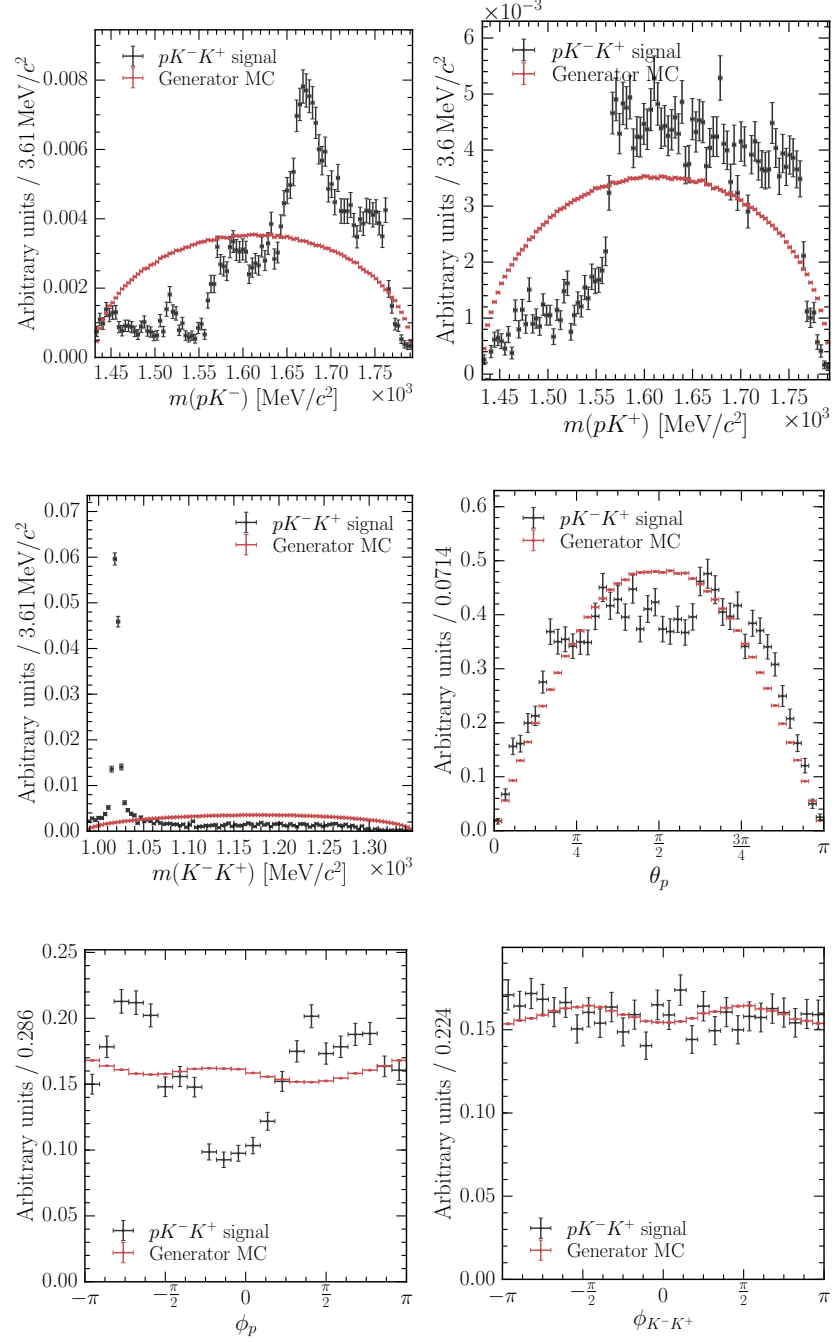


Figure 20.1: One dimensional projections of the $\Lambda_c^+ \rightarrow p K^- K^+$ phase space in the fully selected 2012 magnet down data, weighted by signal s Weights, along with the respective one-dimensional projections from the generator-level MC, described in Section 16.1.

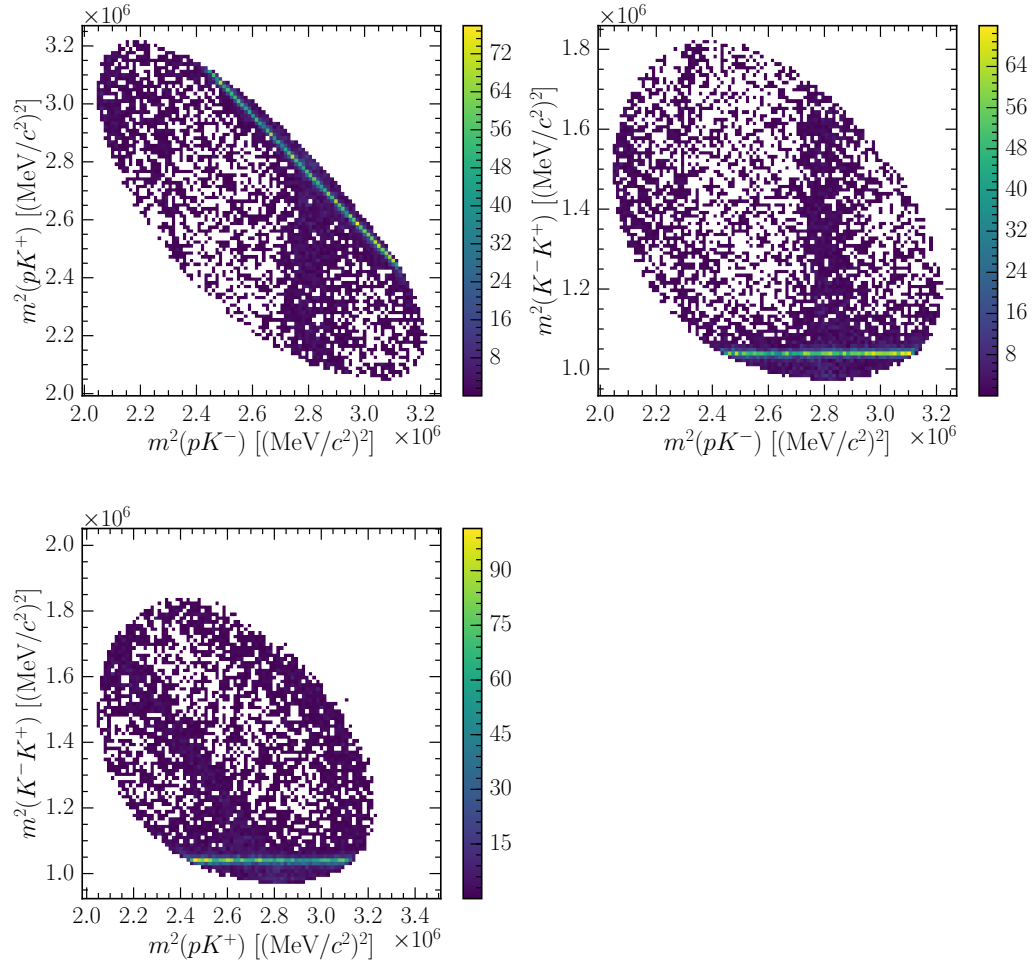


Figure 20.2: Two dimensional projections of the $\Lambda_c^+ \rightarrow pK^-K^+$ phase space in the fully selected 2012 magnet down data, weighted by signal $sWeights$. The contents of bins where the sum of signal $sWeights$ is negative are forced to be zero.

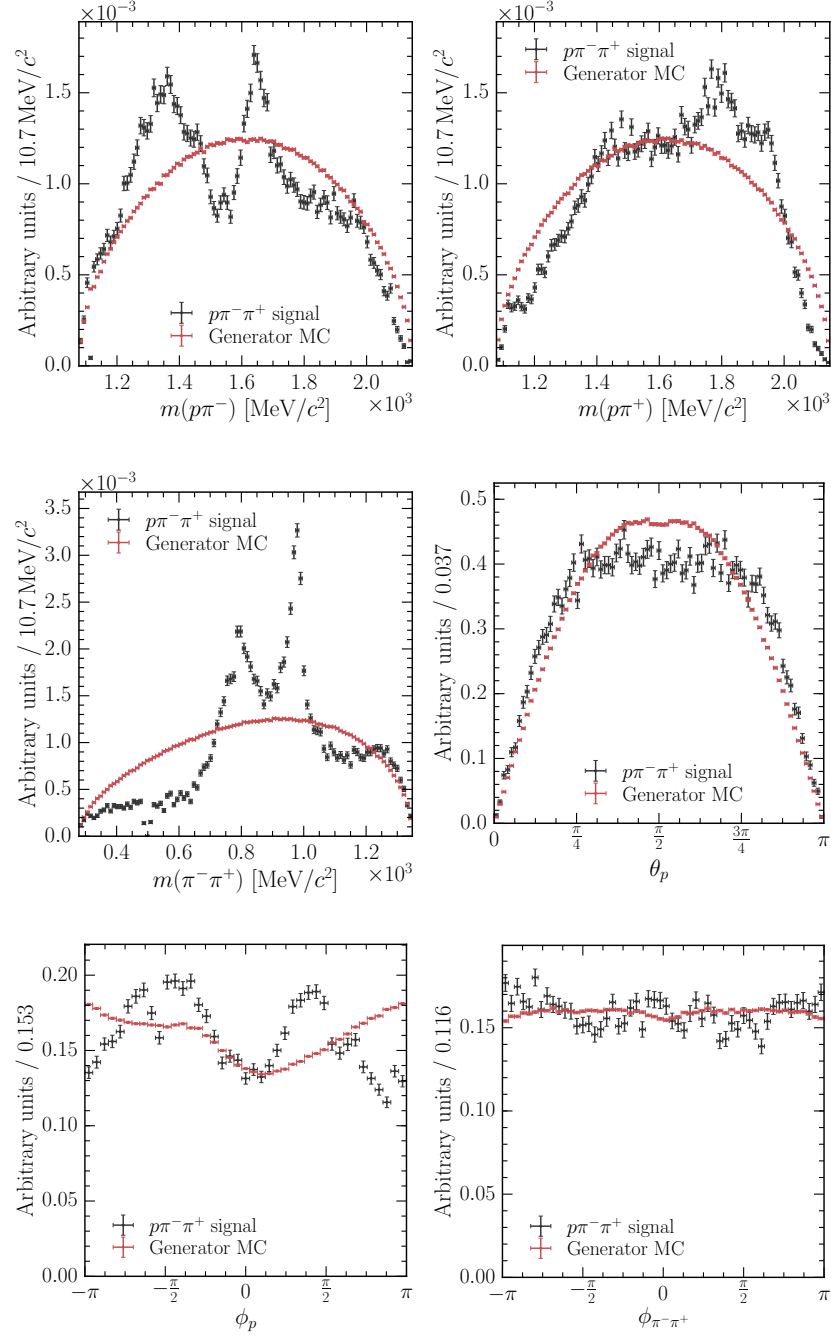


Figure 20.3: One dimensional projections of the $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ phase space in the fully selected 2012 magnet down data, weighted by signal s Weights, along with the respective one-dimensional projections from the generator-level MC, described in Section 16.1.

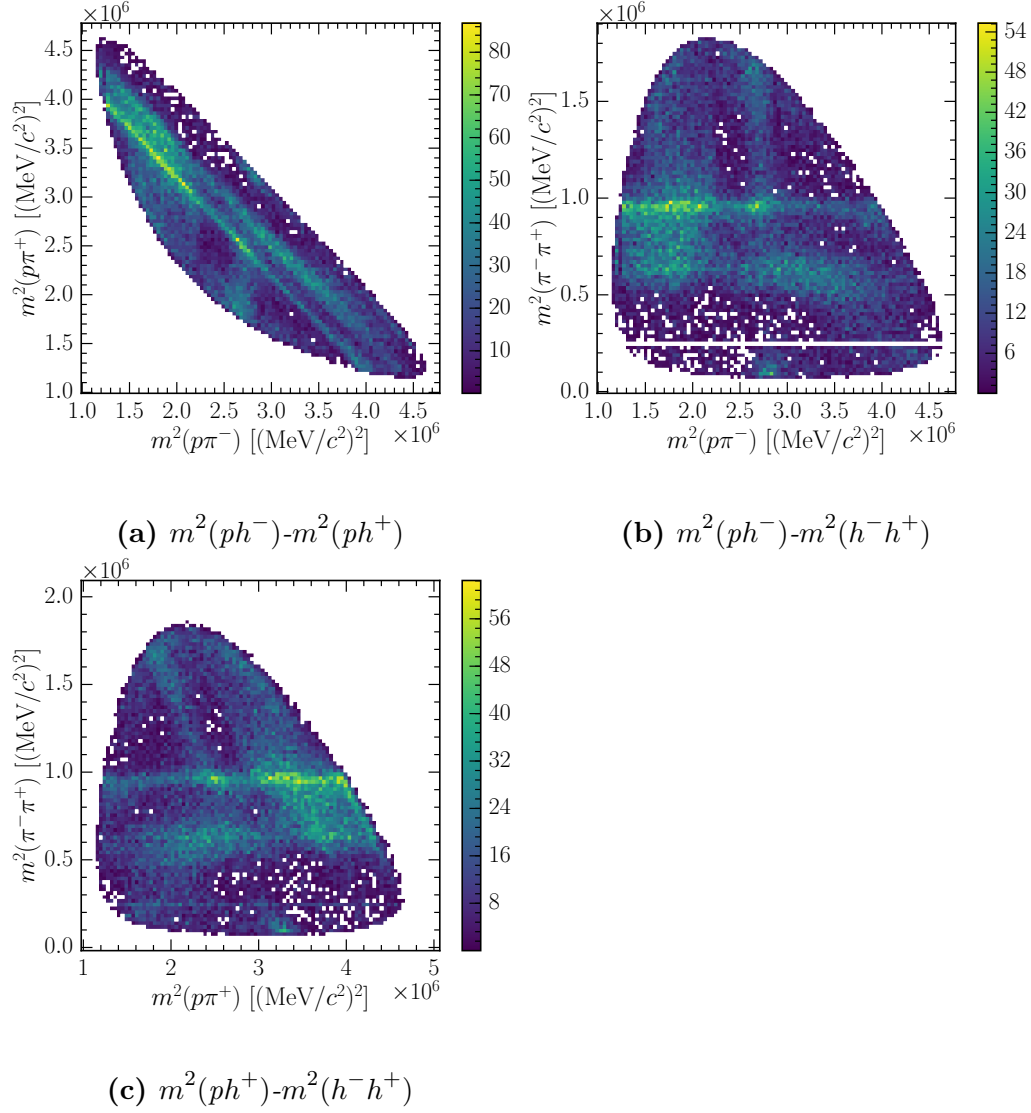


Figure 20.4: Two dimensional projections of the $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ phase space in the fully selected 2012 magnet down data, weighted by signal $sWeights$. The contents of bins where the sum of signal $sWeights$ is negative are forced to be zero.

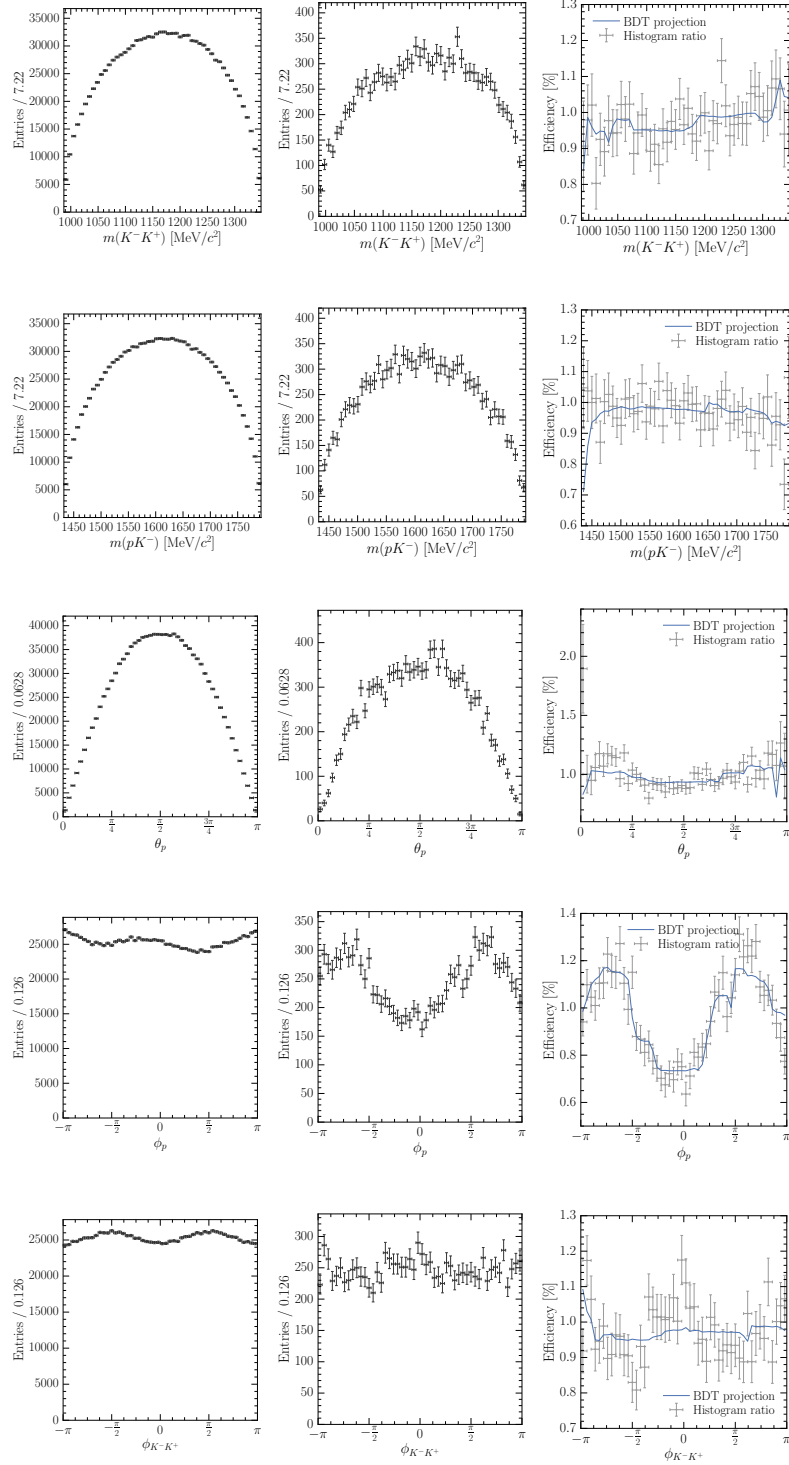


Figure 20.5: One dimensional projections of the $\Lambda_c^+ \rightarrow p K^- K^+$ phase space in the generator-level MC (left column) and in the reconstructed and selected MC (centre column), for the 2012 magnet down simulation. The right column shows a one-dimensional projection efficiency curve learnt by a BDT trained to weight the generator-level MC to look like the reconstructed and selected MC. For comparison, a ratio of the histograms is also shown with the curve.

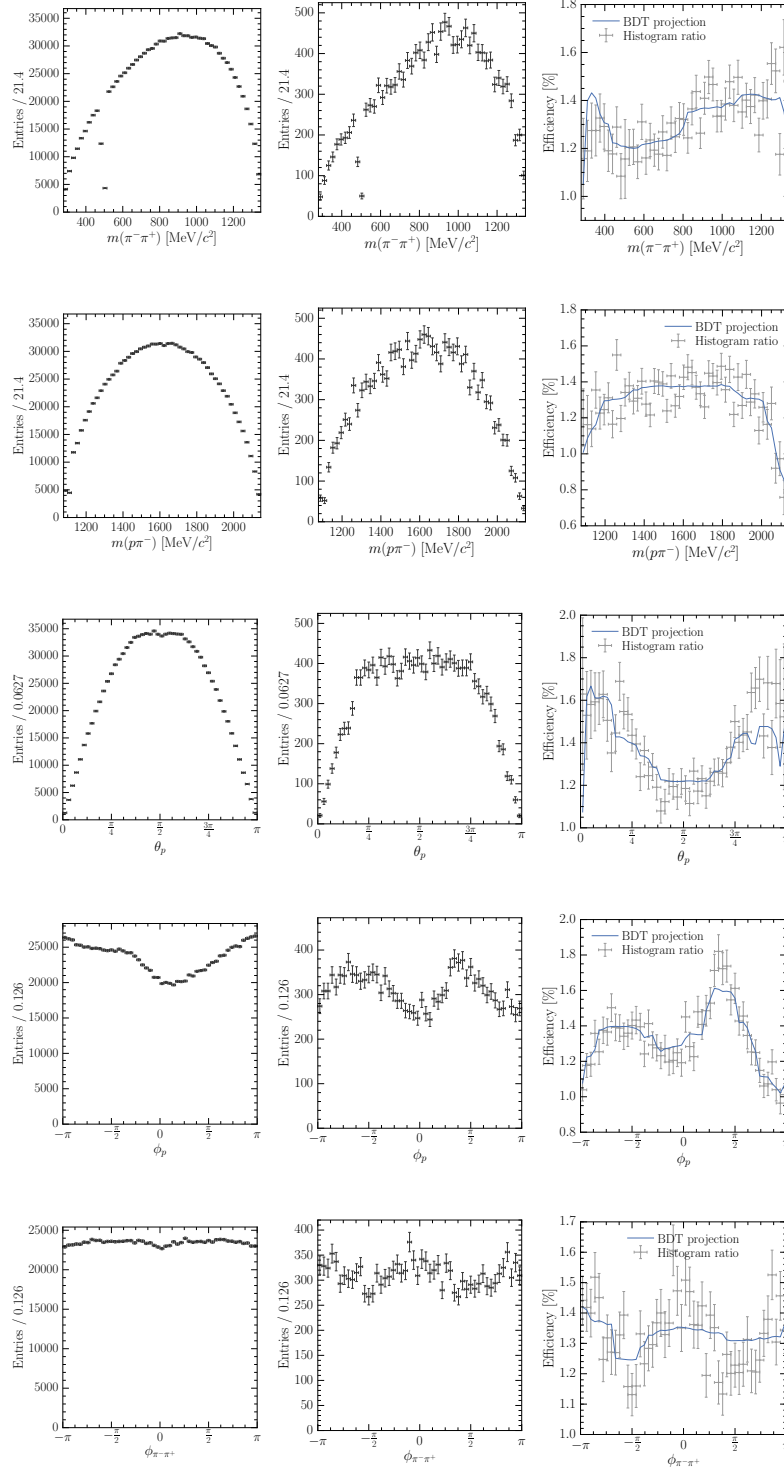


Figure 20.6: One dimensional projections of the $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ phase space in the generator-level MC (left column) and in the reconstructed and selected MC (centre column), for the 2012 magnet down simulation. The right column shows a one-dimensional projection efficiency curve learnt by a BDT trained to weight the generator-level MC to look like the reconstructed and selected MC. For comparison, a ratio of the histograms is also shown with the curve.

21 Measuring A_{CP}^{Raw}

In order to calculate ΔA_{CP} , A_{CP}^{Raw} must first be measured for each mode. This is the asymmetry in the Λ_c^+ and Λ_c^- yields, as in Equation (14.4). The yields are measured with χ^2 fits to the Λ_c^+ and Λ_c^- mass spectra, in a similar manner to the fit to the charge-combined sample discussed in Chapter 18. The weights from the kinematic weighting, from Chapter 19, and the phase space efficiency correction, from Chapter 20, are used in the fits such that the bin contents b_i , to which to fit model is compared, is

$$b_i = \sum_{j=1}^{N_i} \frac{w_j(p_T^{\Lambda_c^+}, \eta^{\Lambda_c^+}, p_T^p, \eta^p)}{\varepsilon_j(m^2(ph^-), m^2(h^-h^+), \theta_p, \phi_p, \phi_{h^-h^+})}, \quad (21.1)$$

where N_i is the number of candidate decays in the i th ph^-h^+ mass bin, w_j is per-candidate kinematic weight for the j th candidate in the i th bin, and ε_j is the efficiency. Each weight is scaled by a factor $B = \sum_i b_i / \sum_i b_i^2$ for the reasons given in Section 19.4. The variance on the bin contents δb_i is given by the effective number of entries, as in Equation (19.8)

$$\delta b_i^2 = N_{\text{Eff.}} = B \sum_i^N b_i. \quad (21.2)$$

For $p\pi^-\pi^+$ the kinematic weights are those described in Chapter 19, whilst for pK^-K^+ they are unity.

In an attempt reduce the effects of experimenter bias, the fitted central values of A_{CP}^{Raw} are blinded. To be able to assess the quality of the fit results without unblinding the result, only the fit to Λ_c^+ data is inspected. For the Λ_c^- data, only the pull distribution of the fit with respect to the data is inspected, to indicate the fit quality. The specific method used to blind the central value is discussed in Section 21.2.

21.1 Simultaneous fit

Fits are performed simultaneously on the charge-separated Λ_c^+ and Λ_c^- samples, measuring A_{CP}^{Raw} by including it as a parameter in the fit. This is advantageous as it increases the statistical sensitivity of the measurement, as all the data are considered at once, and the error analysis is taken care of by the fitter.

The fit model for each charge is created as a clone of a ‘master’ PDF, whose signal and background components are identical to those described in Chapter 18. Before creating the two models, one per charge, it is specified which parameters of the master PDF are to become charge-specific. As each Λ_c^+ decay mode may interact differently with the detector than those of the Λ_c^- , the matter and antimatter states may have different mass resolutions. This motivates the splitting of all of the signal model parameters. Likewise, the background composition can be different for the positively- and negatively-charged final states, and so the background model parameter a_0 is also split. The signal PDF, f_+ for Λ_c^+ and f_- for Λ_c^- , is then

$$f_{\pm}(x; \mu_{\pm}, w_{\pm}, \alpha_{\pm}) = \alpha_{\pm} G(x; \mu_{\pm}, \alpha_{\pm} w_{\pm}) + (1 - \alpha_{\pm}) G(x; \mu_{\pm}, \frac{\alpha_{\pm}}{w_{\pm}}), \quad (21.3)$$

and the corresponding background PDF is

$$g_{\pm}(x) = 1 + a_{\pm,0}x, \quad (21.4)$$

as in Equation (18.1) and Equation (18.3). The width parameters $\sigma_{\pm,1}$ and $\sigma_{\pm,2}$ of the signal model are defined as in Equation (18.2).

The total PDF for each charge is

$$h_{\pm}(x) = N_{\pm,\text{Sig.}} f_{\pm}(x) + N_{\pm,\text{Bkg.}} g_{\pm}(x), \quad (21.5)$$

where the signal and background yields are parameterised as

$$N_{\pm,\text{Sig.}} = \frac{1}{2} N_{\text{Sig.}} (1 \pm A_{CP}^{\text{Raw}}), \quad (21.6)$$

$$N_{\pm,\text{Bkg.}} = \frac{1}{2} N_{\text{Bkg.}} (1 \pm A_{\text{Bkg.}}^{\text{Raw}}). \quad (21.7)$$

Here, $N_{\text{Sig.}}$ and $N_{\text{Bkg.}}$ are the signal and background yields for the *combined* Λ_c^+ and Λ_c^- sample. This parameterisation of the charge-specific yields permits A_{CP}^{Raw} to be measured directly from the fit.

The fully selected data are used, with one simultaneous fit performed for each year and magnet polarity sub-sample. The loss function is defined as a χ^2 , which is minimised with MIGRAD and a covariance matrix is computed with HESSE [108, 109].

21.2 Blinding

The blinding is done by fitting a variable as the *transform* of A_{CP}^{Raw} . The transformation itself can be adding some random offset to the central value, or changing its order of magnitude, but it is applied in such a way as to leave the uncertainty on the parameter estimate unchanged. Only the transformed value is reported, leaving the unblinded value hidden.

This analysis applies a linear blinding offset, transforming the unblinded central value x as

$$x' = x + \beta\xi, \quad (21.8)$$

where ξ is a number sampled from a normal distribution with mean zero and unit width, and β is a known scaling parameter, which is chosen to be 0.25. The number ξ is chosen deterministically from a string in the fitting program, such that the same input string always gives the same ξ , but in a way that the smallest changes in the input change the output unpredictably. In this way, the string is effectively the seed for a random number generator, and ξ is the first number returned by the generator. Different strings are used for different modes, but are the same for the fits to the different data sub-samples.

As the blinding offsets are different for each mode, the difference between $A_{CP}^{\text{Raw}}(pK^-K^+)$ and $A_{CP}^{\text{Raw}}(p\pi^-\pi^+)$, that is ΔA_{CP} , is also blinded. However, the difference between two blinded values of ΔA_{CP} is equal to the difference between the unblinded values. This is useful for systematic studies, which will be described in Chapter 23. Similarly, as the blinding parameters are

Table 21.1: *Model parameters as determined in the preliminary fit to the 2012 magnet down subset of the pK^-K^+ data.*

Parameter	Value
A_{Bkg}	$(23.0 \pm 9.5) \times 10^{-3}$
A_{Raw}^{CP}	$(-7.2 \pm 1.3) \times 10^{-2}$
N_{Bkg}	$(142.4 \pm 1.3) \times 10^2$
N_{Sig}	$(88.9 \pm 1.1) \times 10^2$
$a_{0,\text{Bkg}}^+$	$(-2.7 \pm 2.0) \times 10^{-2}$
α_{Sig}^+	$(75.3 \pm 2.7) \times 10^{-2}$
μ_{Sig}^+	2286.957 ± 0.071
w_{Sig}^+	$(325.8 \pm 6.1) \times 10^{-2}$
$a_{0,\text{Bkg}}^-$	$(-3.9 \pm 2.1) \times 10^{-2}$
α_{Sig}^-	$(73.9 \pm 2.5) \times 10^{-2}$
μ_{Sig}^-	2287.084 ± 0.074
w_{Sig}^-	$(323.6 \pm 6.3) \times 10^{-2}$

the same for all data sub-samples, it is still meaningful to compare blinded values of A_{CP}^{Raw} and ΔA_{CP} between them to check for consistency, as the differences are not blinded.

In addition to blinding the values of A_{CP}^{Raw} and ΔA_{CP} , the plots of the fits to the Λ_c^- data are blind (as comparing Λ_c^+ and Λ_c^- fits side-by-side would allow one to infer the size of the asymmetry), as are plots of the data and model asymmetry in bins of Λ_c^+ mass. The pull plots and χ^2 values are not blinded in the fit presentations as they do not indicate the asymmetry value.

21.3 Fit results

Fits to the pK^-K^+ and $p\pi^-\pi^+$ 2012 magnet down data are shown in Figures 21.1 and 21.2. The corresponding correlation matrices for the fits are given in Figure 21.3, and the parameter values are given in Tables 21.1 and 21.2.

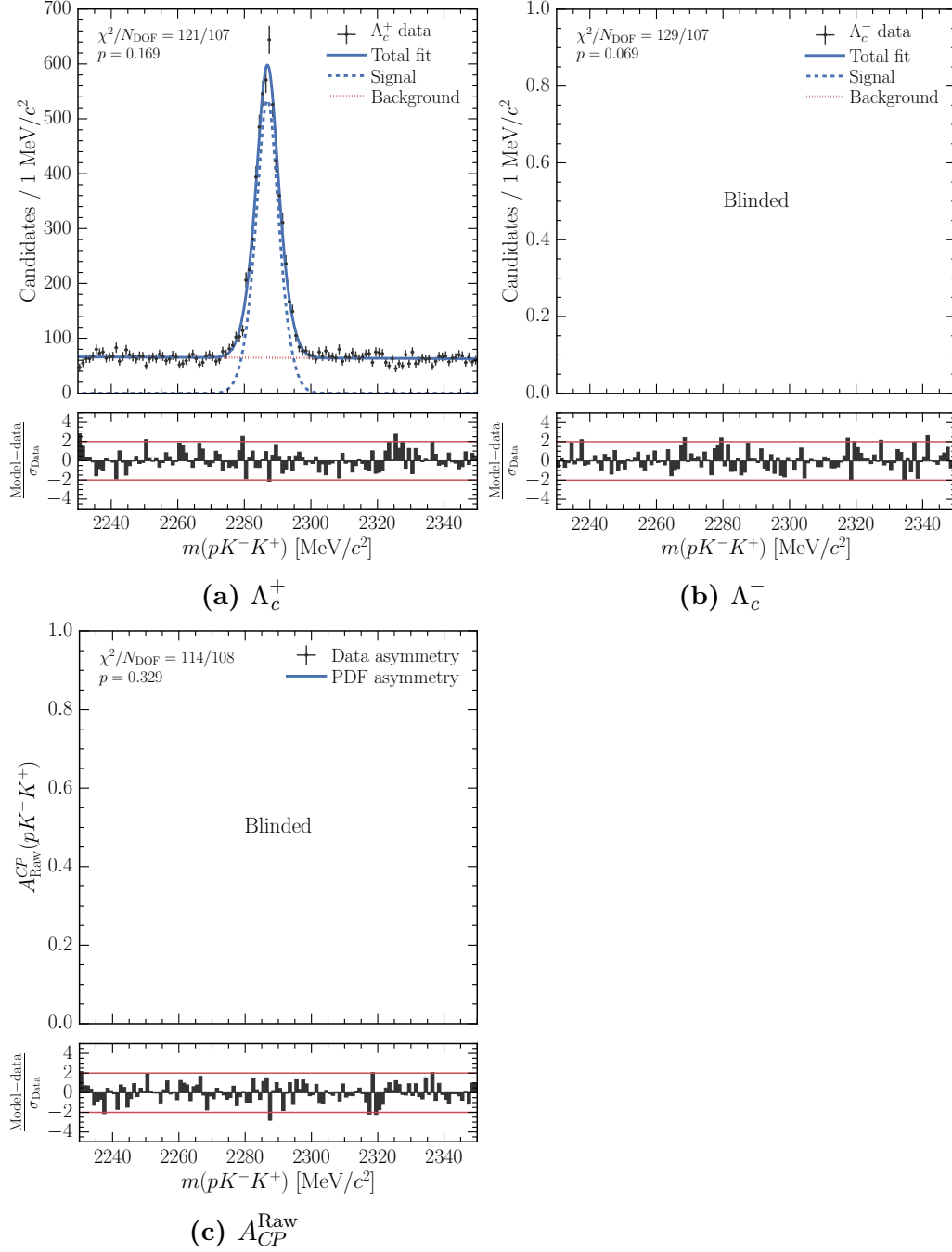


Figure 21.1: Results of the simultaneous fit to the pK^-K^+ 2012 magnet down dataset. The difference between the Λ_c^+ data and model (a) and the Λ_c^- data and model (b) is shown in the bottom plot (c). The full offline selection is applied. The solid blue line is the total fit to the data in black points, and the dotted red and dashed blue lines are the background and signal components, respectively. Below each fit is a pull plot, showing the difference between the total fit model and the data in each bin, normalised by the Poisson uncertainty on the number of entries in that bin.

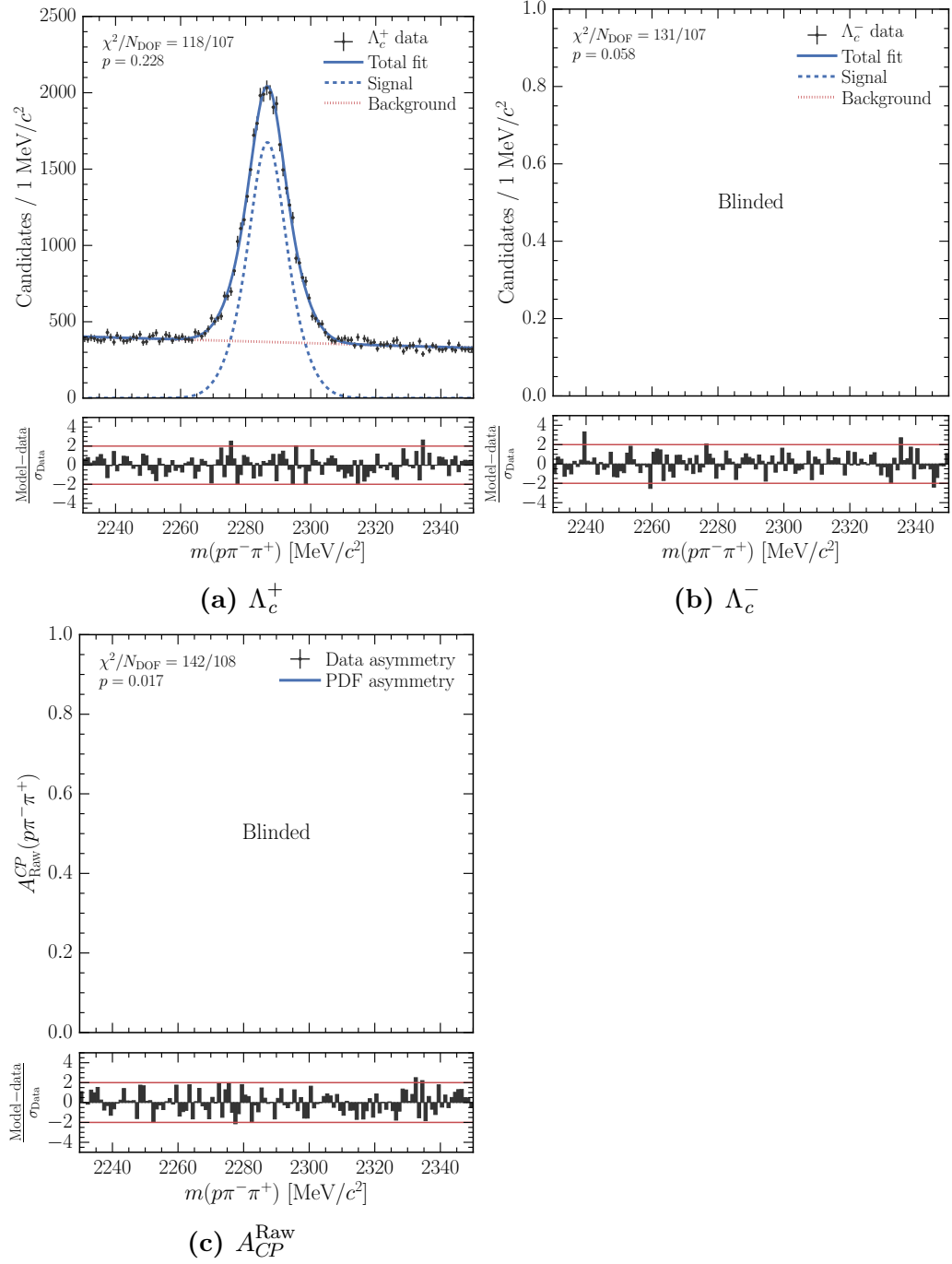


Figure 21.2: Results of the simultaneous fit to the $p\pi^-\pi^+$ 2012 magnet down dataset. The difference between the Λ_c^+ data and model (a) and the Λ_c^- data and model (b) is shown in the bottom plot (c). The full offline selection is applied. The solid blue line is the total fit to the data in black points, and the dotted red and dashed blue lines are the background and signal components, respectively. Below each fit is a pull plot, showing the difference between the total fit model and the data in each bin, normalised by the Poisson uncertainty on the number of entries in that bin.

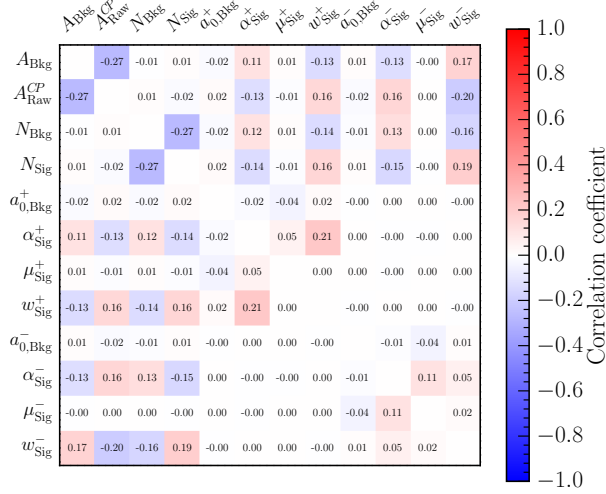
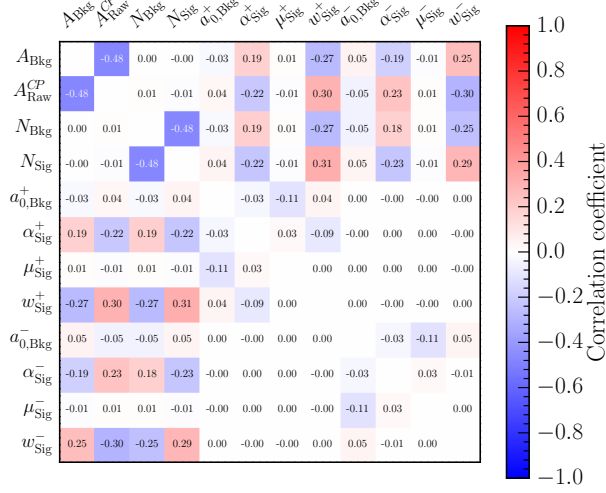
(a) pK^-K^+ (b) $p\pi^-\pi^+$

Figure 21.3: Correlation matrices for the fit parameters used in the simultaneous fit to the Λ_c^+ and Λ_c^- mass spectra in the fully selected 2012 magnet down dataset for pK^-K^+ (a) and $p\pi^-\pi^+$ (b). The corresponding fits are shown in Figures 21.1 and 21.2.

Table 21.2: *Model parameters as determined in the preliminary fit to the 2012 magnet down subset of the $p\pi^-\pi^+$ data.*

Parameter	Value
A_{Bkg}	$(14.1 \pm 5.3) \times 10^{-3}$
A_{Raw}^{CP}	$(-84.9 \pm 8.1) \times 10^{-3}$
N_{Bkg}	$(639.5 \pm 3.4) \times 10^2$
N_{Sig}	$(325.2 \pm 2.6) \times 10^2$
$a_{0,\text{Bkg}}^+$	$(-10.3 \pm 1.0) \times 10^{-2}$
α_{Sig}^+	$(73.6 \pm 1.4) \times 10^{-2}$
μ_{Sig}^+	2286.696 ± 0.07
w_{Sig}^+	$(576.4 \pm 6.9) \times 10^{-2}$
$a_{0,\text{Bkg}}^-$	$(-11.9 \pm 1.0) \times 10^{-2}$
α_{Sig}^-	$(75.2 \pm 1.6) \times 10^{-2}$
μ_{Sig}^-	2286.717 ± 0.075
w_{Sig}^-	$(591.0 \pm 7.2) \times 10^{-2}$

21.4 Validation

It has been shown in Section 18.2 that the measurements of the signal and background yields in the fits to the charge-combined data samples are not biased. It remains to validate the simultaneous fit that has been described in this chapter. This is done in using real data. Using real data, the fits are performed with the charge of the Λ_c^+ candidate being randomly assigned. The true values of A_{CP}^{Raw} and $A_{\text{Bkg}}^{\text{Raw}}$ should then be zero. This checks for mistakes in the implementation of the fitter, and that the uncertainty on those parameters is not under- or overestimated.

Null test

Fits as described in this chapter are performed on real data where the charge of the Λ_c^+ candidate is randomly assigned, with there being an equal probability of any one decay being assigned a positive or a negative charge. One thousand permutations of the charge assignments are tested for each mode and data sub-sample, the fit is performed for each permutation and the value of A_{CP}^{Raw} is recorded. As with the studies described in Section 18.2,

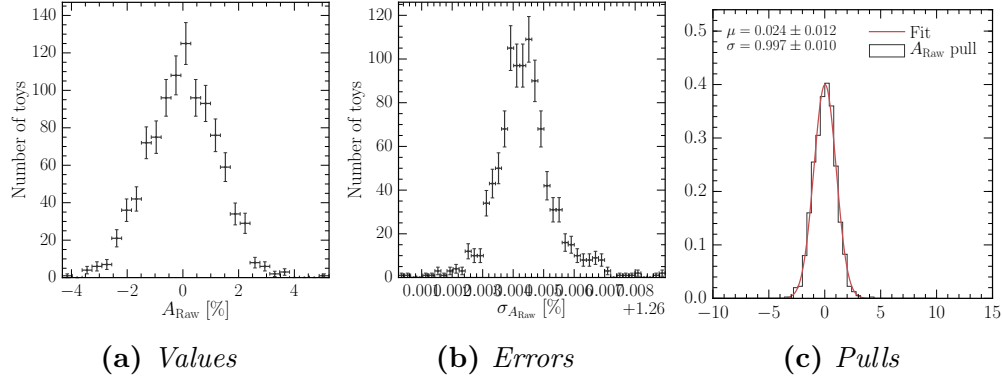


Figure 21.4: Validation of the simultaneous fits to the Λ_c^+ and Λ_c^- mass spectra in the 2012 magnet down pK^-K^+ dataset, using randomly assigned Λ_c^+ charges. The plots show the distribution of the central values (a), uncertainties (b), and pulls (c) of the A_{CP}^{Raw} parameter, assuming that the true value is zero. The pull distribution is overlaid with a fit of a Gaussian distribution with mean μ and width σ .

the pull distributions of A_{CP}^{Raw} should have a width of one, and the unblinded value should have a mean of zero.

The true charge of the Λ_c^+ candidates do not enter the fit, and so the charge permutation experiments are performed without blinding. The results are shown in Figures 21.4 and 21.5 for the 2012 magnet down pK^-K^+ and $p\pi^-\pi^+$ data. The width of the pull distribution of A_{CP}^{Raw} for both is consistent with unity, and the means are consistent with zero, indicating that the fit is unbiased.

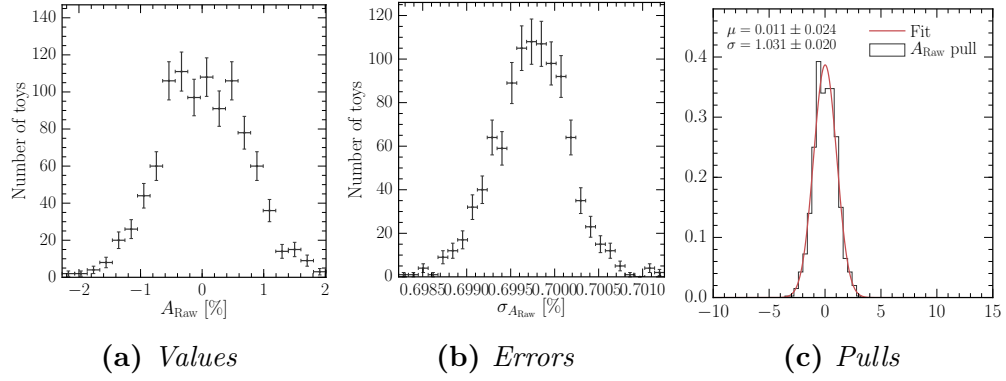


Figure 21.5: Validation of the simultaneous fits to the Λ_c^+ and Λ_c^- mass spectra in the 2012 magnet down $p\pi^-\pi^+$ dataset, using randomly assigned Λ_c^+ charges. The plots show the distribution of the central values (a), uncertainties (b), and pulls (c) of the A_{CP}^{Raw} parameter, assuming that the true value is zero. The pull distribution is overlaid with a fit of a Gaussian distribution with mean μ and width σ .

22 Results

The difference in the raw asymmetries of the pK^-K^+ and $p\pi^-\pi^+$ data defines ΔA_{CP}

$$\Delta A_{CP} = A_{CP}^{\text{Raw}}(pK^-K^+) - A_{CP}^{\text{Raw}}(p\pi^-\pi^+).$$

The values of, and uncertainties on, A_{CP}^{Raw} are measured by the weighted fits described in Chapter 21. In this section, the combinations of A_{CP}^{Raw} are made within and across the data-taking years (2011 and 2012) and magnet polarities (magnet up and down), and the measured values of A_{CP}^{Raw} and ΔA_{CP} are reported.

22.1 Combining measurements

The values of A_{CP}^{Raw} for a given decay mode and data-taking year and polarity are measured in the fits. Within a given year ζ , $\zeta \in [2011, 2012]$, the average value of the A_{CP}^{Raw} for each mode is defined as the arithmetic mean of the two values of A_{CP}^{Raw} measured on the magnet up and magnet down data. The average value of ΔA_{CP} within a year is then defined as

$$\Delta A_{CP}(\zeta) = A_{CP}^{\text{Raw}}(\zeta, pK^-K^+) - A_{CP}^{\text{Raw}}(\zeta, p\pi^-\pi^+). \quad (22.1)$$

Across both years, the values of A_{CP}^{Raw} are computed as the weighted sum of the per-year values, where the weights are the inverses of the squares of the statistical uncertainties σ on the measurements. This quantity can be computed using the polarity-averaged A_{CP}^{Raw} measurements, and it can also be computed per polarity, for example to compute the average magnet down result across 2011 and 2012.

The measured values of $A_{CP}^{\text{Raw}}(pK^-K^+)$, $A_{CP}^{\text{Raw}}(p\pi^-\pi^+)$, and ΔA_{CP} are given in Table 22.1, where the central values are blinded. Plots of the three measurements for each data sub-sample are given in Figure 22.1, where a χ^2 fit of a constant function to the measurements is also shown. The values are compatible across years and magnet polarities.

The value of ΔA_{CP} averaged across all magnet polarities and years is measured to be

$$\Delta A_{CP} = (1.20 \pm 0.92) \%$$

The central value is blinded.

Unweighted asymmetries

The results computed without the inclusion of the kinematic weights and phase space efficiency corrections are given in Table 22.2, where the central values are blinded. The average value of ΔA_{CP} in these results is

$$\Delta A_{CP} = (1.29 \pm 0.85) \%$$

Assuming the weighting does correctly equalise the kinematics, the small change in central value from the weighted measurement indicates that the differences between the modes before the weighting were small. The 8% reduction in the statistical uncertainty is due to the effective increase in the $p\pi^-\pi^+$ sample size when unweighted, and illustrates that the uncertainty on ΔA_{CP} is dominated by the pK^-K^+ sample size.

Table 22.1: *Measured asymmetries, in percent, for each data sub-sample and combination of sub-samples. The computation of the combinations, “2011 + 2012” and “Average”, is defined in Section 22.1. The central values of all measurements are blinded.*

Year	Polarity	$A_{CP}^{\text{Raw}}(pK^-K^+)$	$A_{CP}^{\text{Raw}}(p\pi^-\pi^+)$	ΔA_{CP}
2011	Up	-8.45 ± 2.29	-11.70 ± 1.40	3.25 ± 2.68
2011	Down	-5.82 ± 1.98	-7.02 ± 1.16	1.20 ± 2.29
2011	Average	-7.14 ± 1.51	-9.36 ± 0.91	2.22 ± 1.76
2012	Up	-7.52 ± 1.30	-7.78 ± 0.83	0.26 ± 1.54
2012	Down	-7.17 ± 1.28	-8.49 ± 0.81	1.32 ± 1.52
2012	Average	-7.35 ± 0.91	-8.14 ± 0.58	0.79 ± 1.08
2011 + 2012	Up	-7.75 ± 1.13	-8.80 ± 0.72	1.06 ± 1.33
2011 + 2012	Down	-6.77 ± 1.08	-8.00 ± 0.67	1.23 ± 1.27
2011 + 2012	Average	-7.29 ± 0.78	-8.49 ± 0.49	1.20 ± 0.92

Table 22.2: *Measured asymmetries, in percent, for each data sub-sample and combination of sub-samples without the inclusion of the kinematic weights and phase space efficiency correction. The computation of the combinations, “2011 + 2012” and “Average”, is defined in Section 22.1. The central values of all measurements are blinded.*

Year	Polarity	$A_{CP}^{\text{Raw}}(pK^-K^+)$	$A_{CP}^{\text{Raw}}(p\pi^-\pi^+)$	ΔA_{CP}
2011	Up	-8.73 ± 2.29	-11.88 ± 1.05	3.15 ± 2.52
2011	Down	-5.13 ± 1.97	-6.24 ± 0.89	1.10 ± 2.16
2011	Average	-6.93 ± 1.51	-9.06 ± 0.69	2.13 ± 1.66
2012	Up	-7.98 ± 1.28	-7.78 ± 0.60	-0.20 ± 1.41
2012	Down	-6.77 ± 1.26	-8.94 ± 0.59	2.17 ± 1.39
2012	Average	-7.37 ± 0.90	-8.36 ± 0.42	0.99 ± 0.99
2011 + 2012	Up	-8.16 ± 1.12	-8.79 ± 0.52	0.64 ± 1.23
2011 + 2012	Down	-6.29 ± 1.06	-8.11 ± 0.49	1.82 ± 1.17
2011 + 2012	Average	-7.26 ± 0.77	-8.55 ± 0.36	1.29 ± 0.85

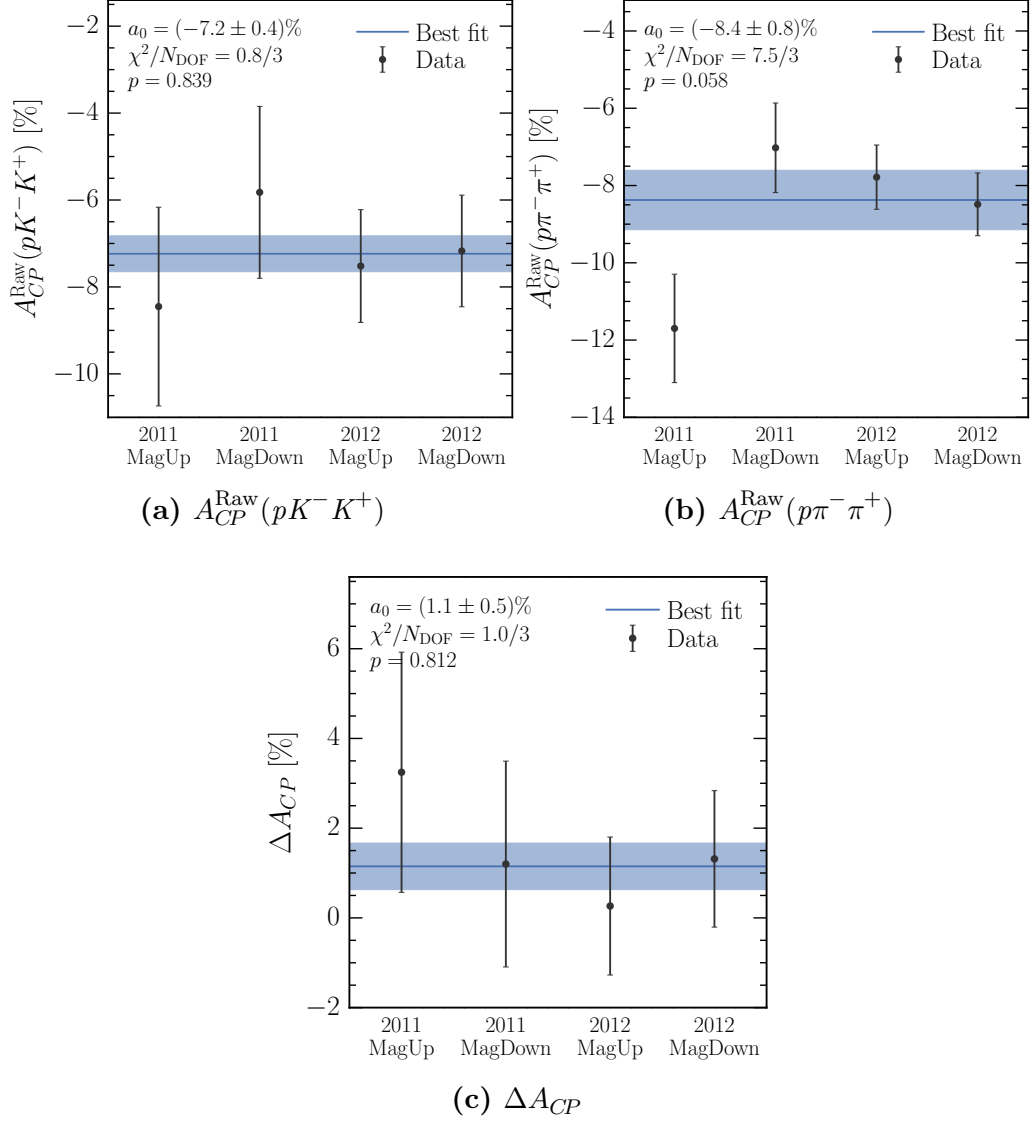


Figure 22.1: Values of the asymmetries $A_{CP}^{\text{Raw}}(pK^-K^+)$ (a), $A_{CP}^{\text{Raw}}(p\pi^-\pi^+)$ (b), and ΔA_{CP} (c), for the four data sub-samples. The central values of the measurements are blinded. For each asymmetry, a χ^2 fit of a constant function to the four data points is also shown, along with the best fit value and uncertainty, χ^2 per degree of freedom (DOF), and p -value.

23 Systematic uncertainties

Three sources of systematic uncertainty are considered in this analysis: the choice of model used to extract the Λ_c^+ and Λ_c^- signal yields; the failure of the kinematic weighting procedure to fully equalise the Λ_b^0 , muon, and proton kinematics between the modes; and the breaking of the assumption that the Λ_c^+ candidates originate from the $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X$ decay. The effect of these on the measurement of ΔA_{CP} will be discussed in this chapter and, where appropriate, systematic uncertainties will be computed and assigned.

23.1 Fit model

The models used for the signal and combinatorial background components in the mass fits, presented in Chapters 18 and 21, do describe the data well. However, other functional forms may also describe the data well but give different yields, which may change the value of ΔA_{CP} . To assess the extent to which different models can affect the measurement, the yields are extracted using the method of sideband subtraction.

Region definitions and method validity

A sideband subtraction is applied to the data as an alternative method to measure the pK^-K^+ and $p\pi^-\pi^+$ signal yields in the charge-separated samples.

The signal region is defined as a window $20 \text{ MeV}/c^2$ wide for the pK^-K^+ data and $30 \text{ MeV}/c^2$ wide for $p\pi^-\pi^+$, centred on the nominal value of the Λ_c^+ mass of $2286.46 \text{ MeV}/c^2$ [12]. The wider $p\pi^-\pi^+$ window accounts for the

wider signal shape, which is itself due to the larger energy release, or Q value, of the $p\pi^-\pi^+$ decay. The two sideband regions, each $10 \text{ MeV}/c^2$ wide for pK^-K^+ and $15 \text{ MeV}/c^2$ wide for $p\pi^-\pi^+$, are chosen such that the centres are $\pm 3/2$ signal window widths away from the signal window centre. The signal and sideband regions are shown in Figure 23.1, with the 2012 magnet down data overlaid for reference.

The assumption of the sideband subtraction technique that the background PDF is linear is justified by the good description of the background by a linear function in the nominal fits. The lack of any visible physics backgrounds after the full selection described in Chapter 17 suggests that the sidebands are purely combinatorial, and that no peaking structures are present in the signal region other than the that of the signal. Finally, the assumption that the signal contribution to the data in the sidebands is negligible is checked using MC, where 0.7 % of the $p\pi^-\pi^+$ truth-matched signal data are in the sidebands, and for pK^-K^+ the fraction of 0.3 %. These fractions are much smaller than the relative statistical uncertainties on the signal yields shown in Tables 18.1 and 18.2, and so the level of contamination is considered to be negligible.

Results

For each data sub-sample, the pK^-K^+ and $p\pi^-\pi^+$ samples are split by the charge of the Λ_c^+ and the sideband subtraction technique is used to measure the yields. The yield asymmetry is then

$$A_{CP}^{\text{Raw}}(f) = \frac{(N_S(\Lambda_c^+) - N_{\text{SB}}(\Lambda_c^+)) - (N_S(\Lambda_c^-) - N_{\text{SB}}(\Lambda_c^-))}{(N_S(\Lambda_c^+) - N_{\text{SB}}(\Lambda_c^+)) + (N_S(\Lambda_c^-) - N_{\text{SB}}(\Lambda_c^-))}, \quad (23.1)$$

where $N_{\text{S(SB)}}(\Lambda_c^+)$ is the number of candidates in the signal (sideband) region in the Λ_c^+ sample, and $N_{\text{S(SB)}}(\Lambda_c^-)$ is the same but in the Λ_c^- sample. The counts are assumed to be Poisson-distributed, and so the uncertainty δN on each count is taken to be $\delta N = \sqrt{N}$.

The difference ΔA_{CP} between A_{CP}^{Raw} for pK^-K^+ and $p\pi^-\pi^+$ is then computed using the values obtain with Equation (23.1), and the difference is computed between those values and the nominal measurements given in

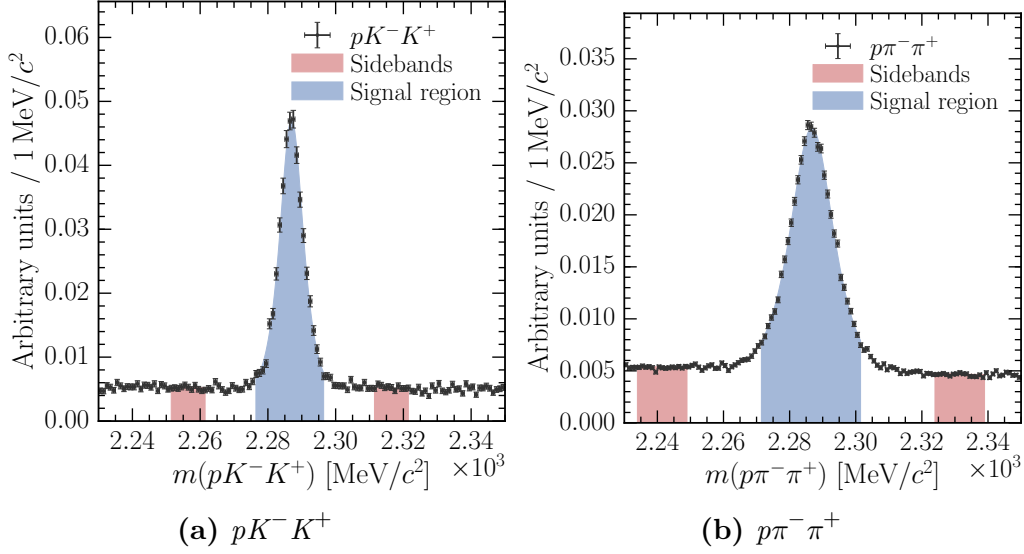


Figure 23.1: Definitions of the signal and sideband regions in the Λ_c^+ mass spectrum for pK^-K^+ (a) and $p\pi^-\pi^+$ (b). The 2012 magnet down data are overlaid for reference.

Table 22.1. The uncertainty on A_{CP}^{Raw} is found by propagating the uncertainties on the candidate counts using an MC error propagation, and the values are assumed to be fully correlated with those from the nominal procedure. The differences for the measurements in all data sub-samples, and their averages across years and magnet polarities, are presented in Table 23.1. Almost all of the deviations from the nominal values are significant, and so they taken as a systematic uncertainties on the measurements of ΔA_{CP} . The systematic uncertainty is 0.26 % on the nominal result quoted in Section 22.1.

23.2 Residual background asymmetries

The kinematic weighting discussed in Chapter 19 does not completely remove all differences between the Λ_b^0 , muon, and proton kinematics between the pK^-K^+ and $p\pi^-\pi^+$ datasets, nor does it remove all K^-/K^+ and π^-/π^+ differences within the decay modes. Therefore, contributions from the respective background asymmetries will still contribute to ΔA_{CP} . A study is performed to quantify this contamination. Firstly, the available measure-

Table 23.1: Differences between asymmetries measured with the nominal fit method and the sideband subtraction, measured for each data sub-sample and combination of sub-samples. The uncertainties quoted assume the values found by the two methods are fully correlated. The computation of the combinations, “2011 + 2012” and “Average”, is defined in Section 22.1.

Year	Polarity	$A_{CP}^{\text{Raw}}(pK^-K^+)$ (%)	$A_{CP}^{\text{Raw}}(p\pi^-\pi^+)$ (%)	ΔA_{CP} (%)
2011	Up	0.73 ± 0.06	-0.95 ± 0.21	1.68 ± 0.03
2011	Down	-0.71 ± 0.09	-0.14 ± 0.16	-0.57 ± 0.02
2011	Average	0.01 ± 0.05	-0.55 ± 0.13	0.56 ± 0.01
2012	Up	-0.00 ± 0.06	-0.29 ± 0.14	0.28 ± 0.01
2012	Down	0.43 ± 0.05	0.48 ± 0.14	-0.05 ± 0.02
2012	Average	0.21 ± 0.04	0.10 ± 0.10	0.12 ± 0.01
2011 + 2012	Up	0.18 ± 0.05	-0.48 ± 0.12	0.66 ± 0.01
2011 + 2012	Down	0.10 ± 0.04	0.32 ± 0.11	-0.22 ± 0.01
2011 + 2012	Average	0.16 ± 0.03	-0.10 ± 0.08	0.26 ± 0.01

ments of the detection and production asymmetries are enumerated, and then their application to this analysis is described.

The asymmetries considered are those that arise from the differences in the matter and antimatter cross-sections with the detector, called detection asymmetries. The trigger asymmetry for muons from b decays is known to be small [110], and is neglected here.

The pion detection asymmetry is measured using $\sqrt{s} = 7$ TeV data [70] by a method of partial reconstruction of D^{*+} -tagged $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decays. In this, the delta mass variable, defined similarly to that in Equation (9.1), provides a good signal-to-background discrimination even if the pion is not required to have been reconstructed. The ratio of the signal yields before and after the explicit reconstruction gives the pion detection efficiency, and this is performed separately for each pion charge in bins of pion kinematics. The asymmetries are shown in Figure 23.2 as a function of pion momentum. The central values of the relative asymmetries in bins of pion momentum are small, generally less than 2 %, and are very close to zero when averaged across the magnet polarities. The integrated ratios are measured to the 0.9914 ± 0.0040 and 1.0045 ± 0.0034 for magnet up and magnet down configurations, respectively. Here, these values are considered to be compatible with zero, and so contributions from a pion detection asymmetry are neglected.

The kaon detection asymmetry is measured using the full Run 1 dataset with $D^+ \rightarrow K^- \pi^+ \pi^+$ and $D^+ \rightarrow \bar{K}^0 \pi^+$ decays [90]. The raw asymmetries between the two decays are subtracted, as is the contribution from the $\bar{K}^0$ interaction asymmetry, to measure $A_D(K^- \pi^+)$. The kaon material cross-section asymmetry is known to be much larger than that for pions [12], and so $A_D(K^- \pi^+)$ is assumed to be dominated by the kaon detection asymmetry. The asymmetry as a function of kaon momentum, averaged across magnet polarities, is shown in Figure 23.3, where it is seen to be significantly larger in magnitude than the pion asymmetry in Figure 23.2 and significantly different from zero.

The muon detection asymmetry is measured using the tag-and-probe technique described in Section 10.2 with $J/\psi \rightarrow \mu^+ \mu^-$ [111]. The track reconstruction efficiency is computed separately for μ^- and μ^+ and an

asymmetry is formed, shown in Figure 23.4. The asymmetries are small, of the order of 0.1 %.

Measurements of the production asymmetry from unequal Λ_b^0 and $\bar{\Lambda}_b^0$ cross-sections have been made by LHCb [97], where the asymmetry as a function of p_T was found to be negligible over both 2011 and 2012, but as a function of rapidity y was found to be

$$A_P^{\Lambda_b^0} = (0.058 \pm 0.014)(y - \langle y \rangle), \quad (23.2)$$

where $\langle y \rangle = 3.1$ is the average Λ_b^0 rapidity in the sample used. These values are compatible between the 2011 and 2012 data. The same analysis also measured the proton detection asymmetry using simulation, assigning a systematic uncertainty to the result based on the relative disagreement between the data and the MC when evaluating the *kaon* detection asymmetry. The results are shown in Table 23.2. Simulations using both the pK^-K^+ and the $p\pi^-\pi^+$ samples defined in Section 16.1 agree with these results, and are shown in Figure 23.5.

Evaluation

The effect on the value of ΔA_{CP} of the residual background asymmetry for a single asymmetry A due to some particle X is given by a shift ΔA , defined as

$$\Delta A^X = \frac{1}{N_{\text{Sig.}}^{pK^-K^+}} \sum_i N_{\text{Sig.}}^{pK^-K^+,i} A_i - \frac{1}{N_{\text{Sig.}}^{p\pi^-\pi^+}} \sum_i N_{\text{Sig.}}^{p\pi^-\pi^+,i} A_i, \quad (23.3)$$

where $N_{\text{Sig.}}^f$ is the total number of signal candidates in the $\Lambda_c^+ \rightarrow f$ sample; i is the bin index of the asymmetry parameter, such as Λ_b^0 rapidity; and $N_{\text{Sig.}}^{f,i}$ is the number of signal candidates in the $\Lambda_c^+ \rightarrow f$ sample that fall in the i th asymmetry bin. The number of signal candidates is computed as the sum of signal sWeights, weighted by kinematic weights and phase space efficiency corrections as in Equation (21.1). One such shift is defined for each of the muon and proton detection asymmetries and the Λ_b^0 production asymmetry, $\Delta A_P^{\Lambda_b^0}$, ΔA_D^p , and $\Delta A_D^{\mu^-}$. The shift due to the residual kaon

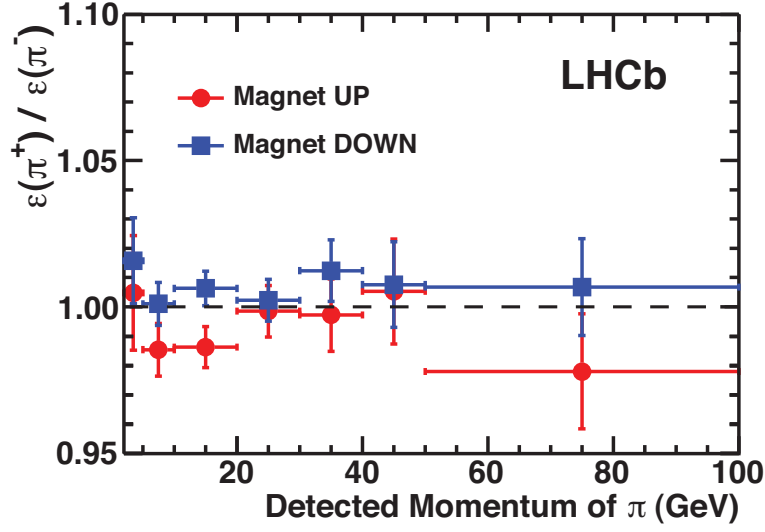


Figure 23.2: Ratio of pion detection asymmetries as a function of pion momentum [70]. The uncertainties on the data points are statistical.

detection asymmetry is similarly defined

$$\Delta A_D^K = \frac{1}{N_{\text{Sig.}}^{pK^-K^+}} \left(\sum_i N_{\text{Sig.}}^{pK^-K^+,K^-} A_i - \sum_i N_{\text{Sig.}}^{pK^-K^+,K^+} A_i \right), \quad (23.4)$$

where $N_{\text{Sig.}}^{pK^-K^+,K^\pm}$ is the number of signal K^+/K^- particles in the i th asymmetry bin.

The total shift of ΔA_{CP} , ΔA is given by the sum of the individual shifts

$$\Delta A = \Delta A_P^{\Lambda_b^0} + \Delta A_D^p + \Delta A_D^{\mu^-} + \Delta A_D^K. \quad (23.5)$$

These shifts for each year and data-taking condition are given in Table 23.3, along with the shifts for the polarity-integrated datasets. All of the values found are below one per mille, with the proton detection asymmetry providing the largest contribution. Conservatively, a 0.1 % systematic uncertainty is assigned to the measurement of ΔA_{CP} based on the residual background uncertainty.

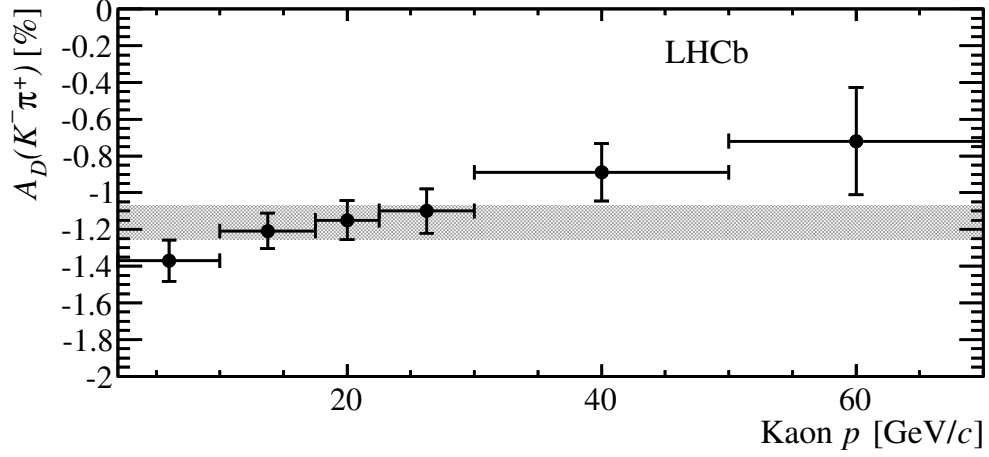


Figure 23.3: Kaon detection asymmetry $A_D(K^- \pi^+)$ as a function of kaon momentum averaged across magnet polarities [90]. The average of the measurements is shown as a grey band. The error bars on the data points include statistical and systematic uncertainties.

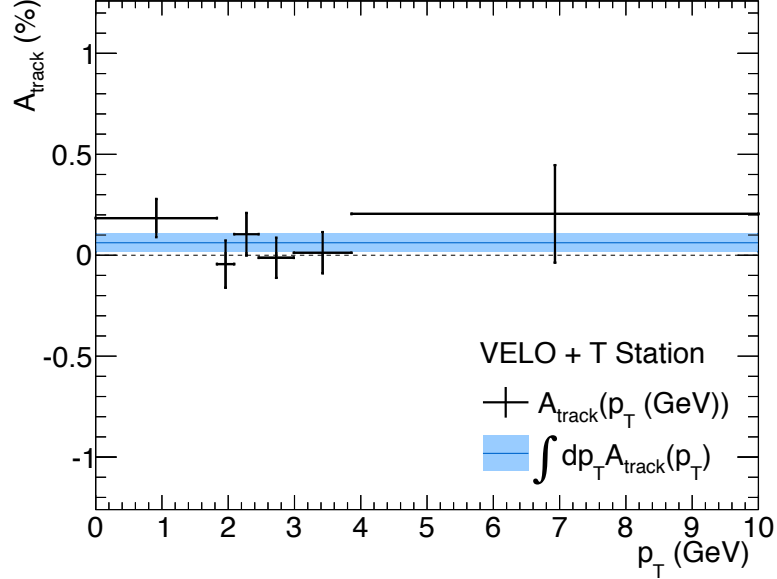


Figure 23.4: Muon detection asymmetry as a function of muon p_T averaged across magnet polarities in $\sqrt{s} = 8$ TeV data [111]. The average of the measurements is shown as a blue band. The uncertainties on the data points are statistical.

Table 23.2: Proton detection asymmetries measured using simulation in bins of proton momentum [97].

Momentum bin (GeV/c)			Asymmetry (%)
0	$< p <$	10	4.38 ± 0.29
10	$< p <$	17.5	2.46 ± 0.31
17.5	$< p <$	22.5	1.90 ± 0.47
22.5	$< p <$	30	1.12 ± 0.47
30	$< p <$	50	1.20 ± 0.42
50	$< p <$	70	0.60 ± 0.68
70	$< p <$	100	-0.23 ± 0.91
100	$< p <$	150	0.4 ± 1.4

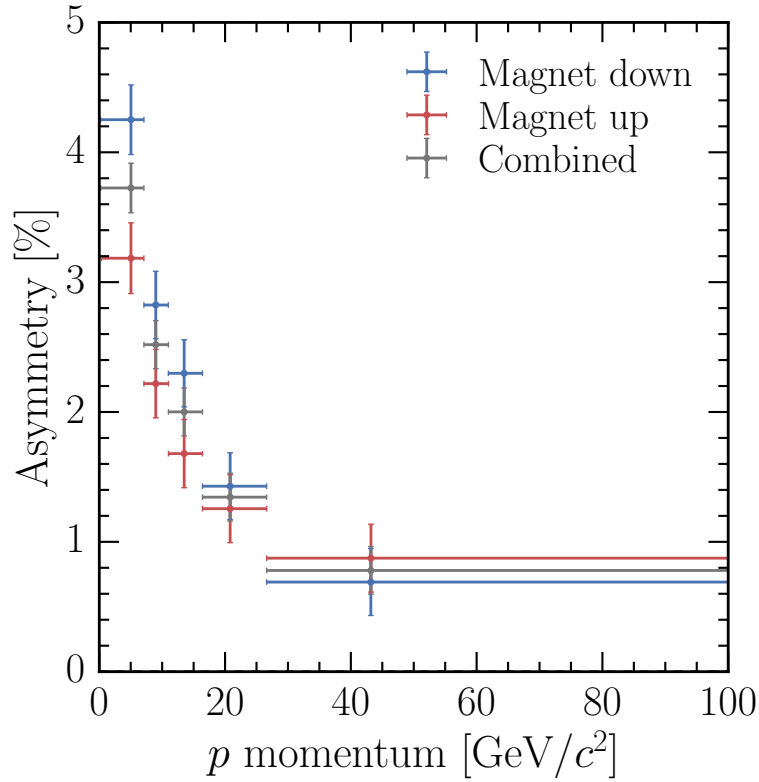
**Figure 23.5:** Asymmetry of the proton detection efficiency as a function of proton momentum in GeV/c. The uncertainties on the data points are statistical.

Table 23.3: Shifts in the measurement of ΔA_{CP} due to background asymmetries not completely removed during the kinematic weighting. The total shift ΔA is the sum of the individual ones, and all shifts are given in 10^{-4} .

Year	Polarity	$A_P^{\Lambda_b^0}$	$A_D^{\mu^-}$	A_D^p	A_D^K	ΔA
2011	Up	0.06 ± 1.51	0.35 ± 0.44	-7.08 ± 1.13	0.61 ± 0.29	-6.06 ± 1.96
2011	Down	-0.97 ± 1.62	0.31 ± 0.59	-8.53 ± 2.05	0.79 ± 0.29	-8.41 ± 2.69
2011	Both	-0.68 ± 1.07	0.33 ± 0.52	-7.80 ± 1.44	0.71 ± 0.27	-7.43 ± 1.89
2012	Up	0.56 ± 0.82	-0.05 ± 0.16	-6.84 ± 1.19	0.47 ± 0.21	-5.86 ± 1.47
2012	Down	9.41 ± 1.55	-0.15 ± 0.22	-11.55 ± 1.56	0.48 ± 0.20	-1.80 ± 2.22
2012	Both	5.06 ± 0.96	-0.10 ± 0.16	-9.26 ± 1.34	0.47 ± 0.20	-3.82 ± 1.67

23.3 Signal decays from other sources

The mass fits count the number of real Λ_c^+ decays in the samples. However, backgrounds to $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X$ can peak in the Λ_c^+ mass spectrum, such as baryonic decays of B mesons or promptly produced Λ_c^+ decays associated with random muons. These backgrounds will have different asymmetries, contaminating ΔA_{CP} .

Selecting pK^-K^+ and $p\pi^-\pi^+$ decays from promptly produced Λ_c^+ baryons will contaminate ΔA_{CP} with the Λ_c^+ production asymmetry. Figure 23.6 shows the signal $\ln \chi_{\text{IP}}^2$ distributions in the 2012 magnet down data sample, in which a peaking structure corresponding to prompt Λ_c^+ is not visible. These distributions can be compared to the $\ln \chi_{\text{IP}}^2$ distributions in samples dominated by prompt charm, given in Chapter 9. Even if there is a prompt Λ_c^+ component in the data used for the ΔA_{CP} measurement, the contribution from a Λ_c^+ production asymmetry should cancel as the Λ_c^+ kinematics are weighted to be equal. To confirm this hypothesis, the data sample is split in two, defined by the candidates passing or failing the cut $\ln \chi_{\text{IP}}^2 > 4$. The value of ΔA_{CP} is recomputed for each sample, and the difference between the two values is compatible with zero within 1σ . No corresponding systematic uncertainty is assigned.

Known alternative sources of Λ_c^+ from beauty decays are considered. These include $B^+ \rightarrow \Lambda_c^- p \pi^+$ decays and $B^+ \rightarrow \bar{\Sigma}_c p \pi \pi$ with $\Sigma_c \rightarrow \Lambda_c^+ \pi$. In total, the inclusive $B \rightarrow \Lambda_c^+ X$ branching fractions are, conservatively, less than 10% for both B^+ and B^0 , and the B_s^0 contribution is negligible [12]. The combined B cross-section is around four times that of the Λ_b^0 [11,97], and the relative rate of $B \rightarrow \Lambda_c^+ X$ is around 3 times that of $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X$, for a total relative rate of $N(B \rightarrow \Lambda_c^+ X)/N(\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X) = 12$. However, the muon triggers and $\Lambda_c^+ \mu^-$ vertex requirements strongly suppress the alternate sources of Λ_c^+ decays. In a simulated sample of one million generated $B^0 \rightarrow \Lambda_c^- p \pi^- \pi^+$ decays, of the order of 10 Λ_c^+ candidates pass the selection requirements defined in Chapter 17. Given this, and that at least the B^0 production asymmetry is compatible with zero below the percent level [112], the contribution from B decays is neglected.

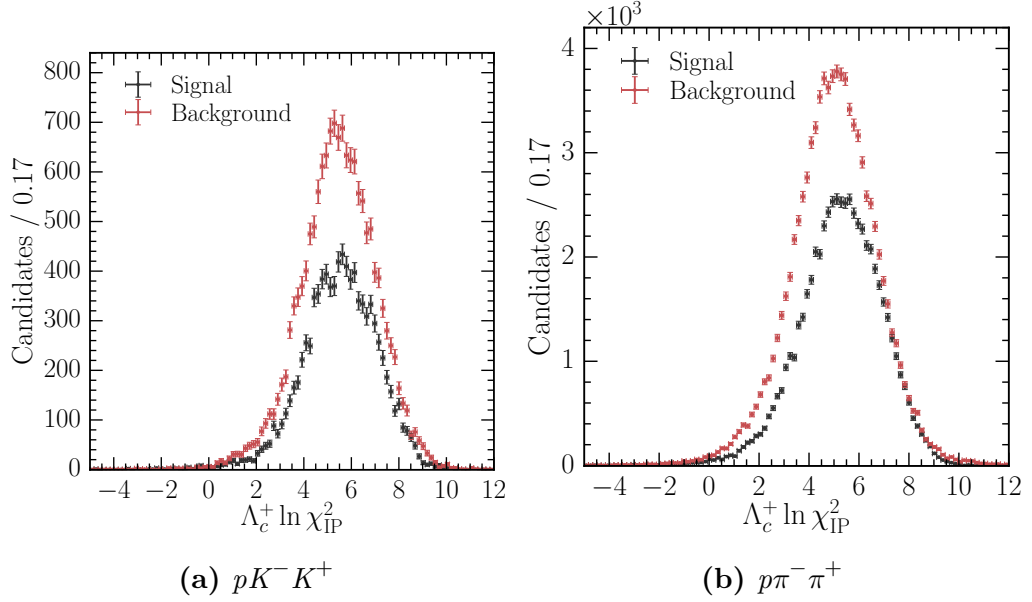


Figure 23.6: Signal (black) and background (red) $\ln \chi_{\text{IP}}^2$ distributions for $pK^- K^+$ (a) and $p\pi^- \pi^+$ (b) 2012 magnet down data samples. The $\ln \chi_{\text{IP}}^2$ requirement made for the systematic studies is $\ln \chi_{\text{IP}}^2 > 3$.

23.4 Summary and combination

Sources of systematic uncertainties on the measurement of ΔA_{CP} have been studied. The effect of the arbitrary choice of fit model is assessed with a sideband subtraction method, with the deviation from the nominal result found to be 0.26 %. Residual asymmetries arising from the failure of the kinematic weighting to completely equalise all kinematic distributions are found to shift the value of ΔA_{CP} by 0.1 %. The contribution of Λ_c^+ candidates from non- Λ_b^0 decays is considered to be negligible.

The two systematic uncertainties are added in quadrature in order to assign a single systematic to ΔA_{CP} , giving 0.28 %. This is small compared with the statistical uncertainty of 0.89 % given in Table 22.1.

24 Summary

The violation of CP symmetry has not yet been observed in charm decays, the last system in the SM where it is predicted to occur. Measuring CP asymmetries can be experimentally challenging as a good understanding of production and detection asymmetries must be obtained, but the construction of the CP asymmetry difference ΔA_{CP} allows for robust measurements that are less sensitive to spurious background effects.

With the analysis presented in this part, the first measurement of direct CP violation in charm baryons at LHCb has been completed. The decays $\Lambda_c^+ \rightarrow pK^-K^+$ and $\Lambda_c^+ \rightarrow p\pi^-\pi^+$ are reconstructed via the decay $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-X$, with the Λ_b^0 and muon providing powerful background discrimination. Backgrounds from charm mesons are removed with mass vetoes. Contributions from background asymmetries from the Λ_b^0 production asymmetry and the pion, kaon, proton, and muon detection asymmetries are suppressed by a weighting between the Λ_c^+ decay modes, employing a boosted decision tree to partition the data in Λ_c^+ and proton kinematics. The acceptance and selection efficiencies are computed as a function of the five-dimensional Λ_c^+ phase space also using a BDT, with an average efficiency of around 1%. The Λ_c^+ signal yields are counted with χ^2 fits, weighting the data with the kinematic weights and the inverse of the phase space efficiencies. Systematic effects due to the signal and background modelling in the fit, the residual background asymmetries after the weighting, and the contribution of Λ_c^+ candidates from sources other than semileptonic Λ_c^+ decays are studied, and are found to be considerably smaller than the statistical uncertainty on the signal yields.

The difference in CP asymmetries in the decays $\Lambda_c^+ \rightarrow pK^-K^+$ and

$\Lambda_c^+ \rightarrow p\pi^-\pi^+$ is measured to be

$$\begin{aligned}\Delta A_{CP} &= A_{CP}(pK^-K^+) - A_{CP}(p\pi^-\pi^+) \\ &= (1.20 \pm 0.92 \text{ (stat.)} \pm 0.28 \text{ (syst.)}) \text{ \%}.\end{aligned}$$

The central value is currently blinded, pending approval from the LHCb collaboration.

Appendices

A Charm production yield, efficiency, and systematic tables

Prompt signal yields measured in charm hadron (p_T , y) bins are given in Tables A.1 to A.4 for D^0 , D^+ , D_s^+ , and D^{*+} mesons, measured using the fit procedure described in Chapter 9.

The following efficiencies are given for D^0 and D^+ : acceptance efficiencies in Tables A.5 and A.6; reconstruction efficiencies in Tables A.7 and A.8; tracking efficiency correction factors in Tables A.9 and A.10; selection efficiencies in Tables A.11 and A.12; particle identification efficiencies in Tables A.13 and A.14; and signal window efficiencies in Tables A.15 and A.16. The computation of the efficiencies is described in Chapter 10.

The following systematic uncertainties for D^0 and D^+ are presented: due to the size of the finite MC samples in Tables A.17 and A.18; due to the finite size of the PID calibration samples in Tables A.19 and A.20; due to the choice of binning scheme in the PID calibration procedure in Tables A.21 and A.22; due to the tracking efficiency correction methods in Tables A.23 and A.24; due to the choice of fit model in Tables A.25 and A.26; and due to the statistical uncertainty on the model parameters used to evaluate the signal window efficiency in Tables A.27 and A.28. The computation of the systematic uncertainties is described in Chapter 11.

Table A.1: Prompt signal yields for $D^0 \rightarrow K^- \pi^+$ measured in $D^0(p_T, y)$ bins. Cells marked with a dash ‘-’ indicate bins where insufficient data were available for a statistically significant prompt signal yield measurement.

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[1400, 1500]	550 $\pm$ 30	610 $\pm$ 30	72 $\pm$ 9	-	-
[1300, 1400]	860 $\pm$ 30	970 $\pm$ 40	180 $\pm$ 20	-	-
[1200, 1300]	1250 $\pm$ 40	1580 $\pm$ 40	420 $\pm$ 30	-	-
[1100, 1200]	1910 $\pm$ 50	2520 $\pm$ 50	890 $\pm$ 30	-	-
[1000, 1100]	2990 $\pm$ 60	3890 $\pm$ 70	1880 $\pm$ 50	32 $\pm$ 6	-
[900, 1000]	4830 $\pm$ 70	6420 $\pm$ 80	3810 $\pm$ 70	170 $\pm$ 20	-
[800, 900]	7570 $\pm$ 90	10900 $\pm$ 200	7250 $\pm$ 90	680 $\pm$ 30	-
[700, 800]	12000 $\pm$ 200	18600 $\pm$ 200	13700 $\pm$ 200	2460 $\pm$ 50	-
[600, 700]	19100 $\pm$ 200	32400 $\pm$ 200	25500 $\pm$ 200	7510 $\pm$ 90	39 $\pm$ 6
[500, 600]	30100 $\pm$ 200	56300 $\pm$ 300	47000 $\pm$ 300	19000 $\pm$ 200	670 $\pm$ 30
[400, 500]	44800 $\pm$ 300	98500 $\pm$ 400	85500 $\pm$ 300	41900 $\pm$ 300	3990 $\pm$ 70
[350, 400]	28000 $\pm$ 200	70900 $\pm$ 300	63600 $\pm$ 300	34100 $\pm$ 200	5290 $\pm$ 80
[300, 350]	30600 $\pm$ 200	85600 $\pm$ 300	80500 $\pm$ 300	44400 $\pm$ 300	8560 $\pm$ 100
[250, 300]	31700 $\pm$ 200	98100 $\pm$ 400	97000 $\pm$ 400	54400 $\pm$ 300	11800 $\pm$ 200
[200, 250]	30100 $\pm$ 200	102200 $\pm$ 400	106700 $\pm$ 400	62200 $\pm$ 300	14400 $\pm$ 200
[150, 200]	27500 $\pm$ 200	96600 $\pm$ 400	107700 $\pm$ 400	64400 $\pm$ 300	15300 $\pm$ 200
[100, 150]	22100 $\pm$ 200	84900 $\pm$ 300	99900 $\pm$ 400	60700 $\pm$ 300	14900 $\pm$ 200
[0, 1000]	20300 $\pm$ 200	85900 $\pm$ 300	108800 $\pm$ 400	69400 $\pm$ 300	17200 $\pm$ 200

Table A.2: Prompt signal yields for $D^+ \rightarrow K^- \pi^+ \pi^+$ measured in D^+ (p_T, y) bins. Cells marked with a dash '-' indicate bins where insufficient data were available for a statistically significant prompt signal yield measurement.

p_T [MeV/c]	y			
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]
[14000, 15000]	780 ± 30	1070 ± 40	420 ± 20	3.2 ± 0.4
[13000, 14000]	1110 ± 40	1590 ± 40	760 ± 30	13 ± 2
[12000, 13000]	1600 ± 40	2380 ± 50	1280 ± 40	61 ± 4
[11000, 12000]	2380 ± 50	3570 ± 70	2330 ± 50	190 ± 20
[10000, 11000]	3360 ± 60	5450 ± 80	3910 ± 70	490 ± 30
[9000, 10000]	5010 ± 80	8730 ± 100	6820 ± 90	1390 ± 40
[8000, 9000]	7530 ± 90	13700 ± 200	11400 ± 200	3300 ± 60
[7000, 8000]	10800 ± 200	22200 ± 200	19400 ± 200	7550 ± 100
[6000, 7000]	16200 ± 200	36000 ± 200	33700 ± 200	16400 ± 200
[5000, 6000]	22700 ± 200	61100 ± 300	59600 ± 300	33200 ± 200
[4000, 5000]	30600 ± 200	99200 ± 400	103800 ± 400	63400 ± 300
[3500, 4000]	17000 ± 200	68300 ± 300	76300 ± 300	48100 ± 300
[3000, 3500]	16800 ± 200	78800 ± 300	93300 ± 400	60600 ± 300
[2500, 3000]	14400 ± 200	83900 ± 400	107000 ± 400	70400 ± 300
[2000, 2500]	9900 ± 200	70600 ± 300	99100 ± 400	67100 ± 300
[1500, 2000]	3870 ± 70	39700 ± 300	63000 ± 300	44300 ± 300
[1000, 1500]	630 ± 30	11300 ± 200	21800 ± 200	16200 ± 200
[0, 1000]	21 ± 2	900 ± 40	2580 ± 70	2030 ± 70
				460 ± 40

Table A.3: Prompt signal yields for $D_s^+ \rightarrow \phi\pi^+$ measured in D_s^+ (p_T , y) bins. Cells marked with a dash ‘-’ indicate bins where insufficient data were available for a statistically significant prompt signal yield measurement.

p_T [MeV/c]	[2, 2.5]	[2.5, 3]	y [3, 3.5]	[3.5, 4]	[4, 4.5]
[1400, 1500]	51 $\pm$ 8	85 $\pm$ 10	28 $\pm$ 6	-	-
[1300, 1400]	79 $\pm$ 9	140 $\pm$ 20	56 $\pm$ 8	2.0 $\pm$ 0.5	-
[1200, 1300]	150 $\pm$ 20	150 $\pm$ 20	110 $\pm$ 20	4 $\pm$ 2	-
[1100, 1200]	190 $\pm$ 20	260 $\pm$ 20	180 $\pm$ 20	10 $\pm$ 3	-
[1000, 1100]	300 $\pm$ 20	400 $\pm$ 30	260 $\pm$ 20	34 $\pm$ 6	-
[900, 1000]	450 $\pm$ 30	610 $\pm$ 30	500 $\pm$ 30	110 $\pm$ 20	-
[800, 900]	660 $\pm$ 30	1000 $\pm$ 40	820 $\pm$ 30	250 $\pm$ 20	-
[700, 800]	990 $\pm$ 40	1590 $\pm$ 50	1310 $\pm$ 40	510 $\pm$ 30	13 $\pm$ 4
[600, 700]	1500 $\pm$ 50	2730 $\pm$ 60	2230 $\pm$ 50	1090 $\pm$ 40	75 $\pm$ 9
[500, 600]	2300 $\pm$ 50	4480 $\pm$ 70	3870 $\pm$ 70	2160 $\pm$ 50	250 $\pm$ 20
[400, 500]	2860 $\pm$ 60	6990 $\pm$ 90	6250 $\pm$ 90	3530 $\pm$ 70	580 $\pm$ 30
[350, 400]	1630 $\pm$ 50	4500 $\pm$ 70	4340 $\pm$ 70	2480 $\pm$ 60	390 $\pm$ 30
[300, 350]	1560 $\pm$ 50	5190 $\pm$ 80	5170 $\pm$ 80	2850 $\pm$ 60	480 $\pm$ 30
[250, 300]	1340 $\pm$ 40	5210 $\pm$ 80	5440 $\pm$ 80	2880 $\pm$ 60	490 $\pm$ 30
[200, 250]	850 $\pm$ 40	3800 $\pm$ 70	4120 $\pm$ 70	2200 $\pm$ 60	310 $\pm$ 30
[150, 200]	290 $\pm$ 20	1720 $\pm$ 50	1940 $\pm$ 50	960 $\pm$ 40	130 $\pm$ 20
[100, 150]	15 $\pm$ 4	240 $\pm$ 20	330 $\pm$ 20	140 $\pm$ 20	12 $\pm$ 4

Table A.4: Prompt signal yields for $D^{*+} \rightarrow D^0 \pi^+$, with $D^0 \rightarrow K^- \pi^+$, measured in $D^{*+}(p_T, y)$ bins. Cells marked with a dash '-' indicate bins where insufficient data were available for a statistically significant prompt signal yield measurement.

$p_T[\text{MeV}/c]$	[2, 2.5]	[2.5, 3]	y [3, 3.5]	[3.5, 4]	[4, 4.5]
[14000, 15000]	121 $\pm$ 9	180 $\pm$ 20	26 $\pm$ 4	-	-
[13000, 14000]	210 $\pm$ 20	250 $\pm$ 20	84 $\pm$ 7	-	-
[12000, 13000]	270 $\pm$ 20	460 $\pm$ 20	158 $\pm$ 10	-	-
[11000, 12000]	400 $\pm$ 20	670 $\pm$ 20	290 $\pm$ 20	4.0 $\pm$ 0.5	-
[10000, 11000]	610 $\pm$ 20	1020 $\pm$ 30	610 $\pm$ 20	19 $\pm$ 4	-
[9000, 10000]	780 $\pm$ 30	1770 $\pm$ 40	1130 $\pm$ 30	81 $\pm$ 7	-
[8000, 9000]	1300 $\pm$ 30	2750 $\pm$ 50	1930 $\pm$ 40	270 $\pm$ 20	-
[7000, 8000]	1790 $\pm$ 40	4420 $\pm$ 60	3460 $\pm$ 50	830 $\pm$ 30	-
[6000, 7000]	2330 $\pm$ 40	7320 $\pm$ 70	5950 $\pm$ 70	1850 $\pm$ 40	24 $\pm$ 4
[5000, 6000]	3110 $\pm$ 50	11330 $\pm$ 90	10820 $\pm$ 90	4470 $\pm$ 60	220 $\pm$ 20
[4000, 5000]	3510 $\pm$ 50	16700 $\pm$ 200	18700 $\pm$ 200	8980 $\pm$ 80	890 $\pm$ 30
[3500, 4000]	1500 $\pm$ 40	10210 $\pm$ 90	13520 $\pm$ 100	6880 $\pm$ 70	1000 $\pm$ 30
[3000, 3500]	1050 $\pm$ 30	9990 $\pm$ 90	16000 $\pm$ 200	8310 $\pm$ 80	1560 $\pm$ 40
[2500, 3000]	570 $\pm$ 20	8810 $\pm$ 80	17600 $\pm$ 200	9600 $\pm$ 90	1960 $\pm$ 40
[2000, 2500]	104 $\pm$ 10	6010 $\pm$ 70	17400 $\pm$ 200	10210 $\pm$ 90	2090 $\pm$ 40
[1500, 2000]	-	2850 $\pm$ 50	12170 $\pm$ 100	7750 $\pm$ 80	1680 $\pm$ 40
[1000, 1500]	-	620 $\pm$ 30	5420 $\pm$ 70	3600 $\pm$ 60	910 $\pm$ 30
[0, 1000]	-	-	160 $\pm$ 30	130 $\pm$ 20	67 $\pm$ 10

Table A.5: Acceptance efficiencies $\epsilon_{\text{Acceptance}}$ for $D^0 \rightarrow K^- \pi^+$ measured in $D^0(p_T, y)$ bins.

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	$76.41^{+0.29}_{-0.29}$	$91.57^{+0.20}_{-0.20}$	$95.30^{+0.15}_{-0.16}$	$91.45^{+0.22}_{-0.22}$	$77.80^{+0.35}_{-0.36}$
[1000, 1500]	$78.50^{+0.34}_{-0.35}$	$93.35^{+0.22}_{-0.22}$	$96.34^{+0.17}_{-0.18}$	$93.02^{+0.25}_{-0.26}$	$80.22^{+0.44}_{-0.45}$
[1500, 2000]	$81.59^{+0.36}_{-0.37}$	$94.56^{+0.22}_{-0.23}$	$97.60^{+0.15}_{-0.17}$	$94.42^{+0.26}_{-0.27}$	$83.44^{+0.47}_{-0.48}$
[2000, 2500]	$84.59^{+0.38}_{-0.39}$	$96.22^{+0.21}_{-0.22}$	$98.30^{+0.15}_{-0.16}$	$95.79^{+0.26}_{-0.27}$	$87.13^{+0.50}_{-0.51}$
[2500, 3000]	$87.75^{+0.40}_{-0.42}$	$97.14^{+0.21}_{-0.23}$	$98.70^{+0.15}_{-0.17}$	$95.95^{+0.29}_{-0.32}$	$89.27^{+0.54}_{-0.57}$
[3000, 3500]	$89.45^{+0.44}_{-0.45}$	$98.31^{+0.19}_{-0.21}$	$98.93^{+0.16}_{-0.18}$	$97.07^{+0.29}_{-0.33}$	$91.56^{+0.56}_{-0.60}$
[3500, 4000]	$91.66^{+0.45}_{-0.48}$	$98.56^{+0.20}_{-0.23}$	$99.70^{+0.09}_{-0.13}$	$98.51^{+0.24}_{-0.29}$	$94.09^{+0.56}_{-0.61}$
[4000, 5000]	$94.34^{+0.33}_{-0.35}$	$99.32^{+0.12}_{-0.14}$	$99.73^{+0.08}_{-0.11}$	$99.19^{+0.16}_{-0.19}$	$96.22^{+0.40}_{-0.44}$
[5000, 6000]	$95.83^{+0.37}_{-0.40}$	$99.87^{+0.06}_{-0.10}$	$99.64^{+0.11}_{-0.17}$	$99.56^{+0.14}_{-0.20}$	$98.10^{+0.38}_{-0.47}$
[6000, 7000]	$96.75^{+0.41}_{-0.46}$	$100.00^{+0.00}_{-0.09}$	$99.90^{+0.07}_{-0.16}$	$99.66^{+0.15}_{-0.27}$	$99.26^{+0.30}_{-0.48}$
[7000, 8000]	$98.62^{+0.33}_{-0.44}$	$99.87^{+0.09}_{-0.21}$	$100.00^{+0.00}_{-0.17}$	$99.80^{+0.14}_{-0.34}$	$99.69^{+0.23}_{-0.53}$
[8000, 9000]	$98.48^{+0.43}_{-0.60}$	$100.00^{+0.00}_{-0.26}$	$100.00^{+0.00}_{-0.30}$	$100.00^{+0.00}_{-0.41}$	$98.51^{+0.68}_{-1.16}$
[9000, 10000]	$99.28^{+0.33}_{-0.57}$	$100.00^{+0.00}_{-0.36}$	$100.00^{+0.00}_{-0.45}$	$100.00^{+0.00}_{-0.62}$	$99.19^{+0.58}_{-1.38}$
[10000, 11000]	$99.65^{+0.25}_{-0.60}$	$100.00^{+0.00}_{-0.63}$	$100.00^{+0.00}_{-0.70}$	$100.00^{+0.00}_{-0.99}$	$100.00^{+0.00}_{-1.50}$
[11000, 12000]	$97.65^{+0.94}_{-1.50}$	$100.00^{+0.00}_{-0.81}$	$100.00^{+0.00}_{-1.13}$	$100.00^{+0.00}_{-1.40}$	$100.00^{+0.00}_{-2.50}$
[12000, 13000]	$100.00^{+0.00}_{-0.88}$	$100.00^{+0.00}_{-1.37}$	$100.00^{+0.00}_{-1.23}$	$100.00^{+0.00}_{-2.04}$	$100.00^{+0.00}_{-4.13}$
[13000, 14000]	$100.00^{+0.00}_{-1.32}$	$100.00^{+0.00}_{-1.91}$	$100.00^{+0.00}_{-2.67}$	$100.00^{+0.00}_{-3.33}$	$100.00^{+0.00}_{-3.86}$
[14000, 15000]	$100.00^{+0.00}_{-1.67}$	$100.00^{+0.00}_{-2.07}$	$100.00^{+0.00}_{-2.35}$	$100.00^{+0.00}_{-5.69}$	$93.75^{+4.30}_{-9.43}$

Table A.6: Acceptance efficiencies $\varepsilon_{\text{Acceptance}}$ for $D^+ \rightarrow K^- \pi^+ \pi^+$ measured in D^+ (p_T, y) bins.

p_T [MeV/c]	y			
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]
[0, 1000]	$67.05^{+0.34}_{-0.34}$	$88.70^{+0.24}_{-0.24}$	$93.18^{+0.19}_{-0.20}$	$83.11^{+0.31}_{-0.32}$
[1000, 1500]	$70.61^{+0.40}_{-0.41}$	$91.98^{+0.25}_{-0.26}$	$95.14^{+0.21}_{-0.22}$	$88.42^{+0.33}_{-0.34}$
[1500, 2000]	$75.54^{+0.42}_{-0.43}$	$94.67^{+0.23}_{-0.24}$	$97.06^{+0.18}_{-0.19}$	$91.77^{+0.32}_{-0.34}$
[2000, 2500]	$80.04^{+0.45}_{-0.46}$	$96.60^{+0.21}_{-0.23}$	$98.23^{+0.16}_{-0.18}$	$94.44^{+0.32}_{-0.33}$
[2500, 3000]	$83.81^{+0.49}_{-0.50}$	$98.08^{+0.18}_{-0.20}$	$98.42^{+0.18}_{-0.20}$	$96.32^{+0.30}_{-0.32}$
[3000, 3500]	$87.14^{+0.50}_{-0.52}$	$98.86^{+0.16}_{-0.18}$	$99.01^{+0.16}_{-0.19}$	$96.96^{+0.32}_{-0.35}$
[3500, 4000]	$89.87^{+0.53}_{-0.55}$	$99.01^{+0.17}_{-0.21}$	$99.31^{+0.15}_{-0.19}$	$97.82^{+0.31}_{-0.35}$
[4000, 5000]	$92.32^{+0.41}_{-0.43}$	$99.59^{+0.09}_{-0.12}$	$99.49^{+0.11}_{-0.15}$	$98.35^{+0.24}_{-0.28}$
[5000, 6000]	$95.96^{+0.38}_{-0.42}$	$99.76^{+0.09}_{-0.13}$	$99.56^{+0.13}_{-0.19}$	$98.64^{+0.28}_{-0.36}$
[6000, 7000]	$96.79^{+0.44}_{-0.50}$	$100.00^{+0.00}_{-0.10}$	$99.50^{+0.18}_{-0.28}$	$99.15^{+0.27}_{-0.39}$
[7000, 8000]	$98.68^{+0.33}_{-0.44}$	$99.87^{+0.10}_{-0.23}$	$99.51^{+0.22}_{-0.38}$	$99.58^{+0.22}_{-0.43}$
[8000, 9000]	$99.46^{+0.24}_{-0.42}$	$99.80^{+0.15}_{-0.35}$	$100.00^{+0.00}_{-0.32}$	$98.31^{+0.62}_{-0.94}$
[9000, 10000]	$99.46^{+0.29}_{-0.55}$	$100.00^{+0.00}_{-0.36}$	$100.00^{+0.00}_{-0.46}$	$99.42^{+0.42}_{-0.98}$
[10000, 11000]	$99.60^{+0.29}_{-0.68}$	$100.00^{+0.00}_{-0.64}$	$100.00^{+0.00}_{-0.67}$	$99.12^{+0.64}_{-1.49}$
[11000, 12000]	$100.00^{+0.00}_{-0.75}$	$100.00^{+0.00}_{-0.88}$	$100.00^{+0.00}_{-1.27}$	$100.00^{+0.00}_{-1.37}$
[12000, 13000]	$100.00^{+0.00}_{-1.05}$	$100.00^{+0.00}_{-1.05}$	$100.00^{+0.00}_{-1.62}$	$100.00^{+0.00}_{-2.61}$
[13000, 14000]	$100.00^{+0.00}_{-1.50}$	$100.00^{+0.00}_{-1.91}$	$100.00^{+0.00}_{-2.04}$	$100.00^{+0.00}_{-3.86}$
[14000, 15000]	$100.00^{+0.00}_{-2.40}$	$100.00^{+0.00}_{-2.40}$	$100.00^{+0.00}_{-3.00}$	$100.00^{+0.00}_{-4.99}$
				$59.35^{+0.44}_{-0.44}$
				$69.86^{+0.53}_{-0.53}$
				$76.51^{+0.56}_{-0.57}$
				$82.68^{+0.59}_{-0.60}$
				$88.16^{+0.59}_{-0.62}$
				$91.10^{+0.60}_{-0.64}$
				$93.66^{+0.61}_{-0.67}$
				$95.20^{+0.48}_{-0.53}$
				$97.23^{+0.47}_{-0.57}$
				$97.21^{+0.61}_{-0.78}$
				$99.67^{+0.24}_{-0.56}$
				$98.91^{+0.58}_{-1.11}$
				$98.25^{+0.94}_{-1.78}$
				$100.00^{+0.00}_{-1.72}$
				$100.00^{+0.00}_{-2.61}$
				$100.00^{+0.00}_{-4.79}$
				$100.00^{+0.00}_{-3.53}$
				$100.00^{+0.00}_{-7.95}$

Table A.7: *Reconstruction efficiencies $\varepsilon_{\text{Reconstruction}}$ for $D^0 \rightarrow K^- \pi^+$ measured in D^0 (p_T , y) bins.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	$18.02^{+0.13}_{-0.13}$	$38.33^{+0.15}_{-0.15}$	$47.13^{+0.16}_{-0.16}$	$45.23^{+0.18}_{-0.18}$	$32.30^{+0.23}_{-0.23}$
[1000, 1500]	$19.83^{+0.15}_{-0.15}$	$39.65^{+0.17}_{-0.17}$	$48.14^{+0.18}_{-0.18}$	$45.81^{+0.19}_{-0.19}$	$32.59^{+0.21}_{-0.21}$
[1500, 2000]	$21.07^{+0.16}_{-0.15}$	$42.66^{+0.18}_{-0.18}$	$50.40^{+0.19}_{-0.19}$	$47.87^{+0.20}_{-0.20}$	$34.09^{+0.22}_{-0.22}$
[2000, 2500]	$23.73^{+0.17}_{-0.17}$	$45.20^{+0.20}_{-0.20}$	$52.59^{+0.21}_{-0.21}$	$50.76^{+0.22}_{-0.22}$	$37.56^{+0.25}_{-0.25}$
[2500, 3000]	$25.73^{+0.20}_{-0.20}$	$47.72^{+0.22}_{-0.22}$	$55.08^{+0.23}_{-0.23}$	$53.32^{+0.25}_{-0.25}$	$40.25^{+0.28}_{-0.28}$
[3000, 3500]	$28.43^{+0.23}_{-0.23}$	$50.24^{+0.25}_{-0.25}$	$56.76^{+0.26}_{-0.26}$	$55.78^{+0.29}_{-0.29}$	$43.01^{+0.33}_{-0.33}$
[3500, 4000]	$30.34^{+0.27}_{-0.27}$	$51.59^{+0.29}_{-0.29}$	$58.42^{+0.31}_{-0.31}$	$56.82^{+0.33}_{-0.33}$	$44.93^{+0.39}_{-0.39}$
[4000, 5000]	$32.73^{+0.24}_{-0.24}$	$53.86^{+0.26}_{-0.26}$	$60.06^{+0.27}_{-0.27}$	$58.29^{+0.30}_{-0.30}$	$46.44^{+0.35}_{-0.35}$
[5000, 6000]	$35.49^{+0.32}_{-0.32}$	$55.24^{+0.34}_{-0.34}$	$61.62^{+0.36}_{-0.37}$	$60.30^{+0.40}_{-0.41}$	$48.14^{+0.48}_{-0.48}$
[6000, 7000]	$36.87^{+0.43}_{-0.42}$	$56.76^{+0.45}_{-0.45}$	$62.26^{+0.48}_{-0.49}$	$62.14^{+0.54}_{-0.55}$	$49.30^{+0.66}_{-0.66}$
[7000, 8000]	$40.13^{+0.56}_{-0.56}$	$59.07^{+0.59}_{-0.60}$	$63.58^{+0.63}_{-0.64}$	$63.37^{+0.72}_{-0.73}$	$49.26^{+0.91}_{-0.91}$
[8000, 9000]	$44.84^{+0.72}_{-0.72}$	$63.71^{+0.74}_{-0.74}$	$67.73^{+0.79}_{-0.80}$	$66.13^{+0.92}_{-0.93}$	$53.24^{+1.13}_{-1.13}$
[9000, 10000]	$42.14^{+0.21}_{-0.21}$	$58.50^{+0.23}_{-0.23}$	$62.46^{+0.25}_{-0.25}$	$61.69^{+0.29}_{-0.30}$	$50.16^{+0.38}_{-0.38}$
[10000, 11000]	$43.57^{+0.22}_{-0.22}$	$59.30^{+0.23}_{-0.23}$	$62.32^{+0.26}_{-0.26}$	$61.79^{+0.31}_{-0.31}$	$49.33^{+0.39}_{-0.39}$
[11000, 12000]	$44.30^{+0.26}_{-0.26}$	$59.05^{+0.29}_{-0.29}$	$60.99^{+0.32}_{-0.32}$	$62.21^{+0.38}_{-0.38}$	$50.12^{+0.50}_{-0.50}$
[12000, 13000]	$44.79^{+0.32}_{-0.32}$	$58.79^{+0.35}_{-0.35}$	$62.00^{+0.39}_{-0.39}$	$61.65^{+0.47}_{-0.47}$	$49.14^{+0.63}_{-0.63}$
[13000, 14000]	$45.09^{+0.38}_{-0.38}$	$58.05^{+0.42}_{-0.42}$	$61.85^{+0.47}_{-0.48}$	$61.14^{+0.58}_{-0.58}$	$50.36^{+0.77}_{-0.77}$
[14000, 15000]	$44.90^{+0.45}_{-0.45}$	$58.84^{+0.49}_{-0.50}$	$61.51^{+0.57}_{-0.57}$	$61.41^{+0.69}_{-0.70}$	$46.16^{+0.94}_{-0.94}$

Table A.8: Reconstruction efficiencies $\varepsilon_{\text{Reconstruction}}$ for $D^+ \rightarrow K^- \pi^+ \pi^+$ measured in D^+ (p_T , y) bins.

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	$4.03^{+0.07}_{-0.07}$	$16.32^{+0.12}_{-0.12}$	$25.00^{+0.14}_{-0.14}$	$23.22^{+0.15}_{-0.15}$	$12.17^{+0.14}_{-0.14}$
[1000, 1500]	$5.24^{+0.09}_{-0.09}$	$19.01^{+0.14}_{-0.14}$	$28.08^{+0.16}_{-0.16}$	$26.46^{+0.17}_{-0.17}$	$14.25^{+0.17}_{-0.16}$
[1500, 2000]	$6.45^{+0.10}_{-0.10}$	$21.87^{+0.15}_{-0.15}$	$31.21^{+0.18}_{-0.18}$	$29.58^{+0.19}_{-0.19}$	$16.69^{+0.18}_{-0.18}$
[2000, 2500]	$8.05^{+0.12}_{-0.12}$	$24.99^{+0.18}_{-0.17}$	$34.40^{+0.20}_{-0.20}$	$32.92^{+0.21}_{-0.21}$	$19.53^{+0.21}_{-0.21}$
[2500, 3000]	$9.67^{+0.14}_{-0.14}$	$27.95^{+0.20}_{-0.20}$	$37.96^{+0.23}_{-0.23}$	$36.31^{+0.25}_{-0.25}$	$22.32^{+0.25}_{-0.24}$
[3000, 3500]	$11.56^{+0.17}_{-0.17}$	$30.52^{+0.24}_{-0.24}$	$40.04^{+0.27}_{-0.27}$	$39.10^{+0.29}_{-0.29}$	$24.56^{+0.29}_{-0.29}$
[3500, 4000]	$13.20^{+0.20}_{-0.20}$	$32.40^{+0.28}_{-0.28}$	$42.36^{+0.31}_{-0.31}$	$41.58^{+0.34}_{-0.34}$	$26.97^{+0.35}_{-0.35}$
[4000, 5000]	$15.34^{+0.18}_{-0.18}$	$35.45^{+0.25}_{-0.25}$	$44.07^{+0.27}_{-0.27}$	$43.21^{+0.30}_{-0.30}$	$29.37^{+0.32}_{-0.32}$
[5000, 6000]	$17.90^{+0.26}_{-0.26}$	$38.05^{+0.33}_{-0.33}$	$45.64^{+0.37}_{-0.37}$	$45.53^{+0.41}_{-0.41}$	$31.44^{+0.45}_{-0.45}$
[6000, 7000]	$20.46^{+0.36}_{-0.35}$	$40.01^{+0.45}_{-0.45}$	$46.55^{+0.50}_{-0.50}$	$45.80^{+0.56}_{-0.56}$	$31.61^{+0.62}_{-0.61}$
[7000, 8000]	$23.55^{+0.48}_{-0.48}$	$41.90^{+0.59}_{-0.59}$	$48.49^{+0.65}_{-0.65}$	$48.03^{+0.73}_{-0.73}$	$34.60^{+0.86}_{-0.85}$
[8000, 9000]	$24.06^{+0.62}_{-0.61}$	$44.13^{+0.76}_{-0.75}$	$49.20^{+0.85}_{-0.85}$	$49.41^{+0.98}_{-0.98}$	$33.08^{+1.14}_{-1.12}$
[9000, 10000]	$31.30^{+0.74}_{-0.73}$	$50.40^{+0.79}_{-0.79}$	$56.39^{+0.80}_{-0.81}$	$52.83^{+0.83}_{-0.83}$	$37.42^{+0.90}_{-0.89}$
[10000, 11000]	$26.73^{+0.18}_{-0.17}$	$44.57^{+0.21}_{-0.21}$	$49.27^{+0.24}_{-0.24}$	$47.65^{+0.29}_{-0.29}$	$33.79^{+0.35}_{-0.35}$
[11000, 12000]	$28.20^{+0.22}_{-0.22}$	$45.31^{+0.26}_{-0.26}$	$48.61^{+0.29}_{-0.29}$	$48.78^{+0.35}_{-0.35}$	$33.95^{+0.42}_{-0.42}$
[12000, 13000]	$28.86^{+0.26}_{-0.26}$	$45.33^{+0.31}_{-0.31}$	$49.59^{+0.36}_{-0.36}$	$49.01^{+0.43}_{-0.43}$	$33.18^{+0.53}_{-0.52}$
[13000, 14000]	$29.96^{+0.32}_{-0.31}$	$45.02^{+0.38}_{-0.38}$	$49.19^{+0.43}_{-0.43}$	$47.71^{+0.52}_{-0.52}$	$33.37^{+0.65}_{-0.64}$
[14000, 15000]	$29.32^{+0.37}_{-0.37}$	$44.56^{+0.44}_{-0.44}$	$48.73^{+0.52}_{-0.52}$	$47.59^{+0.63}_{-0.63}$	$33.91^{+0.79}_{-0.78}$

Table A.9: Tracking efficiency correction factors ρ_{Tracking} for $D^0 \rightarrow K^- \pi^+$ measured in D^0 in (p_T, y) bins.

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	110.26 $\pm$ 0.03	107.14 $\pm$ 0.03	104.04 $\pm$ 0.03	100.75 $\pm$ 0.03	99.40 $\pm$ 0.03
[1000, 1500]	109.14 $\pm$ 0.05	105.85 $\pm$ 0.04	102.95 $\pm$ 0.03	100.30 $\pm$ 0.03	99.03 $\pm$ 0.03
[1500, 2000]	108.04 $\pm$ 0.06	105.26 $\pm$ 0.04	102.58 $\pm$ 0.03	100.17 $\pm$ 0.03	99.00 $\pm$ 0.04
[2000, 2500]	107.21 $\pm$ 0.06	104.85 $\pm$ 0.05	102.24 $\pm$ 0.03	99.95 $\pm$ 0.03	99.16 $\pm$ 0.03
[2500, 3000]	106.94 $\pm$ 0.06	104.58 $\pm$ 0.03	102.02 $\pm$ 0.03	99.70 $\pm$ 0.03	99.52 $\pm$ 0.03
[3000, 3500]	106.48 $\pm$ 0.07	104.32 $\pm$ 0.04	101.76 $\pm$ 0.04	99.48 $\pm$ 0.03	99.93 $\pm$ 0.03
[3500, 4000]	106.47 $\pm$ 0.05	104.05 $\pm$ 0.04	101.44 $\pm$ 0.03	99.38 $\pm$ 0.03	100.30 $\pm$ 0.04
[4000, 5000]	106.12 $\pm$ 0.04	103.79 $\pm$ 0.03	101.14 $\pm$ 0.03	99.40 $\pm$ 0.03	100.78 $\pm$ 0.03
[5000, 6000]	105.82 $\pm$ 0.04	103.59 $\pm$ 0.04	100.87 $\pm$ 0.04	99.87 $\pm$ 0.04	101.40 $\pm$ 0.04
[6000, 7000]	105.35 $\pm$ 0.04	103.45 $\pm$ 0.05	100.69 $\pm$ 0.06	100.53 $\pm$ 0.05	102.00 $\pm$ 0.04
[7000, 8000]	104.99 $\pm$ 0.04	103.37 $\pm$ 0.05	100.84 $\pm$ 0.05	100.91 $\pm$ 0.06	102.26 $\pm$ 0.05
[8000, 9000]	104.53 $\pm$ 0.06	103.36 $\pm$ 0.06	101.12 $\pm$ 0.07	101.38 $\pm$ 0.06	102.61 $\pm$ 0.07
[9000, 10000]	104.37 $\pm$ 0.02	103.25 $\pm$ 0.02	101.47 $\pm$ 0.03	101.73 $\pm$ 0.02	102.83 $\pm$ 0.03
[10000, 11000]	104.15 $\pm$ 0.02	103.14 $\pm$ 0.02	101.73 $\pm$ 0.02	101.97 $\pm$ 0.02	103.00 $\pm$ 0.03
[11000, 12000]	103.98 $\pm$ 0.02	103.12 $\pm$ 0.02	102.01 $\pm$ 0.03	102.18 $\pm$ 0.03	103.13 $\pm$ 0.04
[12000, 13000]	103.81 $\pm$ 0.02	103.09 $\pm$ 0.02	102.18 $\pm$ 0.03	102.36 $\pm$ 0.03	103.30 $\pm$ 0.04
[13000, 14000]	103.65 $\pm$ 0.03	103.17 $\pm$ 0.03	102.37 $\pm$ 0.03	102.48 $\pm$ 0.04	103.45 $\pm$ 0.05
[14000, 15000]	103.55 $\pm$ 0.03	103.17 $\pm$ 0.03	102.54 $\pm$ 0.04	102.74 $\pm$ 0.05	103.52 $\pm$ 0.08

Table A.10: Tracking efficiency correction factors ρ_{Tracking} for $D^+ \rightarrow K^- \pi^+ \pi^+$ measured in D^+ in (p_T, y) bins.

p_T [MeV/c]	y			
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]
[0, 1000]	112.17 $\pm$ 0.15	104.73 $\pm$ 0.08	99.51 $\pm$ 0.06	98.03 $\pm$ 0.06
[1000, 1500]	112.18 $\pm$ 0.16	105.73 $\pm$ 0.08	100.62 $\pm$ 0.06	98.51 $\pm$ 0.06
[1500, 2000]	112.48 $\pm$ 0.14	106.68 $\pm$ 0.08	101.50 $\pm$ 0.06	98.93 $\pm$ 0.05
[2000, 2500]	112.79 $\pm$ 0.13	107.53 $\pm$ 0.07	102.13 $\pm$ 0.06	99.17 $\pm$ 0.06
[2500, 3000]	113.03 $\pm$ 0.12	107.91 $\pm$ 0.08	102.30 $\pm$ 0.06	99.30 $\pm$ 0.05
[3000, 3500]	113.43 $\pm$ 0.11	108.25 $\pm$ 0.08	102.43 $\pm$ 0.06	99.18 $\pm$ 0.06
[3500, 4000]	113.22 $\pm$ 0.10	108.21 $\pm$ 0.08	102.41 $\pm$ 0.07	99.15 $\pm$ 0.05
[4000, 5000]	112.93 $\pm$ 0.07	108.17 $\pm$ 0.06	102.31 $\pm$ 0.05	98.88 $\pm$ 0.05
[5000, 6000]	112.02 $\pm$ 0.07	107.78 $\pm$ 0.06	102.02 $\pm$ 0.06	98.79 $\pm$ 0.05
[6000, 7000]	111.42 $\pm$ 0.07	107.27 $\pm$ 0.07	101.61 $\pm$ 0.09	98.86 $\pm$ 0.06
[7000, 8000]	110.42 $\pm$ 0.07	106.83 $\pm$ 0.09	101.54 $\pm$ 0.09	99.13 $\pm$ 0.08
[8000, 9000]	109.76 $\pm$ 0.07	106.61 $\pm$ 0.09	101.47 $\pm$ 0.11	99.53 $\pm$ 0.10
[9000, 10000]	109.02 $\pm$ 0.09	106.26 $\pm$ 0.08	101.20 $\pm$ 0.10	100.18 $\pm$ 0.08
[10000, 11000]	108.43 $\pm$ 0.02	105.94 $\pm$ 0.03	101.44 $\pm$ 0.03	100.35 $\pm$ 0.03
[11000, 12000]	107.96 $\pm$ 0.02	105.71 $\pm$ 0.03	101.57 $\pm$ 0.04	100.82 $\pm$ 0.03
[12000, 13000]	107.60 $\pm$ 0.03	105.58 $\pm$ 0.03	101.66 $\pm$ 0.05	101.18 $\pm$ 0.04
[13000, 14000]	107.23 $\pm$ 0.03	105.40 $\pm$ 0.04	101.87 $\pm$ 0.05	101.70 $\pm$ 0.07
[14000, 15000]	106.96 $\pm$ 0.04	105.40 $\pm$ 0.04	101.98 $\pm$ 0.06	101.93 $\pm$ 0.06
				104.16 $\pm$ 0.09

Table A.11: Selection efficiencies $\varepsilon_{\text{Selection}}/\varepsilon_{\text{L0}}$ for $D^0 \rightarrow K^- \pi^+$, excluding PID requirements, measured in D^0 (p_{T} , y) bins.

p_{T} [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	$3.99^{+0.16}_{-0.15}$	$7.53^{+0.13}_{-0.13}$	$8.21^{+0.13}_{-0.13}$	$6.82^{+0.14}_{-0.14}$	$3.66^{+0.17}_{-0.16}$
[1000, 1500]	$4.82^{+0.18}_{-0.17}$	$8.92^{+0.16}_{-0.16}$	$9.33^{+0.15}_{-0.15}$	$7.48^{+0.15}_{-0.15}$	$4.19^{+0.16}_{-0.16}$
[1500, 2000]	$6.00^{+0.20}_{-0.19}$	$10.51^{+0.17}_{-0.17}$	$11.20^{+0.17}_{-0.17}$	$8.61^{+0.16}_{-0.16}$	$4.75^{+0.17}_{-0.16}$
[2000, 2500]	$7.63^{+0.23}_{-0.22}$	$12.91^{+0.20}_{-0.20}$	$13.22^{+0.19}_{-0.19}$	$10.80^{+0.20}_{-0.19}$	$6.35^{+0.20}_{-0.20}$
[2500, 3000]	$9.89^{+0.27}_{-0.26}$	$15.94^{+0.24}_{-0.24}$	$16.82^{+0.24}_{-0.23}$	$13.26^{+0.24}_{-0.23}$	$7.09^{+0.24}_{-0.23}$
[3000, 3500]	$11.60^{+0.31}_{-0.30}$	$18.96^{+0.28}_{-0.28}$	$19.41^{+0.28}_{-0.28}$	$15.81^{+0.29}_{-0.28}$	$8.28^{+0.28}_{-0.27}$
[3500, 4000]	$13.98^{+0.37}_{-0.36}$	$22.09^{+0.34}_{-0.33}$	$21.65^{+0.34}_{-0.33}$	$18.12^{+0.35}_{-0.34}$	$7.90^{+0.32}_{-0.31}$
[4000, 5000]	$16.42^{+0.33}_{-0.32}$	$24.63^{+0.30}_{-0.30}$	$24.52^{+0.31}_{-0.30}$	$20.76^{+0.32}_{-0.32}$	$5.37^{+0.23}_{-0.22}$
[5000, 6000]	$18.76^{+0.44}_{-0.43}$	$26.55^{+0.41}_{-0.41}$	$28.94^{+0.44}_{-0.43}$	$20.47^{+0.43}_{-0.43}$	$2.83^{+0.24}_{-0.22}$
[6000, 7000]	$22.24^{+0.61}_{-0.60}$	$29.32^{+0.56}_{-0.55}$	$30.84^{+0.59}_{-0.58}$	$17.26^{+0.55}_{-0.53}$	$0.54^{+0.16}_{-0.12}$
[7000, 8000]	$23.69^{+0.78}_{-0.76}$	$29.99^{+0.73}_{-0.72}$	$30.35^{+0.77}_{-0.76}$	$12.80^{+0.65}_{-0.62}$	$0.00^{+0.08}_{-0.00}$
[8000, 9000]	$20.59^{+0.89}_{-0.86}$	$30.05^{+0.89}_{-0.88}$	$27.36^{+0.93}_{-0.91}$	$5.54^{+0.58}_{-0.53}$	$0.00^{+0.12}_{-0.00}$
[9000, 10000]	$23.94^{+0.28}_{-0.28}$	$32.02^{+0.28}_{-0.28}$	$26.22^{+0.29}_{-0.29}$	$2.95^{+0.13}_{-0.13}$	$0.00^{+0.01}_{-0.00}$
[10000, 11000]	$23.84^{+0.28}_{-0.28}$	$32.00^{+0.29}_{-0.29}$	$22.44^{+0.28}_{-0.28}$	$1.19^{+0.09}_{-0.08}$	$0.00^{+0.02}_{-0.00}$
[11000, 12000]	$25.63^{+0.35}_{-0.35}$	$32.18^{+0.36}_{-0.36}$	$17.53^{+0.32}_{-0.32}$	$0.16^{+0.04}_{-0.04}$	$0.00^{+0.02}_{-0.00}$
[12000, 13000]	$24.78^{+0.42}_{-0.41}$	$31.55^{+0.43}_{-0.43}$	$12.67^{+0.35}_{-0.34}$	$0.00^{+0.02}_{-0.00}$	$0.00^{+0.04}_{-0.00}$
[13000, 14000]	$24.90^{+0.50}_{-0.49}$	$30.32^{+0.51}_{-0.51}$	$9.03^{+0.36}_{-0.35}$	$0.00^{+0.03}_{-0.00}$	$0.00^{+0.06}_{-0.00}$
[14000, 15000]	$25.72^{+0.60}_{-0.59}$	$29.53^{+0.60}_{-0.59}$	$6.23^{+0.37}_{-0.35}$	$0.00^{+0.04}_{-0.00}$	$0.00^{+0.09}_{-0.00}$

Table A.12: Selection efficiencies $\varepsilon_{\text{Selection}}/\varepsilon_{\text{L0}}$ for $D^+ \rightarrow K^- \pi^+ \pi^+$, excluding PID requirements, measured in D^+ (p_{T} , y) bins.

p_{T} [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	$0.03^{+0.06}_{-0.02}$	$0.25^{+0.04}_{-0.04}$	$0.41^{+0.04}_{-0.04}$	$0.53^{+0.06}_{-0.05}$	$0.31^{+0.08}_{-0.06}$
[1000, 1500]	$0.87^{+0.18}_{-0.15}$	$2.72^{+0.14}_{-0.13}$	$3.93^{+0.14}_{-0.13}$	$3.96^{+0.15}_{-0.15}$	$2.91^{+0.22}_{-0.20}$
[1500, 2000]	$3.17^{+0.29}_{-0.27}$	$9.85^{+0.24}_{-0.23}$	$11.77^{+0.22}_{-0.22}$	$11.34^{+0.24}_{-0.24}$	$6.76^{+0.30}_{-0.29}$
[2000, 2500]	$9.19^{+0.45}_{-0.43}$	$18.92^{+0.32}_{-0.31}$	$21.18^{+0.29}_{-0.29}$	$19.92^{+0.32}_{-0.31}$	$13.35^{+0.41}_{-0.40}$
[2500, 3000]	$13.76^{+0.53}_{-0.51}$	$27.24^{+0.38}_{-0.38}$	$28.96^{+0.35}_{-0.35}$	$27.10^{+0.38}_{-0.38}$	$18.22^{+0.49}_{-0.48}$
[3000, 3500]	$18.38^{+0.61}_{-0.59}$	$30.83^{+0.43}_{-0.43}$	$32.71^{+0.40}_{-0.40}$	$30.52^{+0.43}_{-0.43}$	$21.20^{+0.56}_{-0.55}$
[3500, 4000]	$23.76^{+0.70}_{-0.69}$	$35.77^{+0.50}_{-0.50}$	$37.13^{+0.47}_{-0.46}$	$34.28^{+0.51}_{-0.50}$	$22.20^{+0.63}_{-0.62}$
[4000, 5000]	$27.50^{+0.59}_{-0.58}$	$39.06^{+0.42}_{-0.42}$	$40.65^{+0.41}_{-0.41}$	$37.21^{+0.45}_{-0.45}$	$21.44^{+0.53}_{-0.52}$
[5000, 6000]	$31.24^{+0.74}_{-0.73}$	$43.16^{+0.55}_{-0.55}$	$43.89^{+0.55}_{-0.55}$	$40.44^{+0.60}_{-0.60}$	$15.65^{+0.63}_{-0.61}$
[6000, 7000]	$34.11^{+0.92}_{-0.91}$	$47.20^{+0.72}_{-0.72}$	$47.79^{+0.73}_{-0.73}$	$37.78^{+0.81}_{-0.80}$	$9.86^{+0.73}_{-0.68}$
[7000, 8000]	$36.32^{+1.13}_{-1.11}$	$48.82^{+0.92}_{-0.92}$	$48.89^{+0.94}_{-0.93}$	$34.10^{+1.00}_{-0.99}$	$3.45^{+0.60}_{-0.52}$
[8000, 9000]	$37.10^{+1.42}_{-1.40}$	$48.45^{+1.14}_{-1.14}$	$49.11^{+1.21}_{-1.21}$	$28.72^{+1.27}_{-1.24}$	$1.22^{+0.56}_{-0.39}$
[9000, 10000]	$36.17^{+1.37}_{-1.35}$	$45.78^{+1.11}_{-1.11}$	$43.70^{+1.07}_{-1.07}$	$16.48^{+0.86}_{-0.83}$	$0.09^{+0.16}_{-0.07}$
[10000, 11000]	$38.90^{+0.37}_{-0.37}$	$51.04^{+0.32}_{-0.32}$	$46.46^{+0.35}_{-0.34}$	$13.76^{+0.29}_{-0.29}$	$0.00^{+0.02}_{-0.00}$
[11000, 12000]	$40.84^{+0.45}_{-0.44}$	$51.41^{+0.39}_{-0.39}$	$42.87^{+0.42}_{-0.42}$	$7.82^{+0.27}_{-0.26}$	$0.00^{+0.03}_{-0.00}$
[12000, 13000]	$39.70^{+0.53}_{-0.52}$	$50.90^{+0.47}_{-0.47}$	$38.34^{+0.50}_{-0.49}$	$4.96^{+0.27}_{-0.26}$	$0.00^{+0.04}_{-0.00}$
[13000, 14000]	$39.10^{+0.62}_{-0.61}$	$49.92^{+0.57}_{-0.57}$	$31.79^{+0.57}_{-0.57}$	$1.79^{+0.21}_{-0.19}$	$0.00^{+0.07}_{-0.00}$
[14000, 15000]	$39.81^{+0.73}_{-0.73}$	$49.61^{+0.67}_{-0.67}$	$26.61^{+0.66}_{-0.65}$	$1.40^{+0.23}_{-0.20}$	$0.00^{+0.10}_{-0.00}$

Table A.13: Particle identification selection efficiencies ϵ_{PID} for $D^0 \rightarrow K^- \pi^+$, measured in $D^0(p_T, y)$ bins.

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	82.79 ± 0.18	87.38 ± 0.16	92.27 ± 0.10	89.87 ± 0.14	80.50 ± 0.23
[1000, 1500]	82.94 ± 0.18	88.53 ± 0.13	92.36 ± 0.10	89.08 ± 0.12	75.97 ± 0.19
[1500, 2000]	82.67 ± 0.18	89.14 ± 0.10	92.69 ± 0.08	88.54 ± 0.10	70.86 ± 0.19
[2000, 2500]	82.55 ± 0.17	89.75 ± 0.09	92.89 ± 0.06	86.89 ± 0.09	65.13 ± 0.21
[2500, 3000]	82.82 ± 0.16	90.46 ± 0.08	93.13 ± 0.05	84.08 ± 0.09	62.42 ± 0.23
[3000, 3500]	83.27 ± 0.15	91.18 ± 0.07	93.15 ± 0.05	80.62 ± 0.10	59.19 ± 0.25
[3500, 4000]	83.90 ± 0.14	91.80 ± 0.06	92.95 ± 0.05	77.58 ± 0.11	56.20 ± 0.25
[4000, 5000]	85.25 ± 0.13	92.18 ± 0.06	91.55 ± 0.05	72.98 ± 0.14	50.63 ± 0.25
[5000, 6000]	86.12 ± 0.13	91.67 ± 0.06	88.16 ± 0.08	68.45 ± 0.16	43.58 ± 0.25
[6000, 7000]	85.85 ± 0.14	90.19 ± 0.07	83.81 ± 0.11	64.67 ± 0.19	43.96 ± 0.27
[7000, 8000]	83.73 ± 0.16	87.34 ± 0.10	80.01 ± 0.15	61.07 ± 0.22	-
[8000, 9000]	81.68 ± 0.19	83.08 ± 0.15	76.91 ± 0.19	57.29 ± 0.26	-
[9000, 10000]	78.78 ± 0.22	77.73 ± 0.24	74.72 ± 0.22	53.34 ± 0.28	-
[10000, 11000]	73.28 ± 0.31	74.36 ± 0.29	71.52 ± 0.25	56.66 ± 0.45	-
[11000, 12000]	72.42 ± 0.37	69.90 ± 0.48	68.96 ± 0.30	60.38 ± 0.65	-
[12000, 13000]	69.61 ± 0.44	64.19 ± 0.66	66.18 ± 0.33	-	-
[13000, 14000]	65.73 ± 0.50	58.92 ± 0.96	64.56 ± 0.36	-	-
[14000, 15000]	60.94 ± 0.80	57.57 ± 1.20	60.32 ± 0.41	-	-

Table A.14: Particle identification selection efficiencies ε_{PID} for $D^+ \rightarrow K^- \pi^+ \pi^+$, measured in D^0 (p_T, y) bins.

p_T [MeV/c]	[2, 2.5]	[2.5, 3]	y [3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	79.65 $\pm$ 0.24	82.98 $\pm$ 0.25	85.74 $\pm$ 0.32	84.03 $\pm$ 0.44	74.82 $\pm$ 0.47
[1000, 1500]	75.24 $\pm$ 0.22	81.55 $\pm$ 0.24	83.74 $\pm$ 0.39	81.50 $\pm$ 0.38	74.91 $\pm$ 0.34
[1500, 2000]	73.41 $\pm$ 0.23	80.45 $\pm$ 0.22	84.61 $\pm$ 0.32	82.24 $\pm$ 0.29	73.73 $\pm$ 0.28
[2000, 2500]	71.38 $\pm$ 0.24	80.27 $\pm$ 0.19	85.90 $\pm$ 0.21	83.11 $\pm$ 0.21	71.76 $\pm$ 0.24
[2500, 3000]	71.27 $\pm$ 0.25	80.84 $\pm$ 0.17	87.57 $\pm$ 0.15	83.36 $\pm$ 0.16	69.21 $\pm$ 0.21
[3000, 3500]	70.99 $\pm$ 0.25	82.05 $\pm$ 0.15	88.97 $\pm$ 0.12	83.04 $\pm$ 0.13	66.50 $\pm$ 0.21
[3500, 4000]	71.48 $\pm$ 0.24	83.44 $\pm$ 0.13	89.96 $\pm$ 0.10	82.02 $\pm$ 0.12	63.79 $\pm$ 0.22
[4000, 5000]	71.94 $\pm$ 0.23	85.04 $\pm$ 0.11	90.72 $\pm$ 0.08	79.72 $\pm$ 0.12	59.99 $\pm$ 0.23
[5000, 6000]	72.97 $\pm$ 0.22	86.40 $\pm$ 0.10	90.27 $\pm$ 0.07	76.76 $\pm$ 0.13	54.40 $\pm$ 0.26
[6000, 7000]	73.77 $\pm$ 0.20	87.10 $\pm$ 0.09	88.81 $\pm$ 0.08	72.86 $\pm$ 0.15	50.45 $\pm$ 0.27
[7000, 8000]	74.17 $\pm$ 0.20	86.97 $\pm$ 0.08	86.61 $\pm$ 0.09	69.86 $\pm$ 0.17	46.69 $\pm$ 0.28
[8000, 9000]	73.55 $\pm$ 0.20	85.55 $\pm$ 0.09	83.80 $\pm$ 0.11	67.11 $\pm$ 0.19	42.68 $\pm$ 0.35
[9000, 10000]	73.13 $\pm$ 0.21	83.65 $\pm$ 0.11	81.96 $\pm$ 0.14	65.24 $\pm$ 0.21	45.24 $\pm$ 0.54
[10000, 11000]	71.57 $\pm$ 0.22	80.85 $\pm$ 0.15	79.59 $\pm$ 0.16	61.67 $\pm$ 0.24	-
[11000, 12000]	69.75 $\pm$ 0.24	77.95 $\pm$ 0.25	77.79 $\pm$ 0.18	58.54 $\pm$ 0.26	-
[12000, 13000]	68.33 $\pm$ 0.27	75.09 $\pm$ 0.30	75.25 $\pm$ 0.21	58.82 $\pm$ 0.30	-
[13000, 14000]	66.76 $\pm$ 0.32	74.77 $\pm$ 0.22	73.82 $\pm$ 0.24	50.41 $\pm$ 0.46	-
[14000, 15000]	64.81 $\pm$ 0.41	72.99 $\pm$ 0.39	72.70 $\pm$ 0.26	50.29 $\pm$ 0.49	-

Table A.15: *Signal window efficiencies $\epsilon_{\text{Sig. win.}}$ for $D^0 \rightarrow K^- \pi^+$, measured in $D^0(p_T, y)$ bins.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	98.02 $\pm$ 0.04	97.92 $\pm$ 0.04	97.56 $\pm$ 0.04	96.79 $\pm$ 0.06	95.38 $\pm$ 0.17
[1000, 1500]	97.99 $\pm$ 0.04	97.86 $\pm$ 0.04	97.46 $\pm$ 0.04	96.60 $\pm$ 0.06	95.28 $\pm$ 0.15
[1500, 2000]	97.97 $\pm$ 0.04	97.82 $\pm$ 0.04	97.36 $\pm$ 0.04	96.26 $\pm$ 0.06	93.70 $\pm$ 1.90
[2000, 2500]	97.91 $\pm$ 0.04	97.76 $\pm$ 0.04	97.28 $\pm$ 0.04	96.11 $\pm$ 0.06	93.61 $\pm$ 0.17
[2500, 3000]	97.84 $\pm$ 0.04	97.72 $\pm$ 0.04	97.17 $\pm$ 0.04	95.93 $\pm$ 0.06	93.74 $\pm$ 0.13
[3000, 3500]	97.81 $\pm$ 0.04	97.68 $\pm$ 0.04	97.09 $\pm$ 0.04	95.93 $\pm$ 0.07	93.67 $\pm$ 0.08
[3500, 4000]	97.77 $\pm$ 0.04	97.60 $\pm$ 0.04	97.00 $\pm$ 0.05	95.82 $\pm$ 0.07	93.63 $\pm$ 0.09
[4000, 5000]	97.68 $\pm$ 0.04	97.53 $\pm$ 0.04	96.89 $\pm$ 0.05	95.66 $\pm$ 0.07	93.61 $\pm$ 0.07
[5000, 6000]	97.58 $\pm$ 0.04	97.41 $\pm$ 0.04	96.72 $\pm$ 0.05	95.72 $\pm$ 0.09	93.33 $\pm$ 0.20
[6000, 7000]	97.44 $\pm$ 0.05	97.20 $\pm$ 0.05	96.47 $\pm$ 0.07	95.53 $\pm$ 0.14	93.76 $\pm$ 0.34
[7000, 8000]	97.41 $\pm$ 0.05	97.09 $\pm$ 0.06	96.32 $\pm$ 0.08	95.28 $\pm$ 0.23	97.79 $\pm$ 0.04
[8000, 9000]	97.35 $\pm$ 0.06	96.84 $\pm$ 0.07	96.16 $\pm$ 0.11	94.62 $\pm$ 0.45	97.79 $\pm$ 0.04
[9000, 10000]	96.92 $\pm$ 0.10	96.78 $\pm$ 0.09	96.19 $\pm$ 0.15	93.76 $\pm$ 0.16	97.79 $\pm$ 0.04
[10000, 11000]	96.78 $\pm$ 0.14	96.73 $\pm$ 0.12	95.68 $\pm$ 0.24	92.90 $\pm$ 1.08	97.79 $\pm$ 0.04
[11000, 12000]	96.58 $\pm$ 0.18	96.53 $\pm$ 0.16	96.01 $\pm$ 0.31	96.30 $\pm$ 0.10	97.79 $\pm$ 0.04
[12000, 13000]	96.62 $\pm$ 0.21	96.09 $\pm$ 0.23	93.90 $\pm$ 0.28	97.79 $\pm$ 0.04	97.79 $\pm$ 0.04
[13000, 14000]	95.98 $\pm$ 0.32	96.24 $\pm$ 0.26	93.58 $\pm$ 0.25	97.79 $\pm$ 0.04	97.79 $\pm$ 0.04
[14000, 15000]	95.64 $\pm$ 0.41	95.22 $\pm$ 0.42	97.60 $\pm$ 0.33	97.79 $\pm$ 0.04	97.79 $\pm$ 0.04

Table A.16: Signal window efficiencies $\varepsilon_{\text{Sig. win.}}$ for $D^+ \rightarrow K^- \pi^+ \pi^+$, measured in D^+ (p_T, y) bins.

p_T [MeV/c]	[2, 2.5]	[2.5, 3]	y [3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	97.15 $\pm$ 0.50	97.28 $\pm$ 0.05	97.03 $\pm$ 0.08	96.29 $\pm$ 0.23	94.67 $\pm$ 0.88
[1000, 1500]	97.33 $\pm$ 0.04	97.30 $\pm$ 0.04	97.06 $\pm$ 0.04	96.14 $\pm$ 0.09	93.61 $\pm$ 0.09
[1500, 2000]	97.32 $\pm$ 0.04	97.30 $\pm$ 0.04	97.08 $\pm$ 0.04	96.34 $\pm$ 0.06	93.75 $\pm$ 0.08
[2000, 2500]	97.32 $\pm$ 0.04	97.31 $\pm$ 0.04	97.13 $\pm$ 0.04	96.26 $\pm$ 0.05	94.63 $\pm$ 0.14
[2500, 3000]	97.32 $\pm$ 0.04	97.31 $\pm$ 0.04	97.12 $\pm$ 0.04	96.28 $\pm$ 0.05	94.09 $\pm$ 0.12
[3000, 3500]	97.32 $\pm$ 0.04	97.30 $\pm$ 0.04	97.08 $\pm$ 0.04	96.28 $\pm$ 0.05	93.89 $\pm$ 0.12
[3500, 4000]	97.31 $\pm$ 0.04	97.29 $\pm$ 0.04	97.04 $\pm$ 0.04	96.16 $\pm$ 0.05	93.86 $\pm$ 0.13
[4000, 5000]	97.31 $\pm$ 0.04	97.28 $\pm$ 0.04	96.99 $\pm$ 0.04	95.94 $\pm$ 0.05	93.78 $\pm$ 0.07
[5000, 6000]	97.30 $\pm$ 0.04	97.25 $\pm$ 0.04	96.89 $\pm$ 0.04	95.76 $\pm$ 0.06	93.78 $\pm$ 1.62
[6000, 7000]	97.29 $\pm$ 0.04	97.22 $\pm$ 0.04	96.77 $\pm$ 0.05	95.43 $\pm$ 0.09	93.75 $\pm$ 0.12
[7000, 8000]	97.25 $\pm$ 0.04	97.19 $\pm$ 0.04	96.65 $\pm$ 0.05	95.31 $\pm$ 0.13	94.03 $\pm$ 0.38
[8000, 9000]	97.22 $\pm$ 0.04	97.10 $\pm$ 0.04	96.43 $\pm$ 0.07	95.58 $\pm$ 0.17	96.33 $\pm$ 1.07
[9000, 10000]	97.18 $\pm$ 0.05	97.01 $\pm$ 0.05	96.35 $\pm$ 0.09	95.78 $\pm$ 0.24	92.94 $\pm$ 2.08
[10000, 11000]	97.21 $\pm$ 0.05	96.97 $\pm$ 0.06	96.31 $\pm$ 0.11	95.67 $\pm$ 0.36	97.30 $\pm$ 0.04
[11000, 12000]	97.05 $\pm$ 0.06	96.90 $\pm$ 0.07	96.11 $\pm$ 0.15	96.01 $\pm$ 0.57	97.30 $\pm$ 0.04
[12000, 13000]	96.97 $\pm$ 0.09	96.63 $\pm$ 0.11	95.89 $\pm$ 0.23	96.34 $\pm$ 0.72	97.30 $\pm$ 0.04
[13000, 14000]	96.96 $\pm$ 0.10	96.69 $\pm$ 0.12	96.11 $\pm$ 0.26	97.09 $\pm$ 0.65	97.30 $\pm$ 0.04
[14000, 15000]	96.74 $\pm$ 0.15	96.14 $\pm$ 0.22	96.17 $\pm$ 0.35	96.58 $\pm$ 1.32	97.30 $\pm$ 0.04

Table A.17: *Relative uncertainty on the D^0 cross-section, in D^0 (p_T , y) bins, due to the finite size of the Monte Carlo sample used to assess the selection efficiency.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	3.91	1.81	1.62	2.07	4.66
[1000, 1500]	3.77	1.83	1.63	2.02	3.87
[1500, 2000]	3.39	1.69	1.54	1.96	3.57
[2000, 2500]	3.02	1.61	1.52	1.85	3.31
[2500, 3000]	2.83	1.56	1.47	1.84	3.40
[3000, 3500]	2.82	1.58	1.52	1.89	3.52
[3500, 4000]	2.82	1.62	1.64	2.01	4.06
[4000, 5000]	2.14	1.31	1.32	1.62	4.34
[5000, 6000]	2.53	1.68	1.62	2.20	8.15
[6000, 7000]	2.96	2.06	2.05	3.23	24.57
[7000, 8000]	3.56	2.61	2.70	5.06	-
[8000, 9000]	4.55	3.17	3.56	10.00	-
[9000, 10000]	1.39	1.03	1.24	4.52	-
[10000, 11000]	1.44	1.08	1.45	7.33	-
[11000, 12000]	1.74	1.37	2.11	-	-
[12000, 13000]	2.00	1.75	3.18	-	-
[13000, 14000]	2.54	2.28	4.38	-	-
[14000, 15000]	2.85	2.96	6.51	-	-

Table A.18: *Relative uncertainty on the D^+ cross-section, in D^+ (p_T , y) bins, due to the finite size of the Monte Carlo sample used to assess the selection efficiency.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	55.08	15.49	9.88	9.94	21.48
[1000, 1500]	18.04	4.93	3.43	3.84	7.29
[1500, 2000]	8.85	2.53	1.98	2.25	4.53
[2000, 2500]	5.00	1.83	1.51	1.75	3.29
[2500, 3000]	4.05	1.59	1.36	1.58	2.91
[3000, 3500]	3.64	1.60	1.40	1.62	2.95
[3500, 4000]	3.34	1.64	1.46	1.72	3.16
[4000, 5000]	2.47	1.29	1.19	1.41	2.73
[5000, 6000]	2.78	1.56	1.50	1.78	4.25
[6000, 7000]	3.21	1.89	1.88	2.48	7.39
[7000, 8000]	3.69	2.36	2.36	3.31	16.04
[8000, 9000]	4.55	2.91	3.01	4.87	34.43
[9000, 10000]	4.46	2.89	2.87	5.40	-
[10000, 11000]	1.29	0.93	1.03	2.50	-
[11000, 12000]	1.44	1.17	1.51	3.71	-
[12000, 13000]	1.79	1.40	1.92	5.94	-
[13000, 14000]	2.22	2.00	2.50	11.93	-
[14000, 15000]	2.87	2.44	3.53	-	-

Table A.19: *Relative uncertainty on the D^0 cross-section, in D^0 (p_T , y) bins, due to the finite size of the PID calibration sample.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	0.22	0.19	0.11	0.16	0.29
[1000, 1500]	0.21	0.15	0.11	0.13	0.25
[1500, 2000]	0.21	0.12	0.09	0.11	0.27
[2000, 2500]	0.21	0.10	0.07	0.10	0.33
[2500, 3000]	0.20	0.08	0.06	0.11	0.37
[3000, 3500]	0.18	0.07	0.06	0.12	0.42
[3500, 4000]	0.17	0.07	0.05	0.14	0.45
[4000, 5000]	0.16	0.06	0.06	0.19	0.50
[5000, 6000]	0.15	0.06	0.08	0.24	0.57
[6000, 7000]	0.17	0.08	0.13	0.29	0.62
[7000, 8000]	0.19	0.12	0.19	0.36	-
[8000, 9000]	0.23	0.17	0.24	0.44	-
[9000, 10000]	0.28	0.31	0.29	0.51	-
[10000, 11000]	0.41	0.38	0.35	0.79	-
[11000, 12000]	0.50	0.65	0.44	-	-
[12000, 13000]	0.63	1.01	0.50	-	-
[13000, 14000]	0.75	1.57	0.56	-	-
[14000, 15000]	1.20	2.01	0.67	-	-

Table A.20: *Relative uncertainty on the D^+ cross-section, in D^+ (p_T , y) bins, due to the finite size of the PID calibration sample.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	0.30	0.30	0.36	0.52	0.62
[1000, 1500]	0.29	0.29	0.44	0.44	0.46
[1500, 2000]	0.31	0.27	0.35	0.34	0.37
[2000, 2500]	0.34	0.23	0.24	0.24	0.33
[2500, 3000]	0.35	0.20	0.17	0.19	0.30
[3000, 3500]	0.35	0.18	0.13	0.16	0.31
[3500, 4000]	0.34	0.16	0.11	0.15	0.34
[4000, 5000]	0.32	0.13	0.08	0.14	0.39
[5000, 6000]	0.29	0.11	0.08	0.16	0.47
[6000, 7000]	0.28	0.10	0.09	0.20	0.53
[7000, 8000]	0.27	0.09	0.11	0.24	0.59
[8000, 9000]	0.28	0.11	0.14	0.29	0.81
[9000, 10000]	0.28	0.13	0.16	0.33	-
[10000, 11000]	0.31	0.18	0.20	0.39	-
[11000, 12000]	0.35	0.31	0.23	0.44	-
[12000, 13000]	0.38	0.38	0.28	0.50	-
[13000, 14000]	0.47	0.30	0.32	0.91	-
[14000, 15000]	0.63	0.52	0.36	-	-

Table A.21: *Relative uncertainty on the D^0 cross-section, in D^0 (p_T , y) bins, due to the choice of PID calibration binning.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	1.15	0.09	0.23	0.74	2.13
[1000, 1500]	1.43	0.98	0.33	0.26	1.86
[1500, 2000]	0.80	0.44	0.11	0.34	0.00
[2000, 2500]	0.52	0.11	0.14	0.88	0.05
[2500, 3000]	0.56	0.39	0.31	2.00	2.28
[3000, 3500]	0.88	0.53	0.51	2.95	3.90
[3500, 4000]	1.26	0.51	0.73	2.79	7.21
[4000, 5000]	1.58	0.40	1.28	1.41	10.83
[5000, 6000]	1.89	0.15	2.10	1.65	17.34
[6000, 7000]	1.93	0.20	1.83	5.37	44.91
[7000, 8000]	2.26	0.49	0.67	10.51	-
[8000, 9000]	2.54	0.38	0.37	12.97	-
[9000, 10000]	3.22	0.05	0.68	24.15	-
[10000, 11000]	4.47	0.34	3.82	30.49	-
[11000, 12000]	4.96	0.29	3.91	-	-
[12000, 13000]	7.81	0.32	4.69	-	-
[13000, 14000]	7.47	0.00	11.38	-	-
[14000, 15000]	6.43	3.10	13.40	-	-

Table A.22: *Relative uncertainty on the D^+ cross-section, in D^+ (p_T , y) bins, due to the choice of PID calibration binning.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	0.00	3.88	1.29	1.61	6.62
[1000, 1500]	9.07	8.28	2.51	2.57	4.27
[1500, 2000]	6.64	5.45	2.23	1.72	2.78
[2000, 2500]	3.94	3.03	1.32	1.04	1.69
[2500, 3000]	2.82	1.05	0.57	0.18	2.04
[3000, 3500]	2.01	0.20	0.13	0.61	1.99
[3500, 4000]	1.71	0.32	0.04	1.36	2.15
[4000, 5000]	1.50	0.79	0.44	1.72	1.57
[5000, 6000]	1.34	1.00	0.98	1.15	2.61
[6000, 7000]	1.19	0.84	1.32	0.38	0.00
[7000, 8000]	1.29	0.57	1.08	2.03	8.74
[8000, 9000]	1.57	0.32	0.75	3.19	0.00
[9000, 10000]	1.74	0.27	0.35	5.55	-
[10000, 11000]	2.00	0.00	0.00	7.63	-
[11000, 12000]	2.58	0.00	0.00	4.19	-
[12000, 13000]	3.08	0.40	0.00	11.46	-
[13000, 14000]	3.51	0.63	0.00	0.00	-
[14000, 15000]	3.53	1.31	0.02	-	-

Table A.23: *Relative uncertainty on the D^0 cross-section, in D^0 (p_T , y) bins, due to the finite calibration sample size, hadronic interaction uncertainty, and event multiplicity weighting in the tracking efficiency correction factor analysis.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	4.76	4.59	3.96	4.04	3.39
[1000, 1500]	5.05	4.79	4.15	4.24	3.74
[1500, 2000]	5.29	4.59	3.95	3.97	3.72
[2000, 2500]	5.07	4.22	3.75	3.65	3.57
[2500, 3000]	4.60	3.83	3.45	3.39	3.30
[3000, 3500]	4.19	3.57	3.24	3.30	3.13
[3500, 4000]	3.74	3.40	3.21	3.15	3.04
[4000, 5000]	3.43	3.25	3.10	3.06	2.99
[5000, 6000]	3.06	3.09	2.92	2.92	3.06
[6000, 7000]	2.91	2.98	3.00	2.88	3.13
[7000, 8000]	2.87	2.94	2.86	2.92	-
[8000, 9000]	2.86	2.93	2.87	3.01	-
[9000, 10000]	2.82	3.00	2.82	3.03	-
[10000, 11000]	2.81	3.04	2.82	3.07	-
[11000, 12000]	2.82	3.09	2.80	-	-
[12000, 13000]	2.84	3.10	2.81	-	-
[13000, 14000]	2.88	3.10	2.81	-	-
[14000, 15000]	2.93	3.05	2.83	-	-

Table A.24: *Relative uncertainty on the D^+ cross-section, in D^+ (p_T , y) bins, due to the finite calibration sample size, hadronic interaction uncertainty, and event multiplicity weighting in the tracking efficiency correction factor analysis.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	8.97	16.37	18.31	14.14	9.66
[1000, 1500]	8.91	14.45	15.30	12.29	9.26
[1500, 2000]	8.81	12.37	12.77	10.59	8.65
[2000, 2500]	8.48	10.16	10.63	9.16	7.89
[2500, 3000]	7.87	8.68	9.14	7.90	7.27
[3000, 3500]	6.99	7.37	7.92	7.13	6.48
[3500, 4000]	6.43	6.56	6.89	6.49	5.85
[4000, 5000]	5.88	5.68	6.28	6.02	5.46
[5000, 6000]	5.42	5.08	5.46	5.45	5.03
[6000, 7000]	5.05	4.74	5.17	5.05	4.74
[7000, 8000]	4.80	4.59	4.79	4.80	4.85
[8000, 9000]	4.64	4.44	4.66	4.64	4.77
[9000, 10000]	4.52	4.41	4.59	4.62	4.76
[10000, 11000]	4.45	4.42	4.46	4.63	-
[11000, 12000]	4.41	4.44	4.56	4.59	-
[12000, 13000]	4.39	4.46	4.50	4.62	-
[13000, 14000]	4.36	4.48	4.49	4.63	-
[14000, 15000]	4.36	4.52	4.45	4.71	-

Table A.25: *Relative uncertainty on the D^0 cross-section, in D^0 (p_T , y) bins, due to the arbitrary choice of fit model.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	0.12	0.14	0.21	0.37	0.33
[1000, 1500]	0.04	0.10	0.22	0.40	0.59
[1500, 2000]	0.04	0.11	0.25	0.45	0.54
[2000, 2500]	0.04	0.12	0.28	0.48	0.51
[2500, 3000]	0.04	0.15	0.31	0.39	0.46
[3000, 3500]	0.07	0.17	0.34	0.38	0.42
[3500, 4000]	0.08	0.22	0.38	0.38	0.40
[4000, 5000]	0.14	0.27	0.33	0.36	0.37
[5000, 6000]	0.20	0.36	0.33	0.33	0.57
[6000, 7000]	0.25	0.32	0.37	0.34	0.46
[7000, 8000]	0.35	0.33	0.37	0.39	-
[8000, 9000]	0.47	0.33	0.34	0.46	-
[9000, 10000]	0.54	0.33	0.33	1.14	-
[10000, 11000]	0.41	0.37	0.43	6.34	-
[11000, 12000]	0.52	0.48	0.47	-	-
[12000, 13000]	0.50	0.48	0.68	-	-
[13000, 14000]	0.56	0.59	1.57	-	-
[14000, 15000]	0.64	0.47	1.13	-	-

Table A.26: *Relative uncertainty on the D^+ cross-section, in D^+ (p_T , y) bins, due to the arbitrary choice of fit model.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	2.93	0.17	0.25	0.24	0.67
[1000, 1500]	0.39	0.02	0.38	0.44	0.18
[1500, 2000]	0.21	0.18	0.32	0.62	1.03
[2000, 2500]	0.10	0.19	0.44	0.86	1.04
[2500, 3000]	0.01	0.29	0.59	1.06	1.27
[3000, 3500]	0.05	0.37	0.79	1.19	1.38
[3500, 4000]	0.07	0.52	0.88	1.37	1.35
[4000, 5000]	0.27	0.70	1.09	1.41	1.47
[5000, 6000]	0.49	0.90	1.29	1.54	3.69
[6000, 7000]	0.67	1.17	1.41	1.80	1.64
[7000, 8000]	0.90	1.37	1.70	1.58	2.28
[8000, 9000]	1.01	1.72	1.62	1.62	4.26
[9000, 10000]	0.94	1.81	1.68	1.61	6.95
[10000, 11000]	1.21	1.44	1.37	1.77	-
[11000, 12000]	1.72	1.42	1.34	1.87	-
[12000, 13000]	1.29	1.43	1.58	3.26	-
[13000, 14000]	1.82	1.41	1.66	1.59	-
[14000, 15000]	1.72	1.37	1.54	0.59	-

Table A.27: *Relative uncertainty on the D^0 cross-section, in D^0 (p_T , y) bins, due to the statistical uncertainty parameters of the fit model used to compute the signal window fraction.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	0.04	0.04	0.04	0.06	0.18
[1000, 1500]	0.04	0.04	0.04	0.06	0.16
[1500, 2000]	0.04	0.04	0.04	0.06	2.01
[2000, 2500]	0.04	0.04	0.04	0.06	0.18
[2500, 3000]	0.04	0.04	0.04	0.07	0.14
[3000, 3500]	0.04	0.04	0.05	0.07	0.09
[3500, 4000]	0.04	0.04	0.05	0.08	0.10
[4000, 5000]	0.04	0.04	0.05	0.07	0.07
[5000, 6000]	0.04	0.04	0.05	0.09	0.21
[6000, 7000]	0.05	0.05	0.07	0.14	0.36
[7000, 8000]	0.06	0.06	0.09	0.24	-
[8000, 9000]	0.07	0.08	0.12	0.47	-
[9000, 10000]	0.10	0.09	0.15	0.17	-
[10000, 11000]	0.14	0.12	0.25	1.15	-
[11000, 12000]	0.19	0.17	0.33	-	-
[12000, 13000]	0.22	0.23	0.30	-	-
[13000, 14000]	0.33	0.27	0.27	-	-
[14000, 15000]	0.43	0.44	0.33	-	-

Table A.28: *Relative uncertainty on the D^+ cross-section, in D^+ (p_T , y) bins, due to the statistical uncertainty parameters of the fit model used to compute the signal window fraction.*

p_T [MeV/c]	y				
	[2, 2.5]	[2.5, 3]	[3, 3.5]	[3.5, 4]	[4, 4.5]
[0, 1000]	0.51	0.05	0.08	0.24	0.92
[1000, 1500]	0.04	0.04	0.05	0.09	0.09
[1500, 2000]	0.04	0.04	0.04	0.06	0.09
[2000, 2500]	0.04	0.04	0.04	0.05	0.14
[2500, 3000]	0.04	0.04	0.04	0.05	0.13
[3000, 3500]	0.04	0.04	0.04	0.05	0.12
[3500, 4000]	0.04	0.04	0.04	0.05	0.14
[4000, 5000]	0.04	0.04	0.04	0.05	0.07
[5000, 6000]	0.04	0.04	0.04	0.07	1.72
[6000, 7000]	0.04	0.04	0.05	0.09	0.13
[7000, 8000]	0.04	0.04	0.05	0.13	0.40
[8000, 9000]	0.04	0.04	0.07	0.17	1.11
[9000, 10000]	0.05	0.05	0.09	0.25	-
[10000, 11000]	0.05	0.06	0.11	0.38	-
[11000, 12000]	0.06	0.07	0.16	0.59	-
[12000, 13000]	0.09	0.11	0.24	0.74	-
[13000, 14000]	0.10	0.13	0.27	0.67	-
[14000, 15000]	0.16	0.23	0.36	-	-

B Decision trees

Decision trees [113, 114] are part of the set of multivariate analysis tools that also includes k -nearest neighbours and artificial neural networks. In high-energy physics, such tools are commonly used for the task of *classification*, where one wishes to discriminate between several species that are mixed together, usually in unknown proportions, in the data. Most generally, the algorithms try to estimate some function whose parameters are a set of variables $\vec{x}$, or *features*, and when used for binary classification the function is the decision boundary, either side of which is a pure sample of each of the two species. In practice, the generating PDFs for the species overlap to some degree, or the algorithm cannot distinguish them with certainty across the entire feature space. For a given set of values of the features, the algorithm then returns a probability that the point $\vec{x}$ belongs to a particular species. Throughout this appendix, only the problem of binary classification will be considered.

Rectangular cuts are a crude attempt at estimating the multidimensional decision boundary that separates signal from background, with the decision being either zero if at least one cut is not passed, or one if all cuts are passed. A decision tree refines this approach by defining cuts which are conditional on the outcome of those preceding it. This begins with a single *node*, the *root*, which defines a cut on a single feature x , say $x > y$ for some value y , by which the input sample is split in two. The sample passing the cut is passed to a node in the next *layer*, whilst the failing sample is passed to another node in that layer. Each of these nodes defines some other cut, possibly on a different feature, and the samples are further split and passed to nodes in the next layer. Each of the N layers can have up to 2^{n-1} nodes, where n

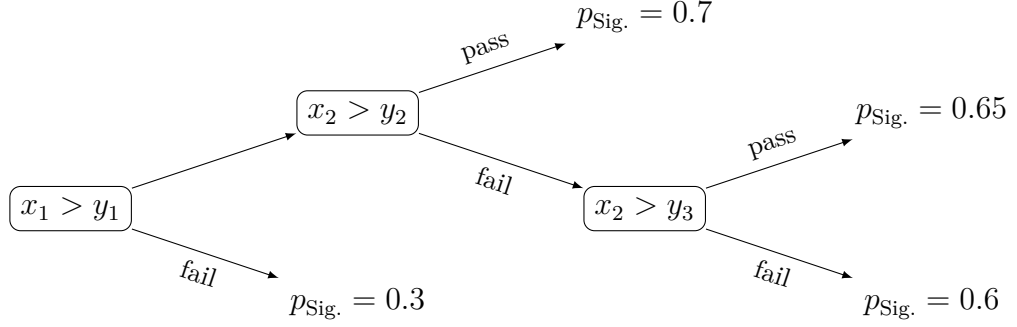


Figure B.1: A decision tree, computing signal probabilities $p_{\text{Sig.}}$ in the leaf nodes using a set of features x_i and cut values on those features y_i .

is the *depth* or layer number. Eventually the sub-samples can proceed no further and reside in a *leaf*, or terminal, node at which point a per-species probability is assigned to them. An example decision tree is illustrated in Figure B.1.

The construction, or growth, of a single decision tree is the process to determine which features and cut values are used in what nodes, and what the assigned probabilities should be in the terminal nodes. There are many parameters available for tuning during construction, including: the maximum depth; the figure of merit the cut at each node should optimise; and the stopping criteria that determines whether additional nodes can be created. Both tuning of these *hyper-parameters* and of the features and cuts to use is a complex optimisation problem [114], the details of which are not necessary for understanding the classification problem at hand. The important feature of the construction is that it is performed using an ensemble of individual *events*, feature vectors $\vec{x}$, called the *training* sample, in which the true species of each event is known. This is typically a combination of some simulated MC sample, known to be signal through truth-matching, and some sample from a control region in the data which is known to contain no signal, such as mass sidebands. The decision tree then tries to fit the set of cuts and values such that the samples can be classified as cleanly as possible, given the hyper-parameters. The probability of the terminal nodes can then be taken as, for example, the ratio of the

two species that remain after the training sample has been passed through the final tree.

The discriminating power of a decision tree is driven by two factors: how discriminatory the set of features is between the two species; and what are the values of the hyper-parameters. The exact choice of these is dependent on the nature of the classification problem, however the hyper-parameters should be chosen such that the decision tree is not *over-fitted* to the training data. This occurs when the tree is too sensitive to the statistical fluctuations in the features in the training sample. It is possible that a tree with infinitely many layers would be able to perfectly classify the training sample, for example, but a new signal sample not present in the training may be classified as background, as the values of the features do not exactly match those in the training sample. In general, over-fitting results in the performance of the tree being worse on an independent *testing* sample, within which the true species are also known, than on the training sample. When tuning the set of hyper-parameters, the performance of the tree should then be evaluated on the testing sample. This is comparable with the normal fitting of functions to data with maximum likelihood or χ^2 methods, in that more complex models can describe the data well, to an arbitrary degree for higher complexities, but the models begin to offer less general insight, instead being tailored to a specific dataset that happens to have been observed. Measures of performance include the *accuracy*, the number of correctly classified events relative to the total number, and the *error rate*, the number of wrong predictions relative to the total number. In the worst case, a classifier does no better at guessing event species than uniform random guessing.

The benefit of using decision trees over some other multivariate analysis (MVA) algorithms is that they can be intuitively interpreted: the tree can be visualised as such, and the decisions can be reasoned with by seeing how different features influence the lower depths of the tree. However, a tree that fits the training sample well is in general very deep, as only one feature enters the decision in each node. Deep trees are particularly susceptible to over-fitting, and so *ensemble* methods are often used.

Ensembles of trees and boosting

The error rate of a decision tree, the fraction of incorrectly classified events, is due to two properties of the classifier: the bias, the consistent deviation away from the true values, and the variance, the spread around the mean value [115]. A complex model, prone to over-fitting, may have a low bias but a high variance, whereas a simple model may have a high bias but a low variance.

To reduce the variance of a tree-based classifier, ensembles of complex trees can be created, where each tree is grown using a subset of the training data [116]. The classification for an event is then determined using the set of classifications produced by the ensemble, such as by taking the mean of that set. This does not increase the bias of the model if the predictions of the individual trees are uncorrelated. To guard against this possibility, the technique of *random forests* [114, 117] can be used, in which each node splitting can only use information from a random subset of the features, in addition to the random sub-sampling of the training data used to grow each tree in the ensemble.

The bias can be decreased by employing *boosting*, for which ADABOOST [107] is a commonly used algorithm. Rather than training an ensemble of complex models in parallel, ADABOOST trains a *sequence* of *weak learners* (decisions trees here), whose predictions are only slightly better than random guessing, and returns a weighted sum of their predictions as the final classification. This begins by growing a single, shallow tree where the i th vector in the training sample is weighted with a weight $w_i = 1/N$. The training sample is fed through the tree, and events which were classified incorrectly are assigned a larger weight, and those which were classified correctly are assigned a smaller weight. A new tree is grown using the new weights, and the procedure is iterated. With each iteration, events that are consistently incorrectly classified are given a greater importance in the training.

The generalisation of the ADABOOST algorithm is *gradient boosting* [118]. With ADABOOST, additional weak learners compensate for previous incor-

rect classifications by being trained with specially weighted data. In gradient boosting, additional weak learners try to correct the classification of the previous learners by fitting a regression model $f(\vec{x})$ to the residuals $(\hat{y}_i - f(\vec{x}_i))^2$ of the predicted classification $f(\vec{x}_i) = y_i$ and the true classification $\hat{y}_i$. In this way, gradient boosting tries to minimise the *loss function* that characterises how poorly the classifier is performing, and can be generalised by considering an arbitrary loss function $L(\hat{y}, f(x))$, which can be the residual.

References

- [1] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2016) 159, [arXiv:1510.0170](#), [Erratum: JHEP09,013(2016)].
- [2] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 5$ TeV*, [arXiv:1610.0223](#).
- [3] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021.
- [4] BaBar collaboration, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, [arXiv:hep-ex/0107013](#).
- [5] Belle collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, [arXiv:hep-ex/0107061](#).
- [6] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in $B^\pm \rightarrow DK^\pm$ decays*, Phys. Lett. **B712** (2012) 203, [arXiv:1203.3662](#), [Erratum: Phys. Lett. **B713**,351(2012)].
- [7] LHCb collaboration, R. Aaij *et al.*, *First observation of CP violation in the decays of B_s^0 mesons*, Phys. Rev. Lett. **110** (2013), no. 22 221601, [arXiv:1304.6173](#).
- [8] LHCb collaboration, R. Aaij *et al.*, *Measurement of matter-antimatter differences in beauty baryon decays*, [arXiv:1609.0521](#).

- [9] Y. Grossman, A. L. Kagan, and Y. Nir, *New physics and CP violation in singly Cabibbo suppressed D decays*, Phys. Rev. **D75** (2007) 036008, [arXiv:hep-ph/0609178](#).
- [10] 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](#).
- [11] LHCb collaboration, R. Aaij *et al.*, *Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s} = 7$ TeV*, JHEP **08** (2013) 117, [arXiv:1306.3663](#).
- [12] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [13] Belle, S. K. Choi *et al.*, *Observation of a resonance-like structure in the $\pi^\pm\psi'$ mass distribution in exclusive $B \rightarrow K\pi^\pm\psi'$ decays*, in *Lepton and photon interactions at high energies. Proceedings, 23rd International Symposium, LP2007, Daegu, South Korea, August 13-18, 2007*, vol. 100, p. 142001, 2008. [arXiv:0708.1790](#). doi: 10.1103/PhysRevLett.100.142001.
- [14] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014), no. 22 222002, [arXiv:1404.1903](#).
- [15] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, [arXiv:1507.0341](#).
- [16] C. S. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413.
- [17] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.
- [18] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, [arXiv:1207.7214](#).

- [19] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, [arXiv:1207.7235](#).
- [20] ATLAS and CMS collaborations, G. Aad *et al.*, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, [arXiv:1606.0226](#).
- [21] The CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005), no. 1 1, [arXiv:hep-ph/0406184](#), and updated results and plots available at <http://ckmfitter.in2p3.fr>.
- [22] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.
- [23] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of $B \rightarrow DK$ analyses*, LHCb-CONF-2016-001.
- [24] J. Brod and J. Zupan, *The ultimate theoretical error on γ from $B \rightarrow DK$ decays*, JHEP **01** (2014) 051, [arXiv:1308.5663](#).
- [25] J. Brod, A. Lenz, G. Tetlalmatzi-Xolocotzi, and M. Wiebusch, *New physics effects in tree-level decays and the precision in the determination of the quark mixing angle*, Phys. Rev. **D92** (2015), no. 3 033002, [arXiv:1412.1446](#).
- [26] O. S. Bruning *et al.*, *The LHC Main Ring*, LHC Design Report **1** (2004).
- [27] ATLAS collaboration, M. Aaboud *et al.*, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, [arXiv:1606.0262](#).
- [28] CMS collaboration, *Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV*, CMS-PAS-FSQ-15-005 (2016).

- [29] LHCb collaboration, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [30] LHCb collaboration, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015), no. 07 1530022, [arXiv:1412.6352](#).
- [31] CDF collaboration, A. Abulencia *et al.*, *Observation of $B_s^0 - \bar{B}_s^0$ Oscillations*, Phys. Rev. Lett. **97** (2006) 242003, [arXiv:hep-ex/0609040](#).
- [32] LHCb collaboration, R. Aaij *et al.*, *Observation of $D^0 - \bar{D}^0$ oscillations*, Phys. Rev. Lett. **110** (2013), no. 10 101802, [arXiv:1211.1230](#).
- [33] LHCb and CMS collaborations, V. Khachatryan *et al.*, *Observation of the rare $B_s^0 \rightarrow \mu^+ \mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature **522** (2015) 68, [arXiv:1411.4413](#).
- [34] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, [arXiv:1512.0444](#).
- [35] LHCb collaboration, R. Aaij *et al.*, *Measurement of b -hadron branching fractions for two-body decays into charmless charged hadrons*, JHEP **10** (2012) 037, [arXiv:1206.2794](#).
- [36] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, [arXiv:1408.1251](#).
- [37] G. Dujany and B. Storaci, *Real-time alignment and calibration of the LHCb Detector in Run II*, Journal of Physics: Conference Series **664** (2015), no. 8 082010.
- [38] R. W. Forty and O. Schneider, *RICH pattern recognition*, LHCb-98-040.
- [39] LHCb collaboration, P. Perret and X. Vilasís-Cardona, *Performance of the LHCb calorimeters during the period 2010-2012*, J. Phys. Conf. Ser. **587** (2015), no. 1 012012.

- [40] M Frank and C Gaspar and E v Herwijnen and B Jost and N Neufeld, *Deferred High Level Trigger in LHCb: A Boost to CPU Resource Utilization*, Journal of Physics: Conference Series **513** (2014), no. 1 012006.
- [41] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, LHCb-PUB-2011-016.
- [42] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, LHCb-PUB-2011-016.
- [43] S. Benson, V. V. Gligorov, M. A. Vesterinen, and J. M. Williams, *The LHCb Turbo Stream*, J. Phys. : Conf. Ser. **664** (2015) 082004. 8 p.
- [44] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **10** (2015) 172, [arXiv:1509.0077](#).
- [45] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, [arXiv:0710.3820](#).
- [46] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, LHCb-PROC-2011-005.
- [47] LHCb collaboration, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [48] V. V. Gligorov, *Reconstruction of the Channel $B_d^0 \rightarrow D^+\pi^-$ and Background Classification at LHCb (revised)*, LHCb-2007-044.
- [49] CLEO collaboration, J. P. Alexander *et al.*, *Absolute Measurement of Hadronic Branching Fractions of the D_s^+ Meson*, Phys. Rev. Lett. **100** (2008) 161804, [arXiv:0801.0680](#).
- [50] Particle Data Group, C. Amsler *et al.*, *Review of Particle Physics*, Phys. Lett. **B667** (2008) 1.

- [51] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1, [arXiv:hep-ph/0409313](#).
- [52] S. Forte and G. Watt, *Progress in the Determination of the Partonic Structure of the Proton*, Ann. Rev. Nucl. Part. Sci. **63** (2013) 291, [arXiv:1301.6754](#).
- [53] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438, [Yad. Fiz.15,781(1972)].
- [54] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641, [Zh. Eksp. Teor. Fiz.73,1216(1977)].
- [55] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298.
- [56] T. Kneesch, B. A. Kniehl, G. Kramer, and I. Schienbein, *Charmed-meson fragmentation functions with finite-mass corrections*, Nucl. Phys. **B799** (2008) 34, [arXiv:0712.0481](#).
- [57] ZEUS and H1 collaborations, H. Abramowicz *et al.*, *Combination and QCD Analysis of Charm Production Cross Section Measurements in Deep-Inelastic ep Scattering at HERA*, Eur. Phys. J. **C73** (2013), no. 2 2311, [arXiv:1211.1182](#).
- [58] PROSA collaboration, O. Zenaiev *et al.*, *Impact of heavy-flavour production cross sections measured by the LHCb experiment on parton distribution functions at low x* , Eur. Phys. J. **C75** (2015), no. 8 396, [arXiv:1503.0458](#).
- [59] ALICE collaboration, B. Abelev *et al.*, *Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 2.76$ TeV*, JHEP **07** (2012) 191, [arXiv:1205.4007](#).

- [60] ALICE collaboration, B. Abelev *et al.*, D_s^+ meson production at central rapidity in proton–proton collisions at $\sqrt{s} = 7$ TeV, Phys. Lett. **B718** (2012) 279, [arXiv:1208.1948](#).
- [61] ALICE collaboration, B. Abelev *et al.*, Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 7$ TeV, JHEP **01** (2012) 128, [arXiv:1111.1553](#).
- [62] ATLAS collaboration, G. Aad *et al.*, Measurement of $D^{*\pm}$, $D^\pm$ and $D_s^\pm$ meson production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, Nucl. Phys. **B907** (2016) 717, [arXiv:1512.0291](#).
- [63] R. Gauld, J. Rojo, L. Rottoli, and J. Talbert, Charm production in the forward region: constraints on the small- x gluon and backgrounds for neutrino astronomy, JHEP **11** (2015) 009, [arXiv:1506.0802](#).
- [64] IceCube collaboration, A. Achterberg *et al.*, First Year Performance of The IceCube Neutrino Telescope, Astropart. Phys. **26** (2006) 155, [arXiv:astro-ph/0604450](#).
- [65] A. Bhattacharya *et al.*, Perturbative charm production and the prompt atmospheric neutrino flux in light of RHIC and LHC, JHEP **06** (2015) 110, [arXiv:1502.0107](#).
- [66] B. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, Inclusive charmed-meson production at the CERN LHC, Eur. Phys. J. **C72** (2012) 2082, [arXiv:1202.0439](#).
- [67] M. Cacciari, M. L. Mangano, and P. Nason, Gluon PDF constraints from the ratio of forward heavy quark production at the LHC at $\sqrt{s} = 7$ and 13 TeV, [arXiv:1507.0619](#).
- [68] LHCb collaboration, R. Aaij *et al.*, Precision luminosity measurements at LHCb, JINST **9** (2014) P12005, [arXiv:1410.0149](#).
- [69] LHCb collaboration, R. Aaij *et al.*, Search for CP violation in $D^+ \rightarrow K^- K^+ \pi^+$ decays, Phys. Rev. **D84** (2011) 112008, [arXiv:1110.3970](#).

- [70] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $D_s^+ - D_s^-$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B713** (2012) 186, [arXiv:1205.0897](#).
- [71] M. Schmelling, *Event loss rates and readout chips*, LHCb-98-057.
- [72] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, JINST **10** (2015), no. 02 P02011, [arXiv:1411.5528](#).
- [73] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, PhD thesis, Cracow, INP, 1986.
- [74] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. **A362** (1995) 487.
- [75] M. De Cian, *Track Reconstruction Efficiency and Analysis of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ at the LHCb Experiment*, PhD thesis, Zurich U., 2013-09-23.
- [76] LHCb RICH group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759](#).
- [77] M. Pivk and F. R. Le Diberder, *$_s\text{Plot}$: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](#).
- [78] M. Paterno, *Calculating efficiencies and their uncertainties*, , FERMILAB-TM-2286-CD.
- [79] B. A. C. Alan Agresti, *Approximate is better than "exact" for interval estimation of binomial proportions*, The American Statistician **52** (1998), no. 2 119.
- [80] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP violation parameter A_Γ with 2012 data*, LHCb-CONF-2016-010.

- [81] M. Lisovyi, A. Verbytskyi, and O. Zenaiev, *Combined analysis of charm-quark fragmentation-fraction measurements*, Eur. Phys. J. **C76** (2016), no. 7 397, [arXiv:1509.0106](#).
- [82] CLEO collaboration, M. Artuso *et al.*, *Charm meson spectra in e^+e^- annihilation at 10.5 GeV center of mass energy*, Phys. Rev. **D70** (2004) 112001, [arXiv:hep-ex/0402040](#).
- [83] Belle collaboration, R. Seuster *et al.*, *Charm hadrons from fragmentation and B decays in e^+e^- annihilation at $\sqrt{s} = 10.6$ GeV*, Phys. Rev. **D73** (2006) 032002, [arXiv:hep-ex/0506068](#).
- [84] BaBar collaboration, B. Aubert *et al.*, *Measurement of D_s^+ and D_s^{*+} production in B meson decays and from continuum e^+e^- annihilation at $\sqrt{s} = 10.6$ GeV*, Phys. Rev. **D65** (2002) 091104, [arXiv:hep-ex/0201041](#).
- [85] L. Lyons, D. Gibaut, and P. Clifford, *How to combine correlated estimates of a single physical quantity*, Nucl. Instrum. Meth. **A270** (1988) 110.
- [86] LHCb collaboration, R. Aaij *et al.*, *Study of cold nuclear matter effects using prompt D^0 meson production in pPb collisions at LHCb*, LHCb-CONF-2016-003.
- [87] NA48 collaboration, J. R. Batley *et al.*, *A Precision measurement of direct CP violation in the decay of neutral kaons into two pions*, Phys. Lett. **B544** (2002) 97, [arXiv:hep-ex/0208009](#).
- [88] LHCb collaboration, R. Aaij *et al.*, *Evidence for CP violation in time-integrated $D^0 \rightarrow h^- h^+$ decay rates*, Phys. Rev. Lett. **108** (2012) 111602, [arXiv:1112.0938](#).
- [89] A. Lenz, *What did we learn in theory from the ΔA_{CP} -saga?*, in *Proceedings, 6th International Workshop on Charm Physics (Charm 2013): Manchester, UK, August 31-September 4, 2013*, 2013. [arXiv:1311.6447](#).

- [90] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \rightarrow K^- K^+$ and $D^0 \rightarrow \pi^- \pi^+$ decays*, JHEP **07** (2014) 041, [arXiv:1405.2797](#).
- [91] LHCb collaboration, R. Aaij *et al.*, *Measurement of the difference of time-integrated CP asymmetries in $D^0 \rightarrow K^- K^+$ and $D^0 \rightarrow \pi^- \pi^+$ decays*, Phys. Rev. Lett. **116** (2016), no. 19 191601, [arXiv:1602.0316](#).
- [92] Heavy Flavor Averaging Group (HFAG), Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of summer 2014*, [arXiv:1412.7515](#), and online updates at <http://www.slac.stanford.edu/xorg/hfag>.
- [93] M. Gersabeck *et al.*, *On the interplay of direct and indirect CP violation in the charm sector*, J. Phys. **G39** (2012) 045005, [arXiv:1111.6515](#).
- [94] E791 collaboration, E. M. Aitala *et al.*, *Multidimensional resonance analysis of $\Lambda_c^+ \rightarrow p K^- \pi^+$* , Phys. Lett. **B471** (2000) 449, [arXiv:hep-ex/9912003](#).
- [95] I. I. Bigi, *Probing CP Asymmetries in Charm Baryons Decays*, [arXiv:1206.4554](#).
- [96] BESIII collaboration, M. Ablikim *et al.*, *Measurement of Singly Cabibbo-Suppressed Decays $\Lambda_c^+ \rightarrow p \pi^+ \pi^-$ and $\Lambda_c^+ \rightarrow p K^+ K^-$* , [arXiv:1608.0040](#).
- [97] LHCb collaboration, R. Aaij *et al.*, *Study of the production of Λ_b^0 and $\bar{B}^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \rightarrow J/\psi p K^-$ branching fraction*, Chin. Phys. **C40** (2016), no. 1 011001, [arXiv:1509.0029](#).
- [98] J. L. Rosner, *Asymmetry in Λ_b and $\bar{\Lambda}_b$ production*, Phys. Rev. D **90** (2014) 014023.
- [99] M. Vesterinen, *Considerations on the LHCb dipole magnet polarity reversal*, LHCb-PUB-2014-006. On behalf of the LHCb collaboration.

- [100] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, [arXiv:physics/0503191](#).
- [101] J. Brehmer, J. Albrecht, and P. Seyfert, *Ghost probability: an efficient tool to remove background tracks*, LHCb-INT-2012-025.
- [102] CDF collaboration, T. Aaltonen *et al.*, *Evidence for the charmless annihilation decay mode $B_s^0 \rightarrow \pi^+ \pi^-$* , Phys. Rev. Lett. **108** (2012) 211803, [arXiv:1111.0485](#).
- [103] M. Artuso *et al.*, *Measurement of the CP violating asymmetry a_{sl}^s* , LHCb-ANA-2016-004.
- [104] P. Koppenburg, *Dealing with Multiple Candidates*, LHCb-INT-2011-009.
- [105] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits*, [arXiv:1210.7141](#).
- [106] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, [arXiv:1608.0580](#).
- [107] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, Journal of Computer and System Sciences **55** (1997), no. 1 119 .
- [108] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.
- [109] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, .
- [110] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in $B_s^0 - \bar{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016), no. 6 061803, [arXiv:1605.0976](#).
- [111] M. Stahl and M. Vesterinen, *Tracking asymmetry measurement with $J/\psi \rightarrow \mu^+ \mu^-$ decays*, LHCb-INT-2015-012.

- [112] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\bar{B}^0 - B^0$ and $\bar{B}_s^0 - B_s^0$ production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218, [arXiv:1408.0275](#).
- [113] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*, CRC press, 1984.
- [114] G. Louppe, *Understanding Random Forests: From Theory to Practice*, [arXiv:1407.7502](#).
- [115] L. Breiman, *Bias, variance, and arcing classifiers*, tech. rep., Statistics Department, University of California, Apr, 1996.
- [116] L. Breiman, *Bagging predictors*, Machine Learning **24** (1996), no. 2 123.
- [117] L. Breiman, *Random forests*, Machine Learning **45** (2001), no. 1 5.
- [118] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, Ann. Statist. **29** (2001) 1189.