

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

P. d'Argent¹, E. Gersabeck², M. Kecke¹, M. Schiller³

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 $B_s^0 \rightarrow D_s K \pi\pi$ decays using proton-proton collision data corresponding to an integrated luminosity of 7 fb^{-1} recorded by the LHCb detector.

Contents

1	Introduction	1
2	Formalism	2
2.1	Decay rates and CP -observables	2
2.2	Amplitude model	4
2.2.1	Form Factors and Resonance Lineshapes	4
2.2.2	Spin Densities	6
2.3	Fit implementation	8
2.4	Validation	9
3	Data samples and event selection	12
3.1	Stripping and Trigger selection	12
3.2	Offline selection	12
3.2.1	Phase space region	13
3.2.2	Physics background vetoes	14
3.2.3	Training of multivariate classifier	17
3.2.4	Final selection	19
3.3	Simulation	22
4	Yields determination	23
4.1	Signal model	23
4.2	Background models	24
4.3	Results	25
5	Decay-time Resolution	27
5.1	Calibration for Run-I data	28
5.2	Calibration for Run-II data	29
6	Acceptance	32
6.1	MC corrections	32
6.1.1	Truth matching of simulated candidates	32
6.1.2	Correction of data-simulation differences	32
6.2	Decay-time acceptance	33
6.3	Phase space acceptance	38
7	Flavour Tagging	41
7.1	OS tagger combination	41
7.2	Tagging performance	43
8	Production and Detection Asymmetries	45
8.1	B_s Production Asymmetry	45
8.2	$K^-\pi^+$ Detection Asymmetry	46
9	Decay-time fit	49
9.1	Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data	49
9.2	Fit to $B_s^0 \rightarrow D_sK\pi\pi$ data	50

10 Time-dependent amplitude fit	52
10.1 Signal Model Construction	52
10.2 Results	55
11 Systematic uncertainties	58
11.1 Fit bias	58
11.2 Background subtraction	58
11.3 Decay-time acceptance	59
11.4 Decay-time resolution and tagging	59
11.5 Production, detection asymmetries and mixing frequency	60
11.6 Multiple candidates	60
11.7 Length and momentum scales	60
11.8 Phase space acceptance	61
11.9 Resonance description	61
11.10 Alternative amplitude models	62
A Stripping and Trigger cuts	66
B Details of multivariate classifier	68
C Detailed mass fits	72
D Decay-time Resolution fits	76
E Comparison of time-acceptance in subsamples	79
F Spin Amplitudes	80
G Considered Decay Chains	81
H Data-simulation comparisson	82
I Data distributions	83
I.1 Comparison of signal and calibration channels	83
I.2 Comparison of Run-I and Run-II data	84
I.3 Comparison of D_s final states	84
I.4 Comparison of trigger categories	85
References	86

1 Introduction

This note presents the first measurement of the CKM angle $\gamma \equiv \arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$ using $B_s^0 \rightarrow D_s K\pi\pi$ 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 $b \rightarrow c$ and $b \rightarrow u$ transitions achieved through $B_s^0 - \bar{B}_s^0$ 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.

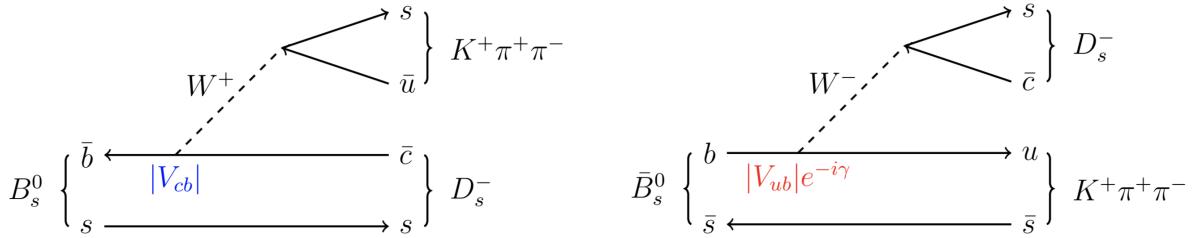


Figure 1.1: Feynman diagram for $B_s^0/\bar{B}_s^0 \rightarrow D_s^- K^+ \pi^+ \pi^-$ decays.

14 2 Formalism

15 2.1 Decay rates and *CP*-observables

16 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
 17 given by:

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

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

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

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

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

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

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

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

⁴¹ Assuming there is no direct CP violation in the B_s decay implies for the CP conjugate
⁴² transition amplitudes:

$$\mathcal{A}(\bar{B}_s^0 \rightarrow \bar{f}) = \mathcal{A}_f^c(\mathbf{x}) = \mathcal{A}_f^c(\bar{\mathbf{x}}) \quad (2.5)$$

$$\mathcal{A}(B_s^0 \rightarrow \bar{f}) = \mathcal{A}_f^u(\mathbf{x}) = \mathcal{A}_f^u(\bar{\mathbf{x}}) \quad (2.6)$$

⁴³ where the CP -conjugate phase space point $\bar{\mathbf{x}}$ is defined such that it is mapped onto \mathbf{x} by
⁴⁴ the interchange of final state charges, and the reversal of three-momenta.

⁴⁵ The phenomenological description of the intermediate state amplitudes is discussed
⁴⁶ in Sec. 2.2. For a model-independent measurement, the differential decay rate can be
⁴⁷ integrated over the phase space:

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

⁴⁸ where the same convention for the CP coefficients as for the $B_s \rightarrow D_s K$ analysis is used:

$$C = \frac{1 - r^2}{1 + r^2} \quad (2.8)$$

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

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

⁴⁹ The coherence factor κ , the strong phase difference δ and the ratio of the suppressed
⁵⁰ ($b \rightarrow u$) over favored ($b \rightarrow c$) decay mode, averaged over the phase space, are defined as:

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

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

⁵¹ The coherence factor dilutes the sensitivity to the weak phase γ due to the integration
⁵² over the interfering amplitudes across the phase space. The value of κ is bounded between
⁵³ zero and unity. The latter corresponds to the limit of only one contributing intermediate
⁵⁴ state in which case the same sensitivity as in $B_s \rightarrow D_s K$ decays is reached, while $\kappa = 0$
⁵⁵ would result in no sensitivity to γ at all.

56 2.2 Amplitude model

57 To construct the intermediate state amplitudes $A_i(\mathbf{x})$, the isobar approach is used, which
 58 assumes that the decay process can be factorized into subsequent two-body decay am-
 59 plitudes [6–8]. This gives rise to two different decay topologies; quasi two-body decays
 60 $B_s \rightarrow (R_1 \rightarrow h_1 h_2) (R_2 \rightarrow h_3 h_4)$ or cascade decays $B_s \rightarrow h_1 [R_1 \rightarrow h_2 (R_2 \rightarrow h_3 h_4)]$. In
 61 either case, the intermediate state amplitude is parameterized as a product of orbital
 62 angular momentum, L , dependent form factors B_L , included for each vertex of the decay
 63 tree, Breit-Wigner propagators T_R , included for each resonance R , and an overall angular
 64 distribution represented by a spin factor S ,

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

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

67 2.2.1 Form Factors and Resonance Lineshapes

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

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

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

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

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

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

$$\Gamma_{R \rightarrow AB}^{(2)}(s) = \Gamma_0 \frac{m_0}{\sqrt{s}} \left(\frac{q}{q_0} \right)^{2L+1} \frac{B_L(q)^2}{B_L(q_0)^2}, \quad (2.16)$$

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

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

$$\Gamma_{R \rightarrow ABC}^{(3)}(s) = \frac{1}{2\sqrt{s}} \int |A_{R \rightarrow ABC}|^2 d\Phi_3, \quad (2.17)$$

and therefore requires knowledge of the resonant substructure. The three-body amplitude $A_{R \rightarrow ABC}$ 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,

$$\Gamma(s) = \sum_i g_i \Gamma_i(s), \quad (2.18)$$

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 $\rho(770)^0 \rightarrow \pi\pi$ propagator [14]. For the decay chain $K_1(1270) \rightarrow \rho(770)K$, 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]:

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

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.

¹¹⁹ **2.2.2 Spin Densities**

¹²⁰ The spin amplitudes are phenomenological descriptions of decay processes that are required
¹²¹ to be Lorentz invariant, compatible with angular momentum conservation and, where
¹²² appropriate, parity conservation. They are constructed in the covariant Zemach (Rarita-
¹²³ Schwinger) tensor formalism [22–24]. At this point, we briefly introduce the fundamental
¹²⁴ objects of the covariant tensor formalism which connect the particle’s four-momenta to
¹²⁵ the spin dynamics of the reaction and give a general recipe to calculate the spin factors
¹²⁶ for arbitrary decay trees. Further details can be found in Refs. [25, 26].

¹²⁷ A spin- S particle with four-momentum p , and spin projection λ , is represented by the
¹²⁸ polarization tensor $\epsilon_{(S)}(p, \lambda)$, which is symmetric, traceless and orthogonal to p . These
¹²⁹ so-called Rarita-Schwinger conditions reduce the a priori 4^S elements of the rank- S tensor
¹³⁰ to $2S + 1$ independent elements in accordance with the number of degrees of freedom of a
¹³¹ spin- S state [23, 27].

¹³² The spin projection operator $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$, for a resonance R , with spin $S =$
¹³³ $\{0, 1, 2\}$, and four-momentum p_R , is given by [26]:

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

¹³⁴ where $g^{\mu\nu}$ is the Minkowski metric. Contracted with an arbitrary tensor, the projection
¹³⁵ operator selects the part of the tensor which satisfies the Rarita-Schwinger conditions.

¹³⁶ For a decay process $R \rightarrow AB$, with relative orbital angular momentum L , between
¹³⁷ particle A and B , the angular momentum tensor is obtained by projecting the rank- L
¹³⁸ tensor $q_R^{\nu_1} q_R^{\nu_2} \dots q_R^{\nu_L}$, constructed from the relative momenta $q_R = p_A - p_B$, onto the spin- L
¹³⁹ subspace,

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

¹⁴⁰ Their $|\vec{q}_R|^L$ dependence accounts for the influence of the centrifugal barrier on the transition
¹⁴¹ amplitudes. For the sake of brevity, the following notation is introduced,

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

¹⁴² Following the isobar approach, a four-body decay amplitude is described as a product
¹⁴³ of two-body decay amplitudes. Each sequential two-body decay $R \rightarrow A B$, with relative
¹⁴⁴ orbital angular momentum L_{AB} , and total intrinsic spin S_{AB} , contributes a term to the
¹⁴⁵ overall spin factor given by

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

¹⁴⁶ where

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

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

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

155 where $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita symbol and j refers to the arguments of X defined in
 156 Eqs. 2.23 and 2.24. Its antisymmetric nature ensures the correct parity transformation
 157 behavior of the amplitude. The spin factor for a whole decay chain, for example $R \rightarrow$
 158 $(R_1 \rightarrow AB)(R_2 \rightarrow CD)$, is obtained by combining the two-body terms and performing a
 159 sum over all unobservable, intermediary spin projections

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

160 where $\lambda_R = \lambda_A = \lambda_B = \lambda_C = \lambda_D = 0$, $S_{AB} = S_{CD} = 0$ and $S_{R_1 R_2} = L_{R_1 R_2}$, as only
 161 pseudoscalar initial/final states are involved.

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

¹⁶⁴ 2.3 Fit implementation

¹⁶⁵ The hadronic amplitudes are renormalized prior to the amplitude fit such that

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

¹⁶⁶ This allows us to set more intuitive starting values as the amplitude coefficients are all on
¹⁶⁷ a comparable scale. Moreover, the total amplitudes $\mathcal{A}_f^{c(u)}(\mathbf{x})$ are renormalized on-the-fly
¹⁶⁸ such that

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

¹⁶⁹ As a result, the average amplitude ratio and strong phase difference between the $b \rightarrow u$ and
¹⁷⁰ $b \rightarrow c$ transitions can be introduced as direct fit parameters instead of derived quantities
¹⁷¹ that have to be calculated from Equation 2.11 after the fit. For the differential decay rate
¹⁷² follows:

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

¹⁷³ This renormalization procedure was found to be crucial for the fit stability since it reduces
¹⁷⁴ the correlation between the a_i^c and a_i^u amplitude coefficients significantly. Due to the
¹⁷⁵ overall normalization, one of the complex amplitude coefficients a_i^c can be fixed to unity
¹⁷⁶ and since r and δ are included as fit parameters one of the complex amplitude coefficient
¹⁷⁷ a_i^u can be additionally fixed to unity.

¹⁷⁸ We force strong decays in the cascade topology to have the same pattern in $b \rightarrow c$
¹⁷⁹ and $b \rightarrow u$ transitions by the sharing of couplings between related quasi-two-body final
¹⁸⁰ states. For example, given the two a_i^c parameters required for $B_s \rightarrow D_s^- K_1(1270)^+$
¹⁸¹ with $K_1(1270)^+ \rightarrow \rho(770)^0 \pi^+$ and $K_1(1270)^+ \rightarrow K^*(892) \pi^+$, the amplitude $\bar{B}_s \rightarrow$
¹⁸² $D_s^- K_1(1270)^+$ with $K_1(1270)^+ \rightarrow \rho(770)^0 \pi^+$ and $K_1(1270)^+ \rightarrow K^*(892) \pi^+$ only requires
¹⁸³ one additional global complex parameter to represent the different production processes
¹⁸⁴ of $B_s \rightarrow D_s^- K_1(1270)^+$ and $\bar{B}_s \rightarrow D_s^- K_1(1270)^+$, while the relative magnitude and phase
¹⁸⁵ of $K_1(1270)^+ \rightarrow \rho(770)^0 \pi^+$ and $K_1(1270)^+ \rightarrow K^*(892) \pi^+$ are the same regardless of
¹⁸⁶ the production mechanism. For this purpose, multiple decay amplitudes of a three-body
¹⁸⁷ resonance are defined relative to a given reference channel.

188 2.4 Validation

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

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

204 The CP coefficients $C, D, \bar{D}, S, \bar{S}$ are the fit parameters in case of the phasespace-
 205 integrated pdf, which are converted after to the fit to the physical observables r, κ, δ and γ
 206 using the **GammaCombo** package [30]. The obtained 1-CL contours are shown in Fig. 2.3.
 207 The full pdf determines r, δ and γ directly. As shown in Tab. 2.2 and 2.3, the fit results
 208 are in excellent agreement with the generated input values. The phasespace-integrated fit
 209 is, in addition, performed with the **B2DX** fitter used for the $B_s \rightarrow D_s K$ analysis yielding
 210 identical results. Note though that some parts of the **B2DX** fitter have been taken over to
 211 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 ps^{-1}
Δm_s	17.757 ps^{-1}
r	0.37
κ	1
δ	10.0°
γ	71.1°
β_s	0.0°

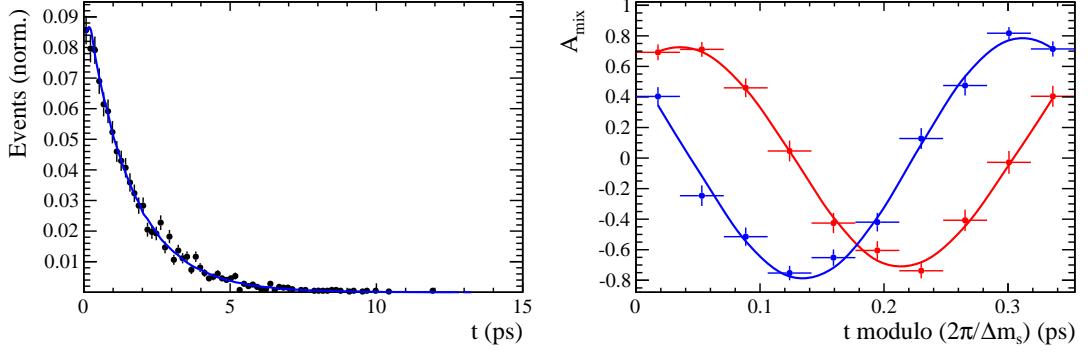


Figure 2.1: Left: Time distribution of $B_s \rightarrow D_s K \pi \pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (solid line). Right: Time-dependent asymmetry between mixed and unmixed events folded into one oscillation period for $D_s^- K^+ \pi \pi$ (red) and $D_s^+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

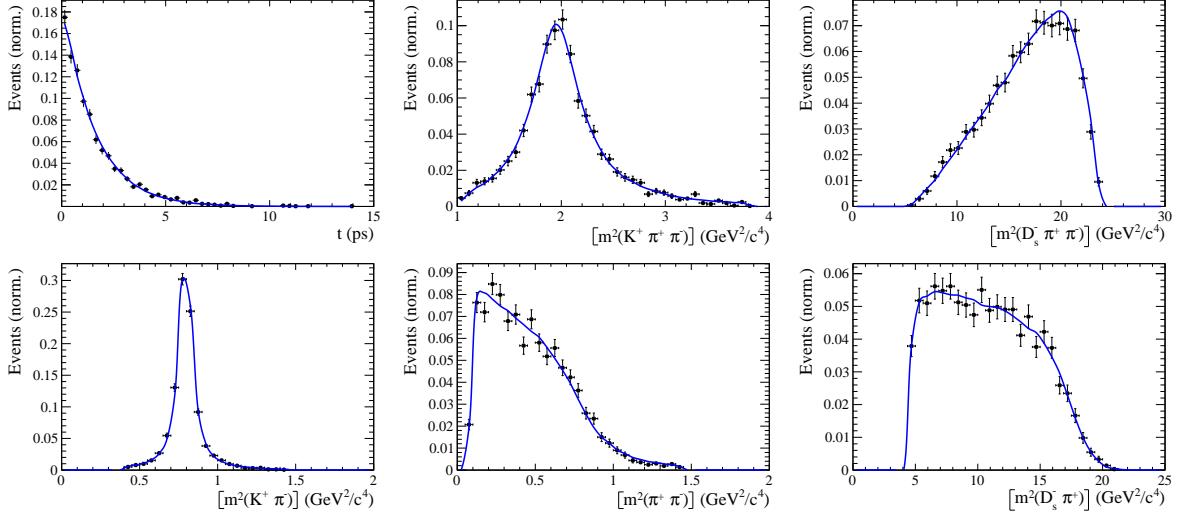


Figure 2.2: Time and invariant mass distributions of $B_s \rightarrow D_s K \pi \pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

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

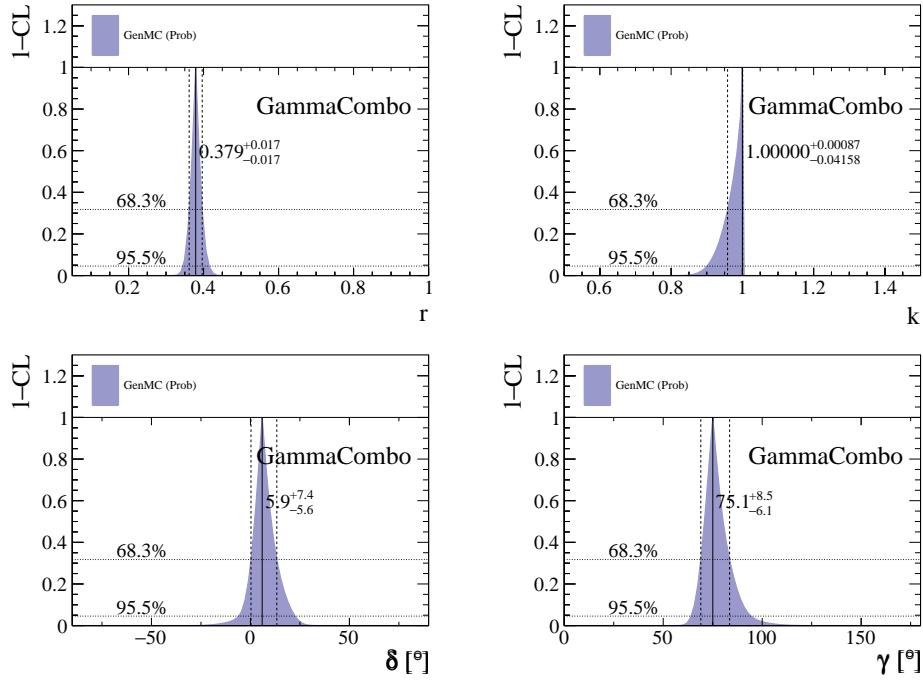


Figure 2.3: The 1-CL contours for the physical observable r, κ, δ and γ obtained with the phasespace-integrated fit to the EVTGEN toy sample.

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

212 3 Data samples and event selection

213 3.1 Stripping and Trigger selection

214 The dataset used for this analysis corresponds to 1 fb^{-1} of proton-proton collision data col-
215 lected in 2011 with a centre of mass energy $\sqrt{s} = 7 \text{ TeV}$, 2 fb^{-1} collected in 2012 with $\sqrt{s} =$
216 7 TeV and 4 fb^{-1} collected in 2015/2016/2017 with $\sqrt{s} = 13 \text{ TeV}$. Candidate $B_s^0 \rightarrow D_s K \pi \pi$
217 ($B_s^0 \rightarrow D_s \pi \pi \pi$) decays are reconstructed using the `B02DKPiPiD2HHHPIDBeauty2CharmLine`
218 (`B02DPiPiD2HHHPIDBeauty2CharmLine`) line of the `BHadronCompleteEvent` stream of
219 `Stripping21r1` (2011), `Stripping21` (2012), `Stripping24r1` (2015) and `Stripping28r1p1` (2016)
220 and `Stripping29r2` (2017). Both stripping lines employ the same selection cuts, listed in
221 Appendix A, except for the PID requirement on the bachelor kaon/pion.

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

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

234 3.2 Offline selection

235 The offline selection, in particular the requirements on the D_s hadron, are guided by
236 the previous analyses of $B_s \rightarrow D_s K/\pi$, $B_d \rightarrow D^0 \pi$ as well as the branching fraction
237 measurement of $B_s^0 \rightarrow D_s K \pi \pi$ decays. Tables 3.1 and 3.2 summarize all selection
238 requirements which are described in the following. Throughout the note, we abbreviate
239 $B_s^0 \rightarrow D_s X_s (\rightarrow K \pi \pi)$ and $B_s^0 \rightarrow D_s X_d (\rightarrow \pi \pi \pi)$.

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

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

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

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

259 We reconstruct the $B_s^0 \rightarrow D_s h\pi\pi$ decay through three different final states of the
260 D_s meson: $D_s \rightarrow KK\pi$, $D_s \rightarrow \pi\pi\pi$ and $D_s \rightarrow K\pi\pi$. Of those, $D_s \rightarrow KK\pi$ is the
261 most prominent one, while $\mathcal{BR}(D_s \rightarrow \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$ and $\mathcal{BR}(D_s \rightarrow$
262 $K\pi\pi) \approx 0.12 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$ holds for the others. For the $KK\pi$ final state we make
263 use of the well known resonance structure; the decay proceeds either via the narrow ϕ
264 resonance, the broader K^{*0} resonance or (predominantly) non-resonant. Within the ϕ
265 resonance region the sample is already sufficiently clean after the stripping so that we
266 do not impose additional criteria on the D_s daughters. For the K^{*0} and non-resonant
267 regions consecutively tighter requirements on the particle identification and the D_s flight-
268 distance are applied. We apply global requirements for the other final states. All cuts are
269 summarized in Table 3.1.

270 3.2.1 Phase space region

271 Due to the comparably low masses of the final state particles with respect to the B_s
272 mass, there is, in contrast to for example b-hadron to charmonia or charm decays, a
273 huge phase-space available for the $B_s^0 \rightarrow D_s K\pi\pi$ decay. For the invariant mass of
274 the $K\pi\pi$ subsystem it extends up to 3.4 GeV. It has however been observed that the
275 decay proceeds predominantly through the low lying axial vector states $K_1(1270)$ and
276 $K_1(1400)$, while the combinatorial background is concentrated at high $K\pi\pi$ invariant
277 masses ($m(K\pi\pi) > 2000$ MeV). Moreover, the strange hadron spectrum above 2 GeV
278 is poorly understood experimentally such that a reliable extraction of the strong phase
279 motion in that region is not possible. We consequently choose to limit the considered
280 phase space region to $m(K\pi\pi) < 1950$ MeV, which is right below the charm-strange
281 threshold ($B_s^0 \rightarrow D_s^+ D_s^-$).

282 **3.2.2 Physics background vetoes**

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

287 1. $D_s^- \rightarrow K^+ K^- \pi^-$

288 (a) $D^- \rightarrow K^+ \pi^- \pi^-$:

289 Possible with $\pi^- \rightarrow K^-$ misidentification, vetoed by requiring $m(K^+ K_\pi^- \pi^-) \neq$
 290 $m(D^-) \pm 40$ MeV or the K^- has to fulfill more stringent PID criteria depending
 291 on the resonant region (see Table 3.1).

292 (b) $\Lambda_c^- \rightarrow K^+ \bar{p} \pi^-$:

293 Possible with $\bar{p} \rightarrow K^-$ misidentification, vetoed by requiring $m(K^+ K_p^- \pi^-) \neq$
 294 $m(\Lambda_c^-) \pm 40$ MeV or the K^- has to fulfill more stringent PID criteria depending
 295 on the resonant region (see Table 3.1).

296 (c) $D^0 \rightarrow KK$:

297 D^0 combined with a random π can fake a $D_s \rightarrow KK\pi$ decay, vetoed by
 298 requiring $m(KK) < 1840$ MeV.

299 2. $D_s^- \rightarrow \pi^+ \pi^- \pi^-$

300 (a) $D^0 \rightarrow \pi\pi$:

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

303 3. $D_s^- \rightarrow K^- \pi^+ \pi^-$

304 (a) $D^- \rightarrow \pi^- \pi^+ \pi^-$:

305 Possible with $\pi^- \rightarrow K^-$ misidentification, vetoed by requiring $m(K_\pi^- \pi^+ \pi^-) \neq$
 306 $m(D^-) \pm 40$ MeV or $\text{PIDK}(K^+) > 15$.

307 (b) $\Lambda_c^- \rightarrow \bar{p} \pi^+ \pi^-$:

308 Possible with $\bar{p} \rightarrow K^-$ misidentification, vetoed by requiring $m(K_p^- \pi^+ \pi^-) \neq$
 309 $m(\Lambda_c^-) \pm 40$ MeV or $\text{PIDK}(K^-) - \text{PIDp}(K^-) > 5$.

310 (c) $D^0 \rightarrow K\pi$:

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

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

- 317 1. $X_s^+ \rightarrow K^+\pi^+\pi^-$:
- 318 (a) $B_s^0 \rightarrow D_s\pi\pi\pi$:
 Possible with $\pi^+ \rightarrow K^+$ misidentification, suppressed with $\text{PIDK}(K^+) > 10$.
- 319 320 (b) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^+\pi^-)$:
 Outside of considered phase-space region, see Sec. 3.2.1.
- 321 322 (c) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+K^-\pi^+)$:
 To suppress $B_s^0 \rightarrow D_s^-K^+K^-\pi^+$ background, possible with $K^- \rightarrow \pi^-$ misiden-
 tification, we require $\text{PIDK}(\pi^-) < 0$. 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
 $\text{PIDK}(\pi^-) < -5$.
- 323 328 2. $X_d^+ \rightarrow \pi^+\pi^+\pi^-$:
- 329 (a) $B_s^0 \rightarrow D_sK\pi\pi$:
 Possible with single missID of $K^+ \rightarrow \pi^+$, suppressed with $\text{PIDK}(\pi^+) < 0$.
- 330 331 (b) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow \pi^+\pi^+\pi^-)$:
 Outside of considered phase-space region, see Sec. 3.2.1.
- 332 333 (c) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^-\pi^+)$:
 Possible with single missID of $K^+ \rightarrow \pi^+$, vetoed by requiring $m(\pi^+\pi_K^+\pi^-) \neq$
 $m(D_s) \pm 20$ MeV or $\text{PIDK}(\pi^+) < -5$ for both π^+ .

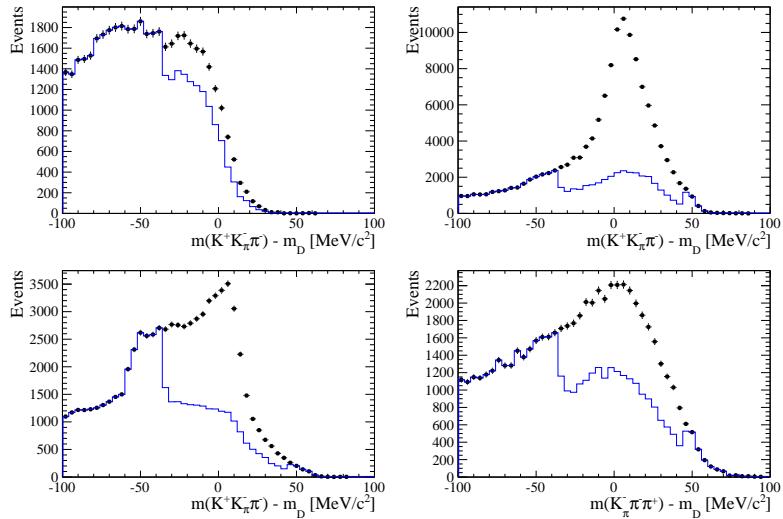


Figure 3.1: Background contributions from D^- decays where the π^- is misidentified as K^- . The D_s invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*(892)K$, $D_s \rightarrow KK\pi$ (non-resonant) and $D_s \rightarrow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

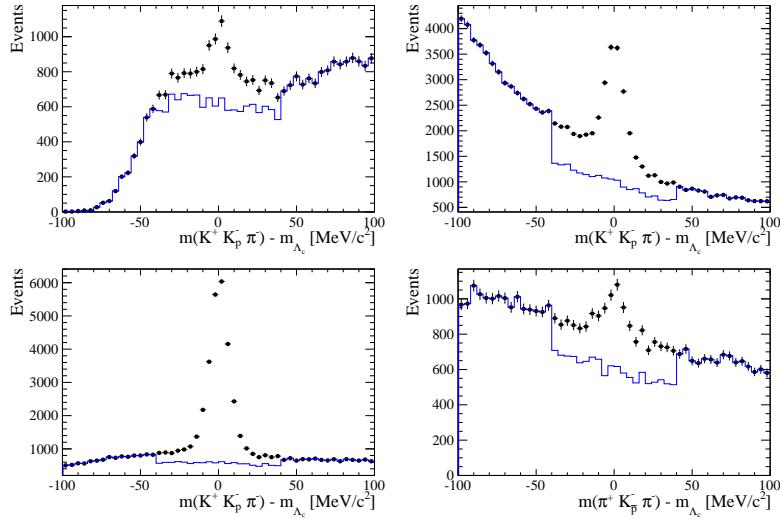


Figure 3.2: Background contributions from Λ_c decays where the \bar{p} is misidentified as K^- . The D_s invariant mass is recomputed applying the proton mass hypothesis to the kaon and shown for the $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*(892)K$, $D_s \rightarrow KK\pi$ (non-resonant) and $D_s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

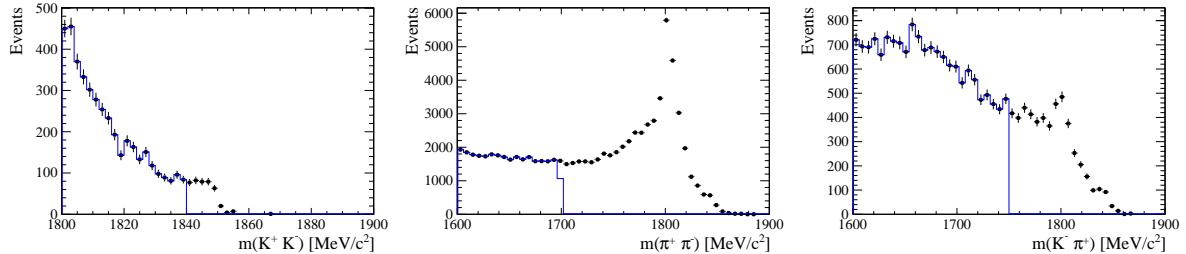


Figure 3.3: Background contributions to $D_s \rightarrow KK\pi$ (left), $D_s \rightarrow \pi\pi\pi$ (middle) and $D_s \rightarrow K\pi\pi$ (right) from $D^0 \rightarrow hh$ decays combined with a random pion.

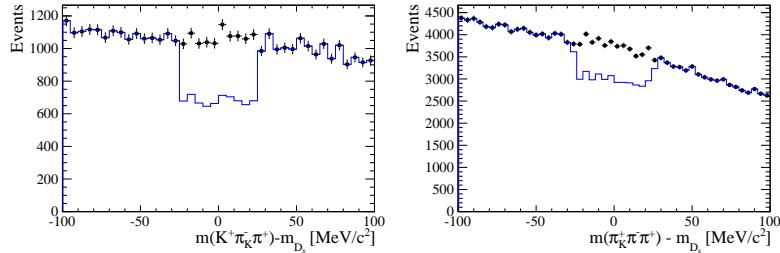


Figure 3.4: Background contributions to $B_s \rightarrow D_s K\pi\pi$ (left) and $B_s \rightarrow D_s \pi\pi\pi$ (right) from $B_s \rightarrow D_s D_s$ decays where the kaon is misidentified as pion. The $X_{s,d}$ invariant mass is recomputed applying the kaon mass hypothesis to the pion and shown without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

336 **3.2.3 Training of multivariate classifier**

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

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

350 The following discriminating variables are used for the BDTG training:

- 351 • logarithm of the B_s impact-parameter χ^2 , $B_s \log(\chi_{IP}^2)$
- 352 • logarithm of the cosine of the B_s direction angle, $\log(\text{DIRA})$
- 353 • fit quality of the DTF with PV constrain, χ_{DTF}^2/ndf
- 354 • logarithm of the minimal vertex quality difference for adding one extra track,
355 $\log(\Delta\chi_{add-track}^2)$
- 356 • the asymmetry between the transverse momentum of the B_s - candidate and the
357 transverse momentum of all the particles reconstructed with a cone of radius
358 $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2} < 1$ rad around the B_s - candidate, $B_s A_{pT}^{\text{cone}}$
- 359 • largest ghost probability of all tracks, $\max(\text{ghostProb})$
- 360 • logarithm of the the smallest X_s daughter impact-parameter χ^2 , $X_s \log(\min(\chi_{IP}^2))$
- 361 • largest distance of closest approach of the X_s daughters, $\max(\text{DOCA})$
- 362 • cosine of the largest opening angle between the D_s and another bachelor track h_i in
363 the plane transverse to the beam, $\cos(\max \theta_{D_s h_i})$
- 364 • logarithm of the the smallest D_s daughter impact-parameter χ^2 , $D_s \log(\min(\chi_{IP}^2))$
- 365 • logarithm of the D_s flight-distance significance, $D_s \log(\chi_{FD}^2)$
- 366 • logarithm of the D_s radial flight-distance, $D_s \log(RFD)$

367 Loose cuts on the variables χ_{DTF}^2/ndf , $\Delta\chi_{add-track}^2$ and $\cos(\max \theta_{D_s h_i})$ are applied prior
368 to the training which are expected to be 100% signal efficient. Figure 3.6 shows the
369 distributions of the input variables for signal and background. As shown in Appendix B,
370 these distributions differ between data-taking period and trigger category. In particular
371 variables depending on the B_s kinematics and the event multiplicity are affected (e.g.
372 $\theta_{D_s h_i}$ or A_{pT}^{cone}). The BDTG is consequently trained separately for these categories. The
373 resulting classifier response is shown in Fig. 3.7 for each category (even and odd test
374 samples combined) and in Appendix B for each training.

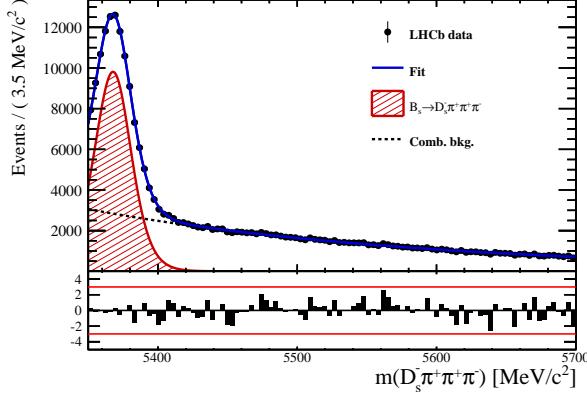


Figure 3.5: Reconstructed B_s mass for $B_s \rightarrow D_s \pi\pi\pi$ events that pass the preselection. The fitted PDF is shown in blue, the signal component in red and the background component in black.

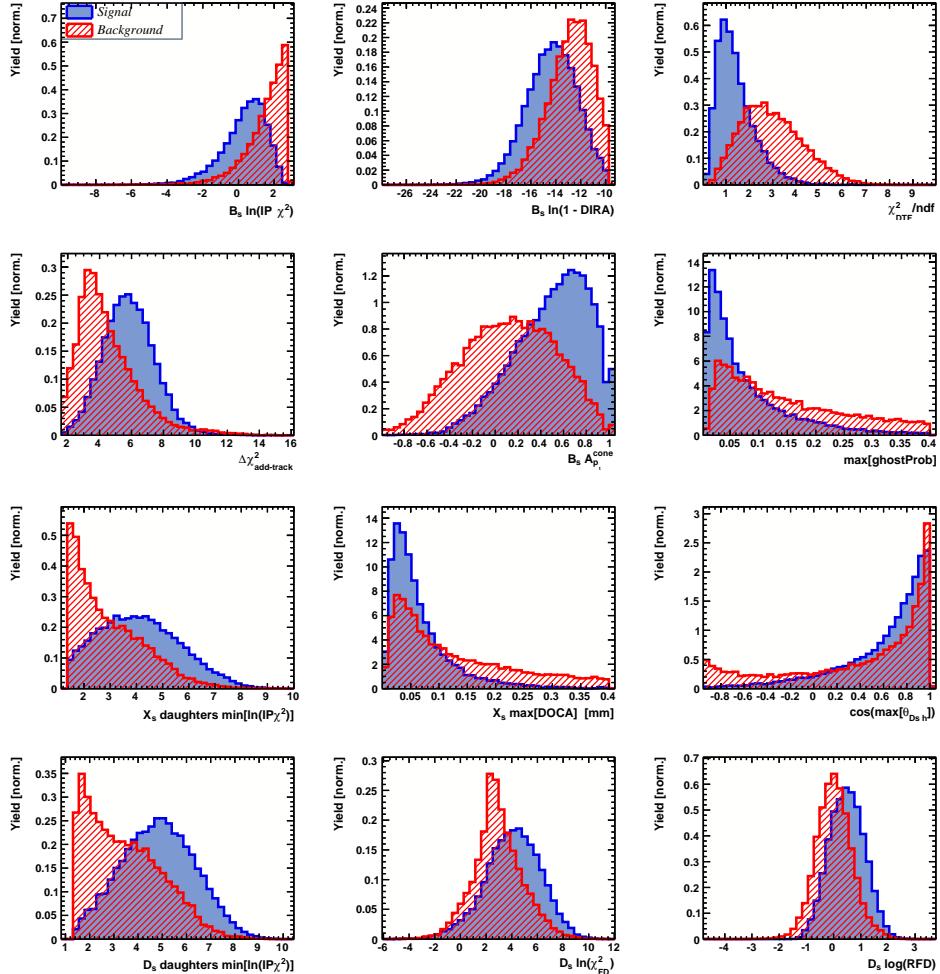


Figure 3.6: Discriminating variables used to train the BDTG for all data categories combined.

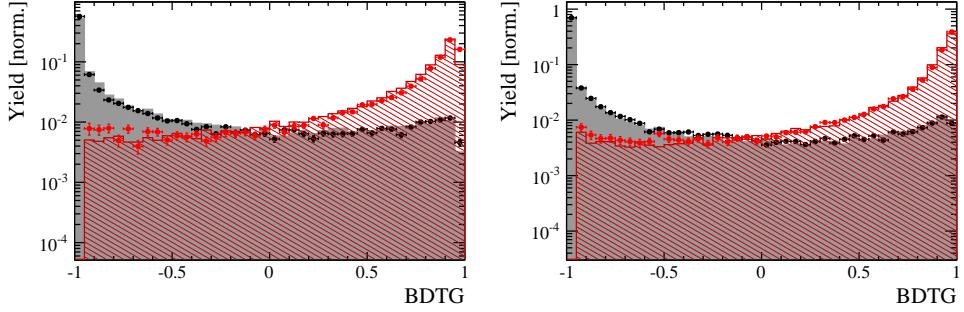


Figure 3.7: Signal (red) and background (black) distributions for the BDTG response for Run-I (left) and Run-II (right) data. Filled histograms (data points) show the BDTG response for the L0-TOS (L0-TIS) category. Even and odd test samples are combined.

3.2.4 Final selection

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

$$\text{FOM}(\text{BDTG}) = \frac{N_s(\text{BDTG})}{\sqrt{N_s(\text{BDTG}) + N_b(\text{BDTG})}} \quad (3.1)$$

where $N_s(\text{BDTG})$ is the $B_s \rightarrow D_s K\pi\pi$ signal yield for a given BDTG cut and $N_b(\text{BDTG})$ is the combinatorial background yield in the signal region ($m(D_s K\pi\pi) = m_{B_s} \pm 40 \text{ MeV}$). 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 $\text{BDTG} > 0$ 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:

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

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

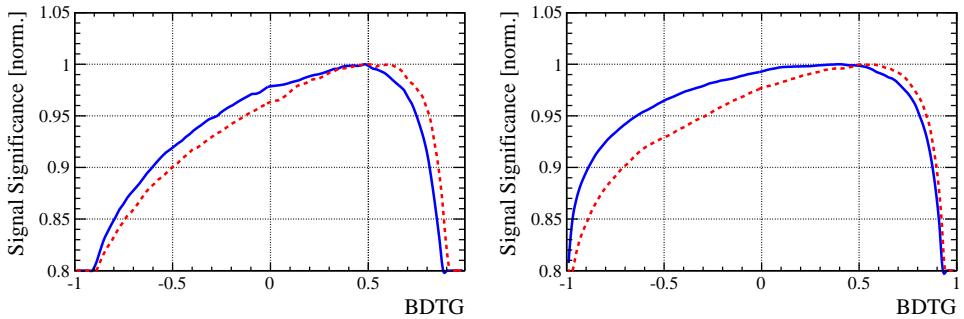


Figure 3.8: Signal significance as function of the applied cut on the BDTG response for Run-I (left) and Run-II (right) data. The scans for the L0-TOS (L0-TIS) category are shown in blue (red). The signal significance is normalized to be 1 at the optimal BDTG cut value.

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

	Description	Requirement
$D_s \rightarrow hh\bar{h}$	$m(hh\bar{h})$	$= m_{D_s} \pm 25$ MeV
$D_s^- \rightarrow KK\pi^-$	D^0 veto	$m(KK) < 1840$ MeV
$D_s^- \rightarrow \phi\pi^-$	$m(KK)$ PIDK(K^+) PIDK(K^-) PIDK(π^-) χ_{FD}^2 FD in z D^- veto Λ_c veto	$= m_\phi \pm 12$ MeV > -10 > -10 < 20 > 0 > -1 $m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 40$ MeV PIDK(K^-) > 5 $m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 40$ MeV PIDK(K^-) – PIDp(K^-) > 2
$D_s^- \rightarrow K^*(892)K^-$	$m(KK)$ $m(K^+\pi^-)$ PIDK(K^+) PIDK(K^-) PIDK(π^-) χ_{FD}^2 FD in z D^- veto Λ_c veto	$\neq m_\phi \pm 12$ MeV $= m_{K^*(892)} \pm 75$ MeV > -10 > -5 < 10 > 0 > 0 $m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 40$ MeV PIDK(K^-) > 15 $m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 40$ MeV PIDK(K^-) – PIDp(K^-) > 5
$D_s^- \rightarrow (KK\pi^-)_{NR}$	$m(KK)$ $m(K^+\pi^-)$ PIDK(K^+) PIDK(K^-) PIDK(π^-) χ_{FD}^2 FD in z D^- veto Λ_c veto	$\neq m_\phi \pm 12$ MeV $\neq m_{K^*(892)} \pm 75$ MeV > 5 > 5 < 10 > 4 > 0 $m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 40$ MeV PIDK(K^-) > 15 $m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 40$ MeV PIDK(K^-) – PIDp(K^-) > 5
$D_s \rightarrow \pi\pi\pi$	PIDK(π) PIDp(π) D^0 veto χ_{FD}^2 FD in z	< 10 < 20 $m(\pi^+\pi^-) < 1700$ MeV > 9 > 0
$D_s^- \rightarrow K^-\pi^+\pi^-$	PIDK(K) PIDK(π) PIDp(π) D^0 veto χ_{FD}^2 FD in z D^- veto Λ_c veto	> 8 < 5 < 20 $m(K^-\pi^+) < 1750$ MeV > 9 > 0 $m(K_\pi^-\pi^+\pi^-) \neq m(D^-) \pm 40$ MeV PIDK(K^-) > 15 $m(K_p^-\pi^+\pi^-) \neq m(\Lambda_c) \pm 40$ MeV PIDK(K^-) – PIDp(K^-) > 5

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

	Description	Requirement
$B_s \rightarrow D_s h\pi\pi$	$m(D_s h\pi\pi)$	$> 5200 \text{ MeV}$
	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	$> 0.4 \text{ ps}$
	δt	$< 0.15 \text{ ps}$
	Phasespace region	$m(h\pi\pi) < 1.95 \text{ GeV}$ $m(h\pi) < 1.2 \text{ GeV}$ $m(\pi\pi) < 1.2 \text{ GeV}$
Wrong PV veto		$\text{nPV} = 1 \parallel \min(\Delta\chi^2_{IP}) > 20$
BDTG		$> 0.45 \text{ [Run-I,L0-TOS]}$ $> 0.50 \text{ [Run-I,L0-TIS]}$ $> 0.35 \text{ [Run-II,L0-TOS]}$ $> 0.50 \text{ [Run-II,L0-TIS]}$
$X_s^+ \rightarrow K^+\pi^+\pi^-$	PIDK(K)	> 10
	PIDK(π^+)	< 10
	PIDK(π^-)	< 0
	D_s veto	$m(K^+\pi^+\pi_K^-) \neq m(D_s) \pm 20 \text{ MeV} \parallel \text{PIDK}(\pi^-) < -5$
$X_d^+ \rightarrow \pi^+\pi^+\pi^-$	PIDK(π^+)	< 0
	PIDK(π^-)	< 10
	D_s veto	$m(\pi^+\pi_K^+\pi^-) \neq m(D_s) \pm 20 \text{ MeV} \parallel \text{PIDK}(\pi^+) < -5$
All tracks	hasRich	$= 1$

386 3.3 Simulation

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

393

394

All samples indicated with 'Requested' have been requested in Dec. 17 and are ex-
 395 pected 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	
$B_s \rightarrow (D_s \rightarrow KK\pi)K\pi\pi$	13266007	08i	1.2 M	1.2 M	-	-	Generator Level
$B_s \rightarrow (D_s \rightarrow KK\pi)K\pi\pi$	13266008	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow K\pi\pi)K\pi\pi$	13266058	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow \pi\pi\pi)K\pi\pi$	13266038	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow KK\pi)\pi\pi\pi$	13266006	08i	1.2 M	1.2 M	-	-	Generator Level
$B_s \rightarrow (D_s \rightarrow KK\pi)\pi\pi\pi$	13266068	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow K\pi\pi)\pi\pi\pi$	13266088	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow \pi\pi\pi)\pi\pi\pi$	13266078	09c	Requested	Requested	Requested	Requested	Generator Level, Stripping
$B_s \rightarrow D_s^* \pi\pi\pi, D_s \rightarrow KK\pi$	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: $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*(892)\pi$, $D_s \rightarrow \pi\pi\pi$ and $D_s \rightarrow Kh\pi$ to account for different signal purities. The $D_s \rightarrow Kh\pi$ category combines the two D_s decay channels with the lowest statistics: $D_s \rightarrow KK\pi$ (non-resonant) and $D_s \rightarrow K\pi\pi$. This amounts to 16 categories in total.

4.1 Signal model

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

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

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

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

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

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

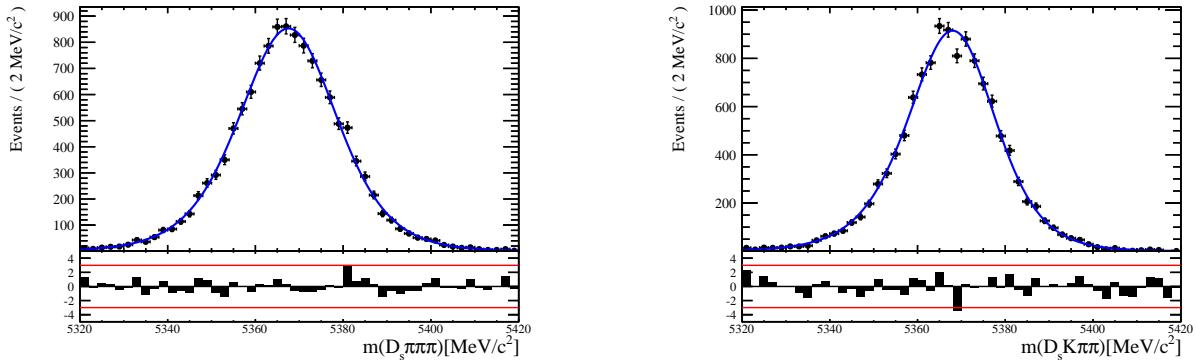


Figure 4.1: Invariant mass distributions of simulated (left) $B_s^0 \rightarrow D_s \pi\pi\pi$ and (right) $B_s^0 \rightarrow D_s K\pi\pi$ events. A fit with a Johnson's SU PDF is overlaid.

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

419 4.2 Background models

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

422 Combinatorial background

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

427 Peaking B_d background

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

431 Partially reconstructed background

433 Partially reconstructed $B_s^0 \rightarrow D_s^*\pi\pi\pi$ decays, with $D_s^* \rightarrow D_s\gamma$ or $D_s^* \rightarrow D_s\pi^0$, are expected
434 to be peaking lower than signal in the $m(D_s\pi\pi\pi)$ spectrum with large tails due to the
435 momentum carried away by the not reconstructed π^0 or γ . An empirical description for
436 the shape of this contribution is derived from a $B_s^0 \rightarrow D_s^*\pi\pi\pi$ MC sample subject to
437 the nominal $B_s^0 \rightarrow D_s\pi\pi\pi$ selection. Figure 4.2 (left) shows the respective reconstructed
438 $m(D_s\pi\pi\pi)$ distribution. A sum of three bifurcated Gaussian functions is used to describe
439 it. In the fit to data, all parameters are fixed to the ones obtained from MC except for
440 the parameter which describes the width of the right tail of the distribution to account for
441 data-simulation differences in mass resolution. The equivalent $B_s^0 \rightarrow D_s^*K\pi\pi$ component
442 contributing to the $B_s^0 \rightarrow D_sK\pi\pi$ data sample is described by the same PDF with the
443 right tail fixed to the $B_s^0 \rightarrow D_s\pi\pi\pi$ result.

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

446 Misidentified background

448 A small fraction of $B_s \rightarrow D_s^-\pi^+\pi^+\pi^-$ and $B_s^0 \rightarrow D_s^*\pi^+\pi^+\pi^-$ decays, where one of the
449 pions is misidentified as a kaon, contaminate the $B_s^0 \rightarrow D_sK^+\pi^+\pi^-$ sample. To determine
450 the corresponding background shapes, we use simulated events passing the nominal
451 selection except for the PID cuts on the bachelor π^+ tracks. The **PIDCalib** package
452 is used to determine the p_T, η -dependent $\pi^+ \rightarrow K^+$ misidentification probability for
453 each pion. We change the particle hypothesis from pion to kaon for the pion with the
454 higher misidentification probability and recompute the invariant B_s^0 mass, $m(D_s^-\pi_K^+\pi^+\pi^-)$.
455 Similarly, the invariant masses $m(\pi_K^+\pi^+\pi^-)$ and $m(\pi_K^+\pi^-)$ are recomputed and required
456 to be within the considered phasespace region. The background distributions are shown
457 in Fig. 4.2 (middle,right) and modeled by the sum of two Crystal Ball functions. The
458 expected yield of misidentified $B_s^0 \rightarrow D_s\pi\pi\pi$ ($B_s^0 \rightarrow D_s^*\pi^+\pi^+\pi^-$) candidates in the
459 $B_s^0 \rightarrow D_sK\pi\pi$ sample is computed by multiplying the fake rate (within the considered
460 B_s mass range) of 0.47% (0.61%) by the $B_s^0 \rightarrow D_s\pi\pi\pi$ ($B_s^0 \rightarrow D_s^*\pi^+\pi^+\pi^-$) yield as

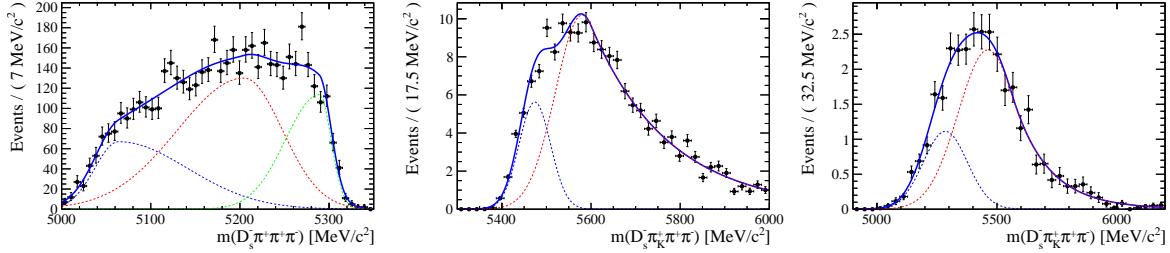


Figure 4.2: Left: Invariant mass distribution of simulated $B_s^0 \rightarrow D_s^* \pi\pi\pi$ events, where the γ/π^0 is excluded from the reconstruction. Middle: Invariant mass distribution of simulated $B_s^0 \rightarrow D_s \pi\pi\pi$ events, where one of the pions is reconstructed as a kaon taking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

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

465 We consider the $B_s^0 \rightarrow D_s K\pi\pi$ and $B_s^0 \rightarrow D_s^* K\pi\pi$ components contributing to the
 466 $B_s^0 \rightarrow D_s \pi\pi\pi$ data sample to be negligible due to the low branching fractions and the
 467 tight PID cuts on the bachelor pions.

468 4.3 Results

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

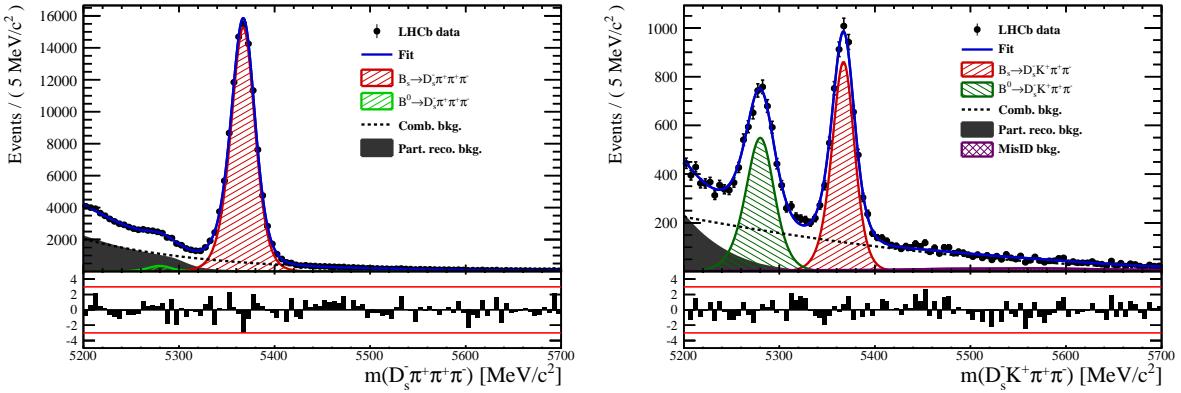


Figure 4.3: Invariant mass distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ (left) and $B_s^0 \rightarrow D_s K\pi\pi$ (right) candidates.

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

Component	Yield		
$B_s \rightarrow D_s\pi\pi\pi$	101289 ± 348		
$B^0 \rightarrow D_s\pi\pi\pi$	2318 ± 1763		
Partially reconstructed bkg.	29817 ± 530		
Combinatorial bkg.	52256 ± 603		
D_s final state	Signal yield		
$D_s^- \rightarrow \phi^0(1020)\pi^-$	34563 ± 197		
$D_s^- \rightarrow K^{*0}(892)K^-$	28472 ± 189		
$D_s^- \rightarrow (K^-h^+\pi^-)$	21047 ± 160		
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	17208 ± 145		

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

Component	Yield		
$B_s \rightarrow D_sK\pi\pi$	5125 ± 86		
$B^0 \rightarrow D_sK\pi\pi$	4190 ± 92		
Partially reconstructed bkg.	1707 ± 90		
Misidentified bkg.	683 ± 0		
Combinatorial bkg.	9686 ± 162		
D_s final state	Signal yield		
$D_s^- \rightarrow \phi^0(1020)\pi^-$	1613 ± 47		
$D_s^- \rightarrow K^{*0}(892)K^-$	1527 ± 46		
$D_s^- \rightarrow (K^-h^+\pi^-)$	1128 ± 40		
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	857 ± 37		

473 5 Decay-time Resolution

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

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

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

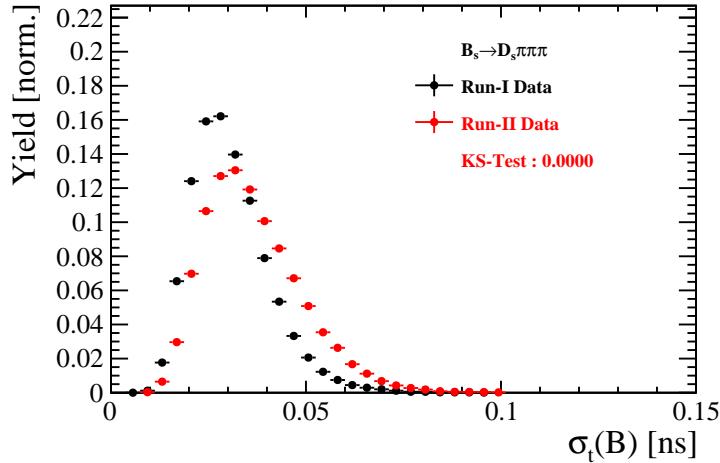


Figure 5.1: Distribution of the decay time error for $B_s^0 \rightarrow D_s\pi\pi\pi$ signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

496 5.1 Calibration for Run-I data

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

$$505 \quad \mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s^2 / 2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s^2 / 2}, \quad (5.1)$$

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

$$508 \quad \sigma_{eff} = \sqrt{(-2/\Delta m_s^2) \ln \mathcal{D}}, \quad (5.2)$$

508 which converts the resolution into a single-Gaussian function with an effective resolution
 509 that causes the same damping effect on the magnitude of the B_s oscillation. For the Run-I
 510 $B_s^0 \rightarrow D_s K \pi\pi$ MC sample the effective average resolution is found to be $\sigma_{eff} = 39.1 \pm 0.3$ fs.

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

$$519 \quad \sigma_{eff}^{MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.232 \pm 0.010) \sigma_t \quad (5.3)$$

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

$$521 \quad \sigma_{eff}^{D_s K, MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.201 \pm 0.013) \sigma_t. \quad (5.4)$$

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

$$523 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (10.26 \pm 1.52) \text{ fs} + (1.280 \pm 0.042) \sigma_t \quad (5.5)$$

523 can be used for our analysis. The following calibration functions were used in the
 524 $B_s^0 \rightarrow D_s K$ analysis to estimate the systematic uncertainty on the decay-time resolution:

$$525 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (-0.568 \pm 1.570) \text{ fs} + (1.243 \pm 0.044) \sigma_t \quad (5.6)$$

$$525 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.772 \pm 0.012) \sigma_t \quad (5.7)$$

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

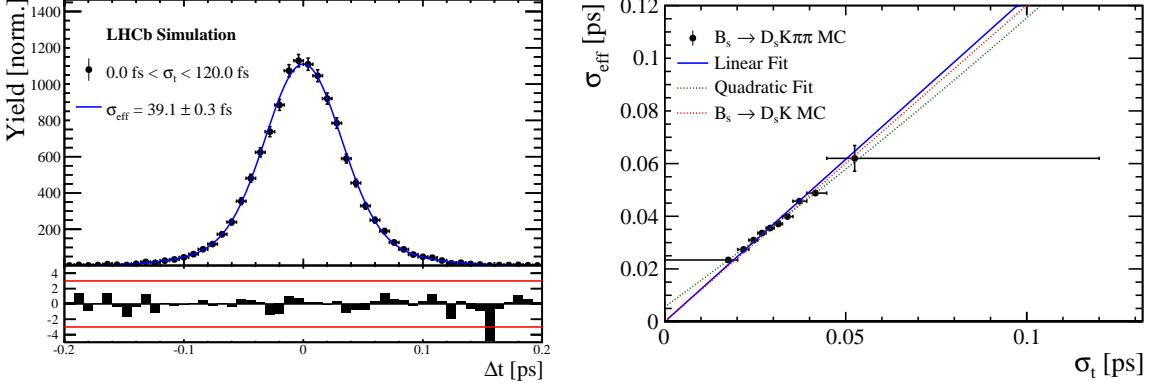


Figure 5.2: (Left) Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K\pi\pi$ MC candidates. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for $B_s \rightarrow D_s K\pi\pi$ MC (Run-I). The fitted calibration curve is shown in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

528 5.2 Calibration for Run-II data

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

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

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

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

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

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

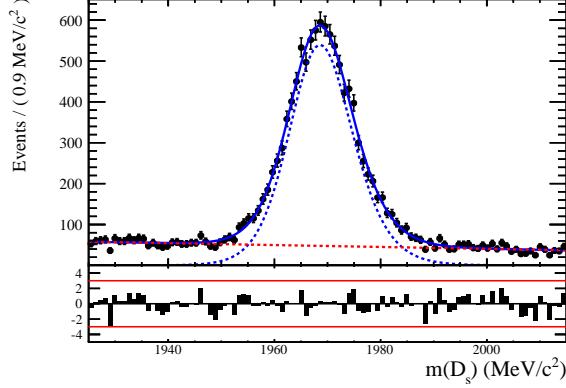


Figure 5.3: The invariant mass distribution for prompt D_s candidates.

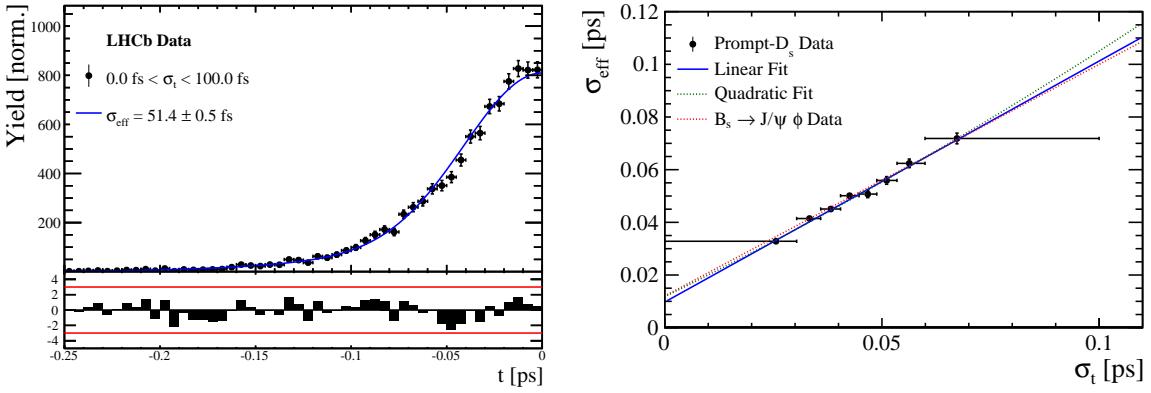


Figure 5.4: (Left) Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data). The 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
$B_s \rightarrow D_s K\pi\pi$	χ^2_{vtx}/ndof	< 8
	χ^2_{DTF}/ndof	< 15
	t	< 0 ps
$D_s \rightarrow hhh$	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	$nPV = 1 \parallel \min(\Delta\chi^2_{IP}) > 20$
$D_s^- \rightarrow KK\pi^-$	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
$D_s^- \rightarrow \phi\pi^-$	$m(KK)$	$= m_\phi \pm 20$ MeV
	PIDK(K^+)	> -10
	PIDK(K^-)	> -10
	PIDK(π^-)	< 20
$D_s^- \rightarrow K^*(892)K^-$	$m(KK)$	$\neq m_\phi \pm 20$ MeV
	$m(K^+\pi^-)$	$= m_{K^*(892)} \pm 75$ MeV
	PIDK(K^+)	> -10
	PIDK(K^-)	> -5
	PIDK(π^-)	< 20
$D_s^- \rightarrow (KK\pi^-)_{NR}$	$m(KK)$	$\neq m_\phi \pm 20$ MeV
	$m(K^+\pi^-)$	$\neq m_{K^*(892)} \pm 75$ MeV
	PIDK(K^+)	> 5
	PIDK(K^-)	> 5
	PIDK(π^-)	< 10
$D_s \rightarrow \pi\pi\pi$	PIDK(h)	< 10
	PIDp(h)	< 10
	D^0 veto	$m(\pi^+\pi^-) < 1700$ MeV
$X_s \rightarrow K\pi\pi$	$\chi^2_{IP}(h)$	< 40
	PIDK(K)	> 10
	PIDK(π)	< 5
	isMuon(h)	False
All tracks	p_T	> 500 MeV

556 6 Acceptance

557 6.1 MC corrections

558 6.1.1 Truth matching of simulated candidates

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

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

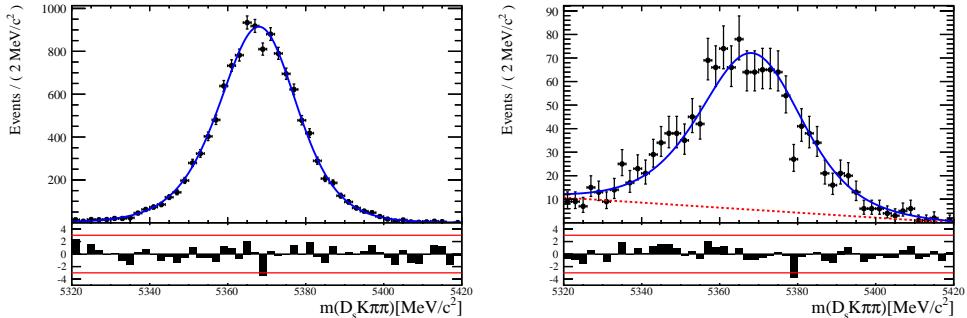


Figure 6.1: The reconstructed B_s mass distribution for simulated $B_s \rightarrow D_s K\pi\pi$ decays classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

577 6.1.2 Correction of data-simulation differences

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

586 6.2 Decay-time acceptance

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

592 We use a combination of control channels to derive the acceptance function $\epsilon(t)$,
 593 because for $B_s^0 \rightarrow D_s K\pi\pi$ decays the decay-time acceptance is strongly correlated with
 594 the CP -observables which we aim to measure. Therefore, extracting the CP -observables
 595 and the acceptance shape at the same time is not possible. A fit to the decay-time
 596 distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates is performed and the obtained acceptance shape
 597 is corrected for the small difference observed between the $B_s^0 \rightarrow D_s K\pi\pi$ and $B_s^0 \rightarrow D_s \pi\pi\pi$
 598 MC samples. In addition, we include the control channel $B^0 \rightarrow D_s K\pi\pi$ to increase
 599 the statistical precision. A simultaneous fit to the four datasets ($B_s^0 \rightarrow D_s \pi\pi\pi$ data,
 600 $B^0 \rightarrow D_s K\pi\pi$ data, $B_s^0 \rightarrow D_s K\pi\pi$ MC and $B_s^0 \rightarrow D_s \pi\pi\pi$ MC) is performed to allow for
 601 a straightforward propagation of uncertainties. In each case, a PDF of the following form

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

602 is used to describe the decay-time distribution. For real data, the values for $\Gamma_{s,d}$ and
 603 $\Delta\Gamma_{s,d}$ are fixed to the latest HFAG results [38], while for simulated data, the generated
 604 values are used. A single Gaussian resolution function $\mathcal{R}(t - t', \delta t)$ is used where the
 605 decay-time error estimate is scaled with the respective calibration functions determined in
 606 Sec. 5. Each decay-time acceptance $\epsilon(t)$ is modeled using cubic splines, allowing for the
 607 analytical computation of the decay-time integrals appearing in the PDF [39]. The splines
 608 are parametrized by so-called knots (t_0, t_1, \dots, t_N) which determine their boundaries. Two
 609 knots are located by default at the lower and upper edge of the interval allowed for the
 610 decay time, the remaining ones are chosen such that there is an approximately equal
 611 amount of data in-between two consecutive knots. In the basis of cubic b-splines, $b_i(t)$,
 612 the acceptance is then constructed as:

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

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

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

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

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

- 619 • $B_s^0 \rightarrow D_s K\pi\pi$ MC: $\epsilon_{D_s K\pi\pi}^{MC}(t)$
620 • $B_s^0 \rightarrow D_s \pi\pi\pi$ MC: $\epsilon_{D_s \pi\pi\pi}^{MC}(t) = R(t) \cdot \epsilon_{D_s K\pi\pi}^{MC}(t)$
621 • $B_s^0 \rightarrow D_s \pi\pi\pi$ data: $\epsilon_{D_s \pi\pi\pi}^{Data}(t) = R(t) \cdot \epsilon_{D_s K\pi\pi}^{Data}(t)$
622 • $B^0 \rightarrow D_s K\pi\pi$ data: $\epsilon_{D_s K\pi\pi}^{Data}(t)$

623 where $\epsilon_{D_s K\pi\pi}^{MC}(t)$ represents the acceptance in $B_s^0 \rightarrow D_s K\pi\pi$ MC, $R(t)$ represents the
624 ratio of acceptances in $B_s^0 \rightarrow D_s \pi\pi\pi$ and $B_s^0 \rightarrow D_s K\pi\pi$ MC and the final acceptance in
625 $B_s^0 \rightarrow D_s K\pi\pi$ data is represented by $\epsilon_{D_s K\pi\pi}^{Data}(t)$.

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

629

630

631 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
 $B^0 \rightarrow D_s K\pi\pi$ data sample in order to remove the MC dependency. The final strat-
egy 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	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ 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	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ 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	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ 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	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ 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)

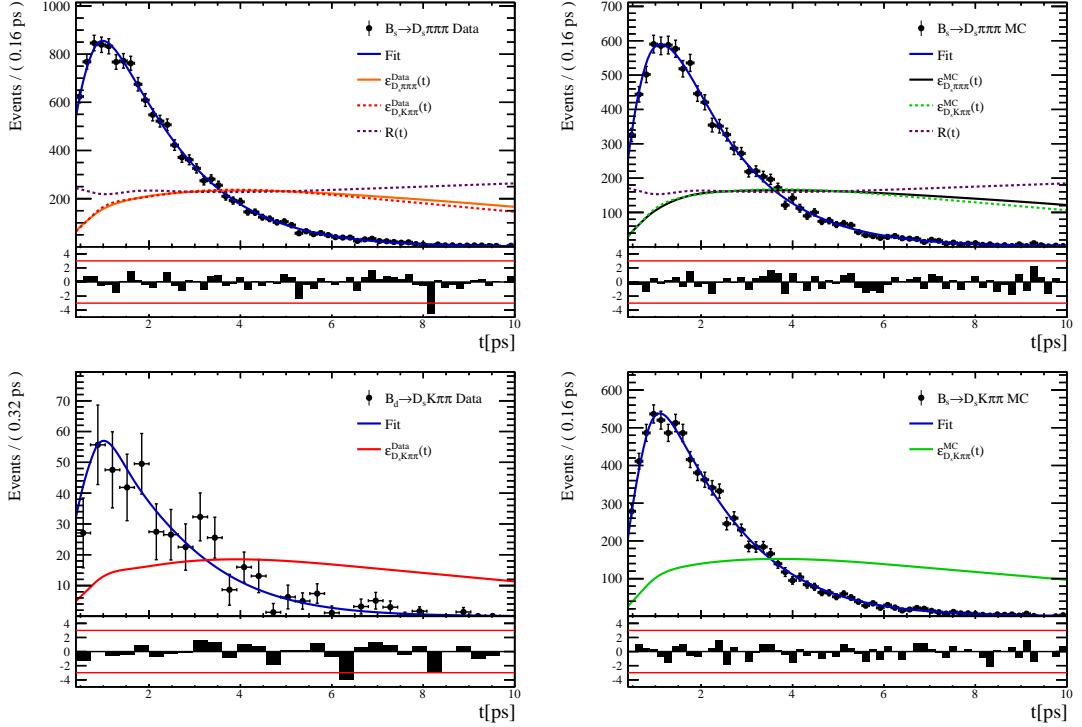


Figure 6.2: Decay-time fit projections for $B_s^0 \rightarrow D_s\pi\pi\pi$ data (top-left), $B_s^0 \rightarrow D_s\pi\pi\pi$ MC (top-right), $B_s^0 \rightarrow D_sK\pi\pi$ data (bottom-left) and $B_s^0 \rightarrow D_sK\pi\pi$ MC (bottom-right) in category [Run-I,L0-TOS]. The respective acceptance function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

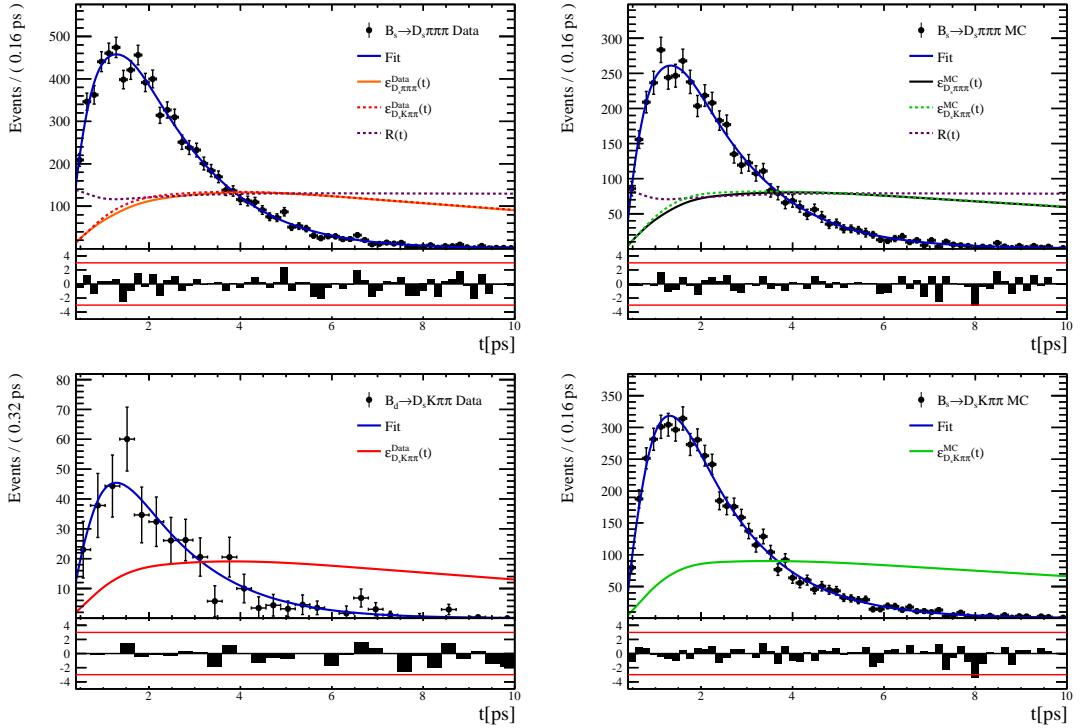


Figure 6.3: Decay-time fit projections for $B_s^0 \rightarrow D_s\pi\pi\pi$ data (top-left), $B_s^0 \rightarrow D_s\pi\pi\pi$ MC (top-right), $B_s^0 \rightarrow D_sK\pi\pi$ data (bottom-left) and $B_s^0 \rightarrow D_sK\pi\pi$ MC (bottom-right) in category [Run-I,L0-TIS]. The respective acceptance function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

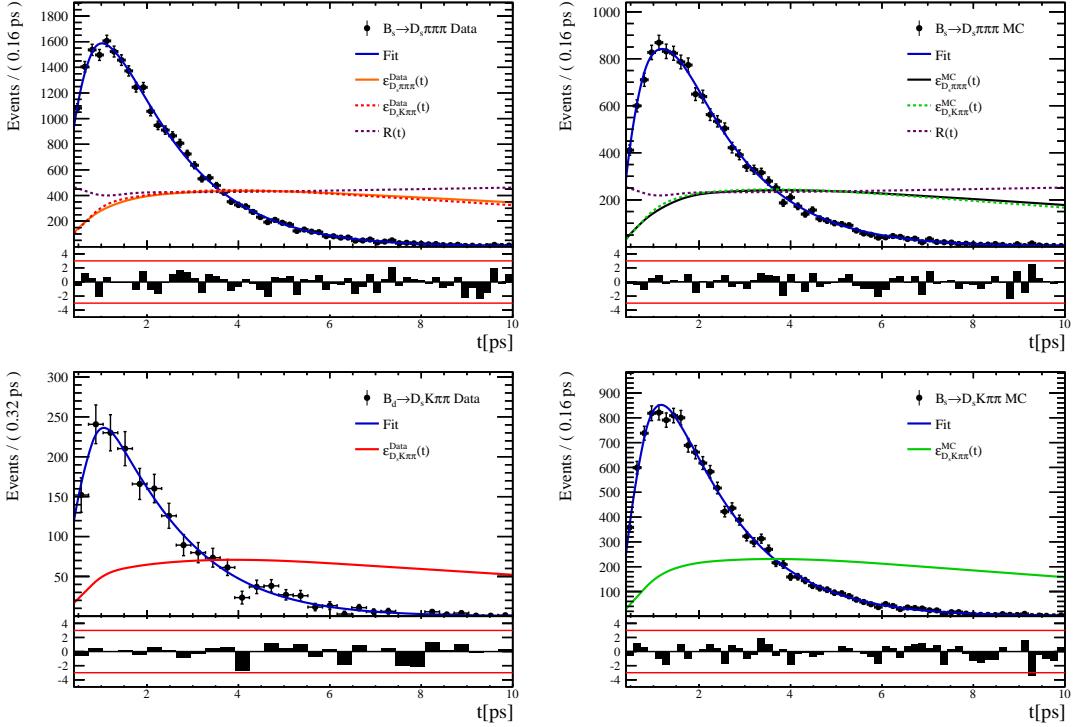


Figure 6.4: Decay-time fit projections for $B_s^0 \rightarrow D_s \pi\pi\pi$ data (top-left), $B_s^0 \rightarrow D_s \pi\pi\pi$ MC (top-right), $B^0 \rightarrow D_s K\pi\pi$ data (bottom-left) and $B_s^0 \rightarrow D_s K\pi\pi$ MC (bottom-right) in category [Run-II,L0-TOS]. The respective acceptance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

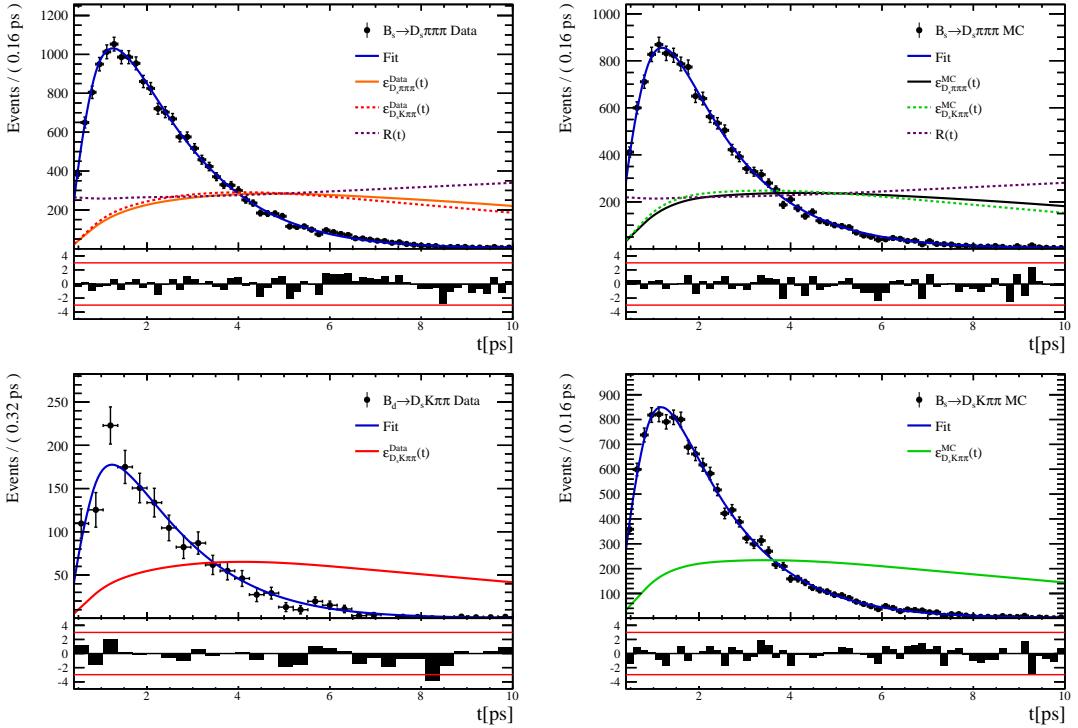


Figure 6.5: Decay-time fit projections for $B_s^0 \rightarrow D_s \pi\pi\pi$ data (top-left), $B_s^0 \rightarrow D_s \pi\pi\pi$ MC (top-right), $B^0 \rightarrow D_s K\pi\pi$ data (bottom-left) and $B_s^0 \rightarrow D_s K\pi\pi$ MC (bottom-right) in category [Run-II,L0-TIS]. The respective acceptance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

632 6.3 Phase space acceptance

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

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

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

$$\int |\mathcal{A}_f^c(\mathbf{x})|^2 \epsilon(\mathbf{x}) d\Phi_4 \approx \frac{1}{N_{MC}} \sum_k^{N_{MC}} \frac{|\mathcal{A}_f^c(\mathbf{x}_k)|^2}{|A'(\mathbf{x}_k)|^2} \quad (6.5)$$

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

646

647

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.

648

649 We use a BDTG to map the five-dimensional phase space to an one-dimensional distribution
 650 [40]. The BDTG is trained to learn the differences between the selected MC and a generator
 651 level MC sample. As discriminating variables, five invariant mass combinations are used
 652 as shown in Fig. 6.6. Based on the classifier output distributions, shown in Fig. 6.7, an
 653 efficiency as function of the BDTG response is derived.

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

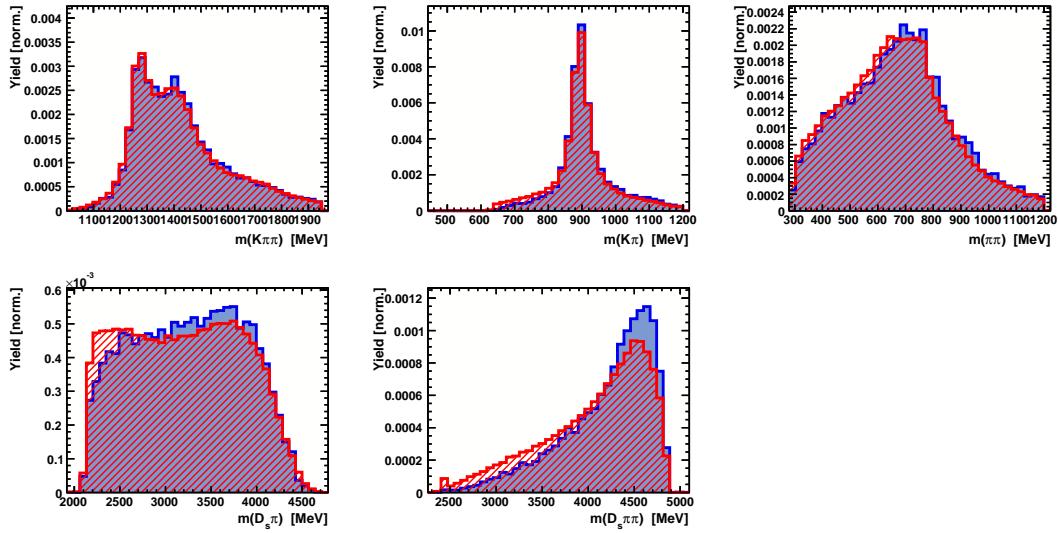


Figure 6.6: Discriminating variables used to train the BDTG. The selected MC sample is shown in blue and the generator MC sample in red.

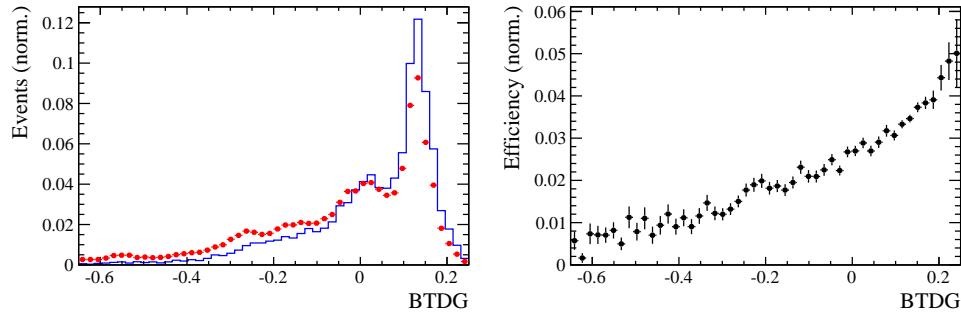


Figure 6.7: Left: Output distributions of the BDTG for the simulated MC sample (blue) and the generator level sample (red). Right: Phase space efficiency as function of the BDTG response computed as the ratio of selected and generated decays.

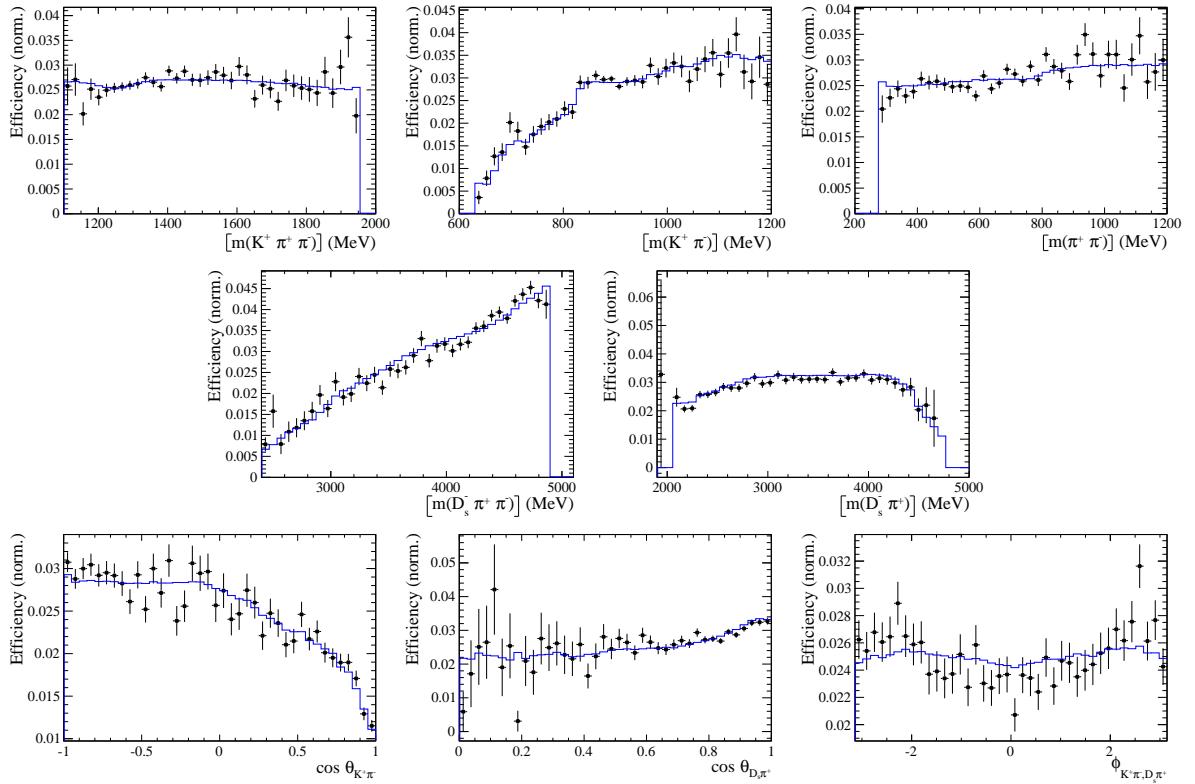


Figure 6.8: Efficiency variation as a function of the phase-space variables obtained from the ratio of selected and generated MC events (data points) and efficiency obtained from a reweighted toy MC sample (blue).

661 7 Flavour Tagging

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

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

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

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

680 For each B_s^0/\bar{B}_s^0 candidate, the tagging algorithms provide, besides a flavour tag
 681 $q = 1, -1, 0$ (for an initial B_s^0 , \bar{B}_s^0 or no tag), a prediction for the mistag probability η
 682 based on the output of multivariate classifiers. These are trained on simulated samples
 683 of flavour specific control channels ($B_s^0 \rightarrow D_s^- \pi^+$ (SS algorithm) and $B^+ \rightarrow J/\psi K^+$ (OS
 684 algorithms)) and are optimized for highest ϵ_{eff} on data. Utilizing flavour-specific final
 685 states, the estimated mistag η of each tagger has to be calibrated to match the actual
 686 mistag probability ω . For the calibration, a linear model

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

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

692 7.1 OS tagger combination

693 First, the OS electron, muon, neural net kaon and the secondary vertex charge taggers
 694 are individually calibrated and then combined into a single OS-Combo tagger using the
 695 `EspressoPerformanceMonitor` tool. We choose the flavour specific decay $B_s \rightarrow D_s \pi \pi \pi$ as
 696 calibration mode since it is very similar to the signal decay $B_s \rightarrow D_s K \pi \pi$. The calibration
 697 is performed separately for Run-I and Run-II data. Where available the latest Run-II
 698 tuning is used for Run-II data, otherwise the Run-I tuning of the taggers is used. Figures
 699 7.1 and 7.2 show the fitted calibration functions and Tables 7.1 and 7.2 list the measured
 700 tagging performances.

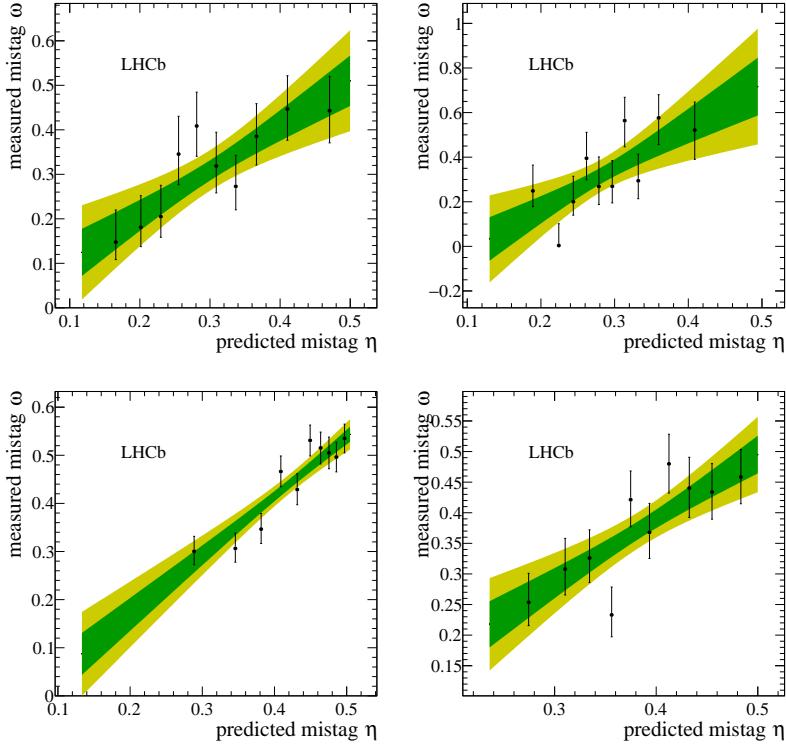


Figure 7.1: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run-I. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

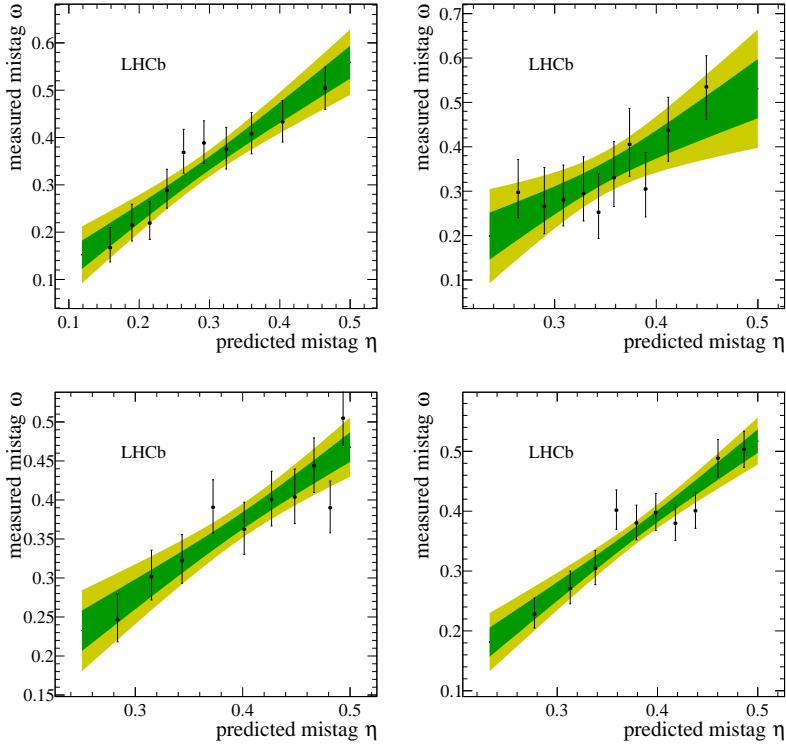


Figure 7.2: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run-II. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

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 $B_s \rightarrow D_s \pi\pi\pi$ decay-time distribution as discussed in Sec. 9. The predicted mistag probabilities η_{OS} and η_{SS} , shown Fig. 7.3 for $B_s \rightarrow D_s \pi\pi\pi$ and $B_s \rightarrow D_s K\pi\pi$ 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.

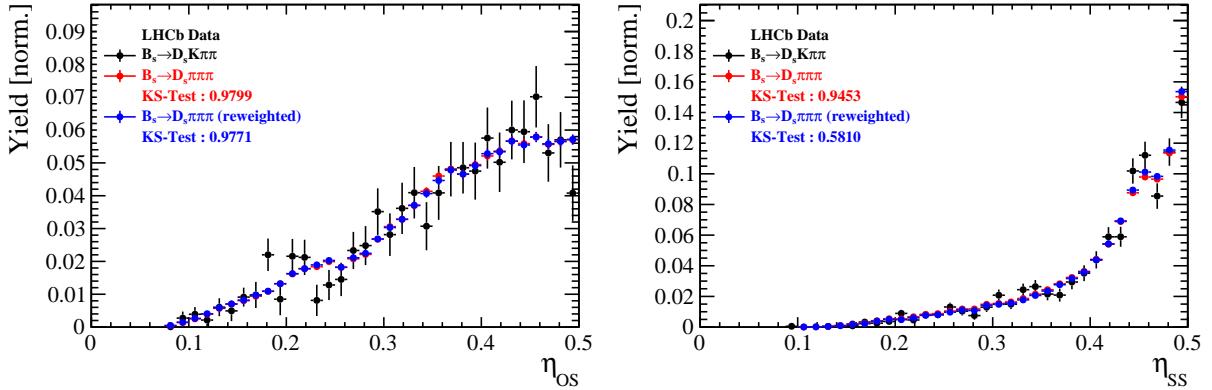


Figure 7.3: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \rightarrow D_s K\pi\pi$ (black) and $B_s^0 \rightarrow D_s \pi\pi\pi$ (red) data samples.

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.

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
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.

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
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:

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

where σ are the corresponding production cross-section. This asymmetry was measured by LHCb in pp collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV by means of a time-dependent analysis of $B_s \rightarrow D_s^- \pi^+$ 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 $B_s \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s K\pi\pi$ 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:

$$A_p(B_s^0)_{2011} = (-0.506 \pm 1.90)\% \quad (8.2)$$

$$A_p(B_s^0)_{2012} = (-0.164 \pm 1.30)\% \quad (8.3)$$

$$A_p(B_s^0)_{\text{Run-I}} = (-0.045 \pm 1.04)\%. \quad (8.4)$$

As for Run-II data no measurement is available yet, we determine the production asymmetry from $B_s \rightarrow D_s \pi\pi\pi$ 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	$A_p(B_s^0)_{\sqrt{s}=7 \text{ TeV}}$	$A_p(B_s^0)_{\sqrt{s}=8 \text{ TeV}}$
(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)	$-0.0586 \pm 0.0648 \pm 0.0150$	$0.0537 \pm 0.0447 \pm 0.0124$

724 8.2 $K^-\pi^+$ Detection Asymmetry

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

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

733 is used.

734 This asymmetry can be measured from the difference in asymmetries in the $D^+ \rightarrow$
 735 $K^-\pi^+\pi^+$ and $D^+ \rightarrow K_s^0\pi^+$ modes [45]:

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

736 where possible CP violation in the $D^+ \rightarrow K_s^0\pi^+$ mode is predicted to be smaller than
 737 10^{-4} in the Standard Model [46]. The asymmetry in the neutral kaon system, $A(K^0)$, has
 738 to be taken into account as a correction.

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

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

750 After the samples are weighted, fits are performed to the invariant
 751 $m(K^-\pi^+\pi^+)/m(K^+\pi^-\pi^-)$ and $m(K_s^0\pi^+)/m(K_s^0\pi^-)$ distributions to determine
 752 $A^{det}(K^-\pi^+)$. The PDFs used to describe the invariant mass distributions consist of
 753 gaussian functions for the signal component and exponentials describing the residual
 754 background.

755 The detection asymmetry is determined separately for every year and (since it is a
 756 charge asymmetry effect) magnet polarity. Serving as an example for Run-I and Run-
 757 II, the fits used to determine $N(D^+ \rightarrow K^-\pi^+\pi^+)/N(D^- \rightarrow K^+\pi^-\pi^-)$ and $N(D^+ \rightarrow$
 758 $K_s^0\pi^+)/N(D^- \rightarrow K_s^0\pi^-)$ for 2011, magnet up data and 2015, magnet up data are shown
 759 in Fig. 8.1 and 8.2 respectively. The obtained values of $A^{det}(K^-\pi^+) + A(K^0)$ for all years
 760 and polarities are shown in Table 8.2.

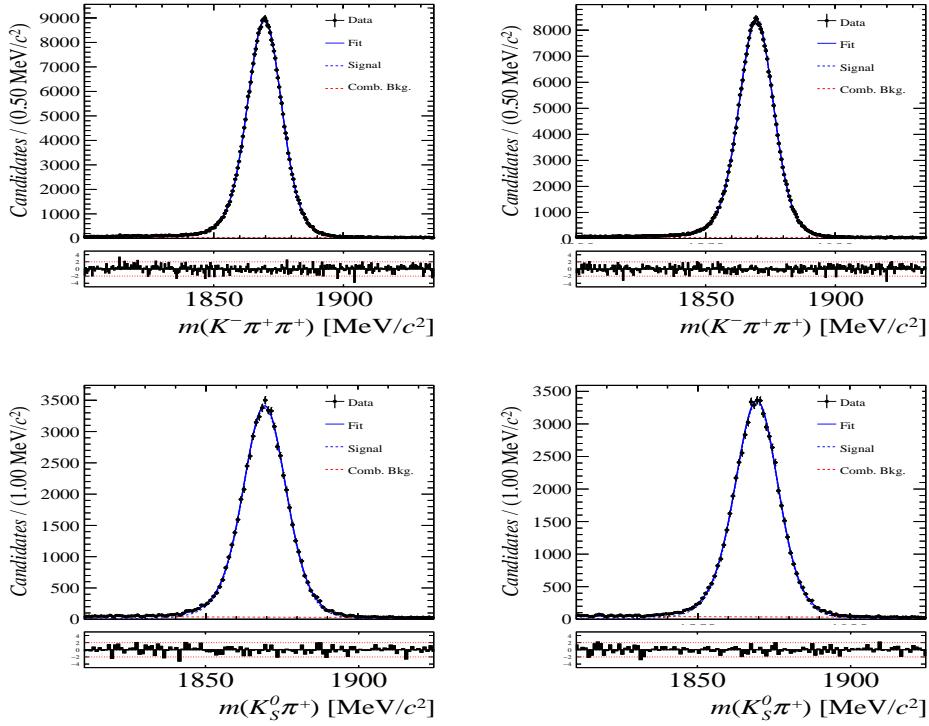


Figure 8.1: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_S^0 \pi^\pm$ candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

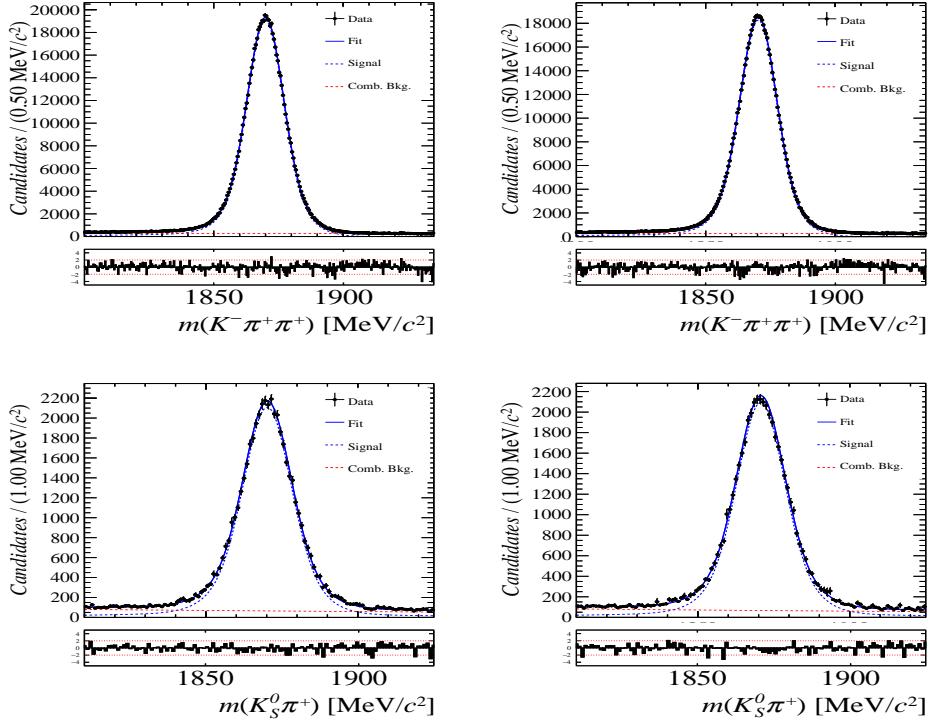


Figure 8.2: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_S^0 \pi^\pm$ candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

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 $B_s^0 \rightarrow D_s h\pi\pi$ 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 ($i = \text{OS,SS}$) and the decay-time error δt :

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

where $p(t|q_{\text{OS}}, \eta_{\text{OS}}, q_{\text{SS}}, \eta_{\text{SS}})$ 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 $B_s^0 \rightarrow D_s\pi\pi\pi$ data

Since the decay $B_s^0 \rightarrow D_s\pi\pi\pi$ 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 $B_s \rightarrow D_s\pi\pi\pi$ decay-time distribution. The uncertainties are statistical and systematic, respectively.

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

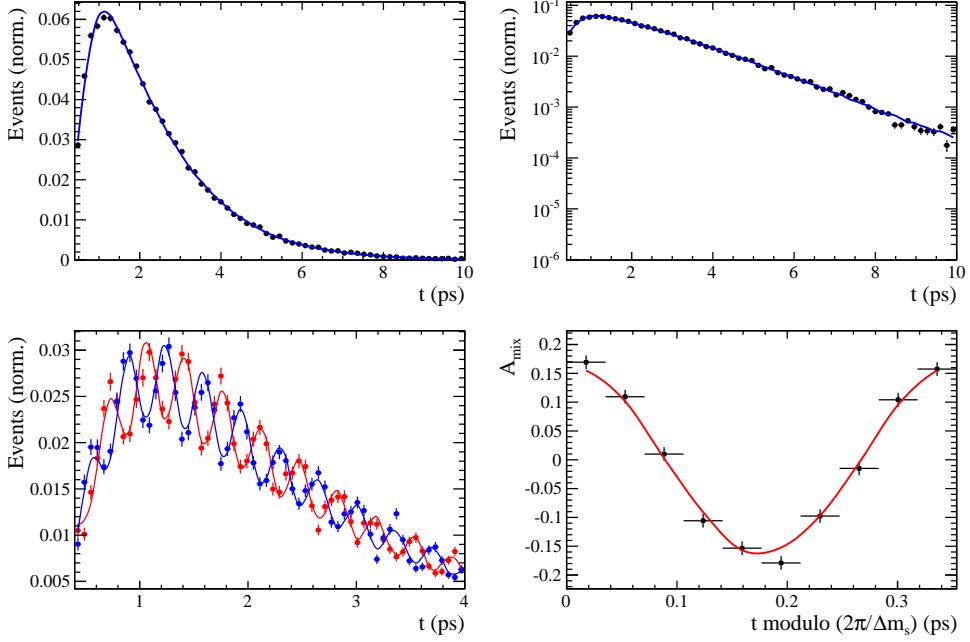


Figure 9.1: Top: Flavour averaged decay time distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates. Bottom-left: Tagged decay time distribution of mixed (red) and unmixed (blue) signal candidates. Bottom-right: Time-dependent asymmetry A_{mix} between mixed and unmixed B_s^0 candidates folded into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

780 9.2 Fit to $B_s^0 \rightarrow D_s K\pi\pi$ data

781 The measured CP coefficients $C, D_f, D_{\bar{f}}, S_f$ and $S_{\bar{f}}$ extracted from a fit to the
 782 $B_s \rightarrow D_s K\pi\pi$ decay-time distribution are reported in Table 9.2. The fit projection is
 783 shown in Fig. 9.2. We included Gaussian-constraints for the tagging calibration parameters
 784 with the central values and uncertainties determined in Sec. 9.1.

785

786

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

787

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

788

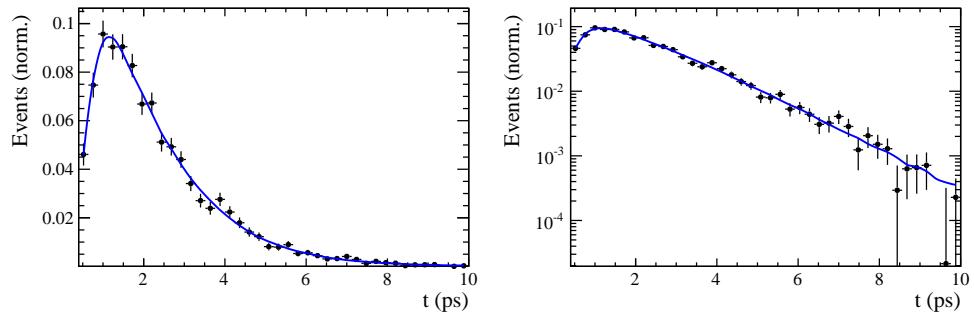


Figure 9.2: Decay-time distribution of $B_s^0 \rightarrow D_s K\pi\pi$ signal candidates with the fit projection overlaid.

Table 9.2: CP coefficients determined from a fit to the $B_s \rightarrow D_s K\pi\pi$ 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

789 10 Time-dependent amplitude fit

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

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

791 where $p(\mathbf{x}, t | q_{OS}, \eta_{OS}, q_{SS}, \eta_{SS})$ is given the differential decay rate in Equation 2.29 taking
 792 the tagging dilution into account. The phase space efficiency $\epsilon(\mathbf{x})$ is only included in the
 793 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
 794 of the amplitude components is described in the following Section. The remaining fitting
 795 strategy is exactly the same as for the decay-time fits, see Sec. 9.

796 10.1 Signal Model Construction

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

803 First, we fit the time-integrated and flavour averaged phase-space distribution of
 804 $B_s \rightarrow D_s K\pi\pi$ decays. In this case, a single total amplitude can be used:

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

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

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

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

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

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

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

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

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

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 $b \rightarrow c$ and $b \rightarrow u$ transitions and the likelihood is extended as follows:

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

Figure 10.1(right) shows a plot of the complexity factor λ , against the resulting BIC values. The final set of $b \rightarrow c$ and $b \rightarrow u$ 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 [%]
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	8.56 ± 1.43
$B_s \rightarrow K(1)(1400)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	43.72 ± 2.80
$B_s \rightarrow K(1460)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	3.25 ± 0.69
$B_s \rightarrow K^*(1410)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	15.33 ± 1.13
$B_s \rightarrow (D_s^- \pi^+)_P K^*(892)^0 (\rightarrow K^+ \pi^-)$	4.63 ± 0.69
$B_s \rightarrow K^*(1410)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	5.58 ± 0.62
$B_s \rightarrow (D_s^- K^+)_P \rho(770)^0 (\rightarrow \pi^+ \pi^-)$	1.49 ± 0.40
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K(0)^*(1430)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	4.72 ± 0.54
$B_s \rightarrow K(1)(1270)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	14.20 ± 1.56
Sum	101.47 ± 3.86

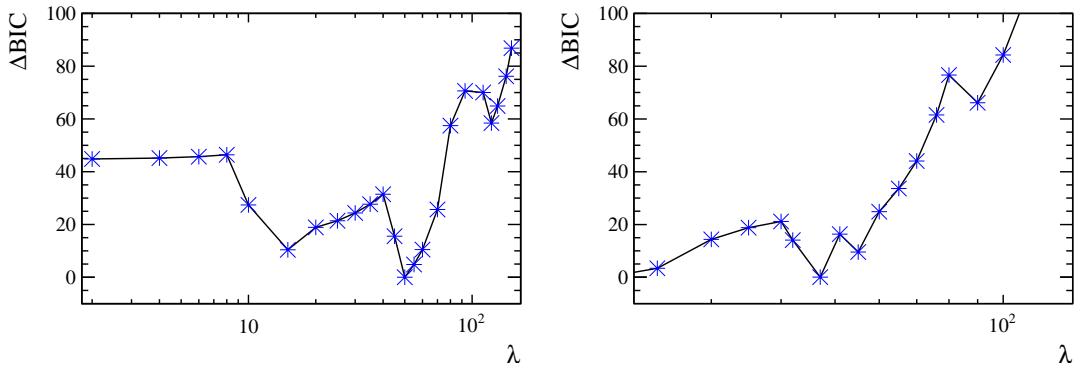


Figure 10.1: Difference in the BIC value from its minimum as function of the LASSO parameter λ for step 1 (left) and step 2 (right) of the model selection.

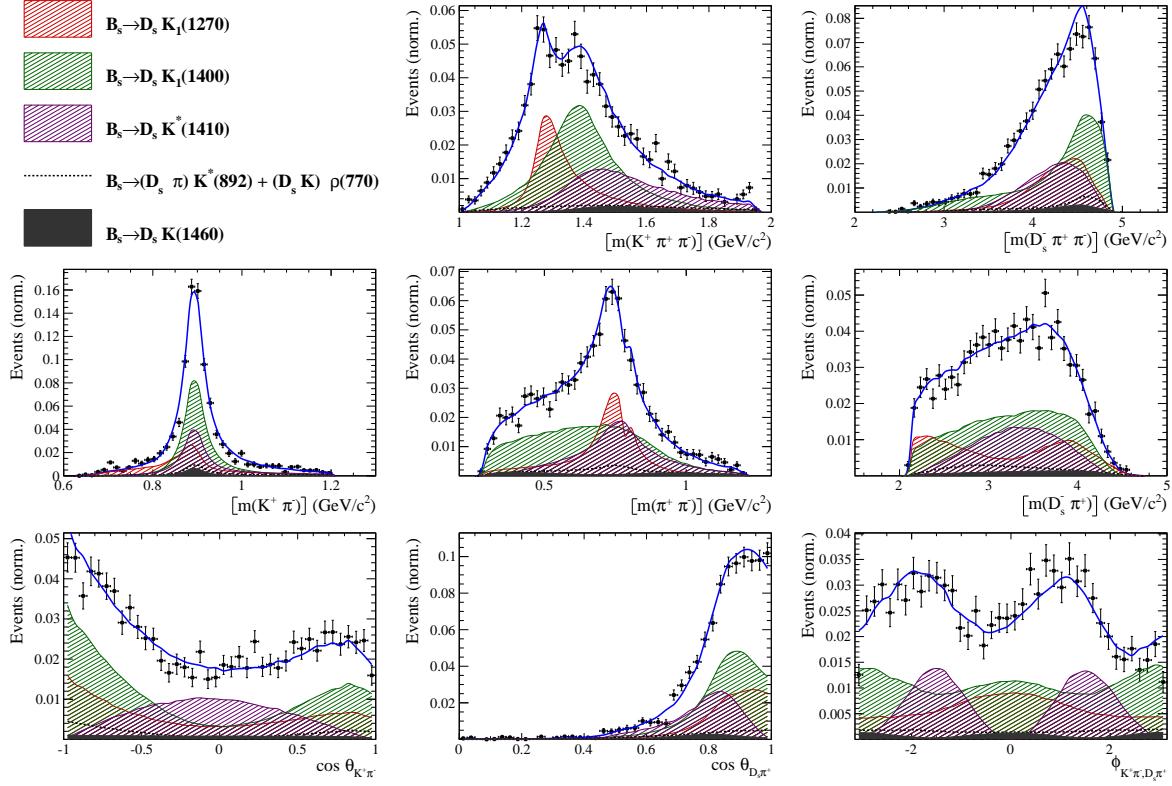


Figure 10.2: Projections of the fit result to the time-integrated and flavour averaged phase-space distribution of $B_s \rightarrow D_s K \pi \pi$ decays.

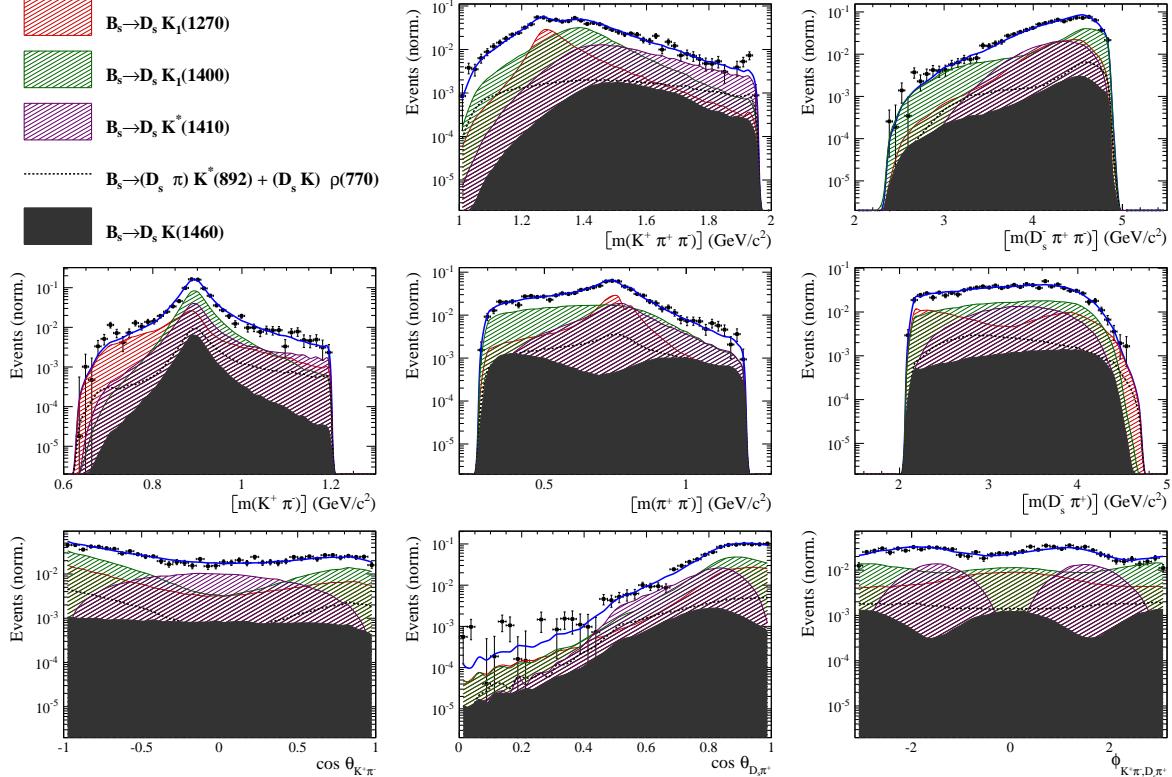


Figure 10.3: Projections of the fit result to the time-integrated and flavour averaged phase-space distribution of $B_s \rightarrow D_s K \pi \pi$ decays in logarithmic scale.

827 10.2 Results

828 Table 10.2 lists the modulus and phases of the complex amplitude coefficients a_i^c and a_i^u ,
 829 obtained by fitting the LASSO model to the data. The corresponding fit fractions for the
 830 $b \rightarrow c$ and $b \rightarrow u$ amplitudes are individually normalized

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

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

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

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

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

844 where N_b is the number of data events in a given bin, N_b^{exp} is the event count predicted
 845 by the fitted PDF and N_{bins} is the number of bins. An adaptive binning is used to ensure
 846 sufficient statistics in each bin for a robust χ^2 calculation [51]. At least 25 events per
 847 bin are required. The number of degrees of freedom ν , in an unbinned fit is bounded by
 848 $N_{\text{bins}} - 1$ and $(N_{\text{bins}} - 1) - N_{\text{par}}$, where N_{par} is the number of free fit parameters. We use
 849 the χ^2 value divided by $\nu = (N_{\text{bins}} - 1) - N_{\text{par}}$ as a conservative estimate. For the LASSO
 850 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 $b \rightarrow c$ and $b \rightarrow u$ 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	$A_{b \rightarrow c}$		$A_{b \rightarrow u}$	
	$ a_i $	$\arg(a_i)[^\circ]$	$ a_i $	$\arg(a_i)[^\circ]$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	1.0	0.0	1.0	0.0
$K_1(1270) \rightarrow K^*(892) \pi$	$0.76 \pm 0.11 \pm 0.16$	$60.9 \pm 9.6 \pm 14.1$		
$K_1(1270) \rightarrow K_0^*(1430) \pi$	$0.68 \pm 0.06 \pm 0.34$	$116.5 \pm 5.1 \pm 43.5$		
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$2.53 \pm 0.27 \pm 0.57$	$12.9 \pm 7.4 \pm 8.2$	$0.67 \pm 0.20 \pm 0.51$	$-76.3 \pm 16.9 \pm 22.9$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$1.28 \pm 0.12 \pm 0.25$	$54.9 \pm 5.6 \pm 10.0$		
$K^*(1410) \rightarrow K \rho(770)$	$0.66 \pm 0.04 \pm 0.04$	$-172.9 \pm 5.0 \pm 6.7$		
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$			$0.77 \pm 0.11 \pm 0.62$	$-93.6 \pm 11.2 \pm 12.6$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$1.02 \pm 0.13 \pm 0.41$	$-28.4 \pm 8.0 \pm 10.5$	$0.79 \pm 0.18 \pm 0.36$	$3.7 \pm 12.5 \pm 15.1$
$B_s \rightarrow (D_s K)_P \rho(770)$			$0.61 \pm 0.08 \pm 0.26$	$36.4 \pm 7.7 \pm 14.5$
Fit parameter	Value			
$m_{K_1(1400)} [\text{MeV}]$	$1394.9 \pm 8.8 \pm 12.7 \pm 21.2$			
$\Gamma_{K_1(1400)} [\text{MeV}]$	$224.0 \pm 15.9 \pm 22.3 \pm 20.9$			
$m_{K^*(1410)} [\text{MeV}]$	$1419.6 \pm 10.8 \pm 26.9 \pm 24.1$			
$\Gamma_{K^*(1410)} [\text{MeV}]$	$342.4 \pm 23.5 \pm 51.7 \pm 52.9$			
r	$xx.xx \pm 0.04 \pm 0.05 \pm 0.04$			
$\delta [^\circ]$	$xx.xx \pm 16.1 \pm 6.8 \pm 6.8$			
$\gamma - 2\beta_s [^\circ]$	$xx.xx \pm 16.1 \pm 11.6 \pm 6.2$			

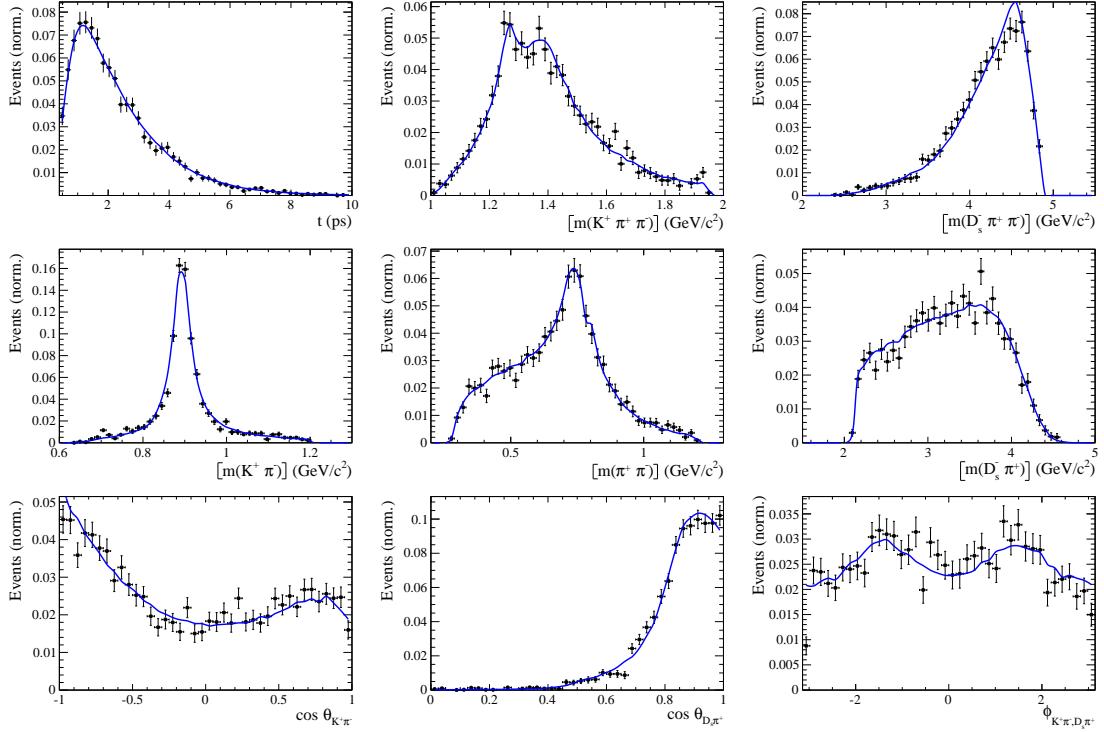


Figure 10.4: Projections of the full time-dependent amplitude fit.

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

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

851 11 Systematic uncertainties

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

858 11.1 Fit bias

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

863 11.2 Background subtraction

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

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

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

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

887 11.3 Decay-time acceptance

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

896 To improve the coverage of the multi-dimensional parameter space, a Cholesky decom-
 897 position [54] is used to generate a set of uncorrelated vectors from the covariance matrix
 898 $\text{cov}(\lambda_i, \lambda_j)$, where the vector λ includes the parameters Γ_s , $\Delta\Gamma_s$ and the $N = 4$ spline
 899 coefficients for each category of the simultaneous fit. The correlations between Γ_s ($\Delta\Gamma_s$)
 900 and the spline coefficients are measured by rerunning the acceptance fits described in
 901 Sec. 6.2 with the values of Γ_s ($\Delta\Gamma_s$) varied by $\pm 1\sigma$ and measuring the shift in the spline
 902 coefficients as a fraction of their uncertainty. For the correlation between Γ_s and $\Delta\Gamma_s$ we
 903 use the HFAG value [38].

904 11.4 Decay-time resolution and tagging

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

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

915 For each case, a new scaling function is derived:

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

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

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

918 Due to the high correlation between the decay-time resolution and the tagging calibra-
 919 tion, their systematic uncertainty has to be studied simultaneously. First, the decay-time
 920 fits to $B_s \rightarrow D_s \pi \pi$ data are repeated using the alternative decay-time error scaling
 921 functions. New tagging calibration parameters are obtained which are then used (together
 922 with the respective decay-time error scaling function) in the fits to $B_s \rightarrow D_s K \pi \pi$ data
 923 to define the Gaussian-constraints as discussed in Sec. 9. For the width of the Gaussians
 924 only the statistical error of the tagging calibration parameters are used since systematic
 925 uncertainties (except the systematic arising from the decay-time resolution which is already

926 included by the procedure described above) are found to be negligible, see Table 11.1.
 927 Finally, we take the biggest change in fit central value as the systematic for each parameter
 928 of the $B_s \rightarrow D_s K\pi\pi$ fits.

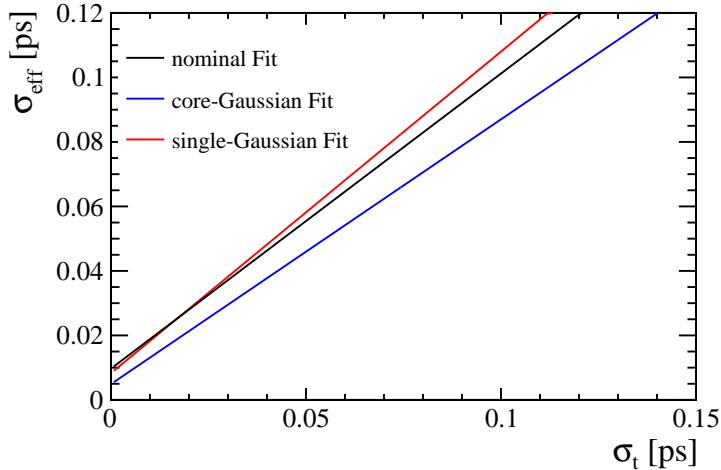


Figure 11.1: The measured resolution scaling function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data) for (black line) the nominal scaling, (blue line) only using the narrow gaussian width of the double gaussian fit model or (red line) when determining the resolution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

929 11.5 Production, detection asymmetries and mixing frequency

930 The systematic from the production, detection asymmetries and Δm_s (in case of $B_s \rightarrow$
 931 $D_s K\pi\pi$ decays) which are fixed in the fit are evaluated by means of a toy study similar
 932 to the procedure performed for the time-acceptance. The parameters are assumed to be
 933 uncorrelated.

934 11.6 Multiple candidates

935 The fraction of events with multiple candidates has been found to be very small, it is
 936 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
 937 candidates, while a systematic uncertainty is assigned by repeating the fit randomly
 938 keeping only one candidate when multiple ones are founds. No shifts in the fit central
 939 values are observed.

940 11.7 Length and momentum scales

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

944 11.8 Phase space acceptance

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

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

- 952 • No BDT cut is applied
- 953 • A tighter BDT requirement is used ($\text{BDTG} > 0.6$)
- 954 • No reweighting is applied
- 955 • Instead of the PID responses obtained from the `PIDCorr` tool, we use the `PIDGen`
956 tool to resample the PID variables [37]
- 957 • The raw MC PID variables are used
- 958 • Candidates with `BKGAT= 60` are removed

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

961 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.

962 11.9 Resonance description

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

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

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

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

982 11.10 Alternative amplitude models

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

- 988 • All amplitudes selected by Stage 1 of the model selection are included for both
989 $b \rightarrow c$ and $b \rightarrow u$ transitions
- 990 • The decay channel $K_1(1270)[D] \rightarrow K^*(892)\pi$ is added where the $K^*(892)\pi$ system
991 is in relative a D-wave state
- 992 • The decay channel $K_1(1400) \rightarrow K\rho(770)$ is added
- 993 • The decay channels $K_2^*(1430) \rightarrow K\rho(770)$ and $K_2^*(1430) \rightarrow K^*(892)\pi$ are added
- 994 • The decay channels $K(1460) \rightarrow K\rho(770)$ and $K(1460) \rightarrow K\sigma$ are added
- 995 • The $K(1460)$ resonance is removed
- 996 • The decay channels $K^*(1680) \rightarrow K\rho(770)$ and $K^*(1680) \rightarrow K^*(892)\pi$ are added
- 997 • The decay channels $K_2(1770) \rightarrow K\rho(770)$ and $K_2(1770) \rightarrow K^*(892)\pi$ are added
- 998 • The amplitudes $B_s \rightarrow (D_s\pi)_P K^*(892)$ and $B_s \rightarrow (D_sK)_P \rho(770)$ are replaced by
999 $B_s \rightarrow (D_s\pi)_S K^*(892)$ and $B_s \rightarrow (D_sK)_S \rho(770)$
- 1000 • Higher orbital angular momentum states are added for the amplitudes: $B_s[S, P, D] \rightarrow$
1001 $(D_s\pi)_P K^*(892)$ and $B_s[S, P, D] \rightarrow (D_sK)_P \rho(770)$
- 1002 • The amplitudes $B_s \rightarrow (D_s\pi)_P K^*(892)$ and $B_s \rightarrow (D_sK)_P \rho(770)$ are removed
- 1003 • The amplitudes $B_s \rightarrow (D_sK)_P \sigma$, $B_s \rightarrow (D_sK)_P f_2(1270)$ and $B_s \rightarrow$
1004 $(D_sK)_P f_0(1370)$ are added
- 1005 • The amplitudes $B_s \rightarrow (D_s\pi)_P K_0^*(1430)$ and $B_s \rightarrow (D_sK)_S K_2^*(1430)$ are added

1006 In total 20 different sets of amplitudes are fitted. In some cases, the fit fractions of
1007 additionally added amplitudes turn out to be exactly zero. These model are effectively
1008 not distinguishable from the baseline LASSO model and are not considered further. From
1009 the remaining 15 models, we compute the sample variance for each observable and take it
1010 as model uncertainty.

Table 11.1: Systematic uncertainties on the fit parameters of the fit to $B_s \rightarrow D_s\pi\pi\pi$ 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
ϵ_{tag}^{OS} Run-I	0.02	0.00	0.00	0.01	0.09	0.00		0.09
$\Delta \epsilon_{tag}^{OS}$ 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
ϵ_{tag}^{SS} Run-I	0.02	0.00	0.00	0.01	0.01	0.00		0.03
$\Delta \epsilon_{tag}^{SS}$ 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
ϵ_{tag}^{OS} Run-II	0.02	0.00	0.00	0.00	0.04	0.00		0.05
$\Delta \epsilon_{tag}^{OS}$ 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
ϵ_{tag}^{SS} Run-II	0.00	0.00	0.00	0.00	0.04	0.00		0.04
$\Delta \epsilon_{tag}^{SS}$ 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 $B_s \rightarrow D_s K\pi\pi$ 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 $B_s \rightarrow D_s K\pi\pi$ 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
$B_s \rightarrow D_s(K_1(1270) \rightarrow K^*(892)\pi)$ Mag	0.04	0.17	0.01	0.01	0.02	0.15	1.30	0.28	0.42	0.06		1.42
$B_s \rightarrow D_s(K_1(1270) \rightarrow K^*(892)\pi)$ Phase	0.08	0.20	0.03	0.01	0.16	0.06	0.85	0.31	0.20	1.10		1.47
$B_s \rightarrow D_s(K_1(1270) \rightarrow K_0^*(1430)\pi)$ Mag	0.07	0.17	0.02	0.01	0.11	0.25	3.96	3.69	0.45	2.20		5.87
$B_s \rightarrow D_s(K_1(1270) \rightarrow K_0^*(1430)\pi)$ Phase	0.24	0.16	0.02	0.01	0.18	0.15	7.28	0.21	0.51	4.47		8.57
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Mag($b \rightarrow c$)	0.08	0.13	0.02	0.03	0.43	0.27	1.38	0.28	0.38	1.44		2.12
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Phase($b \rightarrow c$)	0.07	0.24	0.01	0.03	0.13	0.28	0.66	0.25	0.32	0.69		1.11
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Mag($b \rightarrow u$)	0.21	0.19	0.02	0.04	0.06	0.19	0.83	0.24	0.56	2.27		2.52
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Phase($b \rightarrow u$)	0.01	0.16	0.04	0.10	0.15	0.36	0.79	0.43	0.25	0.88		1.36
$B_s \rightarrow D_s(K^*(1410) \rightarrow K^*(892)\pi)$ Mag($b \rightarrow c$)	0.32	0.13	0.03	0.05	0.19	0.18	1.08	0.28	1.60	0.09		2.00
$B_s \rightarrow D_s(K^*(1410) \rightarrow K^*(892)\pi)$ Phase($b \rightarrow c$)	0.25	0.23	0.01	0.01	0.21	0.10	1.42	0.22	0.75	0.62		1.79
$B_s \rightarrow D_s(K^*(1410) \rightarrow K\rho(770))$ Mag	0.49	0.20	0.01	0.01	0.12	0.17	0.60	0.18	0.19	0.15		0.88
$B_s \rightarrow D_s(K^*(1410) \rightarrow K\rho(770))$ Phase	0.23	0.22	0.01	0.01	0.10	0.13	0.34	0.12	0.29	1.22		1.35
$B_s \rightarrow D_s(K(1460) \rightarrow K^*(892)\pi)$ Mag($b \rightarrow u$)	0.03	0.24	0.02	0.04	0.23	0.22	0.68	0.76	5.39	1.96		5.84
$B_s \rightarrow D_s(K(1460) \rightarrow K^*(892)\pi)$ Phase($b \rightarrow u$)	0.02	0.30	0.03	0.04	0.13	0.21	0.64	0.40	0.48	0.55		1.12
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Mag($b \rightarrow c$)	0.15	0.16	0.02	0.02	0.35	0.24	1.28	0.20	2.66	0.99		3.16
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Phase($b \rightarrow c$)	0.01	0.20	0.01	0.01	0.20	0.47	0.95	0.18	0.34	0.59		1.30
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Mag($b \rightarrow u$)	0.15	0.14	0.04	0.03	0.37	0.13	0.47	0.27	1.73	0.68		1.99
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Phase($b \rightarrow u$)	0.01	0.26	0.05	0.03	0.88	0.28	0.56	0.21	0.42	0.10		1.21
$B_s \rightarrow (D_sK)_P \rho(770)$ Mag($b \rightarrow u$)	0.45	0.24	0.01	0.05	0.83	0.49	1.34	0.38	2.81	0.33		3.34
$B_s \rightarrow (D_sK)_P \rho(770)$ Phase($b \rightarrow u$)	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
$\Gamma_{K_1(1400)}$	0.05	0.22	0.02	0.01	0.29	0.13	1.23	0.12	0.25	0.46		1.31
$m_{K^*(1410)}$	0.08	0.19	0.01	0.01	0.51	0.11	1.69	0.27	1.63	0.51		2.22
$\Gamma_{K^*(1410)}$	0.30	0.17	0.01	0.01	0.10	0.18	1.17	0.59	1.71	0.15		2.25
r	0.07	0.19	0.05	0.10	0.38	0.29	1.02	0.20	0.18	0.58		1.64
δ	0.02	0.17	0.04	0.06	0.03	0.10	0.24	0.07	0.14	0.23		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		0.82

1011 A Stripping and Trigger cuts

1012 The following text describes variables which are used in Table 1.1 and might be ambiguous,
 1013 or which benefits are not straight forward. Where noted, different cut values are applied
 1014 for Run-I and Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest
 1015 approach. This variable is used to ensure that all D_s and $X_{s,d}$ daughters originate from
 1016 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 $B_s^0 \rightarrow D_s K \pi \pi$ 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	$m_{D_s} \pm 100$ MeV	$m_{D_s} \pm 80$ MeV
D_s Daughter p_T	$\sum_{i=1}^3 p_{t,i} > 1800$ 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 MeV	< 3500 MeV
$X_{s,d}$ Daughter p	> 2 GeV/ c	
$X_{s,d}$ Daughter DOCA	< 0.4 mm	
$X_{s,d}$ Daughter p_T	$\sum_{i=1}^3 p_{t,i} > 1250$ 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] MeV/ c^2	[5000, 6000] 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	

1017

1018 Table 1.2 summarizes the trigger requirements imposed by the Hlt1 line used in this
 1019 analysis for Run-I. At least one of the six decay particles must pass the listed requirements
 1020 in order for the event to be stored for further analysis. For Run-II, this trigger line was
 1021 updated and uses a multivariate classifier which takes the variables listed in Table 1.2 as
 1022 input, rather than directly cutting on them.

1023 The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the
 1024 b-hadron p_T , its flight distance χ^2 from the nearest PV and the sum of the B_s^0 and D_s
 1025 vertex χ^2 divided by the sum of their number of degrees of freedom. Table 1.3 summarizes
 1026 the cuts applied by the inclusive ϕ trigger, which requires that a $\phi \rightarrow KK$ 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 $\phi \rightarrow KK$ 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
ϕp_T	> 2.5 GeV/c
ϕ vertex χ^2/nDoF	< 20
ϕ IP χ^2 to any PV	> 5

1027

B Details of multivariate classifier

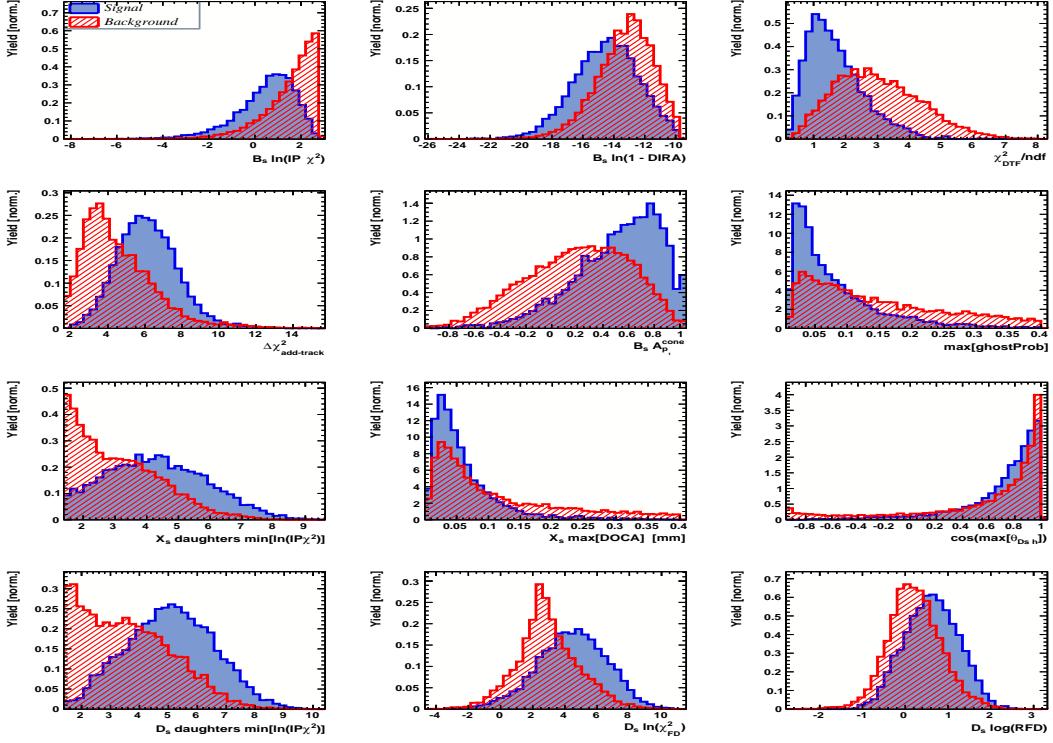


Figure A.1: Variables used to train the BDTG for category [Run-I,L0-TOS].

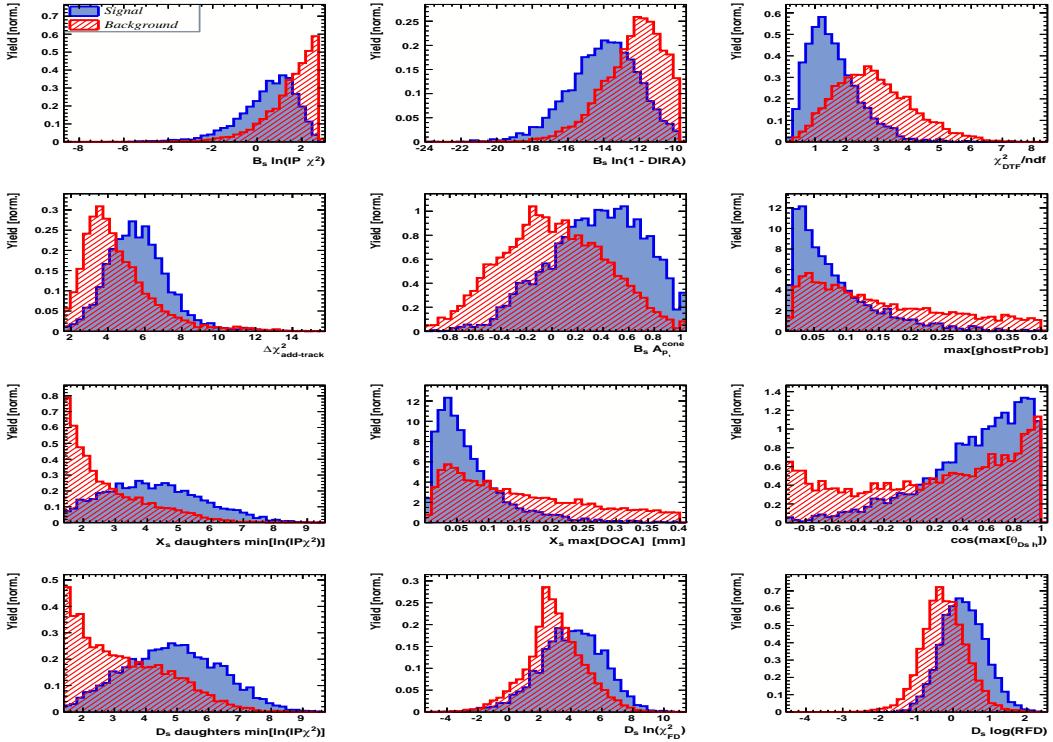


Figure A.2: Variables used to train the BDTG for category [Run-I,L0-TIS].

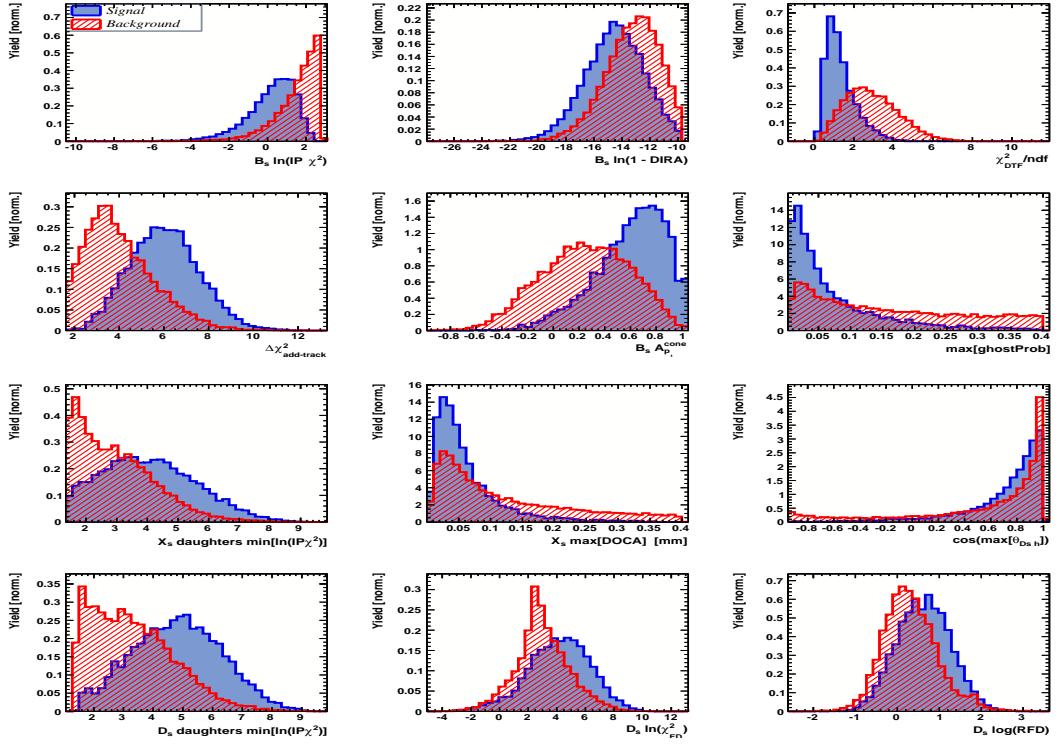


Figure A.3: Variables used to train the BDTG for category [Run-II,L0-TOS].

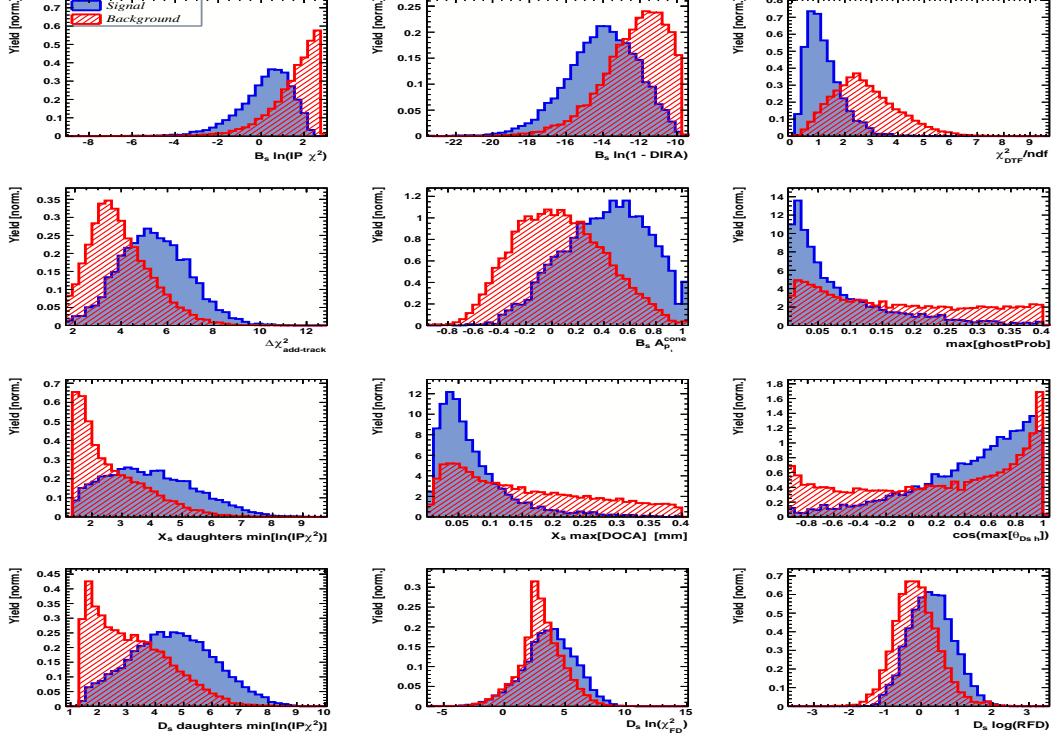


Figure A.4: Variables used to train the BDTG for category [Run-II,L0-TIS].

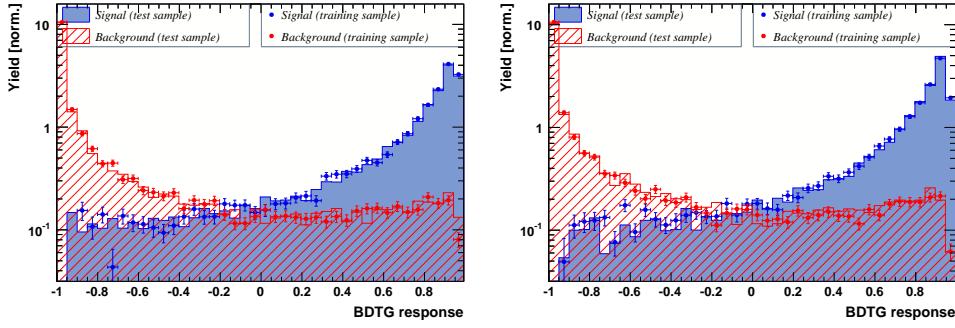


Figure A.5: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].

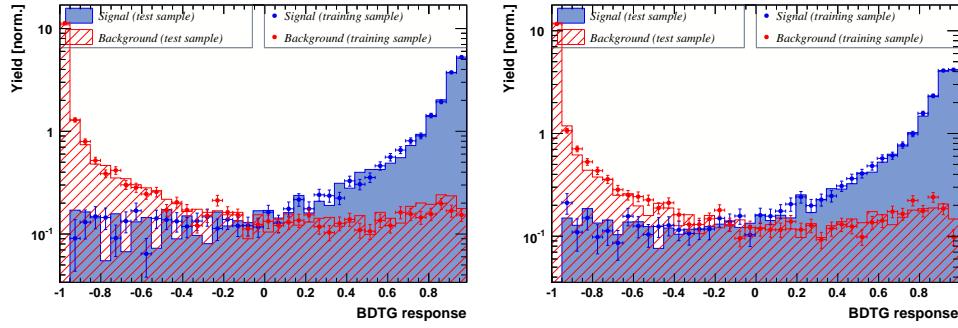


Figure A.6: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].

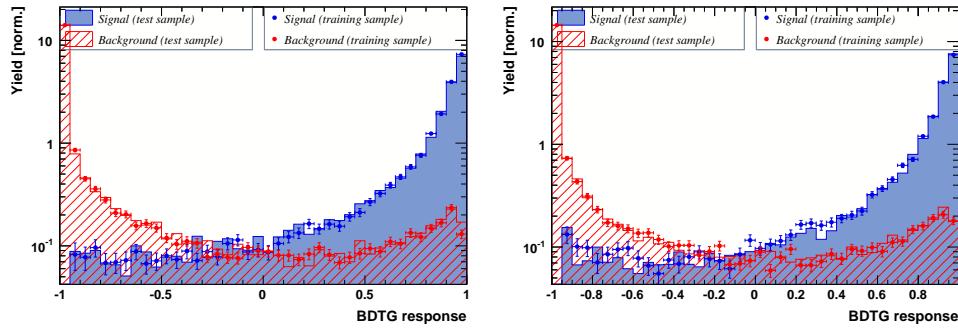


Figure A.7: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].

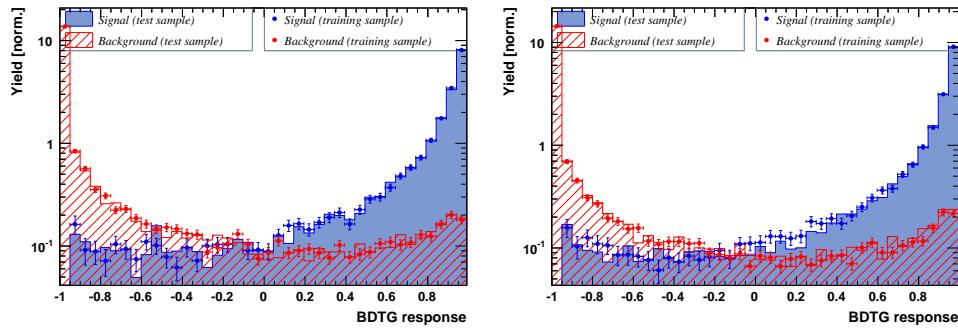


Figure A.8: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].

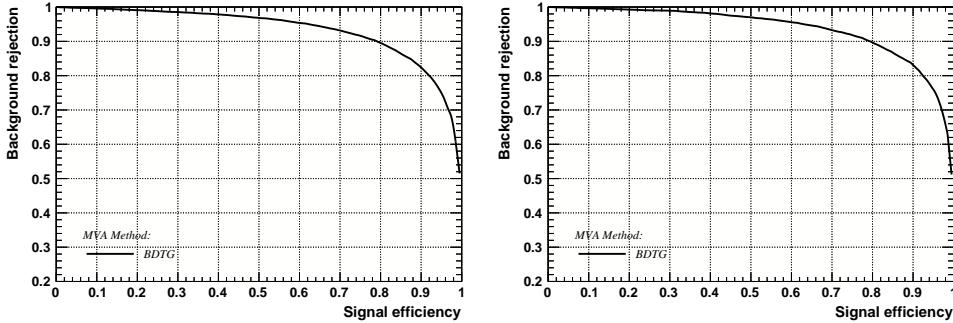


Figure A.9: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].

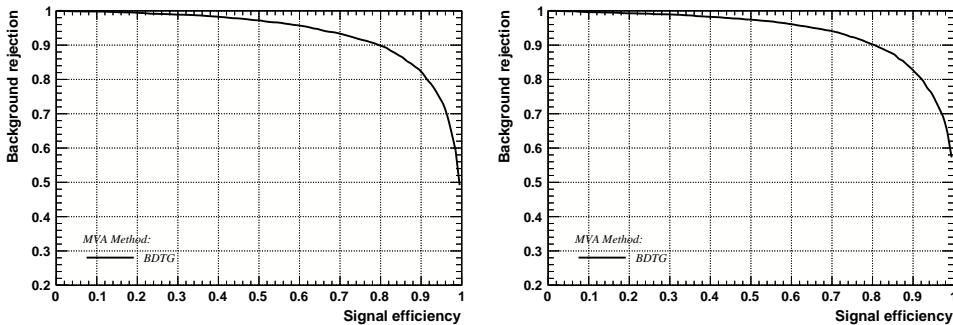


Figure A.10: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TIS].

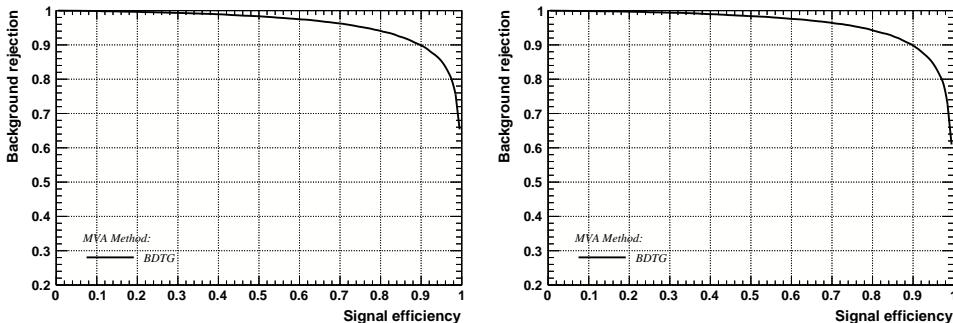


Figure A.11: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TOS].

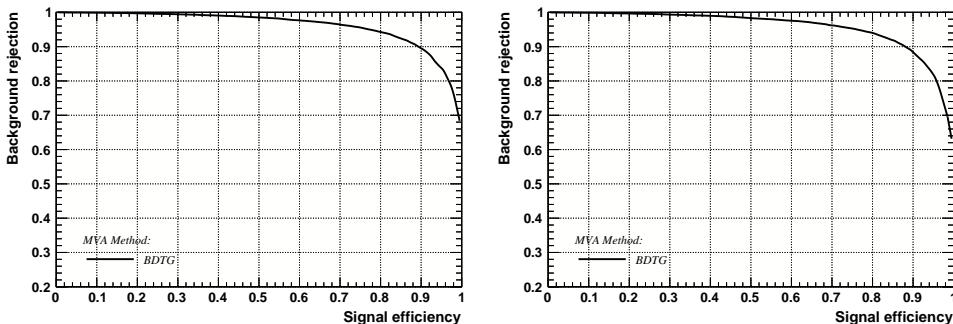


Figure A.12: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TIS].

1029 C Detailed mass fits

1030 In this section, all fits to the mass distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ and $B_s^0 \rightarrow D_sK\pi\pi$
 1031 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-
 1032 II), D_s final state ($D_s \rightarrow KK\pi$ non-resonant, $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*K$, or $D_s \rightarrow \pi\pi\pi$) and
 1033 L0 trigger category.

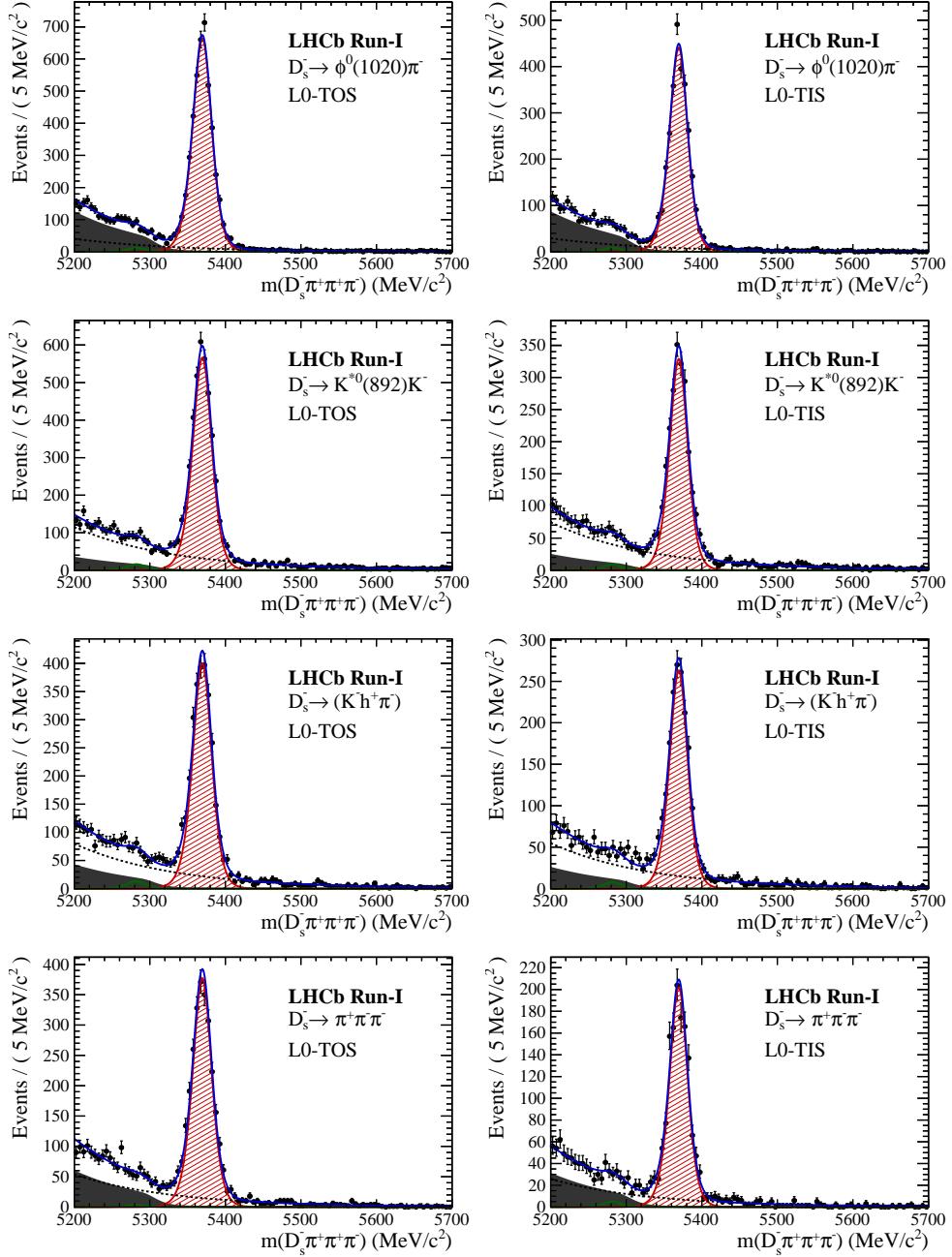


Figure B.1: Invariant mass distributions of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates for Run-I data.

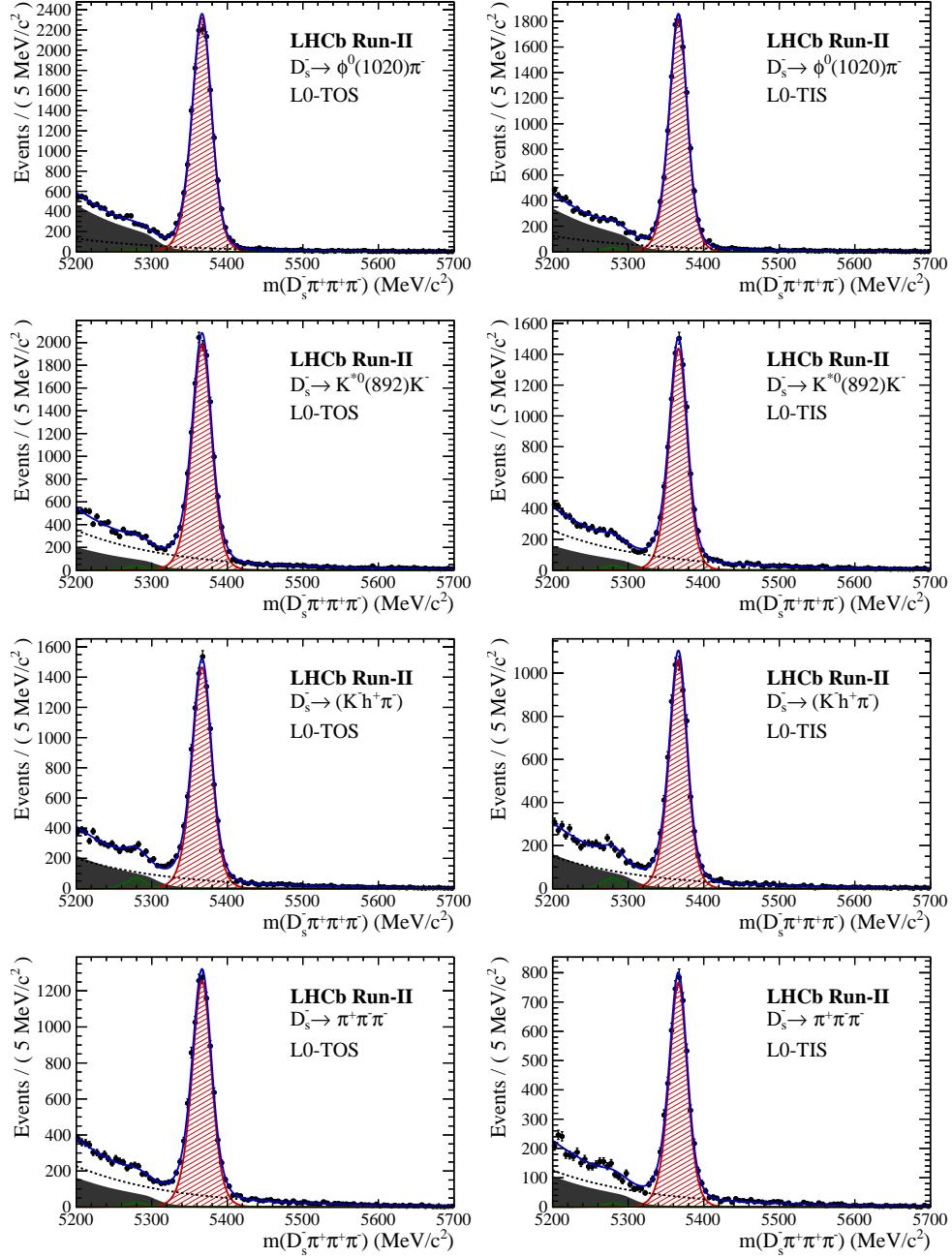


Figure B.2: Invariant mass distributions of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates for Run-II data.

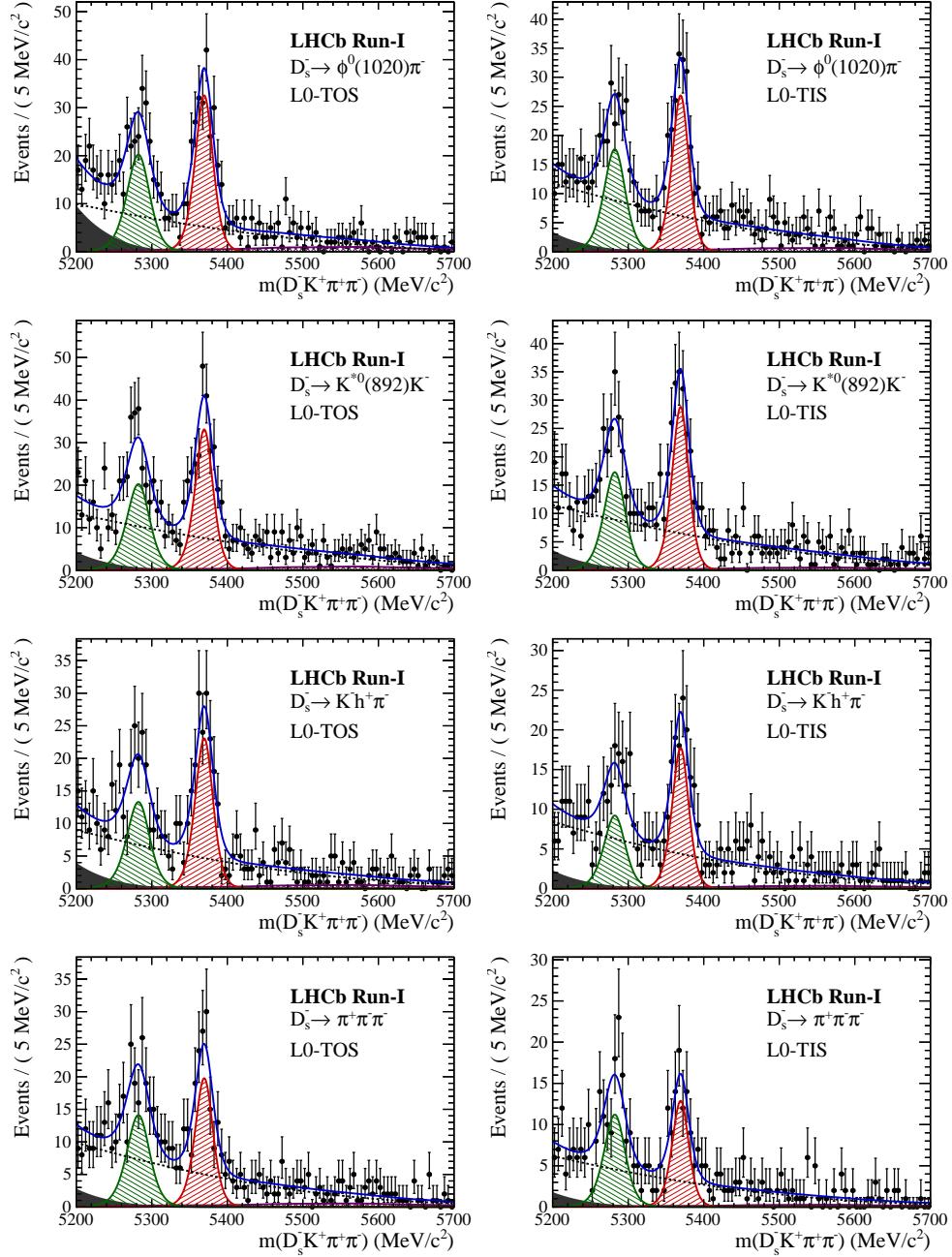


Figure B.3: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi\pi$ candidates for Run-I data.

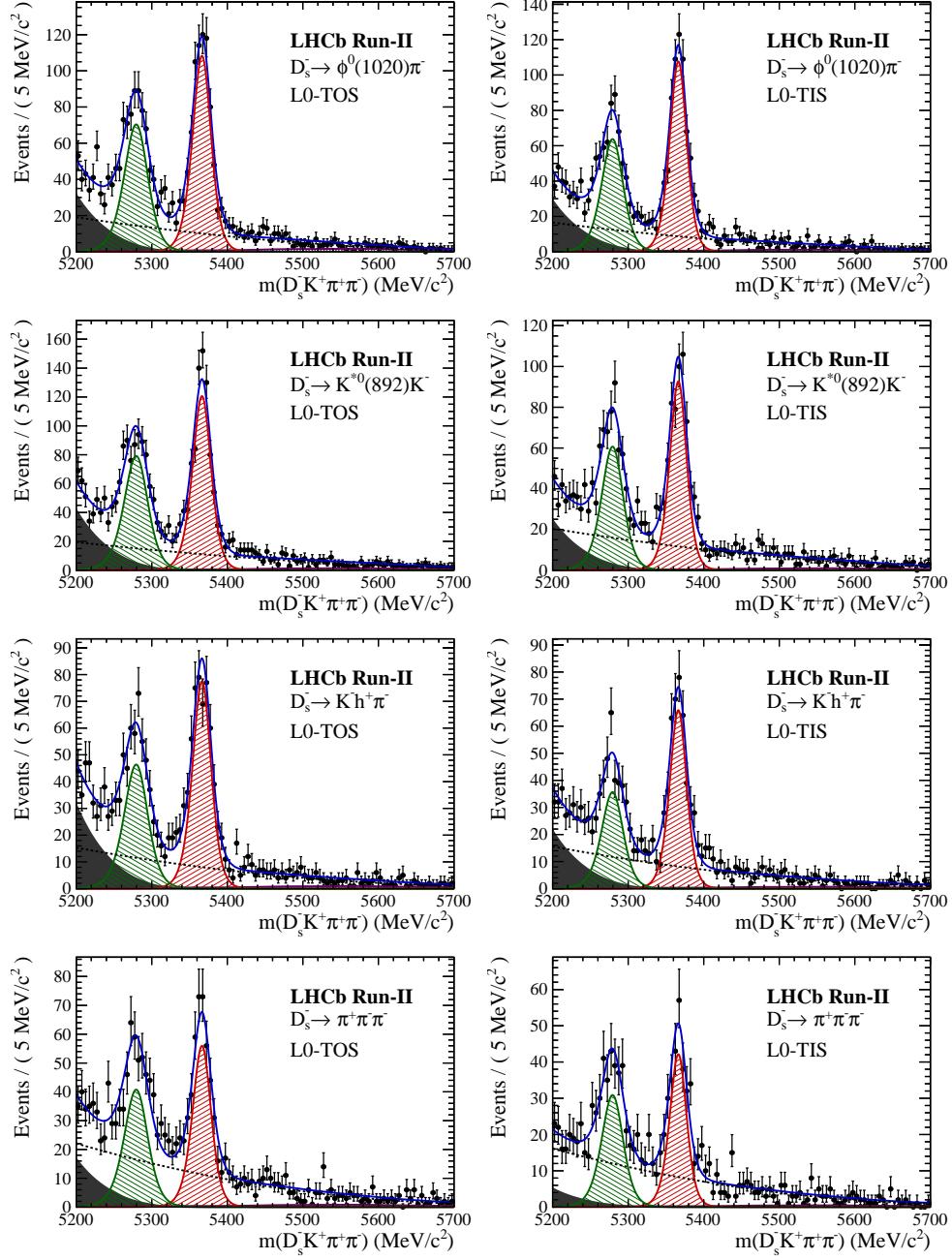


Figure B.4: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-II data.

1034 D Decay-time Resolution fits

1035 This section contains all fits to the distributions of the decay time difference Δt between
 1036 the true and the reconstructed decay time of the truth-matched B_s^0 candidates on MC.
 1037 The fits are performed in bins of the decay time error σ_t , where an adaptive binning
 1038 scheme is used to ensure that approximately the same number of events are found in each
 1039 bin.

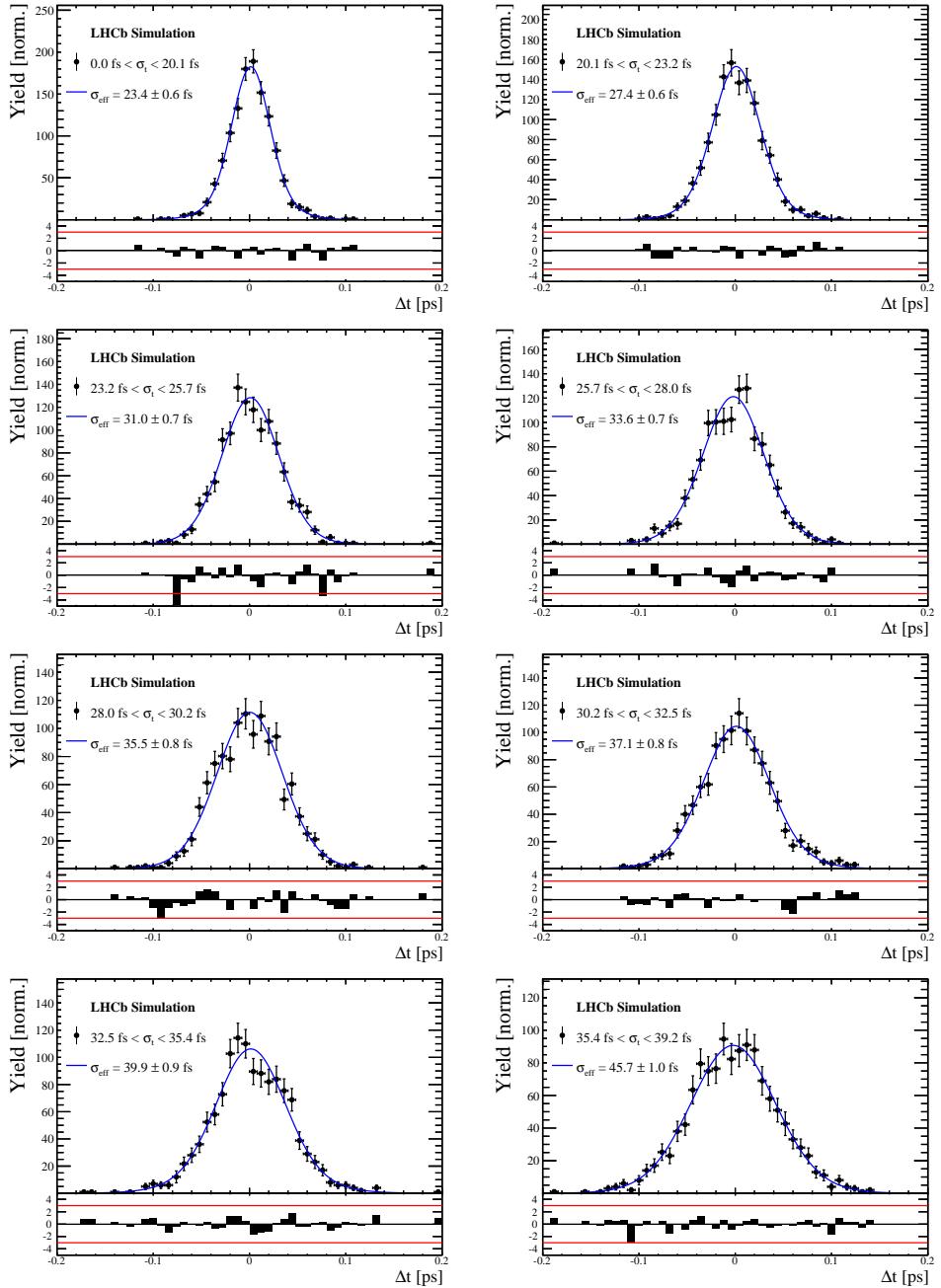


Figure C.1: Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K \pi\pi$ MC candidates in bins of the per-event decay time error estimate..

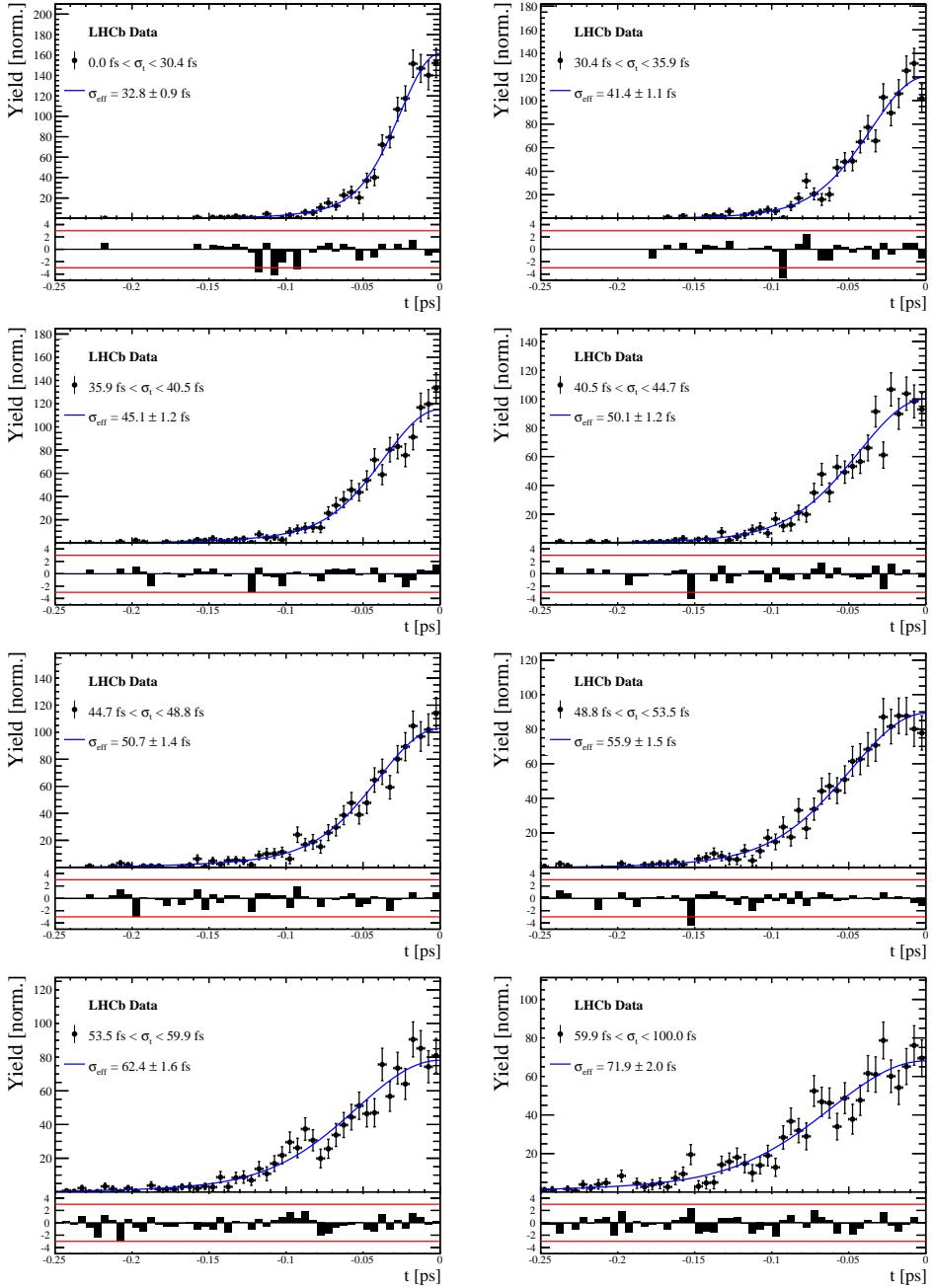


Figure C.2: Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks, for bins in the per-event decay time error estimate.

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 $B_s \rightarrow D_s K\pi\pi$ 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.

1040 E Comparison of time-acceptance in subsamples

1041 Figure C.1 shows the spline coefficients obtained by fitting the decay-time distribution of
 1042 $B_s^0 \rightarrow D_s\pi\pi\pi$ data candidates in different subsamples. Sufficient agreement is observed
 1043 within a given data-taking period, while the acceptance shapes for Run-I and Run-II
 1044 data differ significantly. The fitted splines for the different D_s final states are in a good
 1045 agreement. The largest deviations are observed between the different L0 categories.

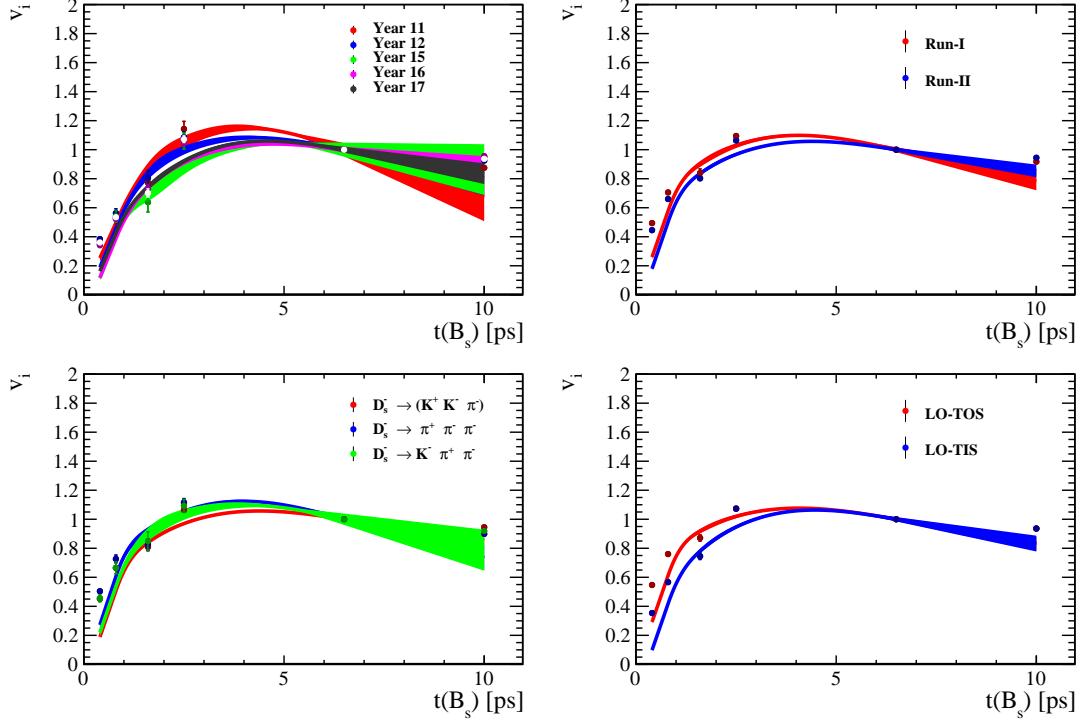


Figure C.1: Comparison of the spline coefficients (point with error bars) obtained from time-dependent fits to the $B_s^0 \rightarrow D_s\pi\pi\pi$ decay-time for different subsamples: (top-left) different years of data-taking; (top-right) different data-taking periods; (bottom-left) different D_s final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

¹⁰⁴⁶ **F Spin Amplitudes**

¹⁰⁴⁷ The spin factors used for $B \rightarrow P_1 P_2 P_3 P_4$ 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	$B \rightarrow (P P_1)$, $P \rightarrow (S P_2)$, $S \rightarrow (P_3 P_4)$	1
2	$B \rightarrow (P P_1)$, $P \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(P) L_{(1)}^\alpha(V)$
3	$B \rightarrow (A P_1)$, $A \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) P_{(1)}^{\alpha\beta}(A) L_{(1)\beta}(V)$
4	$B \rightarrow (A P_1)$, $A[D] \rightarrow (P_2 V)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(2)}^{\alpha\beta}(A) L_{(1)\beta}(V)$
5	$B \rightarrow (A P_1)$, $A \rightarrow (S P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(1)}^\alpha(A)$
6	$B \rightarrow (A P_1)$, $A \rightarrow (T P_2)$, $T \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(1)\beta}(A) L_{(2)}^{\alpha\beta}(T)$
7	$B \rightarrow (V_1 P_1)$, $V_1 \rightarrow (V_2 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$L_{(1)\mu}(B) P_{(1)}^{\mu\alpha}(V_1) \epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\beta(V_1) p_{V_1}^\gamma L_{(1)}^\delta(V_2)$
8	$B \rightarrow (PT P_1)$, $PT \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) P_{(2)}^{\alpha\beta\gamma\delta}(PT) L_{(1)\gamma}(PT) L_{(1)\delta}(V)$
9	$B \rightarrow (PT P_1)$, $PT \rightarrow (S P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}(PT)^{\alpha\beta}$
10	$B \rightarrow (PT P_1)$, $PT \rightarrow (T P_2)$, $T \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) P_{(2)}^{\alpha\beta\gamma\delta}(PT) L_{(2)\gamma\delta}(T)$
11	$B \rightarrow (T P_1)$, $T \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(2)\mu\nu}(B) P_{(2)}^{\mu\nu\rho\alpha}(T) \epsilon_{\alpha\beta\gamma\delta} L_{(2)\rho}^\beta(T) p_T^\gamma P_{(1)}^{\delta\sigma}(T) L_{(1)\sigma}(V)$
12	$B \rightarrow (T_1 P_1)$, $T_1 \rightarrow (T_2 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\mu\nu}(B) P_{(2)}^{\mu\nu\rho\alpha}(T_1) \epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\beta(T_1) p_{T_1}^\gamma L_{(2)\rho}^\delta(T_2)$
13	$B \rightarrow (S_1 S_2)$, $S_1 \rightarrow (P_1 P_2)$, $S_2 \rightarrow (P_3 P_4)$	1
14	$B \rightarrow (V S)$, $V \rightarrow (P_1 P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(1)}^\alpha(V)$
15	$B \rightarrow (V_1 V_2)$, $V_1 \rightarrow (P_1 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(V_1) L_{(1)}^\alpha(V_2)$
16	$B[P] \rightarrow (V_1 V_2)$, $V_1 \rightarrow (P_1 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$\epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\alpha(B) L_{(1)}^\beta(V_1) L_{(1)}^\gamma(V_2) p_B^\delta$
17	$B[D] \rightarrow (V_1 V_2)$, $V_1 \rightarrow (P_1 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(1)}^\alpha(V_1) L_{(1)}^\beta(V_2)$
18	$B \rightarrow (T S)$, $T \rightarrow (P_1 P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\beta}(T)$
19	$B \rightarrow (V T)$, $T \rightarrow (P_1 P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(2)}^{\alpha\beta}(T) L_{(1)\beta}(V)$
20	$B[D] \rightarrow (TV)$, $T \rightarrow (P_1 P_2)$, $V \rightarrow (P_3 P_4)$	$\epsilon_{\alpha\beta\delta\gamma} L_2^{\alpha\mu}(B) L_{2\mu}^\beta L_{(1)}^\gamma(V) p_B^\delta$
21	$B \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(T_1) L_{(2)}^{\alpha\beta}(T_2)$
22	$B[P] \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$\epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\alpha(B) L_{(2)}^{\beta\mu}(T_1) L_{(2)\mu}^\gamma(T_2) p_B^\delta$
23	$B[D] \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\gamma}(T_1) L_{(2)\gamma}^\beta(T_2)$

1048 G Considered Decay Chains

1049 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
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ \rightarrow \pi^+ K^*(1430)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \omega(782)]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ \kappa]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \sigma]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow \sigma^0(D_s^- K^+)_S$
$B_s [S, P, D] \rightarrow \sigma^0(D_s^- K^+)_V$
$B_s \rightarrow \rho(770)^0(D_s^- K^+)_S$
$B_s [S, P, D] \rightarrow \rho(770)^0(D_s^- K^+)_V$
$B_s \rightarrow K^*(892)^0(D_s^- \pi^+)_S$
$B_s [S, P, D] \rightarrow K^*(892)^0(D_s^- \pi^+)_V$
$B_s \rightarrow (D_s^- K^+)_S (\pi^+ \pi^-)_S$

H Data-simulation comparisson

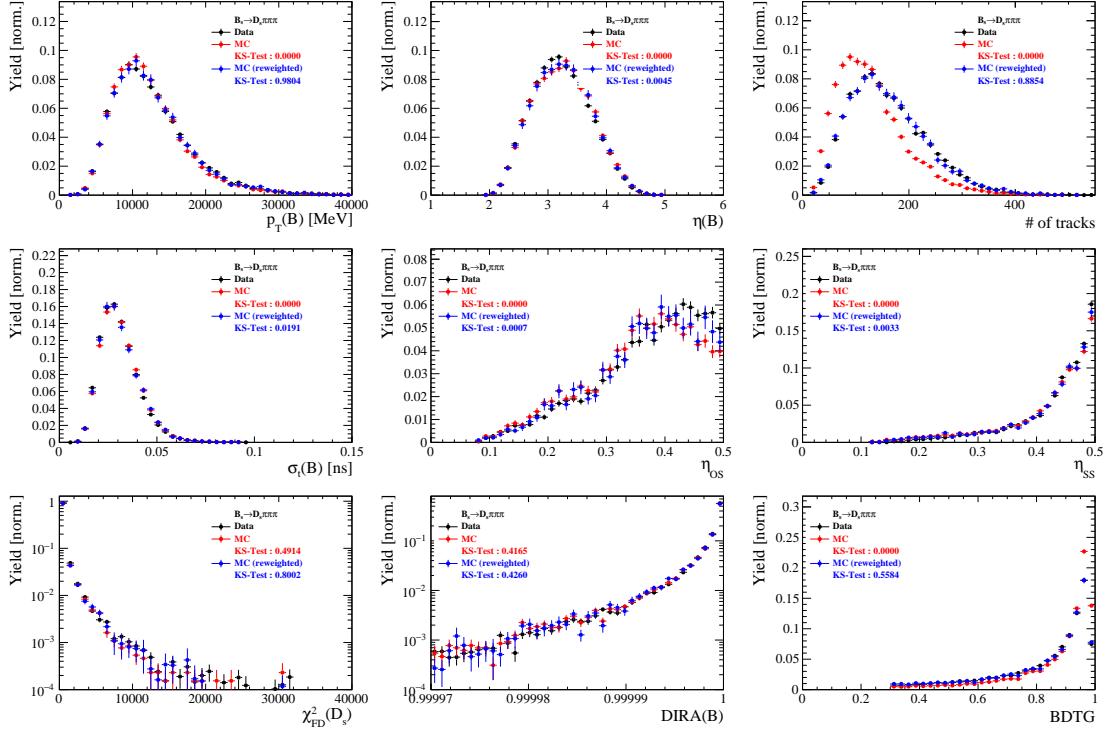


Figure C.1: Comparison between data and MC of selected variables for $B_s \rightarrow D_s \pi\pi\pi$ decays.

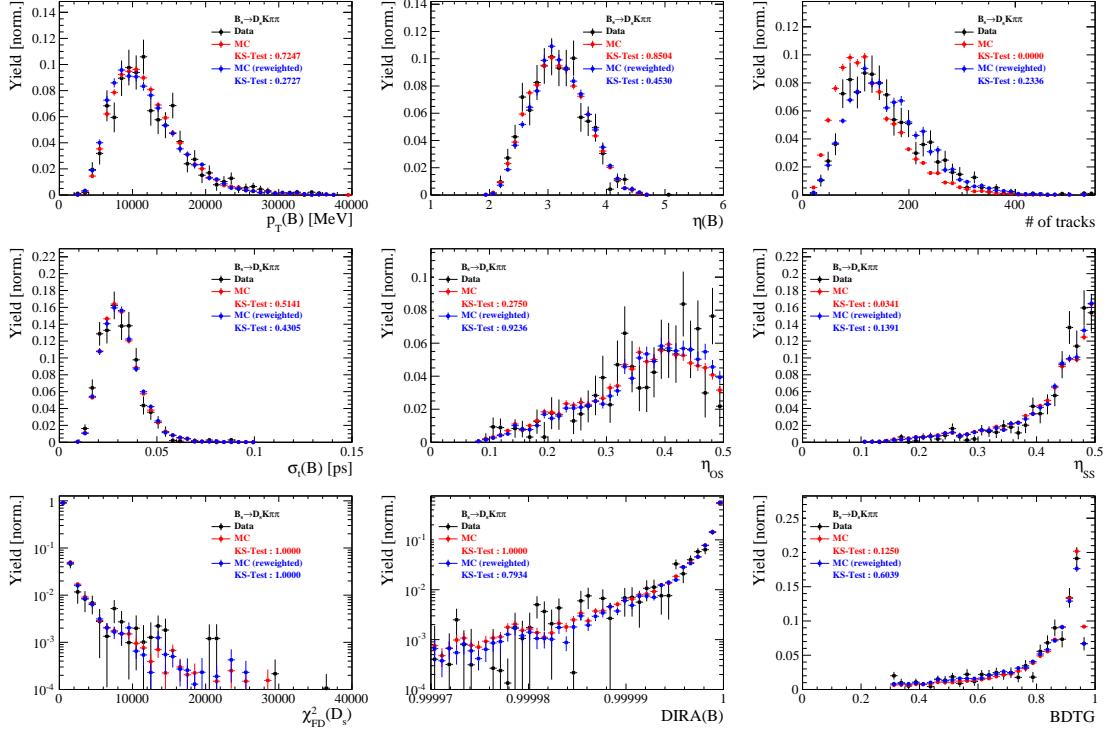


Figure C.2: Comparison between data and MC of selected variables for $B_s \rightarrow D_s K\pi\pi$ decays.

1051 I Data distributions

1052 I.1 Comparison of signal and calibration channels

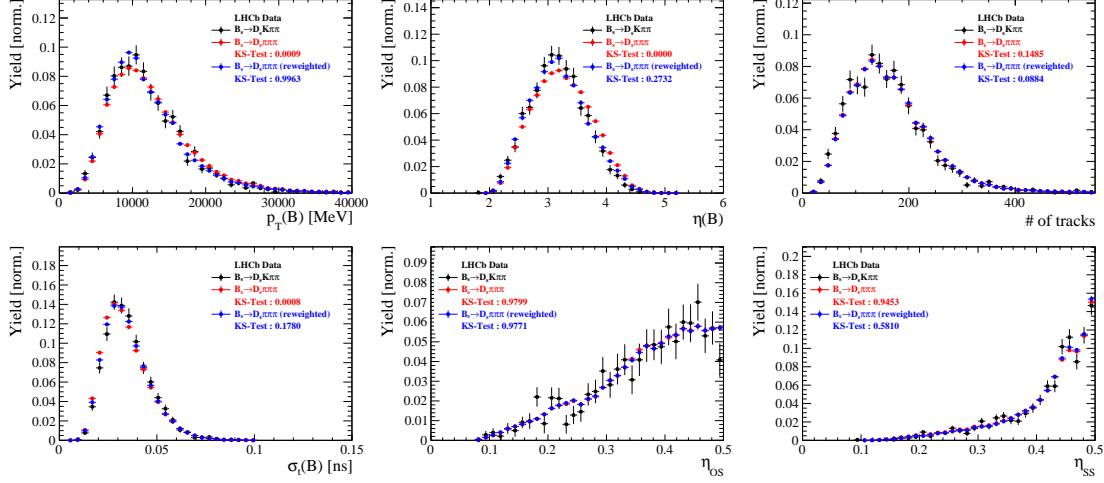


Figure C.1: Comparison between $B_s \rightarrow D_s K\pi\pi$ and $B_s \rightarrow D_s \pi\pi\pi$ decays for selected variables.

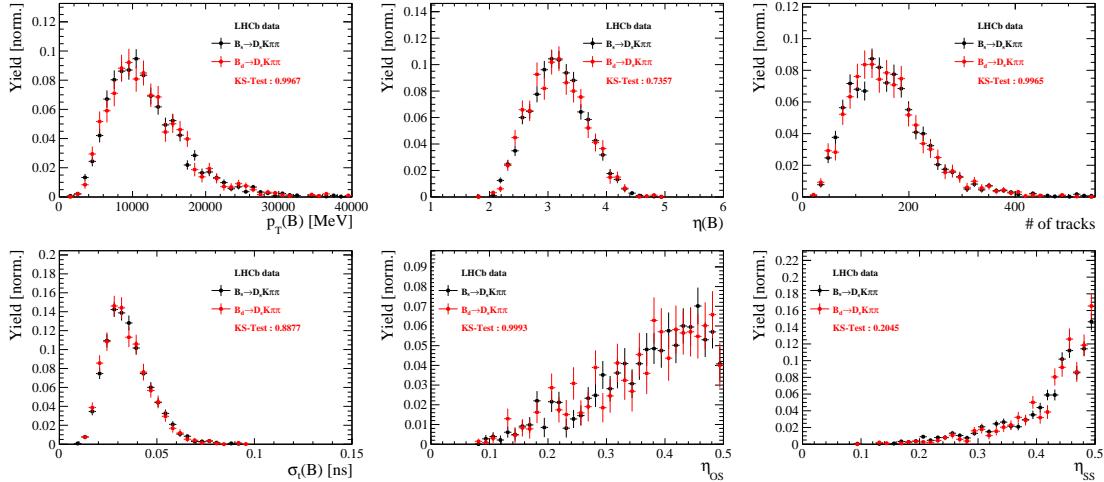


Figure C.2: Comparison between $B_s \rightarrow D_s K\pi\pi$ and $B_d \rightarrow D_s K\pi\pi$ decays for selected variables.

1053 I.2 Comparison of Run-I and Run-II data

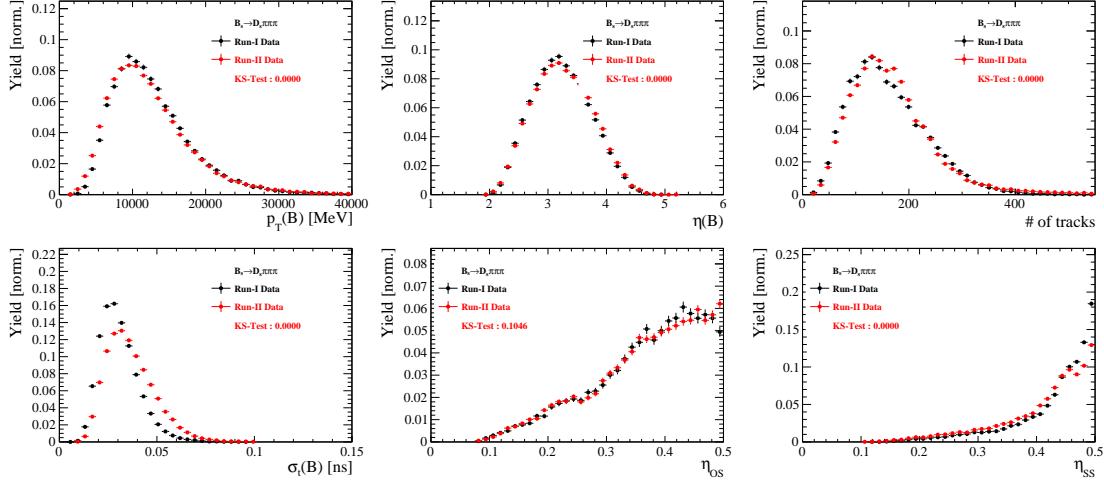


Figure C.3: Comparison of selected variables for Run-I and Run-II data.

1054 I.3 Comparison of D_s final states

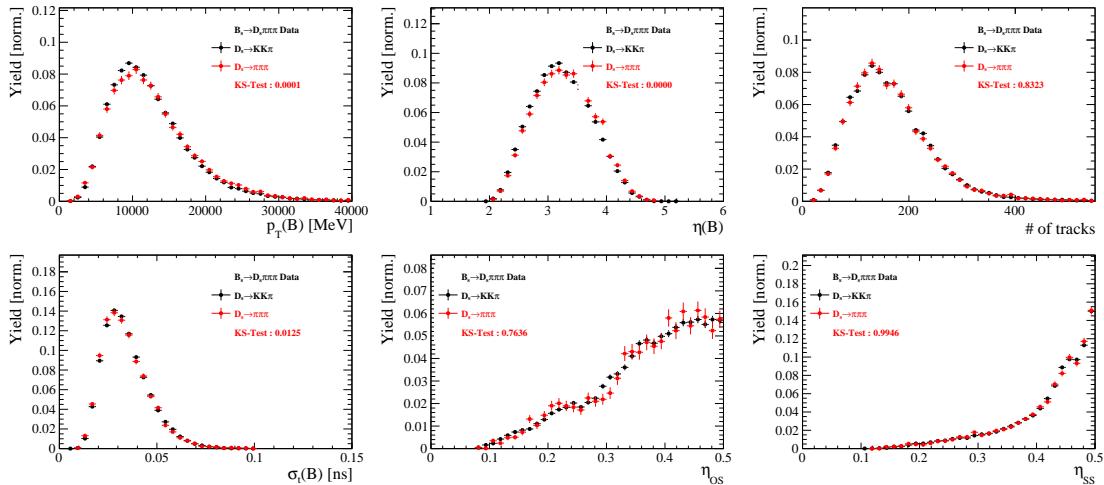


Figure C.4: Comparison of selected variables for different D_s final states.

1055 I.4 Comparison of trigger categories

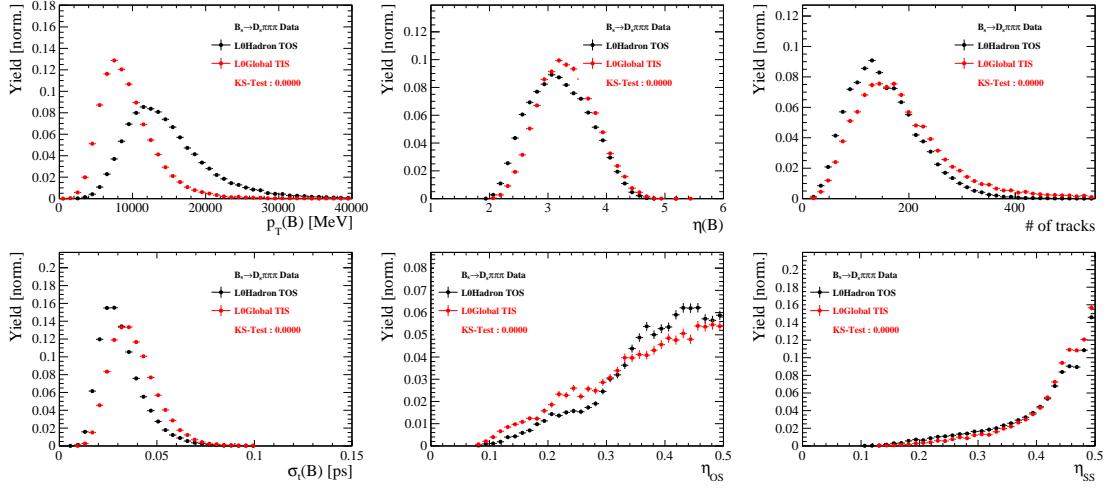


Figure C.5: Comparison of selected variables for different trigger categories.

1056 References

- 1057 [1] S. Blusk, *First observations and measurements of the branching fractions for the*
1058 *decays $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$,* .
- 1059 [2] LHCb, S. Blusk, *Measurement of the CP observables in $\bar{B}_s^0 \rightarrow D_s^+ K^-$ and first obser-*
1060 *vation of $\bar{B}_{(s)}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}_s^0 \rightarrow D_{s1}(2536)^+ \pi^-$,* 2012. [arXiv:1212.4180](#).
- 1061 [3] R. Fleischer, *New strategies to obtain insights into CP violation through $B_s \rightarrow$*
1062 *$D_s^\pm K^\mp, D_s^{*\pm} K^\mp, \dots$ and $B_d \rightarrow D^\pm \pi^\mp, D^{*\pm} \pi^\mp, \dots$ decays,* Nucl. Phys. **B671** (2003) 459,
1063 [arXiv:hep-ph/0304027](#).
- 1064 [4] K. De Bruyn *et al.*, *Exploring $B_s \rightarrow D_s^{(*)\pm} K^\mp$ Decays in the Presence of a Sizable*
1065 *Width Difference $\Delta\Gamma_s$,* Nucl. Phys. **B868** (2013) 351, [arXiv:1208.6463](#).
- 1066 [5] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 1067 [6] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation*
1068 *of production processes*, Annals of Physics **18** (1962), no. 2 198 .
- 1069 [7] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*,
1070 Phys. Rev. D **11** (1975) 3165.
- 1071 [8] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of
1072 Physics **108** (1977), no. 2 454 .
- 1073 [9] P. d'Argent *et al.*, *Amplitude Analyses of $D^0 \rightarrow \pi^+ \pi^- \pi^+ \pi^-$ and $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$*
1074 *Decays*, JHEP **05** (2017) 143, [arXiv:1703.08505](#).
- 1075 [10] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay*
1076 *widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.
- 1077 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo
1078 Cimento Series 10 **34** (1964), no. 6 1644.
- 1079 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys.
1080 **C40** (2016), no. 10 100001.
- 1081 [13] D. V. Bugg, *The mass of the σ pole*, Journal of Physics G Nuclear Physics **34** (2007)
1082 151, [arXiv:hep-ph/0608081](#).
- 1083 [14] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-*
1084 *dominance prediction for $\rho \rightarrow e^+ e^-$* , Phys. Rev. Lett. **21** (1968) 244.
- 1085 [15] M. S. et al. *Search for CP violation in the $D^0 \rightarrow KK\pi\pi$ decay through a full*
1086 *amplitude analysis*, LHCb-ANA-2017-064.
- 1087 [16] S. M. Flatté, *Coupled-channel analysis of the $\pi\eta$ and KK systems near KK threshold*,
1088 Physics Letters B **63** (1976), no. 2 224 .
- 1089 [17] BES Collaboration, M. Ablikim *et al.*, *Resonances in $J/\psi \rightarrow \phi \pi^+ \pi^-$ and $\phi K^+ K^-$* ,
1090 Phys. Lett. **B607** (2005) 243, [arXiv:hep-ex/0411001](#).

- [1091] [18] W. Dunwoodie, Fits to $K\pi$ $I = \frac{1}{2}$ S -wave amplitude and phase data.
- [1092] [19] D. Aston *et al.*, *A Study of $K^-\pi^+$ Scattering in the Reaction $K^-p \rightarrow K^-\pi^+n$ at $11\text{GeV}/c$* , Nucl. Phys. **B296** (1988) 493.
- [1094] [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)].
- [1097] [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.
- [1099] [22] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- [1100] [23] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys. Rev. **60** (1941) 61.
- [1102] [24] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys. Rev. D **57** (1998) 431.
- [1104] [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.
- [1106] [26] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. **D51** (1995) 2247.
- [1108] [27] J.-J. Zhu, *Explicit expressions of spin wave functions*, arXiv:hep-ph/9906250.
- [1109] [28] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput. Phys. Commun. **180** (2009) 1847, arXiv:0805.2956.
- [1111] [29] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [1113] [30] M. Karbach and M. Kenzie, *Gammacombo package*, <http://gammacombo.hepforge.org/web/HTML/index.html>, 2014.
- [1115] [31] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.
- [1117] [32] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.
- [1119] [33] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [1121] [34] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.
- [1123] [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.

- 1125 [36] LHCb collaboration, L. Zhang, *Measurements of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$*
 1126 *decays in the low $K^+ K^-$ mass range with 13 TeV data*, LHCb-ANA-2017-028.
- 1127 [37] A. Poluektov, *Correction of simulated particle identification response in LHCb using*
 1128 *kernel density estimation*, LHCb-INT-2017-007.
- 1129 [38] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and*
 1130 *τ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots
 1131 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 1132 [39] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*
 1133 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 1134 [40] LHCb Collaboration, R. Aaij *et al.*, *Measurement of angular and cp asymmetries*
 1135 *in $D^0 \rightarrow \pi^+ \pi^- \mu^+ \mu^-$ and $D^0 \rightarrow K^+ K^- \mu^+ \mu^-$ decays*, Phys. Rev. Lett. **121** (2018)
 1136 091801.
- 1137 [41] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the*
 1138 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- 1139 [42] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at*
 1140 *LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- 1141 [43] LHCb, R. Aaij *et al.*, *Measurement of B^0 , B_s^0 , B^+ and Λ_b^0 production asymmetries in 7*
 1142 *and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- 1143 [44] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of*
 1144 *the $K\pi$ Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-
 1145 2012-027, CERN, Geneva, Feb, 2013.
- 1146 [45] A. Davis *et al.*, *Measurement of the instrumental asymmetry for $K^- \pi^+$ -pairs at LHCb*
 1147 *in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN,
 1148 Geneva, Mar, 2018.
- 1149 [46] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly*
 1150 *forbidden transitions in $D \rightarrow K(S)$, $K(L) + \pi$'s decays*, Phys. Lett. **B349** (1995)
 1151 363, arXiv:hep-ph/9502238.
- 1152 [47] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*,
 1153 Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- 1154 [48] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal
 1155 Statistical Society, Series B **58** (1994) 267.
- 1156 [49] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude*
 1157 *analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- 1158 [50] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- 1159 [51] CLEO Collaboration, M. Artuso *et al.*, *Amplitude analysis of $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$* ,
 1160 Phys. Rev. **D85** (2012) 122002, arXiv:1201.5716.

- 1161 [52] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime*
1162 *and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986,
1163 DESY-F31-86-02.
- 1164 [53] D. Hill, M. John, and P. Gandini, *A study of partially reconstructed $B^\pm \rightarrow D^{*0} h^\pm$*
1165 *decays using the $D^0 \rightarrow K\pi, KK, \pi\pi$ final states*, .
- 1166 [54] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins
1167 University Press, Baltimore, MD, USA, 1996.
- 1168 [55] L. collaboration, *Precision measurement of the B_s oscillation frequency with the decay*
1169 $B_s \rightarrow D_s\pi$, LHCb-ANA-2012-053.