Alma Mater Studiorum · University of Bologna

DOTTORATO DI RICERCA IN FISICA
Ciclo XXXVI

Settore concorsuale: 02/A1 - Fisica sperimentale delle interazioni fondamentali
Settore scientifico disciplinare: FIS/01 - Fisica sperimentale

Measurement of integrated CP asymmetries in $\Lambda_b^0 \to ph^-$ decays with the full Run 1+2 dataset from LHCb

Presentata da: Andrea Villa

Coordinatore dottorato:
Prof. Alessandro Gabrielli

Supervisore:
Prof. Angelo Carbone

Co-supervisori:
Dr. Stefano Perazzini
Dr. Fabio Ferrari

Esame finale anno 2024

Abstract

In this thesis, a search for *CP* violation in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays using the full Run 1+2 dataset gathered by the LHCb experiment is presented. The sample corresponds to an integrated luminosity of $9\,\text{fb}^{-1}$ of proton-proton collisions at centre-of-mass energies of 7, 8 and 13 TeV, collected between 2011 and 2018. Several corrections accounting for particle identification, trigger, Λ_b^0 production, as well as detection asymmetries of final-state particles are needed to extract the physical *CP* asymmetries and are determined in this thesis. The *CP* asymmetries corresponding to the Run 1 dataset are measured to be

$$\begin{aligned} A_{CP}(pK^-) &= (0.09 \pm 1.53 \pm 0.72)\%, \\ A_{CP}(p\pi^-) &= (-0.52 \pm 1.89 \pm 0.56)\%, \end{aligned}$$

where the first uncertainty is statistical and the second systematic. The results are compatible with the ones from a previous iteration of the analysis and supersede them thanks to a significant reduction in the systematic uncertainties. The *CP* asymmetries corresponding to the Run 2 dataset are measured to be

$$\begin{aligned} A_{CP}(pK^-) &= (-1.45 \pm 0.75 \pm 0.43)\%, \\ A_{CP}(p\pi^-) &= (0.34 \pm 0.95 \pm 0.43)\%. \end{aligned}$$

When combining the values from Run 1 and Run 2 the following values are found

$$\begin{aligned} A_{CP}(pK^-) &= (-1.12 \pm 0.67 \pm 0.36)\%, \\ A_{CP}(p\pi^-) &= (0.15 \pm 0.85 \pm 0.36)\%, \end{aligned}$$

which are the world-best measurements of these asymmetries from a single experiment to date, and are compatible with the *CP* symmetry conservation hypothesis within 1.5σ and 0.2σ, respectively.

Contents

Introduction

The violation of the *CP* symmetry (CPV) is described in the Standard Model (SM) through the presence of a complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix, that describes the transitions between up- and down-type quarks [1, 2]. However, the amount of CPV currently observed in the SM is not sufficient to justify the observed baryon asymmetry of the Universe, which can be explained if three conditions are met [3]:

B **non-conservation** obviously, the baryon number must not be a conserved quantity for an asymmetry between matter and antimatter to appear; currently, the conservation of B is a weak symmetry in the SM, justified simply by the lack of evidence of its violation;

C **and** *CP* **violation** the need for C symmetry violation is clear, as any process $X \to Y + Z$ would otherwise be counterbalanced by the C-conjugate $\overline{X} \to \overline{Y} + \overline{Z}$; but *CP* violation is also needed: if we consider as an example the decay of a particle X to a pair of left-handed (right-handed) quarks $q_L q_L$ ($q_R q_R$), then *CP* conservation would imply $\Gamma(X \to q_L q_L) + \Gamma(X \to q_R q_R) = \Gamma(\overline{X} \to \overline{q}_L \overline{q}_L) + \Gamma(\overline{X} \to \overline{q}_R \overline{q}_R)$, *i.e.* no baryon asymmetry would be created;

Loss of thermal equilibrium this is needed because if the Universe was at thermal equilibrium any process that could introduce an imbalance between baryons and antibaryons would be countered by the opposite process, thus restoring equilibrium;

The first observation of CPV dates back to the 1960s in neutral kaon decays [4], and much later was also observed in the B^0 [5,6], B_s^0 [7], and B^+ systems [8]. In 2019 the LHCb collaboration reported also the first observation of CPV in charm decays [9]. To date, however, CPV is still unobserved in the baryon sector. In 2016 the LHCb collaboration reported an evidence of CPV in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays [10] with a significance of 3.3σ, but the updated measurement did not confirm the evidence [11].

The decays of the Λ_b^0 (bud) baryon to two-body charmless final states, pK^- or $p\pi^-$, are a promising place to search for CPV that may arise from the interference between tree and penguin topologies that contribute to the decay amplitudes, given a nonzero strong and weak relative phases. Examples of Feynman diagrams contributing to the $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays are shown in Fig. 1. Theoretical computations of branching fractions and *CP* asymmetries for these two decays have been performed only in recent times, and they tend to predict values for A_{CP} of the order of a couple of percents for both channels, although with large uncertainties in some cases. Experimentally, the

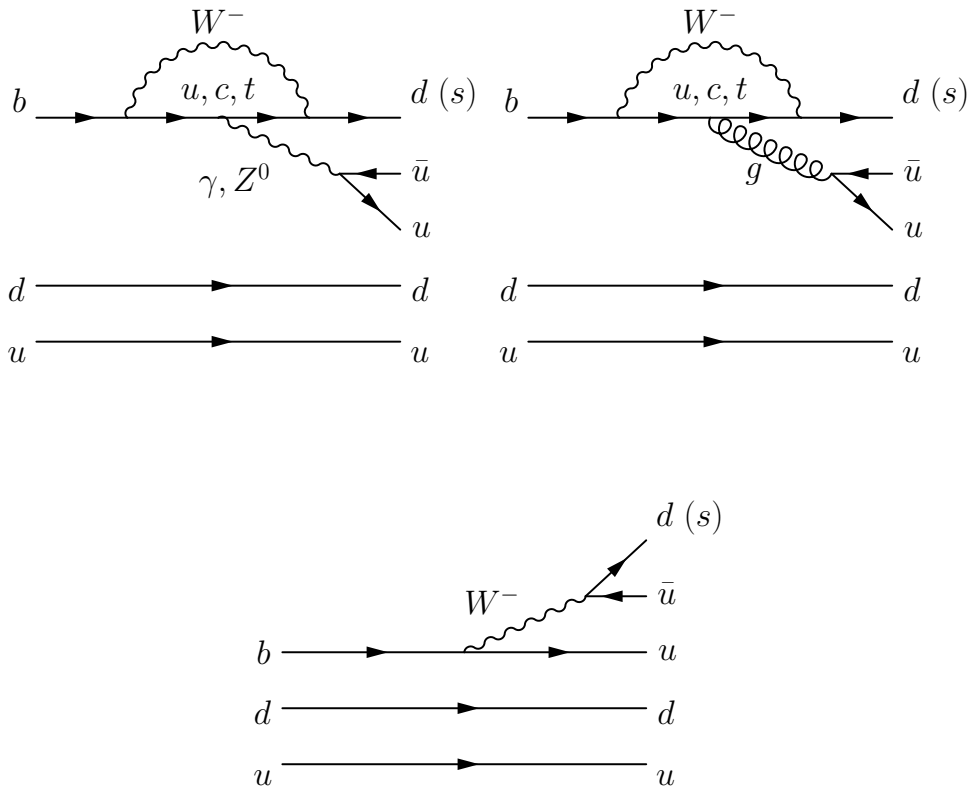

Figure 1: Examples of Feynman diagrams representing (top left) the penguin EW topology, (top right) the penguin QCD topology and (bottom) the tree-level topology contributing to $\Lambda_b^0 \to pK^-$ ($b \to su\bar{u}$ transition) and $\Lambda_b^0 \to p\pi^-$ decays ($b \to du\bar{u}$ transition).

Table 1: Current knowledge of direct *CP* asymmetries in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays. The measurements are performed by the CDF [12] and LHCb [13] collaborations. The last column reports the world averages computed by the Particle Data Group [14].

	CDF	LHCb	PDG average
$A_{CP}(pK^-)$	$(-10 \pm 8 \pm 4)\%$	$(-2.0 \pm 1.3 \pm 1.9)\%$	$(-2.5 \pm 2.5)\%$
$A_{CP}(p\pi^-)$	$(\ \ 6 \pm 7 \pm 3)\%$	$(-3.5 \pm 1.7 \pm 2.0)\%$	$(-2.5 \pm 2.2)\%$

CP asymmetries have been measured by the CDF [12] and LHCb [13] collaborations, as summarised in Table 1; the latter result dominates the world average, which is compatible with the no-CPV hypothesis. With the addition of Run2 statistics and new developments in the determination of nuisance asymmetries, a significant reduction of the uncertainties can be achieved, opening the possibility to observe CPV. Even in the case CPV would not be observed, more precise determinations of the *CP* asymmetries in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays remain interesting as an input to the theoretical models used to deal with the effect of residual strong interaction in hadronic Λ_b^0 decays, that represent the main source of theoretical uncertainties.

In this document, a measurement of $A_{CP}(pK^-)$ and $A_{CP}(p\pi^-)$ using the full Run

1+2 dataset collected by the LHCb experiment will be presented, with a sample size corresponding to $9\,\text{fb}^{-1}$ of pp collisions at centre-of-mass energies of 7, 8 and 13 TeV. The document is structured in three chapters: in Chapter 1, a theoretical description of CP violation in the SM will be presented, together with a focus on two-body charmless decays of b-hadrons that constitute the topic of the analysis; in Chapter 2, the LHCb detector will be described in detail, explaining the various components that allow it to reconstruct the decay products of b- and c-hadrons, as well as the dedicated variables and trigger components that will be used in the analysis; finally, in Chapter 3 the analysis will be discussed, starting from the strategy and going through the data selection, the calibration of the particle identification efficiencies, the definition of the invariant-mass models used to extract the signal yields with maximum-likelihood fits, the computation of all the nuisance experimental asymmetries, and the determination of the final result.

Chapter 1

Theory of *CP* violation in the Standard Model

The Standard Model (SM) of particle physics describes three of the four fundamental interactions with the exception of gravity, *i.e.* the electromagnetic, weak, and strong interaction. The SM is formulated as a quantum field theory describing the dynamics of all particles, both elementary and composite.

There are 25 fundamental particles in the SM: 6 quarks (u, d, c, s, b, t), 6 leptons (e, ν_e, μ, ν_μ, τ, ν_τ), 8 massless bosons of the strong interaction (g), 3 massive bosons of the weak interaction (W^+, W^-, Z^0), 1 massless boson of the electromagnetic interaction (γ), and 1 massive Higgs boson (H), responsible for the masses of all particles interacting with its field. Fermions, which include quarks and leptons, are grouped into multiplets based on the symmetries of the underlying Yang–Mills theories, as shown in Table 1.1. Each interaction if based on a symmetry group: SU(3) for the strong inteaction, SU(2) for the weak interaction, and U(1) for the electromagnetic interaction.

The electric charge of each particle can be obtained from the Gell-Mann–Nishijima [15, 16] formula $Q = T_3 + Y/2$, where T_3 is the third component of the weak isospin. The electric charges come out to be $-|q_e|$ for leptons, 0 for neutrinos, $+2/3|q_e|$ for up-quarks and $-1/3|q_e|$ for down-quarks.

1.1 The GWS model of electroweak interactions

In the SM, the electromagnetic and weak interaction are unified into the electroweak interaction by the Glashow–Weinberg–Salam (GWS) model [17–19], which describes the dynamics and the interactions with other particles of the 4 electroweak gauge bosons. The model can be showcased by writing the following Lagrangian [20, 21]:

$$\mathcal{L} = \mathcal{L}_{\mathrm{K}} + \mathcal{L}_{\mathrm{g}} + \mathcal{L}_{\mathrm{H}} + \mathcal{L}_{\mathrm{Y}}, \tag{1.1}$$

where $\mathcal{L}_{\mathrm{K}}$ is the kinetic term, $\mathcal{L}_{\mathrm{g}}$ the gauge term, $\mathcal{L}_{\mathrm{H}}$ the Higgs term, and $\mathcal{L}_{\mathrm{Y}}$ the Yukawa term. The kinetic term starts in the form of a free propagator of massless particles:

$$\mathcal{L}_{\mathrm{K}} = i\overline{\psi}\partial^\mu\gamma_\mu\psi, \tag{1.2}$$

where γ_μ are the four Dirac matrices, $\overline{\psi} = \psi^\dagger\gamma^0$, and the spinor fields ψ, of which there are three generations, consist of the five representations shown in Table 1.1. Equation (1.2)

Table 1.1: Arrangement of the SM fundamental particles in multiplets. The superscript i spans over the three generations of particles.

Multiplet	Symbol	SU(3) charge	SU(2) charge	Hypercharge
Left-handed fermions	ℓ_L^i	1	2	-1
Right-handed fermions	ℓ_R^i	1	1	-2
Left-handed quarks	q_L^i	3	2	1/3
Right-handed up-quarks	u_R^i	3	1	4/3
Right-handed down-quarks	d_R^i	3	1	-2/3

can be made gauge-invariant by substituting the partial derivative with the covariant derivative, defined as

$$D^\mu = \partial^\mu + ig_s G_a^\mu \lambda_a + igW_b^\mu \sigma_b + ig' B^\mu Y, \tag{1.3}$$

where λ_a are the 8 Gell-Mann matrices, σ_b the 3 Pauli matrices, G_a^μ, W_b^μ, and B^μ are respectively the fields for the gluons, the weak bosons, and the hypercharge boson, while g_s, g, and g' are three coupling constants. With this modification, the kinetic term also acquires interaction terms with the gauge bosons; for example, we can write explicitly the coupling to the left-handed quarks:

$$\begin{aligned}\mathcal{L}_{\mathrm{K}}(q_L) &= i\overline{q_L} D^\mu \gamma_\mu q_L \\ &= i\overline{q_L}(\partial^\mu + ig_s G_a^\mu \lambda_a + igW_b^\mu \sigma_b + ig' B^\mu Y)\gamma_\mu q_L.\end{aligned} \tag{1.4}$$

The gauge term encodes the interaction amogn the gauge bosons:

$$\mathcal{L}_{\mathrm{g}} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} W^b_{\mu\nu} W^{b\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{1.5}$$

where

$$G^{a\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu + g_s f^{abc} G_a^\mu G_c^\nu, \tag{1.6}$$

$$W^{b\mu\nu} = \partial^\mu W_b^\nu - \partial^\nu W_b^\mu + g\varepsilon^{bcd} W_c^\mu G_d^\nu, \tag{1.7}$$

$$B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu, \tag{1.8}$$

are the strentgh tensors of the three gauge fields, and f^{abc} and ε^{bcd} are respectively the structure constants of SU(3) and SU(2).

The third piece is the Higgs field term, which is responsible for the mass of all particles; it is defined as

$$\begin{aligned}\mathcal{L}_{\mathrm{H}} &= (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \\ &= (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2,\end{aligned} \tag{1.9}$$

with $\phi = (\phi^+, \phi^0)$ being a scalar isospin doublet. The potential $V(\phi)$ has an infinite number of minimum states satisfying $\phi^\dagger \phi = \mu^2/\lambda = v^2/2$, with $v \approx 246\,\mathrm{GeV}$ the vacuum expectation value of ϕ. Thanks to this, the number of degrees of freedom of the scalar field can be reduced to one by expanding it around one particular minimum state:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \tag{1.10}$$

A massive scalar field H, called the Higgs boson, appears as a result of the spontaneous breaking of the SU(2) symmetry due to the particular choice of the vacuum state. This process takes the name of the Higgs mechanism, although it was simultaneous developed also by Englert, Brout, Guralnik, Hagen, and Kibble [22–24]. With this new addition, the first term in Equation (1.9) can be modified to show an interaction between the Higgs field and four new gauge bosons, obtained as the following combinations of the initially massless W_b^μ and B^μ fields:

$$W_\mu^\pm = \frac{W_1^\mu \mp i W_2^\mu}{\sqrt{2}}, \tag{1.11}$$

$$\begin{pmatrix} Z^\mu \\ A^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_3^\mu \\ B^\mu \end{pmatrix}, \tag{1.12}$$

where $\sin\theta_W = g/\sqrt{g^2+g'^2}$. These four bosons are respectively the $W^\pm$, the Z^0, and the photon γ, and they acquire masses of $vg/2$, $v\sqrt{g^2+g'^2}/2$, and 0, respectively.

Finally, the last piece of the SM Lagrangian is the Yukawa term, which contains the interaction between the fermions and the Higgs field:

$$\mathcal{L}_\mathrm{Y} = -(Y_{ij}^d \overline{q_{Li}} \phi d_{Rj} + Y_{ij}^u \overline{q_{Li}} \tilde{\phi} u_{Rj} + Y_{ij}^\ell \overline{\ell_{Li}} \phi \ell_{Rj} + \mathrm{h.c.}), \tag{1.13}$$

where $\tilde{\phi} = i\sigma_2 \phi^*$, "h.c." stands for *hermitian conjugate*, and Y_{ij}^d, Y_{ij}^u, and Y_{ij}^ℓ are the 3×3 complex Yukawa matrices: After the breaking of the electroweak symmetry and the substitution of ϕ from Equation (1.10),the Yukawa term becomes (we focus just on the quarks from now on):

$$\begin{aligned} \mathcal{L}_\mathrm{Y} &= -(Y_{ij}^d \overline{q_{Li}} \phi d_{Rj} + Y_{ij}^u \overline{q_{Li}} \tilde{\phi} u_{Rj} + \mathrm{h.c.}) \\ &= -(\frac{v}{\sqrt{2}} Y_{ij}^d \overline{d_{Li}} d'_{Rj} + \frac{v}{\sqrt{2}} Y_{ij}^u \overline{u_{Li}} u_{Rj} + \mathrm{h.c.} + \text{interaction terms}), \end{aligned} \tag{1.14}$$

where mass terms for the fermions have appeared; the physical states can then be made visible by diagonalizing the Yukawa matrices with four unitary matrices:

$$M^d = \frac{v}{\sqrt{2}} V_L^d Y^d V_R^{d\dagger}, \qquad M^u = \frac{v}{\sqrt{2}} V_L^u Y^u V_R^{u\dagger}. \tag{1.15}$$

If we now write explicitly the two isospin components of the left-handed quarks, *i.e.* $q_L = (u_L, d_L)$, Equation (1.14) becomes (exploiting the unitarity condition $V_{L,R}^{u,d\dagger} V_{L,R}^{u,d} = \mathbb{1}$):

$$\begin{aligned} \mathcal{L}_\mathrm{Y} &= -(\overline{d_{Li}} V_L^{d\dagger} V_L^d M_{ij}^d V_R^{d\dagger} V_R^d d_{Rj} + \overline{u_{Li}} V_L^{u\dagger} V_L^u M_{ij}^u V_R^{u\dagger} V_R^u u_{Rj} + \mathrm{h.c.} + \dots) \\ &= -(\overline{d'_{Li}} M_{ij}^d d'_{Rj} + \overline{u'_{Li}} M_{ij}^u u'_{Rj} + \mathrm{h.c.} + \dots), \end{aligned} \tag{1.16}$$

where we defined the quark mass eigenstates

$$d'_{Li} = V_{Lij}^d d_{Lj}, \qquad d_{Ri} = V_{Rij}^d d_{Rj} \tag{1.17}$$

$$u'_{Li} = V_{Lij}^u u_{Lj}, \qquad u_{Ri} = V_{Rij}^u u_{Rj}. \tag{1.18}$$

If we express the SM Lagrangian in terms of the quark mass eigentstates instead of the weak interaction eigenstates, a fundamental change happens in the interaction mediated

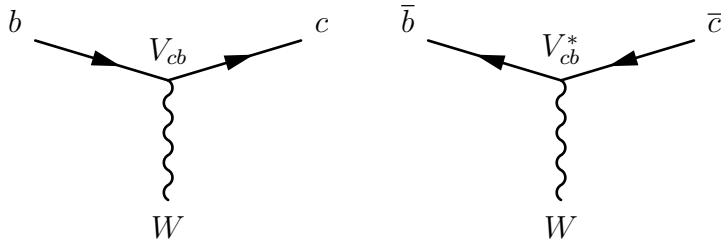

Figure 1.1: Feynman diagram describing the role of the CKM matrix entries in a charged-current quark transition.

by the $W^{\pm}$ boson (also called *charged-current interaction*):

$$\begin{aligned}\mathcal{L}_{\mathrm{K}}^{cc} &= \frac{g}{\sqrt{2}}\overline{u_{Li}}\gamma^{\mu}W_{\mu}^{-}d_{Li} + \frac{g}{\sqrt{2}}\overline{d_{Li}}\gamma^{\mu}W_{\mu}^{+}u_{Li} + \dots \\ &= \frac{g}{\sqrt{2}}\overline{u'_{Li}}(V_L^u V_L^{d\dagger})\gamma^{\mu}W_{\mu}^{-}d'_{Li} + \frac{g}{\sqrt{2}}\overline{d_L i}(V_L^d V_L^{u\dagger})\gamma^{\mu}W_{\mu}^{+}u'_{Li} + \dots,\end{aligned} \tag{1.19}$$

i.e. the possibility to mix between quark families opens up. The probability of the transition is encoded in the so called 3×3 Cabibbo–Kobayashi–Maskawa (CKM) quark-mixing matrix [1, 2]

$$V_{\mathrm{CKM}} = V_L^u V_L^{d\dagger}. \tag{1.20}$$

It is chosen by convention to set the interaction and mass eigenstates equal for up-type quark, while for down-type quarks these are related by:

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \tag{1.21}$$

This notation makes clear that the elements of the CKM matrix (1.21) act as proportionality constants of the amplitude of a charged-current transition between a down-type quark and an up-type quark, while the amplitudes related to antiquarks are proportional to the entries of V^*_{CKM}. This is shown in Figure 1.1.

1.2 The CKM matrix

The magnitudes of the elements of the CKM matrix V_{ij} correspond to the probabilities of a charged-current transition from a quark i to a quark j. The elements can be parametrised with a certain number of free parameters. Any $N \times N$ complex unitary matrix can be completely described with $\frac{1}{2}N(N-1)$ angles and $\frac{1}{2}(N-1)(N-2)$ phases. For $N = 3$, this means 3 Euler angles θ_1, θ_2, θ_3, and a single phase δ. Using the notation $s_{ij} = \sin\theta_{ij}$,

$c_{ij} = \cos\theta_{ij}$, a possible parameterization of the CKM matrix is:

$$V_{\text{CKM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \quad (1.22)$$

The magnitude of the elements of the CKM matrix can be measured by studying SM processes involving the corresponding quarks [25] The latest results can be summarised as follows:

$|V_{ud}|$ can be obtained from meaurements of the neutron lifetime and branching fractions of pion decays; the currently most precise results come from superallowed $0^+ \to 0^+$ nuclear beta-decays [26] and yield

$$|V_{ud}| = 0.97373 \pm 0.00031;$$

$|V_{us}|$ can be obtained in several ways: by studying semileptonic K-decays such as $K^0_{\text{L}} \to \pi e\nu$, $K^0_{\text{L}} \to \pi\mu\nu$, $K^\pm \to \pi^0 e^\pm \nu$, and $K^\pm \to \pi^0\mu^\pm\nu$, which yield the product of V_{us} and the kaon form factors at $q^2 = 0$ [27]; and by computing the kaon and pion decay constants to get $|V_{us}/V_{ud}|$ [28]; the current best estimate comes from averaging these two results:

$$|V_{us}| = 0.2243 \pm 0.0008;$$

$|V_{ub}|$ can be obtained in three ways: with a study of inclusive $\overline{B} \to X_u\ell\overline{\nu}_\ell$ decays; by measuring such decays exclusively, the most promising being $\overline{B} \to \pi\ell\overline{\nu}_\ell$; and by combining measurements of different branching fractions to get $|V_{ub}|/|V_{cb}|$; the current experimental average combines the exclusive and inclusive results, while using the ratio information as a cross-check [29], resulting in

$$|V_{ub}| = (3.82 \pm 0.20) \times 10^{-3};$$

$|V_{cd}|$ was originally measured by analysing charm production in neutrino and antineutrino scattering on nuclei [30]; nowadays, more direct results can be obtain from semileptonic and leptonic charm decays, *e.g.* $D \to (K, \pi)\ell\nu$ and $D^+ \to (\mu^+, \tau^+)\nu$; all the three methods can be combined to obtain the most precise estimate [29, 31]

$$|V_{cd}| = 0.221 \pm 0.004;$$

$|V_{cs}|$ governs the dominant decay mode of the charm quark, and its magnitude can be measured with semileptonic D decays or leptonic D_s decays using external inputs for the form factors and decay constants; both these methods are averaged for maximum precision [29]:

$$|V_{cs}| = 0.975 \pm 0.006;$$

$|V_{cb}|$ can be obtained, much like $|V_{ub}|$, by measuring exclusive and inclusive decays of B-mesons to open charm final states and combining the results to get the value [32]

$$|V_{cb}| = (40.8 \pm 1.4) \times 10^{-3};$$

$|V_{td}|$, $|V_{ts}|$ cannot be precisely measured at tree-level due to the extremely low chance of the top quark to decay to anything other than a b quark. therefore their values must be inferred with indirect measurements; some examples of processes allowing access to V_{td} and V_{ts} are B–$\overline{B}$ oscillations [33], $B \to X_s\gamma$ decays [34, 35], and the ratios of branching fractions of $B \to \rho\gamma$ and $B \to K^*\gamma$ decays; the current best estimates are

$$|V_{td}| = (8.6 \pm 0.2) \times 10^{-3}, \quad |V_{ts}| = (41.5 \pm 0.9) \times 10^{-3};$$

$|V_{tb}|$ limits where set at the Tevatron computing the ratios $\mathcal{B}(t \to Wb)/\mathcal{B}(t \to Wq)$ [36, 37], while it is now measured directly at the LHC by studying single top quark production [38]; the combination of the two methods is used as the world average:

$$|V_{tb}| = 1.014 \pm 0.029.$$

A global picture of the CKM matrix can be built by performing a global fit of all available measurements, and including the unitarity condition as a constraint. The final result is the following [32]

$$|V_{\text{CKM}}| = \begin{pmatrix} 0.97435(16) & 0.22500(67) & 0.00369(11) \\ 0.22486(67) & 0.97349(16) & 0.04182^{(+85)}_{(-74)} \\ 0.00857^{(+20)}_{(-18)} & 0.04110^{(+83)}_{(-72)} & 0.999118^{(+31)}_{(-36)} \end{pmatrix} \tag{1.23}$$

The presence of the complex phase δ in Equation (1.22) allows for *CP* symmetry to be violated ($N = 3$ is the lowest possible number of fermion generations to have a phase appear in the CKM matrix), but it is not a sufficient condition; a relation between the quark masses and the elementes of the matrix must also hold [39]:

$$J_{CP} \left(m_t^2 - m_c^2\right) \left(m_t^2 - m_u^2\right) \left(m_c^2 - m_u^2\right) \left(m_b^2 - m_s^2\right) \left(m_b^2 - m_d^2\right) \left(m_s^2 - m_d^2\right) \neq 0, \tag{1.24}$$

where

$$J_{CP} = \left|\text{Im}\left(V_{i\alpha} V_{j\beta} V_{i\beta}^* V_{j\alpha}^*\right)\right| \quad (i \neq j, \alpha \neq \beta) \tag{1.25}$$

is called the Jarlskog invariant. Equation (1.24) implicitly contains all the necessary conditions for *CP* violation to arise in the SM:

1. $m_u \neq m_c \neq m_t$, $m_d \neq m_s \neq m_b$, *i.e.* there should not be mass degeneracy within up- and down-type quarks;
2. The three mixing angles θ_1, θ_2, and θ_3 should not be 0 or $\frac{\pi}{2}$;
3. The phase δ should not be 0 or π.

Inspecting the CKM matrix reported in (1.23), it can be seen that the elements on the diagonal are close to 1, and that they get smaller the further away from the main diagonal, indicating a less common transition between different quark families. This hierarchy can be made more evident by employing the so-called Wolfenstein parameterisation[1] [40]. It is based on the four parameters A, λ, ρ, and η, defined by

$$\lambda = s_{12} = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \approx 0.22, \tag{1.26}$$

$$A\lambda^2 = s_{23} = \lambda \left| \frac{V_{cb}}{V_{ud}} \right| \approx 0.04, \tag{1.27}$$

$$A\lambda^3(\rho + i\eta) = s_{13}e^{i\delta} = V_{ub}^* \approx 4 \times 10^{-3}. \tag{1.28}$$

With these definitions, the CKM matrix can be written as a power series of λ. For example, if we go up to third order we get

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4); \tag{1.29}$$

the hierarchy is clearly visible now: the elements on the main diagonal are 1 at leading order, while we have to reach third order to see the first non-vanishing elements in the off-diagonal places.

The unitarity condition of the CKM matrix, $V_{ij}V_{ik}^* = \delta_{jk}$ and $V_{ij}V_{kj}^* = \delta_{ik}$, leads to 12 equations relating the matrix entries, three sums equalling 1 in the main diagonal and the remaining ones summing to 0. The latter can be geometrically interpreted as the vanishing sums of three complex numbers, *i.e.* as triangles in the complex plane which all share the same area of $J_{CP}/2$. Two of these equations are particularly interesting, as they contain terms of the same order in λ:

$$\underbrace{V_{ud}V_{ub}^*}_{A\lambda^3(\rho+i\eta)} + \underbrace{V_{cd}V_{cb}^*}_{-A\lambda^3} + \underbrace{V_{td}V_{tb}^*}_{A\lambda^3(1-\rho-i\eta)} = 0, \tag{1.30}$$

$$\underbrace{V_{ud}^*V_{td}}_{A\lambda^3(1-\rho-i\eta)} + \underbrace{V_{us}^*V_{ts}}_{-A\lambda^3} + \underbrace{V_{ub}^*V_{tb}}_{A\lambda^3(\rho+i\eta)} = 0. \tag{1.31}$$

The triangle defined by Equation (1.30) is used as a benchmark to constrain the parameters of the CKM matrix and is referred to as the Unitary Triangle (UT). If we divide all its sides by $V_{cd}V_{cb}^*$, its vertices move to $(0,0)$, $(1,0)$, and $(\overline{\rho},\overline{\eta}) = (\rho(1-\lambda^2/2), \eta(1-\lambda^2/2))$, while the internal angles are:

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \approx \arg\left(-\frac{1-\rho-i\eta}{\rho+i\eta}\right), \tag{1.32}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \approx \arg\left(-\frac{1}{1-\rho-i\eta}\right), \tag{1.33}$$

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \approx \arg\left(\rho+i\eta\right). \tag{1.34}$$

[1] Bear in mind that any parameterisation of the CKM matrix is arbitrary and none is intrinsically better; on the other hand, some parameterisations can be more insightful than others from a phenomenological point of view.

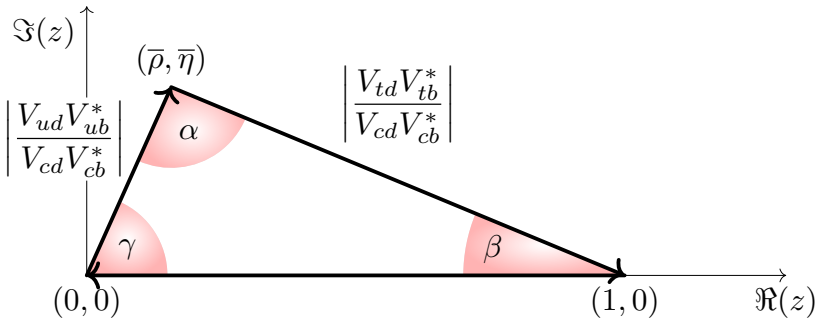

Figure 1.2: Graphical representation of the unitary triangle defined by Equation (1.30) in the complex plane.

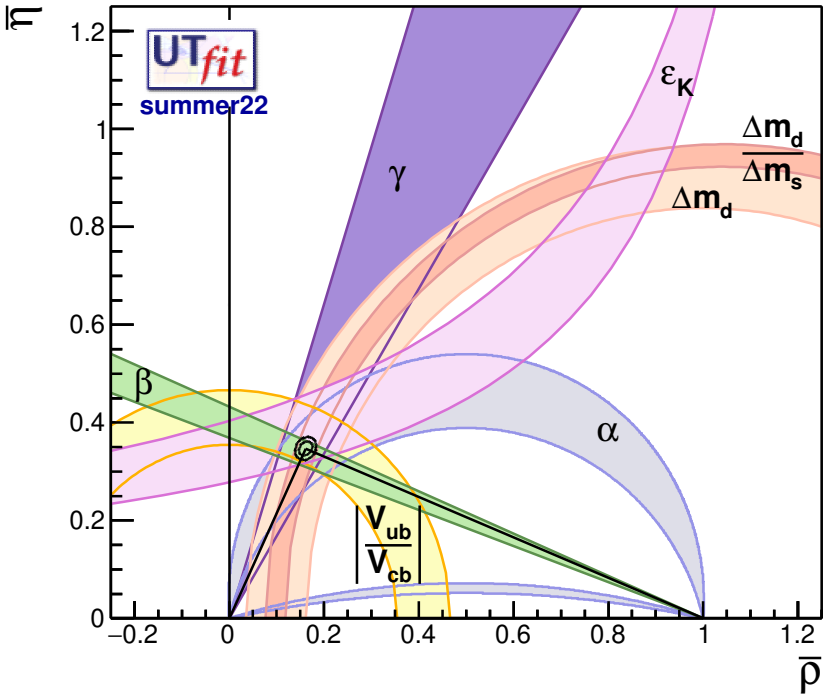

Figure 1.3: Current experimental status of the properties of the unitary triangle [41].

The UT is shown in Figure 1.2, while the most recent experimental constraints, obtained combining the measurements of several different observables, are reported in Figure 1.3.

The presence of the complex phase in the CKM matrix opens the possibility for *CP* symmetry to be violated in the charged-current quark interactions. On a more general basis [42], *CP* violation depends on the existence of more than one amplitude for the same process, with different weak and strong phases. The name "weak" and "strong" do not refer to the interaction responsible for it, but only to their behaviour under a *CP* transformation: we call *weak* the *CP*-odd phases and *strong* the *CP*-even ones.

Weak phases can originate as complex couplings in the Lagrangian, as they do for

example in the Fermi theory of the neutron decay with the vector and axial components of the hadronic current [43]. Another possibility is to appear as a complex phase in the CKM matrix, as the δ parameter in Equation (1.22) does. Strong phases, on the other hand, have two possible sources: one as the trace of a product of an even number of γ matrices together with γ^5 entering the Lagrangian, and the other comes from final-state interaction (FSI), where an initial state i decays weakly to an intermediate state f' which then scatters to a final state f via a strong or electromagnetic amplitude.

It is important to notice that these phases do not have any physical meaning by themselves, since it is always possible to change them at will with a phase transformation. Instead, the relative phase of two competing contributions, *i.e.* the difference of the phases of the two amplitudes, does not change with a rephasing and therefore has a physical importance. To understand why *CP* can only occur when at least two different amplitudes with different weak and strong phases contribute to the same decay, we can start with the simple example of a single-amplitude process $i \to f$ and its *CP*-conjugate equivalent $\overline{i} \to \overline{f}$, mediated by a Hamiltonian $\mathcal{H}$:

$$\langle f|\mathcal{H}|i\rangle = Ae^{i(\delta+\phi)}, \tag{1.35}$$

$$\left\langle \overline{f}\middle|\mathcal{H}\middle|\overline{i}\right\rangle = Ae^{i(\delta-\phi+\theta)}, \tag{1.36}$$

where A is a real positive number, ϕ is a weak *CP*-odd phase, δ a *CP*-even strong phase, and θ a spurious phase introduced by the *CP* transformation. It can be immediately seen that in this scenario *CP* in conserved, since $|\langle f|\mathcal{H}|i\rangle| - \left|\left\langle \overline{f}\middle|\mathcal{H}\middle|\overline{i}\right\rangle\right| = |A| - |A| = 0$.

Let us now introduce an additional contribution to the decay amplitude instead:

$$\langle f|\mathcal{H}|i\rangle = A_1e^{i(\delta_1+\phi_1)} + A_2e^{i(\delta_2+\phi_2)}, \tag{1.37}$$

$$\left\langle \overline{f}\middle|\mathcal{H}\middle|\overline{i}\right\rangle = A_1e^{i(\delta_1-\phi_1+\theta)} + A_2e^{i(\delta_2-\phi_2+\theta)}, \tag{1.38}$$

in this case, the difference of the two *CP*-conjugated amplitudes does not cancel out:

$$|\langle f|\mathcal{H}|i\rangle|^2 - \left|\left\langle \overline{f}\middle|\mathcal{H}\middle|\overline{i}\right\rangle\right|^2 = -4A_1A_2\sin(\delta_1-\delta_2)\sin(\phi_1-\phi_2), \tag{1.39}$$

therefore *CP* violation can occur if $\delta_1 \neq \delta_2$ and $\phi_1 \neq \phi_2$ (and obviously $A_1 \neq 0$, $A_2 \neq 0$). A common experimental observable to search for *CP* violation in this case is the *CP* asymmetry

$$A_{CP} = \frac{|\langle f|\mathcal{H}|i\rangle|^2 - \left|\left\langle \overline{f}\middle|\mathcal{H}\middle|\overline{i}\right\rangle\right|^2}{|\langle f|\mathcal{H}|i\rangle|^2 + \left|\left\langle \overline{f}\middle|\mathcal{H}\middle|\overline{i}\right\rangle\right|^2} = \frac{-2A_1A_2\sin(\delta_1-\delta_2)\sin(\phi_1-\phi_2)}{A_1^2+A_2^2+2A_1A_2\cos(\delta_1-\delta_2)\cos(\phi_1-\phi_2)} \tag{1.40}$$

1.3 Two-body b-hadron decays

Hadronic two-body decays of beauty hadrons provide an interesting testbench for studying *CP* violation in the SM, thanks to the number of different Feynman topologies that can contribute to each decay. The relatively high contribution of loop transitions can potentially enhance the effect of new particles and interactions beyond the SM.

A general two-body B decay is mediated by the process $b \to q_1\overline{q}_2d(s)$, with $q_1, q_2 \in \{u, d, c, s\}$. The Feynman diagrams contributing to this transition can be divided into two categories: tree-level and loop (also called penguin) topologies. The flavour content of the final state restricts the possible combinations:

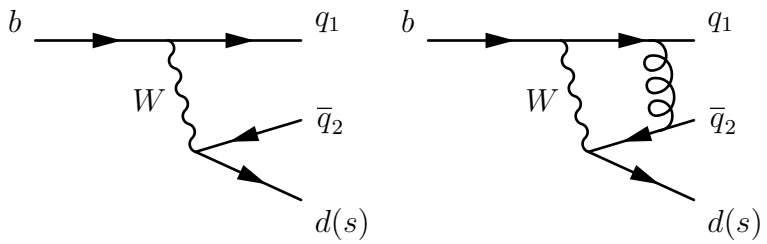

Figure 1.4: Example of Feynman diagrams for a $b \to q_1\overline{q}_2 d(s)$ transistion at tree-level (left) and with a first-order QCD correction (right).

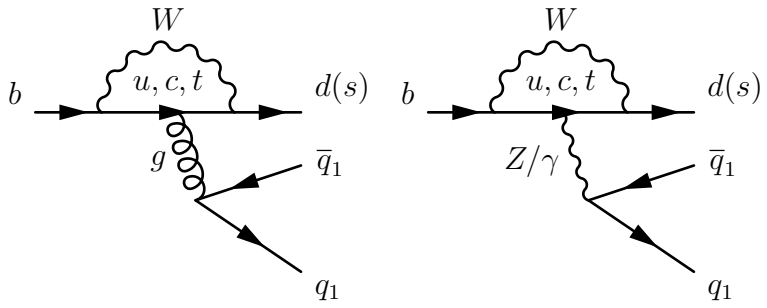

Figure 1.5: Example of Feynman diagrams for a $b \to q_1\overline{q}_2 d(s)$ QCD (left) and EW (right) penguin transition.

- $q_1 \neq q_2 \in \{u, c\}$: only possible at tree-level;
- $q_1 = q_2 \in \{u, c\}$: possible both at tree-level and via loops;
- $q_1 = q_2 \in \{d, s\}$: only possible via loops;

examples of such a diagram can be seen in Figures 1.4 and 1.5.

In order to describe the weak decays of b-hadrons, the strong interaction between the constituent quarks must also be taken into account. The typical approach is to treat it under the framework of operator product expansion (OPE), in which a separation of the energy scales of short-distance (high-energy) and long-distance (low-energy) contributions is defined by a parameter μ [44–47]. Any contributions from particles heavier than μ enter only through the Wilson coefficients $C_i(\mu)$, which account for all short-distance effects; long-distance strong-interaction contributions, instead, are included in the corresponding matrix elements of local operators $Q_i(\mu)$ that need to be evaluated with non-perturbative techniques. This factorization of the two energy regimes allows an effective weak Hamiltonian to be written in the following way [48]:

$$\langle f|\mathcal{H}_{\text{eff}}|B\rangle = \frac{G_F}{\sqrt{2}} \sum_i V_i^{\text{CKM}} C_i(\mu) \langle f|Q_i(\mu)|B\rangle \,, \tag{1.41}$$

where $G_F/\sqrt{2} = g^2/(8M_W^2)$ is the Fermi constant and V_i^{CKM} are factors related to the CKM matrix elements that set the strentgh with which each operator enters the Hamiltonian. As

the amplitude shouldn't depend on the energy scale μ, it is required that the μ dependence of the $C_i(\mu)$ cancels the dependence of the $Q_i(\mu)$.

To illustrate the use of OPE, let us first start with the simple example of a purely weak tree-level transition. Neglecting QCD corrections for the moment, the effective Hamiltonian can be written as [42, 49]

$$\begin{aligned}\mathcal{H}_{\text{eff}} &= \frac{G_F}{\sqrt{2}} V_{q_2 r}^* V_{q_1 b} \frac{M_W^2}{k^2 - M_W^2} (\overline{r}^i q_2^i)_{V-A} (\overline{q}_1^j b^j)_{V-A} \\ &= -\frac{G_F}{\sqrt{2}} V_{q_2 r}^* V_{q_1 b} (\overline{r}^i q_2^i)_{V-A} (\overline{q}_1^j b^j)_{V-A} + \mathcal{O}\left(\frac{k^2}{M_W^2}\right), \end{aligned} \tag{1.42}$$

where $r \in \{d, s\}$, the indices i and j label the quark colour, and the terms of $\mathcal{O}\left(\frac{k^2}{M_W^2}\right)$ can be neglected given the smallness of the transferred momentum k, that is of the order of the quark masses, with respect to M_W. The label V and A refer to the vector and axial current respectively, *i.e.*

$$(\overline{q} q')_{V \pm A} = \overline{q} \gamma^\mu (1 \pm \gamma_5) q'. \tag{1.43}$$

Since no QCD correction is considered yet, Equation (1.42) is equivalent to having a Wilson coefficient $C_2(\mu) = 1$ and a single matrix element

$$Q_2 = (\overline{r}^i q_2^i)_{V-A} (\overline{q}_1^j b^j)_{V-A}. \tag{1.44}$$

If we now add a simple QCD correction such as a gluon connecting two quark currents (like in Figure 1.4 on the right), this must necessarily mix different color charges, therefore we are required to add an operator of the form

$$Q_1 = (\overline{r}^i q_2^j)_{V-A} (\overline{q}_1^j b^i)_{V-A} \tag{1.45}$$

and a corresponding Wilson coefficient $C_1(\mu)$. The effective Hamiltonian then becomes

$$\mathcal{H}_{\text{eff}} = -\frac{G_F}{\sqrt{2}} V_{q_2 r}^* V_{q_1 b} (C_1(\mu) Q_1 + C_2(\mu) Q_2). \tag{1.46}$$

Generalising now the procedure to all possible topologies of Feynman diagrams that can contribute to a two-body hadronic B decay, we can write the effective Hamiltonian as

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \left[V_{ur}^* V_{ub} \sum_{k=1}^{2} C_k(\mu) Q_k^{ur} + V_{cr}^* V_{cb} \sum_{k=1}^{2} C_k(\mu) Q_k^{cr} - V_{tr}^* V_{tb} \sum_{k=3}^{10} C_k(\mu) Q_k^r \right], \tag{1.47}$$

where the 10 operators (and relative Wilson coefficients) are grouped into three categories:

- Current-current tree-level operators ($q \in \{u, c\}$)

$$Q_1^{qr} = (\overline{r}^i q^j)_{V-A} (\overline{q}^j b^i)_{V-A} \tag{1.48}$$

$$Q_2^{qr} = (\overline{r}^i q^i)_{V-A} (\overline{q}^j b^j)_{V-A} \tag{1.49}$$

- QCD penguins (*e.g.* Figure 1.5, left):

$$Q_3^r = (\overline{r}^i b^i)_{V-A} \sum_{q\in\{u,d,c,s,b\}} (\overline{q}^j q^j)_{V-A}, \tag{1.50}$$

$$Q_4^r = (\overline{r}^i b^j)_{V-A} \sum_{q\in\{u,d,c,s,b\}} (\overline{q}^j q^i)_{V-A}, \tag{1.51}$$

$$Q_5^r = (\overline{r}^i b^i)_{V-A} \sum_{q\in\{u,d,c,s,b\}} (\overline{q}^j q^j)_{V+A}, \tag{1.52}$$

$$Q_6^r = (\overline{r}^i b^j)_{V-A} \sum_{q\in\{u,d,c,s,b\}} (\overline{q}^j q^i)_{V+A}, \tag{1.53}$$

- Electroweak penguins (*e.g.* Figure 1.5, left):

$$Q_7^r = \frac{3}{2}(\overline{r}^i b^i)_{V-A} \sum_{q\in\{u,d,c,s,b\}} e_q(\overline{q}^j q^j)_{V+A}, \tag{1.54}$$

$$Q_8^r = \frac{3}{2}(\overline{r}^i b^j)_{V-A} \sum_{q\in\{u,d,c,s,b\}} e_q(\overline{q}^j q^i)_{V+A}, \tag{1.55}$$

$$Q_9^r = \frac{3}{2}(\overline{r}^i b^i)_{V-A} \sum_{q\in\{u,d,c,s,b\}} e_q(\overline{q}^j q^j)_{V-A}, \tag{1.56}$$

$$Q_{10}^r = \frac{3}{2}(\overline{r}^i b^j)_{V-A} \sum_{q\in\{u,d,c,s,b\}} e_q(\overline{q}^j q^i)_{V-A}. \tag{1.57}$$

This summarizes the OPE approach to computing decay amplitudes. Provided the necessary changes are applied to the hadronix matrix elements in Equations (1.48) to (1.57) for different exclusive final states, this approach can be considered "universal", meaning that all the results shown above can be applied equally to every B decay governed by the same quark-level transition $b \to q_1\overline{q}_2 d(s)$. The difficult task is then to actually compute each matrix element; there are different methods available to do so, each with its own tradeoffs between ease of computation, need for experimental input, and size of the theoretical uncertainties. Since a detailed explanation of the various methods is beyond the scope of this thesis, we will just give a brief description of them:

QCD Factorization (QCDF): this approach is based on the assumption that factorization holds when hadrons containing quarks of mass m_q larger than the strong interaction scale Λ_{QCD} are involved, which allows to compute amplitudes at leading order in Λ_{QCD}/m_q [50–53].

Perturbative QCD (pQCD): starting on the same footing as QCDF, it goes further into computing higher-order QCD corrections thanks to a resummation of Sudakov effects [54–56]. In this way, less experimental inputs are needed compared to QCDF.

Soft Collinear Effective Theory (SCET): a more recent technique in which the emission of soft and collinear gluons is studied, allowing factorization to be generalised at all orders of α_s [57–60].

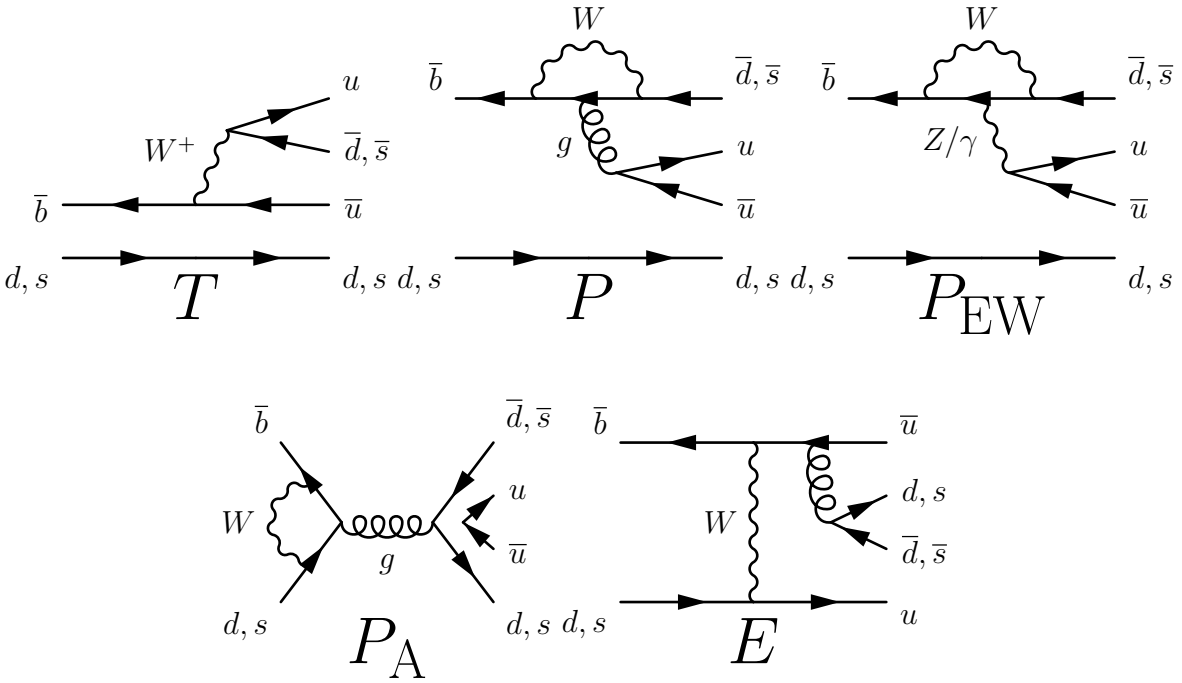

Figure 1.6: Feynman diagrams of processes contributing to two-body charmless B meson decays: tree-level (T), QCD penguin (P), electroweak penguin (P_{EW}), penguin annihilation (P_{A}), and exchange (E). The equivalent diagrams for baryon decays can be easily obtained by adding one spectator quark line.

QCD Light-Cone Sum-Rules: computations of hadronic two-body B decay amplitudes have been shown to be feasible with QCD sum rules as well [61–63].

Let us now focus on the phenomenology of charmless two-body b-hadron decays, which we will generally refer to as $H_b \to h^+h'^-$. This category comprises the decays $B^0 \to K^+K^-$, $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B_s^0 \to K^+K^-$, $B_s^0 \to \pi^+\pi^-$, $B_s^0 \to \pi^+K^-$, $\Lambda_b^0 \to pK^-$, and $\Lambda_b^0 \to p\pi^-$[2]. Many different observables can be studied for these decays, such as branching fractions, integrated CP asymmetries, as well as time-dependent CP asymmetries for neutral B mesons. This group of decays have been extensively studied both at past and present accelerator facilities [7,13,64–80]. The discussion above regarding the computation of the decay amplitudes for hadronic two-body decays of b-hadrons naturally applies also to the subgroup of charmless decays, but in this case the interpretation of the experimental results is not simple, because of the sizable contributions from penguin topologies that these decays contain alongside the tree-level transitions, making it harder to evaluate the underlying CKM phases. On the other hand, the higher relative size of loop diagrams to the total decay amplitude allows for possible New Physics (NP) contributions to amplify CP-violating effects.

In the SM, the topologies contributing to the $B \to h^+h'^-$ decay amplitudes [81] are tree-level (T), QCD penguin (P), electroweak penguin (P_{EW}), penguin annihilation (P_{A}), and exchange (E), which are shown in Figure 1.6. The T, P, and P_{EW} topologies have in common the presence of a spectator quark, while in P_{A} and E all initial- and final-state quarks are involved. The latter are expected to be suppressed by a factor $f_B/m_B \approx 5\%$,

[2]From now on, the inclusion of CP-conjugate decays is implied throughout this document.

Table 1.2: Scheme of the contribution of Feynman diagrams from Figure 1.6 to each $B \to h^+h'^-$ decay. The current knowledge of the branching fraction of each decay is also reported [32].

Decay	T	P	P_{EW}	P_{A}	E	$\mathcal{B}$ $[10^{-6}]$
$B_s^0 \to K^+K^-$	✓	✓	✓	✓	✓	26.6(22)
$B^0 \to K^+\pi^-$	✓	✓	✓	✗	✗	19.6(5)
$B_s^0 \to \pi^+K^-$	✓	✓	✓	✗	✗	5.8(7)
$\Lambda_b^0 \to pK^-$	✓	✓	✓	✗	✗	5.4(10)
$B^0 \to \pi^+\pi^-$	✓	✓	✓	✓	✓	5.12(19)
$\Lambda_b^0 \to p\pi^-$	✓	✓	✓	✗	✗	4.5(8)
$B_s^0 \to \pi^+\pi^-$	✗	✗	✗	✓	✓	0.7(1)
$B^0 \to K^+K^-$	✗	✗	✗	✓	✓	0.078(15)

where f_B is the decay constant of the B meson [82]. The individual contributions of these diagrams to each decay is reported in Table 1.2. Some decays are related by a simple exchange of two quarks, that can be summarised as follows:

$$\underbrace{B^0 \to \pi^+\pi^-}_{T+P+2/3P_{\text{EW}}+P_{\text{A}}+E} \quad d \longleftrightarrow s \quad \underbrace{B_s^0 \to K^+K^-}_{T+P+2/3P_{\text{EW}}+P_{\text{A}}+E}\,, \tag{1.58}$$

$$\underbrace{B^0 \to K^+\pi^-}_{T+P+2/3P_{\text{EW}}} \quad d \longleftrightarrow s \quad \underbrace{B_s^0 \to \pi^+K^-}_{T+P+2/3P_{\text{EW}}}\,, \tag{1.59}$$

$$\underbrace{B^0 \to K^+\pi^-}_{T+P+2/3P_{\text{EW}}} \quad d \overset{\text{spect.}}{\longleftrightarrow} s \quad \underbrace{B_s^0 \to K^+K^-}_{T+P+2/3P_{\text{EW}}+P_{\text{A}}+E}\,, \tag{1.60}$$

$$\underbrace{B^0 \to \pi^+\pi^-}_{T+P+2/3P_{\text{EW}}+P_{\text{A}}+E} \quad d \overset{\text{spect.}}{\longleftrightarrow} s \quad \underbrace{B_s^0 \to \pi^+K^-}_{T+P+2/3P_{\text{EW}}}\,, \tag{1.61}$$

$$\underbrace{\Lambda_b^0 \to p\pi^-}_{T+P+2/3P_{\text{EW}}} \quad d \longleftrightarrow s \quad \underbrace{\Lambda_b^0 \to pK^-}_{T+P+2/3P_{\text{EW}}}\,, \tag{1.62}$$

where the notation $d \overset{(\text{spect.})}{\longleftrightarrow} s$ indicates a symmetry with respect to the exchange between a d and an s interacting (spectating) quark.

Regarding the $\Lambda_b^0 \to p\pi^-$ and $\Lambda_b^0 \to pK^-$ decays, theoretical computations of the branching fractions and *CP* asymmetries have been performed only in recent times. Several different approaches have been used, such as QCDF [83, 84], pQCD [85], and the MIT bag model [86]. The results of these computations, together with the experimental measurements done by CDF [70] and LHCb [13] collaborations, and their average [32], are reported in Table 1.3 and shown in Figure 1.7. Theoretical computations for the $\Lambda_b^0 \to pK^-$ decay tend to predict a large positive value for the *CP* asymmetry, which has been instead measured as negative, though with large uncertainties and hence still compatible with 0. This indicates that newer and more precise measurements are needed in order to improve the knowledge on *CP* violation in baryon decays.

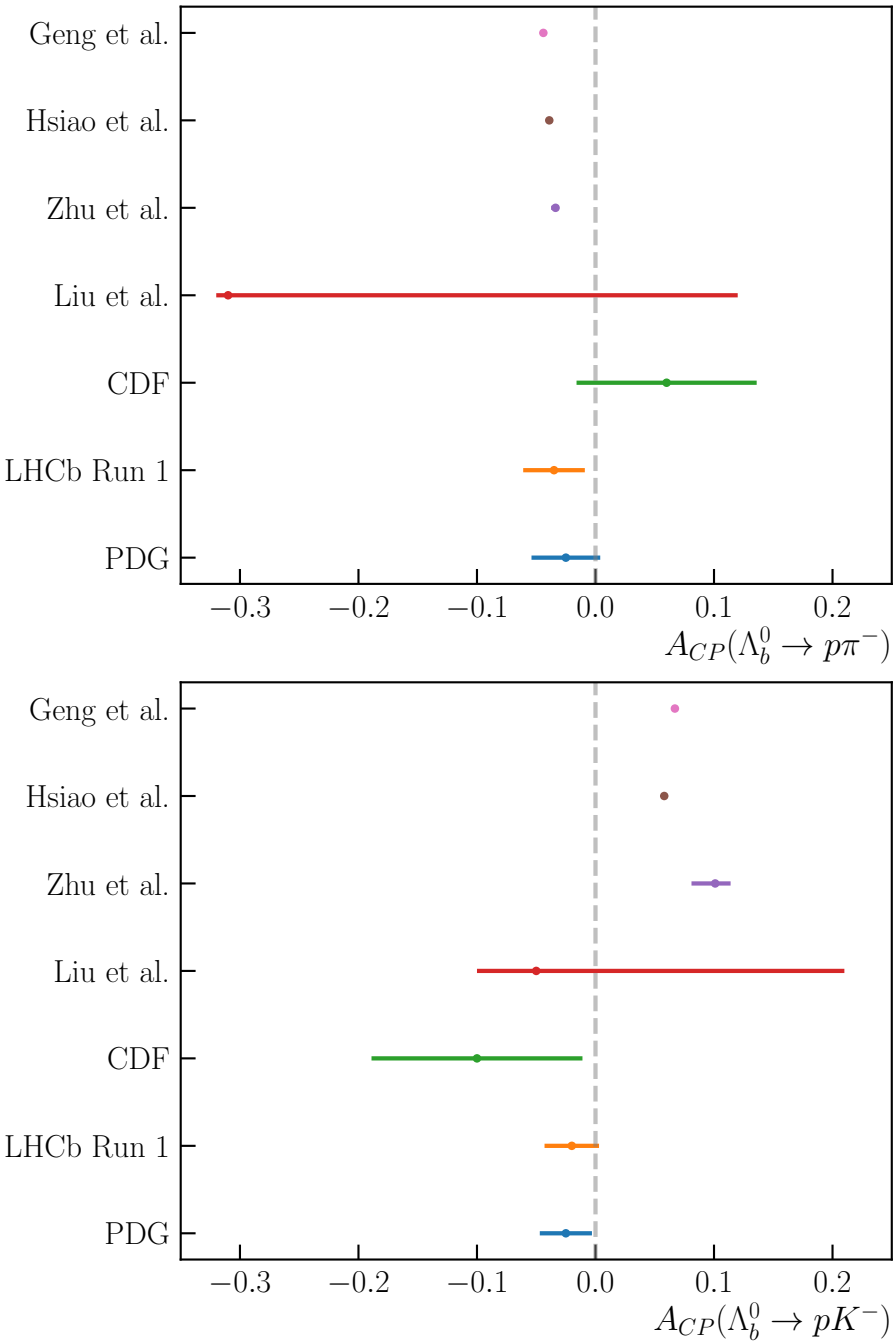

Figure 1.7: Comparison between the theoretical estimates of $A_{CP}(\Lambda_b^0 \to p\pi^-)$ (top) and $A_{CP}(\Lambda_b^0 \to pK^-)$ (bottom) and the values measured by the CDF and LHCb collaborations, together with the PDG average of the two.

Table 1.3: Current theoretical and experimental knowledge of the *CP*-violating properties of the decays $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$.

	$A_{CP}(\Lambda_b^0 \to p\pi^-)$ [%]	$A_{CP}(\Lambda_b^0 \to pK^-)$ [%]
Theoretical		
[85] (pQCD)	-32^{+49}_{-1}	-3^{+25}_{-4}
[84] (NLO QCDF)	$-3.37^{+0.29}_{-0.37}$	$10.1^{+1.3}_{-2.0}$
[83] (QCDF)	-3.9 ± 0.4	$6.7^{+0.3}_{-0.2} \pm 0.3$
[86] (MIT bag model)	-4.4 ± 0.1	6.7 ± 0.0
Experimental		
CDF [70]	$6 \pm 7 \pm 3$	$-10 \pm 8 \pm 4$
LHCb Run 1 [13]	$-3.5 \pm 1.7 \pm 2.0$	$-2.0 \pm 1.3 \pm 1.9$
PDG [32]	-2.5 ± 2.9	-2.5 ± 2.2

Chapter 2

The LHCb detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [87], situated approximately 100 meters underground near Geneva and spanning a length of 26.7 km, is a hadron synchrotron that occupies the tunnel once utilized by the Large Electron Positron collider. Engineered for proton collisions with a center-of-mass energy of 14 TeV and an instantaneous luminosity exceeding $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, as well as lead ion collisions at 2.76 TeV per nucleon with an instantaneous luminosity of $10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, the LHC directs hadron beams in opposite directions within two accelerating rings. These beams converge at four interaction points, where the experiments ALICE, ATLAS, CMS, and LHCb are installed.

Due to the impracticality of accelerating protons directly to 7 TeV from a quasi-rest state, a series of preaccelerators — Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS) — are employed. The process begins with Linac2 accelerating protons to 50 MeV, injecting them into the PSB, which elevates their energy to 1.4 GeV. Subsequently, the PS increases their energy to 25 GeV, forming proton bunches, each comprising about 10^{11} protons. These bunches are then transferred to the SPS, where they are accelerated to 450 GeV. Finally, the proton bunches are injected from the SPS into the LHC, where their energy is raised to the target value. A schematic representation of the CERN accelerator complex is depicted in Figure 2.1.

During the initial run (Run 1) of data collection, the LHC collided protons at center-of-mass energies of 7 and 8 TeV in 2011 and 2012, respectively. In the second run (Run 2), spanning the years 2015-2018, the collision energy increased to 13 TeV.

2.2 The LHCb detector

The LHCb experiment [89, 90] was originally designed to study processes involving beauty and charm hadrons. The core objectives of the collaboration include probing physics processes within and beyond the Standard Model by studying *CP* violation in b or c decays, observation of rare processes, heavy-quark spectroscopy, and searches for exotic states. During the years the physics programme of the experiment has widened beyond the original design to include electroweak physics, heavy ions, and fixed-target measurements.

LHCb is designed as a single-arm forward spectrometer with a geometrical acceptance

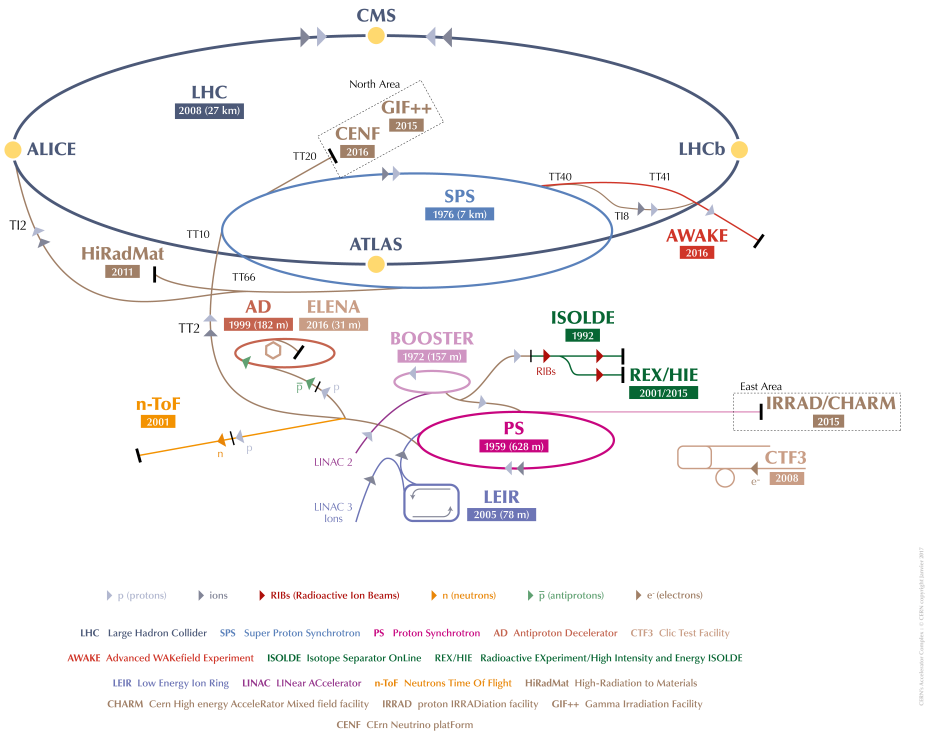

Figure 2.1: The CERN accelerator complex [88].

spanning $[10, 250]$ mrad in the vertical plane and $[10, 300]$ mrad in the horizontal plane. The unique geometry of the LHCb detector is driven by the fact that the majority of $b\bar{b}$ pairs resulting from pp collisions hadronize into b hadrons with a significant boosts towards the forward (or backward) beam direction, maintaining a small angle with respect to the beam axis, as illustrated in Fig. 2.2.

The LHCb collaboration has measured the $b\bar{b}$ production cross-section in the pseudorapidity[1] range $2 < \eta < 5$ at center-of-mass energies of 7 and 13 TeV, yielding $(72.0 \pm 0.3 \pm 6.8)$ mb and $(144 \pm 1 \pm 21)$ mb, respectively [91]. The LHCb experiment is also well-suited for the investigation of charm physics, with the added advantage that the $c\bar{c}$ production cross-section is notably higher than that of $b\bar{b}$, specifically $(1419 \pm 12 \pm 116 \pm 65)$ mb [92] and $(2369 \pm 3 \pm 152 \pm 118)$ mb [93] at $\sqrt{s} = 7$ and 13 TeV, respectively.

To maintain a low pile-up and minimize radiation damage to detector elements near the beam pipe, the luminosity at the interaction point of LHCb is deliberately kept below the maximum achievable from the LHC by adjusting the transverse separation between the beams. This is achieved through a luminosity leveling technique [94], where the beams are progressively brought closer to each other in the transverse plane to maintain a constant collision rate throughout the beam's lifetime. In 2011, the instantaneous luminosity was set at approximately $3 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, followed by $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in 2012, 2015, 2016,

[1]The pseudorapidity of a particle traveling at a polar angle θ with respect to the beam axis is defined as $\eta = -\log\left(\tan\frac{\theta}{2}\right)$.

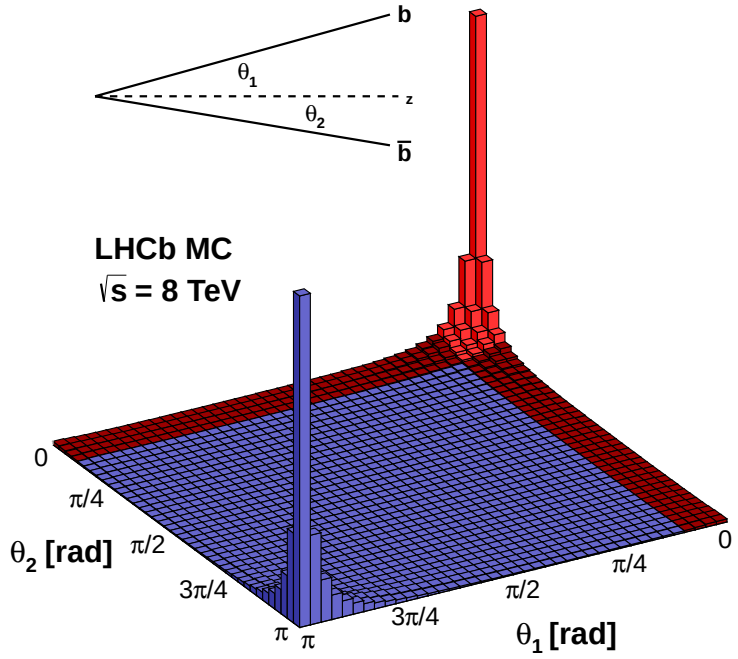

Figure 2.2: Production angles with respect to the beam direction of $b\bar{b}$ pairs in simulated pp collisions at $\sqrt{s} = 8\,\text{TeV}$. The red area represent the geometrical acceptance of LHCb.

and 2017, and about $5 \times 10^{32}\,\text{cm}^{-2}\,\text{s}^{-1}$ in 2018. The total integrated luminosity collected by the LHCb experiment up to now is summarized in Figure 2.3.

In Fig. 2.4 a diagram of the full LHCb detector that operated in Run 1 and 2 is shown; a right-handed coordinate system is defined such that the z axis is oriented along the beam and the y axis along the vertical pointing upwards. The apparatus comprised multiple sub-detectors, each with a design optimized to measure a specific property of particles produced in the pp collisions. These detectors can be broadly categorized into two classes based on their purposes:

- the Tracking system, consisting of the Vertex Locator (VELO), the magnet, the Trigger Tracker (TT), and three tracking stations (T1–T3);
- the Particle Identification (PID) system, including two Ring Imaging Cherenkov detectors (RICH1 and RICH2), the Electromagnetic Calorimeter (SPD, PS, and ECAL), the Hadronic Calorimeter (HCAL), and five muon stations (M1–M5);

Each sub-detector will be described in detail in the following sections.

2.3 The LHCb tracking system

The tracking system of the LHCb experiment is engineered to reconstruct the trajectories of charged particles and measure their momentum starting from the VELO, situated in

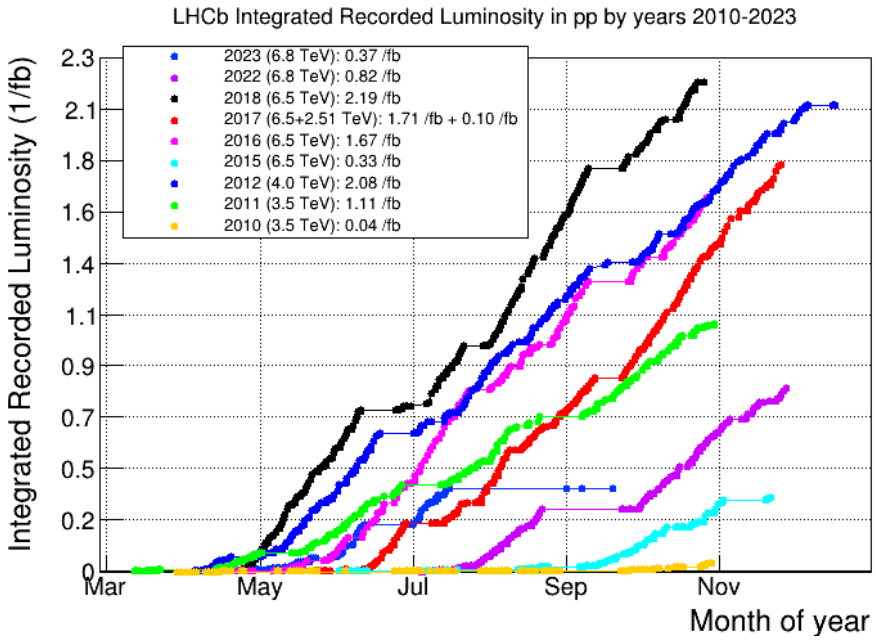

Figure 2.3: Integrated luminosity collected at LHCb, divided by years of data taking.

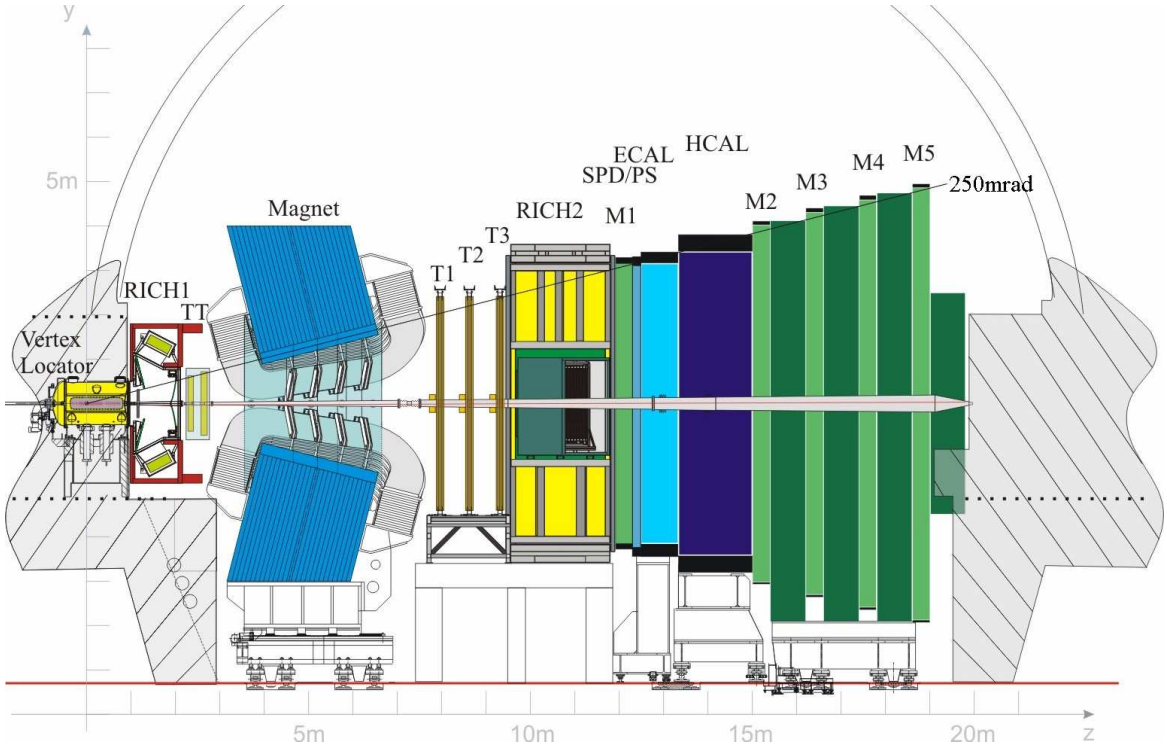

Figure 2.4: Side view of the LHCb detector in the (y, z) plane. The various sub-detectors are also shown and labelled.

close proximity to the interaction point. Silicon trackers are employed to capture the coordinates of particles traveling in close proximity to the beam axis, while the outer tracker, consisting of straw-tubes, covers the broader acceptance region externally. The inclusion of a dipole magnet enables the precise measurement of particle's momentum.

2.3.1 The Vertex Locator

Beauty and charm hadrons, which decay weakly, travel approximately 1 cm within the LHCb detector before decaying. Consequently, the presence of displaced secondary vertices is a distinctive characteristic of b- and c-hadron decays. The precise reconstruction of these vertices at a micro-metric scale is crucial to select signal events while effectively rejecting the majority of background events. A high spatial resolution is also essential to conduct time-dependent analyses, relying on the accurate measurement of the particles' decay time.

The VELO detector [95, 96], positioned closest to the interaction point, meets these requirements with the following key features (see Figure 2.5):

- The resolution on the position of the primary vertex (PV), when reconstructed with 25 tracks, is 71 µm along the z-axis (beam direction) and 13 µm in the transverse plane;
- The resolution on the impact parameter (IP) of charged tracks, calculated with respect to the PV along x or y, is below 35 µm for particles with $p_{\mathrm{T}} > 1$ GeV/c and decreases asymptotically to 12 µm at high transverse momentum;
- The decay time of heavy-flavoured hadrons is measured with a typical resolution of about 50 fs;

Comprising 23 modules placed orthogonally to the beam, the VELO detector employs radiation-tolerant silicon strip sensors, each 300 µm thick. Split into left and right parts, each station measures the radial coordinate and azimuthal angle of charged tracks using two distinct types of modules, called R and ϕ sensors, respectively. Two stations equipped solely with radial sensors are positioned upstream of the interaction point to gauge the number of interactions per collision and provide pile-up information for the trigger.

During LHC injection, the two sides are moved to a safety distance of 3 cm from the beam, while during collisions of stable beams, they are brought closer to a distance of only 5 mm. The modules are housed in a vacuum vessel separated from the beam vacuum by a 0.5 mm aluminum sheet. A schematic representation of the VELO is illustrated in Figure 2.6.

2.3.2 The Trigger Tracker

The Trigger Tracker (TT) is a four-layer detection system positioned after the RICH1 sub-detector and before the magnetic dipole, located approximately 2.4 m from the beam interaction region. The system consists of two stations, TTa and TTb, each containing two layers, with a separation of about 30 cm. Employing silicon microstrip technology, the TT uses 500 µm thick vertical micro-strips, with pitch of 200 µm and lengths up to 38 cm. In the first and fourth layers, the strips are oriented parallel to the vertical axis,

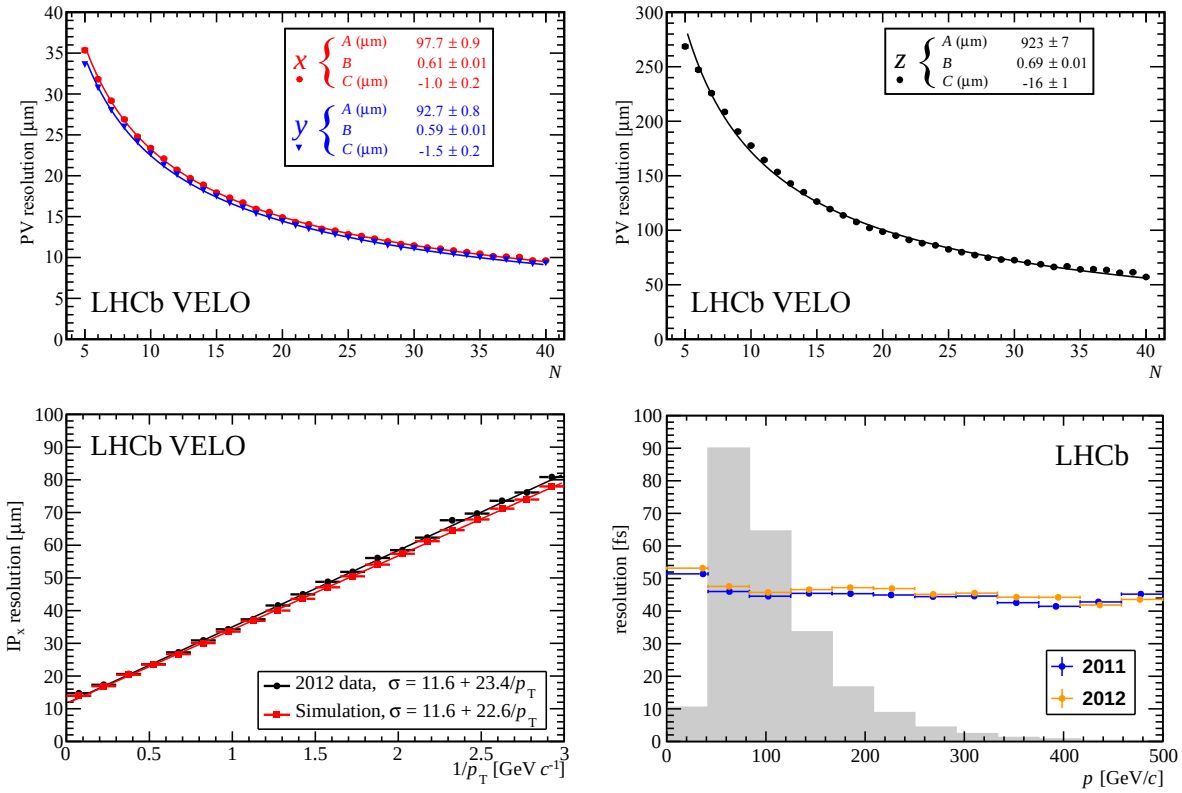

Figure 2.5: Summary of the performance of the VELO: x, y, and z PV position resolution versus track multiplicity (top), IP resolution versus $1/p_{\mathrm{T}}$ (bottom left), decay time resolution versus momentum (bottom right).

while in the second and third layers, they are tilted by $+5°$ and $-5°$, respectively; this design, called x–u–v–x, allows both the x and y coordinates of the hits in the detector (see Figure 2.7) to be reconstructed.

The role of the TT is to supply reference segments, enhancing the resolution of momentum and trajectory by combining tracks from the VELO and tracking stations, consequently reducing the occurrence of fake tracks. It also serves the purpose of providing transverse momentum information of high-p_{T} tracks for the High Level Trigger, as well as for low-momentum tracks that are bent outside of the detector acceptance before reaching the tracking stations. The single-hit resolution of the TT during Run 1 was approximately 52 µm with a hit efficiency above 99.7% (as depicted in Fig. 2.8).

2.3.3 The tracking stations

The tracking stations T1, T2, and T3 are positioned behind the magnet, are divided in two parts that employ different technologies: the Inner Tracker (IT), in the region closer to the beam pipe, and the Outer Tracker (OT), covering the outer part of the detector.

The IT [97] features silicon micro-strip sensors in four cross-shaped detection layers (see Figure 2.9, left); in each layer, the top and bottom modules contain one row of seven sensors, while the left and right modules contain two rows of seven sensors. The sensors have a pitch of 198 µm, providing a single-hit resolution of about 50 µm with an efficiency above 99.8%, as shown in Figure 2.10. The overall dimensions of the IT sub-detector are

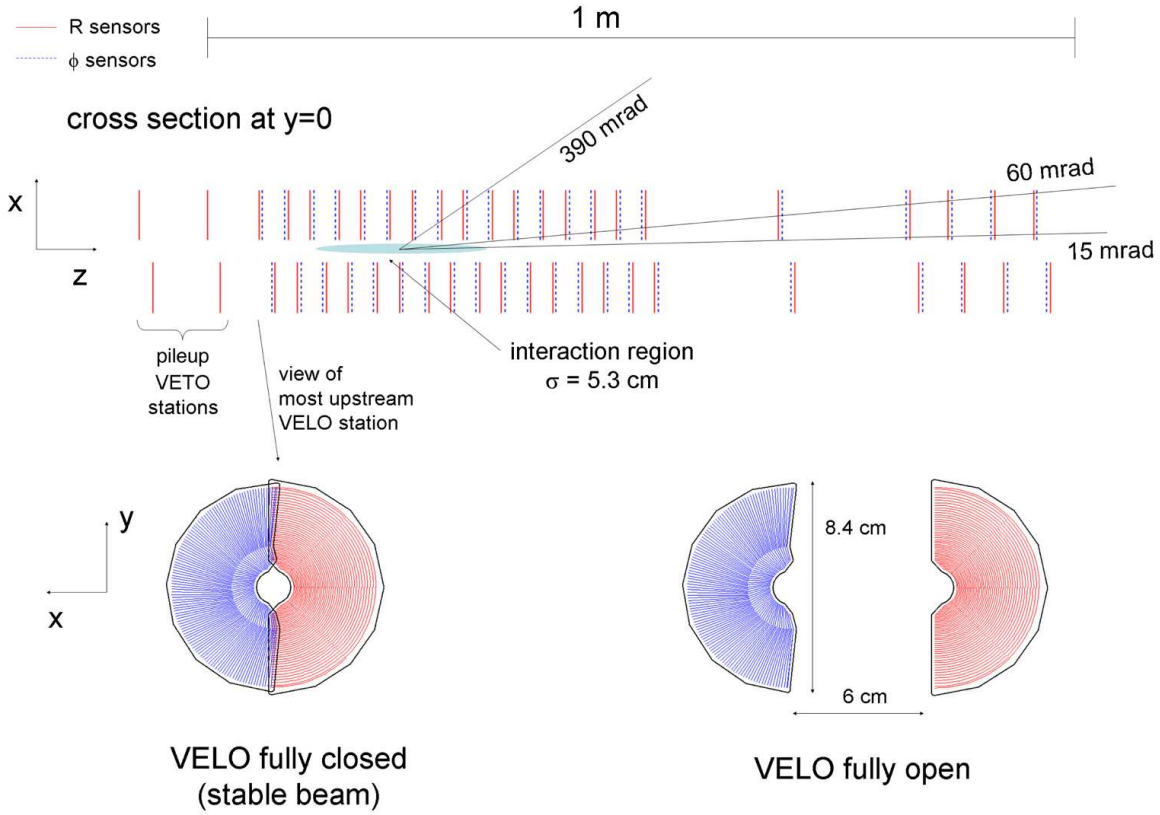

Figure 2.6: Cross-section in the (x, z) plane of the VELO silicon sensors, at $y = 0$, with the detector in the fully closed position. The front face of the first modules is illustrated in both the closed and open positions. The two pile-up veto stations are located upstream of the VELO sensors.

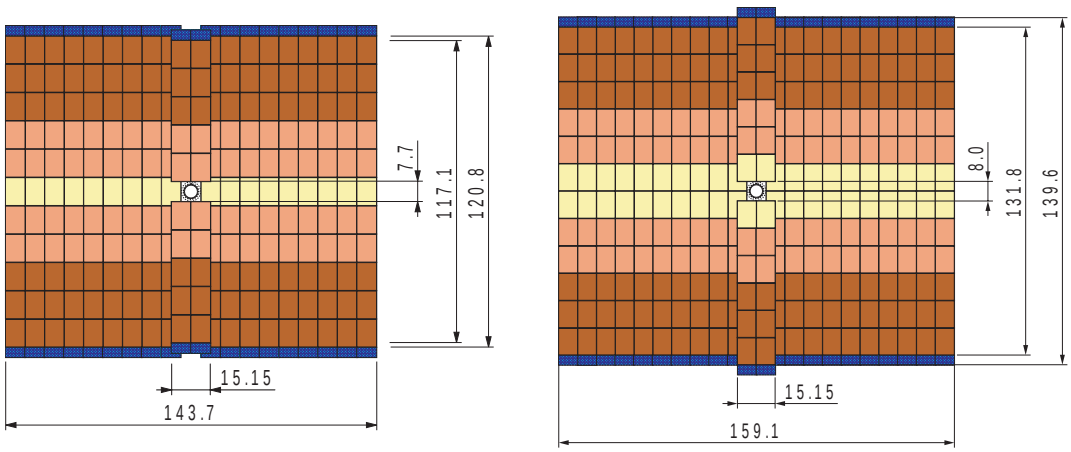

Figure 2.7: Layout of the x-layers of the TTa (left) and TTb (right) stations. Dimensions are reported in centimeters.

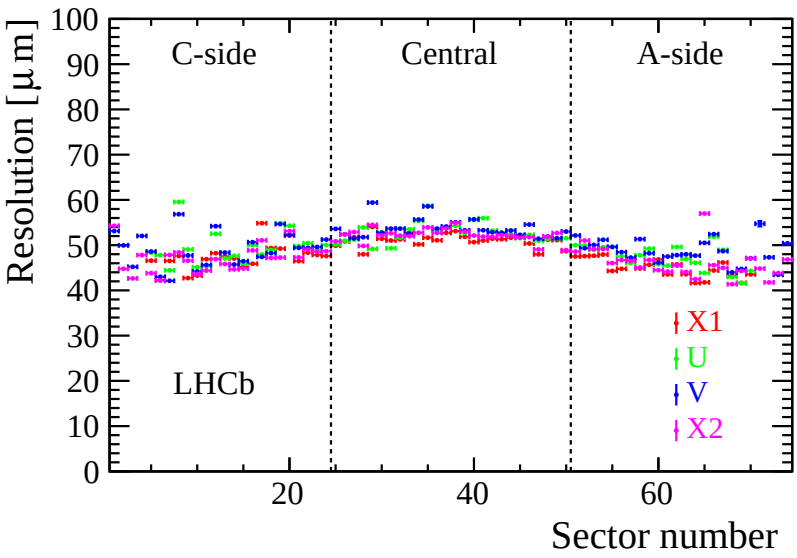

Figure 2.8: Hit resolution of the TT measured with 2012 data. The resolution of the four layers x–u–v–$x$is shown separately. The "Sector number" label on the horizontal axis approximately corresponds to the x direction.

approximately 1.2 m in the bending (x) plane and 40 cm in the vertical (y) plane.

The OT [98] comprises three stations made of four layers arranged in a x–u–v–$x$layout; the total active area of the detector is $5971 \times 4850\,\text{mm}^2$ (Figure 2.9, right). Each layer is made of straw tubes, arranged in two staggered layers of 64 tubes in total, each with a diameter of 4.9 mm and filled with a 70/30% mixture of Argon and CO_2 to achieve a drift time below 50 ns and a drift-coordinate resolution of 205 µm. The hit efficiency is very high for tracks passing close to the center of the straw, while it falls rapidly when a particle passes closer to the edge of the tube due to the longer drift time, as it can be seen in Figure 2.11.

2.3.4 The magnet

The magnetic field in the LHCb experiment is generated by a warm dipole magnet [99] positioned between the (TT) and the first tracking station T1. The magnetic field allows the momentum of charged particles to be inferred by measuring the bending angle of tracks before and after the magnet.

The magnet is composed of a 1450 t yoke and two inclined coils weighting 54 t, with a geometry adapted to the detector acceptance (see Figure 2.12); the maximum magnetic field strength in the y direction is 1 T, and the field integral $\int B\,\mathrm{d}z$ over the total detector length is approximately 4 T m. The polarity of the magnet is regularly flipped during data acquisition to compensate any left-right asymmetry introduced by the detector, as positively and negatively charged particles experience opposite deflections. The nominal current in the coils is 5.85 kA with a peak capacity of 6.6 kA, yielding an electric power dissipation of 4.2 MW.

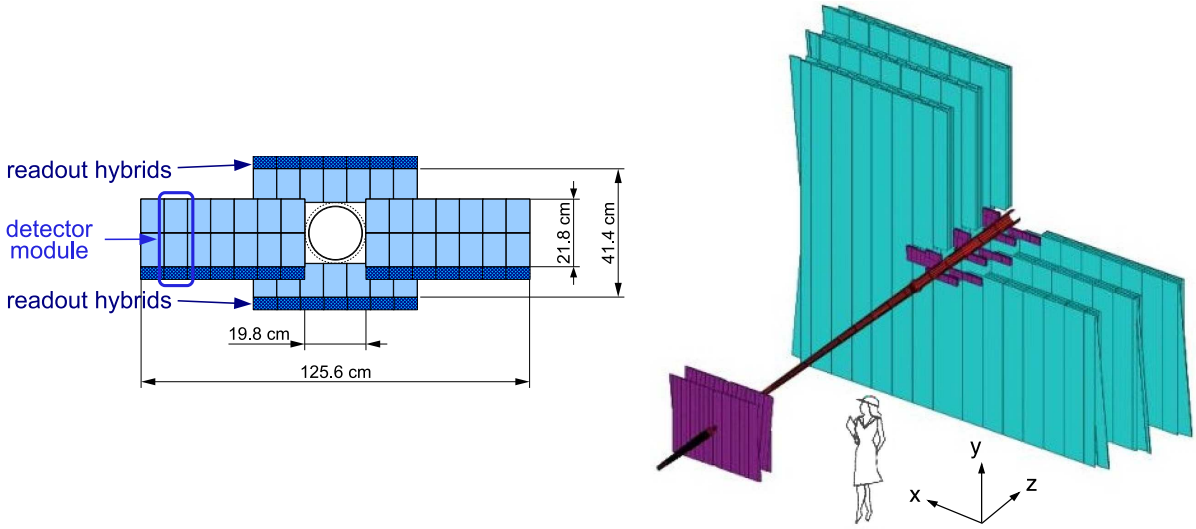

Figure 2.9: Left: view of one layer of the IT detector, with the four detector modules in a cross shape aroud the beam pipe. Right: arrangement of the OT modules and layers in cyan, with the IT stations in purple.

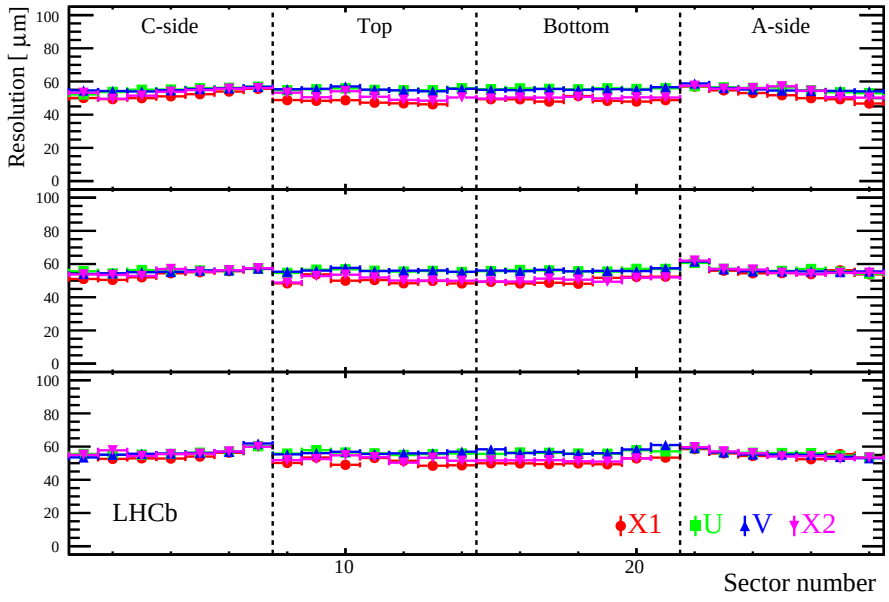

Figure 2.10: Hit resolution in the four modules of IT1 (bottom), IT2 (middle), and IT3 (top). The resolution in shown separately for each of the four x–u–v–$x$layers. The "Sector number" label approximately corresponds to the x direction.

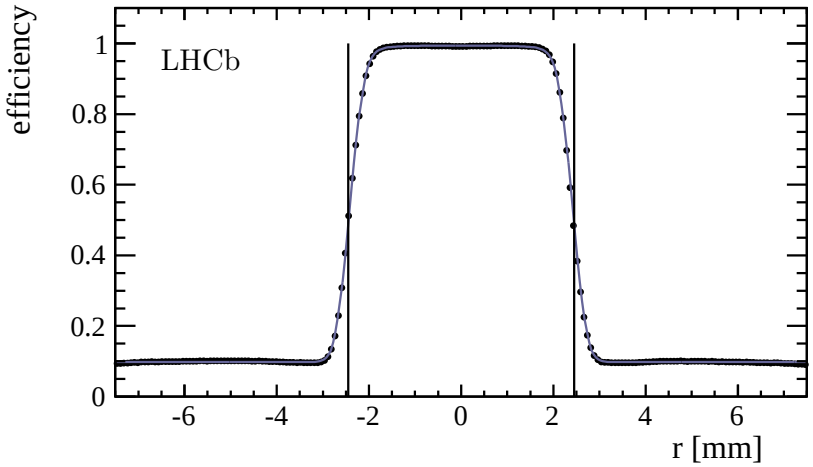

Figure 2.11: Hit efficiency in the OT modules as a function of the track distance from the center of the straw.

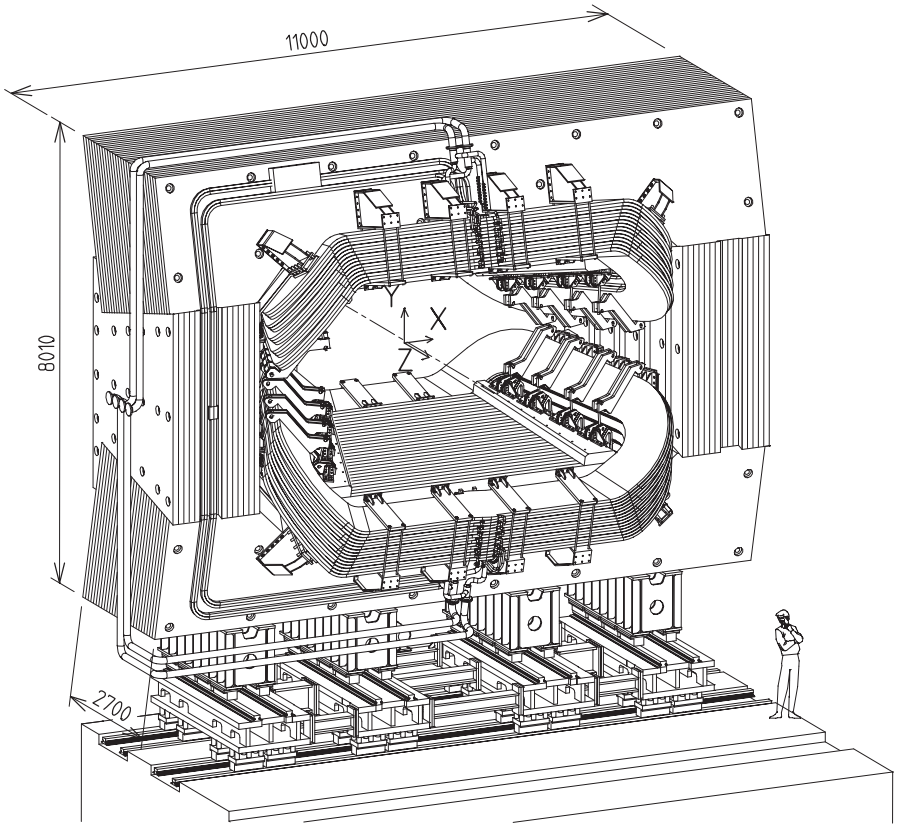

Figure 2.12: Perspective view of the LHCb dipole magnet.

2.3.5 Track reconstruction

Track reconstruction [100] involves reconstructing the trajectories of charged particles as they traverse the tracking system, which includes the VELO, TT, IT, and OT detectors. Different track types are defined based on their paths through the spectrometer, each serving specific purposes:

- Long tracks traverse the entire system, providing precise momentum estimates crucial for physics analyses.
- Upstream tracks pass through VELO and TT, often utilized to understand backgrounds in RICH1.
- Downstream tracks pass through TT and T stations, aiding in the reconstruction of long-lived particles such as K^0_S and Λ decaying outside VELO's reach.
- VELO tracks traverse only the VELO, and are useful for primary vertex reconstruction.
- T tracks pass only through T stations, mainly due to particles produced in secondary interactions, and are valuable for RICH2 data analysis.

Figure 2.13 illustrates tracks reconstructed in a typical event. The long track reconstruction starts in the VELO, with two complementary algorithms integrating information from downstream tracking stations: the track matching algorithm combines VELO tracks with segments from T stations, while the forward tracking algorithm utilizes VELO and a single T station hit to determine particle momentum and trajectory. Downstream and upstream tracks are found by extrapolating T and VELO tracks, respectively, incorporating hits from TT. The final step involves fitting the tracks using a Kalman filter [101], considering factors like multiple scattering and energy loss. The quality of the reconstructed track is assessed by the χ^2 per degree of freedom of the fit, and state vectors at specific z-positions represent the reconstructed tracks. During Run 1 and Run 2, the average tracking efficiency was above 96%, going slighlty below that only in high-multiplicity events (more than 200 tracks). The relative momentum resolution depends on the particle's momentum, going from 0.5% for low-momentum tracks to about 1% for $p > 150$ GeV/c. The performance of the tracking algorithm can be seen in Fig. 2.14.

2.4 The LHCb particle identification system

Particle identification is a fundamental task of the LHCb detector, to ensure that the final states of b- and c-hadron decays are correctly reconstructed. The indentification of pions, kaons, protons, electrons, photons, and muons is done using informations from the RICH system, the calorimenters, and the muon stations. We will now describe in detail each one of these sub-dectectors.

2.4.1 The RICH detectors

The two Ring Imaging Cherenkov (RICH) detectors [102], RICH1 located between the VELO and the TT, and RICH2 between T3 and the first station of the muon detector,

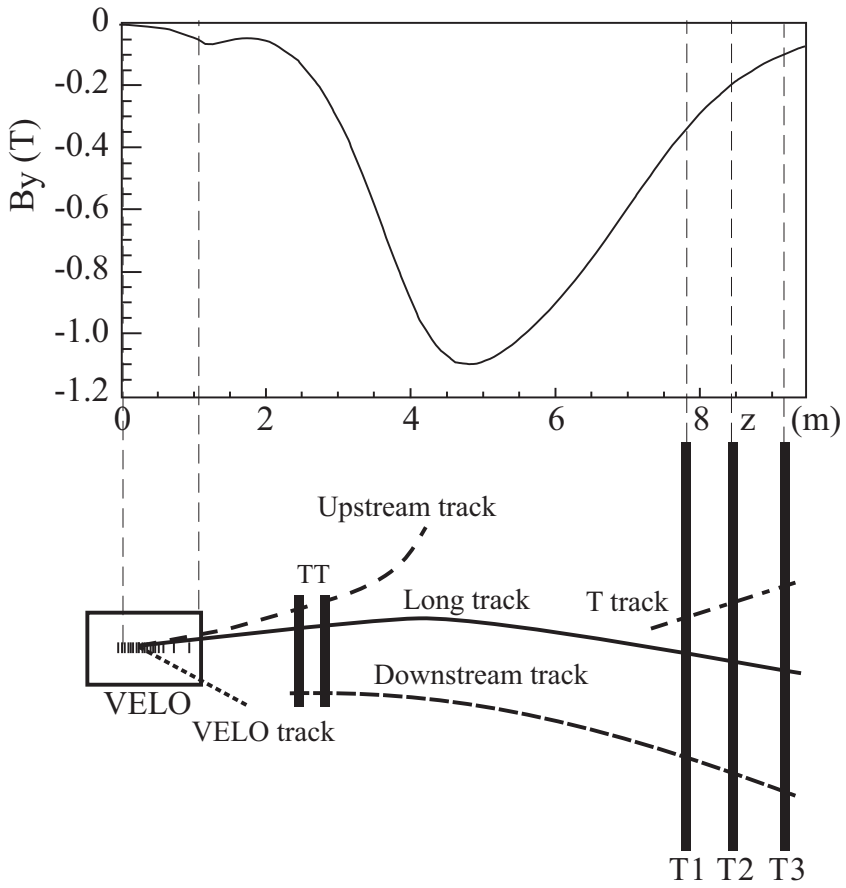

Figure 2.13: Types of reconstructible tracks at LHCb. In the top part s plot of the magnetic field strentgh as a function of the distance from the interaction region is shown.

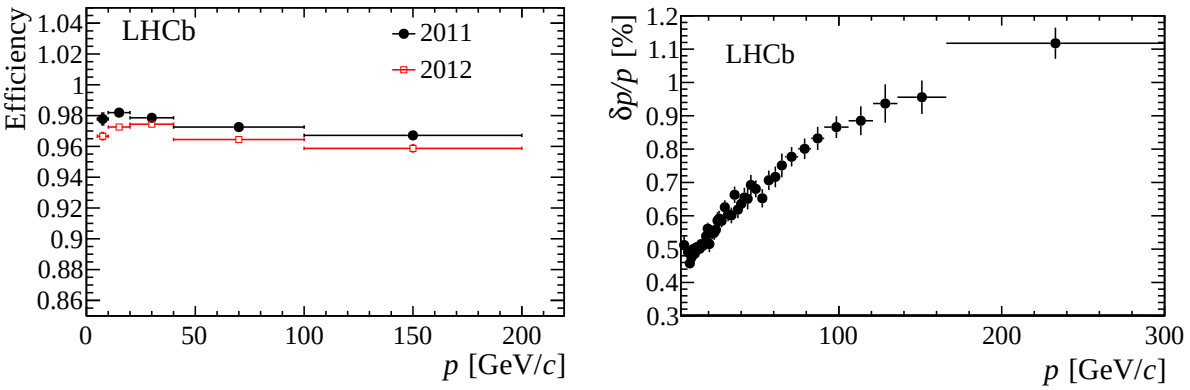

Figure 2.14: Tracking efficiency (left) and momentum resolution (right) as a function of particle momentum.

are dedicated to the discrimination of charged light hadrons. Their structure is presented in Figure 2.15. Their working principle is based on Cherenkov radiation, which is the emission of light from a charged particle travelling through a medium with a velocity greater than the speed of light in that medium. Under this condition, Cherenkov photons are emitted on a conical wave-front of aperture angle θ_c, given by

$$\cos\theta_c = \frac{1}{\beta n}, \tag{2.1}$$

where n is the refractive index of the medium and β the velocity of the particles as a fraction of the speed of light c. By measuring the Cherenkov angle and combining it with the momentum of the particle it is then possible to infer the particle mass and thus identify it.

In order to cover complementary momentum ranges, the two RICH use different radiators. RICH1 uses two radiators: a 5 cm layer of aerogel ($n = 1.03$), suitable for low momentum particles, and fluorobutane (C_4F_{10}, $n = 1.0014$), which fills the remaining 85 cm of the tank and is needed for higher momentum particles; these two radiators provide $\pi - K$ discrimination up to 50 GeV/c. The layer of aerogel has been removed between Run 1 and Run 2. For RICH2, tetrafluoromethane (CF_4, $n = 1.0005$) has been chosen as radiator, allowing to cover the high momentum range up to 100 GeV/c (and slighlty above), with a limited angular acceptance of [15, 120] mrad in the horizontal plane and up to 100 mrad in the vertical plane.

Both RICH1 and RICH2 reflect the Cherenkov light with a combination of spherical and flat mirrors, projecting the photons onto a grid of Hybrid Photon Detectors (HPDs), located out of the spectrometer acceptance in a region shielded against the residual magnetic field. The HPDs are vacuum photon detectors where a photon impacting on a photocathode releases photoelectrons, which are then are accelerated and multiplied by a voltage of typically 10 to 20 kV onto a silicon detector.

Figure 2.16 shows the dependence of the Cherenkov angle on the momentum for different particles. The angular resolution of the two detectors was measured to be (1.618 ± 0.002) mrad for RICH1 and (0.68 ± 0.02) mrad for RICH2. Due to resolution effects, the presence of photons coming from other particles in the event, and the complexity of the problem, a dedicated algorithm is adopted to obtain optimal performances [103]. The first step is assigning the pion mass hypothesis to all reconstructed tracks in the event, since the most abundant particles produced in pp collisions are pions. Then, a global event likelihood is computed from the distributions of the photon yield of each track under this first hypothesis. After that, the particle hypotheses are changed, one track at a time, to one of the other possible choices (K, p) while leaving all other particles unchanged. The particle type which yields the largest increase in the global likelihood is identified, and the corresponding mass hypothesis is set for that track. The procedure is repeated for all tracks, until no significant change in the event likelihood is observed.

After the optimal configuration of particle types has been found, PID variables are computed to be stored and used during the analysis to select final-state particles. For each particle in an event, the following DLL variables are defined:

$$\mathrm{DLL}_{K\pi}(h) = \log\mathcal{L}_K(h) - \log\mathcal{L}_\pi(h), \tag{2.2}$$

$$\mathrm{DLL}_{p\pi}(h) = \log\mathcal{L}_p(h) - \log\mathcal{L}_\pi(h), \tag{2.3}$$

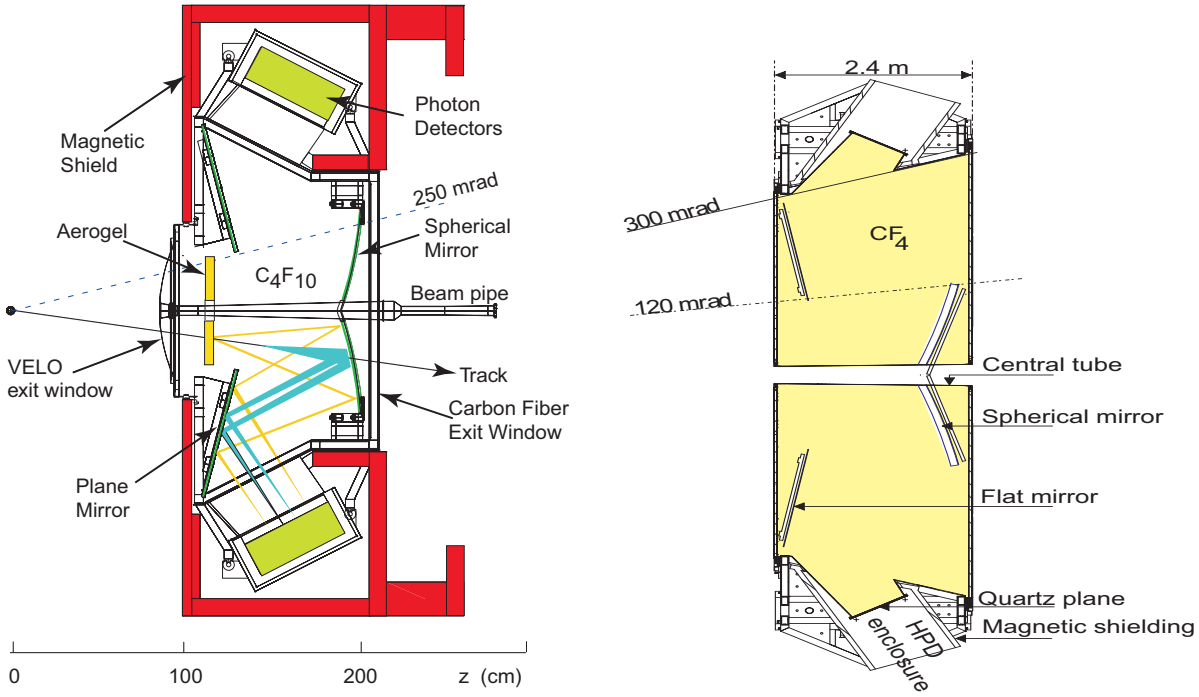

Figure 2.15: Side view of the (left) RICH1 detector and (right) RICH2 detector.

meaning that a negative value of the DLL represent a higher chance of the particle being a pion, and either a kaon or a proton otherwise. The performances of the DLL variables for pion, kaons and protons are reported in Figure 2.17.

2.4.2 The calorimeters

The LHCb calorimeter system [104] identifies hadrons, electrons, and photons while providing information on their transverse energy or momentum fast enough to be used at the first stage of the trigger. The calorimeter system is placed between the first and the second muon station and is composed of a scintillating pad detector (SPD), a thin lead converter, a preshower (PS), an electromagnetic calorimeter (ECAL) and an hadronic calorimeter (HCAL), each one of them made of alternating layers of scintillating and absorbing material. Upon hitting the absorbing material, particles lose energy and produce electromagnetic or hadronic cascades which excite the molecules of the scintillating material, that in turn emit electromagnetic radiation in an amount proportional to the initial energy of the impinging particle. The light is carried away by optical fibres with wavelength-shifting (WLS) properties to photon detectors, respectively multi-anode photomultipliers (MAPMTs) in the SPD and PS and phototubes in ECAL and HCAL. The WLS is needed to shift the wavelength of the radiation such that the light reaching the PMT is at the optimal wavelength for collection. Since the hit density varies by two orders of magnitude over the surface of the calorimeters, they are split in cells with a variable lateral segmentation, as shown in Figure 2.18; ECAL and SPD/PS are divided into three different sections, HCAL in two.

The SPD/PS system consists of a 15 mm thick (2.5 radiation lengths, X_0) lead con-

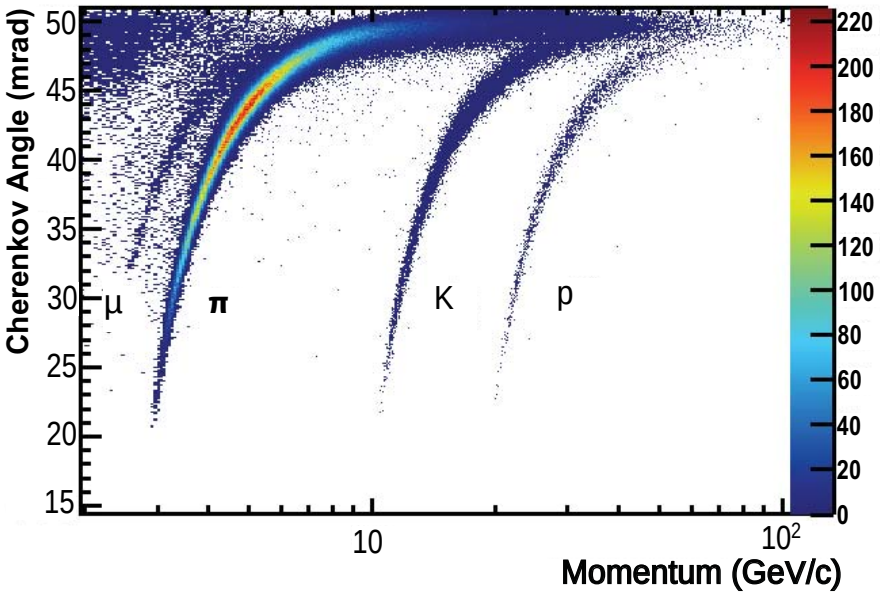

Figure 2.16: Reconstructed Cherenkov angle as a function of particle momentum in RICH1 for μ, π, K and p. The data was collected using only isolated tracks, for which the photon rings can be clearly reconstructed.

verter sandwiched between two planes of 15 mm polystyrene-based scintillating pads. The SPD allows charged and neutral particles to be distinguished, since only charged particles produce light in the scintillator material, providing π^0 background rejection. The discrimination of charged pions from electrons is enhanced by the PS, where the latter release a larger amount of energy with respect of the former. The sensitive area of these two detectors is $7.6 \times 6.2\,\mathrm{m}^2$.

In the beam direction, the ECAL is made of 66 modules of 2 mm of lead followed by 4 mm thick scintillator tiles, amounting to a total of 25 X_0, sufficient for a full containment of the electromagnetic showers. The energy resolution of the ECAL is given by (see Figure 2.19)

$$\frac{\sigma(E)}{E} = \frac{(13.5 \pm 0.7)\%}{\sqrt{E}} \oplus \frac{(0.320 \pm 0.030)}{E} \oplus (5.2 \pm 0.1)\% \tag{2.4}$$

where E is expressed in GeV and $\oplus$ indicates a sum in quadrature.

Since the trigger requirements for HCAL are not as stringent as those for ECAL, the total thickness of the HCAL is 1.2 m due to space limitations, corresponding to 5.6 interaction lengths; The HCAL alternates 4 mm scintillator planes to 16 mm iron plates, with the special feature of having the scintillating tiles installed parallel to the beam, to enhance light collection. The energy resolution of the HCAL (see Figure 2.19) is

$$\frac{\sigma(E)}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%, \tag{2.5}$$

with E expressed in GeV.

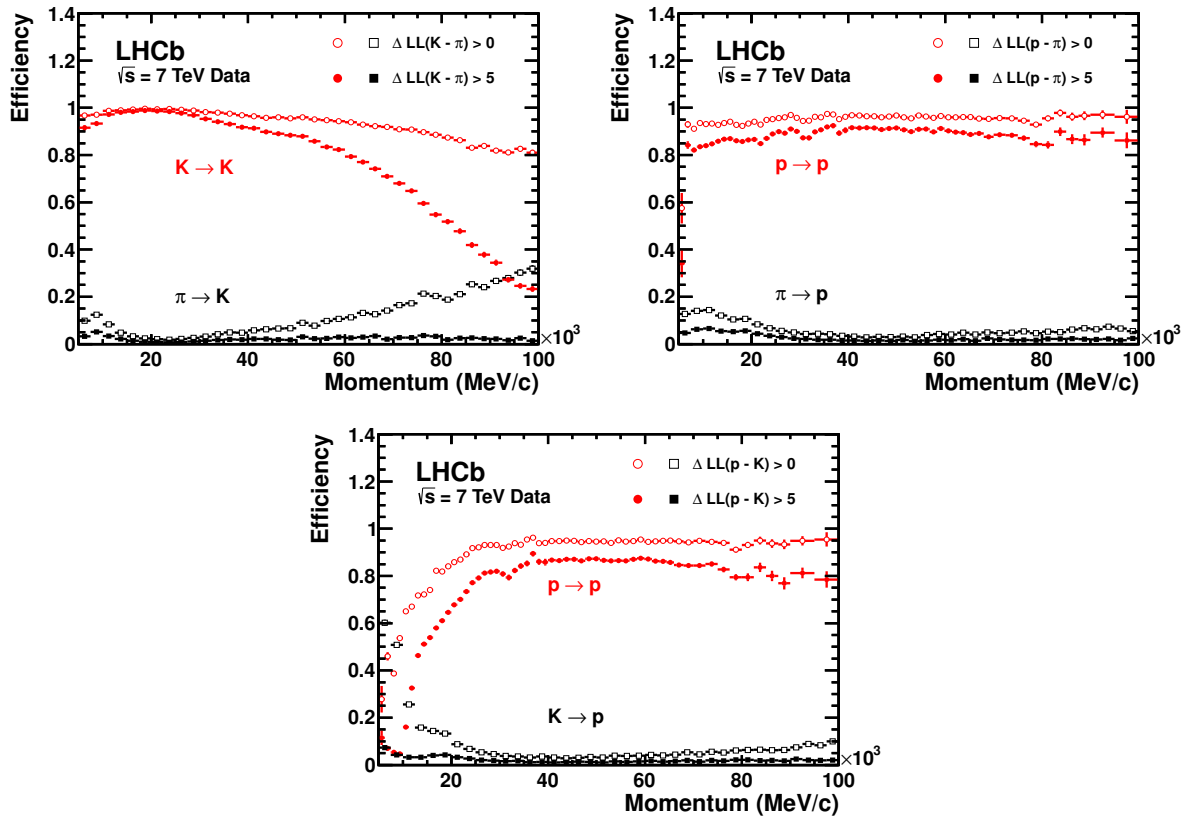

Figure 2.17: Efficiencies of correctly and incorrectly identifying hadrons from the RICH detectors while requiring different values of the DLL variables (described in the text). Top left: K identification and $\pi - K$ misidentification. Top right: p identification and $\pi - p$ misidentification. Bottom: p identification and $K - p$ misidentification.

With the use of the informations coming from the SPD and PS, the calorimeter system is able to provide γ/π^0 discrimination, as shown in Figure 2.20. In addition, similarly to what happens for RICH, a DLL variable is defined in order to discriminate between hadrons and electrons, which gets information from the ECAL, the HCAL and the PS and correlates it with the momentum of the track, obtaining a likelihood for the electron and hadron hypothesis:

$$\mathrm{DLL}_{eh}^{\mathrm{CALO}} = \mathrm{DLL}_{eh}^{\mathrm{ECAL}} + \mathrm{DLL}_{eh}^{\mathrm{HCAL}} + \mathrm{DLL}_{eh}^{\mathrm{PS}}. \tag{2.6}$$

The performance of $\mathrm{DLL}_{eh}^{\mathrm{CALO}}$ is illustrated in Figure 2.21; for example, by requiring $\mathrm{DLL}_{eh}^{\mathrm{CALO}} > 2$, an electron efficiency of 90% can be achieved while keeping the $e \to h$ mis-identification rate below 3%.

2.4.3 The muon system

The muon identification system of the LHCb experiment [105] is a crucial component designed to accurately identify and track muons, playing a pivotal role in the L0 trigger of the experiment. Comprising five stations (M1-M5) of rectangular shape, this system employs advanced detectors and strategic placement along the beam axis to ensure comprehensive coverage and effective identification of muons.

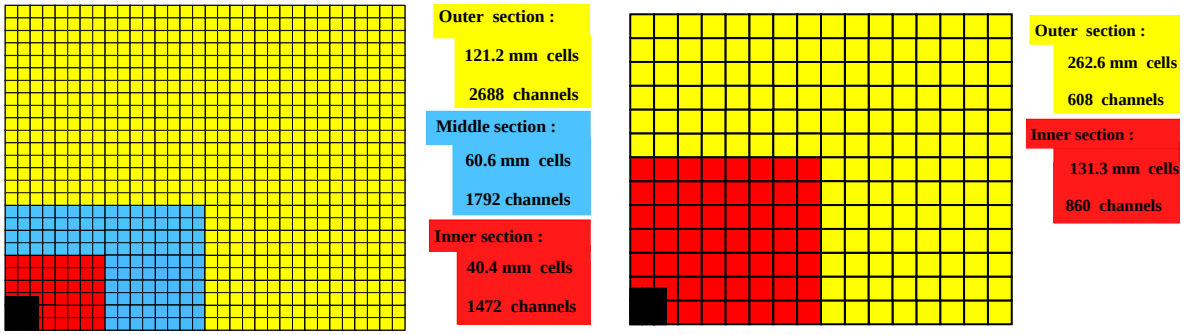

Figure 2.18: Lateral segmentation of one quarter of the (left) SPD, PS, and ECAL and (right) HCAL. The cell dimension must be reduced by about 1.5% for SPD and PS.

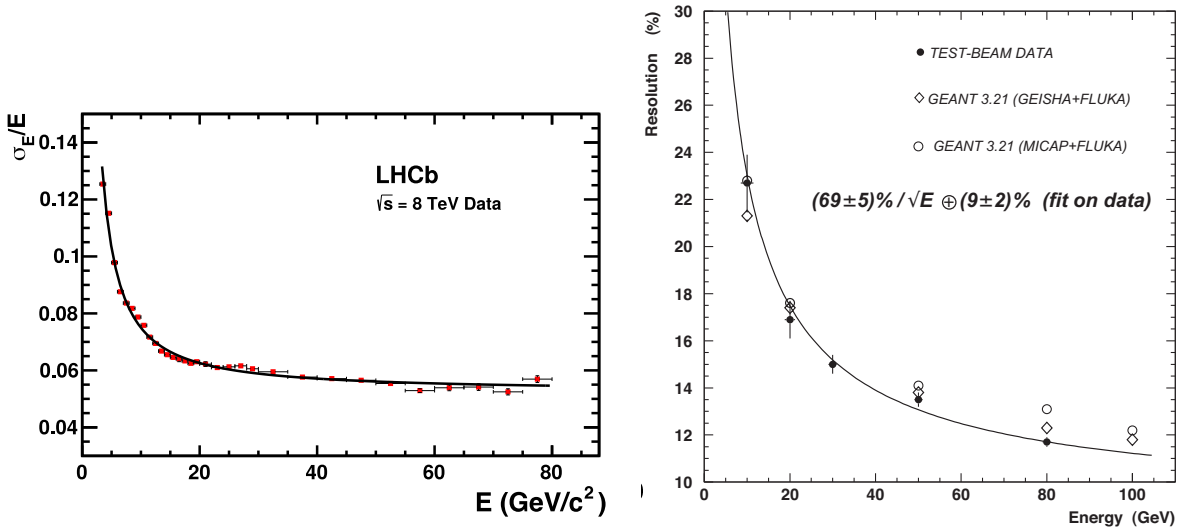

Figure 2.19: Energy resolution of (left) ECAL and (right) HCAL.

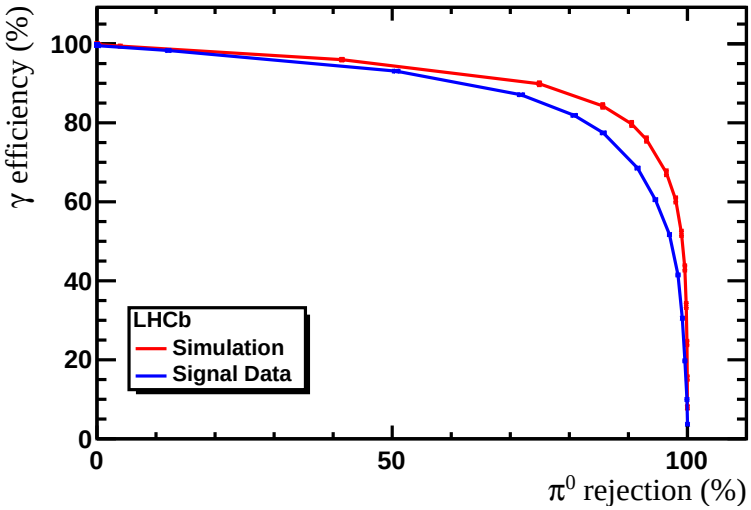

Figure 2.20: Photon identification efficiency as a function of π^0 rejection efficiency measured from simulation (red) and data (blue).

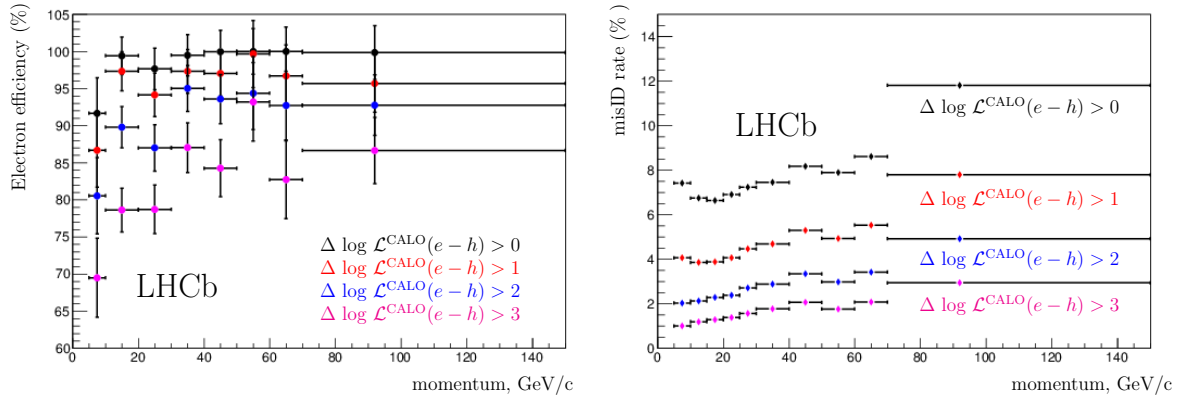

Figure 2.21: Electron identification rate (left) and $e \to h$ mis-identification (right) as a function of momentum, for different requirements on $\mathrm{DLL}_{eh}^{\rm CALO}$.

The full muon system spans a total area of 435 m^2 and consists of 1380 chambers. Stations M2 to M5 are strategically positioned downstream of the calorimeters and are interspersed with 80 cm thick iron absorbers. This arrangement serves the dual purpose of selecting penetrating muons and enhancing the efficiency of muon identification. The minimum momentum required for a muon to traverse all five stations is approximately 6 GeV/c, considering a total absorber thickness of approximately 20 interaction lengths, including the calorimeters.

Station M1, located in front of the calorimeters, plays a crucial role in improving the transverse momentum measurement in the first level hardware trigger. The geometry of the five stations is projective, meaning that the transverse dimensions of each station scale with the distance from the interaction point. This projective geometry ensures optimal coverage and sensitivity across the entire angular spectrum.

The detection elements primarily consist of Multi-Wire Proportional Chambers (MWPC), with the exception of the highest rate region in M1, where triple Gas Electron Multiplier (GEM) detectors are employed. The choice of detectors is carefully calibrated to meet the specific demands of different regions within the muon system.

The detectors provide space point measurements of the muon tracks, supplying a binary yes/no decision to both the trigger processor and the Data Acquisition (DAQ) system. The system achieves this by partitioning the detectors into rectangular logical pads, and the dimensions of these pads define the (x, y) resolution in the plane perpendicular to the beam axis.

Each station is further divided into four regions (R1 to R4) with increasing distance from the beam axis. The linear dimensions of these regions and their segmentation follow a specific ratio (1:2:4:8), ensuring that the channel occupancies are comparable across each region of a given station (see Fig. 2.22). This design approach optimizes the performance and reliability of the muon identification system.

The majority of the muon chambers are equipped with MWPC using $\mathrm{Ar/CO_2/CF_4}$ (40/55/5 %) as the gas mixture. Only the inner part of the first station (M1) is instrumented with triple-GEM detectors filled with $\mathrm{Ar/CO_2/CF_4}$ (45/15/40 %).

Each muon station is designed to operate with an efficiency above 99% in a 20 ns time

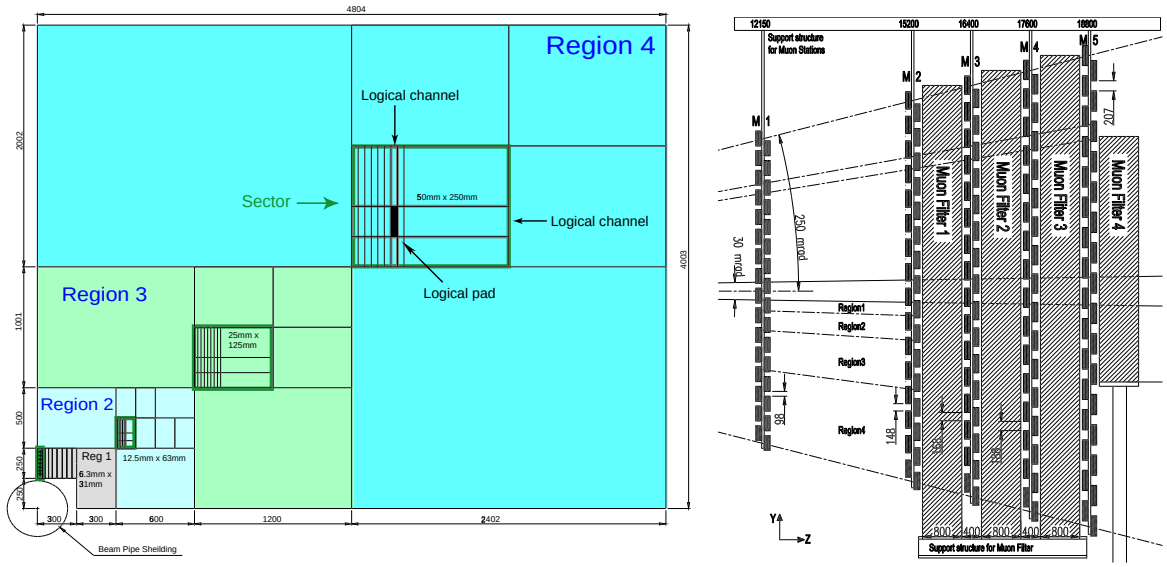

Figure 2.22: Left: front view of one stations of the muon system, where the segmentation in four regions can be seen. Right: side view of the muon system in the (y, z) plane.

window, maintaining a noise rate below 1 kHz per physical channel. These performance criteria were successfully achieved during operation, as documented in [106]. Overall, the LHCb muon identification system stands as a sophisticated and reliable tool for the precise identification and tracking of muons in the demanding environment of high-energy particle collisions.

The muon identification procedure [107] is divided in three steps. The first step, which is based on the penetration of a candidate into the muon system, produces a binary variable called `IsMuon`; the higher the momentum of the track, the higher the number of stations traversed that are required to provide a positive decision. The second step is the computation of a likelihodd for the muon and non-muon hypothesis, taking into account the pattern of the hits around the tracks extrapolated from the tracking system to the muon chambers; the logarithm of the ratio between the muon and non-muon hypotheses, called, `muDLL`, is used as discriminating variable. Finally, a combination of all the informations from the RICH, the calorimenters and the muon stations is combined to provide DLL variables for electron, muon, kaons, and protons, as they were defined in Eqs. (2.2) and (2.3). The performances of `IsMuon` and `muDLL` are illustrated in Fig. 2.23.

2.5 The LHCb trigger

The trigger algorithm of the LHCb experiment is a sophisticated system designed to efficiently select and record relevant events for physics analysis, operating within the challenging environment of the LHC. The LHCb experiment operated, during Run 1 and Run 2, at an average luminosity of $2 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, significantly lower than the LHC's maximum design luminosity. This deliberate choice not only makes radiation damage more manageable but also ensures that the number of interactions per crossing is dominated by single interactions, simplifying triggering and reconstruction processes with low channel

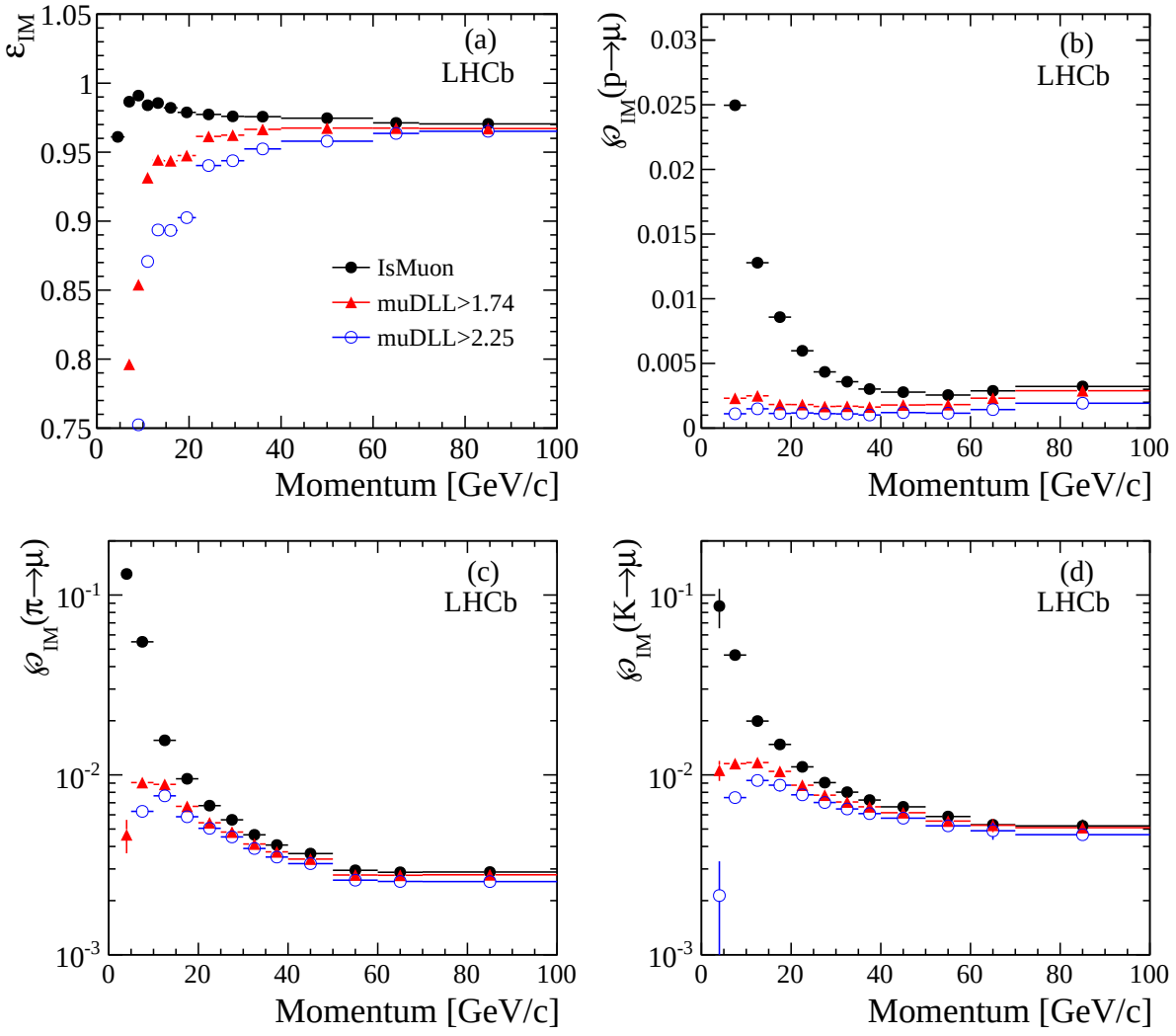

Figure 2.23: Muon identification efficiency (a) and misidentification efficiencies for protons (b), pions (c), and kaons (d) as a function of the track momentum. The efficiencies after requiring `IsMuon` and with the addition of `muDLL` > 1.74 and `muDLL` > 2.25 are shown.

occupancy.

To achieve the necessary reduction in event rate from the initial frequency of about 10 MHz to a few hundred Hz, the LHCb trigger [108–110] employs a three-level hierarchical system: Level-0 (L0), High Level Trigger 1 (HLT1), and High Level Trigger 2 (HLT2). L0, implemented in custom electronics, operates on the hardware level and relies on information from the calorimeter and muon systems to make rapid decisions, reducing the event rate to below 1 MHz, allowing the entire detector to be read out. The HLT1 runs on a cluster of CPUs, called the Event Filter Farm (EFF), to perform a partial event reconstruction and select candidates, reducing the input rate to around 40 kHz after the first stage. After that, HLT2 runs a more complete reconstruction and applies a final selection before saving the event to storage, with a final output rate of 2-5 kHz. We will now describe in more detail each stage of the trigger.

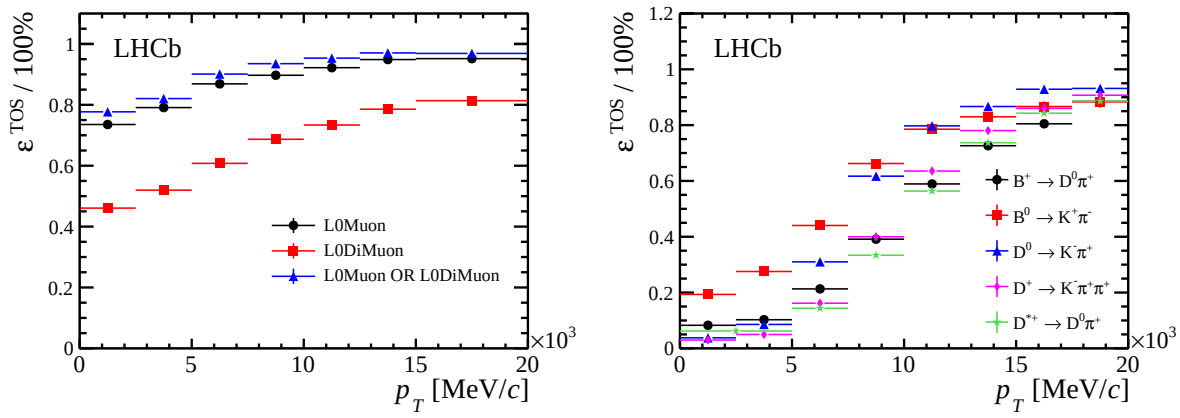

Figure 2.24: Performance of the Level-0 (left) muon trigger and (right) hadron trigger as a function of transverse momentum.

2.5.1 Level-0 trigger

The L0 trigger consists of three independent units: the Calorimeter trigger, the Muon trigger, and the Pile-Up trigger. Operating synchronously with the 40 MHz bunch crossing rate of the LHC, they help identifying the distinctive features of b-hadron decays, *i.e.* leptons, hadrons, or photons with a high transverse momentum.

Calorimeter trigger High-E_{T} deposits are searched for in the calorimeter system by inspecting clusters formed by 2×2 cells, large enough to contain most of the shower's energy but small enough to avoid overlap between different particles. The total E_{T} deposited in the HCAL and the SPD multiplicity is computed at this stage; the latter is also used to reject event with no visible interactions. Different particle types (e, γ, h) can be identified using informations from the SPD, PS, and the calorimeters; this way, three different types of candidates are built: `L0Hadron` is the cluster with the highest E_{T} in the HCAL; `L0Photon` is the cluster wit the highest E_{T} in the ECAL with corresponding hits in the PS and no hit in the SPD; `L0Electron` has the same requirement as `L0Photon`, with the addition of hits in the SPD as well. The E_{T} deposit in each cluster is computed, and the trigger fires only if there is at least one candidate above a certain threshold in the event.

Muon trigger High-p_{T} tracks in the muon stations are searched for, by subdividing the detector in 192 towers pointing to the interaction region. The two tracks with the highest p_{T} in each quadrant of the system are identified, and a selection is applied based on the highest p_{T} (`L0Muon`) or on the product $p_{\mathrm{T}}^{\text{largest}} p_{\mathrm{T}}^{\text{2nd largest}}$ (`L0DiMuon`).

Pile-Up trigger This part of the trigger employs the upstream stations of the VELO (see Fig. 2.6) to veto *empty-beam* (eb) bunch crossings, where no particles are produced in the detector acceptance, and to get a rough estimate of the event multiplicity, both for triggering and luminosity measurement purposes [111].

The performances of the L0 muon and hadron triggers are shown in Fig. 2.24

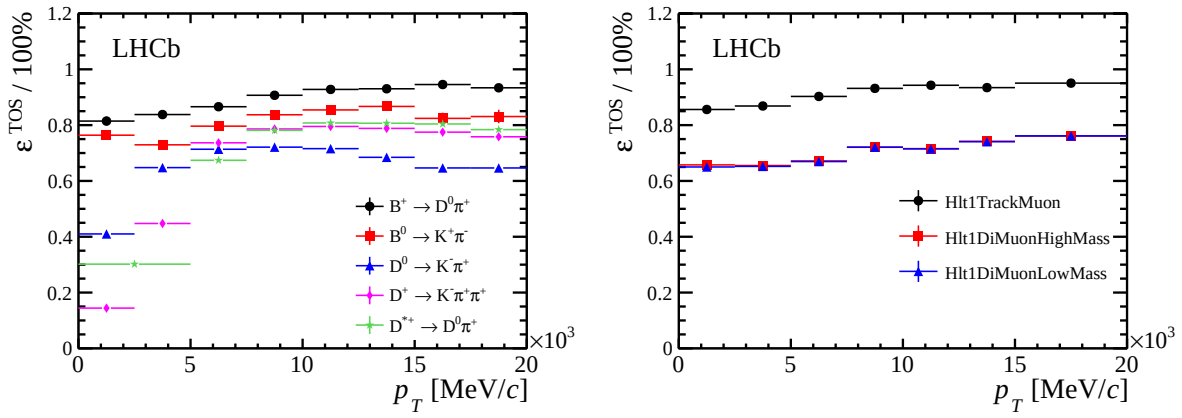

Figure 2.25: Left: HLT1 efficiencies of selected D and B decays. Right: HLT1 muon trigger efficiency from $B^+ \to J/\psi K^+$ decays as a function of B^+ p_{T}.

2.5.2 High Level Trigger 1

When events are accepted by the L0 trigger, they are transmitted to the EFF, where they are further scrutinised by the High Level Trigger, which is designed to efficiently process and reduce the event rate, ensuring that only the most relevant events are stored for subsequent analysis.

The reconstruction algorithms used in the HLT are similar to those employed offline, with some simplifications necessary to meet the stringent time constraints. For example, the VELO reconstruction algorithm, which performs a full 3D pattern recognition, is adapted to run on all events entering the HLT, meaning that certain offline-specific features, such as a second pass on unused hits to enhance efficiency for tracks pointing away from the beam-line, are omitted due to CPU constraints. The HLT1 employs various criteria to select VELO tracks, limiting their number and applying quality cuts based on the impact parameter (IP) and expected number of hits. In cases where muon identification is required, a fast muon identification algorithm is also applied.

The HLT1 trigger algortihms (also called *lines*) cover a range of physics channels, including inclusive beauty and charm triggers, single and dimuon triggers, and specific triggers for high transverse momentum electrons, di-protons, displaced vertices, or high ET jets. Each trigger line consists of a sequence of reconstruction algorithms and selections, returning an accept or reject decision. Additionally, the HLT includes lines dedicated to luminosity measurements, physics triggers with looser cuts, low multiplicity events, large transverse momentum jets, and various monitoring and calibration purposes. The performances of a selection of HLT1 lines as a function of particle kinematics are shown in Fig. 2.25.

Of particular interest for the analysis presented in this document are the lines `Hlt1TrackAllL0` for Run 1 and `Hlt1TrackMVA` for Run 2: they are general-purpose lines for selecting hadrons that are significantly displaced from the PV. The requirements of these two lines are reported in Table 2.1

Table 2.1: Requirements of the (top) `Hlt1TrackAllL0` and (bottom) `Hlt1TrackMVA` trigger lines, running respectively during Run 1 and Run 2. The value of C varied between 1.1 and 1.2 throughout the years.

`Hlt1TrackAllL0`		
Variable		Cut
Track IP	$>$	0.1 mm
VELO hits/track	$>$	9
Missed VELO hits/track	$<$	3
OT+IT hits/track	$>$	16
Track χ^2_{IP}	$>$	16
Track p_{T}	$>$	1.8 GeV/c
Track p	$>$	10 GeV/c
Track χ^2/ndf	$<$	2.5
`Hlt1TrackMVA`		
Variable		Cut
p_{T}	$\in$	$[1, 25]$ GeV/c
$\min(\chi^2_{\text{IP}})$	$>$	7.4
Track χ^2/ndf	$<$	2.5
`GhostProb`	$<$	0.2 (added in 2016)
$\log\left(\frac{\chi^2_{\text{IP}}}{7.4}\right)$	$>$	$\frac{1}{(p_{\text{T}}-1)^2} + \frac{C}{25}(25 - p_{\text{T}})$ (p_{T} in GeV/c)

2.5.3 High Level Trigger 2

The second level of the High Level Trigger follows the initial reduction of events rate achieved by HLT1. Operating on an event rate of about 43 kHz, HLT2 performs more detailed reconstruction, focusing on specific physics goals of the LHCb experiment. In the HLT2 stage, forward tracking of all (VELO) tracks is feasible given the reduced event rate, but while offline reconstruction employs two tracking algorithms, HLT2 simplifies by using an algorithm based only on VELO seeds. This simplification, however, leads to a lower tracking efficiency of 1-2% per track compared to offline reconstruction. The search is further limited to tracks with momenta greater than 5 GeV/c and transverse momenta greater than 0.5 GeV/c, effectively reducing the search windows to manage processing time. Muon identification in HLT2 involves applying the offline muon identification algorithm to all tracks from the forward tracking. Tracks are also associated with ECAL clusters to identify electrons. During Run 2 PID requirements for hadrons, determined from the information provided by the RICH detectors, were also exploited in the HLT2 algorithms.

The output rate of HLT2, set at 3 kHz, encompasses various trigger lines, with a significant portion dedicated to "topological" trigger lines designed to inclusively capture b- and c-hadron decays, targeting hadrons with at least two charged particles in the final state and a displaced decay vertex. The topological trigger lines make decisions based on combinations of 2, 3, or 4 “Topo-Tracks”, which are a subset of HLT2 tracks with additional requirements on track fit quality, impact parameter (IP), and muon or electron

identification. N-body candidates are constructed based on the distance of closest approach (DOCA) between Topo-Tracks. The mass window is defined based on the corrected mass, defined as

$$m_{\text{corr}} = \sqrt{m^2 + |p_{\text{Tmiss}}|^2} + |p_{\text{Tmiss}}|, \tag{2.7}$$

where p_{Tmiss} is the missing energy in the transverse direction with respect to the flight direction of the B meson. This variable allows for inclusive selections without a loose mass window.

Together with topological or inclusive lines, HLT2 also contains exclusive trigger lines designed for specific physics goals. For example most c-hadron decays, for which the rate would exceed the allowance of an inclusive line, are selected by defining exclusive lines tailored for each decay. These lines apply stringent cuts on invariant mass and angles to ensure exclusive selections. Specific lines, such as those for $D^{*+} \to D^0\pi^+$, can be selected inclusively, while others, targeting hadronic two-body and three-body decays, require full reconstruction of decay products. The selection efficiency of a handful of HLT2 lines can be seen in Fig. 2.26.

Additionally, HLT2 incorporates various exclusive and technical trigger lines for specialized purposes, including luminosity measurements, prescaled physics lines, low multiplicity events, and monitoring for quality feedback on the data. These lines contribute to the overall flexibility and efficiency of the LHCb trigger system, ensuring that a diverse range of physics processes is captured for subsequent analysis. For the scope of this thesis, a dedicated HLT2 line was developed to identify and select $H_b \to h^+h'^-$ decays. More details on the algorithm sequence and applied requirements are provided in Section 3.2

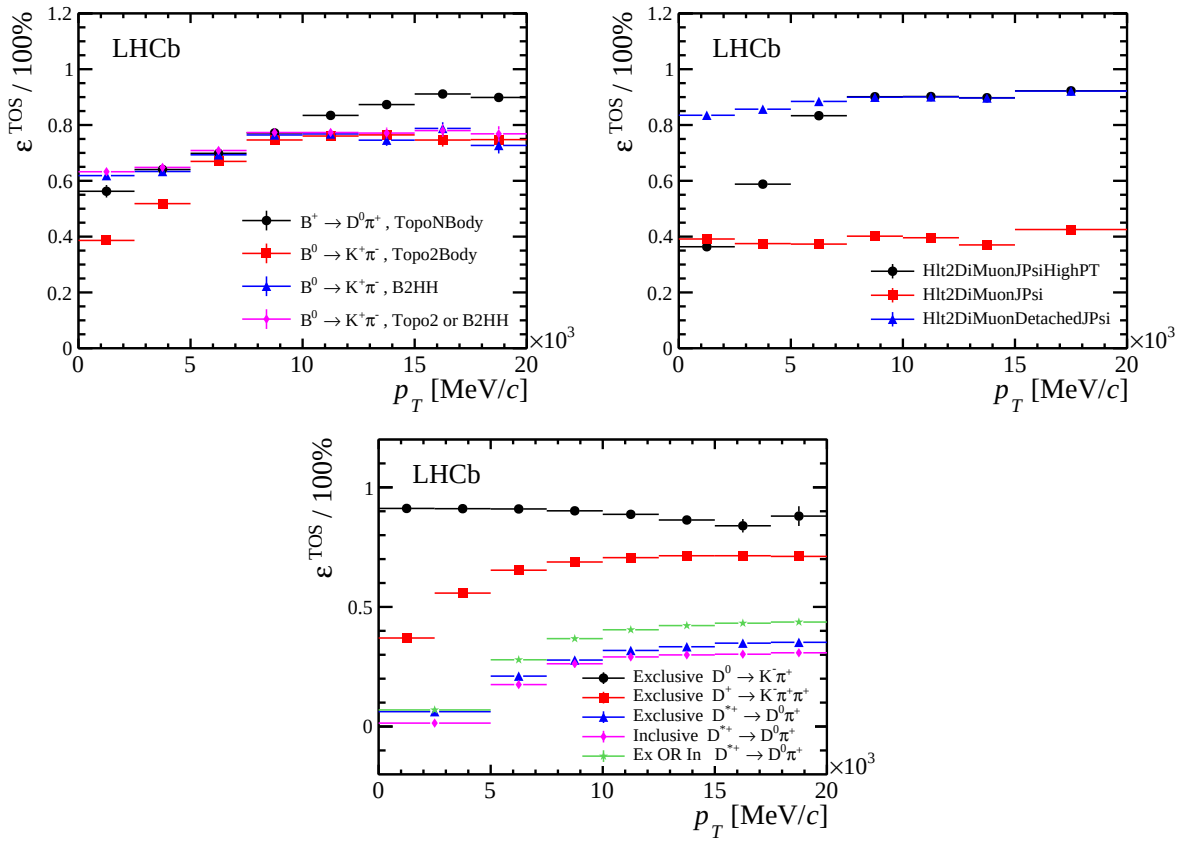

Figure 2.26: Efficiencies of selected HLT2 lines as a function of p_{T}: beauty lines (top left), J/ψ lines (top right), and charm lines (bottom).

Chapter 3

Measurement of $A_{CP}(pK^-)$ and $A_{CP}(p\pi^-)$

In this chapter, the procedure to extract the *CP* asymmetries of $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays from the Run 1 and 2 dataset collected by the LHCb experiment is described.

3.1 Analysis strategy

The direct *CP* asymmetries in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays are defined as

$$A_{CP}(pK^-) \equiv \frac{\Gamma(\Lambda_b^0 \to pK^-) - \Gamma(\overline{\Lambda}_b^0 \to \overline{p}K^+)}{\Gamma(\Lambda_b^0 \to pK^-) + \Gamma(\overline{\Lambda}_b^0 \to \overline{p}K^+)}, \tag{3.1}$$

$$A_{CP}(p\pi^-) \equiv \frac{\Gamma(\Lambda_b^0 \to p\pi^-) - \Gamma(\overline{\Lambda}_b^0 \to \overline{p}\pi^+)}{\Gamma(\Lambda_b^0 \to p\pi^-) + \Gamma(\overline{\Lambda}_b^0 \to \overline{p}\pi^+)}, \tag{3.2}$$

where Γ is the instantaneous decay rate of the process between parentheses. From the experimental point of view the *CP* asymmetries can be expressed as the sum of various contributions

$$A_{CP}(pK^-) \approx A_{\rm raw}(pK^-) - A_{\rm D}(p) - A_{\rm D}(K^-) - A_{\rm PID}(pK^-) - A_{\rm P}(\Lambda_b^0) - A_{\rm trig}(pK^-), \tag{3.3}$$

$$A_{CP}(p\pi^-) \approx A_{\rm raw}(p\pi^-) - A_{\rm D}(p) - A_{\rm D}(\pi^-) - A_{\rm PID}(p\pi^-) - A_{\rm P}(\Lambda_b^0) - A_{\rm trig}(p\pi^-), \tag{3.4}$$

where $A_{\rm raw}(pK^-)$ and $A_{\rm raw}(p\pi^-)$ are the raw asymmetries between the yields observed in data of the two charge-conjugated modes for the $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays, respectively. The other terms appearing in Eqs. (3.3) and (3.4) are nuisance asymmetries introduced by experimental effects: $A_{\rm D}(h)$ is the detection asymmetry for a given hadron, with $h \in \{K, \pi, p\}$, $A_{\rm PID}(ph)$ are the asymmetries due to the particle identification (PID) requirements imposed on the final-state particles, $A_{\rm P}(\Lambda_b^0)$ is the Λ_b^0 production asymmetry, arising from the different production cross-sections of Λ_b^0 and $\overline{\Lambda}_b^0$ baryons, and $A_{\rm trig}(ph)$ are the asymmetries introduced by the requirements imposed by the L0 and HLT1 triggers.

The raw asymmetries are measured by means of a simultaneous maximum-likelihood fit to the invariant mass of the eight final-state samples in which the data are divided using PID requirements: $K^+\pi^-$, π^+K^-, K^+K^-, $\pi^+\pi^-$, pK^-, $K^+\overline{p}$, $p\pi^+$, and $\pi^+\overline{p}$. Fitting the eight spectra simultaneously is needed to determine the contamination of mis-identified

$B \to h^+h'^-$ decays. The raw asymmetries are defined as

$$A_{\rm raw}(pK^-) = \frac{N(\Lambda_b^0 \to pK^-) - N(\overline{\Lambda}_b^0 \to \overline{p}K^+)}{N(\Lambda_b^0 \to pK^-) + N(\overline{\Lambda}_b^0 \to \overline{p}K^+)}, \tag{3.5}$$

$$A_{\rm raw}(p\pi^-) = \frac{N(\Lambda_b^0 \to p\pi^-) - N(\overline{\Lambda}_b^0 \to \overline{p}\pi^+)}{N(\Lambda_b^0 \to p\pi^-) + N(\overline{\Lambda}_b^0 \to \overline{p}\pi^+)}, \tag{3.6}$$

where N is the signal yield of a given decay mode as determined from the fits. The selection of the events will be described in Section 3.2, including the trigger and offline cuts, as well as the description of the simulated samples used for the analysis. The calibration of the Particle Identification (PID) performances, needed to correctly treat the contamination from cross-feed decays and the potential impact on the raw asymmetries, is described in Section 3.3. The models used to fit the data, covering all signal and background components, are described in Section 3.4. A multivariate BDT classfier is employed to further improve the purity of the sample, and a simultaneous optimization of the BDT and PID cuts is performed to achieve the best sensitivity on the *CP* asymmetries, as will be described in Section 3.5. In Section 3.6, more refinements of the mass models are presented, as needed to take into account the effect of the optimised selection. The final results of the fitting procedure are presented and discussed in Section 3.7, while the discussion of the systematic uncertainties related to the fit model is reported in Section 3.8.

The detection asymmetry for a particle h can be written as

$$A_{\rm D}(h^+) \equiv \frac{\varepsilon_{\rm det}(h^+) - \varepsilon_{\rm det}(h^-)}{\varepsilon_{\rm det}(h^+) + \varepsilon_{\rm det}(h^-)}, \tag{3.7}$$

where $\varepsilon_{\rm det}(h^\pm)$ is the detection efficiency of a positive or negative particle. This analysis requires measuring three different asymmetries:

- the proton detection asymmetry, $A_{\rm D}(p)$, that is determined by means of the same strategy used in Ref. [112] relying on fully simulated events and validated with data;
- the kaon detection asymmetry, $A_{\rm D}(K^-)$, is determined by correcting for the π detection asymmetry the $K^+\pi^-$ detection asymmetry. The latter is measured using $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K^0_{\rm S}\pi^+$ decays as described in Refs. [113, 114];
- the pion detection asymmetry, $A_{\rm D}(\pi^-)$, is taken from two independent studies for Run 1 and Run 2: the former was performed in Ref. [115] by studying partially reconstructed $D^{*+} \to (D^0 \to K^-\pi^+\pi^-\pi^+)\pi^+$ decays, while the latter comes from studies on $D^* \to (D^0 \to K^0_{\rm S}\pi^+\pi^-)\pi^+$ decays done in Ref. [116];

all these measurements will be presented in Section 3.9.

The PID asymmetries are defined as

$$A_{\rm PID}(pK^-) \equiv \frac{\varepsilon_{\rm PID}(pK^-) - \varepsilon_{\rm PID}(\overline{p}K^+)}{\varepsilon_{\rm PID}(pK^-) + \varepsilon_{\rm PID}(\overline{p}K^+)}, \tag{3.8}$$

$$A_{\rm PID}(p\pi^-) \equiv \frac{\varepsilon_{\rm PID}(p\pi^-) - \varepsilon_{\rm PID}(\overline{p}\pi^+)}{\varepsilon_{\rm PID}(p\pi^-) + \varepsilon_{\rm PID}(\overline{p}\pi^+)}, \tag{3.9}$$

where the PID efficiencies $\epsilon_{\rm PID}$ are those obtained from the PID calibration procedure described in Section 3.3. The efficiency evaluation for the $\Lambda_b^0 \to ph^-$ sample is presented in Section 3.10.

The trigger asymmetry is the sum of the L0 and HLT1 trigger asymmetries. When particles are responsible for the affirmative decision of a trigger algorithm (TOS or `Trigger On Signal`), the induced asymmetry is determined using inclusive samples of $\Lambda_b^0 \to \Lambda_c^+(pK^-\pi^+)\mu^-\overline{\nu}_\mu X$ (for protons) and $B^+ \to D^0(K^-\pi^+)\mu^-\overline{\nu}_\mu X$ decays (for kaons and pions). When the trigger algorithms select an event independently on the signal candidates (TIS or `Trigger Independent from Signal`) the induced asymmetry is calibrated by means of $B^+ \to J/\psi(\mu^+\mu^-)K^+$ decays. The procedure used to determine these corrections is described in Section 3.11.

The production asymmetry of the Λ_b^0 baryon was only measured directly in Run 1, and since it is expected to decrease with increasing colliding energies, it can only be used on the Run 1 sample, for which the results are reported in Section 3.12. To remove the contribution from the Λ_b^0 production asymmetry in the Run 2 dataset, a control sample of $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\pi^-$ decays is used. The raw asymmetry for this decay can be written as

$$\begin{aligned} A_{\text{raw}}(\Lambda_b^0 \to \Lambda_c^+\pi^-) \approx\, & A_{CP}(\Lambda_b^0 \to \Lambda_c^+\pi^-) + A_{\text{D}}(p) + A_{\text{D}}(K^-) + A_{\text{D}}(\pi^+) \\ & + A_{\text{D}}(\pi^-) + A_{\text{PID}}(pK^-\pi^+\pi^-) + A_{\text{P}}(\Lambda_b^0) + A_{\text{trig}}(pK^-\pi^+\pi^-), \end{aligned} \tag{3.10}$$

where the detection, PID, and trigger asymmetries will be measured with the methods just introduced and presented in Section 3.13. The *CP* asymmetry of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay, being a Cabibbo-favoured decay, is expected to be compatible with 0 to a precision much higher than the statistical power of this analysis, therefore we can neglect it in Eq. (3.10). The integrated value of $A_{\text{P}}(\Lambda_b^0)$ is expected to differ between the $\Lambda_b^0 \to ph^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ samples, since it was shown in Ref. [112] that it depends at least on the pseudorapidity of the baryon. Therefore, we will perform a reweighting of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ sample to equalise the momentum and pseudorapidity distributions to those of the $\Lambda_b^0 \to ph^-$ sample. Everything considered, we can then combine Eq. (3.10) with Eqs. (3.3) and (3.4) to obtain $A_{CP}(pK^-)$ and $A_{CP}(p\pi^-)$:

$$\begin{aligned} A_{CP}(pK^-) \approx\, & A_{\text{raw}}(pK^-) - A_{\text{D}}(p|\Lambda_b^0 \to pK^-) - A_{\text{D}}(K^-|\Lambda_b^0 \to pK^-) \\ & - A_{\text{PID}}(pK^-) - A_{\text{trig}}(pK^-) - A_{\text{raw}}(\Lambda_b^0 \to \Lambda_c^+\pi^-) + A_{\text{D}}(p|\Lambda_b^0 \to \Lambda_c^+\pi^-) \\ & + A_{\text{D}}(K^-|\Lambda_c^+ \to pK^-\pi^+) + A_{\text{D}}(\pi^+|\Lambda_c^+ \to pK^-\pi^+) \\ & + A_{\text{D}}(\pi^-|\Lambda_b^0 \to \Lambda_c^+\pi^-) + A_{\text{PID}}(\Lambda_b^0 \to \Lambda_c^+\pi^-) + A_{\text{trig}}(\Lambda_b^0 \to \Lambda_c^+\pi^-), \end{aligned} \tag{3.11}$$

$$\begin{aligned} A_{CP}(p\pi^-) \approx\, & A_{\text{raw}}(p\pi^-) - A_{\text{D}}(p|\Lambda_b^0 \to p\pi^-) - A_{\text{D}}(K^-|\Lambda_b^0 \to p\pi^-) \\ & - A_{\text{PID}}(p\pi^-) - A_{\text{trig}}(p\pi^-) - A_{\text{raw}}(\Lambda_b^0 \to \Lambda_c^+\pi^-) + A_{\text{D}}(p|\Lambda_b^0 \to \Lambda_c^+\pi^-) \\ & + A_{\text{D}}(K^-|\Lambda_c^+ \to pK^-\pi^+) + A_{\text{D}}(\pi^+|\Lambda_c^+ \to pK^-\pi^+) \\ & + A_{\text{D}}(\pi^-|\Lambda_b^0 \to \Lambda_c^+\pi^-) + A_{\text{PID}}(\Lambda_b^0 \to \Lambda_c^+\pi^-) + A_{\text{trig}}(\Lambda_b^0 \to \Lambda_c^+\pi^-), \end{aligned} \tag{3.12}$$

where we introduced the notation $A_{\text{D}}(h^\pm|\Lambda_b^0 \to f)$ to indicate the detection asymmetry of a particle $h^\pm$ measured on the sample $\Lambda_b^0 \to f$.

As the reader may have noted, some of the samples used to determine the nuisance asymmetries are used multiple times, hence there are relevant correlations among the various corrections that must be taken into account. The procedure to do so, based on running a high number of pseudexperiments, is described in Section 3.14.

Finally, in Section 3.15, the final results for the *CP* asymmetries are presented and discussed. The measurement is carried out for the total sample as well as divided by

Table 3.1: Total integrated luminosity corresponding to the pp collisions collected by LHCb, separated by year and magnet polarity. The corresponding centre-of-mass energy at which collisions occurred in each year are also reported.

Year of data taking	2011	2012	2015	2016	2017	2018
$\sqrt{s}$ [TeV]	7	8	13	13	13	13
Integrated Luminosity [fb^{-1}]	1.11	2.08	0.29	1.63	1.47	2.02

data-taking year and magnet polarity, in order to check that the value of A_{CP} is consistent across the subsamples, proving the robustness of the analysis method.

3.2 Data set and event selection

The data used in this analysis is taken from the pp collisions collected with the LHCb detector during the Run 1 (2011 and 2012) and Run 2 (2015, 2016, 2017, and 2018) of the LHC. The collisions occured at different centre-of-mass energies ($\sqrt{s} = 7,\ 8$ and 13 TeV) in different data-taking periods, corresponding to a total integrated luminosity corresponding to about 9 fb^{-1}. In Tab. 3.1, the breakdown of the centre-of-mass energies and total collected luminosities, separated by year of data taking, is presented.

3.2.1 Stripping preselection

The collected sample is firstly filtered with a preselection (also called *stripping*) named `StrippingB2HHBDT`, aiming at retaining as much signals as possible while reducing the total sample size to a manageable level. The `StrippingB2HHBDT` algorithm combines pairs of oppositely charged tracks in order to form $H_b \to h^+h'^-$ candidates, where the pion mass hypothesis is assumed for the final-state particles. For historical reasons, small differences are present between the requirements applied by the algorithm used during Run 1 and Run 2. In both versions, only good-quality tracks, *i.e.* with small normalised χ^2 (χ^2/ndf) and small probability to be a *ghost track* (`GhostProb`), are used. Then, those with large transverse momentum ($p_{\rm T}$) and incompatible to come from any of the primary pp-collision vertices (primary vertices or PVs) are retained for the next steps. The incompatibility for a track to come from any PV is obtained requiring large impact parameter ($d_{\rm IP}$) with respect to all the PVs for Run 1 data, while for Run 2 data, the same effect is obtained by requiring a large $\chi^2_{\rm IP}$[1]. For Run 1 data, pairs of tracks with a small distance of closest approach ($d_{\rm CA}$) are fitted to a common vertex in order to form the H_b candidate, while for Run 2 data a small normalised χ^2 for the hypothesis of the two tracks to come from the same vertex is used ($\chi^2_{\rm DOCA}$/ndf) in addition to requiring a large scalar sum of the $p_{\rm T}$ of the two tracks ($p^+_{\rm T} + p^-_{\rm T}$). For the Run 1 preselection, only the candidates with a large transverse momentum $p^{H_b}_{\rm T}$, a small impact parameter with respect to the associated PV

[1]The $\chi^2_{\rm IP}$ is defined as the difference in the vertex-fit χ^2 of a given PV reconstructed with and without the track under consideration.

Table 3.2: Summary of the values of the requirements used to form the $H_b \to h^+h'^-$ candidates by the `B2HHBDT` stripping line in Run 1 (left) and Run 2 (right). The meaning of the various symbols is explained in the text.

Variable	Cut
Track χ^2/ndf	< 3
Track `GhostProb`	< 0.5
Track p_{T} [GeV/c]	> 1.0
Track d_{IP} [µm]	> 120
d_{CA} [µm]	< 100
$d_{\mathrm{IP}}^{H_b}$ [µm]	< 120
$t_{\pi\pi}$ [ps]	> 0.6
$p_{\mathrm{T}}^{H_b}$ [GeV/c]	> 1.2
BDT output	> -0.3
$m_{\pi^+\pi^-}$ [GeV/c^2]	$\in [4.6, 6.4]$

Variable	Cut
Track χ^2/ndf	< 4
Track `GhostProb`	< 3
Track p_{T} [GeV/c]	> 1.0
Track χ^2_{IP}	> 16
$p_{\mathrm{T}}^+ + p_{\mathrm{T}}^-$ [GeV/c]	> 4.5
χ^2_{DOCA}/ndf (H_b)	< 9
χ^2_{FD} (H_b)	> 100
χ^2_{IP} (H_b)	< 9
DIRA(H_b)	> 0.99
$m_{\pi^+\pi^-}$ [GeV/c^2]	$\in [4.8, 6.2]$

$(d_{\mathrm{IP}}^{H_b})^2$ and a large decay time ($t_{\pi\pi}$, computed assuming decay into the $\pi^+\pi^-$ final-state) are selected by the Run 1 version of the stripping preselection. The purity of the sample is further improved for the Run 1 algorithm by means of a multivariate Boosted Decision Tree (BDT) classifier. The BDT algorithm discriminates between signal and combinatorial background on the basis of the smallest and largest p_{T} of the two tracks, the smallest and largest impact parameter (d_{IP}) of the two tracks, the distance of closest approach (d_{CA}) between the two tracks, the quality of the common vertex fit of the two tracks (χ^2_{vtx}), the p_{T} and d_{IP} of the H_b candidate ($p_{\mathrm{T}}^{H_b}$ and $d_{\mathrm{IP}}^{H_b}$), and the flight distance (FD) of the H_b candidate with respect to the associated PV. Background-like candidates are characterised by smaller values of the BDT output while signal-like candidates show larger values for the BDT output. The Run-2 version of the algorithm, instead, imposes limits on the values of the χ^2 of the impact parameter (χ^2_{IP}), distance of closest approach (χ^2_{DOCA}), and flight distance (χ^2_{FD}) of the H_b candidate with respect to the associated PV, as well as on the cosine of the angle between the flight and momentum direction of the H_b candidate (DIRA). No further refinement is applied in addition to this cut-based selection for the Run-2 version of the algorithm. In Tab. 3.2 the values of the requirements applied by the two versions of the stripping algorithm are reported.

Finally, as it will be described in Sec. 3.3, fiducial requirements are imposed on the momentum (p) and pseudrapidity (η) of the final-state tracks of the $H_b \to h^+h'^-$ candidates. The fiducial region is defined as the area inside the polygon with vertices of coordinates in the $(p\,[\mathrm{GeV}/c], \eta)$ plane $(0, 2)$, $(25, 2)$, $(120, 3.747)$, $(123, 4.2)$, and $(0, 4.2)$. The application of these fiducial requirements is needed to guarantee a proper coverage of the phase space by the PID calibrations samples. Their effects on the analysis are discussed in details in Sec. 3.3.1.

[2]The PV that fits best to the flight direction of the H_b candidate is taken as the associated PV.

3.2.2 Trigger selection

Among the candidates surviving the stripping preselection only those that have been acquired by specific trigger algorithms are retained. At the level of the hardware trigger, the final-state particles of the H_b candidates are required to be responsible for the affirmative decision of the `L0Hadron` trigger algorithm (`L0HadronTOS`), that means that at least one of the clusters in the HCAL associated to the final-state particles must have a transverse energy, E_{T}, larger than a given threshold. In addition, also the H_b candidates in event where the affirmative decision of the L0 trigger is due to particles not used to form the H_b candidates are stored. This category goes under the naming of `L0GloblaTIS` (the distrinction between `TIS` and `TOS` trigger decisions is explained in Section 3.1). At the first stage of software trigger (HLT1) the H_b candidates are required to be responsible for the affirmative decision of the `Hlt1TrackAllL0` (in Run 1) or `Hlt1TrackMVA` (in Run 2) algorithms, as described in Section 2.5. At the final stage (HLT2), the H_b candidates are required to be responsible for the affirmative decision of the `Hlt2B2HH` algorithm, that is a simplified version of the stripping algorithms described in Sec. 3.2.1. The requirements applied by the `Hlt2B2HH` algorithm are reported in Tab. 3.2.

3.2.3 Simulated samples

Fully simulated samples of all $H_b \to h^+h'^-$ decays, produced with the LHCb simulation framework for Run 1 and Run 2, are used in several parts of this analysis. In LHCb, pp collisions are simulated by Pythia [117, 118] with an LHCb-specific configuration, while EvtGen [119] describes the decay of unstable particles, with Photos [120] managing QED final-state radiation. The Geant4 toolkit [121, 122] is used to simulate the interactions between the generated particles and the detector. In these samples the data taking conditions, trigger, reconstruction and stripping corresponding to the different years have been reproduced in order to have events as similar as possible to real data. The trigger conditions used for the largest part of the collected luminosity during the years are simulated. The total amount of simulated events is generated such that the proportion between data collected with the two magnet polarities is reproduced. In Table 3.3 we report the amount of simulated events for all the $H_b \to h^+h'^-$ decays. The reconstructed candidates are required to pass all the preselections described in Sections 3.2.1 and 3.2.2. In addition, only those associated with a true $H_b \to h^+h'^-$ decay are retained.

3.3 PID calibration

The proper determination of the efficiencies introduced by the requirements on PID variables is of the utmost importance for this analysis. All the different $H_b \to h^+h'^-$ decay modes represent one of the main sources of background to each other, since their invariant-mass distributions with one or both final-state particles misidentified peak very close to where also the correctly identified modes do. The discriminating variables chosen in this analysis to distinguish between pions, kaons, and protons are the DLL variables described in Section 2.4.1.

Table 3.3: Number of generated events, separated by data taking year.

Decay	2011 [10^3]	2012 [10^3]	2015 [10^3]	2016 [10^3]	2017 [10^3]	2018 [10^3]
$B^0 \to K^+\pi^-$	368	692	1109	4722	7631	7621
$B^0 \to \pi^+\pi^-$	392	704	565	2361	3832	3806
$B^0 \to K^+K^-$	230	472	164	960	1693	1600
$B_s^0 \to K^+K^-$	374	700	523	2359	3834	3815
$B_s^0 \to \pi^+K^-$	353	709	292	1918	3161	3216
$B_s^0 \to \pi^+\pi^-$	255	489	167	965	1761	1574
$\Lambda_b^0 \to pK^-$	167	305	142	976	1576	1536
$\Lambda_b^0 \to p\pi^-$	176	326	284	563	577	651

3.3.1 Calibration samples

In order to calibrate the efficiencies introduced by PID requirements applied on pions and kaons, high-statistics and high-purity samples of $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays are used. In the case of protons, instead, samples of $\Lambda \to p\pi^-$ and $\Lambda_c^+ \to pK^-\pi^+$ decays are utilised. The kinematic features of these decays enable the identity of the final-state particles to be determined without using any PID information. The residual background contamination is removed from the samples using the *sPlot* technique [123]. The weighted samples are provided centrally as part of the `PIDCalibTool` package of the LHCb Collaboration [124, 125]. As anticipated in Sec. 3.2.1, fiducial requirements are applied to the H_b candidates out of the stripping selection. This is done in order to remove the H_b candidates having final-state particles in regions of the phase-space that are not covered by the PID-calibration samples. In Fig. 3.1 the (p, η) distributions of kaons, pions and protons are shown for both PID-calibration samples and fully simulated $H_b \to h^+h'^-$ decays, with the fiducial region highlighted. From the distributions shown in Fig. 3.1, it is clear that a large portion of $H_b \to h^+h'^-$ candidates is removed when applying the fiducial requirements to both the final state particles. The removed candidates amount to about 30% of the total sample. However, it is important to note two relevant aspects:

- the regions of the phase space removed by the fiducial requirements are regions where PID efficiencies are small, hence the events rejected by the fiducial cut would most probably be removed also by any PID requirement;

- the sample lying outside of the fiducial region was kept in the previous version of this analysis [13]. That decision implied a large systematic uncertainty in the determination of the asymmetry introduced by PID requirements, due to the lack of calibration protons in that region of the phase space. Considering the much improved statistical precision expected in this analysis, it is fundamental to reduce accordingly all the systematic uncertainties. Therefore, the loss of statistics introduced by the fiducial cut is counterbalanced by the reduction in the systematic uncertainty due to the chosen PID requirements.

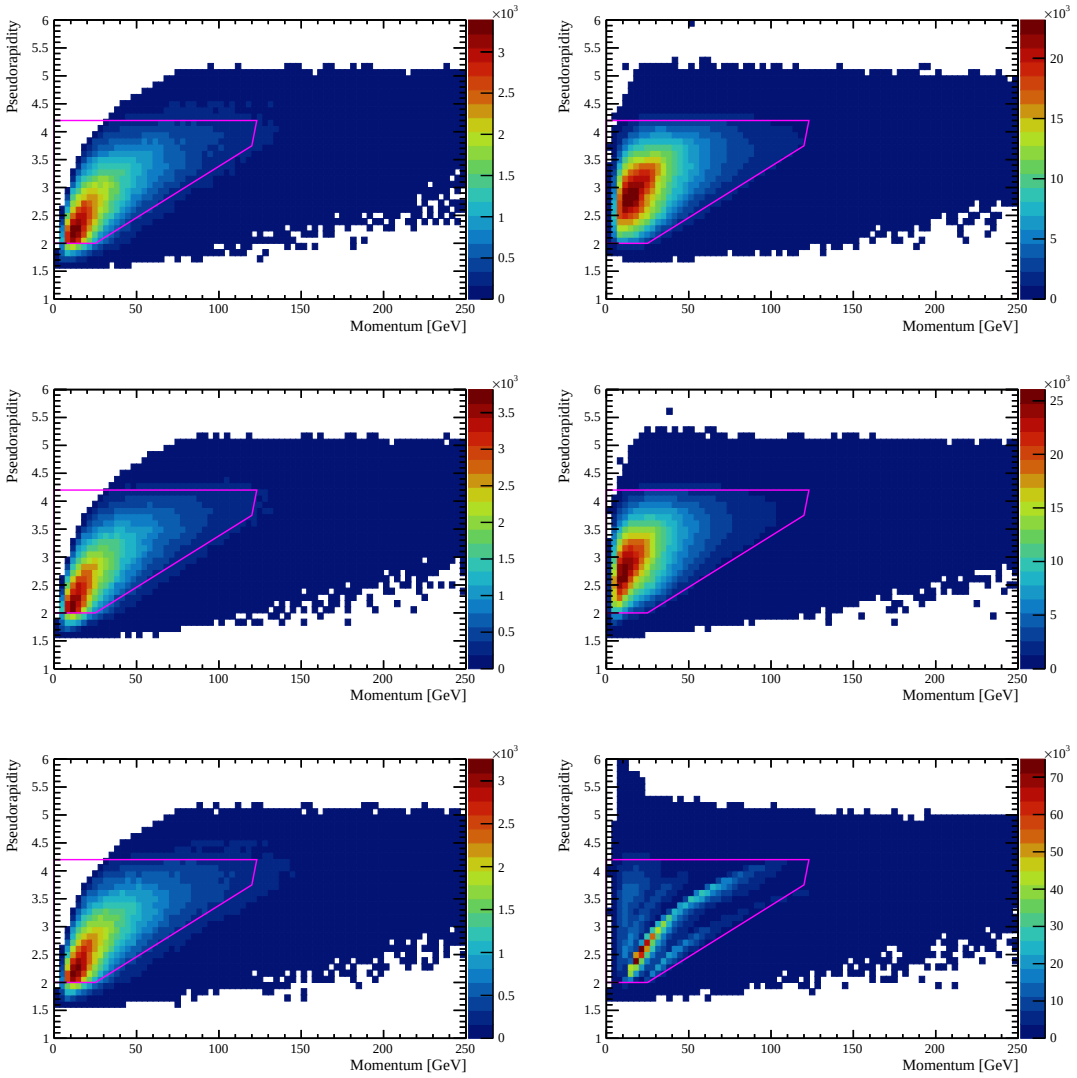

Figure 3.1: Distribution in the (p, η) plane of (top) kaons, (middle) pions and (bottom) protons coming from (left) $H_b \to h^+h'^-$ decays and (right) PID-calibration samples. The distribution related to $H_b \to h^+h'^-$ decays are taken from fully simulated events, while those related to PID calibration samples are obtained from background-subtracted events, as explained in the text. In all the cases, a magenta line is drawn to identify the fiducial region defined in Section 3.2.1.

3.3.2 Calibration procedure

The PID calibration procedure has been developed taking into account the following considerations:

- in order to distinguish between three types of particles (pions, kaons and protons) it is necessary to utilise two DLL variables for each final-state particle. For example, in order to select kaons, requirements on both $\mathrm{DLL}_{K\pi}$ (in order to reject pions) and DLL_{Kp} (in order to reject protons) must be applied. If a particle satisfies the criteria defined to select kaons, then the kaon hypothesis is assigned to that particle;

- the value of DLL depends directly on the momentum of the particle through its

relation with the emission angle of Cherenkov photons, as well as on its pseudorapidity due to the different length of radiator material traversed at different angles. In addition, since the two RICH detectors have different angular acceptances and have radiators optimised for different momentum regions, the DLL values show a dependence also on the pseudorapidity of the particle;

- since the RICH performance degrades with increased occupancy [103], the effect is taken into account by studying the dependence of DLL with respect to an occupancy figure Θ, chosen to be the number of tracks in the event (nTracks) in Run 1 and the number of hits registered by the SPD detector (nSPDHits) in Run 2.

As a first step, for a given set of PID requirements, maps of PID efficiencies in bins of p, η and Θ are determined. For example, in a given region of p, η and Θ, the efficiency of a PID requirement applied to kaons, is given by the number of calibration kaons satisfying that requirement divided by the total number of calibration kaons in that region. The binning scheme used to divide the phase space is

Track momentum:
2 bins for $0 < p < 10\,\mathrm{GeV}/c$;
45 bins for $10 < p < 100\,\mathrm{GeV}/c$;
20 bins for $100 < p < 150\,\mathrm{GeV}/c$;
4 bins for $150 < p < 500\,\mathrm{GeV}/c$;

Track pseudorapidity:
10 bins for $1 < \eta < 6$;

Number of tracks:
4 bins for $0 < \mathrm{nTracks} < 400$;
1 bin for $400 < \mathrm{nTracks} < 600$;

Number of SPD hits:
3 bins for $0 < \mathrm{nSPDHits} < 450$;
3 bins for $450 < \mathrm{nSPDHits} < 1000$;

Since the event occupancy and the kinematic of a particle are independent quantities, the dependency of the efficiency from Θ is integrated out. Assuming the possibility of analytically describing the PID efficiency, ε, as a function of p, η and Θ and the distribution of SPD hits for the $H_b \to h^+h'^-$ sample $f(\Theta)$, the procedure could be formalized by the following equation

$$\bar{\varepsilon}(p,\eta) = \int \varepsilon(p,\eta,\Theta) \cdot f(\Theta)\,\mathrm{d}\Theta, \tag{3.13}$$

where $\bar{\varepsilon}(p,\eta)$ is the PID efficiency as a function of p and η for a particle in the same occupancy regime that we observe in the $H_b \to h^+h'^-$ data sample. However, in reality the PID efficiency and the distribution of Θ can not be expressed in an analytical form. Hence, the integration of Eq. (3.13) is replaced by the sum

$$\bar{\varepsilon}(p_i,\eta_j) = \frac{1}{N}\sum_{k=1}^{N} \varepsilon(p_i,\eta_j,\Theta_k), \tag{3.14}$$

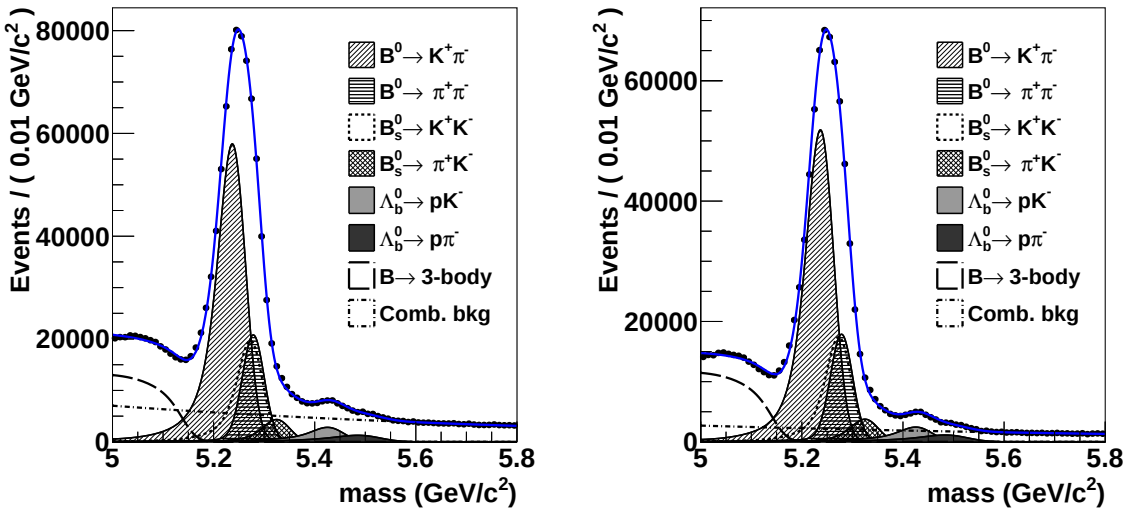

Figure 3.2: Distributions of the invariant mass of $H_b \to h^+h'^-$ under the $\pi\pi$ hypothesis without PID cuts. A cut of (left) 0.08 and (right) 0.12 on the BDT classifier described in Section 3.5 is applied. The results of fits with a model describing the signal, combinatorial, and partially reconstructed background are superimposed.

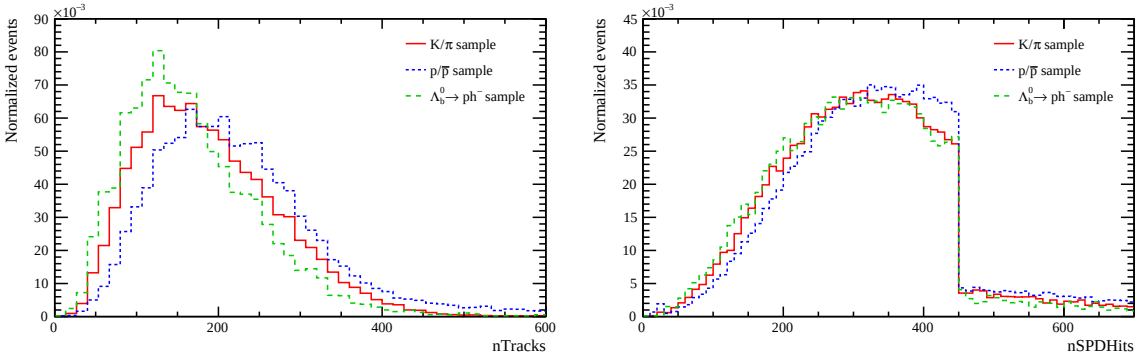

Figure 3.3: Distribution of (left) nTracks and (right) nSPDHits for PID-calibration and background-subtracted $\Lambda_b^0 \to ph^-$ samples from Run 1 and Run 2, respectively.

where $\bar{\varepsilon}(p_i, \eta_j)$ is the final PID efficiency corresponding to the i-th bin of particle momentum and j-th bin of particle pseudorapidity; $\varepsilon(p_i, \eta_j, \Theta_k)$ is the PID efficiency corresponding to the i-th bin of particle momentum, the j-th bin of particle pseudorapidity, and k-th bin of Θ; N is a number large enough to avoid statistical fluctuations in the average (set to $N = 1000000$). For each term of the sum the value of Θ_k has been randomly extracted according to the distribution of Θ obtained from the background-subtracted $H_b \to h^+h'^-$ sample.

The background subtraction of $H_b \to h^+h'^-$ events has been performed using the sPlot technique [126], by fitting the invariant mass computed assuming both final state particles to be pions ($m_{\pi\pi}$). Events are selected applying the offline selection described

in Section 3.2 apart from the PID requirements that are not applied. The shapes of signal contributions have been parameterised applying a Kernel Estimation Method [127] to the distribution of $m_{\pi\pi}$ for fully simulated events, computed assuming perfect invariant-mass resolution. The obtained non-parametric distributions have then been convolved with a Gaussian resolution model with free mean and width. The relative fractions between the various $H_b \to h^+h'^-$ decays have been fixed to the values measured by LHCb in Ref. [128]. In the case of Λ_b^0 decays we used the world averages of the absolute branching fractions computed by the Heavy Flavour Averaging Group (HFLAV) [29]. The contribution due to combinatorial background has been parameterized with an exponential function, while the component coming from partially reconstructed 3-body B decays has been described using an ARGUS function [129] convolved with the same Gaussian resolution model used for the signal shapes and described in Section 3.4. The $m_{\pi\pi}$ distributions with the two BDT cuts used for the selection of $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays are reported in Fig. 3.2 with the result of the fit overlaid. The resulting background-subtracted distributions of the occupancy variables for Run 1 and Run 2 are reported in Fig. 3.3 together with those of the calibration samples.

The computation from Eq. (3.14) produces PID efficiency maps in bins of p and η for particles coming from $H_b \to h^+h'^-$ decays. As a reference, in Fig. 3.4 we report the PID efficiency maps for pions, kaons and protons following three different choices of PID requirements.

Finally, the efficiency of a PID requirement applied on a $H_b \to h^+h'^-$ decay is estimated using the following equation:

$$\hat{\varepsilon}_{h^+h'^-} = \frac{1}{N}\sum_{i=1}^{N} \bar{\varepsilon}_{h^+}\left(p_i^+, \eta_i^+\right) \cdot \bar{\varepsilon}_{h'^-}\left(p_i^-, \eta_i^-\right), \tag{3.15}$$

where N is the number of $H_b \to h^+h'^-$ candidates, $\bar{\varepsilon}_{h^+}$ and $\bar{\varepsilon}_{h'^-}$ are the efficiencies as a function of p and η as determined from Eq. (3.14), $p_i^{+(-)}$ and $\eta_i^{+(-)}$ are the momentum and the pseudorapidity of the positive (negative) particle of the i-th candidate. Candidates from fully simulated events are used.

3.4 Invariant-mass models

The strategy we adopted to optimise the event selection is based on the knowledge of the model used to fit the invariant-mass spectra of selected events. In this Section the studies performed in order to determine the various probability density functions (p.d.f.s) used to parameterise all contributions to the spectra are presented. Four componets are identified:

Signal: $H_b \to h^+h'^-$ decays where the final-state particles are correctly identified.

Cross-feed background: $H_b \to h^+h'^-$ decays in which the identity of one or both the final-state particles is wrongly assigned. This background is particularly dangerous since it peaks at the signal distribution.

Partially reconstructed background: $H_b \to h^+h'^-X$ multibody decays where only the two h^+ and h'^- hadrons have been used to form the parent H_b.

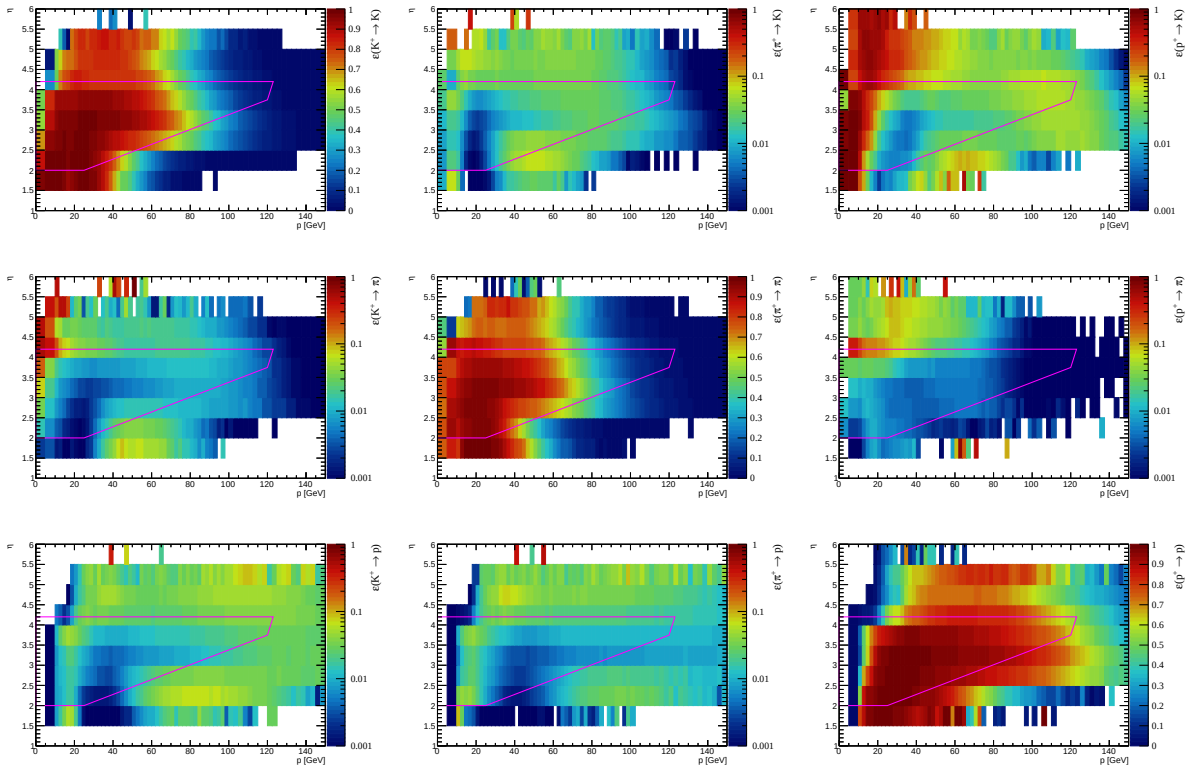

Figure 3.4: (p_{T},η) maps of PID efficiencies of reconstructing kaons (left column), pions (middle column), and protons (right column) as kaons (top row), pions (middle row), and protons (bottom row). The PID requirements are $\mathrm{DLL}_{K\pi} > 5$ and $\mathrm{DLL}_{Kp} > -5$ for the kaon hypothesis (top row), $\mathrm{DLL}_{K\pi} > -5$ and $\mathrm{DLL}_{p\pi} > -5$ for the pion hypothesis (middle row), and $\mathrm{DLL}_{p\pi} > -5$ and $\mathrm{DLL}_{pK} > 5$ for the proton hypothesis (bottom row). The fiducial region is shown as a magenta line.

Combinatorial background: candidates composed by pairs of oppositely charged tracks not coming from the same decay chain.

The shapes chosen to model the invariant-mass distributions of all the sources above will be described in detail in the following.

3.4.1 Signal model

The signal distribution is modelled by the sum of a Johnson S_U function and one or two Gaussian functions. The Johnson distribution is defined by the formula [130]

$$S_U(m;\mu,\sigma,\gamma,\delta) = \frac{\delta}{\sigma\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{m-\mu}{\sigma}\right)^2}} \exp\left[-\frac{1}{2}\left(\gamma + \delta \sinh^{-1}\left(\frac{m-\mu}{\sigma}\right)^2\right)\right], \quad (3.16)$$

where m is the reconstructed invariant-mass and μ is the peak of the considered H_b hadron taken from [32]. One Gaussian function with the same peak and width is added to every signal component, yielding the total p.d.f.:

$$\mathrm{pdf}_s(m) = f S_U(m;\mu,\sigma,\gamma,\delta) + (1-f)G(m;\mu,\sigma), \quad (3.17)$$

where f is the relative fraction of the two components. For decays with a relatively large branching ratio, *i.e.* $B^0 \to K^+\pi^-$, $\overline{B}^0 \to \pi^+K^-$, $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ a second Gaussian is added with the same mean and a separate width to improve the fit quality.

The values for the shape parameters of the Johnson, γ and δ, are fixed to those obtained from fits to the simulated samples, separately for each decay channel. In Figs 3.5 and 3.6 we report the invariant-mass distribution for all the simulated $H_b \to h^+h'^-$ decays with the result of the best fit superimposed.

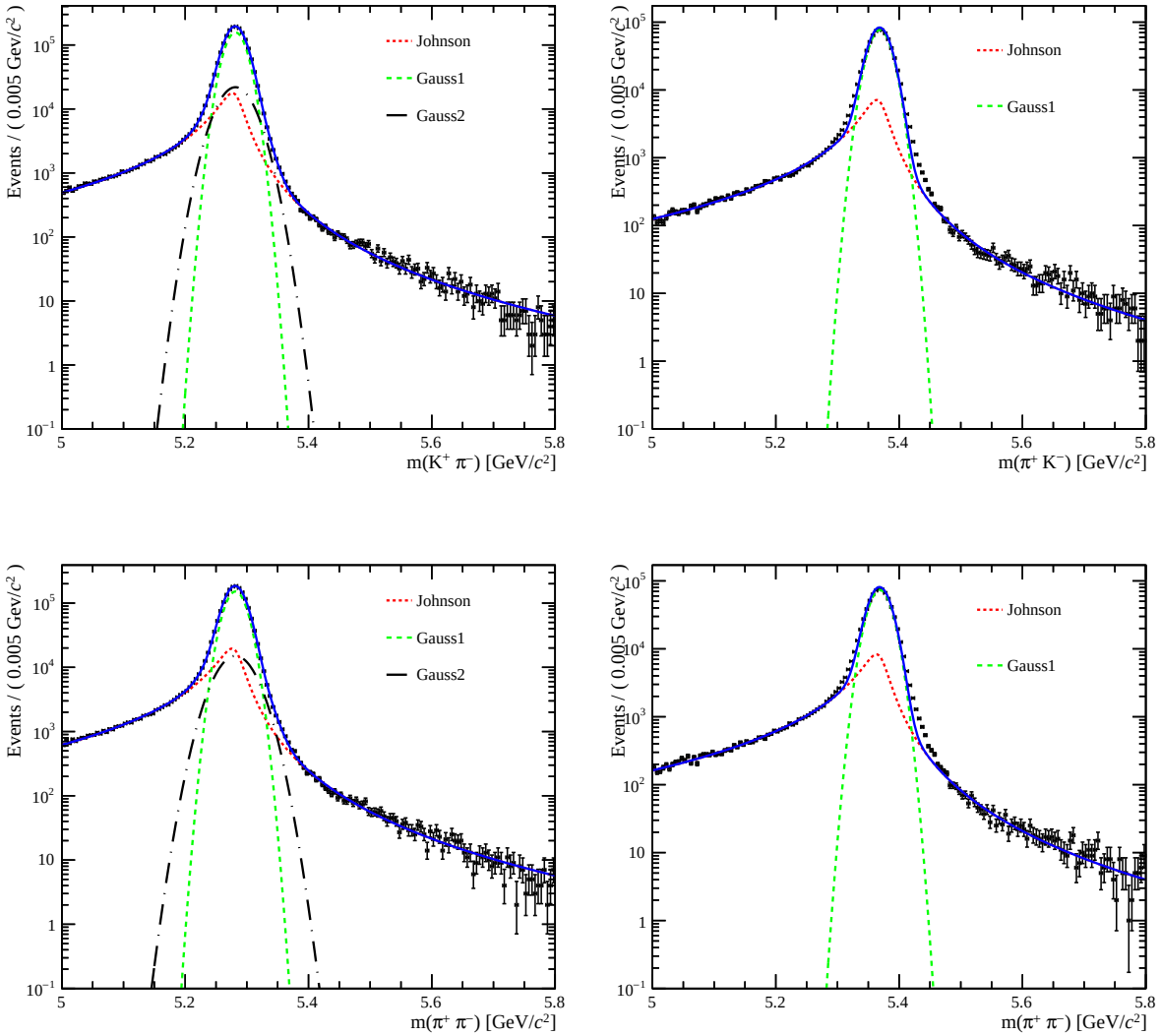

Figure 3.5: Invariant-mass distributions of fully simulated (top left) $B^0 \to K^+\pi^-$, (top right) $B_s^0 \to \pi^+K^-$, (bottom left) $B^0 \to \pi^+\pi^-$ and (bottom right) $B_s^0 \to \pi^+\pi^-$ decays passing the preselection described in Section 3.2. The result of the best fit using the model described in the text in Section 3.4.1 is overlaid.

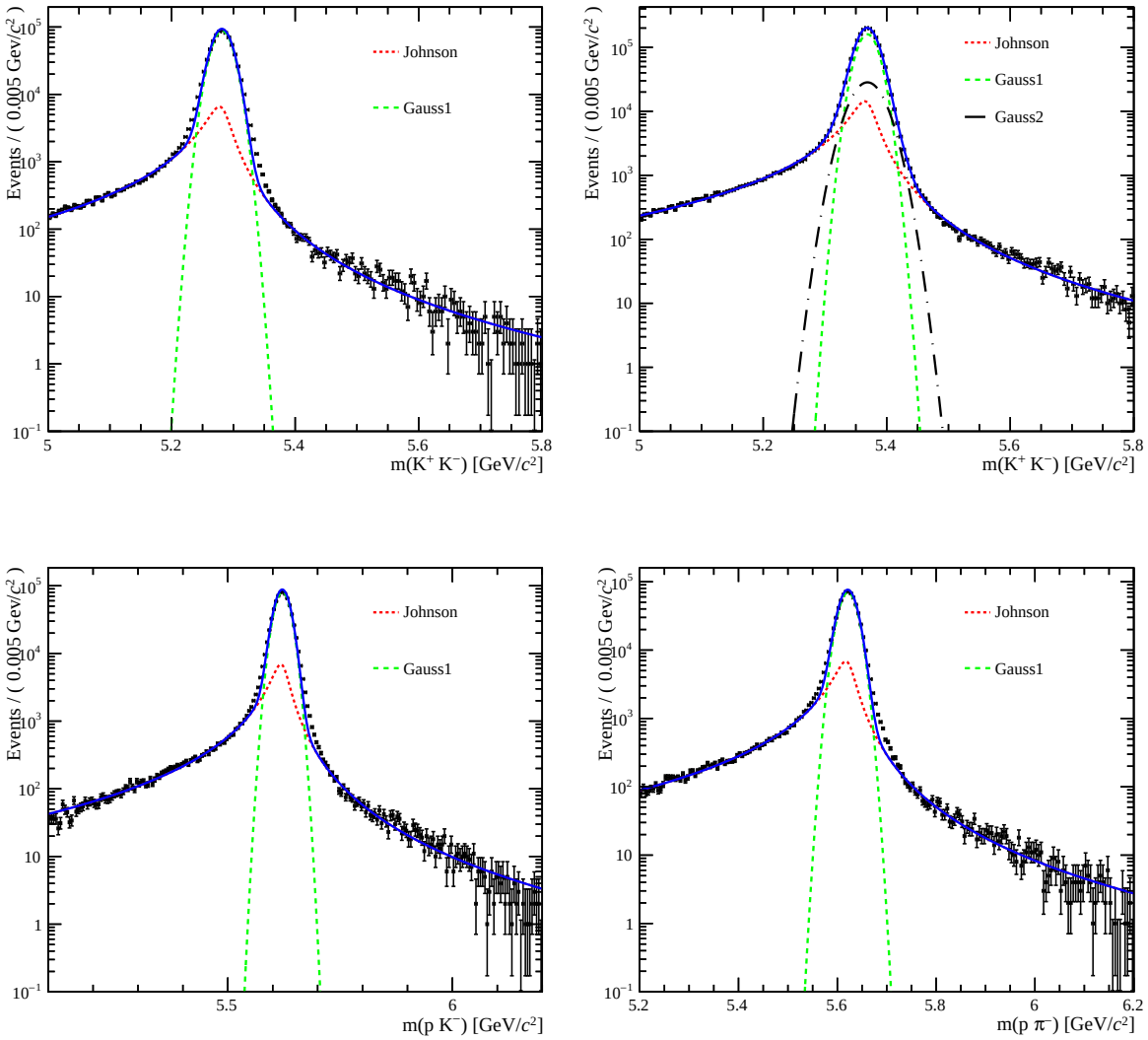

Figure 3.6: Distribution of invariant-mass for fully simulated (top left) $B^0 \to K^+K^-$, (top right) $B_s^0 \to K^+K^-$, (bottom left) $\Lambda_b^0 \to pK^-$ and (bottom right) $\Lambda_b^0 \to p\pi^-$ decays passing the preselection described in Section 3.2. The result of the best fit using the model described in the text in Section 3.4.1 is overlaid.

3.4.2 Cross-feed background model

The parameterisation of the model used to describe the cross-feed backgrounds is studied using fully simulated signal decays. The procedure consists of two steps: a dataset containing the invariant-mass computed under the wrong hypothesis is produced and then the shape is built by applying a kernel density estimation (KDE) method [127] to the dataset.

The reconstructed invariant mass of any two-body decay under a different $h^+h'^-$ final-state hypothesis can be written as

$$m(h^+h'^-) = \sqrt{m_{h^+}^2 + m_{h'^-}^2 + 2\left(\sqrt{(m_{h^+}^2 + p_+^2)(m_{h'^-}^2 + p_-^2)} - \vec{p}_+ \cdot \vec{p}_-\right)}, \tag{3.18}$$

where m_{h^+} and $m_{h'^-}$ are the masses corresponding to the hypothesis, $\vec{p}_{+(-)}$ is the momentum of the positive (negative) particle obtained from the simulation and $p_{+(-)}$ stands for the module of $\vec{p}_{+(-)}$.

The datasets to which the KDE is applied are created from simulated $H_b \to h^+h'^-$ decays, computing for each event the wrong invariant-mass using the true momenta of the final-state particles (obtained accessing Monte Carlo truth information). In order to describe the cross-feed mass shapes it is necessary to take into account the effect of PID requirements. Since their application alters the momentum distribution of tracks, they have the effect of deforming the invariant-mass distributions obtained through the procedure described above. To address this, a weight to each simulated event is applied, corresponding to

$$w_i = \varepsilon_{h^+}\left(p_i^+, \eta_i^+\right) \varepsilon_{h'^-}\left(p_i^-, \eta_i^-\right), \tag{3.19}$$

where $\varepsilon_{h^\pm}$ are the PID efficiencies of the positive and negative particles obtained from the efficiency maps presented in Section 3.3, and $p_i^\pm$ and $\eta_i^\pm$ are the momenta and pseudorapidities in the i-th event. The kernel estimation method is applied to these weighted datasets. As an example, invariant-mass models obtained from the application of the kernel density estimation to the samples are shown in Fig. 3.7. Finally, in the invariant-mass fits, the obtained non-parametric p.d.f. are convolved with a Gaussian function of width equal to the one used for the signal shape, to account for invariant-mass resolution effects.

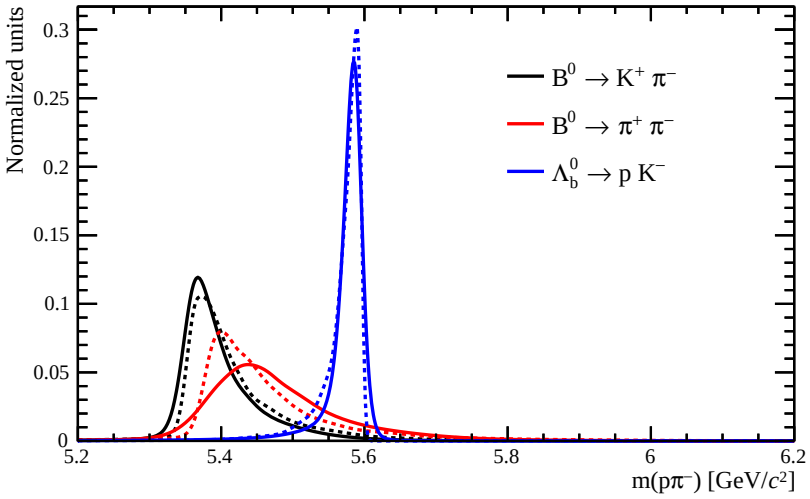

Figure 3.7: Invariant mass distributions of (black) $B^0 \to K^+\pi^-$, (red) $B^0 \to \pi^+\pi^-$, and (blue) $\Lambda_b^0 \to pK^-$ decays reconstructed in the $p\pi^-$ mass hypothesis calculated by means of Eq. (3.18) and with a kernel estimation technique applied. The dashed lines show the distributions without taking into account the deformation induced by PID requirements, while the solid lines represent the distributions weighted using the PID efficiencies.

3.4.3 Partially reconstructed multi-body H_b decays

This type of background originates from partially reconstructed decays where one or more final-state particles are not reconstructed. Usually, an ARGUS function [129] convolved

with a resolution model provides a good empirical modelling of this background component. The p.d.f. of the ARGUS function is

$$p(m; m_0, c) = m\sqrt{1 - \frac{m^{\,2}}{m_0}}\exp\left[c\left(1 - \left(\frac{m}{m_0}\right)^2\right)\right], \tag{3.20}$$

where m is the mass of the reconstructed candidates, m_0 is the threshold value and c is the parameter governing the shape of the funcion. Since the lightest particle that can be missed in the reconstruction of the candidate is a π^0, the end point of the ARGUS functions is fixed to $m_{B^0} - m_{\pi^0}$ and $m_{B_s^0} - m_{\pi^0}$ for partially reconstructed backgrounds coming from B^0, and B_s^0 decays, respectively. Another possible source of partially reconstructed background is the component due to three-body decays of the B^+ meson (like $B^+ \to h^+h'^-\pi^+$). This component is not parameterised explicitly as its shape is almost equal to that of partially reconstructed B^0-meson decays.

For partially reconstructed decays in the pK^- and $p\pi^-$ mass spectra, instead, a more careful study was carried out. Large samples of the relevant decays were produced with the fast simulation software RapidSim [131] to study the shape of the invariant-mass distribution of the visible final-state particles. The considered decays are chosen to be the ones that are expected to be more abundant in the signal mass window, *i.e.* $\Lambda_b^0 \to p(K^{*-} \to K^-\pi^0)$ for the pK^- channel and $\Lambda_b^0 \to p(\rho^- \to \pi^-\pi^0)$ for the $p\pi^-$ channel. The resulting distributions were used as a template in the fit and are shown on the left of Fig. 3.8. The shape of these distributions are affected by the BDT cut (described in Section 3.5) but the distortion cannot be studied directly since RapidSim does not reproduce all the variables used to train the model, so another approach was used, *i.e.* measuring the efficiency of various BDT cuts as a function of the DIRA of the two tracks using fully simulated $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays. Since this type of background is mainly due to 3-body Λ_b^0 decays containing a π^0 that goes undetected, the reconstructed invariant mass of the remaining pair of hadrons is by construction lower than the nominal Λ_b^0 mass. The higher the momentum carried away by the neutral pion, the worse the agreement between the sum of the momenta of the two hadrons and the vector joining the primary and the decay vertex of the Λ_b^0; the DIRA has a value of 1 when these two vectors are aligned, and it is lower when they are not, therefore it can be used as a discriminating variable for identifying events where an additional particle went missing in the reconstruction of the H_b candidate. This efficiency was then applied to the RapidSim samples as a weight to model the sculpting of the distribution under the different BDT cuts. The shape of the distributions for various BDT requirements are shown on the right of Fig. 3.8. The histograms are used in the fitting model to describe the shape of this component.

3.4.4 Combinatorial background model

The combinatorial background component has been modelled with an exponential function

$$c_f(m) = e^{-k_f m}, \tag{3.21}$$

where k_f is left free to vary in the fit and is different for each spectrum corresponding to the final states $\pi^+\pi^-$, $K^+\pi^-$, K^+K^-, and $p\pi^-$.

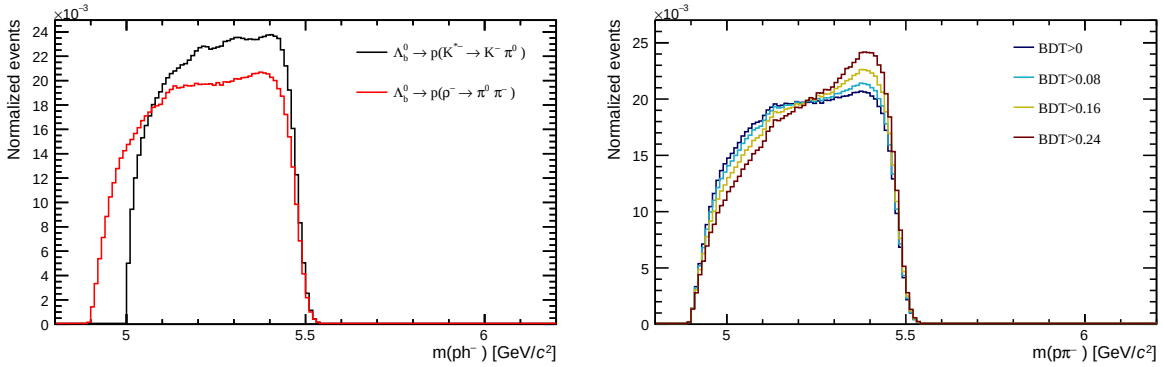

Figure 3.8: Left: invariant-mass distributions of the RapidSim samples of $\Lambda_b^0 \to p(K^{*-} \to K^-\pi^0)$ (black) and $\Lambda_b^0 \to p(\rho^- \to \pi^-\pi^0)$ (red). Right: effect of various BDT cuts on the invariant-mass distribution of $\Lambda_b^0 \to p(\rho^- \to \pi^-\pi^0)$ decays.

In the pK^- final state, the combinatorial background was found to have a turn-on shape at low mass, making it difficult for a pure exponential to fit well, especially at low BDT cuts where the sample is much more contaminated by combinatorial background. To find a more effective function to use, data in the upper sideband $m_{\pi\pi} > 5.6$ GeV/c^2 were selected to obtain an almost pure combinatorial sample; its invariant mass was recomputed under the pK^- hypothesis and shifted back by 0.8 GeV/c^2, reproducing the shape that characterizes the distribution at low invariant mass. This comes from the fact that the Stripping line has a cut at $m_{\pi\pi} > 4.8$ GeV/c^2, which results in a border effect at low mass when recomputing it in a different hypothesis (such as pK^-). Therefore, when cutting $m_{\pi\pi} > 5.6$ GeV/c^2 and recomputing the mass under the pK^- hypothesis, the shape should be reproduced by shifting the mass back by $5.6 - 4.8 = 0.8$ GeV/c^2. The resulting distribution was fitted with several test functions to find the most suitable one. The final choice was the following function:

$$c_{pK}(m) = (1 + \tanh(b(m - \delta)))e^{-k_{pK}m}, \tag{3.22}$$

in which δ is a mass shift parameter and b governs the slope of the rising edge of the function, fixed in the nominal fit to the value found with this procedure. A fit to the distribution with this p.d.f. is shown in Fig. 3.9.

3.5 Offline selection optimization

For the Run 1 analysis, an additional selection step was added offline in order to obtain the best statistical sensitivity on the *CP* asymmetries; the same strategy is applied on the Run 2 sample as well. The offline selection is composed of two distinct parts:

- a kinematic and geometrical selection applied to all the decay channels and based on a Boosted Decision Tree (BDT) multivariate algorithm;
- a specific final-state selection based on the application of PID requirements.

Note that both the selection criteria must be optimized simultaneously to obtain the combination of criteria that yields the smallest uncertainty for each *CP* asymmetry.

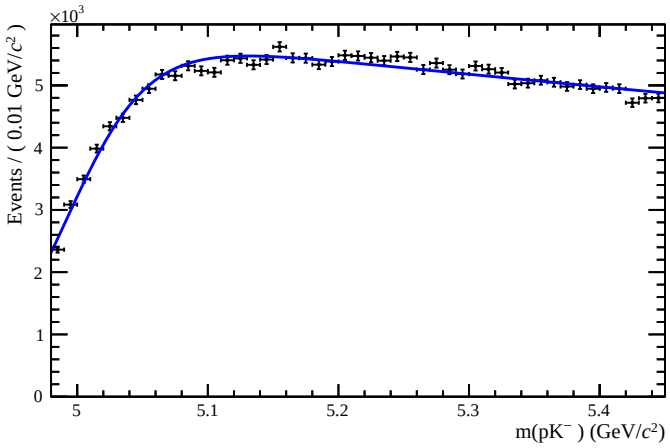

Figure 3.9: Distribution of m_{pK^-} in the upper sideband $m_{\pi\pi} > 5.6$ GeV. A fit with the function defined in Eq. (3.22) is superimposed.

Before describing the offline procedure used to optimise the offline selection criteria, it is appropriate to make some considerations:

- The kinematic and geometrical requirements imposed by the BDT selection reduce the amount of combinatorial background present in each invariant-mass spectrum, while PID requirements are needed in order to decrease the number of cross-feed (misidentified) background candidates. However, PID requirements also modify the composition and the amount of combinatorial background.

- For each set of BDT and PID requirements we need to determine the number of signal, cross-feed background, partially reconstructed background and combinatorial background candidates. The grid of BDT and PID requirements is reported in Table 3.4.

The procedure is the following: first of all, we train a BDT for each set of PID requirements, chosen in order to cover a wide region of the DLL distributions. Secondly, we select different samples of fully reconstructed pK^- and $p\pi^-$ final states, one for each combination of BDT and PID requirements. Then we perform maximum-likelihood fits to the invariant-mass spectra, determining the relevant parameters of the model. Finally, 50 pseudoexperiments for each set of requirements are performed, generating and fitting the data. The average uncertainty for each set of pseudoexperiments is then computed and the set of requirements chosen are those that give the smallest averaged uncertainty. The optimisation procedure is performed separately and independently for the $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays, as described in the next sections. In the following we will refer to the selection optimised for $\Lambda_b^0 \to pK^-$ decays as Selection A, while to that optimised for $\Lambda_b^0 \to p\pi^-$ decays as Selection B.

3.5.1 BDT training

The training of the BDT algorithm is performed considering that the application of PID modifies the amount and composition of combinatorial background. For this reason a

Table 3.4: List of the PID and BDT requirements explored during the optimisation procedure for protons, kaons, and pions.

Variables		Values used	Used in Selection
PID1 = $\mathrm{DLL}_{p\pi}(p)$	$>$	$1 \rightarrow 13$, step size 2	A,B
PID2 = $\mathrm{DLL}_{pK}(p)$	$>$	$1 \rightarrow 9$, step size 2	A,B
PID3 = $\mathrm{DLL}_{K\pi}(K)$	$>$	$1 \rightarrow 7$, step size 2	A
PID4 = $\mathrm{DLL}_{Kp}(K)$	$>$	$-\mathrm{DLL}_{pK}(p) \rightarrow -1$, step size 2	A
PID3 = $\mathrm{DLL}_{K\pi}(\pi)$	$<$	$-7 \rightarrow -1$, step size 2	B
PID4 = $\mathrm{DLL}_{p\pi}(\pi)$	$<$	$1 \rightarrow \mathrm{DLL}_{p\pi}(p)$, step size 2	B
BDT	$>$	$0 \rightarrow 0.4$, step size 0.04	A,B

training is made for each configuration of explored PID requirements. The signal sample is taken from fully simulated $\Lambda_b^0 \rightarrow pK^-$ and $\Lambda_b^0 \rightarrow p\pi^-$ events. The background sample is extracted from real data, selecting events passing the PID requirements used to isolate pK^- and $p\pi^-$ final states with an invariant mass under the $\pi\pi$ hypothesis larger than 5.6 GeV/c^2. This is done to have cleaner background samples, since cross-feed backgrounds have a long tails to the right of the pK^- and $p\pi^-$ invariant-mass spectra and thus cross-feed contributions would contaminate the right-hand sideband. In the $\pi^+\pi^-$ hypothesis, however, the physical decays are pushed to the left part of the invariant-mass spectrum and so the right-hand sideband is populated only by combinatorial background. Indeed, as can be seen in Fig. 3.10, there are no $H_b \rightarrow h^+h'^-$ decays reconstructed under the $\pi^+\pi^-$ invariant-mass hypothesis above 5.6 GeV/c^2 [79].

The variables used to train the classifier are:

- the minimum and maximum p_T of the two tracks;
- the minimum and maximum impact parameter of the two tracks computed with respect to all the PVs (d_IP);
- the minimum and maximum χ^2 of the impact parameter($\chi^2(d_\mathrm{IP})$);
- the distance of closest approach (DOCA) of the two tracks (d_CA);
- the χ^2 of the H_b candidate decay-vertex fit (χ^2_{vtx});
- the p_T of the H_b candidate ($p_\mathrm{T}^{H_b}$);
- the χ^2 of the impact parameter of the H_b candidate with respect to the associated PV ($\chi^2(d_\mathrm{IP}^{H_b})$);
- the flight distance of the H_b candidate with respect to the associated PV (FD(H_b));
- the χ^2 of the flight distance of the H_b candidate ($\chi^2(\mathrm{FD}(H_b))$).

The distributions and correlations of these variables are reported in Figs. 3.11 to 3.13 for both background and signal events with a particular set of PID requirements for illustration purposes. The BDT is trained and tested with a 3-fold validation, *i.e.* the

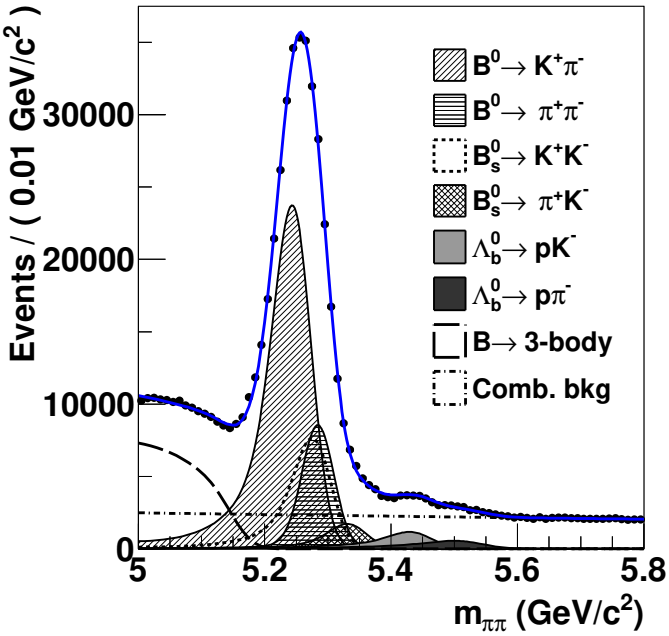

Figure 3.10: Invariant-mass distribution of $H_b \to h^+h'^-$ candidates reconstructed under the $\pi^+\pi^-$ invariant-mass hypothesis. The results of the best fit are superimposed. No PID requirements are imposed on the events, wherease a cut on the BDT classifier is requested. No $H_b \to h^+h'^-$ events are present above 5.6 GeV/c^2.

traning dataset is split into 3 parts, three identical copies of the classifier are trained on each subsample, and then they are tested on a different subsample from the one they were trained on. We report in Fig. 3.14 the distributions of the BDT score relative to the training and testing samples with a particular set of PID requirements for illustration purposes.

3.5.2 Optimization procedure

The first step of the optimisation procedure consists in determining the amount of signal, cross-feed, partially-reconstructed, and combinatorial background events surviving each combination of PID and BDT requirements. The PID requirements used to select protons for the pK^- and $p\pi^-$ final states require the DLL_{pK} and $\mathrm{DLL}_{p\pi}$ variables to be greater than a given threshold, while to choose kaons we ask $\mathrm{DLL}_{K\pi}$ and DLL_{Kp} to be greater than a given threshold. Finally, we impose $\mathrm{DLL}_{K\pi}$ and $\mathrm{DLL}_{p\pi}$ to be smaller than a given threshold when discriminating pions from kaons and protons. Note that the value of the DLL_{Kp} and $\mathrm{DLL}_{p\pi}$ variables used to select kaons and pions for pK^- and $p\pi^-$ final states are constrained to be mutually exclusive with respect to the $\mathrm{DLL}_{p\pi}$ and DLL_{pK} requirements employed to select protons. This is done in order to avoid double counting among the different final states.

We perform maximum-likelihood fits to the selected samples in order to obtain the various yields and the other relevant parameters of the fitting model; note that in this step the two *CP*-conjugate final states of each decay are not distinguished, hence no asymmetry

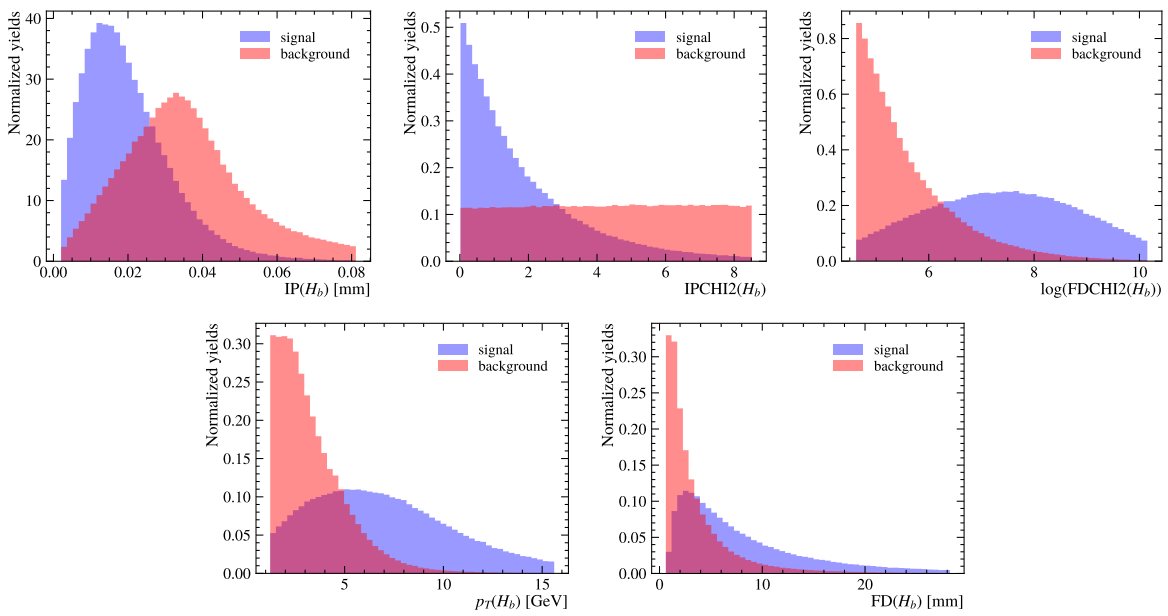

Figure 3.11: Distribution of part of the variables used in the BDT training for (blue) signal and (red) background events. The dataset corresponds to a particular set of PID cuts for Selection A.

is measured. The model used to describe the data is the one described in Section 3.4. As an example, we report in Figs. 3.15 and 3.16 two normalisation fits corresponding to the pK^- and $p\pi^-$ invariant-mass spectra.

The yields of signal, partially-reconstructed background and combinatorial background events are left free to vary in the fit procedure. The number of cross-feed background events is calculated in a different way. First of all, we consider only cross-feed background contributions to the pK^- and $p\pi^-$ invariant-mass spectra where just one final state particle is misidentified, since the amount of decays where the identity of both final state particles is wrongly assigned is expected to be negligible even with the softest PID requirements scrutinised. Thus, the cross-feed backgrounds considered in the fit model are

- $\overline{B}^0 \to \pi^+K^-$, $B_s^0 \to \pi^+K^-$, $B_s^0 \to K^+K^-$, and $\Lambda_b^0 \to p\pi^-$ decays for the $\Lambda_b^0 \to pK^-$ invariant-mass spectrum;
- $B^0 \to K^+\pi^-$, $\overline{B}_s^0 \to K^+\pi^-$, $B^0 \to \pi^+\pi^-$, and $\Lambda_b^0 \to pK^-$ decays for the $\Lambda_b^0 \to p\pi^-$ invariant-mass spectrum.

We determine the number of $B^0 \to K^+\pi^-$ ($\overline{B}^0 \to \pi^+K^-$) decays directly from the fits to the pK^- ($p\pi^-$) invariant-mass spectrum. The yields of the other cross-feed backgrounds coming from B mesons, *i.e.* $B_s^0 \to \pi^+K^-$ and $B_s^0 \to K^+K^-$ ($B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$), are constrained to the $B^0 \to K^+\pi^-$ ($B_s^0 \to \pi^+K^-$) yield, while the yields of cross-feed backgrounds coming from the other Λ_b^0 decay are constrained to the yields of the signal. The relation used to constrain the yields is

$$N_i = N_j \frac{\mathcal{B}_i}{\mathcal{B}_j} \frac{f_i}{f_j} \frac{\varepsilon_i}{\varepsilon_j}, \tag{3.23}$$

where N_i represents the yield of the considered cross-feed background, N_j represents the yield of the reference decay, $\mathcal{B}$ stands for the branching ratio, f is the hadronisation

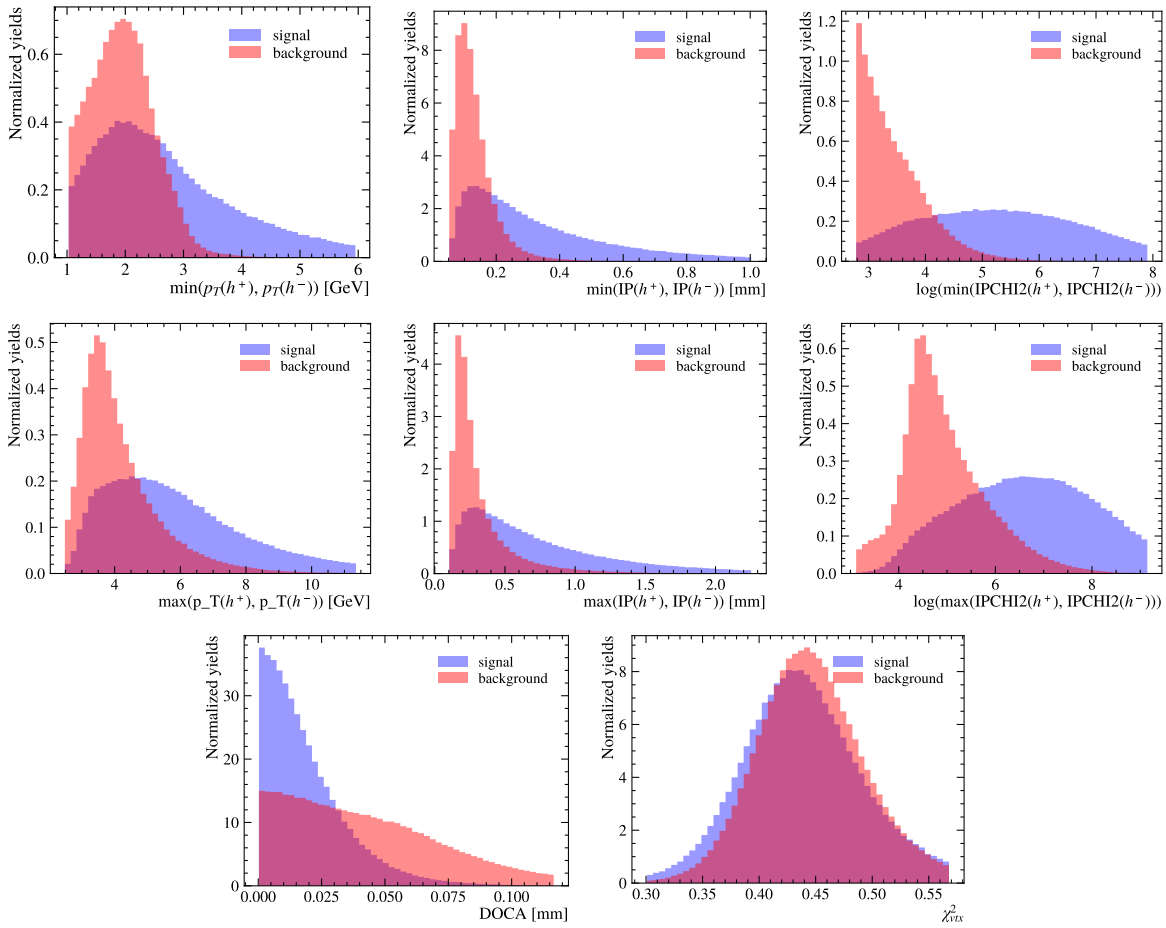

Figure 3.12: Distribution of part of the variables used in the BDT training for (blue) signal and (red) background events. The dataset corresponds to a particular set of PID cuts for Selection A.

Table 3.5: Values used in Eq. (3.23) taken from Refs. [32, 132].

Quantity	Value
$\mathcal{B}(B^0 \to K^+\pi^-)$	$(19.6 \pm 0.5) \times 10^{-6}$
$\mathcal{B}(B^0 \to \pi^+\pi^-)$	$(5.12 \pm 0.19) \times 10^{-6}$
$\mathcal{B}(B_s^0 \to K^+K^-)$	$(26.6 \pm 2.2) \times 10^{-6}$
$\mathcal{B}(B_s^0 \to \pi^+K^-)$	$(5.8 \pm 0.7) \times 10^{-6}$
$\mathcal{B}(\Lambda_b^0 \to pK^-)$	$(5.4 \pm 1.0) \times 10^{-6}$
$\mathcal{B}(\Lambda_b^0 \to p\pi^-)$	$(4.5 \pm 0.8) \times 10^{-6}$
f_s/f_d (13 TeV)	0.2539 ± 0.0079

fraction of the b hadron, and ε is the PID efficiency of the decay. The values of the branching ratios and of f_s/f_d are taken as an external input from [32] and [132], and they are reported in Table 3.5. In Fig. 3.17 are shown the values of the PID efficiencies obtained from the calibration procedure described in Section 3.3 using the best PID and BDT requirements found for the pK^- and $p\pi^-$ final states by means of the optimisation procedure.

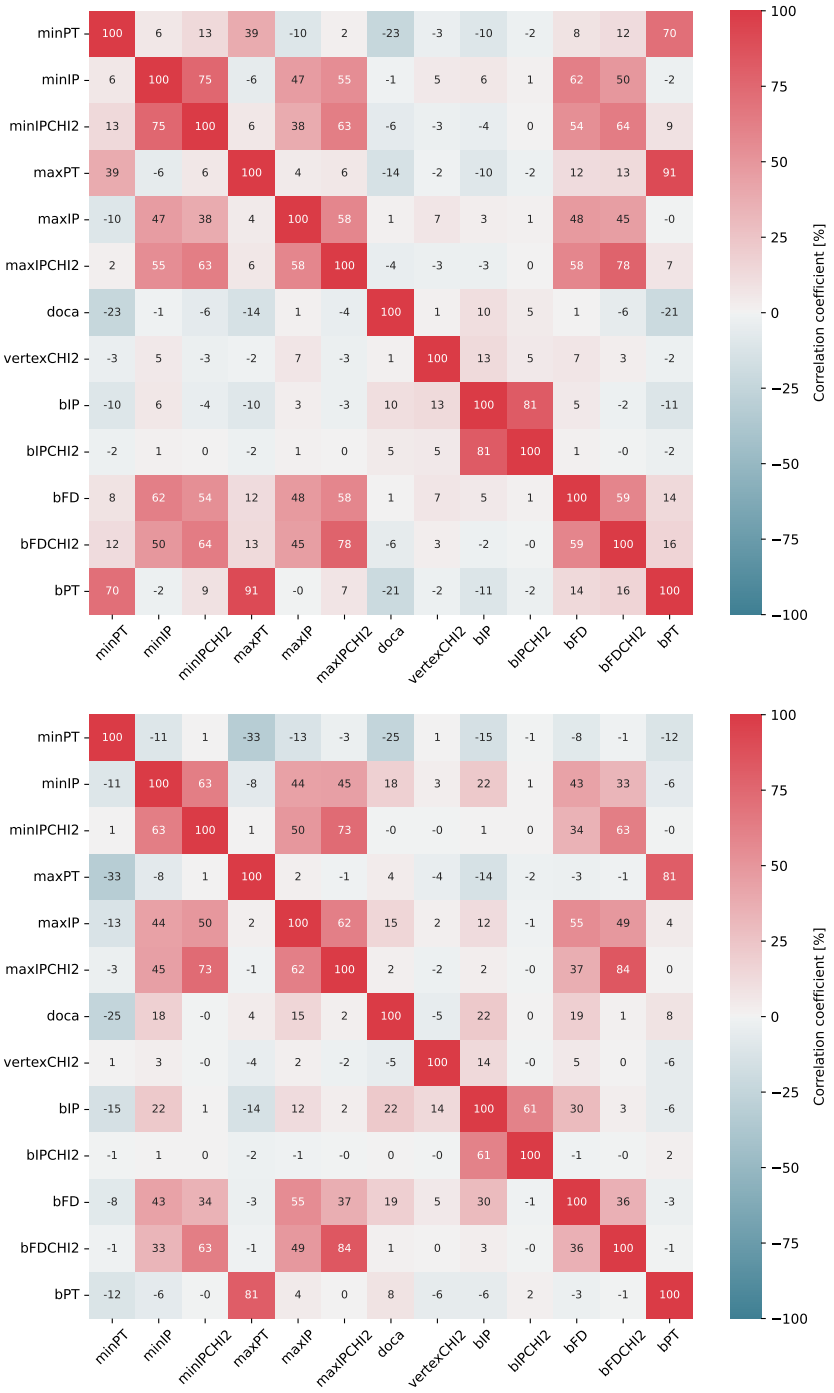

Figure 3.13: Correlation matrices of the variables used to train the BDT for (top) signal and (bottom) background samples. The dataset corresponds to a particular set of cuts for Selection A.

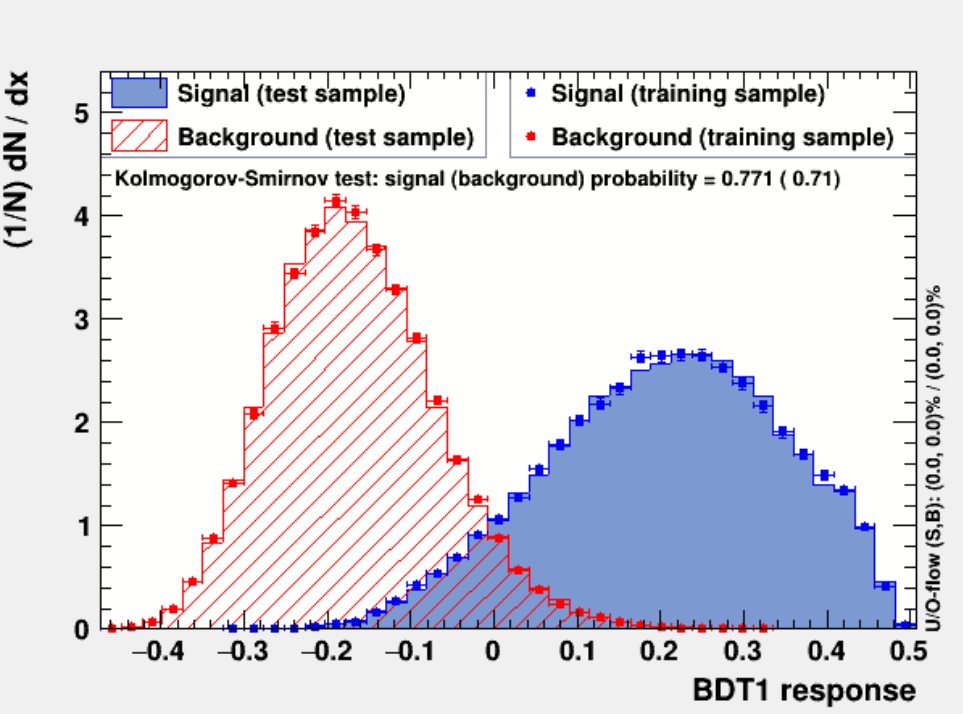

Figure 3.14: Plots showing the distributions of the BDT score relative to the (dots) training and (lines) testing samples, divided for (blue) signal and (red) background events. The dataset corresponds to a particular set of cuts for Selection A.

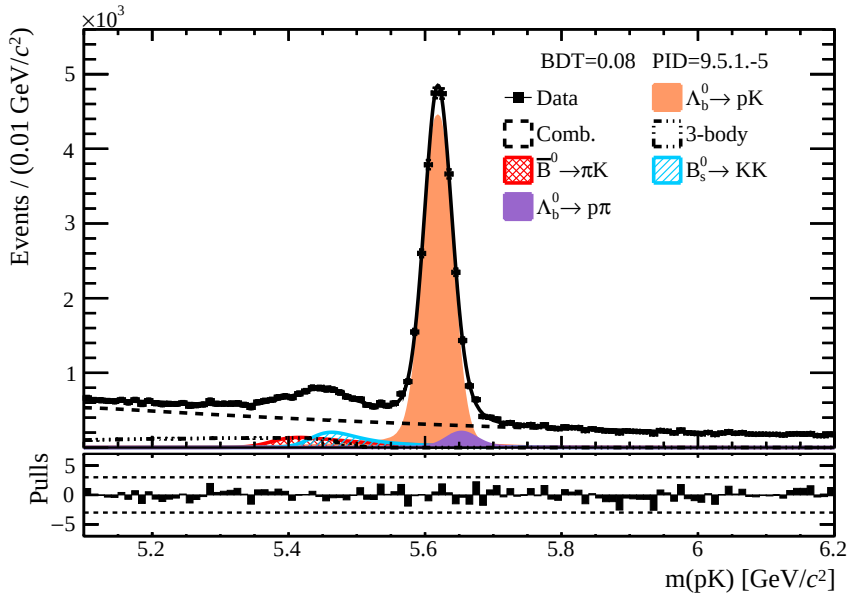

Figure 3.15: pK invariant-mass spectrum fitted with the model defined in Section 3.4 and selected requiring $\mathrm{DLL}_{p\pi}(p) > 9$, $\mathrm{DLL}_{pK}(p) > 5$, $\mathrm{DLL}_{K\pi}(K) > 1$, $\mathrm{DLL}_{Kp}(K) > -5$ and BDT > 0.08. The results of the binned maximum-likelihood fit are superimposed.

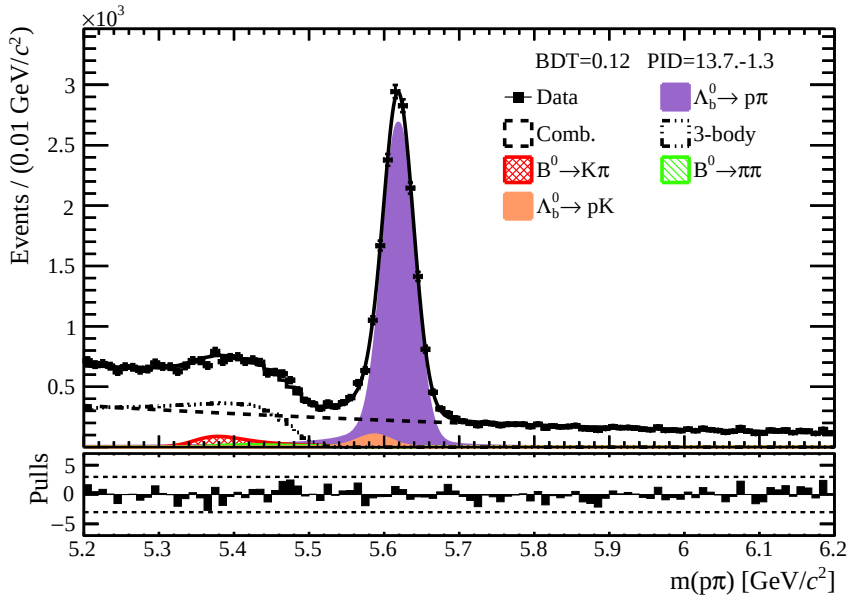

Figure 3.16: $p\pi$ invariant-mass spectrum fitted with the model defined in Section 3.4 and selected requiring $\mathrm{DLL}_{p\pi}(p) > 13$,$\mathrm{DLL}_{pK}(p) > 7$, $\mathrm{DLL}_{K\pi}(\pi) < -1$, $\mathrm{DLL}_{p\pi}(\pi) < 3$ and BDT > 0.12. The results of the binned maximum-likelihood fit are superimposed.

Then, we perform 50 pseudoexperiments for each combination of BDT and PID requirements, generating and then fitting each sample. Note that in this step we introduce the asymmetry between the two *CP*-conjugate modes for each component, generating a dataset with a null injected asymmetry and then leaving it free to vary in the fit. Toy studies are used to check that the precision on the asymmetry does not depend strongly on the value of the asymmetry itself for asymmetries lower than 10%. Indeed, toys show that, if any effect exists, it is of the order of 0.01%, that is approximately 100 time smaller of the precision we will have on the measured raw asymmetries. Finally, we take the average of the fifty uncertainties on each signal raw asymmetry for each set of BDT and PID requirements and we identify the criteria that give the smallest average of the statistical uncertainties on the asymmetry. The optimal values of the requirements found for the $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays are listed in Table 3.6.

For Run 1, we will use the same values which were found in Ref. [13] with the same strategy, and that are reported in Table 3.7.

In Fig. 3.18 are shown the histograms of the 50 values of the asymmetry errors from the toys for the two selections with the optimal cuts from Table 3.6. As it can be seen, the value of the asymmetry error is stable across the toy samples, with an RMS of less than 0.01%. We also show in Fig. 3.19 the dependence of the predicted statistical uncertainties of the two raw asymmetries on the PID and BDT requirements.

The optimised requirements for $\mathrm{DLL}_{K\pi}(K)$ and $\mathrm{DLL}_{K\pi}(\pi)$ are on the border of the probed region, and hence suggest to further loosen the requirement. However, a larger contribution of $\Lambda_b^0 \to pK^-$ decays at the $\Lambda_b^0 \to p\pi^-$ peak (and vice versa) may lead to larger systematic uncertainties not yet estimated at this level. As a consequence, the requirement on the $\Lambda_b^0 \to p\pi^-$ PID variable is not further loosened.

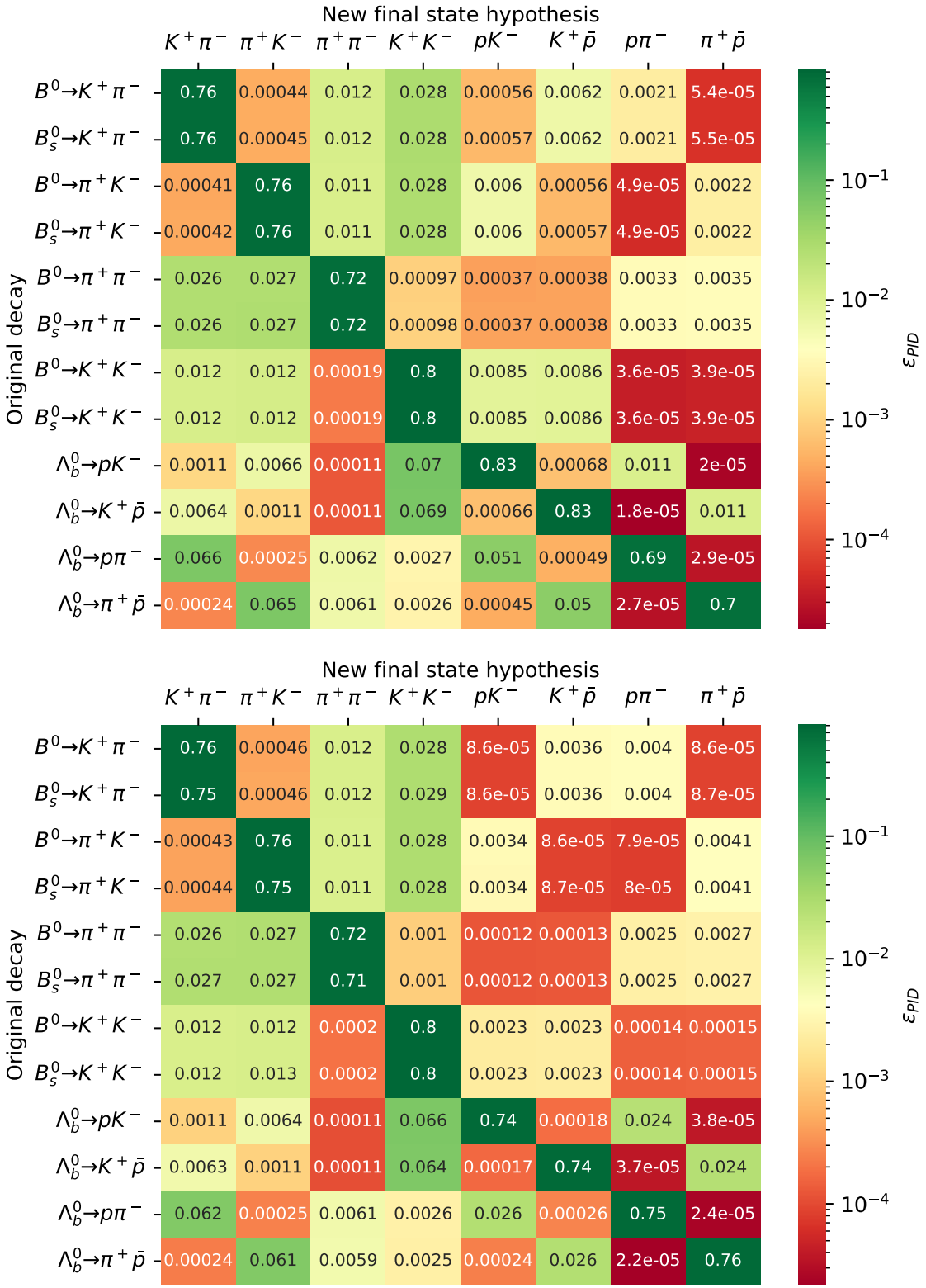

Figure 3.17: PID efficiencies of (mis)identifying a two-body decay (y-axis) as a particular final state (x-axis), computed with the method described in Section 3.3. The top and bottom tables correspond to the optimal values for Selection A and B, respectively, which are reported in Table 3.6.

Table 3.6: Optimal PID and BDT requirements found by the optimisation procedure for the $\Lambda_b^0 \to pK^-$ (Selection A) and $\Lambda_b^0 \to p\pi^-$ (Selection B) decays in Run 2, and predicted value of the raw asymmetry uncertainty.

Selection A		Selection B	
Cut	Value found	Cut	Value found
PID1 = $\mathrm{DLL}_{p\pi}(p)$	> 9	PID1 = $\mathrm{DLL}_{p\pi}(p)$	> 13
PID2 = $\mathrm{DLL}_{pK}(p)$	> 5	PID2 = $\mathrm{DLL}_{pK}(p)$	> 7
PID3 = $\mathrm{DLL}_{K\pi}(K)$	> 1	PID3 = $\mathrm{DLL}_{K\pi}(\pi)$	< -1
PID4 = $\mathrm{DLL}_{Kp}(K)$	> -5	PID4 = $\mathrm{DLL}_{p\pi}(\pi)$	< 3
BDT	> 0.08	BDT	> 0.12
Average asymmetry error	0.75%		0.99%

Table 3.7: Optimal PID and BDT requirements found by the optimisation procedure for the $\Lambda_b^0 \to pK^-$ (Selection A) and $\Lambda_b^0 \to p\pi^-$ (Selection B) decays in Run 1, as previously obtained in [13].

Selection A		Selection B	
Cut	Value found	Cut	Value found
PID1 = $\mathrm{DLL}_{p\pi}(p)$	> 11	PID1 = $\mathrm{DLL}_{p\pi}(p)$	> 11
PID2 = $\mathrm{DLL}_{pK}(p)$	> 7	PID2 = $\mathrm{DLL}_{pK}(p)$	> 7
PID3 = $\mathrm{DLL}_{K\pi}(K)$	> 0	PID3 = $\mathrm{DLL}_{K\pi}(\pi)$	< 0
PID4 = $\mathrm{DLL}_{Kp}(K)$	> -7	PID4 = $\mathrm{DLL}_{p\pi}(\pi)$	< 9
BDT	> 0.16	BDT	> 0.2

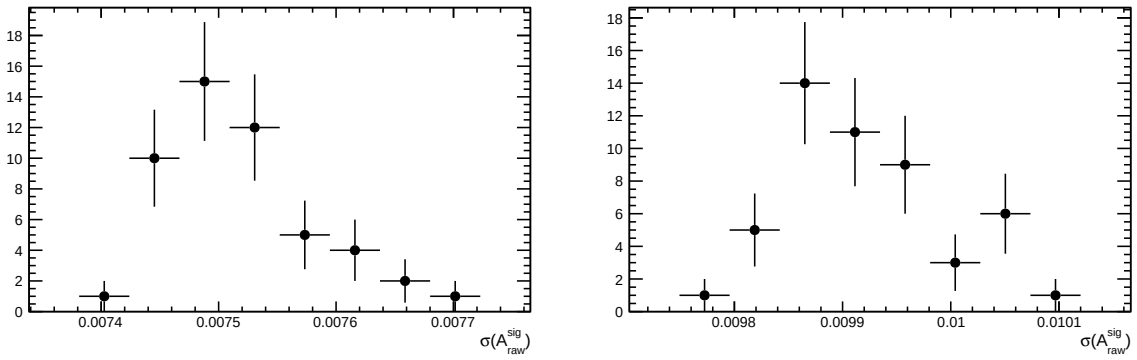

Figure 3.18: Histograms of the 50 asymmetry errors obtained from the toys for (left) Selection A and (right) Selection B with the optimal cuts.

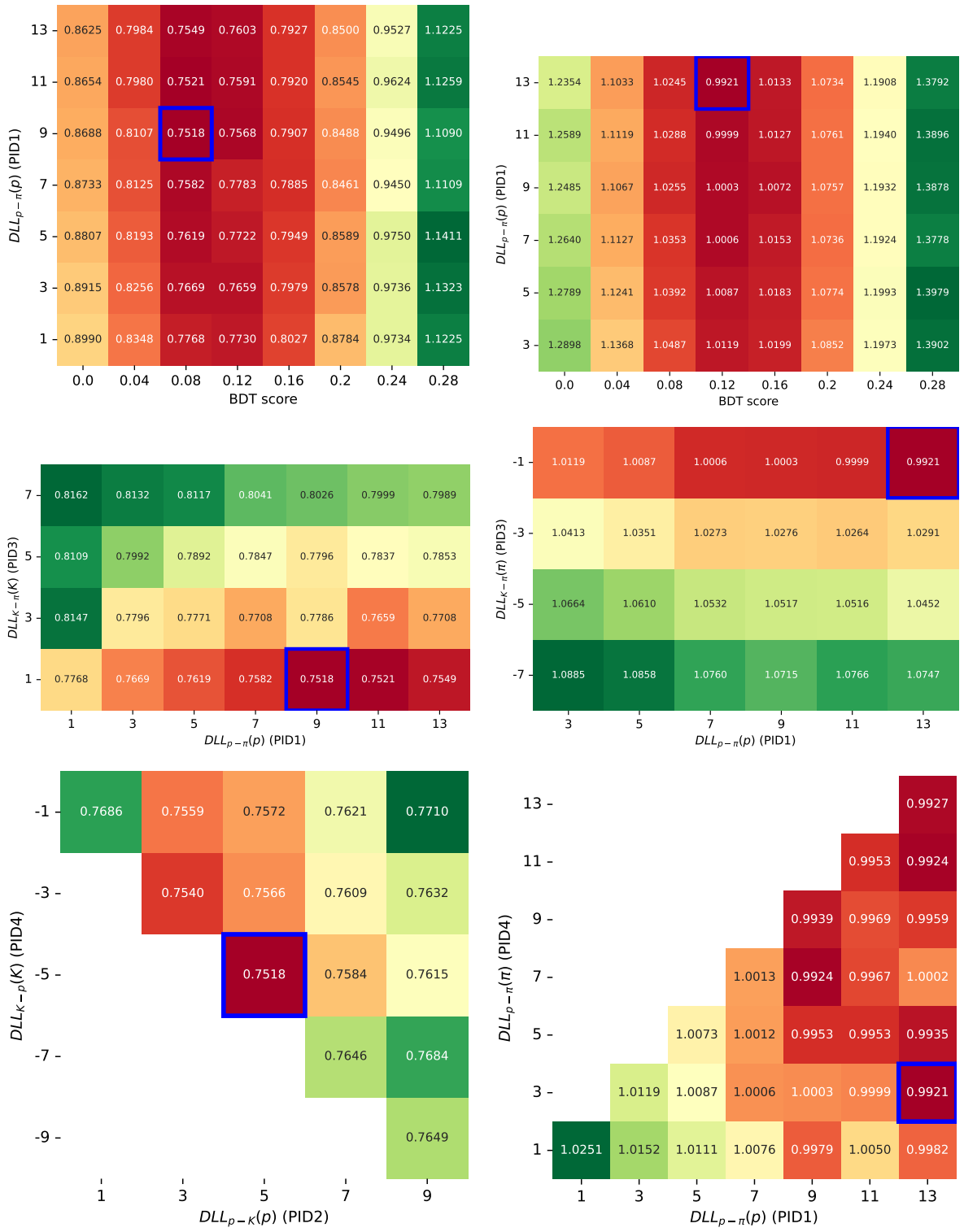

Figure 3.19: Plots showing the predicted uncertainty on (left) $A_{\text{raw}}(pK^-)$ and (right) $A_{\text{raw}}(p\pi^-)$ as a function of different combinations of BDT and PID requirements. The highlighted square indicates the optimal value found in the optimization.

Table 3.8: Fractions of the Gaussian functions contributing to the invariant-mass resolution model as used in the final fits. The ratios between the widths of the two Gaussian functions contributing to the invariant-mass resolution model for different final states, with respect to those determined for the K^+K^- spectrum, are obtained from fully simulated events.

Selection A	
Variable	Value
$f_1(B_s^0 \to K^+K^-)$	0.121 ± 0.001
$f_2(B_s^0 \to K^+K^-)$	0.69 ± 0.01
$f_1(B^0 \to K^+\pi^-)$	0.149 ± 0.002
$f_2(B^0 \to K^+\pi^-)$	0.700 ± 0.008
$f_1(B^0 \to \pi^+\pi^-)$	0.175 ± 0.002
$f_2(B^0 \to \pi^+\pi^-)$	0.709 ± 0.006
$f(\Lambda_b^0 \to pK^-)$	0.177 ± 0.002
$f(\Lambda_b^0 \to p\pi^-)$	0.213 ± 0.002
$\sigma_1(K^+\pi^-)/\sigma_1(K^+K^-)$	1.006 ± 0.005
$\sigma_2(K^+\pi^-)/\sigma_2(K^+K^-)$	1.06 ± 0.02
$\sigma_1(\pi^+\pi^-)/\sigma_1(K^+K^-)$	1.033 ± 0.005
$\sigma_2(\pi^+\pi^-)/\sigma_2(K^+K^-)$	1.033 ± 0.005
$\sigma_1(pK^-)/\sigma_1(K^+K^-)$	1.113 ± 0.005
$\sigma_1(p\pi^-)/\sigma_1(K^+K^-)$	1.116 ± 0.005

Selection B	
Variable	Value
$f_1(B_s^0 \to K^+K^-)$	0.124 ± 0.001
$f_2(B_s^0 \to K^+K^-)$	0.69 ± 0.01
$f_1(B^0 \to K^+\pi^-)$	0.151 ± 0.002
$f_2(B^0 \to K^+\pi^-)$	0.701 ± 0.009
$f_1(B^0 \to \pi^+\pi^-)$	0.180 ± 0.002
$f_2(B^0 \to \pi^+\pi^-)$	0.705 ± 0.007
$f(\Lambda_b^0 \to pK^-)$	0.178 ± 0.002
$f(\Lambda_b^0 \to p\pi^-)$	0.217 ± 0.002
$\sigma_1(K^+\pi^-)/\sigma_1(K^+K^-)$	1.006 ± 0.005
$\sigma_2(K^+\pi^-)/\sigma_2(K^+K^-)$	1.07 ± 0.02
$\sigma_1(\pi^+\pi^-)/\sigma_1(K^+K^-)$	1.032 ± 0.005
$\sigma_2(\pi^+\pi^-)/\sigma_2(K^+K^-)$	1.032 ± 0.005
$\sigma_1(pK^-)/\sigma_1(K^+K^-)$	1.109 ± 0.005
$\sigma_1(p\pi^-)/\sigma_1(K^+K^-)$	1.116 ± 0.005

3.6 Modification to the invariant-mass model

The invariant-mass model used to fit the samples surviving the optimised selection is the same described in Section 3.4. We apply a few tunings in order to achieve a better stability in the final fits to the data.

3.6.1 Constraints to the invariant-mass resolution model

In order to constrain some of the parameters governing the invariant-mass resolution model we use fully simulated events. We fix the value of the fraction ($f_{1,2}$) between the two Gaussians of the resolution model to that observed in simulated $B_s^0 \to K^+K^-$ events, and this parameter is common between all the spectra. In addition, in order to take into account the dependence of σ_1 and σ_2 from the PID requirements, we fix the ratios $\sigma_{1(2)}(K^+\pi^-)/\sigma_{1(2)}(K^+K^-)$, $\sigma_{1(2)}(\pi^+\pi^-)/\sigma_{1(2)}(K^+K^-)$, $\sigma_1(pK^-)/\sigma_1(K^+K^-)$ and $\sigma_1(p\pi^-)/\sigma_1(K^+K^-)$ to the values determined from simulation. The values of the fractions and the ratio of the widths of the Gaussian functions obtained from the fit to fully simulated events are reported in Table 3.8.

3.6.2 Determination of the yields of cross-feed background

In contrast to the fits presented in Section 3.5, now all the cross-feed components are considered in the mass model. The amount of these contributions is constrained to the amount of the same decay where the final state is correctly identified. For example, the

yield of misidentified $B^0 \to K^+\pi^-$ decays in the K^+K^- spectrum is obtained from

$$N\left(B^0 \to K^+\pi^- | K^+K^-\right) = N\left(B^0 \to K^+\pi^- | K^+\pi^-\right) \frac{\varepsilon\left(B^0 \to K^+\pi^- | K^+K^-\right)}{\varepsilon\left(B^0 \to K^+\pi^- | K^+\pi^-\right)}, \quad (3.24)$$

where $N\left(B^0 \to K^+\pi^- | K^+K^-\right)$ is the number of $B^0 \to K^+\pi^-$ decays present in the K^+K^- mass spectrum, $N\left(B^0 \to K^+\pi^- | K^+\pi^-\right)$ is the number of correctly identified $B^0 \to K^+\pi^-$ decays in the $K^+\pi^-$ spectrum, $\varepsilon\left(B^0 \to K^+\pi^- | K^+K^-\right)$ is the probability to assign the K^+K^- hypothesis to a $B^0 \to K^+\pi^-$ decay, and $\varepsilon\left(B^0 \to K^+\pi^- | K^+\pi^-\right)$ is the probability to assign the correct mass hypothesis to a $B^0 \to K^+\pi^-$ decay. The various PID efficiencies used to compute the yields of cross-feed decays are shown in Fig. 3.17.

3.7 Results of invariant-mass fits

In this Section the results of the simultaneous invariant-mass fits to the $H_b \to h^+h'^-$ samples performed to measure $A_{\rm raw}(pK^-)$ and $A_{\rm raw}(p\pi^-)$ are presented. The binned maximum-likelihood fits features 189 parameters, of which 44 are free:

- four raw asymmetries for the $H_b \to h^+h'^-$ modes: $A_{\rm raw}(B^0 \to K^+\pi^-)$, $A_{\rm raw}(B_s^0 \to \pi^+K^-)$, $A_{\rm raw}(\Lambda_b^0 \to pK^-)$, $A_{\rm raw}(\Lambda_b^0 \to p\pi^-)$;
- three raw asymmetries for the combinatorial backgrounds relative to the $K^+\pi^-$, pK^- and $p\pi^-$ final states;
- four raw asymmetries for the three-body partially reconstructed background components in the $K^+\pi^-$, pK^-, $p\pi^-$ invariant-mass spectra; the number of parameters results to be greater than the number of final states since for the $K^+\pi^-$ sample two sources of partially reconstructed backgrounds are considered, one from B^0 decays and the other from B_s^0 decays;
- eight yields for the signal $H_b \to h^+h'^-$ decays: $B^0 \to K^+\pi^-$, $B^0 \to K^+K^-$, $B^0 \to \pi^+\pi^-$, $B_s^0 \to \pi^+K^-$, $B_s^0 \to K^+K^-$, $B_s^0 \to \pi^+\pi^-$, $\Lambda_b^0 \to pK^-$, $\Lambda_b^0 \to p\pi^-$;
- five yields for the combinatorial background relative to $K^+\pi^-$, K^+K^-, $\pi^+\pi^-$, pK^- and $p\pi^-$ final states;
- six yields for the three-body partially reconstructed background components contributing to the $K^+\pi^-$, K^+K^-, $\pi^+\pi^-$, pK^- and $p\pi^-$ final states; as for the raw asymmetries, the $K^+\pi^-$ spectrum is parameterized using two yields instead of one;
- three parameters governing the exponential tail of the ARGUS p.d.f. that models the three-body partially reconstructed background in the $K^+\pi^-$, K^+K^-, $\pi^+\pi^-$ invariant-mass spectra; in this case the parameter governing the ARGUS shape in the $K^+\pi^-$ is in common between B^0 and B_s^0 modes;
- three mean values (defined as μ in Eq. (3.16)), of the Gaussian functions describing the invariant mass resolution; the means are three since we are considering three different kinds of b hadrons: B^0, B_s^0 and Λ_b^0;

Table 3.9: PID selection criteria for the $K^+\pi^-$, K^+K^-, and $\pi^+\pi^-$ mass hypotheses. To obtain the charge conjugate final states, h^+ and h^- must be exchanged.

$K^+\pi^-$		K^+K^-		$\pi^+\pi^-$	
Cut	Value	Cut	Value	Cut	Value
$\mathrm{DLL}_{K\pi}(h^+)$	> 3	$\mathrm{DLL}_{K\pi}(h^+)$	> 3	$\mathrm{DLL}_{K\pi}(h^+)$	< -3
$\mathrm{DLL}_{Kp}(h^+)$	> -5	$\mathrm{DLL}_{Kp}(h^+)$	> -5	$\mathrm{DLL}_{p\pi}(h^+)$	< 5
$\mathrm{DLL}_{K\pi}(h^-)$	< -3	$\mathrm{DLL}_{K\pi}(h^-)$	> 3	$\mathrm{DLL}_{K\pi}(h^-)$	< -3
$\mathrm{DLL}_{p\pi}(h^-)$	< 5	$\mathrm{DLL}_{Kp}(h^-)$	> -5	$\mathrm{DLL}_{p\pi}(h^-)$	< 5

Table 3.10: PID selection criteria applied for the pK^- and $p\pi^-$ mass hypotheses when the selection is optimised for the other final state. To obtain the charge conjugate final states, h^+ and h'^- must be exchanged.

Selection B		Selection A	
pK^-		$p\pi^-$	
Cut	Value	Cut	Value
$\mathrm{DLL}_{p\pi}(h^+)$	> 10	$\mathrm{DLL}_{p\pi}(h^+)$	> 10
$\mathrm{DLL}_{pK}(h^+)$	> 10	$\mathrm{DLL}_{pK}(h^+)$	> 10
$\mathrm{DLL}_{K\pi}(h'^-)$	> 3	$\mathrm{DLL}_{K\pi}(h'^-)$	< -3
$\mathrm{DLL}_{Kp}(h'^-)$	> -5	$\mathrm{DLL}_{p\pi}(h'^-)$	< 5

- two standard deviations relative to the Gaussian functions composing the mass resolution for the K^+K^- invariant-mass spectra;
- five exponential slopes for the combinatorial background relative to the $K^+\pi^-$, K^+K^-, $\pi^+\pi^-$, pK^- and $p\pi^-$;
- one additional parameter (defined as b in Eq. (3.22)) to describe the rise of the turn-on point of the combinatorial background in the pK^- spectrum at low invariant mass;

The PID selection criteria for the $K^+\pi^-$, K^+K^- and $\pi^+\pi^-$ final states are taken from the previous analysis [13] and are reported in Table 3.9. The PID requirements used for the pK^- and $p\pi^-$ final states when the selection is optimised for the $p\pi^-$ and pK^- final states are the same as the previous analysis [13] and reported in Table 3.10. After the fit, the sPlot technique [123] is employed to obtain signal weights for later use.

In Table 3.11 the signal yields and the raw asymmetries for all the $H_b \to h^+h'^-$ decays are reported as obtained from the invariant-mass fits to the data sample selected using Selection A on Run 1 and 2 data. In Figs. 3.20 and 3.21 the fits to all the invariant-mass spectra for Run 1 and 2 after the requirements imposed by selection A are shown.

In Table 3.12 the signal yields and the raw asymmetries for all the $H_b \to h^+h'^-$ decays are reported, as obtained from the invariant-mass fits to the data sample selected using Selection B on Run 1 and Run 2 data. In Figs. 3.22 and 3.23 the fits to all the invariant-mass spectra in Run 1 and 2 after the requirements imposed by selection B are

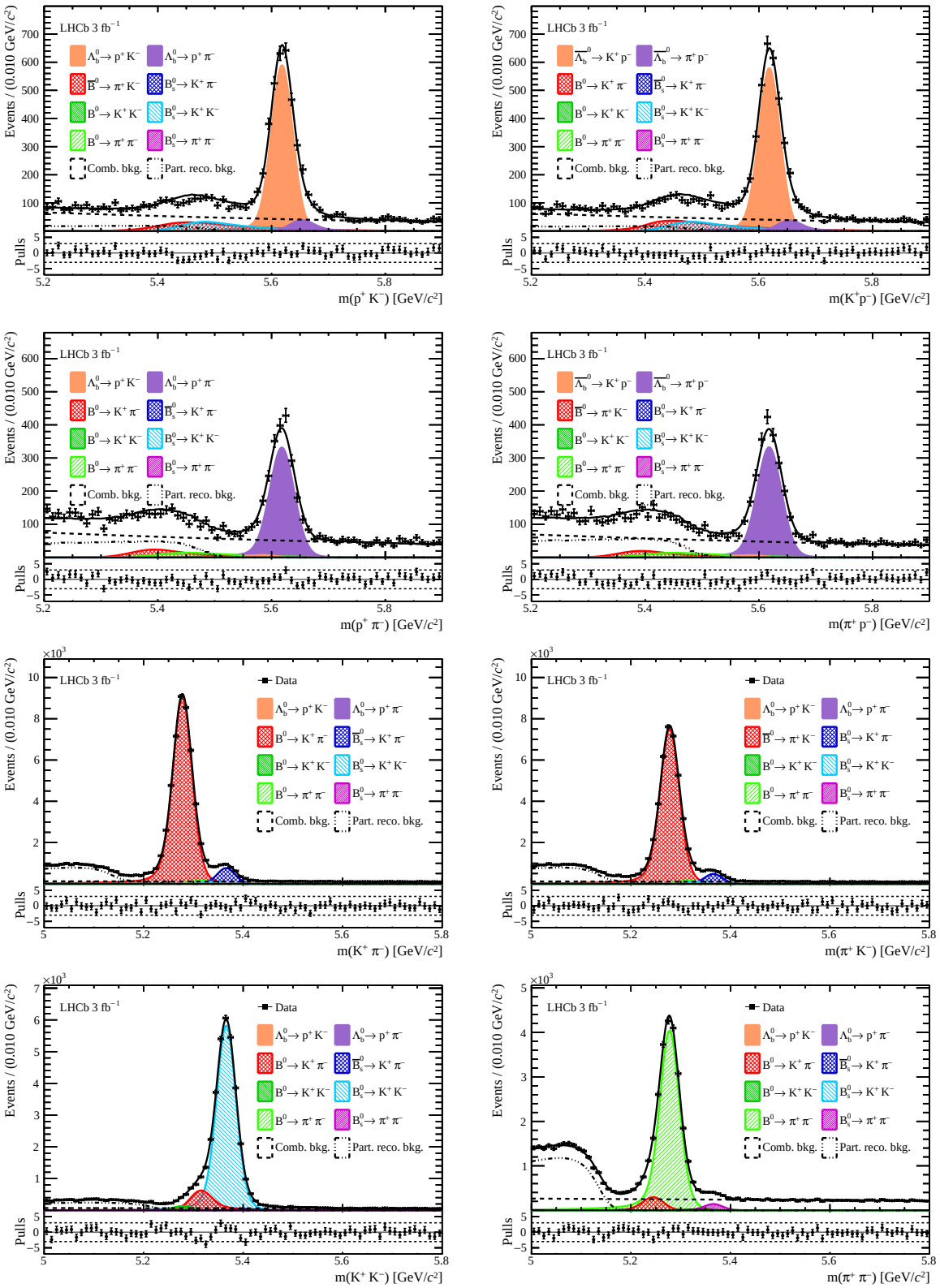

Figure 3.20: Run 1 invariant mass distributions of the (top left) pK^-, (top right) $K^+\overline{p}$, (second row left) $p\pi^-$, (second row right) $\pi^+\overline{p}$, (third row left) $K^+\pi^-$, (third row right) π^+K^-, (bottom left) K^+K^-, and (bottom right) $\pi^+\pi^-$ spectra after selection A. The curves superimposed to the data points represent the result of the best fit.

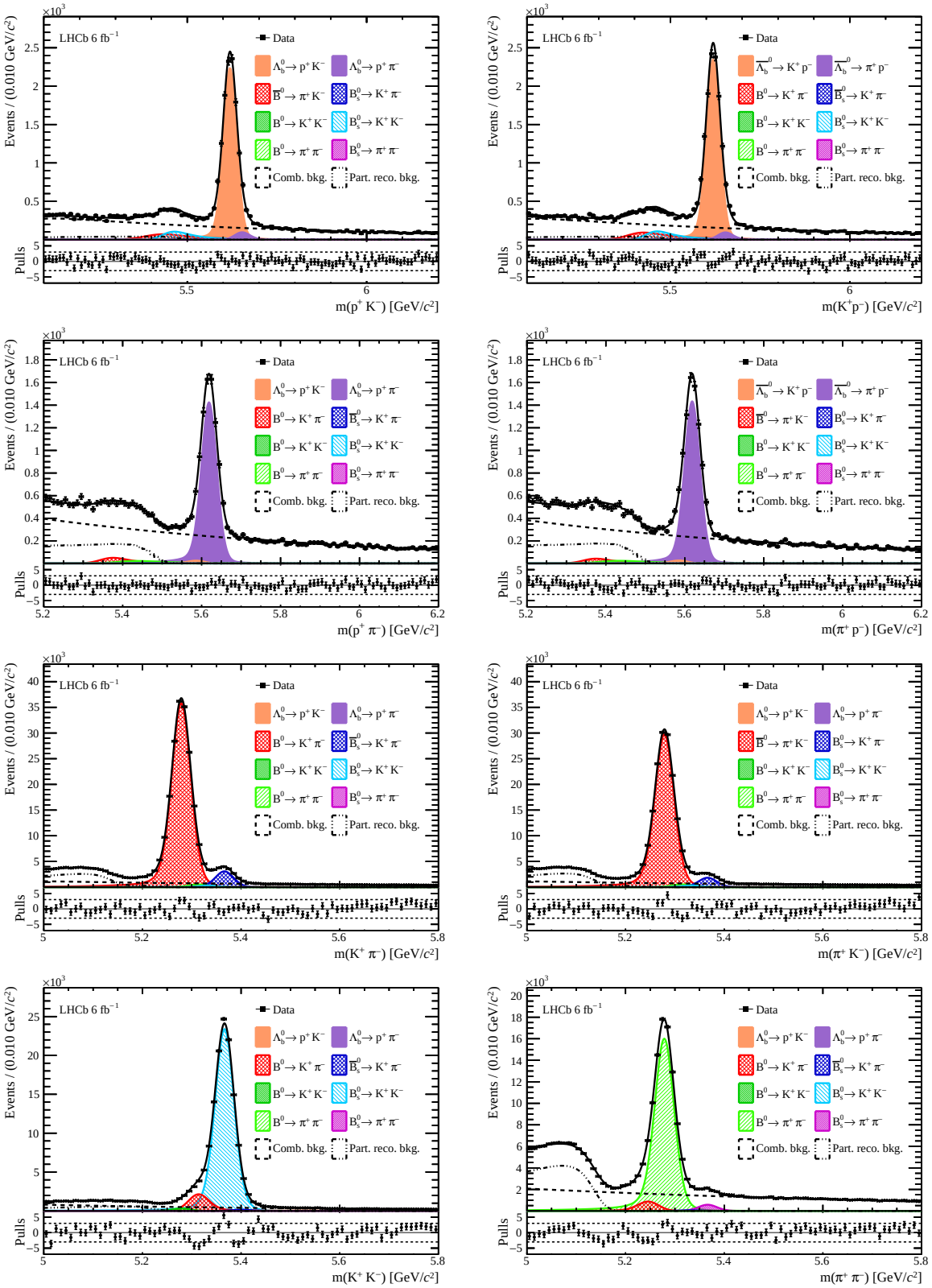

Figure 3.21: Run 2 invariant mass distributions of the (top left) pK^-, (top right) $K^+\overline{p}$, (second row left) $p\pi^-$, (second row right) $\pi^+\overline{p}$, (third row left) $K^+\pi^-$, (third row right) π^+K^-, (bottom left) K^+K^-, and (bottom right) $\pi^+\pi^-$ spectra after selection A. The curves superimposed to the data points represent the result of the best fit.

Table 3.11: Values of signal yields and raw asymmetries obtained from the invariant-mass fits with Selection A in (left) Run 1 and (right) Run 2 data.

	Run 1	Run 2
Parameter	Fit result	
Signal yields		
$N_{\rm sig}(B^0 \to K^+\pi^-)$	87161 ± 330	343917 ± 680
$N_{\rm sig}(B^0 \to K^+K^-)$	667 ± 66	1631 ± 138
$N_{\rm sig}(B^0 \to \pi^+\pi^-)$	22178 ± 178	87081 ± 364
$N_{\rm sig}(B_s^0 \to \pi^+K^-)$	6063 ± 118	24730 ± 257
$N_{\rm sig}(B_s^0 \to K^+K^-)$	30631 ± 193	121273 ± 389
$N_{\rm sig}(B_s^0 \to \pi^+\pi^-)$	728 ± 62	2928 ± 137
$N_{\rm sig}(\Lambda_b^0 \to pK^-)$	5867 ± 92	23150 ± 179
$N_{\rm sig}(\Lambda_b^0 \to p\pi^-)$	3887 ± 79	15275 ± 157
Asymmetries [%]		
$A_{\rm raw}(B^0 \to K^+\pi^-)$	-9.200 ± 0.358	-9.213 ± 0.183
$A_{\rm raw}(B_s^0 \to \pi^+K^-)$	22.65 ± 1.71	25.309 ± 0.895
$A_{\rm raw}(\Lambda_b^0 \to pK^-)$	0.86 ± 1.53	-2.424 ± 0.749
$A_{\rm raw}(\Lambda_b^0 \to p\pi^-)$	-0.13 ± 1.94	-0.351 ± 0.984

shown. In Fig. 3.24 the raw asymmetries in all year and magnet polarity subsamples are shown, both for Selection A and B.

3.7.1 Fast toy studies

To validate the fit, we repeat it 1000 times on a dataset sampled from the original distribution and measure the mean and standard deviation of the resulting $A_{\rm raw}$ distribution to check the presence of any bias. In Fig. 3.25 the distributions of the pulls of the 1000 bootstrapped $A_{\rm raw}(pK^-)$ and $A_{\rm raw}(p\pi^-)$ are shown for the corresponding selections. As it can be seen, the fit model does not introduce any bias on the measured value of $A_{\rm raw}$.

3.8 Fit model systematic uncertainties

In this Section the studies performed in order to assess the main systematic uncertainties affecting the determination of the raw asymmetries of $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays are presented. All uncertainties are estimated by generating 1000 pseudoexperiments sampled from the original distribution, fitting them both with the baseline model and with an alternative model, and looking at the distribution of the differences between the raw asymmetries obtained with the two models; the sum in quadrature of the mean and standard deviation of each distribution is taken as the systematic uncertainty related to the corresponding model modification. The pseudoexperiments are performed running together all the subsamples, hence the systematic uncertainties will be applied to the final average of A_{CP}.

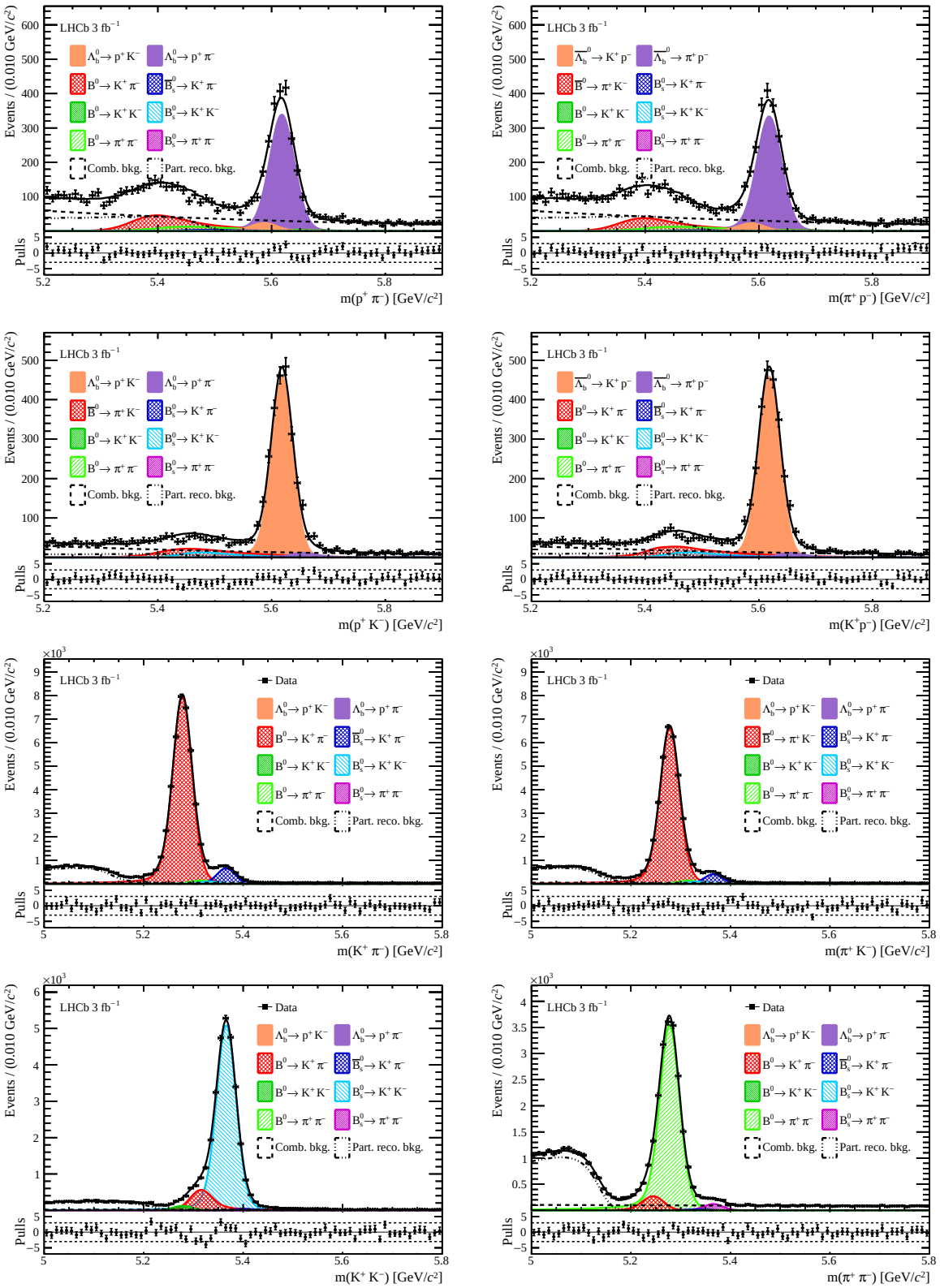

Figure 3.22: Run 1 invariant mass distributions of the (top left) $p\pi^-$, (top right) $\pi^+\overline{p}$, (second row left) pK^-, (second row right) $K^+\overline{p}$, (third row left) $K^+\pi^-$, (third row right) π^+K^-, (bottom left) K^+K^-, and (bottom right) $\pi^+\pi^-$ spectra after selection B. The curves superimposed to the data points represent the result of the best fit.

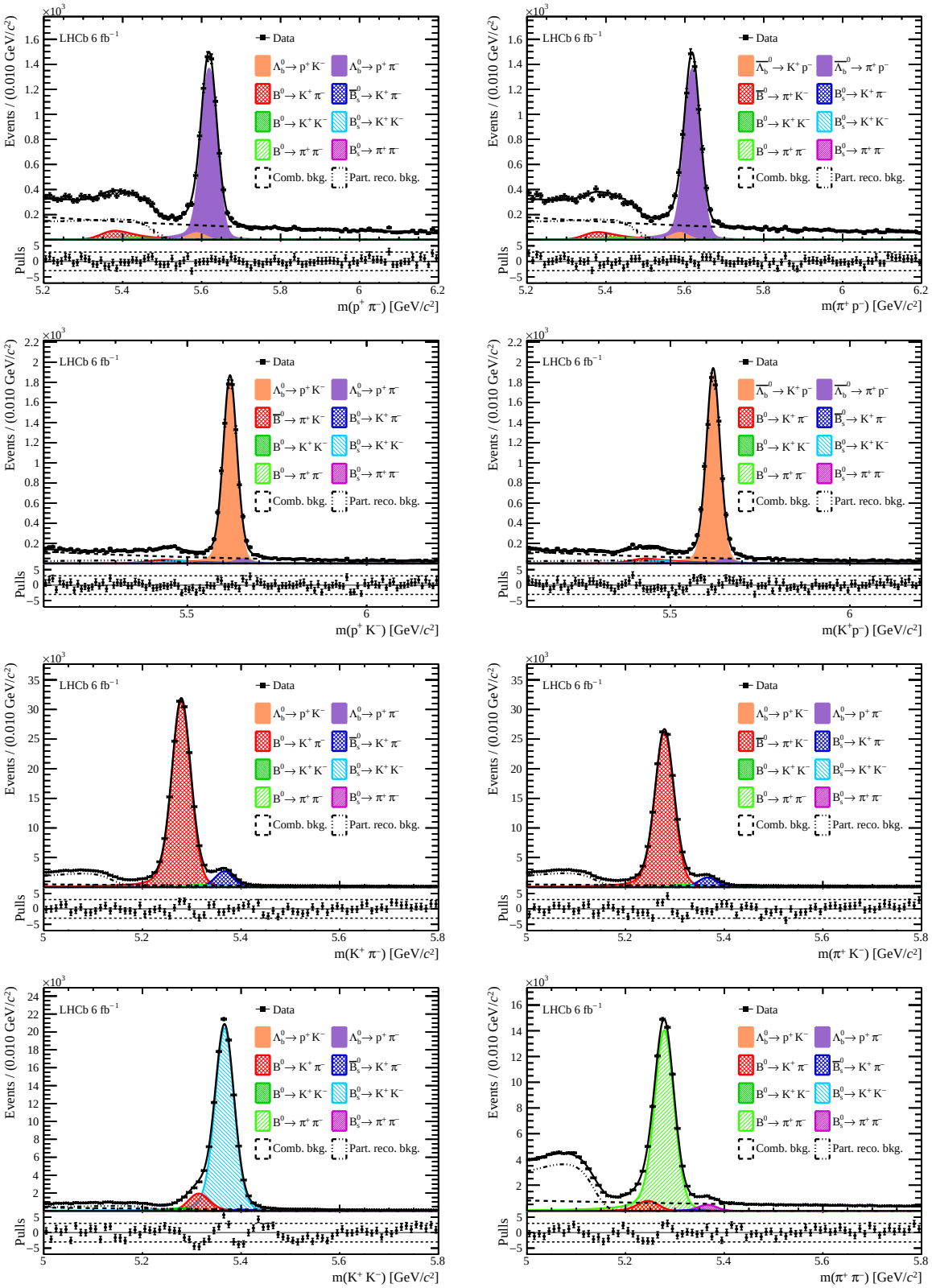

Figure 3.23: Run 2 invariant mass distributions of the (top left) $p\pi^-$, (top right) $\pi^+\overline{p}$, (second row left) pK^-, (second row right) $K^+\overline{p}$, (third row left) $K^+\pi^-$, (third row right) π^+K^-, (bottom left) K^+K^-, and (bottom right) $\pi^+\pi^-$ spectra after selection B. The curves superimposed to the data points represent the result of the best fit.

Table 3.12: Values of signal yields and raw asymmetries obtained from the invariant-mass fits with Selection B in (left) Run 1 and (right) Run 2 data.

	Run 1	Run 2
Parameter	Fit result	
Signal yields		
$N_{\rm sig}(B^0 \to K^+\pi^-)$	76962 ± 302	303217 ± 610
$N_{\rm sig}(B^0 \to K^+K^-)$	608 ± 55	1567 ± 119
$N_{\rm sig}(B^0 \to \pi^+\pi^-)$	19552 ± 160	76572 ± 322
$N_{\rm sig}(B^0_s \to \pi^+K^-)$	5320 ± 103	21879 ± 219
$N_{\rm sig}(B^0_s \to K^+K^-)$	27030 ± 178	106413 ± 359
$N_{\rm sig}(B^0_s \to \pi^+\pi^-)$	625 ± 47	2531 ± 103
$N_{\rm sig}(\Lambda^0_b \to pK^-)$	4606 ± 75	18521 ± 148
$N_{\rm sig}(\Lambda^0_b \to p\pi^-)$	3971 ± 77	14641 ± 144
Asymmetries [%]		
$A_{\rm raw}(B^0 \to K^+\pi^-)$	-9.052 ± 0.377	-9.050 ± 0.192
$A_{\rm raw}(B^0_s \to \pi^+K^-)$	23.70 ± 1.72	24.418 ± 0.869
$A_{\rm raw}(\Lambda^0_b \to pK^-)$	-0.16 ± 1.62	-1.941 ± 0.794
$A_{\rm raw}(\Lambda^0_b \to p\pi^-)$	0.90 ± 1.89	0.143 ± 0.949

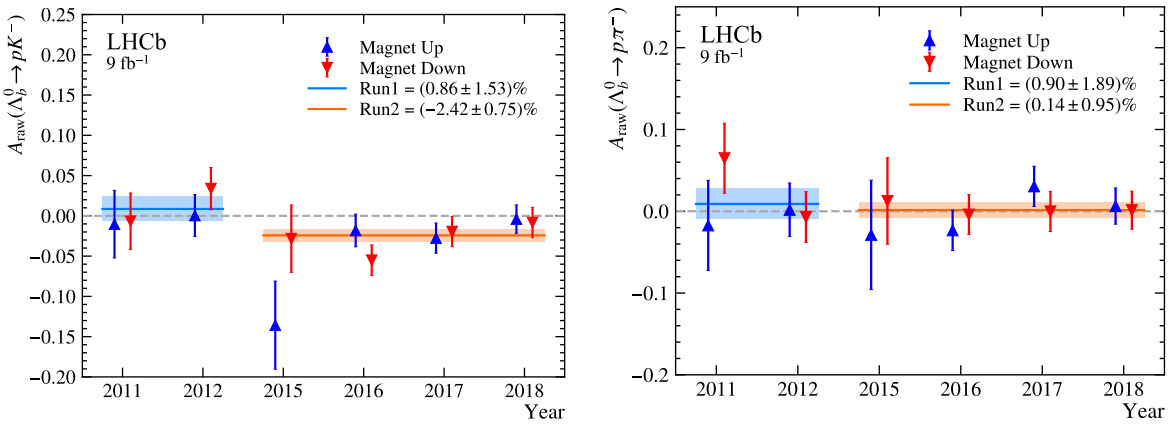

Figure 3.24: Values of the raw asymmetries of (left) $\Lambda^0_b \to pK^-$ and (right) $\Lambda^0_b \to p\pi^-$ decays, divided by year and magnet polarity.

Signal shape To assess the impact of the choice of the function to model signal events, we modify the p.d.f. by switching from the sum of a Johnson S_U function and one or two Gaussians to the sum of a double-sided Crystal Ball function and one Gaussian. The shape parameters of the Crystal Ball are obtained from fits to simulated sample, as done for the baseline Johnson function, and fixed in the fit;

Combinatorial shape To assess the impact of the choice of the function to model combinatorial background events, we modify the p.d.f. by switching from an exponential function to a second-order Chebychev polynomial;

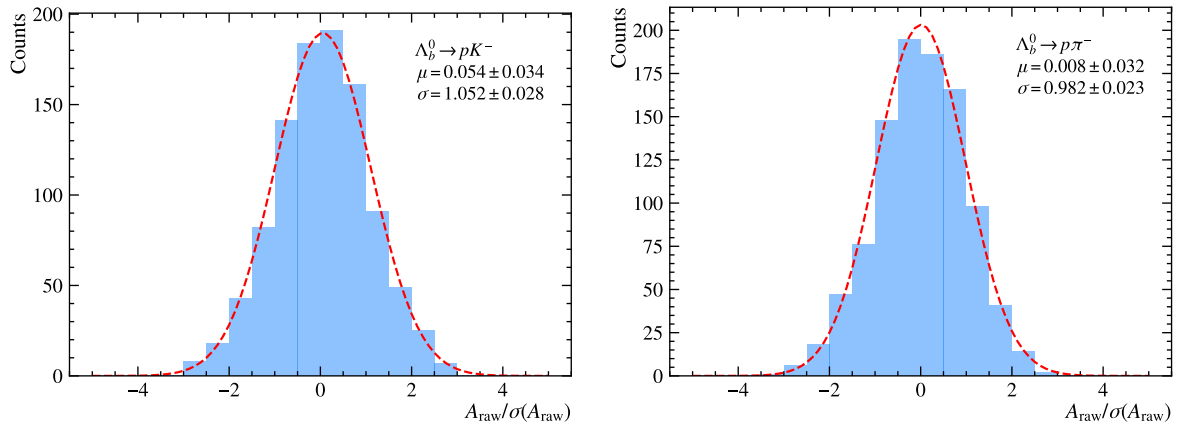

Figure 3.25: Distribution of the 1000 values of $A_{\rm raw}$ from the toy studies for (left) $\Lambda_b^0 \to pK^-$ and (right) $\Lambda_b^0 \to p\pi^-$ decays. The distributions are shifted by the nominal value of $A_{\rm raw}$.

Cross-feed shape To assess the impact of the choice of the function to model cross-feed background events, we fit the generated data with the same templates used in the baseline model without applying the weights to correct for the PID requirements (see Fig. 3.7 for a comparison of the shapes with and without the weights);

Partially reconstructed shape To assess the impact of the choice of the function to model partially reconstructed background events, we modify the p.d.f. by switching from the templates obtained from fast simulation (see Section 3.4.3) to an ARGUS function. This change is only applied to the pK^-, $K^+\overline{p}$, $p\pi^-$, and $\pi^+\overline{p}$ invariant-mass spectra, as ARGUS functions are already used in the other 4 spectra; since modifying the shape of partially reconstructed decays in non-signal spectra is not expected to change significantly the yields of signal events in the target spectra, we don't apply any modification to the shapes of partially reconstructed decays in the K^+K^-, $\pi^+\pi^-$, $K^+\pi^-$, and π^+K^- spectra;

PID efficiencies To assess the impact of the use of PID efficiencies to scale the yields of cross-feed decays from those of signal events in all invariant-mass spectra, we fit the generated samples with a set of PID efficiencies sampled from Gaussian functions of means equal to the nominal values and widths equal to the nominal errors;

The results of the pseudoexperiments on the $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ samples are shown in Figs. 3.26 and 3.27 and reported in Table 3.13, together with the sum in quadrature of each term, to be used as a total systematic uncertainty related to the fit model.

3.9 Interaction asymmetries

In order to get the physical *CP* asymmetries, one needs to subtract from the raw asymmetries different sources of experimental asymmetries, as evidenced in Eqs. (3.3) and (3.4). In this section we will discuss only the determination of interaction asymmetries ($A_{\rm D}$) of protons, kaons, and pions. We will describe the PID-induced and the trigger-induced asymmetries in more detail in Sections 3.10 and 3.11, respectively.

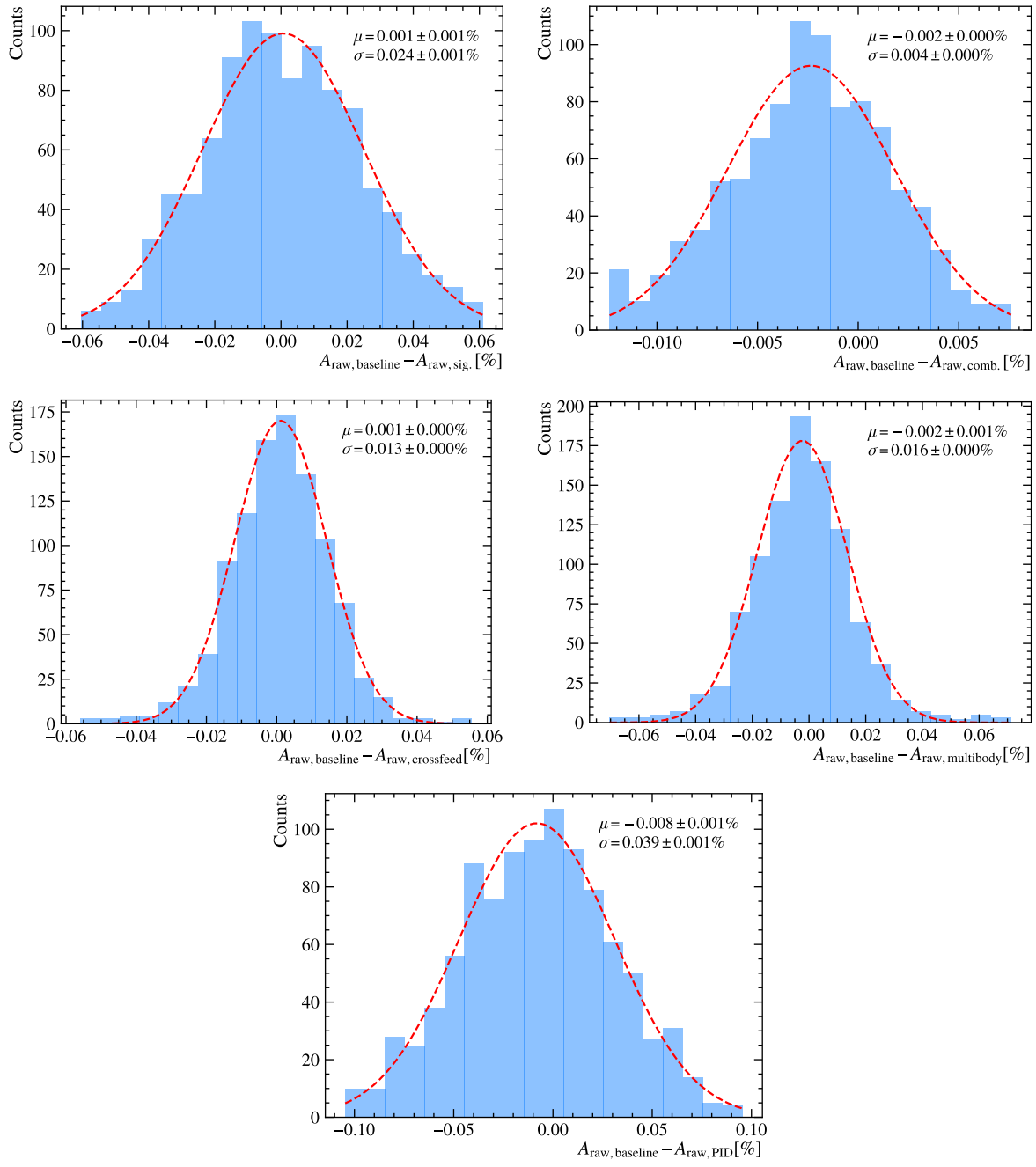

Figure 3.26: Distributions of the difference of the 1000 baseline and modified $A_{\text{raw}}(pK^-)$ from the pseudoexperiments for the estimation of systematic uncertainties: (top left) signal shape, (top right) combinatorial shape, (middle left) crossfeed shape, (middle right) partially reconstructed shape, and (bottom) PID efficiencies.

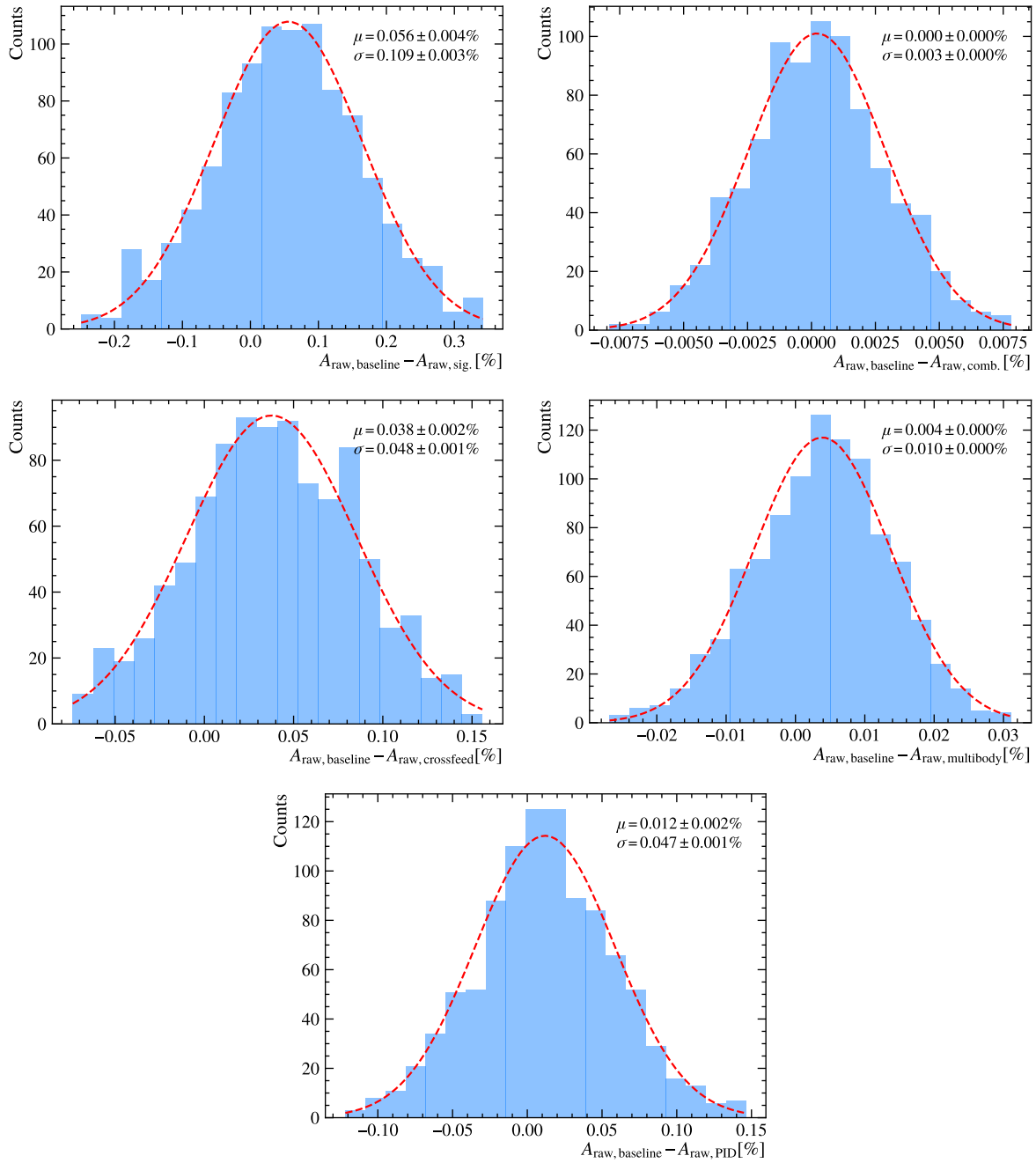

Figure 3.27: Distributions of the difference of the 1000 baseline and modified $A_{\text{raw}}(p\pi^-)$ from the pseudoexperiments for the estimation of systematic uncertainties: (top left) signal shape, (top right) combinatorial shape, (middle left) crossfeed shape, (middle right) partially reconstructed shape, and (bottom) PID efficiencies.

Table 3.13: Values of the systematic uncertainties due to the fit model estimated with $\Lambda_b^0 \to pK^-$ decays, divided by source. The last row reports the sum in quadrature of the terms, providing a total value of the systematic uncertainty.

Source	$\Lambda_b^0 \to pK^-$	$\Lambda_b^0 \to p\pi^-$
Signal shape	0.024%	0.123%
Combinatorial shape	0.004%	0.003%
Cross-feed shape	0.013%	0.061%
Multibody shape	0.016%	0.011%
PID efficiencies	0.040%	0.049%
Sum in quadrature	0.051%	0.146%

The interaction asymmetry of a charged particle arises from the difference in the interaction cross-sections of positively- and negatively- charged particles with the detector material; as these can lead to a different proportion of positive and negative particles being detected by the experiment, it is important to estimate them.

It must be noted that the LHCb detector did not get any changes[3] between Run 1 and Run 2, therefore the material encountered by particles traversing has remained the same, and so must be the interaction asymmetries. If that is the case, it should be sufficient to measure the interaction asymmetry of a particle in either Run 1 or Run 2 to have a valid measurement for both runs.

To test this assumption, we will compare two different methods used for computing $A_D(\pi^-)$ in Run 1 and Run 2 in the past and confirm that they yield similar results; we will also measure directly $A_D(K^-)$ for both Run 1 and Run 2 to ensure that they are compatible, at least to the current level of precision achievable; these evidences will then allow us to use a measurement of $A_D(p)$ available only for Run 1 on the Run 2 sample as well.

3.9.1 Pion detection asymmetry

The pion detection asymmetry is defined as

$$A_{\mathrm{D}}(\pi^+) = \frac{\varepsilon_{\mathrm{reco}}(\pi^+) - \varepsilon_{\mathrm{reco}}(\pi^-)}{\varepsilon_{\mathrm{reco}}(\pi^+) + \varepsilon_{\mathrm{reco}}(\pi^-)}, \tag{3.25}$$

where $\varepsilon_{\mathrm{reco}}$ stands for the reconstruction efficiency of the given particle. The pion detection asymmetry has been measured on two separate occasions in Run 1 and Run 2, with two different methods that are not part of this thesis; we will now briefly describe them nonetheless.

Run 1 the measurement was carried out during the a_{sl}^s Run 1 analysis [115] by studying partially reconstructed $D^{*+} \to (D^0 \to K^-\pi^+\pi^-\pi^+)\pi_s^+$ decays. The procedure involves selecting decays in which one of the final-state π^+ is undetected by studying the mass difference $\Delta M = m(\pi_s^+ K^-\pi^+\pi^-) - m(K^-\pi^+\pi^-)$, that thanks to the

[3]The only modification being the removal of aerogel from RICH1, but its lightness suggests the effect of its removal on the total material budget is negligible.

kinematic constraints is sufficient to infer the momentum of the missing pion using the Lagrange multiplier method. By selecting the fully reconstructed events as well, a reconstruction efficiency can be defined as

$$\varepsilon_{\text{reco}}(\pi^{\pm}) = \frac{N_{\text{partial}}(\pi^{\pm})}{N_{\text{full}}(\pi^{\pm})}. \tag{3.26}$$

The efficiency can be measured as a function of pion momentum to allow a comparison between the kinematics of pions from different samples, To correct the momentum resolution of the missing pion in the partial reconstruction, which is worse than the resolution of a detected pion, an unfolding was performed to obtain an efficiency correction as a function of the true momentum of the charged particle. From the efficiencies for positive and negative particles, then, the asymmetry can be computed using Eq. (3.25).

Run 2 the measurement was carried out during the bin-flip $D^0 \to K^0_S\pi^+\pi^-$ analysis [116]. The method involves reconstructing prompt $K^0_S \to \pi^+\pi^-$ decays by combining one VELO track and a long track; the missing momentum, and therefore the invariant mass of the K^0_S candidate, can be inferred by constraining the origin of the K^0_S candidate to the primary vertex. After that, it is checked if a matching long track is found to fully reconstruct the decay. The efficiency to reconstruct a pion is then determined by mass fits to the tag (long track) and the probe (VELO track) sample and the combining the yields according to

$$\varepsilon^{\pm}_{\text{reco}} = \frac{N(\pi^{\pm} \text{ VELO tracks matched to long tracks})}{N(\pi^{\pm} \text{ VELO tracks})} \tag{3.27}$$

The efficiencies are obtained as function of momentum, η and charge and the asymmetry is calculated using Eq. (3.25).

Table 3.14 and Fig. 3.28 report and show the values of $A_{\text{D}}(\pi^+)$ in Run 1 as a function of pion momentum. The measurement was carried out up to 100 GeV/c^2 of pion momentum, so in order to ensure coverage of higher-momentum pions in the $\Lambda^0_b \to p\pi^-$ sample one last bin was added with central value equal to the previous one and twice the uncertainty. The Run 2 values of $A_D(\pi^+)$ are shown in Fig. 3.29 and reported in Table 3.15, as a function of pion momentum and pseudorapidity. The statistics was not sufficient to perform the measurement in several bins, in which a value of $(0 \pm 1)\%$ is set; this is justified by noticing that these bin mostly occur at high momentum in Fig. 3.29, where detection asymmetries are expected to become smaller in magnitude. This measurement was only performed for 2016, 2017, and 2018; for 2015 data, we will use the map for 2016 and inflate the errors according to the ratio of the integrated luminosities collected in the two years ($\sqrt{\mathcal{L}(16)/\mathcal{L}(15)} \approx 2.4$), to simulate the difference in data sample sizes.

Table 3.14: Values, in percent, of the pion interaction asymmetry in 2011 and 2012, separated by magnet polarity.

Momentum	2011		2012	
$[\,\mathrm{GeV}/c^2]$	Down	Up	Down	Up
$p \in [2, 6]$	-0.59 ± 0.36	-0.45 ± 0.43	0.32 ± 0.22	-1.21 ± 0.21
$p \in [6, 15]$	0.34 ± 0.24	-0.47 ± 0.29	0.00 ± 0.15	-0.52 ± 0.15
$p \in [15, 20]$	0.14 ± 0.34	-0.22 ± 0.42	-0.12 ± 0.21	0.08 ± 0.21
$p \in [20, 30]$	0.18 ± 0.37	-0.31 ± 0.45	-0.12 ± 0.22	0.04 ± 0.22
$p \in [30, 40]$	0.04 ± 0.56	-0.11 ± 0.68	-0.73 ± 0.33	0.15 ± 0.33
$p \in [40, 50]$	0.04 ± 0.80	0.88 ± 0.96	-0.50 ± 0.48	0.15 ± 0.48
$p \in [50, 100]$	-0.49 ± 0.88	0.6 ± 1.1	-1.07 ± 0.51	0.62 ± 0.51
$p \in [100, 150]$	-0.5 ± 1.8	0.6 ± 2.1	-1.1 ± 1.0	0.6 ± 1.0

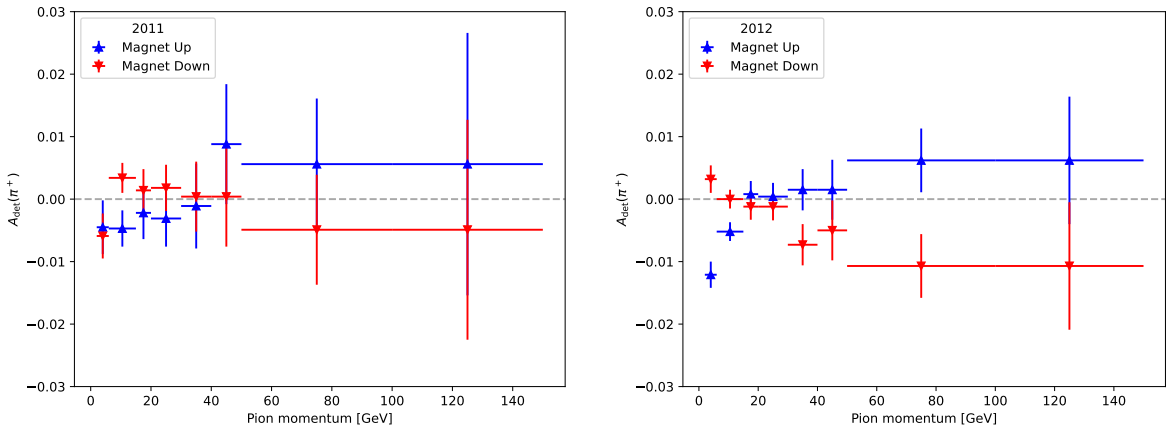

Figure 3.28: Values of the pion interaction asymmetry in (left) 2011 and (right) 2012, separated by magnet polarity.

Table 3.15: Values of the pion interaction asymmetry (in percent) in 2016, 2017, and 2018, separated by magnet polarity.

Pseudorapidity	Momentum	2016		2017		2018	
	$[\mathrm{GeV}/c^2]$	Down	Up	Down	Up	Down	Up
$\eta \in [2.0, 3.0]$	$p \in [2, 5]$	1.01 ± 0.04	-1.27 ± 0.04	1.16 ± 0.03	-1.59 ± 0.03	1.03 ± 0.03	-1.32 ± 0.02
$\eta \in [2.0, 3.0]$	$p \in [5, 8]$	0.11 ± 0.04	-0.44 ± 0.04	0.02 ± 0.03	-0.45 ± 0.04	-0.04 ± 0.04	-0.55 ± 0.04
$\eta \in [2.0, 3.0]$	$p \in [8, 10]$	0.18 ± 0.06	-0.35 ± 0.06	0.03 ± 0.05	-0.26 ± 0.05	-0.09 ± 0.05	-0.28 ± 0.06
$\eta \in [2.0, 3.0]$	$p \in [10, 12]$	0.25 ± 0.09	-0.36 ± 0.09	-0.02 ± 0.07	-0.20 ± 0.07	-0.03 ± 0.08	-0.07 ± 0.09
$\eta \in [2.0, 3.0]$	$p \in [12, 15]$	0.1 ± 0.1	-0.1 ± 0.1	0.1 ± 0.1	-0.3 ± 0.1	-0.1 ± 0.1	-0.2 ± 0.1
$\eta \in [2.0, 3.0]$	$p \in [15, 20]$	-0.2 ± 0.2	-0.2 ± 0.2	0.0 ± 0.1	-0.1 ± 0.1	0.0 ± 0.1	0.3 ± 0.2
$\eta \in [2.0, 3.0]$	$p \in [20, 25]$	0.2 ± 0.4	-0.3 ± 0.4	0.1 ± 0.3	0.3 ± 0.3	-0.3 ± 0.3	0.8 ± 0.4
$\eta \in [2.0, 3.0]$	$p \in [25, 32]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1
$\eta \in [2.0, 3.0]$	$p \in [32, 40]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1
$\eta \in [2.0, 3.0]$	$p \in [40, 50]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1
$\eta \in [2.0, 3.0]$	$p \in [50, 75]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1
$\eta \in [3.0, 4.0]$	$p \in [2, 5]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1
$\eta \in [3.0, 4.0]$	$p \in [5, 8]$	0.37 ± 0.06	-0.67 ± 0.06	0.21 ± 0.04	-0.63 ± 0.05	0.19 ± 0.03	-0.47 ± 0.03
$\eta \in [3.0, 4.0]$	$p \in [8, 10]$	0.11 ± 0.04	-0.37 ± 0.04	-0.02 ± 0.03	-0.36 ± 0.03	0.11 ± 0.03	-0.35 ± 0.04
$\eta \in [3.0, 4.0]$	$p \in [10, 12]$	0.12 ± 0.05	-0.35 ± 0.05	0.02 ± 0.03	-0.30 ± 0.04	0.07 ± 0.04	-0.40 ± 0.04
$\eta \in [3.0, 4.0]$	$p \in [12, 15]$	0.08 ± 0.06	-0.28 ± 0.06	-0.01 ± 0.04	-0.24 ± 0.04	0.16 ± 0.05	-0.20 ± 0.05
$\eta \in [3.0, 4.0]$	$p \in [15, 20]$	0.16 ± 0.05	-0.27 ± 0.05	-0.04 ± 0.04	-0.24 ± 0.04	-0.01 ± 0.04	-0.22 ± 0.05
$\eta \in [3.0, 4.0]$	$p \in [20, 25]$	0.11 ± 0.07	-0.29 ± 0.07	0.01 ± 0.05	-0.12 ± 0.06	-0.05 ± 0.07	-0.19 ± 0.08
$\eta \in [3.0, 4.0]$	$p \in [25, 32]$	0.0 ± 0.1	-0.1 ± 0.1	0.05 ± 0.09	-0.1 ± 0.1	0.1 ± 0.1	-0.1 ± 0.1
$\eta \in [3.0, 4.0]$	$p \in [32, 40]$	-0.1 ± 0.2	0.2 ± 0.2	0.1 ± 0.1	-0.0 ± 0.1	-0.5 ± 0.2	-0.3 ± 0.2
$\eta \in [3.0, 4.0]$	$p \in [40, 50]$	0.2 ± 0.4	0.1 ± 0.4	-0.1 ± 0.3	-0.2 ± 0.3	0.0 ± 0.3	-0.2 ± 0.3
$\eta \in [3.0, 4.0]$	$p \in [50, 75]$	0.6 ± 0.5	-0.4 ± 0.5	-0.0 ± 0.3	-0.2 ± 0.3	-0.5 ± 0.4	0.3 ± 0.5
$\eta \in [4.0, 5.0]$	$p \in [2, 5]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1
$\eta \in [4.0, 5.0]$	$p \in [5, 8]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1
$\eta \in [4.0, 5.0]$	$p \in [8, 10]$	-0.27 ± 0.09	0.12 ± 0.09	-0.40 ± 0.07	-0.13 ± 0.08	-0.11 ± 0.06	-0.28 ± 0.07
$\eta \in [4.0, 5.0]$	$p \in [10, 12]$	-0.6 ± 0.1	0.2 ± 0.1	-0.4 ± 0.1	0.2 ± 0.1	-0.49 ± 0.07	-0.245 ± 0.003
$\eta \in [4.0, 5.0]$	$p \in [12, 15]$	-0.2 ± 0.2	-0.1 ± 0.2	-0.6 ± 0.1	0.1 ± 0.1	-0.35 ± 0.08	0.246 ± 0.006
$\eta \in [4.0, 5.0]$	$p \in [15, 20]$	-0.5 ± 0.1	-0.4 ± 0.1	-0.3 ± 0.1	-0.0 ± 0.1	-0.45 ± 0.07	0.00 ± 0.08
$\eta \in [4.0, 5.0]$	$p \in [20, 25]$	-0.2 ± 0.2	-0.1 ± 0.2	-0.2 ± 0.1	0.0 ± 0.2	-0.9 ± 0.1	0.0 ± 0.1
$\eta \in [4.0, 5.0]$	$p \in [25, 32]$	-0.3 ± 0.2	0.2 ± 0.2	0.1 ± 0.2	-0.1 ± 0.2	-0.5 ± 0.2	-0.4 ± 0.2
$\eta \in [4.0, 5.0]$	$p \in [32, 40]$	0.1 ± 0.2	-0.1 ± 0.2	-0.2 ± 0.1	-0.1 ± 0.2	-0.5 ± 0.2	-0.2 ± 0.2
$\eta \in [4.0, 5.0]$	$p \in [40, 50]$	0.2 ± 0.2	0.0 ± 0.2	-0.1 ± 0.2	0.1 ± 0.2	-0.4 ± 0.2	0.1 ± 0.2
$\eta \in [4.0, 5.0]$	$p \in [50, 75]$	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1	0 ± 1

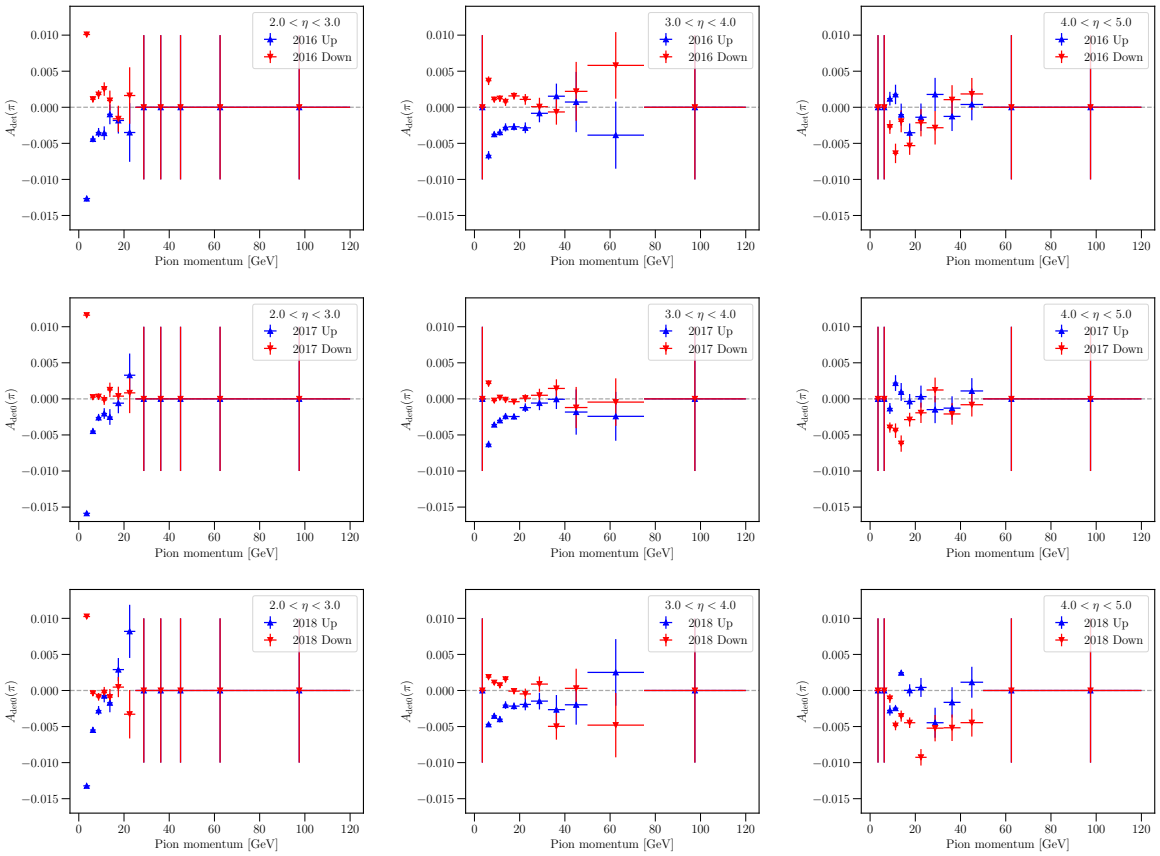

Figure 3.29: Values of the pion interaction asymmetry in (top) 2016, (middle) 2017, and (bottom) 2018, for increasing pseudorapidity from left to right, separated by magnet polarity.

To obtain a value of $A_D(\pi^+)$ to be subtracted from the raw asymmetry in the $\Lambda_b^0 \to p\pi^-$ sample, a reweighting of this asymmetries is done by computing:

$$A_D(\pi^+) = \sum_{i=1}^{N} f_i A_{D,i}, \tag{3.28}$$

where $f_i = w_i / \sum_i^N w_i$, with w_i the sum of the $\Lambda_b^0 \to p\pi^-$ s–Weights in the i-th momentum (and pseudorapidity for Run 2 data) bin, and $A_{D,i}$ is the value of the detection asymmetry in the same bin. The results for both Run 1 and Run 2 are shown in Fig. 3.30. The average values are found to be $(-0.10 \pm 0.11)\%$ for Run 1 and $(-0.02 \pm 0.07)\%$ for Run 2, compatible between each other.

3.9.2 Kaon detection asymmetry

The kaon detection asymmetry is defined as

$$A_{\mathrm{D}}(K^-) = \frac{\varepsilon_{\mathrm{reco}}(K^-) - \varepsilon_{\mathrm{reco}}(K^+)}{\varepsilon_{\mathrm{reco}}(K^-) + \varepsilon_{\mathrm{reco}}(K^+)}, \tag{3.29}$$

where $\varepsilon_{\mathrm{reco}}$ stands for the reconstruction efficiency of the given particle. The kaon detection asymmetry used in this thesis is obtained by measuring the two-body detection asymmetry

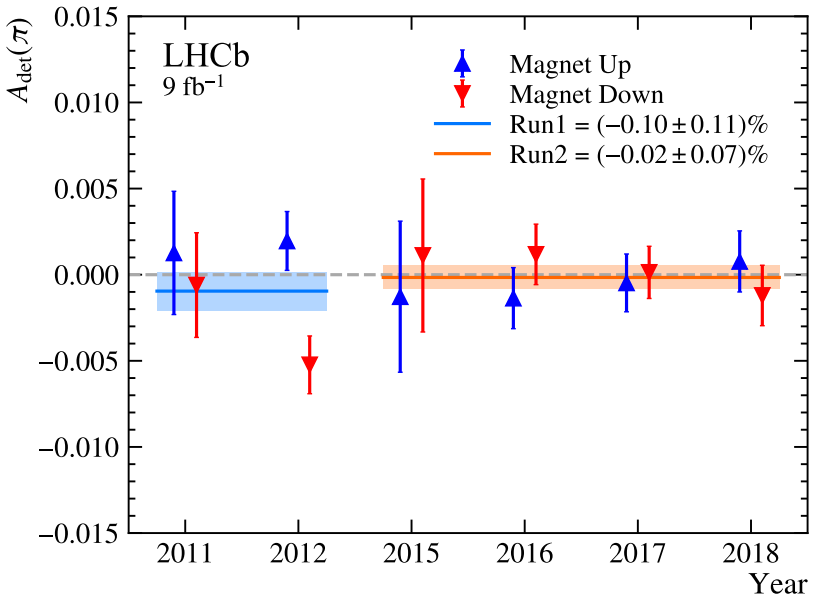

Figure 3.30: Values of $A_D(\pi^+)$ reweighted over the signal $\Lambda_b^0 \to p\pi^-$ sample's kinematics, separated by year and magnet polarity.

$A_D(K^-\pi^+)$ with the use of $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K^0_S\pi^+$ control modes [113, 114]. Here we summarize the main points:

- $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K^0_S\pi^+$ decays are selected from Run 1 and Run 2 data, applying the requirements reported in Table 3.16;
- the samples are divided in bins of kaon momentum with edges (in GeV/c) $[5, 10, 15, 20, 25, 40, 60, 80]$;
- in each bin, a kinematic reweighting is applied to equalize the kinematic of the two decays, in order to cancel any residual nuisance asymmetry between them; to ensure a successful reweighting, a fiducial cut defined by

  $$p_{\rm T}(D^+) \leq \min\left(36 - 5.48\eta(D^+),\ 20\tanh\big(4(\eta(D^+) - 1.8)\big),\ 100 - 19.5\eta(D^+)\right)$$

  is added to both sample to avoid regions of the phase-space not sufficiently covered by either sample ($p_{\rm T}(D^+)$ in GeV/c);
- a maximum-likelihood fit is performed in each bin to the reweighted D^+ and D^- invariant-mass distributions to obtain the raw asymmetries, which can be expressed as:

  $$A_{\rm raw}(D^+ \to K^-\pi^+\pi^+) = A_D(K^-\pi^+) + A_D(\pi^+) + A_{\rm exp}(D^+ \to K^-\pi^+\pi^+), \quad (3.30)$$
  $$A_{\rm raw}(D^+ \to K^0_S\pi^+) = A(K^0) + A_D(\pi^+) + A_{\rm exp}(D^+ \to K^0_S\pi^+), \quad (3.31)$$

  where CP asymmetries are ignored as they are expected to be negligible for Cabibbo-favoured decays, and $A(K^0)$ includes the effects of CP violation in $K^0_S \to \pi^+\pi^-$ decays and interaction of a K^0 with the detector material, and was last measured

Table 3.16: Offline selection of (top) $D^+ \to K^-\pi^+\pi^+$ and (bottom) $D^+ \to K^0_S\pi^+$ decays. The numbers in round brackets correspond to the values used in 2016, while those in square brackets were used only for the Run 1 sample.

Variable		Cut	Units
Track χ^2/ndf of each daughter	$<$	3	
$p_{\rm T}$ of each daughter	$>$	250	MeV/c
At least one daughter with $p_{\rm T}$	$>$	1 (0.2)	GeV/c
At least two daughters with $p_{\rm T}$	$>$	0.4 (0.2)	GeV/c
$\chi^2_{\rm IP}$ of each daughter	$>$	4 (25)	
At least one daughter with $\chi^2_{\rm IP}$	$>$	50 (100)	
At least two daughters with $\chi^2_{\rm IP}$	$>$	10 (49)	
DLL$_{K\pi}(K)$	$>$	1 [0]	
DLL$_{Kp}(K)$	$>$	-5 [-7]	
DLL$_{K\pi}(\pi)$	$<$	5	
Scalar sum of daughters' $p_{\rm T}$	$>$	3 (1)	GeV/c
$p_T(D^+)$	$>$	2.5	GeV/c
D^+ decay time	$>$	0.4 (0.5)	ps
D^+ vertex fit χ^2/ndf	$<$	6 (1)	
D^+ pointing angle (a.k.a. arccos(DIRA))	$<$	10 (14.1)	mrad
D^+ flight-distance χ^2	$>$	150 (125)	
$m(D^+)$	$\in$	$[1.80, 1.94]$	GeV/c^2
$p_{\rm T}$ of each daughter	$>$	250	MeV/c
$\chi^2_{\rm IP}$ of each daughter	$>$	36	
DLL$_{K\pi}$ of the bachelor pion	$<$	5	
$\overline{K}^0$ vertex fit χ^2/ndf	$<$	30	
$m(\overline{K}^0)$	$\in$	$[463, 533]$	MeV/c^2
$\overline{K}^0$ decay time (w.r.t. best PV)	$>$	2	ps
z position of the $\overline{K}^0$ vertex	$\in$	$[-100, 500]$	mm
Scalar sum of daughters' $p_{\rm T}$	$>$	2	GeV/c
$p_{\rm T}(D^+)$	$>$	2.5	GeV/c
D^+ pointing angle (a.k.a. arccos(DIRA))	$<$	17.3	mrad
D^+ flight-distance χ^2	$>$	30	
D^+ decay time	$>$	0.4	ps
$m(D^+)$	$\in$	$[1.80, 1.94]$	GeV/c^2

in [133]. The remaining experimental asymmetries, which we called $A_{\rm exp}(D^+ \to f)$ and that include the D^+ production asymmetry, the D^+ bachelor pion detection asymmetry, and the trigger-induced asymmetries, are expected to be equal and thus cancel out after the kinematic reweighting. Therefore, the $K^-\pi^+$ interaction asymmetry can be computed as:

$$A_D(K^-\pi^+) = A_{\rm raw}(D^+ \to K^-\pi^+\pi^+) - A_{\rm raw}(D^+ \to K^0_S\pi^+) + A(K^0); \qquad (3.32)$$

- finally, the kaon detection asymmetry can be obtained by subtracting the pion

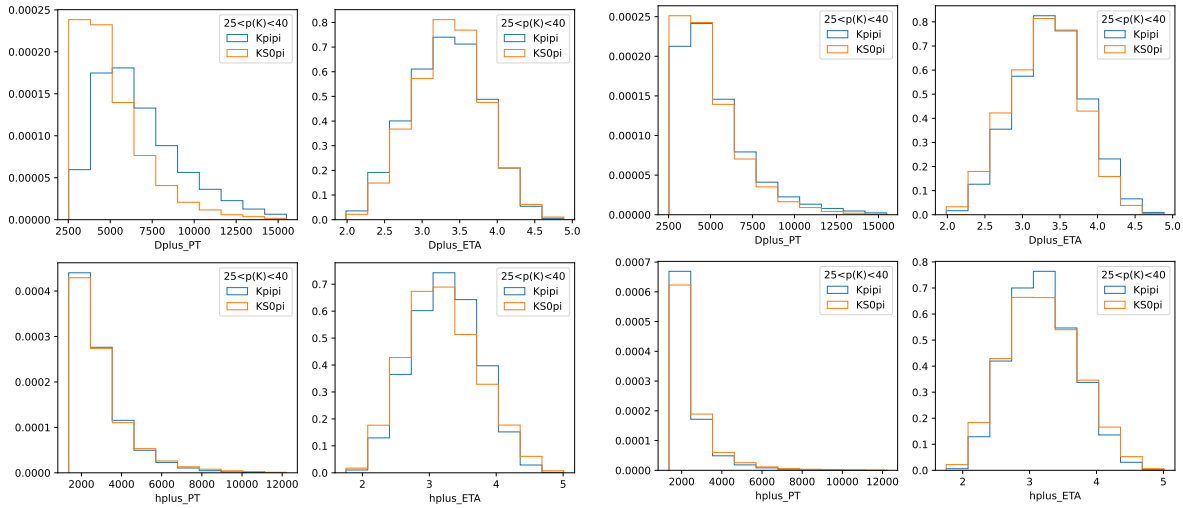

Figure 3.31: Kinematic distributions of the $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K^0_S\pi^+$ samples before (left) and after (right) the kinematic reweighting. The plots correspond to the fifth momentum bin ($25 < p < 40$ GeV/c^2) of the 2018, magnet down sample.

interaction asymmetry as measured in Section 3.9.1:

$$A_D(K^-) = A_D(K^-\pi^+) - A_D(\pi^+). \tag{3.33}$$

In this thesis, we applied the procedure separately to the sub-samples divided by year and magnet polarity.The kinematic reweighting is performed with the `hep-ml` package [134], to equalize the p_T and η distribution of the D^+ and the tag pion. The latter is defined as the π^+ from the $D^+ \to K^0_S\pi^+$ decay, and the one pion firing the `Hlt1TrackMVA` line (or `Hlt1TrackAllL0` in Run 1) in the $D^+ \to K^-\pi^+\pi^+$ decay, or one random pion in case both activate the trigger. In Fig. 3.31 the kinematic distributions in one momentum bin before and after the reweighting are shown.

After the reweighting, a maximum-likelihood fit is performed to extract the raw asymmetries. The signal shape is modelled by the sum of a Crystal Ball [135] function and one or two gaussians for $D^+ \to K^0_S\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ decays, respectively. The combinatorial background is modelled with an exponential function. The plots of the fits in all subsamples are shown in Section 3.15.

After subtracting the $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K^0_S\pi^+$ raw asymmetries and using the value $A(K^0) = (-5.1 \pm 0.6) \times 10^{-4}$ from Ref. [133], the values for the two-body interaction asymmetry $A_D(K^-\pi^+)$ are obtained. In Ref. [133] the value of $A(K^0)$ was measured using two distinct decay channels, namely $D^+ \to K^0_S\pi^+$ and $D^+_s \to K^0_S K^+$. The number used in this analysis comes from the former, while the latter returned the result $A(K^0) = (-8.5 \pm 1.3) \times 10^{-4}$, differing from the other one by about 2.5σ due to the different decay-time distribution of the K^0_S meson. To make a conservative estimate of a possible systematic effect due to the different kinematic distribution of the K^0_S meson from the $D^+ \to K^0_S\pi^+$ channel, we will add in quadrature an uncertainty equal to three times the difference between the two values of $A(K^0)$ quoted in Ref. [133], *i.e.* $3 \times (8.5 - 5.1) \times 10^{-4} = 1.2 \times 10^{-3}$, resulting in a value of $A(K^0) = (-5.1 \pm 1.3) \times 10^{-4}$.

To get the single-track $A_D(K^-)$ asymmetry, in each bin of kaon momentum the single pion asymmetry maps from Tables 3.14 and 3.15 are integrated over the non-tag pion

Table 3.17: Values of the (left) $K^-\pi^+$ and (right) K^- interaction asymmetries obtained from the Run 1 and Run 2 samples.

	$K^-\pi^+$		K^-	
Momentum [GeV/c^2]	Run 1 [%]	Run 2 [%]	Run 1 [%]	Run 2 [%]
$5 < p < 10$	-3.0 ± 2.4	-1.34 ± 0.20	-2.8 ± 2.4	-1.23 ± 0.20
$10 < p < 15$	-0.55 ± 0.30	-0.868 ± 0.093	-0.36 ± 0.31	-0.763 ± 0.095
$15 < p < 20$	-0.82 ± 0.22	-0.639 ± 0.090	-0.65 ± 0.23	-0.526 ± 0.092
$20 < p < 25$	-0.90 ± 0.24	-0.652 ± 0.095	-0.77 ± 0.25	-0.545 ± 0.097
$25 < p < 40$	-1.04 ± 0.16	-0.840 ± 0.082	-0.90 ± 0.17	-0.742 ± 0.085
$40 < p < 60$	-0.60 ± 0.21	-0.81 ± 0.12	-0.47 ± 0.22	-0.71 ± 0.12
$60 < p < 80$	-0.54 ± 0.36	-0.65 ± 0.14	-0.41 ± 0.37	-0.55 ± 0.14

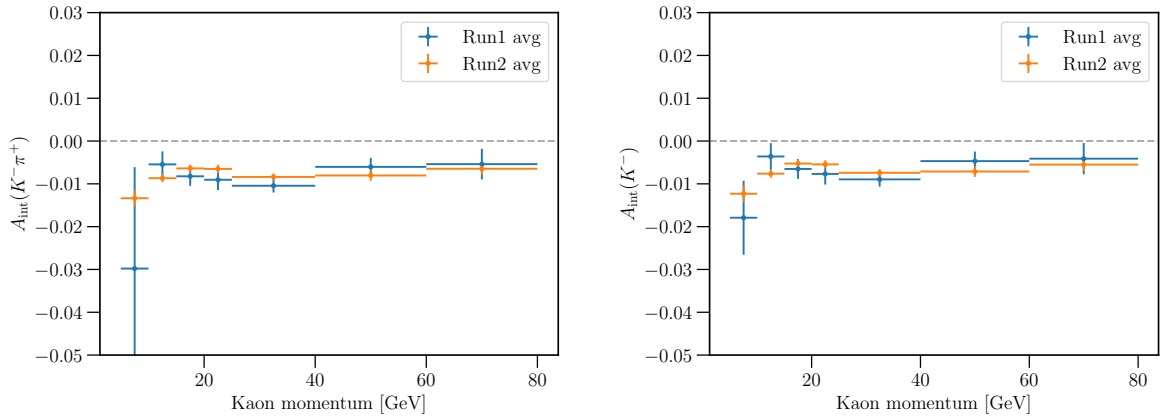

Figure 3.32: Values of the (left) $K^-\pi^+$ and (right) K^- interaction asymmetries obtained from the Run 1 and Run 2 samples.

kinematic distributions to get a total asymmetry value per-bin, which is then subtracted from the two-track asymmetry. The values of both $A_D(K^-\pi^+)$ and $A_D(K^-)$ are reported in Table 3.17 and shown in Fig. 3.32. It can be seen how in each momentum bin the two quantities are compatible with each other.

To obtain a value of $A_D(K^-)$ to be subtracted from the raw asymmetry in the $\Lambda_b^0 \to pK^-$ sample, a reweighting of this asymmetries is done with a formula equivalent to Eq. (3.28). The results for both Run 1 and Run 2 are shown in Fig. 3.33. The average values are found to be $(-0.69 \pm 0.10)\%$ for Run 1 and $(-0.69 \pm 0.05)\%$ for Run 2, compatible between each other.

3.9.3 Proton detection asymmetry

Given the evidence for the compatibility between Run 1 and Run 2 values of the interaction asymmetries of pions and kaons illustrated above, we deem it sufficient to use the available Run 1 proton interaction asymmetry for both Runs. The measurement was performed during the Run 1 Λ_b^0 production asymmetry analysis [112] that is not part of this thesis

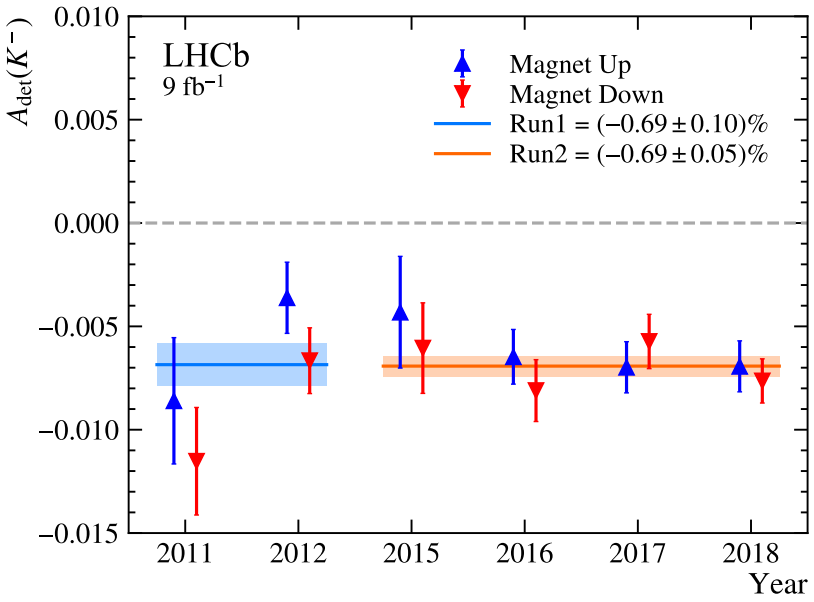

Figure 3.33: Values of $A_D(K^-)$ reweighted over the signal $\Lambda_b^0 \to pK^-$ sample's kinematics, separated by year and magnet polarity.

and for which we summarise the procedure:

- For every proton (or antiproton) track in the $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\mu\nu X$ sample, the total detector material encountered in the flight path, $d^\pm$, is computed using a simulation of the detector and expressed in units of the nuclear collision length $\lambda^\pm$. The latter is in turn defined as the typical length that a hadron travels before it undergoes an (in)elastic scatter with a nucleus in the material;

- Assuming that the dependence of $\lambda^\pm$ on the momentum of the particle for the detector material scales equally as the collision length for a deuterium target, for which there are some available measurements [136–138], then the collision length for any atomic number A and momentum p can be factorised as

$$\lambda^\pm(A,p) = \lambda_{\text{avg}}(A, 20\ \text{GeV}/c)\frac{\sigma_{\text{avg}}(^2\text{H}, 20\ \text{GeV})}{\sigma^\pm(^2\text{H}, p)}, \tag{3.34}$$

  where σ indicates the hadronic interaction cross-section and the "avg" subscript indicates a charge-averaged quantity;

- computing the traversed thickness for every (anti)proton, accounting for every material collision length at a given momentum, a detection efficiency can be evaluated as

$$\varepsilon^\pm \propto \exp\left(-\frac{d^\pm}{\lambda^\pm}\right), \tag{3.35}$$

  from which a detection asymmetry can be readily computed:

$$A_{\text{D}}(p) = \frac{\exp(-d^+/\lambda^+) - \exp(-d^-/\lambda^-)}{\exp(-d^+/\lambda^+) + \exp(-d^-/\lambda^-)} \tag{3.36}$$

Table 3.18: Values of the proton interaction asymmetry in 2011 and 2012, separated by magnet polarity.

Momentum bin	2011		2012	
$[\,\mathrm{GeV}/c^2]$	Down	Up	Down	Up
$10 < p < 15$	$(2.23 \pm 0.28)\%$	$(2.08 \pm 0.26)\%$	$(2.20 \pm 0.28)\%$	$(2.07 \pm 0.26)\%$
$15 < p < 25$	$(1.79 \pm 0.24)\%$	$(1.61 \pm 0.22)\%$	$(1.77 \pm 0.24)\%$	$(1.60 \pm 0.22)\%$
$25 < p < 40$	$(1.38 \pm 0.22)\%$	$(1.27 \pm 0.21)\%$	$(1.37 \pm 0.22)\%$	$(1.25 \pm 0.20)\%$
$40 < p < 60$	$(1.08 \pm 0.20)\%$	$(1.02 \pm 0.19)\%$	$(1.07 \pm 0.20)\%$	$(1.02 \pm 0.19)\%$
$60 < p < 80$	$(0.89 \pm 0.18)\%$	$(0.86 \pm 0.18)\%$	$(0.88 \pm 0.18)\%$	$(0.85 \pm 0.18)\%$
$80 < p < 125$	$(0.74 \pm 0.17)\%$	$(0.74 \pm 0.17)\%$	$(0.75 \pm 0.17)\%$	$(0.71 \pm 0.16)\%$

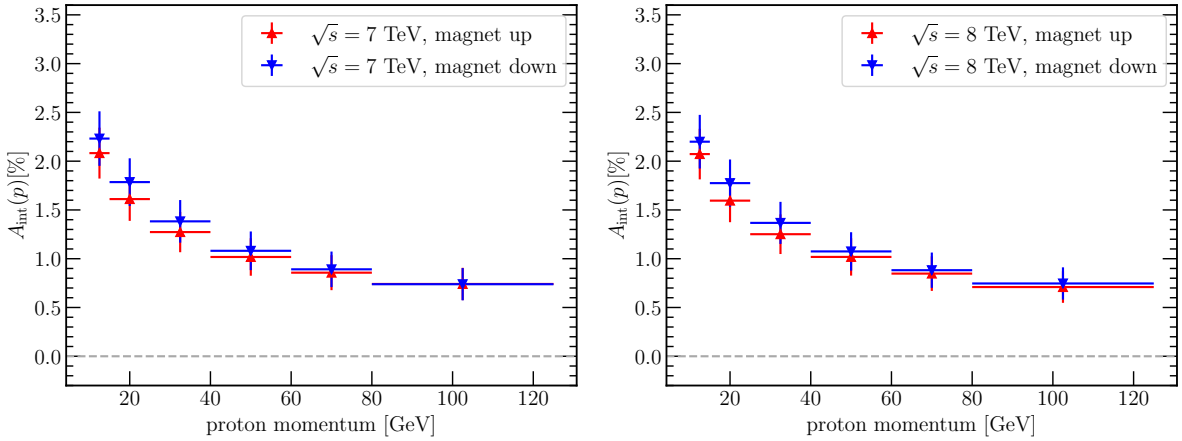

Figure 3.34: Values of the proton interaction asymmetry for (left) 2011 and (right) 2012, separated by magnet polarity.

- Finally, the method is validated using partially reconstructed $\Lambda \to p\pi^-$ decays, for which the protons can be reconstructed as an upstream or VELO track. To do so, as a cross-check the detection asymmetries are also measured for pions from $K^0_S \to \pi^+\pi^-$, which can be measured with a much higher precision, and compared to those for protons. Any residual discrepancies between the results from data and simulation are taken as systematic uncertainties for the measurement.

The measurement was performed both as a function of Λ_b^0 kinematics and proton momentum; we will use the latter so that we can reweight it to the kinematics of the protons in the signal $\Lambda_b^0 \to ph^-$ sample. The momentum distribution of the proton interaction asymmetry in reported in Table 3.18 and shown in Fig. 3.34 for 2011 and 2012. The asymmetry maps are summed over the proton momentum of signal $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays with the same procedure described in Eq. (3.28). For Run 1 the maps will be used as they are, while for the Run 2 samples the 2012 map will be used for all years. The values for each year and magnet polarities are shown in Fig. 3.35. The average values for $\Lambda_b^0 \to pK^-$ decays are found to be $(1.45 \pm 0.08)\%$ for Run 1 and $(1.45 \pm 0.06)\%$ for Run 2, while for $\Lambda_b^0 \to p\pi^-$ they are $(1.44 \pm 0.08)\%$ for Run 1 and $(1.43 \pm 0.06)\%$ for Run 2.

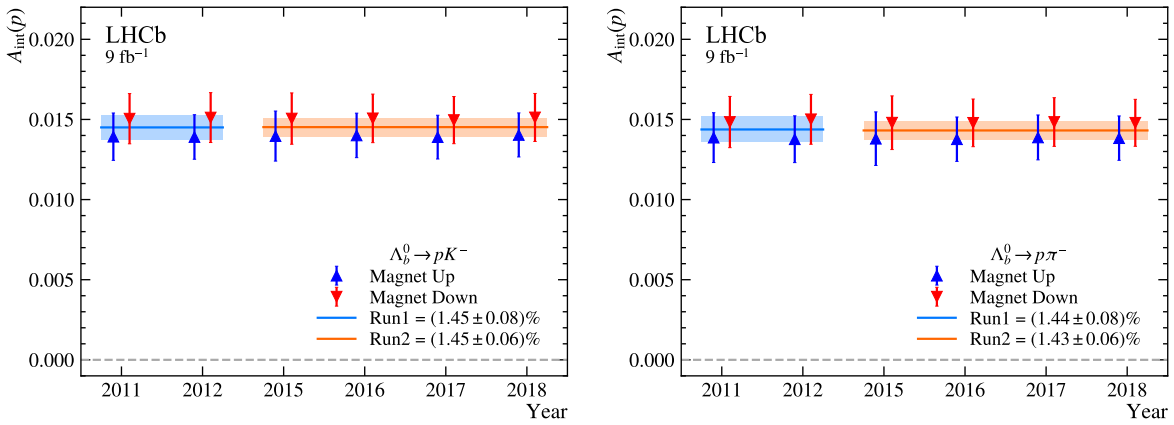

Figure 3.35: Values of $A_D(p)$ reweighted to the proton kinematics of signal (left) $\Lambda_b^0 \to pK^-$ and (right) $\Lambda_b^0 \to p\pi^-$

3.10 PID asymmetries

Particle identification requirements on the the pK^- and $p\pi^-$ final states can induce asymmetries that need to be taken into account. The PID asymmetries are defined as

$$A_{\rm PID}(pK^-) = \frac{\varepsilon_{\rm PID}(pK^-) - \varepsilon_{\rm PID}(\overline{p}K^+)}{\varepsilon_{\rm PID}(pK^-) + \varepsilon_{\rm PID}(\overline{p}K^+)}, \tag{3.37}$$

$$A_{\rm PID}(p\pi^-) = \frac{\varepsilon_{\rm PID}(p\pi^-) - \varepsilon_{\rm PID}(\overline{p}\pi^+)}{\varepsilon_{\rm PID}(p\pi^-) + \varepsilon_{\rm PID}(\overline{p}\pi^+)}, \tag{3.38}$$

where $\varepsilon_{\rm PID}(f)$ is the PID efficiency for the final state f. Since the application of fiducial cuts described in Section 3.3 rejects phase-space regions where the PID calibration samples do not provide enough coverage, the PID efficiencies computed on the signal samples can be used directly in Eqs. (3.37) and (3.38).

The effect due to the choice of kinematic bins used to compute the efficiency maps is evaluated as a systematic uncertainty on the PID asymmetry. In order to assess the size of the effect, we alternatively varied the binning of each variable and we determined again the PID efficiencies from Eq. (3.14) for all the $H_b \to h^+h'^-$ decays. Furthermore, an additional variable, the azimuthal angle of the track, ϕ, is introduced for the binning, for a total of 125 different binning schemes. The largest variation between the baseline asymmetry and the asymmetry computed from the new efficiencies is eventually taken as systematic uncertainty. When including the systematic uncertainty, the values shown in Fig. 3.36 and reported in Table 3.19 are obtained. The Run-average values of $A_{\rm PID}$ for $\Lambda_b^0 \to pK^-$ are $(-0.40 \pm 0.56)\%$ for Run 1 and $(-0.10 \pm 0.28)\%$ for Run 2, while for $\Lambda_b^0 \to p\pi^-$ they are $(-0.36 \pm 0.59)\%$ for Run 1 and $(-0.20 \pm 0.30)\%$ for Run 2.

3.11 Determination of trigger-induced asymmetries

In this section the determination of the asymmetries introduced by the trigger requirements is described. Since the candidates are required to pass some trigger selection, additional sources of asymmetry can be introduced in the sample if the performance of the trigger

Table 3.19: Values of the PID efficiencies and asymmetries for $\Lambda_b^0 \to ph^-$ and $\overline{\Lambda}_b^0 \to h^+\overline{p}$ decays in Run 1 and Run2. The uncertainties for the asymmetries include the statistical and systematic suorces.

Quantity	$\Lambda_b^0 \to pK^-$		$\Lambda_b^0 \to p\pi^-$	
	Run1	Run 2	Run1	Run 2
$\varepsilon(\Lambda_b^0)$	$(76.37 \pm 0.34)\,\%$	$(83.00 \pm 0.13)\,\%$	$(72.90 \pm 0.34)\,\%$	$(75.34 \pm 0.12)\,\%$
$\varepsilon(\overline{\Lambda}_b^0)$	$(76.97 \pm 0.36)\,\%$	$(83.18 \pm 0.13)\,\%$	$(73.43 \pm 0.36)\,\%$	$(75.64 \pm 0.12)\,\%$
A_{PID}	$(-0.40 \pm 0.56)\,\%$	$(-0.10 \pm 0.28)\,\%$	$(-0.36 \pm 0.59)\,\%$	$(-0.20 \pm 0.30)\,\%$

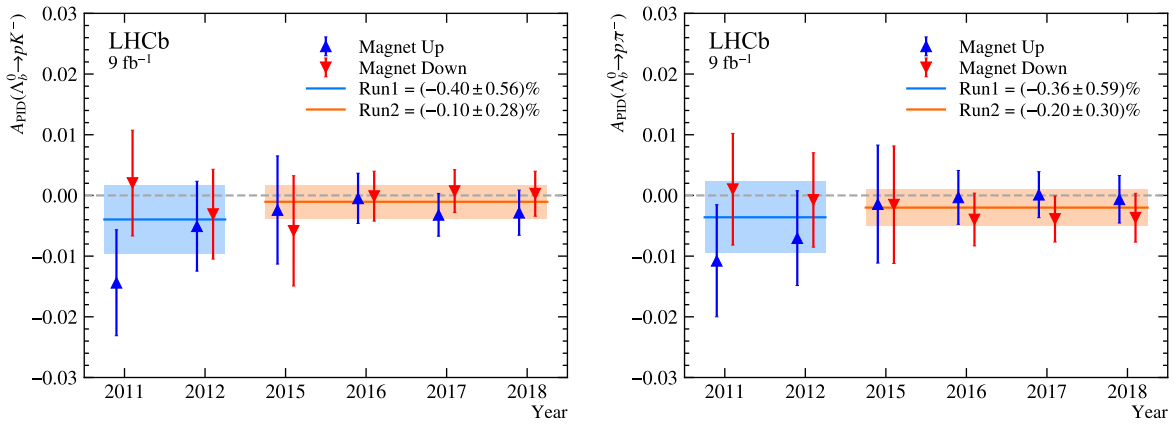

Figure 3.36: Values of the PID asymmetries obtained for (left) $\Lambda_b^0 \to pK^-$ and (right) $\Lambda_b^0 \to p\pi^-$ decays, separated by magnet polarity. The uncertainties inlcude statistical and systematics.

differs for oppositely charged pK and $p\pi$ pairs. In particular, the trigger selection of this analysis are `Hadron_TOS OR Global_TIS` on L0 and `TrackMVA_TOS` on `HLT1`. Therefore, we need to evaluate the effect of three components: the `TOS` part of L0, the `TIS` part of L0, and the `TOS` part of HLT1. As the HLT2 selection is applied after the full decay chain has been reconstructed, and the cuts from the `Hlt2_B2HHBDT` line do not discriminate between positively and negatively charged particles (see Table 3.2), no asymmetry is expected at this stage, therefore we won't need to compute any correction. The `TIS` part will be estimated by means of $B^+ \to J/\psi K^+$ decays, while for the `TOS` contribution a tag-and-probe method was developed using semileptonic $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\mu\nu X$ and $B^0 \to (D^0 \to K^+\pi^-)\mu\nu X$ decays. The followed procedure will be explained in the following sections.

3.11.1 TIS asymmetry

A potential asymmetry from `L0Global_TIS` can arise from the fact that `TIS` events are typically triggered by the other beauty hadron in the event, which preferably decays semileptonically in a high-p_{T} muon that can trigger L0. Since the charge of the muon is related to the flavour of the b-hadron, an asymmetry between CP conjugate final states can be introduced. The `L0Global_TIS` term of the selection is not expected to give a noticeable effect on the raw asymmetry observed in the sample, as this trigger is activated

Table 3.20: Selection of $B^+ \to J/\psi K^+$ candidates for the TIS asymmetry computation (from [139]).

Particle	Variable		Cut
$B^+ \to J/\psi K^+$	$m(J/\psi K^+)$	$\in$	$[5150, 5450]$ MeV/c^2
	t	$>$	0.2 ps
	$\chi^2_{\text{vtx}}/\text{ndf}$	$<$	10
$J/\psi \to \mu^+\mu^-$	$m(\mu^+\mu^-)$	$\in$	$[3016.9, 3176.9]$ MeV/c^2
	$p_{\text{T}}(\mu^\pm)$	$>$	500 MeV/c
	$\chi^2_{\text{DOCA}}(\mu^\pm)$	$<$	20
	$\chi^2_{\text{vtx}}/\text{ndf}$	$<$	16
	$\text{DLL}_{\mu\pi}(\mu^\pm)$	$>$	0
K^+	p_{T}	$>$	500 MeV/c
	$\chi^2_{\text{track}}/\text{ndf}$	$<$	5
	$\text{DLL}_{K\pi}$	$>$	0

independently of the signal decay by other particles coming form the pp collision; we will evaluate this correction nonetheless.

In principle, any b-hadron decay with a large enough statistics should serve the purpose of computing this asymmetry, since we have no way of knowing which decay was responsible for the trigger decision. We decided to use $B^+ \to J/\psi K^+$ decays, and we will perform the measurement as a function of the B^+ transverse momentum, to account for possible dependences of the asymmetry of the parent particle's kinematics. The total value of the asymmetry on the signal $\Lambda_b^0 \to ph^-$ sample will then be obtained by reweigthing the distribution over the Λ_b^0 transverse momentum spectrum. The procedure for computing the TIS asymmetry is the following:

- Apply a selection to the $B^+ \to J/\psi K^+$ sample, selected from the stripping line `StrippingBetaSBu2JpsiKDetachedLine`; the selection is taken from [139] and is reported in Table 3.20;
- Divide the sample in 5 bins of transverse momentum, as well as by charge of the B^+ and trigger category:
  - TIS: `L0Global_TIS && L0Hadron_TOS`;
  - !TIS: `!L0Global_TIS && L0Hadron_TOS`;

  (note that `L0Hadron_TOS` is required in order to have a common denominator for the two samples firing or not the `L0Global_TIS` trigger.)
- For each subsample, perform a maximum likelihood invariant-mass fit to extract the signal yields. The signal is modelled with the sum of two Gaussians and the combinatorial background with an exponential function; the plots of the fits in all subsamples are shown in Chapter A;

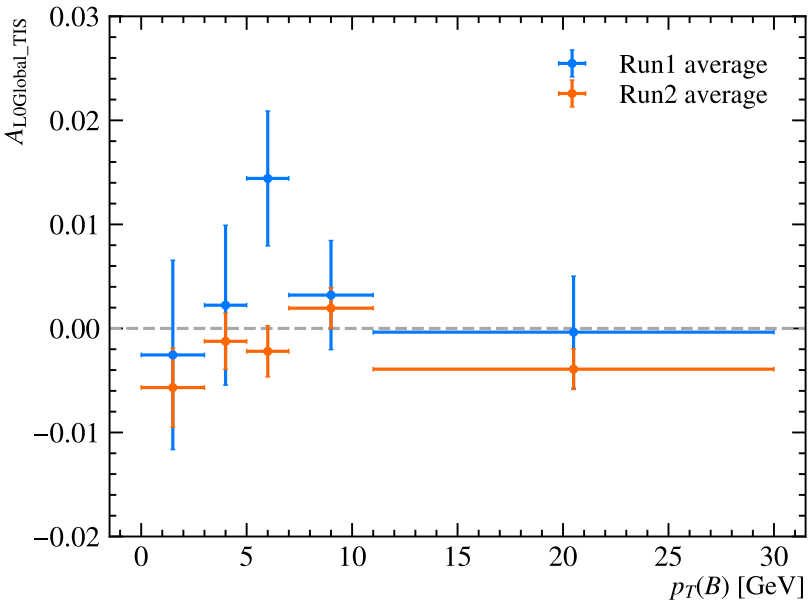

Figure 3.37: Values of A_{TIS} as a function of the B^+ transverse momentum for Run1 and Run2.

- Compute the `TIS` efficiencies and asymmetry as:

$$\varepsilon^+_{\texttt{TIS}} = \frac{N(\texttt{TIS}, B^+)}{N(\texttt{TIS}, B^+) + N(\texttt{!TIS}, B^+)} \tag{3.39}$$

$$\varepsilon^-_{\texttt{TIS}} = \frac{N(\texttt{TIS}, B^-)}{N(\texttt{TIS}, B^-) + N(\texttt{!TIS}, B^-)} \tag{3.40}$$

$$A_{\text{TIS}} = \frac{\varepsilon^-_{\texttt{TIS}} - \varepsilon^+_{\texttt{TIS}}}{\varepsilon^-_{\texttt{TIS}} + \varepsilon^+_{\texttt{TIS}}}, \tag{3.41}$$

  where the sign in Eq. (3.41) reflects the fact that a B^+ meson contains a $\bar{b}$ quark, while a Λ_b^0 baryon contains a b quark.

The resulting A_{TIS} values as a function of the B^+ transverse momentum are shown in Fig. 3.37 for Run 1 and Run 2. A total value to be subtracted from the Λ_b^0 raw asymmetries observed in data is computed by integrating the values over the transverse momentum distribution of signal $\Lambda_b^0 \to ph^-$ decays, following Eq. (3.28). The values of A_{TIS} in each subsample are shown in Fig. 3.38. The total corrections amount to $(0.46 \pm 0.35)\%$ for Run 1 and $(-0.13 \pm 0.13)\%$ for Run 2, respectively, on the $\Lambda_b^0 \to pK^-$ sample, while for the $\Lambda_b^0 \to p\pi^-$ they are $(0.66 \pm 0.36)\%$ for Run 1 and $(-0.17 \pm 0.13)\%$ for Run 2, respectively.

3.11.2 TOS asymmetry

The TOS part of the trigger comprises two different algorithms: L0 and HLT1; we developed a new method for evaluating them. Large statistics samples of protons, kaons, and pions are obtained selecting $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\mu\nu X$ decays for protons and $B^0 \to (D^0 \to K^+\pi^-)\mu\nu X$ decays for kaons and pions; the latter sample is split randomly

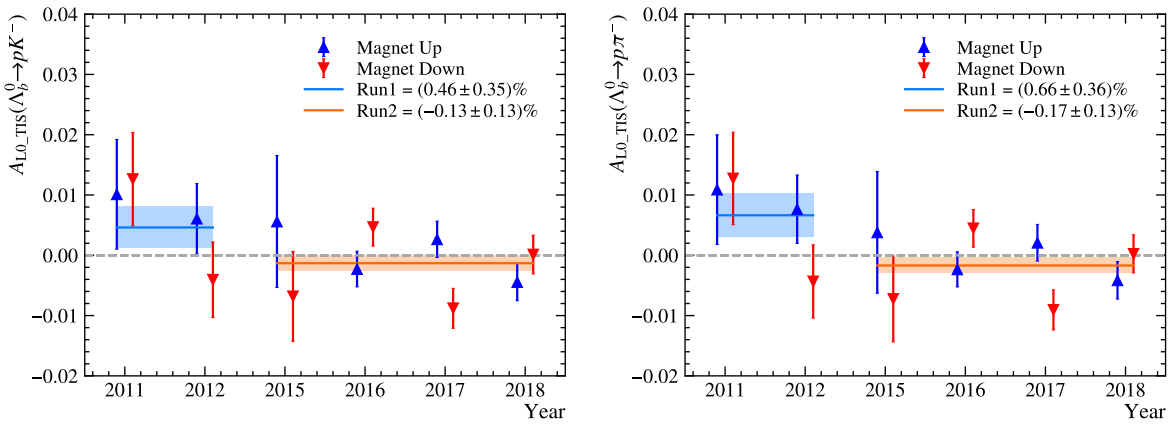

Figure 3.38: Values of A_{TIS} reweighted to the kinematic of signal (left) $\Lambda_b^0 \to pK^-$ and (right) $\Lambda_b^0 \to p\pi^-$ decays.

in half to decorrelate the correction for pions and kaons. Despite being very similar, we will describe the procedures for L0 and HLT1 separately.

L0 trigger

The Level 0 hadron trigger selects events in which a significant amount of transverse energy is deposited in the hadronic calorimeter. In order to evaluate the performance of the trigger on positively and negatively charged particles, we need to get an efficiency map as a function of the transverse energy deposited in the calorimeter. The steps are the following:

- To avoid any overlap between the energy clusters from the different particles in the event, the semileptonic samples are selected by requiring that all energy clusters in the hadron calorimeter are well separated; explicitly, this is done by asking that the variable `Trigger_ET`, which stores the energy of the cluster that activated the hadron trigger, should not be the same for any two of the particles in the event[4];
- The E_{T} of a cluster depends not only on the particle generating it, but also on the underlying event whose particles generate energy deposits that may feed into the considered cluster. To verify that this effect is the same in the $\Lambda_b^0 \to ph^-$ and calibration sample, the distributions of $E_{\text{T}}/p_{\text{T}}$ of the two samples are compared in Fig. 3.39 for protons, kaons, and pions in different bins of p_{T}. In the relativistic regime $E_{\text{T}} \approx p_{\text{T}}$, therefore the distribution of their ratio should peak around 1, unless particles from the underlying event hit the same calorimeter cell and increase the energy deposit, making the ratio deviate from 1; it can be seen that the distributions for reweighted semileptonic samples and background-subtracted signal events are sufficiently similar and peak around 1, proving that the pollution from the underlying event is the same in both samples. This ensures that the efficiencies determined from the calibration sample can be used for the $\Lambda_b^0 \to ph^-$ sample;

[4]This requirement will also need to be applied to the $\Lambda_b^0 \to ph^-$ sample: it results in removing about 6% of candidates.

- Combinatorial background events can then be easily removed with a sideband subtraction, given the high purity of the samples and the flatness of the invariant-mass distributions away from the signal peak; the signal regions are defined as windows of width 16 and 26 MeV/c^2 centered around 2286 and 1865 MeV/c^2 for Λ_c^+ and D^0, respectively, while the two sidebands regions have half the size each and are centered around 2239 and 2335 MeV/c^2 for Λ_c^+ and 1816 and 1923 MeV/c^2 for D^0; the invariant-mass distribution of the two semileptonic samples are shown in Fig. 3.40;

- An efficiency map is computed as a function of transverse energy by triggering on the muon and probing the hadron with the following formula:

$$\varepsilon_h^{\pm}(E_{\mathrm{T}}) = \frac{N(\texttt{L0Hadron\_TOS}(h^{\pm}) \;\&\; \texttt{L0Muon\_TOS}(\mu), E_{\mathrm{T}})}{N(\texttt{L0Muon\_TOS}(\mu), E_{\mathrm{T}})}, \tag{3.42}$$

  where h represents either a proton, a kaon, or a pion. This efficiency is also calculated separately for the inner and outer region of the HCAL, as they have different segmentations and thus likely different performances. The plots of the efficiencies for all subsamples are shown in Chapter B. The E_{T} bin edges are 0, 2, 2.8, 3.2, 3.6, 4, 4.4, 4.8, 5.2, 5.6, 6, 6.4, 6.8, 7.2, 7.6, 8, 8.4, 8.8, 9.4 and 10 GeV/c.

- On the $\Lambda_b^0 \to ph^-$ sample, in each E_{T} bin and HCAL region a two-body efficiency is computed; we will omit the E_{T} and region dependence in the following to simplify the notation and call $\varepsilon_h^{\pm}$ the probability that a charged particle $h \in \{\pi, K, p\}$ with a given E_{T} activates the TOS trigger by hitting a cell in the inner/outer HCAL region, and likewise $1 - \varepsilon_h^{\pm}$ the probability that the same particle does not activate the trigger; the two-body TOS efficiency can then be written as

$$\begin{aligned} \varepsilon_{\Lambda_b^0} &= \varepsilon_p^+(1-\varepsilon_h^-) + (1-\varepsilon_p^+)\varepsilon_h^- + \varepsilon_p^+\varepsilon_h^-, \\ &= 1 - (1-\varepsilon_p^+)(1-\varepsilon_h^-), \end{aligned} \tag{3.43}$$

$$\begin{aligned} \varepsilon_{\bar{\Lambda}_b^0} &= \varepsilon_p^-(1-\varepsilon_h^+) + (1-\varepsilon_p^-)\varepsilon_h^+ + \varepsilon_p^-\varepsilon_h^+, \\ &= 1 - (1-\varepsilon_p^-)(1-\varepsilon_h^+), \end{aligned} \tag{3.44}$$

  from which a bin-by-bin asymmetry is obtained:

$$A_{\Lambda_b^0,i} = \frac{\varepsilon_{\Lambda_b^0} - \varepsilon_{\bar{\Lambda}_b^0}}{\varepsilon_{\Lambda_b^0} + \varepsilon_{\bar{\Lambda}_b^0}}, \tag{3.45}$$

- Finally, the value of the L0 asymmetry is computed for the whole $\Lambda_b^0 \to ph^-$ sample computing the sum:

$$A_{\mathrm{TOS}} = \sum_{i=1}^{N} f_i A_{\Lambda_b^0,i}, \tag{3.46}$$

  where i is the index of each E_{T} and HCAL region bin, $f_i = w_i / \sum_i^N w_i$, and w_i is the sum of s-Weights in the same bin. To assign a systematic uncertainty on the TOS asymmetry, the efficiencies from Eq. (3.42) were also produced applying the method to simulated samples; the results are compared to those obtained from the MC truth the simulation. The difference between the two numbers is taken as a

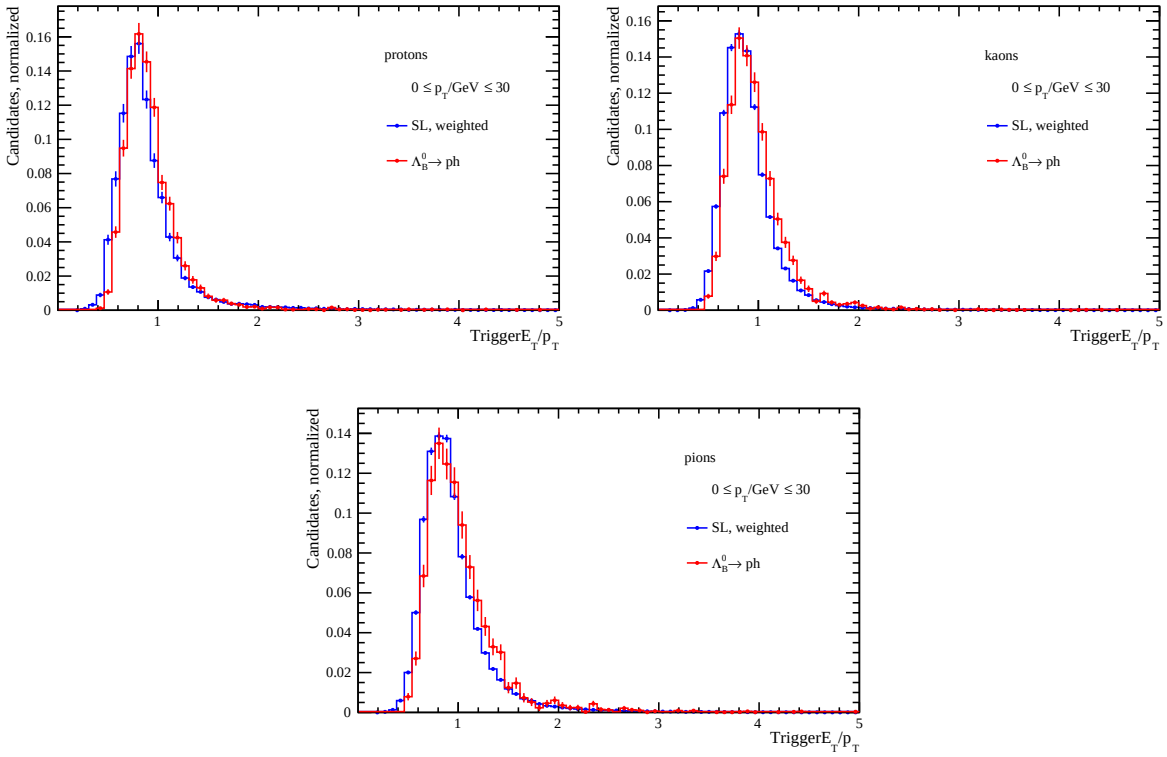

Figure 3.39: Plot of $E_{\mathrm{T}}/p_{\mathrm{T}}$ for (top left) protons, (top right) kaons, and (bottom) pions in the inner HCAL region and for $0 < p_T < 30$ GeV/c. The blue line corresponds to the reweighted semileptonic samples, while the red line to background-subtracted $\Lambda_b^0 \to ph^-$ decays.

systematic uncertainty. The values of A_{TOS} for each year and magnet polarity are reported in Fig. 3.41; for $\Lambda_b^0 \to pK^-$ decays the total values are $(-1.14 \pm 0.19)\%$ for Run 1 and $(-1.29 \pm 0.08)\%$ for Run 2, while for $\Lambda_b^0 \to p\pi^-$ they are $(-1.12 \pm 0.20)\%$ for Run 1 and $(-1.16 \pm 0.09)\%$ for Run 2.

Once A_{TIS} and A_{TOS} are computed, they are combined to get A_{L0} according to the respective fraction of trigger category observed in data:

$$A_{\mathrm{L0}} = A_{\mathrm{TIS}} \frac{N(\texttt{TIS\&\&!TOS})}{N} + A_{\mathrm{TOS}} \frac{N(\texttt{TOS})}{N}; \tag{3.47}$$

as an example, the $\Lambda_b^0 \to pK^-$ sample contains about 39% `TIS` events and 61% `TOS` events. The values shown in Figs. 3.38 and 3.41 and reported in the text are already weighted according to these fractions.

HLT1

The first level of the High Level Trigger performs a partial event reconstruction and selects promising candidates based on some preliminary information. In particular, the `Hlt1TrackMVA` line employed in this analysis makes a selection based on a two-dimensional cut in the $(p_T, \log(\chi^2_{\mathrm{IP}}))$ plane of the tracks. In order to evaluate the performance of the HLT1 trigger on positively and negatively charged particles, we need to get an efficiency

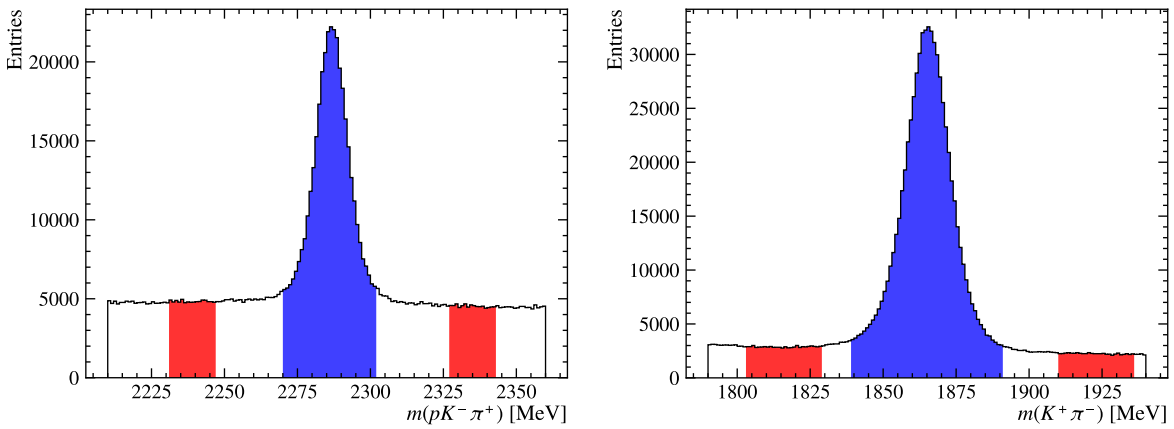

Figure 3.40: Invariant-mass distribution of the semileptonic samples $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\mu\nu X$ (left) and $B^0 \to (D^0 \to K^+\pi^-)\mu\nu X$ (right). The blue and red areas indicate respectively the selected signal and background regions for the sideband subtraction.

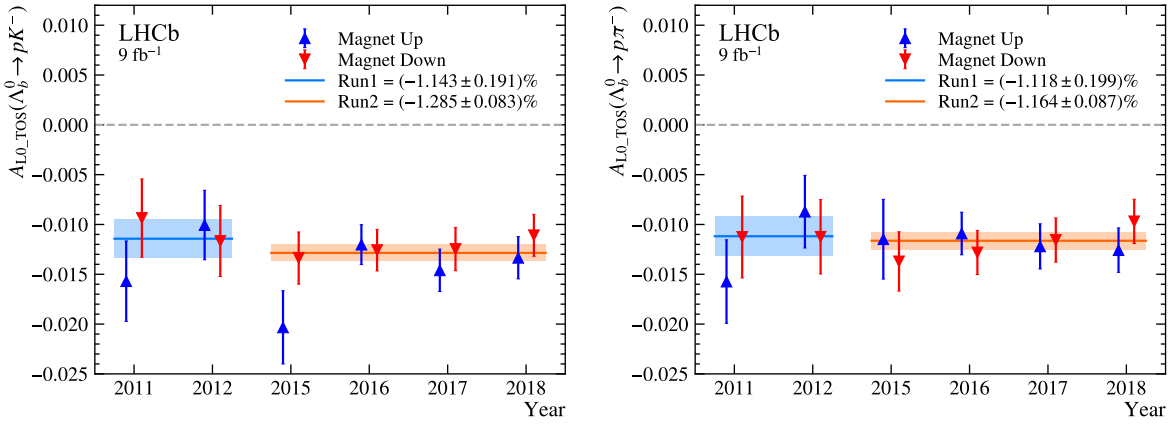

Figure 3.41: Values of A_{TOS} for (left) $\Lambda_b^0 \to pK^-$ and (right) $\Lambda_b^0 \to p\pi^-$ decays, separated by year and magnet polarity.

map as a function of the transverse momentum and the $\log(\chi^2_{\text{IP}})$ of positive and negative particles. The steps are the following:

- The semileptonic samples are selected with the same sideband subtraction as described for the `L0Hadron_TOS` trigger, minus the energy cell overlap cut;
- An efficiency map is computed as a function of transverse momentum and χ^2_{IP} by triggering on the muon and probing the hadron with the following formula:

$$\varepsilon_h^{\pm}(p_T, \log(\chi^2_{\text{IP}})) = \frac{N(\texttt{Hlt1\_TOS}(h^{\pm}) \;\&\; \texttt{Hlt1TrackMuon\_TOS}(\mu) \;\&\; \texttt{L0\_TOS}(\mu))}{N(\texttt{Hlt1TrackMuon\_TOS}(\mu) \;\&\; \texttt{L0\_TOS}(\mu))}, \tag{3.48}$$

where h represents either a proton, a kaon, or a pion. `Hlt1_TOS` and `L0_TOS` are shorthands for `Hlt1TrackMVA_TOS` and `L0Muon_TOS`, respectively. The plots of the efficiencies in all subsamples are shown in Chapter C. The p_{T} bin edges are 0, 1.65, 1.8, 1.92, 2.04, 2.16, 2.28, 2.52, 2.76, 2.88, 3.24, 3.6, 4.08, 4.68, 6 and 12 GeV/c^2.

The $\log(\chi^2_{\rm IP})$ bin edges are 2, 3.5, 4, 4.5, 4.9, 5.3, 5.6, 5.9, 6.1, 6.4, 6.7, 7, 7.3, 7.6, 8, 8.6 and 12.

- On the $\Lambda_b^0 \to ph^-$ sample, in each $p_{\rm T}$ bin and $\log(\chi^2_{\rm IP})$ bin we have a chance $\varepsilon_h^\pm$ that a charged $h^\pm$ particle with a given $p_{\rm T}$ and $\log(\chi^2_{\rm IP})$ activates the HLT1 trigger, and a chance $1-\varepsilon_h^\pm$ of not doing so; then the two-body efficiency can be computed as (we omit the p_T and $\log(\chi^2_{\rm IP})$ dependence in the following to simplify the notation):

$$\begin{aligned}\varepsilon_{\Lambda_b^0} &= \varepsilon_p^+(1-\varepsilon_h^-) + (1-\varepsilon_p^+)\varepsilon_h^- + \varepsilon_p^+\varepsilon_h^- \\ &= 1-(1-\varepsilon_p^+)(1-\varepsilon_h^-) \qquad (3.49)\\ \varepsilon_{\bar{\Lambda}_b^0} &= \varepsilon_p^-(1-\varepsilon_h^+) + (1-\varepsilon_p^-)\varepsilon_h^+ + \varepsilon_p^-\varepsilon_h^+ \\ &= 1-(1-\varepsilon_p^-)(1-\varepsilon_h^+), \qquad (3.50)\end{aligned}$$

  from which a bin-by-bin asymmetry is obtained:

$$A_{\Lambda_b^0,i} = \frac{\varepsilon_{\Lambda_b^0} - \varepsilon_{\bar{\Lambda}_b^0}}{\varepsilon_{\Lambda_b^0} + \varepsilon_{\bar{\Lambda}_b^0}} \qquad (3.51)$$

- Finally, the value of the HLT1 asymmetry is computed for the whole $\Lambda_b^0 \to ph^-$ sample computing the sum:

$$A_{\rm HLT1} = \sum_{i=1}^{N} f_i A_{\Lambda_b^0,i}, \qquad (3.52)$$

  where i is the bin index in the $(p_{\rm T}, \log(\chi^2_{\rm IP}))$ plane, $f_i = w_i / \sum_i^N w_i$, and w_i is the sum of s-Weights in the same bin. Similarly to the L0 asymmetry, a systematic uncertainty is added by doing the same computation with efficiencies obtained from simulated samples and comparing it with the value obtained from MC truth. The difference between the two numbers is taken as a systematic uncertainty. The values of $A_{\rm HLT1}$ for each year and magnet polarity are reported in Fig. 3.42; for $\Lambda_b^0 \to pK^-$ decays the total values are $(0.01 \pm 0.35)\%$ for Run 1 and $(0.01 \pm 0.16)\%$ for Run 2, while for $\Lambda_b^0 \to p\pi^-$ they are $(0.02 \pm 0.35)\%$ for Run 1 and $(0.00 \pm 0.16)\%$ for Run 2.

3.12 Run 1 Λ_b^0 production asymmetry

The production asymmetry of a Λ_b^0 baryon was measured in Run 1 studying $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays [112], with $\Lambda_c^+ \to pK^-\pi^+$. The measurement was performed as a function of Λ_b^0 rapidity and transverse momentum, making it easy to apply the resulting maps (shown in Fig. 3.43) to a sample with a different kinematic distributions. We chose to use the rapidity-dependent values, reported in Table 3.21, to compute the production asymmetry on the signal $\Lambda_b^0 \to ph^-$ sample in Run 1. The computation is performed with the following equation:

$$A_{\rm P}(\Lambda_b^0) = \sum_i w_i A_{{\rm P},i}, \qquad (3.53)$$

where i labels the rapidity bin and w_i is the sum of the s-Weights of the $\Lambda_b^0 \to ph^-$ sample in bin i. The results of the computation are shown in Fig. 3.44; the Run 1 values are measured to be $(1.18 \pm 0.21)\%$ for $\Lambda_b^0 \to pK^-$ and $(1.12 \pm 0.21)\%$ for $\Lambda_b^0 \to p\pi^-$.

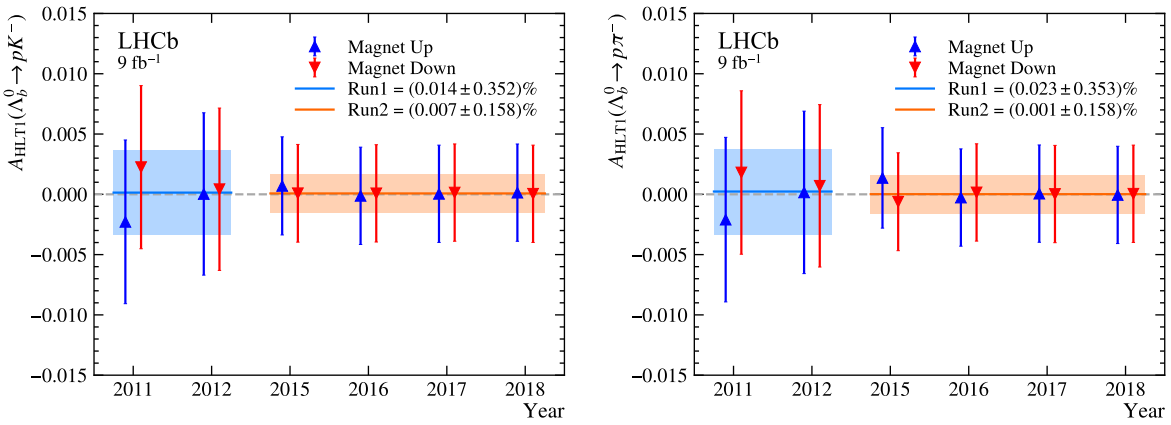

Figure 3.42: Values of $A_{\rm HLT1}$ for (left) $\Lambda_b^0 \to pK^-$ and (right) $\Lambda_b^0 \to p\pi^-$ decays, separated by year and magnet polarity.

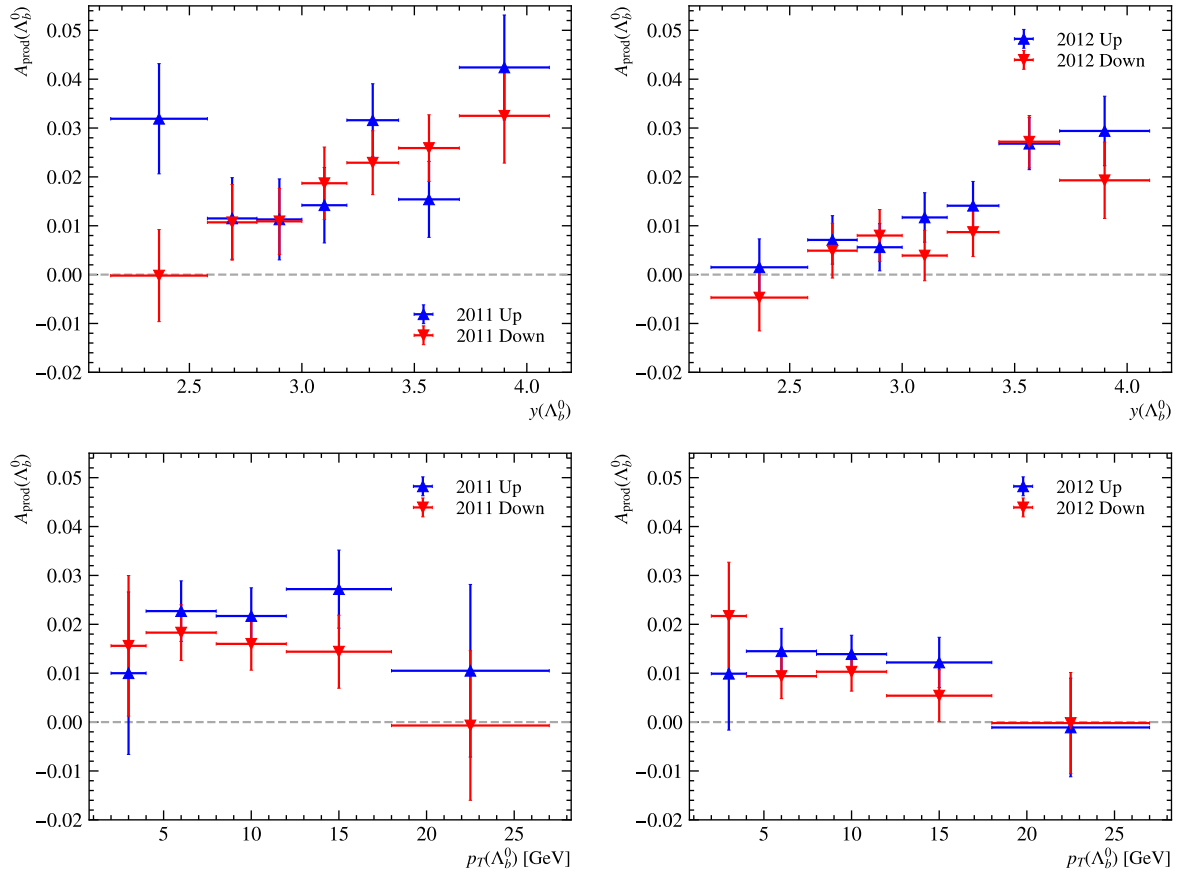

Figure 3.43: Values of the Λ_b^0 production asymmetry as a function of (top) rapidity and (bottom) transverse momentum. Left values are for 2011, right values are for 2012.

Table 3.21: Values, in percent, of the Λ_b^0 production asymmetry in 2011 and 2012, separated by magnet polarity.

Rapidity	2011		2012	
	Down	Up	Down	Up
$y \in [2.15, 2.58]$	-0.02 ± 0.94	3.19 ± 1.13	-0.47 ± 0.68	0.15 ± 0.58
$y \in [2.58, 2.80]$	1.07 ± 0.77	1.15 ± 0.83	0.49 ± 0.56	0.71 ± 0.50
$y \in [2.80, 3.00]$	1.09 ± 0.68	1.13 ± 0.83	0.8 ± 0.53	0.56 ± 0.48
$y \in [3.00, 3.20]$	1.87 ± 0.74	1.42 ± 0.77	0.39 ± 0.52	1.17 ± 0.51
$y \in [3.20, 3.43]$	2.29 ± 0.65	3.16 ± 0.75	0.87 ± 0.50	1.41 ± 0.50
$y \in [3.43, 3.70]$	2.59 ± 0.68	1.54 ± 0.78	2.72 ± 0.54	2.68 ± 0.53
$y \in [3.70, 4.10]$	3.25 ± 0.97	4.24 ± 1.07	1.93 ± 0.78	2.94 ± 0.71

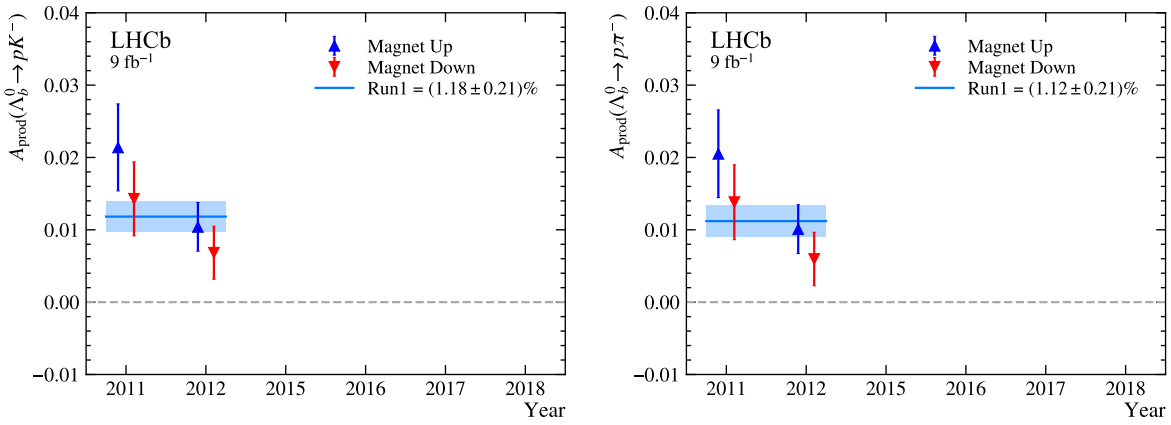

Figure 3.44: Values of the Λ_b^0 production asymmetry on the (left) $\Lambda_b^0 \to pK^-$ and (right) $\Lambda_b^0 \to p\pi^-$ samples in Run 1.

3.13 Asymmetries of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ sample

As anticipated in Section 3.1, we will use a control sample of $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\pi^-$ decays to remove the production asymmetry on the Run 2 sample by subtracting the raw asymmetries and computing the remaining instrumental asymmetries. This means that PID, trigger, and detection asymmetries for all particles in the control sample must be computed; they will be measured the same way as they were done for the signal channel, with the changes needed to account for a 4-body decay. We will describe and show in this section the computation of all nuisance asymmetries in the control sample.

The first step in this procedure is to select $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\pi^-$ decays: we adopted the same selection used in [140], which is summarised in Table 3.22, with two differences: the invariant-mass window is restricted to the range $[5.54, 6.15]$ GeV/c^2 to exclude most of the peaking background contributions from cross-feed decays, and the PID requirement ProbNNp > 0.6 on the proton is changed to $\mathrm{DLL}_{pK} > 7$, which has a similar efficiency and will allow us to employ the methods described in Section 3.3 for computing the PID asymmetry on this sample.

Next, a reweighting is performed to ensure that the production asymmetry of the Λ_b^0

Table 3.22: Selection requirements of $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\pi^-$ decays (from [140]).

Particle	Variable		Cut	Units
All	nTracks	$\in$	$[15, 500]$	
	η	$\in$	$[2, 5]$	
	`hasRich`	==	1	
	`isMuon`	==	1	
Λ_b^0	$m(\Lambda_c^+\pi^-)$	$\in$	$[5540, 6150]$	MeV/c^2
	`L0`	`Hadron_TOS OR Global_TIS`		
	`HLT1`	`TrackMVA_TOS`		
Λ_c^+	$m(pK^+\pi^-)$	$\in$	$[2266, 2306]$	MeV/c^2
	τ	$>$	0	ps
	Flight Distance (FD)	$>$	0	
	χ^2_{FD}	$>$	2	
p	p	$\in$	$[10, 150]$	GeV/c
	p_{T}	$\in$	$[1, 45]$	GeV/c
	DLL_{pK}	$>$	7	
K	p	$\in$	$[2, 150]$	GeV/c
	p_{T}	$\in$	$[0.25, 150]$	GeV/c
	$\text{DLL}_{K\pi}$	$>$	0	
π from Λ_c^+	p	$\in$	$[2, 150]$	GeV/c
	p_{T}	$\in$	$[0.25, 150]$	GeV/c
	$\text{DLL}_{K\pi}$	$<$	5	
π from Λ_b^0	$\text{DLL}_{K\pi}$	$<$	0	

baryon in the two samples cancels out; this is done by training a BDT classifier using the `hep_ml` [134] package to make the kinematics of the control sample match with the signal sample; the variables used for the training are the momentum and pseudorapidity of the Λ_b^0, and the reweighting is done without splitting it by year and magnet polarity as the kinematic distributions of all the subsamples within a given Run period were found to be similar, as show in Fig. 3.45. The distributions before and after reweighting are shown in Fig. 3.46.

On this reweighted samples, raw asymmetries are extracted with simultaneous maximum-likelihood fits to the invariant-mass distributions of *CP*-conjugated final states. The signal events are modelled with a Crystal Ball distribution [135]; the combinatorial events are modelled with an exponential function; the cross-feed events from $\Lambda_b^0 \to \Lambda_c^+ K^-$ decays are modelled with a Gaussian function with width in common with the signal shape and mean restricted to the range $[5.56, 5.58]$ GeV/c^2. The fits are performed separately for each year and magnet polarity subsample. The plots of the fits in all subsamples are shown in Chapter D, while the raw asymmetries are shown in Figure 3.47.

After extracting the raw asymmetries, the pion, kaon, and proton detection, as well as the trigger and PID asymmetries must be measured; they are computed with the same procedure described in Sections 3.9 to 3.11, with the difference that in this case

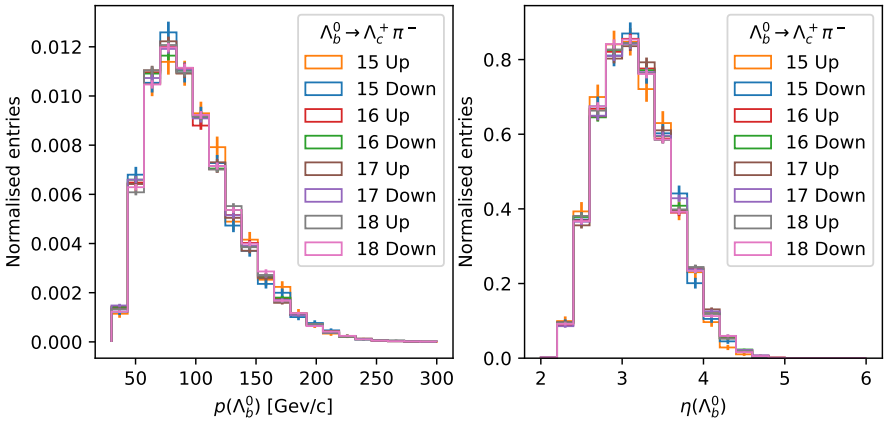

Figure 3.45: Kinematic distributions of the $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ sample in Run 2 divided by year and magnet polarity.

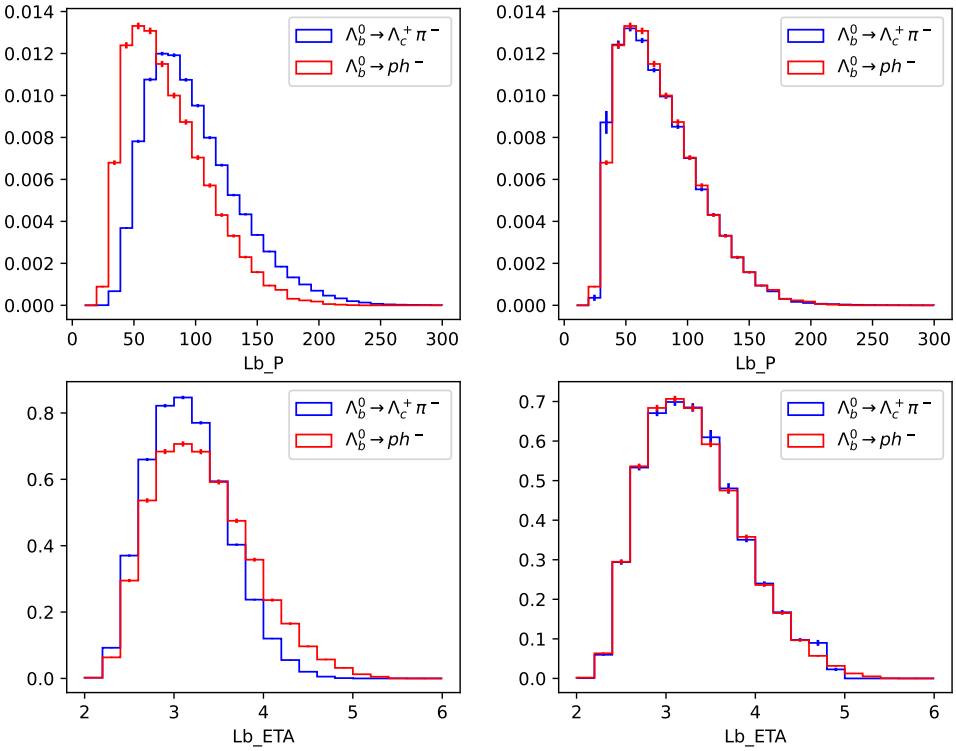

Figure 3.46: Kinematic distribution of Run 2 $\Lambda_b^0 \to p h^-$ and $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ before (left) and after (right) reweighting.

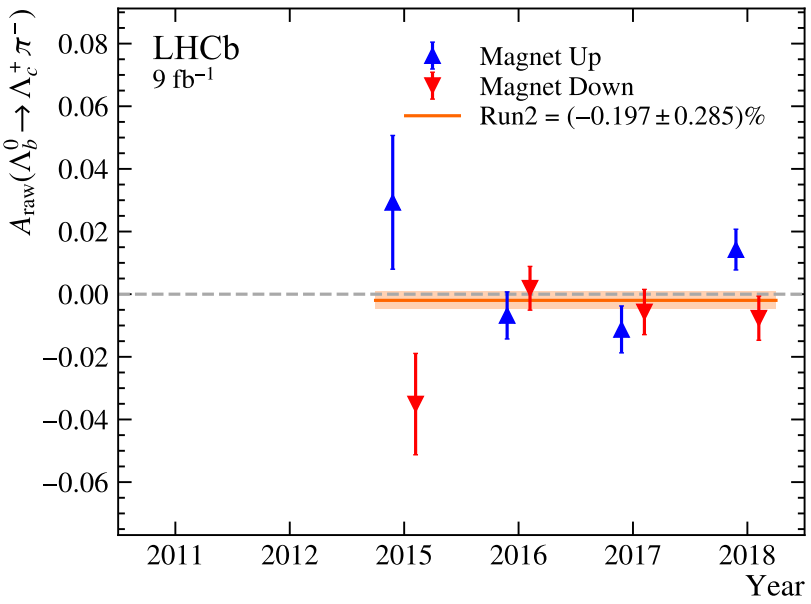

Figure 3.47: Raw asymmetries of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ sample, divided by year and magnet polarity.

the final state is made of four particles instead of two. For example, the efficiencies in Equations (3.43) and (3.44) become[5]:

$$\varepsilon(\Lambda_b^0) = 1 - (1-\varepsilon_p^+)(1-\varepsilon_K^-)(1-\varepsilon_\pi^+)(1-\varepsilon_\pi^-), \tag{3.54}$$

$$\varepsilon(\overline{\Lambda}_b^0) = 1 - (1-\varepsilon_p^-)(1-\varepsilon_K^+)(1-\varepsilon_\pi^-)(1-\varepsilon_\pi^+). \tag{3.55}$$

The instrumental asymmetries for $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays are shown in Figure 3.48, divided by year and magnet polarity. In the plots, the label π_1 indicates the pion from the decay of the Λ_b^0, while π_2 indicates the pion from the decay of the Λ_c^+. For illustration purposes, the average values for Run 2 are also shown, although they are not used in the averaging procedure to obtain the final results due to the correlations between all corrections, as will be explained in the next section.

3.14 Combination of corrections to the raw asymmetries

In order to combine all the corrections to the raw asymmetries into a single value, their correlations must be taken into account. Such correlations arise from the fact that some information used to determine the corrections are used multiple times. For example, the efficiency tables used to determine the PID asymmetries (see Sections 3.3 and 3.10) are used to determine the correction for both the $\Lambda_b^0 \to ph^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays, as it can be seen by looking at Eqs. (3.11) and (3.12).

Another correlation to be taken into account is that among the different data taking periods. In this case, for example, the correlation comes from the proton detection asymmetry (see Section 3.9.3), where the correction determined in bins of proton momentum from 2012 data is used as well for all the Run 2 samples.

[5]The formulas shown here are shorthands for requesting that at least one of the four particles be `TOS` on `L0_Hadron`.

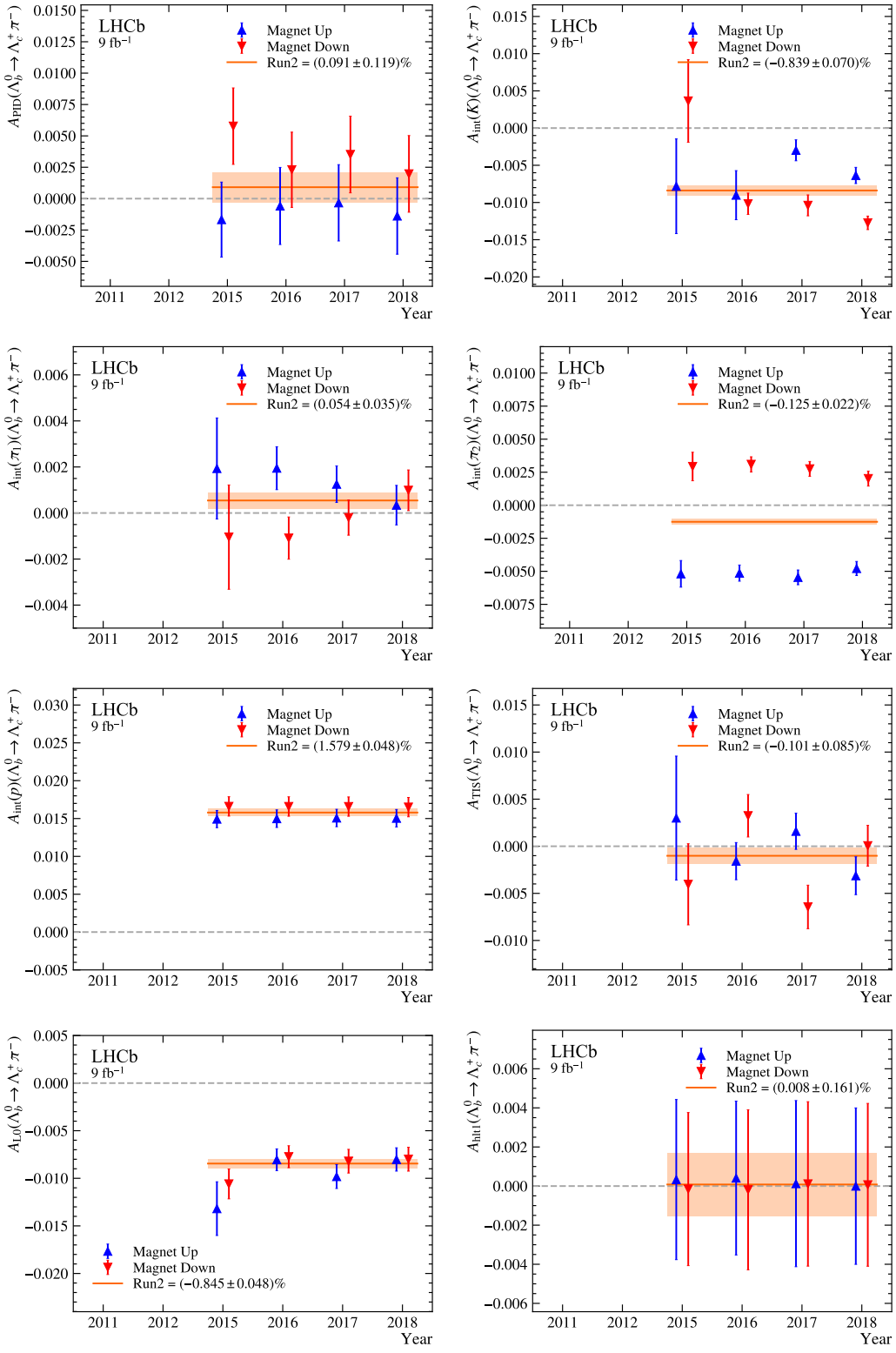

Figure 3.48: Instrumental asymmetries of $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays. From top to bottom and from left to right: PID, K detection, π^- detection, π^+ detection, proton detection, `TIS`, `L0_Hadron`, `Hlt1`. For illustration purposes, the average values for Run 1 and Run 2 are also shown, although they are not used in the averaging procedure to obtain the final results due to the correlations between al corrections.

Hence, in order to properly combine all the corrections, a set of pseudoexperiments have been run. The strategy to generate the pseudoexperiments starts with the identifications of the basic parts of the measurement that are completely independent from each other; they are:

$K^-\pi^+$ detection asymmetries the detection asymmetry of a $K^-\pi^+$ pair in bins of kaon kinematic from Fig. 3.32 is completely uncorrelated from any other corrections, but the same maps are used twice for computing the kaon detection asymmetry both for the signal and control sample;

Pion detection asymmetries the pion detection asymmetries in bins of pion kinematic from Figs. 3.28 and 3.29 is used twice to compute the correction for the signal and control sample; in addition, it is also used to obtain the kaon detection asymmetry from the $K^-\pi^+$ detection asymmetry, introducing an additional correlation;

Raw asymmetries of $\Lambda_b^0 \to \Lambda_c^+\pi^-$ the raw asymmetries obtained from Fig. 3.47 are uncorrelated from everything else;

Proton detection asymmetry the proton detection asymmetry as a function of proton momentum in Table 3.18 and Fig. 3.34 is only available for Run 1; for all the Run 2 subsamples, the 2012 map will be reused;

PID asymmetries the PID-efficiency tables from Section 3.3 (see Fig. 3.4) are independent for all the years, but are used twice to compute the correction on the signal and control sample.

L0 `TOS` asymmetries the tables with L0_TOS asymmetries as a function of E_{T} and HCAL region from Chapter B are independent for all the years, but are used twice to compute the correction on the signal and control sample.

L0 `TIS` asymmetries the tables with L0_TIS asymmetries as a function of the p_{T} of the Λ_b^0 from Fig. 3.37 are independent for all the years, but are used twice to compute the correction on the signal and control sample.

HLT1 asymmetries the tables with the Hlt1 efficiencies as a function of p_{T} and χ^2_{IP} of particles from Chapter C are independent for all the years, but are used twice to compute the correction on the signal and control sample.

Λ_b^0 production asymmetries the Run 1 values of the Λ_b^0 production asymmetry are independent for all years and magnet polarities, but the are correlated to the corresponding proton interaction asymmetry maps, as they are measured from the same sample.

In the next step of the procedure, the information summarised in the list above, in form of tables, is bootstrapped and new tables are created. The central values in each bin of the tables are randomly extracted using Gaussian functions with mean and width corresponding to central values and uncertainties from the original tables. At this point, a new set of base informations necessary to determine all the corrections needed to compute the *CP* asymmetries for each year and magnet polarity are available, and are combined according to Eqs. (3.3) and (3.4) for Run 1 and Eqs. (3.11) and (3.12) for Run 2. Finally,

the bootstrapping is repeated 1000 times, each time determining the final values for the *CP* asymmetries.

To obtain a final value of $A_{CP}(pK^-)$ and $A_{CP}(p\pi^-)$, we will average the 12 measurements in each year and magnet subsample using the prescription from Ref. [141], which we summarise here: we define two 12×12 covariance matrices, one containing only the statistical uncertainties of the 12 measurements, which are completely uncorrelated to all other uncertainties and to each other, and one containing only the systematic uncertainties, including their correlations estimated from the distribution of the 1000 bootstrapped *CP* asymmetries:

$$V_{\text{stat}} = \begin{pmatrix} \left(\sigma^{\text{1U}}_{\text{stat}}\right)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \left(\sigma^{\text{8D}}_{\text{stat}}\right)^2 \end{pmatrix}, \tag{3.56}$$

$$V_{\text{syst}} = \begin{pmatrix} \left(\sigma^{\text{1U}}_{\text{syst}}\right)^2 & \cdots & \rho_{\text{1U8D}}\sigma^{\text{1U}}_{\text{syst}}\sigma^{\text{8D}}_{\text{syst}} \\ \vdots & \ddots & \vdots \\ \rho_{\text{1U8D}}\sigma^{\text{1U}}_{\text{syst}}\sigma^{\text{8D}}_{\text{syst}} & \cdots & \left(\sigma^{\text{18D}}_{\text{syst}}\right)^2 \end{pmatrix}, \tag{3.57}$$

where we used the notation σ^{YM}, with $\text{Y} \in \{1, 2, 5, 6, 7, 8\}$ and $\text{M} \in \{\text{U}, \text{D}\}$ to indicate the A_{CP} uncertainty of the 201Y magnet M subsample, and similarly $\rho_{\text{YMY'M'}}$ to indicate the correlation coefficient between the *CP* asymmetry measured in the two YM and Y'M' subsamples. The correlation matrices of the 12 measurements are shown in Fig. 3.49 for $A_{CP}(pK^-)$ and $A_{CP}(p\pi^-)$.

With the statistical and systematic matrices, we define the total covariance matrix $V = V_{\text{stat}} + V_{\text{syst}}$, which we can use to find the optimal averages of all the 12 measurements of $A_{CP}(pK^-)$ and $A_{CP}(p\pi^-)$ as

$$\overline{A}_{CP} = \left(\sum_i \sum_j \left(V^{-1}\right)_{ij}\right)^{-1} \left(\sum_i \sum_j \left(V^{-1}\right)_{ij} A_{CPj}\right), \tag{3.58}$$

with uncertainty given by

$$\sigma^2\left(\overline{A}_{CP}\right) = \left(\sum_i \sum_j \left(V^{-1}\right)_{ij}\right)^{-1}. \tag{3.59}$$

To obtain separate values of A_{CP} for Run 1 and Run 2, the procedure is applied only on the corresponding subset of the 12 measurements.

3.15 Results

In this section we present the results of the averaging procedure described in Section 3.14. The results of the 1000 pseudoexperiments are shown in Fig. 3.50 for all 12 subsamples, while the Run 1 and 2 values of all the experimental asymmetries contributing to A_{CP} are reported in Tables 3.23 and 3.24. The results for the Run 1 sample are:

$$A_{CP}(pK^-) = (0.09 \pm 1.53 \pm 0.72)\%,$$
$$A_{CP}(p\pi^-) = (-0.52 \pm 1.89 \pm 0.56)\%,$$

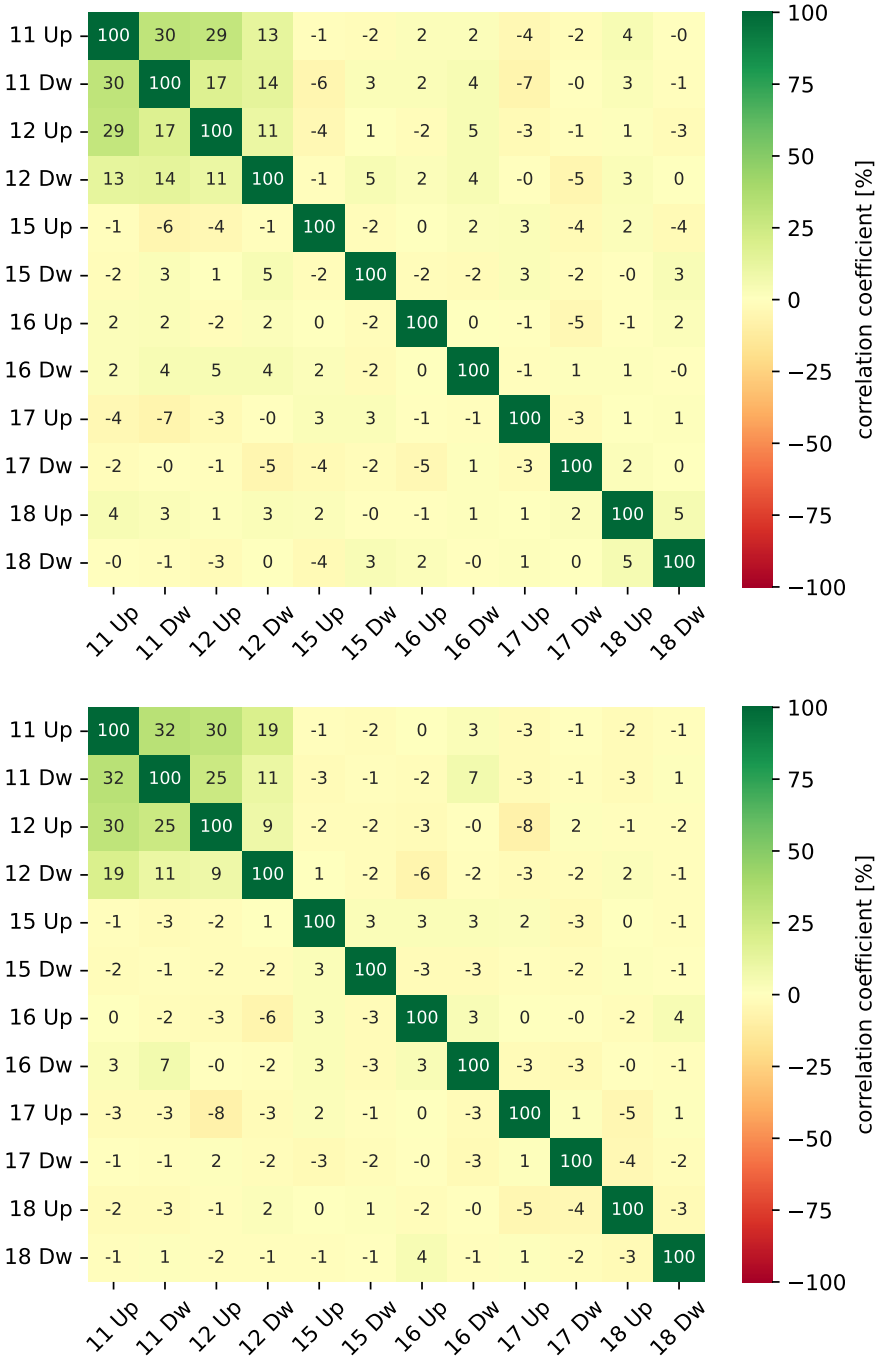

Figure 3.49: Correlation matrices of (top) $A_{CP}(p\pi^-)$ and (bottom) $A_{CP}(pK^-)$ measured in each Run 1+2 year and magnet polarity. Only systematic uncertainties are considered, *i.e.* these correspond to the matrices defined in Eq. (3.57), excluding the magnitudes of the systematic errors.

Table 3.23: Numerical results of all the experimental asymmetries contributing to A_{CP} for Run 1. The numbers are just for illustration purposes, as the values of A_{CP} shown on the bottom are not the sums of these corrections, but come from the combination of the measurements in the 4 Run 1 subsamples, according to Eqs. (3.58) and (3.59).

	$\Lambda_b^0 \to pK^-$	$\Lambda_b^0 \to p\pi^-$
A_{raw}	$(+0.86 \pm 1.53)\%$	$(+0.90 \pm 1.89)\%$
A_{PID}	$(-0.40 \pm 0.56)\%$	$(-0.36 \pm 0.59)\%$
A_{TOS}	$(-1.14 \pm 0.19)\%$	$(-1.12 \pm 0.20)\%$
A_{TIS}	$(+0.46 \pm 0.35)\%$	$(+0.66 \pm 0.36)\%$
A_{hlt1}	$(+0.01 \pm 0.35)\%$	$(+0.02 \pm 0.35)\%$
$A_{\text{D}}(p)$	$(+1.45 \pm 0.08)\%$	$(+1.44 \pm 0.08)\%$
$A_{\text{D}}(K^-)$	$(-0.69 \pm 0.12)\%$	-
$A_{\text{D}}(\pi^-)$	-	$(+0.10 \pm 0.11)\%$
$A_{\text{P}}(\Lambda_b^0)$	$(1.18 \pm 0.21)\%$	$(1.12 \pm 0.21)\%$
A_{CP}	$(0.09 \pm 1.69)\%$	$(-0.52 \pm 1.97)\%$

with a correlation of 3%, where the first uncertainty is statistical and the second is systematic. They are compatible with the previous results [13], that are respectively $(-2.0 \pm 1.3 \pm 1.9)\%$ and $(-3.5 \pm 1.7 \pm 2.0)\%$. It can be seen how the systematic uncertainty have been reduced by a factor 2 thanks to the improvements in the analysis, which means that these numbers supersede the previous ones.
The results for the Run 2 sample are:

$$A_{CP}(pK^-) = (-1.45 \pm 0.75 \pm 0.43)\%,$$
$$A_{CP}(p\pi^-) = (0.34 \pm 0.95 \pm 0.43)\%,$$

with a correlation of 15%. The combination with the Run 1 values yield:

$$A_{CP}(pK^-) = (-1.12 \pm 0.67 \pm 0.36)\%,$$
$$A_{CP}(p\pi^-) = (0.15 \pm 0.85 \pm 0.36)\%,$$

with a total correlation of 10%. These results are the most precise measurements of the CP asymmetries in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays from a single experiment, and they are compatible with the CP symmetry conservation hypothesis within 1.5σ and 0.2σ, respectively.

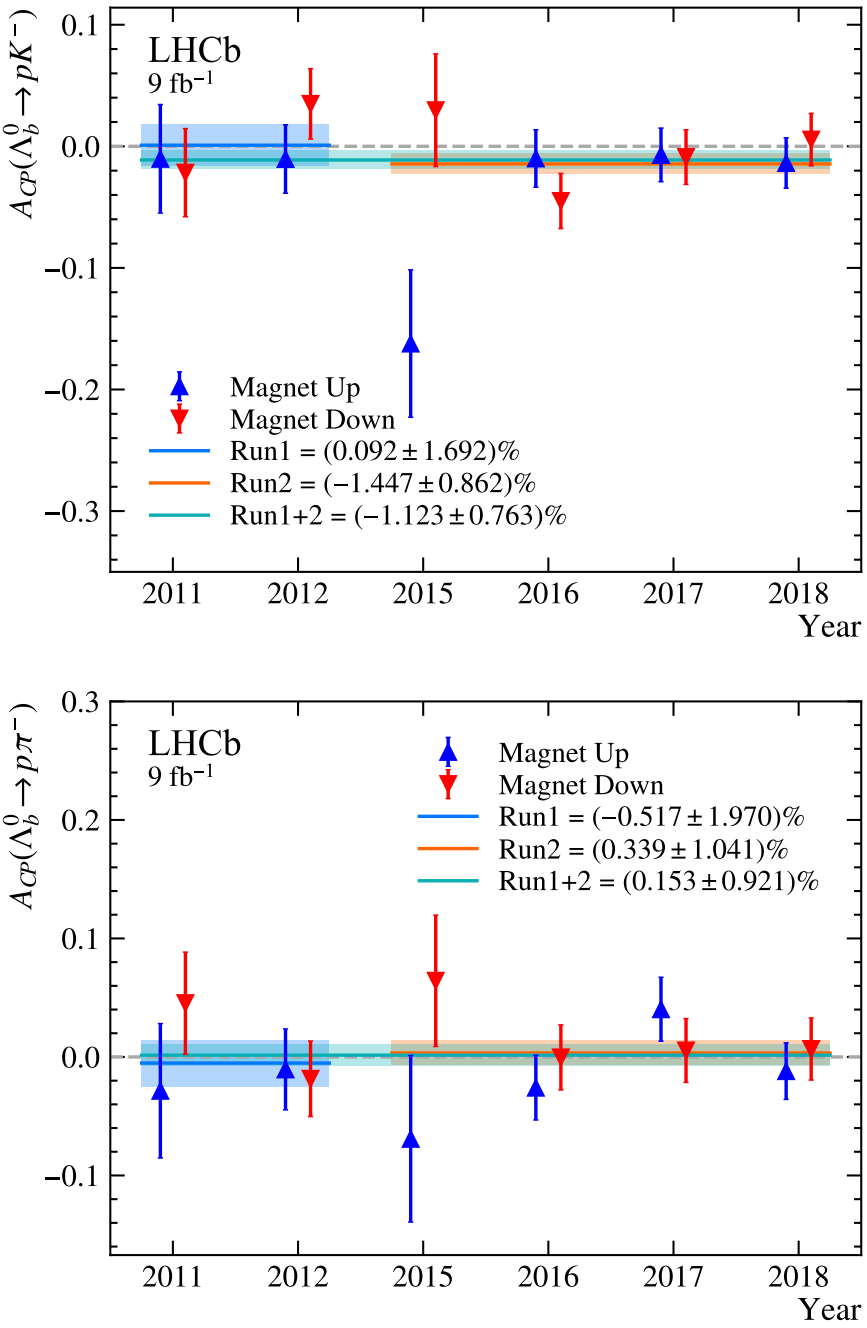

Figure 3.50: Values of $A_{CP}(pK^-)$ (top) and $A_{CP}(p\pi^-)$ (bottom), divided by year and magnet polarity, including statistical and systematic uncertainties. The average values for Run 1, Run 2, and their combination, computed according to Eqs. (3.58) and (3.59), are also shown.

Table 3.24: Numerical results of all the experimental asymmetries contributing to A_{CP} for Run 2. The numbers are just for illustration purposes, as the values of A_{CP} shown on the bottom are not the sums of these corrections, but come from the combination of the measurements in the 8 Run 2 subsamples, according to Eqs. (3.58) and (3.59).

	$\Lambda_b^0 \to pK^-$	$\Lambda_b^0 \to p\pi^-$	$\Lambda_b^0 \to \Lambda_c^+\pi^-$
A_{raw}	$(-0.86 \pm 0.75)\%$	$(+0.14 \pm 0.95)\%$	$(-0.20 \pm 0.29)\%$
A_{PID}	$(+0.03 \pm 0.12)\%$	$(-0.20 \pm 0.30)\%$	$(+0.09 \pm 0.12)\%$
A_{TOS}	$(-1.29 \pm 0.08)\%$	$(-1.16 \pm 0.09)\%$	$(-0.85 \pm 0.05)\%$
A_{TIS}	$(-0.05 \pm 0.05)\%$	$(-0.17 \pm 0.13)\%$	$(-0.10 \pm 0.09)\%$
A_{hlt1}	$(+0.01 \pm 0.16)\%$	$(+0.00 \pm 0.16)\%$	$(+0.01 \pm 0.16)\%$
$A_{\text{D}}(p)$	$(+1.31 \pm 0.04)\%$	$(+1.43 \pm 0.06)\%$	$(+1.58 \pm 0.05)\%$
$A_{\text{D}}(K^-)$	$(-0.68 \pm 0.05)\%$	-	$(-0.84 \pm 0.07)\%$
$A_{\text{D}}(\pi^-)$	-	$(+0.02 \pm 0.07)\%$	$(+0.05 \pm 0.04)\%$
$A_{\text{D}}(\pi^+)$	-	-	$(-0.13 \pm 0.02)\%$
A_{CP}	$(-1.45 \pm 0.86)\%$	$(0.34 \pm 1.04)\%$	-

Conclusions

In this thesis, a measurement of the *CP* asymmetries of $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays using the full Run 1+2 dataset collected by the LHCb experiment has been presented. The data sample corresponds to an integrated luminosity of $9\,\text{fb}^{-1}$ of pp collisions at centre-of-mass energies of 7, 8 and 13 TeV. The Run 1 analysis [13] was repeated to profit from the developments in the estimation of the Λ_b^0 production asymmetry [112] and from the use of a novel technique to compute the trigger-induced asymmetries.

After a multivariate selection aimed at finding the optimal selection requirements to minimise the uncertainties on the A_{CP} observables, the signal yields have been computed from maximum-likelihood fits to the invariant-mass spectra of the eight $H_b \to h^+h'^-$ decays. A simultaneous invariant-mass fit allows the yields of mis-identified decays in the target spectrum signal window to be determined. Asymmetries due to PID requirements have been estimated with calibration samples, building 2-dimensional mis-ID efficiency maps in the (p, η) plane that have been integrated over the signal kinematic distributions. The effect of the trigger selection on the raw asymmetries has been evaluated from control samples of $B^+ \to J/\psi K^+$, $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\mu\nu X$, and $B^0 \to (D^0 \to K^+\pi^-)\mu\nu X$ decays, used respectively to compute the impact on the asymmetry of the specific trigger algorithms used to collect the data. Detection asymmetries for final-state particles have been either computed directly (for kaons) or taken from already available measurements done by the collaboration (for pions and protons).

The production asymmetry of the Λ_b^0 baryon was measured in Run 1 but not in Run 2, therefore for the latter sample we used a control sample of $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\pi^-$ decays to remove it by subtracting the raw asymmetries between the signal and control sample, and computing the other instrumental asymmetries also on the latter. This has introduced correlations between the various terms, since the efficiencies needed to compute every correction were used identically both on the signal and control sample; such correlations have been taken into account when computing the average of the measurements of A_{CP} in all the year and magnet polarity subsamples.

The results of the analysis on the Run 1 sample are:

$$\begin{aligned} A_{CP}(pK^-) &= (0.09 \pm 1.53 \pm 0.72)\%, \\ A_{CP}(p\pi^-) &= (-0.52 \pm 1.89 \pm 0.56)\%, \end{aligned}$$

where the first uncertainty is statistical and the second is systematic, and treir mutual correlation is 3%. They are compatible with the previous results [13] while having the systematic uncertainties reduced by a factor 2, thus superseding them. The results on the

Run 2 sample are:

$$A_{CP}(pK^-) = (-1.45 \pm 0.75 \pm 0.43)\%,$$
$$A_{CP}(p\pi^-) = (0.34 \pm 0.95 \pm 0.43)\%,$$

with a correlation of 15%. Upon combination with the Run 1 results they yield the final values:

$$A_{CP}(pK^-) = (-1.12 \pm 0.67 \pm 0.36)\%,$$
$$A_{CP}(p\pi^-) = (0.15 \pm 0.85 \pm 0.36)\%.$$

with a total correlation of 10%. These are the most precise measurements of the *CP* asymmetries in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays from a single experiment. They are compatible with 0 within 1.5σ and 0.2σ, respectively, indicating *CP* symmetry conservation.

Fits for the $K^-\pi^+$ asymmetry

$K^-\pi^+\pi^+$ sample

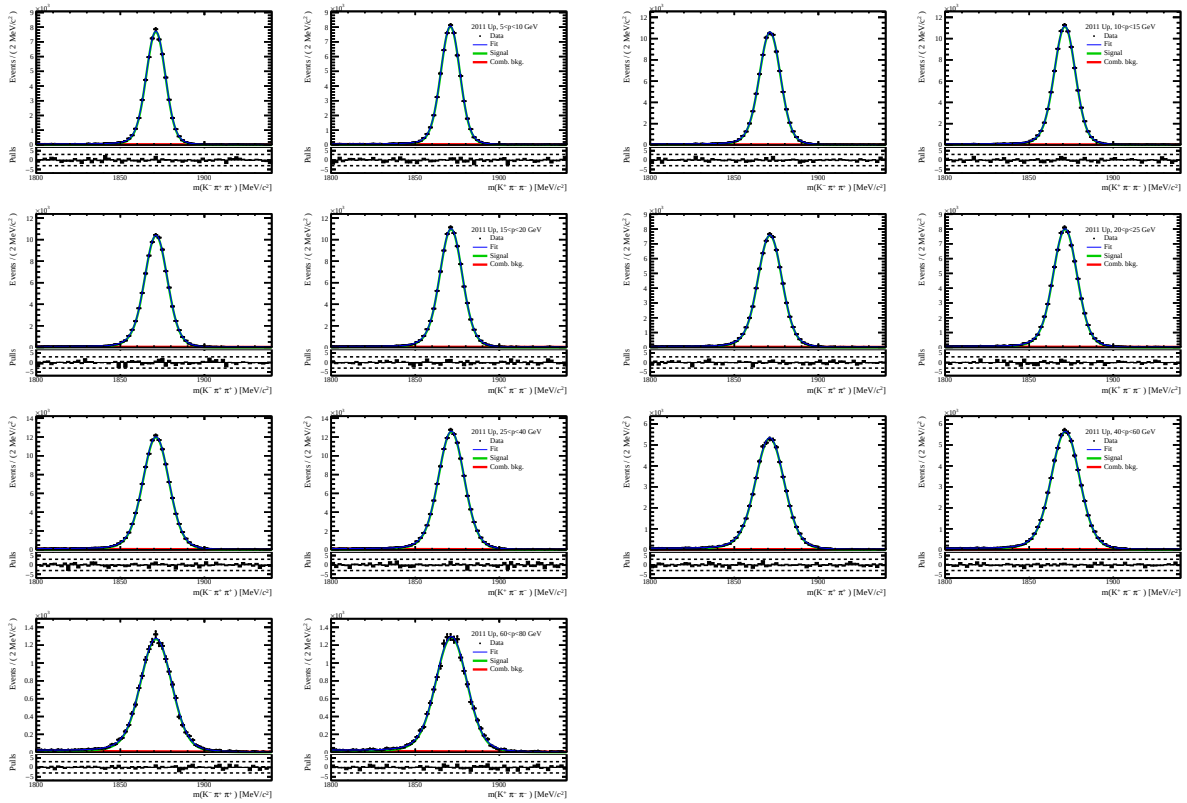

Figure 51: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2011 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

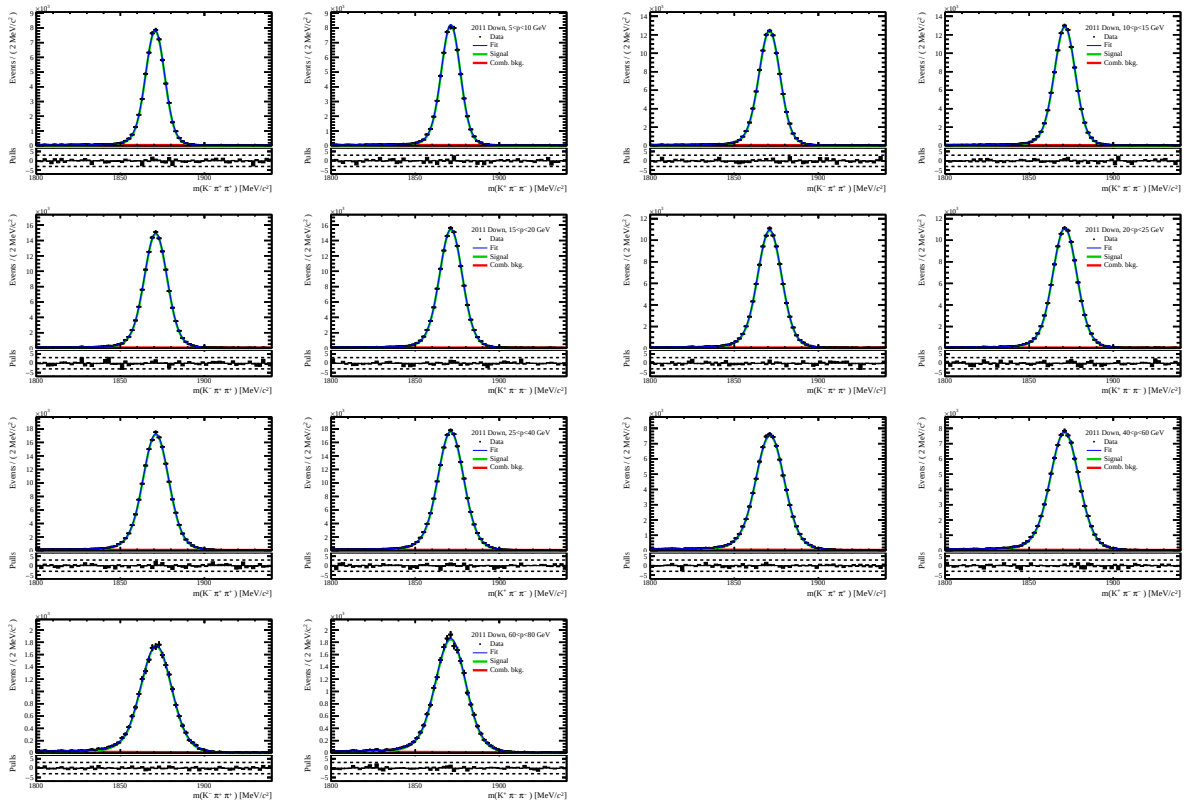

Figure 52: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2011 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

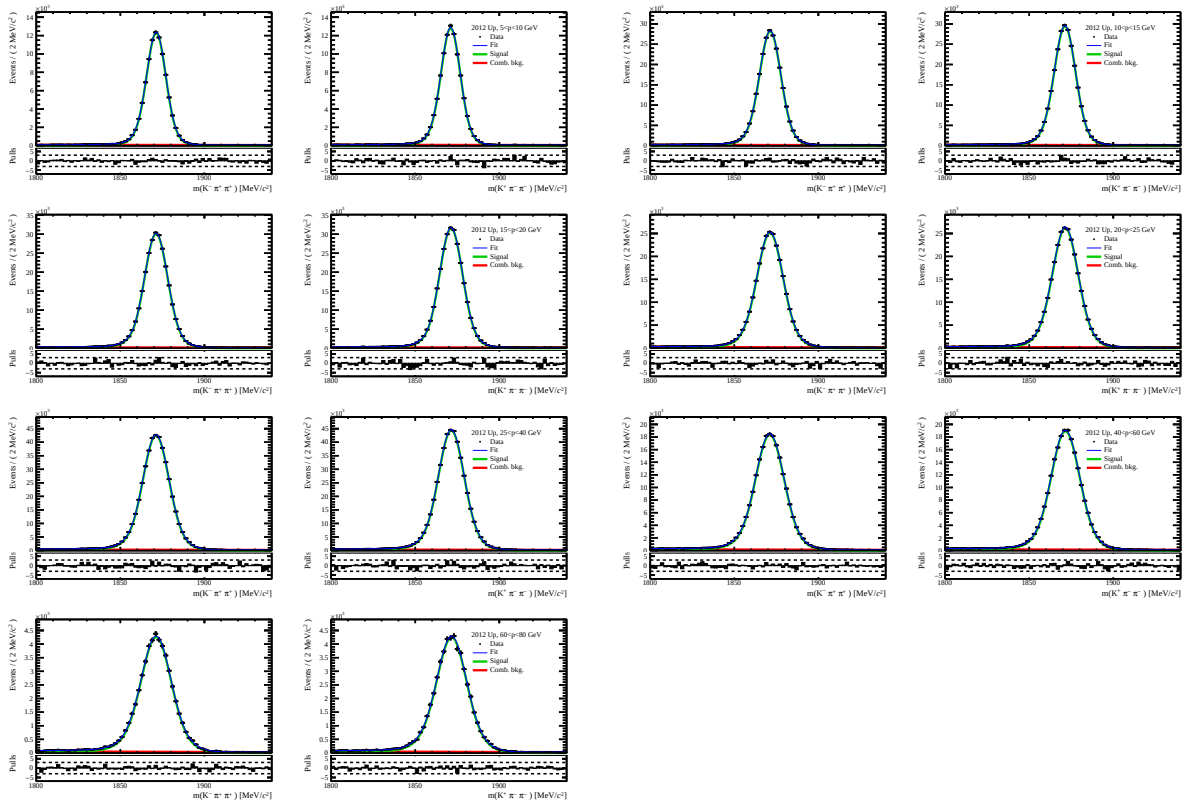

Figure 53: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2012 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

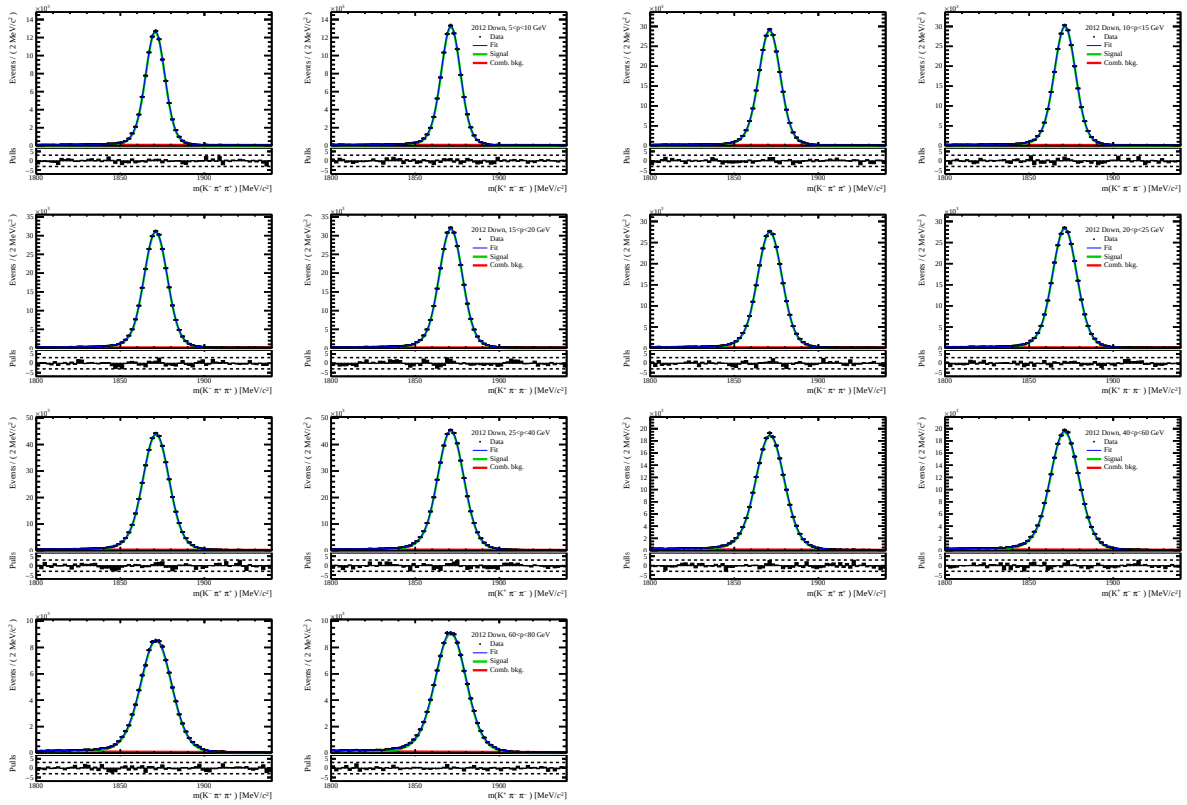

Figure 54: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2012 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

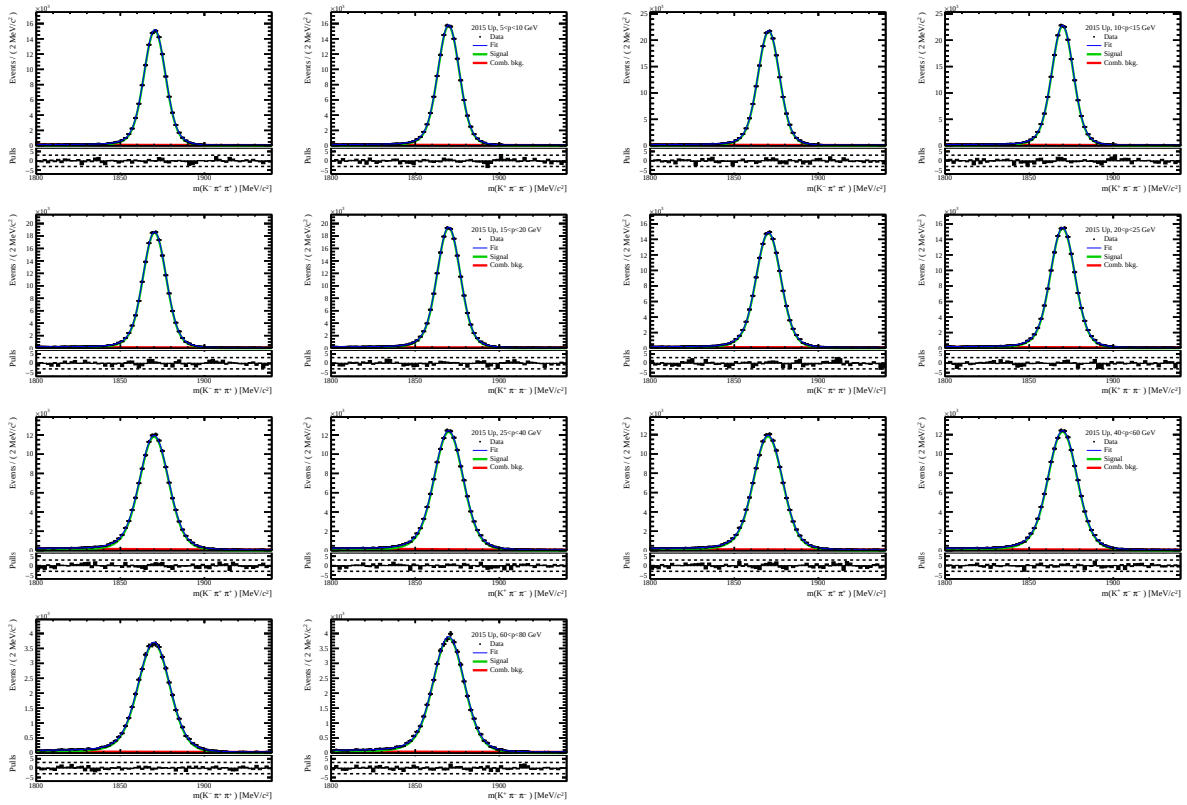

Figure 55: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2015 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

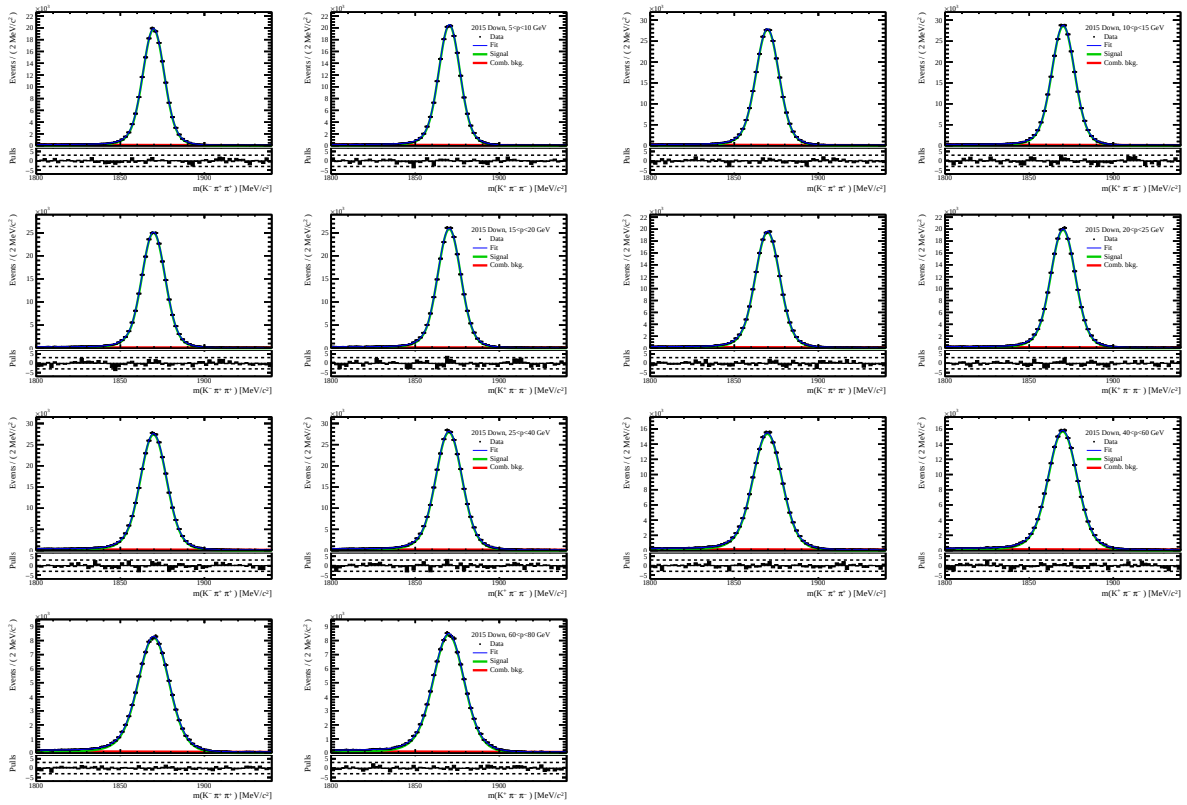

Figure 56: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2015 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

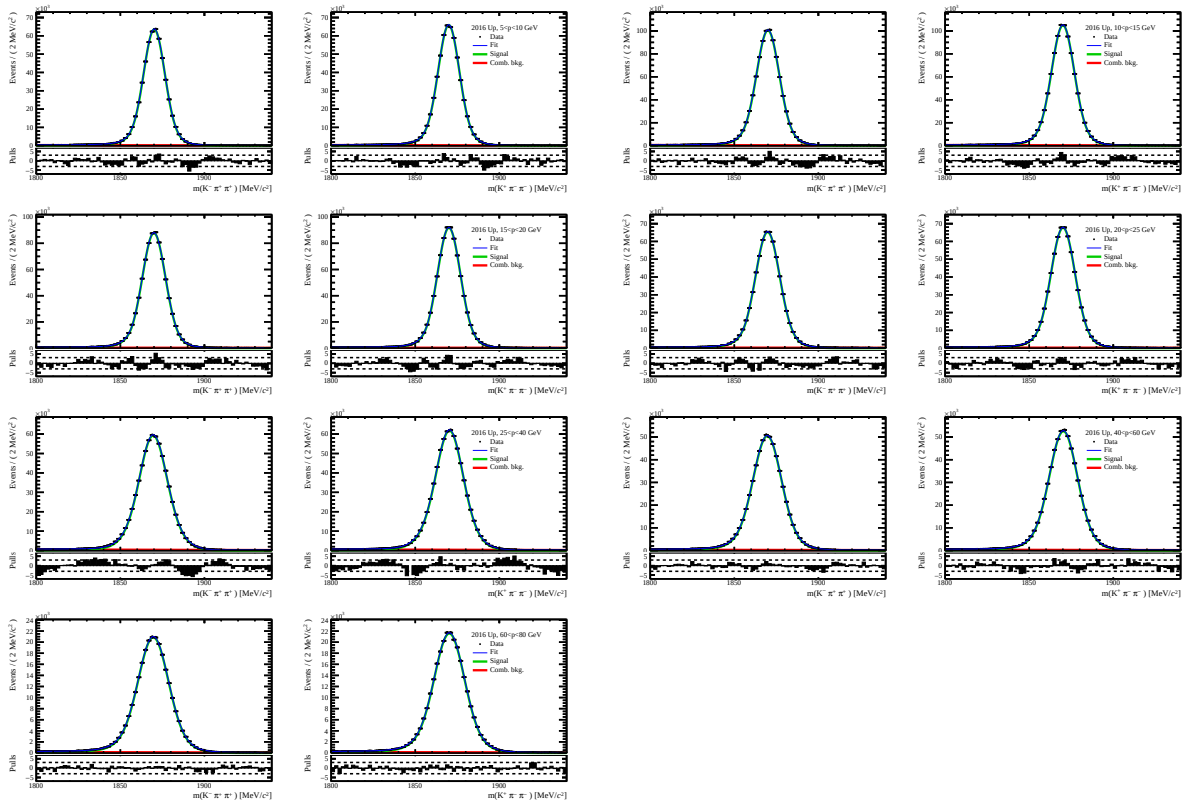

Figure 57: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2016 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

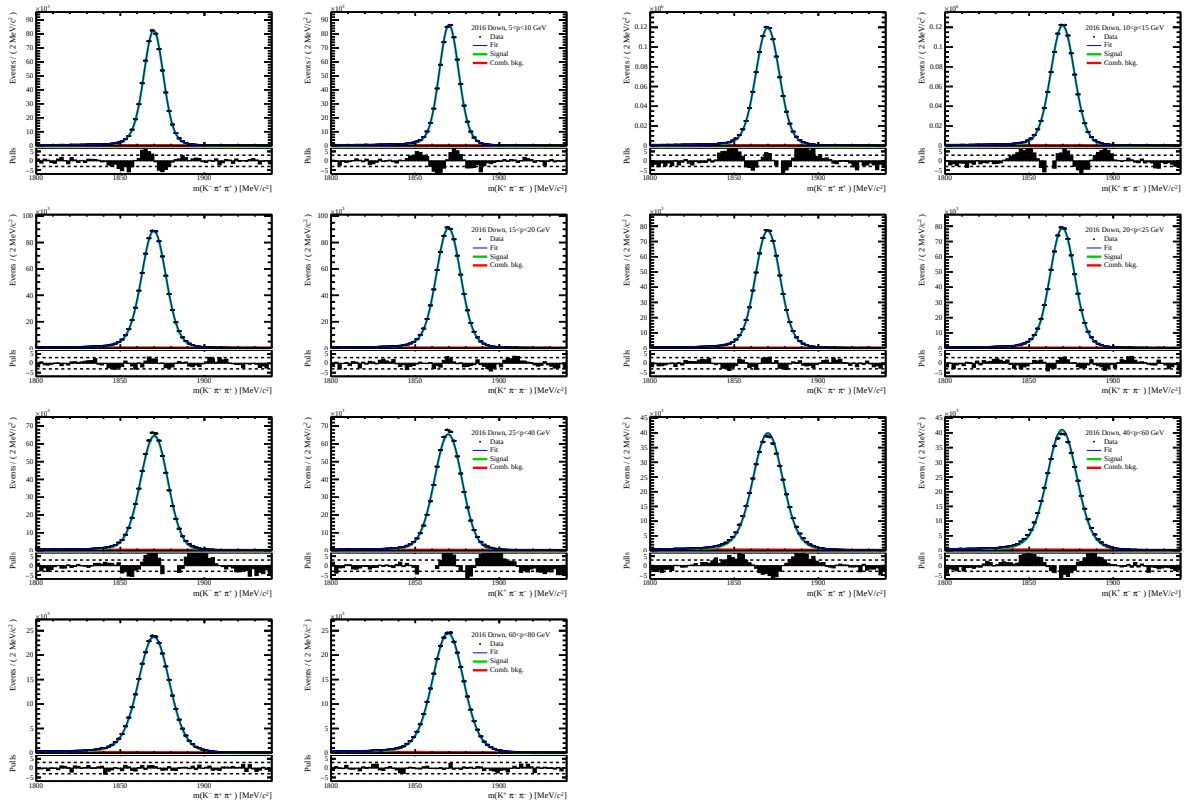

Figure 58: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2016 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

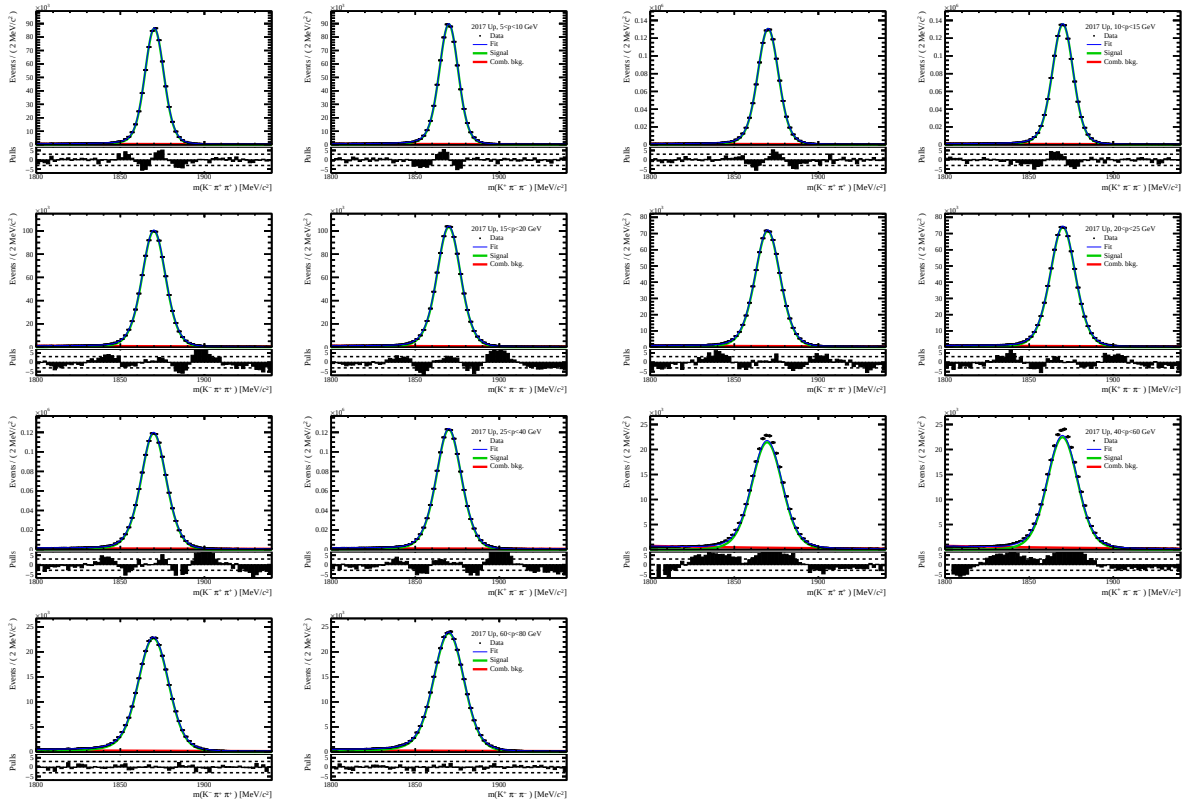

Figure 59: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2017 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

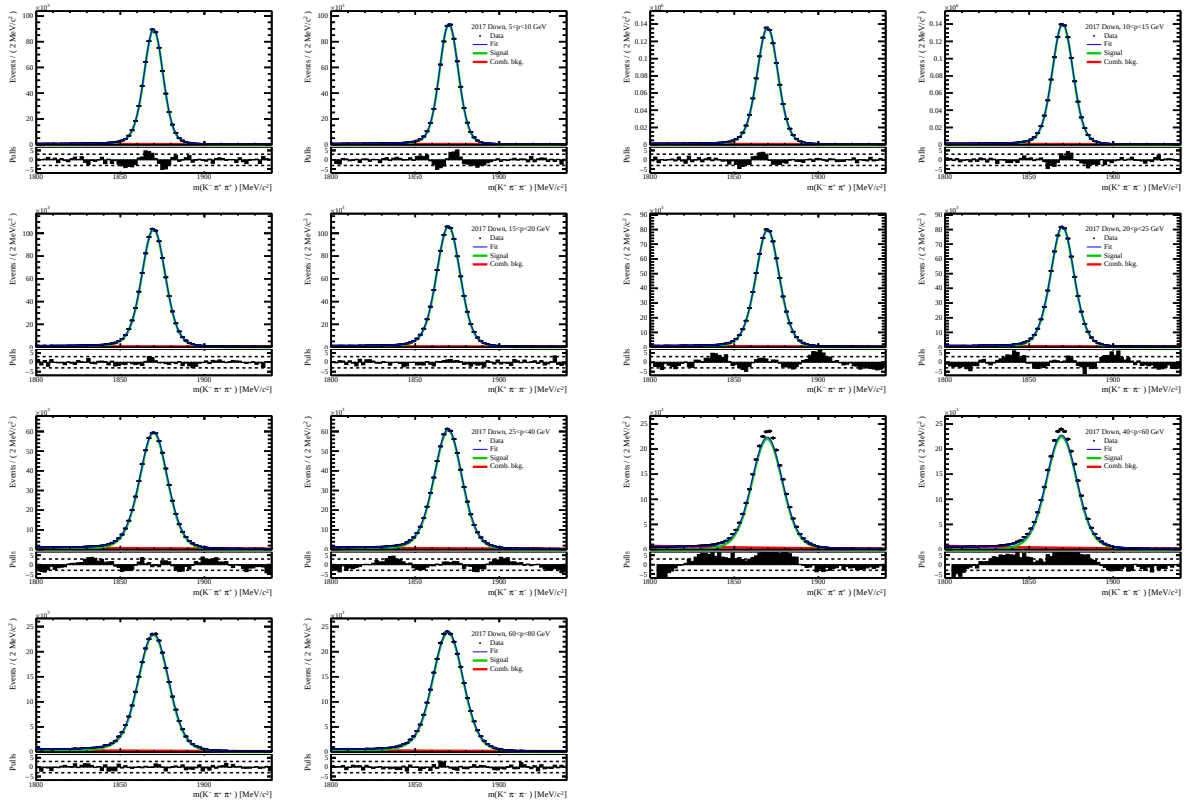

Figure 60: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2017 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

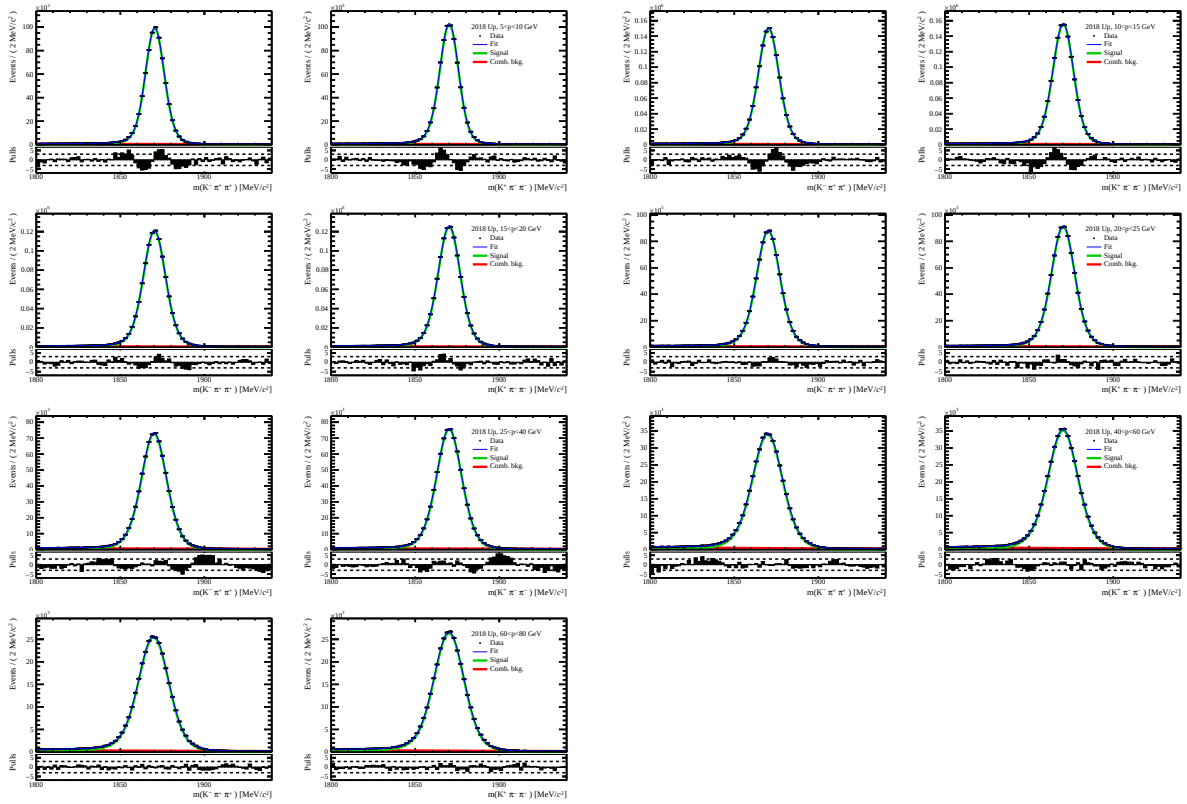

Figure 61: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2018 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

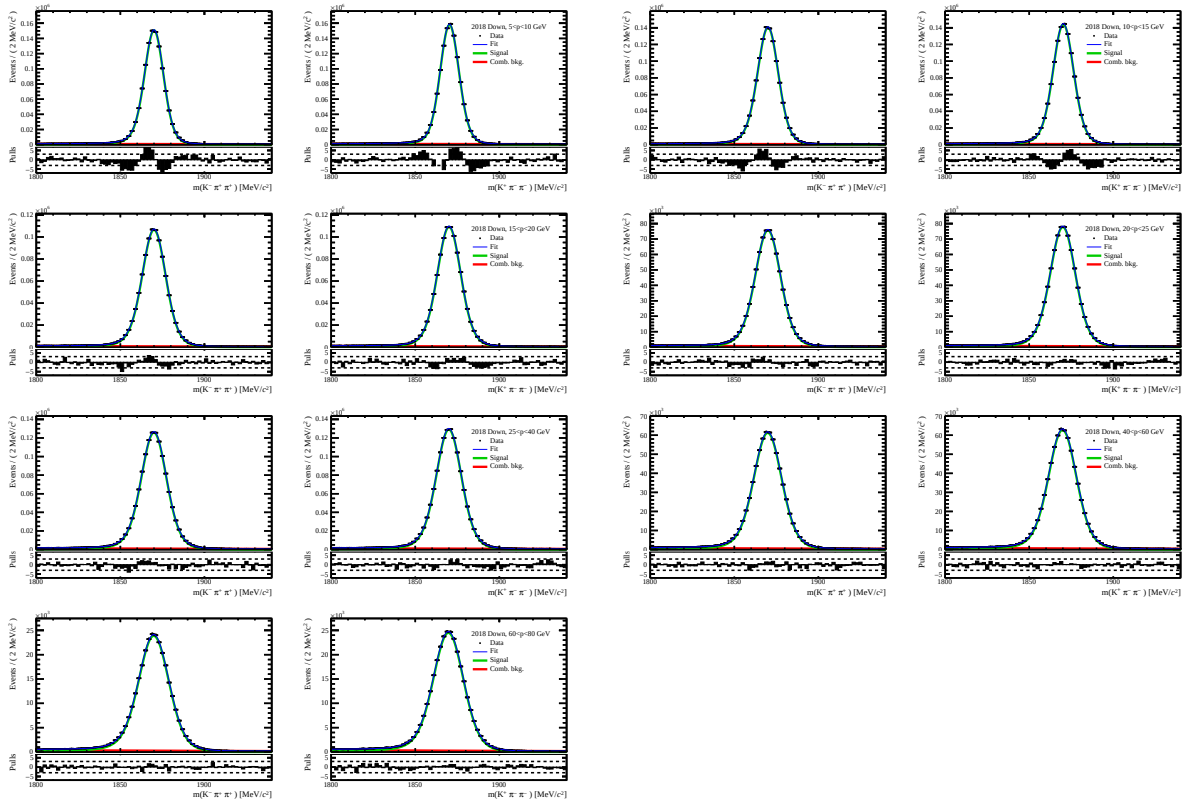

Figure 62: Fits to the invariant-mass distributions of (left of the pair) $K^-\pi^+\pi^+$ and (right of the pair) $K^+\pi^-\pi^-$ decays in the year 2018 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

$K^0_S\pi^+$ sample

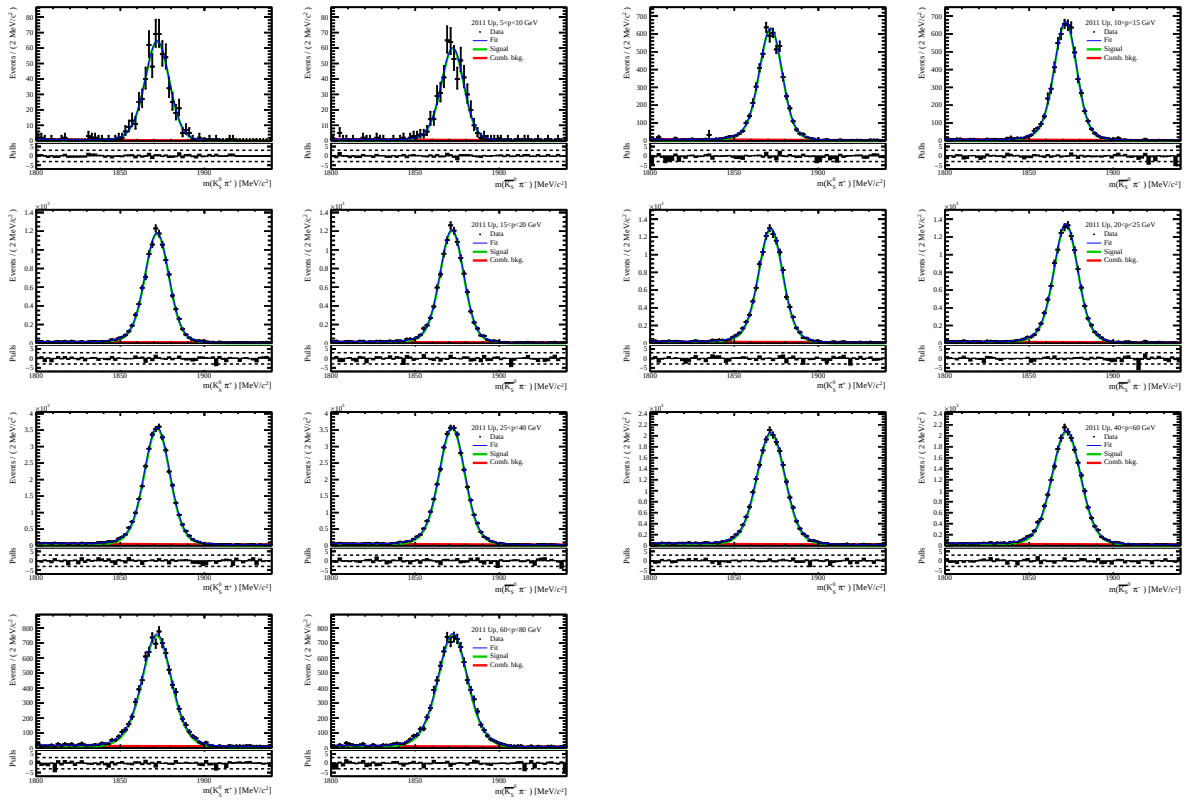

Figure 63: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2011 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

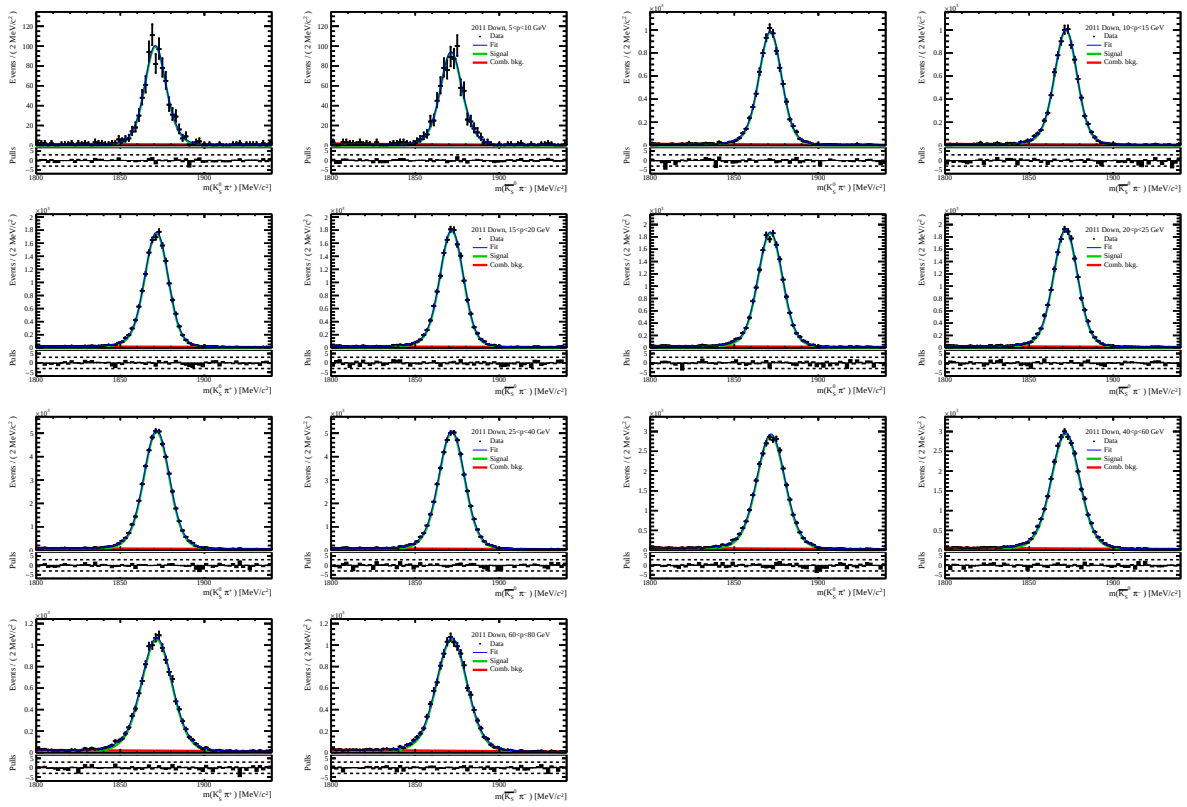

Figure 64: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2011 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

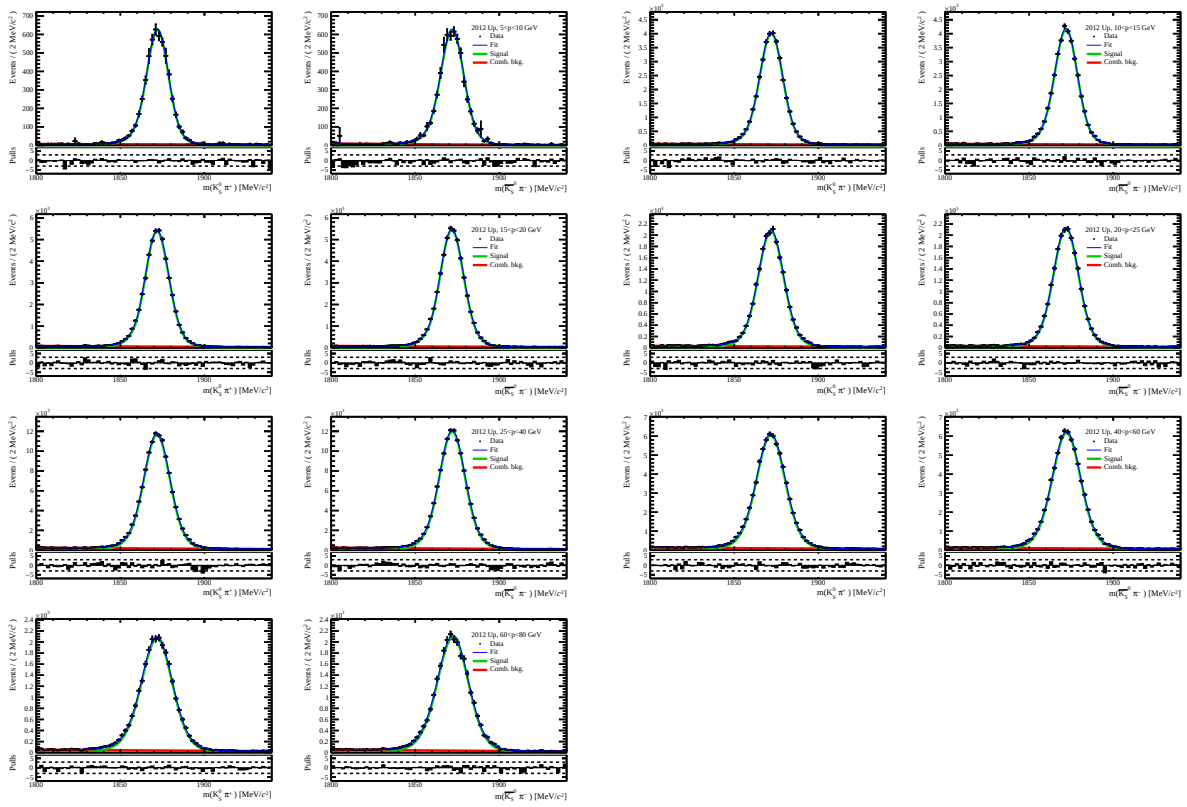

Figure 65: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2012 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

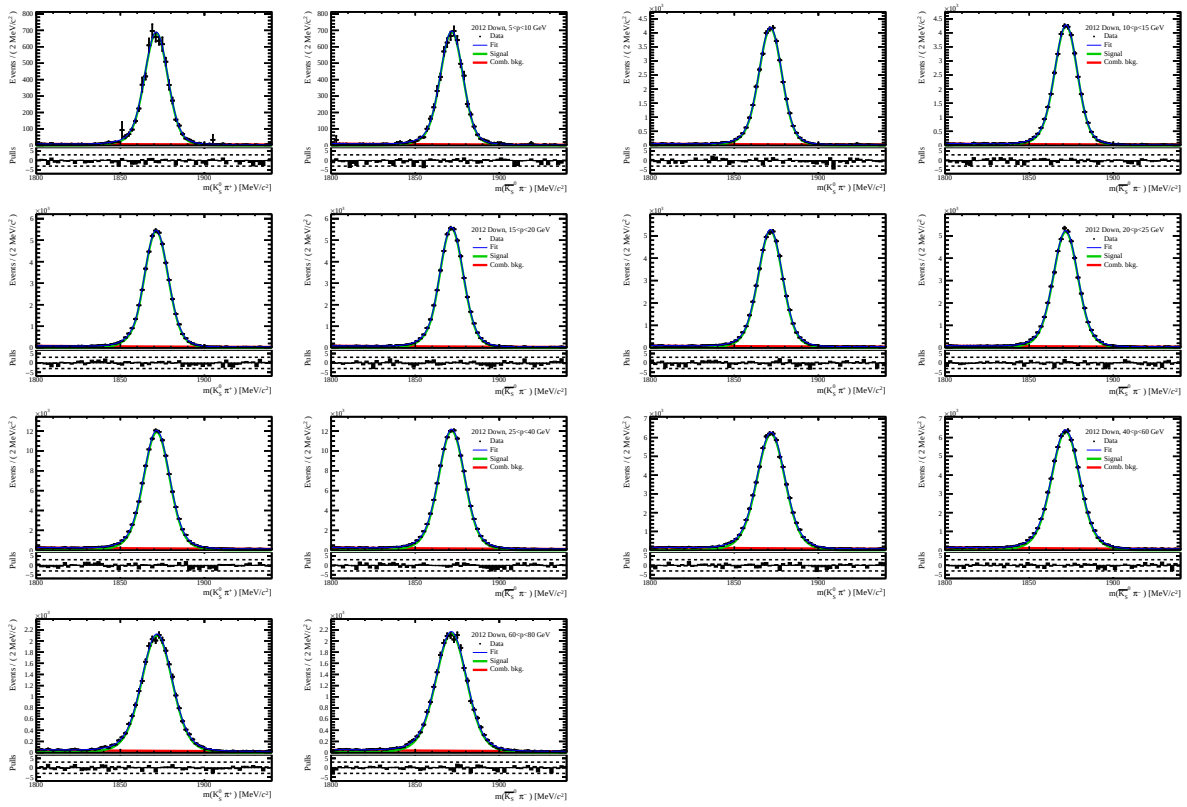

Figure 66: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2012 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

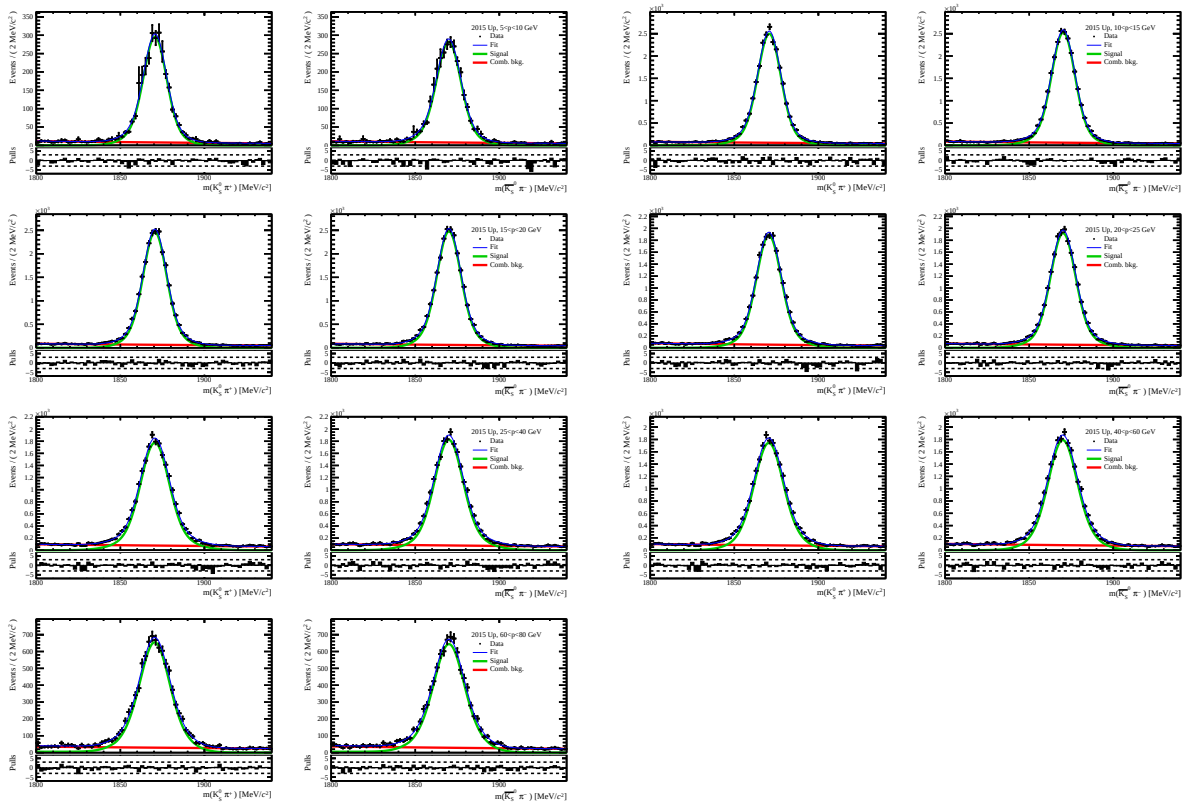

Figure 67: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2015 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

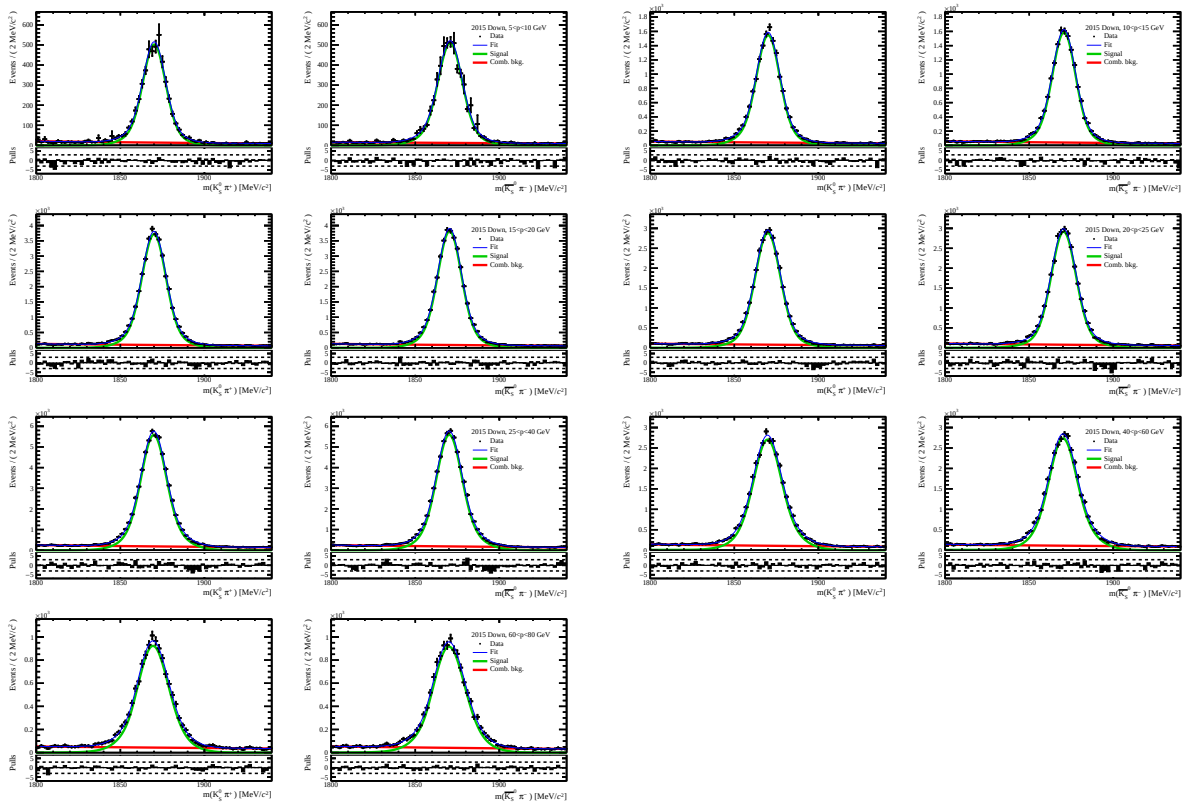

Figure 68: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2015 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

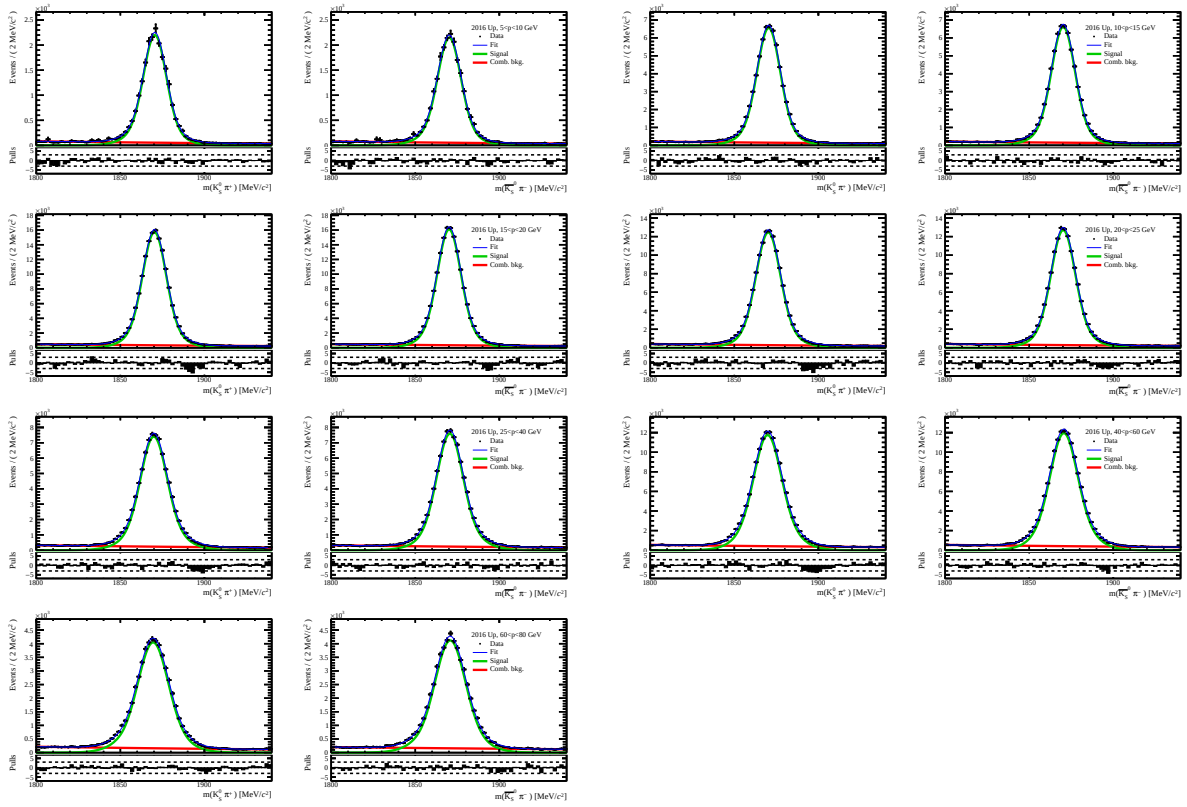

Figure 69: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2016 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

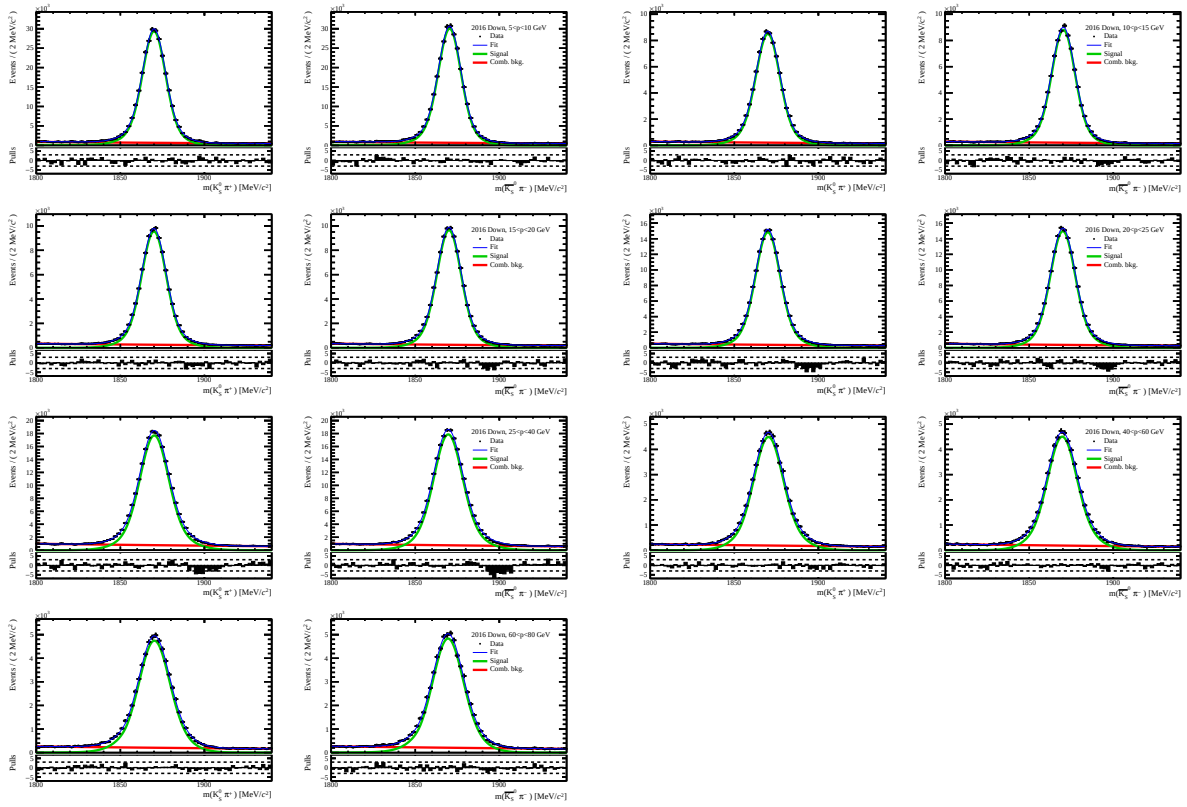

Figure 70: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2016 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

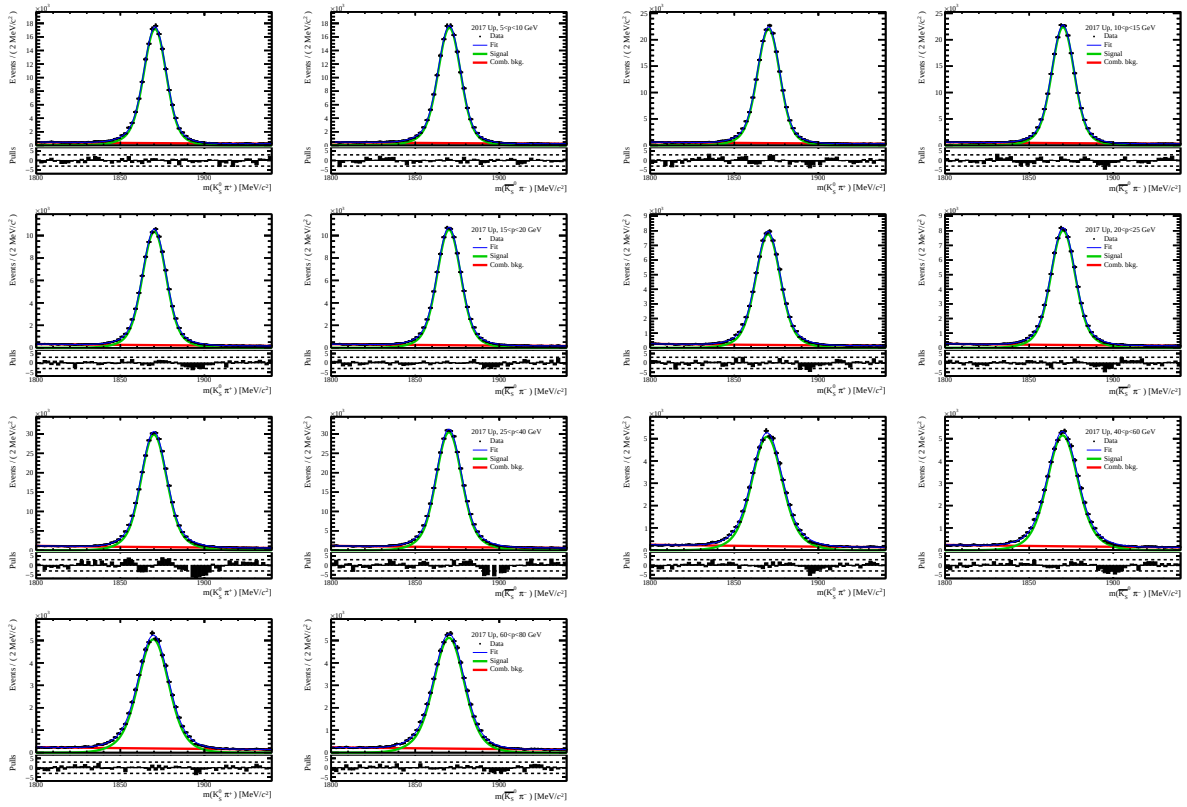

Figure 71: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2017 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

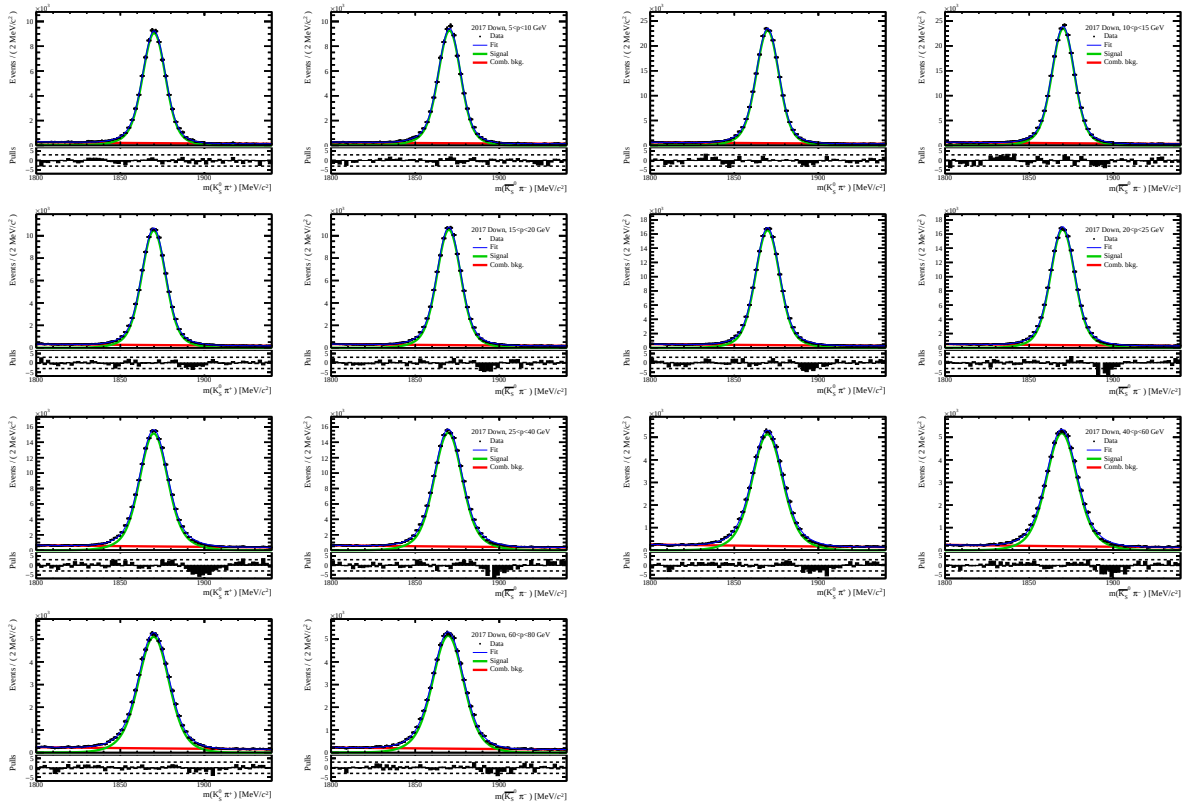

Figure 72: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2017 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

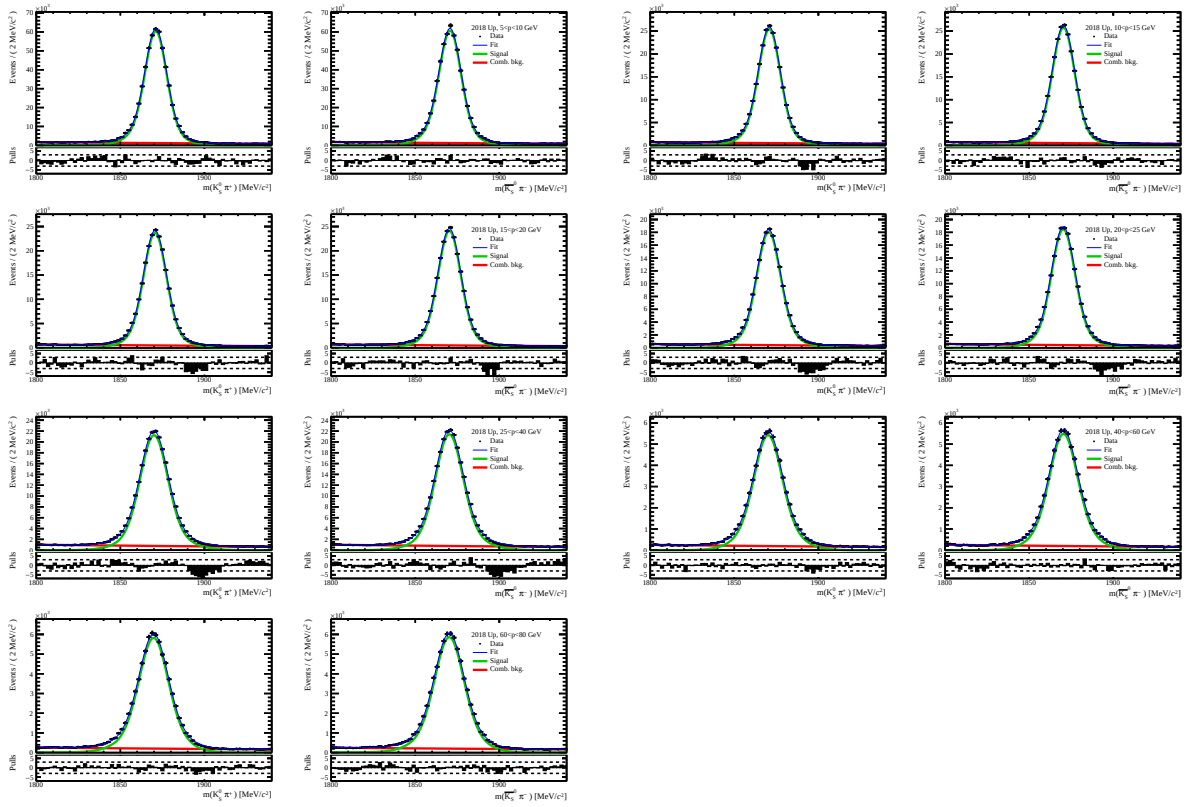

Figure 73: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2018 and magnet polarity up sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

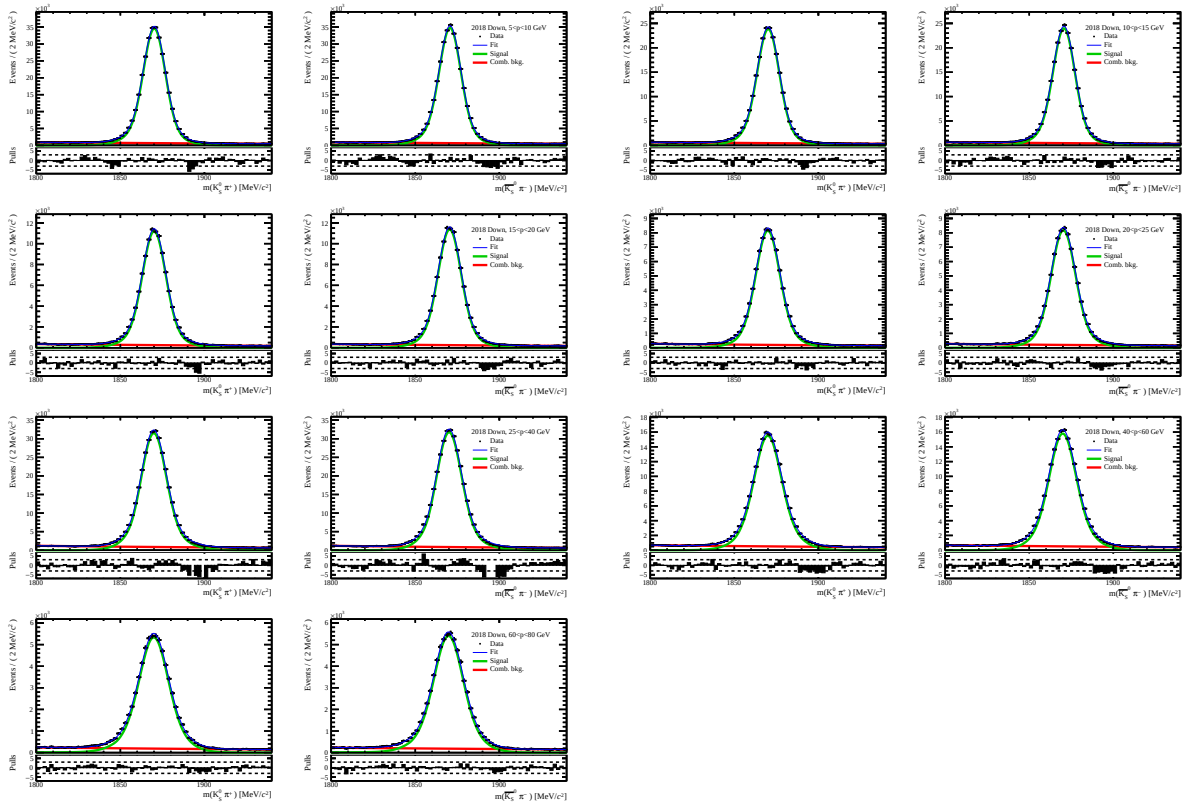

Figure 74: Fits to the invariant-mass distributions of (left of the pair) $K^0_S\pi^+$ and (right of the pair) $\overline{K^0_S}\pi^-$ decays in the year 2018 and magnet polarity down sample. The samples are divided in bins of kaon momentum, from lowest to highest from the top to the bottom and from left to right.

Appendix A

Fits for the TIS asymmetry

`TIS & TOS` trigger category

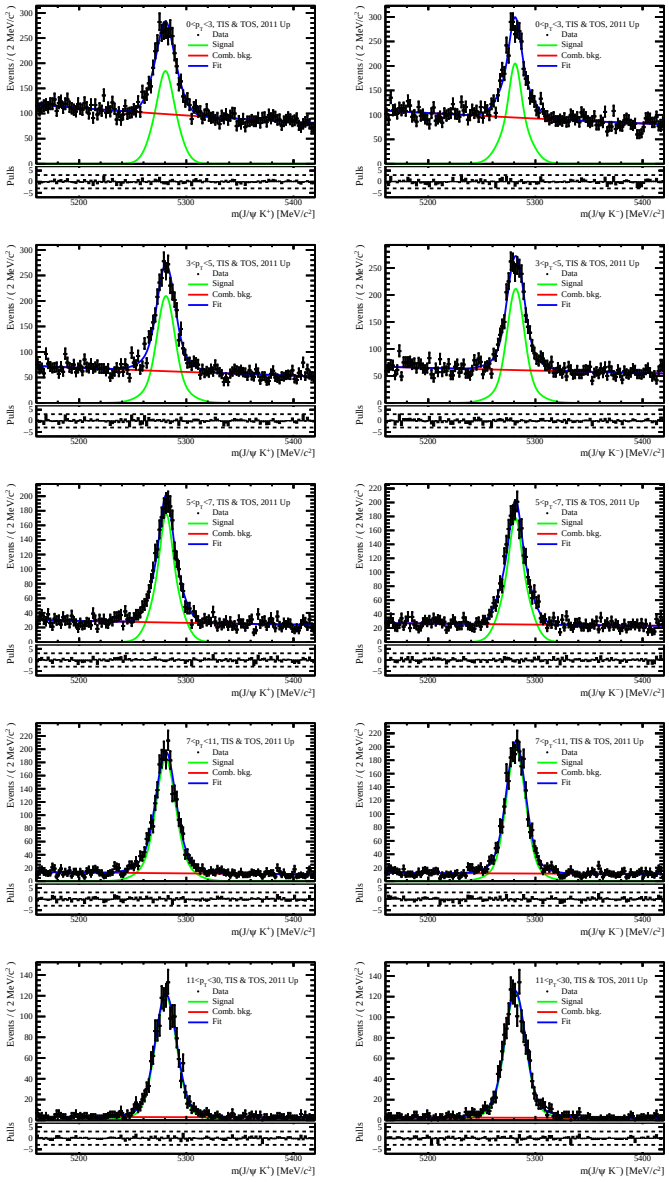

Figure A.1: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2011, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

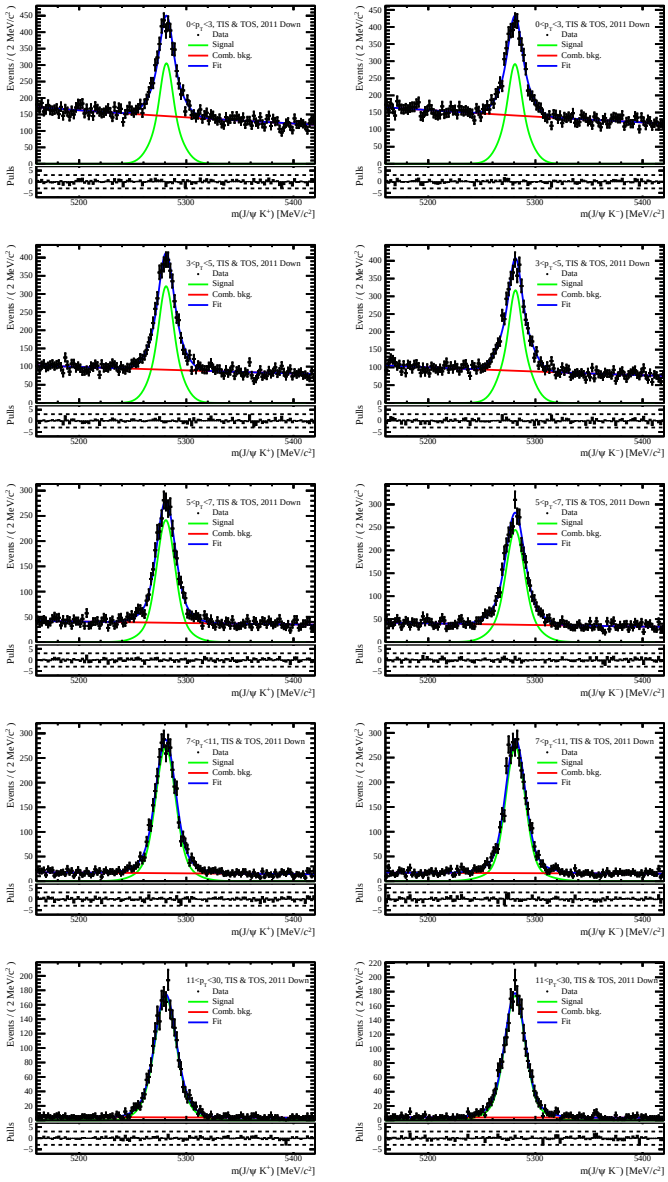

Figure A.2: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2011, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

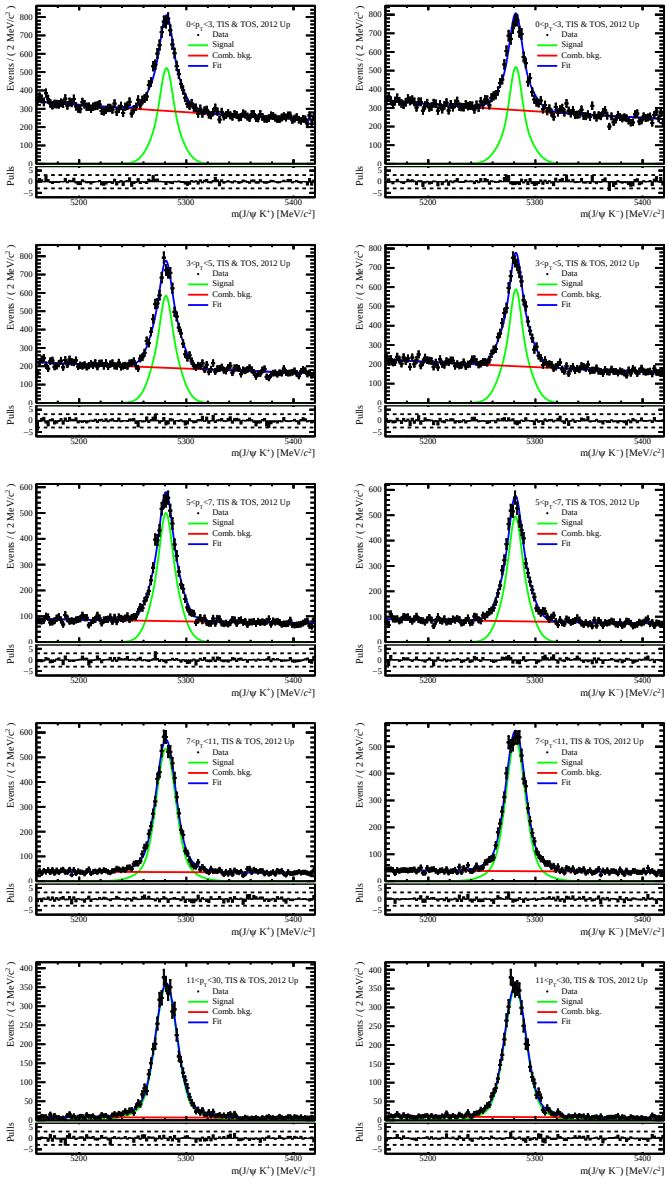

Figure A.3: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2012, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

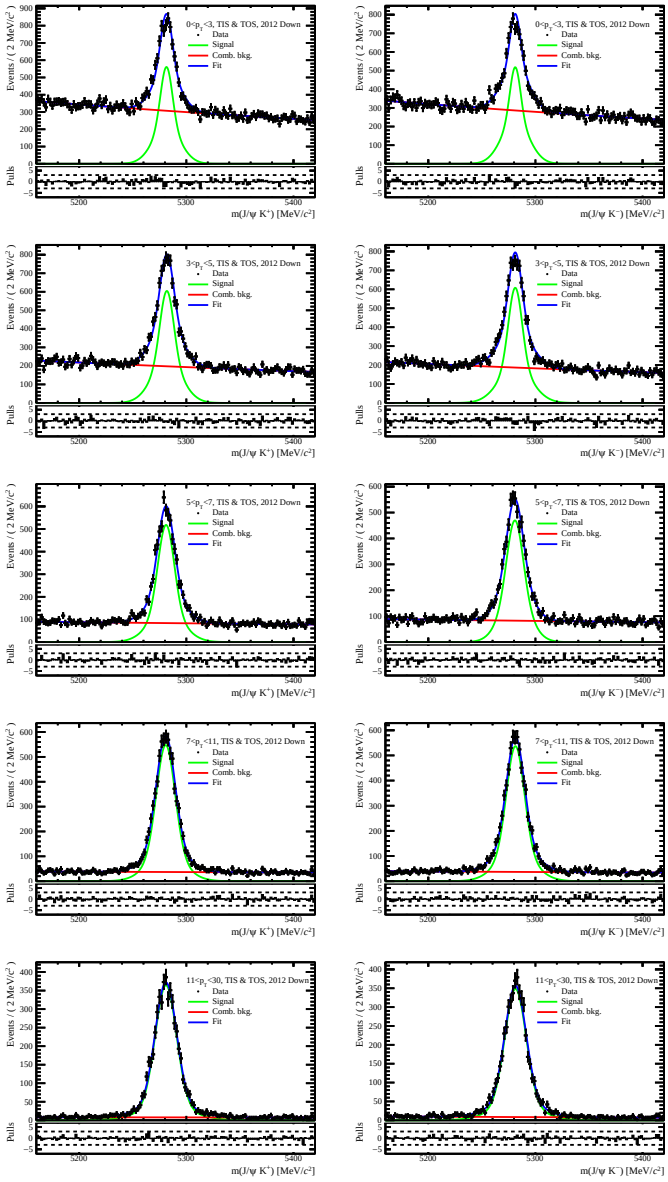

Figure A.4: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2012, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

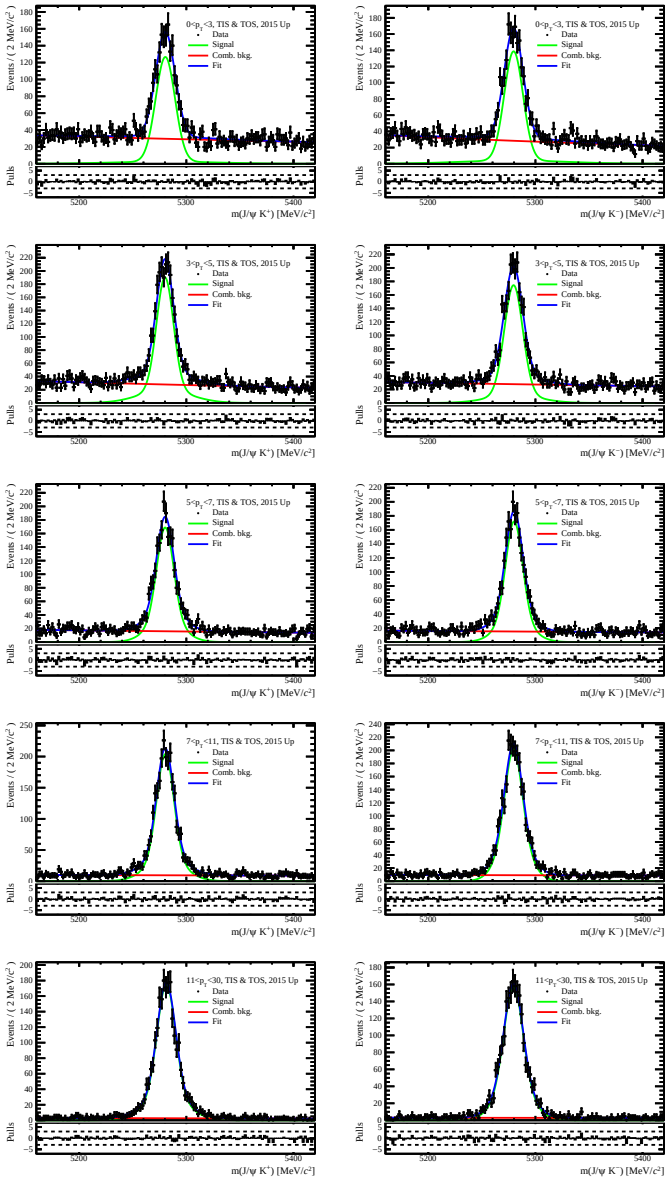

Figure A.5: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2015, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

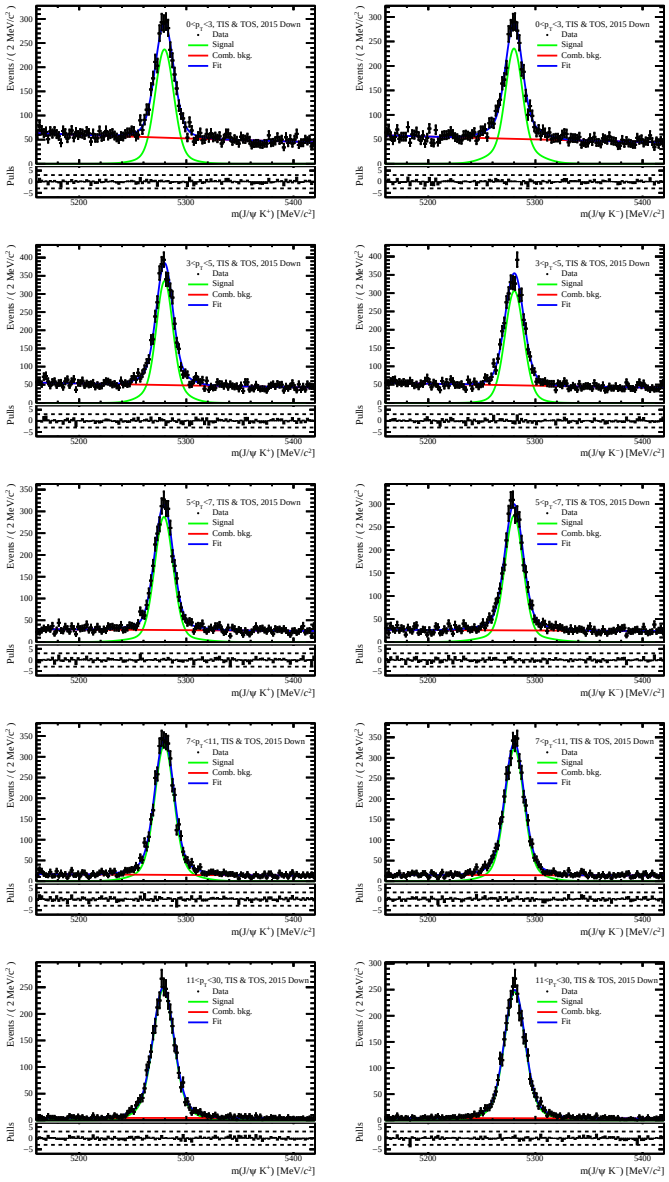

Figure A.6: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2015, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

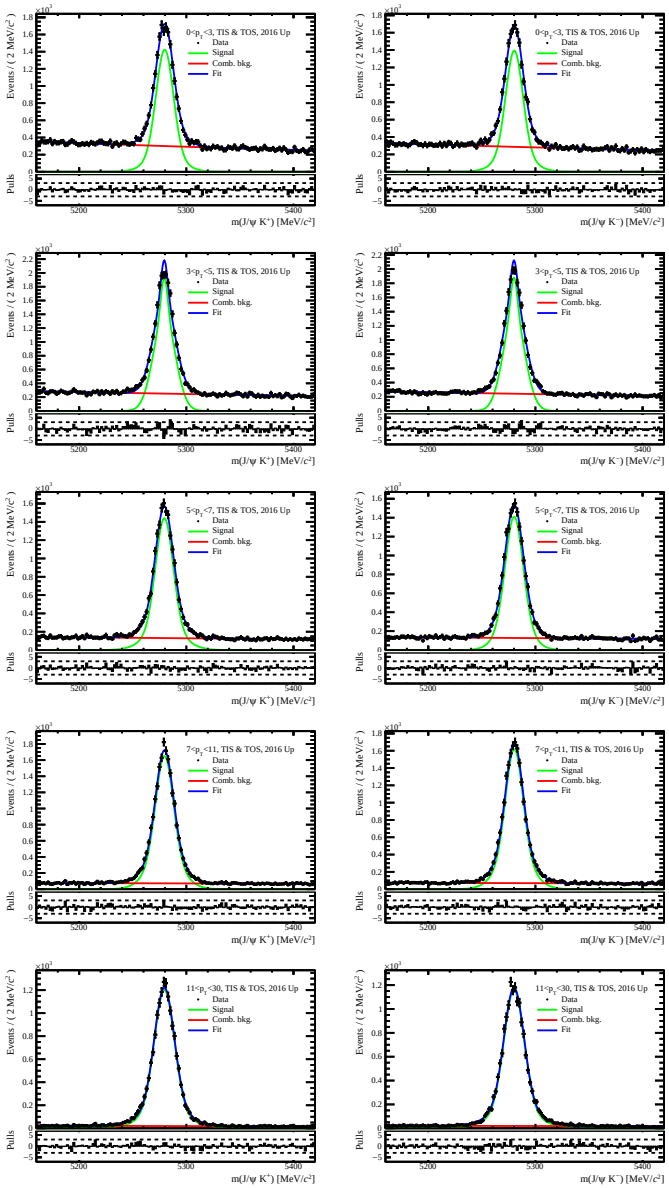

Figure A.7: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2016, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

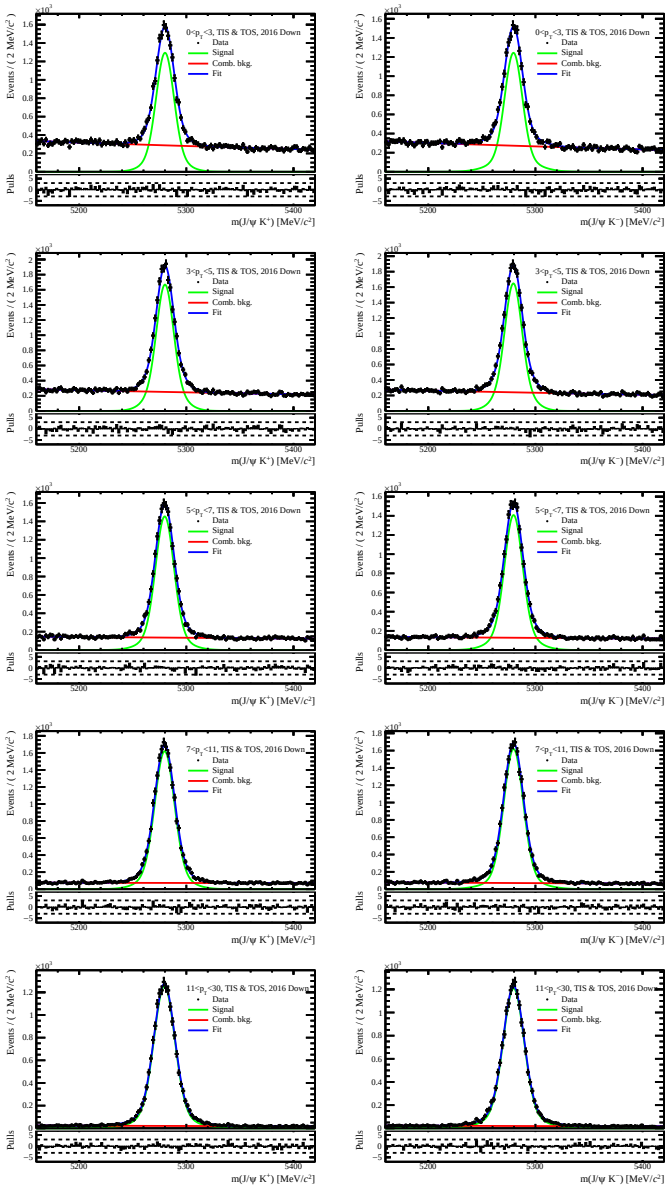

Figure A.8: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2016, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

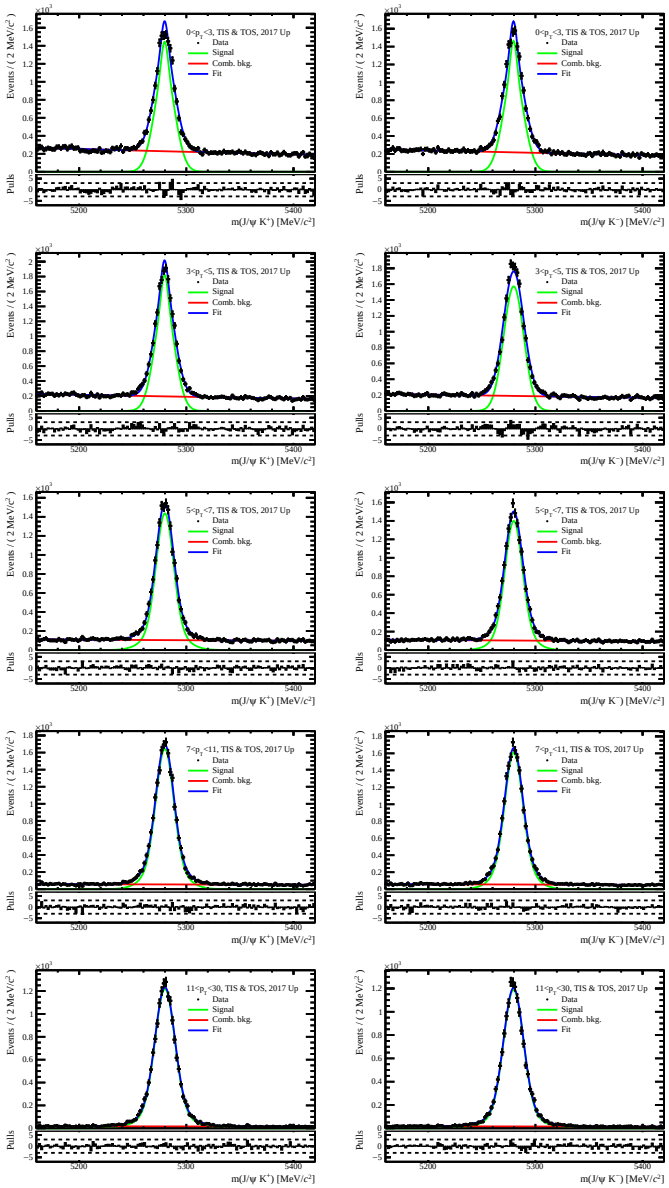

Figure A.9: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2017, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

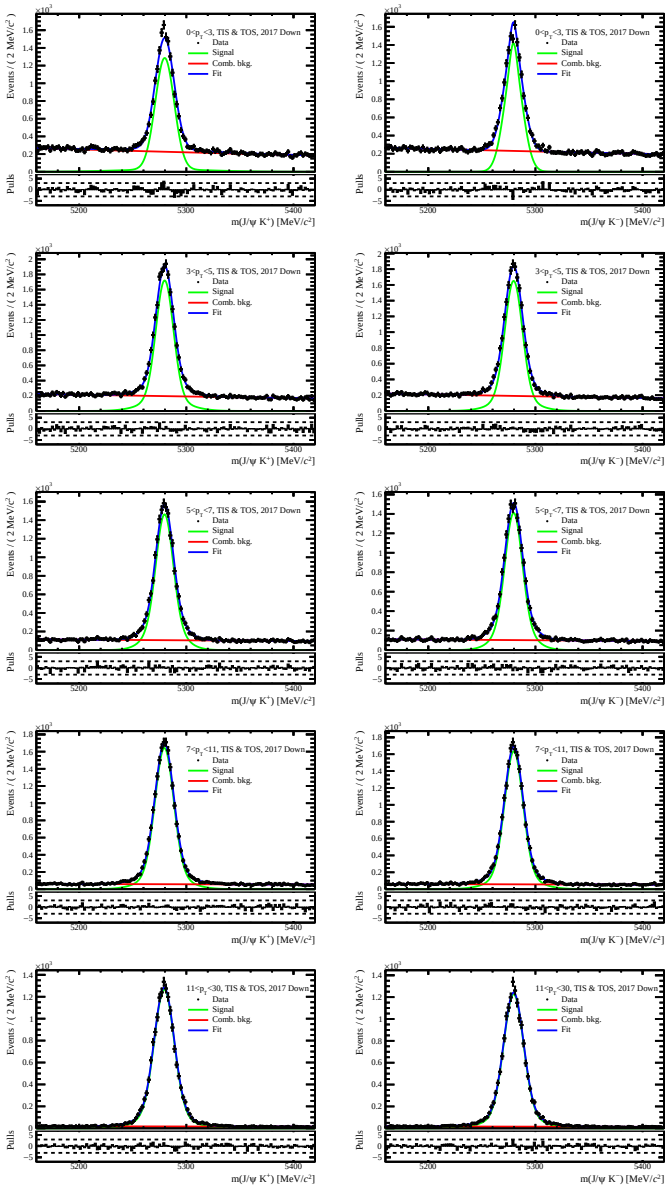

Figure A.10: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2017, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

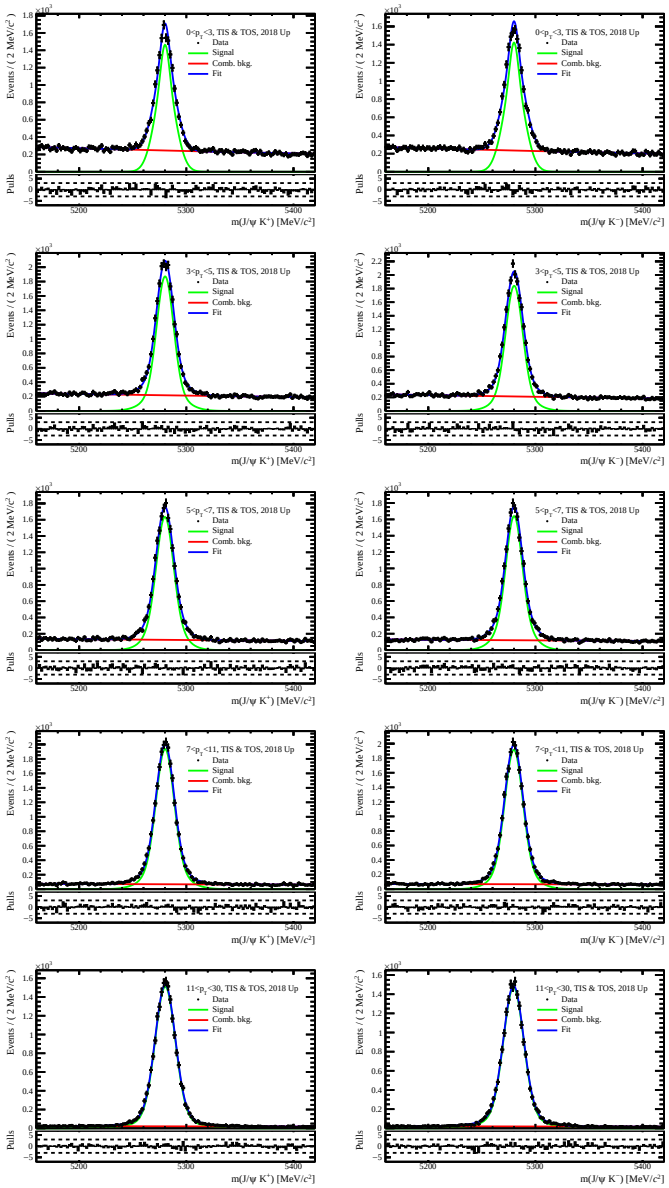

Figure A.11: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2018, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

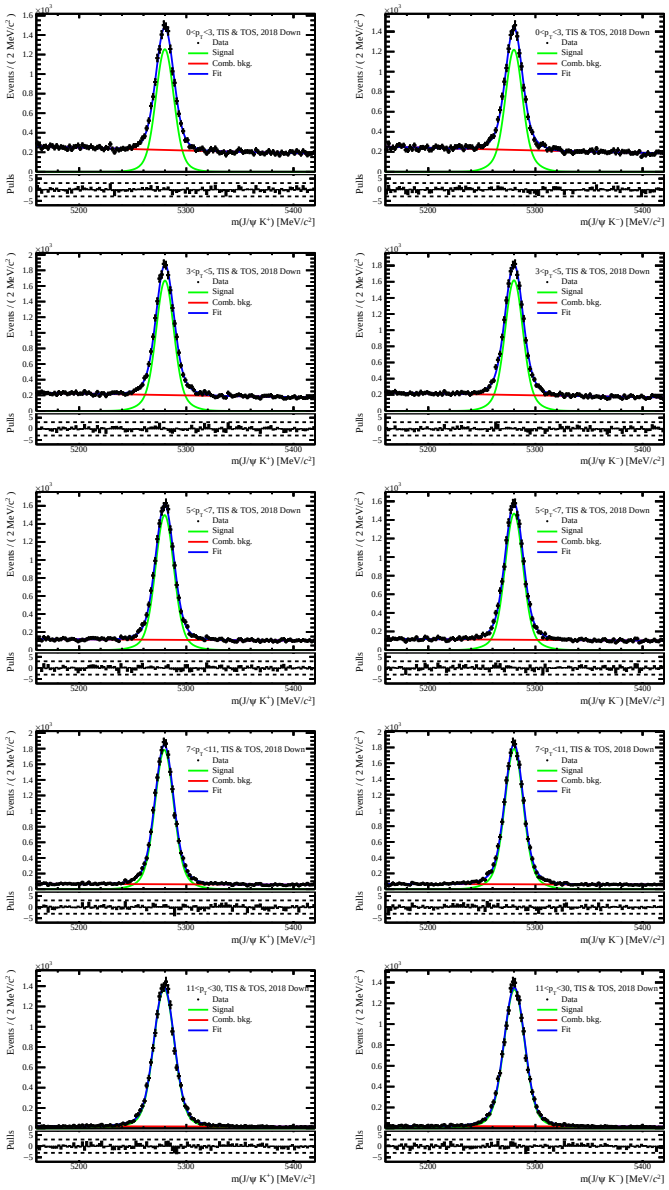

Figure A.12: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `TIS & TOS`, year 2018, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

!TIS & TOS trigger category

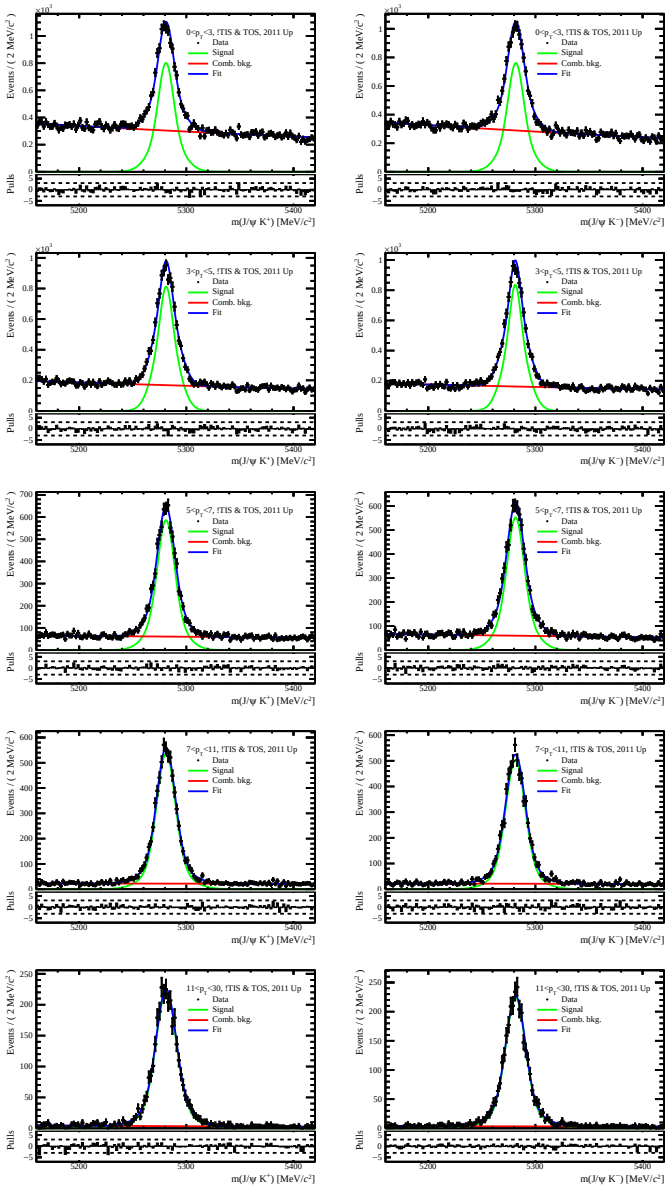

Figure A.13: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category !TIS & TOS, year 2011, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

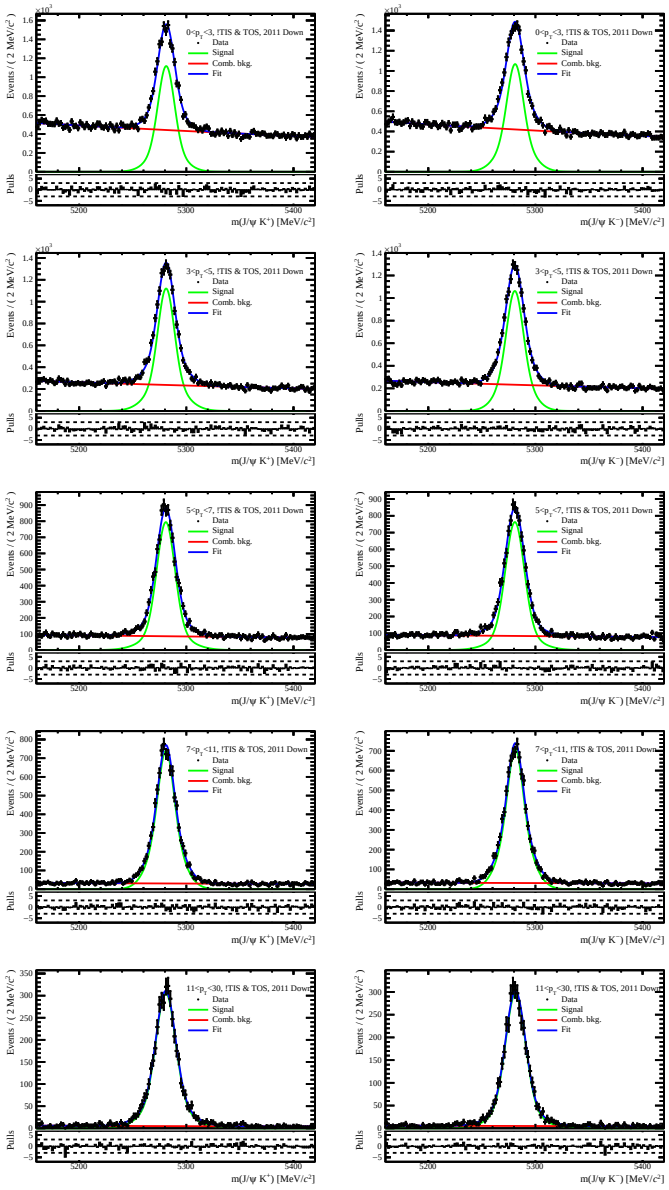

Figure A.14: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2011, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

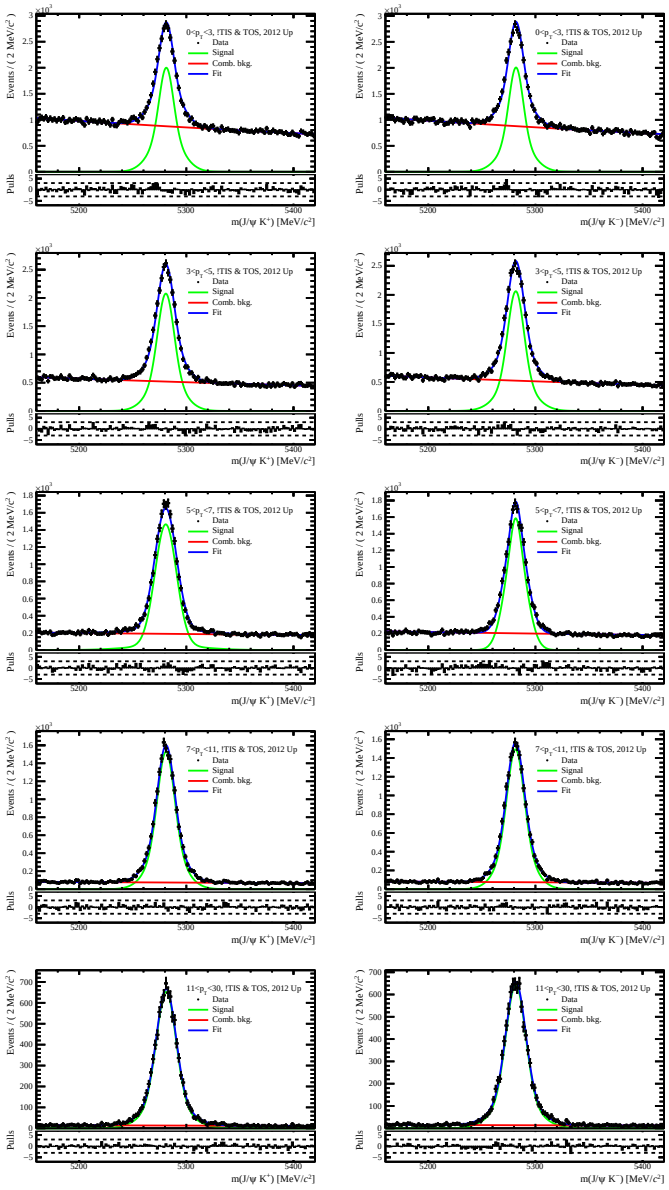

Figure A.15: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2012, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

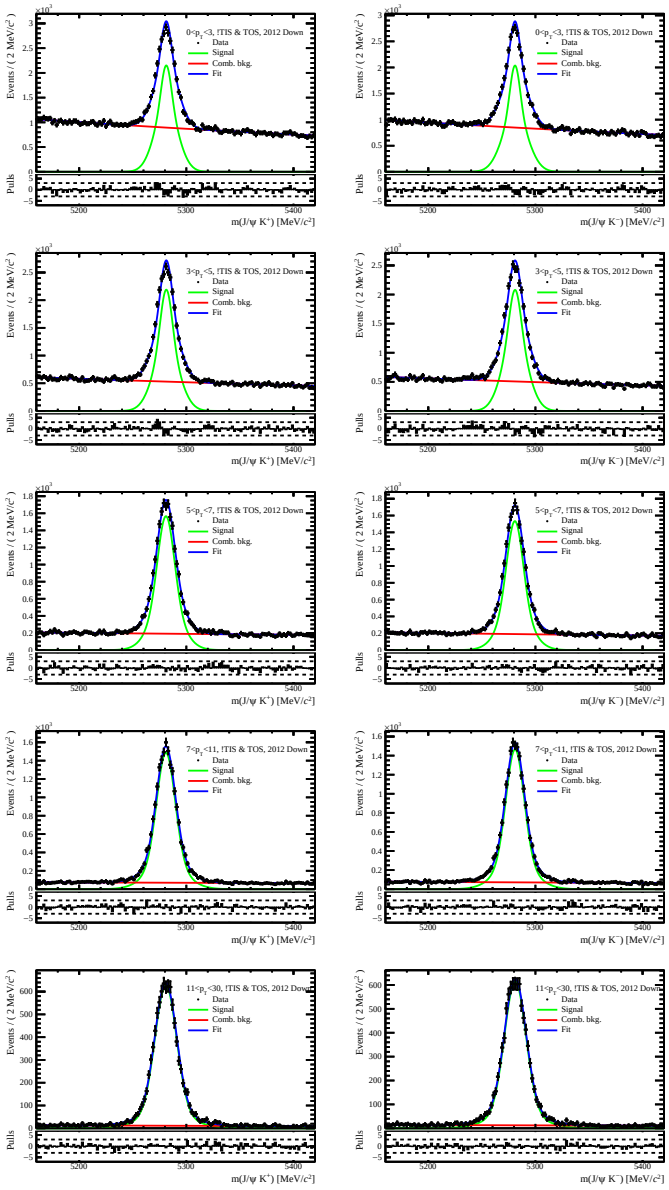

Figure A.16: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2012, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

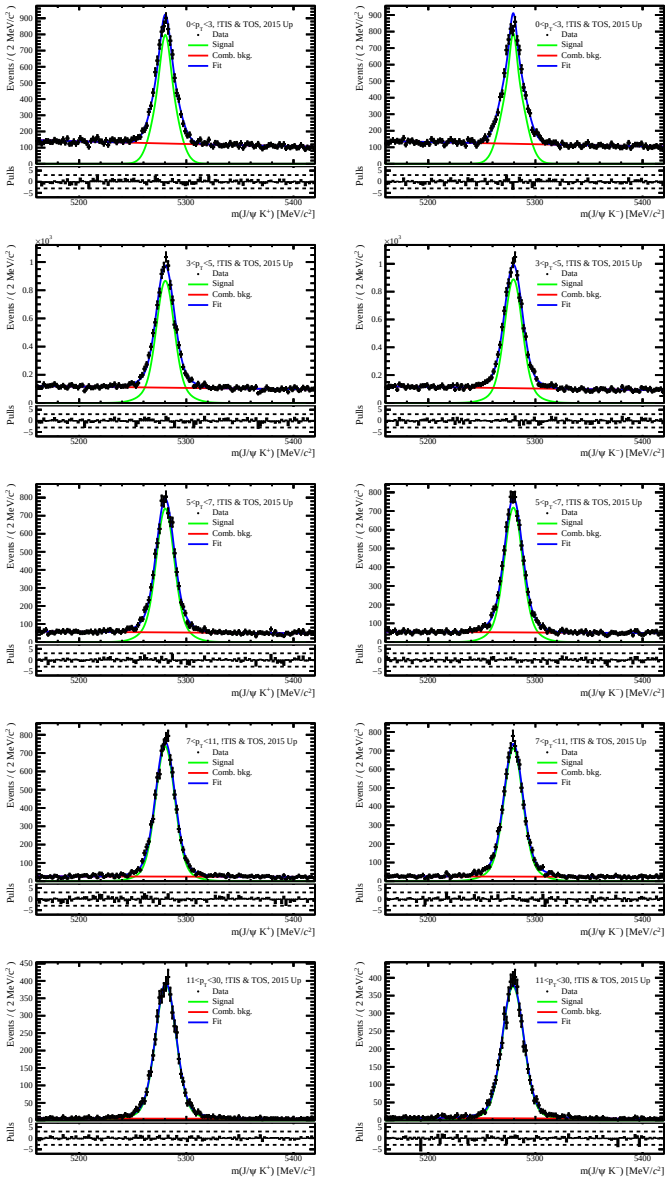

Figure A.17: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2015, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

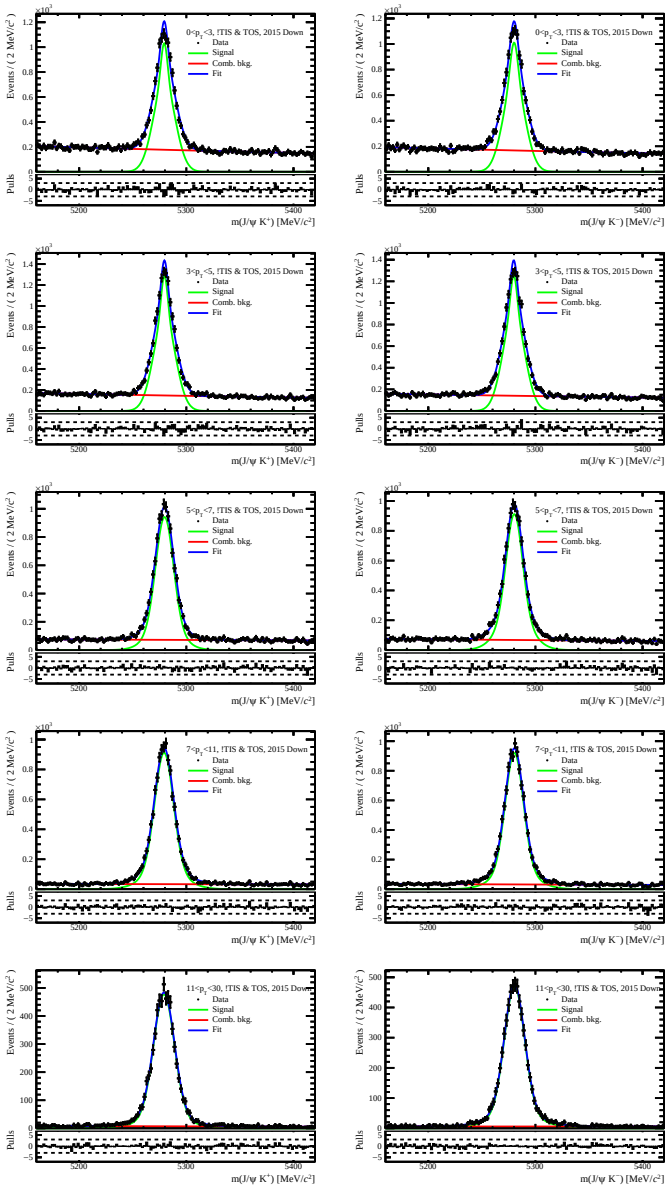

Figure A.18: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2015, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

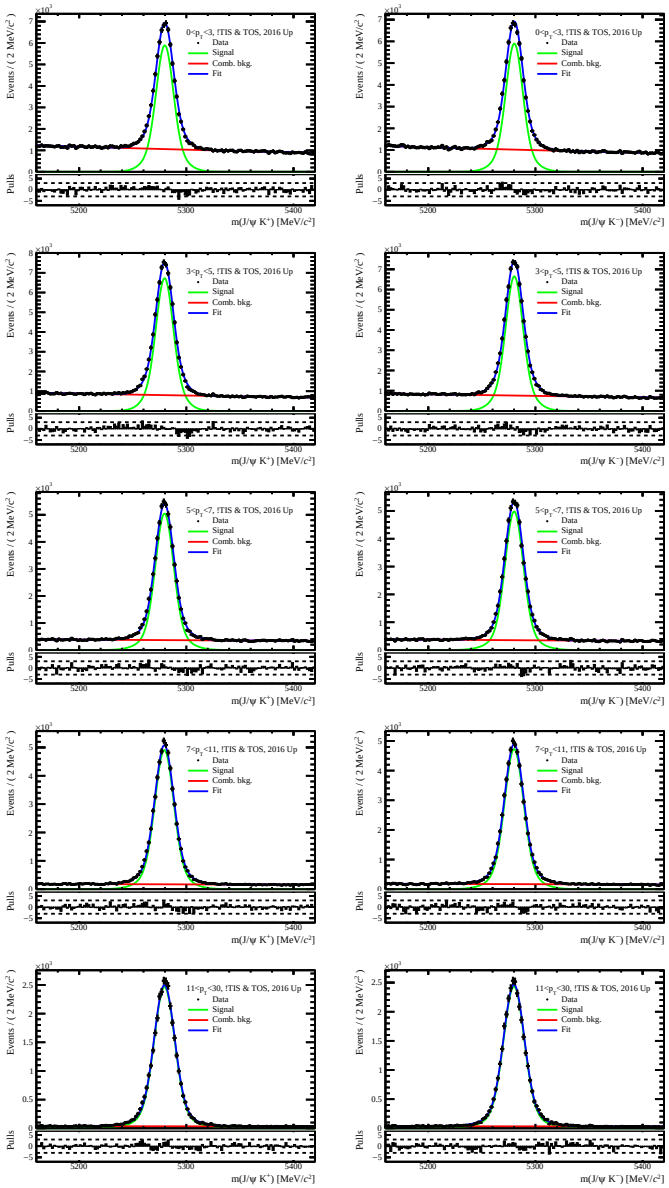

Figure A.19: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2016, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

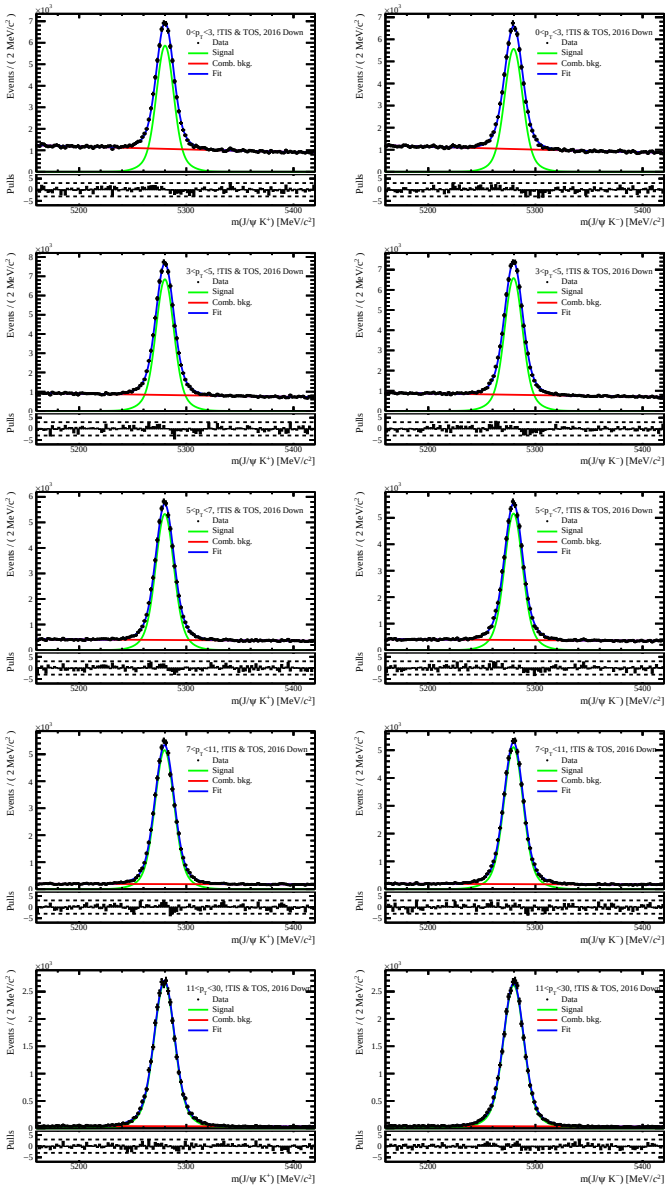

Figure A.20: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2016, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

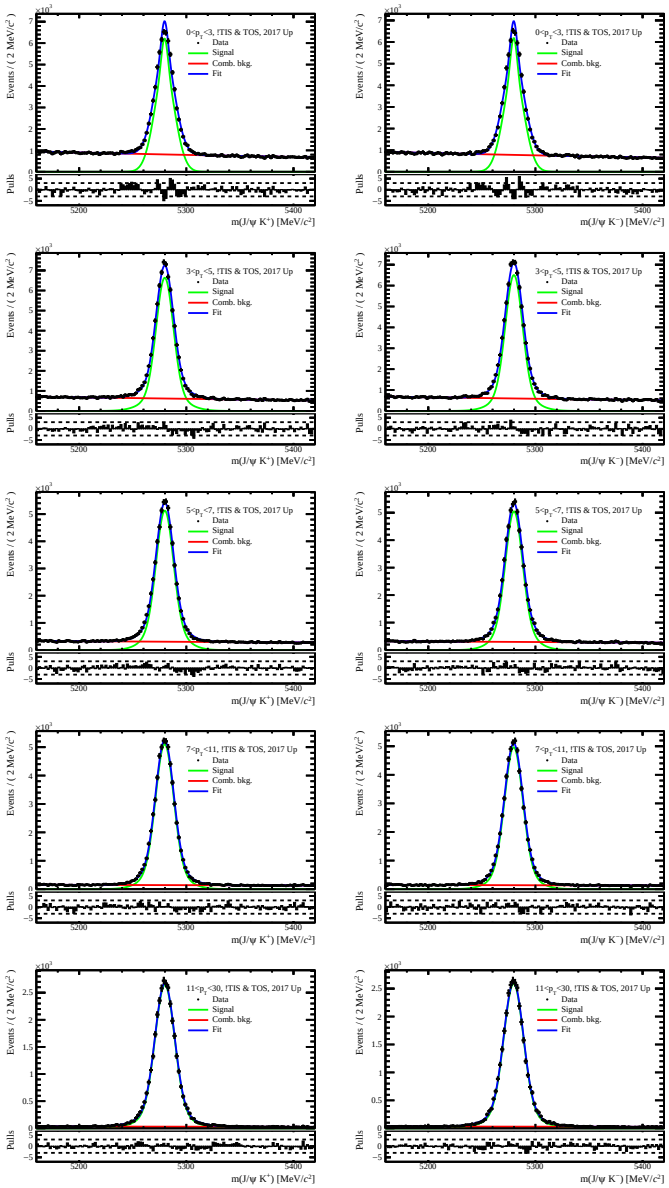

Figure A.21: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2017, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

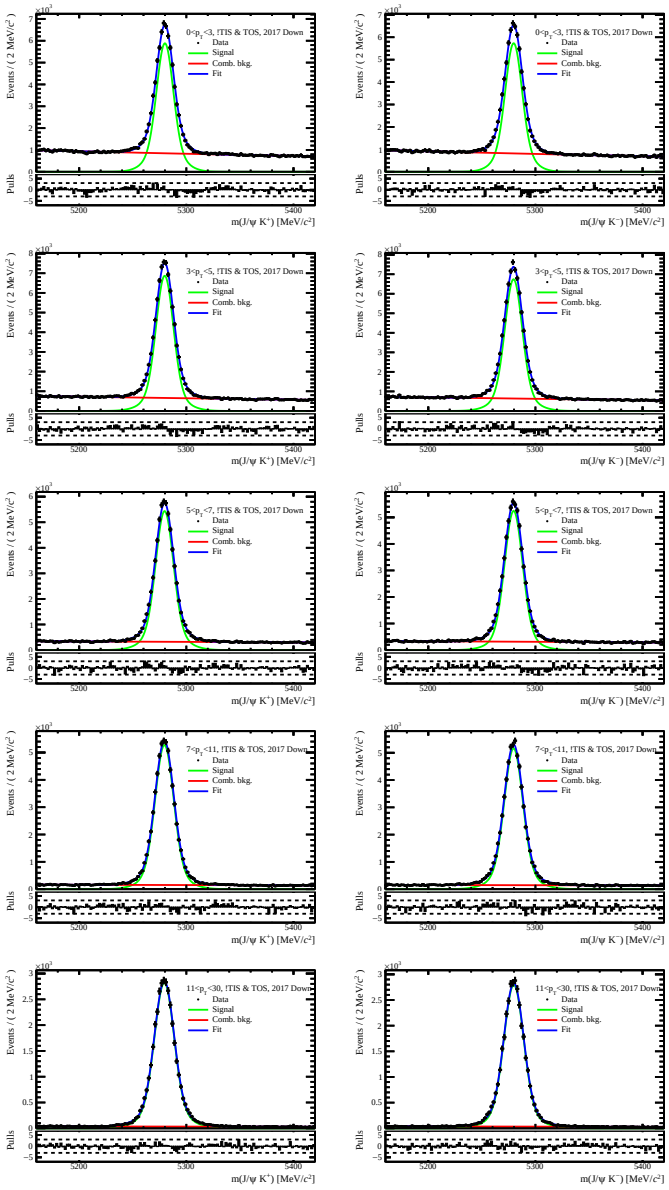

Figure A.22: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2017, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

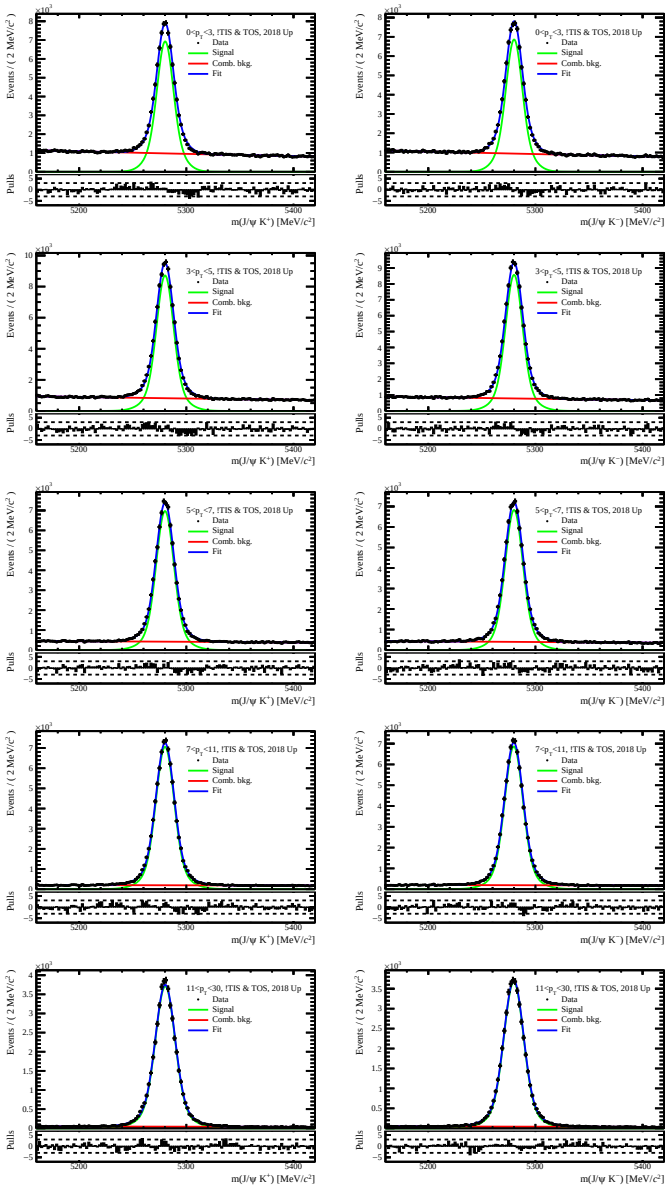

Figure A.23: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2018, and magnet polarity up. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

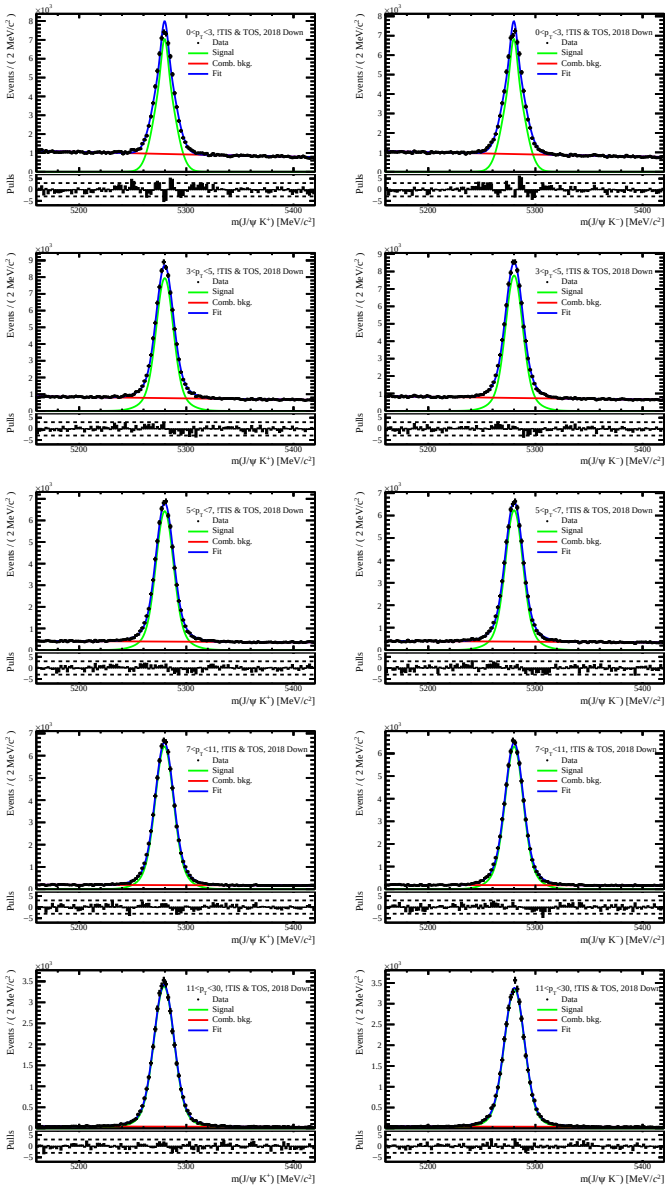

Figure A.24: Fits to the invariant-mass distributions of (left) $J/\psi K^+$ and (right) $J/\psi K^-$ decays in the trigger category `!TIS & TOS`, year 2018, and magnet polarity down. The samples are divided in bins of B^+ transverse momentum, from lowest to highest from the top to the bottom.

Appendix B

L0 E_{T} efficiencies

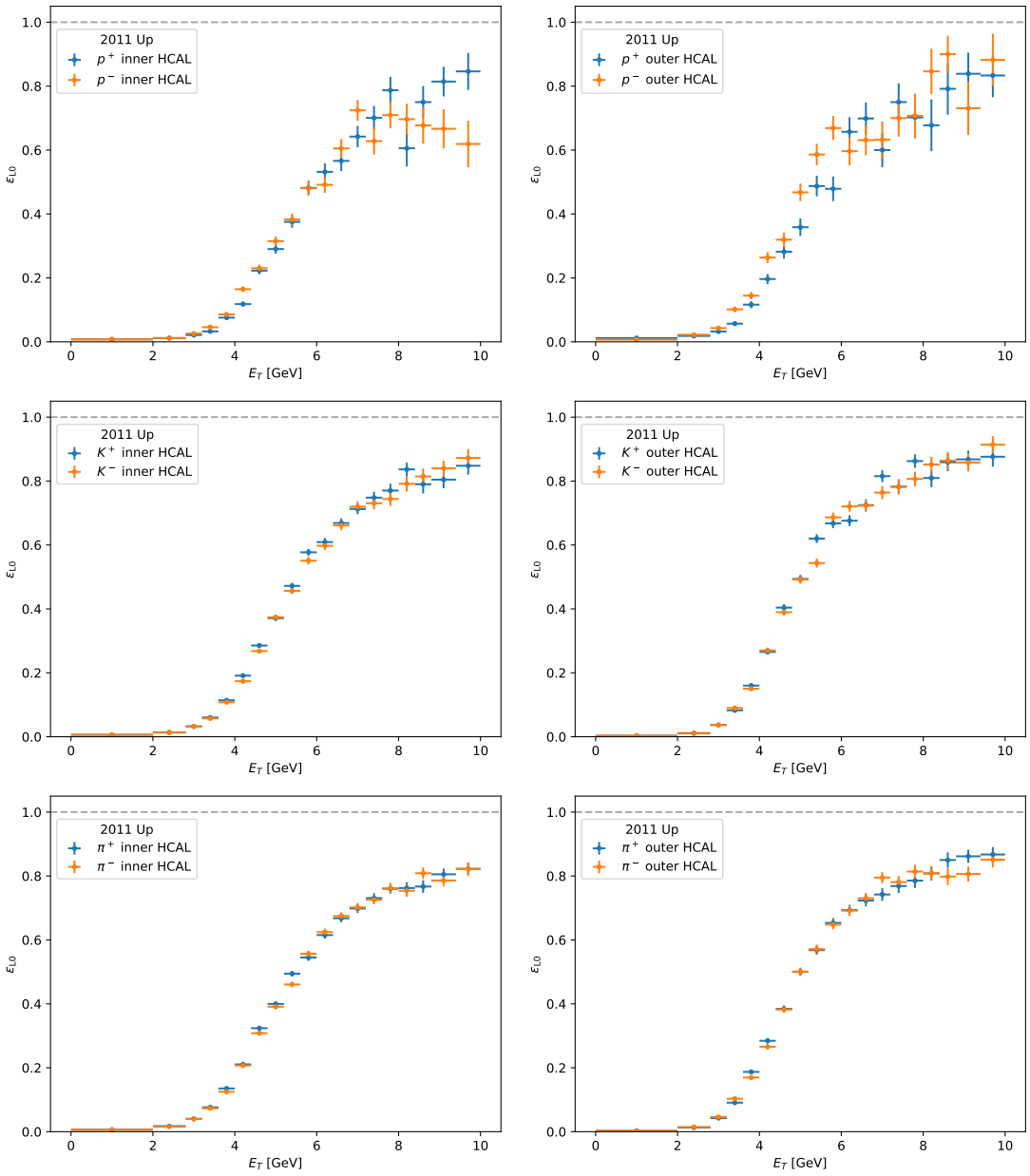

Figure B.1: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2011 magnet Up sample.

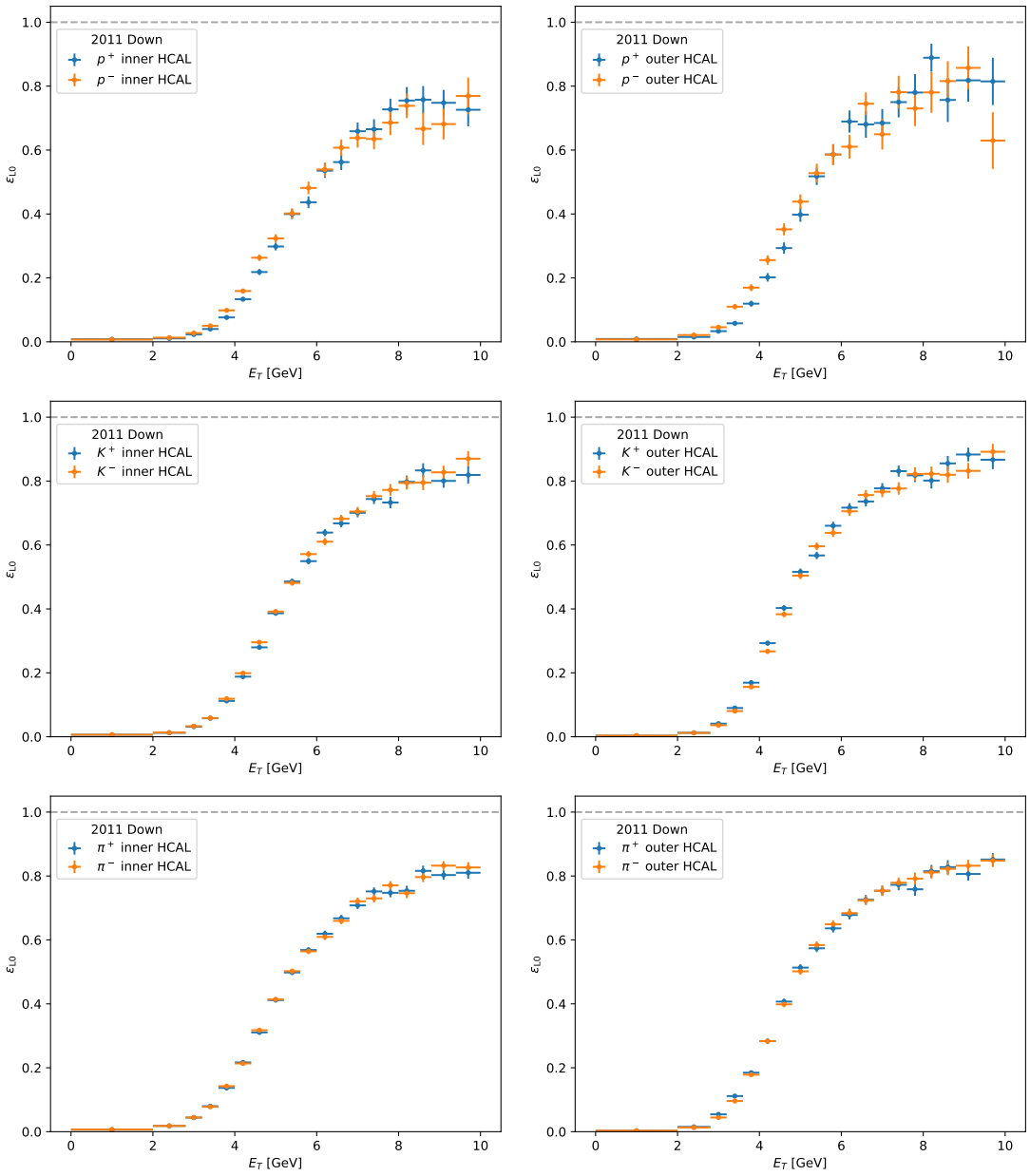

Figure B.2: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2011 magnet Down sample.

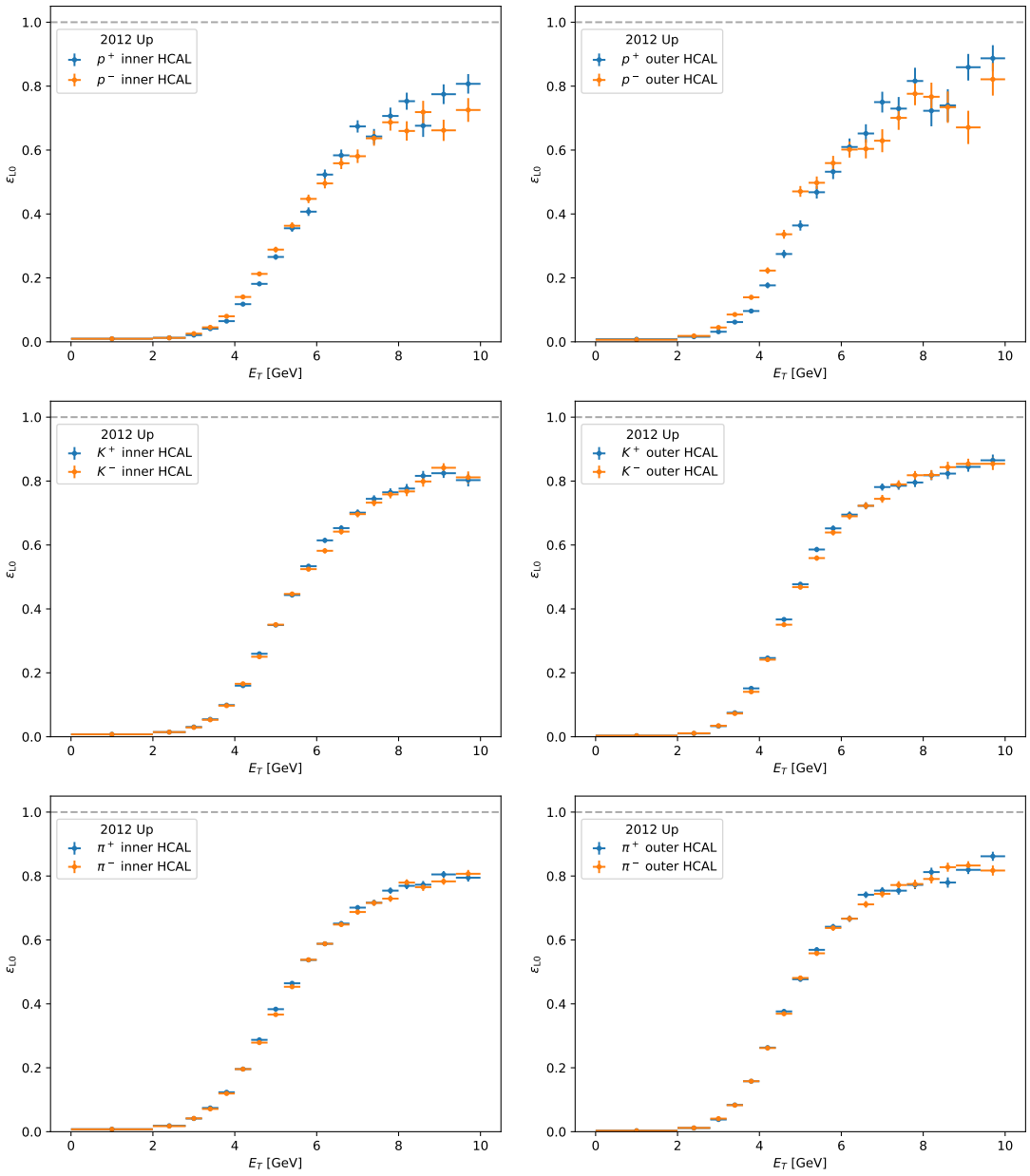

Figure B.3: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2012 magnet Up sample.

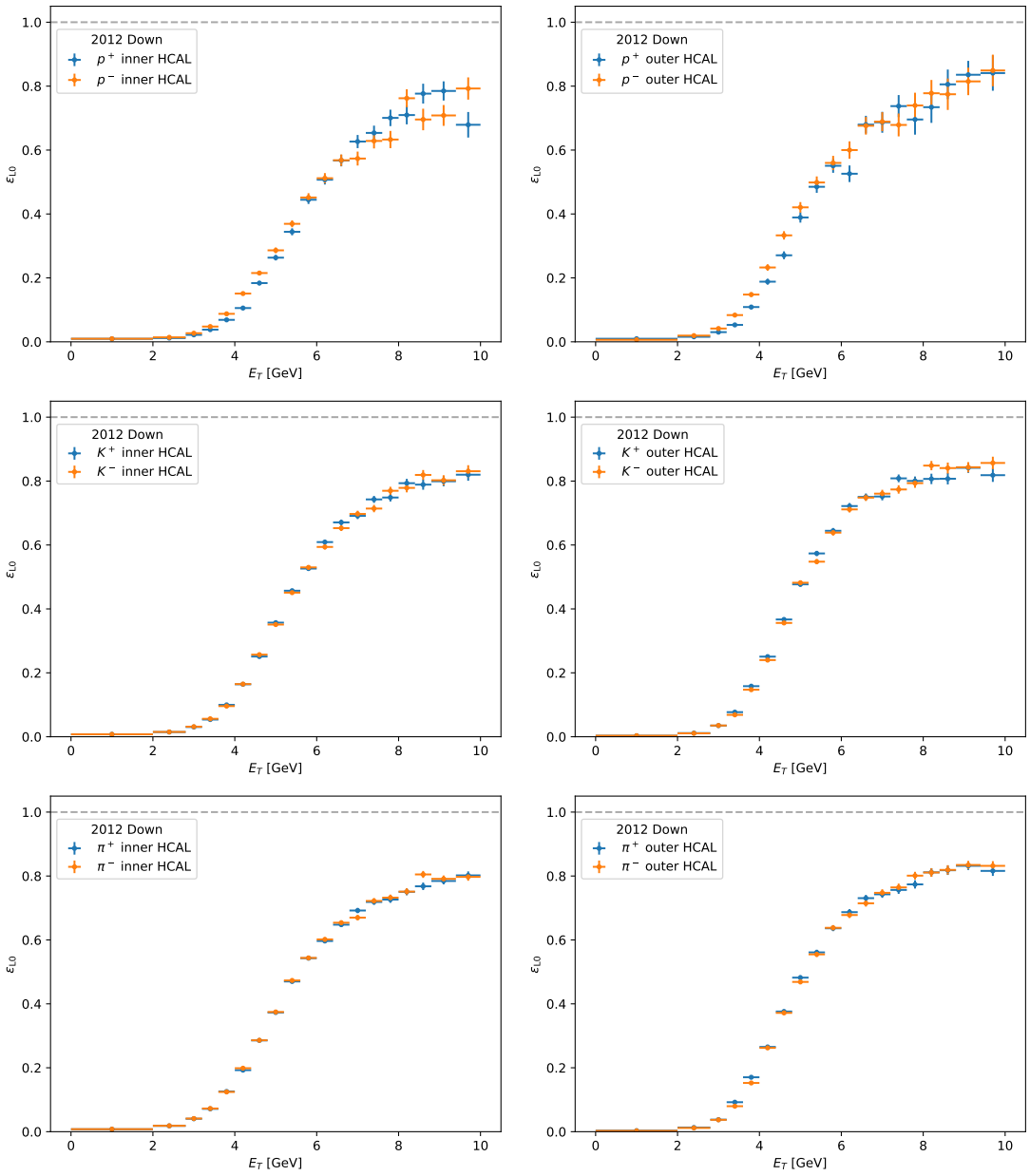

Figure B.4: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2012 magnet Down sample.

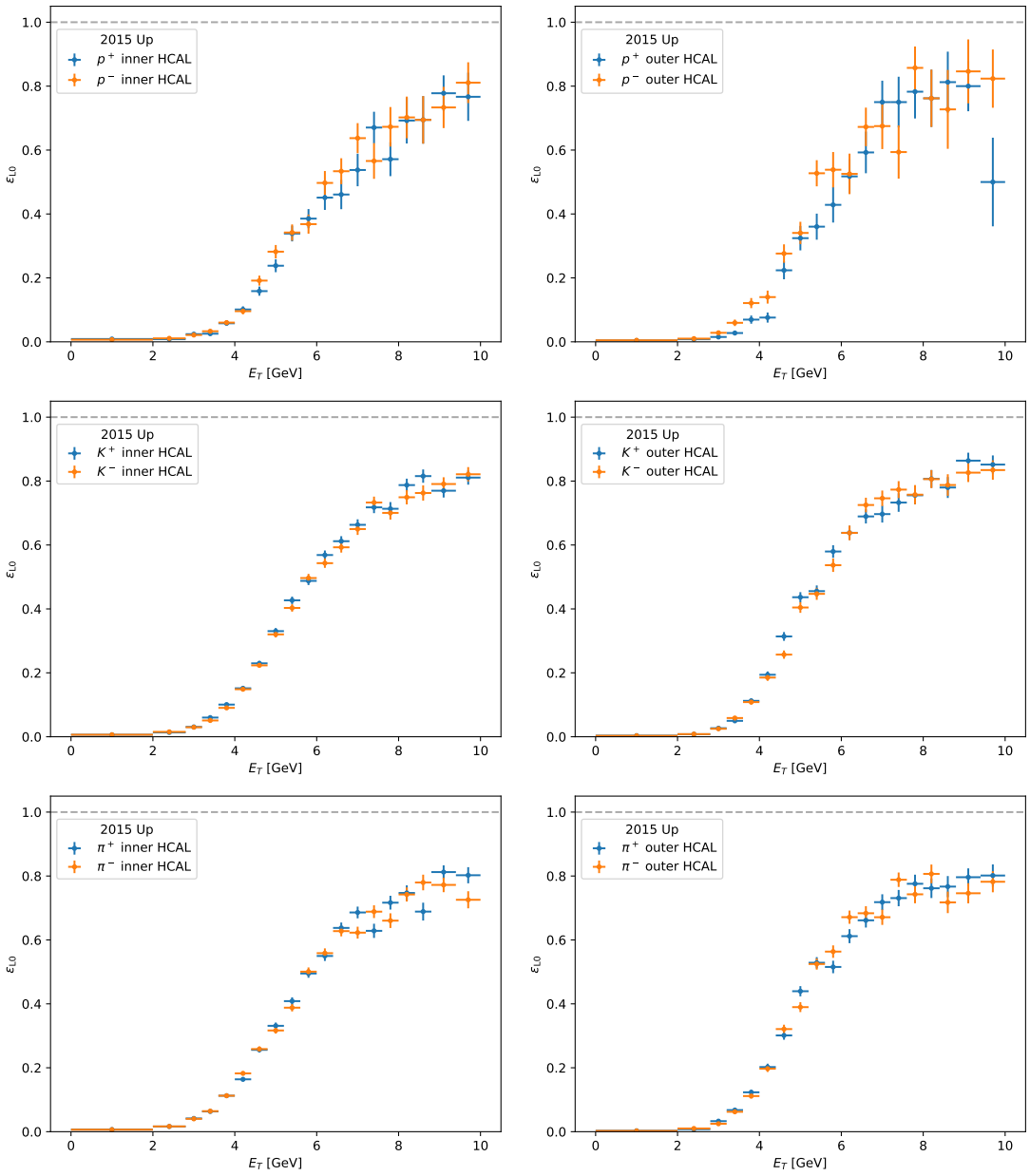

Figure B.5: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2015 magnet Up sample.

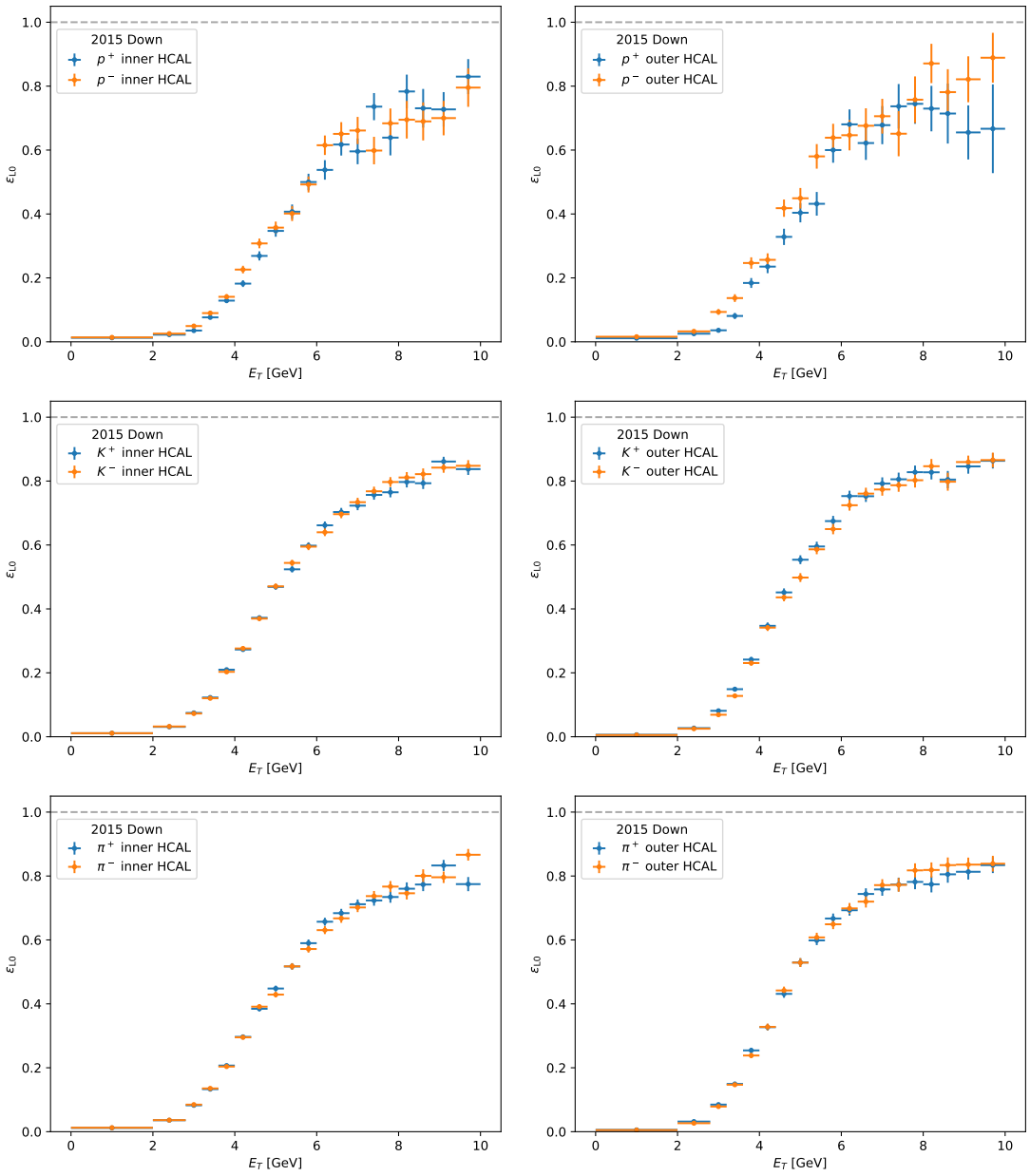

Figure B.6: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2015 magnet Down sample.

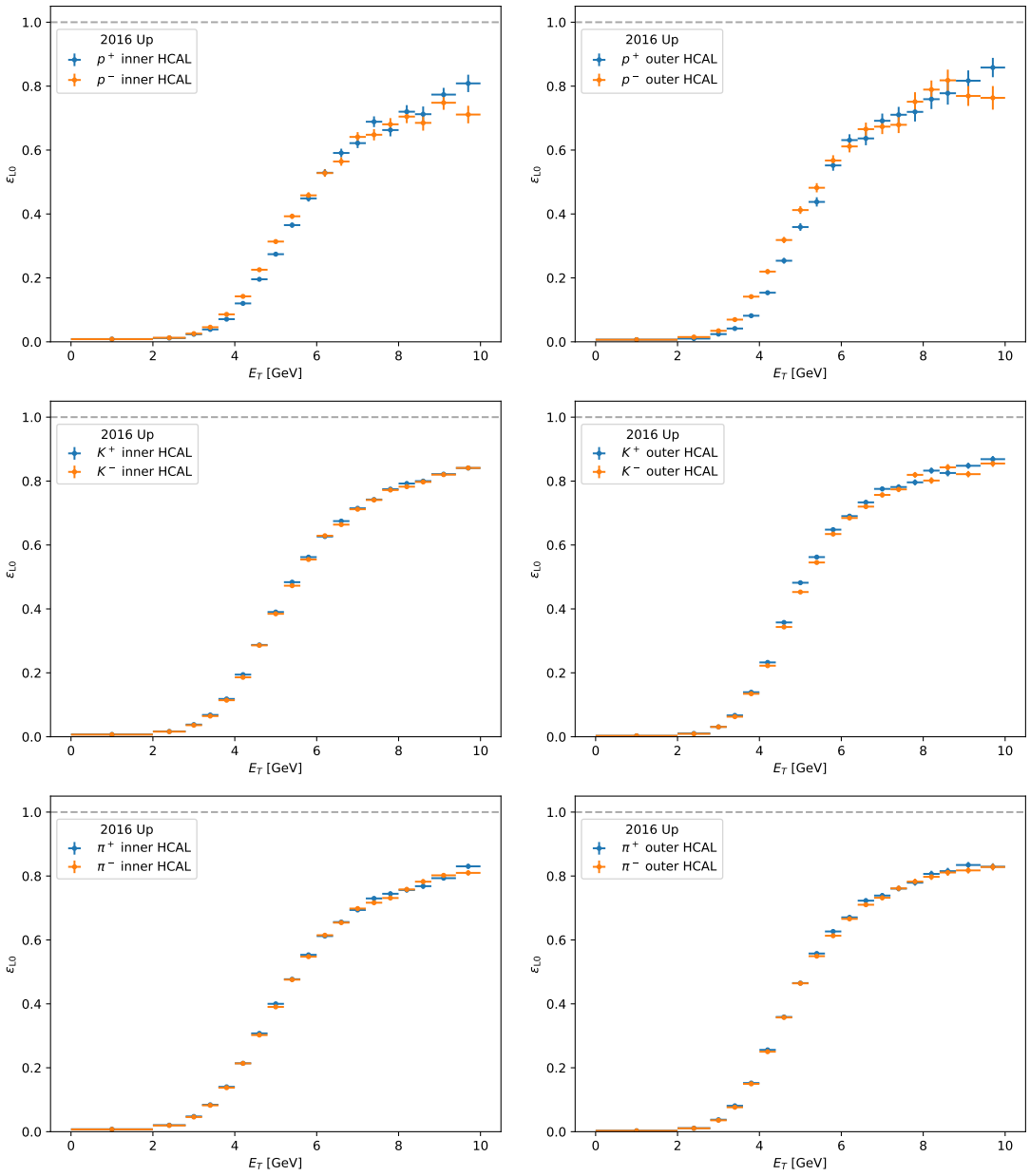

Figure B.7: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2016 magnet Up sample.

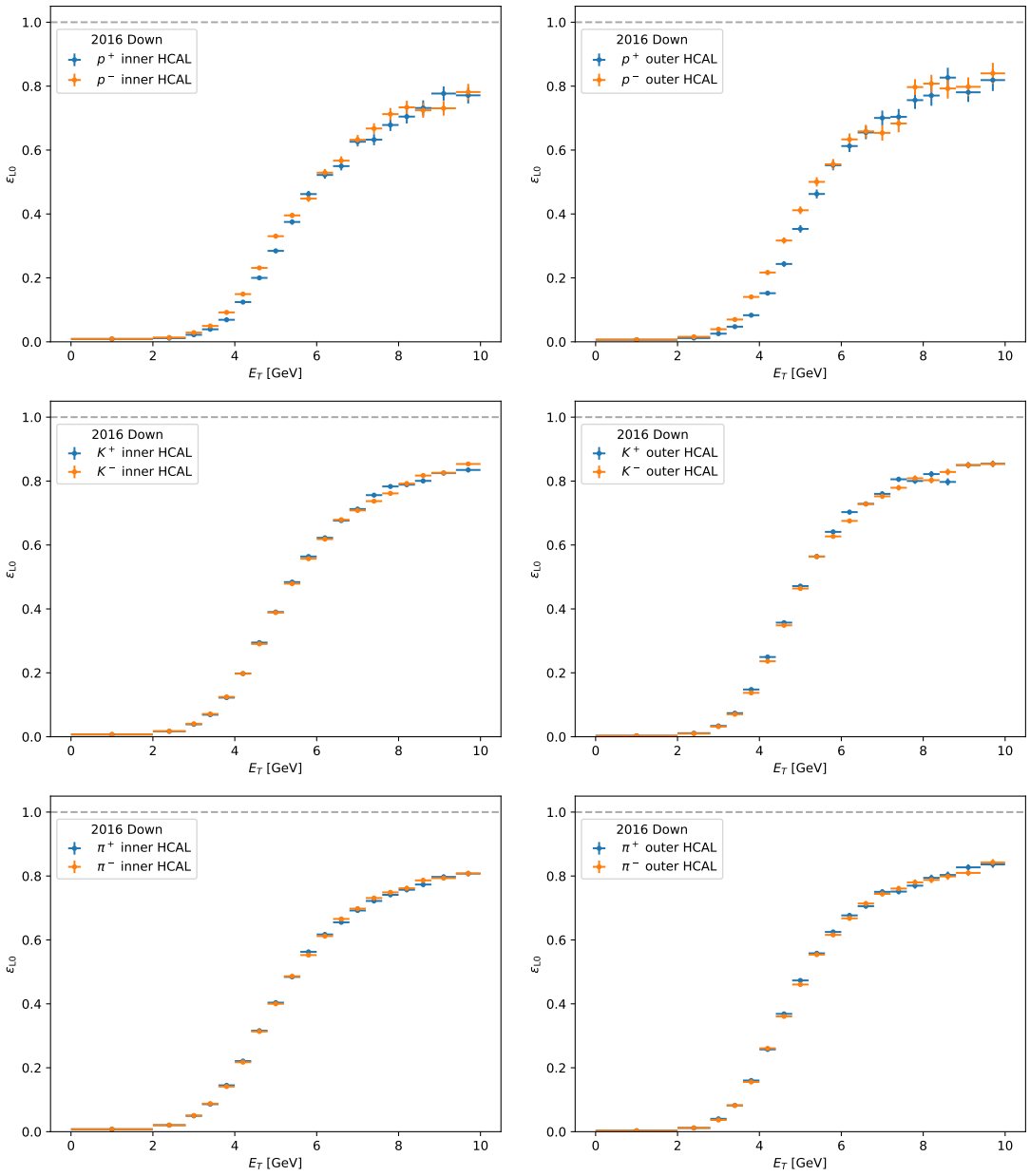

Figure B.8: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2016 magnet Down sample.

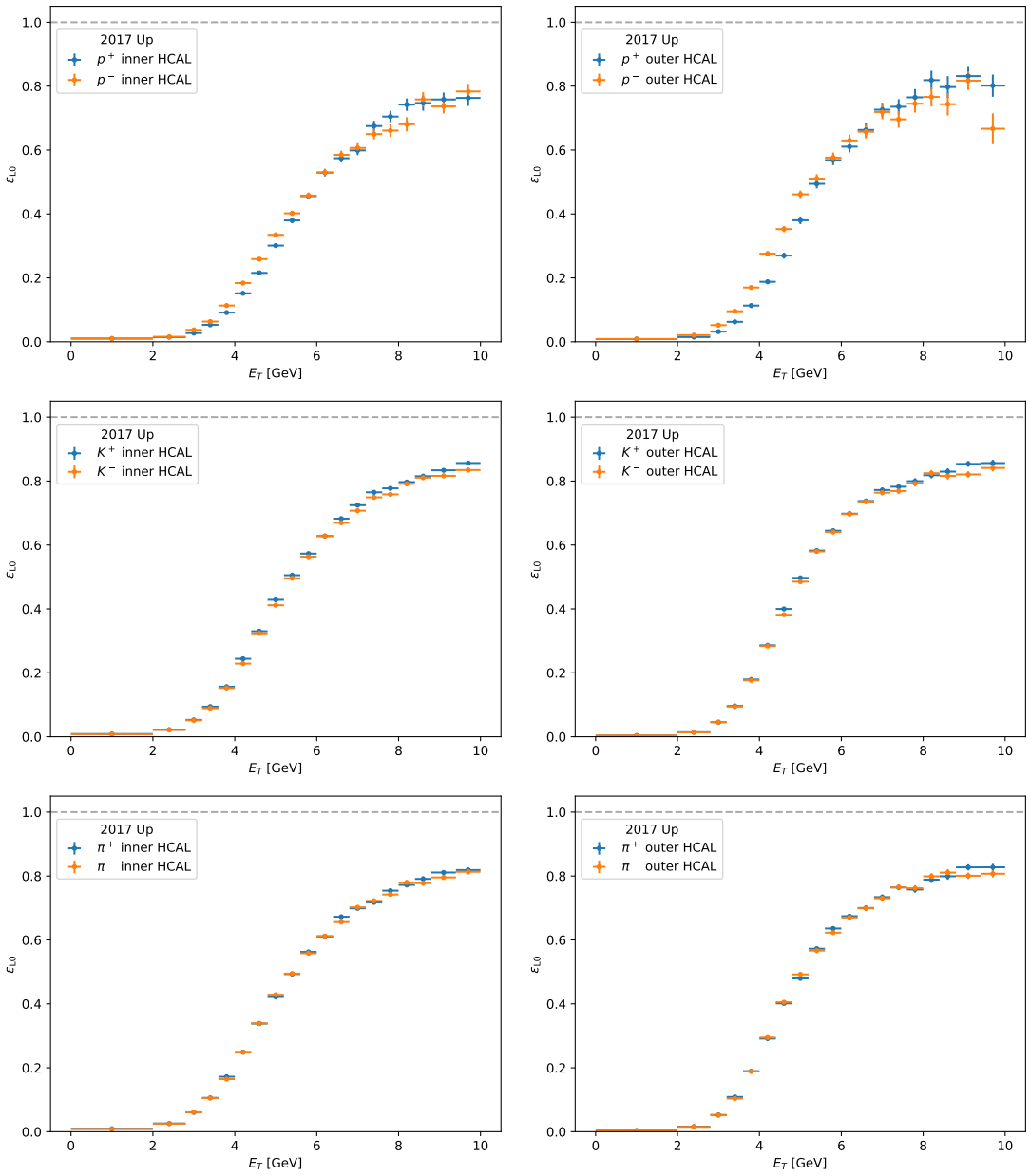

Figure B.9: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2017 magnet Up sample.

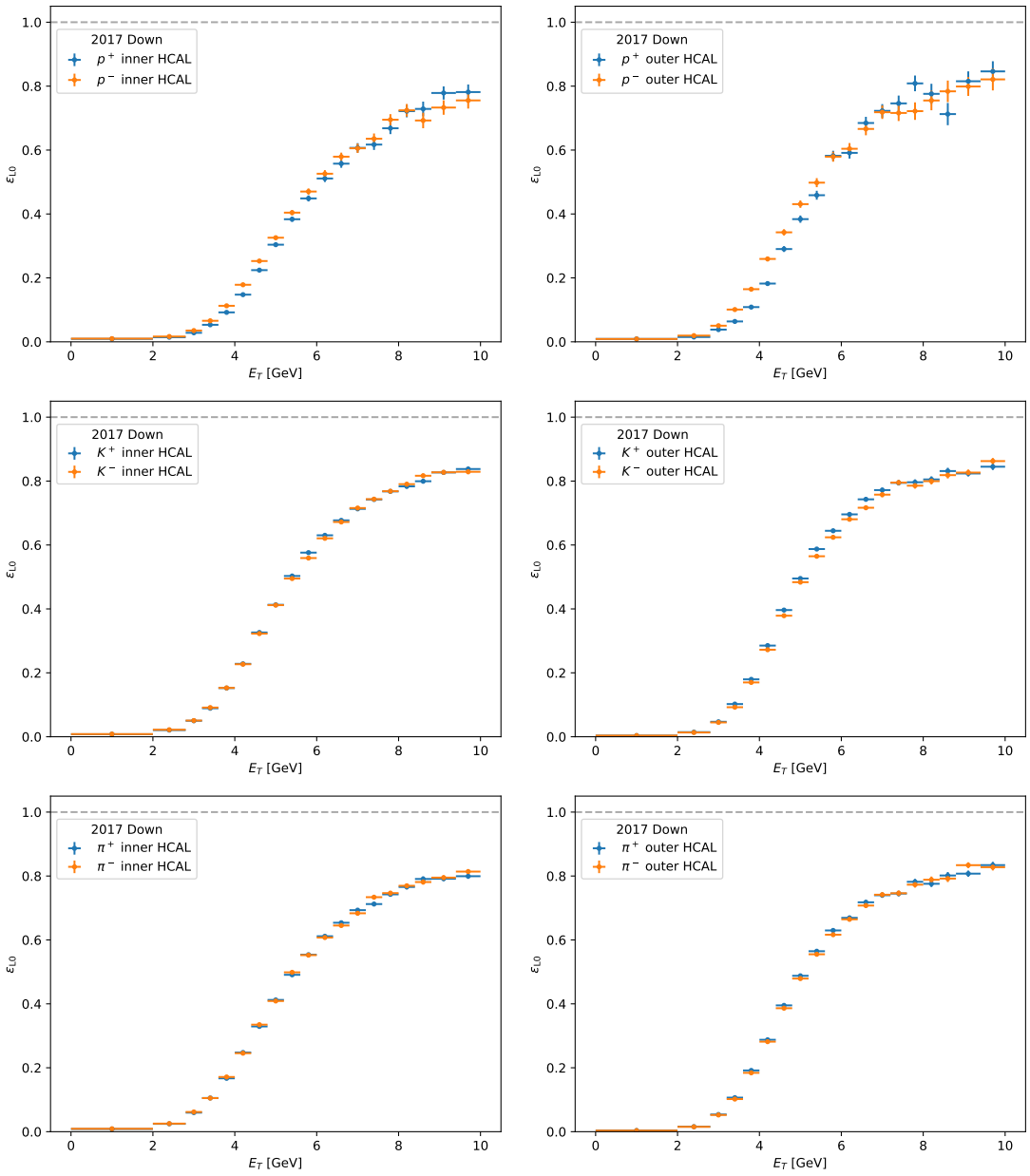

Figure B.10: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2017 magnet Down sample.

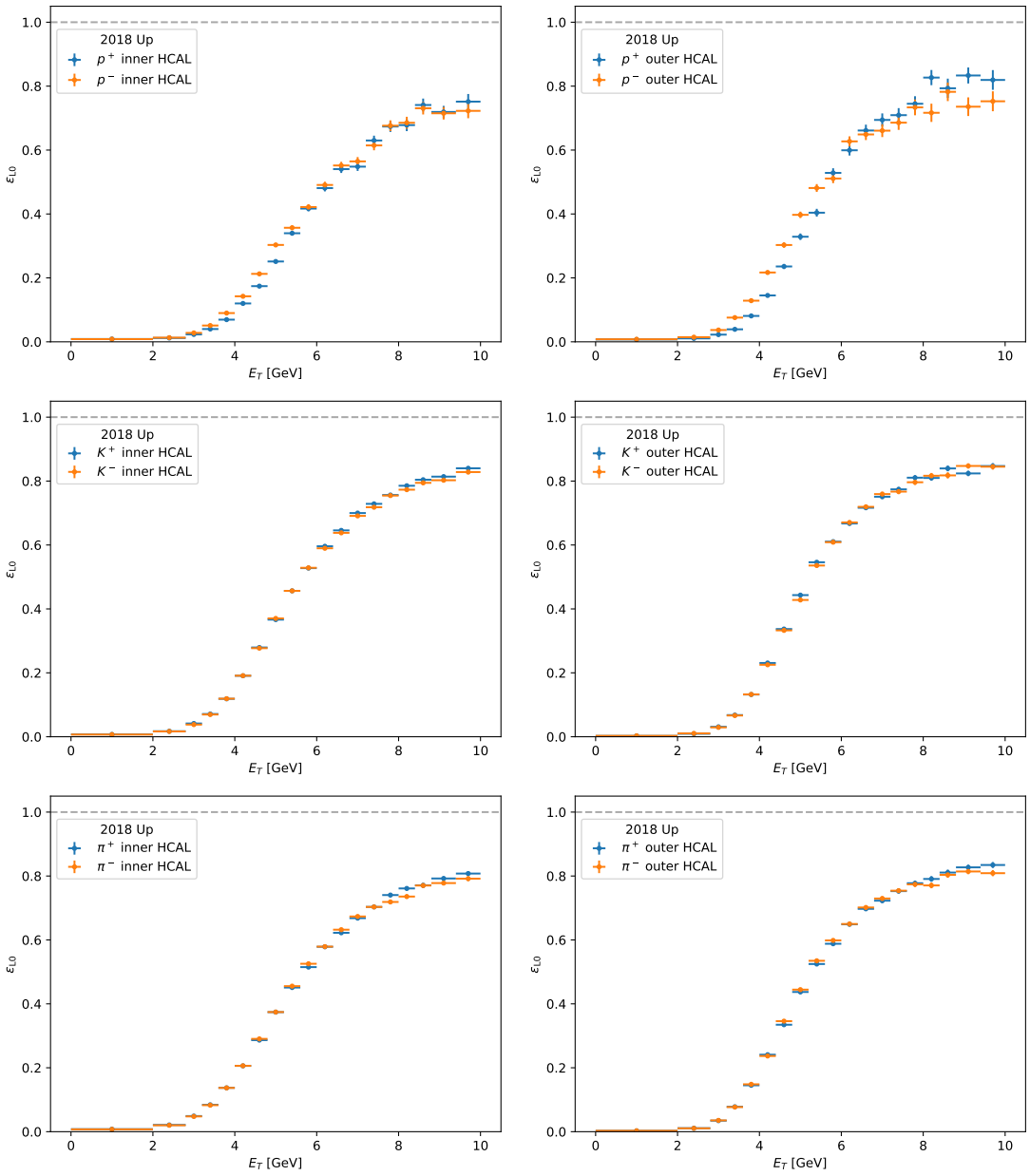

Figure B.11: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2018 magnet Up sample.

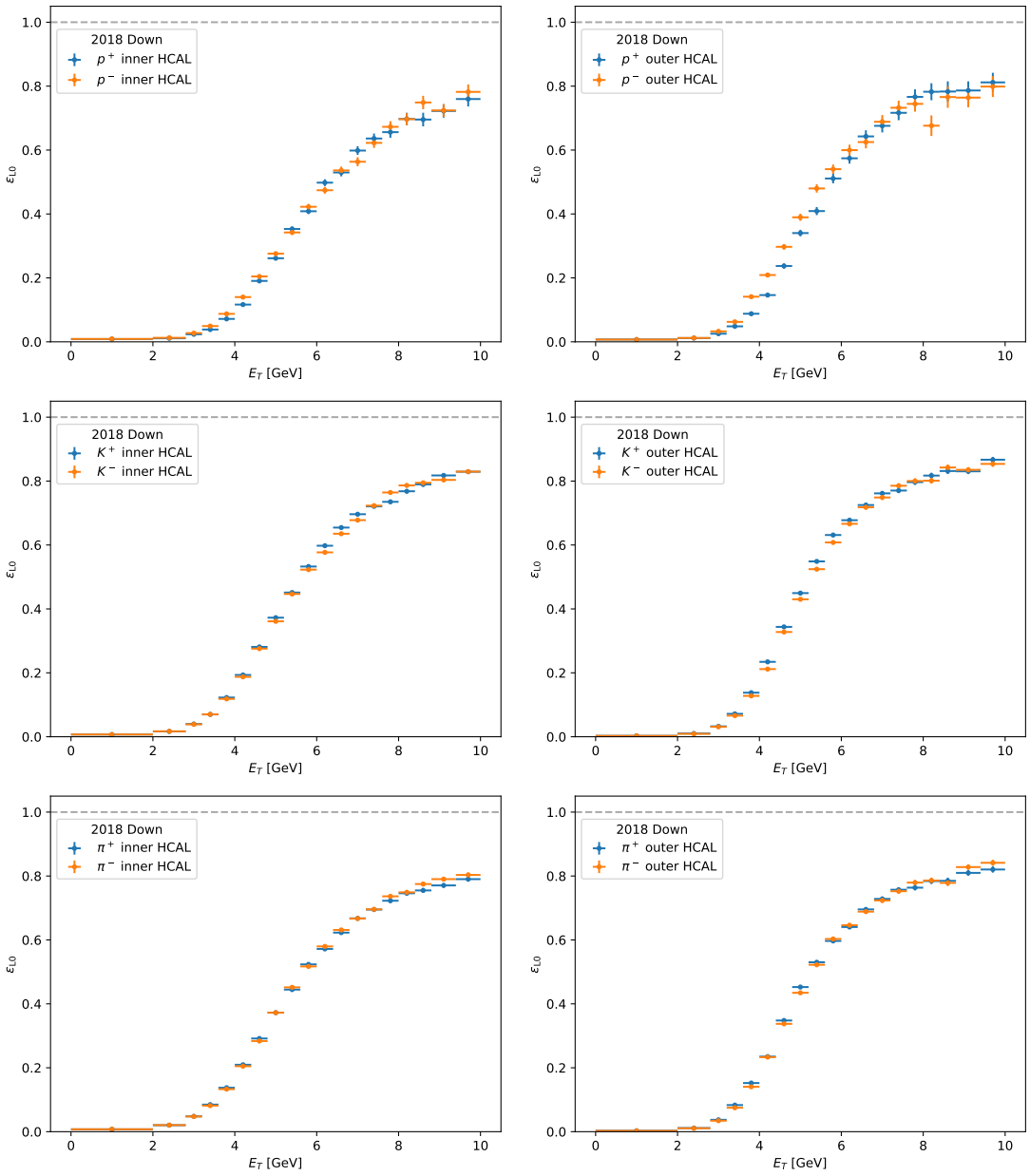

Figure B.12: `L0_Hadron` efficiency as a function of E_{T} for (top row) protons, (middle row) kaons, and (bottom row) pions hitting the (left) inner and (right) outer region of HCAL. These are the plots for the 2018 magnet Down sample.

Appendix C

HLT1 $p_{\mathrm{T}} - \log(\chi^2_{\mathrm{IP}})$ efficiencies

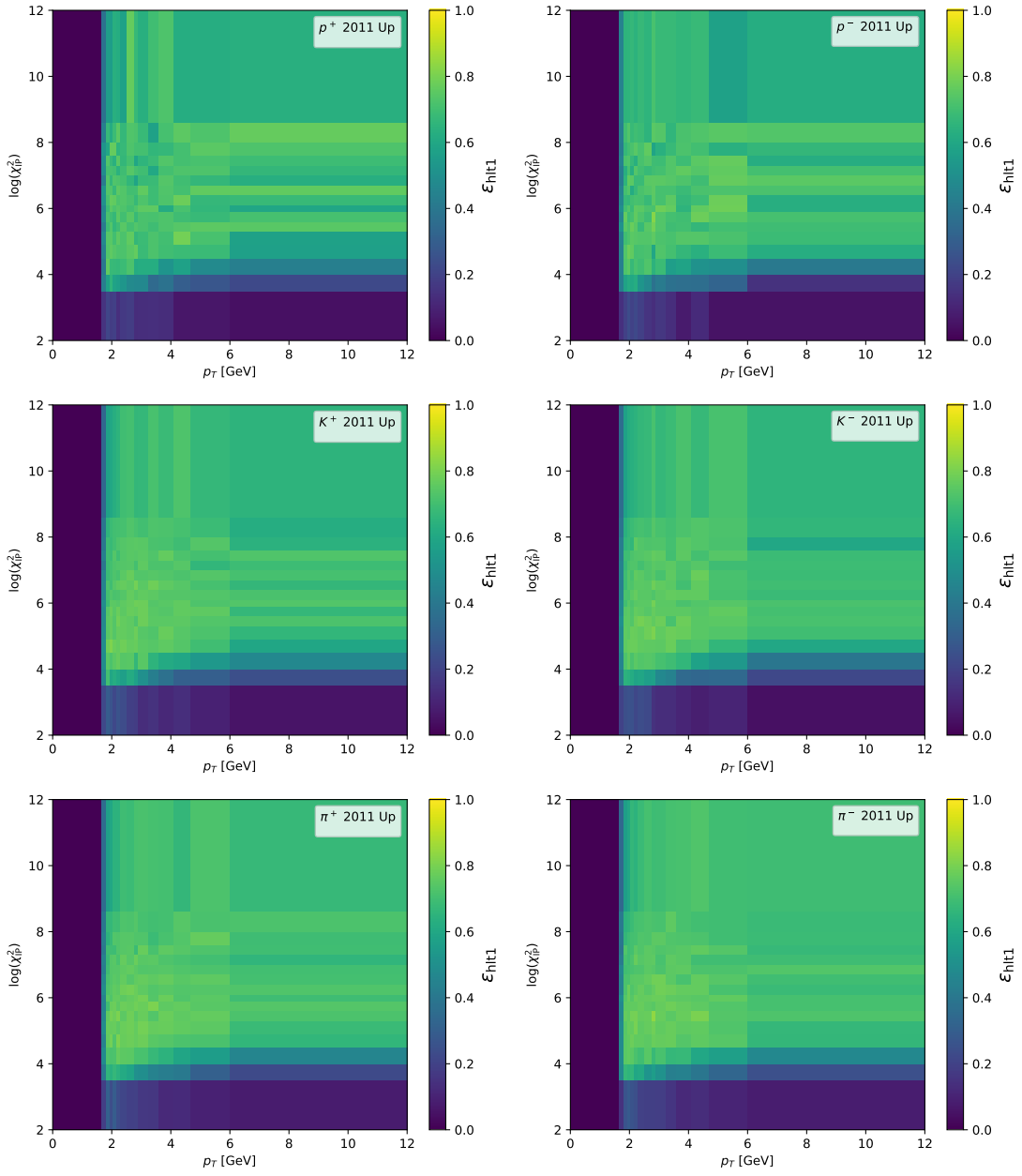

Figure C.1: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2011 magnet Up sample.

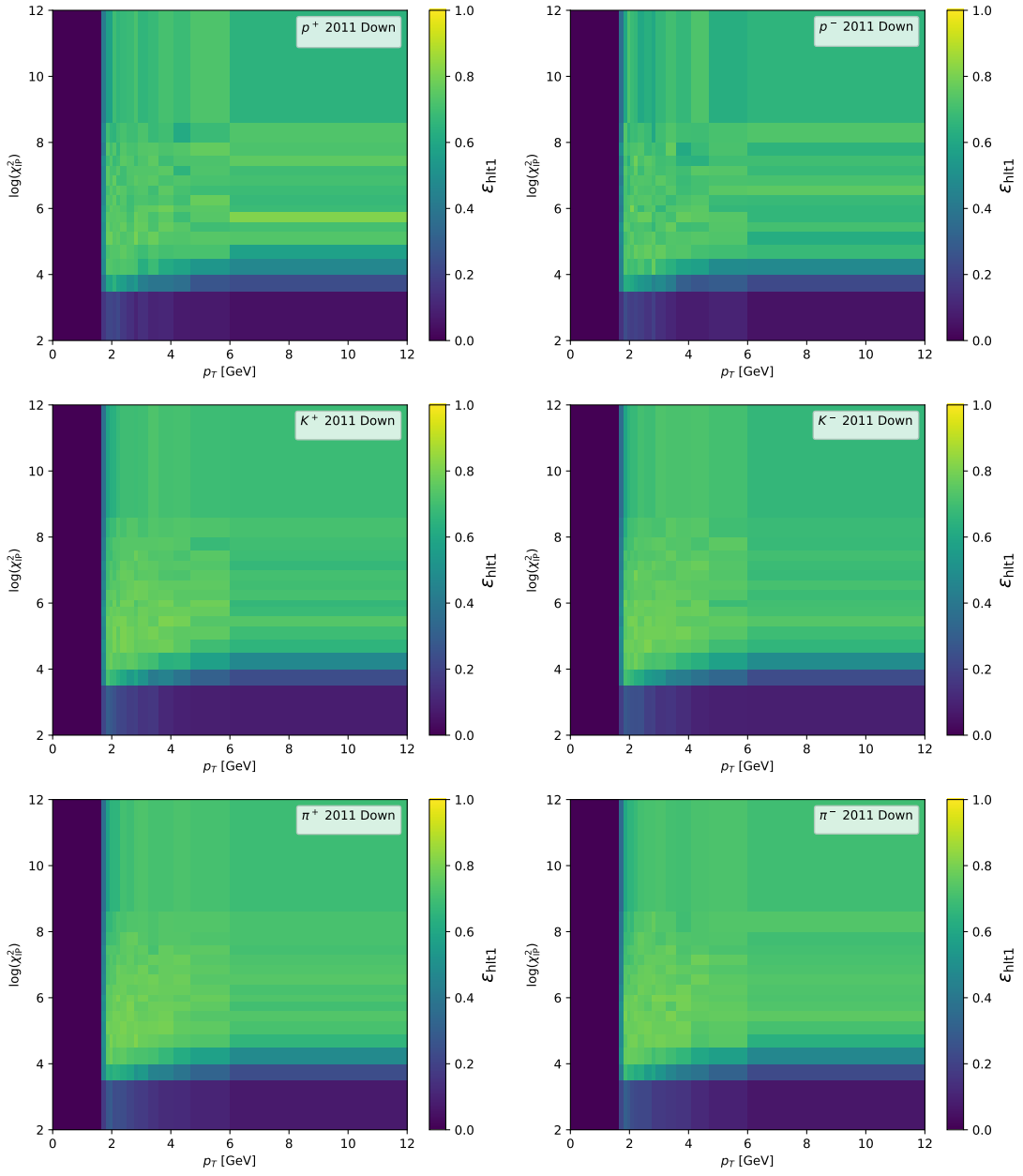

Figure C.2: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2011 magnet Down sample.

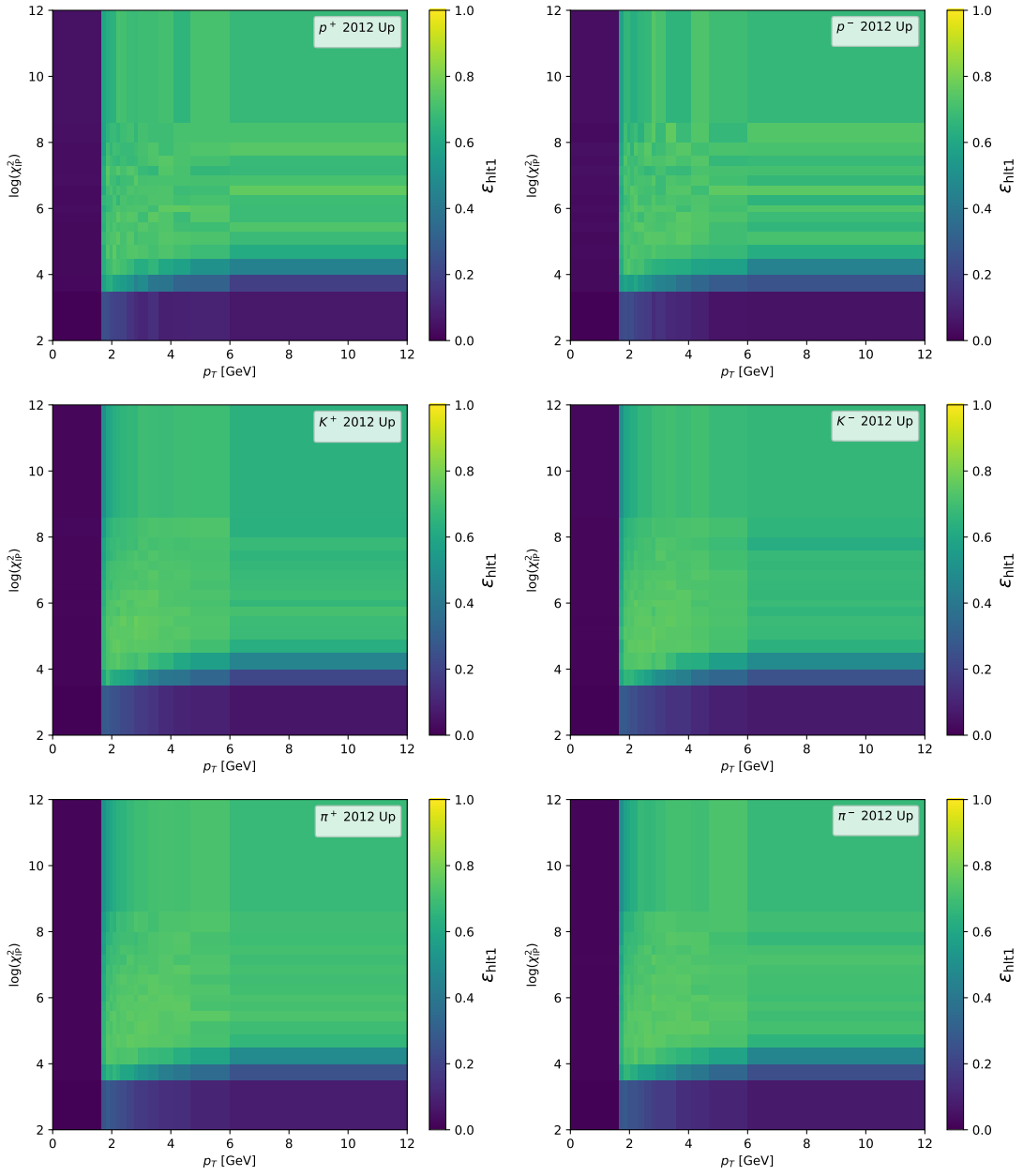

Figure C.3: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2012 magnet Up sample.

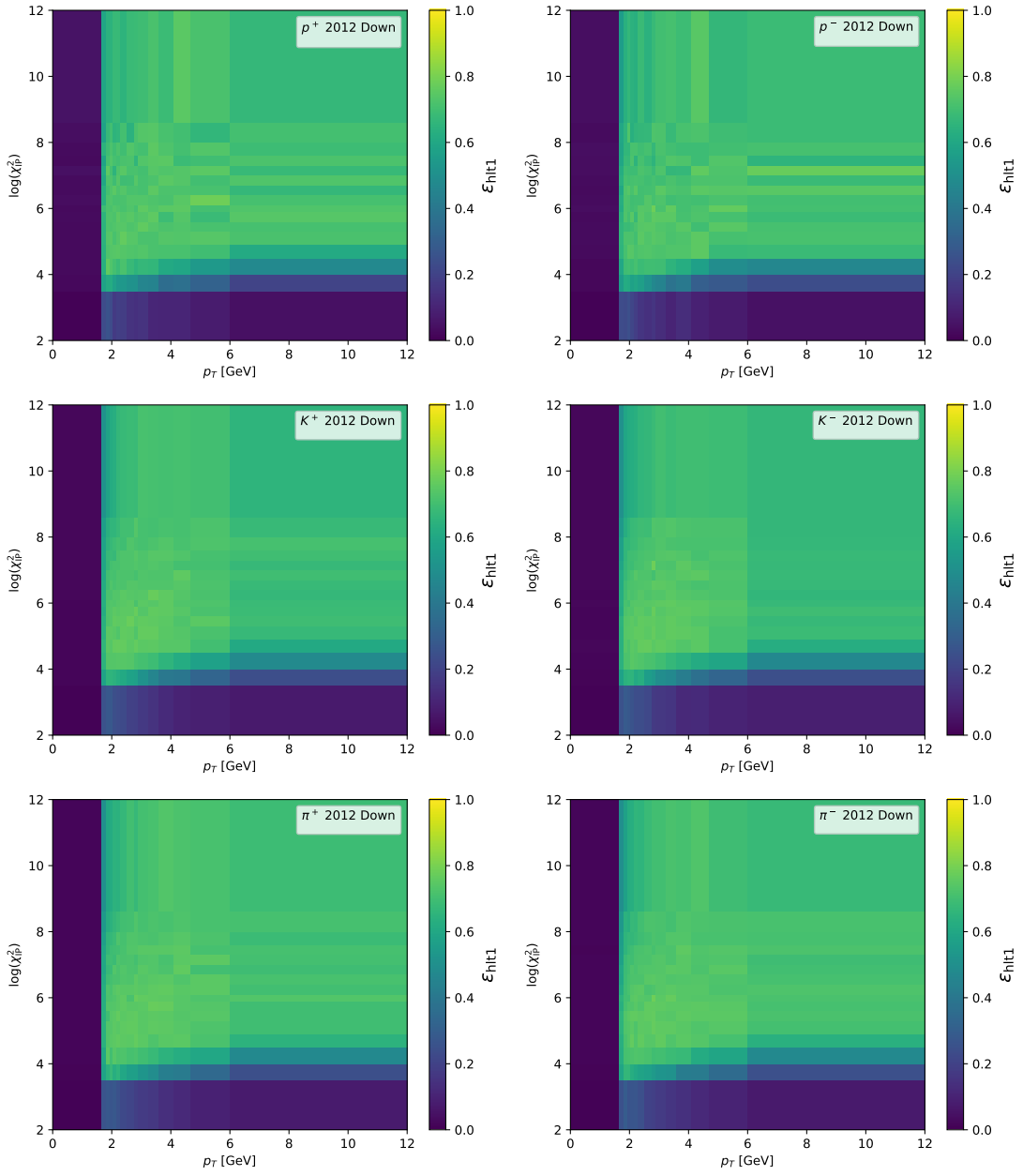

Figure C.4: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2012 magnet Down sample.

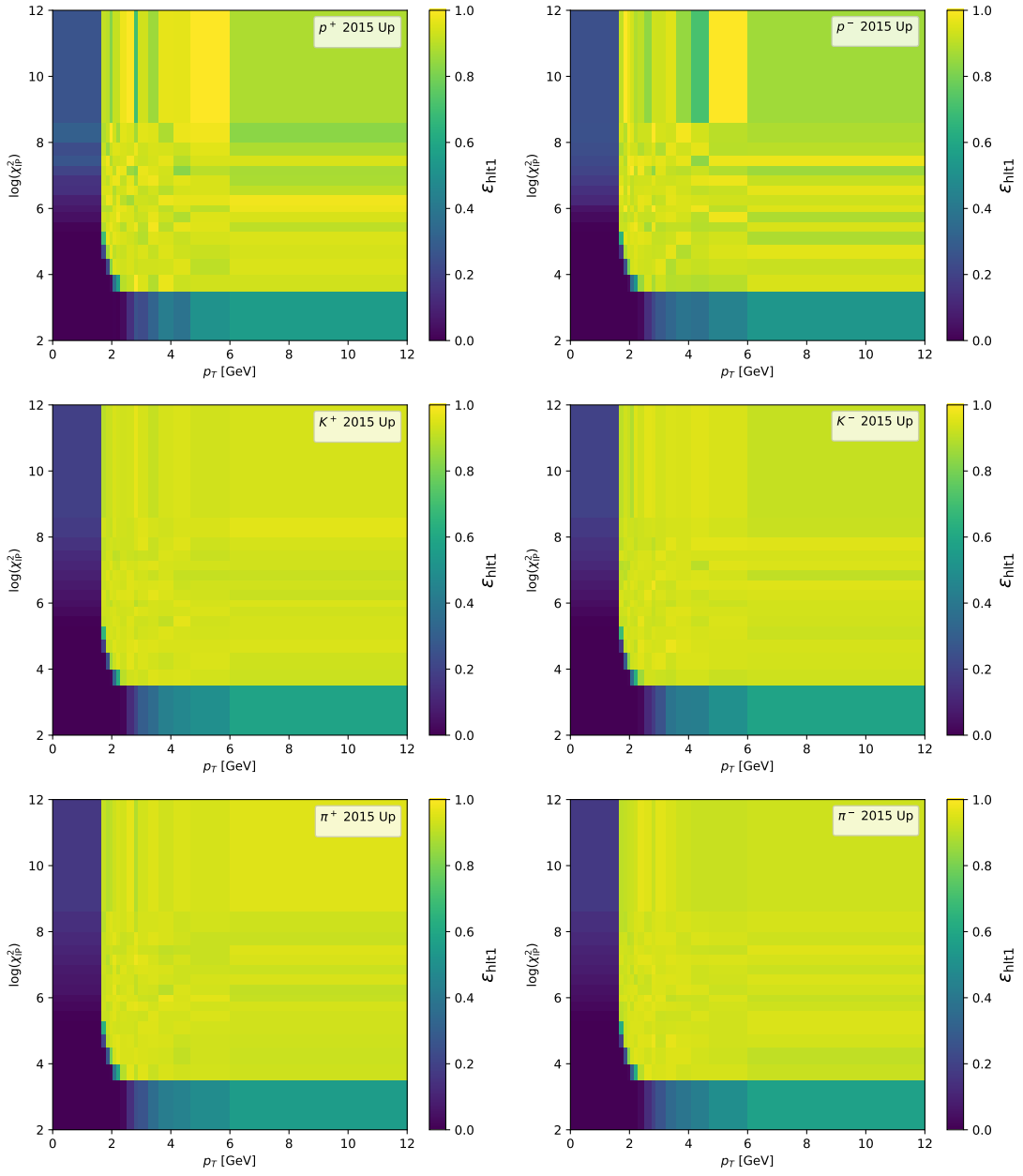

Figure C.5: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2015 magnet Up sample.

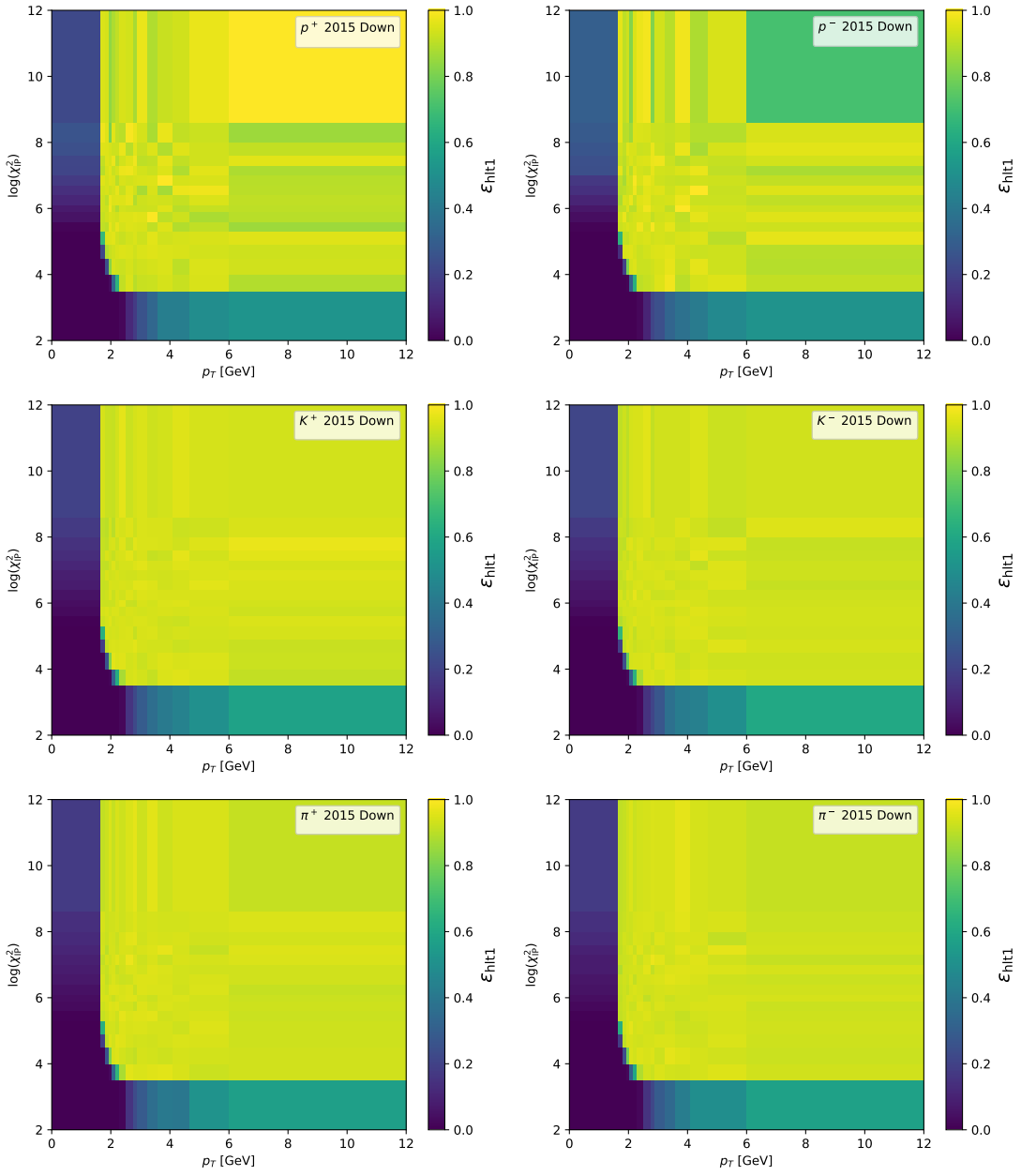

Figure C.6: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2015 magnet Down sample.

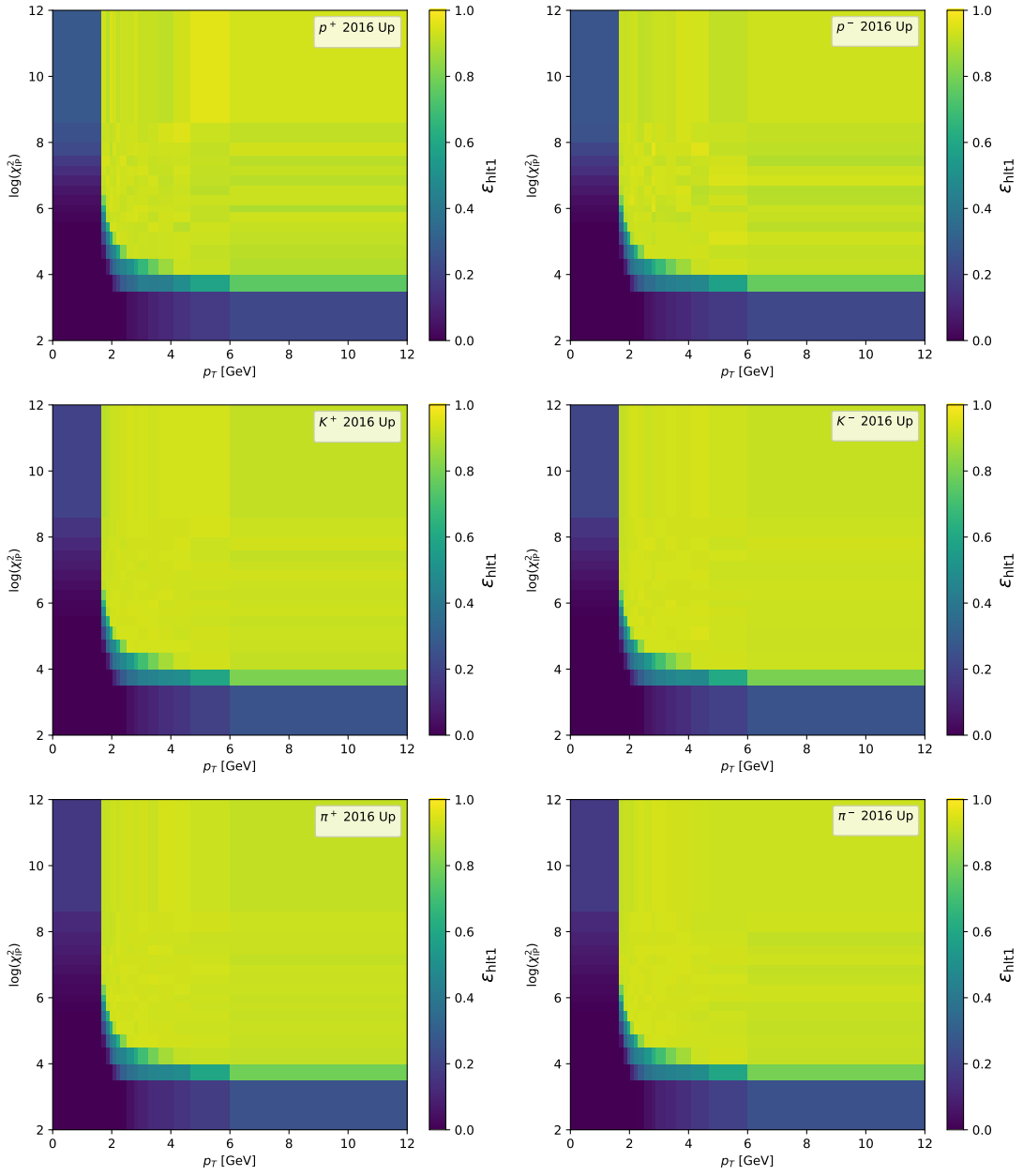

Figure C.7: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2016 magnet Up sample.

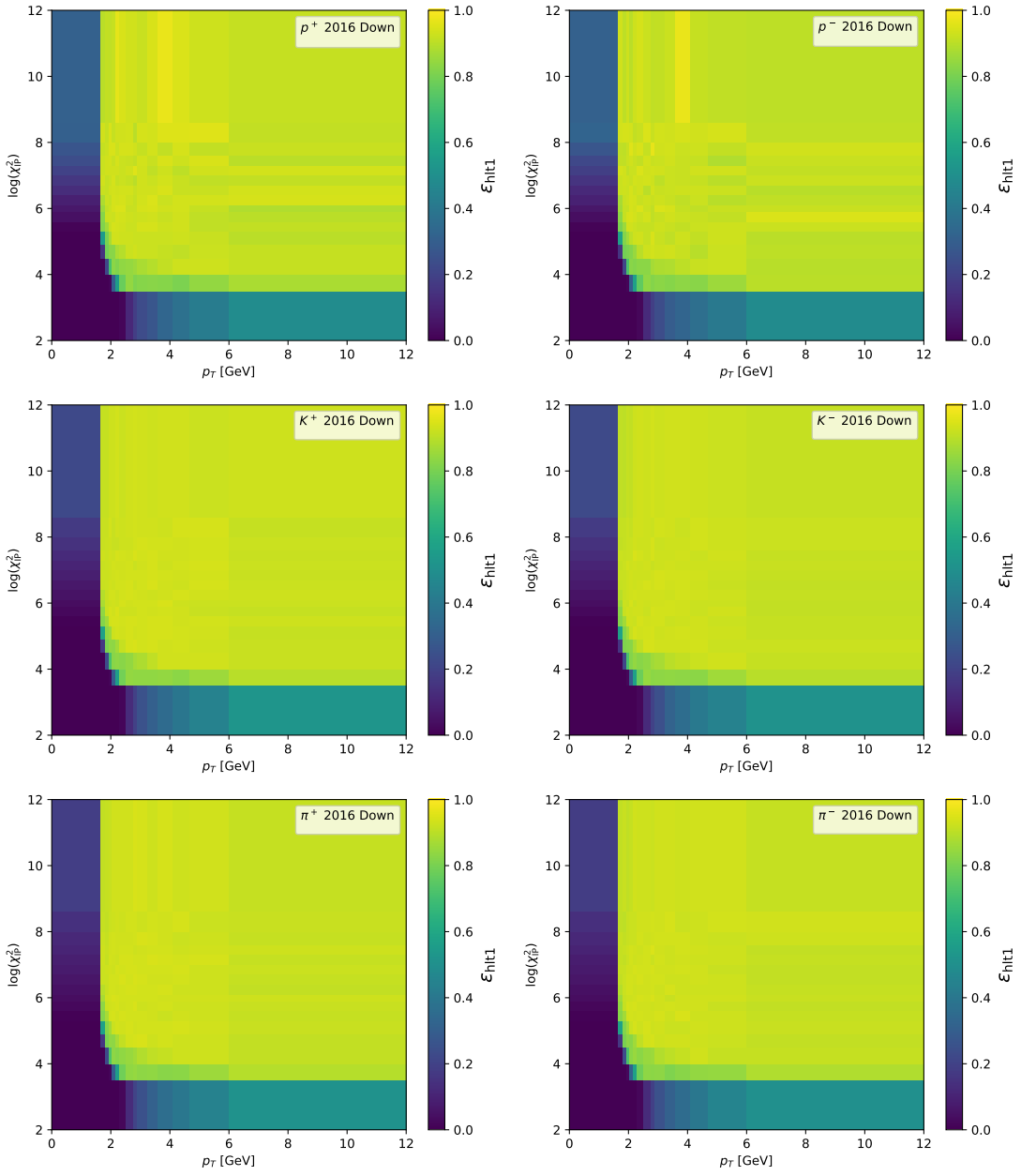

Figure C.8: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2016 magnet Down sample.

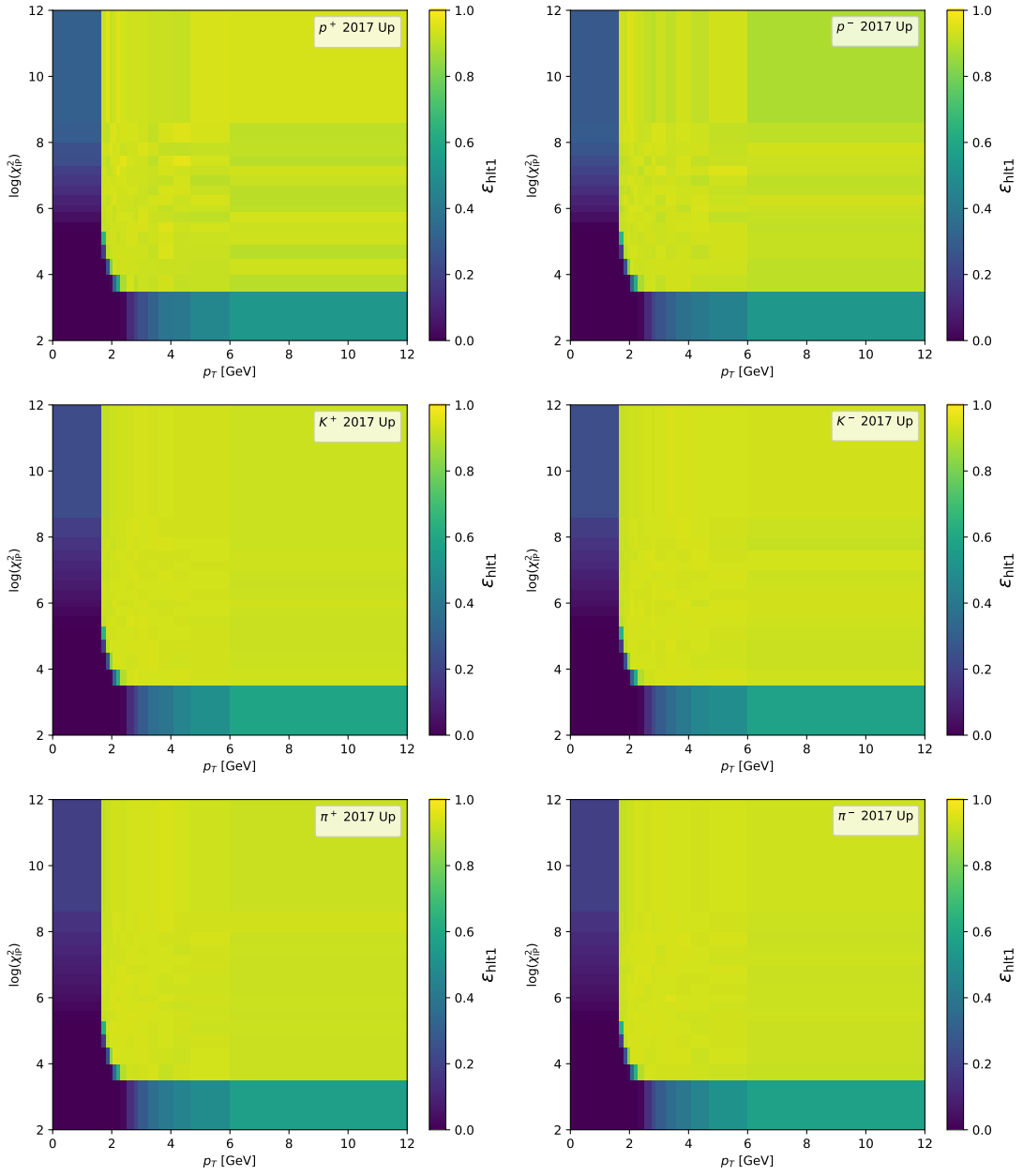

Figure C.9: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2017 magnet Up sample.

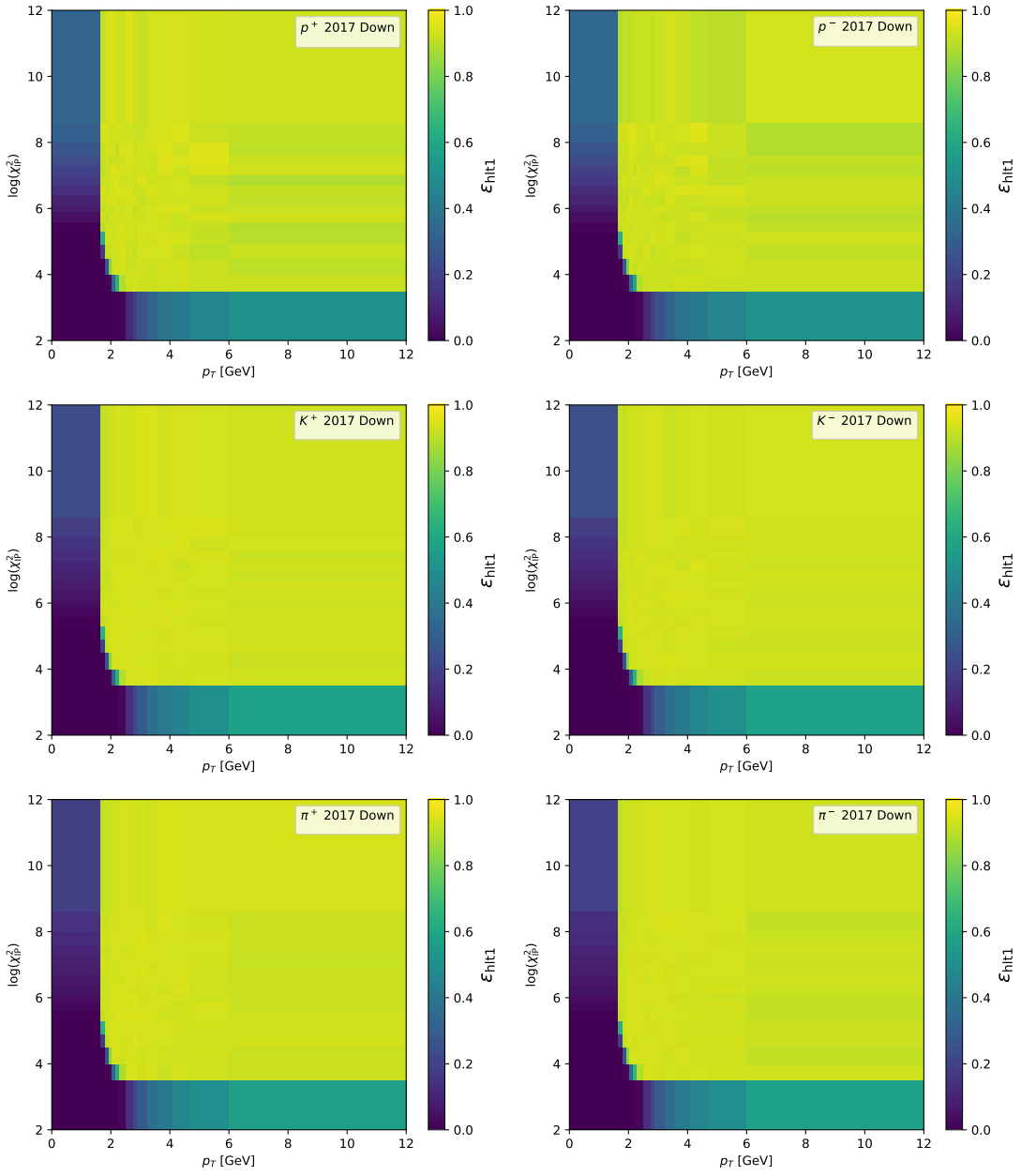

Figure C.10: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2017 magnet Down sample.

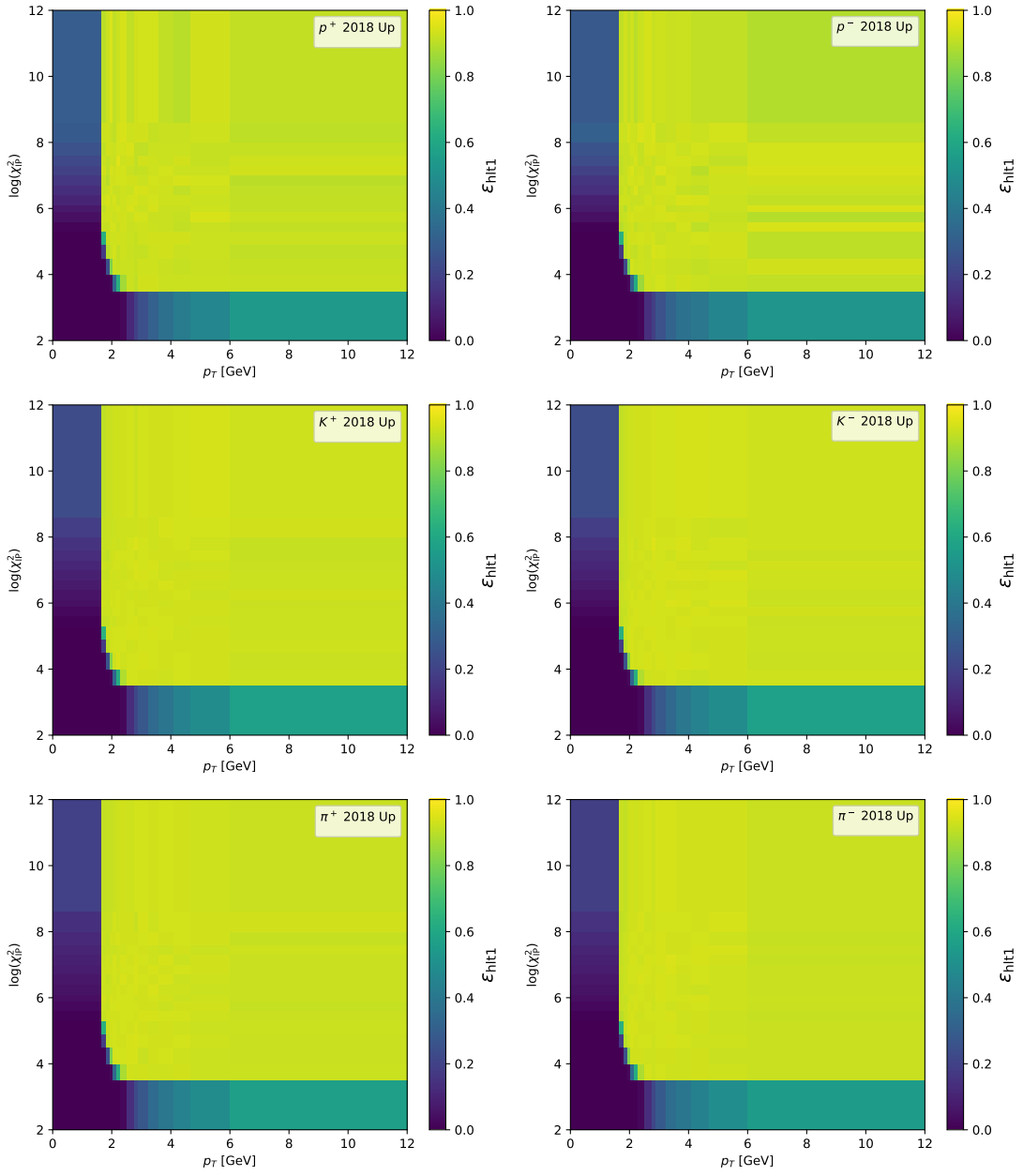

Figure C.11: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2018 magnet Up sample.

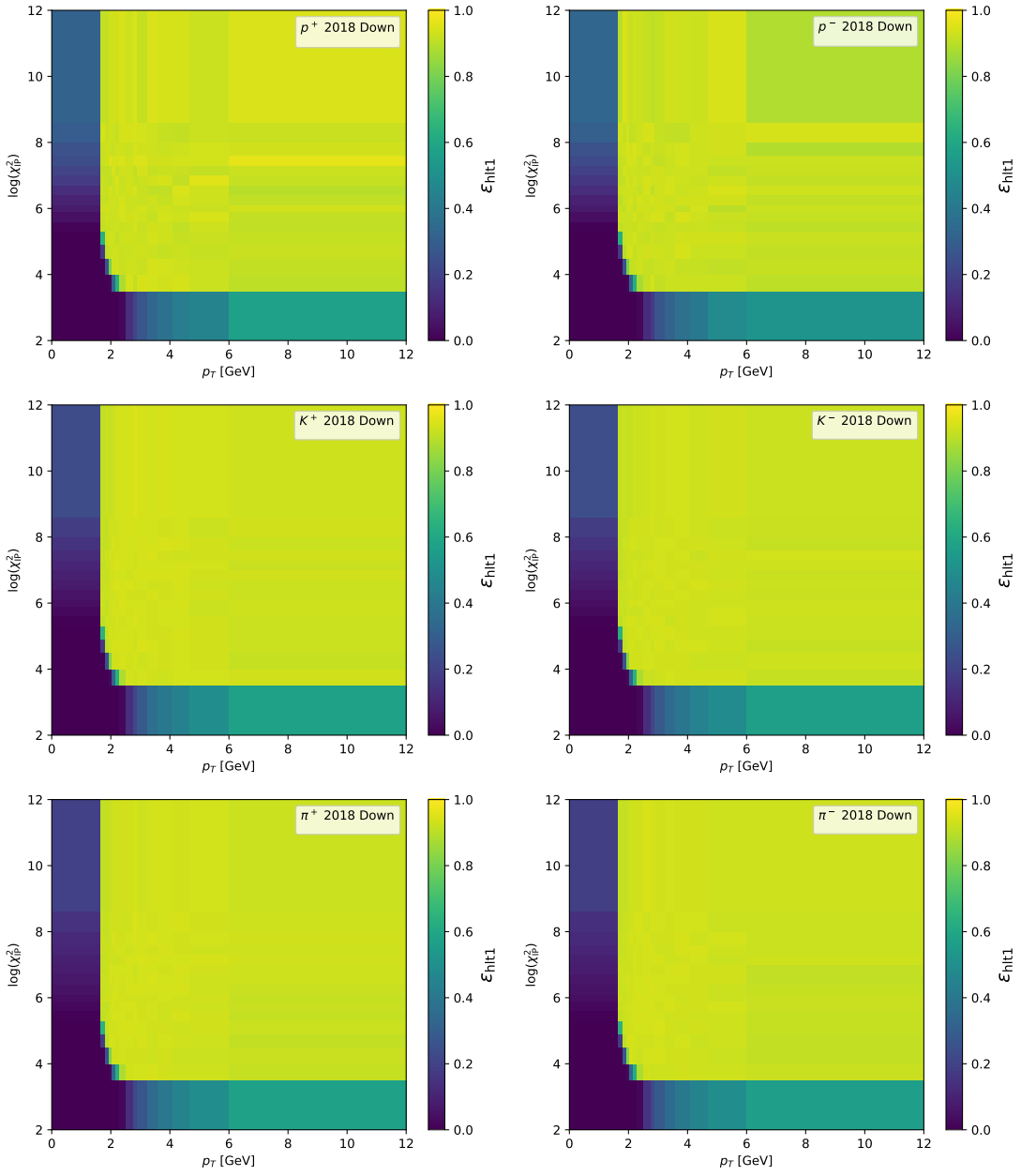

Figure C.12: `Hlt1_TrackMVA` efficiency as a function of p_{T} and $\log\left(\chi^2_{\mathrm{IP}}\right)$ for (top row) protons, (middle row) kaons, and (bottom row) pions of (left) positive and (right) negative charge. These are the plots for the 2018 magnet Down sample.

Appendix D

$\Lambda_b^0 \to \Lambda_c^+ \pi^-$ fits

Linear scale

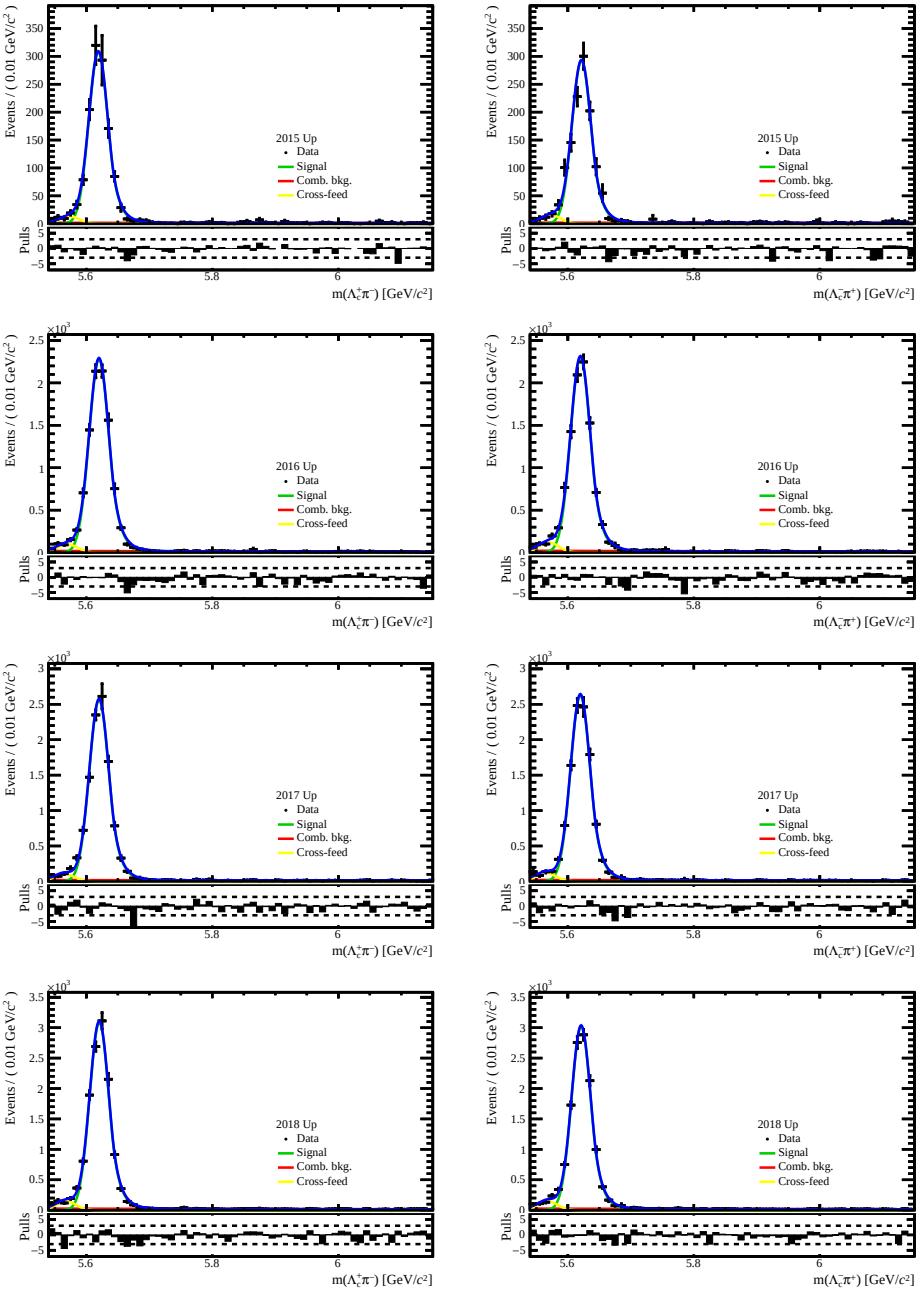

Figure D.1: Fits to the invariant-mass distributions of (left) $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ and (right) $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays for the years from 2015 to 2018 magnet Up from top to bottom; the plots are in linear scale.

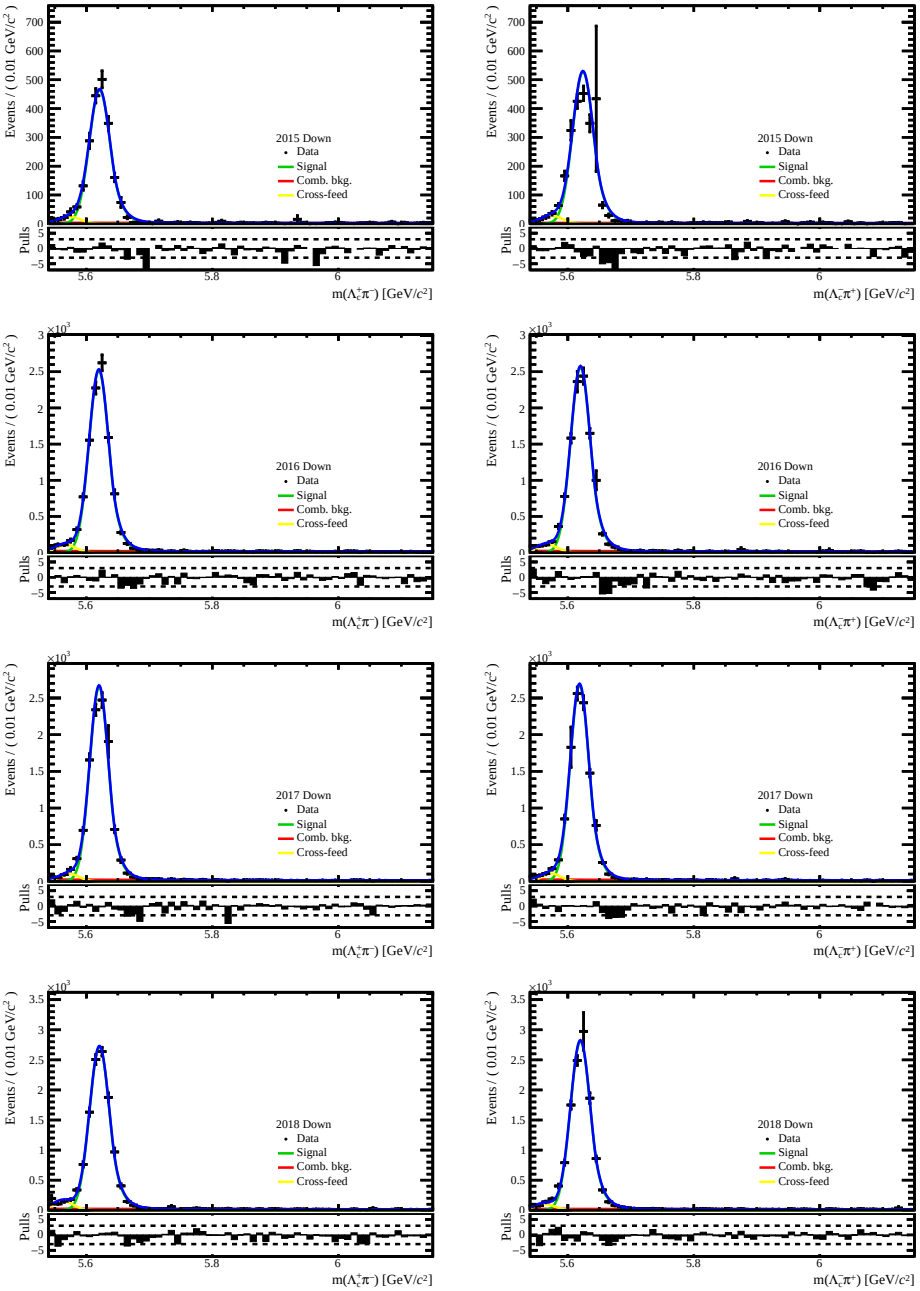

Figure D.2: Fits to the invariant-mass distributions of (left) $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ and (right) $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays for the years from 2015 to 2018 magnet Down from top to bottom; the plots are in linear scale.

Logarithmic scale

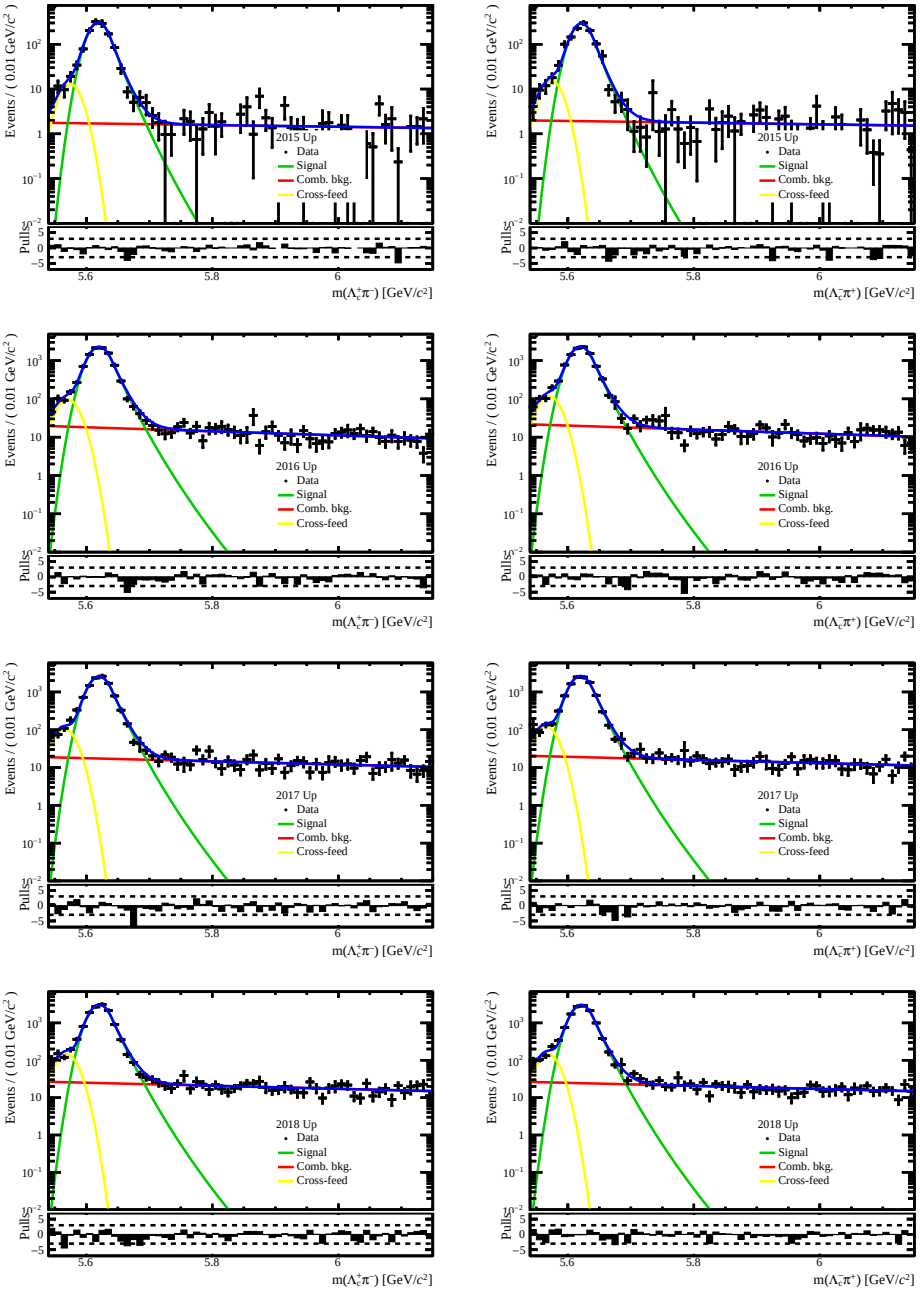

Figure D.3: Fits to the invariant-mass distributions of (left) $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ and (right) $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays for the years from 2015 to 2018 magnet Up from top to bottom; the plots are in logarithmic scale.

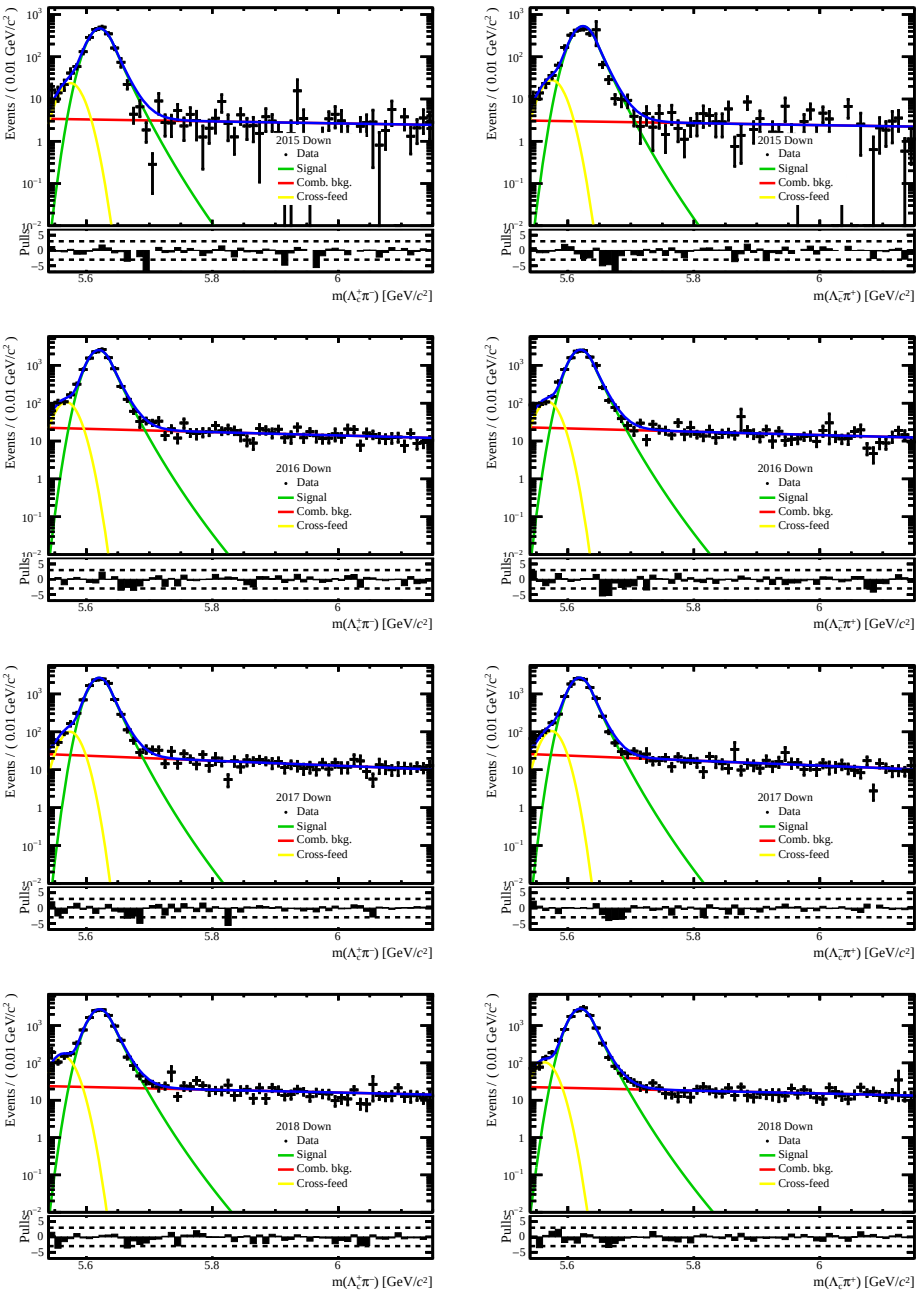

Figure D.4: Fits to the invariant-mass distributions of (left) $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ and (right) $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays for the years from 2015 to 2018 magnet Down from top to bottom; the plots are in logarithmic scale.

Appendix E

A_{CP} scatter plots

$\Lambda_b^0 \to pK^-$

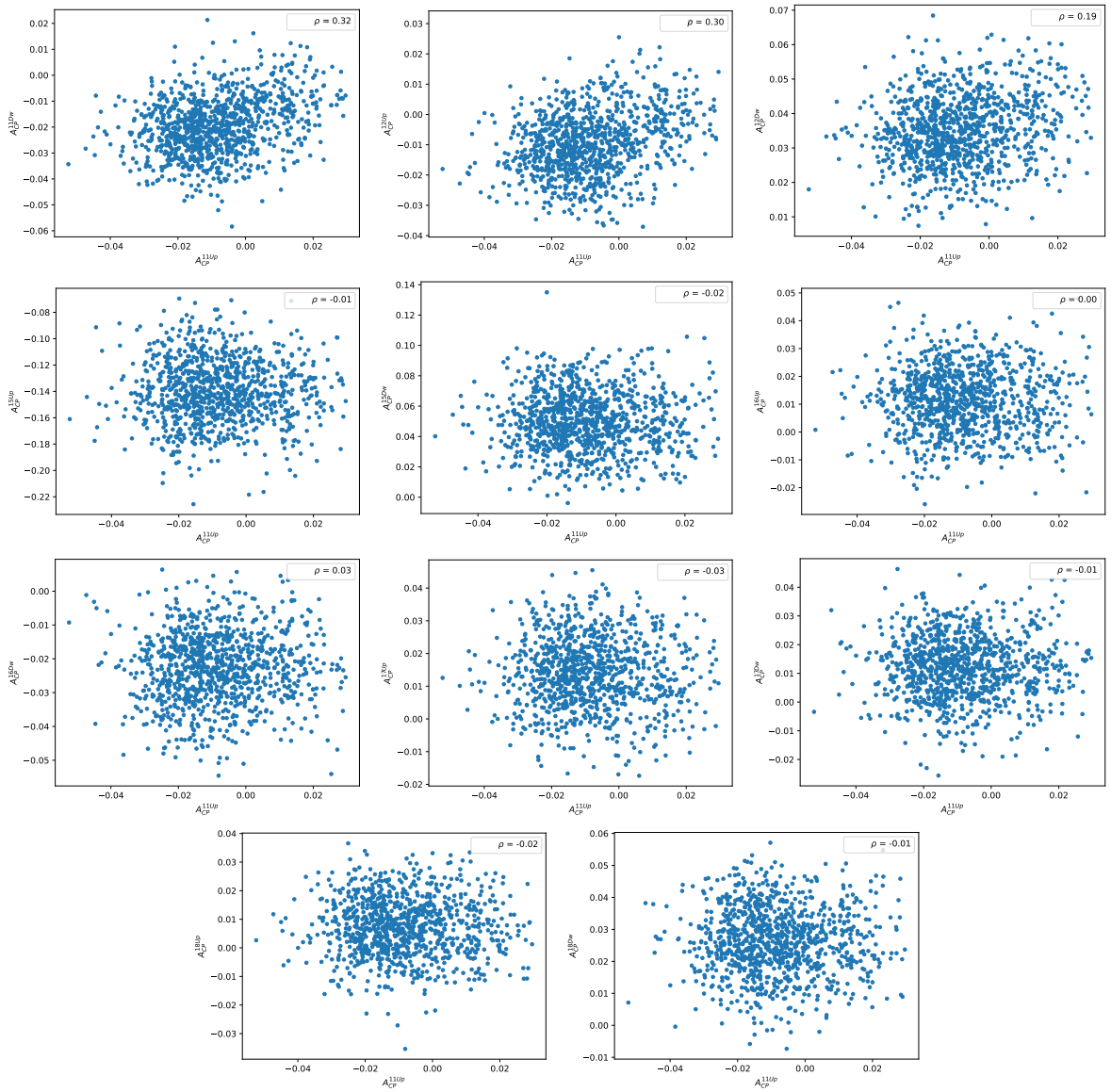

Figure E.1: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2011 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

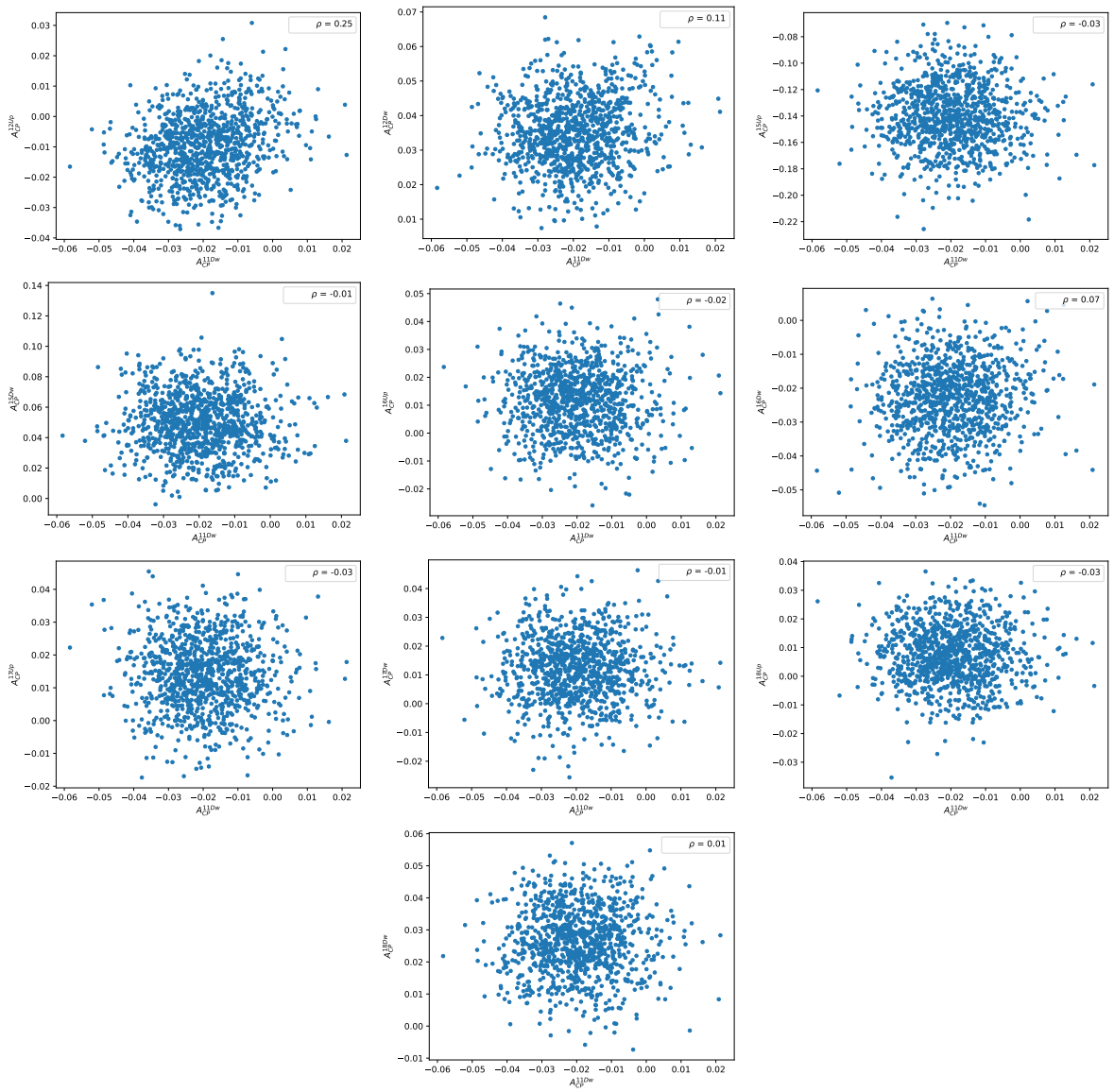

Figure E.2: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2011 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

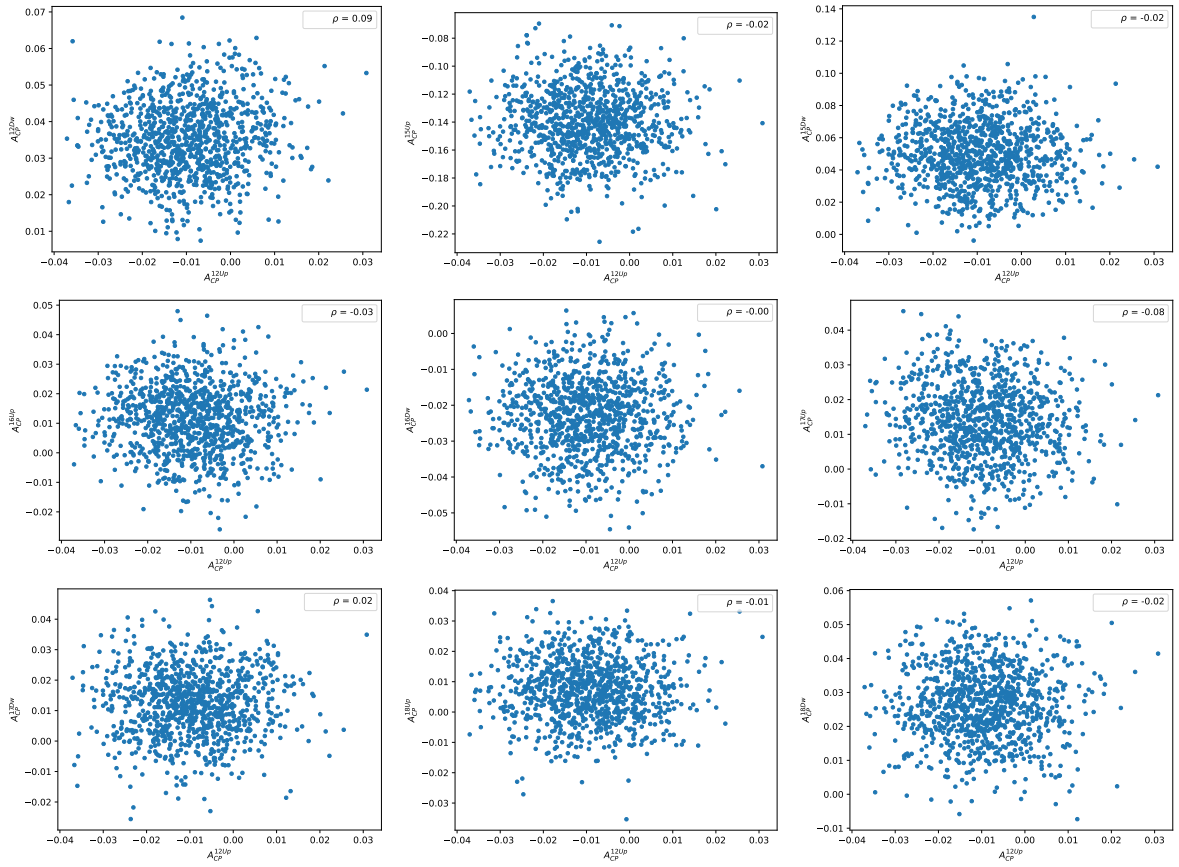

Figure E.3: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2012 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

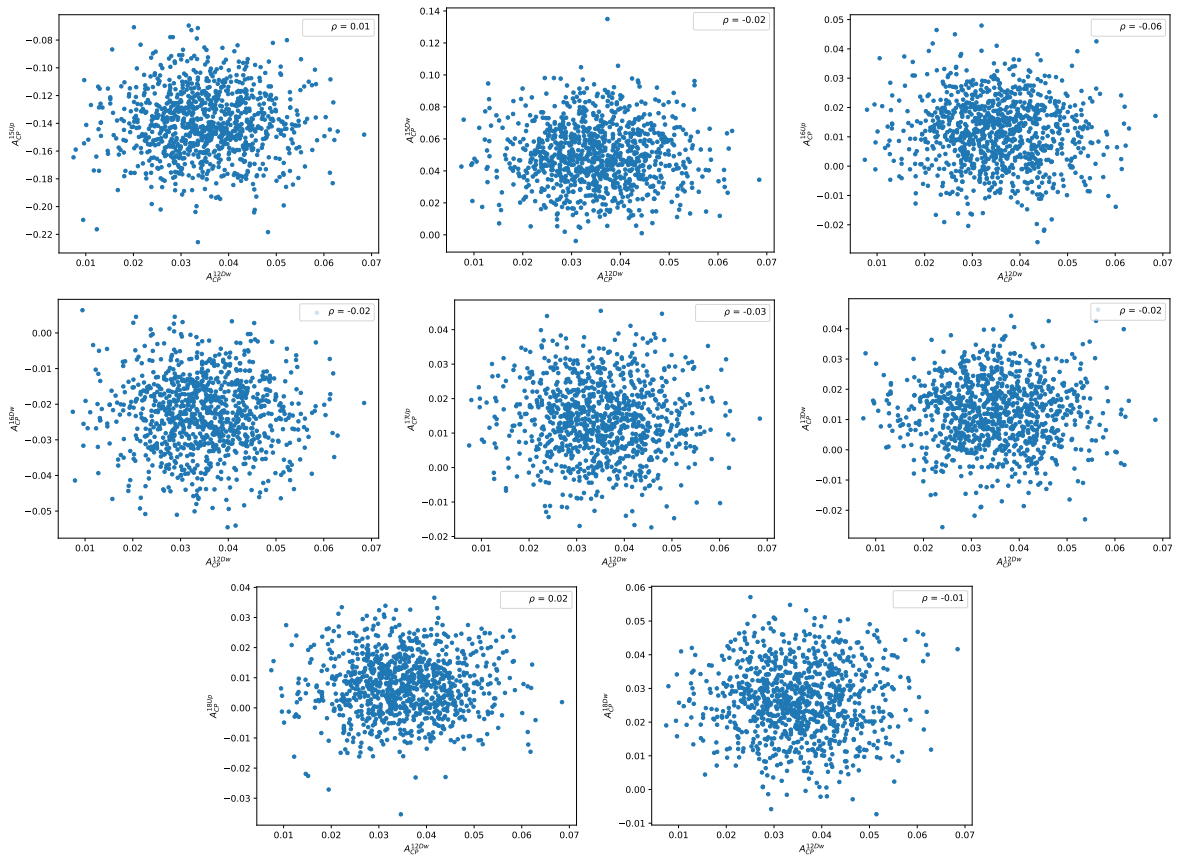

Figure E.4: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2012 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

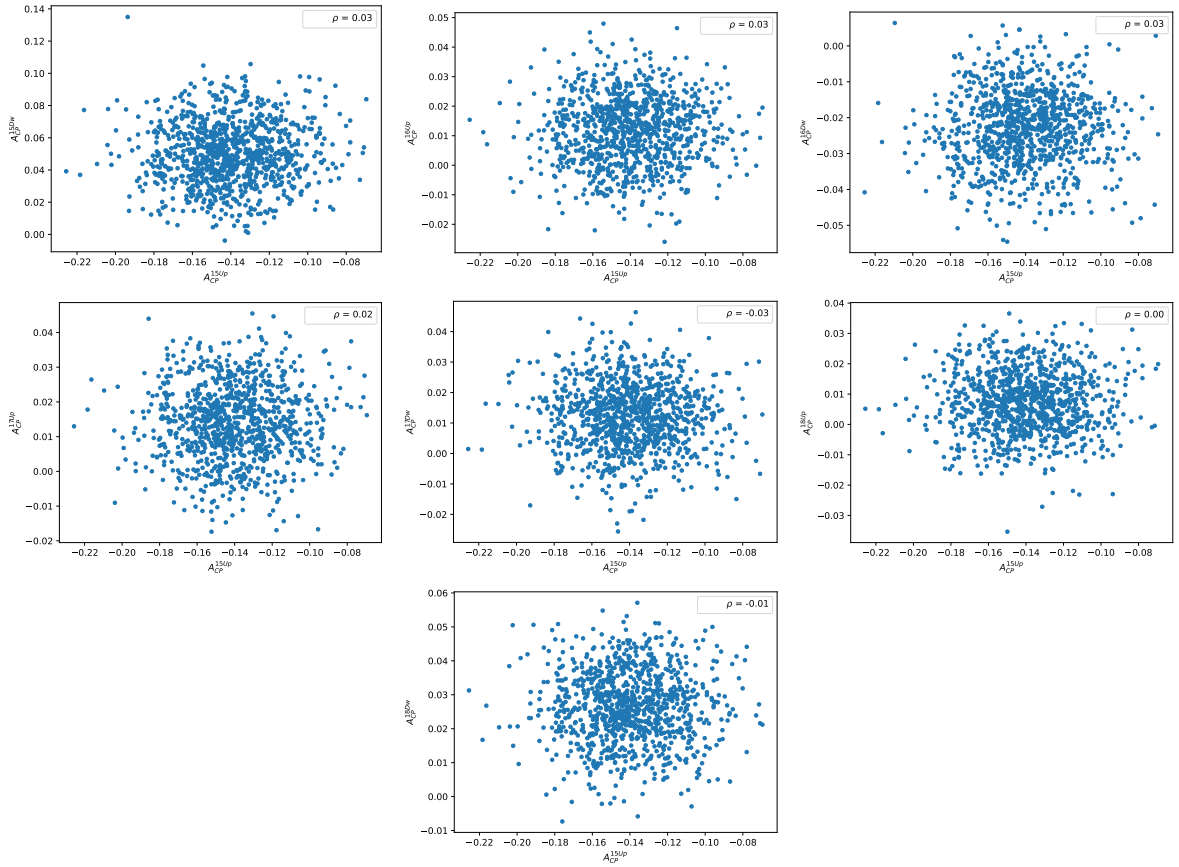

Figure E.5: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2015 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

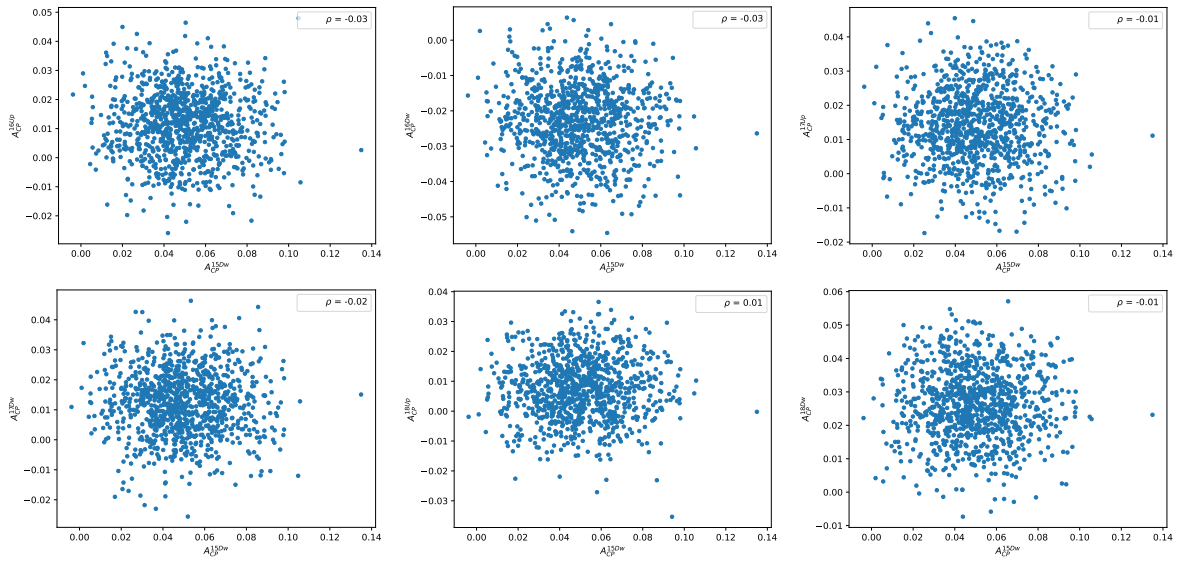

Figure E.6: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2015 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

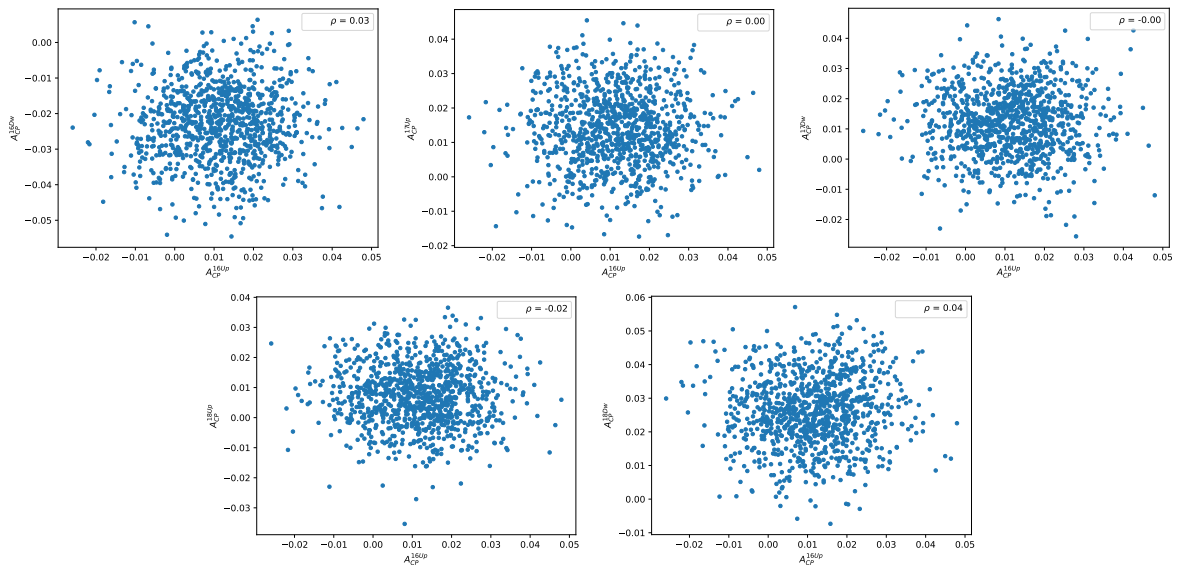

Figure E.7: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2016 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

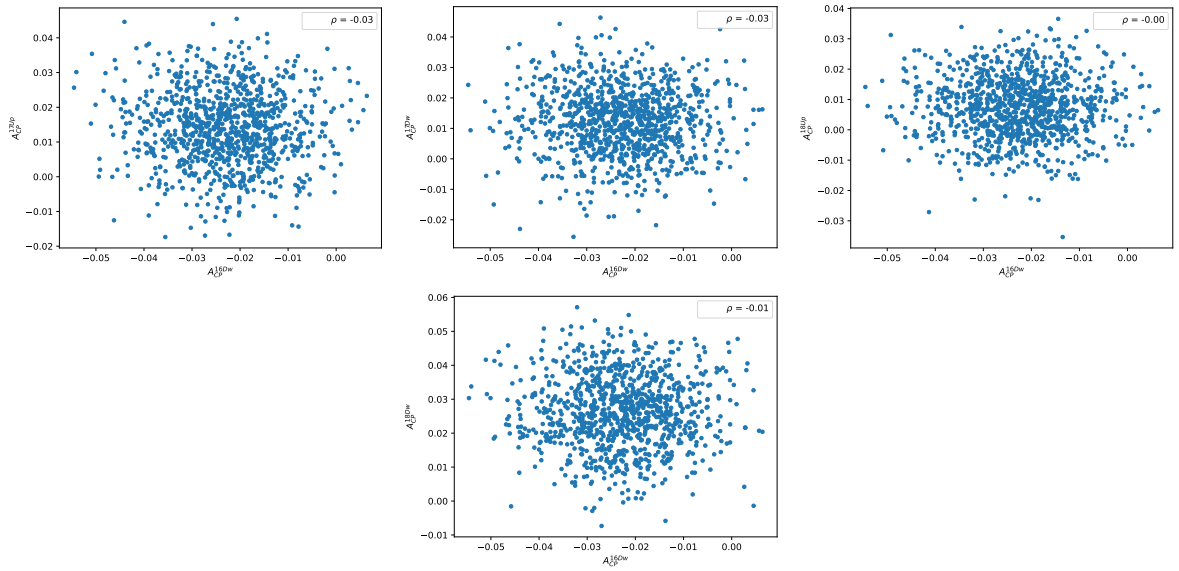

Figure E.8: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2016 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

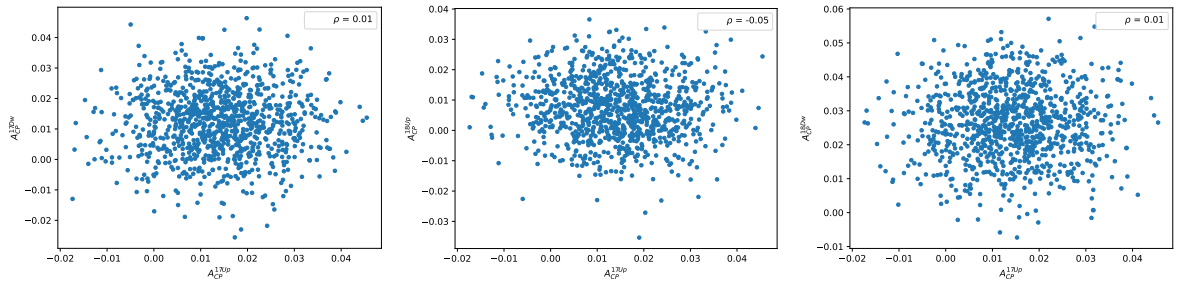

Figure E.9: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2017 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

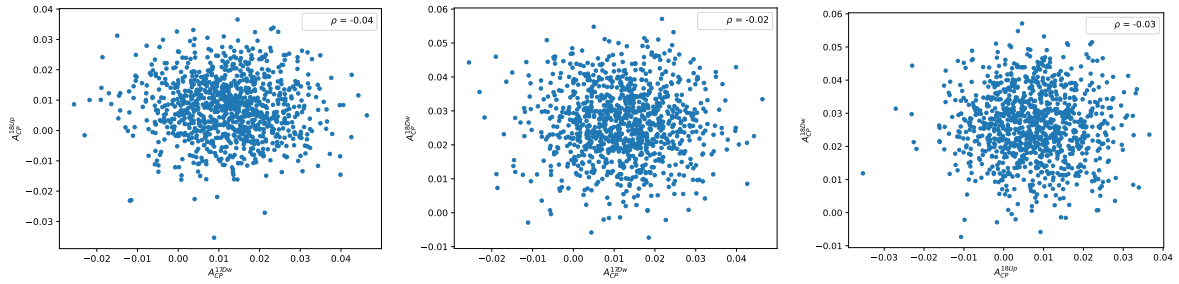

Figure E.10: Scatter plots of the 1000 values of $A_{CP}(pK^-)$ obtained from the pseudoexperiments in the 2017 magnet Down sample and 2018 magnet Up and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

$\Lambda_b^0 \to p\pi^-$

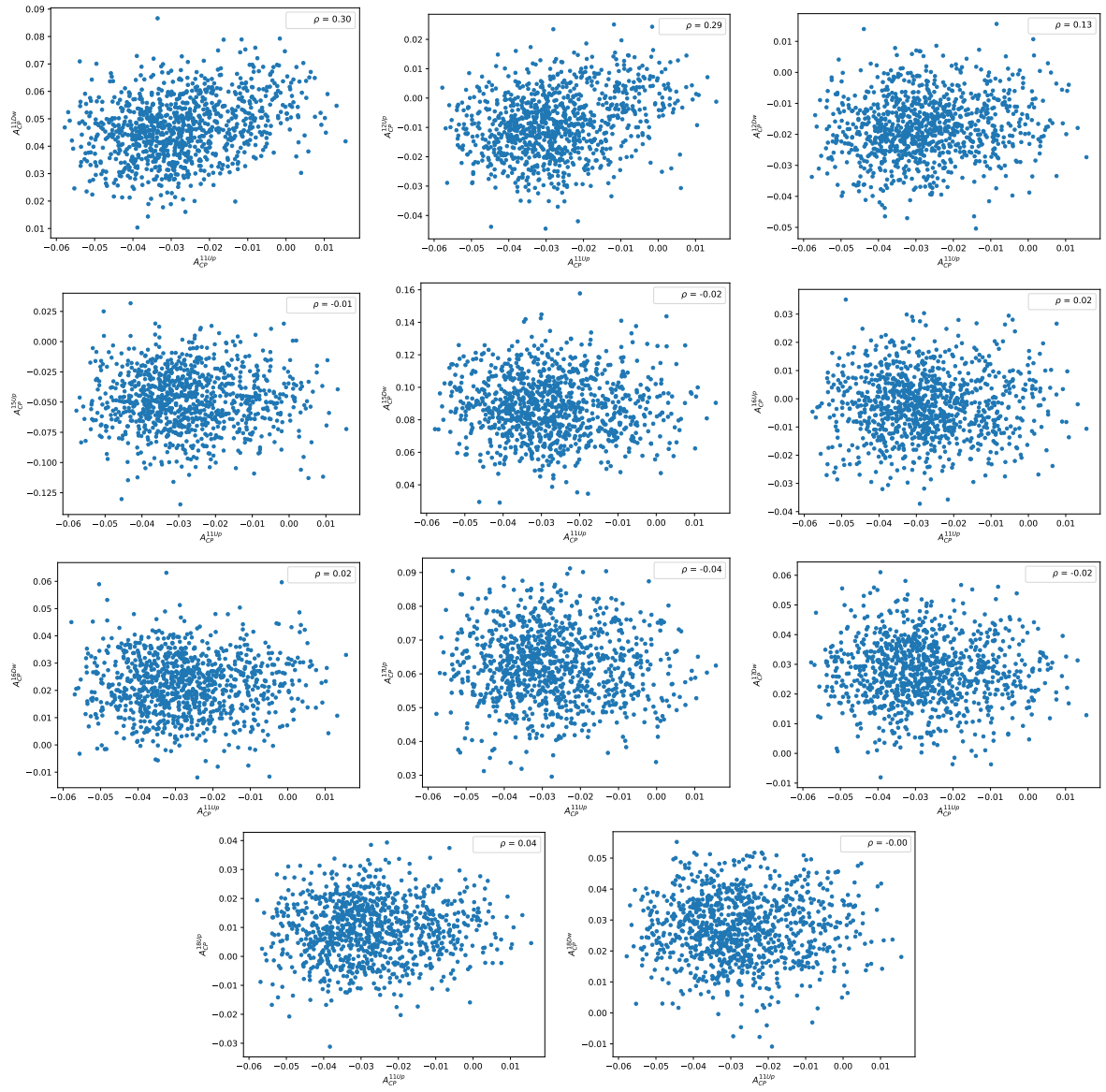

Figure E.11: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2011 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

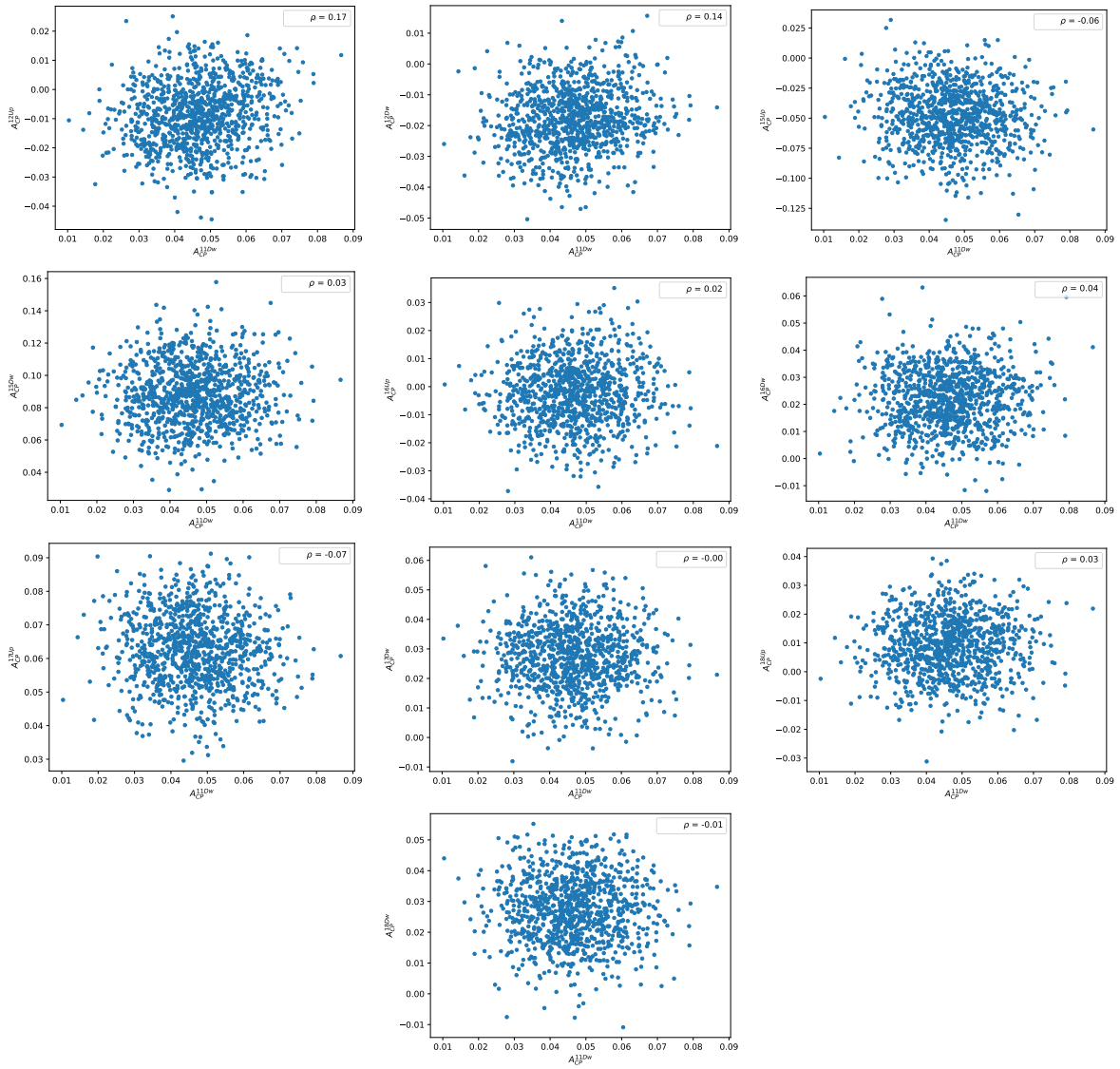

Figure E.12: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2011 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

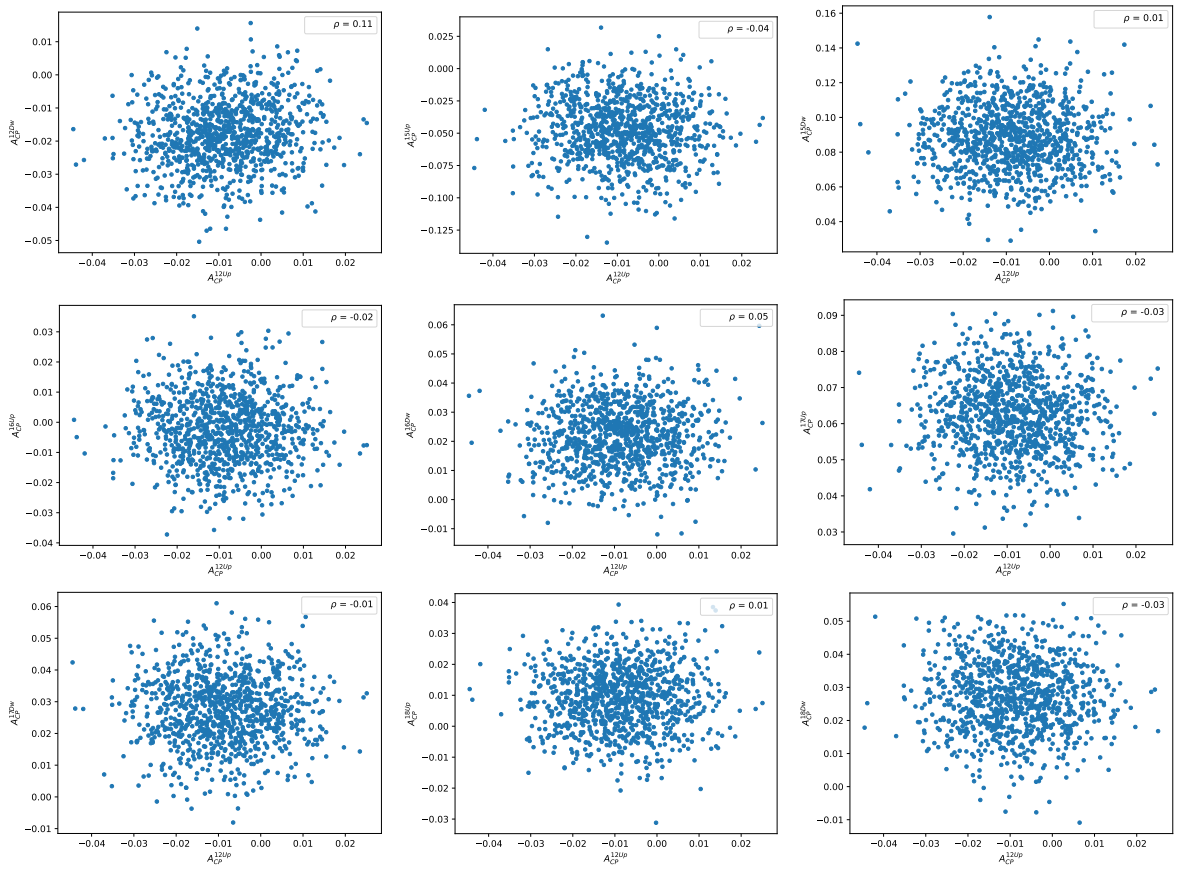

Figure E.13: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2012 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

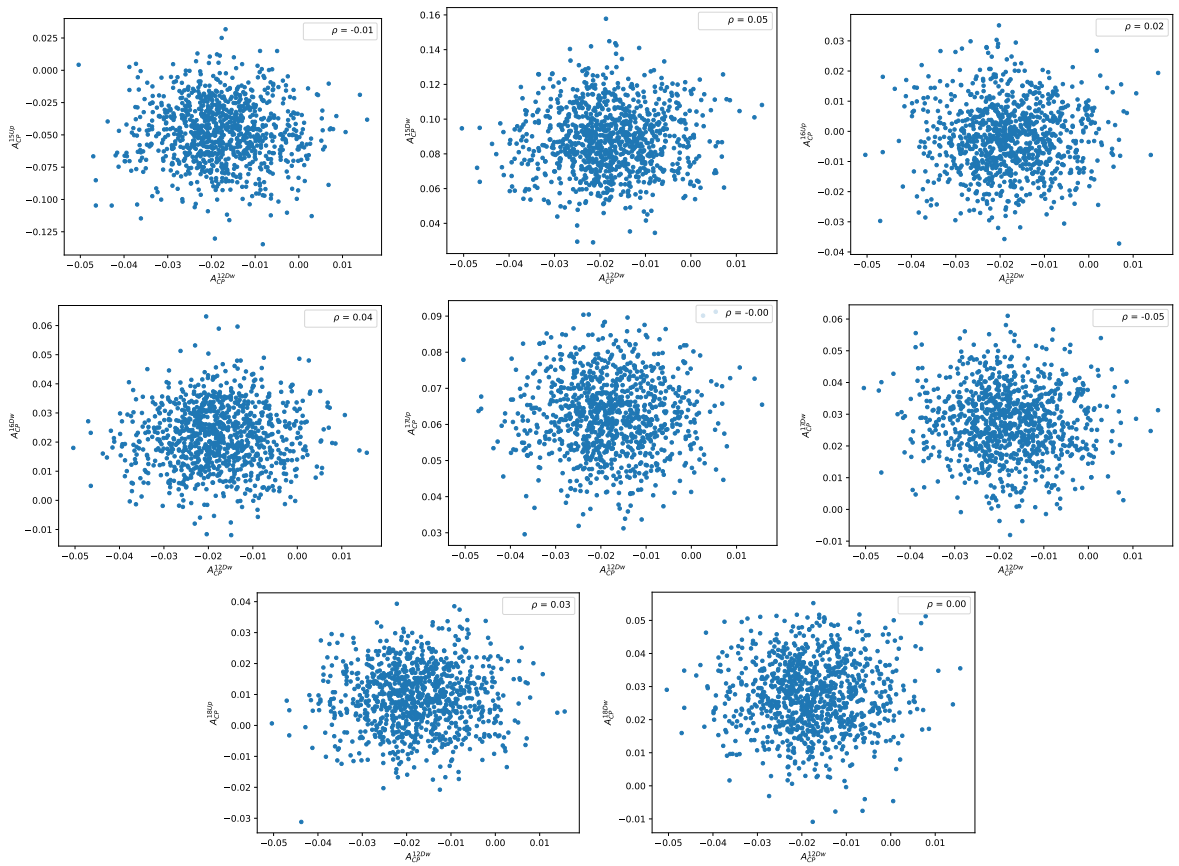

Figure E.14: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2012 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

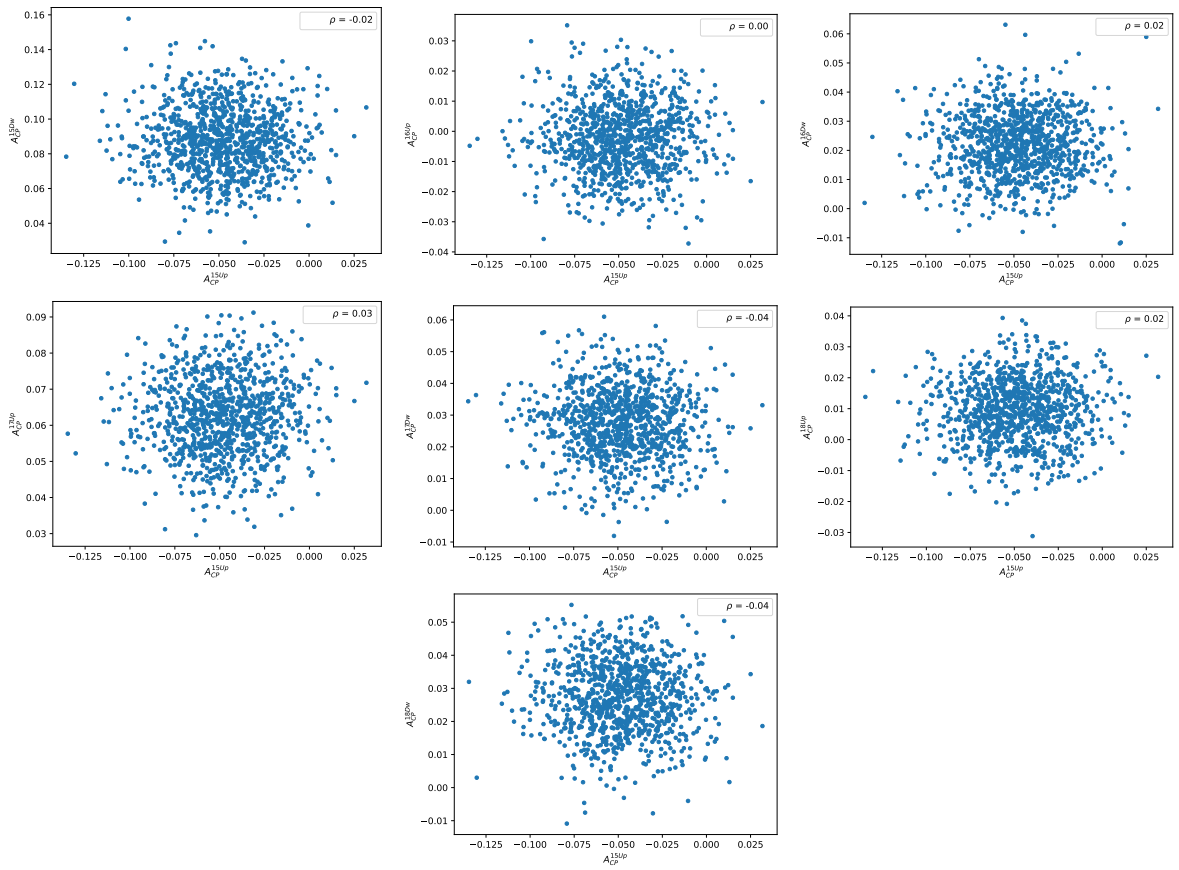

Figure E.15: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2015 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

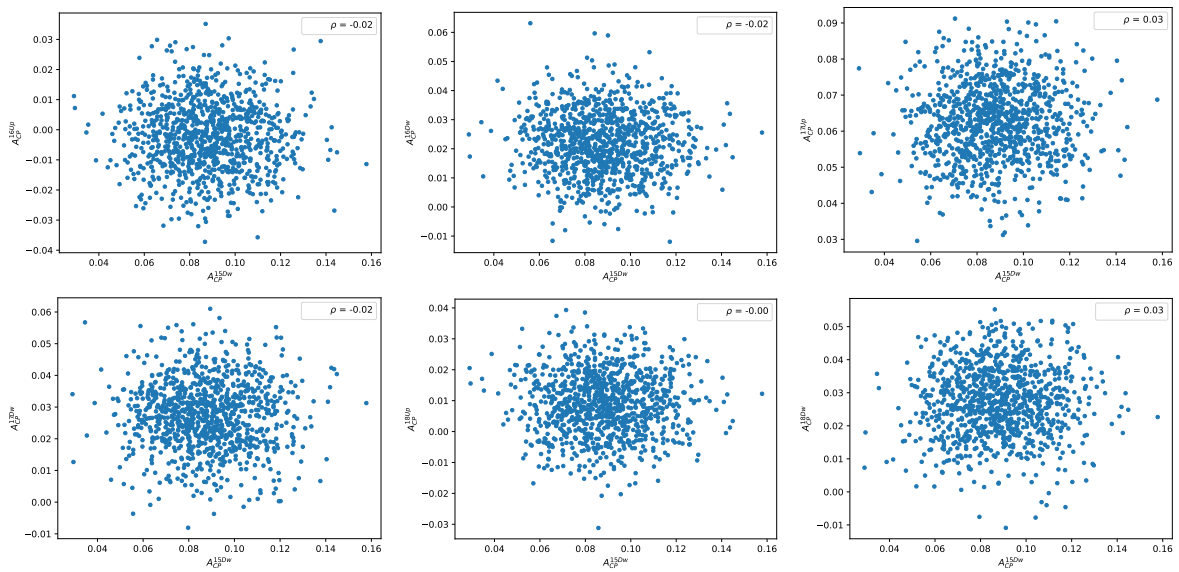

Figure E.16: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2015 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

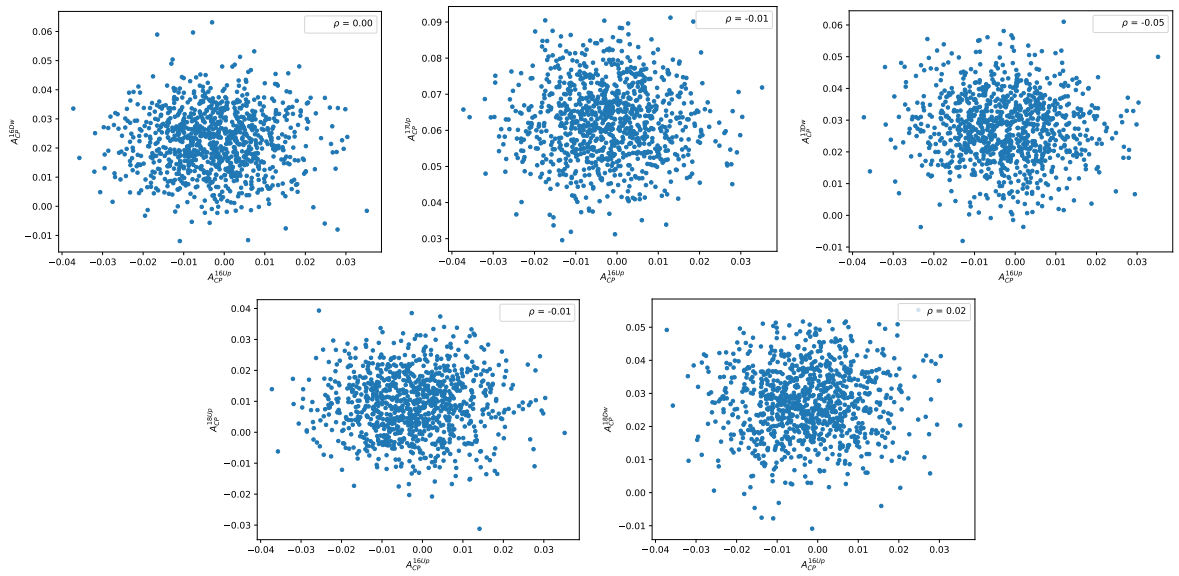

Figure E.17: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2016 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

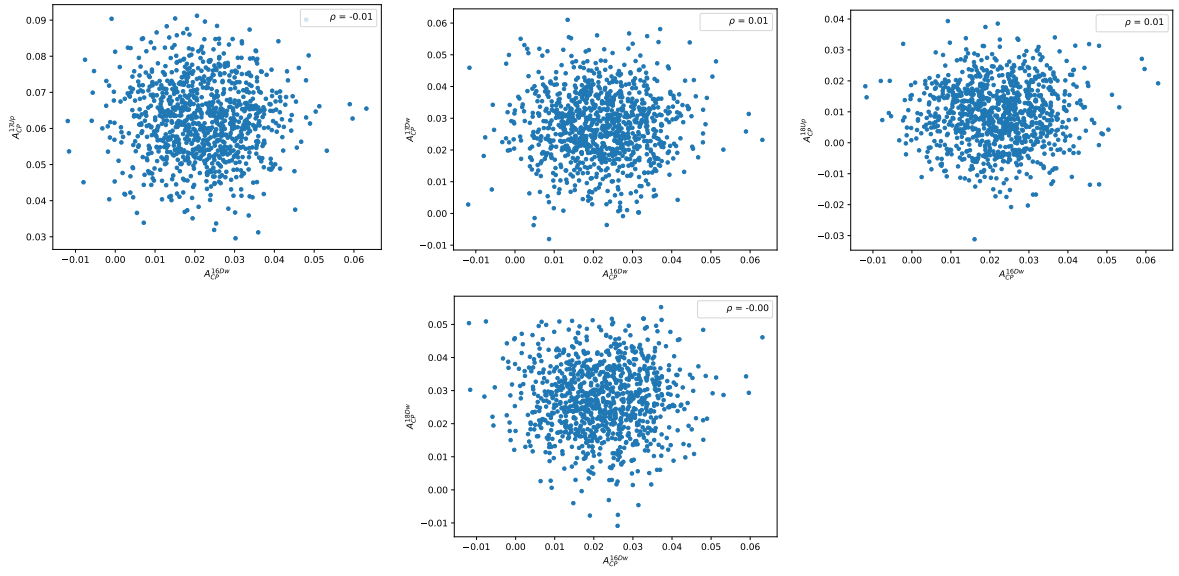

Figure E.18: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2016 magnet Down sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

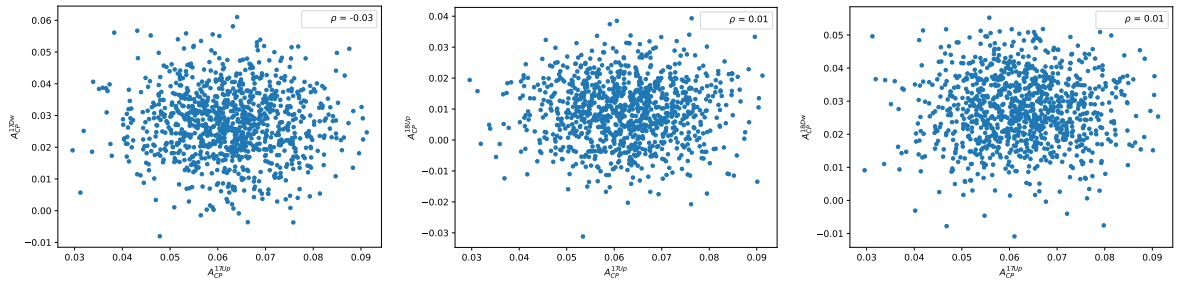

Figure E.19: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2017 magnet Up sample and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

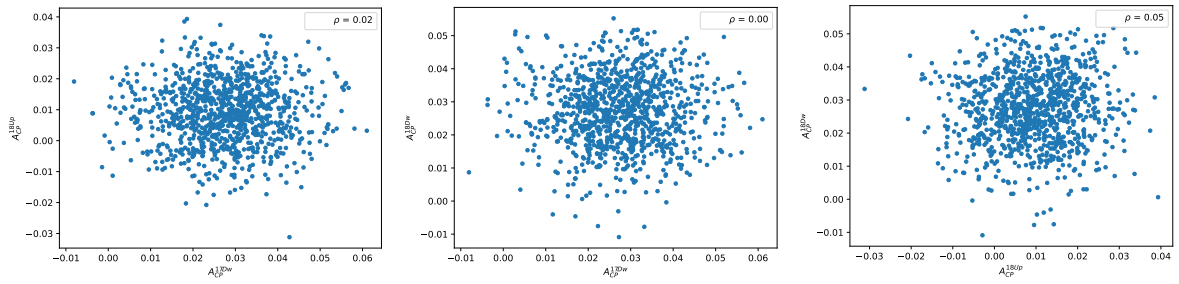

Figure E.20: Scatter plots of the 1000 values of $A_{CP}(p\pi^-)$ obtained from the pseudoexperiments in the 2017 magnet Down sample and 2018 magnet Up and the other subsamples. The correlation coefficient is shown in the upper right corner of each plot.

Bibliography

[1] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

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

[3] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[4] 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.

[5] 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.

[6] 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.

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

[8] LHCb collaboration, R. Aaij *et al.*, *Observation of several sources of CP violation in* $B^+ \to \pi^+\pi^+\pi^-$ *decays*, Phys. Rev. Lett. **124** (2020) 031801, arXiv:1909.05211.

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

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

[11] 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. **D102** (2020) 051101, arXiv:1912.10741.

[12] CDF collaboration, T. A. Aaltonen *et al.*, *Measurements of direct CP-violating asymmetries in charmless decays of bottom baryons*, Phys. Rev. Lett. **113** (2014) 242001, arXiv:1403.5586.

[13] LHCb collaboration, R. Aaij *et al.*, *Search for CP violation in* $\Lambda_b^0 \to pK^-$ *and* $\Lambda_b^0 \to p\pi^-$ *decays*, Phys. Lett. **B784** (2018) 101, arXiv:1807.06544.

[14] Particle Data Group, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[15] K. Nishijima, *Charge Independence Theory of V Particles*, Progress of Theoretical Physics **13** (1955) 285.

[16] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, Il Nuovo Cimento (1955-1965) **4** (1956) 848.

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

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

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

[20] M. E. Peskin and D. Schroeder, *An introduction to Quantum Field Theory*, Addison-Wesley, 1995.

[21] F. Halzen and A. Martin, *Quarks and Leptons: an Introductory Course in Modern Particle Physics*, John Wiley and Sons, 1984.

[22] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[23] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[24] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[25] M. Antonelli *et al.*, *Flavor Physics in the Quark Sector*, Phys. Rept. **494** (2010) 197, `arXiv:0907.5386`.

[26] J. C. Hardy and I. S. Towner, *Superallowed $0^+ \to 0^+$ nuclear β decays: 2020 critical survey, with implications for V_{ud} and CKM unitarity*, Phys. Rev. C **102** (2020) 045501.

[27] FlaviaNet Working Group on Kaon Decays, M. Antonelli *et al.*, *An Evaluation of $|V_{us}|$ and precise tests of the Standard Model from world data on leptonic and semileptonic kaon decays*, Eur. Phys. J. C **69** (2010) 399, `arXiv:1005.2323`.

[28] W. J. Marciano, *Precise determination of $|V_{us}|$ from lattice calculations of pseudoscalar decay constants*, Phys. Rev. Lett. **93** (2004) 231803.

[29] HFLAV collaboration, Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021*, Phys. Rev. D **107** (2023) 052008, `arXiv:2206.07501`.

[30] CHARM II collaboration, P. Vilain *et al.*, *Leading order QCD analysis of neutrino induced dimuon events*, Eur. Phys. J. C **11** (1999) 19.

[31] Flavour Lattice Averaging Group (FLAG), Y. Aoki *et al.*, *FLAG Review 2021*, Eur. Phys. J. C **82** (2022) 869, arXiv:2111.09849.

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

[33] LHCb collaboration, R. Aaij *et al.*, *Precise determination of the B_s^0-$\overline{B}_s^0$ oscillation frequency*, Nature Physics **18** (2022) 1, arXiv:2104.04421.

[34] M. Misiak *et al.*, *Updated NNLO QCD predictions for the weak radiative B-meson decays*, Phys. Rev. Lett. **114** (2015) 221801, arXiv:1503.01789.

[35] M. Czakon *et al.*, *The $(Q_7, Q_{1,2})$ contribution to $\overline{B} \to X_s\gamma$ at $\mathcal{O}(\alpha_s^2)$*, JHEP **04** (2015) 168, arXiv:1503.01791.

[36] CDF collaboration, D. Acosta *et al.*, *Measurement of $B(t \to Wb)/B(t \to Wq)$ at the Collider Detector at Fermilab*, Phys. Rev. Lett. **95** (2005) 102002, arXiv:hep-ex/0505091.

[37] DØ collaboration, V. M. Abazov *et al.*, *Precision measurement of the ratio* $\mathrm{B}(t \to Wb)/\mathrm{B}(t \to Wq)$ *and Extraction of V_{tb}*, Phys. Rev. Lett. **107** (2011) 121802, arXiv:1106.5436.

[38] ATLAS, CMS collaboration, M. Aaboud *et al.*, *Combinations of single-top-quark production cross-section measurements and $|f_{LV}V_{tb}|$ determinations at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments*, JHEP **05** (2019) 088, arXiv:1902.07158.

[39] I. I. Bigi and A. I. Sanda, *CP Violation*, Cambridge University Press, 2 ed., 2009.

[40] L. Wolfenstein, *Parametrization of the Kobayashi–Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[41] UTfit collaboration, M. Bona *et al.*, *New UTfit Analysis of the Unitarity Triangle in the Cabibbo-Kobayashi-Maskawa scheme*, Rend. Lincei Sci. Fis. Nat. **34** (2023) 37, arXiv:2212.03894.

[42] G. C. Branco, L. Lavoura, and J. P. Silva, *CP violation*, no. 103 in International Series of Monographies on Physics, Oxford University Press, 1999.

[43] F. L. Wilson, *Fermi's theory of beta decay*, American Journal of Physics **36** (1968) 1150.

[44] K. G. Wilson, *Non-lagrangian models of current algebra*, Phys. Rev. **179** (1969) 1499.

[45] K. G. Wilson, *Renormalization group and strong interactions*, Phys. Rev. D **3** (1971) 1818.

[46] E. Witten, *Short distance analysis of weak interactions*, Nuclear Physics B **122** (1977) 109.

[47] G. Altarelli and L. Maiani, *Octet enhancement of non-leptonic weak interactions in asymptotically free gauge theories*, Physics Letters B **52** (1974) 351.

[48] A. J. Buras and L. Silvestrini, *Non-leptonic two-body b decays beyond factorization*, Nuclear Physics B **569** (2000) 3.

[49] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125, `arXiv:hep-ph/9512380`.

[50] M. Bauer, B. Stech, and M. Wirbel, *Exclusive Nonleptonic Decays of D, D_s, and B Mesons*, Z. Phys. C **34** (1987) 103.

[51] M. J. Dugan and B. Grinstein, *QCD basis for factorization in decays of heavy mesons*, Phys. Lett. B **255** (1991) 583.

[52] H. D. Politzer and M. B. Wise, *Perturbative corrections to factorization in anti-B decay*, Phys. Lett. B **257** (1991) 399.

[53] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *QCD factorization for $B \to \pi\pi$ decays: Strong phases and CP violation in the heavy quark limit*, Phys. Rev. Lett. **83** (1999) 1914, `arXiv:hep-ph/9905312`.

[54] H.-n. Li and H.-L. Yu, *Perturbative QCD analysis of B meson decays*, Phys. Rev. D **53** (1996) 2480, `arXiv:hep-ph/9411308`.

[55] Y.-Y. Keum, H.-n. Li, and A. I. Sanda, *Fat penguins and imaginary penguins in perturbative QCD*, Phys. Lett. B **504** (2001) 6, `arXiv:hep-ph/0004004`.

[56] C.-H. Chen, Y.-Y. Keum, and H.-n. Li, *Perturbative QCD analysis of $B \to \phi K^*$ decays*, Phys. Rev. D **66** (2002) 054013, `arXiv:hep-ph/0204166`.

[57] C. W. Bauer and I. W. Stewart, *Invariant operators in collinear effective theory*, Phys. Lett. B **516** (2001) 134, `arXiv:hep-ph/0107001`.

[58] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *An Effective field theory for collinear and soft gluons: Heavy to light decays*, Phys. Rev. D **63** (2001) 114020, `arXiv:hep-ph/0011336`.

[59] C. W. Bauer, I. Z. Rothstein, and I. W. Stewart, *SCET analysis of $B \to K\pi$, $B \to K\overline{K}$, and $B \to \pi\pi$ decays*, Phys. Rev. D **74** (2006) 034010, `arXiv:hep-ph/0510241`.

[60] A. R. Williamson and J. Zupan, *Two body B decays with isosinglet final states in SCET*, Phys. Rev. D **74** (2006) 014003, `arXiv:hep-ph/0601214`.

[61] V. M. Braun and I. E. Filyanov, *QCD Sum Rules in Exclusive Kinematics and Pion Wave Function*, Z. Phys. C **44** (1989) 157.

[62] A. Khodjamirian, *$B \to \pi\pi$ decay in QCD*, Nucl. Phys. B **605** (2001) 558, `arXiv:hep-ph/0012271`.

[63] A. Khodjamirian, T. Mannel, and P. Urban, *Gluonic penguins in $B \to \pi\pi$ from QCD light cone sum rules*, Phys. Rev. D **67** (2003) 054027, `arXiv:hep-ph/0210378`.

[64] BABAR Collaboration, B. Aubert *et al.*, *Measurement of branching fractions and search for CP-violating charge asymmetries in charmless two-body B decays into pions and kaons*, Phys. Rev. Lett. **87** (2001) 151802.

[65] BaBar collaboration, J. P. Lees *et al.*, *Measurement of CP Asymmetries and Branching Fractions in Charmless Two-Body B-Meson Decays to Pions and Kaons*, Phys. Rev. D **87** (2013) 052009, arXiv:1206.3525.

[66] CDF collaboration, A. Abulencia *et al.*, *Observation of B^0 (s) $\to K^+K^-$ and Measurements of Branching Fractions of Charmless Two-body Decays of B^0 and B_s^0 Mesons in $\bar{p}p$ Collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Rev. Lett. **97** (2006) 211802, arXiv:hep-ex/0607021.

[67] CDF collaboration, T. Aaltonen *et al.*, *Observation of New Charmless Decays of Bottom Hadrons*, Phys. Rev. Lett. **103** (2009) 031801, arXiv:0812.4271.

[68] CDF collaboration, T. Aaltonen *et al.*, *Measurements of Direct CP Violating Asymmetries in Charmless Decays of Strange Bottom Mesons and Bottom Baryons*, Phys. Rev. Lett. **106** (2011) 181802, arXiv:1103.5762.

[69] CDF collaboration, T. Aaltonen *et al.*, *Evidence for the charmless annihilation decay mode $B_s^0 \to \pi^+\pi^-$*, Phys. Rev. Lett. **108** (2012) 211803, arXiv:1111.0485.

[70] CDF collaboration, T. A. Aaltonen *et al.*, *Measurements of Direct CP-Violating Asymmetries in Charmless Decays of Bottom Baryons*, Phys. Rev. Lett. **113** (2014) 242001, arXiv:1403.5586.

[71] Belle collaboration, K. Abe *et al.*, *Measurement of branching fractions for $B \to \pi\pi$, $K\pi$ and KK decays*, Phys. Rev. Lett. **87** (2001) 101801, arXiv:hep-ex/0104030.

[72] Belle collaboration, K. Abe *et al.*, *Study of CP violating asymmetries in $B^0 \to \pi^+\pi^-$ decays*, Phys. Rev. Lett. **89** (2002) 071801, arXiv:hep-ex/0204002.

[73] Belle collaboration, K. Abe *et al.*, *Evidence for CP violating asymmetries $B^0 \to \pi^+\pi^-$ decays and constraints on the CKM angle ϕ_2*, Phys. Rev. D **68** (2003) 012001, arXiv:hep-ex/0301032.

[74] Belle collaboration, Y. Chao *et al.*, *Improved measurements of partial rate asymmetry in $B \to hh$ decays*, Phys. Rev. D **71** (2005) 031502, arXiv:hep-ex/0407025.

[75] LHCb collaboration, R. Aaij *et al.*, *First evidence of direct CP violation in charmless two-body decays of B_s^0 mesons*, Phys. Rev. Lett. **108** (2012) 201601, arXiv:1202.6251.

[76] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron branching fractions for two-body decays into charmless charged hadrons*, JHEP **10** (2012) 037, arXiv:1206.2794.

[77] LHCb collaboration, R. Aaij *et al.*, *First measurement of time-dependent CP violation in $B_s^0 \to K^+K^-$ decays*, JHEP **10** (2013) 183, arXiv:1308.1428.

[78] LHCb collaboration, R. Aaij *et al.*, *Observation of the annihilation decay mode* $B^0 \to K^+K^-$, Phys. Rev. Lett. **118** (2017) 081801, arXiv:1610.08288.

[79] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetries in two-body* $B^0_{(s)}$*-meson decays to charged pions and kaons*, Phys. Rev. **D98** (2018) 032004, arXiv:1805.06759.

[80] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in two-body* B^0_s*-meson decays to charged pions and kaons*, JHEP **03** (2021) 075, arXiv:2012.05319.

[81] M. Gronau, O. F. Hernández, D. London, and J. L. Rosner, *Decays of b mesons to two pseudoscalars in broken* SU(3) *symmetry*, Phys. Rev. D **52** (1995) 6356.

[82] M. Gronau, O. F. Hernández, D. London, and J. L. Rosner, *Decays of b mesons to two light pseudoscalars*, Phys. Rev. D **50** (1994) 4529.

[83] Y. K. Hsiao, Y. Yao, and C. Q. Geng, *Charmless two-body antitriplet b-baryon decays*, Phys. Rev. D **95** (2017) 093001.

[84] J. Zhu, Z.-T. Wei, and H.-W. Ke, *Semileptonic and nonleptonic weak decays of* Λ_b^0, Phys. Rev. D **99** (2019) 054020.

[85] C.-D. Lü *et al.*, *Anatomy of the perturbative qcd approach to the baryonic decays* $\Lambda_b \to p\pi$*,* pk, Phys. Rev. D **80** (2009) 034011.

[86] C.-Q. Geng, C.-W. Liu, and T.-H. Tsai, *Nonleptonic two-body weak decays of* Λ_b *in a modified mit bag model*, Phys. Rev. D **102** (2020) 034033.

[87] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.

[88] E. Mobs, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, , General Photo.

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

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

[91] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.

[92] LHCb collaboration, R. Aaij *et al.*, *Prompt charm production in pp collisions at* $\sqrt{s} = 7\,TeV$, Nucl. Phys. **B871** (2013) 1, arXiv:1302.2864.

[93] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at* $\sqrt{s} = 13\,TeV$, JHEP **03** (2016) 159, Erratum ibid. **09** (2016) 013, Erratum ibid. **05** (2017) 074, arXiv:1510.01707.

[94] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, in *ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders*, 183–187, 2014, arXiv:1410.3667.

[95] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011, 2001.

[96] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[97] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, 2002.

[98] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, 2001.

[99] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007, 2000.

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

[101] R. Frühwirth, *Application of kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) 444.

[102] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037, 2000.

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

[104] C. Abellan Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, arXiv:2008.11556, submitted to JINST.

[105] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010, 2001.

[106] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[107] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[108] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031, 2003.

[109] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[110] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, arXiv:1812.10790.

[111] LHCb collaboration, R. Aaij *et al.*, *Absolute luminosity measurements with the LHCb detector at the LHC*, JINST **7** (2012) P01010, arXiv:1110.2866.

[112] LHCb collaboration, R. Aaij *et al.*, *Observation of a $\Lambda_b^0 - \overline{\Lambda}_b^0$ production asymmetry in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV*, JHEP **10** (2021) 060, arXiv:2107.09593.

[113] LHCb collaboration, R. Aaij *et al.*, *Measurement of B^0, B_s^0, B^+ and Λ_b^0 production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.

[114] L. Dufour, *High-precision measurements of charge asymmetries at LHCb*, 2019. doi: https://research.rug.nl/en/publications/high-precision-measurements-of-charge-asymmetries-at-lhcb.

[115] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in B_s^0–$\overline{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, arXiv:1605.09768.

[116] LHCb collaboration, R. Aaij *et al.*, *Observation of the mass difference between neutral charm-meson eigenstates*, Phys. Rev. Lett. **127** (2021) 111801, arXiv:2106.03744.

[117] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[118] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

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

[120] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and Physics Documentation*, Comput. Phys. Commun. **199** (2016) 86, arXiv:1011.0937.

[121] J. Allison, K. Amako, J. Apostolakis *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[122] S. Agostinelli, J. Allison, K. Amako *et al.*, *Geant4—a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

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

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

[125] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2019) 1, arXiv:1803.00824.

[126] M. Pivk and F. R. L. Diberder, *splot: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (2005) 356.

[127] K. Cranmer, *Kernel estimation in high-energy physics*, Computer Physics Communications **136** (2001) 198.

[128] LHCb collaboration, R. Aaij *et al.*, *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. **D85** (2012) 032008, arXiv:1111.2357.

[129] ARGUS collaboration, H. Albrecht, R. Gläser, G. Harder *et al.*, *Search for hadronic $b \to u$ decays*, Physics Letters B **241** (1990) 278.

[130] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949) 149.

[131] G. A. Cowan, D. C. Craik, and M. D. Needham, *Rapidsim: An application for the fast simulation of heavy-quark hadron decays*, Computer Physics Communications **214** (2017) 239.

[132] LHCb collaboration, R. Aaij *et al.*, *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B^0_s decay branching fractions*, Phys. Rev. **D104** (2021) 032005, arXiv:2103.06810.

[133] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-integrated CP asymmetry in $D^0 \to K^-K^+$ decays*, arXiv:2209.03179, to appear in Phys. Rev. Lett.

[134] A. Rogozhnikov *et al.*, *arogozhnikov/hep_ml: hep_ml v0.7.2*, 2023. doi: 10.5281/zenodo.7717895.

[135] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nuclear Instruments and Methods in Physics Research **764** (2014) 150.

[136] P. V. R. Murthy *et al.*, *Neutron total cross sections on nuclei at Fermilab energies*, Nuclear Physics B **92** (1975) 269.

[137] J. H. Atkinson, W. N. Hess, V. Perez-Mendez, and R. Wallace, *5-Bev neutron cross sections in hydrogen and other elements*, Phys. Rev. **123** (1961) 1850.

[138] G. Bellettini *et al.*, *Proton-nuclei cross sections at 20 GeV*, Nuclear Physics **79** (1966) 609.

[139] LHCb collaboration, R. Aaij *et al.*, *Improved measurement of CP violation parameters in $B^0_s \to J/\psi K^+K^-$ decays in the vicinty of the $\phi(1020)$ resonance*, arXiv:2308.01468, submitted to Phys. Rev. Lett.

[140] LHCb collaboration, R. Aaij *et al.*, *First observation and branching fraction measurement of the $\Lambda^0_b \to D^-_s p$ decay*, JHEP **07** (2023) 075, arXiv:2212.12574.

[141] M. Schmelling, *Averaging correlated data*, Physica Scripta **51** (1995) 676.