Syracuse University

SURFACE at Syracuse University

Dissertations - ALL

SURFACE at Syracuse University

Winter 12-22-2021

First Measurement of B(ξc0 → Π−λ+c)

Michael Kent Wilkinson
Syracuse University

Follow this and additional works at: https://surface.syr.edu/etd

Part of the Physics Commons

Recommended Citation

Wilkinson, Michael Kent, "First Measurement of B(ξc0 → Π−λ+c)" (2021). *Dissertations - ALL*. 1365.
https://surface.syr.edu/etd/1365

This Dissertation is brought to you for free and open access by the SURFACE at Syracuse University at SURFACE at Syracuse University. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.

Abstract

The first measurement of the branching fraction of the Ξ_c^0 baryon in its decay to π^- and Λ_c^+ hadrons, $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$, is presented. This measurement was conducted with the LHCb detector, located on the Large Hadron Collider, using 13 TeV proton-proton collisions. There have been a number of theoretical predictions for the size of this branching fraction, spanning more than an order of magnitude; this large difference is primarily due to differing assumptions about the internal interactions contributing to the decay. The branching fraction measured here provides the first experimental evidence for the nature of these interactions; it is consistent with the larger predictions and inconsistent with the smaller ones, indicating a constructive interference effect. In addition, the second measurement of the branching fraction of the Ξ_c^+ baryon in its decay to p, K^-, and π^+ hadrons, $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$, is presented and found to be consistent with the Belle collaboration's previous measurement while carrying a lower relative uncertainty.

First Measurement of $\boldsymbol{\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)}$

by

Michael K. Wilkinson

B.S. in Physics, the University of Oklahoma, 2013

B.A., the University of Oklahoma, 2013

Dissertation

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

Syracuse University

December 2021

Copyright © Michael K. Wilkinson 2021

All Rights Reserved

Acknowledgements

My years in graduate school have been some of the most challenging and rewarding of my life, and the production of this dissertation would not have been possible without the advice and support of people too numerous to list here exhaustively. I will nonetheless attempt to highlight a few of the primary contributors.

First and foremost, I would like to posthumously thank my advisor, Sheldon Stone, whose guidance and scientific insight have been essential to my graduate studies and whose loss I feel every day.

Many thanks also to my advisor, Steve Blusk, for his extraordinary support and generosity and for stepping in during a difficult time. His advice and help have been crucial to the conclusion of my graduate studies.

Thanks to my undergraduate advisor, Mike Strauss, for starting me on this path and for giving good advice when I needed it most.

I would also like to thank Marina Artuso, Vanya Belyaev, Lars Eklund, Michael Eppard, and Jennifer Ross for their help navigating physics.

Thanks to Sergey Barsuk, Arantza Campos, Lars Eklund, Tara Nanut, and the rest of the Charm Working Group for reviewing this measurement.

Thanks to Ray Mountain for his guidance in all things hardware, support of my career ambitions, and willingness to discuss the finer points of English grammar.

Thanks to Andy Beiter and Arvind Venkateswaran for creating the best office I've ever worked in; to the whole Syracuse LHCb group for being excellent scientists and collaborators; and to Scott Bassler, Kyle Bishop, Shelby Buffington, Sarah Callihan, Wayne Chou, Alex Vincent Jannini, and Monica Ripp for being themselves.

Thanks to my parents, Kent and Cheryl Wilkinson, and to my sister, Rachel Daniels, for making all this possible.

Finally, thanks to Sarah Fram. There's no one I'd rather do this with.

For Clair & Griz

Contents

Lists of illustrative material **xv**

List of tables xv

List of figures xxvii

1 Introduction **1**

2 The Standard Model of particle physics **2**

2.1 Overview 2

2.2 Feynman diagrams 7

2.3 Isospin and SU(3) symmetry 8

2.4 Partially conserved axial current 10

2.5 Heavy-quark effective theory 11

2.6 Dalitz plots 12

3 The LHCb experiment **15**

3.1 The Large Hadron Collider 15

3.2 The LHCb detector 16

3.2.1 Magnet 18

3.2.2 Vertex Locator 19

3.2.3 Tracker Turicensis 20

3.2.4 Inner Tracker 22

3.2.5 Outer Tracker 23

3.2.6 Calorimeters 25

3.2.7 Ring Imaging Cherenkov detectors 26

3.2.8 Muon system 26

3.2.9 Trigger 27

3.3 The LHCb data flow 30

4 First branching fraction measurement of the suppressed decay $\Xi_c^0 \to \pi^- \Lambda_c^+$ 33

4.1 Introduction 33

4.2 Data sample 37

4.3 Event selections 38

4.3.1 Turbo selections 39

4.3.2 Trigger selections 39

4.3.3 Offline selections 41

4.3.4 Additional simulation selections 42

4.3.5 BDT selections 45

4.3.6 BDT performance 48

4.3.7 BDT cut selection 48

4.4 Efficiencies 52

4.4.1 Correlations between datasets 54

4.4.2 Binning in p_{T} and η 55

4.5 Signal yield 56

4.5.1 Fit model 56

4.5.2 Fit results 59

4.6 Results 66

4.6.1 Intrinsic results 66

4.6.2 Derived results 70

4.7 Systematic uncertainties 76

4.7.1 Heavy-quark symmetry 76

4.7.2 Charm-baryon lifetimes 77

4.7.3 Particle identification 77

4.7.4 Tracking 78

4.7.5 Fit model 78

4.7.6 Ξ_c^0 fit model 79

4.7.7 Λ_c^+ fit model 79
4.7.8 Ξ_c^+ fit model 84
4.7.9 Dalitz weighting 87
4.7.10 BDT input parameter differences in data and simulation 87
4.7.11 b-baryon decay contamination 90
4.7.12 Differences in integrated luminosity 96
4.7.13 Summary 97
4.8 Comparison with experiment 99
4.9 Comparison with theory 100

5 Conclusions **102**

A $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ Appendices **104**
A.1 Background categories in simulation 104
A.2 BDT response in each fold 108
A.3 BDT input variable rankings 113
A.4 BDT input parameters 114
A.5 Wrong-mass discrimination 123
A.6 Fits in each bin of p_{T} and η 126
A.7 Fitted parameter values 141
A.8 Λ_c^+ and Ξ_c^+ momentum distributions 147
A.9 Ξ_c^0 in Λ_c^+ sidebands 149
A.10 Dalitz weighting 150
A.11 TISTOS trigger efficiencies 150
A.12 Updated Gronau and Rosner predictions 151

References **159**

Vita **160**

List of Tables

1 Software used to generate simulation for the Ξ_c^0 and Λ_c^+ samples. The software for the Ξ_c^+ sample sometimes differs; where this is the case, it is given in parentheses. 38

2 Simulation samples (2016, 2017, and 2018 combined) used for BDT training and efficiency studies before any selections have been applied. See Section 4.4 for explanations of the generator-level cuts. 38

3 Λ_c^+ decay models used. The rates listed are from the PDG [1], but they were not adjusted to account for the branching fractions of the intermediate states, *e.g.*, $\overline{K}^{*0} \to K^-\pi^+$. Since we weight the Dalitz spectra to match data, however, this should have a negligible impact on the results; see Sections 4.3.4 and 4.7.9. 39

4 Selections made in the Turbo stream. See text. H_c is either the signal Λ_c^+ or the signal Ξ_c^+ baryon, as appropriate. χ^2_{IP} is the difference in the PV vertex fit χ^2 with and without the p, K^-, and π^+ tracks. GhostProb is the probability the track is a ghost. ndf is the number of degrees of freedom in the fit. HasRich indicates whether the track has information in the RICH. $\text{DLL}_{p\pi}$ is the difference in log likelihood that a particle is a proton rather than a π meson, similarly for other particles. $\vec{x}$ is the position vector pointing from the PV to the particle's decay vertex; $\frac{\vec{x}(H_c)\cdot\vec{p}(H_c)}{x(H_c)\cdot p(H_c)}$ is thus equivalent to the cosine of the angle between the particle's momentum and its decay vertex. See Section 4.3.2 for an explanation of the trigger requirements. 40

5 Selections made to produce Ξ_c^0 candidates. See text. χ^2_{vtx} is the fit χ^2 of the decay vertex. 41

6 Trigger requirements. The signal charm-baryon must pass at least one of the listed trigger lines for each stage of the trigger (TOS). The “signal charm-baryon” is defined as the signal Λ_c^+ baryon candidate for the $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ and $\Lambda_c^+ \to pK^-\pi^+$ samples and as the signal Ξ_c^+ baryon for the $\Xi_c^+ \to pK^-\pi^+$ sample. `L0HadronDecision` is based on the response of the hadronic calorimeters; the `Hlt1TrackMVA` lines select events with quality tracks formed in HLT1. 41

7 Offline selections. See text. Selections requiring a parameter to be > 0 ensure DTF completed the fit successfully, as negative values would indicate a failed fit. 42

8 Charm-baryon lifetime values. Measurements are taken from Reference [2], where uncertainties are statistical, systematic, and due to the uncertainty on the lifetime of the D^+ meson, respectively. The simulated values are those used by EvtGen; see text. 43

9 BDT input parameters. See text. ProbNN$_{\text{PID}}$ denotes the output of a probabilistic neural-network deciding how likely it is that a given track has been assigned the correct mass. The input distributions can be found in Appendix A.4. 46

10 Number of candidates used for BDT training and testing. The choice to use 99,999 background candidates is arbitrary; we avoid using the whole dataset in order to speed up processing time. All candidates have passed the Turbo, offline, and trigger selections. None of the weights described in Section 4.3.4 have been applied; these are raw candidate numbers. 48

11 Selection efficiencies, in the range described in Section 4.4.2, only statistical uncertainties shown. Turbo and `DaughtersInLHCb` are given relative to generator-level. Offline and BDT are each given relative to the previous one (left-to-right). "Ghost fix" is the efficiency adjustment for ghost tracks (see text). Trigger is given relative to BDT and is determined using the data-driven TISTOS method (see text), while all others are determined using simulation. Total is the product of all others with correlations taken into account. A factor of 10^{-1} is shown in the Total Λ_c^+ efficiency to account for the retention of only 10% of the candidates in data (see Section 4.3.1). 53

12 Summary of results. See text. $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$ are defined in Equations 24, 25, and 26. N is the efficiency-corrected yield; Ξ_c^0 is implied to decay as $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$, Λ_c^+ as $\Lambda_c^+ \to pK^-\pi^+$, and Ξ_c^+ as $\Xi_c^+ \to pK^-\pi^+$. Statistical errors only. No integrated luminosity or secondary decay corrections have been applied. 68

13 Relevant properties of Ξ_b baryon and π meson states. M is the mass. $\mathrm{CG}^i_{\Xi_b^{**}}$ indicates the squared Clebsch-Gordon coefficient for $\Xi_b^{**} \to \Xi_b^i\pi$. All values come from the PDG [1]. $M(\Xi_b'^0)$ has not been measured. We list $\mathrm{CG}^0_{\Xi_b'^0}$ as 1 and $\mathrm{CG}^-_{\Xi_b'^0}$ as 0 because $\Xi_b'^0 \to \Xi_b^-\pi^+$ is not observed (see text) [3]. 74

14 Calculated RPH values used in Equation 54. 74

15 Calculated values for the expressions in Equations 50, 51, 52, and 53. 74

16 Uncertainties due to charm baryon lifetimes, given in percent. 77

17 Simulation samples used for secondary-discrimination studies. $\Lambda_c^+ \to pK^-\pi^+$ or $\Xi_c^+ \to pK^-\pi^+$ decays are used in all cases. N_{gen} is the number of generated signal charm-baryons (either Λ_c^+ or Ξ_c^+ baryons). 91

18 Relative systematic uncertainties 98

19 Theory predictions for $\mathcal{B}(\varXi_c^0 \to \pi^- \varLambda_c^+)$ from Voloshin [4], Gronau and Rosner [5], Faller and Mannel [6], and Cheng *et al.* [7]. + or − indicates an assumption of positive or negative interference, respectively, between the SUUD and WS amplitudes (see Section 4.1). Appendix A.12 describes our updates to Gronau and Rosner's predictions. 101

20 BDT input parameters, ranked according to signal, background separation. See Table 9. The input distributions can be found in Appendix A.4. 113

21 Fit parameters for the $\varXi_c^0$ fit in the TOS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n is the n parameter of the Crystal Ball function, and c_1 is the first-order coefficient of the Chebychev polynomial. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty. 141

22 Fit parameters for the $\varXi_c^0$ fit in the TIS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n is the n parameter of the Crystal Ball function, and c_1 is the first-order coefficient of the Chebychev polynomial. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty. 141

23 Fit parameters for the $\varXi_c^0$ fit in the TISTOS dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n is the n parameter of the Crystal Ball function, and c_1 is the first-order coefficient of the Chebychev polynomial. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty. 142

24 Fit parameters for the Λ_c^+ fit in the TOS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty. 142

25 Fit parameters for the Λ_c^+ fit in the TIS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty. 142

26 Fit parameters for the Λ_c^+ fit in the TISTOS dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty. 143

27 Fit parameters for the Ξ_c^+ fit in the TOS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty. 143

28 Fit parameters for the Ξ_c^+ fit in the TIS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 143

29 Fit parameters for the Ξ_c^+ fit in the TISTOS dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 144

30 Fit parameters for the Ξ_c^0 fit in the truth-matched simulation dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, and n is the n parameter of the Crystal Ball function. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 144

31 Fit parameters for the Ξ_c^0 fit in the truth-matched-plus-ghosts simulation dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, and n is the n parameter of the Crystal Ball function. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 144

32 Fit parameters for the Λ_c^+ fit in the truth-matched simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 145

33 Fit parameters for the Λ_c^+ fit in the truth-matched-plus-ghosts simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 145

34 Fit parameters for the Ξ_c^+ fit in the truth-matched simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 145

35 Fit parameters for the Ξ_c^+ fit in the truth-matched-plus-ghosts simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty. 146

36 Old world-averages of charm-baryon lifetime measurements [8]. 152

List of Figures

1 The Standard Model of particle physics [9]. The brown loops indicate couplings between (red) bosons and (purple and green) fermions. Masses listed are from Reference [10]. Charges are given as fractions of e, where the electron charge is $-e$. Antiparticles are not pictured; for each particle shown, there exists an antiparticle with the same mass but opposite charge; some particles, such as photons, are their own antiparticles. The t and b quarks are sometimes referred to as the “truth” and “beauty” quarks. 3

2 A Feynman diagram depicting $\Lambda_c^+ \to pK^-\pi^+$ decay (left) with and (right) without an example gluon interaction shown. 8

3 The baryon octet in the SM [11]. s indicates strangeness (the number of s quarks multiplied by -1); q indicates charge in multiples of e, where the electron charge is $-e$. Particles along the same downward-sloping line have the same charge; those along the same horizontal line have the same strangeness. 9

4 Dalitz plot for $\Lambda_c^+ \to pK^-\pi^+$ decay. See Figure 19 in Section 4.3.4. 14

5 The CERN accelerator complex as of 2019 [12]. 16

6 The LHCb detector, shown schematically in profile [13]. 17

7 Drawing of the LHCb magnet [13]. 19

8 Perspective illustration of the VELO sensors [13]. 21

9 Cross-sectional drawing of the VELO [13]. 22

10 Illustration of the VELO vacuum system [13]. 23

11 A schematic of the third TT layer [13]. 24

12 A schematic of one IT layer [13]. 24

13 Layout of the LHCb tracking system with a cross-sectional cut-out in the $-x, +y$ quadrant of the three downstream tracking stations [13]. The OT is shown in teal while the TT and IT are shown in purple. See Figure 6. . . . 25

14 Separation of muons, pions, kaons, and protons as a function of momentum in data collected by the RICH1 detector [14]. 27

15 Display of typical simulated event in RICH1 [13]. 28

16 Layout of the muon stations [13]. 29

17 The LHCb data flow in Run II [15]. 32

18 Decay diagrams for $\Xi_c^0 \to \pi^-\Lambda_c^+$ transitions via (a) the SUUD amplitude and (b) the WS amplitude. 33

19 Dalitz plots for signal (left) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ and (right) $\mathit{\Xi}_c^+ \to pK^-\pi^+$ for (a & b) data, (c & d) simulation (without Dalitz weights applied), and (e & f) data divided by simulation (weight histograms). 44

20 Plots of mass distributions of candidates used for BDT training for (a & b) $\mathit{\Xi}_c^0 \to \pi^- (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, (c & d) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and (e & f) $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays for (left) signal simulation, for which we apply truth-matching and make no mass selections, and (right) background data, defined by the hatched red regions. The (b) $\mathit{\Xi}_c^0$ data sample shown is WS (a $\mathit{\Lambda}_c^+$ baryon paired with a π^+ meson), while all other samples shown use daughters with the correct charges. Note that the data distributions shown here represent the entire datasets; we use only a small fraction to train the BDTs (see Table 10). . . 49

21 BDT performance, as demonstrated by (left) its response and (right) its receiver operating characteristic (ROC) curve for (a & b) $\mathit{\Xi}_c^0 \to \pi^- (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, (c & d) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and (e & f) $\mathit{\Xi}_c^+ \to pK^-\pi^+$. 50

22 BDT selection figure of merit for (a) $\mathit{\Xi}_c^0 \to \pi^- (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, (b) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, (c) and $\mathit{\Xi}_c^+ \to pK^-\pi^+$ samples. See Section 4.3.7. The black histograms represent the fraction of candidates in data. The green histograms represent the fraction of signal candidates in simulation. The blue curves represent the normalized values of $\frac{\epsilon_{\text{signal}}}{\sqrt{N}}$ when the BDT response is $> x$. We select the red highlighted regions. 51

23 Efficiency as a function of p_{T} and η for (a) $\mathit{\Xi}_c^0 \to \pi^- (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, (b) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and (c) $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays. 55

24 Fit to $\mathrm{DTF}\left(m\left(\mathit{\Xi}_c^0\right)\right) - \mathrm{DTF}\left(m\left(\mathit{\Lambda}_c^+\right)\right) + m_{\mathrm{PDG}}\left(\mathit{\Lambda}_c^+\right)$ in the $\mathit{\Xi}_c^0$ data dataset using a Crystal Ball signal shape and a 1st-order Chebychev polynomial background shape; see text. 57

25 Fit to DTF $(m\,(pK^-\pi^+))$ in the $\mathit{\Lambda}_c^+$ data dataset using a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text. 58

26 Fit to DTF $(m\,(pK^-\pi^+))$ in the $\mathit{\Xi}_c^+$ data dataset using a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text. 59

27 DTF $(m\,(\mathit{\Xi}_c^0))$ $-$DTF $(m\,(\mathit{\Lambda}_c^+))$ $+m_{\mathrm{PDG}}\,(\mathit{\Lambda}_c^+)$ distribution from the $\mathit{\Xi}_c^0$ data (a) TIS and (b) TISTOS datasets fit to a Crystal Ball signal shape and a 1st-order Chebychev polynomial background shape; see text. 60

28 DTF mass distribution of $pK^-\pi^+$ from the $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ data (a) TIS and (b) TISTOS datasets fit to a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text. . 61

29 DTF mass distribution of $pK^-\pi^+$ from the $\mathit{\Xi}_c^+ \to pK^-\pi^+$ data (a) TIS and (b) TISTOS datasets fit to a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text. 62

30 DTF $(m\,(\mathit{\Xi}_c^0))$ $-$DTF $(m\,(\mathit{\Lambda}_c^+))$ $+m_{\mathrm{PDG}}\,(\mathit{\Lambda}_c^+)$ distribution from the $\mathit{\Xi}_c^0$ simulation (a) truth-matched and (b) truth-matched-plus-ghosts datasets fit to a Crystal Ball signal shape and a flat background shape; see text. 63

31 DTF mass distribution of $pK^-\pi^+$ from the $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ simulation (a) truth-matched and (b) truth-matched-plus-ghosts datasets fit to a double-Gaussian signal shape (both Gaussians are constrained to have the same mean) and a flat background shape; see text. 64

32 DTF mass distribution of $pK^-\pi^+$ from the $\mathit{\Xi}_c^+ \to pK^-\pi^+$ simulation (a) truth-matched and (b) truth-matched-plus-ghosts datasets fit to a double-Gaussian signal shape (both Gaussians are constrained to have the same mean) and a flat background shape; see text. 65

33 Efficiency-corrected yields as functions of p_{T} and η for (a) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$, (b) $\Lambda_c^+ \to pK^-\pi^+$, and (c) $\Xi_c^+ \to pK^-\pi^+$ decays. Only the statistical errors (from the fits and efficiencies) are shown here. The Λ_c^+ yields have been corrected to account for the 10% prescale. 66

34 Intrinsic results as functions of p_{T} and η for (a) $\mathcal{R}_1$, (b) $\mathcal{R}_2$, and (c) $\mathcal{R}_3$, defined in Equations 24, 25, and 26. Only statistical uncertainties (from both data and simulation) are shown. 69

35 $\mathrm{DTF}\left(m\left(\Xi_c^0\right)\right) - \mathrm{DTF}\left(m\left(\Lambda_c^+\right)\right) + m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ distribution from the Ξ_c^0 TOS data dataset fit using alternative models: The α and n parameters of the Crystal Ball are fixed to, respectively, (a) -5 and 100, (b) -25 and 5, (c) -20 and 22, and (d) -10 and 22; (e) an exponential background shape; (f) a Gaussian signal shape; (g) a double-Gaussian signal shape, where the Gaussians are constrained to have the same mean; and (h) a Gaussian signal shape with an exponential background shape. 80

36 $\mathrm{DTF}\left(m\left(\Xi_c^0\right)\right) - \mathrm{DTF}\left(m\left(\Lambda_c^+\right)\right) + m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ distributions from the Ξ_c^0 (a) TIS and (b) TISTOS data datasets fit using a model that is the same as the nominal, but with a double-Gaussian signal shape, where the Gaussians are constrained to have the same mean. 81

37 DTF mass distribution of $pK^-\pi^+$ from the Λ_c^+ TOS data dataset fit using alternative models: (a) the left and right n parameters of the double-sided Crystal Ball are fixed to 15 and 120, respectively; (b) the left and right n parameters of the double-sided Crystal Ball are fixed to 5.96 and 155, respectively, and the fraction of the signal yield assigned to the double-sided Crystal Ball is allowed to float; (c) the width, α, and left and right n parameters of the double-sided Crystal Ball are fixed to 7.34, 2, 5.96, and 155, respectively; (d) the width and left and right n parameters of the double-sided Crystal Ball are fixed to 7.34, 5.96, and 155, respectively, and the α parameter is allowed to float, though it still constrained to be the same for both left and right; (e) an exponential background shape; and (f) a linear background shape. 82

38 DTF mass distribution of $pK^-\pi^+$ from the Λ_c^+ (a) TIS and (b) TISTOS data datasets fit using a model that is the same as the nominal, but the width, α, and left and right n parameters of the double-sided Crystal Ball are fixed to 7.34, 2, 5.96, and 155, respectively. 83

39 DTF mass distribution of $pK^-\pi^+$ from the Ξ_c^+ TOS data dataset fit using alternative models: The left and right n parameters of the double-sided Crystal Ball are fixed to, respectively, (a) 15 and 120; (b) 5.5 and 155, and the fraction of the signal yield assigned to the double-sided Crystal Ball is allowed to float; (c) 5.5 and 155 and the width and α to 8.09 and 5, respectively; and (d) 5.5 and 155 and the width to 8.09, and the α parameter is allowed to float, though it is still constrained to be the same for both left and right. (e) An exponential background shape. (f) A double-Gaussian signal shape, where the Gaussians are constrained to have the same mean. (g) A double-Gaussian signal shape, where the Gaussians are constrained to have the same mean, with an exponential background shape. (h) A linear background shape. . . 85

40 DTF mass distribution of $pK^-\pi^+$ from the $\mathit{\Xi}_c^+$ (a) TIS and (b) TISTOS data datasets fit using a model that is the same as the nominal, but with a double-Gaussian signal shape, where the Gaussians are constrained to have the same mean. 86

41 $\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(H_c)\cdot\vec{p}(H_c)}{x(H_c)\cdot p(H_c)}\right)\right)\right)$ distributions for signal (left) $\mathit{\Xi}_c^0 \to \pi^- \left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)$, (center) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and (right) $\mathit{\Xi}_c^+ \to pK^-\pi^+$ in (a, b, & c) data, (d, e, & f) simulation, and (g, h, & i) data divided by simulation (weight histograms). 89

42 Comparison of $\mathcal{B}_1$, $\mathcal{B}_2$, and $\langle\mathcal{B}_{1,2}\rangle$ with Faller and Mannel's upper limit (FM) [6], the updated predictions of Gronau and Rosner (GR) [5], Voloshin's lower limit (V) [4], and Cheng *et al.*'s prediction (C) [7]. See Table 19. + or − indicates an assumption of positive or negative interference, respectively, between the SUUD and WS amplitudes (see Section 4.1). 101

43 Mass distributions for (a) $\mathit{\Xi}_c^0 \to \pi^- \left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)$, (b) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and (c) $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays in simulation. The purple distributions have truth-matching applied. The red distributions fail the truth-matching criteria. None of these plots has any of the weighting described in Section 4.3.4 applied. . 106

44 Mass distributions from $\mathit{\Xi}_c^0 \to \pi^- \left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)$ decays in simulation for candidates (a) failing the truth-matching criteria described in Section 4.3.4 that (purple) are and (red) are not matched to a generated $\mathit{\Xi}_c^0$ and (b) falling into various background categories. None of these plots has any of the weighting described in Section 4.3.4 applied. 106

45 Plots from the $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ simulation dataset: (a) The background category of other candidates appearing in an event with a candidate with background category 30 and (b) the mass distribution of background category 30 candidates which are not identified with a background category 0 or 10 candidate in the same event. None of these plots has any of the weighting described in Section 4.3.4 applied. 107

46 Mass distributions for (a) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$, (b) $\Lambda_c^+ \to pK^-\pi^+$, and (c) $\Xi_c^+ \to pK^-\pi^+$ decays in simulation, background category 60 (ghosts) only. None of these plots has any of the weighting described in Section 4.3.4 applied. 107

47 BDT performance in each fold, as demonstrated by (left) its response and (right) its receiver operating characteristic (ROC) curve for (a & b) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$, (c & d) $\Lambda_c^+ \to pK^-\pi^+$, and (e & f) $\Xi_c^+ \to pK^-\pi^+$. . . . 109

48 Overtraining check for the $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ BDT in each of its 10 folds. 110

49 Overtraining check for the $\Lambda_c^+ \to pK^-\pi^+$ BDT in each of its 10 folds. 111

50 Overtraining check for the $\Xi_c^+ \to pK^-\pi^+$ BDT in each of its 10 folds. 112

51 BDT input variable distributions (after the application of offline and trigger selections) for the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ BDTs. Solid blue shows the distributions in the signal sample; hatched red shows that in the background sample. 115

52 Input distributions (after the application of the offline and trigger selections) for the (left) Ξ_c^0, (center) Λ_c^+, and (right) Ξ_c^+ BDTs. The (black) data distributions are the result of a side-band subtraction using the fits shown in Figure 54 (c & d). The (green) simulation distributions are truth matched. 116

53 The output of a probablistic neural-network determining how likely a given final-state daughter has been assigned the correct mass after Turbo, offline, and trigger selections have been applied for (left) Ξ_c^0, (middle) Λ_c^+, and (right) Ξ_c^+ data and simulation datasets. The (black) data distribution is the result of a side-band subtraction using the fits shown in Figure 54, where the signal region corresponds to 99.7300% of the total integral symetric about the mean and the background region is taken from 99.9937% to 99.9999%. The (green) raw simulation distribution shows the uncorrected PID distribution, while the (blue) corrected simulation distribution shows the PID distribution produced by PIDGen. The simulation distributions have been truth matched. 117

54 Binned fits to the relevant invariant masses in the (top) Ξ_c^0, (middle) Λ_c^+, and (bottom) Ξ_c^+ data datsets, using the shapes described in Section 4.5.1, after the BDT input selections have been applied. The blue (red) regions are considered signal (background) for sideband subtraction; the Ξ_c^0 signal region is taken from 2467.5 MeV to 2475.0 MeV, and its background region is taken from 2465.5 MeV to 2466.3 MeV and from 2476.2 MeV to 2499.5 MeV; the Λ_c^+ signal region is taken from 2271.5 MeV to 2303.0 MeV, and its background region is taken from 2238.1 MeV to 2256.7 MeV and from 2317.8 MeV to 2336.5 MeV; the Ξ_c^+ signal region is taken from 2450.9 MeV to 2485.0 MeV, and its background region is taken from 2418.0 MeV to 2434.3 MeV and from 2501.6 MeV to 2517.9 MeV; 118

55 BDT input parameter distributions after the application of the BDTs for the (left) Ξ_c^0, (center) Λ_c^+, and (right) Ξ_c^+ BDTs. The (black) data distributions are the result of the application of the *sPlot* method [16] using the fits shown in Figures 24, 25, and 26. The (green) simulation distributions are truth matched. 119

56 BDT response distributions after the application of the BDTs for the (left) Ξ_c^0, (center) Λ_c^+, and (right) Ξ_c^+ BDTs. The (black) data distributions are the result of the application of the *sPlot* method [16] using the fits shown in Figures 24, 25, and 26. The (green) simulation distributions are truth matched. These plots do not appear in Reference [17]. 120

57 BDT input distributions (after the application of the Turbo, offline, and trigger selections) for the (black) side-band and (green) WS Ξ_c^0 data distributions. The sidebands are taken from 2465.46 MeV to 2466.46 MeV and from 2475.46 to 2482.46 MeV in the $\mathrm{DTF}\,(m\,(\Xi_c^0)) - \mathrm{DTF}\,(m\,(\Lambda_c^+)) + m_{\mathrm{PDG}}\,(\Lambda_c^+)$ distribution, and the WS are restricted to 2465.87 MeV to 2475.87 MeV. 121

58 BDT input distributions (after the application of the Turbo, offline, and trigger selections) for the (purple) low-mass and (orange) high-mass side-band (left) Λ_c^+ and (right) Ξ_c^+ data distributions. The side-bands are described in Section 4.3.5. 122

59 Invariant mass of the Ξ_c^0 baryon's daughters in data with the mass hypothesis of its daughter π^- meson changed to a K^- meson. (DTF is not applied.) . 123

60 Invariant mass of the Λ_c^+ or Ξ_c^+ baryon's daughters in the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ data datasets with the mass hypothesis of the proton and the π^+ meson changed to a K^+ meson and DTF re-applied. 124

61 Invariant mass of the Λ_c^+ or Ξ_c^+ baryon's daughters in the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ data datasets with the mass hypothesis of the proton changed to a K^+ meson and DTF re-applied. 124

62 Invariant mass of the Λ_c^+ or Ξ_c^+ baryon's daughters in the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ data datasets with the mass hypothesis of the proton changed to a π^+ meson and DTF re-applied. 125

63 Nominal mass fits in bins of p_{T} and η in the $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ TOS dataset. The range is given in the x-axis title of the pull distribution of each plot. 126

64 Nominal mass fits in bins of p_{T} and η in the $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ TIS dataset. The range is given in the x-axis title of the pull distribution of each plot. 127

65 Nominal mass fits in bins of p_{T} and η in the $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ TISTOS dataset. The range is given in the x-axis title of the pull distribution of each plot. 128

66 Nominal mass fits in bins of p_{T} and η in the $\Lambda_c^+ \to pK^-\pi^+$ TOS data dataset. The range is given in the x-axis title of the pull distribution of each plot. . . 129

67 Nominal mass fits in bins of p_{T} and η in the $\Lambda_c^+ \to pK^-\pi^+$ TIS data dataset. The range is given in the x-axis title of the pull distribution of each plot. . . 130

68 Nominal mass fits in bins of p_{T} and η in the $\Lambda_c^+ \to pK^-\pi^+$ TISTOS data dataset. The range is given in the x-axis title of the pull distribution of each plot. 131

69 Nominal mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ TOS data dataset. The range is given in the x-axis title of the pull distribution of each plot. . . 132

70 Nominal mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ TIS data dataset. The range is given in the x-axis title of the pull distribution of each plot. . . 133

71 Nominal mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ TISTOS data dataset. The range is given in the x-axis title of the pull distribution of each plot. 134

72 Mass fits in bins of p_{T} and η in the $\Xi_c^0 \to \pi^- \Lambda_c^+$ MC truth-matched dataset. The range is given in the x-axis title of the pull distribution of each plot. . . 135

73 Mass fits in bins of p_{T} and η in the $\Xi_c^0 \to \pi^- \Lambda_c^+$ MC truth-matched-plus-ghosts dataset. The range is given in the x-axis title of the pull distribution of each plot. 136

74 Mass fits in bins of p_{T} and η in the $\Lambda_c^+ \to pK^-\pi^+$ MC truth-matched dataset. The range is given in the x-axis title of the pull distribution of each plot. . . 137

75 Mass fits in bins of p_{T} and η in the $\Lambda_c^+ \to pK^-\pi^+$ MC truth-matched-plus-ghosts dataset. The range is given in the x-axis title of the pull distribution of each plot. 138

76 Mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ MC truth-matched dataset. The range is given in the x-axis title of the pull distribution of each plot. . . 139

77 Mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ MC truth-matched-plus-ghosts dataset. The range is given in the x-axis title of the pull distribution of each plot. 140

78 Momentum distributions of the Λ_c^+ or Ξ_c^+ baryon, as appropriate, and its final-state daughters for the (purple) Ξ_c^0, (green) Λ_c^+, and (red) Ξ_c^+ (left) data and (right) simulation datasets after the application of Turbo, offline, and trigger selections. The (left) data distributions are the result of a side-band subtraction using the fits shown in Figure 54. The (right) simulation distributions are truth matched. 148

79 $\mathrm{DTF}\left(m\left(\Xi_c^0\right)\right) - \mathrm{DTF}\left(m\left(\Lambda_c^+\right)\right) + m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ in the Ξ_c^0 data dataset after the application of the Turbo, offline, and trigger selections (purple) without any modifications and (red) with the Λ_c^+ mass required to be outside ±40 MeV of its value in the PDG [1]. 149

80 The squared invariant-mass of (top) the proton and π^+ meson and (bottom) the K^- meson and π^+ meson for (a & d) Ξ_c^0, (b & e) Λ_c^+, and (c & f) Ξ_c^+ (black) data and (green) simulation datasets after the application of all selections and weights. The (black) data distributions are the result of the application of the *sPlot* method [16] using the fits shown in Figures 24, 25, and 26. The (green) simulation distributions are truth matched. 150

81 Ratios of trigger-efficiency ratios, derived using the TISTOS method, in simulation to those in data for (a) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ to $\Lambda_c^+ \to pK^-\pi^+$, (b) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ to $\Xi_c^+ \to pK^-\pi^+$, and (c) $\Xi_c^+ \to pK^-\pi^+$ to $\Lambda_c^+ \to pK^-\pi^+$. Only statistical errors shown. 151

1 Introduction

A longstanding goal of the natural sciences has been to explain the fundamental building blocks that make up the world around us. For example, Thales of Miletus, one of the pre-Socratic Greek philosophers, postulated that the world was made up of water [18], while Heraclitus, another pre-Socratic, postulated the converse, that fire was the fundamental principle [19]. These are among the earliest in a long line of philosophers who attempted to explain the unobservable elements of the world around them by analogy to observable ones. Our descriptions and means of observation have become considerably more sophisticated over the millennia, but arguably the fundamental sciences today share the same goal as the early natural philosophers: describing the world in terms humans can understand.

These days, we believe the universe to be made up not of fire nor of water but of something much stranger: particles. These bits of matter appear to be infinitesimally small and to have indeterminate positions and momenta. This counterintuitive nature notwithstanding, modeling the universe as the interactions of fundamental particles has been extraordinarily successful both descriptively and predictively. This model is so successful and so widely accepted, in fact, that it has become known as the Standard Model of particle physics (SM). Despite its many successes, the SM also has a number of glaring shortcomings, lacking explanations for, *e.g.*, gravity, the prevalence of matter over antimatter, and dark matter, and retaining ambiguities that can only be resolved by experimentation.

The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) is the largest machine dedicated to the study of particle physics built to date. The LHCb detector, one of four experiments built at the LHC, was designed for the study of b quarks, but it is also a general-purpose detector capable of studying other particles, including c quarks. The measurement presented here takes advantage of the abundance of c quarks produced in the LHCb detector to perform novel measurements of the behavior of SM particles and provide experimental feedback on the accuracy of theoretical models.

This dissertation is organized into sections: Section 2 introduces the SM and describes its relevant aspects; Section 3 introduces the LHC and the LHCb detector; Section 4 describes measurements of the branching fractions of the Ξ_c^0 baryon in its decay to π^- and Λ_c^+ hadrons and of the Ξ_c^+ baryon in its decay to p, K^-, and π^+ hadrons; and Section 5 summarizes the findings of the preceding chapters. Natural units are used throughout, *i.e.*, $\hbar = c = 1$, where $\hbar$ is the reduced Planck constant and c is the speed of light, though some figures show factors of c.

2 The Standard Model of particle physics

What follows is a brief, introductory-level description of relevant concepts in the Standard Model. The advanced reader is encouraged to consult the referenced works for more information.

2.1 Overview

The Standard Model of particle physics (SM), first articulated in the second half of the twentieth century [20–23], is a description of fundamental particles and their interactions. One of the most successful scientific theories ever put forward, its predictions have been found to be very robust. As shown in Figure 1, the SM can be pictured as something like the periodic table, familiar from chemistry, organizing elementary particles into useful categories.

All SM particles are either fermions or bosons. Generally speaking, one can conceive of the elementary fermions as making up matter and the elementary bosons as mediating their interactions. The fermions are further divided into quarks (q) and leptons (ℓ); the most familiar of these are the up and down quarks, found in protons and neutrons, and electrons, which orbit protons and neutrons in atoms and whose movements constitute electric current. The bosons mediate the four fundamental forces of nature and, in the case of the Higgs, give other particles mass. The fundamental forces are the strong force, mediated by the gluon (g); the electromagnetic force, mediated by the photon (γ); the weak force, mediated by the W

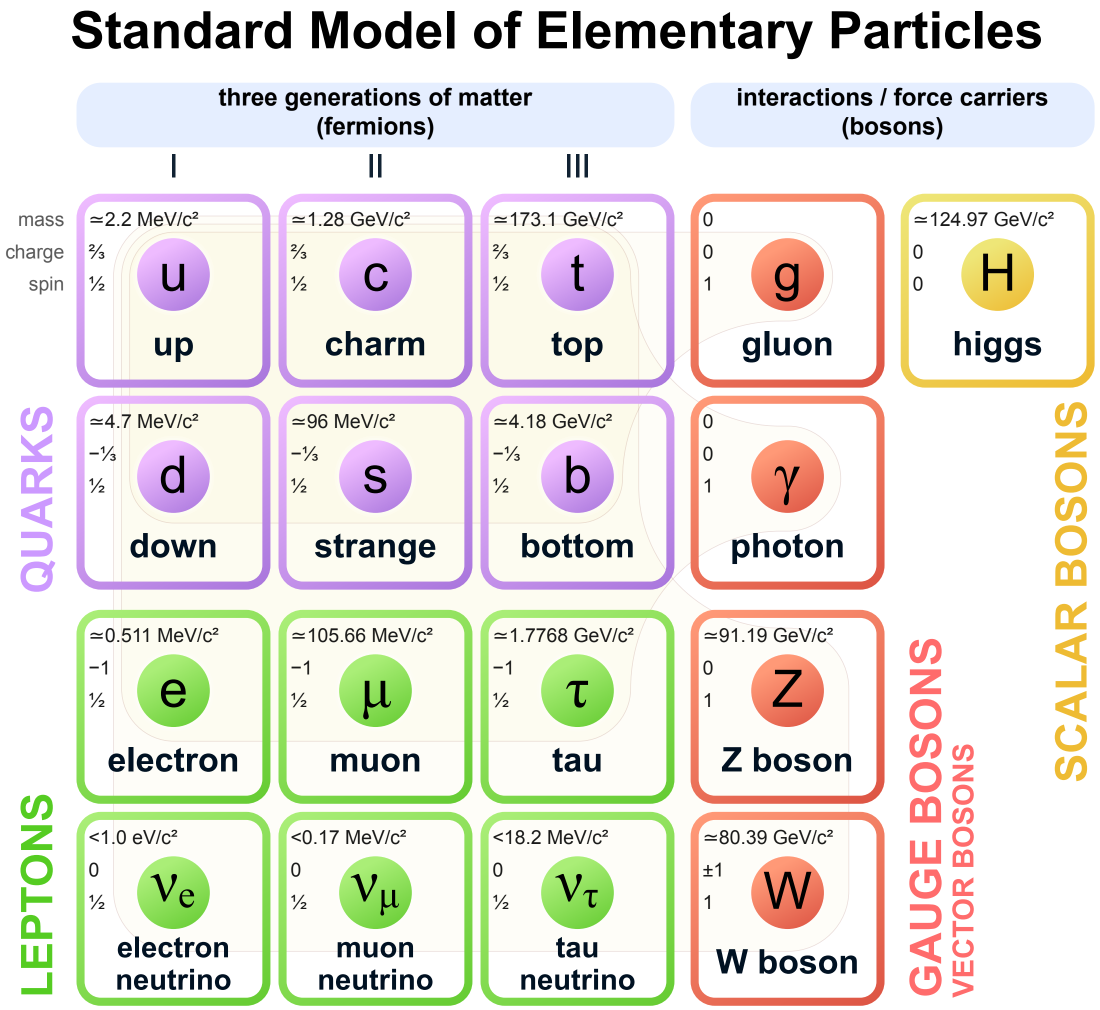

Standard Model of Elementary Particles
three generations of matter (fermions)
interactions / force carriers (bosons)
I
II
III
mass
charge
spin
≃2.2 MeV/c²
⅔
½
u
up
≃1.28 GeV/c²
⅔
½
c
charm
≃173.1 GeV/c²
⅔
½
t
top
0
0
1
g
gluon
≃124.97 GeV/c²
0
0
H
higgs
≃4.7 MeV/c²
−⅓
½
d
down
≃96 MeV/c²
−⅓
½
s
strange
≃4.18 GeV/c²
−⅓
½
b
bottom
0
0
1
γ
photon
≃0.511 MeV/c²
−1
½
e
electron
≃105.66 MeV/c²
−1
½
μ
muon
≃1.7768 GeV/c²
−1
½
τ
tau
≃91.19 GeV/c²
0
1
Z
Z boson
<1.0 eV/c²
0
½
νe
electron neutrino
<0.17 MeV/c²
0
½
νμ
muon neutrino
<18.2 MeV/c²
0
½
ντ
tau neutrino
≃80.39 GeV/c²
±1
1
W
W boson
QUARKS
LEPTONS
GAUGE BOSONS
VECTOR BOSONS
SCALAR BOSONS

Figure 1: The Standard Model of particle physics [9]. The brown loops indicate couplings between (red) bosons and (purple and green) fermions. Masses listed are from Reference [10]. Charges are given as fractions of e, where the electron charge is $-e$. Antiparticles are not pictured; for each particle shown, there exists an antiparticle with the same mass but opposite charge; some particles, such as photons, are their own antiparticles. The t and b quarks are sometimes referred to as the "truth" and "beauty" quarks.

and Z bosons; and gravity, which is not included in the SM.

An important limit on the behavior of quarks is a phenomenon called "quark confinement" [24]. This phenomenon prevents quarks from existing in isolation; they are always found near (approximately within the diameter of a proton) another quark (q) or antiquark ($\overline{q}$). These composite configurations of quarks, called "hadrons", are held together by the strong force. Similar to electric charge in electromagnetism, quarks carry a property called "color charge" that governs how they interact via the strong force. Color charge comes in six types, red, green, and blue, plus their anti-matter equivalents, and hadrons must be color neutral, *e.g.*, contain one of each color, contain a color and its opposite, *etc.*[1] This means that hadrons containing three quarks must have one of each color, one red, one green, and one blue;[2] hadrons containing three quarks are called "baryons". Antiquarks carry anti-color charge, anti-red, anti-green, or anti-blue, which means one quark and one antiquark can also combine to form a color-neutral hadron, *e.g.*, a red d-quark with an anti-red $\overline{u}$-antiquark; hadrons containing one quark and one antiquark are called "mesons". Quarks can also combine in more exotic ways; for example, two red quarks, a blue quark, a green quark, and an anti-red antiquark may combine to form a color-neutral hadron called a "pentaquark". The theory describing the behavior of the strong force, gluons, and color charge is called "Quantum ChromoDynamics" (QCD).

Most hadrons do not last long; in fact, the only hadron believed to be stable is the proton,[3] a baryon comprising two up-quarks and one down-quark (uud). All other hadrons last for a finite amount of time (called a "lifetime") before decaying into other particles,[4] generally through the effects of one or more of the fundamental forces; usually, the heavier the particle, the shorter its lifetime. While this is called "decay", the particles are generally not breaking apart into their constituents but transforming into other particles with lower masses. There are a variety of rules governing which decays are allowed, one of which is that only weak decays (mediated by the weak force) can change the flavor of quarks; "flavor" is a term of art referring to the six different types of quarks pictured in Figure 1: up (u), down (d), charm (c), strange (s), top (t), and bottom (b). Thus, when we observe a particle containing, *e.g.*, a c quark decaying into particles with no c quarks, we know that decay is mediated by the weak force.

---

[1] This nomenclature is a nod to the color wheel of visible light, where red, green, and blue combine to make white.

[2] Quarks are constantly interacting via the strong force by the exchange of gluons, which changes their color charge, *e.g.*, a blue u-quark may become a green u-quark, but the other quarks in the hadron simultaneously change colors in a way that ensures the overall hadron remains color neutral during this process.

[3] The proton and antiproton are both believed to be stable, or rather, neither has been observed to decay.

[4] The lifetime of a particle is a property that characterizes the probable length of time before decay in isolation. It does not mean that every neutron, for example, lasts $879.4 \pm 0.6\,\mathrm{s}$ [25].

Hadrons generally have a variety of ways they are allowed to decay and a variety of particles they can decay into. The decaying hadron is referred to as the "mother" particle and particles produced in the decay are called "daughter" particles. A set of specific daughter particles is called a "decay mode", and the probability of a mother particle decaying to a certain mode is called a "branching fraction" ($\mathcal{B}$). Because they are governed by the interactions of hadronic components, measuring the sizes of various branching fractions can be a useful way to test models of hadronic behavior. Each set of interactions leading to a certain decay mode carries an "amplitude" characterizing its likelihood; the branching fraction is proportional to the square of the sum of these amplitudes.[5] One might think that having more available internal interactions always increases the branching fraction for a given mode, but this is not the case: The different possibilities can interfere with each other and *reduce* the overall branching fraction. Discerning whether the different amplitudes increase (constructive interference) or decrease (destructive interference) the branching fraction for a given mode can be difficult and often requires experimentation.

The total lifetime of a particle is governed by the decay modes available to it. On an intuitive level, it makes some sense that a particle is more likely to decay if it has more possibilities for decay available; after all, if no decay modes are available, it won't ever decay.[6] The available decay modes are thus said to contribute "partial widths" to a particle's "total width". This terminology is explained by the confluence of the Heisenberg Uncertainty Principle and Einstein's most famous equation,

$$E = m$$

---

[5] The idea that the probability of an event is proportional to the square of a sum of amplitudes will be familiar to anyone who has studied quantum-mechanical wave functions.

[6] Note that a "decay mode" is distinct from an "amplitude". A particular decay mode may have many amplitudes contributing to it, corresponding to the many processes that can lead to the same final state.

(recall that we are using natural units, where $c = \hbar = 1$). The Heisenberg Uncertainty Principle was a revolutionary idea when first proposed and now forms one of the pillars of modern physics. Its fundamental insight is that there are irreducible limits to the precision with which certain parameters can be known, and it can be written as

$$\Delta E \Delta t \geq \frac{1}{2},$$

where ΔE is the uncertainty on the particle's energy and Δt that on the time it takes to decay. Invoking Einstein's equation and realizing that Δt is just the particle's lifetime (τ), we obtain

$$\Delta m \geq \frac{1}{2\tau}.$$

Okay, so where does the idea of a "width" come in? Well, Δm characterizes a range of masses a given particle may take, and if the probability that a particle will take a certain mass is plotted on a graph, it appears as a curve with a characteristic width, $2 \times \Delta m$, often written as Γ. Thus,

$$\Gamma \propto \frac{1}{\tau},$$

so the shorter the particle's lifetime, the larger its (total) width. A given branching fraction, $\mathcal{B}_i$, can be written as

$$\mathcal{B}_i = \frac{\Gamma_i}{\Gamma},$$

where $\sum_i \Gamma_i = \Gamma$. Each Γ_i is referred to as a "partial width" and Γ the "total width", and we see that each additional decay mode increases the particle's total width and decreases its lifetime.

The same fundamental forces that mediate particle decay also mediate particle collisions. When protons collide at high energies (see Section 3.1), their constituent quarks and the gluons that bind them interact with each other primarily via the strong force.[7] These interactions transform the energy of the collision into more quarks and gluons, which recombine to form new hadrons in a process called "hadronization". These hadrons fly out from the point of collision (called a "primary vertex" [PV]) at high speed and the unstable ones then decay.

---

[7] Protons can also interact via the electromagnetic or even weak force, which can result in the production of other types of particles as well, not just quarks and gluons.

2.2 Feynman diagrams

Decays and collisions are commonly visualized using Feynman diagrams, which are helpful tools for understanding the allowed interactions between particles. The rules for making them are relatively simple, and any introductory particle physics textbook will explain them fully,[8] but here is presented a basic summary of how to read them. The first thing to note is that time flows from left to right, so that the particles shown on the left are in the initial state and those on the right are in the final state. Lines are then drawn connecting the initial state to the final state; these lines summarize the physical processes that lead from one to the other. In general, only the initial and final states (the line endings) are observable.

Figure 2 shows an example. On the left are the u, d, and c quarks that make up the Λ_c^+ baryon. On the right are the proton, K^- meson, and π^+ meson that the Λ_c^+ baryon decays to. The wavy line marked W^+ represents the emission of a W^+ boson by the c quark (weak decay); this transforms the c quark to an s quark, and the W^+ boson decays to a u quark and a $\overline{d}$ antiquark, which hadronize to form a π^+ meson. Simultaneously, the d quark emits a gluon, depicted as a curly line in the left-hand diagram in Figure 2, which decays to a u quark and a $\overline{u}$ antiquark. The two u quarks and the d quark hadronize to form a proton, while the $\overline{u}$ antiquark and s quark hadronize into a K^- meson. The arrows distinguish matter from antimatter; right-facing arrows indicate quarks, while left-facing arrows indicate antiquarks. There are frequently many Feynman diagrams available for a given set of initial and final states, corresponding to the many internal interactions that can lead to a given final state, but there are so many potential strong-force interactions available in any decay that the selection of which ones to show becomes arbitrary, and it is common not to draw them at all.

---

[8] See, *e.g.*, Chapter 2 and Appendix D of Reference [26] for a description of the rules of Feynman diagrams.

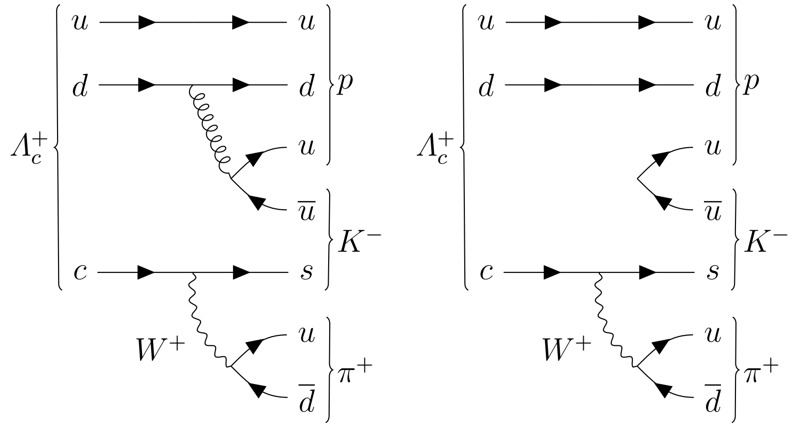


Figure 2: A Feynman diagram depicting $\Lambda_c^+ \to pK^-\pi^+$ decay (left) with and (right) without an example gluon interaction shown.

2.3 Isospin and SU(3) symmetry

Symmetries are among the most basic and useful features of physics as we practice it today, and the SM involves a number of them. One of the most striking has to do with how the strong force interacts with quarks. In 1932, Heisenberg proposed that, given the similarities in mass between the well-known proton and neutron (938.27 MeV and 939.57 MeV, respectively [25]), they be considered two states (called "isospin" states) of the same "particle", the nucleon [27]. In this way of thinking, the proton (quark content *uud*) is a nucleon with isospin "up", and the neutron (*udd*) is one with isospin "down".[9] The "symmetry" part of this comes with the realization that the strong force treats all quark flavors the same (apart from their masses—see below), leading to similarities between the properties of protons and neutrons, a phenomenon known as "isospin symmetry".[10] In the language of group theory, this is an SU(2) symmetry. Group theory has a remarkable ability to clarify aspects of hadronic behavior, but there is not space for a satisfactory explanation here.[11] For now, it is enough to note that SU(2) symmetry is a symmetry under the exchange of *two* quarks, the u and d

---

[9] The total isospin of the nucleon is $I = 1/2$. When its third component, I_3, is $+1/2$, it is a proton; when $I_3 = -1/2$, it is a neutron. $I_3 = +1/2, -1/2$ are sometimes referred to as isospin "up" and "down". This notation will be familiar to anyone who has studied quantum-mechanical spin, and it is likely to be confusing to anyone who hasn't; the details are not required to understand this section.

[10] Isospin symmetry applies to all isospin partners, *e.g.*, the Ξ_c^0 and Ξ_c^+ baryons, not just protons and neutrons.

[11] For a simple introduction to group theory in the context of particle physics, see Chapter 4 of Reference [26].

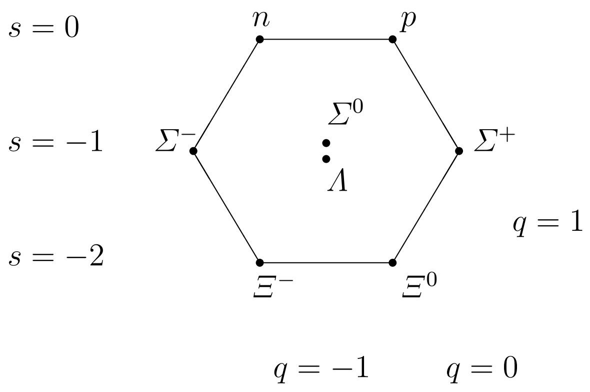


Figure 3: The baryon octet in the SM [11]. s indicates strangeness (the number of s quarks multiplied by -1); q indicates charge in multiples of e, where the electron charge is $-e$. Particles along the same downward-sloping line have the same charge; those along the same horizontal line have the same strangeness.

quarks.

Having noticed this symmetry, particle physicists are able to calculate certain properties of particle phenomenology, notably, the relative rates of certain decays, *e.g.*, $\mathcal{B}(\Xi_b^{*0} \to \Xi_b^0 \pi^0)/\mathcal{B}(\Xi_b^{*0} \to \Xi_b^- \pi^+)$. These relative rates are governed by the Clebsch-Gordan coefficients, which are useful for determining all kinds of relative phenomena that obey SU(2) symmetry, and the available phase-space, which corresponds to the energy and momentum available to the daughter particles.

In the 1950s, particle physics was facing the Eight Baryon Problem. The p (uud), n (udd), Ξ^- (dds), Ξ^0 (uds), Σ^+ (uus), Σ^0 (uds), Σ^- (dds), and Λ (uds) baryons had all been observed, and it seemed that the Ξ, Σ, and Λ baryons should belong to a grouping analogous to the nucleon (the proton and neutron group), but the symmetry wasn't clear. Gell-Mann's solution to this was the Eightfold Way [28], which elucidated an SU(3) symmetry under the exchange of *three* quarks, the u, d, and s quarks, and implied similarities among some of their properties. This theory posited that these eight baryons belonged to a "baryon octet", shown in Figure 3, with the proton and neutron belonging to the nucleon doublet; the Ξ^- and Ξ^0 baryons belonging to the Ξ doublet; the Σ^+, Σ^0, and Σ^- baryons belonging to the Σ triplet; and the Λ baryon belonging to the Λ singlet.

Of course, these days we know of six quarks, and their collective symmetry is SU(6). These symmetries, however, are not exact, because the quarks have different masses. Recall that Heisenberg initially proposed the idea of a nucleon doublet because of the similarity between the proton and neutron masses; since their masses are only *similar*, not *the same*, SU(2) symmetry doesn't quite hold. This "symmetry breaking" is worse for SU(3) because of the larger difference in the mass of the s quark vs. the u and d quarks (see Figure 1), and it is terrible in SU(6) due to the extremely large mass of the t quark, which doesn't form hadrons at all.[12] Thus one must be careful when invoking these symmetries, and large uncertainties must sometimes be assigned.

2.4 Partially conserved axial current

Another SM symmetry that can be exploited to predict hadronic behavior is chiral symmetry. The idea comes from the observation that each quark can be left-handed (with spin antiparallel to its momentum) or right-handed (with spin parallel to its momentum).[13] This orientation is called "helicity", and chirality may be thought of as its frame-independent analogue.[14] For massless particles, chirality is a permanent feature, so in the limit of massless quarks (where we pretend the quark masses are zero), left- and right-handed quarks begin to look like distinct particles. As discussed in Section 2.3, the u and d quarks (approximately) obey an

---

[12] The t quark does not form hadrons because of its extremely short lifetime, $\approx 5 \times 10^{-25}$ s, which is not enough time for it to hadronize with other quarks [25].

[13] Determining the orientation of a spin vector is traditionally done using one's hands. Curl the fingers of your right hand in the direction of rotation with your thumb sticking straight out; the direction your thumb is pointing is the direction of the spin vector. For example, if you look at an analog clock and curl your right fingers clockwise, your right thumb should point directly into the clock face. Any introductory physics textbook will describe this convention, as well as vector algebra generally, in detail.

[14] For massless particles traveling at the speed of light, helicity is identical to chirality. This is a consequence of Einstein's Special Theory of Relativity, in which massless particles travel at the speed of light and nothing may travel faster than that. The simplest visualization is perhaps to imagine driving parallel to a train; if the train is moving faster than you are, you see it traveling forward, but if you accelerate to move faster than the train, you see it slow down and then reverse, traveling backward relative to your own motion. By moving faster than the train, therefore, you have changed the direction of its momentum in your frame of reference. So it is with particles: for a massive particle, there is always a frame of reference in which its momentum, and therefore its helicity, is reversed. Massless particles travel at the speed of light, however, and there is therefore no reference frame which reverses their momenta; thus their helicity is independent of reference frame and identical to their chirality.

SU(2) symmetry (“isospin”). If they are treated as massless, they are no longer two quarks but four, two with each chirality, and obey the chiral symmetry, $\mathrm{SU_L}(2) \otimes \mathrm{SU_R}(2)$. Chiral symmetry within the SM is characterized by an axial[15] current, and this current is conserved in the limit of massless quarks.[16] Of course, quarks are not actually massless, and one with left-handed chirality may change to right-handed chirality and vice-versa, explicitly breaking chiral symmetry. Since quarks have mass, chiral symmetry is not exact, and its current is only a *Partially* Conserved Axial Current (PCAC). PCAC can be used to make various predictions for the behaviors of hadrons using current algebra, which focuses on mathematical manipulations of currents in the SM [30–33].

2.5 Heavy-quark effective theory

Calculating predictions for the properties of hadrons using QCD can be challenging if not impossible. It is therefore useful to employ a variety of tricks to make approximate predictions. One such trick is the heavy-quark expansion employed in Heavy-Quark Effective Theory (HQET) [34,35]. The fundamental insight of HQET is that, for hadrons containing one heavy (c or b) quark, the actual mass of the heavy quark becomes irrelevant to the behavior of the light (u, d, or s) quarks within the hadron, allowing for calculations of the hadron’s behavior.

The “expansion” part of the heavy-quark expansion in HQET refers to a perturbative expansion around the infinite-mass limit of the heavy quark. Put simply, the behavior of a hadron is described by a summation of terms where the largest term is independent of the mass of the heavy quark. The remaining terms are multiplied by inverse-powers of the heavy-quark mass and contribute less and less at higher and higher inverse-powers

---

[15] “Axial vectors” are invariant under parity transformations (which carry every point in a coordinate system through the origin to their diametrically opposite positions) and are sometimes called “pseudovectors”; this is in contrast with “polar vectors”, which change sign under parity. The classic example of an axial vector is a cross-product: Consider a vector $\vec{c} = \vec{a} \times \vec{b}$, where $\vec{a}$ and $\vec{b}$ are polar vectors. Under a parity transformation P, therefore, $P(\vec{c}) = -\vec{a} \times -\vec{b} = \vec{c}$.

[16] According to Noether’s Theorem, physical symmetries imply conservation laws [29], *e.g.*, the invariance of physical laws over time implies the well-known conservation of energy. Conserved quantities in Noether’s Theorem can be described using currents; these are mathematically similar to electric currents, which are also conserved. Indeed, the conservation of electric charge follows from Noether’s Theorem.

(“higher-order terms”). If the quark mass is taken to be infinite, these higher-order terms reduce to 0, leaving only the first-order term and simplifying the calculation significantly. Expressed in over-simplified math,

$$\begin{aligned}
\text{hadron behavior} &= \sum_{i=0}^{\infty} c_i \times \left(\frac{1}{m_Q}\right)^i \\
&= c_0 + c_1 \times \frac{1}{m_Q} + c_2 \times \left(\frac{1}{m_Q}\right)^2 + c_3 \times \left(\frac{1}{m_Q}\right)^3 + \cdots \\
\implies \lim_{m_Q \to \infty} \text{hadron behavior} &= c_0,
\end{aligned}$$

where the c_i are independent of the mass of the heavy quark, m_Q. Of course, heavy quarks have finite masses, so in practice some terms in the heavy-quark expansion must be calculated, but it can be easily seen that the heavier the heavy quark, the less the higher-order terms (“perturbations”) contribute to the hadron’s behavior.

The heaviest quark capable of forming hadrons is the b quark [25], and HQET works well in predicting various properties of hadrons containing a single b quark. More challenging is the other heavy quark capable of forming hadrons, the c quark. Since it is so much lighter than the b quark (see Figure 1), its behavior is more sensitive to contributions from higher-order terms, which must therefore be calculated with greater precision to ensure accurate predictions using HQET. That said, the beautiful thing about HQET is that the higher-order terms usually contribute less than the lower-order ones, so to a decent approximation, one expects hadrons that differ only by the exchange of a b and a c quark to exhibit similar behaviors. This symmetry under the exchange of a heavy quark is called “heavy-quark symmetry”.

2.6 Dalitz plots

We often cannot observe short-lived mother particles directly and instead examine the properties of their daughters to glean information. A common technique for extracting information about particles that decay to three daughters was developed by R. H. Dalitz in the 1950s [36]. The basic idea exploits the conservation of energy, the conservation of momentum, and Einstein's famous equation,

$$E^2 = m^2 + p^2,$$

where E is the particle's energy, m its mass, and $\vec{p}$ its momentum, and which reduces to its popular form, $E = m$, for particles with zero momentum. An accessible derivation of the technique can be found in Chapter 4.3 of Reference [37], but the takeaway is that, thanks to these relations, plotting the squared mass of any two pairs of daughter particles against one another allows us to see certain otherwise-hidden dynamics of the decay.

Consider Figure 4, which shows a Dalitz plot for $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ decay. The horizontal axis shows the squared mass of the proton and pion pair,

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

and the vertical axis shows that for the kaon and pion pair, $m^2(K^-, \pi^+)$. The non-zero region of the plot takes the shape of a rounded triangle, whose area is determined by the available phase-space of the decay, and a number of linear enhancements can be seen within it. These linear enhancements correspond to the presence of intermediate resonances in $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ decay; for example, the horizontal band at $\approx 0.8\,\mathrm{GeV}^2$ is due to the $K^*(892)^0$ resonance, where the decay is $\mathit{\Lambda}_c^+ \to p(K^*(892)^0 \to K^-\pi^+)$ and $m^2(K^*(892)^0) \approx 0.8\,\mathrm{GeV}^2$ [25], and the diagonal bands correspond to pK^- resonances. Dalitz plots thus encode information about the intermediate resonant structure of three-body decays, which in turn influence daughter-particle dynamics, *e.g.*, the opening angles between the daughters.[17]

---

[17] An "opening angle" is the angle between the momentum vectors of two created particles.

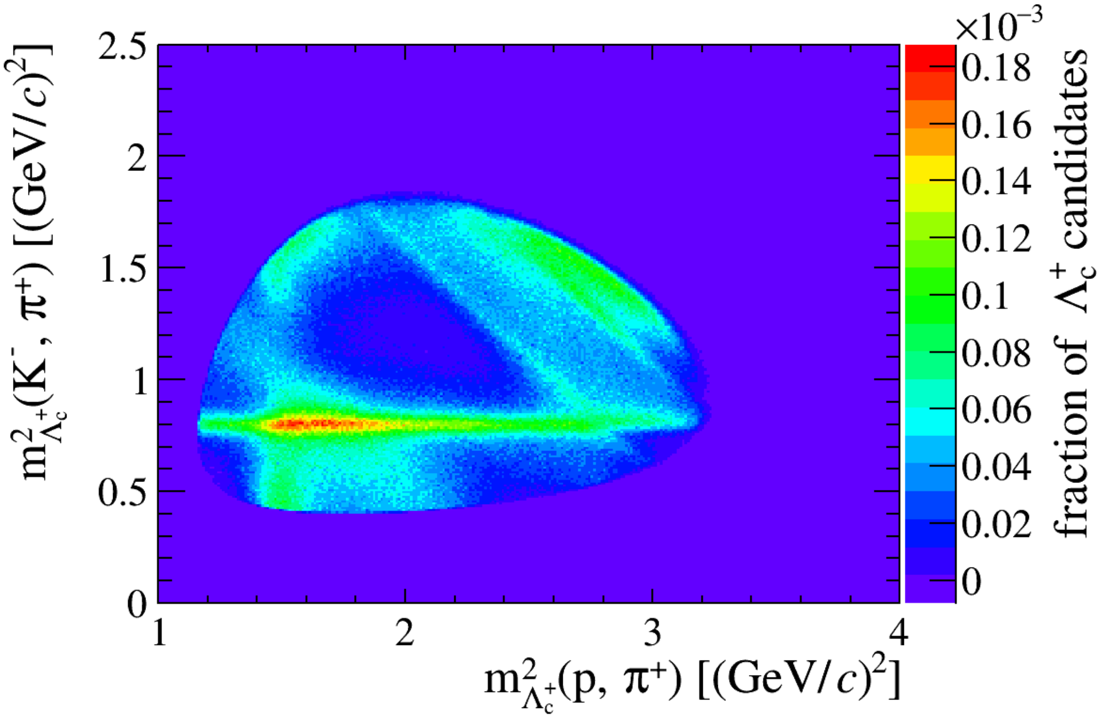


Figure 4: Dalitz plot for $\Lambda_c^+ \to pK^-\pi^+$ decay. See Figure 19 in Section 4.3.4.

3 The LHCb experiment

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest machine dedicated to the study of particle physics built to date. It is located at the European Organization for Nuclear Research (CERN) on the border between France and Switzerland and has a circumference of 27 km. The purpose of the LHC is to accelerate hadrons to high energies and force them to collide.

The LHC is the last stage of a complex of accelerators at CERN, as shown in Figure 5 [38]. The acceleration process begins with Linear Accelerator 4 (LINAC4), which accelerates negative hydrogen ions to 160 MeV. These ions are then fed into the Proton Synchrotron Booster (PSB), during which the electrons are stripped from the ions, leaving only protons. The PSB brings the protons up to 2 GeV before injecting them into the Proton Synchrotron (PS), which accelerates them to 26 GeV before passing them to the Super Proton Synchrotron (SPS), which accelerates them to 450 GeV. Now the protons can be injected into the LHC itself.

The LHC circulates two beams of protons, one clockwise and one anti-clockwise, held in separate rings where they are each accelerated to 6.5 TeV. They are then forced to collide at four different interaction points around the ring, the locations of the ALICE, ATLAS, CMS, and LHCb detectors, with a total center-of-mass collision energy of 13 TeV. The protons are separated into discrete bunches to help control the collisions; a bunch in one beam is brought into an interaction point at the same time as a bunch in the other beam, allowing the protons in the bunches to collide. Bunch crossings occur once every 25 ns. Bunch crossings where protons collide are referred to as “events”.

The preceding description is for the accelerator complex at the time of publication, not as it existed during data-taking for the work presented in Section 4, during which, for example, Linear Accelerator 2 rather than LINAC4 served as the starting point of the acceleration process. The overall picture is roughly the same, however, including the final collision energy

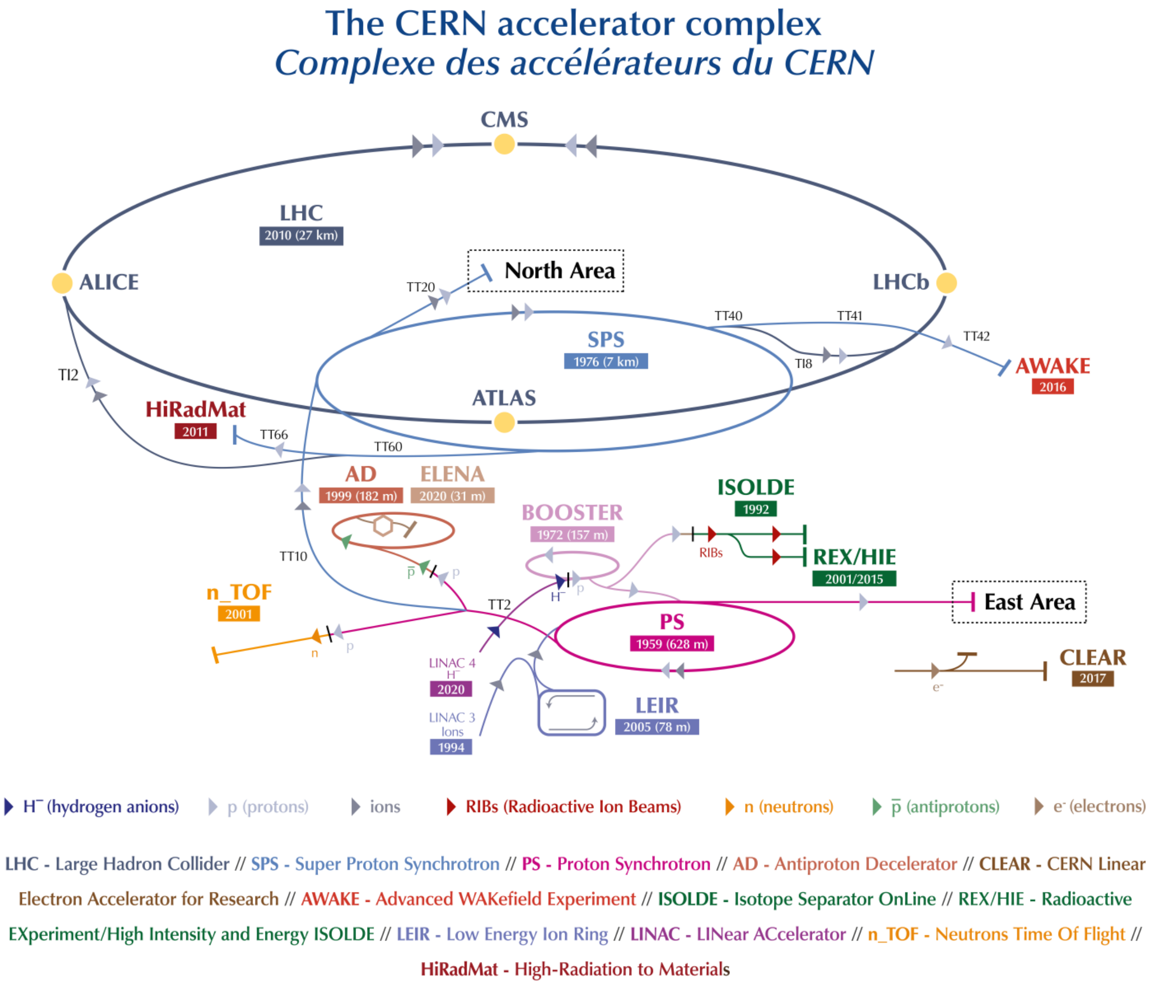

The CERN accelerator complex
Complexe des accélérateurs du CERN
CMS
LHC
2010 (27 km)
ALICE
North Area
LHCb
TT20
TT40
TT41
TT42
SPS
1976 (7 km)
TI8
AWAKE
2016
TI2
HiRadMat
2011
TT66
ATLAS
TT60
AD
1999 (182 m)
ELENA
2020 (31 m)
ISOLDE
1992
BOOSTER
1972 (157 m)
RIBs
REX/HIE
2001/2015
TT10
n_TOF
2001
TT2
East Area
PS
1959 (628 m)
CLEAR
2017
LINAC 4
H⁻
2020
LINAC 3
Ions
1994
LEIR
2005 (78 m)
H⁻ (hydrogen anions)
p (protons)
ions
RIBs (Radioactive Ion Beams)
n (neutrons)
p̄ (antiprotons)
e⁻ (electrons)
LHC - Large Hadron Collider // SPS - Super Proton Synchrotron // PS - Proton Synchrotron // AD - Antiproton Decelerator // CLEAR - CERN Linear Electron Accelerator for Research // AWAKE - Advanced WAKefield Experiment // ISOLDE - Isotope Separator OnLine // REX/HIE - Radioactive EXperiment/High Intensity and Energy ISOLDE // LEIR - Low Energy Ion Ring // LINAC - LINear ACcelerator // n_TOF - Neutrons Time Of Flight // HiRadMat - High-Radiation to Materials

Figure 5: The CERN accelerator complex as of 2019 [12].

of the proton beams.

3.2 The LHCb detector

The Large Hadron Collider beauty (LHCb) detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, described in detail in References [13, 39] and shown schematically in Figure 6. What follows is an introductory-level description of some of its most-relevant features.

Designed to collect the largest sample of $b\bar{b}$ pairs in the world, the LHCb detector has a

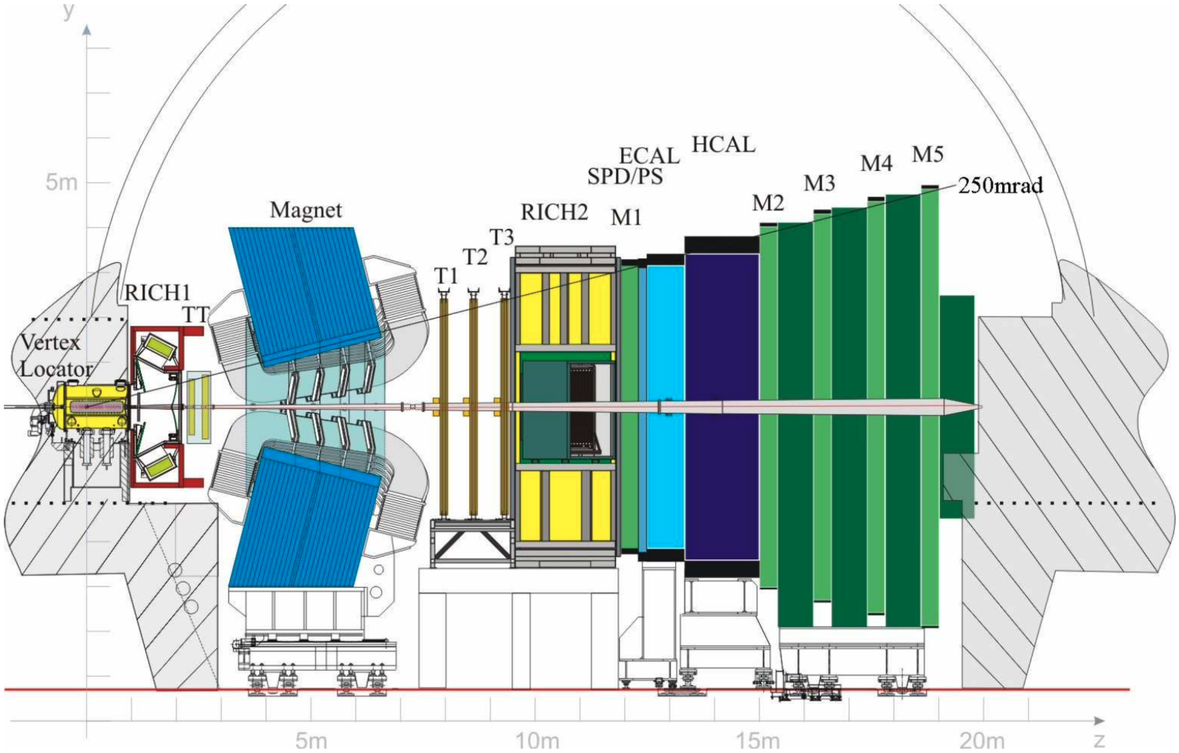


Figure 6: The LHCb detector, shown schematically in profile [13].

unique design, especially among detectors operating on the LHC. Most obviously, it operates only in the forward region. Proton-proton (pp) collisions occur on the left-hand side of Figure 6 and produce particles in all directions, but LHCb only detects those in a conical region extending to the right (the "forward" region, along the positive z-axis). This is the pseudorapidity range $2 < \eta < 5$, where pseudorapidity is defined as

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$

and θ is the angle with respect to (w.r.t.) the z-axis. The reason LHCb operates in such a restricted region has to do with the production properties of $b\bar{b}$ pairs in pp collisions, specifically that they are overwhelmingly produced in a forward and backward cone. Rather than build a double detector in each direction, only one region (the forward cone) is chosen for simplicity, exploiting the symmetry of the collision.

LHCb does not use the full capabilities of the LHC in that it uses a process called "luminosity leveling" to reduce the number of proton collisions. Luminosity ($\mathcal{L}$) is the number of particles passing along the beam line per unit time per unit area; it is roughly analogous to the likelihood of collision. The LHC can deliver $\mathcal{L} \approx 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, but the p beams are adjusted before colliding in the LHCb detector so that it receives $\mathcal{L} = 2 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, only $\approx 1\%$ of the LHC's capacity. This leveling still allows for the production of trillions of $b\bar{b}$ pairs annually and provides some advantages, including lower radiation damage to the detector and easier event analysis, since there is usually a single pp interaction per bunch crossing. LHCb first collected data from 2010 to 2012, known as "Run I", during which roughly $3\,\mathrm{fb}^{-1}$ of data were collected at a pp center-of-mass energy of $7\,\mathrm{TeV}$.[18] Run II began in 2015 and ran through 2018, during which an additional $\approx 5.5\,\mathrm{fb}^{-1}$ of data were collected with a center-of-mass energy of $13\,\mathrm{TeV}$.

The LHCb detector comprises a number of subdetectors, described below, dedicated to specific tasks. Unless otherwise stated, the facts and figures in the following sections are taken from References [13, 39].

3.2.1 Magnet

LHCb makes use of a dipole magnet to measure the momenta of charged particles. (Charged particles moving perpendicular to a magnetic field have their trajectories bent at an angle proportional to their momenta.) It consists of two saddle-shaped coils in a window-frame yoke, for a total weight of 1600 tons, and is shown in Figure 7. It was designed to expose passing particles to an integrated magnetic field of 4 Tm for particles traversing 10 m while respecting the low-field needs of the RICH and creating a field as large as possible between the VELO and TT (see Sections 3.2.2, 3.2.3, and 3.2.7). In addition, the magnet polarity is occasionally switched during LHCb data-taking to help with the cancelation of certain

---

[18] The amount of data is given in terms of "integrated luminosity" ($\int \mathcal{L}$), the instantaneous luminosity integrated over the data-taking periods. It is, by convention, given in terms of inverse barns (b^{-1}), a whimsically-named unit of area equivalent to $10^{-28}\,\mathrm{m}^{-2}$.

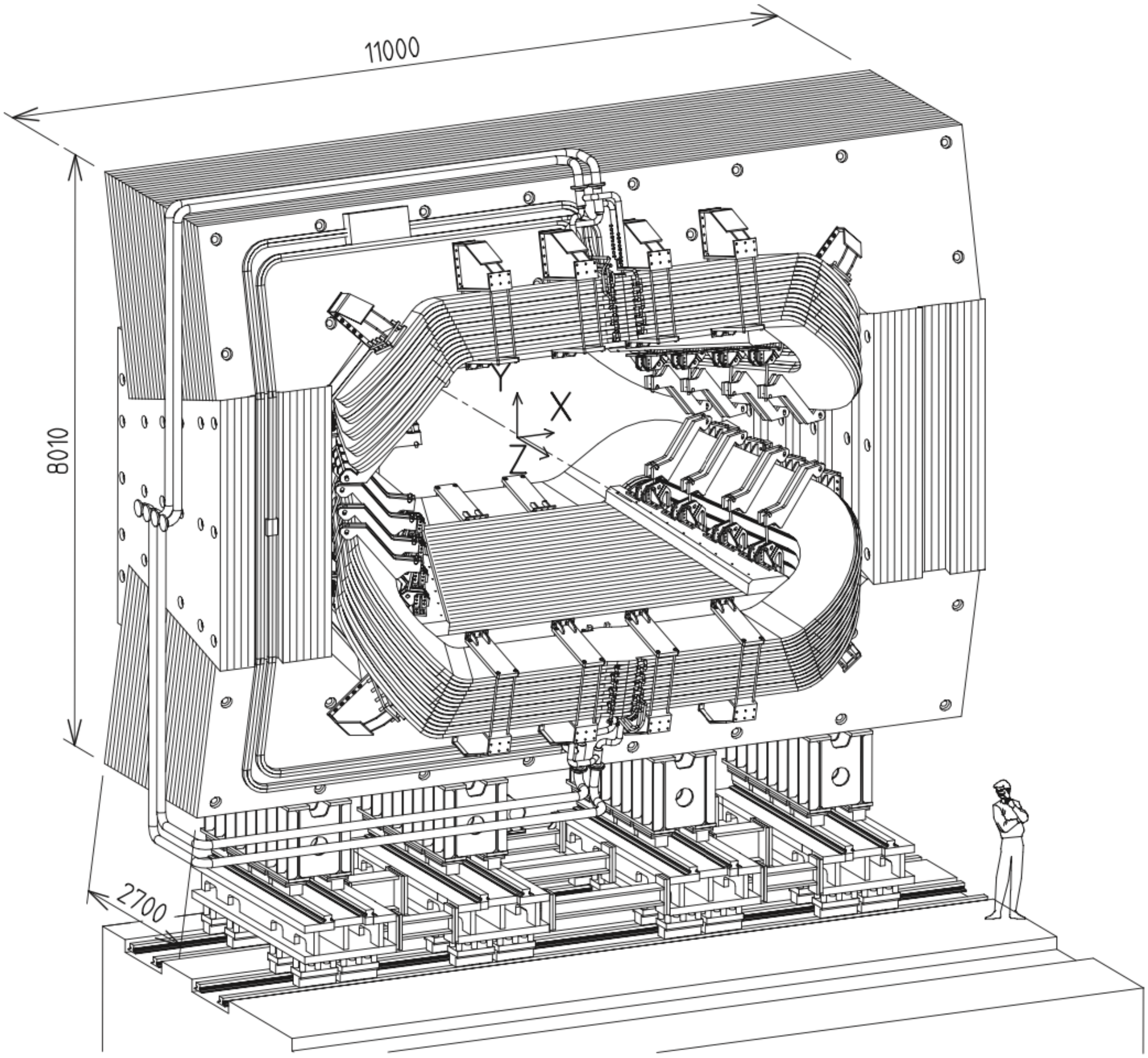


Figure 7: Drawing of the LHCb magnet [13].

systematic uncertainties, so the magnet needed to perform identically in both cases (magnet polarity up and down). The magnet has been found to provide a uniform field (±1%) and satisfy the field strength requirements to a precision of 4×10^{-4}.

3.2.2 Vertex Locator

Proton collisions in LHCb occur in the Vertex Locator (VELO), located around the interaction point. It is a silicon-strip tracking detector that provides precise positional information about particles that pass through it. Since b and c hadrons are relatively long-lived, one of the primary functions of the VELO is to distinguish the PV from the secondary decay vertices of b and c hadrons; that is, both proton collisions and hadron decays produce particles that leave tracks in the detector, and the VELO helps to differentiate the precise origins of those tracks. The VELO is made up of silicon modules, arranged along the beam line, that measure r and ϕ coordinates (see Figure 8).[19] It sits so close to the beam that it has to be retractable to accommodate the size of the beam during injection (see Figure 9). (The LHC focuses the beam over time.) Its sensors are situated in a vacuum separated from the LHC vacuum by a thin aluminum sheet (see Figure 10).

The VELO was subject to several design constraints, including a signal to noise ratio (S/N) greater than 14 : 1, high signal efficiency ($\geq 99\%$ at $S/N > 5$), and a resolution around 0.004 mm. The use of $r\phi$ coordinates was chosen to enable faster reconstruction of tracks in the trigger (see Section 3.2.9); the sensors are arranged in pairs, one of which provides r information with the other providing ϕ information, and the arrangement of the pairs along the z-axis provides z information, thus giving three-dimensional spatial information about the location of hits in the VELO. It also needed to withstand an extreme amount of radiation, the equivalent of 1 MeV neutrons with a flux of $1.3 \times 10^{14} n_{\mathrm{eq}}/\mathrm{cm}^2$ annually, initially for 3 years. Besides the general radiation-hardness requirements this put on the materials, the high radiation environment also necessitated a sophisticated cooling system capable of keeping the sensors between -10 and 0°C.

3.2.3 Tracker Turicensis

The Tracker Turicensis (TT)[20] sits just upstream of the dipole magnet (see Section 3.2.1) and is a silicon-strip tracking detector. It is 150 cm wide by 130 cm tall, covering the full detector acceptance, and comprises four layers, with the silicon strips of the outer two layers arranged vertically and those of the inner two rotated at $-5°$ and $5°$, thus providing stereo tracking. To aid in the track reconstruction process, they are arranged in two pairs separated by

---

[19] r is the distance from the beam line, and ϕ is the angle from the x-axis.

[20] The Tracker Turicensis is the new name for what was formerly known as the Trigger Tracker.

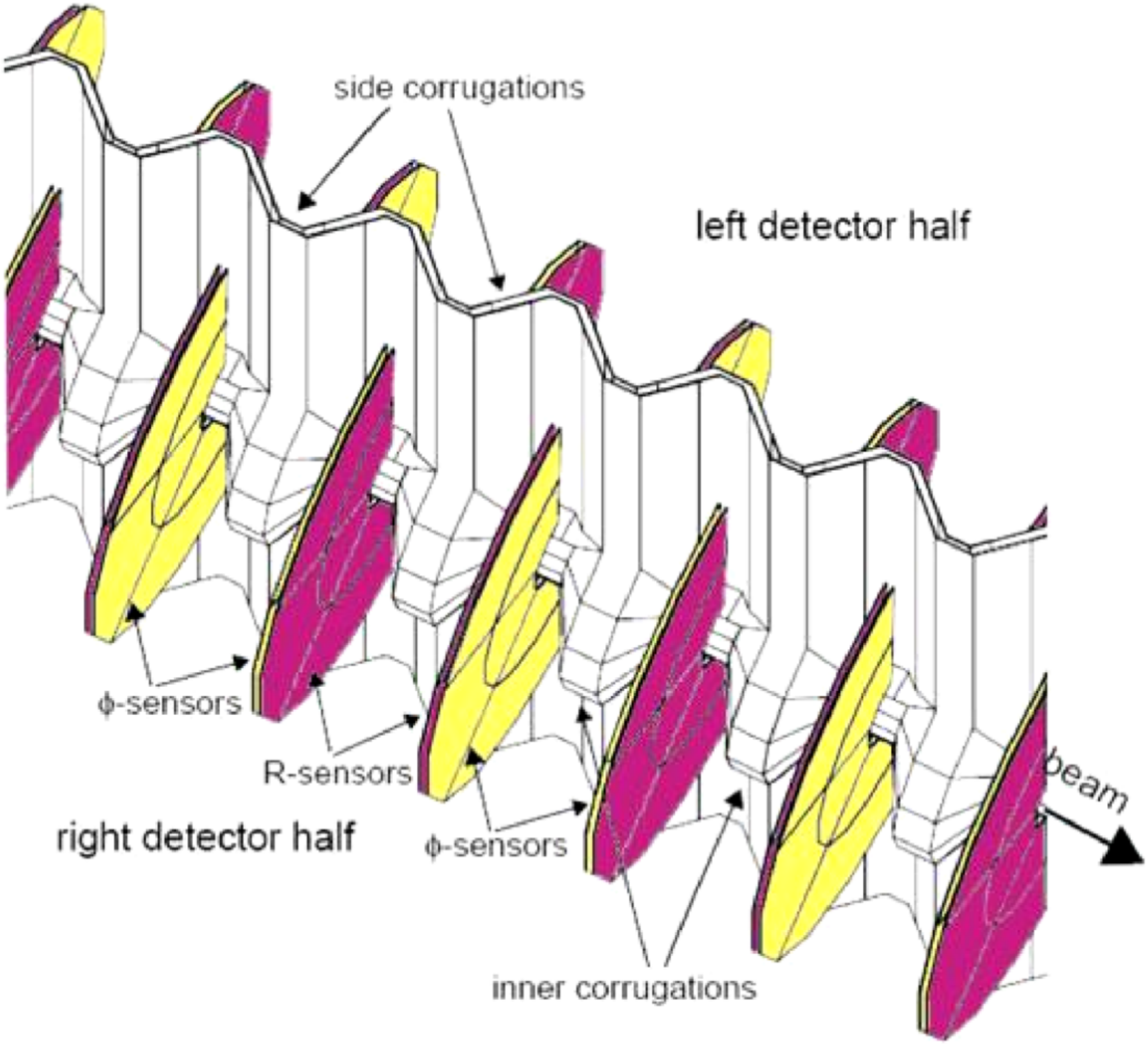

side corrugations
left detector half
ϕ-sensors
R-sensors
right detector half
ϕ-sensors
beam
inner corrugations

Figure 8: Perspective illustration of the VELO sensors [13].

≈ 27 cm along the beam line. It was designed to achieve a single-hit resolution of ≈ 0.050 mm, $S/N > 12:1$, and to withstand $5 \times 10^{14} n_{\mathrm{eq}}/\mathrm{cm}^2$ of 1 MeV neutron equivalent fluences over ten years.

The layers of the TT are thermally and electrically insulated in a single light-tight detector volume, which is kept below 5°C to maintain detector performance and flushed with dry nitrogen to prevent condensation. The sensors are arranged in half-modules that are half the height of the detector, each of which holds seven sensors. The sensors in a given layer are staggered in x and z to avoid gaps. The readout electronics and cooling support structures

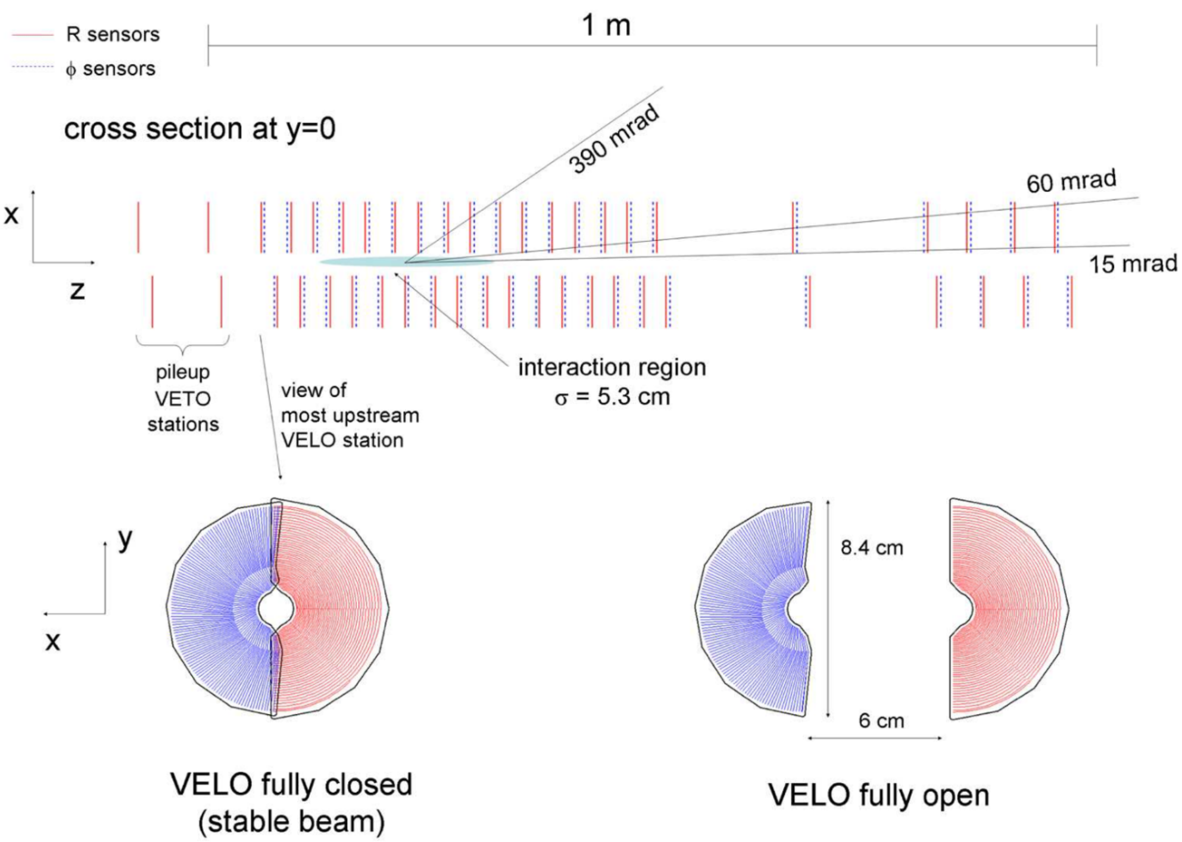

R sensors
ϕ sensors
1 m
cross section at y=0
390 mrad
60 mrad
15 mrad
x
z
pileup
VETO
stations
view of
most upstream
VELO station
interaction region
σ = 5.3 cm
y
x
8.4 cm
6 cm
VELO fully closed
(stable beam)
VELO fully open

Figure 9: Cross-sectional drawing of the VELO [13].

are located outside the detector acceptance. Figure 11 shows the layout of one TT layer.

3.2.4 Inner Tracker

The Inner Tracker (IT) is a silicon-strip tracking detector comprising a 120 cm wide and 40 cm tall cross-shaped region in the middle of the three tracking stations downstream of the LHCb magnet (T1, T2, and T3 in Figure 6). It was designed to achieve a single-hit resolution of $\approx 0.050\,\text{mm}$, $S/N > 12:1$, and to withstand $9 \times 10^{12} n_{\text{eq}}/\text{cm}^2$ of 1 MeV neutron equivalent fluences over ten years. Each of its three layers is arranged into four separate detector boxes, which are light tight, electrically and thermally insulated, maintained below 5°C, and flushed with nitrogen, similar to the TT (see Section 3.2.3). Also like the TT, the modules are staggered in x and z to avoid gaps. Unlike the TT, there are only one or two sensors per module. Figure 12 shows the layout of one IT layer.

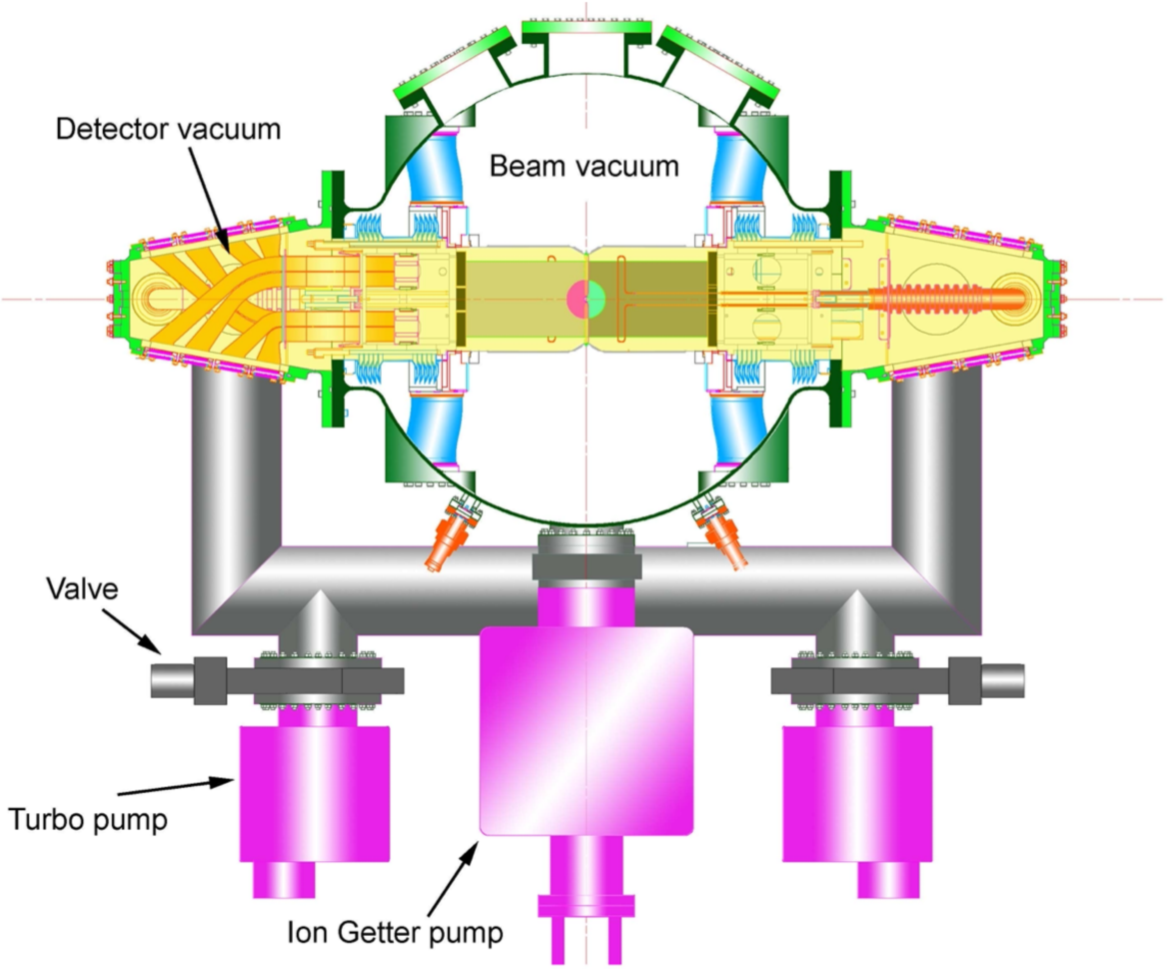


Figure 10: Illustration of the VELO vacuum system [13].

3.2.5 Outer Tracker

The Outer Tracker (OT) is a drift-time detector occupying the three tracking stations downstream of the LHCb magnet (T1, T2, and T3 in Figure 6) on the outside of the IT (see Section 3.2.4). It was designed to achieve a momentum resolution of $\Delta p/p \approx 0.4\%$. It comprises an array of gas-tight straw-tube modules containing two layers of drift tubes, each of which has an inner diameter of 4.9 mm. Each tracking station has four OT layers, with the outer two having their straw tubes arranged vertically and the inner two having theirs rotated by -5° and 5°, similar to the TT (see Section 3.2.3). Figure 13 shows the layout of the OT.

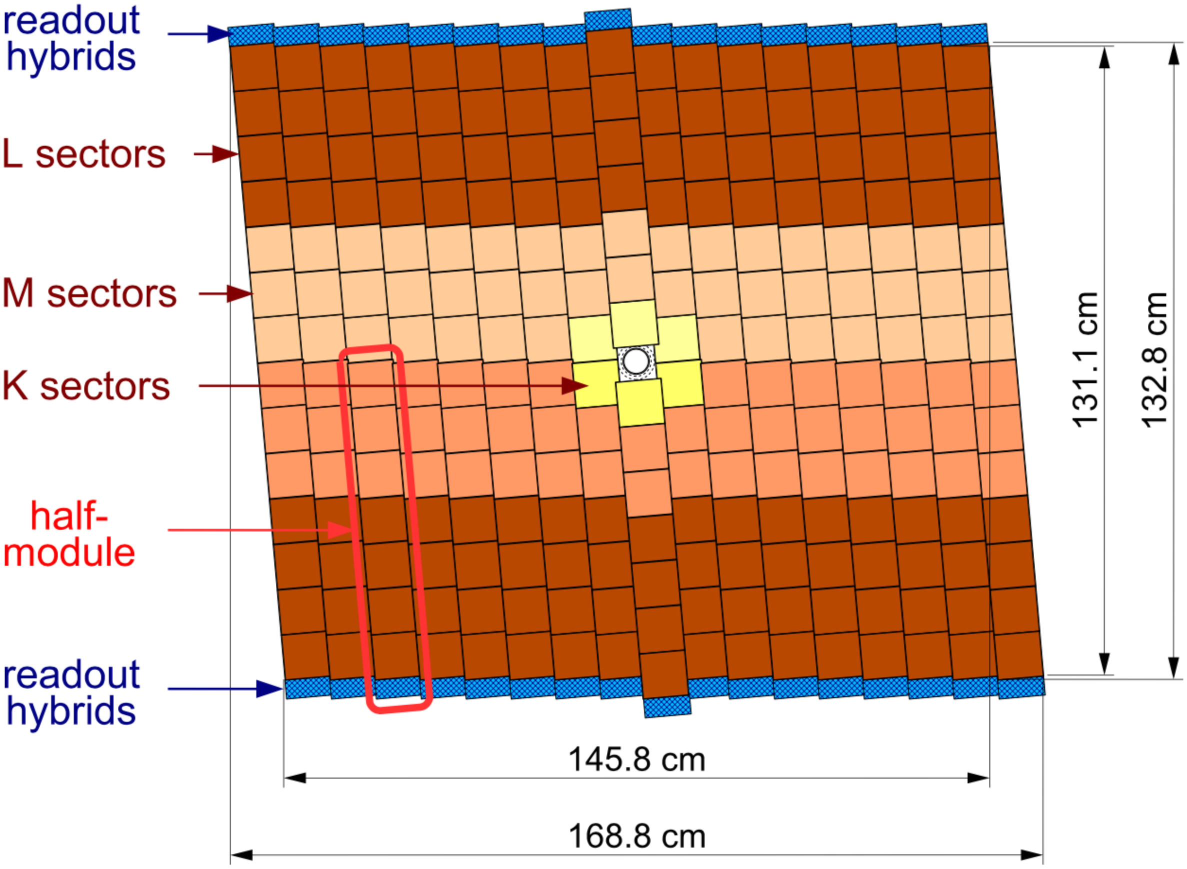

readout hybrids
L sectors
M sectors
K sectors
half-module
readout hybrids
131.1 cm
132.8 cm
145.8 cm
168.8 cm

Figure 11: A schematic of the third TT layer [13].

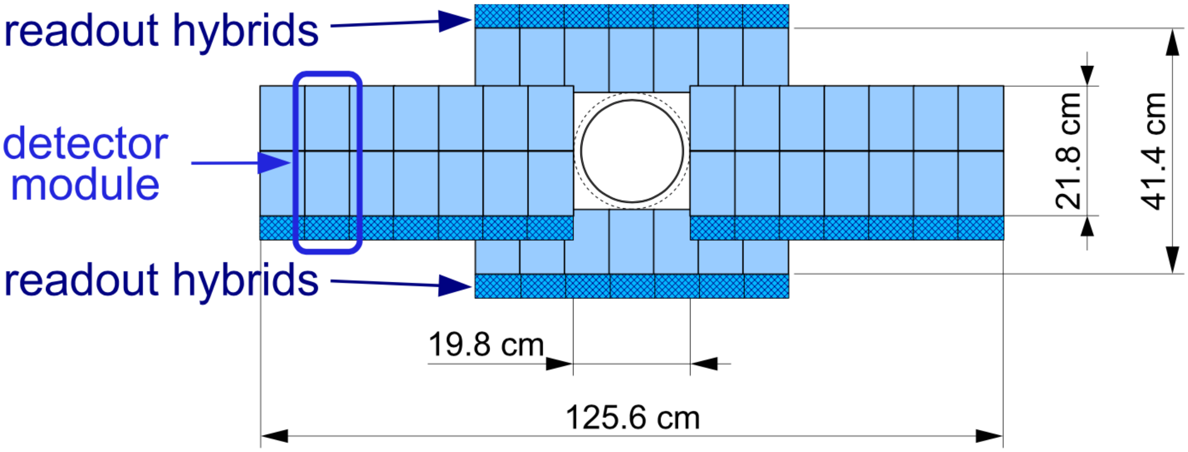

readout hybrids
detector module
readout hybrids
21.8 cm
41.4 cm
19.8 cm
125.6 cm

Figure 12: A schematic of one IT layer [13].

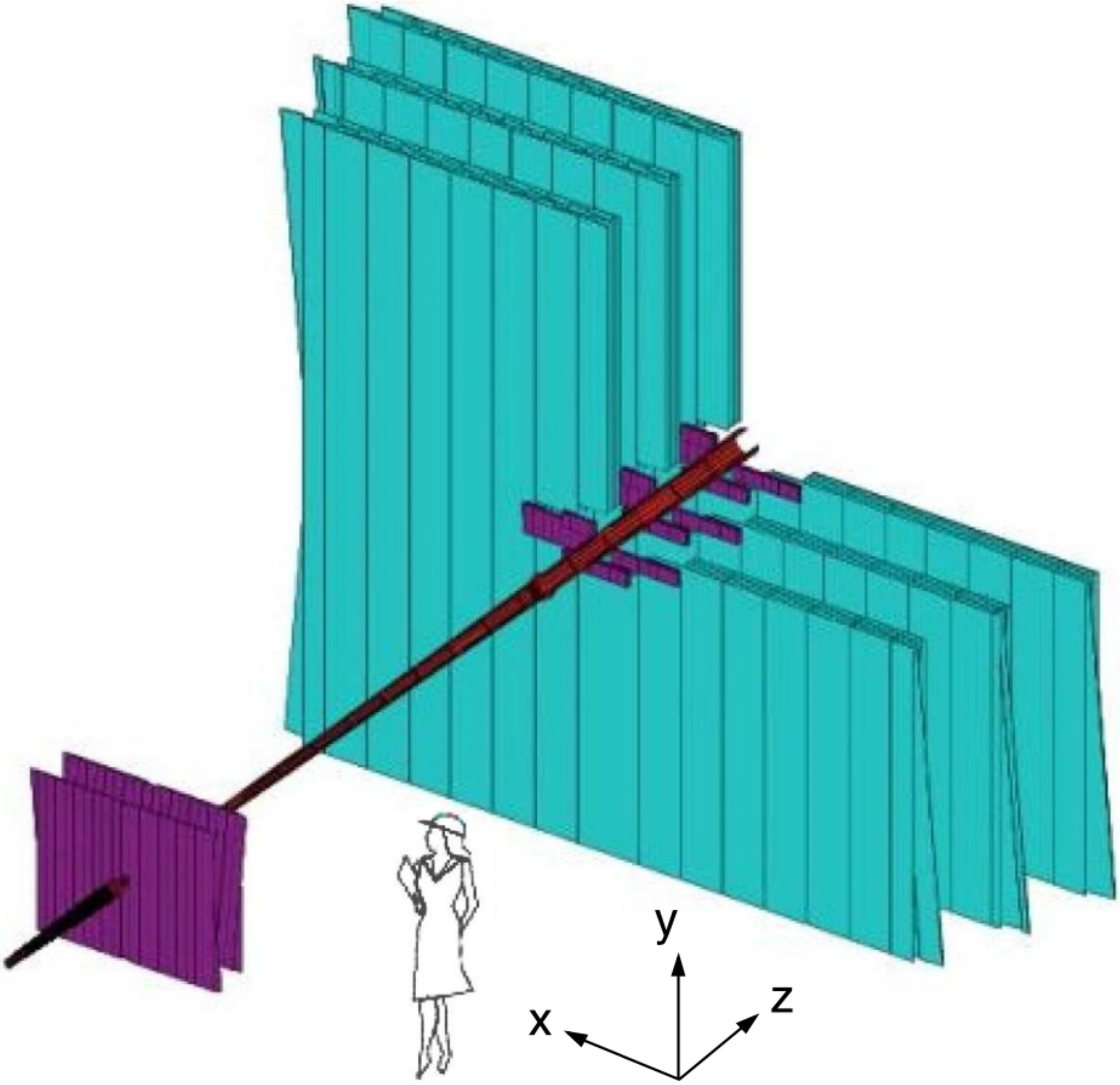


Figure 13: Layout of the LHCb tracking system with a cross-sectional cut-out in the $-x, +y$ quadrant of the three downstream tracking stations [13]. The OT is shown in teal while the TT and IT are shown in purple. See Figure 6.

3.2.6 Calorimeters

The calorimeter system, situated between the first and second muon stations (see Section 3.2.8), is essential to the functioning of the LHCb trigger system (see Section 3.2.9), providing electron, photon, and hadron identification and measuring their energies and positions. It comprises an electromagnetic calorimeter (ECAL) followed by a hadronic calorimeter (HCAL). To improve the trigger performance of the ECAL, a preshower detector (PS) is placed just upstream of the main section of the ECAL to help reject backgrounds from charged pions. To help reject neutral pions, a scintillator pad detector (SPD) is placed in front of the PS to select charged particles. All of the calorimeters operate by the transmission of scintillation light to a photomultiplier by wavelength-shifting fibers; the basic idea is that particle interactions within the detector bulk release energy in the form of light, which is picked up and transmitted to the electronics.

3.2.7 Ring Imaging Cherenkov detectors

One of the less common and more essential aspects of the LHCb detector is its particle identification (PID) capabilities. The two Ring Imaging Cherenkov (RICH) detectors are largely responsible for these capabilities. RICH1 is located upstream of the magnet, covers the full detector acceptance, and is suitable for low momentum particles; RICH2 is located downstream of the magnet, covers only a narrow acceptance range, and is suitable for high momentum particles, which are preferentially produced closer to the beam pipe. The RICH detectors are named for Cherenkov radiation, which is a cone of light produced as a charged particle passes through certain media—the angular opening of the cone is proportional to the velocity of the particle. Taken together with the momentum measured by the tracking system, this allows for an estimate of the mass of the particle, as shown in Figure 14. Complicating its performance is the presence of multiple particles passing through the detector simultaneously; Figure 15 shows an example. The RICH system achieved an angle resolution of $(1.66 \pm 0.03)\,\text{mrad}$ in test beam studies, and the system readout allows for a determination of, *e.g.*, whether a reconstructed particle is more likely to be a kaon than a pion.

3.2.8 Muon system

The muon systems provide muon triggering (see Section 3.2.9) and identification capabilities. They comprise five stations, shown in Figures 6 and 16, located downstream of the magnet

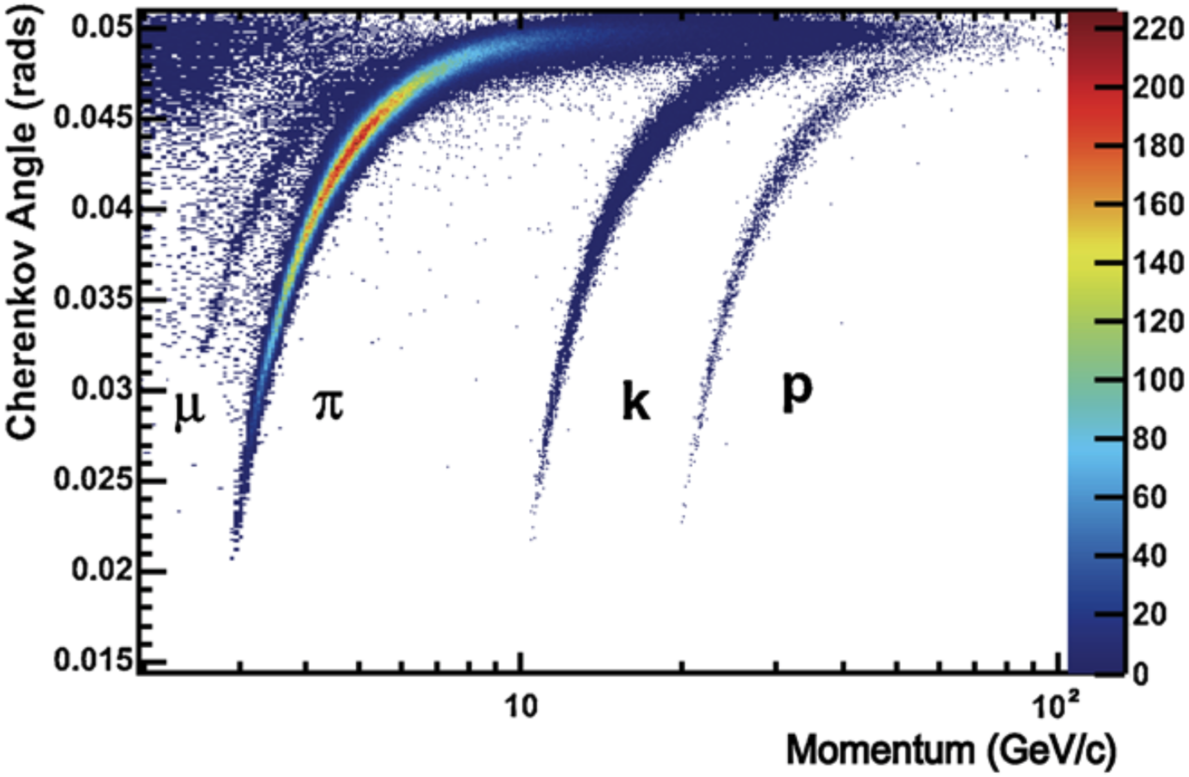


Figure 14: Separation of muons, pions, kaons, and protons as a function of momentum in data collected by the RICH1 detector [14].

at the end of the detector. The detection capabilities are provided by a combination of multi-wire proportional chambers and triple gas electron multipliers. Iron absorbers 80 cm thick are placed after most of the stations; they help block most non-muon particles and impose high momentum requirements that help with the trigger system. The first three stations have good spatial resolution and are used to measure the muon momentum transverse to the beam line (p_{T}); the last two stations are mostly designed to indicate whether a particle penetrated the iron filters, rejecting non-muons.

3.2.9 Trigger

As discussed above, LHCb is designed to operate at a much lower luminosity than the design luminosity of the LHC. Bunch crossings occur in the LHC every 25 ns, but the frequency of events that produce at least two charged particles that can be reconstructed by the VELO

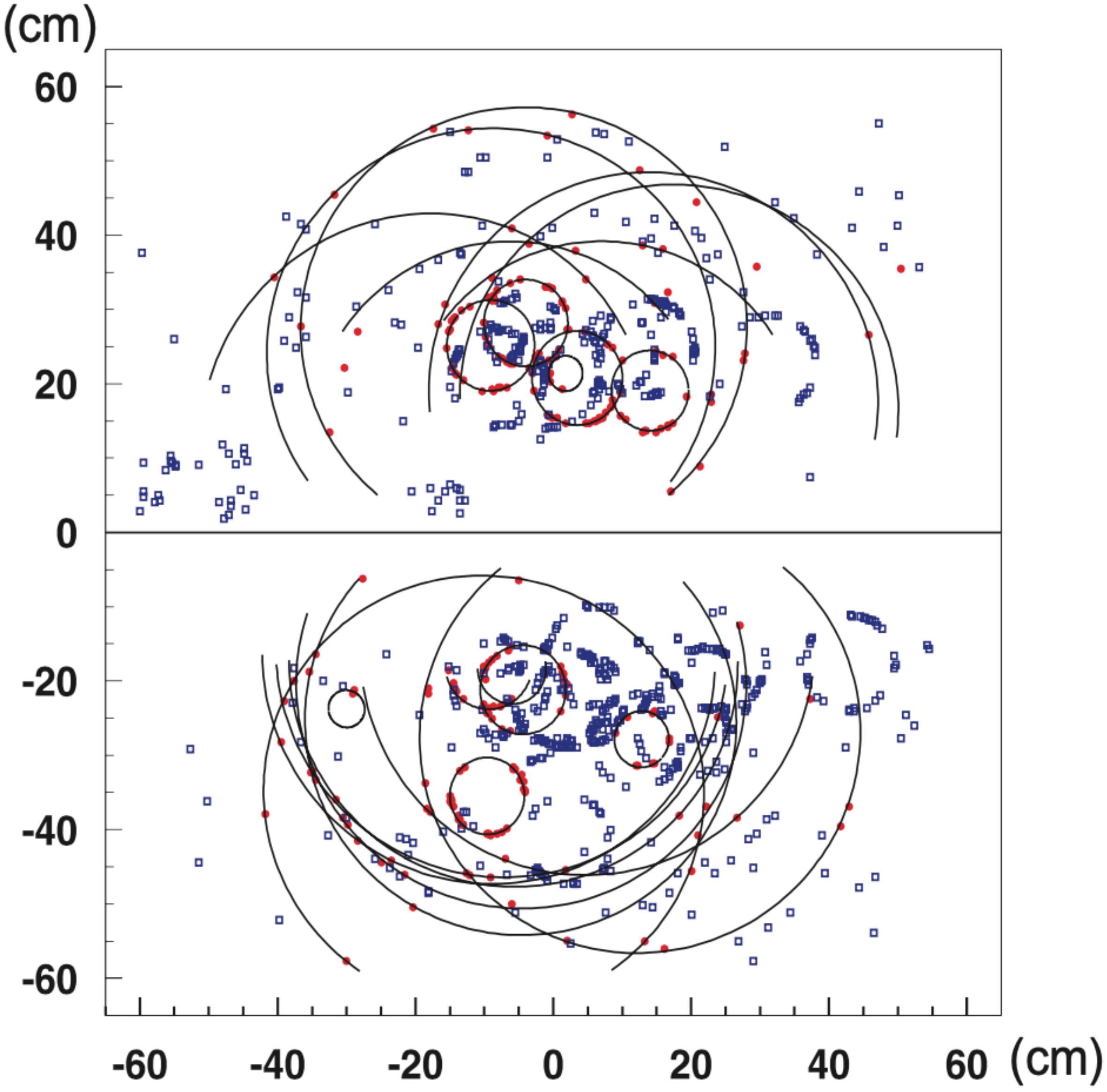


Figure 15: Display of typical simulated event in RICH1 [13].

and tracking stations (the minimum requirements to be visible to the detector) is about 10 MHz. This must be further reduced by the trigger to a level suitable for writing to storage for offline analysis.[21] The trigger system is thus tasked with rejecting uninteresting events and saving interesting ones for further study.

---

[21] Paritcle physics jargon distinguishes between "online" and "offline" event processing. "Online" refers to processing that occurs during data taking, roughly in real time. "Offline" refers to processing and analysis that occurs after an event has been saved and can therefore be done at any point after data-taking is complete.

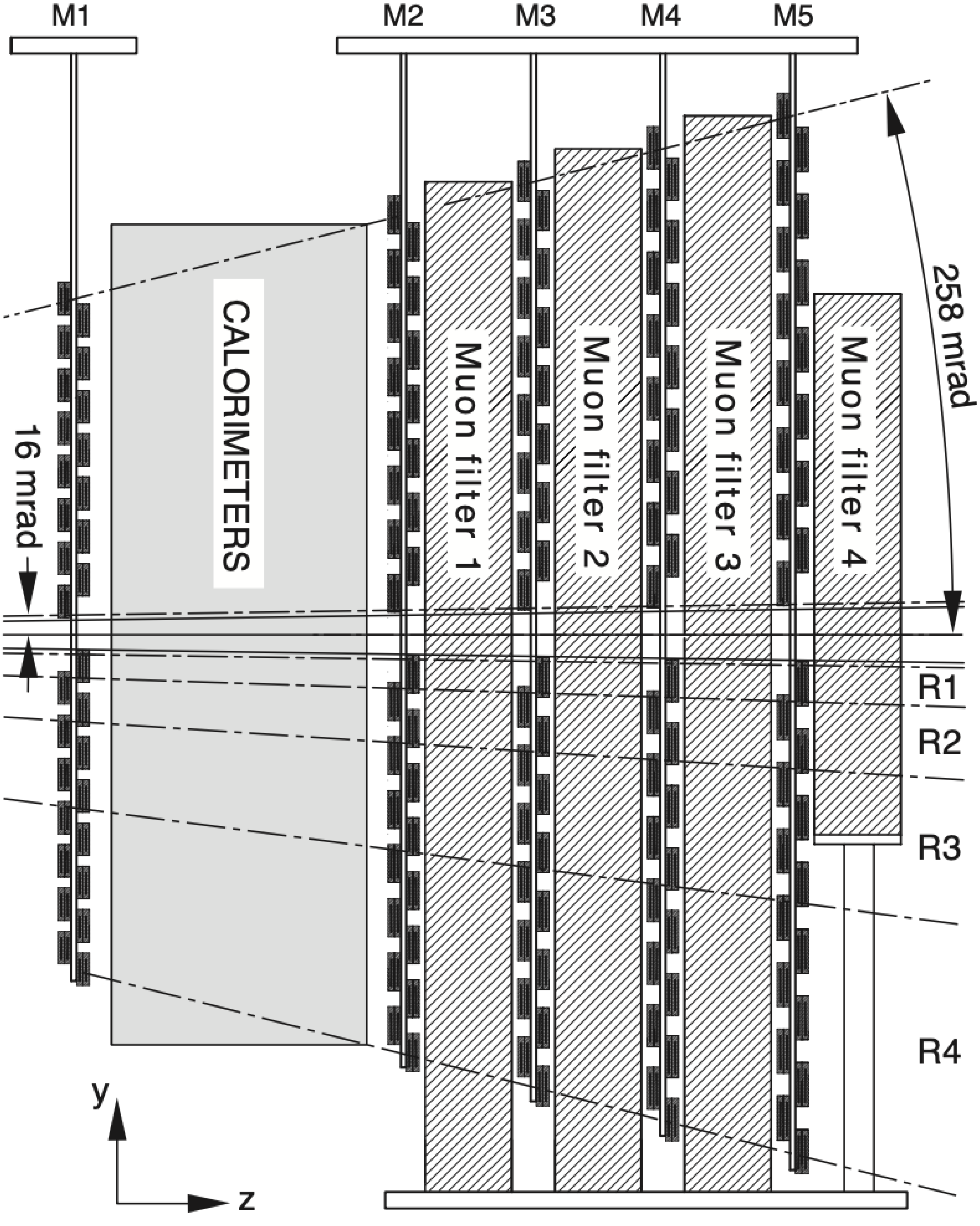

M1
M2
M3
M4
M5
CALORIMETERS
Muon filter 1
Muon filter 2
Muon filter 3
Muon filter 4
258 mrad
16 mrad
R1
R2
R3
R4
y
z

Figure 16: Layout of the muon stations [13].

The LHCb trigger system is split into two levels: Level-0 (L0), the hardware stage, and the High Level Trigger (HLT), the software stage. The L0 stage is executed synchronously (with the 40 MHz LHC bunch crossing) by electronics and brings the 10 MHz visible interaction rate down to 1 MHz primarily by considering the particles with the highest transverse momentum (p_{T}) and energy (E_{T}) in an event and estimating the total number of tracks. These pieces of information help it select events likely to contain b hadrons, whose daughters tend to have high p_{T}, without being bogged down by combinatoric problems from high track-multiplicity.

The HLT stage runs asynchronously on a processor farm and uses all the available data in an event to bring down the event rate. It is split into two parts. HLT1 uses partial event reconstruction, relying on the VELO and tracking stations, to confirm the L0 results and provide further refinement, reducing the rate to about 30 kHz. HLT2 then performs a full event reconstruction and makes sophisticated decisions about which events to save. In Run I, the HLT2 performed a simpler, faster reconstruction than could be done offline (see Section 3.3); in Run II, a major effort was undertaken to improve the HLT2 so that it performed identically to the offline reconstruction, eliminating this discrepancy.

3.3 The LHCb data flow

The raw data collected by the detector, not to mention generated in simulation, is processed by a number of software applications to make it suitable for analysis. These applications are based on the Gaudi framework [40] and serve a variety of purposes.

Data from the detector is initially selected by the trigger, described in Section 3.2.9, which is handled by MOORE [41]. A more sophisticated offline reconstruction is then performed by BRUNEL [42]. BRUNEL is run in centralized reconstruction campaigns, which occur only rarely due to their computational intensity. After reconstruction, the various electrical readouts from the detector have been thoroughly transformed into physics objects in software, allowing decay chains, *e.g.*, $\mathit{\Xi}_c^0 \to \pi^- \, (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, to be constructed, but the dataset is still quite large and not suited to access by analysts. LHCb therefore conducts Stripping campaigns to further reduce the size of the dataset beyond the simple trigger selections; these are centrally run data-reduction campaigns, which occur more frequently than the reconstruction campaigns, that produce output suitable for analysts. The stripping campaigns are handled by DaVinci [42], which is also typically used to query the stripped datasets to create n-tuples for use in analyses.[22]

There is also a parallel data flow used in LHCb called the "Turbo stream". Since the reconstruction executed in HLT2 (using Moore) and in Brunel perform identically in Run II, as mentioned in Section 3.2.9, analysts have the choice of skipping the reconstruction and stripping steps and accessing data directly from HLT2 in the Turbo stream. Due to rate considerations, the data available in this stream is generally less comprehensive and more restricted than in the "full stream", but it can be convenient if no more information is required.

Physics simulation in LHCb is handled by Gauss, which coordinates and runs several different applications to produce effective simulation [43]. It uses Pythia to simulate pp collisions [44, 45], EvtGen to simulate particle decay [46], and Geant4 to model particle interactions with the detector [47, 48]. The response of the detector itself is simulated by Boole [42], and from this point on, simulation and detector data follow the same processing path. The jargon for simulated data is "Monte Carlo" (MC), for the popular gambling destination, while data collected by the detector is generally referred to simply as "data".

Figure 17 illustrates the LHCb data flow for data and MC.

---

[22] An n-tuple is an ordered list of sets of values. For example, one might record the x-, y-, and z-momenta of a series of Λ_c^+ hadrons as an n-tuple.

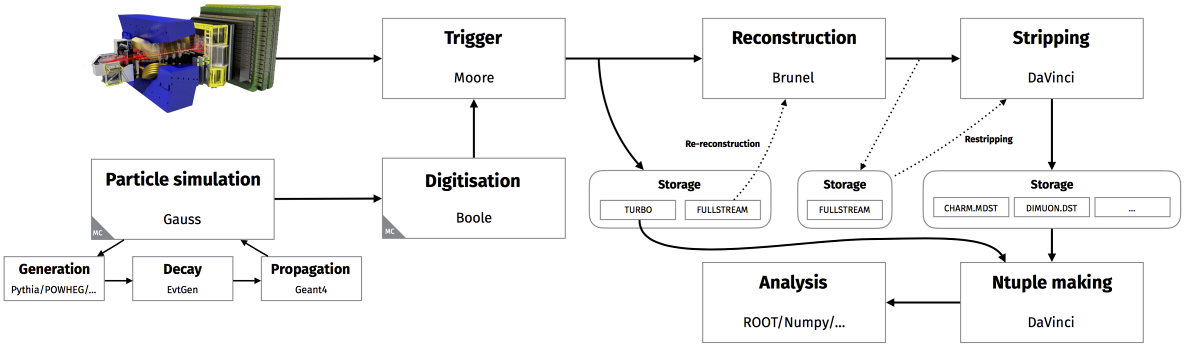

Trigger
Moore
Reconstruction
Brunel
Stripping
DaVinci
Re-reconstruction
Restripping
Particle simulation
Gauss
MC
Digitisation
Boole
MC
Storage
TURBO
FULLSTREAM
Storage
FULLSTREAM
Storage
CHARM.MDST
DIMUON.DST
...
Generation
Pythia/POWHEG/...
Decay
EvtGen
Propagation
Geant4
Analysis
ROOT/Numpy/...
Ntuple making
DaVinci

Figure 17: The LHCb data flow in Run II [15].

4 First branching fraction measurement of the suppressed decay $\Xi_c^0 \to \pi^- \Lambda_c^+$

This section is an adaptation of an analysis note written with Sheldon Stone, published as Reference [17], which formed the basis for a paper published with the LHCb collaboration as Reference [49]. All of the figures used are taken from References [17, 49]. Charge-conjugation is implied throughout this section.[23]

4.1 Introduction

Baryons containing both an s quark and a heavy c or b quark, denoted as Q, usually decay to lighter particles via the decay of the heavy quark. There is, however, the possibility of the s quark causing the transformation. Theoretical predictions concerning the decay widths of $\Xi_Q \to \pi \Lambda_Q$ are based on the size of the s-quark decay amplitude $s \to u(\overline{u}d)$ (SUUD) and the "weak scattering" (WS) amplitude $Qs \to dQ$ [50]; the Feynman diagrams corresponding to these amplitudes are shown in Figure 18 for Ξ_c^0 decay.

The WS amplitude is not present in all such decays, *e.g.*, $\Xi_b^- \to \pi^- \Lambda_b^0$ decay. This allows interesting comparisons to be made using the measured rates. Studies of these Ξ_Q baryon decays connect theories concerning hyperon[24] decays to those for the heavy b and c quarks.

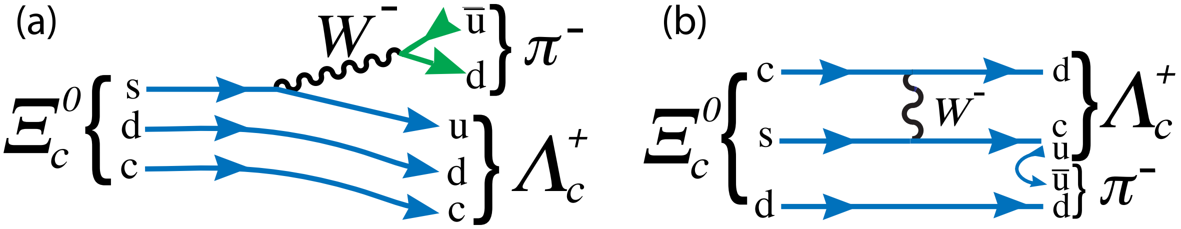


Figure 18: Decay diagrams for $\Xi_c^0 \to \pi^- \Lambda_c^+$ transitions via (a) the SUUD amplitude and (b) the WS amplitude.

---

[23] Whenever a decay is referenced in this section, the same decay with matter exchanged for anti-matter and vice versa should be understood to be referenced also.

[24] "Hyperon" is jargon for a baryon containing at least one s quark but no c or b (or t) quarks, *e.g.*, the Λ baryon (uds).

The former use partially conserved axial currents (PCAC) and SU(3) symmetry, whereas the latter apply more modern approaches, including heavy-quark effective theory (HQET). These transitions can result in decays into a Λ_b^0 baryon and either a π^0 or a π^- meson; here we discuss only decays involving the π^- meson. As the Ξ_b^- baryon consists of b, s, and d quarks, the WS amplitude is not present in $\Xi_b^- \to \pi^- \Lambda_b^0$ decay [51], so the measurement of that decay rate can be used to determine the SUUD amplitude. This information can be used to predict the $\Xi_c^0 \to \pi^- \Lambda_c^+$ decay rate that, in principle, involves both amplitudes.

The well-known Ξ_c^0 baryon consists of the c, s, and d quarks. Its lifetime has been measured to be $154.5 \pm 1.7 \pm 1.6 \pm 1.0\,\text{fs}$ [2]. Both the SUUD and WS s-quark decay processes are allowed [50], but the explicit decay $\Xi_c^0 \to \pi^- \Lambda_c^+$ has not been previously observed.[25] Several authors have made predictions using the measured Ξ_b^- SUUD amplitude and the measured lifetimes of the SU(3) triplet baryons, Ξ_c^0, Λ_c^+, and Ξ_c^+, as input for determining the WS amplitude. This method was pioneered by Voloshin where he used SU(3) symmetry, PCAC, and the heavy-quark limit to determine an upper limit on $\Gamma(\Xi_b^- \to \pi^- \Lambda_b^0)$ [50]. In a subsequent paper, he uses the input from the LHCb measurement of $\mathcal{B}(\Xi_b^- \to \pi^- \Lambda_b^0) = (0.60 \pm 0.18)\%$ [51] and updated values for the charmed-baryon lifetimes to find the SUUD rate and then calculates the WS amplitude, predicting $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) \gtrapprox (0.25 \pm 0.15) \times 10^{-3}$ after assuming negative interference between the two strangeness-changing amplitudes [4].

Gronau and Rosner, using the same approach as Voloshin, predict two possible branching fractions for $\Xi_c^0 \to \pi^- \Lambda_c^+$ decay, depending on the sign of the interference between the two decay amplitudes [5]. Based on the measured $\mathcal{B}(\Xi_b^- \to \pi^- \Lambda_b^0)$ [51] and using the available charmed-baryon lifetimes [8], they predict $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = (0.19 \pm 0.07)\%$ for constructive interference and $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) \lessapprox 0.01\%$ for destructive interference between the SUUD and WS contributions. We have redone their calculation using updated lifetime measurements from LHCb [2], finding $\mathcal{B}\,(\Xi_c^0 \to \pi^- \Lambda_c^+) = (0.14 \pm 0.07)\%$ for constructive interference and $\mathcal{B}\,(\Xi_c^0 \to \pi^- \Lambda_c^+) \lessapprox (0.018 \pm 0.015)\%$ for destructive; see Appendix A.12.

---

[25] A Belle paper shows a peak near the Ξ_c^0 mass for this decay [52]. However, they quote neither a signal significance, nor a branching fraction.

Faller and Mannel, on the other hand, predict $\mathcal{B}\,(\Xi_c^0 \!\rightarrow \pi^- \Lambda_c^+) < 0.3\%$, an upper limit arrived at by assuming constructive interference [6]. Finally, Cheng *et al.* predict $\mathcal{B}\,(\Xi_c^0 \!\rightarrow \pi^- \Lambda_c^+) \sim 0.0087\%$, assuming negative interference [7]. We have not updated these predictions; the effect would be to lower Faller and Mannel's positive interference prediction and raise Cheng *et al.*'s negative one, giving somewhat better agreement with Gronau and Rosner's predictions.

Measurement of $\mathcal{B}(\Xi_c^0 \rightarrow \pi^- \Lambda_c^+)$ is hampered by the lack of precisely measured Ξ_c^0 branching fractions [10]. The Belle collaboration measured $\mathcal{B}(\Xi_c^0 \rightarrow \Xi^- \pi^+) = (1.80 \pm 0.50 \pm 0.14)\%$, $\mathcal{B}(\Xi_c^0 \rightarrow \Lambda K^- \pi^+) = (1.17 \pm 0.37 \pm 0.09)\%$, and $\mathcal{B}(\Xi_c^0 \rightarrow p K^- K^- \pi^+) = (0.58 \pm 0.23 \pm 0.05)\%$, but these either involve hyperons, which have low reconstruction efficiencies in LHCb, or carry $> 40\%$ relative uncertainties [53].[26] We therefore take a somewhat unusual approach to measure $\mathcal{B}(\Xi_c^0 \!\rightarrow \pi^- \Lambda_c^+)$. First, we note that

$$\frac{N(\Xi_c^0)}{N(\Lambda_c^+)} = \frac{f_{\Xi_c^0}}{f_{\Lambda_c^+}}, \tag{1}$$

where N indicates the number of (efficiency-corrected) particles in a dataset and f_i indicates the relative production fraction of charm-baryon i. (All baryon yields refer only to those promptly produced from the initial *pp* collision; we assign a systematic uncertainty due to this assumption in Section 4.7.11.) The branching fraction is defined as

$$\mathcal{B}(\Xi_c^0 \!\rightarrow \pi^- \Lambda_c^+) = N(\Xi_c^0 \!\rightarrow \pi^- \Lambda_c^+)/N(\Xi_c^0), \tag{2}$$

so, from Equation 1,

$$\frac{N(\Xi_c^0 \!\rightarrow \pi^- \Lambda_c^+)}{N(\Lambda_c^+)} = \frac{f_{\Xi_c^0}}{f_{\Lambda_c^+}} \times \mathcal{B}(\Xi_c^0 \!\rightarrow \pi^- \Lambda_c^+). \tag{3}$$

---

[26] The PDG fit value for $\mathcal{B}(\Xi_c^0 \rightarrow p K^- K^- \pi^+) = (0.48 \pm 0.12)\%$, carrying a 25% relative uncertainty, was published shortly before the measurement presented here, well after this analysis was complete [25, 49].

To obtain the ratio of production fractions, we use the heavy-quark symmetry relation,

$$\frac{f_{\Xi_c^0}}{f_{\Lambda_c^+}} = \mathcal{C} \times \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}}, \tag{4}$$

where $\mathcal{C}$ is a correction factor for differences in excited states between the charm and bottom sectors, which we calculate in Section 4.6.2. $\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}}$ has been measured by LHCb in 13 TeV pp collisions to be $(8.2 \pm 0.7 \pm 2.6)\%$ [54]; an additional 5% will be added to this already large relative uncertainty for the heavy-quark symmetry assumption (see Section 4.7.1). In this analysis, we use only $\Lambda_c^+ \to pK^-\pi^+$ decays, so in summary,

$$\mathcal{B}_1 \equiv \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = \frac{N(\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+))}{N(\Lambda_c^+ \to pK^-\pi^+)} \times \frac{1}{\mathcal{C}} \times \frac{f_{\Lambda_b^0}}{f_{\Xi_b^-}}. \tag{5}$$

The Belle collaboration measured $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$ to be $(0.45 \pm 0.21 \pm 0.07)\%$ [55]. Using an approach similar to that above, we can use this to make an alternative measurement of $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$. We note that

$$\frac{N(\Xi_c^0)}{N(\Xi_c^+)} = \frac{f_{\Xi_c^0}}{f_{\Xi_c^+}} \tag{6}$$

$$\frac{N(\Xi_c^0 \to \pi^- \Lambda_c^+)}{N(\Xi_c^+)} = \frac{f_{\Xi_c^0}}{f_{\Xi_c^+}} \times \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) \tag{7}$$

and that

$$\frac{f_{\Xi_c^0}}{f_{\Xi_c^+}} \approx 1 \tag{8}$$

due to isospin symmetry.[27] In this analysis, we use only $\Lambda_c^+ \to pK^-\pi^+$ and $\Xi_c^+ \to pK^-\pi^+$ decays, so in summary,

$$\mathcal{B}_2 \equiv \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = \frac{N(\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+))}{N(\Xi_c^+ \to pK^-\pi^+)} \times \frac{\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)}{\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)}, \tag{9}$$

where $\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+) = (6.23 \pm 0.33)\%$ [1].

---

[27] See Section 84.2 of Reference [10].

In addition, by equating Equations 5 and 9 and solving for $\mathcal{B}(\mathit{\Xi}_c^+ \to pK^-\pi^+)$, we find

$$\mathcal{B}_3 \equiv \mathcal{B}(\mathit{\Xi}_c^+ \to pK^-\pi^+) = \frac{N(\mathit{\Xi}_c^+ \to pK^-\pi^+)}{N(\mathit{\Lambda}_c^+ \to pK^-\pi^+)} \times \frac{1}{\mathcal{C}} \times \frac{f_{\mathit{\Lambda}_b^0}}{f_{\mathit{\Xi}_b^-}} \times \mathcal{B}(\mathit{\Lambda}_c^+ \to pK^-\pi^+), \tag{10}$$

allowing a determination of $\mathcal{B}(\mathit{\Xi}_c^+ \to pK^-\pi^+)$ without measuring any additional parameters.

4.2 Data sample

This analysis is based on the data sample collected by the LHCb detector at a 13 TeV center-of-mass pp collision energy during the Run II years 2017 and 2018, which correspond to integrated luminosities of $1.6\,\mathrm{fb}^{-1}$ and $2.2\,\mathrm{fb}^{-1}$, respectively. We use data passing specific L0, HLT1, and HLT2 (Turbo) trigger lines (see Section 4.3). The decays $\mathit{\Xi}_c^0 \to \pi^- \, (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and $\mathit{\Xi}_c^+ \to pK^-\pi^+$ are reconstructed using `Turbo04` in 2017 and `Turbo05` in 2018. They are then processed into useable n-tuples using DaVinci v45r2.

The software versions to produce simulation are summarized in Table 1. The simulation data are produced with the Gauss framework, which employs Pythia 8 for simulating the fragmentation and hadronisation processes for the signal $\mathit{\Xi}_c^0$, $\mathit{\Lambda}_c^+$, and $\mathit{\Xi}_c^+$ baryons; EvtGen for simulating the decay of the signal baryons; and Geant4 for simulating the detector response. Digitization is performed with Boole, and from this point on, simulation undergoes the same processing as data: The trigger is applied using Moore; Brunel is used for event reconstruction; and DaVinci is used to apply the Turbo and stripping. (See Section 3.3.) Track smearing is performed to ensure that the simulated data have similar momentum and mass resolutions as real data.

The simulation samples used in this analysis are listed in Table 2. The $\mathit{\Xi}_c^0$, $\mathit{\Lambda}_c^+$, and $\mathit{\Xi}_c^+$ baryons decay to their immediate daughters via a phase-space model. The intermediate resonances in $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ proceed via various models, described in Table 3. (The intermediate spectra of $\mathit{\Xi}_c^+ \to pK^-\pi^+$ has not been measured; see Section 4.7.9.) The generator-level selection applied to the $\mathit{\Xi}_c^+ \to pK^-\pi^+$ sample must be accounted for, since

Table 1: Software used to generate simulation for the Ξ_c^0 and Λ_c^+ samples. The software for the Ξ_c^+ sample sometimes differs; where this is the case, it is given in parentheses.

Package	2016 version	2017 version	2018 version
GAUSS	v49r13 (v49r12)	v49r13 (v49r12)	v49r13 (v49r12)
BOOLE	v30r3	v30r4 (v30r3)	v30r4
MOORE	v25r4	v26r6p1	v28r3p1
BRUNEL	v50r5	v52r6p1	v54r2 (v54r1)
DAVINCI (Turbo)	v41r4p3	v42r8p3	v44r7
DAVINCI (Stripping)	v41r4p4	v42r7p3	v44r7

it requires final-state daughters to be within the coverage of the LHCb detector, whereas the generator-level selections applied to the other samples only require the mother charm baryon to be within coverage (see Section 4.4).[28] We use 2016 simulation in addition to 2017 and 2018 in order to maximize the statistics; there is no difference relevant to this analysis between the simulation samples for these years.

Table 2: Simulation samples (2016, 2017, and 2018 combined) used for BDT training and efficiency studies before any selections have been applied. See Section 4.4 for explanations of the generator-level cuts.

Decay	Statistics	Generator-Level Cuts
$\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$	11.98M	`LHCbAcceptance`
$\Lambda_c^+ \to pK^-\pi^+$	12.21M	`LHCbAcceptance`
$\Xi_c^+ \to pK^-\pi^+$	12.06M	`DaughtersInLHCb`

4.3 Event selections

We use a two-step process to maximize the statistical significance of the datasets. First, we apply loose selections with large signal efficiencies and good background suppression. Then, we apply three different boosted decision trees (BDT) [56, 57], one for each baryon decay, to further separate signal from background.

---

[28] In LHCb, generator-level cuts are selections applied directly to simulated events, before they are reconstructed by the detector, whereas most selections are applied to reconstructed candidates. See Section 3.3.

Table 3: Λ_c^+ decay models used. The rates listed are from the PDG [1], but they were not adjusted to account for the branching fractions of the intermediate states, *e.g.*, $\overline{K}^{*0} \to K^-\pi^+$. Since we weight the Dalitz spectra to match data, however, this should have a negligible impact on the results; see Sections 4.3.4 and 4.7.9.

Decay	Rate	Intermediate resonance decay model
$\Lambda_c^+ \to p(\overline{K}^{*0} \to K^-\pi^+)$	1.94%	VSS
$\Lambda_c^+ \to K^-(\Delta^{++} \to p\pi^+)$	1.07%	PHSP
$\Lambda_c^+ \to \pi^+(\Lambda^*(1520) \to pK^-)$	2.20%	PHSP
$\Lambda_c^+ \to pK^-\pi^+$ non-resonant	3.40%	

4.3.1 Turbo selections

The $\Lambda_c^+ \to pK^-\pi^+$ and $\Xi_c^+ \to pK^-\pi^+$ data samples comes from the LHCb Turbo stream, and Table 4 summarizes their selections. We then combine the Λ_c^+ baryon candidates with π^- meson candidates persisted in the Turbo stream to create $\Xi_c^0 \to \pi^-\ (\Lambda_c^+ \to pK^-\pi^+)$ candidates, as summarized in Table 5. Unless they are used to build a Ξ_c^0 candidate, we keep only 10% of the Λ_c^+ candidates in data in order to improve processing speed and save storage space; we account for this by multiplying the final yield by 10.

For simulation, we do not use the Turbo stream directly owing to the presence of particle identification (PID) requirements. Instead, we rebuild the Turbo selections by hand using general sources of p, K^-, and π^+ candidates, excluding the PID requirements (e.g., $\mathrm{DLL}_{p\pi} > 5$), which we reproduce offline using an independent data sample (see Section 4.3.4).

4.3.2 Trigger selections

In the Ξ_c^0 and Λ_c^+ samples, we require each signal Λ_c^+ baryon to satisfy certain trigger requirements (TOS[29]), and we apply identical requirements to the signal Ξ_c^+ baryon in the Ξ_c^+ sample, as described in Table 6. This approach minimizes any biasing effects from differing trigger selections.

---

[29] "Trigger On Signal", or "TOS", indicates that the candidate in question passed a given set of trigger selections, opposed to simply being in an event where *any* particle passed the selections.

Table 4: Selections made in the Turbo stream. See text. H_c is either the signal $\varLambda_c^+$ or the signal $\varXi_c^+$ baryon, as appropriate. χ^2_{IP} is the difference in the PV vertex fit χ^2 with and without the p, K^-, and π^+ tracks. GhostProb is the probability the track is a ghost. ndf is the number of degrees of freedom in the fit. HasRich indicates whether the track has information in the RICH. $\text{DLL}_{p\pi}$ is the difference in log likelihood that a particle is a proton rather than a π meson, similarly for other particles. $\vec{x}$ is the position vector pointing from the PV to the particle's decay vertex; $\frac{\vec{x}(H_c)\cdot\vec{p}(H_c)}{x(H_c)\cdot p(H_c)}$ is thus equivalent to the cosine of the angle between the particle's momentum and its decay vertex. See Section 4.3.2 for an explanation of the trigger requirements.

Property	Value
after refitting PVs	
$m\,(\varLambda_c^+)$	$> 2211\,\text{MeV}, < 2362\,\text{MeV}$
$m\,(\varXi_c^+)$	$> 2392\,\text{MeV}, < 2543\,\text{MeV}$
before refitting PVs	
$m\,(H_c)$	$> 2201\,\text{MeV}, < 2553\,\text{MeV}$
$\chi^2_{\text{IP}}(p)$ or $\chi^2_{\text{IP}}(K^-)$ or $\chi^2_{\text{IP}}(\pi^+)$	> 16
at least two of $\chi^2_{\text{IP}}(p)$, $\chi^2_{\text{IP}}(K^-)$, $\chi^2_{\text{IP}}(\pi^+)$	> 9
$\chi^2_{\text{IP}}(p)$ and $\chi^2_{\text{IP}}(K^-)$ and $\chi^2_{\text{IP}}(\pi^+)$	> 6
$p_{\text{T}}(p)$ or $p_{\text{T}}(K^-)$ or $p_{\text{T}}(\pi^+)$	$> 1\,\text{GeV}$
at least two of $p_{\text{T}}(p)$, $p_{\text{T}}(K^-)$, $p_{\text{T}}(\pi^+)$	$> 400\,\text{MeV}$
$p_{\text{T}}(p)$ and $p_{\text{T}}(K^-)$ and $p_{\text{T}}(\pi^+)$	$> 200\,\text{MeV}$
$\sum_{p,K^-,\pi^+} p_{\text{T}}$	$> 3\,\text{GeV}$
$p(K^-)$ and $p(\pi^+)$	$> 1\,\text{GeV}$
$p(p)$	$> 10\,\text{GeV}$
$\chi^2_{\text{vtx}}/\text{ndf}(H_c)$	< 10
$\frac{\vec{x}(H_c)\cdot\vec{p}(H_c)}{x(H_c)\cdot p(H_c)}$	$> \cos(10\,\text{mrad})$
$\tau(H_c)$	$> 150\,\text{fs}$
$\chi^2_{\text{track}}/\text{ndf}(p)$ and $\chi^2_{\text{track}}/\text{ndf}(K^-)$ and $\chi^2_{\text{track}}/\text{ndf}(\pi^+)$	< 3
GhostProb(p) and GhostProb(K^-) and GhostProb(π^+)	< 0.4
$\text{DLL}_{p\pi}(p) - \text{DLL}_{K\pi}(p)$	> 5
$\text{DLL}_{p\pi}(p)$	> 5
$\text{DLL}_{K\pi}(K^-)$	> 5
$\text{DLL}_{K\pi}(\pi^+)$	< 5
HasRich(p) and HasRich(K^-) and HasRich(π^+)	True
any `Track` HLT1 line TOS H_c	True

Table 5: Selections made to produce Ξ_c^0 candidates. See text. χ^2_{vtx} is the fit χ^2 of the decay vertex.

Λ_c^+, π^- combination	
Property	Value
after refitting PVs	
$\chi^2_{\text{vtx}}(\Xi_c^0)$	< 25
$p_{\text{T}}(\Xi_c^0)$	$> 1\,\text{GeV}$
$c \times \tau\,(\Xi_c^0)$	$> 0.05\,\text{mm}$
before refitting PVs	
$p_{\text{T}}(\Xi_c^0)$	$> 950\,\text{MeV}$
$m\,(\Xi_c^0) - m\,(\Lambda_c^+)$	$< 240\,\text{MeV}$

Table 6: Trigger requirements. The signal charm-baryon must pass at least one of the listed trigger lines for each stage of the trigger (TOS). The "signal charm-baryon" is defined as the signal Λ_c^+ baryon candidate for the $\Xi_c^0 \to \pi^-\ (\Lambda_c^+ \to pK^-\pi^+)$ and $\Lambda_c^+ \to pK^-\pi^+$ samples and as the signal Ξ_c^+ baryon for the $\Xi_c^+ \to pK^-\pi^+$ sample. `L0HadronDecision` is based on the response of the hadronic calorimeters; the `Hlt1TrackMVA` lines select events with quality tracks formed in HLT1.

Stage	Trigger Lines
L0	`L0HadronDecision`
HLT1	`Hlt1TwoTrackMVADecision`
	`Hlt1TrackMVADecision`

4.3.3 Offline selections

Before passing the data and simulation datasets to the BDTs (see Section 4.3.5), we make a few basic selections, summarized in Table 7. DecayTreeFitter (DTF) recalculates selected parameters based on certain assumptions [58]; in this case, it assumes the signal Ξ_c^0, Λ_c^+, or Ξ_c^+ baryon came from the primary vertex (PV) and that all daughters form vertices with their mothers. (We do not apply any mass constraints.) Several of the offline selections, most notably those requiring a parameter to be > 0, are meant primarily to ensure DTF completed the fit successfully. (Negative values would indicate a failed fit.)

Table 7: Offline selections. See text. Selections requiring a parameter to be > 0 ensure DTF completed the fit successfully, as negative values would indicate a failed fit.

Property	Value
$\Xi_c^0 \to \pi^- \left(\Lambda_c^+ \to pK^-\pi^+\right)$	
$\mathrm{DTF}\left(m\left(\Xi_c^0\right)\right) - \mathrm{DTF}\left(m\left(\Lambda_c^+\right)\right)$	$> 100\,\mathrm{MeV},\ < 300\,\mathrm{MeV}$
$m\left(\Lambda_c^+\right)$	$> 2266.46\,\mathrm{MeV},\ < 2306.46\,\mathrm{MeV}$
$\mathrm{DTF}\left(m\left(\Xi_c^0\right)\right)$	$> 0\,\mathrm{MeV}$
$\chi^2(\mathrm{DTF}\left(\tau\left(\Xi_c^0\right)\right))$	> 0
$\mathrm{DTF}\left(\frac{\vec{x}(\Xi_c^0)\cdot\vec{p}(\Xi_c^0)}{x(\Xi_c^0)\cdot p(\Xi_c^0)}\right)$	> 0.999
$\frac{\vec{x}(\Xi_c^0)\cdot\vec{p}(\Xi_c^0)}{x(\Xi_c^0)\cdot p(\Xi_c^0)}$	> 0.999
$\mathrm{GhostProb}(\pi^-, p, K^-, \pi^+)$	< 0.3
$p_{\mathrm{T}}(\Xi_c^0)$	$> 5\,\mathrm{GeV}$
$\Lambda_c^+ \to pK^-\pi^+$	
$m\left(\Lambda_c^+\right)$	$> 2220\,\mathrm{MeV},\ < 2350\,\mathrm{MeV}$
$\chi^2(\mathrm{DTF}\left(\tau\left(\Lambda_c^+\right)\right))$	> 0
$\mathrm{DTF}\left(\frac{\vec{x}(\Lambda_c^+)\cdot\vec{p}(\Lambda_c^+)}{x(\Lambda_c^+)\cdot p(\Lambda_c^+)}\right)$	> 0.999
$\frac{\vec{x}(\Lambda_c^+)\cdot\vec{p}(\Lambda_c^+)}{x(\Lambda_c^+)\cdot p(\Lambda_c^+)}$	> 0.999
$\mathrm{GhostProb}(p, K^-, \pi^+)$	< 0.3
$p_{\mathrm{T}}(\Lambda_c^+)$	$> 5\,\mathrm{GeV}$
$\Xi_c^+ \to pK^-\pi^+$	
$\mathrm{DTF}\left(m\left(\Xi_c^+\right)\right)$	$> 2410\,\mathrm{MeV},\ < 2540\,\mathrm{MeV}$
$\chi^2(\mathrm{DTF}\left(\tau\left(\Xi_c^+\right)\right))$	> 0
$\mathrm{DTF}\left(\frac{\vec{x}(\Xi_c^+)\cdot\vec{p}(\Xi_c^+)}{x(\Xi_c^+)\cdot p(\Xi_c^+)}\right)$	> 0.999
$\frac{\vec{x}(\Xi_c^+)\cdot\vec{p}(\Xi_c^+)}{x(\Xi_c^+)\cdot p(\Xi_c^+)}$	> 0.999
$\mathrm{GhostProb}(p, K^-, \pi^+)$	< 0.3
$p_{\mathrm{T}}(\Xi_c^+)$	$> 5\,\mathrm{GeV}$

4.3.4 Additional simulation selections

The simulation serves two purposes: to calculate the signal-detection efficiencies and to provide an independent signal dataset with which to train the BDTs (see Section 4.3.5). To achieve both these ends, we require all reconstructed signal candidates in simulation to correspond to their generated signal counterparts. This ensures that we feed the BDTs a pure signal sample and that we calculate the *signal* efficiencies explicitly. See Appendix A.1 for details.

We weight the generated lifetimes of the simulated $\mathit{\Xi}_c^0$, $\mathit{\Lambda}_c^+$, and $\mathit{\Xi}_c^+$ baryons to match the values from recent LHCb measurements [2] using the expression

$$\exp\left[-\left(\frac{t\left(H_c\right)}{\tau_{\text{meas}}\left(H_c\right)}-\frac{t\left(H_c\right)}{\tau_{\text{gen}}\left(H_c\right)}\right)\right], \tag{11}$$

where H_c is the charm baryon in question, $t\left(H_c\right)$ is the decay time of the simulated particle, $\tau_{\text{meas}}\left(H_c\right)$ is the measured lifetime of the baryon, and $\tau_{\text{gen}}\left(H_c\right)$ is the lifetime used in the simulation by EvtGen, shown in Table 8 (see Section 4.2).[30] We assign a systematic uncertainty due to the uncertainty of the measurements in Table 8 (see Section 4.7.2).

Table 8: Charm-baryon lifetime values. Measurements are taken from Reference [2], where uncertainties are statistical, systematic, and due to the uncertainty on the lifetime of the D^+ meson, respectively. The simulated values are those used by EvtGen; see text.

Particle	Measured [fs]	Simulated [fs]
$\mathit{\Xi}_c^0$	$154.5 \pm 1.7 \pm 1.6 \pm 1.0$	155
$\mathit{\Lambda}_c^+$	$203.5 \pm 1.0 \pm 1.3 \pm 1.4$	200
$\mathit{\Xi}_c^+$	$456.8 \pm 3.5 \pm 2.9 \pm 3.1$	442

To ensure we correctly reproduce the intermediate spectra and opening angles of the $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ and $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays, we weight their Dalitz spectra according to their distributions in data. After applying all selections (Turbo, offline, BDT, and trigger, described elsewhere in this section), we use the *sPlot* method [16] to assign weights to the data sample from the total fits shown in Figures 25 and 26 in Section 4.5.1. We then generate the appropriate Dalitz plots using the $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ and $\mathit{\Xi}_c^+ \to pK^-\pi^+$ data and simulation datasets and normalize them to one. We take the ratio of the normalized histograms and use them to weight all simulation accordingly, as shown in Figure 19. Appendix A.10 shows the results of the weighting. We assign a systematic uncertainty due to this weighting; see Section 4.7.9.

---

[30] The lifetime of a particle, τ, is just the inverse of the decay constant in the common formula for exponential decay.

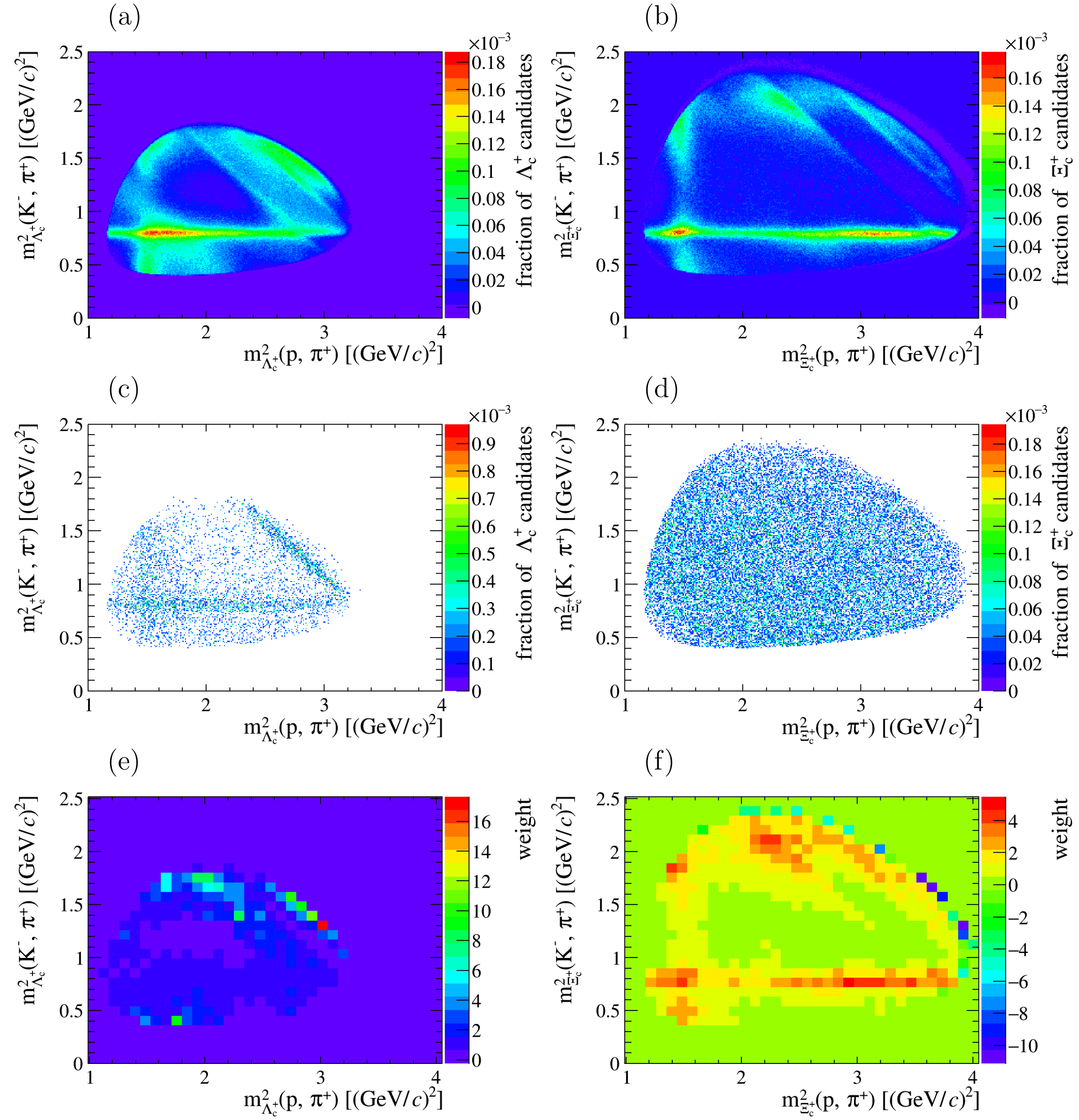


Figure 19: Dalitz plots for signal (left) $\Lambda_c^+ \to pK^-\pi^+$ and (right) $\Xi_c^+ \to pK^-\pi^+$ for (a & b) data, (c & d) simulation (without Dalitz weights applied), and (e & f) data divided by simulation (weight histograms).

Simulation is known not to produce the correct reconstruction efficiency for tracks in the detector. We therefore weight the reconstructed simulation using histograms generated by the `TrackCalib` package,[31] which uses an independent dataset to determine the tracking efficiency in the detector, according to the momentum and pseudorapidity of each track. We assign a systematic uncertainty from the use of this tool; see Section 4.7.4.

Simulation does not give a precise enough estimate of particle identification (PID) efficiencies for the final-state daughters; therefore, we use the `PIDCalib` package,[32] which uses an independent dataset to determine the PID efficiency of the detector, to estimate the PID response [59]. We do this in two steps: First, we use `PIDCalib` to reproduce the effects of the PID selections in Turbo (see Section 4.3.1) and produce per-event weights that represent the efficiency of the selections. Second, we replace the PID variables used in BDT training (see Table 9); we do this by sampling the PID response of the `PIDCalib` dataset as a function of the p_{T} and η of the simulated tracks and using this instead of the simulated response.[33] We assign systematic uncertainties to the PID efficiencies; see Section 4.7.3.

To avoid pollution of the signal sample with decays proceeding from secondary vertices, we ignore all simulated decays proceeding from a b quark, including $\mathit{\Lambda}_b^0$ baryons, B^- mesons, etc., by making cuts on the signal $\mathit{\Xi}_c^0$, $\mathit{\Lambda}_c^+$, or $\mathit{\Xi}_c^+$ baryon's ancestor (mother, grand mother, great-grand mother) particles' PID numbers. (We rely on the BDTs to remove decays from secondary vertices in data; see Sections 4.3.5 and 4.7.11.)

Unless otherwise stated, all simulation statistics referenced and figures shown have these selections, weights, and substitutions applied.

4.3.5 BDT selections

To better distinguish the signal from background, we train a boosted decision-tree (BDT) for each of the $\mathit{\Xi}_c^0 \to \pi^- \, (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays using the

---

[31] `https://twiki.cern.ch/twiki/bin/view/LHCb/TrackCalib`

[32] `https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage`

[33] `https://twiki.cern.ch/twiki/bin/view/LHCb/MeerkatPIDResampling`

Toolkit for Multivariate Data Analysis (TMVA) inside ROOT [60].[34] A BDT is a type of decision tree that compares the responses of several variables to determine how signal-like or background-like a given candidate is, allowing for improved discrimination between signal and background in a dataset, especially when compared to rectangular cuts like those applied above [56, 57]. As inputs, we give the BDT the parameters listed in Table 9 and apply the Turbo, offline, and trigger selections described above. The input distributions can be found in Appendix A.4.

Table 9: BDT input parameters. See text. $\mathrm{ProbNN}_{\mathrm{PID}}$ denotes the output of a probabilistic neural-network deciding how likely it is that a given track has been assigned the correct mass. The input distributions can be found in Appendix A.4.

$\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$
$\ln\left(\chi^2/\mathrm{ndf}(\mathrm{DTF}\,(\Xi_c^0 \to \pi^- \,(\Lambda_c^+ \to pK^-\pi^+)))\right)$
$\ln\left(\lVert\vec{x}(\Lambda_c^+) - \vec{x}(\Xi_c^0)\rVert\right)$
$\ln\left(\mathrm{DTF}\,(\tau\,(\Xi_c^0))\right)$
$\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(\Xi_c^0)\cdot\vec{p}(\Xi_c^0)}{x(\Xi_c^0)\cdot p(\Xi_c^0)}\right)\right)\right)$
$\ln\left(\arccos\left(\frac{\vec{x}(\Lambda_c^+)\cdot\vec{p}(\Lambda_c^+)}{x(\Lambda_c^+)\cdot p(\Lambda_c^+)}\right)\right)$
$\ln\left(\chi^2_{\mathrm{IP}}(\Xi_c^0)\right)$
$\ln\left(\chi^2_{\mathrm{IP}}(\Lambda_c^+)\right)$
$\ln\left(\chi^2_{\mathrm{IP}}(\pi^-)\right)$
$\mathrm{ProbNN}_{\pi,p,K,\pi}(\pi^-, p, K^-, \pi^+)$
$\Lambda_c^+ \to pK^-\pi^+$
$\ln\left(\chi^2/\mathrm{ndf}(\mathrm{DTF}\,(\Lambda_c^+ \to pK^-\pi^+))\right)$
$\ln\left(\mathrm{DTF}\,(\tau\,(\Lambda_c^+))\right)$
$\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(\Lambda_c^+)\cdot\vec{p}(\Lambda_c^+)}{x(\Lambda_c^+)\cdot p(\Lambda_c^+)}\right)\right)\right)$
$\ln\left(\chi^2_{\mathrm{IP}}(\Lambda_c^+)\right)$
$\mathrm{ProbNN}_{p,K,\pi}(p, K^-, \pi^+)$
$\Xi_c^+ \to pK^-\pi^+$
$\ln\left(\chi^2/\mathrm{ndf}(\mathrm{DTF}\,(\Xi_c^+ \to pK^-\pi^+))\right)$
$\ln\left(\mathrm{DTF}\,(\tau\,(\Xi_c^+))\right)$
$\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(\Xi_c^+)\cdot\vec{p}(\Xi_c^+)}{x(\Xi_c^+)\cdot p(\Xi_c^+)}\right)\right)\right)$
$\ln\left(\chi^2_{\mathrm{IP}}(\Xi_c^+)\right)$
$\mathrm{ProbNN}_{p,K,\pi}(p, K^-, \pi^+)$

---

[34] We use separate BDTs for $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$ and $\Lambda_c^+ \to pK^-\pi^+$ decays since the Λ_c^+ baryons in the former come from secondary vertices while those in the latter are promptly produced.

To train the $\varXi_c^0$ BDT, we use wrong-sign (WS) data, wherein the daughter $\varLambda_c^+$ baryon is paired with a π^+ meson instead of a π^- meson, as a background sample, whereas for the $\varLambda_c^+$ and $\varXi_c^+$ BDTs we use sideband data as background samples. (We choose to use WS data for the $\varXi_c^0$ BDT since most of the background comes from combinations of real $\varLambda_c^+$ baryons with random π mesons; see Appendix A.4.) We use truth-matched simulation as signal samples in all cases. We require

$$\left|\mathrm{DTF}\left(m\left(\varXi_c^0\right)\right) - \mathrm{DTF}\left(m\left(\varLambda_c^+\right)\right) + m_{\mathrm{PDG}}\left(\varLambda_c^+\right) - m_{\mathrm{PDG}}\left(\varXi_c^0\right)\right| < 5\,\mathrm{MeV} \tag{12}$$

for the $\varXi_c^0$ background sample since there are no resonances in this region,

$$\left|\mathrm{DTF}\left(m\left(\varLambda_c^+\right)\right) - m_{\mathrm{PDG}}\left(\varLambda_c^+\right)\right| > 40\,\mathrm{MeV} \tag{13}$$

for the $\varLambda_c^+$ background sample in order to avoid signal $\varLambda_c^+$ baryons, and

$$\left|\mathrm{DTF}\left(m\left(\varXi_c^+\right)\right) - m_{\mathrm{PDG}}\left(\varXi_c^+\right)\right| > 40\,\mathrm{MeV} \tag{14}$$

for the $\varXi_c^+$ background sample in order to avoid signal $\varXi_c^+$ baryons. To decrease training time, we use only a small fraction of the total data sample for background training. Figure 20 and Table 10 summarize the training inputs.

Due to the small size of the simulation sample surviving `L0HadronDecision`, we use the kFolding technique to maximize the amount of simulation available for BDT training and efficiency calculation.[35] We use 10 folds, where the fold assigned to a particular event is determined by the expression

$$\left[(2 * \texttt{EventNumber} + 3 * \texttt{RunNumber} + 5 * \texttt{nTracks} + 7 * \texttt{nSPDHits})/10\right] \%10, \tag{15}$$

---

[35] See, e.g., Section 3.2 in `https://root.cern.ch/root/htmldoc/guides/tmva/TMVAUsersGuide.pdf` for a description of the kFolding technique.

where `EventNumber` references the event within a given run wherein the candidate was generated, `RunNumber` references the run, `nTracks` is the total number of tracks in the event, `nSPDHits` is the number of hits in the Scintillating Pad Detector, and % is the modulo operator. Each fold is used as the testing sample for a BDT trained on the events in the other 9 folds. Thus we actually train 10 BDTs per signal charm-baryon and use the aggregate response to discriminate signal from background.

Table 10: Number of candidates used for BDT training and testing. The choice to use 99,999 background candidates is arbitrary; we avoid using the whole dataset in order to speed up processing time. All candidates have passed the Turbo, offline, and trigger selections. None of the weights described in Section 4.3.4 have been applied; these are raw candidate numbers.

Decay	Signal	Background
Ξ_c^0	4,833	99,999
Λ_c^+	5,406	99,999
Ξ_c^+	39,644	99,999

4.3.6 BDT performance

We find all BDTs discriminate signal from background with acceptable efficiency and without overtraining, as shown in Figure 21 and Appendix A.2.[36] The Ξ_c^0 BDT performs rather worse than its counterparts; this is because there are real Λ_c^+ baryons in both its signal and its background samples, which it must distinguish from fake Λ_c^+ baryons in addition to ensuring the candidates vertex with π^- mesons.

4.3.7 BDT cut selection

To determine the BDT value to select, we construct a figure of merit (FoM)

$$\mathrm{FoM}(x) = \frac{\epsilon_{\mathrm{signal}}}{\sqrt{N}}, \tag{16}$$

---

[36] We trained additional BDTs with different hyperparameters, selecting a final BDT based on its performance in the overtraining check.

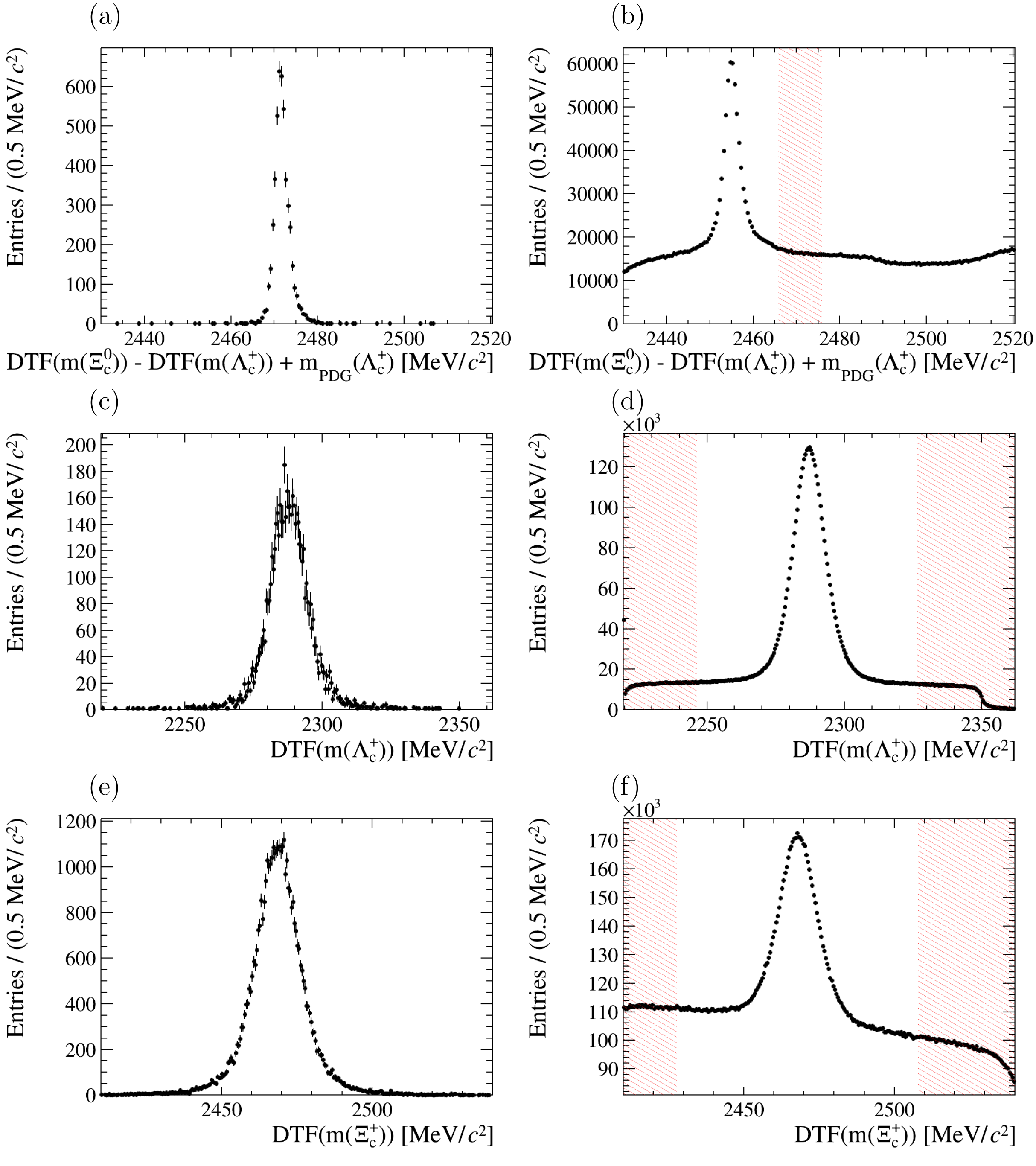


Figure 20: Plots of mass distributions of candidates used for BDT training for (a & b) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$, (c & d) $\Lambda_c^+ \to pK^-\pi^+$, and (e & f) $\Xi_c^+ \to pK^-\pi^+$ decays for (left) signal simulation, for which we apply truth-matching and make no mass selections, and (right) background data, defined by the hatched red regions. The (b) Ξ_c^0 data sample shown is WS (a Λ_c^+ baryon paired with a π^+ meson), while all other samples shown use daughters with the correct charges. Note that the data distributions shown here represent the entire datasets; we use only a small fraction to train the BDTs (see Table 10).

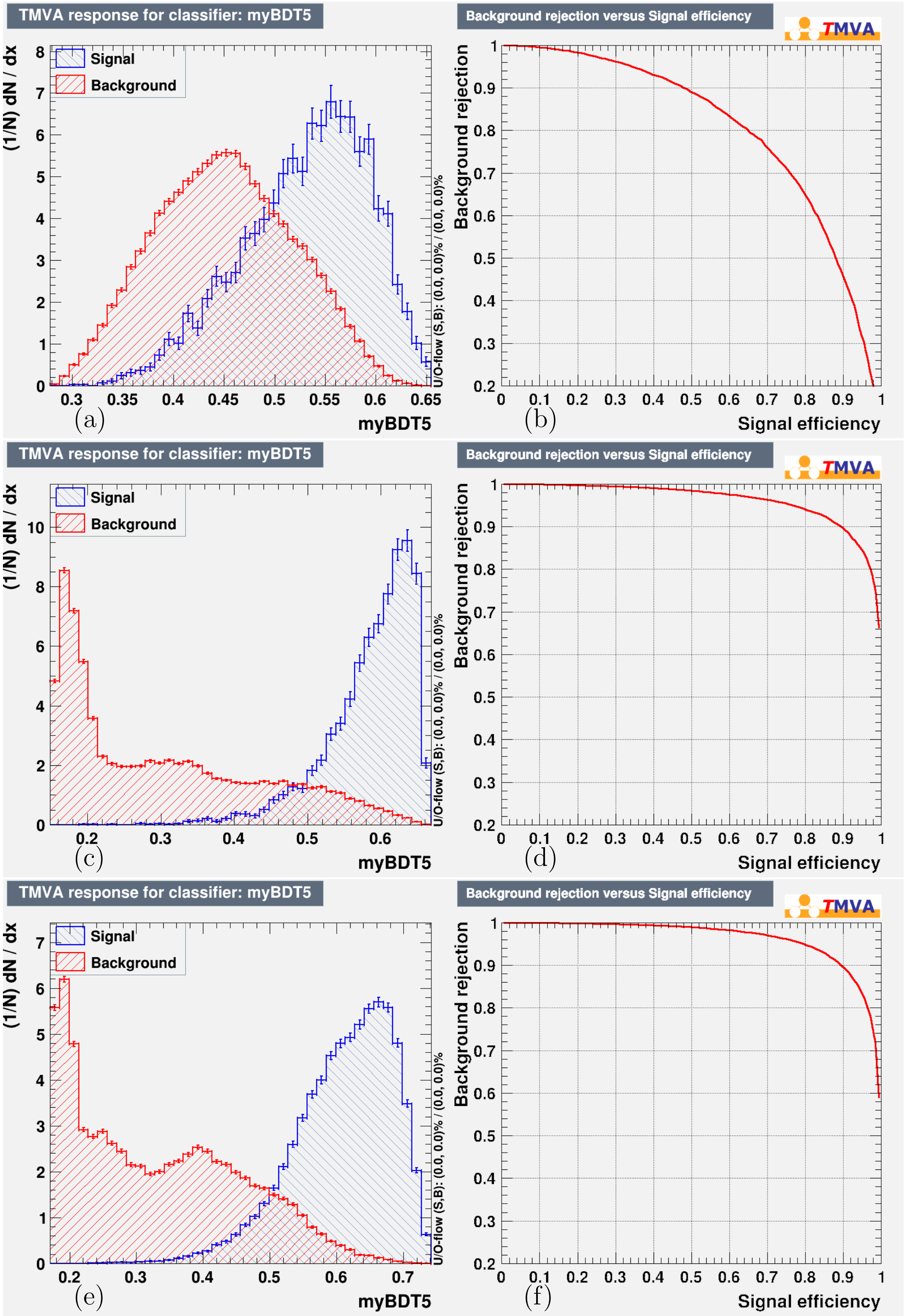


Figure 21: BDT performance, as demonstrated by (left) its response and (right) its receiver operating characteristic (ROC) curve for (a & b) $\Xi_c^0 \rightarrow \pi^- \ (\Lambda_c^+ \rightarrow pK^-\pi^+)$, (c & d) $\Lambda_c^+ \rightarrow pK^-\pi^+$, and (e & f) $\Xi_c^+ \rightarrow pK^-\pi^+$.

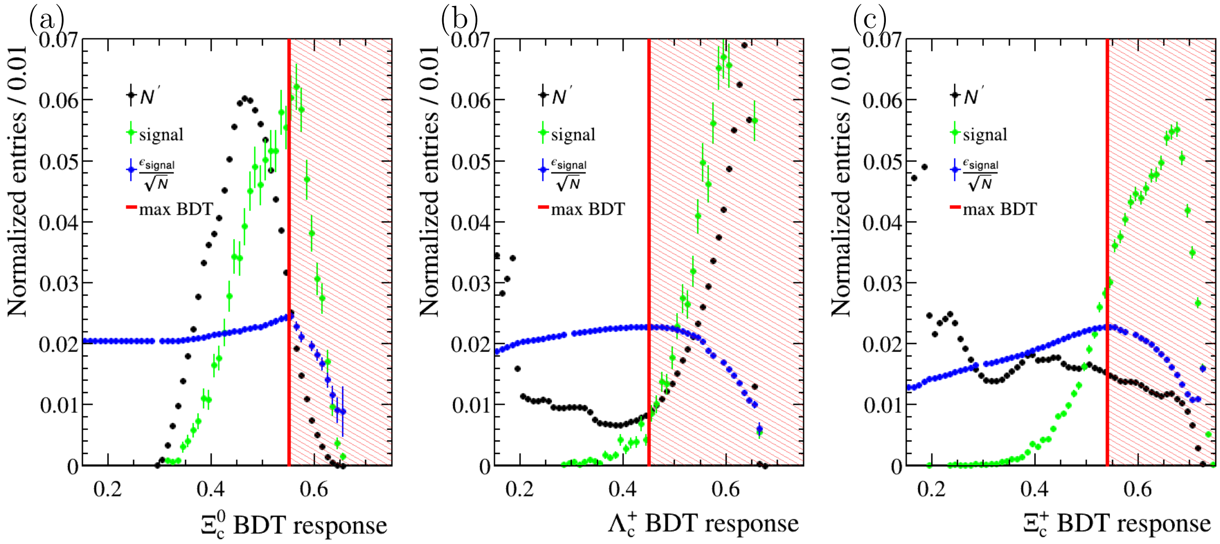


Figure 22: BDT selection figure of merit for (a) $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$, (b) $\Lambda_c^+ \to pK^-\pi^+$, (c) and $\Xi_c^+ \to pK^-\pi^+$ samples. See Section 4.3.7. The black histograms represent the fraction of candidates in data. The green histograms represent the fraction of signal candidates in simulation. The blue curves represent the normalized values of $\frac{\epsilon_{\mathrm{signal}}}{\sqrt{N}}$ when the BDT response is $> x$. We select the red highlighted regions.

where $\epsilon_{\mathrm{signal}}$ is efficiency of the BDT $> x$ selection relative to the Turbo, trigger, and offline selections; N is the number of data entries surviving BDT $> x$; and x varies from 0.15 to 0.75 in steps of 0.01. In computing the FoM, we consider only candidates whose mass falls within a certain range of its PDG value: 4 MeV for the Ξ_c^0, 20 MeV for the Λ_c^+, or 20 MeV for the Ξ_c^+ baryon, ≈ 2.5 times their fitted widths (see Section 4.5). We do not apply any weighting to the intermediate spectra of the Λ_c^+ or Ξ_c^+ baryons. Figure 22 summarizes the results. We select the values of x that maximize the FoM, leading to minimum allowed values of 0.55, 0.45, and 0.54 for the Ξ_c^0, Λ_c^+, and Ξ_c^+ datasets, respectively.

4.4 Efficiencies

We determine the efficiencies of the non-trigger selections using simulation. The total efficiency of all selections, across both 2017 and 2018, is summarized in Table 11.

We use the data-driven TISTOS method [61] to determine the efficiency of the trigger selections.[37] We apply the Turbo, offline, and BDT selections to the data sample, then select candidates in events that pass the trigger selections independently of the signal (TIS) and fit the mass distributions of those candidates to determine the signal yields. We then apply the trigger selections to those same events (TISTOS) and repeat the fits. The ratio of the number of candidates passing both TIS and TOS trigger selections (TISTOS) to those passing TIS selections is equivalent to the trigger selection efficiency. Sections 4.5.1 and 4.5.2 describe the fit models and summarize yields, and Section 4.4.1 describes the treatment of the correlation between the TIS and TISTOS samples.

Table 2 shows that the $\Xi_c^+ \to pK^-\pi^+$ simulation sample uses `DaughtersInLHCb` generator-level cuts, which requires final-state daughters to be within the coverage of the LHCb detector, while the other samples use `LHCbAcceptance`, which requires only the charm baryon to be within coverage. This means that the reconstruction efficiency (incorporated into "Turbo" in Table 11) given by the $\Xi_c^+ \to pK^-\pi^+$ simulation sample does not account for signal loss due to final-state daughters falling outside the detector acceptance, while the reconstruction efficiencies given by the $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ and $\Lambda_c^+ \to pK^-\pi^+$ simulation samples do. We account for this difference by simulating an additional sample of 100,000 $\Xi_c^+ \to pK^-\pi^+$ events at the generator level (Gauss stage only) with `LHCbAcceptance` cuts, which we apply again offline to ensure every candidate satisfies them, and then applying `DaughtersInLHCb` offline to each candidate. The ratio of the number of candidates surviving `DaughtersInLHCb` to the number surviving `LHCbAcceptance` is given in the "`DInLHCb`" column of Table 11. We use Gauss v49r12 to generate the events and DaVinci v44r8 to construct useable n-tuples.

---

[37] Our approach follows that described in Reference [61], but in this case, $N_{Trig} = N_{\mathrm{TOS}}$ (see Section 4.3.2), and since we are interested in c-quarks, which are much lighter than b-quarks, we treat "the rest of the event" as independent of the signal candidates.

Ghosts, reconstructed tracks that cannot be associated to generated particles, cannot be truth-matched but may still contain signal candidates. Since the truth-matching criteria necessarily exclude them (see Section 4.3.4), we correct the signal efficiency by the increase in signal associated with including ghosts. We do this by comparing the signal yields in fits to the signal mass spectra in truth-matched simulation samples to the signal yields in the combined truth-matched-plus-ghosts samples; Section 4.5.1 describes the fit models we use to extract the yields; both samples have the Turbo, offline, and BDT selections applied; and we do not apply any of the weights described in Section 4.3.4 to the datasets used in these fits. The correction factor is equal to the signal yield in the truth-matched-plus-ghosts sample divided by that in the truth-matched sample. Table 11 summarizes the corrections in the "Ghost fix" column, and Section 4.4.1 describes the treatment of the correlation between the truth-matched-plus-ghosts and truth-matched samples.

Table 11: Selection efficiencies, in the range described in Section 4.4.2, only statistical uncertainties shown. Turbo and `DaughtersInLHCb` are given relative to generator-level. Offline and BDT are each given relative to the previous one (left-to-right). "Ghost fix" is the efficiency adjustment for ghost tracks (see text). Trigger is given relative to BDT and is determined using the data-driven TISTOS method (see text), while all others are determined using simulation. Total is the product of all others with correlations taken into account. A factor of 10^{-1} is shown in the Total Λ_c^+ efficiency to account for the retention of only 10% of the candidates in data (see Section 4.3.1).

%	`DInLHCb`	Turbo	Offline	BDT	Ghost fix	Trigger
Ξ_c^0	–	0.96 ± 0.01	93.7 ± 2.0	32.2 ± 1.0	105.2 ± 2.0	35.6 ± 4.4
Λ_c^+	–	1.05 ± 0.01	99.0 ± 2.0	95.4 ± 1.9	105.9 ± 0.3	33.4 ± 0.1
Ξ_c^+	97.9 ± 2.1	5.08 ± 0.03	99.2 ± 0.7	80.6 ± 0.6	104.2 ± 0.1	28.4 ± 0.1

%	Total
Ξ_c^0	0.11 ± 0.01
Λ_c^+	$(0.35 \pm 0.01)/10$
Ξ_c^+	1.18 ± 0.03

4.4.1 Correlations between datasets

As described in Reference [61], the TIS and TISTOS samples are not independent, since the TIS set necessarily includes the events in the TISTOS set. Using the notation in Reference [61],

$$N_{\mathrm{TISTOS}} = d \tag{17}$$

$$N_{\mathrm{TIS}} = c + d \tag{18}$$

$$\epsilon_{Trig} = \frac{N_{\mathrm{TISTOS}}}{N_{\mathrm{TIS}}} = \frac{d}{c+d}. \tag{19}$$

The covariance between N_{TISTOS} and N_{TIS}, $\mathrm{cov}(N_{\mathrm{TISTOS}}, N_{\mathrm{TIS}})$, is thus equivalent to $\mathrm{cov}(d, c+d)$ and, from the distributive property of covariance,

$$\mathrm{cov}(d, c+d) = \mathrm{cov}(d, c) + \mathrm{cov}(d, d) \tag{20}$$

$$= 0 + \sigma_d^2 = \sigma_d^2, \tag{21}$$

since c and d are independent samples. So,

$$\mathrm{cov}(N_{\mathrm{TISTOS}}, N_{\mathrm{TIS}}) = \sigma^2_{N_{\mathrm{TISTOS}}}. \tag{22}$$

A similar consideration applies to the covariance between N_{TOS} and each of N_{TIS} and N_{TISTOS}, but we treat N_{TOS} as independent since its statistics are so much larger; recalculation of the results taking this covariance into account confirms that the difference is negligible.

The same consideration applies to the covariance between the truth-matched and truth-matched-plus-ghosts simulation samples; the calculation is the same as above, but set $d = N_{\mathrm{truth-matched}}$ and $c + d = N_{\mathrm{truth-matched-plus-ghosts}}$. Thus,

$$\mathrm{cov}(N_{\mathrm{truth-matched}}, N_{\mathrm{truth-matched-plus-ghosts}}) = \sigma^2_{N_{\mathrm{truth-matched}}}. \tag{23}$$

We take all these covariances into account using the Uncertainties package in python [62].

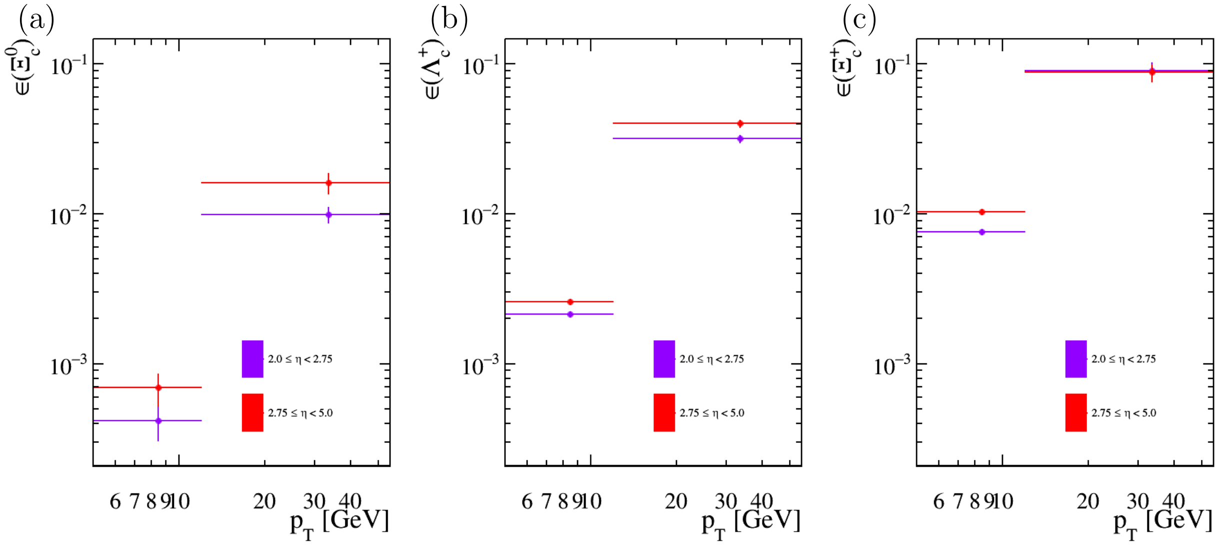


Figure 23: Efficiency as a function of p_{T} and η for (a) $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$, (b) $\Lambda_c^+ \to pK^-\pi^+$, and (c) $\Xi_c^+ \to pK^-\pi^+$ decays.

4.4.2 Binning in p_{T} and η

The overall efficiency is not flat in either the transverse momentum (p_{T}) or pseudorapidity (η) of the signal Ξ_c^0, Λ_c^+, and Ξ_c^+ baryons, and we calculate the efficiency in two dimensions: $\epsilon(p_{\mathrm{T}}, \eta)$. We use simulation to calculate the non-trigger efficiency by taking the ratio in each bin of the number of signal candidates surviving all non-trigger selections to the number of signal particles generated. We use data to calculate the trigger efficiency by taking the ratio in each bin of the signal yield surviving TISTOS trigger selections to that surviving TIS trigger selections, as described above. The total efficiency is the product of the trigger and non-trigger efficiencies. Figure 23 shows the total efficiency in each bin.

The binning is determined by ensuring that there are sufficient entries in each bin to perform satisfactory fits to data (see Section 4.5) and such that enough simulation events in each bin (≥ 10) survive all the selections to maintain sufficient statistics for the errors on the efficiencies.

4.5 Signal yield

We extract signal yields (without efficiency correction) from the data distributions by performing unbinned, extended maximum log-likelihood fits to relevant invariant mass spectra using the RooFit package within Root [63]. All fits use MIGRAD to perform the minimization and use a sum-of-weights-squared (`SumW2`) strategy to calculate their errors in HESSE.

4.5.1 Fit model

In this section, we present the models used to fit the data distributions. To calculate the results, we do the fits in regions of p_{T} and η corresponding to the binning used to calculate the efficiencies described in Section 4.4.2; Appendix A.6 shows these fits, and Section 4.5.2 describes the results. Figures 24, 25, and 26 show the fits to the aggregate (TOS) datasets, in which parameter values are allowed to float (within certain limits). For the fits in each region of p_{T} and η, however, we constrain the signal parameters to Gaussian distributions with means and widths equal to the values and errors, respectively, of the parameters determined by the fits to the aggregate datasets.

Figure 24 shows $\mathrm{DTF}\left(m\left(\mathit{\Xi}_c^0\right)\right) - \mathrm{DTF}\left(m\left(\mathit{\Lambda}_c^+\right)\right) + m_{\mathrm{PDG}}\left(\mathit{\Lambda}_c^+\right)$ for the $\mathit{\Xi}_c^0$ data dataset fitted with a Crystal Ball signal shape and a 1st-order Chebychev polynomial background shape.[38] We fix the α and n parameters to -14.9 and 22, respectively, to encourage convergence; see Section 4.7.6 for the systematic uncertainties we assign to this choice. We find 6316 ± 229 $\mathit{\Xi}_c^0 \to \pi^- \left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)$ signal candidates.

Figure 25 shows $\mathrm{DTF}\left(m\left(pK^-\pi^+\right)\right)$ for the $\mathit{\Lambda}_c^+$ data dataset fitted with a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background.[39] We fix the fraction of the signal yield assigned to the double-sided Crystal Ball shape to 0.58 and fix the α parameter on both sides to 1.85 to enable the fit to converge; see Section 4.7.7 for the systematic uncertainties we assign to this choice. We find 2667193 ± 3279

---

[38] Reference [64] describes the Crystal Ball shape.

[39] A double-sided Crystal Ball is modified to have both low- and high-mass power-law tails; see footnote 38.

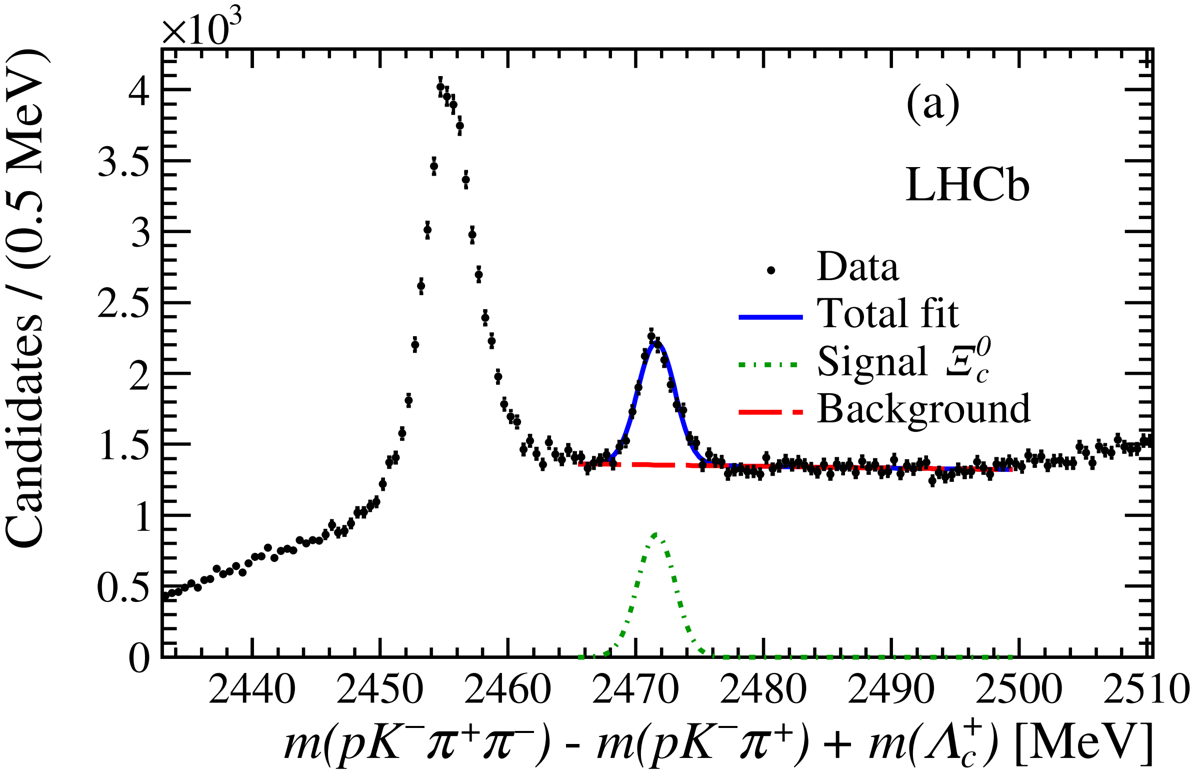


Figure 24: Fit to DTF $\left(m\left(\Xi_c^0\right)\right)$ $-$DTF $\left(m\left(\Lambda_c^+\right)\right)$ $+m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ in the Ξ_c^0 data dataset using a Crystal Ball signal shape and a 1st-order Chebychev polynomial background shape; see text.

$\Lambda_c^+ \to pK^-\pi^+$ signal candidates; recall that we prescale the $\Lambda_c^+ \to pK^-\pi^+$ sample by a factor of 10 (see Section 4.3.1).

Figure 26 shows DTF $\left(m\left(pK^-\pi^+\right)\right)$ for the Ξ_c^+ data dataset fitted with the same basic shapes used for the Λ_c^+ data dataset described above. We find 1613027 ± 3544 $\Xi_c^+ \to pK^-\pi^+$ signal candidates and assign systematic uncertainties described in Section 4.7.8.

We also perform fits to determine the TIS and TISTOS (see Section 4.4) yields in data, using the same approach and shapes described above. We find 876 ± 100 TIS and 312 ± 55 TISTOS $\Xi_c^0 \to \pi^-\left(\Lambda_c^+ \to pK^-\pi^+\right)$ signal candidates, 431239 ± 909 TIS and 144153 ± 568 TISTOS $\Lambda_c^+ \to pK^-\pi^+$ signal TIS candidates (recall that we prescale the $\Lambda_c^+ \to pK^-\pi^+$ sample by a factor of 10 [see Section 4.3.1]), and 249895 ± 1272 TIS and 70983 ± 817 TISTOS $\Xi_c^+ \to pK^-\pi^+$ signal candidates. Figures 27, 28, and 29 show the fits.

We also perform fits to determine the ghost correction factors for the efficiencies (see

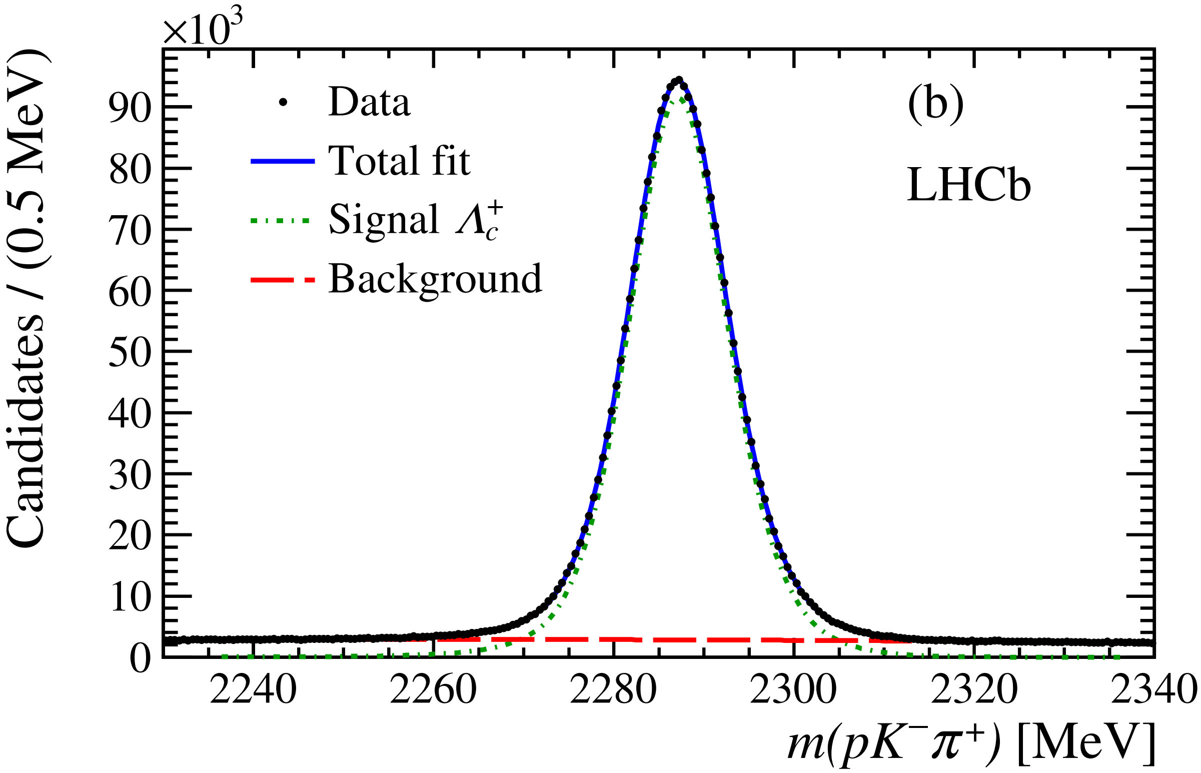


Figure 25: Fit to DTF $(m\,(pK^-\pi^+))$ in the Λ_c^+ data dataset using a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text.

Section 4.4). We use the same approach described above: floating parameters for overall fits, followed by fits in each region of p_{T} and η with constrained parameters, but we use different fit shapes, described below. These datasets have the Turbo, offline, and BDT selections applied (no trigger), and we do not apply any of the weights described in Section 4.3.4.

Figure 30 shows DTF $(m\,(\Xi_c^0))$ $-$DTF $(m\,(\Lambda_c^+))$ $+m_{\mathrm{PDG}}\,(\Lambda_c^+)$ for the Ξ_c^0 simulation dataset with truth-matched or truth-matched-plus-ghosts applied, fitted with a Crystal Ball signal shape and a flat background shape. We fix the value of n in the Crystal Ball shape to 50 to enable the fit to converge; this is the only parameter we fix. We find $2,639 \pm 16$ and $2,775 \pm 52$ signal candidates in the truth-matched and truth-matched-plus-ghosts datasets, respectively.

Figure 31 shows DTF $(m\,(pK^-\pi^+))$ for the Λ_c^+ simulation dataset with truth-matched or truth-matched-plus-ghosts applied, fitted with a double-Gaussian signal shape (both

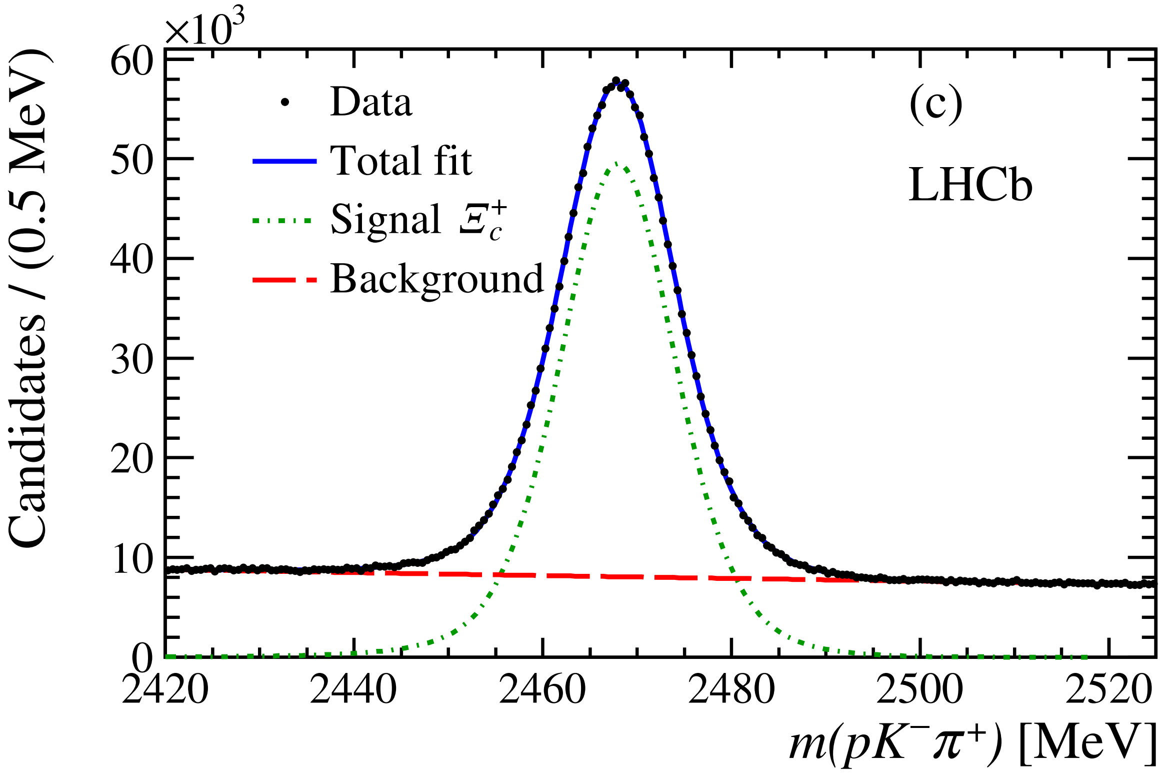


Figure 26: Fit to DTF $(m\,(pK^-\pi^+))$ in the Ξ_c^+ data dataset using a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text.

Gaussians are constrained to have the same mean) and a flat background shape. We find 8245±94 and 8736±100 signal candidates in the truth-matched and truth-matched-plus-ghosts datasets, respectively.

Figure 32 shows DTF $(m\,(pK^-\pi^+))$ for the Ξ_c^+ simulation dataset with truth-matched or truth-matched-plus-ghosts applied, fitted with a double-Gaussian signal shape (both Gaussians are constrained to have the same mean) and a flat background shape. We find 59342 ± 252 and 61810 ± 261 signal candidates in the truth-matched and truth-matched-plus-ghosts datasets, respectively.

4.5.2 Fit results

We perform fits in ranges of p_{T} and η using the models discussed in Section 4.5.1 and extract signal yields. Figure 33 shows the efficiency-corrected results, and the individual fits can be

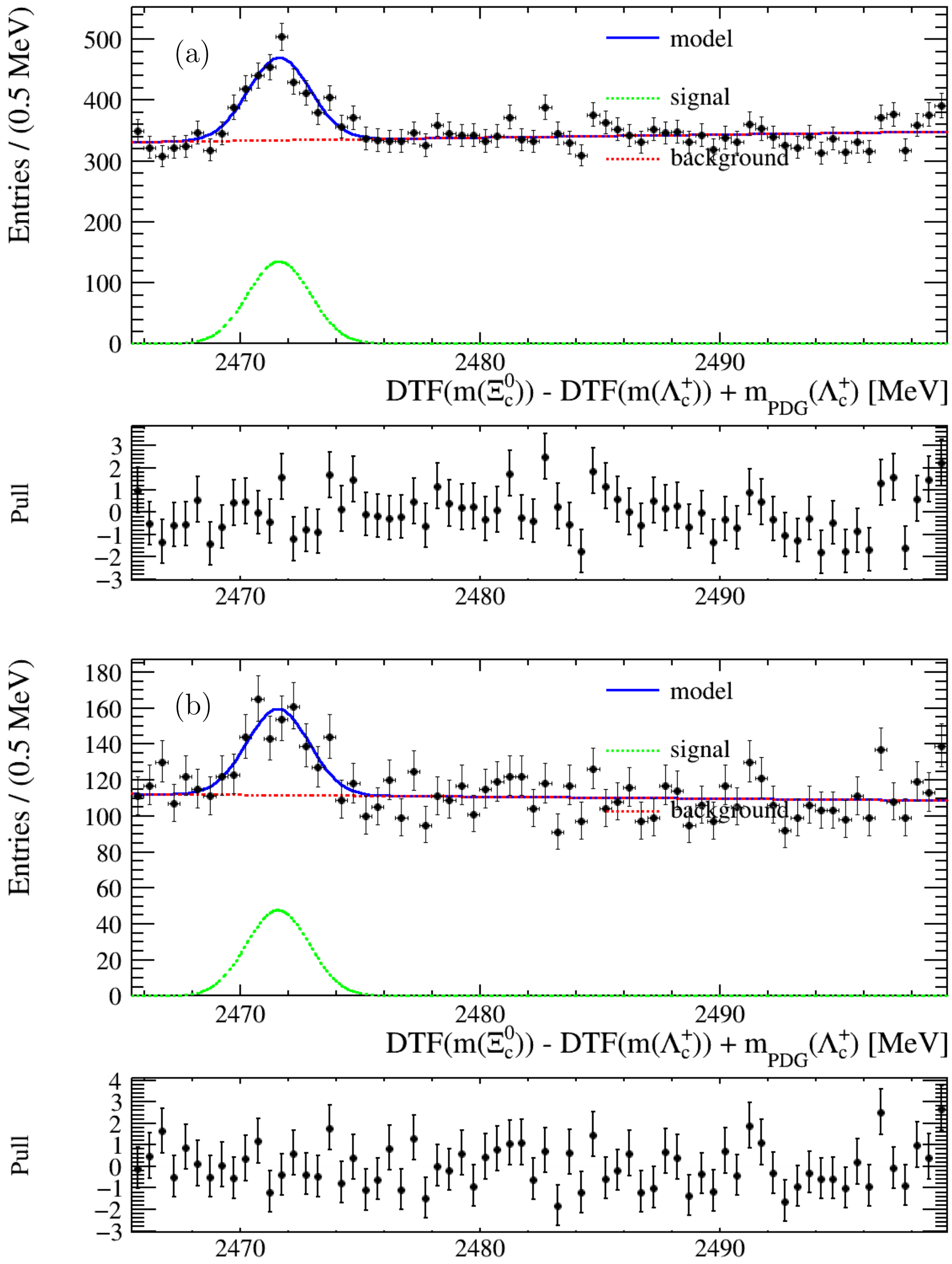


Figure 27: DTF $\left(m\left(\Xi_c^0\right)\right)$ $-$DTF $\left(m\left(\Lambda_c^+\right)\right)$ $+m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ distribution from the Ξ_c^0 data (a) TIS and (b) TISTOS datasets fit to a Crystal Ball signal shape and a 1st-order Chebychev polynomial background shape; see text.

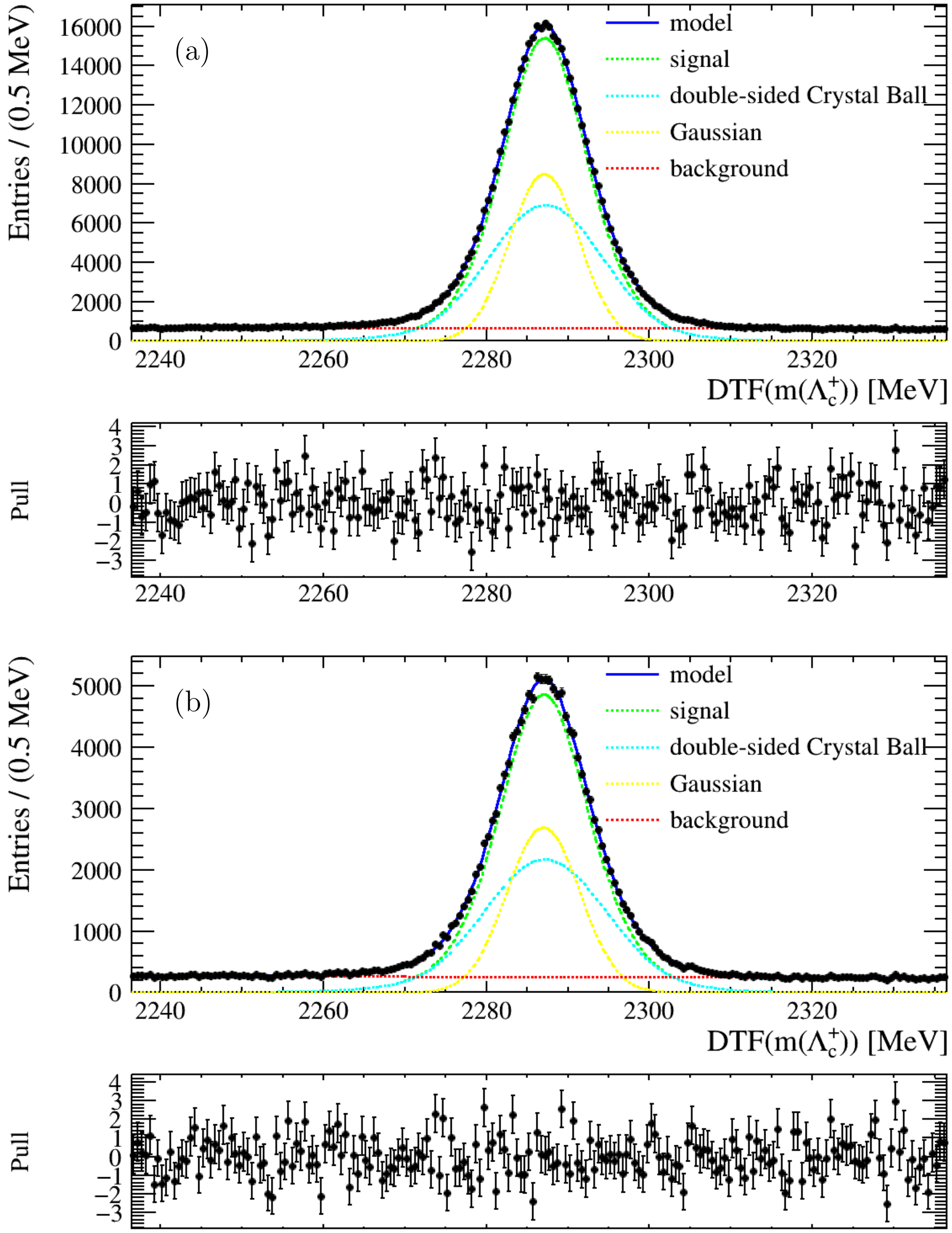


Figure 28: DTF mass distribution of $pK^-\pi^+$ from the $\Lambda_c^+ \to pK^-\pi^+$ data (a) TIS and (b) TISTOS datasets fit to a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text.

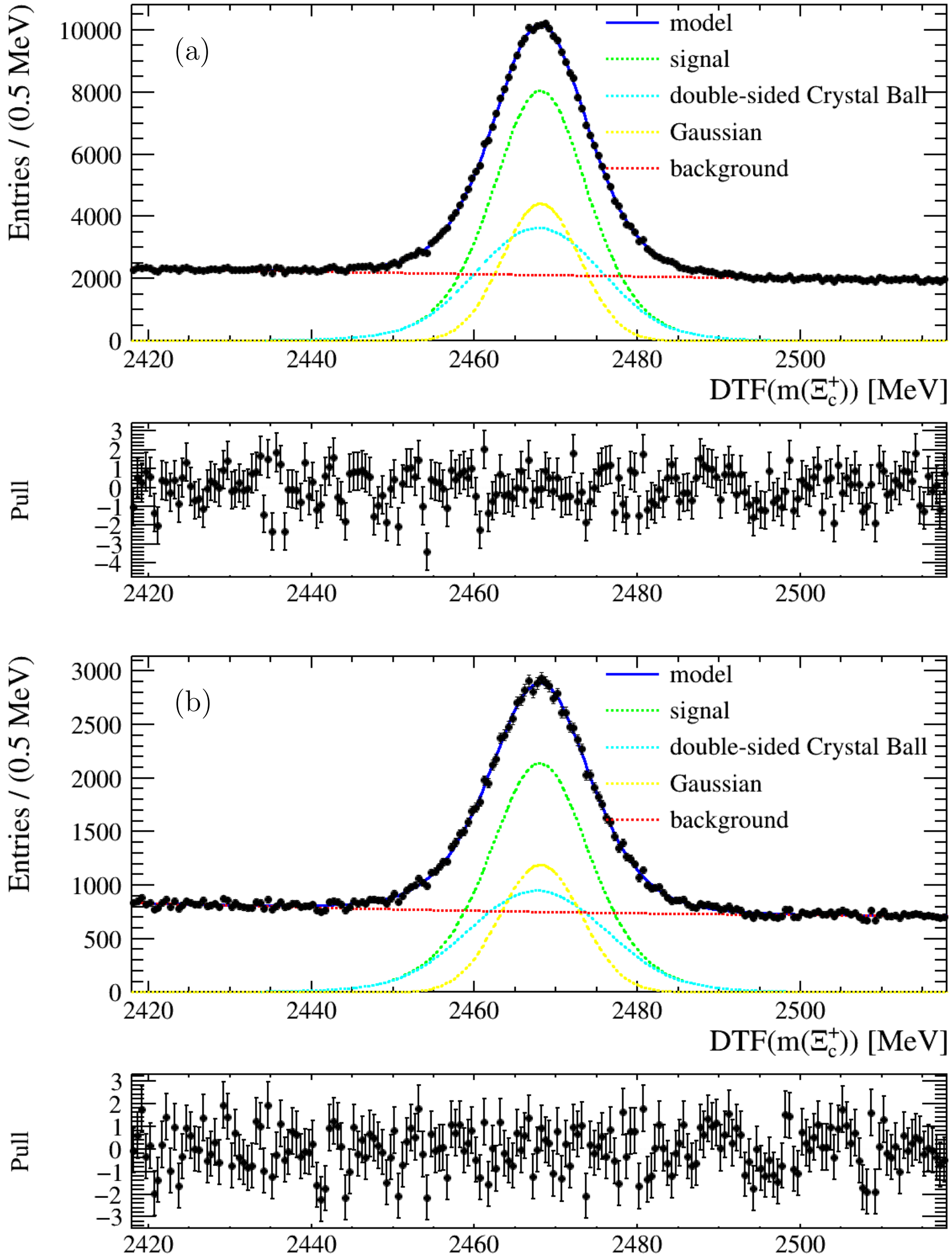


Figure 29: DTF mass distribution of $pK^-\pi^+$ from the $\Xi_c^+ \to pK^-\pi^+$ data (a) TIS and (b) TISTOS datasets fit to a double-sided Crystal Ball added to a Gaussian signal shape and a 2nd-order Chebychev polynomial background shape; see text.

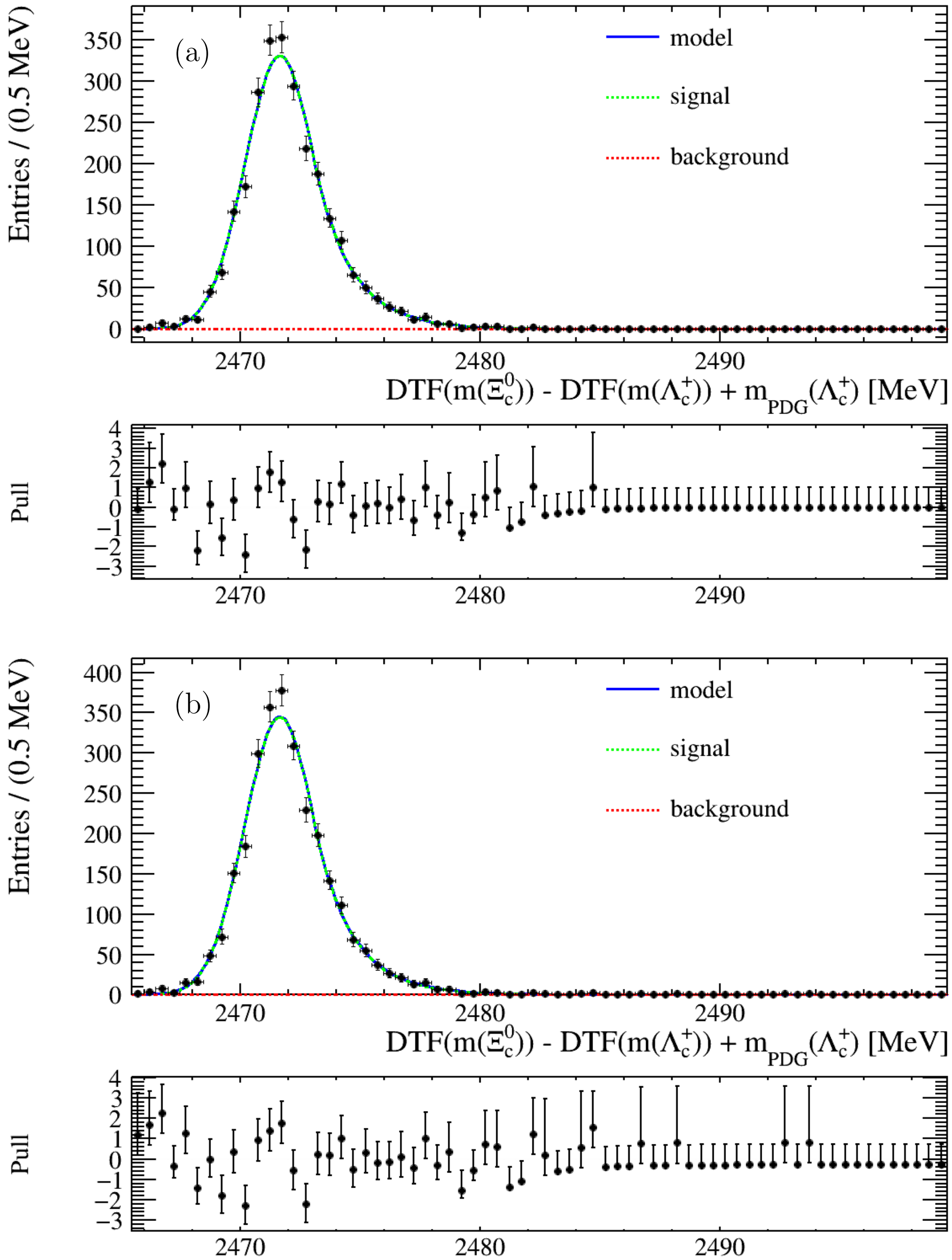


Figure 30: DTF $\left(m\left(\Xi_c^0\right)\right)$ $-$DTF $\left(m\left(\Lambda_c^+\right)\right)$ $+m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ distribution from the Ξ_c^0 simulation (a) truth-matched and (b) truth-matched-plus-ghosts datasets fit to a Crystal Ball signal shape and a flat background shape; see text.

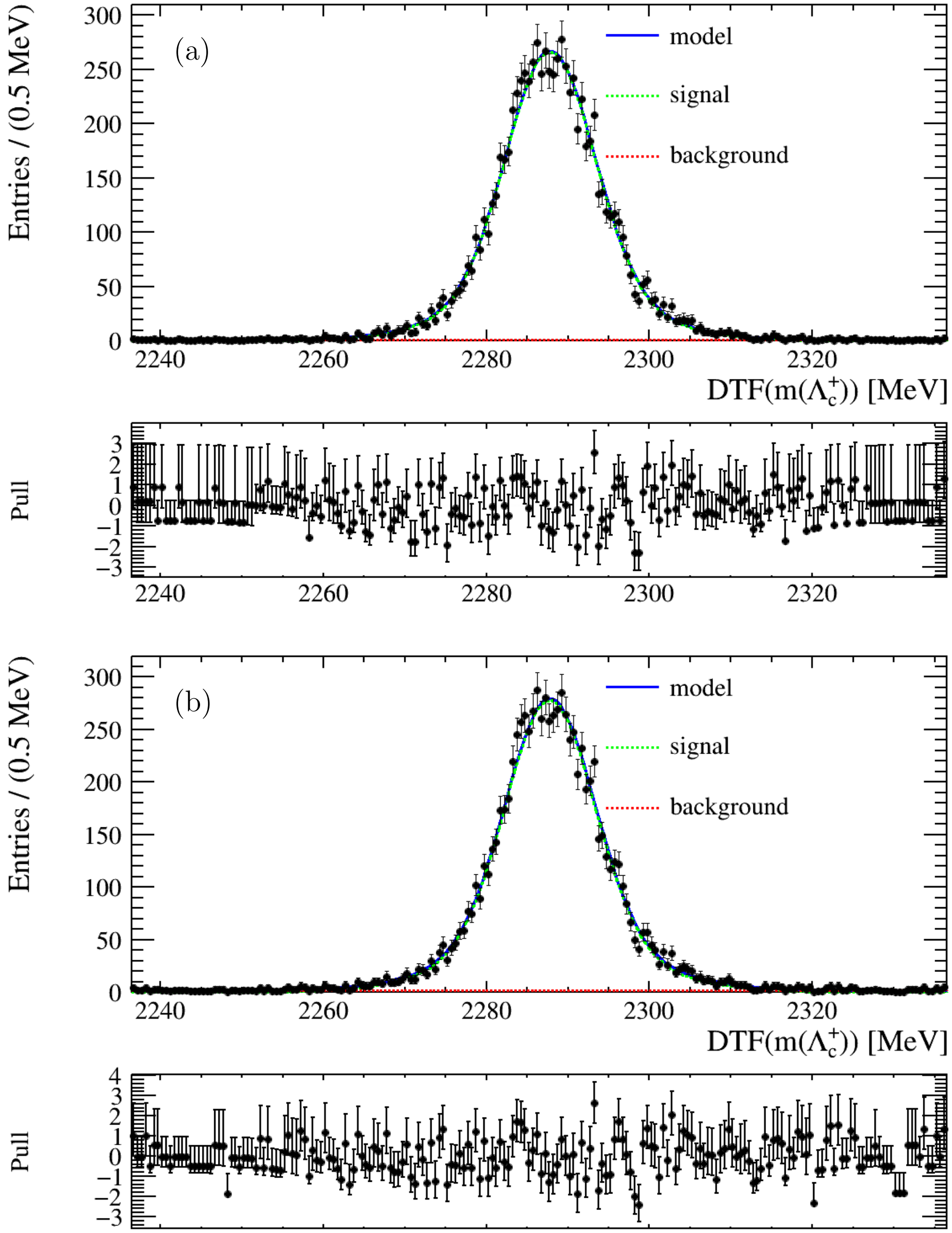


Figure 31: DTF mass distribution of $pK^-\pi^+$ from the $\Lambda_c^+ \to pK^-\pi^+$ simulation (a) truth-matched and (b) truth-matched-plus-ghosts datasets fit to a double-Gaussian signal shape (both Gaussians are constrained to have the same mean) and a flat background shape; see text.

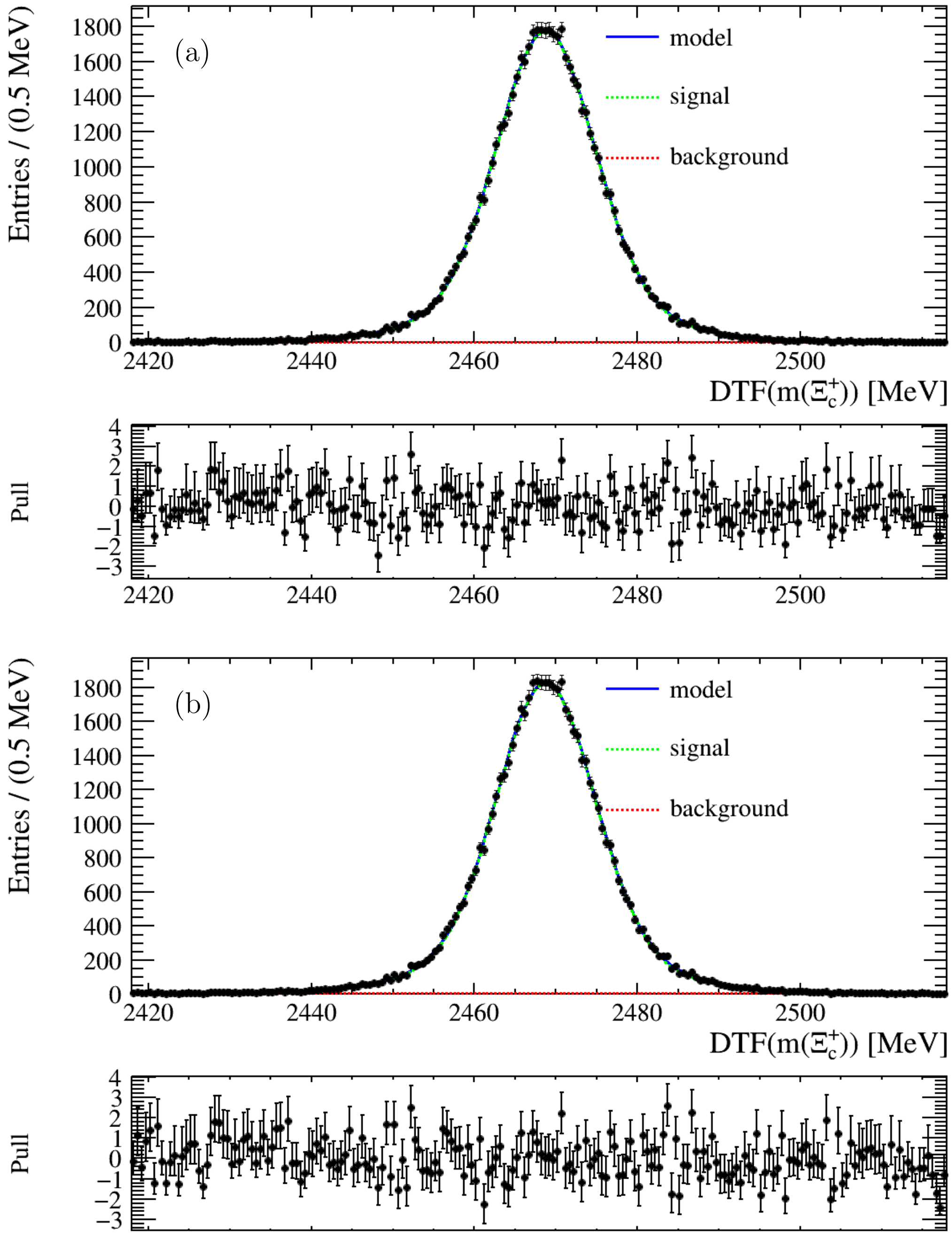


Figure 32: DTF mass distribution of $pK^-\pi^+$ from the $\Xi_c^+ \to pK^-\pi^+$ simulation (a) truth-matched and (b) truth-matched-plus-ghosts datasets fit to a double-Gaussian signal shape (both Gaussians are constrained to have the same mean) and a flat background shape; see text.

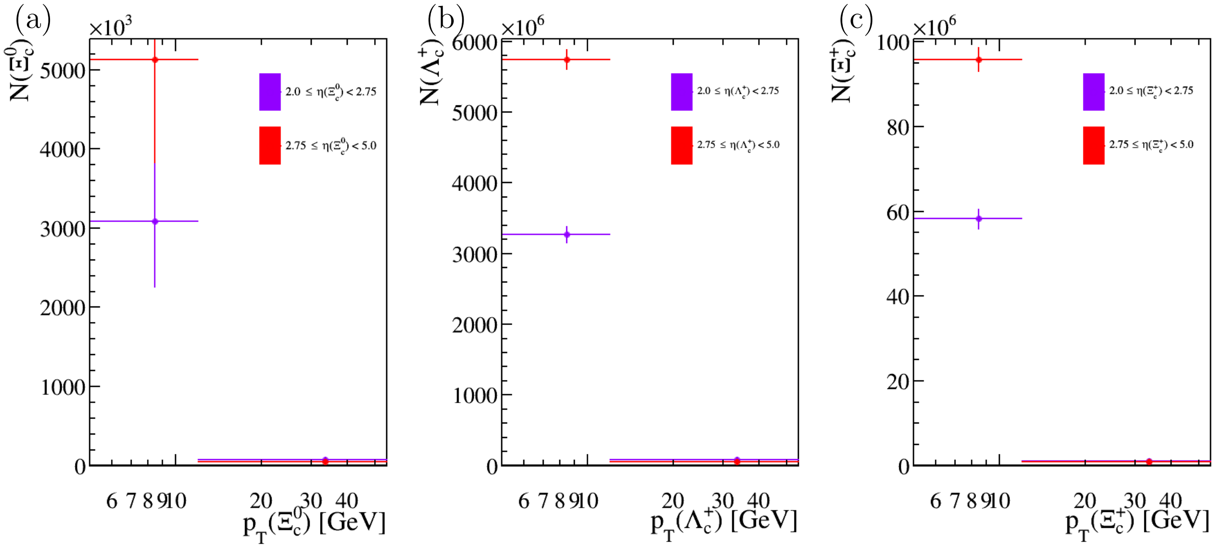


Figure 33: Efficiency-corrected yields as functions of p_{T} and η for (a) $\mathit{\Xi}_c^0 \to \pi^- \, (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, (b) $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and (c) $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays. Only the statistical errors (from the fits and efficiencies) are shown here. The $\mathit{\Lambda}_c^+$ yields have been corrected to account for the 10% prescale.

found in Appendix A.6.

4.6 Results

4.6.1 Intrinsic results

The intrinsic results are the ratios of efficiency-corrected yields in $\mathit{\Xi}_c^0 \to \pi^- \, (\mathit{\Lambda}_c^+ \to pK^-\pi^+)$, $\mathit{\Lambda}_c^+ \to pK^-\pi^+$, and $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays (see Section 4.1). These are equivalent to the determination of

$$
\begin{aligned}
\mathcal{R}_1 &\equiv \frac{f_{\mathit{\Xi}_c^0}}{f_{\mathit{\Lambda}_c^+}} \times \mathcal{B}(\mathit{\Xi}_c^0 \to \pi^- \mathit{\Lambda}_c^+) \\
&= \frac{N(\mathit{\Xi}_c^0 \to \pi^-\, (\mathit{\Lambda}_c^+ \to pK^-\pi^+))}{N(\mathit{\Lambda}_c^+ \to pK^-\pi^+)} \\
&= (0.095 \pm 0.003 \pm 0.012)\%, \qquad (24) \\
\mathcal{R}_2 &\equiv \frac{f_{\mathit{\Xi}_c^0}}{f_{\mathit{\Xi}_c^+}} \times \frac{\mathcal{B}(\mathit{\Lambda}_c^+ \to pK^-\pi^+)}{\mathcal{B}(\mathit{\Xi}_c^+ \to pK^-\pi^+)} \times \mathcal{B}(\mathit{\Xi}_c^0 \to \pi^- \mathit{\Lambda}_c^+) \\
&= \frac{N(\mathit{\Xi}_c^0 \to \pi^-\, (\mathit{\Lambda}_c^+ \to pK^-\pi^+))}{N(\mathit{\Xi}_c^+ \to pK^-\pi^+)} \\
&= (5.70 \pm 0.19 \pm 0.77)\%, \qquad (25) \\
\text{and } \mathcal{R}_3 &\equiv \frac{f_{\mathit{\Xi}_c^+}}{f_{\mathit{\Lambda}_c^+}} \times \frac{\mathcal{B}(\mathit{\Xi}_c^+ \to pK^-\pi^+)}{\mathcal{B}(\mathit{\Lambda}_c^+ \to pK^-\pi^+)} \\
&= \frac{N(\mathit{\Xi}_c^+ \to pK^-\pi^+)}{N(\mathit{\Lambda}_c^+ \to pK^-\pi^+)} \\
&= (1.753 \pm 0.003 \pm 0.107)\%, \qquad (26)
\end{aligned}
$$

where the first uncertainties are statistical and the second systematic. (All uncertainties, both statistical and systematic, and their correlations have been calculated using a combination of the ROOT Data Analysis Framework and the Uncertainties package in python [62, 65].) We derive the values by evaluating the rates in each bin of p_{T} and η, taking the inverse-variance weighted mean,[40] multiplying by a relative-luminosity correction factor described in Section 4.7.12, and multiplying by the secondary-decay correction factors described in Section 4.7.11. Table 12 summarizes the values used in these calculations. Their correlation matrix is

---

[40] The inverse-variance weighted mean is an analysis technique to characterize the average of a set of values while taking their uncertainties into account. Mathematically, it is given by $\frac{\sum_i y_i/\sigma_i^2}{\sum_i 1/\sigma_i^2}$, where $y_i \in y_1, y_2, ..., y_N$, y_i is a measurement of some parameter, N is the number of measurements, and σ_i is the error of the measurement. In this way, more precise values are given more weight than less precise values.

$$\left(\begin{array}{cccc} & \mathcal{R}_1 & \mathcal{R}_2 & \mathcal{R}_3 \\ \mathcal{R}_1 & 1 & 0.71 & 0.15 \\ \mathcal{R}_2 & ... & 1 & -0.18 \\ \mathcal{R}_3 & ... & ... & 1 \end{array}\right) \tag{27}$$

The systematic uncertainties will be discussed in Section 4.7. Figure 34 demonstrates that the intrinsic results are invariant as functions of the p_T and η of the signal baryons, as they should be.

Table 12: Summary of results. See text. $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$ are defined in Equations 24, 25, and 26. N is the efficiency-corrected yield; $\varXi_c^0$ is implied to decay as $\varXi_c^0 \to \pi^- (\varLambda_c^+ \to pK^-\pi^+)$, $\varLambda_c^+$ as $\varLambda_c^+ \to pK^-\pi^+$, and $\varXi_c^+$ as $\varXi_c^+ \to pK^-\pi^+$. Statistical errors only. No integrated luminosity or secondary decay corrections have been applied.

η	p_T [GeV]	$N(\varXi_c^0)$ [10^3]	$N(\varLambda_c^+)$ [10^6]	$\mathcal{R}_1$ [%]
[2, 2.75)	[5, 12)	3084.2 ± 827.6	3271.0 ± 119.6	0.09 ± 0.03
	[12, 55)	73.5 ± 10.3	83.5 ± 5.0	0.09 ± 0.01
[2.75, 5)	[5, 12)	5132.4 ± 1301.5	5748.0 ± 144.4	0.09 ± 0.02
	[12, 55)	49.8 ± 8.6	54.0 ± 3.0	0.09 ± 0.02

η	p_T [GeV]	$N(\varXi_c^0)$ [10^3]	$N(\varXi_c^+)$ [10^4]	$\mathcal{R}_2$ [%]
[2, 2.75)	[5, 12)	3084.2 ± 827.6	5822.6 ± 235.8	5.30 ± 1.44
	[12, 55)	73.5 ± 10.3	116.0 ± 15.1	6.34 ± 1.21
[2.75, 5)	[5, 12)	5132.4 ± 1301.5	9585.4 ± 289.5	5.35 ± 1.37
	[12, 55)	49.8 ± 8.6	91.8 ± 12.1	5.42 ± 1.18

η	p_T [GeV]	$N(\varXi_c^+)$ [10^4]	$N(\varLambda_c^+)$ [10^6]	$\mathcal{R}_3$ [%]
[2, 2.75)	[5, 12)	5822.6 ± 235.8	3271.0 ± 119.6	1.78 ± 0.10
	[12, 55)	116.0 ± 15.1	83.5 ± 5.0	1.39 ± 0.20
[2.75, 5)	[5, 12)	9585.4 ± 289.5	5748.0 ± 144.4	1.67 ± 0.07
	[12, 55)	91.8 ± 12.1	54.0 ± 3.0	1.70 ± 0.24

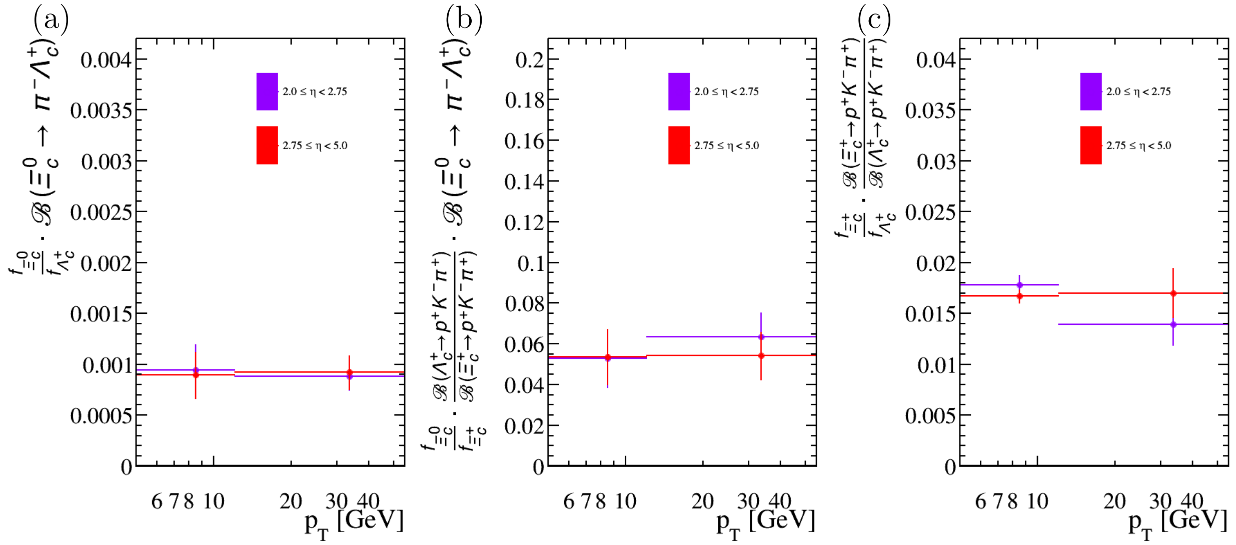


Figure 34: Intrinsic results as functions of p_T and η for (a) $\mathcal{R}_1$, (b) $\mathcal{R}_2$, and (c) $\mathcal{R}_3$, defined in Equations 24, 25, and 26. Only statistical uncertainties (from both data and simulation) are shown.

4.6.2 Derived results

Here we discuss additional calculations necessary to derive the final results.

In principle we want to assume that

$$\frac{f_{\Xi_c^0}}{f_{\Lambda_c^+}} = \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \tag{28}$$

in order to calculate $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ and $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$; however, we need to correct this relation because, in the charm case, $f_{\Xi_c^0} = f_{\Xi_c^+}$ to within ±1% from isospin symmetry, but $f_{\Xi_b^-} \neq f_{\Xi_b^0}$ primarily because the $\Xi_b'^0$ state is not seen to decay into $\Xi_b^-\pi^+$ [3], whereas both the $\Xi_b'^-$ and Ξ_b^{*-} states are seen to decay into $\Xi_b\pi$ [66].[41] (The $\Xi_b'^0$ state is presumed to be below threshold for decays to $\Xi_b^-\pi^+$ and thus only allowed transitions to Ξ_b^0 [66].[42]) That is,

$$\mathcal{B}(\Xi_b'^- \to \Xi_b\pi) = \mathcal{B}(\Xi_b^{*-} \to \Xi_b\pi) = \mathcal{B}(\Xi_b^{*0} \to \Xi_b\pi) = \mathcal{B}(\Xi_b'^0 \to \Xi_b^0 X) = 1 \tag{29}$$

and

$$\mathcal{B}(\Xi_b'^0 \to \Xi_b^- X) = 0. \tag{30}$$

Thus, the relation we need is

$$\frac{f_{\Xi_c^0}}{f_{\Lambda_c^+}} = \frac{0.5 \times \left(f_{\Xi_b^-} + f_{\Xi_b^0}\right)}{f_{\Lambda_b^0}} = \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \times 0.5 \times \left(1 + \frac{f_{\Xi_b^0}}{f_{\Xi_b^-}}\right). \tag{31}$$

Since $\frac{f_{\Xi_b^0}}{f_{\Xi_b^-}}$ has not been measured, we must derive an expression using measured values.

---

[41] The production fractions, *e.g.*, $f_{\Xi_b^-}$, are measured for particles produced promptly in the *pp* collision. "Promptly", in this case, effectively means that they appear to originate from the PV. This excludes particles that are produced in b decays, since they are characteristically long-lived and their decay vertices are easily separated from the PV. The decays of excited states, however, such as $\Xi_b'^- \to \Xi_b^-\pi^0$, cannot be separated from the PV, since they are characteristically short-lived. Thus, production fractions include particles produced directly in the collision (hadronizing without any intermediate states) and daughter particles resulting from the decays of directly-produced excited states.

[42] In order to decay to a given final state, a mother particle must have a mass at least as great as the sum of the masses of the daughters in that final state. This minimum required mass is called "threshold", and particles that do not have enough mass to decay to a given final state are said to be "below threshold".

In the following, we use these symbols and definitions:

$$\mathcal{C} \equiv 0.5 \times \left(1 + \frac{f_{\Xi_b^0}}{f_{\Xi_b^-}}\right) \tag{32}$$

$$= 0.5 \times \left(1 + \frac{\sigma(pp\to \Xi_b^0 X)}{\sigma(pp\to \Xi_b^- X)}\right) \tag{33}$$

$$r \equiv \frac{\sigma(pp\to \Xi_b^0 X)}{\sigma(pp\to \Xi_b^- X)} \tag{34}$$

$$\Xi_b^{**} \in \left\{\Xi_b'^0, \Xi_b'^-, \Xi_b^{*0}, \Xi_b^{*-}\right\} \tag{35}$$

$$i \in \{0, -\} \tag{36}$$

$$\Sigma^i \equiv \sum_{\Xi_b^{**}} \sigma(pp\to \Xi_b^{**} X) \times \mathcal{B}(\Xi_b^{**} \to \Xi_b^i X) \tag{37}$$

$$f^i_{\Xi_b^{**}} \equiv \frac{\sigma(pp\to \Xi_b^{**} X) \times \mathcal{B}(\Xi_b^{**} \to \Xi_b^i X)}{\sigma(pp\to \Xi_b^i X)} \tag{38}$$

First notice that we can rewrite r as

$$r = 1 + \frac{\sigma(pp\to \Xi_b^0 X) - \sigma(pp\to \Xi_b^- X)}{\sigma(pp\to \Xi_b^- X)}. \tag{39}$$

Then

$$\sigma(pp\to \Xi_b^i X) = \sigma(pp \xrightarrow{\text{direct}} \Xi_b^i X) + \Sigma^i + \Sigma^i_{\text{other}}, \tag{40}$$

where Σ^i_{other} is the sum of contributions from other excited states. Since

$$\sigma(pp \xrightarrow{\text{direct}} \Xi_b^0 X) = \sigma(pp \xrightarrow{\text{direct}} \Xi_b^- X) \tag{41}$$

and

$$\Sigma^0_{\text{other}} = \Sigma^-_{\text{other}} \tag{42}$$

due to isospin symmetry and because these states are far from threshold, we find

$$\begin{aligned}
r &= 1 + \frac{\left[\sigma(pp \xrightarrow{\text{direct}} \Xi_b^0 X) + \Sigma^0 + \Sigma^0_{\text{other}}\right] - \left[\sigma(pp \xrightarrow{\text{direct}} \Xi_b^- X) + \Sigma^- + \Sigma^-_{\text{other}}\right]}{\sigma(pp \to \Xi_b^- X)} \\
&= 1 + \frac{\sigma(pp \to \Xi_b^0 X)}{\sigma(pp \to \Xi_b^- X)} \times \frac{\Sigma^0}{\sigma(pp \to \Xi_b^0 X)} - \frac{\Sigma^-}{\sigma(pp \to \Xi_b^- X)} \\
&= 1 + r \times \sum_{\Xi_b^{**}} f^0_{\Xi_b^{**}} - \sum_{\Xi_b^{**}} f^-_{\Xi_b^{**}} \qquad (43)
\end{aligned}$$

using Equation 39.

The following relative production fractions have been measured in References [3, 66]:

$$f^-_{\Xi_b^{*0}} = \frac{\sigma(pp \to \Xi_b^{*0} X) \times \mathcal{B}(\Xi_b^{*0} \to \Xi_b^- \pi^+)}{\sigma(pp \to \Xi_b^- X)} = (28 \pm 3 \pm 1)\% \qquad (44)$$

$$f^0_{\Xi_b^{\prime -}} = \frac{\sigma(pp \to \Xi_b^{\prime -} X) \times \mathcal{B}(\Xi_b^{\prime -} \to \Xi_b^0 \pi^-)}{\sigma(pp \to \Xi_b^0 X)} = (11.8 \pm 1.7 \pm 0.7)\% \qquad (45)$$

$$f^0_{\Xi_b^{*-}} = \frac{\sigma(pp \to \Xi_b^{*-} X) \times \mathcal{B}(\Xi_b^{*-} \to \Xi_b^0 \pi^-)}{\sigma(pp \to \Xi_b^0 X)} = (20.7 \pm 3.2 \pm 1.5)\%. \qquad (46)$$

We use Clebsch-Gordon coefficients and relative phase-space factors to calculate the $f^i_{\Xi_b^{**}}$ that have not been measured. The relative phase-space for a decay is given by

$$\text{RPH}(\Xi_b^{**} \to \Xi_b^i \pi^j / \Xi_b^{**} \to \Xi_b^k \pi^l) \equiv \frac{\Phi\left(M_{\Xi_b^i}/M_{\Xi_b^{**}}, M_{\pi^j}/M_{\Xi_b^{**}}\right)}{\Phi\left(M_{\Xi_b^k}/M_{\Xi_b^{**}}, M_{\pi^l}/M_{\Xi_b^{**}}\right)}, \qquad (47)$$

where the function Φ is the usual phase-space function,[43]

$$\Phi(x, y) = \sqrt{[1 - (x + y)^2][1 - (x - y)^2)]}, \qquad (48)$$

and the Ms are masses. (Table 13 lists the relevant Clebsch-Gordon coefficients and masses.)

---

[43] See Section 48.4.2 of Reference [25].

We then note that

$$\mathcal{B}(\varXi_b^{**}\to \varXi_b^{i}\pi^{j}) = \mathcal{B}(\varXi_b^{**}\to \varXi_b^{k}\pi^{l}) \times \frac{\mathrm{CG}^{i}_{\varXi_b^{**}}}{\mathrm{CG}^{k}_{\varXi_b^{**}}} \times \mathrm{RPH}(\varXi_b^{**}\to \varXi_b^{i}\pi^{j}/\varXi_b^{**}\to \varXi_b^{k}\pi^{l}), \tag{49}$$

where the $\mathrm{CG}^{i}_{\varXi_b^{**}}$ indicate the squared Clebsch-Gordon coefficients given in Table 13. Therefore:

$$\begin{align}
f^{0}_{\varXi_b^{*0}} &= \frac{1}{r} \times f^{-}_{\varXi_b^{*0}} \times \frac{\mathrm{CG}^{0}_{\varXi_b^{*0}}}{\mathrm{CG}^{-}_{\varXi_b^{*0}}} \times \mathrm{RPH}(\varXi_b^{*0}\to \varXi_b^{0}\pi^{0}/\varXi_b^{*0}\to \varXi_b^{-}\pi^{+}) \tag{50}\\
f^{0}_{\varXi_b'^{0}} &= \frac{\sigma(pp\to \varXi_b'^{0}X)\times \mathcal{B}(\varXi_b'^{0}\to \varXi_b^{0}X)}{\sigma(pp\to \varXi_b^{0}X)} = \frac{\sigma(pp\to \varXi_b'^{0}X)\times 1}{\sigma(pp\to \varXi_b^{0}X)} = \frac{\sigma(pp\to \varXi_b'^{-}X)\times 1}{\sigma(pp\to \varXi_b^{0}X)} \notag\\
&= \frac{\sigma(pp\to \varXi_b'^{-}X)\times \left[\mathcal{B}(\varXi_b'^{-}\to \varXi_b^{0}\pi^{-}) + \mathcal{B}(\varXi_b'^{-}\to \varXi_b^{-}\pi^{0})\right]}{\sigma(pp\to \varXi_b^{0}X)} \notag\\
&= f^{0}_{\varXi_b'^{-}} + \frac{1}{r}\times f^{-}_{\varXi_b'^{-}} \tag{51}\\
f^{-}_{\varXi_b^{*-}} &= r \times f^{0}_{\varXi_b^{*-}} \times \frac{\mathrm{CG}^{-}_{\varXi_b^{*-}}}{\mathrm{CG}^{0}_{\varXi_b^{*-}}} \times \mathrm{RPH}(\varXi_b^{*-}\to \varXi_b^{-}\pi^{0}/\varXi_b^{*-}\to \varXi_b^{0}\pi^{-}) \tag{52}\\
f^{-}_{\varXi_b'^{-}} &= r \times f^{0}_{\varXi_b'^{-}} \times \frac{\mathrm{CG}^{-}_{\varXi_b'^{-}}}{\mathrm{CG}^{0}_{\varXi_b'^{-}}} \times \mathrm{RPH}(\varXi_b'^{-}\to \varXi_b^{-}\pi^{0}/\varXi_b'^{-}\to \varXi_b^{0}\pi^{-}), \tag{53}
\end{align}$$

where we have taken $\sigma(pp\to \varXi_b'^{0}X) = \sigma(pp\to \varXi_b'^{-}X)$ in Equation 51 due to isospin symmetry.

Substituting the $f^{i}_{\varXi_b^{**}}$ into Equation 43, we find, after a little algebra,

$$\begin{align}
r &= \frac{1 + f^{-}_{\varXi_b^{*0}} \times \left[\frac{\mathrm{CG}^{0}_{\varXi_b^{*0}}}{\mathrm{CG}^{-}_{\varXi_b^{*0}}} \times \mathrm{RPH}(\varXi_b^{*0}\to \varXi_b^{0}\pi^{0}/\varXi_b^{*0}\to \varXi_b^{-}\pi^{+}) - 1\right]}{1 - 2\times f^{-}_{\varXi_b'^{-}} - f^{0}_{\varXi_b^{*-}} \times \left[1 - \frac{\mathrm{CG}^{-}_{\varXi_b^{*-}}}{\mathrm{CG}^{0}_{\varXi_b^{*-}}} \times \mathrm{RPH}(\varXi_b^{*-}\to \varXi_b^{-}\pi^{0}/\varXi_b^{*-}\to \varXi_b^{0}\pi^{-})\right]} \tag{54}\\
&= 1.37 \pm 0.09 \tag{55}\\
\implies \mathcal{C} &= 1.18 \pm 0.04, \tag{56}
\end{align}$$

where the uncertainty is due to the uncertainties on the measured values of the M and $f^{i}_{\varXi_b^{**}}$ and the relative phase-space factors are shown in Table 14. Table 15 shows the calculated

Table 13: Relevant properties of Ξ_b baryon and π meson states. M is the mass. $\mathrm{CG}^i_{\Xi_b^{**}}$ indicates the squared Clebsch-Gordon coefficient for $\Xi_b^{**} \to \Xi_b^i \pi$. All values come from the PDG [1]. $M(\Xi_b'^0)$ has not been measured. We list $\mathrm{CG}^0_{\Xi_b'^0}$ as 1 and $\mathrm{CG}^-_{\Xi_b'^0}$ as 0 because $\Xi_b'^0 \to \Xi_b^- \pi^+$ is not observed (see text) [3].

State	M (MeV)	$\mathrm{CG}^0_{\Xi_b^{**}}$	$\mathrm{CG}^-_{\Xi_b^{**}}$
Ξ_b^-	5797.00 ± 0.90	–	–
Ξ_b^0	5791.90 ± 0.50	–	–
$\Xi_b'^0$	–	1	0
$\Xi_b'^-$	5935.02 ± 0.05	2/3	1/3
Ξ_b^{*0}	5952.30 ± 0.90	1/3	2/3
Ξ_b^{*-}	5955.33 ± 0.13	2/3	1/3
$\pi^\pm$	139.57061 ± 0.00024	–	–
π^0	134.977 ± 0.0005	–	–

values of the unmeasured $f^i_{\Xi_b^{**}}$.

Table 14: Calculated RPH values used in Equation 54.

Parameter	Value
$\mathrm{RPH}(\Xi_b^{*0} \to \Xi_b^0 \pi^0 / \Xi_b^{*0} \to \Xi_b^- \pi^+)$	1.27 ± 0.04
$\mathrm{RPH}(\Xi_b^{*-} \to \Xi_b^- \pi^0 / \Xi_b^{*-} \to \Xi_b^0 \pi^-)$	0.97 ± 0.02

Table 15: Calculated values for the expressions in Equations 50, 51, 52, and 53.

Parameter	Value [%]
$f^0_{\Xi_b^{*0}}$	13.0 ± 1.9
$f^0_{\Xi_b'^0}$	17.1 ± 2.8
$f^-_{\Xi_b^{*-}}$	13.8 ± 2.8
$f^-_{\Xi_b'^-}$	7.3 ± 1.9

We can now calculate $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ from Equation 5 using $f_{\Xi_b^-}/f_{\Lambda_b^0}$, which has been measured by LHCb as $(8.2 \pm 0.7 \pm 2.6)\%$ [54], and applying the correction in Equation 56:

$$\mathcal{B}_1 = \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = (0.98 \pm 0.04 \pm 0.35)\%. \tag{57}$$

We also calculate $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ from Equation 9 using Belle's measurement of $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+) = (0.45 \pm 0.21 \pm 0.07)\%$ [55] and the PDG value of $\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+) = (6.23 \pm 0.33)\%$ [1]:

$$\mathcal{B}_2 = \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = (0.41 \pm 0.01 \pm 0.21)\%. \tag{58}$$

$\mathcal{B}_1$ and $\mathcal{B}_2$ are correlated through the measurement of $N(\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+))$. We calculate $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$ from Equation 10 using the same value and correction for $f_{\Xi_b^-}/f_{\Lambda_b^0}$ as we did in Equation 57 and using $\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)$, which is listed in the PDG as $(6.23 \pm 0.33)\%$ [1]:

$$\mathcal{B}_3 = \mathcal{B}(\Xi_c^+ \to pK^-\pi^+) = (1.135 \pm 0.002 \pm 0.387)\%. \tag{59}$$

This is the correlation matrix for the derived results:

$$\left(\begin{array}{cccc} & \mathcal{B}_1 & \mathcal{B}_2 & \mathcal{B}_3 \\ \mathcal{B}_1 & 1 & 0.07 & 0.92 \\ \mathcal{B}_2 & ... & 1 & -0.02 \\ \mathcal{B}_3 & ... & ... & 1 \end{array}\right) \tag{60}$$

We calculate the mean of $\mathcal{B}_1$ and $\mathcal{B}_2$ and provide the extended correlation matrix in Equations 106 and 107 in Section 4.8.

4.7 Systematic uncertainties

The major sources of uncertainties in this analysis are systematic. All sources are listed in Table 18.

4.7.1 Heavy-quark symmetry

In addition to the uncertainties on the measurements used in the calculation of $\mathcal{C}$ in Equation 56, we assign an uncertainty due to the assumption of heavy-quark symmetry in Equation 4. We estimate this uncertainty by examining an analogous ratio of production fractions between charm and beauty states

$$\frac{f_{D_s^+}}{f_{D^0} + f_{D^+}} = \frac{f_{B_s^0}}{f_{B^0} + f_{B^+}}. \tag{61}$$

Both have been measured in 13 TeV pp collisions. The beauty ratio over a range $4\,\text{GeV} < p_{\text{T}} < 25\,\text{GeV}$ is 0.122 ± 0.006 [67]. Most of the 4.3% relative systematic uncertainty is due to the uncertainty on $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$. The charm ratio is measured for $0\,\text{GeV} < p_{\text{T}} < 15\,\text{GeV}$. Using the total cross-sections reported in Reference [68], we find

$$\frac{f_{D_s^+}}{f_{D^0} + f_{D^+}} = \frac{\sigma(pp \to D_s^+ X)}{\sigma(pp \to D^0 X) + \sigma(pp \to D^+ X)} \approx 0.121, \tag{62}$$

almost identical to the beauty ratio. The statistical uncertainty on the charm ratio is negligible. We note that the beauty measurement uses the CLEO measurement of $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$, while the charm measurement uses a derived result in the CLEO paper of $\mathcal{B}(D_s^+ \to \phi\pi^+)$, so the systematic uncertainties on these measurements mostly cancel in the ratio. The rest of the systematic uncertainties, tracking, particle identification, luminosity, *etc.*, also mostly cancel, leaving us with a few percent uncertainty in the comparison of the ratios. The p_{T} distribution ratios fall linearly with p_{T} in the beauty case and are flatter in the charm case (see Figure 3 in Reference [67] and Section 4 in Reference [68]). Taking this into account, a

5% uncertainty due to the heavy-quark symmetry assumption appears reasonable. (This is in addition to the uncertainty on the correction given in Equation 56.)

4.7.2 Charm-baryon lifetimes

We estimate the dependence on the uncertainty of the charm-baryon lifetimes by varying $\tau_{\text{meas}}\left(H_c\right)$ in Equation 11 within its uncertainty and recalculating the efficiencies. We take the variation from the nominal result to be a systematic uncertainty, which results in the uncertainties shown in Table 16.

Table 16: Uncertainties due to charm baryon lifetimes, given in percent.

[%]	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$
$\tau(\Xi_c^0)$	2.3	2.3	-
$\tau(\Lambda_c^+)$	1.2	0.7	2.1
$\tau(\Xi_c^+)$	-	1.1	1.3

4.7.3 Particle identification

We assign an uncertainty due to the use of the `PIDCalib` package, which uses an independent dataset to determine the efficiency of the PID selections in the Turbo lines as a function of p and η (see Sections 4.3.4 and 4.3.1). We account for the limited statistics of the `PIDCalib` sample by varying the weight given coherently within its uncertainty and recalculating the efficiencies, which results in a conservative uncertainty of 4×10^{-5} on $\mathcal{R}_1$, 1.5×10^{-5} on $\mathcal{R}_2$, and 2.7×10^{-5} on $\mathcal{R}_3$; these uncertainties are quite small due to the large statistics of the samples used by `PIDCalib`.

We assign an additional 1% uncertainty due to the use of the offline “HasRich” variable in simulation; data uses the online version, but a bug in the calibration samples prevents us from using it for simulation.

To be conservative, we assign an additional 1% uncertainty for particle identification to the additional π meson in $\Xi_c^0 \to \pi^- \left(\Lambda_c^+ \to pK^-\pi^+\right)$ decay.

4.7.4 Tracking

The tracking uncertainties largely cancel in the ratios, but we must account for the additional π meson in $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$. The tracking efficiency comes from the `TrackCalib` package, which uses an independent dataset to determine the tracking efficiency as a function of p and η (see Section 4.3.4). We account for the limited statistics of the `TrackCalib` sample by varying the weight given to the π meson coherently within its uncertainty and treating the variation as a source of bin-by-bin uncertainty in the result, which results in a conservative overall uncertainty of 0.8% on $\mathcal{R}_1$ and 0.9% on $\mathcal{R}_2$.

Since `TrackCalib` relies on a sample of μ leptons, whereas we are interested in the tracking efficiency for π mesons, we must assign an additional uncertainty for hadronic interactions. Hadronic interactions prevent about 14% of π mesons from being reconstructed, according to simulation, and there is a 10% material-budget uncertainty; therefore, we assign an additional 1.4% uncertainty to the use of `TrackCalib`.[44]

Finally, the `TrackCalib` method itself caries a 0.8% systematic uncertainty;[45] we round this up to 1% to account for differences between the simulation versions used in the analysis and those used in the `TrackCalib` datasets.

4.7.5 Fit model

We must also account for the systematic uncertainty introduced by the choice of fit model. To this end, we fit a variety of alternative models to the (TOS) data datasets and assign the largest yield variation from the nominal model as the systematic uncertainty on the yield due to the fit model. We then fit the model with the largest yield variation in TOS data to the TIS and TISTOS data datasets and assign the variation in the ratio of TISTOS to TIS yields as the systematic uncertainty on the trigger efficiency due to the fit model (see Section 4.4). Sections 4.7.6, 4.7.7, and 4.7.8 describe the alternative shapes that we consider and the systematic uncertainties that we assign. Note that we treat the uncertainties on fit models as independent from one another; we only apply the correlations described in Section 4.4.1 to the nominal yields.

---

[44] See "Tracking systematics for hadrons" at the link in footnote 31 and Section 9 of Reference [69].

[45] The method used to arrive at the 0.8% value for the systematic uncertainty is described in Section 8 of Reference [69].

4.7.6 Ξ_c^0 fit model

The nominal fit model for the $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$ TOS, TIS, and TISTOS datasets is a Crystal Ball signal + linear background shape (see Section 4.5.1). We choose to consider the alternative fit models shown in Figure 35. The double-Gaussian signal shape has the largest variation, yielding 6662 ± 360 signal candidates, a variation of 5.5% from the nominal, and we therefore assign a systematic uncertainty of 5.5%. We fit also the TIS and TISTOS data datasets using the double-Gaussian signal shape, shown in Figure 36, yielding 889 ± 110 and 321 ± 146 signal candidates, respectively, and leading to a trigger efficiency of 36.1%, a 1.5% deviation from the nominal result of 35.6%. We therefore assign a systematic uncertainty of 1.5%.

4.7.7 Λ_c^+ fit model

The nominal fit model for the $\Lambda_c^+ \to pK^-\pi^+$ TOS, TIS, and TISTOS datasets is a double-sided Crystal Ball signal + 2nd-order Chebychev background shape (see Section 4.5.1). We choose to consider the alternative fit models shown in Figure 37. The third alternative signal shape (width, α, and left and right n fixed to 7.34, 2, 5.96, and 155, respectively) has the largest variation, yielding 2647553 ± 1840 signal candidates, a variation of 0.7% from the nominal, and we therefore assign a systematic uncertainty of 0.7%. We fit also the TIS and TISTOS data datasets using this signal shape, shown in Figure 38, yielding 432496 ± 763 and 141973 ± 450 signal candidates, respectively, and leading to a trigger efficiency of 32.8%, a 1.8% deviation from the nominal result of 33.4%. We therefore assign a systematic uncertainty of 1.8%.

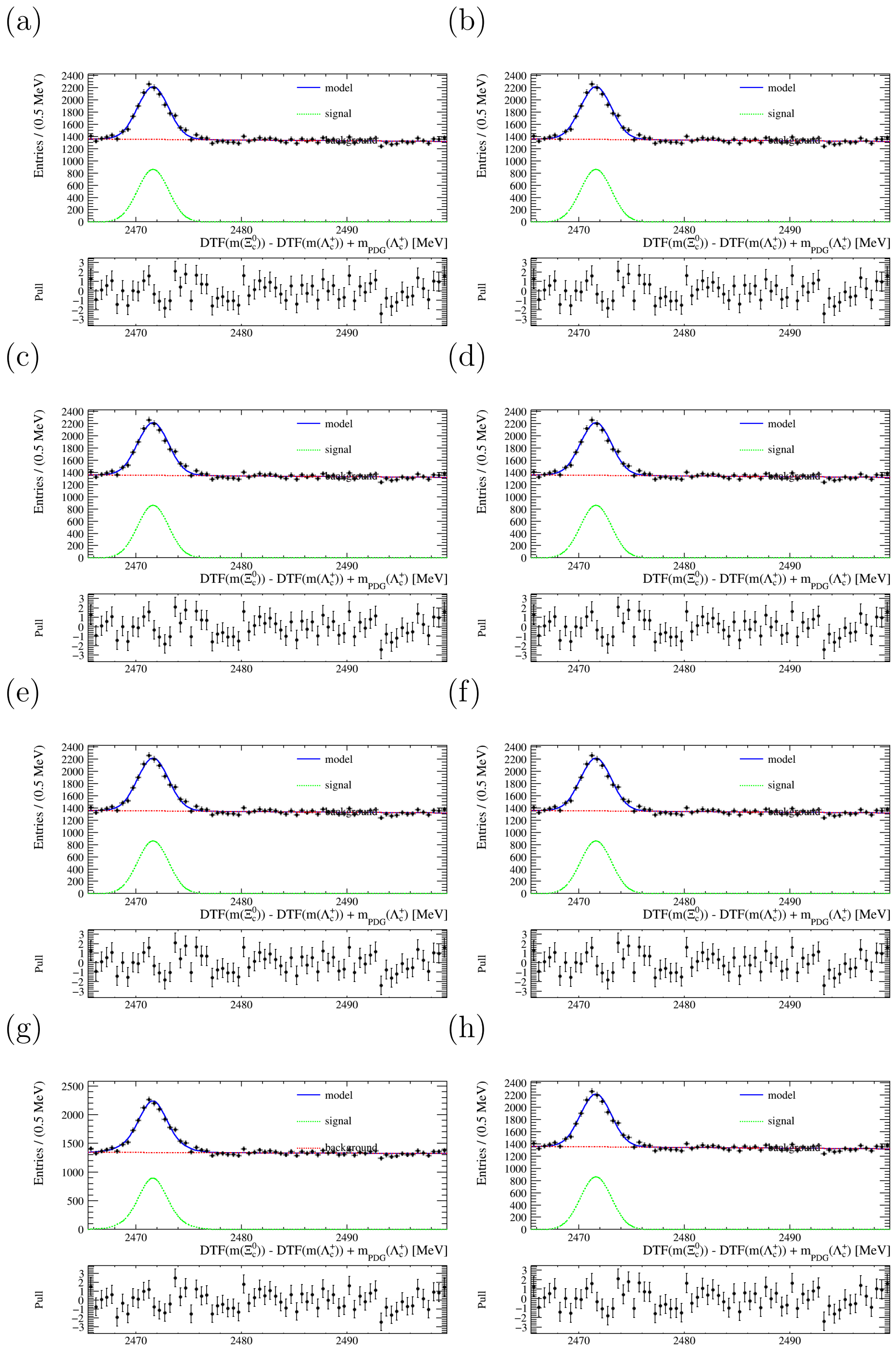

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Figure 35: DTF $\left(m\left(\Xi_c^0\right)\right)$ $-$DTF $\left(m\left(\Lambda_c^+\right)\right)$ $+m_{\text{PDG}}\left(\Lambda_c^+\right)$ distribution from the Ξ_c^0 TOS data dataset fit using alternative models: The α and n parameters of the Crystal Ball are fixed to, respectively, (a) -5 and 100, (b) -25 and 5, (c) -20 and 22, and (d) -10 and 22; (e) an exponential background shape; (f) a Gaussian signal shape; (g) a double-Gaussian signal shape, where the Gaussians are constrained to have the same mean; and (h) a Gaussian signal shape with an exponential background shape.

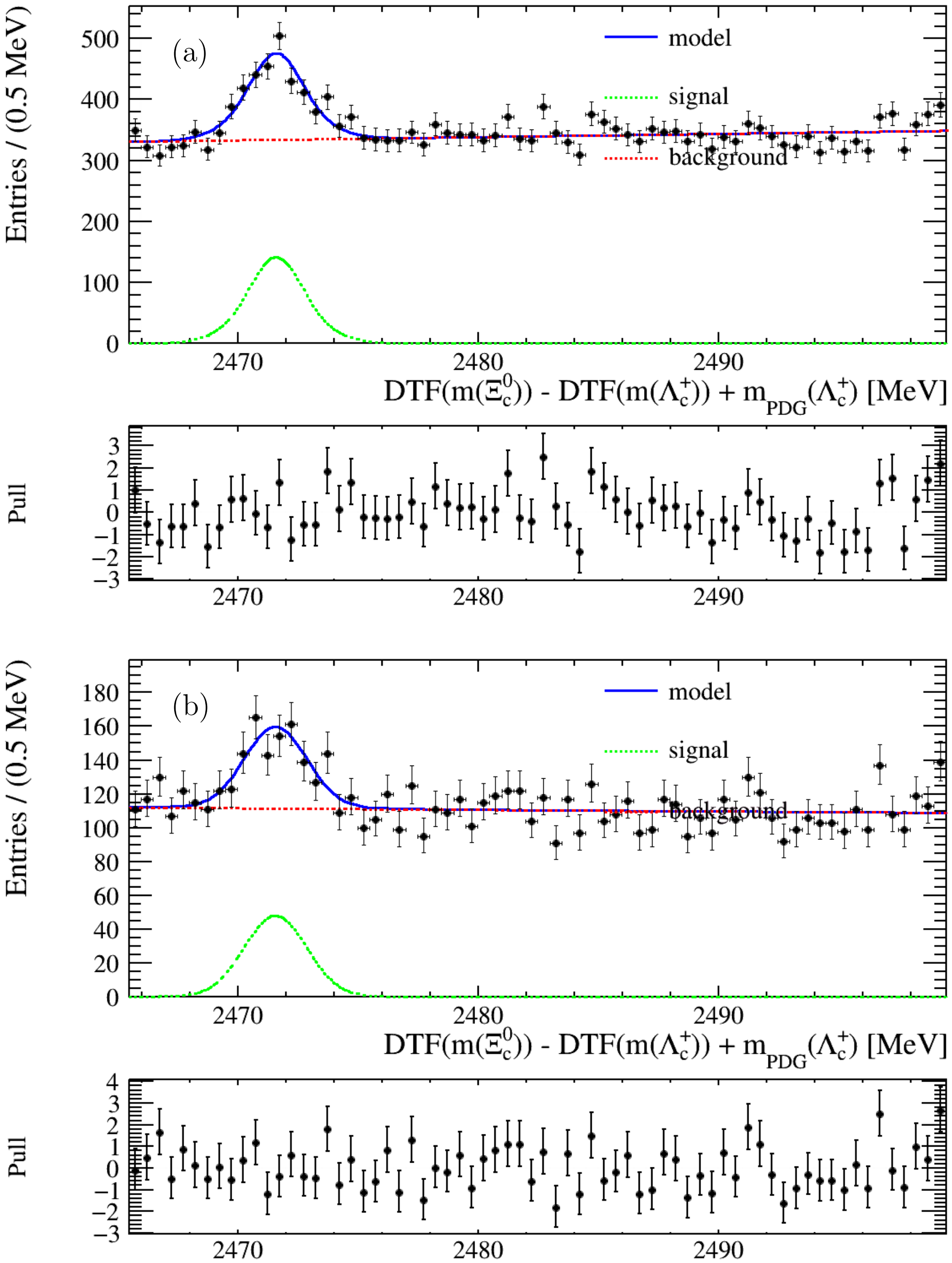


Figure 36: DTF $\left(m\left(\Xi_c^0\right)\right)$ $-$DTF $\left(m\left(\Lambda_c^+\right)\right)$ $+m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ distributions from the Ξ_c^0 (a) TIS and (b) TISTOS data datasets fit using a model that is the same as the nominal, but with a double-Gaussian signal shape, where the Gaussians are constrained to have the same mean.

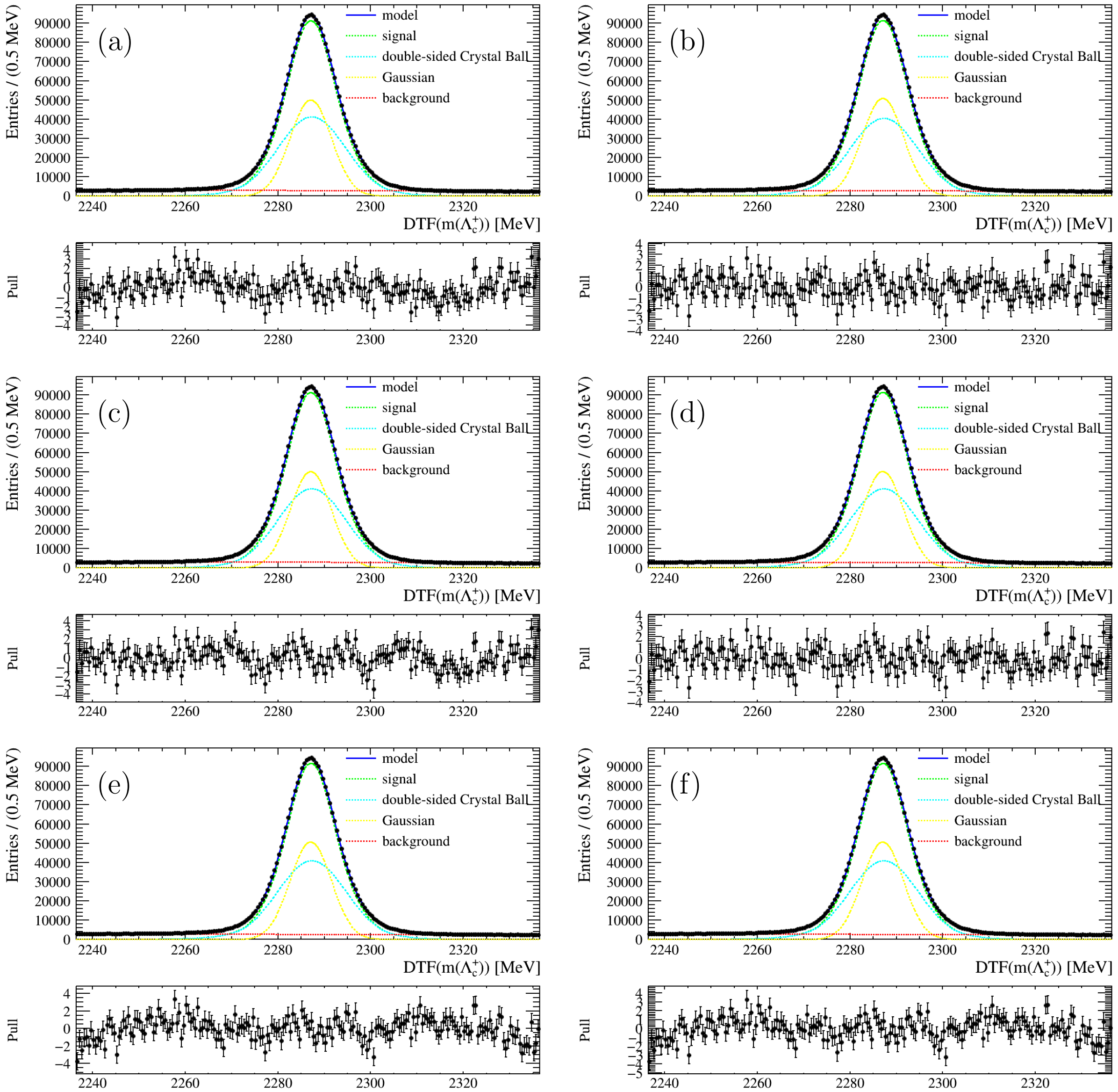


Figure 37: DTF mass distribution of $pK^-\pi^+$ from the Λ_c^+ TOS data dataset fit using alternative models: (a) the left and right n parameters of the double-sided Crystal Ball are fixed to 15 and 120, respectively; (b) the left and right n parameters of the double-sided Crystal Ball are fixed to 5.96 and 155, respectively, and the fraction of the signal yield assigned to the double-sided Crystal Ball is allowed to float; (c) the width, α, and left and right n parameters of the double-sided Crystal Ball are fixed to 7.34, 2, 5.96, and 155, respectively; (d) the width and left and right n parameters of the double-sided Crystal Ball are fixed to 7.34, 5.96, and 155, respectively, and the α parameter is allowed to float, though it still constrained to be the same for both left and right; (e) an exponential background shape; and (f) a linear background shape.

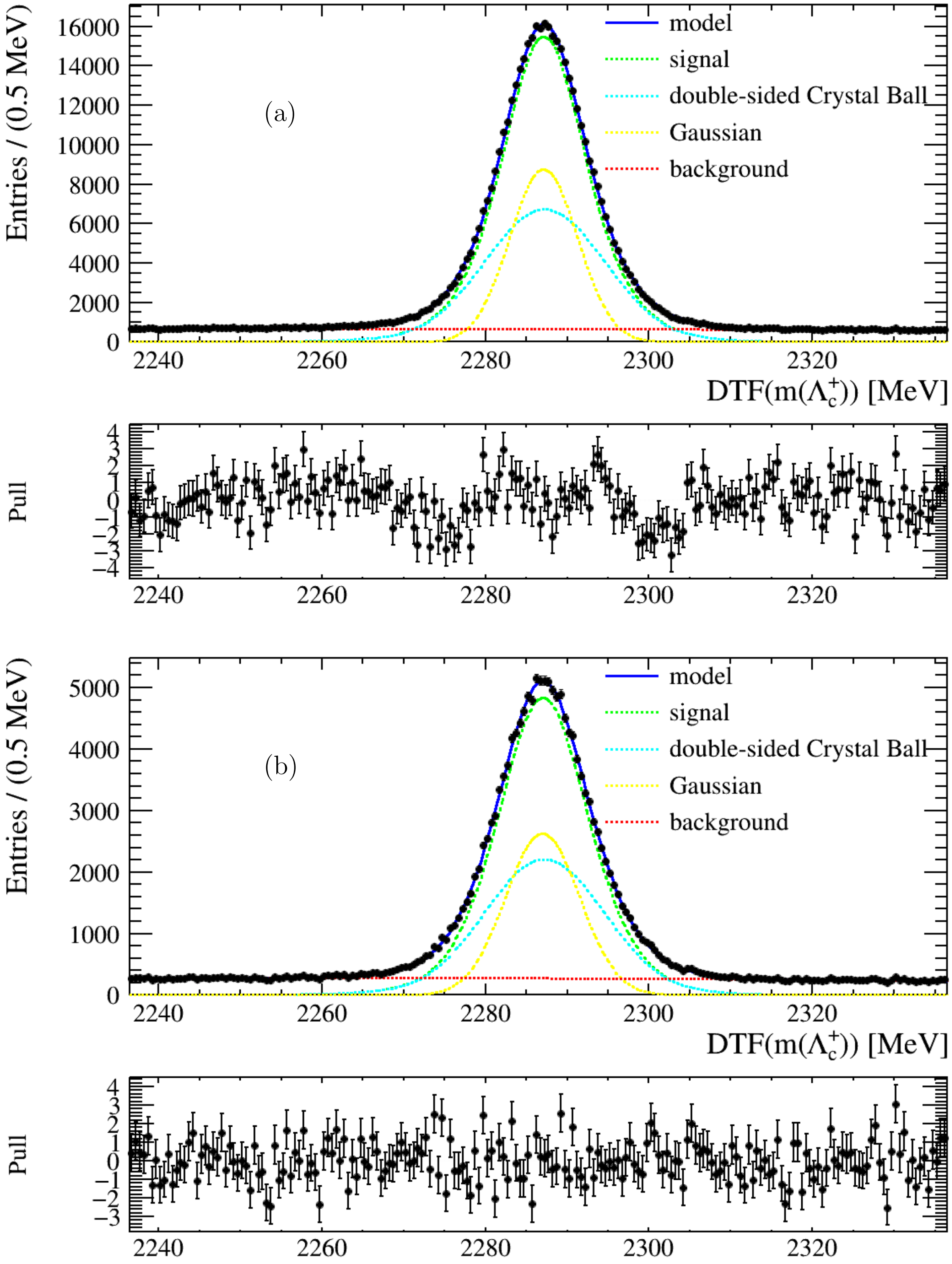


Figure 38: DTF mass distribution of $pK^-\pi^+$ from the Λ_c^+ (a) TIS and (b) TISTOS data datasets fit using a model that is the same as the nominal, but the width, α, and left and right n parameters of the double-sided Crystal Ball are fixed to 7.34, 2, 5.96, and 155, respectively.

4.7.8 Ξ_c^+ fit model

The nominal fit model for the $\Xi_c^+ \to pK^-\pi^+$ TOS, TIS, and TISTOS datasets is a double-sided Crystal Ball signal + 2nd-order Chebychev background shape. We choose to consider the alternative fit models shown in Figure 39. The double-Gaussian signal shape has the largest variation, yielding 1570725 ± 2637 signal candidates, a variation of 2.6% from the nominal, and we therefore assign a systematic uncertainty of 2.6%. We fit also the TIS and TISTOS data datasets using this signal shape, shown in Figure 40, yielding 244821 ± 2140 and 69181 ± 1255 signal candidates, respectively, and leading to a trigger efficiency of 28.3%, a 0.5% deviation from the nominal result of 28.4%. We therefore assign a systematic uncertainty of 0.5%.

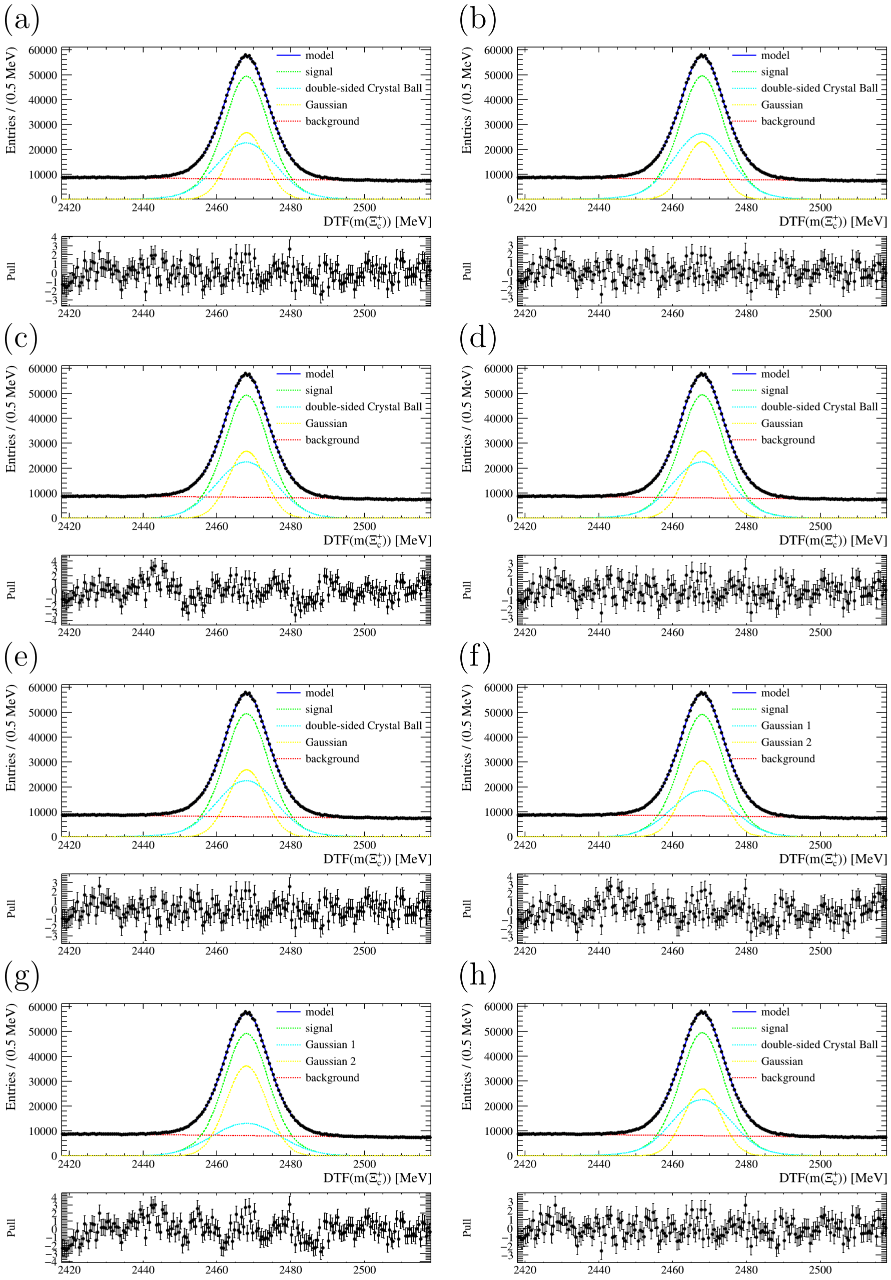


Figure 39: DTF mass distribution of $pK^-\pi^+$ from the Ξ_c^+ TOS data dataset fit using alternative models: The left and right n parameters of the double-sided Crystal Ball are fixed to, respectively, (a) 15 and 120; (b) 5.5 and 155, and the fraction of the signal yield assigned to the double-sided Crystal Ball is allowed to float; (c) 5.5 and 155 and the width and α to 8.09 and 5, respectively; and (d) 5.5 and 155 and the width to 8.09, and the α parameter is allowed to float, though it is still constrained to be the same for both left and right. (e) An exponential background shape. (f) A double-Gaussian signal shape, where the Gaussians are constrained to have the same mean. (g) A double-Gaussian signal shape, where the Gaussians are constrained to have the same mean, with an exponential background shape. (h) A linear background shape.

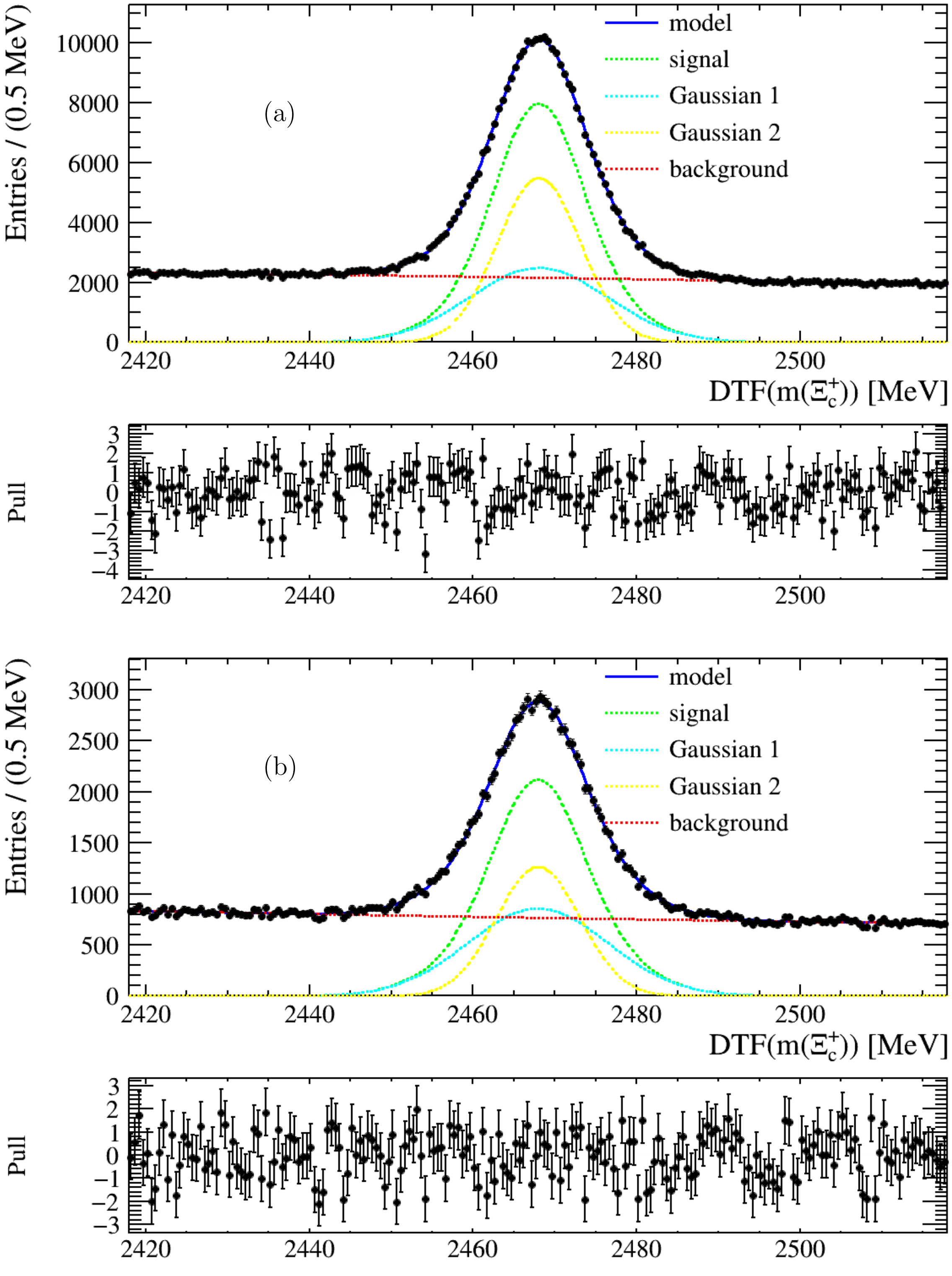


Figure 40: DTF mass distribution of $pK^-\pi^+$ from the Ξ_c^+ (a) TIS and (b) TISTOS data datasets fit using a model that is the same as the nominal, but with a double-Gaussian signal shape, where the Gaussians are constrained to have the same mean.

4.7.9 Dalitz weighting

We estimate the uncertainty on the simulation-derived efficiencies due to the weighting of the intermediate resonance spectra in $\Lambda_c^+ \to pK^-\pi^+$ and $\Xi_c^+ \to pK^-\pi^+$ decays by varying the value of each bin in the scaling histograms (Figure 19 [e & f]) stochastically within its statistical uncertainty. We do this by sampling from a Gaussian distribution centered on the bin's nominal value with a standard deviation equal to its error. We do this 250 times for each histogram, thus finding 250 alternate weighting histograms for each decay spectrum. We then recalculate each of the results using weights taken from each combination of the alternate histograms; so since $\mathcal{R}_1$ (Equation 24) depends on $N(\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+))$ and $N(\Lambda_c^+ \to pK^-\pi^+)$, we calculate 250 alternate results; since $\mathcal{R}_2$ (Equation 25) depends on $N(\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+))$ and $N(\Xi_c^+ \to pK^-\pi^+)$, we calculate $250 \times 250 = 62500$ alternate results; and since $\mathcal{R}_3$ (Equation 26) depends on $N(\Xi_c^+ \to pK^-\pi^+)$ and $N(\Lambda_c^+ \to pK^-\pi^+)$, we calculate $250 \times 250 = 62500$ alternate results. Finally, in each case, we take the arithmetic mean of all 250 or 62500 alternate results and compare that to the nominal result. We assign systematic uncertainties equal to the standard deviations of the alternate results.

This method results in systematic uncertainties of 2.0% on $\mathcal{R}_1$, 1.8% on $\mathcal{R}_2$, and 1.7% on $\mathcal{R}_3$.

4.7.10 BDT input parameter differences in data and simulation

If the distributions of the variables used to train the BDTs vary between the simulation and data datasets, the efficiencies in simulation could be incorrect. We therefore weight a representative distribution in truth-matched simulation to match that in background-subtracted data and consider the subsequent change in results to represent the uncertainty on BDT efficiency due to differences between data and simulation. After applying all selections (Turbo, offline, BDT, and trigger), we use the *sPlot* method [16] to assign weights to the data sample using the signal fits shown in Figures 24, 25, and 26. We then compare the distributions of all BDT training variables in truth-matched simulation to those in s-weighted data and choose the $\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(H_c)\cdot\vec{p}(H_c)}{x(H_c)\cdot p(H_c)}\right)\right)\right)$ distributions as representative of the relevant differences.[46] (See Appendix A.4 for plots of all the BDT training variables.) The ratio of the normalized histograms can be used as a weighting distribution to align simulation with data, as shown in Figure 41. We apply these weights to the post-BDT simulation sample and recalculate the efficiencies, leading to a change in the nominal results of 2.8% on $\mathcal{R}_1$, 3.0% on $\mathcal{R}_2$, and 0.6% on $\mathcal{R}_3$, which we round up to 1% to be conservative.

---

[46] The caption of Table 4 includes an explanation of this variable.

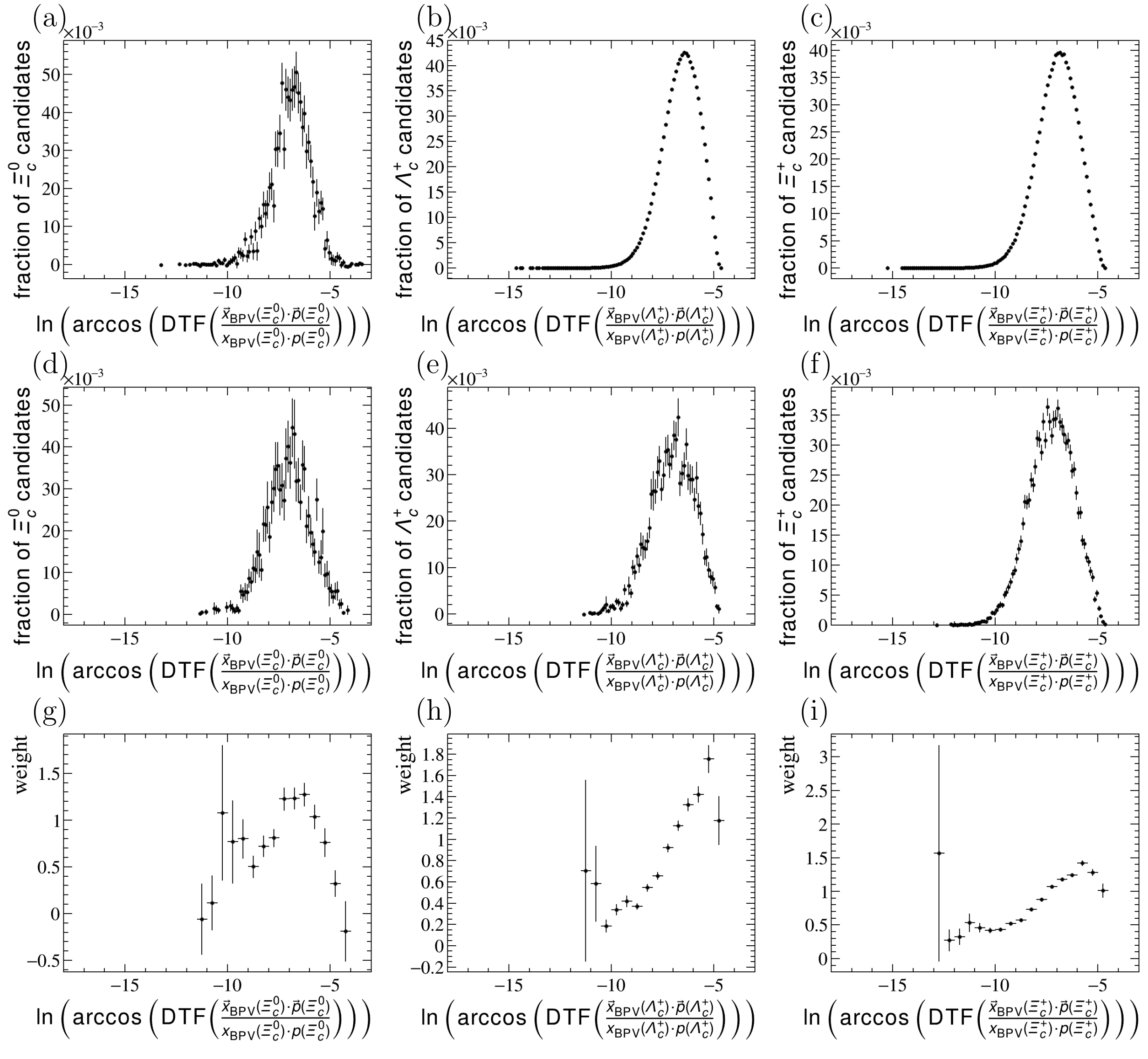


Figure 41: $\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(H_c)\cdot\vec{p}(H_c)}{x(H_c)\cdot p(H_c)}\right)\right)\right)$ distributions for signal (left) $\Xi_c^0\to\pi^-\ (\Lambda_c^+\to pK^-\pi^+)$, (center) $\Lambda_c^+\to pK^-\pi^+$, and (right) $\Xi_c^+\to pK^-\pi^+$ in (a, b, & c) data, (d, e, & f) simulation, and (g, h, & i) data divided by simulation (weight histograms).

4.7.11 b-baryon decay contamination

The approach detailed in Section 4.1 only applies if the relevant baryons are promptly produced; it is therefore important that we exclude baryons from secondary vertices from the signal yields. We rely on the BDTs to discriminate charm baryons produced promptly in the proton collision from those produced by the decay of b-baryons, called "secondary decays" here, in the data datasets. (We explicitly exclude secondaries from the simulation datasets; see Section 4.3.4.) We estimate here the fractional contamination in each of the charmed baryon (H_c) samples, and then we evaluate the changes in the ratios we are measuring. The relative secondary contamination in a given data dataset is given by

$$f_s(H_c) = \frac{\epsilon_s^{\text{tot}}(H_c \to f)}{\epsilon_p^{\text{tot}}(H_c \to f)} \times \frac{\sigma(pp \to H_b X)}{\sigma(pp \to H_c X)} \times \mathcal{B}(H_b \to H_c X), \tag{63}$$

where ϵ^{tot} is the total efficiency of all selections determined from simulation, the s subscript indicates the secondaries sample, the p subscript indicates the prompt sample, H_b is the bottom baryon in question (which decays into a given charm baryon H_c), and f refers to the particular final state of the decay of the H_c baryon.

We estimate the level of contamination from secondaries by applying the selections to a sample of simulated $\mathit{\Lambda}_b^0 \to (\mathit{\Lambda}_c^+ \to pK^-\pi^+)X$ and $\mathit{\Xi}_b^0 \to (\mathit{\Xi}_c^+ \to pK^-\pi^+)X$ decays; there are no $H_b \to (\mathit{\Xi}_c^0 \to \pi^-\mathit{\Lambda}_c^+)X$ simulation samples available. Table 17 details the simulation samples we use. We reconstruct $\mathit{\Lambda}_c^+ \to pK^-\pi^+$ and $\mathit{\Xi}_c^+ \to pK^-\pi^+$ decays from these samples and apply the selections (excluding secondary rejection; see Section 4.3). We can then evaluate ϵ^{tot} for each dataset. We find

$$\epsilon_p^{\text{tot}}(\mathit{\Lambda}_c^+ \to pK^-\pi^+) = (0.586 \pm 0.010)\% \tag{64}$$

$$\epsilon_s^{\text{tot}}(\mathit{\Lambda}_c^+ \to pK^-\pi^+) = (0.637 \pm 0.008)\% \tag{65}$$

$$\epsilon_p^{\text{tot}}(\mathit{\Xi}_c^+ \to pK^-\pi^+) = (2.089 \pm 0.016)\% \tag{66}$$

$$\epsilon_s^{\text{tot}}(\mathit{\Xi}_c^+ \to pK^-\pi^+) = (0.717 \pm 0.009)\%. \tag{67}$$

$\epsilon_p^{\text{tot}}(\Xi_c^+ \to pK^-\pi^+) > \epsilon_p^{\text{tot}}(\Lambda_c^+ \to pK^-\pi^+)$ due to the longer lifetime of the Ξ_c^+ baryon (see Table 8).

Table 17: Simulation samples used for secondary-discrimination studies. $\Lambda_c^+ \to pK^-\pi^+$ or $\Xi_c^+ \to pK^-\pi^+$ decays are used in all cases. N_{gen} is the number of generated signal charm-baryons (either Λ_c^+ or Ξ_c^+ baryons).

Decay	Year	N_{gen}
$\Lambda_b^0 \to \pi^- \Lambda_c^+$	2016	1.5M
	2017	0.4M
	2018	0.5M
$\Lambda_b^0 \to K^- \Lambda_c^+$	2017	0.5M
	2018	0.5M
$\Lambda_b^0 \to K^- K^+ \pi^- \Lambda_c^+$	2016	2.0M
$\Lambda_b^0 \to \overline{\nu}_\mu \mu^- \Lambda_c^+$	2016	2.0M
	2017	1.0M
	2018	4.2M
$\Xi_b^0 \to \pi^- \Xi_c^+$	2016	4.0M
$\Xi_b^{*-} \to \pi^-(\Xi_b^0 \to \overline{\nu}_\mu \mu^- \Xi_c^+)$	2016	1.5M

We must also determine the relative production cross-sections for H_b to H_c baryons. At 7 TeV, LHCb has measured the Λ_c^+ baryon production cross-section to be $(233 \pm 26 \pm 71 \pm 14)\,\mu\text{b}$ [70] and that for the $B^\pm$ meson to be $(43 \pm 0.2 \pm 2.5 \pm 1.7)\,\mu\text{b}$ [71]. We must then consider the b-baryon production fraction versus that of the b-meson; LHCb has measured $f_{\Lambda_b^0}/(f_u + f_d) = 0.259 \pm 0.018$ at 13 TeV [67], and we double this value to account for $B^0 + \overline{B}^0$. Putting this all together, we find

$$\begin{aligned}\frac{\sigma(pp \to \Lambda_b^0 X)}{\sigma(pp \to \Lambda_c^+ X)} &= \frac{\sigma(pp \to B^\pm X)}{\sigma(pp \to \Lambda_c^+ X)} \times \frac{f_{\Lambda_b^0}}{\frac{1}{2}(f_u + f_d)} \\ &= (9.6 \pm 3.3)\%. \end{aligned} \tag{68}$$

We must calculate $\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ X)$, since it has not been measured. To this end, we assume

$$\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ X) + \mathcal{B}(\Lambda_b^0 \to D^0 p X) + \mathcal{B}(\Lambda_b^0 \to D^+ n X) + \mathcal{B}(\Lambda_b^0 \to u X) = 1, \tag{69}$$

where $\mathit{\Lambda}_b^0 \to uX$ indicates decays where the b-quark in the $\mathit{\Lambda}_b^0$ baryon decays to a u-quark instead of a c-quark. In Reference [67], LHCb measures

$$N_{\text{raw}}(\mathit{\Lambda}_b^0 \to (\mathit{\Lambda}_c^+ \to pK^-\pi^+)\mu X\overline{\nu}_\mu) = 1753600 \tag{70}$$

$$N_{\text{raw}}(\mathit{\Lambda}_b^0 \to (\mathit{\Lambda}_c^+(2860) \to (D^0 \to K^-\pi^+)p)\mu X\overline{\nu}_\mu) = 6120 \pm 460 \tag{71}$$

$$N_{\text{raw}}(\mathit{\Lambda}_b^0 \to (\mathit{\Lambda}_c^+(2880) \to (D^0 \to K^-\pi^+)p)\mu X\overline{\nu}_\mu) = 2200 \pm 200 \tag{72}$$

$$N_{\text{raw}}(\mathit{\Lambda}_b^0 \to (\mathit{\Lambda}_c^+(2940) \to (D^0 \to K^-\pi^+)p)\mu X\overline{\nu}_\mu) = 1200 \pm 260 \tag{73}$$

$$N_{\text{raw}}(\mathit{\Lambda}_b^0 \xrightarrow{\text{nonresonant}} (D^0 \to K^-\pi^+)p\mu X\overline{\nu}_\mu) = 29770 \pm 690, \tag{74}$$

where N_{raw} is the yield without correcting for the efficiency. Noticing that, since $b \to u$ decays comprise $\approx 1\%$ of semileptonic b-hadron decays [72],

$$99\% \times N(\mathit{\Lambda}_b^0) = N(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X) + N(\mathit{\Lambda}_b^0 \to D^0 pX) + N(\mathit{\Lambda}_b^0 \to D^+ nX) \tag{75}$$

and assuming

$$\frac{N(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X)}{N(\mathit{\Lambda}_b^0)} = \frac{N(\mathit{\Lambda}_b^0 \xrightarrow{\text{SL}} \mathit{\Lambda}_c^+ X)}{N(\mathit{\Lambda}_b^0 \xrightarrow{\text{SL}} X)}, \tag{76}$$

where $\xrightarrow{\text{SL}}$ indicates semileptonic decay, we calculate

$$\begin{aligned}
\mathcal{B}\left(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X\right) &= \frac{N\left(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X\right)}{N\left(\mathit{\Lambda}_b^0 \to X\right)} \\
&= \frac{N\left(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X\right) \times 99\%}{N\left(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X\right) + N\left(\mathit{\Lambda}_b^0 \to D^0 pX\right) + N\left(\mathit{\Lambda}_b^0 \to D^+ nX\right)} \\
&= \frac{N\left(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X\right) \times 99\%}{N\left(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X\right) + 2 \times N\left(\mathit{\Lambda}_b^0 \to D^0 pX\right)} \\
&= \frac{\frac{N\left(\mathit{\Lambda}_b^0 \to \left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)X\right)}{\mathcal{B}\left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)} \times 99\%}{\frac{N\left(\mathit{\Lambda}_b^0 \to \left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)X\right)}{\mathcal{B}\left(\mathit{\Lambda}_c^+ \to pK^-\pi^+\right)} + 2 \times \frac{N\left(\mathit{\Lambda}_b^0 \to (D^0 \to K^-\pi^+)pX\right)}{\mathcal{B}(D^0 \to K^-\pi^+)}} \\
&= (92.4 \pm 2.8)\%,
\end{aligned} \tag{77}$$

where we have assumed that $N(\mathit{\Lambda}_b^0 \to D^0 p X) = N(\mathit{\Lambda}_b^0 \to D^+ n X)$ due to isospin symmetry and that the relevant efficiencies cancel in the ratio. The latter assumption comes from all the measured decays having the same final-state daughters; we assign a 3% uncertainty to this assumption, which has been incorporated into Equation 77 above.

Now, we can put this all together to calculate

$$\begin{aligned} f_s(\mathit{\Lambda}_c^+) &= \frac{\epsilon_s^{\mathrm{tot}}(\mathit{\Lambda}_c^+)}{\epsilon_p^{\mathrm{tot}}(\mathit{\Lambda}_c^+)} \times \frac{\sigma(pp \to \mathit{\Lambda}_b^0 X)}{\sigma(pp \to \mathit{\Lambda}_c^+ X)} \times \mathcal{B}(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X) \\ &= (108.6 \pm 2.3)\% \times (9.6 \pm 3.3)\% \times (92.4 \pm 2.8)\% \\ &= (9.6 \pm 3.3)\%. \end{aligned} \tag{78}$$

For $\mathit{\Xi}_c^+$, we assume that

$$\frac{\sigma(pp \to \mathit{\Xi}_c^+ X)}{\sigma(pp \to \mathit{\Lambda}_c^+ X)} = s_{\mathrm{corr}} = \frac{f_{\mathit{\Xi}_b^-}}{f_{\mathit{\Lambda}_b^0}} \times \mathcal{C} \tag{79}$$

where $\mathcal{C}$ is given by Equations 32 and 56 and s_{corr} is a factor accounting for the difference in quark content between the $\mathit{\Lambda}_c^+$ baryon (udc) and the $\mathit{\Xi}_c^+$ baryon (usc) or $\mathit{\Xi}_c^0$ baryon (dsc).

Thus, we derive

$$
\begin{aligned}
\frac{\sigma(pp\to \mathit{\Xi}_b^0 X)}{\sigma(pp\to \mathit{\Xi}_c^+ X)} &= \frac{f_{\mathit{\Xi}_b^0}}{f_{\mathit{\Xi}_c^+}} && (80)\\
&= \frac{f_{\mathit{\Xi}_b^-}}{f_{\mathit{\Xi}_c^+}} \times \frac{f_{\mathit{\Xi}_b^0}}{f_{\mathit{\Xi}_b^-}} && (81)\\
&= \frac{f_{\mathit{\Xi}_b^-}}{f_{\mathit{\Xi}_c^+}} \times (2\times\mathcal{C}-1) && (82)\\
&= (2\times\mathcal{C}-1) \times \frac{\sigma(pp\to B^\pm X)}{\sigma(pp\to \mathit{\Xi}_c^+ X)} \times \frac{f_{\mathit{\Xi}_b^-}}{\frac{1}{2}(f_u+f_d)} && (83)\\
&= (2\times\mathcal{C}-1) \times \frac{\sigma(pp\to B^\pm X)}{s_{\mathrm{corr}} \times \sigma(pp\to \mathit{\Lambda}_c^+ X)} \times \frac{f_{\mathit{\Lambda}_b^0} \times \frac{s_{\mathrm{corr}}}{\mathcal{C}}}{\frac{1}{2}(f_u+f_d)} && (84)\\
&= \frac{2\times\mathcal{C}-1}{\mathcal{C}} \times \frac{\sigma(pp\to B^\pm X)}{\sigma(pp\to \mathit{\Lambda}_c^+ X)} \times \frac{f_{\mathit{\Lambda}_b^0}}{\frac{1}{2}(f_u+f_d)} && (85)\\
&= \frac{2\times\mathcal{C}-1}{\mathcal{C}} \times \frac{\sigma(pp\to \mathit{\Lambda}_b^0 X)}{\sigma(pp\to \mathit{\Lambda}_c^+ X)} && (86)\\
&= (1.15\pm0.03) \times \frac{\sigma(pp\to \mathit{\Lambda}_b^0 X)}{\sigma(pp\to \mathit{\Lambda}_c^+ X)} && (87)\\
&= (11.0\pm3.9)\%, && (88)
\end{aligned}
$$

where we have incorporated the 5% relative uncertainty on the heavy-quark symmetry assumption (see Section 4.7.1), and therefore

$$
\begin{aligned}
f_s(\mathit{\Xi}_c^+) &= \frac{\epsilon_s^{\mathrm{tot}}(\mathit{\Xi}_c^+)}{\epsilon_p^{\mathrm{tot}}(\mathit{\Xi}_c^+)} \times \frac{\sigma(pp\to \mathit{\Xi}_b^0 X)}{\sigma(pp\to \mathit{\Xi}_c^+ X)} \times \mathcal{B}(\mathit{\Xi}_b^0\to \mathit{\Xi}_c^+ X)\\
&= (34.4\pm0.5)\% \times (11.0\pm3.9)\% \times (92.4\pm2.8)\%\\
&= (3.5\pm1.2)\%, && (89)
\end{aligned}
$$

where we have assumed $\mathcal{B}(\mathit{\Xi}_b^0\to \mathit{\Xi}_c^+ X) = \mathcal{B}(\mathit{\Lambda}_b^0\to \mathit{\Lambda}_c^+ X)$.

As mentioned above, there are no $H_b\to (\mathit{\Xi}_c^0\to \pi^-\mathit{\Lambda}_c^+)X$ simulation samples available,[47] so we must find a way to approximate $\frac{\epsilon_s^{\mathrm{tot}}(\mathit{\Xi}_c^0)}{\epsilon_p^{\mathrm{tot}}(\mathit{\Xi}_c^0)}$. We assume the charm-baryon lifetime is the driving factor differentiating $\frac{\epsilon_s^{\text{tot}}(\Lambda_c^+)}{\epsilon_p^{\text{tot}}(\Lambda_c^+)}$ from $\frac{\epsilon_s^{\text{tot}}(\Xi_c^+)}{\epsilon_p^{\text{tot}}(\Xi_c^+)}$ (see Table 8), since the b-baryon lifetimes are nearly the same [1]. We further assume $\frac{\epsilon_s^{\text{tot}}}{\epsilon_p^{\text{tot}}}$ is linear as a function of lifetime (τ). We can then use Equations 64, 65, 66, and 67 to determine

---

[47] There were no $H_b\to (\mathit{\Xi}_c^0\to \pi^-\mathit{\Lambda}_c^+)X$ simulation samples available at the time this analysis was completed. Some have been produced in the intervening time, but the approximation described in this section still suffices.

$$\frac{\epsilon_s^{\text{tot}}}{\epsilon_p^{\text{tot}}}(\tau) = (-2931 \pm 117)\,\text{ns}^{-1} \times \tau + (1.68 \pm 0.05), \tag{90}$$

which allows us to calculate

$$\frac{\epsilon_s^{\text{tot}}(\Xi_c^0)}{\epsilon_p^{\text{tot}}(\Xi_c^0)} = \frac{\epsilon_s^{\text{tot}}}{\epsilon_p^{\text{tot}}}\left(\tau\left(\Xi_c^0\right) + \tau\left(\Lambda_c^+\right)\right) = (63.3 \pm 2.0)\%, \tag{91}$$

where we have assumed $\frac{\epsilon_s^{\text{tot}}}{\epsilon_p^{\text{tot}}}$ is driven by the sum of the Ξ_c^0 and Λ_c^+ baryon lifetimes, since we reconstruct the Ξ_c^0 baryon by adding a π^- meson to an already reconstructed Λ_c^+ baryon (see Section 4.3.1); we have rounded the error from 1.6% to 2.0% to account for this assumption. We then assume that

$$\frac{\sigma(pp \rightarrow \Xi_c^0 X)}{\sigma(pp \rightarrow \Lambda_c^+ X)} = s_{\text{corr}} = \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \times \mathcal{C}, \tag{92}$$

similar to the Ξ_c^+ case above, and derive

$$\begin{align}
\frac{\sigma(pp \rightarrow \Xi_b^- X)}{\sigma(pp \rightarrow \Xi_c^0 X)} &= \frac{f_{\Xi_b^-}}{f_{\Xi_c^0}} \tag{93}\\
&= \frac{\sigma(pp \rightarrow B^\pm X)}{\sigma(pp \rightarrow \Xi_c^0 X)} \times \frac{f_{\Xi_b^-}}{\frac{1}{2}(f_u + f_d)} \tag{94}\\
&= \frac{\sigma(pp \rightarrow B^\pm X)}{s_{\text{corr}} \times \sigma(pp \rightarrow \Lambda_c^+ X)} \times \frac{f_{\Lambda_b^0} \times \frac{s_{\text{corr}}}{\mathcal{C}}}{\frac{1}{2}(f_u + f_d)} \tag{95}\\
&= \frac{1}{\mathcal{C}} \times \frac{\sigma(pp \rightarrow B^\pm X)}{\sigma(pp \rightarrow \Lambda_c^+ X)} \times \frac{f_{\Lambda_b^0}}{\frac{1}{2}(f_u + f_d)} \tag{96}\\
&= (0.85 \pm 0.03) \times \frac{\sigma(pp \rightarrow \Lambda_b^0 X)}{\sigma(pp \rightarrow \Lambda_c^+ X)} \tag{97}\\
&= (8.1 \pm 2.8)\%, \tag{98}
\end{align}$$

where we have again incorporated the 5% relative uncertainty on the heavy-quark symmetry assumption. This allows us to calculate

$$\begin{aligned} f_s(\mathit{\Xi}_c^0) &= \frac{\epsilon_s^{\mathrm{tot}}(\mathit{\Xi}_c^0)}{\epsilon_p^{\mathrm{tot}}(\mathit{\Xi}_c^0)} \times \frac{\sigma(pp\to \mathit{\Xi}_b^- X)}{\sigma(pp\to \mathit{\Xi}_c^0 X)} \times \mathcal{B}(\mathit{\Xi}_b^- \to \mathit{\Xi}_c^0 X) \\ &= (63.3 \pm 2.0)\% \times (8.1 \pm 2.8)\% \times (92.4 \pm 2.8)\% \\ &= (4.7 \pm 1.7)\%, \end{aligned} \tag{99}$$

where we have assumed $\mathcal{B}(\mathit{\Xi}_b^- \to \mathit{\Xi}_c^0 X) = \mathcal{B}(\mathit{\Lambda}_b^0 \to \mathit{\Lambda}_c^+ X)$.

We can now calculate the corrections:

$$\mathcal{C}_{\mathcal{R}_1} \equiv \frac{1 - f_s(\mathit{\Xi}_c^0)}{1 - f_s(\mathit{\Lambda}_c^+)} = \frac{(95.3 \pm 1.7)\%}{(90.4 \pm 3.3)\%} = (105.4 \pm 2.1)\% \tag{100}$$

$$\mathcal{C}_{\mathcal{R}_2} \equiv \frac{1 - f_s(\mathit{\Xi}_c^0)}{1 - f_s(\mathit{\Xi}_c^+)} = \frac{(95.3 \pm 1.7)\%}{(96.5 \pm 1.2)\%} = (98.7 \pm 0.6)\% \tag{101}$$

$$\mathcal{C}_{\mathcal{R}_3} \equiv \frac{1 - f_s(\mathit{\Xi}_c^+)}{1 - f_s(\mathit{\Lambda}_c^+)} = \frac{(96.5 \pm 1.2)\%}{(90.4 \pm 3.3)\%} = (106.7 \pm 2.6)\%, \tag{102}$$

where all correlations between the f_s have been taken into account. These corrections have been applied to the results given in Section 4.6, and the estimated errors are taken as systematic uncertainties.

4.7.12 Differences in integrated luminosity

The Turbo stream stores $\mathit{\Lambda}_c^+$ and $\mathit{\Xi}_c^+$ candidates in different sets of files, which we must run over separately when producing n-tuples for analysis. Since the retrieval script inevitably fails to process 100% of the files, we must account for differences in integrated luminosity, $\int \mathcal{L}$, between the data datasets. (The simulation samples are not stored separately.) Unfortunately, the integrated luminosity reported by DaVinci is not calibrated, but since $> 99.5\%$ of files are successfully processed in all cases, we can estimate the integrated luminosity handled in each case based on the mean fraction of uncalibrated integrated luminosity handled by each instance of the retrieval script. To do so, we recognize that each instance of the retrieval script handles the same number of files, modulo some remainder; fill a histogram of the integrated luminosity reported by each instance; and find the mean and standard deviation of this histogram. This is the mean (uncalibrated) integrated luminosity, $\left\langle\int \mathcal{L}_i\right\rangle$, handled by each instance of the retrieval script for a given year for either $\mathit{\Lambda}_c^+$ or $\mathit{\Xi}_c^+$. We then calculate the integrated luminosity:

$$\int \mathcal{L}_s = \sum_y \left\{ \int \mathcal{L}_y \times \left[1 - \frac{(N_{s,y}^t - N_{s,y}^c) \times \left\langle\int \mathcal{L}_i\right\rangle}{(N_{s,y}^t - N_{s,y}^c) \times \left\langle\int \mathcal{L}_i\right\rangle + \sum_i^{N_{s,y}^c} \int \mathcal{L}_i} \right] \right\}, \tag{103}$$

where s is either the $\mathit{\Lambda}_c^+$ or $\mathit{\Xi}_c^+$ file set; y is the year (2017 or 2018); $\int \mathcal{L}_y$ is the value of the total integrated luminosity for a given year ($\int \mathcal{L}_{2017} = 1.6\,\mathrm{fb}^{-1}$, $\int \mathcal{L}_{2018} = 2.2\,\mathrm{fb}^{-1}$); $N_{s,y}^t$ is the total number of instances of the retrieval script for a given file set for a given year; $N_{s,y}^c$ is the number of successful instances of the retrieval script for a given file set for a given year; and $\int \mathcal{L}_i$ is the (uncalibrated) integrated luminosity reported by each instance of the retrieval script. We find

$$\frac{\int \mathcal{L}_{\mathit{\Lambda}_c^+}}{\int \mathcal{L}_{\mathit{\Xi}_c^+}} = 1.024 \pm 0.007, \tag{104}$$

where the uncertainty comes from the standard deviation of the means and the statistical error on $\sum_i^{N_{s,y}^c} \int \mathcal{L}_i$. We multiply by this factor in the determination of $\mathcal{R}_2$ and divide by it in the determination of $\mathcal{R}_3$ (see Section 4.6.1).

4.7.13 Summary

Table 18 summarizes the sources of systematic uncertainty.

Table 18: Relative systematic uncertainties

Source	Estimate [%]		
	$\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$		$\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$
	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$
$\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$	–	49	–
$\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)$	–	5	5
$f_{\Xi_c^0}/f_{\Xi_c^+} = 1$	–	1	1
$f_{\Xi_b^-}/f_{\Lambda_b^0}$	32	–	32
$f_{\Xi_c^0}/f_{\Lambda_c^+} = \mathcal{C} \cdot f_{\Xi_b^-}/f_{\Lambda_b^0}$	6	–	6
Lifetimes	3	3	2
TISTOS statistics	7	8	1
Simulation statistics	4	3	2
`DInLHCb` correction	–	6	2
PID	1	1	0
`PIDCalib`	1	1	1
`TrackCalib`	2	2	0
Fit model	6	6	3
TISTOS fit model	2	2	2
Ghost correction	2	2	0
Dalitz correction	2	2	2
BDT training variables	3	3	1
Secondaries correction	2	0	2
$\int \mathcal{L}$	–	1	1
Sum of external	33	49	33
Sum of intrinsic	12	13	6
Sum of all	35	51	34

4.8 Comparison with experiment

We measure $\mathcal{B}_1 = \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)\%$ using Equation 5 in Equation 57 and ($\mathcal{B}_2$) using Equation 9 in Equation 58. Taking all correlations into account, we find

$$\mathcal{B}_1 - \mathcal{B}_2 = (0.57 \pm 0.39)\%. \tag{105}$$

That is, they differ by 1.4 standard deviations. Given that each result depends on different external inputs, this difference is not surprising. The best linear unbiased estimate (BLUE) of their mean is

$$\langle \mathcal{B}_{1,2} \rangle = \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = (0.55 \pm 0.02 \pm 0.18)\%, \tag{106}$$

and this is their correlation matrix, an extended version of the one given in Equation 60:

$$\left(\begin{array}{ccccc} & \langle \mathcal{B}_{1,2} \rangle & \mathcal{B}_1 & \mathcal{B}_2 & \mathcal{B}_3 \\ \langle \mathcal{B}_{1,2} \rangle & 1 & 0.53 & 0.88 & 0.41 \\ \mathcal{B}_1 & ... & 1 & 0.07 & 0.92 \\ \mathcal{B}_2 & ... & ... & 1 & -0.02 \\ \mathcal{B}_3 & ... & ... & ... & 1 \end{array}\right) \tag{107}$$

We determine the BLUE of the mean following the method given by Nisius in Reference [73],

$$\langle \mathcal{B}_{1,2} \rangle = (1 - \beta) \times \mathcal{B}_2 + \beta \times \mathcal{B}_1, \tag{108}$$

where

$$\beta = \frac{1 - \rho z}{1 - 2 \times \rho z + z^2} = 25.2\% \tag{109}$$

$$\rho = 6.86\% \text{ (the correlation between } \mathcal{B}_1 \text{ and } \mathcal{B}_2\text{)} \tag{110}$$

$$z = \frac{\sigma(\mathcal{B}_1)}{\sigma(\mathcal{B}_2)} = 165\%. \tag{111}$$

Belle measures $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$ to be $(0.45 \pm 0.21 \pm 0.07)\%$ [55], whereas we measure $\mathcal{B}_3 = \mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$ using Equation 10 in Equation 59, finding

$$\mathcal{B}_3 - \mathcal{B}(\Xi_c^+ \to pK^-\pi^+)_{\mathrm{Belle}} = (0.69 \pm 0.45)\%. \tag{112}$$

That is, the measurements differ by 1.5 standard deviations, consistent with one another. The inverse-variance weighted mean of these results is[48]

$$(0.62 \pm 0.19)\%, \tag{113}$$

and our result has better relative precision (34% vs. 49%).

4.9 Comparison with theory

As discussed in Section 4.1, there have been several theoretical predictions for $\mathcal{B}(\Xi_c^0 \to \pi^-\Lambda_c^+)$, summarized in Table 19. We take $\langle\mathcal{B}_{1,2}\rangle$, given in Equation 106, as our nominal result. It is well above Voloshin's lower limit and consistent with Faller and Mannel's upper limit. It favors Gronau and Rosner's constructive prediction and disfavors Gronau and Rosner's and Cheng *et al.*'s destructive predictions. Figure 42 shows our results compared with the predictions.

---

[48] The PDG average of $\mathcal{B}_3$ with $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)_{\mathrm{Belle}}$ is $(0.62 \pm 0.30)\%$, since they have chosen to multiply the error by a scaling factor of 1.5 [25]; see Section 5.2.2 of Reference [25] for an explanation of this procedure.

Table 19: Theory predictions for $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ from Voloshin [4], Gronau and Rosner [5], Faller and Mannel [6], and Cheng *et al.* [7]. + or − indicates an assumption of positive or negative interference, respectively, between the SUUD and WS amplitudes (see Section 4.1). Appendix A.12 describes our updates to Gronau and Rosner's predictions.

Theorist	Prediction
Voloshin (-)	$\gtrapprox (0.025 \pm 0.015)\%$
Gronau and Rosner (+)	$(0.19 \pm 0.07)\%$
Gronau and Rosner (-)	$\lessapprox 0.01\%$
Gronau and Rosner (+), updated	$(0.14 \pm 0.07)\%$
Gronau and Rosner (-), updated	$\lessapprox 0.018 \pm 0.015\%$
Faller and Mannel (+)	$< 0.3\%$
Cheng *et al.* (-)	$\sim 0.0087\%$

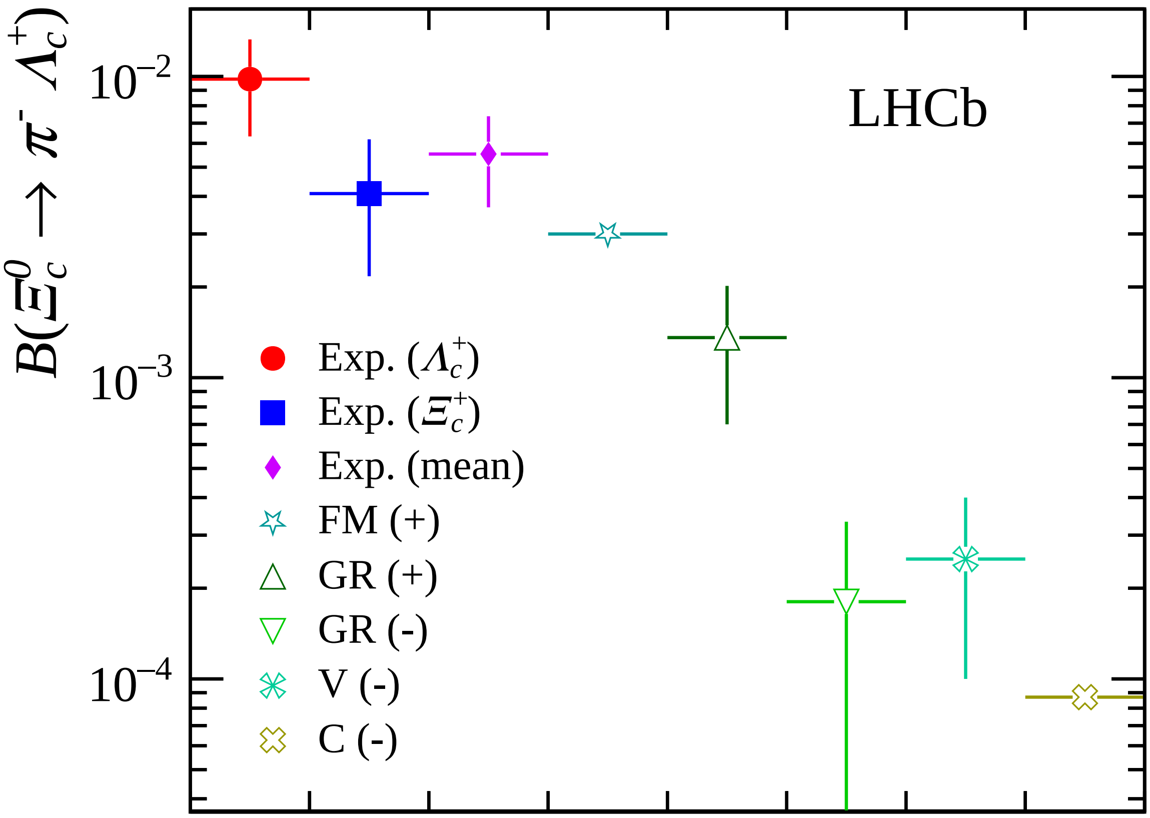


Figure 42: Comparison of $\mathcal{B}_1$, $\mathcal{B}_2$, and $\langle \mathcal{B}_{1,2} \rangle$ with Faller and Mannel's upper limit (FM) [6], the updated predictions of Gronau and Rosner (GR) [5], Voloshin's lower limit (V) [4], and Cheng *et al.*'s prediction (C) [7]. See Table 19. + or − indicates an assumption of positive or negative interference, respectively, between the SUUD and WS amplitudes (see Section 4.1).

5 Conclusions

We measure

$$
\begin{aligned}
\frac{f_{\Xi_c^0}}{f_{\Lambda_c^+}} \times \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) &= (0.095 \pm 0.003 \pm 0.012)\% \\
\frac{f_{\Xi_c^0}}{f_{\Xi_c^+}} \times \frac{\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)}{\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)} \times \mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) &= (5.70 \pm 0.19 \pm 0.77)\% \\
\frac{f_{\Xi_c^+}}{f_{\Lambda_c^+}} \times \frac{\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)}{\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)} &= (1.753 \pm 0.003 \pm 0.107)\%
\end{aligned}
$$

(Equations 24, 25, and 26), from which we calculate

$$
\begin{aligned}
\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) &= (0.98 \pm 0.04 \pm 0.35)\% \\
\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) &= (0.41 \pm 0.01 \pm 0.21)\% \\
\mathcal{B}(\Xi_c^+ \to pK^-\pi^+) &= (1.135 \pm 0.002 \pm 0.387)\%
\end{aligned}
$$

(Equations 57, 58, and 59), respectively; the best linear unbiased estimate (BLUE) of the mean of our two results for $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ is

$$\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = (0.55 \pm 0.02 \pm 0.18)\%$$

(Equation 106). These results for $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ and $\mathcal{B}(\Xi_c^+ \to pK^-\pi^+)$ have been published in References [25, 49].

This is the first observation of the suppressed decay $\Xi_c^0 \to \pi^- \Lambda_c^+$ and the first measurement of its branching fraction, providing a way to distinguish among theoretical predictions. This result is somewhat larger than any of them, but it is consistent with the constructive-interference predictions of Gronau and Rosner and of Faller and Mannel [5, 6], and it is inconsistent with the destructive-interference predictions of Gronau and Rosner and of Cheng *et al.* [5, 7], from which we conclude that the SUUD and WS amplitudes interfere constructively, the first experimental evidence that this is the case. In addition, this is the second measurement of the branching fraction of the suppressed decay $\Xi_c^+ \to pK^-\pi^+$, and it carries the best relative precision to date [25, 55].

Appendices

A $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ Appendices

All of the figures used in these appendices are taken from Reference [17] unless otherwise noted.

A.1 Background categories in simulation

The simulation serves two purposes: to calculate the signal-detection efficiencies and to provide an independent signal dataset with which to train the BDTs (see Section 4.3.5). To achieve both these ends, we require all reconstructed simulated particles to correspond to their generated counterparts by applying background category selections.[49] This ensures that we feed the BDTs a pure signal sample and that we calculate the *signal* efficiencies explicitly. Given the intermediate resonances generated in $\Lambda_c^+ \to pK^-\pi^+$ MC, the appropriate background category selections are 0 (signal) and 10 (quasi-signal) for the $\Xi_c^0 \to \pi^-$ ($\Lambda_c^+ \to pK^-\pi^+$) and $\Lambda_c^+ \to pK^-\pi^+$ samples. We also treat background category 50 (low mass) as signal for the $\Xi_c^0 \to \pi^-$ ($\Lambda_c^+ \to pK^-\pi^+$) sample; see below. For the $\Xi_c^+ \to pK^-\pi^+$ sample, since Ξ_c^+ is named a Λ_c^+ baryon by the Turbo stream, the appropriate background category selection is 20 (fully reconstructed physical background), but since this would still allow non-Ξ_c^+ decays to p, K^-, and π^+ hadrons, we also require the absolute value of the true PID of the Ξ_c^+ candidate to be 4232, following the PDG Monte Carlo Particle Numbering Scheme.[50] All three samples have some signal present in background category 60 (ghosts); we apply a correction to the efficiency due to these contributions, described in Section 4.4.

The inclusion of background category 50 as "truth-matched" stems from a thorough investigation of the decay chains of the candidates appearing in this category. All of the candidates in this category are associated with a generated Ξ_c^0 baryon, and all of the intermediate and final-state daughters have been assigned the correct PIDs. (It is likely that these candidates come from the random hadronization and decay of the other simulated c-quark in the decay to $\Xi_c^0 \to \Lambda_c^+ \to pK^-\pi^+X$.) Therefore, they appear to be signal, and we count them as such.

---

[49] `https://twiki.cern.ch/twiki/bin/view/LHCb/TupleToolMCBackgroundInfo`

[50] `http://pdg.lbl.gov/2002/montecarlorpp.pdf`

It is theoretically possible for signal to appear outside the "truth-matching" selections we apply to simulation (see Section 4.3.4). Figure 43 compares the distributions of truth-matched simulation to that failing the truth-matching criteria, excluding ghosts. (We will address ghosts separately.) Only the $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$ sample has any non-ghost peaking contribution outside truth-matching, and Figure 44 (a) shows that the peaking contribution comes only from candidates matched to a generated Ξ_c^0 baryon. Figure 44 (b) shows that this peaking contribution comes from background category 30 (reflection) (no other non-ghost background categories contribute).

In category 30, a real Ξ_c^0 may have been found, as in Figure 44 (a), but the algorithm thinks one or more of the final-state daughters has been assigned the wrong mass hypothesis, based on the generated particles in the event. A real Ξ_c^0 may thus be found but appear as two candidates: one with the correct mass hypotheses, one without. If this peaking structure is from multiple candidates, it is important to exclude it from "signal" to avoid double-counting; additionally, the weight provided by `PIDCalib` is only valid for particles with correct mass hypotheses.

Careful investigation reveals that the peak in background category 30 is indeed an artifact of multiple candidates in an event. Figure 45 (a) shows that almost all the background 30 candidates appear in events with a candidate in a "signal" category. Figure 45 (b) demonstrates only a small peaking contribution remains once the duplicate candidates are weeded out; if we included all of these candidates, not just the ones in the peak, it would result in a 1.2% increase in the number of signal candidates surviving selections. Given the size of the other systematic uncertainties (see Section 4.7), we exclude them.

As for ghosts, Figure 46 demonstrates that each simulation dataset has a non-negligible

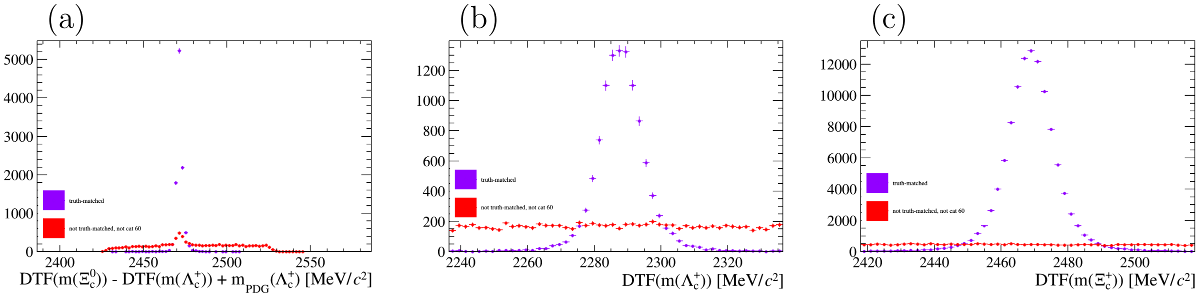


Figure 43: Mass distributions for (a) $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$, (b) $\Lambda_c^+ \to pK^-\pi^+$, and (c) $\Xi_c^+ \to pK^-\pi^+$ decays in simulation. The purple distributions have truth-matching applied. The red distributions fail the truth-matching criteria. None of these plots has any of the weighting described in Section 4.3.4 applied.

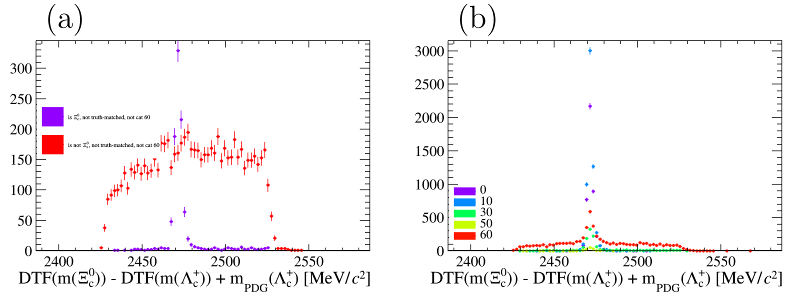


Figure 44: Mass distributions from $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$ decays in simulation for candidates (a) failing the truth-matching criteria described in Section 4.3.4 that (purple) are and (red) are not matched to a generated Ξ_c^0 and (b) falling into various background categories. None of these plots has any of the weighting described in Section 4.3.4 applied.

peaking ghost contribution. We therefore estimate the amount of signal in background category 60 (ghosts) and correct the efficiencies accordingly, as described in Section 4.4.

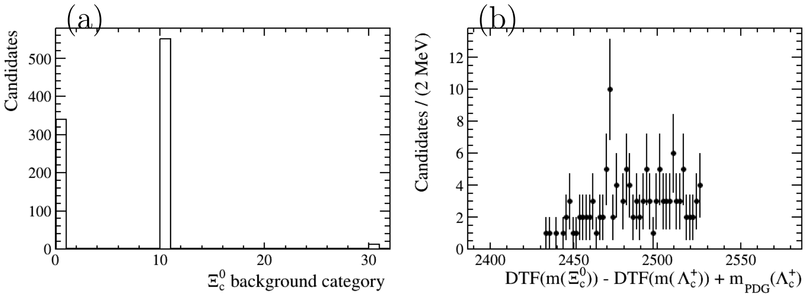


Figure 45: Plots from the $\Xi_c^0 \to \pi^-$ ($\Lambda_c^+ \to pK^-\pi^+$) simulation dataset: (a) The background category of other candidates appearing in an event with a candidate with background category 30 and (b) the mass distribution of background category 30 candidates which are not identified with a background category 0 or 10 candidate in the same event. None of these plots has any of the weighting described in Section 4.3.4 applied.

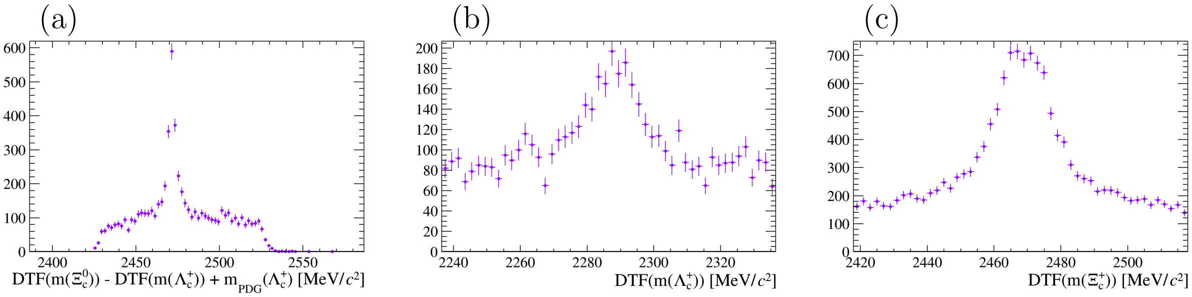


Figure 46: Mass distributions for (a) $\Xi_c^0 \to \pi^-$ ($\Lambda_c^+ \to pK^-\pi^+$), (b) $\Lambda_c^+ \to pK^-\pi^+$, and (c) $\Xi_c^+ \to pK^-\pi^+$ decays in simulation, background category 60 (ghosts) only. None of these plots has any of the weighting described in Section 4.3.4 applied.

A.2 BDT response in each fold

We use a 10-fold BDT-technique to maximize the size of the datasets available for training and testing (see Section 4.3.5). We present the responses of the BDTs in each fold in Figure 47, and we present the overtraining checks for each BDT in Figures 48, 49, and 50. The response in each fold strongly agrees (within errors) with the responses in each other fold, and the responses in the training and testing samples strongly agree (within errors).

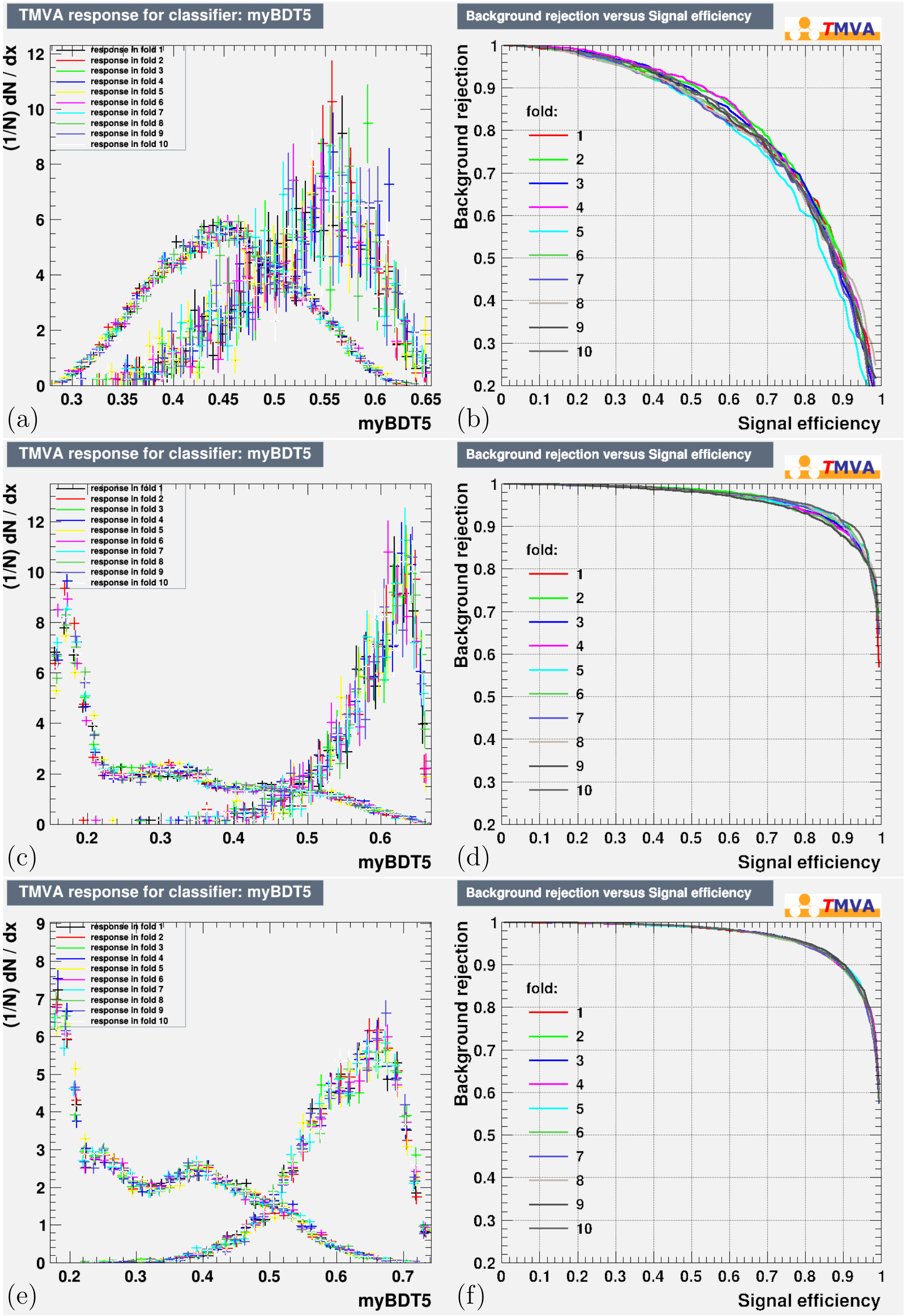
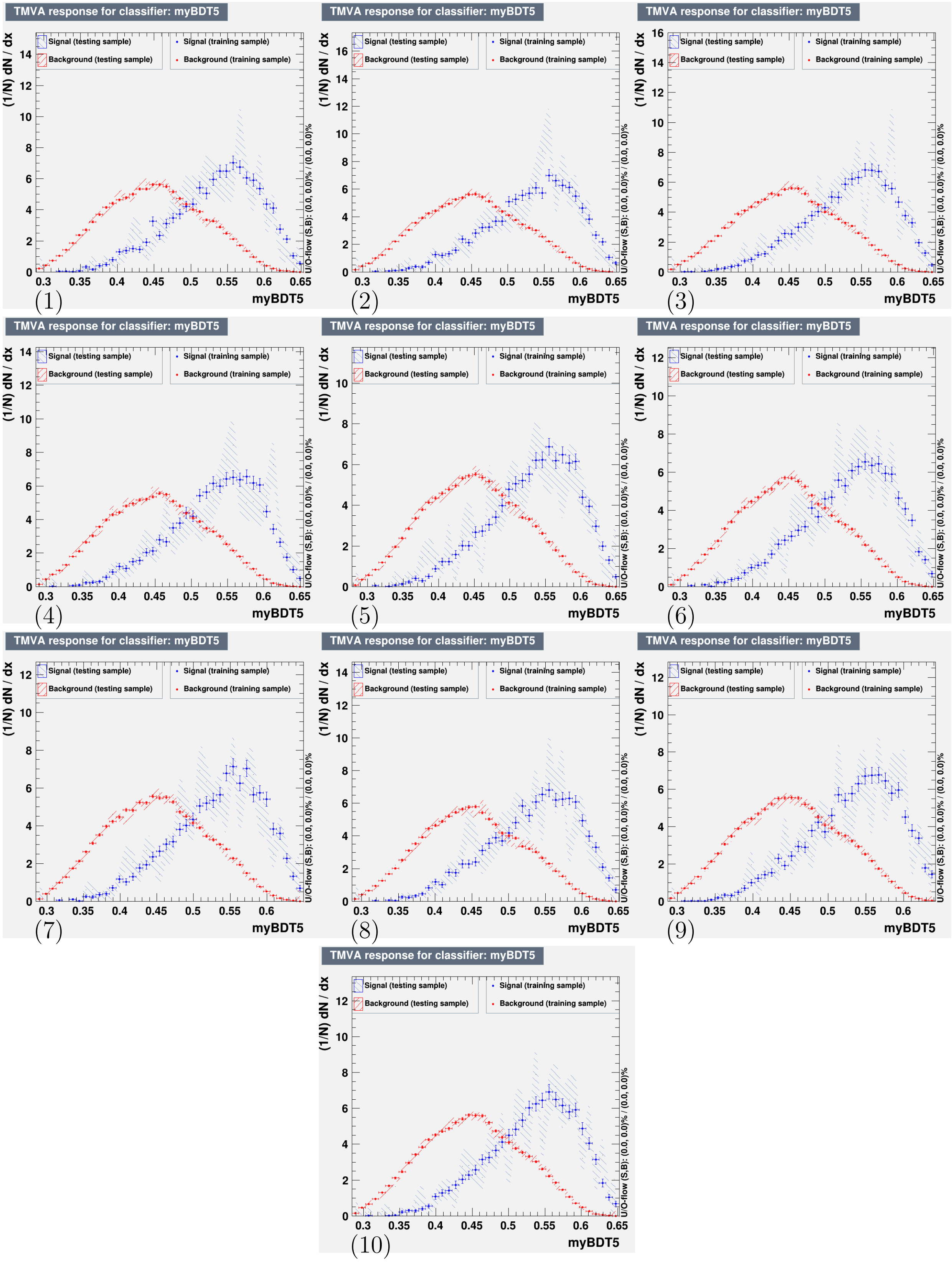
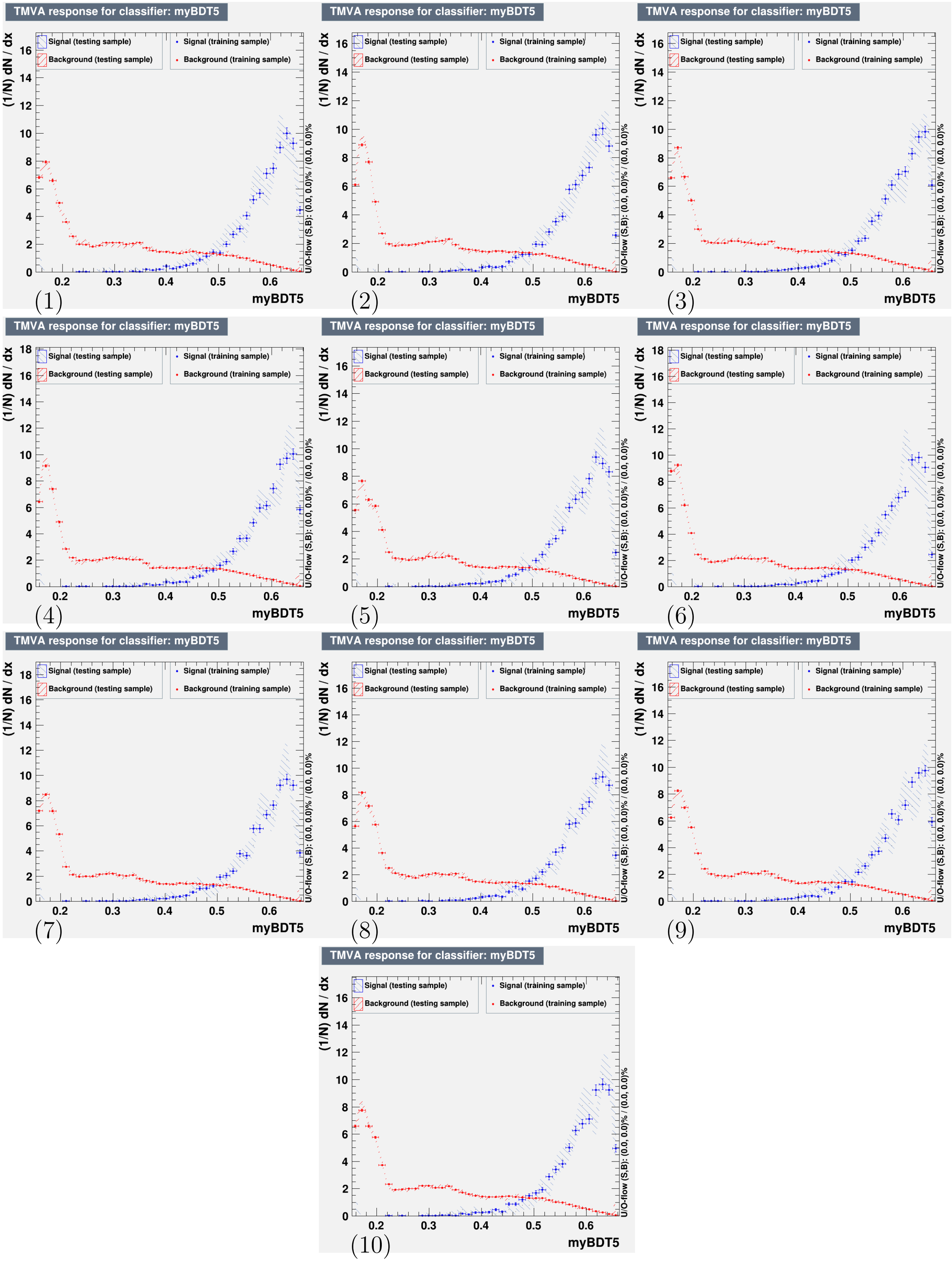


Figure 47: BDT performance in each fold, as demonstrated by (left) its response and (right) its receiver operating characteristic (ROC) curve for (a & b) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$, (c & d) $\Lambda_c^+ \to pK^-\pi^+$, and (e & f) $\Xi_c^+ \to pK^-\pi^+$.



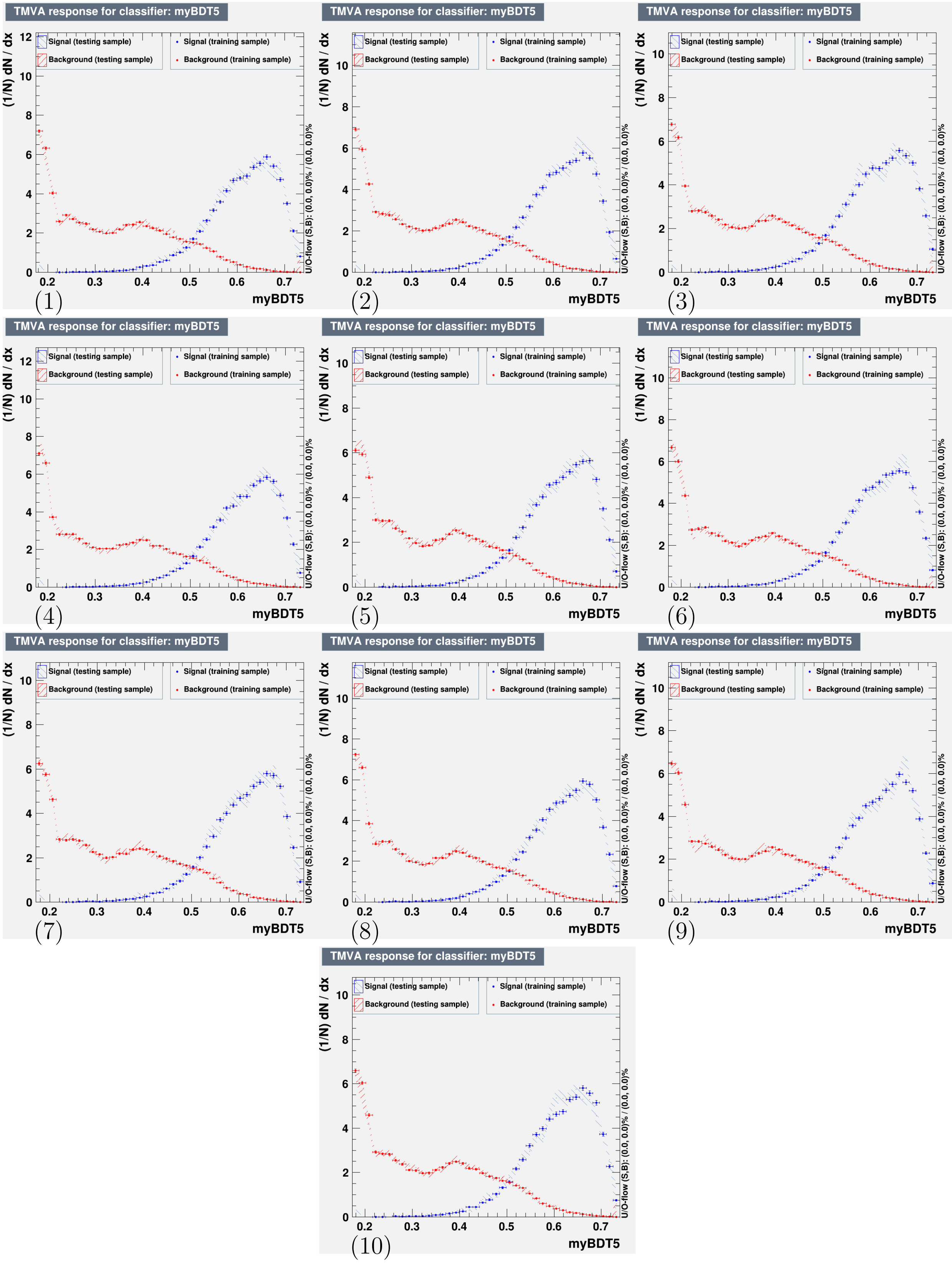
(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

Figure 48: Overtraining check for the $\Xi_c^0 \to \pi^- \; (\Lambda_c^+ \to pK^-\pi^+)$ BDT in each of its 10 folds.



(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

Figure 49: Overtraining check for the $\Lambda_c^+ \to pK^-\pi^+$ BDT in each of its 10 folds.



(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

Figure 50: Overtraining check for the $\Xi_c^+ \to pK^-\pi^+$ BDT in each of its 10 folds.

A.3 BDT input variable rankings

Table 20 shows the discrimination power of the input variables for each BDT.

Table 20: BDT input parameters, ranked according to signal, background separation. See Table 9. The input distributions can be found in Appendix A.4.

Rank	Variable	Separation
	$\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$	
1	$\mathrm{ProbNN}_\pi(\pi^-)$	1.473×10^{-1}
2	$\ln(\chi^2_{\mathrm{IP}}(\pi^-))$	4.188×10^{-2}
3	$\ln(\chi^2_{\mathrm{IP}}(\Xi_c^0))$	3.965×10^{-2}
4	$\mathrm{ProbNN}_K(K^-)$	3.491×10^{-2}
5	$\mathrm{ProbNN}_p(p)$	3.084×10^{-2}
6	$\ln(\chi^2_{\mathrm{IP}}(\Lambda_c^+))$	2.853×10^{-2}
7	$\ln\left(\arccos\left(\frac{\vec{x}(\Lambda_c^+)\cdot\vec{p}(\Lambda_c^+)}{x(\Lambda_c^+)\cdot p(\Lambda_c^+)}\right)\right)$	2.678×10^{-2}
8	$\ln(\mathrm{DTF}\,(\tau\,(\Xi_c^0)))$	2.488×10^{-2}
9	$\ln(\chi^2/\mathrm{ndf}(\mathrm{DTF}\,(\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+))))$	2.422×10^{-2}
10	$\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(\Xi_c^0)\cdot\vec{p}(\Xi_c^0)}{x(\Xi_c^0)\cdot p(\Xi_c^0)}\right)\right)\right)$	2.175×10^{-2}
11	$\mathrm{ProbNN}_\pi(\pi^+)$	9.168×10^{-3}
12	$\ln(\lVert\vec{x}(\Lambda_c^+) - \vec{x}(\Xi_c^0)\rVert)$	5.198×10^{-3}
	$\Lambda_c^+ \to pK^-\pi^+$	
1	$\ln(\chi^2_{\mathrm{IP}}(\Lambda_c^+))$	4.984×10^{-1}
2	$\ln(\chi^2/\mathrm{ndf}(\mathrm{DTF}\,(\Lambda_c^+ \to pK^-\pi^+)))$	4.534×10^{-1}
3	$\ln(\mathrm{DTF}\,(\tau\,(\Lambda_c^+)))$	4.118×10^{-1}
4	$\mathrm{ProbNN}_p(p)$	2.387×10^{-1}
5	$\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(\Lambda_c^+)\cdot\vec{p}(\Lambda_c^+)}{x(\Lambda_c^+)\cdot p(\Lambda_c^+)}\right)\right)\right)$	9.819×10^{-2}
6	$\mathrm{ProbNN}_\pi(\pi^+)$	6.251×10^{-2}
7	$\mathrm{ProbNN}_K(K^-)$	5.096×10^{-2}
	$\Xi_c^+ \to pK^-\pi^+$	
1	$\ln(\chi^2_{\mathrm{IP}}(\Xi_c^+))$	4.897×10^{-1}
2	$\ln(\chi^2/\mathrm{ndf}(\mathrm{DTF}\,(\Xi_c^+ \to pK^-\pi^+)))$	4.709×10^{-1}
3	$\mathrm{ProbNN}_p(p)$	2.684×10^{-1}
4	$\ln(\mathrm{DTF}\,(\tau\,(\Xi_c^+)))$	1.931×10^{-1}
5	$\ln\left(\arccos\left(\mathrm{DTF}\left(\frac{\vec{x}(\Xi_c^+)\cdot\vec{p}(\Xi_c^+)}{x(\Xi_c^+)\cdot p(\Xi_c^+)}\right)\right)\right)$	1.484×10^{-1}
6	$\mathrm{ProbNN}_\pi(\pi^+)$	7.170×10^{-2}
7	$\mathrm{ProbNN}_K(K^-)$	5.117×10^{-2}

A.4 BDT input parameters

Figure 51 shows the training distributions for each BDT's input parameters. (See Section 4.3.5.)

Figure 52 shows the signal distributions of each of the BDT input parameters, as listed in Section 4.3.5, in both simulation and side-band subtracted data, excluding the PID parameters, which are shown in Figure 53. The signal distributions in data are the result of a side-band subtraction using the fits shown in Figure 54. The distributions in the $\mathit{\Lambda}_c^+$ and $\mathit{\Xi}_c^+$ data and simulation datasets do not align as well as those for the $\mathit{\Xi}_c^0$ datasets; this is due to the presence of secondaries in the $\mathit{\Lambda}_c^+$ and $\mathit{\Xi}_c^+$ data datasets, which are a source of background not present in the simulation datasets (see Section 4.7.11). The distributions shown are restricted to the signal p_{T} and η range (see Section 4.4.2).

Figure 55 shows the signal distributions of each of the BDT input parameters, as listed in Section 4.3.5, excluding the PID parameters, *after the BDT has been applied* in both simulation and signal data, as determined by the *sPlot* method [16]. Notice the improved alignment between $\mathit{\Lambda}_c^+$ and $\mathit{\Xi}_c^+$ data and simulation datasets as secondaries are eliminated. Figure 56 shows the BDT distributions. Figures 24, 25, and 26 show the fits used for the *sPlot* method [16]. The distributions shown are restricted to the signal p_{T} and η range (see Section 4.4.2).

We use a wrong-sign (WS) $\mathit{\Xi}_c^0$ dataset as the background input for the BDT training (see Section 4.3.5). Figure 57 demonstrates that the WS distributions are sufficiently similar to those of right-sign background.

We use side-band data as background inputs for the $\mathit{\Lambda}_c^+$ and $\mathit{\Xi}_c^+$ BDT training (see Section 4.3.5). Figure 58 demonstrates that the low-mass and high-mass side-bands have similar BDT responses.

(a)

(b)

(c)

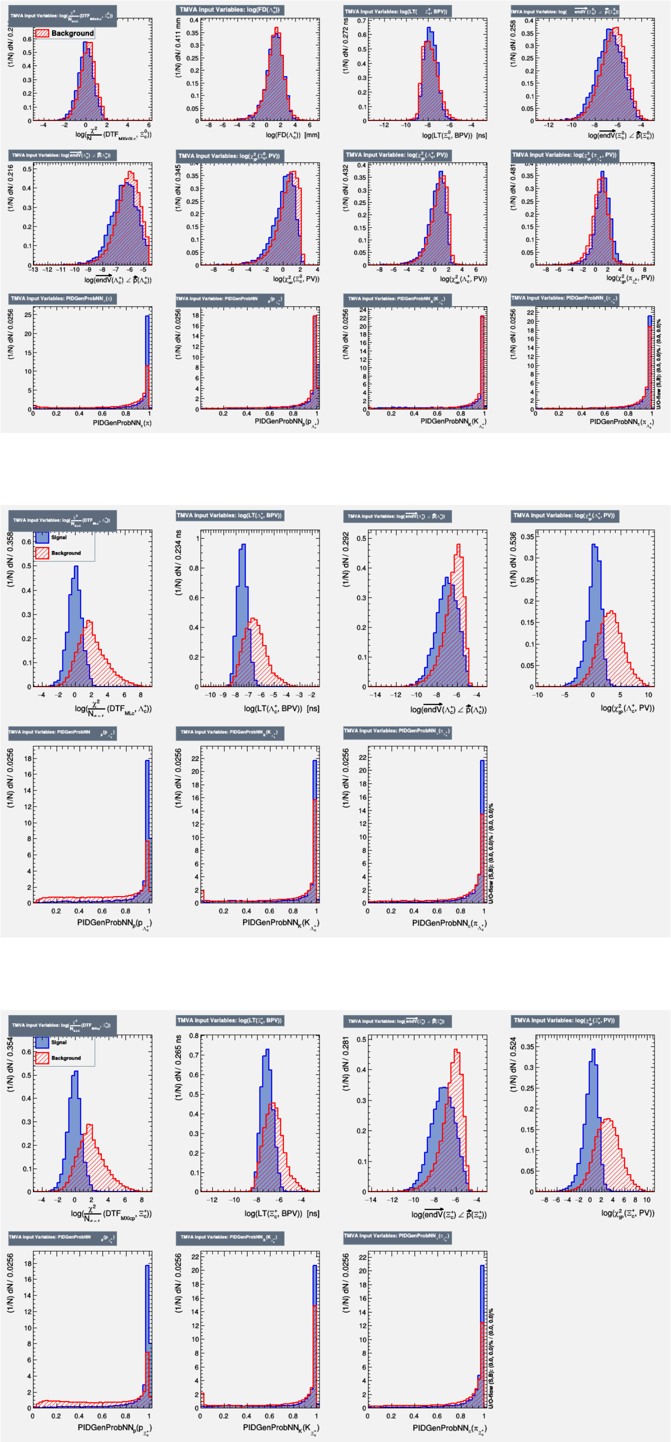


Figure 51: BDT input variable distributions (after the application of offline and trigger selections) for the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ BDTs. Solid blue shows the distributions in the signal sample; hatched red shows that in the background sample.

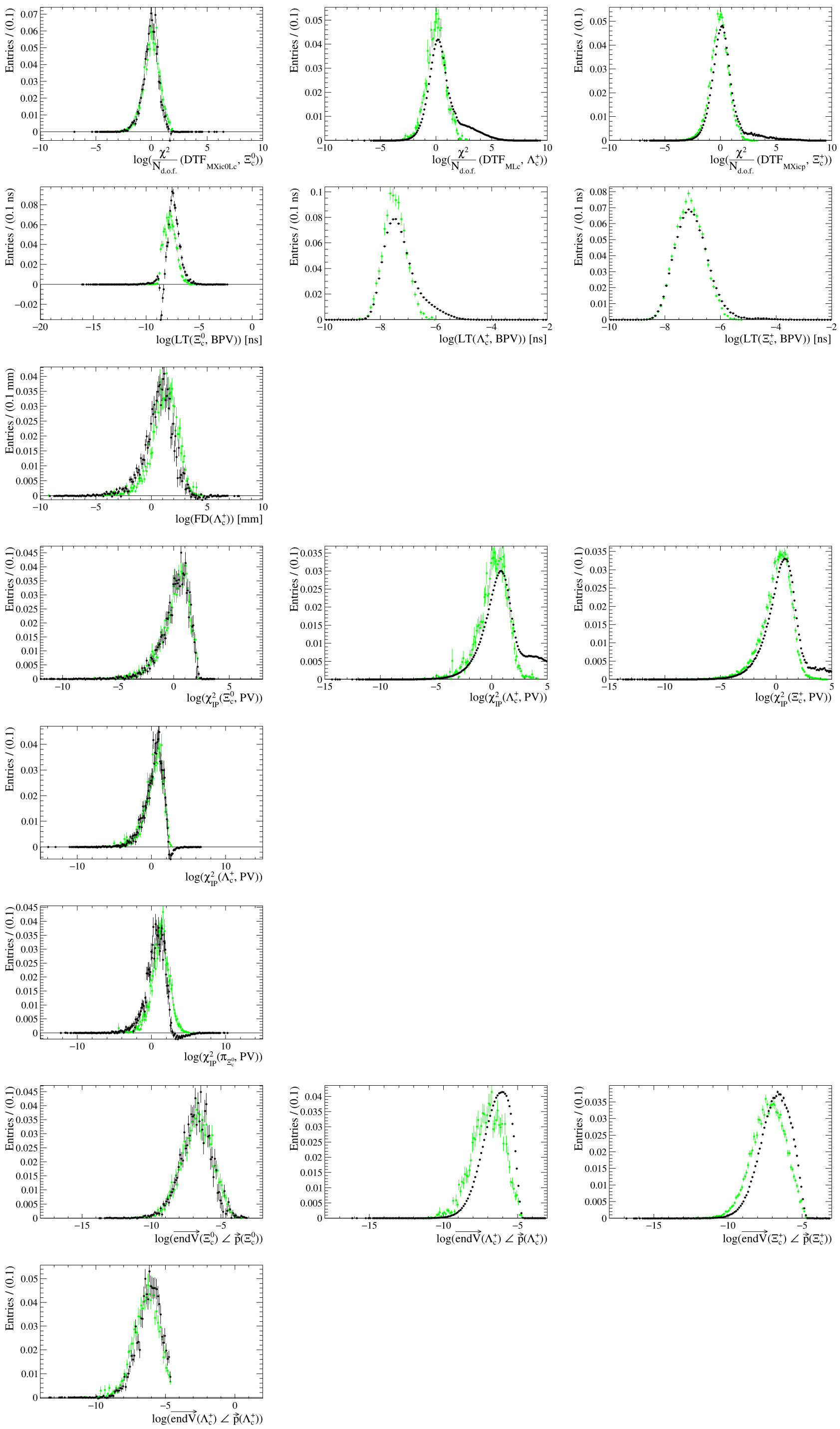


Figure 52: Input distributions (after the application of the offline and trigger selections) for the (left) Ξ_c^0, (center) Λ_c^+, and (right) Ξ_c^+ BDTs. The (black) data distributions are the result of a side-band subtraction using the fits shown in Figure 54 (c & d). The (green) simulation distributions are truth matched.

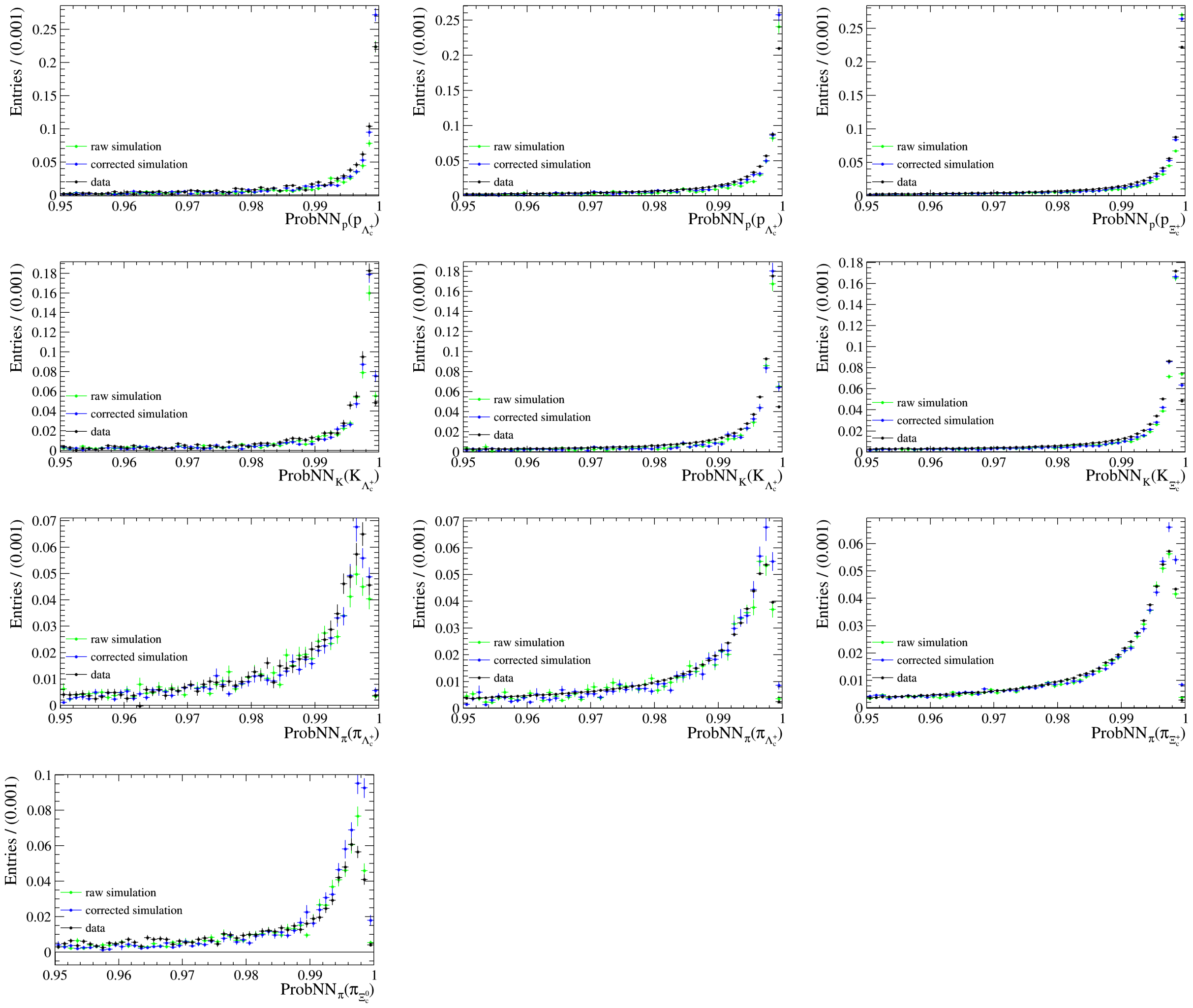


Figure 53: The output of a probablistic neural-network determining how likely a given final-state daughter has been assigned the correct mass after Turbo, offline, and trigger selections have been applied for (left) Ξ_c^0, (middle) Λ_c^+, and (right) Ξ_c^+ data and simulation datasets. The (black) data distribution is the result of a side-band subtraction using the fits shown in Figure 54, where the signal region corresponds to 99.7300% of the total integral symetric about the mean and the background region is taken from 99.9937% to 99.9999%. The (green) raw simulation distribution shows the uncorrected PID distribution, while the (blue) corrected simulation distribution shows the PID distribution produced by PIDGen. The simulation distributions have been truth matched.

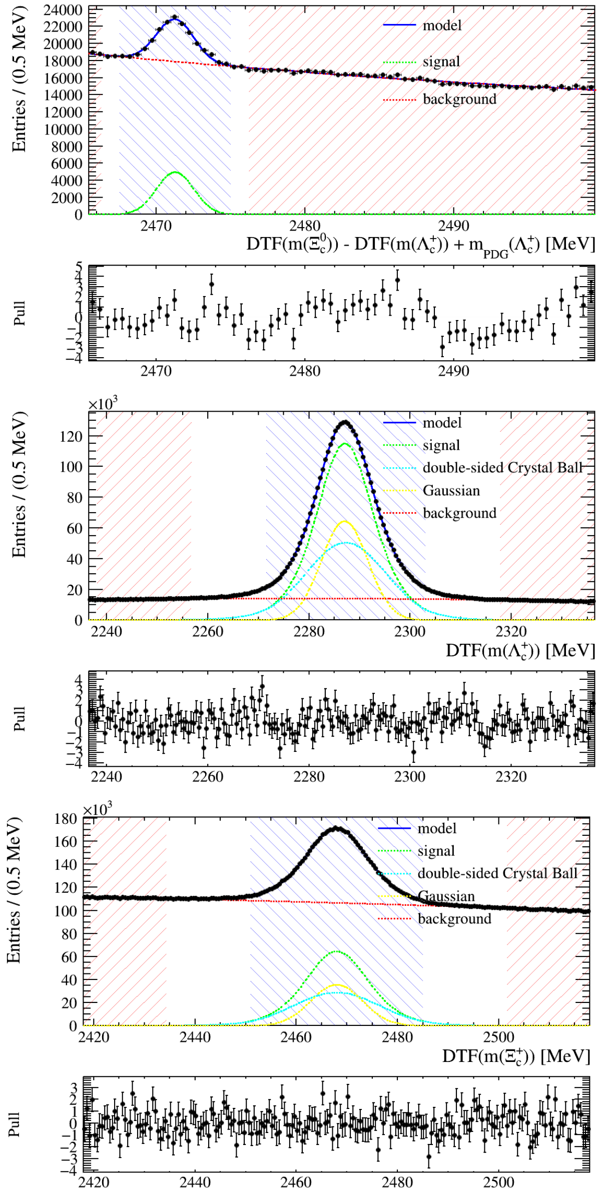


Figure 54: Binned fits to the relevant invariant masses in the (top) $\mathit{\Xi}_c^0$, (middle) $\mathit{\Lambda}_c^+$, and (bottom) $\mathit{\Xi}_c^+$ data datsets, using the shapes described in Section 4.5.1, after the BDT input selections have been applied. The blue (red) regions are considered signal (background) for sideband subtraction; the $\mathit{\Xi}_c^0$ signal region is taken from 2467.5 MeV to 2475.0 MeV, and its background region is taken from 2465.5 MeV to 2466.3 MeV and from 2476.2 MeV to 2499.5 MeV; the $\mathit{\Lambda}_c^+$ signal region is taken from 2271.5 MeV to 2303.0 MeV, and its background region is taken from 2238.1 MeV to 2256.7 MeV and from 2317.8 MeV to 2336.5 MeV; the $\mathit{\Xi}_c^+$ signal region is taken from 2450.9 MeV to 2485.0 MeV, and its background region is taken from 2418.0 MeV to 2434.3 MeV and from 2501.6 MeV to 2517.9 MeV;

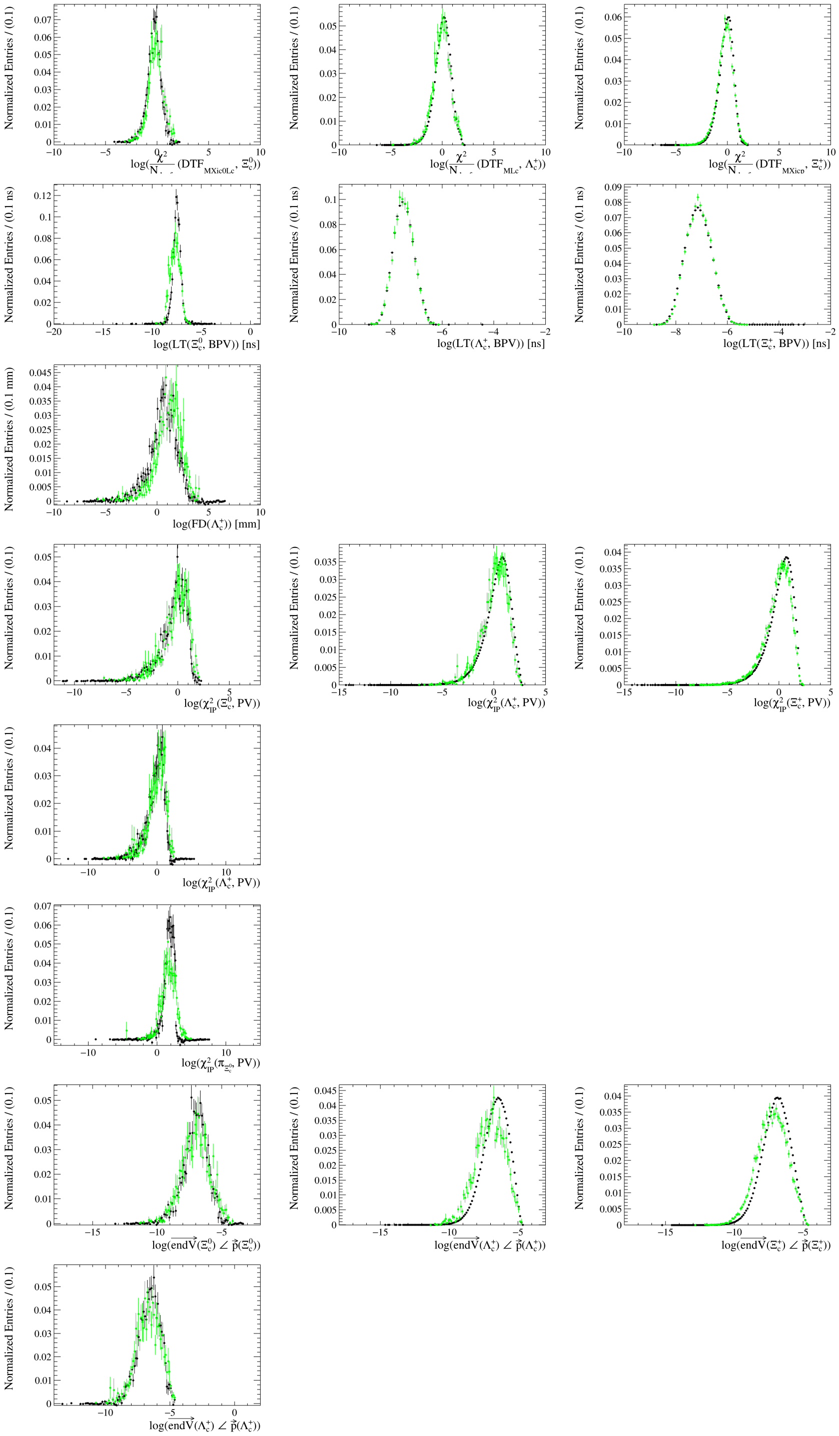


Figure 55: BDT input parameter distributions after the application of the BDTs for the (left) Ξ_c^0, (center) Λ_c^+, and (right) Ξ_c^+ BDTs. The (black) data distributions are the result of the application of the *sPlot* method [16] using the fits shown in Figures 24, 25, and 26. The (green) simulation distributions are truth matched.

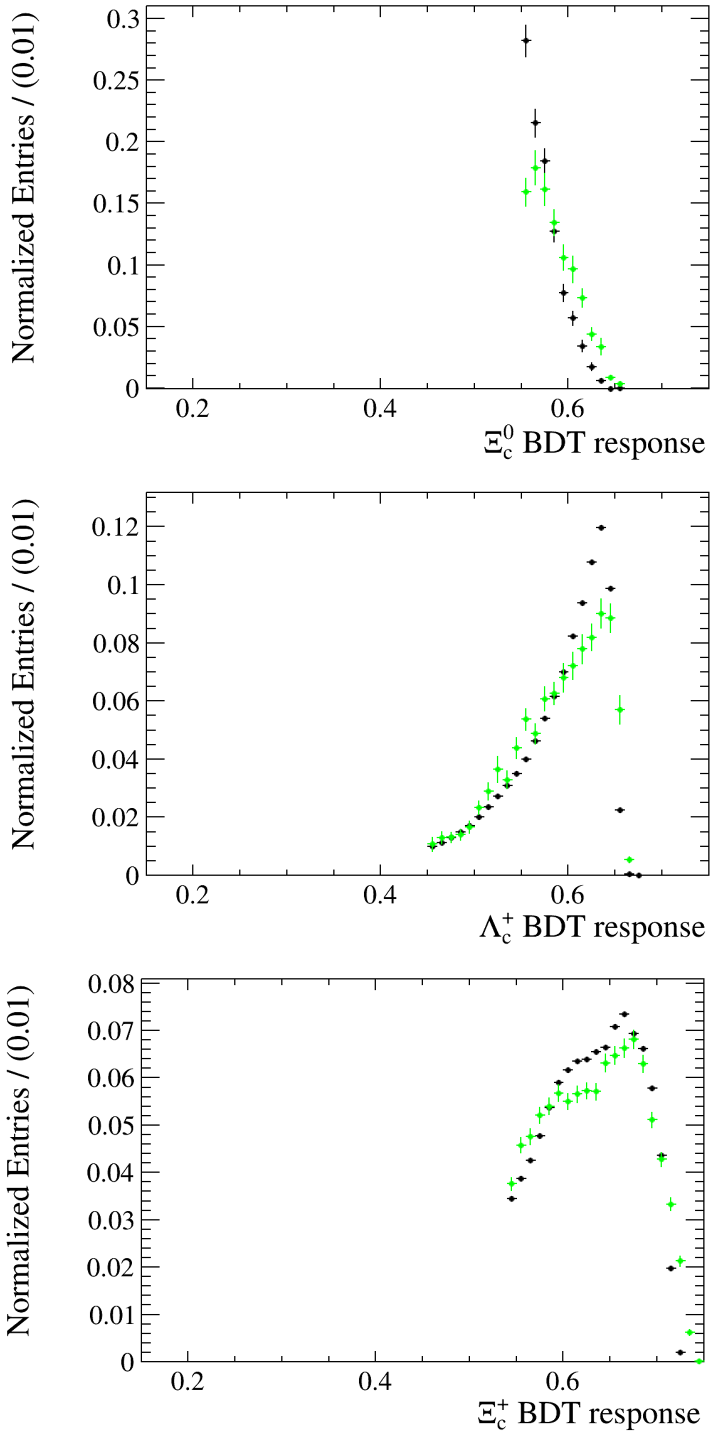


Figure 56: BDT response distributions after the application of the BDTs for the (left) Ξ_c^0, (center) Λ_c^+, and (right) Ξ_c^+ BDTs. The (black) data distributions are the result of the application of the *sPlot* method [16] using the fits shown in Figures 24, 25, and 26. The (green) simulation distributions are truth matched. These plots do not appear in Reference [17].

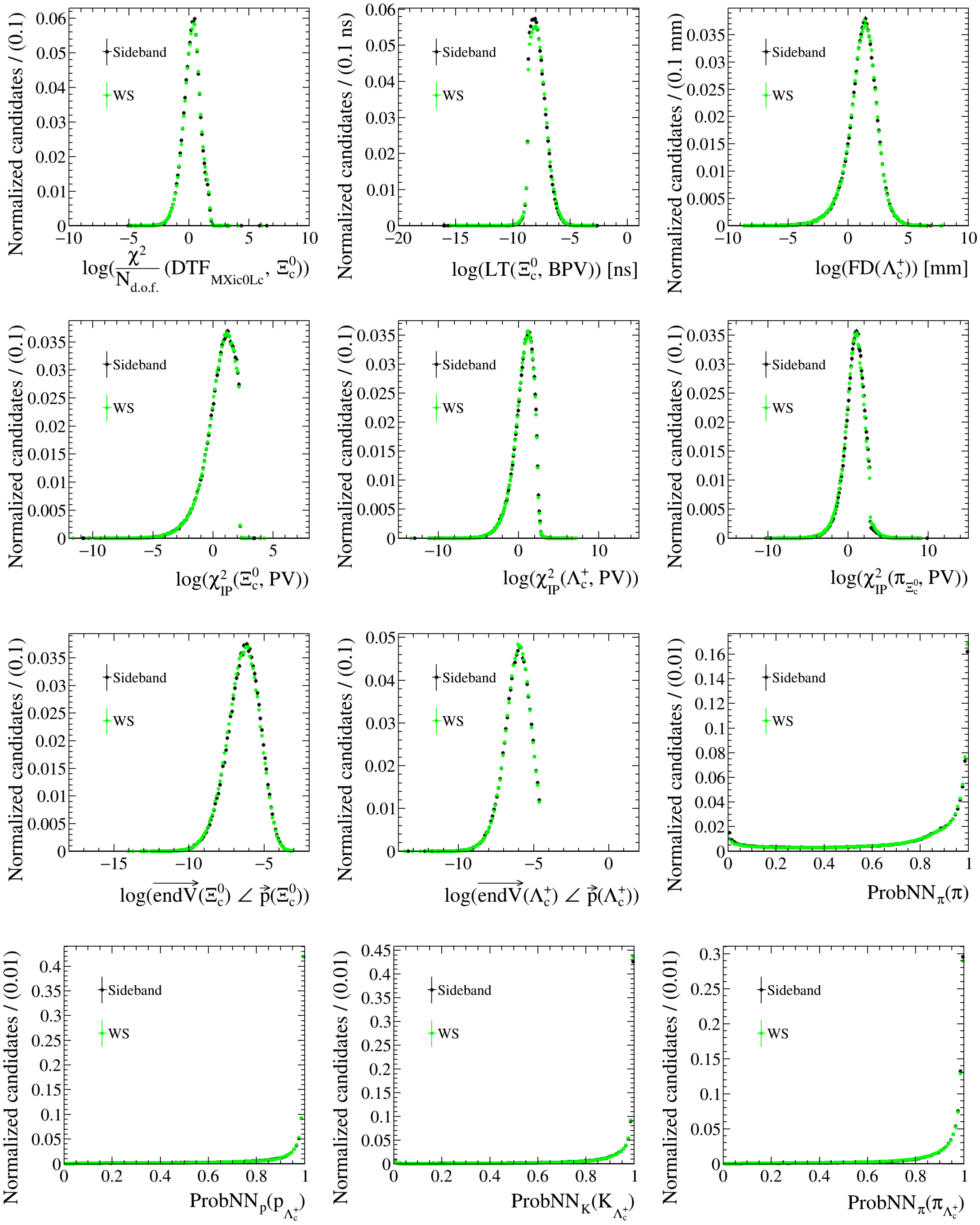


Figure 57: BDT input distributions (after the application of the Turbo, offline, and trigger selections) for the (black) side-band and (green) WS Ξ_c^0 data distributions. The sidebands are taken from 2465.46 MeV to 2466.46 MeV and from 2475.46 to 2482.46 MeV in the DTF $\left(m\left(\Xi_c^0\right)\right)$ −DTF $(m\,(\Lambda_c^+))$ $+m_{\text{PDG}}\,(\Lambda_c^+)$ distribution, and the WS are restricted to 2465.87 MeV to 2475.87 MeV.

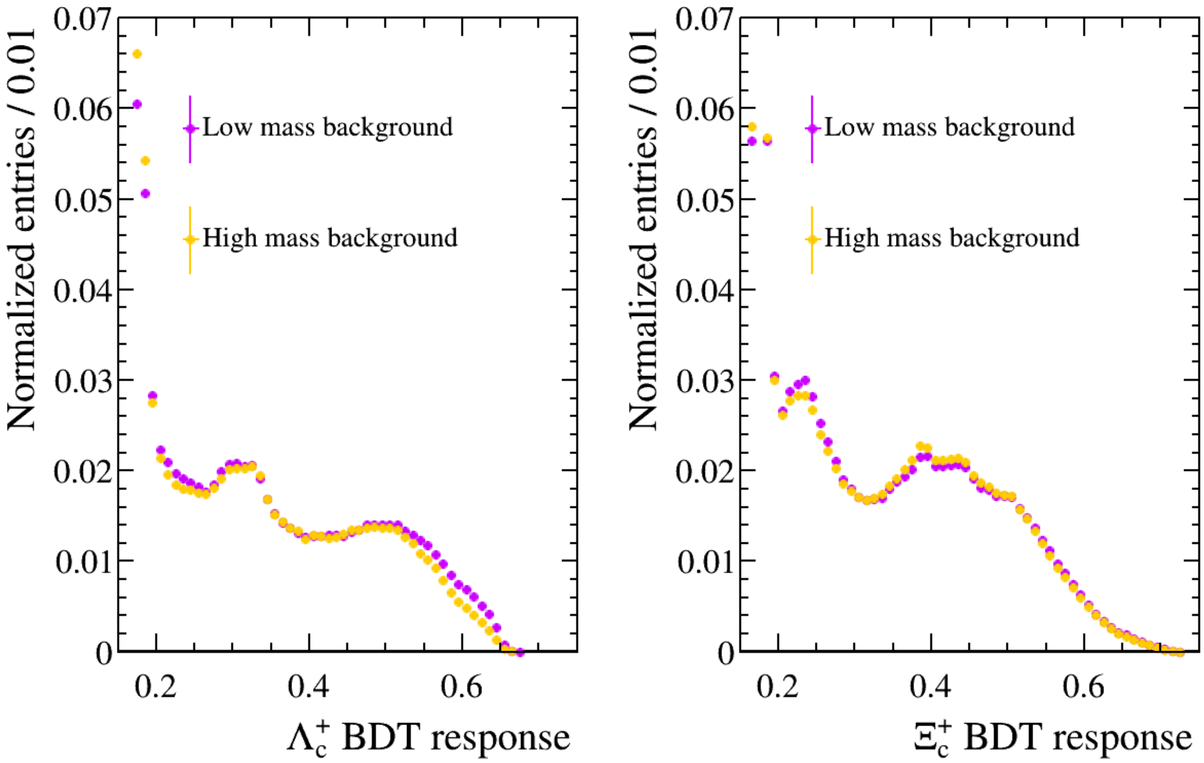


Figure 58: BDT input distributions (after the application of the Turbo, offline, and trigger selections) for the (purple) low-mass and (orange) high-mass side-band (left) Λ_c^+ and (right) Ξ_c^+ data distributions. The side-bands are described in Section 4.3.5.

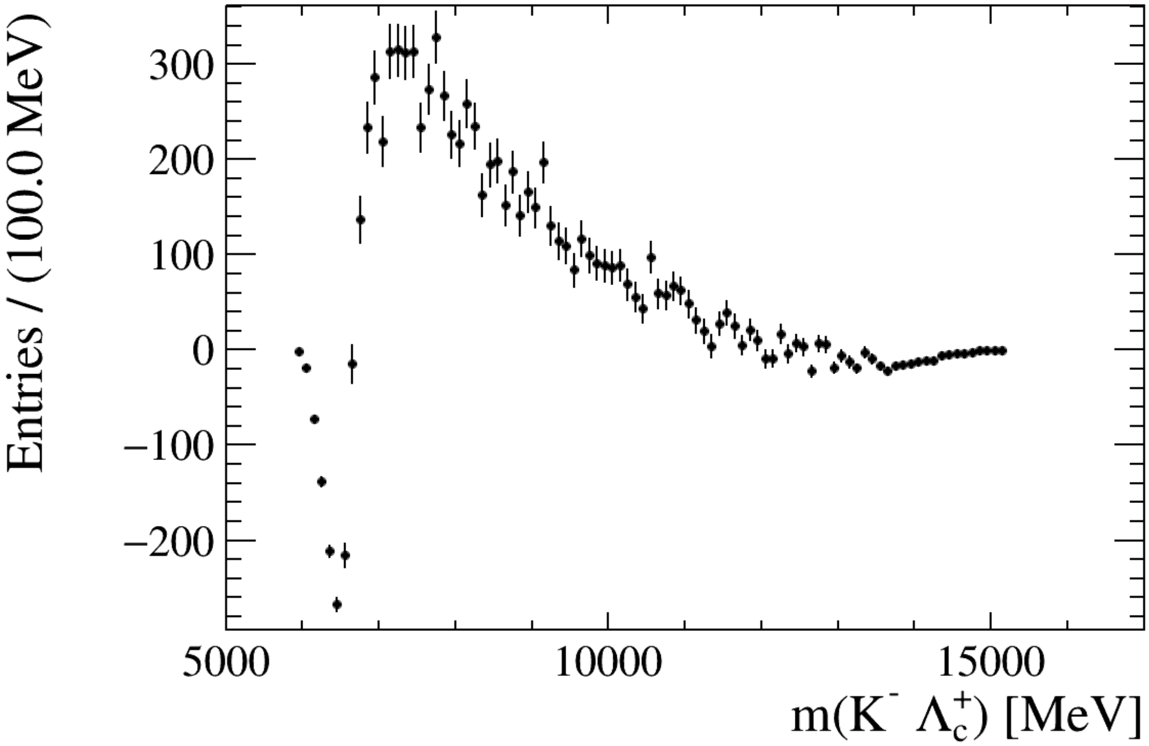


Figure 59: Invariant mass of the Ξ_c^0 baryon's daughters in data with the mass hypothesis of its daughter π^- meson changed to a K^- meson. (DTF is not applied.)

A.5 Wrong-mass discrimination

There exists the possibility of picking up background decays through particle misidentification. To ensure this is not the case, we apply the *sPlot* method [16] using the signal fits shown in Section 4.5.1 and then plot mass distributions with the various mass hypotheses of the final-state daughters changed. Figures 59, 60, 61, and 62 show the results; we do not find evidence of significant contribution to the signal selection from background resonances.

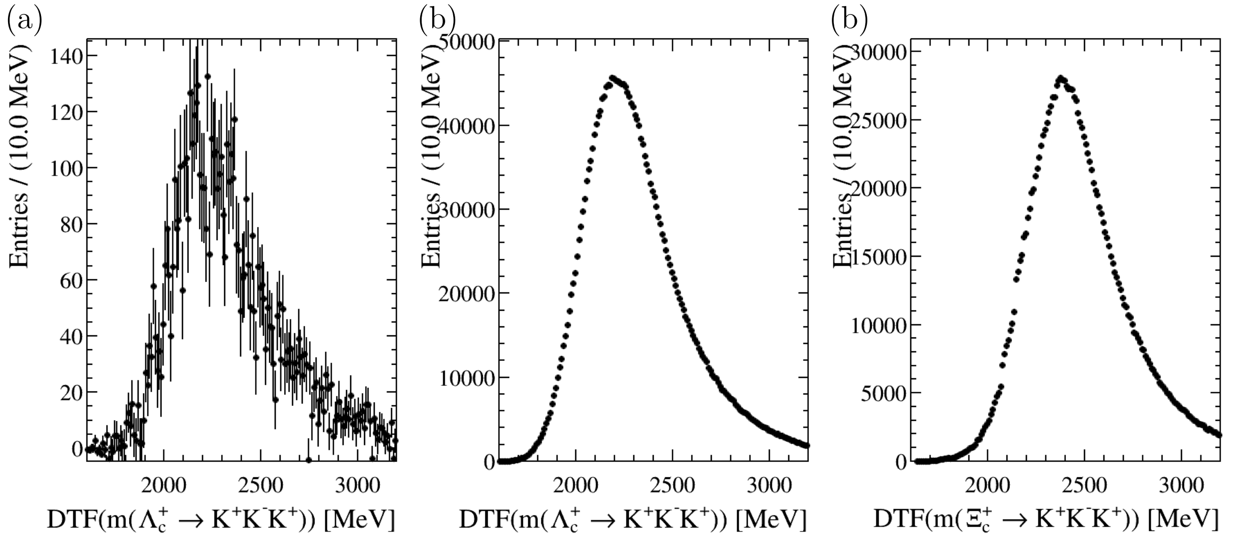


Figure 60: Invariant mass of the Λ_c^+ or Ξ_c^+ baryon's daughters in the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ data datasets with the mass hypothesis of the proton and the π^+ meson changed to a K^+ meson and DTF re-applied.

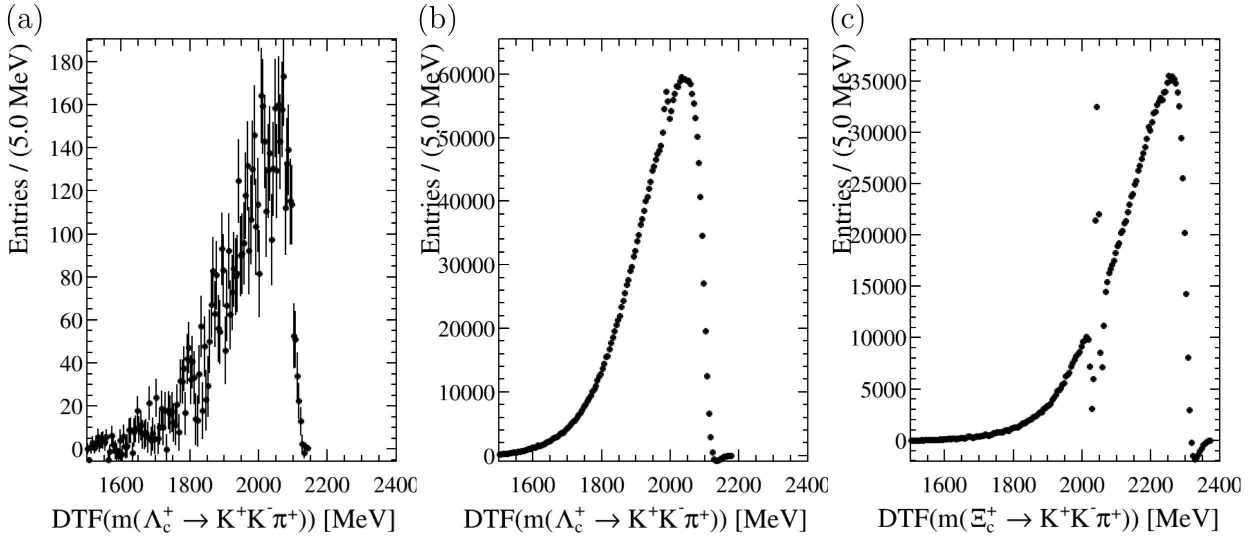


Figure 61: Invariant mass of the Λ_c^+ or Ξ_c^+ baryon's daughters in the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ data datasets with the mass hypothesis of the proton changed to a K^+ meson and DTF re-applied.

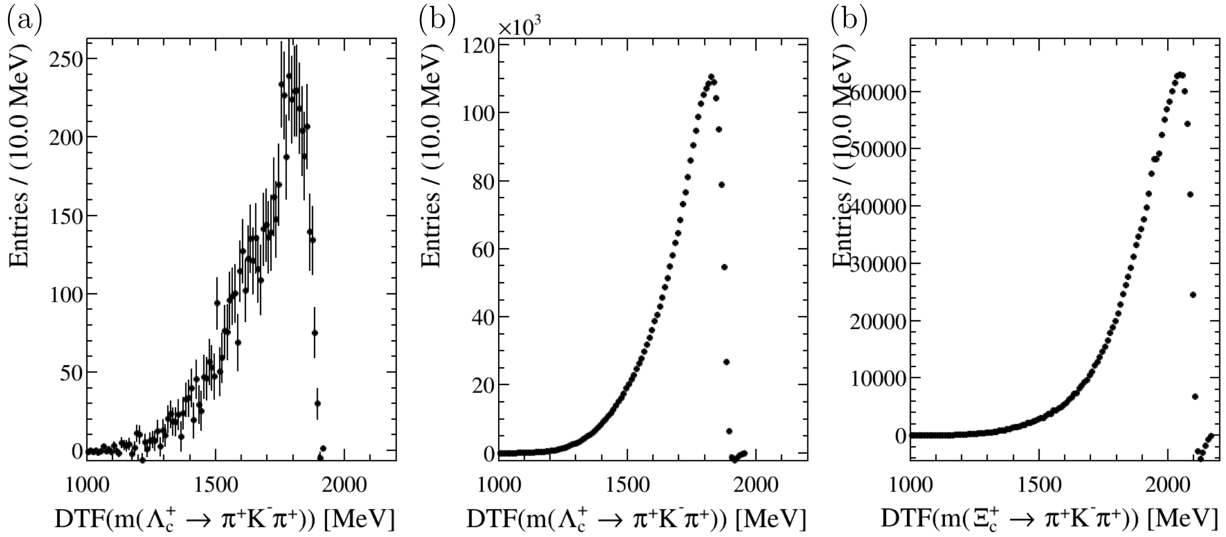


Figure 62: Invariant mass of the Λ_c^+ or Ξ_c^+ baryon's daughters in the (a) Ξ_c^0, (b) Λ_c^+, and (c) Ξ_c^+ data datasets with the mass hypothesis of the proton changed to a π^+ meson and DTF re-applied.

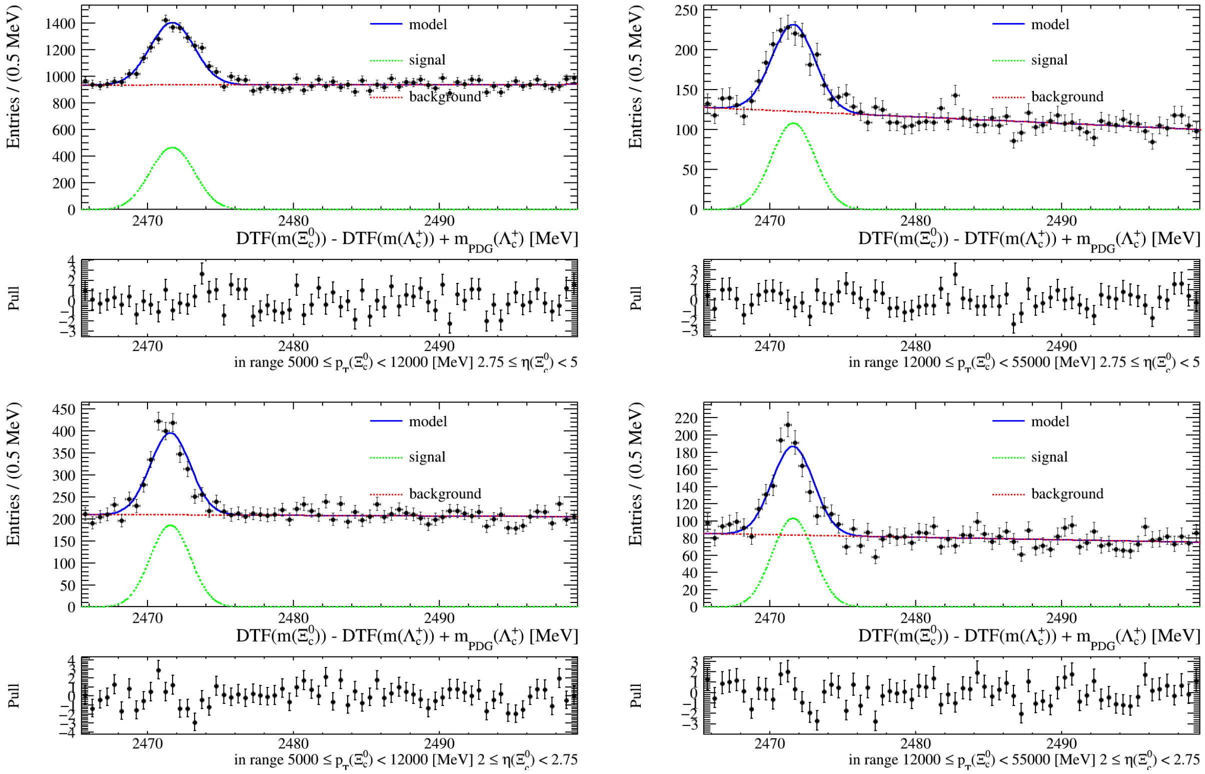


Figure 63: Nominal mass fits in bins of p_T and η in the $\Xi_c^0 \to \pi^-$ $(\Lambda_c^+ \to pK^-\pi^+)$ TOS dataset. The range is given in the x-axis title of the pull distribution of each plot.

A.6 Fits in each bin of p_T and η

As described in 4.5.2, we calculate the result by performing fits in regions of p_T and η. Figures 63, 66, 69, 64, 65, 67, 68, 70, 71, 72, 73, 74, 75, 76, and 77, show the fits in each bin.

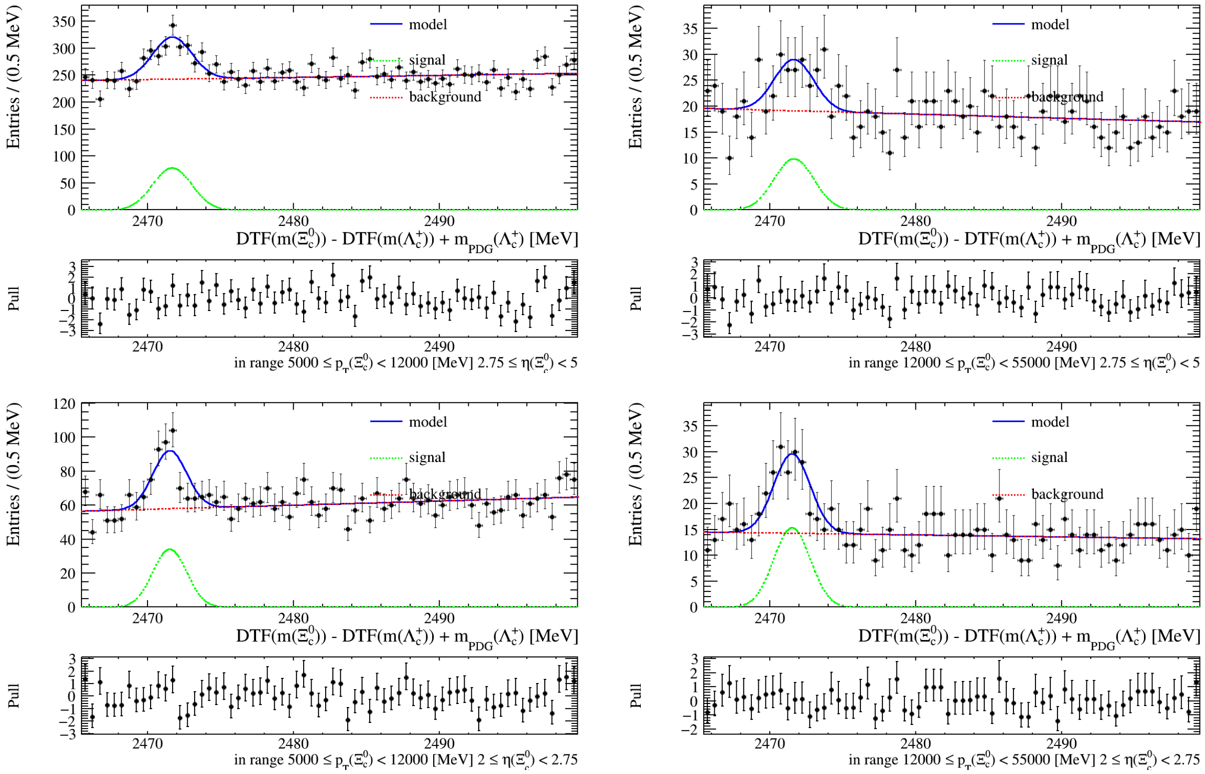


Figure 64: Nominal mass fits in bins of p_T and η in the $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$ TIS dataset. The range is given in the x-axis title of the pull distribution of each plot.

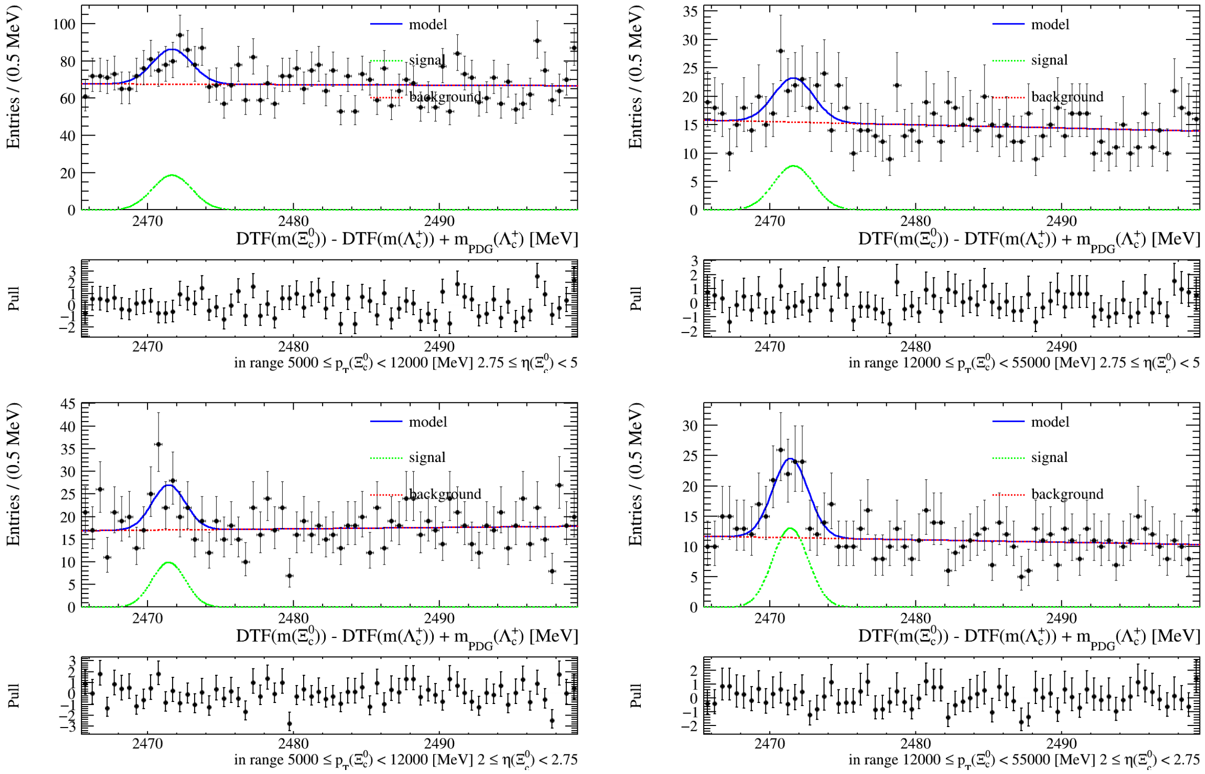


Figure 65: Nominal mass fits in bins of p_T and η in the $\Xi_c^0 \to \pi^- \, (\Lambda_c^+ \to pK^-\pi^+)$ TISTOS dataset. The range is given in the x-axis title of the pull distribution of each plot.

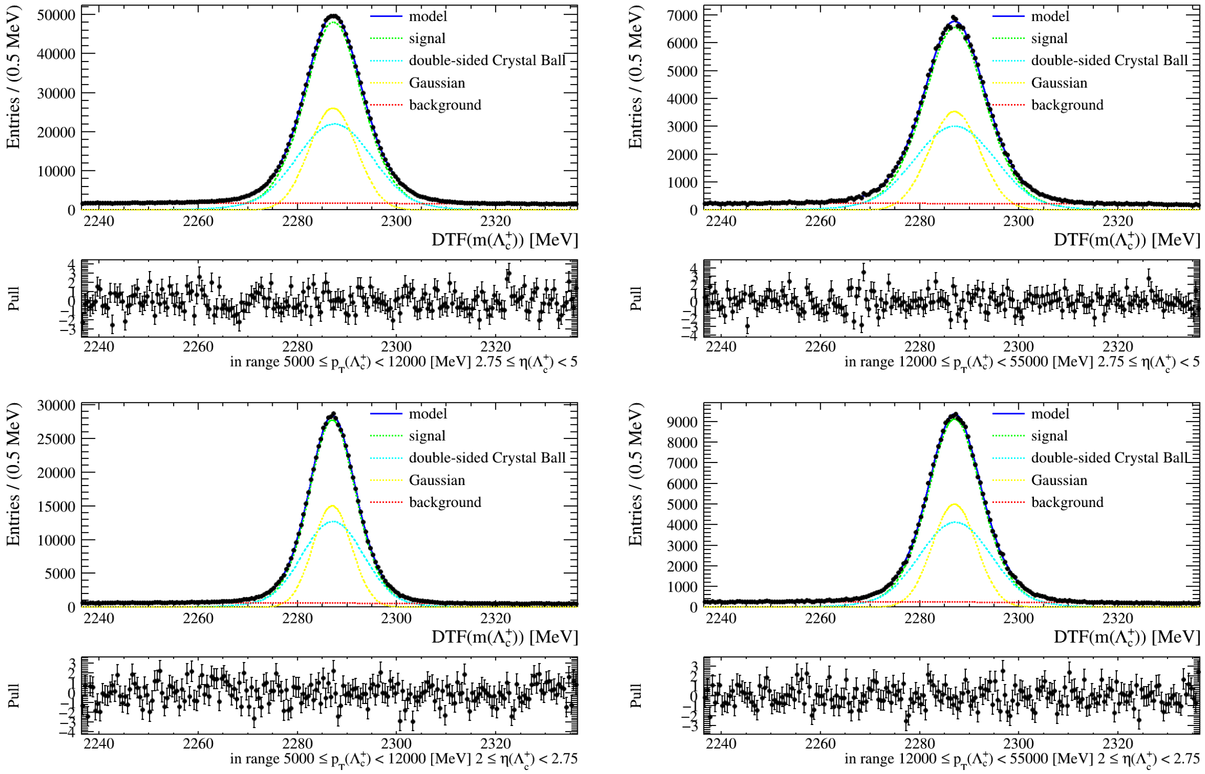


Figure 66: Nominal mass fits in bins of p_{T} and η in the $\Lambda_c^+ \to pK^-\pi^+$ TOS data dataset. The range is given in the x-axis title of the pull distribution of each plot.

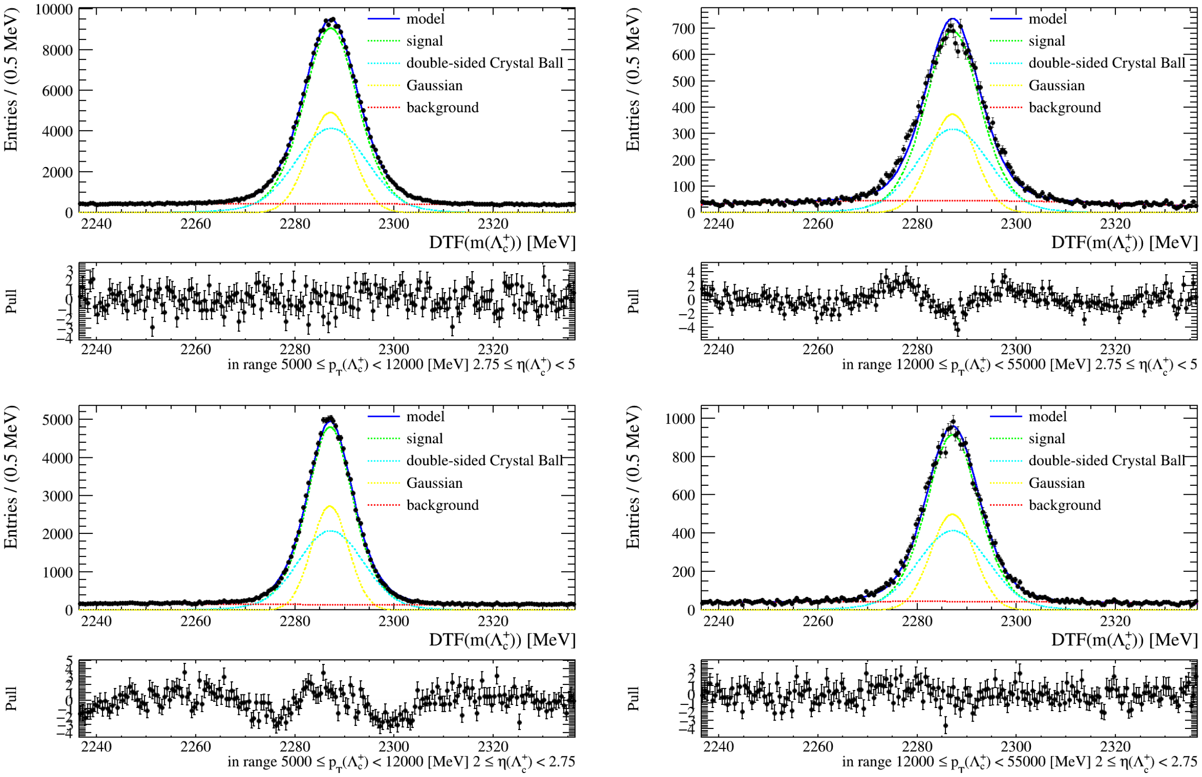


Figure 67: Nominal mass fits in bins of p_{T} and η in the $\Lambda_c^+ \to pK^-\pi^+$ TIS data dataset. The range is given in the x-axis title of the pull distribution of each plot.

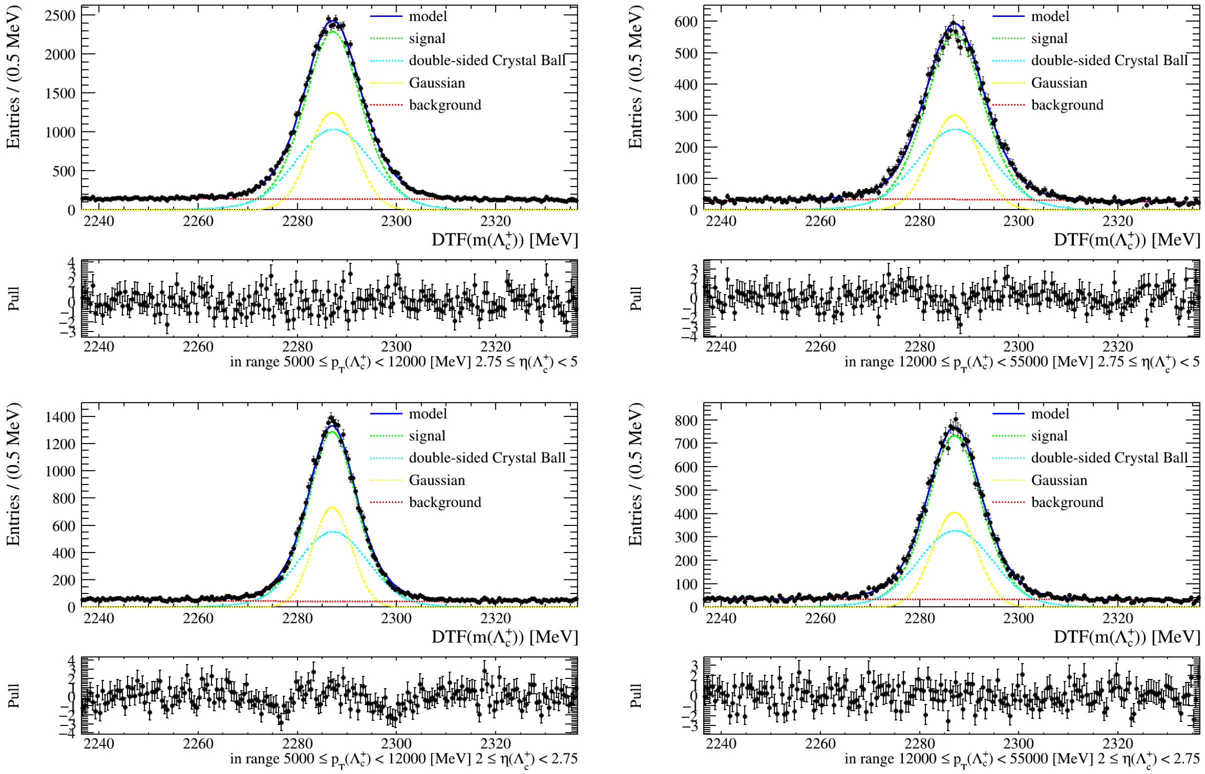


Figure 68: Nominal mass fits in bins of p_T and η in the $\Lambda_c^+ \to pK^-\pi^+$ TISTOS data dataset. The range is given in the x-axis title of the pull distribution of each plot.

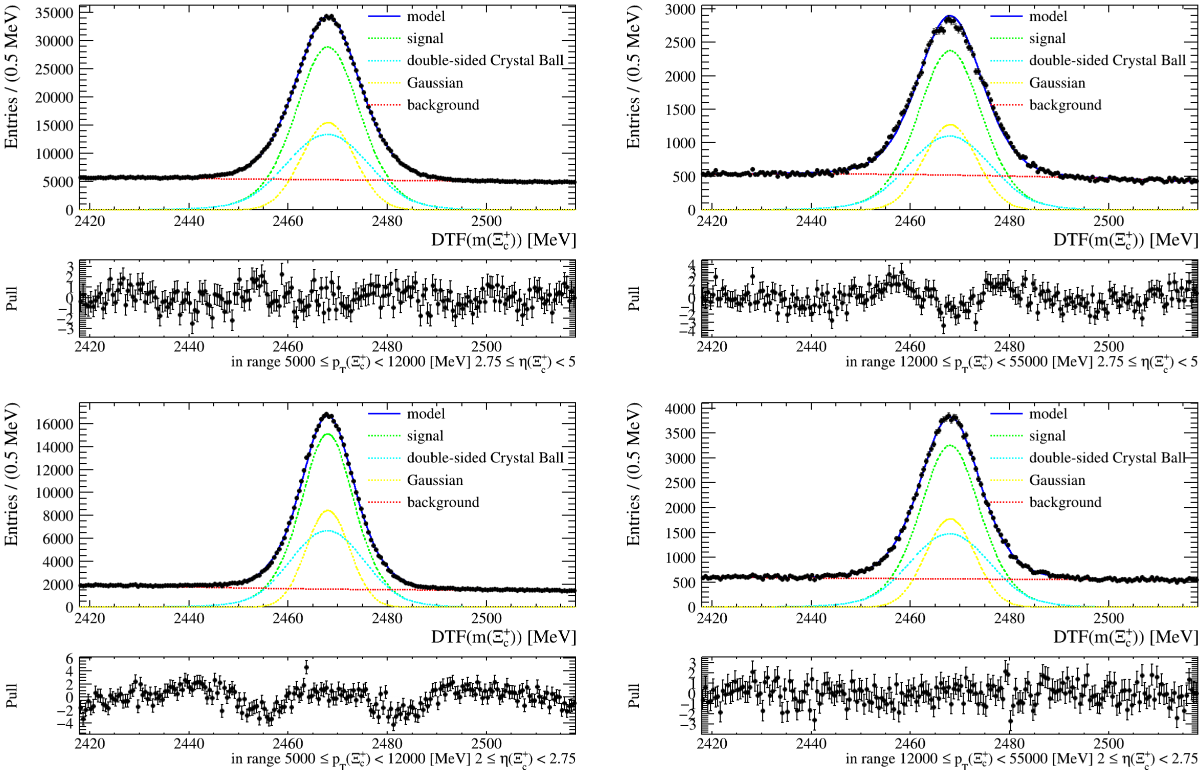


Figure 69: Nominal mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ TOS data dataset. The range is given in the x-axis title of the pull distribution of each plot.

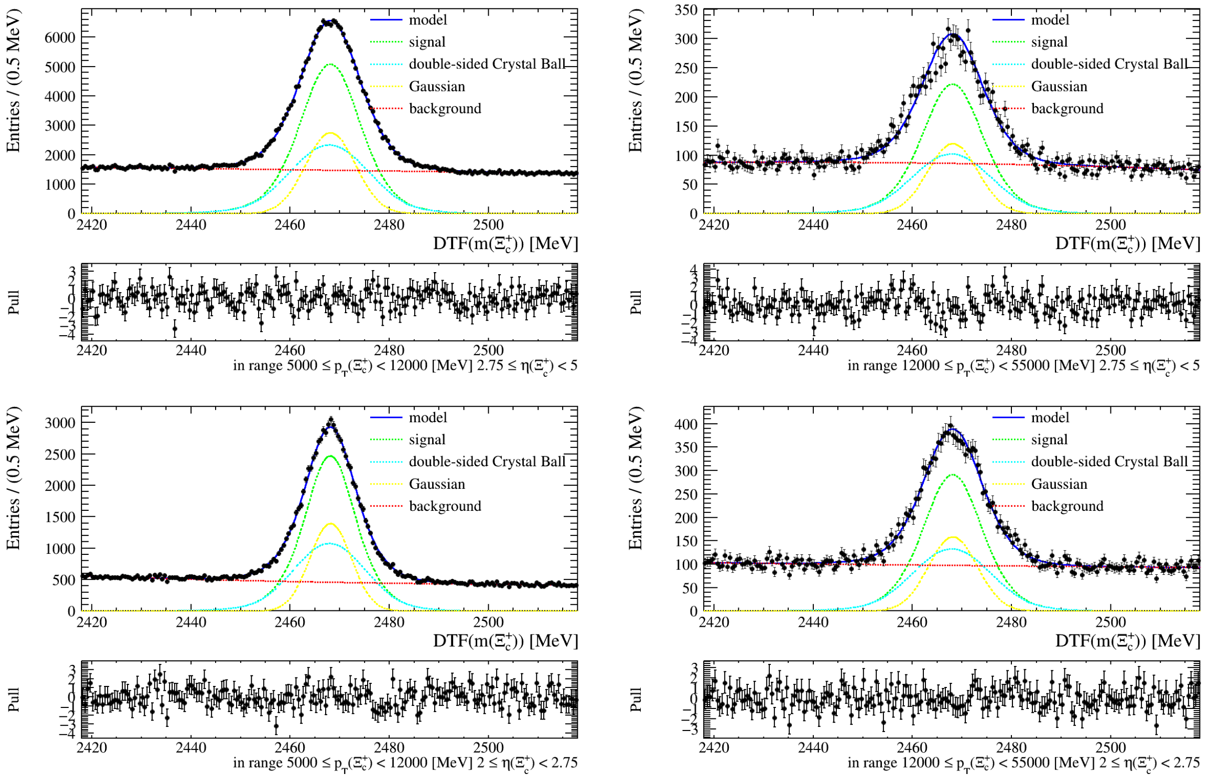


Figure 70: Nominal mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ TIS data dataset. The range is given in the x-axis title of the pull distribution of each plot.

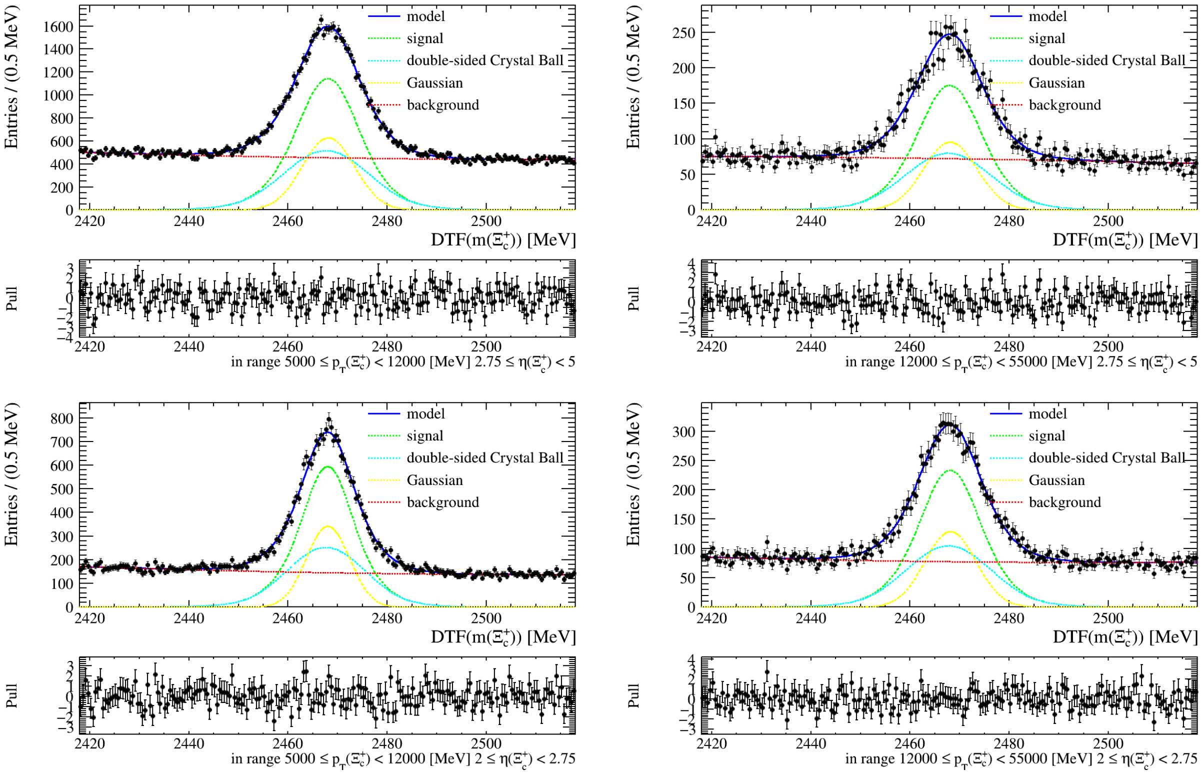


Figure 71: Nominal mass fits in bins of p_{T} and η in the $\Xi_c^+ \to pK^-\pi^+$ TISTOS data dataset. The range is given in the x-axis title of the pull distribution of each plot.

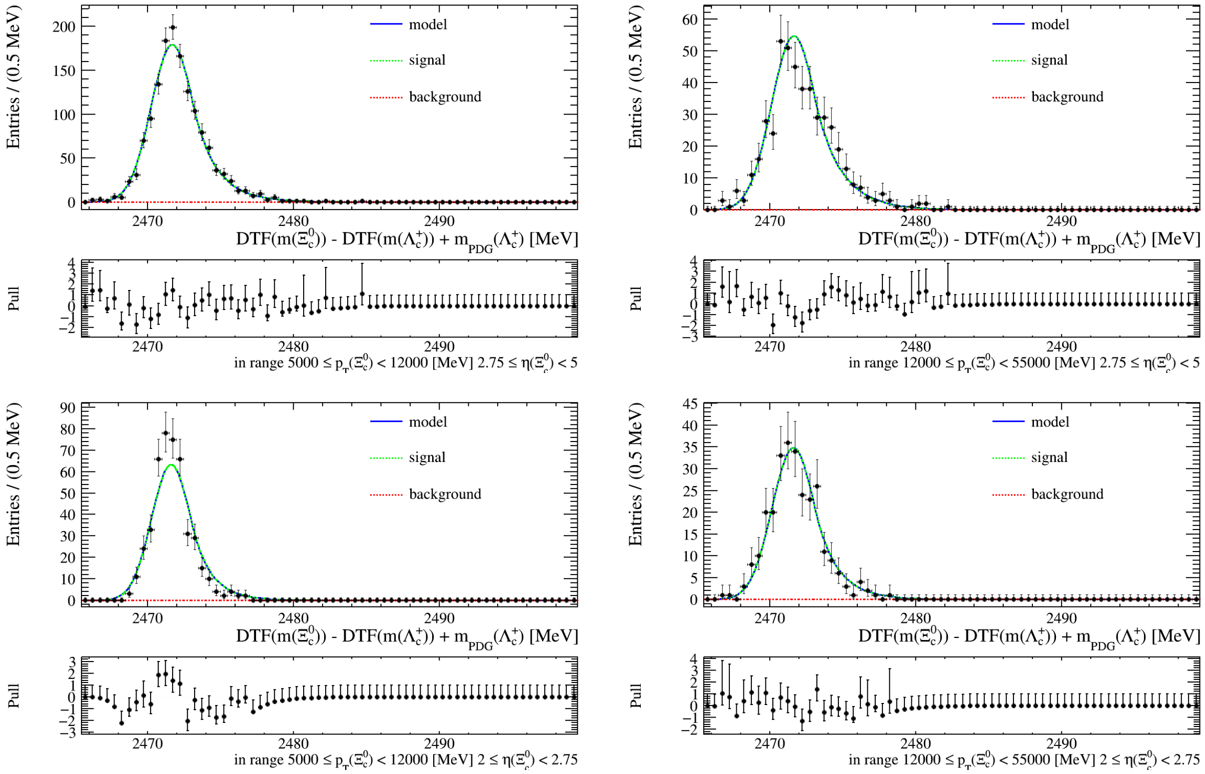


Figure 72: Mass fits in bins of p_{T} and η in the $\mathit{\Xi}_c^0 \to \pi^- \mathit{\Lambda}_c^+$ MC truth-matched dataset. The range is given in the x-axis title of the pull distribution of each plot.

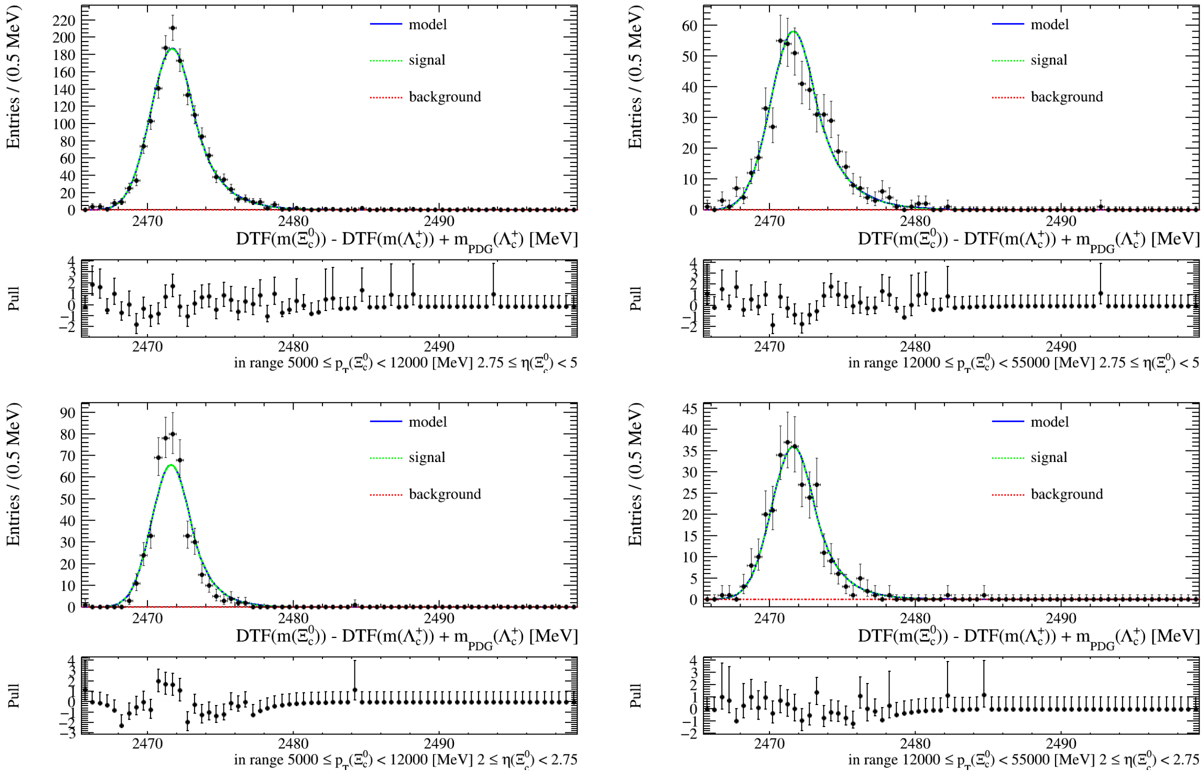


Figure 73: Mass fits in bins of p_T and η in the $\Xi_c^0 \rightarrow \pi^- \Lambda_c^+$ MC truth-matched-plus-ghosts dataset. The range is given in the x-axis title of the pull distribution of each plot.

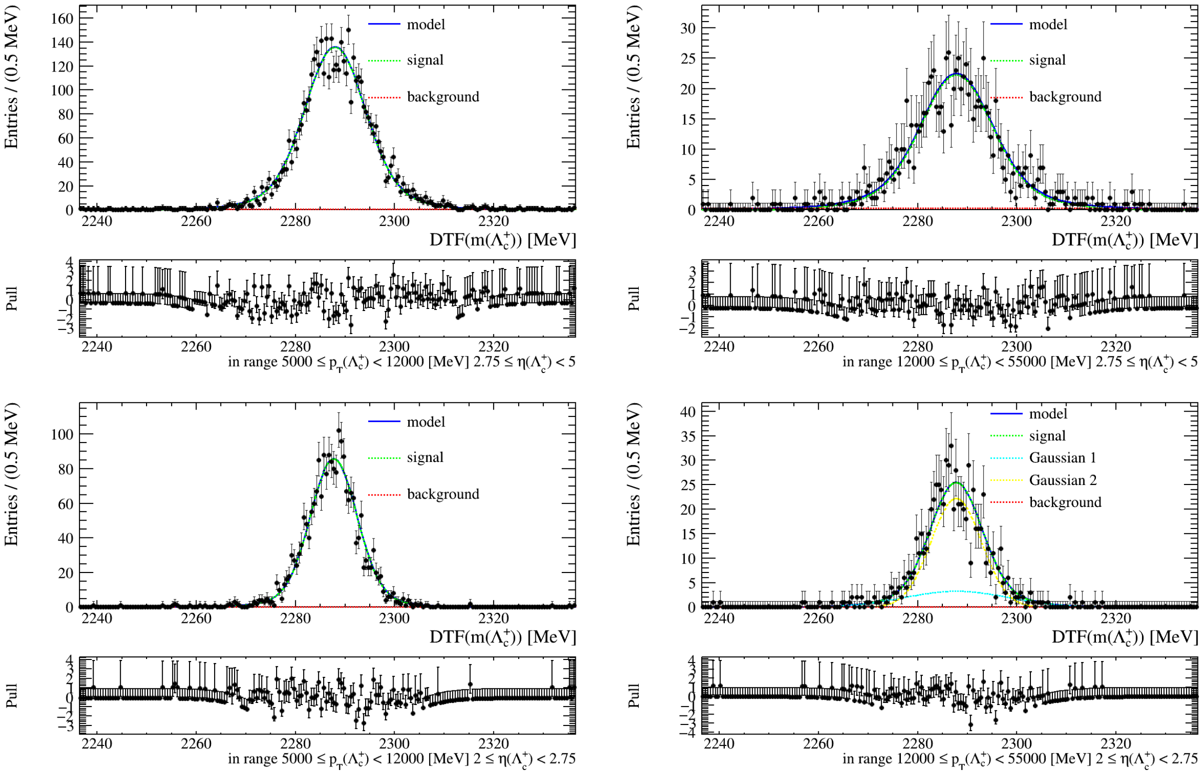


Figure 74: Mass fits in bins of p_T and η in the $\Lambda_c^+ \to pK^-\pi^+$ MC truth-matched dataset. The range is given in the x-axis title of the pull distribution of each plot.

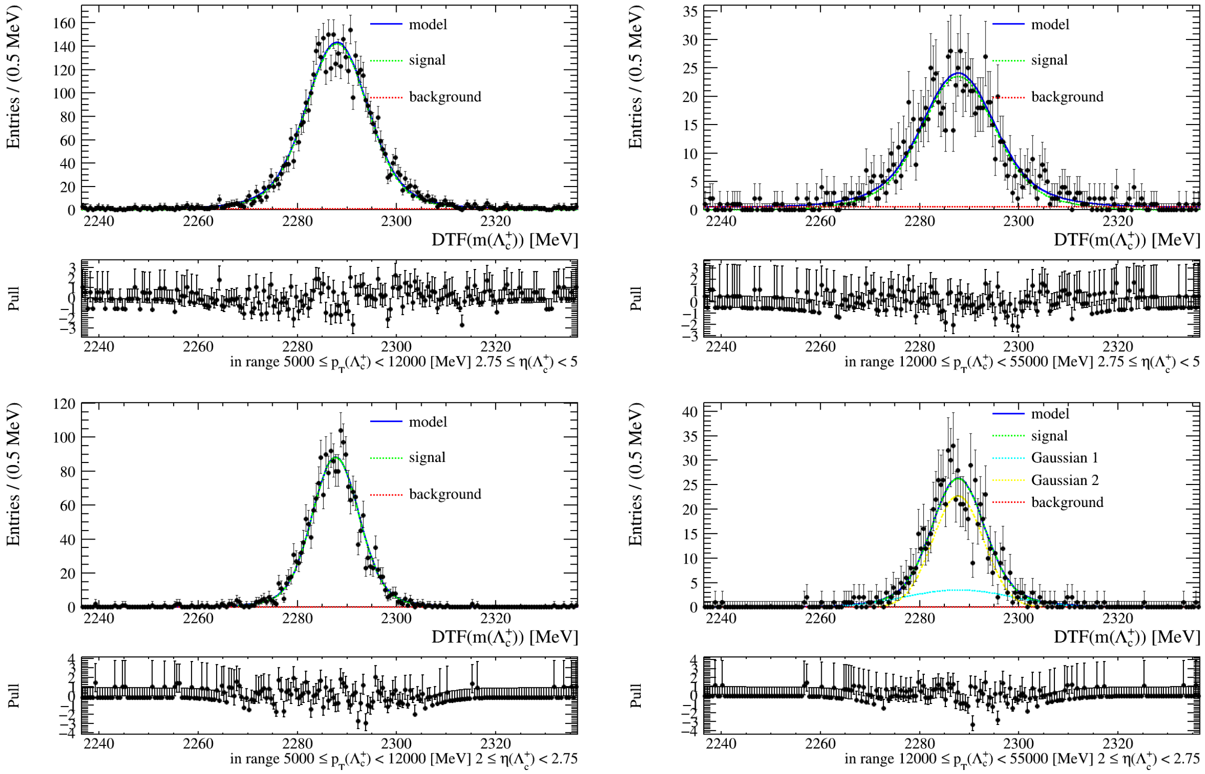


Figure 75: Mass fits in bins of p_T and η in the $\Lambda_c^+ \to pK^-\pi^+$ MC truth-matched-plus-ghosts dataset. The range is given in the x-axis title of the pull distribution of each plot.

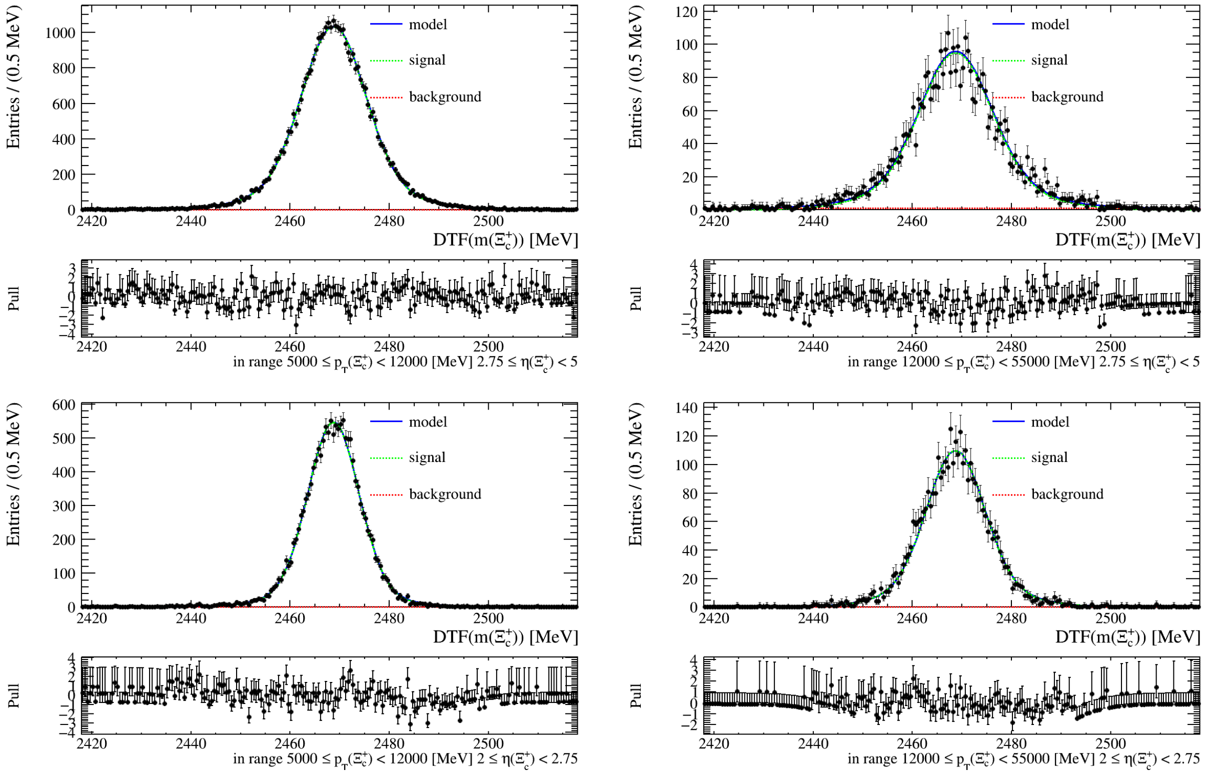


Figure 76: Mass fits in bins of p_T and η in the $\mathit{\Xi}_c^+ \to pK^-\pi^+$ MC truth-matched dataset. The range is given in the x-axis title of the pull distribution of each plot.

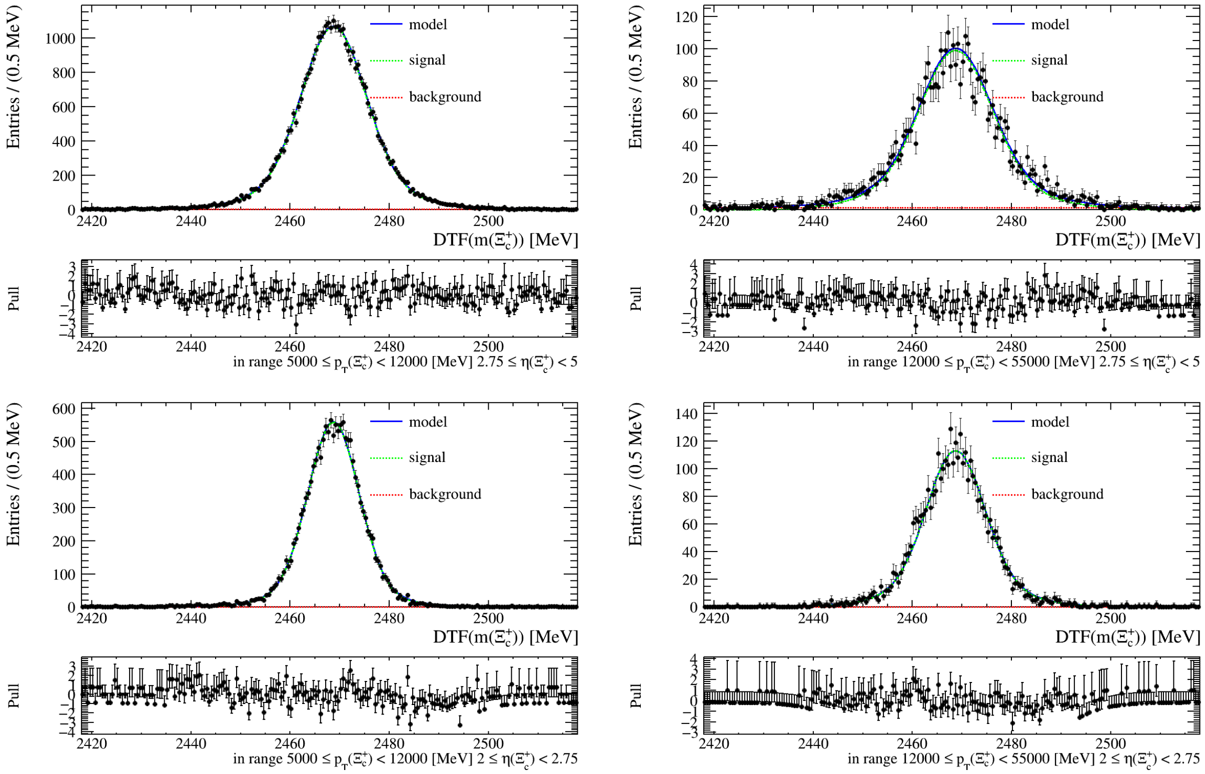


Figure 77: Mass fits in bins of p_{T} and η in the $\varXi_c^+ \to pK^-\pi^+$ MC truth-matched-plus-ghosts dataset. The range is given in the x-axis title of the pull distribution of each plot.

A.7 Fitted parameter values

Tables 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, and 35 show the values of the fitted shape parameters in each fit.

Table 21: Fit parameters for the Ξ_c^0 fit in the TOS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n is the n parameter of the Crystal Ball function, and c_1 is the first-order coefficient of the Chebychev polynomial. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n	c_1 [10^{-2}]
Total		2472 ± 0	1.5 ± 0.1	-14.9	22	-2 ± 1
[2, 2.75)	[5, 12)	2472 ± 0	1.4 ± 0.0	-14.9	22	-1 ± 2
	[12, 55)	2472 ± 0	1.4 ± 0.0	-14.9	22	-6 ± 3
[2.75, 5)	[5, 12)	2472 ± 0	1.5 ± 0.0	-14.9	22	0 ± 1
	[12, 55)	2472 ± 0	1.5 ± 0.0	-14.9	22	-12 ± 2

Table 22: Fit parameters for the Ξ_c^0 fit in the TIS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n is the n parameter of the Crystal Ball function, and c_1 is the first-order coefficient of the Chebychev polynomial. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n	c_1 [10^{-2}]
Total		2472 ± 0	1.3 ± 0.2	-14.9	22	3 ± 1
[2, 2.75)	[5, 12)	2471 ± 0	1.2 ± 0.1	-14.9	22	7 ± 3
	[12, 55)	2472 ± 0	1.2 ± 0.1	-14.9	22	-5 ± 6
[2.75, 5)	[5, 12)	2472 ± 0	1.3 ± 0.1	-14.9	22	3 ± 2
	[12, 55)	2472 ± 0	1.3 ± 0.1	-14.9	22	-7 ± 6

Table 23: Fit parameters for the Ξ_c^0 fit in the TISTOS dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n is the n parameter of the Crystal Ball function, and c_1 is the first-order coefficient of the Chebychev polynomial. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n	c_1 [10^{-2}]
Total		2472 ± 0	1.3 ± 0.2	-14.9	22	-2 ± 2
[2, 2.75)	[5, 12)	2471 ± 0	1.1 ± 0.2	-14.9	22	3 ± 6
	[12, 55)	2471 ± 0	1.2 ± 0.2	-14.9	22	-6 ± 7
[2.75, 5)	[5, 12)	2472 ± 0	1.3 ± 0.2	-14.9	22	-1 ± 3
	[12, 55)	2472 ± 0	1.4 ± 0.2	-14.9	22	-6 ± 6

Table 24: Fit parameters for the Λ_c^+ fit in the TOS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n_L	n_R
Total		2287 ± 0	7.3 ± 0.0	1.85	6 ± 1	155 ± 2
[2, 2.75)	[5, 12)	2287 ± 0	6.2 ± 0.0	1.85	6 ± 0	155 ± 0
	[12, 55)	2287 ± 0	7.3 ± 0.0	1.85	7 ± 1	145 ± 6
[2.75, 5)	[5, 12)	2287 ± 0	7.6 ± 0.0	1.85	6 ± 1	146 ± 3
	[12, 55)	2287 ± 0	8.1 ± 0.1	1.85	7 ± 0	151 ± 5

Table 25: Fit parameters for the Λ_c^+ fit in the TIS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n_L	n_R
Total		2287 ± 0	7.1 ± 0.0	1.85	6 ± 0	155 ± 1
[2, 2.75)	[5, 12)	2287 ± 0	6.6 ± 0.0	1.85	6 ± 0	155 ± 0
	[12, 55)	2287 ± 0	7.1 ± 0.0	1.85	6 ± 0	155 ± 0
[2.75, 5)	[5, 12)	2287 ± 0	7.2 ± 0.0	1.85	6 ± 0	155 ± 0
	[12, 55)	2287 ± 0	7.2 ± 0.0	1.85	6 ± 0	155 ± 0

Table 26: Fit parameters for the Λ_c^+ fit in the TISTOS dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_T [GeV]	μ [MeV]	σ [MeV]	α	n_L	n_R
Total		2287 ± 0	7.5 ± 0.1	1.85	6 ± 0	155 ± 3
[2, 2.75)	[5, 12)	2287 ± 0	7.0 ± 0.0	1.85	6 ± 0	155 ± 0
	[12, 55)	2287 ± 0	7.5 ± 0.0	1.85	6 ± 0	155 ± 0
[2.75, 5)	[5, 12)	2287 ± 0	7.7 ± 0.0	1.85	6 ± 0	155 ± 0
	[12, 55)	2287 ± 0	7.8 ± 0.0	1.85	6 ± 0	155 ± 0

Table 27: Fit parameters for the Ξ_c^+ fit in the TOS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_T [GeV]	μ [MeV]	σ [MeV]	α	n_L	n_R
Total		2468 ± 0	8.1 ± 0.0	1.85	6 ± 0	155 ± 1
[2, 2.75)	[5, 12)	2468 ± 0	7.5 ± 0.0	1.85	6 ± 0	155 ± 0
	[12, 55)	2468 ± 0	8.1 ± 0.0	1.85	6 ± 0	155 ± 0
[2.75, 5)	[5, 12)	2468 ± 0	8.3 ± 0.0	1.85	6 ± 0	155 ± 0
	[12, 55)	2468 ± 0	8.3 ± 0.0	1.85	6 ± 0	155 ± 0

Table 28: Fit parameters for the Ξ_c^+ fit in the TIS data dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_T [GeV]	μ [MeV]	σ [MeV]	α	n_L	n_R
Total		2468 ± 0	7.8 ± 0.1	1.85	6 ± 0	151 ± 5
[2, 2.75)	[5, 12)	2468 ± 0	7.3 ± 0.1	1.85	6 ± 0	151 ± 5
	[12, 55)	2468 ± 0	7.9 ± 0.1	1.85	6 ± 0	151 ± 5
[2.75, 5)	[5, 12)	2468 ± 0	8.0 ± 0.1	1.85	6 ± 0	151 ± 5
	[12, 55)	2468 ± 0	8.0 ± 0.1	1.85	6 ± 0	151 ± 5

Table 29: Fit parameters for the Ξ_c^+ fit in the TISTOS dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, n_L is the Left n parameter of the double-sided Crystal Ball function, and n_R is the Right n parameter of the double-sided Crystal Ball function. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n_L	n_R
Total		2468 ± 0	8.5 ± 0.2	1.85	7 ± 0	151 ± 5
[2, 2.75)	[5, 12)	2468 ± 0	7.9 ± 0.1	1.85	7 ± 0	151 ± 5
	[12, 55)	2468 ± 0	8.6 ± 0.2	1.85	7 ± 0	151 ± 5
[2.75, 5)	[5, 12)	2468 ± 0	8.7 ± 0.1	1.85	7 ± 0	151 ± 5
	[12, 55)	2468 ± 0	8.7 ± 0.2	1.85	7 ± 0	151 ± 5

Table 30: Fit parameters for the Ξ_c^0 fit in the truth-matched simulation dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, and n is the n parameter of the Crystal Ball function. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n
Total		2472 ± 0	1.5 ± 0.0	-0.97 ± 0.05	50
[2, 2.75)	[5, 12)	2472 ± 0	1.3 ± 0.0	-1.07 ± 0.04	50
	[12, 55)	2472 ± 0	1.5 ± 0.0	-1.01 ± 0.04	50
[2.75, 5)	[5, 12)	2472 ± 0	1.5 ± 0.0	-0.95 ± 0.03	50
	[12, 55)	2472 ± 0	1.5 ± 0.0	-0.89 ± 0.04	50

Table 31: Fit parameters for the Ξ_c^0 fit in the truth-matched-plus-ghosts simulation dataset. μ is the mean, σ is the width of the Crystal Ball function, α is the α parameter of the Crystal Ball function, and n is the n parameter of the Crystal Ball function. Parameters without error have been fixed, and ± 0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	σ [MeV]	α	n
Total		2472 ± 0	1.5 ± 0.0	-0.98 ± 0.05	50
[2, 2.75)	[5, 12)	2472 ± 0	1.3 ± 0.0	-1.07 ± 0.04	50
	[12, 55)	2472 ± 0	1.5 ± 0.0	-1.01 ± 0.05	50
[2.75, 5)	[5, 12)	2472 ± 0	1.5 ± 0.0	-0.97 ± 0.04	50
	[12, 55)	2472 ± 0	1.6 ± 0.0	-0.90 ± 0.04	50

Table 32: Fit parameters for the Λ_c^+ fit in the truth-matched simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_T [GeV]	μ [MeV]	f_1 [%]	σ_1 [MeV]	$\frac{\sigma_2}{\sigma_1}$ [%]
Total		2288 ± 0	26 ± 4	10.8 ± 0.7	50 ± 2
[2, 2.75)	[5, 12)	2288 ± 0	16 ± 3	9.4 ± 0.5	49 ± 2
	[12, 55)	2288 ± 0	23 ± 4	10.5 ± 0.5	48 ± 2
[2.75, 5)	[5, 12)	2288 ± 0	25 ± 3	11.1 ± 0.4	52 ± 2
	[12, 55)	2288 ± 0	34 ± 4	12.8 ± 0.5	53 ± 2

Table 33: Fit parameters for the Λ_c^+ fit in the truth-matched-plus-ghosts simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_T [GeV]	μ [MeV]	f_1 [%]	σ_1 [MeV]	$\frac{\sigma_2}{\sigma_1}$ [%]
Total		2288 ± 0	27 ± 4	11.0 ± 0.7	49 ± 2
[2, 2.75)	[5, 12)	2288 ± 0	17 ± 3	9.5 ± 0.5	48 ± 2
	[12, 55)	2288 ± 0	25 ± 4	10.7 ± 0.5	48 ± 2
[2.75, 5)	[5, 12)	2288 ± 0	26 ± 3	11.4 ± 0.4	51 ± 2
	[12, 55)	2288 ± 0	36 ± 4	12.9 ± 0.5	52 ± 2

Table 34: Fit parameters for the Ξ_c^+ fit in the truth-matched simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_T [GeV]	μ [MeV]	f_1 [%]	σ_1 [MeV]	$\frac{\sigma_2}{\sigma_1}$ [%]
Total		2469 ± 0	25 ± 2	11.8 ± 0.3	50 ± 1
[2, 2.75)	[5, 12)	2469 ± 0	15 ± 1	10.5 ± 0.2	48 ± 1
	[12, 55)	2469 ± 0	23 ± 1	11.5 ± 0.2	50 ± 1
[2.75, 5)	[5, 12)	2469 ± 0	25 ± 1	12.2 ± 0.2	51 ± 1
	[12, 55)	2469 ± 0	31 ± 1	13.4 ± 0.2	52 ± 1

Table 35: Fit parameters for the Ξ_c^+ fit in the truth-matched-plus-ghosts simulation dataset. μ is the mean, f_1 is the fraction of the signal yield covered by the first Gaussian, σ_1 is the width of the first Gaussian, and $\frac{\sigma_2}{\sigma_1}$ is the width of the second Gaussian divided by that of the first. Parameters without error have been fixed, and ±0 indicates a negligible uncertainty.

η	p_{T} [GeV]	μ [MeV]	f_1 [%]	σ_1 [MeV]	$\frac{\sigma_2}{\sigma_1}$ [%]
Total		2469 ± 0	25 ± 2	12.1 ± 0.3	49 ± 1
[2, 2.75)	[5, 12)	2469 ± 0	16 ± 1	10.7 ± 0.2	47 ± 1
	[12, 55)	2469 ± 0	24 ± 1	11.8 ± 0.2	49 ± 1
[2.75, 5)	[5, 12)	2469 ± 0	25 ± 1	12.5 ± 0.2	50 ± 1
	[12, 55)	2469 ± 0	31 ± 1	13.7 ± 0.2	51 ± 1

A.8 Λ_c^+ and Ξ_c^+ momentum distributions

The decay modes of interest for both Ξ_c^0 and Λ_c^+ baryons have similar topologies in that they both include $\Lambda_c^+ \to pK^-\pi^+$. Figure 78 shows the momentum distributions of the Λ_c^+ baryon and its final-state daughters after applying Turbo, offline, and trigger selections, as for the input samples to the BDT (see Section 4.3.5). The distributions shown are restricted to the signal p_{T} and η range (see Section 4.4.2). Since the decay mode of interest of the Ξ_c^+ baryon ($\Xi_c^+ \to pK^-\pi^+$) is quite similar to that of the Λ_c^+ baryon, its momentum distribution and that of its daughters is also shown.

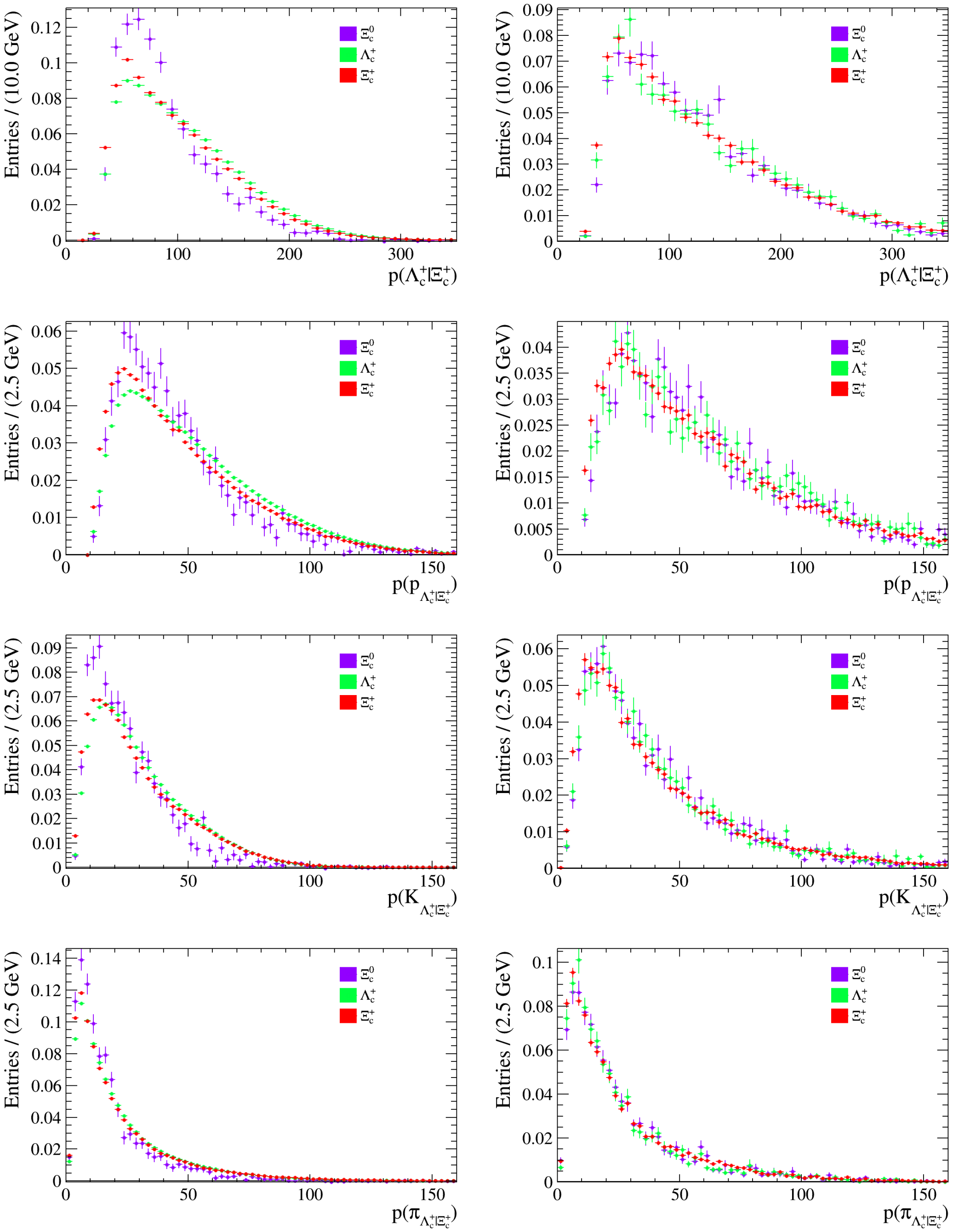


Figure 78: Momentum distributions of the $\mathit{\Lambda}_c^+$ or $\mathit{\Xi}_c^+$ baryon, as appropriate, and its final-state daughters for the (purple) $\mathit{\Xi}_c^0$, (green) $\mathit{\Lambda}_c^+$, and (red) $\mathit{\Xi}_c^+$ (left) data and (right) simulation datasets after the application of Turbo, offline, and trigger selections. The (left) data distributions are the result of a side-band subtraction using the fits shown in Figure 54. The (right) simulation distributions are truth matched.

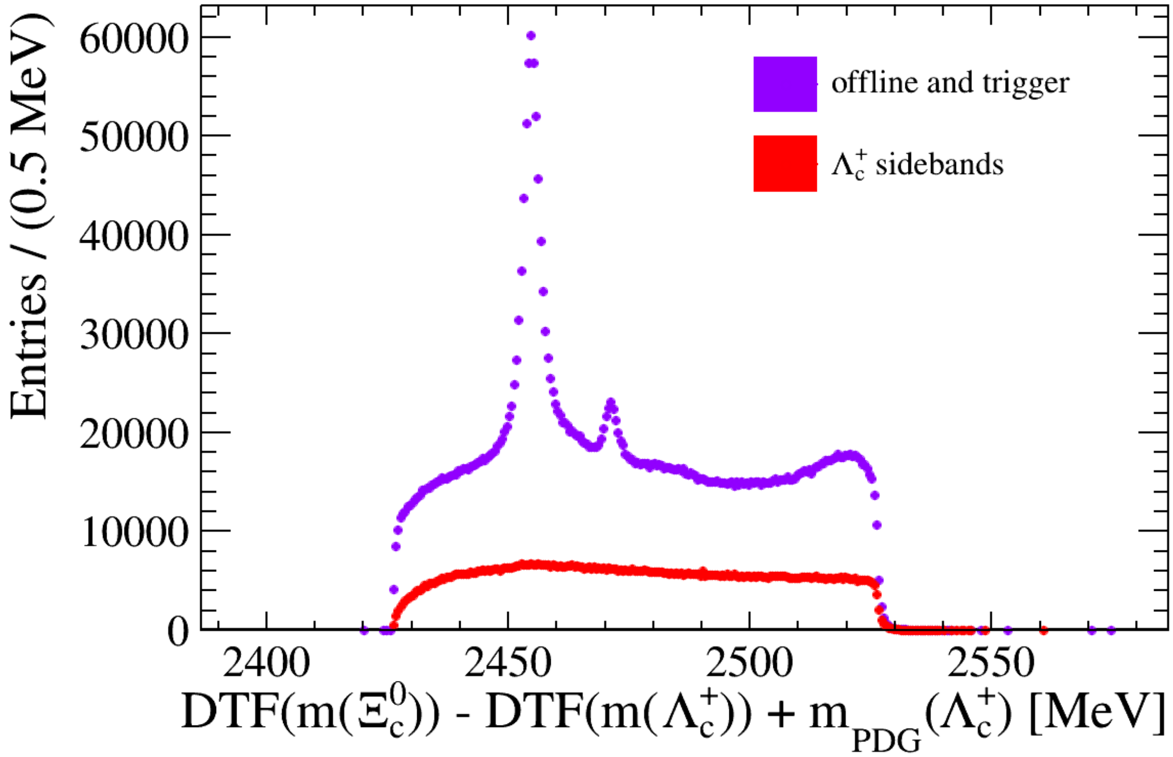


Figure 79: $\mathrm{DTF}\left(m\left(\Xi_c^0\right)\right) - \mathrm{DTF}\left(m\left(\Lambda_c^+\right)\right) + m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ in the Ξ_c^0 data dataset after the application of the Turbo, offline, and trigger selections (purple) without any modifications and (red) with the Λ_c^+ mass required to be outside ±40 MeV of its value in the PDG [1].

A.9 Ξ_c^0 in Λ_c^+ sidebands

The decay $\Xi_c^0 \to pK^-\pi^+\pi^-$ may proceed in the Standard Model via strong gluon interactions and the weak decay of the charm quark. Since these are the same final-state daughters as in the decay of interest, $\Xi_c^0 \to \pi^-\left(\Lambda_c^+ \to pK^-\pi^+\right)$, we check whether such decays contribute to the signal peak in $\mathrm{DTF}\left(m\left(\Xi_c^0\right)\right) - \mathrm{DTF}\left(m\left(\Lambda_c^+\right)\right) + m_{\mathrm{PDG}}\left(\Lambda_c^+\right)$ (see Figure 24). Figure 79 demonstrates that there is no peaking contribution from this decay in the signal distribution after the application of the Turbo, offline, and trigger selections (see Sections 4.3.1, 4.3.3, and 4.3.2).

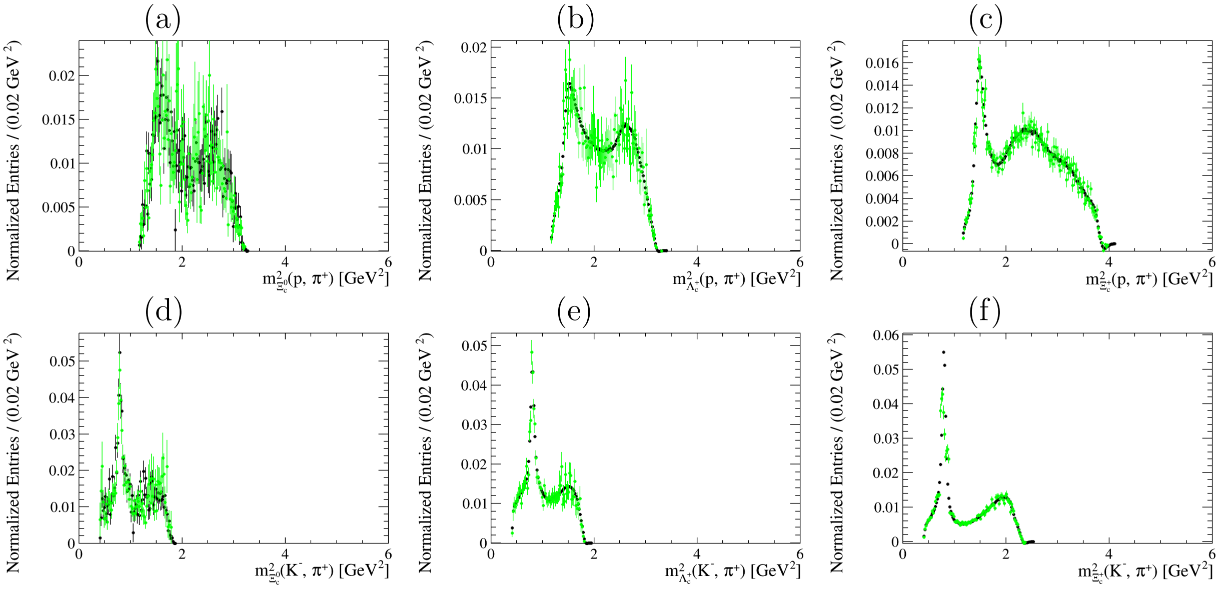


Figure 80: The squared invariant-mass of (top) the proton and π^+ meson and (bottom) the K^- meson and π^+ meson for (a & d) $\mathit{\Xi}_c^0$, (b & e) $\mathit{\Lambda}_c^+$, and (c & f) $\mathit{\Xi}_c^+$ (black) data and (green) simulation datasets after the application of all selections and weights. The (black) data distributions are the result of the application of the *sPlot* method [16] using the fits shown in Figures 24, 25, and 26. The (green) simulation distributions are truth matched.

A.10 Dalitz weighting

We weight the Dalitz spectra of the $\mathit{\Lambda}_c^+$ and $\mathit{\Xi}_c^+$ baryons in simulation to match data (see Section 4.3.4). Figure 80 shows the results of this weighting, compared to data.

A.11 TISTOS trigger efficiencies

We cross-check the efficiency given to us in data by the TISTOS method by applying the TISTOS method to simulation and comparing the two results for each efficiency ratio. We calculate the trigger efficiency for each decay in simulation using the TISTOS method as described in Section 4.4, but, just as with other simulation-determined efficiencies, we simply count candidates (with appropriate weighting and truth-matching criteria applied [see Section 4.3.4]) to determine the signal yields rather than using fits. We then calculate the relevant efficiency ratios in simulation and data and compare them. Figure 81 shows the comparison. Ratios consistent with 1 indicate agreement between data and simulation.

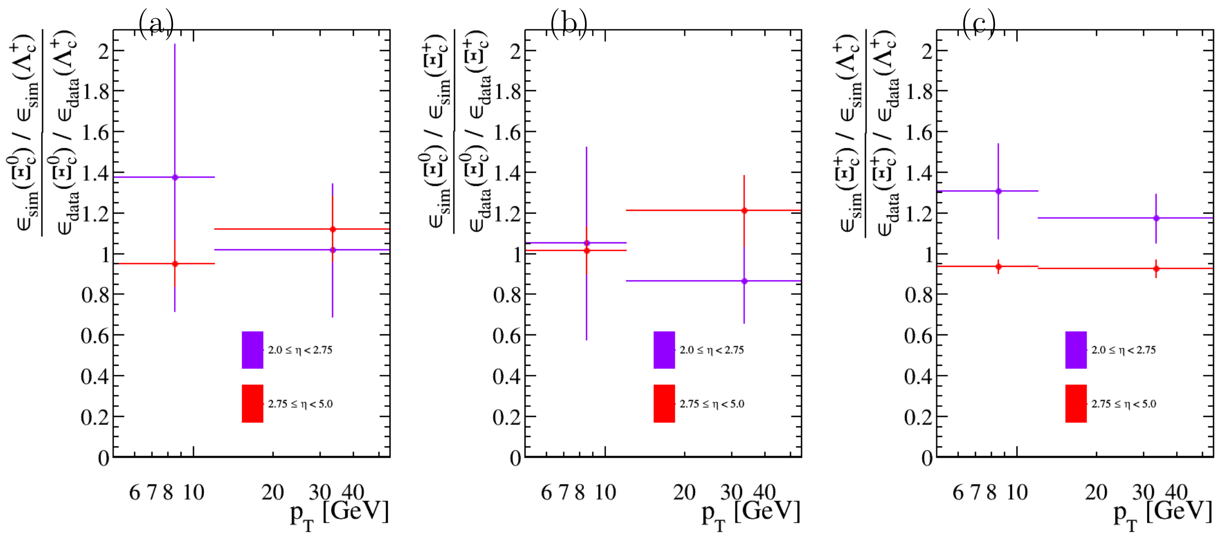


Figure 81: Ratios of trigger-efficiency ratios, derived using the TISTOS method, in simulation to those in data for (a) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ to $\Lambda_c^+ \to pK^-\pi^+$, (b) $\Xi_c^0 \to \pi^- (\Lambda_c^+ \to pK^-\pi^+)$ to $\Xi_c^+ \to pK^-\pi^+$, and (c) $\Xi_c^+ \to pK^-\pi^+$ to $\Lambda_c^+ \to pK^-\pi^+$. Only statistical errors shown.

Correlations between TIS and TISTOS datasets are taken into account, as described in Section 4.4.1.

Most ratios are consistent with 1 within errors, though the trigger efficiency given by simulation in the Ξ_c^+ case differs somewhat from data, showing a greater efficiency at low η and a lesser efficiency at high η. Given the size of the error and that only the statistical error is shown in the figure, Figure 81 demonstrates that there is good agreement between the TISTOS method in simulation and that in data.

A.12 Updated Gronau and Rosner predictions

As stated in Section 4.1, we calculate updated predictions for Gronau and Rosner using the recent LHCb measurement of charm-baryon lifetimes [2]. To do this, we rewrite Equation 14 in Reference [5] as

$$A_{cs\to cd}(\Xi_c^0 \to \pi^- \Lambda_c^+) = C \times \left(0.44[\Gamma(\Xi_c^0) - \Gamma(\Lambda_c^+)] + 0.05[\Gamma(\Lambda_c^+) - \Gamma(\Xi_c^+)]\right). \quad (114)$$

We can then solve for C using Gronau and Rosner's input charm lifetime values from Reference [8], enumerated in Table 36, and find $C = -0.9917$. We can then use Equation 114 to re-calculate $A_{cs\to cd}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ using the updated lifetimes listed in Table 8. Next, we rewrite Equation 2 in Reference [5] as

$$\Gamma(B_1 \to B_2\pi) = D \times |A|^2, \tag{115}$$

where we have set B to 0, as described elsewhere in Reference [5]. We then notice

$$\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+) = \frac{D \times |A|^2}{\Gamma(\Xi_c^0)} \tag{116}$$

and calculate $D = (0.13 \pm 0.07)\%$ using Gronau and Rosner's constructive prediction for $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ and for A (Equation 18 in Reference [5]).[51] Finally, we can use Equation 116 and the updated Ξ_c^0 lifetime listed in Table 8 to calculate the updated predictions for $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ listed in Table 19.

Table 36: Old world-averages of charm-baryon lifetime measurements [8].

Particle	Lifetime [fs]
Ξ_c^0	112 ± 13
Λ_c^+	200 ± 6
Ξ_c^+	442 ± 26

We have not updated the predictions of Faller and Mannel or Cheng *et al.*; the effect would be to lower Faller and Mannel's positive interference prediction and raise the Cheng *et al.* negative one, giving somewhat better agreement with Gronau and Rosner's predictions.

---

[51] Gronau and Rosner only provide an upper limit for their destructive prediction of $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$, hence the decision to use only the *constructive* values for A and $\mathcal{B}(\Xi_c^0 \to \pi^- \Lambda_c^+)$ to calculate D.

References

[1] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[2] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the Λ_c^+, Ξ_c^+, and Ξ_c^0 baryon lifetimes*, Phys. Rev. **D100** (2019) 032001, `arXiv:1906.08350`.

[3] LHCb collaboration, R. Aaij *et al.*, *Measurement of the properties of the Ξ_b^{*0} baryon*, JHEP **05** (2016) 161, `arXiv:1604.03896`.

[4] M. B. Voloshin, *Update on splitting of lifetimes of c and b hyperons within the heavy quark expansion and decays $\Xi_Q \to \Lambda_Q\pi$*, Phys. Rev. D **100** (2019) 114030, `arXiv:1911.05730`.

[5] M. Gronau and J. L. Rosner, *From $\Xi_b \to \Lambda_b^0\pi$ to $\Xi_c \to \Lambda_c^+\pi$*, Phys. Lett. **B757** (2016) 330, `arXiv:1603.07309`.

[6] S. Faller and T. Mannel, *Light-quark decays in heavy hadrons*, Phys. Lett. **B750** (2015) 653, `arXiv:1503.06088`.

[7] H.-Y. Cheng *et al.*, *Heavy-flavor-conserving hadronic weak wecays of heavy baryons*, JHEP **03** (2016) 028, `arXiv:1512.01276`.

[8] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

[9] C. MissMJ, *Standard model of elementary particles*, `https://commons.wikimedia.org/w/index.php?curid=4286964`.

[10] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001, and 2019 update.

[11] Laurascudder, *Baryon octet*, `https://commons.wikimedia.org/wiki/File:Baryon_octet.png`. Slight modifications have been made to the original; this modified image (and

it alone) is released for use by others according to the Creative Commons Attribution-Share Alike 3.0 Unported license.

[12] CERN, *The cern accelerator complex - 2019*, `https://cds.cern.ch/images/CERN-GRAPHICS-2019-002-1`.

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

[14] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[15] LHCb, *The lhcb data flow during run ii*, `https://lhcb.github.io/starterkit-lessons/first-analysis-steps/dataflow.html`, accessed 4 October 2021.

[16] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[17] S. Stone and M. Wilkinson, *Observation of the suppressed decay* $\Xi_c^0 \to \pi^- \Lambda_c^+$, LHCb-ANA-2020-016, 2020.

[18] Aristotle and C. D. C. Reeve, *Metaphysics*, Hackett Publishing Company, 2016.

[19] J. Barnes, *The Presocratic Philosophers*, Routledge, 1999.

[20] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579.

[21] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[22] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

[23] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear Physics B **44** (1972) 189.

[24] R. L. Jaffe, *Quark Confinement*, Nature **268** (1977) 201.

[25] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.

[26] D. Griffiths, *Introduction to elementary particles*, 2008.

[27] W. Heisenberg, *On the structure of atomic nuclei*, Z. Phys. **77** (1932) 1.

[28] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry*, doi: 10.2172/4008239.

[29] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918** (1918) 235, `arXiv:physics/0503066`.

[30] Y. Nambu, *Axial vector current conservation in weak interactions*, Phys. Rev. Lett. **4** (1960) 380.

[31] M. Gell-Mann, *Symmetries of baryons and mesons*, Phys. Rev. **125** (1962) 1067.

[32] M. Gell-Mann, *The Symmetry group of vector and axial vector currents*, Physics Physique Fizika **1** (1964) 63.

[33] S. Treiman, *Commutator Algebra and the PCAC Hypothesis*, Comments Nucl. Part. Phys. **1** (1967) 13, text can be found on INSPIRE.

[34] ALPHA collaboration, J. Heitger and R. Sommer, *Nonperturbative heavy quark effective theory*, JHEP **02** (2004) 022, `arXiv:hep-lat/0310035`.

[35] M. Neubert, *Heavy quark effective theory*, Subnucl. Ser. **34** (1997) 98, `arXiv:hep-ph/9610266`.

[36] R. H. Dalitz, *On the analysis of τ-meson data and the nature of the τ-meson*, Phil. Mag. Ser. 7 **44** (1953) 1068.

[37] A. Bettini, *Introduction to Elementary Particle Physics, Second Edition*, Cambridge University Press, New York, 2014.

[38] CERN, *Cern's accelerator complex*, `https://home.cern/science/accelerators/accelerator-complex`, accessed 1 October 2021.

[39] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[40] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.

[41] LHCb, *Moore*, `https://lhcbdoc.web.cern.ch/lhcbdoc/moore/master/index.html`, accessed 4 October 2021.

[42] G. Corti *et al.*, *Software for the LHCb experiment*, IEEE Trans. Nucl. Sci. **53** (2006) 1323.

[43] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[44] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`.

[45] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

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

[47] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[48] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[49] LHCb collaboration, R. Aaij *et al.*, *First branching fraction measurement of the suppressed decay $\Xi_c^0 \to \pi^- \Lambda_c^+$*, Phys. Rev. **D102** (2020) 071101(R), `arXiv:2007.12096`.

[50] M. B. Voloshin, *Weak decays $\Xi(Q) \to \Lambda(Q)\pi$*, Phys. Lett. **B476** (2000) 297, `arXiv:hep-ph/0001057`.

[51] LHCb collaboration, R. Aaij *et al.*, *Evidence for the strangeness-changing weak decay $\Xi_b^- \to \Lambda_b^0 \pi^-$*, Phys. Rev. Lett. **115** (2015) 241801, `arXiv:1510.03829`.

[52] Belle collaboration, S. H. Lee *et al.*, *Measurements of the masses and widths of the $\Sigma_c(2455)^{0/++}$ and $\Sigma_c(2520)^{0/++}$ baryons*, Phys. Rev. **D89** (2014) 091102, `arXiv:1404.5389`.

[53] Belle collaboration, Y. B. Li *et al.*, *First Measurements of Absolute Branching Fractions of the Ξ_c^0 Baryon at Belle*, Phys. Rev. Lett. **122** (2019) 082001, `arXiv:1811.09738`.

[54] LHCb collaboration, R. Aaij *et al.*, *Measurement of the mass and production rate of Ξ_b^- baryons*, Phys. Rev. **D99** (2019) 052006, `arXiv:1901.07075`.

[55] Belle collaboration, Y. B. Li *et al.*, *First measurements of absolute branching fractions of the Ξ_c^+ baryon at Belle*, Phys. Rev. **D100** (2019) 031101, `arXiv:1904.12093`.

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

[57] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, `arXiv:1210.6861`.

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

[59] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.

[60] A. Hoecker *et al.*, *TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide.*, `arXiv:physics/0703039`.

[61] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, 2014.

[62] E. O. LEBIGOT, *Uncertainties: a python package for calculations with uncertainties*, `http://pythonhosted.org/uncertainties/`.

[63] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, `arXiv:physics/0306116`, [,186(2003)].

[64] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[65] R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86. See also `http://root.cern.ch/`.

[66] LHCb collaboration, R. Aaij *et al.*, *Observation of two new Ξ_b^- baryon resonances*, Phys. Rev. Lett. **114** (2015) 062004, `arXiv:1411.4849`.

[67] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13 TeV pp collisions*, Phys. Rev. **D100** (2019) 031102(R), `arXiv:1902.06794`.

[68] 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, Erratum ibid. **09** (2016) 013, Erratum ibid. **05** (2017) 074, `arXiv:1510.01707`.

[69] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[70] LHCb collaboration, R. Aaij *et al.*, *Measurement of the* $B^{\pm}$ *production cross-section in* pp *collisions at* $\sqrt{s}=7$ *and 13 TeV*, JHEP **12** (2017) 026, `arXiv:1710.04921`.

[71] LHCb collaboration, R. Aaij *et al.*, *Measurement of the* $B^{\pm}$ *production cross-section in* pp *collisions at* $\sqrt{s}=$ *7 and 13 TeV*, JHEP **12** (2017) 026, `arXiv:1710.04921`.

[72] M. Artuso, E. Barberio, and S. Stone, *B Meson Decays*, PMC Phys. A **3** (2009) 3, `arXiv:0902.3743`.

[73] R. Nisius, *On the combination of correlated estimates of a physics observable*, The European Physical Journal C **74** (2014) 3004.

Vita

NAME OF AUTHOR: Michael Kent Wilkinson

PLACE OF BIRTH: Oklahoma City, Oklahoma, USA

DATE OF BIRTH: May 30, 1990

DEGREES AWARDED: Bachelor of Science in Physics (Summa cum Laude) and Bachelor of Arts (Summa cum Laude) from the University of Oklahoma

SELECTED PUBLICATIONS:

> The LHCb collaboration, *First branching fraction measurement of the suppressed decay* $\Xi_c^0 \to \pi^- \Lambda_c^+$, Phys. Rev. D102 (2020) 071101(R), arXiv:2007.12096.
>
> The LHCb collaboration, *Isospin amplitudes in* $\Lambda_b^0 \to J/\psi \Lambda(\Sigma^0)$ *and* $\Xi_b^0 \to J/\psi \Xi^0(\Lambda)$ *decays*, Phys. Rev. Lett. 124 (2020) 111802, arXiv:1912.02110.
>
> M. Artuso *et al.*, *First beam test of UT sensors with the SALT 3.0 readout ASIC*, CERN-LHCb-PUB-2019-009, CERN, Geneva, 2019.
>
> The LHCb collaboration, *Measurement of b-hadron fractions in 13 TeV pp collisions*, Phys. Rev. D100 (2019) 031102(R), arXiv:1902.06794.
>
> The LHCb collaboration, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. 118 (2017) 052002, Erratum ibid. 119 (2017) 169901, arXiv:1612.05140.

FULL PUBLICATION LIST: `https://orcid.org/0000-0001-6561-2145`