LHCb-ANA-2018-021
September 17, 2018
v1.0

Measurement of the CKM angle γ using $B_s^0 \to D_s K\pi\pi$ decays

P. d'Argent[1], E. Gersabeck[2], M. Kecke[1], M. Schiller[3]

[1] *Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany*
[2] *School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom*
[3] *School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom*

Abstract

We present the first measurement of the weak phase $\gamma - 2\beta_s$ obtained from a time-dependent (amplitude) analysis of $B_s^0 \to D_s K\pi\pi$ decays using proton-proton collision data corresponding to an integrated luminosity of $7\,\mathrm{fb}^{-1}$ recorded by the LHCb detector.

Contents

1 Introduction

This note presents the first measurement of the CKM angle $\gamma \equiv \arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$ using $B_s^0 \to D_s K\pi\pi$ decays, where the $K\pi\pi$ subsystem is dominated by excited kaon states such as the $K_1(1270)$ and $K_1(1400)$ resonances [1,2]. In these decays, sensitivity to the weak phase results from the interference between $b \to c$ and $b \to u$ transitions achieved through $B_s^0 - \overline{B}_s^0$ mixing [3,4]. The amplitudes for both processes are of the same order in the Wolfenstein parameters λ, $\mathcal{O}(\lambda^3)$, so that interference effects are expected to be large. The corresponding Feynman diagrams are shown in Fig. 1.1. Due to the interference between mixing and decay amplitudes, the physical *CP* violating observables in these decays are functions of a combination of γ and the mixing phase β_s, namely $\gamma - 2\beta_s$. To account for the non-constant strong phase across the phase space, one can either perform a time-dependent amplitude fit or select a suitable phase-space region and introduce a coherence factor as additional hadronic parameter to the decay-time fit.

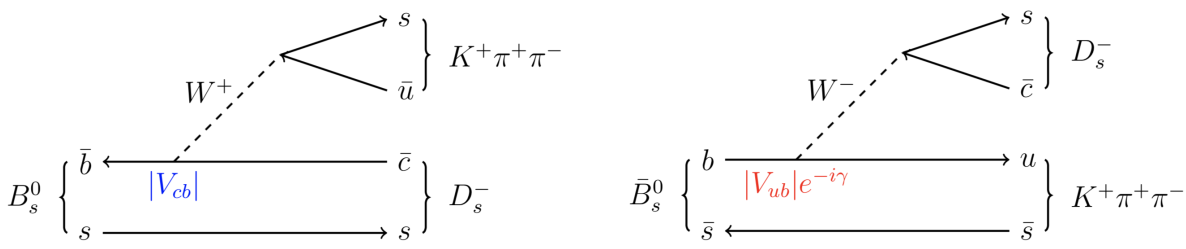

Figure 1.1: Feynman diagram for $B_s^0/\bar{B}_s^0 \to D_s^- K^+\pi^+\pi^-$ decays.

2 Formalism

2.1 Decay rates and *CP*-observables

The differential decay rate of B_s^0 or $\bar{B}_s^0$ decays to the final state $D_s^- K^+ \pi\pi$ or $D_s^+ K^- \pi\pi$ is given by:

$$\begin{aligned}\frac{\mathrm{d}\Gamma(\mathbf{x}, t, q, f)}{e^{-\Gamma_s t}\,\mathrm{d}t\,\mathrm{d}\Phi_4} \propto &\left(|\mathcal{A}_f^c(\mathbf{x})|^2 + |\mathcal{A}_f^u(\mathbf{x})|^2\right)\,\cosh\left(\frac{\Delta\Gamma_s\,t}{2}\right) \\ &+ q\,f\,\left(|\mathcal{A}_f^c(\mathbf{x})|^2 - |\mathcal{A}_f^u(\mathbf{x})|^2\right)\,\cos\left(\Delta m_s\,t\right) \\ &- 2\mathrm{Re}\left(\mathcal{A}_f^c(\mathbf{x})^*\,\mathcal{A}_f^u(\mathbf{x})\,e^{-if(\gamma-2\beta_s)}\right)\,\sinh\left(\frac{\Delta\Gamma_s\,t}{2}\right) \\ &- 2\,q\,f\,\mathrm{Im}\left(\mathcal{A}_f^c(\mathbf{x})^*\,\mathcal{A}_f^u(\mathbf{x})\,e^{-if(\gamma-2\beta_s)}\right)\,\sin\left(\Delta m_s\,t\right)\end{aligned} \tag{2.1}$$

where $q = +1$ (-1) refers to an initially produced B_s^0 ($\bar{B}_s^0$) flavour eigenstate, $q = 0$ to an undetermined initial flavour, $f = +1$ or -1 denotes $D_s^- K^+\pi\pi$ or $D_s^+ K^-\pi\pi$ final states and Γ_s, $\Delta\Gamma_s$ and Δm_s are the width average, the width difference and the mass difference of the two B_s mass eigenstates. We choose a convention in which $\Delta\Gamma_s < 0$ and $\Delta m_s > 0$. We further assume $|q/p| = 1$ for the complex coefficients p and q which relate the B_s meson mass eigenstates to the flavour eigenstates (no *CP* violation in the mixing). The CKM angle γ can be extracted from the *CP* violating phase associated to the interference between mixing and decay, $\gamma - 2\beta_s$, since the $B_s^0 - \bar{B}_s^0$ mixing phase, β_s, is well constrained from $B_s \to J/\psi\,\phi$ and related modes.

The static total decay amplitudes $\mathcal{A}_f^c(\mathbf{x})$ and $\mathcal{A}_f^c(\mathbf{x})$ are given by the coherent sum over all intermediate state amplitudes $A_i(\mathbf{x})$, each weighted by a complex coefficient to be determined from data,

$$\mathcal{A}(B_s^0 \to D_s^- K^+\pi\pi) \equiv \mathcal{A}_f^c(\mathbf{x}) = \sum_i a_i^c\,A_i(\mathbf{x}) \tag{2.2}$$

$$\mathcal{A}(\bar{B}_s^0 \to D_s^- K^+\pi\pi) \equiv \mathcal{A}_f^u(\mathbf{x}) = \sum_i a_i^u\,A_i(\mathbf{x}) \tag{2.3}$$

where the superscript c (u) indicates a $b \to c$ ($b \to u$) quark-level transition and $\mathbf{x}$ represents a unique set of kinematic conditions within the five-dimensional phase space of the decay. Convenient choices for the kinematic observables include the invariant mass combinations of the final state particles or acoplanarity and helicity angles. In practice, we do not need to choose a particular five-dimensional basis, but use the full four-vectors of the decay in our analysis. The dimensionality is handled by the phase space element which can be written in terms of any set of five independent kinematic observables, $\mathbf{x} = (x_1, \ldots, x_5)$, as

$$\mathrm{d}\Phi_4 = \phi_4(\mathbf{x})\,\mathrm{d}^5 x, \tag{2.4}$$

where $\phi_4(\mathbf{x}) = \left|\frac{\partial\Phi_4}{\partial(x_1,\ldots x_5)}\right|$ is the phase space density. In contrast to three-body decays, the four-body phase space density function is not flat in the usual kinematic variables. Therefore, an analytic expression for ϕ_4 is taken from Ref. [5].

Assuming there is no direct CP violation in the B_s decay implies for the CP conjugate transition amplitudes:

$$\mathcal{A}(\bar{B}_s^0 \to \bar{f}) = \mathcal{A}_{\bar{f}}^c(\mathbf{x}) = \mathcal{A}_f^c(\overline{\mathbf{x}}) \tag{2.5}$$

$$\mathcal{A}(B_s^0 \to \bar{f}) = \mathcal{A}_{\bar{f}}^u(\mathbf{x}) = \mathcal{A}_f^u(\overline{\mathbf{x}}) \tag{2.6}$$

where the CP-conjugate phase space point $\overline{\mathbf{x}}$ is defined such that it is mapped onto $\mathbf{x}$ by the interchange of final state charges, and the reversal of three-momenta.

The phenomenological description of the intermediate state amplitudes is discussed in Sec. 2.2. For a model-independent measurement, the differential decay rate can be integrated over the phase space:

$$\begin{aligned}\int \frac{\mathrm{d}\Gamma(x,t,q,f)}{e^{-\Gamma_s t}\,\mathrm{d}t\,\mathrm{d}\Phi_4}\,\mathrm{d}\Phi_4 \propto\ & \cosh\left(\frac{\Delta\Gamma_s\, t}{2}\right) + q\, f\, C\cos\left(\Delta m_s\, t\right) \\ & + D_f \sinh\left(\frac{\Delta\Gamma_s\, t}{2}\right) - q\, S_f \sin\left(\Delta m_s\, t\right)\end{aligned} \tag{2.7}$$

where the same convention for the CP coefficients as for the $B_s \to D_s K$ analysis is used:

$$C = \frac{1-r^2}{1+r^2} \tag{2.8}$$

$$D_f = -\ \frac{2\, r\, \kappa \cos\left(\delta - q_f\left(\gamma - 2\beta_s\right)\right)}{1+r^2} \tag{2.9}$$

$$S_f = f\ \frac{2\, r\, \kappa \sin\left(\delta - q_f\left(\gamma - 2\beta_s\right)\right)}{1+r^2} \tag{2.10}$$

The coherence factor κ, the strong phase difference δ and the ratio of the suppressed $(b \to u)$ over favored $(b \to c)$ decay mode, averaged over the phase space, are defined as:

$$\kappa\, e^{i\delta} \equiv \frac{\int \mathcal{A}_f^c(x)^*\, \mathcal{A}_f^u(x)\,\mathrm{d}\Phi_4}{\sqrt{\int |\mathcal{A}_f^c(x)|^2\,\mathrm{d}\Phi_4}\sqrt{\int |\mathcal{A}_f^u(x)|^2\,\mathrm{d}\Phi_4}} \tag{2.11}$$

$$r \equiv \frac{\sqrt{\int |\mathcal{A}_f^u(x)|^2\,\mathrm{d}\Phi_4}}{\sqrt{\int |\mathcal{A}_f^c(x)|^2\,\mathrm{d}\Phi_4}}. \tag{2.12}$$

The coherence factor dilutes the sensitivity to the weak phase γ due to the integration over the interfering amplitudes across the phase space. The value of κ is bounded between zero and unity. The latter corresponds to the limit of only one contributing intermediate state in which case the same sensitivity as in $B_s \to D_s K$ decays is reached, while $\kappa = 0$ would result in no sensitivity to γ at all.

2.2 Amplitude model

To construct the intermediate state amplitudes $A_i(\mathbf{x})$, the isobar approach is used, which assumes that the decay process can be factorized into subsequent two-body decay amplitudes [6–8]. This gives rise to two different decay topologies; quasi two-body decays $B_s \to (R_1 \to h_1\, h_2)\,(R_2 \to h_3\, h_4)$ or cascade decays $B_s \to h_1\, [R_1 \to h_2\,(R_2 \to h_3\, h_4)]$. In either case, the intermediate state amplitude is parameterized as a product of orbital angular momentum, L, dependent form factors B_L, included for each vertex of the decay tree, Breit-Wigner propagators T_R, included for each resonance R, and an overall angular distribution represented by a spin factor S,

$$A_i(\mathbf{x}) = B_{L_{B_s}}(\mathbf{x})\,[B_{L_{R_1}}(\mathbf{x})\,T_{R_1}(\mathbf{x})]\,[B_{L_{R_2}}(\mathbf{x})\,T_{R_2}(\mathbf{x})]\,S_i(\mathbf{x})\,. \tag{2.13}$$

The following description of the individual components is adapted from Ref. [9] and only included for completeness.

2.2.1 Form Factors and Resonance Lineshapes

To account for the finite size of the decaying resonances, the Blatt-Weisskopf penetration factors, derived in Ref. [10] by assuming a square well interaction potential with radius r_{BW}, are used as form factors, B_L. They depend on the breakup momentum q, and the orbital angular momentum L, between the resonance daughters. Their explicit expressions are

$$\begin{aligned} B_0(q) &= 1, \\ B_1(q) &= 1/\sqrt{1+(q\,r_{\mathrm{BW}})^2}, \\ B_2(q) &= 1/\sqrt{9+3\,(q\,r_{\mathrm{BW}})^2+(q\,r_{\mathrm{BW}})^4}. \end{aligned} \tag{2.14}$$

Resonance lineshapes are described as function of the energy-squared, s, by Breit-Wigner propagators

$$T(s) = \frac{1}{m_0^2 - s - i\,m_0\,\Gamma(s)}\,, \tag{2.15}$$

where the total width, $\Gamma(s)$, is normalized to give the nominal width, Γ_0, when evaluated at the nominal mass m_0.

For a decay into two stable particles $R \to AB$, the energy dependence of the decay width can be described by

$$\Gamma^{(2)}_{R\to AB}(s) = \Gamma_0\,\frac{m_0}{\sqrt{s}}\left(\frac{q}{q_0}\right)^{2L+1}\frac{B_L(q)^2}{B_L(q_0)^2}\,, \tag{2.16}$$

where q_0 is the value of the breakup momentum at the resonance pole [11].

The energy-dependent width for a three-body decay $R \to ABC$, on the other hand, is considerably more complicated and has no analytic expression in general. However, it can be obtained numerically by integrating the transition amplitude-squared over the phase space,

$$\Gamma^{(3)}_{R\to ABC}(s) = \frac{1}{2\sqrt{s}}\int |A_{R\to ABC}|^2\,\mathrm{d}\Phi_3, \tag{2.17}$$

and therefore requires knowledge of the resonant substructure. The three-body amplitude $A_{R\to ABC}$ can be parameterized similarly to the four-body amplitude in Eq. (2.13). In particular, it includes form factors and propagators of intermediate two-body resonances.

Both Eq. (2.16) and Eq. (2.17) give only the partial width for the decay into a specific channel. To obtain the total width, a sum over all possible decay channels has to be performed,

$$\Gamma(s) = \sum_i g_i \, \Gamma_i(s), \tag{2.18}$$

where the coupling strength to channel i, is given by g_i.

The treatment of the lineshape for various resonances considered in this analysis is described in what follows. The nominal masses and widths of the resonances are taken from the PDG [12] with the exceptions described below.

For the broad scalar resonance σ, the model from Bugg is used [13]. We use the Gounaris-Sakurai parametrization for the $\rho(770)^0 \to \pi\pi$ propagator [14]. For the decay chain $K_1(1270) \to \rho(770)K$, we include $\rho - \omega$ mixing with the relative magnitude and phase between ρ and ω fixed to the values determined in Ref. [15]. The energy-dependent width of the $f_0(980)$ resonance is given by the sum of the partial widths into the $\pi\pi$ and KK channels [16], where the coupling constants as well as the mass and width are taken from a measurement performed by the BES Collaboration [17]. For the $f_2(1270)$ and the $f_0(1370)$ mesons we use the total decay widths calculated in Ref. [9] which take the channels $\pi\pi, KK, \eta\eta$ and $\pi\pi\pi\pi$ into account. The Lass parameterization is used to model the $K\pi$ $S-$wave contribution. It consists of the $K_0^*(1430)$ resonance together with an effective range non-resonant component [18–20]:

$$T_{Lass}(s) = \frac{\sqrt{s}}{q\cot\delta_L - iq} + e^{2i\,\delta_L} \frac{m_0\,\Gamma_0 \frac{m_0}{q_0}}{m_0^2 - s - i\, m_0\,\Gamma_0 \, \frac{m_0}{\sqrt{s}} \, \frac{q}{q_0}} \tag{2.19}$$

with $\cot\delta_L = \frac{1}{aq} + \frac{1}{2}rq$. We use the values for the scattering length a and effective range parameter r from Ref. [18, 19]. Equation (2.16) is used for all other resonances decaying into a two-body final state.

For the resonances $K_1(1270)$ and $K(1460)$, the energy-dependent widths as well as the nominal mass and width are taken from Ref. [21]. We further use the energy-dependent widths for the $K_1(1400)$, $K^*(1410)$ and $K^*(1680)$ mesons from Ref. [9]. For all other resonances decaying into a three-body final state, an energy-dependent width distribution is derived from Equation 2.17 assuming an uniform phase space population.

Some particles may not originate from a resonance but are in a state of relative orbital angular momentum. We denote such non-resonant states by surrounding the particle system with brackets and indicate the partial wave state with an subscript; for example $(\pi\pi)_S$ refers to a non-resonant di-pion S-wave. The lineshape for non-resonant states is set to unity.

2.2.2 Spin Densities

The spin amplitudes are phenomenological descriptions of decay processes that are required to be Lorentz invariant, compatible with angular momentum conservation and, where appropriate, parity conservation. They are constructed in the covariant Zemach (Rarita-Schwinger) tensor formalism [22–24]. At this point, we briefly introduce the fundamental objects of the covariant tensor formalism which connect the particle's four-momenta to the spin dynamics of the reaction and give a general recipe to calculate the spin factors for arbitrary decay trees. Further details can be found in Refs. [25, 26].

A spin-S particle with four-momentum p, and spin projection λ, is represented by the polarization tensor $\epsilon_{(S)}(p,\lambda)$, which is symmetric, traceless and orthogonal to p. These so-called Rarita-Schwinger conditions reduce the a priori 4^S elements of the rank-S tensor to $2S+1$ independent elements in accordance with the number of degrees of freedom of a spin-S state [23, 27].

The spin projection operator $P^{\mu_1\ldots\mu_S\nu_1\ldots\nu_S}_{(S)}(p_R)$, for a resonance R, with spin $S = \{0,1,2\}$, and four-momentum p_R, is given by [26]:

$$\begin{aligned}
P^{\mu\nu}_{(0)}(p_R) &= 1 \\
P^{\mu\nu}_{(1)}(p_R) &= -\,g^{\mu\nu} + \frac{p_R^\mu\, p_R^\nu}{p_R^2} \\
P^{\mu\nu\alpha\beta}_{(2)}(p_R) &= \frac{1}{2}\left[P^{\mu\alpha}_{(1)}(p_R)\,P^{\nu\beta}_{(1)}(p_R) + P^{\mu\beta}_{(1)}(p_R)\,P^{\nu\alpha}_{(1)}(p_R)\right] - \frac{1}{3}\,P^{\mu\nu}_{(1)}(p_R)\,P^{\alpha\beta}_{(1)}(p_R)\,,
\end{aligned} \tag{2.20}$$

where $g^{\mu\nu}$ is the Minkowski metric. Contracted with an arbitrary tensor, the projection operator selects the part of the tensor which satisfies the Rarita-Schwinger conditions.

For a decay process $R \to AB$, with relative orbital angular momentum L, between particle A and B, the angular momentum tensor is obtained by projecting the rank-L tensor $q_R^{\nu_1}\, q_R^{\nu_2}\,\ldots\, q_R^{\nu_L}$, constructed from the relative momenta $q_R = p_A - p_B$, onto the spin-L subspace,

$$L_{(L)\mu_1\ldots\mu_L}(p_R, q_R) = (-1)^L\, P_{(L)\mu_1\ldots\mu_L\nu_1\ldots\nu_L}(p_R)\, q_R^{\nu_1}\,\ldots\, q_R^{\nu_L}\,. \tag{2.21}$$

Their $|\vec{q}_R|^L$ dependence accounts for the influence of the centrifugal barrier on the transition amplitudes. For the sake of brevity, the following notation is introduced,

$$\begin{aligned}
\varepsilon_{(S)}(R) &\equiv \varepsilon_{(S)}(p_R, \lambda_R), \\
P_{(S)}(R) &\equiv P_{(S)}(p_R), \\
L_{(L)}(R) &\equiv L_{(L)}(p_R, q_R).
\end{aligned} \tag{2.22}$$

Following the isobar approach, a four-body decay amplitude is described as a product of two-body decay amplitudes. Each sequential two-body decay $R \to A\,B$, with relative orbital angular momentum L_{AB}, and total intrinsic spin S_{AB}, contributes a term to the overall spin factor given by

$$\begin{aligned}
S_{R\to AB}(\mathbf{x}|L_{AB}, S_{AB}; \lambda_R, \lambda_A, \lambda_B) = \varepsilon_{(S_R)}(R)\, X(S_R, L_{AB}, S_{AB})\, L_{(L_{AB})}(R) \\
\times\, \Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B),
\end{aligned} \tag{2.23}$$

where

$$\Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B) = P_{(S_{AB})}(R)\, X(S_{AB}, S_A, S_B)\, \varepsilon^*_{(S_A)}(A)\, \varepsilon^*_{(S_B)}(B)\,. \tag{2.24}$$

Here, a polarization vector is assigned to the decaying particle and the complex conjugate vectors for each decay product. The spin and orbital angular momentum couplings are described by the tensors $P_{(S_{AB})}(R)$ and $L_{(L_{AB})}(R)$, respectively. Firstly, the two spins S_A and S_B, are coupled to a total spin-S_{AB} state, $\Phi(\mathbf{x}|S_{AB})$, by projecting the corresponding polarization vectors onto the spin-S_{AB} subspace transverse to the momentum of the decaying particle. Afterwards, the spin and orbital angular momentum tensors are properly contracted with the polarization vector of the decaying particle to give a Lorentz scalar. This requires in some cases to include the tensor $\varepsilon_{\alpha\beta\gamma\delta}\, p_R^\delta$ via

$$X(j_a, j_b, j_c) = \begin{cases} 1 & \text{if } j_a + j_b + j_c \text{ even} \\ \varepsilon_{\alpha\beta\gamma\delta}\, p_R^\delta & \text{if } j_a + j_b + j_c \text{ odd} \end{cases}, \tag{2.25}$$

where $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita symbol and j refers to the arguments of X defined in Eqs. 2.23 and 2.24. Its antisymmetric nature ensures the correct parity transformation behavior of the amplitude. The spin factor for a whole decay chain, for example $R \to (R_1 \to AB)\,(R_2 \to CD)$, is obtained by combining the two-body terms and performing a sum over all unobservable, intermediary spin projections

$$\sum_{\lambda_{R_1},\lambda_{R_2}} S_{R\to R_1R_2}(\mathbf{x}|L_{R_1R_2};\lambda_{R_1},\lambda_{R_2})\, S_{R_1\to AB}(\mathbf{x}|L_{AB};\lambda_{R_1})\, S_{R_2\to CD}(\mathbf{x}|L_{CD};\lambda_{R_2}), \tag{2.26}$$

where $\lambda_R = \lambda_A = \lambda_B = \lambda_C = \lambda_D = 0$, $S_{AB} = S_{CD} = 0$ and $S_{R_1R_2} = L_{R_1R_2}$, as only pseudoscalar initial/final states are involved.

The spin factors for all decay topologies considered in this analysis are explicitly given in Appendix F.

2.3 Fit implementation

The hadronic amplitudes are renormalized prior to the amplitude fit such that

$$\int |A_i(\mathbf{x})|^2 \, \mathrm{d}\Phi_4 = 1. \tag{2.27}$$

This allows us to set more intuitive starting values as the amplitude coefficients are all on a comparable scale. Moreover, the total amplitudes $\mathcal{A}_f^{c(u)}(\mathbf{x})$ are renormalized on-the-fly such that

$$\begin{aligned} \int \left|\mathcal{A}_f^{c(u)}(\mathbf{x})\right|^2 \, \mathrm{d}\Phi_4 = 1 \\ \arg\left(\int \mathcal{A}_f^{c}(\mathbf{x})^* \, \mathcal{A}_f^{u}(\mathbf{x}) \, \mathrm{d}\Phi_4\right) = 0\,. \end{aligned} \tag{2.28}$$

As a result, the average amplitude ratio and strong phase difference between the $b \to u$ and $b \to c$ transitions can be introduced as direct fit parameters instead of derived quantities that have to be calculated from Equation 2.11 after the fit. For the differential decay rate follows:

$$\begin{aligned} \frac{\mathrm{d}\Gamma(\mathbf{x},t,q,f)}{e^{-\Gamma_s t}\,\mathrm{d}t\,\mathrm{d}\Phi_4} \propto & \left(|\mathcal{A}_f^c(\mathbf{x})|^2 + r^2\,|\mathcal{A}_f^u(\mathbf{x})|^2\right) \cosh\left(\frac{\Delta\Gamma_s\,t}{2}\right) \\ & + q\,f\,\left(|\mathcal{A}_f^c(\mathbf{x})|^2 - r^2\,|\mathcal{A}_f^u(\mathbf{x})|^2\right) \cos\left(\Delta m_s\,t\right) \\ & - 2\,r\mathrm{Re}\left(\mathcal{A}_f^c(\mathbf{x})^*\,\mathcal{A}_f^u(\mathbf{x})\,e^{i\delta - if(\gamma-2\beta_s)}\right) \sinh\left(\frac{\Delta\Gamma_s\,t}{2}\right) \\ & - 2\,q\,f\,r\,\mathrm{Im}\left(\mathcal{A}_f^c(\mathbf{x})^*\,\mathcal{A}_f^u(\mathbf{x})\,e^{i\delta - if(\gamma-2\beta_s)}\right) \sin\left(\Delta m_s\,t\right) \end{aligned} \tag{2.29}$$

This renormalization procedure was found to be crucial for the fit stability since it reduces the correlation between the a_i^c and a_i^u amplitude coefficients significantly. Due to the overall normalization, one of the complex amplitude coefficients a_i^c can be fixed to unity and since r and δ are included as fit parameters one of the complex amplitude coefficient a_i^u can be additionally fixed to unity.

We force strong decays in the cascade topology to have the same pattern in $b \to c$ and $b \to u$ transitions by the sharing of couplings between related quasi-two-body final states. For example, given the two a_i^c parameters required for $B_s \to D_s^- \, K_1(1270)^+$ with $K_1(1270)^+ \to \rho(770)^0 \, \pi^+$ and $K_1(1270)^+ \to K^*(892) \, \pi^+$, the amplitude $\bar{B}_s \to D_s^- \, K_1(1270)^+$ with $K_1(1270)^+ \to \rho(770)^0 \, \pi^+$ and $K_1(1270)^+ \to K^*(892) \, \pi^+$ only requires one additional global complex parameter to represent the different production processes of $B_s \to D_s^- \, K_1(1270)^+$ and $\bar{B}_s \to D_s^- \, K_1(1270)^+$, while the relative magnitude and phase of $K_1(1270)^+ \to \rho(770)^0 \, \pi^+$ and $K_1(1270)^+ \to K^*(892) \, \pi^+$ are the same regardless of the production mechanism. For this purpose, multiple decay amplitudes of a three-body resonance are defined relative to a given reference channel.

2.4 Validation

The amplitude fit is performed by using the amplitude analysis tool MINT2. It was previously applied to analyze $D^0 \to 4\pi$ and $D^0 \to KK\pi\pi$ decays [9] which have an identical general spin structure (*i.e.* scalar to four scalar decay) to $B_s \to D_s K\pi\pi$ decays. In the course of the $D^0 \to hhhh$ analysis, the implementation of the amplitudes were extensively cross-checked against other available tools such as qft++ [28], AmpGen [21] and where possible EVTGEN [29]. Since no additional line shapes or spin factors are needed for this analysis, we consider the amplitude calculation as fully validated.

This does, however, not apply to the full time-dependent amplitude pdf which is newly implemented for this analysis. To cross-check it, we use EVTGEN to generate toy events with time-dependent CP violation according to the SSD_CP event model [29]. Since this event model does not allow for multiple interfering resonances, we generate only the decay chain $B_s \to D_s\,(K_1(1400) \to K^*(892)\pi)$. Table 2.1 lists the generated input parameters. The toy data set is fitted with our MINT2 implementation of the full time-dependent amplitude pdf and the phasespace-integrated pdf. The fit projections are shown in Figs. 2.1 and 2.2.

The CP coefficients $C, D, \bar{D}, S, \bar{S}$ are the fit parameters in case of the phasespace-integrated pdf, which are converted after to the fit to the physical observables r, κ, δ and γ using the GammaCombo package [30]. The obtained 1-CL contours are shown in Fig. 2.3. The full pdf determines r, δ and γ directly. As shown in Tab. 2.2and 2.3, the fit results are in excellent agreement with the generated input values. The phasespace-integrated fit is, in addition, performed with the B2DX fitter used for the $B_s \to D_s K$ analysis yielding identical results. Note though that some parts of the B2DX fitter have been taken over to our MINT2 fitter, such that the implementations are not fully independent.

Table 2.1: Input values used to generate EVTGEN toy events according to the SSD_CP event model.

τ	1.5 ps
$\Delta\Gamma$	$-0.1\,\text{ps}^{-1}$
Δm_s	$17.757\,\text{ps}^{-1}$
r	0.37
κ	1
δ	$10.0°$
γ	$71.1°$
β_s	$0.0°$

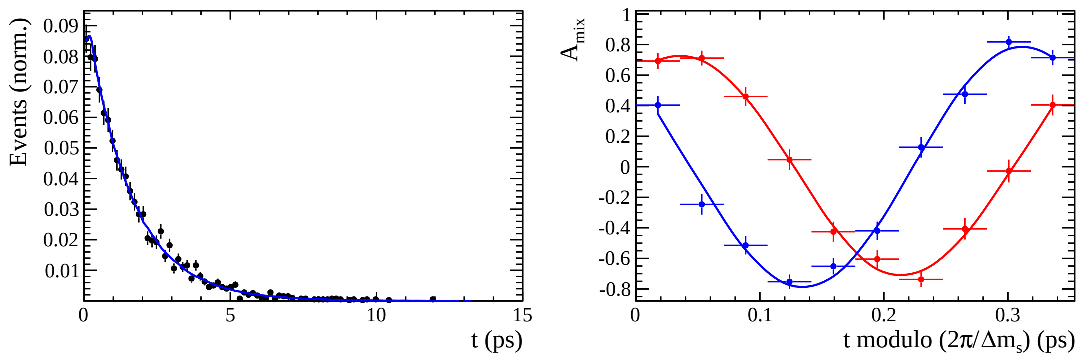

Figure 2.1: Left: Time distribution of $B_s \to D_s K\pi\pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (solid line). Right: Time-dependent asymmetry between mixed and unmixed events folded into one oscillation period for $D_s^- K^+\pi\pi$ (red) and $D_s^+ K^-\pi\pi$ (blue) final states. The data points show events generated with EVTGEN, while the solid lines show the MINT2 fit projections.

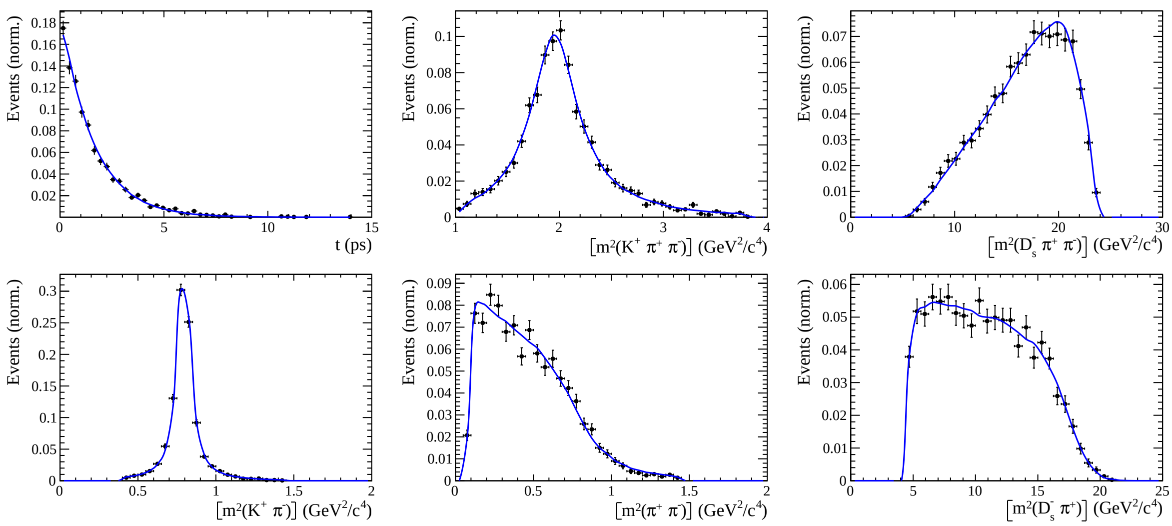

Figure 2.2: Time and invariant mass distributions of $B_s \to D_s K\pi\pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

Table 2.2: Result of the phasespace-integrated fit to EVTGEN toy events.

	Generated	Fit result	Pull(σ)
C	0.759	0.763 ± 0.026	0.2
D	-0.314	-0.376 ± 0.227	-0.3
$\bar{D}$	-0.101	-0.261 ± 0.246	-0.7
S	-0.570	-0.626 ± 0.035	1.6
$\bar{S}$	-0.643	-0.669 ± 0.035	-0.7

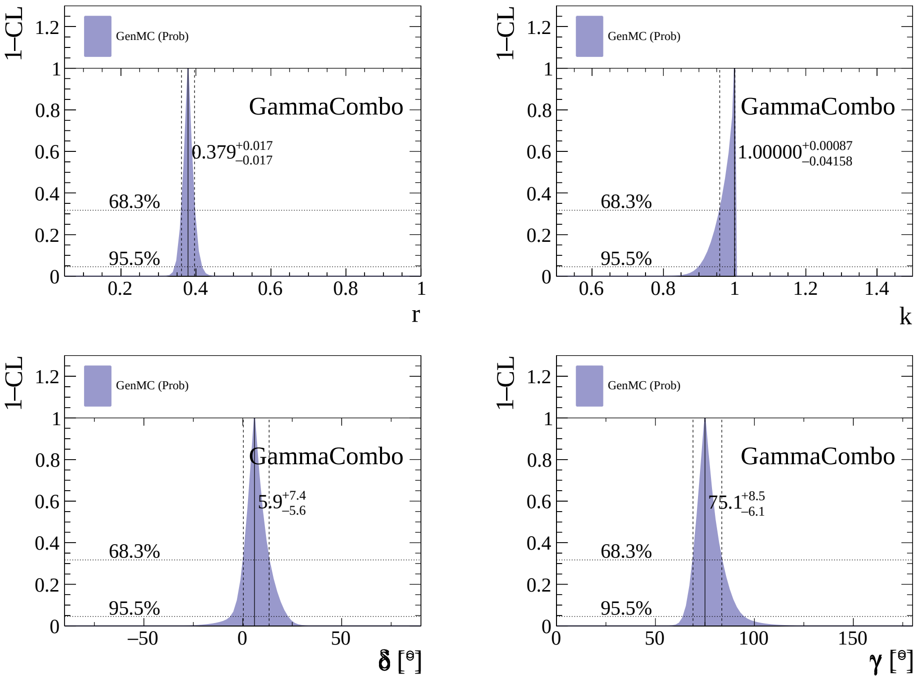

Figure 2.3: The 1-CL contours for the physical observable r, κ, δ and γ obtained with the phasespace-integrated fit to the EVTGEN toy sample.

Table 2.3: Results for the physical observables obtained from fits to EVTGEN toy events.

	Generated	Full PDF	Phasespace-integrated
r	0.370	0.379 ± 0.021	0.379 ± 0.017
κ	1.0	1.0	1.000 ± 0.059
δ	$10.0°$	9.0 ± 5.1	5.9 ± 6.0
γ	$71.1°$	67.3 ± 5.9	75.1 ± 6.9

3 Data samples and event selection

3.1 Stripping and Trigger selection

The dataset used for this analysis corresponds to $1\,\text{fb}^{-1}$ of proton-proton collision data collected in 2011 with a centre of mass energy $\sqrt{s} = 7\,\text{TeV}$, $2\,\text{fb}^{-1}$ collected in 2012 with $\sqrt{s} = 7\,\text{TeV}$ and $4\,\text{fb}^{-1}$ collected in 2015/2016/2017 with $\sqrt{s} = 13\,\text{TeV}$. Candidate $B_s^0 \to D_s K\pi\pi$ ($B_s^0 \to D_s\pi\pi\pi$) decays are reconstructed using the B02DKPiPiD2HHHPIDBeauty2CharmLine (B02DPiPiPiD2HHHPIDBeauty2CharmLine) line of the BHadronCompleteEvent stream of Stripping21r1 (2011), Stripping21 (2012), Stripping24r1 (2015) and Stripping28r1p1 (2016) and Stripping29r2 (2017). Both stripping lines employ the same selection cuts, listed in Appendix A, except for the PID requirement on the bachelor kaon/pion.

Events that pass the stripping selection are further required to fulfill the following trigger requirements: at the hardware stage, the B_s^0 candidates are required to be TOS on the L0Hadron trigger or TIS on L0Global; at Hlt1, B_s^0 candidates are required to be TOS on the Hlt1TrackAllL0 (Hlt1TrackMVA or Hlt1TwoTrackMVA) trigger line for Run-I (Run-II) data; at Hlt2, candidates have to be TOS on either one of the 2, 3 or 4-body topological trigger lines or the inclusive ϕ trigger. More details on the used HLT lines are given in Appendix A.

Due to a residual kinematic dependence on whether the event is triggered by L0Hadron or not and on the data taking condition, the analysis is performed in four disjoint categories: [Run-I,L0-TOS], [Run-I,L0-TIS], [Run-II,L0-TOS] and [Run-II,L0-TIS], where for simplicity we denote L0Hadron-TOS as L0-TOS and (L0Global-TIS and not L0Hadron-TOS) as L0-TIS.

3.2 Offline selection

The offline selection, in particular the requirements on the D_s hadron, are guided by the previous analyses of $B_s \to D_s K/\pi$, $B_d \to D^0\pi$ as well as the branching fraction measurement of $B_s^0 \to D_s K\pi\pi$ decays. Tables 3.1 and 3.2 summarize all selection requirements which are described in the following. Throughout the note, we abbreviate $B_s^0 \to D_s X_s(\to K\pi\pi)$ and $B_s^0 \to D_s X_d(\to \pi\pi\pi)$.

Given the high number of pp interactions per bunch crossing, a large fraction of events have more than one reconstructed PV. We choose the 'best' PV to be the one to which the B_s candidate has the smallest χ^2_{IP}. In instances where the association of the B_s candidate to the best PV is wrong, the decay time of this candidate will be incorrect. These wrongly associated candidates are rejected by requiring that the B_s χ^2_{IP} with respect to any other PV is sufficiently higher than with respect to the best PV ($\Delta\chi^2_{IP} = \chi^2_{\text{IP,SECONDBEST}} - \chi^2_{\text{IP,BEST}} > 20$). Events with only a single PV are not affected.

In order to clean up the sample and to align the Run-I to the slightly tighter Run-II stripping selection, we apply the following loose cuts to the b-hadron:

- DIRA > 0.99994
- min IP $\chi^2 < 16$ to the best PV,
- FD $\chi^2 > 100$ to the best PV,
- Vertex $\chi^2/\text{nDoF} < 8$.

The cut on the B_s decay-time is tightened with respect to the stripping selection ($t > 0.2\,\mathrm{ps}$) because, while offline we use the decay-time determined for a DTF in which the PV position, the D_s and the B_s mass are constrained, in the stripping the simple decay-time returned by a kinematic fit is used. The difference between these two decay-times extends up to 0.1 ps, thus cutting at 0.4 ps avoids any boundary effects and simplifies the time-acceptance studies. We further remove outliers with poorly estimated decay times ($\delta t < 0.15\,\mathrm{ps}$).

We reconstruct the $B_s^0 \to D_s h\pi\pi$ decay through three different final states of the D_s meson: $D_s \to KK\pi$, $D_s \to \pi\pi\pi$ and $D_s \to K\pi\pi$. Of those, $D_s \to KK\pi$ is the most prominent one, while $\mathcal{BR}(D_s \to \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \to KK\pi)$ and $\mathcal{BR}(D_s \to K\pi\pi) \approx 0.12 \cdot \mathcal{BR}(D_s \to KK\pi)$ holds for the others. For the $KK\pi$ final state we make use of the well known resonance structure; the decay proceeds either via the narrow ϕ resonance, the broader K^{*0} resonance or (predominantly) non-resonant. Within the ϕ resonance region the sample is already sufficiently clean after the stripping so that we do not impose additional criteria on the D_s daughters. For the K^{*0} and non-resonant regions consecutively tighter requirements on the particle identification and the D_s flight-distance are applied. We apply global requirements for the other final states. All cuts are summarized in Table 3.1.

3.2.1 Phase space region

Due to the comparably low masses of the final state particles with respect to the B_s mass, there is, in contrast to for example b-hadron to charmonia or charm decays, a huge phase-space available for the $B_s^0 \to D_s K\pi\pi$ decay. For the invariant mass of the $K\pi\pi$ subsystem it extends up to $3.4\,\mathrm{GeV}$. It has however been observed that the decay proceeds predominantly through the low lying axial vector states $K_1(1270)$ and $K_1(1400)$, while the combinatorial background is concentrated at high $K\pi\pi$ invariant masses ($m(K\pi\pi) > 2000\,\mathrm{MeV}$). Moreover, the strange hadron spectrum above $2\,\mathrm{GeV}$ is poorly understood experimentally such that a reliable extraction of the strong phase motion in that region is not possible. We consequently choose to limit the considered phase space region to $m(K\pi\pi) < 1950\,\mathrm{MeV}$, which is right below the charm-strange threshold ($B_s^0 \to D_s^+ D_s^-$).

3.2.2 Physics background vetoes

We veto various physical backgrounds, which have either the same final state as our signal decay, or can contribute via a single misidentification of $K \leftrightarrow \pi$, $K \leftrightarrow p$ or $\pi \leftrightarrow p$. Depending on the D_s final state different vetoes are applied in order to account for peaking backgrounds originating from charm meson or charmed baryon decays.

1. ${D_s}^- \to K^+K^-\pi^-$

   (a) $D^- \to K^+\pi^-\pi^-$:
   Possible with $\pi^- \to K^-$ misidentification, vetoed by requiring $m(K^+K^-_{\pi}\pi^-) \neq m(D^-) \pm 40$ MeV or the K^- has to fulfill more stringent PID criteria depending on the resonant region (see Table 3.1).

   (b) $\Lambda_c^- \to K^+\bar{p}\pi^-$:
   Possible with $\bar{p} \to K^-$ misidentification, vetoed by requiring $m(K^+K^-_{p}\pi^-) \neq m(\Lambda_c) \pm 40$ MeV or the K^- has to fulfill more stringent PID criteria depending on the resonant region (see Table 3.1).

   (c) $D^0 \to KK$:
   D^0 combined with a random π can fake a $D_s \to KK\pi$ decay, vetoed by requiring $m(KK) < 1840\,\text{MeV}$.

2. ${D_s}^- \to \pi^+\pi^-\pi^-$

   (a) $D^0 \to \pi\pi$:
   D^0 combined with a random π can fake a $D_s \to \pi\pi\pi$ decay, vetoed by requiring both possible combinations to have $m(\pi\pi) < 1700\,\text{MeV}$.

3. ${D_s}^- \to K^-\pi^+\pi^-$

   (a) $D^- \to \pi^-\pi^+\pi^-$:
   Possible with $\pi^- \to K^-$ misidentification, vetoed by requiring $m(K^-_{\pi}\pi^+\pi^-) \neq m(D^-) \pm 40$ MeV or PIDK(K^+) > 15.

   (b) $\Lambda_c^- \to \bar{p}\pi^+\pi^-$:
   Possible with $\bar{p} \to K^-$ misidentification, vetoed by requiring $m(K^-_{p}\pi^+\pi^-) \neq m(\Lambda_c) \pm 40$ MeV or PIDK(K^-) $-$ PIDp(K^-) > 5.

   (c) $D^0 \to K\pi$:
   D^0 combined with a random π can fake a $D_s \to K\pi\pi$ decay, vetoed by requiring both possible combinations to have $m(K\pi) < 1750\,\text{MeV}$.

The effects of these veto cuts are illustrated in Figs. 3.1,3.2 and 3.3. To reduce cross-feed from our calibration channel into the signal channel and vice-versa we require tight PID cuts on the ambiguous bachelor kaon (PIDK(K) > 10)/pion (PIDK(K) < 0). In addition, we veto $B_s^0 \to D_s^- D_s^+$ decays which is illustrated in Fig. 3.4.

1. $X_s^+ \to K^+\pi^+\pi^-$:

   (a) $B_s^0 \to D_s\pi\pi\pi$:
   Possible with $\pi^+ \to K^+$ misidentification, suppressed with PIDK(K^+) > 10.

   (b) $B_s^0 \to D_s^-(D_s^+ \to K^+\pi^+\pi^-)$:
   Outside of considered phase-space region, see Sec. 3.2.1.

   (c) $B_s^0 \to D_s^-(D_s^+ \to K^+K^-\pi^+)$:
   To suppress $B_s^0 \to D_s^- K^+K^-\pi^+$ background, possible with $K^- \to \pi^-$ misidentification, we require PIDK(π^-) < 0. In case the invariant mass of the $K^+\pi^+\pi$ system recomputed applying the kaon mass hypothesis to the pion is close to the D_s mass ($m(K^+\pi^+\pi^-_K) = m(D_s) \pm 20$ MeV), we further tighten the cut to PIDK(π^-) < -5.

2. $X_d^+ \to \pi^+\pi^+\pi^-$:

   (a) $B_s^0 \to D_sK\pi\pi$:
   Possible with single missID of $K^+ \to \pi^+$, suppressed with PIDK(π^+) < 0.

   (b) $B_s^0 \to D_s^-(D_s^+ \to \pi^+\pi^+\pi^-)$:
   Outside of considered phase-space region, see Sec. 3.2.1.

   (c) $B_s^0 \to D_s^-(D_s^+ \to K^+\pi^-\pi^+)$:
   Possible with single missID of $K^+ \to \pi^+$, vetoed by requiring $m(\pi^+\pi^+_K\pi^-) \neq m(D_s) \pm 20$ MeV or PIDK(π^+) < -5 for both π^+.

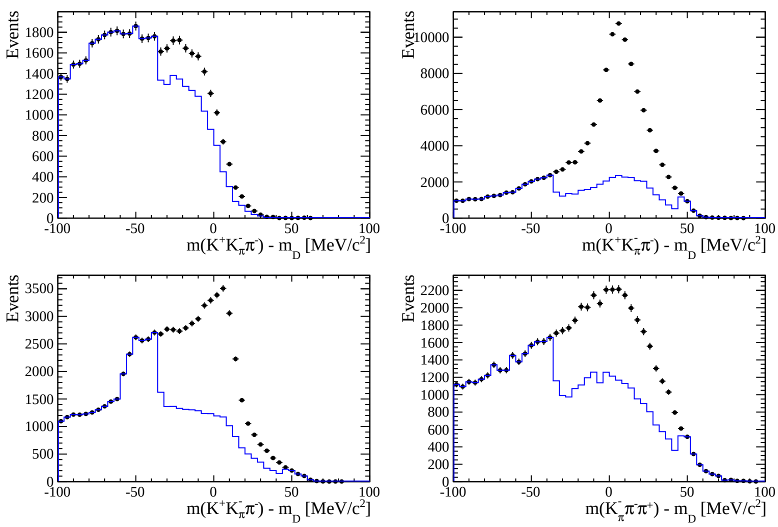

Figure 3.1: Background contributions from D^- decays where the π^- is misidentified as K^-. The D_s invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the $D_s \to \phi\pi$, $D_s \to K^*(892)K$, $D_s \to KK\pi$ (non-resonant) and $D_s \to K\pi\pi$ final state categories from top-left to bottom-right. The distributions are shown without (black) and with (blue) the D^--veto applied.

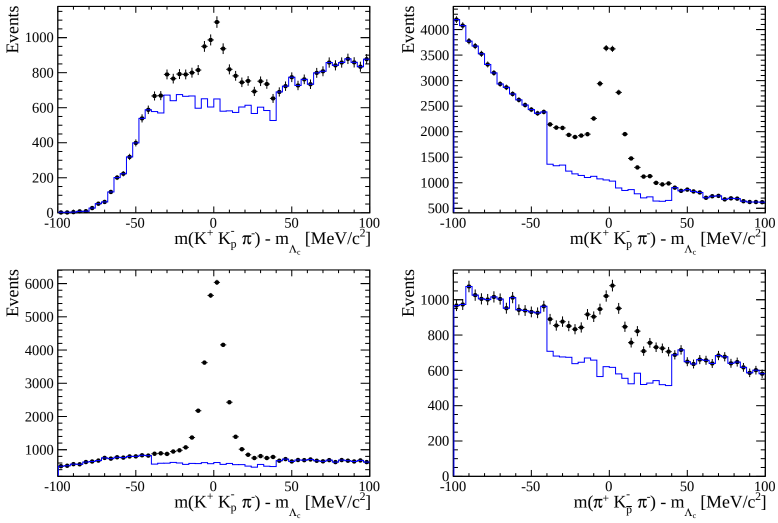

Figure 3.2: Background contributions from Λ_c decays where the $\bar{p}$ is misidentified as K^-. The D_s invariant mass is recomputed applying the proton mass hypothesis to the kaon and shown for the $D_s \to \phi\pi$, $D_s \to K^*(892)K$, $D_s \to KK\pi$ (non-resonant) and $D_s \to K\pi\pi$ final state categories from top-left to bottom-right. The distributions are shown without (black) and with (blue) the Λ_c-veto applied.

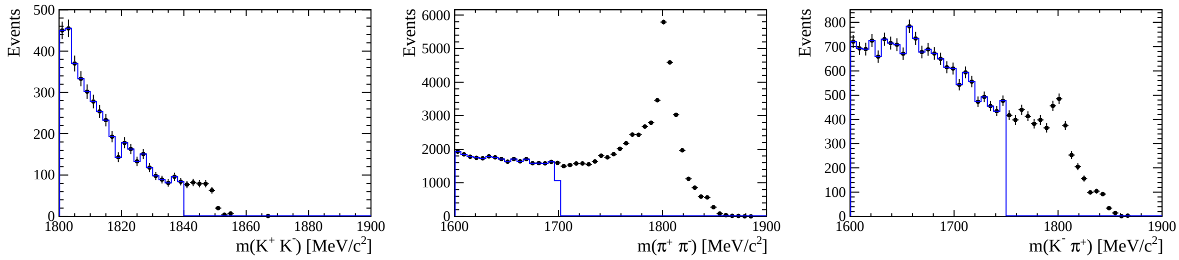

Figure 3.3: Background contributions to $D_s \to KK\pi$ (left), $D_s \to \pi\pi\pi$ (middle) and $D_s \to K\pi\pi$ (right) from $D^0 \to hh$ decays combined with a random pion.

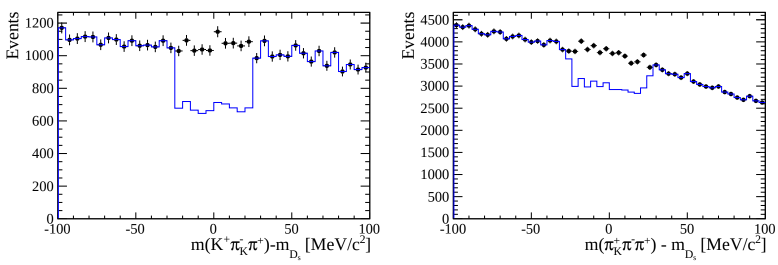

Figure 3.4: Background contributions to $B_s \to D_s K\pi\pi$ (left) and $B_s \to D_s\pi\pi\pi$ (right) from $B_s \to D_s D_s$ decays where the kaon is misidentified as pion. The $X_{s,d}$ invariant mass is recomputed applying the kaon mass hypothesis to the pion and shown without (black) and with (blue) the D_s-veto applied.

3.2.3 Training of multivariate classifier

The Toolkit for Multivariate Analysis (TMVA [31]) is used to train a multivariate classifier (BDT with gradient boosting, BDTG) discriminating signal and combinatorial background. We use $B_s \to D_s\pi\pi\pi$ data that passes the preselection as signal proxy. The background is statistically subtracted by applying sWeights based on the fit to the reconstructed B_s mass shown in Fig. 3.5. This is a simplified version (performed in a reduced mass range) of the final mass fits described in Sec. 4. The sideband data ($m(B_s) > 5500\,\text{MeV}$) is used as background proxy.

Training the classifier on a sub-sample which is supposed to be used in the final analysis might cause a bias, as the classifier selects, in case of overtraining, the training events more efficiently. As overtraining can not be completely avoided, we split the signal and the background training samples into two disjoint subsamples according to whether the event number is even or odd. We then train the classifier on the even sample and apply it to the odd one, and vice-versa (cross-training).

The following discriminating variables are used for the BDTG training:

- logarithm of the B_s impact-parameter χ^2, B_s $\log(\chi^2_{IP})$
- logarithm of the cosine of the B_s direction angle, log(DIRA)
- fit quality of the DTF with PV constrain, χ^2_{DTF}/ndf
- logarithm of the minimal vertex quality difference for adding one extra track, $\log(\Delta\chi^2_{add-track})$
- the asymmetry between the transverse momentum of the B_s- candidate and the transverse momentum of all the particles reconstructed with a cone of radius $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2} < 1$ rad around the B_s- candidate, B_s $A^{cone}_{p_T}$
- largest ghost probability of all tracks, max(ghostProb)
- logarithm of the the smallest X_s daughter impact-parameter χ^2, X_s $\log(\min(\chi^2_{IP}))$
- largest distance of closest approach of the X_s daughters, max(DOCA)
- cosine of the largest opening angle between the D_s and another bachelor track h_i in the plane transverse to the beam, $\cos(\max\theta_{D_s h_i})$
- logarithm of the the smallest D_s daughter impact-parameter χ^2, D_s $\log(\min(\chi^2_{IP}))$
- logarithm of the D_s flight-distance significance, D_s $\log(\chi^2_{FD})$
- logarithm of the D_s radial flight-distance, D_s $\log(RFD)$

Loose cuts on the variables χ^2_{DTF}/ndf, $\Delta\chi^2_{add-track}$ and $\cos(\max\theta_{D_s h_i})$ are applied prior to the training which are expected to be 100% signal efficient. Figure 3.6 shows the distributions of the input variables for signal and background. As shown in Appendix B, these distributions differ between data-taking period and trigger category. In particular variables depending on the B_s kinematics and the event multiplicity are affected (*e.g.* $\theta_{D_s h_i}$ or $A^{cone}_{p_T}$). The BDTG is consequently trained separately for these categories. The resulting classifier response is shown in Fig. 3.7 for each category (even and odd test samples combined) and in Appendix B for each training.

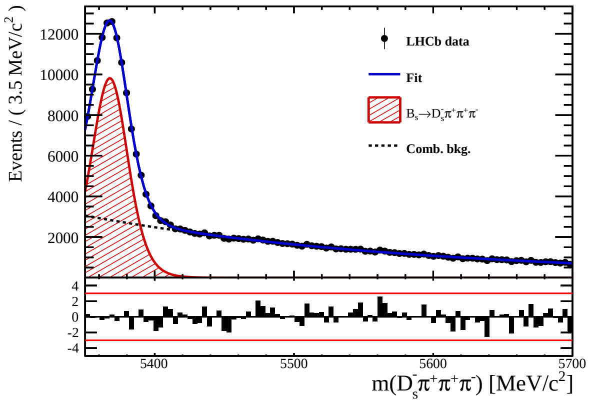

Figure 3.5: Reconstructed B_s mass for $B_s \to D_s\pi\pi\pi$ events that pass the preselection. The fitted PDF is shown in blue, the signal component in red and the background component in black.

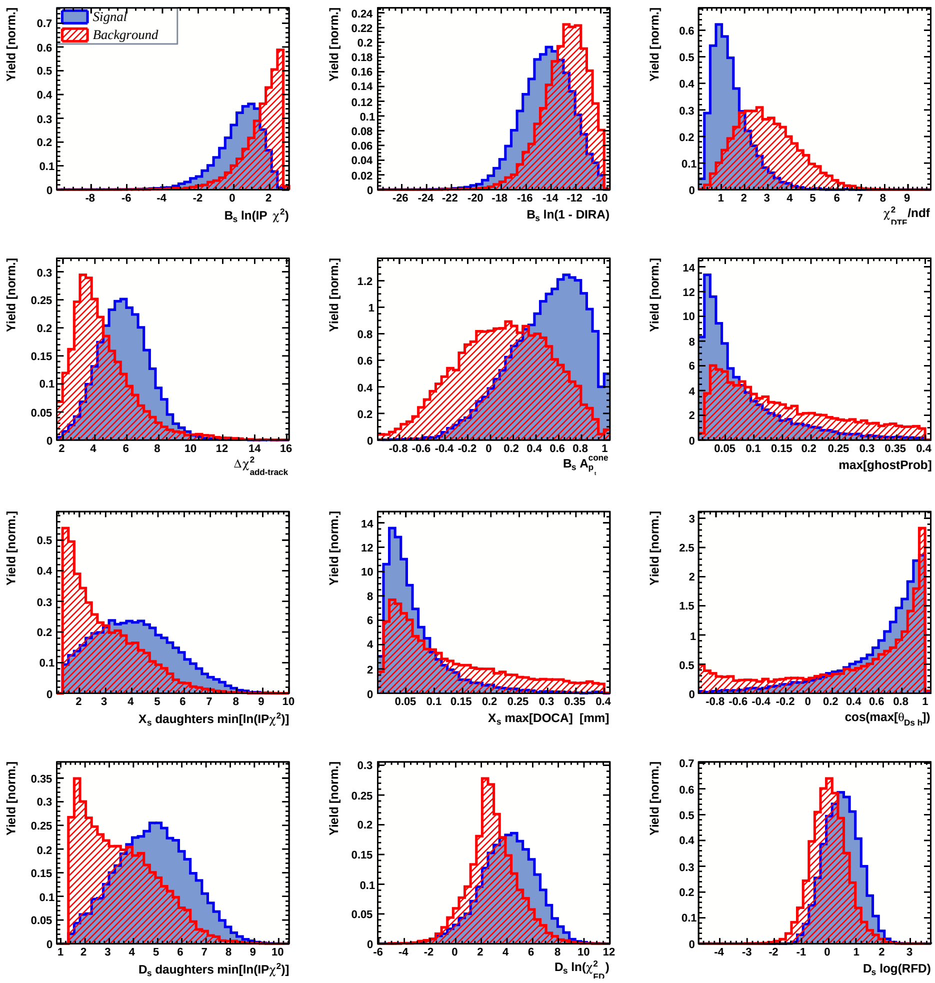

Figure 3.6: Discriminating variables used to train the BDTG for all data categories combined.

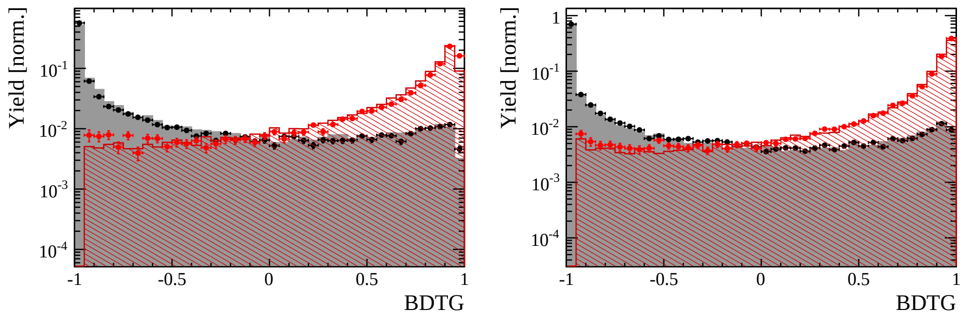

Figure 3.7: Signal (red) and background (black) distributions for the BDTG response for Run-I (left) and Run-II (right) data. Filled histograms (data points) show the BDTG response for the L0-TOS (L0-TIS) category. Even and odd test samples are combined.

3.2.4 Final selection

The cut on the BDTG output is optimized using the signal significance as figure of merit:

$$\mathrm{FOM(BDTG)} = \frac{N_s(\mathrm{BDTG})}{\sqrt{N_s(\mathrm{BDTG}) + N_b(\mathrm{BDTG})}} \tag{3.1}$$

where $N_s(\mathrm{BDTG})$ is the $B_s \to D_s K\pi\pi$ signal yield for a given BDTG cut and $N_b(\mathrm{BDTG})$ is the combinatorial background yield in the signal region ($m(D_s K\pi\pi) = m_{B_s} \pm 40\,\mathrm{MeV}$). To compute the yields as function of the BDTG cut, we use the BDTG efficiencies, $\epsilon_{s,b}$, evaluated on the corresponding test samples. To fix the overall scale, it is required to know the yields at (at least) one point of the scanned range. We arbitrarily choose this fix point to be BDTG > 0 and perform a fit to the reconstructed B_s mass as described in Sec. 4 to obtain the yields $N_{s,b}(0)$. These yields are then efficiency corrected to calculate the yields for a given BDTG cut:

$$N_{s,b}(\mathrm{BDTG}) = N_{s,b}(0) \cdot \frac{\epsilon_{s,b}(\mathrm{BDTG})}{\epsilon_{s,b}(0)}. \tag{3.2}$$

Figure 3.8 shows the resulting BDTG scans for each training category.

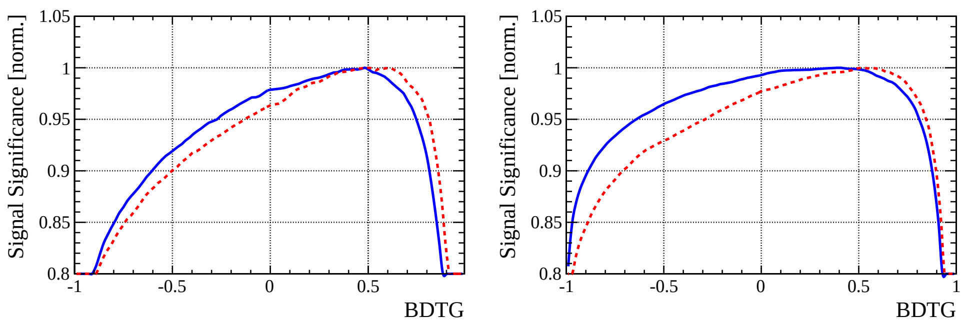

Figure 3.8: Signal significance as function of the applied cut on the BDTG response for Run-I (left) and Run-II (right) data. The scans for the L0-TOS (L0-TIS) category are shown in blue (red). The signal significance is normalized to be 1 at the optimal BDTG cut value.

Table 3.1: Offline selection requirements for $D_s \to 3h$ candidates.

	Description	Requirement
$D_s \to hhh$	$m(hhh)$	$= m_{D_s} \pm 25\,\text{MeV}$
$D_s^- \to KK\pi^-$	D^0 veto	$m(KK) < 1840$ MeV
$D_s^- \to \phi\pi^-$	$m(KK)$	$= m_\phi \pm 12$ MeV
	PIDK(K^+)	> -10
	PIDK(K^-)	> -10
	PIDK(π^-)	< 20
	χ^2_{FD}	> 0
	FD in z	> -1
	D^- veto	$m(K^+K^-_\pi\pi^-) \neq m(D^-) \pm 40$ MeV || PIDK(K^-) > 5
	Λ_c veto	$m(K^+K^-_p\pi^-) \neq m(\Lambda_c) \pm 40$ MeV || PIDK(K^-) $-$ PIDp(K^-) > 2
$D_s^- \to K^*(892)K^-$	$m(KK)$	$\neq m_\phi \pm 12$ MeV
	$m(K^+\pi^-)$	$= m_{K^*(892)} \pm 75$ MeV
	PIDK(K^+)	> -10
	PIDK(K^-)	> -5
	PIDK(π^-)	< 10
	χ^2_{FD}	> 0
	FD in z	> 0
	D^- veto	$m(K^+K^-_\pi\pi^-) \neq m(D^-) \pm 40$ MeV || PIDK(K^-) > 15
	Λ_c veto	$m(K^+K^-_p\pi^-) \neq m(\Lambda_c) \pm 40$ MeV || PIDK(K^-) $-$ PIDp(K^-) > 5
$D_s^- \to (KK\pi^-)_{NR}$	$m(KK)$	$\neq m_\phi \pm 12$ MeV
	$m(K^+\pi^-)$	$\neq m_{K^*(892)} \pm 75$ MeV
	PIDK(K^+)	> 5
	PIDK(K^-)	> 5
	PIDK(π^-)	< 10
	χ^2_{FD}	> 4
	FD in z	> 0
	D^- veto	$m(K^+K^-_\pi\pi^-) \neq m(D^-) \pm 40$ MeV || PIDK(K^-) > 15
	Λ_c veto	$m(K^+K^-_p\pi^-) \neq m(\Lambda_c) \pm 40$ MeV || PIDK(K^-) $-$ PIDp(K^-) > 5
$D_s \to \pi\pi\pi$	PIDK(π)	< 10
	PIDp(π)	< 20
	D^0 veto	$m(\pi^+\pi^-) < 1700$ MeV
	χ^2_{FD}	> 9
	FD in z	> 0
$D_s^- \to K^-\pi^+\pi^-$	PIDK(K)	> 8
	PIDK(π)	< 5
	PIDp(π)	< 20
	D^0 veto	$m(K^-\pi^+) < 1750$ MeV
	χ^2_{FD}	> 9
	FD in z	> 0
	D^- veto	$m(K^-_\pi\pi^+\pi^-) \neq m(D^-) \pm 40$ MeV || PIDK(K^-) > 15
	Λ_c veto	$m(K^-_p\pi^+\pi^-) \neq m(\Lambda_c) \pm 40$ MeV || PIDK(K^-) $-$ PIDp(K^-) > 5

Table 3.2: Offline selection requirements for $B_s \to D_s K\pi\pi (D_s\pi\pi\pi)$ candidates.

	Description	Requirement
$B_s \to D_s h\pi\pi$	$m(D_s h\pi\pi)$	$> 5200\,\mathrm{MeV}$
	χ^2_{vtx}/ndof	< 8
	DIRA	> 0.99994
	χ^2_{FD}	> 100
	χ^2_{IP}	< 16
	χ^2_{DTF}/ndof	< 15
	$\Delta\chi^2_{add-track}$	> 2
	$\cos(\max\theta_{D_s h_i})$	> -0.9
	t	$> 0.4\,\mathrm{ps}$
	δt	$< 0.15\,\mathrm{ps}$
	Phasespace region	$m(h\pi\pi) < 1.95\,\mathrm{GeV}$
		$m(h\pi) < 1.2\,\mathrm{GeV}$
		$m(\pi\pi) < 1.2\,\mathrm{GeV}$
	Wrong PV veto	nPV $= 1$ || $\min(\Delta\chi^2_{IP}) > 20$
	BDTG	> 0.45 [Run-I,L0-TOS]
		> 0.50 [Run-I,L0-TIS]
		> 0.35 [Run-II,L0-TOS]
		> 0.50 [Run-II,L0-TIS]
$X_s^+ \to K^+\pi^+\pi^-$	PIDK(K)	> 10
	PIDK(π^+)	< 10
	PIDK(π^-)	< 0
	D_s veto	$m(K^+\pi^+\pi^-_K) \neq m(D_s) \pm 20$ MeV || PIDK(π^-) < -5
$X_d^+ \to \pi^+\pi^+\pi^-$	PIDK(π^+)	< 0
	PIDK(π^-)	< 10
	D_s veto	$m(\pi^+\pi^+_K\pi^-) \neq m(D_s) \pm 20$ MeV || PIDK(π^+) < -5
All tracks	hasRich	$= 1$

3.3 Simulation

Several Monte Carlo (MC) samples are used in the analysis for acceptance and background studies. A full list of them is given in Tab. 3.3. In each case, the decay model includes a mixture of non-interfering resonances contributing to the $X_s \to K\pi\pi$ or $X_d \to \pi\pi\pi$ bachelor system and a non-resonant (phase-space) component. All MC samples are generated using Pythia8, reconstructed using Reco14c, Reco15 and Reco16 for Run-I,15 and 16 data and selected using the same criteria as in data.

All samples indicated with 'Requested' have been requested in Dec. 17 and are expected to be available soon.

Table 3.3: List of simulated samples used in the analysis.

Decay	Event Type	Sim	Statistics				Filter
			11	12	15	16	
$B_s \to (D_s \to KK\pi)K\pi\pi$	13266007	08i	1.2 M	1.2 M	-	-	Generator Level
$B_s \to (D_s \to KK\pi)K\pi\pi$	13266008	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \to (D_s \to K\pi\pi)K\pi\pi$	13266058	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \to (D_s \to \pi\pi\pi)K\pi\pi$	13266038	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \to (D_s \to KK\pi)\pi\pi\pi$	13266006	08i	1.2 M	1.2 M	-	-	Generator Level
$B_s \to (D_s \to KK\pi)\pi\pi\pi$	13266068	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \to (D_s \to K\pi\pi)\pi\pi\pi$	13266088	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \to (D_s \to \pi\pi\pi)\pi\pi\pi$	13266078	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \to D_s^{*}\pi\pi\pi, D_s \to KK\pi$	13266201	08i	1.2 M	1.2 M	-	-	Generator Level

4 Yields determination

An extended unbinned maximum likelihood fit to the reconstructed B_s mass of the selected events is performed in order to determine the signal and background yields. The invariant mass $m(D_s h\pi\pi)$ is determined from a DTF constraining the mass of the D_s to the PDG value and the position of the PV. The probability density functions (PDFs) used to describe the signal and background components are described in the following.

Due to different mass resolutions, we perform the invariant mass fits simultaneously for each data-taking period and each trigger category. We further introduce four D_s final state categories: $D_s \to \phi\pi$, $D_s \to K^*(892)\pi$, $D_s \to \pi\pi\pi$ and $D_s \to Kh\pi$ to account for different signal purities. The $D_s \to Kh\pi$ category combines the two D_s decay channels with the lowest statistics: $D_s \to KK\pi$ (non-resonant) and $D_s \to K\pi\pi$. This amounts to 16 categories in total.

4.1 Signal model

The signal B_s-mass distribution of both $B_s^0 \to D_s K\pi\pi$ and $B_s^0 \to D_s\pi\pi\pi$ is modeled using a Johnson's SU function [32], which results from a variable transformation of a normal distribution to allow for asymmetric tails:

$$\mathcal{J}(x|\mu,\sigma,\nu,\tau) = \frac{e^{-\frac{1}{2}r^2}}{2\pi\cdot c\cdot\sigma\cdot\tau\cdot\sqrt{z^2+1}} \tag{4.1}$$

$$r = -\nu + \frac{\operatorname{asinh}(z)}{\tau} \tag{4.2}$$

$$z = \frac{x - (\mu - c\cdot\sigma\cdot e^{\tau}\sinh(\nu\cdot\tau))}{c\cdot\tau} \tag{4.3}$$

$$c = \frac{e^{\tau^2}-1}{2\sqrt{e^{\tau^2}\cdot\cosh(2\nu\cdot\tau)+1}}. \tag{4.4}$$

It is conveniently expressed in terms of the central moments up to order four: The mean of the distribution μ, the standard deviation σ, the skewness ν and the kurtosis τ. The tail parameters ν and τ are fixed to the values obtained by a fit to the invariant mass distribution of simulated events shown in Fig 4.1. To account for differences between

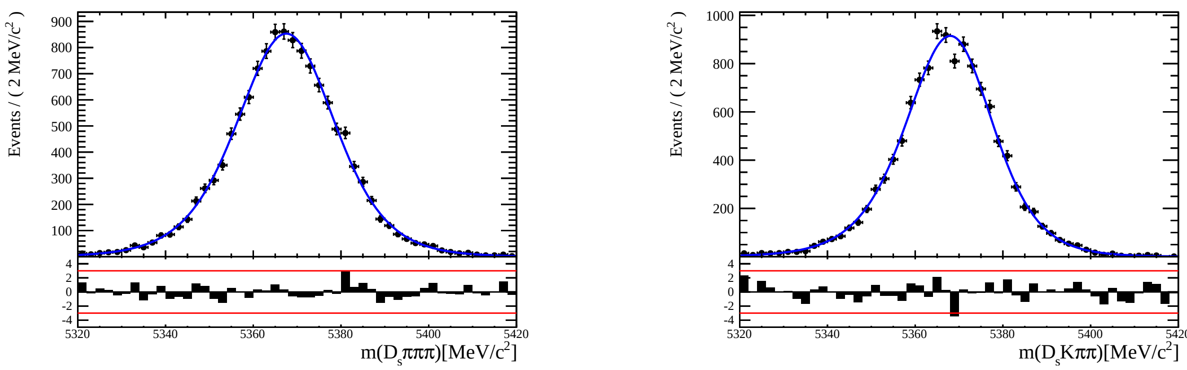

Figure 4.1: Invariant mass distributions of simulated (left) $B_s^0 \to D_s\pi\pi\pi$ and (right) $B_s^0 \to D_s K\pi\pi$ events. A fit with a Johnson's SU PDF is overlaid.

simulation and real data, linear scaling factors for the mean μ and width σ are determined in the fit to $B_s^0 \to D_s\pi\pi\pi$ data and later fixed in the fit to $B_s^0 \to D_sK\pi\pi$ decays. The scale factors are determined separately for each data-taking period and each trigger category.

4.2 Background models

After the full selection the following residual background components have to be accounted for:

Combinatorial background
The combinatorial background is described by a second order polynomial, whose parameters are determined, for each D_s final state separately, in the fit to data. For systematic studies an exponential PDF is used.

Peaking B_d background
Decays of B_d mesons into the $D_sh\pi\pi$ final state are described by the B_s signal PDF where the mean is shifted by the known mass difference $m_{B_s} - m_{B_d}$ [12].

Partially reconstructed background
Partially reconstructed $B_s^0 \to D_s^*\pi\pi\pi$ decays, with $D_s^* \to D_s\gamma$ or $D_s^* \to D_s\pi^0$, are expected to be peaking lower than signal in the $m(D_s\pi\pi\pi)$ spectrum with large tails due to the momentum carried away by the not reconstructed π^0 or γ. An empirical description for the shape of this contribution is derived from a $B_s^0 \to D_s^*\pi\pi\pi$ MC sample subject to the nominal $B_s^0 \to D_s\pi\pi\pi$ selection. Figure 4.2 (left) shows the respective reconstructed $m(D_s\pi\pi\pi)$ distribution. A sum of three bifurcated Gaussian functions is used to describe it. In the fit to data, all parameters are fixed to the ones obtained from MC except for the parameter which describes the width of the right tail of the distribution to account for data-simulation differences in mass resolution. The equivalent $B_s^0 \to D_s^*K\pi\pi$ component contributing to the $B_s^0 \to D_sK\pi\pi$ data sample is described by the same PDF with the right tail fixed to the $B_s^0 \to D_s\pi\pi\pi$ result.

Contributions from $B^0 \to D_s^*K\pi\pi$ decays are modeled with the $B_s^0 \to D_s^*K\pi\pi$ PDF shifted by $m_{B_s^0} - m_{B^0}$.

Misidentified background
A small fraction of $B_s \to D_s^-\pi^+\pi^+\pi^-$ and $B_s^0 \to D_s^*\pi^+\pi^+\pi^-$ decays, where one of the pions is misidentified as a kaon, contaminate the $B_s^0 \to D_sK^+\pi^+\pi^-$ sample. To determine the corresponding background shapes, we use simulated events passing the nominal selection except for the PID cuts on the bachelor π^+ tracks. The PIDCalib package is used to determine the p_T, η-dependent $\pi^+ \to K^+$ misidentification probability for each pion. We change the particle hypothesis from pion to kaon for the pion with the higher misidentification probability and recompute the invariant B_s^0 mass, $m(D_s^-\pi_K^+\pi^+\pi^-)$. Similarly, the invariant masses $m(\pi_K^+\pi^+\pi^-)$ and $m(\pi_K^+\pi^-)$ are recomputed and required to be within the considered phasespace region. The background distributions are shown in Fig. 4.2 (middle,right) and modeled by the sum of two Crystal Ball functions. The expected yield of misidentified $B_s^0 \to D_s\pi\pi\pi$ ($B_s^0 \to D_s^*\pi^+\pi^+\pi^-$) candidates in the $B_s^0 \to D_sK\pi\pi$ sample is computed by multiplying the fake rate (within the considered B_s mass range) of 0.47% (0.61%) by the $B_s^0 \to D_s\pi\pi\pi$ ($B_s^0 \to D_s^*\pi^+\pi^+\pi^-$) yield as

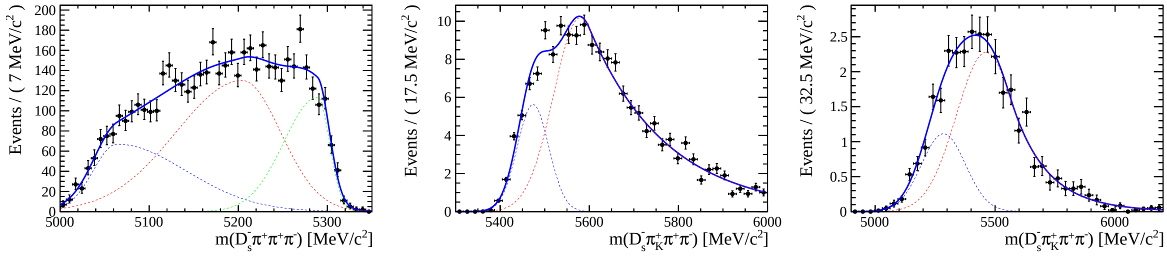

Figure 4.2: Left: Invariant mass distribution of simulated $B_s^0 \to D_s^* \pi\pi\pi$ events, where the γ/π^0 is excluded from the reconstruction. Middle: Invariant mass distribution of simulated $B_s^0 \to D_s \pi\pi\pi$ events, where one of the pions is reconstructed as a kaon taking the misidentification probability into account. Right: Invariant mass distribution for simulated $B_s^0 \to D_s^* \pi\pi\pi$ events, where the γ/π^0 from the D_s^* is excluded from reconstruction and one of the pions is reconstructed as a kaon taking the misidentification probability into account. The fitted PDF is shown in blue.

determined in the mass fit to the $B_s^0 \to D_s \pi\pi\pi$ data sample which is corrected for the PID(π^+)< 0 requirement. The $B_s^0 \to D_s^* \pi^+\pi^+\pi^-$ yield is additionally corrected for the efficiency of the cut $m(D_s K\pi\pi) > 5200\,\text{MeV}$ evaluated on MC. In the fit to data, the misidentified background yields are fixed to the predicted ones.

We consider the $B_s^0 \to D_s K\pi\pi$ and $B_s^0 \to D_s^* K\pi\pi$ components contributing to the $B_s^0 \to D_s \pi\pi\pi$ data sample to be negligible due to the low branching fractions and the tight PID cuts on the bachelor pions.

4.3 Results

Figure 4.3 shows the invariant mass distribution for $B_s^0 \to D_s \pi\pi\pi$ and $B_s^0 \to D_s K\pi\pi$ candidates passing all selection criteria. The projections for all categories of the simultaneous fit are shown in Appendix C together with the results for all fitted parameters. The integrated signal and background yields are listed in Tables 4.1 and 4.2.

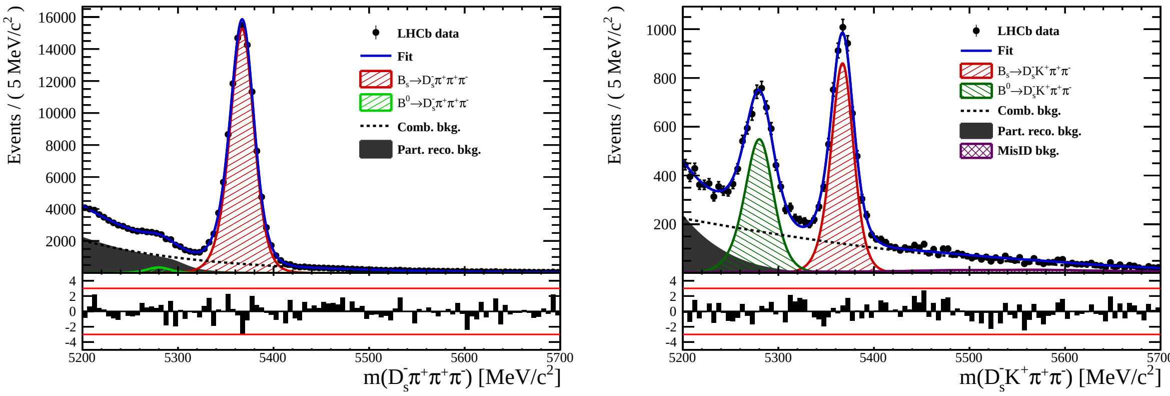

Figure 4.3: Invariant mass distribution of $B_s^0 \to D_s \pi\pi\pi$ (left) and $B_s^0 \to D_s K\pi\pi$ (right) candidates.

Table 4.1: Total signal and background yields for the $B_s \to D_s\pi\pi\pi$ sample (left) and signal yield for the different D_s final states contributing to the $B_s \to D_s\pi\pi\pi$ sample (right).

Component	Yield
$B_s \to D_s\pi\pi\pi$	101289 ± 348
$B^0 \to D_s\pi\pi\pi$	2318 ± 1763
Partially reconstructed bkg.	29817 ± 530
Combinatorial bkg.	52256 ± 603

D_s final state	Signal yield
$D_s^- \to \phi^0(1020)\pi^-$	34563 ± 197
$D_s^- \to K^{*0}(892)K^-$	28472 ± 189
$D_s^- \to (K^-h^+\pi^-)$	21047 ± 160
$D_s^- \to \pi^+\pi^-\pi^-$	17208 ± 145

Table 4.2: Total signal and background yields for the $B_s \to D_sK\pi\pi$ sample (left) and signal yield for the different D_s final states contributing to the $B_s \to D_sK\pi\pi$ sample (right).

Component	Yield
$B_s \to D_sK\pi\pi$	5125 ± 86
$B^0 \to D_sK\pi\pi$	4190 ± 92
Partially reconstructed bkg.	1707 ± 90
Misidentified bkg.	683 ± 0
Combinatorial bkg.	9686 ± 162

D_s final state	Signal yield
$D_s^- \to \phi^0(1020)\pi^-$	1613 ± 47
$D_s^- \to K^{*0}(892)K^-$	1527 ± 46
$D_s^- \to (K^-h^+\pi^-)$	1128 ± 40
$D_s^- \to \pi^+\pi^-\pi^-$	857 ± 37

5 Decay-time Resolution

The observed oscillation of B mesons is prone to dilution, if the detector resolution is of similar magnitude as the oscillation period. In the B_s^0 system, considering that the measured oscillation frequency of the B_s^0 [33] and the average LHCb detector resolution [34] are both $\mathcal{O}(50\,\text{fs}^{-1})$, this is the case. Therefore, it is crucial to correctly describe the decay time resolution in order to avoid a bias on the measurement of time dependent CP violation. Since the time resolution depends on the particular event, especially the decay time itself, the sensitivity on γ can be significantly improved by using an event dependent resolution model rather than an average resolution. For this purpose, we use the per-event decay time error that is estimated based on the uncertainty obtained from the global kinematic fit of the momenta and vertex positions (DTF) with additional constraints on the PV position and the D_s mass. In order to apply it correctly, it has to be calibrated. The raw decay time error distributions for $B_s^0 \to D_s\pi\pi\pi$ signal candidates are shown in Figure 5.1 for Run-I and Run-II data. Significant deviations between the two different data taking periods are observed due to the increase in center of mass energy from Run-I to Run-II, as well as (among others) new tunings in the pattern and vertex reconstruction. The decay time error calibration is consequently performed separately for both data taking periods.

For Run-I data, we use the calibration from the closely related $B_s^0 \to D_sK$ analysis which was performed on a data sample of prompt-D_s candidates combined with a random pion track from the PV. We verify the portability to our decay channel on MC.

For Run-II data, a new lifetime unbiased stripping line (LTUB) has been implemented which selects prompt-D_s candidates combined with random $K\pi\pi$ tracks from the PV.

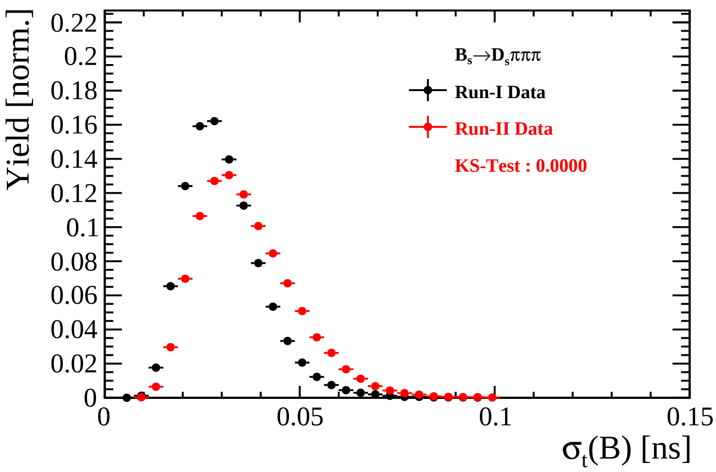

Figure 5.1: Distribution of the decay time error for $B_s^0 \to D_s\pi\pi\pi$ signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

5.1 Calibration for Run-I data

For simulated $B_s^0 \to D_s K\pi\pi$ events, the spread of the differences between reconstructed decay time and true decay time, $\Delta t = t - t_{true}$, is a direct measure of the decay time resolution. The sum of two Gaussian pdfs with common mean but different widths is used to fit the distribution of the decay time difference Δt as shown in Fig. 5.2. The effective damping of the CP amplitudes due to the finite time resolution is described by the dilution $\mathcal{D}$. In the case of infinite precision, there would be no damping and therefore $\mathcal{D} = 1$ would hold, while for a resolution that is much larger than the B_s^0 oscillation frequency, $\mathcal{D}$ would approach 0. For a double-Gaussian resolution model, the dilution is given by [35]

$$\mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s{}^2/2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s{}^2/2}, \tag{5.1}$$

where σ_1 and σ_2 are the widths of the Gaussians, f_1 is the relative fraction of events described by the first Gaussian relative to the second and Δm_s is the oscillation frequency of B_s^0 mesons. An effective single Gaussian width is calculated from the dilution as,

$$\sigma_{eff} = \sqrt{(-2/\Delta m_s{}^2) \ln \mathcal{D}}, \tag{5.2}$$

which converts the resolution into a single-Gaussian function with an effective resolution that causes the same damping effect on the magnitude of the B_s oscillation. For the Run-I $B_s^0 \to D_s K\pi\pi$ MC sample the effective average resolution is found to be $\sigma_{eff} = 39.1 \pm 0.3$ fs.

To analyze the relation between the per-event decay time error δ_t and the actual resolution σ_t, the simulated $B_s^0 \to D_s K\pi\pi$ sample is divided into equal-statistics slices of δ_t. For each slice, the effective resolution is determined as described above. Details of the fit results in each slice are shown in appendix D. Figure 5.2 shows the obtained values for σ_{eff} as a function of the per-event decay time error σ_t. To account for the variable binning, the bin values are not placed at the bin center but at the weighted mean of the respective per-event-error bin. A linear function is used to parametrize the distribution. The obtained values are

$$\sigma_{eff}^{MC}(\sigma_t) = (0.0 \pm 0.0)\,\text{fs} + (1.232 \pm 0.010)\,\sigma_t \tag{5.3}$$

where the offset is fixed to 0. For comparison, the calibration function found for $B_s^0 \to D_s K$ MC is also shown in Figure 5.2 [35]:

$$\sigma_{eff}^{D_sK,MC}(\sigma_t) = (0.0 \pm 0.0)\,\text{fs} + (1.201 \pm 0.013)\,\sigma_t. \tag{5.4}$$

Due to the good agreement between the scale factors for $B_s^0 \to D_s K$ and $B_s^0 \to D_s K\pi\pi$ MC, we conclude that the resolution calibration for $B_s^0 \to D_s K$ data:

$$\sigma_{eff}^{D_sK}(\sigma_t) = (10.26 \pm 1.52)\,\text{fs} + (1.280 \pm 0.042)\,\sigma_t \tag{5.5}$$

can be used for our analysis. The following calibration functions were used in the $B_s^0 \to D_s K$ analysis to estimate the systematic uncertainty on the decay-time resolution:

$$\sigma_{eff}^{D_sK}(\sigma_t) = (-0.568 \pm 1.570)\,\text{fs} + (1.243 \pm 0.044)\,\sigma_t \tag{5.6}$$

$$\sigma_{eff}^{D_sK}(\sigma_t) = (0.0 \pm 0.0)\,\text{fs} + (1.772 \pm 0.012)\,\sigma_t \tag{5.7}$$

The difference of the scale factors observed on $B_s^0 \to D_s K$ and $B_s^0 \to D_s K\pi\pi$ MC is assigned as additional systematic uncertainty.

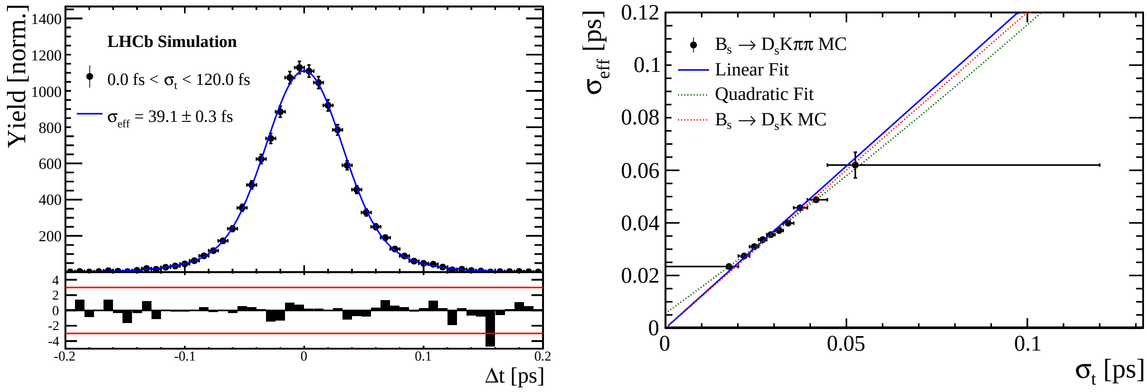

Figure 5.2: (Left) Difference of the true and measured decay time of $B_s^0 \to D_s K\pi\pi$ MC candidates. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for $B_s \to D_s K\pi\pi$ MC (Run-I). The fitted calibration curve is shown in blue.

5.2 Calibration for Run-II data

For the resolution calibration of Run-II data, a sample of promptly produced D_s candidates is selected using the B02DsKPiPiLTUBD2HHHBeauty2CharmLine stripping line. This lifetime-unbiased stripping line does not apply selection requirements related to lifetime or impact parameter, allowing for a study of the resolution. In order to reduce the rate of this sample it is pre-scaled in the stripping. Each D_s candidate is combined with a random kaon track and two random pion tracks which originate from the PV to obtain a sample of fake B_s candidates with a known true decay-time of $t_{true} = 0$. The difference of the measured decay time, t, of these candidates with respect to the true decay time is attributed to the decay time resolution.

The offline selection of the fake B_s candidates is summarized in Tab. 5.1. The selection is similar than the one for real data with the important difference that the D_s candidate is required to come from the PV by cutting on the impact parameter significance. Even after the full selection, a significant number of multiple candidates is observed. These are predominantly fake B_s candidates that share the same D_s candidate combined with different random tracks from the PV. We select one of these multiple candidates randomly which retains approximately 20% of the total candidates. The invariant mass distribution of the selected D_s candidates is shown in Fig. 5.3. To separate true D_s candidates from random combinations, the sPlot method is used to statistically subtract combinatorial background from the sample.

Figure 5.4 shows the sWeighted decay-time distribution for fake B_s candidates. Similar as in the previous section, the decay-time distribution is fitted with a double-Gaussian resolution model in slices of the per-event decay time error. Since some D_s candidates might actually originate from true B_s decays, the decay-time distribution of the fake B_s candidates might show a bias towards positive decay times. Therefore, we determine the decay-time resolution from the negative decay-time distribution only. Details of the fit results in each slice are shown in appendix D. The resulting calibration function:

$$\sigma_{eff}^{Data}(\sigma_t) = (9.7 \pm 1.7)\,\text{fs} + (0.915 \pm 0.040)\,\sigma_t \tag{5.8}$$

is in good agreement with the one obtained for the $B_s \to J/\psi\phi$ (Run-II) analysis [36].

$$\sigma_{eff}^{J/\psi\phi}(\sigma_t) = (12.25 \pm 0.33)\,\mathrm{fs} + (0.8721 \pm 0.0080)\,\sigma_t \tag{5.9}$$

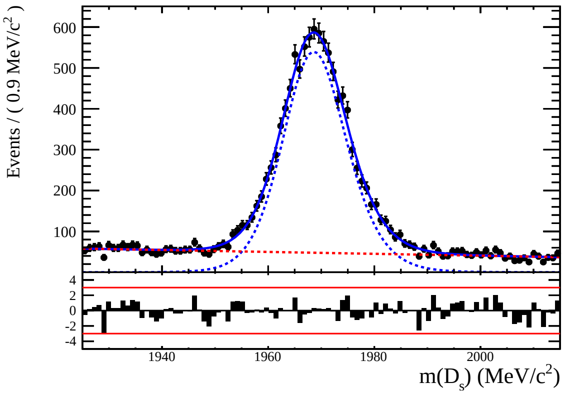

Figure 5.3: The invariant mass distribution for prompt D_s candidates.

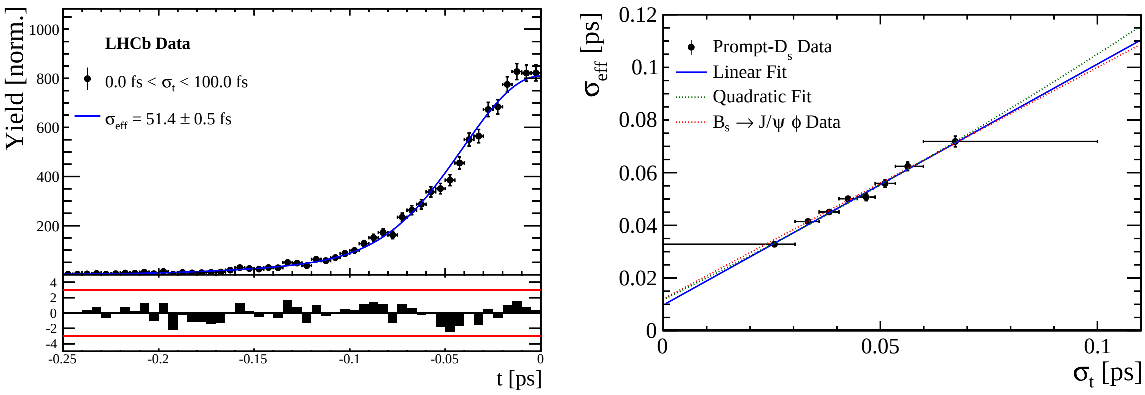

Figure 5.4: (Left) Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data). The fitted calibration curve is shown in blue.

Table 5.1: Offline selection requirements for fake B_s candidates from promptly produced D_s candidates combined with random prompt $K\pi\pi$ bachelor tracks.

	Description	Requirement
$B_s \to D_s K\pi\pi$	χ^2_{vtx}/ndof	< 8
	χ^2_{DTF}/ndof	< 15
	t	$< 0\,$ps
$D_s \to hhh$	χ^2_{vtx}/ndof	< 5
	DIRA	> 0.99994
	χ^2_{FD}	> 9
	p_T	> 1800 MeV
	χ^2_{IP}	< 9
	$\chi^2_{IP}(h)$	> 5
	Wrong PV veto	nPV $= 1 \mid\mid \min(\Delta\chi^2_{IP}) > 20$
$D_s^- \to KK\pi^-$	D^0 veto	$m(KK) < 1840$ MeV
	D^- veto	$m(K^+K^-_\pi\pi^-) \neq m(D^-) \pm 30$ MeV
	Λ_c veto	$m(K^+K^-_p\pi^-) \neq m(\Lambda_c) \pm 30$ MeV
$D_s^- \to \phi\pi^-$	$m(KK)$	$= m_\phi \pm 20$ MeV
	PIDK(K^+)	> -10
	PIDK(K^-)	> -10
	PIDK(π^-)	< 20
$D_s^- \to K^*(892)K^-$	$m(KK)$	$\neq m_\phi \pm 20$ MeV
	$m(K^+\pi^-)$	$= m_{K^*(892)} \pm 75$ MeV
	PIDK(K^+)	> -10
	PIDK(K^-)	> -5
	PIDK(π^-)	< 20
$D_s^- \to (KK\pi^-)_{NR}$	$m(KK)$	$\neq m_\phi \pm 20$ MeV
	$m(K^+\pi^-)$	$\neq m_{K^*(892)} \pm 75$ MeV
	PIDK(K^+)	> 5
	PIDK(K^-)	> 5
	PIDK(π^-)	< 10
$D_s \to \pi\pi\pi$	PIDK(h)	< 10
	PIDp(h)	< 10
	D^0 veto	$m(\pi^+\pi^-) < 1700$ MeV
$X_s \to K\pi\pi$	$\chi^2_{IP}(h)$	< 40
	PIDK(K)	> 10
	PIDK(π)	< 5
	isMuon(h)	False
All tracks	p_T	> 500 MeV

6 Acceptance

6.1 MC corrections

6.1.1 Truth matching of simulated candidates

We use the BackgroundCategory tool to truth match our simulated candidates. Candidates with background category (BKGCAT) 20 or 50 are considered to be true signal. Background category 60 is more peculiar since it contains both badly reconstructed signal candidates and ghost background. This is due to the fact that the classification algorithms identifies all tracks for which less than 70% of the reconstructed hits are matched to generated truth-level quantities as ghosts. In particular for signal decays involving many tracks (as in our case), this arbitrary cutoff is not 100% efficient. Moreover, the efficiency is expected to depend on the kinematics which would lead to a biased acceptance determination if candidates with BKGCAT= 60 would be removed. We therefore include BKGCAT= 60 and statistically subtract the ghost background by using the sPlot technique. The sWeights are calculated from a fit to the reconstructed B_s mass. The signal contribution is modeled as described in Sec. 4.1 and the background with a polynomial. The fit is performed simultaneously in two categories; the first includes events with BKGCAT = 20 or 50 and the second events with BKGCAT = 60. To account for the different mass resolution we use a different σ for each category, while the mean and the tail parameters are shared between them. The background component is only included for the second category.

A significant fraction of 8% of the true signal candidates are classified as ghosts, while only 20% of the events classified with BKGCAT= 60 are indeed genuine ghosts.

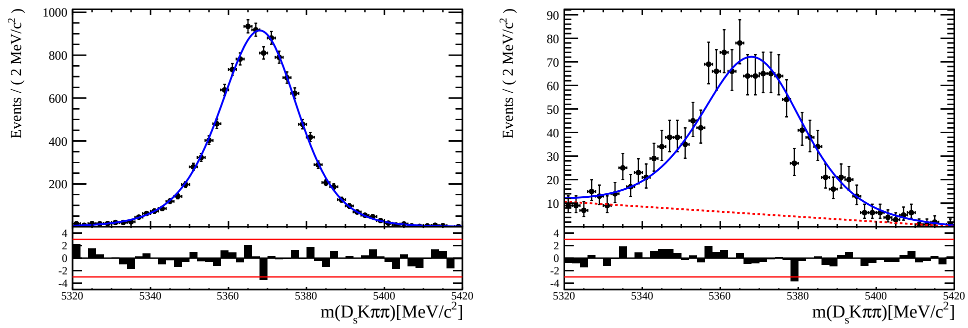

Figure 6.1: The reconstructed B_s mass distribution for simulated $B_s \to D_s K\pi\pi$ decays classified with BKGCAT = 20 or 50 (left) and BKGCAT = 60 (right) after the full selection.

6.1.2 Correction of data-simulation differences

For the evaluation of phase space efficiency and to a lesser extend also the decay-time efficiency we rely on simulated data as discussed in the following sections. A number of data-driven corrections are applied to the MC samples to account for known data-simulation differences. The MC sample is reweighted to match the two-dimensional p_T and η distribution observed on real data. An additional reweighting of the track multiplicity is applied on top of that. The distributions before and after reweighting are shown in Appendix H. We use the PIDCorr tool to correct the simulated PID responses based on PID calibration samples [37].

6.2 Decay-time acceptance

The decay-time distribution of the B_s^0 mesons is sculpted due to the geometry of the LHCb detector and the applied selection cuts, which are described in Section 3. In particular, any requirement on the flight distance, the impact parameter or the direction angle (DIRA) of the B_s^0 mesons, as well as the direct cut on the proper-time, will lead to a decay-time dependent efficiency $\epsilon(t)$.

We use a combination of control channels to derive the acceptance function $\epsilon(t)$, because for $B_s^0 \to D_s K\pi\pi$ decays the decay-time acceptance is strongly correlated with the *CP*-observables which we aim to measure. Therefore, extracting the *CP*-observables and the acceptance shape at the same time is not possible. A fit to the decay-time distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates is performed and the obtained acceptance shape is corrected for the small difference observed between the $B_s^0 \to D_s K\pi\pi$ and $B_s^0 \to D_s\pi\pi\pi$ MC samples. In addition, we include the control channel $B^0 \to D_s K\pi\pi$ to increase the statistical precision. A simultaneous fit to the four datasets ($B_s^0 \to D_s\pi\pi\pi$ data, $B^0 \to D_s K\pi\pi$ data, $B_s^0 \to D_s K\pi\pi$ MC and $B_s^0 \to D_s\pi\pi\pi$ MC) is performed to allow for a straightforward propagation of uncertainties. In each case, a PDF of the following form

$$\mathcal{P}(t, \delta t) = \left[e^{-\Gamma t} \cdot \cosh\left(\frac{\Delta\Gamma\, t'}{2}\right) \otimes \mathcal{R}(t - t^{'}, \delta t) \right] \cdot \epsilon(t), \tag{6.1}$$

is used to describe the decay-time distribution. For real data, the values for $\Gamma_{s,d}$ and $\Delta\Gamma_{s,d}$ are fixed to the latest HFAG results [38], while for simulated data, the generated values are used. A single Gaussian resolution function $\mathcal{R}(t - t^{'}, \delta t)$ is used where the decay-time error estimate is scaled with the respective calibration functions determined in Sec. 5. Each decay-time acceptance $\epsilon(t)$ is modeled using cubic splines, allowing for the analytical computation of the decay-time integrals appearing in the PDF [39]. The splines are parametrized by so-called knots $(t_0, t_1, \ldots, t_N)$ which determine their boundaries. Two knots are located by default at the lower and upper edge of the interval allowed for the decay time, the remaining ones are chosen such that there is an approximately equal amount of data in-between two consecutive knots. In the basis of cubic b-splines, $b_i(t)$, the acceptance is then constructed as:

$$\epsilon(t) = \sum_{i=0}^{N} v_i\, b_i(t) \tag{6.2}$$

where the spline coefficients v_i are determined from the fit. We fix coefficient v_{N-1} to unity in order to normalize the overall acceptance function. To stabilize the upper decay-time acceptance, v_N is fixed by a linear extrapolation from the two previous coefficients:

$$v_N = v_{N-1} + \frac{v_{N-2} - v_{N-1}}{t_{N-2} - t_{N-1}} \cdot (t_N - t_{N-1}). \tag{6.3}$$

It was found that at least $N = 6$ knots are necessary for a sufficient fit quality.

Three distinct splines are used in the following combinations to describe the acceptances for the four datasets:

- $B_s^0 \to D_s K\pi\pi$ MC: $\epsilon^{MC}_{D_s K\pi\pi}(t)$
- $B_s^0 \to D_s \pi\pi\pi$ MC: $\epsilon^{MC}_{D_s \pi\pi\pi}(t) = R(t) \cdot \epsilon^{MC}_{D_s K\pi\pi}(t)$
- $B_s^0 \to D_s \pi\pi\pi$ data: $\epsilon^{Data}_{D_s \pi\pi\pi}(t) = R(t) \cdot \epsilon^{Data}_{D_s K\pi\pi}(t)$
- $B^0 \to D_s K\pi\pi$ data: $\epsilon^{Data}_{D_s K\pi\pi}(t)$

where $\epsilon^{MC}_{D_s K\pi\pi}(t)$ represents the acceptance in $B_s^0 \to D_s K\pi\pi$ MC, $R(t)$ represents the ratio of acceptances in $B_s^0 \to D_s \pi\pi\pi$ and $B_s^0 \to D_s K\pi\pi$ MC and the final acceptance in $B_s^0 \to D_s K\pi\pi$ data is represented by $\epsilon^{Data}_{D_s K\pi\pi}(t)$.

The acceptances are determined separately for each data-taking period and each trigger category as discussed in more detail in Appendix E. The fit results are shown in Figs. 6.2 to 6.5 and the fitted parameters are summarized in Tables 6.1 to 6.4.

As currently there are no Run-II MC samples available, we use the Run-I MC samples also for the Run-II fits. An alternative approach would be to fit only the $B^0 \to D_s K\pi\pi$ data sample in order to remove the MC dependency. The final strategy will be fixed, depending on the MC availability, during the review.

Table 6.1: Time acceptance parameters for events in category [Run-I,L0-TOS].

Knot position	Coefficient	$B_s^0 \to D_s K\pi\pi$ data	$B_s^0 \to D_s K\pi\pi$ MC	Ratio
0.4	v_0	0.592 ± 0.038	0.542 ± 0.021	0.972 ± 0.056
0.8	v_1	0.805 ± 0.057	0.781 ± 0.033	0.915 ± 0.064
1.6	v_2	0.852 ± 0.077	0.917 ± 0.051	1.034 ± 0.080
2.5	v_3	1.117 ± 0.042	1.108 ± 0.029	0.955 ± 0.045
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.898 (interpolated)	0.906 (interpolated)	1.039 (interpolated)

Table 6.2: Time acceptance parameters for events in category [Run-I,L0-TIS].

Knot position	Coefficient	$B_s^0 \to D_s K\pi\pi$ data	$B_s^0 \to D_s K\pi\pi$ MC	Ratio
0.4	v_0	0.417 ± 0.038	0.415 ± 0.021	0.948 ± 0.077
0.8	v_1	0.623 ± 0.060	0.654 ± 0.035	0.873 ± 0.080
1.6	v_2	0.901 ± 0.097	0.976 ± 0.061	0.909 ± 0.087
2.5	v_3	1.095 ± 0.052	1.076 ± 0.035	1.003 ± 0.051
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.917 (interpolated)	0.933 (interpolated)	0.998 (interpolated)

Table 6.3: Time acceptance parameters for events in category [Run-II,L0-TOS].

Knot position	Coefficient	$B_s^0 \to D_s K\pi\pi$ data	$B_s^0 \to D_s K\pi\pi$ MC	Ratio
0.4	v_0	0.568 ± 0.028	0.496 ± 0.015	0.965 ± 0.044
0.8	v_1	0.787 ± 0.043	0.738 ± 0.024	0.892 ± 0.049
1.6	v_2	0.899 ± 0.061	0.943 ± 0.039	0.984 ± 0.059
2.5	v_3	1.079 ± 0.030	1.093 ± 0.021	0.979 ± 0.030
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.931 (interpolated)	0.919 (interpolated)	1.018 (interpolated)

Table 6.4: Time acceptance parameters for events in category [Run-II,L0-TIS].

Knot position	Coefficient	$B_s^0 \to D_s K\pi\pi$ data	$B_s^0 \to D_s K\pi\pi$ MC	Ratio
0.4	v_0	0.389 ± 0.009	0.506 ± 0.015	0.908 ± 0.031
0.8	v_1	0.592 ± 0.013	0.744 ± 0.024	0.896 ± 0.035
1.6	v_2	0.798 ± 0.052	0.965 ± 0.041	0.927 ± 0.054
2.5	v_3	1.111 ± 0.035	1.112 ± 0.023	0.941 ± 0.039
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.903 (interpolated)	0.902 (interpolated)	1.052 (interpolated)

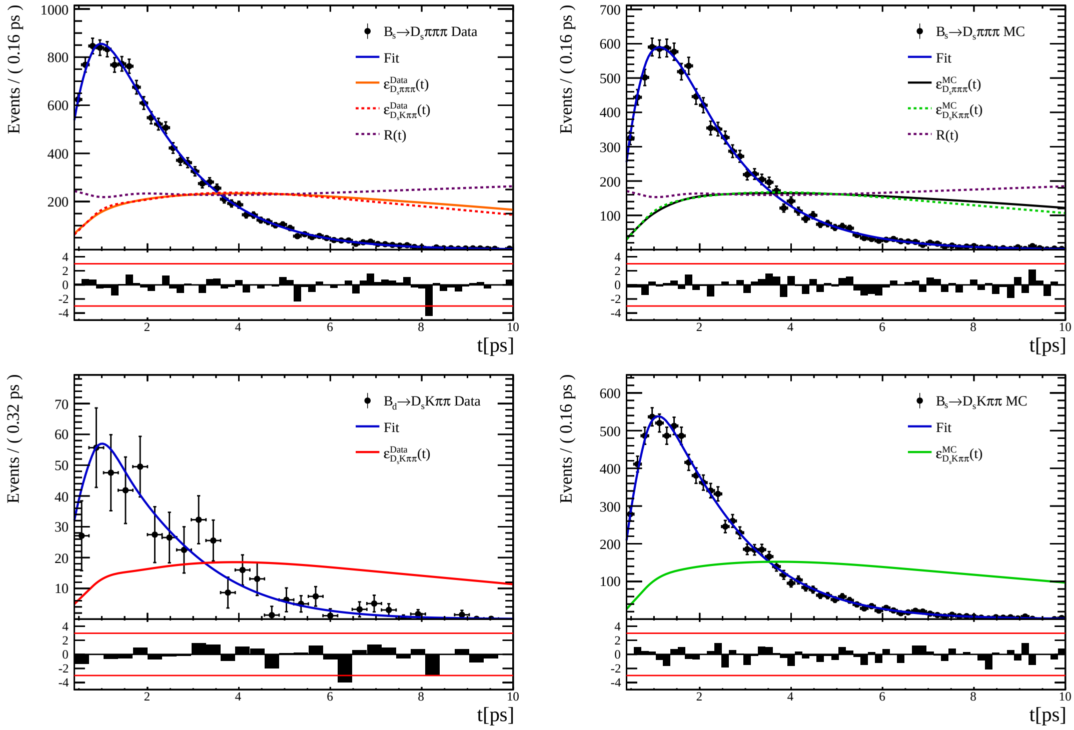

Figure 6.2: Decay-time fit projections for $B_s^0 \to D_s\pi\pi\pi$ data (top-left), $B_s^0 \to D_s\pi\pi\pi$ MC (top-right), $B^0 \to D_sK\pi\pi$ data (bottom-left) and $B_s^0 \to D_sK\pi\pi$ MC (bottom-right) in category [Run-I,L0-TOS]. The respective acceptance function is overlaid in an arbitrary scale.

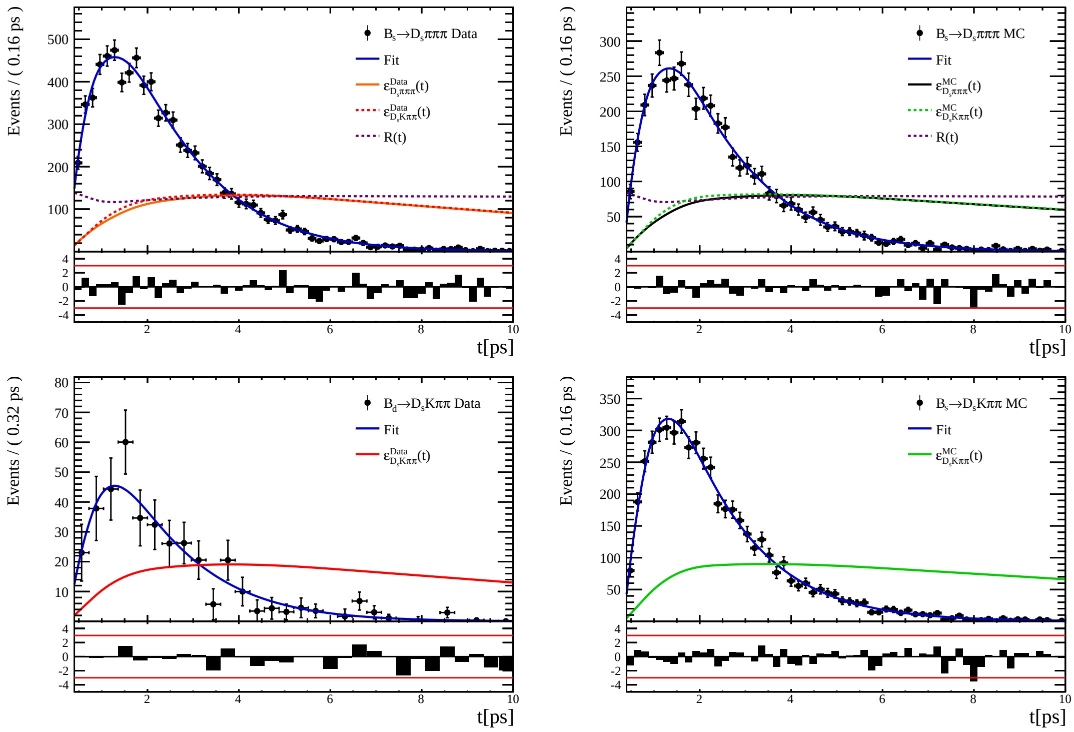

Figure 6.3: Decay-time fit projections for $B_s^0 \to D_s\pi\pi\pi$ data (top-left), $B_s^0 \to D_s\pi\pi\pi$ MC (top-right), $B^0 \to D_sK\pi\pi$ data (bottom-left) and $B_s^0 \to D_sK\pi\pi$ MC (bottom-right) in category [Run-I,L0-TIS]. The respective acceptance function is overlaid in an arbitrary scale.

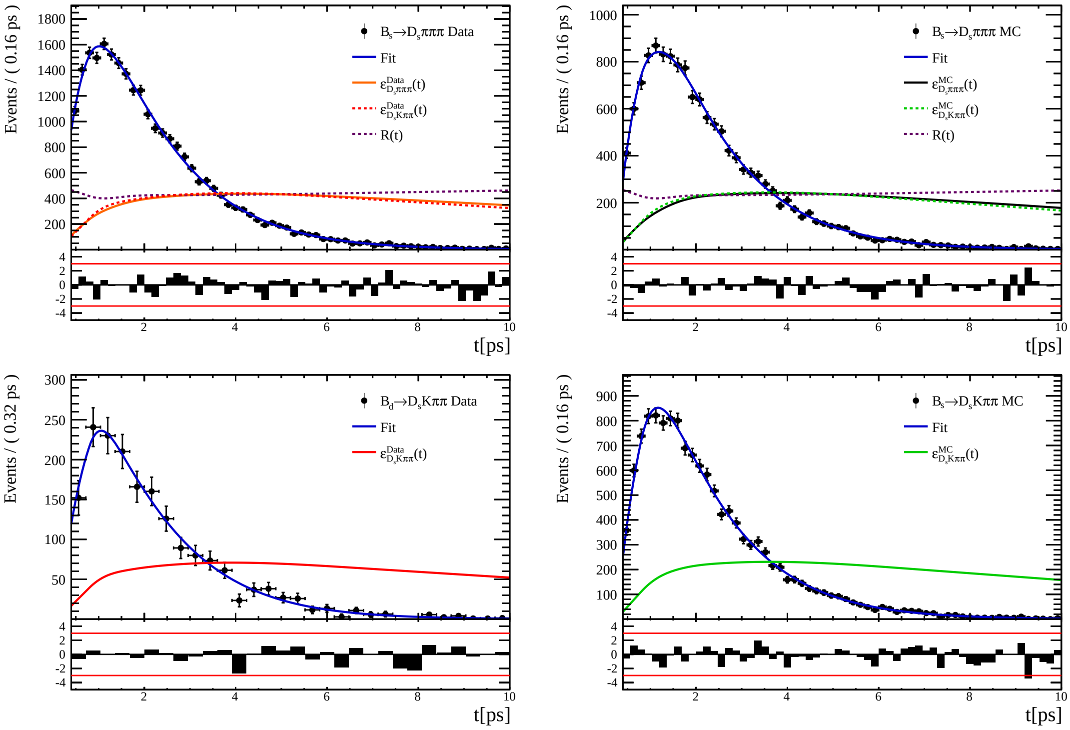

Figure 6.4: Decay-time fit projections for $B_s^0 \to D_s\pi\pi\pi$ data (top-left), $B_s^0 \to D_s\pi\pi\pi$ MC (top-right), $B^0 \to D_s K\pi\pi$ data (bottom-left) and $B_s^0 \to D_s K\pi\pi$ MC (bottom-right) in category [Run-II,L0-TOS]. The respective acceptance function is overlaid in an arbitrary scale.

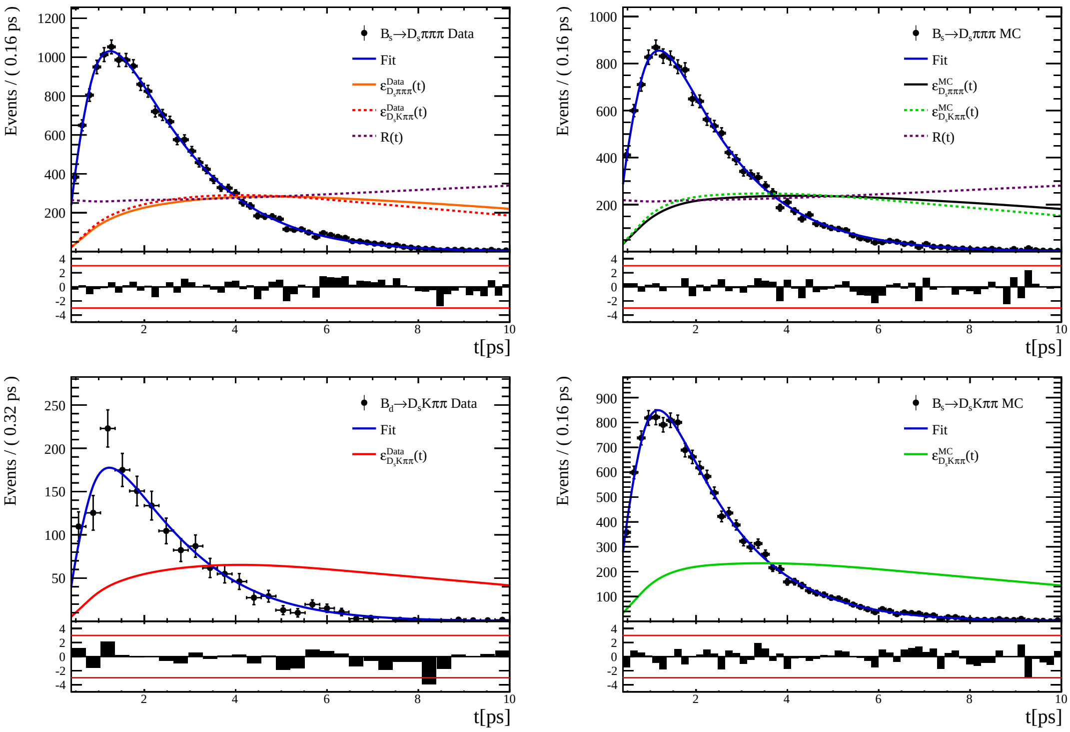

Figure 6.5: Decay-time fit projections for $B_s^0 \to D_s\pi\pi\pi$ data (top-left), $B_s^0 \to D_s\pi\pi\pi$ MC (top-right), $B^0 \to D_s K\pi\pi$ data (bottom-left) and $B_s^0 \to D_s K\pi\pi$ MC (bottom-right) in category [Run-II,L0-TIS]. The respective acceptance function is overlaid in an arbitrary scale.

6.3 Phase space acceptance

The signal PDF used for the full time-dependent amplitude fit can be written in terms of the differential decay rate from Equation 2.29 as

$$\mathcal{P}(\mathbf{x}, t, g, f) = \frac{\left(\frac{\mathrm{d}\Gamma(\mathbf{x},t,q,f)}{\mathrm{d}t\,\mathrm{d}\Phi_4}\right) \cdot \epsilon(\mathbf{x}) \cdot \epsilon(t)}{\int \sum_{q,f} \left(\frac{\mathrm{d}\Gamma(\mathbf{x},t,q,f)}{\mathrm{d}t\,\mathrm{d}\Phi_4}\right) \cdot \epsilon(\mathbf{x}) \cdot \epsilon(t)\,\mathrm{d}t\,\mathrm{d}\Phi_4} \tag{6.4}$$

where $\epsilon(\mathbf{x})$ is the phase-space efficiency. Note that the efficiency in the numerator appears as an additive constant in the $\log \mathcal{L}$ that does not depend on any fit parameters such that it can be ignored. However, the efficiency function still enters via the normalization integrals. In contrast to the time integrals which can be performed analytically as discussed in Sec. 6.2, the phase-space integrals are determined numerically. For this purpose, we use simulated events generated with EVTGEN, pass them through the full detector simulation and apply the same selection criteria as for data in order to perform the MC integrals. As an example, the integral of the total $b \to c$ amplitude squared can be approximated as

$$\int \left|\mathcal{A}_f^c(\mathbf{x})\right|^2 \epsilon(\mathbf{x})\,\mathrm{d}\Phi_4 \approx \frac{1}{N_{\mathrm{MC}}} \sum_{k}^{N_{\mathrm{MC}}} \frac{\left|\mathcal{A}_f^c(\mathbf{x_k})\right|^2}{\left|A'(\mathbf{x_k})\right|^2} \tag{6.5}$$

where A' labels the amplitude model used for the generation and x_k is the k-th MC event. As a result, the phase-space efficiency can be included in the fit without explicitly modeling it.

Disclaimer: At the moment there is only a small Run-I MC sample available where a DecFile (EventType: 13266007) was used from which we were not able to reproduce the generator pdf A'. We can therefore not follow our preferred procedure described above. An alternative, provisionally method is briefly described in the following.

We use a BDTG to map the five-dimensional phase space to an one-dimensional distribution [40]. The BDTG is trained to learn the differences between the selected MC and a generator level MC sample. As discriminating variables, five invariant mass combinations are used as shown in Fig. 6.6. Based on the classifier output distributions, shown in Fig. 6.7, an efficiency as function of the BDTG response is derived.

A large toy MC sample is generated (500 k events) according to a preliminary amplitude model $A'(\mathbf{x})$ and for each event a weight, depending on the BDTG response, is assigned to account for the efficiency variation across phase space. This reweighted toy MC sample is then effectively distributed as $A'(\mathbf{x}) \cdot \epsilon(\mathbf{x})$ and can be used to calculate the normalization integrals in Equation 6.5. Figure 6.8 compares the phase space efficiency obtained from the reweigted toy MC sample with the 'true' efficiency given by the ratio of selected and generated MC events. A fairly good agreement is observed in all dimensions.

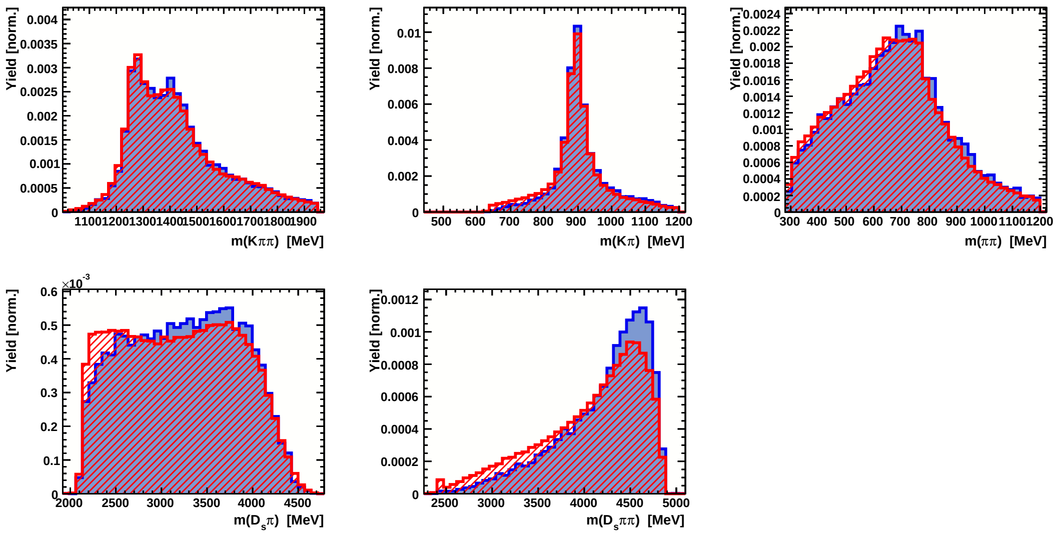

Figure 6.6: Discriminating variables used to train the BDTG. The selected MC sample is shown in blue and the generator MC sample in red.

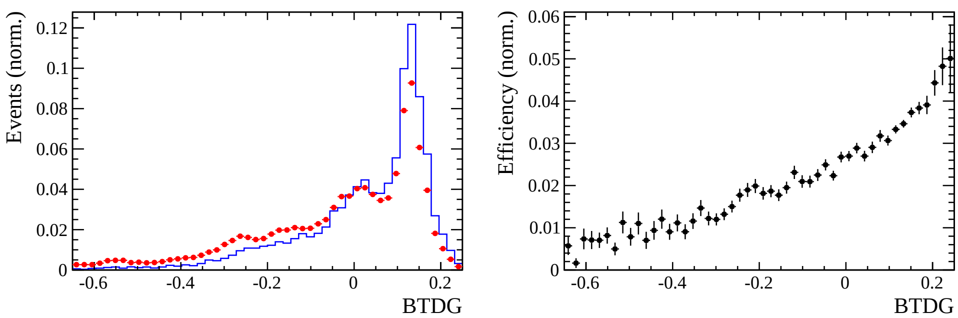

Figure 6.7: Left: Output distributions of the BDTG for the simulated MC sample (blue) and the generator level sample (red). Right: Phase space efficiency as function of the BDTG response computed as the ratio of selected and generated decays.

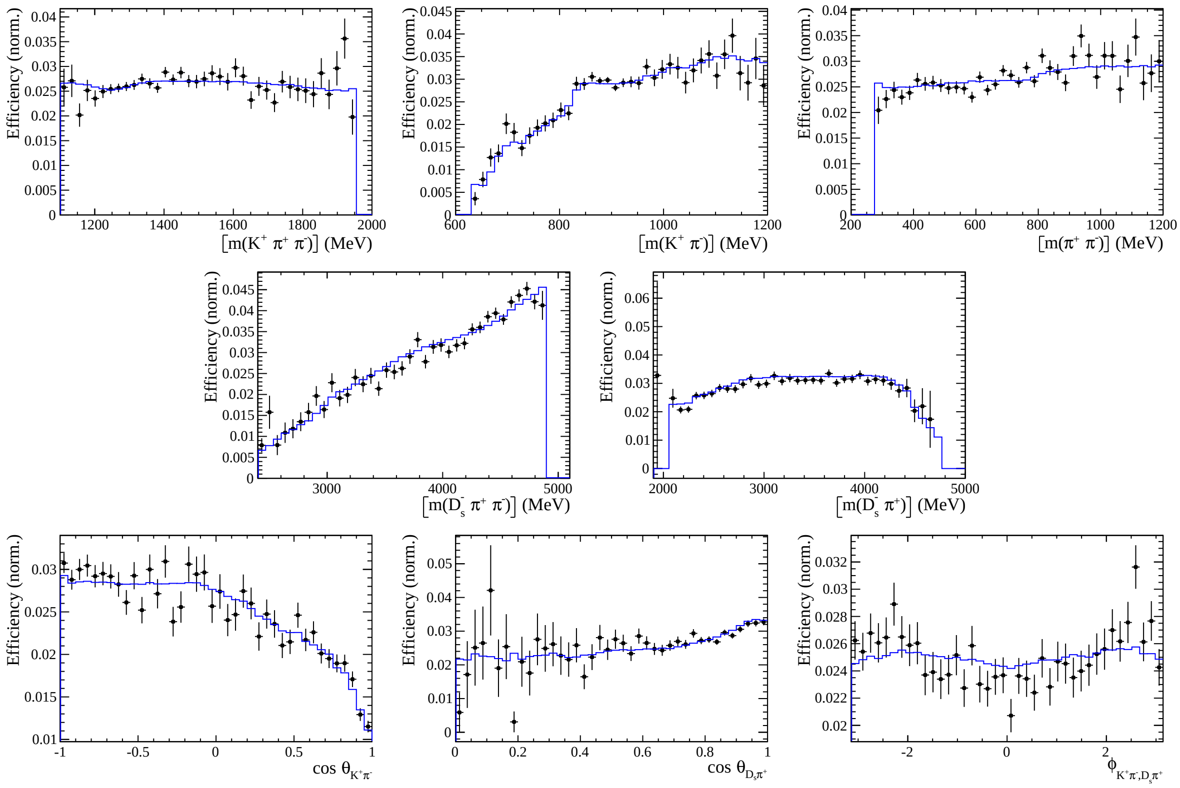

Figure 6.8: Efficiency variation as a function of the phase-space variables obtained from the ratio of selected and generated MC events (data points) and efficiency obtained from a reweigted toy MC sample (blue).

7 Flavour Tagging

To identify the initial flavour state of the B_s^0 meson, a number of flavour tagging algorithms are used that either determine the flavour of the non-signal b-hadron produced in the event (opposite site, OS [41]) or use particles produced in the fragmentation of the signal candidate $B_s^0/\overline{B}_s^0$ (same side, SS [42]). For the same side, the algorithm searching for the charge of an additional kaon that accompanies the fragmentation of the signal candidate is used (SS-Kaon). For the opposite site, four different taggers are chosen: The algorithms that use the charge of an electron or a muon from semi-leptonic B decays (OS-e,μ), the tagger that uses the charge of a kaon from a b $\to$ c $\to$ s decay chain (OS-nnetKaon) and the algorithm that determines the $B_s^0/\overline{B}_s^0$ candidate flavour from the charge of a secondary vertex, reconstructed from the OS b decay product (OS-VtxCharge).

Every tagging algorithm is prone to misidentify the signal candidate at a certain mistag rate ω. This might be caused by particle misidentification, flavour oscillation of the neutral opposite site B-meson or by tracks that are wrongly picked up from the underlying event. An imperfect determination of the B_s^0 production flavor dilutes the observed CP asymmetry by a factor $D_{tag} = 1 - 2\omega$. This means that the statistical precision, with which the CP asymmetry can be measured, scales as the inverse square root of the effective tagging efficiency:

$$\epsilon_{eff} = \epsilon_{tag}(1 - 2\omega)^2, \tag{7.1}$$

where ϵ_{tag} is the fraction of tagged candidates.

For each $B_s^0/\overline{B}_s^0$ candidate, the tagging algorithms provide, besides a flavour tag $q = 1, -1, 0$ (for an initial B_s^0, $\overline{B}_s^0$ or no tag), a prediction for the mistag probability η based on the output of multivariate classifiers. These are trained on simulated samples of flavour specific control channels ($B_s^0 \to D_s^- \pi^+$ (SS algorithm) and $B^+ \to J/\psi K^+$ (OS algorithms)) and are optimized for highest ϵ_{eff} on data. Utilizing flavour-specific final states, the estimated mistag η of each tagger has to be calibrated to match the actual mistag probability ω. For the calibration, a linear model

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - \langle\eta\rangle), \tag{7.2}$$

is used where $\langle\eta\rangle$ is the average estimated mistag probability. A perfectly calibrated tagger would lead to $\omega(\eta) = \eta$ and one would expect $p_1 = 1$ and $p_0 = \langle\eta\rangle$. Due to the different interaction cross-sections of oppositely charged particles, the tagging calibration parameters depend on the initial state flavour of the B_s^0. Therefore, the flavour asymmetry parameters Δp_0, Δp_1 and $\Delta\epsilon_{tag}$ are introduced.

7.1 OS tagger combination

First, the OS electron, muon, neural net kaon and the secondary vertex charge taggers are individually calibrated and then combined into a single OS-Combo tagger using the EspressoPerformanceMonitor tool. We choose the flavour specific decay $B_s \to D_s\pi\pi\pi$ as calibration mode since it is very similar to the signal decay $B_s \to D_s K\pi\pi$. The calibration is performed separately for Run-I and Run-II data. Where available the latest Run-II tuning is used for Run-II data, otherwise the Run-I tuning of the taggers is used. Figures 7.1 and 7.2 show the fitted calibration functions and Tables 7.1 and 7.2 list the measured tagging performances.

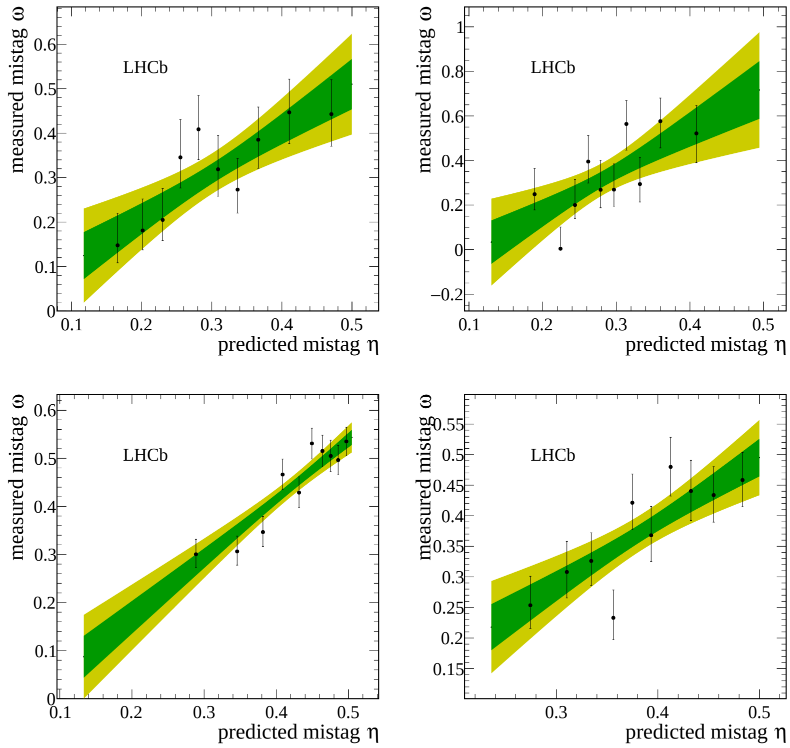

Figure 7.1: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run-I. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

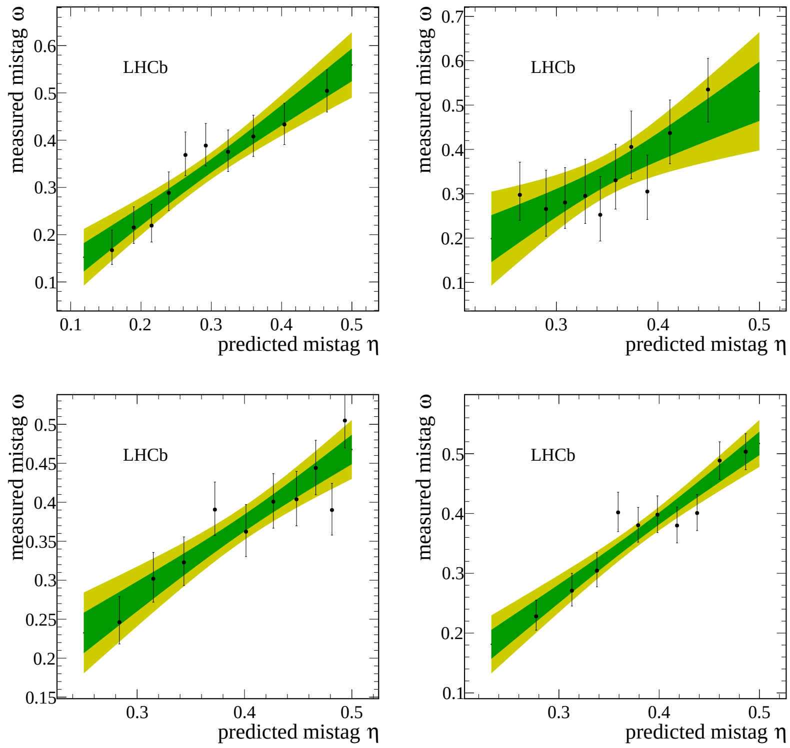

Figure 7.2: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run-II. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

Table 7.1: The flavour tagging performances for the used OS taggers for Run-I data.

Tagger	ϵ	ω	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$
OS μ	$(8.775 \pm 0.207)\%$	$(28.935 \pm 0.180(\text{stat}) \pm 2.288(\text{cal}))\%$	$(1.558 \pm 0.045(\text{stat}) \pm 0.338(\text{cal}))\%$
OS e	$(3.191 \pm 0.129)\%$	$(28.778 \pm 0.366(\text{stat}) \pm 3.636(\text{cal}))\%$	$(0.575 \pm 0.031(\text{stat}) \pm 0.197(\text{cal}))\%$
OS K NN	$(32.099 \pm 0.342)\%$	$(38.405 \pm 0.094(\text{stat}) \pm 1.152(\text{cal}))\%$	$(1.726 \pm 0.033(\text{stat}) \pm 0.343(\text{cal}))\%$
Vertex Charge	$(21.797 \pm 0.302)\%$	$(35.672 \pm 0.092(\text{stat}) \pm 1.480(\text{cal}))\%$	$(1.790 \pm 0.034(\text{stat}) \pm 0.370(\text{cal}))\%$

Table 7.2: The flavour tagging performances for the used OS taggers for Run-II data.

Tagger	ϵ	ω	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$
OS μ	$(8.904 \pm 0.146)\%$	$(30.119 \pm 0.119(\text{stat}) \pm 1.477(\text{cal}))\%$	$(1.408 \pm 0.029(\text{stat}) \pm 0.209(\text{cal}))\%$
OS e	$(3.284 \pm 0.091)\%$	$(32.834 \pm 0.166(\text{stat}) \pm 2.367(\text{cal}))\%$	$(0.387 \pm 0.013(\text{stat}) \pm 0.107(\text{cal}))\%$
OS K NN	$(16.709 \pm 0.191)\%$	$(35.960 \pm 0.075(\text{stat}) \pm 1.076(\text{cal}))\%$	$(1.317 \pm 0.021(\text{stat}) \pm 0.202(\text{cal}))\%$
Vertex Charge	$(20.605 \pm 0.208)\%$	$(34.625 \pm 0.077(\text{stat}) \pm 0.967(\text{cal}))\%$	$(1.948 \pm 0.028(\text{stat}) \pm 0.245(\text{cal}))\%$

7.2 Tagging performance

The OS-Combo and SS-Kaon taggers are calibrated simultaneously by fitting the $B_s \to D_s\pi\pi\pi$ decay-time distribution as discussed in Sec. 9. The predicted mistag probabilities η_{OS} and η_{SS}, shown Fig. 7.3 for $B_s \to D_s\pi\pi\pi$ and $B_s \to D_sK\pi\pi$ data, are included as per-event observables, effectively giving a larger weight to the events that have a lower mistag probability. The tagger responses are combined into a single response on an event-by-event basis during the fit. Tables 7.3 and 7.4 report the tagging performances for the OS and SS combination considering three mutually exclusive categories of tagged events: OS only, SS only and both OS and SS.

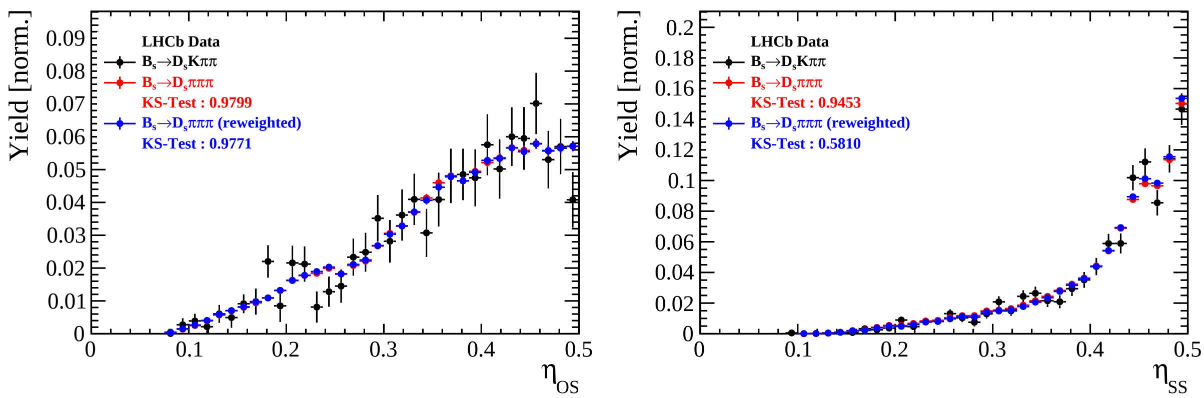

Figure 7.3: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \to D_sK\pi\pi$ (black) and $B_s^0 \to D_s\pi\pi\pi$ (red) data samples.

Table 7.3: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-I data.

$B_s \to D_s\pi\pi\pi$	$\epsilon_{tag}[\%]$	$\langle\omega\rangle[\%]$	$\epsilon_{eff}[\%]$
Only OS	14.75 ± 0.11	39.03 ± 0.82	1.27 ± 0.17
Only SS	35.46 ± 0.18	44.15 ± 0.64	1.10 ± 0.19
Both OS-SS	32.92 ± 0.30	37.18 ± 0.76	3.48 ± 0.35
Combined	83.12 ± 0.37	40.48 ± 0.72	5.85 ± 0.43

Table 7.4: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-II data.

$B_s \to D_s\pi\pi\pi$	$\epsilon_{tag}[\%]$	$\langle\omega\rangle[\%]$	$\epsilon_{eff}[\%]$
Only OS	10.92 ± 0.05	36.56 ± 0.56	1.09 ± 0.07
Only SS	43.80 ± 0.11	42.44 ± 0.37	1.99 ± 0.15
Both OS-SS	26.08 ± 0.14	34.87 ± 0.45	3.44 ± 0.17
Combined	80.80 ± 0.19	39.20 ± 0.42	6.52 ± 0.23

8 Production and Detection Asymmetries

8.1 B_s Production Asymmetry

The production rates of b and $\bar{b}$ hadrons in pp collisions are not expected to be identical, therefore this effect must be taken into account when computing CP asymmetries. The production asymmetry for B_s mesons is defined as:

$$A_p(B_s^0) = \frac{\sigma(\bar{B}_s^0) - \sigma(B_s^0)}{\sigma(\bar{B}_s^0) + \sigma(B_s^0)} \tag{8.1}$$

where σ are the corresponding production cross-section. This asymmetry was measured by LHCb in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$ by means of a time-dependent analysis of $B_s \to D_s^- \pi^+$ decays [43]. The results in bins of p_T and η of the B_s meson are shown in Table 8.1. To correct for the different kinematics of $B_s \to D_s^- \pi^+$ and $B_s^0 \to D_s K\pi\pi$ decays, the measured B_s production asymmetries $A_p(p_T, \eta)$ are folded with the sWeighted p_T, η distribution of our signal channel. The resulting effective production asymmetries are:

$$A_p(B_s^0)_{2011} = (-0.506 \pm 1.90)\% \tag{8.2}$$
$$A_p(B_s^0)_{2012} = (\;\;0.164 \pm 1.30)\% \tag{8.3}$$
$$A_p(B_s^0)_{\text{Run-I}} = (-0.045 \pm 1.04)\%. \tag{8.4}$$

As for Run-II data no measurement is available yet, we determine the production asymmetry from $B_s \to D_s\pi\pi\pi$ data together with the tagging parameters.

Table 8.1: B_s production asymmetries in kinematic bins for 2011 and 2012 data. [43]

p_{T} [GeV/c]	η	$A_{\text{P}}(B_s^0)_{\sqrt{s}=7\text{ TeV}}$	$A_{\text{P}}(B_s^0)_{\sqrt{s}=8\text{ TeV}}$
$(2.00, 7.00)$	$(2.10, 3.00)$	$0.0166 \pm 0.0632 \pm 0.0125$	$0.0412 \pm 0.0416 \pm 0.0150$
$(2.00, 7.00)$	$(3.00, 3.30)$	$0.0311 \pm 0.0773 \pm 0.0151$	$-0.0241 \pm 0.0574 \pm 0.0079$
$(2.00, 7.00)$	$(3.30, 4.50)$	$-0.0833 \pm 0.0558 \pm 0.0132$	$0.0166 \pm 0.0391 \pm 0.0092$
$(7.00, 9.50)$	$(2.10, 3.00)$	$0.0364 \pm 0.0479 \pm 0.0068$	$0.0482 \pm 0.0320 \pm 0.0067$
$(7.00, 9.50)$	$(3.00, 3.30)$	$0.0206 \pm 0.0682 \pm 0.0127$	$0.0983 \pm 0.0470 \pm 0.0155$
$(7.00, 9.50)$	$(3.30, 4.50)$	$0.0058 \pm 0.0584 \pm 0.0089$	$-0.0430 \pm 0.0386 \pm 0.0079$
$(9.50, 12.00)$	$(2.10, 3.00)$	$-0.0039 \pm 0.0456 \pm 0.0121$	$0.0067 \pm 0.0303 \pm 0.0063$
$(9.50, 12.00)$	$(3.00, 3.30)$	$0.1095 \pm 0.0723 \pm 0.0179$	$-0.1283 \pm 0.0503 \pm 0.0171$
$(9.50, 12.00)$	$(3.30, 4.50)$	$0.1539 \pm 0.0722 \pm 0.0212$	$-0.0500 \pm 0.0460 \pm 0.0104$
$(12.00, 30.00)$	$(2.10, 3.00)$	$-0.0271 \pm 0.0336 \pm 0.0061$	$-0.0012 \pm 0.0222 \pm 0.0050$
$(12.00, 30.00)$	$(3.00, 3.30)$	$-0.0542 \pm 0.0612 \pm 0.0106$	$0.0421 \pm 0.0416 \pm 0.0162$
$(12.00, 30.00)$	$(3.30, 4.50)$	$-0.0586 \pm 0.0648 \pm 0.0150$	$0.0537 \pm 0.0447 \pm 0.0124$

8.2 $K^-\pi^+$ Detection Asymmetry

The presented measurement of the CKM-angle γ using $B_s^0 \to D_s K\pi\pi$ decays is sensitive to a possible charge asymmetry of the kaon. Kaons are known to have a nuclear cross-section which is asymmetrically dependent on the sign of their charge. It is indispensable to determine the charge asymmetry of the kaon, as fitting without taking this effect into account would introduce a 'fake' CP violation. Instead of determining the single track detection asymmetry of a kaon, it is found that the combined two track asymmetry of a kaon-pion pair is much easier to access [44] . Therefore, the two track asymmetry defined as

$$A^{det}(K^-\pi^+) = \frac{\epsilon^{det}(K^-\pi^+) - \epsilon^{det}(K^+\pi^-)}{\epsilon^{det}(K^-\pi^+) + \epsilon^{det}(K^+\pi^-)}, \tag{8.5}$$

is used.

This asymmetry can be measured from the difference in asymmetries in the $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K_S^0\pi^+$ modes [45]:

$$\begin{aligned} A^{det}(K^-\pi^+) =& \frac{N(D^+ \to K^-\pi^+\pi^+) - N(D^- \to K^+\pi^-\pi^-)}{N(D^+ \to K^-\pi^+\pi^+) + N(D^- \to K^+\pi^-\pi^-)} \\ &- \frac{N(D^+ \to K_S^0\pi^+) - N(D^- \to K_S^0\pi^-)}{N(D^+ \to K_S^0\pi^+) + N(D^- \to K_S^0\pi^-)} \\ &- A(K^0), \end{aligned} \tag{8.6}$$

where possible CP violation in the $D^+ \to K_S^0\pi^+$ mode is predicted to be smaller than 10^{-4} in the Standard Model [46]. The asymmetry in the neutral kaon system, $A(K^0)$, has to be taken into account as a correction.

We use a dedicated LHCb tool to determine $A^{det}(K^-\pi^+)$ for all data taking periods used in this analysis. A detailed description can be found in [45]. The tool provides large calibration samples of $D^\pm \to K^\pm\pi^\pm\pi^\pm$ and $D^\pm \to K_S^0\pi^\pm$ decays, which are used to determine the asymmetry following Eq. 8.6. Several weighting steps are performed to match the kinematics of the calibration samples to our signal decay sample:

First, weights are assigned to the $K^\pm$ and $\pi^\pm$ of the $D^\pm \to K^\pm\pi^\pm\pi^\pm$ sample, using p, η of the $K^\pm$ and p_T, η of the $\pi^\pm$ from our $B_s^0 \to D_s K\pi\pi$ signal decay. Then, weights are assigned to the $D^\pm$ (p_T, η) and the $\pi^\pm$ (p_T) of the $D^\pm \to K_S^0\pi^\pm$ sample to match the corresponding, weighted distributions of the $D^\pm \to K^\pm\pi^\pm\pi^\pm$ sample. In a last step, weights are assigned to match the bachelor pions ϕ distributions between the two calibration samples.

After the samples are weighted, fits are performed to the invariant $m(K^-\pi^+\pi^+)/m(K^+\pi^-\pi^-)$ and $m(K_S^0\pi^+)/m(K_S^0\pi^-)$ distributions to determine $A^{det}(K^-\pi^+)$. The PDFs used to describe the invariant mass distributions consist of gaussian functions for the signal component and exponentials describing the residual background.

The detection asymmetry is determined separately for every year and (since it is a charge asymmetry effect) magnet polarity. Serving as an example for Run-I and Run-II, the fits used to determine $N(D^+ \to K^-\pi^+\pi^+)/N(D^- \to K^+\pi^-\pi^-)$ and $N(D^+ \to K_S^0\pi^+)/N(D^- \to K_S^0\pi^-)$ for 2011, magnet up data and 2015, magnet up data are shown in Fig. 8.1 and 8.2 respectively. The obtained values of $A^{det}(K^-\pi^+) + A(K^0)$ for all years and polarities are shown in Table 8.2.

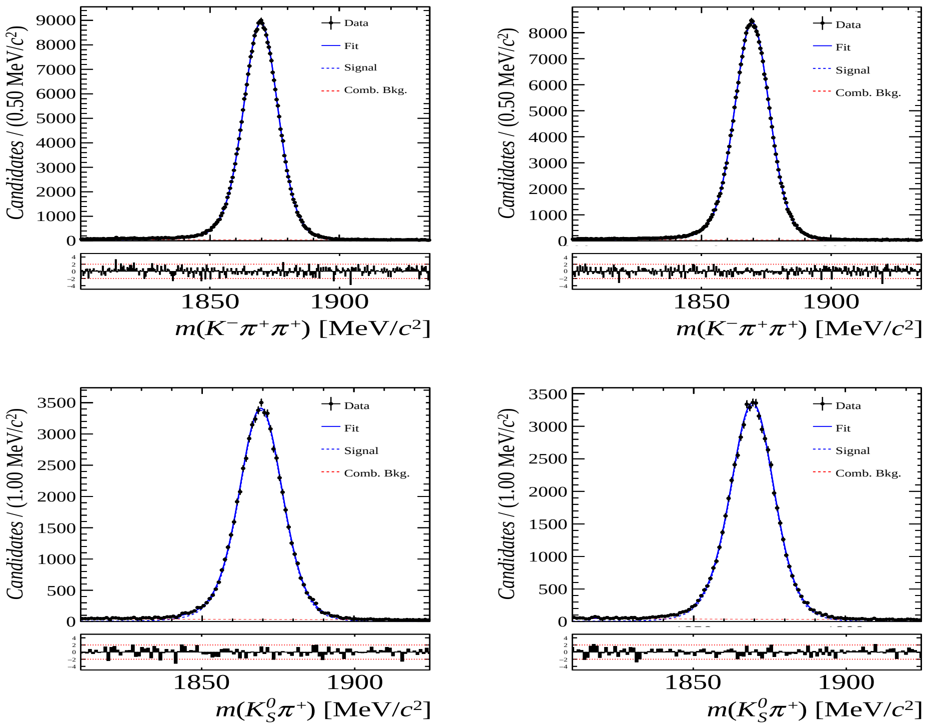

Figure 8.1: Distributions of the invariant mass of (top) $D^\pm \to K^\pm\pi^\pm\pi^\pm$ and (bottom) $D^\pm \to K^0_S\pi^\pm$ candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

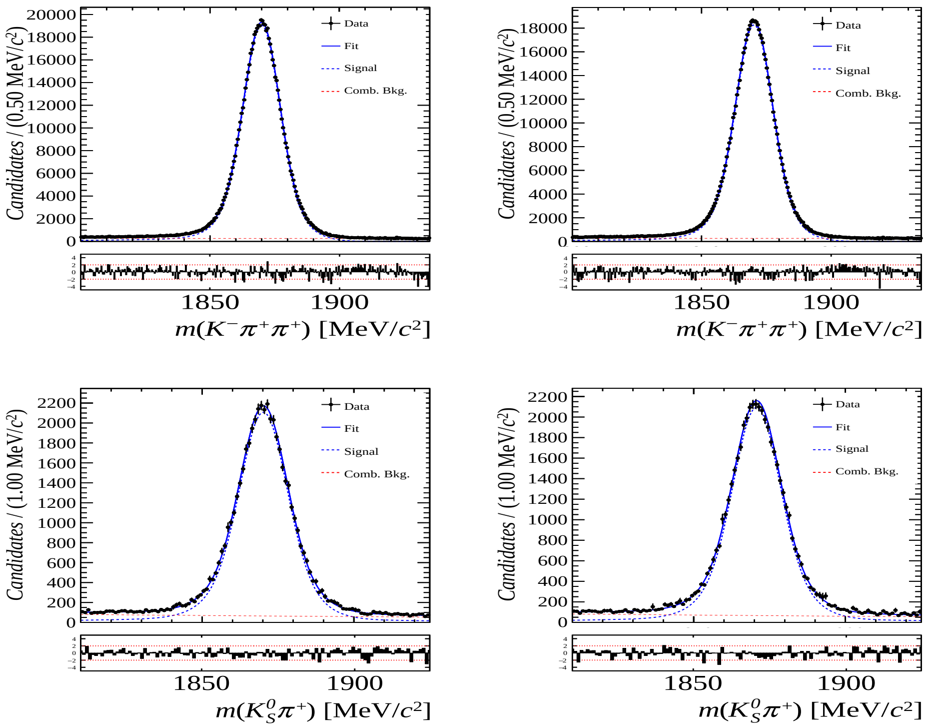

Figure 8.2: Distributions of the invariant mass of (top) $D^\pm \to K^\pm\pi^\pm\pi^\pm$ and (bottom) $D^\pm \to K^0_S\pi^\pm$ candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

Data sample	$A^{det}(K^-\pi^+) + A(K^0)$ [%]
Run-I	
2011, mag. up	-2.01 $\pm$ 0.32
2011, mag. down	-0.16 $\pm$ 0.28
2011, average	-1.09 $\pm$ 0.21
2012, mag. up	-0.90 $\pm$ 0.20
2012, mag. down	-1.01 $\pm$ 0.22
2012, average	-0.96 $\pm$ 0.15
Run-II	
mag. up	-1.16 $\pm$ 0.34
mag. down	-0.65 $\pm$ 0.27
average	-0.91 $\pm$ 0.22

Table 8.2: Summary of the $K^-\pi^+$ detection asymmetry obtained from the fits to the Run-I and Run-II calibration samples.

9 Decay-time fit

This section covers the (phase space integrated) decay-time fits to $B_s^0 \to D_s h\pi\pi$ data. We use the sFit technique [47] to statistically subtract the background, leaving only the signal PDF to describe the decay-time. The sWeights are calculated based on the fit to the reconstructed B_s mass distribution described in Sec. 4. The signal PDF is conditional on the tagging decisions q_i, the mistag estimates η_i (i =OS,SS) and the decay-time error δt:

$$\mathcal{P}(t|\delta t, q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS}) \propto \left[p(t'|q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS}) \otimes \mathcal{R}(t-t', \delta t)\right] \cdot \epsilon(t) \tag{9.1}$$

where $p(t|q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS})$ is given by Equation 2.7 taking the tagging dilution into account. The decay-time acceptance $\epsilon(t)$ (Sec. 6) and the Gaussian time-resolution function $\mathcal{R}(t-t', \delta t)$ (Sec. 5) are fixed to the values obtained by the dedicated studies. We fix the values of Γ_s and $\Delta\Gamma_s$ to the latest HFAG results [38].

The unbinned maximum likelihood fits are performed simultaneously in four categories: [Run-I,L0-TOS],[Run-I,L0-TIS],[Run-II,L0-TOS] and [Run-II,L0-TIS].

9.1 Fit to $B_s^0 \to D_s\pi\pi\pi$ data

Since the decay $B_s^0 \to D_s\pi\pi\pi$ is flavour specific, the *CP* coefficients can be fixed to $C = 1$ and $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$. The fit determines the calibration parameters for the OS-Combo and SS-Kaon taggers, the B_s^0 production asymmetry for Run-II data as well as the mixing frequency Δm_s. Table 9.1 summarizes the fitted parameters. The sWeighted decay-time distribution and the time-dependent asymmetry A_{mix} between mixed and unmixed B_s^0 candidates are shown in Fig. 9.1 along with the fit projections.

Table 9.1: Parameters determined from a fit to the $B_s \to D_s\pi\pi\pi$ decay-time distribution. The uncertainties are statistical and systematic, respectively.

Fit Parameter	Run-I	Run-II
p_0^{OS}	0.397 ± 0.010 ± 0.010	0.367 ± 0.005 ± 0.009
p_1^{OS}	0.908 ± 0.087 ± 0.090	0.772 ± 0.046 ± 0.063
Δp_0^{OS}	0.030 ± 0.011 ± 0.002	0.006 ± 0.006 ± 0.000
Δp_1^{OS}	0.010 ± 0.094 ± 0.015	0.085 ± 0.054 ± 0.003
$\epsilon_{\text{tag}}^{OS}$ [%]	47.667 ± 0.365 ± 0.032	37.018 ± 0.181 ± 0.009
$\Delta\epsilon_{\text{tag}}^{OS}$ [%]	0.087 ± 1.249 ± 0.093	0.185 ± 0.582 ± 0.127
p_0^{SS}	0.443 ± 0.008 ± 0.004	0.426 ± 0.004 ± 0.004
p_1^{SS}	0.974 ± 0.110 ± 0.066	0.800 ± 0.041 ± 0.050
Δp_0^{SS}	-0.019 ± 0.009 ± 0.001	-0.017 ± 0.005 ± 0.000
Δp_1^{SS}	0.057 ± 0.125 ± 0.018	0.038 ± 0.048 ± 0.004
$\epsilon_{\text{tag}}^{SS}$ [%]	0.684 ± 0.003 ± 0.000	0.699 ± 0.002 ± 0.000
$\Delta\epsilon_{\text{tag}}^{SS}$ [%]	-0.003 ± 0.012 ± 0.001	-0.003 ± 0.006 ± 0.000
A_P [%]	-0.045 (fixed)	-0.150 ± 0.618 ± 0.090
Δm_s [ps^{-1}]	xx.xx ± 0.009 ± 0.006	

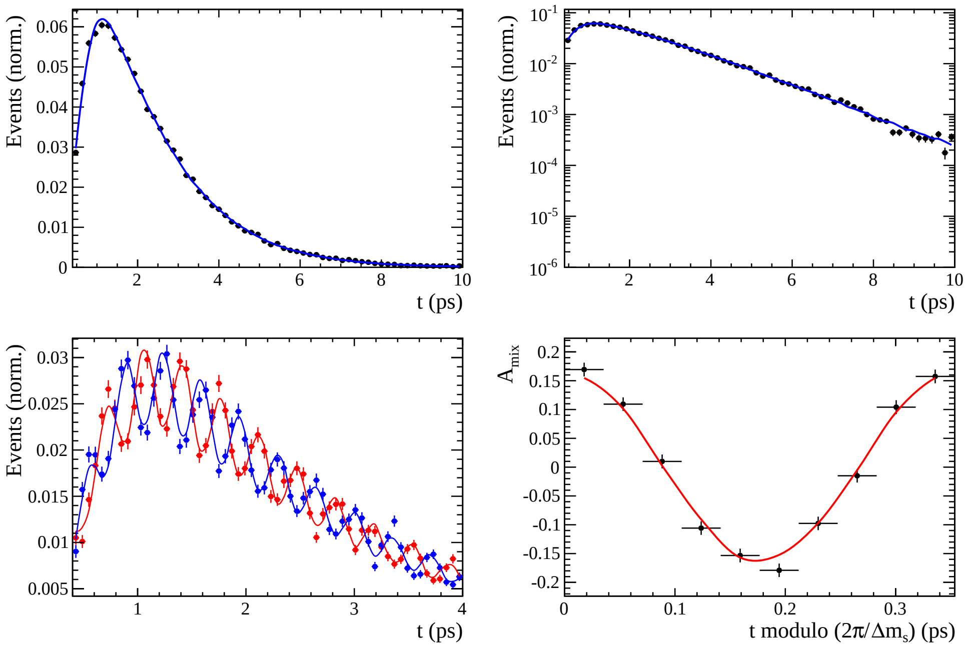

Figure 9.1: Top: Flavour averaged decay time distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates. Bottom-left: Tagged decay time distribution of mixed (red) and unmixed (blue) signal candidates. Bottom-right: Time-dependent asymmetry A_{mix} between mixed and unmixed B_s^0 candidates folded into one oscillation period.

9.2 Fit to $B_s^0 \to D_sK\pi\pi$ data

The measured *CP* coefficients $C, D_f, D_{\bar{f}}, S_f$ and $S_{\bar{f}}$ extracted from a fit to the $B_s \to D_sK\pi\pi$ decay-time distribution are reported in Table 9.2. The fit projection is shown in Fig. 9.2. We included Gaussian-constrains for the tagging calibration parameters with the central values and uncertainties determined in Sec. 9.1.

The *CP* coefficients will be converted to the observables $r, \kappa, \delta, \gamma$ using the Gamma-Combo package after unblinding.

Currently the mixing frequency is fixed to the HFAG value. We intend to update the fit after unblinding our result from the $B_s^0 \to D_s\pi\pi\pi$ fit since our precision is significantly higher.

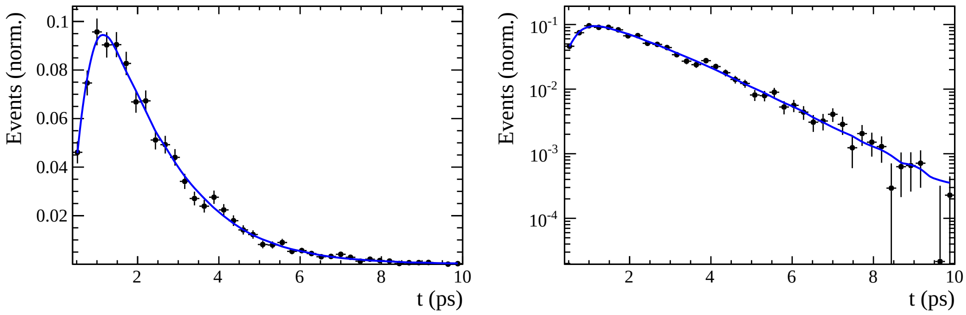

Figure 9.2: Decay-time distribution of $B_s^0 \to D_s K\pi\pi$ signal candidates with the fit projection overlaid.

Table 9.2: *CP* coefficients determined from a fit to the $B_s \to D_s K\pi\pi$ decay-time distribution. The uncertainties are statistical and systematic, respectively.

Fit Parameter	Value
C	x.xx $\pm$ 0.11 $\pm$ 0.02
D	x.xx $\pm$ 0.29 $\pm$ 0.08
$\bar{D}$	x.xx $\pm$ 0.27 $\pm$ 0.09
S	x.xx $\pm$ 0.16 $\pm$ 0.05
$\bar{S}$	x.xx $\pm$ 0.16 $\pm$ 0.04

10 Time-dependent amplitude fit

The signal PDF used for the full time-dependent fit is defined as

$$\mathcal{P}(\mathbf{x}, t|\delta t, q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS}) \propto \left[p(\mathbf{x}, t^{'}|q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS}) \otimes \mathcal{R}(t - t^{'}, \delta t)\right] \cdot \epsilon(t) \quad (10.1)$$

where $p(\mathbf{x}, t|q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS})$ is given the differential decay rate in Equation 2.29 taking the tagging dilution into account. The phase space efficiency $\epsilon(\mathbf{x})$ is only included in the normalization of $\mathcal{P}(\mathbf{x}, t|\delta t, q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS})$ as discussed in Sec 6.3. The model selection of the amplitude components is described in the following Section. The remaining fitting strategy is exactly the same as for the decay-time fits, see Sec. 9.

10.1 Signal Model Construction

The light meson spectrum comprises multiple resonances which are expected to contribute to $B_s \to D_s K\pi\pi$ decays as intermediate states. Apart from clear contributions coming from resonances such as $K_1(1270)$, $K_1(1400)$, $\rho(770)$ and $K^*(892)^0$, the remaining structure is impossible to infer due to the cornucopia of broad, overlapping and interfering resonances within the phase space boundary. We follow the LASSO [48, 49] approach to limit the model complexity in two steps.

First, we fit the time-integrated and flavour averaged phase-space distribution of $B_s \to D_s K\pi\pi$ decays. In this case, a single total amplitude can be used:

$$\mathcal{A}_f^{eff}(\mathbf{x}) = \sum_i a_i^{eff} A_i(\mathbf{x}) \quad (10.2)$$

which effectively describes the incoherent superposition of the $b \to c$ and $b \to u$ amplitudes:

$$|A_f^{eff}(\mathbf{x})|^2 = |A_f^c(\mathbf{x})|^2 + |A_f^u(\mathbf{x})|^2. \quad (10.3)$$

This significantly simplifies the fitting procedure and allows us to include the whole pool of considered intermediate state amplitudes A_i which can be found in Appendix G. The LASSO penalty term added to the likelihood function

$$-2\,\log\mathcal{L} \to -2\,\log\mathcal{L} + \lambda\sum_i \sqrt{\int |a_i^{eff} A_i(\mathbf{x})|^2\,\mathrm{d}\Phi_4}, \quad (10.4)$$

shrinks the amplitude coefficients towards zero. The amount of shrinkage is controlled by the parameter λ, to be tuned on data. Higher values for λ encourage sparse models, *i.e.* models with only a few non-zero amplitude coefficients. The optimal value for λ is found by minimizing the Bayesian information criteria [50] (BIC),

$$\mathrm{BIC}(\lambda) = -2\,\log\mathcal{L} + r\,\log N_{\mathrm{Sig}}, \quad (10.5)$$

where N_{Sig} is the number of signal events and r is the number of amplitudes with a decay fraction above a certain threshold. The fit fractions are defined as

$$F_i \equiv \frac{\int \left|a_i^{eff}\,A_i(\mathbf{x})\right|^2\,\mathrm{d}\Phi_4}{\int \left|\mathcal{A}_f^{eff}(\mathbf{x})\right|^2\,\mathrm{d}\Phi_4}, \quad (10.6)$$

and are a measure of the relative strength between the different transitions. Figure 10.1(left) shows the distribution of BIC values obtained by scanning over λ where we choose the decay fraction threshold to be 0.5%. At the optimal value of $\lambda = 50$, the set of amplitudes with a decay fraction above the threshold are considered further for step two of the model selection. The selected amplitudes and their fractions are summarized in Table 10.1. The fit projections are shown in Fig. 10.2.

In Stage 2, the LASSO procedure is again performed by fitting the full time-dependent amplitude PDF. The components selected by Stage 1 are included for both $b \to c$ and $b \to u$ transitions and the likelihood is extended as follows:

$$-2\log\mathcal{L} \to -2\log\mathcal{L} + \lambda \sum_i \sqrt{\int |a_i^c A_i(\mathbf{x})|^2 \, \mathrm{d}\Phi_4} + \lambda \sum_i \sqrt{\int |a_i^u A_i(\mathbf{x})|^2 \, \mathrm{d}\Phi_4} \quad (10.7)$$

Figure 10.1(right) shows a plot of the complexity factor λ, against the resulting BIC values. The final set of $b \to c$ and $b \to u$ amplitudes is selected using the optimal value of $\lambda = 28$, and is henceforth called the LASSO model.

Table 10.1: Fit fractions of the amplitudes selected by Stage 1 of the model selection procedure.

Decay channel	Fraction [%]
$B_s \to K(1)(1270)^+(\to K^*(892)^0(\to K^+\pi^-)\pi^+)\,D_s^-$	8.56 ± 1.43
$B_s \to K(1)(1400)^+(\to K^*(892)^0(\to K^+\pi^-)\pi^+)\,D_s^-$	43.72 ± 2.80
$B_s \to K(1460)^+(\to K^*(892)^0(\to K^+\pi^-)\pi^+)\,D_s^-$	3.25 ± 0.69
$B_s \to K^*(1410)^+(\to K^*(892)^0(\to K^+\pi^-)\pi^+)\,D_s^-$	15.33 ± 1.13
$B_s \to (D_s^-\pi^+)_P\,K^*(892)^0(\to K^+\pi^-)$	4.63 ± 0.69
$B_s \to K^*(1410)^+(\to \rho(770)^0(\to \pi^+\pi^-)K^+)\,D_s^-$	5.58 ± 0.62
$B_s \to (D_s^-K^+)_P\,\rho(770)^0(\to \pi^+\pi^-)$	1.49 ± 0.40
$B_s \to K(1)(1270)^+(\to K(0)^*(1430)^0(\to K^+\pi^-)\pi^+)\,D_s^-$	4.72 ± 0.54
$B_s \to K(1)(1270)^+(\to \rho(770)^0(\to \pi^+\pi^-)K^+)\,D_s^-$	14.20 ± 1.56
Sum	101.47 ± 3.86

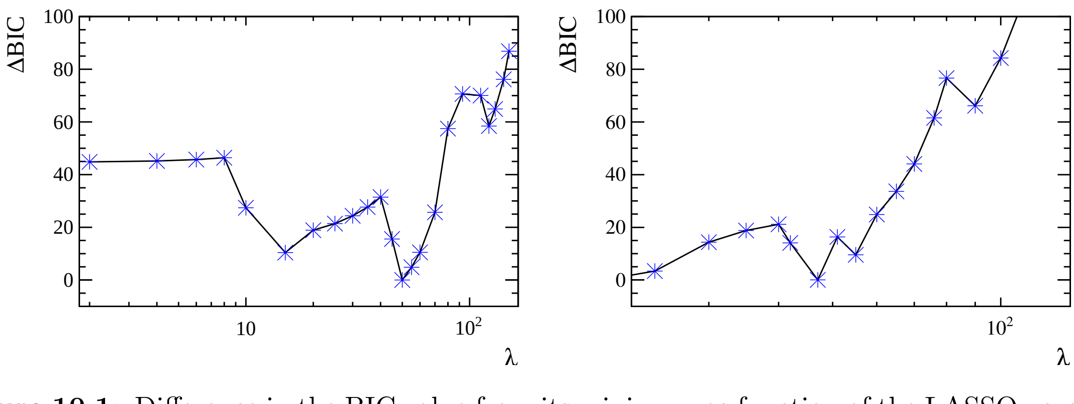

Figure 10.1: Difference in the BIC value from its minimum as function of the LASSO parameter λ for step 1 (left) and step 2 (right) of the model selection.

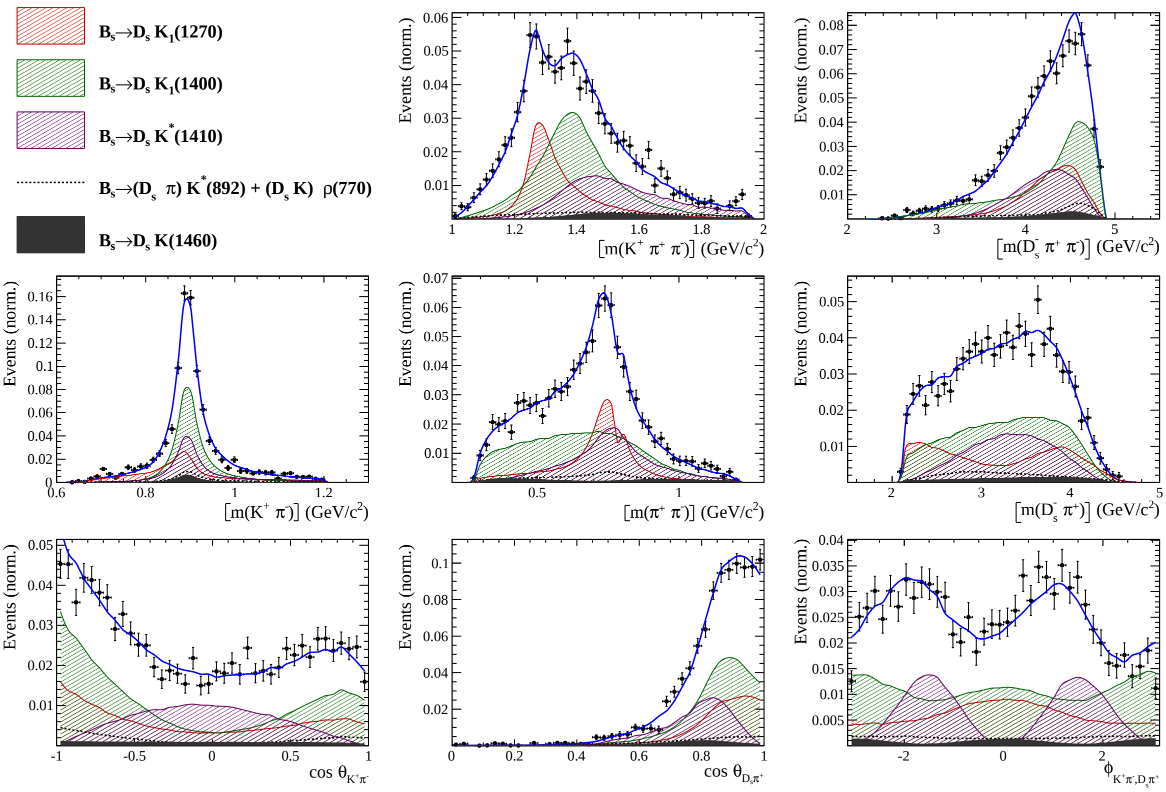

Figure 10.2: Projections of the fit result to the time-integrated and flavour averaged phase-space distribution of $B_s \to D_s K\pi\pi$ decays.

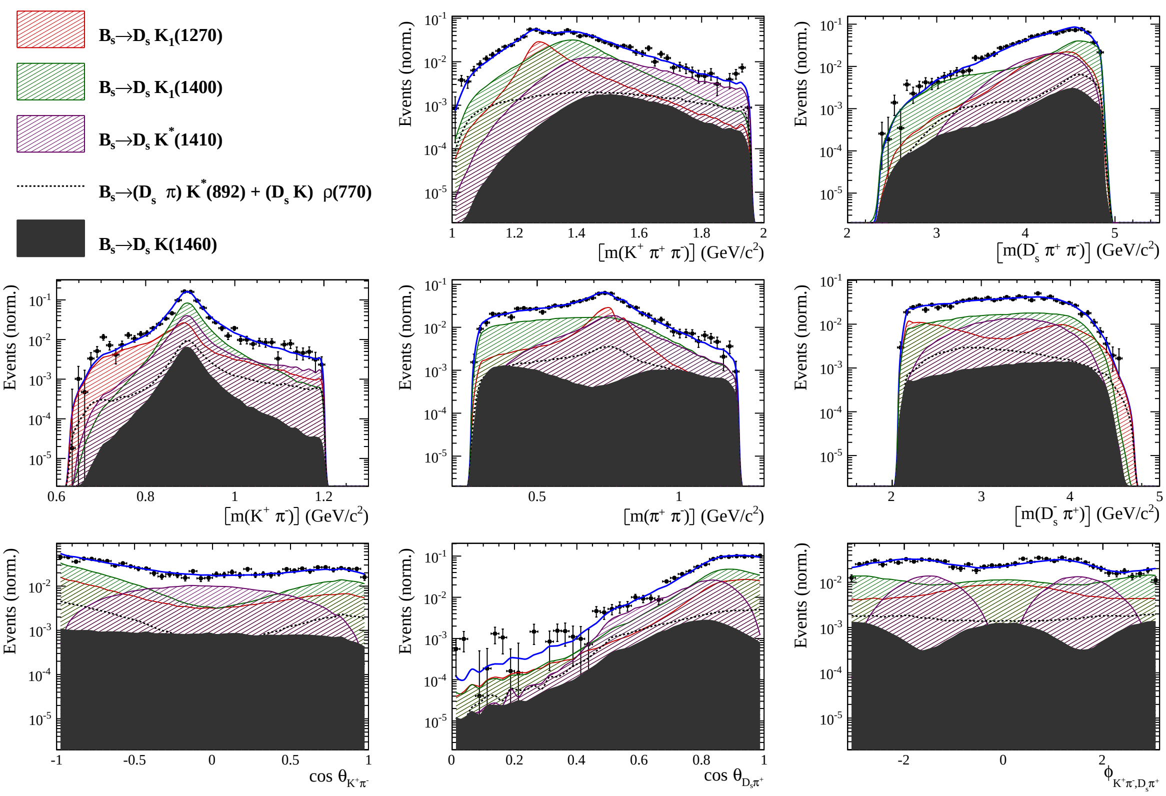

Figure 10.3: Projections of the fit result to the time-integrated and flavour averaged phase-space distribution of $B_s \to D_s K\pi\pi$ decays in logarithmic scale.

10.2 Results

Table 10.2 lists the modulus and phases of the complex amplitude coefficients a_i^c and a_i^u, obtained by fitting the LASSO model to the data. The corresponding fit fractions for the $b \to c$ and $b \to u$ amplitudes are individually normalized

$$F_i^{c,u} \equiv \frac{\int |a_i^{c,u} A_i(\mathbf{x})|^2 \, \mathrm{d}\Phi_4}{\int \left|\mathcal{A}_f^{c,u}(\mathbf{x})\right|^2 \, \mathrm{d}\Phi_4} \tag{10.8}$$

and shown in Table 10.3. In addition to the amplitude coefficients, the amplitude ratio and the strong and weak phase differences between the $b \to c$ and $b \to u$ decays are determined. Moreover, the masses and widths of the $K_1(1400)$ and $K^*(1410)$ resonances are fitted.

Figure 10.4 shows the distributions of selected phase space observables, which demonstrate reasonable agreement between data and the fit model. We also project into the transversity basis to demonstrate good description of the overall angular structure. The acoplanarity angle χ, is the angle between the two decay planes formed by the $K^+\pi^-$ system and the $D_s^-\pi^+$ system in the B_s rest frame; boosting into the rest frames of the two-body systems defining these decay planes, the two helicity variables are defined as the cosine of the angle, θ, of the K^+ or D_S^- momentum with the B_s flight direction.

In order to quantify the quality of the fit in the five-dimensional phase space, a χ^2 value is determined by binning the data;

$$\chi^2 = \sum_{b=1}^{N_{\text{bins}}} \frac{(N_b - N_b^{\text{exp}})^2}{N_b^{\text{exp}}}, \tag{10.9}$$

where N_b is the number of data events in a given bin, N_b^{exp} is the event count predicted by the fitted PDF and N_{bins} is the number of bins. An adaptive binning is used to ensure sufficient statistics in each bin for a robust χ^2 calculation [51]. At least 25 events per bin are required. The number of degrees of freedom ν, in an unbinned fit is bounded by $N_{\text{bins}} - 1$ and $(N_{\text{bins}} - 1) - N_{\text{par}}$, where N_{par} is the number of free fit parameters. We use the χ^2 value divided by $\nu = (N_{\text{bins}} - 1) - N_{\text{par}}$ as a conservative estimate. For the LASSO model, this amounts to $\chi^2/\nu = 1.40$ indicating a decent fit quality.

Table 10.2: Modulus and phases of the amplitudes contributing to $b \to c$ and $b \to u$ decays. In case of multiple decay modes of three-body resonances, the amplitude coefficients are defined relative to the one listed first. Additional fit parameters are listed below. The first quoted uncertainty is statistical, while the second arises from systematic sources. The third uncertainty arises from the alternative models considered.

Decay Channel	$A_{b\to c}$		$A_{b\to u}$	
	$\vert a_i\vert$	$arg(a_i)[^\circ]$	$\vert a_i\vert$	$arg(a_i)[^\circ]$
$B_s \to D_s\,(K_1(1270) \to K\,\rho(770))$	1.0	0.0	1.0	0.0
$K_1(1270) \to K^*(892)\,\pi$	0.76 ± 0.11 ± 0.16	60.9 ± 9.6 ± 14.1		
$K_1(1270) \to K_0^*(1430)\,\pi$	0.68 ± 0.06 ± 0.34	116.5 ± 5.1 ± 43.5		
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$	2.53 ± 0.27 ± 0.57	12.9 ± 7.4 ± 8.2	0.67 ± 0.20 ± 0.51	-76.3 ± 16.9 ± 22.9
$B_s \to D_s\,(K^*(1410) \to K^*(892)\,\pi)$	1.28 ± 0.12 ± 0.25	54.9 ± 5.6 ± 10.0		
$K^*(1410) \to K\,\rho(770)$	0.66 ± 0.04 ± 0.04	-172.9 ± 5.0 ± 6.7		
$B_s \to D_s\,(K(1460) \to K^*(892)\,\pi)$			0.77 ± 0.11 ± 0.62	-93.6 ± 11.2 ± 12.6
$B_s \to (D_s\,\pi)_P\ K^*(892)$	1.02 ± 0.13 ± 0.41	-28.4 ± 8.0 ± 10.5	0.79 ± 0.18 ± 0.36	3.7 ± 12.5 ± 15.1
$B_s \to (D_s\,K)_P\ \rho(770)$			0.61 ± 0.08 ± 0.26	36.4 ± 7.7 ± 14.5

Fit parameter	Value
$m_{K_1(1400)}$ [MeV]	1394.9 ± 8.8 ± 12.7 ± 21.2
$\Gamma_{K_1(1400)}$ [MeV]	224.0 ± 15.9 ± 22.3 ± 20.9
$m_{K^*(1410)}$ [MeV]	1419.6 ± 10.8 ± 26.9 ± 24.1
$\Gamma_{K^*(1410)}$ [MeV]	342.4 ± 23.5 ± 51.7 ± 52.9
r	xx.xx ± 0.04 ± 0.05 ± 0.04
$\delta\,[^\circ]$	xx.xx ± 16.1 ± 6.8 ± 6.8
$\gamma - 2\beta_s\,[^\circ]$	xx.xx ± 16.1 ± 11.6 ± 6.2

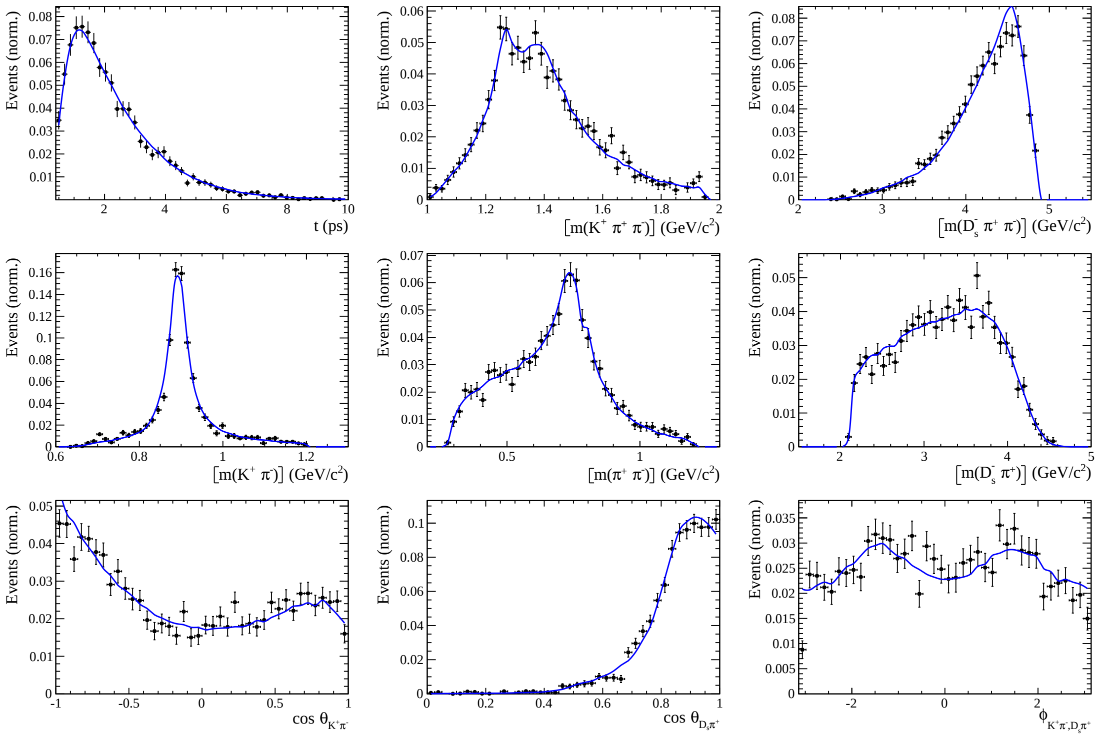

Figure 10.4: Projections of the full time-dependent amplitude fit.

Table 10.3: Fit fractions of the amplitudes contributing to $b \to c$ and $b \to u$ decays.

Decay Channel	$F_{b\to c}[\%]$	$F_{b\to u}[\%]$
$B_s \to D_s\,(K_1(1270) \to K^*(892)\,\pi)$	5.2 ± 1.2	17.6 ± 4.2
$B_s \to D_s\,(K_1(1270) \to K\,\rho(770))$	9.4 ± 1.1	32.0 ± 4.9
$B_s \to D_s\,(K_1(1270) \to K_0^*(1430)\,\pi)$	4.5 ± 0.6	15.2 ± 2.4
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$	59.9 ± 5.0	16.6 ± 8.5
$B_s \to D_s\,(K^*(1410) \to K^*(892)\,\pi)$	14.8 ± 0.9	
$B_s \to D_s\,(K^*(1410) \to K\,\rho(770))$	7.0 ± 0.6	
$B_s \to D_s\,(K(1460) \to K^*(892)\,\pi)$		18.8 ± 4.2
$B_s \to (D_s\,\pi)_P\;K^*(892)$	9.7 ± 1.8	21.8 ± 7.2
$B_s \to (D_s\,K)_P\;\rho(770)$		13.5 ± 4.0
Sum	110.5 ± 5.4	135.4 ± 10.3

11 Systematic uncertainties

The systematic uncertainties on the measured observables are summarized in Table 11.1 for the decay-time fit to $B_s \to D_s\pi\pi\pi$, in Table 11.2 for the decay-time fit to $B_s \to D_sK\pi\pi$ and in Table 11.3 for the full time-dependent amplitude fit to $B_s \to D_sK\pi\pi$ decays. A description of each systematic effect is given in the following subsections starting with the ones common to all fits. Afterwards, systematic effect specific to the amplitude description are discussed.

11.1 Fit bias

Pseudo-experiments are performed, where a signal toy sample of the same size as the number of observed signal data events is generated according to the nominal fit model and subsequently fitted with the same model. The means of the pull distributions are taken as systematic uncertainties of the fit parameters.

11.2 Background subtraction

The statistical subtraction of the residual background [47], left after the full selection, relies on the correct description of the invariant B_s^0 mass distribution. Since the choice of signal and background models is not unique, alternative parameterizations are tested:

- The Johnson's SU function which is used as nominal signal model is replaced by the sum of two Crystal Ball functions [52].
- For the combinatorial background, the nominal second order polynomial is replaced by an exponential function.
- For the description of the partially reconstructed background, a combination of the RooHILLdini and RooHORNsdini model [53] is used instead of the nominal model of three bifurcated gaussians.
- For the shape of the mis-ID background, the nominal approach is to use a simulated sample of $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ or $B_s^0 \to D_s^{*-}\pi^+\pi^-\pi^+$ decays and flip the mass hypothesis of the π^+ with the higher misidentification probability (see Sec. 4). Two alternative approaches are considered: we flip the mass hypothesis of the π^+ candidate with the lower probability of beeing misidentified; we randomly flip the mass hypothesis of a π^+ candidate.

To evaluate the possible source of systematic uncertainty arising from the fixed yields of the mis-ID backgrounds, the yields are fixed to zero or doubled.

In total 15 (7) different combinations of the modifications discussed above are tested for the fit to the $D_sK\pi\pi$ ($D_s\pi\pi\pi$) mass distribution. For each case, new signal sWeights are calculated and the sFits to data are repeated. The sample variance of the obtained differences to the nominal fit value are used as systematic uncertainty due to the background subtraction.

11.3 Decay-time acceptance

The systematic uncertainty related to the decay-time efficiency as well as Γ_s and $\Delta\Gamma_s$ are studied simultaneously. We generate toys in the nominal configuration and fit back in both this nominal configuration and a configuration in which we have randomized the acceptance parameters together with Γ_s and $\Delta\Gamma_s$ within their uncertainties. For each toy, a pull is calculated by dividing the difference between the fitted values of the nominal and shifted configurations by the uncertainty in the nominal toy. We add the bias in the mean of this pull to its width, in quadrature, in order to arrive at the final systematic uncertainty.

To improve the coverage of the multi-dimensional parameter space, a Cholesky decomposition [54] is used to generate a set of uncorrelated vectors from the covariance matrix $\mathrm{cov}(\lambda_i, \lambda_j)$, where the vector λ includes the parameters Γ_s, $\Delta\Gamma_s$ and the $N = 4$ spline coefficients for each category of the simultaneous fit. The correlations between Γ_s ($\Delta\Gamma_s$) and the spline coefficients are measured by rerunning the acceptance fits described in Sec. 6.2 with the values of Γ_s ($\Delta\Gamma_s$) varied by $\pm 1\sigma$ and measuring the shift in the spline coefficients as a fraction of their uncertainty. For the correlation between Γ_s and $\Delta\Gamma_s$ we use the HFAG value [38].

11.4 Decay-time resolution and tagging

To study systematic effects originating from the scaling of the decay-time error estimate, two different approaches which either slightly overestimate or underestimate the resolution are used:

- A double Gaussian is fit to the decay-time distributions of fake B_s^0 candidates, but only the width of the core Gaussian is considered to represent the time resolution in the respective bin. Therefore the resolution is slightly underestimated in this case.
- A single Gaussian is fit to the decay-time distributions of fake B_s^0 candidates in a wide range of $[-3\sigma_t : 1.5\sigma_t]$. Due to the tails of the distribution, which broaden the width of the Gaussian function, this method slightly overestimates the decay-time resolution.

For each case, a new scaling function is derived:

$$\sigma_{eff}^{core-Gauss}(\sigma_t) = (4.9 \pm 2.0)\,\mathrm{fs} + (0.821 \pm 0.050)\,\sigma_t \tag{11.1}$$

$$\sigma_{eff}^{single-Gauss}(\sigma_t) = (8.3 \pm 1.5)\,\mathrm{fs} + (0.997 \pm 0.037)\,\sigma_t \tag{11.2}$$

which are compared to the nominal result in Fig. 11.1.

Due to the high correlation between the decay-time resolution and the tagging calibration, their systematic uncertainty has to be studied simultaneously. First, the decay-time fits to $B_s \to D_s\pi\pi\pi$ data are repeated using the alternative decay-time error scaling functions. New tagging calibration parameters are obtained which are then used (together with the respective decay-time error scaling function) in the fits to $B_s \to D_s K\pi\pi$ data to define the Gaussian-constrains as discussed in Sec. 9. For the width of the Gaussians only the statistical error of the tagging calibration parameters are used since systematic uncertainties (except the systematic arising from the decay-time resolution which is already included by the procedure described above) are found to be negligible, see Table 11.1. Finally, we take the biggest change in fit central value as the systematic for each parameter of the $B_s \to D_s K\pi\pi$ fits.

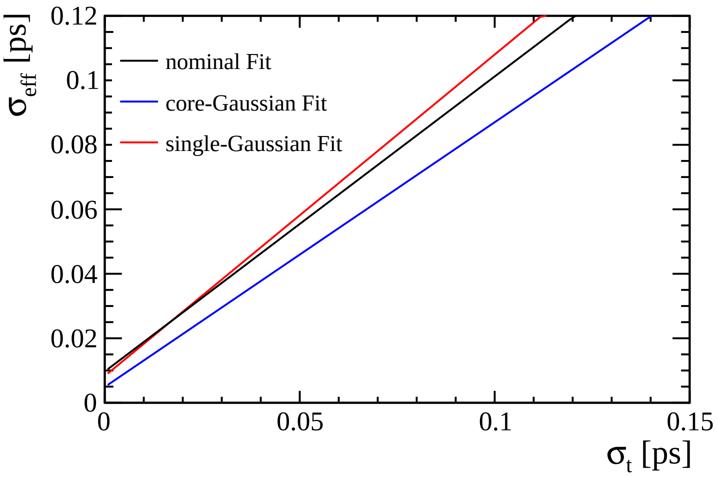

Figure 11.1: The measured resolution scaling function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data) for (black line) the nominal scaling, (blue line) only using the narrow gaussian width of the double gaussian fit model or (red line) when determining the resolution using a single gaussian model.

11.5 Production, detection asymmetries and mixing frequency

The systematic from the production, detection asymmetries and Δm_s (in case of $B_s \to D_s K\pi\pi$ decays) which are fixed in the fit are evaluated by means of a toy study similar to the procedure performed for the time-acceptance. The parameters are assumed to be uncorrelated.

11.6 Multiple candidates

The fraction of events with multiple candidates has been found to be very small, it is 1.6% for $D_s K\pi\pi$ and 1.5% for $D_s\pi\pi\pi$. Thus the nominal result is obtained keeping all candidates, while a systematic uncertainty is assigned by repeating the fit randomly keeping only one candidate when multiple ones are founds. No shifts in the fit central values are observed.

11.7 Length and momentum scales

The uncertainty on the LHCb length scale is estimated to be at most 0.020% [55], which translates directly in an uncertainty on Δm_s of 0.020% with other parameters being unaffected. The momentum scale uncertainty is at most 0.022%.

11.8 Phase space acceptance

For the phase space acceptance we rely on simulated data. The integration error due to the limited size of the MC sample used to normalize the signal PDF is evaluated by bootstrapping the MC sample and repeating the full time-dependent amplitude fit.

To asses the uncertainty due to possible data-simulation differences, we determine alternative phase space efficiencies by varying the selection requirements on quantities that are expected not to be well described by the simulation. In particular, we consider the following variations:

- No BDT cut is applied
- A tighter BDT requirement is used (BDTG > 0.6)
- No reweighting is applied
- Instead of the PID responses obtained from the PIDCorr tool, we use the PIDGen tool to resample the PID variables [37]
- The raw MC PID variables are used
- Candidates with BKGCAT$= 60$ are removed

We assign the sample variance of the fitted values using the alternative phase space acceptances as systematic.

This will be done when the final MC samples are available. We expect the integration error to be negligible and the systematic error from data-simulation differences to be small. At the moment we estimate a systematic by assuming a flat phase space acceptance. The resulting uncertainties shown in Table 11.3 should be considered as upper limit and illustrate that we are not highly sensitive to the details of the phase space acceptance shape.

11.9 Resonance description

The following alternative line shape parameterizations are considered as part of the systematic studies:

- The Lass description for the $K\pi$ S-wave is replaced by a relativistic Breit-Wigner propagator (Equation 2.16)
- The Gounaris-Sakurai description for the $\rho(770)$ is replaced by a relativistic Breit-Wigner propagator (Equation 2.16)
- The ω contribution to the decay channel $K_1(1270) \to K\,\rho(770)/\omega$ is set to zero
- For the decay channel $K^*(1410) \to K\,\rho(770)$, we include $\rho(770) - \omega$ mixing with a relative magnitude and phase determined from data
- Instead of taking the energy-dependent widths of the three-body resonances from Refs. [9, 21], we derive them from Equation 2.17 assuming an uniform phase space population.

The data fits are repeated for each alternative model and the shifts of the central values are taken as systematic uncertainties. They are added in quadrature.

The uncertainties due to fixed masses and widths of resonances are evaluated from toys where we vary them one-by-one within their quoted errors. In our nominal fit, the Blatt-Weisskopf radial parameter is set to $r_{BW} = 1.5\,\mathrm{GeV}^{-1}$. Again, toys are generated according to this nominal configuration and then fitted whereby the radial parameter is uniformly varied within the interval $[0, 3]\,\mathrm{GeV}^{-1}$.

11.10 Alternative amplitude models

We tested several modifications of the LASSO model to assign an additional model uncertainty to the measured observables r, δ and $\gamma - 2\beta_s$ as well as to the measured masses and widths of the $K_1(1400)$ and $K^*(1410)$ resonances. The amplitude coefficients are by definition parameters of a given model which is why we do not evaluate a model uncertainty for them.

- All amplitudes selected by Stage 1 of the model selection are included for both $b \to c$ and $b \to u$ transitions
- The decay channel $K_1(1270)[D] \to K^*(892)\,\pi$ is added where the $K^*(892)\,\pi$ system is in relative a D-wave state
- The decay channel $K_1(1400) \to K\,\rho(770)$ is added
- The decay channels $K_2^*(1430) \to K\,\rho(770)$ and $K_2^*(1430) \to K^*(892)\,\pi$ are added
- The decay channels $K(1460) \to K\,\rho(770)$ and $K(1460) \to K\,\sigma$ are added
- The $K(1460)$ resonance is removed
- The decay channels $K^*(1680) \to K\,\rho(770)$ and $K^*(1680) \to K^*(892)\,\pi$ are added
- The decay channels $K_2(1770) \to K\,\rho(770)$ and $K_2(1770) \to K^*(892)\,\pi$ are added
- The amplitudes $B_s \to (D_s\,\pi)_P\,K^*(892)$ and $B_s \to (D_s\,K)_P\,\rho(770)$ are replaced by $B_s \to (D_s\,\pi)_S\,K^*(892)$ and $B_s \to (D_s\,K)_S\,\rho(770)$
- Higher orbital angular momentum states are added for the amplitudes: $B_s[S,P,D] \to (D_s\,\pi)_P\,K^*(892)$ and $B_s[S,P,D] \to (D_s\,K)_P\,\rho(770)$
- The amplitudes $B_s \to (D_s\,\pi)_P\,K^*(892)$ and $B_s \to (D_s\,K)_P\,\rho(770)$ are removed
- The amplitudes $B_s \to (D_s\,K)_P\,\sigma$, $B_s \to (D_s\,K)_P\,f_2(1270)$ and $B_s \to (D_s\,K)_P\,f_0(1370)$ are added
- The amplitudes $B_s \to (D_s\,\pi)_P\,K_0^*(1430)$ and $B_s \to (D_s\,K)_S\,K_2^*(1430)$ are added

In total 20 different sets of amplitudes are fitted. In some cases, the fit fractions of additionally added amplitudes turn out to be exactly zero. These model are effectively not distinguisbale from the baseline LASSO model and are not considered further. From the remaining 15 models, we compute the sample variance for each observable and take it as model uncertainty.

Table 11.1: Systematic uncertainties on the fit parameters of the fit to $B_s \to D_s\pi\pi\pi$ data in units of statistical standard deviations.

Fit Parameter	Fit-bias	Acceptance	Resolution	Asymmetries	Background	Mult.-Cand.	Mom./z-Scale	Total
p_0^{OS} Run-I	0.04	0.00	0.99	0.01	0.04	0.00		0.99
p_1^{OS} Run-I	0.01	0.00	1.03	0.00	0.05	0.00		1.03
Δp_0^{OS} Run-I	0.03	0.00	0.02	0.15	0.02	0.00		0.16
Δp_1^{OS} Run-I	0.02	0.00	0.03	0.16	0.02	0.00		0.16
$\epsilon_{\text{tag}}^{OS}$ Run-I	0.02	0.00	0.00	0.01	0.09	0.00		0.09
$\Delta\epsilon_{\text{tag}}^{OS}$ Run-I	0.03	0.00	0.07	0.01	0.02	0.00		0.07
p_0^{SS} Run-I	0.01	0.00	0.55	0.00	0.03	0.00		0.55
p_1^{SS} Run-I	0.04	0.00	0.60	0.01	0.03	0.00		0.60
Δp_0^{SS} Run-I	0.00	0.00	0.00	0.10	0.01	0.00		0.10
Δp_1^{SS} Run-I	0.07	0.00	0.01	0.12	0.03	0.00		0.15
$\epsilon_{\text{tag}}^{SS}$ Run-I	0.02	0.00	0.00	0.01	0.01	0.00		0.03
$\Delta\epsilon_{\text{tag}}^{SS}$ Run-I	0.04	0.00	0.05	0.01	0.02	0.00		0.07
p_0^{OS} Run-II	0.01	0.01	1.65	0.00	0.10	0.00		1.65
p_1^{OS} Run-II	0.01	0.00	1.37	0.00	0.10	0.00		1.38
Δp_0^{OS} Run-II	0.05	0.00	0.06	0.00	0.03	0.00		0.08
Δp_1^{OS} Run-II	0.02	0.00	0.03	0.00	0.04	0.00		0.05
$\epsilon_{\text{tag}}^{OS}$ Run-II	0.02	0.00	0.00	0.00	0.04	0.00		0.05
$\Delta\epsilon_{\text{tag}}^{OS}$ Run-II	0.01	0.00	0.21	0.00	0.04	0.00		0.22
p_0^{SS} Run-II	0.00	0.00	1.06	0.00	0.03	0.00		1.06
p_1^{SS} Run-II	0.07	0.00	1.22	0.00	0.03	0.00		1.22
Δp_0^{SS} Run-II	0.00	0.00	0.02	0.00	0.03	0.00		0.04
Δp_1^{SS} Run-II	0.07	0.00	0.03	0.00	0.03	0.00		0.08
$\epsilon_{\text{tag}}^{SS}$ Run-II	0.00	0.00	0.00	0.00	0.04	0.00		0.04
$\Delta\epsilon_{\text{tag}}^{SS}$ Run-II	0.02	0.00	0.05	0.00	0.02	0.00		0.06
A_P Run-II	0.10	0.00	0.10	0.01	0.03	0.00		0.14
Δm_s	0.01	0.00	0.15	0.03	0.06	0.00	0.61	0.63

Table 11.2: Systematic uncertainties on the fit parameters of the phase-space integrated fit to $B_s \to D_s K\pi\pi$ data in units of statistical standard deviations.

Fit Parameter	Fit bias	Acceptance	Resolution	Δm_s	Asymmetries	Background	Total
C	0.02	0.04	0.07	0.06	0.03	0.09	0.14
D	0.04	0.26	0.00	0.02	0.05	0.11	0.29
$\bar{D}$	0.05	0.26	0.01	0.02	0.05	0.16	0.32
S	0.01	0.02	0.03	0.24	0.03	0.15	0.29
$\bar{S}$	0.04	0.03	0.06	0.23	0.03	0.13	0.27

Table 11.3: Systematic uncertainties on the fit parameters of the full time-dependent amplitude fit to $B_s \to D_s K \pi\pi$ data in units of statistical standard deviations.

Fit Parameter	Fit bias	Time-Acc.	Resolution	Δm_s	Asymmetries	Background	Lineshapes	Resonances m, Γ	Form-Factors	Phsp-Acc.	Amp. Model	Total
$B_s \to D_s\,(K_1(1270) \to K^*(892)\,\pi)$ Mag	0.04	0.17	0.01	0.01	0.02	0.15	1.30	0.28	0.42	0.06		1.42
$B_s \to D_s\,(K_1(1270) \to K^*(892)\,\pi)$ Phase	0.08	0.20	0.03	0.01	0.16	0.06	0.85	0.31	0.20	1.10		1.47
$B_s \to D_s\,(K_1(1270) \to K_0^*(1430)\,\pi)$ Mag	0.07	0.17	0.02	0.01	0.11	0.25	3.96	3.69	0.45	2.20		5.87
$B_s \to D_s\,(K_1(1270) \to K_0^*(1430)\,\pi)$ Phase	0.24	0.16	0.02	0.01	0.18	0.15	7.28	0.21	0.51	4.47		8.57
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Mag$(b \to c)$	0.08	0.13	0.02	0.03	0.43	0.27	1.38	0.28	0.38	1.44		2.12
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Phase$(b \to c)$	0.07	0.24	0.01	0.03	0.13	0.28	0.66	0.25	0.32	0.69		1.11
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Mag$(b \to u)$	0.21	0.19	0.02	0.04	0.06	0.19	0.83	0.24	0.56	2.27		2.52
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Phase$(b \to u)$	0.01	0.16	0.04	0.10	0.15	0.36	0.79	0.43	0.25	0.88		1.36
$B_s \to D_s\,(K^*(1410) \to K^*(892)\,\pi)$ Mag$(b \to c)$	0.32	0.13	0.03	0.05	0.19	0.18	1.08	0.28	1.60	0.09		2.00
$B_s \to D_s\,(K^*(1410) \to K^*(892)\,\pi)$ Phase$(b \to c)$	0.25	0.23	0.01	0.01	0.21	0.10	1.42	0.22	0.75	0.62		1.79
$B_s \to D_s\,(K^*(1410) \to K\,\rho(770))$ Mag	0.49	0.20	0.01	0.01	0.12	0.17	0.60	0.18	0.19	0.15		0.88
$B_s \to D_s\,(K^*(1410) \to K\,\rho(770))$ Phase	0.23	0.22	0.01	0.01	0.10	0.13	0.34	0.12	0.29	1.22		1.35
$B_s \to D_s\,(K(1460) \to K^*(892)\,\pi)$ Mag$(b \to u)$	0.03	0.24	0.02	0.04	0.23	0.22	0.68	0.76	5.39	1.96		5.84
$B_s \to D_s\,(K(1460) \to K^*(892)\,\pi)$ Phase$(b \to u)$	0.02	0.30	0.03	0.04	0.13	0.21	0.64	0.40	0.48	0.55		1.12
$B_s \to (D_s\,\pi)_P\;K^*(892)$ Mag$(b \to c)$	0.15	0.16	0.02	0.02	0.35	0.24	1.28	0.20	2.66	0.99		3.16
$B_s \to (D_s\,\pi)_P\;K^*(892)$ Phase$(b \to c)$	0.01	0.20	0.01	0.01	0.20	0.47	0.95	0.18	0.34	0.59		1.30
$B_s \to (D_s\,\pi)_P\;K^*(892)$ Mag$(b \to u)$	0.15	0.14	0.04	0.03	0.37	0.13	0.47	0.27	1.73	0.68		1.99
$B_s \to (D_s\,\pi)_P\;K^*(892)$ Phase$(b \to u)$	0.01	0.26	0.05	0.03	0.88	0.28	0.56	0.21	0.42	0.10		1.21
$B_s \to (D_s\,K)_P\;\rho(770)$ Mag$(b \to u)$	0.45	0.24	0.01	0.05	0.83	0.49	1.34	0.38	2.81	0.33		3.34
$B_s \to (D_s\,K)_P\;\rho(770)$ Phase$(b \to u)$	0.31	0.31	0.02	0.03	0.24	0.66	0.25	0.60	0.71	1.37		1.87
$m_{K_1(1400)}$	0.04	0.18	0.02	0.01	0.36	0.17	1.15	0.16	0.33	0.66	2.41	2.81
$\Gamma_{K_1(1400)}$	0.05	0.22	0.02	0.01	0.29	0.13	1.23	0.12	0.25	0.46	1.31	1.92
$m_{K^*(1410)}$	0.08	0.19	0.01	0.01	0.51	0.11	1.69	0.27	1.63	0.51	2.22	3.33
$\Gamma_{K^*(1410)}$	0.30	0.17	0.01	0.01	0.10	0.18	1.17	0.59	1.71	0.15	2.25	3.14
r	0.07	0.19	0.05	0.10	0.38	0.29	1.02	0.20	0.18	0.58	0.98	1.64
δ	0.02	0.17	0.04	0.06	0.03	0.10	0.24	0.07	0.14	0.23	0.42	0.60
$\gamma - 2\beta_s$	0.01	0.11	0.05	0.07	0.28	0.25	0.30	0.29	0.06	0.42	0.39	0.82

A Stripping and Trigger cuts

The following text describes variables which are used in Table 1.1 and might be ambiguous, or which benefits are not straight forward. Where noted, different cut values are applied for Run-I and Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest approach. This variable is used to ensure that all D_s and $X_{s,d}$ daughters originate from the same vertex. DIRA is the abbreviation for the cosine of the angle θ between the hadron's flight direction $\vec{x}$ and it's corresponding momentum vector $\vec{p}$, $\cos\theta_{\vec{x}-\vec{p}}$.

Table 1.1: Summary of the stripping selections for $B_s^0 \to D_s K\pi\pi$ decays. The requirements for Run-II are only given if they have been changed.

Variable	Stripping Cut (Run-I)	Stripping Cut (Run-II)
Track χ^2/nDoF	< 3	
Track p	> 1000 MeV/c	
Track p_{T}	> 100 MeV/c	
Track IP χ^2	> 4	
Track ghost-prob.	< 0.4	
D_s mass	$m_{D_s} \pm 100\,\mathrm{MeV}$	$m_{D_s} \pm 80\,\mathrm{MeV}$
D_s Daughter p_{T}	$\sum_{i=1}^{3} p_{t,i} > 1800\,\mathrm{MeV}/c$	
D_s Daughter DOCA	< 0.5 mm	
D_s Vertex χ^2/nDoF	< 10	
D_s χ^2-separation from PV	> 36	
D_s daughter PID(π)	< 20	
D_s daughter PID(K)	> -10	
$X_{s,d}$ mass	$< 4000\,\mathrm{MeV}$	$< 3500\,\mathrm{MeV}$
$X_{s,d}$ Daughter p	> 2 GeV/c	
$X_{s,d}$ Daughter DOCA	< 0.4 mm	
$X_{s,d}$ Daughter p_{T}	$\sum_{i=1}^{3} p_{t,i} > 1250\,\mathrm{MeV}/c$	
$X_{s,d}$ Vertex χ^2/nDoF	< 8	
$X_{s,d}$ χ^2-separation from PV	> 16	
$X_{s,d}$ DIRA	> 0.98	
$X_{s,d}$ $\Delta\rho$	> 0.1 mm	
$X_{s,d}$ Δz	> 2.0 mm	
$X_{s,d}$ daughter PID(π)	< 10	
X_s daughter PID(K)	> -2	> 4
B_s^0 mass	$[4750, 7000]\,\mathrm{MeV}/c^2$	$[5000, 6000]\,\mathrm{MeV}/c^2$
B_s^0 DIRA	> 0.98	> 0.99994
B_s^0 min IP χ^2	< 25	< 20
B_s^0 Vertex χ^2/nDoF	< 10	< 8
B_s^0 $\tau_{B_s^0}$	> 0.2 ps	

Table 1.2 summarizes the trigger requirements imposed by the Hlt1 line used in this analysis for Run-I. At least one of the six decay particles must pass the listed requirements in order for the event to be stored for further analysis. For Run-II, this trigger line was updated and uses a multivariate classifier which takes the variables listed in Table 1.2 as input, rather than directly cutting on them.

The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the b-hadron p_{T}, its flight distance χ^2 from the nearest PV and the sum of the B_s^0 and D_s vertex χ^2 divided by the sum of their number of degrees of freedom. Table 1.3 summarizes the cuts applied by the inclusive ϕ trigger, which requires that a $\phi \to KK$ candidate can be formed out of two tracks present in the event.

Table 1.2: Summary of the cuts applied by the Hlt1TrackAllL0 trigger for Run-I. At least one of the six decay particles must pass this requirements, in order for the event to be accepted.

Quantity	Hlt1TrackAllL0 requirement
Track IP [mm]	> 0.1
Track IP χ^2	> 16
Track χ^2/nDoF	< 2.5
Track p_{T}	> 1.7 GeV/c
Track p	> 10 GeV/c
Number VELO hits/track	> 9
Number missed VELO hits/track	< 3
Number OT+IT $\times$ 2 hits/track	> 16

Table 1.3: Summary of the cuts applied by the Hlt2 inclusive ϕ trigger. A $\phi \to KK$ candidate, formed by two tracks in the event, must pass this requirements in order for the event to be accepted.

Quantity	Hlt2IncPhi requirement
ϕ mass	$m_\phi \pm 12$ MeV/c^2 of PDG value
ϕ p_{T}	> 2.5 GeV/c
ϕ vertex χ^2/nDoF	< 20
ϕ IP χ^2 to any PV	> 5

B Details of multivariate classifier

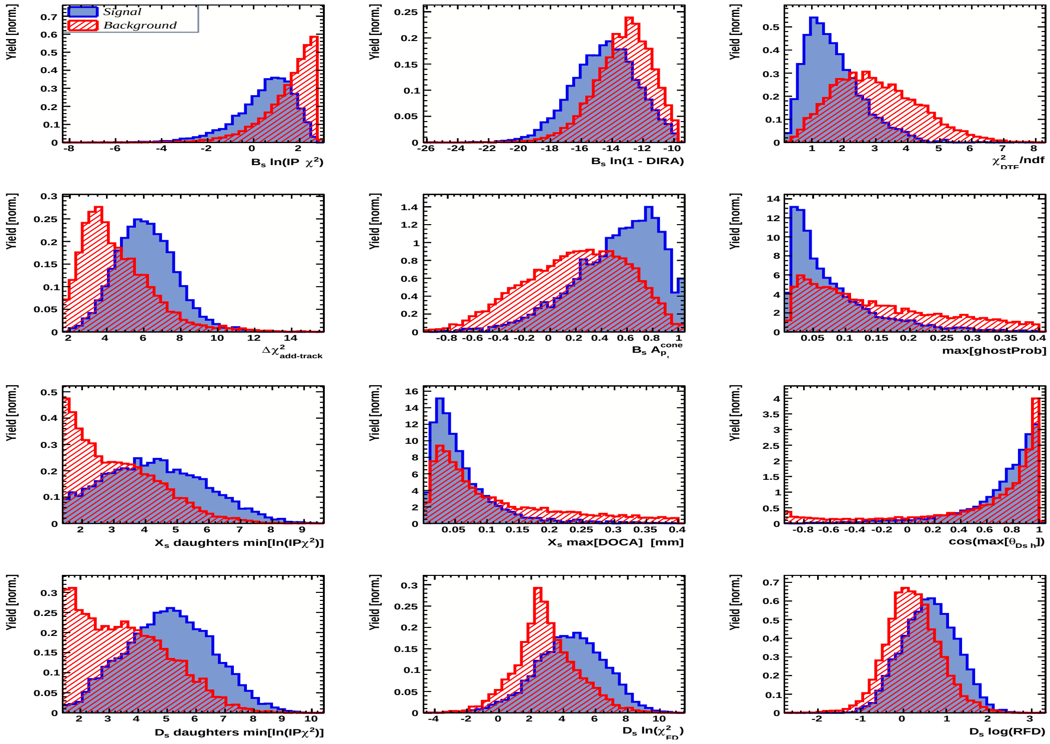

Figure A.1: Variables used to train the BDTG for category [Run-I,L0-TOS].

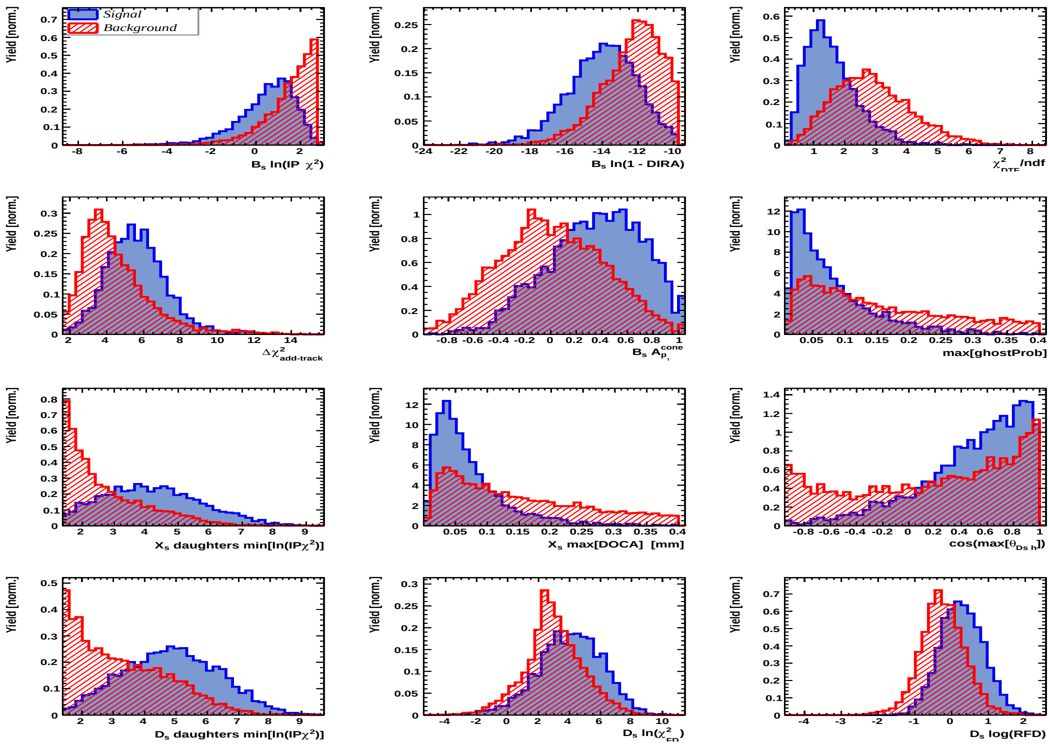

Figure A.2: Variables used to train the BDTG for category [Run-I,L0-TIS].

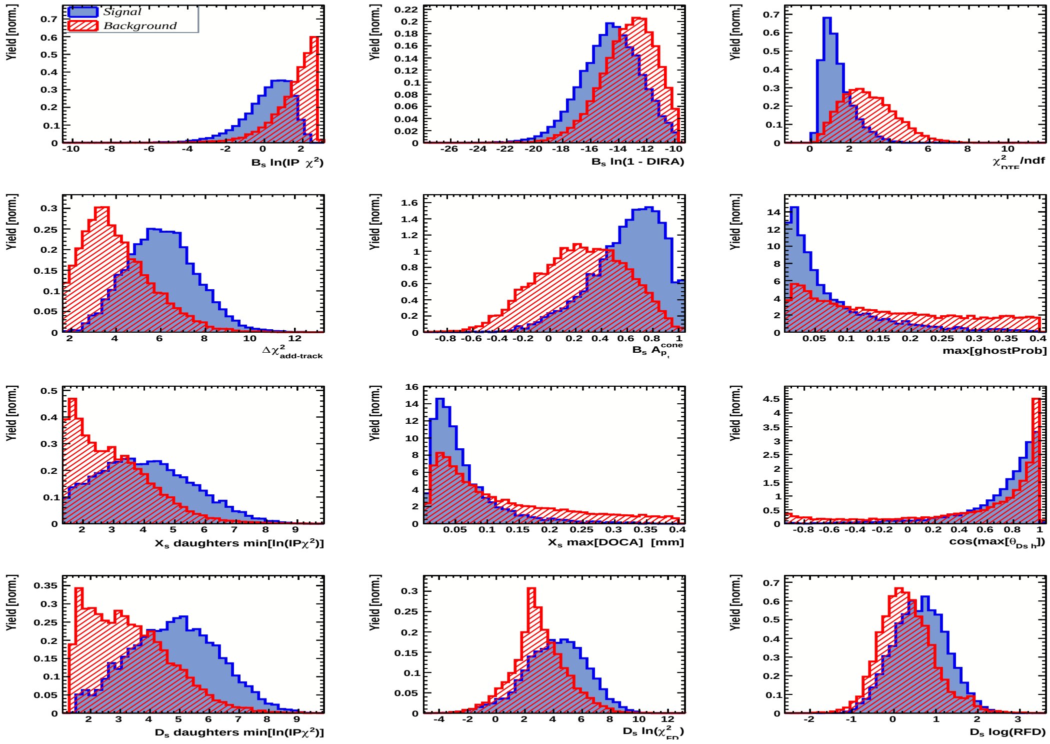

Figure A.3: Variables used to train the BDTG for category [Run-II,L0-TOS].

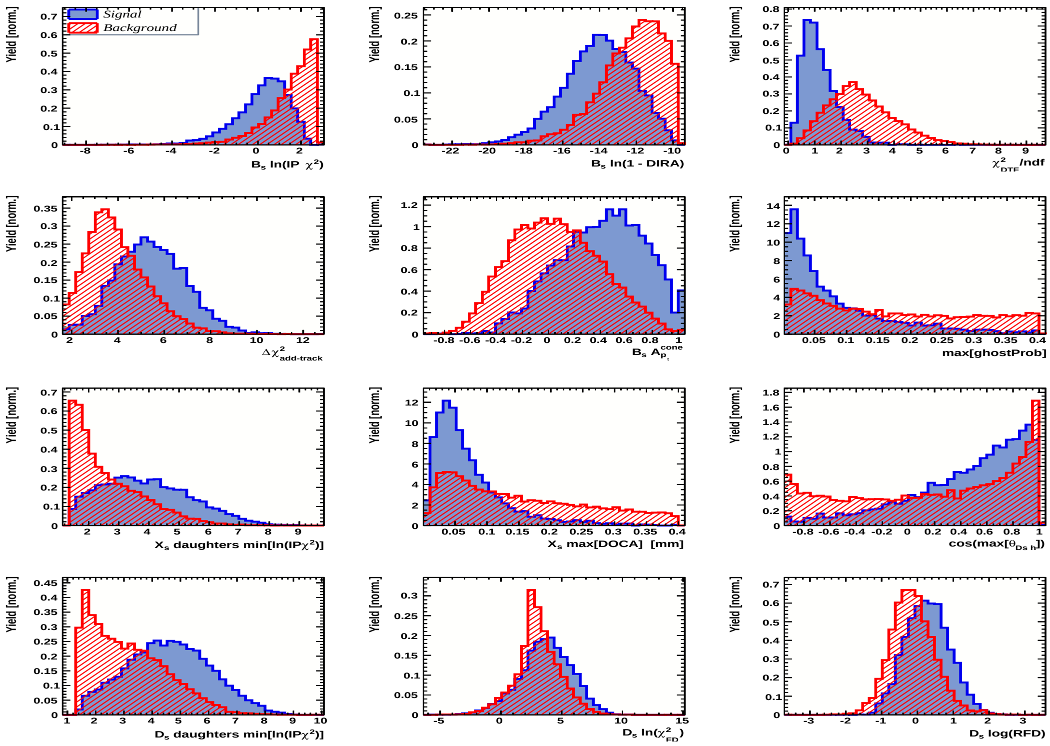

Figure A.4: Variables used to train the BDTG for category [Run-II,L0-TIS].

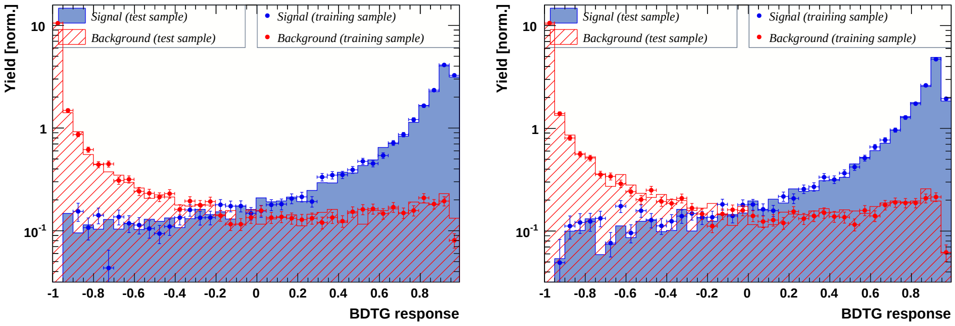

Figure A.5: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].

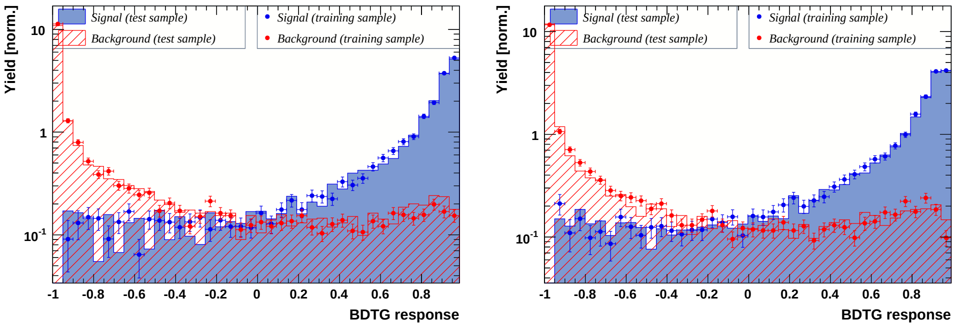

Figure A.6: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].

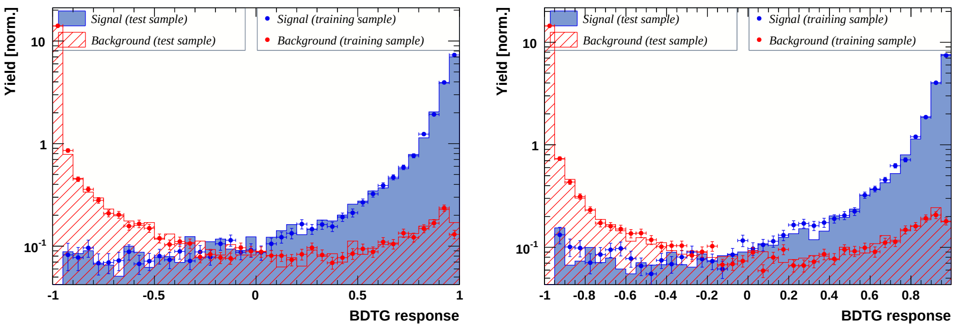

Figure A.7: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].

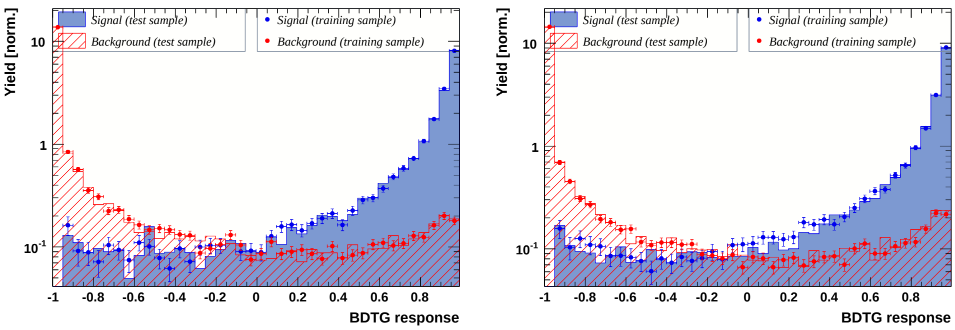

Figure A.8: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].

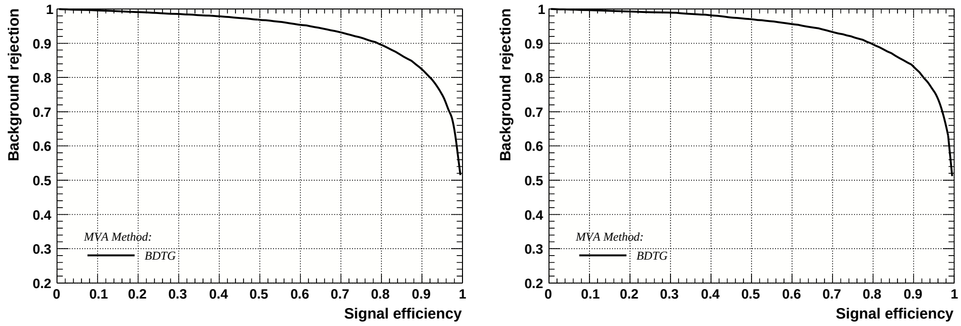

Figure A.9: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].

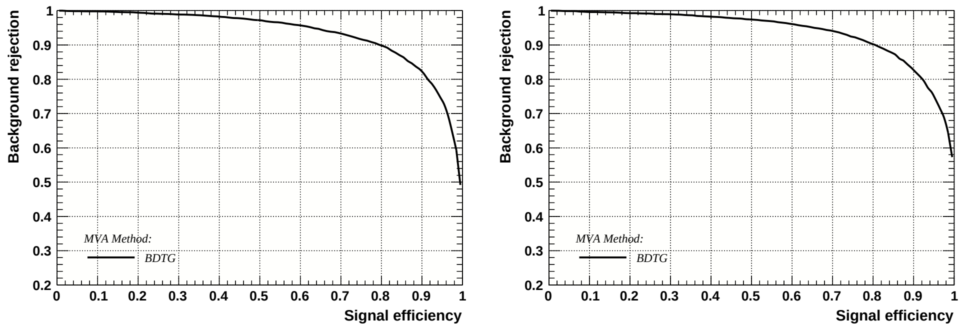

Figure A.10: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TIS].

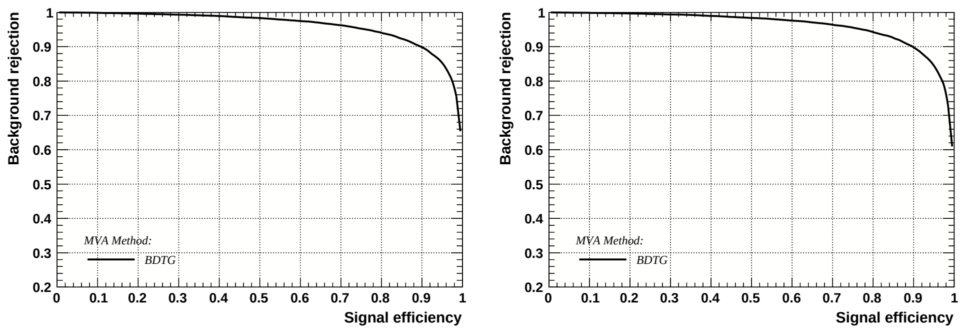

Figure A.11: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TOS].

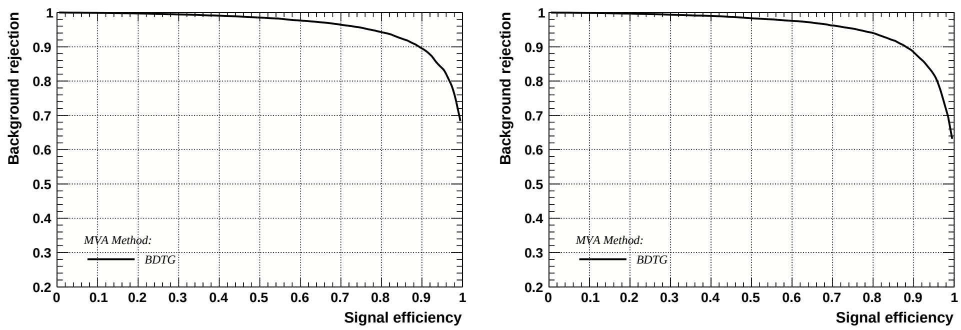

Figure A.12: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category[Run-II,L0-TIS].

C Detailed mass fits

In this section, all fits to the mass distribution of $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_sK\pi\pi$ candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-II), D_s final state ($D_s \to KK\pi$ non-resonant, $D_s \to \phi\pi$, $D_s \to K^*K$, or $D_s \to \pi\pi\pi$) and L0 trigger category.

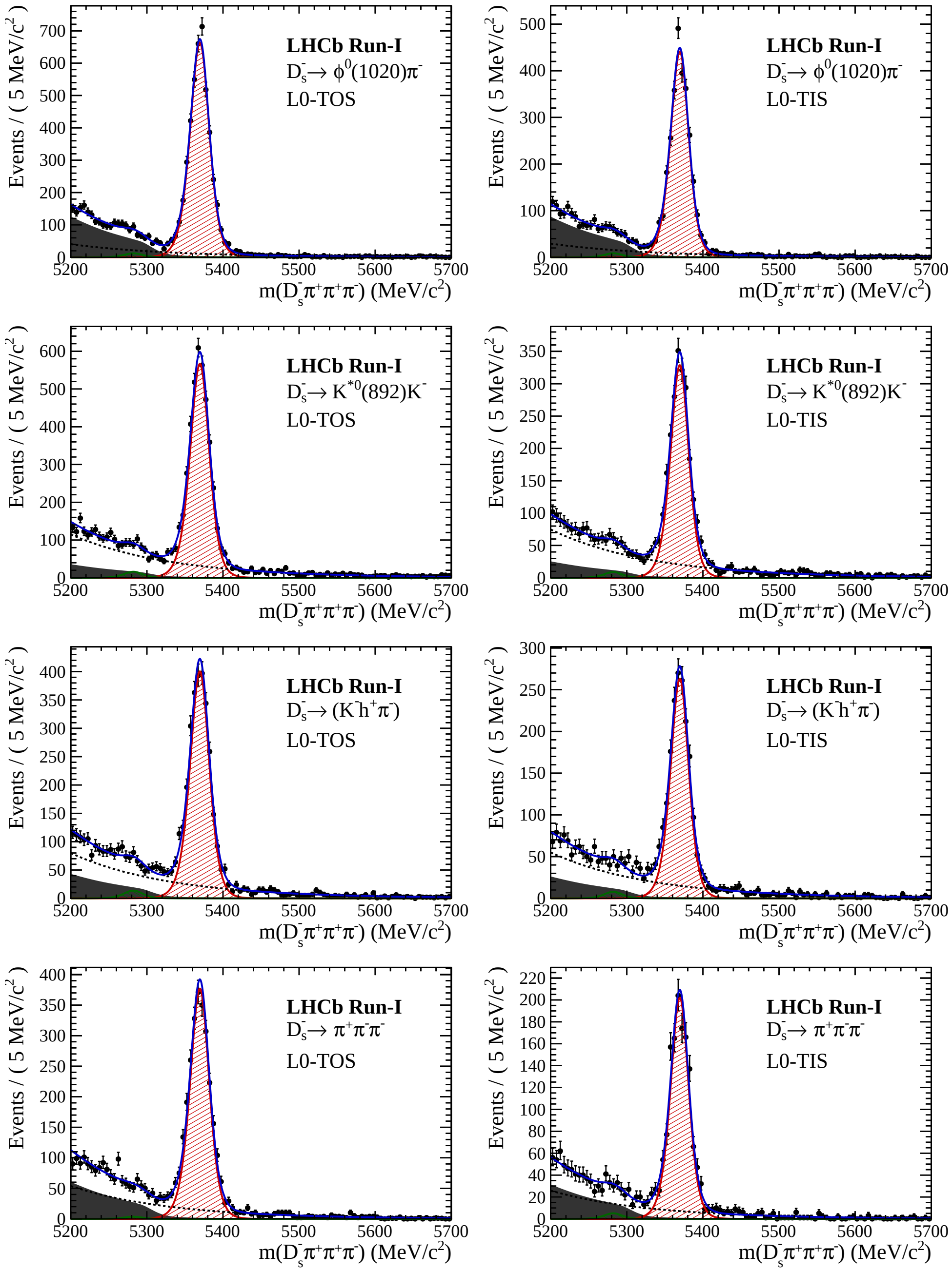

Figure B.1: Invariant mass distributions of $B_s^0 \to D_s\pi\pi\pi$ candidates for Run-I data.

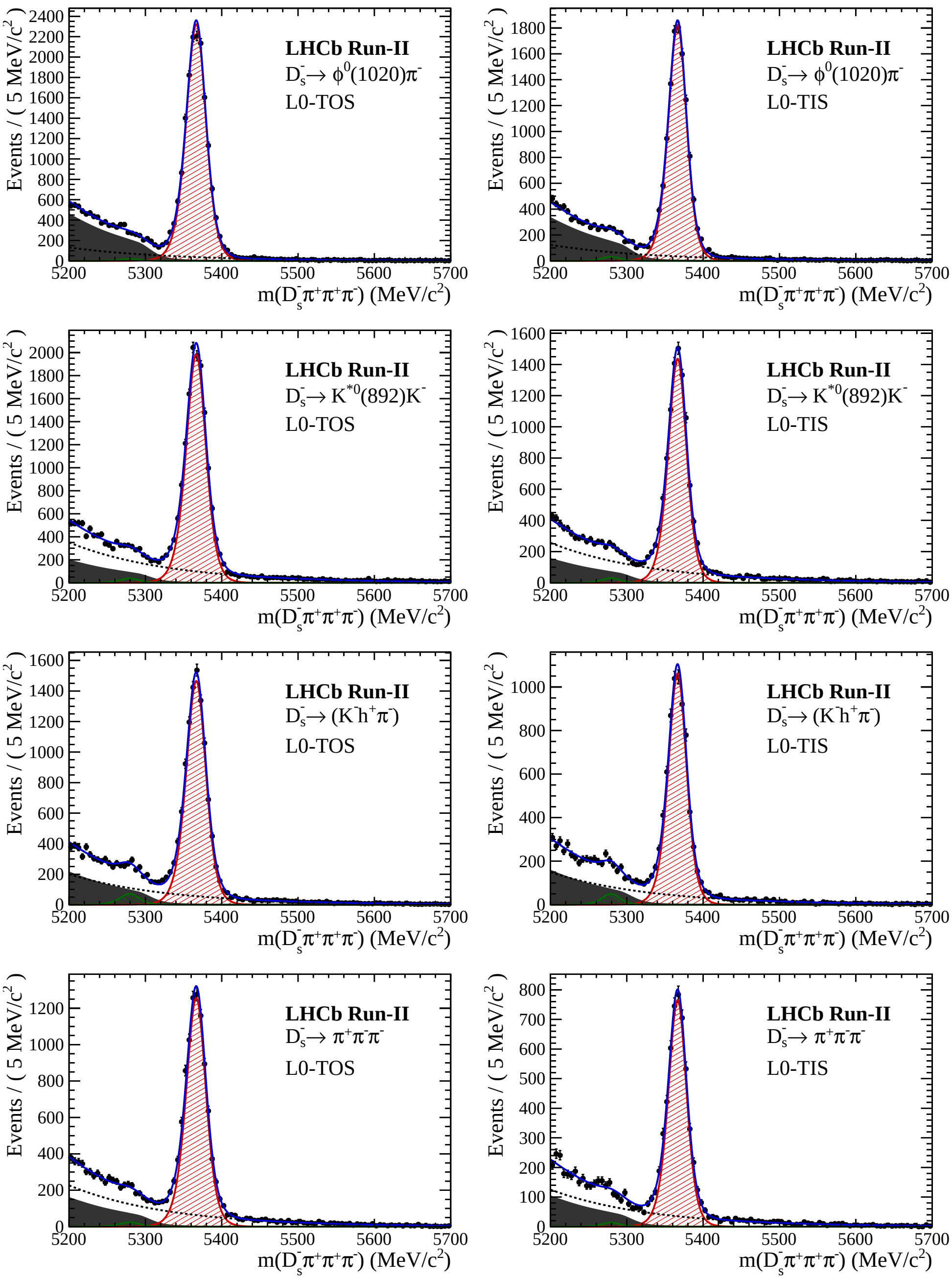

Figure B.2: Invariant mass distributions of $B_s^0 \to D_s\pi\pi\pi$ candidates for Run-II data.

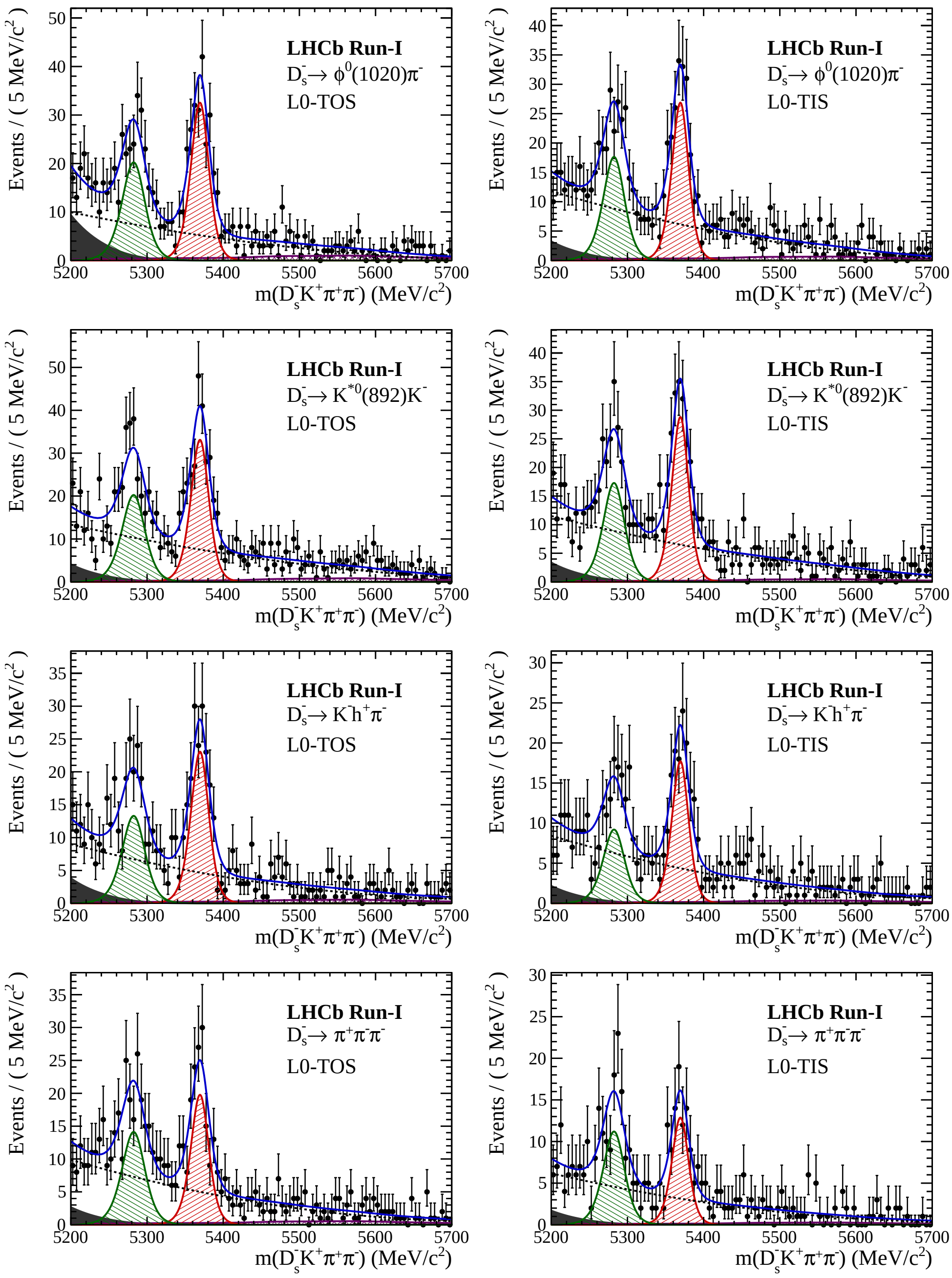

Figure B.3: Invariant mass distributions of $B_s^0 \to D_s K\pi\pi$ candidates for Run-I data.

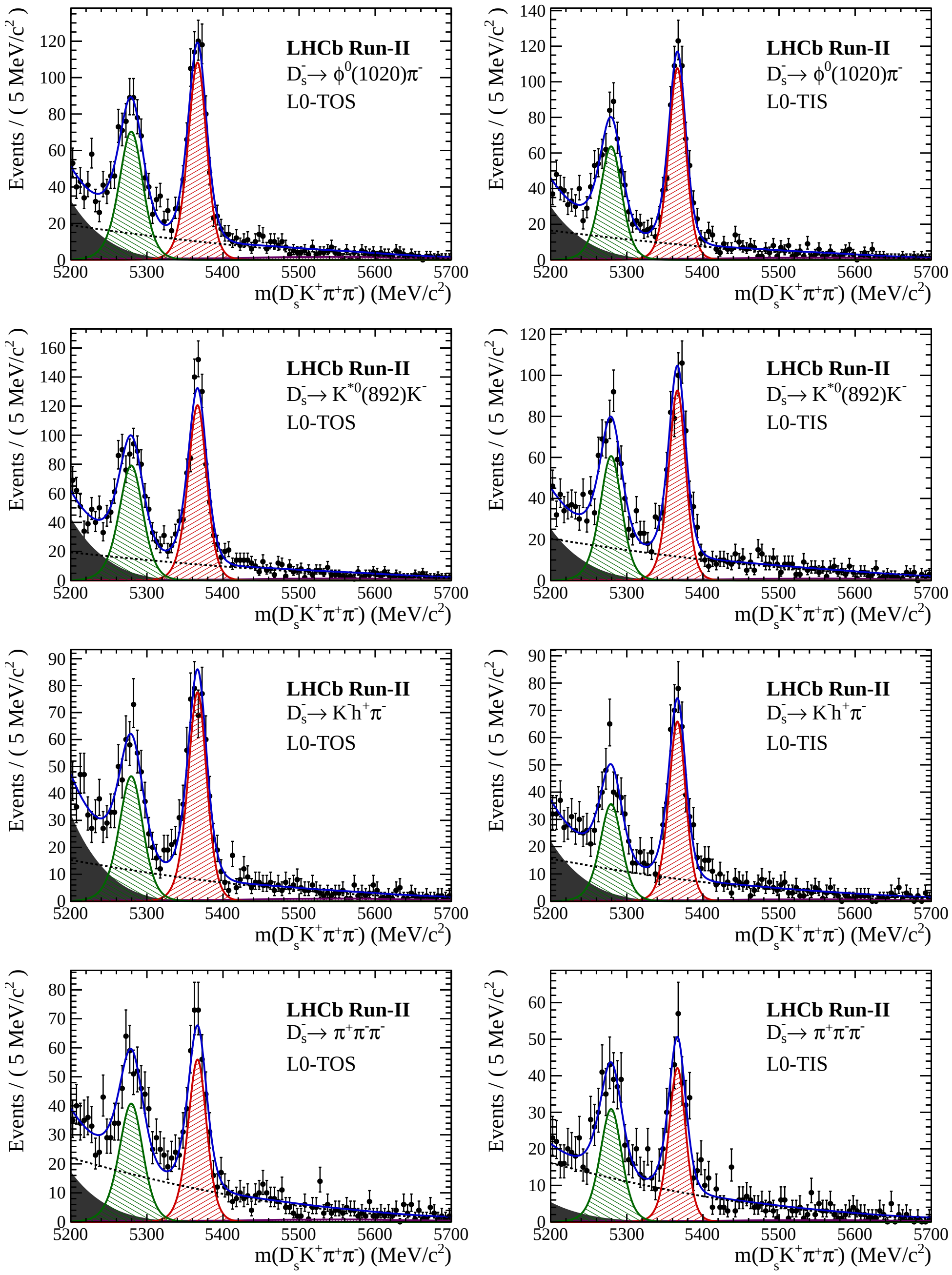

Figure B.4: Invariant mass distributions of $B_s^0 \to D_s K\pi\pi$ candidates for Run-II data.

D Decay-time Resolution fits

This section contains all fits to the distributions of the decay time difference Δt between the true and the reconstructed decay time of the truth-matched B_s^0 candidates on MC. The fits are performed in bins of the decay time error σ_t, where an adaptive binning scheme is used to ensure that approximately the same number of events are found in each bin.

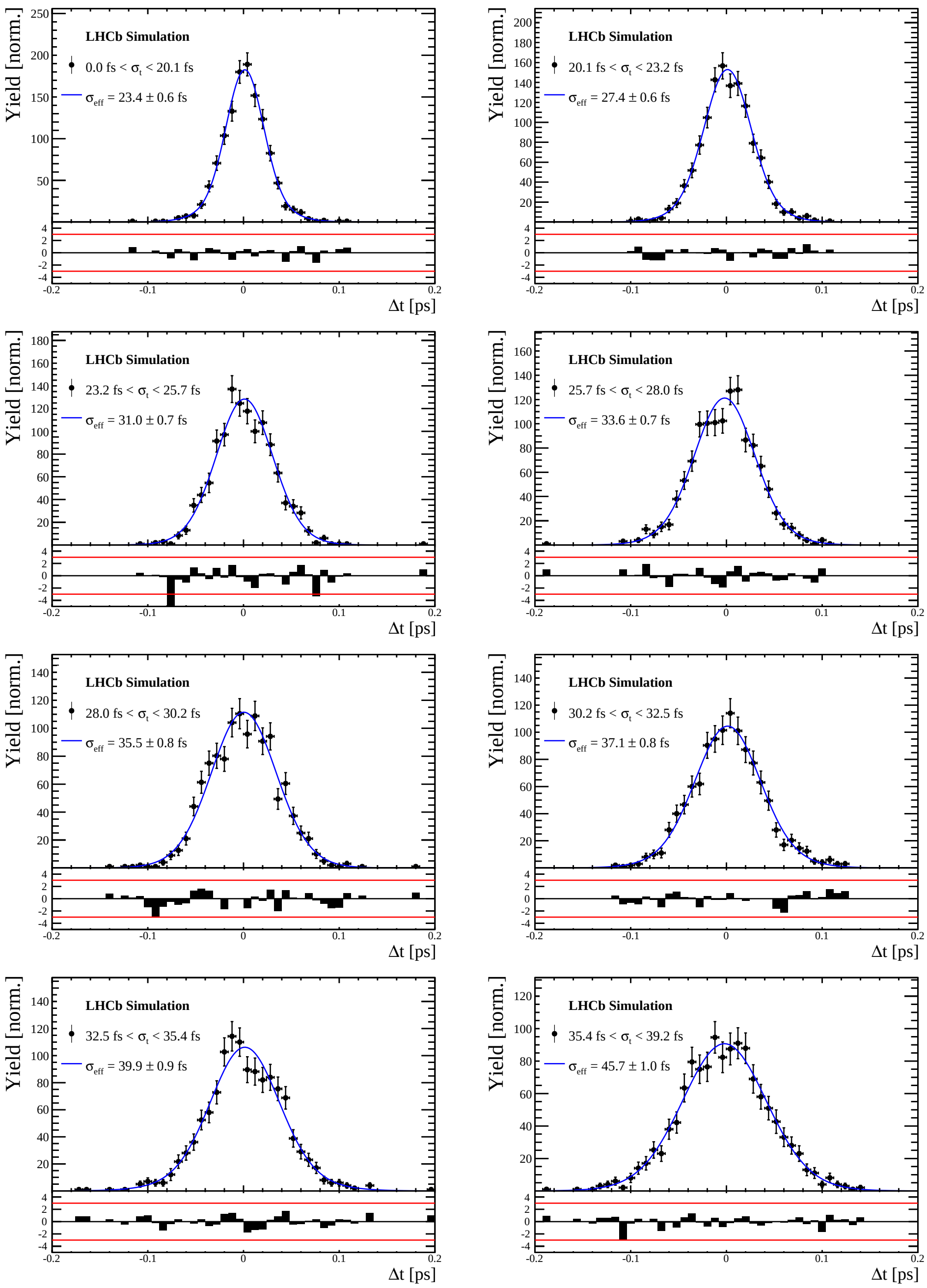

Figure C.1: Difference of the true and measured decay time of $B_s^0 \to D_s K\pi\pi$ MC candidates in bins of the per-event decay time error estimate..

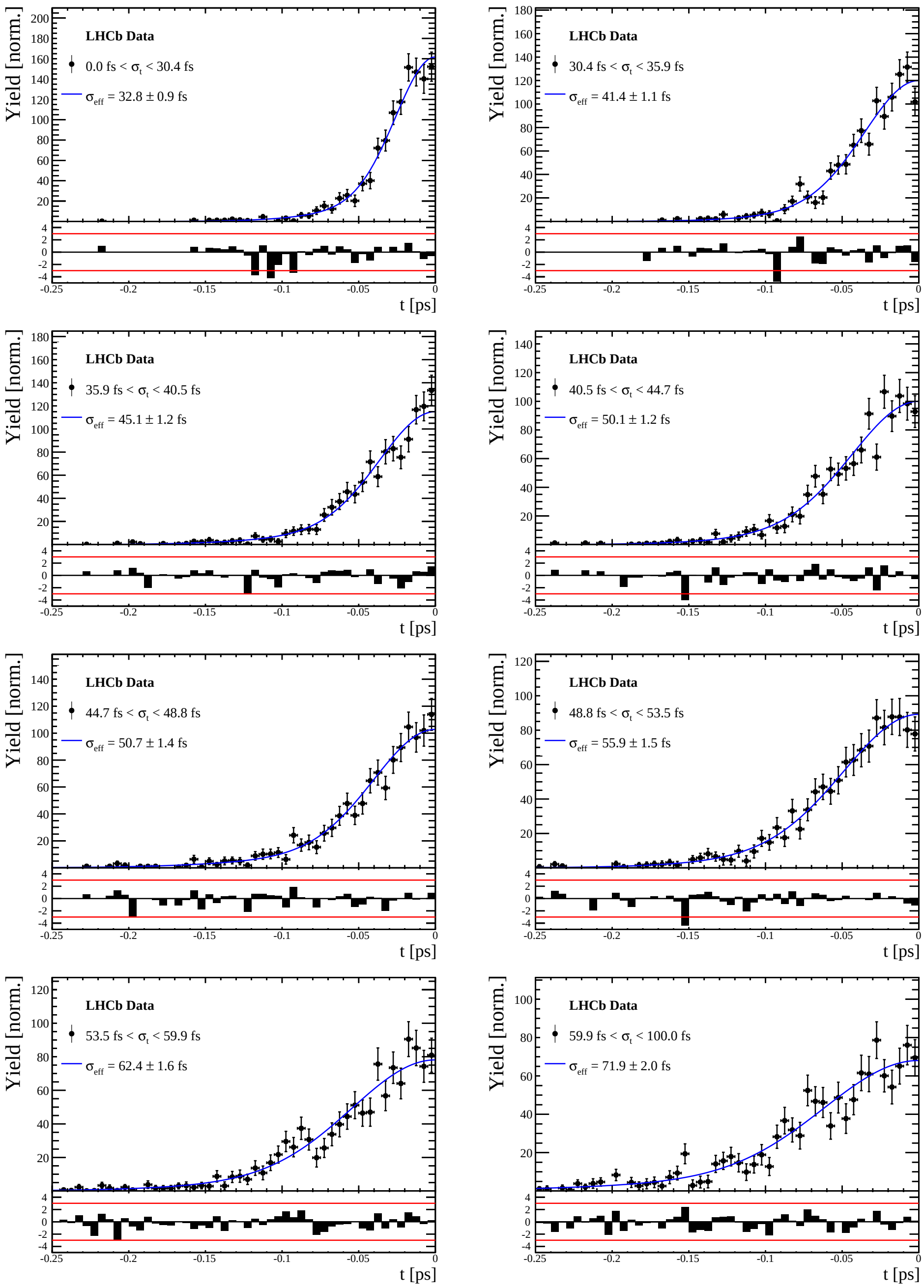

Figure C.2: Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks, for bins in the per-event decay time error estimate.

σ_t Bin [fs]	σ_1 [fs]	σ_2 [fs]	f_1	D	σ_{eff} [fs]
0.0 - 20.1	19 ± 0.675	33.8 ± 1.77	0.75 ± 0	0.917 ± 0.00406	23.4 ± 0.599
20.1 - 23.2	23.4 ± 0.86	37.4 ± 1.95	0.75 ± 0	0.888 ± 0.00477	27.4 ± 0.621
23.2 - 25.7	28.1 ± 1.02	38.7 ± 2.32	0.75 ± 0	0.86 ± 0.00563	31 ± 0.671
25.7 - 28.0	30.1 ± 1.12	43.2 ± 2.56	0.75 ± 0	0.837 ± 0.00651	33.6 ± 0.734
28.0 - 30.2	32.4 ± 1.12	44.2 ± 2.59	0.75 ± 0	0.819 ± 0.00694	35.5 ± 0.756
30.2 - 32.5	32.6 ± 1.38	49.2 ± 3.04	0.75 ± 0	0.805 ± 0.00792	37.1 ± 0.841
32.5 - 35.4	34.4 ± 1.19	54.7 ± 2.85	0.75 ± 0	0.778 ± 0.0086	39.9 ± 0.879
35.4 - 39.2	41.9 ± 1.8	56.9 ± 4.18	0.75 ± 0	0.719 ± 0.00997	45.7 ± 0.962
39.2 - 44.7	42.2 ± 1.56	68.1 ± 4.01	0.75 ± 0	0.687 ± 0.0114	48.8 ± 1.08
44.7 - 120.0	55.5 ± 2.59	83 ± 14.7	0.75 ± 0	0.546 ± 0.0521	62 ± 4.89

Table 4.1: Measured time resolution for $B_s \to D_s K\pi\pi$ MC in bins of the per-event decay time error estimate.

σ_t Bin [fs]	σ_1 [fs]	σ_2 [fs]	f_1	D	σ_{eff} [fs]
0.0 - 30.4	25.4 ± 1.03	50.7 ± 2.77	0.75 ± 0	0.844 ± 0.00822	32.8 ± 0.942
30.4 - 35.9	34.5 ± 1.46	60.2 ± 3.48	0.75 ± 0	0.763 ± 0.0108	41.4 ± 1.08
35.9 - 40.5	35.6 ± 1.35	71.3 ± 3.84	0.75 ± 0	0.726 ± 0.0121	45.1 ± 1.18
40.5 - 44.7	42.3 ± 1.65	73.3 ± 4.21	0.75 ± 0	0.673 ± 0.0132	50.1 ± 1.24
44.7 - 48.8	39.6 ± 1.64	84.8 ± 5.07	0.75 ± 0	0.666 ± 0.0145	50.7 ± 1.36
48.8 - 53.5	47.6 ± 1.94	82.4 ± 5.48	0.75 ± 0	0.611 ± 0.0157	55.9 ± 1.46
53.5 - 59.9	53 ± 2.15	95.3 ± 6.84	0.75 ± 0	0.541 ± 0.0174	62.4 ± 1.63
59.9 - 100.0	60.5 ± 2.8	125 ± 14	0.75 ± 0	0.443 ± 0.0204	71.9 ± 2.03

Table 4.2: Measured time resolution for prompt-D_s data in bins of the per-event decay time error estimate.

E Comparison of time-acceptance in subsamples

Figure C.1 shows the spline coefficients obtained by fitting the decay-time distribution of $B_s^0 \to D_s\pi\pi\pi$ data candidates in different subsamples. Sufficient agreement is observed within a given data-taking period, while the acceptance shapes for Run-I and Run-II data differ significantly. The fitted splines for the different D_s final states are in a good agreement. The largest deviations are observed between the different L0 categories.

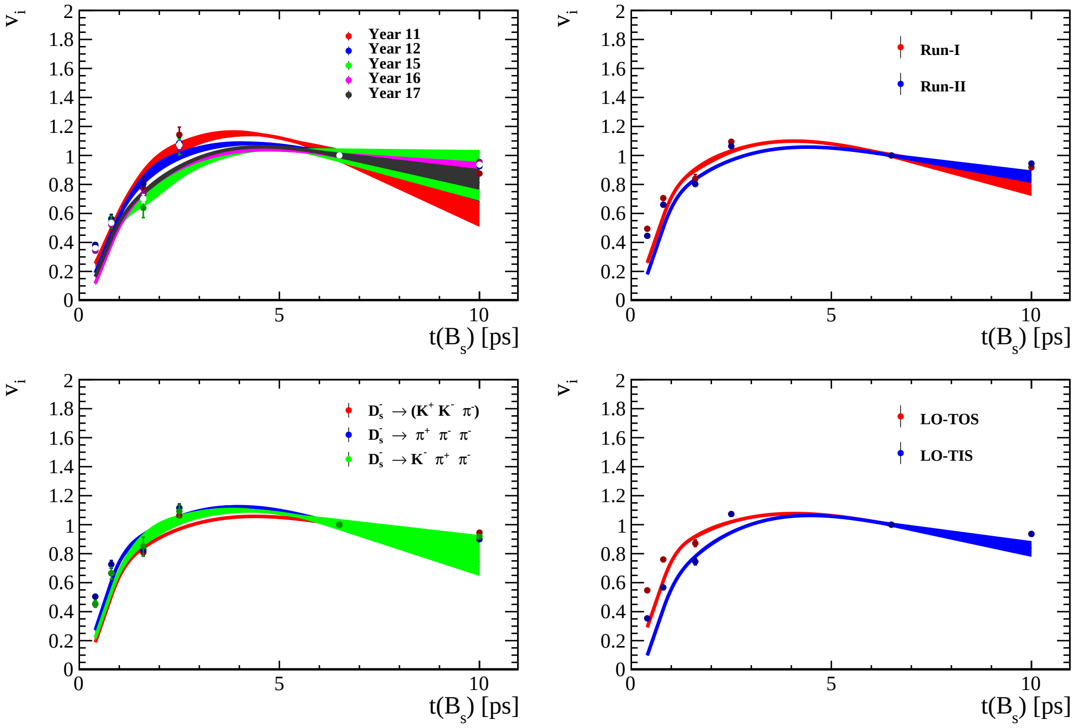

Figure C.1: Comparison of the spline coefficients (point with error bars) obtained from time-dependent fits to the $B_s^0 \to D_s\pi\pi\pi$ decay-time for different subsamples: (top-left) different years of data-taking; (top-right) different data-taking periods; (bottom-left) different D_s final states; (bottom-right) different trigger categories. The interpolated splines are overlaid.

F Spin Amplitudes

The spin factors used for $B \to P_1\, P_2\, P_3\, P_4$ decays are given in Table 6.1.

Table 6.1: Spin factors for all topologies considered in this analysis. In the decay chains, S, P, V, A, T and PT stand for scalar, pseudoscalar, vector, axial vector, tensor and pseudotensor, respectively. If no angular momentum is specified, the lowest angular momentum state compatible with angular momentum conservation and, where appropriate, parity conservation, is used.

Number	Decay chain	Spin amplitude
1	$B \to (P\,P_1),\ P \to (S\,P_2),\ S \to (P_3\,P_4)$	1
2	$B \to (P\,P_1),\ P \to (V\,P_2),\ V \to (P_3\,P_4)$	$L_{(1)\alpha}(P)\ L^{\alpha}_{(1)}(V)$
3	$B \to (A\,P_1),\ A \to (V\,P_2),\ V \to (P_3\,P_4)$	$L_{(1)\alpha}(B)\,P^{\alpha\beta}_{(1)}(A)\,L_{(1)\beta}(V)$
4	$B \to (A\,P_1),\ A[D] \to (P_2\,V),\ V \to (P_3\,P_4)$	$L_{(1)\alpha}(B)\,L^{\alpha\beta}_{(2)}(A)\,L_{(1)\beta}(V)$
5	$B \to (A\,P_1),\ A \to (S\,P_2),\ S \to (P_3\,P_4)$	$L_{(1)\alpha}(B)\,L^{\alpha}_{(1)}(A)$
6	$B \to (A\,P_1),\ A \to (T\,P_2),\ T \to (P_3\,P_4)$	$L_{(1)\alpha}(B)\,L_{(1)\beta}(A)\,L^{\alpha\beta}_{(2)}(T)$
7	$B \to (V_1\,P_1),\ V_1 \to (V_2\,P_2),\ V_2 \to (P_3\,P_4)$	$L_{(1)\mu}(B)\,P^{\mu\alpha}_{(1)}(V_1)\,\epsilon_{\alpha\beta\gamma\delta}\,L^{\beta}_{(1)}(V_1)\,p^{\gamma}_{V_1}\ L^{\delta}_{(1)}(V_2)$
8	$B \to (PT\,P_1),\ PT \to (V\,P_2),\ V \to (P_3\,P_4)$	$L_{(2)\alpha\beta}(B)\,P^{\alpha\beta\gamma\delta}_{(2)}(PT)\,L_{(1)\gamma}(PT)\,L_{(1)\delta}(V)$
9	$B \to (PT\,P_1),\ PT \to (S\,P_2),\ S \to (P_3\,P_4)$	$L_{(2)\alpha\beta}(B)\,L_{(2)}(PT)^{\alpha\beta}$
10	$B \to (PT\,P_1),\ PT \to (T\,P_2),\ T \to (P_3\,P_4)$	$L_{(2)\alpha\beta}(B)\,P^{\alpha\beta\gamma\delta}_{(2)}(PT)\,L_{(2)\gamma\delta}(T)$
11	$B \to (T\,P_1),\ T \to (V\,P_2),\ V \to (P_3\,P_4)$	$L_{(2)\mu\nu}(B)\,P^{\mu\nu\rho\alpha}_{(2)}(T)\,\epsilon_{\alpha\beta\gamma\delta}\,L^{\beta}_{(2)\rho}(T)\,p^{\gamma}_{T}\,P^{\delta\sigma}_{(1)}(T)\,L_{(1)\sigma}(V)$
12	$B \to (T_1\,P_1),\ T_1 \to (T_2\,P_2),\ T_2 \to (P_3\,P_4)$	$L_{(2)\mu\nu}(B)\,P^{\mu\nu\rho\alpha}_{(2)}(T_1)\,\epsilon_{\alpha\beta\gamma\delta}\,L^{\beta}_{(1)}(T_1)\,p^{\gamma}_{T_1}\ L^{\delta}_{(2)\rho}(T_2)$
13	$B \to (S_1\,S_2),\ S_1 \to (P_1\,P_2),\ S_2 \to (P_3\,P_4)$	1
14	$B \to (V\,S),\ V \to (P_1\,P_2),\ S \to (P_3\,P_4)$	$L_{(1)\alpha}(B)\,L^{\alpha}_{(1)}(V)$
15	$B \to (V_1\,V_2),\ V_1 \to (P_1\,P_2),\ V_2 \to (P_3\,P_4)$	$L_{(1)\alpha}(V_1)\,L^{\alpha}_{(1)}(V_2)$
16	$B[P] \to (V_1\,V_2),\ V_1 \to (P_1\,P_2),\ V_2 \to (P_3\,P_4)$	$\epsilon_{\alpha\beta\gamma\delta}\,L^{\alpha}_{(1)}(B)\,L^{\beta}_{(1)}(V_1)\,L^{\gamma}_{(1)}(V_2)\,p^{\delta}_{B}$
17	$B[D] \to (V_1\,V_2),\ V_1 \to (P_1\,P_2),\ V_2 \to (P_3\,P_4)$	$L_{(2)\alpha\beta}(B)\,L^{\alpha}_{(1)}(V_1)\,L^{\beta}_{(1)}(V_2)$
18	$B \to (T\,S),\ T \to (P_1\,P_2),\ S \to (P_3\,P_4)$	$L_{(2)\alpha\beta}(B)\,L^{\alpha\beta}_{(2)}(T)$
19	$B \to (V\,T),\ T \to (P_1\,P_2),\ V \to (P_3\,P_4)$	$L_{(1)\alpha}(B)\,L^{\alpha\beta}_{(2)}(T)\,L_{(1)\beta}(V)$
20	$B[D] \to (T\,V),\ T \to (P_1\,P_2),\ V \to (P_3\,P_4)$	$\epsilon_{\alpha\beta\delta\gamma}\,L^{\alpha\mu}_{2}(B)\,L^{\beta}_{2\mu}\,L^{\gamma}_{(1)}(V)\,p^{\delta}_{B}$
21	$B \to (T_1\,T_2),\ T_1 \to (P_1\,P_2),\ T_2 \to (P_3\,P_4)$	$L_{(2)\alpha\beta}(T_1)\,L^{\alpha\beta}_{(2)}(T_2)$
22	$B[P] \to (T_1\,T_2),\ T_1 \to (P_1\,P_2),\ T_2 \to (P_3\,P_4)$	$\epsilon_{\alpha\beta\gamma\delta}\,L^{\alpha}_{(1)}(B)\,L^{\beta\mu}_{(2)}(T_1)\,L^{\gamma}_{(2)\mu}(T_2)\,p^{\delta}_{B}$
23	$B[D] \to (T_1\,T_2),\ T_1 \to (P_1\,P_2),\ T_2 \to (P_3\,P_4)$	$L_{(2)\alpha\beta}(B)\,L^{\alpha\gamma}_{(2)}(T_1)\,L^{\beta}_{(2)\gamma}(T_2)$

G Considered Decay Chains

The various decay channels considered in the model building are listed in Table 7.1.

Table 7.1: Decays considered in the LASSO model building.

Decay channel
$B_s \to D_s^- \, [K_1(1270)^+[S,D] \to \pi^+ \, K^*(892)^0]$
$B_s \to D_s^- \, [K_1(1270)^+ \to \pi^+ \, K^*(1430)^0]$
$B_s \to D_s^- \, [K_1(1270)^+[S,D] \to K^+ \, \rho(770)^0]$
$B_s \to D_s^- \, [K_1(1270)^+[S,D] \to K^+ \, \omega(782)]$
$B_s \to D_s^- \, [K_1(1400)^+[S,D] \to \pi^+ \, K^*(892)^0]$
$B_s \to D_s^- \, [K_1(1400)^+[S,D] \to K^+ \, \rho(770)^0]$
$B_s \to D_s^- \, [K(1460)^+ \to \pi^+ \, \kappa]$
$B_s \to D_s^- \, [K(1460)^+ \to K^+ \, \sigma]$
$B_s \to D_s^- \, [K(1460)^+ \to \pi^+ \, K^*(892)^0]$
$B_s \to D_s^- \, [K(1460)^+ \to K^+ \, \rho(770)^0]$
$B_s \to D_s^- \, [K^*(1410)^+ \to \pi^+ \, K^*(892)^0]$
$B_s \to D_s^- \, [K^*(1410)^+ \to K^+ \, \rho(770)^0]$
$B_s \to D_s^- \, [K_2^*(1430)^+ \to \pi^+ \, K^*(892)^0]$
$B_s \to D_s^- \, [K_2^*(1430)^+ \to K^+ \, \rho(770)^0]$
$B_s \to D_s^- \, [K^*(1680)^+ \to \pi^+ \, K^*(892)^0]$
$B_s \to D_s^- \, [K^*(1680)^+ \to K^+ \, \rho(770)^0]$
$B_s \to D_s^- \, [K_2(1770)^+ \to \pi^+ \, K^*(892)^0]$
$B_s \to D_s^- \, [K_2(1770)^+ \to K^+ \, \rho(770)^0]$
$B_s \to \sigma^0 (D_s^- K^+)_S$
$B_s[S,P,D] \to \sigma^0 (D_s^- K^+)_V$
$B_s \to \rho(770)^0 (D_s^- K^+)_S$
$B_s[S,P,D] \to \rho(770)^0 \, (D_s^- K^+)_V$
$B_s \to K^*(892)^0 \, (D_s^- \pi^+)_S$
$B_s[S,P,D] \to K^*(892)^0 \, (D_s^- \pi^+)_V$
$B_s \to (D_s^- K^+)_S \, (\pi^+ \pi^-)_S$

H Data-simulation comparisson

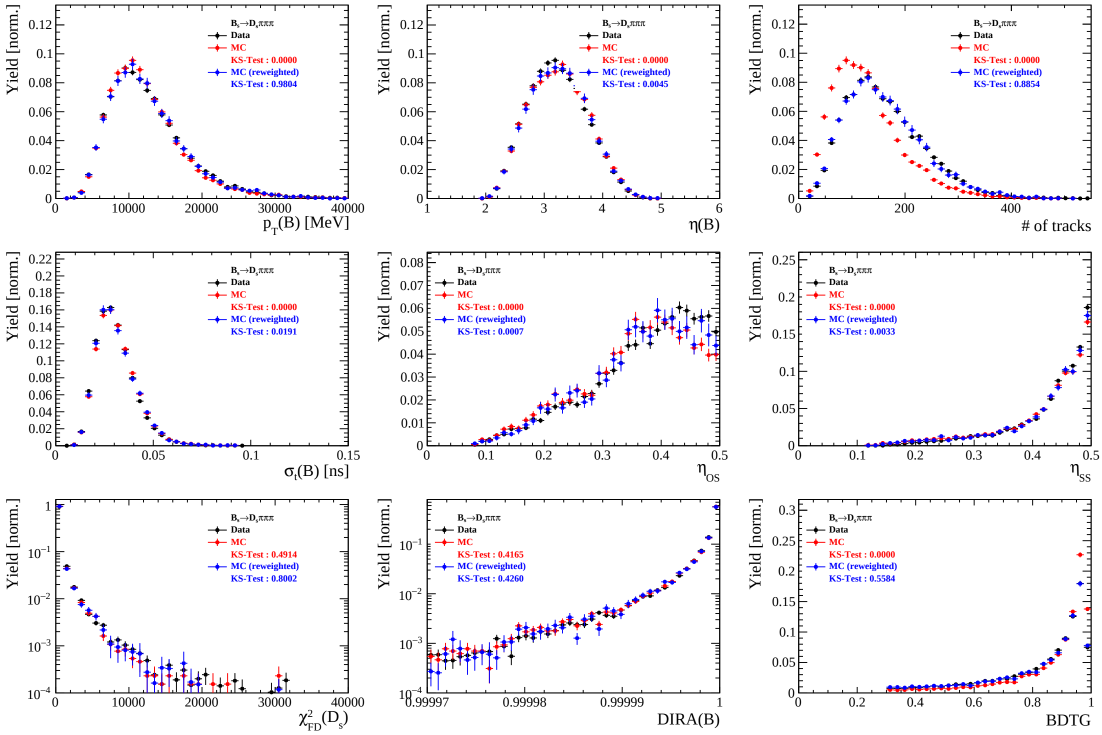

Figure C.1: Comparison between data and MC of selected variables for $B_s \to D_s\pi\pi\pi$ decays.

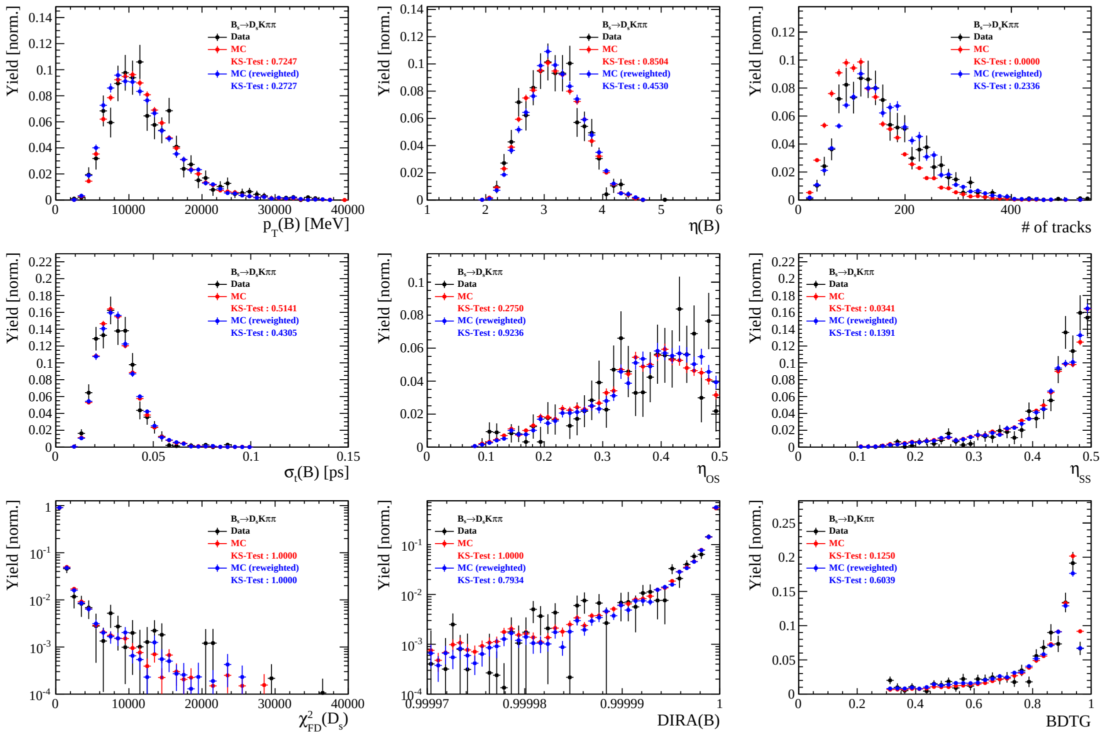

Figure C.2: Comparison between data and MC of selected variables for $B_s \to D_sK\pi\pi$ decays.

I Data distributions

I.1 Comparison of signal and calibration channels

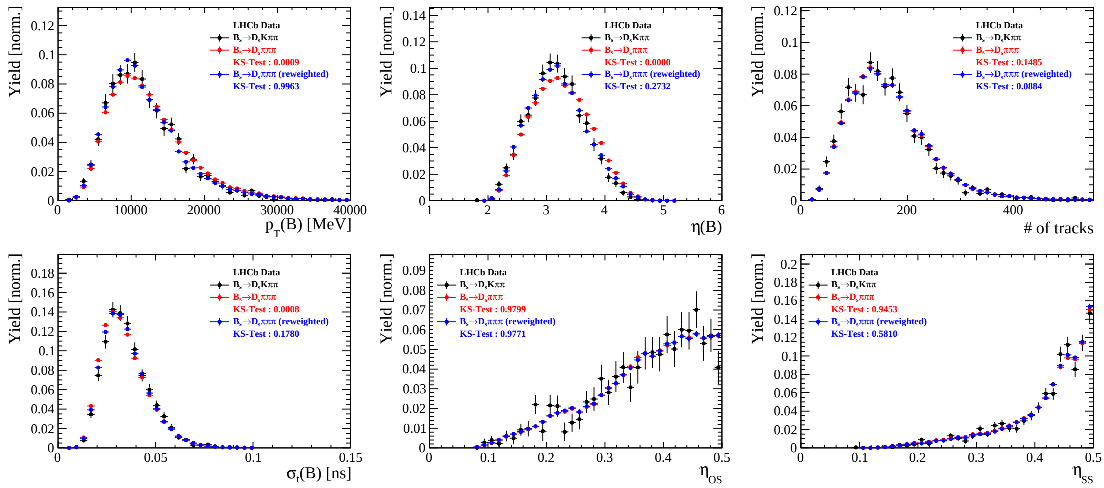

Figure C.1: Comparison between $B_s \to D_s K\pi\pi$ and $B_s \to D_s\pi\pi\pi$ decays for selected variables.

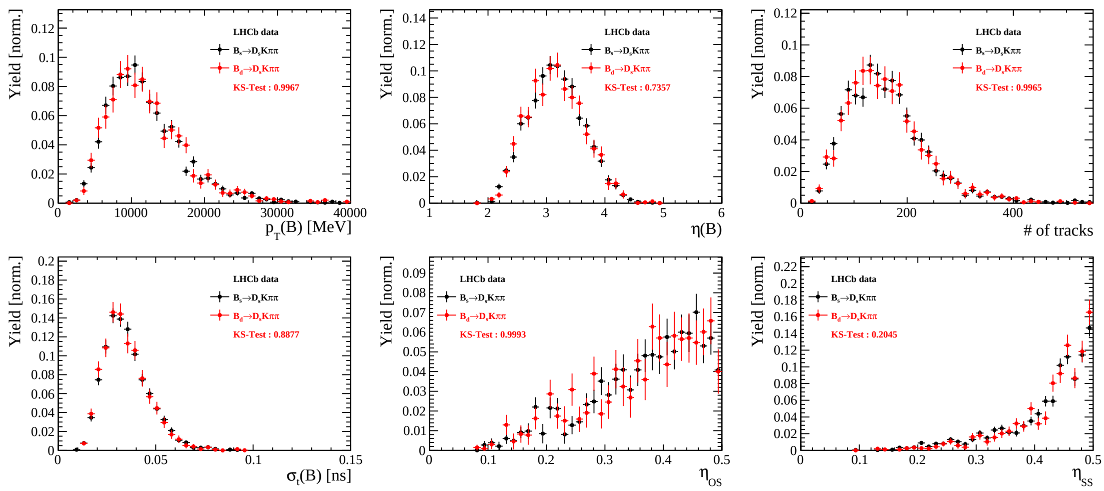

Figure C.2: Comparison between $B_s \to D_s K\pi\pi$ and $B_d \to D_s K\pi\pi$ decays for selected variables.

I.2 Comparison of Run-I and Run-II data

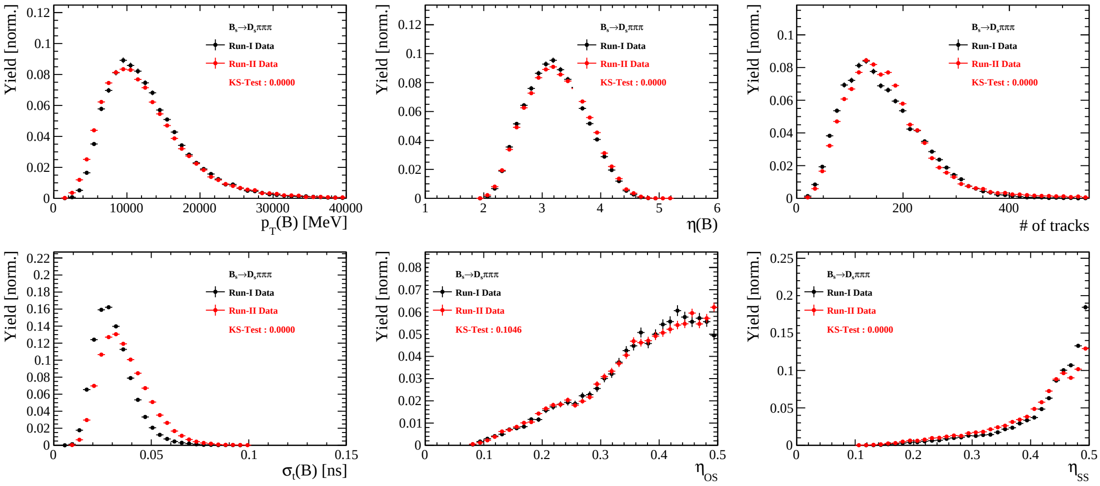

Figure C.3: Comparison of selected variables for Run-I and Run-II data.

I.3 Comparison of D_s final states

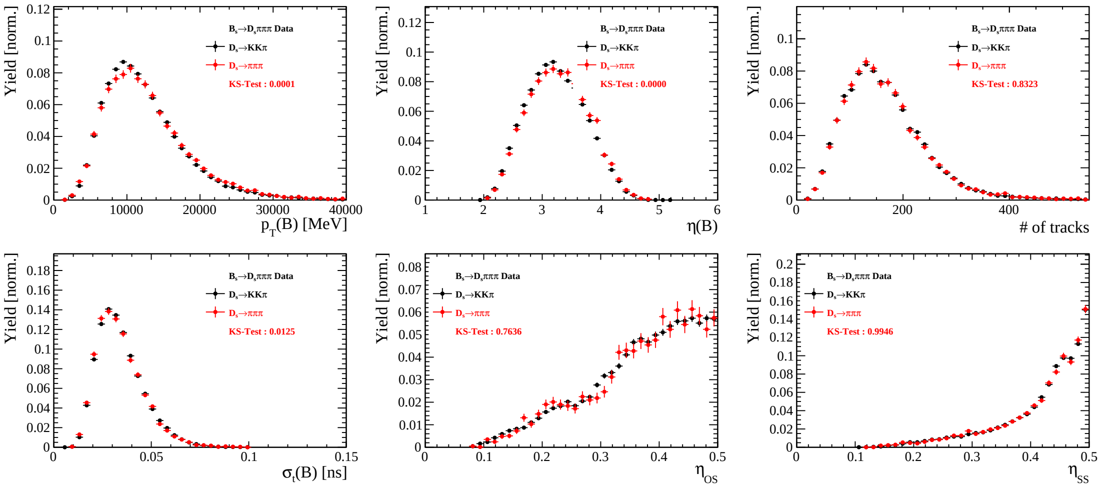

Figure C.4: Comparison of selected variables for different D_s final states.

I.4 Comparison of trigger categories

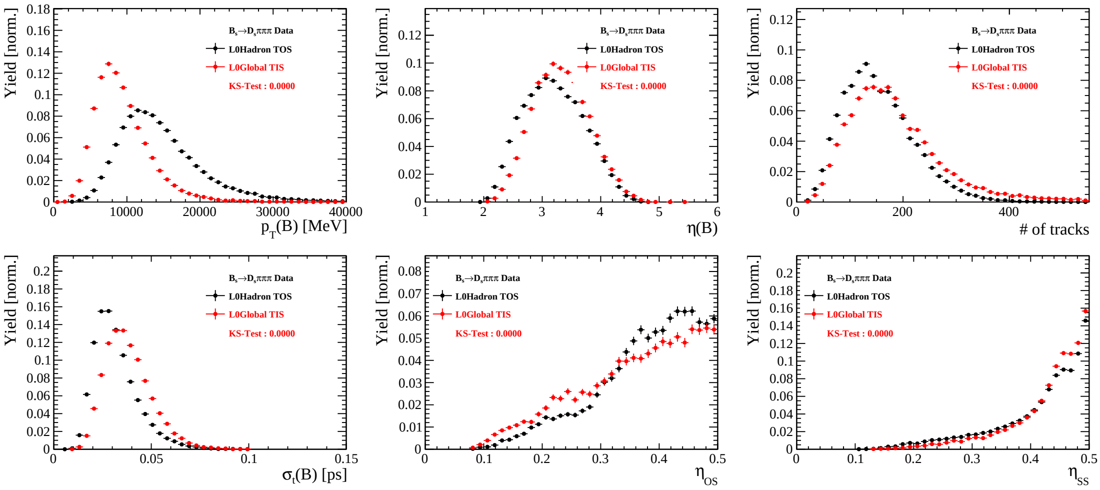

Figure C.5: Comparison of selected variables for different trigger categories.

References

[1] S. Blusk, *First observations and measurements of the branching fractions for the decays* $\bar{B}_s^0 \to D_s^+ K^- \pi^+ \pi^-$ *and* $\bar{B}^0 \to D_s^+ K^- \pi^+ \pi^-$, .

[2] LHCb, S. Blusk, *Measurement of the CP observables in* $\bar{B}_s^0 \to D_s^+ K^-$ *and first observation of* $\bar{B}_{(s)}^0 \to D_s^+ K^- \pi^+ \pi^-$ *and* $\bar{B}_s^0 \to D_{s1}(2536)^+ \pi^-$, 2012. `arXiv:1212.4180`.

[3] R. Fleischer, *New strategies to obtain insights into CP violation through* $B_s \to D_s^\pm K^\mp, D_s^{*\pm} K^\mp, ...andB_d \to D^\pm \pi^\mp, D^{*\pm} \pi^\mp, ...$ *decays*, Nucl. Phys. **B671** (2003) 459, `arXiv:hep-ph/0304027`.

[4] K. De Bruyn *et al.*, *Exploring* $B_s \to D_s^{(*)\pm} K^\mp$ *Decays in the Presence of a Sizable Width Difference* $\Delta\Gamma_s$, Nucl. Phys. **B868** (2013) 351, `arXiv:1208.6463`.

[5] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.

[6] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation of production processes*, Annals of Physics **18** (1962), no. 2 198 .

[7] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*, Phys. Rev. D **11** (1975) 3165.

[8] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of Physics **108** (1977), no. 2 454 .

[9] P. d'Argent *et al.*, *Amplitude Analyses of* $D^0 \to \pi^+ \pi^- \pi^+ \pi^-$ *and* $D^0 \to K^+ K^- \pi^+ \pi^-$ *Decays*, JHEP **05** (2017) 143, `arXiv:1703.08505`.

[10] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.

[11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo Cimento Series 10 **34** (1964), no. 6 1644.

[12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.

[13] D. V. Bugg, *The mass of the* σ *pole*, Journal of Physics G Nuclear Physics **34** (2007) 151, `arXiv:hep-ph/0608081`.

[14] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-dominance prediction for* $\rho \to e^+ e^-$, Phys. Rev. Lett. **21** (1968) 244.

[15] M. S. et al. *Search for CP violation in the* $D^0 \to KK\pi\pi$ *decay through a full amplitude analysis*, LHCb-ANA-2017-064.

[16] S. M. Flatté, *Coupled-channel analysis of the* $\pi\eta$ *and* KK *systems near* KK *threshold*, Physics Letters B **63** (1976), no. 2 224 .

[17] BES Collaboration, M. Ablikim *et al.*, *Resonances in* $J/\psi \to \phi \pi^+ \pi^-$ *and* $\phi K^+ K^-$, Phys. Lett. **B607** (2005) 243, `arXiv:hep-ex/0411001`.

[18] W. Dunwoodie. Fits to $K\pi$ $I = \frac{1}{2}$ $S-$wave amplitude and phase data.

[19] D. Aston *et al.*, *A Study of $K^-\pi^+$ Scattering in the Reaction $K^-p \rightarrow K^-\pi^+n$ at $11 GeV/c$*, Nucl. Phys. **B296** (1988) 493.

[20] BaBar, B. Aubert *et al.*, *Dalitz-plot analysis of the decays $B^\pm \rightarrow K^\pm\pi^\mp\pi^\pm$*, Phys. Rev. **D72** (2005) 072003, `arXiv:hep-ex/0507004`, [Erratum: Phys. Rev.D74,099903(2006)].

[21] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in $D^0 \rightarrow K^\mp\pi^\pm\pi^\pm\pi^\mp$ decays*, Eur. Phys. J. **C78** (2018), no. 6 443, `arXiv:1712.08609`.

[22] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.

[23] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys. Rev. **60** (1941) 61.

[24] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys. Rev. D **57** (1998) 431.

[25] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of ψ decay to mesons*, Eur. Phys. J. **A16** (2003) 537, `arXiv:hep-ph/0211457`.

[26] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. **D51** (1995) 2247.

[27] J.-J. Zhu, *Explicit expressions of spin wave functions*, `arXiv:hep-ph/9906250`.

[28] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput. Phys. Commun. **180** (2009) 1847, `arXiv:0805.2956`.

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

[30] M. Karbach and M. Kenzie, *Gammacombo package*, http://gammacombo.hepforge.org/web/HTML/index.html, 2014.

[31] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[32] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.

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

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

[35] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*, Submitted to: JHEP (2017) `arXiv:1712.07428`.

[36] LHCb collaboration, L. Zhang, *Measurements of CP violation in $B_s^0 \to J/\psi K^+K^-$ decays in the low K^+K^- mass range with 13 TeV data*, LHCb-ANA-2017-028.

[37] A. Poluektov, *Correction of simulated particle identification response in LHCb using kernel density estimation*, LHCb-INT-2017-007.

[38] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, `arXiv:1412.7515`, updated results and plots available at `http://www.slac.stanford.edu/xorg/hfag/`.

[39] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, `arXiv:1407.0748`.

[40] LHCb Collaboration, R. Aaij *et al.*, *Measurement of angular and cp asymmetries in $D^0 \to \pi^+\pi^-\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$ decays*, Phys. Rev. Lett. **121** (2018) 091801.

[41] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, `arXiv:1202.4979`.

[42] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016), no. 05 P05010, `arXiv:1602.07252`.

[43] LHCb, 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`.

[44] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of the $K\pi$ Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-2012-027, CERN, Geneva, Feb, 2013.

[45] A. Davis *et al.*, *Measurement of the instrumental asymmetry for $K^-\pi^+$-pairs at LHCb in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN, Geneva, Mar, 2018.

[46] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly forbidden transitions in D —¿ K(S), K(L) + pi's decays*, Phys. Lett. **B349** (1995) 363, `arXiv:hep-ph/9502238`.

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

[48] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal Statistical Society, Series B **58** (1994) 267.

[49] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude analysis*, JINST **10** (2015), no. 09 P09002, `arXiv:1505.05133`.

[50] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.

[51] CLEO Collaboration, M. Artuso *et al.*, *Amplitude analysis of $D^0 \to K^+K^-\pi^+\pi^-$*, Phys. Rev. **D85** (2012) 122002, `arXiv:1201.5716`.

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

[53] D. Hill, M. John, and P. Gandini, *A study of partially reconstructed $B^{\pm} \to D^{*0}h^{\pm}$ decays using the $D^0 \to K\pi, KK, \pi\pi$ final states*, .

[54] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[55] L. collaboration, *Precision measurement of the B_s oscillation frequency with the decay $B_s \to D_s\pi$*, LHCb-ANA-2012-053.