Measurement of prompt Λ_c^+ baryon production cross-section and Λ_c^+/D^0 production ratio in pPb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16\,\mathrm{TeV}$ in LHCb

Chenxi Gu[1], Giulia Manca[2], Yiheng Luo[3], Jiayin Sun[2], Jianqiao Wang[4], Di Yang[4], Xianglei Zhu[4]

[1] *Laboratoire Leprince-Ringuet, Palaiseau, France*
[2] *Peking University, Beijing, China*
[3] *Universita e INFN, Cagliari, Italy*
[4] *Tsinghua University, Beijing, China*

Abstract

Open heavy charmed baryon Λ_c^+ is a good probe to nuclear matter effects in heavy-ion collisions. Due to the large charm quark mass, perturbative QCD is applicable in calculating their production cross sections in the initial hard parton scatterings. With the data of proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16\,\mathrm{TeV}$ collected by LHCb during Run 2 in 2016, the double-differential production cross section of Λ_c^+ baryon is measured in the transverse momentum region of $2 - 15$ GeV/c, corresponding to a center-of-mass rapidity region of $1.5 < y^* < 4.0$ for forward configurations while $-5.0 < y^* < -2.5$ for backward configurations. Open charm baryon over meson ratio Λ_c^+/D^0 is presented to study hadronisation mechanism and compared with results measured by ALICE in proton-proton and proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ and results obtained by LHCb in proton-lead and peripheral lead-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$.

History

- 0.0, Feb. 8th 2024:

Contents

1 Introduction

To study the creation and property of the quark-gluon plasma (QGP) is one of the most important goals of high energy nuclear physics, i.e. heavy-ion collision physics. The QGP can be created and studied in laboratory through high energy nucleus-nucleus (AA) collisions. In central AA collisions, which are also called 'large system', lots of energy is deposited in a sizeable volume, where the vacuum is heated up and the phase transition from hadron gas to QGP is expected to occur. However in proton-proton and proton-nucleus collisions (*a.k.a.* 'small systems'), the system volume is too small, and the deconfined QGP matter is not supposed to be formed. These collisions are usually treated as the baseline for the studies of QGP in AA collisions.

With the formation of the QGP in the early stage of heavy-ion collisions, the production yields and momentum spectra of those final-state hadrons might be modified compared to the expectations based on the measurements in proton-proton collisions. These QGP effects are called hot nuclear matter (HNM) effects.

One of the HNM effects that can be studied is how the major feature of QGP, deconfinement of quarks and gluons, influences the hadronisation mechanism. Hadronisation, the process in which quarks/gluons are transformed into hadrons, are quite different depending on whether a QGP is formed or not. In principle, hadronisation should be calculated with first-principle QCD, however phenomenological functions are widely used due to the non-perturbative nature of hadronization. This is called fragmentation hadronisation, and these functions are applicable no matter a QGP is formed or not. When QGP is formed, quarks can also hadronise into hadrons through recombination with other quarks in QGP, and this is called recombination or coalescence. The recombination/coalescence explains the enhancement of the baryon-to-meson ratio at intermediate transverse momentum, and is not supposed to occur without the deconfined medium.

The so-called 'baryon enhancement' phenomena were first observed in anti-protons and pions in gold-gold collisions at RHIC in 2000s and inspired the proposal of coalescence hadronization mechanism. Thereafter similar phenomena are observed in strange and charm hadrons in AA collisions. However, the enhancement of Λ_c^+/D^0 ratio at intermediate transverse momentum, are also observed by ALICE experiment in proton-proton and proton-lead collisions at the nucleon-nucleon center-of-mass energy of 5.02TeV, where no QGP is expected to be formed. The data can't be reproduced by fragmentation models but are consistent with some coalescence models, and this is named as 'Λ_c^+ puzzle' and calls for deeper understanding and more precise measurement of hadronisation in proton-lead collisions. The Λ_c^+/D^0 in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV also is measured with LHCb, and also shows enhancement at intermediate p_{T}, which is described well by coalescence models.

Besides hot nuclear matter effects, the production yields and momentum spectra of final-state hadrons in heavy-ion collisions may also be affected by those mechanisms not related to the QGP, which are called cold nuclear matter effects (CNM). To determine whether QGP is formed in heavy-ion collisions, it is essential to distinguish cold nuclear matter effects and hot nuclear matter effects. These CNM effects include the nuclear modification of parton distribution function, which is called nuclear parton distribution function (nPDF), and rescatterings of final-state hadrons produced in heavy-ion collisions.

Cold nuclear matter effects can be studied through proton-nucleus (pA) collisions where the HNM effects are supposed to be non-dominant. These collisions, with smaller

1

size of systems compared to AA collisions and lager number of produced hadrons compared to proton-proton collisions, are ideal for the study of CNM effects. For example, the nPDFs can be probed or constrained by the production asymmetry of final-state particles between forward and backward rapidities.

Open heavy charmed baryon Λ_c^+ is a good probe to nuclear matter effects in heavy-ion collisions. Due to the large charm quark mass, perturbative QCD is applicable in calculating their production cross sections in the initial hard parton scatterings. The charm quarks produced in QGP thermally or produced by rescattering of final-state hadrons can be neglected since the medium temperature is much less than the charm quark mass.

The LHCb experiment provides unique measurements of Λ_c^+ production in proton-lead collisions in forward rapidities. Designed as a detector whose primary goals is to study heavy-flavour physics in hadron collisions, LHCb has more precise tracking system and particle identification system compared to other heavy-ion experiments, well suitable for precise charm hadron reconstructions. With its pseudo-rapidity coverage of $2.0 < \eta < 5.0$, very small Bjoken-x regions ($\sim 10^{-5}$) can be probed with open charm hadrons, significantly beyond the coverage of other heavy-ion experiments such as ALICE and STAR, giving unique information for constraining nPDFs.

2 Data sets and selections

2.1 Data sets

The data for this analysis was pPbdata collected by LHCb detector at late 2016, including two different configurations due to the asymmetry of collisions: forward (Fwd) collision(p beam coming from upstream of VELO) and backward (Bwd) collision(Pb beam coming from upstream of VELO). The configurations, corresponding to positive and negative rapidity regions respectively, are illustrated in the two pannels of Fig. 1: The integrated
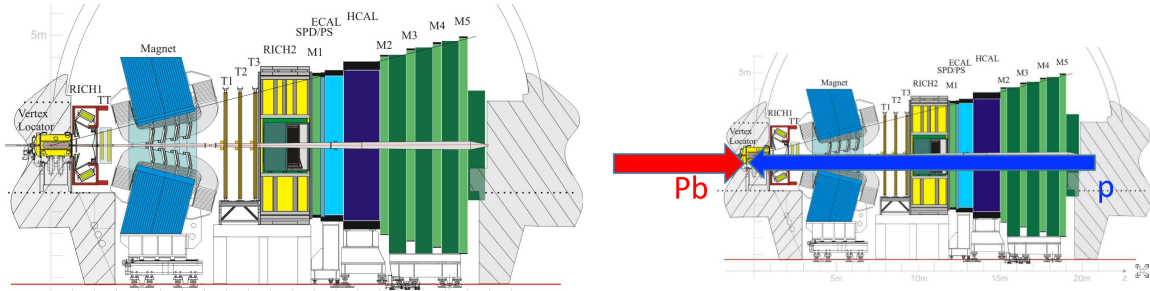

Figure 1: Skeches of the two beam configurations of pPb data taking, left for Fwd(pPb) and right for Bwd(Pbp).

luminosity for Fwd collision is 12.18 ± 0.32 nb^{-1} and 18.57 ± 0.46 nb^{-1} for Bwd. The centre-of-mass energy per nucleon pair $\sqrt{s_{\mathrm{NN}}}$ corresponds to 8.16 TeV. The bookkeeping

Table 1: Online (L0, HLT1) tigger on Λ_c^+ baryons

Quantity	Selection
	L0
nSPDHits	> 0
	HLT1TrackMVA
$\chi^2/\mathrm{ndf}(\mathrm{track})$	< 4.0
ProbNNghost(track)	< 0.3
$\chi^2_{\mathrm{IP}}(\mathrm{track})$ for $p_{\mathrm{T}} > 10.0\,\mathrm{GeV}/c$	> 6.0
$\chi^2_{\mathrm{IP}}(\mathrm{track})$ for $0.5 < p_{\mathrm{T}} < 10.0\,\mathrm{GeV}/c$	$> 6.0 \cdot \exp\left(\frac{0.3}{p_{\mathrm{T}}^2} + 0.2 \cdot \left(1 - \frac{p_{\mathrm{T}}}{10.0}\right) \right)$
	HLT1TwoTrackMVA
$p_{\mathrm{T}}(\mathrm{track})$	$> 0.3\,\mathrm{GeV}/c$
$p(\mathrm{track})$	$> 2\,\mathrm{GeV}/c$
$\chi^2/\mathrm{ndf}(\mathrm{track})$	< 4.0
$\chi^2_{\mathrm{IP}}(\mathrm{track})$	> 4.0
$\chi^2_{\mathrm{IP}}(\mathrm{vtx})$ for $2 < \eta(\mathrm{track}) < 5\,\mathrm{GeV}/c$	< 10
$\cos(\mathrm{DIRA})$ for $M_{\mathrm{corr}} > 0.5\,\mathrm{GeV}/c$	> 0.0

paths for Fwd and Bwd are:

/LHCb/Protonion16/Beam6500GeV-VeloClosed-MagDown/Real Data
/Turbo03pLead/94000000/TURBO.MDST;
/LHCb/Ionproton16/Beam6500GeV-VeloClosed-MagDown/Real Data
/Turbo03pLead/94000000/TURBO.MDST.

The data samples are triggered online in three levels. The first level (Level-0) is a hardware level trigger L0SPD for pPb data, which requires at least ones hit in the SPD detector. Thus it is an minmum biased trigger with $\varepsilon_{\mathrm{L0}} \approx 1$. The (software) high-level triggers HLT1 and HLT2 provide real time reconstruction and selection of tracks and particles. For hadron-final-state channels, such as $D^0 \to K^-\pi^+$ and $\Lambda_c^+ \to pK^-\pi^+$, the HLT1 triggers are Hlt1TrackMVADecision_TOS or Hlt1TwoTrackMVADecision_TOS==1, which require either one detached long track or two detached long tracks to originate from a common vertex in an event. The selections are cut-based and the explicit expressions are listed in Table ?? from Ξ_c^+ production [1] note (LHCb-ANA-2022-039). Hlt2 triggers serve as online selections on different particle, in order to save more signal cadidates as possible. The Λ_c^+ candidates are reconstructed by final-state p, K^- and π^+ tracks. The HLT2 selections for Λ_c^+ candidates in the HLT2 line Hlt2CharmHadLc2KPPi_XSecTurbo are shown in Table 2.

The simulation samples include $\sim$36M $\Lambda_c^+ \to pK^-\pi^+$ & c.c. decays for both rapidities, used for obtaining information of prompt Λ_c^+ baryon and for calculating efficiencies. The event type is 25103000 and the production tag is sim09c/j/k. For sim09k simulation samples (16M), the multiplicity distributions is fixed by increasing the number of pill-up.

Table 2: Hlt2 (Turbo) selection on Λ_c^+ baryons

Quantity	Selections
$p_{\text{T}}(\text{track})$	$> 200\,\text{MeV}/c$
$\chi^2_{\text{IP}}(\text{track})$	> 4
$\chi^2/\text{ndf}(\text{track})$	< 3
$p(\text{track})$	$> 1\,\text{GeV}/c$
$p(p)$	$> 10\,\text{GeV}/c$
$\text{DLL}_{K\pi}(K)$	> 5
$\text{DLL}_{K\pi}(\pi)$	< 5
$\text{DLL}_{p\pi}(p)$	> 5
$\text{DLL}_{pK}(p)$	> 5
$N(p_{\text{T}}(\text{track}) > 400\,\text{MeV}/c)$	≥ 2
$N(p_{\text{T}}(\text{track}) > 1000\,\text{MeV}/c)$	≥ 1
$N(\chi^2_{\text{IP}}(\text{track}) > 10)$	≥ 2
$N(\chi^2_{\text{IP}}(\text{track}) > 50)$	≥ 1
$m(\Lambda_c^+)$	$2210 < m(\Lambda_c^+) < 2543\,\text{MeV}/c^2$
DIRA	$< 34.6\text{mrad}$
$\chi^2/\text{ndf}(\text{vtx})$	< 25
Lifetime	$\tau > 0.075\text{ps}$

The paths are:

/MC/2016/pPb-Beam6500GeV-2560GeV-2016-MagDown-Fix1-Epos
/Sim09c(j,k)/Trig0x61421621/Reco16pLead/Turbo03/25103000/DST;
/MC/2016/Pbp-Beam2560GeV-6500GeV-2016-MagDown-Fix1-Epos
/Sim09c(j,k)/Trig0x61421621/Reco16pLead/Turbo03/25103000/DST.

The Λ_c^+ baryon from b hadrons are simulated by 2M $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ for both configurations. The paths are:

/MC/2016/pPb-Beam6500GeV-2560GeV-2016-MagDown-Fix1-Epos
/Sim09c(h)/Trig0x61421621/Reco16pLead/Turbo03/12163001/DST;
/MC/2016/Pbp-Beam2560GeV-6500GeV-2016-MagDown-Fix1-Epos
/Sim09c(h)/Trig0x61421621/Reco16pLead/Turbo03/12163001/DST.

2.2 Offline selections

To further improve the signal purity of the samples, tighter selections are applied offline as listed in Table 3, mainly following the analysis of Λ_c^+ production at 5.02 TeV in pPb [2]. The transeverse momentum lower limit on 400 MeV/c, pseudo-rapidity region of $2-5$ and the selection of $\text{ProbNN}_{\text{ghost}}(\text{track}) < 0.3$ are introduced to improve the final-track quality. The direction angle and χ^2/ndf of vertex fit cuts help to obtain better Λ_c^+ decay vertices. Since the lifetime of Λ_c^+ baryon is approximately 0.2 ps, the vertex displacement (VD) and decay time cuts are applied to exclude some of the background and Λ_c^+ baryons from beauty decays. More tight PID cuts are applied to reduce misID backgrounds. A signal

4

Table 3: Offline selections on Λ_c^+ baryons

Quantity	Selections
$p_{\mathrm{T}}(\mathrm{track})$	$> 400\,\mathrm{MeV}/c$
$\eta(\mathrm{track})$	$2 < \eta < 5$
$\mathrm{ProbNNGhost}(\mathrm{track})$	< 0.3
$p(\mathrm{track})$	$3.2 < p < 100\,\mathrm{GeV}/c$
$\mathrm{DLL}_{K\pi}(\pi)$	< 0
$\mathrm{DLL}_{pK}(p)$	> 15
$\cos(\mathrm{DIRA})$	> 0.99975
$\chi^2_{\mathrm{VD}}(\Lambda_c^+)$	> 50
$\chi^2/\mathrm{ndf}(\mathrm{vtx})$	< 6
$m(\Lambda_c^+)$	$2237 < m(\Lambda_c^+) < 2337\,\mathrm{MeV}/c^2$
$\log_{10}\chi^2_{\mathrm{IP}}(\Lambda_c^+)$	$-5 < \log_{10}\chi^2_{\mathrm{IP}}(\Lambda_c^+) < 5$
Lifetime	$0.2 < \tau < 1.2\mathrm{ps}$

window of invariant-mass varied by $50\,\mathrm{MeV}/c^2$ around $2287\,\mathrm{MeV}/c^2$ and $\log_{10}\chi^2_{\mathrm{IP}}$ $[-5, 5]$ for Λ_c^+ is set for convenience of mass and $\log_{10}\chi^2_{\mathrm{IP}}$ fit. An extra selection on track momemtum is introduced due to the fiducial regions of tracking and PID calibration tables.

3 Analysis strategy

For this analysis, the key measurement is the Λ_c^+ production cross-section as a function $p_{\mathrm{T}}(\Lambda_c^+)$ and $y^*(\Lambda_c^+)$, where y^* is the rapidity defined in the nucleon-nucleon centre-of-mass frame. Here prompt Λ_c^+ baryons refer to those produced directly from collisions or from the strong decay of other charm hadrons (e.g. $\Sigma_c^{++} \to \Lambda_c^+ \pi^+$). Whereas non-prompt Λ_c^+ baryon, or Λ_c^+-from-b, refer to those from the weak decay of B hadrons (e.g. $\Lambda_b^0 \to \Lambda_c^+ \pi^-$). The centre-of-mass frame does not coincide with laboratory frame due to the asymmetry of the collision, so y^* is shifted by a constant value with respect to the rapidity in the laboratory frame:

$$y^* = y - \delta y \ , \tag{1}$$

where $\delta y = 0.5 \log(A_{\mathrm{Pb}}/Z_{\mathrm{Pb}}) = 0.465$. The direction of proton beam is defined as the positive z-axis.

the double differential cross-section is defined as:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_{\mathrm{T}}\mathrm{d}y^*} = \frac{N(\Lambda_c^+ \to pK^-\pi^+)}{\mathcal{L} \times \epsilon_{\mathrm{tot}} \times \mathcal{B}(\Lambda_c^+ \to pK^-\pi^+) \times \Delta p_{\mathrm{T}} \times \Delta y^*} \ , \tag{2}$$

- $N(\Lambda_c^+ \to K^\mp\pi^\pm)$ is the prompt Λ_c^+ signal candidates reconstructed through $\Lambda_c^+ \to K^\mp\pi^\pm$ decay channels, including their charge conjugate channels. It is measured in Section 4.

- $\mathcal{L}$ is the integrated luminosity, for Fwd $\mathcal{L}_{\mathrm{Fwd}} = 0.01218 \pm 0.00032\,\mathrm{pb}^{-1}$, for Bwd $\mathcal{L}_{\mathrm{Bwd}} = 0.01857 \pm 0.00046\,\mathrm{pb}^{-1}$, which is determined as the Ref. [?] described.

- ϵ_{tot} is the total efficiency in each (p_{T}, y^*) bin, evaluated in Section 5.

- $\mathcal{B}(\Lambda_c^+ \to K^\mp \pi^\pm) = (6.26 \pm 0.29)\%$ is the branching fraction of decay $\Lambda_c^+ \to pK^-\pi^+$, obtained from PDG 2022 [3].

- Δp_T is the bin width of the Λ_c^+ transverse momentum, with a p_T range of $[0, 15]$ GeV/c.

- $\Delta y^* = 0.5$ is the bin width of the Λ_c^+ rapidity, for Fwd $1.5 < y^* < 4.0$, for Bwd $-5.0 < y^* < -2.5$.

Then the forward-backward production ratio can be derived from the cross-section as

$$R_{\mathrm{FB}}(p_T, y^*) \equiv \frac{\mathrm{d}^2\sigma_{p\mathrm{Pb}}(p_T, +|y^*|)/\mathrm{d}p_T\mathrm{d}y^*}{\mathrm{d}^2\sigma_{\mathrm{Pb}p}(p_T, -|y^*|)/\mathrm{d}p_T\mathrm{d}y^*} , \tag{3}$$

which is calculated among common rapidity bins $2.5 < |y^*| < 4.0$. By comparing the production cross-section with that D^0 meson, the Λ_c^+/D^0 production ratio $R_{\Lambda_c^+/D^0}$ can be given as

$$R_{\Lambda_c^+/D^0} \equiv \frac{\sigma_{\Lambda_c^+}}{\sigma_{D^0}} . \tag{4}$$

$R_{\Lambda_c^+/D^0}$ can also be measured as functions of (p_T, y^*), as well as multiplicity variables such as $N_{\mathrm{PV}}^{\mathrm{tracks}}$ and $N_{\mathrm{VELO}}^{\mathrm{Clusters}}$, in order to study the system size dependence of hadronisation processes.

4 Signal yield

Two steps of fit are performed to determine the prompt signal yield in each kinematic intervals, which are both extended unbinned maximum likelihood fit. The first step is performed on the invariant mass $M(pK^-\pi^+)$ distribution in the selected signal window. Following previous analyses [4, 5], a Crystal Ball (CB) function [6] plusk a Gaussian function is used to describe the signal shape, where

$$f \times \mathrm{CB} + (1 - f) \times \mathrm{Guass} , \tag{5}$$

$$f_{\mathrm{CB}}(x; M, \sigma, \alpha, n) = \begin{cases} \dfrac{\left(\frac{n}{|\alpha|}\right)^n e^{-\frac{1}{2}\alpha^2}}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-M}{\sigma}\right)^n}, & \text{if } \frac{x-M}{\sigma} < -|\alpha| , \\ \exp\left(-\frac{1}{2}\left(\frac{x-M}{\sigma}\right)^2\right), & \text{if } \frac{x-M}{\sigma} \geq -|\alpha| . \end{cases} \tag{6}$$

A linear function is used to describe the shape of background. In this PDF, n of CB function is always fixed to 1 from physical constraint, and CB and Gaussian function share a common mean value. Due to limited statistics in some kinematic intervals, fit results will be unstable or hard to converge if all parameters. Thus, a global fit is performed to fix α in CB, the ratio between width of Gaussian over that of CB r_{width}, and the fraction of the CB function f. To ensure the stability of the fits, the mass fits are repeated with 50 times, varying the initial values of all parameters randomly. The fit result with the smallest minimal negative likelihood (NLL) is chosen to be the final result. Such *best* results for two rapidities are shown in Fig. 2.

This invariant mass fit is performed in the signal window around Λ_c^+ PDG mass $M(\Lambda_c^+) \pm 50\,\mathrm{MeV}$ as listed in Table 3. Here, *best* fit results are also selected following the
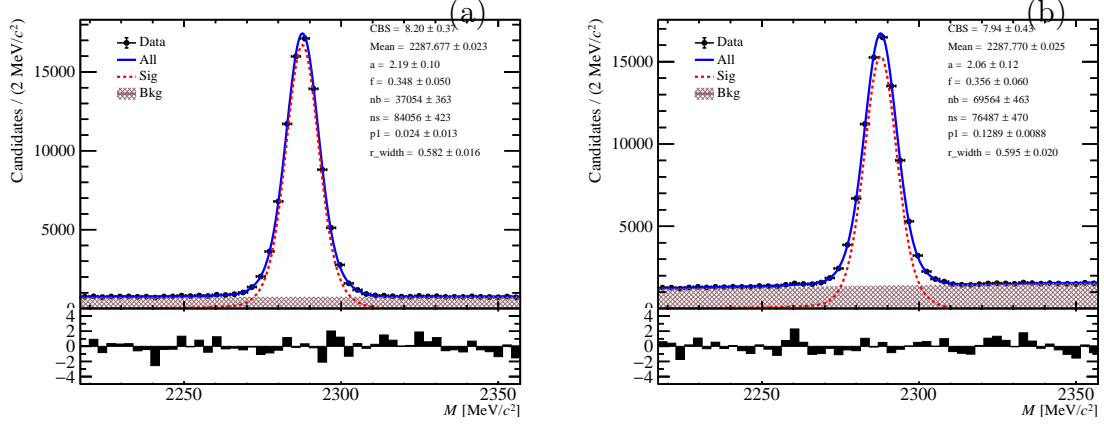
Figure 2: Fit on kinematic-unbinned $M(K\pi)$ distribution for Fwd (left) and Bwd (right) rapidities.
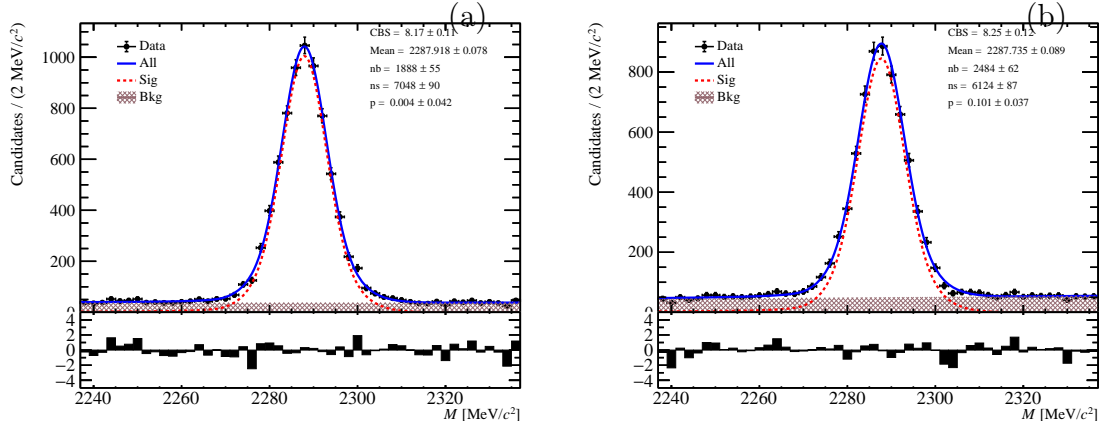

Figure 3: Fit on $M(K\pi)$ distribution for Fwd ($4 < p_{\mathrm{T}}(\Lambda_c^+) < 5\,\mathrm{GeV}/c$ and $2.5 < y^* < 3.0$, left) and Bwd ($4 < p_{\mathrm{T}}(\Lambda_c^+) < 5\,\mathrm{GeV}/c$ and $-4.0 < y^* < -3.5$, right) rapidities.

method above. Two examples for Fwd and Bwd rapidities are shown in Fig. 3, and all results are summarized in Appendix **??**.

A second step of fit on the $\log_{10}\chi^2_{\mathrm{IP}}$ (Λ_c^+) is performed to get prompt yields from total yields, following previous analyses [2, 5] Here χ^2_{IP} is defined as the difference in the vertex-fit χ^2 of a given PV reconstructed with and without the candidate under consideration. which is approximate to the significance of IP significance $\mathrm{IP}/\sigma(\mathrm{IP})$. So non-prompt Λ_c^+ baryons have larger IP than prompt ones due to the decay length of B hadrons, and the two different components can be distinguished with this parameter. To suppress the background component, a *sPlot* method [**?**] is performed using the fit result from $M(K\pi)$ fit. So the $\log_{10}\chi^2_{\mathrm{IP}}$ distribution of weighted data contains only prompt Λ_c^+ component and non-prompt component. The PDF describing the shapes is a Bukin
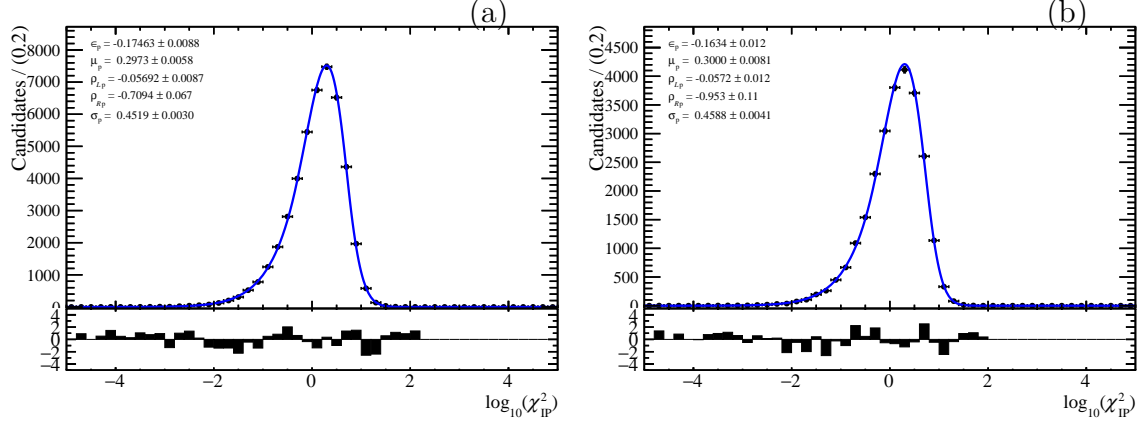
Figure 4: Global fit on $\log_{10} \chi^2_{\mathrm{IP}}$ distribution of prompt Λ_c^+ simulation for Fwd (left) and Bwd (right) rapidities.
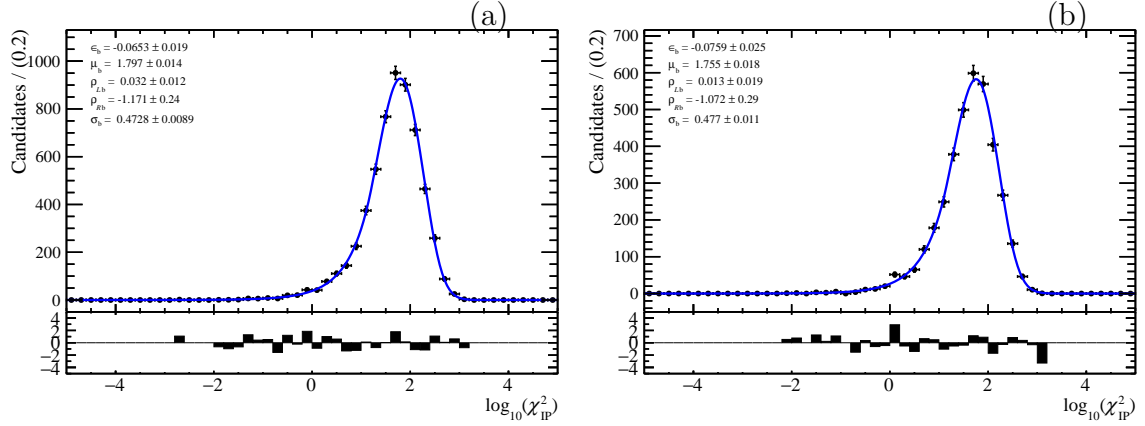

Figure 5: Global fit on $\log_{10} \chi^2_{\mathrm{IP}}$ distribution of from-b Λ_c^+ simulation for Fwd (left) and Bwd (right) rapidities.

function [7] as follows:

$$
\mathcal{P}(x; \mu, \sigma, \epsilon, \rho_L, \rho_R) = \begin{cases} \exp\left\{ \dfrac{(x-x_1)\epsilon\sqrt{\epsilon^2+1}\sqrt{2\ln 2}}{\sigma\left(\sqrt{\epsilon^2+1}-\epsilon\right)^2 \ln\left(\sqrt{\epsilon^2+1}+\epsilon\right)} + \rho_L\left(\dfrac{x-x_1}{\mu-x_1}\right)^2 - \ln 2 \right\} & x \leq x_1, \\[2em] \exp\left\{ -\left[\dfrac{\ln\left(1+2\epsilon\sqrt{\epsilon^2+1}\,\frac{x-\mu}{\sigma\sqrt{2\ln 2}}\right)}{\ln\left(1+2\epsilon^2-2\epsilon\sqrt{\epsilon^2+1}\right)}\right]^2 \times \ln 2 \right\} & x_1 < x < x_2, \\[2em] \exp\left\{ \dfrac{(x-x_2)\epsilon\sqrt{\epsilon^2+1}\sqrt{2\ln 2}}{\sigma\left(\sqrt{\epsilon^2+1}-\epsilon\right)^2 \ln\left(\sqrt{\epsilon^2+1}+\epsilon\right)} + \rho_R\left(\dfrac{x-x_2}{\mu-x_2}\right)^2 - \ln 2 \right\} & x \geq x_2. \end{cases}
$$
(7)

$$
x_1 = \mu + \sigma\sqrt{2\ln 2}\left(\frac{\epsilon}{\sqrt{\epsilon^2+1}} - 1\right), x_2 = \mu + \sigma\sqrt{2\ln 2}\left(\frac{\epsilon}{\sqrt{\epsilon^2+1}} + 1\right).
$$

It is a asymmetric gaussian function, with ϵ discribing the asymmetry and two ρs discrbing the left and right tail length. For both configurations, ρ_L and ρ_R of prompt component are fixed while ϵ, ρ_L and ρ_R of non-prompt component are fixed, all using simulation results of from-b ($\Lambda_b^0 \to \Lambda_c^+ \pi^-$) simulation fit as shown Fig. 4 and 5.

From the fit of second step, the prompt Λ_c^+ yield yield in signal window $[1815, 1915]\,\mathrm{MeV}/c^2 \times [-5, 5]$ can be directly obtained, as summarized in Fig. 6. There

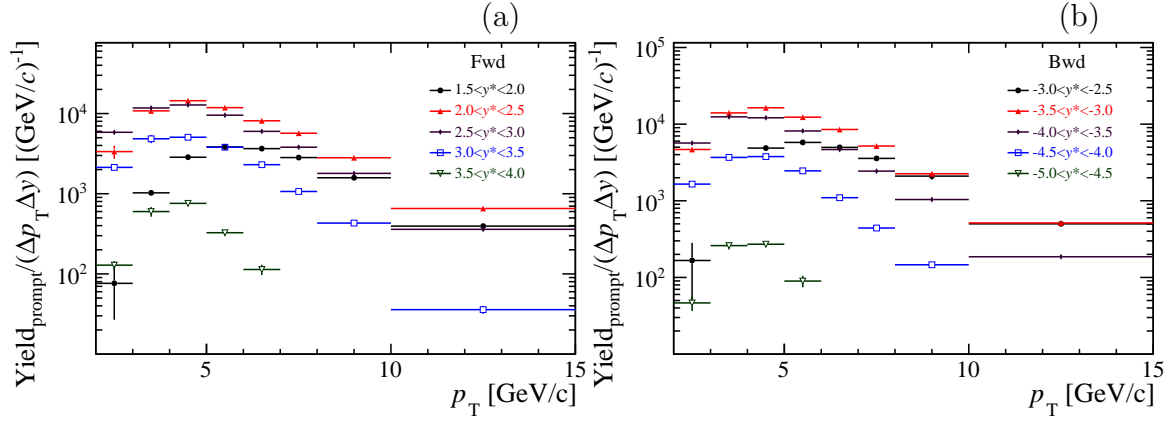Figure 6: Prompt yields obtained from $\log_{10} \chi_{IP}^2$ fit for Fwd (left) and Bwd (right), statistical uncertainties only.
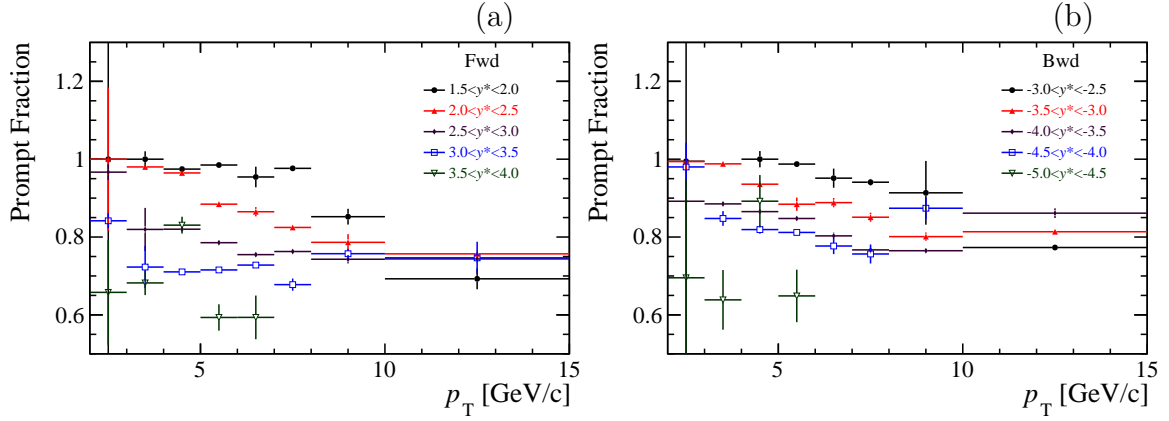

Figure 7: The fraction of prompt Λ_c^+ component from fit for Fwd (left) and Bwd (right), statistical uncertainties only.

175 may be some Λ_c^+ signals outside the window so its effect is considered in the evaluation of
176 selection efficiencies. The fraction of prompt component is also given from as Fig. 7. It
177 is however not multiplied with raw yield because their correlation should be considered
178 while calculating the uncertainty of prompt yield, which is of more difficulties.

9

5 Efficiencies

In this analysis, total efficiency includes four parts, the geometrical acceptance efficiency $\varepsilon_{\mathrm{acc}}$, the reconstruction and selection efficiency $\varepsilon_{\mathrm{rec\&sel}}$ (without PID efficiency), the PID efficiency $\varepsilon_{\mathrm{PID}}$ and trigger efficiency $\varepsilon_{\mathrm{trigger}}$. The total efficiency is obtained by multiplying these four parts:

$$\varepsilon_{\mathrm{tot}} = \varepsilon_{\mathrm{acc}} \times \varepsilon_{\mathrm{rec\&sel}} \times \varepsilon_{\mathrm{PID}} \times \varepsilon_{\mathrm{trigger}} \ . \tag{8}$$

The four efficiencies are measured in each kinematic bin. The steps are introduced in the next sections separately.

5.1 Corrections

To have a more precise measurement of efficiencies, it is necessary to correct the difference between data and simulation samples. The main differences lie between the PID, multiplicity variables and kinematic variables, originating from the poor decription of the corresponding detectors. The Dalitz plots are also different, as the Λ_c^+ MC generator is a phase space model, while there are three intermediate states $\Lambda(1520), \overline{K}^*(892)^0, \Delta(1232)^{++}$ in the $\Lambda_c^+ \to pK^-\pi^+$ decay

For the PID efficiency, data diven methods will be employed instead, which will be discussed in the **PID efficiency** section. For other variables, a k-fold Gradient Boosting (GB) [8] reweight method is employed to weight these variables simultaneously. Since different species of variables are independent, $(N_{\mathrm{VELO}}^{\mathrm{Clusters}}, N_{\mathrm{Long}}^{\mathrm{Tracks}}, N_{\mathrm{TT}}^{\mathrm{Clusters}})$, $(p_{\mathrm{T}}(\Lambda_c^+), y^*(\Lambda_c^+))$ and invariant masses of intermediate states are weighted to sWeighted data separately. and final weight are obtained by multiplying these weights. One-dimesional plots in this correction for multiplicity and kinematic variables are shown in Fig. 8 and 9 respectively. The Dalitz plots are also shown in Fig 10. Throughout the efficiency calculation, the weights are applied for both full and generator-level simulation samples.

5.2 Geometrical acceptance efficiency

The $\varepsilon_{\mathrm{acc}}$ is given by Equation 5.2

$$\varepsilon_{\mathrm{acc}} \equiv \frac{\Lambda_c^+ \text{ with } pK^-\pi^+ \text{ in LHCb acceptance}}{\text{Generated } \Lambda_c^+} .$$

In this equation the acceptance refers to a range of polar angle θ of $[10, 400]$ mrad with respect to the beam direction. So a generator-level simulation sample without geometrical acceptance requirement is used to measure $\varepsilon_{\mathrm{acc}}$ for both forward and backward rapidities. The $\varepsilon_{\mathrm{acc}}$ as a function of p_{T} and y^* for is shown in Fig. 11. The results are summarized in Table **??** and **??** in Appendix **??**.

5.3 Reconstruction and selection efficiency

The reconstruction and selection efficiency is defined as

$$\varepsilon_{\mathrm{rec\&sel}} = \frac{\sum \Lambda_c^+ \text{in acceptance, reconstructed and selected}}{\Lambda_c^+ \text{ with } pK^-\pi^+ \text{ in LHCb acceptance}} \ . \tag{9}$$
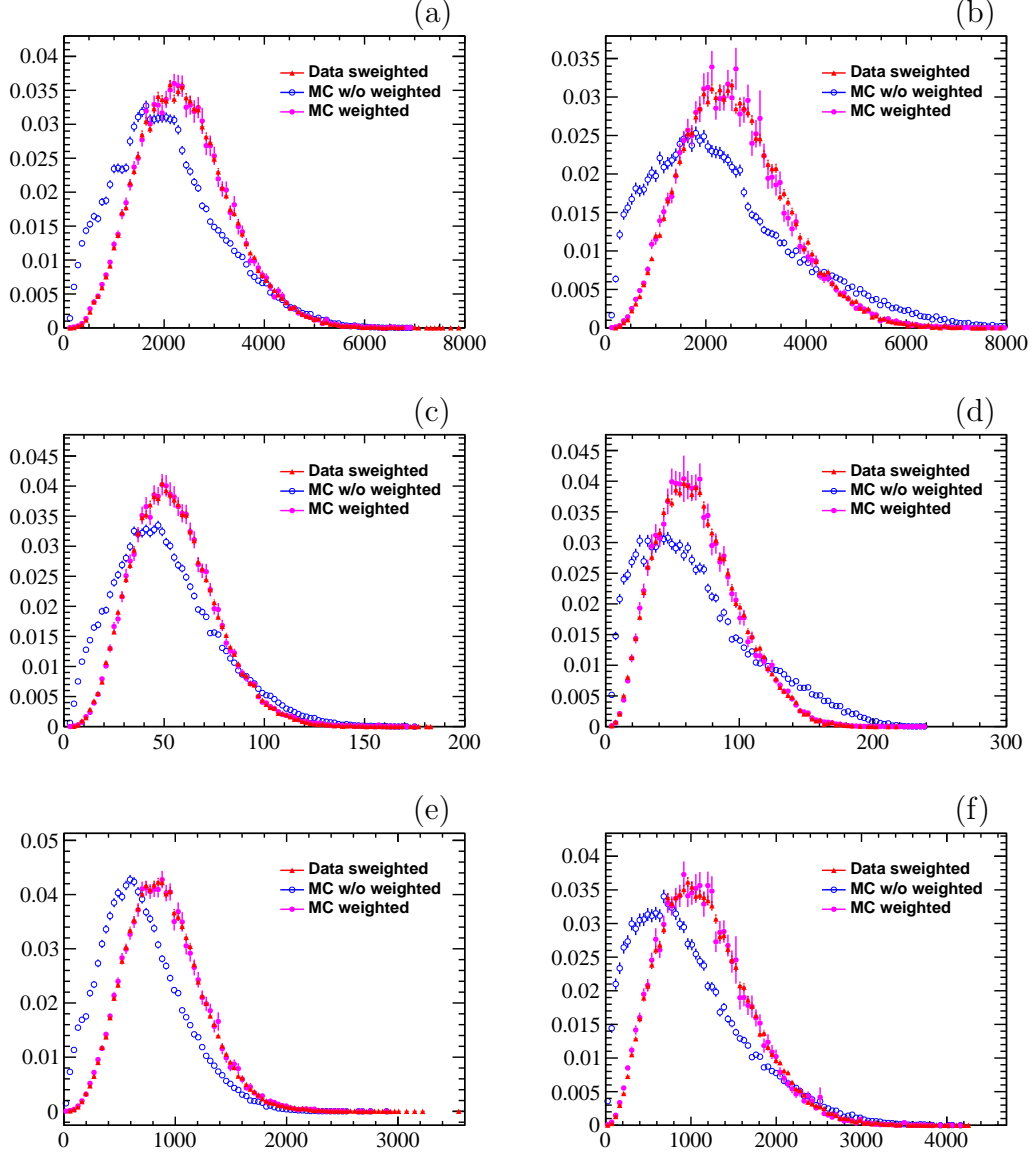
10

Figure 8: Distributions of multiplicity variables $N_{\text{VELO}}^{\text{Clusters}}$ (top), $N_{\text{Long}}^{\text{Tracks}}$ (middle) and $N_{\text{TT}}^{\text{Clusters}}$ (bottom) for orignal MC, corrected MC and sWeighted data in (left) forward and (right) backward regions.

This efficiency include two parts: the efficiency of reconstructing the three long tracks and the refinement of the Λ_c^+ signals. The selections are listed in Tables 2 and 3 without PID requirements. As the HLT1 selections for tracks in Table 1 are also cut-based, so they are also included in this section. The sample for calculating this efficiency is truth matched $\Lambda_c^+ \to pK^-\pi^+$ decays in pPbfull simulation sample. The *truth matching* requires that particle IDs are in agreement with there PDG IDs and the background category (BKGCAT) equals to 0 (signal) or 50 (low-mass background). There are two extra corrections to be considered in $\varepsilon_{\text{rec\&sel}}$calculation, which are truth matching inefficiency and track finding efficiency.
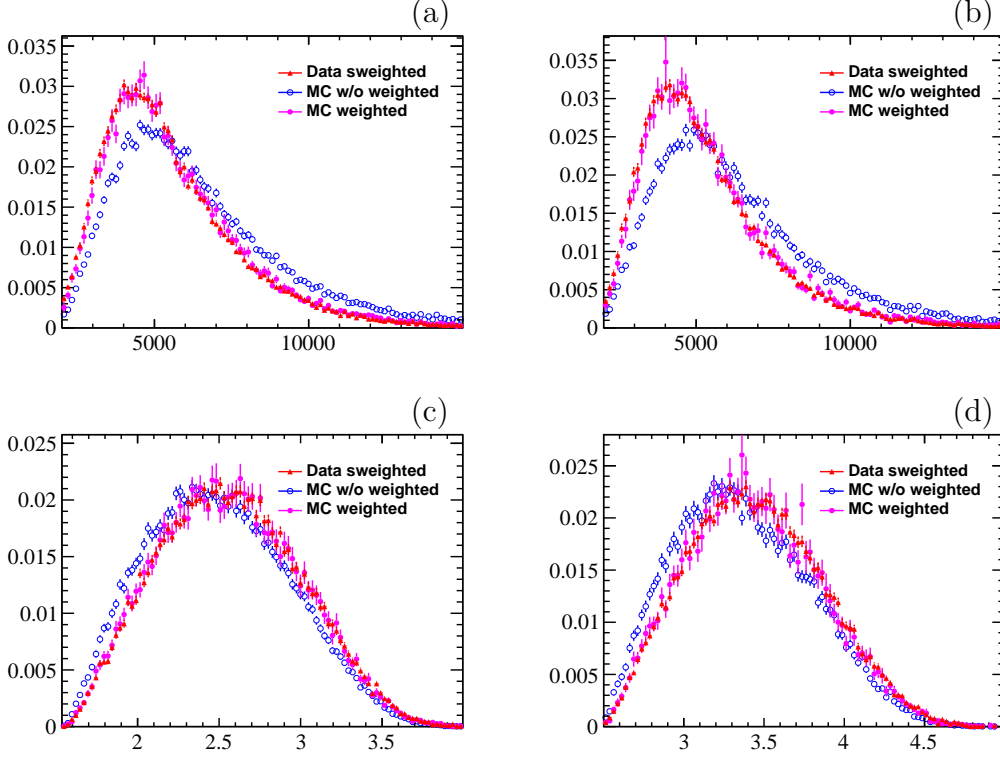
Figure 9: Distributions of multiplicity variables $p_{\mathrm{T}}(\varLambda_c^+)$ (top) and $y^*(\varLambda_c^+)$ (bottom) for orignal MC, corrected MC and sWeighted data in (left) forward and (right) backward regions.

5.3.1 Truth matching inefficiency

The signals in the simulation are picked out by truth matching requirements, but the truth matching algorithm occasionally flags the signal track as a ghost. This effect can be seen by plotting the $M(pK^-\pi^+)$ distribution of not truth matched passing the selections. A peak around $\varLambda_c^+$ mass can be seen from the blue points in Fig. 12. The ratio of the number of $\varLambda_c^+$ signal in this peak over the total number (truth matched) of $\varLambda_c^+$ signal. A Gaussian signal plus a linear background is used to fit the mass spectrum. The fits give a global fraction of 0.01% for both rapidities. The effect would result in an underestimate of the efficiency, thus the total efficiency should be multiplied by a factor of 1.0001 for Fwd and 1.0001 for Bwd. As the fraction is quite small, uncertainties and kinematic dependence are also negligible.

5.3.2 Tracking correction

The track finding efficiency for data and simulation is different. Calibrations are performed using a tag-and-probe method [?], with $J/\psi \to \mu^+\mu^-$ and $K_{\mathrm{S}}^0 \to \pi^+\pi^-$ decays. The tables are the same as that from the $\varLambda_c^+$ production in pPb at $\sqrt{s_{\mathrm{NN}}} = 8.16\,\mathrm{TeV}$ [5](LHCb-ANA-2019-039), given in histograms of track momemtum p and pseudo-rapidity η. When performing this correction, the weights are applied to selected candidates only, defined as

$$\varepsilon_{\mathrm{rec\&sel}} = \frac{\sum_{\varLambda_c^+ \text{ reconstructed and selected}} w_i(p_p, \eta_p) \times w_i(p_{K^-}, \eta_{K^-}) \times w_i(p_{\pi^+}, \eta_{\pi^+})}{\varLambda_c^+ \text{ with } pK^-\pi^+ \text{ in LHCb acceptance}} \ . \tag{10}$$
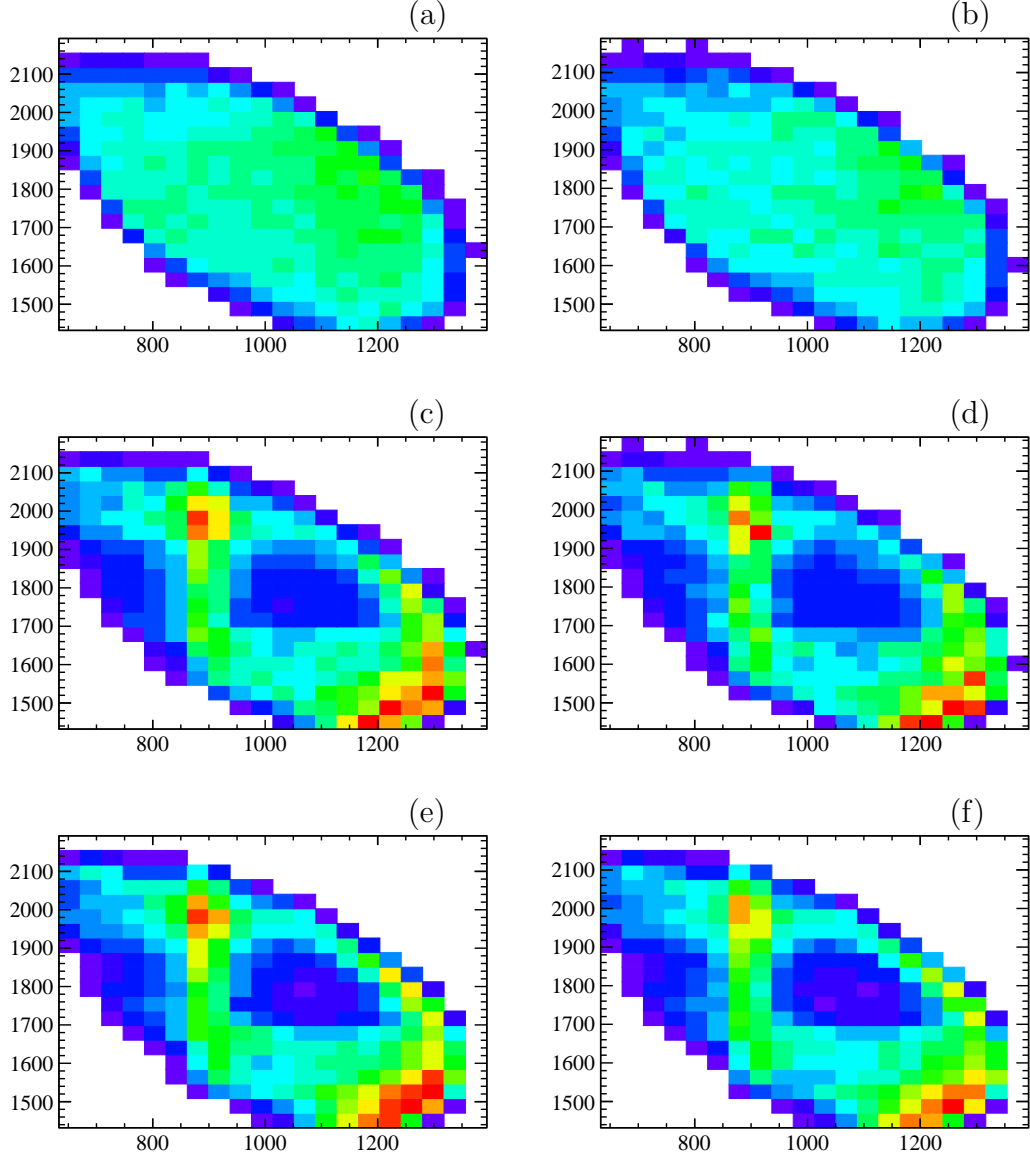
Figure 10: Distributions of Dalitz plots for orignal MC (top), corrected MC (middle) and sWeighted data (bottom) in (left) forward and (right) backward regions.
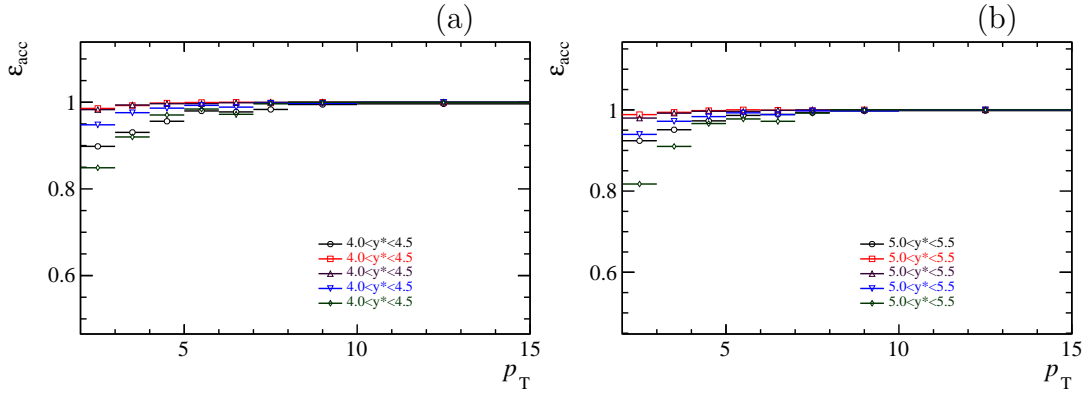

Figure 11: The geometrical acceptance efficiency ε_{acc} as a function of p_T and y^* of prompt Λ_c^+ baryon for (left) forward and (right) backward rapidities. Statistical uncertainties only.
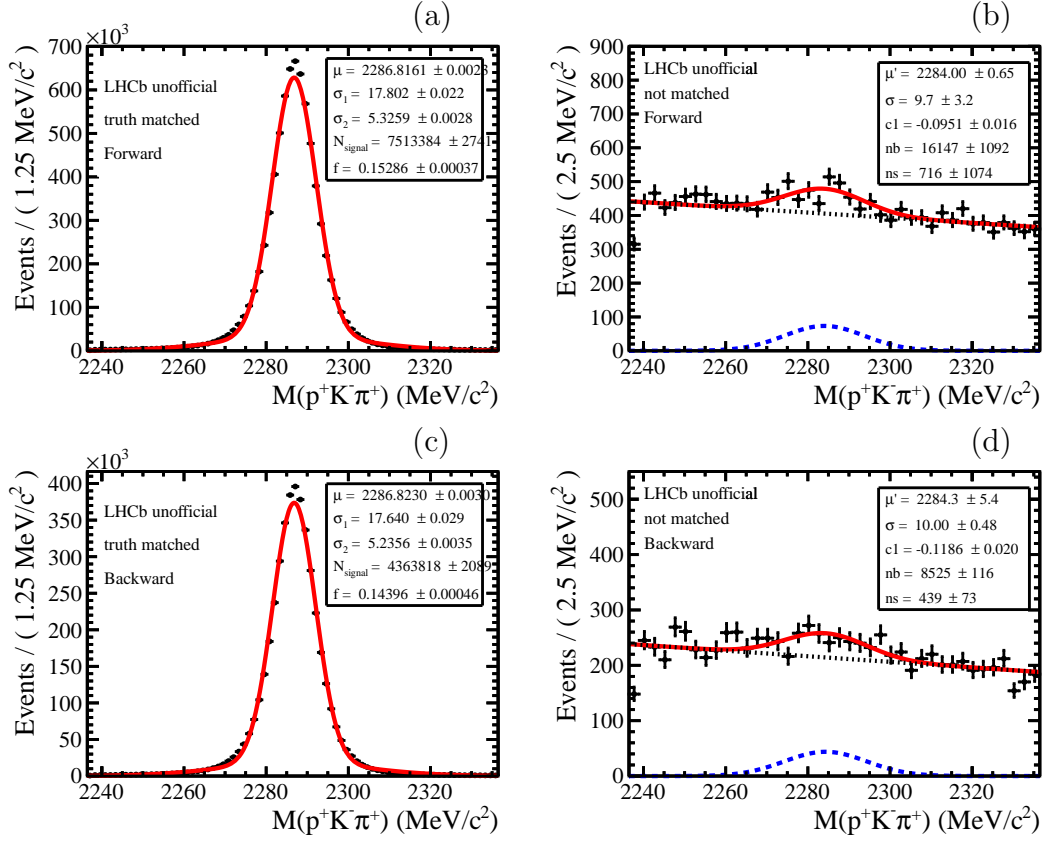
Figure 12: The $M(K\pi)$ distribution of (left) truth matched and (right) not matched Λ_c^+ baryons for (top) foward and (bottom) backward simulations.
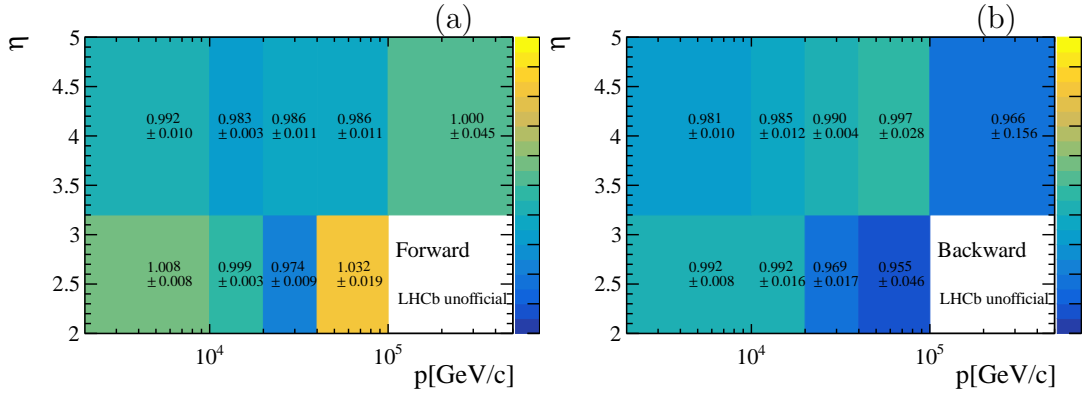
Figure 13: Tracking calibration table of (left) forward and (right) backward rapidities.

shown in Fig. 14 and incorporated with all corrections. The numerical values are listed in the Table **??** and **??**.
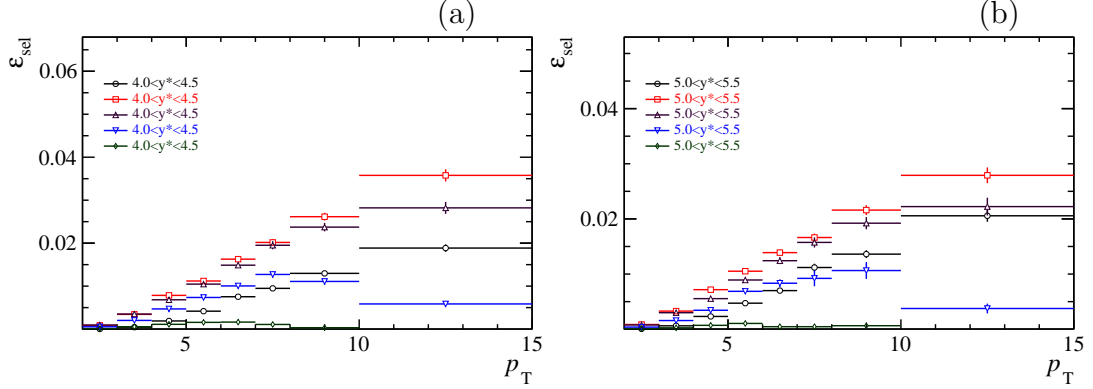
Figure 14: The selection efficiency $\varepsilon_{\text{rec\&sel}}$ as a function of p_{T} and y^* of prompt Λ_c^+ baryon for Fwd (top) and Bwd (bottom) configurations with all corrections considered. Statistical uncertainties only.

5.4 PID efficiency

The $\varepsilon_{\text{PID}}$of Λ_c^+ particle is evaluated with single track PID efficiencies with the formula

$$\varepsilon_{\text{PID}} \equiv \frac{\sum_i^N \varepsilon_p \times \varepsilon_K \times \varepsilon_\pi}{\sum_i^N \Lambda_c^+ \text{ reconstructed and selected}} \ , \tag{11}$$

where the sum goes over the selected Λ_c^+ simulation candidates. In Table 2 and 3, the selections are $\text{DLL}_{p\pi}(p) > 15$ and $\text{DLL}_{pK}(p) > 5$. and $\text{DLL}_{K\pi}(K^-) > 5$. and $\text{DLL}_{K\pi}(\pi^+) < 0$. The single track PID efficiencies are given as a function of $(p, \eta, N_{\text{VELO}}^{\text{Clusters}})$ and are estimated with data-driven methods. The calibration tables can be easily obtained from the package PIDCalib of URANIA, in the form of three-dimensional histograms, seeing https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage. For the K and π tables, the statistics are less than 150k, which limits the bin numbers of the histogram. Thus, the three-dimensional efficiencies can be obtained by multiplying the efficiency in different axses, that is

$$\varepsilon_{\text{PID}}(p, \eta, \text{nVeluClusters}) = \varepsilon_{\text{PID}}(p, \eta) \times \varepsilon(N_{\text{VELO}}^{\text{Clusters}})/\overline{\varepsilon}_{\text{PID}} \ . \tag{12}$$

The independence between the kinematic variables and multiplicity variables need to be examined before. Thus, the $N_{\text{VELO}}^{\text{Clusters}}$ distributions in different kinematic regions are shown in Figs. 15, 16 and 17. This independence holds for K and π samples, while does not hold for p sample.

The $\varepsilon_{\text{PID}}(p, \eta)$ and ε_{PID} for K and π are shown in Fig. 18 and Fig. 20 respectively. For p, the correlation between kinematic and multiplicity variables does exist. Fortunately, the corresponding calibration samples have more than 8M candidates, so three-dimensional tables can be directly obtained from these samples. Projections of the table to the kinematic and multiplicity axes are also shown in Fig. **??** to see an approximate value in each regions. As the $\varepsilon_{\text{PID}}(p)$ shows strong dependence on $p(p)$, the bin widths there are set much smaller. With the calibration tables, $\varepsilon_{\text{PID}}(\Lambda_c^+)$ can be derived via Equation 11. Fig. 21 show the results while the numerical values are listed in **??** in Appendix **??**.
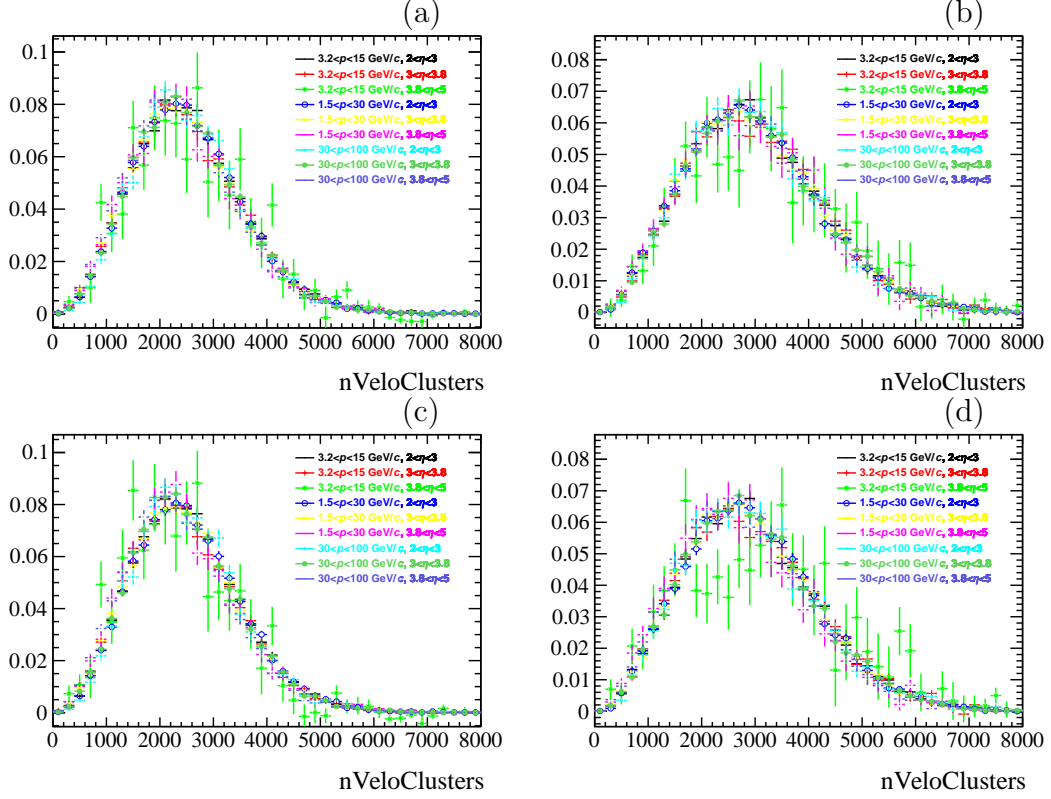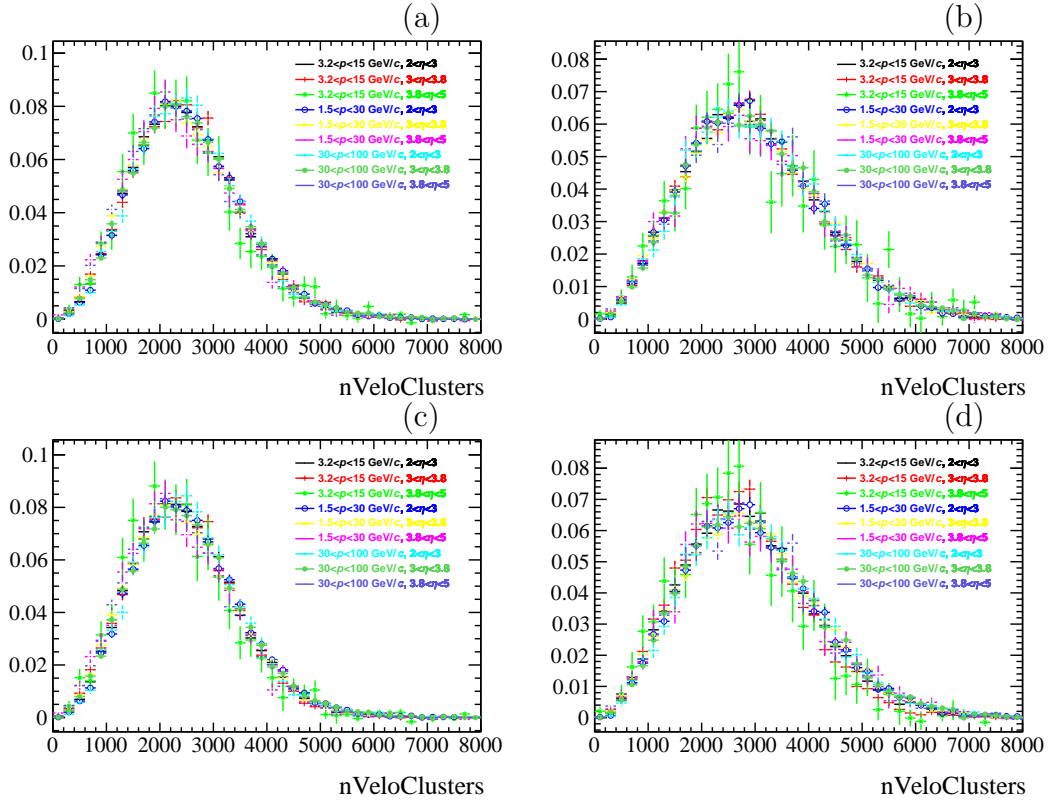
Figure 15: Normalised $N_{\text{VELO}}^{\text{Clusters}}$ distribution of K PID calibration sample (top) before and (bottom) after PID selections in (left) forward and (right) backward rapidites in different kinematic regions.

5.5 Trigger efficiency

As mentioned above, the triggers in this sample include L0, HLT1 and HLT2. The LOSPD is a minimum-biased trigger, so the efficiency is 1. The other two selections are considered in $\varepsilon_{\text{rec\&sel}}$.

5.6 Total efficiency

The total efficiencies are obtained directly from the multiplication of the efficiencies above as Equation 8. The results are plotted on Fig. 22 and listed in Table **??** and **??**,

Figure 16: Normalised $N_{\text{VELO}}^{\text{Clusters}}$ distribution of π PID calibration sample (top) before and (bottom) after PID selections in (left) forward and (right) backward rapidites in different kinematic regions.
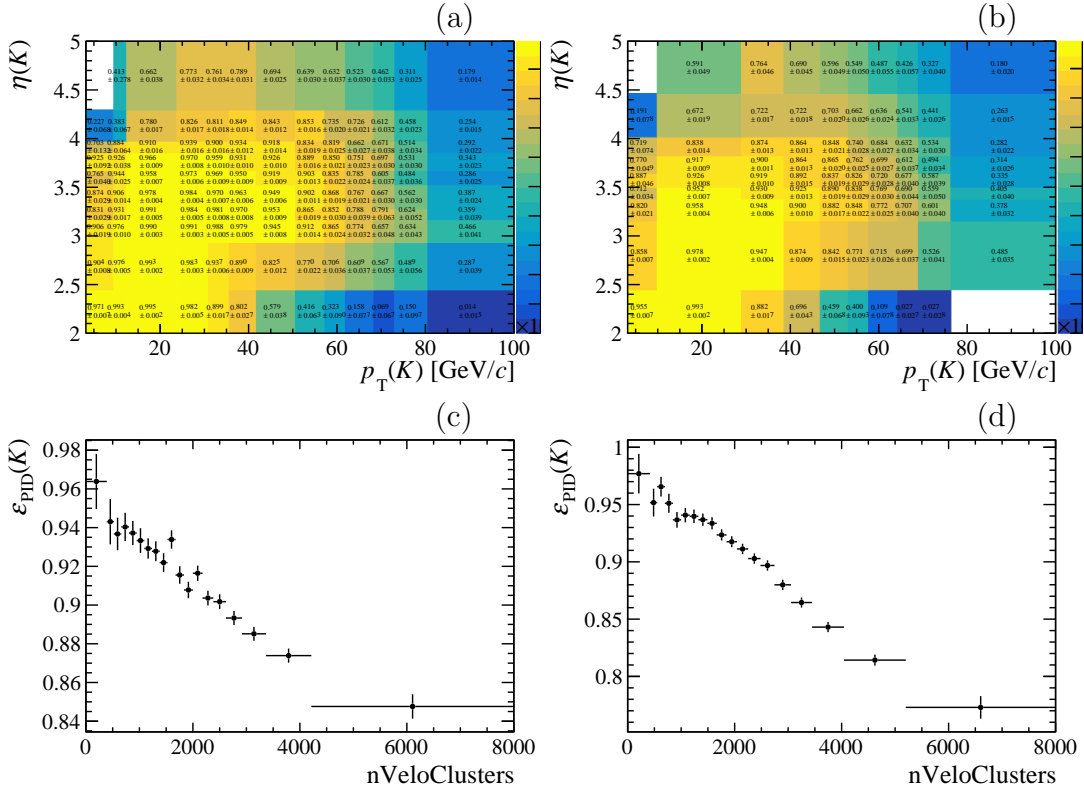
17
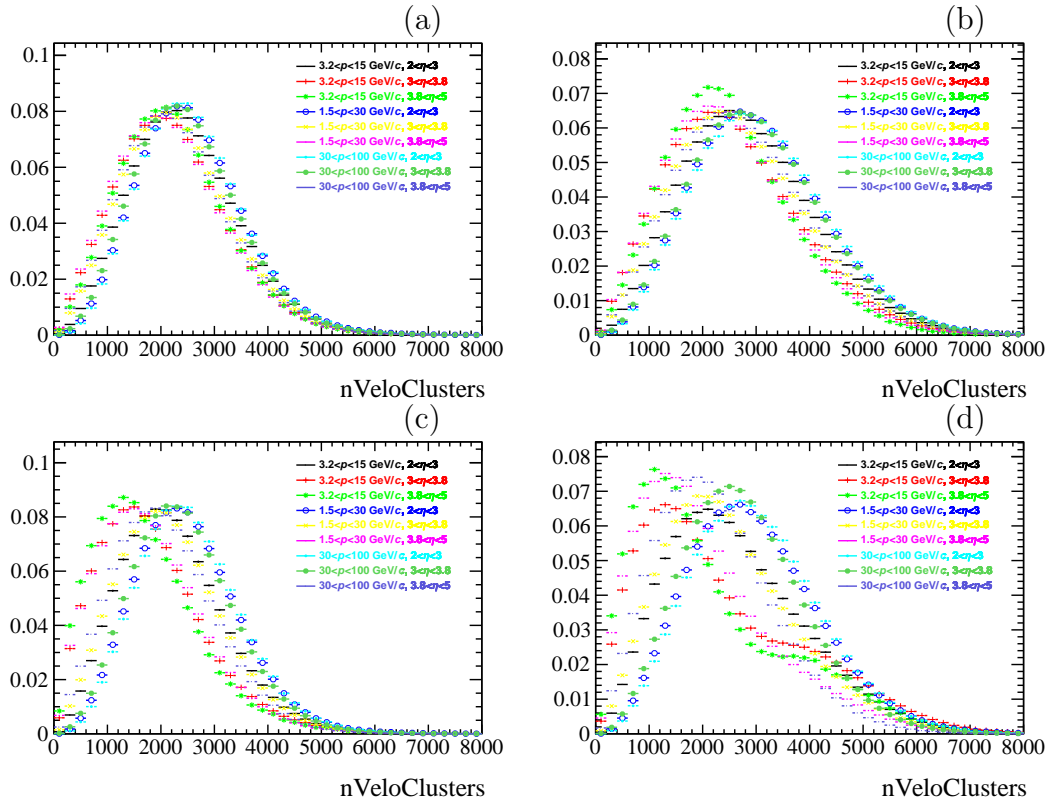
Figure 17: Normalised $N_{\mathrm{VELO}}^{\mathrm{Clusters}}$ distribution of p PID calibration sample (top) before and (bottom) after PID selections in (left) forward and (right) backward rapidites in different kinematic regions.

(a)

(b)

(c)

(d)

Figure 18: $\varepsilon_{\mathrm{PID}}(K)$ as functions of (top) (p,η) and (bottom) $N_{\mathrm{VELO}}^{\mathrm{Clusters}}$ in (left) forward and (right) backward rapidites.
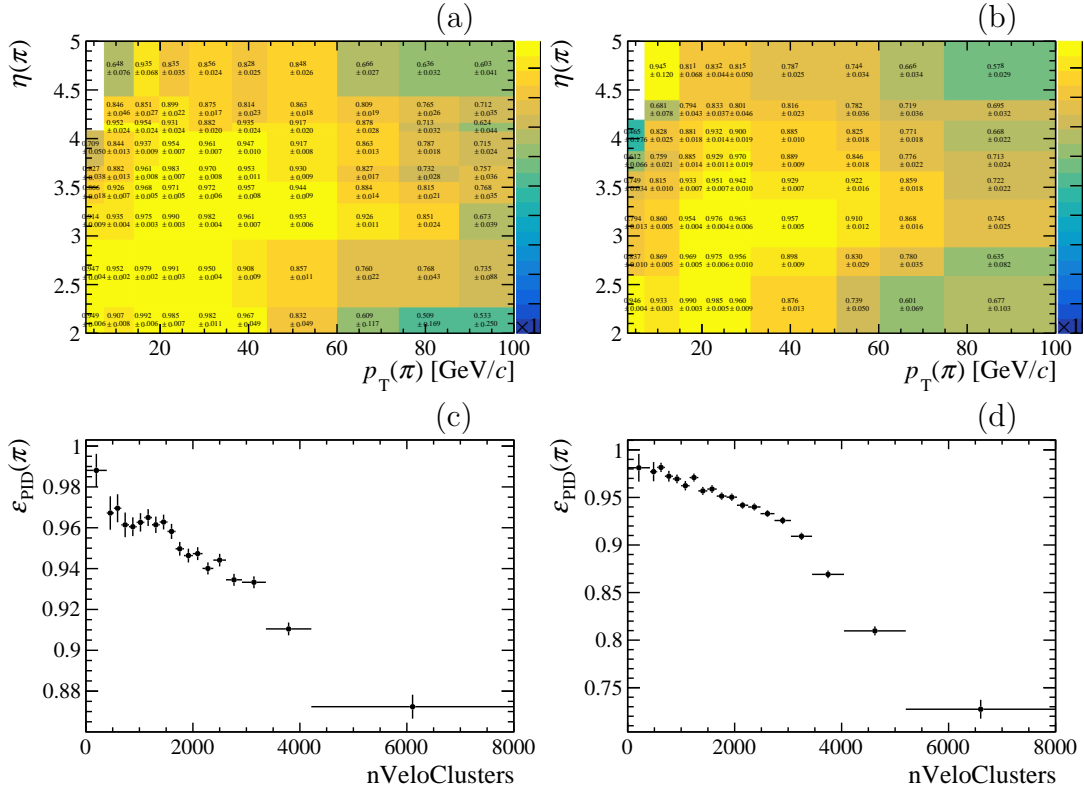
19

Figure 19: $\varepsilon_{\mathrm{PID}}(\pi)$ as functions of (top) (p, η) and (bottom) $N_{\mathrm{VELO}}^{\mathrm{Clusters}}$ in (left) forward and (right) backward rapidites.
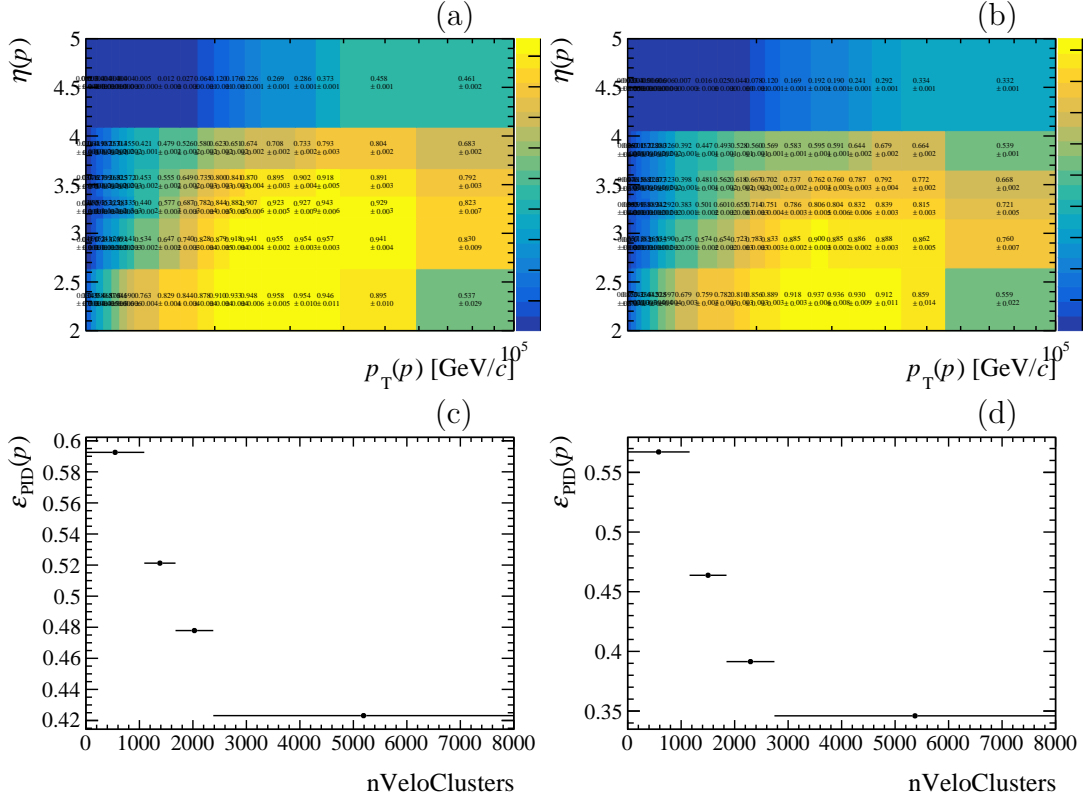
Figure 20: $\varepsilon_{\mathrm{PID}}(p)$ as functions of (top) (p, η) and (bottom) $N_{\mathrm{VELO}}^{\mathrm{Clusters}}$ in (left) forward and (right) backward rapidites.

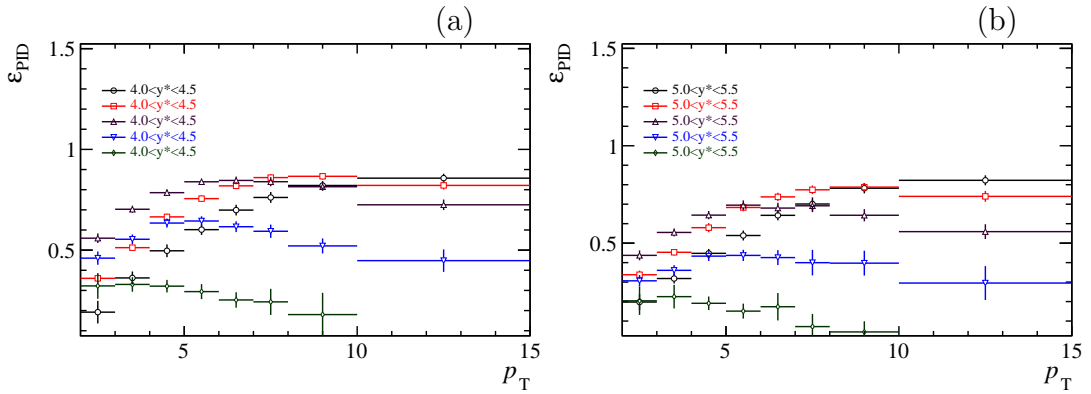Figure 21: The PID efficiency $\varepsilon_{\mathrm{PID}}$ as a function of p_{T} and y^* of prompt Λ_c^+ baryon for (left) forward and (right) backward rapidities. Statistical uncertainties only.
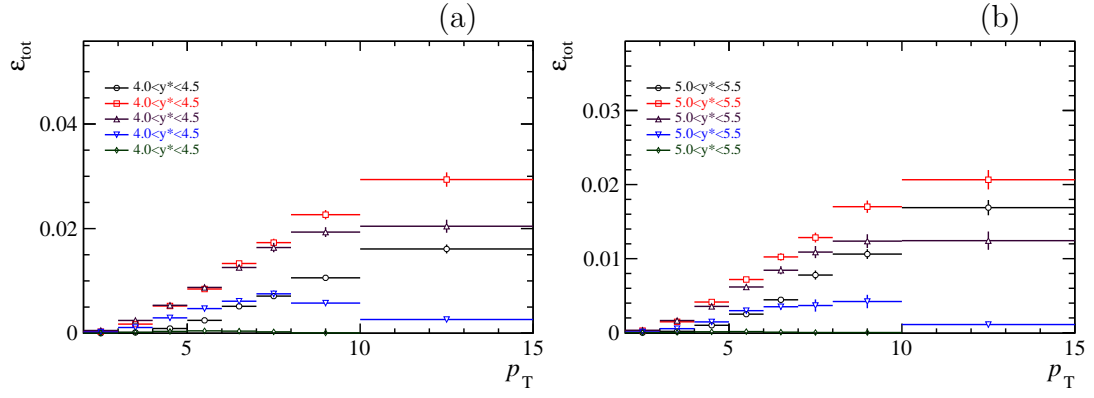
Figure 22: The total efficiency ε_{tot} as a function of p_{T} and y^* of prompt Λ_c^+ baryon for (left) forward and (right) backward rapidities.

References

[1] LHCb collaboration, R. Aaij *et al.*, *Measurement of Ξ_c^+ production in pPb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV at LHCb*, arXiv:2305.06711, submitted to Phys. Rev. Lett.

[2] LHCb collaboration, R. Aaij *et al.*, *Prompt Λ_c^+ production in pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, JHEP **02** (2019) 102, arXiv:1809.01404.

[3] Particle Data Group, R. L. Workman *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2022** (2022) 083C01.

[4] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP observables in the process $B^0 \to DK^{*0}$ with two- and four-body D decays*, JHEP **08** (2019) 041, arXiv:1906.08297.

[5] LHCb collaboration, R. Aaij *et al.*, *Measurement of the prompt D^0 nuclear modification factor in pPb collisions at $\sqrt{s_{NN}} = 8.16$ TeV*, Phys. Rev. Lett. **131** (2023) 102301, arXiv:2205.03936.

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

[7] A. D. Bukin, *Fitting function for asymmetric peaks*, arXiv:0711.4449.

[8] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, Annals of Statistics **29** (2001) 1189.