

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 $2\beta + \gamma$ 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 5 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
4	Yields determination	22
4.1	Signal model	22
4.2	Background models	23
4.3	Results	24
5	Decay-time Resolution	26
5.1	Calibration for Run-I data	27
5.2	Calibration for Run-II data	28
6	Acceptance	31
6.1	MC corrections	31
6.1.1	Truth matching of simulated candidates	31
6.2	Decay-time acceptance	32
6.3	Phase space acceptance	37
7	Flavour Tagging	40
7.1	OS tagging calibration	41
7.2	SS tagging calibration	42
7.3	Tagging performance comparison between the signal and normalization channel	43
7.4	Combination of OS and SS taggers	44
8	Production and Detection Asymmetries	46
8.1	B_s Production Asymmetry	46
8.2	$K^-\pi^+$ Detection Asymmetry	47

9 Time dependent fit	50
9.1 sFit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data	50
9.2 sFit to $B_s^0 \rightarrow D_sK\pi\pi$ data	53
10 Time dependent amplitude fit	54
10.1 Signal Model Construction	54
10.2 Results	57
11 Systematic uncertainties	60
11.1 Models for B_s^0 mass distribution	60
11.1.1 Signal model	60
11.1.2 Background model	60
11.1.3 Description of misidentified background	61
11.1.4 Systematic effect on observables	61
11.2 Decay-time acceptance	62
11.2.1 Varition of knot positions	62
11.2.2 Variation of spline coefficients	62
11.3 Decay-time resolution	63
11.4 Tagging calibration	64
11.5 Summary of systematic uncertainties	64
A Stripping and Trigger cuts	68
B Details of multivariate classifier	70
C Detailed mass fits	74
D Decay-time Resolution fits	78
E Comparison of time-acceptance in subsamples	82
F Spin Amplitudes	83
G Considered Decay Chains	84
H MC corrections	85
I Data distributions	90
I.1 Comparison of signal and calibration channel	90
I.2 Comparison of Run-I and Run-II data	92
I.3 Comparison of D_s final states	94
I.4 Comparison of trigger categories	96
I.5 Comparison of B_s and B_d decays	98
References	99

1 Introduction

The weak phase γ is the least well known angle of the CKM unitary triangle. A key channel to measure γ is the time-dependent analysis of $B_s^0 \rightarrow D_s K$ decays [1, 2].

To measure the weak CKM phase $\gamma \equiv \arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$, a decay with interference between $b \rightarrow c$ and $b \rightarrow u$ transitions is needed [1]. This note present the first measurement of γ 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. To account for the non-constant strong phase across the phasespace, 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 fit. This analysis is based on the first observation of the $B_s^0 \rightarrow D_s K\pi\pi$ decay by LHCb [3, 4], where the branching ratio is measured relative to $B_s^0 \rightarrow D_s \pi\pi\pi$.

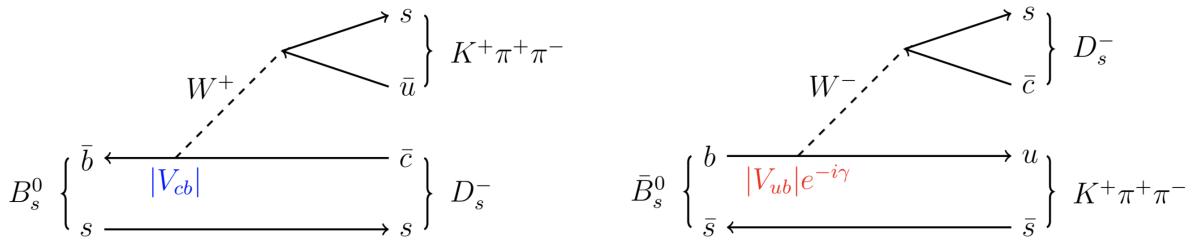


Figure 1.1: Feynman diagram for $B_s^0/\bar{B}_s^0 \rightarrow D_s^- K^+ \pi^+ \pi^-$ decays.

13 2 Formalism

14 2.1 Decay rates and *CP*-observables

15 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
 16 given by:

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

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

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

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

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

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

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

37 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,
 38 the four-body phase space density function is not flat in the usual kinematic variables.
 39 Therefore, an analytic expression for ϕ_4 is taken from Ref. [5].

⁴⁰ Assuming there is no direct CP violation in the B_s decay implies for the CP conjugate
⁴¹ transition amplitudes:

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

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

⁴² where the CP -conjugate phase space point $\bar{\mathbf{x}}$ is defined such that it is mapped onto \mathbf{x} by
⁴³ the interchange of final state charges, and the reversal of three-momenta.

⁴⁴ The phenomenological description of the intermediate state amplitudes is discussed
⁴⁵ in Sec. 2.2. For a model-independent measurement, the differential decay rate can be
⁴⁶ integrated over the phase space:

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

⁴⁷ where the same convention for the CP coefficients as for the $B_s \rightarrow D_s K$ analysis is used:

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

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

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

⁴⁸ The coherence factor κ , the strong phase difference δ and the ratio of the suppressed
⁴⁹ ($b \rightarrow u$) over favored ($b \rightarrow c$) decay mode, averaged over the phase space, are defined as:

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

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

⁵⁰ The coherence factor dilutes the sensitivity to the weak phase γ due to the integration
⁵¹ over the interfering amplitudes across the phase space. The value of κ is bounded between
⁵² zero and unity. The latter corresponds to the limit of only one contributing intermediate
⁵³ state in which case the same sensitivity as in $B_s \rightarrow D_s K$ decays is reached, while $\kappa = 0$
⁵⁴ would result in no sensitivity to γ at all.

55 2.2 Amplitude model

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

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

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

66 2.2.1 Form Factors and Resonance Lineshapes

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

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

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

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

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

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

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

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

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

¹¹⁸ **2.2.2 Spin Densities**

¹¹⁹ The spin amplitudes are phenomenological descriptions of decay processes that are required
¹²⁰ to be Lorentz invariant, compatible with angular momentum conservation and, where
¹²¹ appropriate, parity conservation. They are constructed in the covariant Zemach (Rarita-
¹²² Schwinger) tensor formalism [22–24]. At this point, we briefly introduce the fundamental
¹²³ objects of the covariant tensor formalism which connect the particle’s four-momenta to
¹²⁴ the spin dynamics of the reaction and give a general recipe to calculate the spin factors
¹²⁵ for arbitrary decay trees. Further details can be found in Refs. [25, 26].

¹²⁶ A spin- S particle with four-momentum p , and spin projection λ , is represented by the
¹²⁷ polarization tensor $\epsilon_{(S)}(p, \lambda)$, which is symmetric, traceless and orthogonal to p . These
¹²⁸ so-called Rarita-Schwinger conditions reduce the a priori 4^S elements of the rank- S tensor
¹²⁹ to $2S + 1$ independent elements in accordance with the number of degrees of freedom of a
¹³⁰ spin- S state [23, 27].

¹³¹ The spin projection operator $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$, for a resonance R , with spin $S =$
¹³² $\{0, 1, 2\}$, and four-momentum p_R , is given by [26]:

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

¹³³ where $g^{\mu\nu}$ is the Minkowski metric. Contracted with an arbitrary tensor, the projection
¹³⁴ operator selects the part of the tensor which satisfies the Rarita-Schwinger conditions.

¹³⁵ For a decay process $R \rightarrow AB$, with relative orbital angular momentum L , between
¹³⁶ particle A and B , the angular momentum tensor is obtained by projecting the rank- L
¹³⁷ 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
¹³⁸ subspace,

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

¹³⁹ Their $|\vec{q}_R|^L$ dependence accounts for the influence of the centrifugal barrier on the transition
¹⁴⁰ amplitudes. For the sake of brevity, the following notation is introduced,

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

¹⁴¹ Following the isobar approach, a four-body decay amplitude is described as a product
¹⁴² of two-body decay amplitudes. Each sequential two-body decay $R \rightarrow A B$, with relative
¹⁴³ orbital angular momentum L_{AB} , and total intrinsic spin S_{AB} , contributes a term to the
¹⁴⁴ overall spin factor given by

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

¹⁴⁵ where

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

¹⁴⁶ Here, a polarization vector is assigned to the decaying particle and the complex conjugate
¹⁴⁷ vectors for each decay product. The spin and orbital angular momentum couplings are
¹⁴⁸ described by the tensors $P_{(S_{AB})}(R)$ and $L_{(L_{AB})}(R)$, respectively. Firstly, the two spins S_A
¹⁴⁹ and S_B , are coupled to a total spin- S_{AB} state, $\Phi(\mathbf{x}|S_{AB})$, by projecting the corresponding
¹⁵⁰ polarization vectors onto the spin- S_{AB} subspace transverse to the momentum of the
¹⁵¹ decaying particle. Afterwards, the spin and orbital angular momentum tensors are
¹⁵² properly contracted with the polarization vector of the decaying particle to give a Lorentz
¹⁵³ scalar. This requires in some cases to include the tensor $\varepsilon_{\alpha\beta\gamma\delta} p_R^\delta$ via

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

¹⁵⁴ where $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita symbol and j refers to the arguments of X defined in
¹⁵⁵ Eqs. 2.23 and 2.24. Its antisymmetric nature ensures the correct parity transformation
¹⁵⁶ behavior of the amplitude. The spin factor for a whole decay chain, for example $R \rightarrow$
¹⁵⁷ $(R_1 \rightarrow AB)(R_2 \rightarrow CD)$, is obtained by combining the two-body terms and performing a
¹⁵⁸ sum over all unobservable, intermediary spin projections

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

¹⁵⁹ 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
¹⁶⁰ pseudoscalar initial/final states are involved.

¹⁶¹ The spin factors for all decay topologies considered in this analysis are explicitly given
¹⁶² in Appendix F.

163 2.3 Fit implementation

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

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

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

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

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

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

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

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

187 2.4 Validation

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

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

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

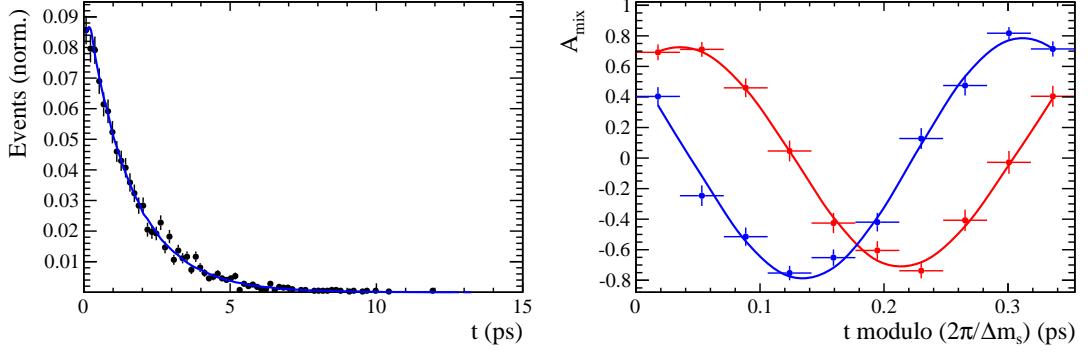


Figure 2.1: Left: Time distribution of $B_s \rightarrow D_s K \pi \pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (solid line). Right: Time-dependent asymmetry between mixed and unmixed events folded into one oscillation period for $D_s^- K^+ \pi \pi$ (red) and $D_s^+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

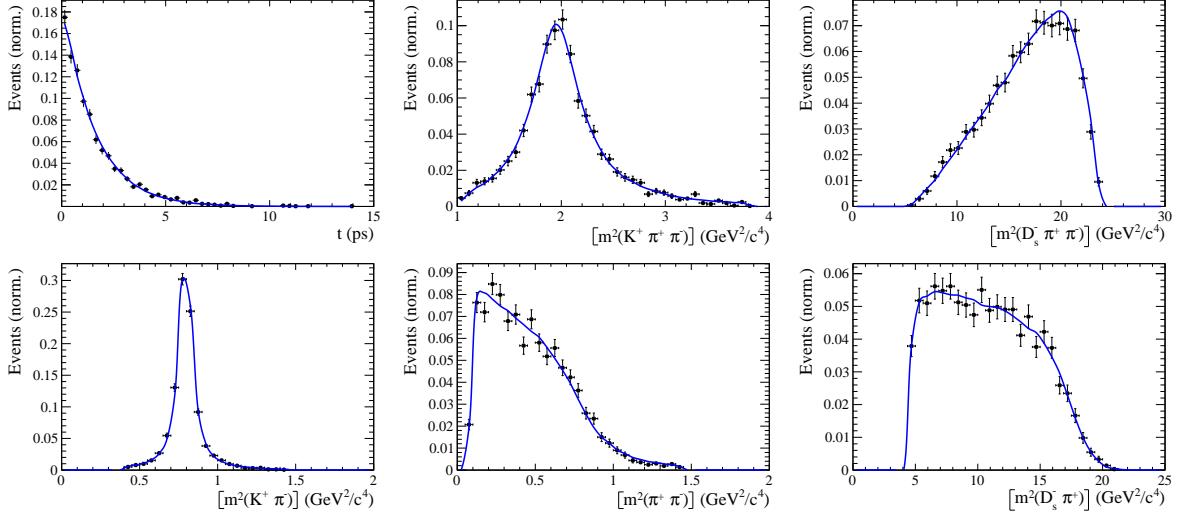


Figure 2.2: Time and invariant mass distributions of $B_s \rightarrow D_s K \pi \pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

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

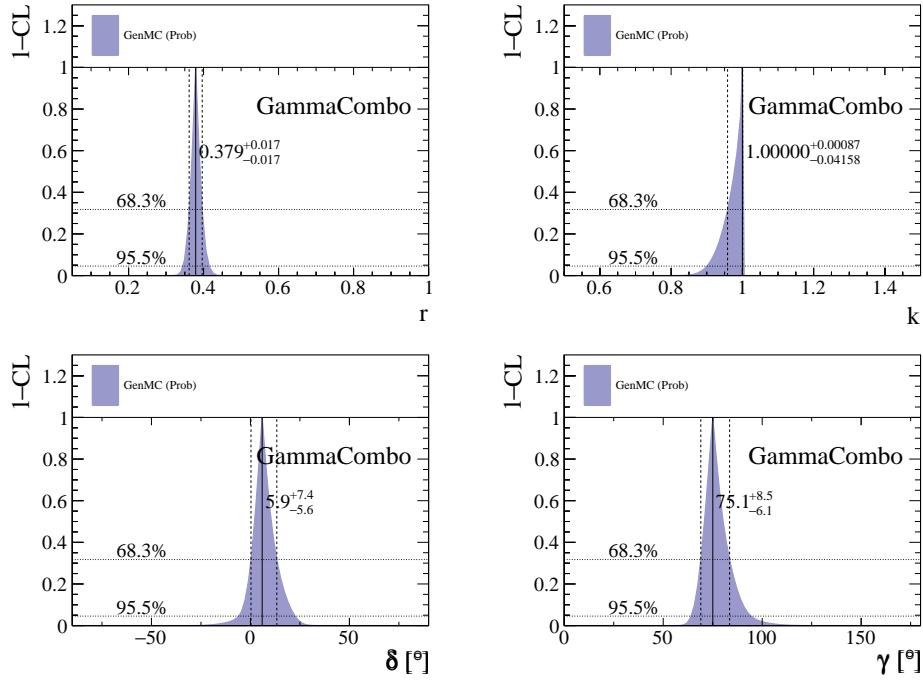


Figure 2.3: The 1-CL contours for the physical observable r, κ, δ and γ obtained with the phasespace-integrated fit to the EVTGEN toy sample.

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

211 3 Data samples and event selection

212 3.1 Stripping and Trigger selection

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

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

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

233 3.2 Offline selection

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

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

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

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

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

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

269 3.2.1 Phase space region

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

281 **3.2.2 Physics background vetoes**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

- 316 1. $X_s^+ \rightarrow K^+\pi^+\pi^-$:
- 317 (a) $B_s^0 \rightarrow D_s\pi\pi\pi$:
318 Possible with $\pi^+ \rightarrow K^+$ misidentification, suppressed with $\text{PIDK}(K^+) > 10$.
- 319 (b) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^+\pi^-)$:
320 Outside of considered phase-space region, see Sec. 3.2.1.
- 321 (c) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+K^-\pi^+)$:
322 To suppress $B_s^0 \rightarrow D_s^-K^+K^-\pi^+$ background, possible with $K^- \rightarrow \pi^-$ misiden-
323 tification, we require $\text{PIDK}(\pi^-) < 0$. In case the invariant mass of the $K^+\pi^+\pi^-$
324 system recomputed applying the kaon mass hypothesis to the pion is close to
325 the D_s mass ($m(K^+\pi^+\pi_K^-) = m(D_s) \pm 20$ MeV), we further tighten the cut to
326 $\text{PIDK}(\pi^-) < -5$.
- 327 2. $X_d^+ \rightarrow \pi^+\pi^+\pi^-$:
- 328 (a) $B_s^0 \rightarrow D_s K\pi\pi$:
329 Possible with single missID of $K^+ \rightarrow \pi^+$, suppressed with $\text{PIDK}(\pi^+) < 0$.
- 330 (b) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow \pi^+\pi^+\pi^-)$:
331 Outside of considered phase-space region, see Sec. 3.2.1.
- 332 (c) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^-\pi^+)$:
333 Possible with single missID of $K^+ \rightarrow \pi^+$, vetoed by requiring $m(\pi^+\pi_K^+\pi^-) \neq$
334 $m(D_s) \pm 20$ MeV or $\text{PIDK}(\pi^+) < -5$ for both π^+ .

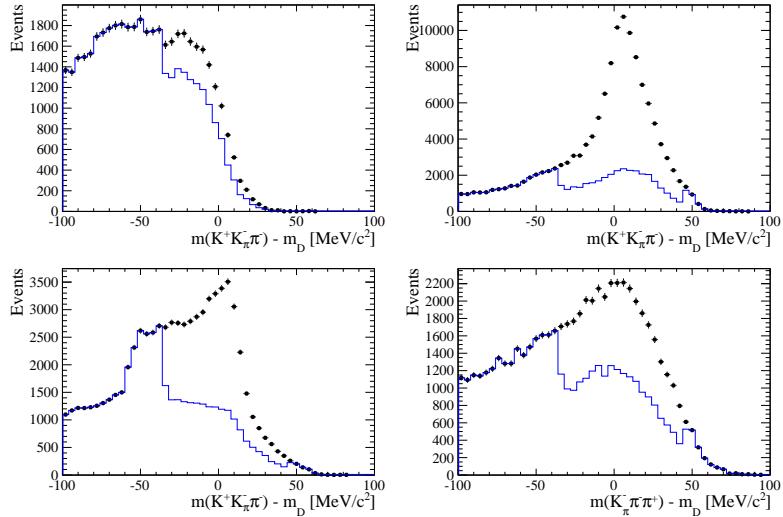


Figure 3.1: Background contributions from D^- decays where the π^- is misidentified as K^- . The D_s invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*(892)K$, $D_s \rightarrow KK\pi$ (non-resonant) and $D_s \rightarrow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

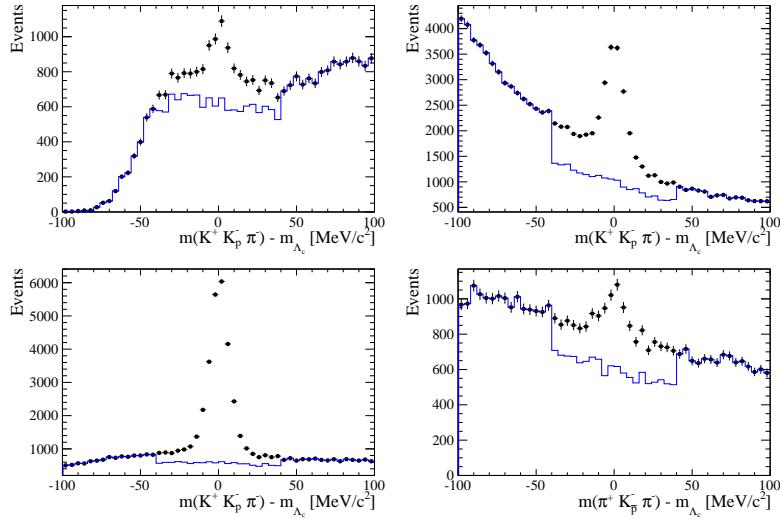


Figure 3.2: Background contributions from Λ_c decays where the \bar{p} is misidentified as K^- . The D_s invariant mass is recomputed applying the proton mass hypothesis to the kaon and shown for the $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*(892)K$, $D_s \rightarrow KK\pi$ (non-resonant) and $D_s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

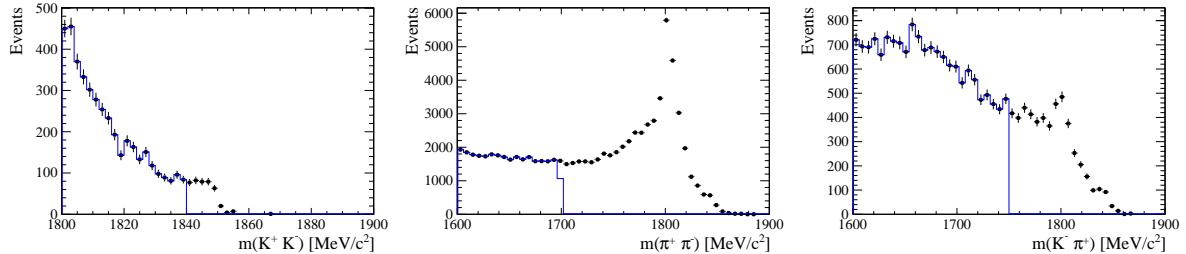


Figure 3.3: Background contributions to $D_s \rightarrow KK\pi$ (left), $D_s \rightarrow \pi\pi\pi$ (middle) and $D_s \rightarrow K\pi\pi$ (right) from $D^0 \rightarrow hh$ decays combined with a random pion.

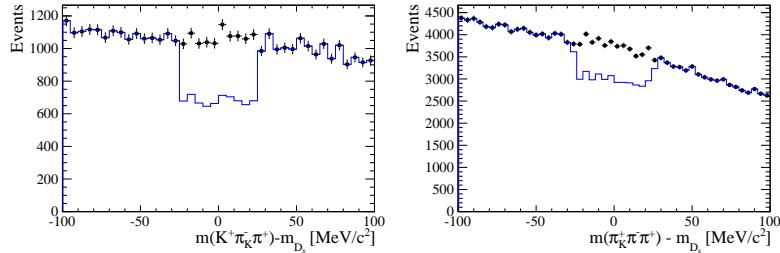


Figure 3.4: Background contributions to $B_s \rightarrow D_s K\pi\pi$ (left) and $B_s \rightarrow D_s \pi\pi\pi$ (right) from $B_s \rightarrow D_s D_s$ decays where the kaon is misidentified as pion. The $X_{s,d}$ invariant mass is recomputed applying the kaon mass hypothesis to the pion and shown without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

335 **3.2.3 Training of multivariate classifier**

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

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

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

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

362 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
 363 to the training which are expected to be 100% signal efficient. Figure 3.6 shows the
 364 distributions of the input variables for signal and background. As shown in Appendix B,
 365 these distributions differ between data-taking period and trigger category. In particular
 366 variables depending on the B_s kinematics and the event multiplicity are affected (e.g.
 367 $\theta_{D_s h_i}$ or A_{pT}^{cone}). The BDTG is consequently trained separately for these categories. The
 368 resulting classifier response is shown in Fig. 3.7 for each category (even and odd test
 369 samples combined) and in Appendix B for each training.

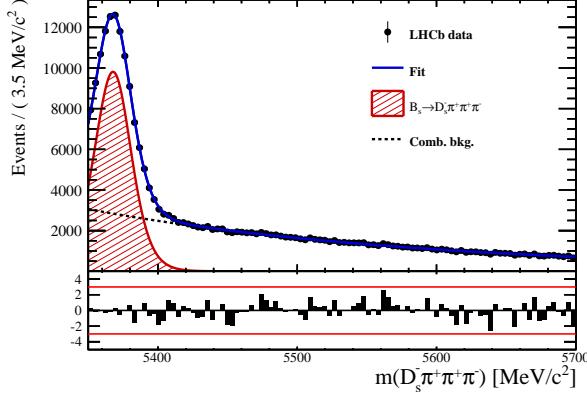


Figure 3.5: Reconstructed B_s mass for $B_s \rightarrow D_s \pi\pi\pi$ events that pass the preselection. The fitted PDF is shown in blue, the signal component in red and the background component in black.

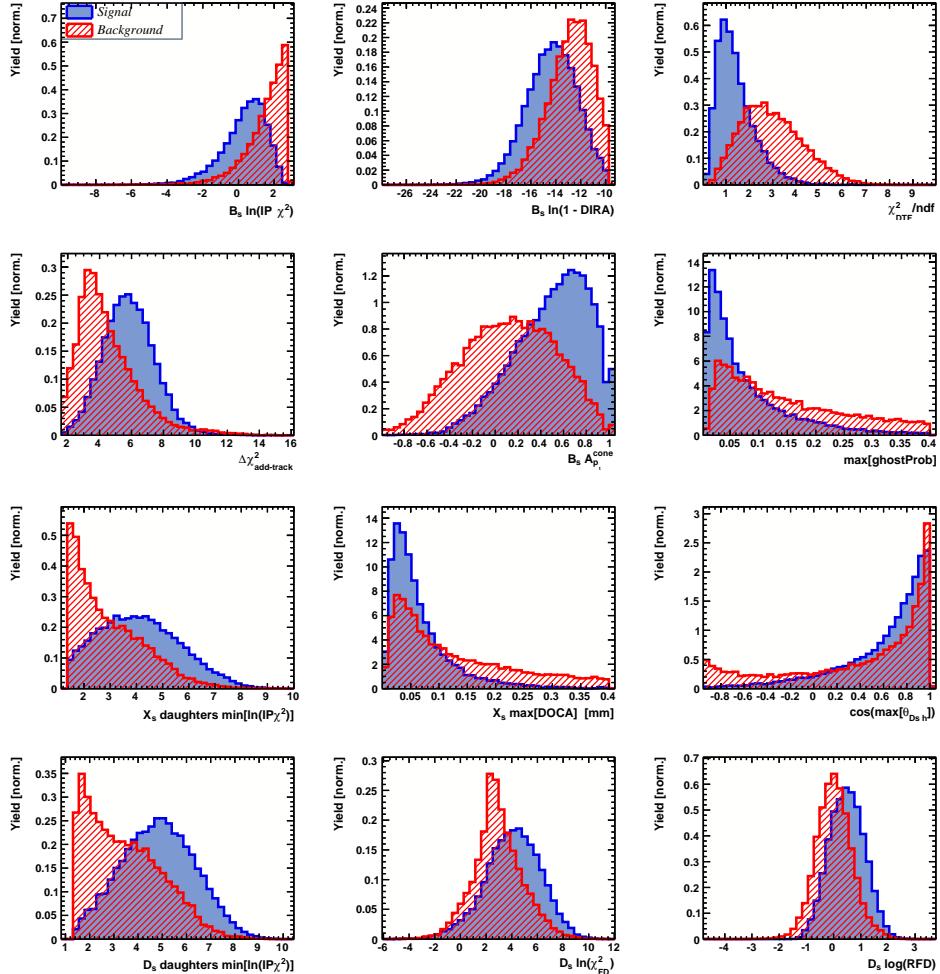


Figure 3.6: Discriminating variables used to train the BDTG for all data categories combined.

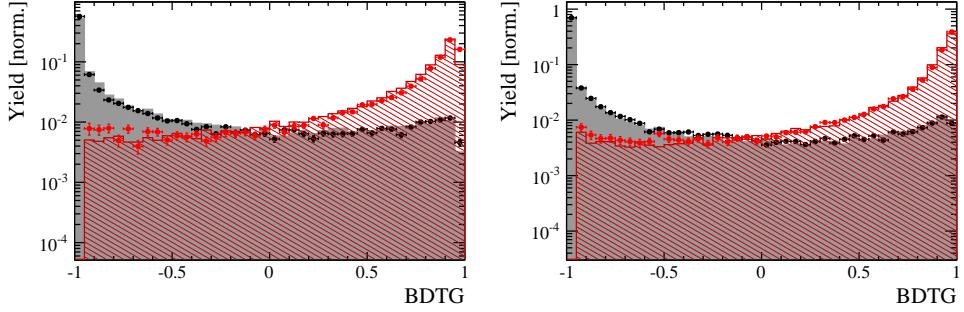


Figure 3.7: Signal (red) and background (black) distributions for the BDTG response for Run-I (left) and Run-II (right) data. Filled histograms (data points) show the BDTG response for the L0-TOS (L0-TIS) category. Even and odd test samples are combined.

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.

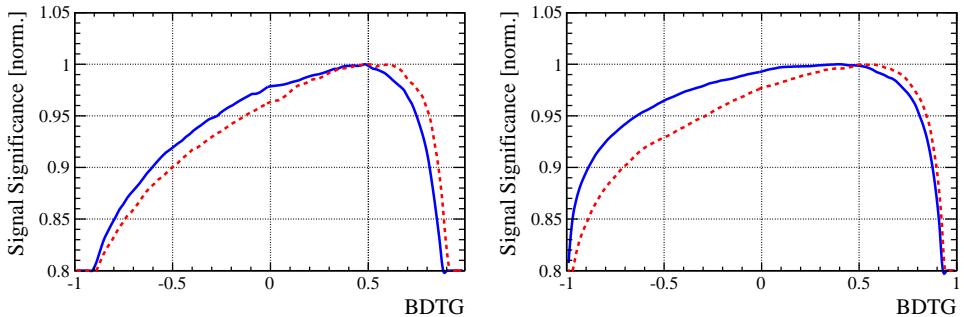


Figure 3.8: Signal significance as function of the applied cut on the BDTG response for Run-I (left) and Run-II (right) data. The scans for the L0-TOS (L0-TIS) category are shown in blue (red). The signal significance is normalized to be 1 at the optimal BDTG cut value.

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	$= 1$

385 4 Yields determination

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

391 Due to different mass resolutions, we perform the invariant mass fits simultaneously
 392 for each data-taking period and each trigger category. We further introduce four D_s final
 393 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
 394 different signal purities. The $D_s \rightarrow Kh\pi$ category combines the two D_s decay channels
 395 with the lowest statistics: $D_s \rightarrow KK\pi$ (non-resonant) and $D_s \rightarrow K\pi\pi$. This amounts to
 396 16 categories in total.

397 4.1 Signal model

398 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
 399 using a Johnson's SU function [32], which results from a variable transformation of a
 400 normal distribution to allow for asymmetric tails:

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

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

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

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

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

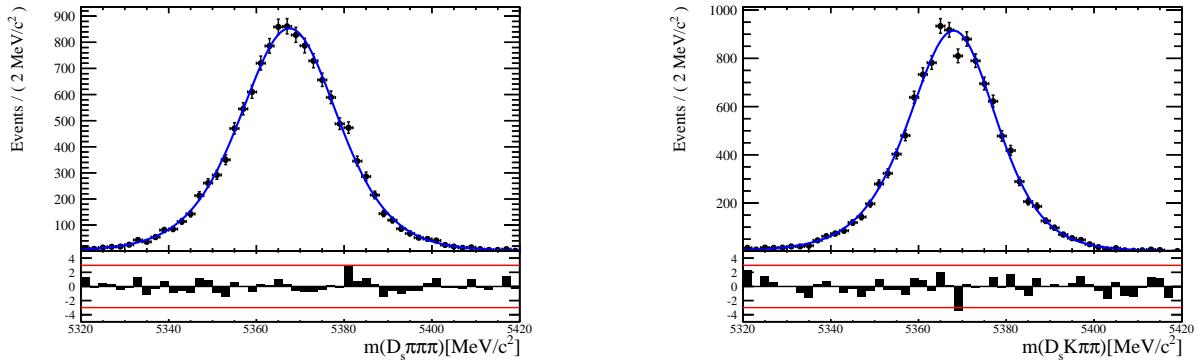


Figure 4.1: Invariant mass distributions of simulated (left) $B_s^0 \rightarrow D_s \pi\pi\pi$ and (right) $B_s^0 \rightarrow D_s K\pi\pi$ events. A fit with a Johnson's SU PDF is overlaid.

405 simulation and real data, linear scaling factors for the mean μ and width σ are determined
406 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
407 factors are determined separately for each data-taking period and each trigger category.

408 4.2 Background models

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

411 Combinatorial background

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

415 Peaking B_d background

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

418 Partially reconstructed background

419 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
420 to be peaking lower than signal in the $m(D_s\pi\pi\pi)$ spectrum with large tails due to the
421 momentum carried away by the not reconstructed π^0 or γ . An empirical description for
422 the shape of this contribution is derived from a $B_s^0 \rightarrow D_s^*\pi\pi\pi$ MC sample subject to
423 the nominal $B_s^0 \rightarrow D_s\pi\pi\pi$ selection. Figure 4.2 (left) shows the respective reconstructed
424 $m(D_s\pi\pi\pi)$ distribution. A sum of three bifurcated Gaussian functions is used to describe
425 it. In the fit to data, all parameters are fixed to the ones obtained from MC except for
426 the parameter which describes the width of the right tail of the distribution to account for
427 data-simulation differences in mass resolution. The equivalent $B_s^0 \rightarrow D_s^*K\pi\pi$ component
428 contributing to the $B_s^0 \rightarrow D_sK\pi\pi$ data sample is described by the same PDF with the
429 right tail fixed to the $B_s^0 \rightarrow D_s\pi\pi\pi$ result.

430 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
431 shifted by $m_{B_s^0} - m_{B^0}$.

432 Misidentified background

433 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
434 pions is misidentified as a kaon, contaminate the $B_s^0 \rightarrow D_sK^+\pi^+\pi^-$ sample. To determine
435 the corresponding background shapes, we use simulated events passing the nominal
436 selection except for the PID cuts on the bachelor π^+ tracks. The **PIDCalib** package
437 is used to determine the p_T, η -dependent $\pi^+ \rightarrow K^+$ misidentification probability for
438 each pion. We change the particle hypothesis from pion to kaon for the pion with the
439 higher misidentification probability and recompute the invariant B_s^0 mass, $m(D_s^-\pi_K^+\pi^+\pi^-)$.
440 Similarly, the invariant masses $m(\pi_K^+\pi^+\pi^-)$ and $m(\pi_K^+\pi^-)$ are recomputed and required
441 to be within the considered phasespace region. The background distributions are shown
442 in Fig. 4.2 (middle,right) and modeled by the sum of two Crystal Ball functions. The
443 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
444 $B_s^0 \rightarrow D_sK\pi\pi$ sample is computed by multiplying the fake rate (within the considered
445 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

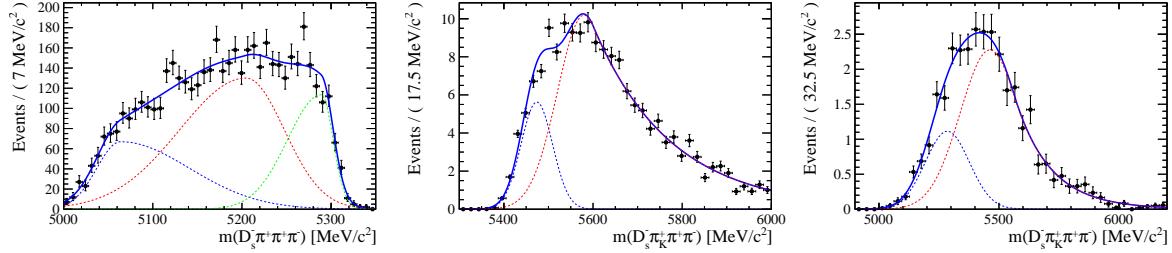


Figure 4.2: Left: Invariant mass distribution of simulated $B_s^0 \rightarrow D_s^* \pi\pi\pi$ events, where the γ/π^0 is excluded from the reconstruction. Middle: Invariant mass distribution of simulated $B_s^0 \rightarrow D_s \pi\pi\pi$ events, where one of the pions is reconstructed as a kaon taking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

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

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

4.3 Results

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

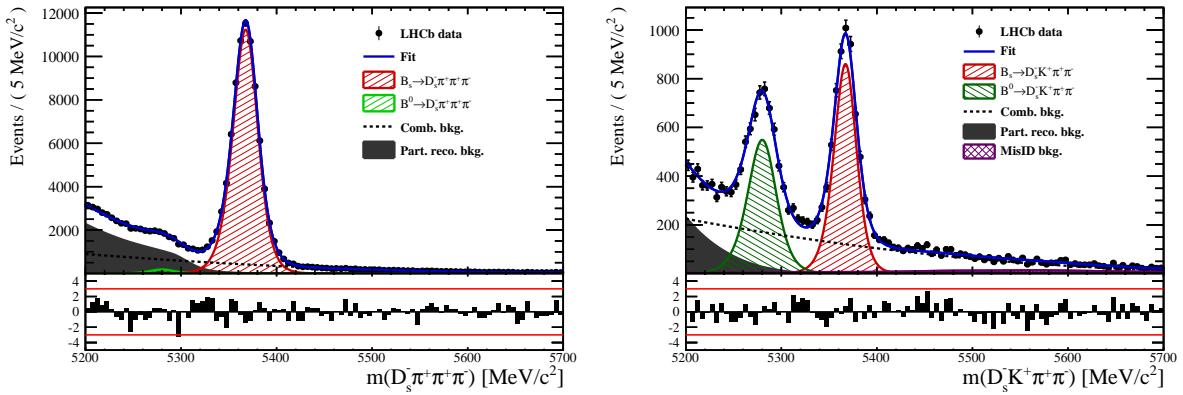


Figure 4.3: Invariant mass distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ (left) and $B_s^0 \rightarrow D_s K\pi\pi$ (right) candidates.

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		
$B_s \rightarrow D_s\pi\pi\pi$	77225 ± 304		
$B^0 \rightarrow D_s\pi\pi\pi$	1263 ± 454		
Partially reconstructed bkg.	31805 ± 351		
Combinatorial bkg.	32821 ± 393		
D_s final state	Signal yield		
$D_s^- \rightarrow \phi^0(1020)\pi^-$	26458 ± 172		
$D_s^- \rightarrow K^{*0}(892)K^-$	23105 ± 170		
$D_s^- \rightarrow (K^-h^+\pi^-)$	15201 ± 136		
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	12461 ± 122		

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		
$B_s \rightarrow D_sK\pi\pi$	5376 ± 88		
$B^0 \rightarrow D_sK\pi\pi$	4384 ± 101		
Partially reconstructed bkg.	1796 ± 96		
Misidentified bkg.	808 ± 0		
Combinatorial bkg.	9376 ± 177		
D_s final state	Signal yield		
$D_s^- \rightarrow \phi^0(1020)\pi^-$	1706 ± 49		
$D_s^- \rightarrow K^{*0}(892)K^-$	1712 ± 49		
$D_s^- \rightarrow (K^-h^+\pi^-)$	1145 ± 41		
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	814 ± 36		

462 5 Decay-time Resolution

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

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

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

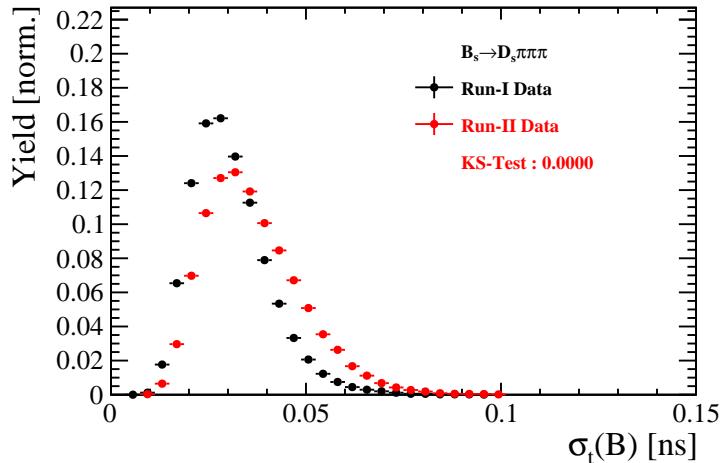


Figure 5.1: Distribution of the decay time error for $B_s^0 \rightarrow D_s\pi\pi\pi$ signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

485 5.1 Calibration for Run-I data

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

$$494 \quad \mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s^2 / 2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s^2 / 2}, \quad (5.1)$$

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

$$497 \quad \sigma_{eff} = \sqrt{(-2/\Delta m_s^2) \ln \mathcal{D}}, \quad (5.2)$$

497 which converts the resolution into a single-Gaussian function with an effective resolution
 498 that causes the same damping effect on the magnitude of the B_s oscillation. For the Run-I
 499 $B_s^0 \rightarrow D_s K \pi\pi$ MC sample the effective average resolution is found to be $\sigma_{eff} = 39.1 \pm 0.3$ fs.

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

$$508 \quad \sigma_{eff}^{MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.232 \pm 0.010) \sigma_t \quad (5.3)$$

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

$$509 \quad \sigma_{eff}^{D_s K, MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.201 \pm 0.013) \sigma_t. \quad (5.4)$$

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

$$512 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (10.26 \pm 1.52) \text{ fs} + (1.280 \pm 0.042) \sigma_t \quad (5.5)$$

512 can be used for our analysis. The following calibration functions were used in the
 513 $B_s^0 \rightarrow D_s K$ analysis to estimate the systematic uncertainty on the decay-time resolution:

$$514 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (-0.568 \pm 1.570) \text{ fs} + (1.243 \pm 0.044) \sigma_t \quad (5.6)$$

$$514 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.772 \pm 0.012) \sigma_t \quad (5.7)$$

515 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
 516 assigned as additional systematic uncertainty.

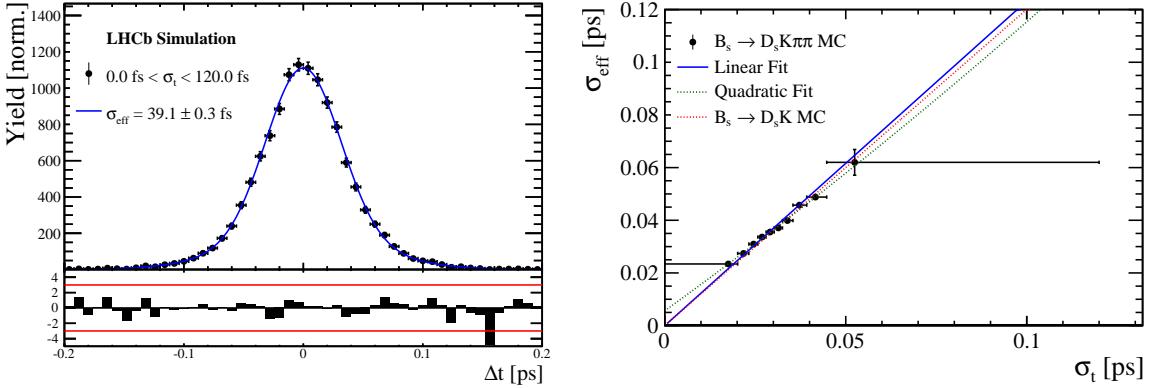


Figure 5.2: (Left) Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K\pi\pi$ MC candidates. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for $B_s \rightarrow D_s K\pi\pi$ MC (Run-I). The fitted calibration curve is shown in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

517 5.2 Calibration for Run-II data

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

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

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

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

⁵⁴⁴ is in good agreement with the one obtained for the $B_s \rightarrow J/\psi\phi$ (Run-II) analysis [36].

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

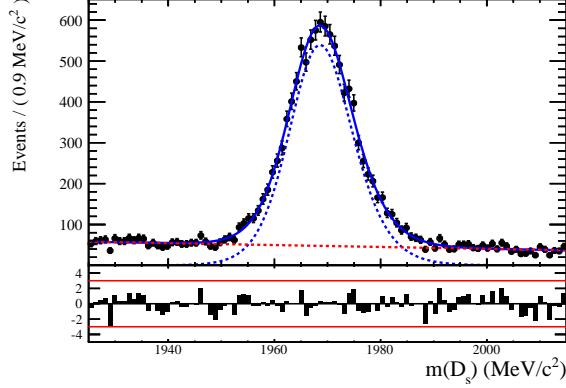


Figure 5.3: The invariant mass distribution for prompt D_s candidates.

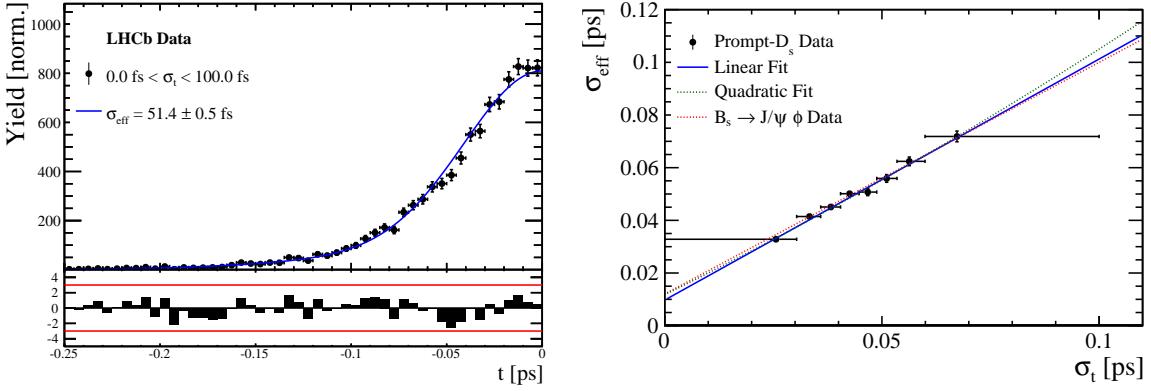


Figure 5.4: (Left) Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data). The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

545 **6 Acceptance**

546 **6.1 MC corrections**

547 **6.1.1 Truth matching of simulated candidates**

548 We use the `BackgroundCategory` tool to truth match our simulated candidates. Candidates
 549 with background category (`BKGCAT`) 20 or 50 are considered to be true signal. Background
 550 category 60 is more peculiar since it contains both badly reconstructed signal candidates
 551 and ghost background. This is due to the fact that the classification algorithms identifies
 552 all tracks for which less than 70% of the reconstructed hits are matched to generated
 553 truth-level quantities as ghosts. In particular for signal decays involving many tracks (as
 554 in our case), this arbitrary cutoff is not 100% efficient. Moreover, the efficiency is expected
 555 to depend on the kinematics which would lead to a biased acceptance determination if
 556 candidates with `BKGCAT`= 60 would be removed.

557 We therefore include `BKGCAT`= 60 and statistically subtract the ghost background by
 558 using the `sPlot` technique. The `sWeights` are calculated from a fit to the reconstructed B_s
 559 mass. The signal contribution is modeled as described in Sec. 4.1 and the background
 560 with a polynomial. The fit is performed simultaneously in two categories; the first includes
 561 events with `BKGCAT` = 20 or 50 and the second events with `BKGCAT` = 60. To account
 562 for the different mass resolution we use a different σ for each category, while the mean
 563 and the tail parameters are shared between them. The background component is only
 564 included for the second category.

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

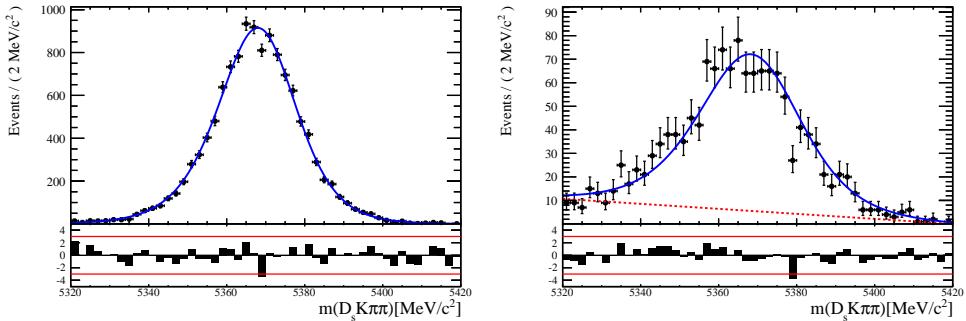


Figure 6.1: The reconstructed B_s mass distribution for simulated $B_s \rightarrow D_s K\pi\pi$ decays
 classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

567 **6.2 Decay-time acceptance**

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

573 We use a combination of control channels to derive the acceptance function $\epsilon(t)$,
 574 because for $B_s^0 \rightarrow D_s K\pi\pi$ decays the decay-time acceptance is strongly correlated with
 575 the CP -observables which we aim to measure. Therefore, extracting the CP -observables
 576 and the acceptance shape at the same time is not possible. A fit to the decay-time
 577 distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates is performed and the obtained acceptance shape
 578 is corrected for the small difference observed between the $B_s^0 \rightarrow D_s K\pi\pi$ and $B_s^0 \rightarrow D_s \pi\pi\pi$
 579 MC samples. In addition, we include the control channel $B^0 \rightarrow D_s K\pi\pi$ to increase
 580 the statistical precision. A simultaneous fit to the four datasets ($B_s^0 \rightarrow D_s \pi\pi\pi$ data,
 581 $B^0 \rightarrow D_s K\pi\pi$ 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
 582 a straightforward propagation of uncertainties. In each case, a PDF of the following form

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

583 is used to describe the decay-time distribution. For real data, the values for $\Gamma_{s,d}$ and
 584 $\Delta\Gamma_{s,d}$ are fixed to the latest HFAG results [37], while for simulated data, the generated
 585 values are used. A single Gaussian resolution function $\mathcal{R}(t - t', \delta t)$ is used where the
 586 decay-time error estimate is scaled with the respective calibration functions determined in
 587 Sec. 5. Each decay-time acceptance $\epsilon(t)$ is modeled using cubic splines, allowing for the
 588 analytical computation of the decay-time integrals appearing in the PDF [38]. The splines
 589 are parametrized by so-called knots (t_0, t_1, \dots, t_N) which determine their boundaries. Two
 590 knots are located by default at the lower and upper edge of the interval allowed for the
 591 decay time, the remaining ones are chosen such that there is an approximately equal
 592 amount of data in-between two consecutive knots. In the basis of cubic b-splines, $b_i(t)$,
 593 the acceptance is then constructed as:

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

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

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

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

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

- 600 • $B_s^0 \rightarrow D_s K\pi\pi$ MC: $\epsilon_{D_s K\pi\pi}^{MC}(t)$
601 • $B_s^0 \rightarrow D_s \pi\pi\pi$ MC: $\epsilon_{D_s \pi\pi\pi}^{MC}(t) = R(t) \cdot \epsilon_{D_s K\pi\pi}^{MC}(t)$
602 • $B_s^0 \rightarrow D_s \pi\pi\pi$ data: $\epsilon_{D_s \pi\pi\pi}^{Data}(t) = R(t) \cdot \epsilon_{D_s K\pi\pi}^{Data}(t)$
603 • $B^0 \rightarrow D_s K\pi\pi$ data: $\epsilon_{D_s K\pi\pi}^{Data}(t)$

604 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
605 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
606 $B_s^0 \rightarrow D_s K\pi\pi$ data is represented by $\epsilon_{D_s K\pi\pi}^{Data}(t)$.

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

610

611

612 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)

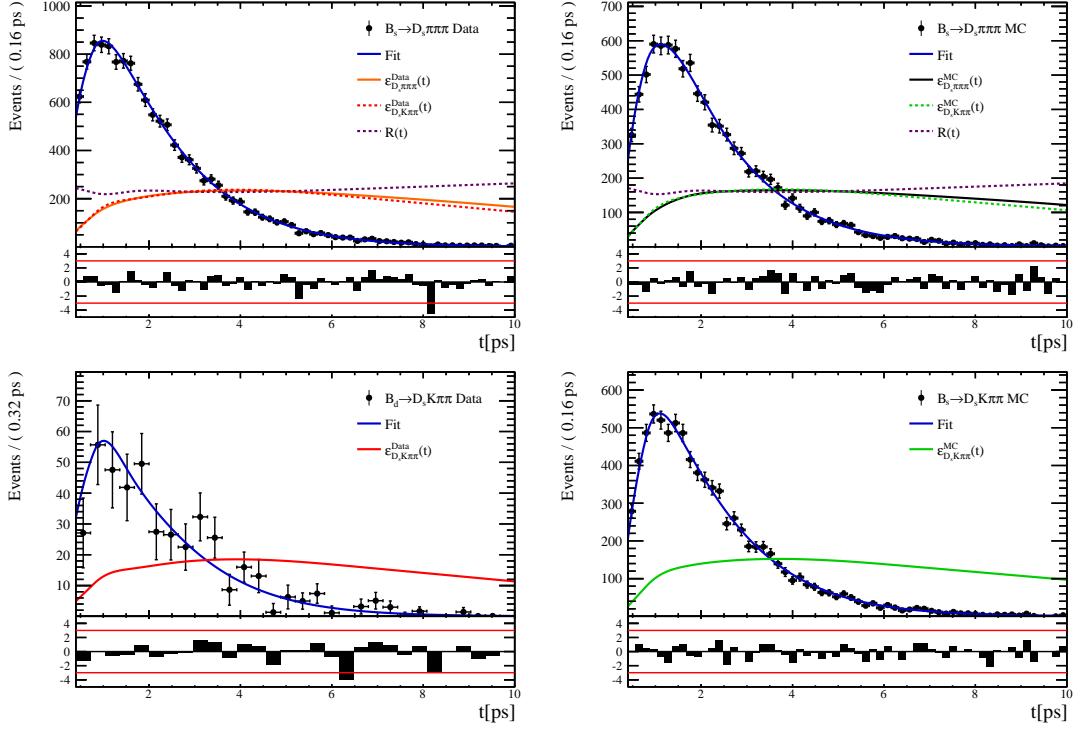


Figure 6.2: Decay-time fit projections for $B_s^0 \rightarrow D_s\pi\pi\pi$ data (up-left), $B_s^0 \rightarrow D_s\pi\pi\pi$ MC (up-right), $B_s^0 \rightarrow D_sK\pi\pi$ data (down-left) and $B_s^0 \rightarrow D_sK\pi\pi$ MC (down-right) in category [Run-I,L0-TOS]. The respective acceptance function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

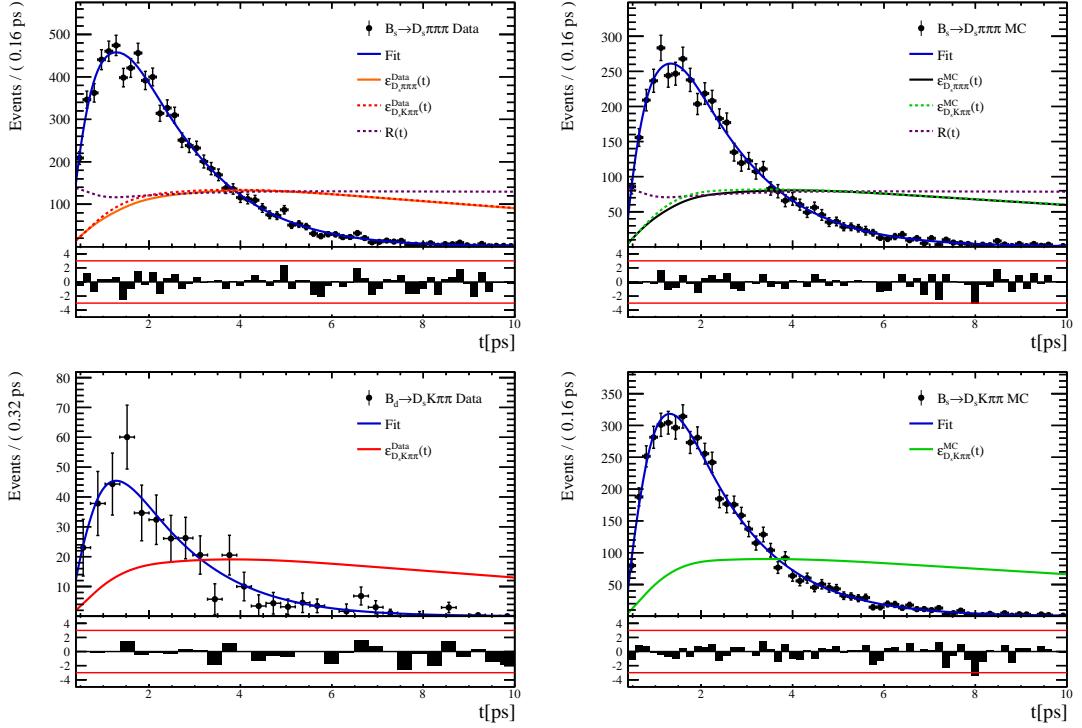


Figure 6.3: Decay-time fit projections for $B_s^0 \rightarrow D_s\pi\pi\pi$ data (up-left), $B_s^0 \rightarrow D_s\pi\pi\pi$ MC (up-right), $B_s^0 \rightarrow D_sK\pi\pi$ data (down-left) and $B_s^0 \rightarrow D_sK\pi\pi$ MC (down-right) in category [Run-I,L0-TIS]. The respective acceptance function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

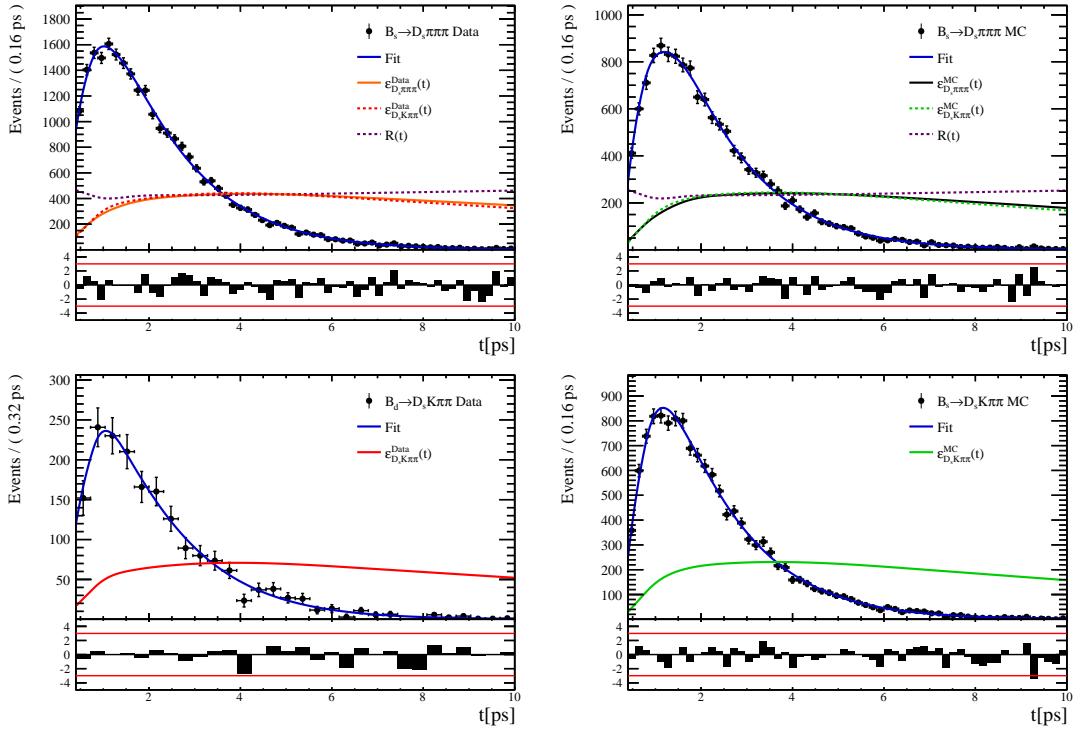


Figure 6.4: Decay-time fit projections for $B_s^0 \rightarrow D_s\pi\pi\pi$ data (up-left), $B_s^0 \rightarrow D_s\pi\pi\pi$ MC (up-right), $B^0 \rightarrow D_sK\pi\pi$ data (down-left) and $B_s^0 \rightarrow D_sK\pi\pi$ MC (down-right) in category [Run-II,L0-TOS]. The respective acceptance function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

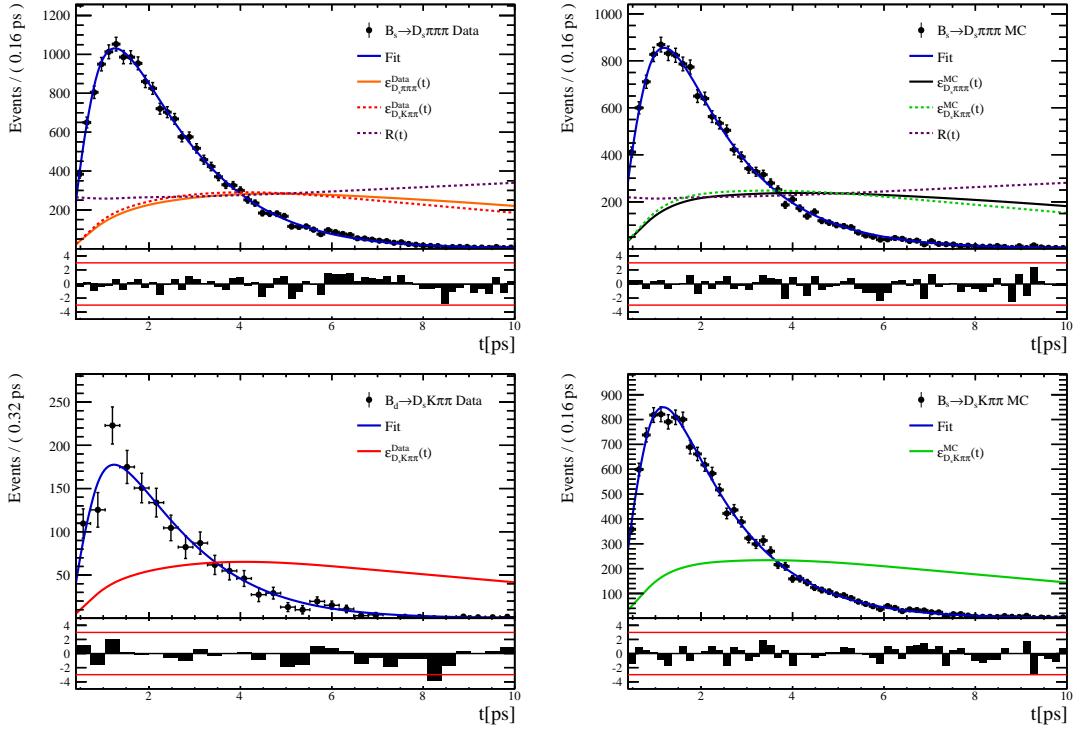


Figure 6.5: Decay-time fit projections for $B_s^0 \rightarrow D_s\pi\pi\pi$ data (up-left), $B_s^0 \rightarrow D_s\pi\pi\pi$ MC (up-right), $B^0 \rightarrow D_sK\pi\pi$ data (down-left) and $B_s^0 \rightarrow D_sK\pi\pi$ MC (down-right) in category [Run-II,L0-TIS]. The respective acceptance function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

613 6.3 Phase space acceptance

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

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

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

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

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

627

628

Disclaimer: At the moment there is only a small Run-I MC sample available where
 a DecFile (EventType:) 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.

629

630 We use a BDTG to map the five-dimensional phase space to an one-dimensional distribution
 631 [39]. The BDTG is trained to learn the differences between the selected MC and a generator
 632 level MC sample. As discriminating variables, five invariant mass combinations are used
 633 as shown in Fig. 6.6. Based on the classifier output distributions, shown in Fig. 6.7, an
 634 efficiency as function of the BDTG response is derived.

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

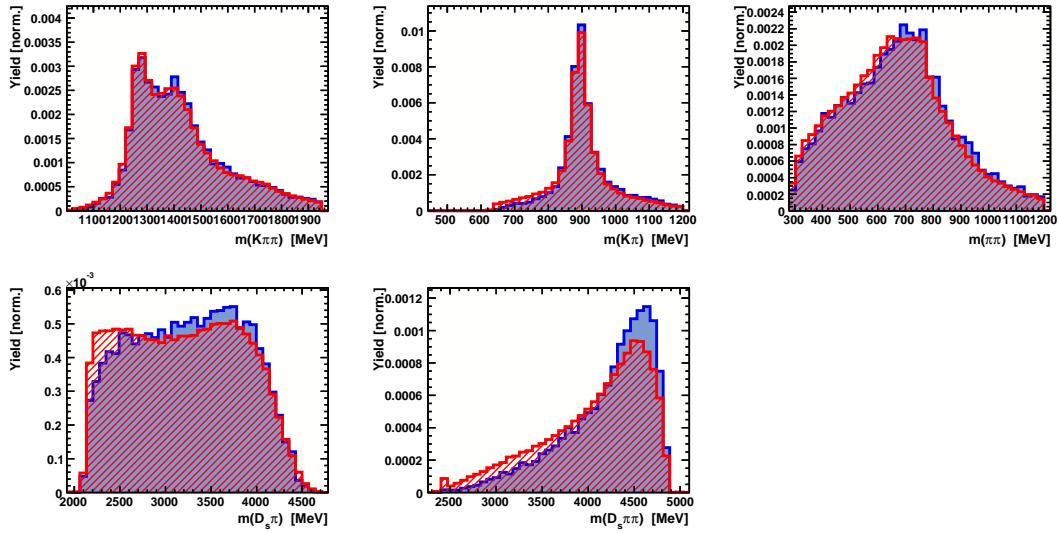


Figure 6.6: Discriminating variables used to train the BDTG. The selected MC sample is shown in blue and the generator MC sample in red.

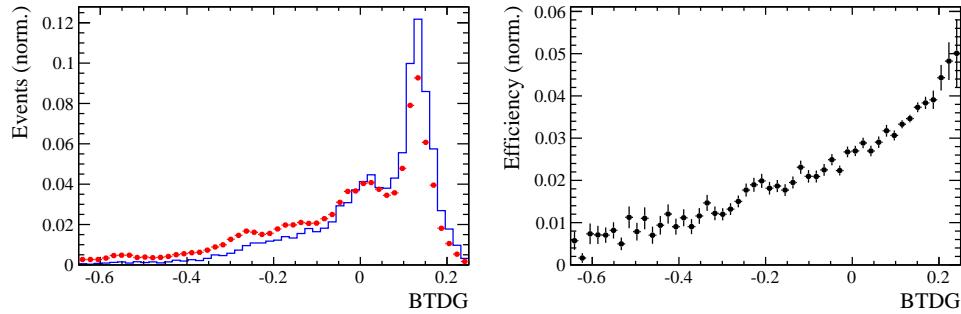


Figure 6.7: Left: Output distributions of the BDTG for the simulated MC sample (blue) and the generator level sample (red). Right: Phase space efficiency as function of the BDTG response computed as the ratio of selected and generated decays.

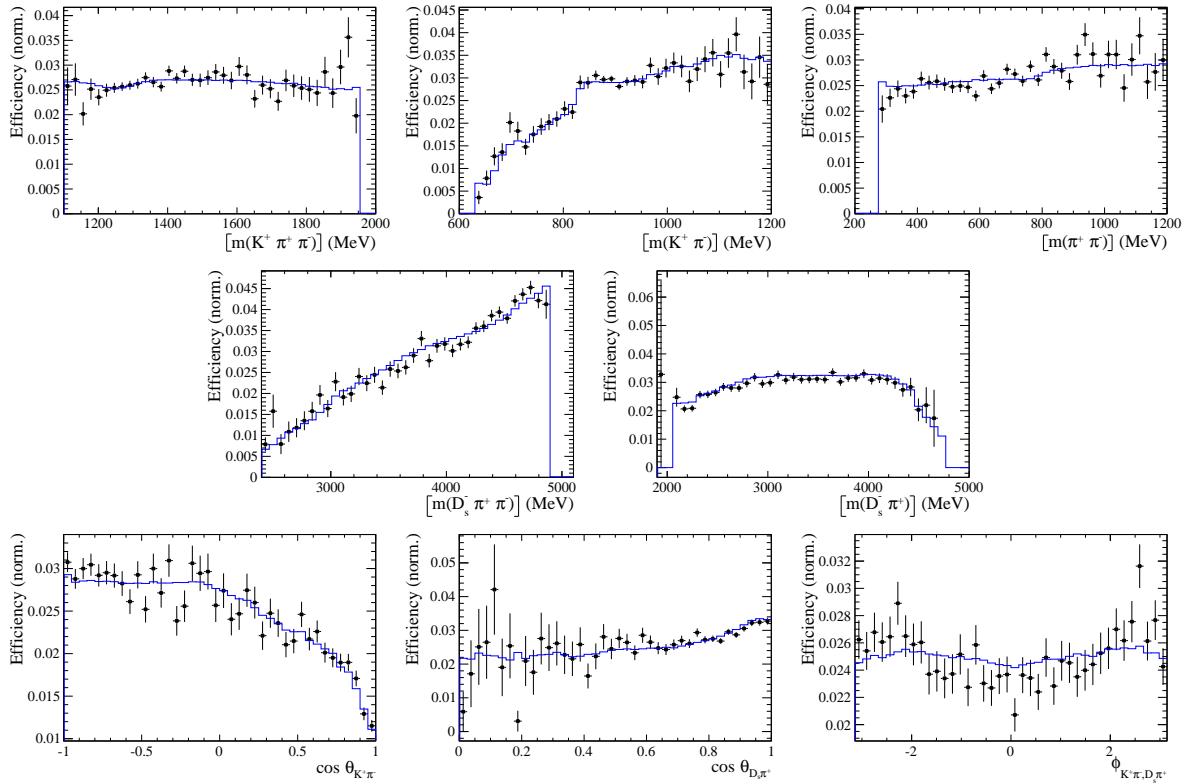


Figure 6.8: Efficiency variation as a function of the phase-space variables obtained from the ratio of selected and generated MC events (data points) and efficiency obtained from a reweighted toy MC sample (blue).

642 7 Flavour Tagging

643 To identify the initial flavour state of the B_s^0 meson, a number of flavour tagging algorithms
 644 are used that either determine the flavour of the non-signal b-hadron produced in the
 645 event (opposite site, OS) or use particles produced in the fragmentation of the signal
 646 candidate B_s^0/\bar{B}_s^0 (same side, SS).

647 For the same side, the algorithm searching for the charge of an additional kaon that
 648 accompanies the fragmentation of the signal candidate is used (SS-nnetKaon). For the
 649 opposite site, four different taggers are chosen: The algorithms that use the charge of an
 650 electron or a muon from semileptonic B decays (OS- e,μ), the tagger that uses the charge
 651 of a kaon from a $b \rightarrow c \rightarrow s$ decay chain (OS-nnetKaon) and the algorithm that determines
 652 the B_s^0/\bar{B}_s^0 candidate flavour from the charge of a secondary vertex, reconstructed from
 653 the OS b decay product (OS-VtxCharge). All four taggers are then combined into a single
 654 OS tagger.

655 Every single tagging algorithm is prone to misidentify the signal candidate at a certain
 656 mistag rate $\omega = (wrongtags)/(alltags)$. This might be caused by particle misidentification,
 657 flavour oscillation of the neutral opposite site B-meson or by tracks that are wrongly picked
 658 up from the underlying event. For every signal B_s^0/\bar{B}_s^0 candidate, each tagging algorithm
 659 predicts a mistag probability η , which is calculated using a combination of inputs such
 660 as the kinematics of the tagging particles. The inputs are then combined to a predicted
 661 mistag using neural networks. These are trained on simulated samples of $B_s^0 \rightarrow D_s^- \pi^+$
 662 (SS algorithm) and $B^+ \rightarrow J/\psi K^+$ (OS algorithms) decays. For the presented analysis, the
 663 measurable CP-violating coefficients are damped by the tagging dilution D , that depends
 664 on the mistag rate:

$$D = 1 - 2\omega. \quad (7.1)$$

665 This means that the statistical precision, with which these coefficients can be measured,
 666 scales as the inverse square root of the effective tagging efficiency,

$$\epsilon_{eff} = \epsilon_{tag}(1 - 2\omega)^2, \quad (7.2)$$

667 where ϵ_{tag} is the fraction of events that have a tagging decision. The flavour tagging
 668 algorithms are optimized for highest ϵ_{eff} on data, using the $B_s^0 \rightarrow D_s^- \pi^+$ and $B^+ \rightarrow J/\psi K^+$
 669 samples.

670 Utilizing flavour-specific final states, the predicted mistag η of each tagger has to be
 671 calibrated to match the observed mistag ω on the data sample. For the calibration, a
 672 linear model of the form

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - <\eta>), \quad (7.3)$$

673 where the values of p_0 and p_1 are determined using the $B_s^0 \rightarrow D_s \pi \pi \pi$ normalization mode
 674 and $<\eta>$ is the average estimated mistag probability $<\eta> = \sum_{i=1}^{N_{cand}} (\eta_i)/N_{cand}$ is used.
 675 Following this model, a perfectly calibrated tagger would lead to $\omega(\eta) = \eta$ and one would
 676 expect $p_1 = 1$ and $p_0 = <\eta>$. Due to the different interaction cross-sections of oppositely
 677 charged particles, the tagging calibration parameters depend on the initial state flavour of
 678 the B_s^0 . Therefore, the flavour asymmetry parameters Δp_0 , Δp_1 and $\Delta \epsilon_{tag}$ are introduced.
 679 For this analysis, the calibrated mistag is treated as per-event variable, giving a larger
 680 weight to events that are less likely to have an incorrect tag. This adds one additional
 681 observable to the time- and amplitude-dependent fit.

682 The tagging calibration is determined using a time-dependent fit to the full $B_s^0 \rightarrow D_s \pi \pi \pi$

sample, where the mixing frequency Δm_s is fixed to the nominal PDG value [33]. The calibration procedure for the OS tagging algorithms (Sec.7.1) and the SS kaon tagger (Sec.7.2) is applied on the full Run I and 2015 and 2016 Run II $B_s^0 \rightarrow D_s\pi\pi\pi$ data sample, which is selected following the steps described in Sec. 3. The similar selection ensures as close as possible agreement between the $B_s^0 \rightarrow D_s\pi\pi\pi$ and $B_s^0 \rightarrow D_sK\pi\pi$ samples in terms of the decay kinematics, which are crucial for the flavour tagging. Section 7.3 shows the compatibility of both samples. After applying the calibration, the response of the OS and SS taggers are combined, which is shown in Sec. 7.4.

7.1 OS tagging calibration

The responses of the OS electron, muon, neural net kaon and the secondary vertex charge taggers are combined for the mistag calibration. The distributions of the predicted OS mistag for signal candidates from $B_s^0 \rightarrow D_s\pi\pi\pi$ is shown in (Run 1) Figure 7.1 and (Run 2) Figure 7.2. The measured tagging power for the OS combination is $\epsilon_{eff,OS} = 4.81\%$.

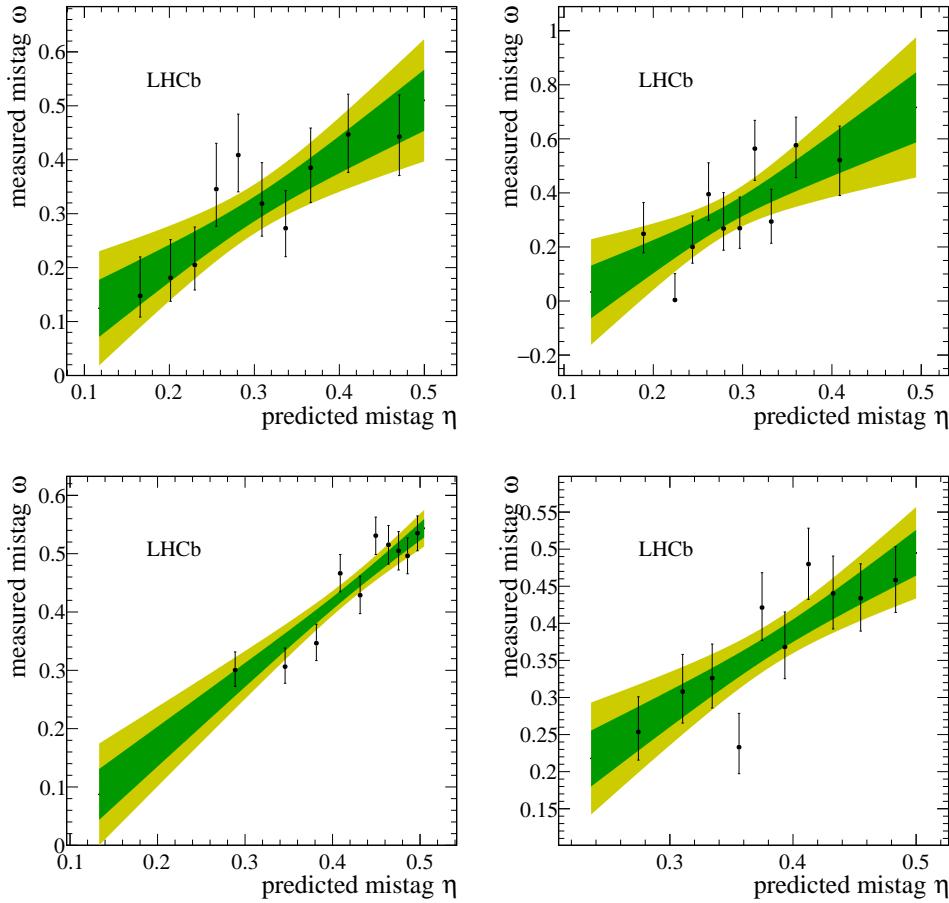


Figure 7.1: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run 1. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

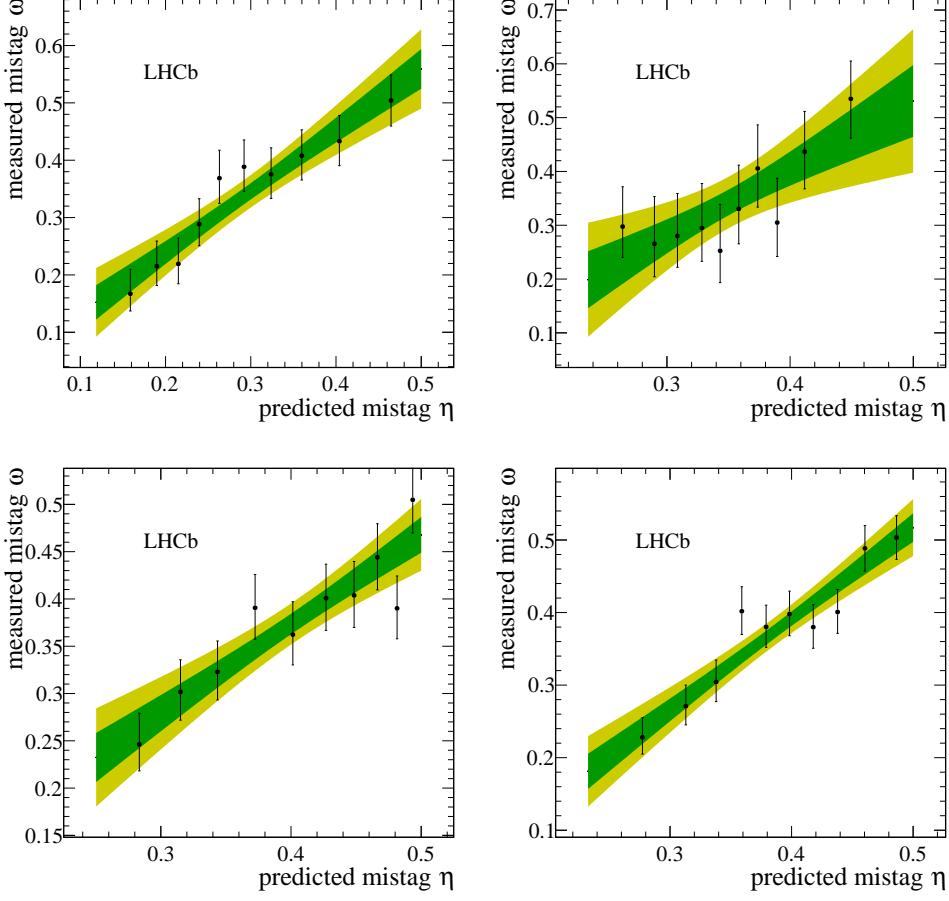


Figure 7.2: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run 2. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

696 7.2 SS tagging calibration

697 The SS neural net kaon tagger can be calibrated using the flavour-specific $B_s^0 \rightarrow D_s\pi\pi\pi$
 698 decay. Its development, performance and calibration is described in detail in [40]. Figure
 699 7.3 shows the distribution of the predicted mistag of the neural net kaon tagger. The
 700 extracted calibration parameters and tagging asymmetries are summarized in Table 7.1
 701 and the measured tagging power for this algorithm is $\epsilon_{eff,SS} = 3.22\%$.

p_0	p_1	$\langle \eta \rangle$	ϵ_{tag}	Δp_0	Δp_1	$\epsilon_{eff} [\%]$
0.008 ± 0.004	1.086 ± 0.059	0.381	0.571 ± 0.002	-0.017 ± 0.004	0.135 ± 0.058	$3.22 \pm 0.03 \text{ (stat)} \pm 0.26 \text{ (cal)}$

Table 7.1: Calibration parameters and tagging asymmetries of the SS tagger extracted from $B_s^0 \rightarrow D_s\pi\pi\pi$ decays.

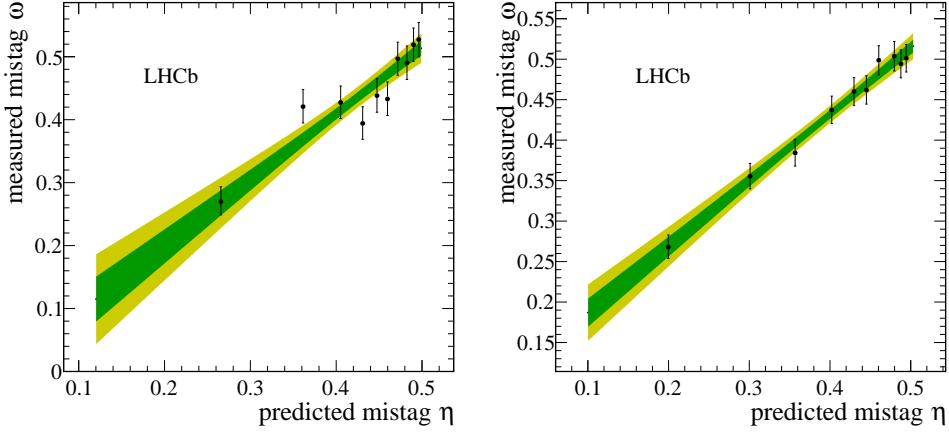


Figure 7.3: Predicted versus measured mistag probability for the SS neural net kaon tagger for (left) Run 1 and (right) Run 2. A linear fit, including the 1σ and 2σ error bands is overlaid for both distributions.

7.3 Tagging performance comparison between the signal and normalization channel

To justify the usage of the tagging calibration, obtained using the $B_s^0 \rightarrow D_s\pi\pi\pi$ sample, for our signal decay, the performance of the taggers in the two decay channels needs to be compatible. This is verified using both, simulated signal samples of both decays and sweighted data, to compare the similarity of the mistag probabilities, tagging decisions and kinematic observables that are correlated with the tagging response, on simulation and data.

The data distributions of the predicted mistag probability η for the OS combination and the SS kaon tagger are shown Fig. 7.4.

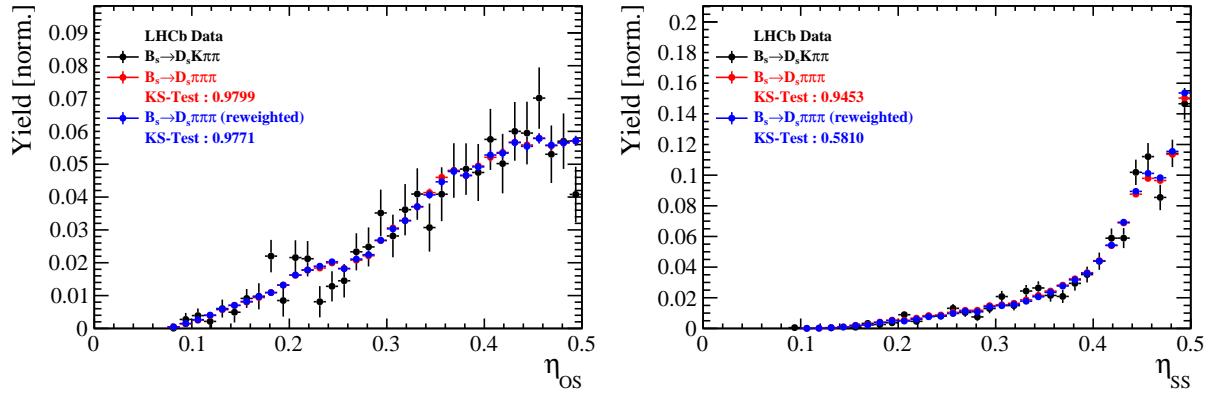


Figure 7.4: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \rightarrow D_s K\pi\pi$ (black) and $B_s^0 \rightarrow D_s \pi\pi\pi$ (red) data samples.

Both, data and simulated samples, show good agreement between the signal and normalization channel.

To justify the portability of the flavour tagging calibration obtained from $B_s^0 \rightarrow D_s\pi\pi\pi$ to the $B_s^0 \rightarrow D_sK\pi\pi$ channel, besides the good agreement of the distributions shown above, the dependence of the measured mistag ω on the predicted mistag η has to be compatible in both channels. This dependence is shown in Fig. 7.5 for simulated signal events of both channels, where good agreement is observed.

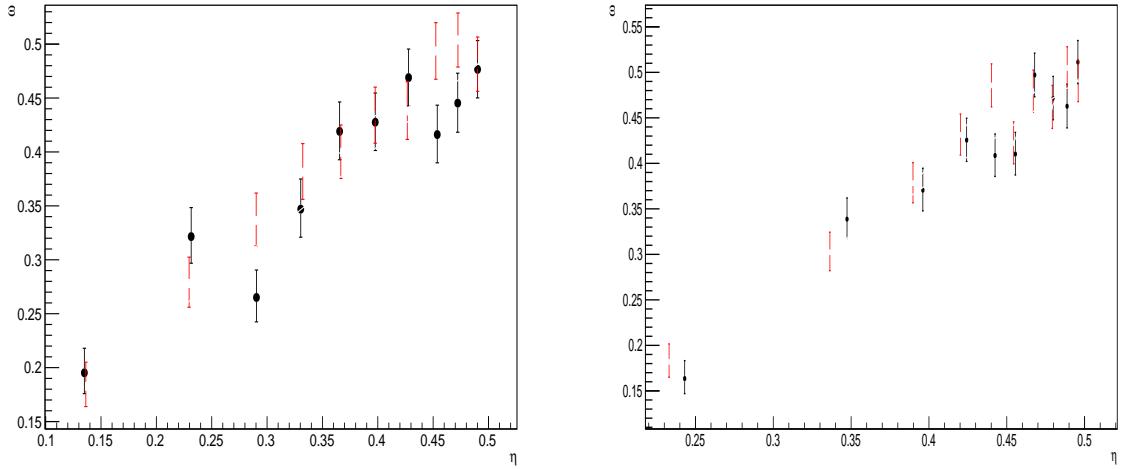


Figure 7.5: Dependence of the observed mistag ω on the predicted mistag η for the (left) OS combination and the (right) SS kaon tagger, found in the simulated $B_s^0 \rightarrow D_sK\pi\pi$ (black) and $B_s^0 \rightarrow D_s\pi\pi\pi$ (red) signal samples.

7.4 Combination of OS and SS taggers

In the time- and amplitude-dependent fit to $B_s^0 \rightarrow D_sK\pi\pi$ data, the obtained tagging responses of the OS and SS tagger will be combined after the calibration described in the previous sections is applied. Events that acquire a mistag probability greater than 0.5 after the calibration will have their tagging decision flipped. For events where only one of the two taggers fired, the combination of the tagging decision is trivial. In those events where both taggers made a decision, we use the standard combination of taggers [41] provided by the flavour tagging group. In the nominal fit, the calibrated mistags ω are combined event by event for the OS and SS tagger, thus adding one observable to the fit procedure. This ensures that the uncertainties of the OS and SS tagging calibration parameters are propagated properly to the combined tagging response for each event.

The tagging performance for the combined tagger in the categories SS tagged only, OS tagged only and SS+OS tagged, are shown in (Run 1) Tab. 7.2 and (Run 2) Tab. 7.2.

Table 7.2: 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	11.32 ± 0.09	37.91 ± 1.02	0.98 ± 0.14
Only SS	41.66 ± 0.18	43.78 ± 0.53	1.54 ± 0.23
Both OS-SS	27.17 ± 0.25	36.68 ± 0.81	2.91 ± 0.29
Combined	80.15 ± 0.32	40.55 ± 0.72	5.43 ± 0.40

Table 7.3: 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	10.51 ± 0.07	35.32 ± 0.77	1.25 ± 0.11
Only SS	43.27 ± 0.14	43.29 ± 0.44	1.58 ± 0.17
Both OS-SS	24.77 ± 0.18	35.14 ± 0.61	3.19 ± 0.22
Combined	78.55 ± 0.24	39.65 ± 0.55	6.02 ± 0.30

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 [42]. 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. [42]

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)	$-0.0586 \pm 0.0648 \pm 0.0150$	$0.0537 \pm 0.0447 \pm 0.0124$

746 8.2 $K^-\pi^+$ Detection Asymmetry

747 The presented measurement of the CKM-angle γ using $B_s^0 \rightarrow D_s K\pi\pi$ decays is sensitive to
 748 a possible charge asymmetry of the kaon. Kaons are known to have a nuclear cross-section
 749 which is asymmetrically dependent on the sign of their charge. It is indispensable to
 750 determine the charge asymmetry of the kaon, as fitting without taking this effect into
 751 account would introduce a 'fake' CP violation. Instead of determining the single track
 752 detection asymmetry of a kaon, it is found that the combined two track asymmetry of a
 753 kaon-pion pair is much easier to access [43]. Therefore, the two track asymmetry defined
 754 as

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

755 is used.

756 This asymmetry can be measured from the difference in asymmetries in the $D^+ \rightarrow$
 757 $K^-\pi^+\pi^+$ and $D^+ \rightarrow K_s^0\pi^+$ modes [44]:

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

758 where possible CP violation in the $D^+ \rightarrow K_s^0\pi^+$ mode is predicted to be smaller than
 759 10^{-4} in the Standard Model [45]. The asymmetry in the neutral kaon system, $A(K^0)$, has
 760 to be taken into account as a correction.

761 We use a dedicated LHCb tool to determine $A^{det}(K^-\pi^+)$ for all data taking periods
 762 used in this analysis. A detailed description can be found in [44]. The tool provides
 763 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
 764 determine the asymmetry following Eq. 8.6. Several weighting steps are performed to
 765 match the kinematics of the calibration samples to our signal decay sample:

766 First, weights are assigned to the K^\pm and π^\pm of the $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$ sample, using
 767 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
 768 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
 769 the corresponding, weighted distributions of the $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$ sample. In a last
 770 step, weights are assigned to match the bachelor pions ϕ distributions between the two
 771 calibration samples.

772 After the samples are weighted, fits are performed to the invariant
 773 $m(K^-\pi^+\pi^+)/m(K^+\pi^-\pi^-)$ and $m(K_s^0\pi^+)/m(K_s^0\pi^-)$ distributions to determine
 774 $A^{det}(K^-\pi^+)$. The PDFs used to describe the invariant mass distributions consist of
 775 gaussian functions for the signal component and exponentials describing the residual
 776 background.

777 The detection asymmetry is determined separately for every year and (since it is a
 778 charge asymmetry effect) magnet polarity. Serving as an example for Run-I and Run-
 779 II, the fits used to determine $N(D^+ \rightarrow K^-\pi^+\pi^+)/N(D^- \rightarrow K^+\pi^-\pi^-)$ and $N(D^+ \rightarrow$
 780 $K_s^0\pi^+)/N(D^- \rightarrow K_s^0\pi^-)$ for 2011, magnet up data and 2015, magnet up data are shown
 781 in Fig. 8.1 and 8.2 respectively. The obtained values of $A^{det}(K^-\pi^+) + A(K^0)$ for all years
 782 and polarities are shown in Table 8.2.

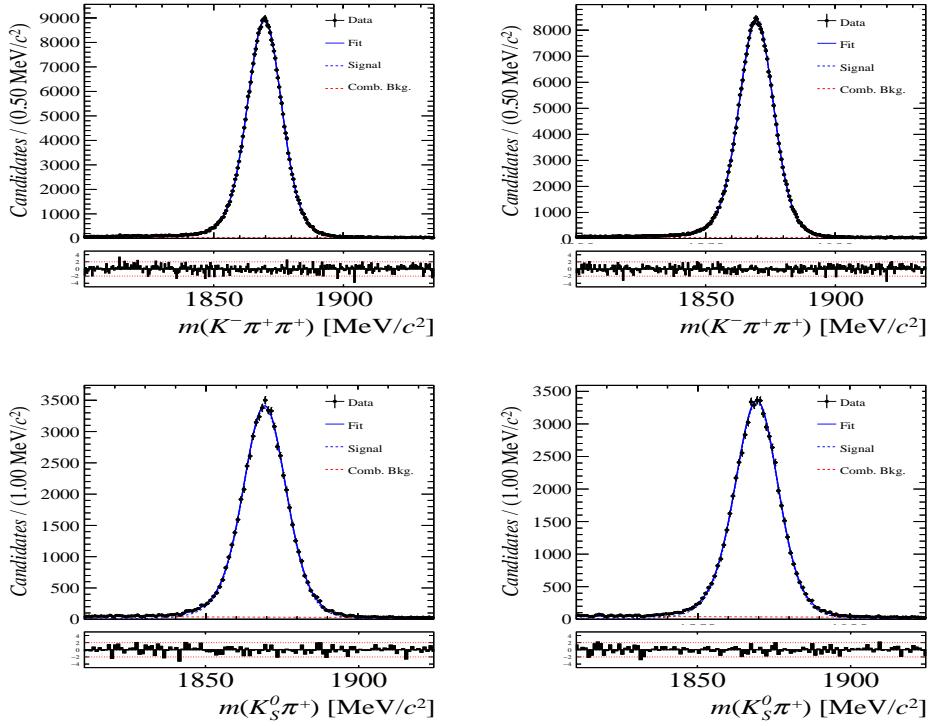


Figure 8.1: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_S^0 \pi^\pm$ candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

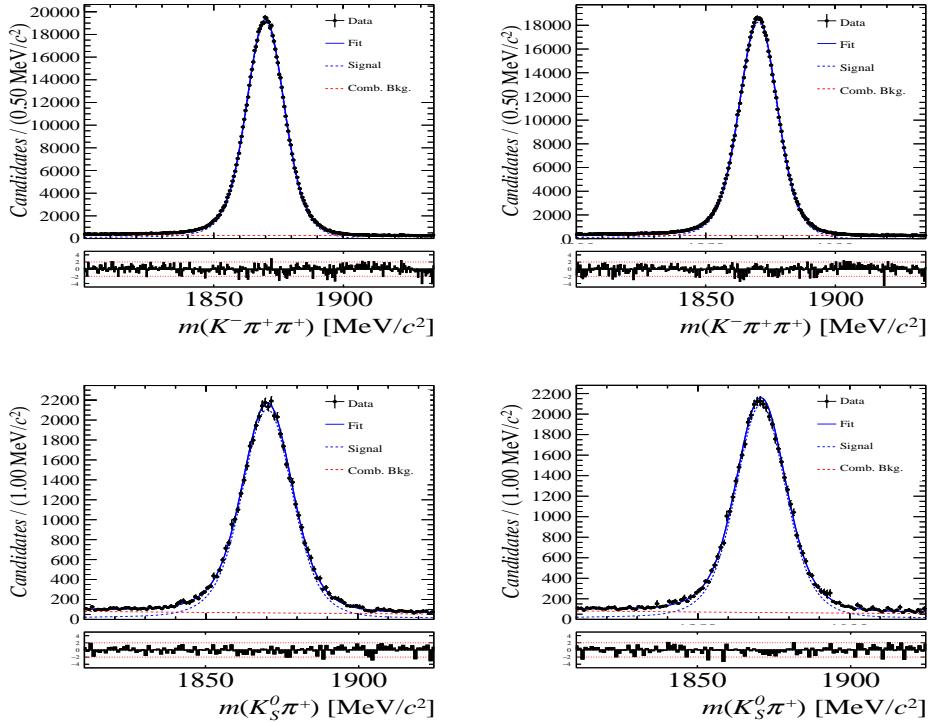


Figure 8.2: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_S^0 \pi^\pm$ candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

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.

783 9 Time dependent fit

784 This section covers the phasespace integrated, time-dependent fit to $B_s^0 \rightarrow D_s h\pi\pi$ data.
 785 We use the **sFit** technique [46] to statistically remove background from the decay time fit,
 786 leaving only the signal PDF to describe the decay time. The **sWeights** are calculated based
 787 on the fit to the reconstructed B_s mass distribution described in Sec. 4. As additional
 788 input to the fit, the tagging information (Sec. 7), as well as the decay time acceptance
 789 (Sec. 6) and resolution (Sec. 5) is used and fixed to the values obtained by the dedicated
 790 studies. Taking all inputs into account, the final time dependent fit PDF is given by

$$791 \quad \mathcal{P}\mathcal{D}\mathcal{F}(t, \vec{\lambda}) = \left(\epsilon(t) \cdot \int P(x, t, q_t, q_f) dx \right) \times \mathcal{R}(t - t'), \quad (9.1)$$

791 where $\int P(x, t, q_t, q_f) dx$ is the PDF given by Eq. 2.7, $\epsilon(t)$ is the efficiency due to the time
 792 acceptance effects and $\mathcal{R}(t - t')$ is the Gaussian time resolution function.

793 9.1 sFit to $B_s^0 \rightarrow D_s \pi\pi\pi$ data

794 The phasespace-integrated, time-dependent fit is performed to the full sWeighted sample
 795 of selected candidates from Run I and 2015+2016 Run II data, containing both possible
 796 magnet polarities and all D_s final states. In the fit, the values of Γ_s and $\Delta\Gamma_s$ are fixed to
 797 the latest PDG report [12]. All tagging parameters are fixed to the central values found in
 798 the tagging calibration, described in Sec. 7. Due to the fact that the $B_s^0 \rightarrow D_s \pi\pi\pi$ decay
 799 is flavour specific, the CP-coefficients can be fixed to $C = 1$ and $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$,
 800 reducing Eq. 2.7 to

$$799 \quad \int P(x, t, q_t, q_f) dx = [\cosh\left(\frac{\Delta\Gamma}{2}t\right) + q_t q_f \cos(m_s t)] e^{-\Gamma t}. \quad (9.2)$$

801 Note that in this case, the dependence on the coherence factor κ is dropped and the
 802 same relation as found for $B_s^0 \rightarrow D_s \pi$ decays is recovered. Therefore, the only free fit
 803 parameter left is Δm_s . The data distribution with the overlaid fit is shown in Fig. 9.1.

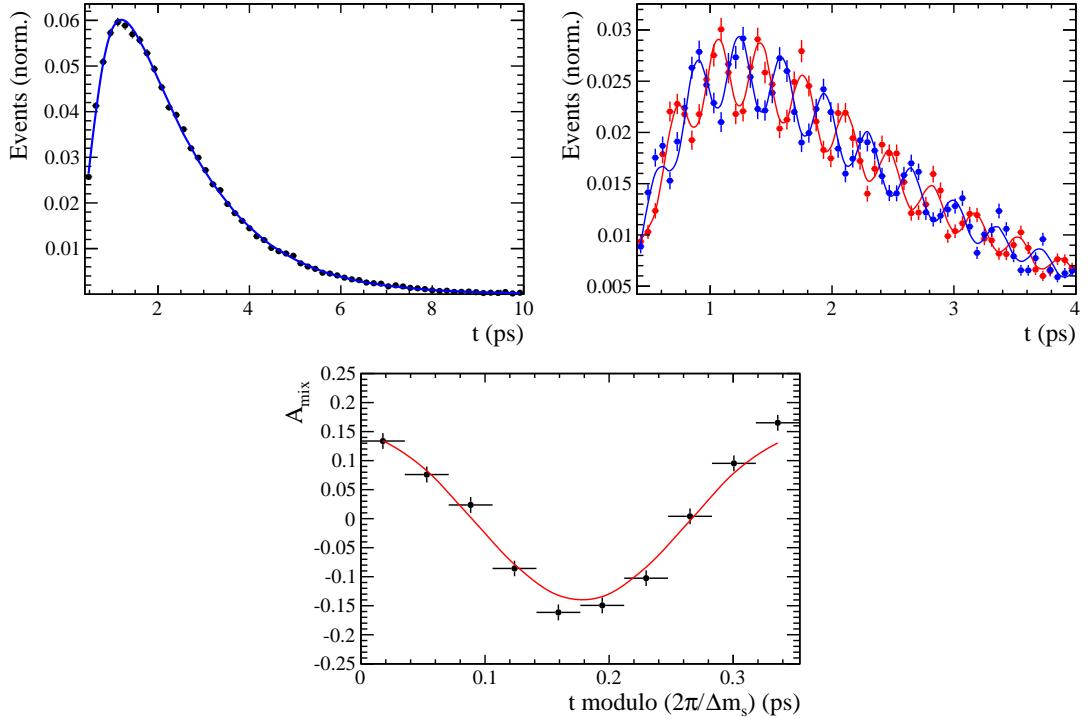


Figure 9.1: Top left: Flavour averaged decay time distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates with the averaged fit overlaid. Top right: Tagged decay time distribution of mixed (red) and unmixed (blue) signal candidates with the fit described in the text overlaid. Bottom: Time-dependent asymmetry A_{mix} between mixed and unmixed B_s^0 candidates in bins of $t/(2\pi\Delta m_s)$.

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}	$0.397 \pm 0.010 \pm 0.010$	$0.369 \pm 0.007 \pm 0.009$
p_1^{OS}	$0.898 \pm 0.086 \pm 0.090$	$0.867 \pm 0.062 \pm 0.063$
Δp_0^{OS}	$0.029 \pm 0.011 \pm 0.002$	$0.010 \pm 0.008 \pm 0.000$
Δp_1^{OS}	$0.015 \pm 0.094 \pm 0.015$	$0.041 \pm 0.075 \pm 0.003$
$\epsilon_{tag}^{OS} [\%]$	$47.739 \pm 0.364 \pm 0.032$	$36.848 \pm 0.243 \pm 0.009$
$\Delta \epsilon_{tag}^{OS} [\%]$	$-0.032 \pm 1.317 \pm 0.093$	$0.987 \pm 0.806 \pm 0.127$
p_0^{SS}	$0.444 \pm 0.008 \pm 0.004$	$0.433 \pm 0.005 \pm 0.004$
p_1^{SS}	$0.950 \pm 0.109 \pm 0.066$	$0.800 \pm 0.055 \pm 0.050$
Δp_0^{SS}	$-0.019 \pm 0.009 \pm 0.001$	$-0.018 \pm 0.006 \pm 0.000$
Δp_1^{SS}	$0.064 \pm 0.123 \pm 0.018$	$0.025 \pm 0.066 \pm 0.004$
$\epsilon_{tag}^{SS} [\%]$	$0.684 \pm 0.003 \pm 0.000$	$0.696 \pm 0.002 \pm 0.000$
$\Delta \epsilon_{tag}^{SS} [\%]$	$-0.000 \pm 0.012 \pm 0.001$	$0.008 \pm 0.008 \pm 0.000$
$A_P [\%]$		-0.045 (fixed)
$\Delta m_s [\text{ps}^{-1}]$		$\text{xx.xx} \pm 0.012 \pm 0.006$

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

The time-dependent fit to the sWeighted sample of $B_s^0 \rightarrow D_s K\pi\pi$ signal candidates is performed simultaneously in the four bins defined in Sec. ??, splitting the data into Run I & II and trigger category 0 (L0Hadron TOS) & 1 (L0Hadron TIS). In these four bins, the respective description of the decay-time acceptance (Sec. 6) is used as an input. As further input the decay-time resolution scaling relation, found separately for Run I & II in Sec. 5, is used in the simultaneous fit. The full fit model is given in Eq. 9.1, where $\int P(x, t, q_t, q_f)$ is:

$$\int P(x, t, q_t, q_f) dx \propto [\cosh\left(\frac{\Delta\Gamma t}{2}\right) + q_t q_f C \cos(m_s t) + \kappa D_{q_f} \sinh\left(\frac{\Delta\Gamma t}{2}\right) - q_t \kappa S_{q_f} \sin(m_s t)] e^{-\Gamma t}. \quad (9.3)$$

Note that the integration over the available phase space x gives rise to the coherence factor κ , which dilutes the sensitivity to the CP coefficients D & S and with that, also to the CKM phase γ . All input parameters from the tagging, time acceptance and resolution are fixed in the fit. The CP coefficients, as well as κ , are therefore the only parameters left floating. The data distribution and the overlaid fit is shown in Fig. 9.2.

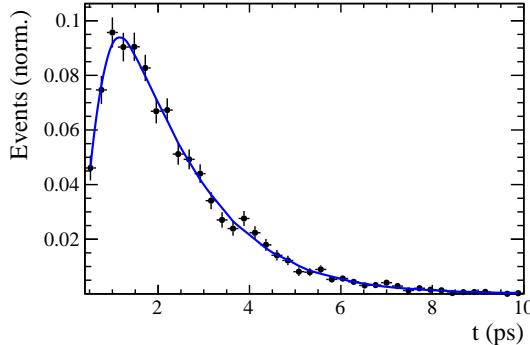


Figure 9.2: Tagged and sWeighted decay-time distribution of $B_s^0 \rightarrow D_s K\pi\pi$ signal candidates. The fit described in the text is overlaid.

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}	$x.xx \pm 0.16 \pm 0.04$

817 10 Time dependent amplitude fit

818 10.1 Signal Model Construction

819 The light meson spectrum comprises multiple resonances which are expected to contribute
 820 to $B_s \rightarrow D_s K\pi\pi$ decays as intermediate states. Apart from clear contributions coming
 821 from resonances such as $K_1(1270)$, $K_1(1400)$ $\rho(770)$ and $K^*(892)^0$, the remaining structure
 822 is impossible to infer due to the cornucopia of broad, overlapping and interfering resonances
 823 within the phase space boundary. The complete list of considered amplitudes can be
 824 found in Appendix G.

825 To build the amplitude model, one could successively add amplitudes on top of one
 826 another until a reasonable agreement between data and fit was achieved. However, this
 827 step-wise approach is not particularly suitable for amplitude analyses as discussed in
 828 Ref. [47]. Instead, we include the whole pool of amplitudes in the first instance and use
 829 the Least Absolute Shrinkage and Selection Operator [47, 48] (LASSO) approach to limit
 830 the model complexity. In this method, the event likelihood is extended by a penalty term

$$-2 \log \mathcal{L} \rightarrow -2 \log \mathcal{L} + \lambda \sum_i \sqrt{\int |a_i A_i(\mathbf{x})|^2 d\Phi_4}, \quad (10.1)$$

831 which shrinks the amplitude coefficients towards zero. The amount of shrinkage is
 832 controlled by the parameter λ , to be tuned on data. Higher values for λ encourage sparse
 833 models, *i.e.* models with only a few non-zero amplitude coefficients. The optimal value
 834 for λ is found by minimizing the Bayesian information criteria [49] (BIC),

$$BIC(\lambda) = -2 \log \mathcal{L} + r \log N_{Sig}, \quad (10.2)$$

835 where N_{Sig} is the number of signal events and r is the number of amplitudes with a decay
 836 fraction above a certain threshold. In this way, the optimal λ balances the fit quality
 837 ($-2 \log \mathcal{L}$) against the model complexity. The LASSO penalty term is only used to select
 838 the model. Afterwards, this term must be discarded in the final amplitude fit with the
 839 selected model, otherwise the parameter uncertainties would be biased.

840 The set of amplitudes is selected using the optimal value of $\lambda = 28$, and is henceforth
 841 called the LASSO model; Figure ??(a) shows the distribution of BIC values obtained by
 842 scanning over λ where we choose the decay fraction threshold to be 0.5%. In addition, we
 843 repeated the model selection procedure under multiple different conditions:

- 844 1. The fit fraction threshold for inclusion in the final model was varied within the
 845 interval [0.05, 5]%. The set of selected amplitudes is stable for thresholds between
 846 0.1% and 1%. Other choices result in marginally different models containing one
 847 component more or less.
- 848 2. Instead of BIC, the Akaike information criteria ($AIC(\lambda) = -2 \log \mathcal{L} + 2r$ [50]) was
 849 used to optimize λ . For a given threshold, the AIC method tends to prefer lower
 850 λ values. However, the set of models obtained varying the threshold within the
 851 interval [0.05, 5]% is identical to the BIC method.
- 852 3. The amplitudes selected under nominal conditions were excluded one-by-one from
 853 the set of all amplitudes considered.

⁸⁵⁴ From that we obtained a set of alternative models shown in Appendix ??.

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

⁸⁵⁵

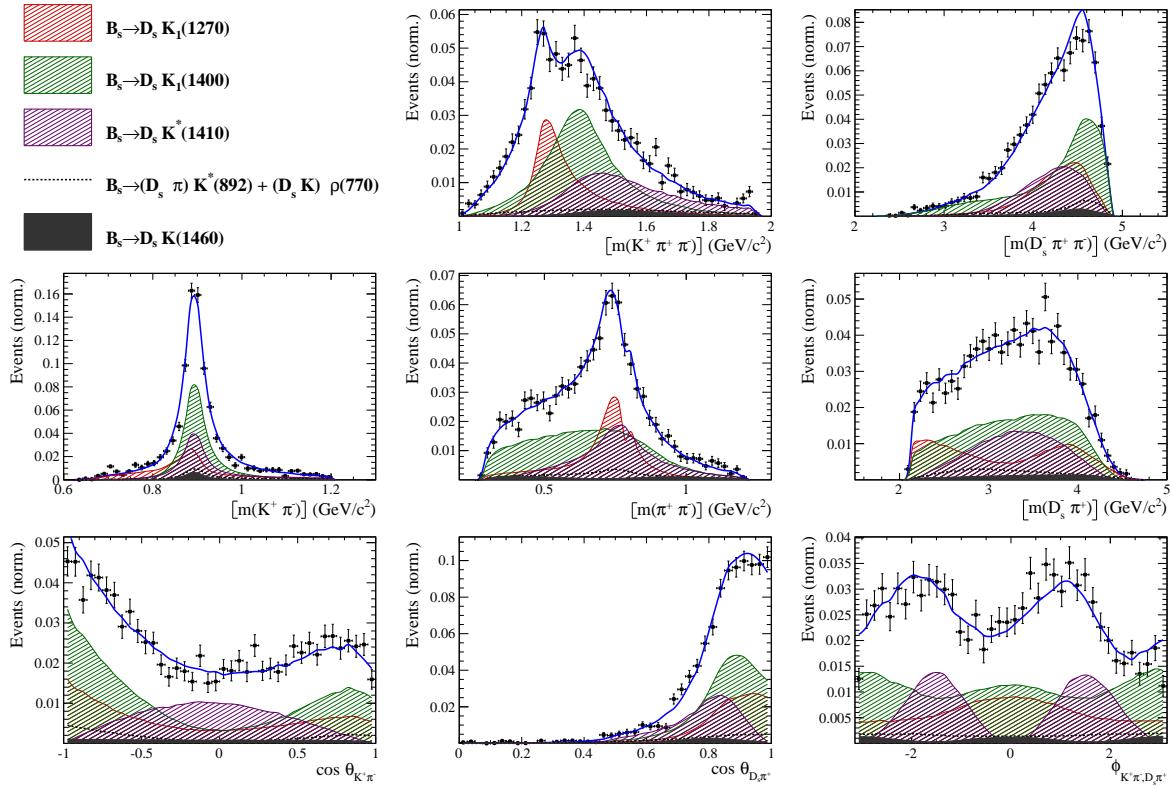


Figure 10.1

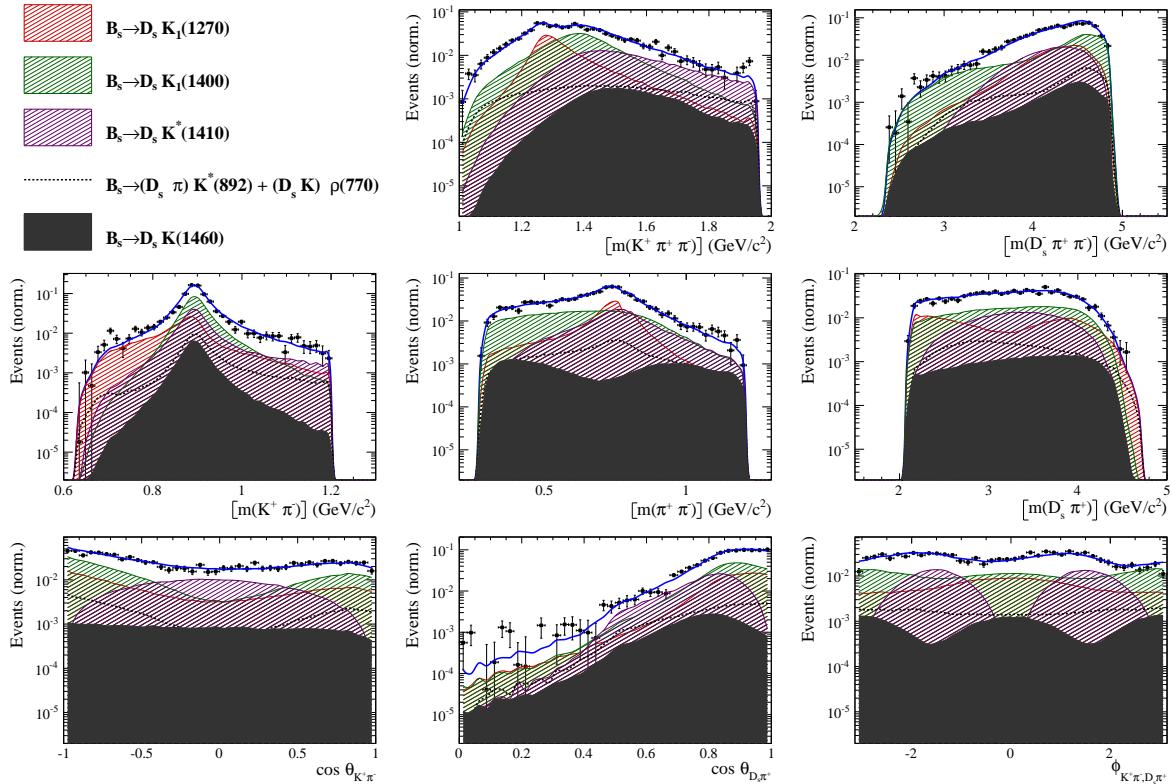


Figure 10.2

856 10.2 Results

857 Table ?? lists the real and imaginary part of the complex amplitude coefficients a_i ,
 858 obtained by fitting the LASSO model to the data, along with the corresponding fit
 859 fractions. The letters in square brackets refer to the relative orbital angular momentum of
 860 the decay products. If no angular momentum is specified, the lowest angular momentum
 861 state consistent with angular momentum conservation and, where appropriate, parity
 862 conservation is used. In order to provide implementation-independent measurements in
 863 addition to the complex coefficients a_i , we define two quantities. Firstly, the fit fractions

$$F_i \equiv \frac{\int |a_i A_i(\mathbf{x})|^2 d\Phi_4}{\int |A_{D^0}(\mathbf{x})|^2 d\Phi_4}, \quad (10.3)$$

864 which are a measure of the relative strength between the different transitions. Secondly,
 865 the interference fractions are given by

$$I_{ij} \equiv \frac{\int 2 \Re[a_i a_j^* A_i(\mathbf{x}) A_j^*(\mathbf{x})] d\Phi_4}{\int |A_{D^0}(\mathbf{x})|^2 d\Phi_4}, \quad (10.4)$$

866 which measures the interference effects between amplitude pairs. Constructive interference
 867 leads to $I_{ij} > 0$, while destructive interference leads to $I_{ij} < 0$. Note that $\sum_i F_i +$
 868 $\sum_{j < k} I_{j,k} = 1$.

869 Figure ?? shows the distributions of selected phase space observables, which demon-
 870 strate reasonable agreement between data and the fit model. We also project into the
 871 transversity basis to demonstrate good description of the overall angular structure in
 872 Fig. ??: The acoplanarity angle χ , is the angle between the two decay planes formed by
 873 the $\pi^+ \pi^-$ combination with minimum invariant mass, $\min[m(\pi^+ \pi^-)]$, and the remaining
 874 $\pi^+ \pi^-$ combination in the D rest frame; boosting into the rest frames of the two-body
 875 systems defining these decay planes, the two helicity variables are defined as the cosine of
 876 the angle, θ , of each π^+ momentum with the D flight direction.

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

$$\chi^2 = \sum_{b=1}^{N_{\text{bins}}} \frac{(N_b - N_b^{\text{exp}})^2}{N_b^{\text{exp}}}, \quad (10.5)$$

879 where N_b is the number of data events in a given bin, N_b^{exp} is the event count predicted
 880 by the fitted PDF and N_{bins} is the number of bins. An adaptive binning is used to ensure
 881 sufficient statistics in each bin for a robust χ^2 calculation [51]. At least 25 events per
 882 bin are required. The number of degrees of freedom ν , in an unbinned fit is bounded by
 883 $N_{\text{bins}} - 1$ and $(N_{\text{bins}} - 1) - N_{\text{par}}$, where N_{par} is the number of free fit parameters. We use
 884 the χ^2 value divided by $\nu = (N_{\text{bins}} - 1) - N_{\text{par}}$ as a conservative estimate. For the LASSO
 885 model, this amounts to $\chi^2/\nu = 1.40$ indicating a decent fit quality.

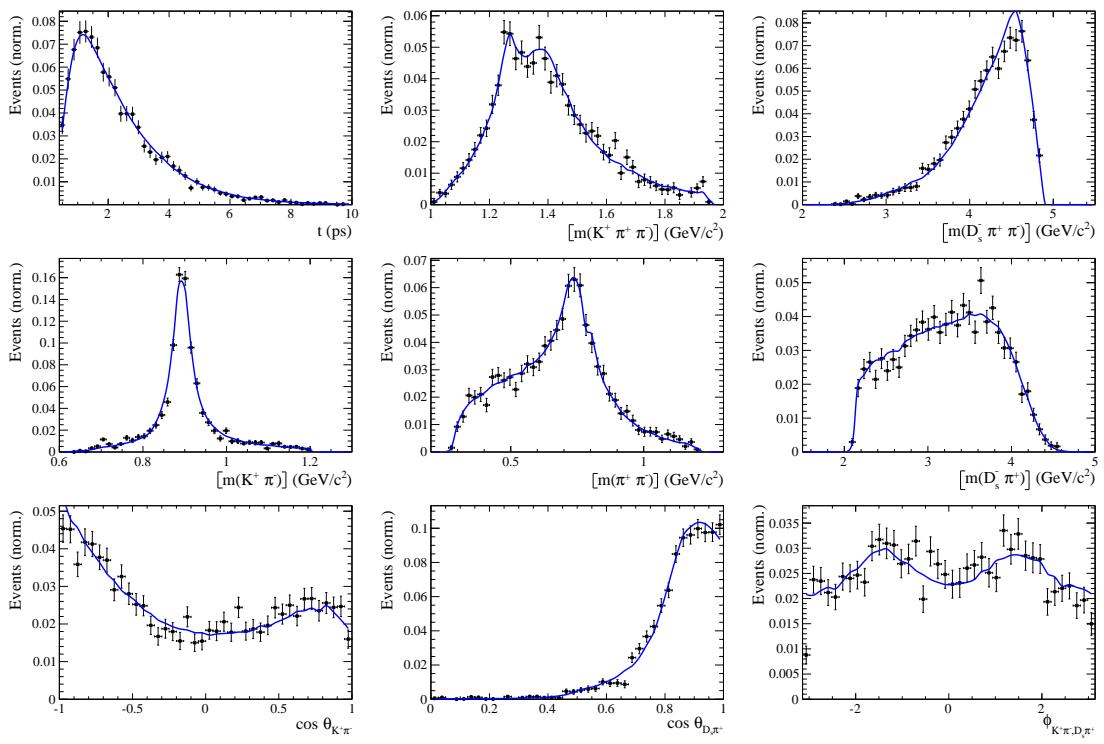


Figure 10.3

Table 10.1: 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			$\text{xx.xx} \pm 0.04 \pm 0.05 \pm 0.04$	
$\delta [^\circ]$			$\text{xx.xx} \pm 16.1 \pm 6.8 \pm 6.8$	
$\gamma - 2\beta_s [^\circ]$			$\text{xx.xx} \pm 16.1 \pm 11.6 \pm 6.2$	

Table 10.2: 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

886 11 Systematic uncertainties

887 This section covers all relevant systematic uncertainties on the measured observables.
 888 In particular, the model dependent description of the invariant B_s^0 mass spectrum, the
 889 parametrization of the time acceptance using cubic splines, as well as the scaling of the
 890 time resolution and tagging calibration are potential sources of systematic errors. The
 891 largest contribution of systematic uncertainty is expected to appear in the choice of
 892 amplitudes entering the model to describe the 5 dimensional phase space, discussed in
 893 Section 10.

894 11.1 Models for B_s^0 mass distribution

895 The statistical subtraction of the residual background [46], left after the full selection,
 896 relies on the correct description of the invariant B_s^0 mass distribution. Since the choice
 897 of signal and background models is not unique, alternative descriptions which lead to
 898 slightly different yields for the signal and background components are available. The
 899 difference in yields could result in shifted values for the measured observables and are
 900 therefore treated as systematic uncertainty.

901

902 11.1.1 Signal model

903 The Johnson's SU function which is used as nominal signal model is replaced by a double
 904 Crystal Ball [52]. The crystal ball model is given by a gaussian core with an exponential
 905 tail on one side. Choosing a double Crystal Ball allows for asymmetric tails in a slightly
 906 different way compared to the Johnson's SU function.

907 11.1.2 Background model

908 For the description of the partially reconstructed background, a combination of the
 909 RooHILLdini and RooHORNsdini model [53] is used instead of the nominal model of
 910 three bifurcated gaussians. The HORNsdini model is used to describe the $B_s^0 \rightarrow D_s^* [\rightarrow$
 911 $D_s(\pi^0)] X_{s/d}$ decay, where the brackets around the π^0 indicate that it is missed in the
 912 reconstruction. The $D_s^* \rightarrow D_s\pi^0$ decay is a Vector \rightarrow Scalar-Scalar ($1^- \rightarrow 0^-0^-$)
 913 transition. Using the helicity of the D_s , one can show that this results in a double-peak
 914 structure in the reconstructed B_s^0 mass. Therefore, the HORNsdini shape consists of a
 915 gaussian-like double-peak structure:

$$916 HORN S(m_{B_s^0}) = \int_a^b dm_{B_s^0} \left(m_{B_s^0} - \frac{a+b}{2} \right)^2 \mathcal{D}\mathcal{G}(m_{B_s^0} | \mu, \sigma, f_G) \left(\frac{1-\zeta}{b-a} m_{B_s^0} + \frac{b\zeta - a}{b-a} \right), \quad (11.1)$$

916 where a and b are the kinematic endpoints of the distribution and ζ is the positive,
 917 real fraction of the two peak heights. Additionally, the shape is convoluted with a gaussian
 918 to account for resolution effects.

919 The HILLdini model parametrizes the invariant mass shape of $B_s^0 \rightarrow D_s^* [\rightarrow D_s(\gamma)] X_{s/d}$
 920 candidates, where the γ is not reconstructed. Contrary to the previously discussed process,
 921 the $Ds^* \rightarrow D_s\gamma$ is a Vector \rightarrow Scalar-Vector ($1^- \rightarrow 0^-1^-$) transition. From helicity

arguments, the expected shape in the mass distribution of B_s^0 candidates follows a parabolic curve without any peaking structure. To accommodate for this shape, the HILLdini model consists of a parabolic curve between the kinematic endpoints a & b:

$$HILL(m_{B_s^0}) = \begin{cases} -(m_{B_s^0} - a)(m_{B_s^0} - b), & \text{for } a < m_{B_s^0} < b \\ 0, & \text{otherwise.} \end{cases} \quad (11.2)$$

This shape is convoluted with the same gaussian resolution function used for the HORNSdini model.

To study systematic uncertainties originating from the description of the combinatorial background, the nominal second order polynomial is replaced by an exponential function.

11.1.3 Description of misidentified background

The fixed shape and yield of the mis-ID background in the $m(D_s K\pi\pi)$ spectrum is another source of systematic uncertainty. To evaluate this possible source arising from the description of the single mis-ID of $B_s^0 \rightarrow D_s^{(*)}\pi_K\pi\pi$ candidates, we vary the yield of this component as follows:

- We fix the yield of the mis-ID components to zero.
- We double the yield of the mis-ID components.
- We quadruple the yield of the mis-ID components.

For the shape of the mis-ID background, the nominal approach is to use a simulated 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 hypothesis of the π^+ with the higher misidentification probability (see Sec. 4). The resulting $m(D_s^{(*)}\pi_K\pi\pi)$ distribution is then modelled and the shape obtained from the fit is used in the nominal mass fit to signal. This approach is modified as follows:

- We flip the mass hypothesis of the π^+ candidate with the lower probability of being misidentified.
- We randomly flip the mass hypothesis of a π^+ candidate.

For the five variations of the misidentified background component, new signal sWeights are generated and the time dependent fit is reiterated.

11.1.4 Systematic effect on observables

The shift of the central values of the observables in the full fit when using sWeights obtained from a combination of alternative models, as well as using only one alternative model for the signal/comb.background/part.reco.background and keeping the nominal model for the other parts, is shown in Table ??.

952 **11.2 Decay-time acceptance**

953 To investigate the systematic uncertainty related to the decay-time dependent efficiency,
 954 we vary our parametrization of the acceptance using cubic splines. This is explicitly
 955 done by choosing slightly different knot positions, varying the spline coefficients at the
 956 nominal positions within their statistical uncertainties and adding/subtracting knots
 957 in the range $0.4 \text{ ps} < t < 11 \text{ ps}$. Additionally, an adaptive binning scheme is used to
 958 determine the knot positions in a way that roughly equal amounts of data is covered
 959 between two knots. Strictly speaking, the variation of the spline coefficients within their
 960 uncertainty gives the statistical uncertainty of the decay-time acceptance parametrization.
 961 For the presented measurement, this is done using the Cholesky decomposition [54] of
 962 the covariance matrix of coefficients c_i , generating toy splines with randomized coefficient
 963 values $c_{i,toy}$ from this decomposition and refitting using the toy spline. Furthermore, the
 964 fit to the decay-time distribution of signal $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates, used to determine
 965 the spline parametrization, is reiterated with varying fixed/constrained values for $\Delta\Gamma_s$.

966 **11.2.1 Varition of knot positions**

967 The nominal knot positions are changed to be:

$$k_{alt1}(t) = [0.5 1 1.5 2 3 6 9.5], k_{alt2}(t) = [0.5 1 1.5 2 3 9 11], k_{adaptive}(t) = [0.7 1.2 1.7 2.2 6.3]$$

968 The variation of knot positions is found to give a neglectable effect when compared to
 969 the variation of spline coefficients.

970 **11.2.2 Variation of spline coefficients**

971 Due to the sizeable correlation of the spline coefficients c_i determined in Chapter 6.2, the
 972 variations of the observables in the amplitude fit when changing one spline coefficient can
 973 not be added up in quadrature for all coefficients. To simplify the problem, a Cholesky
 974 decomposition [54] is used to generate a set of uncorrelated vectors from the covariance
 975 matrix A_{cov} . It can be shown that every Hermitian positive-definite matrix, such as A_{cov} ,
 976 has a unique Cholesky decomposition of the form:

$$A_{cov} = L \cdot L^T, \quad (11.3)$$

977 where L is a lower triangular matrix with real and positive diagonal entries and L^T
 978 denotes the transpose of L .

979 Given the four free spline coefficients which are determined from the fit described in 6,
 980 A_{cov} is a 4×4 matrix. Therefore, the lower triangular matrix L is of the form:

$$L = \begin{pmatrix} v_{11} & 0 & 0 & 0 \\ v_{12} & v_{22} & 0 & 0 \\ v_{13} & v_{23} & v_{33} & 0 \\ v_{14} & v_{24} & v_{34} & v_{44} \end{pmatrix}, \quad (11.4)$$

982 where v_{ij} are real and positive numbers. L contains four row vectors, which are by
 983 construction the four decorrelated modes of the covariant matrix A_{cov} . From this modes,
 984 one can form variations for each of the spline coefficients:

$$c_i = c_{nom,i} + \sum_j (r_j \cdot v_{ij}), \quad (11.5)$$

where $i = 1..4$, c_i is the i-th generated coefficient of the toy spline, $c_{nom,i}$ is the i-th coefficient determined from the nominal decay-time dependent fit to $B_s^0 \rightarrow D_s \pi\pi\pi$, r_j are normally distributed real random numbers from a distribution of unit width and v_{ij} are the components of L (where i is the row index and j the column index).

We now generate four sets of 100 toy splines, where one of the four spline coefficients is varied each time using Eq. 11.5. Thus, the time-dependent amplitude fit is repeated in total 400 times with a generated toy spline and the shift of the mean value of the physics observables over each of the $4 \cdot 100$ sets is quoted as uncertainty arising from $c_{i=1..4}$. The uncertainties are then added in quadrature to form the overall uncertainty due to the spline coefficients.

11.3 Decay-time resolution

To study systematic effects originating from the scaling of the decay-time resolution σ_t , the decay-time distribution of fake B_s^0 candidates using prompt D_s is described by single Gaussian function. The results of the single Gaussians in the different bins of the per-event decay-time error can then be used to derive the scaling function in a straightforward way. Since the distribution of the fake B_s^0 decay time does not follow a perfect Gaussian distribution, two different approaches which either slightly overestimate or underestimate the decay time error are used:

- A double Gaussian is fit to the decay-time distributions of fake B_s^0 candidates, but only the narrow width of the core Gaussian is considered to represent the time resolution in the respective bin. This method assumes that the other, broader Gaussian component does not represent the decay-time resolution of the signal B_s^0 sample. Therefore the resolution is slightly underestimated in this case.
- A single Gaussian is fit to the decay-time distributions of fake B_s^0 candidates in a wide range of $[-3\sigma_t : 1.5\sigma_t]$. Due to the tails of the distribution, which broaden the width of the Gaussian function, this method slightly overestimates the decay-time resolution.

The widths of the single Gaussians from the fits performed with the two methods in bins of the per-event decay-time error is studied and a new resolution scaling function is derived for both cases:

$$\sigma_{eff}^{core-Gauss}(\sigma_t) = (4.9 \pm 2.0) \text{ fs} + (0.821 \pm 0.050) \sigma_t \quad (11.6)$$

$$\sigma_{eff}^{single-Gauss}(\sigma_t) = (8.3 \pm 1.5) \text{ fs} + (0.997 \pm 0.037) \sigma_t \quad (11.7)$$

The scaling functions are shown in Fig. 11.1 and the systematic uncertainty to the CP-observables is summarized in Table ?? .

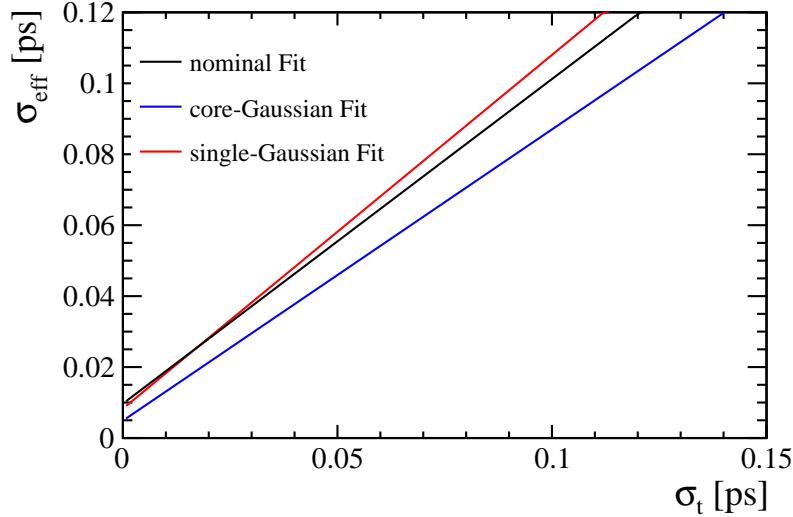


Figure 11.1: The measured resolution scaling function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data) for (black line) the nominal scaling, (blue line) only using the narrow gaussian width of the double gaussian fit model or (red line) when determining the resolution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

1018 11.4 Tagging calibration

1019 Systematic uncertainties arise from the statistical precision of the tagging parameters
 1020 determined from the calibration, discussed in Sec. 7. These uncertainties are accounted
 1021 for by the inclusion of Gaussian constrains in the nominal fit. The width of the respective
 1022 constrain for the tagging parameter p_i is chosen to be Δp_i . In this way, the systematic
 1023 uncertainty due to the tagging calibration is included in the statistical uncertainty of the
 1024 time dependent fit.

1025 11.5 Summary of systematic uncertainties

1026 All contributing systematic uncertainties are summarized in Table ???. The individual
 1027 uncertainties are summed in quadrature to arrive at the total systematic uncertainty
 1028 for the respective CP observable. Their total magnitude ranges from (30-40)% of the
 1029 statistical uncertainty of the fit.

Table 11.1: 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 11.2: 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	1.01	0.01	0.04	0.00		1.01
p_1^{OS} Run-I	0.01	0.00	1.04	0.00	0.05	0.00		1.04
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	0.06	0.01	0.02	0.00		0.07
p_0^{SS} Run-I	0.01	0.00	0.56	0.00	0.03	0.00		0.56
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	0.13	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	0.00	1.22	0.00	0.07	0.00		1.22
p_1^{OS} Run-II	0.01	0.00	1.00	0.00	0.07	0.00		1.01
Δp_0^{OS} Run-II	0.04	0.00	0.04	0.00	0.02	0.00		0.06
Δp_1^{OS} Run-II	0.01	0.00	0.02	0.00	0.03	0.00		0.04
ϵ_{tag}^{OS} Run-II	0.01	0.00	0.00	0.00	0.03	0.00		0.04
$\Delta \epsilon_{tag}^{OS}$ Run-II	0.01	0.00	0.15	0.00	0.03	0.00		0.16
p_0^{SS} Run-II	0.00	0.00	0.79	0.00	0.02	0.00		0.79
p_1^{SS} Run-II	0.05	0.00	0.90	0.00	0.02	0.00		0.90
Δp_0^{SS} Run-II	0.00	0.00	0.01	0.00	0.02	0.00		0.03
Δp_1^{SS} Run-II	0.05	0.00	0.02	0.00	0.02	0.00		0.06
ϵ_{tag}^{SS} Run-II	0.00	0.00	0.00	0.00	0.03	0.00		0.03
$\Delta \epsilon_{tag}^{SS}$ Run-II	0.01	0.00	0.04	0.00	0.02	0.00		0.04
A_P Run-II	0.07	0.00	0.07	0.00	0.02	0.00		0.10
Δm_s	0.01	0.00	0.12	0.02	0.05	0.00	0.48	0.50

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	0.02	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
$B_s \rightarrow (D_s \pi)_P K^*(892)$ Phase($b \rightarrow u$)	0.45	0.24	0.01	0.05	0.83	0.49	1.34	0.38	2.81	0.33		3.34
$B_s \rightarrow (D_s K)_P \rho(770)$ Mag($b \rightarrow u$)	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

1030 A Stripping and Trigger cuts

1031 The following text describes variables which are used in Table 1.1 and might be ambiguous,
 1032 or which benefits are not straight forward. Where noted, different cut values are applied
 1033 for Run-I and Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest
 1034 approach. This variable is used to ensure that all D_s and $X_{s,d}$ daughters originate from
 1035 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	$> 1000 \text{ MeV}/c$	
Track p_T	$> 100 \text{ MeV}/c$	
Track IP χ^2	> 4	
Track ghost-prob.	< 0.4	
D_s mass	$m_{D_s} \pm 100 \text{ MeV}$	$m_{D_s} \pm 80 \text{ MeV}$
D_s Daughter p_T	$\sum_{i=1}^3 p_{t,i} > 1800 \text{ MeV}/c$	
D_s Daughter DOCA	$< 0.5 \text{ mm}$	
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	$< 4000 \text{ MeV}$	$< 3500 \text{ MeV}$
$X_{s,d}$ Daughter p	$> 2 \text{ GeV}/c$	
$X_{s,d}$ Daughter DOCA	$< 0.4 \text{ mm}$	
$X_{s,d}$ Daughter p_T	$\sum_{i=1}^3 p_{t,i} > 1250 \text{ MeV}/c$	
$X_{s,d}$ Vertex χ^2/nDoF	< 8	
$X_{s,d}$ χ^2 -separation from PV	> 16	
$X_{s,d}$ DIRA	> 0.98	
$X_{s,d}$ $\Delta\rho$	$> 0.1 \text{ mm}$	
$X_{s,d}$ Δz	$> 2.0 \text{ mm}$	
$X_{s,d}$ daughter PID(π)	< 10	
X_s daughter PID(K)	> -2	> 4
B_s^0 mass	$[4750, 7000] \text{ MeV}/c^2$	$[5000, 6000] \text{ MeV}/c^2$
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
$B_s^0 \tau_{B_s^0}$	$> 0.2 \text{ ps}$	

1036

1037 Table 1.2 summarizes the trigger requirements imposed by the Hlt1 line used in this
 1038 analysis for Run-I. At least one of the six decay particles must pass the listed requirements
 1039 in order for the event to be stored for further analysis. For Run-II, this trigger line was
 1040 updated and uses a multivariate classifier which takes the variables listed in Table 1.2 as
 1041 input, rather than directly cutting on them.

1042 The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the
 1043 b-hadron p_T , its flight distance χ^2 from the nearest PV and the sum of the B_s^0 and D_s
 1044 vertex χ^2 divided by the sum of their number of degrees of freedom. Table 1.3 summarizes
 1045 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

1046

B Details of multivariate classifier

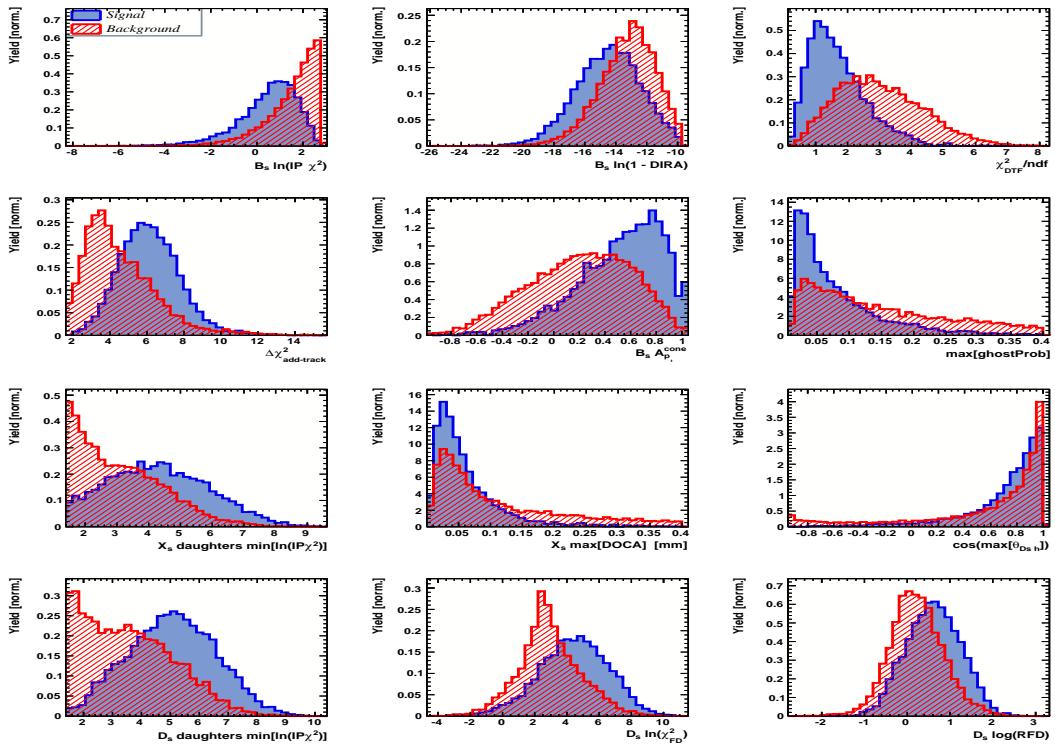


Figure A.1: Variables used to train the BDTG for category [Run-I,L0-TOS].

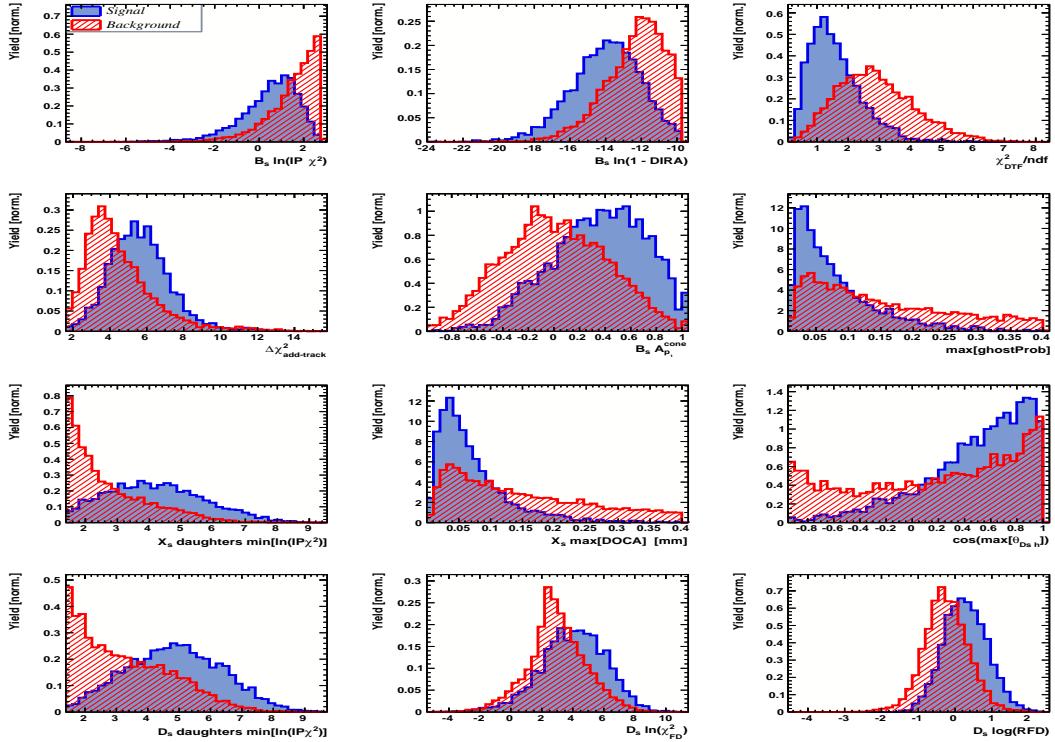


Figure A.2: Variables used to train the BDTG for category [Run-I,L0-TIS].

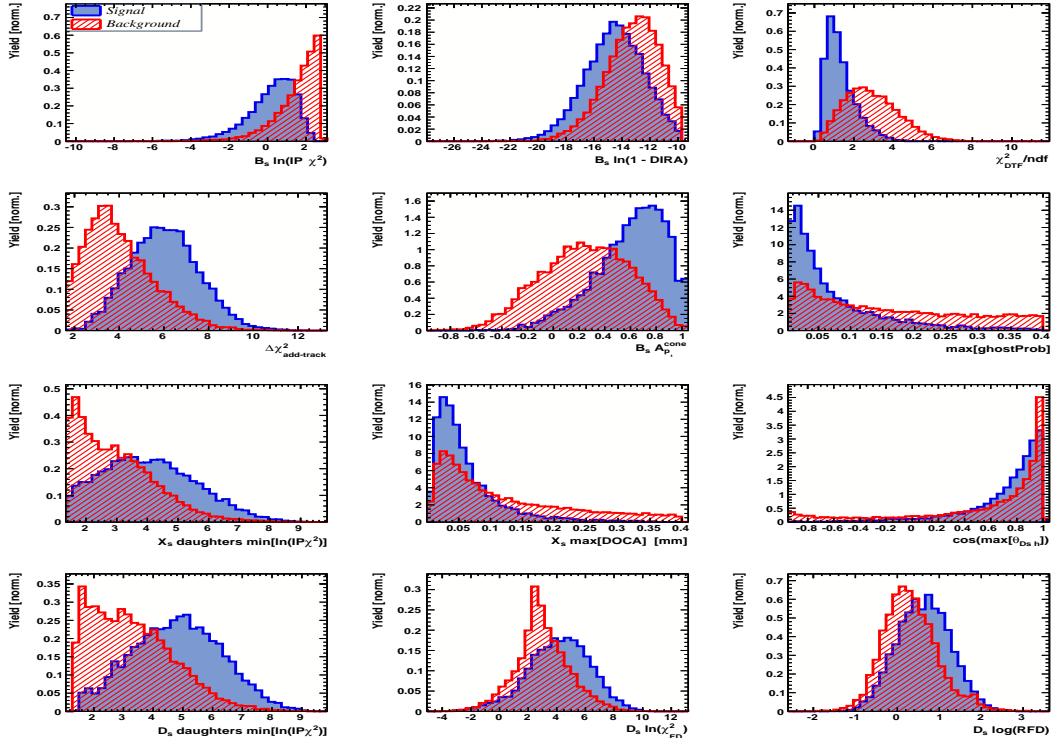


Figure A.3: Variables used to train the BDTG for category [Run-II,L0-TOS].

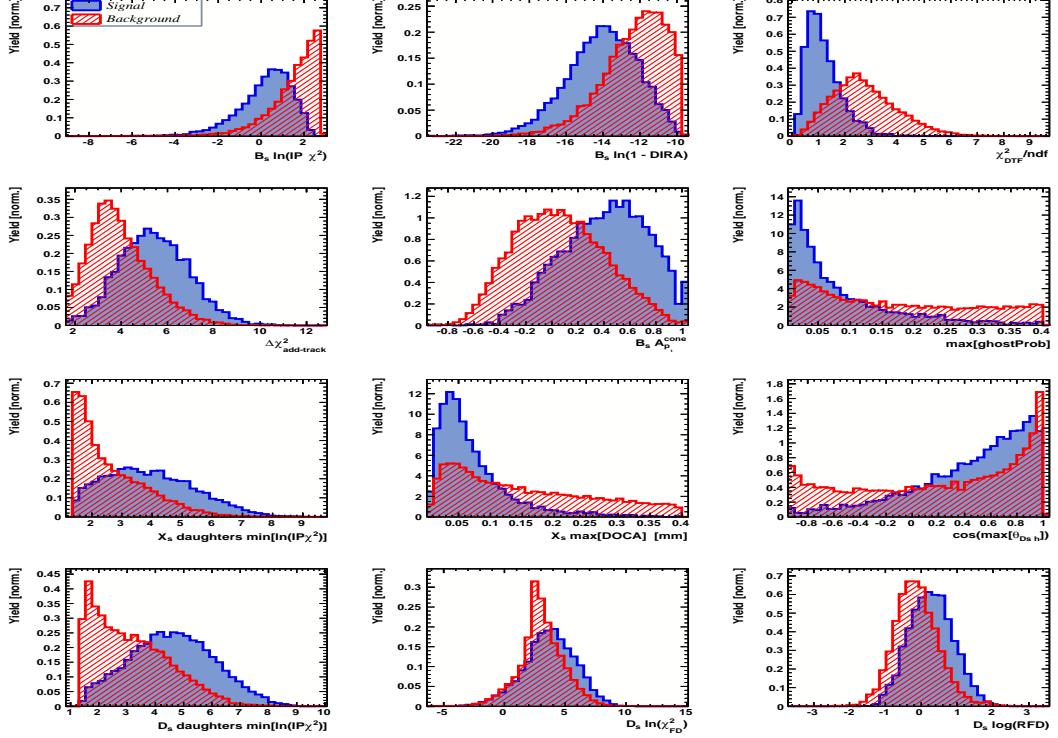


Figure A.4: Variables used to train the BDTG for category [Run-II,L0-TIS].

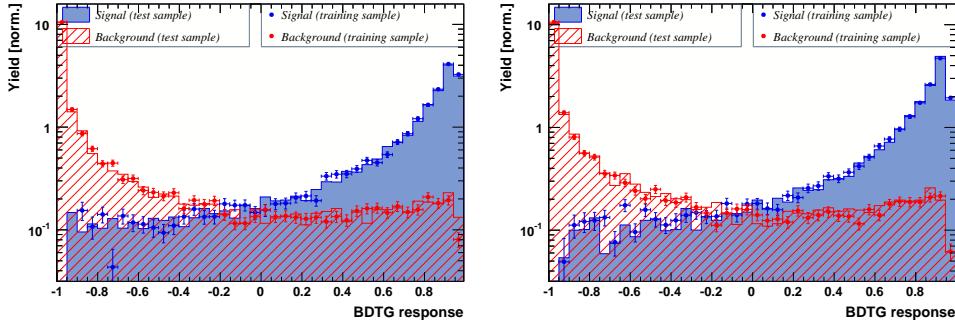


Figure A.5: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].

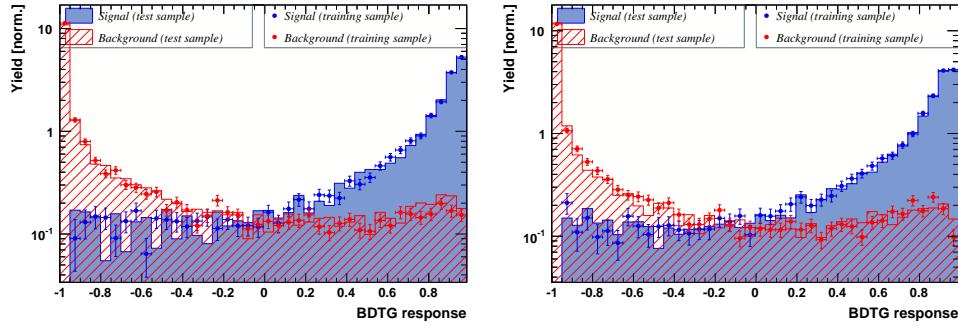


Figure A.6: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].

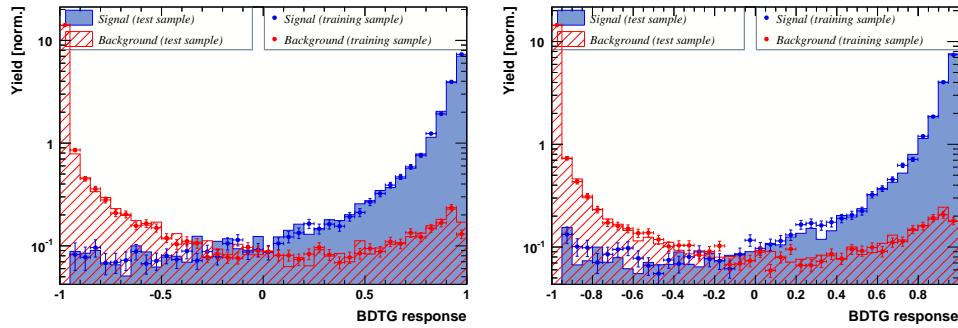


Figure A.7: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].

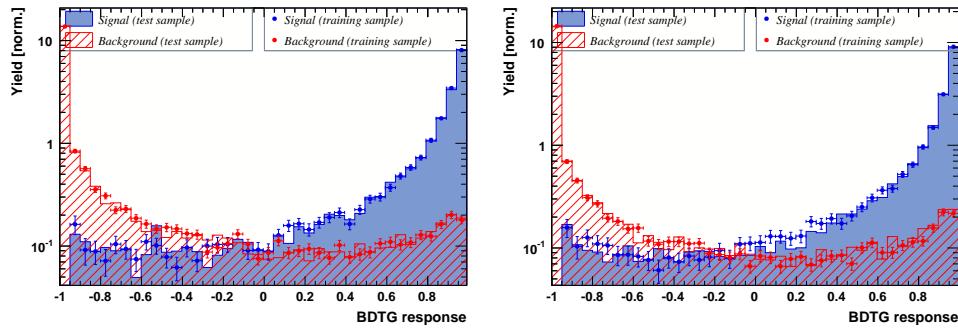


Figure A.8: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].

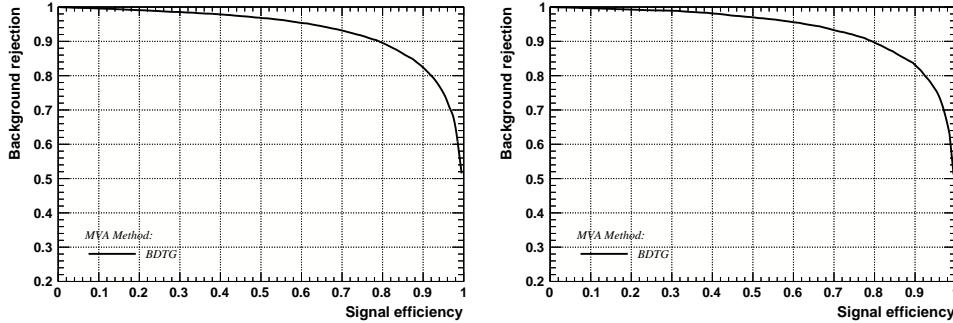


Figure A.9: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].

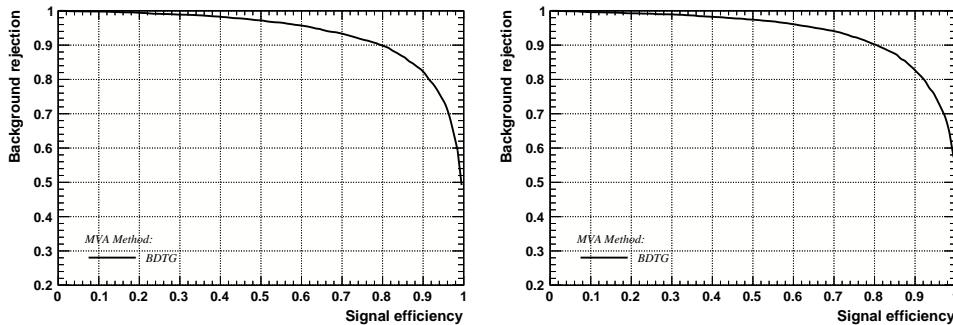


Figure A.10: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TIS].

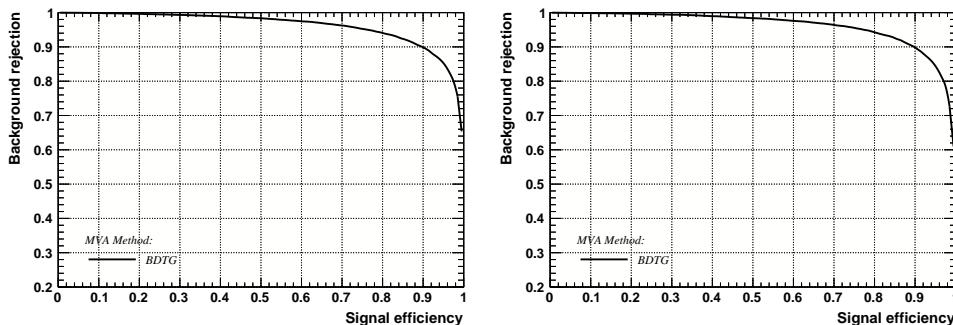


Figure A.11: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TOS].

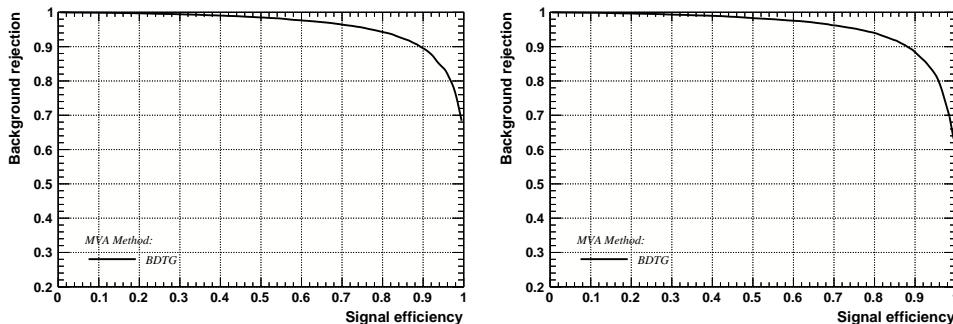


Figure A.12: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TIS].

1048 C Detailed mass fits

1049 In this section, all fits to the mass distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ and $B_s^0 \rightarrow D_sK\pi\pi$
 1050 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-
 1051 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
 1052 L0 trigger category.

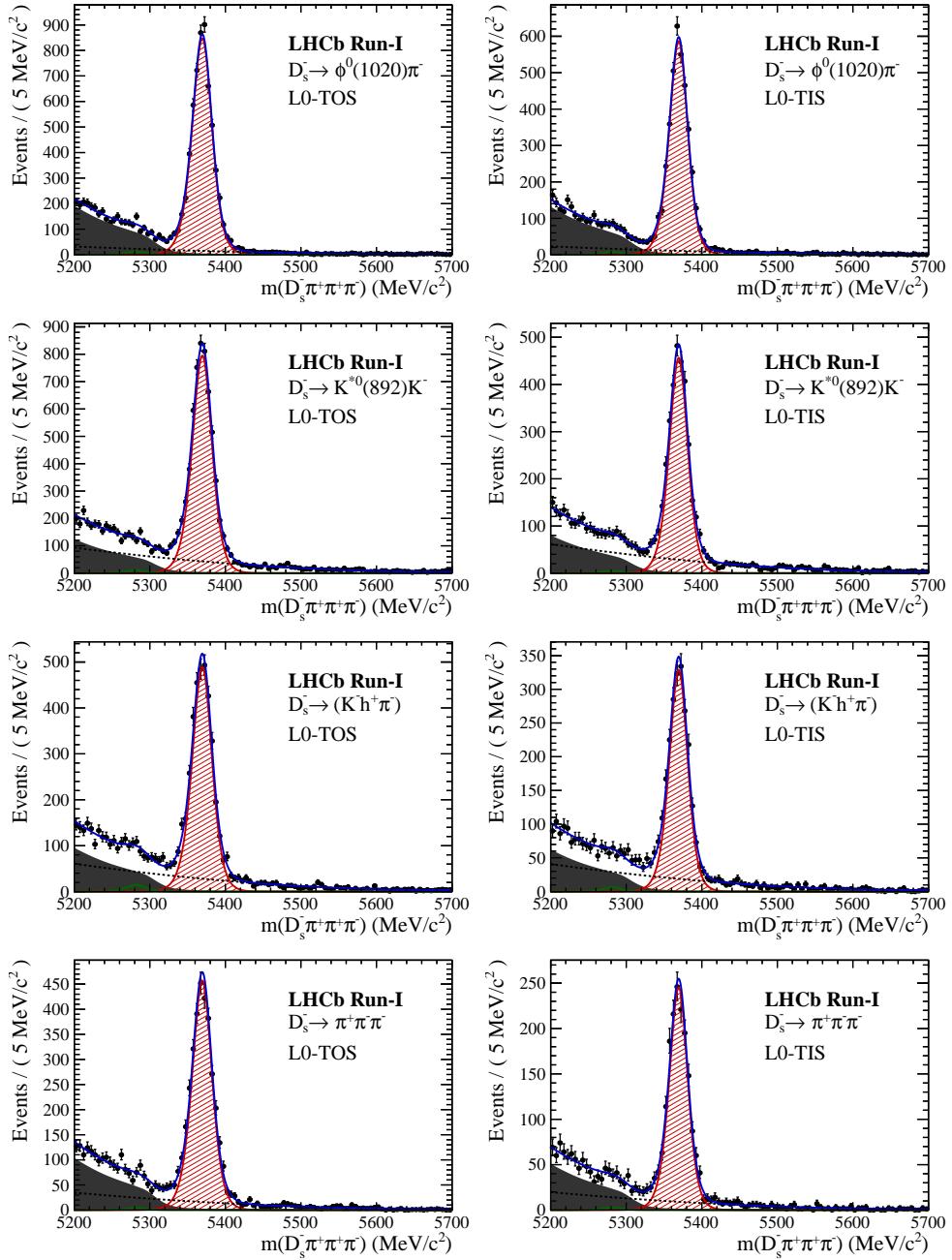


Figure B.1: Invariant mass distributions of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates for Run-I data.

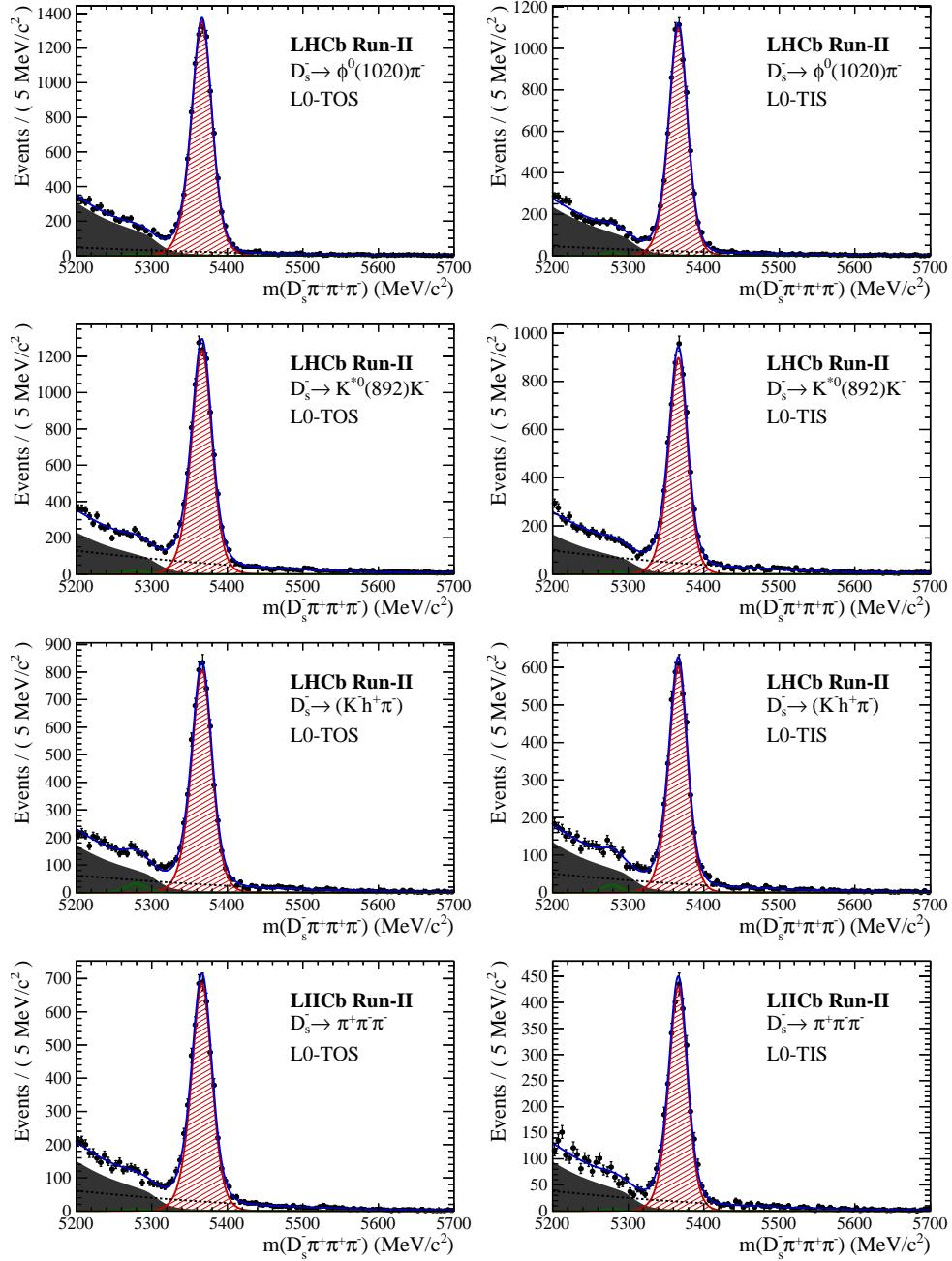


Figure B.2: Invariant mass distributions of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates for Run-II data.

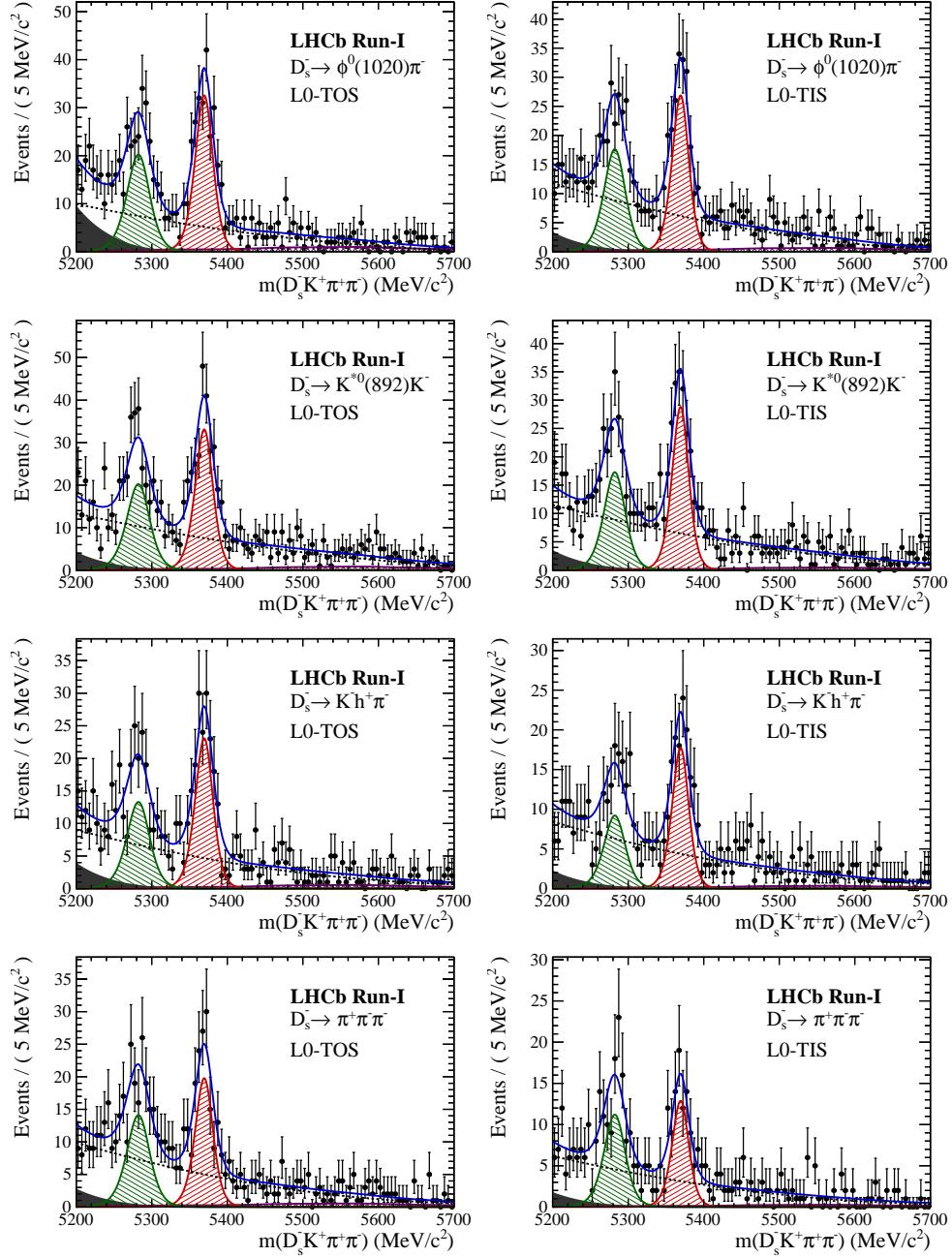


Figure B.3: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-I data.

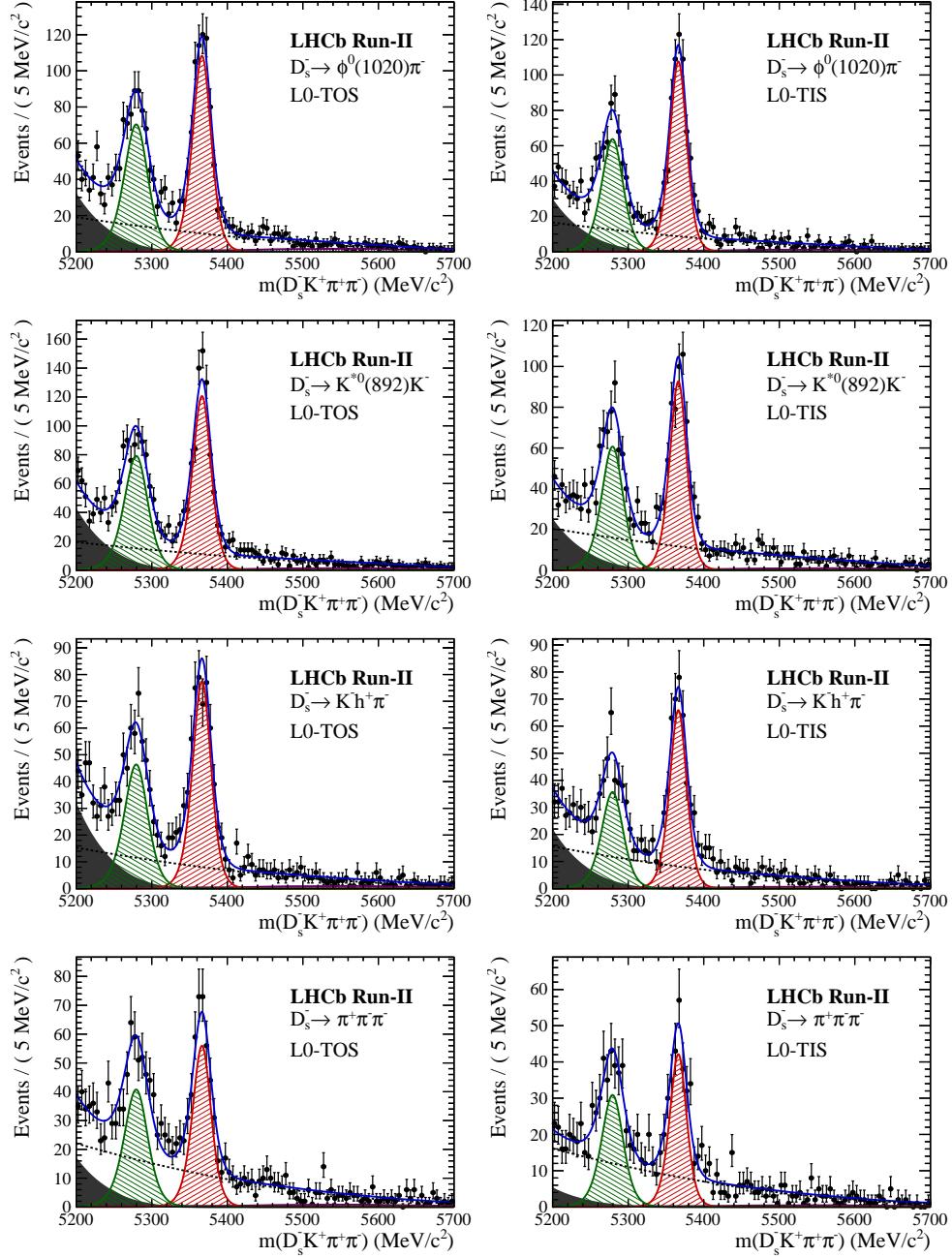


Figure B.4: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-II data.

1053 D Decay-time Resolution fits

1054 This section contains all fits to the distributions of the decay time difference Δt between
 1055 the true and the reconstructed decay time of the truth-matched B_s^0 candidates on MC.
 1056 The fits are performed in bins of the decay time error σ_t , where an adaptive binning
 1057 scheme is used to ensure that approximately the same number of events are found in each
 1058 bin.

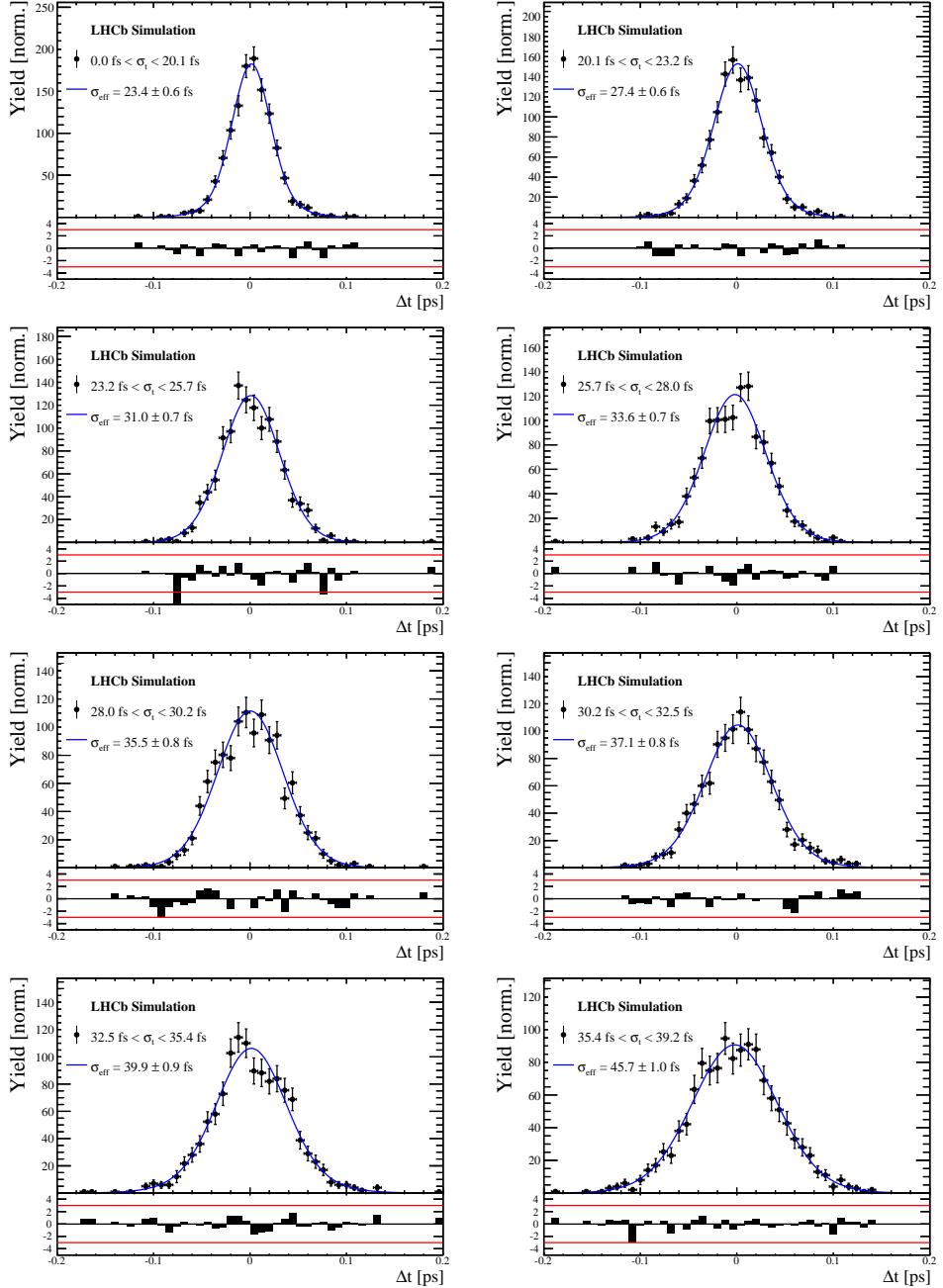


Figure C.1: Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K \pi \pi$ MC candidates in bins of the per-event decay time error estimate..

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.

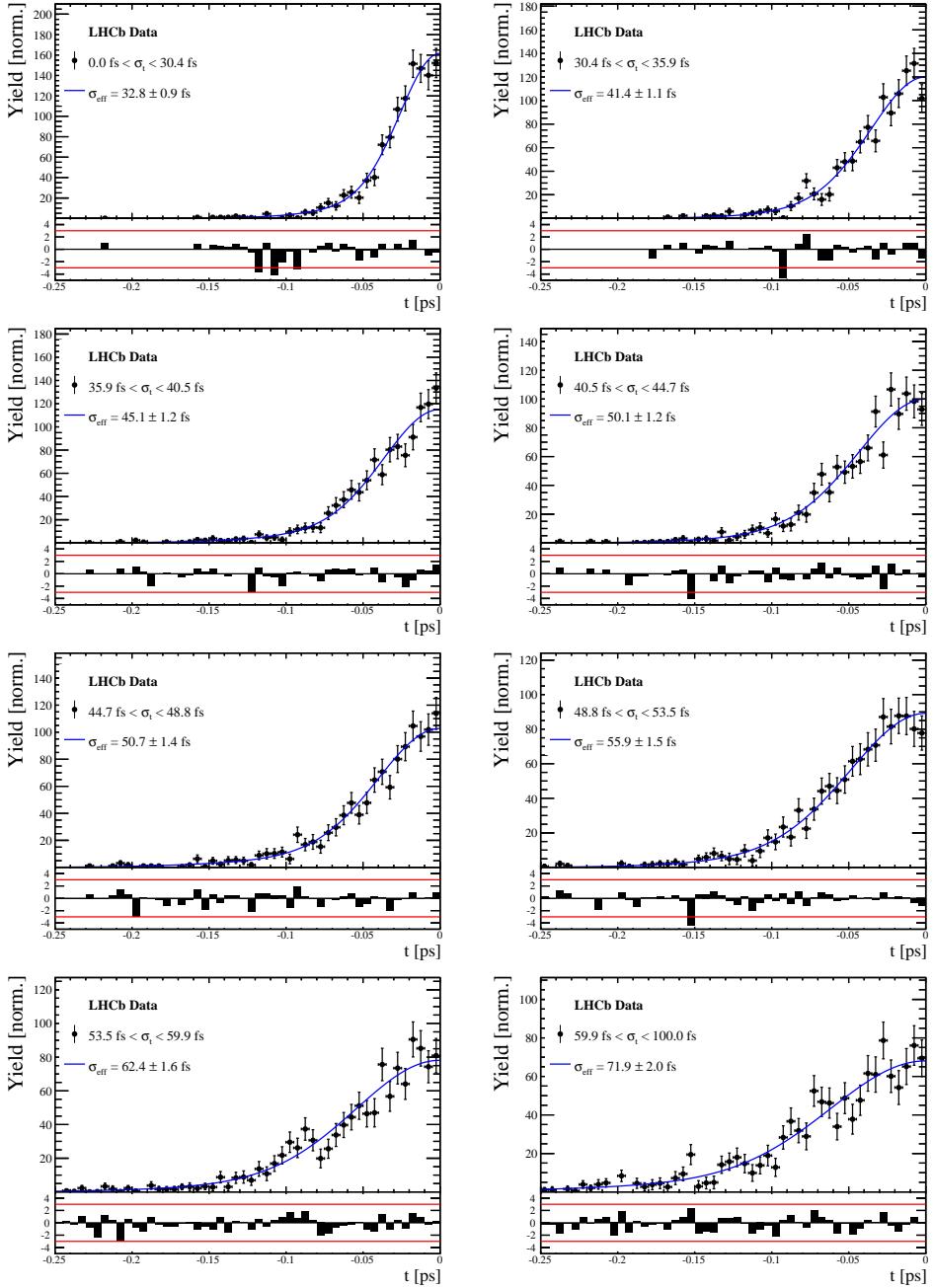


Figure C.2: Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks, for bins in the per-event decay time error estimate.

σ_t Bin [fs]	σ_1 [fs]	σ_2 [fs]	f_1	D	σ_{eff} [fs]
0.0 - 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.

1059 E Comparison of time-acceptance in subsamples

1060 Figure C.1 shows the spline coefficients obtained by fitting the decay-time distribution of
 1061 $B_s^0 \rightarrow D_s \pi\pi\pi$ data candidates in different subsamples. Sufficient agreement is observed
 1062 within a given data-taking period, while the acceptance shapes for Run-I and Run-II
 1063 data differ significantly. The fitted splines for the different D_s final states are in a good
 1064 agreement. The largest deviations are observed between the different L0 categories.

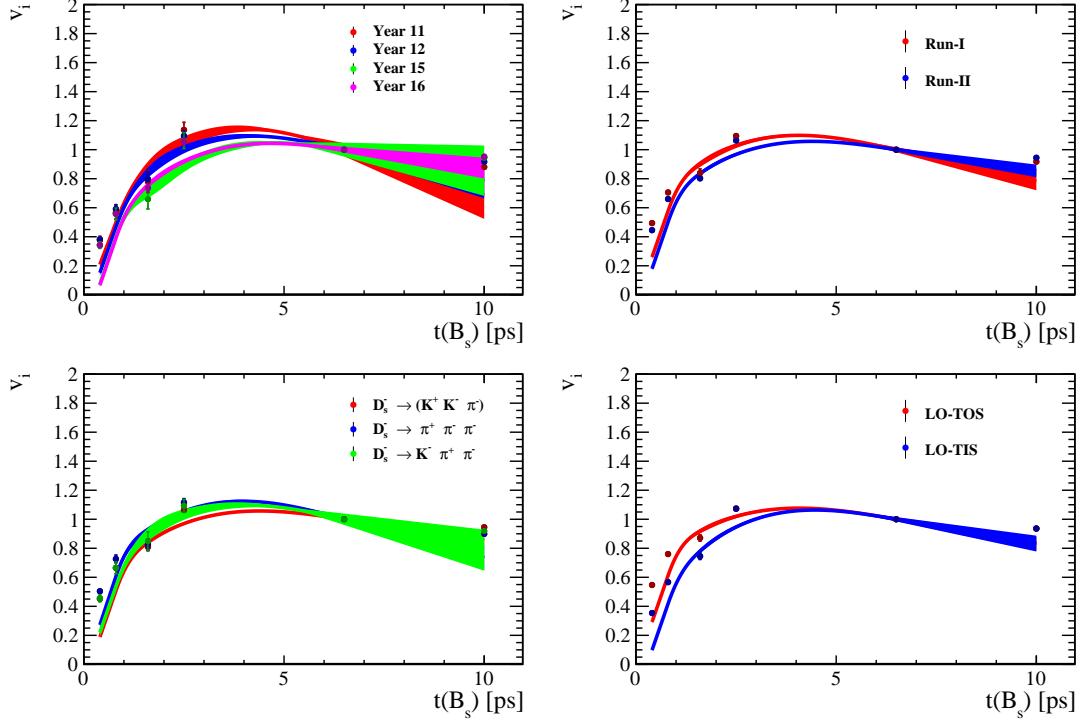


Figure C.1: Comparison of the spline coefficients (point with error bars) obtained from time-dependent fits to the $B_s^0 \rightarrow D_s \pi\pi\pi$ decay-time for different subsamples: (up-left) different years of data-taking; (up-right) different data-taking periods; (down-left) different D_s final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

1065 **F Spin Amplitudes**

1066 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	$B[D] \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\gamma}(T_1) L_{(2)\gamma}^\beta(T_2)$

1067 G Considered Decay Chains

1068 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
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ \rightarrow \pi^+ K^*(1430)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \omega(782)]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ \kappa]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \sigma]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow \sigma^0(D_s^- K^+)_S$
$B_s [S, P, D] \rightarrow \sigma^0(D_s^- K^+)_V$
$B_s \rightarrow \rho(770)^0(D_s^- K^+)_S$
$B_s [S, P, D] \rightarrow \rho(770)^0(D_s^- K^+)_V$
$B_s \rightarrow K^*(892)^0(D_s^- \pi^+)_S$
$B_s [S, P, D] \rightarrow K^*(892)^0(D_s^- \pi^+)_V$
$B_s \rightarrow (D_s^- K^+)_S (\pi^+ \pi^-)_S$

¹⁰⁶⁹ H MC corrections

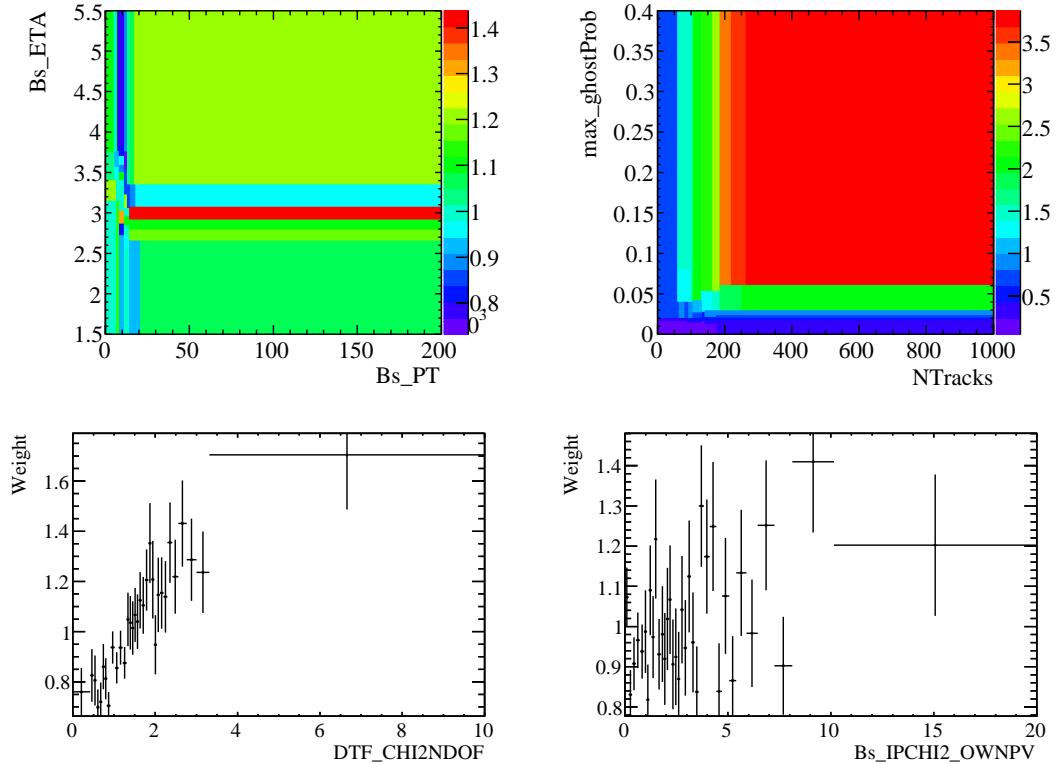


Figure C.1: Weights applied to correct for Data/MC differences.

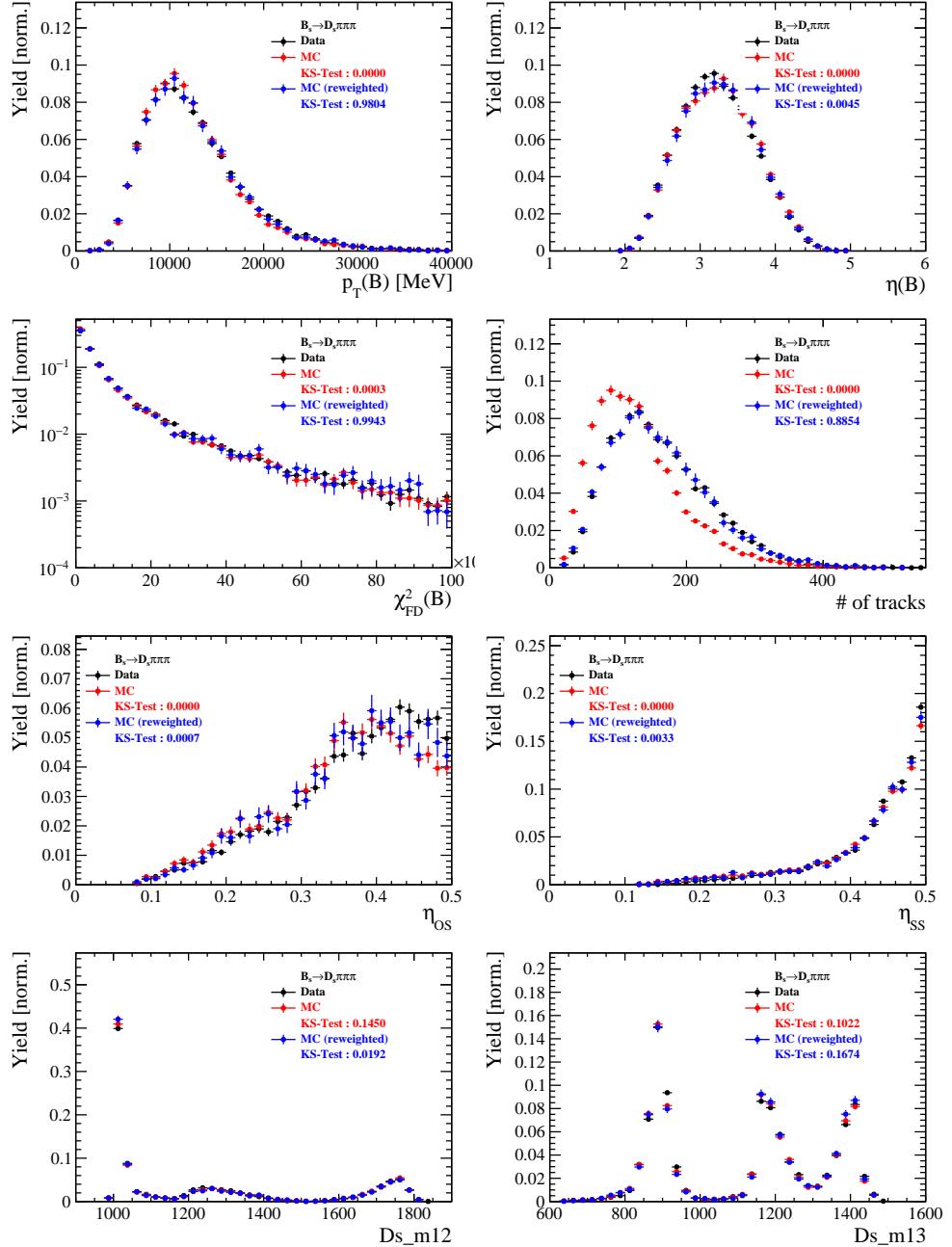


Figure C.2: Comparison of selected variables for $B_s \rightarrow D_s \pi\pi\pi$ decays.

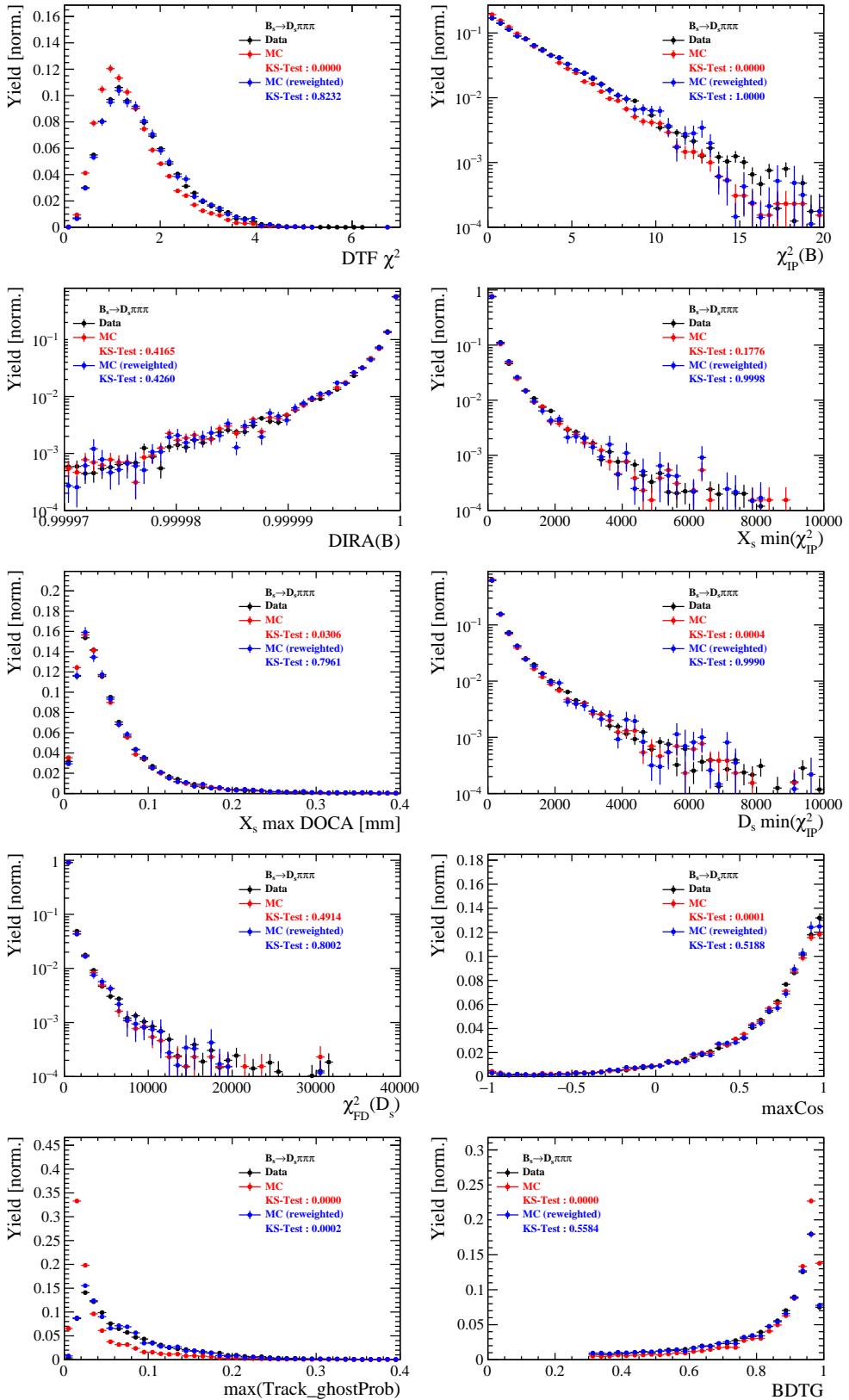


Figure C.3: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s \pi\pi\pi$ decays.

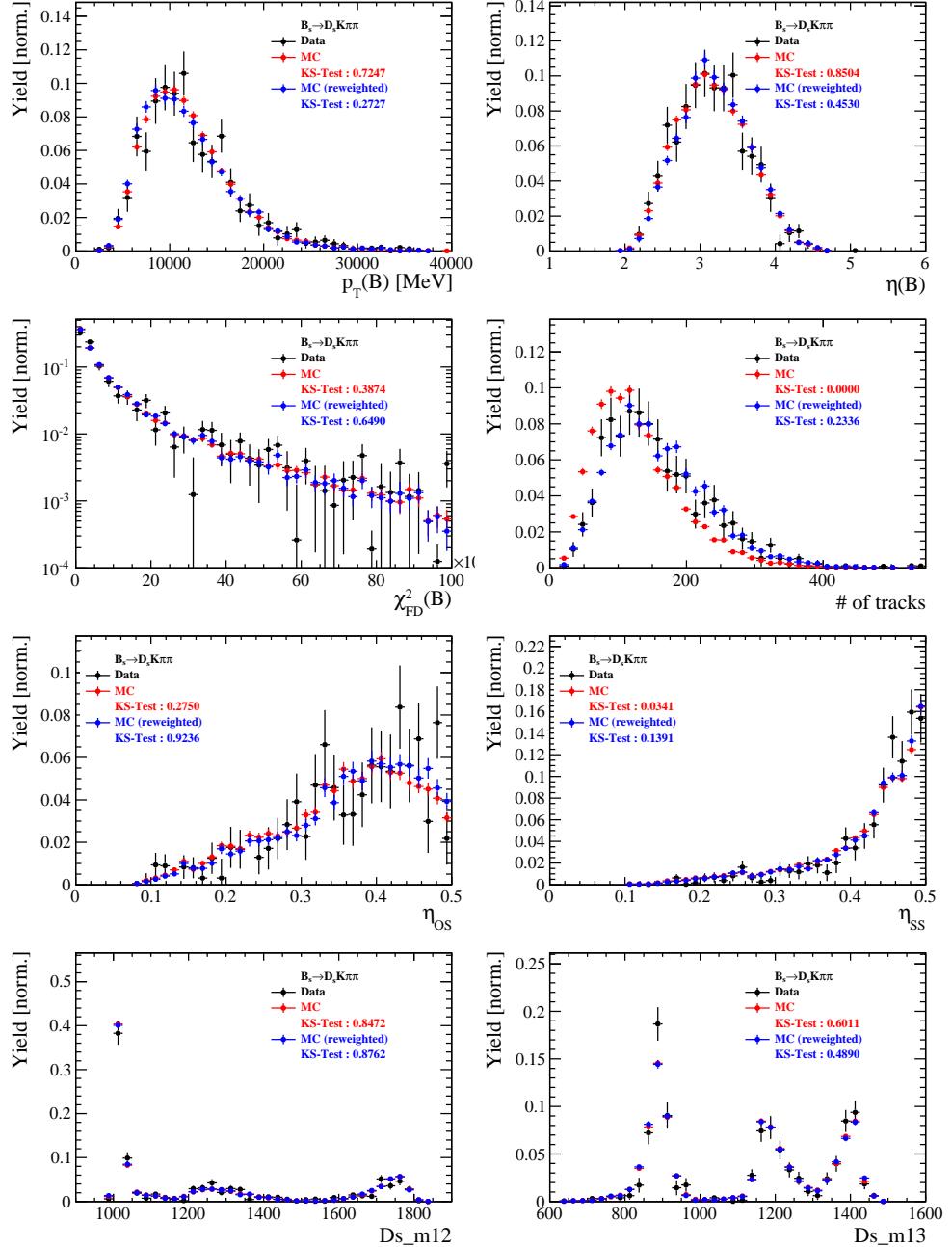


Figure C.4: Comparison of selected variables for $B_s \rightarrow D_s K\pi\pi$ decays.

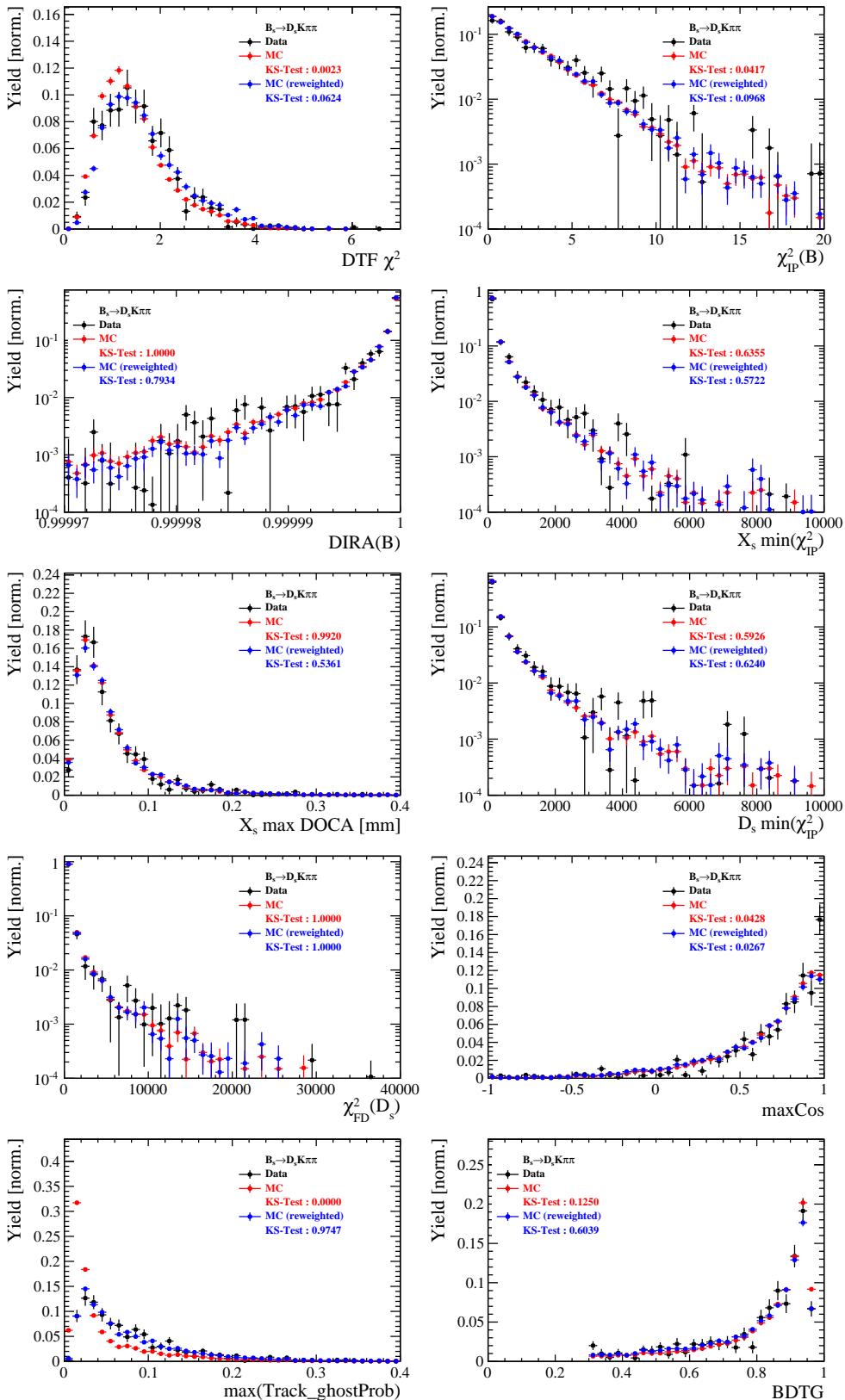


Figure C.5: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s K \pi \pi$ decays.

1070 I Data distributions

1071 I.1 Comparison of signal and calibration channel

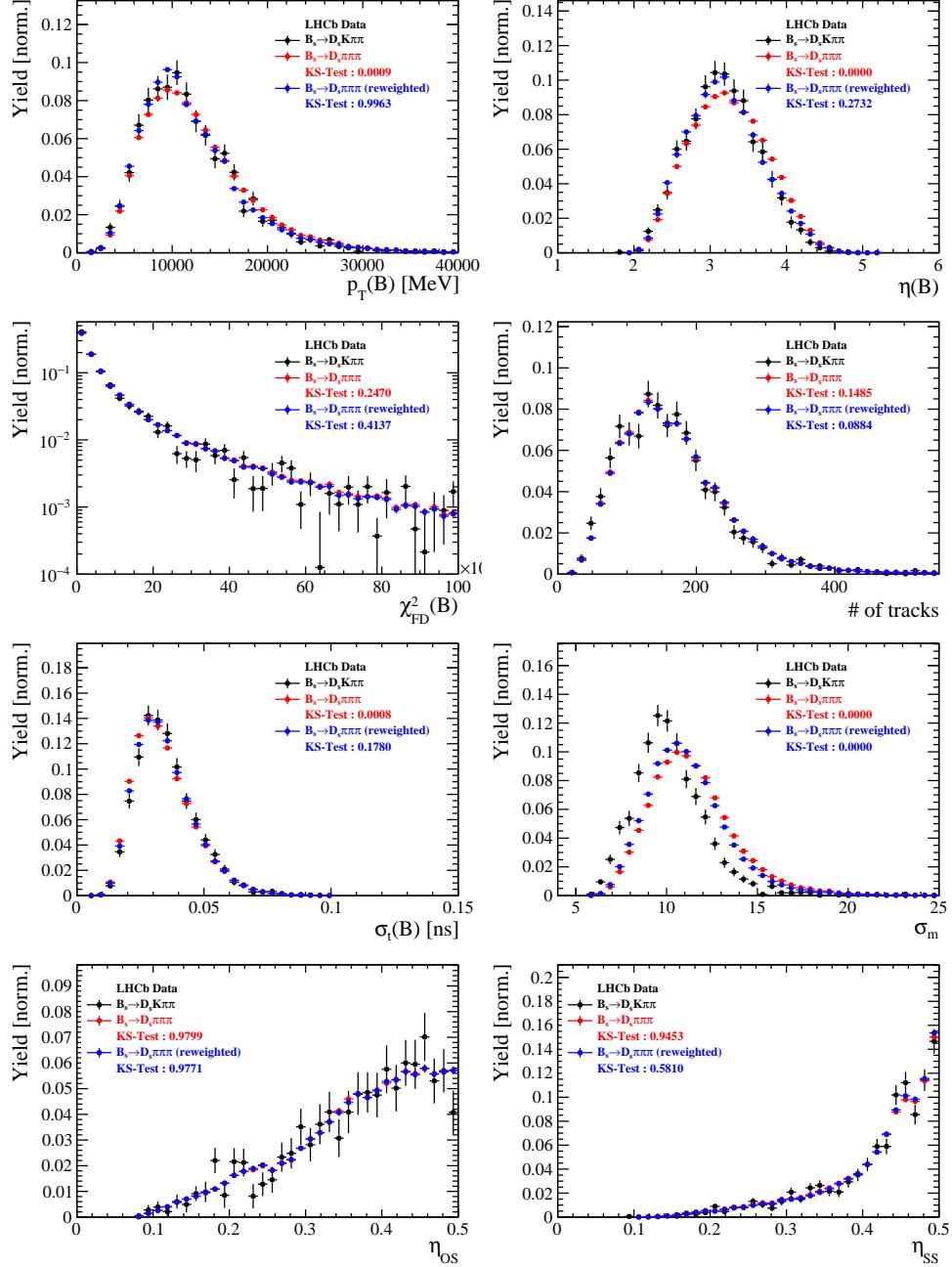


Figure C.1: Comparison of selected variables.

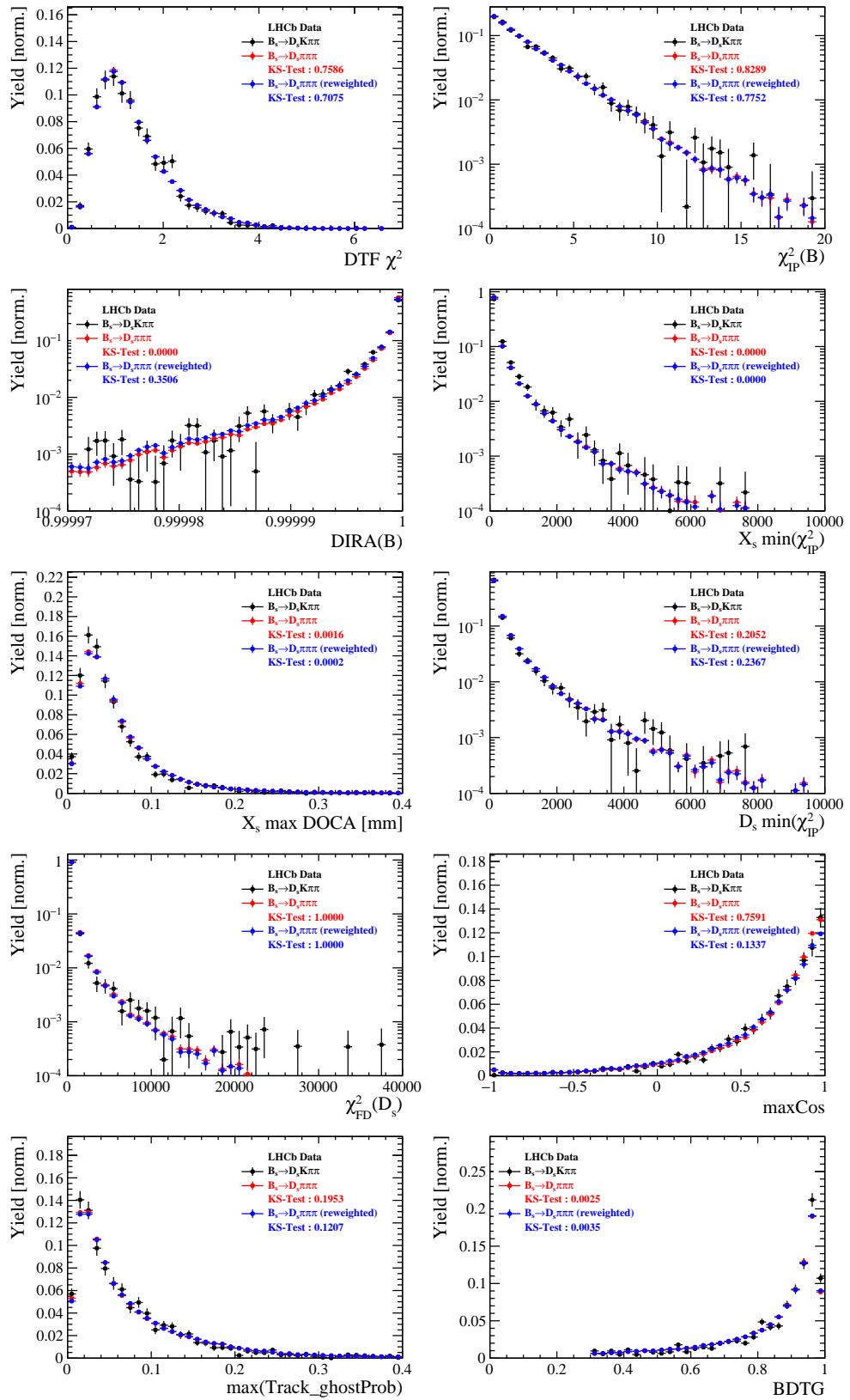


Figure C.2: Comparison of BDTG input variables and classifier response.

1072 I.2 Comparison of Run-I and Run-II data

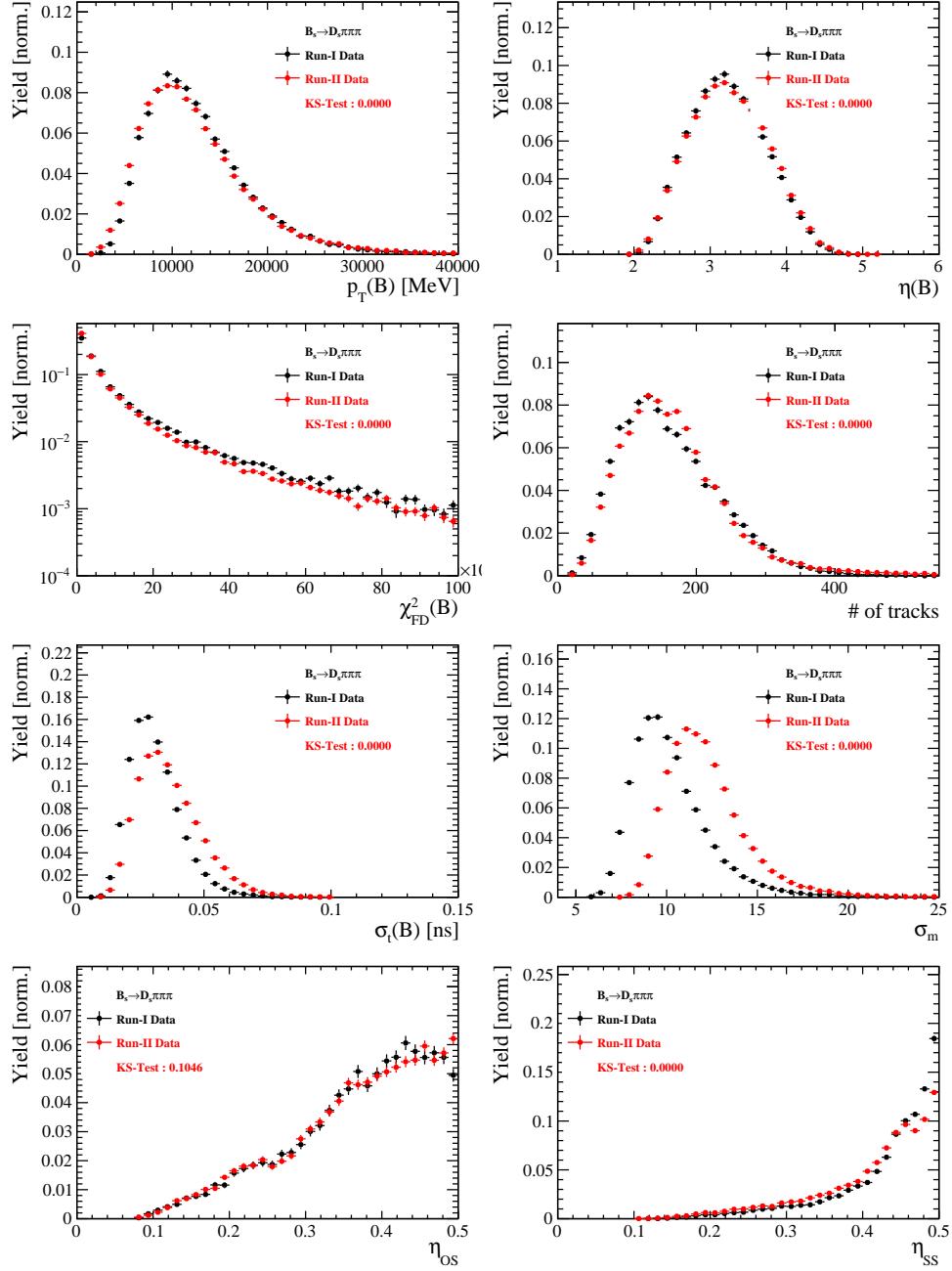


Figure C.3: Comparison of selected variables.

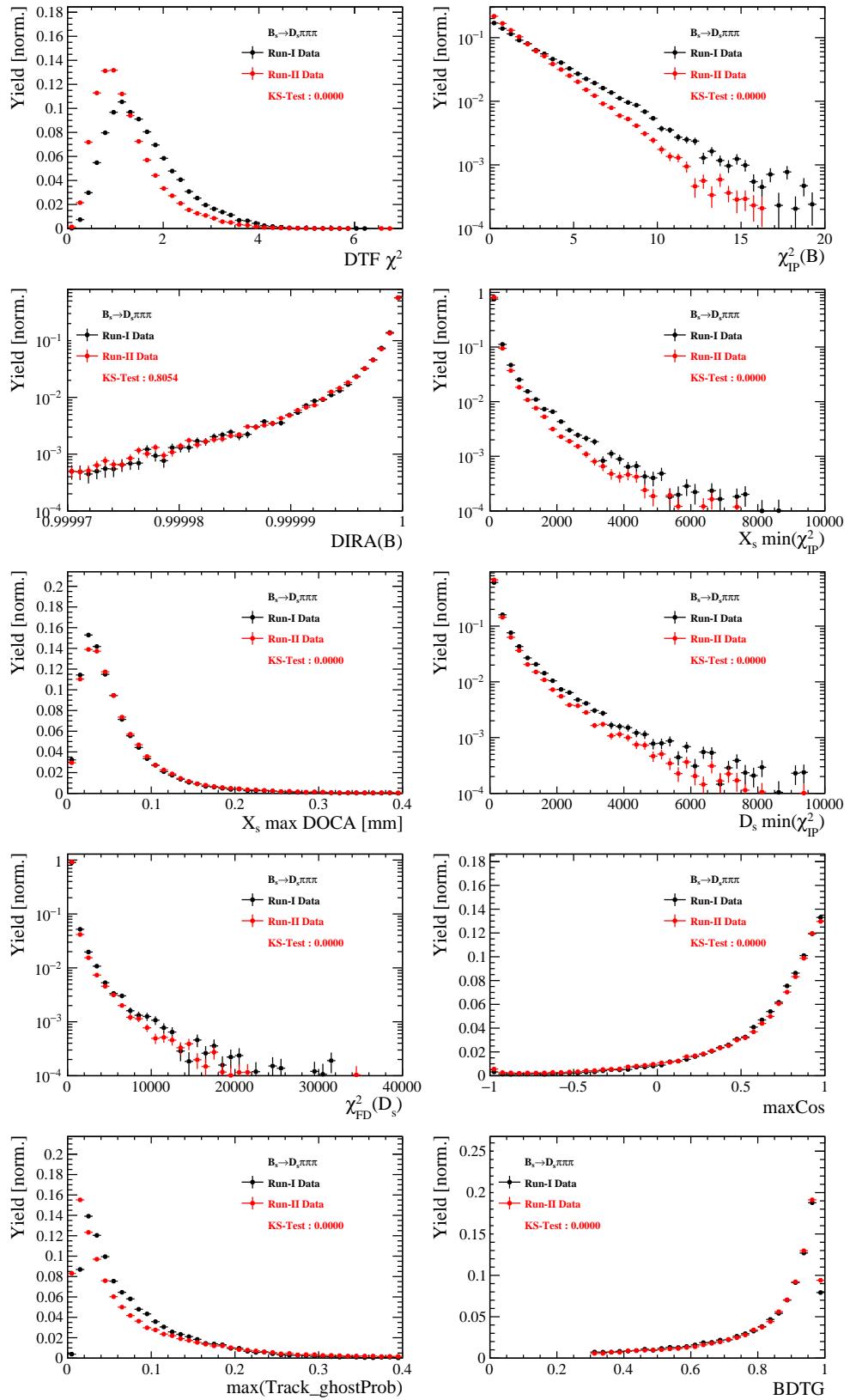


Figure C.4: Comparison of BDTG input variables and classifier response.

1073 I.3 Comparison of D_s final states

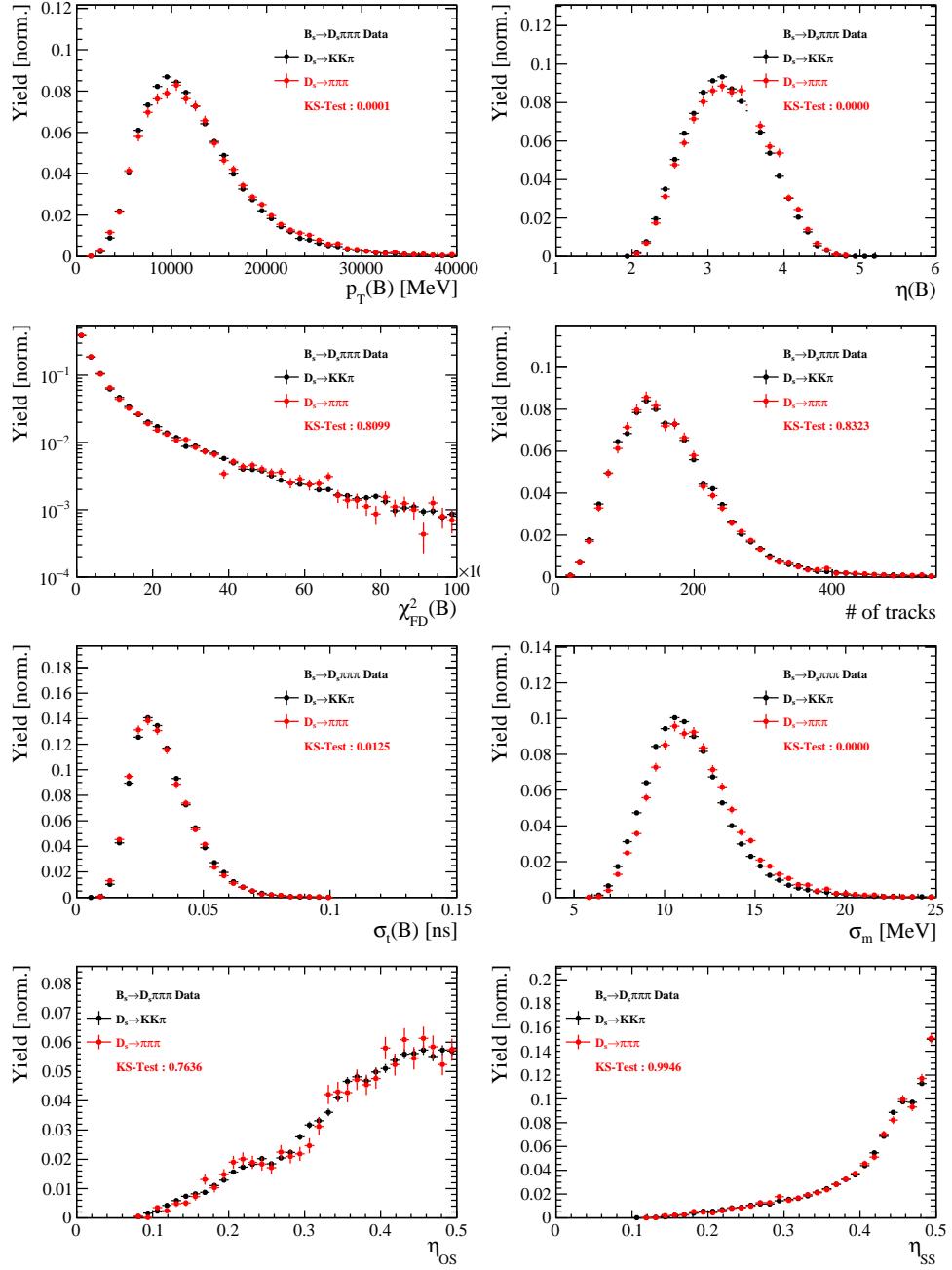


Figure C.5: Comparison of selected variables.

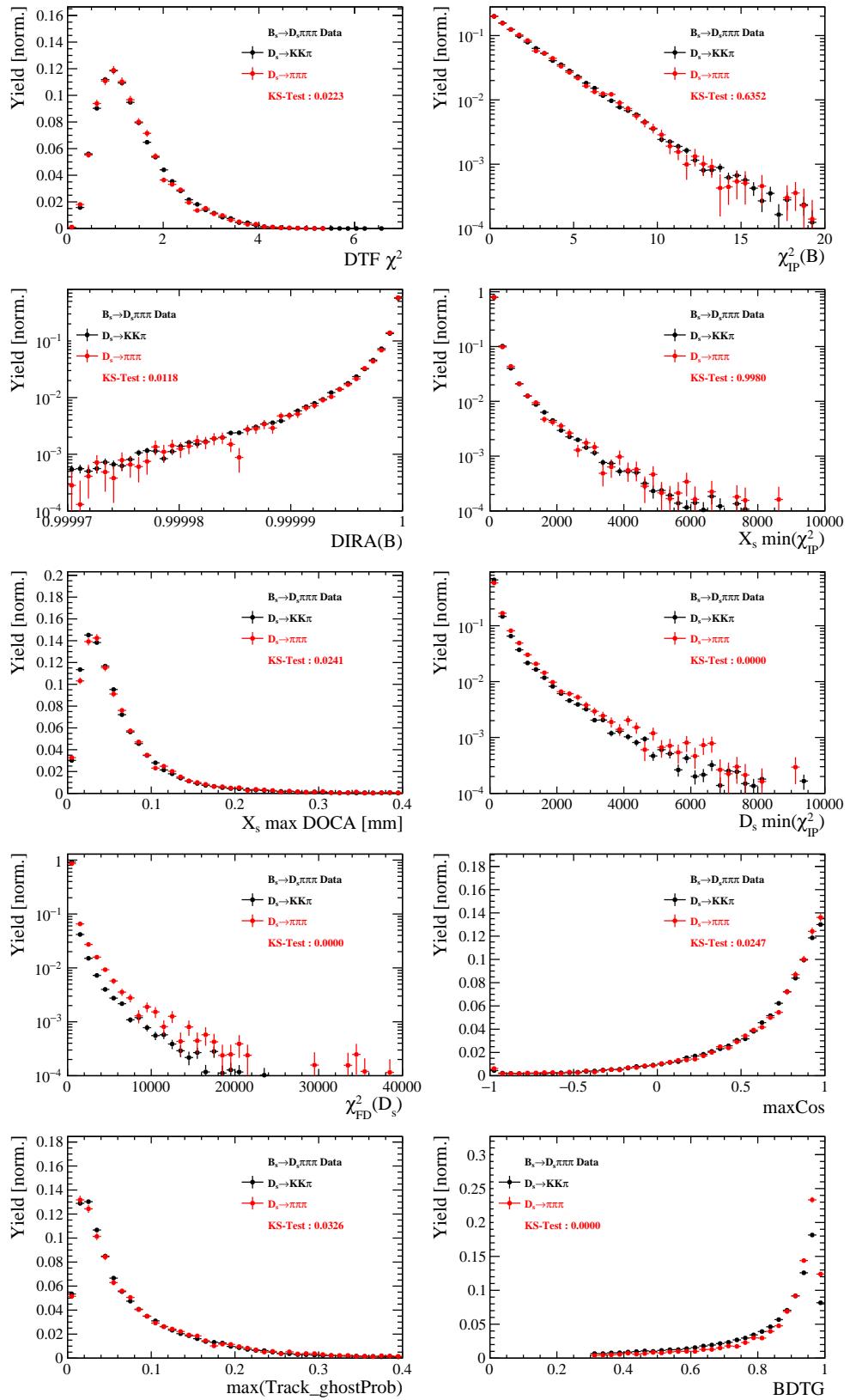


Figure C.6: Comparison of BDTG input variables and classifier response.

1074 I.4 Comparison of trigger categories

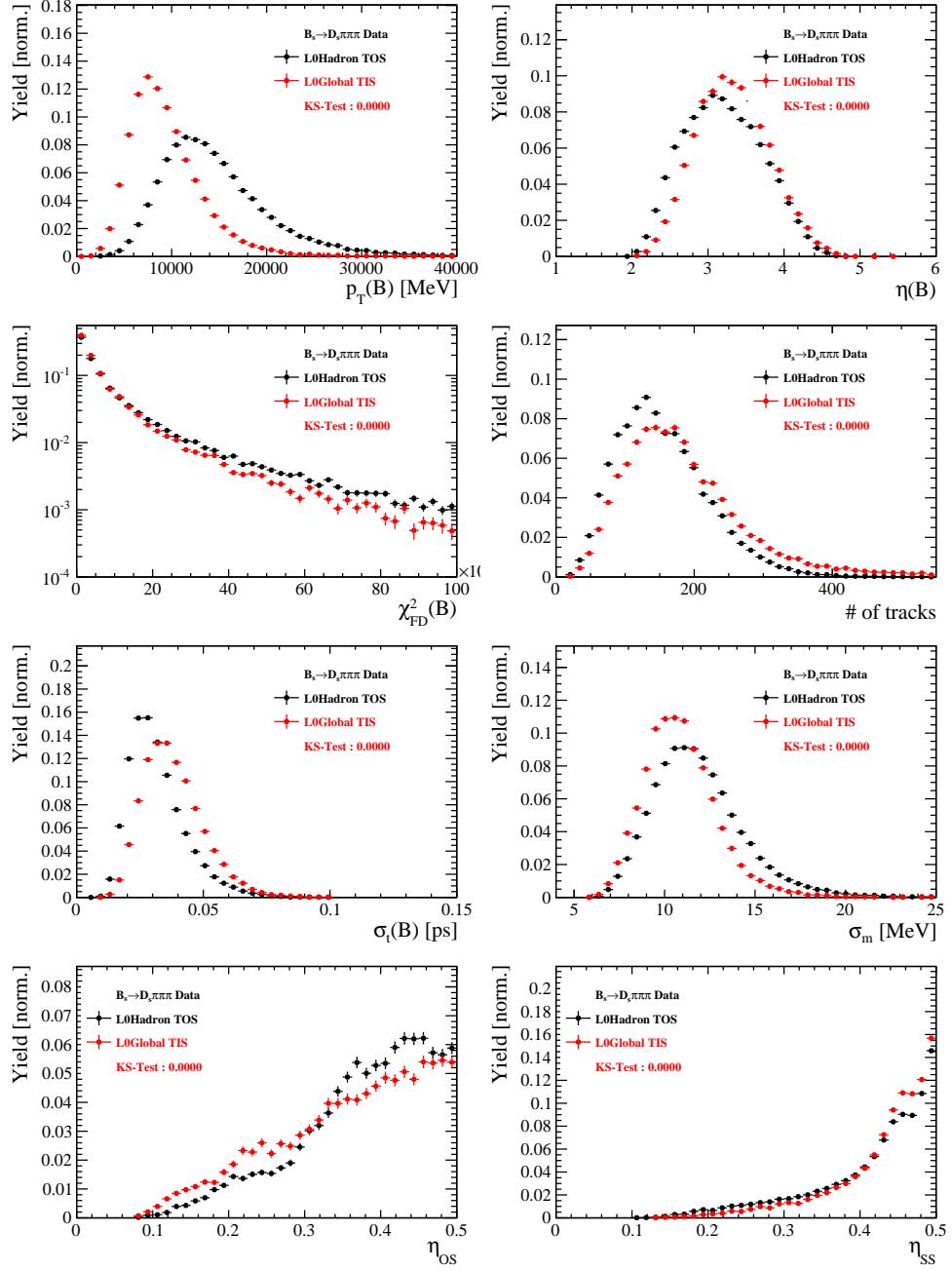


Figure C.7: Comparison of selected variables.

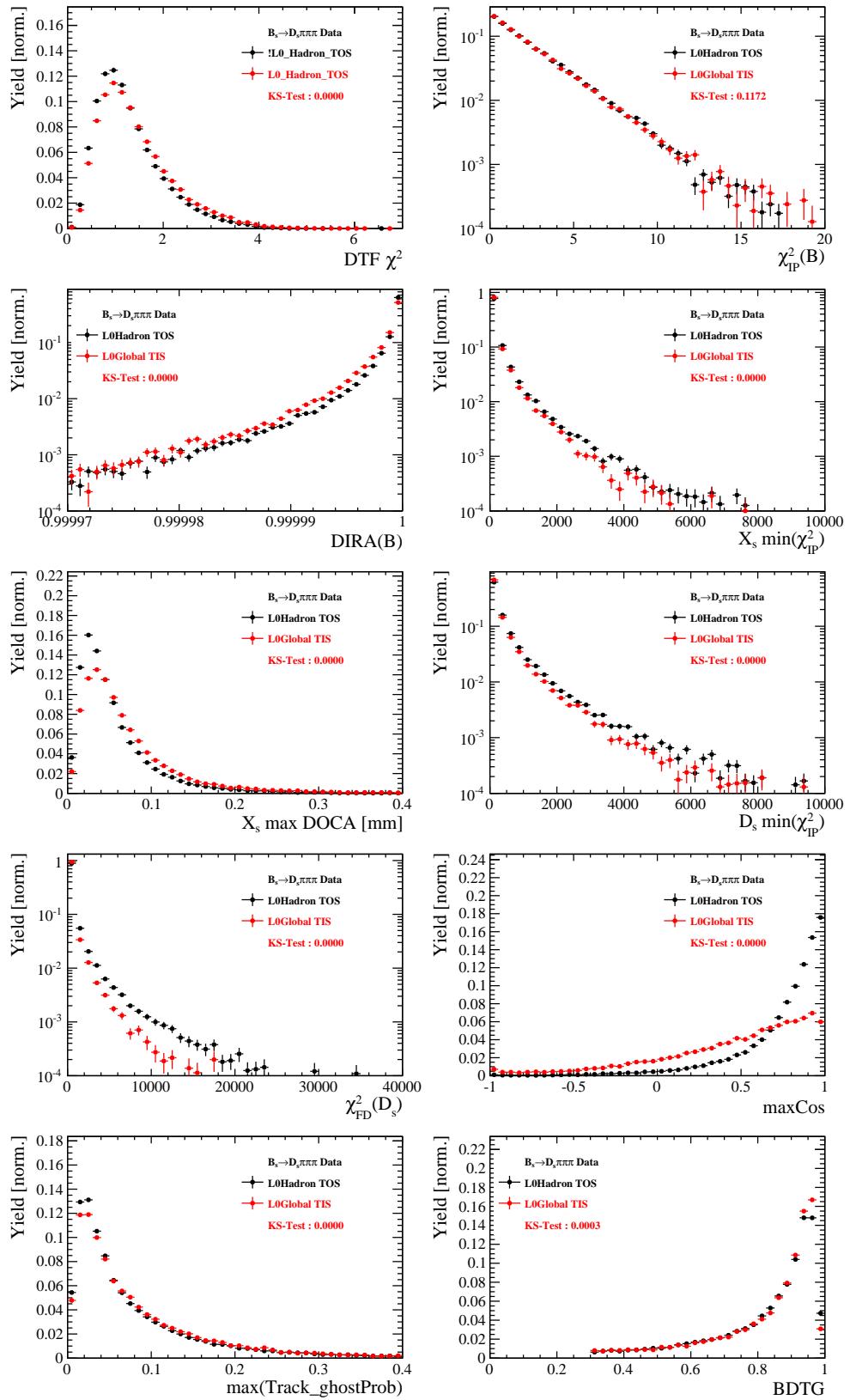


Figure C.8: Comparison of BDTG input variables and classifier response.

1075 I.5 Comparison of B_s and B_d decays

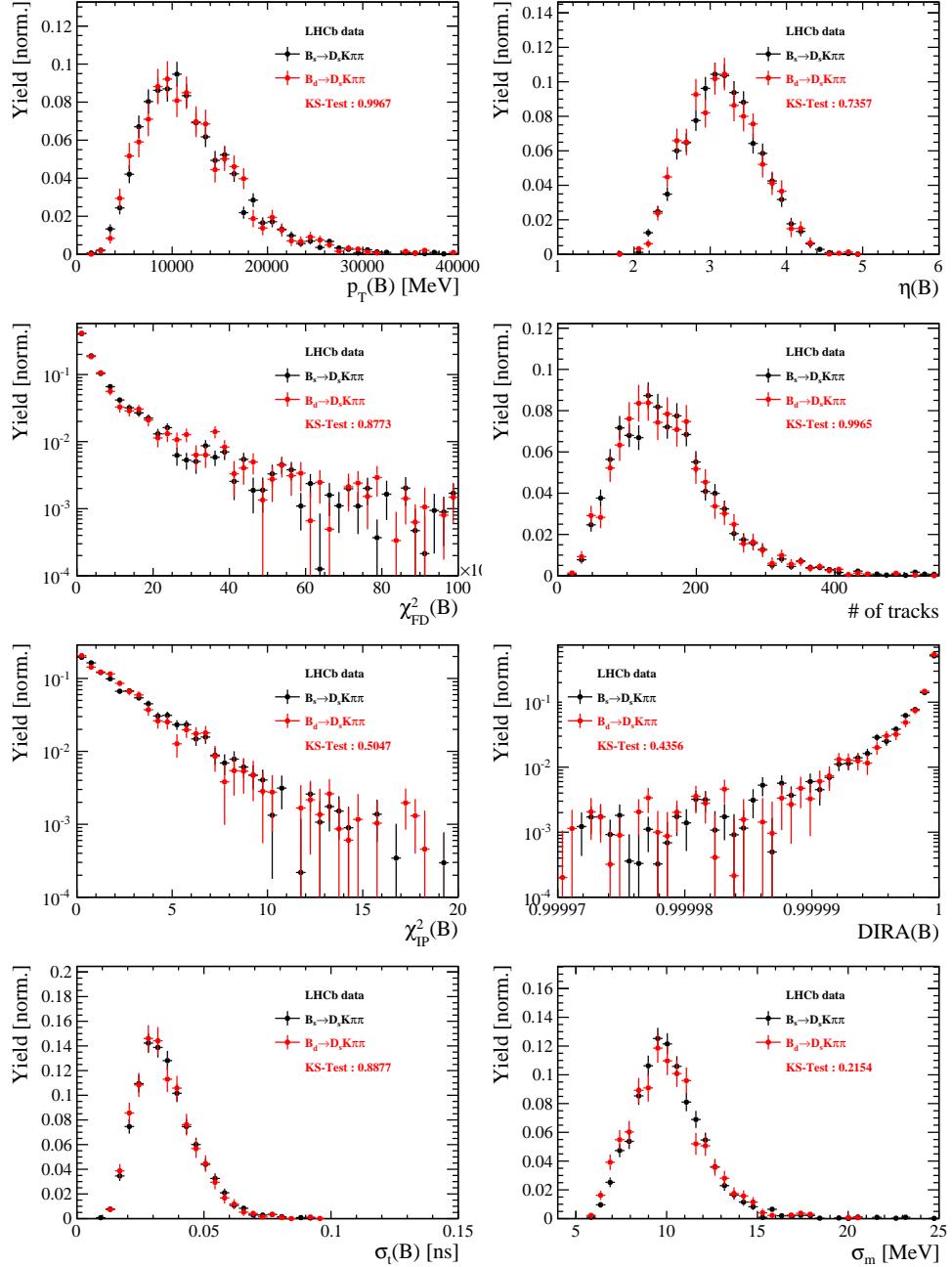


Figure C.9: Comparison of selected variables.

1076 **References**

- 1077 [1] R. Fleischer, *New strategies to obtain insights into CP violation through $B(s) \rightarrow D(s) \rightarrow K\pi$, $D(s)^* \rightarrow K\pi$, ... and $B(d) \rightarrow D \rightarrow \pi\pi$, $D^* \rightarrow \pi\pi$, ... decays*, Nucl.
1078 Phys. **B671** (2003) 459, arXiv:hep-ph/0304027.
- 1080 [2] K. De Bruyn *et al.*, *Exploring $B_s \rightarrow D_s^{(*)\pm} K^\mp$ Decays in the Presence of a Sizable Width Difference $\Delta\Gamma_s$* , Nucl. Phys. **B868** (2013) 351, arXiv:1208.6463.
- 1082 [3] S. Blusk, *First observations and measurements of the branching fractions for the decays $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* , .
- 1084 [4] LHCb, S. Blusk, *Measurement of the CP observables in $\bar{B}_s^0 \rightarrow D_s^+ K^-$ and first observation 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.
- 1086 [5] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 1087 [6] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation of production processes*, Annals of Physics **18** (1962), no. 2 198 .
- 1089 [7] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*, Phys. Rev. D **11** (1975) 3165.
- 1091 [8] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of Physics **108** (1977), no. 2 454 .
- 1093 [9] P. d'Argent *et al.*, *Amplitude Analyses of $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$ and $D^0 \rightarrow K^+K^-\pi^+\pi^-$ Decays*, JHEP **05** (2017) 143, arXiv:1703.08505.
- 1095 [10] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.
- 1097 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo Cimento Series 10 **34** (1964), no. 6 1644.
- 1099 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.
- 1101 [13] D. V. Bugg, *The mass of the σ pole*, Journal of Physics G Nuclear Physics **34** (2007) 151, arXiv:hep-ph/0608081.
- 1103 [14] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-dominance prediction for $\rho \rightarrow e^+e^-$* , Phys. Rev. Lett. **21** (1968) 244.
- 1105 [15] M. S. et al. *Search for CP violation in the $D^0 \rightarrow KK\pi\pi$ decay through a full amplitude analysis*, LHCb-ANA-2017-064.
- 1107 [16] S. M. Flatté, *Coupled-channel analysis of the $\pi\eta$ and KK systems near KK threshold*, Physics Letters B **63** (1976), no. 2 224 .
- 1109 [17] BES Collaboration, M. Ablikim *et al.*, *Resonances in $J/\psi \rightarrow \phi\pi^+\pi^-$ and ϕK^+K^-* , Phys. Lett. **B607** (2005) 243, arXiv:hep-ex/0411001.

- [18] W. Dunwoodie, Fits to $K\pi$ $I = \frac{1}{2}$ S -wave amplitude and phase data.
- [19] D. Aston *et al.*, *A Study of K - π + Scattering in the Reaction K - $p \rightarrow K$ - π + n at 11- GeV/c* , Nucl. Phys. **B296** (1988) 493.
- [20] BaBar, B. Aubert *et al.*, *Dalitz-plot analysis of the decays $B^\pm \rightarrow K^\pm \pi^\mp \pi^\pm$* , Phys. Rev. **D72** (2005) 072003, arXiv:hep-ex/0507004, [Erratum: Phys. Rev.D74,099903(2006)].
- [21] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in $D^0 \rightarrow K^\mp \pi^\pm \pi^\pm \pi^\mp$ decays*, Eur. Phys. J. **C78** (2018), no. 6 443, arXiv:1712.08609.
- [22] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- [23] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys. Rev. **60** (1941) 61.
- [24] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys. Rev. D **57** (1998) 431.
- [25] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of ψ decay to mesons*, Eur. Phys. J. **A16** (2003) 537, arXiv:hep-ph/0211457.
- [26] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. **D51** (1995) 2247.
- [27] J.-J. Zhu, *Explicit expressions of spin wave functions*, arXiv:hep-ph/9906250.
- [28] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput. Phys. Commun. **180** (2009) 1847, arXiv:0805.2956.
- [29] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [30] M. Karbach and M. Kenzie, *Gammacombo package*, <http://gammacombo.hepforge.org/web/HTML/index.html>, 2014.
- [31] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.
- [32] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.
- [33] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [34] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.
- [35] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*, Submitted to: JHEP (2017) arXiv:1712.07428.

- 1145 [36] LHCb collaboration, L. Zhang, *Measurements of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$*
 1146 *decays in the low $K^+ K^-$ mass range with 13 TeV data*, LHCb-ANA-2017-028.
- 1147 [37] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and*
 1148 *τ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots
 1149 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 1150 [38] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*
 1151 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 1152 [39] LHCb Collaboration, R. Aaij *et al.*, *Measurement of angular and cp asymmetries*
 1153 *in $D^0 \rightarrow \pi^+ \pi^- \mu^+ \mu^-$ and $D^0 \rightarrow K^+ K^- \mu^+ \mu^-$ decays*, Phys. Rev. Lett. **121** (2018)
 1154 091801.
- 1155 [40] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at*
 1156 *LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- 1157 [41] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the*
 1158 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- 1159 [42] LHCb, R. Aaij *et al.*, *Measurement of B^0 , B_s^0 , B^+ and Λ_b^0 production asymmetries in 7*
 1160 *and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- 1161 [43] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of*
 1162 *the $K\pi$ Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-
 1163 2012-027, CERN, Geneva, Feb, 2013.
- 1164 [44] A. Davis *et al.*, *Measurement of the instrumental asymmetry for $K^- \pi^+$ -pairs at LHCb*
 1165 *in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN,
 1166 Geneva, Mar, 2018.
- 1167 [45] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly*
 1168 *forbidden transitions in $D \rightarrow K(S)$, $K(L) + \pi$'s decays*, Phys. Lett. **B349** (1995)
 1169 363, arXiv:hep-ph/9502238.
- 1170 [46] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*,
 1171 Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- 1172 [47] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude*
 1173 *analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- 1174 [48] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal
 1175 Statistical Society, Series B **58** (1994) 267.
- 1176 [49] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- 1177 [50] H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on
 1178 Automatic Control **19** (1974) 716.
- 1179 [51] CLEO Collaboration, M. Artuso *et al.*, *Amplitude analysis of $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$* ,
 1180 Phys. Rev. **D85** (2012) 122002, arXiv:1201.5716.

- 1181 [52] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime*
1182 *and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986,
1183 DESY-F31-86-02.
- 1184 [53] D. Hill, M. John, and P. Gandini, *A study of partially reconstructed $B^\pm \rightarrow D^{*0} h^\pm$*
1185 *decays using the $D^0 \rightarrow K\pi, KK, \pi\pi$ final states*, .
- 1186 [54] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins
1187 University Press, Baltimore, MD, USA, 1996.