AGH

AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY

Field of Science: Natural Sciences

Scientific discipline: Physical Sciences

DOCTORAL THESIS

New approaches for searching for *CP* violation effects in the $\Xi_c^+ \to pK^-\pi^+$ decays in the LHCb experiment

Author: Jakub Ryżka

Supervisor: prof. dr hab. inż. Tomasz Szumlak
Co-supervisor: dr. hab. Artur Ukleja, prof. AGH

Completed in: AGH University of Science and Technology, Department of Particle Interactions and Detection Techniques, Krakow, 2023

Declaration of the Author of this dissertation

Aware of legal responsibility for making untrue statements I hereby declare that I have written this dissertation myself and all the contents of the dissertation have been obtained by legal means.

signature date

Declaration of the Supervisor of this dissertation

This dissertation is ready to be reviewed.

signature of thesis Supervisor date

Contents

Streszczenie

Symetria przestrzenno-ładunkowa (*CP*) oznacza, że prawa fizyki są takie same, jeśli cząstka zostanie zamieniona z jej antycząstką (symetria sprzężenia ładunkowego), a jej współrzędne przestrzenne ulegną odwróceniu (symetria parzystości). Jednak symetria w transformacji *CP* może zostać naruszona w przypadku oddziaływań słabych. W rozpadach mezonów dziwnych i pięknych efekt łamania *CP* jest dobrze potwierdzony eksperymentalnie. Kilka lat temu efekt ten został także potwierdzony w rozpadach mezonów powabnych. Niemniej jednak, aby w pełni wyjaśnić proces bariogenezy w naszym Wszechświecie, potrzebny jest większy poziom łamania *CP* niż ten obserwowany w modelu standardowym (MS), co implikuje konieczność obecności dodatkowych źródeł łamania *CP* poza tymi znanymi w MS.

Po dziesięcioleciach poszukiwań eksperymentalnych udało się zaobserwować łamanie *CP* w sektorze powabnym. Asymetrie *CP* zostały zmierzone w rozpadach $D^0 \to K^+K^-$ i $D^0 \to \pi^+\pi^-$ przy użyciu dużej próbki powabnych mezonów zebranych przez eksperyment LHCb podczas LHC Run 1 i Run 2. Dokonując pełnego połączenia z wcześniejszymi wynikami LHCb, uzyskano następującą wartość: $\Delta A_{CP} = (-15.4 \pm 2.9) \times 10^{-4}$. Niedawno, w 2022 roku, fizycy ze Współpracy LHCb po raz pierwszy zmierzyli asymetrię *CP* w pojedynczym rozpadzie mezonu powabnego: $a^{dir}_{\pi^-\pi^+} = (23,2 \pm 6,1) \times 10^{-4}$.

Odkrycie łamania *CP* w rozpadach mezonów powabnych otwiera okno na poszukiwanie nowych możliwych źródeł asymetrii *CP* w cząstkach powabnych. Jedną ze szczególnie interesujących grup rozpadów, w których oczekuje się złamania symetrii *CP*, są rozpady barionów powabnych $\Xi_c^+ \to pK^-\pi^+$. Dane z Run 1 zostały już przeanalizowane. W jednym z regionów rozkładu Dalitza obserwuje się zgodność z brakiem łamania symetrii *CP* wynoszącą 2.7 standardowych odchyleń. Dlatego kluczowa jest dalsza analiza tych rozpadów z większą statystyką danych zebranych w Run 2 w celu potwierdzenia istnienia łamania *CP* z tych rozpadach. Trudno jest określić, czy obecnie stosowane metody nie są wystarczająco czułe, czy też łamanie *CP* jest zbyt małe, aby można je było zmierzyć, ponieważ zgodnie z przewidywaniami MS oczekuje się, że wielkość tego łamania w rozpadach cząstek powabnych będzie rzędu $10^{-4} - 10^{-3}$ i mniej.

Niniejsza rozprawa doktorska przedstawia szczegóły poszukiwań efektów łamania *CP* w trzyciałowych rozpadach barionów $\Xi_c^+ \to pK^-\pi^+$ produkowanych bezpośrednio w zderzeniach *pp* w akceleratorze LHC i zebranych w eksperymencie LHCb w Run 2, odpowiadających scałkowanej świetlności niecałych 6 fb^{-1}. Poszukiwania są prowadzone przy użyciu różnych technik w sposób niezależny od modelu. Techniki te to: metoda zależna od podziału na przedziały S_{CP}, estymator gęstości jądra (ang. Kernel Density Estimator (KDE)) i metoda testu energetycznego (ang. Energy Test (ET)).

Abstract

The charge conjugation parity (CP) symmetry states that laws of physics should be the same if a particle is changed with its antiparticle (charge conjugation symmetry) and its spatial coordinates undergo inversion (parity symmetry). However, the symmetry under CP transformation can be violated in case of weak interactions. In the decays of strange and beauty mesons, the CP violation effect is well-established experimentally. In last years, the effect is also confired in charm mesons. Nonetheless, more than the size of CP violation in the Standard Model (SM) is needed to fully explain the baryogenesis process in our universe, implying the necessity of the presence of additional sources of CP violation beyond those known in the SM.

After decades of experimental searches, the CP violation in the charm sector has been observed. The CP asymmetries were measured in $D^0 \to K^+K^-$, and $D^0 \to \pi^+\pi^-$ decays using a large sample of charm mesons collected by the LHCb experiment during the LHC Run 1 and Run 2. By making a full combination of LHCb data, the following value was obtained: $\Delta A_{CP} = (15.4 \pm 2.9) \times 10^{-4}$. Recently, in 2022, physicists from LHCb collaboration measured, for the first time, the CP asymmetry in a single charm meson decay: $a^{dir}_{\pi^-\pi^+} = (23.2 \pm 6.1) \times 10^{-4}$.

Discoveries of CP violation in the decays of charmed mesons open a window to the search for new possible sources of CP asymmetries in charm particles. The particularly interesting decays wherein CP symmetry is suspected to be violated are charm baryons $\Xi_c^+ \to pK^-\pi^+$. The Run 1 data have already been analysed. In one of the regions of the Dalitz plot, agreement with the lack of CP symmetry violation is observed of 2.7 standard deviations. Therefore, further analysis of these decays is crucial with more data statistics collected in Run 2 to confirm the existence of CP violation in such decays. It is hard to determine whether currently used methods are not sensitive enough or CP violation is too small to be measured as, according to the SM, the size of this asymmetry in charm decays is expected to be of the order of $10^4 - 10^3$ or less.

This dissertation presents the details of searches for CP violation effects in the Cabibbo suppressed charm baryon $\Xi_c^+ \to pK^-\pi^+$ decays produced directly in pp collisions at LHC accelerator and collected in the LHCb experiment in Run 2, corresponding to an integrated luminosity of about 6 fb^{-1}. The searches are performed with different techniques in model independent ways. These techniques are: the binned S_{CP} method, the Kernel Density Estimator and the unbinned Energy Test method.

Acknowledgments

Sincere thanks to my supervisors Tomasz Szumlak and Artur Ukleja for all their help, guidance, care and patience during the creation of this dissertation.

Many thanks to Wojciech and Paweł for their valuable insights and advice.

My gratitude also goes to my wife and my parents, who have always supported me throughout my doctoral studies.

Author's contribution

The first step of the analysis was data selection. This selection involved choosing cases with Ξ_c^+ decays in such a way as to retain as many events as possible while reducing the number of background events. To be sure that the sample was not biased during this process, I checked the lifetime of the Ξ_c^+ using the sPlot method. It turned out to be consistent with what is known in the PDG. Being certain that the sample was correctly selected, I proceeded to analyse the effects of reconstruction and detector asymmetry to determine additional cuts to reduce pollution asymmetries and to eliminate false *CP* violation signals. Apart from data selection I have performed sensitivity analysis of the methods used in this analysis to search for *CP* asymmetry (S_{CP}, Energy Test) and created an optimized version of the Kernel Density Estimation technique. The results of all the above activities are included in this thesis.

Talks given by the Author and posters presented at the international and national conferences:

- J. Ryżka, *Discovery of CP violation in charm decays at LHCb experiment and prospects for Run III and Run IV*, 3rd Jagiellonian Symposium on Fundamental and Applied Subatomic Physics, Kraków, Poland, 23-28 June, 2019, poster, plenary
- J. Ryżka, *CP violation in charm*, The XIVth Quark Confinement and the Hadron Spectrum Conference, Norway, 2-6, August, 2021
- J. Ryżka, *Charmed baryons at LHCb*, The XIV International Conference on Beauty, Charm and Hyperon Hadrons (BEACH 2022), Krakow, Poland, 5-11, June, 2022, talk, plenary
- J. Ryżka, *The charm of charm*, The 4th Jagiellonian Symposium on Advances in Particle Physics and Medicine, Krakow, Poland, 10-15, July, 2022, talk, plenary

The Author published the following publications and proceedings:

- J. Ryżka, *Discovery of CP violation in charm decays at LHCb experiment and prospects for Run III and Run IV*, Acta Phys.Polon. B Vol. 51 (2020) 301-307
- J. Ryżka, *CP violation in charm*, EPJ Web of Conferences 258, 06006 (2022)
- J. Ryżka, *Charmed baryons at LHCb*, Acta Phys. Polon. B, Proceedings Series, Vol. 16, article 3-A12
- J. Ryżka, *The charm of charm*, Acta Phys. Polon. B, Proceedings Series, Vol. 15 (2022) article 4-A2

1 Introduction

In modern particle physics and cosmology, one of the most crucial problem is how to understand the the existed asymmetry between matter and antimatter in the universe. It is believed that the baryogenesis mechanism can solve the problem, which requires three Sakharov criteria: baryon number violation, charge and charge-parity (*CP*) violation, and out of equilibrium. In the Standard Model (SM) of the particle physics, *CP* violation (*CPV*) is directly generated from a weak phase in the quark mixing matrix in the Kobayashi-Maskawa mechanism. However, in the SM, *CPV* is unfortunately far smaller than the requirement of explaining the matter-dominant universe. It implies that it is needed for a new *CPV* sources of new physics beyond the SM. *CPV* has only been observed in the K, B and D meson systems but has never been established in the baryon sectors yet. Besides, the visible matter of the universe is dominantly made of baryons. Therefore, studying baryon *CPV* provides a good opportunity to test the SM and search for new physics.

The conditions for the symmetry-violating process between matter and antimatter are currently possible only in particle accelerators. Such experiments require a lot of work and the use of the latest technical and technological solutions. Currently, the European Centre for Nuclear Research in Geneve, Switzerland (CERN) is one of the leading centres for physics research and observations of this kind. In addition, this facility carries out the measurements and implements a program directly related to the asymmetry between matter and antimatter called the Large Hadron Collider beauty experiment (LHCb). This experiment was launched in 2008, and since then, many studies have been carried out within its framework, which has allowed, among other things, the characterization of *CPV* in beauty and charm meson and baryons decays.

Over the past sixty years, *CPV* has been observed in the K, D, and B-meson systems in several experiments [1–9]. In the charm mesons, the recent observation of *CPV* [10] stimulates a wide discussion to understand its nature. The discovery measurement of *CPV* in neutral charm meson decays uses the difference between two times integrated *CP*-violating asymmetries of Cabibbo-suppressed D^0 decays ($D^0 \to \pi^-\pi^+$ and $D^0 \to K^-K^+$) found to be $(15.4 \pm 2.9) \times 10^{-4}$. This observable does not establish the amount of *CPV* in the individual decays. In 2022, the LHCb experiment investigated $D^0 \to K^-K^+$ decays again to measure *CP* asymmetry in a single decay [11]. The obtained value was combined with the previous measurement [10] to calculate *CP* asymmetry in the $D^0 \to \pi^-\pi^+$ decay. It was measured to be $a^{dir}_{\pi^-\pi^+} = (23.2 \pm 6.1) \times 10^{-4}$. This asymmetry value deviates from zero at $3.8\,\sigma$ level. It provides the first evidence of direct *CPV* in a single charm meson decay.

Huge data samples containing information about the decays of charm particles have been collected in the LHCb experiment during two data-taking periods: Run 1 (2011-2012) and Run 2 (2016-2018). These data sets enable searches for new sources of *CPV* in the charm sector, with a previously unavailable sensitivity. So far, it has not been possible to confirm the existence of *CPV* in any beauty and charm baryon decays.

This thesis aims to present the current state of knowledge concerning the phenomenon of *CPV* in decays of charm baryons $\Xi_c^+ \to pK^-\pi^+$. The search for this phenomenon and the effects of the new physics in the LHCb experiment is discussed in detail in this thesis. The results are obtained using the binned S_{CP} method [12], the unbinned Energy Test method (ET) [13–17] and the Kernel Density Estimation technique (KDE) [18]. The latter two have never before been used to analyze baryon decays.

This dissertation is organised as follows: in Sec. 2, the experimental and theoretical aspects of the *CPV* phenomenon in the SM are presented. The known ways of *CPV* as well as the predicted values of them, and the experimental status is also discussed. In Sec. 3, the experimental framework of the LHCb experiment is described. In Sec. 4, the data selection, the detector effects and the possible background sources are presented. Then, in Sec. 5, the used methods for searching for *CPV* effects are discussed. The tests of the using methods in the control $\Lambda_c^+ \to pK^-\pi^+$ decays, as well as in the sidebands of the $\Xi_c^+ \to pK^-\pi^+$ decays are shown in Sec. 7. In the same section, the sensitivity of the methods, checked in pseudo-experiments, is also presented.

2 Experimental and theoretical aspects

This section contains basic information about *CPV* phenomenon: its definition, known mechanisms, expected values in the SM and the importance of planned searches for new physics effects.

2.1 Standard Model

In the SM, there are three types of elementary particles: leptons, quarks and intermediate bosons. On the other hand, the elementary particles are grouped into three families (generations). Each lepton family contains a pair: one massive charged particle (e^-, μ^-, τ^-) and one neutral neutrino (ν_e, ν_μ, ν_τ), i.e. e^- and ν_e etc. More than twenty years ago, neutrinos were treated as massless, but non-zero differences of mass between them are measured [19, 20].

In addition to three lepton generations, there are six quarks with fractional electric charge grouping into three families as well. In a similar way to leptons, each quark family is composed of pair: up-down, strange-charm and beauty-top. There are also bosons: two massive charged W (W^+ and W^-), one massive neutral Z^0, massless photon, gluon and massive Higgs discovered in 2012 in ATLAS and CMS experiments [21].

Although, the SM describes the existing data in amazing satisfactory way, there are phenomena which are not properly accommodated therein. For example, the observed domination of matter over antimatter in our universe is not fully understood. This asymmetry between matter and antimatter is characterised by the baryon-to-entropy ratio:

$$Y_B \equiv \frac{n_B - n_{\bar{B}}}{s} \approx 10^{-11}, \tag{1}$$

where n_B and $n_{\bar{B}}$ are baryon and antibaryon densities, respectively, and s is the entropy density. There are three necessary conditions, called Sakharov's criteria (reference), and the first of them is to produce a baryon asymmetry in the early universe. The second condition is the loss of the thermal equilibrium. According to the third condition, both C (charge conjugation) and CP (composition of charge and parity) symmetries are violated. The first and second conditions require to observe the baryon asymmetry, the Y_B parameter must take a non-zero value and thermal balance implies that rates of the opposing processes in which baryons would be created and destroyed would be equal.

2.2 Phenomenon of *CP* violation

The *CP* symmetry states that laws of physics are the same if a particle is changed with its antiparticle (charge conjugation symmetry) and its spatial coordinates are inversed (parity symmetry). However, in the case of weak interactions, the matter-antimatter symmetry is violated under the *CP* transformation. The weak interactions of the quarks are described by Cabibbo-Kobayashi-Masakawa (CKM) matrix [22, 23]. In the SM, *CPV* is included through a complex weak phase in the CKM matrix.

There are three ways of *CPV* which are known in the SM. The first one is induced in the neutral mesons mixing and is related to the difference in the transition probability of meson and antimeson, i.e. $P(M^0(t) \to \bar{M}^0) \neq P(\bar{M}^0(t) \to M^0)$, where M^0 and $\bar{M}^0$ are flavour eigenstates, measured in the experiments. For those particles, the two mass eigenstates $M_{1,2}$ can be expressed as a linear combination of the flavour eigenstates:

$$|M_{1,2}\rangle = p|M^0\rangle \pm q|\bar{M}^0\rangle, \tag{2}$$

where the p and q are complex coefficients which satisfy relation $|p|^2 + |q|^2 = 1$. The case $|p/q| \neq 1$ means CPV existence. This type of CP violation is independent of the decay mode.

Secondly, CP symmetry can be violated directly in the decay amplitudes (A_f) of a particle decaying to a final state f in comparison to that of its own antiparticle decay amplitude $(\bar{A}_{\bar{f}})$ to the CP conjugate final state $\bar{f}$. In this case, the CP asymmetry can be defined as:

$$A_{CP} \equiv \frac{|A_f|^2 - |\bar{A}_{\bar{f}}|^2}{|A_f|^2 + |\bar{A}_{\bar{f}}|^2} \tag{3}$$

and CPV is present when $|\bar{A}_{\bar{f}}|/|A_f| \neq 1$. This type of CP violation depends on the decay mode and concerns both neutral and charged particles.

The third way is related to the interference between direct decay and its corresponding decay processing via mixing [24] and it concerns neutral mesons only. In order to CPV to be present the following condition must be fulfilled:

$$arg(\lambda_f) + arg(\lambda_{\bar{f}}) \neq 0, \tag{4}$$

where

$$\lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f}. \tag{5}$$

The combination of CPV in the interference and in mixing is also called indirect CPV.

Baryon-number conservation prevents the mixing of baryons and antibaryons, such that only the second way (so-called direct CPV) is allowed in baryon decays.

To observe CPV at least two amplitudes contributing to a particular process should interfere with different weak phases. For example, one can consider $\zeta \to f$decay with amplitude A_f and its corresponding CP conjugate process $\bar{\zeta} \to \bar{f}$ with amplitude $\bar{A}_{\bar{f}}$. There are two types of phases in those decay amplitudes. The first kind appears in A_f and $\bar{A}_{\bar{f}}$ with opposite signs and occurs in couplings of $W^{\pm}$ bosons, hence it's called the "weak phase". This is the phase responsible for CPV existence. The second type is the same for both decay amplitudes and because it is mostly related to the strong interactions it is named "strong phase". It is worth noting that the strong interactions preserve the CP symmetry.

It is convenient to write out each contribution of the decay amplitude in three parts: its magnitude $|a_i|$, its weak phase ϕ, and its strong phase δ_i:

$$\begin{aligned} A_f &= |a_1|e^{i(\delta_1+\phi_1)} + |a_2|e^{i(\delta_2+\phi_2)}, \\ \bar{A}_{\bar{f}} &= |a_1|e^{i(\delta_1-\phi_1)} + |a_2|e^{i(\delta_2-\phi_2)}. \end{aligned} \tag{6}$$

After combining amplitudes from Eq. 6 with Eq. 3, CP asymmetry (A_{CP}) can be described as:

$$A_{CP} = \frac{2|a_1||a_2|sin(\delta_2 - \delta_1)sin(\phi_2 - \phi_1)}{|a_1|^2 + |a_2|^2 + 2|a_1||a_2|cos(\delta_2 - \delta_1)cos(\phi_2 - \phi_1)}. \tag{7}$$

The above Eq. 7 shows that to observe CPV two or more amplitudes have to contribute to the decay and their strong and weak phases must be different. The existence of a non-zero strong phase difference is one of the key conditions necessary for the CPV to occur.

The CPV expressed in Eq. (7) can be only measured in the SM in the so-called Singly Cabibbo-Suppressed (SCS) decays. It is due to the fact that only those decays proceed in tree and penguin transitions, leading to the interference of at least two amplitudes: one related to three diagrams and one related to the penguin diagram. If weak and strong phases of tree and penguin diagrams are different, then CP symmetry between particles and antiparticles is violated. An example of charm baryon SCS decay is the $\Xi_c^+ \to pK^-\pi^+$, discussed in this thesis. In this case, the a_1 and a_2 are magnitudes of the tree (A_T) and penguin (A_P) amplitudes, respectively, given by the following expressions:

$$\begin{aligned} A_T &= \lambda T, \\ A_P &= \lambda P_d + \lambda P_s + \lambda^5 P_b, \end{aligned} \tag{8}$$

where λ is one of the Wolfenstein parameters of the CKM matrix. Its value is 0.22500 ± 0.00067 [25].

The $\Xi_c^+ \to pK^-\pi^+$ charm baryon decay diagrams are presented in the Figs. 1 and 2 as tree and penguin, respectively. The penguin transitions can occur in three ways as there is a loop in which c quark changes its flavour, by exchanging W^+ boson, and one d, s or b quark.

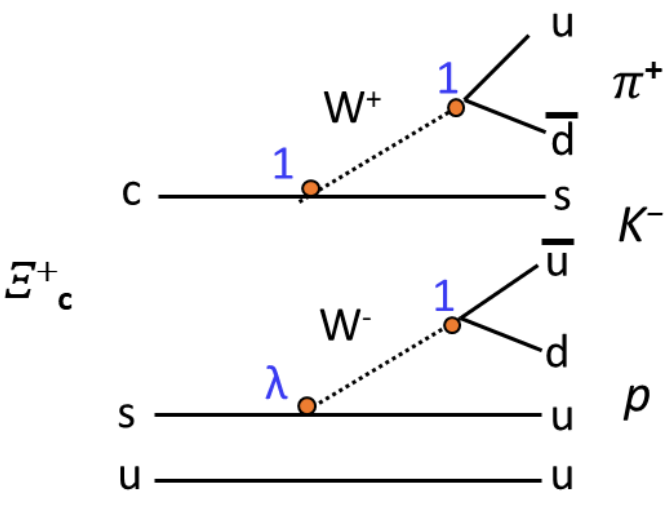

Figure 1: The tree diagram of $\Xi_c^+ \to pK^-\pi^+$ decay.

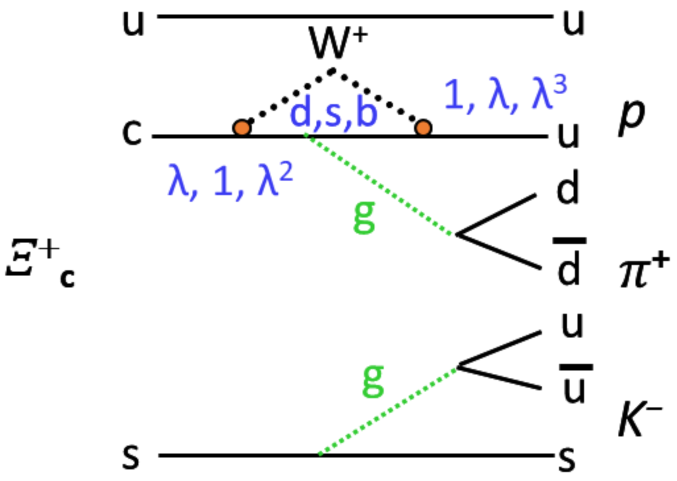

Figure 2: The penguin diagram of $\Xi_c^+ \to pK^-\pi^+$ decay.

In contrast to SCS processes, there are processes described only using three diagrams, which are called Cabibbo favoured (CF). Processes occur through penguin diagrams only are called doubly Cabibbo-Suppressed (DCS). Both CF and DCS processes are those in which the SM *CP* asymmetry is expected to be negligible.

The CF $\Lambda_c^+ \to pK^-\pi^+$ decay is used in this analysis as a control process because as was mentioned above the *CP* asymmetry in those decays in the SM is negligible and the final states of Λ_c^+ and Ξ_c^+ are the same. In the control processes all methods for searching for *CPV* are tested in aspects that they do not generate fake sygnals of *CPV* as well as the detector effects are estimated.

2.3 Expected value of *CP* violation

In the SM, the theory predicts the value of the *CPV* for charm mesons to be 10^{-3} or less and *CP* asymmetry of the order of magnitude of 10^{-4} in the charm sector has already been observed [10, 11, 26–35], what is significantly smaller than measured in the beauty mesons [36, 37], where even extremely large 75% of *CPV* is found [38]. Predictions of the magnitude of *CP* asymmetry for charm baryons are of the same order as for charm mesons [39, 40]. However, it is essential to note that predictions for mixing mesons are subject to significant uncertainties up to a few orders of magnitude [41]. It is because contributions from non-perturbative long-range diagrams, which are hard to determine, impact their values.

The searches for *CPV* are interpreted as searches for physics beyond the SM. The charm sector is particularly promising in that due to the small contribution from the SM (the expectation of *CPV* is smaller than per mille). There are calculations which show that new particles can increase the size of *CPV* by an order of magnitude [42, 43] . Since *CPV* in CF and DCS decays is predicted to be negligible in the SM any non-zero signal of

CP asymmetry in these decays implies the existence of new physics.

2.4 Experimental status

The existence of *CPV* in the decays of strange and beauty mesons is very well established experimentally. The violation of *CP* symmetry was implied for the first time in an experiment in 1964 at Brookhaven Laboratory Christenson [44], which showed that the long-lived neutral kaon can decay into two pions, and latter was detected in BaBar and Belle collaborations in 2001 in B^0 mesons [45,46]. After decades of experimental searches, *CPV* in the charm particle decays has been observed for the first time in the LHCb experiment [10]. In this measurement, the existence of *CPV* is confirmed via non-zero difference of the *CP* asymmetries between the two-body charm meson decays ($D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$):

$$\Delta A_{CP} = (-15.4 \pm 2.9) \times 10^{-4}, \tag{9}$$

where both statistical and systematic uncertainties are incorporated into one. The significance of the deviation from zero is 5.3 standard deviations. The difference in *CP* asymmetries is a powerful variable, and in case of decays of D^0 into charged pions or kaons, two contributing asymmetries were believed to be of the same magnitude and opposite signs [24]. However, recent findings by theorists indicate that both asymmetries $A(K^-K^+)$ and $A(\pi^-\pi^+)$ magnitudes are significantly different [47]. Nonetheless, to properly determine and investigate the source of potential *CP* violation, one has to examine the single asymmetry. And recently, in 2022, the LHCb experiment has investigated $D^0 \to K^-K^+$ decays again to measure *CP* asymmetry in a single decay [11]. Data were taken at the centre of mass energy of 13 TeV, corresponding to an integrated luminosity of 5.9 fb^{-1}. Instead of calculating ΔA_{CP} this time physicists considered several control channels with negligible *CP* asymmetries to remove nuisance asymmetries and obtained the *CP* asymmetry in the $D^0 \to K^-K^+$ decay only. This value is measured to be

$$A_{CP}(K^-K^+) = (6.8 \pm 5.4 \pm 1.6) \times 10^{-4}. \tag{10}$$

Combining above value with previous result [10], the direct *CP* asymmetry $a^{dir}_{K^-K^+}$ is determined to be

$$a^{dir}_{K^-K^+} = (7.7 \pm 5.7) \times 10^{-4}. \tag{11}$$

The above value is consistent with the *CP* symmetry but allows to measure a single value of direct *CP* asymmetry in $D^0 \to \pi^-\pi^+$ mode $a^{dir}_{\pi^-\pi^+}$ by combining several components: the LHCb measurement of $A_{CP}(K^-K^+)$ [48, 49], previous *CP* asymmetry difference ΔA_{CP} [10, 48, 50], the current LHCb average of ΔY (the slope of the time-dependent asymmetry of the decay rates of D^0 and $\bar{D}^0$ mesons) [51], the world average of the D^0 lifetime [25] and values of reconstructed mean time decays of the $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$. The measured $a^{dir}_{\pi^-\pi^+}$ is

$$a^{dir}_{\pi^-\pi^+} = (23.2 \pm 6.1) \times 10^{-4}. \tag{12}$$

This time, the *CP* asymmetry in the $D^0 \to \pi^-\pi^+$ decay deviates from zero at 3.8 standard deviations. This is the first evidence of direct *CPV* in a single charm meson decay.

In regards to the *CPV* measurements in charm baryons, the difference of the *CP* asymmetry (ΔA_{CP}^{baryon}) between $\Lambda_c^+ \to pK^-K^+$ and $\Lambda_c^+ \to p\pi^-\pi^+$ is obtained to be [52]

$$\Delta A_{CP}^{baryon} = (0.30 \pm 0.91 \pm 0.61)\%. \tag{13}$$

This value is consistent with zero in agreement with the *CP* symmetry.

So far, *CPV* is not observed in any baryon decays, charm as well as beauty. In 2017 the LHCb collaboration observed a 3.3 standard deviation from *CP* symmetry in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays [53]. They proceed via tree and loop diagrams with similar contributions and through various intermediate resonances, increasing the possibility for *CPV* observation. However, this result was not confirmed after introducing 5.6 fb^{-1} of Run 2 data [54].

The first search of the *CPV* in the $\Xi_c^+ \to pK^-\pi^+$ decays was conducted using data collected in Run 1 [55]. Two model-independent techniques were applied: binned S_{CP} method and the k-nearest neighbour unbinned method. The results are consistent with *CP* symmetry. Although that, the $\Xi_c^+ \to pK^-\pi^+$ decays are promising in the first observation of *CPV* in charm baryons since in one region of the phase space the local agreement with *CP* symmetry was measured to be 2.7 standard deviations. Thus, it is worth continuing these measurements with data from Run 2.

In this thesis, the searches for *CPV* in the $\Xi_c^+ \to pK^-\pi^+$ baryon decays are presented. This time, data taken in Run 2 are analysed. Additionally, the new techniques are implemented which can improve sensitivities for searches for *CPV* and outperform the χ^2 test under conditions typical for this kind of studies [15].

3 Experimental framework

The CERN laboratory in Geneva, Switzerland, hosts one of the largest and most powerful particle accelerators in the world, namely the Large Hadron Collider (LHC). The LHC collider is an impressive engineering and scientific achievement, capable of accelerating protons to very high energies and then colliding these protons to generate new particles and analyze the properties of matter. Particles resulting from these collisions leave traces in detectors, and the analysis of these traces assists scientists in understanding the fundamental components of the universe and the interactions between them. The LHCb detector, one of the main experiments conducted at the LHC, focuses on studying the asymmetry between matter and antimatter. Its main goal is to better understand why the universe is predominantly composed of matter rather than equal amounts of matter and antimatter. In this section, it is described how the LHC collider enables experiments on an unprecedented scale of energy, and the LHCb detector allows for precise investigation of processes involving particles of different heavy flavours as well as the measured data statistics.

3.1 LHC accelerator

The LHC collider is a circular accelerator located at CERN. It boosts the particles in a 27-kilometre ring of superconducting magnets and accelerating structures at an energy of 6.5 TeV, generating collisions at an energy of 13 TeV. The two high-energy particle beams travel close to the speed of light in opposite directions in separate tubes kept in an ultrahigh vacuum. Then the proton beams are collided at four points around the accelerator ring, corresponding to the locations of four particle detectors – ATLAS, CMS, ALICE and LHCb. The ATLAS and the CMS (Compact Muon Solenoid) are two general-purpose detectors. Their main objective is to investigate a broad range of physics, from studying the Standard Model to searching for particles that could constitute dark matter. The third detector ALICE (A Large Ion Collider Experiment), is dedicated to studying quark-gluon plasma physics. The fourth LHCb detector aims to provide precise measurements of *CP* symmetry violation in decays of beauty and charm particles and analyse rare decays of these hadrons, including the indirect searches for new physics effects.

The LHC started accelerating protons on 10 September 2008. Till the end of 2011, it operated at center-of-mass energy of $\sqrt{s} = 7$ TeV. In 2012, this energy was increased to $\sqrt{s} = 8\,\text{TeV}$. In 2013-2015, the LHC was modernised and restarted in 2015 with $\sqrt{s} = 13\,\text{TeV}$ energy at center-of-mass. In this shape, the LHC operated till the end of 2018. After that once again the accelerator went to major upgrade, as well as all detectors. This time the energy was increased to the level of $\sqrt{s} = 14\,\text{TeV}$ at center-to-mass (originally planned in 1994). In 2022, the first tests of upgraded accelerating system have been started. The accelerator is expected to operate intermittently until 2035.

Period of time from 2010 to 2012 is called Run 1, whilst years 2015-2018 belong to Run 2. The cycle started in 2022 is called Run 3.

The expected number of events can be calculated using the following formula:

$$N = \mathcal{L}\sigma, \tag{14}$$

where $\mathcal{L}$ is the accelerator luminosity and σ is the cross section of a given process. For $\sqrt{s} = 13$ TeV energy at the center of mass, the cross-section for proton-proton interaction is estimated to be 75 mb and integrated luminosities in Run 1 and Run 2 were measured to be about 3 fb^{-1} and 6 fb^{-1}, respectively.

3.2 LHCb experiment

The LHCb experiment is one of the four large experiments operating at LHC. The construction of the LHCb detector [56–59] was designed to investigate the discrepancies between matter and antimatter in the decays of heavy particles containing b and/or c quarks as well as their rare interactions. The detector is a single-arm forward spectrometer and its scheme is presented in Fig. 3.

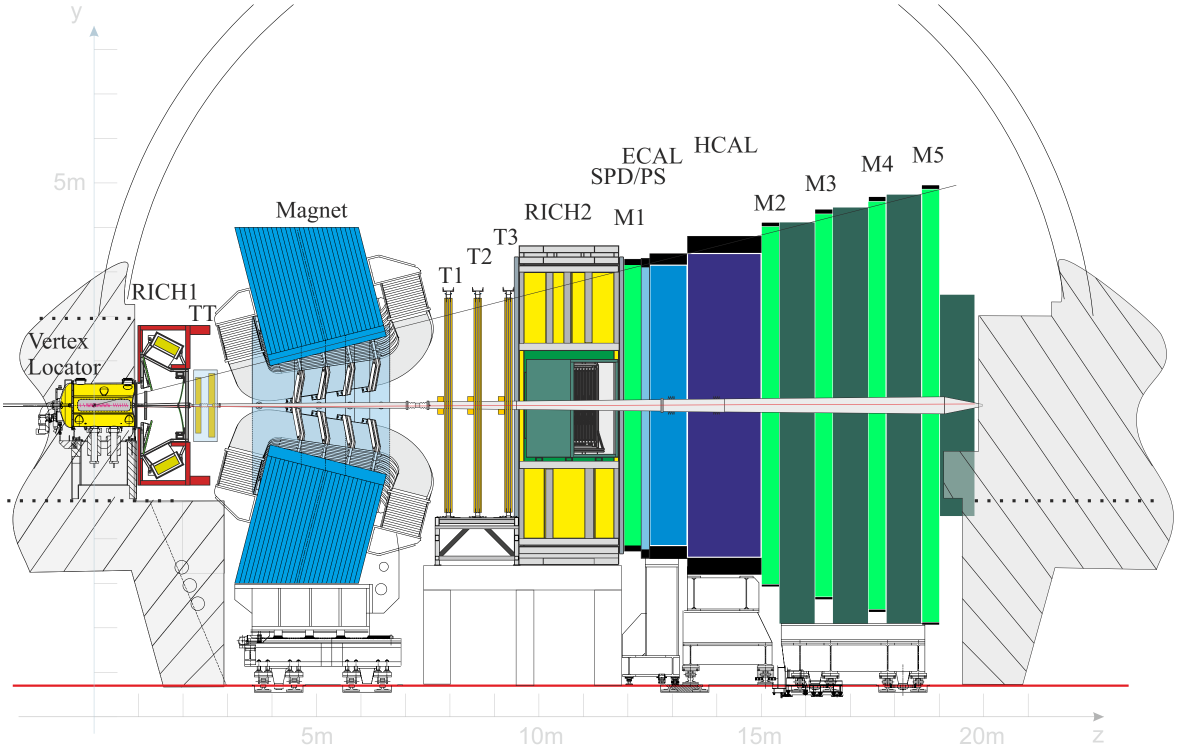

Figure 3: A scheme of the LHCb detector. The distances are given in respect to nominal pp interaction point.

The right-handed coordinate system adopted has the y-axis along the vertical and the z-axis along the beam. During the proton-proton collision, heavy quarks are produced through two gluons fusion. The pairs of $b\bar{b}$ and $c\bar{c}$ are predominantly made in a forward cone with very small θ and in a backward cone with large θ, where θ is the polar angle, defined as the angle between particle flight direction and beam axis. Due tothe LHCb detector geometry, it can only register half of the created particles. The other half is used in one of the trigger systems called VETO, and part of its signal is recorded in the vertex detector. The LHCb spectrometer covers the range of pseudorapidity $2 < \eta < 5$, where η is defined as:

$$\eta = -\ln \mathrm{tg} \frac{\theta}{2}. \tag{15}$$

Point where two proton beams collide is called the primary vertex (PV). This is also a point of production beauty and charm particles called as prompt. Newly created heavy hadrons travel from a PV a few millimetres to 1–2 centimetres, then decay. This decay point is called the secondary vertex (SV). It is also possible for several decays to happen. Should this occur, every secondary vertex must be separated from each other and treated individually. The detector component used for distinction between primary and secondary vertexes is VErtex LOcator (VELO), which surrounds the PV. The VELO measures so-called impact parameter (IP) which is defined as the closest distance from the primary vertex of the collision in a collider to a given track. For particles with transverse momentum $p_{\mathrm{T}} > 1\,\mathrm{GeV}/c$, IP is determined with precision greater than 35 μm. Impact parameter for particles produced at the secondary vertex is high as oppose to IP measured for particles created at primary vertex, which are expected to be close to 0.

The LHCb detector comprises a number of sub-detectors designed for measuring different characteristics of the particles produced by the colliding protons. Components which are essential for measuring *CP* asymmetry in $\Xi_c^+ \to pK^-\pi^+$ are described below. This characterisation refers to detector and its operation during the two data collection cycles Run 1 and Run 2.

3.2.1 Tracking system

The tracking system provides the measurements of the momenta (p) of charged particles with a relative uncertainty that varies from 0.5% at the lowest momentum values to 1.0% at momentum values of $200\,\mathrm{GeV}/c$. The minimum distance of a track to a PV, the IP is measured with a resolution of $(15 + 29/p_{\mathrm{T}})\,\mu\mathrm{m}$. The average tracking reconstruction efficiency is above 95% [58].

The tracking system consists of:

- VELO,
- the dipole magnet, which bends the traces of charged particles in the generated magnetic field with a curvature capacity of 4 Tm.
- the Tracker Turicensis (TT) located upstream of the magnet,
- the three tracking stations T1–T3 situated downstream of the magnet.

The VELO is a position sensitive silicon detector. It contains of 42 modules arranged along the beam, each providing a measure of the radial r and azimuth ϕ coordinates. The pitch within a module varies from 38 μm at the inner radius of 8.2 mm, increasing linearly to 102 μm at the outer radius of 42 mm. The VELO sensors are placed at a radial distance from the beam which is smaller than the aperture required by the LHC during injection. Hence, the modules are retracted by 29 mm in the horizontal direction during injection of the LHC beams and are subsequently moved back, following a fully automated procedure once stable conditions have been declared.

The TT uses silicon microstrip sensors with a strip pitch of 183 μm. The TT is about 150 cm wide and 130 cm high, with a total active area of around 8 m^2.

Three tracking stations T1-T3 consist of parts close to beam pipe (Inner Tracker, IT) and external sensor (Outer Tracker, OT). Former use silicon microstrips with a strip pitch of 198 μm and latter is drift-gas detector. The IT covers a 120 cm wide and 40 cm high cross-shaped region in the centre of the three tracking stations T1–T3. The total active area of the IT is approximately 4 m^2. The OT consists of approximately 200 gas-tight straw-tube modules with a drift-time read-out. Each module contains two staggered layers of drift-tubes with an inner diameter of 4.9 mm. A counting gas is a mixture of Argon (70%), $\mathrm{CO_2}$ (28.5%) and $\mathrm{O_2}$ (1.5%), which guarantees a drift time below 50 ns and a spatial resolution of 200 μm.

The magnetic field deflects oppositely charged particles in opposite directions. The configuration with the magnetic field pointing upwards (downwards), *MagUp* (*MagDown*), bends positively (negatively) charged particles in the horizontal plane towards the centre of the LHC ring. The polarity is periodically reverse throughout the data-taking to study and cancel detection asymmetries, which can be of the same or greater magnitude than the expected *CP* asymmetry.

3.2.2 Particle identification

Depending on type of particle, its identification is based on different subdetector, which are:

- two Cherenkov detectors,
- the calorimeter system,
- the muon chambers.

The two ring-imaging Cherenkov detectors are RICH 1 (located between VELO and magnet) and RICH 2 (situated just after three tracking stations T1–T3 and before calorimeters). Based on information from RICH 1 and RICH2 the protons, kaons and pions are identified. Their missidentification depends on momenta and is smaller than 5% for kaons and smaller than 3% for pions [57].

The calorimeter system is composed of sequentially placed: a Scintillating Pad Detector (SPD), a Preshower (PS), a shashlik type Electromagnetic Calorimeter (ECAL) and a Hadronic Calorimeter (HCAL). This system provides the identification of electrons, photons and hadrons along with their energies and positions, and selects candidates with high transverse energy (E_{T}) for the first trigger level (described in the next sections). The SPD is constructed from two planes of almost identical rectangular scintillators pads placed on either side of a 15 mm thick layer of lead. The ECAL has a sampling lead structure with a total thickness of 25 radiation lengths. In addition, a segmentation into three different sections has been chosen for the ECAL with a corresponding projective geometry for the SPD and PS detectors, meaning that all of their transverse dimensions scale with the distance from the interaction point. The HCAL is a device made from alternating layers of iron and scintillating tiles, as absorber and active material, respectively. Given the dimensions of the hadronic showers, the HCAL is divided transversely into two zones with different lateral dimensions: 130 mm width near the beam and 260 mm in the outer areas. The thickness of the HCAL is limited to 5.6 nuclear interaction lengths due to space constraints.

The muon chambers [57] provide an identification of muons as its name indicates. Their signals contribute to the first level trigger (described in the next sections). It is composed of five stations (M1—M5) of rectangular shape equipped predominantly with multiwire proportional chambers. Apart from the highest rate region of M1, where triple gas electron multiplier detectors are used, the four stations M2–M5 are placed sequentially after HCAL closing the detector. The M1 station is placed between RICH 2 and SPD. The full system comprises 1380 chambers and covers a total area of 435 m^2. The M2-M4 chambers are interleaved with 80 cm thick iron absorbers to select penetrating muons. The minimum momentum that a muon must have to travel through the five stations is about 6 GeV/c. The minimum p_{T} value for a single muon is required to be in the range of 1.48 to 1.76 GeV/c, while for a pair of two muons, their minimum combined p_{T}^2 should be in the range from 1.69 to 2.56 $(\mathrm{GeV}/c)^2$. The efficiency of muon identification is of the order of 97% [57].

3.2.3 Trigger

The trigger system used in Run 1 [56, 57] performed a two-level version of the full offline reconstruction. In the first stage, only charged particles with at least $1\,\mathrm{GeV}/c$ of p_{T} displaced from the PV were available. The muons requirements were slightly different: the threshold was lower, and they didn't have to be displaced. This first reconstruction phase enabled the bunch crossing rate to be decreased efficiently by approximately one order of magnitude. In the following second stage, charged particles with $p_T \geq 300\,\mathrm{MeV}/c$ were available to classify the bunch crossings.

Whole trigger system was redesigned in Run 2 to implement full offline event reconstruction [59]. A schematic diagram illustrating the trigger data flow in Run 2 is presented in Fig. 4. The LHCb trigger is constructed to allow data taking with minimal dead time at the total LHC bunch crossing rate of 40 MHz. A system of field-programmable gate arrays with a fixed latency of 4 s (the L0 trigger) makes decisions based on information from the calorimeter and muon systems to reduce the event rate to below 1 MHz.

The HLT is is fully programmable and frequently called as a software trigger. It is divided into two sublevels: HLT1 and HLT2. The HLT1 performs an inclusive selection of events based on one- or two-track signatures, on the presence of muon tracks displaced from the PVs, or on dimuon combinations in the event. Selected events are then buffered to disk storage in the online system. They can be further processed during inter-fill cycles, and the detector can be calibrated and aligned run-by-run before the HLT2 stage. Once the alignment and calibration processes are finished, events are passed to HLT2, where a full event reconstruction is performed. After the HLT, events are stored and later processed with a more accurate alignment and calibration of the sub-detectors. This part of the reconstruction and subsequent selection of interesting events is called offline reconstruction and particle selection.

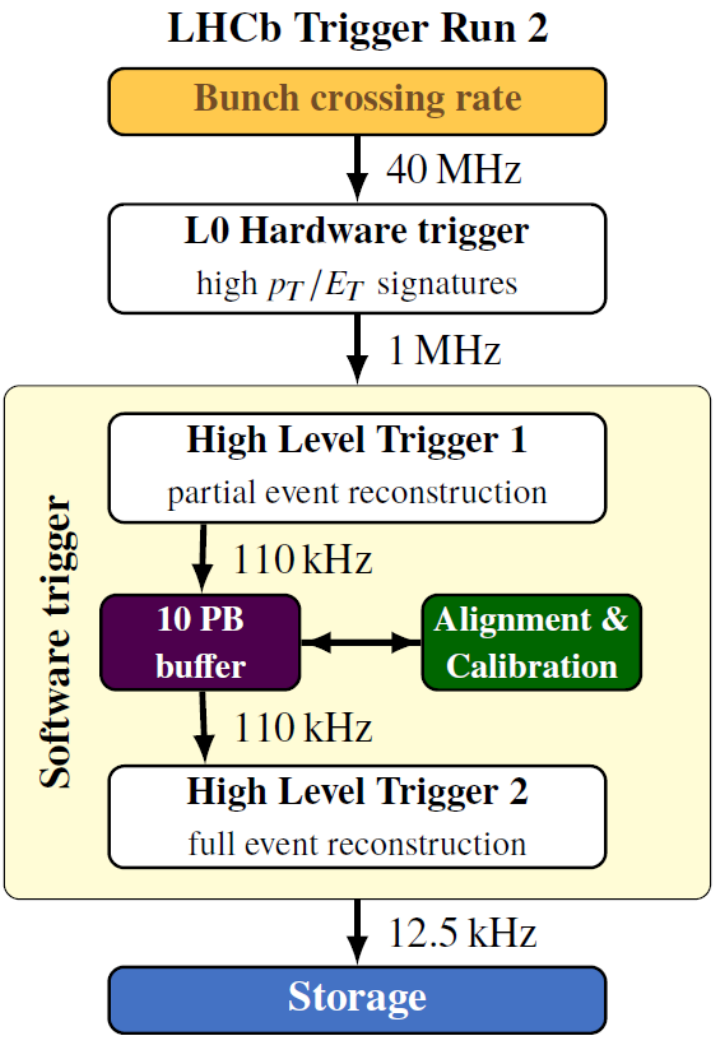

Figure 4: Overview of the LHCb trigger system in Run 2.

3.3 Data statistics

The cross section of $b\bar{b}$ production measured at LHCb for centre-of-mass energy equal to 7 TeV is about 75 μb [60]. The cross section of $c\bar{c}$ production measured for the same energy is about 20 times bigger and is around 1400 μb [61]. However, for centre-of-mass energy equal to 13 TeV accessible in Run 2 cross section of $c\bar{c}$ production was measured to be about 2400 μb [62] which is 30 times higher than $b\bar{b}$ production cross section. Hence, one can expect data samples consisting extremely huge amount of charm mesons and barions, which are never collected in world till now. These data samples allow for precise searches for *CP* violation, which were never achieved before.

As far as charm baryon decays are concerned, the most accessible are ones containing a single proton and several kaons or pions in the final state, such as CF $\Lambda_c^+ \to pK^-\pi^+$ and SCS $\Xi_c^+ \to pK^-\pi^+$ decays with identical final states. An identification of such decays is an experimental challenge, due to the relatively short lifetimes of them. It reduces the efficiency of the selection based on separating the decay vertex of the charm particle from the other vertices, increasing the combinatorical background, which is mainly made up of pions and kaons of low momenta. Although, the selection of charm baryons is mostly determined by good discrimination of charged particle tracks and by precise reconstruction of the secondary vertex. The presence of a proton in the final state is both valuable in event selection and experimentally challenging for *CP* violation measurements. This is because one has to additionally consider the proton-antiproton detection asymmetry, which is related to the different reconstruction efficiency of protons and antiprotons. That asymmetry can be on the order of one-two percent and, therefore, is more significant than the expected *CP* asymmetry.

3.4 The upgrade and future of LHCb

The LHCb proved itself to be a general purpose forward physics detector with a rich physics program extending significantly beyond its primary goals. It was designed to collect data for five years at $\mathcal{L} = 2 \times 10^{32}\mathrm{cm}^{-1}\mathrm{s}^{-1}$ yet this value was exceeded at the end of Run 2 and detector operated successfully at twice the previous luminosity: $4 \times 10^{32}\mathrm{cm}^{-1}\mathrm{s}^{-1}$. To overcome experimental challenges due to increased luminosity and collect more data LHCb undergone a major upgrade, so-called LHCb Upgrade, during the "Long Shutdown 2" of LHC (LS2), which started in 2018 and ended in 2022, when LHCb detector has been relaunched. The aim is to collect 50 fb^{-1} of data till 2028. The target peak luminosity $\mathcal{L} = 2 \times 10^{33}\ \mathrm{cm}^{-1}\mathrm{s}^{-1}$ is five times higher than luminosity recorded in Run 2. It is estimated that the total luminosity in Run 3 (2022-2024) will be 23 fb^{-1}, and in Run 4 (2026-2030) will be 50 fb^{-1}.

During the LHCb Upgrade, almost all of the subsystems were redesigned. One of the most significant changes of the upgrade is the 40 MHz readout ability by all of the LHCb readout electronics. This readout is also prepared to manage the multi-TB/s data stream. The VELO detector was replaced by a new subdetector constructed of 52 rectangular modules perpendicular to the particle beam with 55×55 μm pixel sensors resulting in timing resolution of 25 ns [63]. The RICH detectors and tracking stations are also upgraded to maintain performance for higher luminosities. Total error on the Cherenkov angle for the RICH1 decreased from 0.99 mrad to 0.80 mrad and the total material budget in the RICH1 has been reduced from $\sim 7.6\%$ to $\sim 4.8\%$ radiation lengths [64]. The L0 hardware-based trigger is removed, and the new trigger system is entirely software-based. The calorimeters and muon chambers are not undergoing substantial upgrades, except for replacing the front-end electronics.

4 Data sets and selection criteria

The $\Lambda_c^+ \to pK^-\pi^+$ and $\Xi_c^+ \to pK^-\pi^+$ decay candidates are selected online by the dedicated software trigger lines. These selections are not sufficient and more precise requirements of data are necessary due to topology of producing Λ_c^+ and Ξ_c^+ baryons. This further requirements are called offline selections and aims to reject background events as the best as possible while preserving as many Λ_c^+ and Ξ_c^+ decays as possible. The stages of this strategy are described in this section. The criteria for elimination of detector effects are also presented in aspect the effective measurements of *CP* violation effects.

4.1 Data sets

The *CP* violation searches are based on data collected in *pp* collisions at $\sqrt{s} = 13$ TeV by the LHCb detector in 2016, 2017 and 2018, commonly known as Run 2 data. These data samples correspond approximately to 4.0 fb^{-1} in total. The specific values of the luminosity in each year of data taking are presented in Table 1.

Table 1: Luminosity in each year of data taking divided into two types of magnet polarization: "Up" and "Down".

	Down	Up
2016	845 pb^{-1}	795 pb^{-1}

	Down	Up
2017	525 pb^{-1}	500 pb^{-1}

	Down	Up
2018	618 pb^{-1}	612 pb^{-1}

The luminosities for both types of magnet polarization are in similar orders. The luminosity of the data collected in 2016 is the highest due to the existence of the HLT2 trigger which is not used in 2017 and 2018. The data samples have already been preselected at the trigger stage. However, this selection is not enough and further offline selection cuts are used. Its goal is to reduce background events while keeping as many Λ_c^+ and Ξ_c^+ decays as possible.

4.2 Selection strategy

The choice of each offline selection cut is based on the improvement of the Figure of Merit parameter (FoM), which is defined as follows:

$$\text{FoM} = \frac{S}{\sqrt{S+B}}, \tag{16}$$

where S and B are considered as the numbers of candidates below signal (Λ_c^+ or Ξ_c^+ decays) and background functions, respectively. In addition to the FoM parameter, for each set of cuts for a given variable, efficiency (defined as the ratio of the number of events left after the cut to the number of events before the cut) is monitored. The S and B numbers are taken from fitted model to the mass distributions obtained after each cut.

The fitted model to the mass distribution consists of two functions: a single Gauss for the signal or double Gauss to better describe the mass peak if needed and a Chebyshev polynomial function to describe the background. A Gaussian function has two free parameters: the mean (μ) and the width (σ). The mass (below represented by x parameter) is the observable. In case of double Gauss, it is a linear sum of two Gaussian functions. Then the μ parameter is in common one between the two Gaussian functions, while the widths σ_1 and σ_2 may differ. The single Gauss is defined as:

$$G(x, \mu, \sigma) = exp(-\frac{(x-\mu)^2}{2\sigma^2}) \tag{17}$$

and then the double Gauss is defined as:

$$f(x) = sig1frac \cdot G(x, \mu, \sigma_1) + (1 - sig1frac) \cdot G(x, \mu, \sigma_2), \tag{18}$$

where the $sig1frac$ parameter determines fraction of the first Gaussian function.

The Chebyshev function has two coefficients: p_0 and p_1, and is defined as:

$$g(x) = p_0 + p_1 \cdot x. \tag{19}$$

The signal and background functions are combined as a linear sum to form a total model $h(x)$, where each component is weighted by their respective yields $nsig$ and $nbkg$, corresponding to the number of signal candidates and the number of background candidates, respectively:

$$h(x) = nsig \cdot f(x) + nbkg \cdot g(x). \tag{20}$$

A χ^2 fit is performed in the data sets by minimizing a normalized version of $h(x)$ (as a probability density function). The χ^2 is minimized with MIGRAD algorithm, inside MINUIT package, with all parameters allowed to float, and the covariance matrix is computed with HESSE function, all implemented in RooFit framework [65].

The reconstructed mass (M) of $pK^-\pi^+$ selected online by the dedicated software trigger for Ξ_c^+ and Λ_c^+ for each year of data taking are presented in Figs 5 and 6, respectively. These selections are based on three charged particles identified as proton, kaon and pion which are combined to form a good-quality decay vertex, detached from any PV. The PV with the smallest value of χ^2 associated to the decay candidate, where IP χ^2 defined as the difference in the vertex-fit χ^2 of the PV reconstructed with and without the particle under consideration, in this case the Ξ_c^+ or Λ_c^+ candidate. The models are not superimposed because of the difficulty of obtaining a good fit to large and contaminated samples of data not yet subjected to the offline selection. The models are fitted to the final data samples in Sec. 4.7. There are no significant differences in reconstructed Ξ_c^+ (or Λ_c^+) mass distributions for different years of data taking. The observed peaks correspond to the Ξ_c^+ and Λ_c^+ decays. The reconstructed background is relatively large (the background in Λ_c^+ sample is significantly smaller than in Ξ_c^+ sample). It shows that further offline selection cuts are neccessary aiming to reduce the background simultaneously increasing the efficiency of baryon reconstruction. At this first stage in total, there are around 447 million Λ_c^+ candidates and 252 million Ξ_c^+ candidates including background events.

Further requirements are applied among others on the Ξ_c^+/Λ_c^+ decay time; on the angle between the reconstructed Ξ_c^+/Λ_c^+ momentum vector and the vector connecting the PV to the decay vertex; on the χ^2 of the Ξ_c^+/Λ_c^+ decay vertex fit; and on the momentum, the transverse momentum and the χ^2 of IP of the the Ξ_c^+/Λ_c^+ candidate and its decay products. Below describes in detail how the offline cuts are selected. Since there are no significant differences between the years of data taking and the goal is to optimize the selection to choose as many the Ξ_c^+ decays as possible, the cuts are developed using sample with the Ξ_c^+ decays collected in 2018.

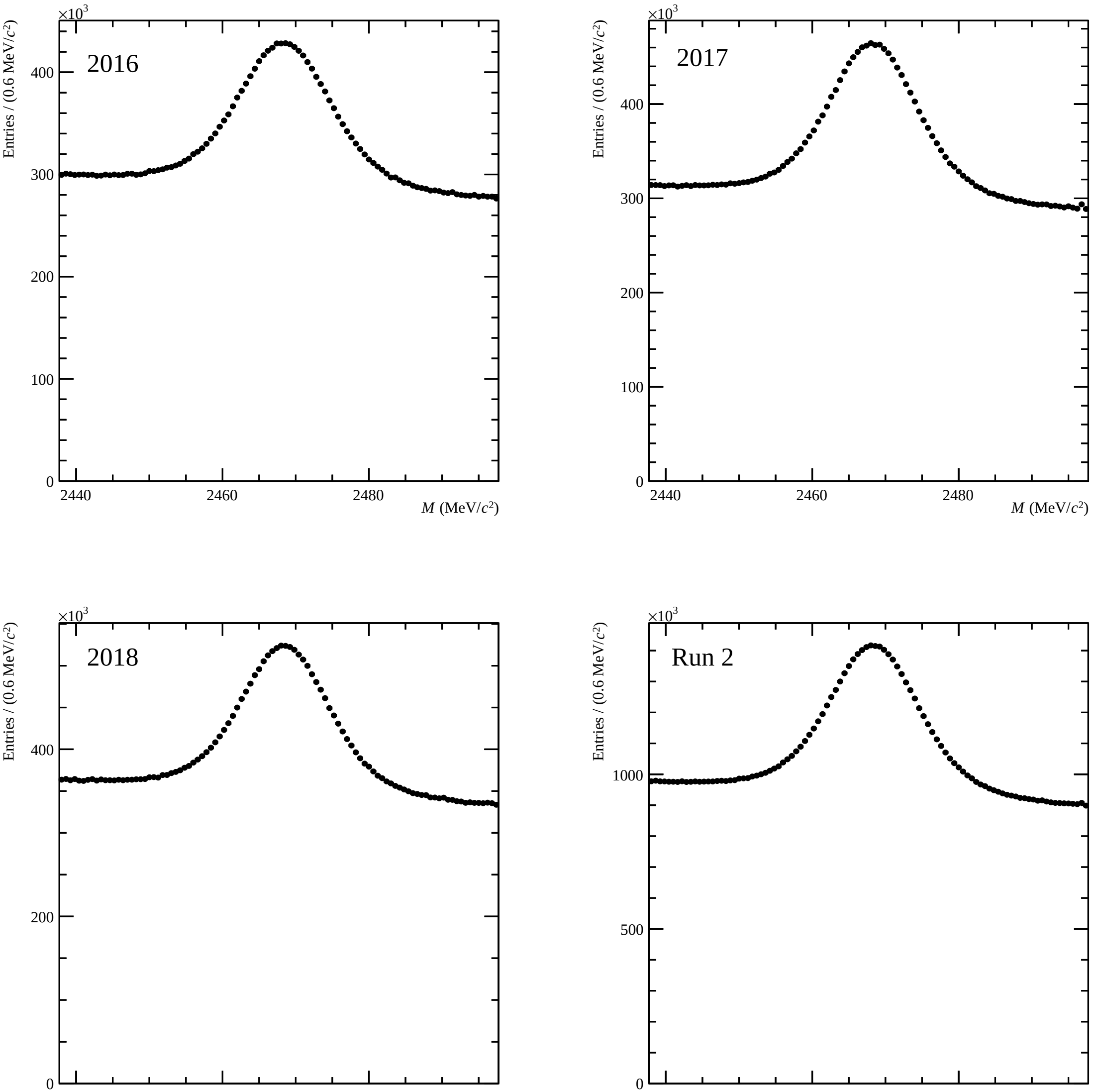

Figure 5: Reconstructed mass (M) of $pK^-\pi^+$ before offline cuts for Ξ_c^+ samples in 2016 (top left), 2017 (top right), 2018 (bottom left) and whole Run 2 data sets (bottom right).

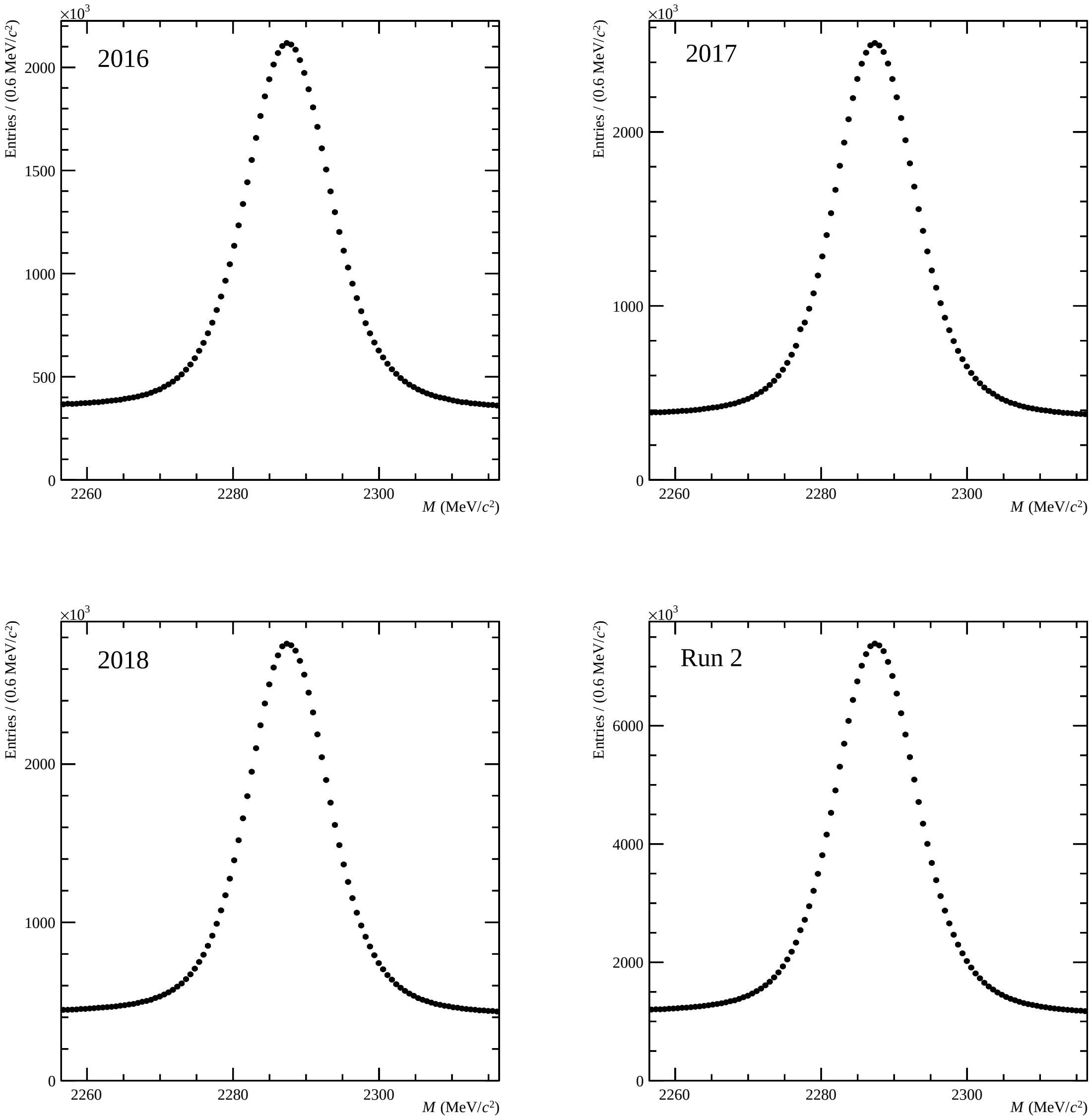

Figure 6: Reconstructed mass (M) of $pK^-\pi^+$ before offline cuts for Λ_c^+ samples in 2016 (top left), 2017 (top right), 2018 (bottom left) and whole Run 2 (bottom right) data sets.

Determination of each offline cut is based on calculation efficiency of this cut, which is defined as:

$$\text{efficiency} = \frac{\text{number of entries after cut}}{\text{number of all entries before cut}}. \tag{21}$$

Figures 11-21 show the distributions of variables for which the selection cuts are applied, and plots of its efficiency as a function of these variables. For better readability, the Y axis on each variable distribution is drawn in a logarithmic scale. Ranges of the distributions for which efficiency is a constant indicate that online cuts have previously been applied to the variable. The vertical red and blue lines indicate the selected value of cuts. Several potential cuts are considered for each variable. For all of them, the model is fitted to the mass distribution and from this fit the number of signal and background candidates are taken and the value of the FoM parameter is calculated. Based on these four numbers and the efficiency, the final cuts are selected.

The first considered variables are those related to particle identification. There are two methods to identify particles combining the information from the sub-detectors described in Sec. 3.2.2. One method is based on calculation quantity which is the change in log-likelihood ($\Delta log\mathcal{L}(h-h')$) between the h and h' hypotheses for the particle under consideration (PID). The PID is calculated with respect to the pion. The second method is based on neural network quantity which is the output of multivariate techniques created by combining tracking and PID information (ProbNN). The ProbNN results in a single probability value for each particle hypothesis. The PID for protons is shown in Fig. 7, whilst the ProbNN is shown in Figs 8, 9 and 10 for protons, kaons and pions, respectively. The PID for kaons and pions are already implemented in the online selection process and further for them the ProbNN variables are considered only. For protons, the PID is chosen to be higher than value 10 corresponding to the efficiency about 70% and the ProbNN is chosen to be higher than 0.5 corresponding to the efficiency about 60%. For both kaons and pions, the ProbNN is higher than 0.1 keeping similar efficiency about 95%. After these cuts the FoM improves by 25% reducing combinatorial background by about 50%. This test shows that the bigger problem is with proton reconstruction than with kaon and pion. This problem is also evident in other LHCb analyses. Therefore, one should expect greater problems with removing the detector effects associated with protons than with kaons and pions, which are described in Sec. 4.6.

The next considered cuts are the IPχ^2 for protons, kaons, pions and charm baryon candidates. Their distributions are shown in Figs. 11, 12, 13 and 14, respectively. Due to the decay topology of prompt Ξ_c^+ baryon with respect to the primary vertex its IP should be close to zero originating from the primary vertex, whilst its daughters (protons, kaons and pions) should originate from the secondary vertex, i.e. vertex of the Ξ_c^+ decay. The associated IPχ^2 carries the values of IP but cutting to IPχ^2 is more effective in reducing the background. Therefore, the cut limits the maximum IPχ^2 value for charm baryon and for daughter particles IP restrictions are set from below as protons, kaons and pions should originate from the secondary vertex. The cut on the IPχ^2 discards the most background events and have the greatest impact on the statistics of the final sample. For example, the cut on Ξ_c^+ IP$\chi^2 < 12$ removes around 60% background events and only 5% signal candidates. The IPχ^2 and efficiency distributions for daughter particles are shown on Figs.

11-13. For protons, pions and kaons the IP$\chi^2 > 9$ is chosen. The efficiency for this cut is around 90%. The IPχ^2 cuts on proton, kaon and pion remove around 15% signal and 15% background events and FoM improves by about 35%. The signal candidates (S), the background candidates (B), FoM and efficiency values for the Ξ_c^+ IPχ^2 is presented in Table 2. The IP$\chi^2 < 15$ cut keeps the highest number of background events. For the IP$\chi^2 < 10$ the FoM is the lowest and the number of signal events is too reduced. Finally, the IP$\chi^2 < 12$ is chosen since of its efficiency and the number of signal events larger than the IP$\chi^2 < 8$ cut.

Table 2: Number of signal (S) and background (B) events, FoM and efficiency for different cuts on Ξ_c^+ IPχ^2.

cut	< 8	< 10	< 12	< 15
S	37k	35k	38k	38k
B	244k	270k	288k	313k
FoM	69	63	67	64
efficiency	0.65	0.70	0.76	0.80

One of very effective cut in background elimination is related with flight distance (FD) of the Ξ_c^+ shown in Fig. 15. The S and B numbers, the FoM and the efficiency values for this variable are presented in Table 3. The FD $\chi^2 < 2000$ is chosen as it reduces the number of background events to the greatest extent. The FoM parameter for all considered cuts is very similar and it is improved by about 17%. The efficiency for the selected cut is 84%.

Table 3: Number of signal (S) and background (B) events, FoM and efficiency for different cuts on Ξ_c^+ FDχ^2.

cut	< 2000	< 4000	< 6000
S	34k	37k	37k
B	471k	528k	552k
FoM	48.5	48.8	48.7
efficiency	0.84	0.92	0.96

The next very effective cuts in terms of background reduction are the ones associated with momenta of daughter particles. The proton, kaon and pion momenta and their efficiency distributions are shown in Figs 16, 17 and 18, respectively. The S and B numbers, the FoM and the efficiency values for proton momentum are presented in Table 4. Two of the analysed cuts ($10 < p < 80$ and $10 < p < 100$) remove too few background events. The $15 < p < 100$ GeV/c is chosen as the number of signal events, the FoM and the efficiency are larger for it. This cut removes roughly 25% of the background events. The kaons and pions are easier to reconstruct than protons and therefore looser constraints are imposed on them than on the proton. In this case, the selected cut on momentum is $3 < p < 150$ GeV/c. The chosen cut reduces background events by around 10%. The signal events are reduced by 5% and the FoM improves by around 8%.

Table 4: Number of signal (S) and background (B) events, FoM and efficiency for different cuts on proton momentum (p). There are two efficiency values: the first corresponds to limitations from below and the second to the limit on the maximum value.

cut [GeV/c]	$10 < p < 80$	$15 < p < 80$	$10 < p < 100$	$15 < p < 100$
S	36k	33k	37.5	34k
B	627k	533k	651k	556k
FoM	44.3	43.5	45	44.5
efficiency	1 ; 0.95	0.90 ; 0.95	1 ; 0.98	0.90 ; 0.98

One of the quantities analysed for the charm baryon which is also effective in removing background events is the transverse momentum p_T shown in the Fig. 19. The S and B numbers, the FoM and the efficiency values are presented in Table 5. For Ξ_c^+, the cut $3 < p_T < 16$ GeV/c rejects too little background events. The cut $4 < p_T < 16$ GeV/c is chosen because it leaves more signal events and has a higher efficiency. The signal events are reduced by around 20% and the background events are reduced more by about 40% and the FoM is slightly improved by 2%. This cut eliminates mostly combinatorial background, i.e. number of candidates per event.

Table 5: Number of signal (S) and background (B) events, FoM efficiency for different cuts on Ξ_c^+ p_T. There are two efficiency values: the first corresponds to limitations from below and the second to the limit on the maximum value.

cut [GeV/c]	$4 < p_T < 15$	$3 < p_T < 16$	$4 < p_T < 16$
S	28k	36k	29k
B	439k	620k	440k
FoM	42	45	42
efficiency	0.99 ; 0.99	0.98 ; 1	0.99 ; 1

The next considered variable is cosine of the angle between the reconstructed Ξ_c^+ momentum vector and the vector connecting the PV to the decay vertex (DIRA) shown in the Fig. 20). The final cut on this variable is chosen to be DIRA > 0.99995. It removes about 10% of the background entries and the signal events and the FoM are almost intact. The efficiency of this cut is 99%.

In the case of charm baryons, cuts on pseudorapidity (η) are also considered. The η and efficiency distributions are shown in Figs. 21. The efficiency of the chosen cuts $\eta > 2$ and $\eta < 4.5$ is 99%. The signal and background events are reduced by about 8%.

The last considered variable is the Ξ_c^+ Vertex χ^2 $/ndof$ shown in Fig. 22. The S and B numbers, the FoM and the efficiency values are presented in Table 6. Two cuts are analysed Vertex $\chi^2/ndof < 8$ and Vertex $\chi^2/ndof < 10$. The former removes around 15% background entries and improves the FoM by about 15% and the latter reduces background entries by 9% and improves the FoM by 10%, while the number of signal events after applying either of the cuts is almost the same. Ultimately the cut Vertex $\chi^2/ndof < 8$ is chosen.

Table 6: Number of signal (S) and background (B) events, FoM efficiency for different cuts on Ξ_c^+ Vertex χ^2 $/ndof$.

cut	Vertex χ^2 $/ndof < 8$	Vertex χ^2 $/ndof < 8$
S	38.7k	38.5k
B	630k	670k
FoM	47	45
efficiency	0.95	1

The cut on χ^2 separation from related PV > 20 is based on other LHCb analyses. They remove around 8% signal events and 14% background events.

The list of final offline selection cuts is presented in Table 7 and the reconstructed mass (M) of $pK^-\pi^+$ in the Ξ_c^+ and Λ_c^+ 2018 samples are shown in Figs. 23 and 24. Comparing the obtained mass distribution for 2018 data sample with the one before the offline selection (Figs. 5 and 6) for the same year of the data taking, it can be observed that the background events for Ξ_c^+ sample are reduced by about 90% and the FoM is increased by about 80% and in case of Λ_c^+ background are reduced by around 90% and FoM is increased by about 10%.

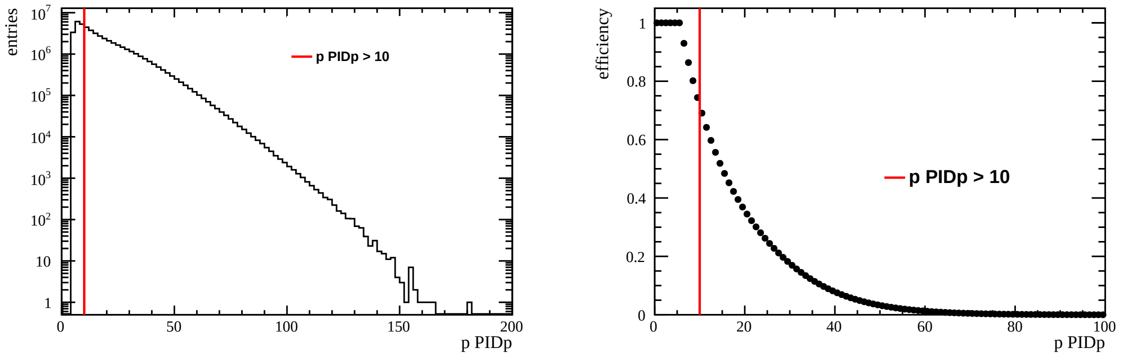

Figure 7: The PID distribution for **proton** (p) for 2018 Ξ_c^+ Down data sample (left) and corresponding efficiency as a function of PID (right). The chosen cut PID(p) > 10 is shown by vertical red line.

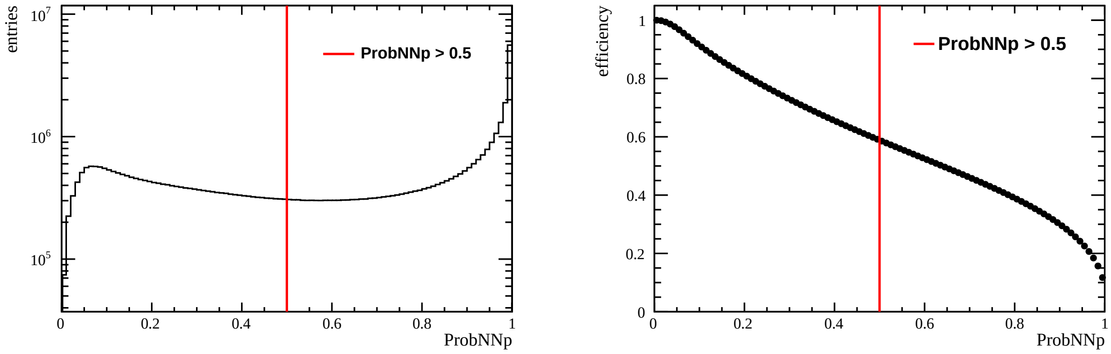

Figure 8: The ProbNN distribution for **proton** (p) for 2018 Ξ_c^+ Down data sample (left) and corresponding efficiency as a function of ProbNN (right). The chosen cut ProbNN(p) > 0.5 is shown by vertical red line.

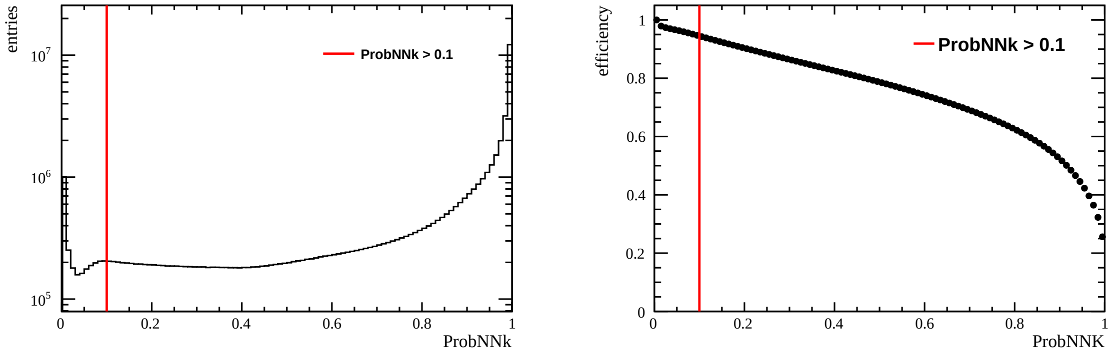

Figure 9: The ProbNN distribution for **kaon** (K) for 2018 Ξ_c^+ Down data sample (left) and corresponding efficiency as a function of ProbNN (right). The chosen cut ProbNN(K) > 0.1 is shown by vertical red line.

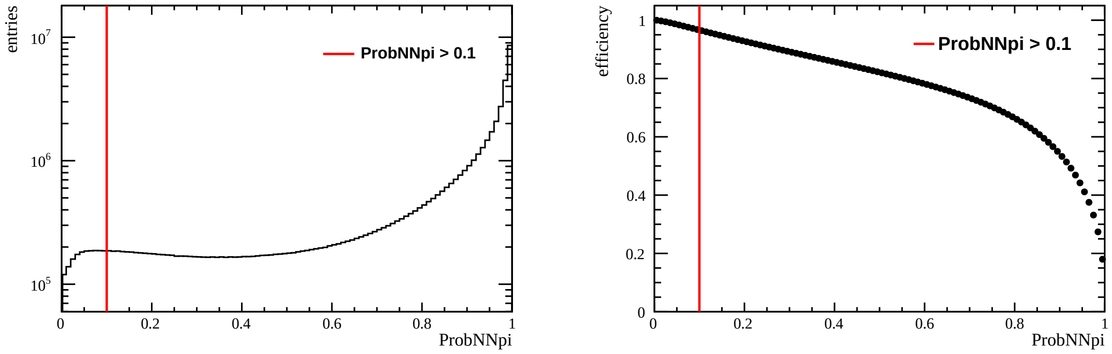

Figure 10: The ProbNN distribution for **pion** (π) for 2018 Ξ_c^+ Down data sample (left) and corresponding efficiency as a function of ProbNN (right). The chosen cut ProbNN(π) > 0.1 is shown by vertical red line.

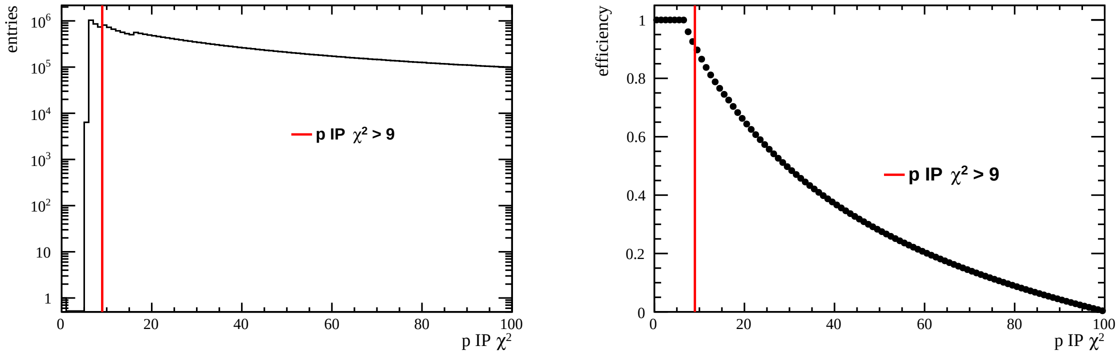

Figure 11: The IPχ^2 distribution for **proton** (p) for 2018 Ξ_c^+ Down data sample (left) and corresponding efficiency as a function of IPχ^2 (right). The chosen cut IP$\chi^2 > 9$ is shown by vertical red line.

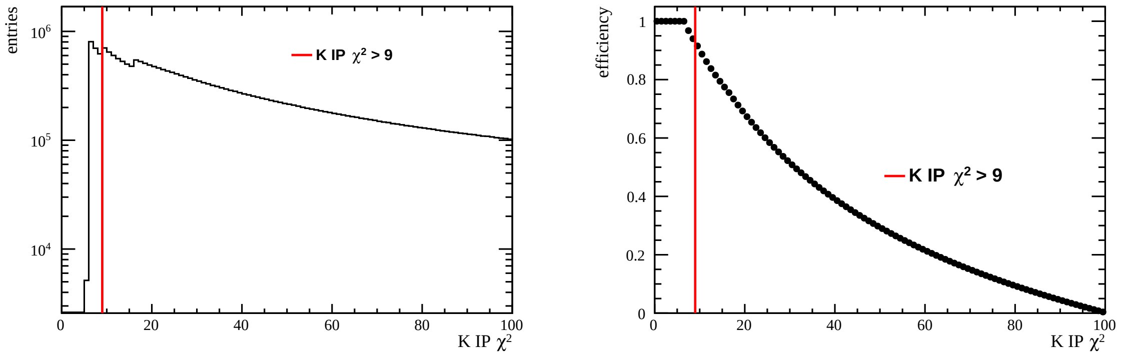

Figure 12: The IPχ^2 distribution for **kaon** (K) for 2018 Ξ_c^+ Down data sample (left) and corresponding efficiency as a function of IPχ^2 (right). The chosen cut IP$\chi^2 > 9$ is shown by vertical red line.

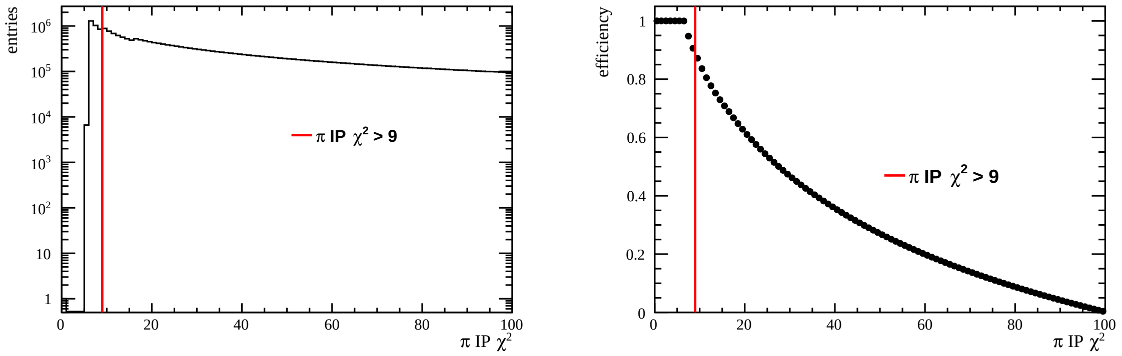

Figure 13: The IPχ^2 distribution for **pion** for 2018 Ξ_c^+ Down data sample (left) and corresponding efficiency as a function of IPχ^2 (right). The chosen cut IP$\chi^2 > 9$ is shown by vertical red line.

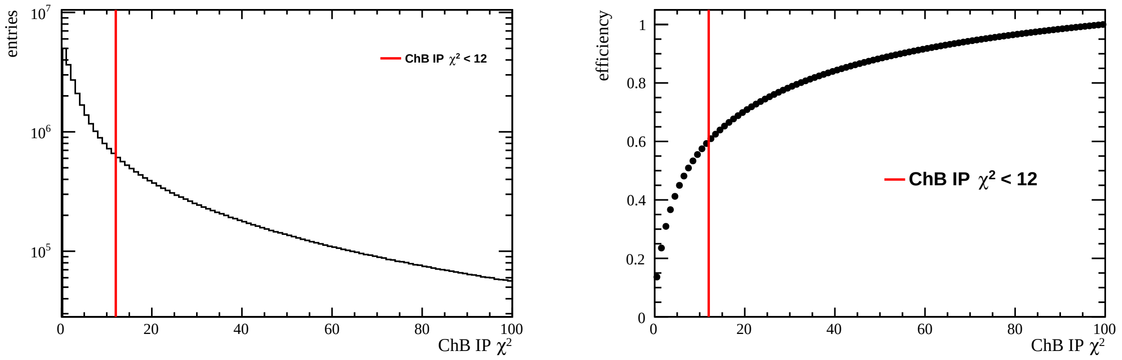

Figure 14: The IPχ^2 distribution of Ξ_c^+ measured in 2018 Down data sample (left) and corresponding efficiency as a function of IPχ^2 (right). The chosen cut IP$\chi^2 < 12$ is shown as vertical red line.

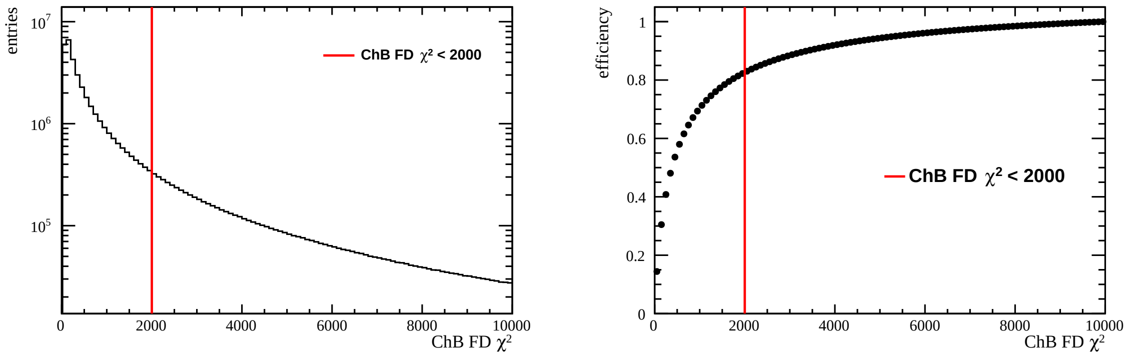

Figure 15: The FD χ^2 distribution of Ξ_c^+ measured in 2018 Down data sample (left) and corresponding efficiency as a function of FDχ^2 (right). The chosen cut FD$\chi^2 < 2000$ is shown as vertical red line.

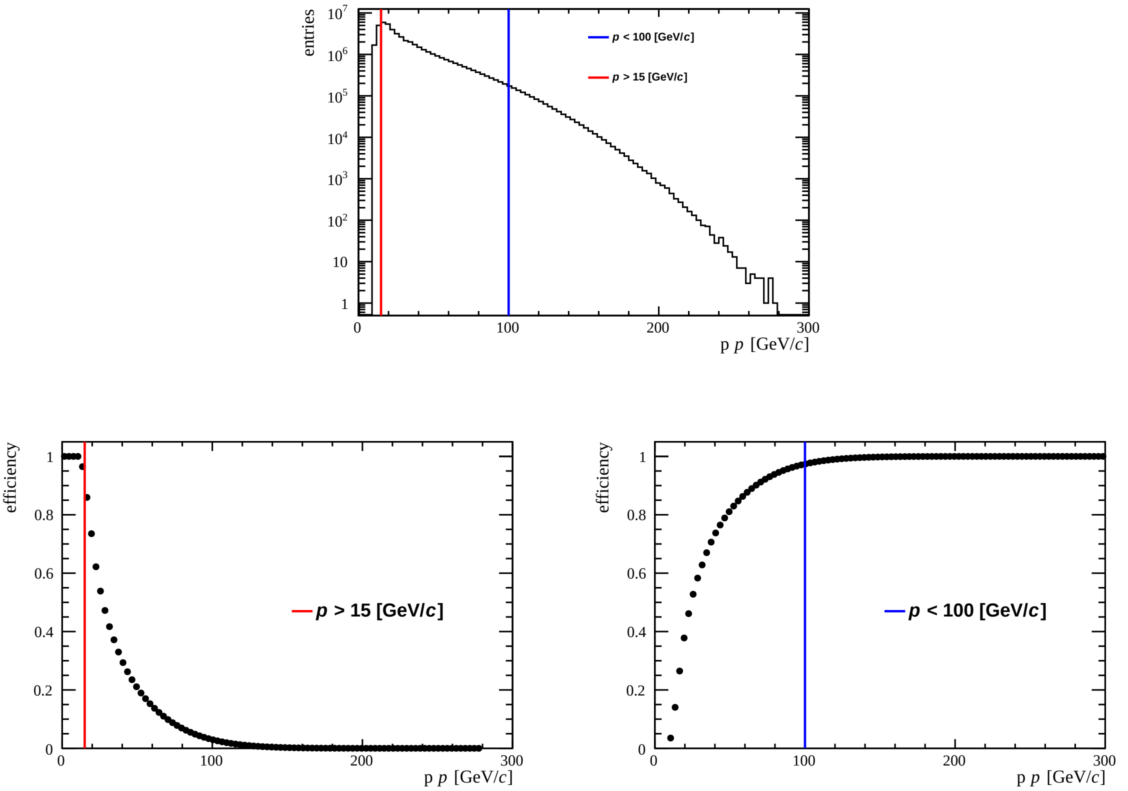

Figure 16: The p distribution for **proton** for 2018 Ξ_c^+ Down data sample (top) and corresponding efficiency as a function of p for minimum limit (bottom left) and maximum limit (bottom right). The chosen cuts $p > 15\,\text{GeV}/c$ and $p < 100\,\text{GeV}/c$ are shown as vertical red and blue lines, respectively.

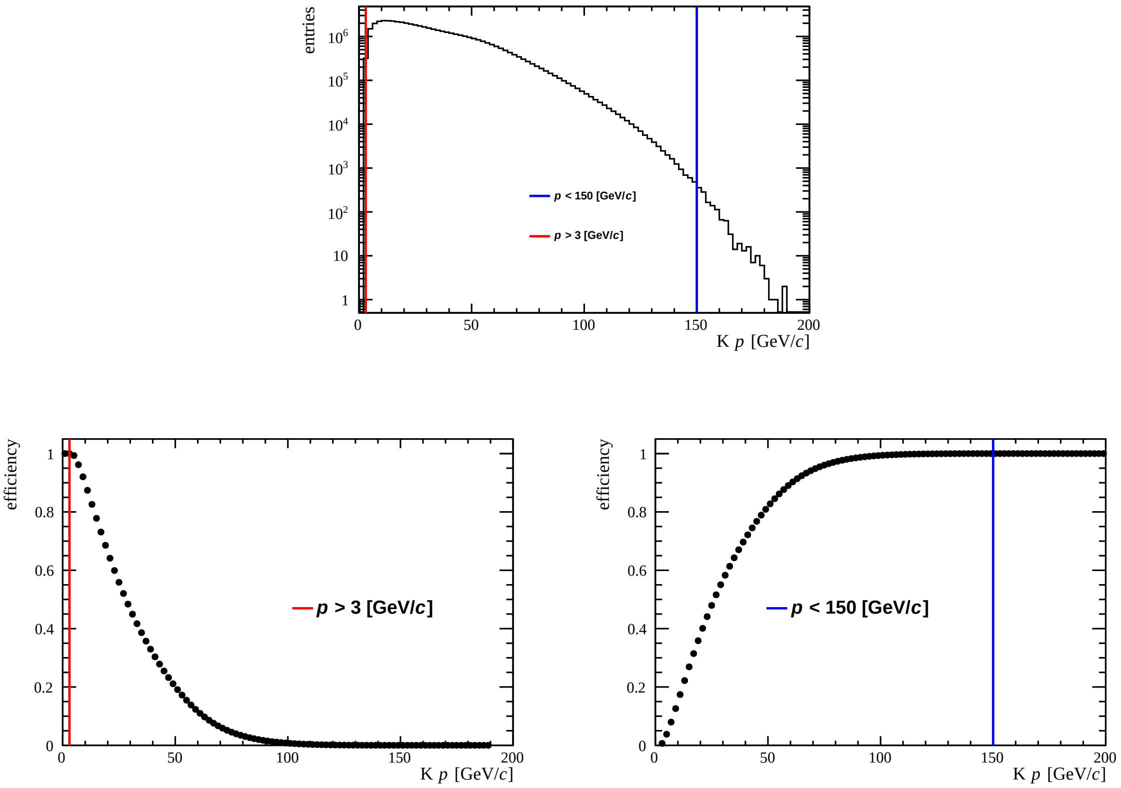

Figure 17: The p distribution for **kaon** for 2018 Ξ_c^+ Down data sample (top) and corresponding efficiency as a function of p for minimum limit (bottom left) and maximum limit (bottom right). The chosen cuts $p > 3\,\text{GeV}/c$ and $p < 150\,\text{GeV}/c$ are shown as vertical red and blue lines, respectively.

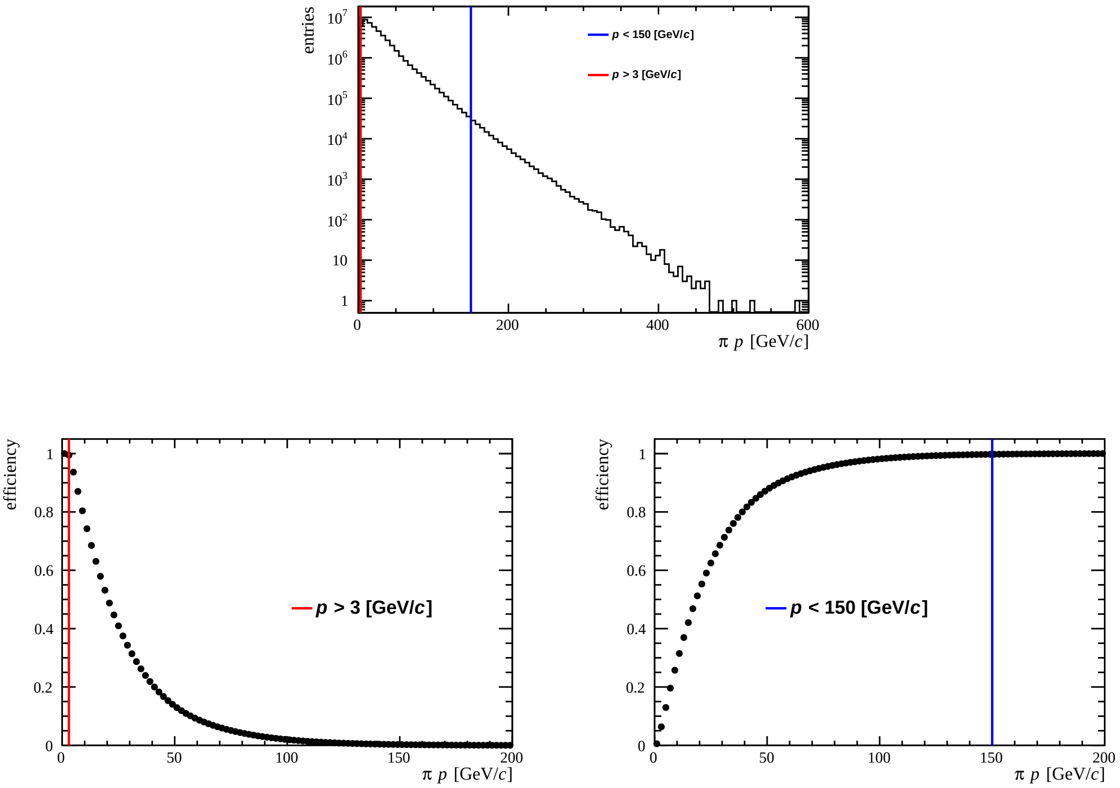

Figure 18: The p distribution for **pion** for 2018 Ξ_c^+ Down data sample (top) and corresponding efficiency as a function of p for minimum limit (bottom left) and maximum limit (bottom right). The chosen cuts $p > 3\,\text{GeV}/c$ and $p < 150\,\text{GeV}/c$ are shown as vertical red and blue lines, respectively.

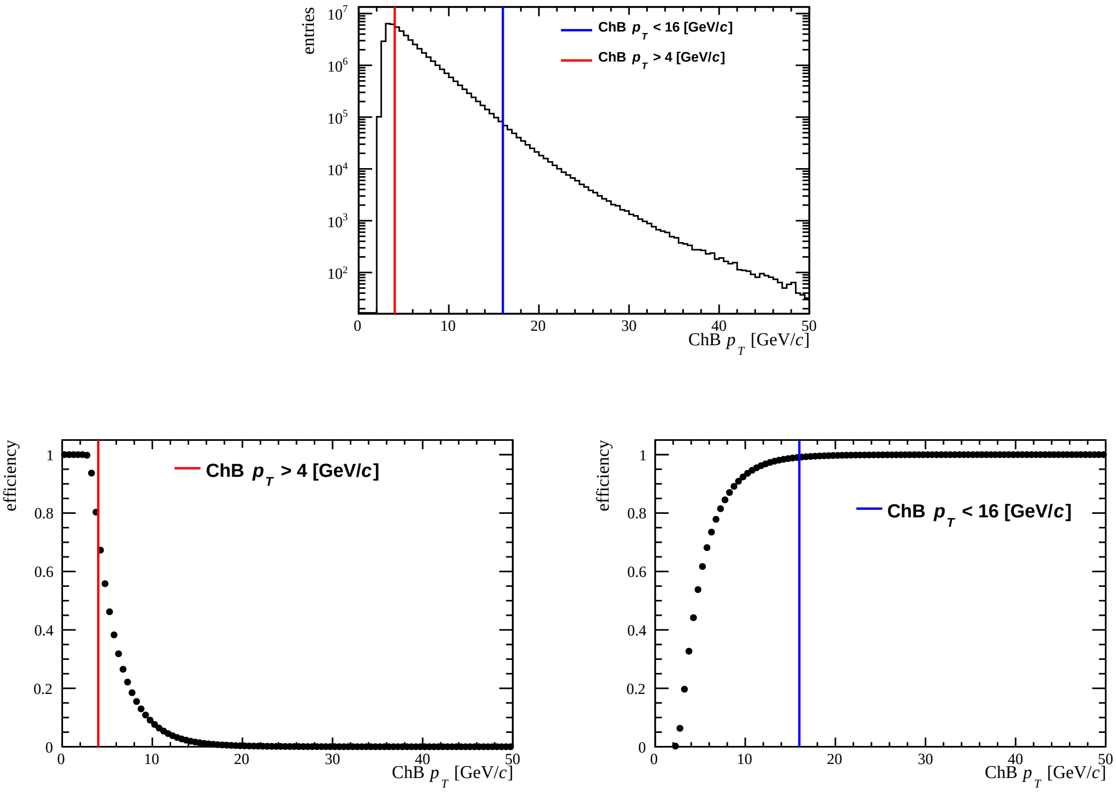

Figure 19: The transverse momentum (p_T) distribution of Ξ_c^+ measured in 2018 Down data sample (top) and corresponding efficiency as a function of p_T for minimum limit (bottom left) and maximum limit (bottom right). The chosen cuts $p_T > 4\,\text{GeV}/c$ and $p_T < 16\,\text{GeV}/c$ are shown as vertical red and blue lines, respectively.

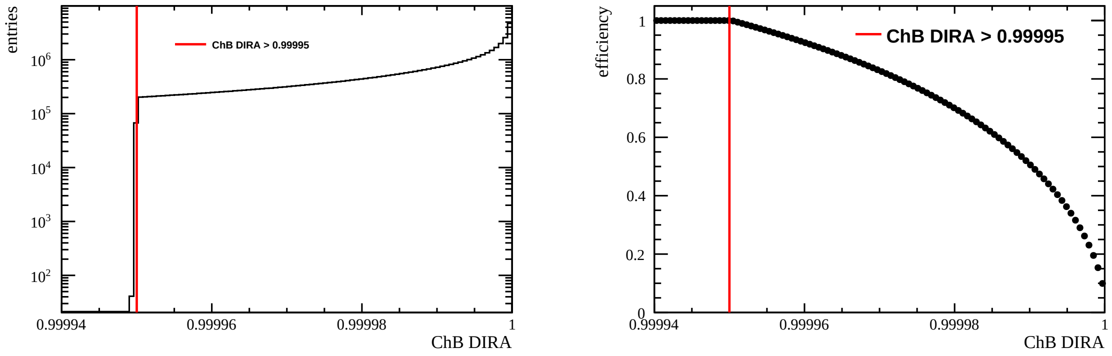

Figure 20: The DIRA distribution for Ξ_c^+ in 2018 Down data sample (left) and corresponding efficiency as a function of DIRA (right). The chosen cut DIRA > 0.99995 is shown as vertical red line.

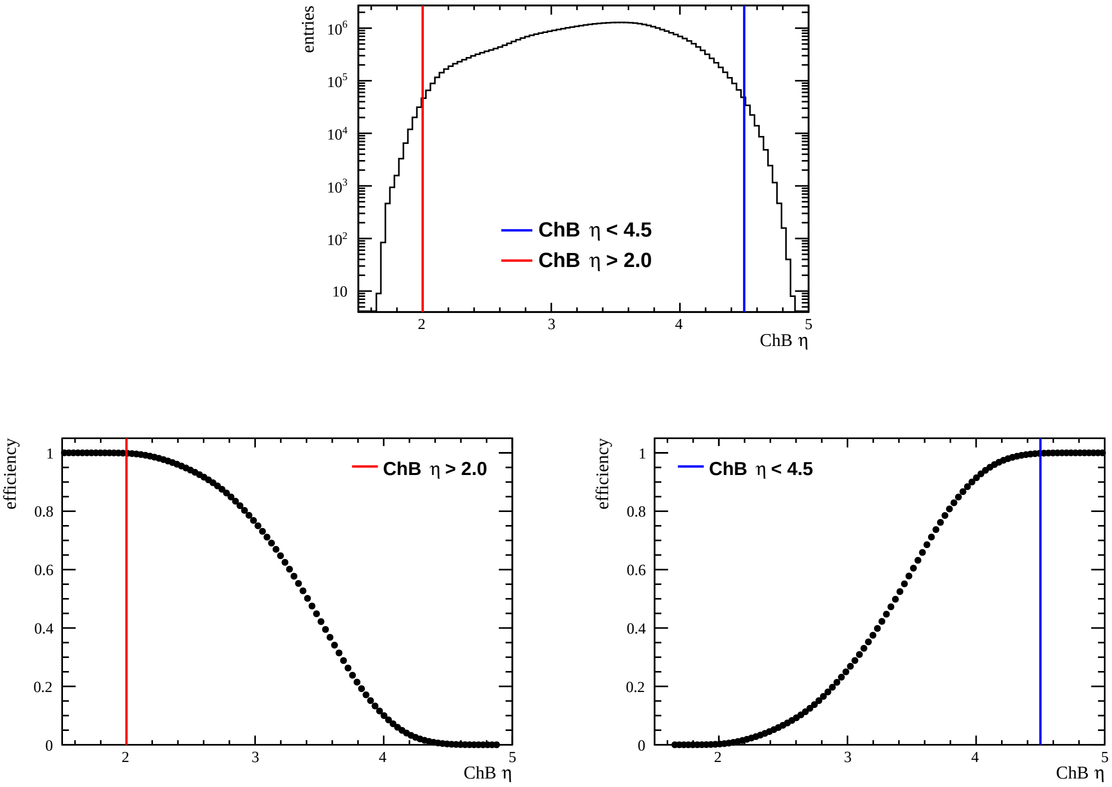

Figure 21: The pseudorapidity (η) distribution of Ξ_c^+ measured in 2018 Down data sample (top) and corresponding efficiency as a function of η for minimum limit (bottom left) and maximum limit (bottom right). The chosen cuts $\eta > 2$ and $\eta < 4.5$ are shown as vertical red and blue lines, respectively.

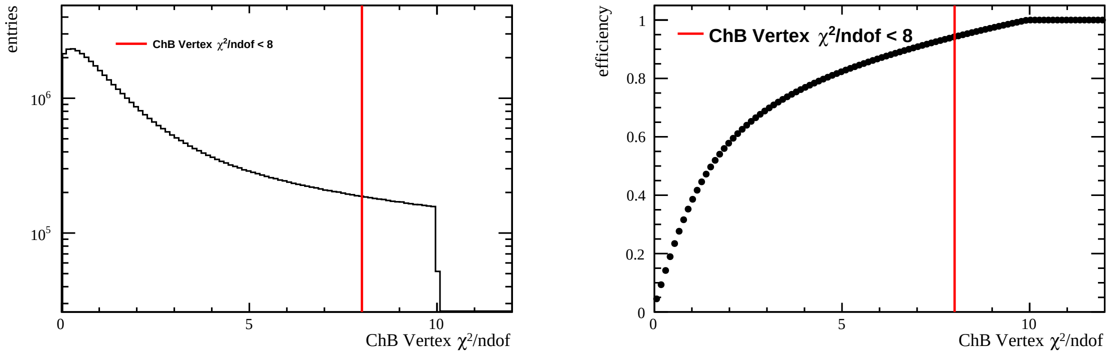

Figure 22: The vertex $\chi^2/ndof$ distribution of Ξ_c^+ measured in 2018 Down data sample (left) and corresponding efficiency as a function of Vertex $\chi^2/ndof$ (right). The chosen cut Vertex $\chi^2/ndof < 8$ is shown as vertical red line.

Table 7: The offline selection cuts.

variable	cut
p	
PID(p)	> 10
ProbNN(p)	> 0.5
IPχ^2	> 9
momentum	$15 < p < 100\,\mathrm{GeV}/c$
K	
ProbNN(K)	> 0.1
IPχ^2	> 9
momentum	$3 < p < 150\,\mathrm{GeV}/c$
π	
ProbNN(π)	> 0.1
IPχ^2	> 9
momentum	$3 < p < 150\,\mathrm{GeV}/c$
$\mathit{\Lambda}_c^+, \mathit{\Xi}_c^+$	
IPχ^2	< 12
χ^2 separation from related PV	> 20
Vertex $\chi^2/ndof$	< 8
DIRA	> 0.99995
pesudorapidity (η)	$(2; 4.5)$
transverse momentum	$4 < p_T < 16\,\mathrm{GeV}/c$
FDχ^2	< 2000

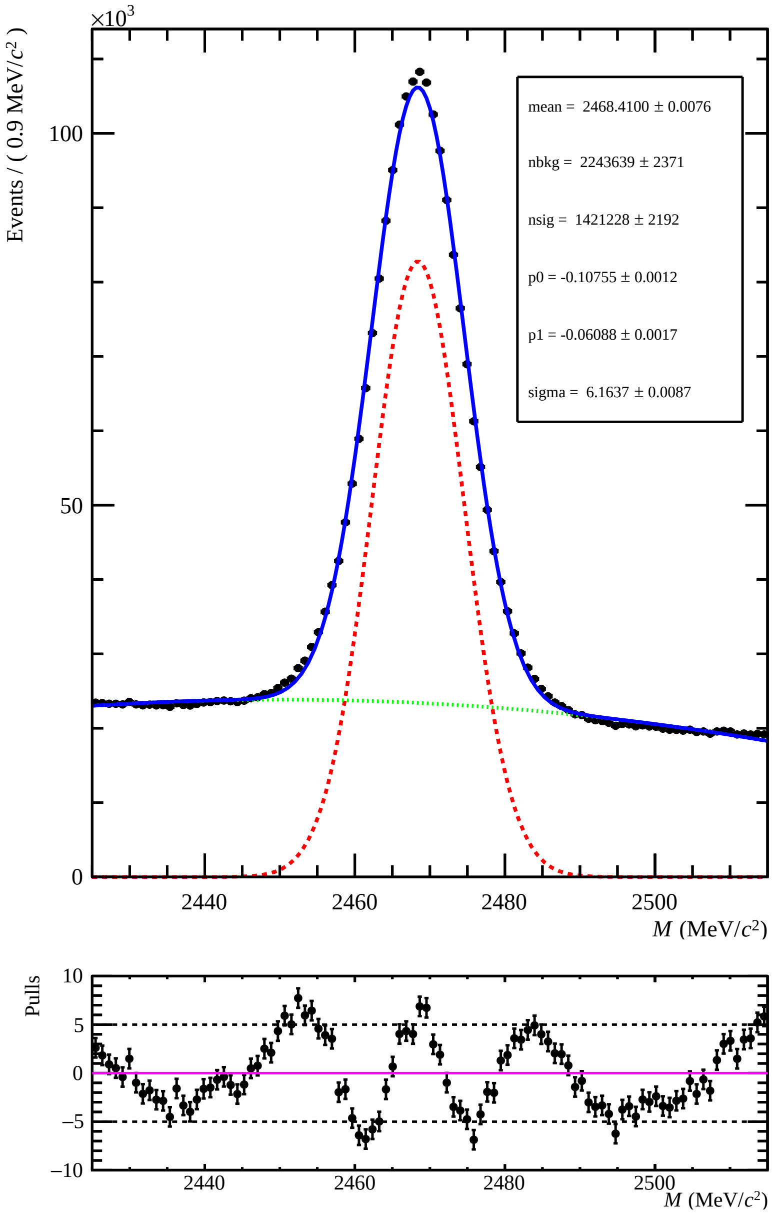

Figure 23: The reconstructed mass (M) of $pK^-\pi^+$ **after offline cuts** for Ξ_c^+ 2018 sample. The fitted mass model is overlaid (blue line). It is composed of single Gaussian function (red dashed line), with mean μ and width σ, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The $nsig$ and $nbkg$ are the numbers of signal and background events obtained from the fit, respectively. The shown pull distribution is a difference between data and fitted model.

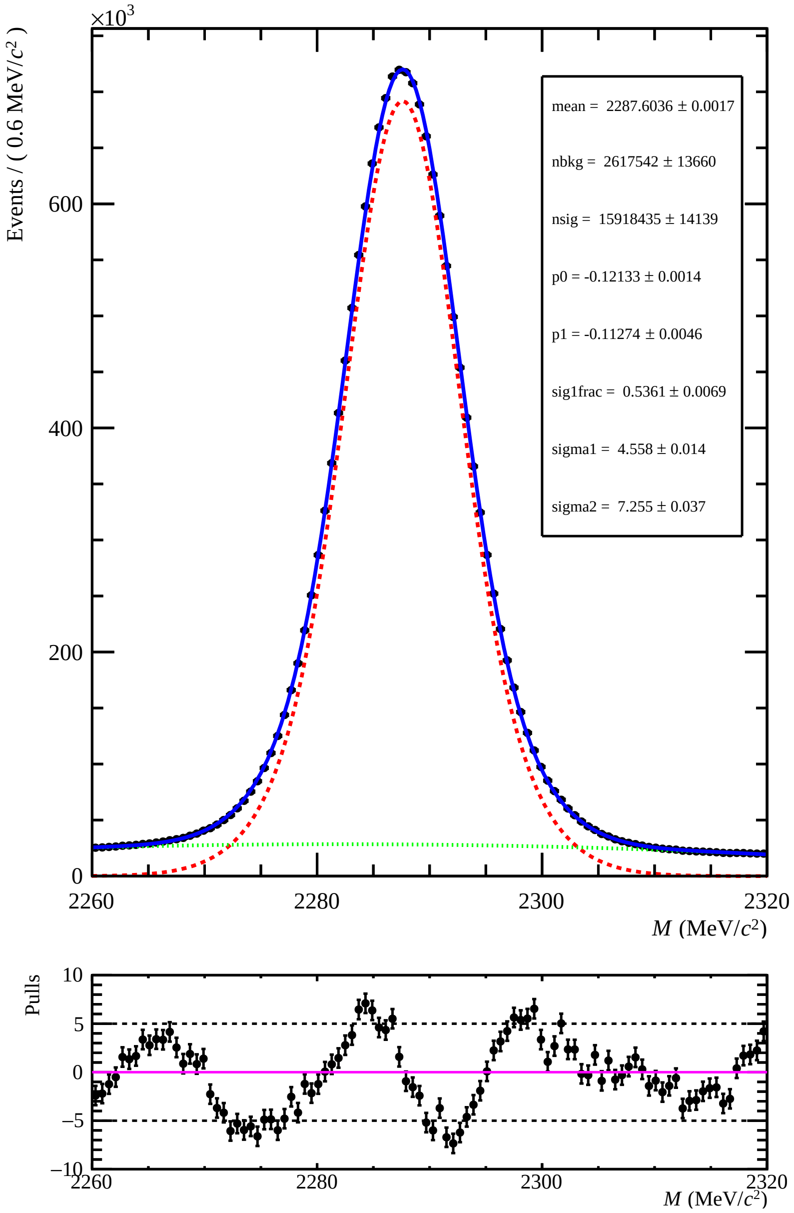

Figure 24: The reconstructed mass (M) of $pK^-\pi^+$ **after offline cuts** for Λ_c^+ 2018 sample. The fitted mass model is overlaid (blue line). It is composed of double Gaussian function (red dashed line), with common mean μ and different widths σ_1 and σ_2, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The $nsig$ and $nbkg$ are the numbers of signal and background events obtained from the fit, respectively. The shown pull distribution is a difference between data and fitted model.

4.3 Wrong mass hypothesis

Despite the fact that particle identification requirements are imposed on the proton, kaon and pion candidates, it is possible that a true meson decay is misidentified as Λ_c^+ or Ξ_c^+. For example, a true $D^+ \to \pi^+K^-\pi^+$ decay could be misidentified as $\Lambda_c^+; \Xi_c^+ \to pK^-\pi^+$ if the proton mass hypothesis is assigned to the pion and passes the particle identification cuts, or a true $D_s^+ \to K^+K^-\pi^+$ decay might be misinterpreted as $\Lambda_c^+; \Xi_c^+ \to pK^-\pi^+$ if the proton mass hypothesis is appointed to be kaon. Such misidentified decays can contribute to the Λ_c^+ and Ξ_c^+ mass spectra and the measured asymmetries may have a contribution from production asymmetry of D^+ or D_s^+ .

An approach to test the presence of misidentified decays is to inspect three-body mass distribution when each particle is assigned different mass hypotheses. For instance, one can assign the pion mass to the proton candidate in the $pK^-\pi^+$ mode and investigate the invariant mass of the $\pi^+K^-\pi^+$ combination. Any excess in the area close to the D^+ mass would indicate a contamination of D^+ candidates in the data etc. These exercises are performed for both control Λ_c^+ and signal Ξ_c^+ decays after applying all offline selection cuts to those data samples. Data with various different hypothesis sets for Λ_c^+ decays are shown in Figs 25 and 26, where there is not proton mass hypothesis in any set (expected background from D^+ and D_s^+ decays) and where each set does include the proton mass hypothesis (expected background from other Λ_c^+ decays than $\Lambda_c^+ \to pK^-\pi^+$). There are no mass peaks corresponding to the decays of D^+ and D_s^+ and also Λ_c^+ from decays other than $\Lambda_c^+ \to pK^-\pi^+$. The contamination from wrongly identified particles is insignificant.

As far as four-body D_s^+ decays are concerned they can be only partially reconstructed and in the result the peak should be observed below the expected mass, as for example $D^+ \to hhh\pi^0$, π^0 is not reconstructed and for $D_s^+ \to hhhh$ one h is missing. In addition to the partially reconstructed event, a misidentification may also occur. For such decays the peak could be observed in Fig. 25. The nominal D^+, D_s^+ and Λ_c^+ masses are $1869.6\,\mathrm{MeV}/c^2$, $1968.3\,\mathrm{MeV}/c^2$ and $2286.5\,\mathrm{MeV}/c^2$, respectively, and no contamination from those particle decays can be found.

To sum up, in every data set, the contamination from wrongly identified particles is negligible as well as from partially reconstructed four-body $D_{(s)}^+$. The conclusion is that there is only production asymmetry from Λ_c^+ or Ξ_c^+ and the contributions from the other decays are not needed to be considered.

4.4 Multiple candidates

Due to the assembly of daughter particles into one particle, it is possible that two or more physical decay candidates are reconstructed in a single event. Such multiple candidates could be caused by nonphysical means such as cloned tracks. The fully selected data are checked for multiple candidates. This is done by counting the number of candidates which share the same run number and event number, the value of which is unique. For events with duplicate Λ_c^+ or Ξ_c^+ candidates, the candidate with the best χ^2 of vertex is chosen, removing all others. After final offline selection cuts, it is found about 5% of events with multiple candidates in all analysed Run 2 data.

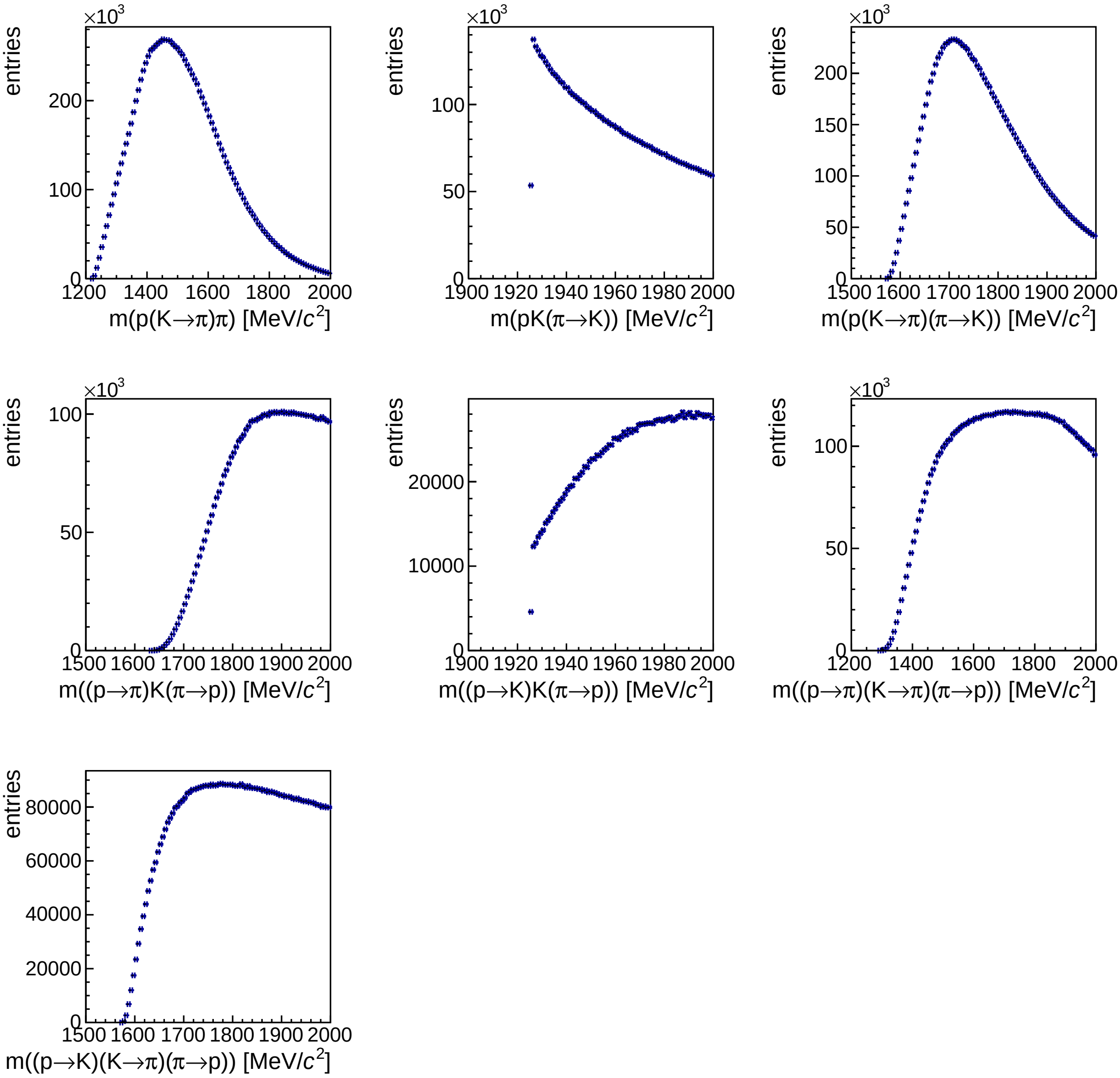

Figure 25: Wrong-mass hypothesis distributions obtained by changing mass in fully selected $pK^-\pi^+$ decay mode, where no child is appointed the proton mass. For example, the $(p \to \pi^+)K^-\pi^+$ distribution is where the nominal proton candidate has been assigned the pion mass hypothesis. In presented mass window, the D^+ and D_s^+ decays can be manifested as the PDG mass of D^+ equals 1869.6 MeV and the PDG mass of D_s^+ equals $1968.3\,\text{MeV}/c^2$ [25]. The presented distributions are obtained in whole Run 2 data.

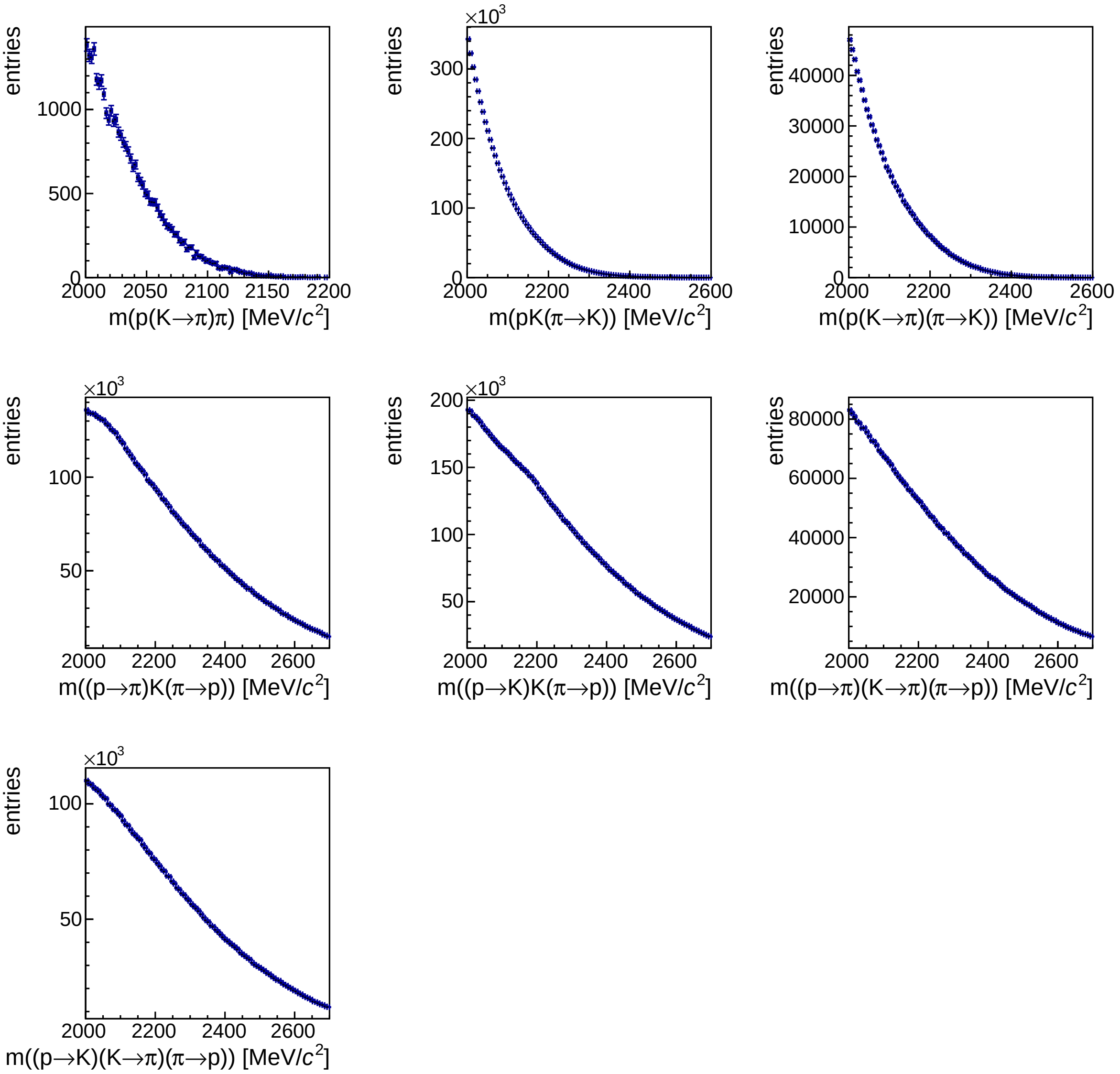

Figure 26: Wrong-mass hypothesis distributions obtained by changing mass in fully selected $pK\pi$ decay mode, where no child is appointed the proton mass. For example, the $(p \to \pi)K\pi$ distribution is where the nominal proton candidate has been assigned the pion mass hypothesis. In presented mass window the other Ξ_c^+ and Λ_c^+ decay modes than $\Lambda_c^+; \Xi_c^+ \to pK^-\pi^+$ ones can be manifested. The PDG mass of Λ_c^+ equals 2286.5 MeV/c^2 and the PDG mass of Λ_c^+ equals 2467.8 MeV/c^2. The presented distributions are obtained in whole Run 2 data.

4.5 Lifetime distribution test

In order to check whether the selected cuts do not bias the sample, the baryon lifetime is measured. To do it the sPlot technique [66] is used. Since data typically consists of event from different species, such as background and signal merged in one sample, the sPlot technique allows to separate them statistically. Each event is described by the set of variables divided into two components. The first one contains variables for which distributions of all the sources of events are known. These variables are referred as to discriminating variables. The second class are called control variables for which distributions of some events are unknown. The sPlot technique allows to estimate the distributions for control variables without the use of any prior knowledge of those variables. Only information accessible for the discriminating variable is applied to understand the behaviour of the individual sources of events with respect to the control variable. A critical assumption to sPlot to work properly is that control variables are not correlated with discriminating values. To unfold data distributions for control variables for each species weights called sWeights are used to fill the histogram bins. In this analysis discriminating variable is the Ξ_c^+ baryon mass and the control value is the lifetime of this baryon.

The process of the sPlot technique consists of several steps. Firstly, the sample on which sPlot technique is applied is created and it is about 5% of the Down 2018 data sample. Then the maximum likelihood fit is performed to baryon mass distribution of Ξ_c^+ and sWeights are calculated. In the next step, those sWeights are imposed on lifetime distribution of Ξ_c^+ to separate statistically the signal part from the background part. Finally, the exponential function is fitted to the statistically separated signal part of the lifetime distribution of Ξ_c^+. Technically the fit is performed using the RooFit tool [65]. Following function is fitted:

$$f(x, \lambda) = N \cdot exp(x \cdot \lambda), \tag{22}$$

where N is a normalization constant that depends on the range and values of the arguments and λ is the decay constant which can be translated to the mean lifetime τ of the Ξ_c^+ as:

$$\tau = \frac{1}{\lambda}. \tag{23}$$

The uncertainty of τ $(u(\tau))$ is calculated using following formula:

$$u(\tau) = \frac{u(\lambda)}{\lambda^2}, \tag{24}$$

where $u(\lambda)$ is the measured uncertainty on λ.

The re-weighted lifetime distributions is presented in Fig. 27. The obtained Ξ_c^+ lifetime $\tau = 450 \pm 8$ fs is in agreement with average τ of Ξ_c^+ 453 ± 5 fs given in PDG [25]. The conclusion is that any contamination from other decays such as secondary Λ_b^0 or Ξ_b seems to be negligible.

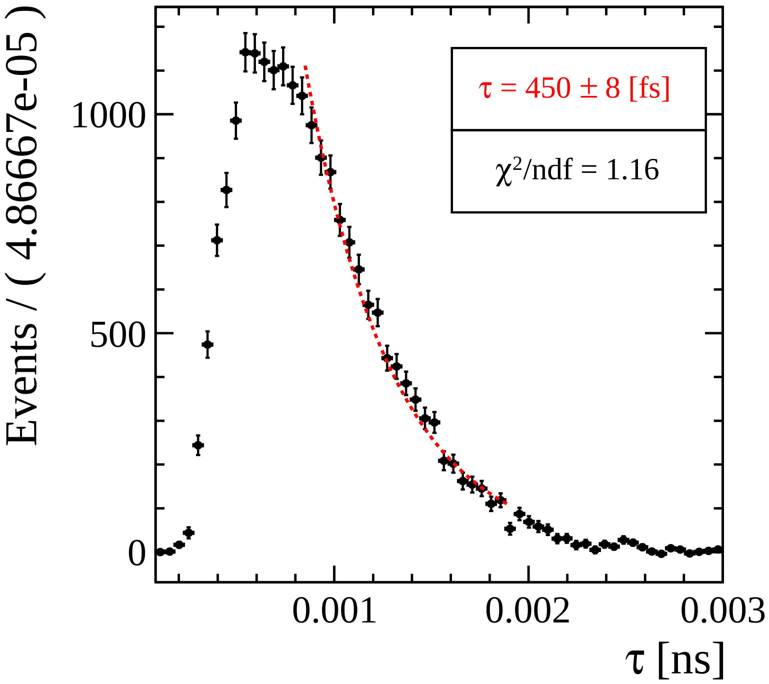

Figure 27: Lifetime (τ) distribution of Ξ_c^+ re-weighted using the sPlot technique. Dashed red line represents section of τ distribution in which exponential function $f(x, \lambda) = N \cdot exp(x \cdot \lambda)$ is performed.

4.6 Detector effects

The pollution asymmetries can affect the searches for *CP* violation by changing this asymmetry. Hence, it is necessary to identify all pollution asymmetries and subtract their from final *CP* asymmetry or eliminate all pollution asymmetries if their disentangle from *CP* asymmetry is not possible. For this reason the physics effects as well as effects related with the geometry of the detector are taken into account. For further consideration the events in the mass windows defined as $\pm$ 20 MeV/c^2 around the expected mass of Λ_c^+ and Ξ_c^+ are chosen. The mass distribution with that range contains the most signal events relative to the number of background events. The purity of Λ_c^+ and Ξ_c^+ distributions are 93% and 76% respectively.

The pollution asymmetries can be disentangle into production and detector asymmetries. The reason for the production asymmetry is different branching ratio for production particles and antiparticles in *pp* interactions what is characteristic for LHC (the *pp* interactions are charge asymmetric). In addition, the production asymmetry depends on particle momenta. The detector asymmetries are related to all plausible detector effects, as coming from geometry, acceptance, reconstruction etc. For example, any inconsistencies in geometry of the detector can cause large detector asymmetries which do not cancel after mixing data with two types of magnet polarization.

4.6.1 Fiducial cuts

During the construction of the detector, a lot of effort is made to make it as symmetrical as possible, but any irregularities in its geometry may lead to detector asymmetry. These asymmetries can affect the *CP* asymmetry. Since *CP* asymmetry in charm sector is expected to be smaller than a per mille, even very small differences in the detector geometry can cause larger effects than this expected small value. In particular, significant detector asymmetries are expected at the edges of the detector: external and internal close to the beam axis. To investigate that effects two-dimensional histograms are created: distributions of the x versus the z component of each particle momentum and the raw asymmetry A_{raw} is calculated using the following definition:

$$A_{raw} = \frac{N_+ - N_-}{N_+ + N_-}, \tag{25}$$

where N_+ and N_- are the numbers of accounted particles and antiparticles, respectively.

Calculations of A_{raw} are performed in Up and Down samples individually. In case of no geometrical differences of the detector, alteration of the polarity should change the sign of the A_{raw} only and after mixing Up and Down samples their non zero values should cancel. However, any slight differences in the detector geometry could cause that residual detector asymmetries not cancel totally after Up and Down samples are mixed. Such residual detector asymmetries have to be reduced by introducing so called fiducial cuts. Two categories of fiducial cuts are applied: the external and the internal. The former are related to outer regions of the detector and latter to the area close to the beam pipe.

For the proton, kaon and pion measured A_{raw} in Down and Up 2018 samples are presented in Figs 28, 29 and 30, respectively, for Ξ_c^+ decays and in Figs 31, 32 and 33 for Λ_c^+ decays. The substantial detector asymmetries are clearly observed. For all particles - protons, kaons, and pions - large asymmetry values (close to one) are measured in the outer parts of the detector and medium values (about twice as small) in the vicinity of the accelerator tube. These asymmetries change sign when the magnetic field changes (the colours reverse). Overall the fiducial cuts are introduced for all particles reconstructed in the final state, i.e. for protons, kaons and pions. The results show that outer areas should be treated with $|p_x| > 10\,\text{GeV}$ and $|p_x/p_z| > 0.2$ cuts to get rid of those significant detection asymmetries and rejecting events which satisfying the conditions $|p_y/p_z| < 0.02$ and $|p_x/p_z| < 0.025$ is enough to remove asymmetries in the surroundings of the beam axis. Lines representing cuts (considered and chosen) are drawn in Figs. 28-33. The chosen cuts discarded 2% of all events.

The total A_{raw} is measured in 2018 Up, Down and mixed samples after implementing offline selection cuts and fiducial cuts. The results for Ξ_c^+ and Λ_c^+ decays for protons, kaons and pions are shown in Tables 8-10 and 11-13, respectively. The individual raw asymmetries in the Up and Down samples are different from zero, which is particularly noticeable in the Λ_c^+ samples . In contrast, when Up and Down samples are mixed, these asymmetries cancel out and the measured A_{raw} are consistent with zero.

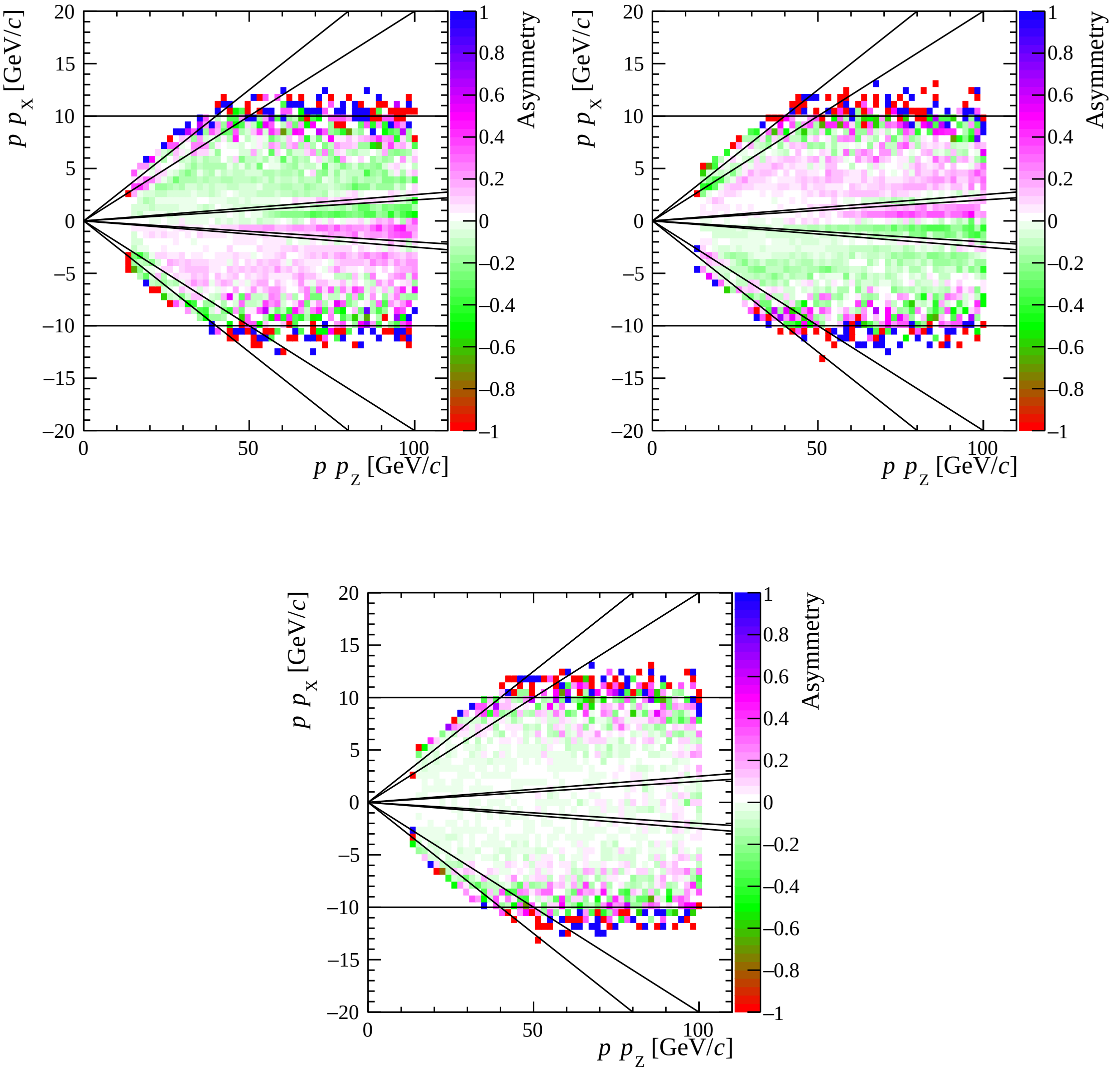

Figure 28: Measured raw asymmetry (A_{raw}) of **protons** (p) in $p_x - p_z$ plane of the mass windows of ± 20 MeV of $\Xi_c^+ \to pK^-\pi^+$ decays in 2018 samples: **Down** (top left), **Up** (top right) and **Down+Up** (bottom). The black lines represent different considered external and internal fiducial cuts. For the external fiducial cuts, the outer lines correspond to, in order of outermost: $|p_x/p_z| = 0.25$ and $|p_x/p_z| = 0.2$. For the internal fiducial cuts the outer line corresponds to $|p_x/p_z| = 0.025$ and the internal line corresponds to $|p_x/p_z| = 0.02$. The horizontal lines correspond to $|p_x| = 10\,\text{GeV}/c$.

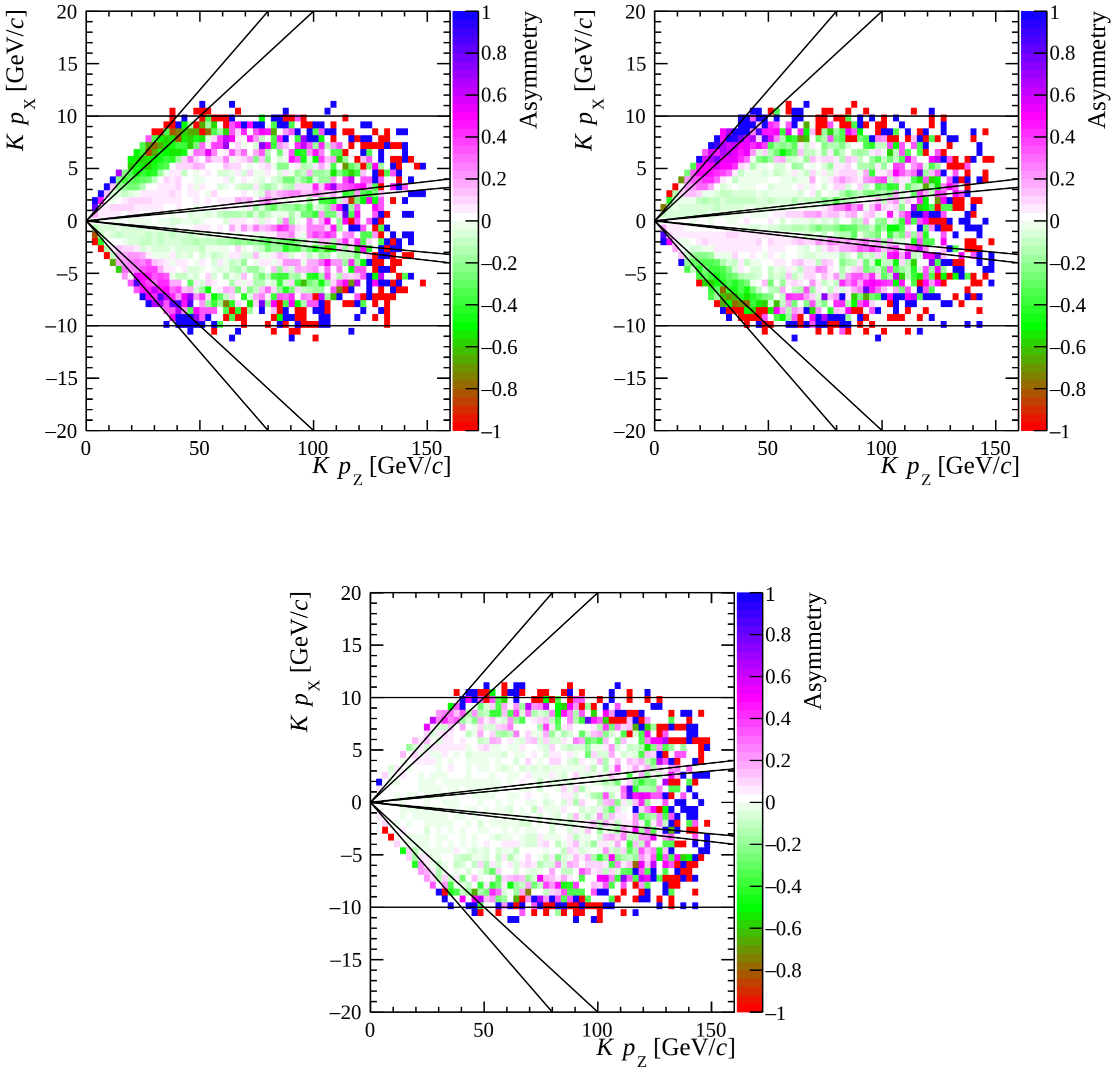

Figure 29: Measured raw asymmetry A_{raw} of **kaons** (K) in $p_x - p_z$ plane of the mass windows of ±20 MeV of $\Xi_c^+ \to pK^-\pi^+$ decays in 2018 samples: **Down** (top left), **Up** (top right) and **Down+Up** (bottom). The black lines represent different considered external and internal fiducial cuts. For the external fiducial cuts, the outer lines correspond to, in order of outermost: $|p_x/pz| = 0.25$ and $|p_x/pz| = 0.2$. For the internal fiducial cuts the outer line corresponds to $|p_x/p_z| = 0.025$ and the internal line corresponds to $|p_x/p_z| = 0.02$. The horizontal line represents the $|p_x| < 10\,\text{GeV}/c$ cut.

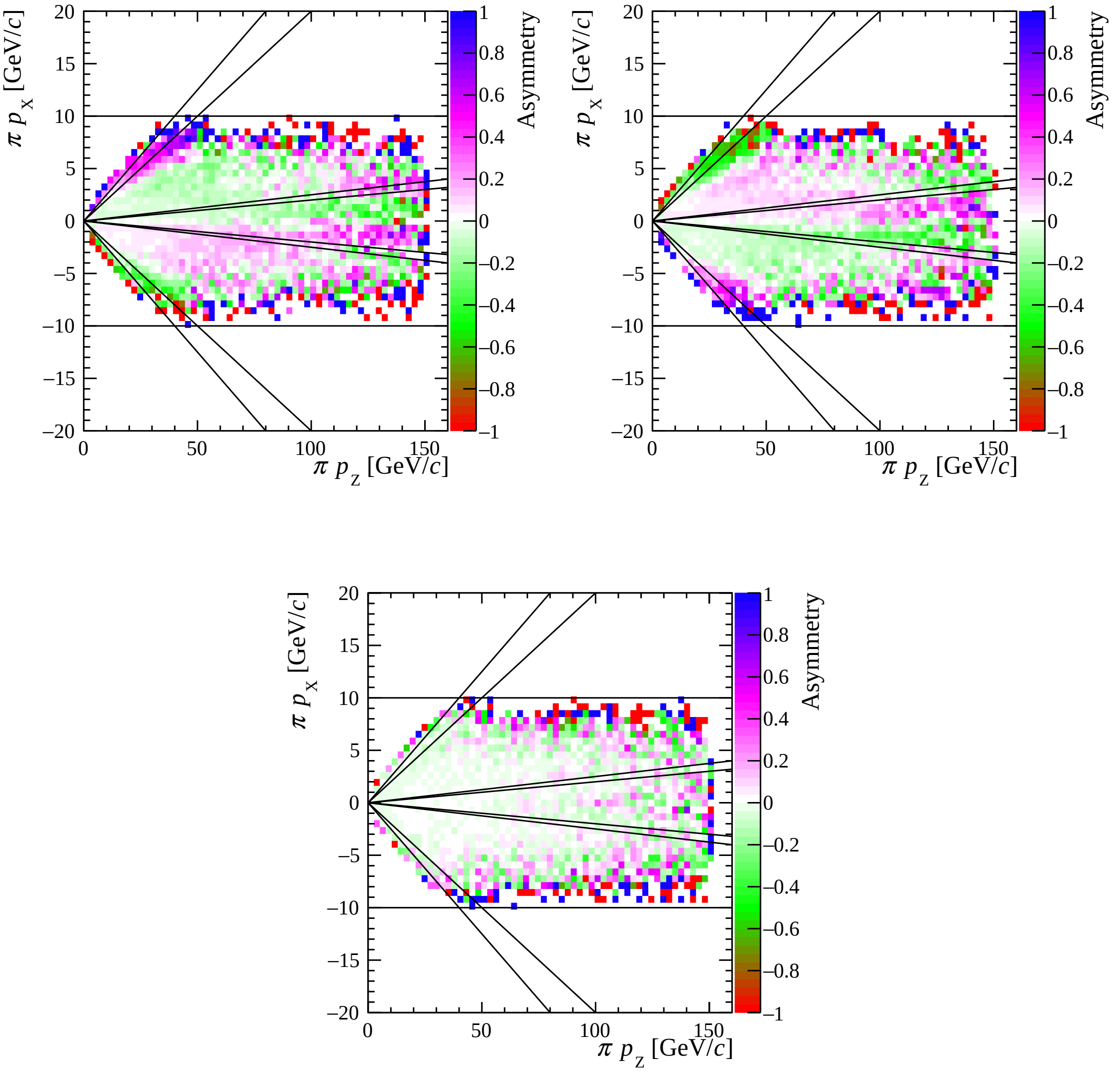

Figure 30: Measured raw asymmetry A_{raw} of **pions** (π) in $p_x - p_z$ plane of the mass windows of ±20 MeV of $\Xi_c^+ \to pK^-\pi^+$ decays in 2018 samples: **Down** (top left), **Up** (top right) and **Down+Up** (bottom). The black lines represent different considered external and internal fiducial cuts. For the external fiducial cuts, the outer lines correspond to, in order of outermost: $|p_x/pz| = 0.25$ and $|p_x/pz| = 0.2$. For the internal fiducial cuts the outer line corresponds to $|p_x/p_z| = 0.025$ and the internal line corresponds to $|p_x/p_z| = 0.02$. The horizontal line represents the $|p_x| < 10\,\mathrm{GeV}/c$ cut.

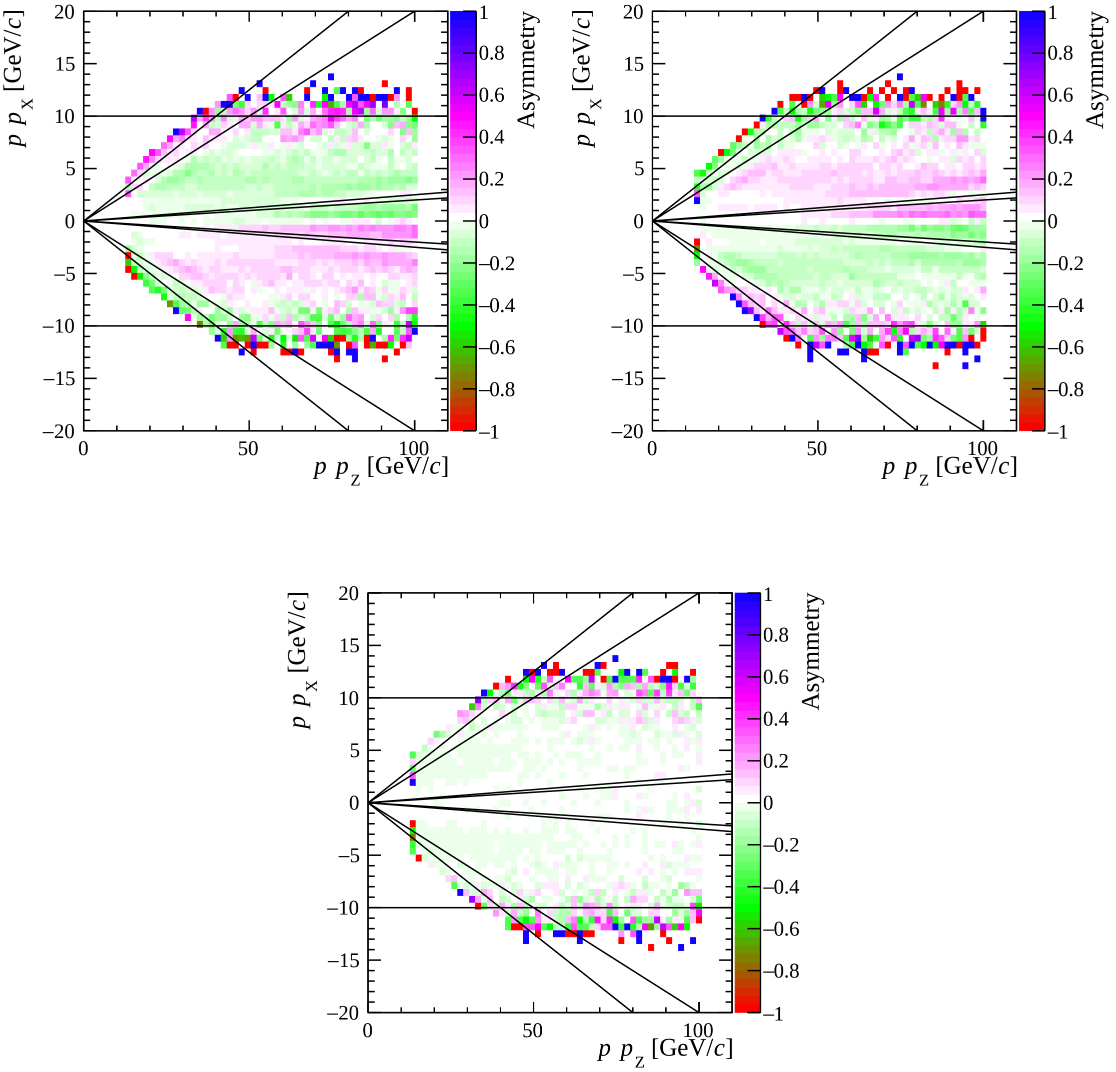

Figure 31: Measured raw asymmetry A_{raw} of **protons** (p) in $p_x - p_z$ plan of the mass windows of ±20 MeV of $\Lambda_c^+ \to pK^-\pi^+$ decays in 2018 samples: **Down** (top left), **Up** (top right) and **Down+Up** (bottom). The black lines represent different considered external and internal fiducial cuts. For the external fiducial cuts, the outer lines correspond to, in order of outermost: $|p_x/pz| = 0.25$ and $|p_x/pz| = 0.2$. For the internal fiducial cuts the outer line corresponds to $|p_x/p_z| = 0.025$ and the internal line corresponds to the cut $|p_x/p_z| = 0.02$. The horizontal line represents the $|p_x| < 10\,\text{GeV}/c$ cut.

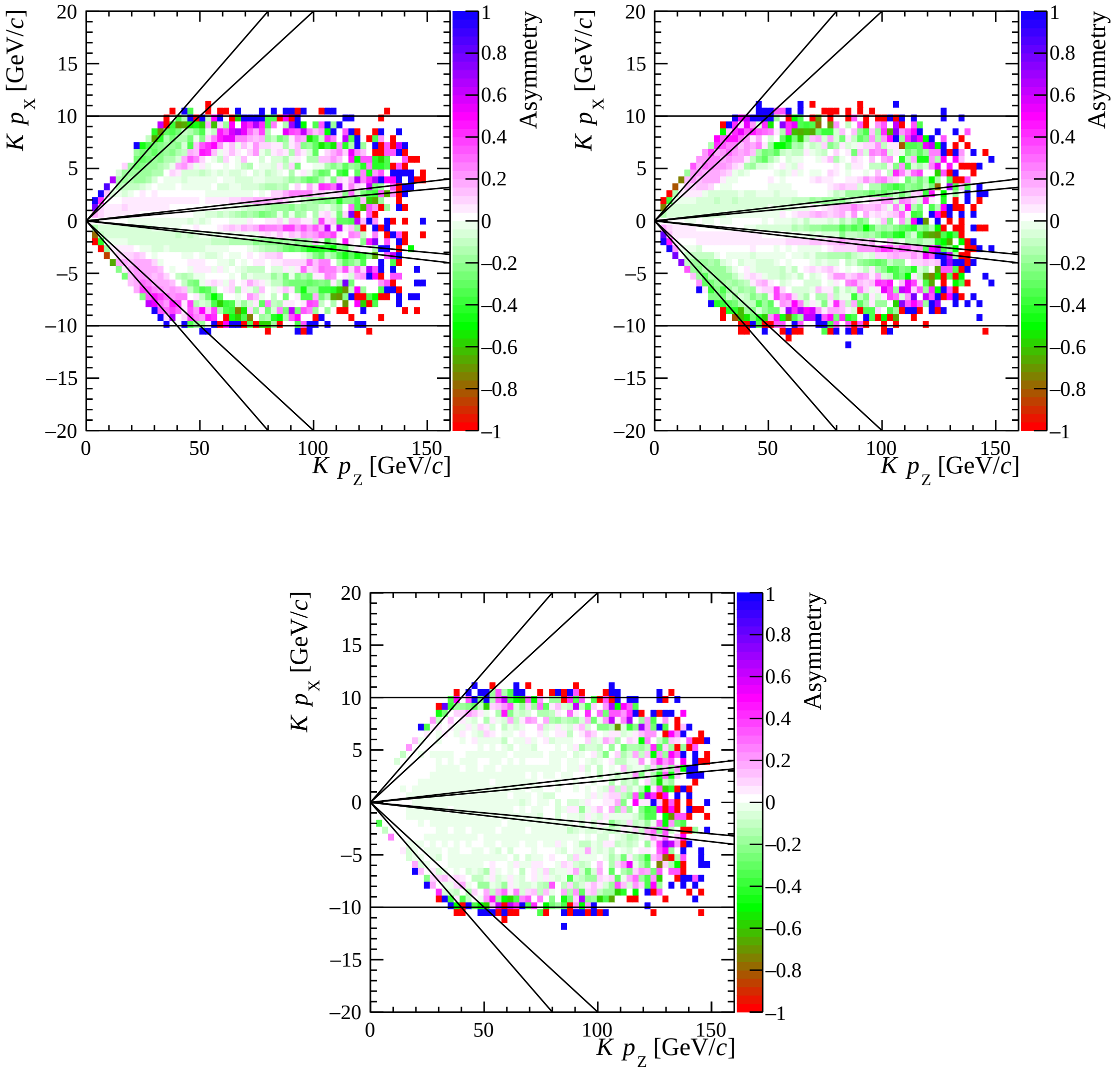

Figure 32: Measured raw asymmetry A_{raw} of **kaons** (K) in $p_x - p_z$ plan of the mass windows of ± 20 MeV of $\Lambda_c^+ \to pK^-\pi^+$ 2decays in 2018 samples: **Down** (top left), **Up** (top right) and **Down+Up** (bottom). The black lines represent different considered external and internal fiducial cuts. For the external fiducial cuts, the outer lines correspond to, in order of outermost: $|p_x/pz| = 0.25$ and $|p_x/pz| = 0.2$. For the internal fiducial cuts the outer line corresponds to $|p_x/p_z| = 0.025$ and the internal line corresponds to $|p_x/p_z| = 0.02$. The horizontal line represents the $|p_x| < 10\,\mathrm{GeV}/c$ cut.

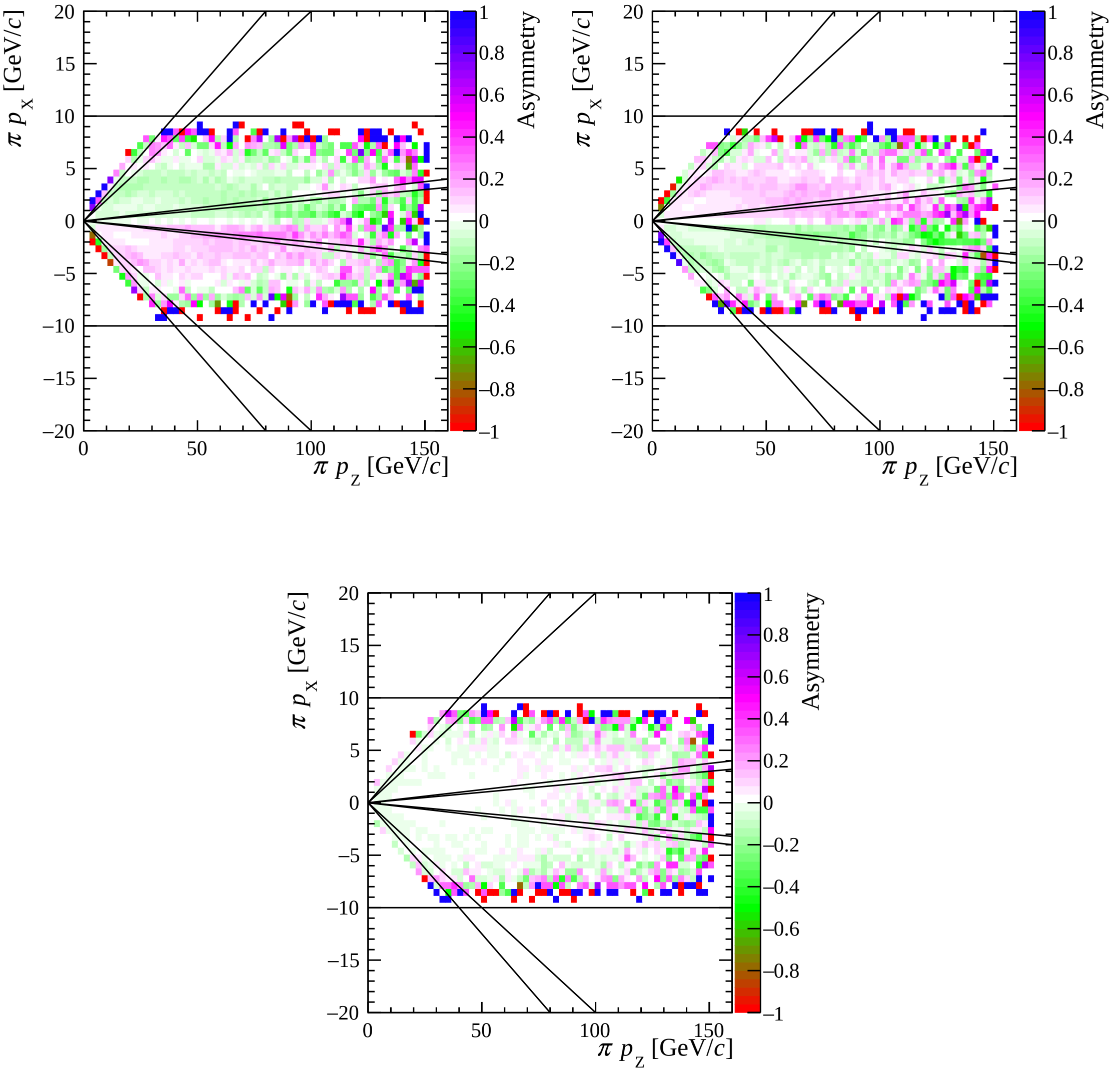

Figure 33: Measured raw asymmetry A_{raw} of **pions** (π) in $p_x - p_z$ plan of the mass windows of ±20 MeV of $\Lambda_c^+ \to pK^-\pi^+$ decays in 2018 samples: **Down** (top left), **Up** (top right) and **Down+Up** (bottom). The black lines represent different considered external and internal fiducial cuts. For the external fiducial cuts, the outer lines correspond to, in order of outermost: $|p_x/pz| = 0.25$ and $|p_x/pz| = 0.2$. For the internal fiducial cuts the outer line corresponds to $|p_x/p_z| = 0.025$ and the internal line corresponds to $|p_x/p_z| = 0.02$. The horizontal line represents the $|p_x| < 10\,\mathrm{GeV}/c$ cut.

Table 8: Measured raw asymmetries (A_{raw}) and the number of events in the mass windows of ± 20 MeV of $\Xi_c^+ \to pK^-\pi^+$ decays for **protons** after implementing offline selection cuts and fiducial cuts in 2018 data sets.

Sample	Number of events	A_{raw}
Up	1 182 444	-0.00286 ± 0.00092
Down	1 218 724	0.00182 ± 0.00091
Down + Up	2 401 168	-0.0010 ± 0.0013

Table 9: Measured raw asymmetries (A_{raw}) and the number of events in the mass windows of ± 20 MeV of $\Xi_c^+ \to pK^-\pi^+$ decays for **kaons** after implementing offline selection cuts and fiducial cuts in 2018 data sets.

Sample	Number of events	A_{raw}
Up	1 178 412	-0.00222 ± 0.00092
Down	1 214 680	0.00139 ± 0.00091
Down + Up	2 393 092	-0.0008 ± 0.0013

Table 10: Measured raw asymmetries (A_{raw}) and the number of events in the mass windows of ± 20 MeV of $\Xi_c^+ \to pK^-\pi^+$ decays for **pions** after implementing offline selection cuts and fiducial cuts in 2018 data sets.

Sample	Number of events	A_{raw}
Up	1 173 381	-0.00306 ± 0.00092
Down	1 209 572	0.00209 ± 0.00091
Down + Up	2 382 953	-0.0010 ± 0.0013

Table 11: Measured raw asymmetries (A_{raw}) and the number of events in the mass windows of $\pm 20\,\mathrm{MeV}/c^2$ of $\Lambda_c^+ \to pK^-\pi^+$ decays for **protons** after implementing all offline selection cuts and fiducial cuts in 2018 data sets.

2018	Number of events	A_{raw}
Up	8 537 057	-0.00384 ± 0.00034
Down	8 781 389	0.00469 ± 0.00034
Down + Up	17 318 446	0.00086 ± 0.00048

Table 12: Measured raw asymmetries (A_{raw}) and the number of events in the mass windows of $\pm 20\,\mathrm{MeV}/c^2$ of $\Lambda_c^+ \to pK^-\pi^+$ decays for **kaons** after implementing all offline selection cuts and fiducial cuts in 2018 data sets.

2018	Number of events	A_{raw}
Up	8 509 581	-0.00366 ± 0.00034
Down	8 752 972	0.00439 ± 0.00034
Down + Up	17 262 553	0.00073 ± 0.00048

Table 13: Measured raw asymmetries (A_{raw}) and the number of events in the mass windows of $\pm 20\,\mathrm{MeV}/c^2$ of $\Lambda_c^+ \to pK^-\pi^+$ decays for **pions** after implementing all offline selection cuts and fiducial cuts in 2018 data sets.

2018	Number of events	A_{raw}
Up	8 480 884	-0.00410 ± 0.00034
Down	8 722 938	0.00481 ± 0.00034
Down + Up	17 203 822	0.00071 ± 0.00048

4.6.2 Reconstruction effects

Due to different cross section for interacting with the material of the detector, the particles and antiparticles can be reconstructed disparately which leads to reconstruction asymmetries. These asymmetries could be bigger than potential *CP* asymmetry and therefore bias the possible signal of *CPV*. Since cross section depends on momentum of interacting particle, to probe the magnitude of these effects, the reconstruction asymmetries between Down and Up samples are calculated as a function of the particle momentum. This asymmetry is defined as follows:

$$A = \frac{N_U - N_D}{N_U + N_D}, \tag{26}$$

where N_U is the number of events in Up sample and N_D is the number of events in Down sample. The asymmetries are calculated separately for all particles and antiparticles reconstructed in the final state, i.e. for protons, kaons and pions. Figs 34-36 show the measured A as a function of momenta for protons and antiprotons, kaons and antikaons, pions and antipions, respectively, in $\Lambda_c^+ \to pK^-\pi^+$ 2018 samples, while Figs 37 -39 show the results obtained in $\Xi_c^+ \to pK^-\pi^+$ 2018 samples. The differences in reconstruction particles and antiparticles are clearly seen. The largest differences are observed for protons in the smallest values of measured momenta. It is apparent that differences in reconstruction of protons and antiprotons emerge for the smallest and largest momenta. To remedy this issue, additional cuts are introduced for proton and antiproton momenta to choose region where their reconstructions are the same. The protons and antiprotons with $p > 25\,\mathrm{GeV}/c$ and $p < 80\,\mathrm{GeV}/c$ are chosen for further analysis. For kaons and antikaons, pions and antipions the chosen region is with $p > 15\,\mathrm{GeV}/c$. These additional requirements induced on momentum combined with fiducial cuts, discussed in previous section 4.6.1 successfully remove disparities between particles and antiparticles. The final samples contain events, which passed through all introduced cuts: offline, fiducial and $80\,\mathrm{GeV}/c > p(p) > 25\,\mathrm{GeV}/c$, $p(K) > 15\,\mathrm{GeV}/c$, $p(\pi) > 15\,\mathrm{GeV}/c$. The fiducial cuts and cuts on momenta are throwing away about 75% of background and signal events. This is the cost for getting rid of detector effects, as they cannot be separated in any other way and subtracted from the final *CP* asymmetry.

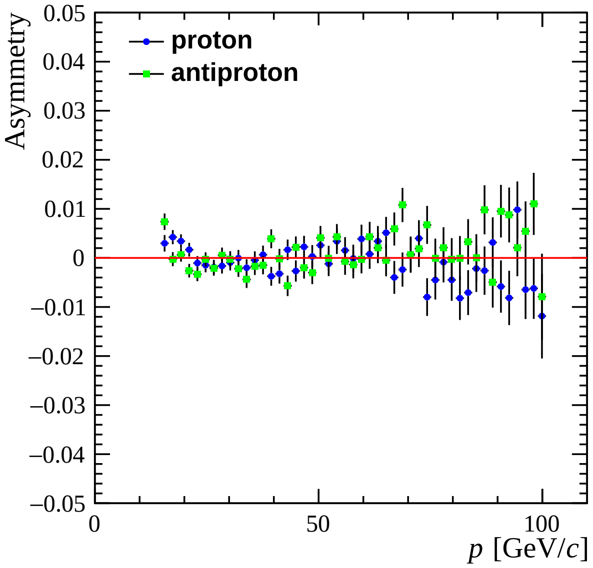

Figure 34: Reconstruction asymmetry as a function of momentum for proton (blue markers) and antiproton (green markers) measured in Λ_c^+ 2018 samples.

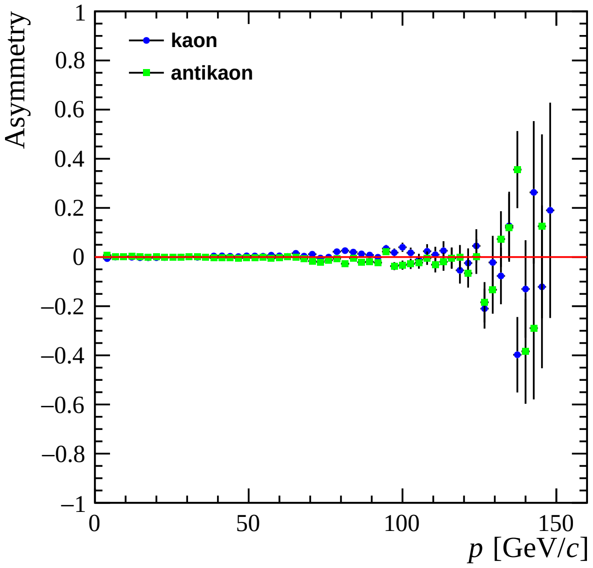

Figure 35: Reconstruction asymmetry as a function of momentum for kaon (blue markers) and antikaon (green markers) measured in Λ_c^+ 2018 samples.

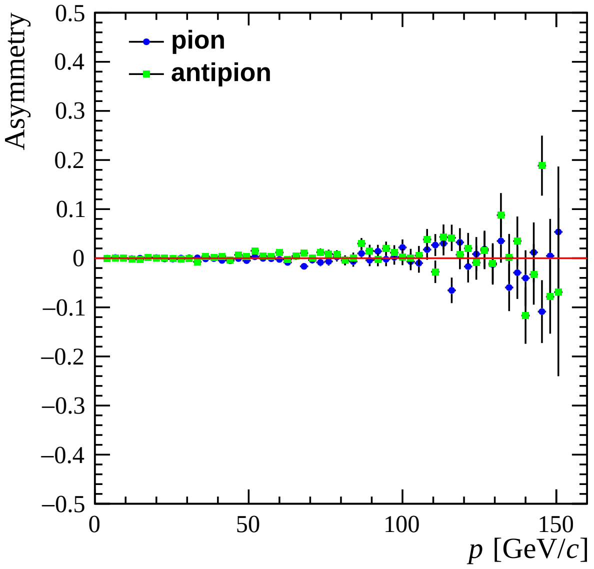

Figure 36: Reconstruction asymmetry as a function of momentum for pion (blue markers) and antipion (green markers) measured in Λ_c^+ 2018 samples.

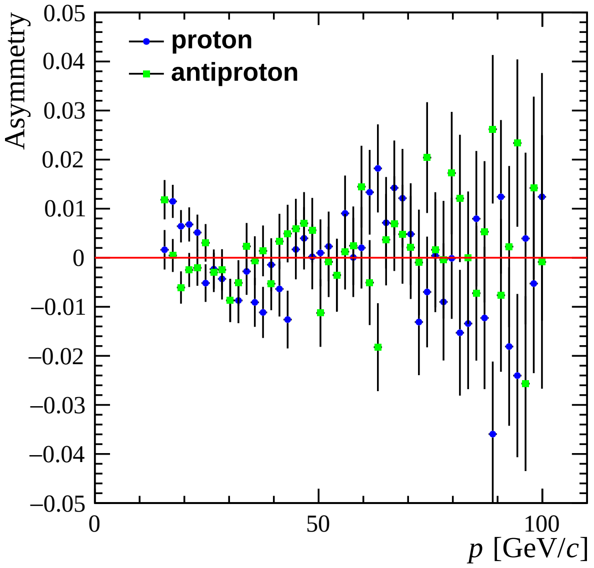

Figure 37: Reconstruction asymmetry as a function of momentum for proton (blue markers) and antiproton (green markers) measured in Ξ_c^+ 2018 samples.

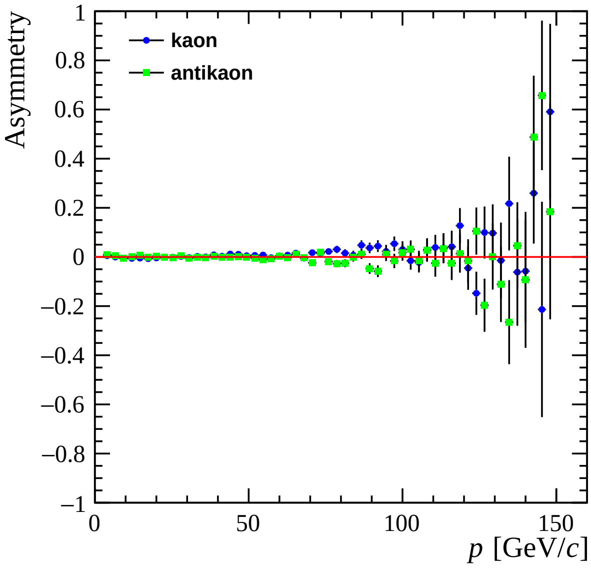

Figure 38: Reconstruction asymmetry as a function of momentum for kaon (blue markers) and antikaon (green markers) measured in Ξ_c^+ 2018 samples.

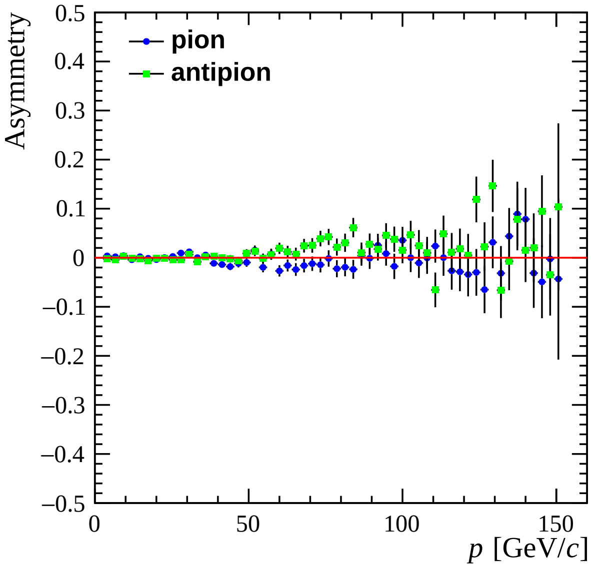

Figure 39: Reconstruction asymmetry as a function of momentum for pion (blue markers) and antipion (green markers) measured in Ξ_c^+ 2018 samples.

4.7 Final statistics

After implementig all cuts, the reconstructed masses in whole Run 2 are presented in Figs 40 and 41 for Λ_c^+ and Ξ_c^+, respectively. To data the model described in section 4.2 is fitted. The measured numbers of Λ_c^+ and Ξ_c^+ candidates obtained from the fit are shown in Table 14. Finaly, there are $10\,800 \pm 16$ thousand of Λ_c^+ candidates and 945 ± 2 thousand of Ξ_c^+ candidates in whole Run 2. The measured masses are $2287.5609 \pm 0.0021\,\mathrm{MeV}/c^2$ and $2468.281 \pm 0.010\,\mathrm{MeV}/c^2$ for Λ_c^+ and Ξ_c^+, respectively. These values are in agreement with the PDG values, which are $2286.46 \pm 0.14\,\mathrm{MeV}/c^2$ for Λ_c^+ and $2467.71 \pm 0.23\,\mathrm{MeV}/c^2$ for Ξ_c^+. Noteworthy is the fact that the accuracy of mass measurement is better than known in the PDG. The mass distributions with fitted models for each year of data taking are in Appendix 9.

Since it is not achievable to completely separate signal and background events (only statistical separation is possible using the sPlot technique), the test of *CP* asymmetry will be performed for the candidates reconstructed in the mass window defined as $\pm 20\,\mathrm{MeV}$ around the expected mass of Λ_c^+ and Ξ_c^+. It corresponds to around 2.5σ and 3σ (for single Gaussian function σ is taken directly from fitted model and for double Gaussian function with different widths σ_1 and σ_2, σ is calculated as $\sqrt{\sigma_1^2 + \sigma_2^2}$) for Λ_c^+ and Ξ_c^+, respectively. Finally, there are **12 029 640** Λ_c^+ decays (signal plus background) and **1 670 068** Ξ_c^+ decays (signal plus background). The individual numbers of selected events with Λ_c^+ and Ξ_c^+ in each year are shown Table 15. The purities of the final Ξ_c^+ and Λ_c^+ samples are 75% and 96%, respectively.

Table 14: Final reconstructed numbers of Λ_c^+ and Ξ_c^+ candidates obtained from the fit to the mass distribution in Down and Up samples in 2016, 2017 and 2018. Likewise, the total yield is received from the fit to the mass spectra of mixture of Down and Up data sets.

Λ_c^+	2016	2017	2018
Down	$1\,679\,400 \pm 7\,400$	$1\,863\,600 \pm 5\,000$	$2\,008\,400 \pm 6\,400$
Up	$1\,536\,000 \pm 3\,700$	$1\,795\,900 \pm 5\,500$	$1\,947\,100 \pm 5\,100$
Total	$3\,212\,000 \pm 8\,700$	$3\,659\,200 \pm 8\,900$	$3\,954\,800 \pm 7\,800$

Ξ_c^+	2016	2017	2018
Down	$146\,680 \pm 880$	$162\,470 \pm 840$	$172\,030 \pm 880$
Up	$141\,760 \pm 840$	$153\,780 \pm 820$	$168\,480 \pm 880$
Total	$288\,400 \pm 1\,200$	$313\,700 \pm 1\,300$	$336\,700 \pm 1\,300$

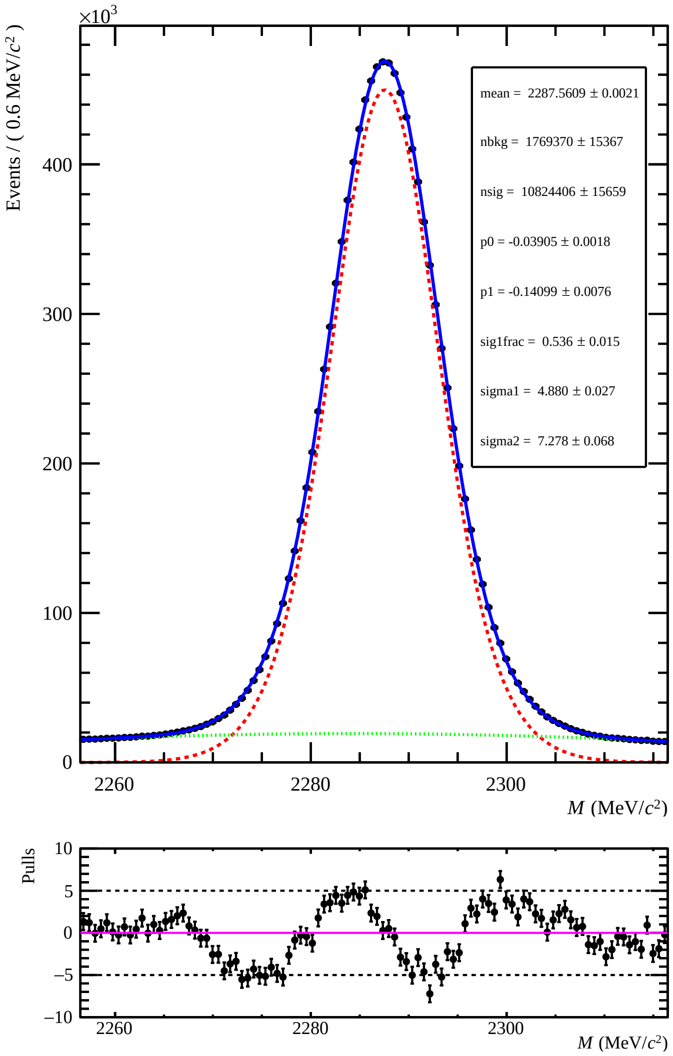

Figure 40: The reconstructed mass of $pK\pi$ **after all implemented cuts** for Λ_c^+ candidates in whole Run 2 data set. The fitted mass model is overlaid (blue line). It is composed of double Gaussian function (red dashed line), with common mean μ and different widths σ_1 and σ_2, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The *nsig* and *nbkg* are the numbers of signal and background events obtained from the fit, respectively. The shown pull distribution is a difference between data and fitted model.

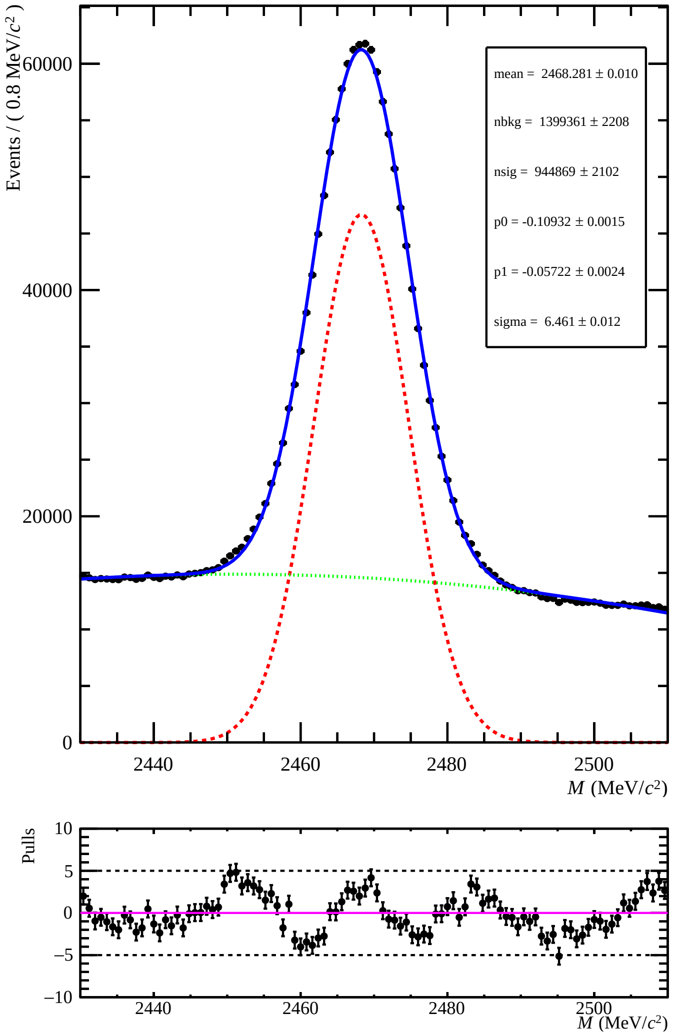

Figure 41: The reconstructed mass of $pK\pi$ **after all implemented cuts** for Ξ_c^+ candidates in whole Run 2 data set. The fitted mass model is overlaid (blue line). It is composed of single Gaussian function (red dashed line), with mean μ and width σ, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (red dotted line) to describe background, where x is the mass of the $pK\pi$. The *nsig* and *nbkg* are the numbers of signal and background events obtained from the fit, respectively. The shown pull distribution is a difference between data and fitted model.

Table 15: Final numbers of $\mathit{\Lambda}_c^+$ and $\mathit{\Xi}_c^+$ events (signal candidates including background events) reconstructed in the mass windows defined as $\pm 20\,\mathrm{MeV}/c^2$ around the expected mass of $\mathit{\Lambda}_c^+$ or $\mathit{\Xi}_c^+$, respectively. Results are given for Down and Up samples in 2016, 2017 and 2018 data sets and whole Run 2.

$\mathit{\Lambda}_c^+$	2016	2017	2018	Run 2
Down	1 866 997	2 061 910	2 233 518	6 162 425
Up	1 710 738	1 986 452	2 170 025	5 867 215
Total	3 577 735	4 048 362	4 403 543	12 029 640

$\mathit{\Xi}_c^+$	2016	2017	2018	Run 2
Down	268 751	278 955	306 871	854 577
Up	248 817	267 304	299 388	815 509
Total	517 568	546 259	606 259	1 670 086

5 Methods

The analysis is based on Dalitz plots representing three-body decays. This phase space is described by the squares of the invariant masses of two pairs of the decay products. This approach has several advantages. First, the investigated process's dynamics are illustrated by any observed structure since the phase space density of the Dalitz plot is uniform. Secondly, strong resonances can be easily spotted in certain mass areas, where events population is intensified, and interference between them is marked as mass band distortion.

In the absence of a *CPV* signal, it is crucial to first confirm this phenomenon. The search can be conducted in a model-dependent and model-independent manner. The latter is much preferred, as search for *CP* asymmetry is more open-ended and exploratory in nature and model-independent approach helps to exploit unknown territories beyond the SM, where new sources of *CP* asymmetry could potentially be found. There are binned and unbinned variants of this approach. The former is based on dividing the Dalitz plot into bins, while the latter version has no such constraints. In this section the methods used to search for *CP* violation are described . The first one is the known and commonly applied binned S_{CP} method [12], the second is unbinned Energy Test method [13–17] and the third is the unbinned Kernel Density Estimation (KDE) technique. The last one has never been used in this type of analysis.

5.1 Binned S_{CP} method

The binned S_{CP} (also called *Miranda*) method is based on dividing the whole phase space (the Dalitz plot) into a given number of bins. For each i^{th} bin, the significance of the difference between particles (N_+^i) and antiparticles (N_-^i) is calculated using the following expression:

$$S_{CP}^i = \frac{N_+^i - \alpha N_-^i}{\sqrt{\alpha(N_+^i + N_-^i)}}, \tag{27}$$

where α is a normalization factor defined as

$$\alpha = \frac{N_+}{N_-} \tag{28}$$

and is introduced in the Eq. 27 to reduce global effects such as production asymmetry which can lead to false signal of *CPV*.

The production asymmetry results in a difference between the total numbers of produced particles and antiparticles. The reason for that is that the LHC is the charge asymmetric *pp* collider. The production asymmetry is a global effect leading to an overall charge asymmetry, which should be constant throughout the Dalitz plot. When computing the S_{CP} value, this effect is taken into account by introducing the constant α as a normalization factor.

After calculating the S_{CP} value in each bin, all measured values of S_{CP} are presented in a plot and used to calculate the corresponding p-value in the χ^2 test ($\chi^2 \equiv \sum(S_{CP})^2$). To avoid statistical fluctuations, only those bins with a minimum occupancy of 10 events for both particles and antiparticles are taken into consideration. The p-value is the probability of obtaining, for a given number of degrees of freedom, the value of the test statistic at least as extreme as the one calculated using the data sample, assuming that the null hypothesis (*CP* is conserved) is correct. It measures the degree to which we are confident that the differences between particles and antiparticles in the Dalitz plots are driven only by statistical fluctuations. Large (> 0.01) p-values make the null hypothesis look good, and small ($< 10^{-7}$) suggests rejection.

If *CPV* is observed, the p-value could be converted into significance for a signal using Gaussian statistics. Without local asymmetries, the S_{CP} distribution is reproduced by the Gauss one with mean $\mu = 0$ and sigma $\sigma = 1$. However, in the scenario that no *CPV* is found, there is no model-independent mechanism for setting limits on *CPV* in the Dalitz plot.

5.2 Unbinned KDE technique

Kernel Density Estimation (KDE) [67] is a non-parametric way to estimate the probability density function $\hat{f}$ of a random variable. KDE is a fundamental data smoothing technique where inferences about population are made based on the finite data sample:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\omega(x - x_i, h) \tag{29}$$

where:

$$\omega(t, h) = \frac{1}{h}K\left(\frac{t}{h}\right) \tag{30}$$

Is weighting function. K is a kernel, which determines the shape of the weighting function, and h is a smoothing parameter, often called kernel bandwidth. In general, function K has to meet the following conditions:

$$\int_{-\infty}^{+\infty} K(x)dx = 1 \tag{31}$$

$$\int_{-\infty}^{+\infty} xK(x)dx = 0 \tag{32}$$

$$\int_{-\infty}^{+\infty} x^2K(x)dx > 0 \tag{33}$$

Essentially, K function needs to be normalized (Eq. 31), its first and second moment are equal to 0 (Eq. 32) and greater than 0 respectively (Eq. 33).

In this analysis, two Triangle kernel function was used, which can be described as follows:

$$w(t, h) = \begin{cases} \frac{1}{h}(1 - |t|/h) & \text{for } |t| < h \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

A critical issue related to the kernel method is the optimal selection of the bandwidth parameter. It has a significant impact on the estimation results. For invariant functions, one can use globally determined bandwidth, often called fixed bandwidth, which is described by the formula:

$$h = \kappa\hat{S}N^{-\frac{1}{5}} \tag{35}$$

where $\hat{S}$ is the sample standard deviation, N is its size and κ is so called correction parameter. It controls the degree of smoothing of the final estimate. It takes a value from 1.06 to 1.44. However, for a more complicated shape, a fixed bandwidth may lead to a large drop in the sensitivity to the estimated distribution. It should depend on the local features of the data. Hence, the adaptive bandwidth parameter h_i^{opt} is estimated using the analysed data's properties. For instance, one can use an approach known as square root law [18]:

$$h_i^{opt} = \frac{h}{\sqrt{f(x_i)}} \tag{36}$$

It is evaluated for each point x_i individually. For the first iteration, one uses globally defined h to determine p.d.f $f(x_i)$ and then optimises the bandwidth parameter. This procedure can be repeated several times till the best value of the smoothing parameter is reached. Since we use a kernel with finite support, supplementary boundary correction is required to maintain proper probability normalisation on the respective edges of the data range.

5.3 Energy Test

The statistical method called the Energy Test [13–17] was developed recently (in 2004) to address one of the main problem in the data science. That problem is to determine whether two samples, measured in a multi-dimensional space, are drawn from the same underlying population. In the Energy Test the test statistics is defined as on analogue of the potential energy in the field of charges characterised by density functions. It correlates the difference between the $X \to abc$ and its *CP*-conjugate p.d.f.s denoted by $f(\vec{x})$ and $\bar{f}(\vec{x})$, in the multivariate space:

$$\begin{aligned} T &= \frac{1}{2}\iint (f(\vec{x}) - \bar{f}(\vec{x}))(f(\vec{x}') - \bar{f}(\vec{x}'))\psi(|\vec{x} - \vec{x}'|)d\vec{x}d\vec{x}' \\ &= \frac{1}{2}\iint [f(\vec{x})f(\vec{x}') + \bar{f}(\vec{x})\bar{f}(\vec{x}') - 2f(\vec{x})\bar{f}(\vec{x}')]\psi(|\vec{x} - \vec{x}'|)d\vec{x}d\vec{x}' \end{aligned} \tag{37}$$

where $\psi(|\vec{x} - \vec{x}'|)$ is a weighting function usually taken as a Gaussian of Euclidean squared distance d^2 between two points in phase space:

$$\psi(\Delta\vec{x}_{ij}) = e^{-\Delta\vec{x}_{ij}^2/2\sigma^2} \tag{38}$$

The distance function ψ drops fast with distance, which increases sensitivity to local asymmetries. The area in which that asymmetry is measured is determined by the adjustable parameter σ.

One can estimate the test statistics T without any knowledge about f or $\bar{f}$:

$$T = \frac{1}{n(n-1)}\sum_{i,j>1}^{n}\psi(\Delta\vec{x}_{ij}) + \frac{1}{\bar{n}(\bar{n}-1)}\sum_{i,j>1}^{\bar{n}}\psi(\Delta\vec{x}_{ij}) - \frac{1}{n\bar{n}}\sum_{i,j}^{n,\bar{n}}\psi(\Delta\vec{x}_{ij}) \tag{39}$$

The energy test essentially calculates the mean distance of events within the n events of the first sample and the $\bar{n}$ events of the second sample, respectively (first two terms). The third term corresponds to the weighted average distance of events in one flavour sample to events of the opposite flavour sample. The fact that the same events are used many times is considered. In the standard test the two samples are those of different flavours, namely particles and anti-particles.

In the case of the null hypothesis, when two samples come from the same population, the expected value of T equals 0. We denote such statistics as T_0. Significant, positive values of T indicate a difference between the samples. However, the distribution of T_0 is unknown and therefore needs to be estimated to measure p-value and interpret the results correctly. The most common way to find the distribution of T-values is to perform a permutation test. It involves preparing data samples by randomly labeling them as particle or antiparticle decays. For each such pair of samples, a T-value is determined. The process is repeated n times and finally a set of permuted T-values $\{T_1, T_2, ..., T_{nperm}\}$ is obtained. Then the p-value is computed as the fraction of the elements in the set that are larger than nominal T calculated for the original samples.

The Energy Test was used several times for CP violation studies in LHCb experiment, but only in mesons decays [68–70].

6 Optimization of the KDE technique

As mentioned in the previous Sec. 5.2, the value of the bandwidth parameter h has the greatest impact on the performance of the KDE method. It is therefore crucial to properly optimise this parameter, especially at the edges of the distribution under consideration. This section presents the optimisation process of the smoothing parameter h, which is carried out on a toy sample with statistics from Run 1.

6.1 1D scenario

This thesis presents the results of the KDE method applied to one-dimensional samples. This includes both sensitivity tests and calculations on control Λ_c^+ samples. The one-dimensionality of the data is based on the fact that the two-dimensional Dalitz spaces were divided into 4 kinematic regions, each about 0.49 $m^2_{K\pi}$ wide, and projected onto the x-axis, where the m^2_{pK} values are distributed. The Dalitz plots for 0% toy sample with the four regions marked with a red line are shown in Fig. 42.

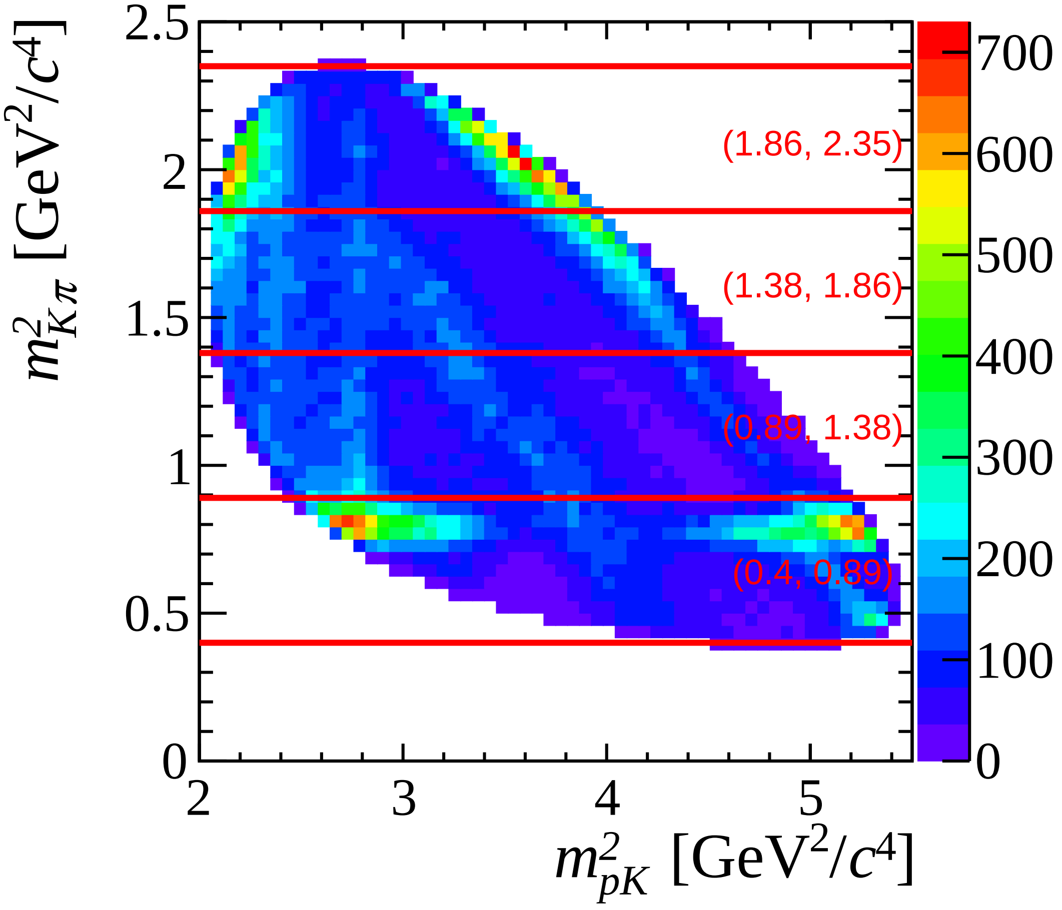

Figure 42: The Dalitz plot for 0% toy sample split into four kinematic regions.

In the next step, estimation and optimization of probability density functions for particles and antiparticles were performed on the one-dimensional data samples prepared in this way.

6.2 Bandwidth optimisation

As it was mentioned in the Sec. 5.2, the choice of the bandwidth parameter has a significant impact on the KDE performance. Therefore before any sensitivity tests can be run for this unbinned technique, the smoothing parameter h must be correctly determined for each entry in each data sample. This procedure has been executed in several steps. First one was to set h as the global value, using Eq. 35. Then adaptive bandwidth parameter h_{opt} was introduced by applying Eq. 36. The estimated functions $\hat{f}(x)$, where x denotes the square of invariant mass $m^2(pK)$, at this stage of optimization, along with the original distributions of x for the 0%, 5% and 20% *CP* toy samples, are shown in the Figs. 43-45 respectively.

For further optimization, only those parts of the estimated function $\hat{f}(x)$, that differed from the distribution of x by a value greater than the so-called threshold value were selected. The value of the threshold was set as 0.04. The suboptimisation for the intervals selected in this manner was as follows. Three points (except for the edges) were chosen for each such section. The first, called x_{max}, was chosen at the point of the largest difference between the estimated function and the previously mentioned histogram. The other two points are located to the left and right of the first point at the places where the discrepancy between the pdf and the histogram is the smallest and are named x_{min}^{left} and x_{min}^{right}, respectively. Then the h_{opt} values from the interval determined by the two minima were replaced by values from quadratic functions interpolated on that interval whose common point and the parabolas' vertex is the x_{max} point. The value of h_{opt} at the vertex was reduced by 90% for the edges and by 80% for the other points relative to the value of h_{opt}. Only two points were selected on the edges: x_{max} and x_{min}^{right}, and x_{max} and x_{min}^{left} for the left and right edges, respectively. The newly created set of $h_{opt}^{new}(x)$ values is finally used to determine the new function $\hat{f}_{new}(x)$. The new density estimates for toy data samples with 0%, 5%, 20% *CP* are presented on the Figs. 46, 47, 48 respectively along with the distribution of the square of the invariant mass $m^2(pK)$.

6.3 The Master KDE

After optimizing the bandwidth parameter and obtaining the new probability density estimate $\hat{f}_{new}(x)$, the next step was to determine the distribution $\hat{f}_{master}(x)$, referred to, in this work, as the Master KDE. Firstly, 100 subsamples were drawn from toy samples, each containing 20000 entries. Then, the KDE technique was applied to every subsample. As a result, a set of 100 estimated probability density distributions was obtained for each point in each sample. The sets of estimated pdfs for 10th, 30th and 50th bins are displayed in the Fig. 49. All distributions are in the shape of the normal distribution, which is the expected result and proves that values of probability density functions estimated for the each bin are consistent for each subsample. Hence, standard deviations extracted from mentioned sets of 100 estimated probability density functions can serve as uncertainty for each bin of the Master KDE.

Finally, the $\hat{f}_{master}(x)$ is determined as the average of the previously estimated values of the probability density function. The Master KDE for 0%, 5% and 20% toy samples are presented in the Figs. 50-55 respectively. For each bin, the formerly calculated uncertainty is marked.

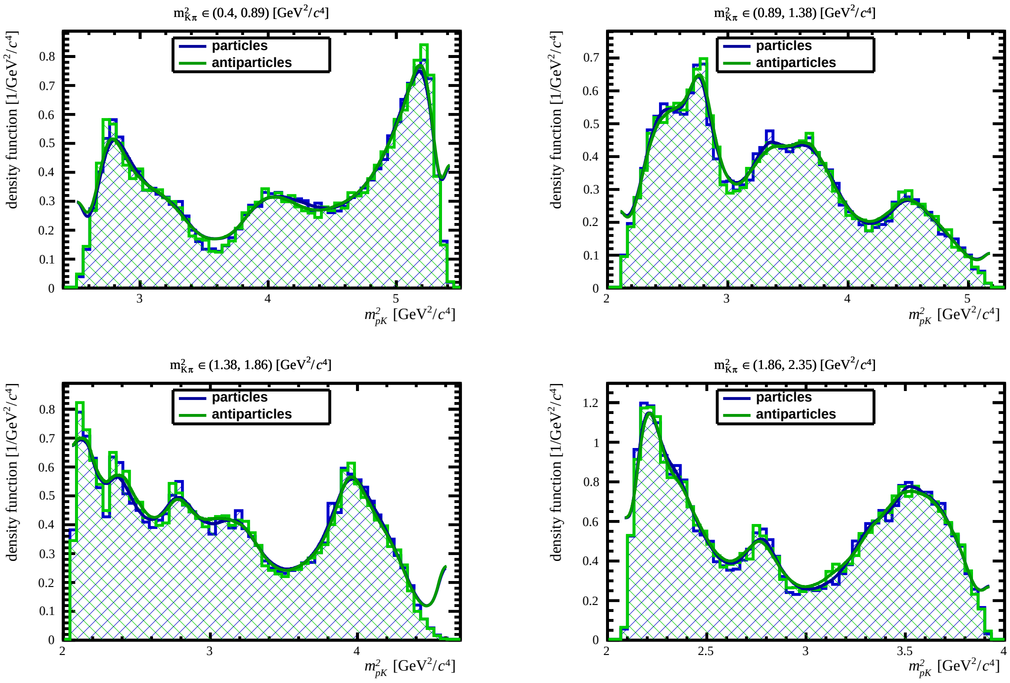

Figure 43: First density estimate $\hat{f}(x)$ for particles (blue line) and antiparticles (green line) for 0% toy sample for four regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

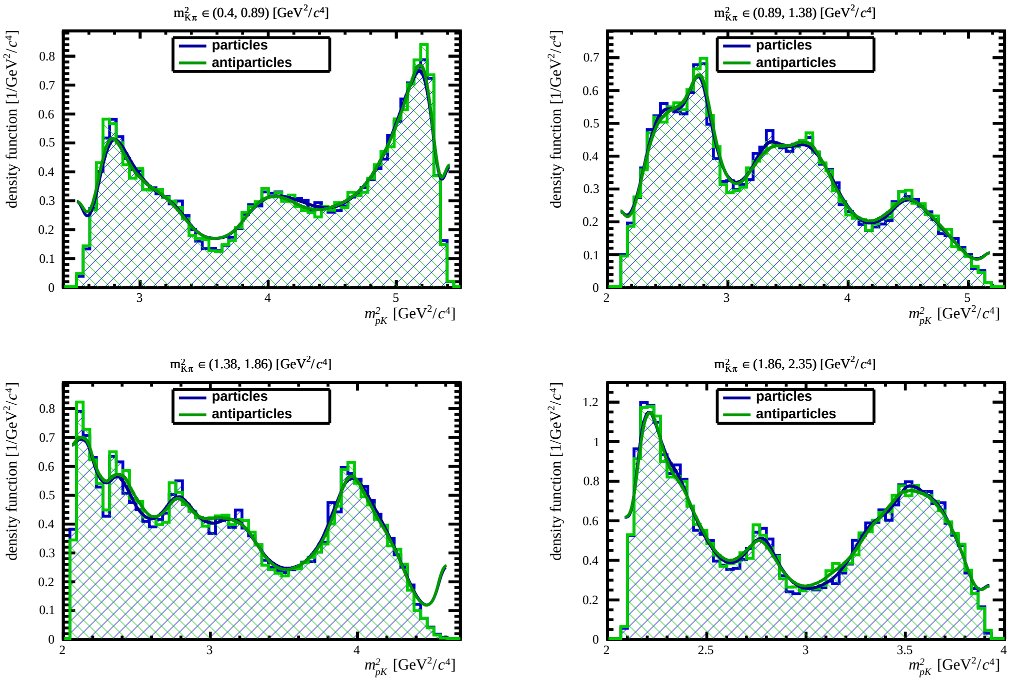

Figure 44: First density estimate $\hat{f}(x)$ for particles (blue line) and antiparticles (green line) for 5% toy sample for four regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

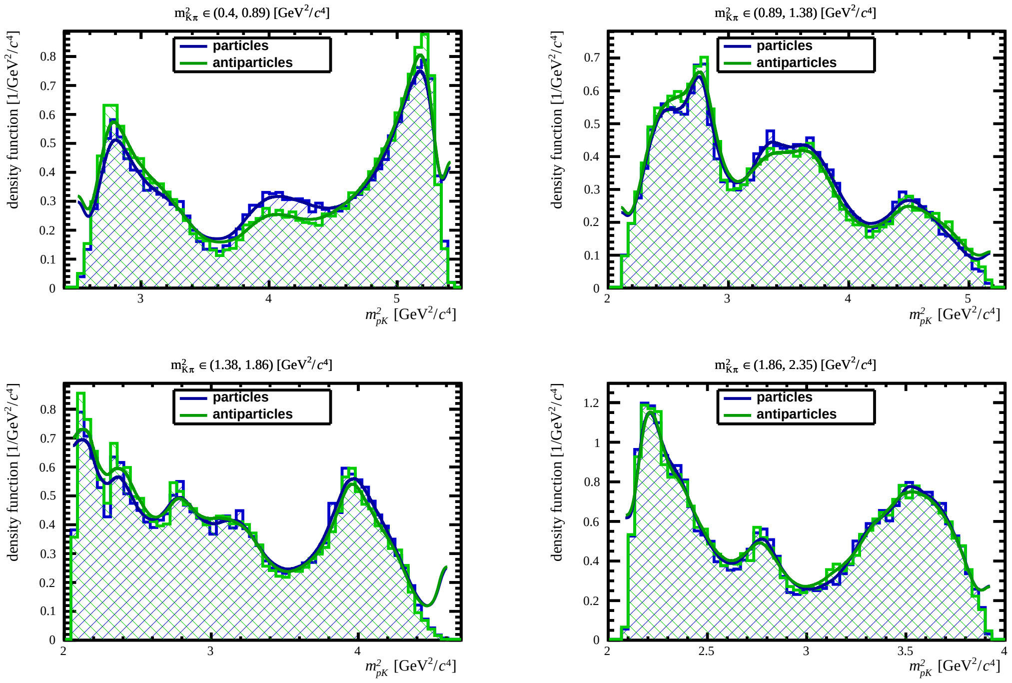

Figure 45: First density estimate $\hat{f}(x)$ for particles (blue line) and antiparticles (green line) for 20% toy sample for four regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

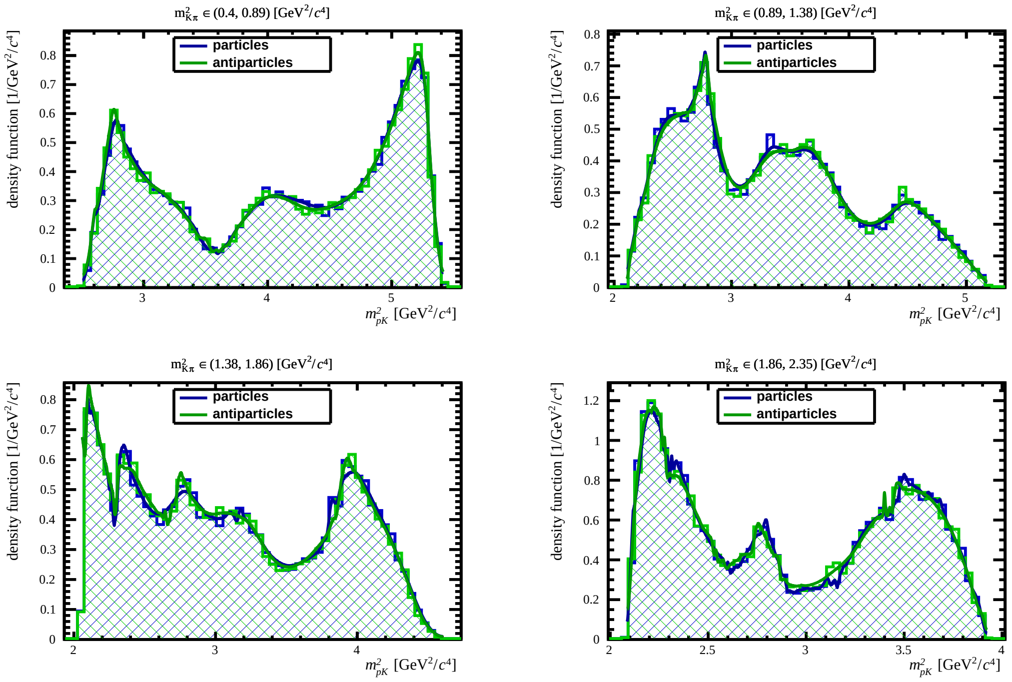

Figure 46: Optimised density estimate $\hat{f}(x)$ for particles (blue line) and antiparticles (green line) for 0% toy sample for four regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

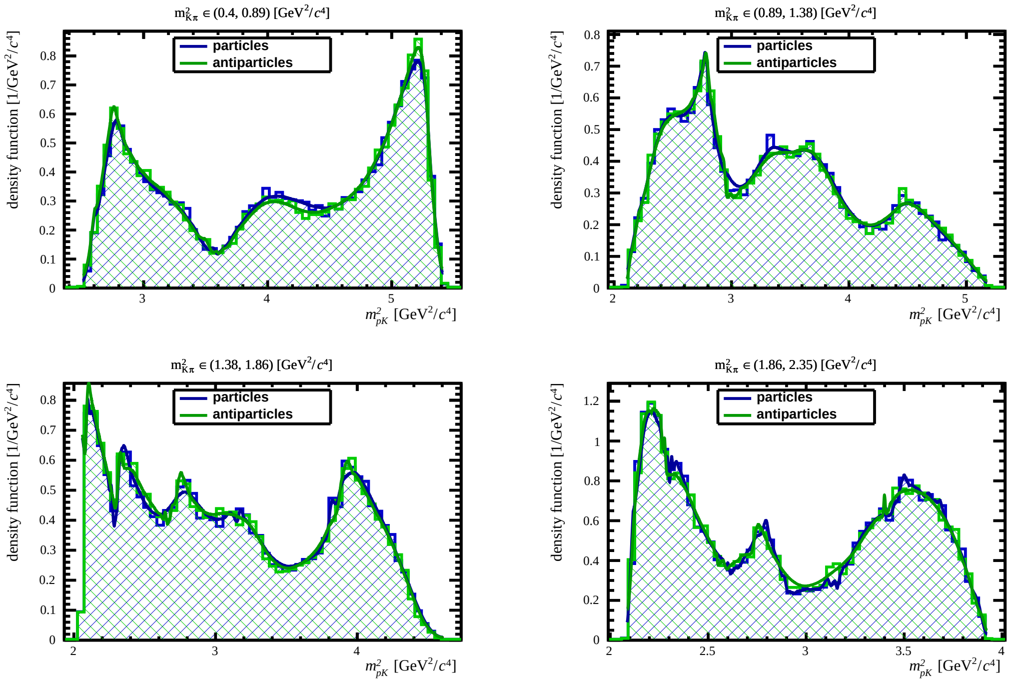

Figure 47: Optimised density estimate $\hat{f}(x)$ for particles (blue line) and antiparticles (green line) for 5% toy sample for four regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

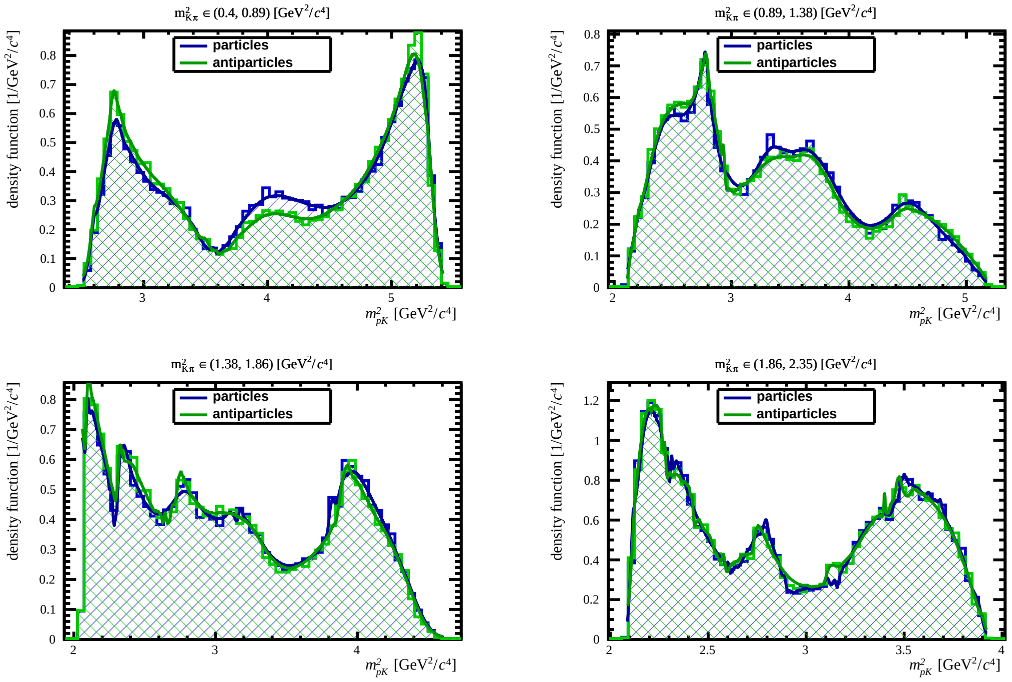

Figure 48: Optimised density estimate $\hat{f}(x)$ for particles (blue line) and antiparticles (green line) for 20% toy sample for four regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

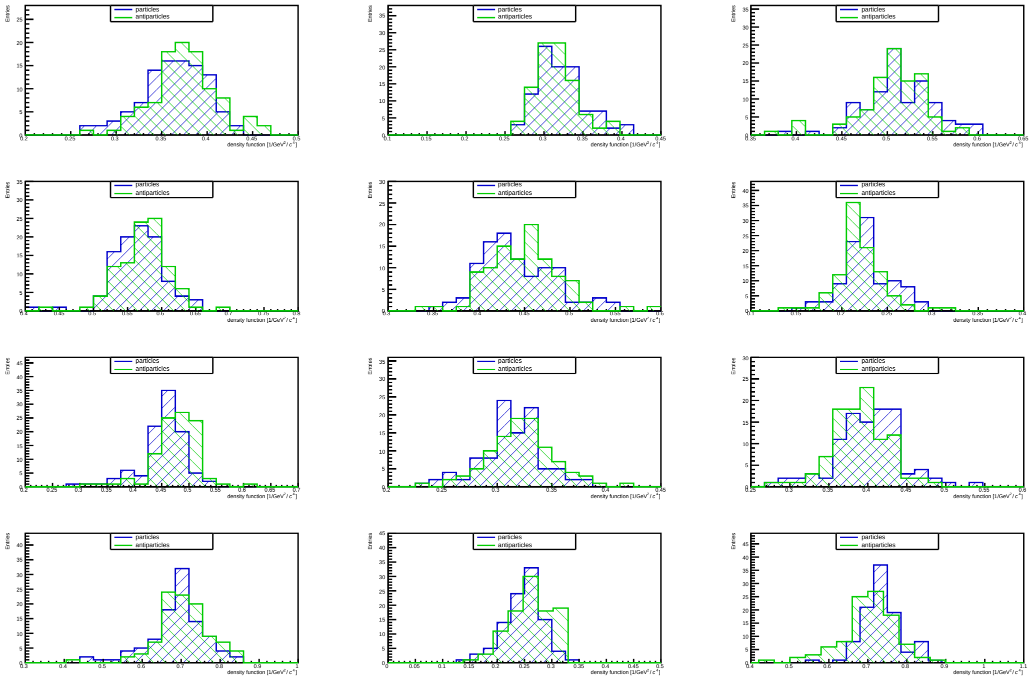

Figure 49: Distributions of values of density functions estimated for the 10th (left side), 30th (middle) and 50th (right side) bins for each region of the Dalitz plot for sample with 0% CP asymmetry. Plots for region no. 1 are in the 1st row, plots for region no. 2 are in 2nd row, plots for region no. 3 are in 3rd row and plots for region no. 4 are in 4th row. Distributions for particles are drawn with blue lines and distributions for antiparticles are drawn with green lines.

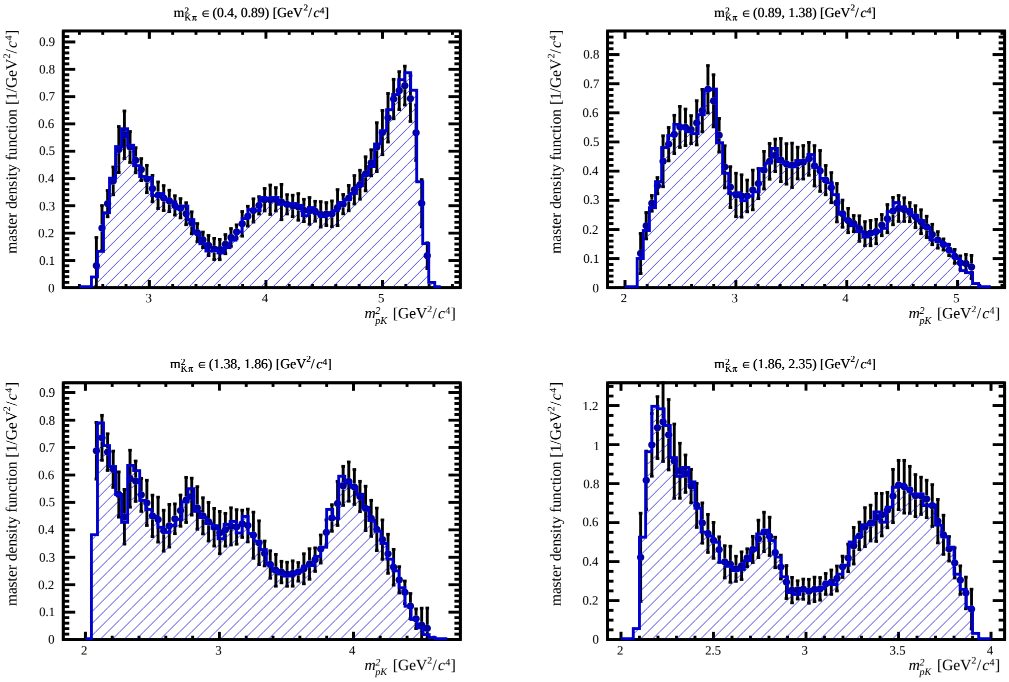

Figure 50: The Master KDE $\hat{f}_{master}(x)$ density function for particles for 0% toy sample.

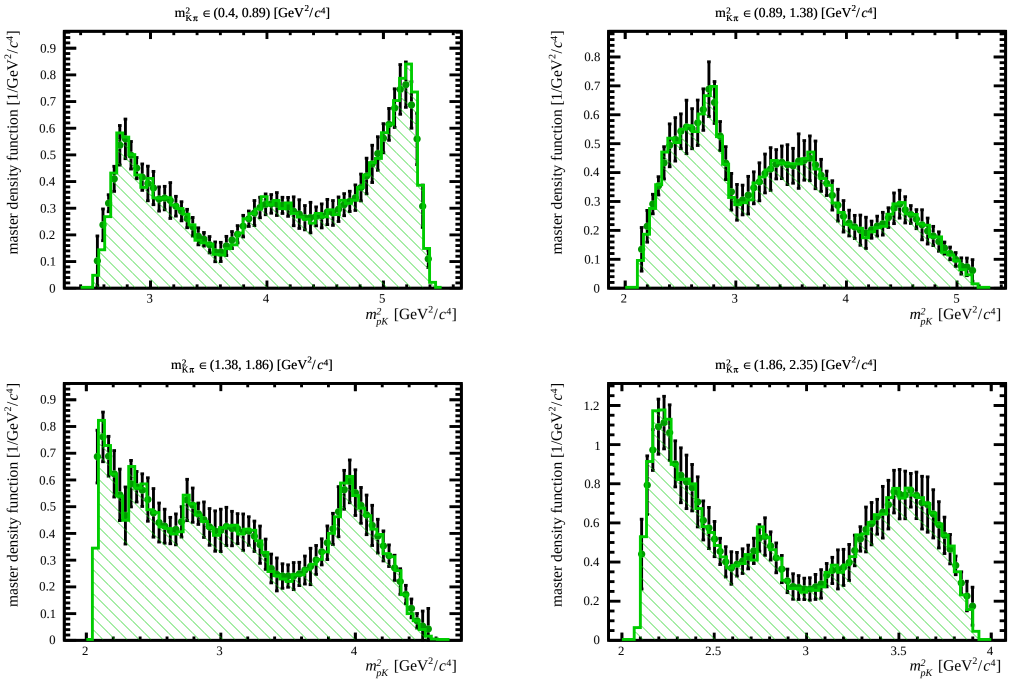

Figure 51: The Master KDE $\hat{f}_{master}(x)$ density function for antiparticles for 0% toy sample.

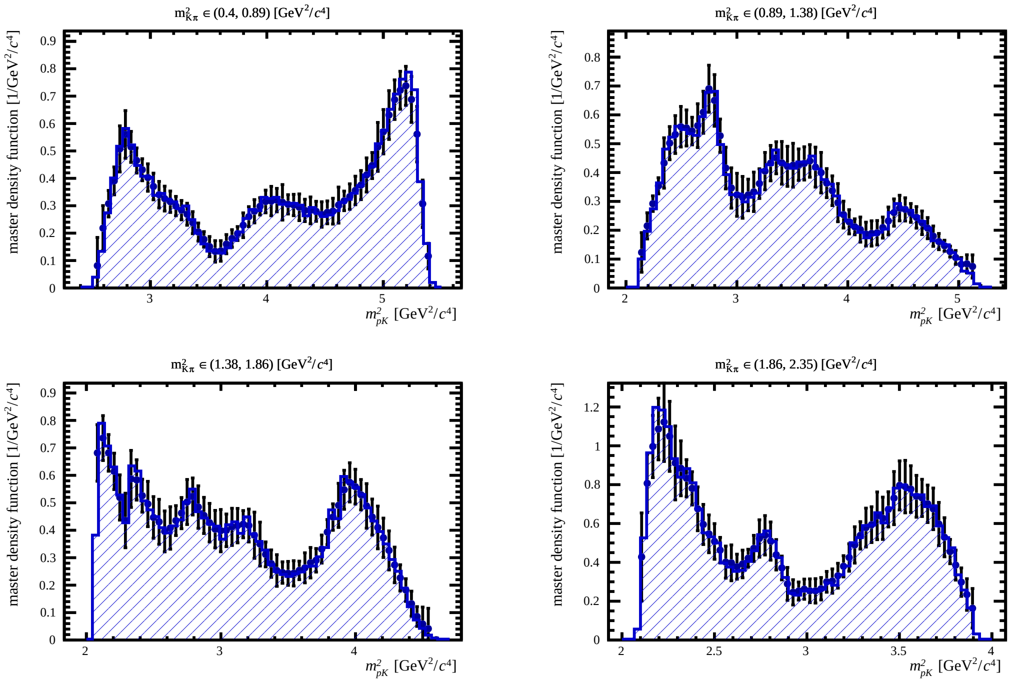

Figure 52: The Master KDE $\hat{f}_{master}(x)$ density function for particles for 5% toy sample.

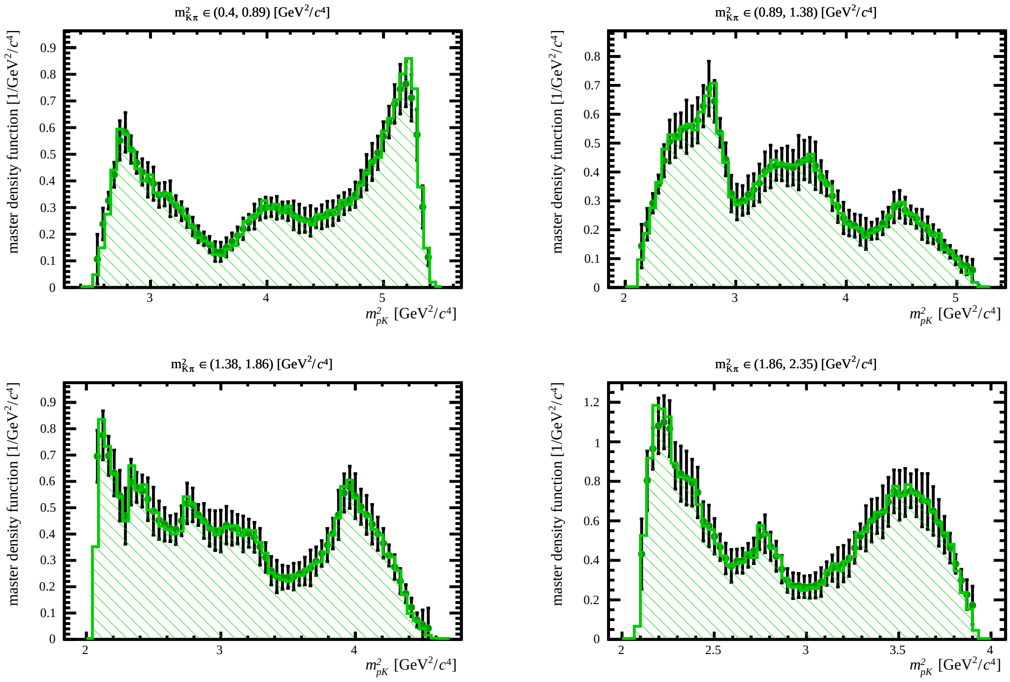

Figure 53: The Master KDE $\hat{f}_{master}(x)$ density function for antiparticles for 5% toy sample.

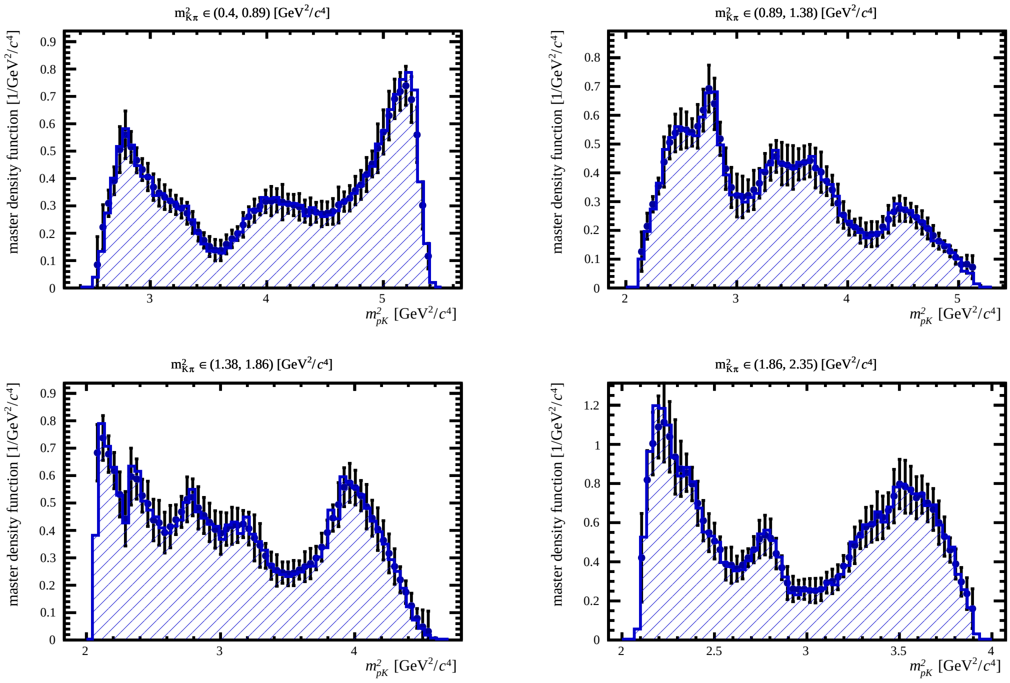

Figure 54: The Master KDE $\hat{f}_{master}(x)$ density function for particles for 20% toy sample.

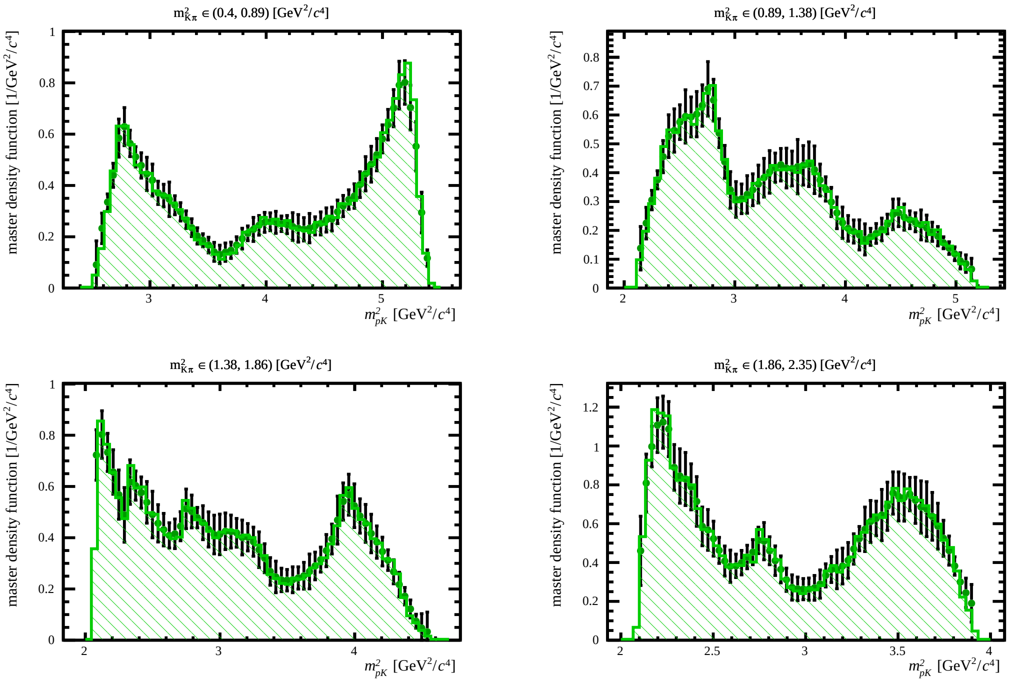

Figure 55: The Master KDE $\hat{f}_{master}(x)$ density function for antiparticles for 20% toy sample.

The determined Master KDE functions agree with the corresponding distributions determined directly from the data and presented as histograms. It is the last step of the optimization process. Very similar procedure is then applied to real data samples and it is described in the Subsec. 8.3.

7 Method sensitivities

In the absence of observations of *CP* violation in charm baryons, it is essential to carry out so-called sensitivity studies to determine which values of *CP* asymmetry are detectable by the methods used for a given data statistics.

In this section sensitivity tests of the binned S_{CP} method and the Energy test method are performed. These sensitivities are studied using Monte Carlo samples. They are generated with different value of the *CP* asymmetry. The generated data without *CP* asymmetry are taken into account as a reference point. Two aspects of tested methods are checked: the method should not generate the signal of *CP* asymmetry if this asymmetry does not exist and the method should confirm signal of *CP* asymmetry when such asymmetry exists.

7.1 Toy MC generator

To test the methods and study their sensitivity and power, a dedicated Monte Carlo toy experiments with the $\Xi_c^+ \to pK^-\pi^+$ decays are generated. In this model, the Dalitz-plot is constructed using the isobar formalism as follows:

$$M(\vec{x}) = a_{nr}e^{i\phi_{nr}} + \sum_r a_r e^{i\phi_r} A_r(\vec{x}), \tag{40}$$

where $\vec{x}$ represents the position in the Dalitz plot, which is described by squared of invariant masses $m^2(pK)$ and $m^2(K\pi)$ and $ae^{i\phi}$ accounts for complex amplitude. A small admixture of continuum (nonresonant) event is denoted by nr. The resonant amplitudes, denoted A_r contain contributions from relativistic Breit-Wigner lineshapes, to describe the propagators, multiplied by angular barrier terms and corrected by Blatt-Weisskopf form factors [71] and spin factors obtained using the Zemach formalism [72, 73] . For example this model is used in [74]. The final probability distribution function defined over the Dalitz phase space is calculated using the formula:

$$f(\vec{x}) = \frac{|M(\vec{x})|^2}{\int |M(\vec{x})|^2 d\vec{x}}. \tag{41}$$

The toy model for $\Xi_c^+ \to pK\pi$ decays is built using the resonances which can be spotted in real data, as: K^*, $K^*(1410)$, $K^*(1430)$, $\Lambda(1520)$, $\Lambda(1600)$, $\Lambda(1670)$, $\Lambda(1690)$, $\Lambda(1710)$, $\Lambda(1800)$, $\Lambda(1810)$, $\Lambda(1820)$, $\Lambda(1830)$, $\Lambda(1890)$, $\Delta(1232)$, $\Delta(1600)$, $\Delta(1620)$ and $\Delta(1700)$. Some of the resonances are very difficult to resolve since they are very close together, for example: $K^*(1410)$ and $K^*(1430)$ or $\Lambda(1800)$, $\Lambda(1810)$, $\Lambda(1820)$ and $\Lambda(1830)$. Using all above resonances and the background assumed to be constant over the phase space, several toy model is created and generated toy data sets contain number of events comparable to real data. The spins of resonances, mother particle Ξ_c^+ and proton in the final state are included in the creation of the model.

First three toy samples containing 200 000 events (100 000 particle and 100 000 antiparticles events) are generated according to a *CP*-conserving model and the other with a moderate amount of *CP*-violation (20%) in the vector resonance K^*. Size of those toy samples correspond to the amount of events collected in Run 1. These data are used for KDE bandwidth optimization discussed in Sec. 6.

The other generated toy samples contain four different magnitudes of *CP* asymmetry: 1%, 2%, 5%, 10% and the fifth is generated according to the *CP*-conserving model. There are 1 700 000 events (850 000 particles and 850 000 antiparticles) in each sample and the purity is set to 75%, which correspond to the size and purity of the samples obtained in Run 2.

The generated Dalitz plots with 200 000 and 1 700 000 events without differences between particles and antiparticles are shown in Fig. 56. The main resonances seen in data are reproduced.

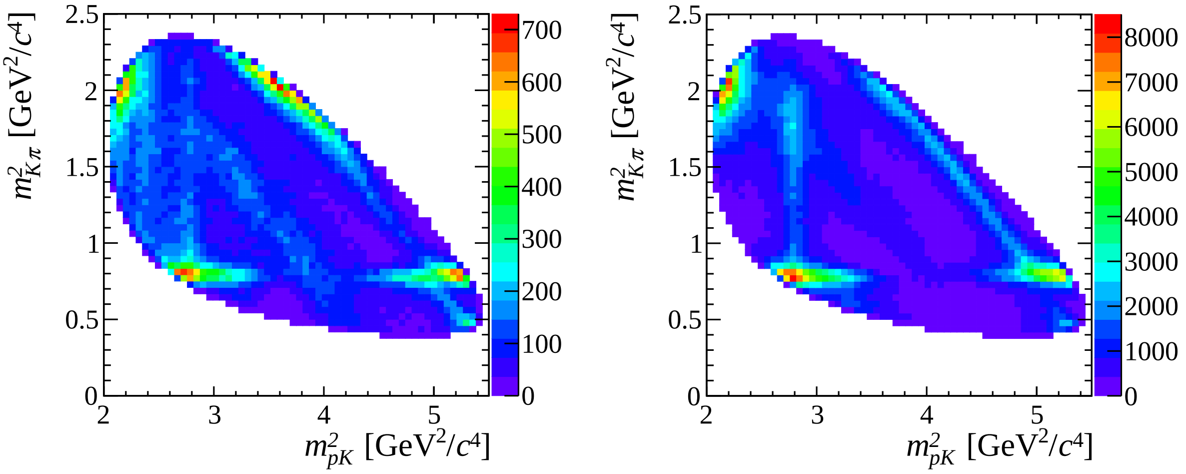

Figure 56: Generated Dalitz plots with 200 000 (left) and 1 700 000 (right) $\Xi_c^+ \to pK^-\pi^+$ events.

7.2 Sensitivity of the $\mathrm{S_{CP}}$ method

The sensitivity tests of the the S_{CP} method are performed in toy samples with *CP* asymmetry (1%, 2%, 5%, 10%) in the amplitude of the resonance K^* and one generated according to a *CP*-conserving model. Each sample contain a total of 1 700 000 events (850 000 particles and 850 000 antiparticles) and 75% purity which correspond to the size and purity of the samples obtained in Run 2. The Dalitz plots are divided into 121 bins in total (11 bins along x and y axes) leading to around 58 bins with more than 10 events. The Dalitz plots with S_{CP} values in bins and the S_{CP} distributions are presented in Figs 57-61 without *CP* asymmetry and with 1%, 2%, 5%, 10% *CP* asymmetry, respectively. The corresponding p-values are shown in Table 16. A Gaussian function is fitted to each the S_{CP} distribution. The results of the fits are displayed in the right upper corner of each plot and also listed in Table 16. For toy sample without *CP* asymmetry both mean (μ) and width (σ), obtained from fit and calculated from data itself, agree with 0 and 1, respectively. The measured p-value corresponds to a significance smaller than 3σ agreeing with no observation of *CP* asymmetry. For samples with 1%, 2%, 5% ans 10% *CP* asymmetry, the means agree with 0 within 3 uncertainties. For samples with 1% and 2% *CP* asymmetry, the widths agree with 0 within 3 uncertainties. For samples with 5% and 10% *CP* asymmetry widths do not match 0. The p-value calculated for sample with 1% *CP* asymmetry corresponds to a significance smaller than 3σ; for sample with 2% *CP* asymmetry the p-value corresponds to 3σ significance level; for samples with 5% and 10%

the p-values are effectively equal to 0, which means it corresponds to the significances much greater than 5σ.

It can be concluded that there are no signs of false CP asymmetry in toy sample without CP asymmetry and null hypothesis of no CP asymmetry cannot be rejected for sample with 1% CP asymmetry. In the other samples with (5%, 10%) CP asymmetry in the amplitude of K^* resonance is observed, whilst 2% seems to be the limit value from which CP asymmetry may become apparent.

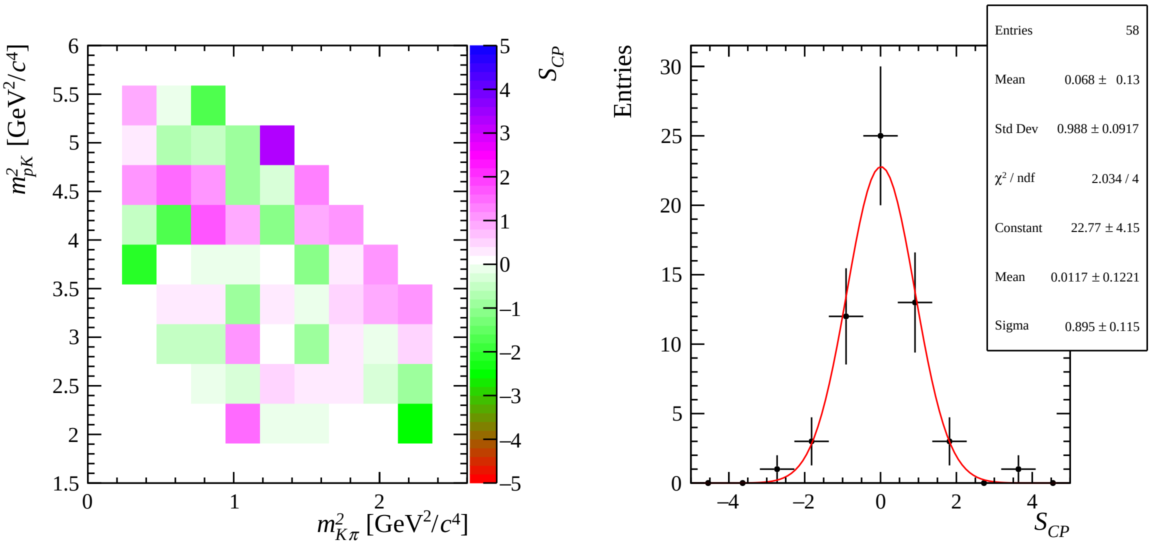

Figure 57: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Xi_c^+ \to pK^-\pi^+$ toy sample without CP asymmetry. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the right plot is shown by the red line.

Table 16: The mean (μ) and width (σ) obtained from the histograms (the first two rows) and obtained from the fit (third and fourth rows) and corresponding p-values (last row) in toy samples with different amount of CP asymmetry for $\Xi_c^+ \to pK^-\pi^+$ decays obtained from χ^2 test. The results correspond to Figs 57-61.

Ξ_c^+	0%	1%	2%	5%	10%
μ	0.07 ± 0.13	0.02 ± 0.16	0.07 ± 0.18	0.22 ± 0.31	0.48 ± 0.57
σ	0.988 ± 0.092	1.18 ± 0.11	1.39 ± 0.13	2.39 ± 0.22	4.35 ± 0.40
μ (fit)	0.01 ± 0.12	0.05 ± 0.24	0.43 ± 0.26	0.88 ± 0.24	1.81 ± 0.31
σ (fit)	0.90 ± 0.12	1.54 ± 0.29	1.52 ± 0.37	1.66 ± 0.24	2.13 ± 0.26
p-value	0.481	0.018	$1.511 \cdot 10^{-5}$	$1.210 \cdot 10^{-40}$	$2.786 \cdot 10^{-195}$

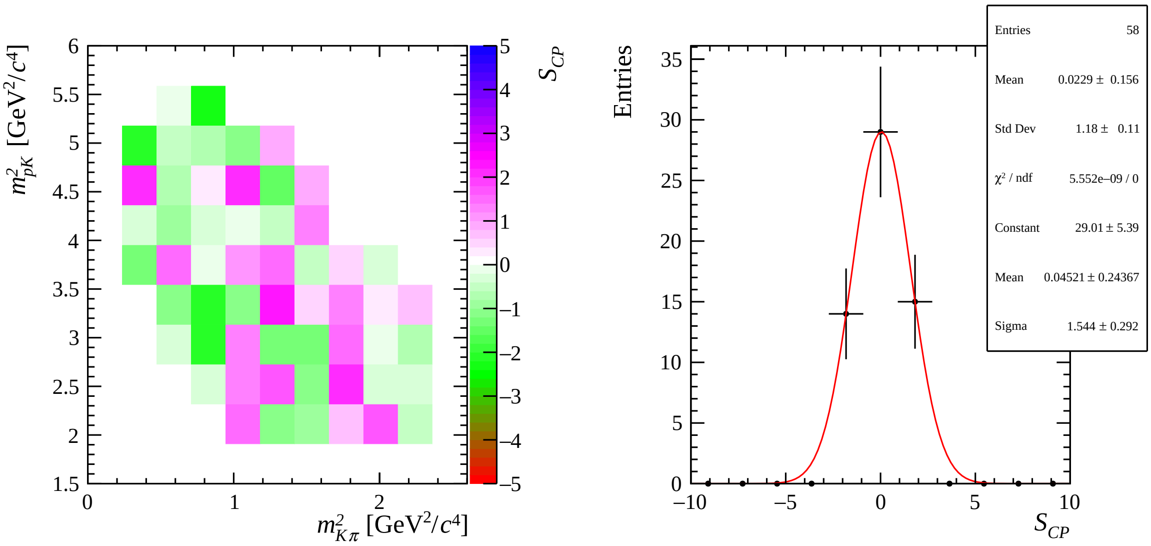

Figure 58: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Xi_c^+ \to pK^-\pi^+$ toy sample with 1% CP asymmetry. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the right plot is shown by the red line.

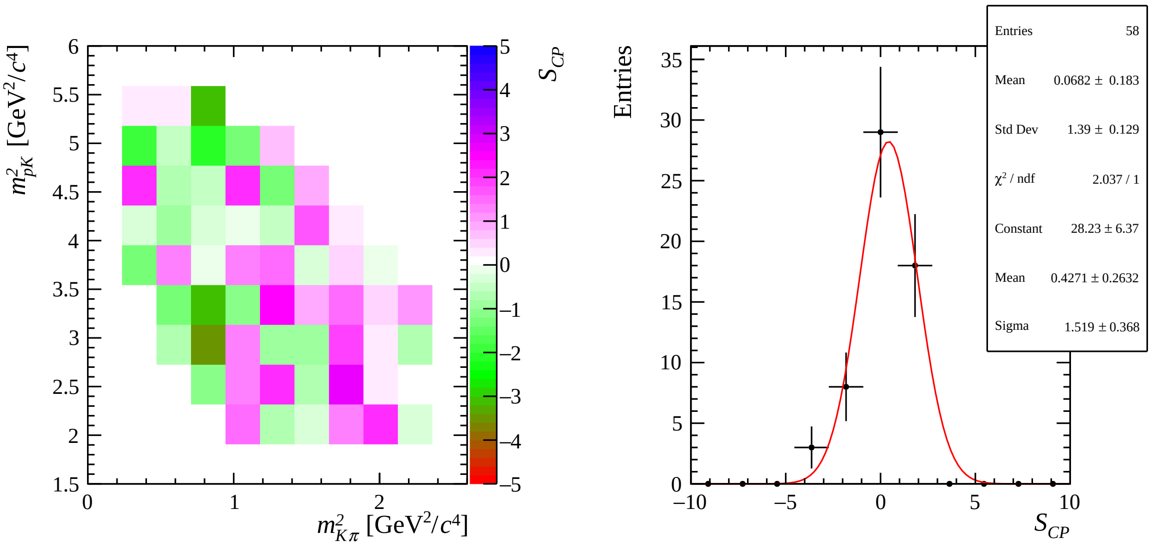

Figure 59: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Xi_c^+ \to pK^-\pi^+$ toy sample with 2% CP asymmetry. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the right plot is shown by the red line.

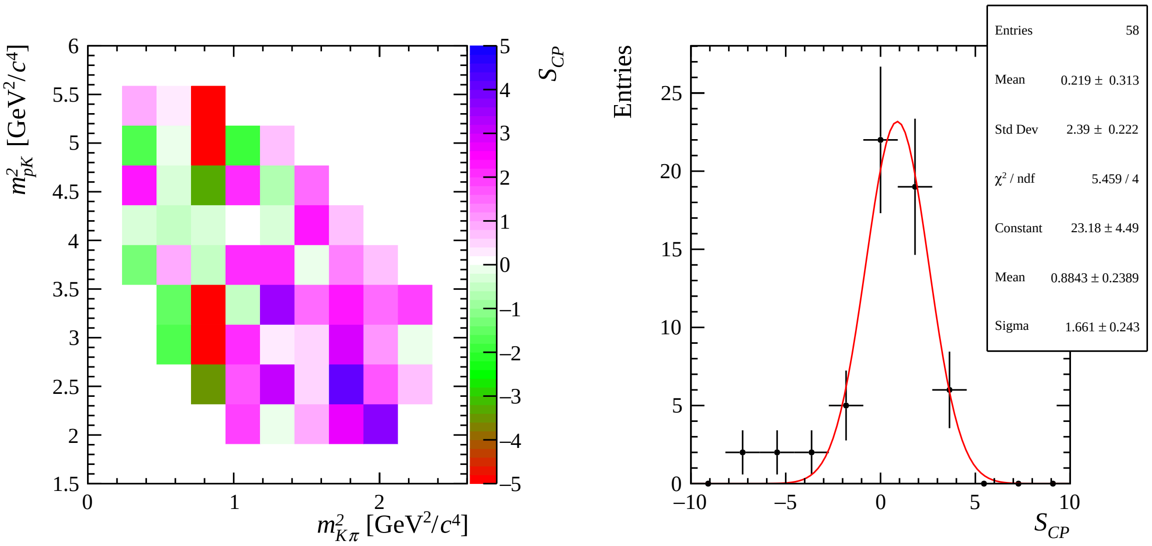

Figure 60: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Xi_c^+ \to pK^-\pi^+$ toy sample with 5% CP asymmetry. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the right plot is shown by the red line.

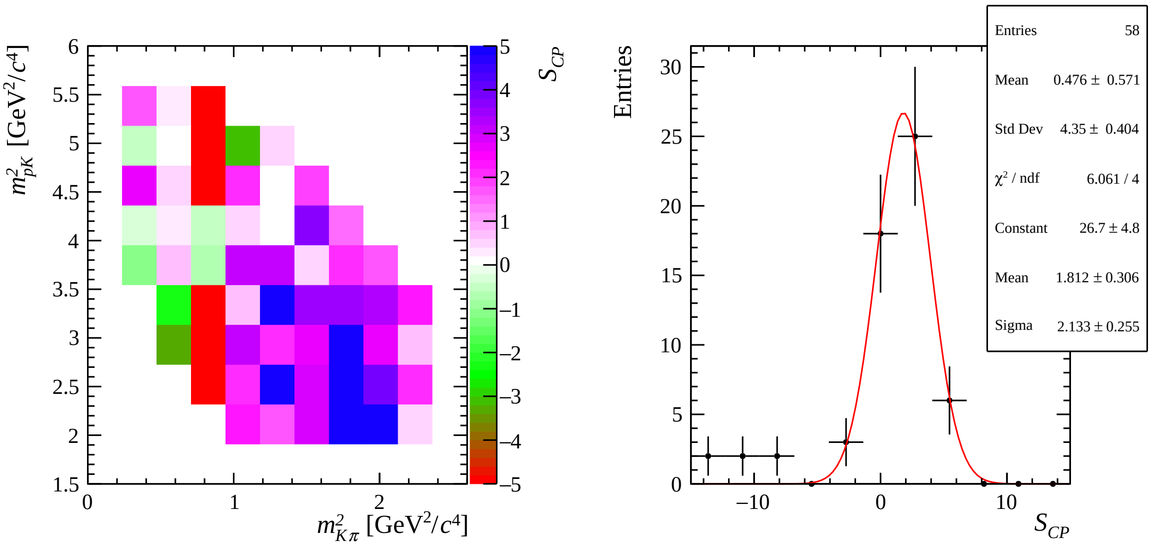

Figure 61: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Xi_c^+ \to pK^-\pi^+$ toy sample with 10% CP asymmetry. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the right plot is shown by the red line.

7.3 Sensitivity of the Energy Test method

The sensitivity tests of the Energy Test method are performed on the same samples on which the sensitivity of the S_{CP} method is tested. Details about those samples are provided in the Subsec. 7.2.

The T- values and the corresponding p-values are shown in Table 17. As it was explained in Subsec. 5.3 the obtain p-values the sets of permuted T-values were prepared. They contain 200 values and were produced with 10 000 permutations in case of toy sample with 0% and 1% CPV, 2 mln for toy sample with 2% CPV and more than 10 mln for samples with 5% and 10% CPV.

Table 17: T values and corresponding p-values for toy samples with different amount of CP violation in K^* amplitudes.

Ξ_c^+	0%	1%	2%	5%	10%
T	$-2.42804 \cdot 10^{-7}$	$9.24189 \cdot 10^{-7}$	$5.09088 \cdot 10^{-6}$	$3.17314 \cdot 10^{-5}$	0.000126783
p-value	0.6325	0.0524	$2 \cdot 10^{-6}$	0.0	0.0

It can be concluded that there are no signs of false CP asymmetry in toy sample without CP asymmetry and null hypothesis of no CP asymmetry cannot be rejected for sample with 1% CP asymmetry. In the sample with 2% CPV p-value corresponds to the 5σ significance level and in the samples with 5% and 10% CP asymmetry p-value is equal to 0, which means that for those three samples the CP asymmetry in the amplitude of K^* resonance the is observed. In summary, the sensitivity of the Energy Test method is similar to the binned S_{CP} method.

8 Results in Λ_c^+ and in sidebands of Ξ_c^+

The analysis is blind meaning that before applying any method in data, where *CP* symmetry can be potentially violated, the calculations are performed in data, where *CP* asymmetry is not expected to check the correctness the methods. Thus, the answer of methods is tested in data with Λ_c^+ control decays and in sidebands of the Ξ_c^+ decays. Once all of the test and cross-checks are completed, request for approval to perform methods in Ξ_c^+ decays is sent to LHCb collaboration. In the section, the results of the binned S_{CP}, the KDE and the Energy Test obtained in control Λ_c^+ and sidebands (the regions outside the $20\,\mathrm{MeV}/c^2$ mass peak) of the Ξ_c^+ decays are discussed.

8.1 S_{CP} results

The S_{CP} method (described in section 5.1) is performed in two scenarios. In the first approach, the method is applied in whole samples for each year of data taking. In the second scenario the S_{CP} method is used in 100 subsamples shosen randomly from total samples in each year of data taking and in whole Run 2 data sample.

In the first scenario, the Dalitz plots are divided into 121 bins in total (11 bins along x and y axes) leading to around 60 bins with more than 10 entries, which are then used in the S_{CP} method. The Dalitz plots with the S_{CP} values in bins and the S_{CP} distributions obtained in control Λ_c^+ data samples are presented in Figs 62-65 in 2016, 2017, 2018 and whole Run 2, respectively. The corresponding p-values are shown in Table 18. A Gaussian function is fitted to each the S_{CP} distribution. The results of the fits are displayed in the right upper corner of each plot and also listed in Table 18. In all cases, the both mean (μ) and width (σ), obtained from fit and calculated from data itself, agree with 0 and 1, respectively. The measured p-values correspond to a significance smaller than 3σ. It can be concluded that there are no signs of false *CP* asymmetries in the Λ_c^+ control decays as it is not predicted in the SM.

The other approach is performed to check the second order effects. The available Λ_c^+ data are divided into 100 subsamples drawn randomly from total samples in each year of data taking and in whole Run 2 data sample. Samples are chosen on a permutation principle with repetitions occurring only between samples. Then, the S_{CP} method is performed in each subsample. The size of each subsample reflects the size of corresponding Ξ_c^+ data sample. There are around 518 000, 546 000, 606 000 and 167 0000 events in each subsample drawn in 2016, 2017, 2018 and whole Run 2 samples, respectively. The phase space in every subsample is divided into about 60 bins with more than 10 events each. The obtained p-value and significance distributions are presented in Figs 66-69 in 2016, 2017, 2018 and whole Run 2, respectively. The Gaussian function is fitted to the latter and its marked as red line. Fit results and histograms parameters can be found in the right upper corner of each plot and listed in Table 19. In all cases, the individual points in the p-value distribution agree with the mean value, represented by the blue line, within the range of the three uncertainties, which means that the p-value distributions are flat. The significance distributions follow the normal distributions with $\mu = 0$ and $\sigma = 1$. All these results show that there are no signs of fake *CP* asymmetry in the Λ_c^+ decays.

Table 18: The mean (μ) and width (σ) obtained from the histograms (the first two rows) and obtained from the fit (third and fourth rows) and corresponding p-values (last row) in each year of data taking for $\Lambda_c^+ \to pK^-\pi^+$ decays obtained from χ^2 test. The results correspond to Figs 62-65.

Λ_c	2016	2017	2018	whole Run 2
μ	0.02 ± 0.14	-0.02 ± 0.13	-0.04 ± 0.14	-0.028 ± 0.140
σ	1.05 ± 0.10	0.962 ± 0.089	1.04 ± 0.10	1.050 ± 0.097
μ (fit)	0.12 ± 0.14	-0.16 ± 0.14	-0.02 ± 0.14	-0.10 ± 0.19
σ (fit)	1.05 ± 0.12	1.01 ± 0.16	1.08 ± 0.12	1.21 ± 0.15
p-value	0.25	0.60	0.26	0.26

Table 19: The mean (μ) and width (σ) obtained from the histograms (the first two rows) and from the fit (last two rows) in each year of data taking in case of dividing samples into 100 subsamples. The results correspond to Figs 66-69.

Λ_c	2016	2017	2018	whole Run 2
μ	0.056 ± 0.096	-0.10 ± 0.08	0.17 ± 0.10	0.061 ± 0.085
σ	0.963 ± 0.068	0.834 ± 0.059	1.01 ± 0.07	0.85 ± 0.06
μ (fit)	-0.036 ± 0.110	-0.14 ± 0.11	0.11 ± 0.12	0.02 ± 0.11
σ (fit)	1.043 ± 0.084	0.933 ± 0.084	1.105 ± 0.099	0.990 ± 0.083

In the case of the sidebands for Ξ_c^+, the Dalitz plots are also divided into 121 bins (11 bins along the x and y axes), resulting in about 60 bins with a number of events above 10. The Dalitz plots with S_{CP} values in bins and the S_{CP} distributions obtained in Ξ_c^+ sidebands are presented in Figs 70-73 in 2016, 2017, 2018 and whole Run 2, respectively. The results of the fits are displayed in the right upper corner of each plot and also listed in Table 20. In all cases, the both mean (μ) and width (σ), obtained from fit and calculated from data itself, agree with 0 and 1, respectively. The measured p-values correspond to a significance smaller than 3σ. It can be concluded that there are no signs of false *CP* asymmetries in the sidebands of Ξ_c^+ decays.

As a cross-check, the S_{CP} is also performed in Dalitz plots divided into 144 (12 bins along x and y axes) and 169 (13 bins along x and y axes) bins corresponding to 70 and 80 bins with more than 10 events, respectively. In all cases, the measured S_{CP} distributions follow the normal distributions with $\mu = 0$ and $\sigma = 1$ and the obtained p-values correspond to significances within 3σ range.

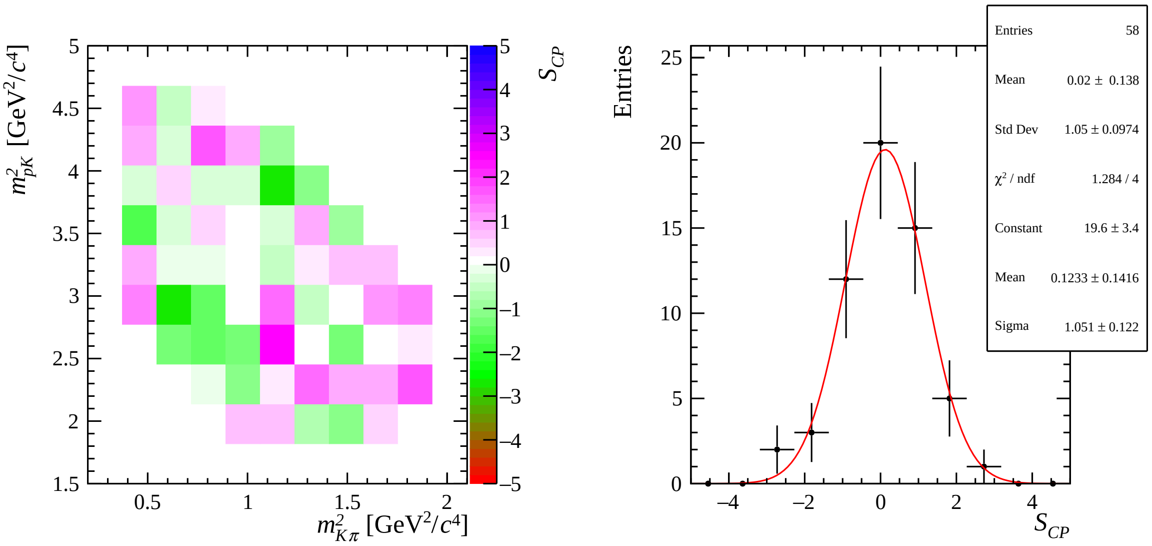

Figure 62: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ in 2016 data sample. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the left plot is shown by the red line.

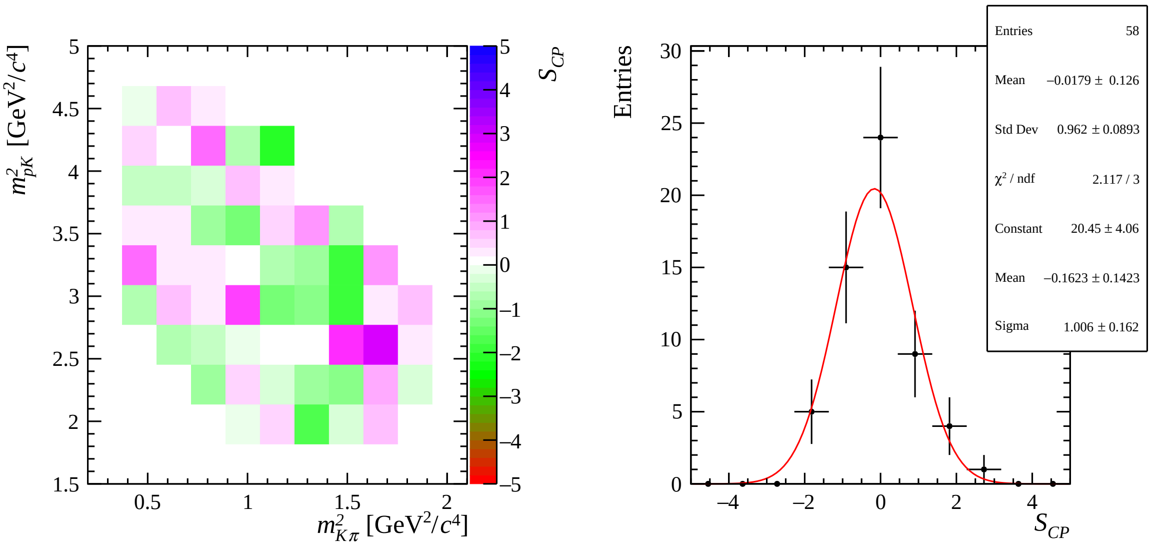

Figure 63: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ in 2017 data sample. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the left plot is shown by the red line.

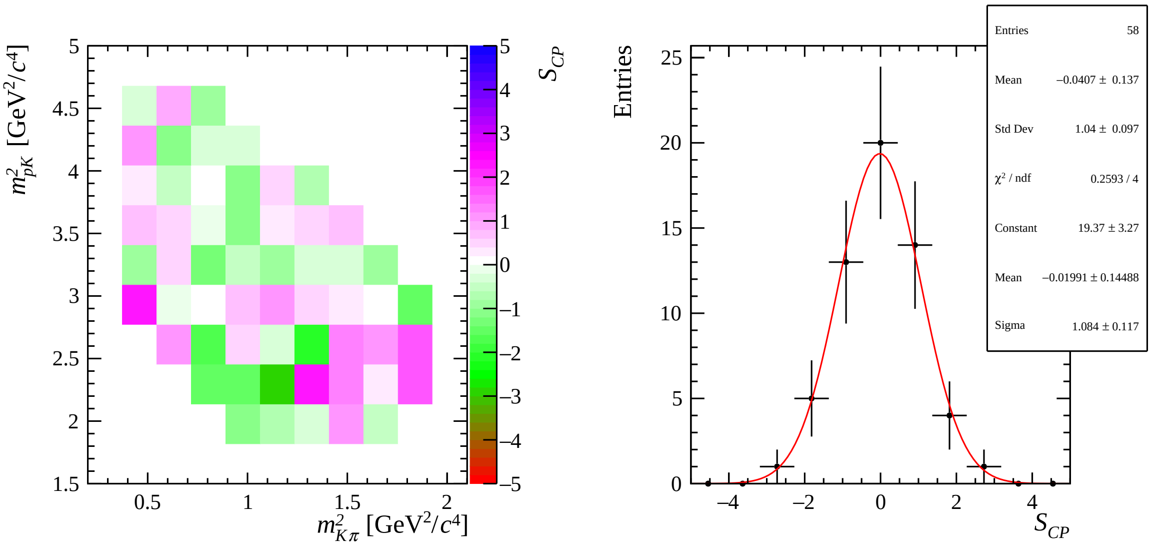

Figure 64: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ in 2018 data sample. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the left plot is shown by the red line.

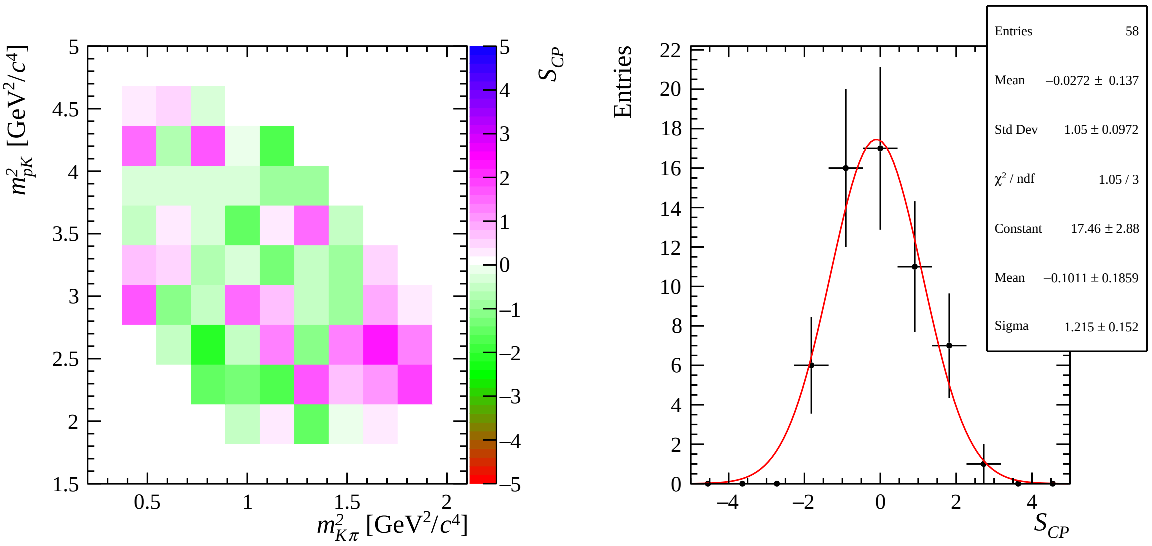

Figure 65: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ in whole Run 2 data sample. The Dalitz plot is divided into 58 bins. The fitted Gaussian function in the left plot is shown by the red line.

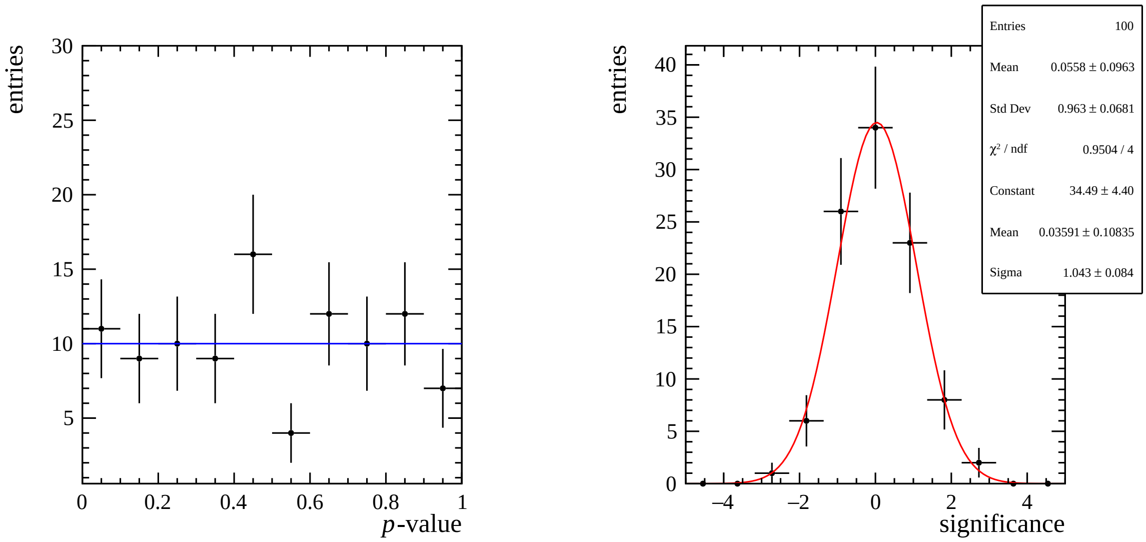

Figure 66: The p-value distribution (left) and corresponding significance distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ decays obtained in 100 subsamples formed from 2016 data. In the left plot the mean value of all bins is overlaid (blue line) and in the right plot Gauss is fitted (red line).

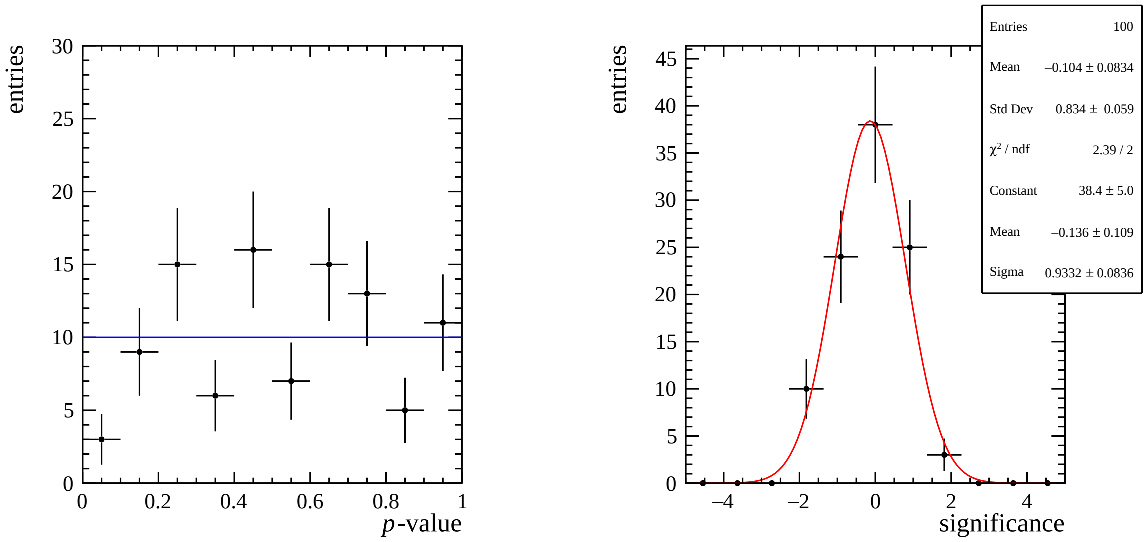

Figure 67: The p-value distribution (left) and corresponding significance distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ decays obtained in 100 subsamples formed from 2017 data. In the left plot the mean value of all bins is overlaid (blue line) and in the right plot Gauss is fitted (red line).

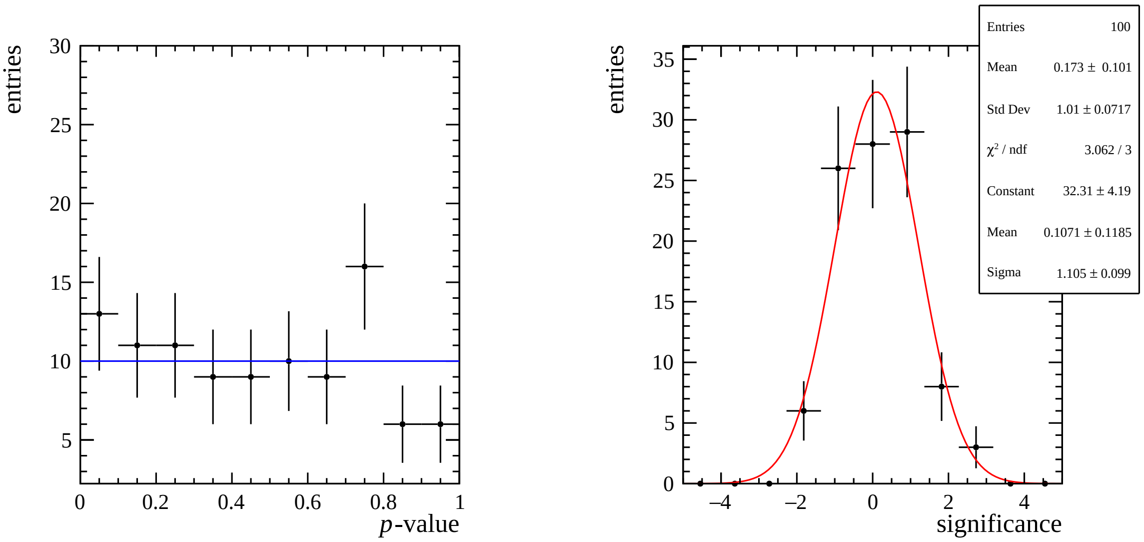

Figure 68: The p-value distribution (left) and corresponding significance distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ decays obtained in 100 subsamples formed from 2018 data.In the left plot the mean value of all bins is overlaid (blue line) and in the right plot Gauss is fitted (red line).

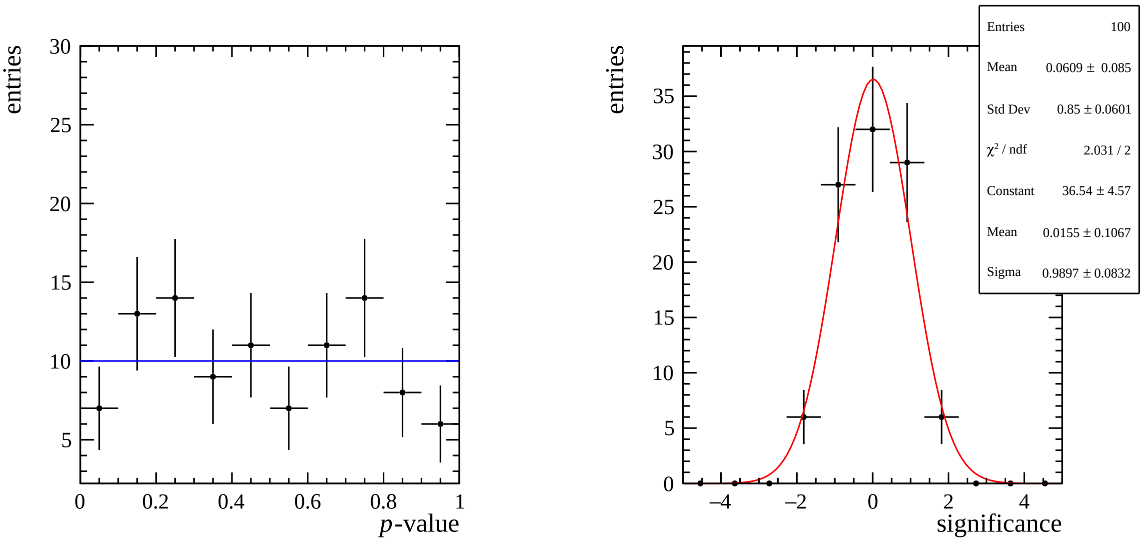

Figure 69: The p-value distribution (left) and corresponding significance distribution (right) for $\Lambda_c^+ \to pK^-\pi^+$ decays obtained in 100 subsamples formed from the whole Run 2 data. In the left plot mean value of all bins is overlaid (blue line) and in the right plot Gauss is fitted (red line).

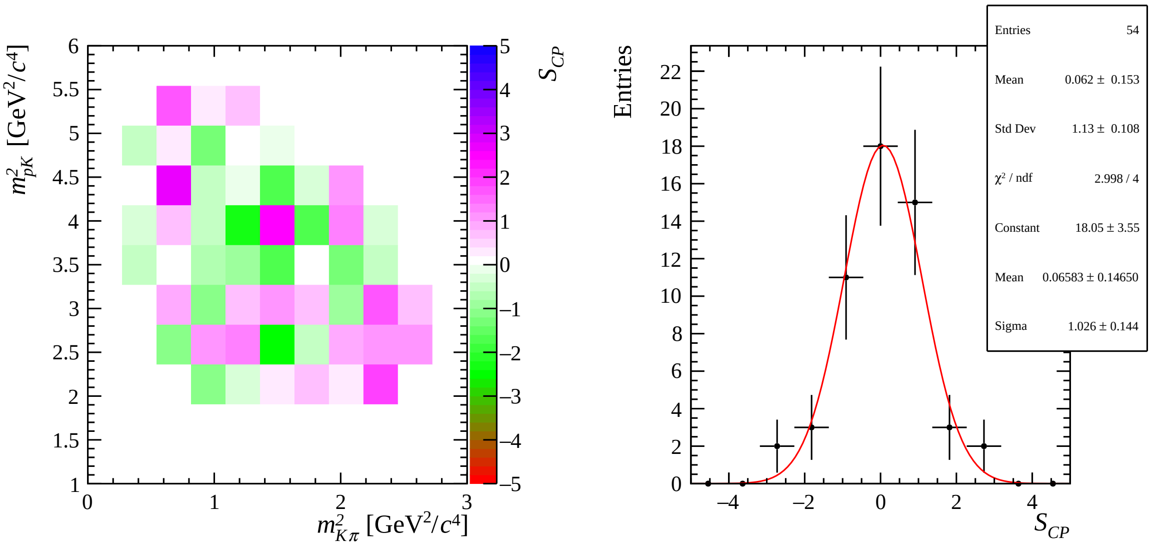

Figure 70: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for sidebands of $\Xi_c^+ \to pK^-\pi^+$ in the 2016 data sample. The Dalitz plot is divided into 54 bins. The fitted Gaussian function in the left plot is shown by the red line.

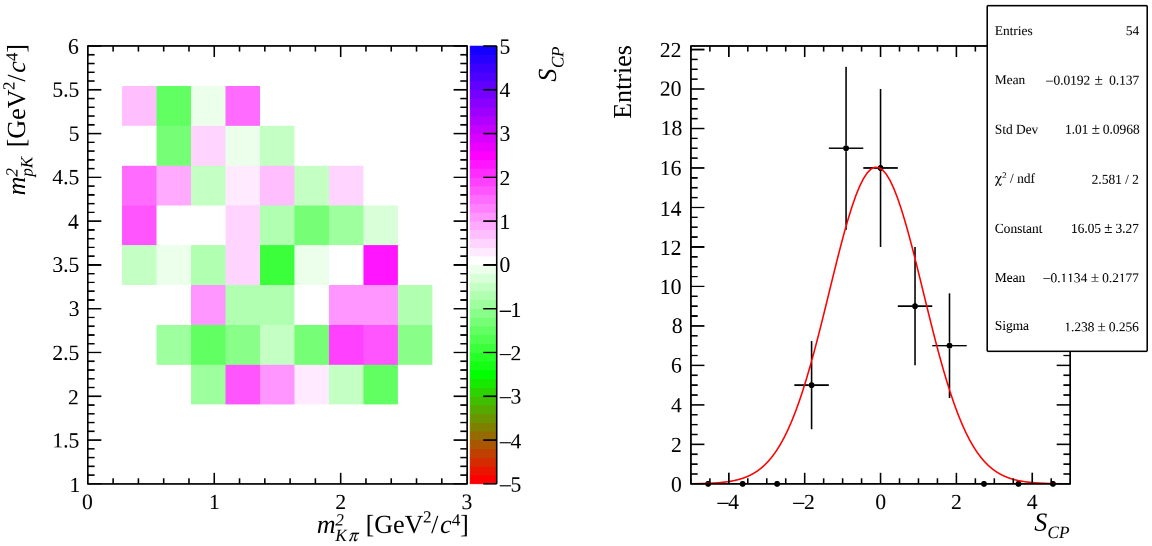

Figure 71: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for sidebands of $\Xi_c^+ \to pK^-\pi^+$ in the 2017 data sample. The Dalitz plot is divided into 54 bins. The fitted Gaussian function in the left plot is shown by the red line.

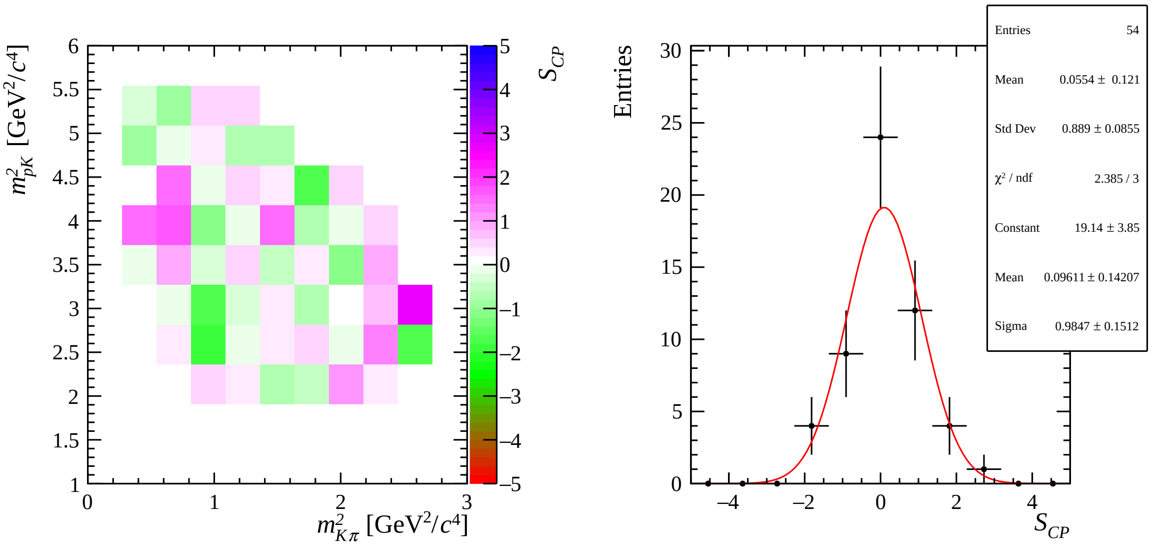

Figure 72: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for sidebands of $\Xi_c^+ \to pK^-\pi^+$ in the 2018 data sample. The Dalitz plot is divided into 54 bins. The fitted Gaussian function in the left plot is shown by the red line.

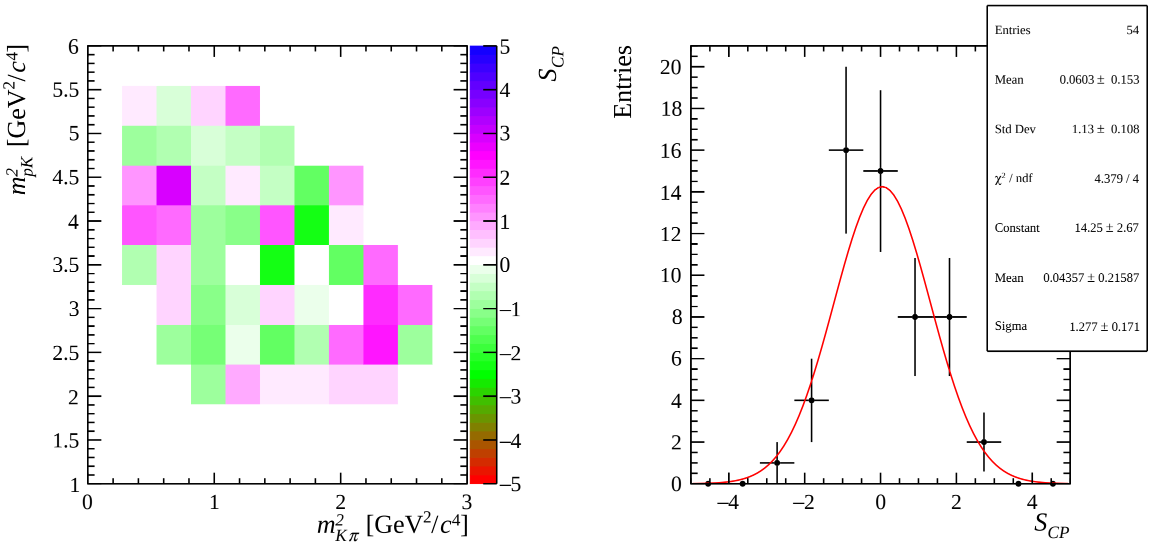

Figure 73: Measured S_{CP} values in bins of the Dalitz plot (left) and corresponding one-dimensional S_{CP} distribution (right) for sidebands of $\Xi_c^+ \to pK^-\pi^+$ in the whole Run 2 data sample. The Dalitz plot is divided into 54 bins. The fitted Gaussian function in the left plot is shown by the red line.

Table 20: The mean (μ) and width (σ) obtained from the histograms (the first two rows) and obtained from the fit (third and fourth rows) and corresponding p-values (last row) in each year of data taking for $\Xi_c^+ \to pK^-\pi^+$ sidebands obtained from χ^2 test. The results correspond to Figs 70-73.

Λ_c	2016	2017	2018	whole Run 2
μ	0.06 ± 0.15	-0.02 ± 0.14	0.06 ± 0.12	0.06 ± 0.15
σ	1.13 ± 0.11	1.01 ± 0.10	0.889 ± 0.085	1.13 ± 0.11
μ (fit)	0.07 ± 0.15	-0.11 ± 0.22	0.10 ± 0.14	0.04 ± 0.22
σ (fit)	1.03 ± 0.14	1.24 ± 0.26	0.98 ± 0.15	1.28 ± 0.17
p-value	0.07	0.41	0.84	0.07

8.2 Energy Test results

Similarly to the S_{CP} method the Energy Test method is firstly performed in the whole Λ_c^+ samples for each year of data taking and secondly in 100 subsamples drawn randomly from total samples in each year of data taking and in whole Run 2. The sizes of subsamples reflect the sizes of corresponding Ξ_c^+ data samples.

For the first test, obtained T values and corresponding p-values are listed in Table 21. The sets of permuted T values consist of 200 values and are produced with 10 000 permutations, which according to [17] is sufficient to reach p-values corresponding to significance of the level of 5σ (procedure of preparing set of permuted T values is described in section 5.3). Further increasing the size and number of permuted samples would only extend the calculation time. The obtained p-values are from 0.015 to 0.37, which correspond to the significances within 3σ range. Those results indicate that *CP* symmetry is conserved in Λ_c^+ decays.

Table 21: T values and corresponding p-values in Λ_c^+ samples for each year of data taking and whole Run 2 data.

Λ_c^+	2016	2017	2018	Run 2
T	$3.61479 \cdot 10^{-7}$	$9.24366 \cdot 10^{-9}$	$2.2938 \cdot 10^{-7}$	$2.16578 \cdot 10^{-7}$
p-value	0.077	0.37	0.10	0.015

In the next step, the Energy Test method is applied in 100 subsamples with Λ_c^+ decays, the same ones previously used in the S_{CP} method. The sets of permuted T-values for the samples are created in the same way as for the original samples - using 10 000 permutations consisting of 200 events. The obtained p-value distributions are shown in Fig. 74. The various points in the p-value distribution are in agreement with the mean value, marked with the blue line, over a range of three uncertainties, which means that p-value distribution is uniform. It indicates that *CP* symmetry in tested subsamples is conserved. It is the expected result since Energy Test method is performed in Λ_c^+ decays for which *CPV* in the SM is negligible. No false signal of *CP* asymmetry has been found.

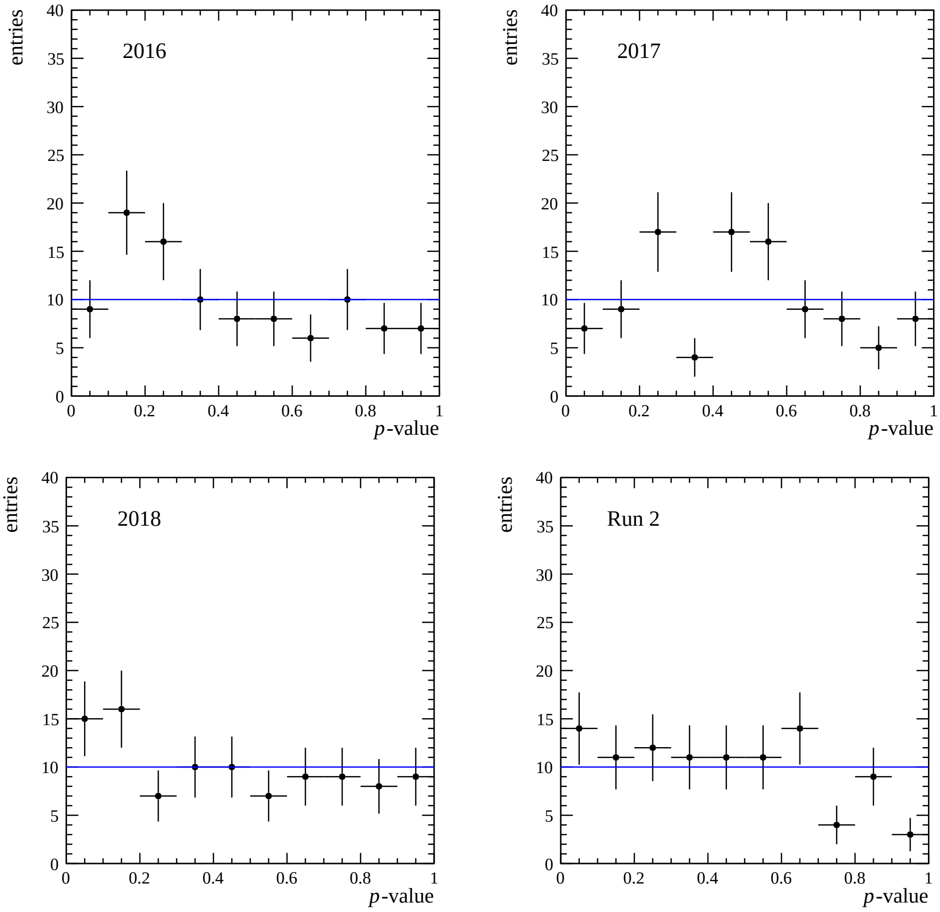

Figure 74: The p-value distributions obtained using the Energy Test method for $\Lambda_c^+ \to pK^-\pi^+$ 100 subsamples formed from data collected in 2016, 2017, 2018 and whole Run 2. Blue lines represent the mean value of all bins.

8.3 KDE results in Λ_c^+ control sample

This section presents the results of the optimisation process of the smoothing parameter h and distribution of the Master KDE in the Λ_c^+ control sample.

8.3.1 1D scenario

Data from the Λ_c^+ control decays were divided into four one-dimensional subsamples, as was the case with the toy samples. The process of preparing the one-dimensional samples is described in the Subsec. 6.1 and the phase space subsample of Λ_c control sample, containing 1000000 events, with the four regions, each about 0.37 $m_{K\pi}^2$ wide, marked with a red line is shown in Fig. 75.

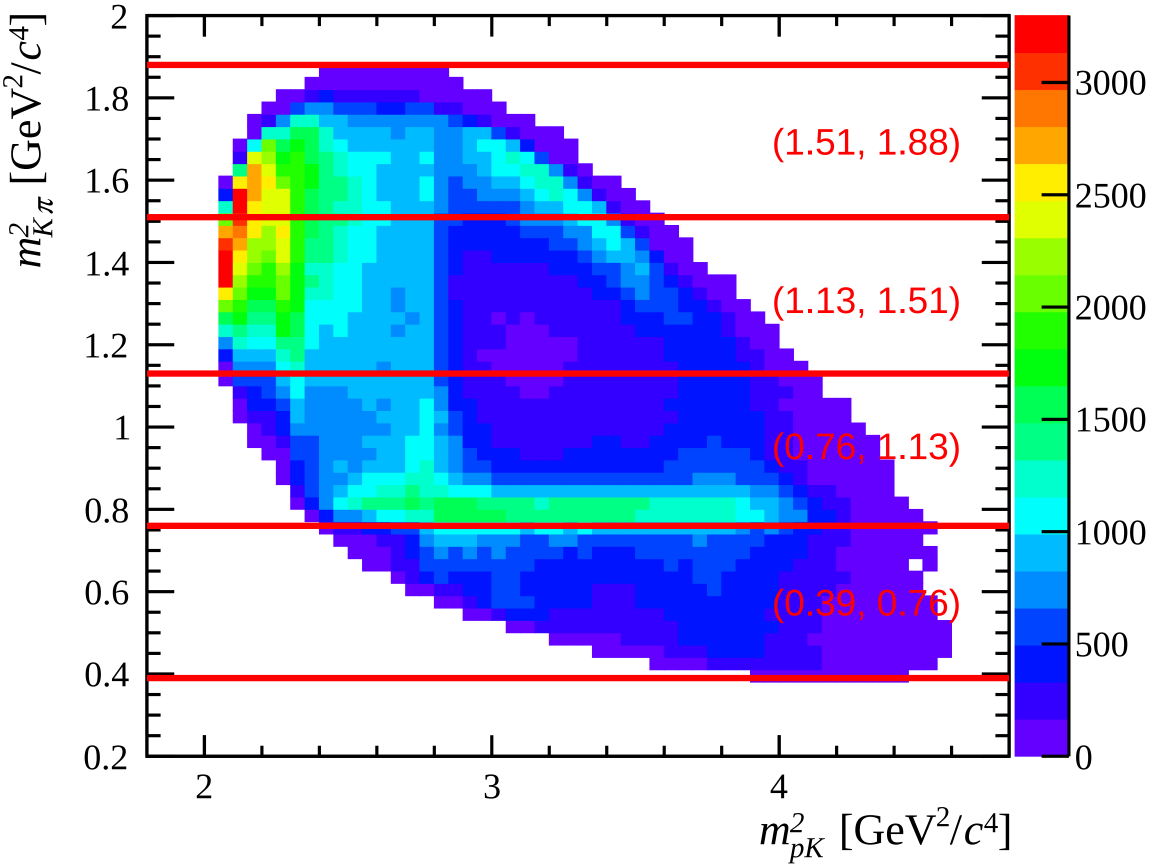

Figure 75: The Dalitz plot of subsample of Λ_c control sample, containing 1000000 events, splitted into four kinematic regions.

As in the case of Monte Carlo samples, each of the four regions is projected onto the horizontal axis resulting in four distributions for particles and antiparticles.

8.3.2 Bandwidth optimisation

The process of optimizing h parameters in control samples starts in the same way as the procedure carried out in toy samples and is described in Subsec. 6.2. Only the threshold value is increased from 0.04 to 0.08. However, it became apparent that for the control sample, additional optimisation was still needed at the edges of the distributions themselves. Two iterations of edge correction are performed, following the same procedure as for the entire mass range, making changes only for a few starting and ending points of a given distribution. For the first threshold is set to 0.05 and for the second 0.025. Plots of the optimized probability density functions for the Λ_c^+ control samples are shown in Fig. 76 below. The optimized particle and antiparticle density functions agree with their corresponding distributions shown as histograms.

8.3.3 The Master KDE

The process of obtaining the Master KDE function in control samples is the same as for toy samples and is described in Subsec. 6.3. Subsample used in the procedure consist of 100000 events. Plots of the optimized probability density functions for the Λ_c^+ control samples are shown in Fig. 76 below and the distribution of the Master KDE is presented in the Figs. 77 and 78 for particles and antiparticles respectively. For each point of the Master KDE distributions, an error is assigned as the standard uncertainty of the pdf distributions determined from the KDE calculations for each subsample. The process of obtaining those uncertainties is described in the Subsec. 6.3 and the histograms of estimated density functions for 10th, 30th and 50th bins are displayed in the Fig. 79. All distributions are in the shape of the normal distribution, which is the expected result.

The Master KDE with errors overlaid for both particle and antiparticle density functions match their corresponding distributions calculated directly from data amd presented as histograms.

The density function optimized the way presented in this work, can in the future be used as part of the Energy Test, in its form not requiring Monte Carlo approximation. It is also a step towards applying the KDE technique for the two-dimensional case, that is, to the whole Dalitz space, without splitting into smaller regions and projections. This, in turn, will make it possible to compare the sensitivity of the KDE with that of the S_{CP} binned method.

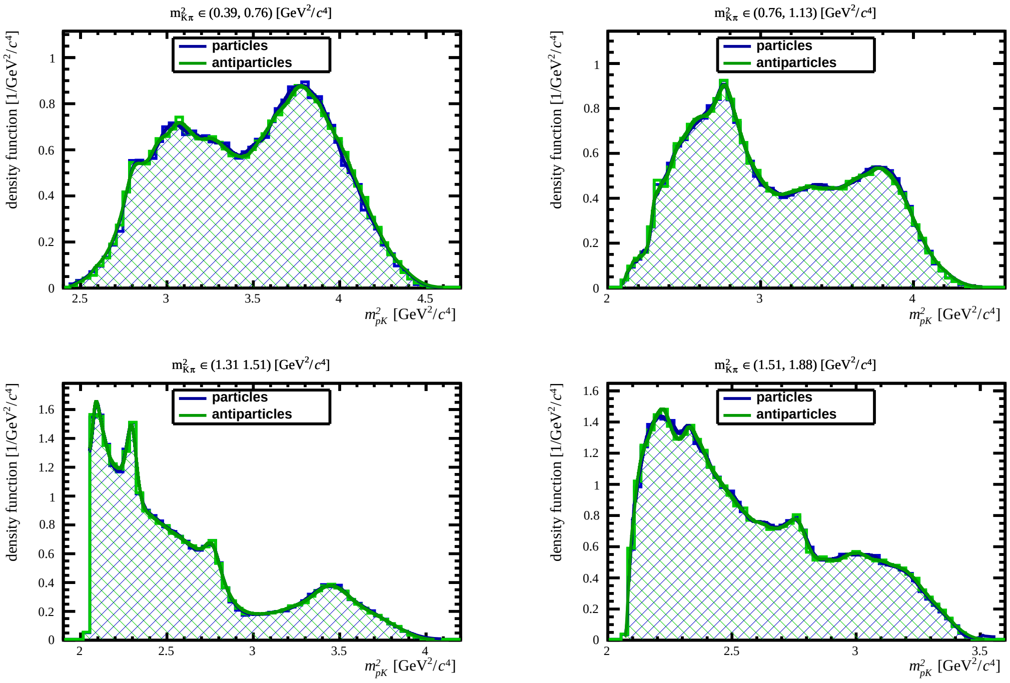

Figure 76: Optimised density estimate $\hat{f}(x)$ for particles (blue line) and antiparticles (green line) for one subsample of control Λ_c^+ sample with 1000000 entries for 4 regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

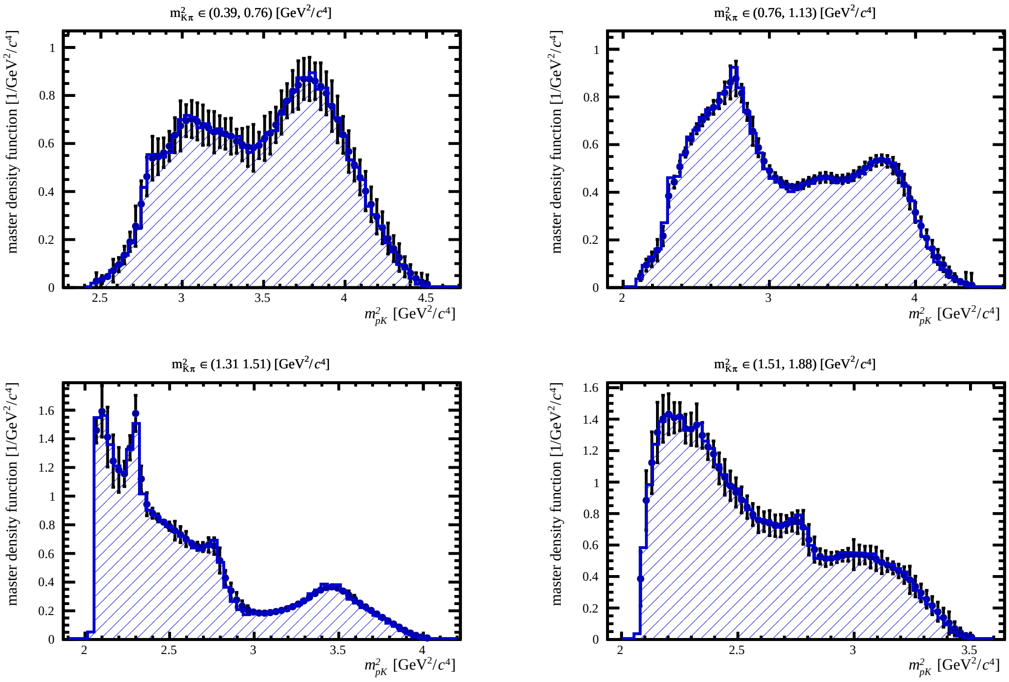

Figure 77: The Master KDE $\hat{f}_{Master}(x)$ density function for particles of control Λ_c^+ sample with 1000000 entries for 4 regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

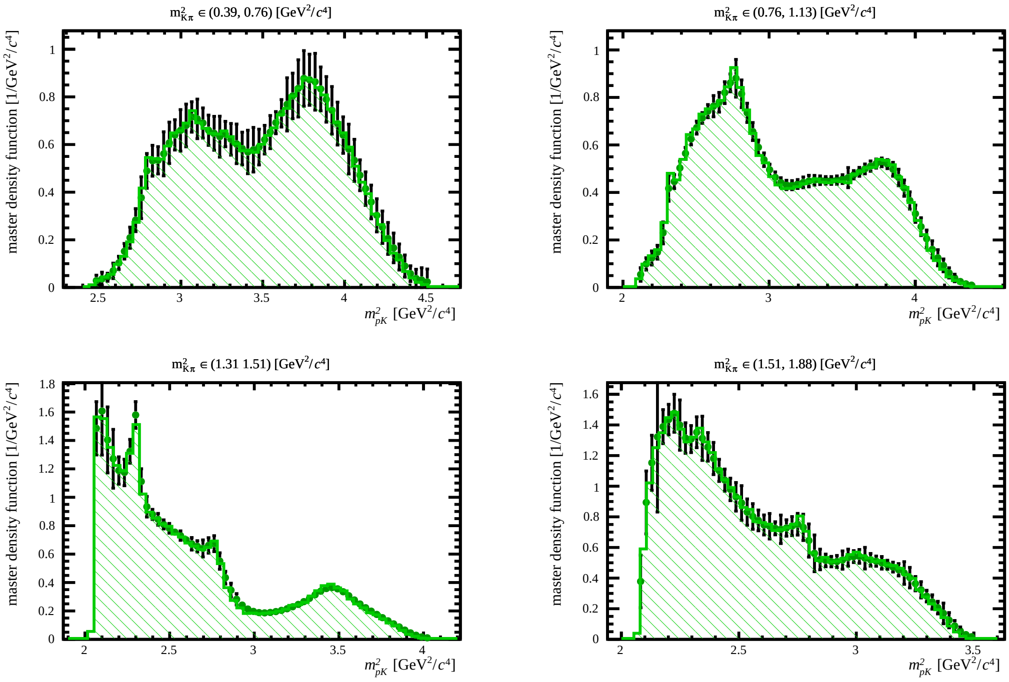

Figure 78: The Master KDE $\hat{f}_{Master}(x)$ density function for antiparticles of control Λ_c^+ sample with 1000000 entries for 4 regions overlaid on the normalized distribution of square of invariant mass $m^2(pK)$.

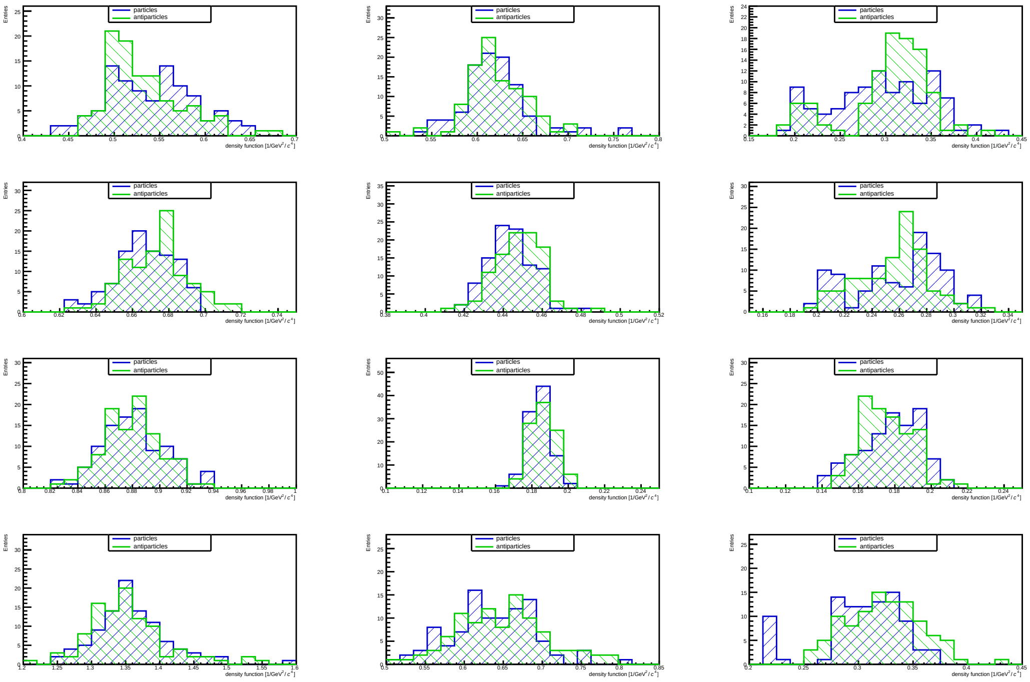

Figure 79: Distributions of values of density functions estimated for the 10th (left side), 30th (middle) and 50th (right side) bins for each region of the Dalitz plot. Plots for region no. 1 are in the 1st row, plots for region no. 2 are in 2nd row, plots for region no. 3 are in 3rd row and plots for region no. 4 are in 4th row. Distributions for particles are drawn with blue lines and distributions for antiparticles are drawn with green lines.

9 Conclusions

The searches for *CP* symmetry violation in decays of charm baryons Ξ_c^+ are developed. These kinds of analysis are a very meaningful test for the Standard Model, because if a discrepancy is found between the measured and expected quantities, it will indicate the existence of physics beyond the Standard Model. Therefore, this type of research is widely recognized as an indirect search for the effects of new physics.

One of the biggest challenges in studying *CP* charge-space symmetry violation in decays of charm baryons is that the expected magnitude of this effect in the Standard Model is of the order of $10^{-4} - 10^{-3}$ or less. The sensitivity of the measurements can be increased by using a larger data sample and by employing new statistical analysis techniques. The presented data and analysed in this thesis were collected by the LHCb detector in 2016, 2017 and 2018 years, commonly known as Run 2 period, while the first search for *CP* asymmetry in Ξ_c^+ decays was performed with data from the so-called Run 1, i.e., collected in 2011 and 2012. Due to the improvement of the LHCb detector between these data collection cycles, the abundance of data produced in Run 2 is several times larger than that obtained in Run 1. Only recently, the LHCb detector underwent another upgrade process and another data collection cycle (Run 3) was started. It is expected to generate data samples with even higher statistics than those obtained in previous years.

In terms of the statistical tests needed to measure *CP* asymmetry, three model-independent methods were used in this analysis: the binned S_{CP} and two ubinned methods: the Energy Test and the Kernel Density Estimation. The first technique (S_{CP}) is a well-known method that was also used in the first search for *CP* violation in Ξ_c^+ decays. The Eenergy Test method has not been used in analyses of baryon decays before, and the Kernel Density Estimation technique represents a completely new approach in the search for *CP* asymmetries.

It should be pointed out that the analysis is blinded, so all calculations were performed on generated Monte Carlo samples (otherwise known as toy samples), with Run 2 data statistics and using samples of Λ_c^+ control decays. The Λ_c^+ decays were chosen because the *CP* asymmetry in them as measured in the Standard Model is negligible and the final states of the Λ_c^+ and Ξ_c^+ decays are the same.

Toy samples in which *CP* asymmetries of different magnitudes were applied were used to test and compare the sensitivity of the S_{CP} and Energy Test methods. On the other hand, calculations performed on the control sample Λ_c^+ allowed the validity of the methods to be checked. This meant veryfing that, for samples with conserved *CP* symmetry, the result does not indicate a false *CP* violation signal and that, for samples with a non-zero *CP* asymmetry value, it is detected. In addition, control samples were used to measure and remove contaminating asymmetries (production and detector) that could lead to a false *CP* violation signal. The results indicated that higher sensitivity was shown by Energy Test and the absence of *CP* asymmetries in the control samples, which was the expected outcome.

In the case of the Kernel Density Estimation technique, the Monte Carlo samples made it possible to carry out an optimisation of the so-called bandwidth parameter, on which the result of the estimation performed with the Kernel Density Estimation technique is most dependent. As mentioned earlier, the Kernel Density Estimation technique has not been used in analyses such as the one presented in this work. The distributions of mass squares to which the probability density function was estimated were much more complicated than those to which the Kernel Density Estimation technique is customarily applied. Therefore, an optimisation procedure was carried out and the correct probability function was finally determined. This can be utilised in future used as part of the Energy Test, in a form that does not require Monte Carlo approximation. This is also a step towards the application of the Kernel Density Estimation technique for the two-dimensional case, i.e. to the entire Dalitz space. This, in turn, will make it possible to have a comparison between the sensitivity of the Kernel Density Estimation and that of the S_{CP}.

The next step in the search for CP asymmetries in charm decays is to complete the so-called Analysis Note, which is required to obtain permission from the LHCb collaboration to unveil Ξ_c^+ signal decays. There are also still unexplored possibilities of the Kernel Density Estimation technique and its combinations with other methods employed in this analysis. Moreover, as written earlier, the LHCb detector is in the middle of another data collection cycle (Run 3), which may in the near future result in finding, for the first time, CP asymmetries in charm baryons or even in the discovery of new physics beyond the Standard Model.

Appendix A

Λ_c^+ and Ξ_c^+ mass distributions after offline selection and fiducial cuts

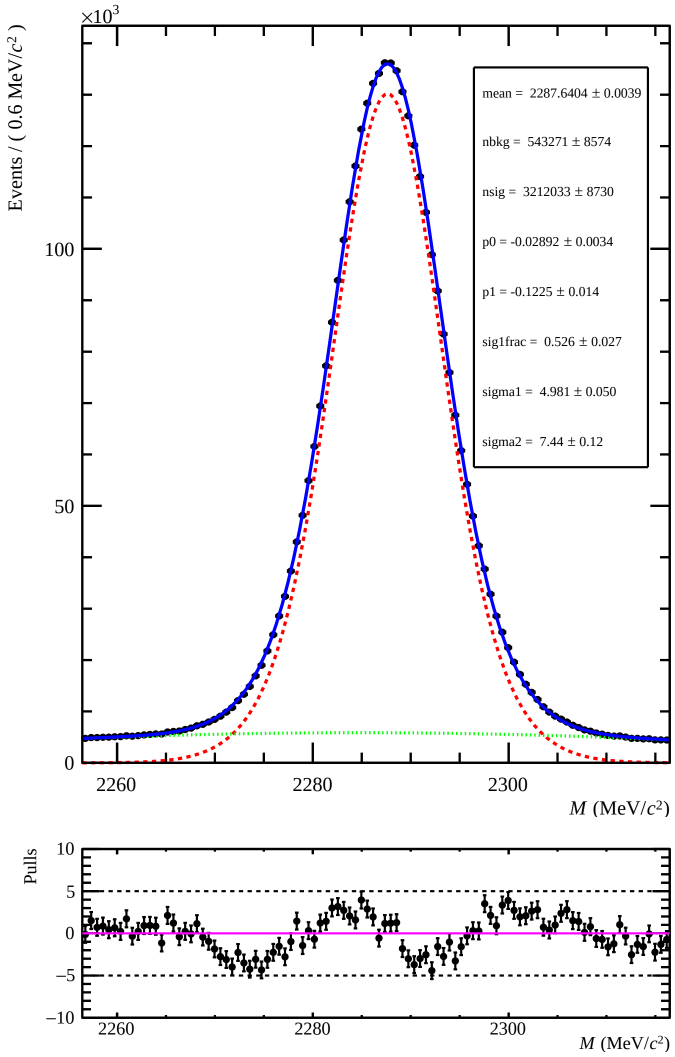

Figure 80: Final reconstructed mass of $pK\pi$ in Λ_c^+ 2016 sample. The fitted mass model is overlaid (blue line). It composed of double Gaussian function (red dashed line), with common mean μ and different widths σ_1 and σ_2 for mass peak and Chebychev polynomial function $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The $nsig$ and $nbkg$ are the numbers of signal and background candidates, respectively. The pull distribution is shown as a difference between data and fitted model.

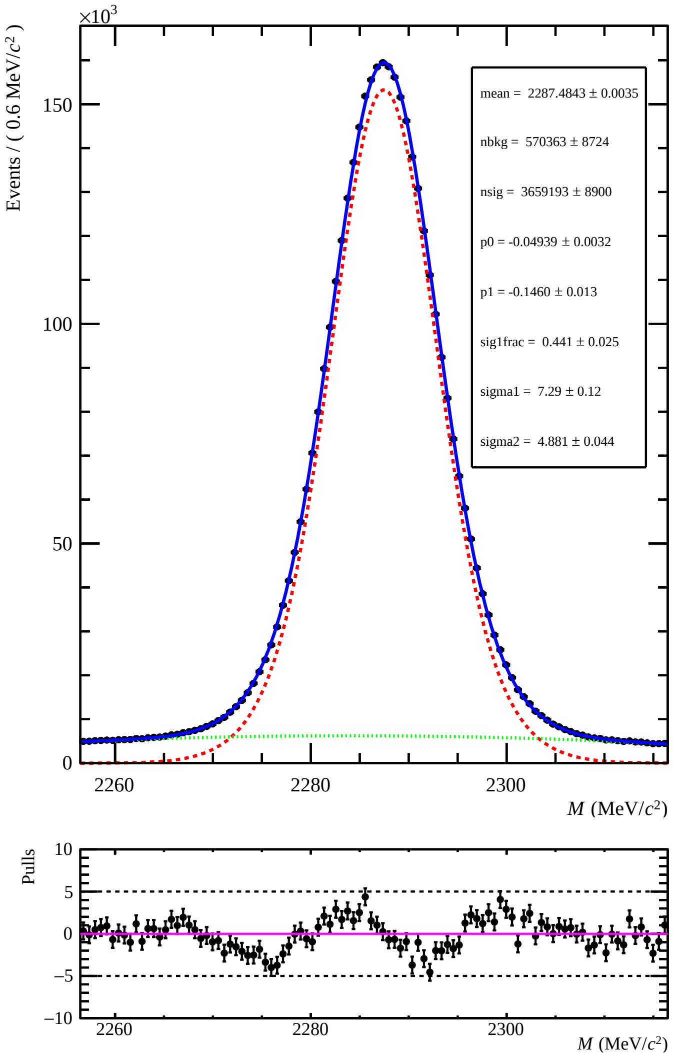

Figure 81: Final reconstructed mass of $pK\pi$ in Λ_c^+ 2017 data set. The fitted mass model is overlaid (blue line). It is composed of double Gaussian function (red dashed line), with common mean μ and different widths σ_1 and σ_2, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The *nsig* and *nbkg* are the number of signal and background events obtained from the fit, respectively. The pull distribution is shown as a difference between data and fitted model.

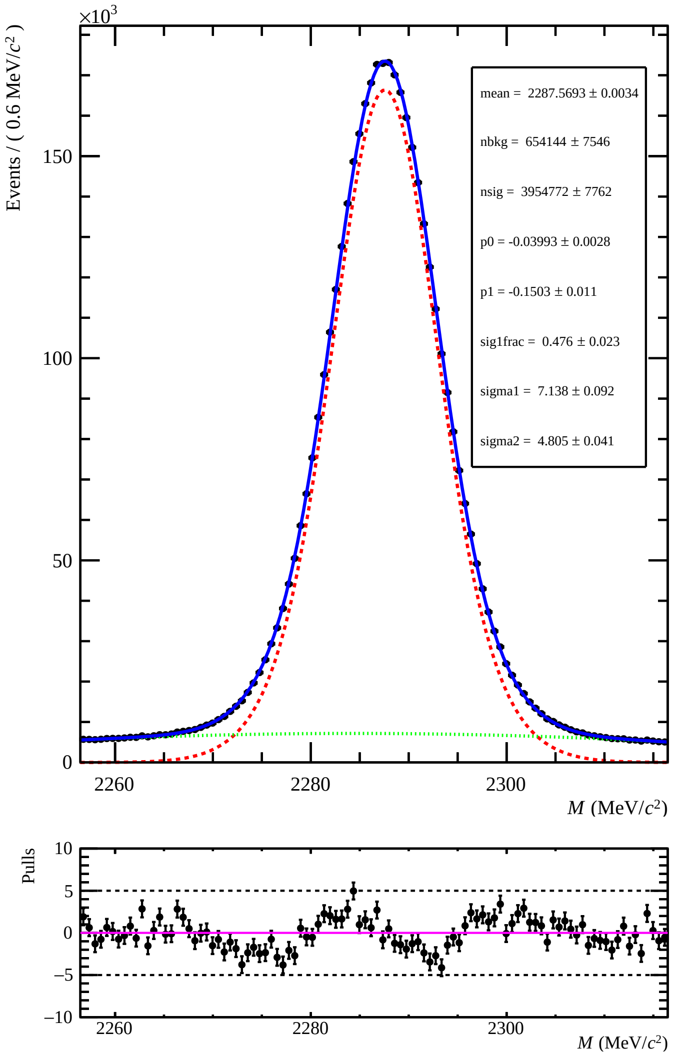

Figure 82: Final reconstructed mass of $pK\pi$ in Λ_c^+ 2018 data set. The fitted mass model is overlaid (blue line). It is composed of double Gaussian function (red dashed line), with common mean μ and different widths σ_1 and σ_2, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The *nsig* and *nbkg* are the number of signal and background events obtained from the fit, respectively. The pull distribution is shown as a difference between data and fitted model.

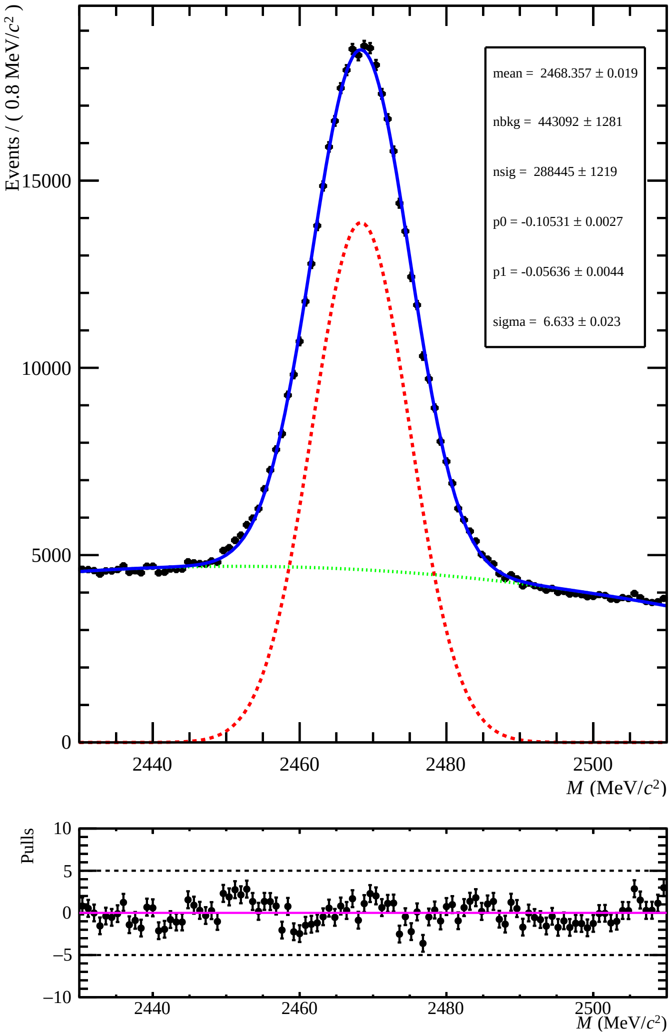

Figure 83: Final reconstructed mass of $pK\pi$ in Ξ_c^+ 2016 data set. The fitted mass model is overlaid (blue line). It is composed of single Gaussian function (red dashed line), with mean μ and width σ, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The *nsig* and *nbkg* are the number of signal and background events obtained from the fit, respectively. The pull distribution is shown as a difference between data and fitted model.

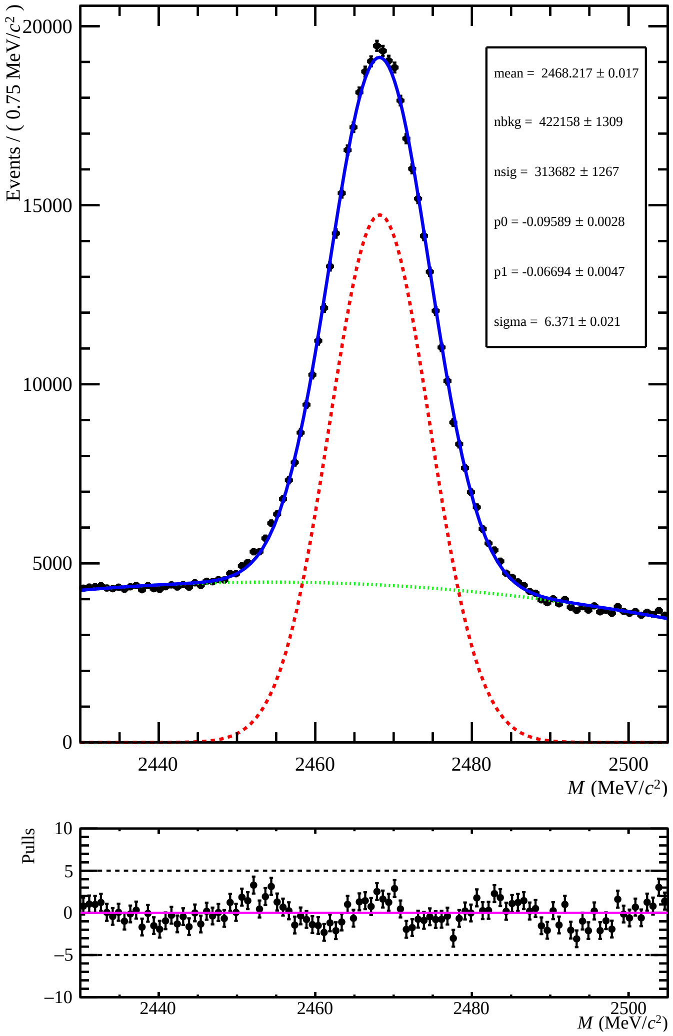

Figure 84: Final reconstructed mass of $pK\pi$ in Ξ_c^+ 2017 data. The fitted mass model is overlaid (blue line). It is composed of single Gaussian function (red dashed line), with mean μ and width σ, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The *nsig* and *nbkg* are the number of signal and background events obtained from the fit, respectively. The pull distribution is shown as a difference between data and fitted model.

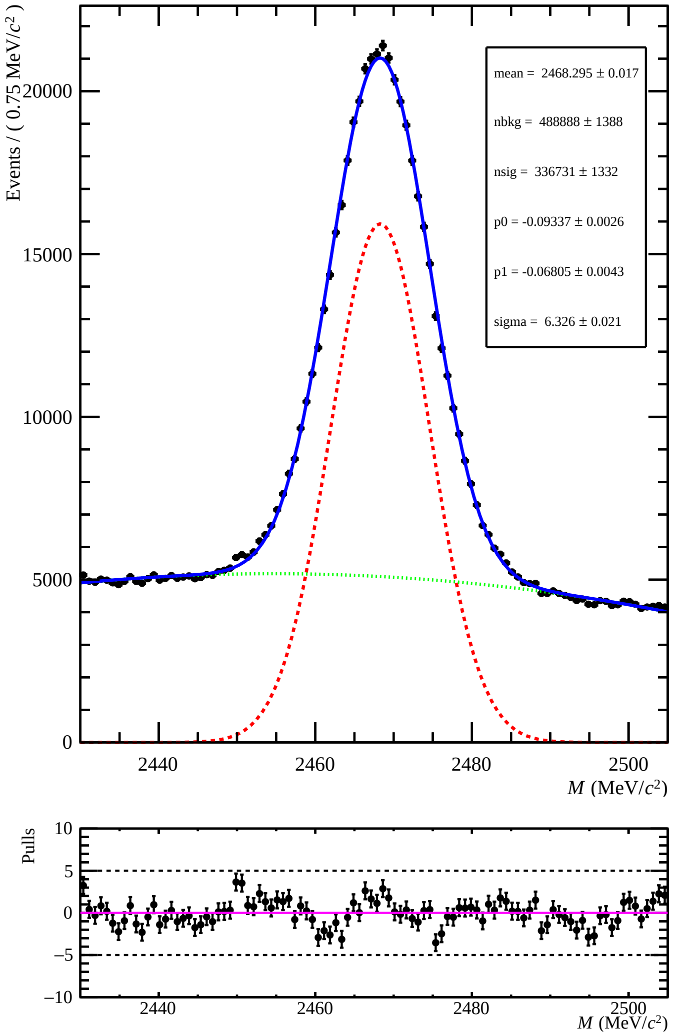

Figure 85: Final reconstructed mass of $pK\pi$ in Ξ_c^+ 2018 data set. The fitted mass model is overlaid (blue line). It is composed of single Gaussian function (red dashed line), with mean μ and width σ, for signal and Chebychev polynomial $g(x) = p_0 + p_1 \cdot x$ (green dotted line) to describe background, where x is the mass of the $pK\pi$. The *nsig* and *nbkg* are the number of signal and background events obtained from the fit, respectively. The pull distribution is shown as a difference between data and fitted model.

References

[1] 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.

[2] KTeV collaboration, A. Alavi-Harati et al., *Observation of Direct CP Violation in* $K_{S,L} \to \pi\pi$ *Decays*, Phys. Rev. Lett. 83 (1999) 22, arXiv:hep-ex/9905060.

[3] NA48 collaboration, A. Lai et al., *A precise measurement of the direct CP violation parameter Re*(ϵ'/ϵ), Eur. Phys. J. C 22 (2001) 231, arXiv:hep-ex/0110019.

[4] BaBar collaboration, B. Aubert et al., *Observation of CP violation in the* B^0 *meson system*, Phys. Rev. Lett. 87 (2001) 091801, arXiv:hep-ex/0107013.

[5] Belle collaboration, K. Abe et al., *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. 87 (2001) 091802, arXiv:hep-ex/0107061.

[6] BaBar collaboration, B. Aubert et al., *Direct CP Asymmetry in* $B^0 \to K^+\pi^-$ *Decays*, Phys. Rev. Lett. 93 (2004) 131801, arXiv:hep-ex/0407057.

[7] Belle collaboration, Y. Chao et al., *Evidence for Direct CP Violation in* $B^0 \to K^+\pi^-$ *Decays*, Phys. Rev. Lett. 93 (2004) 191802, arXiv:hep-ex/0408100.

[8] LHCb collaboration, R. Aaij et al., *First observation of CP violation in the decays of* B_s^0 *mesons*, Phys. Rev. Lett. 110 (2013) 221601, arXiv:1304.6173.

[9] LHCb collaboration, R. Aaij et al., *Observation of CP violation in* $B^\pm \to DK^\pm$ *decays*, Phys. Lett. B 712 (2012) 203, Erratum ibid. B713 (2012) 351, arXiv:1203.3662.

[10] LHCb collaboration, R. Aaij et al., *Observation of CP violation in charm decays*, Phys. Rev. Lett. 122 (2019) 211803, arXiv:1903.08726.

[11] LHCb collaboration, R. Aaij et al., *Measurement of the time-integrated CP asymmetry in* $D^0 \to K^+K^-$ *decays*, LHCb-PAPER-2022-024, arXiv:2209.03179.

[12] I. Bediaga et al., *On a CP anisotropy measurement in the Dalitz plot*, Phys. Rev. D 80 (2009) 096006, arXiv:0905.4233.

[13] B. Aslan and G. Zech, *New test for the multivariate two-sample problem based on the concept of minimum energy*, J. Stat. Comput. Simul. 75 (2005) 109.

[14] B. Aslan and G. Zech, *Statistical energy as a tool for binning-free, multivariate goodness-of-fit tests, two-sample comparison and unfolding*, Nucl. Instrum. Meth. A 537 (2005) 626.

[15] M. Williams, *Observing CP violation in many-body decays*, Phys. Rev. D 84 (2011) 054015, arXiv:1105.5338.

[16] C. Parkes et al., *On model-independent searches for direct CP violation in multi-body decays*, J. Phys. G 44 (2017) 085001, arXiv:1612.04705.

[17] W. Barter, C. Burr, C. Parkes, *Calculating p-values and their significances with the Energy Test for large datasets*, JINST 13 (2018) P04011, arXiv:1801.05222.

[18] T. Szumlak, *Performance of the LHCb Vertex locator and the measurement of the forward-backward asymmetry in* $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ *decay channel as a probe of New Physics*, Krakow, (2013).

[19] S. A. Thomas, F. B. Abdalla, O. Lahav, *Upper Bound of 0.28 eV on Neutrino Masses from the Largest Photometric Redshift Survey*, Phys. Rev. Lett. 105 (2010) 031301, arXiv:0911.5291.

[20] R. A. Battye, A. Moss, *Evidence for massive neutrinos from cosmic microwave background and lensing observations*, Phys. Rev. Lett. 112 (2014) 051303, arXiv:1308.5870.

[21] Atlas collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B716 (2012) 1-29, arXiv:1207.7214.

[22] Cabibbo N., *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. 10 (1963) 531.

[23] Kobayashi M. and Maskawa T., *CP-violation in the re-normalizable theory of weak interaction*, Prog. Theor. Phys. 49 (1973) 652.

[24] Gersabeck M., *Introduction to Charm Physics*, arXiv:1503.00032 (2015).

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

[26] M. Chala, A. Lenz, A. V. Rusov, and J. Scholtz, ΔA_{CP} *within the Standard Model and beyond*, JHEP 07 (2019) 161, arXiv:1903.10490.

[27] F. Buccella et al., *Nonleptonic weak decays of charmed mesons*, Phys. Rev. D 51 (1995) 3478, arXiv:hep-ph/9411286.

[28] S. Bianco, F. L. Fabbri, D. Benson, and I. Bigi, *A Cicerone for the physics of charm*, Riv. Nuovo Cim. 26N7 (2003) 1, arXiv:hep-ex/0309021.

[29] A. A. Petrov, *Hunting for CP violation with untagged charm decays*, Phys. Rev. D 69 (2004) 111901, arXiv:hep-ph/0403030.

[30] A. Khodjamirian and A. A. Petrov, *Direct CP asymmetry in* $D \to \pi^-\pi^+$ *and* $D \to K^-K^+$ *in QCD-based approach*, Phys. Lett. B 774 (2017) 235, arXiv:1706.07780.

[31] H.-Y. Cheng and C.-W. Chiang, *Revisiting CP violation in* $D \to PP$ *and* VP *decays*, Phys. Rev. D 100 (2019) 093002, arXiv:1909.03063.

[32] H.-N. Li, C.-D. Lü, and F.-S. Yu, *Implications on the first observation of charm CPV at LHCb*, arXiv:1903.10638.

[33] S. Müller, U. Nierste, and S. Schacht, *Sum rules of charm CP asymmetries beyond the SU(3)F limit*, Phys. Rev. Lett. 115 (2015) 251802, arXiv:1506.04121.

[34] H.-n. Li, C.-D. Lü, Q. Qin, and F.-S. Yu, *Branching ratios and direct CP asymmetries in $D \to PV$ decays*, Phys. Rev. D 89 (2014) 054006, arXiv:1305.7021.

[35] I. I. Bigi, A. Paul, and S. Recksiegel, *Conclusions from CDF results on CP violation in $D^0 \to \pi^+\pi^-, K^+K^-$ and future tasks*, JHEP 06 (2011) 089, arXiv:1103.5785.

[36] LHCb collaboration, R. Aaij et al., *Observation of CP violation in $B^\pm \to DK^\pm$ decays*, Phys. Lett. B 712 (2012) 203, arXiv:1203.3662.

[37] LHCb collaboration, R. Aaij et al., *First observation of CP violation in the decays of B^0_s mesons*, Phys. Rev. Lett. 110 (2013) 221601, arXiv:1304.6173.

[38] LHCb collaboration, R. Aaij et al., *Direct CP violation in charmless three-body decays of $B^\pm$ mesons*, High Energy Physics - Experiment (hep-ex), FOS: Physical sciences, FOS: Physical sciences (2022) arXiv:2206.07622.

[39] I. Bigi, *Probing CP asymmetries in charm baryons decays*, arXiv:1206.4554.

[40] X.-D. Shi et al., *Prospects for CP and P violation in Λ_c^+ decays at super tau charm facility*, Phys. Rev. D 100 (2019) 113002, arXiv:1904.12415.

[41] A. A. Petrov, *Charm mixing in the Standard Model and beyond*, Int. J. Mod. Phys. A 21 (2006) 5686, arXiv:hep-ph/0611361.

[42] I. Shipsey, *Status of Charm Flavor Physics*, Int. J. Mod. Phys. A 21 (2006) 5381-5403, arXiv:hep-ex/0607070.

[43] M. Artuso, B. Meadows, A. A. Petrov, *Charm Meson Decays*, Ann. Rev. Nucl. Part. Sci. 58 (2008) 249 - 291.

[44] J. H. Christenson, J. W. Cronin, V. L.Fitch, R. Turlay, *Evidence for the 2π Decay of the K^0_2 Meson*, Phys. Rev. Lett. 13 (1964) 138.

[45] BABAR collaboration, B. Aubert et al, *Observation of CP Violation in the B^0 Meson System*, Phys. Rev. Lett. 87 (2001) 091801.

[46] Belle collaboration, K. Abe et al., *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. 87 (2001) 091802.

[47] E. Solomonidi, *CP violation in D decays to two pseudoscalars: A SM-based calculation*, BEACH 2022 Conference, 5.06 - 11.06, Krakow.

[48] LHCb collaboration, R. Aaij et al., *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, JHEP 07 (2014) 041, arXiv:1405.2797.

[49] LHCb collaboration, R. Aaij et al., *Measurement of CP asymmetry in $D^0 \to K^-K^+$ decays*, Phys. Lett. B767 (2017) 177, arXiv:1610.09476.

[50] LHCb collaboration, R. Aaij et al., *Measurement of the difference of time-integrated CP asymmetries in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, Phys. Rev. Lett. 116 (2016) 191601, arXiv:1602.03160.

[51] LHCb collaboration, R. Aaij et al., *Search for time-dependent CP violation in* $D^0 \to K^- K^+$ *and* $D^0 \to \pi^- \pi^+$ *decays*, Phys. Rev. D 104 (2021) 072010, arXiv:2105.09889.

[52] LHCb collaboration, R. Aaij et al., *A measurement of the CP asymmetry difference in* $\Lambda_c^+ \to pK^- K^+$ *and* $\Lambda_c^+ \to p\pi^- \pi^+$ *decays*, JHEP 03 (2018) 182, arXiv:1712.07051.

[53] LHCb collaboration, R. Aaij, et al., *Measurement of matter-antimatter differences in beauty baryon decays.* Nat. Phys. 13 (2017) 391, arXiv:1609.05216.

[54] LHCb collaboration, R. Aaij, et al., *Search for CP violation and observation of P violation in* $\Lambda_b^0 \to p\pi^- \pi^+ \pi^-$ *decays*, Phys. Rev. D 102 (2020) 051101, arXiv:1912.10741

[55] LHCb collaboration, R. Aaij, et al., *Search for CP violation in* $\Xi_c \to pK^- \pi^+$ *decays with model-independent techniques*, Eur. Phys. J. C 80 (2020) 986, arXiv:2006.03145.

[56] LHCb collaboration, R. Aaij et al., *The LHCb detector at the LHC*, JINST 3 (2008) S08005.

[57] LHCb collaboration, R. Aaij et al., *LHCb Detector Performance*, Int. J. Mod. Phys. A 30 (2015) 1530022, arXiv:1412.6352.

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

[59] 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 (2019) P04013, arXiv:1812.10790.

[60] LHCb collaboration, R. Aaij et al., *Measurement of* $\sigma(pp \to b\bar{b}X)$ *at* $\sqrt{s} = 7\ TeV$ *in the forward region*, Phys. Lett. B 694 (2010) 209, arXiv:1009.2731.

[61] LHCb collaboration, R. Aaij et al., *Prompt charm production in pp collisions at* $\sqrt{s} = 7\ TeV$, Nuclear Physics, Section B 871 (2013) 1-20, arXiv:1302.2864.

[62] LHCb collaboration, R. Aaij et al., *Erratum to: Measurements of prompt charm production cross-sections in pp collisions at* $\sqrt{s} = 13\ TeV$, Energ. Phys. 74 (2017).

[63] LHCb collaboration, R. Aaij et al., *LHCb VELO Upgrade Technical Design Report*, Tech. rep. CERNLHCC-2013-021. LHCB-TDR-013. 2013. URL: https://cds.cern.ch/record/1624070.

[64] LHCb collaboration, R. Aaij et al., *LHCb PID Upgrade Technical Design Report*, Tech. rep. CERNLHCC- 2013-022. LHCB-TDR-014. 2013. URL: https://cds.cern.ch/record/1624074.

[65] W. Verkerke, D. Kirkby, *RooFit Users Manual v2.91*, (2008).

[66] M. Pivk and F. R. Le Diberder, *sPlot: A Statistical tool to unfold data distributions*, Nucl. Instrum. Meth. A 555 (2005) 356–369.

[67] P. Kulczycki, *Estymatory jadrowe w analizie systemowej*, WNT, ISBN: 83-204-3147-6 (2005).

[68] LHCb collaboration, R. Aaij et al., *Search for CP violation in* $D^0 \to \pi^-\pi^+\pi^0$ *decays with the energy test*, Phys. Lett. B 740 (2015) 158, arXiv:1410.4170.

[69] LHCb collaboration, R. Aaij et al., *Search for CP violation in the phase space of* $D^0 \to \pi^+\pi^-\pi^+\pi^-$ *decays*, Phys. Lett. B 769 (2017) 345, arXiv:1612.03207.

[70] LHCb collaboration, R. Aaij et al., *Search for CP violation and observation of P violation in* $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ *decays* Phys. Rev. D 102 (2020) 051101(R), arXiv:1912.10741.

[71] J. Blatt and V. E. Weisskopf, *Theoretical Nuclear Physics* (J. Wiley, New York, 1952).

[72] C. Zemach, *Three-Pion Decays of Unstable Particles*, Phys. Rev. 133 (1964) B1201.

[73] C. Zemach, *Use of Angular-Momentum Tensors*, Phys. Rev. 140 (1965) B97.

[74] M. Williams, *How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics*, JINST 5 (2010) P09004.