

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
6.2.1	Comparison of acceptance in subsamples	33
6.2.2	Results	34
6.3	Phasespace acceptance	38
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	56
11 Systematic uncertainties	59
11.1 Models for B_s^0 mass distribution	59
11.1.1 Signal model	59
11.1.2 Background model	59
11.1.3 Description of misidentified background	60
11.1.4 Systematic effect on observables	60
11.2 Decay-time acceptance	61
11.2.1 Varition of knot positions	61
11.2.2 Variation of spline coefficients	61
11.3 Decay-time resolution	62
11.4 Tagging calibration	63
11.5 Summary of systematic uncertainties	63
A Stripping and Trigger cuts	67
B Details of multivariate classifier	69
C Detailed mass fits	73
D Decay-time Resolution fits	77
E Spin Amplitudes	81
F Considered Decay Chains	82
G MC corrections	83
H Data distributions	88
H.1 Comparison of signal and calibration channel	88
H.2 Comparison of Run-I and Run-II data	90
H.3 Comparison of D_s final states	92
H.4 Comparison of trigger categories	94
H.5 Comparison of B_s and B_d decays	96
References	97
References	97

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 $B_s^0 \rightarrow D_s \pi\pi\pi$.

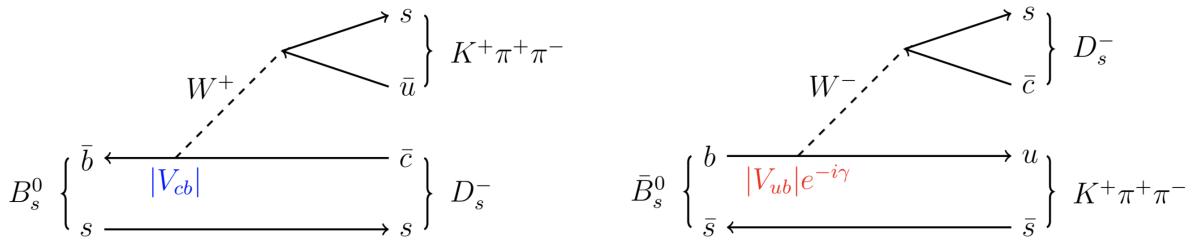


Figure 1.1: Feynman diagram for $B_s^0/\bar{B}_s^0 \rightarrow D_s^- K^+ \pi^+ \pi^-$ decays.

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:

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

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

$$\begin{aligned} m_{ij}^2 &= (p_i + p_j)^2, \\ m_{ijk}^2 &= (p_i + p_j + p_k)^2 \end{aligned} \quad (2.4)$$

³³ or acoplanarity and helicity angles. It is however important to take into account that,
³⁴ while m_{12}^2, m_{23}^2 are sufficient to fully describe a three-body decay, the obvious extension
³⁵ to four-body decays with m_{ij}^2, m_{ijk}^2 requires additional care, as these variables alone are
³⁶ insufficient to describe the parity-odd moments possible in four-body kinematics. In
³⁷ practice, we do not need to choose a particular five-dimensional basis, but use the full
³⁸ four-vectors of the decay in our analysis. The dimensionality is handled by the phase
³⁹ space element which can be written in terms of any set of five independent kinematic
⁴⁰ observables, $\mathbf{x} = (x_1, \dots, x_5)$, as

$$d\Phi_4 = \phi_4(\mathbf{x}) d^5 x, \quad (2.5)$$

41 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,
 42 the four-body phase space density function is not flat in the usual kinematic variables.
 43 Therefore, an analytic expression for ϕ_4 is taken from Ref. [5].

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

$$\mathcal{A}(\bar{B}_s^0 \rightarrow \bar{f}) = \mathcal{A}_{\bar{f}}^c(\mathbf{x}) = \mathcal{A}_f^c(\bar{\mathbf{x}}) \quad (2.6)$$

$$\mathcal{A}(B_s^0 \rightarrow \bar{f}) = \mathcal{A}_{\bar{f}}^u(\mathbf{x}) = \mathcal{A}_f^u(\bar{\mathbf{x}}) \quad (2.7)$$

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

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

$$\int \frac{\Gamma(x, t, q, f)}{e^{-\Gamma_s t} dt d\Phi_4} d\Phi_4 \propto \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + q f C \cos(\Delta m_s t) \\ + D_f \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - q S_f \sin(\Delta m_s t) \quad (2.8)$$

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

$$C = \frac{1 - r^2}{1 + r^2} \quad (2.9)$$

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

$$S_f = f \frac{2 r \kappa \sin(\delta - q_f (\gamma - 2\beta_s))}{1 + r^2} \quad (2.11)$$

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

$$\kappa e^{i\delta} \equiv \frac{\int \mathcal{A}_f^c(x)^* \mathcal{A}_f^u(x) d\Phi_4}{\sqrt{\int |\mathcal{A}_f^c(x)|^2 d\Phi_4} \sqrt{\int |\mathcal{A}_f^u(x)|^2 d\Phi_4}} \quad (2.12)$$

$$r \equiv \frac{\sqrt{\int |\mathcal{A}_f^u(x)|^2 d\Phi_4}}{\sqrt{\int |\mathcal{A}_f^c(x)|^2 d\Phi_4}}. \quad (2.13)$$

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

59 **2.2 Amplitude model**

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

$$A_i(\mathbf{x}) = B_{L_{B_s}}(\mathbf{x}) [B_{L_{R_1}}(\mathbf{x}) T_{R_1}(\mathbf{x})] [B_{L_{R_2}}(\mathbf{x}) T_{R_2}(\mathbf{x})] S_i(\mathbf{x}). \quad (2.14)$$

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

70 **2.2.1 Form Factors and Resonance Lineshapes**

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

$$\begin{aligned} B_0(q) &= 1, \\ B_1(q) &= 1/\sqrt{1 + (q r_{\text{BW}})^2}, \\ B_2(q) &= 1/\sqrt{9 + 3(q r_{\text{BW}})^2 + (q r_{\text{BW}})^4}. \end{aligned} \quad (2.15)$$

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

$$T(s) = \frac{1}{m_0^2 - s - i m_0 \Gamma(s)}, \quad (2.16)$$

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

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

$$\Gamma_{R \rightarrow AB}^{(2)}(s) = \Gamma_0 \frac{m_0}{\sqrt{s}} \left(\frac{q}{q_0} \right)^{2L+1} \frac{B_L(q)^2}{B_L(q_0)^2}, \quad (2.17)$$

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

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

$$\Gamma_{R \rightarrow ABC}^{(3)}(s) = \frac{1}{2\sqrt{s}} \int |A_{R \rightarrow ABC}|^2 d\Phi_3, \quad (2.18)$$

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. (2.14). In particular, it includes form factors and propagators of intermediate two-body resonances.

Both Eq. (2.17) and Eq. (2.18) 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,

$$\Gamma(s) = \sum_i g_i \Gamma_i(s), \quad (2.19)$$

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. [21]. 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 [15], where the coupling constants as well as the mass and width are taken from a measurement performed by the BES Collaboration [16]. 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 [17–19]:

$$T_{Lass}(s) = \frac{\sqrt{s}}{q \cot \delta_L - iq} + e^{2i\delta_L} \frac{m_0 \Gamma_0 \frac{m_0}{q_0}}{m_0^2 - s - i m_0 \Gamma_0 \frac{m_0}{\sqrt{s}} \frac{q}{q_0}} \quad (2.20)$$

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. [17, 18]. Equation (2.17) 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. [20]. 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 2.18 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.

122 2.2.2 Spin Densities

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

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

135 The spin projection operator $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$, for a resonance R , with spin $S =$
 136 $\{0, 1, 2\}$, and four-momentum p_R , is given by [26]:

$$\begin{aligned} P_{(0)}^{\mu\nu}(p_R) &= 1 \\ P_{(1)}^{\mu\nu}(p_R) &= -g^{\mu\nu} + \frac{p_R^\mu p_R^\nu}{p_R^2} \\ P_{(2)}^{\mu\nu\alpha\beta}(p_R) &= \frac{1}{2} \left[P_{(1)}^{\mu\alpha}(p_R) P_{(1)}^{\nu\beta}(p_R) + P_{(1)}^{\mu\beta}(p_R) P_{(1)}^{\nu\alpha}(p_R) \right] - \frac{1}{3} P_{(1)}^{\mu\nu}(p_R) P_{(1)}^{\alpha\beta}(p_R), \end{aligned} \quad (2.21)$$

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

139 For a decay process $R \rightarrow AB$, with relative orbital angular momentum L , between
 140 particle A and B , the angular momentum tensor is obtained by projecting the rank- L
 141 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
 142 subspace,

$$L_{(L)\mu_1 \dots \mu_L}(p_R, q_R) = (-1)^L P_{(L)\mu_1 \dots \mu_L \nu_1 \dots \nu_L}(p_R) q_R^{\nu_1} \dots q_R^{\nu_L}. \quad (2.22)$$

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

$$\begin{aligned} \varepsilon_{(S)}(R) &\equiv \varepsilon_{(S)}(p_R, \lambda_R), \\ P_{(S)}(R) &\equiv P_{(S)}(p_R), \\ L_{(L)}(R) &\equiv L_{(L)}(p_R, q_R). \end{aligned} \quad (2.23)$$

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

$$\begin{aligned} S_{R \rightarrow AB}(\mathbf{x}|L_{AB}, S_{AB}; \lambda_R, \lambda_A, \lambda_B) &= \varepsilon_{(S_R)}(R) X(S_R, L_{AB}, S_{AB}) L_{(L_{AB})}(R) \\ &\times \Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B), \end{aligned} \quad (2.24)$$

149 where

$$\Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B) = P_{(S_{AB})}(R) X(S_{AB}, S_A, S_B) \varepsilon_{(S_A)}^*(A) \varepsilon_{(S_B)}^*(B). \quad (2.25)$$

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

$$X(j_a, j_b, j_c) = \begin{cases} 1 & \text{if } j_a + j_b + j_c \text{ even} \\ \varepsilon_{\alpha\beta\gamma\delta} p_R^\delta & \text{if } j_a + j_b + j_c \text{ odd} \end{cases}, \quad (2.26)$$

158 where $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita symbol and j refers to the arguments of X defined in
 159 Eqs. 2.24 and 2.25. Its antisymmetric nature ensures the correct parity transformation
 160 behavior of the amplitude. The spin factor for a whole decay chain, for example $R \rightarrow$
 161 $(R_1 \rightarrow AB)(R_2 \rightarrow CD)$, is obtained by combining the two-body terms and performing a
 162 sum over all unobservable, intermediary spin projections

$$\sum_{\lambda_{R_1}, \lambda_{R_2}} S_{R \rightarrow R_1 R_2}(\mathbf{x}|L_{R_1 R_2}; \lambda_{R_1}, \lambda_{R_2}) S_{R_1 \rightarrow AB}(\mathbf{x}|L_{AB}; \lambda_{R_1}) S_{R_2 \rightarrow CD}(\mathbf{x}|L_{CD}; \lambda_{R_2}), \quad (2.27)$$

163 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
 164 pseudoscalar initial/final states are involved.

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

167 2.3 Fit implementation

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

$$\int |A_i(\mathbf{x})|^2 d\Phi_4 = 1. \quad (2.28)$$

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

$$\begin{aligned} & \int \left| \mathcal{A}_f^{c(u)}(\mathbf{x}) \right|^2 d\Phi_4 = 1 \\ & \arg \left(\int \mathcal{A}_f^c(\mathbf{x})^* \mathcal{A}_f^u(\mathbf{x}) d\Phi_4 \right) = 0. \end{aligned} \quad (2.29)$$

172 As a result, the average amplitude ratio and strong phase difference between the $b \rightarrow u$ and
173 $b \rightarrow c$ transitions can be introduced as direct fit parameters instead of derived quantities
174 that have to be calculated from Equation 2.12 after the fit. For the differential decay rate
175 follows:

$$\begin{aligned} \frac{\Gamma(x, t, q, f)}{e^{-\Gamma_s t} dt d\Phi_4} \propto & (|\mathcal{A}_f^c(\mathbf{x})|^2 + r^2 |\mathcal{A}_f^u(\mathbf{x})|^2) \cosh \left(\frac{\Delta\Gamma_s t}{2} \right) \\ & + q f (|\mathcal{A}_f^c(\mathbf{x})|^2 - r^2 |\mathcal{A}_f^u(\mathbf{x})|^2) \cos (\Delta m_s t) \\ & - 2 r \operatorname{Re} (\mathcal{A}_f^c(\mathbf{x})^* \mathcal{A}_f^u(\mathbf{x}) e^{i\delta - if(\gamma - 2\beta_s)}) \sinh \left(\frac{\Delta\Gamma_s t}{2} \right) \\ & - 2 q f r \operatorname{Im} (\mathcal{A}_f^c(\mathbf{x})^* \mathcal{A}_f^u(\mathbf{x}) e^{i\delta - if(\gamma - 2\beta_s)}) \sin (\Delta m_s t) \end{aligned} \quad (2.30)$$

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

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

191 2.4 Validation

192 The amplitude fit is performed by using the amplitude analysis tool **MINT2**. It was
 193 previously applied to analyze $D^0 \rightarrow 4\pi$ and $D^0 \rightarrow KK\pi\pi$ decays [9] which have an
 194 identical general spin structure (*i.e.* scalar to four scalar decay) to $B_s \rightarrow D_s K\pi\pi$ decays.
 195 In the course of the $D^0 \rightarrow hhhh$ analysis, the implementation of the amplitudes were
 196 extensively cross-checked against other available tools such as **qft++** [28], **AmpGen** [20]
 197 and where possible **EVTGEN** [29]. Since no additional line shapes or spin factors are
 198 needed for this analysis, we consider the amplitude calculation as fully validated.

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

207 The CP coefficients $C, D, \bar{D}, S, \bar{S}$ are the fit parameters in case of the phasespace-
 208 integrated pdf, which are converted after to the fit to the physical observables r, κ, δ and γ
 209 using the **GammaCombo** package [30]. The obtained 1-CL contours are shown in Fig. 2.3.
 210 The full pdf determines r, δ and γ directly. As shown in Tab. 2.2 and 2.3, the fit results
 211 are in excellent agreement with the generated input values. The phasespace-integrated fit
 212 is, in addition, performed with the **B2DX** fitter used for the $B_s \rightarrow D_s K$ analysis yielding
 213 identical results. Note though that some parts of the **B2DX** fitter have been taken over to
 214 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°

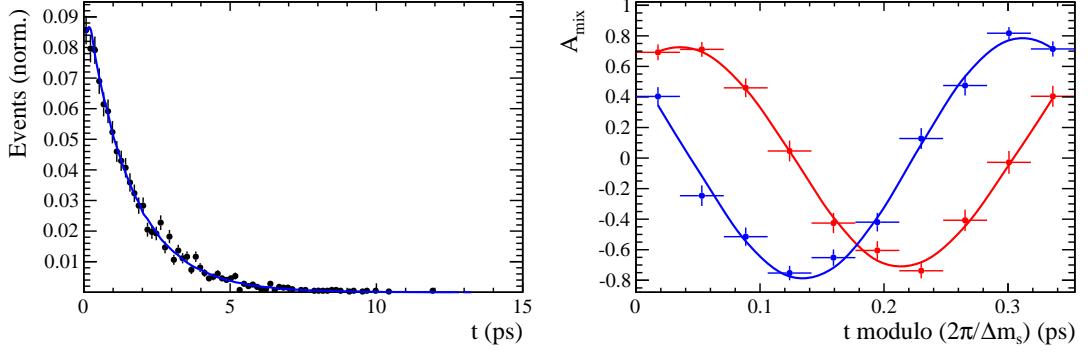


Figure 2.1: Left: Time distribution of $B_s \rightarrow D_s K \pi \pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (solid line). Right: Time-dependent asymmetry between mixed and unmixed events folded into one oscillation period for $D_s^- K^+ \pi \pi$ (red) and $D_s^+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

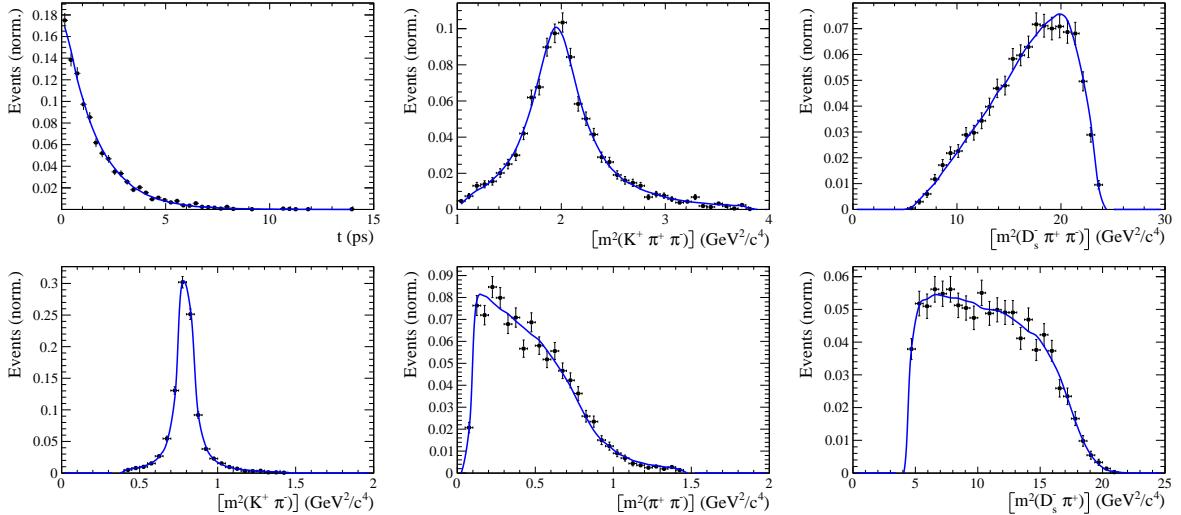


Figure 2.2: Time and invariant mass distributions of $B_s \rightarrow D_s K \pi \pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

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

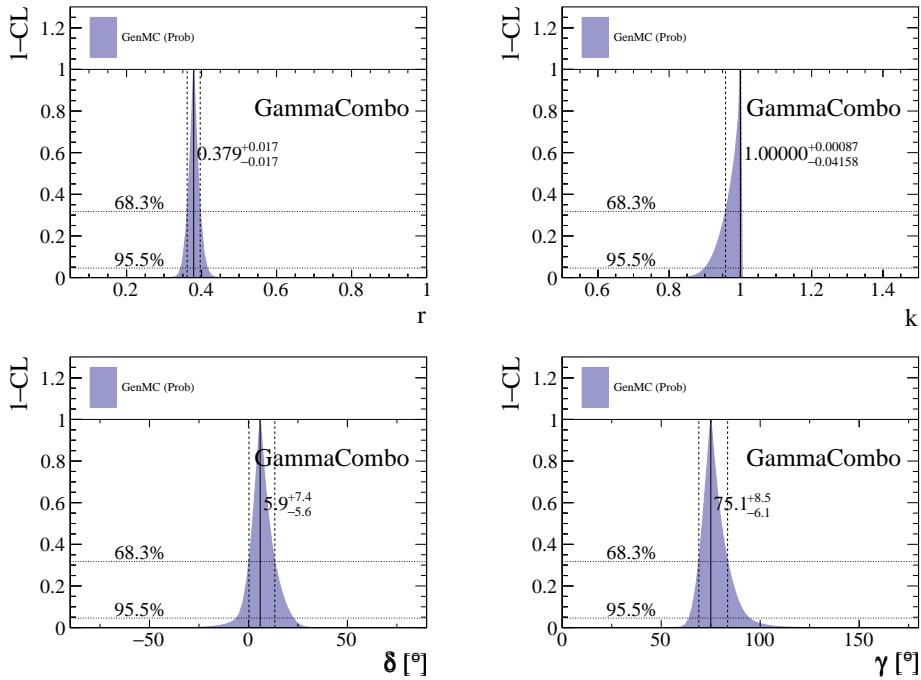


Figure 2.3: The 1-CL contours for the physical observable r, κ, δ and γ obtained with the phasespace-integrated fit to the EVTGEN toy sample.

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

215 3 Data samples and event selection

216 3.1 Stripping and Trigger selection

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

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

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

237 3.2 Offline selection

238 The offline selection, in particular the requirements on the D_s hadron, are guided by
239 the previous analyses of $B_s \rightarrow D_s K/\pi$, $B_d \rightarrow D^0 \pi$ as well as the branching fraction
240 measurement of $B_s^0 \rightarrow D_s K \pi \pi$ decays. Tables 3.1 and 3.2 summarize all selection
241 requirements which are described in the following. Throughout the note, we abbreviate
242 $B_s^0 \rightarrow D_s X_s (\rightarrow K \pi \pi)$ and $B_s^0 \rightarrow D_s X_d (\rightarrow \pi \pi \pi)$.

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

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

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

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

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

273 3.2.1 Phase space region

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

285 **3.2.2 Physics background vetoes**

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

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

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

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

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

296 Possible with $\bar{p} \rightarrow K^-$ misidentification, vetoed by requiring $m(K^+ K_p^- \pi^-) \neq$
 297 $m(\Lambda_c^-) \pm 40$ MeV or the K^- has to fulfill more stringent PID criteria depending
 298 on the resonant region (see Table 3.1).

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

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

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

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

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

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

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

308 Possible with $\pi^- \rightarrow K^-$ misidentification, vetoed by requiring $m(K_\pi^- \pi^+ \pi^-) \neq$
 309 $m(D^-) \pm 40$ MeV or $\text{PIDK}(K^+) > 15$.

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

311 Possible with $\bar{p} \rightarrow K^-$ misidentification, vetoed by requiring $m(K_p^- \pi^+ \pi^-) \neq$
 312 $m(\Lambda_c^-) \pm 40$ MeV or $\text{PIDK}(K^-) - \text{PIDp}(K^-) > 5$.

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

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

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

- 320 1. $X_s^+ \rightarrow K^+\pi^+\pi^-$:
- 321 (a) $B_s^0 \rightarrow D_s\pi\pi\pi$:
322 Possible with $\pi^+ \rightarrow K^+$ misidentification, suppressed with $\text{PIDK}(K^+) > 10$.
- 323 (b) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^+\pi^-)$:
324 Outside of considered phase-space region, see Sec. 3.2.1.
- 325 (c) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+K^-\pi^+)$:
326 To suppress $B_s^0 \rightarrow D_s^-K^+K^-\pi^+$ background, possible with $K^- \rightarrow \pi^-$ misiden-
327 tification, we require $\text{PIDK}(\pi^-) < 0$. In case the invariant mass of the $K^+\pi^+\pi^-$
328 system recomputed applying the kaon mass hypothesis to the pion is close to
329 the D_s mass ($m(K^+\pi^+\pi_K^-) = m(D_s) \pm 20$ MeV), we further tighten the cut to
330 $\text{PIDK}(\pi^-) < -5$.
- 331 2. $X_d^+ \rightarrow \pi^+\pi^+\pi^-$:
- 332 (a) $B_s^0 \rightarrow D_sK\pi\pi$:
333 Possible with single missID of $K^+ \rightarrow \pi^+$, suppressed with $\text{PIDK}(\pi^+) < 0$.
- 334 (b) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow \pi^+\pi^+\pi^-)$:
335 Outside of considered phase-space region, see Sec. 3.2.1.
- 336 (c) $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^-\pi^+)$:
337 Possible with single missID of $K^+ \rightarrow \pi^+$, vetoed by requiring $m(\pi^+\pi_K^+\pi^-) \neq$
338 $m(D_s) \pm 20$ MeV or $\text{PIDK}(\pi^+) < -5$ for both π^+ .

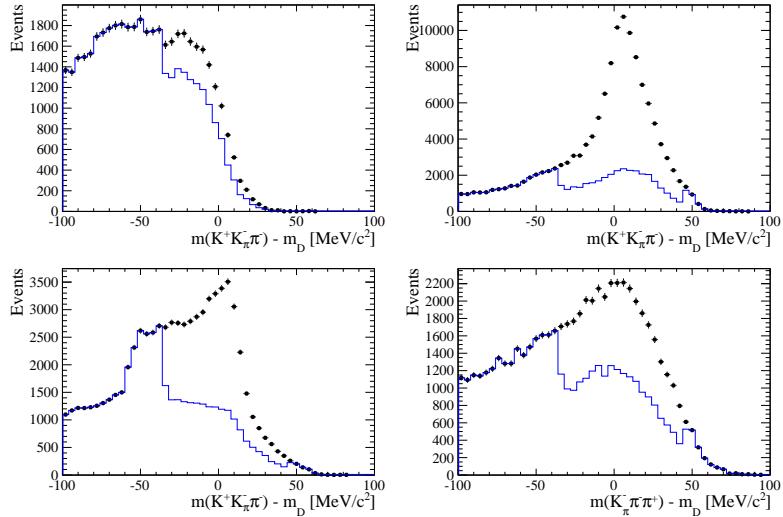


Figure 3.1: Background contributions from D^- decays where the π^- is misidentified as K^- . The D_s invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*(892)K$, $D_s \rightarrow KK\pi$ (non-resonant) and $D_s \rightarrow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

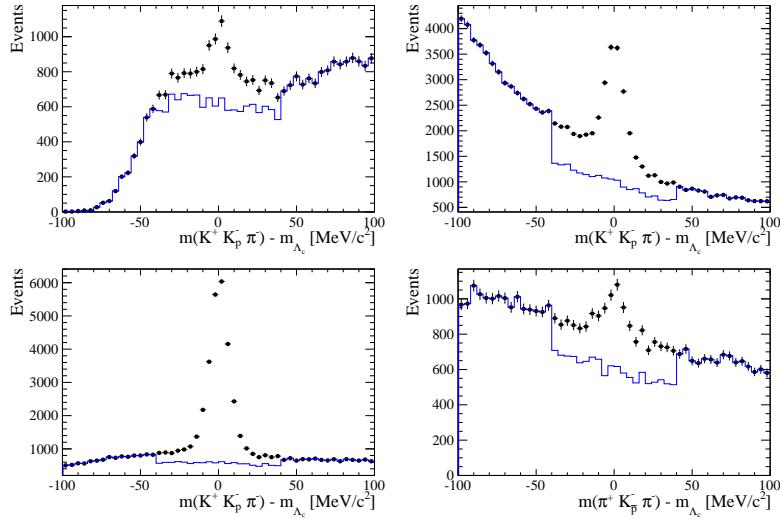


Figure 3.2: Background contributions from Λ_c decays where the \bar{p} is misidentified as K^- . The D_s invariant mass is recomputed applying the proton mass hypothesis to the kaon and shown for the $D_s \rightarrow \phi\pi$, $D_s \rightarrow K^*(892)K$, $D_s \rightarrow KK\pi$ (non-resonant) and $D_s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

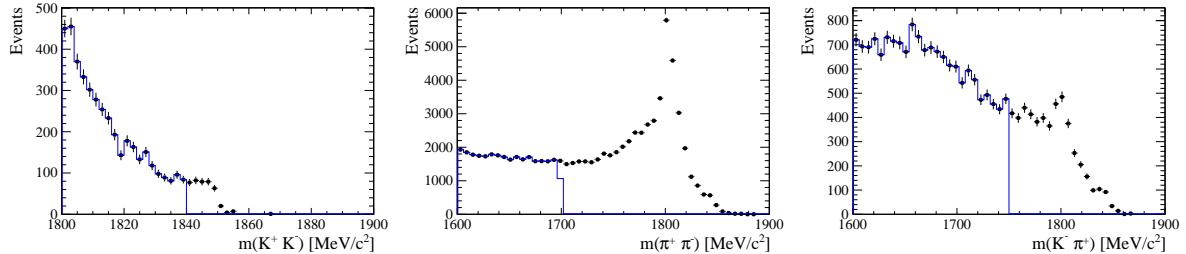


Figure 3.3: Background contributions to $D_s \rightarrow KK\pi$ (left), $D_s \rightarrow \pi\pi\pi$ (middle) and $D_s \rightarrow K\pi\pi$ (right) from $D^0 \rightarrow hh$ decays combined with a random pion.

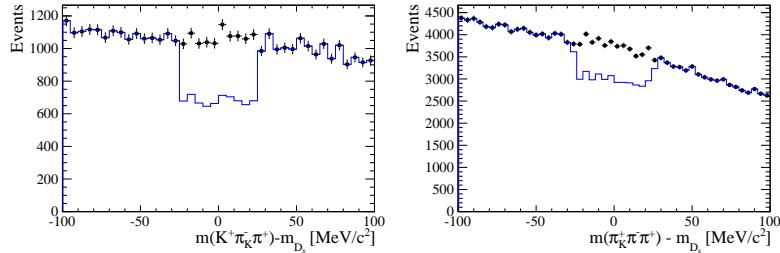


Figure 3.4: Background contributions to $B_s \rightarrow D_s K\pi\pi$ (left) and $B_s \rightarrow D_s \pi\pi\pi$ (right) from $B_s \rightarrow D_s D_s$ decays where the kaon is misidentified as pion. The $X_{s,d}$ invariant mass is recomputed applying the kaon mass hypothesis to the pion and shown without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

339 **3.2.3 Training of multivariate classifier**

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

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

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

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

370 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
371 to the training which are expected to be 100% signal efficient. Figure 3.6 shows the
372 distributions of the input variables for signal and background. As shown in Appendix B,
373 these distributions differ between data-taking period and trigger category. In particular
374 variables depending on the B_s kinematics and the event multiplicity are affected (e.g.
375 $\theta_{D_s h_i}$ or A_{pT}^{cone}). The BDTG is consequently trained separately for these categories. The
376 resulting classifier response is shown in Fig. 3.7 for each category (even and odd test
377 samples combined) and in Appendix B for each training.

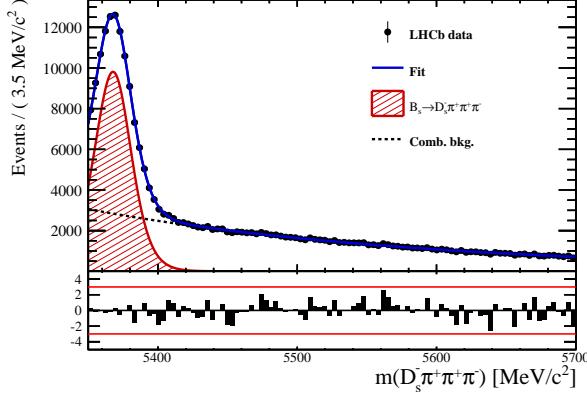


Figure 3.5: Reconstructed B_s mass for $B_s \rightarrow D_s \pi\pi\pi$ events that pass the preselection. The fitted PDF is shown in blue, the signal component in red and the background component in black.

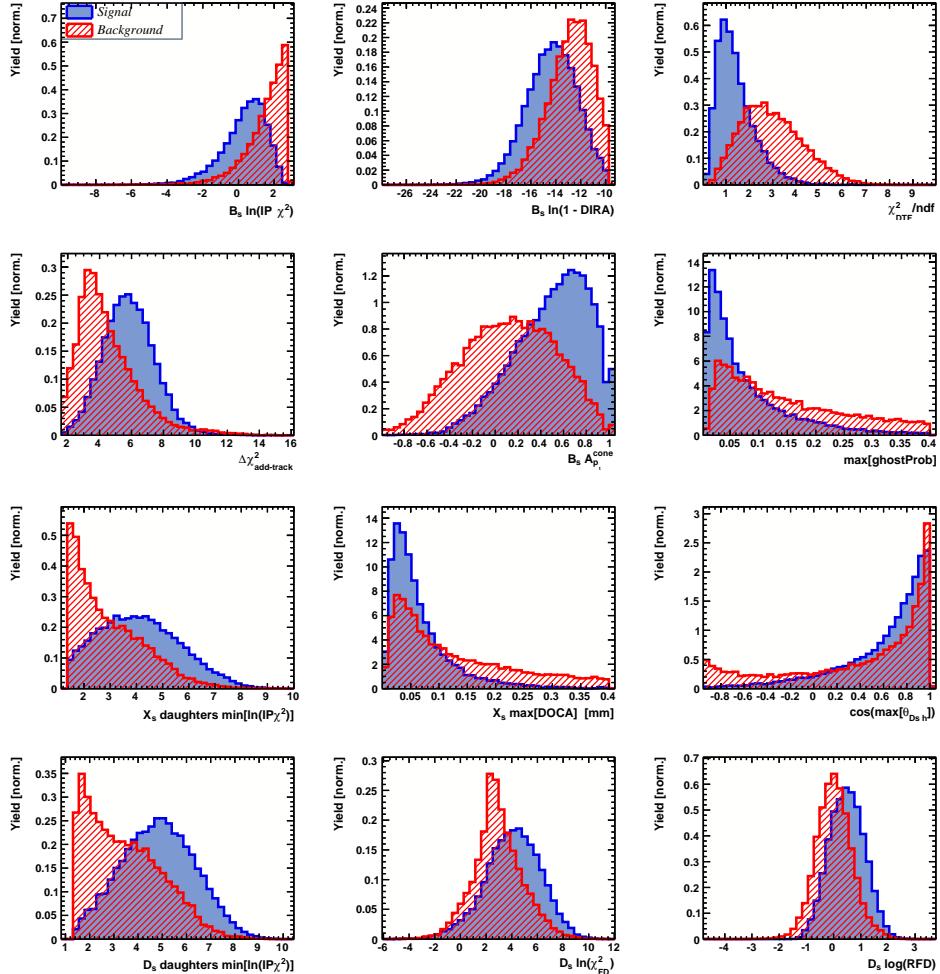


Figure 3.6: Discriminating variables used to train the BDTG for all data categories combined.

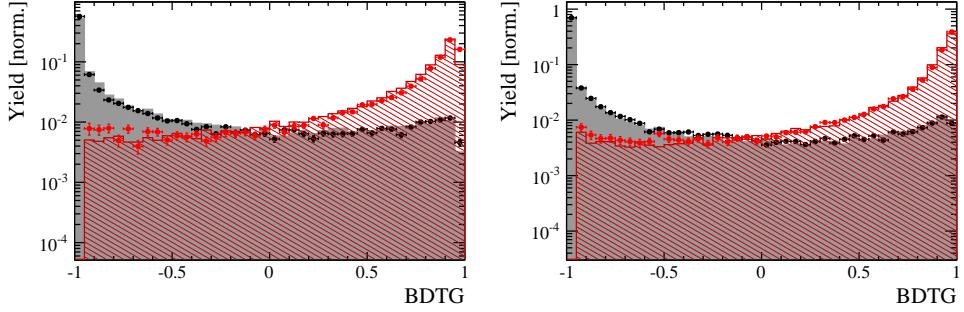


Figure 3.7: Signal (red) and background (black) distributions for the BDTG response for Run-I (left) and Run-II (right) data. Filled histograms (data points) show the BDTG response for the L0-TOS (L0-TIS) category. Even and odd test samples are combined.

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.

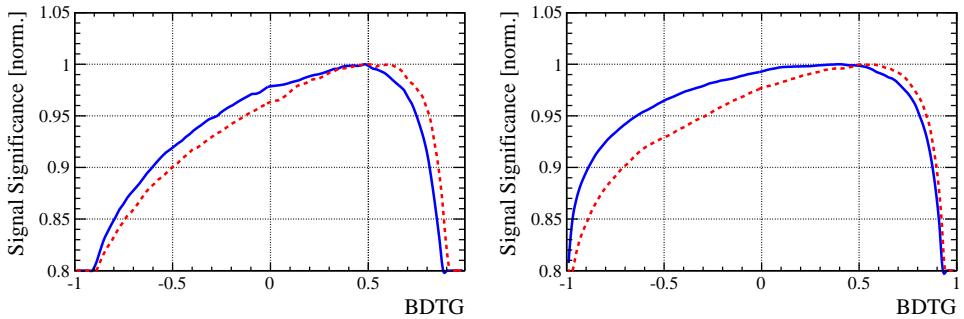


Figure 3.8: Signal significance as function of the applied cut on the BDTG response for Run-I (left) and Run-II (right) data. The scans for the L0-TOS (L0-TIS) category are shown in blue (red). The signal significance is normalized to be 1 at the optimal BDTG cut value.

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	$= 1$

389 4 Yields determination

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

395 Due to different mass resolutions, we perform the invariant mass fits simultaneously
 396 for each data-taking period and each trigger category. We further introduce four D_s final
 397 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
 398 different signal purities. The $D_s \rightarrow Kh\pi$ category combines the two D_s decay channels
 399 with the lowest statistics: $D_s \rightarrow KK\pi$ (non-resonant) and $D_s \rightarrow K\pi\pi$. This amounts to
 400 16 categories in total.

401 4.1 Signal model

402 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
 403 using a Johnson's SU function [32], which results from a variable transformation of a
 404 normal distribution to allow for asymmetric tails:

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

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

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

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

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

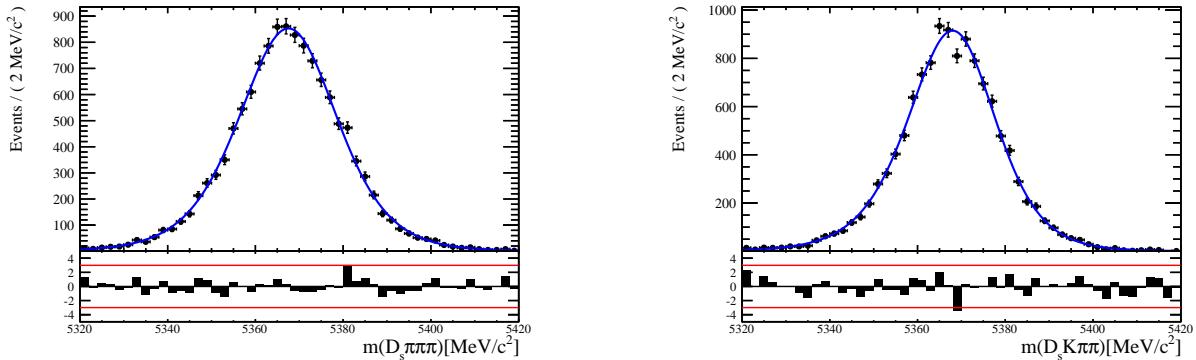


Figure 4.1: Invariant mass distributions of simulated (left) $B_s^0 \rightarrow D_s \pi\pi\pi$ and (right) $B_s^0 \rightarrow D_s K\pi\pi$ events. A fit with a Johnson's SU PDF is overlaid.

409 simulation and real data, linear scaling factors for the mean μ and width σ are determined
410 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
411 factors are determined separately for each data-taking period and each trigger category.

412 4.2 Background models

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

415

416 Combinatorial background

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

420

421 Peaking B_d background

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

424

425 Partially reconstructed background

426 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
427 to be peaking lower than signal in the $m(D_s\pi\pi\pi)$ spectrum with large tails due to the
428 momentum carried away by the not reconstructed π^0 or γ . An empirical description for
429 the shape of this contribution is derived from a $B_s^0 \rightarrow D_s^*\pi\pi\pi$ MC sample subject to
430 the nominal $B_s^0 \rightarrow D_s\pi\pi\pi$ selection. Figure 4.2 (left) shows the respective reconstructed
431 $m(D_s\pi\pi\pi)$ distribution. A sum of three bifurcated Gaussian functions is used to describe
432 it. In the fit to data, all parameters are fixed to the ones obtained from MC except for
433 the parameter which describes the width of the right tail of the distribution to account for
434 data-simulation differences in mass resolution. The equivalent $B_s^0 \rightarrow D_s^*K\pi\pi$ component
435 contributing to the $B_s^0 \rightarrow D_sK\pi\pi$ data sample is described by the same PDF with the
436 right tail fixed to the $B_s^0 \rightarrow D_s\pi\pi\pi$ result.

437 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
438 shifted by $m_{B_s^0} - m_{B^0}$.

439

440 Misidentified background

441 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
442 pions is misidentified as a kaon, contaminate the $B_s^0 \rightarrow D_sK^+\pi^+\pi^-$ sample. To determine
443 the corresponding background shapes, we use simulated events passing the nominal
444 selection except for the PID cuts on the bachelor π^+ tracks. The **PIDCalib** package
445 is used to determine the p_T, η -dependent $\pi^+ \rightarrow K^+$ misidentification probability for
446 each pion. We change the particle hypothesis from pion to kaon for the pion with the
447 higher misidentification probability and recompute the invariant B_s^0 mass, $m(D_s^-\pi_K^+\pi^+\pi^-)$.
448 Similarly, the invariant masses $m(\pi_K^+\pi^+\pi^-)$ and $m(\pi_K^+\pi^-)$ are recomputed and required
449 to be within the considered phasespace region. The background distributions are shown
450 in Fig. 4.2 (middle,right) and modeled by the sum of two Crystal Ball functions. The
451 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
452 $B_s^0 \rightarrow D_sK\pi\pi$ sample is computed by multiplying the fake rate (within the considered
453 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

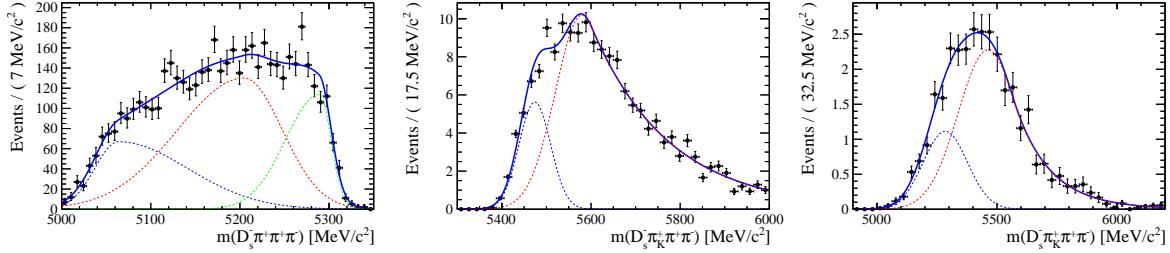


Figure 4.2: Left: Invariant mass distribution of simulated $B_s^0 \rightarrow D_s^* \pi\pi\pi$ events, where the γ/π^0 is excluded from the reconstruction. Middle: Invariant mass distribution of simulated $B_s^0 \rightarrow D_s \pi\pi\pi$ events, where one of the pions is reconstructed as a kaon taking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

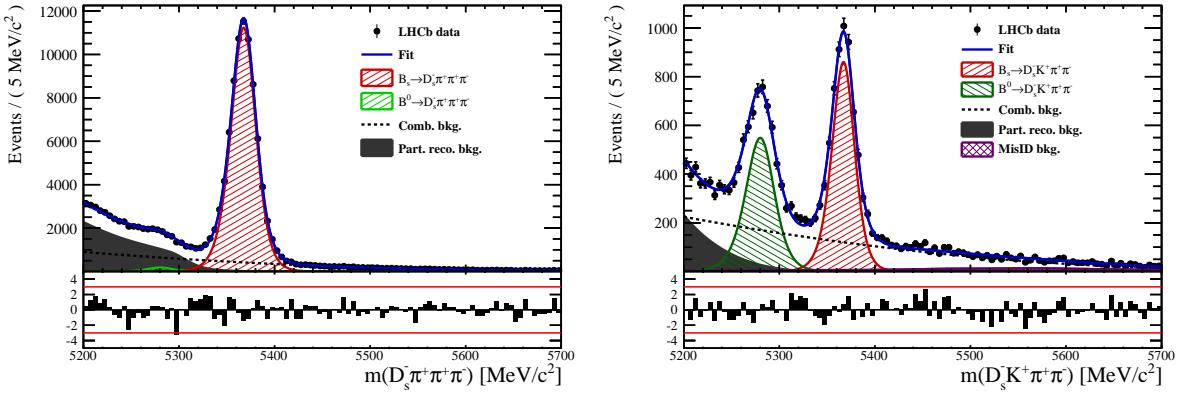


Figure 4.3: Invariant mass distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ (left) and $B_s^0 \rightarrow D_s K \pi\pi$ (right) candidates.

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	

466 5 Decay-time Resolution

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

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

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

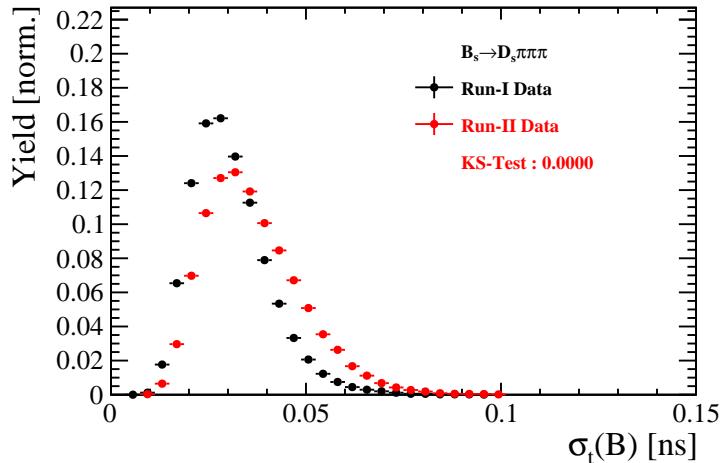


Figure 5.1: Distribution of the decay time error for $B_s^0 \rightarrow D_s\pi\pi\pi$ signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

489 5.1 Calibration for Run-I data

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

$$498 \quad \mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s^2 / 2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s^2 / 2}, \quad (5.1)$$

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

$$503 \quad \sigma_{eff} = \sqrt{(-2/\Delta m_s^2) \ln \mathcal{D}}, \quad (5.2)$$

504 which converts the resolution into a single-Gaussian function with an effective resolution
 505 that causes the same damping effect on the magnitude of the B_s oscillation. For the Run-I
 506 $B_s^0 \rightarrow D_s K \pi\pi$ MC sample the effective average resolution is found to be $\sigma_{eff} = 39.1 \pm 0.3$ fs.

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

$$515 \quad \sigma_{eff}^{MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.232 \pm 0.010) \sigma_t \quad (5.3)$$

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

$$518 \quad \sigma_{eff}^{D_s K, MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.201 \pm 0.013) \sigma_t. \quad (5.4)$$

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

$$521 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (10.26 \pm 1.52) \text{ fs} + (1.280 \pm 0.042) \sigma_t \quad (5.5)$$

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

$$524 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (-0.568 \pm 1.570) \text{ fs} + (1.243 \pm 0.044) \sigma_t \quad (5.6)$$

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

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

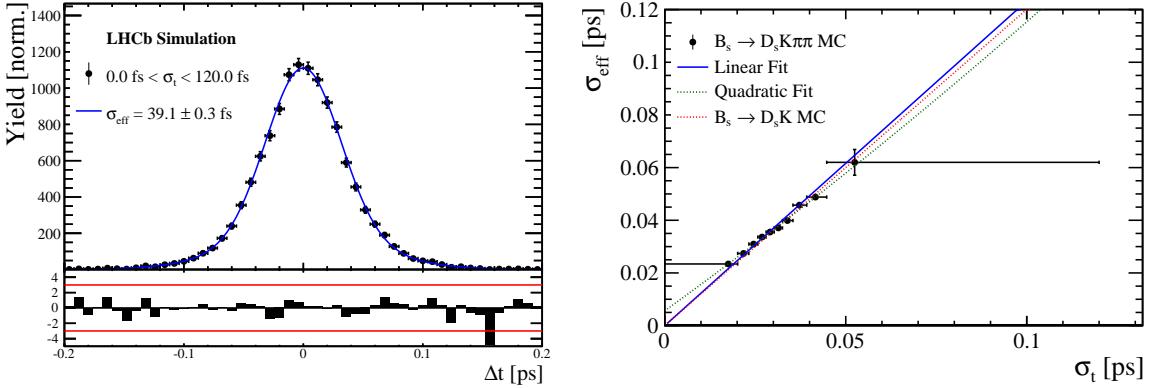


Figure 5.2: (Left) Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K\pi\pi$ MC candidates. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for $B_s \rightarrow D_s K\pi\pi$ MC (Run-I). The fitted calibration curve is shown in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

5.2 Calibration for Run-II data

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

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

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

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

⁵⁴⁸ is in good agreement with the one obtained for the $B_s \rightarrow J/\psi\phi$ (Run-II) analysis [36].

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

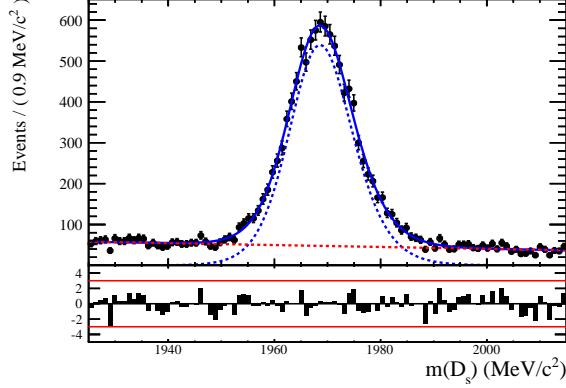


Figure 5.3: The invariant mass distribution for prompt D_s candidates.

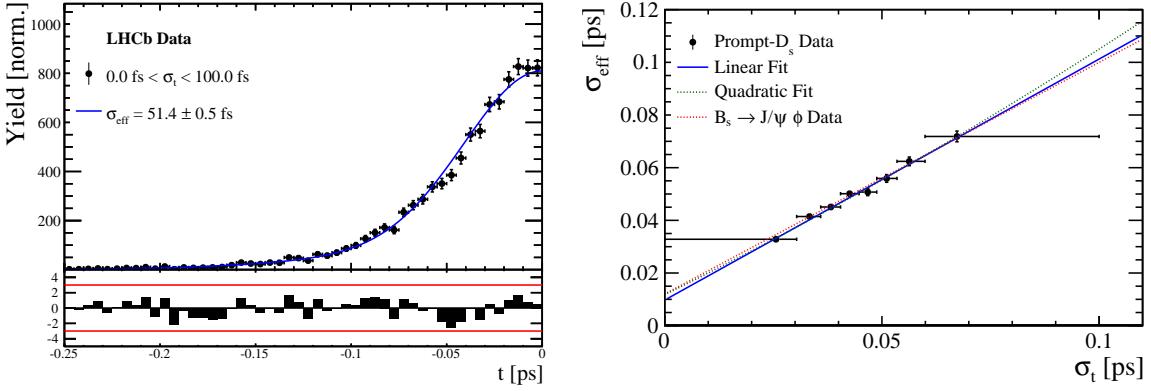


Figure 5.4: (Left) Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data). The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

549 6 Acceptance

550 6.1 MC corrections

551 6.1.1 Truth matching of simulated candidates

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

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

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

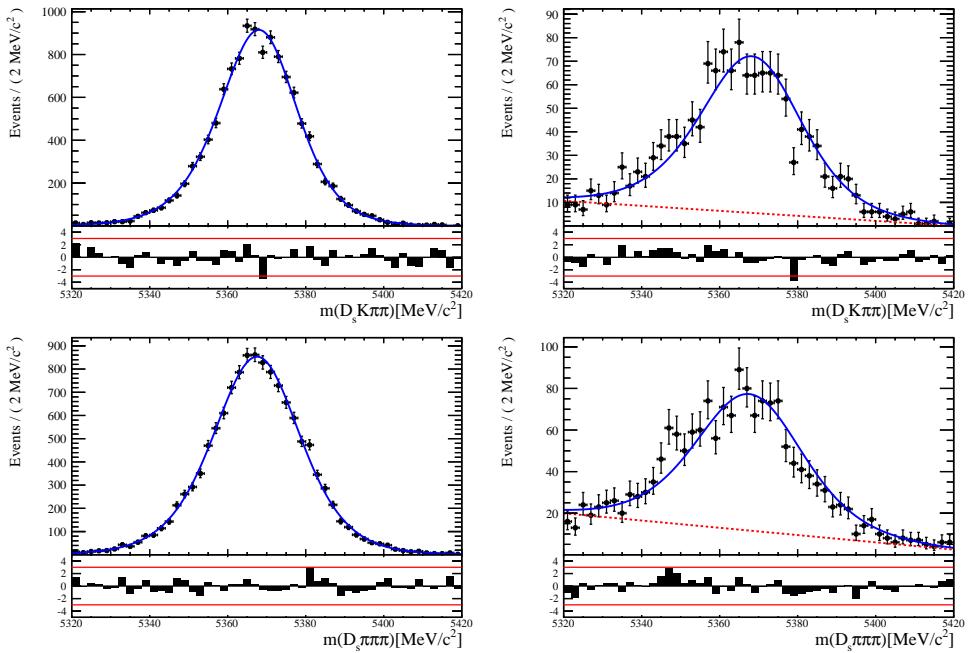


Figure 6.1: The reconstructed B_s mass distribution for simulated $B_s \rightarrow D_s K\pi\pi$ (top) and $B_s \rightarrow D_s \pi\pi\pi$ (bottom) decays classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

571 **6.2 Decay-time acceptance**

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

577 We use our control channel for this measurement, because for $B_s^0 \rightarrow D_s K\pi\pi$ decays
 578 the decay-time acceptance is correlated with the CP -observables which we aim to measure.
 579 Therefore, floating the CP -observables and the acceptance shape at the same time is
 580 not possible. Hence, a fit to the decay-time distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates is
 581 performed and the obtained acceptance shape is corrected by the difference in shape found
 582 for the $B_s^0 \rightarrow D_s K\pi\pi$ and $B_s^0 \rightarrow D_s \pi\pi\pi$ MC.

583 A PDF of the form

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

584 is fit to the decay time distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates in data. Since the
 585 fit is performed untagged, the PDF shown in Eq. 6.1 contains no terms proportional
 586 to Δm_s . The values for Γ_s and $\Delta\Gamma_s$ are fixed to the latest HFAG results [37]. The
 587 decay-time acceptance $\epsilon(t', \vec{\lambda})$ is modelled using the sum of cubic polynomials $v_i(t)$, so
 588 called Splines [38]. The polynomials are parametrised by so-called knots which determine
 589 their boundaries. Knots can be set across the fitted distribution to account for local
 590 changes in the acceptance shape. Using more knots is equivalent to using more base
 591 splines which are defined on a smaller sub-range. In total, $n + 2$ base splines $v_i(t)$ are
 592 needed to describe an acceptance shape which is parametrised using n knots.

593 For fits shown in the following, the knots have been placed at $t =$
 594 $[0.5, 1.0, 1.5, 2.0, 3.0, 9.5] ps$. To accommodate these 6 knot positions, 8 basic splines
 595 v_i , $i = [1, \dots, 8]$ are used. Since a rapid change of the decay time acceptance at low
 596 decay times due to the turn-on effect generated by the lifetime and other selection cuts is
 597 expected, more knots are placed in that regime. At higher decay times we expect linear
 598 behavior, with a possible small effect due to the VELO reconstruction. Therefore fewer
 599 knots are used. Furthermore, v_7 is fixed to 1 in order to normalize the overall acceptance
 600 function. To stabilise the last spline, v_8 is fixed by a linear extrapolation from the two
 601 previous splines:

$$v_N = v_{N-1} + \frac{v_{N-2} - v_{N-1}}{t_{N-2} - t_{N-1}} \cdot (t_N - t_{N-1}). \quad (6.2)$$

602 Here, $N = 8$ and t_{N-1} corresponds to the knot position associated with v_{N-1} .

6.2.1 Comparison of acceptance in subsamples

It is possible that the decay-time dependent efficiency deviates in different subsamples of our data. In particular, the acceptance could differentiate in subsamples with different final state kinematics, such as the run I & run II sample, the various D_s final states and the ways an event is triggered at the L0 stage. To investigate possible deviations, the full selected $B_s^0 \rightarrow D_s\pi\pi\pi$ sample is split into subsamples according to the categories mentioned above (run, D_s state, L0 trigger). For each subsample, the fit procedure described at the beginning of this chapter, using the pdf given by Eq. 6.1, is repeated and the obtained values for the spline coefficients v_i are compared. Figure ?? shows the comparison of the obtained spline coefficients for the different D_s final states.

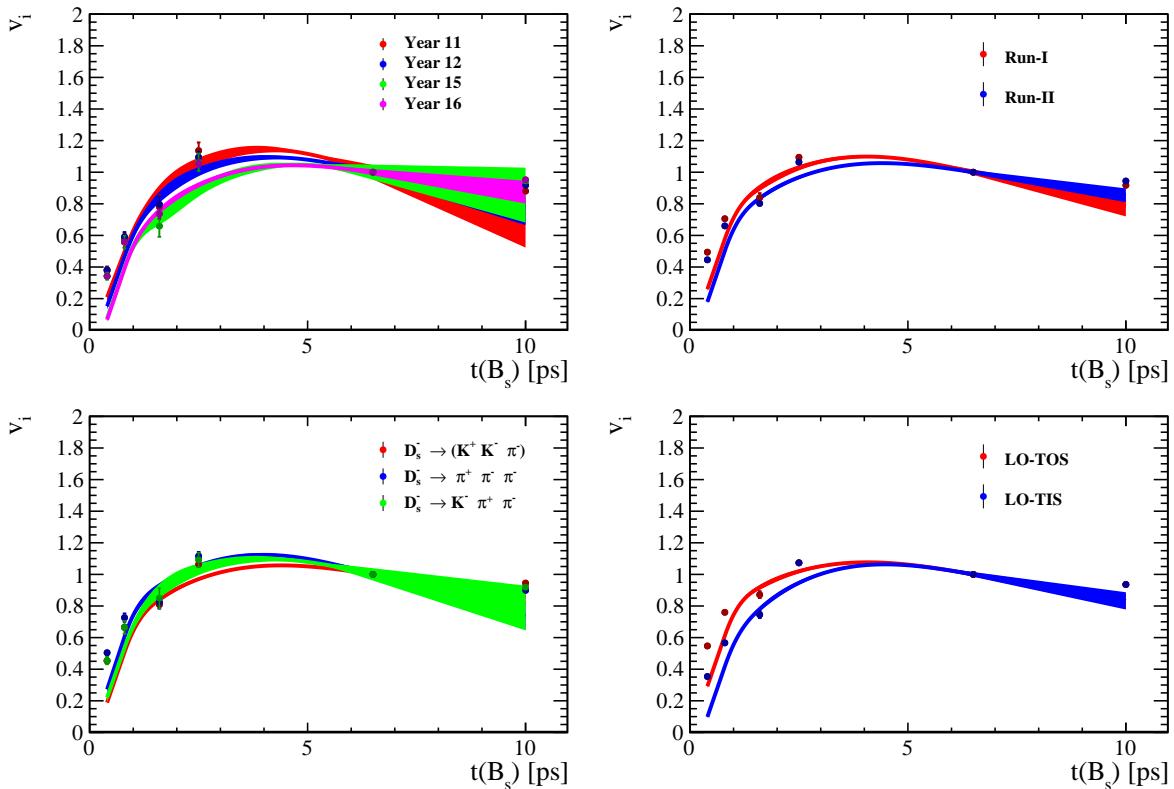


Figure 6.2: Comparison of the spline coefficients obtained from time-dependent fits to the $B_s^0 \rightarrow D_s\pi\pi\pi$ subsamples of (left) the different runs and (right) L0 trigger categories.

Investigating the obtained spline coefficients from different D_s final states, good agreement is observed between all four channels and no need to distinguish between different final states in the time-dependent amplitude fit is found. The comparison between spline coefficients for the different runs and L0 trigger categories is shown in Figure 6.2.

Significant deviations between spline coefficients obtained from the two different runs and L0 trigger categories can be observed. The deviations are most pronounced in the $(0 - 5)$ ps region, where the majority of statistics is found. Therefore, the time-dependent efficiency has to be treated separately for the runs and L0 categories. This is achieved by implementing a simultaneous fit, where the acceptance description is allowed to vary in the subsamples.

624 **6.2.2 Results**

625 The nominal fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data using this configuration is shown in Figure ??.
626 Note that the normalization of the splines in the following figures is not in scale. The fit
627 parameters obtained from the described fits to data and simulation are summarised in
628 Table 6.4.

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	0.576 ± 0.021	0.537 ± 0.017	1.005 ± 0.051
0.8	v_1	0.841 ± 0.024	0.788 ± 0.026	0.894 ± 0.040
1.6	v_2	0.845 ± 0.068	0.914 ± 0.047	1.044 ± 0.074
2.5	v_3	1.113 ± 0.040	1.107 ± 0.028	0.956 ± 0.043
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.901 (interpolated)	0.907 (interpolated)	1.038 (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	0.372 ± 0.036	0.402 ± 0.021	1.046 ± 0.099
0.8	v_1	0.598 ± 0.057	0.640 ± 0.034	0.898 ± 0.075
1.6	v_2	0.917 ± 0.089	0.982 ± 0.057	0.905 ± 0.080
2.5	v_3	1.091 ± 0.053	1.077 ± 0.035	1.007 ± 0.051
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.921 (interpolated)	0.932 (interpolated)	0.994 (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	0.533 ± 0.028	0.486 ± 0.015	1.021 ± 0.054
0.8	v_1	0.789 ± 0.041	0.733 ± 0.024	0.898 ± 0.046
1.6	v_2	0.904 ± 0.059	0.942 ± 0.038	0.990 ± 0.058
2.5	v_3	1.083 ± 0.030	1.096 ± 0.021	0.975 ± 0.031
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.927 (interpolated)	0.916 (interpolated)	1.022 (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	0.345 ± 0.039	0.489 ± 0.028	0.985 ± 0.088
0.8	v_1	0.572 ± 0.065	0.726 ± 0.043	0.940 ± 0.087
1.6	v_2	0.807 ± 0.065	0.966 ± 0.046	0.921 ± 0.063
2.5	v_3	1.111 ± 0.038	1.112 ± 0.024	0.943 ± 0.041
6.5	v_4	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	v_5	0.903 (interpolated)	0.902 (interpolated)	1.050 (interpolated)

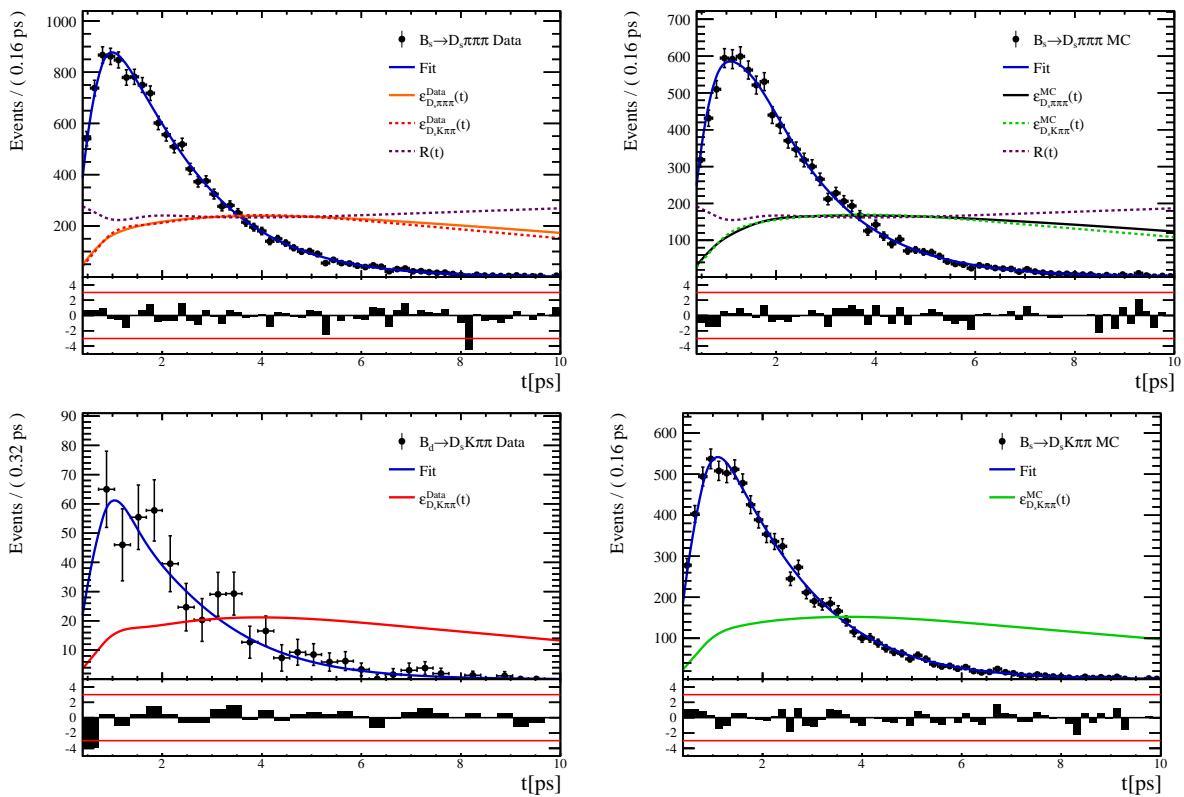


Figure 6.3

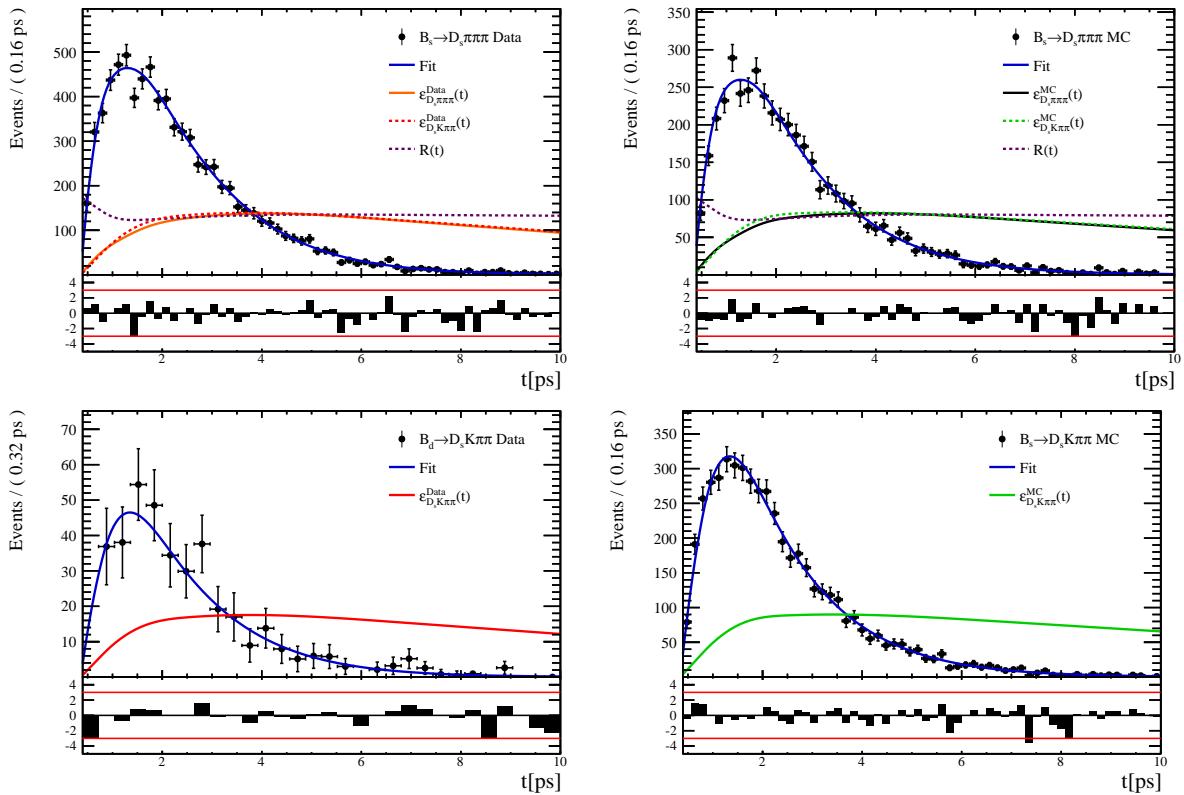


Figure 6.4

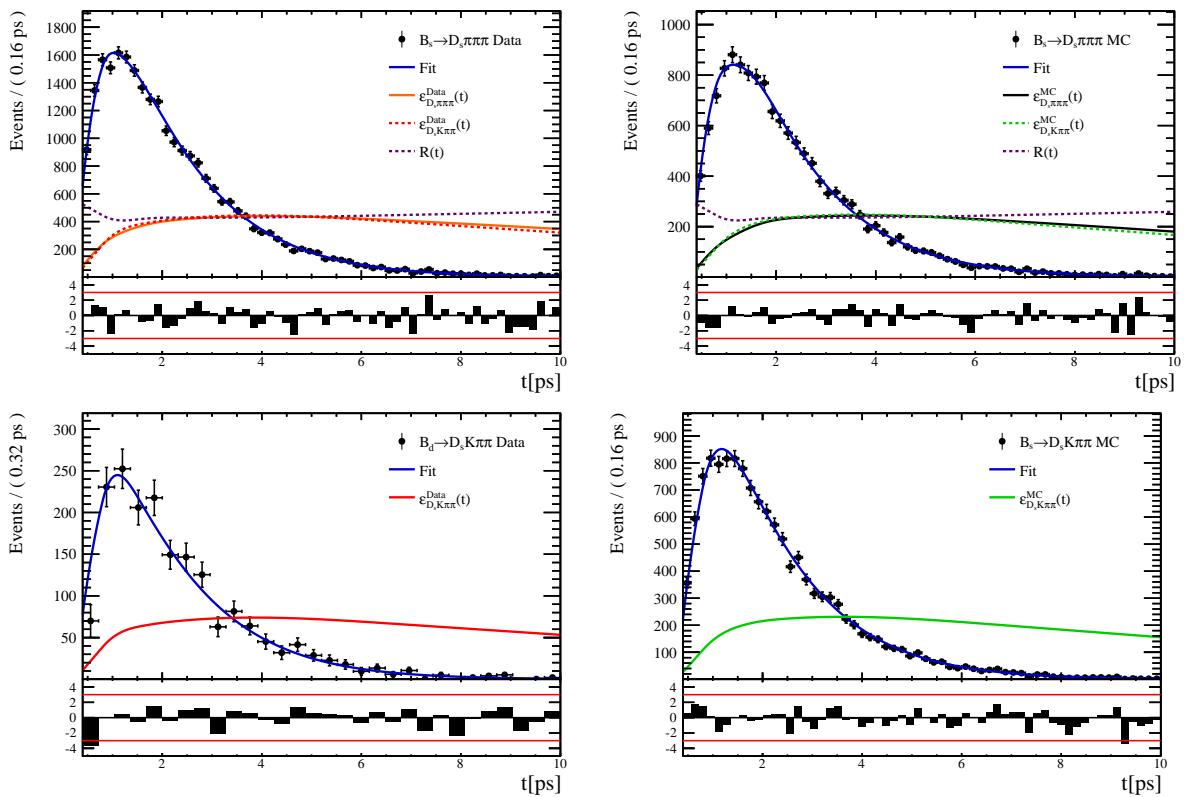


Figure 6.5

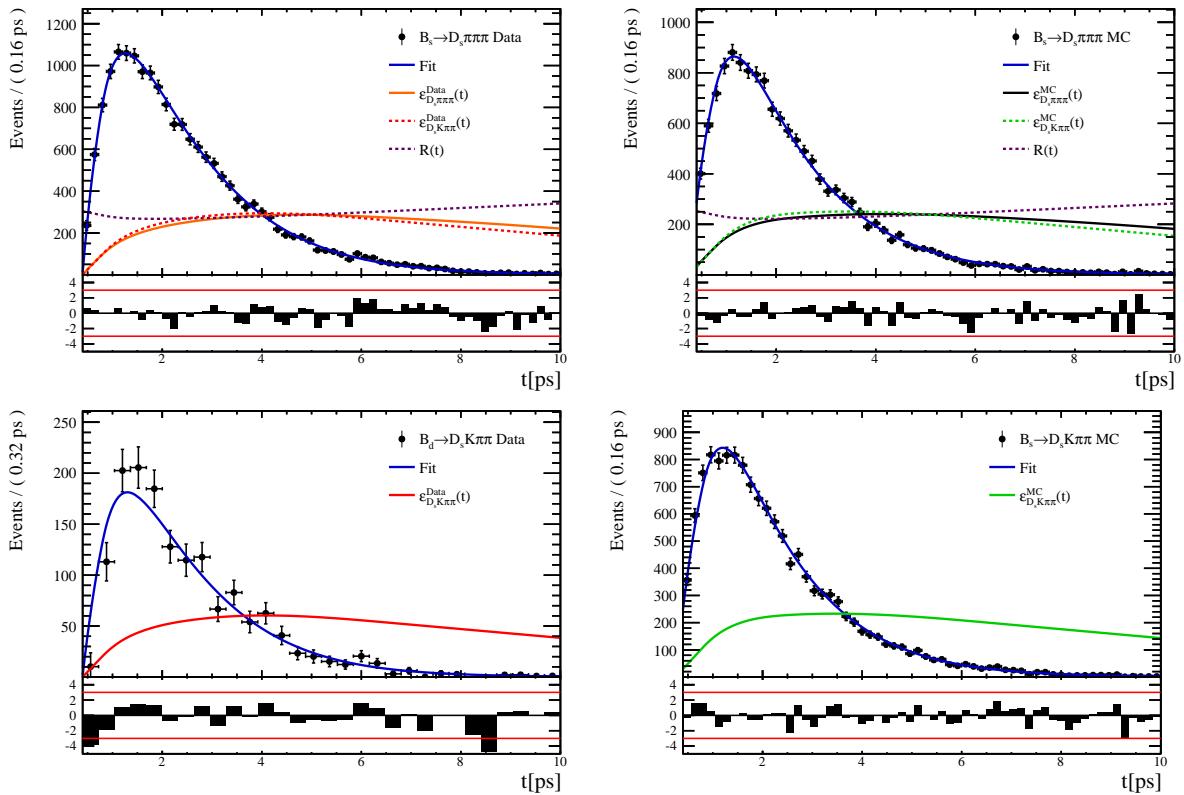


Figure 6.6

6.3 Phasespace acceptance

Disclaimer: ...

In general, the signal PDF for events tagged as $D^0 \rightarrow h^+h^-\pi^+\pi^-$ is given by

$$\mathcal{P}_{\text{Sig}}(\mathbf{x}) = \frac{[(1-w)|A_{D^0}(\mathbf{x})|^2 + w|A_{\bar{D}^0}(\mathbf{x})|^2]\epsilon_{\text{Sig}}(\mathbf{x})\phi_4(\mathbf{x})}{\int [|A_{D^0}(\mathbf{x})|^2 + |A_{\bar{D}^0}(\mathbf{x})|^2]\epsilon_{\text{Sig}}(\mathbf{x})d\Phi_4}, \quad (6.3)$$

where $\epsilon_{\text{Sig}}(\mathbf{x})$ is the phase-space efficiency.

Note that the efficiency in the numerator appears as an additive constant in the log \mathcal{L} that does not depend on any fit parameters such that it can be ignored. However, the efficiency function still enters via the normalization integrals. These normalization terms are determined numerically by a MC integration technique. For this purpose, we use simulated events generated according to a preliminary model, pass them through the full detector simulation and apply the same selection criteria as for data in order to perform the MC integrals. For example, the first integral in Eq. (6.3) can be approximated as

$$\int |A_{D^0}(\mathbf{x})|^2 \epsilon_{\text{Sig}}(\mathbf{x}) d\Phi_4 \approx \frac{1}{N_{\text{MC}}} \sum_k^{N_{\text{MC}}} \frac{|A_{D^0}(\mathbf{x}_k)|^2}{|A'_{D^0}(\mathbf{x}_k)|^2} \quad (6.4)$$

where A'_{D^0} labels the preliminary amplitude model and x_k is the k -th MC event. As a result, the efficiency can be included in the amplitude fit without explicitly modeling it. For $D^0 \rightarrow \pi^+\pi^-\pi^+\pi^-$, we use a sample of $N_{\text{MC}} = 600000$ MC events to ensure that the uncertainty on the integral is less than 0.5%.

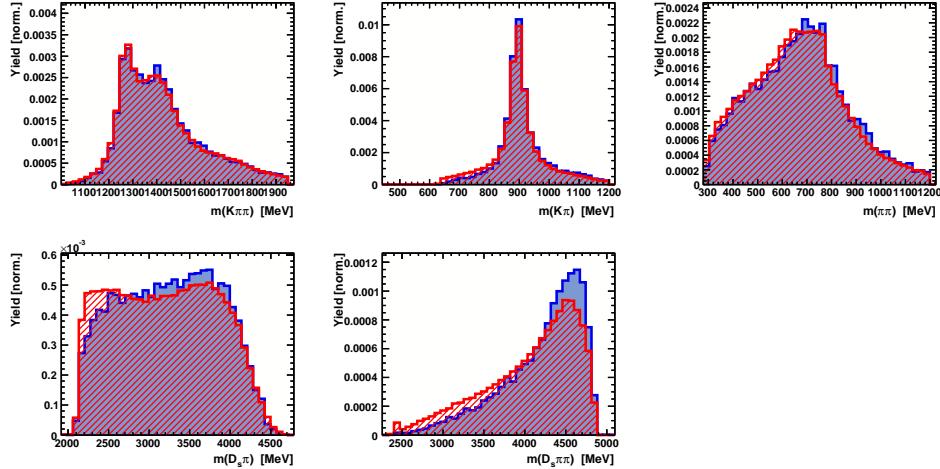


Figure 6.7

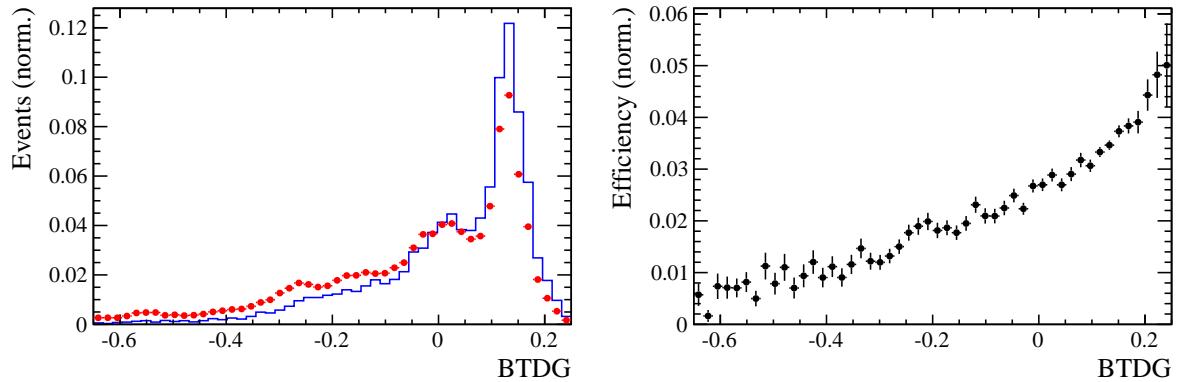


Figure 6.8

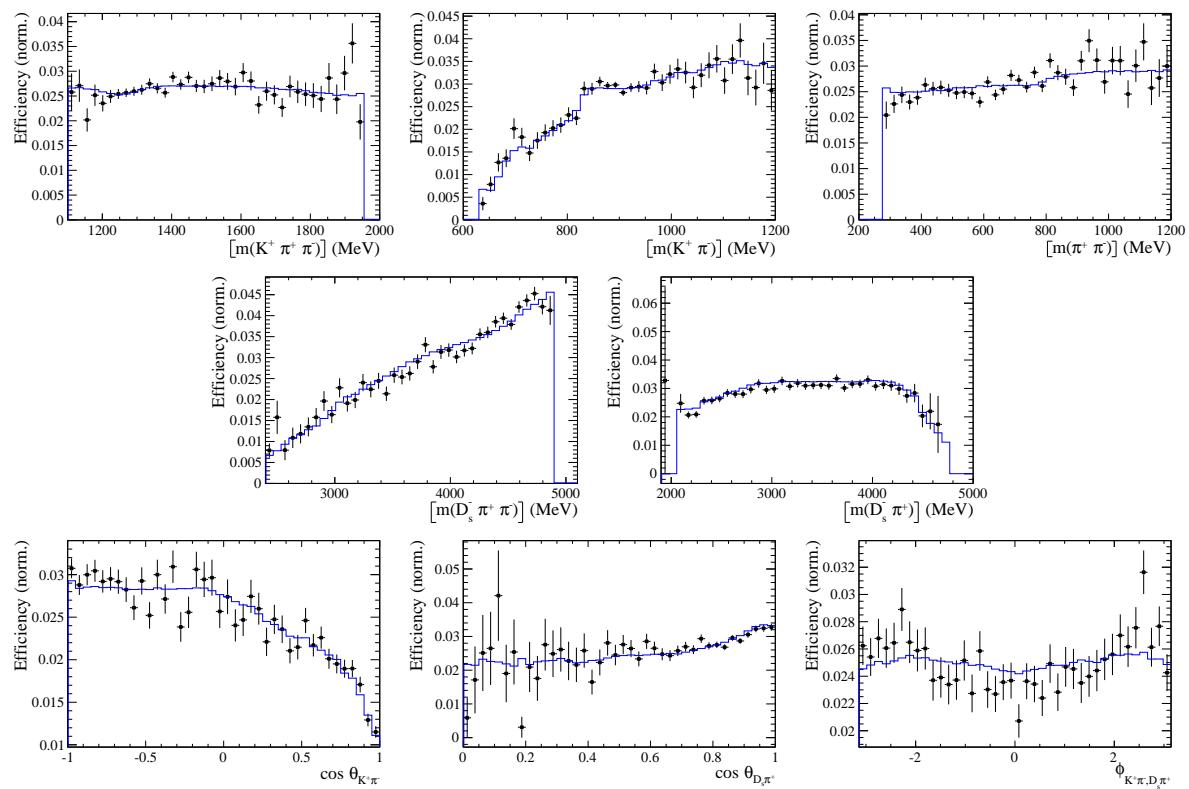


Figure 6.9

644 7 Flavour Tagging

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

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

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

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

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

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

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

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

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

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

684 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 $\epsilon_{eff,OS} = 4.81\%$.

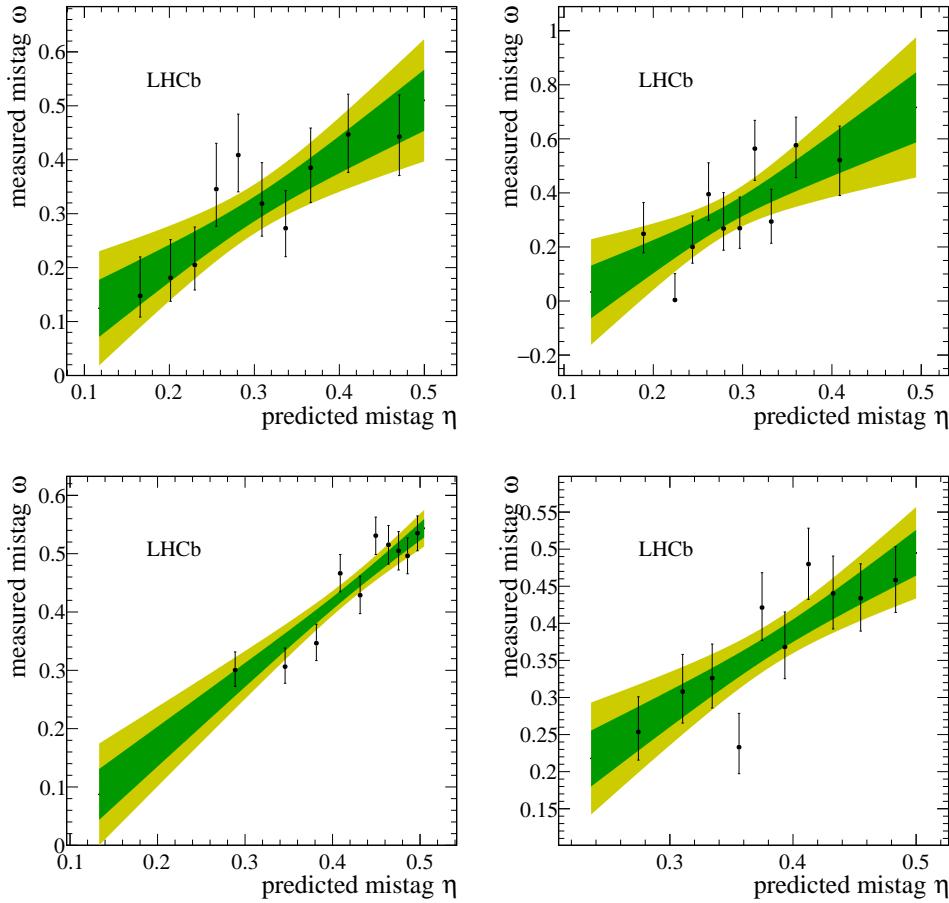


Figure 7.1: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run 1. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

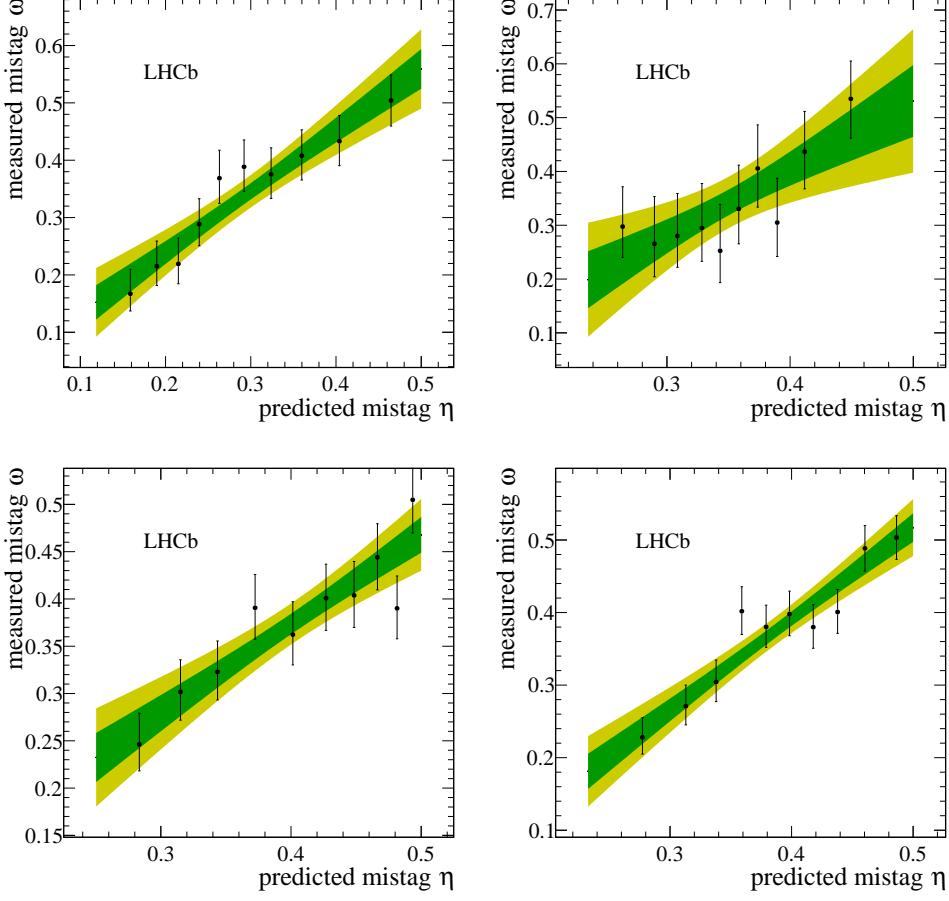


Figure 7.2: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run 2. A linear fit, including the 1σ and 2σ error bands is overlaid for each tagger.

698 7.2 SS tagging calibration

699 The SS neural net kaon tagger can be calibrated using the flavour-specific $B_s^0 \rightarrow D_s\pi\pi\pi$
700 decay. Its development, performance and calibration is described in detail in [39]. Figure
701 7.3 shows the distribution of the predicted mistag of the neural net kaon tagger. The
702 extracted calibration parameters and tagging asymmetries are summarized in Table 7.1
703 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.

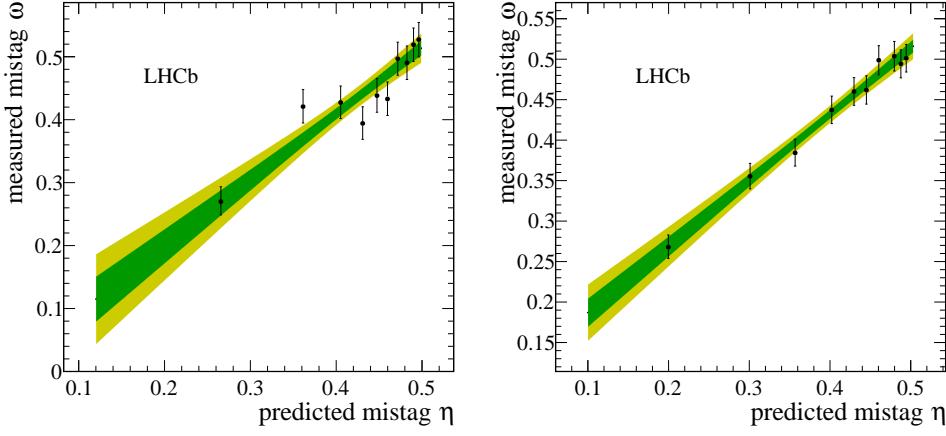


Figure 7.3: Predicted versus measured mistag probability for the SS neural net kaon tagger for (left) Run 1 and (right) Run 2. A linear fit, including the 1σ and 2σ error bands is overlaid for both distributions.

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.

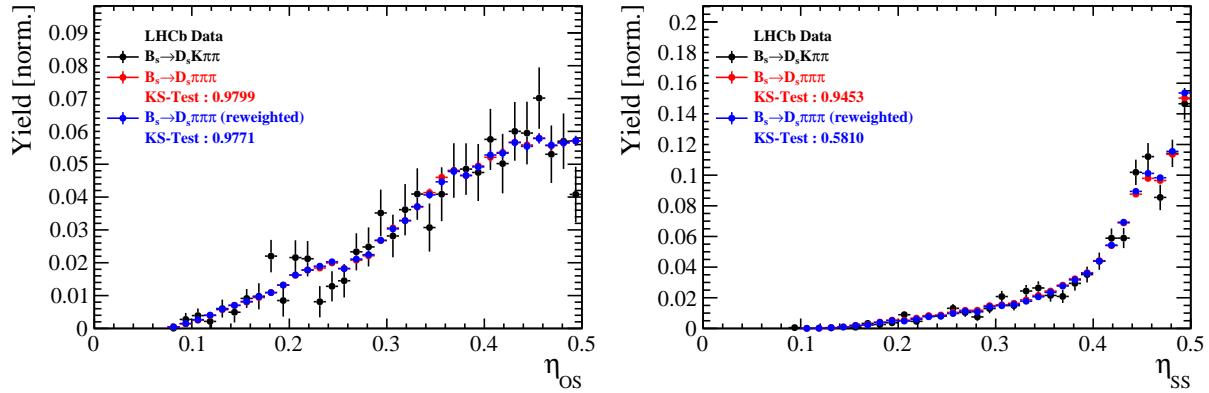


Figure 7.4: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \rightarrow D_s K\pi\pi$ (black) and $B_s^0 \rightarrow D_s \pi\pi\pi$ (red) data samples.

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.

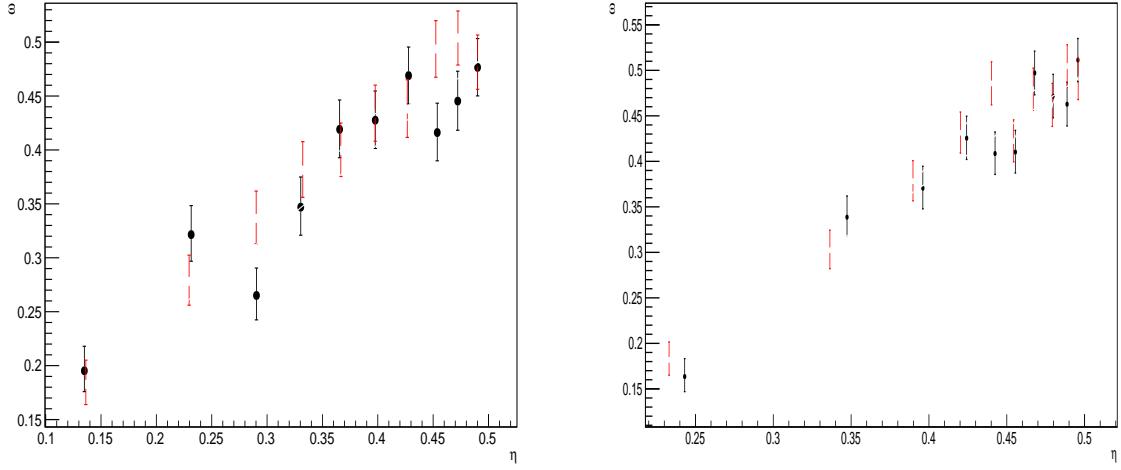


Figure 7.5: Dependence of the observed mistag ω on the predicted mistag η for the (left) OS combination and the (right) SS kaon tagger, found in the simulated $B_s^0 \rightarrow D_sK\pi\pi$ (black) and $B_s^0 \rightarrow D_s\pi\pi\pi$ (red) signal samples.

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 [40] 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 [41]. 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. [41]

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

748 8.2 $K^-\pi^+$ Detection Asymmetry

749 The presented measurement of the CKM-angle γ using $B_s^0 \rightarrow D_s K\pi\pi$ decays is sensitive to
 750 a possible charge asymmetry of the kaon. Kaons are known to have a nuclear cross-section
 751 which is asymmetrically dependent on the sign of their charge. It is indispensable to
 752 determine the charge asymmetry of the kaon, as fitting without taking this effect into
 753 account would introduce a 'fake' CP violation. Instead of determining the single track
 754 detection asymmetry of a kaon, it is found that the combined two track asymmetry of a
 755 kaon-pion pair is much easier to access [42]. Therefore, the two track asymmetry defined
 756 as

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

757 is used.

758 This asymmetry can be measured from the difference in asymmetries in the $D^+ \rightarrow$
 759 $K^-\pi^+\pi^+$ and $D^+ \rightarrow K_s^0\pi^+$ modes [43]:

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

760 where possible CP violation in the $D^+ \rightarrow K_s^0\pi^+$ mode is predicted to be smaller than
 761 10^{-4} in the Standard Model [44]. The asymmetry in the neutral kaon system, $A(K^0)$, has
 762 to be taken into account as a correction.

763 We use a dedicated LHCb tool to determine $A^{det}(K^-\pi^+)$ for all data taking periods
 764 used in this analysis. A detailed description can be found in [43]. The tool provides
 765 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
 766 determine the asymmetry following Eq. 8.6. Several weighting steps are performed to
 767 match the kinematics of the calibration samples to our signal decay sample:

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

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

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

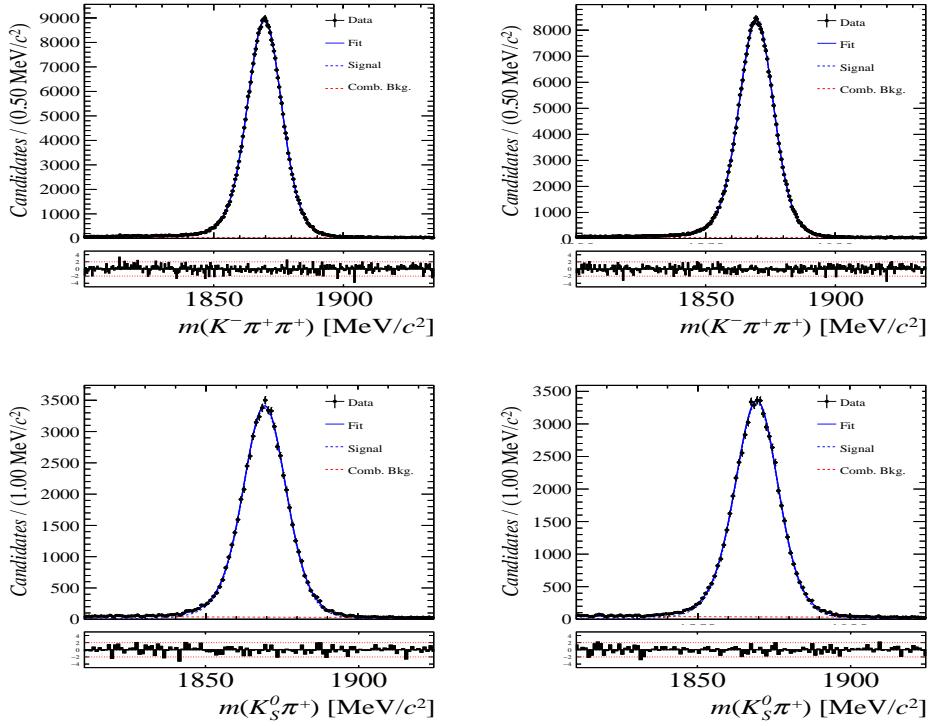


Figure 8.1: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_S^0 \pi^\pm$ candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

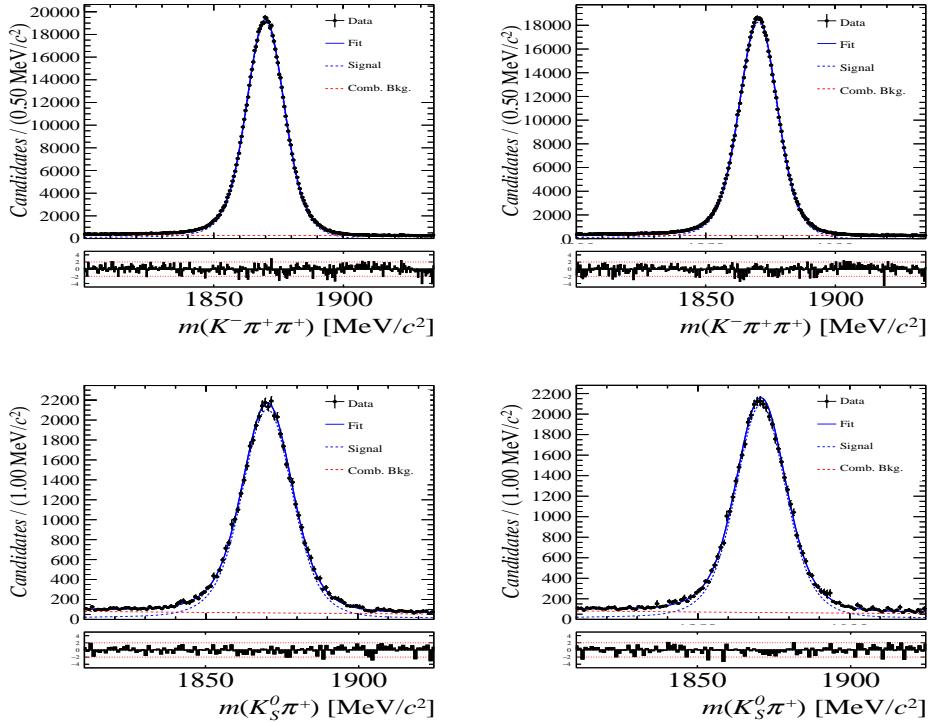


Figure 8.2: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_S^0 \pi^\pm$ candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

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	
2015, mag. up	-1.36 \pm 0.36
2015, mag. down	-0.96 \pm 0.24
2015, average	-1.16 \pm 0.22
2016, mag. up	0.50 \pm 0.88
2016, mag. down	1.23 \pm 0.72
2016, average	0.87 \pm 0.57

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

785 9 Time dependent fit

786 This section covers the phasespace integrated, time-dependent fit to $B_s^0 \rightarrow D_s h\pi\pi$ data.
 787 We use the **sFit** technique [45] to statistically remove background from the decay time fit,
 788 leaving only the signal PDF to describe the decay time. The **sWeights** are calculated based
 789 on the fit to the reconstructed B_s mass distribution described in Sec. 4. As additional
 790 input to the fit, the tagging information (Sec. 7), as well as the decay time acceptance
 791 (Sec. 6) and resolution (Sec. 5) is used and fixed to the values obtained by the dedicated
 792 studies. Taking all inputs into account, the final time dependent fit PDF is given by

$$793 \quad \mathcal{P}\mathcal{D}\mathcal{F}(t, \vec{\lambda}) = \left(\epsilon(t) \cdot \int P(x, t, q_t, q_f) dx \right) \times \mathcal{R}(t - t'), \quad (9.1)$$

793 where $\int P(x, t, q_t, q_f) dx$ is the PDF given by Eq. 2.8, $\epsilon(t)$ is the efficiency due to the time
 794 acceptance effects and $\mathcal{R}(t - t')$ is the Gaussian time resolution function.

795 9.1 sFit to $B_s^0 \rightarrow D_s \pi\pi\pi$ data

796 The phasespace-integrated, time-dependent fit is performed to the full sWeighted sample
 797 of selected candidates from Run I and 2015+2016 Run II data, containing both possible
 798 magnet polarities and all D_s final states. In the fit, the values of Γ_s and $\Delta\Gamma_s$ are fixed to
 799 the latest PDG report [12]. All tagging parameters are fixed to the central values found in
 800 the tagging calibration, described in Sec. 7. Due to the fact that the $B_s^0 \rightarrow D_s \pi\pi\pi$ decay
 801 is flavour specific, the CP-coefficients can be fixed to $C = 1$ and $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$,
 802 reducing Eq. 2.8 to

$$793 \quad \int P(x, t, q_t, q_f) dx = [\cosh\left(\frac{\Delta\Gamma}{2}t\right) + q_t q_f \cos(m_s t)] e^{-\Gamma t}. \quad (9.2)$$

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

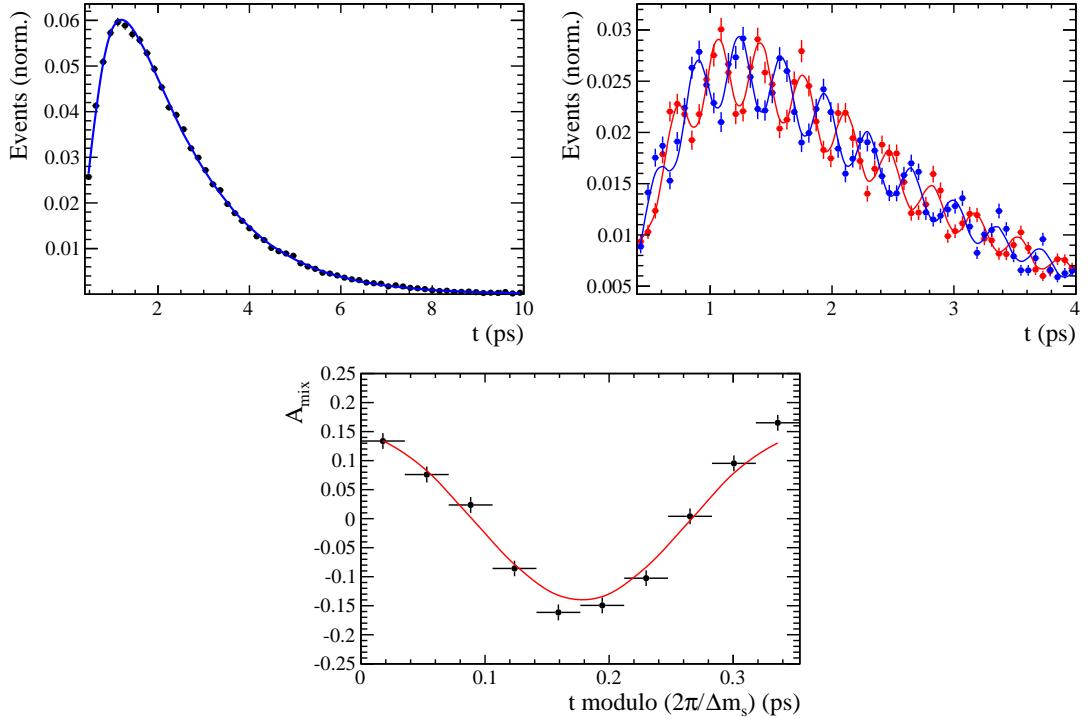


Figure 9.1: Top left: Flavour averaged decay time distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates with the averaged fit overlaid. Top right: Tagged decay time distribution of mixed (red) and unmixed (blue) signal candidates with the fit described in the text overlaid. Bottom: Time-dependent 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. 6.2.1, 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.

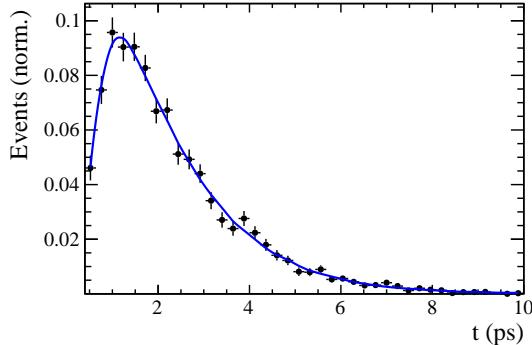


Figure 9.2: Tagged and sWeighted decay-time distribution of $B_s^0 \rightarrow D_s K\pi\pi$ signal candidates. The fit described in the text is overlaid.

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

819 10 Time dependent amplitude fit

820 10.1 Signal Model Construction

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

827 To build the amplitude model, one could successively add amplitudes on top of one
828 another until a reasonable agreement between data and fit was achieved. However, this
829 step-wise approach is not particularly suitable for amplitude analyses as discussed in
830 Ref. [46]. Instead, we include the whole pool of amplitudes in the first instance and use
831 the Least Absolute Shrinkage and Selection Operator [46, 47] (LASSO) approach to limit
832 the model complexity. In this method, the event likelihood is extended by a penalty term

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

833 which shrinks the amplitude coefficients towards zero. The amount of shrinkage is
834 controlled by the parameter λ , to be tuned on data. Higher values for λ encourage sparse
835 models, *i.e.* models with only a few non-zero amplitude coefficients. The optimal value
836 for λ is found by minimizing the Bayesian information criteria [48] (BIC),

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

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

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

- 846 1. The fit fraction threshold for inclusion in the final model was varied within the
847 interval [0.05, 5]%. The set of selected amplitudes is stable for thresholds between
848 0.1% and 1%. Other choices result in marginally different models containing one
849 component more or less.
 - 850 2. Instead of BIC, the Akaike information criteria ($\text{AIC}(\lambda) = -2 \log \mathcal{L} + 2r$ [49]) was
851 used to optimize λ . For a given threshold, the AIC method tends to prefer lower
852 λ values. However, the set of models obtained varying the threshold within the
853 interval [0.05, 5]% is identical to the BIC method.
 - 854 3. The amplitudes selected under nominal conditions were excluded one-by-one from
855 the set of all amplitudes considered.
- 856 From that we obtained a set of alternative models shown in Appendix ??.

Table 10.1: Fit fractions for $B_s \rightarrow D_s K\pi\pi$ data.

Decay channel	Fraction [%]
$B_s \rightarrow K(1)(1400)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	34.70 ± 2.24
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K(0)^*(1430)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	6.85 ± 0.94
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	13.08 ± 1.70
$B_s \rightarrow K(1)(1270)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+\pi^-) K^+) D_s^-$	9.25 ± 0.60
$B_s \rightarrow K(1460)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	0.99 ± 0.06
$BuggB_s \rightarrow K(1460)^+ (\rightarrow \sigma(\rightarrow \pi^+\pi^-) K^+) D_s^-$	3.42 ± 1.49
$B_s \rightarrow K^*(1410)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	16.40 ± 1.06
$B_s \rightarrow K^*(1410)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+\pi^-) K^+) D_s^-$	4.88 ± 0.68
$B_s \rightarrow NonResS0 (\rightarrow D_s^- \pi^+) K^*(892)^0 (\rightarrow K^+\pi^-)$	4.60 ± 1.44
$BuggB_s \rightarrow NonResS0 (\rightarrow D_s^- K^+) \sigma(\rightarrow \pi^+\pi^-)$	4.96 ± 0.68
Sum	99.13 ± 5.87

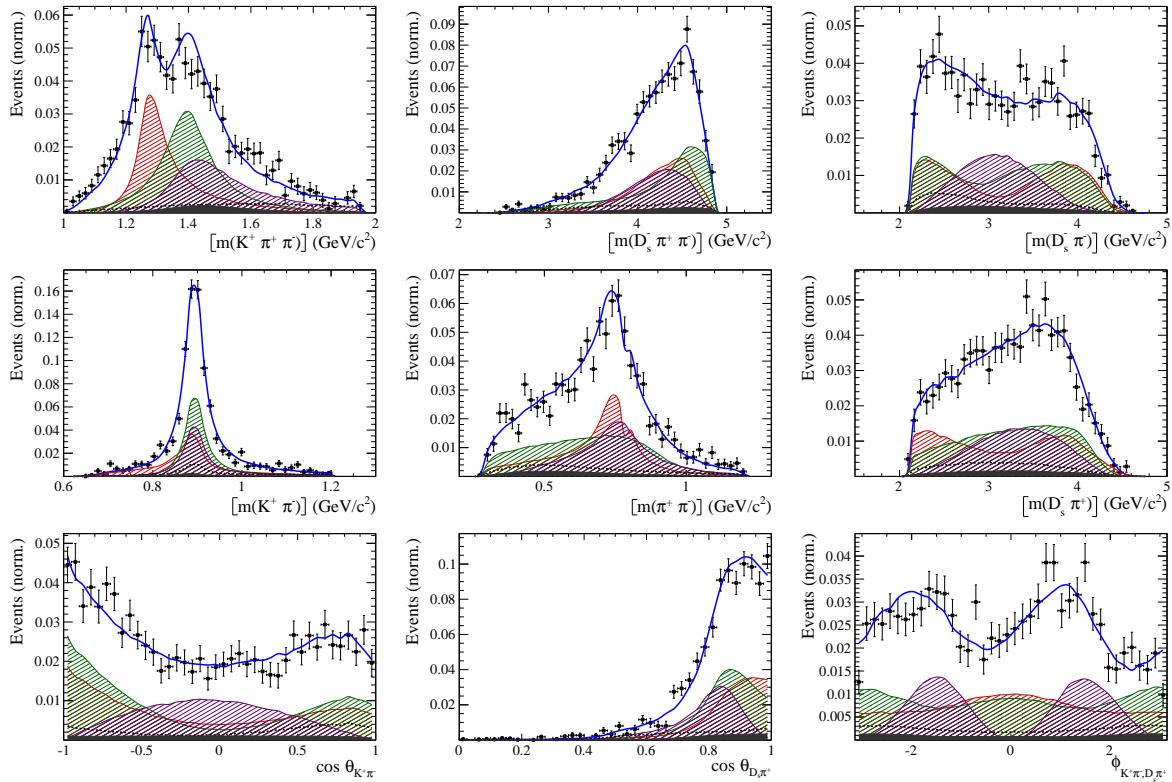


Figure 10.1

857 10.2 Results

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

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

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

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

867 which measures the interference effects between amplitude pairs. Constructive interference
 868 leads to $I_{ij} > 0$, while destructive interference leads to $I_{ij} < 0$. Note that $\sum_i F_i +$
 869 $\sum_{j < k} I_{j,k} = 1$.

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

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

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

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

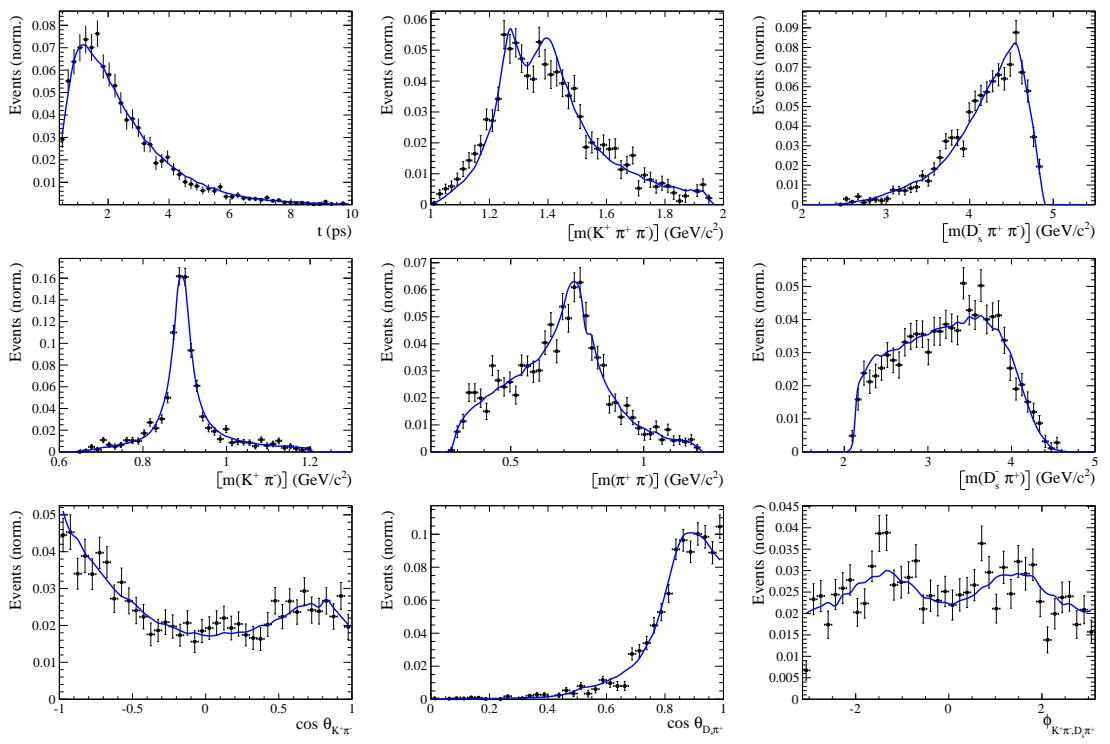


Figure 10.2

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

Decay Channel	$A_{b \rightarrow c}$		$A_{b \rightarrow u}$	
	$ a_i $	$\arg(a_i)[^\circ]$	$ a_i $	$\arg(a_i)[^\circ]$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	1.0	0.0	1.0	0.0
$K_1(1270) \rightarrow K^*(892) \pi$	$0.76 \pm 0.11 \pm 0.16$	$60.9 \pm 9.6 \pm 14.0$		
$K_1(1270) \rightarrow K_0^*(1430) \pi$	$0.68 \pm 0.06 \pm 0.34$	$116.5 \pm 5.1 \pm 43.5$		
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$2.53 \pm 0.27 \pm 0.57$	$12.9 \pm 7.4 \pm 8.0$	$0.67 \pm 0.20 \pm 0.51$	$-76.3 \pm 16.9 \pm 22.8$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$1.28 \pm 0.12 \pm 0.24$	$54.9 \pm 5.6 \pm 9.8$		
$K^*(1410) \rightarrow K \rho(770)$	$0.66 \pm 0.04 \pm 0.03$	$-172.9 \pm 5.0 \pm 6.5$		
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$			$0.77 \pm 0.11 \pm 0.62$	$-93.6 \pm 11.2 \pm 12.1$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$1.02 \pm 0.13 \pm 0.41$	$-28.4 \pm 8.0 \pm 10.4$	$0.79 \pm 0.18 \pm 0.35$	$3.7 \pm 12.5 \pm 14.8$
$B_s \rightarrow (D_s K)_P \rho(770)$			$0.61 \pm 0.08 \pm 0.26$	$36.4 \pm 7.7 \pm 14.1$
Fit parameter				
			Value	
$m_{K_1(1400)}$ [MeV]		1394.9	$\pm 8.8 \pm 12.6 \pm 21.2$	
$\Gamma_{K_1(1400)}$ [MeV]		224.0	$\pm 15.9 \pm 22.0 \pm 20.9$	
$m_{K^*(1410)}$ [MeV]		1419.6	$\pm 10.8 \pm 26.8 \pm 24.1$	
$\Gamma_{K^*(1410)}$ [MeV]		342.4	$\pm 23.5 \pm 51.0 \pm 52.9$	
r		xx.xx	$\pm 0.04 \pm 0.05 \pm 0.04$	
$\delta [^\circ]$		xx.xx	$\pm 16.1 \pm 6.2 \pm 6.8$	
$\gamma - 2\beta_s [^\circ]$		xx.xx	$\pm 16.1 \pm 11.4 \pm 6.2$	

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

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

887 11 Systematic uncertainties

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

895 11.1 Models for B_s^0 mass distribution

896 The statistical subtraction of the residual background [45], left after the full selection,
 897 relies on the correct description of the invariant B_s^0 mass distribution. Since the choice
 898 of signal and background models is not unique, alternative descriptions which lead to
 899 slightly different yields for the signal and background components are available. The
 900 difference in yields could result in shifted values for the measured observables and are
 901 therefore treated as systematic uncertainty.

902

903 11.1.1 Signal model

904 The Johnson's SU function which is used as nominal signal model is replaced by a double
 905 Crystal Ball [50]. The crystal ball model is given by a gaussian core with an exponential
 906 tail on one side. Choosing a double Crystal Ball allows for asymmetric tails in a slightly
 907 different way compared to the Johnson's SU function.

908 11.1.2 Background model

909 For the description of the partially reconstructed background, a combination of the
 910 RooHILLdini and RooHORNsdini model [51] is used instead of the nominal model of
 911 three bifurcated gaussians. The HORNsdini model is used to describe the $B_s^0 \rightarrow D_s^* [\rightarrow$
 912 $D_s(\pi^0)] X_{s/d}$ decay, where the brackets around the π^0 indicate that it is missed in the
 913 reconstruction. The $D_s^* \rightarrow D_s\pi^0$ decay is a Vector \rightarrow Scalar-Scalar ($1^- \rightarrow 0^-0^-$)
 914 transition. Using the helicity of the D_s , one can show that this results in a double-peak
 915 structure in the reconstructed B_s^0 mass. Therefore, the HORNsdini shape consists of a
 916 gaussian-like double-peak structure:

$$917 HORN S(m_{B_s^0}) = \int_a^b dm_{B_s^0} \left(m_{B_s^0} - \frac{a+b}{2} \right)^2 \mathcal{D}\mathcal{G}(m_{B_s^0} | \mu, \sigma, f_G) \left(\frac{1-\zeta}{b-a} m_{B_s^0} + \frac{b\zeta - a}{b-a} \right), \quad (11.1)$$

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

920 The HILLdini model parametrizes the invariant mass shape of $B_s^0 \rightarrow D_s^* [\rightarrow D_s(\gamma)] X_{s/d}$
 921 candidates, where the γ is not reconstructed. Contrary to the previously discussed process,
 922 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 ??.

953 **11.2 Decay-time acceptance**

954 To investigate the systematic uncertainty related to the decay-time dependent efficiency,
 955 we vary our parametrization of the acceptance using cubic splines. This is explicitly
 956 done by choosing slightly different knot positions, varying the spline coefficients at the
 957 nominal positions within their statistical uncertainties and adding/subtracting knots
 958 in the range $0.4 \text{ ps} < t < 11 \text{ ps}$. Additionally, an adaptive binning scheme is used to
 959 determine the knot positions in a way that roughly equal amounts of data is covered
 960 between two knots. Strictly speaking, the variation of the spline coefficients within their
 961 uncertainty gives the statistical uncertainty of the decay-time acceptance parametrization.
 962 For the presented measurement, this is done using the Cholesky decomposition [52] of
 963 the covariance matrix of coefficients c_i , generating toy splines with randomized coefficient
 964 values $c_{i,toy}$ from this decomposition and refitting using the toy spline. Furthermore, the
 965 fit to the decay-time distribution of signal $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates, used to determine
 966 the spline parametrization, is reiterated with varying fixed/constrained values for $\Delta\Gamma_s$.

967 **11.2.1 Varition of knot positions**

968 The nominal knot positions are changed to be:

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

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

971 **11.2.2 Variation of spline coefficients**

972 Due to the sizeable correlation of the spline coefficients c_i determined in Chapter 6.2, the
 973 variations of the observables in the amplitude fit when changing one spline coefficient can
 974 not be added up in quadrature for all coefficients. To simplify the problem, a Cholesky
 975 decomposition [52] is used to generate a set of uncorrelated vectors from the covariance
 976 matrix A_{cov} . It can be shown that every Hermitian positive-definite matrix, such as A_{cov} ,
 977 has a unique Cholesky decomposition of the form:

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

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

980

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

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

983 where v_{ij} are real and positive numbers. L contains four row vectors, which are by
 984 construction the four decorrelated modes of the covariant matrix A_{cov} . From this modes,
 985 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 ?? .

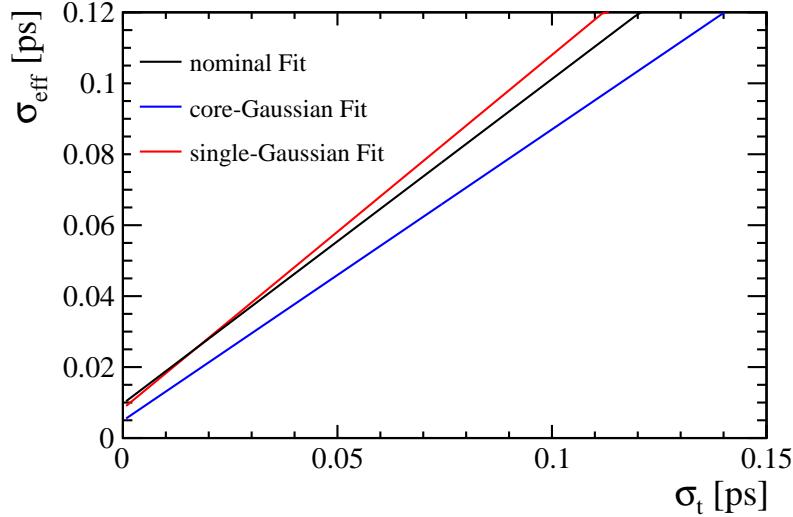


Figure 11.1: The measured resolution scaling function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data) for (black line) the nominal scaling, (blue line) only using the narrow gaussian width of the double gaussian fit model or (red line) when determining the resolution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

1019 11.4 Tagging calibration

1020 Systematic uncertainties arise from the statistical precision of the tagging parameters
 1021 determined from the calibration, discussed in Sec. 7. These uncertainties are accounted
 1022 for by the inclusion of Gaussian constrains in the nominal fit. The width of the respective
 1023 constrain for the tagging parameter p_i is chosen to be Δp_i . In this way, the systematic
 1024 uncertainty due to the tagging calibration is included in the statistical uncertainty of the
 1025 time dependent fit.

1026 11.5 Summary of systematic uncertainties

1027 All contributing systematic uncertainties are summarized in Table ???. The individual
 1028 uncertainties are summed in quadrature to arrive at the total systematic uncertainty
 1029 for the respective CP observable. Their total magnitude ranges from (30-40)% of the
 1030 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 $B_s \rightarrow D_s \pi\pi$ 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	Resolution	Δm_s	Asymmetries	Background	Lineshapes	Resonances m, Γ	Form-Factors	Phsp-Acc.	Amp. Model	Total
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892)\pi) \text{Mag}$	0.01	0.01	0.02	0.15	1.30	0.28	0.42	0.06		1.41
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892)\pi) \text{Phase}$	0.03	0.01	0.16	0.06	0.85	0.31	0.20	1.10		1.45
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430)\pi) \text{Mag}$	0.02	0.01	0.11	0.25	3.96	3.69	0.45	2.20		5.87
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430)\pi) \text{Phase}$	0.02	0.01	0.18	0.15	7.28	0.21	0.51	4.47		8.57
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi) \text{Mag}(b \rightarrow c)$	0.02	0.03	0.43	0.27	1.38	0.28	0.38	1.44		2.12
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi) \text{Phase}(b \rightarrow c)$	0.01	0.03	0.13	0.28	0.66	0.25	0.32	0.69		1.08
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi) \text{Mag}(b \rightarrow u)$	0.02	0.04	0.06	0.19	0.83	0.24	0.56	2.27		2.50
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi) \text{Phase}(b \rightarrow u)$	0.04	0.10	0.15	0.36	0.79	0.43	0.25	0.88		1.35
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892)\pi) \text{Mag}(b \rightarrow c)$	0.03	0.05	0.19	0.18	1.08	0.28	1.60	0.09		1.97
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892)\pi) \text{Phase}(b \rightarrow c)$	0.01	0.01	0.21	0.10	1.42	0.22	0.75	0.62		1.76
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770)) \text{Mag}$	0.01	0.01	0.12	0.17	0.60	0.18	0.19	0.15		0.70
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770)) \text{Phase}$	0.01	0.01	0.10	0.13	0.34	0.12	0.29	1.22		1.31
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892)\pi) \text{Mag}(b \rightarrow u)$	0.02	0.04	0.23	0.22	0.68	0.76	5.39	1.96		5.84
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892)\pi) \text{Phase}(b \rightarrow u)$	0.03	0.04	0.13	0.21	0.64	0.40	0.48	0.55		1.08
$B_s \rightarrow (D_s \pi)_P K^*(892) \text{Mag}(b \rightarrow c)$	0.02	0.02	0.35	0.24	1.28	0.20	2.66	0.99		3.15
$B_s \rightarrow (D_s \pi)_P K^*(892) \text{Phase}(b \rightarrow c)$	0.01	0.01	0.20	0.47	0.95	0.18	0.34	0.59		1.29
$B_s \rightarrow (D_s \pi)_P K^*(892) \text{Mag}(b \rightarrow u)$	0.04	0.03	0.37	0.13	0.47	0.27	1.73	0.68		1.98
$B_s \rightarrow (D_s \pi)_P K^*(892) \text{Phase}(b \rightarrow u)$	0.05	0.03	0.88	0.28	0.56	0.21	0.42	0.10		1.18
$B_s \rightarrow (D_s K)_P \rho(770) \text{Mag}(b \rightarrow u)$	0.01	0.05	0.83	0.49	1.34	0.38	2.81	0.33		3.30
$B_s \rightarrow (D_s K)_P \rho(770) \text{Phase}(b \rightarrow u)$	0.02	0.03	0.24	0.66	0.25	0.60	0.71	1.37		1.82
$m_{K_1(1400)}$	0.02	0.01	0.36	0.17	1.15	0.16	0.33	0.66	2.41	2.80
$\Gamma_{K_1(1400)}$	0.02	0.01	0.29	0.13	1.23	0.12	0.25	0.46	1.31	1.91
$m_{K^*(1410)}$	0.01	0.01	0.51	0.11	1.69	0.27	1.63	0.51	2.22	3.32
$\Gamma_{K^*(1410)}$	0.01	0.01	0.10	0.18	1.17	0.59	1.71	0.15	2.25	3.12
r	0.05	0.10	0.38	0.29	1.02	0.20	0.18	0.58	0.98	1.63
δ	0.04	0.06	0.03	0.10	0.24	0.07	0.14	0.23	0.42	0.57
$\gamma - 2\beta_s$	0.05	0.07	0.28	0.25	0.30	0.29	0.06	0.42	0.39	0.81

1031 A Stripping and Trigger cuts

1032 The following text describes variables which are used in Table 1.1 and might be ambiguous,
 1033 or which benefits are not straight forward. Where noted, different cut values are applied
 1034 for Run-I and Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest
 1035 approach. This variable is used to ensure that all D_s and $X_{s,d}$ daughters originate from
 1036 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
$B_s^0 \tau_{B_s^0}$	$> 0.2 \text{ ps}$	

1038 Table 1.2 summarizes the trigger requirements imposed by the Hlt1 line used in this
 1039 analysis for Run-I. At least one of the six decay particles must pass the listed requirements
 1040 in order for the event to be stored for further analysis. For Run-II, this trigger line was
 1041 updated and uses a multivariate classifier which takes the variables listed in Table 1.2 as
 1042 input, rather than directly cutting on them.

1043 The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the
 1044 b-hadron p_T , its flight distance χ^2 from the nearest PV and the sum of the B_s^0 and D_s
 1045 vertex χ^2 divided by the sum of their number of degrees of freedom. Table 1.3 summarizes
 1046 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

1047

B Details of multivariate classifier

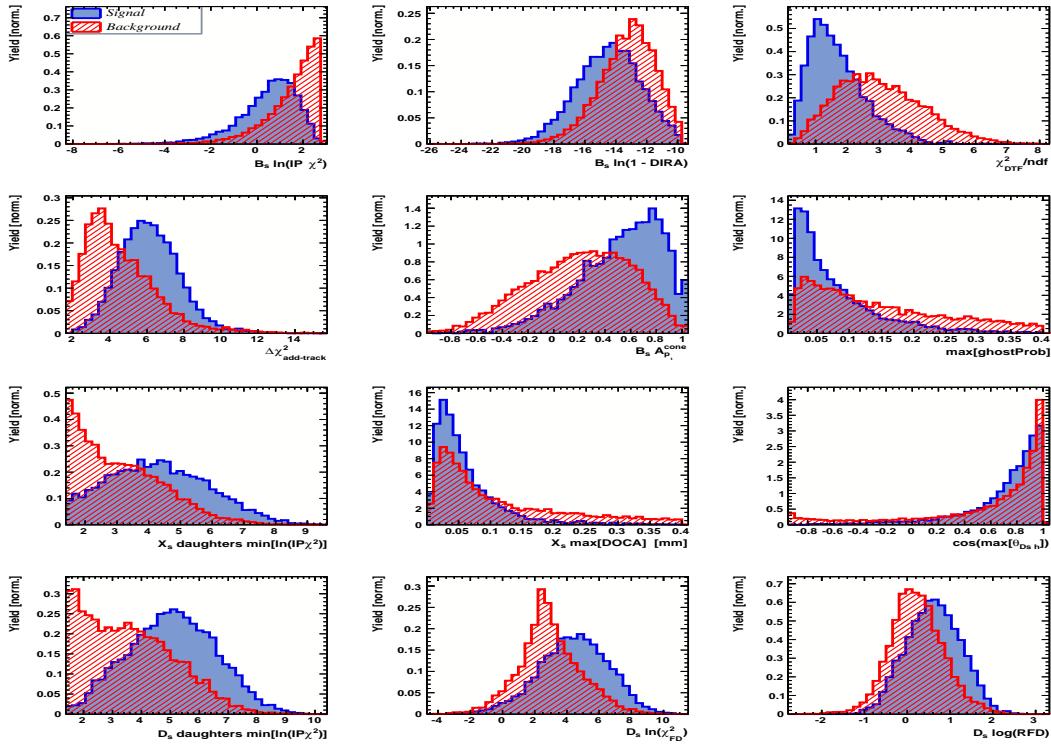


Figure A.1: Variables used to train the BDTG for category [Run-I,L0-TOS].

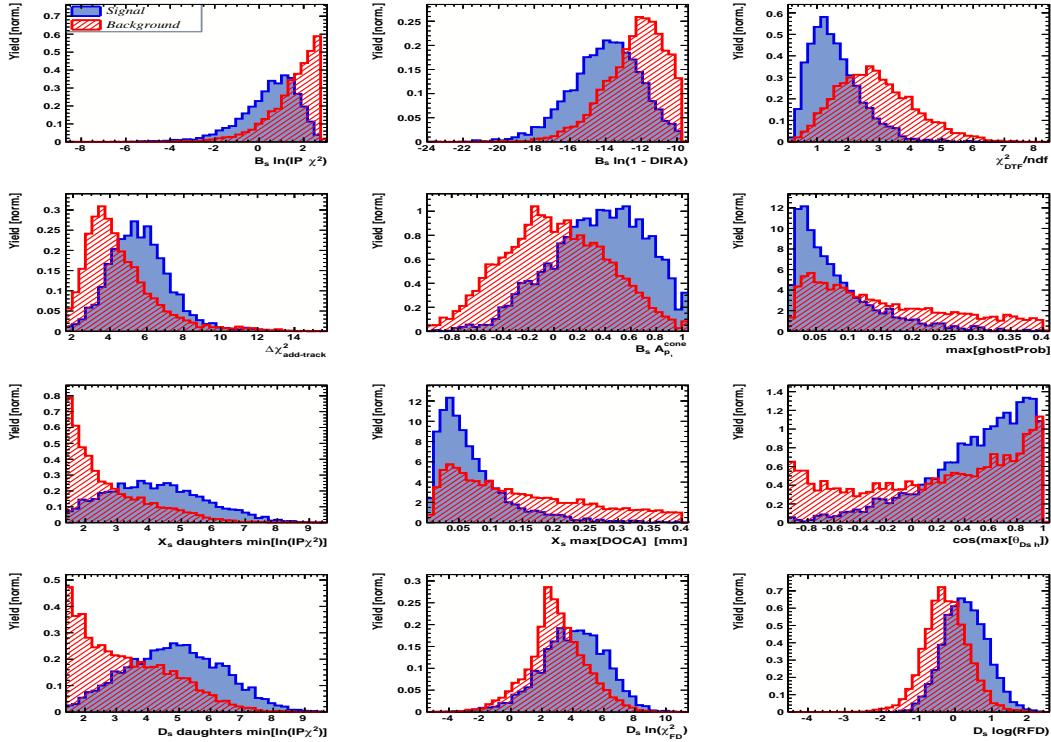


Figure A.2: Variables used to train the BDTG for category [Run-I,L0-TIS].

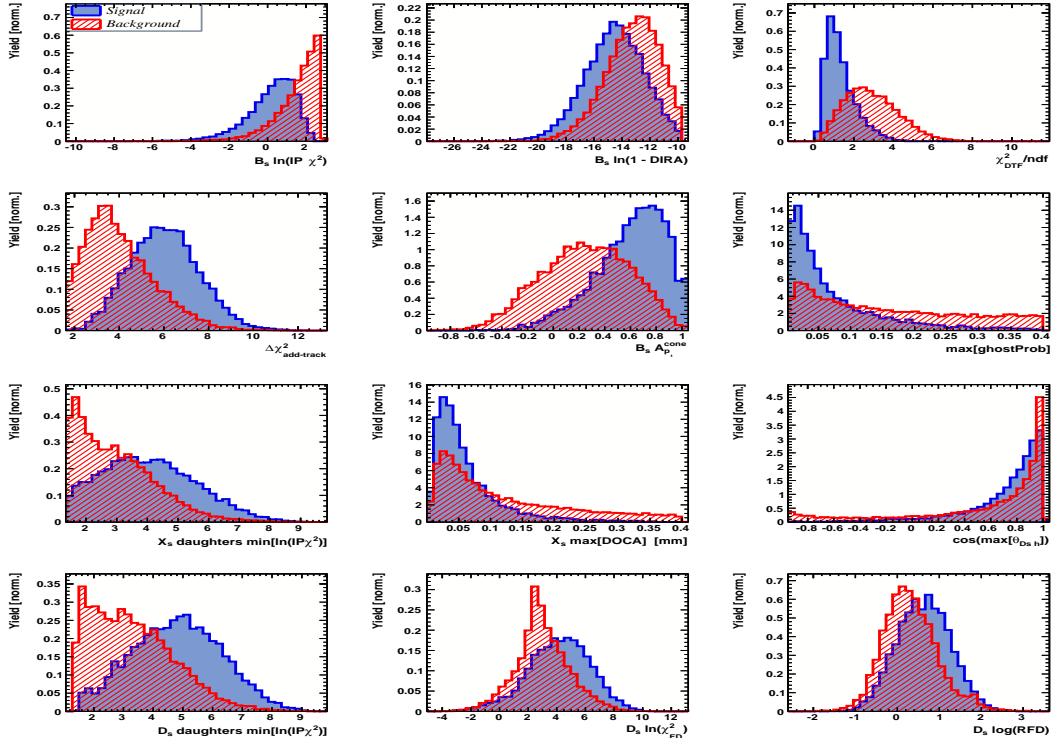


Figure A.3: Variables used to train the BDTG for category [Run-II,L0-TOS].

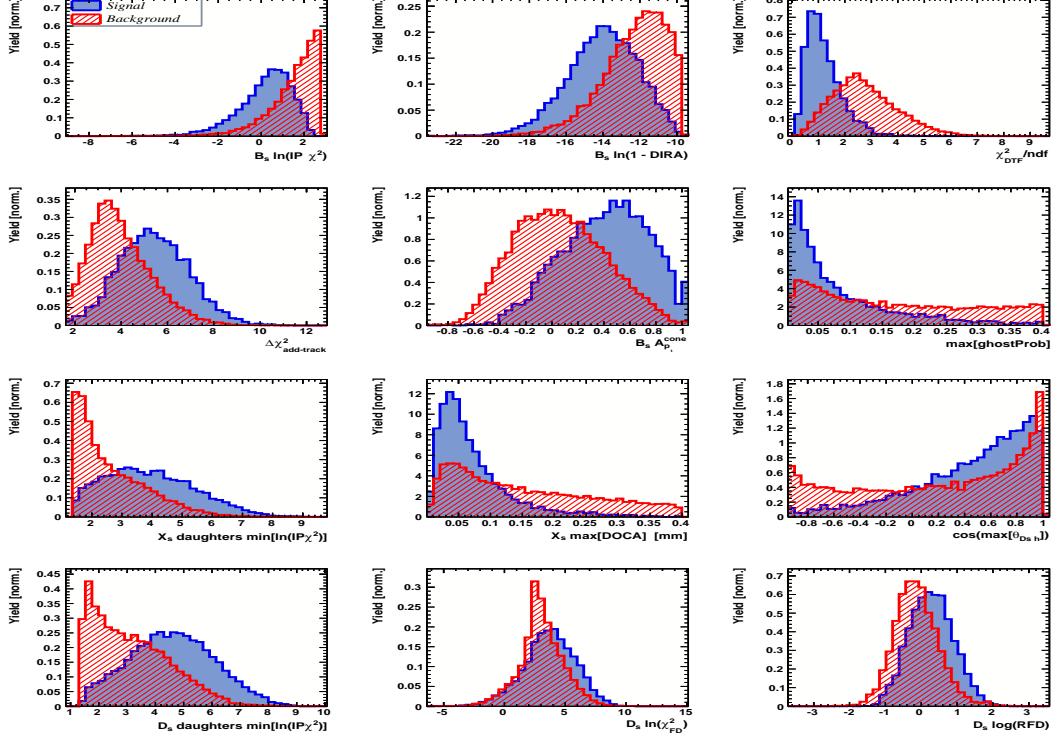


Figure A.4: Variables used to train the BDTG for category [Run-II,L0-TIS].

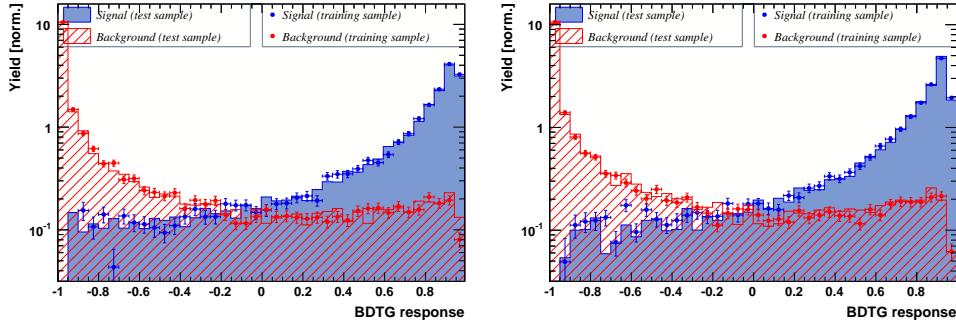


Figure A.5: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].

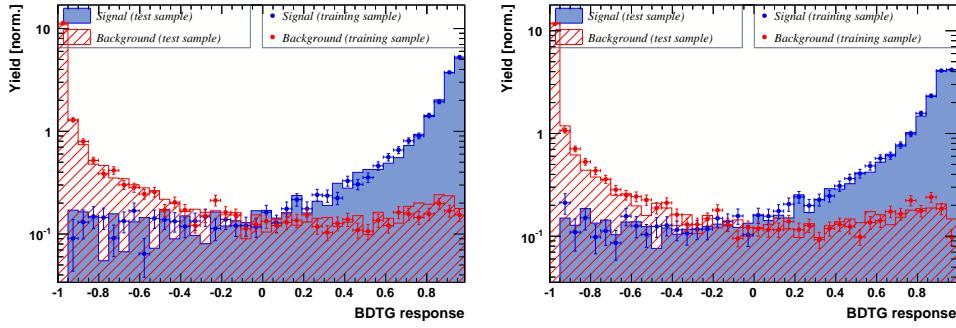


Figure A.6: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].

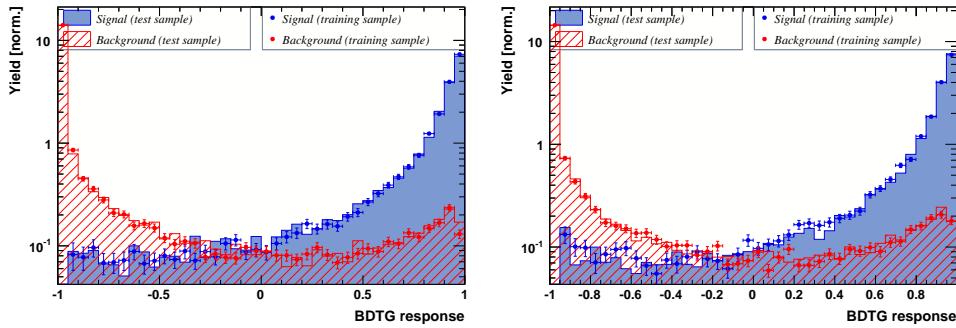


Figure A.7: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].

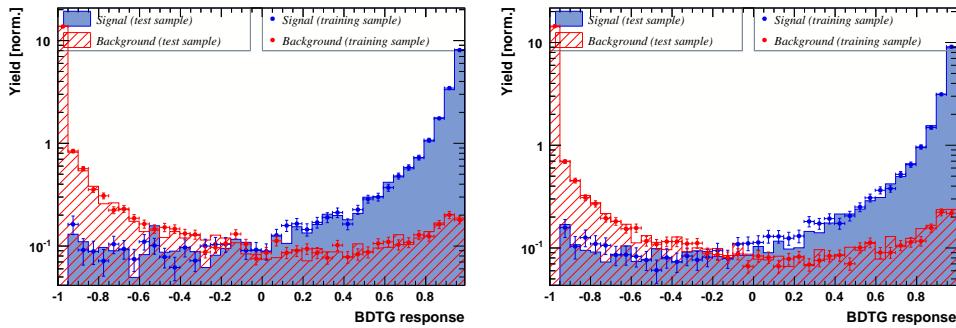


Figure A.8: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].

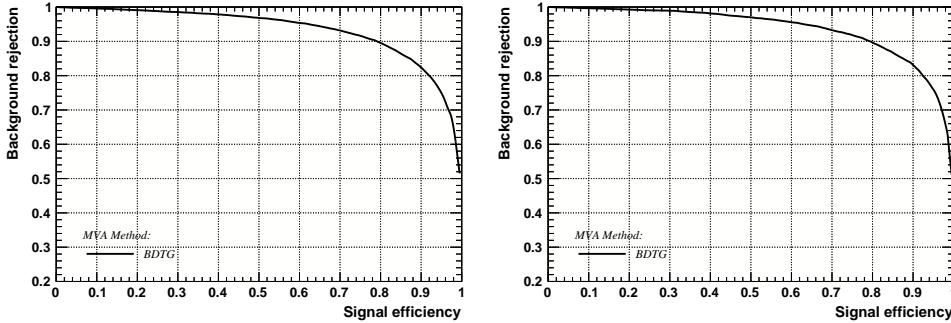


Figure A.9: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].

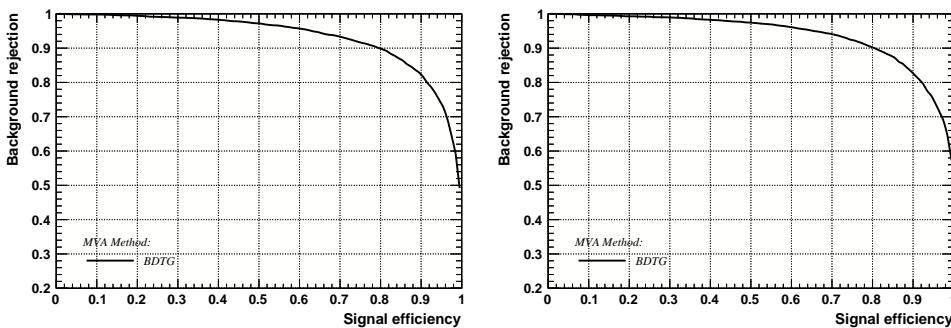


Figure A.10: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TIS].

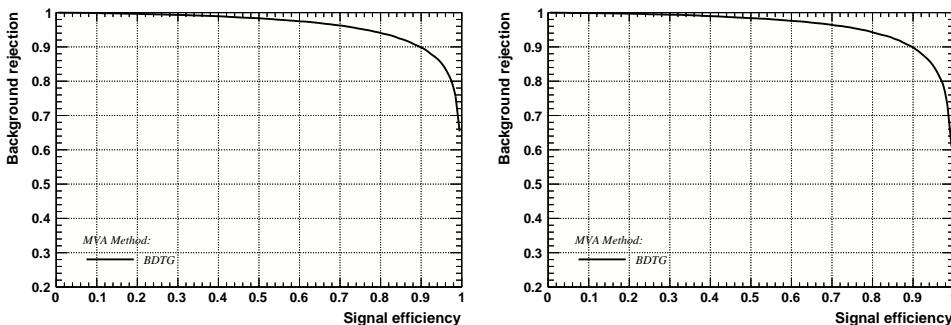


Figure A.11: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TOS].

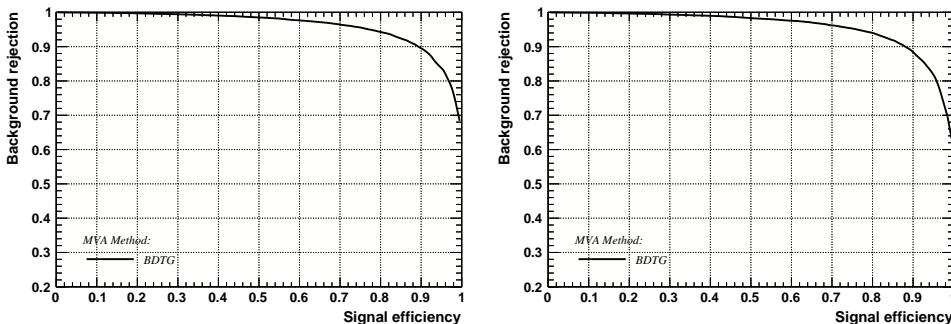


Figure A.12: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TIS].

1049 C Detailed mass fits

1050 In this section, all fits to the mass distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ and $B_s^0 \rightarrow D_sK\pi\pi$
 1051 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-
 1052 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
 1053 L0 trigger category.

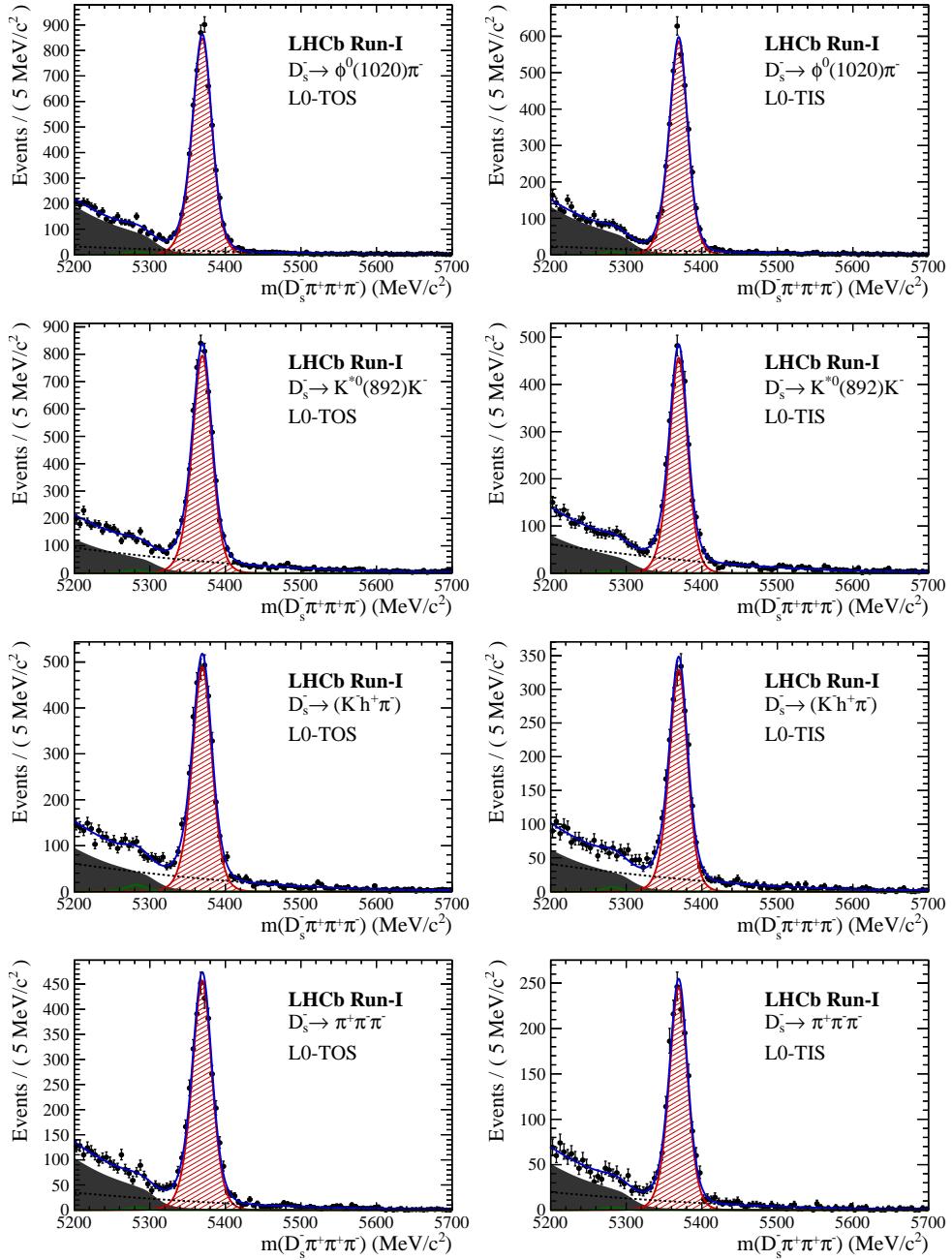


Figure B.1: Invariant mass distributions of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates for Run-I data.

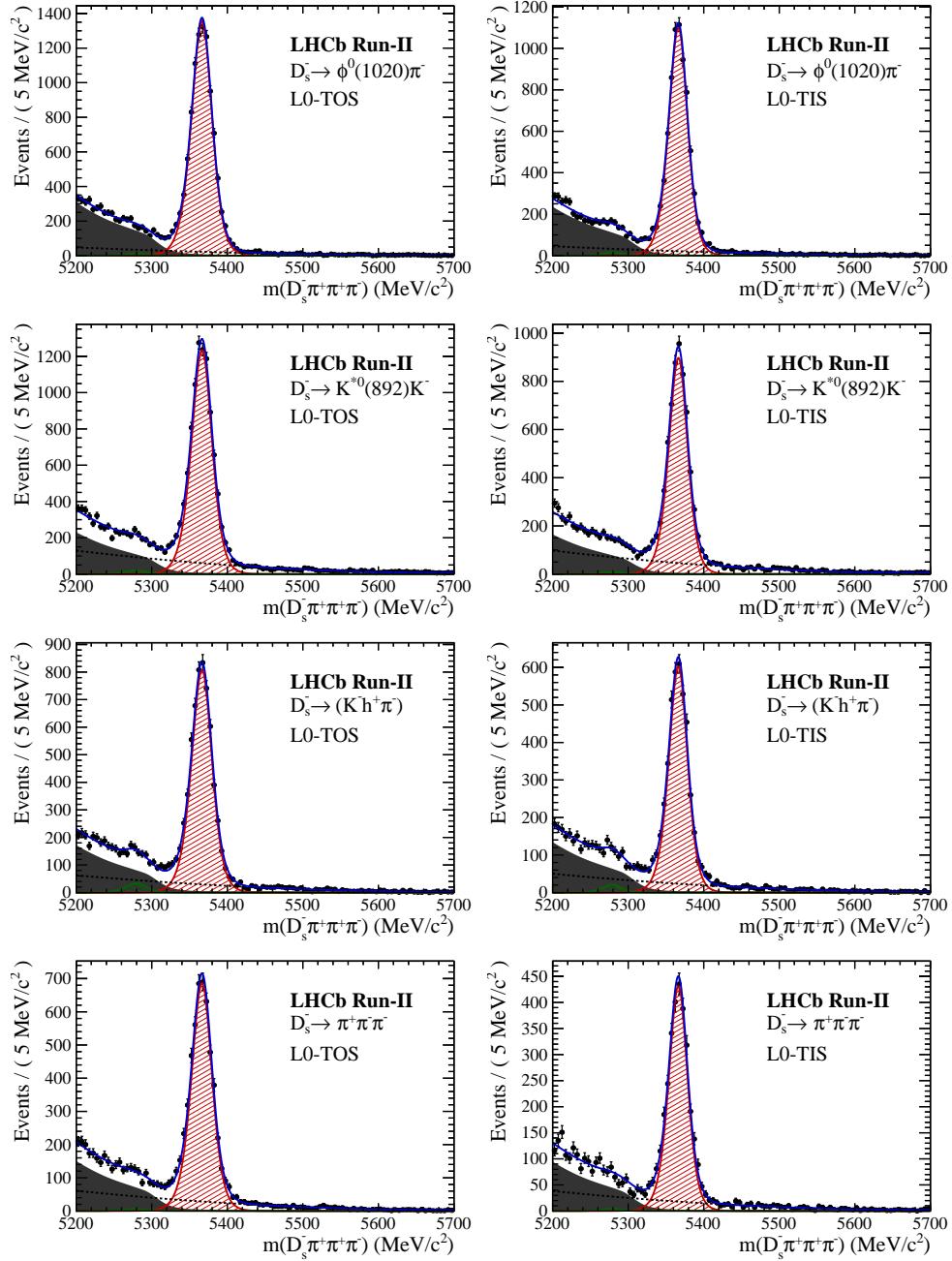


Figure B.2: Invariant mass distributions of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates for Run-II data.

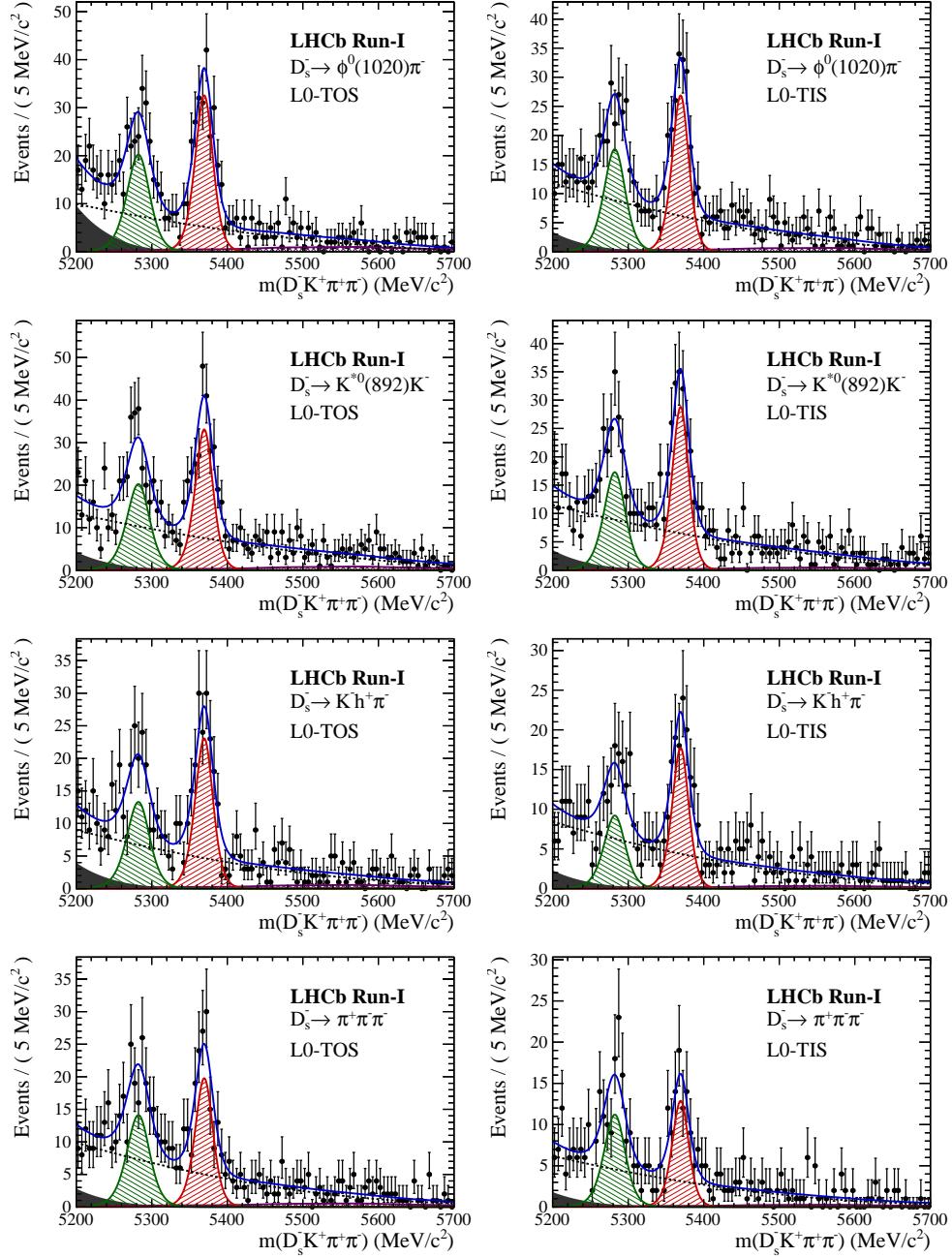


Figure B.3: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-I data.

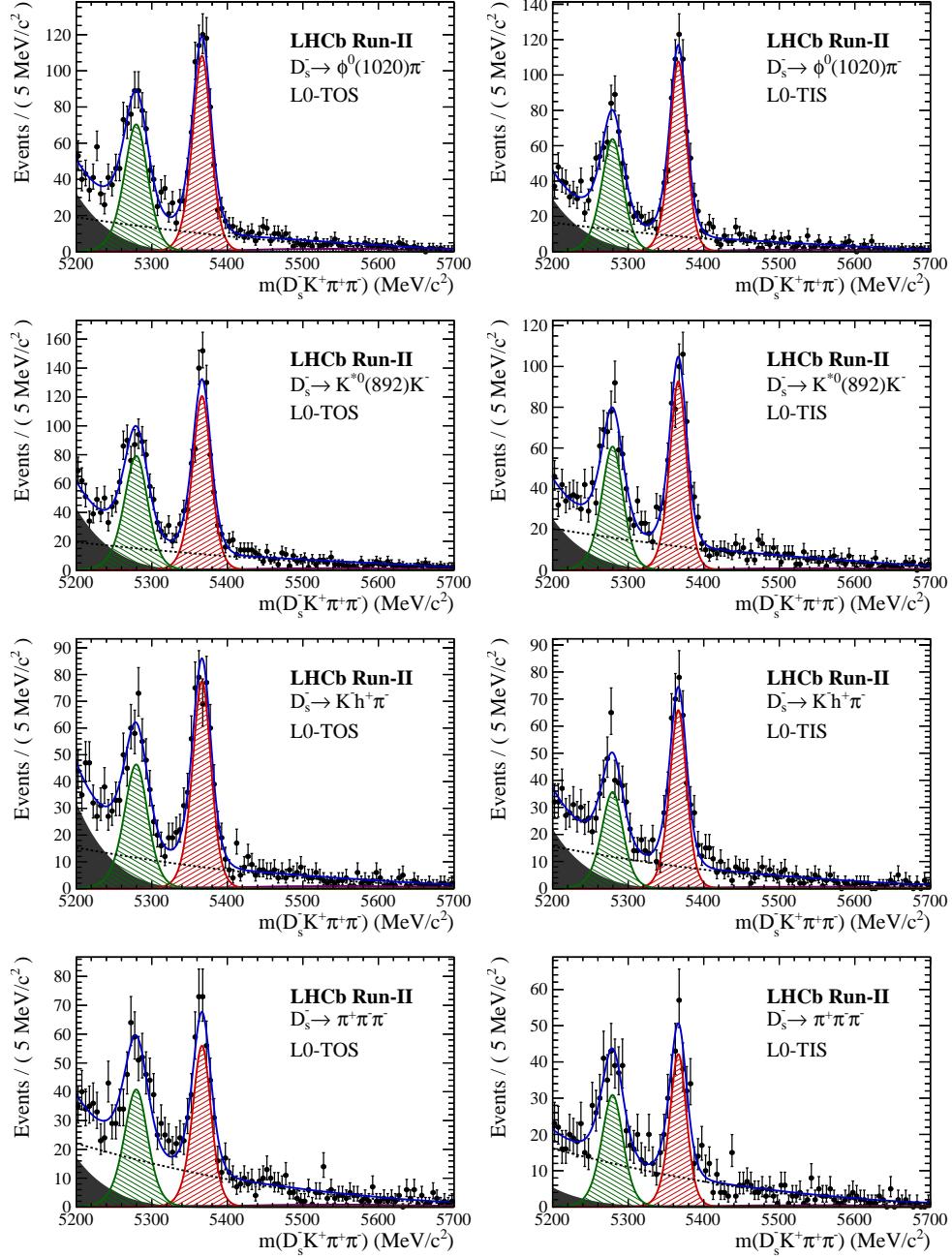


Figure B.4: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-II data.

1054 D Decay-time Resolution fits

1055 This section contains all fits to the distributions of the decay time difference Δt between
 1056 the true and the reconstructed decay time of the truth-matched B_s^0 candidates on MC.
 1057 The fits are performed in bins of the decay time error σ_t , where an adaptive binning
 1058 scheme is used to ensure that approximately the same number of events are found in each
 1059 bin.

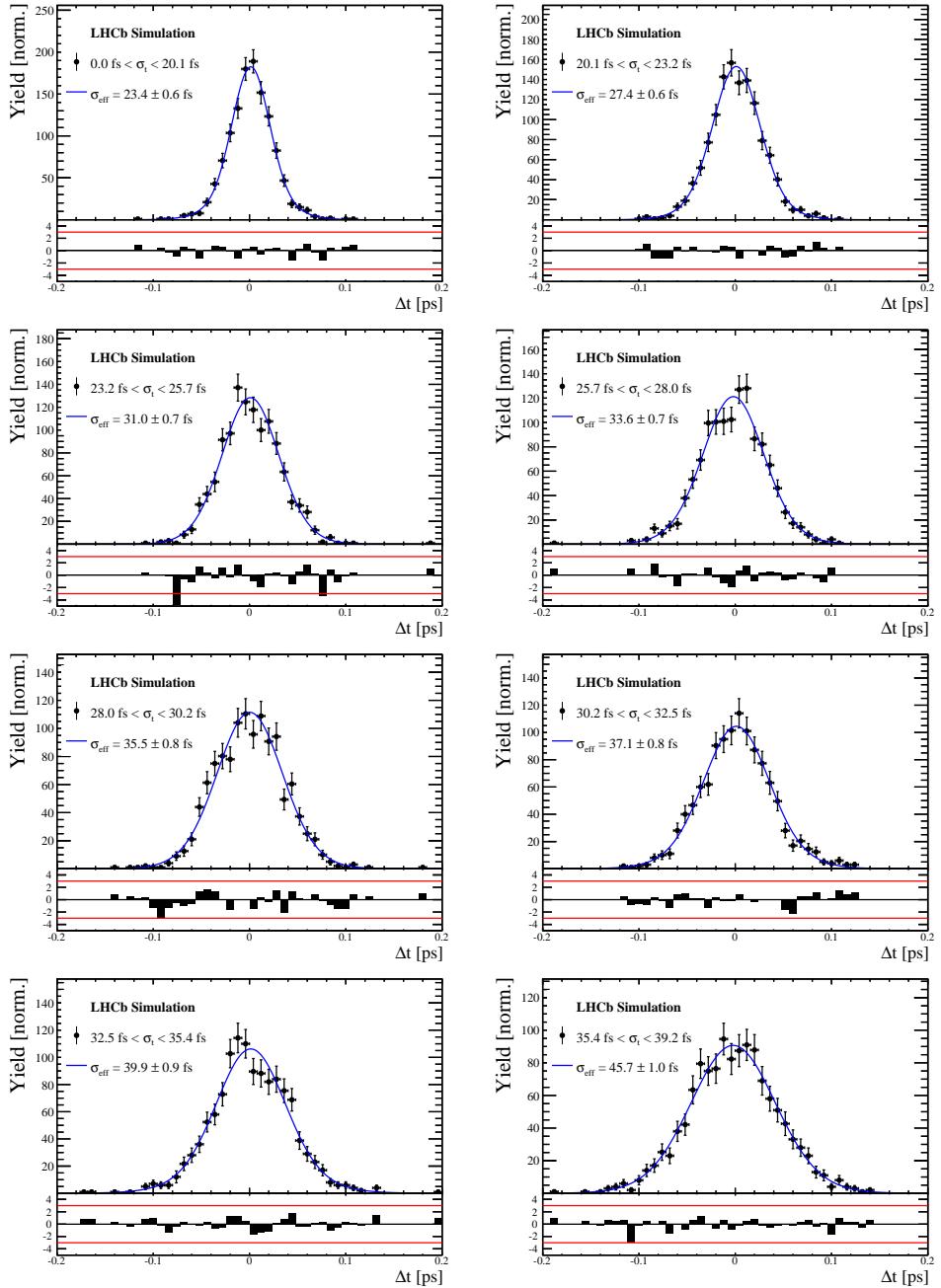


Figure C.1: Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K \pi \pi$ MC candidates in bins of the per-event decay time error estimate..

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.

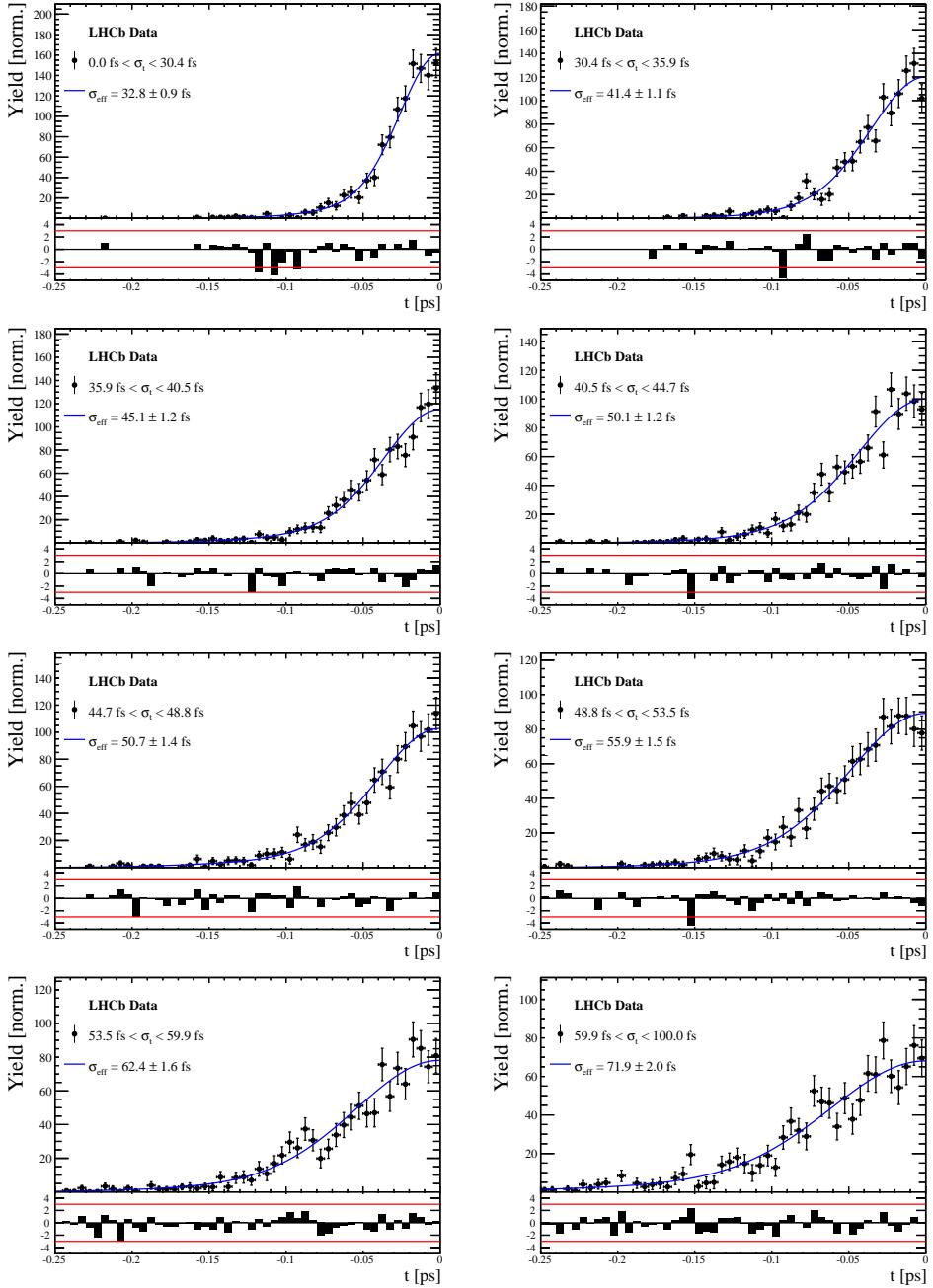


Figure C.2: Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks, for bins in the per-event decay time error estimate.

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.

1060 E Spin Amplitudes

1061 The spin factors used for $B \rightarrow P_1 P_2 P_3 P_4$ decays are given in Table 5.1.

Table 5.1: 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	$B[D] \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\gamma}(T_1) L_{(2)\gamma}^\beta(T_2)$

1062 F Considered Decay Chains

1063 The various decay channels considered in the model building are listed in Table 6.1.

Table 6.1: Decays considered in the LASSO model building.

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

¹⁰⁶⁴ G MC corrections

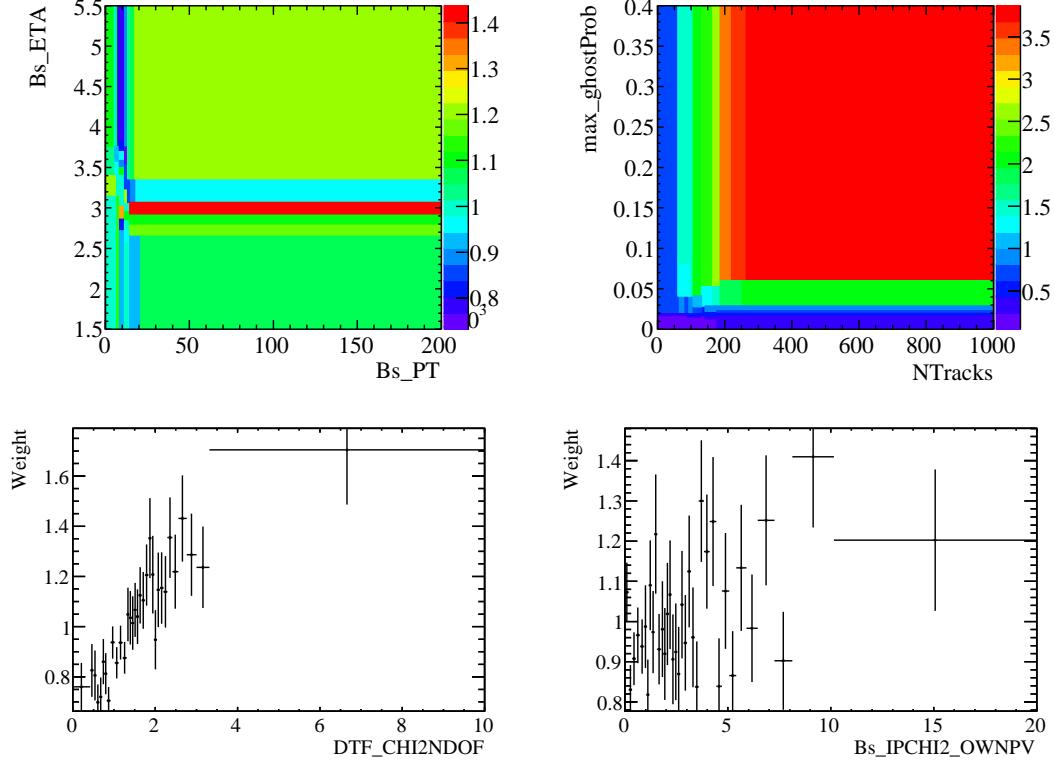


Figure C.1: Weights applied to correct for Data/MC differences.

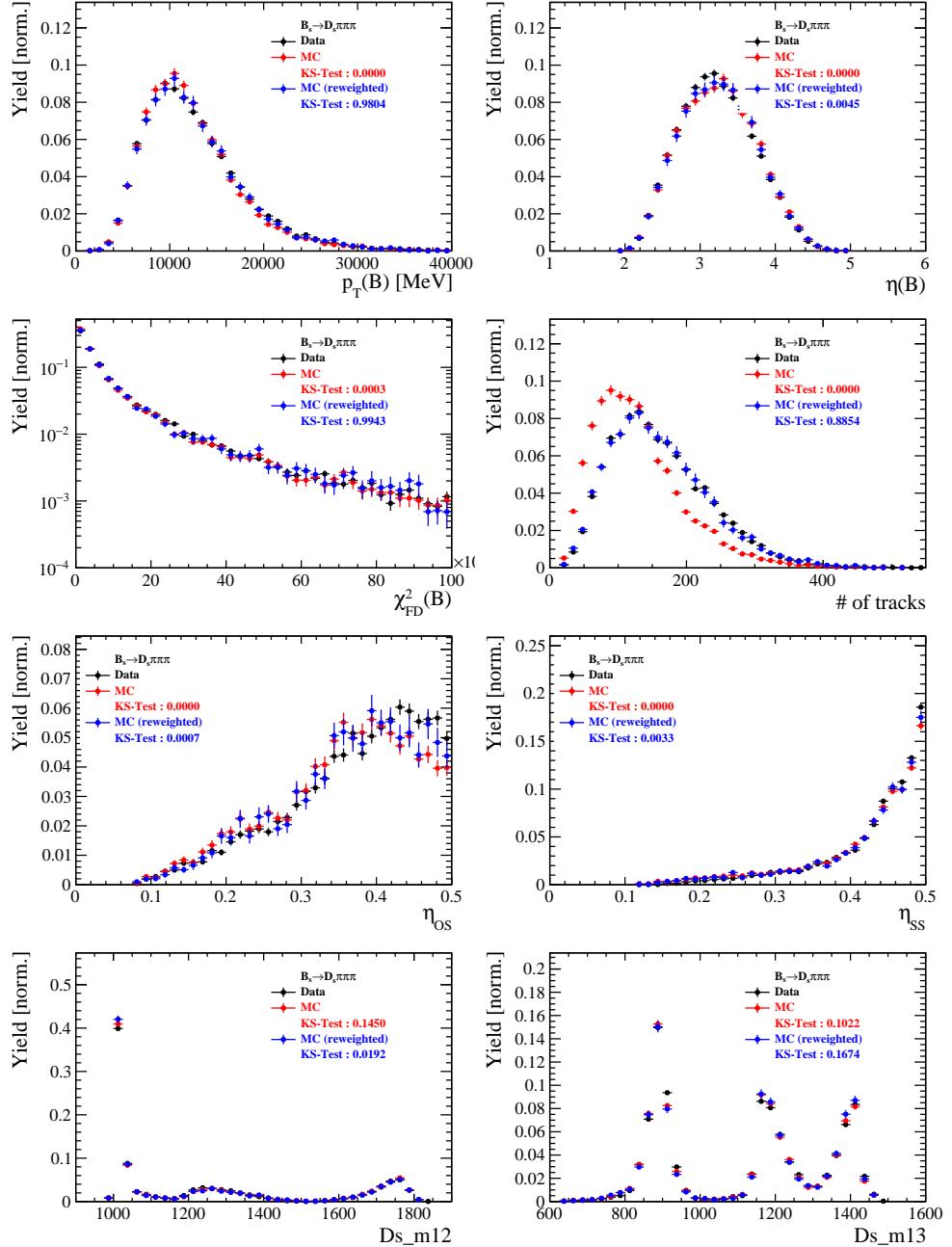


Figure C.2: Comparison of selected variables for $B_s \rightarrow D_s \pi\pi\pi$ decays.

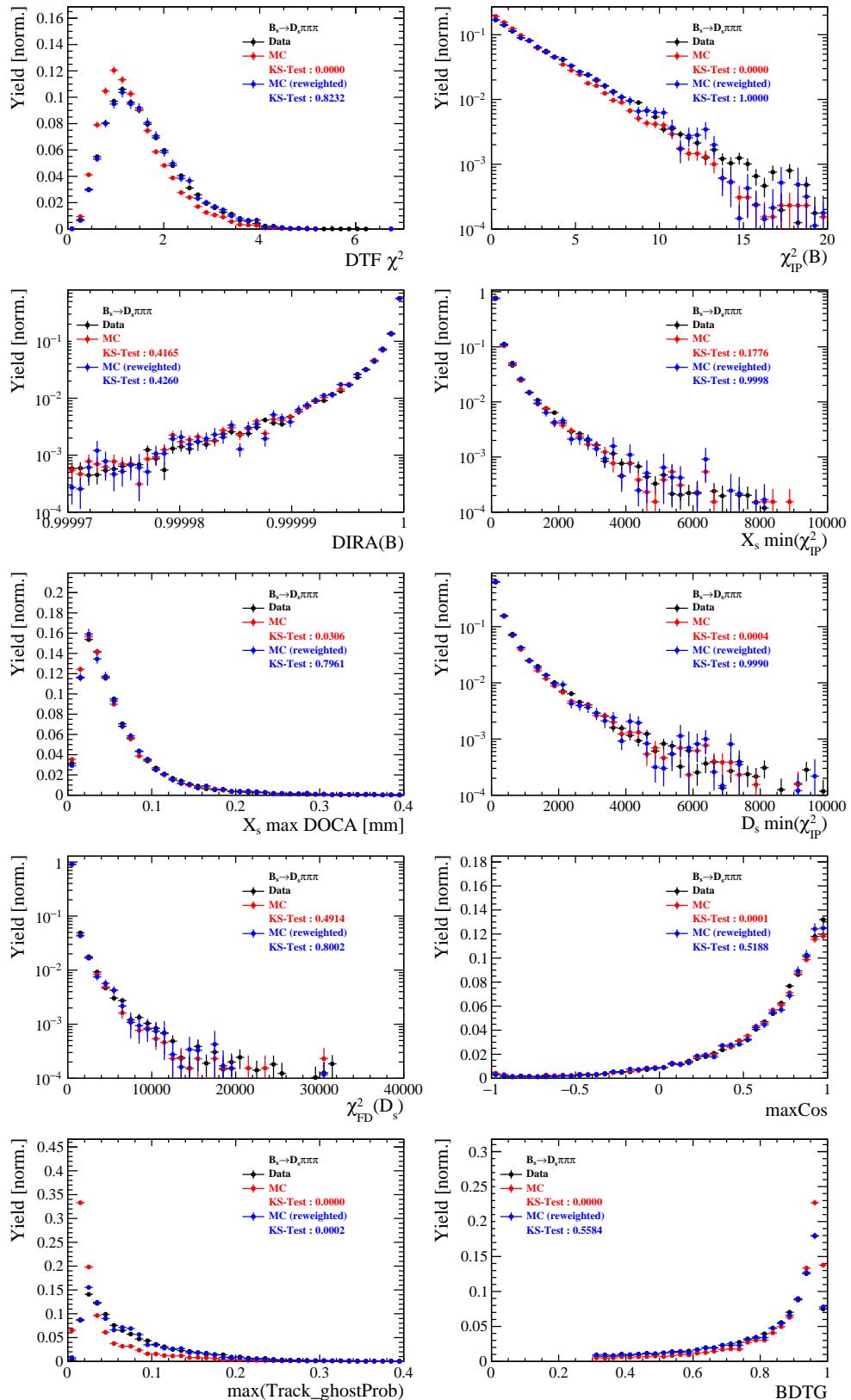


Figure C.3: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s \pi\pi\pi$ decays.

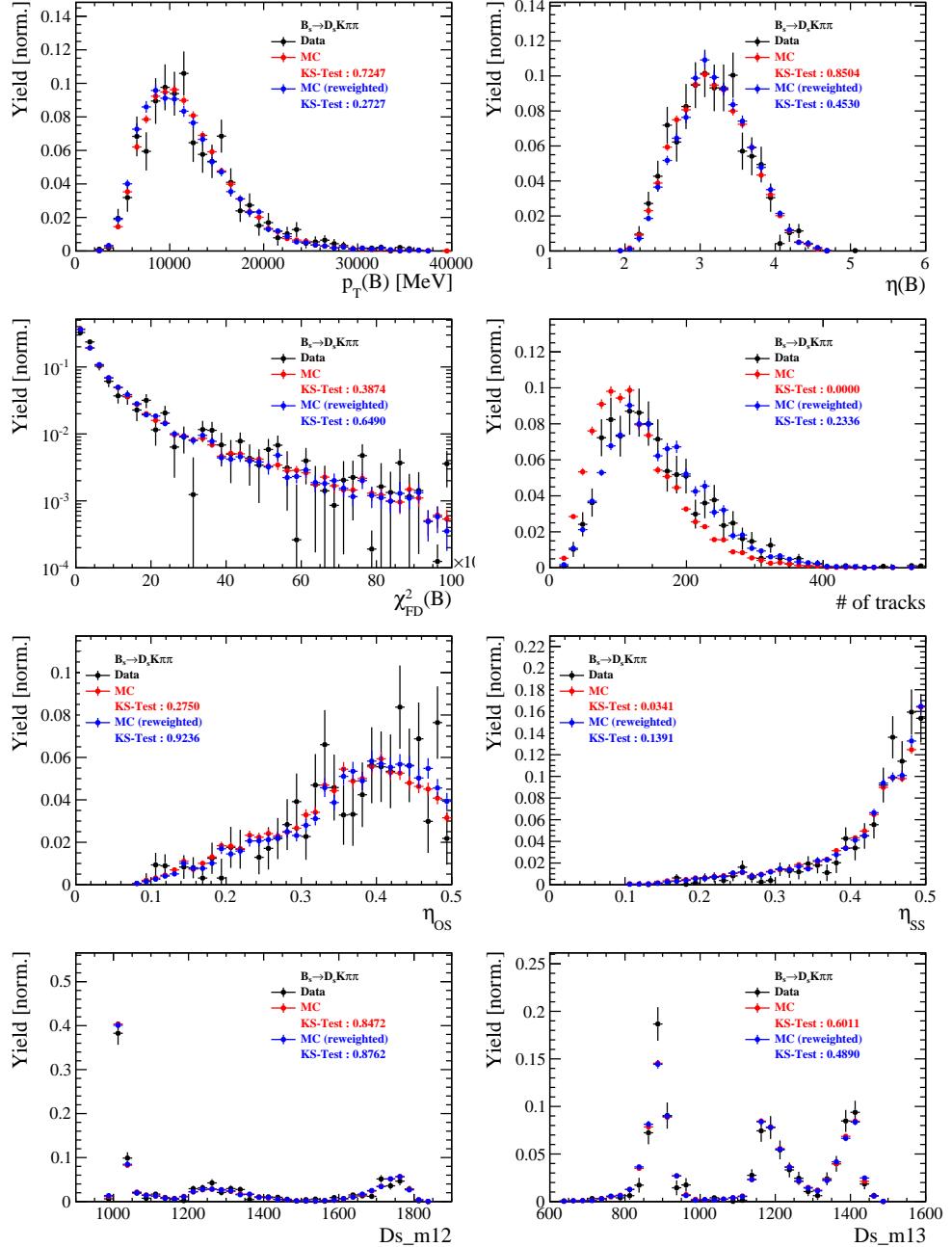


Figure C.4: Comparison of selected variables for $B_s \rightarrow D_s K\pi\pi$ decays.

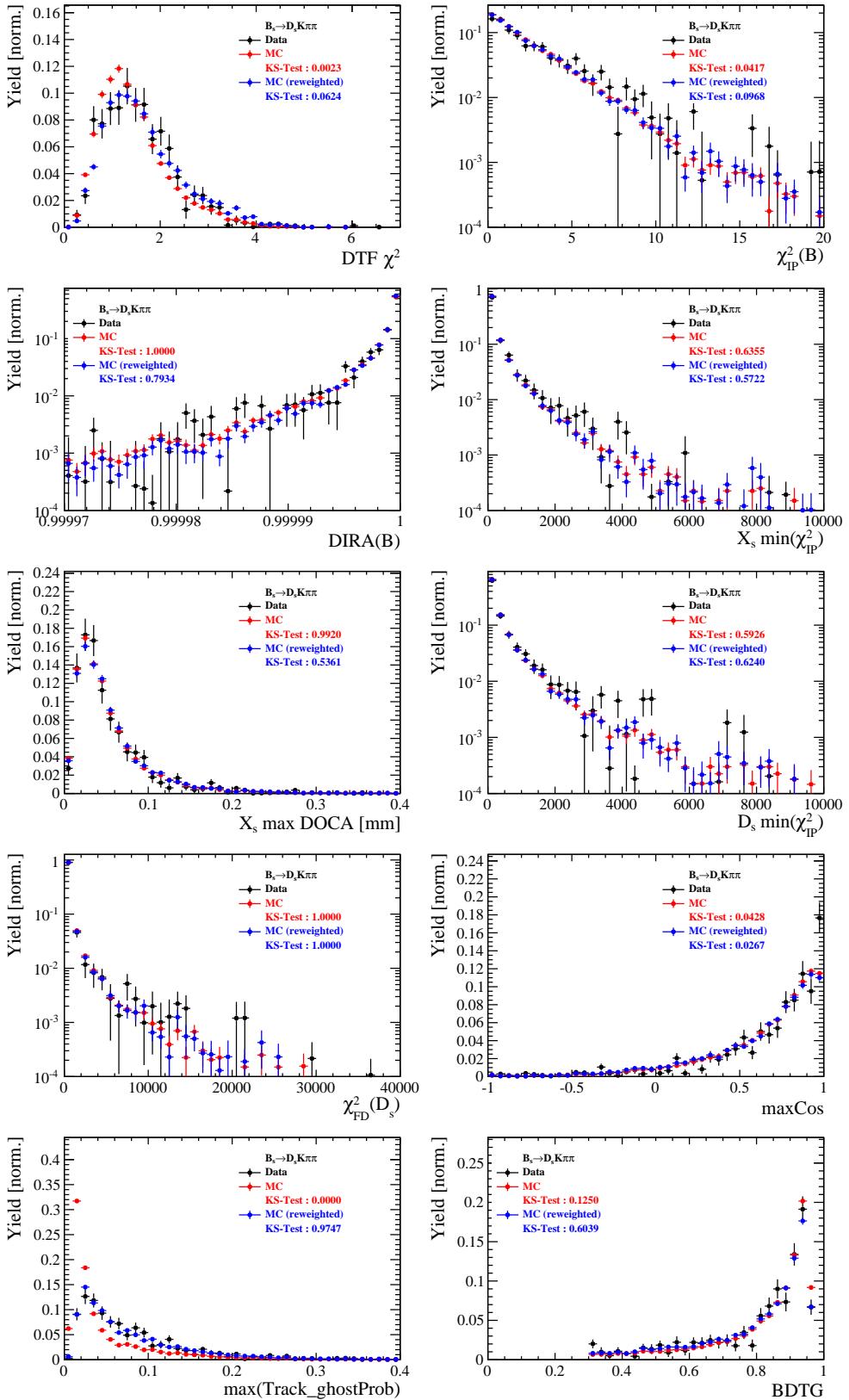


Figure C.5: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s K\pi\pi$ decays.

1065 **H Data distributions**

1066 **H.1 Comparison of signal and calibration channel**

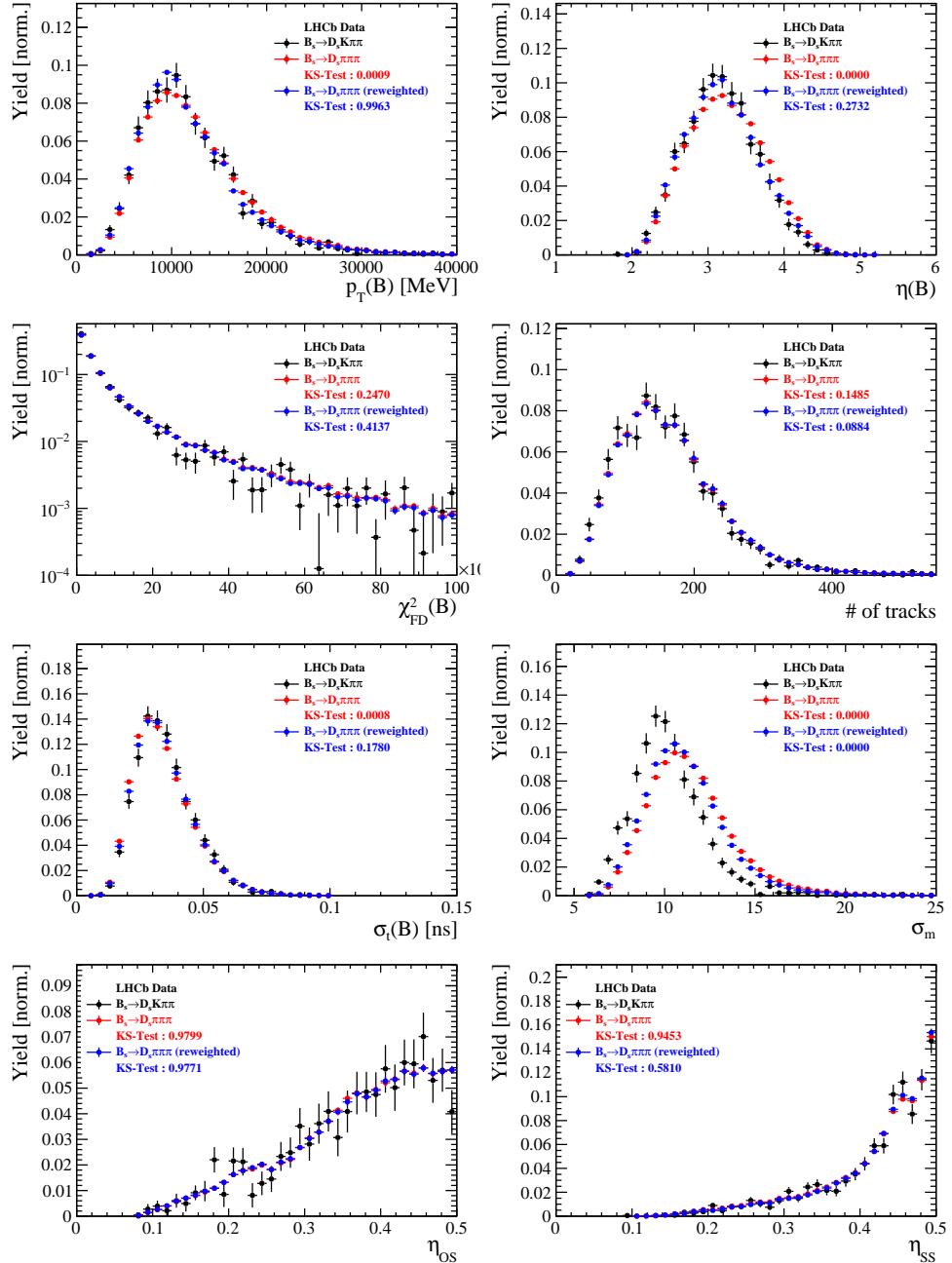


Figure C.1: Comparison of selected variables.

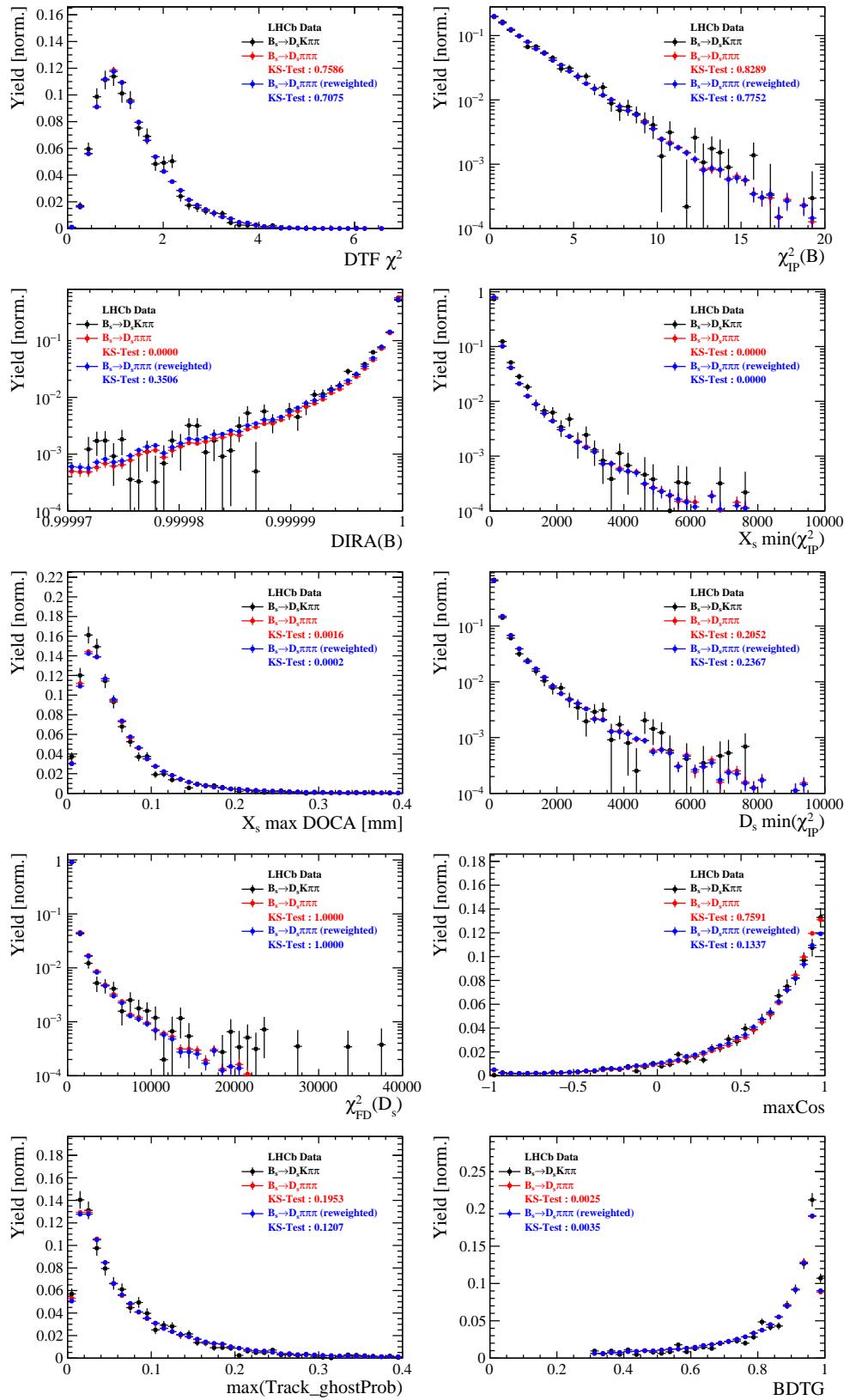


Figure C.2: Comparison of BDTG input variables and classifier response.

1067 H.2 Comparison of Run-I and Run-II data

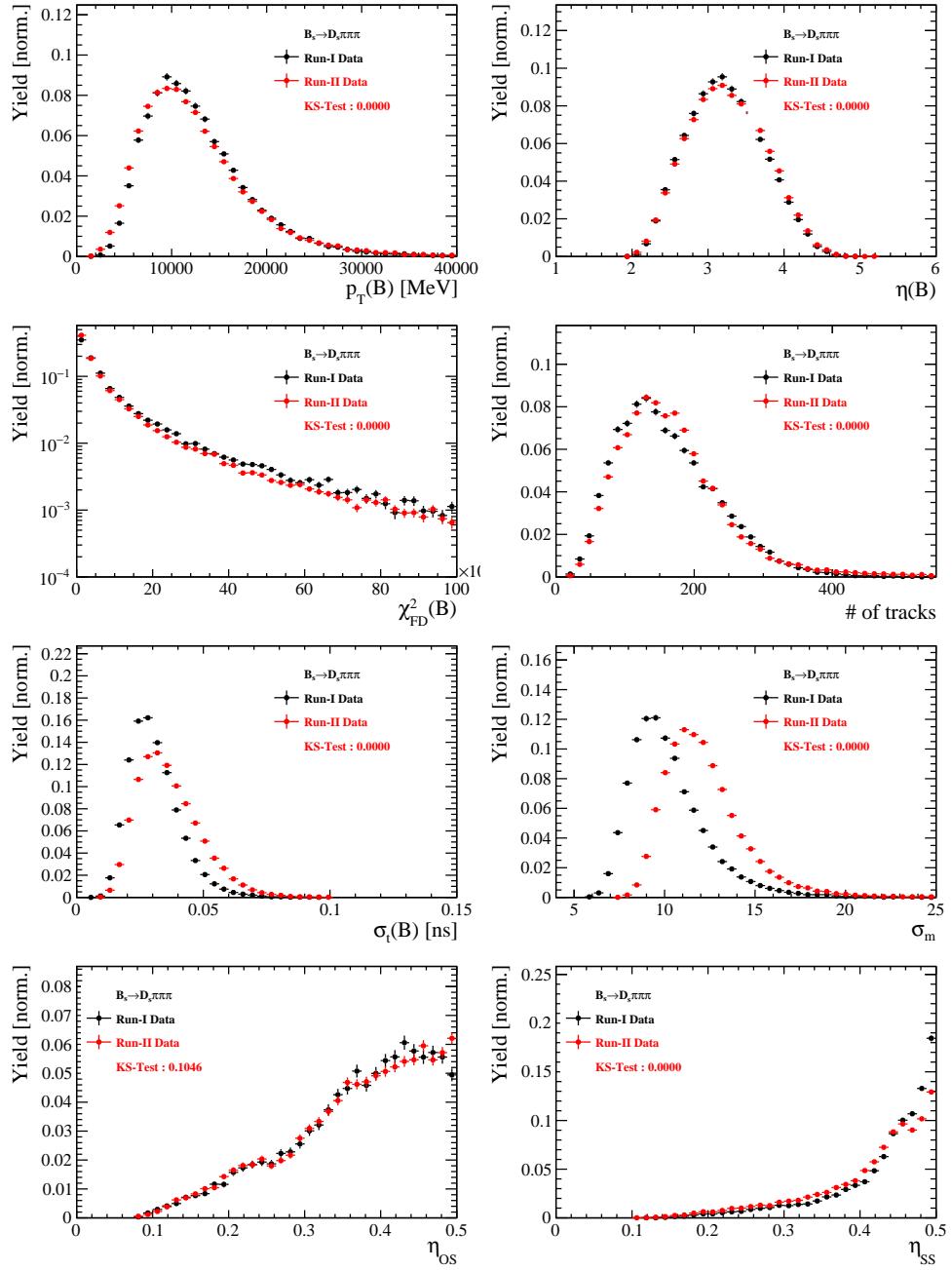


Figure C.3: Comparison of selected variables.

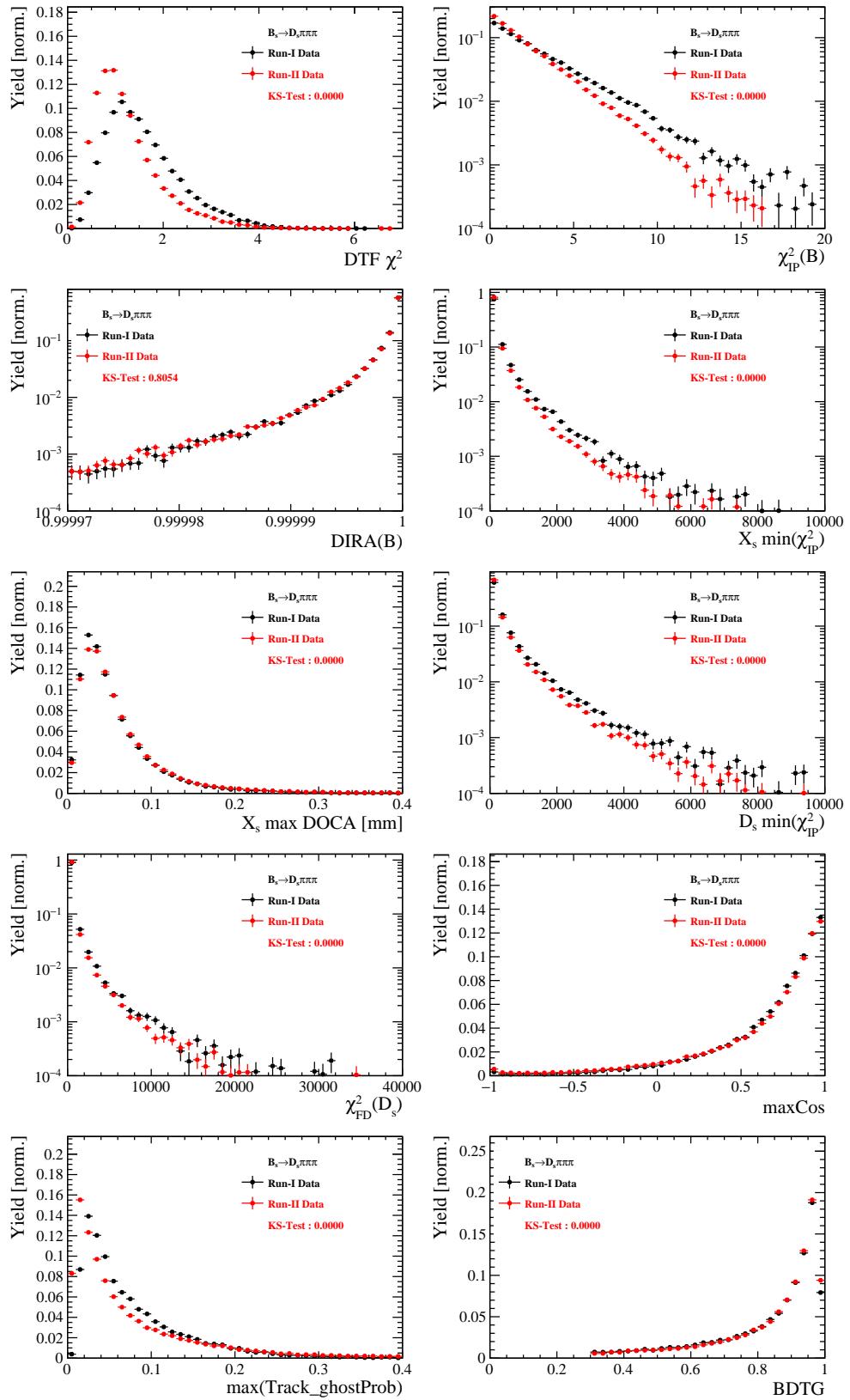


Figure C.4: Comparison of BDTG input variables and classifier response.

1068 H.3 Comparison of D_s final states

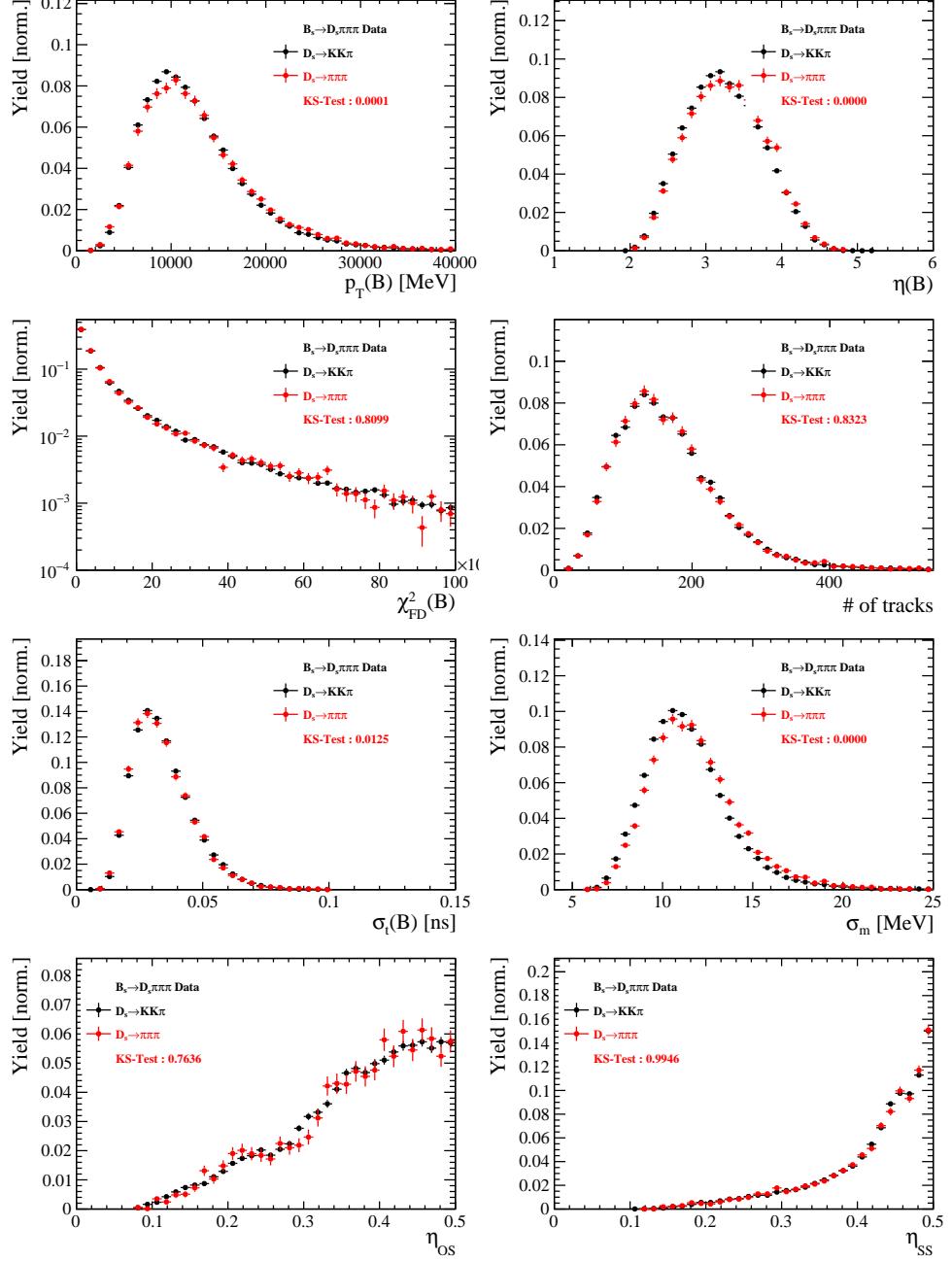


Figure C.5: Comparison of selected variables.

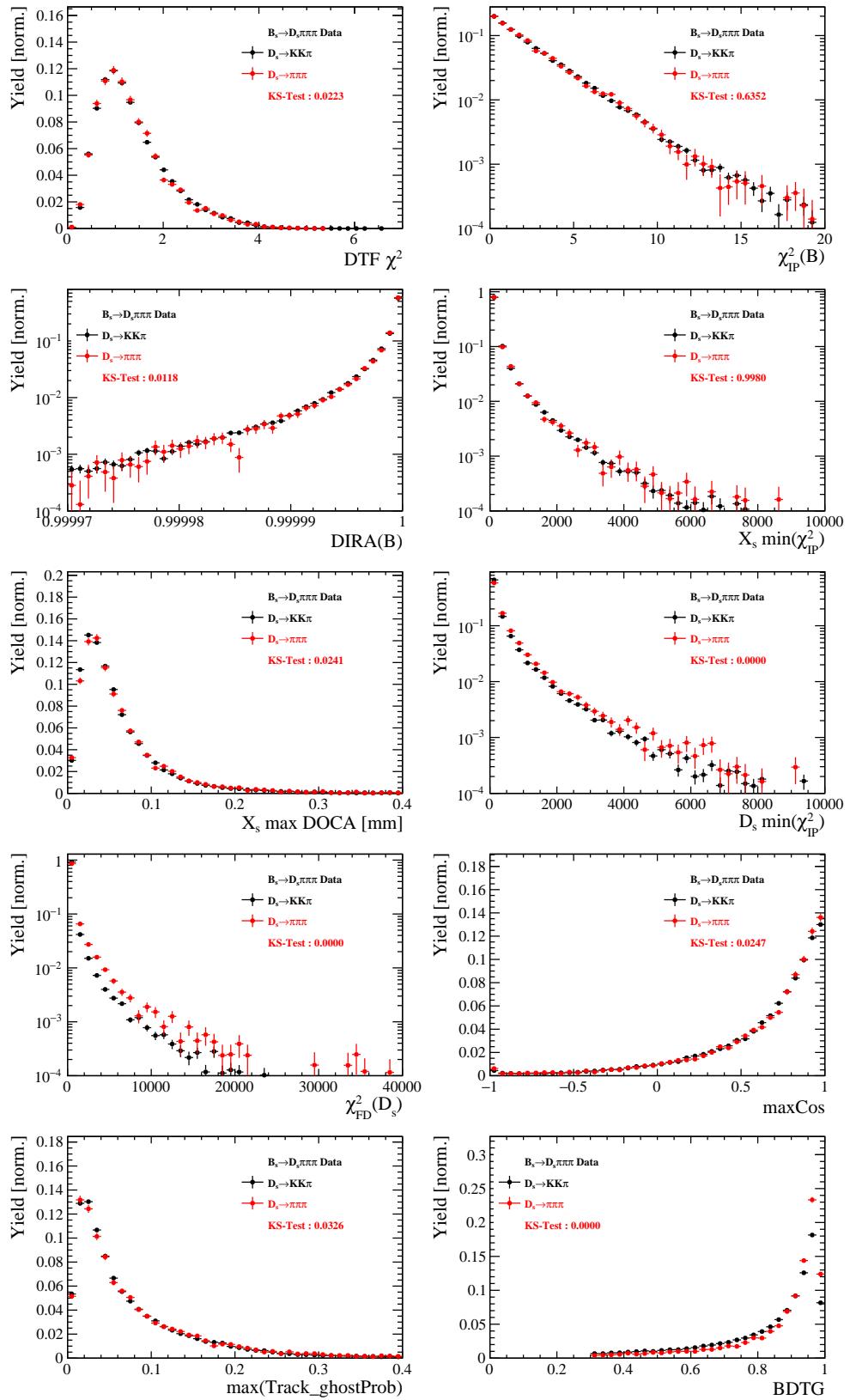


Figure C.6: Comparison of BDTG input variables and classifier response.

1069 H.4 Comparison of trigger categories

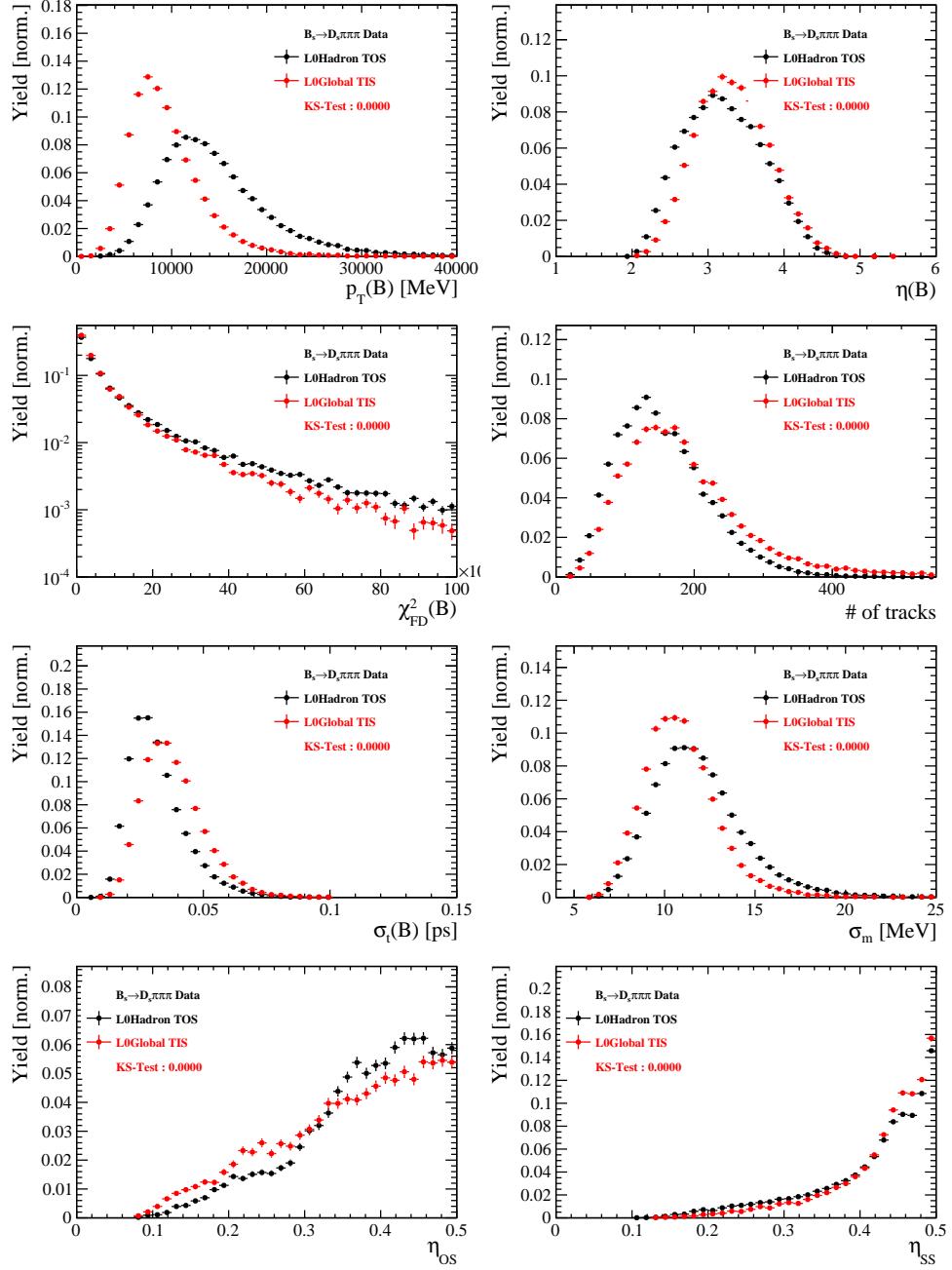


Figure C.7: Comparison of selected variables.

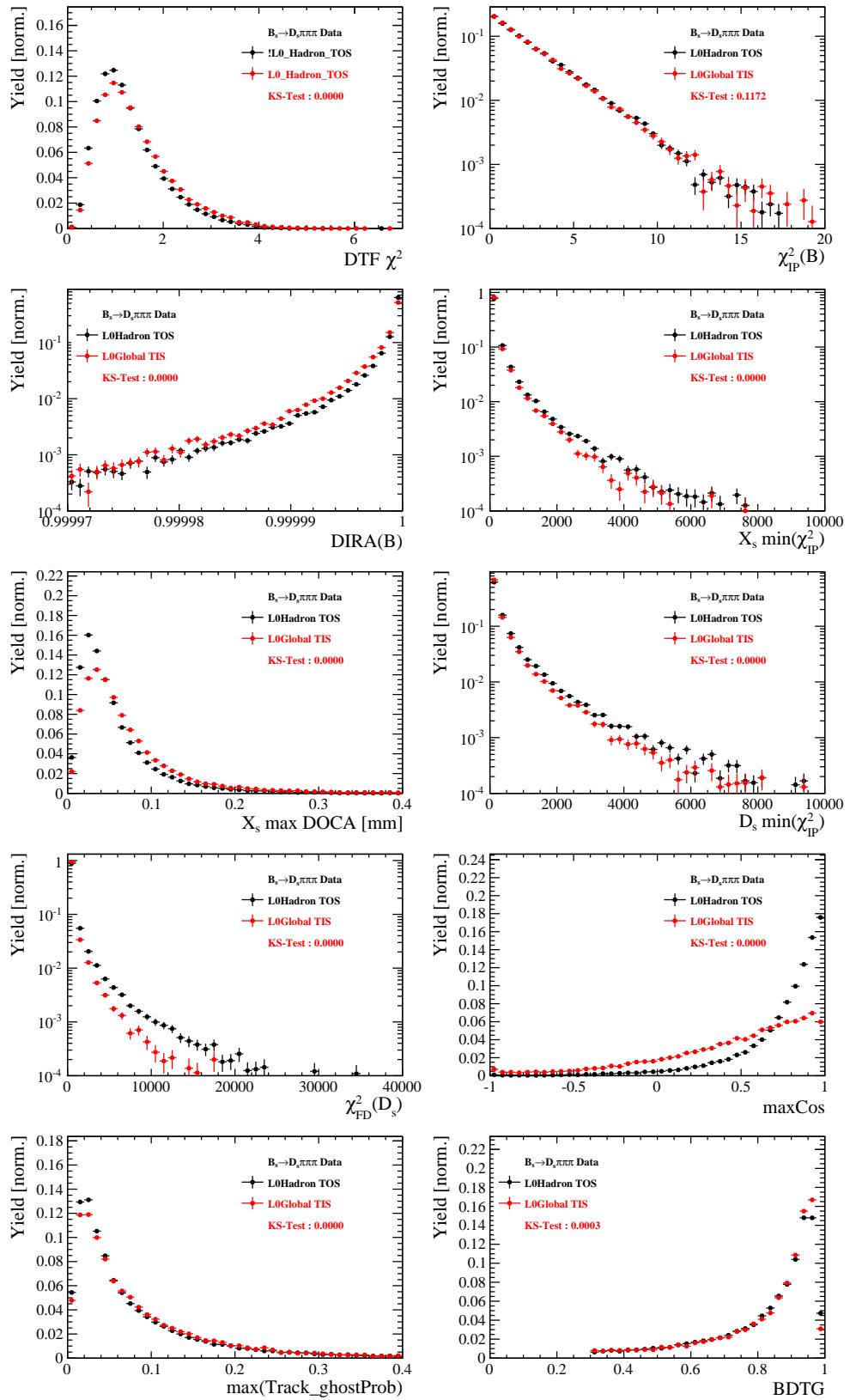


Figure C.8: Comparison of BDTG input variables and classifier response.

1070 H.5 Comparison of B_s and B_d decays

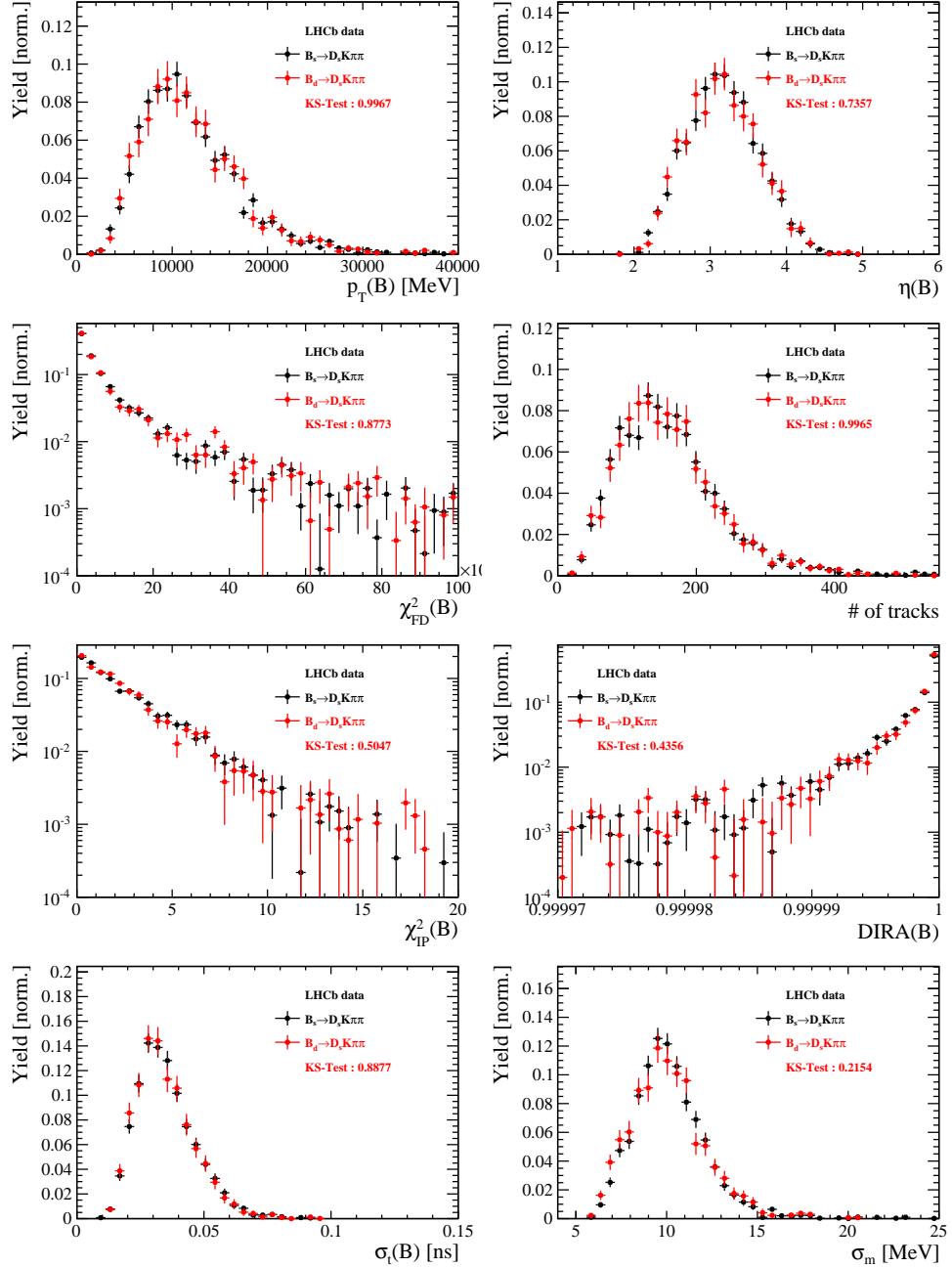


Figure C.9: Comparison of selected variables.

1071 References

- 1072 [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. Phys. **B671** (2003) 459, [arXiv:hep-ph/0304027](#).
- 1073 [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](#).
- 1074 [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 $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* .
- 1075 [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](#).
- 1076 [5] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 1077 [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 .
- 1078 [7] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*, Phys. Rev. D **11** (1975) 3165.
- 1079 [8] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of Physics **108** (1977), no. 2 454 .
- 1080 [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](#).
- 1081 [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.
- 1082 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo Cimento Series 10 **34** (1964), no. 6 1644.
- 1083 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.
- 1084 [13] D. V. Bugg, *The mass of the σ pole*, Journal of Physics G Nuclear Physics **34** (2007) 151, [arXiv:hep-ph/0608081](#).
- 1085 [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.
- 1086 [15] S. M. Flatté, *Coupled-channel analysis of the $\pi\eta$ and KK systems near KK threshold*, Physics Letters B **63** (1976), no. 2 224 .
- 1087 [16] 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](#).
- 1088 [17] W. Dunwoodie. Fits to $K\pi$ $I = \frac{1}{2}$ S -wave amplitude and phase data.

- 1105 [18] D. Aston *et al.*, *A Study of K- pi+ Scattering in the Reaction K- p —> K- pi+ n at*
 1106 *11-GeV/c*, Nucl. Phys. **B296** (1988) 493.
- 1107 [19] BaBar, B. Aubert *et al.*, *Dalitz-plot analysis of the decays $B^\pm \rightarrow K^\pm \pi^\mp \pi^\pm$* ,
 1108 Phys. Rev. **D72** (2005) 072003, arXiv:hep-ex/0507004, [Erratum: Phys.
 1109 Rev.D74,099903(2006)].]
- 1110 [20] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in $D^0 \rightarrow K^\mp \pi^\pm \pi^\pm \pi^\mp$ decays*,
 1111 Eur. Phys. J. **C78** (2018), no. 6 443, arXiv:1712.08609.
- 1112 [21] M. S. et al. *Search for CP violation in the $D^0 \rightarrow KK\pi\pi$ decay through a full*
 1113 *amplitude analysis*, LHCb-ANA-2017-064.
- 1114 [22] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- 1115 [23] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys.
 1116 Rev. **60** (1941) 61.
- 1117 [24] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys.
 1118 Rev. D **57** (1998) 431.
- 1119 [25] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of*
 1120 *ψ decay to mesons*, Eur. Phys. J. **A16** (2003) 537, arXiv:hep-ph/0211457.
- 1121 [26] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spec-*
 1122 *troscopy*, Phys. Rev. **D51** (1995) 2247.
- 1123 [27] J.-J. Zhu, *Explicit expressions of spin wave functions*, arXiv:hep-ph/9906250.
- 1124 [28] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput.
 1125 Phys. Commun. **180** (2009) 1847, arXiv:0805.2956.
- 1126 [29] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth.
 1127 **A462** (2001) 152.
- 1128 [30] M. Karbach and M. Kenzie, *Gammacombo package*,
 1129 <http://gammacombo.hepforge.org/web/HTML/index.html>, 2014.
- 1130 [31] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007)
 1131 040, arXiv:physics/0703039.
- 1132 [32] N. L. Johnson, *Systems of frequency curves generated by methods of translation*,
 1133 Biometrika **36** (1949), no. 1/2 149.
- 1134 [33] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38**
 1135 (2014) 090001.
- 1136 [34] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys.
 1137 **A30** (2015) 1530022, arXiv:1412.6352.
- 1138 [35] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*,
 1139 Submitted to: JHEP (2017) arXiv:1712.07428.

- 1140 [36] LHCb collaboration, L. Zhang, *Measurements of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$*
 1141 *decays in the low $K^+ K^-$ mass range with 13 TeV data*, LHCb-ANA-2017-028.
- 1142 [37] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and*
 1143 *τ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots
 1144 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 1145 [38] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*
 1146 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 1147 [39] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at*
 1148 *LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- 1149 [40] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the*
 1150 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- 1151 [41] LHCb, R. Aaij *et al.*, *Measurement of B^0 , B_s^0 , B^+ and Λ_b^0 production asymmetries in 7*
 1152 *and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- 1153 [42] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of*
 1154 *the $K\pi$ Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-
 1155 2012-027, CERN, Geneva, Feb, 2013.
- 1156 [43] A. Davis *et al.*, *Measurement of the instrumental asymmetry for $K^- \pi^+$ -pairs at LHCb*
 1157 *in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN,
 1158 Geneva, Mar, 2018.
- 1159 [44] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly*
 1160 *forbidden transitions in $D \rightarrow K(S)$, $K(L) + \pi$'s decays*, Phys. Lett. **B349** (1995)
 1161 363, arXiv:hep-ph/9502238.
- 1162 [45] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*,
 1163 Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- 1164 [46] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude*
 1165 *analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- 1166 [47] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal
 1167 Statistical Society, Series B **58** (1994) 267.
- 1168 [48] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- 1169 [49] H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on
 1170 Automatic Control **19** (1974) 716.
- 1171 [50] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime*
 1172 *and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986,
 1173 DESY-F31-86-02.
- 1174 [51] D. Hill, M. John, and P. Gandini, *A study of partially reconstructed $B^\pm \rightarrow D^{*0} h^\pm$*
 1175 *decays using the $D^0 \rightarrow K\pi, KK, \pi\pi$ final states*, .

- ₁₁₇₆ [52] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins
₁₁₇₇ University Press, Baltimore, MD, USA, 1996.