

Syracuse University

SURFACE at Syracuse University

Dissertations - ALL

SURFACE at Syracuse University

12-14-2022

Observation of the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and measurement of $B(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$

Zhuoming Li

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



Part of the [Physics Commons](#)

Recommended Citation

Li, Zhuoming, "Observation of the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and measurement of $B(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$ " (2022). *Dissertations - ALL*. 1589.
<https://surface.syr.edu/etd/1589>

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

We present the first observation of the weak decay $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, which is mediated by an $s \rightarrow u\bar{u}d$ transition within the Ξ_b^- baryon. The analysis uses a pp collision data sample at a center-of-mass energy of 13 TeV corresponding to 5.5 fb^{-1} of integrated luminosity. The sample of Λ_b^0 baryons used in this analysis are reconstructed through their decays to $\Lambda_c^+ \pi^-$ and $\Lambda_c^+ \pi^- \pi^+ \pi^-$. From a fit to the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ mass spectrum, the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay is observed with a significance of 11.3 standard deviations, thus establishing observation of this decay. The relative branching fraction is measured to be

$$\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = (6.74 \pm 0.74 \pm 0.50) \times 10^{-4},$$

where $f_{\Xi_b^-}$ and $f_{\Lambda_b^0}$ are the fragmentation fractions, and $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$ is the branching fraction, and the uncertainties are statistical and experimental systematic. Using the measured value of $\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} = (8.2 \pm 0.7 \pm 0.6 \pm 2.5) \times 10^{-2}$, the branching fraction is measured as

$$\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = (0.82 \pm 0.09 \pm 0.07 \pm 0.25) \times 10^{-2}.$$

For predictions of Ξ_b^- decay width that only consider b-quark decay, this contribution must be added, thus increasing predictions of the Ξ_b^- decay width by about 1%.

Observation of the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and measurement of $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$

by

Zhuoming Li

B.S., Wuhan University, 2014

Dissertation

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Physics

Syracuse University

December 2022

Copyright © Zhuoming Li, 2022

All Rights Reserved

Acknowledgements

I would like to express my deepest appreciation to my advisor, Steven Blusk. His guidance and advice are invaluable and carried me through all the stages of my Ph.D. program. I would also like to give my gratitude to all professors, post-docs, and fellow graduate students in the LHCb group of Syracuse University, for they provided firm support both to my physics analysis work and my UT hardware tasks.

I thank my CERN colleagues involved in the UT development, and also the staff of the Syracuse University physics department, who lent me great help every time I reached out for assistance.

Last but not least, I deeply appreciate the love and understanding from my family and my wife. Despite that we are separated by the Pacific Ocean, they continuously support and encourage me just as they are on my side.

Contents

1	Standard Model	1
2	The LHCb Experiment	6
2.1	Introduction	6
2.2	Dipole Magnet	6
2.3	Tracking system	7
2.3.1	Vertex locator	7
2.3.2	Planar tracking stations	8
2.4	Particle identification	10
2.4.1	RICH	10
2.4.2	Calorimeters	12
2.4.3	Muon system	13
2.5	Trigger	14
2.5.1	Level-0	14
2.5.2	High level trigger	15
3	Analysis of the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay	17
3.1	Overview of the analysis	17
3.2	Data selection	18
3.3	Stripping selections on $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \pi^- \pi^+ \pi^-$	19
3.4	Offline selections common to $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$	19
3.4.1	Veto on clone tracks	19
3.4.2	Veto on other charm final states	20
3.4.3	Additional offline selections for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$	22
3.4.4	Multivariate discriminant for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$	23
3.5	Ξ_b^- selection	24
3.5.1	Multivariate discriminant for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$	25
3.6	Requirements on BDT1 and BDT2	27
4	Inputs from simulation	33
4.1	Mass resolution for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$	33
4.2	Ξ_b^- Signal Mass Resolution Shape	33

4.3	Weighting of the simulation	35
4.3.1	Λ_c^+ Dalitz weights	35
4.3.2	Weighting of the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation	35
4.3.3	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ weighting	35
5	Relative efficiencies	40
5.1	Relative acceptance	40
5.2	Relative Reconstruction efficiency	41
5.3	Relative offline selection efficiency	42
5.4	Relative trigger efficiency	43
5.5	Correction of relative efficiency	43
5.5.1	BDT2 requirement	43
5.5.2	Pion Tracking Efficiency	45
5.5.3	True signal matching in simulation	45
5.6	Summary of relative efficiencies	46
6	Fits to data	47
6.1	Fit to the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ samples	47
6.2	Fit to the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ mass spectra	50
6.2.1	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ Signal shape	50
6.2.2	Σ_b^- and Σ_b^{*-} resonance model	51
6.2.3	Combinatorial background	52
6.2.4	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ mass fit	52
7	Systematic Uncertainties	58
7.1	Yield determination	58
7.1.1	Yield of Λ_b^0	58
7.1.2	Yield of Ξ_b^-	58
7.2	Efficiency determination	59
7.2.1	Efficiency of Λ_b^0 mass window	59
7.2.2	Efficiency uncertainty from $p_T - \eta$ weights	59
7.3	$\Xi_b^- p_T$ spectra in simulation	60
7.4	Multiple Candidates	61

7.5	Pion tracking efficiency	61
7.6	BDT2 efficiency	62
7.7	Systematic uncertainty summary	63
8	Results and summary	64
A	StrippingLine Selections	66
A.1	$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ stripping selections	66
A.2	$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ stripping selections	67
B	Selections applied in DaVinci script	67
C	Veto on other charm final states for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$	68
D	Clone track removal	68
E	BDT1 input variables for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$	70
F	BDT1 input variables for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$	70
G	BDT2 input variables and result for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decay	71
H	Weighting of the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation	71
I	Offline selection efficiencies of each year	72
J	Correction of total relative efficiency	84
K	Full results of Λ_b^0 data fit	87
L	Full results of $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ data fit	93
M	Trigger efficiency	97
N	Multiple candidates plots	98
	References	99

List of Figures

1	Standard Model of fundamental particles, including the quarks, leptons, and interaction force carriers.	2
2	Magnitude of CKM elements, provided by CKMFilter Group [1]	3
3	Feynman diagram of (Left) $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, (Middle) $\Xi_b^- \rightarrow \Lambda_b^0 \mu^- \bar{\nu}_\mu$, and (Right) $\Xi_b^- \rightarrow \Lambda_b^0 e^- \bar{\nu}_e$	4
4	Illustration of a possible layout of the tetraquarks (Left) and pentaquarks (Right) [2]	5
5	Layout of LHC. Two proton beams in the opposite directions (red and blue lines) can collide in one of four detectors: CMS, ALICE, ATLAS or LHCb. [3]	7
6	Simulated angular distribution of $b\bar{b}$ pairs. [4]	7
7	LHCb detector layout, side view [5].	8
8	The perspective view of the LHCb dipole magnet (Left) and the magnetic field component B_y along the z-axis, i.e. the beamline (Right). [6]	8
9	Schematic of VELO, TT and T1-T3's positions relative to the magnetic field B_y and the types of track they monitor. [6]	9
10	The schematic of VELO showing the layout of r and ϕ silicon sensors, and also the arrangement of modules along the z-axis. The bottom shows that the VELO's position when it is fully opened or closed. [6]	9
11	(Left) The geometry of four layers of TT and the silicon sensors arrangement in each layer. (Right) Layout of IT and OT of T1-T3 stations. [7]	10
12	(Left) Side view schematic of RICH1. (Right) Top view schematic of RICH2. [6]	11
13	θ_c as a function of momentum in different medium. [6]	12
14	Layout of the calorimeters. [8]	12
15	(Left) ECAL cell. (Right) HCAL cell. [8]	13
16	Side view of Muon system. [6]	15
17	Overview of LHCb trigger system. [9]	15
18	Cartoon showing the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay (left) and the main background (right), a Λ_b^0 plus a random π^-	18
19	α distribution of all 6 track pairs combined of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$, $\Lambda_c^+ \rightarrow p^+ K^- \pi^+$	20
20	Distributions of the invariant mass ($\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ data): (top left) $m(K^+ K^-)$, (top middle) $m(K^+ K^- \pi^+)$, (top right) $m(K^- \pi^+ \pi^+) - m(K^- \pi^+)$ and (bottom left) $m(K^- \pi^+ \pi^+)$. The black histogram corresponds to data before the veto, and red is after the veto.	21
21	Distribution of the invariant mass (Left) $m(\Lambda_c^+ \pi^-)$ and (Right) $m(\Lambda_c^+ \pi^- \pi^- \pi^+)$ with stripping and offline selections applied as described in the text for the 2016 - 2018 data set.	22

22	Distributions of the output BDT1 response for signal and background for the (Left) $A_b^0 \rightarrow A_c^+ \pi^-$ and (right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$ samples. The test and training samples are both shown.	24
23	Invariant mass spectrum of (left) all $A_b^0 \rightarrow A_c^+ \pi^-$ candidates that pass the rectangular selections, and (middle) candidates with $\text{BTD1} < -0.9$ and (right) $\text{BDT} > -0.9$	24
24	Invariant mass spectrum of (left) all $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$ candidates that pass the rectangular selections, and (middle) candidates with $\text{BTD1} < -0.9$ and (right) $\text{BDT} > -0.9$	25
25	Distributions of the input variables used to train BDT2 for $\Xi_b^- \rightarrow A_b^0 \pi^-$, with $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$; signal and background distributions are superimposed.	27
26	Distributions of the output BDT2 response for $\Xi_b^- \rightarrow A_b^0 \pi^-$, with $A_b^0 \rightarrow A_c^+ \pi^-$ (Left) and $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$ (Right) signals and WS background. The test and training samples are both shown. . .	28
27	(Left) RS (black) and WS (red) BDT response distributions in $10 < \Delta m < 32.89 \text{ MeV}/c^2$ mass window. (Right) The ratio $\frac{N_{RS}}{N_{WS}}$ of each bin of the left plot. This is Run2 $\Xi_b^- \rightarrow A_b^0 \pi^-$, $A_b^0 \rightarrow A_c^+ \pi^-$ data sample. . .	28
28	BDT1 and BDT2 2D scan FOM value for $\Xi_b^- \rightarrow A_b^0 \pi^-$ with (Left) $A_b^0 \rightarrow A_c^+ \pi^-$, and (Right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$	30
29	FOM vs. BDT2 when $\text{BDT1} > 0$ for both data (Left) $A_b^0 \rightarrow A_c^+ \pi^-$ and (Right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.	30
30	Efficiency vs. BDT2 with $\text{BDT1} > 0$ for both simulated (Left) $A_b^0 \rightarrow A_c^+ \pi^-$ and (Right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$. Note that the 100% efficiency is determined at $\text{BDT1} > -1$ and $\text{BDT2} > -1$, thus here, with cut $\text{BDT1} > 0$, the efficiency does not start at 100% but around 95%. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.	31
31	FOM vs. BDT1 with $\text{BDT2} > 0.8$ for both (Left) $A_b^0 \rightarrow A_c^+ \pi^-$ and (Right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.	31
32	Efficiency vs. BDT1 when $\text{BDT2} > 0.8$ for both data (Left) $A_b^0 \rightarrow A_c^+ \pi^-$ and (Right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$. Note again that the 100% efficiency is determined at $\text{BDT1} > -1$ and $\text{BDT2} > -1$, thus here, with cut $\text{BDT2} > 0.8$, the efficiency does not start at 100% but around 50%. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.	32
33	Signal shape from simulated (left) $A_b^0 \rightarrow A_c^+ \pi^-$ and (right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$ signal decays. The results of the fit, described in the text, are superimposed.	33
34	Spectrum of the mass difference, δm , for simulated $\Xi_b^- \rightarrow A_b^0 \pi^-$ decays with (left) $A_b^0 \rightarrow A_c^+ \pi^-$ and (right) $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$, along with the shapes obtained from the fit.	34
35	$A_b^0 p_T$ - η weight distribution, as obtained from the $A_b^0 \rightarrow A_c^+ \pi^-$ sWeighted Run2 data and simulated signal decays.	36

36	$\Lambda_b^0 p_T$ (Left) and η (Right) distributions for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ MC Run2 sample (top) before and (bottom) after the (p_T, η) weight from Fig. 35 is applied.	37
37	(Left) Ratio of the (p_T, η) spectrum in data relative to signal MC for $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ decays, to obtain the Ξ_b^- production spectrum weights. (Right) Ratio of p_T spectrum in data relative to signal MC for $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ decays. The red line is a exponential function fitted to bins from 5 GeV/c to 30 GeV/c	38
38	Projections of the $\Xi_b^- p_T$ and η spectra (top) before and (bottom) after the (p_T, η) weights are applied to $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ MC sample.	39
39	Distributions of BDT2 for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (Right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$, before (black) and after (red) the smearing is applied.	44
40	(Left) Distributions of the BDT1 output for (black) simulation and (red) sWeighted data. (Right) Fraction of events that pass a given BDT1 requirement, along with the fractional difference between data and simulation. The top row is for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode, and the bottom for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode.	46
41	Invariant mass spectrum of (left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the full Run 2 data sample (2016-2018) The fit, as described in the text, is overlaid, with the red dashed line showing the Λ_b^0 signal, green is the combinatorial background, gray shows the partially reconstructed $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- (\pi^+ \pi^-)$ and black shows the misidentified $\Lambda_b^0 \rightarrow \Lambda_c^+ K^- (\pi^+ \pi^-)$	48
42	Invariant mass spectrum of (left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the (top) 2016, (middle) 2107 and (bottom) 2018 data samples The fit, as described in the text, is overlaid, with the red dashed line showing the Λ_b^0 signal, green is the combinatorial background, gray shows the partially reconstructed $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- (\pi^+ \pi^-)$ decays and black shows the misidentified $\Lambda_b^0 \rightarrow \Lambda_c^+ K^- (\pi^- \pi^+)$	49
43	Distributions of δm for $\Lambda_b^0 \pi^-$ candidates with DIRA > 0 (blue) and DIRA < 0 (red).	52
44	Results of the fit to the δm spectrum for the (top) $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and (bottom) Wrong sign $\Lambda_b^0 \pi^+$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the tight BDT2 selection.	54
45	Results of the fit to the δm spectrum for the (top) $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and (bottom) Wrong sign $\Lambda_b^0 \pi^+$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the tight BDT2 selection.	55
46	Results of the fit to the δm spectrum for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the loose BDT2 selection for the r_s measurement.	56
47	Results of the fit to the δm spectrum for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the loose BDT2 selection for the r_s measurement.	57
48	Relative efficiency ϵ^{Rel} of 100 iterations for mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ (Left) and mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ (Right). 2018 MC samples are used.	60

49	Histograms of the number of multiple candidates in one event in simulation (Left) and data samples (Right) for normalization mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ with $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ (Top) and signal mode $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ with $27.8 < \Delta m < 34.8$ (Bottom).	62
50	Distributions of the invariant mass ($\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$ data): (top left) $m(K^+ K^-)$, (top middle) $m(K^+ K^- \pi^+)$, (top right) $m(K^- \pi^+ \pi^+) - m(K^- \pi^+)$ and (bottom left) $m(K^- \pi^+ \pi^+)$. The black histogram corresponds to data before the veto, and red is after the veto.	68
51	Distributions of the opening angle α of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$, in the range from 0 to 2.5 mrad. Pi1/2/3 refers to the $\pi^- \pi^- \pi^+$ from the Λ_b^0 decay.	69
52	A zoom of the histogram of α of all 15 track pairs combined of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$	70
53	Distributions of the opening angle α where PI2 refers to the pion from the Ξ_b^- and Pi1/2/3 refers to the $\pi^- \pi^- \pi^+$ from the Λ_b^0	71
54	Distributions of the opening angle α of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$, in the range from 0 to 2.5 mrad. Pi1 refers to the π^- from the Λ_b^0 decay.	72
55	A zoom of the histogram of α of all 6 track pairs combined of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$	72
56	Distributions of the opening angle α where PI2 refers to the π^- from the Ξ_b^- and Pi1 refers to the π^- from the Λ_b^0	73
57	Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$; signal and background distributions are superimposed.	74
58	Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$; signal and background distributions are superimposed.	75
59	Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$; signal and background distributions are superimposed.	76
60	Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$; signal and background distributions are superimposed.	77
61	Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$; signal and background distributions are superimposed.	78
62	Distributions of the input variables used to train BDT2 for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$; signal and background distributions are superimposed.	79
63	Distributions of the output BDT2 response for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays and WS background. The test and training samples are both shown.	80
64	(Left) RS (black) and WS (red) BDT response distributions in $10 < \Delta m < 32.89 \text{ MeV}/c^2$ mass window. (Right) The ratio $\frac{N_{RS}}{N_{WS}}$ of each bin of the left plot.	80

65	Λ_b^0 p_T - η weight distribution, as obtained from the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ sWeighted data and simulated signal decays.	81
66	Λ_b^0 p_T (Left) and η (Right) distributions before and after weight from Fig. 65	82
67	For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ simulation samples, fitting results of the Ξ_b^- true signal (Top Left) and ghost (Top Right), Λ_b^0 true signal (Bottom Left) and ghost (Bottom Right).	85
68	For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation samples, fitting results of the Ξ_b^- true signal (Top Left) and ghost (Top Right), Λ_b^0 true signal (Bottom Left) and ghost (Bottom Right).	86
69	The N_{TIS} (Left), $N_{TOS TIS}$ (Middle) and $\epsilon_{TOS TIS}$ (Right) versus Ξ_b^- p_T using $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ 2016 simulation sample	97
70	Histograms of the number of multiple candidates in one event in simulation and data samples for normalization mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ (Top) and signal mode $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ (Bottom).	98

List of Tables

1	Two sets of BDT1 and BDT2 requirements applied to MC and data samples. 100% BDT Efficiency and 100% Background Efficiency are set to be with requirements $BDT1 > 0$ or $BDT2 > -1$ applied.	29
2	Fitted signal shape parameters for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays, as obtained from fits to simulated signal decays.	34
3	Fitted signal shape parameters for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ δm shape for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays, as obtained from simulated signal decays.	34
4	Relative acceptance between the signal and normalization modes, as described in the text.	41
5	Relative reconstruction efficiency of the signal and normalization modes. For the final value, the uncertainty is due to finite statistics in the simulated samples.	42
6	Relative offline selection efficiency of the signal and normalization modes. For the final value, the uncertainty is due to finite statistics in the simulated samples.	43
7	Summary of the BDT2 corrections to relative efficiencies.	45
8	Summary of the signal matching corrections to relative efficiencies.	46
9	Summary of the relative efficiencies used to determine r_s . Uncertainties are due to finite statistics in the simulated samples.	47

10	Fitted signal yields inside the (5560, 5680) MeV/c ² mass window and the width scale factor for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ normalization modes for each year and the combined data. The yields of $\Lambda_b^0 \rightarrow \Lambda_c^+ K^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ K^-$ with a misidentified K^- are included.	48
11	Fitted signal yields for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal decay, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the tight BDT2 selection used for observation of the signal, and the loose selection used for the r_s measurement.	53
12	Λ_b^0 yield of the default fitting (bin width 5 MeV/c ² with fixed signal shape), the 2 MeV/c ² bin width fitting with fixed signal shape, and the unfixed signal shape fitting with 5 MeV/c ² bin width.	58
13	Ξ_b yield of the default fitting (inflation coefficient 1.1), and two different coefficient 1.0 and 1.2.	58
14	The efficiencies of 5560 < $m(\Lambda_b^0)$ < 5680 MeV/c ² mass window selection in the simulation and data samples. ϵ^{MC} is the selection efficiency calculated from simulation samples, and f^{Data} is the integral of the normalized fitted signal shape inside mass window 5560 < $m(\Lambda_b^0)$ < 5680 MeV/c ² of data samples.	59
15	Mean and RMS of relative efficiency for 100 iterations. 2018 simulation samples are used.	60
16	the mean number of candidates in one event (See the legends on Fig. 49) after all selections in three-year combined simulation and data samples. For simulation and data, the candidates are limited in the peak region, i.e. 5560 < $m(\Lambda_b^0)$ < 5680 MeV/c ² for normalization modes and 34.8 < Δm < 41.6 MeV/c ² for signal modes.	61
17	Summary of discussed systematic uncertainties.	63
18	Fitted values used in the calculation of r_s . Uncertainties are statistical only.	64
19	Efficiencies and cumulative efficiencies of offline selections for each year.	83
20	For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ simulation samples, summary of true signal and ghost yield, the ratio r and the relative efficiency correction.	84
21	For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation samples, summary of true signal and ghost yield, the ratio r and the relative efficiency correction.	84

1 Standard Model

The Standard Model (SM) of particle physics is the theory describing the electromagnetic, weak and strong interactions and all known elementary particles, shown in Fig. 1. In SM, the elementary particles that form the matter are fermions and classified into two groups: the quarks (u, d, c, s, t, b) and the leptons ($e, \nu_e, \mu, \nu_\mu, \tau, \nu_\tau$). They can be grouped into three generations, as shown. The second and third generations, in the Standard Model, appear to be just replicas of the first with successively larger masses. There is no clear understanding why nature has given us three generations of quarks and leptons. The fundamental interactions are described as the exchanges of gauge bosons (γ, W^\pm, Z, g) as the force carriers between these quarks and leptons.

All quarks and leptons can have weak interactions involving the W^+, W^- or Z bosons, since they all carry weak charge. The electromagnetic interaction between the electrically charged particles is mediated by the photon, γ . These two interactions are unified under the electro-weak (EW) theory, a single gauge theory with the symmetry group $SU(2) \otimes U(1)$.

The strong interaction, however, only happens between quarks, for quarks carry a charge referred to as *colour*. Quarks can strongly interact with other quarks by exchanging gluons g , the force carrier that itself carries color charge (unlike the photon which is electrically neutral under the EM force). Due to the colour confinement, quarks are bounded to form colour-neutral particles called hadrons. The strong interaction is described by quantum chromodynamics (QCD) using the symmetry group $SU(3)$. The Electroweak theory and QCD together describe the fundamental interactions within the Standard Model.

The SM has been very successful to describe experimental data since it was established in the 1970s. The particles it predicted, like the vector bosons W^\pm and Z , the top quark, the τ neutrino, and the Higgs boson, were discovered in a series of high energy particle experiments. But the SM does not account for all of nature's aspects. For example, it does not have a full description of dark matter or gravity, and the CP violation¹ it predicts fails to explain the matter-antimatter asymmetry of the current universe.

The only known source of CP violation in the SM is through the weak interaction. In the SM, the quark's weak interaction eigenstates differ from their flavor eigenstates. A unitary transformation connects the two bases, and is represented as a 3×3 matrix known as the Cabbibo-Kobayashi-Maskawa (CKM) matrix [11].

The transformation between the weak interaction states (d', s', b') and the flavor states (d, s, b) is written

¹ C stands for the charge conjugation operator, which changes a particle to its antiparticle; P stands for the parity operator, which inverts the three spatial coordinate axes; most weak decays are CP invariant but first case of CP violation was observed in 1964 in the neutral kaon system [10]

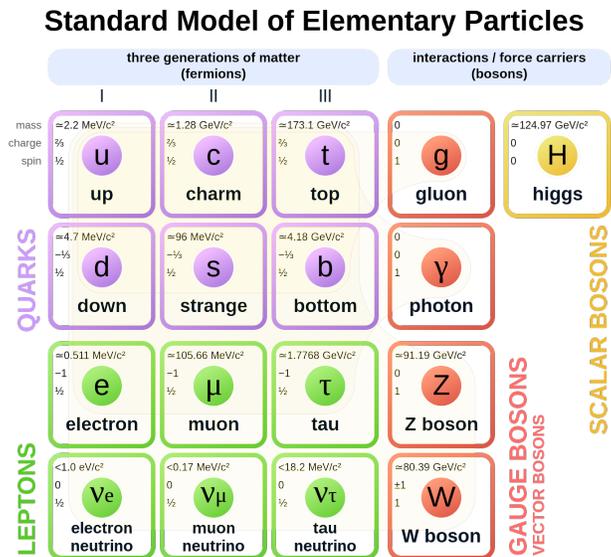


Figure 1: Standard Model of fundamental particles, including the quarks, leptons, and interaction force carriers.

as [12]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1)$$

The 9 complex matrix elements (18 real numbers) can be reduced to four, 3 mixing angles and a single complex phase, by exploiting the nine unitarity constraints, and by a redefinition of the phases of five of the six quark fields. The remaining single complex phase in the CKM matrix allows for CP violation in the Standard Model.

The magnitudes of the CKM elements give the relative strength of the various charged-current weak transitions, and they need to be measured in experiments. Their magnitudes are shown in Fig. 2. The diagonal elements are close to unity, representing large transition strengths for charged-current transitions within a generation of quarks, namely $t \rightarrow b$, $c \rightarrow s$ or $d \rightarrow u$. The off-diagonal elements represent transitions between quarks in generations 1 and 2 (or 2 and 3) and are suppressed by about a factor of five in amplitude (or about a factor of 25 in rate). Weak transitions between generations 1 and 3 (represented by V_{td} and V_{ub}) are clearly suppressed the most. In addition to the magnitudes, the single complex phase needs to be measured precisely to test the explanation of CP violation in the SM (which is insufficient to account for the baryon asymmetry of the Universe).

Measuring the CKM matrix parameters is a key task in flavor physics, and decays of beauty hadrons play

Observable	Central $\pm 1 \sigma$	$\pm 2 \sigma$	$\pm 3 \sigma$
$ V_{ud} $	0.974353 [+0.000049 -0.000056]	0.974353 [+0.000099 -0.000125]	0.97435 [+0.00014 -0.00019]
$ V_{us} $	0.22500 [+0.00024 -0.00021]	0.22500 [+0.00054 -0.00043]	0.22500 [+0.00082 -0.00062]
$ V_{ub} $	0.003667 [+0.000088 -0.000073]	0.00367 [+0.00018 -0.00012]	0.00367 [+0.00028 -0.00017]
$ V_{cd} $	0.22487 [+0.00024 -0.00021]	0.22487 [+0.00054 -0.00043]	0.22487 [+0.00082 -0.00062]
$ V_{cs} $	0.973521 [+0.000057 -0.000062]	0.97352 [+0.00011 -0.00014]	0.97352 [+0.00016 -0.00021]
$ V_{cb} $	0.04145 [+0.00035 -0.00061]	0.04145 [+0.00068 -0.00097]	0.04145 [+0.00094 -0.00122]
$ V_{td} $	0.008519 [+0.000075 -0.000146]	0.00852 [+0.00018 -0.00025]	0.00852 [+0.00026 -0.00035]
$ V_{ts} $	0.04065 [+0.00040 -0.00055]	0.04065 [+0.00075 -0.00084]	0.0406 [+0.0010 -0.0011]
$ V_{tb} $	0.999142 [+0.000018 -0.000023]	0.999142 [+0.000031 -0.000038]	0.999142 [+0.000041 -0.000048]

Figure 2: Magnitude of CKM elements, provided by CKMFilter Group [1]

a central role. Various b -hadron decays are used to measure the magnitudes of four of the CKM elements V_{cb} , V_{ub} , V_{td} , V_{ts} . Other processes can be used to probe the complex phase. So, much can be learned about the CKM matrix through the study of b -hadron decays.

In many cases, getting to the fundamental physics parameters of the SM requires using input from theory. One of the key tools in heavy quark physics is the Heavy Quark Expansion (HQE) [13]. The HQE is a theoretical framework that describes the inclusive decays of hadrons containing a heavy quark, especially for b -hadrons. It includes corrections for the strong-interaction effects between the quarks within a hadron in the computation of weak decay rate, in which the corrections can be expanded in terms of the inverse of the heavy quark mass $1/m_Q$. The HQE can be tested by comparing its predictions to corresponding quantities that can be measured. One such quantity is the total decay width of a particular b -hadron, which is just the inverse of the b -hadron's lifetime τ^{-1} . Thus a measurement of a given b -hadron's lifetime can be used to test the HQE prediction. The predictions become more precise when ratios of b -hadron lifetimes are compared between experiment and theory.

In comparing lifetimes of beauty hadrons to the predictions from the HQE, one must take into account that the HQE predictions only include the decay width of the b -quark plus non-perturbative corrections due to the fact that the b quark is bound inside a hadron [14].

In b -hadrons containing an s quark, such as the Ξ_b^- baryon composed of bds , the decay width predicted in the HQE does not include the contribution where the s -quark decays first, and the b -quark is a spectator.

The decay of the s -quark increases the total decay width, and directly contributes to the lifetime measured in experiments. In particular, $\tau = 1/\Gamma_{tot}$ is measured in experiments, where $\Gamma_{tot} = \Gamma_b + \Gamma_s$, but what is predicted is only Γ_b .

A rough estimate of the contribution of the s quark decay to the Ξ_b^- decay width is obtained by taking the ratio of decay widths, $\Gamma(\Lambda_b^0)/\Gamma(\Xi_b^-) = \tau(\Xi_b^-)/\tau(\Lambda_b^0) \simeq (1.6 \times 10^{-12} \text{ s})/(2.6 \times 10^{-10} \text{ s}) = 0.0062$ [15]. That is, one would naively expect the s -quark to contribute about 0.6% to the total decay width of the Ξ_b^- , if there is no enhancement due to a diquark component in the Ξ_b^- wave function. The remainder of the Ξ_b^- decay width is due to different transitions involving the b quark, with $b \rightarrow c$ being by far the largest contribution.

The three Ξ_b^- decays that could result when the s -quark undergoes the weak decay first are shown in Fig. 3. Because the mass difference between the Λ_b^0 and Ξ_b^- is only about $175 \text{ MeV}/c^2$, there is a limited number of possible final states for the W^- decay. Namely, it can only form a single pion, or the lepton pairs shown.

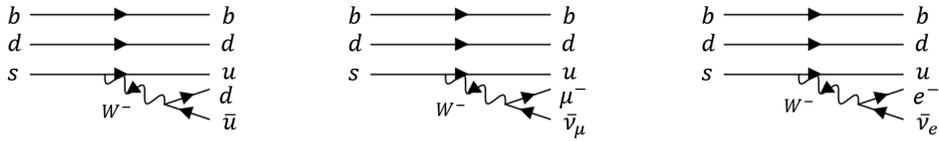


Figure 3: Feynman diagram of (Left) $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, (Middle) $\Xi_b^- \rightarrow \Lambda_b^0 \mu^- \bar{\nu}_\mu$, and (Right) $\Xi_b^- \rightarrow \Lambda_b^0 e^- \bar{\nu}_e$.

It is well known that mesons form due to the strong QCD forces between a quark and an antiquark, It can be shown that, at leading order in QCD, two quarks also exhibit strong binding, half the value as between a quark and antiquark, when the two quarks are in a $J^P = 0^+$ state [16]. Such diquarks can be considered as building blocks to construct baryons as the bound state of a quark and a diquark [17, 18]. Diquarks also provide a convenient way to build up other multi-quark states, including, but not limited to tetraquarks and pentaquarks (see Fig. 4), of which numerous have been seen in recent years ([19, 20]).

In the context of diquarks, it has been shown that there could be a sizeable enhancement to the total decay width of the Ξ_b^- baryon, if the Ξ_b^- baryon wave function is described as a b quark plus a ds diquark [21]. In particular, the contribution Γ_s could be as large as 8% percent, which far exceeds the naive estimate of about 0.6% discussed above. Thus the decay width associated with the s -quark could be enhanced by nearly an order of magnitude over the naive estimate, if the Ξ_b^- baryon is composed of a b and a ds diquark.

A measurement of the size of the s -quark decay rate within the Ξ_b^- baryon could clarify whether there is a sizeable diquark contribution to the Ξ_b^- wave function. This can be achieved by searching for, and measuring, the decay $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$. The decay rate of the two semileptonic decays, $\Xi_b^- \rightarrow \Lambda_b^0 \mu^- \bar{\nu}_\mu$ and $\Xi_b^- \rightarrow \Lambda_b^0 e^- \bar{\nu}_e$,

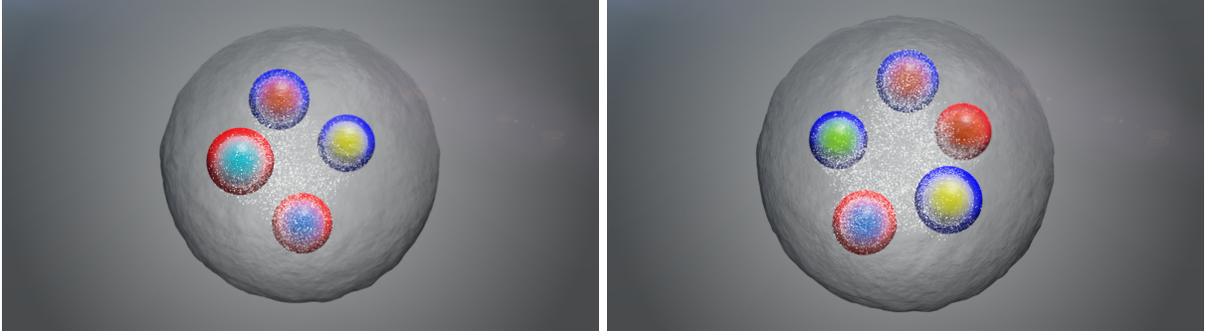


Figure 4: Illustration of a possible layout of the tetraquarks (Left) and pentaquarks (Right) [2]

is expected to be small, considering the ratio of the decay width of the Λ^0 semileptonic decays to that of the Λ^0 pionic decay is in order of 10^{-3} [15]:

$$\frac{\Gamma(\Lambda^0 \rightarrow p\mu^-\bar{\nu}_\mu) + \Gamma(\Lambda^0 \rightarrow pe^-\bar{\nu}_e)}{\Gamma(\Lambda^0 \rightarrow p\pi^-)} = 0.0015 \quad (2)$$

In a previous Run 1 LHCb analysis [22], evidence of the decay, $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ was reported with 3.2σ significance, based a pp collision data sample at $\sqrt{s} = 7$ and 8 TeV corresponding to an integrated luminosity of 3 fb^{-1} . In that analysis, the Λ_b^0 baryons were detected through their decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$. Using the measured ratio, $\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} = (8.2 \pm 0.7 \pm 0.6 \pm 2.5) \times 10^{-2}$ [23], the branching fraction from Run 1 is measured to be $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = (0.7 \pm 0.2 \pm 0.1 \pm 0, 2)\%$.

The previous measurement of $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$ only constituted 3σ evidence for the decay. The goal of this analysis is to observe the decay ($> 5\sigma$) and update the measurement of $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$. The analysis presented in this thesis will benefit not only from having a data sample of $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decays that is about 4 times larger, but also by using an additional Λ_b^0 decay mode, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decays. Thus, a sample of order 6-8 times larger is expected in this analysis. Since $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = \Gamma(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) / \Gamma_{tot}$, this branching fraction measurement is a direct lower bound on the contribution of Γ_s , the s -quark decay contribution to the total Ξ_b decay width.

2 The LHCb Experiment

2.1 Introduction

The European Organization for Nuclear Research, also known as CERN, located near Geneva, Swiss, operates the largest particle physics laboratory in the world. The main function of CERN is to provide particle accelerator facilities, therefore many accelerators were built through the history of CERN since 1954. The Large Hadron Collider (LHC), constructed at CERN, is the current largest and highest-energy particle accelerator of the world.

The Large Hadron Collider (LHC) is a 27-kilometer particle accelerator, which sends two beam of protons in opposite directions and make the proton bunches collided in one of four detectors, which is shown in Fig. 5. In RUN II, from 2015 to 2018, LHC was running at center of mass energy of pp collisions of $\sqrt{s} = 13$ TeV and achieved a peak luminosity 500 Hz/ μb for LHCb in 2018 [24] (for other detectors the luminosity is much higher). The Large Hadron Collider beauty (LHCb) is one of major experiments constructed at LHC and is dedicated to search for the new physics in CP violation and rare decays of b and c quark hadrons with highly precise measurements [5].

In LHC, the $b\bar{b}$ pairs are produced primarily through gluon-gluon fusion, and the angular distributions of those pairs fall into a very small range around the pp beam line in both direction, as shown in Fig. 6. Therefore the LHCb detector was designed as a single-arm forward spectrometer with the forward angular coverage from 10 mrad to 300 mrad in the horizontal magnetic-bending plane and from 10 mrad to 250 mrad in the vertical plane [25]. The side view is in Fig. 7.

LHCb is a complex detector with multiple subsystems which include a dipole magnet, the tracking system, the particle identification (PID) and the trigger system. Each subsystem will be discussed.

2.2 Dipole Magnet

In order to measure the momentum of charged particles, a dipole magnet is used to bend their tracks. It consists of two identical saddle-shaped coils, and they are placed with an open angle to meet the requirements of track acceptance. LHCb experiment measures the momentum of charged particles with a relative uncertainty varying from 0.5% at low momentum to 1.0% at 200 GeV/ [26]. Therefore the dipole magnet provides the 4 T · m of integrated magnetic field in the area between the tracking station TT and T1 (seen in Fig. 7), and meanwhile keeps the field level inside the RICHs envelope (seen in Fig. 7) as low as possible [6]. In RICH1, iron shield and a magnetic distortion calibration system are deployed to mitigate the influence of the fringe

magnetic field. A perspective view of the dipole magnet and the main component of magnetic field B_y are shown in Fig. 8

2.3 Tracking system

The LHCb tracking system consists of multiple tracking stations: the vertex locator system (VELO) surrounding the pp collisions region, the Track Turicensis (TT) upstream of the dipole magnet and T1-T3 downstream of the magnet [25]. The schematic in Fig. 9 presents their relative position to the magnetic field and the types of tracks they are to detect.

2.3.1 Vertex locator

The VELO can precisely measure the coordinates of tracks near the pp collision region and thus identify and locate the primary vertex (PV) and the decay vertices of hadrons containing b and c quarks. To achieve this goal, VELO uses 21 silicon modules in each half, shown in Fig 10. Each module consists of two silicon half discs, one with strips in radial to measure r coordinate of tracks and one in polar to measure ϕ . The discs have only 300 μm thickness in order to reach the minimal multiple scattering rate. Since the VELO is positioned very close to the pp collision region and endures high radiation damage, it is designed to be movable: the sensors are retracted during the beam injection and moved in for data taking when the beams are stable [6].

As a typical example of VELO performance, PV can be measured with resolution of 13 μm in the transverse plane and 71 μm along the beam axis for vertices with 25 tracks, and an impact parameter resolution of less

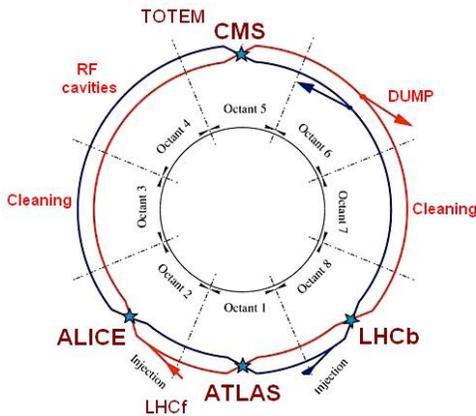


Figure 5: Layout of LHC. Two proton beams in the opposite directions (red and blue lines) can collide in one of four detectors: CMS, ALICE, ATLAS or LHCb. [3]

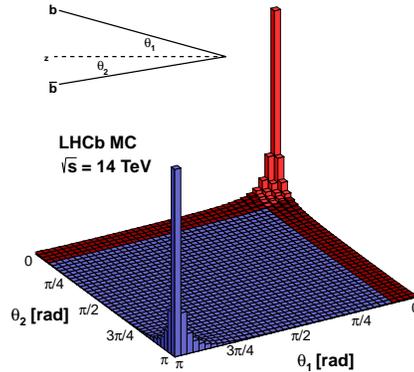


Figure 6: Simulated angular distribution of $b\bar{b}$ pairs. [4]

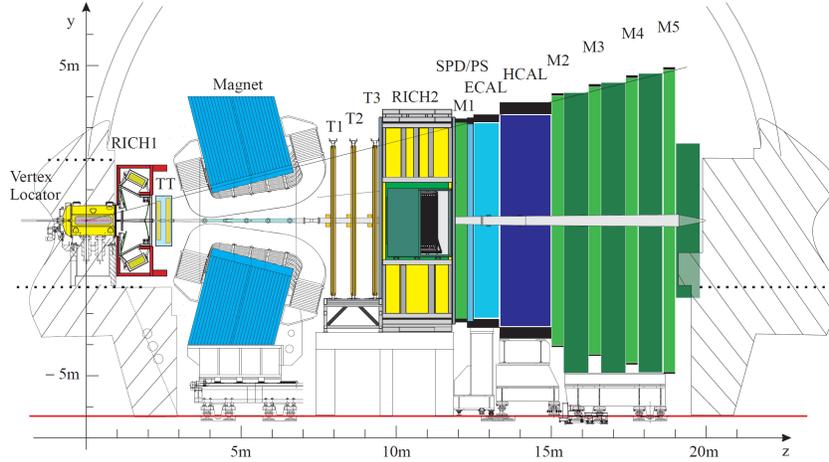


Figure 7: LHCb detector layout, side view [5].

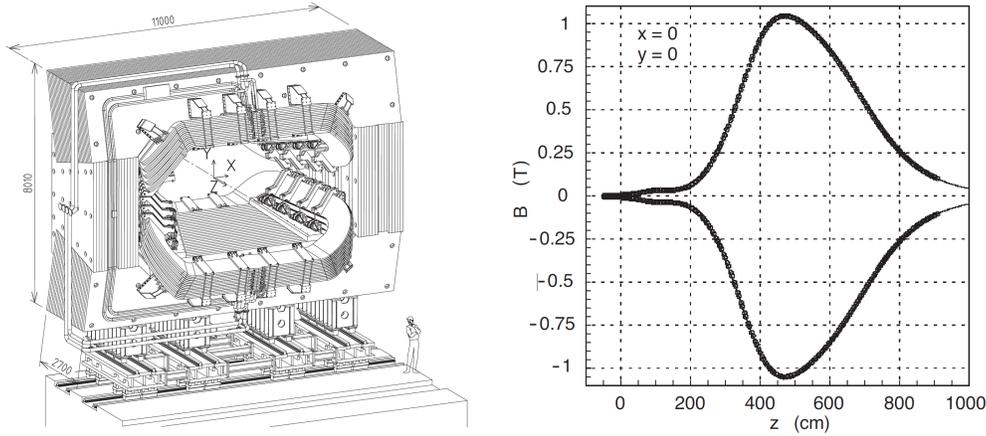


Figure 8: The perspective view of the LHCb dipole magnet (Left) and the magnetic field component B_y along the z -axis, i.e. the beamline (Right). [6]

than $35 \mu\text{m}$ is achieved for particles with transverse momentum greater than $1 \text{ GeV}/c$ [27].

2.3.2 Planar tracking stations

Following VELO, there are four planar tracking stations, the Track Turicensis (TT) positioned upstream of the dipole magnet and T1-T3 downstream of the magnet. Each station of the T1-T3 can be further divided into two parts: the region around the beam pipe is called Inner Tracker (IT) and the rest region is Outer Tracker (OT). The layout of TT, IT and OT can be seen in Fig 11.

TT and IT cover the region of highest particle flux and require the finest resolution, therefore every station is formed by 4 overlapping layers that are made of silicon microstrip sensors with strip pitches of

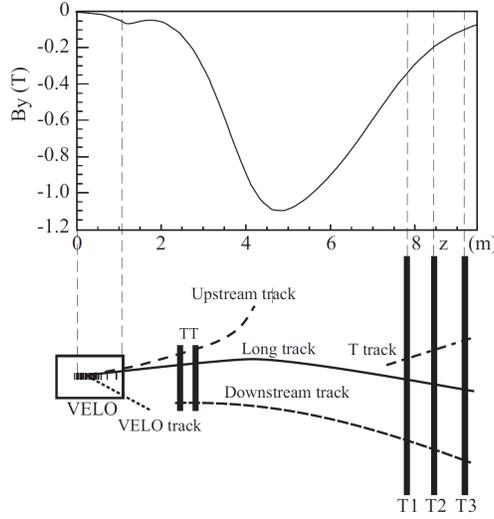


Figure 9: Schematic of VELO, TT and T1-T3's positions relative to the magnetic field B_y and the types of track they monitor. [6]

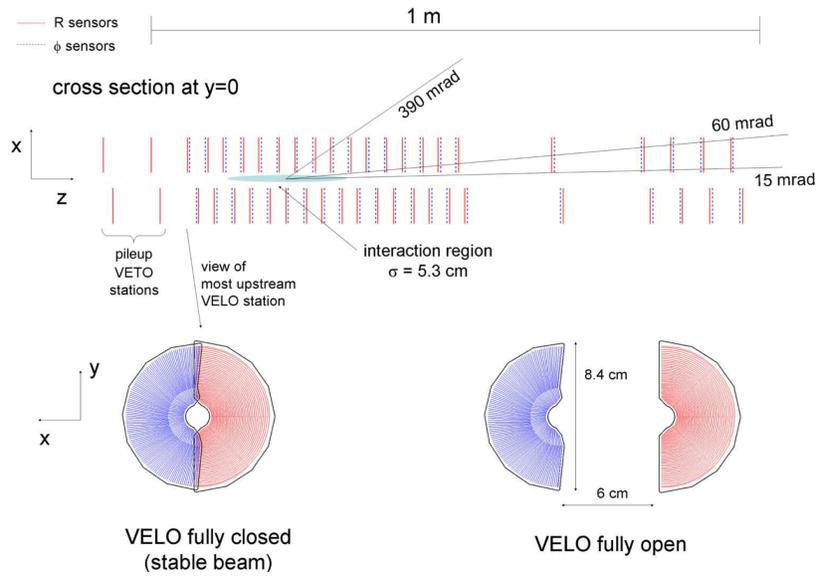


Figure 10: The schematic of VELO showing the layout of r and ϕ silicon sensors, and also the arrangement of modules along the z -axis. The bottom shows that the VELO's position when it is fully opened or closed. [6]

about $200 \mu\text{m}$. When a charged particle flies through the strips, electron-hole pairs are created along its path and separated by the reverse bias voltage. The electrons (holes) drift to the positive (negative) voltage connection and the drift of these charges generates a current pulse which is amplified and collected by external electronics. Same for all silicon track stations, the first and fourth layers are placed vertically and measure the x coordinate of tracks, while the two middle layers are rotated by a stereo angle of $\pm 5^\circ$ respectively. By such method the spacial resolution is approximately $50 \mu\text{m}$.

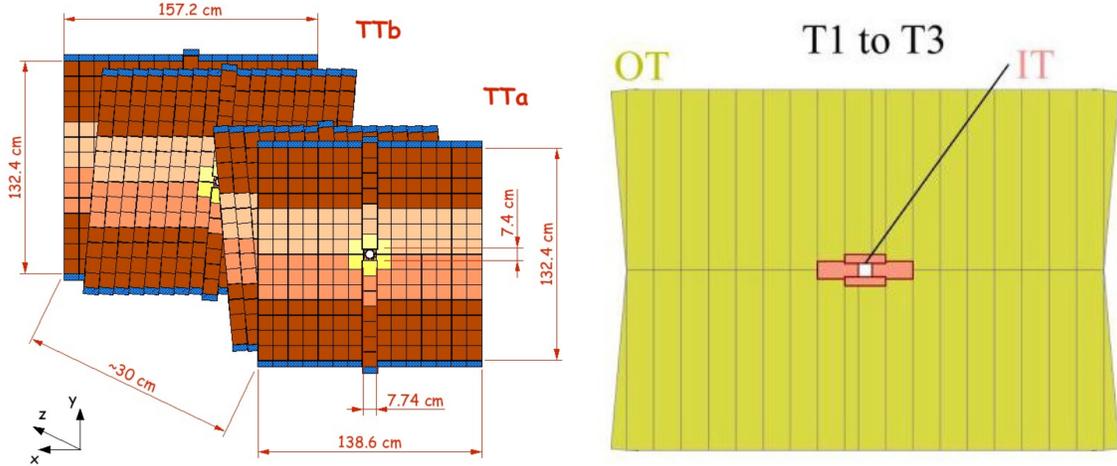


Figure 11: (Left) The geometry of four layers of TT and the silicon sensors arrangement in each layer. (Right) Layout of IT and OT of T1-T3 stations. [7]

The OT covers the large area of the rest of required LHCb track acceptance where the particle flux is smaller. OT is made from approximately 200 straw-tube drift chambers, which are filled with Ar – CO₂ mixed gas, and provides a 205 μm spatial resolution [28].

Combining tracking information from VELO, TT and T1-T3 together, the overall tracking efficiency for "long" tracks (see Fig. 9) is greater than 96% for $5 < p < 200 \text{ GeV}/c$. Long tracks are those tracks that have segments reconstructed through the entire spectrometer, and are the only kind used in this analysis, although other types of tracks are used in other analyses (for example those analyses that analyze the decays of long-lived particles, such as K_S^0 .) The momentum resolution dp/p for long tracks varies from 0.5% to 1.1% as the momentum increases from low range to 240 GeV/c [29].

2.4 Particle identification

Particle identification (PID) is another essential component of LHCb, which sorts out the type of particles whose tracks are recorded by the tracking system. There are three different kind of subdetectors in PID system: two Ring Imaging Cherenkov Detectors (RICH1 and RICH2) which separate K , π and p , two Calorimeters differentiating hadrons, electrons and photons, and a Muon system formed by five stations M1-M5 to identify and track muons.

2.4.1 RICH

In many studies of B hadron decay, the discrimination of K , π and p can provide great a advantage in background rejection. But this can be challenging in LHCb due to the high luminosity and the wide range of

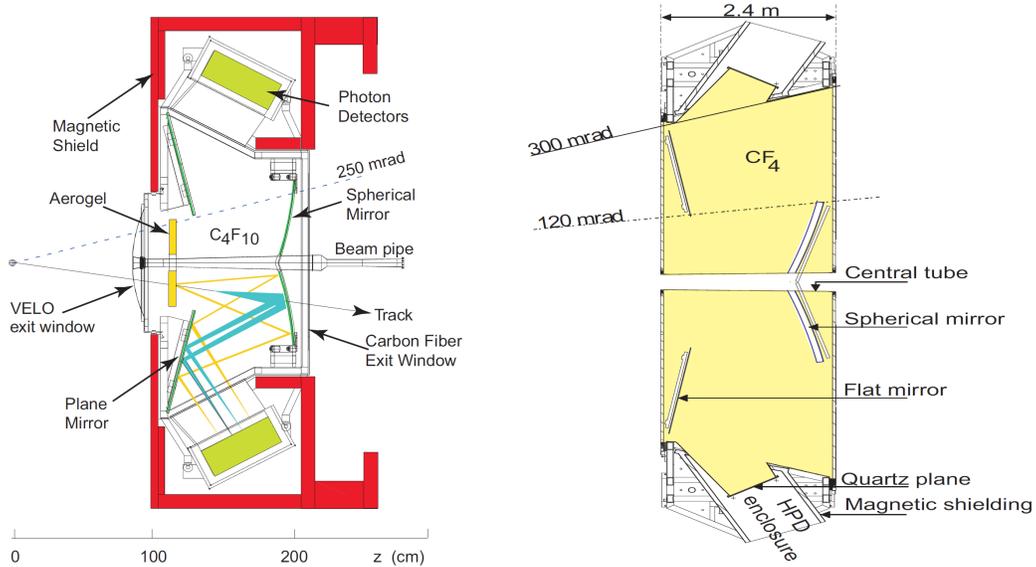


Figure 12: (Left) Side view schematic of RICH1. (Right) Top view schematic of RICH2. [6]

momentum of these final state particles. LHCb employed two Ring Imaging Cherenkov Detectors, RICH1 and RICH2, to cover the charged particle momentum range from 1 GeV/c to more than 100 GeV/c.

RICH1 is located downstream of the VELO and sorts charged particles of low and intermediate momentum: 1 ~ 60 GeV/c. RICH2 is located downstream of T1-T3 and the magnet, covering the high momentum range from ~ 15 GeV/c up to and beyond 100 GeV/c. Since RICH1 is adjacent to VELO, it covers the full acceptance of LHCb. RICH2 is further away downstream, thus it has a limited angular acceptance while still covers the region where high momentum particles are produced.

Inside each RICH, the chamber is filled with a specific type of gas as the medium of travelling particles. When a charged particle moves in a medium with a velocity larger than the local light speed, it emits Cherenkov radiation and the cosine of Cherenkov angle is inversely proportional to the particle's velocity: $\cos(\theta_c) = c/(nv_p)$, where θ_c denotes Cherenkov angle, n is the refractive index and v_p is the particle velocity traveling in the medium. RICH1 and RICH2 collect the Cherenkov photons by reflecting them to Hybrid Photon Detectors (HPD) with flat and spherical mirrors. Fig. 12 shows the inside structure of RICH1 and RICH2. In RICH1 the optical layout is vertical but in RICH2 is horizontal.

Combining the information obtained by RICH and the momentum measured by the magnet and tracking system, the mass of one particle can be identified. The θ_c as a function of momentum in different mediums is shown in left of Fig. 13. RICH1 uses Aerogel (in RUN1) and C₄F₁₀ as they have good performance for low momentum particles and RICH2 uses CF₄ which can separate π and K in the high momentum range [6].

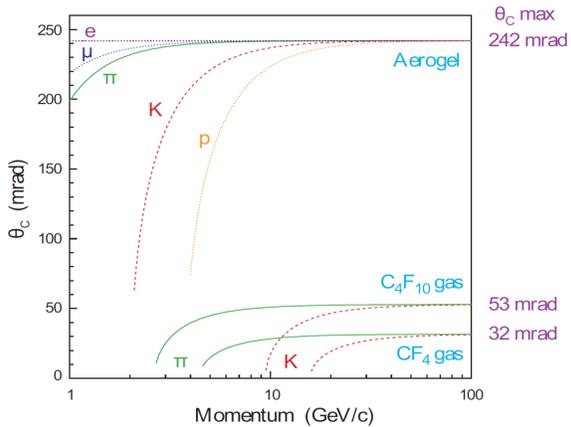


Figure 13: θ_c as a function of momentum in different medium. [6]

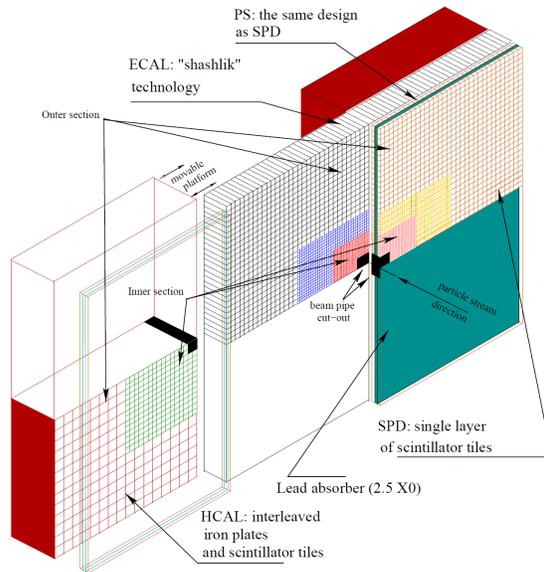


Figure 14: Layout of the calorimeters. [8]

2.4.2 Calorimeters

The Calorimeters provides transverse energy information of various particles to L0 trigger, which will be discussed in the trigger section. It also performs the offline identification on electrons, photons and π^0 , and measure their energies and positions as well. The energy information is also used in the trigger system for fast decisions and this will be explained in the next section. There are 4 subdetectors of Calorimeter system installed in LHCb: the Scintillating Pad Detector (SPD), the Pre-Shower (PS), the Electromagnetic Calorimeter (ECAL), and the Hadronic Calorimeter (HCAL). Their structures are similar: a grid of cells where multiple layers of absorbers and scintillator plates are combined together, seen in Fig. 14. When particle showers interact with absorbers, the scintillation light is collected and transported by the optical fibers to photomultiplier tubes (PMT) and then read out by periphery electronics. Like the design of tracking stations, in order to accommodate the high particle density near the beam line, the calorimeter cells close to the beam are smaller compared to ones in outer region. This can also be seen in Fig. 14 where the size of cells are marked by different colors [8].

To identify particles, each subdetector uses a chosen material (lead or iron) as the absorber. The first station from upstream is SPD/PS, and they both are walls of scintillating pads with similar design but separated by a 2.5 radiation length lead layer which works as the absorber. SPD differentiates electrons, photons and pions, and the PS information is used to determine whether the particle is neutral or charged.

The second station is ECAL which measures the transverse energy of electrons, photons and π^0 for the

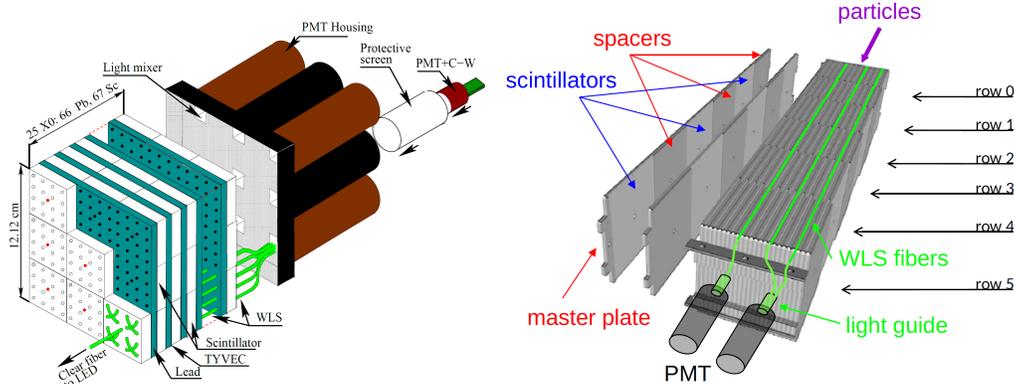


Figure 15: (Left) ECAL cell. (Right) HCAL cell. [8]

L0 trigger. It also uses lead as the absorber but ECAL cell thickness is 25 radiation length, formed by 66 layers of a module consisting of 2 mm lead stacked with 4 mm scintillator, shown in left plot of Fig. 15. This is to ensure the full containment of the high energy electromagnetic showers and to get the optimal energy resolution. For a cell of ECAL, the energy resolution measured in RUN1 and RUN2 is [8].

$$\frac{\sigma(E)}{E} = \frac{(13.5 \pm 0.7)\%}{\sqrt{E_{\text{GeV}}}} \oplus (5.2 \pm 0.1)\% \oplus \frac{(320 \pm 30)\text{MeV}}{E_{\text{GeV}}}$$

where E is the particle energy in GeV, θ is the angle between the beam axis and a line from the LHCb interaction point to the centre of this given ECAL cell.

HCAL sits behind ECAL and measures the transverse energy of hadrons for the L0 trigger. The cells of HCAL are made from iron layers and scintillating tiles which are placed parallel to the beam axis, seen the right plot of Fig. 15. This orientation of scintillating tiles enhances the light collection compared to a perpendicular orientation. Due to the space limitation of LHCb, HCAL thickness is 5.6 radiation length and the resolution measured in test-beam with pions is [30]

$$\frac{\sigma(E)}{E} = \frac{(67 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%$$

2.4.3 Muon system

LHCb installed five stations M1-M5 to identify and track muons, since muon presents in final states of many interesting B decays and the identification of muon is a fundamental requirement of studying such decays. But this analysis does not rely on muon information much, therefore this is a brief introduction of

the muon system.

M1 is located upstream of SPD/PS, and Station M2-M5 are positioned downstream of the calorimeters, separated by 80 cm iron absorbers to stop the redundant hadrons passed calorimeters. This results the minimum momentum for muons crossing all five muon stations to be about 6 GeV/ c . All five stations provide fast p_T information for the L0 trigger. The acceptance of the muon system is $20 \sim 306$ mrad in the bending plane (horizontally) and $16 \sim 258$ mrad in the non-bending plane (vertically). To cover this acceptance, the transverse dimensions of each muon station grow as the station's position is further away from pp collision point, shown in Fig. 16. The complete muon system is mainly formed by 1368 Multi Wire Proportional Chambers, and the inner region of M1 used 12 Triple GEM chambers because the large particle flux in this area may cause ageing problem for the former type of detector [31].

A reconstruction of muon track and the measurement of p_T requires aligned hits in all five stations. Station M1-M3 have a high spatial resolution on x coordinate (bending plane) and measures the track direction and the p_T with a resolution of 20%. M4 and M5 is mainly used to identify the penetrating particles and have a limited spatial resolution.

2.5 Trigger

LHCb operates at a peak luminosity of around 500 Hz/ μb in 2018 [24] and 40 MHz pp bunch crossing rate. However, only a small fraction of pp interaction generates the $b\bar{b}$ pairs, and the processing of the tremendous number of detected signals is also limited by the bandwidth and computing resources LHCb can utilize. Therefore a trigger system that can efficiently select the interesting $b\bar{b}$ events and reject the majority of background collisions is needed. The LHCb trigger has two levels, the Level-0 trigger (L0) and High-level trigger (HLT). Together the two level trigger system reduce the 40 MHz pp interaction event rate to 12.5 kHz so the data of signals can be stored on tape for further offline analysis. The procedures are briefly summarized in Fig. 17.

2.5.1 Level-0

Level-0 (L0) trigger is based on custom electronics board and runs at 40 MHz to process signals in real-time. It reduces the data rate down to 1 MHz by utilizing track's transverse momentum p_T or transverse energy E_T information collected by Calorimeters, Muon system and VELO detectors. L0 contains two components: L0 Calorimeter trigger and L0 Muon trigger.

The L0 calorimeter trigger uses the component of the deposited energy in the SPD/PS, ECAL and HCAL

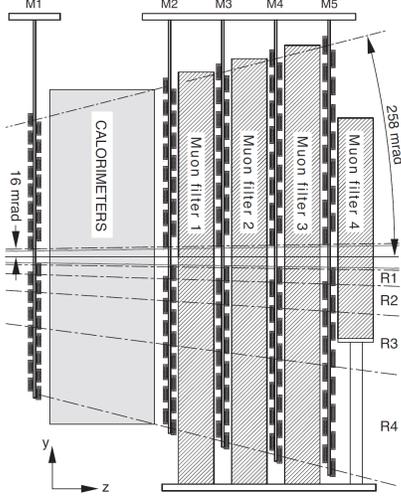


Figure 16: Side view of Muon system. [6]

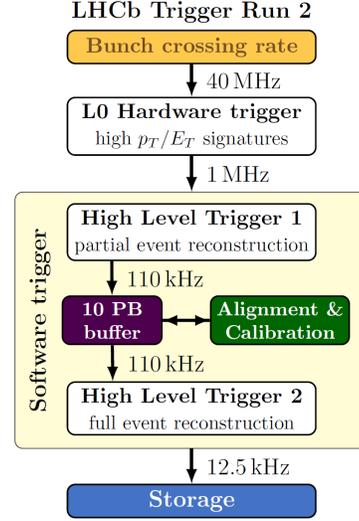


Figure 17: Overview of LHCb trigger system. [9]

transverse to the beamline. The decision to trigger an event depends on the transverse energy E_T deposited in a cluster of 2×2 cells in the ECAL and HCAL, and E_T is defined as [9]

$$E_T = \sum_{i=1}^4 E_i \sin \theta_i$$

where E_i is the energy deposited in cell i and θ is the angle between the z -axis and a line from the cell centre to the average pp interaction point. Particle's identification information from SPD/PS is also used and ultimately all information of the pp interactions that generated highest E_T particle per type (photon, electron, π^0 and hadron) are stored.

The L0 Muon trigger uses p_T and position information from all muon stations, looking for hits forming a straight line through five muon stations and pointing back to the pp interaction point in the $y - z$ plane. The complete muon system is divided into 4 quadrants according to the distance from the beam pipe, which is marked as R1-R4 in Fig. 16, and each quadrant has roughly the same signal occupancy. The two tracks with the largest and second largest p_T in a quadrant are stored. Details can be found in the technical design report [32].

2.5.2 High level trigger

Events passing the L0 selection are processed by the High-level trigger (HLT). HLT process is executed on a computing nodes cluster called Event Filter Farm (EFF) and it is divided into two stages: HLT1 and

HLT2. HLT1 partially reconstructs the event and reduce the 1 MHz L0 rate to 110 kHz. Then the HLT2 further reduces this rate to 12.5 kHz by performing a more complete event reconstruction.

HLT1 determines the location of PV by using the information provided by VELO, and subsequently searches for particles that are inconsistent with having been come directly from any pp interaction in the event. To do this, HLT1 reconstructs trajectories of charged particles traversing the all LHCb tracking stations which have $p_T > 500$ MeV. The hits VELO detects are combined to form a straight-line pointing towards the beam line. This straight-line then get extrapolated to TT region, requiring at least 3 hits in a small area of TT around this possible track from VELO. The extrapolation is also extended to T-stations with the search window defined by the maximum possible magnetic bending effect on charged particles of $p_T > 500$ MeV. All tracks are fitted with a Kalman filter to optimize the estimation and the set of fitted VELO tracks is re-used to determine the position of PVs. By such process, events with a single displaced track with high p_T or a displaced two-track vertex with high p_T are selected. There are also HLT1 muon trigger lines that select for muonic decays of b and c hadrons and an introduction of such lines can be found in Ref. [9].

Based on the reduced rate of HLT1 selected events, HLT2 can utilize full information from all subdetectors of LHCb and perform a full event reconstruction within mainly three steps: reconstructing the track of charged particles, reconstructing the neutral particles, and particle identification (PID). In these steps, same reconstruction executed in HLT1 is repeated with loose p_T constraint. Additional information that are not available to HLT1 is also included, for example, tracks of long lived particles that decay outside the VELO can be reconstructed in HLT2. As a result, high quality long and downstream tracks (seen in Fig. 9) are reconstructed with precise momentum estimation, and PID information from RICH and Calorimeters is applied.

HLT2 inclusive b hadrons lines are important to this analysis, which look for a two-, three-, or four-track vertex with sizable p_T , significant displacement from the PV, and a topology compatible with the decay of a b hadron [9]. Those lines depend on a multivariate selection of the displaced vertex to separate b hadron decays from combinatorial background that mainly consists of displaced c hadron decay vertices or random combinations of charged particles that do not come from a b -hadron.

3 Analysis of the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay

3.1 Overview of the analysis

The signal decay that we want to observe in this analysis is $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$. The strategy is similar to that performed in the Run 1 analysis [22], except here, we consider two Λ_b^0 decay modes, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$. In each case, the discriminating variable for the decay is the mass difference, $\delta m \equiv m(\Xi_b^-) - m(\Lambda_b^0) - M(\pi^-)$, where the small m refers to an invariant mass and the capital M refers to a known mass [33]. From the world-average masses of the Ξ_b^- and Λ_b^0 baryons, a peak at $38.2 \text{ MeV}/c^2$ is expected for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay.

To access the branching fraction $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$, the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal yield is normalized to the inclusive Λ_b^0 yield. For the case where the Λ_b^0 is reconstructed in the $\Lambda_c^+ \pi^-$ final state, the yield ratio can be written as:

$$\frac{N(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-)}{N(\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-)} = \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) \frac{\epsilon_{\Xi_b^-}}{\epsilon_{\Lambda_b^0}}. \quad (3)$$

where $f_{\Xi_b^-}$ and $f_{\Lambda_b^0}$ are the fragmentation fractions for $b \rightarrow \Xi_b^-$ and $b \rightarrow \Lambda_b^0$, and $\epsilon_{\Xi_b^-}$ and $\epsilon_{\Lambda_b^0}$ are the total selection efficiencies. A similar expression applies to the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay mode. Thus, the quantity that is measured in this analysis is

$$r_s \equiv \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-). \quad (4)$$

Thus, the measurement requires the ratio of yields and a ratio of efficiencies. The ratio of fragmentation fractions has been measured at $\sqrt{s} = 13 \text{ TeV}$ based on an independent measurement, and is found to be $\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} = (8.2 \pm 0.7 \pm 0.6 \pm 2.5) \times 10^{-2}$ [23]. Using this value, $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$ is readily obtained. The dominant uncertainty in the measurement of $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$ is the SU(3) assumption used in the measurement of $\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}}$.

Throughout this note, *signal mode* refer to the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay, and *normalization mode* refers to the inclusive Λ_b^0 sample (either $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ or $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$). The contribution of Λ_b^0 decayed from Ξ_b^- is at the level of 1/200 of the inclusive Λ_b^0 , therefore the background level, mainly formed by a Λ_b^0 with a random π^- (seen in Fig. 18), is expected to be high. Exploit of the topology difference between the signal decay and the main background is an efficient method to increase the signal-to-background ratio and will be discussed in later sections.

The basic flow of the analysis is as follows:

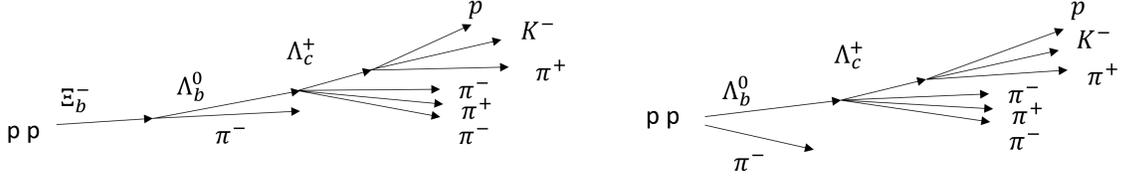


Figure 18: Cartoon showing the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay (left) and the main background (right), a Λ_b^0 plus a random π^- .

- Apply loose selections on $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ candidates;
- Apply a multivariate discriminant to suppress background in the Λ_b^0 sample (BDT1);
- Develop and optimize a second multivariate discriminant (BDT2) to discriminate $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal from $\Lambda_b^0 \pi^-$ combinatorial background.
- Perform a simultaneous fit to the δm distributions using the two Λ_b^0 modes ($\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$) to obtain the Ξ_b^- signal yields in each.
- Obtain the signal yield of all $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decays from a fit to the Λ_b^0 invariant mass spectra;
- Determine weights to bring the kinematics of the simulated Λ_b^0 and Ξ_b^- decays into agreement with data;
- Using the simulation with the weights obtained previously, compute the efficiency ratio, $\frac{\epsilon_{\Xi_b^-}}{\epsilon_{\Lambda_b^0}}$;
- Compute r_s for each Λ_b^0 sample, and compute the weighted average value, and then compute $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$.

Simulated decays are generated with PYTHIA 8 [34,35] and is used to determine the relative selection efficiency and the mass shape of the signal.

3.2 Data selection

The data used for this analysis uses pp collisions at a center of mass energy of 13 TeV, corresponding to an integrated luminosity of 5.5 fb^{-1} , accumulated during the 2016, 2017 and 2018 data runs.

We search for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay by combining a Λ_b^0 candidate with a π^- candidate. The Λ_b^0 baryons are reconstructed in two decay modes, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$, to improve the statistical precision of the measurement. In both cases, the Λ_c^+ baryon is reconstructed in the $pK^- \pi^+$ final state.

Several levels of selection requirements are applied. First, a number of selections are applied in the so-called stripping lines. Stripping lines are a set of algorithms applied to the raw data to select a wide range of final states, most notably those corresponding to specific b -hadron decays. A stripping line typically produces candidates for a single, specific decay, but in some cases it could be several topologically-similar decay modes. They typically suppress the background by at least a factor of 200 while retaining high efficiency for the decay mode of interest (here $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ or $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$). To further suppress combinatorial background, additional relatively loose offline selections are applied. Lastly, a multivariate discriminator is used to further improve the signal-to-background ratio. Where possible and sensible, selection requirements applied to the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ candidates are the same.

3.3 Stripping selections on $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \pi^- \pi^+ \pi^-$

The $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ candidates are pre-selected from two stripping lines. The $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ candidates are reconstructed as $\Lambda_b^0 \rightarrow \Lambda_c^+ a_1^-$, with $a_1^- \rightarrow \pi^- \pi^+ \pi^-$. As a result, the a_1^- pseudo-particle has all the usual properties of a reconstructed composite particle, including vertex-related quantities, invariant mass, etc. Whenever we use the particle notation a_1^- , it refers to the 3π system in this decay mode. There are numerous selections applied in the two stripping lines, and they are listed in Appendix. A.

3.4 Offline selections common to $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$

Substantial background still remains even after the selections imposed in the stripping lines, and needs to be further reduced. There are two specific backgrounds that can be easily identified and removed with some additional selections.

3.4.1 Vetoes on clone tracks

In some cases a set of hits in the VELO (from a single particle) produces two distinct VELO track segments that are almost overlapping in space, and each of these tracks may lead to unique long tracks. Apart from photon conversions in the VELO, such overlapping pairs of tracks are highly improbable for a signal decay. For this reason such pairs are identified and removed by requiring the angle $\alpha = \cos^{-1}(\frac{\vec{P}_1 \cdot \vec{P}_2}{|\vec{P}_1| |\vec{P}_2|})$ between any pair of final-state particles be larger than 0.8 mrad.

For example, in the mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$, $\Lambda_c^+ \rightarrow p^+ K^- \pi^+$, there are six track pairs and the α distribution of all pairs combined is shown in Fig. 19. The spike near $\alpha = 0$ can be easily removed by requiring $\alpha > 0.8$ mrad.

Other relevant plots can be found in the Appendix. D.

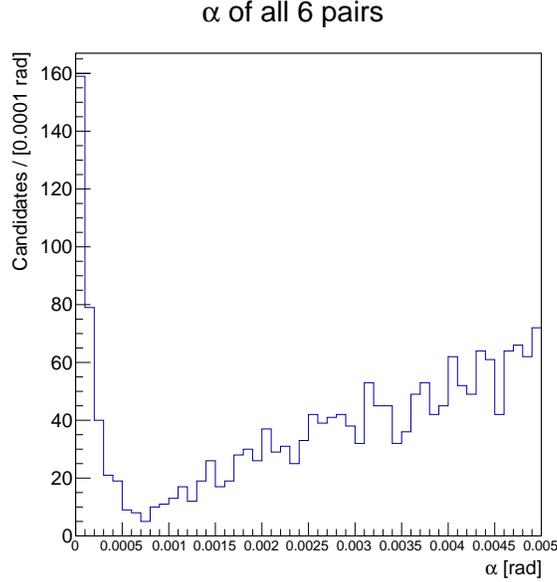


Figure 19: α distribution of all 6 track pairs combined of decay $A_b^0 \rightarrow \Lambda_c^+ \pi^-$, $\Lambda_c^+ \rightarrow p^+ K^- \pi^+$.

3.4.2 Vetoes on other charm final states

The identification of p^+ , K^- and π^+ (and their antiparticles) cannot be perfect, and these hadrons could be incorrectly identified. For example, a K^+ could be misidentified as p^+ or π^+ , which is one of most common cases in our modes of interest. A significant contribution comes from partially reconstructed B^0 , B^- and B_s^0 decays that have a charm meson in the final state. For example, if a π^+ in the $\bar{B}^0 \rightarrow (D^+ \rightarrow K^- \pi^+ \pi^+) \pi^-$ is misidentified as a proton, the D^+ could mimic a Λ_c^+ decay. In these and similar cases, these backgrounds can be identified by narrow peaks in the mass or mass difference spectra for signal candidates when trying alternative mass hypotheses, e.g. proton in Λ_c^+ assigned the K^+ mass when searching for $D_s^+ \rightarrow K^+ K^- \pi^+$ contamination.

This kind of background is removed using a combination of mass windows and stricter PID requirements. We employ a similar strategy to what was done in Refs. [36,37]. Specifically, we require:

- $|m(K^+ K^-) - M(\phi)| > 8 \text{ MeV}/c^2$ or $\text{PROBNN}p > 0.7^2$ for the proton candidate;
- $|m(K^+ K^- \pi^+) - M(D_s^+)| > 25 \text{ MeV}/c^2$ or $\text{PROBNN}p > 0.4$ for the proton candidate;
- $|m(K^+ K^- \pi^+) - M(D^+)| > 15 \text{ MeV}/c^2$ or $\text{PROBNN}p > 0.45$ for the proton candidate;

² $\text{PROBNN}p$ [38,39] is a particle identification (PID) variable and indicates how probable this candidate is a true proton. The larger the value, the candidate is more likely to be a true proton.

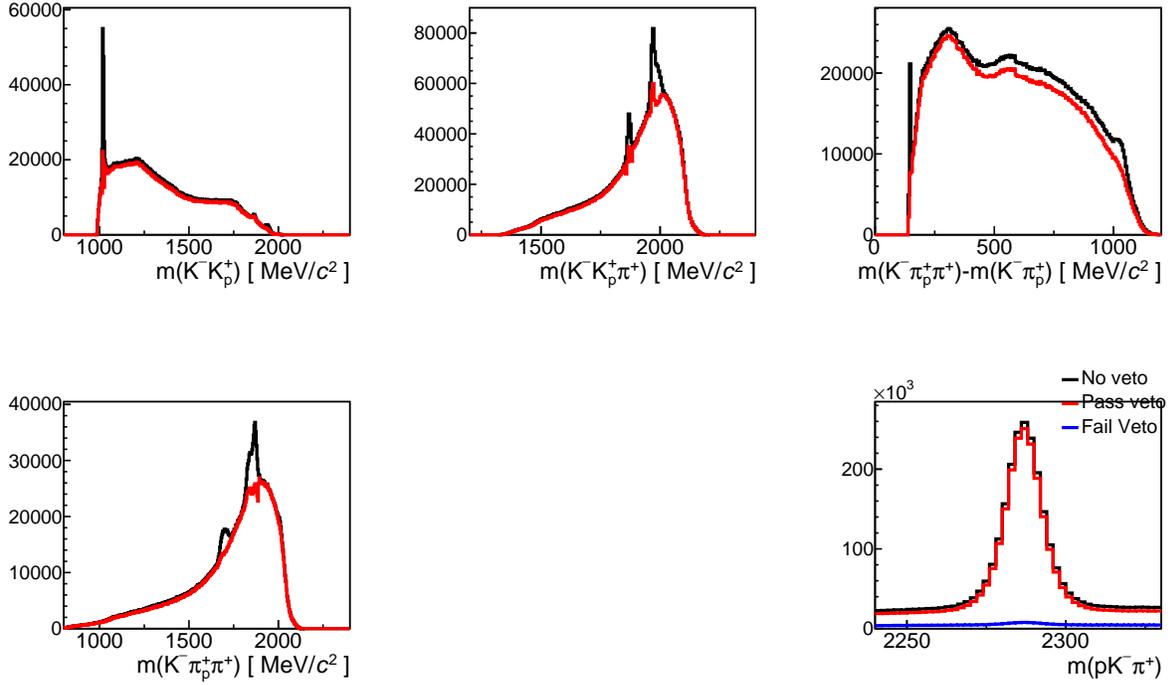


Figure 20: Distributions of the invariant mass ($\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ data): (top left) $m(K^+ K^-)$, (top middle) $m(K^+ K^- \pi^+)$, (top right) $m(K^- \pi^+ \pi^+) - m(K^- \pi^+)$ and (bottom left) $m(K^- \pi^+ \pi^+)$. The black histogram corresponds to data before the veto, and red is after the veto.

- $|m(K^- \pi^+ \pi^+) - m(K^- \pi^+)| > 152 \text{ MeV}/c^2$ or $\text{PROBNnp} > 0.5$ for the proton candidate;
- $|m(K^- \pi^+ \pi^+) - M(D_s^+)| > 20 \text{ MeV}/c^2$ or $\text{PROBNnp} > 0.45$ for the proton candidate;

The mass distributions before and after the veto for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decay are shown in Fig. 20. The distributions for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$ are very similar and can be found in Appendix C. Before the veto, contributions from ϕ , D^+ , D_s^+ and D^{*+} mesons are clearly seen, and after the veto, their contributions have been ameliorated. We also show the $m(K^- \pi^+ \pi^+) - m(K^- \pi^+)$ mass distribution before and after the above vetoes. The third of these vetoes is sufficient to remove contributions from either $D^{*+} \rightarrow D^0 \pi$, with $D^0 \rightarrow K^+ K^-$ or $D^0 \rightarrow K^- \pi^+$, due to the cut on the mass difference.

This veto is highly efficient, and is applied to both the Λ_b^0 normalization mode and the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal mode (as well as the WS background modes, which are discussed later), and therefore the overall efficiency factor cancels in the r_s measurement.

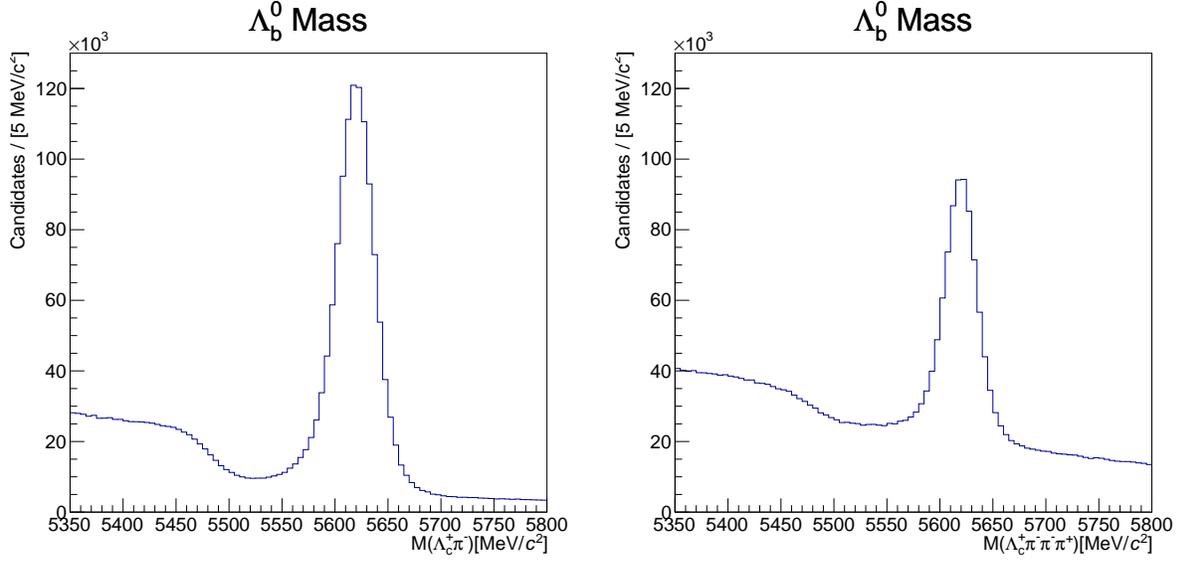


Figure 21: Distribution of the invariant mass (Left) $m(\Lambda_c^+ \pi^-)$ and (Right) $m(\Lambda_c^+ \pi^- \pi^+ \pi^-)$ with stripping and offline selections applied as described in the text for the 2016 - 2018 data set.

3.4.3 Additional offline selections for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$

Further offline selections are applied to the Λ_b^0 candidates to improve the signal-to-background ratio. They include:

- PID: The proton, kaon and all pions must have $\text{PROBN}Np > 0.2$, $\text{PROBN}Nk > 0.1$, and $\text{PROBN}N\pi > 0.01$ respectively;
- Λ_c^+ : $2261 < m(pK^- \pi^+) < 2311 \text{ MeV}/c^2$ and $-0.0005 < \tau_{\Lambda_c^+}$ (lifetime of Λ_c^+) $< 0.005 \text{ ns}$;
- Λ_b^0 : $\text{DIRA} > 0.9999^3$.
- For $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode, $m(\pi^- \pi^- \pi^+) < 2800 \text{ MeV}/c^2$.

The invariant mass spectra of Λ_b^0 candidates passing the stripping and previously mentioned selection requirements discussed above are shown in Fig. 21. Prominent peaks at the known Λ_b^0 mass of $5620 \text{ MeV}/c^2$ are evident, however, there is still significant combinatorial background underneath the peaks, most notably for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay. To further improve the signal-to-background, while retaining high signal efficiency, a multivariate discriminator is employed, as described in the next section.

³DIRA: This is short for directional-angle, and is equal to the cosine of the angle between the momentum vector of the particle and the direction vector from the pp collision point to the decay vertex of the particle.

3.4.4 Multivariate discriminant for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$

To further improve the signal-to-background ratio for the Λ_b^0 candidates, we employ a boosted decision tree discriminant (BDT) [40, 41] via the TMVA package [42]. This first BDT is referred to as BDT1, and is used to improve the S/B ratio in the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$ decays. The selection of variables is motivated by previous studies of these final states [36, 37].

For the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay, the list of variables used in BDT1 includes:

- Λ_b^0 : Vertex fit χ_{vtx}^2 , DIRA angle, flight distance from PV and radial flight distance;
- Λ_c^+ : Vertex fit χ_{vtx}^2 and decay time;
- a_1^- ($\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ only): 3π invariant mass, vertex fit χ_{vtx}^2 and χ_{VS}^2 (the χ^2 of the vertex separation from the primary vertex (PV));
- χ_{IP}^2 ⁴ of the p , K^- and π^+ from the Λ_c^+ and the one (or three) π^\pm from the Λ_b^0 ;
- Momentum (p) and transverse momentum (p_T) of the p , K^- and π^+ from the Λ_c^+ and the one (or three) π^\pm from the Λ_b^0 ;
- PID value (PROBNx) of the p , K^- and π^+ from the Λ_c^+ and the one (or three) π^\pm from the Λ_b^0 .

The Λ_b^0 signal is taken from simulated $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decays. To ensure that we use just a single Λ_b^0 decay per Ξ_b^- decay, and the correct one, we require the Ξ_b^- to be truth-matched, i.e. the decay is recognized as the real decay in simulation. The background is taken from $\Lambda_c^+ \pi^- \pi^+ \pi^-$ combinations in the high Λ_b^0 mass sideband, $5700 < m(\Lambda_c^+ \pi^- \pi^+ \pi^-) < 5850 \text{ MeV}/c^2$. The distributions of the input variables for the signal and background samples are shown in Figs. 57, 58 and 59 in Appendix E for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode. Large discrimination power is evident in several of the variables.

The $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ BDT uses the same variables as listed above, except those related to the a_1 , which is not present in the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decay. The list of input variables and the distributions can be found in Appendix F.

The BDT1 discriminant output is shown in Fig. 22 for both the training sample, and the independent test sample for (left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$. Excellent separation of the signal from the combinatorial background is achieved, as evidenced by the background events piling up at -1, and signal events at +1. Moreover, the distributions are similar for the training and test samples, indicating that the method has not lead to overtraining [43].

⁴the χ^2 of the impact parameter of the particle with respect to some vertex or 3D-point. In this case, it is related to the primary vertex (PV).

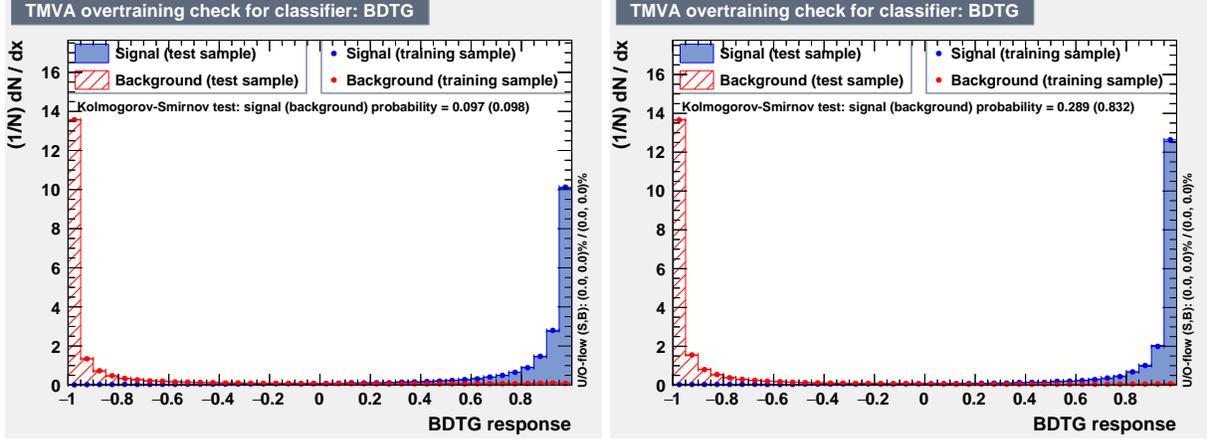


Figure 22: Distributions of the output BDT1 response for signal and background for the (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ samples. The test and training samples are both shown.

A formal optimization of the BDT1 requirement is performed as part of the Ξ_b^- selection. The suppression of the combinatorial background with a loose requirement of $\text{BDT1} > -0.9$ is demonstrated in Fig. 23 and Fig. 24 for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ modes, respectively. Each figure shows the mass spectrum before the BDT1 selection, and the distributions of rejected and accepted candidates. With this loose requirement, about 70% of the combinatorial background is removed with only a few percent signal loss (small bump in middle plots).

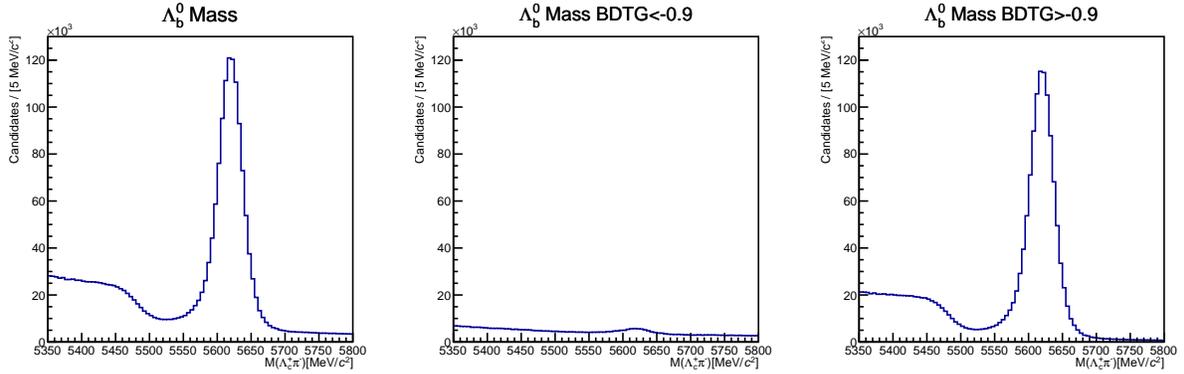


Figure 23: Invariant mass spectrum of (left) all $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ candidates that pass the rectangular selections, and (middle) candidates with $\text{BDT1} < -0.9$ and (right) $\text{BDT1} > -0.9$.

3.5 Ξ_b^- selection

The Ξ_b^- candidates are formed by pairing a Λ_b^0 candidate (that passes all selections and BDT1) with a π^- meson candidate. Only Λ_b^0 candidates that have mass in the range $5560 < m(\Lambda_c^+ \pi^- (\pi^+ \pi^-)) < 5680 \text{ MeV}/c^2$,

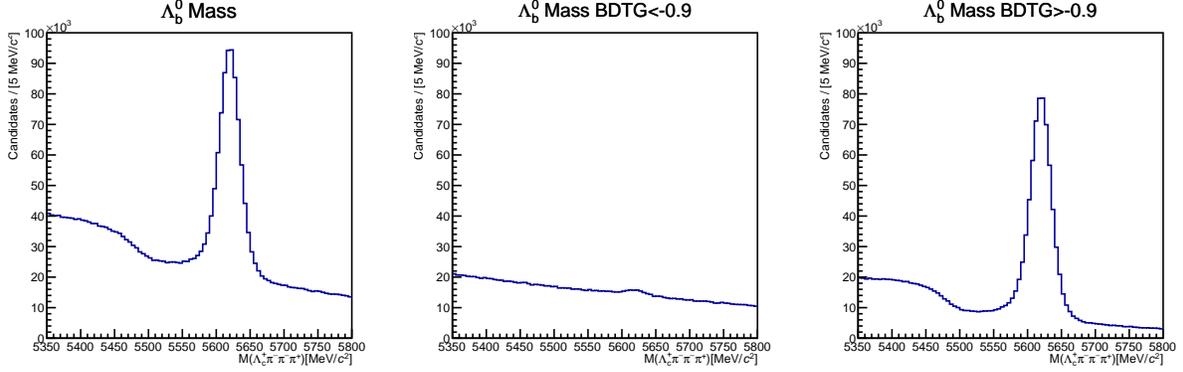


Figure 24: Invariant mass spectrum of (left) all $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ candidates that pass the rectangular selections, and (middle) candidates with $\text{BDT1} < -0.9$ and (right) $\text{BDT1} > -0.9$.

and lifetime in range $-2 < \tau_{\Lambda_b^0} < 12$ ps are used to form Ξ_b^- candidates. In addition Ξ_b candidates are required to point back to the PV by requiring $\Xi_b^- \text{ DIRA} > 0.99955$.

The π^- meson emerging in the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay (hereafter referred to as the bachelor π^-) tends to be low momentum, since the Q -value⁵ of the decay is relatively low and $M(\pi^-) \ll M(\Lambda_b^0)$. A relative low p_T requirement of 100 MeV/c is applied to the bachelor π^- . It is also required to have a ghost probability⁶ of less than 0.4 and a loose PID requirement of $\text{ProbNNpi} > 0.01$. No requirements are imposed on its χ_{IP}^2 value, since the low Q -value does not allow it to be clearly identified as a prompt or secondary particle. Instead, it's χ_{IP}^2 is used in a multivariate discriminant, as discussed below.

The same $\alpha > 0.8$ mrad requirement is imposed to remove the cases where the π^- overlaps with a track from the Λ_b^0 decay (see Sec. 3.4.1 and Appendix D for details). From simulated signal decays, the efficiency of this requirement is found to be 99.5%.

3.5.1 Multivariate discriminant for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$

After a suitable requirement on BDT1 is imposed, the major component of remaining background is from promptly produced Λ_b^0 baryons paired with a random π^- . A second BDT (BDT2) is used to suppress this background. Similar to the previous Run 1 study, we exploit the fact that Λ_b^0 baryons from $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ have larger decay times and flight distances when measured relative to the PV, as compared to those produced promptly. In addition, the π^- from the Ξ_b^- will tend to have higher transverse momentum than promptly-produced π^- from the underlying event of the pp collision. The variables used in the BDT2 training are:

⁵The Q value is the maximum momentum that a decay product can have in rest frame of the decaying particle.

⁶The probability of a track being fake [44].

- DIRA of Ξ_b^- ;
- Flight distance of Ξ_b^- from related PV;
- p_T of Ξ_b^- ;
- Decay time of Λ_b^0 from related PV;
- p and p_T of Λ_b^0 ;
- p and p_T of the bachelor π^- ;
- χ_{IP}^2 of the bachelor π^- ;
- PROBNN π of bachelor pion π^- ;
- difference in pseudorapidity: $d\eta = \eta_\pi - \eta_{\Lambda_b^0}$, where $\eta = -\ln[\tan(\frac{\theta}{2})]$;
- difference in azimuthal angle: $d\phi = |\phi_\pi - \phi_{\Lambda_b^0}|$.

The last two variables, the difference in pseudorapidity and azimuthal angle exploit the fact that the bachelor π^- should be close in phase space to the Λ_b^0 , if it comes from a Ξ_b^- decay. One might have concern that this sculpts the mass distribution and produces an artificial mass peak. Such sculpting is checked by also fitting the wrong-sign (WS) sample. No such artificial peaking is seen.

We use simulated true $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decays to emulate the signal and WS $\Lambda_b^0 \pi^+$ candidates to represent the background. There is no possibility of producing a peak at the Ξ_b^- mass for $\Lambda_b^0 \pi^+$ combinations, and thus it is a good proxy for the $\Lambda_b^0 \pi^-$ background⁷. For the training of BDT2, all previous selections are applied (except for BDT1). In the training, the mass difference, $\delta m = m(\Lambda_b^0 \pi) - m(\Lambda_b^0) - M(\pi)$, is required to be in the interval 28–48 MeV/ c^2 for both signal decays and the background candidates. This helps ensure that the wrong sign background candidates have similar kinematics to the right sign background underneath the Ξ_b^- signal peak.

A 2D optimization of BDT1 and BDT2 is performed and is discussed in Sec. 3.6. The distributions of the input variables to BDT2 for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with $\Lambda_b^0 \rightarrow \pi^-$ are shown in Fig. 25 as an example. Each variable gives some separation power, but exploiting the full set in a single output classifier leads to much better overall separation, as seen in Fig. 26.

To confirm that the WS candidates give a good representation of the RS background, we compare the RS and WS BDT2 distributions from the low δm region, $10 < \delta m < 32.89$ MeV/ c^2 , (see Fig. 27). The BDT

⁷ π^+ and π^- mesons from the underlying event are produced in almost equal numbers in high energy pp collisions with almost identical momentum distributions

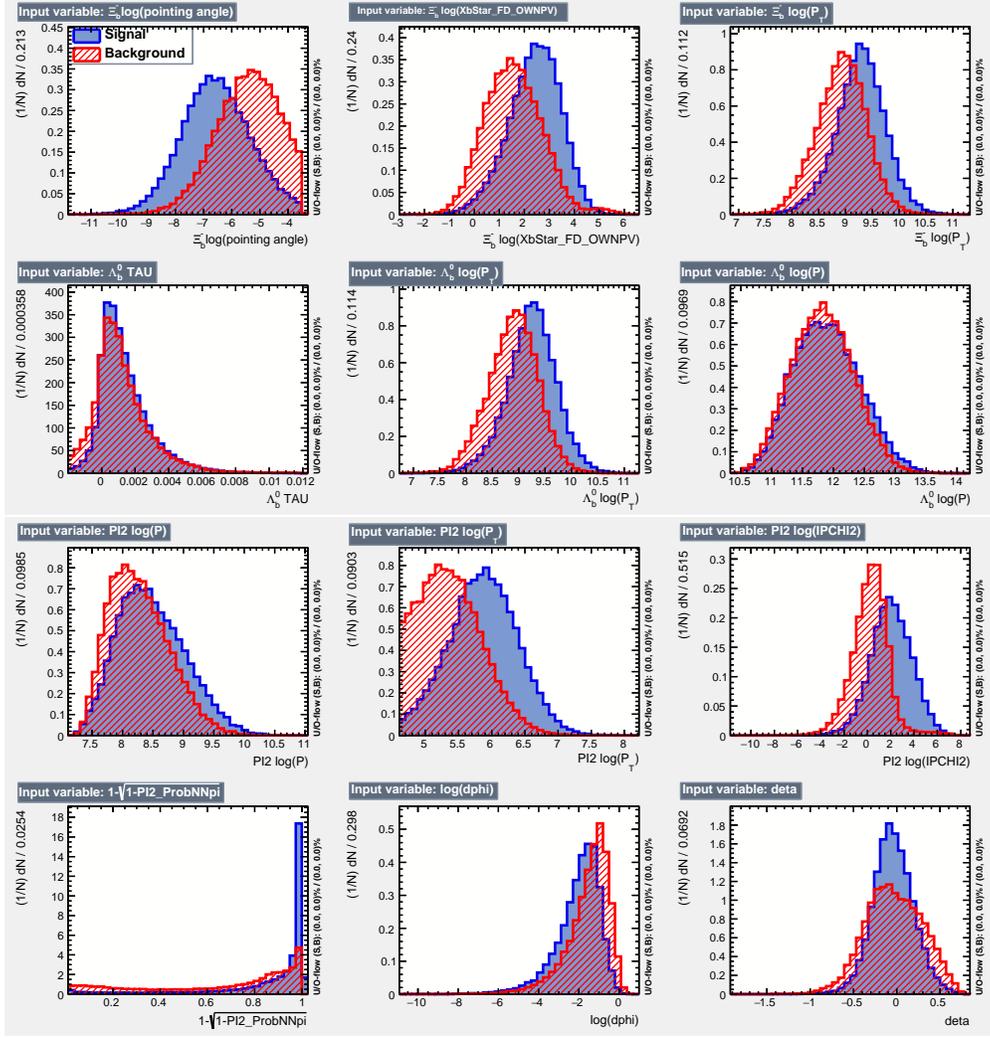


Figure 25: Distributions of the input variables used to train BDT2 for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$; signal and background distributions are superimposed.

response distributions of RS and WS are consistent with each other, and the ratio of $\frac{N_{RS}}{N_{WS}}$ is compatible with unity across the BDT spectrum. For the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay with $\Lambda_b^0 \rightarrow \pi^- \pi^+ \pi^-$, similar figures can be found in Appendix. G;

3.6 Requirements on BDT1 and BDT2

The optimization of the requirements on BDT1 and BDT2 require that the simulation properly models the data. The weighting of the simulation is discussed in Section 4.3, and those weights are used in the optimization discussed here. The optimization of the selection requirements on BDT1 and BDT2 is done by evaluating the expected signal significance, $FOM = S/\sqrt{S+B}$ for each point in a 2D scan of (BDT1,

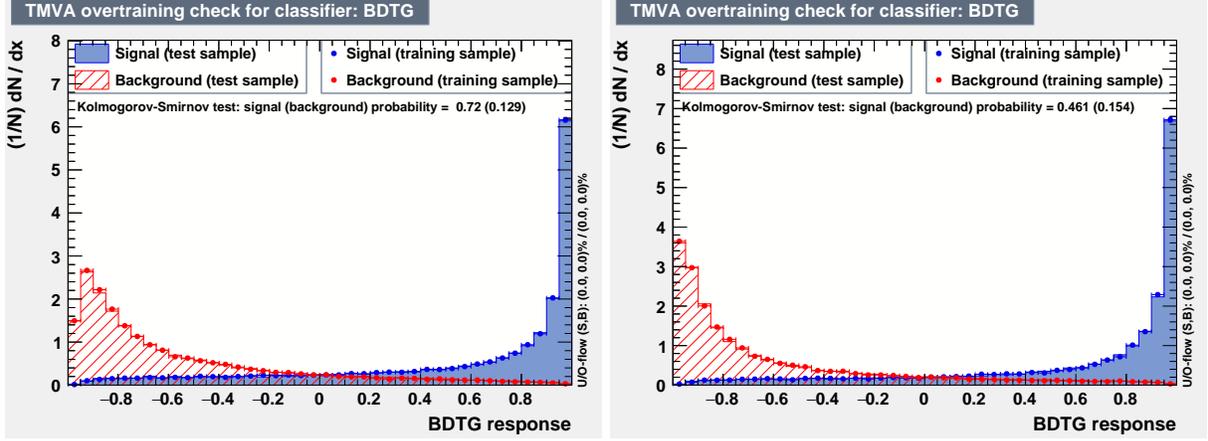


Figure 26: Distributions of the output BDT2 response for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ (Left) and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ (Right) signals and WS background. The test and training samples are both shown.

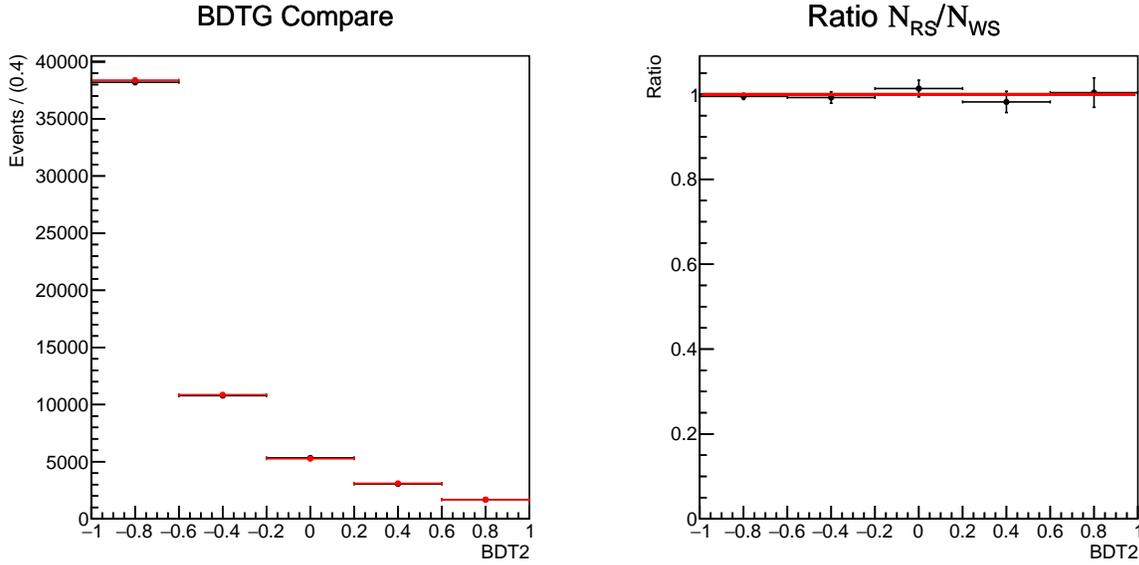


Figure 27: (Left) RS (black) and WS (red) BDT response distributions in $10 < \Delta m < 32.89 \text{ MeV}/c^2$ mass window. (Right) The ratio $\frac{N_{RS}}{N_{WS}}$ of each bin of the left plot. This is Run2 $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ data sample.

BDT2) values. The expected signal yield is given by $S_0 \epsilon_{\text{BDT}}$, where $S_0 = 200$ is the estimated signal yield with no BDT requirement based on the Run 1 result [22], and ϵ_{BDT} is the efficiency of the BDT1 and BDT2 requirements obtained from the weighted simulation. The background yield is determined from the wrong-sign $\Lambda_b^0 \pi^+$ candidate yield in the mass region $34.8 < \delta m < 41.6 \text{ MeV}/c^2$ (a $\pm 2.5\sigma$ window around the expected Ξ_b^- peak position, but in the WS sample). The procedure is carried out separately for both the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ samples. The resulting FOM is shown in Fig. 28.

In general, the FOM versus BDT1 is fairly flat over a wide range of BDT1 cuts. Thus the overall

performance is dominated by the cut on BDT2. With that caveat, the maximum value of FOM for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ is at $\text{BDT1} > 0.6$ and $\text{BDT2} > 0.96$, and with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ it is at $\text{BDT1} > -0.24$ and $\text{BDT2} > 0.92$. While these values are optimal, it assumes the efficiency from simulation gives an almost perfect prediction of what to expect in the data. Any mis-modeling of the signal decay will reduce the efficiency of the BDT selection requirement in data. Since the efficiency decreases sharply near the BDT2 endpoint of 1.0 (see Fig. 30), we relax the requirement on BDT2 and require $\text{BDT2} > 0.9$ for the signal significance determination. The FOM values at $\text{BDT2} > 0.9$ are statistically consistent with those of the optimal values, so little or no loss in significance is expected with this slightly looser requirement.

For the measurement of r_s , we use an even looser requirement on BDT2, namely $\text{BDT2} > 0.8$, so that we are less sensitive to systematic uncertainty in the BDT2 efficiency associated with mis-modeling of the signal input distributions.

Figure 31 and Fig. 32 show projections of the FOM and efficiency as a function of BDT1 cut value. As indicated previously, the FOM is quite flat over a wide range, and we choose to require $\text{BDT1} > 0$, which is exactly the midpoint between the minimum and maximum output values. The two sets of BDT requirements are summarized in Table 1 along with the efficiencies of those requirements on signal and background. With the loose BDT requirements, the efficiency on signal is about 50%. With the loose and tight BDT requirements, the background is suppressed by factors of about 100 and 250-300, respectively.

Table 1: Two sets of BDT1 and BDT2 requirements applied to MC and data samples. 100% BDT Efficiency and 100% Background Efficiency are set to be with requirements $\text{BDT1} > 0$ or $\text{BDT2} > -1$ applied.

	BDT1	BDT2	Signal Efficiency		Background Efficiency	
	(Λ_b^0)	(Ξ_b)	$\Lambda_c^+ \pi^-$	$\Lambda_c^+ \pi^- \pi^+ \pi^-$	$\Lambda_c^+ \pi^-$	$\Lambda_c^+ \pi^- \pi^+ \pi^-$
Loose BDT	0	0.8	0.467	0.565	0.010	0.008
Tight BDT	0	0.9	0.366	0.439	0.004	0.003

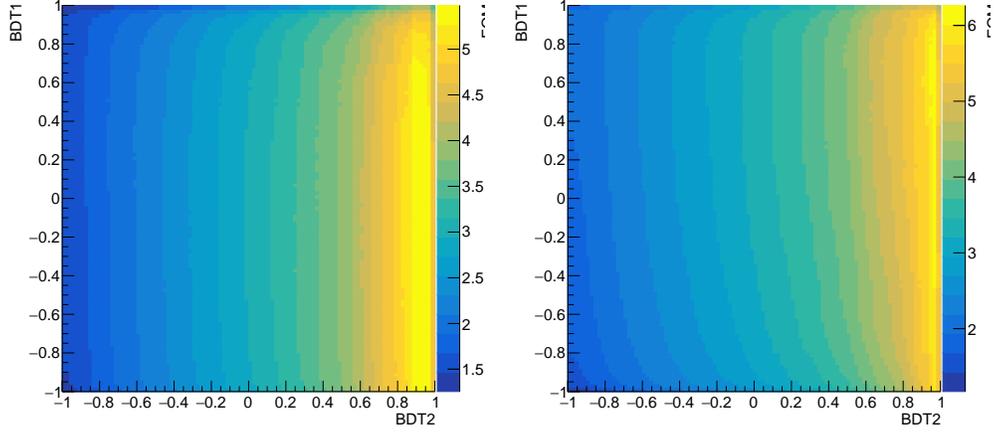


Figure 28: BDT1 and BDT2 2D scan FOM value for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$, and (Right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$

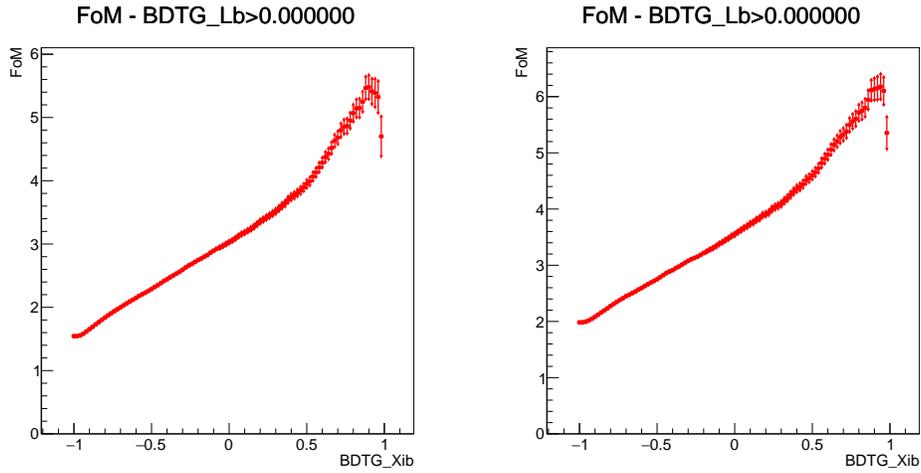


Figure 29: FOM vs. BDT2 when BDT1>0 for both data (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (Right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.

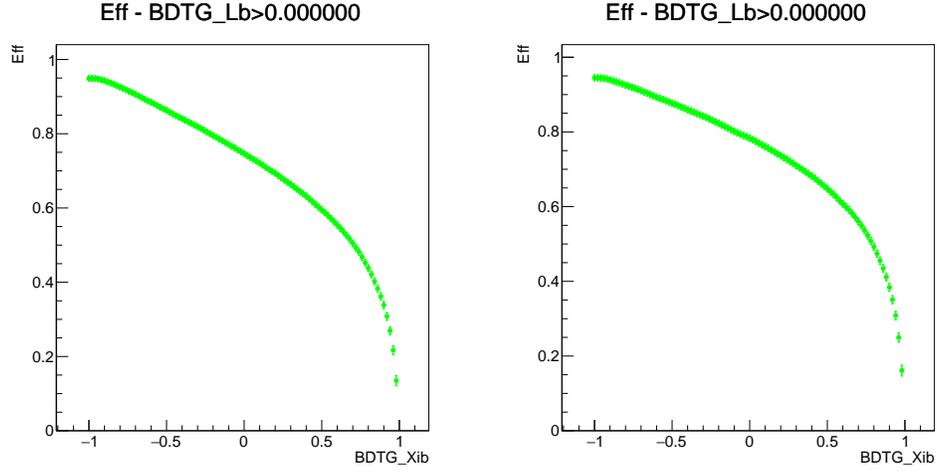


Figure 30: Efficiency vs. BDT2 with BDT1>0 for both simulated (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (Right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$. Note that the 100% efficiency is determined at BDT1 > -1 and BDT2 > -1, thus here, with cut BDT1 > 0, the efficiency does not start at 100% but around 95%. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.

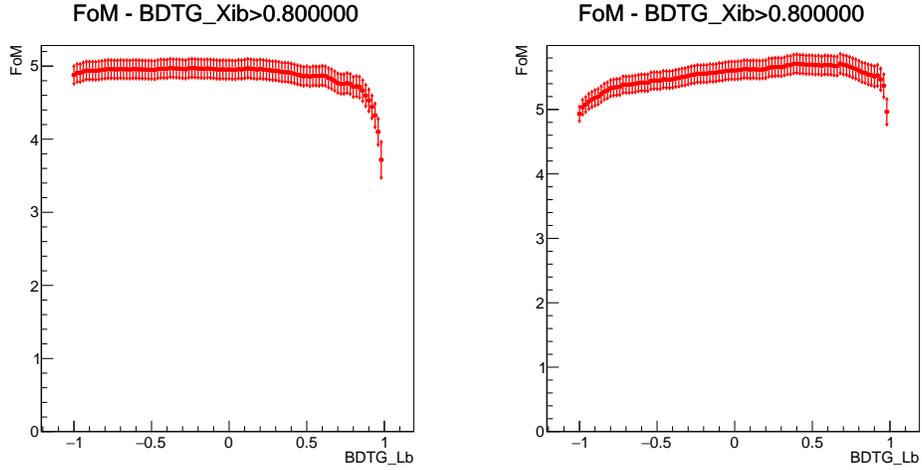


Figure 31: FoM vs. BDT1 with BDT2>0.8 for both (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (Right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.

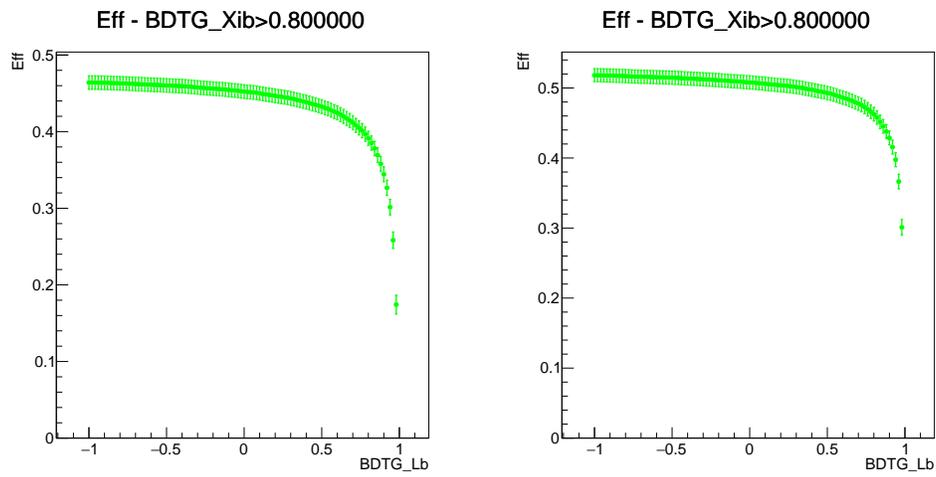


Figure 32: Efficiency vs. BDT1 when BDT2>0.8 for both data (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (Right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$. Note again that the 100% efficiency is determined at BDT1> -1 and BDT2> -1, thus here, with cut BDT2> 0.8, the efficiency does not start at 100% but around 50%. In the above BDTG_Lb refers to BDT1 and BDTG_Xib corresponds to BDT2.

4 Inputs from simulation

4.1 Mass resolution for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$

The shape, or equivalently the mass resolution function, of the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays are obtained from simulation, and are parameterized as the sum of two Crystal Ball functions⁸ as

$$\mathcal{P}_{sig} = f_1 \text{CB}_1(m_0, \sigma_1, \alpha_1, n_1) + (1 - f_1) \text{CB}_2(m_0, r\sigma_1, \alpha_2, n_2) \quad (5)$$

All parameters are freely varied in the fit to simulation. A fit to the simulated $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ true signal decays are shown in Fig. 33, and the shape parameters obtained are shown in Table 2.

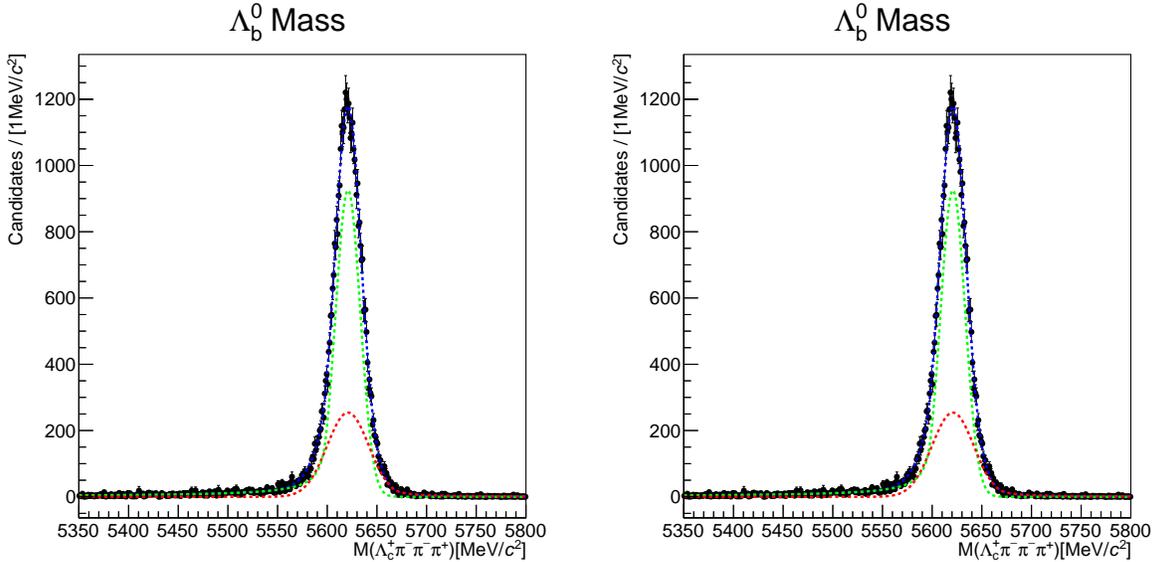


Figure 33: Signal shape from simulated (left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays. The results of the fit, described in the text, are superimposed.

4.2 Ξ_b^- Signal Mass Resolution Shape

The signal shape is determined from simulated $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decays with all of the selections discussed in previous sections applied. The shape of the resolution is described by the sum of two Crystal Ball functions (see Eq. 5). The result of a fit to the δm spectrum for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal decays is shown in Fig. 34.

The fit parameters of the resolution function are shown in Table 3. The average δm resolution is about $1.2 \text{ MeV}/c^2$, owing to the low Q -value in the decay and the excellent momentum resolution of the LHCb

⁸It contains a Gaussian core portion (σ as the width) and a power-law low-end tail (n as the power); m as the mean value and α affects where the Gaussian transform to the power-law tail [45, 46].

Table 2: Fitted signal shape parameters for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays, as obtained from fits to simulated signal decays.

Parameter	$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$	$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$
σ_1 (MeV/c ²)	13.13	12.34
α_1	1.68	1.83
n_1	1.31	1.02
f_1	0.55	0.70
α_2	-2.00	-2.14
n_2	3.22	1.26
r	1.50	1.70

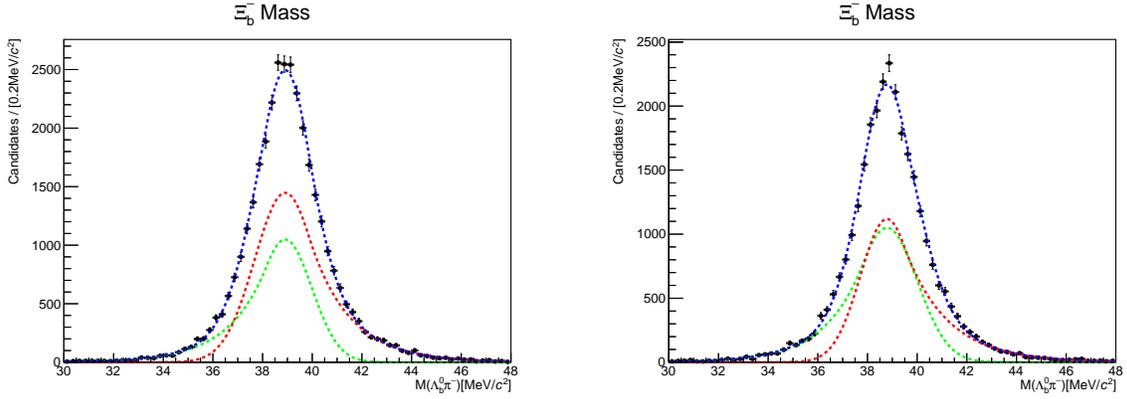


Figure 34: Spectrum of the mass difference, δm , for simulated $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decays with (left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$, along with the shapes obtained from the fit.

spectrometer.

Table 3: Fitted signal shape parameters for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ δm shape for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays, as obtained from simulated signal decays.

Parameter	$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$	$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$
σ_1 (MeV/c ²)	1.07	1.22
α_1	0.96	0.79
n_1	126.78	128.27
f_1	0.40	0.50
α_2	-0.77	-0.62
n_2	135.71	131.50
r	1.15	0.87

4.3 Weighting of the simulation

Before optimizing the selection on BDT1 and BDT2, it is important that the simulation is weighted to reproduce the kinematical spectra of signal decays in the data. Two sets of weights are applied to the kinematical variables in simulated decays to bring them into agreement with the data. They are:

- $\Lambda_c^+ \rightarrow pK^-\pi^+$ Dalitz weights;
- Production kinematics, p_T and η , of the Λ_b^0 and Ξ_b^- baryons

In general the efficiency depends on these kinematical variables, and therefore it is important that the efficiency match the data for any estimation of detection efficiency. These two weighting schemes are discussed below.

4.3.1 Λ_c^+ Dalitz weights

The weighting of the Λ_c^+ Dalitz structure was performed in Ref. [47] (Λ_c^+ , Ξ_c^+ , and Ξ_c^0 lifetime paper) using large samples of semileptonic $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu X$, $\Xi_b^0 \rightarrow \Xi_c^+ \mu^- X$ and $\Xi_b^- \rightarrow \Xi_c^0 \mu^- X$ decays. The weights obtained from that analysis are used here, since the simulation in this analysis uses the same cocktail of resonances for the Λ_c^+ decay that were used in Ref. [47].

4.3.2 Weighting of the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation

It is important that the (p_T, η) spectra in the simulation reproduces that of the data, since signal efficiencies are sensitive to the kinematics, particularly p_T . To perform this weighting, the 2D (p_T, η) distribution in data is extracted using signal weights (or sWeights, for short) obtained from an sPlot [48] using the Λ_b^0 invariant mass as the discriminating variable. A similar 2D distribution is obtained for the simulated decays, and normalized to the yield in the data. The ratio of the data to the simulated distribution for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decay is shown in Fig. 35. The projections onto the p_T and η axes are shown in Fig. 36. The same figures for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode are shown in Appendix. H. The weights are computed from the ratio of the distributions in data to simulation. After the weighting, the simulation matches the data well for both the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay modes.

4.3.3 $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ weighting

Due to the low expected signal yield, it is not possible to use the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal in data to weight the simulation. Instead, we use a $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ (with $\Xi_c^0 \rightarrow pK^- K^- \pi^+$) sample in data and a corresponding

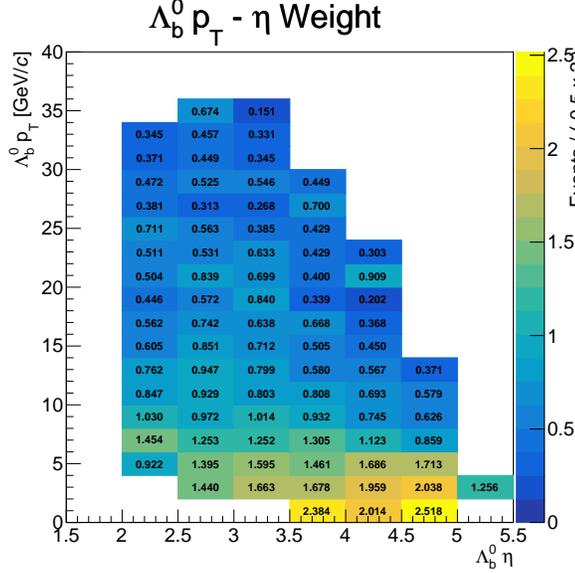


Figure 35: $\Lambda_b^0 p_T$ - η weight distribution, as obtained from the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ sWeighted Run2 data and simulated signal decays.

sample of simulated decays. Weights for the production kinematics could depend weakly on the decay mode (through the detector acceptance), but this control mode is a fully-hadronic mode, and has the same number of final state particles (five) in both cases as the $\Xi_b^- \rightarrow (\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-) \pi^-$ mode. For the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ we consider the modeling of the spectra as an additional systematic uncertainty.

The process of obtaining the (p_T, η) weight of Ξ_b^- follows the same strategy as what is done for the Λ_b^0 (p_T, η) weight. The ratio of the (p_T, η) spectrum in data to simulation is shown in Fig. 37.

Due to the low signal yields in the corners of the phase space of the 2D (p_T, η) distribution, for bins with less than 10 entries, we use 1D weights based on the ratio of p_T spectra in data to simulation, as shown in the right plot of Fig. 37. The ratio is fit to an exponential function, and this function is used instead for those low-statistics bins., For $p_T < 5 \text{ GeV}/c$, the weight is fixed to the value of the function at $p_T = 5 \text{ GeV}/c$.

The projections onto the p_T and η axes before and after application of the weights are shown in Fig. 38. The bottom pair of plots show that the weights bring the simulation into good agreement with the distributions obtained from the data.

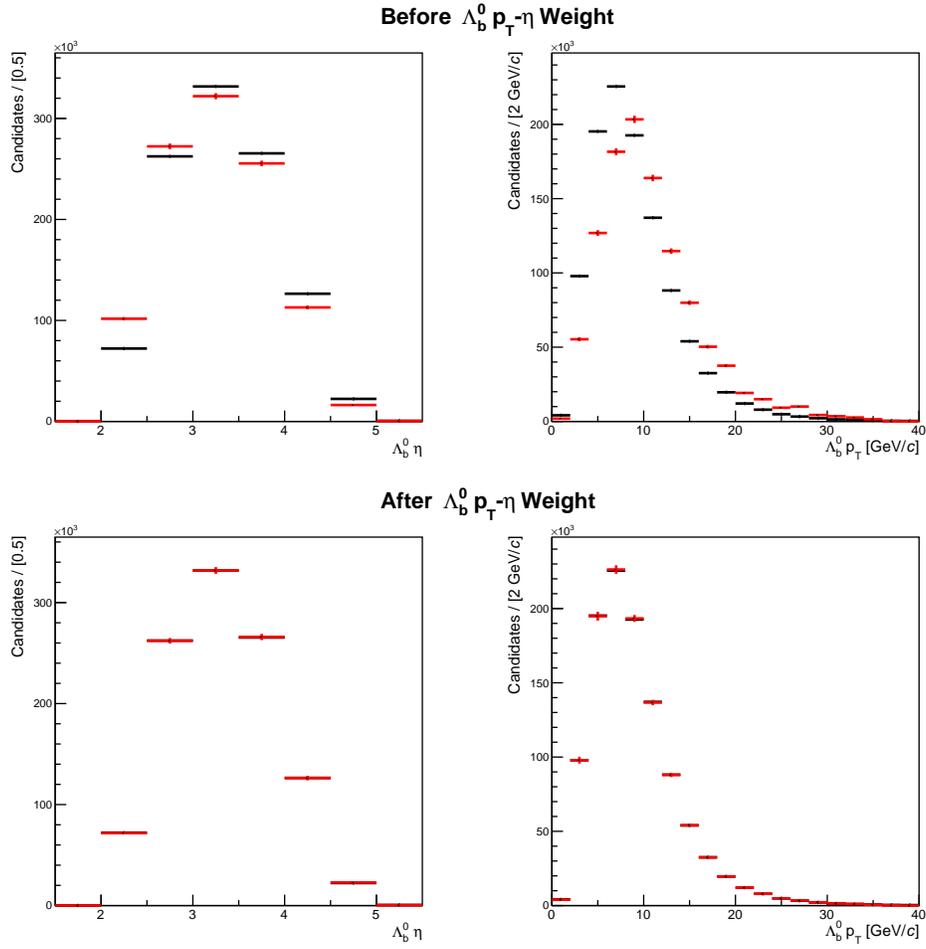


Figure 36: $\Lambda_b^0 p_T$ (Left) and η (Right) distributions for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ MC Run2 sample (top) before and (bottom) after the (p_T, η) weight from Fig. 35 is applied.

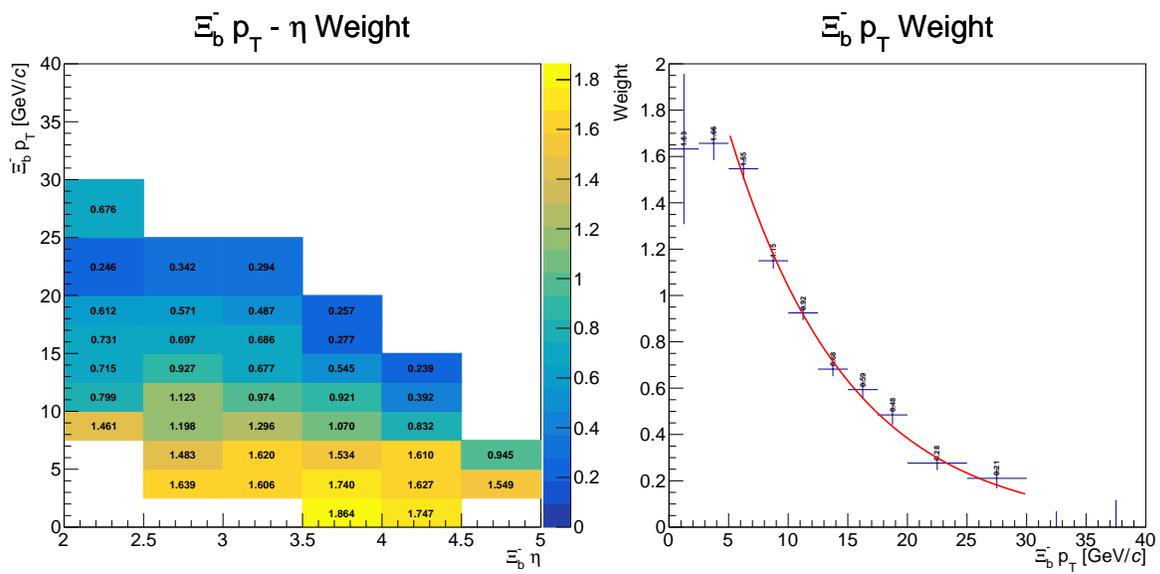


Figure 37: (Left) Ratio of the (p_T, η) spectrum in data relative to signal MC for $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ decays, to obtain the Ξ_b^- production spectrum weights. (Right) Ratio of p_T spectrum in data relative to signal MC for $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ decays. The red line is a exponential function fitted to bins from 5 GeV/c to 30 GeV/c

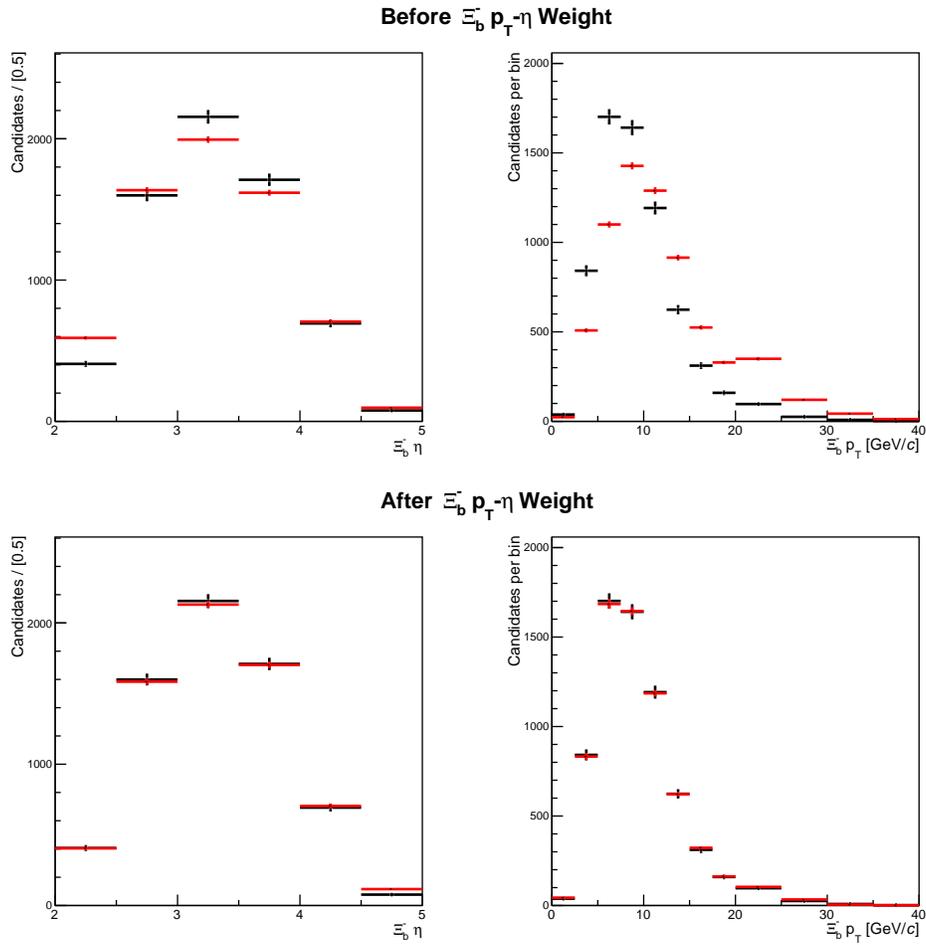


Figure 38: Projections of the $\Xi_b^- p_T$ and η spectra (top) before and (bottom) after the (p_T, η) weights are applied to $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ MC sample.

5 Relative efficiencies

The simulation is also needed to determine the relative efficiency between the signal and normalization modes. To use the measured value of $f_{\Xi_b^-}/f_{\Lambda_b^0}$ in Ref. [23], the Ξ_b^- and Λ_b^0 baryon yields and relative efficiencies must be determined with the same kinematic acceptance requirement of that measurement, which is $p_T < 20 \text{ GeV}/c$ and $2 < \eta < 6$.

The total efficiency is broken up into several factors,

$$\epsilon^{rel} = \epsilon_{acc}^{rel} \epsilon_{rec}^{rel} \epsilon_{sel}^{rel} \epsilon_{trig}^{rel}. \quad (6)$$

Here, ϵ_{acc}^{rel} is the relative acceptance between the Λ_b^0 normalization and Ξ_b^- signal modes, ϵ_{rec}^{rel} is the relative reconstruction efficiency, given the decays are in the acceptance, ϵ_{sel}^{rel} is the relative efficiency of all offline selections given that the decay was reconstructed (including BDT requirement), and ϵ_{trig}^{rel} is the relative trigger efficiency, given that the decay passed all previous selections. Thus, it is important to note that the efficiencies are measured relative to the previously imposed requirements.

5.1 Relative acceptance

The signal and normalization modes are generated with the following generator level requirements:

- $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$: all final state charged tracks to be within 10–400 mrad;
- $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$: all final state charged tracks to be within 10–400 mrad and have $p > 1.6 \text{ GeV}/c$. (Generally, tracks with $p < 1.6 \text{ GeV}/c$ will not make it through the LHCb dipole magnet, and therefore cannot be reconstructed as long tracks.)
- $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- (\pi^+ \pi^-)$: all charged tracks from the Λ_b^0 (not the π^- from the Ξ_b^-) must have $10 < \theta < 400 \text{ mrad}$, $p > 1.5 \text{ GeV}/c$, and $p_T > 90 \text{ MeV}/c$.

On the last item, the stripping line requires $p_T > 100 \text{ MeV}/c$ in all cases, so the decays which have daughter tracks with $p_T < 90 \text{ MeV}/c$ would not be selected, so it would be a waste of CPU resources to generate and simulate them. To make the definitions of the acceptance uniform across all MC samples, we require $p > 1.6 \text{ GeV}/c$ and $p_T > 90 \text{ MeV}/c$ on all final state tracks, and this factor is included in ϵ_{acc} .

Productions of simulated events in LHCb provide the fraction of events that are selected by the generator level cuts relative to the full 4π solid angle, here referred to as $\epsilon_{4\pi}$. However, what is needed is the efficiency of the generator level cuts relative to the kinematic region, $p_T < 20 \text{ GeV}/c$ and $2 < \eta < 6$, for the parent b

Table 4: Relative acceptance between the signal and normalization modes, as described in the text.

$\Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$		$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$	
MC sample	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$	$\Lambda_b^0 \rightarrow f$	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$	$\Lambda_b^0 \rightarrow f$
$\epsilon_{4\pi}$	0.143	0.182	0.0989	0.113
ϵ_{kin}	0.957	0.962	0.958	0.960
f_{kin}	0.248	0.247	0.248	0.247
ϵ_{p,p_T}	0.987	0.853	0.968	0.939
ϵ_{acc}	0.545	0.603	0.370	0.410
ϵ_{acc}^{rel}	1.11 ± 0.02		1.11 ± 0.02	

hadron. To get the proper acceptance, ϵ_{acc} is determined as

$$\epsilon_{acc} = \epsilon_{4\pi} \left(\epsilon_{kin} \epsilon_{p,p_T} / f_{kin} \right). \quad (7)$$

The factor f_{kin} is the fraction of signal decays in which the parent b hadron is in the kinematic region $2 < \eta < 6$ and $p_T < 20 \text{ GeV}/c$, and is obtained from a small EVTGENONLY [49, 50] simulation. The term ϵ_{kin} is determined from the the MC generated events, and applies the kinematic requirement $2 < \eta < 6$ and $p_T < 20 \text{ GeV}/c$, which is not applied in the generation. The factor ϵ_{p,p_T} is the efficiency of the $p > 1.6 \text{ GeV}/c$ and $p_T > 90 \text{ MeV}/c$ requirement on all final state tracks, except for the π^- daughter in the Ξ_b^- decay.

A summary of the factors in Eq. 7 and the resulting acceptance is shown in Table 4. With a uniform definition of the acceptance between the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ modes, we see that the relative acceptance is 1.11 ± 0.02 in both cases. One would expect them to be approximately equal, which is consistent with what is found. The uncertainty is dominated by the uncertainty in f_{kin} .

5.2 Relative Reconstruction efficiency

The reconstruction efficiency is computed as the fraction $\epsilon_{rec} = N_{rec}/N_{gen}$ using simulated decays, where N_{gen} is the number of generated candidates and N_{rec} is number of reconstructed candidates. The weights, discussed in the previous section, for production kinematics and the Λ_c^+ Dalitz are applied to both N_{rec} and N_{gen} . In addition, the Ξ_b^- and Λ_b^0 candidates are required to be within the kinematic region $p_T < 20 \text{ GeV}/c$ and $2 < \eta < 6$, and pass the daughter momentum cuts discussed in Section 5.1. Along with these requirements, some additional selections are applied for consistency with later the calculation of offline selection efficiencies:

- b -hadron parent is truth-matched;
- for normalization mode, the Λ_b^0 invariant mass is in the range (5350, 5800) MeV/c^2

Table 5: Relative reconstruction efficiency of the signal and normalization modes. For the final value, the uncertainty is due to finite statistics in the simulated samples.

$\Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$			$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$		
	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$	$\Lambda_b^0 \rightarrow f$	Relative	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$	$\Lambda_b^0 \rightarrow f$	Relative
ϵ_{rec}^{2016}	0.0282	0.0514	1.83	0.0121	0.0133	1.10
ϵ_{rec}^{2017}	0.0312	0.0534	1.71	0.0124	0.0136	1.10
ϵ_{rec}^{2018}	0.0254	0.0438	1.73	0.0110	0.0112	1.02
ϵ_{rec}^{Run2}	0.0280	0.0491	1.75	0.0118	0.0126	1.07
ϵ_{rec}^{rel}	1.75 ± 0.01			1.07 ± 0.01		

- for the signal mode, $|m_{\Xi_b} - m_{\Lambda_b^0} - M_{\pi^+}|$ in range (5, 130) MeV/ c^2 ;

For both signal mode and normalization mode, the ϵ_{rec} is determined using simulation from each year, and the resulting values obtained are listed in Table 5. The average Run2 ϵ_{rec}^{Run2} is the luminosity-weighted average of the three years

$$\epsilon_{rec}^{Run2} = \frac{\mathcal{L}_{int}^{2016} \epsilon_{rec}^{2016} + \mathcal{L}_{int}^{2017} \epsilon_{rec}^{2017} + \mathcal{L}_{int}^{2018} \epsilon_{rec}^{2018}}{\mathcal{L}_{int}^{Run2}} \quad (8)$$

where $\mathcal{L}_{int}^{2016} = 1.7 \text{ fb}^{-1}$, $\mathcal{L}_{int}^{2017} = 1.7 \text{ fb}^{-1}$, and $\mathcal{L}_{int}^{2018} = 2.2 \text{ fb}^{-1}$, The relative reconstruction efficiency, defined as $\epsilon_{rec}^{rel} = \epsilon_{\Lambda_b^0} / \epsilon_{\Xi_b}$ is given in the bottom row for each Λ_b^0 decay mode.

There are two main competing effects that moderate the efficiency ratio. First, there is the efficiency to reconstruct the π^- from the Ξ_b^- decay in the signal mode, which increases ϵ_{rec}^{rel} . Second, in the Ξ_b^- signal mode, the Λ_b^0 daughters have significantly larger IP due to the combined flight distances of the Ξ_b^- and Λ_b^0 , compared to only the Λ_b^0 for the normalization mode. This results in the signal mode having larger efficiency to pass the impact parameter related requirements, *e.g.* $\chi_{IP}^2 > 4$, which lowers ϵ_{rec}^{rel} . For the second effect, the reduction of ϵ_{rec}^{rel} is much more significant for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ because there are six tracks whose efficiencies are each increased (in the Ξ_b^- signal decay), compared to only four tracks for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode. This is the reason ϵ_{rec}^{rel} differs significantly for the two Λ_b^0 modes.

In summary, $\epsilon_{rec}^{rel} = 1.75 \pm 0.01$ for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode and $\epsilon_{rec}^{rel} = 1.07 \pm 0.01$ for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode.

5.3 Relative offline selection efficiency

The offline efficiency represents the efficiency of all selections applied after the Ξ_b^- reconstruction phase, including the BDT requirements. Table 6 shows the cumulative offline selection efficiencies for each year and the breakdown of efficiencies can be found in Appendix I. The luminosity-weighted average is shown at the

Table 6: Relative offline selection efficiency of the signal and normalization modes. For the final value, the uncertainty is due to finite statistics in the simulated samples.

$\Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$			$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$		
	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$	$\Lambda_b^0 \rightarrow f$	Relative	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$	$\Lambda_b^0 \rightarrow f$	Relative
ϵ_{sel}^{2016}	0.312	0.738	2.36	0.338	0.681	2.02
ϵ_{sel}^{2017}	0.305	0.741	2.24	0.336	0.712	2.12
ϵ_{sel}^{2018}	0.318	0.746	2.35	0.363	0.707	1.94
ϵ_{sel}^{Run2}	0.312	0.741	2.38	0.347	0.701	2.02
ϵ_{sel}^{rel}	2.38 ± 0.01			2.02 ± 0.01		

bottom. Here, the difference in the relative efficiency is not as large for the two Λ_b^0 modes, but there is still a difference, mainly coming from the efficiency of the BDT2 requirement.

The relative offline efficiencies are $\epsilon_{sel}^{rel} = 2.38 \pm 0.01$ for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode and $\epsilon_{sel}^{rel} = 2.02 \pm 0.01$ for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode.

5.4 Relative trigger efficiency

By default, we do not apply any specific trigger selection requirements, other than the event passes the L0, HLT1 and HLT2 triggers. This is possible because we do not expect the π^- from the Ξ_b^- to contribute in any significant way to the trigger efficiency since it is a soft particle with a small χ_{IP}^2 . This expectation is validated with the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ simulation, which shows that there is negligible difference between requiring a trigger-on-signal (TOS) or trigger independent-of-signal (TIS) [51] on the Λ_b^0 or on the Ξ_b^- . Also, the trigger threshold does not impact the final result, and this conclusion will be explained in the systematic uncertainty section.

Thus, the main systematic uncertainty comes from the modeling of the kinematics of the parent b hadron.

5.5 Correction of relative efficiency

5.5.1 BDT2 requirement

The BDT2 requirement has an efficiency of about 50% for $BDT > 0.8$ according to simulation, and by a BDT2 value of 0.9, the distribution is rapidly approaching zero as BDT2 approaches 1.0 (see Fig. 32). To probe whether the BDT2 distribution in data and simulation are compatible, the signal yield ratio, r_{BDT2} for $BDT > 0.8$ to $BDT > 0.9$ is compared. From the yields in Table 11, the ratios are $r_{BDT2}^{1\pi} = 1.37 \pm 0.16$ and $r_{BDT2}^{3\pi} = 1.47 \pm 0.16$, where the expected value is 1.30 for both Λ_b^0 decay modes. Although both ratios in data are consistent within one standard deviation of the expectation from simulation, they are both shifted in a

way that suggests the simulation is more strongly peaked at unity than the data. This would be expected since precise modeling of every aspect of these decays is difficult.

A correction is therefore derived by smearing the BDT2 distribution from simulation with a one-sided Gaussian distribution in order to reproduce the r_{BDT2} values obtained in data, as shown in Fig. 39.

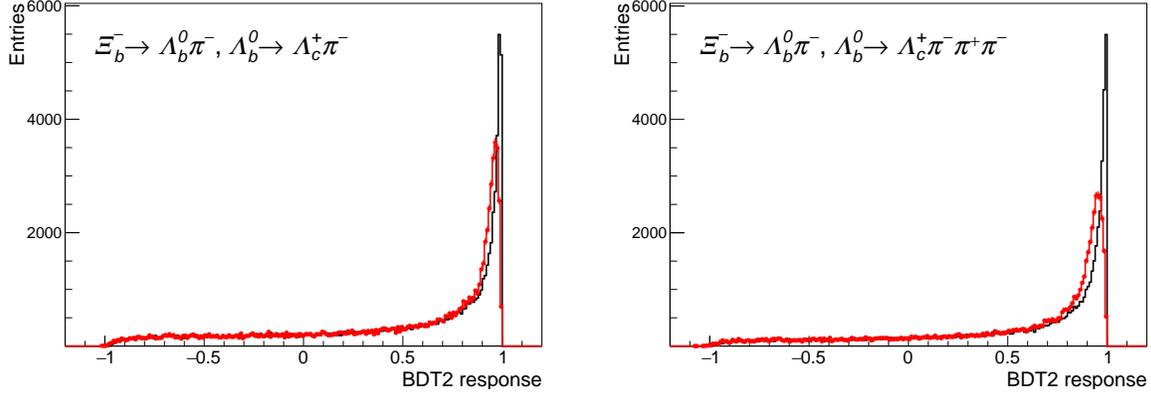


Figure 39: Distributions of BDT2 for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with (Left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (Right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$, before (black) and after (red) the smearing is applied.

The efficiencies for the $\text{BDT2} > 0.8$ before and after the smearing is computed, and the ratio is taken as a correction. We assign half of the difference from unity as the uncertainty in the correction. For the $\text{BDT2} > 0.8$ requirement, this leads to corrections to the signal modes' offline selection efficiency of 0.96 ± 0.02 and 0.94 ± 0.03 for the $\Lambda_c^+ \pi^-$ and $\Lambda_c^+ \pi^- \pi^+ \pi^-$ modes, respectively. Therefore, for the relative efficiency, the corrections are 1.04 ± 0.02 and 1.06 ± 0.03 respectively. For the $\text{BDT2} > 0.9$ selection, the analogous corrections are significantly larger, 0.91 ± 0.05 and 0.82 ± 0.09 , since the efficiency is more steeply decreasing as BDT2 approaches 1.0. The larger correction that is required for the tighter BDT2 selection is the reasoning behind taking the looser BDT2 requirement for measuring r_s . These correction factors are applied to the relative efficiencies.

To gain confidence that these corrections are reasonable, we compare the BDT1 distributions in data and simulation, where there is ample signal yields in data to do a quantitative comparison. The BDT1 spectrum is obtained in data using the *sPlot* technique. All selection requirements are applied, except the BDT1 requirement is relaxed to $\text{BDT1} > -0.9$. The comparison of the BDT1 distributions in data and simulation is shown in Fig. 40. For the nominal BDT1 requirement, $\text{BDT1} > 0$, the difference between data and simulation is at the 1-2% level. Since the BDT1 requirement is applied to both the signal and normalization mode, any data/simulation difference is negligible when computing r_s . On the other hand, BDT2 uses a requirement where the efficiency is not 95%, but rather about 50%. From the BDT1 spectrum, at the 50% efficiency

point, there is a data/simulation difference of about 5% (BDT1>0.88) for the $\Lambda_c^+ \pi^-$ mode and about 10% (BDT1>0.92) for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode.

However, the BDT1 spectrum is significantly more sharply peaked at unity than BDT2, which is quite evident when comparing Figs. 32 and 30, or by noting that the 50% efficiency point for BDT2 is at a value of about 0.8, compared to about 0.9 for BDT1. Thus the BDT2 efficiency should be less sensitive to data/simulation differences, and the corrections of 4% and 6% are reasonable. As these corrections cannot be precisely determined, we assign half of the deviation of the correction from unity as the uncertainty.

Table 7: Summary of the BDT2 corrections to relative efficiencies.

$\Lambda_b^0 \rightarrow f$	$\Lambda_c^+ \pi^-$	$\Lambda_c^+ \pi^- \pi^+ \pi^-$
BDT2 efficiency correction	1.04 ± 0.02	1.06 ± 0.03

5.5.2 Pion Tracking Efficiency

The tracking efficiency is calibrated using large samples of $J/\psi \rightarrow \mu^+ \mu^-$ decays, and provides a data-to-simulation correction for tracks in the momentum range from 5–200 GeV/c and $1.9 < \eta < 4.9$ [52]. The corrections are mostly within 1–2% of unity. About 65% of the π^- mesons from the Ξ_b^- decay have momentum below 5 GeV/c, and for these cases, the correction at 5 GeV/c is used with an uncertainty that is inflated by a factor of two. The luminosity-weighted average correction of the relative efficiency is 1.01 ± 0.03 .

5.5.3 True signal matching in simulation

The total efficiency is obtained by $\epsilon = \frac{N^{sel}}{N^{gen}}$ where N^{sel} is the selected candidates and N^{gen} is the generated candidates. In the efficiency calculations, however, only the true signal candidates are included in the N^{sel} . The ghost candidates should also be considered since they are indistinguishable in the data samples and are included in the data yield. Therefore, a correction should be applied to N^{sel} and thus the total efficiency.

In the simulation samples, the yields of the true signal and the ghost can be obtained by fitting the selected candidates of each with a fixed signal shape while the signal width is allowed to inflate. Table 8 summarized the corrections of total relative efficiencies in each mode and details can be found in Appendix J.

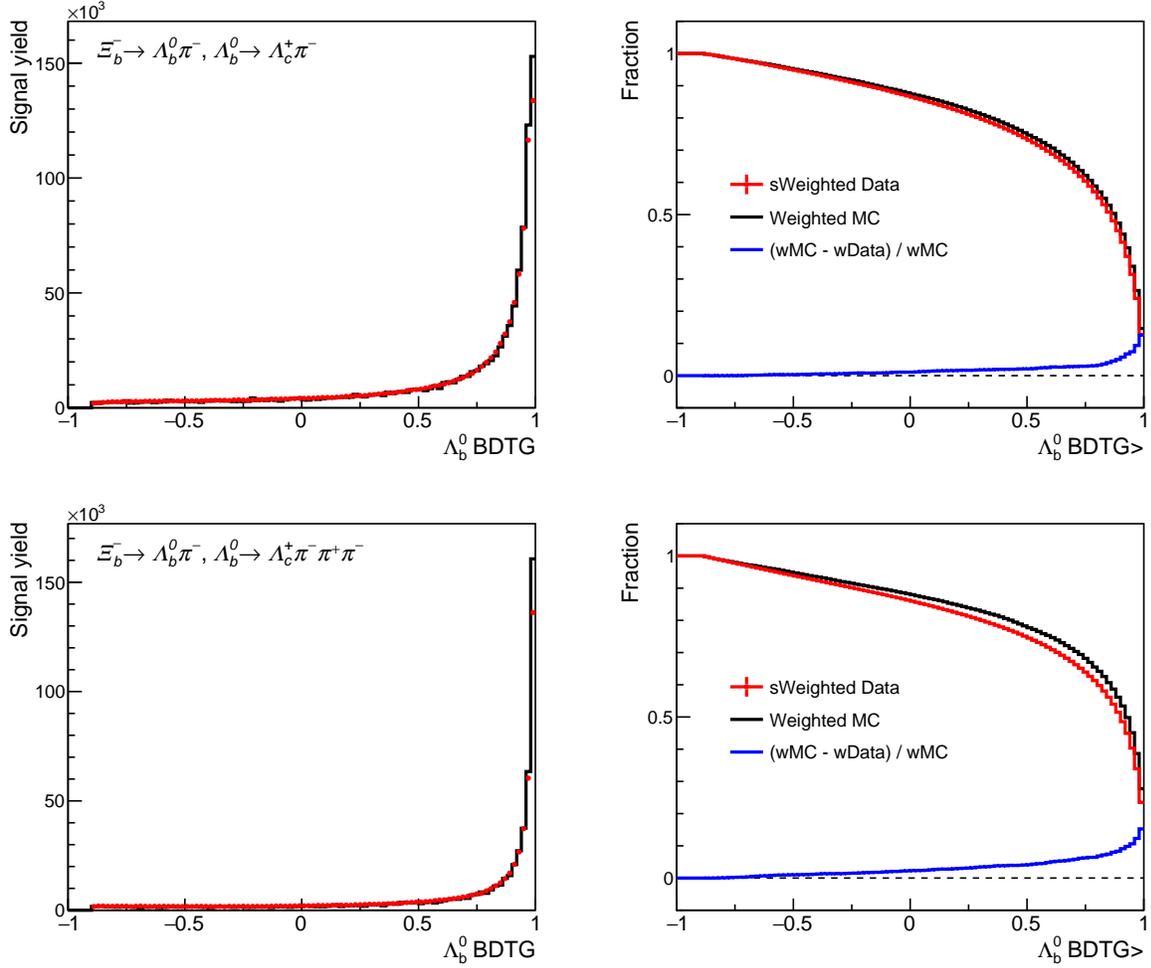


Figure 40: (Left) Distributions of the BDT1 output for (black) simulation and (red) sWeighted data. (Right) Fraction of events that pass a given BDT1 requirement, along with the fractional difference between data and simulation. The top row is for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode, and the bottom for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode.

Table 8: Summary of the signal matching corrections to relative efficiencies.

$\Lambda_b^0 \rightarrow f$	$\Lambda_c^+ \pi^-$	$\Lambda_c^+ \pi^- \pi^+ \pi^-$
correction	0.989	0.986

5.6 Summary of relative efficiencies

A summary of the relative efficiencies is given in Table 9. The bottom row shows the combined relative efficiency needed for the measurement of r_s . The lower value of ϵ^{rel} in the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode indicates that one is about twice as efficient in reconstructing and selecting the Ξ_b^- decay signals when the Λ_b^0 is detected in the $\Lambda_c^+ \pi^- \pi^+ \pi^-$ final state compared to the $\Lambda_c^+ \pi^-$ mode.

Table 9: Summary of the relative efficiencies used to determine r_s . Uncertainties are due to finite statistics in the simulated samples.

$\Lambda_b^0 \rightarrow f$	$\Lambda_c^+ \pi^-$	$\Lambda_c^+ \pi^- \pi^+ \pi^-$
ϵ_{acc}^{rel}	1.11 ± 0.02	1.11 ± 0.02
ϵ_{rec}^{rel}	1.75 ± 0.01	1.07 ± 0.01
ϵ_{sel}^{rel}	2.38 ± 0.01	2.02 ± 0.01
correction	1.041 ± 0.039	1.060 ± 0.048
ϵ^{rel}	4.82 ± 0.09	2.53 ± 0.05

6 Fits to data

To measure r_s , the yields in the signal and normalization modes need to be determined. Fits are performed to the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- (\pi^+ \pi^-)$ normalization modes mass spectra. These fits are described below.

6.1 Fit to the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ samples

A binned extended maximum likelihood fit is performed to the Λ_b^0 candidate invariant mass spectrum using signal and background shapes obtained from simulated signal decays, as described in Section 4.1. An overall scale factor is applied to the Crystal Ball widths to allow for for a slightly larger mass resolution in data compared to simulation. A distribution function formed by two Crystal Ball functions is used to describe the misidentified $\Lambda_b^0 \rightarrow \Lambda_c^+ K^-$ or $\Lambda_b^0 \rightarrow \Lambda_c^+ K^- \pi^+ \pi^-$ component. This component is also selected in the Ξ_b reconstruction therefore the total Λ_b^0 signal yield is the combination of the two components inside the (5560, 5680) MeV/ c^2 mass window. An ARGUS function [53, 54] convolved with a Gaussian mass resolution function describes the partially reconstructed $\Lambda_b^0 \rightarrow \Lambda_c^+ X$ decays with a single missing pion. The shapes for the latter two components are taken from Ref. [22]. An exponential distribution is used to describe the combinatorial background shape with shape parameter that is allowed to freely vary in the fit. The result of the fit to the Run 2 data sample is shown in Fig. 41 for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ with all selection requirements imposed. The mass distributions for each year are shown in Fig. 42.

The yields of $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays selected by the mass window the (5560, 5680) MeV/ c^2 and the corresponding width scale factors are summarized in Table 10. Yields of the misidentified $\Lambda_b^0 \rightarrow \Lambda_c^+ K^-$ or $\Lambda_b^0 \rightarrow \Lambda_c^+ K^- \pi^+ \pi^-$ component are included. About 879,000 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and 483,000 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decays are reconstructed.

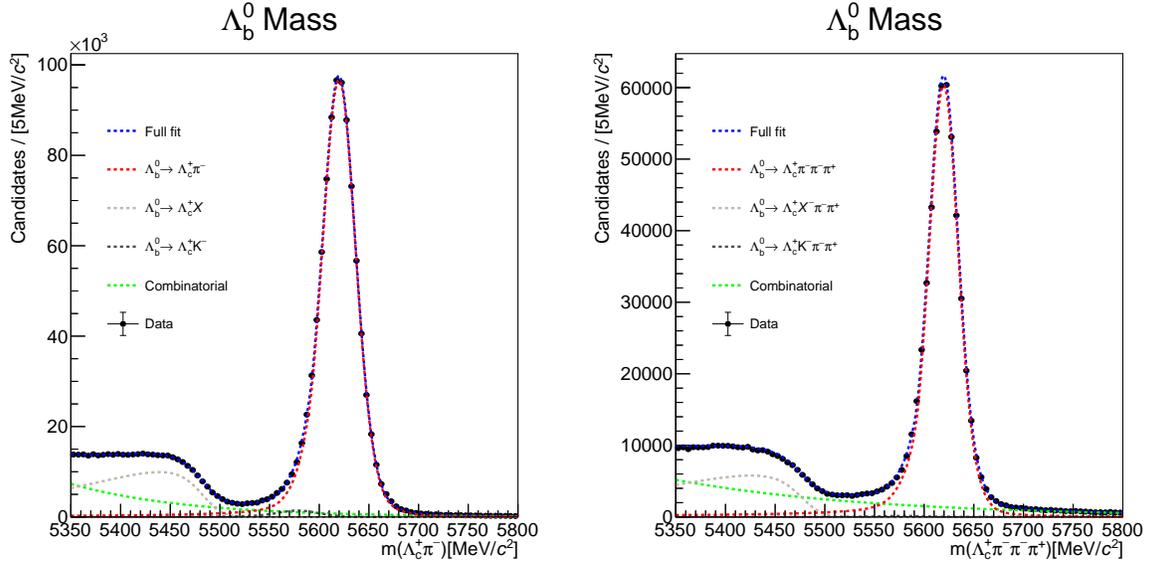


Figure 41: Invariant mass spectrum of (left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the full Run 2 data sample (2016-2018). The fit, as described in the text, is overlaid, with the red dashed line showing the Λ_b^0 signal, green is the combinatorial background, gray shows the partially reconstructed $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- (\pi^+ \pi^-)$ and black shows the misidentified $\Lambda_b^0 \rightarrow \Lambda_c^+ K^- (\pi^+ \pi^-)$.

Table 10: Fitted signal yields inside the (5560, 5680) MeV/c² mass window and the width scale factor for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ normalization modes for each year and the combined data. The yields of $\Lambda_b^0 \rightarrow \Lambda_c^+ K^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ K^-$ with a misidentified K^- are included.

Data set	$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$		$\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$	
	Yield	Scale Factor	Yield	Scale Factor
2016	275701 ± 566	1.15	147122 ± 442	1.14
2017	281217 ± 574	1.12	156236 ± 435	1.10
2018	322113 ± 614	1.12	179363 ± 489	1.10
Run2	879024 ± 1011	1.13	482718 ± 795	1.11
Total	879024 ± 1011		482718 ± 795	

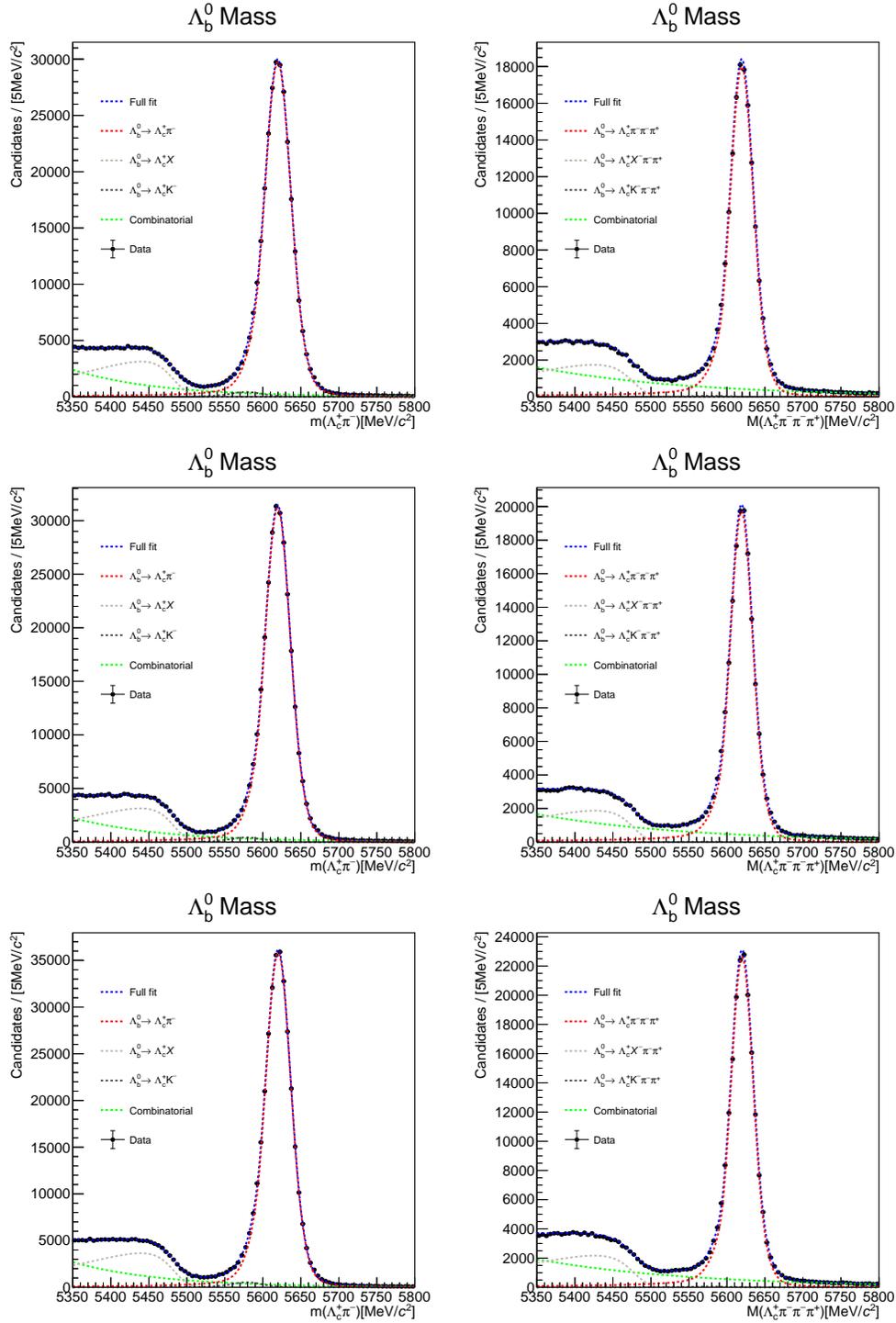


Figure 42: Invariant mass spectrum of (left) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and (right) $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the (top) 2016, (middle) 2017 and (bottom) 2018 data samples. The fit, as described in the text, is overlaid, with the red dashed line showing the Λ_b^0 signal, green is the combinatorial background, gray shows the partially reconstructed $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- (\pi^+ \pi^-) i$ decays and black shows the misidentified $\Lambda_b^0 \rightarrow \Lambda_c^+ K^- (\pi^- \pi^+)$.

6.2 Fit to the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ mass spectra

An unbinned extended maximum likelihood fit to the δm spectra is performed to search for the decay $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$. The $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ samples are fit simultaneously. Two separate fits are performed. In the first fit, we use a BDT2 requirement that provides the optimal FOM in order to establish observation of the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay mode (BDT2 > 0.9). In the second fit, we apply the looser BDT2 requirement (BDT2 > 0.8) for the branching fraction (r_s) measurement. The looser requirement provides a higher selection efficiency, and is therefore less sensitive to the efficiency of the BDT2 selection requirement, which is obtained from simulated decays.

In each fit, and in each subsample, two δm distributions are fit simultaneously, the right-sign (RS) candidates, and the wrong-sign (WS) candidates. The fit to the WS allows us to demonstrate that there are no artificial peaks generated at the expected value of about 38.2 MeV. The main components in the $\Lambda_b^0 \pi$ mass spectra are Ξ_b^- signal, the strongly-decaying $\Sigma_b^{(*)-} \rightarrow \Lambda_b^0 \pi^-$ (in the RS) and $\Sigma_b^{(*)+} \rightarrow \Lambda_b^0 \pi^+$ (in the WS) resonances, and combinatorial background. The RS and WS fits do not share any parameters, since the background shapes could differ. The $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ samples are fitted simultaneously and the peak positions of the Ξ_b^- , $\Sigma_b^{(*)-}$ (RS) and $\Sigma_b^{(*)+}$ (WS) masses are shared between the two Λ_b^0 modes. Below we discuss the shapes used to fit the δm spectra,

6.2.1 $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ Signal shape

The signal shape is taken from the simulation (see Section 4.2). The mean value is allowed to float in the range of $\pm 5 \text{ MeV}/c^2$ around the expected value of 38.2 MeV/ c^2 . The range $\pm 5 \text{ MeV}/c^2$ is about the $\pm 4\sigma$ of the δm resolution in the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay mode, which has a larger mass resolution than the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode. The mean mass value of Ξ_b^- peak is shared by the two modes in the simultaneous fit but unique $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal shapes are used.

6.2.2 Σ_b^- and Σ_b^{*-} resonance model

Following the previous published Run 1 result [22], the Σ_b^- and Σ_b^{*-} resonances are modeled with a relativistic Breit-Wigner (RBW) [55] function. This RBW function is described as:

$$\begin{aligned}
f(m; m_r, \Gamma_r) &= \frac{m_r \Gamma(m)}{(m^2 - m_r^2)^2 + m_r^2 \Gamma^2(m)} \\
\Gamma(m) &= \Gamma_r \left(\frac{m_r}{m}\right) \left(\frac{q}{q_r}\right)^{2L+1} F_r^2 \\
F_r^2 &\equiv \frac{1 + R^2 q_r^2}{1 + R^2 q^2} \text{ (for } L = 1\text{)} \\
m &\equiv m_0 + \delta m \\
q &\equiv \frac{mc}{2} \left(1 - \frac{(m_{\text{exp}}(\Lambda_b^0) + m_{\text{PDG}}(\pi^-))^2}{m^2}\right)^{1/2} \left(1 - \frac{(m_{\text{exp}}(\Lambda_b^0) - m_{\text{PDG}}(\pi^-))^2}{m^2}\right)^{1/2} \\
q_r &\equiv \frac{mc}{2} \left(1 - \frac{(m_{\text{exp}}(\Lambda_b^0) + m_{\text{PDG}}(\pi^-))^2}{m_r^2}\right)^{1/2} \left(1 - \frac{(m_{\text{exp}}(\Lambda_b^0) - m_{\text{PDG}}(\pi^-))^2}{m_r^2}\right)^{1/2} \\
R &= 3.0 (\text{GeV}/c)^{-1} \\
m_0 &\equiv m_{\text{exp}}(\Lambda_b^0) + m_{\text{PDG}}(\pi^-)
\end{aligned}$$

where $m_{\text{exp}}(\Lambda_b^0) = 5619.60 \pm 0.17 \text{ MeV}/c^2$ [33].

The Σ_b and Σ_b^* resonances' mass and width were recently measured precisely by LHCb [56], however, we cannot use the peak mass values from that analysis here, and allow them to vary freely instead. This is because in this analysis, the $\Lambda_b^0 \pi$ candidates are reconstructed assuming there could be a non-zero lifetime, and the Ξ_b^- selections favor $\Lambda_b^0 \pi$ candidates that have a reconstructed vertex downstream of the PV. This biases the mass of promptly decaying resonances, such as the Σ_b and Σ_b^* , toward larger values. This bias is demonstrated in Fig. 43, where the δm distribution for $\Lambda_b^0 \pi^-$ candidates with $\text{DIRA} > 0$ and $\text{DIRA} < 0$ are compared. Positive (negative) DIRA indicates the $\Lambda_b^0 \pi^-$ vertex is downstream (upstream) of the PV. The vertical line shows the nominal value, and one can see that positive DIRA values result in a peak to the right, and negative to the left of the vertical lines. Since the analysis requires $\text{DIRA} > 0.99955$, the peak positions are biased upward, and consequently we let their peak values vary freely in the fit.

In the simultaneous fit, Σ_b^- and Σ_b^{*-} in the RS spectra, and Σ_b^+ and Σ_b^{*+} peaks in the WS spectrum have peak masses and natural widths fixed to the PDG [33], except for an overall shift in the peak value for each, as discussed above.

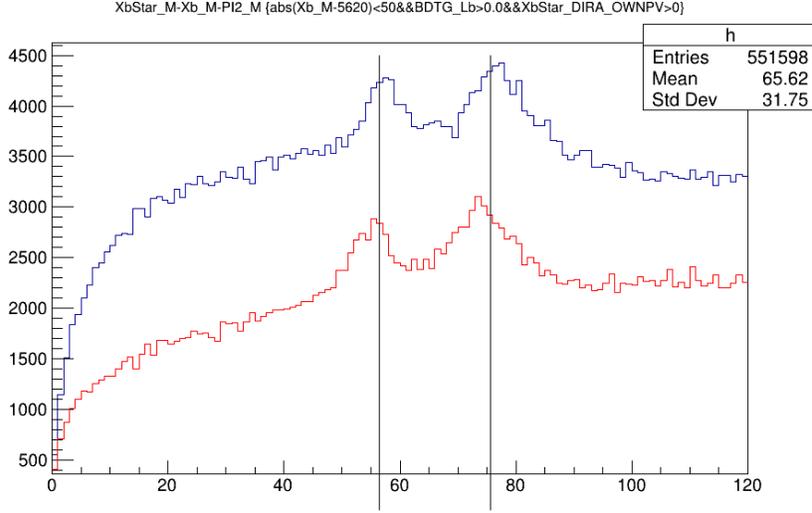


Figure 43: Distributions of δm for $\Lambda_b^0 \pi^-$ candidates with $\text{DIRA} > 0$ (blue) and $\text{DIRA} < 0$ (red).

6.2.3 Combinatorial background

The combinatorial background shape uses the same functional form as in Ref. [57]

$$f_{back}(\delta m) \propto (\delta m)^A (1 - e^{-\delta m/C}), \quad (9)$$

where the parameters A is fixed to 0.1 and C is freely varied in the fit to data. In the simultaneous fit, separate combinatorial background shape parameters for RS and WS spectra are used. For each decay mode, background parameters are allowed to vary independently.

6.2.4 $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ mass fit

The signal and background shapes described above are fitted to the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ δm spectra. The invariant mass spectra and the results of the fit are shown in Fig. 44 (using $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$) and Fig. 45 (using $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$) for the tight BDT2 selection ($\text{BDT2} > 0.9$). Clear and highly significant peaks are seen at about $38.2 \text{ MeV}/c^2$. The significance of the pair of peaks is 11.3σ , as determined from the $\sqrt{2(\Delta \log \mathcal{L})}$ [58]. This establishes the first observation of this strangeness-changing weak decay of the Ξ_b^- baryon.

For the measurement of r_s , we repeat the fit with the looser BDT2 selection requirement ($\text{BDT2} > 0.8$). The resulting fits are shown in Figs. 46 and 47. A summary of the fitted signal yields are shown in Table 11. Fitted parameters are presented in Appendix L.

Table 11: Fitted signal yields for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal decay, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the tight BDT2 selection used for observation of the signal, and the loose selection used for the r_s measurement.

Fit parameter	Tight BDT2	Loose BDT2
$N(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-)$	78.8 ± 12.8	108.6 ± 18.4
$N(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-)$	99.23 ± 14.7	146.0 ± 20.6
Mean value of δm of Ξ_b^- peak (MeV/ c^2)	38.47 ± 0.20	38.45 ± 0.19
$N(\Sigma_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-)$	107.8 ± 23.0	351.6 ± 38.2
$N(\Sigma_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-)$	158.6 ± 25.3	370.5 ± 40.15
$N(\Sigma_b^{*-} \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-)$	224.9 ± 33.6	684.4 ± 54.9
$N(\Sigma_b^{*-} \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-)$	303.4 ± 35.1	760.8 ± 56.8

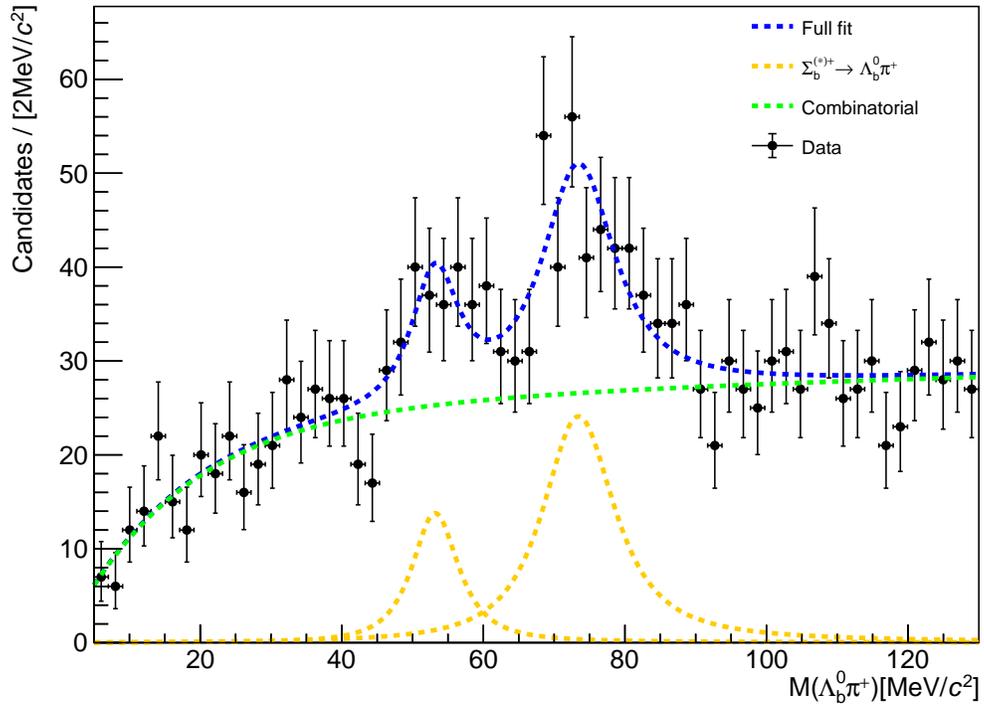
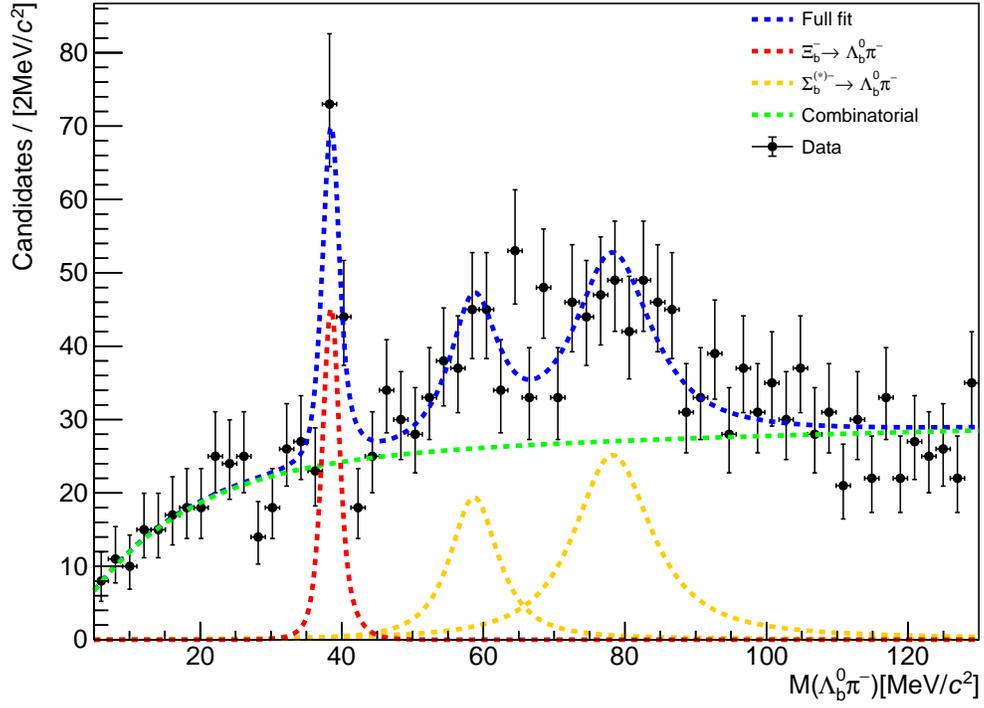


Figure 44: Results of the fit to the δm spectrum for the (top) $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and (bottom) Wrong sign $\Lambda_b^0 \pi^+$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the tight BDT2 selection.

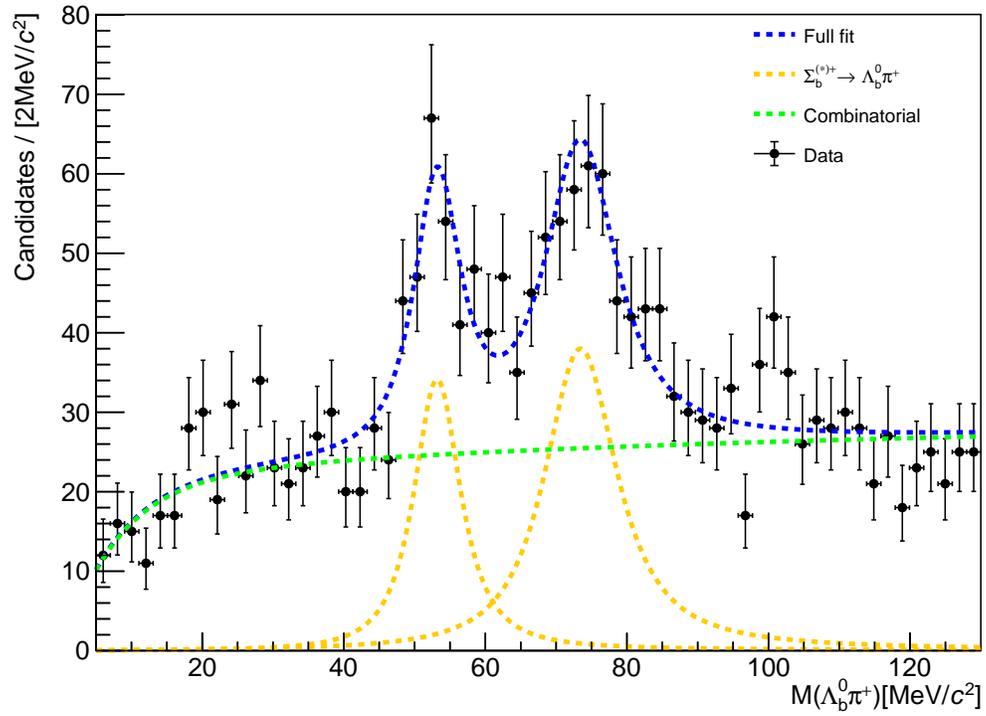
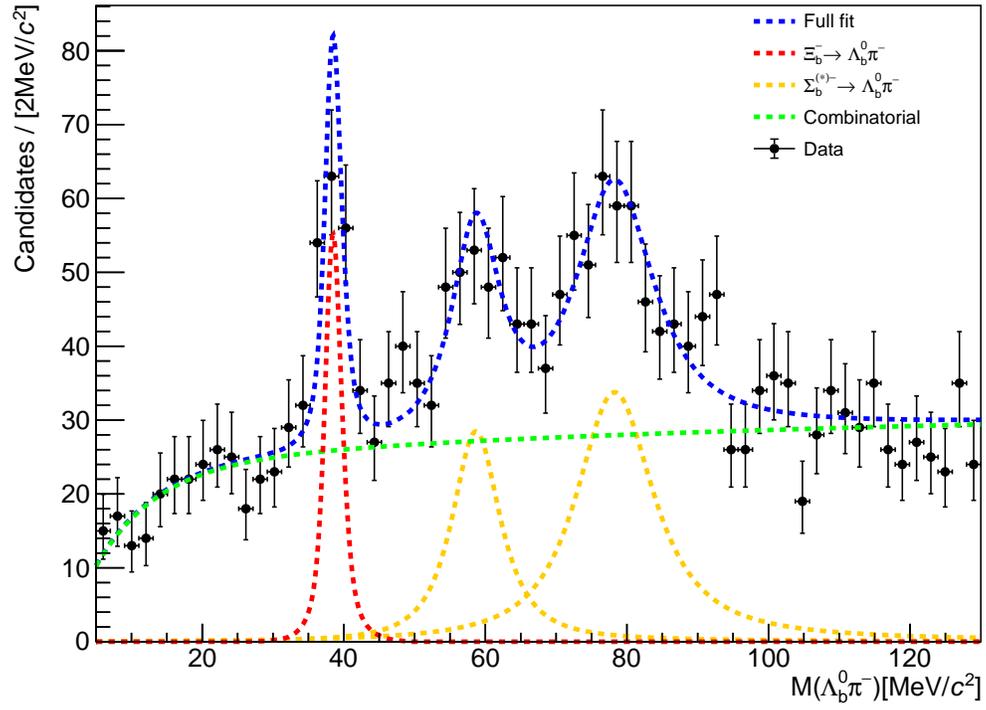


Figure 45: Results of the fit to the δm spectrum for the (top) $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and (bottom) Wrong sign $\Lambda_b^0 \pi^+$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the tight BDT2 selection.

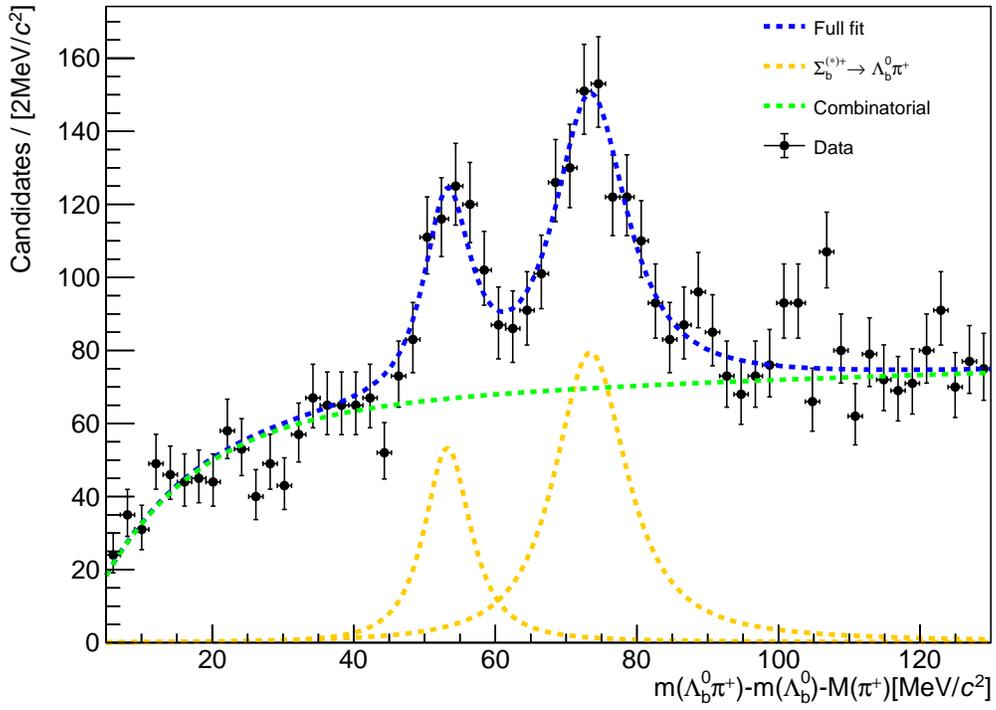
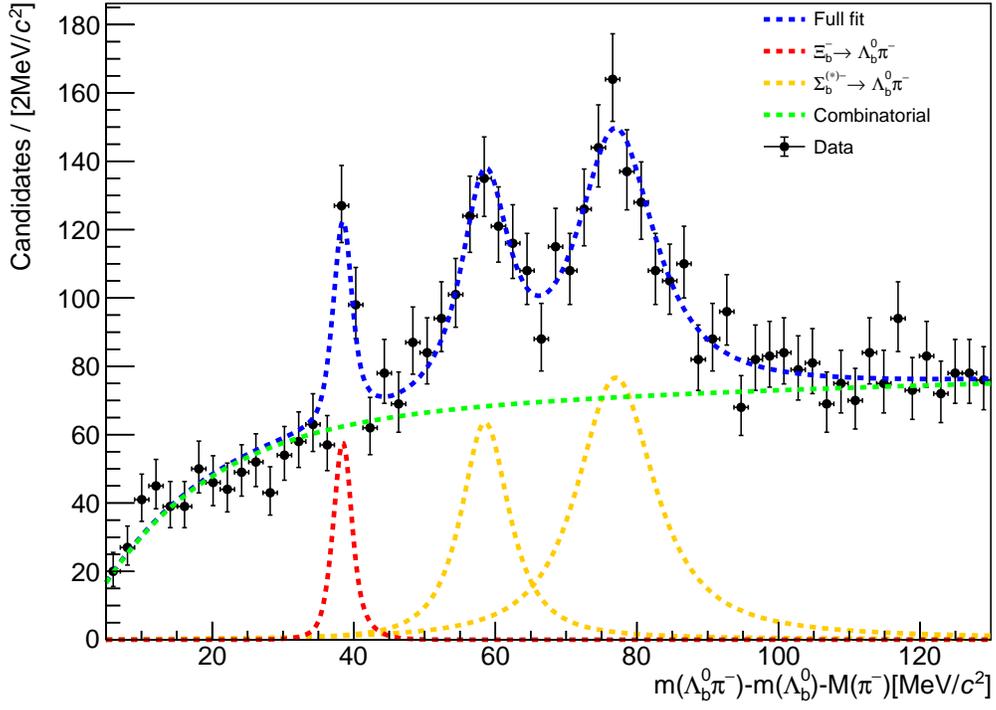


Figure 46: Results of the fit to the δm spectrum for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the loose BDT2 selection for the r_s measurement.

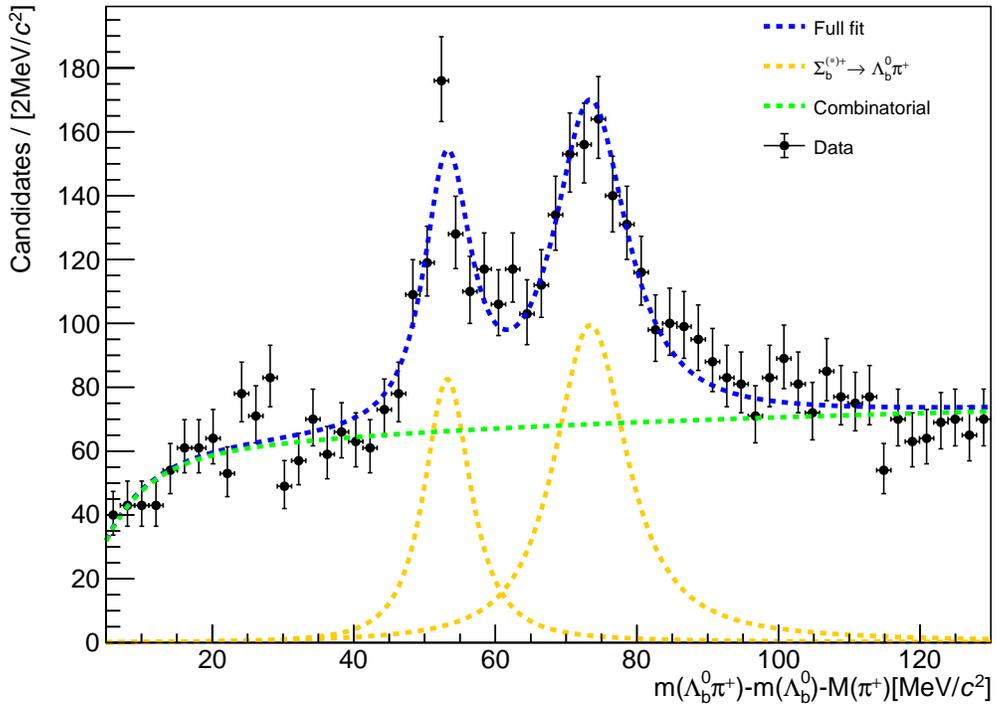
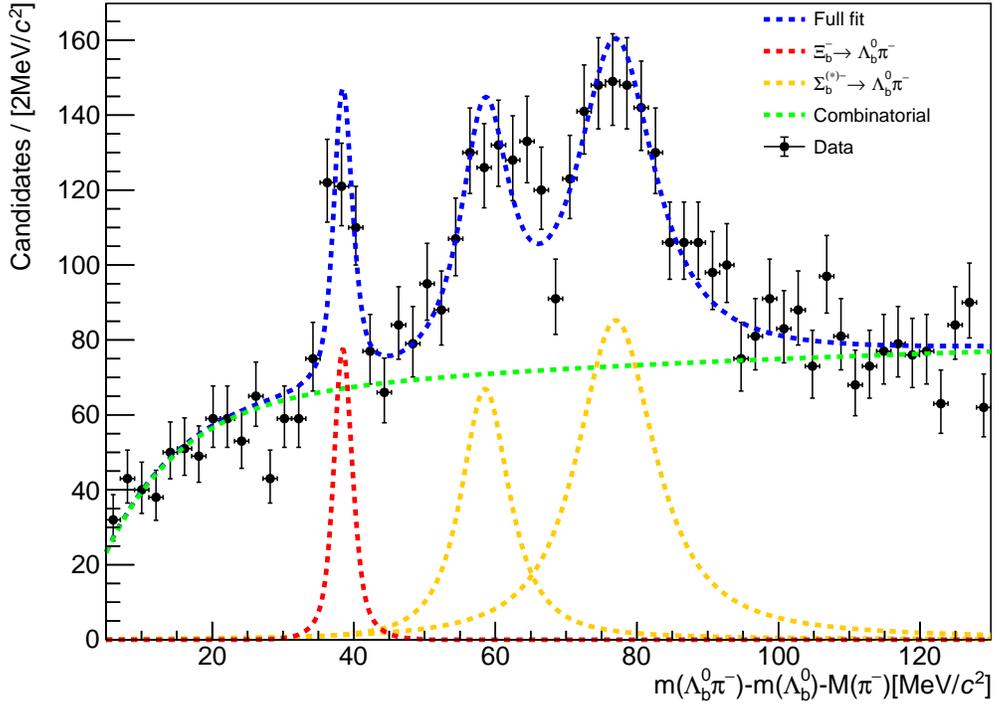


Figure 47: Results of the fit to the δm spectrum for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ for the full Run 2 data sample (5.5 fb^{-1}) with the loose BDT2 selection for the r_s measurement.

7 Systematic Uncertainties

7.1 Yield determination

7.1.1 Yield of Λ_b^0

The fitting of Λ_b^0 mass distribution in normalization mode is a binning fit with fixed parameters. The default bin width is $5 \text{ MeV}/c^2$ and here we changed it to $2 \text{ MeV}/c^2$ to see if the bin width affects the final yield or not. We also allowed the signal shape to vary freely (not fixed to the simulation result).

The result is summarized in Table. 12. The bin width does not affect the final yield of Λ_b^0 signal and the floating parameters of the signal shape only contribute 0.8% systematic uncertainty.

Table 12: Λ_b^0 yield of the default fitting (bin width $5 \text{ MeV}/c^2$ with fixed signal shape), the $2 \text{ MeV}/c^2$ bin width fitting with fixed signal shape, and the unfixed signal shape fitting with $5 \text{ MeV}/c^2$ bin width.

$\Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$		$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$	
	Yield	Difference	Yield	Difference
Default	879024 ± 1011		482718 ± 795	
Bin width $2 \text{ MeV}/c^2$	878977 ± 1011	0%	482696 ± 781	0%
Unfixed signal shape	875686 ± 1093	0.3%	479043 ± 1894	0.8%
Uncertainty assigned	0.3%		0.8%	

7.1.2 Yield of Ξ_b^-

In the default fitting, the signal shape is fixed to the simulation sample with a width inflation coefficient of 1.1. Here we tested two different coefficients 1.05 and 1.15 and check how much the yield changes. The result is summarized in Table. 13.

Table 13: Ξ_b^- yield of the default fitting (inflation coefficient 1.1), and two different coefficient 1.0 and 1.2.

$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$		$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$	
	Yield	Difference	Yield	Difference
Coefficient 1.05	107 ± 18	1.8%	142 ± 20	2.7%
Coefficient 1.10 (Default)	109 ± 18		146 ± 21	
Coefficient 1.15	110 ± 19	0.9%	149 ± 21	2.1%
Uncertainty assigned	1.4%		2.4%	

7.2 Efficiency determination

7.2.1 Efficiency of Λ_b^0 mass window

A $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ mass window selection is applied to the data of both normalization and signal modes, and the efficiency of this selection is calculated from simulation samples of the normalization and signal modes. But the real efficiency in data heavily depends on the actual mass resolution of the Λ_b^0 resonance which may deviate from that in simulation.

For the normalization mode, Table. 14 lists the efficiencies of this mass window selection in the simulation and data samples. In this table, ϵ^{MC} is the selection efficiency calculated from simulation samples, and f^{Data} is the integral of the normalized fitted signal shape inside mass window $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ of data samples. The difference between ϵ^{MC} and f^{Data} is generally around 0.5%. Assuming the signal mode has the same trend and similar scale of this difference, the correction on the relative efficiency $\epsilon^{Norm}/\epsilon^{Sig}$ would be negligible. Therefore we decide not to consider this effect as a source of the systematic uncertainties.

Table 14: The efficiencies of $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ mass window selection in the simulation and data samples. ϵ^{MC} is the selection efficiency calculated from simulation samples, and f^{Data} is the integral of the normalized fitted signal shape inside mass window $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ of data samples.

$\Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$			$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$		
	ϵ^{MC}	f^{Data}	f^{Data}/ϵ^{MC}	ϵ^{MC}	f^{Data}	f^{Data}/ϵ^{MC}
2016	0.964	0.958	0.993	0.942	0.944	1.001
2017	0.962	0.959	0.997	0.946	0.945	0.999
2018	0.962	0.959	0.998	0.951	0.945	.0994

7.2.2 Efficiency uncertainty from $p_T - \eta$ weights

The calculation of $p_T - \eta$ weights discussed in Sec.4.3 introduces a statistical uncertainty to each weight value. To evaluate how this uncertainty influence the relative efficiency $\epsilon^{Rel} = \epsilon^{Norm}/\epsilon^{Sig}$, it is done through the procedures listed below:

- For each weight value in the diagram like Fig. 35 and Fig. 37, an uncertainty value $r \times \sigma$ is added, where r is a normally-distributed random number and σ is the statistical uncertainty of this weight. This generates a new version of weight diagram.
- The uncertainty of 1-D p_T weight in Fig. 37 is fixed to the statistical uncertainty of the $2.5 \sim 5 \text{ GeV}/c$ bin, which should be the worse case compared to the uncertainty of the fitted line.

- 100 different versions of weight diagram are generated and applied to the corresponding Monte-Carlo samples.
- The relative efficiency is re-calculated from new MC samples for 100 iterations. The result of each iteration is filled into a histogram. Fig. 48 shows the histograms of year 2018 MC samples.
- For each mode, the mean value of 100 iterations is compared to the original relative efficiency value, and the RMS is set to be the systematic uncertainty of the relative efficiency.

The result of such process, using 2018 MC samples, are summarized in the Table. 15.

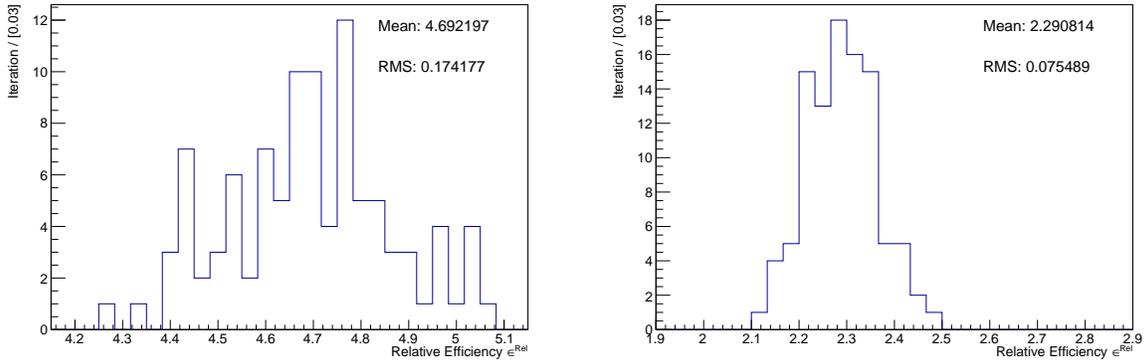


Figure 48: Relative efficiency ϵ^{rel} of 100 iterations for mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ (Left) and mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ (Right). 2018 MC samples are used.

Table 15: Mean and RMS of relative efficiency for 100 iterations. 2018 simulation samples are used.

$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$	$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$
Original ϵ^{rel} (2018)	4.74	2.31
Mean of ϵ^{rel}	4.69	2.29
RMS, $\delta\epsilon$, of ϵ^{rel}	0.17	0.08
Uncertainty assigned $\delta\epsilon/\epsilon^{rel}$	$\pm 3.6\%$	$\pm 3.4\%$

7.3 $\Xi_b^- p_T$ spectra in simulation

The Ξ_b^- efficiencies are obtained from simulated signal decays, where the signal weights are obtained from $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ decays, with $\Xi_c^0 \rightarrow p K^- K^- \pi^+$. Because of differing kinematics between the $\Xi_b^- \rightarrow \Xi_c^0 \pi^-$ and $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ modes, the assigned weights could have some small biases. To estimate the potential magnitude of this bias, the relative efficiencies are re-evaluated, applying the Λ_b^0 weights to the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ simulated decays as the most conservative case. The fractional change in the relative efficiency is assigned

as the systematic uncertainty, 3.2% for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^- \Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode, and 5.6% for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode.

7.4 Multiple Candidates

In reconstructing the parent b -hadrons, the final states in one event may be combined in different ways and result in different parent particles. For example, in $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$, the bachelor π^- can switch with the π^- from Λ_b^0 and form a different Λ_b^0 and Ξ_b^- . This is called multiple candidates. In the ideal case, the rate of multiple candidates in simulation can correctly reflect the rate in data samples but it should be checked in the real case.

Fig. 49 shows the histograms of the number of multiple candidates in one event for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ (see Appendix. N for the other mode). Table. 16 summarized the mean number of candidates in one event (See the legends on Fig. 49) after all selections in three-year combined simulation and data samples. For simulation and data, the candidates are limited in the peak region, i.e. $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ for normalization modes and $34.8 < \Delta m < 41.6 \text{ MeV}/c^2$ for signal modes. For comparison, we also listed the same value in the sideband region of data samples, i.e. $5700 < m(\Lambda_b^0) < 5820 \text{ MeV}/c^2$ for normalization modes and $27.8 < \Delta m < 34.8 \text{ MeV}/c^2$ for signal modes.

Table. 16 shows that the simulation samples roughly have the same multiple candidates rate of data, both in the peak region and sideband region. However, in $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode there is a relatively large deviation. The difference is assigned as a systematic uncertainty.

Table 16: the mean number of candidates in one event (See the legends on Fig. 49) after all selections in three-year combined simulation and data samples. For simulation and data, the candidates are limited in the peak region, i.e. $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ for normalization modes and $34.8 < \Delta m < 41.6 \text{ MeV}/c^2$ for signal modes.

$\Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$			$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$		
	MC	Data - peak	Data - Sideband	MC	Data - peak	Data - Sideband
Normalization mode	1.004	1.002	1.001	1.032	1.035	1.029
Signal mode	1.003	1	1.005	1.026	1.046	1.020
Uncertainty assigned	$\pm 0.5\%$			$\pm 2.6\%$		

7.5 Pion tracking efficiency

The tracking efficiency is calibrated using large samples of $J/\psi \rightarrow \mu^+ \mu^-$ decays, and provides a data-to-simulation correction for tracks in the momentum range from 5–200 GeV/c and $1.9 < \eta < 4.9$ [52]. The corrections are mostly within 1–2% of unity. About 65% of the π^- mesons from the Ξ_b^- decay have momentum

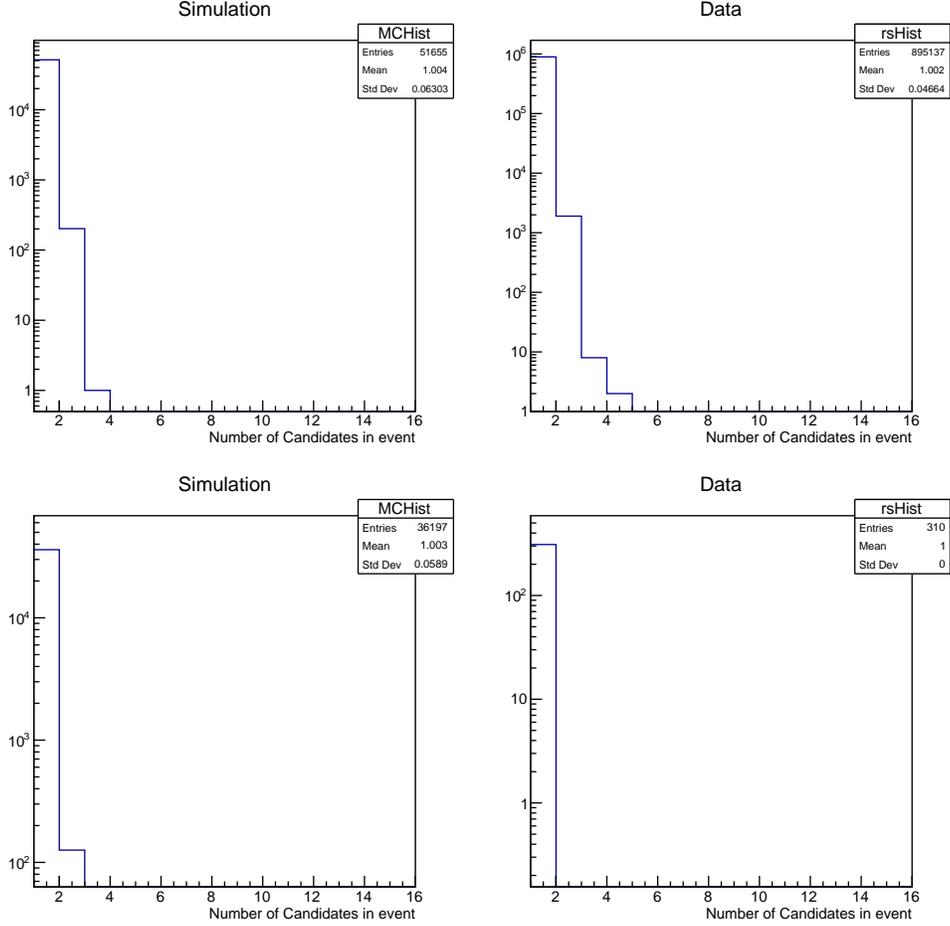


Figure 49: Histograms of the number of multiple candidates in one event in simulation (Left) and data samples (Right) for normalization mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ with $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$ (Top) and signal mode $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ with $27.8 < \Delta m < 34.8$ (Bottom).

below $5 \text{ GeV}/c$, and for these cases, the correction at $5 \text{ GeV}/c$ is used with an uncertainty that is inflated by a factor of two. The luminosity-weighted average correction is 0.99 ± 0.03 . An additional 1% uncertainty is assigned due to a potential difference in the number of hadronic interaction lengths in the simulated and actual detector.

7.6 BDT2 efficiency

As discussed and shown in Section 5.5.1, corrections are applied to the relative efficiencies to account for the less efficient performance of the BDT discriminant on real data than in simulation for the $\text{BDT2} > 0.8$ requirement. Correction factors of 0.96 and 0.94 are applied to the offline selection efficiency for the $\Lambda_c^+ \pi^-$ and $\Lambda_c^+ \pi^- \pi^+ \pi^-$ final states of the Λ_b^0 . For certain, the efficiency in data is not as performant as in simulation,

and we assign half of the difference of the correction from unity as the uncertainty, i.e. 2.0% for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode and 3.0% for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode

7.7 Systematic uncertainty summary

Table. 17 summarized the systematic uncertainties discussed in this section. There could be other sources of systematic uncertainties but here we leave them for further studies in the future.

Table 17: Summary of discussed systematic uncertainties.

$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-, \Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$	$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$
Λ_b^0 yield	0.3%	0.8%
Ξ_b^- yield	1.4%	2.4%
$p_T - \eta$ weight	3.6%	3.4%
$\Xi_b^- p_T$ spectrum	3.2%	5.6%
Multiple candidates	0.5%	2.6%
π tracking efficiency	3.2%	3.2%
BDT2 Efficiency	2.0%	3.0%
Combined	6.3%	8.7%

8 Results and summary

The final set of values used in the calculation of r_s is provided in Table 18.

Table 18: Fitted values used in the calculation of r_s . Uncertainties are statistical only.

$\Lambda_b^0 \rightarrow f$	$f = \Lambda_c^+ \pi^-$	$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$
$N(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$	108.6 ± 18.3	146.0 ± 20.6
$N(\Lambda_b^0 \rightarrow f)$	879024 ± 1011	482718 ± 795
ϵ^{rel}	4.82 ± 0.09	2.53 ± 0.05
Systematic Uncertainty	6.3%	8.7%

With these values, we measure

$$\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = (5.95 \pm 1.01 \pm 0.37) \times 10^{-4} \quad (10)$$

for the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode. The uncertainties are statistical and systematic, respectively. For the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay, we obtain

$$\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = (7.66 \pm 1.09 \pm 0.67) \times 10^{-4} \quad (11)$$

The two results are statistically compatible, with a difference of $1.06\sigma_{\text{stat}}$. Combining the result of two modes, taking the first three uncertainties in Table 17 as uncorrelated, and the remaining four as 100% correlated. Combining the two results, the average value is

$$\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = (6.74 \pm 0.74 \pm 0.50) \times 10^{-4} \quad (12)$$

Using the measured ratio, $\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} = (8.2 \pm 0.7 \pm 0.6 \pm 2.5) \times 10^{-2}$ [23], the branching fraction is

$$\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-) = (0.82 \pm 0.09 \pm 0.07 \pm 0.25_{SU(3)})\% \quad (13)$$

where the last item of uncertainty contains a part inherited from the third uncertainty of $\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}}$, which is an estimate of the typical size of SU(3)-breaking effects and taken to be 30%.

The value is measured with about 11% statistical precision, compared to about 32% statistical uncertainty from the Run 1 measurement. Thus the combination of the two data sets from Run 2 has effectively provided about a nine-fold increase in signal yield.

This branching fraction is consistent with the naive expectation derived from $\Gamma(\Lambda^0)/\Gamma(\Xi_b^-) = 0.0062$. Therefore this result does not indicate any anomalously large $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay rate, as could be the case if the Ξ_b^- wave function is treated as a quark-diquark bound state with the parameter values used in Ref. [21].

As the decay $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ signal is observed with a significance of 11.3σ , the decay width of $\Gamma(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$ should be included in the total decay width (Γ_{tot}) when comparing to measured lifetimes in experiments. This contribution increases the width by about 1% and thus would decrease the theoretical predictions for the Ξ_b^- lifetime by about 1%.

Appendices

A StrippingLine Selections

The $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ candidates are taken from the `StrippingLb2LcPiNoIPLc2PKPiFullDSTBeauty2CharmLine` stripping line using Stripping versions v28r2, v29r2 and v34 for 2016, 2017 and 2018 samples, respectively. The $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ candidates are taken from the `StrippingLb2LcPiPiPiLc2PKPiFullDSTBeauty2CharmLine` stripping line, using v28r2, v29r2p1 and v34r0p1, corresponding to the 2016, 2017 and 2018 data samples.

A.1 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ stripping selections

- $\Lambda_c^+ \rightarrow pK^- \pi^+$:
 - On p , K^- , π^+ : track $\chi^2/\text{ndf} < 4$, $p_T > 100 \text{ MeV}/c$, $p > 1000 \text{ MeV}/c$, $\chi_{\text{IP}}^2 > 4$ and Ghost Probability(TRGHP) < 0.4 ; PID requirement $\text{PID}p > -10.0$, $\text{PID}K > -10.0$, $\text{PID}K < 20$ on p , K^- , π^+ respectively;
 - On Λ_c^+ : $\sum(p_T) > 1800 \text{ MeV}/c$, ACUTDOCA⁹ $< 0.5 \text{ mm}$, at least one daughter with $p_T > 500 \text{ MeV}/c$, $p > 5000 \text{ MeV}/c$ and with track $\chi^2/\text{ndf} < 4$, $\chi_{\text{vtx}}^2/\text{ndf} < 10$, $\chi_{\text{VS}}^2 > 36$, DIRA¹⁰ > 0 , mass within $100 \text{ MeV}/c^2$ of nominal Λ_c^+ mass.
- $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$:
 - Λ_c^+ : mass within $60 \text{ MeV}/c^2$ of nominal Λ_c^+ mass.
 - π^\pm : $\text{PID}K < 10$, $p_T > 100 \text{ MeV}/c$ and $p > 2000 \text{ MeV}/c$;
 - π^- from Λ_b^0 : $\chi^2/\text{ndf} < 4$, $p_T > 500 \text{ MeV}/c$ and $p > 5000 \text{ MeV}/c$.
 - Λ_b^0 : $\chi_{\text{vtx}}^2/\text{ndf} < 10$, at least 1 track with track $\chi^2/\text{ndf} < 4$, $p_T > 1700 \text{ MeV}/c$, $p > 10000 \text{ MeV}/c$ and $\chi_{\text{IP}}^2 > 16$ and $\text{IP} > 100 \mu\text{m}$; at least 2 tracks with $\chi^2/\text{ndf} < 2.5$, $p_T > 500 \text{ MeV}/c$, $p > 5000 \text{ MeV}/c$ and $\chi_{\text{VS}}^2 > 1000$; $t_{\text{decay}} > 0.2 \text{ ps}$; $\chi_{\text{IP}}^2 < 25$; $\text{DIRA} > 0.999$; $5200 < M(\Lambda_b^0) < 7000 \text{ MeV}/c^2$, $\sum(p_T) > 5000 \text{ MeV}/c$.

⁹ACUTDOCA is the maximum distance between any pair of daughter tracks.

¹⁰DIRA is the cosine of the angle between the parent particle's momentum vector, and the line that joins the decay vertex and its associated PV.

A.2 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ stripping selections

- $\Lambda_c^+ \rightarrow pK^- \pi^+$:
 - On p , K^- , π^+ : track $\chi^2/\text{ndf} < 4$, $p_T > 100 \text{ MeV}/c$, $p > 1000 \text{ MeV}/c$, $\chi_{\text{IP}}^2 > 4$ and Ghost Probability(TRGHP) < 0.4 ; PID requirement $\text{PID}_p > -10.0$, $\text{PID}_K > -10.0$, $\text{PID}_K < 20$ on p , K^- , π^+ respectively;
 - On Λ_c^+ : $\sum(p_T) > 1800 \text{ MeV}/c$, ACUTDOCA¹¹ $< 0.5 \text{ mm}$, at least one daughter with $p_T > 500 \text{ MeV}/c$, $p > 5000 \text{ MeV}/c$ and with track $\chi^2/\text{ndf} < 4$, $\chi_{\text{vtx}}^2/\text{ndf} < 10$, $\chi_{\text{VS}}^2 > 36$, DIRA¹² > 0 , mass within $100 \text{ MeV}/c^2$ of nominal Λ_c^+ mass.
- $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$:
 - Λ_c^+ : mass within $60 \text{ MeV}/c^2$ of nominal Λ_c^+ mass.
 - π^\pm : $\text{PID}_K < 10$, $p_T > 100 \text{ MeV}/c$ and $p > 2000 \text{ MeV}/c$;
 - virtual particle $a_1^- \rightarrow \pi^- \pi^- \pi^+$: $\sum(p_T) > 1250 \text{ MeV}/c$ and no more than one π has $p_T < 300 \text{ MeV}/c$, at least one daughter with $\chi^2/\text{ndf} < 4$, $p_T > 500 \text{ MeV}/c$ and $p > 5000 \text{ MeV}/c$, ACUTDOCA $< 0.4 \text{ mm}$.
 $\chi_{\text{vtx}}^2/\text{ndf} < 8$ & (BPVVDCHI2 > 16) & (BPVDIRA > 0.98) & (MIPCHI2DV(PRIMARY) > 0.0) & (BPVVDRHO $> 0.1^* \text{ mm}$) & (BPVVDZ $> 2.0 \text{ mm}$), (MIPCHI2DV(PRIMARY) > 20) & (M < 2800).
 - Λ_b^0 : $\chi_{\text{vtx}}^2/\text{ndf} < 10$, at least 1 track with track $\chi^2/\text{ndf} < 4$, $p_T > 1700 \text{ MeV}/c$, $p > 10000 \text{ MeV}/c$ and $\chi_{\text{IP}}^2 > 16$ and $\text{IP} > 100 \mu\text{m}$; at least 2 tracks with $\chi^2/\text{ndf} < 2.5$, $p_T > 500 \text{ MeV}/c$, $p > 5000 \text{ MeV}/c$ and $\chi_{\text{VS}}^2 > 1000$; $t_{\text{decay}} > 0.2 \text{ ps}$; $\chi_{\text{IP}}^2 < 25$; DIRA > 0.999 ; $5200 < M(\Lambda_b^0) < 7000 \text{ MeV}/c^2$, $\sum(p_T) > 5000 \text{ MeV}/c$.

Both of these stripping lines also require a positive decision on either the TOPO(2,3,4)-BODY BBDT HLT2 trigger or the inclusive ϕ trigger.

B Selections applied in DaVinci script

To reconstruct Ξ_b^- , the selections applied to candidates from the strippings are listed below:

- Λ_c^+ : $|M(pK\pi) - M(\Lambda_c^+)| < 50 \text{ MeV}/c$, Proton $\text{PROBN}p > 0.2$ and Kaon $\text{PROBN}k > 0.1$;

¹¹ACUTDOCA is the maximum distance between any pair of daughter tracks.

¹²DIRA is the cosine of the angle between the parent particle's momentum vector, and the line that joins the decay vertex and its associated PV.

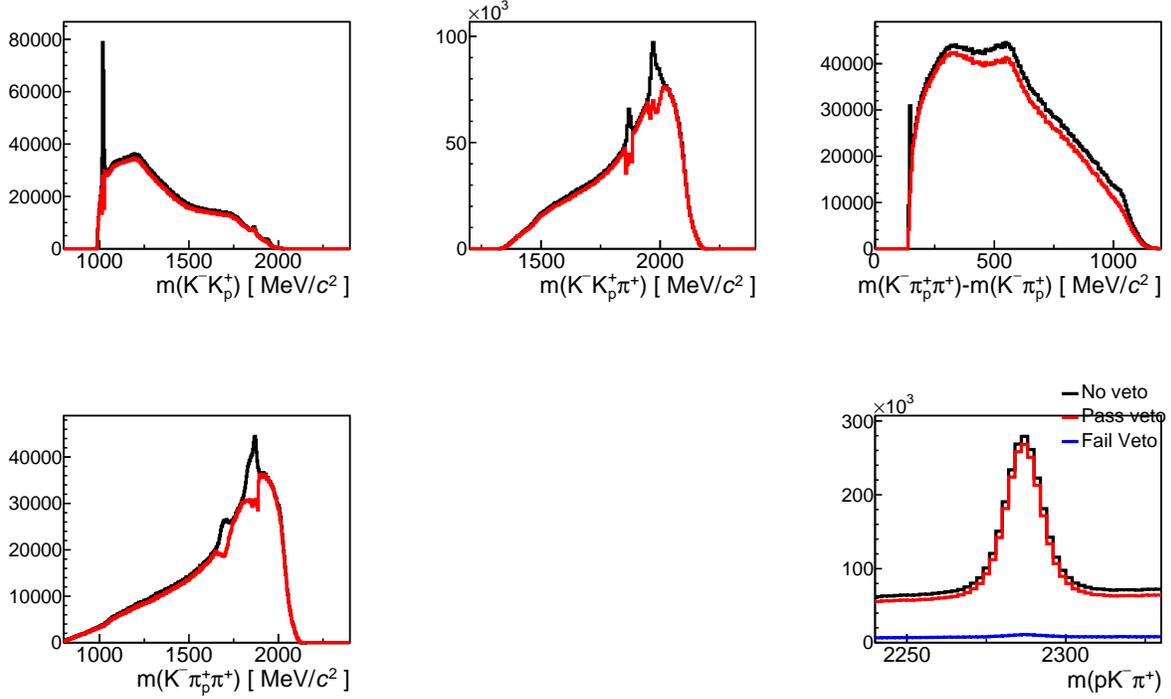


Figure 50: Distributions of the invariant mass ($\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^+$ data): (top left) $m(K^+ K^-)$, (top middle) $m(K^+ K^- \pi^+)$, (top right) $m(K^- \pi^+ \pi^+) - m(K^- \pi^-)$ and (bottom left) $m(K^- \pi^+ \pi^+)$. The black histogram corresponds to data before the veto, and red is after the veto.

- Λ_b^0 : DIRA > 0.9999 and $5350 < M(\Lambda_b^0) < 5950 \text{ MeV}/c^2$;
- Ξ_b^- : $\chi_{\text{vtx}}^2/\text{ndf} < 10$, $5500 < M(\Lambda_b^0) < 5750 \text{ MeV}/c^2$. For the bachelor π^- , it requires $\chi_{\text{track}}^2/\text{ndf} < 4$, GhostProb < 0.4, $p_T > 1000 \text{ MeV}/c$ and to be long track.

C Vetoes on other charm final states for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$

These mass distributions before and after the veto for the data of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ are shown in Fig. 50.

D Clone track removal

Fig. 51 shows the histograms of the opening angle α between a number of the track pairs and shows a spike near $\alpha = 0$ in decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$. Fig. 52 shows a closer look at the angle α between the proton

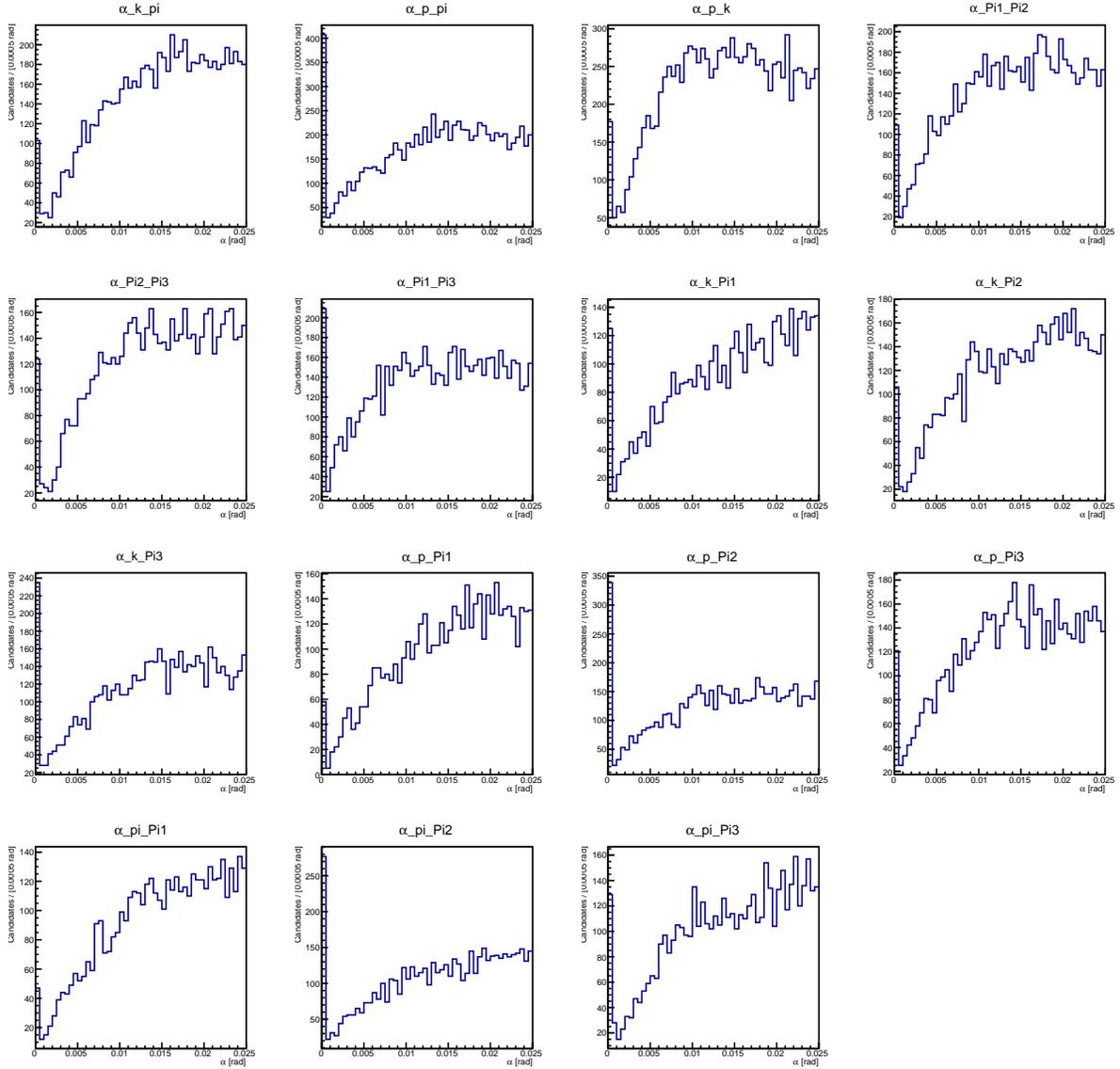


Figure 51: Distributions of the opening angle α of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$, in the range from 0 to 2.5 mrad. Pi1/2/3 refers to the $\pi^- \pi^- \pi^+$ from the Λ_b^0 decay.

momentum and the π^- from the Λ_b^0 . It shows that a requirement of $\alpha > 0.8\text{mrad}$ removes the clone tracks, with a negligible loss of real signal decays.

Fig. 53 below shows the α angle distributions for the bachelor π^- from Ξ_b^- and each of the tracks from the Λ_b^0 . The overlapping pairs of tracks are removed with a requirement that $\alpha > 0.8$ mrad.

This procedure is also applied to decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ and the results are similar which are shown in Fig. 54 to Fig. 56.

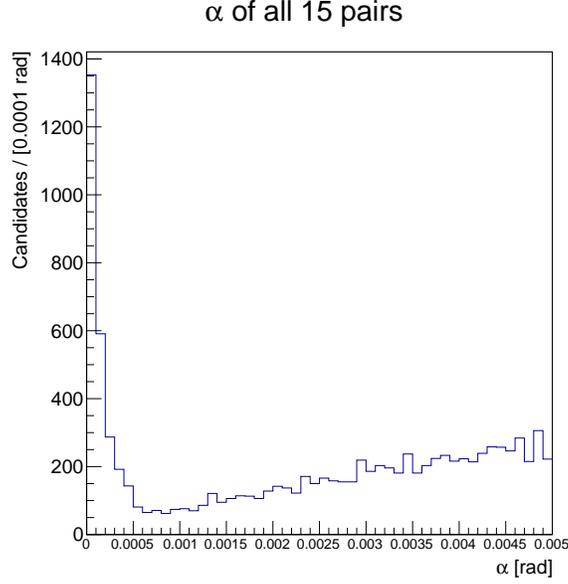


Figure 52: A zoom of the histogram of α of all 15 track pairs combined of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$.

E BDT1 input variables for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$

This appendix shows the input variables used in the BDT1 training for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay (see Figs. 57, 58 and 59).

F BDT1 input variables for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$

This appendix shows the input variables used in the BDT1 training for the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decay (see Figs. 60 and 61). The variables include:

- Λ_b^0 : Vertex fit χ_{vtx}^2 , decay time, DIRA angle, radial flight distance;
- Λ_c^+ : Vertex fit χ_{vtx}^2 and decay time;
- χ_{IP}^2 of the p , K^- and π^+ from the Λ_c^+ and π^- from the Λ_b^0 ;
- Momentum (p) and transverse momentum (p_T) of the p , K^- and π^+ from the Λ_c^+ and the π^- from the Λ_b^0 ;
- PID value (ProbNNx) of the p , K^- and π^+ from the Λ_c^+ and the π^- from the Λ_b^0 .

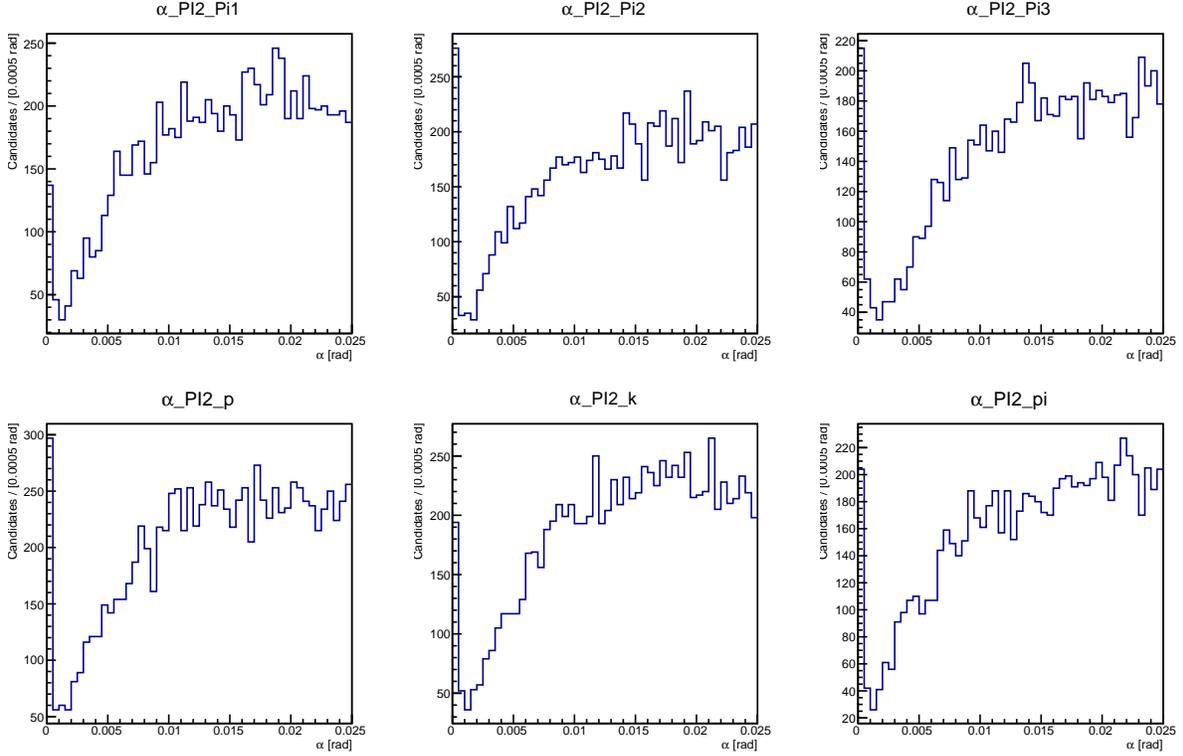


Figure 53: Distributions of the opening angle α where PI2 refers to the pion from the Ξ_b^- and Pi1/2/3 refers to the $\pi^- \pi^- \pi^+$ from the Λ_b^0 .

G BDT2 input variables and result for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decay

The distributions of the input variables to BDT2 for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay are shown in Fig. 62. We observe good separation for many of the input variables. The BDT2 output response is shown in Fig. 63.

Similar to Fig. 27, for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ decay, the RS and WS BDT2 spectra and the ratio from the low δm region, $10 < \Delta m < 32.89 \text{ MeV}/c^2$ are shown in Fig. 64.

H Weighting of the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation

The ratio of the data to simulated distributions of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ is shown in Fig. 65. The projections onto the p_T and η axes are shown in Fig. 66

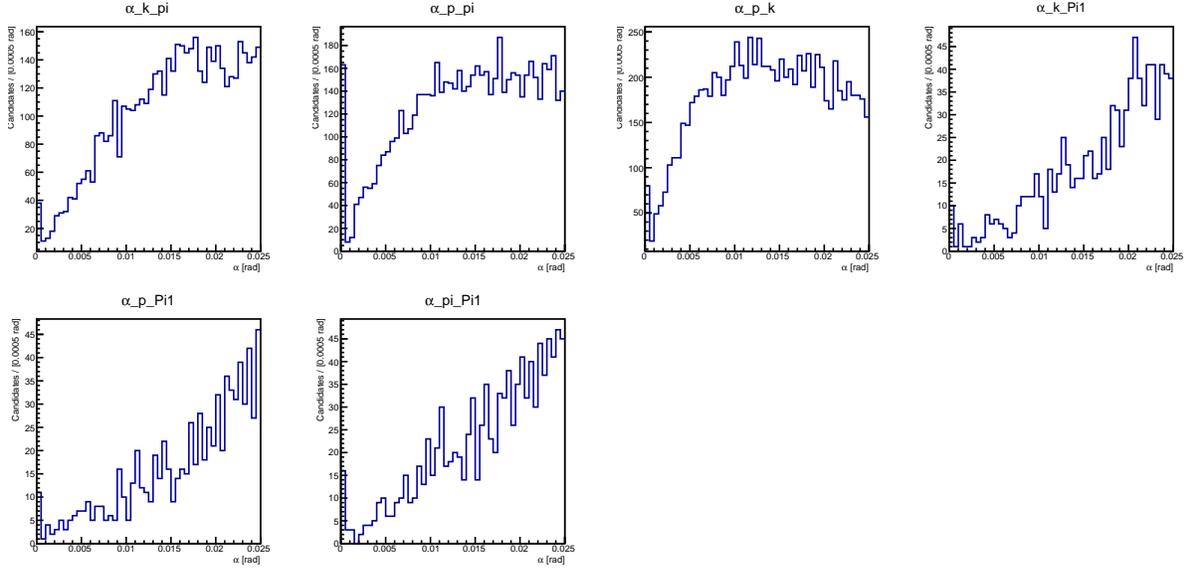


Figure 54: Distributions of the opening angle α of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$, in the range from 0 to 2.5 mrad. Pi1 refers to the π^- from the Λ_b^0 decay.

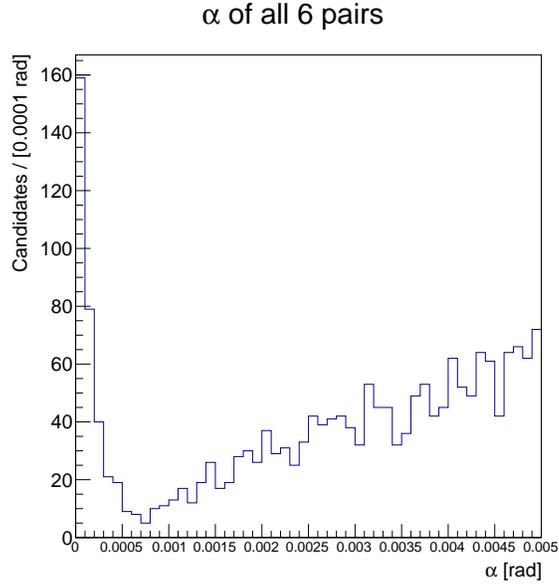


Figure 55: A zoom of the histogram of α of all 6 track pairs combined of decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$.

I Offline selection efficiencies of each year

Table 19 lists the efficiencies of each offline selection step for each year. Each item is explained as below:

- ϵ_{basic} : the efficiency of the basic Λ_b^0 selections which are discussed in Section 3.4;

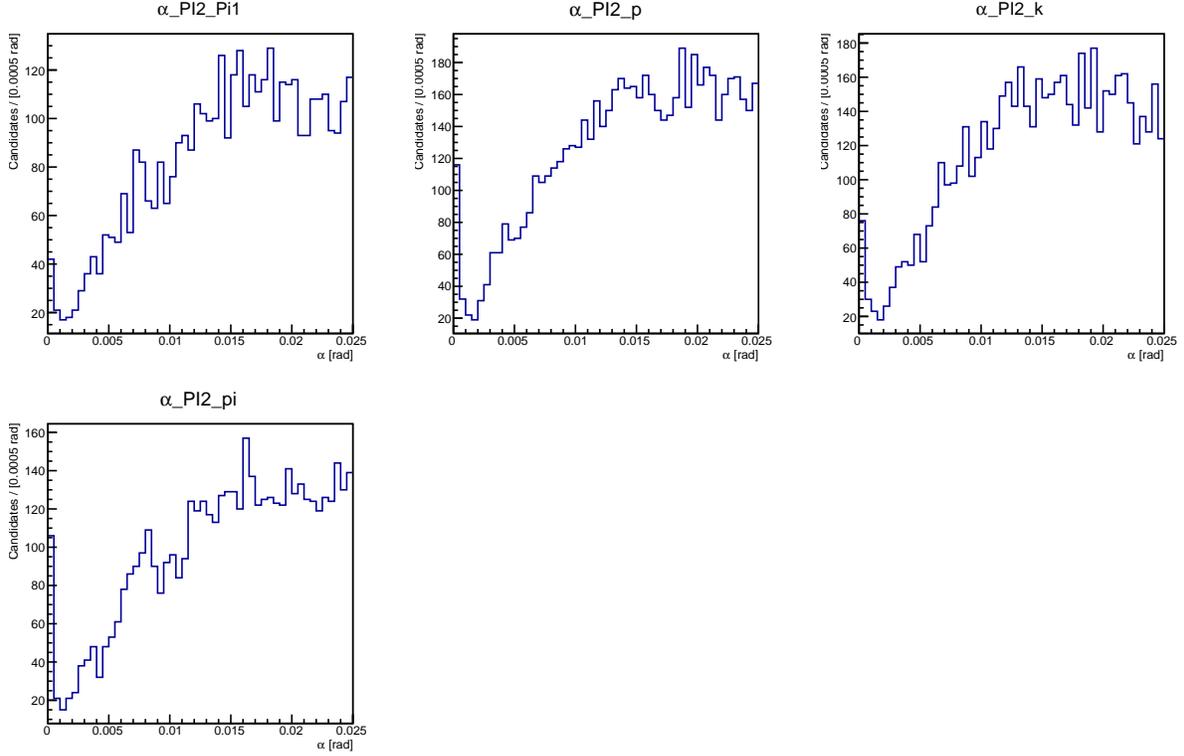


Figure 56: Distributions of the opening angle α where PI2 refers to the π^- from the Ξ_b^- and Pi1 refers to the π^- from the Λ_b^0 .

- ϵ_{BDT1} : the efficiency of $BDT1 > 0$ requirement;
- ϵ_{PI2} : the efficiency of the bachelor π^- $PROBNN\pi > 0.01$;
- $\epsilon_{\Xi_b^-} / \epsilon_{\Lambda_b^0 \text{mass}}$: $\epsilon_{\Xi_b^-}$ is the efficiency of the Ξ_b^- selections which are discussed at the beginning of Section 3.5, and $\epsilon_{\Lambda_b^0 \text{mass}}$ is the efficiency of the Λ_b^0 selection $5560 < m(\Lambda_b^0) < 5680 \text{ MeV}/c^2$.
- ϵ_{BDT2} : the efficiency of $BDT1 > 0.8$ requirement;

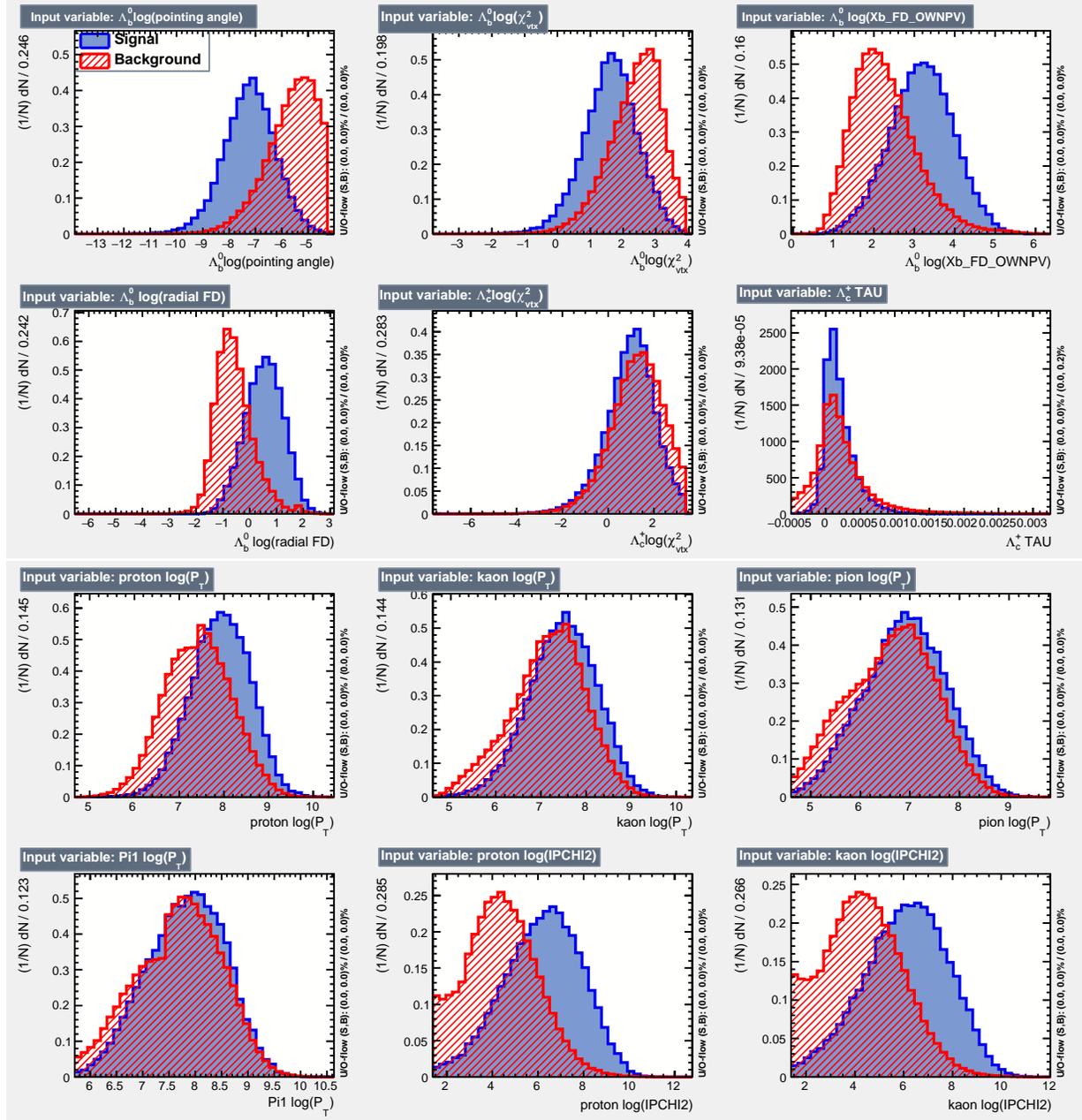


Figure 57: Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^- \pi^+$; signal and background distributions are superimposed.

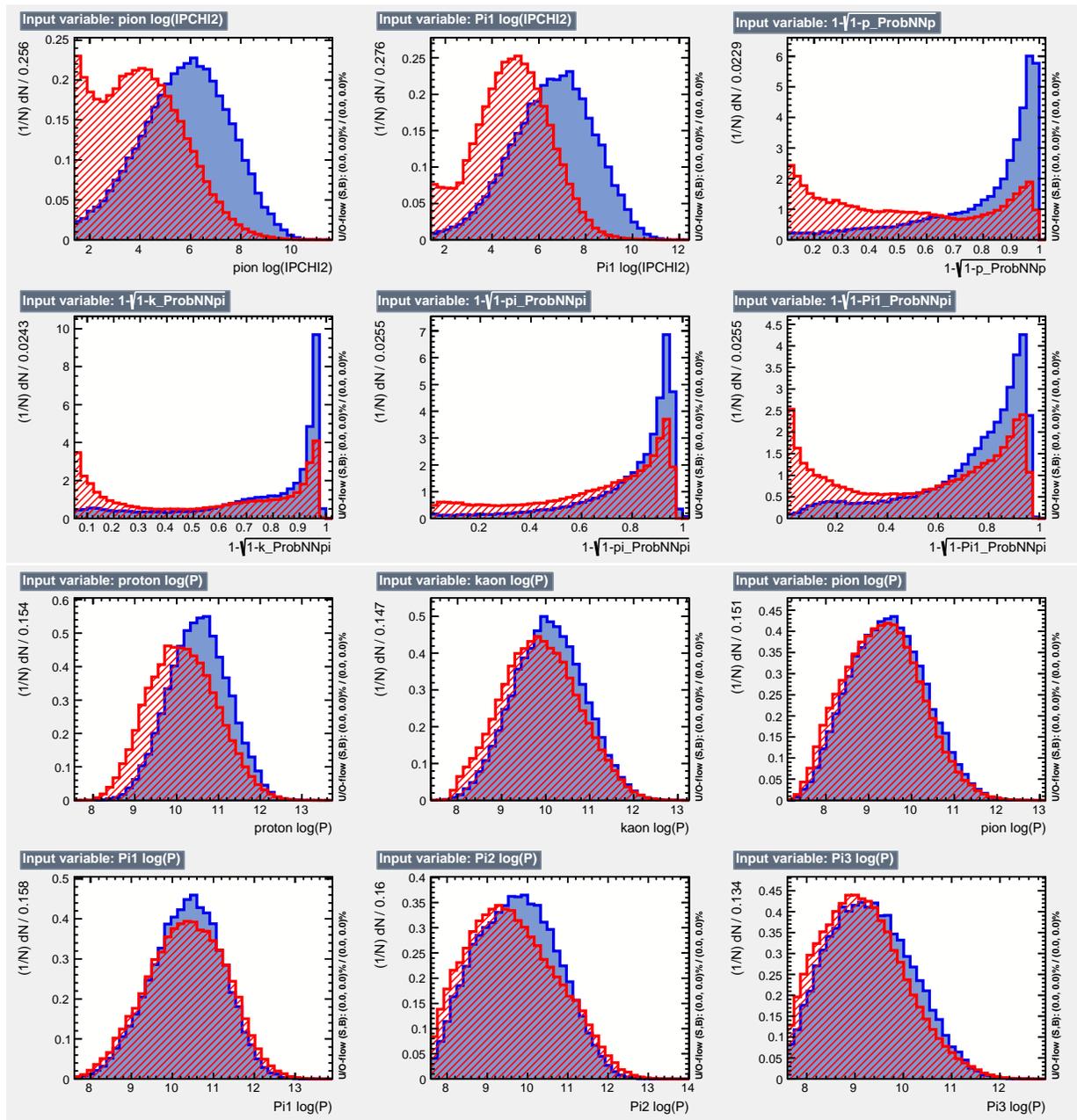


Figure 58: Distributions of the input variables used to train BDT1 for decay $A_b^0 \rightarrow A_c^+ \pi^- \pi^- \pi^+$; signal and background distributions are superimposed.

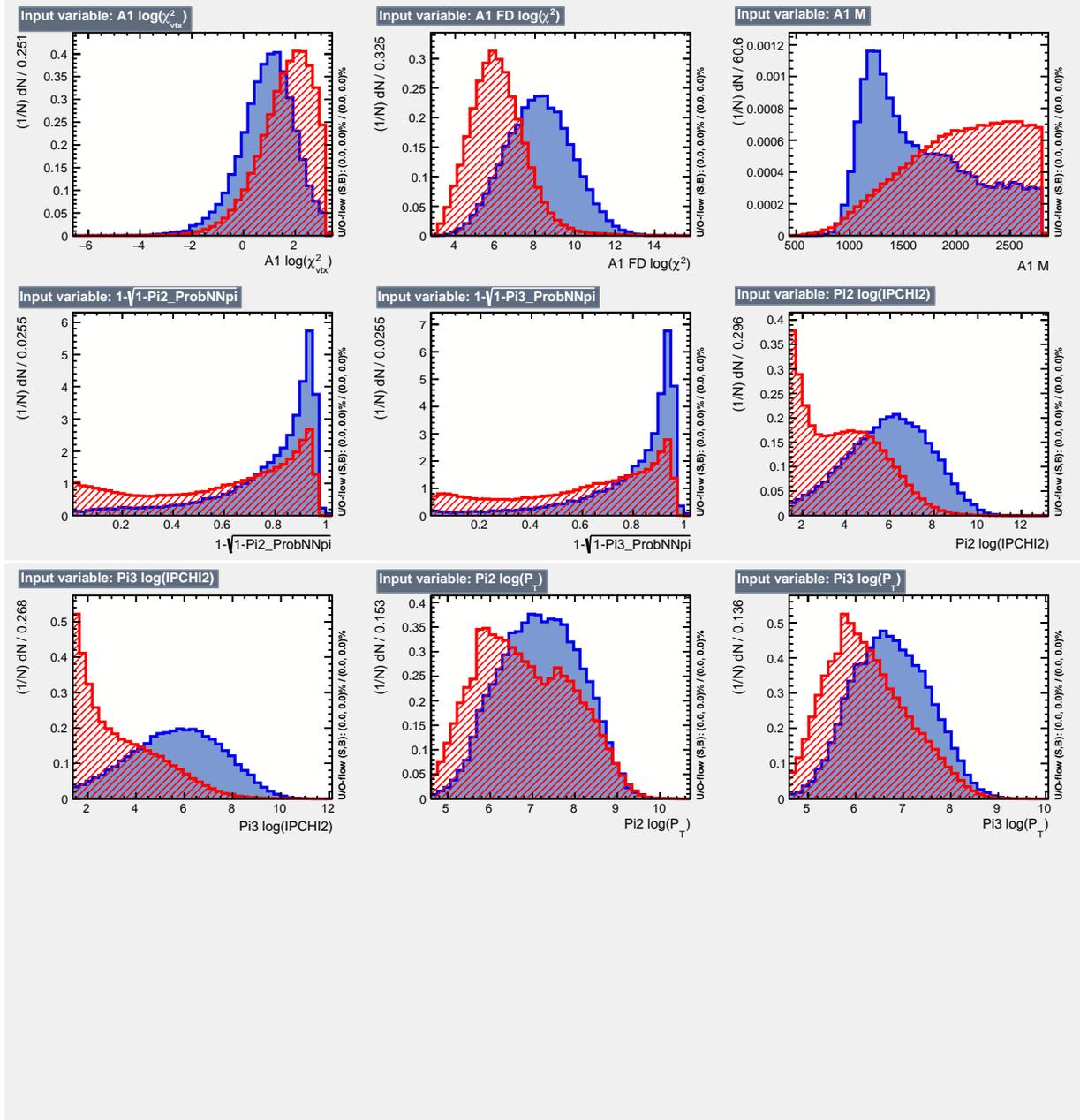


Figure 59: Distributions of the input variables used to train BDT1 for decay $A_b^0 \rightarrow A_c^+ \pi^- \pi^- \pi^+$; signal and background distributions are superimposed.

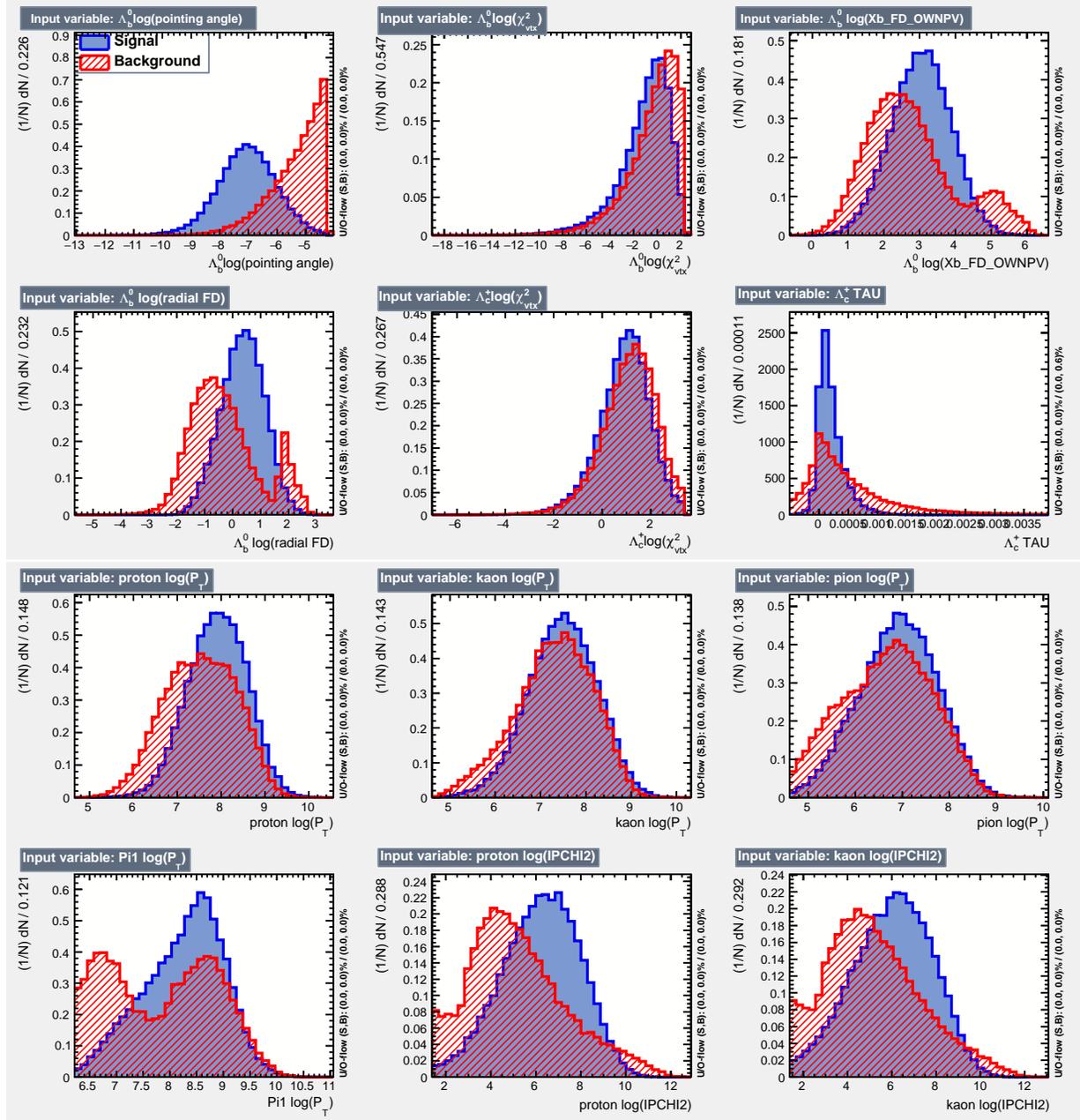


Figure 60: Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$; signal and background distributions are superimposed.

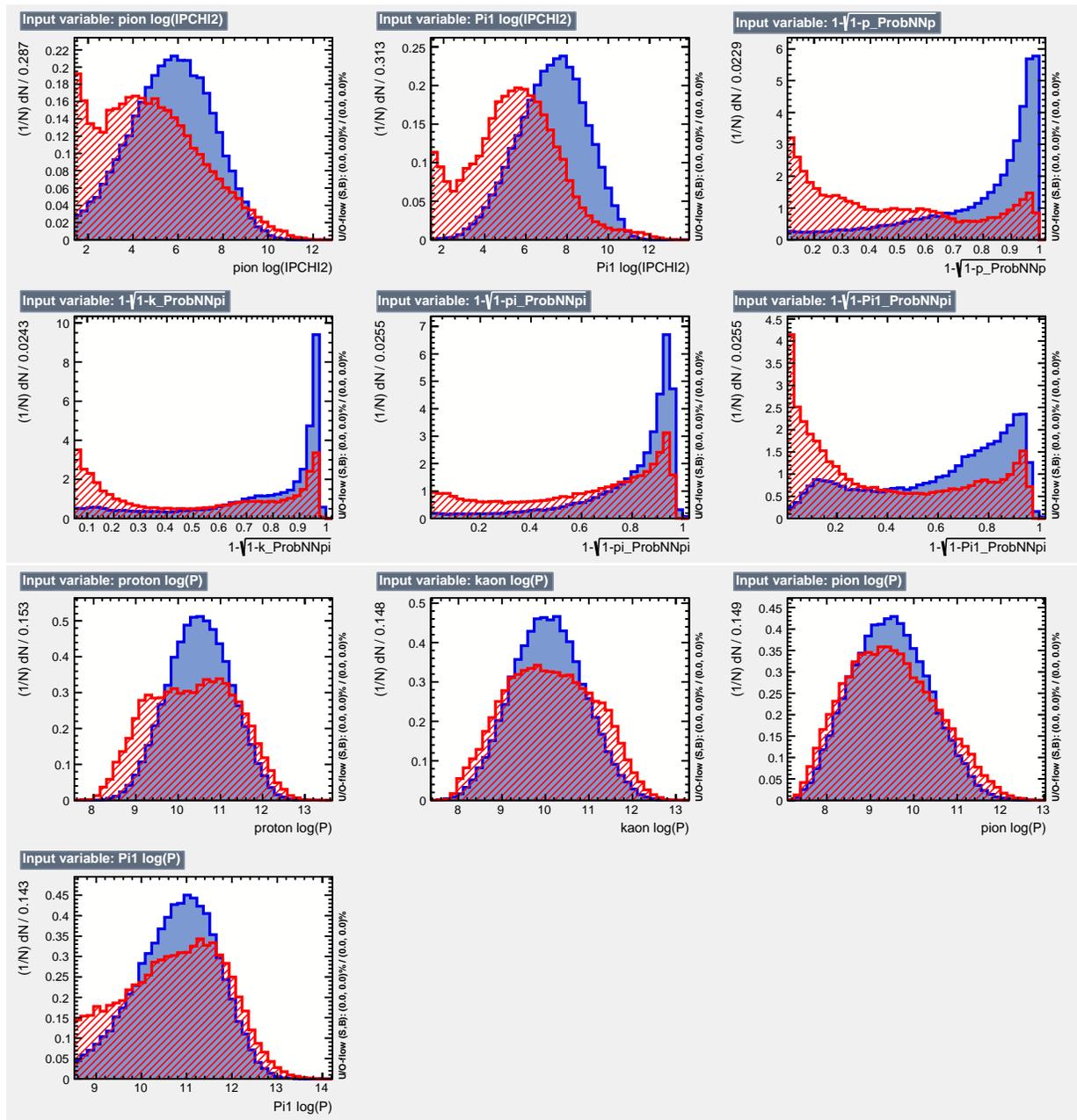


Figure 61: Distributions of the input variables used to train BDT1 for decay $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$; signal and background distributions are superimposed.

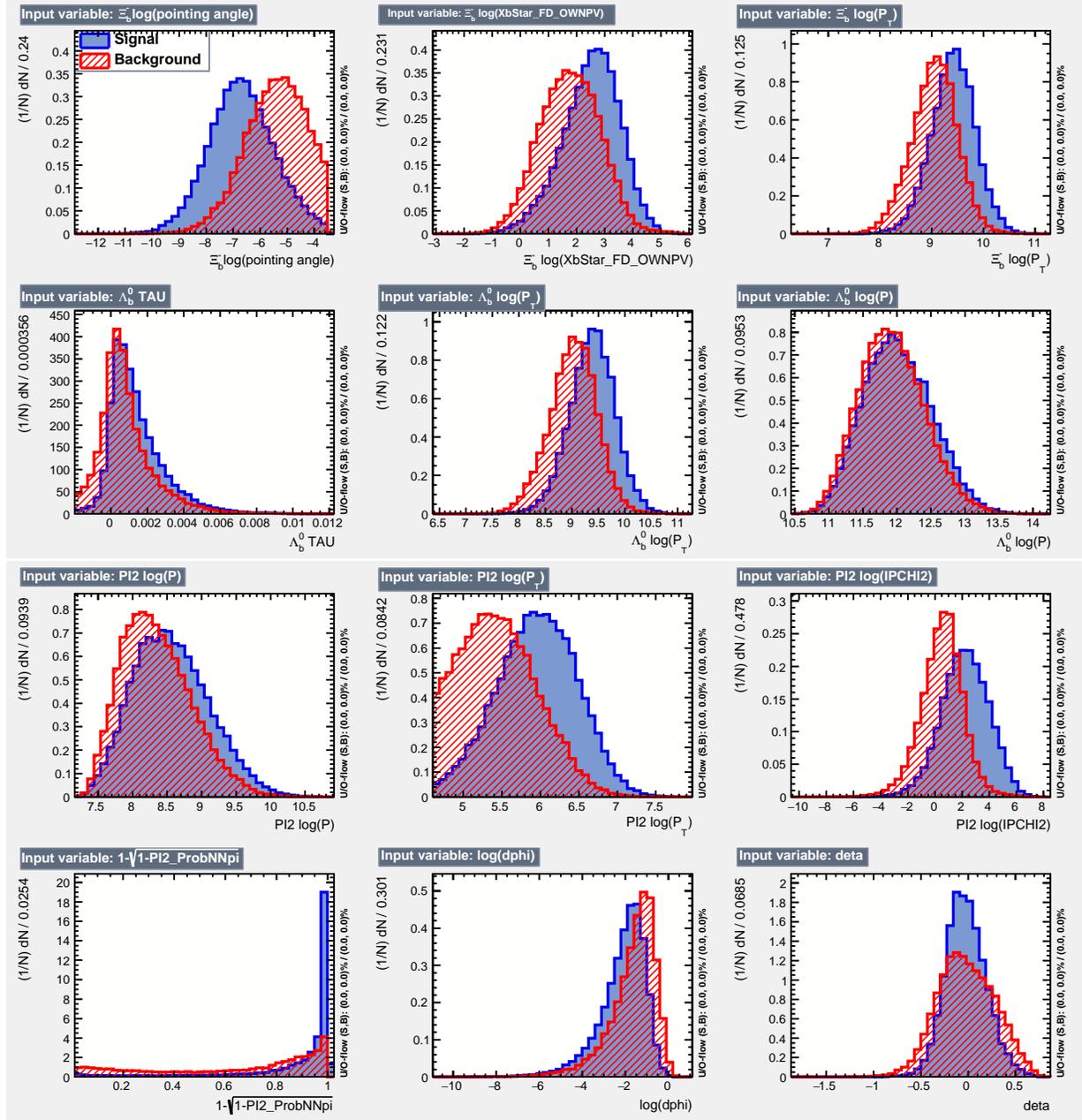


Figure 62: Distributions of the input variables used to train BDT2 for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$; signal and background distributions are superimposed.

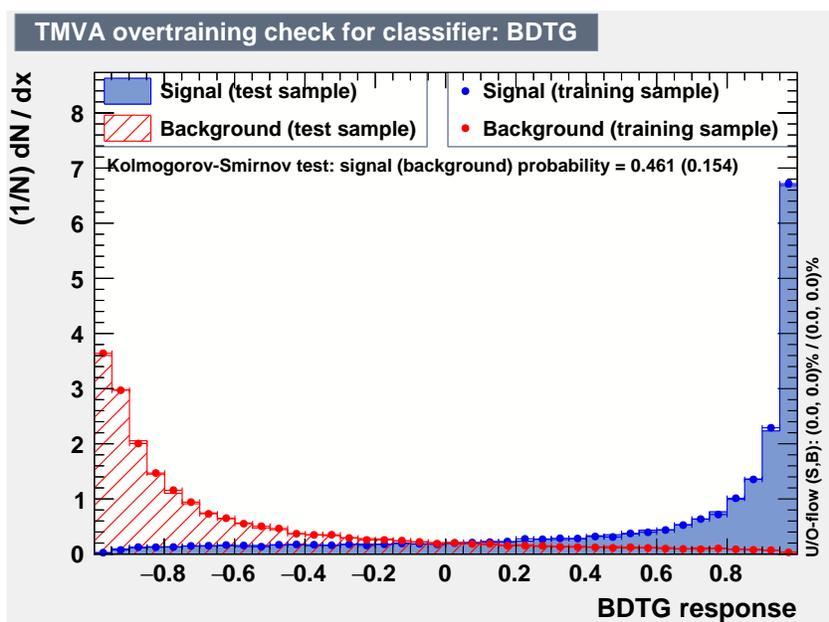


Figure 63: Distributions of the output BDT2 response for $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, with $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ signal decays and WS background. The test and training samples are both shown.

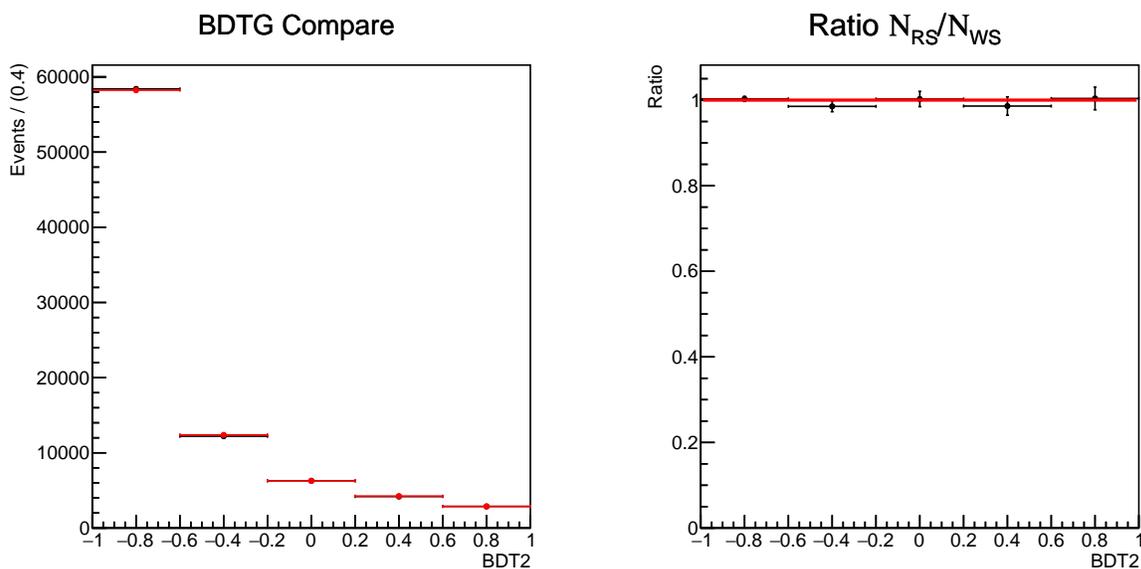


Figure 64: (Left) RS (black) and WS (red) BDT response distributions in $10 < \Delta m < 32.89 \text{ MeV}/c^2$ mass window. (Right) The ratio $\frac{N_{RS}}{N_{WS}}$ of each bin of the left plot.

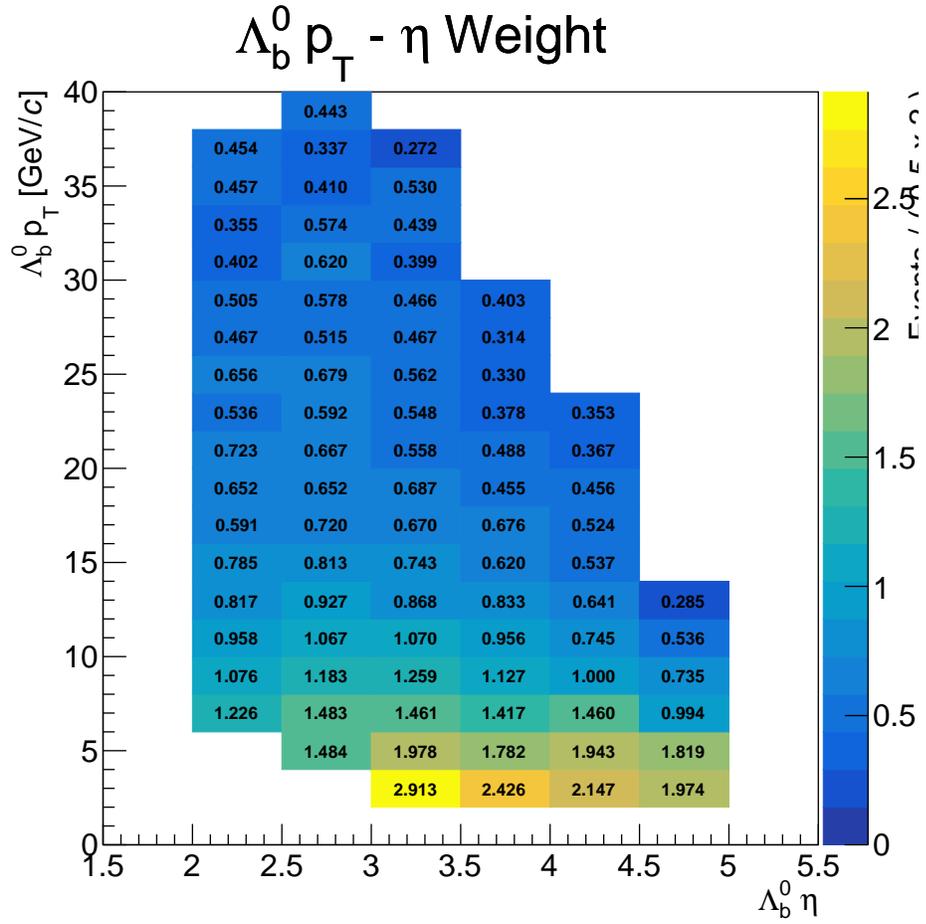


Figure 65: $\Lambda_b^0 p_T$ - η weight distribution, as obtained from the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ sWeighted data and simulated signal decays.

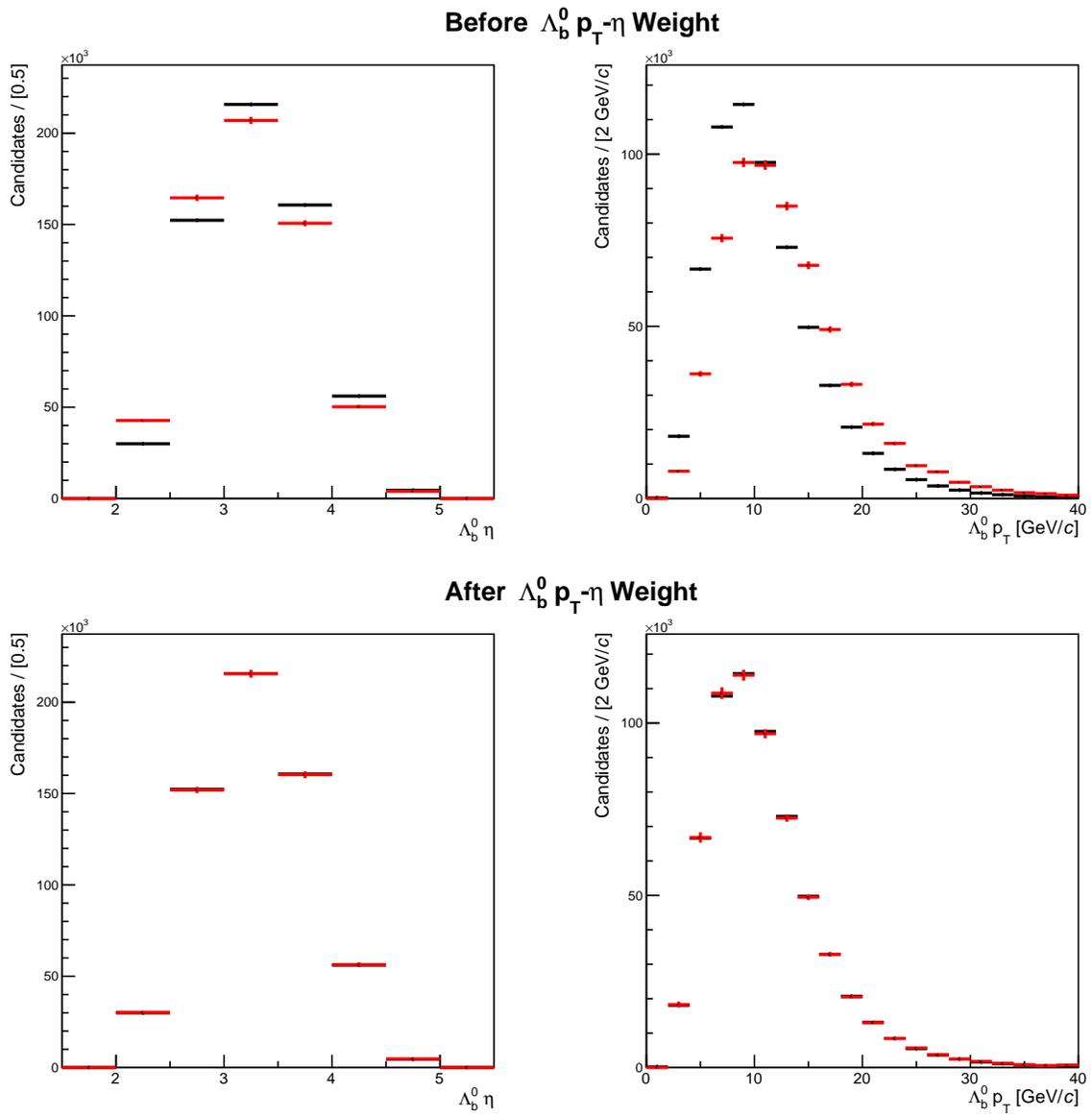


Figure 66: $\Lambda_b^0 p_T$ (Left) and η (Right) distributions before and after weight from Fig. 65

Table 19: Efficiencies and cumulative efficiencies of offline selections for each year.

Λ_b^0 mode	$f = \Lambda_c^+ \pi^-$				$f = \Lambda_c^+ \pi^- \pi^+ \pi^-$			
	$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$		$\Lambda_b^0 \rightarrow f$		$\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$		$\Lambda_b^0 \rightarrow f$	
MC sample	Eff.	Cum. Eff.	Eff.	Cum. Eff.	Eff.	Cum. Eff.	Eff.	Cum. Eff.
2016								
ϵ_{basic}	0.899	0.899	0.883	0.883	0.866	0.866	0.834	0.834
ϵ_{BDT1}	0.941	0.846	0.866	0.764	0.923	0.800	0.864	0.721
ϵ_{PI2}	0.965	0.816	/	/	0.973	0.778	/	/
$\epsilon_{\Xi_b^-} / \epsilon_{\Lambda_b^0 \text{mass}}$	0.825	0.674	0.965	0.738	0.837	0.651	0.945	0.681
ϵ_{BDT2}	0.463	0.312	/	/	0.519	0.338	/	/
ϵ_{sel}^{2016}	0.312 ± 0.002		0.731 ± 0.003		0.338 ± 0.003		0.681 ± 0.003	
2017								
ϵ_{basic}	0.898	0.898	0.887	0.887	0.875	0.875	0.859	0.859
ϵ_{BDT1}	0.938	0.842	0.868	0.770	0.929	0.813	0.874	0.751
ϵ_{PI2}	0.967	0.815	/	/	0.972	0.790	/	/
$\epsilon_{\Xi_b^-} / \epsilon_{\Lambda_b^0 \text{mass}}$	0.825	0.672	0.963	0.741	0.838	0.662	0.948	0.712
ϵ_{BDT2}	0.454	0.305	/	/	0.507	0.336	/	/
ϵ_{sel}^{2017}	0.305 ± 0.002		0.741 ± 0.003		0.336 ± 0.003		0.712 ± 0.004	
2018								
ϵ_{basic}	0.905	0.905	0.898	0.898	0.889	0.889	0.855	0.855
ϵ_{BDT1}	0.935	0.846	0.864	0.775	0.935	0.831	0.869	0.742
ϵ_{PI2}	0.966	0.817	/	/	0.991	0.823	/	/
$\epsilon_{\Xi_b^-} / \epsilon_{\Lambda_b^0 \text{mass}}$	0.830	0.678	0.962	0.746	0.834	0.686	0.952	0.707
ϵ_{BDT2}	0.468	0.317	/	/	0.529	0.363	/	/
ϵ_{sel}^{2018}	0.317 ± 0.002		0.746 ± 0.003		0.363 ± 0.003		0.707 ± 0.005	

J Correction of total relative efficiency

The correction of total relative efficiency is written as:

$$\epsilon_{rel} = \frac{\epsilon_{Norm}}{\epsilon_{Sig}} = \frac{\epsilon'_{Norm}(1 + r_{Norm})}{\epsilon'_{Sig}(1 + r_{Sig})} = \epsilon'_{rel} \frac{1 + r_{Norm}}{1 + r_{Sig}}$$

where $r = N_{Ghost}/N_{True}$, N_{Ghost} is the ghost yield, N_{True} is the true signal yield, and the efficiency with a prime ϵ' is calculated by using only the true signal candidates N_{True} .

Fig. 67 and Fig. 68 show the fitting results of true signal and ghost candidates in the normalization mode and signal mode. The signal shape is fixed to what described in Sec. 4.1 and Sec. 4.2 but the width is allowed to inflate. The background of ghost candidates is described by a flat constant function.

Table 20 and Table 21 summarized the true signal yield, ghost yield, the ratio of them and the total relative efficiency correction.

Table 20: For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ simulation samples, summary of true signal and ghost yield, the ratio r and the relative efficiency correction.

	N_{True}	N_{Ghost}	Ratio r
Normalization mode	51005 ± 227	3422 ± 64	0.067 ± 0.001
Signal mode	34437 ± 188	2720 ± 62	0.079 ± 0.002
Total ϵ^{rel} correction	0.989 ± 0.002		

Table 21: For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation samples, summary of true signal and ghost yield, the ratio r and the relative efficiency correction.

	N_{True}	N_{Ghost}	Ratio r
Normalization mode	45267 ± 216	5252 ± 83	0.116 ± 0.002
Signal mode	29827 ± 182	3929 ± 78	0.132 ± 0.003
Total ϵ^{rel} correction	0.986 ± 0.003		

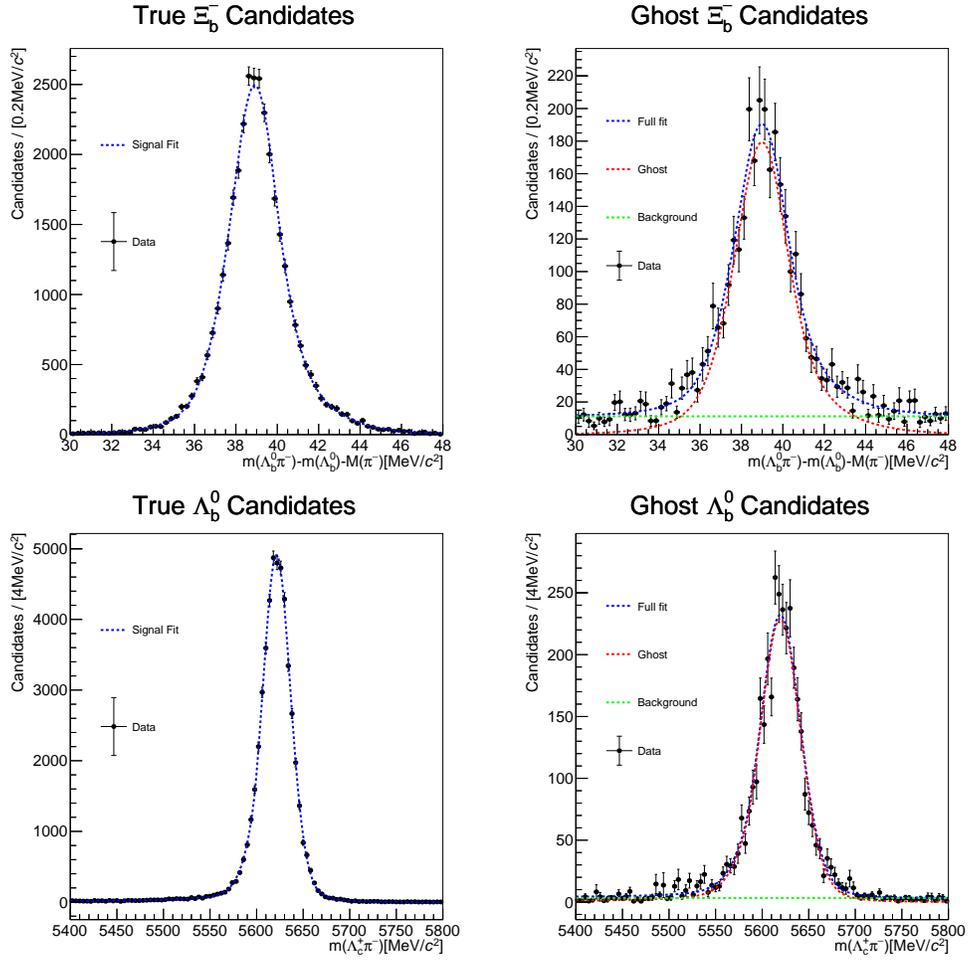


Figure 67: For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ simulation samples, fitting results of the Ξ_b^- true signal (Top Left) and ghost (Top Right), Λ_b^0 true signal (Bottom Left) and ghost (Bottom Right).

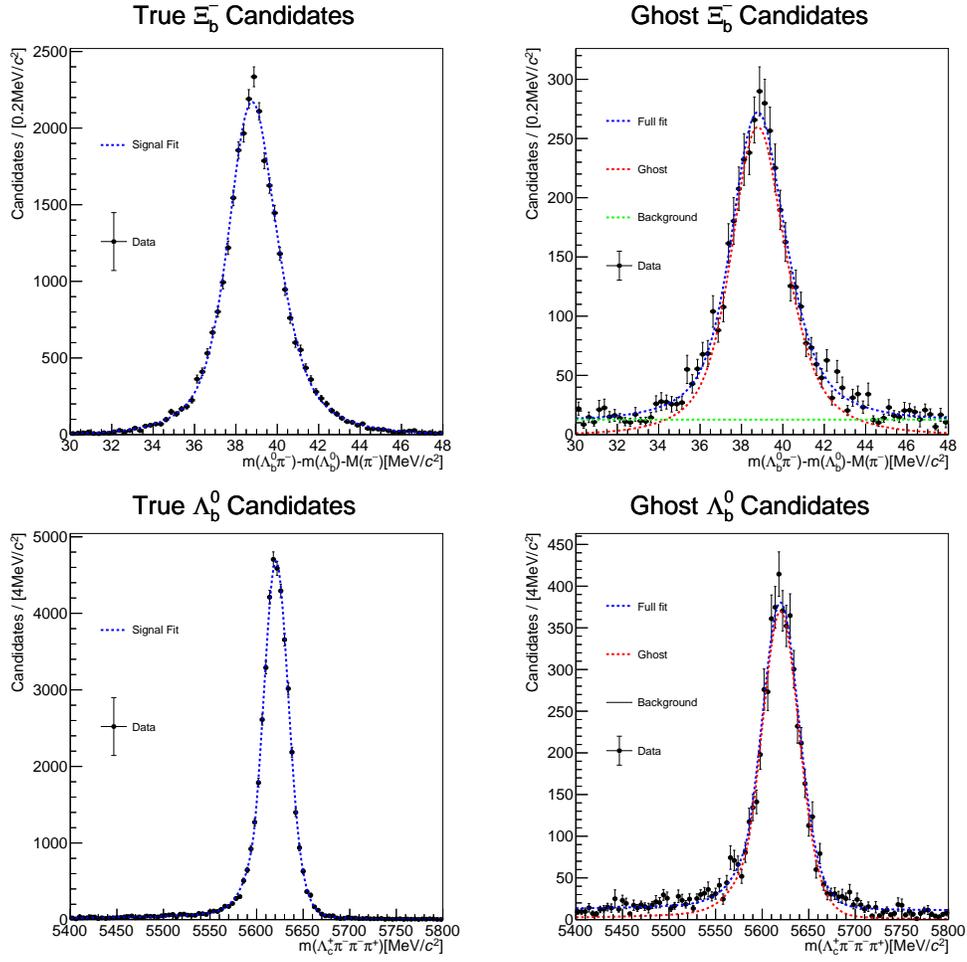


Figure 68: For mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ simulation samples, fitting results of the Ξ_b^- true signal (Top Left) and ghost (Top Right), Λ_b^0 true signal (Bottom Left) and ghost (Bottom Right).

K Full results of Λ_b^0 data fit

Run 2 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ fit result

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-1.02591e+07 FROM HESSE STATUS=OK 106 CALLS 566 TOTAL

EDM=0.00236738 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.30704e+02	3.21393e-01	1.07066e-02	2.34819e-02
2	argpar	-3.01631e+01	7.15814e-01	1.04000e-02	4.23383e-01
3	bExp	-8.56492e-03	9.08218e-05	9.20246e-04	-8.57543e-02
4	frbkg	5.84695e-01	7.54384e-03	1.16374e-02	1.70245e-01
5	frsig	9.81956e-01	7.76073e-04	8.25281e-03	1.30132e+00
6	mg	5.61984e+03	2.62195e-02	2.38902e-02	1.17996e-01
7	nBkg	4.00925e+05	7.76265e+02	1.45027e-03	-8.22305e-01
8	nSig	9.20897e+05	1.05931e+03	1.57027e-03	-3.96366e-01
9	sgbkg	1.42761e+01	2.89905e-01	3.71449e-02	1.42221e-01
10	zero	1.12905e+00	1.47732e-03	1.33548e-02	1.45743e-01

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5z

Run 2 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ fit result:

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-5.95668e+06 FROM HESSE STATUS=OK 100 CALLS 643 TOTAL

EDM=0.000563381 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.35747e+02	4.92852e-01	1.36169e-02	1.92768e-01
2	argpar	-2.49314e+01	6.76801e-01	5.00930e-03	5.42802e-01
3	bExp	-4.96115e-03	5.95836e-05	1.81478e-04	-4.96319e-02
4	frbkg	4.22740e-01	6.05101e-03	3.05120e-03	-1.55172e-01

5	frsig	9.96852e-01	1.03862e-03	1.05809e-02	1.45854e+00
6	mg	5.61957e+03	3.05237e-02	2.25835e-02	-1.49814e-02
7	nBkg	3.18713e+05	7.19102e+02	1.11961e-03	-9.06782e-01
8	nSig	5.11429e+05	8.42387e+02	1.12401e-03	-7.19551e-01
9	sgbkg	1.47327e+01	5.62713e-01	1.99636e-02	1.79244e-01
10	zero	1.11336e+00	2.05166e-03	5.91782e-03	6.68676e-02

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5

2016 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ fit result

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-2.73815e+06 FROM HESSE STATUS=OK 88 CALLS 650 TOTAL

EDM=0.000846882 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.30800e+02	5.38784e-01	3.99643e-04	2.66855e-02
2	argpar	-3.14789e+01	1.29538e+00	3.86568e-04	3.94412e-01
3	bExp	-8.85476e-03	1.60209e-04	3.47883e-05	-8.86637e-02
4	frbkg	5.81305e-01	1.32486e-02	4.35181e-04	1.63369e-01
5	frsig	9.82761e-01	1.46242e-03	3.15683e-04	1.30743e+00
6	mg	5.61969e+03	5.04543e-02	9.01905e-04	4.33850e-02
7	nBkg	1.26229e+05	4.35346e+02	2.84814e-05	-1.15761e+00
8	nSig	2.89145e+05	5.93665e+02	2.84117e-05	-9.39456e-01
9	sgbkg	1.52374e+01	4.98222e-01	1.36169e-03	2.20464e-01
10	zero	1.15463e+00	2.85976e-03	5.25555e-04	2.76685e-01

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5

2016 $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ fit result:

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-1.51537e+06 FROM HESSE STATUS=OK 106 CALLS 737 TOTAL

EDM=2.79714e-05 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.35947e+02	9.34646e-01	1.27503e-02	1.99560e-01
2	argpar	-2.47659e+01	1.27628e+00	4.65897e-03	5.46711e-01
3	bExp	-4.96947e-03	1.08436e-04	1.64889e-04	-4.97152e-02
4	frbkg	4.15904e-01	1.11550e-02	2.80138e-03	-1.68910e-01
5	frsig	9.95902e-01	2.01147e-03	7.02673e-03	1.44267e+00
6	mg	5.61947e+03	5.97018e-02	8.57225e-03	-6.59431e-02
7	nBkg	9.74624e+04	4.01400e+02	2.17366e-04	-1.20833e+00
8	nSig	1.56046e+05	4.68658e+02	5.30750e-04	-1.11061e+00
9	sgbkg	1.50607e+01	1.01705e+00	1.83489e-02	2.05994e-01
10	zero	1.13870e+00	4.12681e-03	5.82024e-03	1.94712e-01

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5

2017 $A_b^0 \rightarrow A_c^+ \pi^-$ fit result

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-2.79069e+06 FROM HESSE STATUS=OK 88 CALLS 632 TOTAL

EDM=0.000577941 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.30434e+02	6.59796e-01	3.92028e-04	1.44523e-02
2	argpar	-2.94597e+01	1.41474e+00	3.84148e-04	4.39025e-01
3	bExp	-8.43917e-03	1.70682e-04	3.39340e-05	-8.44921e-02
4	frbkg	5.87435e-01	1.42177e-02	4.28339e-04	1.75893e-01
5	frsig	9.82814e-01	1.39617e-03	3.08743e-04	1.30785e+00
6	mg	5.61952e+03	4.63044e-02	8.68638e-04	-3.97962e-02

7	nBkg	1.26636e+05	4.40097e+02	2.87901e-05	-1.15694e+00
8	nSig	2.94494e+05	6.01282e+02	2.87152e-05	-9.33438e-01
9	sgbkg	1.35823e+01	5.21403e-01	1.35820e-03	8.63219e-02
10	zero	1.12063e+00	2.60577e-03	4.84857e-04	1.03357e-01

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5

2017 $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$ fit result:

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-1.62011e+06 FROM HESSE STATUS=OK 100 CALLS 697 TOTAL

EDM=0.020271 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.35479e+02	8.10019e-01	1.22071e-02	1.83659e-01
2	argpar	-2.53079e+01	1.10840e+00	1.13914e-02	5.33943e-01
3	bExp	-5.00303e-03	9.51748e-05	4.19440e-04	-5.00512e-02
4	frbkg	4.23286e-01	9.08034e-03	6.99822e-03	-1.53950e-01
5	frsig	9.99590e-01	1.20943e-03	4.70715e-02	1.53031e+00
6	mg	5.61917e+03	4.85736e-02	2.10394e-02	-2.14635e-01
7	nBkg	1.02242e+05	3.85509e+02	5.61451e-04	-1.19945e+00
8	nSig	1.65412e+05	4.60149e+02	5.48862e-04	-1.09674e+00
9	sgbkg	1.40740e+01	9.65757e-01	4.53388e-02	1.25906e-01
10	zero	1.10494e+00	3.33988e-03	1.34362e-02	2.47201e-02

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5

2018 $A_b^0 \rightarrow A_c^+ \pi^-$ fit result

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-3.28165e+06 FROM HESSE STATUS=OK 106 CALLS 868 TOTAL

EDM=0.000613886 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.30937e+02	5.12449e-01	9.96478e-03	3.12536e-02
2	argpar	-2.99897e+01	1.10996e+00	9.68588e-03	4.27228e-01
3	bExp	-8.47287e-03	1.43053e-04	8.41205e-04	-8.48304e-02
4	frbkg	5.82426e-01	1.17361e-02	1.07316e-02	1.55422e+01
5	frsig	9.81261e-01	1.23859e-03	7.53998e-03	1.29615e+00
6	mg	5.62021e+03	4.17946e-02	2.29107e-02	3.10478e-01
7	nBkg	1.48078e+05	4.67598e+02	7.83107e-04	-1.12272e+00
8	nSig	3.37246e+05	6.38505e+02	7.84955e-04	-8.86983e-01
9	sgbkg	1.40454e+01	4.90313e-01	3.52807e-02	-6.15959e+00
10	zero	1.11596e+00	2.34875e-03	1.21451e-02	7.98765e-02

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5

2018 $A_b^0 \rightarrow A_c^+ \pi^- \pi^+ \pi^-$ fit result:

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-1.91228e+06 FROM HESSE STATUS=OK 88 CALLS 748 TOTAL

EDM=0.000191392 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	argoff	1.35846e+02	8.18915e-01	2.52687e-03	1.96116e-01
2	argpar	-2.48869e+01	1.11804e+00	4.60041e-04	5.43853e-01
3	bExp	-4.94075e-03	1.00202e-04	1.68863e-05	-4.94276e-02
4	frbkg	4.25863e-01	1.02797e-02	2.82144e-04	-1.48738e-01
5	frsig	9.96399e-01	1.76644e-03	7.43790e-04	1.45070e+00
6	mg	5.61997e+03	4.92505e-02	8.38214e-04	1.83852e-01
7	nBkg	1.19043e+05	4.44181e+02	2.44616e-05	-1.16971e+00
8	nSig	1.89946e+05	5.17877e+02	2.39692e-05	-1.06208e+00

9	sgbkg	1.49970e+01	9.32977e-01	1.86470e-03	2.00790e-01
10	zero	1.10242e+00	3.24028e-03	5.30854e-04	1.21190e-02

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 25 NPAR= 10 ERR DEF=0.5

L Full results of $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ data fit

This section presents the full fit result of $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ which is discussed in Section 6.2. Parameter names are explained as below:

- 1 or 3: parameters with suffix "1" are used for PDF fitted to data sample of $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ mode, and "3" for $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ mode;
- WS: suffix "WS" indicates this parameter is used in PDF for the wrong sign data samples;
- C1/C3: parameter C used in the combinatorial background PDF which is described in Eq. 9;
- mean.Sig: Mean value of Σ_b^- peak;
- mean.SigP: Mean value of Σ_b^+ peak;
- mean.Sig: Mean value of Σ_b^{*-} peak;
- mean.SigP: Mean value of Σ_b^{*+} peak;
- mg: Mean value of Ξ_b^- peak;
- nbw1/3: yield of Σ_b^- or Σ_b^+ (WS) peak;
- nbwS1/3: yield of Σ_b^{*-} or Σ_b^{*+} peak;
- ncomb1/3: yield of combinatorial background;
- nsig1/3: yield of Ξ_b^- peak;

Fit result with loose BDT requirement:

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-58359.1 FROM HESSE STATUS=OK 376 CALLS 1932 TOTAL

EDM=0.0954592 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	C1	1.36380e+01	1.48584e+00	2.38794e-02	2.20520e+00
2	C1_WS	1.19010e+01	6.18790e-01	2.49232e-02	2.55047e+00
3	C3	9.22834e+00	1.25679e+00	4.19748e-03	-7.04089e-01
4	C3_WS	5.33232e+00	9.88587e-01	2.50277e-02	-3.90983e-01
5	mean_Sig	5.85685e+01	4.17384e-01	5.70344e-03	3.19089e-01
6	mean_SigP	5.32428e+01	3.30053e-01	7.15756e-03	5.21216e-01
7	mean_SigS	7.70844e+01	4.17397e-01	4.01948e-03	3.02354e-01
8	mean_SigSP	7.34980e+01	3.56691e-01	4.71109e-03	9.97703e-02
9	mg	3.84476e+01	1.88732e-01	4.78490e-03	4.95334e-02
10	nbw1	3.51630e+02	3.82361e+01	3.15450e-03	-1.19353e+00
11	nbw1_ws	2.84380e+02	3.51573e+01	1.30363e-03	-1.23191e+00
12	nbw3	3.70507e+02	4.01573e+01	1.26728e-03	-1.18341e+00
13	nbw3_ws	4.41080e+02	3.84601e+01	1.15273e-03	-1.14761e+00
14	nbwS1	6.84439e+02	5.49371e+01	2.99945e-03	5.24178e+00
15	nbwS1_ws	6.61061e+02	4.99750e+01	2.93915e-03	5.23245e+00
16	nbwS3	7.60779e+02	5.68079e+01	2.96308e-03	-1.01191e+00
17	nbwS3_ws	8.27959e+02	5.31019e+01	2.80208e-03	-9.87058e-01
18	ncomb1	3.98943e+03	8.93111e+01	5.93925e-04	-1.16862e+00
19	ncomb1_ws	3.98108e+03	8.28172e+01	5.84158e-04	-1.16905e+00
20	ncomb3	4.22163e+03	9.30550e+01	5.97081e-04	-1.15692e+00
21	ncomb3_ws	4.07106e+03	8.48777e+01	5.94240e-04	-1.16447e+00
22	nsig1	1.08639e+02	1.83801e+01	3.80254e-03	-8.99027e-01
23	nsig3	1.45954e+02	2.06087e+01	3.75937e-03	-7.86793e-01

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 122 NPAR= 23 ERR DEF=0.5

Fit result with tight BDT requirement:

COVARIANCE MATRIX CALCULATED SUCCESSFULLY

FCN=-14385.4 FROM HESSE STATUS=OK 352 CALLS 1548 TOTAL
 EDM=0.000118279 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				INTERNAL	INTERNAL
NO.	NAME	VALUE	ERROR	STEP SIZE	VALUE
1	C1	1.28313e+01	2.22087e+00	1.41878e-03	7.61750e-01
2	C1_WS	1.40627e+01	2.15777e+00	2.12665e-03	2.09424e+00
3	C3	7.51366e+00	1.76558e+00	3.23042e-04	-7.98944e-01
4	C3_WS	6.81087e+00	1.58957e+00	3.74470e-03	-1.70704e-01
5	mean_Sig	5.87057e+01	7.89844e-01	5.42522e-04	3.48120e-01
6	mean_SigP	5.32073e+01	5.54460e-01	5.97810e-04	5.09575e-01
7	mean_SigS	7.83585e+01	7.62547e-01	3.95954e-04	5.00405e-01
8	mean_SigSP	7.35051e+01	6.37921e-01	2.09207e-03	1.01197e-01
9	mg	3.84707e+01	1.98304e-01	1.03827e-01	5.41603e-02
10	nbw1	1.07819e+02	2.30249e+01	6.64086e-05	-1.36275e+00
11	nbw1_ws	7.37241e+01	2.05509e+01	7.31234e-05	-1.39886e+00
12	nbw3	1.58625e+02	2.52574e+01	6.01881e-05	-1.82336e+00
13	nbw3_ws	1.82373e+02	2.40815e+01	5.48979e-05	-1.29988e+00
14	nbwS1	2.24945e+02	3.35844e+01	6.11747e-05	-1.26970e+00
15	nbwS1_ws	2.00569e+02	3.04726e+01	6.10480e-05	-1.28660e+00
16	nbwS3	3.03386e+02	3.51264e+01	5.66027e-05	-1.92094e+00
17	nbwS3_ws	3.16152e+02	3.27227e+01	2.69294e-04	-1.21328e+00
18	ncomb1	1.52440e+03	5.59169e+01	5.79431e-05	-1.32323e+00
19	ncomb1_ws	1.50073e+03	5.13990e+01	5.64667e-05	-1.32517e+00
20	ncomb3	1.62374e+03	5.82192e+01	5.90728e-05	-1.82634e+00
21	ncomb3_ws	1.50243e+03	5.18754e+01	5.84853e-05	-1.32503e+00
22	nsig1	7.87693e+01	1.28226e+01	7.71748e-04	-1.00184e+00
23	nsig3	9.92304e+01	1.46913e+01	7.89263e-04	-9.29865e-01

ERR DEF= 0.5

EXTERNAL ERROR MATRIX. NDIM= 122 NPAR= 23 ERR DEF=0.5

ELEMENTS ABOVE DIAGONAL ARE NOT PRINTED.

M Trigger efficiency

Fig. 69 shows the The N_{TIS} (Left), $N_{TOS|TIS}$ (Middle) and $\epsilon_{TOS|TIS}$ (Right) versus $\Xi_b^- p_T$ using $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ 2016 simulation sample.

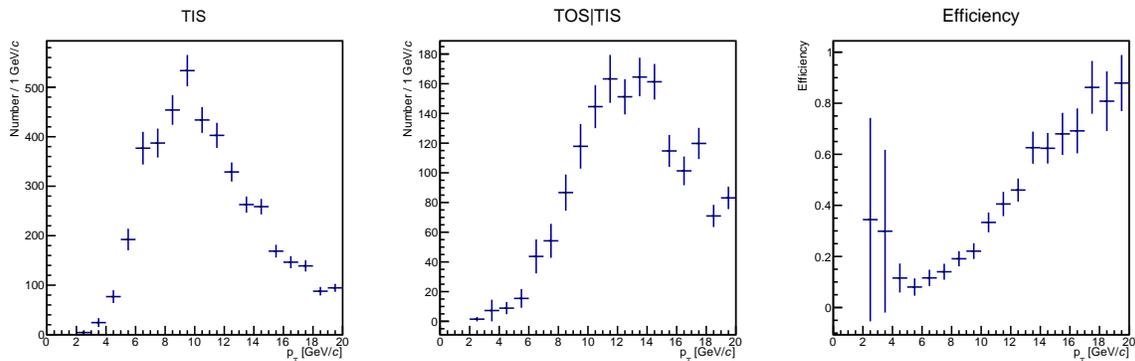


Figure 69: The N_{TIS} (Left), $N_{TOS|TIS}$ (Middle) and $\epsilon_{TOS|TIS}$ (Right) versus $\Xi_b^- p_T$ using $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ 2016 simulation sample

N Multiple candidates plots

Fig. 70 shows the histograms of the number of multiple candidates in one event in simulation and data samples for normalization mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ and signal mode $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$.

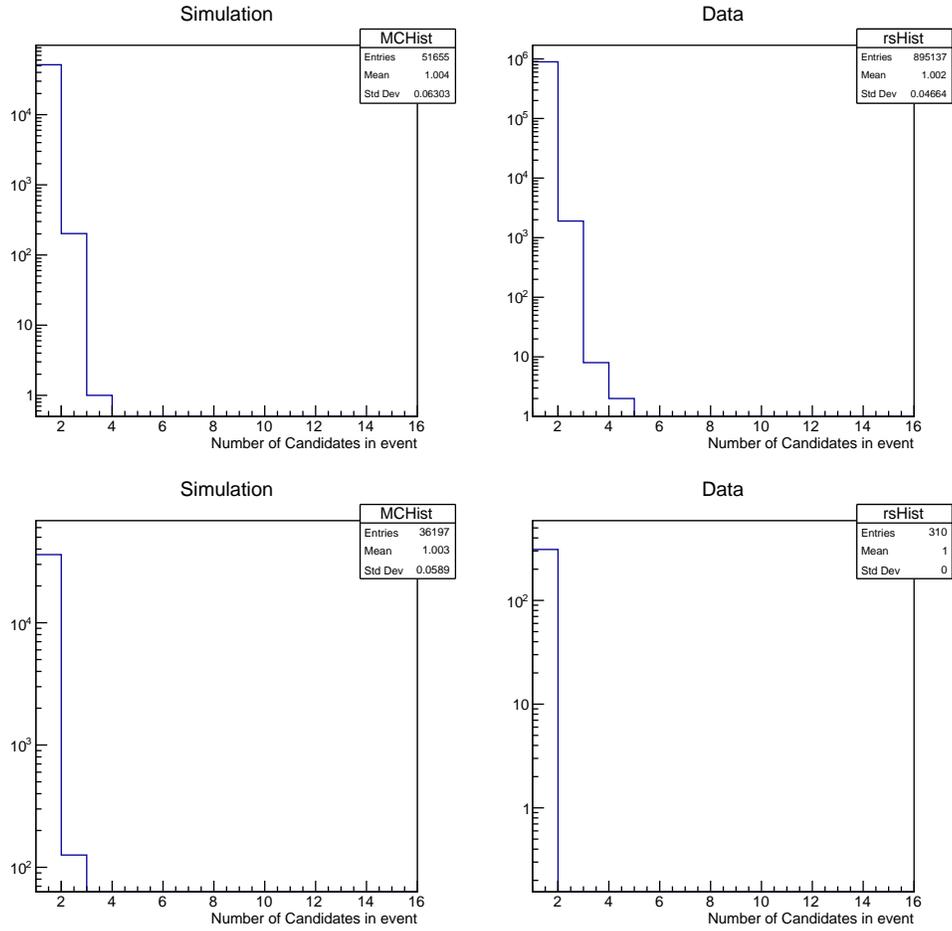


Figure 70: Histograms of the number of multiple candidates in one event in simulation and data samples for normalization mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ (Top) and signal mode $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$, $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^- \pi^+ \pi^-$ (Bottom).

References

- [1] CKMfitter group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, [arXiv:hep-ph/0406184](https://arxiv.org/abs/hep-ph/0406184), updated results and plots available at <http://ckmfitter.in2p3.fr/>.
- [2] D. Dominguez, *Illustrations of the new pentaquark and tetraquarks discovered by LHCb*, , General Photo, <https://cds.cern.ch/record/2814136>.
- [3] LHCb collaboration, R. C. M. Xavier Cid Vida, *LHC Layout*, , <https://bpt.web.cern.ch/lhc/statistics/2018/>.
- [4] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots*, , <https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb.ProductionAngles.html>.
- [5] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. LHCb-TDR-009.
- [6] T. L. Collaboration, *The LHCb detector at the LHC*, Journal of Instrumentation **3** (2008) S08005.
- [7] T. L. Collaboration, *Tracking System*, , <https://lhcb-public.web.cern.ch/en/detector/Trackers2-en.html>.
- [8] C. Abellán Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, tech. rep., Aug, 2020. All figures and tables, along with any supplementary material and additional information, are available at <http://lhcbproject.web.cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-DP-2020-001.html> (LHCb public pages).
- [9] LHCb Collaboration, R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC.*, JINST **14** (2018) P04013. 43 p, [arXiv:1812.10790](https://arxiv.org/abs/1812.10790), 46 pages, 35 figures, 1 table. All figures and tables are available at <https://cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-DP-2019-001.html>.
- [10] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the k_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.
- [11] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652.

- [12] A. Bettini, *Introduction to Elementary Particle Physics*, Cambridge University Press, 2 ed., 2014. The CKM matrix is introduced in Chapter 7.9, doi: 10.1017/CBO9781107279483.
- [13] I. I. Bigi, N. Uraltsev, and A. Vainshtein, *Nonperturbative corrections to inclusive beauty and charm decays. qcd versus phenomenological models*, Physics Letters B **293** (1992), no. 3-4 430.
- [14] A. Lenz, *Lifetimes and hqe*, arXiv: High Energy Physics - Phenomenology (2014).
- [15] Particle Data Group, P. A. Zyla *et al.*, *Review of Particle Physics*, PTEP **2020** (2020), no. 8 083C01, The mean lifetime of Λ^0 and Ξ_b^- can be found on PDG website <https://pdglive.lbl.gov/>.
- [16] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version*, Physics textbook, Wiley, New York, NY, 2008. Section 8.4.1 and 8.4.2.
- [17] R. Jaffe and F. Wilczek, *Diquarks and exotic spectroscopy*, Phys. Rev. Lett. **91** (2003) 232003.
- [18] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.
- [19] S. L. Olsen, T. Skwarnicki, and D. Zieminska, *Nonstandard heavy mesons and baryons: Experimental evidence*, Rev. Mod. Phys. **90** (2018) 015003.
- [20] LHCb Collaboration, *List of hadrons observed at the LHC*, , The update can be found in list <https://www.nikhef.nl/~pkoppenb/particles.html> organized by P. Koppenburg.
- [21] X. Li and M. B. Voloshin, *Decays $\Xi_b \rightarrow \Lambda_b \pi$ and diquark correlations in hyperons*, Phys. Rev. D **90** (2014) 033016.
- [22] LHCb collaboration, R. Aaij *et al.*, *Evidence for the strangeness-changing weak decay $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$* , Phys. Rev. Lett. **115** (2015) 241801 LHCb-PAPER-2015-047 CERN-PH-EP-2015-277, arXiv:1510.03829.
- [23] LHCb collaboration, R. Aaij *et al.*, *Measurement of the mass and production rate of Ξ_b^- baryons*, Phys. Rev. **D99** (2019) 052006 LHCb-PAPER-2018-047 CERN-EP-2018-348, arXiv:1901.07075.
- [24] CERN, CERN, *LHC 2018 Statistics*, , <https://bpt.web.cern.ch/lhc/statistics/2018/>.
- [25] I. Belyaev *et al.*, *The History of LHCb*, .
- [26] J. Andre *et al.*, *Status of the lhcb magnet system*, IEEE Transactions on Applied Superconductivity **12** (2002), no. 1 366.
- [27] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

- [28] L. Collaboration, *Lhcb detector performance*, International Journal of Modern Physics A **30** (2015), no. 07 1530022.
- [29] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](https://arxiv.org/abs/1412.6352).
- [30] C. Coca *et al.*, *The hadron calorimeter prototype beam-test results*, tech. rep., CERN, Geneva, Oct, 2000.
- [31] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010. LHCb-TDR-004.
- [32] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031. LHCb-TDR-010.
- [33] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020), no. 8 083C01.
- [34] PYTHIA collaboration, *PYTHIA*, , <https://pythia.org/>.
- [35] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to pythia 8.1*, Computer Physics Communications **178** (2008), no. 11 852.
- [36] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the mass and lifetime of the Ξ_b^0 baryon*, Phys. Rev. Lett. **113** (2014) 032001 LHCb-PAPER-2014-021 CERN-PH-EP-2014-107, [arXiv:1405.7223](https://arxiv.org/abs/1405.7223).
- [37] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the mass and lifetime of the Ξ_b^- baryon*, Phys. Rev. Lett. **113** (2014) 242002 LHCb-PAPER-2014-048 CERN-PH-EP-2014-226, [arXiv:1409.8568](https://arxiv.org/abs/1409.8568).
- [38] LHCb collaboration, *Charged PID*, , <https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/ChargedPID>.
- [39] LHCb collaboration, *PID Calib Package*, , <https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage>.
- [40] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.
- [41] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.
- [42] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis with ROOT*, PoS **ACAT** (2007) 040.

- [43] A. Hoecker *et al.*, *TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide.*, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [44] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, tech. rep., CERN, Geneva, Mar, 2017.
- [45] Wikipedia, *Crystal Ball Function*, , https://en.wikipedia.org/wiki/Crystal_Ball_function.
- [46] J. E. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'* , other thesis, 8, 1982. Crystal Ball function defined in page 178.
- [47] 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 LHCb-PAPER-2019-008 CERN-EP-2019-122, [arXiv:1906.08350](https://arxiv.org/abs/1906.08350).
- [48] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](https://arxiv.org/abs/physics/0402083).
- [49] T. L. John Back, Michal Kreps, *EvtGen Project*, , <https://evtgen.hepforge.org/>.
- [50] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.
- [51] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, tech. rep., CERN, Geneva, May, 2014.
- [52] T. L. collaboration, *Measurement of the track reconstruction efficiency at lhcb*, Journal of Instrumentation **10** (2015) P02007.
- [53] Wikipedia, *ARGUS distribution*, , https://en.wikipedia.org/wiki/ARGUS_distribution.
- [54] H. Albrecht *et al.*, *Search for hadronic $b \rightarrow u$ decays*, Physics Letters B **241** (1990), no. 2 278.
- [55] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo Cimento (1955-1965) **34** (1964), no. 6 1644.
- [56] LHCb collaboration, R. Aaij *et al.*, *Observation of two resonances in the $\Lambda_b^0 \pi^\pm$ systems and precise measurement of Σ_b^\pm and $\Sigma_b^{*\pm}$ properties*, Phys. Rev. Lett. **122** (2019) 012001 LHCb-PAPER-2018-032 CERN-EP-2018-243, [arXiv:1809.07752](https://arxiv.org/abs/1809.07752).
- [57] LHCb collaboration, R. Aaij *et al.*, *Observation of two new Ξ_b^- baryon resonances*, Phys. Rev. Lett. **114** (2015) 062004 CERN-PH-EP-2014-279 LHCb-PAPER-2014-061, [arXiv:1411.4849](https://arxiv.org/abs/1411.4849).

- [58] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

ZHUOMING LI

Personal Data

NAME	Zhuoming Li
ADDRESS	121 Lafayette Road, Apt.110, Syracuse, New York, 13205
PHONE	+1 315-706-7237
EMAIL	lizmapplication@gmail.com

Work Experience

06/2016–11/2022	Graduate Research Assistant at Syracuse University
08/2014–06/2016	Teaching Assistant at Syracuse University

Education

12/2022	Ph.D in Physics Syracuse University Thesis Title: Observation of the $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and measurement of $\mathcal{B}(\Xi_b^- \rightarrow \Lambda_b^0 \pi^-)$
06/2014	B.S. IN PHYSICS Wuhan University Thesis Title: The redshift and Hubble ratio measured by an off-center observer in a void

Research Experience and Interests

LHCb Collaboration at the Large Hadron Collider

Upstream Tracker (UT Upgrade) 2016–2022

- Planning and execution of UT electronics qualification examination, providing feedback to help the UT ASIC chip design.
- Execution of irradiation experiments to study the performance and stability of UT ASIC chips in real-time under the particle beam and after receiving the high dose of radiation.
- Execution of test beam experiments to study the tracking performance and noise level of UT detector units.
- Preparation and assembly of UT staves.
- Studies of B hadronic decays;

Publications

- Observation of $\Xi_b^- \rightarrow \Lambda_b^0 \pi^-$ decay and measurement of its branching fraction
Z. Li, S. Blusk, Ready to be submitted, 2022
- The SALT — Readout ASIC for Silicon Strip Sensors of Upstream Tracker in the Upgraded LHCb Experiment
Carlos Abellan Beteta, et al. **incl. Z. Li**, *Sensors*, 22(1):107, 2022,
doi:10.3390/s22010107
- First beam test of UT sensors with the SALT 3.0 readout ASIC
M. Artuso, et al. **incl. Z. Li**, Tech. Report, CERN, 2019, doi:10.2172/1568842

Conference Presentations

- Overall description of UT electronic components April APS Meeting 2021, Session B20: Tracking Systems for LHC Upgrades 17-20 April 2020