



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

P. d'Argent<sup>1</sup>, E. Gersabeck<sup>2</sup>, M. Kecke<sup>1</sup>, M. Schiller<sup>3</sup>

<sup>1</sup>*Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany*

<sup>2</sup>*School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom*

<sup>3</sup>*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  $5 \text{ fb}^{-1}$  recorded by the LHCb detector.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Formalism</b>	<b>1</b>
2.1	Decay rates and CP-observables . . . . .	1
2.2	Amplitude model . . . . .	2
2.2.1	Form Factors and Resonance Lineshapes . . . . .	3
2.2.2	Spin Densities . . . . .	5
2.3	Validation . . . . .	7
<b>3</b>	<b>Data samples and event selection</b>	<b>11</b>
3.1	Stripping and Trigger selection . . . . .	11
3.2	Offline selection . . . . .	11
3.2.1	Phase space region . . . . .	12
3.2.2	Physics background vetoes . . . . .	13
3.2.3	Training of multivariate classifier . . . . .	16
3.2.4	Final selection . . . . .	18
<b>4</b>	<b>Yields determination</b>	<b>21</b>
4.1	Signal model . . . . .	21
4.2	Background models . . . . .	22
4.3	Results . . . . .	23
<b>5</b>	<b>Decay-time Resolution</b>	<b>25</b>
5.1	Calibration for Run-I data . . . . .	26
5.2	Calibration for Run-II data . . . . .	27
5.3	Cross-checks . . . . .	30
5.3.1	Kinematic dependence . . . . .	30
5.3.2	DTF constraints . . . . .	30
<b>6</b>	<b>Acceptance</b>	<b>31</b>
6.1	MC corrections . . . . .	31
6.1.1	Truth matching of simulated candidates . . . . .	31
6.1.2	PID efficiencies . . . . .	32
6.1.3	BDT efficiencies . . . . .	33
6.1.4	Tracking efficiencies . . . . .	34
6.2	Decay-time acceptance . . . . .	35
6.2.1	Comparison of acceptance in subsamples . . . . .	36
6.2.2	Results . . . . .	38
6.3	Phasespace acceptance . . . . .	42
<b>7</b>	<b>Flavour Tagging</b>	<b>43</b>
7.1	OS tagging calibration . . . . .	44
7.2	SS tagging calibration . . . . .	44
7.3	Tagging performance comparison between the signal and normalization channel . . . . .	44
7.4	Combination of OS and SS taggers . . . . .	45

<b>8 Production and Detection Asymmetries</b>	<b>47</b>
8.1 $B_s$ Production Asymmetry . . . . .	47
8.2 $K^-\pi^+$ Detection Asymmetry . . . . .	48
<b>9 Time dependent fit</b>	<b>51</b>
9.1 sFit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data . . . . .	51
9.2 sFit to $B_s^0 \rightarrow D_sK\pi\pi$ data . . . . .	54
<b>10 Time dependent amplitude fit</b>	<b>55</b>
10.1 Signal Model Construction . . . . .	55
10.2 Results . . . . .	56
<b>11 Systematic uncertainties</b>	<b>59</b>
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 Systematic effect on observables . . . . .	60
<b>A Stripping and Trigger cuts</b>	<b>61</b>
<b>B Details of multivariate classifier</b>	<b>63</b>
<b>C Detailed mass fits</b>	<b>67</b>
<b>D Decay-time Resolution fits</b>	<b>71</b>
<b>E Spin Amplitudes</b>	<b>75</b>
<b>F Considered Decay Chains</b>	<b>76</b>
<b>G MC corrections</b>	<b>77</b>
<b>H Data distributions</b>	<b>82</b>
H.1 Comparison of signal and calibration channel . . . . .	82
H.2 Comparison of Run-I and Run-II data . . . . .	84
H.3 Comparison of $D_s$ final states . . . . .	86
H.4 Comparison of trigger categories . . . . .	88
H.5 Comparison of $B_s$ and $B_d$ decays . . . . .	90
<b>References</b>	<b>91</b>
<b>References</b>	<b>91</b>

# 1 Introduction

The weak phase  $\gamma$  is the least well known angle of the CKM unitary triangle. A key channel to measure  $\gamma$  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  $\gamma$  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  $B_s^0 \rightarrow D_s \pi\pi\pi$ .

## 2 Formalism

### 2.1 Decay rates and CP-observables

In the following, we choose a convention in which  $\Delta\Gamma_s = \Gamma_L - \Gamma_H < 0$  and  $\Delta m_s = m_H - m_L > 0$ , where the indices  $H$  and  $L$  refer to the heavy and light mass eigenstates of the  $B_s$  meson. We assume  $|q/p| = 1$  for the complex coefficients  $p$  and  $q$  which relate the  $B_s$  meson mass eigenstates to the flavour eigenstates.

$$A(B_s^0 \rightarrow D_s^- K^+ \pi\pi) \equiv A(x) = \sum_i a_i A_i(x) \quad (2.1)$$

$$A(B_s^0 \rightarrow D_s^+ K^- \pi\pi) \equiv \bar{A}(\bar{x}) = \sum_i \bar{a}_i \bar{A}_i(\bar{x}) \quad (2.2)$$

$$A(\bar{B}_s^0 \rightarrow D_s^- K^+ \pi\pi) = \bar{A}(x) \text{ (Assuming no direct CPV)} \quad (2.3)$$

$$A(\bar{B}_s^0 \rightarrow D_s^+ K^- \pi\pi) = A(\bar{x}) \text{ (Assuming no direct CPV)} \quad (2.4)$$

The full time-dependent amplitude pdf is given by:

$$\begin{aligned} P(x, t, q_t, q_f) \propto & [(|A(x)|^2 + |\bar{A}(x)|^2) \cosh\left(\frac{\Delta\Gamma t}{2}\right) \\ & + q_t q_f (|A(x)|^2 - |\bar{A}(x)|^2) \cos(m_s t) \\ & - 2\text{Re}(A(x)^* \bar{A}(x) e^{-iq_f(\gamma-2\beta_s)}) \sinh\left(\frac{\Delta\Gamma t}{2}\right) \\ & - 2q_t q_f \text{Im}(A(x)^* \bar{A}(x) e^{-iq_f(\gamma-2\beta_s)}) \sin(m_s t)] e^{-\Gamma t} \end{aligned} \quad (2.5)$$

where  $q_t = +1, -1, 0$  for events tagged as  $B_s^0$ ,  $\bar{B}_s^0$  or untagged events and  $q_f = +1$  (-1) for  $D_s^- K^+ \pi\pi$  ( $D_s^+ K^- \pi\pi$ ) final states. Integrating over the phasespace, we get

$$\int P(x, t, q_t, q_f) dx \propto [\cosh\left(\frac{\Delta\Gamma t}{2}\right) + q_t q_f C \cos(m_s t) + \kappa D_{q_f} \sinh\left(\frac{\Delta\Gamma t}{2}\right) - q_t \kappa S_{q_f} \sin(m_s t)] e^{-\Gamma t} \quad (2.6)$$

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

$$C = \frac{1 - r^2}{1 + r^2} \quad (2.7)$$

$$D_{q_f} = -\frac{2r \cos(\delta - q_f (\gamma - 2\beta_s))}{1 + r^2} \quad (2.8)$$

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

22 The coherence factor  $\kappa$ , the strong phase difference  $\delta$  and the ratio of the suppressed  
23 ( $b \rightarrow u$ ) over favored ( $b \rightarrow c$ ) decay mode are defined as:

$$\kappa e^{i\delta} \equiv \frac{\int A(x)^* \bar{A}(x) dx}{\sqrt{\int |A(x)|^2 dx} \sqrt{\int |\bar{A}(x)|^2 dx}} \quad (2.10)$$

$$r \equiv \frac{\sqrt{\int |\bar{A}(x)|^2 dx}}{\sqrt{\int |A(x)|^2 dx}}. \quad (2.11)$$

24 In the limit of only one contributing resonance  $\kappa \rightarrow 1$ .

25

## 26 2.2 Amplitude model

27 The differential decay rate of a  $B_s$  meson with mass,  $m_{B_s}$ , decaying into four pseudoscalar  
28 particles with four-momenta  $p_i = (E_i, \vec{p}_i)$  ( $i = 1, 2, 3, 4$ ) is given by

$$d\Gamma = \frac{1}{2 m_{B_s}} |A(\mathbf{x})|^2 d\Phi_4, \quad (2.12)$$

29 where the transition amplitude  $A(\mathbf{x})$ , describes the dynamics of the interaction,  $d\Phi_4$   
30 is the four-body phase space element [5], and  $\mathbf{x}$  represents a unique set of kinematic  
31 conditions within the phase space of the decay. Each final state particle contributes three  
32 observables, manifesting in their three-momentum, summing up to twelve observables in  
33 total. Four of them are redundant due to four-momentum conservation and the overall  
34 orientation of the system can be integrated out. The remaining five independent degrees  
35 of freedom unambiguously determine the kinematics of the decay. Convenient choices  
36 for the kinematic observables include the invariant mass combinations of the final state  
37 particles

$$\begin{aligned} m_{ij}^2 &= (p_i + p_j)^2, \\ m_{ijk}^2 &= (p_i + p_j + p_k)^2 \end{aligned} \quad (2.13)$$

38 or acoplanarity and helicity angles. It is however important to take into account that,  
39 while  $m_{12}^2, m_{23}^2$  are sufficient to fully describe a three-body decay, the obvious extension  
40 to four-body decays with  $m_{ij}^2, m_{ijk}^2$  requires additional care, as these variables alone are  
41 insufficient to describe the parity-odd moments possible in four-body kinematics.

42 In practice, we do not need to choose a particular five-dimensional basis, but use the  
43 full four-vectors of the decay in our analysis. The dimensionality is handled by the phase

44 space element which can be written in terms of any set of five independent kinematic  
 45 observables,  $\mathbf{x} = (x_1, \dots, x_5)$ , as

$$d\Phi_4 = \phi_4(\mathbf{x}) d^5x, \quad (2.14)$$

46 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,  
 47 the four-body phase space density function is not flat in the usual kinematic variables.  
 48 Therefore, an analytic expression for  $\phi_4$  is taken from Ref. [6].

49 The total amplitude for the  $B_s \rightarrow h_1 h_2 h_3 h_4$  decay is given by the coherent sum  
 50 over all intermediate state amplitudes  $A_i(\mathbf{x})$ , each weighted by a complex coefficient  
 51  $a_i = |a_i| e^{i\phi_i}$  to be measured from data,

$$A(\mathbf{x}) = \sum_i a_i A_i(\mathbf{x}). \quad (2.15)$$

52 To construct  $A_i(\mathbf{x})$ , the isobar approach is used, which assumes that the decay process  
 53 can be factorized into subsequent two-body decay amplitudes [7–9]. This gives rise to  
 54 two different decay topologies; quasi two-body decays  $B_s \rightarrow (R_1 \rightarrow h_1 h_2) (R_2 \rightarrow h_3 h_4)$   
 55 or cascade decays  $B_s \rightarrow h_1 [R_1 \rightarrow h_2 (R_2 \rightarrow h_3 h_4)]$ . In either case, the intermediate state  
 56 amplitude is parameterized as a product of form factors  $B_L$ , included for each vertex  
 57 of the decay tree, Breit-Wigner propagators  $T_R$ , included for each resonance  $R$ , and an  
 58 overall angular distribution represented by a spin factor  $S$ ,

$$A_i(\mathbf{x}) = B_{L_{B_s}}(\mathbf{x}) [B_{L_{R_1}}(\mathbf{x}) T_{R_1}(\mathbf{x})] [B_{L_{R_2}}(\mathbf{x}) T_{R_2}(\mathbf{x})] S_i(\mathbf{x}). \quad (2.16)$$

### 59 2.2.1 Form Factors and Resonance Lineshapes

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

$$\begin{aligned} B_0(q) &= 1, \\ B_1(q) &= 1/\sqrt{1 + (q r_{BW})^2}, \\ B_2(q) &= 1/\sqrt{9 + 3(q r_{BW})^2 + (q r_{BW})^4}. \end{aligned} \quad (2.17)$$

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

$$T(s) = \frac{1}{M^2(s) - s - i m_0 \Gamma(s)}, \quad (2.18)$$

67 featuring the energy-dependent mass  $M(s)$  (defined below), and total width,  $\Gamma(s)$ . The  
 68 latter is normalized to give the nominal width,  $\Gamma_0$ , when evaluated at the nominal mass  
 69  $m_0$ , *i.e.*  $\Gamma_0 = \Gamma(s = m_0^2)$ .

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

$$\Gamma_{R \rightarrow AB}^{(2)}(s) = \Gamma_0 \frac{m_0}{\sqrt{s}} \left( \frac{q}{q_0} \right)^{2L+1} \frac{B_L(q)^2}{B_L(q_0)^2}, \quad (2.19)$$

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

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

$$\Gamma_{R \rightarrow ABC}^{(3)}(s) = \frac{1}{2\sqrt{s}} \int |A_{R \rightarrow ABC}|^2 d\Phi_3, \quad (2.20)$$

77 and therefore requires knowledge of the resonant substructure. The three-body amplitude  
 78  $A_{R \rightarrow ABC}$  can be parameterized similarly to the four-body amplitude in Eq. (2.16). In  
 79 particular, it includes form factors and propagators of intermediate two-body resonances.

80 Both Eq. (2.19) and Eq. (2.20) give only the partial width for the decay into a specific  
 81 channel. To obtain the total width, a sum over all possible decay channels has to be  
 82 performed,

$$\Gamma(s) = \sum_i g_i \Gamma_i(s), \quad (2.21)$$

83 where the coupling strength to channel  $i$ , is given by  $g_i$ . Branching fractions  $\mathcal{B}_i$  are related  
 84 to the couplings  $g_i$  via the equation [12]

$$\mathcal{B}_i = \int_{s_{min}}^{\infty} \frac{g_i m_0 \Gamma_i(s)}{|M^2(s) - s - i m_0 \sum_j g_j \Gamma_j(s)|^2} ds. \quad (2.22)$$

85 As experimental values are usually only available for the branching fractions, Eq. (2.22)  
 86 needs to be inverted to obtain values for the couplings. In practice, this is solved by  
 87 minimizing the quantity  $\chi^2(g) = \sum_i [\mathcal{B}_i - \mathcal{I}_i(g)]^2 / \Delta \mathcal{B}_i^2$ , where  $\mathcal{I}_i(g)$  denotes the right-  
 88 hand side of Eq. (2.22).

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

92 For the broad scalar resonance  $\sigma$ , the model from Bugg is used [13]. Besides  $\sigma \rightarrow \pi\pi$   
 93 decays, it includes contributions from the decay modes  $\sigma \rightarrow KK$ ,  $\sigma \rightarrow \eta\eta$  and  $\sigma \rightarrow \pi\pi\pi\pi$   
 94 as well as dispersive effects due to the channel opening of the latter. We use the Gournaris-  
 95 Sakurai parametrization for the  $\rho(770)^0 \rightarrow \pi\pi$  propagator which provides an analytical  
 96 description of the dispersive term,  $M^2(s)$  [14]. The energy-dependent width of the  $f_0(980)$   
 97 resonance is given by the sum of the partial widths into the  $\pi\pi$  and  $KK$  channels [15],

$$\Gamma_{f_0(980)}(s) = g_{\pi\pi} \Gamma_{f_0(980) \rightarrow \pi\pi}^{(2)}(s) + g_{KK} \Gamma_{f_0(980) \rightarrow KK}^{(2)}(s), \quad (2.23)$$

98 where the coupling constants  $g_{\pi\pi}$  and  $g_{KK}$ , as well as the mass and width are taken from  
 99 a measurement performed by the BES Collaboration [16]. The total decay widths for  
 100 both the  $f_2(1270)$  and the  $f_0(1370)$  meson take the channels  $\pi\pi$ ,  $KK$ ,  $\eta\eta$  and  $\pi\pi\pi\pi$  into  
 101 account. While the two-body partial widths are described by Eq. (2.19), a model for  
 102 the partial width for a decay into four pions is taken from Ref. [17]. The corresponding  
 103 branching fractions are taken from the PDG [12]. The nominal mass and width of the  
 104  $f_0(1370)$  resonance are taken from an LHCb measurement [18]. Equation (2.19) is used  
 105 for all other resonances decaying into a two-body final state.

106 Some particles may not originate from a resonance but are in a state of relative orbital  
 107 angular momentum. We denote such non-resonant states by surrounding the particle

108 system with brackets and indicate the partial wave state with an subscript; for example  
109  $(\pi\pi)_S$  refers to a non-resonant di-pion  $S$ -wave. The lineshape for non-resonant states is  
110 set to unity.

### 111 2.2.2 Spin Densities

112 The spin amplitudes are phenomenological descriptions of decay processes that are required  
113 to be Lorentz invariant, compatible with angular momentum conservation and, where  
114 appropriate, parity conservation. They are constructed in the covariant Zemach (Rarita-  
115 Schwinger) tensor formalism [19–21]. At this point, we briefly introduce the fundamental  
116 objects of the covariant tensor formalism which connect the particle’s four-momenta to  
117 the spin dynamics of the reaction and give a general recipe to calculate the spin factors  
118 for arbitrary decay trees. Further details can be found in Refs. [22, 23].

119 A spin- $S$  particle with four-momentum  $p$ , and spin projection  $\lambda$ , is represented by the  
120 polarization tensor  $\epsilon_{(S)}(p, \lambda)$ , which is symmetric, traceless and orthogonal to  $p$ . These  
121 so-called Rarita-Schwinger conditions reduce the a priori  $4^S$  elements of the rank- $S$  tensor  
122 to  $2S + 1$  independent elements in accordance with the number of degrees of freedom of a  
123 spin- $S$  state [20, 24].

124 The spin projection operator  $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$ , for a resonance  $R$ , with spin  $S =$   
125  $\{0, 1, 2\}$ , and four-momentum  $p_R$ , is given by [23]:

$$\begin{aligned} P_{(0)}^{\mu\nu}(p_R) &= 1 \\ P_{(1)}^{\mu\nu}(p_R) &= -g^{\mu\nu} + \frac{p_R^\mu p_R^\nu}{p_R^2} \\ P_{(2)}^{\mu\nu\alpha\beta}(p_R) &= \frac{1}{2} \left[ P_{(1)}^{\mu\alpha}(p_R) P_{(1)}^{\nu\beta}(p_R) + P_{(1)}^{\mu\beta}(p_R) P_{(1)}^{\nu\alpha}(p_R) \right] - \frac{1}{3} P_{(1)}^{\mu\nu}(p_R) P_{(1)}^{\alpha\beta}(p_R), \end{aligned} \quad (2.24)$$

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

128 For a decay process  $R \rightarrow AB$ , with relative orbital angular momentum  $L$ , between  
129 particle  $A$  and  $B$ , the angular momentum tensor is obtained by projecting the rank- $L$   
130 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- $L$   
131 subspace,

$$L_{(L)\mu_1 \dots \mu_L}(p_R, q_R) = (-1)^L P_{(L)\mu_1 \dots \mu_L \nu_1 \dots \nu_L}(p_R) q_R^{\nu_1} \dots q_R^{\nu_L}. \quad (2.25)$$

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

$$\begin{aligned} \varepsilon_{(S)}(R) &\equiv \varepsilon_{(S)}(p_R, \lambda_R), \\ P_{(S)}(R) &\equiv P_{(S)}(p_R), \\ L_{(L)}(R) &\equiv L_{(L)}(p_R, q_R). \end{aligned} \quad (2.26)$$

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

$$\begin{aligned} S_{R \rightarrow AB}(\mathbf{x}|L_{AB}, S_{AB}; \lambda_R, \lambda_A, \lambda_B) &= \varepsilon_{(S_R)}(R) X(S_R, L_{AB}, S_{AB}) L_{(L_{AB})}(R) \\ &\times \Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B), \end{aligned} \quad (2.27)$$

<sup>138</sup> where

$$\Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B) = P_{(S_{AB})}(R) X(S_{AB}, S_A, S_B) \varepsilon_{(S_A)}^*(A) \varepsilon_{(S_B)}^*(B). \quad (2.28)$$

<sup>139</sup> Here, a polarization vector is assigned to the decaying particle and the complex conjugate  
<sup>140</sup> vectors for each decay product. The spin and orbital angular momentum couplings are  
<sup>141</sup> described by the tensors  $P_{(S_{AB})}(R)$  and  $L_{(L_{AB})}(R)$ , respectively. Firstly, the two spins  $S_A$   
<sup>142</sup> and  $S_B$ , are coupled to a total spin- $S_{AB}$  state,  $\Phi(\mathbf{x}|S_{AB})$ , by projecting the corresponding  
<sup>143</sup> polarization vectors onto the spin- $S_{AB}$  subspace transverse to the momentum of the  
<sup>144</sup> decaying particle. Afterwards, the spin and orbital angular momentum tensors are  
<sup>145</sup> properly contracted with the polarization vector of the decaying particle to give a Lorentz  
<sup>146</sup> scalar. This requires in some cases to include the tensor  $\varepsilon_{\alpha\beta\gamma\delta} p_R^\delta$  via

$$X(j_a, j_b, j_c) = \begin{cases} 1 & \text{if } j_a + j_b + j_c \text{ even} \\ \varepsilon_{\alpha\beta\gamma\delta} p_R^\delta & \text{if } j_a + j_b + j_c \text{ odd} \end{cases}, \quad (2.29)$$

<sup>147</sup> where  $\varepsilon_{\alpha\beta\gamma\delta}$  is the Levi-Civita symbol and  $j$  refers to the arguments of  $X$  defined in  
<sup>148</sup> Eqs. 2.27 and 2.28. Its antisymmetric nature ensures the correct parity transformation  
<sup>149</sup> behavior of the amplitude. The spin factor for a whole decay chain, for example  $R \rightarrow$   
<sup>150</sup>  $(R_1 \rightarrow AB)(R_2 \rightarrow CD)$ , is obtained by combining the two-body terms and performing a  
<sup>151</sup> sum over all unobservable, intermediary spin projections

$$\sum_{\lambda_{R_1}, \lambda_{R_2}} S_{R \rightarrow R_1 R_2}(\mathbf{x}|L_{R_1 R_2}; \lambda_{R_1}, \lambda_{R_2}) S_{R_1 \rightarrow AB}(\mathbf{x}|L_{AB}; \lambda_{R_1}) S_{R_2 \rightarrow CD}(\mathbf{x}|L_{CD}; \lambda_{R_2}), \quad (2.30)$$

<sup>152</sup> 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  
<sup>153</sup> pseudoscalar initial/final states are involved.

<sup>154</sup> The spin factors for all decay topologies considered in this analysis are explicitly given  
<sup>155</sup> in Appendix E.

## 156 2.3 Validation

157 The amplitude fit is performed by using the amplitude analysis tool **MINT2**. It was  
 158 previously applied to analyze  $D^0 \rightarrow 4\pi$  and  $D^0 \rightarrow KK\pi\pi$  decays [25] which have an  
 159 identical general spin structure (*i.e.* scalar to four scalar decay) then  $B_s \rightarrow D_s K\pi\pi$  decays.  
 160 In the course of the  $D^0 \rightarrow hhhh$  analysis, the implementation of the amplitudes were  
 161 extensively cross-checked against other available tool such as **qft++** [26], **AmpGen** [27] and  
 162 were possible **EVTGEN** [28]. Since no additional line shapes or spin factors are needed for  
 163 this analysis, we consider the amplitude calculation as fully validated.

164 This does, however, not apply to the full time-dependent amplitude pdf which is newly  
 165 implemented for this analysis. To cross-check it, we use **EVTGEN** to generate toy events  
 166 with time-dependent CP violation according to the **SSD\_Cp** event model [28]. Since this  
 167 event model does not allow for multiple interfering resonances, we generate only the decay  
 168 chain  $B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi)$ . Table 2.1 lists the generated input parameters.  
 169 The toy data set is fitted with our **MINT2** implementation of the full time-dependent  
 170 amplitude pdf and the phasespace-integrated pdf.

171 The CP coefficients  $C, D, \bar{D}, S, \bar{S}$  are the fit parameters in case of the phasespace-  
 172 integrated pdf, while the full pdf determines  $x_{\pm} = r \cos(\delta \pm (\gamma - 2\beta_s))$  and  $y_{\pm} = r \sin(\delta \pm$   
 173  $(\gamma - 2\beta_s))$ . The fit parameters are converted to the physical observables  $r, \kappa, \delta, \gamma$  using  
 174 the **GammaCombo** package [29]. As shown in Tab. 2.2, 2.3 and 2.4, the fit results are  
 175 in excellent agreement with the generated input values. The 1-CL contours are shown  
 176 in Figs. 2.1 and 2.2. The phasespace-integrated fit is, in addition, performed with the  
 177 **B2DX** fitter used for the  $B_s \rightarrow D_s K$  analysis yielding identical results. Note though that  
 178 some parts of the **B2DX** fitter have been taken over to our **MINT2** fitter, such that the  
 179 implementations are not fully independent.

Table 2.1: Input values used to generate **EVTGEN** toy events according to the **SSD\_Cp** event model.

$\tau$	1.5 ps
$\Delta\Gamma$	$-0.1 \text{ ps}^{-1}$
$\Delta m_s$	$17.757 \text{ ps}^{-1}$
$r$	0.37
$\kappa$	1
$\delta$	$10.0^\circ$
$\gamma$	$71.1^\circ$
$\beta_s$	$0.0^\circ$

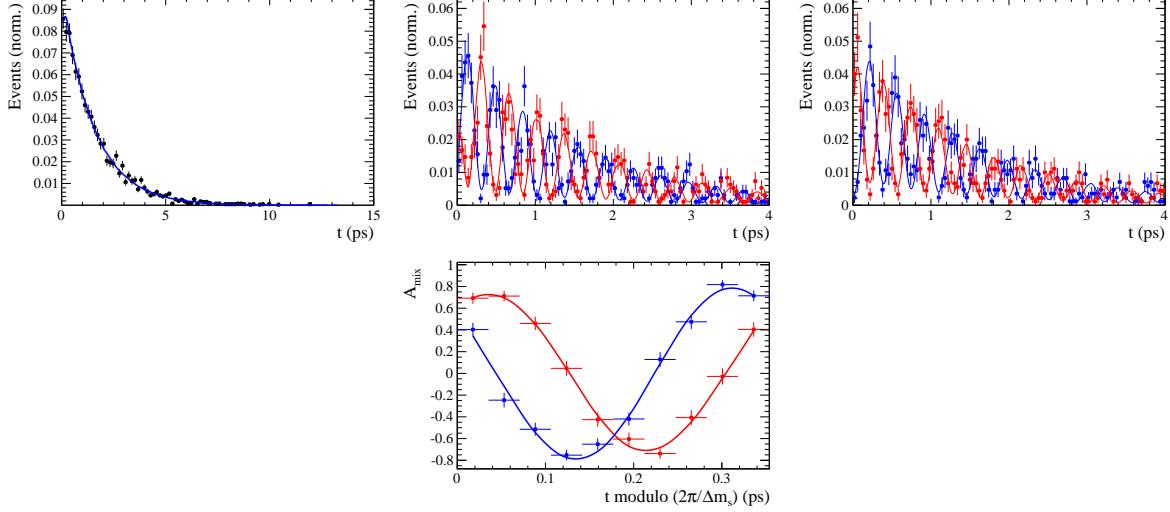


Figure 2.1: Time distribution of  $B_s \rightarrow D_s K \pi\pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

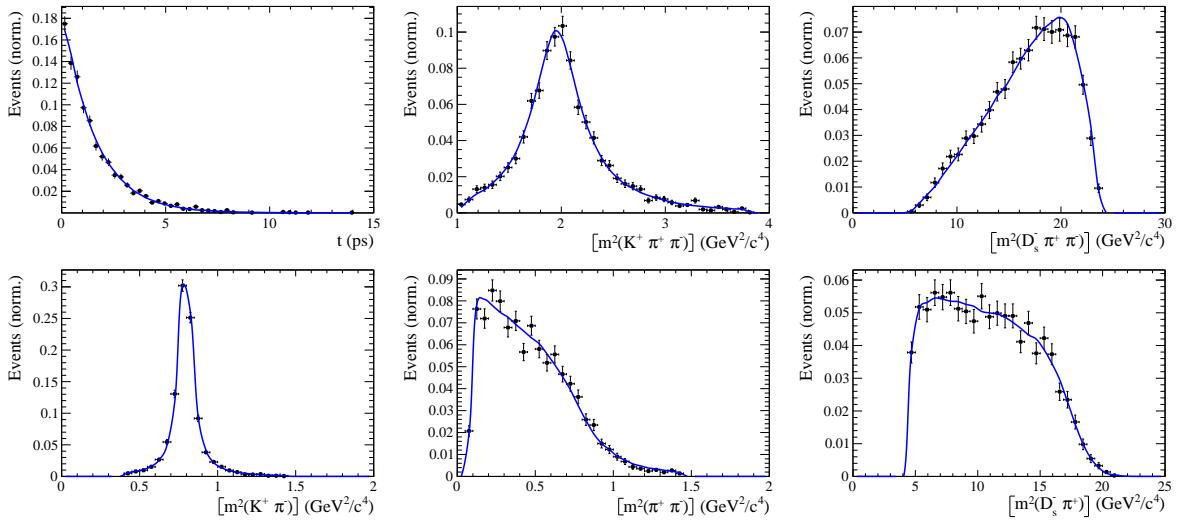


Figure 2.2: Time and invariant mass distributions of  $B_s \rightarrow D_s K \pi\pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

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

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

Table 2.3: Result of the time-dependent amplitude fit to EVTGEN toy events.

	Generated	Fit result	Pull( $\sigma$ )
$x_-$	0.179	$0.135 \pm 0.050$	-0.9
$y_-$	-0.324	$-0.307 \pm 0.022$	0.8
$x_+$	0.057	$0.102 \pm 0.065$	0.6
$y_+$	0.366	$0.394 \pm 0.023$	1.3

Table 2.4: Results for the physical observables obtained from fits to EVTGEN toy events.

	Generated	Full PDF	Phasespace integrated
$r$	0.370	$0.379 \pm 0.021$	$0.379 \pm 0.017$
$\kappa$	1.0	1.0	$1.000 \pm 0.059$
$\delta$	$10.0^\circ$	$9.0 \pm 5.1$	$5.9 \pm 6.0$
$\gamma$	$71.1^\circ$	$67.3 \pm 5.9$	$75.1 \pm 6.9$

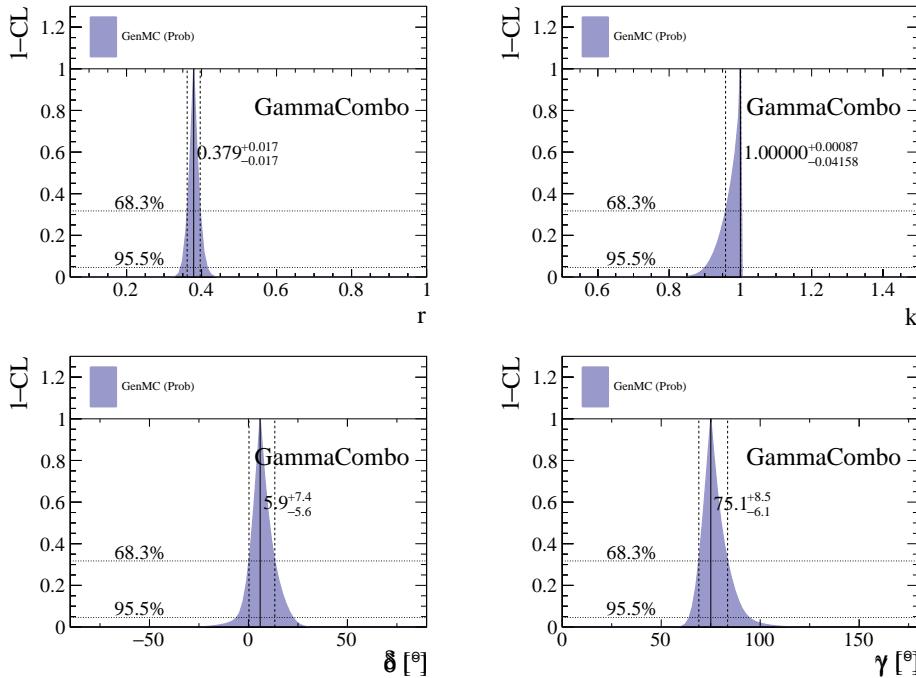


Figure 2.3: The 1-CL contours for the physical observable  $r, \kappa, \delta, \gamma$  obtained with the phasespace integrated fit to the EVTGEN toy sample.

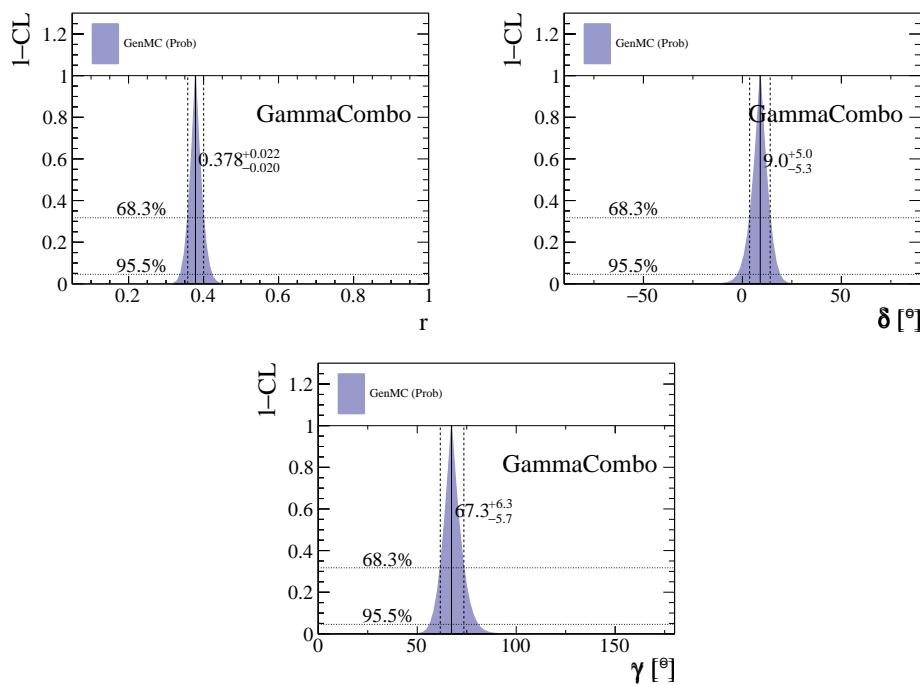


Figure 2.4: The 1-CL contours for the physical observable  $r, \delta, \gamma$  obtained with the time-dependent amplitude fit to the `EVTGEN` toy sample.

## 180 3 Data samples and event selection

### 181 3.1 Stripping and Trigger selection

182 The dataset used for this analysis corresponds to  $1\text{ fb}^{-1}$  of proton-proton collision  
183 data collected in 2011 with a centre of mass energy  $\sqrt{s} = 7\text{ TeV}$ ,  $2\text{ fb}^{-1}$  collected  
184 in 2012 with  $\sqrt{s} = 7\text{ TeV}$  and  $4\text{ fb}^{-1}$  collected in 2015/2016/2017 with  $\sqrt{s} =$   
185  $13\text{ TeV}$ . Candidate  $B_s^0 \rightarrow D_s K \pi \pi$  ( $B_s^0 \rightarrow D_s \pi \pi \pi$ ) decays are reconstructed using the  
186 `B02DKPiPiD2HHHPIDBeauty2CharmLine` (`B02DPiPiD2HHHPIDBeauty2CharmLine`) line  
187 of the `BHadronCompleteEvent` stream of `Stripping21r1` (2011), `Stripping21` (2012), `Strip-`  
188 `ping24r1` (2015) and `Stripping28r1p1` (2016) and `Stripping29r2` (2017). Both stripping lines  
189 employ the same selection cuts, listed in Appendix A, except for the PID requirement on  
190 the bachelor kaon/pion.

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

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

### 203 3.2 Offline selection

204 The offline selection, in particular the requirements on the  $D_s$  hadron, are guided by  
205 the previous analyses of  $B_s \rightarrow D_s K/\pi$ ,  $B_d \rightarrow D^0 \pi$  as well as the branching fraction  
206 measurement of  $B_s^0 \rightarrow D_s K \pi \pi$  decays. Tables 3.1 and 3.2 summarize all selection  
207 requirements which are described in the following.

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

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

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

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

227 We reconstruct the  $B_s^0 \rightarrow D_s h\pi\pi$  decay through three different final states of the  
228  $D_s$  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  
229 most prominent one, while  $\mathcal{BR}(D_s \rightarrow \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$  and  $\mathcal{BR}(D_s \rightarrow$   
230  $K\pi\pi) \approx 0.12 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$  holds for the others. For the  $KK\pi$  final state we make  
231 use of the well known resonance structure; the decay proceeds either via the narrow  $\phi$   
232 resonance, the broader  $K^{*0}$  resonance or (predominantly) non-resonant. Within the  $\phi$   
233 resonance region the sample is already sufficiently clean after the stripping so that we do  
234 not impose additional criteria on the  $D_s$  daughters. For the  $K^{*0}$  and non-resonant regions  
235 consecutively tighter requirements on the particle identification and the  $D_s$  flight-distance  
236 are applied. We apply global requirements for the other final states.

### 237 3.2.1 Phase space region

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

<sup>249</sup> **3.2.2 Physics background vetoes**

<sup>250</sup> We veto various physical backgrounds, which have either the same final state as our  
<sup>251</sup> signal decay, or can contribute via a single misidentification of  $K \leftrightarrow \pi$ ,  $K \leftrightarrow p$  or  $\pi \leftrightarrow p$ .  
<sup>252</sup> Depending on the  $D_s$  final state different vetoes are applied in order to account for peaking  
<sup>253</sup> backgrounds originating from charm meson or charmed baryon decays.

<sup>254</sup> 1.  $D_s^- \rightarrow K^+ K^- \pi^-$

<sup>255</sup> (a)  $D^- \rightarrow K^+ \pi^- \pi^-$ :

<sup>256</sup> Possible with  $\pi^- \rightarrow K^-$  misidentification, vetoed by requiring  $m(K^+ K_\pi^- \pi^-) \neq$   
<sup>257</sup>  $m(D^-) \pm 40$  MeV or the  $K^-$  has to fulfill more stringent PID criteria depending  
<sup>258</sup> on the resonant region.

<sup>259</sup> (b)  $\Lambda_c^- \rightarrow K^+ \bar{p} \pi^-$ :

<sup>260</sup> Possible with  $\bar{p} \rightarrow K^-$  misidentification, vetoed by requiring  $m(K^+ K_p^- \pi^-) \neq$   
<sup>261</sup>  $m(\Lambda_c^-) \pm 40$  MeV or the  $K^-$  has to fulfill more stringent PID criteria depending  
<sup>262</sup> on the resonant region.

<sup>263</sup> (c)  $D^0 \rightarrow KK$ :

<sup>264</sup>  $D^0$  combined with a random  $\pi$  can fake a  $D_s \rightarrow KK\pi$  decay, vetoed by  
<sup>265</sup> requiring  $m(KK) < 1840$  MeV.

<sup>266</sup> 2.  $D_s^- \rightarrow \pi^+ \pi^- \pi^-$

<sup>267</sup> (a)  $D^0 \rightarrow \pi\pi$ :

<sup>268</sup>  $D^0$  combined with a random  $\pi$  can fake a  $D_s \rightarrow \pi\pi\pi$  decay, vetoed by requiring  
<sup>269</sup> both possible combinations to have  $m(\pi\pi) < 1700$  MeV.

<sup>270</sup> 3.  $D_s^- \rightarrow K^- \pi^+ \pi^-$

<sup>271</sup> (a)  $D^- \rightarrow \pi^- \pi^+ \pi^-$ :

<sup>272</sup> Possible with  $\pi^- \rightarrow K^-$  misidentification, vetoed by requiring  $m(K_\pi^- \pi^+ \pi^-) \neq$   
<sup>273</sup>  $m(D^-) \pm 40$  MeV or  $\text{PIDK}(K^+) > 15$ .

<sup>274</sup> (b)  $\Lambda_c^- \rightarrow \bar{p} \pi^+ \pi^-$ :

<sup>275</sup> Possible with  $\bar{p} \rightarrow K^-$  misidentification, vetoed by requiring  $m(K_p^- \pi^+ \pi^-) \neq$   
<sup>276</sup>  $m(\Lambda_c^-) \pm 40$  MeV or  $\text{PIDK}(K^-) - \text{PIDp}(K^-) > 5$ .

<sup>277</sup> (c)  $D^0 \rightarrow K\pi$ :

<sup>278</sup>  $D^0$  combined with a random  $\pi$  can fake a  $D_s \rightarrow K\pi\pi$  decay, vetoed by requiring  
<sup>279</sup> both possible combinations to have  $m(K\pi) < 1750$  MeV.

<sup>280</sup> The effects of these veto cuts are illustrated in Figs. 3.1,3.2 and 3.3. To reduce cross-feed  
<sup>281</sup> from our calibration channel into the signal channel and vice-versa we require tight PID  
<sup>282</sup> cuts on the ambiguous bachelor kaon/pion. In addition, we veto  $B_s^0 \rightarrow D_s^- D_s^+$  decays  
<sup>283</sup> which is illustrated in Fig. 3.4.

- 284    1.  $X_s^+ \rightarrow K^+\pi^+\pi^-$ :
- 285    (a)  $B_s^0 \rightarrow D_s\pi\pi\pi$ :  
 286    Possible with  $\pi^+ \rightarrow K^+$  misidentification, suppressed with  $\text{PIDK}(K^+) > 10$ .
- 287    (b)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^+\pi^-)$ :  
 288    Outside of considered phase-space region, see Sec. 3.2.1.
- 289    (c)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+K^-\pi^+)$ :  
 290    To suppress  $B_s^0 \rightarrow D_s^-K^+K^-\pi^+$  background, possible with  $K^- \rightarrow \pi^-$  misiden-  
 291    tification, we require  $\text{PIDK}(\pi^-) < 0$ . In case the invariant mass of the  $K^+\pi^+\pi^-$   
 292    system recomputed applying the kaon mass hypothesis to the pion is close to  
 293    the  $D_s$  mass ( $m(K^+\pi^+\pi_K^-) = m(D_s) \pm 20$  MeV), we further tighten the cut to  
 294     $\text{PIDK}(\pi^-) < -5$ .
- 295    2.  $X_d^+ \rightarrow \pi^+\pi^+\pi^-$ :
- 296    (a)  $B_s^0 \rightarrow D_sK\pi\pi$ :  
 297    Possible with single missID of  $K^+ \rightarrow \pi^+$ , suppressed with  $\text{PIDK}(\pi^+) < 0$ .
- 298    (b)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow \pi^+\pi^+\pi^-)$ :  
 299    Outside of considered phase-space region, see Sec. 3.2.1.
- 300    (c)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^-\pi^+)$ :  
 301    Possible with single missID of  $K^+ \rightarrow \pi^+$ , vetoed by requiring  $m(\pi^+\pi_K^+\pi^-) \neq$   
 302     $m(D_s) \pm 20$  MeV or  $\text{PIDK}(\pi^+) < -5$  for both  $\pi^+$ .

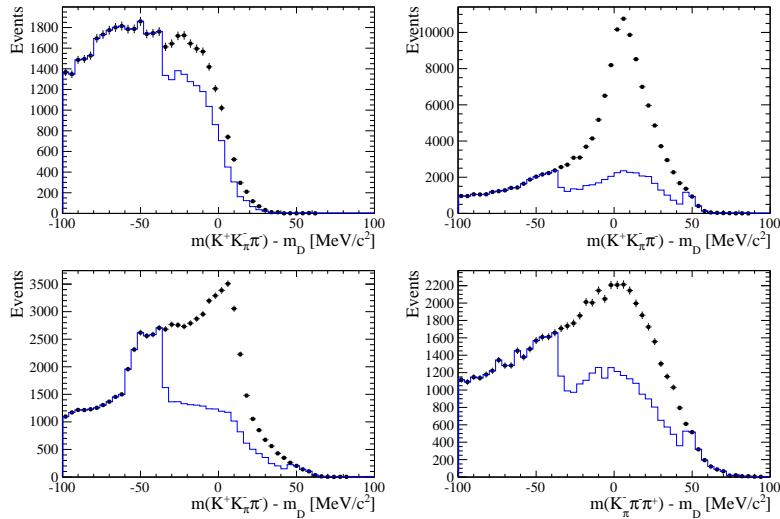


Figure 3.1: Background contributions from  $D^-$  decays where the  $\pi^-$  is misidentified as  $K^-$ . The  $D_s$  invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*(892)K$ ,  $D_s \rightarrow KK\pi$  (non-resonant) 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.

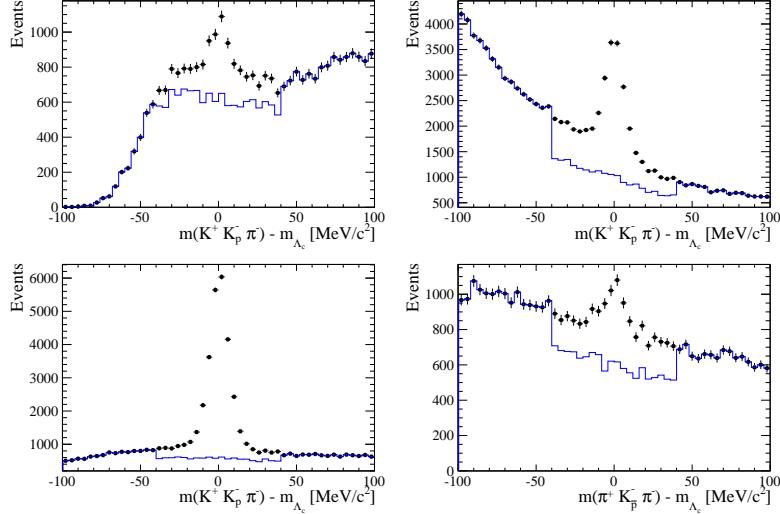


Figure 3.2: Background contributions from  $\Lambda_c$  decays where the  $\bar{p}$  is misidentified as  $K^-$ . The  $D_s$  invariant mass is recomputed applying the proton mass hypothesis to the kaon and shown for the  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*(892)K$ ,  $D_s \rightarrow KK\pi$  (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  $\Lambda_c$ -veto applied.

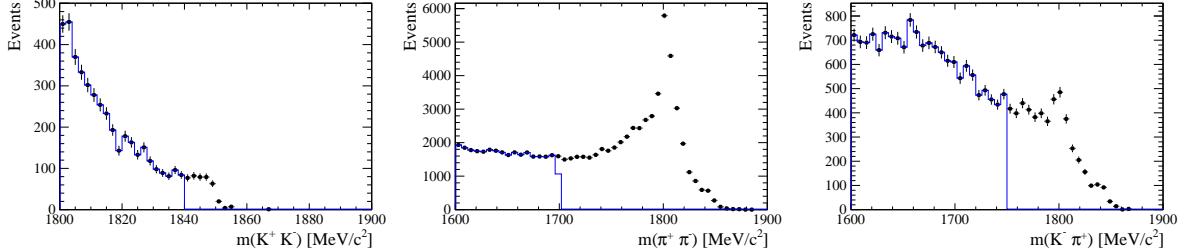


Figure 3.3: Background contributions to  $D_s \rightarrow KK\pi$  (left),  $D_s \rightarrow \pi\pi\pi$  (middle) and  $D_s \rightarrow K\pi\pi$  (right) from  $D^0 \rightarrow hh$  decays combined with a random pion.

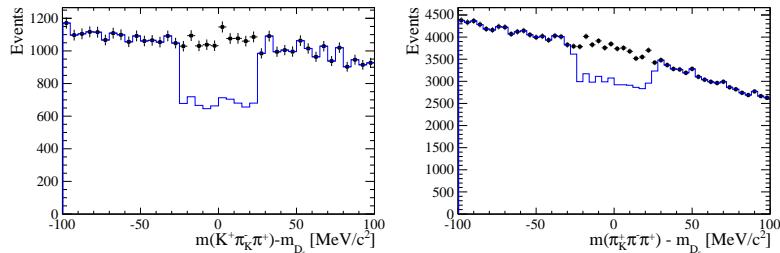


Figure 3.4: Background contributions to  $B_s \rightarrow D_s K\pi\pi$  (left) and  $B_s \rightarrow D_s \pi\pi\pi$  (right) from  $B_s \rightarrow D_s D_s$  decays where the kaon is misidentified as pion. The  $X_{s,d}$  invariant mass is recomputed applying the kaon mass hypothesis to the pion and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

303 **3.2.3 Training of multivariate classifier**

304 The Toolkit for Multivariate Analysis (TMVA [30]) is used to train a multivariate classifier  
 305 (BDT with gradient boosting, BDTG) discriminating signal and combinatorial background.  
 306 We use  $B_s \rightarrow D_s \pi\pi$  data that passes the preselection as signal proxy. The background  
 307 is statistically subtracted by applying `sWeights` based on the fit to the reconstructed  $B_s$   
 308 mass shown in Fig. 3.5. This is a simplified version (performed in a reduced mass range)  
 309 of the final mass fits described in Sec. 4. The sideband data ( $m(B_s) > 5500$  MeV) is used  
 310 as background proxy.

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

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

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

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

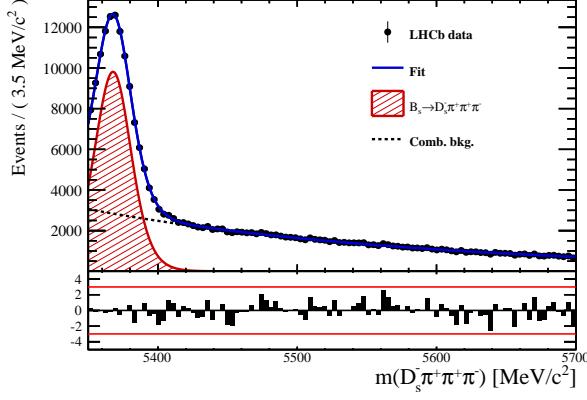


Figure 3.5: Reconstructed  $B_s$  mass for  $B_s \rightarrow D_s\pi\pi\pi$  events that pass the preselection. The fitted PDF is shown in blue, the signal component in red and the background component in black.

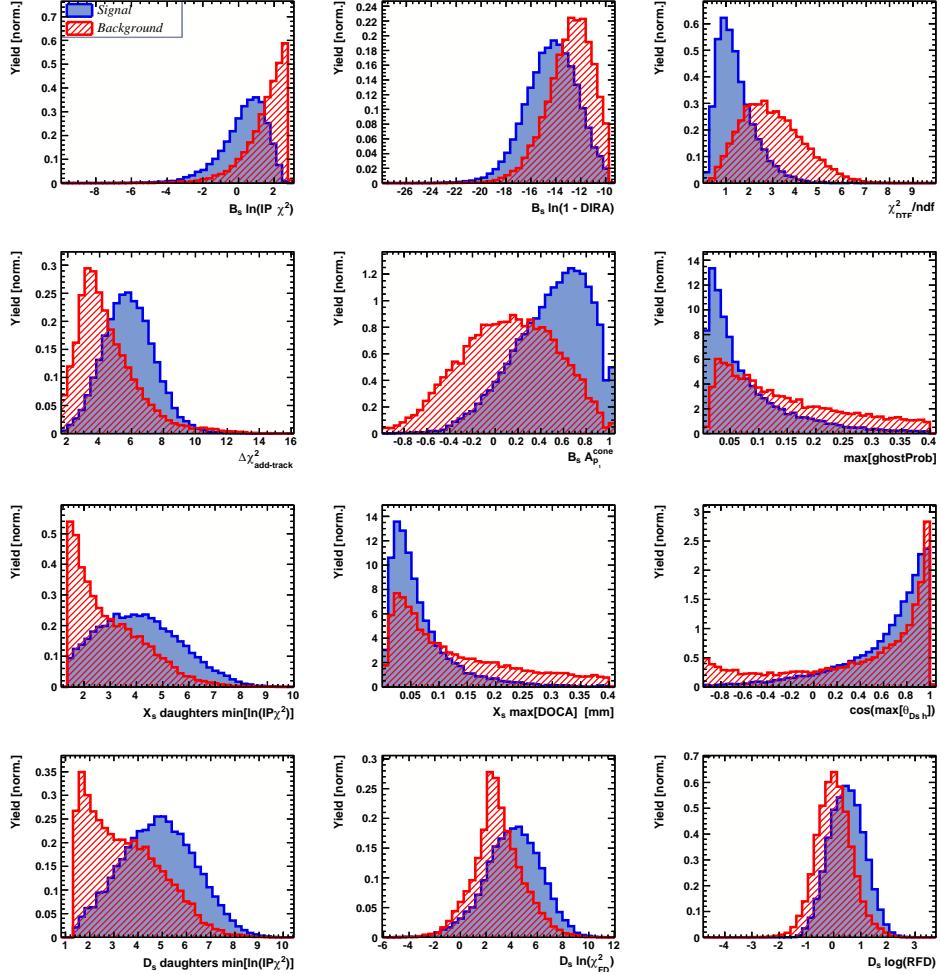


Figure 3.6: Discriminating variables used to train the BDTG for all data categories combined.

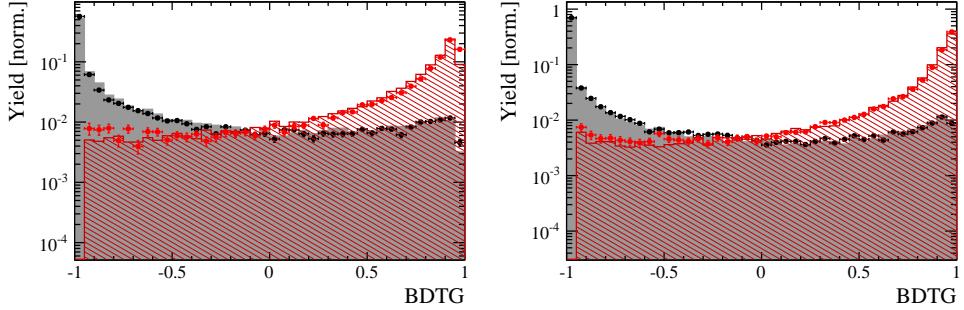


Figure 3.7: Signal (red) and background (black) distributions for the BDTG response for Run-I (left) and Run-II (right) data. Filled histograms (data points) show the BDTG response for the L0-TOS (L0-TIS) category. Even and odd test samples are combined.

### 342 3.2.4 Final selection

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

$$344 \quad \text{FOM}(\text{BDTG}) = \frac{N_s(\text{BDTG})}{\sqrt{N_s(\text{BDTG}) + N_b(\text{BDTG})}} \quad (3.1)$$

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

$$353 \quad N_{s,b}(\text{BDTG}) = N_{s,b}(0) \cdot \frac{\epsilon_{s,b}(\text{BDTG})}{\epsilon_{s,b}(0)}. \quad (3.2)$$

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

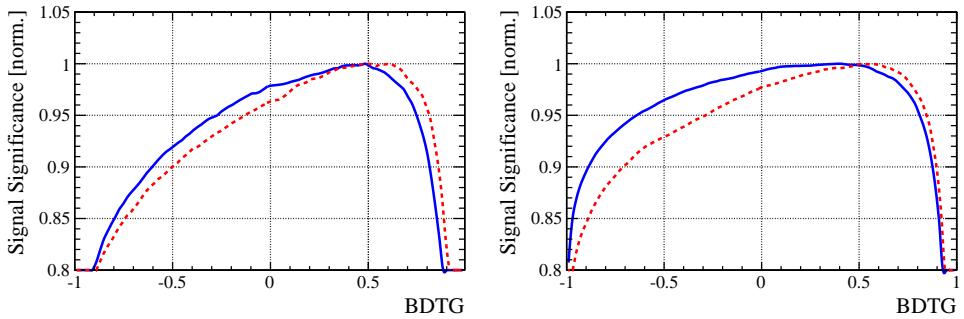


Figure 3.8: Signal significance as function of the applied cut on the BDTG response for Run-I (left) and Run-II (right) data. The scans for the L0-TOS (L0-TIS) category are shown in blue (red). The signal significance is normalized to be 1 at the optimal BDTG cut value.

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( $\pi^-$ ) $\chi_{FD}^2$ FD in $z$ $D^-$ veto $\Lambda_c$ 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( $\pi^-$ ) $\chi_{FD}^2$ FD in $z$ $D^-$ veto $\Lambda_c$ 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( $\pi^-$ ) $\chi_{FD}^2$ FD in $z$ $D^-$ veto $\Lambda_c$ 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( $\pi$ ) PIDp( $\pi$ ) $D^0$ veto $\chi_{FD}^2$ FD in $z$	$< 10$ $< 20$ $m(\pi^+\pi^-) < 1700$ MeV $> 9$ $> 0$
$D_s^- \rightarrow K^-\pi^+\pi^-$	PIDK( $K$ ) PIDK( $\pi$ ) PIDp( $\pi$ ) $D^0$ veto $\chi_{FD}^2$ FD in $z$ $D^-$ veto $\Lambda_c$ 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
$B_s \rightarrow D_s h\pi\pi$	$m(D_s h\pi\pi)$	$> 5200 \text{ MeV}$
	$\chi^2_{vtx}/\text{ndof}$	$< 8$
	DIRA	$> 0.99994$
	$\chi^2_{FD}$	$> 100$
	$\chi^2_{IP}$	$< 16$
	$\chi^2_{DTF}/\text{ndof}$	$< 15$
	$\Delta\chi^2_{add-track}$	$> 2$
	$\cos(\max \theta_{D_s h_i})$	$> -0.9$
	$t$	$> 0.4 \text{ ps}$
	$\delta t$	$< 0.15 \text{ ps}$
	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( $\pi^+$ )	$< 10$
	PIDK( $\pi^-$ )	$< 0$
	$D_s$ veto	$m(K^+\pi^+\pi_K^-) \neq m(D_s) \pm 20 \text{ MeV} \parallel \text{PIDK}(\pi^-) < -5$
$X_s^+ \rightarrow \pi^+\pi^+\pi^-$	PIDK( $\pi^+$ )	$< 0$
	PIDK( $\pi^-$ )	$< 10$
	$D_s$ veto	$m(\pi^+\pi_K^+\pi^-) \neq m(D_s) \pm 20 \text{ MeV} \parallel \text{PIDK}(\pi^+) < -5$
All tracks	hasRich	$= 1$

## 353 4 Yields determination

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

359 Due to different mass resolutions, we perform the invariant mass fits simultaneously  
 360 for each data-taking period and each trigger category. We further introduce four  $D_s$  final  
 361 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  
 362 different signal purities. The  $D_s \rightarrow Kh\pi$  category combines the two  $D_s$  decay channels  
 363 with the lowest statistics:  $D_s \rightarrow KK\pi$  (non-resonant) and  $D_s \rightarrow K\pi\pi$ . This amounts to  
 364 16 categories in total.

### 365 4.1 Signal model

366 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  
 367 using a Johnson's SU function [31], which results from a variable transformation of a  
 368 normal distribution to allow for asymmetric tails:

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

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

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

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

369 It is conveniently expressed in terms of the central moments up to order four: The mean  
 370 of the distribution  $\mu$ , the standard deviation  $\sigma$ , the skewness  $\nu$  and the kurtosis  $\tau$ . The  
 371 tail parameters  $\nu$  and  $\tau$  are fixed to the values obtained by a fit to the invariant mass  
 372 distribution of simulated events shown in Fig 4.1. To account for differences between

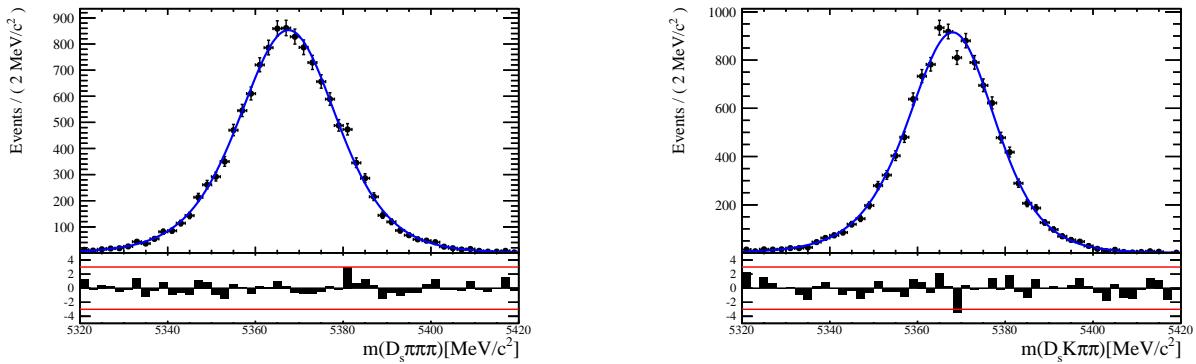


Figure 4.1: Invariant mass distributions of simulated (left)  $B_s^0 \rightarrow D_s \pi\pi\pi$  and (right)  $B_s^0 \rightarrow D_s K\pi\pi$  events. A fit with a Johnson's SU PDF is overlaid.

372  
373 simulation and real data, linear scaling factors for the mean  $\mu$  and width  $\sigma$  are determined  
374 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  
375 factors are determined separately for each data-taking period and each trigger category.

## 376 4.2 Background models

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

### 379 380 Combinatorial background

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

### 384 385 Peaking $B_d$ background

386 Decays of  $B_d$  mesons into the  $D_s h\pi\pi$  final state are described by the  $B_s$  signal PDF  
387 where the mean is shifted by the known mass difference  $m_{B_s} - m_{B_d}$  [?].

### 388 389 Partially reconstructed background

390 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  
391 to be peaking lower than signal in the  $m(D_s\pi\pi\pi)$  spectrum with large tails due to the  
392 momentum carried away by the not reconstructed  $\pi^0$  or  $\gamma$ . An empirical description for  
393 the shape of this contribution is derived from a  $B_s^0 \rightarrow D_s^*\pi\pi\pi$  MC sample subject to  
394 the nominal  $B_s^0 \rightarrow D_s\pi\pi\pi$  selection. Figure 4.2 (left) shows the respective reconstructed  
395  $m(D_s\pi\pi\pi)$  distribution. A sum of three bifurcated Gaussian functions is used to describe  
396 it. In the fit to data, all parameters are fixed to the ones obtained from MC except for  
397 the parameter which describes the width of the right tail of the distribution to account for  
398 data-simulation differences in mass resolution. The equivalent  $B_s^0 \rightarrow D_s^*K\pi\pi$  component  
399 contributing to the  $B_s^0 \rightarrow D_sK\pi\pi$  data sample is described by the same PDF with the  
400 right tail fixed to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  result.

401 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  
402 shifted by  $m_{B_s^0} - m_{B^0}$ .

### 403 404 Misidentified background

405 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  
406 pions is misidentified as a kaon, contaminate the  $B_s^0 \rightarrow D_s K^+ \pi^+ \pi^-$  sample. To determine  
407 the corresponding background shapes, we use simulated events passing the nominal  
408 selection except for the PID cuts on the bachelor  $\pi^+$  tracks. The **PIDCalib** package  
409 is used to determine the  $p_T, \eta$ -dependent  $\pi^+ \rightarrow K^+$  misidentification probability for  
410 each pion. We change the particle hypothesis from pion to kaon for the pion with the  
411 higher misidentification probability and recompute the invariant  $B_s^0$  mass,  $m(D_s^- \pi_K^+ \pi^+ \pi^-)$ .  
412 Similar, the invariant masses  $m(\pi_K^+ \pi^+ \pi^-)$  and  $m(\pi_K^+ \pi^-)$  are recomputed and required  
413 to be within the considered phasespace region. The background distributions are shown  
414 in Fig. 4.2 (middle,right) and modeled by the sum of two Crystal Ball functions. The  
415 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  
416  $B_s^0 \rightarrow D_s K\pi\pi$  sample is computed by multiplying the fake rate (within the considered

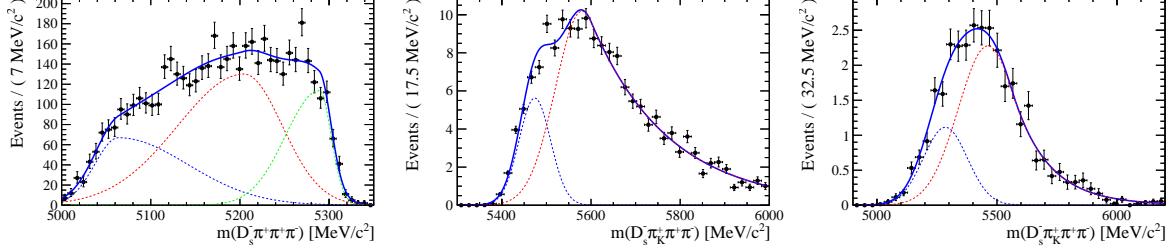


Figure 4.2: Left: Invariant mass distribution of simulated  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  events, where the  $\gamma/\pi^0$  is excluded from the reconstruction. Middle: Invariant mass distribution of simulated  $B_s^0 \rightarrow D_s \pi\pi\pi$  events, where one of the pions is reconstructed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

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

422 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  
423  $B_s^0 \rightarrow D_s \pi\pi\pi$  data sample to be negligible due to the low branching fractions and the  
424 tight PID cuts on the bachelor pions.

### 4.3 Results

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

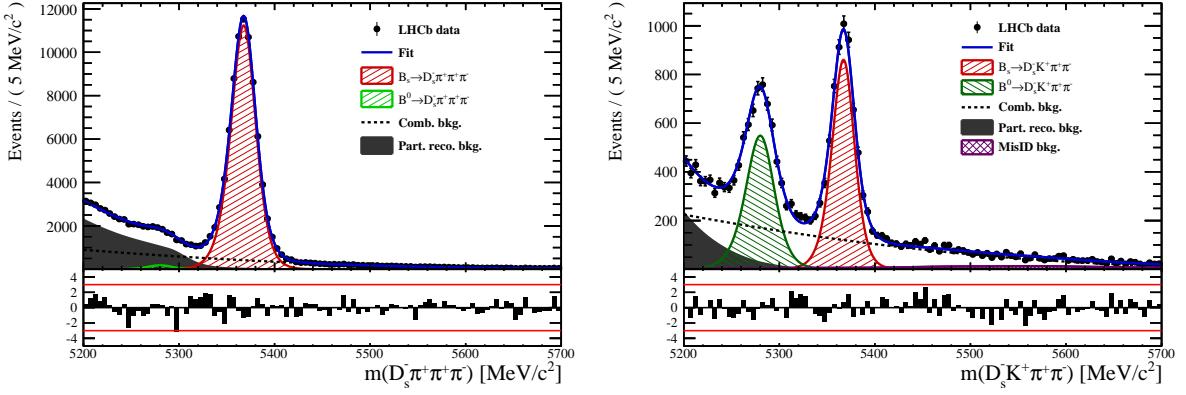


Figure 4.3: Invariant mass distribution of  $B_s^0 \rightarrow D_s \pi\pi\pi$  (left) and  $B_s^0 \rightarrow D_s K\pi\pi$  (right) candidates.

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

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

## 430 5 Decay-time Resolution

431 The observed oscillation of B mesons is prone to dilution, if the detector resolution is  
 432 of similar magnitude as the oscillation period. In the  $B_s^0$  system, considering that the  
 433 measured oscillation frequency of the  $B_s^0$  [32] and the average LHCb detector resolution [33]  
 434 are both  $\mathcal{O}(50 \text{ fs}^{-1})$ , this is the case. Therefore, it is crucial to correctly describe the  
 435 decay time resolution in order to avoid a bias on the measurement of time dependent CP  
 436 violation. Since the time resolution depends on the particular event, especially the decay  
 437 time itself, the sensitivity on  $\gamma$  can be significantly improved by using an event dependent  
 438 resolution model rather than an average resolution. For this purpose, we use the per-event  
 439 decay time error that is estimated based on the uncertainty obtained from the global  
 440 kinematic fit of the momenta and vertex positions (DTF) with additional constraints on  
 441 the PV position and the  $D_s$  mass. In order to apply it correctly, it has to be calibrated.  
 442 The raw decay time error distributions for  $B_s^0 \rightarrow D_s\pi\pi\pi$  signal candidates are shown in  
 443 Figure 5.1 for Run-I and Run-II data. Significant deviations between the two different  
 444 data taking periods are observed due to the increase in center of mass energy from Run-I  
 445 to Run-II, as well as (among others) new tunings in the pattern and vertex reconstruction.  
 446 The decay time error calibration is consequently performed separately for both data taking  
 447 periods.

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

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

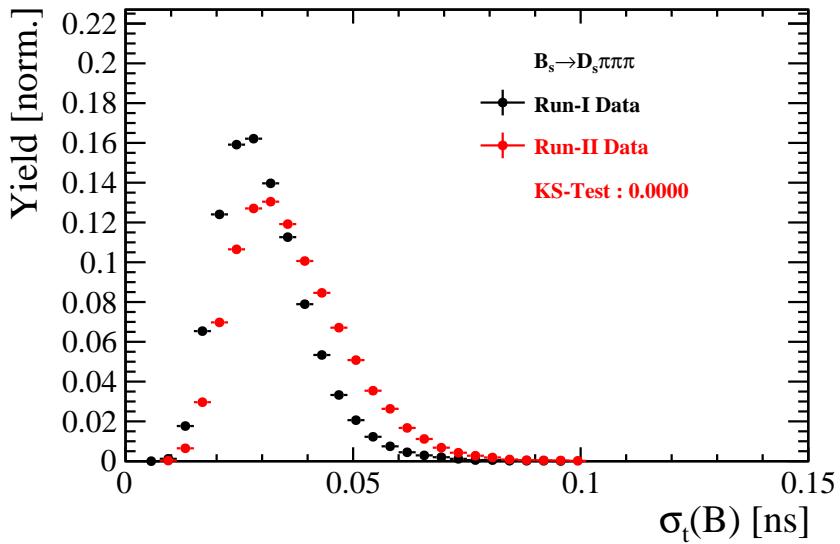


Figure 5.1: Distribution of the decay time error for  $B_s^0 \rightarrow D_s\pi\pi\pi$  signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

## 453 5.1 Calibration for Run-I data

454 For simulated  $B_s^0 \rightarrow D_s K \pi\pi$  events, the spread of the differences between reconstructed  
 455 decay time and true decay time,  $\Delta t = t - t_{true}$ , is a direct measure of the decay time  
 456 resolution. The sum of two Gaussian pdfs with common mean but different widths is used  
 457 to fit the distribution of the decay time difference  $\Delta t$  as shown in Fig. 5.2. The effective  
 458 damping of the CP amplitudes due to the finite time resolution is described by the dilution  
 459  $\mathcal{D}$ . In the case of infinite precision, there would be no damping and therefore  $\mathcal{D} = 1$  would  
 460 hold, while for a resolution that is much larger than the  $B_s^0$  oscillation frequency,  $\mathcal{D}$  would  
 461 approach 0. For a double-Gaussian resolution model, the dilution is given by [34]

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

462 where  $\sigma_1$  and  $\sigma_2$  are the widths of the Gaussians,  $f_1$  is the relative fraction of events  
 463 described by the first Gaussian relative to the second and  $\Delta m_s$  is the oscillation frequency  
 464 of  $B_s^0$  mesons. An effective single Gaussian width is calculated from the dilution as,

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

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

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

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

476 where the offset is fixed to 0. For comparison, the calibration function found for  $B_s^0 \rightarrow D_s K$   
 477 MC is also shown in Figure 5.2 [34]:

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

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

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

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

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

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

483 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  
 484 assigned as additional systematic uncertainty.

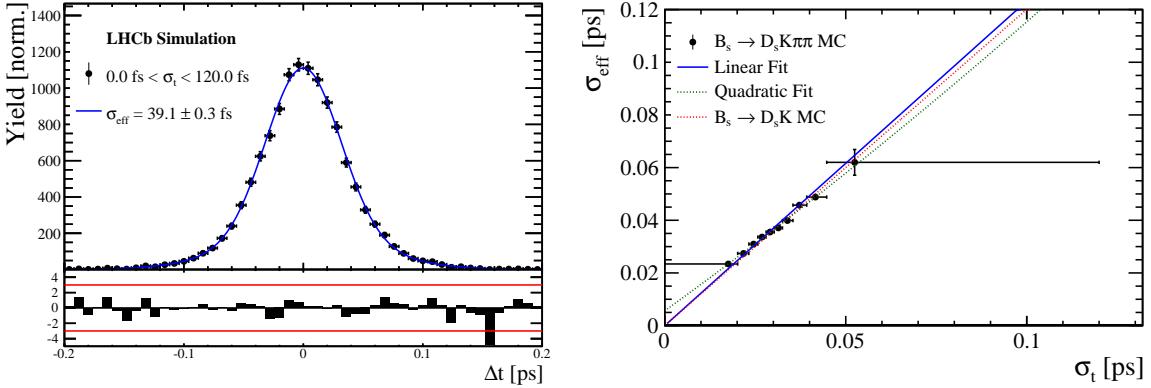


Figure 5.2: (Left) Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K\pi\pi$  MC candidates. (Right) The measured resolution  $\sigma_{eff}$  as function of the per-event decay time error estimate  $\sigma_t$  for  $B_s \rightarrow D_s K\pi\pi$  MC (Run-I). The fitted calibration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

## 485 5.2 Calibration for Run-II data

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

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

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

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

<sup>512</sup> is in good agreement with the one obtained for the  $B_s \rightarrow J/\psi\phi$  (Run-II) analysis.

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

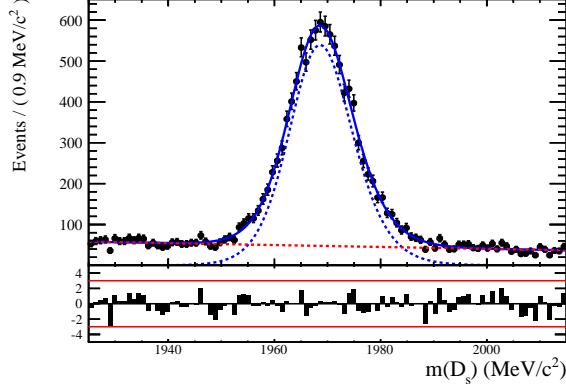


Figure 5.3: The invariant mass distribution for prompt  $D_s$  candidates.

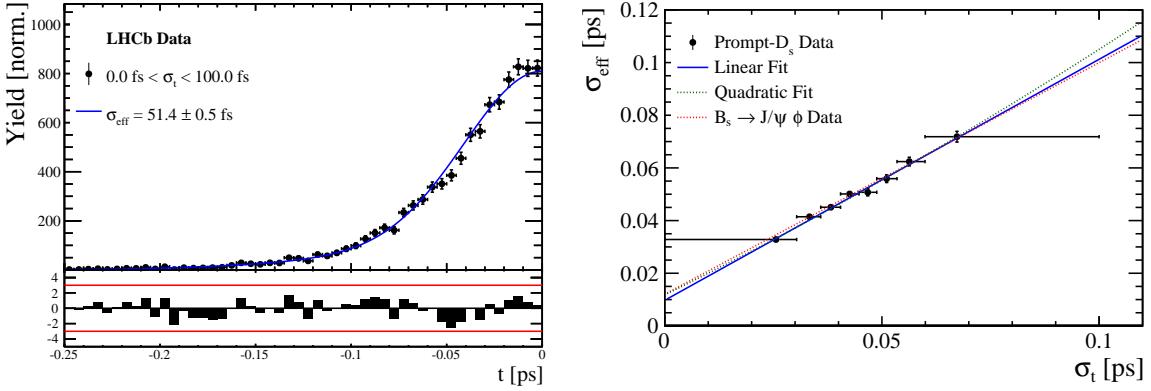


Figure 5.4: (Left) Decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks. (Right) The measured resolution  $\sigma_{eff}$  as function of the per-event decay time error estimate  $\sigma_t$  for fake  $B_s$  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
$B_s \rightarrow D_s K\pi\pi$	$\chi_{vtx}^2/\text{ndof}$	< 8
	$\chi_{DTF}^2/\text{ndof}$	< 15
	$t$	< 0 ps
$D_s \rightarrow hhh$	$\chi_{vtx}^2/\text{ndof}$	< 5
	DIRA	> 0.99994
	$\chi_{FD}^2$	> 9
	$p_T$	> 1800 MeV
	$\chi_{IP}^2$	< 9
	$\chi_{IP}^2(h)$	> 5
$D_s^- \rightarrow KK\pi^-$	Wrong PV veto	$nPV = 1 \parallel \min(\Delta\chi_{IP}^2) > 20$
	$D^0$ veto	$m(KK) < 1840$ MeV
$D_s^- \rightarrow K^*(892)K^-$	$D^-$ veto	$m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 30$ MeV
	$\Lambda_c$ veto	$m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 30$ MeV
	$m(KK)$	$= m_\phi \pm 20$ MeV
$D_s^- \rightarrow \phi\pi^-$	PIDK( $K^+$ )	> -10
	PIDK( $K^-$ )	> -10
	PIDK( $\pi^-$ )	< 20
	$m(KK)$	$\neq m_\phi \pm 20$ MeV
$D_s^- \rightarrow (KK\pi^-)_{NR}$	$m(K^+\pi^-)$	$= m_{K^*(892)} \pm 75$ MeV
	PIDK( $K^+$ )	> -10
	PIDK( $K^-$ )	> -5
	PIDK( $\pi^-$ )	< 20
	$m(KK)$	$\neq m_\phi \pm 20$ MeV
$D_s \rightarrow \pi\pi\pi$	$m(K^+\pi^-)$	$\neq m_{K^*(892)} \pm 75$ MeV
	PIDK( $K^+$ )	> 5
	PIDK( $K^-$ )	> 5
	PIDK( $\pi^-$ )	< 10
$X_s \rightarrow K\pi\pi$	PIDK( $h$ )	< 10
	PIDp( $h$ )	< 10
	$D^0$ veto	$m(\pi^+\pi^-) < 1700$ MeV
$X_s \rightarrow K\pi\pi$	$\chi_{IP}^2(h)$	< 40
	PIDK( $K$ )	> 10
	PIDK( $\pi$ )	< 5
	isMuon( $h$ )	False
All tracks	$p_T$	> 500 MeV

513 **5.3 Cross-checks**

514 **5.3.1 Kinematic dependence**

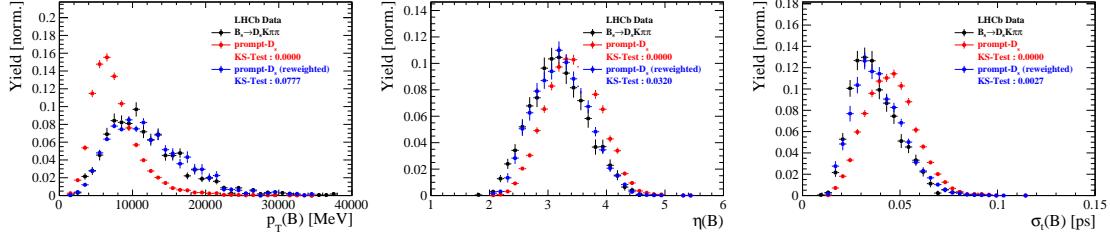


Figure 5.5

515 **5.3.2 DTF constraints**

## 516 6 Acceptance

### 517 6.1 MC corrections

#### 518 6.1.1 Truth matching of simulated candidates

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

528 We therefore include `BKGCAT`= 60 and statistically subtract the ghost background by  
 529 using the `sPlot` technique. The `sWeights` are calculated from a fit to the reconstructed  $B_s$   
 530 mass. The signal contribution is modeled as described in Sec. ?? and the background with  
 531 a polynomial. The fit is performed simultaneously in two categories; the first includes  
 532 events with `BKGCAT` = 20 or 50 and the second events with `BKGCAT` = 60. To account  
 533 for the different mass resolution we use a different  $\sigma$  for each category, while the mean  
 534 and the tail parameters are shared between them. The background component is only  
 535 included for the second category.

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

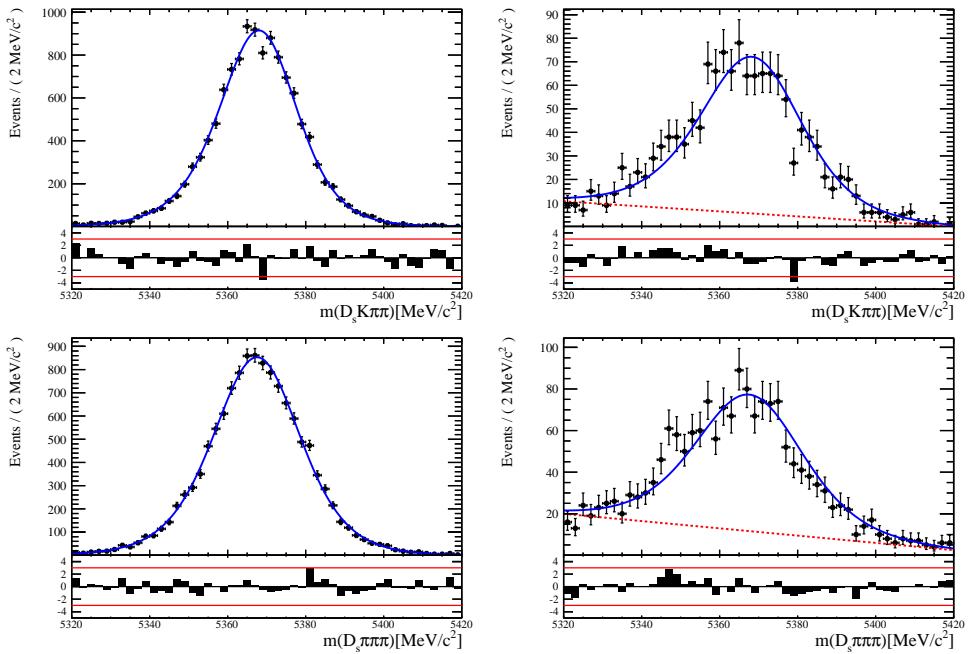


Figure 6.1: The reconstructed  $B_s$  mass distribution for simulated  $B_s \rightarrow D_s K\pi\pi$  (top) and  $B_s \rightarrow D_s \pi\pi\pi$  (bottom) decays classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

538 6.1.2 PID efficiencies

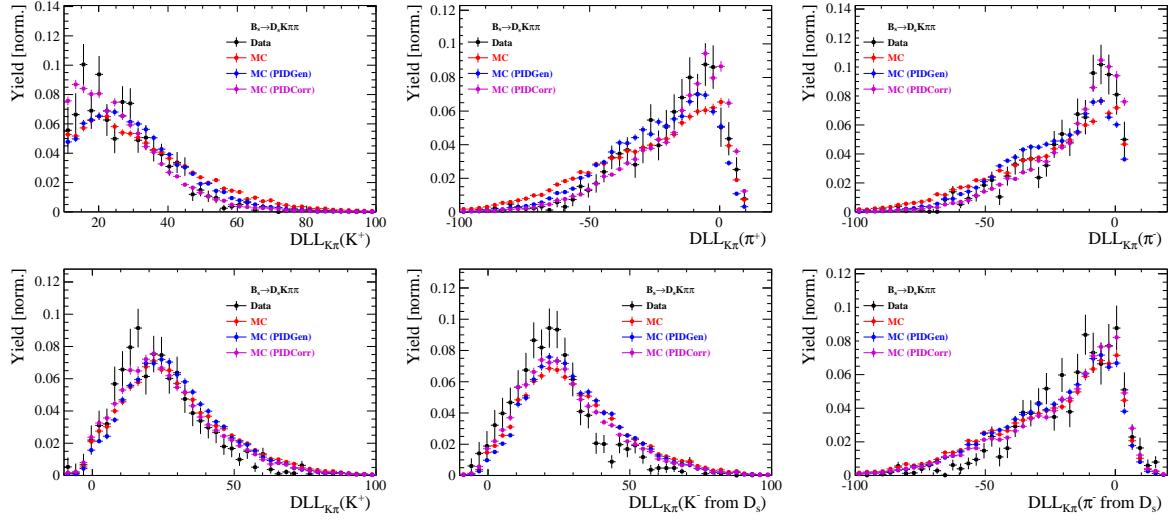


Figure 6.2

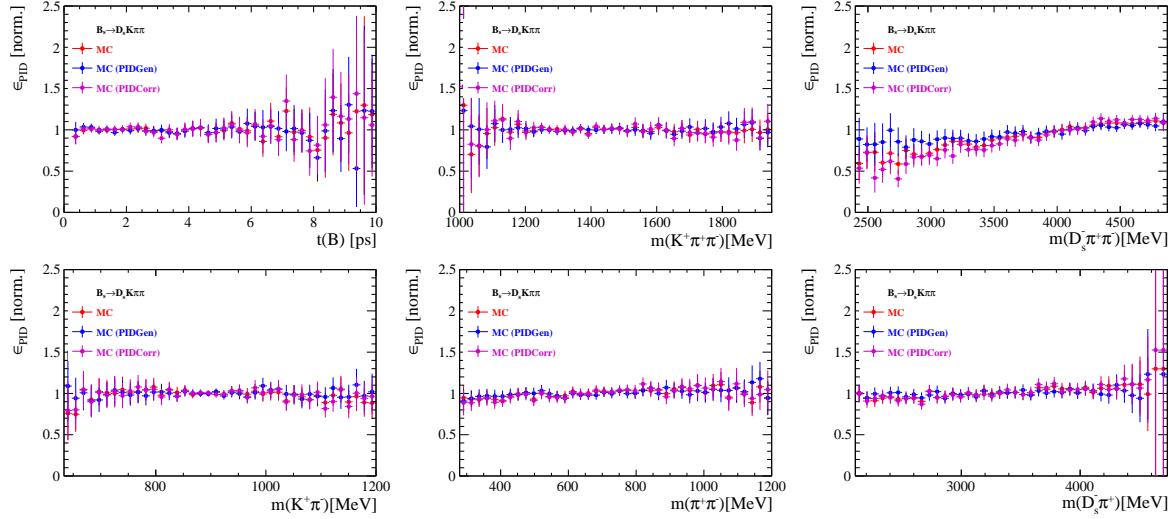


Figure 6.3

539 6.1.3 BDT efficiencies

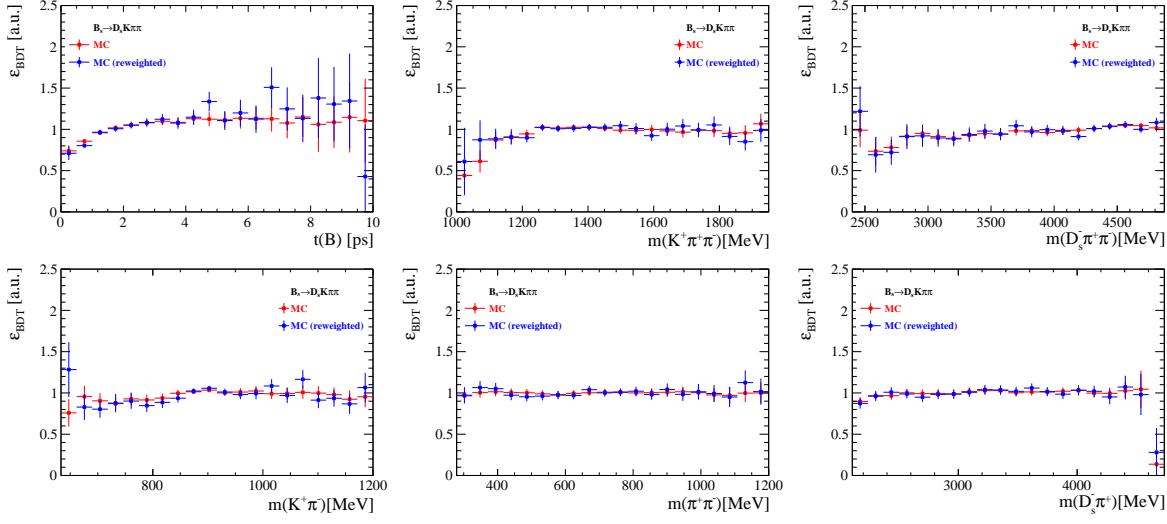


Figure 6.4

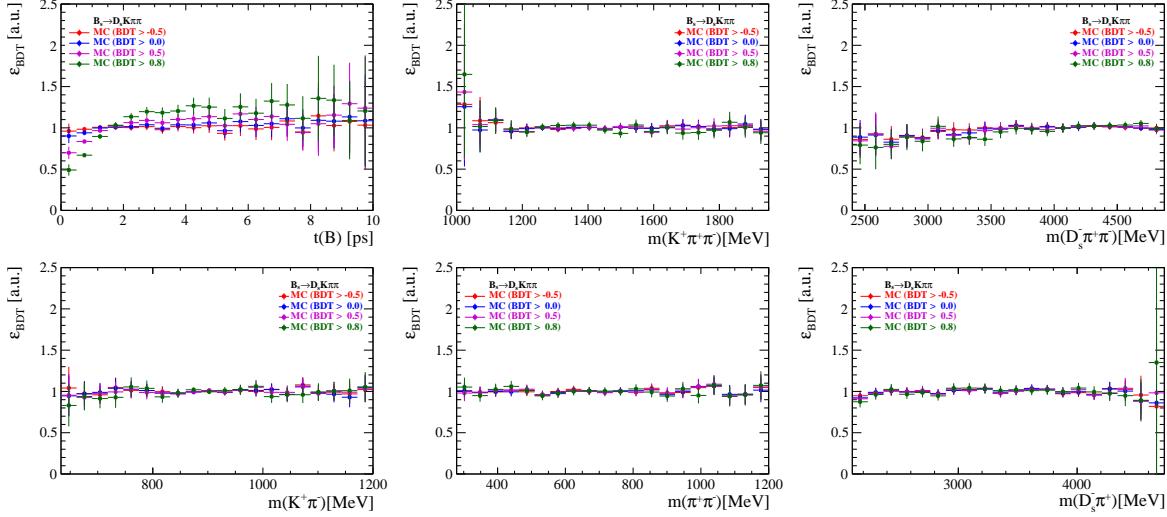


Figure 6.5

<sup>540</sup> 6.1.4 Tracking efficiencies

541 **6.2 Decay-time acceptance**

542 The decay-time distribution of the  $B_s^0$  mesons is sculpted due to the geometry of the LHCb  
 543 detector and the applied selection cuts, which are described in Section 3. In particular,  
 544 any requirement on the flight distance (FD), the impact parameter (IP) or the direction  
 545 angle (DIRA) of the  $B_s^0$  mesons, as well as the direct cut on the lifetime, will lead to a  
 546 decay-time dependent efficiency  $a(t)$ . This efficiency will distort the theoretically expected,  
 547 time-dependent decay rate

$$\frac{\Gamma(t)^{observed}}{dt} = \frac{\Gamma(t)^{theory}}{dt} \cdot a(t), \quad (6.1)$$

548 and has to be modelled correctly, in order to describe the observed decay rate. We  
 549 use our control channel for this measurement, because for  $B_s^0 \rightarrow D_s K\pi\pi$  decays the  
 550 decay-time acceptance is correlated with the CP-observables which we aim to measure.  
 551 Therefore, floating the CP-observables and the acceptance shape at the same time is  
 552 not possible. Hence, a fit to the decay-time distribution of  $B_s^0 \rightarrow D_s \pi\pi\pi$  candidates is  
 553 performed and the obtained acceptance shape is corrected by the difference in shape found  
 554 for the  $B_s^0 \rightarrow D_s K\pi\pi$  and  $B_s^0 \rightarrow D_s \pi\pi\pi$  MC.

555 A PDF of the form

$$\mathcal{P}(t', \vec{\lambda}) = \left[ (e^{\Gamma_s t} \cdot \cosh(\frac{\Delta\Gamma_s t}{2}) \times \mathcal{R}(t - t')) \right] \cdot \epsilon(t', \vec{\lambda}), \quad (6.2)$$

556 is fit to the decay time distribution of  $B_s^0 \rightarrow D_s \pi\pi\pi$  candidates in data. Since the  
 557 fit is performed untagged, the PDF shown in Eq. 6.2 contains no terms proportional  
 558 to  $\Delta m_s$ . The values for  $\Gamma_s$  and  $\Delta\Gamma_s$  are fixed to the latest HFAG results [35]. The  
 559 decay-time acceptance  $\epsilon(t', \vec{\lambda})$  is modelled using the sum of cubic polynomials  $v_i(t)$ , so  
 560 called Splines [36]. The polynomials are parametrised by so-called knots which determine  
 561 their boundaries. Knots can be set across the fitted distribution to account for local  
 562 changes in the acceptance shape. Using more knots is equivalent to using more base  
 563 splines which are defined on a smaller sub-range. In total,  $n + 2$  base splines  $v_i(t)$  are  
 564 needed to describe an acceptance shape which is parametrised using  $n$  knots.

565 For fits shown in the following, the knots have been placed at  $t =$   
 566  $[0.5, 1.0, 1.5, 2.0, 3.0, 9.5] ps$ . To accommodate these 6 knot positions, 8 basic splines  
 567  $v_i$ ,  $i = [1, \dots, 8]$  are used. Since a rapid change of the decay time acceptance at low  
 568 decay times due to the turn-on effect generated by the lifetime and other selection cuts is  
 569 expected, more knots are placed in that regime. At higher decay times we expect linear  
 570 behavior, with a possible small effect due to the VELO reconstruction. Therefore fewer  
 571 knots are used. Furthermore,  $v_7$  is fixed to 1 in order to normalize the overall acceptance  
 572 function. To stabilise the last spline,  $v_8$  is fixed by a linear extrapolation from the two  
 573 previous splines:

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

574 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. 6.2, is repeated and the obtained values for the spline coefficients  $v_i$  are compared. Figure 6.6 shows the comparison of the obtained spline coefficients for the different  $D_s$  final states.

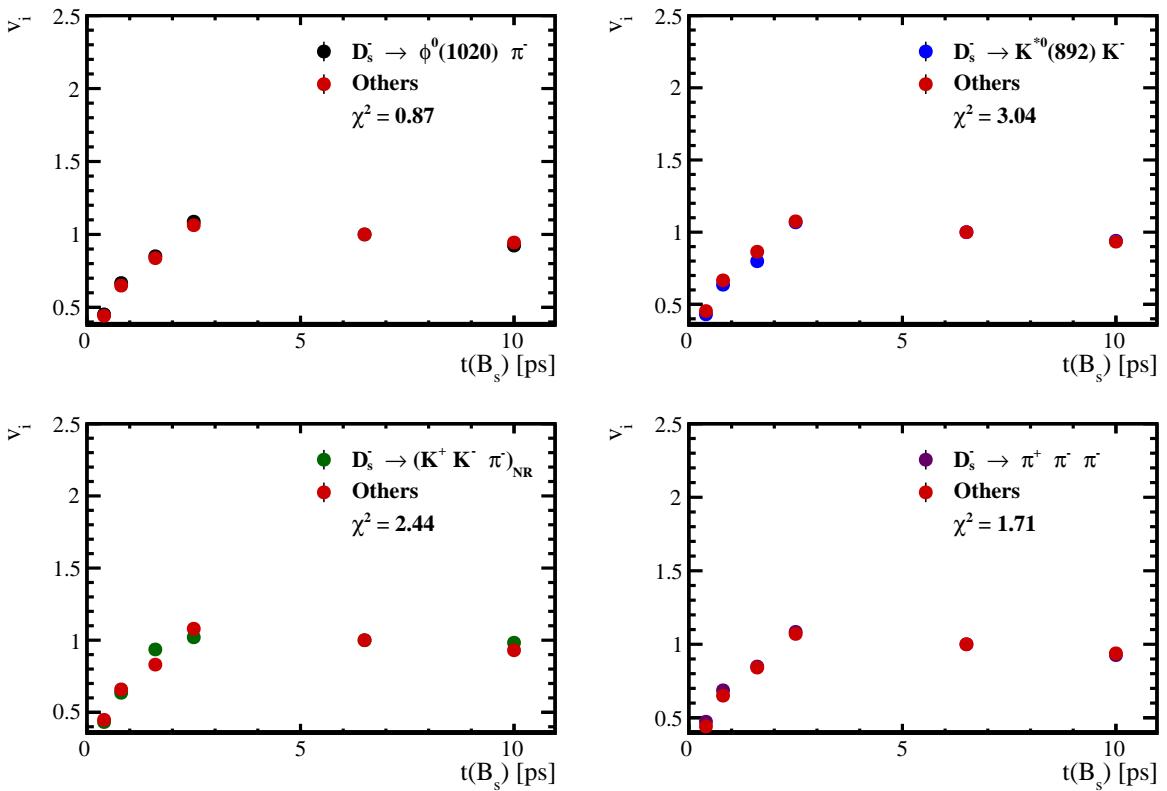


Figure 6.6: Comparison of the spline coefficients obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  subsamples of different  $D_s$  final states. The comparison of one particular  $D_s$  state against all other states is shown.

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 6.7.

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

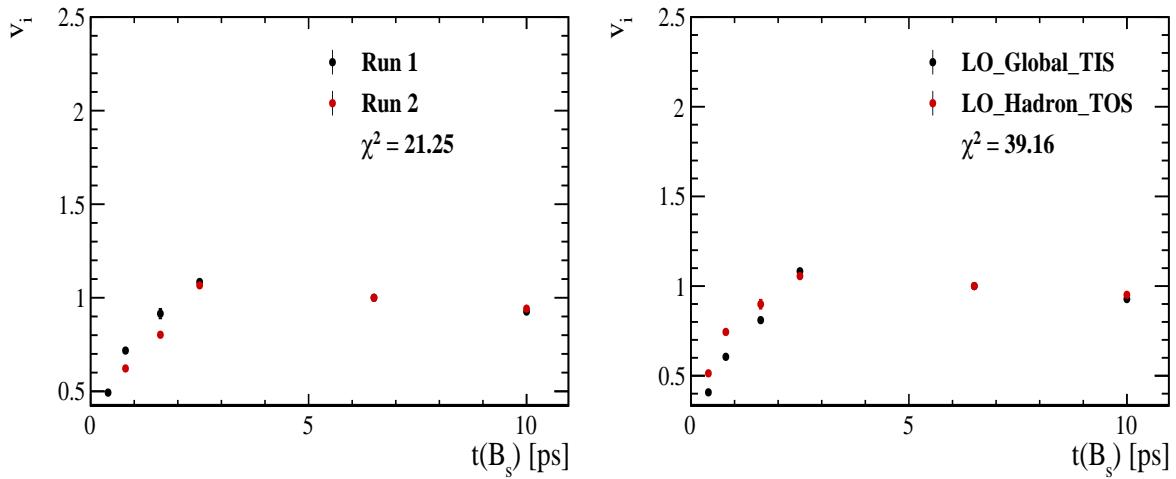


Figure 6.7: Comparison of the spline coefficients obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  subsamples of (left) the different runs and (right) L0 trigger categories.

594 implementing a simultaneous fit, where the acceptance description is allowed to vary in  
 595 the subsamples.

596 **6.2.2 Results**

597 The nominal fit to  $B_s^0 \rightarrow D_s\pi\pi\pi$  data using this configuration is shown in Figure ??.  
598 Note that the normalization of the splines in the following figures is not in scale. The fit  
599 parameters obtained from the described fits to data and simulation are summarised in  
600 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	$v_0$	$0.561 \pm 0.038$	$0.546 \pm 0.022$	$0.953 \pm 0.060$
0.8	$v_1$	$0.826 \pm 0.059$	$0.785 \pm 0.034$	$0.910 \pm 0.066$
1.6	$v_2$	$0.843 \pm 0.087$	$0.905 \pm 0.056$	$1.055 \pm 0.095$
2.5	$v_3$	$1.154 \pm 0.036$	$1.118 \pm 0.028$	$0.930 \pm 0.045$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.866 (interpolated)	0.897 (interpolated)	1.061 (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	$v_0$	$0.368 \pm 0.031$	$0.412 \pm 0.020$	$0.955 \pm 0.077$
0.8	$v_1$	$0.583 \pm 0.050$	$0.648 \pm 0.033$	$0.910 \pm 0.074$
1.6	$v_2$	$0.939 \pm 0.101$	$0.953 \pm 0.061$	$0.947 \pm 0.096$
2.5	$v_3$	$1.052 \pm 0.054$	$1.077 \pm 0.035$	$1.003 \pm 0.051$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.954 (interpolated)	0.932 (interpolated)	0.998 (interpolated)

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

Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.569 \pm 0.028$	$0.496 \pm 0.015$	$0.966 \pm 0.044$
0.8	$v_1$	$0.787 \pm 0.043$	$0.737 \pm 0.024$	$0.893 \pm 0.049$
1.6	$v_2$	$0.899 \pm 0.062$	$0.943 \pm 0.039$	$0.985 \pm 0.060$
2.5	$v_3$	$1.080 \pm 0.030$	$1.093 \pm 0.022$	$0.980 \pm 0.031$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.930 (interpolated)	0.919 (interpolated)	1.018 (interpolated)

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

Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.389 \pm 0.020$	$0.506 \pm 0.015$	$0.909 \pm 0.041$
0.8	$v_1$	$0.593 \pm 0.033$	$0.744 \pm 0.024$	$0.897 \pm 0.048$
1.6	$v_2$	$0.799 \pm 0.052$	$0.965 \pm 0.039$	$0.928 \pm 0.052$
2.5	$v_3$	$1.112 \pm 0.033$	$1.112 \pm 0.022$	$0.941 \pm 0.036$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.902 (interpolated)	0.902 (interpolated)	1.052 (interpolated)

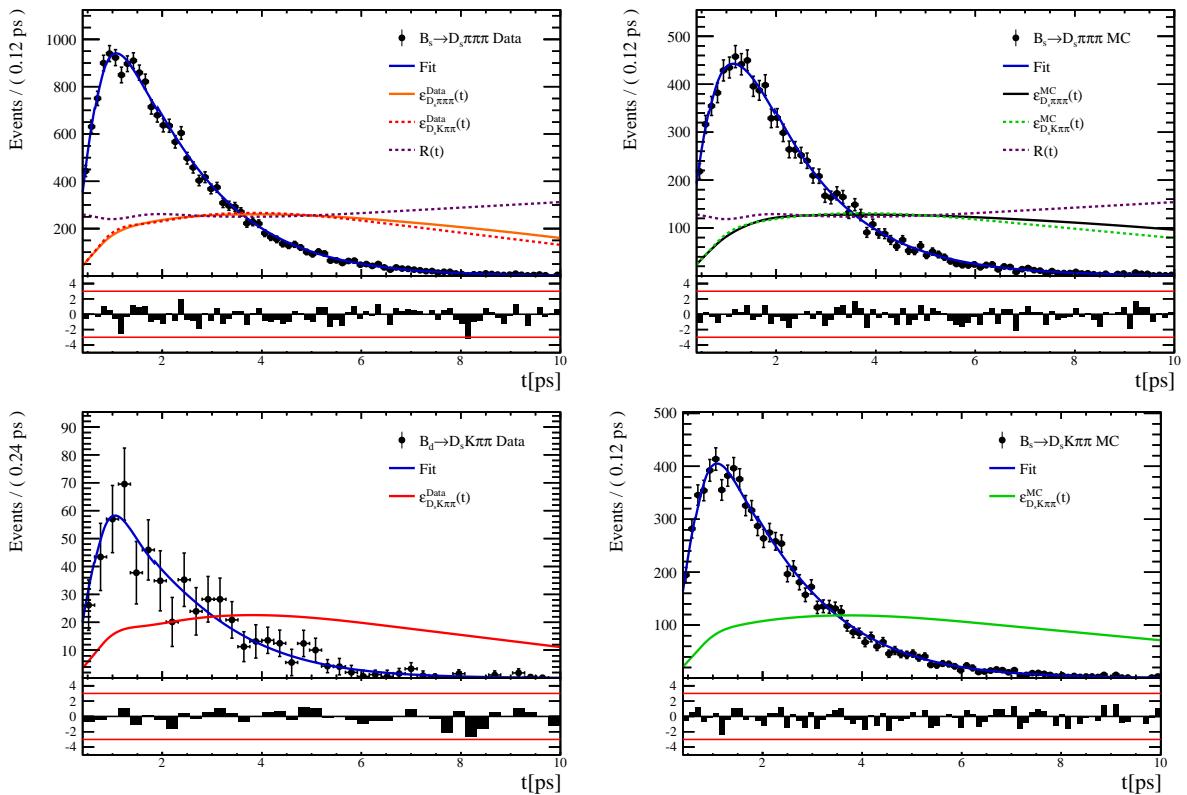


Figure 6.8

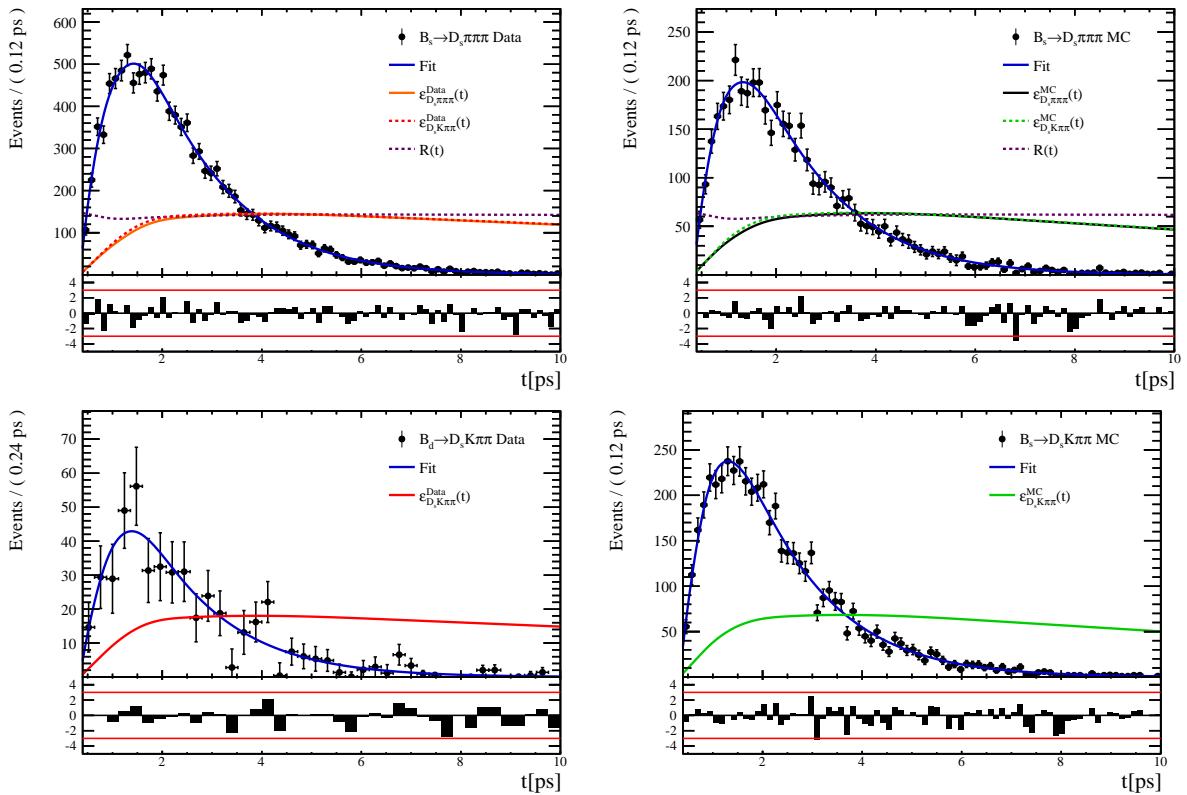


Figure 6.9:

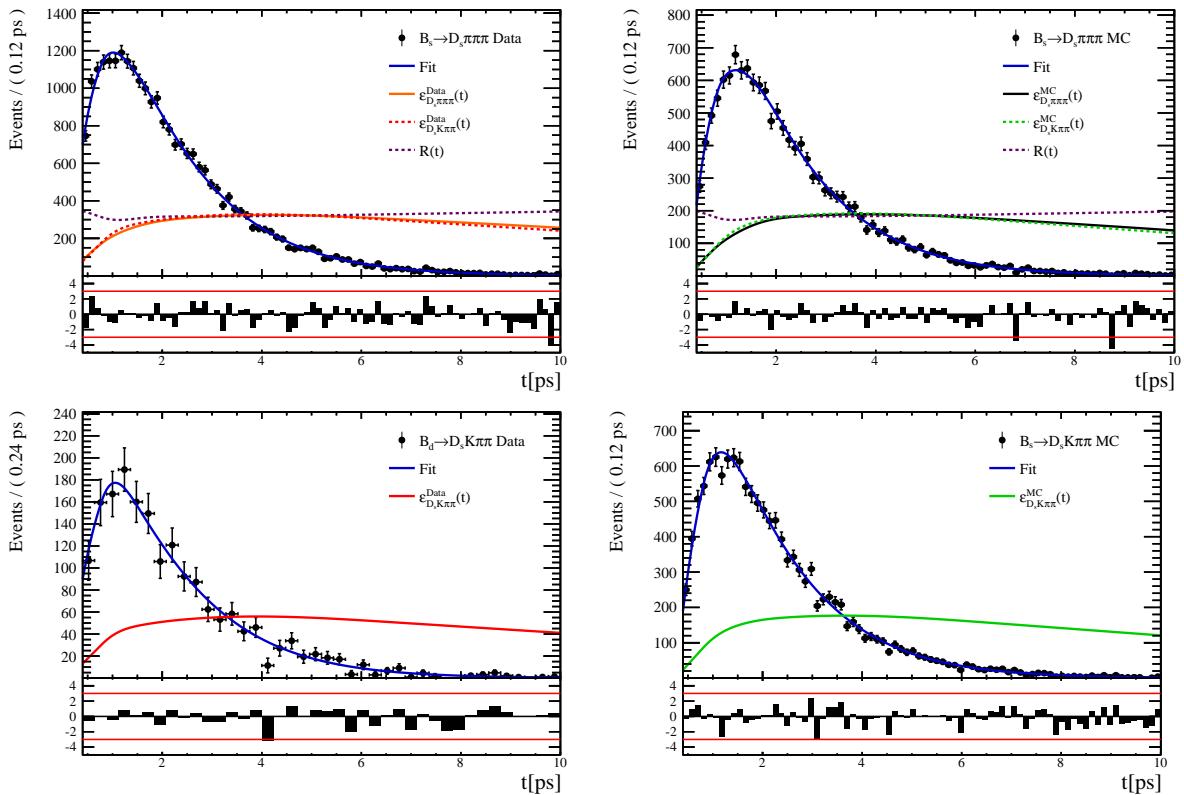


Figure 6.10

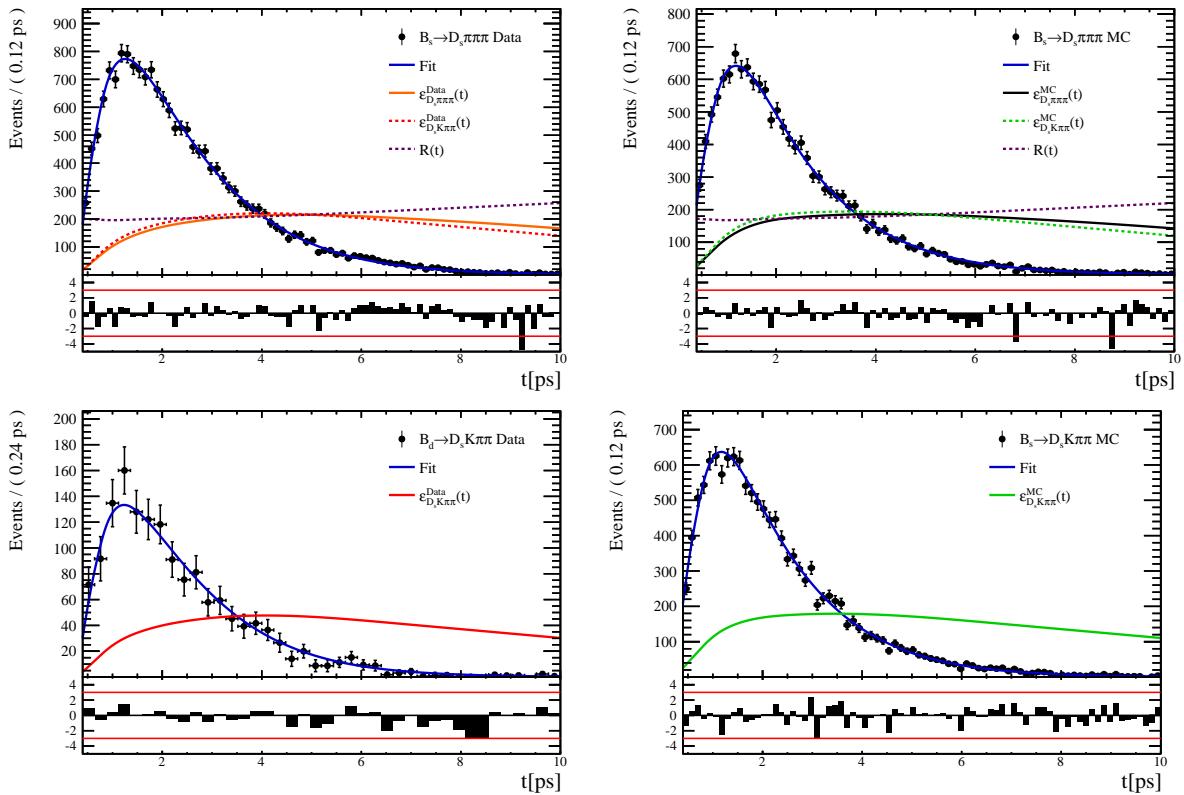


Figure 6.11:

601 **6.3 Phasespace acceptance**

## 602 7 Flavour Tagging

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

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

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

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

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

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

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

630 Utilizing flavour-specific final states, the predicted mistag  $\eta$  of each tagger has to be  
 631 calibrated to match the observed mistag  $\omega$  on the data sample. For the calibration, a  
 632 linear model of the form

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - \langle \eta \rangle), \quad (7.3)$$

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

642 The tagging calibration is determined using a time-dependent fit to the full  $B_s^0 \rightarrow D_s\pi\pi\pi$   
 643 sample, where the mixing frequency  $\Delta m_s$  is fixed to the nominal PDG value [32]. The  
 644 calibration procedure for the OS tagging algorithms (Sec.7.1) and the SS kaon tagger  
 645 (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,  
 646 which is selected following the steps described in Sec. 3. The similar selection ensures  
 647 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  
 648 terms of the decay kinematics, which are crucial for the flavour tagging. Section 7.3 shows  
 649 the compatibility of both samples. After applying the calibration, the response of the OS  
 650 and SS taggers are combined, which is shown in Sec. 7.4.

## 651 7.1 OS tagging calibration

652 The responses of the OS electron, muon, neural net kaon and the secondary vertex charge  
 653 taggers are combined for the mistag calibration. Figure ?? shows the distribution of the  
 654 predicted OS mistag for signal candidates from  $B_s^0 \rightarrow D_s\pi\pi\pi$ . The extracted calibration  
 655 parameters and tagging asymmetries are summarized in Table 7.1 and the measured  
 656 tagging power for the OS combination is  $\epsilon_{eff,OS} = 4.81\%$ .

$p_0$	$p_1$	$\langle \eta \rangle$	$\epsilon_{tag}$	$\Delta p_o$	$\Delta p_1$	$\epsilon_{eff} [\%]$
$0.025 \pm 0.005$	$0.944 \pm 0.048$	$0.347$	$0.517 \pm 0.002$	$0.028 \pm 0.005$	$0.037 \pm 0.045$	$4.81 \pm 0.04 (\text{stat}) \pm 0.37 (\text{cal})$

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

## 657 7.2 SS tagging calibration

658 The SS neural net kaon tagger can be calibrated using the flavour-specific  $B_s^0 \rightarrow D_s\pi\pi\pi$   
 659 decay. Its development, performance and calibration is described in detail in [37]. Figure  
 660 ?? shows the distribution of the predicted mistag of the neural net kaon tagger. The  
 661 extracted calibration parameters and tagging asymmetries are summarized in Table 7.2  
 662 and the measured tagging power for this algorithm is  $\epsilon_{eff,SS} = 3.22\%$ .

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

Table 7.2: Calibration parameters and tagging asymmetries of the SS tagger extracted from  $B_s^0 \rightarrow D_s\pi\pi\pi$  decays.

## 663 7.3 Tagging performance comparison between the signal and 664 normalization channel

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

670 and data.

671 The distributions of the predicted mistag probability  $\eta$  for the OS combination and the  
672 SS kaon tagger are shown in Fig. ?? (simulation) and Fig. 7.1 (data).

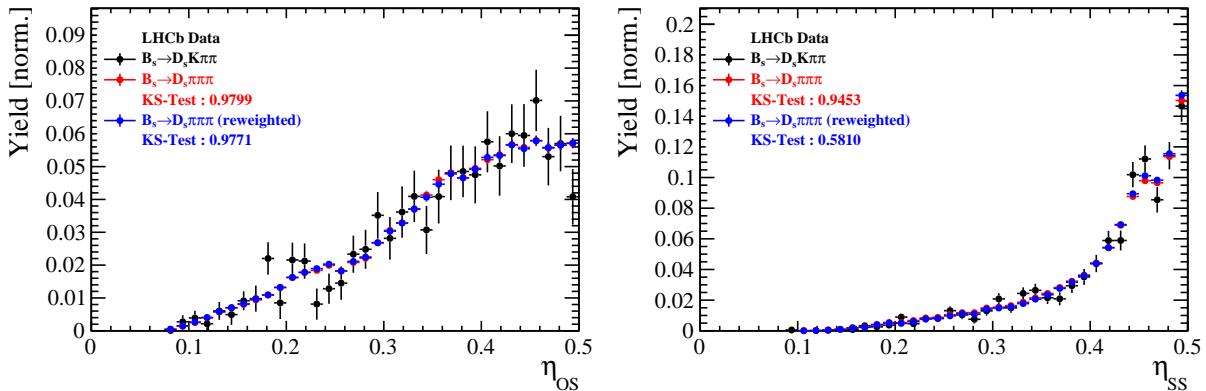


Figure 7.1: Distributions of the predicted mistag  $\eta$  for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the  $B_s^0 \rightarrow D_s K\pi\pi$  (black) and  $B_s^0 \rightarrow D_s \pi\pi\pi$  (red) data samples.

673 Both, data and simulated samples, show good agreement between the signal and  
674 normalization channel. Compatibility is also seen in Fig. ??, which shows the comparison  
675 of the tagging decision distributions of the OS and SS tagger for sweighted data.

676 Fig. ?? shows the signal data distributions of the transverse  $B_s^0$  momentum  $p_T$ , the  
677 pseudorapidity  $\eta$  of the signal candidate and the number of reconstructed tracks per event.  
678 Sufficient agreement is observed.

679 To justify the portability of the flavour tagging calibration obtained from  $B_s^0 \rightarrow D_s \pi\pi\pi$   
680 to the  $B_s^0 \rightarrow D_s K\pi\pi$  channel, besides the good agreement of the distributions shown  
681 above, the dependence of the measured mistag  $\omega$  on the predicted mistag  $\eta$  has to be  
682 compatible in both channel. This dependence is shown in Fig. 7.2 for simulated signal  
683 events of both channels, where good agreement is observed.

## 684 7.4 Combination of OS and SS taggers

685 In the time- and amplitude-dependent fit to  $B_s^0 \rightarrow D_s K\pi\pi$  data, the obtained tagging  
686 responses of the OS and SS tagger will be combined after the calibration described in the  
687 previous sections is applied. Events that aquire a mistag probability greater than 0.5 after  
688 the calibration will have their tagging decision flipped. For events where only one of the  
689 two taggers fired, the combination of the tagging decision is trivial. In those events where  
690 both taggers made a decision, we use the standard combination of taggers [38] provided  
691 by the flavour tagging group. In the nominal fit, the calibrated mistags  $\omega$  are combined  
692 event by event for the OS and SS tager, thus adding one variable to observable to the  
693 fit procedure. This ensures that the uncertainties of the OS and SS tagging calibration  
694 parameters are propagated properly to the combined tagging response for each event.  
695 The taggging performance for the combined tagger in the categories SS tagged only, OS  
696 tagged only and SS+OS tagged, are shown in Tab. ?? for the signal and normalization  
697 channel. The distribution of the observed mistag  $\omega$  as a function of the combined mistag  
698 probability  $\eta$  for  $B_s^0 \rightarrow D_s \pi\pi\pi$  decays is shown in Fig. ??.

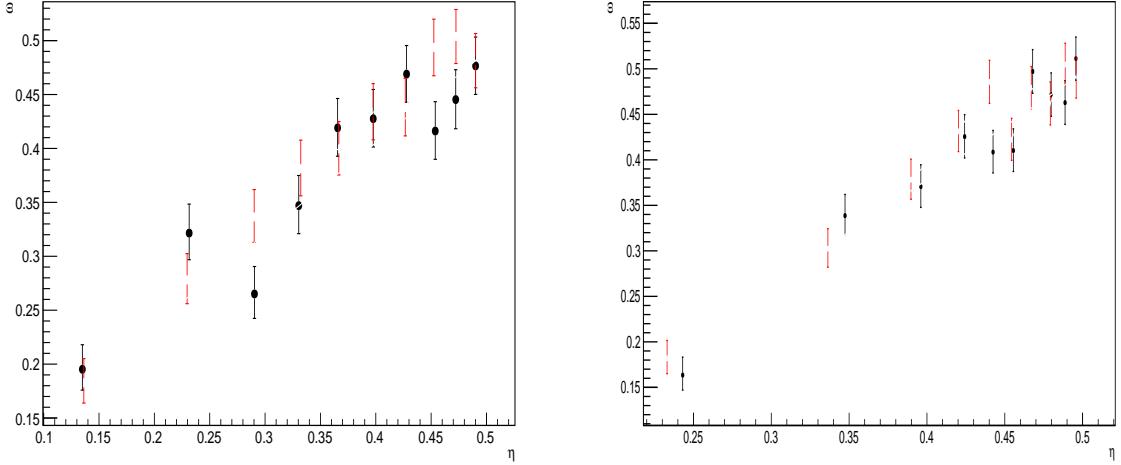


Figure 7.2: Dependence of the observed mistag  $\omega$  on the predicted mistag  $\eta$  for the (left) OS combination and ther (right) SS kaon tagger, found in the simulated  $B_s^0 \rightarrow D_s K\pi\pi$  (black) and  $B_s^0 \rightarrow D_s \pi\pi\pi$  (red) signal samples.

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

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
Only OS	$11.32 \pm 0.09$	$37.91 \pm 1.02$	$0.98 \pm 0.14$
Only SS	$41.66 \pm 0.18$	$43.78 \pm 0.53$	$1.54 \pm 0.23$
Both OS-SS	$27.17 \pm 0.25$	$36.68 \pm 0.81$	$2.91 \pm 0.29$
Combined	$80.15 \pm 0.32$	$40.55 \pm 0.72$	$5.43 \pm 0.40$

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

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
Only OS	$10.51 \pm 0.07$	$35.32 \pm 0.77$	$1.25 \pm 0.11$
Only SS	$43.27 \pm 0.14$	$43.29 \pm 0.44$	$1.58 \pm 0.17$
Both OS-SS	$24.77 \pm 0.18$	$35.14 \pm 0.61$	$3.19 \pm 0.22$
Combined	$78.55 \pm 0.24$	$39.65 \pm 0.55$	$6.02 \pm 0.30$

## 699 8 Production and Detection Asymmetries

### 700 8.1 $B_s$ Production Asymmetry

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

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

704 where  $\sigma$  are the corresponding production cross-section. This asymmetry was measured  
 705 by LHCb in  $pp$  collisions at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV by means of a time-dependent  
 706 analysis of  $B_s \rightarrow D_s^- \pi^+$  decays [39]. The results in bins of  $p_T$  and  $\eta$  of the  $B_s$  meson  
 707 are shown in Table 8.1. To correct for the different kinematics of  $B_s \rightarrow D_s^- \pi^+$  and  
 708  $B_s^0 \rightarrow D_s K \pi \pi$  decays, the measured  $B_s$  production asymmetries  $A_p(p_T, \eta)$  are folded with  
 709 the sWeighted  $p_T, \eta$  distribution of our signal channel. The resulting effective production  
 710 asymmetries are:

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

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

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

711 As for Run-II data no measurement is available yet, we determine the production asym-  
 712 metry 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. [39]

$p_T$ [ GeV/c ]	$\eta$	$A_p(B_s^0)_{\sqrt{s}=7 \text{ TeV}}$	$A_p(B_s^0)_{\sqrt{s}=8 \text{ TeV}}$
(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)	$-0.0586 \pm 0.0648 \pm 0.0150$	$0.0537 \pm 0.0447 \pm 0.0124$

## 713 8.2 $K^-\pi^+$ Detection Asymmetry

714 The presented measurement of the CKM-angle  $\gamma$  using  $B_s^0 \rightarrow D_s K\pi\pi$  decays is sensitive  
 715 to a possible charge asymmetry of the kaon. This effect can be detector induced, because  
 716 kaons are known to have a nuclear cross-section which is asymmetrically dependent on  
 717 the sign of their charge. It is indispensable to determine the detector induced charge  
 718 asymmetry of the kaon, as fitting without taking this effect into account would introduce  
 719 a ‘fake’ CP violation. Instead of determining the single track detection asymmetry of a  
 720 kaon, it is found that the combined two track asymmetry of a kaon-pion pair is much  
 721 easier to access [40]. Therefore the two track asymmetry is used, which is defined as

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

722 This asymmetry can be measured from the difference in asymmetries in the  $D^+ \rightarrow K^-\pi^+\pi^+$   
 723 and  $D^+ \rightarrow K_s^0\pi^+$  modes [41]:

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

724 where possible CP violation in the  $D^+ \rightarrow K_s^0\pi^+$  mode is predicted to be smaller than  
 725  $10^{-4}$  in the Standard Model [42]. The asymmetry in the neutral kaon system,  $A(K^0)$ , has  
 726 to be taken into account as a correction.

727 We use a dedicated LHCb tool to determine  $A^{det}(K^-\pi^+)$  for all data taking periods  
 728 used in this analysis. A detailed description can be found in [41]. The tool provides  
 729 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  
 730 determine the asymmetry following Eq. 8.6. Several weighting steps are performed to  
 731 match the kinematics of the calibration samples to our signal decay sample:

732 First, weights are assigned to the  $K^\pm$  and  $\pi^\pm$  of the  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  sample, using  
 733  $p, \eta$  of the  $K^\pm$  and  $p_T, \eta$  of the  $\pi^\pm$  from our  $B_s^0 \rightarrow D_s K\pi\pi$  signal decay. Then, weights  
 734 are assigned to the  $D^\pm (p_T, \eta)$  and the  $\pi^\pm (p_T)$  of the  $D^\pm \rightarrow K_s^0\pi^\pm$  sample to match  
 735 the corresponding, weighted distributions of the  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  sample. In a last  
 736 step, weights are assigned to match the bachelor pions  $\phi$  distributions between the two  
 737 calibration samples.

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

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

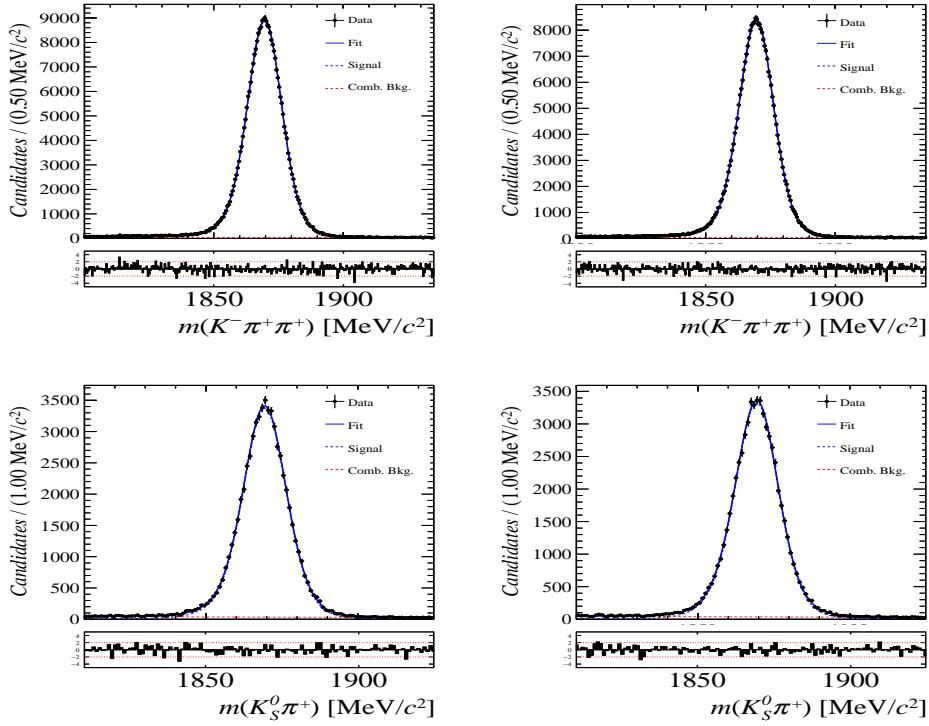


Figure 8.1: Distributions of the invariant mass of (top)  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  and (bottom)  $D^\pm \rightarrow K_s^0\pi^\pm$  candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

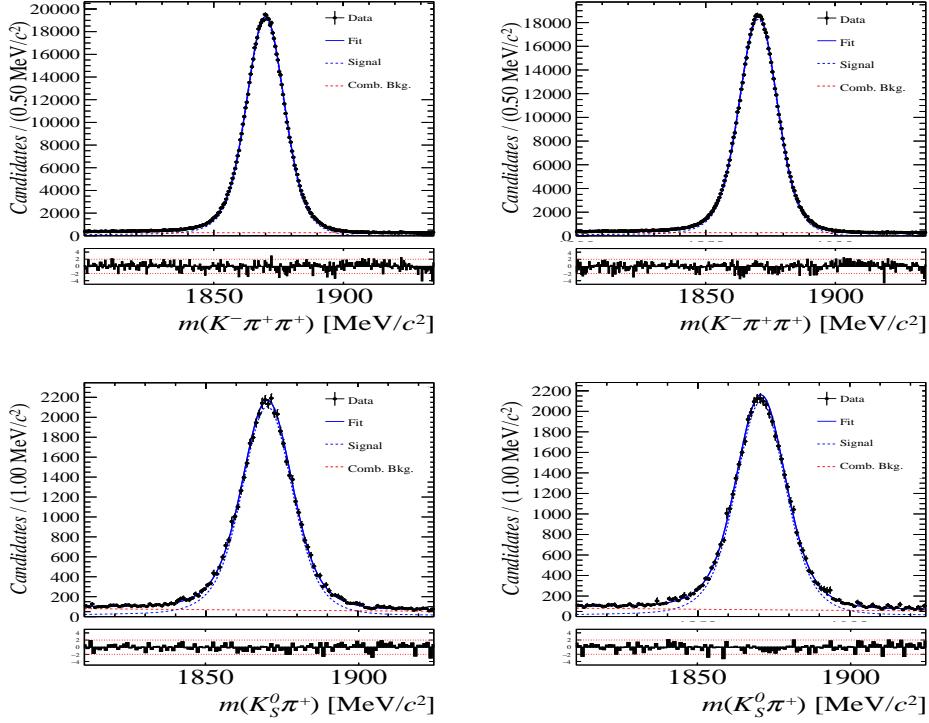


Figure 8.2: Distributions of the invariant mass of (top)  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  and (bottom)  $D^\pm \rightarrow K_s^0\pi^\pm$  candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

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.

## 749 9 Time dependent fit

750 This section covers the phasespace integrated, time-dependent fit to  $B_s^0 \rightarrow D_s h\pi\pi$  data.  
 751 We use the **sFit** technique [43] to statistically remove background from the decay time fit,  
 752 leaving only the signal PDF to describe the decay time. The **sWeights** are calculated based  
 753 on the fit to the reconstructed  $B_s$  mass distribution described in Sec. 4. As additional  
 754 input to the fit, the tagging information (Sec. 7), as well as the decay time acceptance  
 755 (Sec. 6) and resolution (Sec. 5) is used and fixed to the values obtained by the dedicated  
 756 studies. Taking all inputs into account, the final time dependent fit PDF is given by

$$757 \quad \mathcal{P}\mathcal{D}\mathcal{F}(t, \vec{\lambda}) = \left( \epsilon(t) \cdot \int P(x, t, q_t, q_f) dx \right) \times \mathcal{R}(t - t'), \quad (9.1)$$

757 where  $\int P(x, t, q_t, q_f) dx$  is the PDF given by Eq. 2.6,  $\epsilon(t)$  is the efficiency due to the time  
 758 acceptance effects and  $\mathcal{R}(t - t')$  is the Gaussian time resolution function.

### 759 9.1 sFit to $B_s^0 \rightarrow D_s \pi\pi\pi$ data

760 The phasespace-integrated, time-dependent fit is performed to the full sWeighted sample  
 761 of selected candidates from Run I and 2015+2016 Run II data, containing both possible  
 762 magnet polarities and  $D_s$  final states. In the fit, the values of  $\Gamma_s$  and  $\Delta\Gamma_s$  are fixed to the  
 763 latest PDG report. All tagging parameters are fixed to the central values found in the  
 764 tagging calibration, described in Sec. 7. Due to the fact that the  $B_s^0 \rightarrow D_s \pi\pi\pi$  decay is  
 765 flavour specific, the CP-coefficients can be fixed to  $C = 1$  and  $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$ ,  
 766 reducing Eq. 2.6 to

$$767 \quad \int P(x, t, q_t, q_f) dx = [\cosh\left(\frac{\Delta\Gamma t}{2}\right) + q_t q_f C \cos(m_s t)] e^{-\Gamma t}. \quad (9.2)$$

767 Note that in this case, the dependence on the coherence factor  $\kappa$  is dropped and the  
 768 same relation as found for  $B_s^0 \rightarrow D_s \pi$  decays is recovered. Therefore, the only free fit  
 769 parameter left is  $\Delta m_s$ . The data distribution with the overlaid fit is shown in Fig.

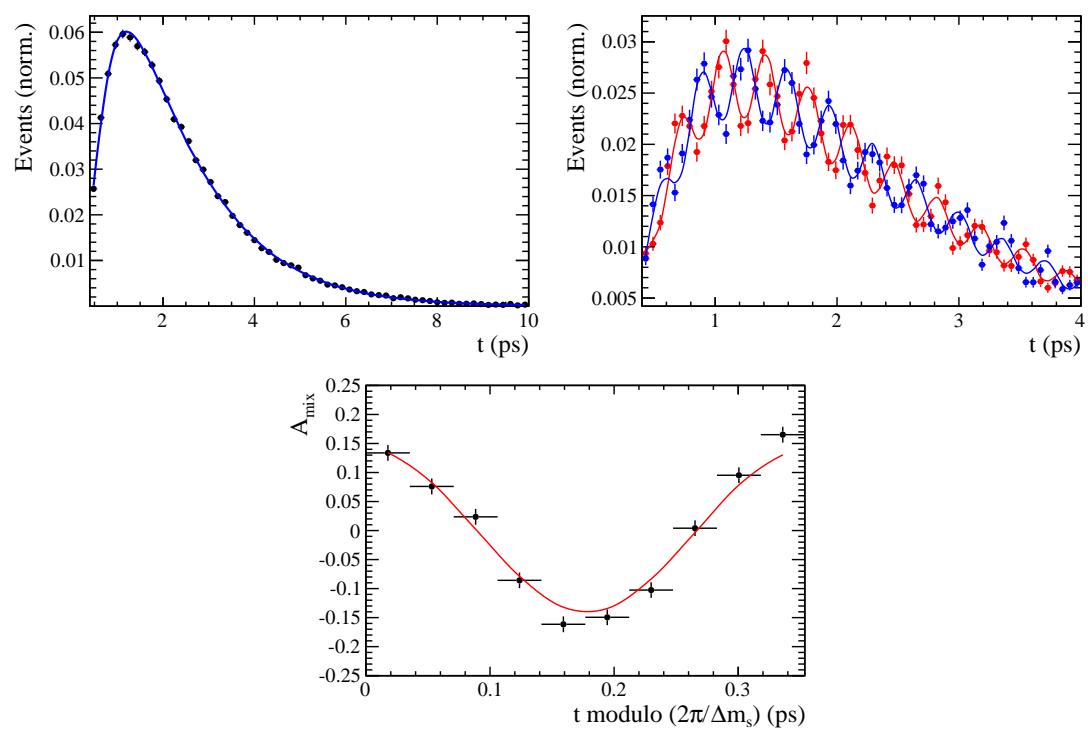


Figure 9.1

Table 9.1: Result of the phase-space integrated fit to  $B_s \rightarrow D_s \pi\pi\pi$  data.

	Fit parameter	Value
Run-I	$p_0^{\text{OS}}$	$0.3896 \pm 0.0101$
	$p_1^{\text{OS}}$	$0.8883 \pm 0.1074$
	$\Delta p_0^{\text{OS}}$	$0.0161 \pm 0.0104$
	$\Delta p_1^{\text{OS}}$	$0.0005 \pm 0.1095$
	$\epsilon_{tag}^{\text{OS}}$	$0.3851 \pm 0.0031$
	$\Delta \epsilon_{tag}^{\text{OS}}$	$0.0069 \pm 0.0123$
	$p_0^{\text{SS}}$	$0.4465 \pm 0.0075$
	$p_1^{\text{SS}}$	$1.0748 \pm 0.1012$
	$\Delta p_0^{\text{SS}}$	$-0.0190 \pm 0.0076$
	$\Delta p_1^{\text{SS}}$	$0.1017 \pm 0.1063$
	$\epsilon_{tag}^{\text{SS}}$	$0.6882 \pm 0.0029$
	$\Delta \epsilon_{tag}^{\text{SS}}$	$-0.0076 \pm 0.0117$
	$A_p$	$-0.0004 \pm 0.0000$
Run-II	$p_0^{\text{OS}}$	$0.3669 \pm 0.0074$
	$p_1^{\text{OS}}$	$0.9298 \pm 0.0761$
	$\Delta p_0^{\text{OS}}$	$0.0118 \pm 0.0085$
	$\Delta p_1^{\text{OS}}$	$0.0234 \pm 0.0855$
	$\epsilon_{tag}^{\text{OS}}$	$0.3525 \pm 0.0023$
	$\Delta \epsilon_{tag}^{\text{OS}}$	$0.0105 \pm 0.0085$
	$p_0^{\text{SS}}$	$0.4532 \pm 0.0055$
	$p_1^{\text{SS}}$	$0.9125 \pm 0.0656$
	$\Delta p_0^{\text{SS}}$	$-0.0123 \pm 0.0060$
	$\Delta p_1^{\text{SS}}$	$0.1374 \pm 0.0757$
	$\epsilon_{tag}^{\text{SS}}$	$0.6804 \pm 0.0023$
	$\Delta \epsilon_{tag}^{\text{SS}}$	$0.0076 \pm 0.0083$
	$A_p$	$-0.0042 \pm 0.0091$
$\Delta m_s$		$\text{xx.xx} \pm 0.0110$

770 **9.2 sFit to  $B_s^0 \rightarrow D_s K\pi\pi$  data**

Table 9.2: Result of the phase-space integrated fit to  $B_s \rightarrow D_s K\pi\pi$  data.

Fit parameter	Value
$C$	xx.xx ± 0.165
$D$	xx.xx ± 0.359
$\bar{D}$	xx.xx ± 0.333
$S$	xx.xx ± 0.248
$\bar{S}$	xx.xx ± 0.218

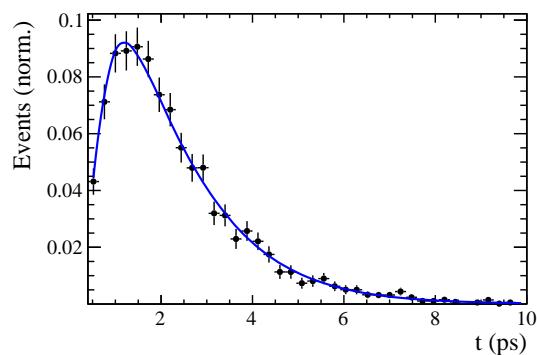


Figure 9.2

## 771 10 Time dependent amplitude fit

### 772 10.1 Signal Model Construction

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

779 To build the amplitude model, one could successively add amplitudes on top of one  
780 another until a reasonable agreement between data and fit was achieved. However, this  
781 step-wise approach is not particularly suitable for amplitude analyses as discussed in  
782 Ref. [44]. Instead, we include the whole pool of amplitudes in the first instance and use  
783 the Least Absolute Shrinkage and Selection Operator [44, 45] (LASSO) approach to limit  
784 the model complexity. In this method, the event likelihood is extended by a penalty term

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

785 which shrinks the amplitude coefficients towards zero. The amount of shrinkage is  
786 controlled by the parameter  $\lambda$ , to be tuned on data. Higher values for  $\lambda$  encourage sparse  
787 models, *i.e.* models with only a few non-zero amplitude coefficients. The optimal value  
788 for  $\lambda$  is found by minimizing the Bayesian information criteria [46] (BIC),

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

789 where  $N_{\text{Sig}}$  is the number of signal events and  $r$  is the number of amplitudes with a decay  
790 fraction above a certain threshold. In this way, the optimal  $\lambda$  balances the fit quality  
791 ( $-2 \log \mathcal{L}$ ) against the model complexity. The LASSO penalty term is only used to select  
792 the model. Afterwards, this term must be discarded in the final amplitude fit with the  
793 selected model, otherwise the parameter uncertainties would be biased.

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

- 798 1. The fit fraction threshold for inclusion in the final model was varied within the  
799 interval [0.05, 5]%. The set of selected amplitudes is stable for thresholds between  
800 0.1% and 1%. Other choices result in marginally different models containing one  
801 component more or less.
  - 802 2. Instead of BIC, the Akaike information criteria ( $\text{AIC}(\lambda) = -2 \log \mathcal{L} + 2r$  [47]) was  
803 used to optimize  $\lambda$ . For a given threshold, the AIC method tends to prefer lower  
804  $\lambda$  values. However, the set of models obtained varying the threshold within the  
805 interval [0.05, 5]% is identical to the BIC method.
  - 806 3. The amplitudes selected under nominal conditions were excluded one-by-one from  
807 the set of all amplitudes considered.
- 808 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 [%]
$B_s \rightarrow K(1)(1400)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	$34.70 \pm 2.24$
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K(0)^*(1430)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	$6.85 \pm 0.94$
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	$13.08 \pm 1.70$
$B_s \rightarrow K(1)(1270)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+\pi^-) K^+) D_s^-$	$9.25 \pm 0.60$
$B_s \rightarrow K(1460)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	$0.99 \pm 0.06$
$BuggB_s \rightarrow K(1460)^+ (\rightarrow \sigma (\rightarrow \pi^+\pi^-) K^+) D_s^-$	$3.42 \pm 1.49$
$B_s \rightarrow K^*(1410)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+\pi^-)\pi^+) D_s^-$	$16.40 \pm 1.06$
$B_s \rightarrow K^*(1410)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+\pi^-) K^+) D_s^-$	$4.88 \pm 0.68$
$B_s \rightarrow NonResS0 (\rightarrow D_s^- \pi^+) K^*(892)^0 (\rightarrow K^+\pi^-)$	$4.60 \pm 1.44$
$BuggB_s \rightarrow NonResS0 (\rightarrow D_s^- K^+) \sigma (\rightarrow \pi^+\pi^-)$	$4.96 \pm 0.68$
Sum	$99.13 \pm 5.87$

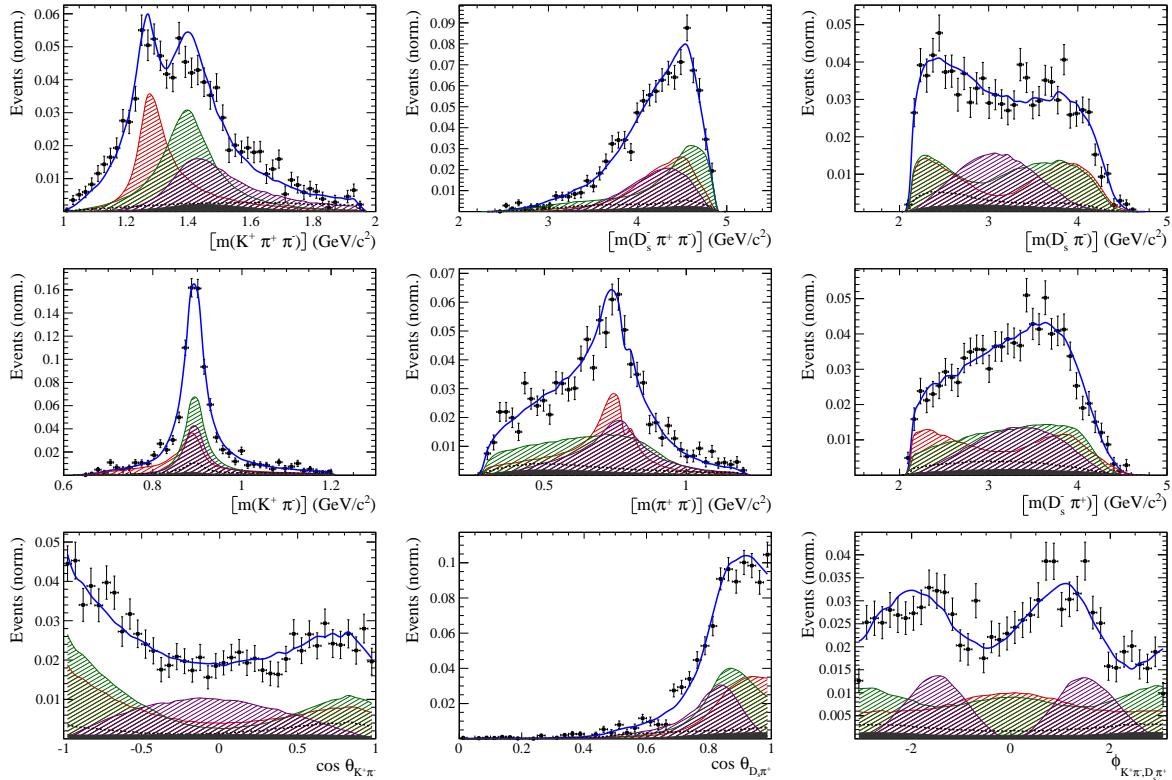


Figure 10.1

## 809 10.2 Results

Table 10.2: Result of the time-dependent amplitude fit to  $B_s \rightarrow D_s K\pi\pi$  data.

Fit parameter	Value
$x_-$	xx.xx $\pm$ 0.119
$y_-$	xx.xx $\pm$ 0.044
$x_+$	xx.xx $\pm$ 0.060
$y_+$	xx.xx $\pm$ 0.038

Table 10.3: Fit fractions for  $B_s \rightarrow D_s K\pi\pi$  data.

Decay channel	Fraction [%]
$B_s \rightarrow K(1)(1400)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	28.77 $\pm$ 0.20
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K(0)^*(1430)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	9.32 $\pm$ 0.06
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	18.13 $\pm$ 0.12
$B_s \rightarrow K(1)(1270)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	12.80 $\pm$ 0.09
$B_s \rightarrow K^*(1410)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	19.78 $\pm$ 0.14
$B_s \rightarrow K^*(1410)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	5.98 $\pm$ 0.04
$BuggB_s \rightarrow NonResS0 (\rightarrow D_s^- K^+) \sigma (\rightarrow \pi^+ \pi^-)$	1.62 $\pm$ 0.73
Sum	96.40 $\pm$ 0.14

Table 10.4: Fit fractions for  $B_s \rightarrow D_s K\pi\pi$  data.

Decay channel	Fraction [%]
$B_s \rightarrow K(1)(1400)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	93.96 $\pm$ 11.84
$B_s \rightarrow K(1460)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	2.55 $\pm$ 0.32
$BuggB_s \rightarrow K(1460)^+ (\rightarrow \sigma (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	8.75 $\pm$ 1.10
$B_s \rightarrow NonResS0 (\rightarrow D_s^- \pi^+) K^*(892)^0 (\rightarrow K^+ \pi^-)$	58.07 $\pm$ 17.15
Sum	163.34 $\pm$ 12.39

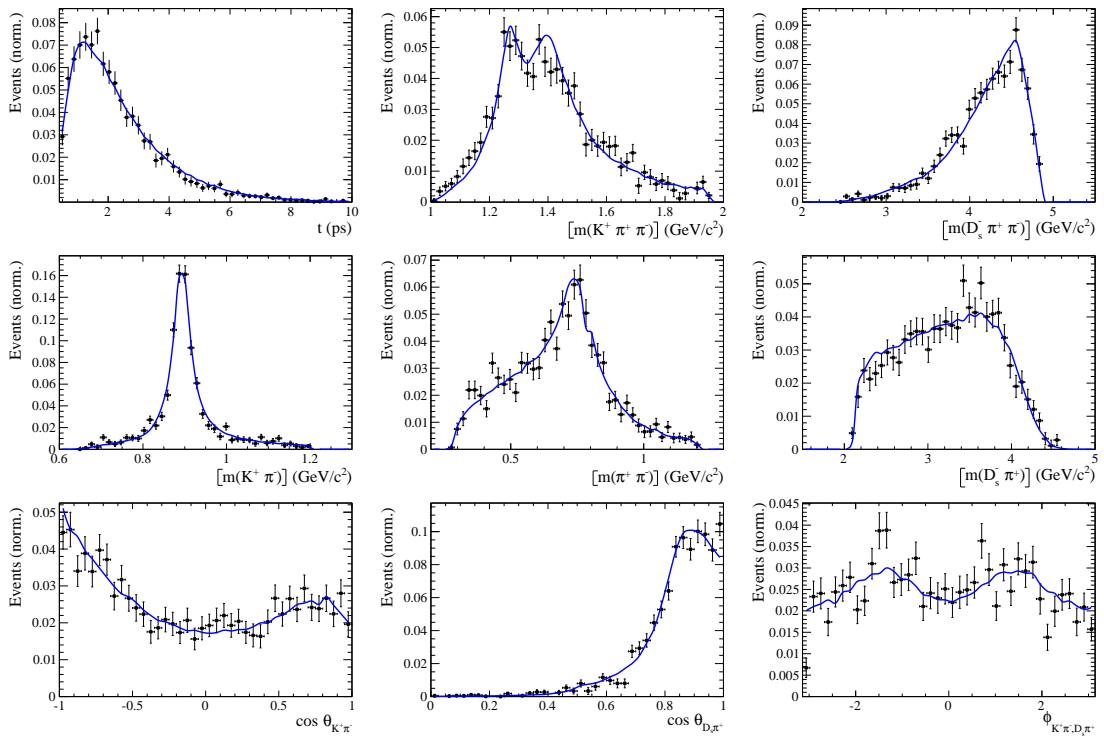


Figure 10.2

## 810 11 Systematic uncertainties

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

### 818 11.1 Models for $B_s^0$ mass distribution

819 The statistical subtraction of the residual background [43], left after the full selection,  
 820 relies on the correct description of the invariant  $B_s^0$  mass distribution. Since the choice  
 821 of signal and background models is not unique, alternative descriptions which lead to  
 822 slightly different yields for the signal and background components are available. The  
 823 difference in yields could result in shifted values for the measured observables and are  
 824 therefore treated as systematic uncertainty.

825

#### 826 11.1.1 Signal model

827 The Johnson's SU function which is used as nominal signal model is replaced by a double  
 828 Crystal Ball [48]. The crystal ball model is given by a gaussian core with an exponential  
 829 tail on one side. Choosing a double Crystal Ball allows for asymmetric tails in a slightly  
 830 different way compared to the Johnson's SU function. Table xXx summarizes the observed  
 831 differences in signal and background yields.

#### 832 11.1.2 Background model

833 For the description of the partially reconstructed background, a combination of the  
 834 RooHILLdini and RooHORNsdini model [REF HERE] is used instead of the nominal  
 835 model of three bifurcated gaussians. The HORNsdini model is used to describe the  
 836  $B_s^0 \rightarrow D_s^* [\rightarrow D_s(\pi^0)] X_{s/d}$  decay, where the brackets around the  $\pi^0$  indicate that it  
 837 is missed in the reconstruction. The  $D_s^* \rightarrow D_s\pi^0$  decay is a Vector  $\rightarrow$  Scalar-Scalar  
 838 ( $1^- \rightarrow 0^-0^-$ ) transition. Using the helicity of the  $D_s$ , one can show that this results in a  
 839 double-peak structure in the reconstructed  $B_s^0$  mass. Therefore, the HORNsdini shape  
 840 consists of a gaussian-like double-peak structure:

$$841 \text{HORN}(m_{B_s^0}) = \int_a^b dm_{B_s^0} \left( m_{B_s^0} - \frac{a+b}{2} \right)^2 \mathcal{D}\mathcal{G}(m_{B_s^0} | \mu, \sigma, f_G) \left( \frac{1-\zeta}{b-a} m_{B_s^0} + \frac{b\zeta-a}{b-a} \right), \quad (11.1)$$

842 where  $a$  and  $b$  are the kinematic endpoints of the distribution and  $\zeta$  is the positive,  
 843 real fraction of the two peak heights. Additionally, the shape is convoluted with a gaussian  
 844 to account for resolution effects.

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

846 the  $Ds^* \rightarrow D_s\gamma$  is a Vector  $\rightarrow$  Scalar-Vector ( $1^- \rightarrow 0^-1^-$ ) transition. From helicity  
 847 arguments, the expected shape in the mass distribution of  $B_s^0$  candidates follows a  
 848 parabolic curve without any peaking structure. To accommodate for this shape, the  
 849 HILLdini model consists of a parabolic curve between the kinematic endpoints a & b:

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

850 This shape is convoluted with the same gaussian resolution function used for the  
 851 HORNSdini model. The resulting differences in yields is shown in Table xXx.

852  
 853 To study systematic uncertainties originating from the description of the combinatorial  
 854 background, the nominal second order polynomial is replaced by an exponential function.  
 855 The changes in signal and background yields after refitting with this alternative shape are  
 856 shown in Table xXx.

857

### 858 11.1.3 Systematic effect on observables

859 The shift of the central values of the observables in the full fit when using sWeights  
 860 obtained from a combination of alternative models, as well as using only one alternative  
 861 model for the signal/comb.background/part.reco.background and keeping the nominal  
 862 model for the other parts, is shown in Table yYy. We conservatively choose the biggest  
 863 variation as systematic uncertainty from the modelling of the invariant  $B_s^0$  mass spectrum.

## 864 A Stripping and Trigger cuts

865 The following text describes variables which are used in Table 1.1 and might be ambiguous,  
 866 or which benefits are not straight forward. Where noted, different cut values are applied  
 867 for Run-I and Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest  
 868 approach. This variable is used to ensure that all  $D_s$  and  $X_{s,d}$  daughters originate from  
 869 the same vertex. DIRA is the abbreviation for the cosine of the angle  $\theta$  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 $\chi^2/\text{nDoF}$	< 3	
Track $p$	> 1000 MeV/ $c$	
Track $p_T$	> 100 MeV/ $c$	
Track IP $\chi^2$	> 4	
Track ghost-prob.	< 0.4	
$D_s$ mass	$m_{D_s} \pm 100$ MeV	$m_{D_s} \pm 80$ MeV
$D_s$ Daughter $p_T$	$\sum_{i=1}^3 p_{t,i} > 1800$ MeV/ $c$	
$D_s$ Daughter DOCA	< 0.5 mm	
$D_s$ Vertex $\chi^2/\text{nDoF}$	< 10	
$D_s$ $\chi^2$ -separation from PV	> 36	
$D_s$ daughter PID( $\pi$ )	< 20	
$D_s$ daughter PID(K)	> -10	
$X_{s,d}$ mass	< 4000 MeV	< 3500 MeV
$X_{s,d}$ Daughter $p$	> 2 GeV/ $c$	
$X_{s,d}$ Daughter DOCA	< 0.4 mm	
$X_{s,d}$ Daughter $p_T$	$\sum_{i=1}^3 p_{t,i} > 1250$ MeV/ $c$	
$X_{s,d}$ Vertex $\chi^2/\text{nDoF}$	< 8	
$X_{s,d}$ $\chi^2$ -separation from PV	> 16	
$X_{s,d}$ DIRA	> 0.98	
$X_{s,d}$ $\Delta\rho$	> 0.1 mm	
$X_{s,d}$ $\Delta z$	> 2.0 mm	
$X_{s,d}$ daughter PID( $\pi$ )	< 10	
$X_s$ daughter PID(K)	> -2	> 4
$B_s^0$ mass	[4750, 7000] MeV/ $c^2$	[5000, 6000] MeV/ $c^2$
$B_s^0$ DIRA	> 0.98	> 0.99994
$B_s^0$ min IP $\chi^2$	< 25	< 20
$B_s^0$ Vertex $\chi^2/\text{nDoF}$	< 10	< 8
$B_s^0 \tau_{B_s^0}$	> 0.2 ps	

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

876 The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the  
 877 b-hadron  $p_T$ , its flight distance  $\chi^2$  from the nearest PV and the sum of the  $B_s^0$  and  $D_s$   
 878 vertex  $\chi^2$  divided by the sum of their number of degrees of freedom. Table 1.3 summarizes  
 879 the cuts applied by the inclusive  $\phi$  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 $\chi^2$	$> 16$
Track $\chi^2/\text{nDoF}$	$< 2.5$
Track $p_T$	$> 1.7 \text{ GeV}/c$
Track $p$	$> 10 \text{ GeV}/c$
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  $\phi$  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
$\phi$ mass	$m_\phi \pm 12 \text{ MeV}/c^2$ of PDG value
$\phi p_T$	$> 2.5 \text{ GeV}/c$
$\phi$ vertex $\chi^2/\text{nDoF}$	$< 20$
$\phi$ IP $\chi^2$ to any PV	$> 5$

## B Details of multivariate classifier

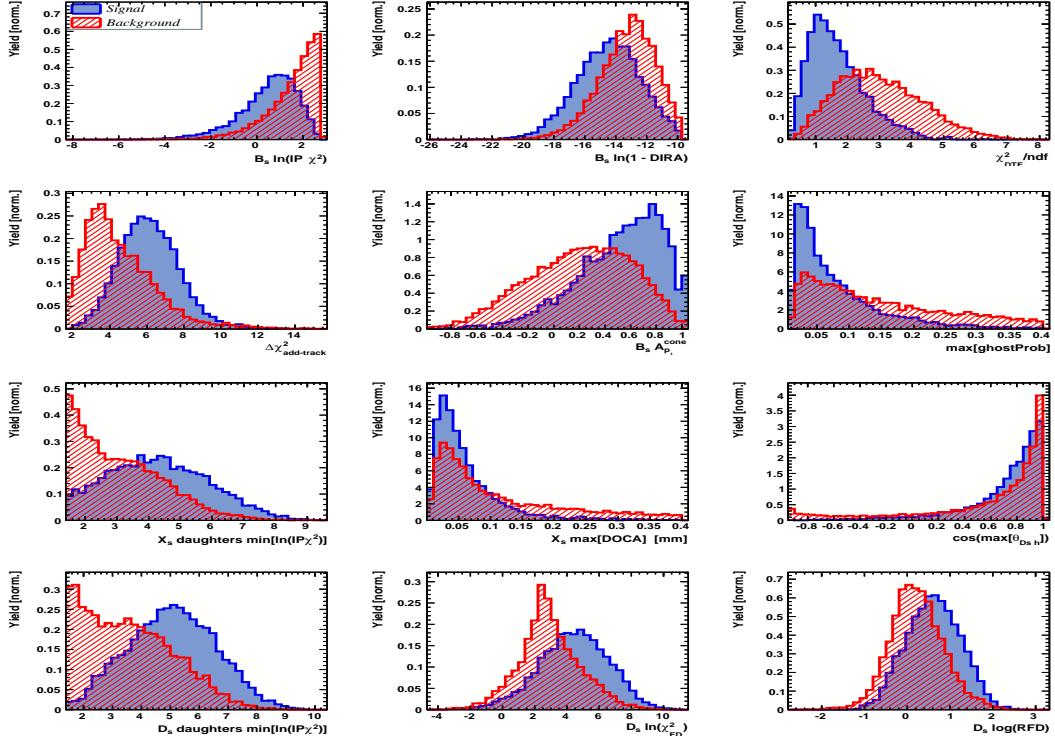


Figure A.1: Variables used to train the BDTG for category [Run-I,L0-TOS].

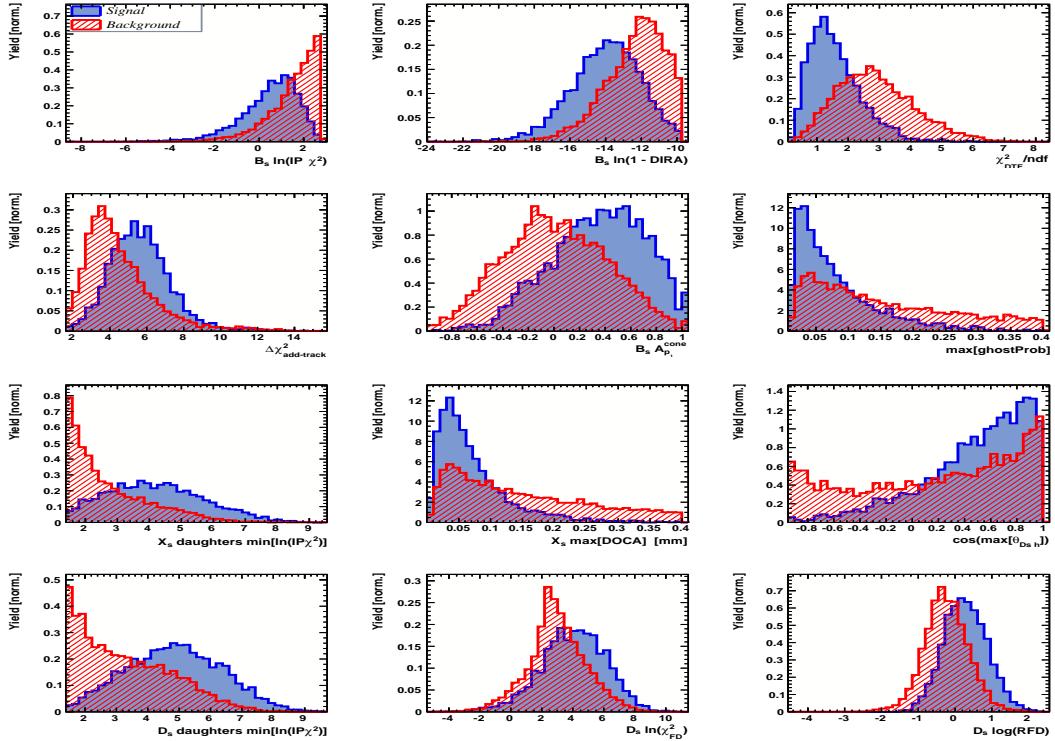


Figure A.2: Variables used to train the BDTG for category [Run-I,L0-TIS].

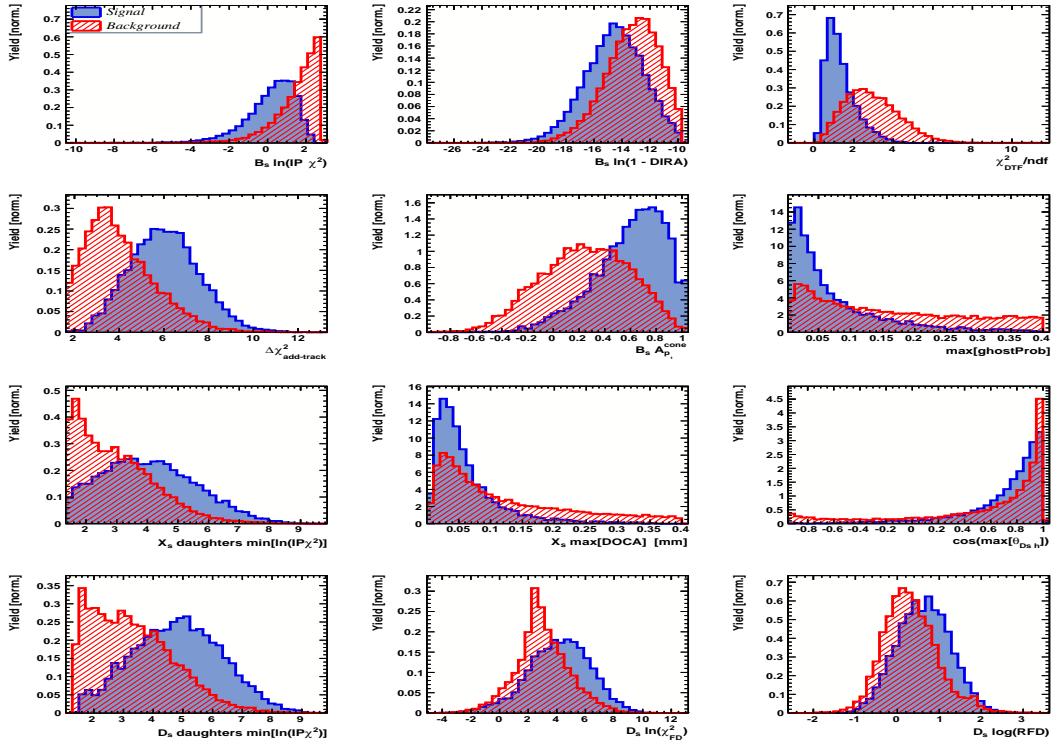


Figure A.3: Variables used to train the BDTG for category [Run-II,L0-TOS].

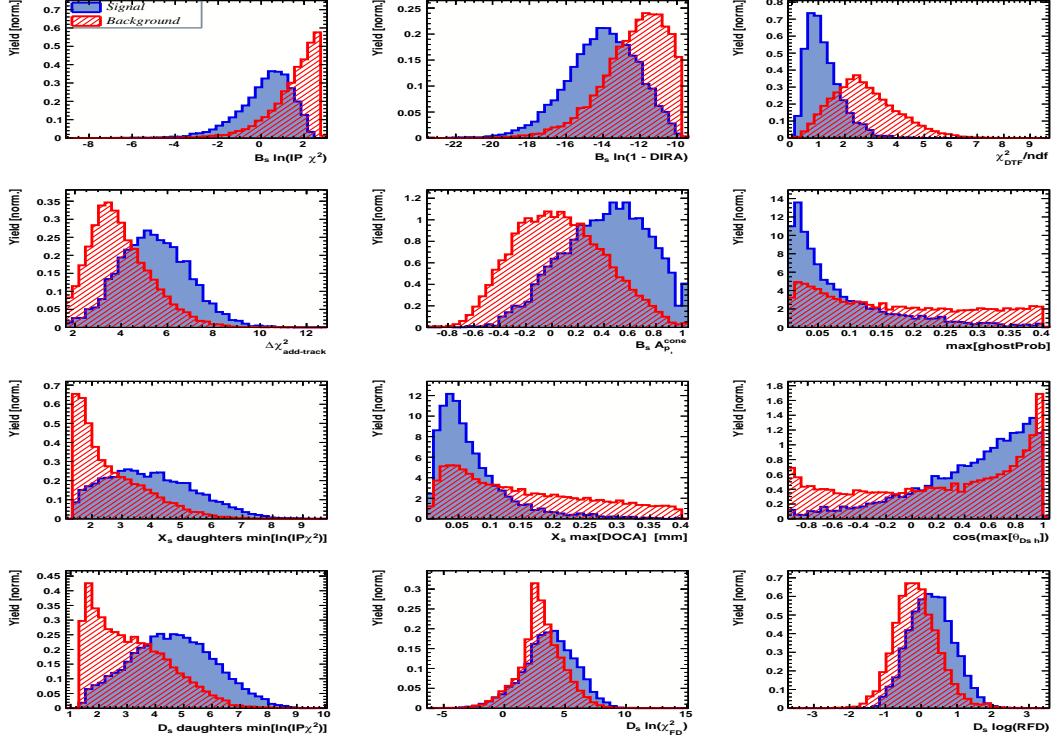


Figure A.4: Variables used to train the BDTG for category [Run-II,L0-TIS].

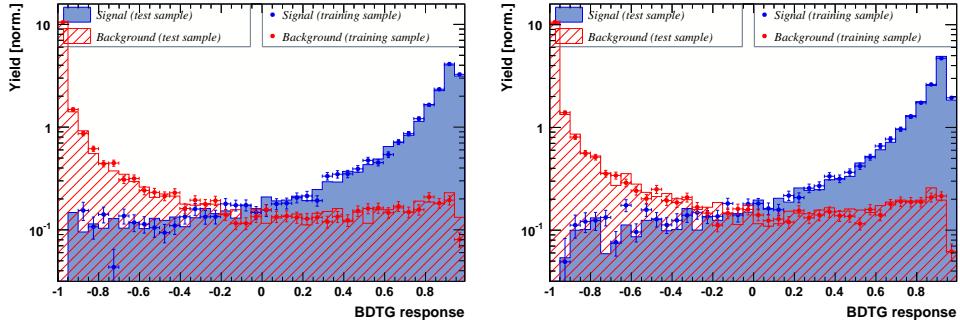


Figure A.5: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].

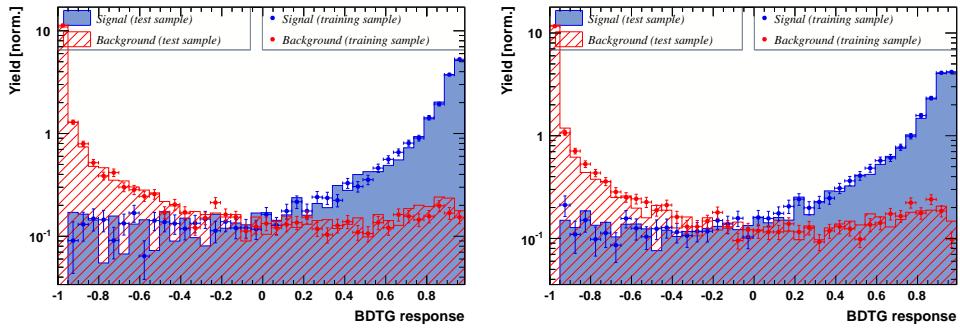


Figure A.6: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].

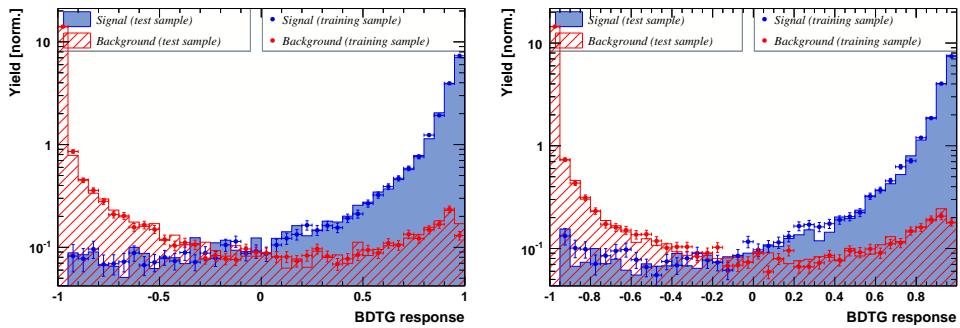


Figure A.7: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].

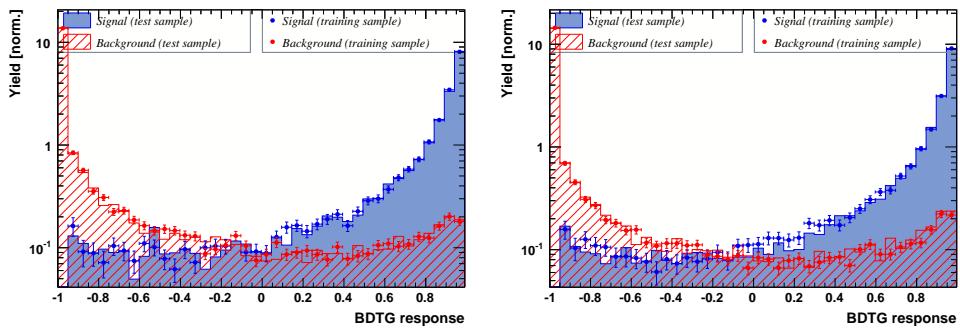


Figure A.8: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].

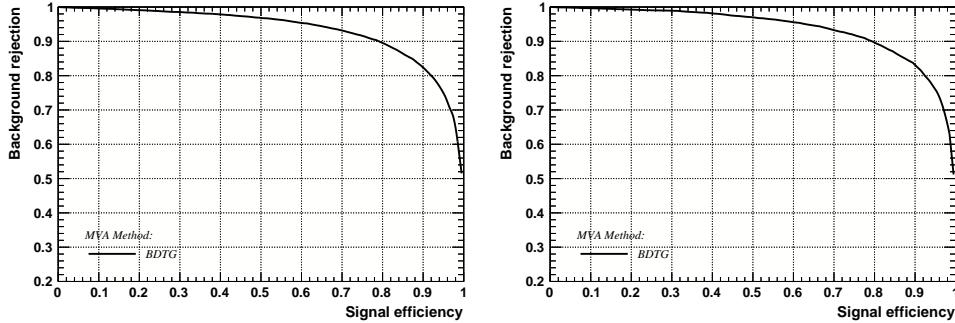


Figure A.9: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].

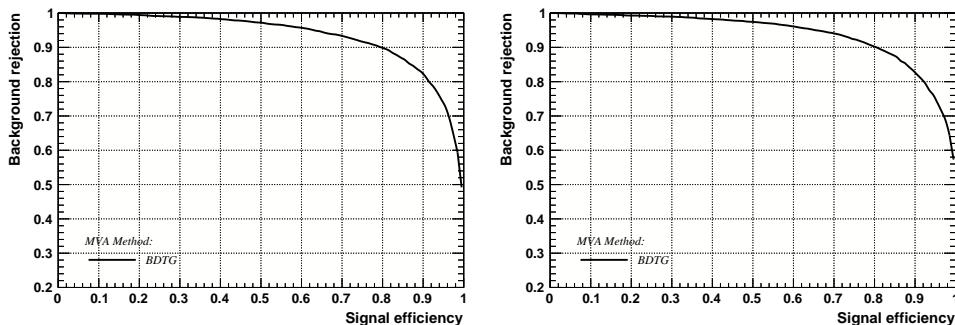


Figure A.10: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TIS].

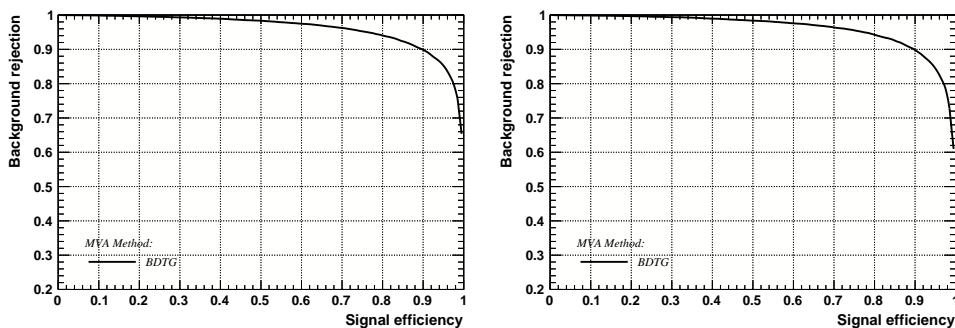


Figure A.11: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TOS].

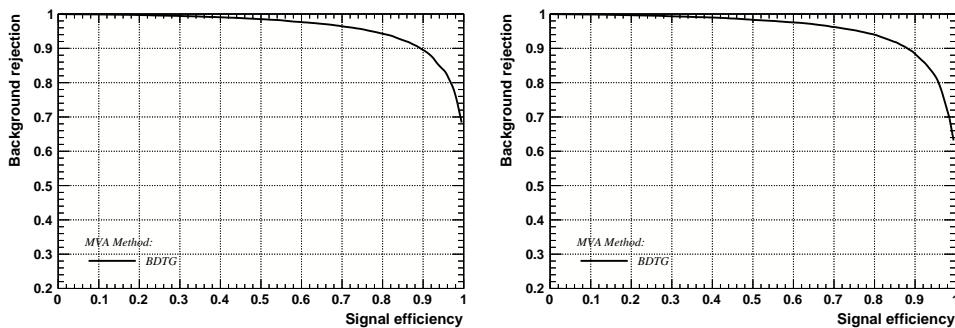


Figure A.12: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L06TIS].

## 882 C Detailed mass fits

883 In this section, all fits to the mass distribution of  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$   
 884 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-  
 885 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  
 886 L0 trigger category.

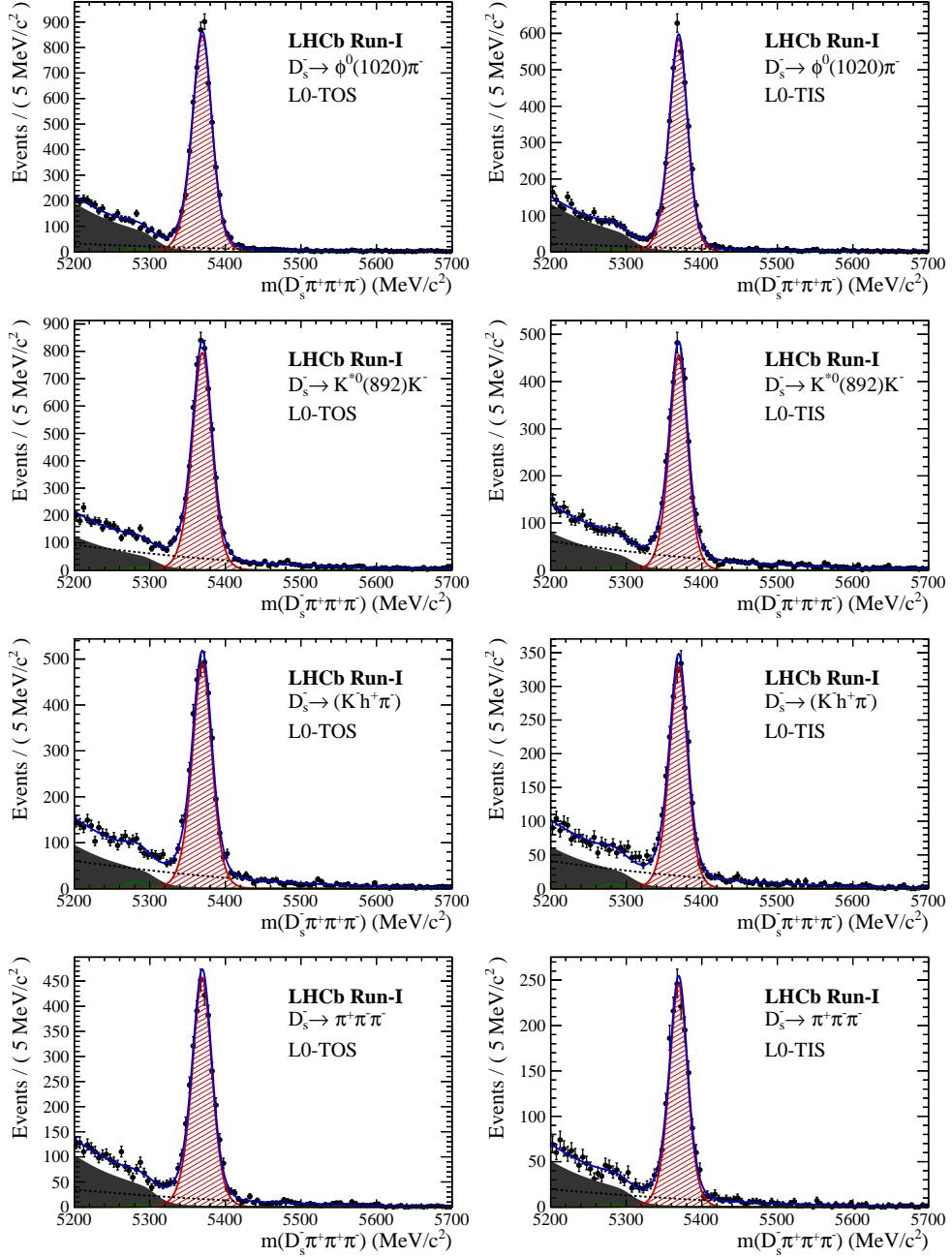


Figure B.1: Invariant mass distributions of  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates for Run-I data.

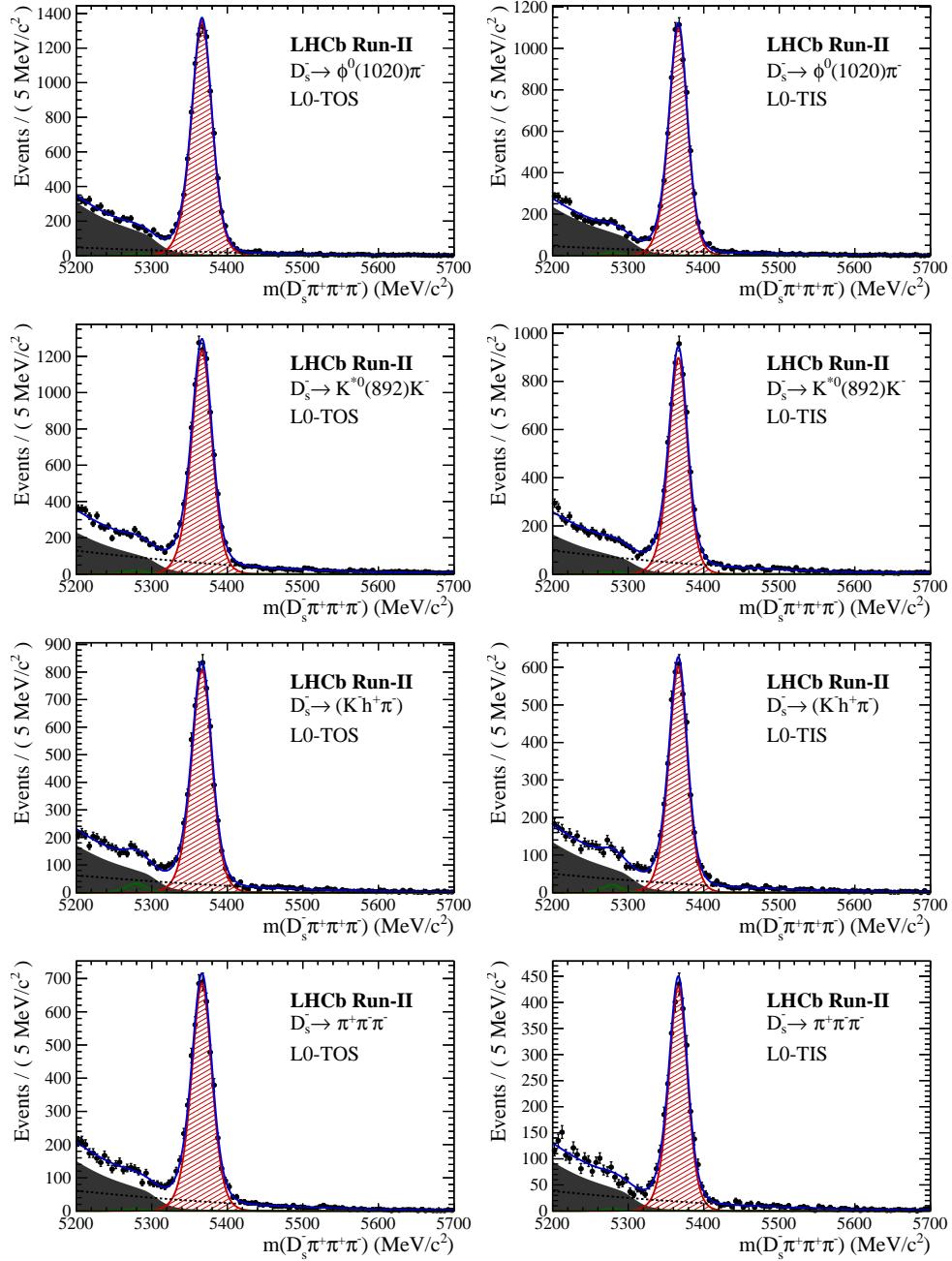


Figure B.2: Invariant mass distributions of  $B_s^0 \rightarrow D_s \pi^+ \pi^- \pi^+$  candidates for Run-II data.

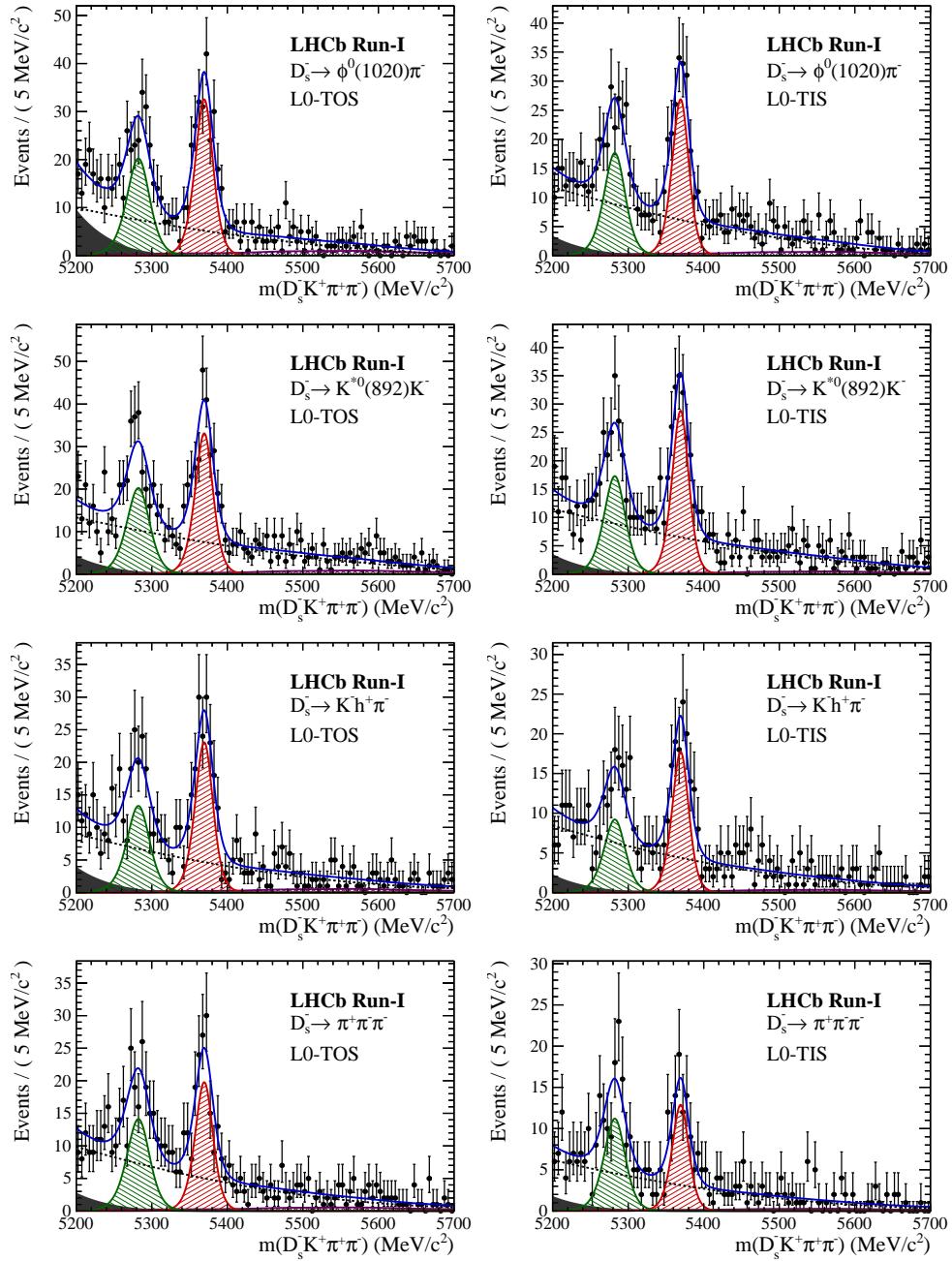


Figure B.3: Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for Run-I data.

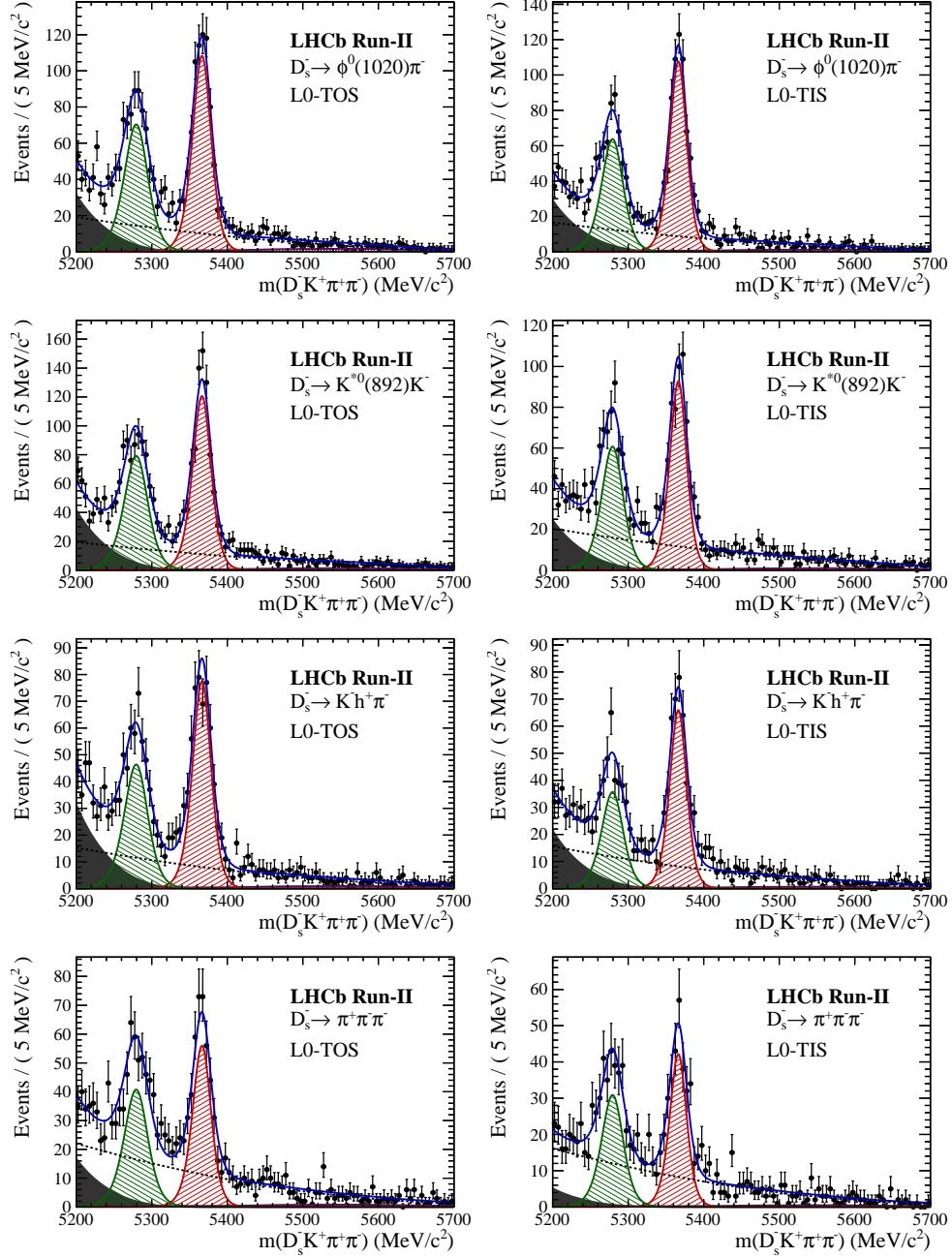


Figure B.4: Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi\pi$  candidates for Run-II data.

## 887 D Decay-time Resolution fits

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

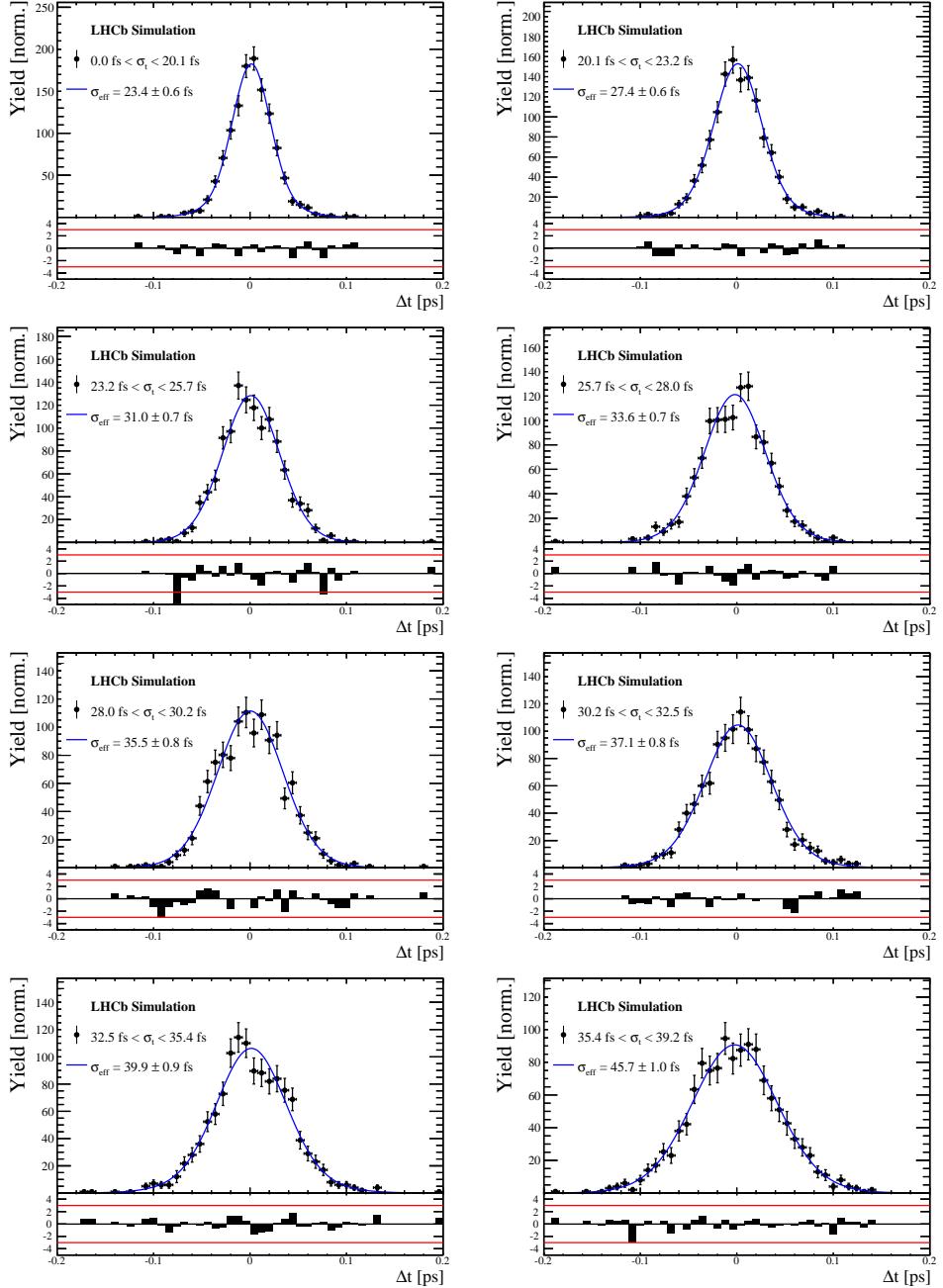


Figure C.1: Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K \pi\pi$  MC candidates in bins of the per-event decay time error estimate..

$\sigma_t$ Bin [fs]	$\sigma_1$ [fs]	$\sigma_2$ [fs]	$f_1$	D	$\sigma_{eff}$ [fs]
0.0 - 20.1	$19 \pm 0.675$	$33.8 \pm 1.77$	$0.75 \pm 0$	$0.917 \pm 0.00406$	$23.4 \pm 0.599$
20.1 - 23.2	$23.4 \pm 0.86$	$37.4 \pm 1.95$	$0.75 \pm 0$	$0.888 \pm 0.00477$	$27.4 \pm 0.621$
23.2 - 25.7	$28.1 \pm 1.02$	$38.7 \pm 2.32$	$0.75 \pm 0$	$0.86 \pm 0.00563$	$31 \pm 0.671$
25.7 - 28.0	$30.1 \pm 1.12$	$43.2 \pm 2.56$	$0.75 \pm 0$	$0.837 \pm 0.00651$	$33.6 \pm 0.734$
28.0 - 30.2	$32.4 \pm 1.12$	$44.2 \pm 2.59$	$0.75 \pm 0$	$0.819 \pm 0.00694$	$35.5 \pm 0.756$
30.2 - 32.5	$32.6 \pm 1.38$	$49.2 \pm 3.04$	$0.75 \pm 0$	$0.805 \pm 0.00792$	$37.1 \pm 0.841$
32.5 - 35.4	$34.4 \pm 1.19$	$54.7 \pm 2.85$	$0.75 \pm 0$	$0.778 \pm 0.0086$	$39.9 \pm 0.879$
35.4 - 39.2	$41.9 \pm 1.8$	$56.9 \pm 4.18$	$0.75 \pm 0$	$0.719 \pm 0.00997$	$45.7 \pm 0.962$
39.2 - 44.7	$42.2 \pm 1.56$	$68.1 \pm 4.01$	$0.75 \pm 0$	$0.687 \pm 0.0114$	$48.8 \pm 1.08$
44.7 - 120.0	$55.5 \pm 2.59$	$83 \pm 14.7$	$0.75 \pm 0$	$0.546 \pm 0.0521$	$62 \pm 4.89$

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.

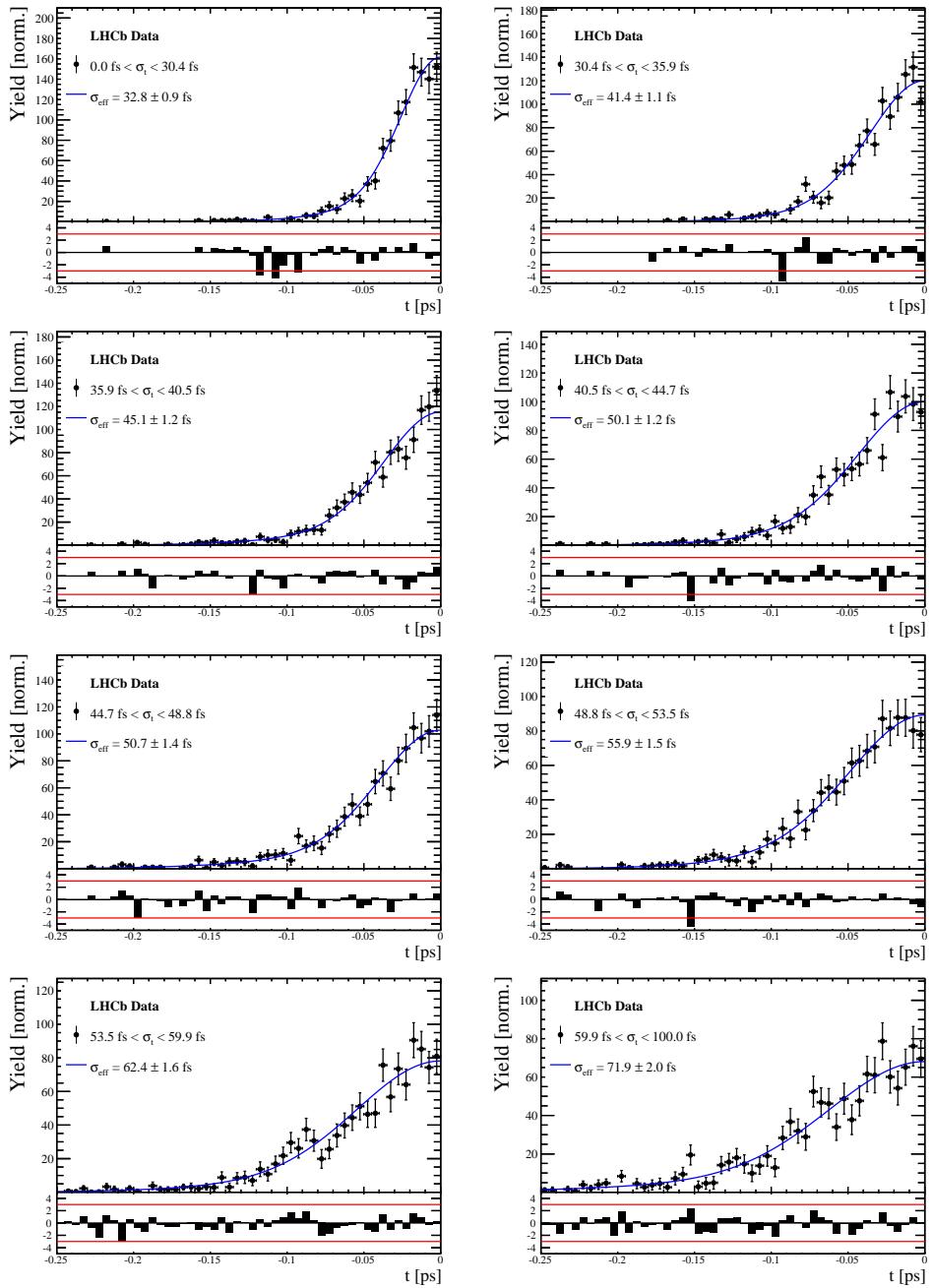


Figure C.2: Decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks, for bins in the per-event decay time error estimate.

$\sigma_t$ Bin [fs]	$\sigma_1$ [fs]	$\sigma_2$ [fs]	$f_1$	D	$\sigma_{eff}$ [fs]
0.0 - 30.4	$25.4 \pm 1.03$	$50.7 \pm 2.77$	$0.75 \pm 0$	$0.844 \pm 0.00822$	$32.8 \pm 0.942$
30.4 - 35.9	$34.5 \pm 1.46$	$60.2 \pm 3.48$	$0.75 \pm 0$	$0.763 \pm 0.0108$	$41.4 \pm 1.08$
35.9 - 40.5	$35.6 \pm 1.35$	$71.3 \pm 3.84$	$0.75 \pm 0$	$0.726 \pm 0.0121$	$45.1 \pm 1.18$
40.5 - 44.7	$42.3 \pm 1.65$	$73.3 \pm 4.21$	$0.75 \pm 0$	$0.673 \pm 0.0132$	$50.1 \pm 1.24$
44.7 - 48.8	$39.6 \pm 1.64$	$84.8 \pm 5.07$	$0.75 \pm 0$	$0.666 \pm 0.0145$	$50.7 \pm 1.36$
48.8 - 53.5	$47.6 \pm 1.94$	$82.4 \pm 5.48$	$0.75 \pm 0$	$0.611 \pm 0.0157$	$55.9 \pm 1.46$
53.5 - 59.9	$53 \pm 2.15$	$95.3 \pm 6.84$	$0.75 \pm 0$	$0.541 \pm 0.0174$	$62.4 \pm 1.63$
59.9 - 100.0	$60.5 \pm 2.8$	$125 \pm 14$	$0.75 \pm 0$	$0.443 \pm 0.0204$	$71.9 \pm 2.03$

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

## 893 E Spin Amplitudes

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

## 895 F Considered Decay Chains

896 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
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ \rightarrow \pi^+ K^*(1430)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \omega(782)]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ \kappa]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \sigma]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow \sigma^0(D_s^- K^+_S)$
$B_s [S, P, D] \rightarrow \sigma^0(D_s^- K^+_V)$
$B_s \rightarrow \rho(770)^0 (D_s^- K^+_S)$
$B_s [S, P, D] \rightarrow \rho(770)^0 (D_s^- K^+_V)$
$B_s \rightarrow K^*(892)^0 (D_s^- \pi^+_S)$
$B_s [S, P, D] \rightarrow K^*(892)^0 (D_s^- \pi^+_V)$
$B_s \rightarrow (D_s^- K^+_S) (\pi^+ \pi^-)_S$

897 G MC corrections

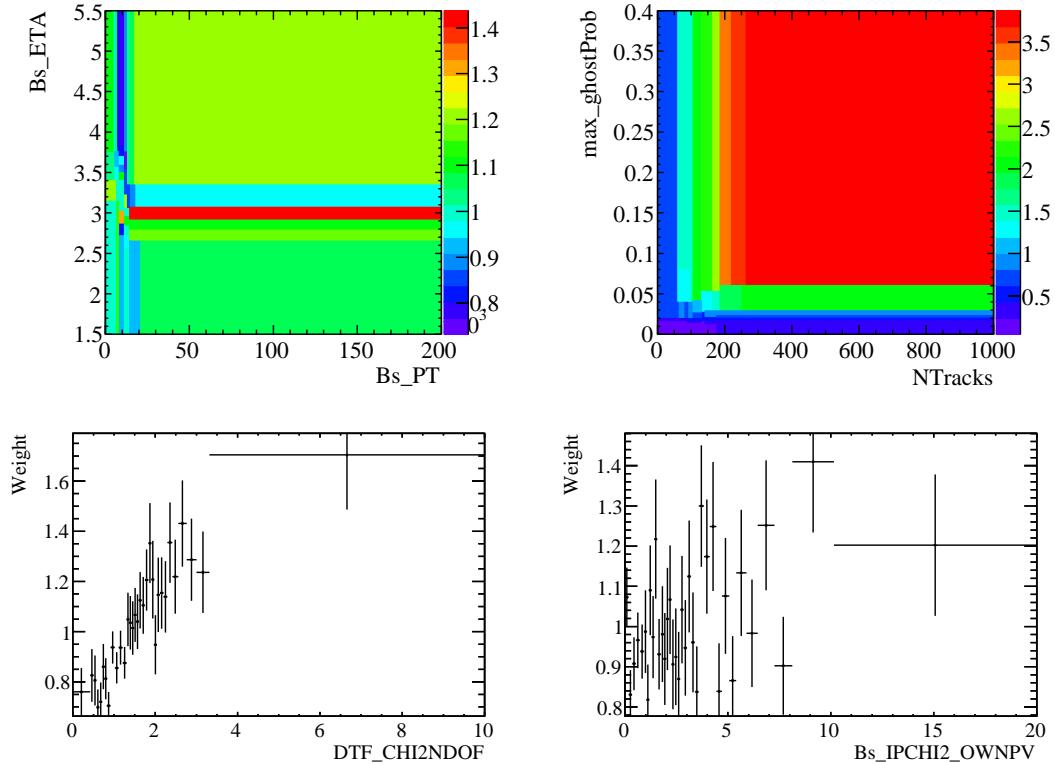


Figure C.1: Weights applied to correct for Data/MC differences.

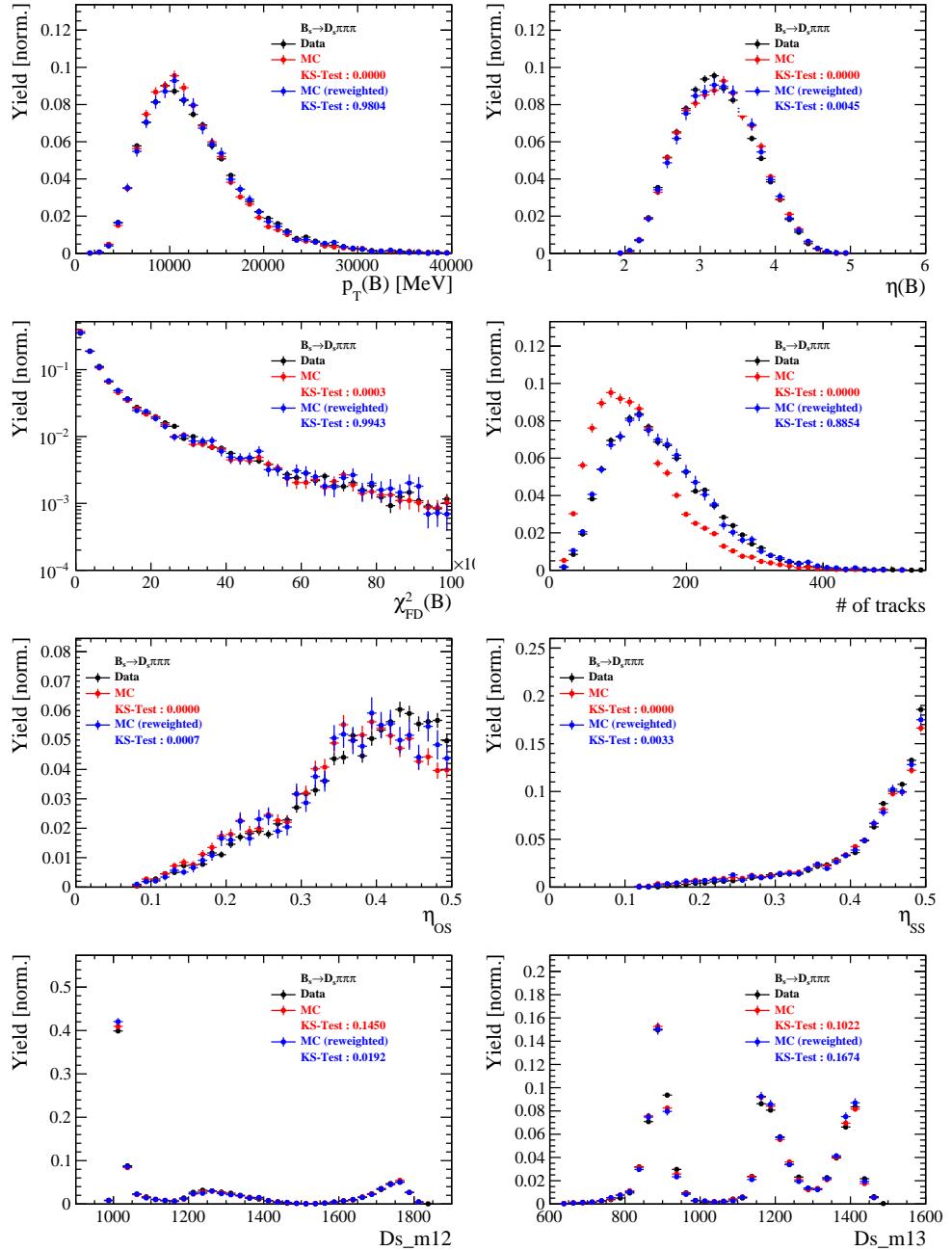


Figure C.2: Comparison of selected variables for  $B_s \rightarrow D_s \pi\pi\pi$  decays.

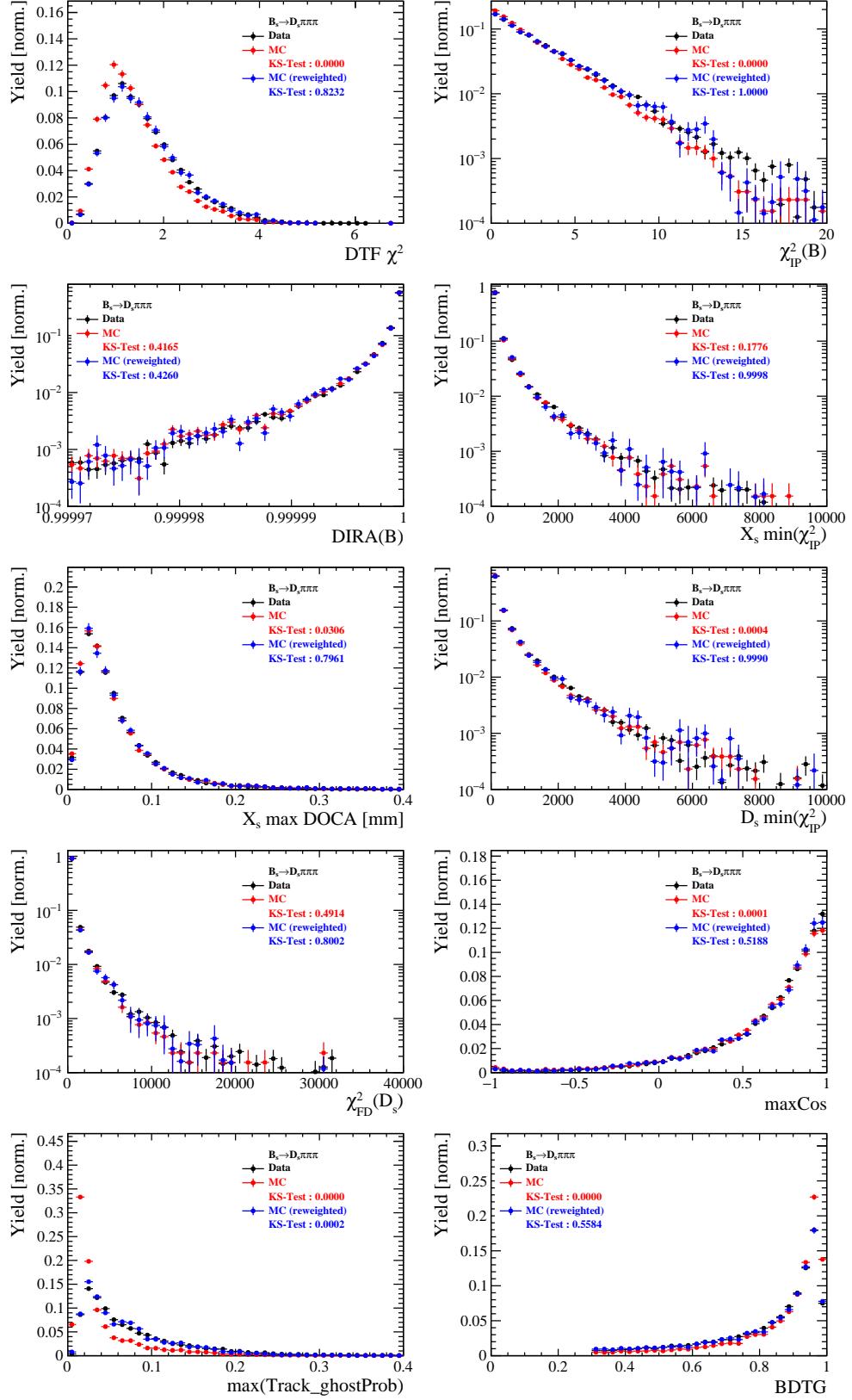


Figure C.3: Comparison of BDTG input variables and classifier response for  $B_s \rightarrow D_s\pi\pi\pi$  decays.

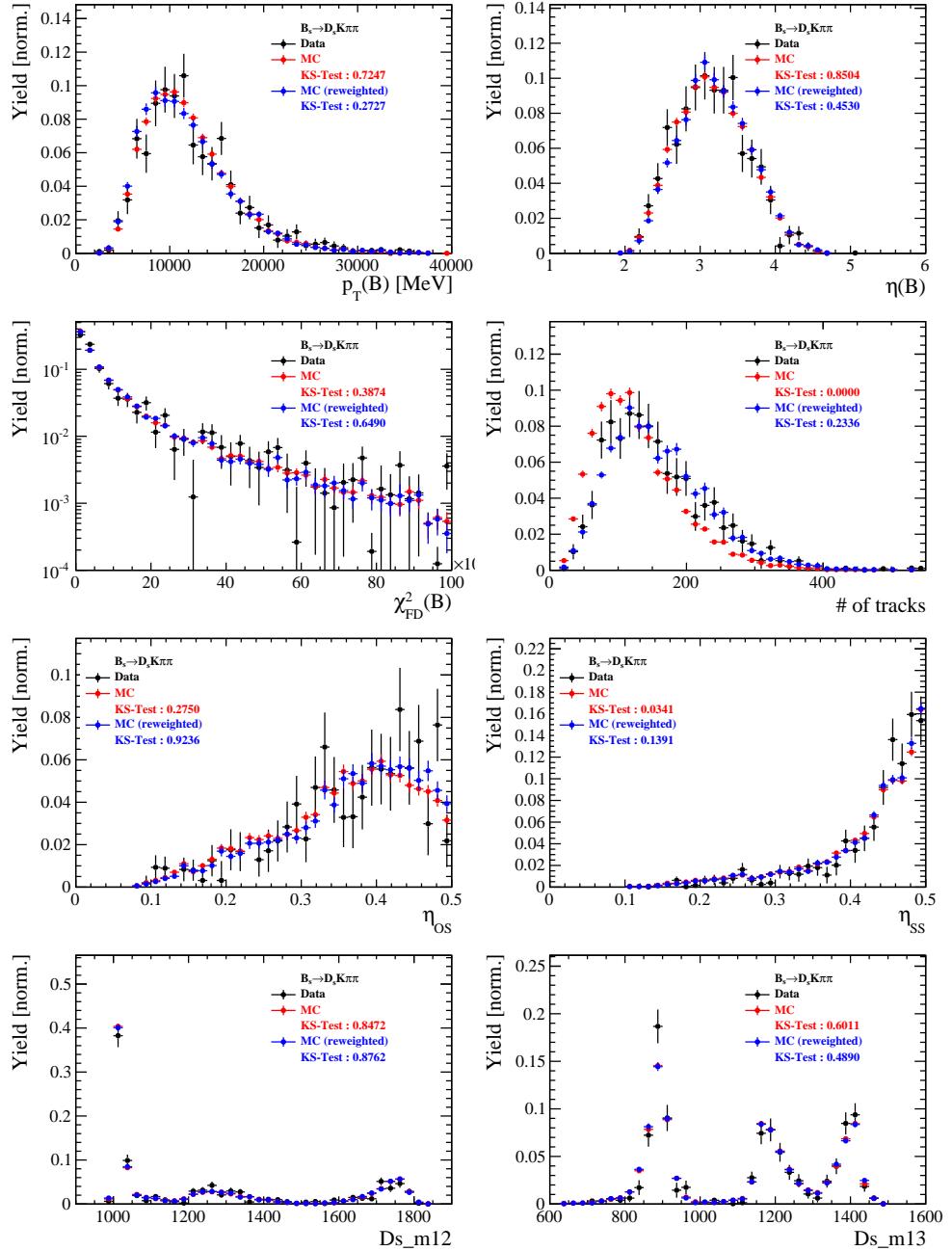


Figure C.4: Comparison of selected variables for  $B_s \rightarrow D_s K\pi\pi$  decays.

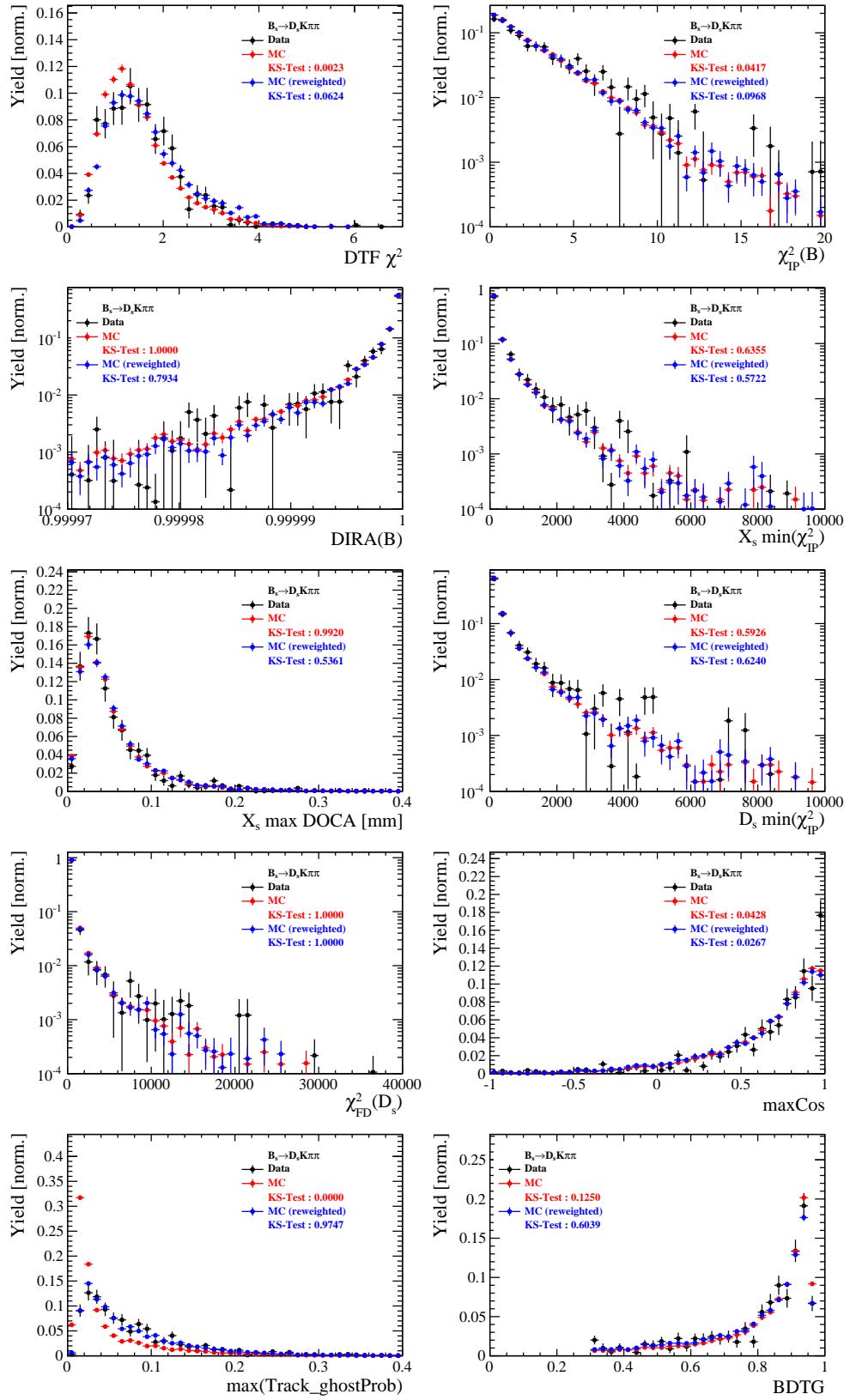


Figure C.5: Comparison of BDTG input variables and classifier response for  $B_s \rightarrow D_s K\pi\pi$  decays.

## 898 H Data distributions

### 899 H.1 Comparison of signal and calibration channel

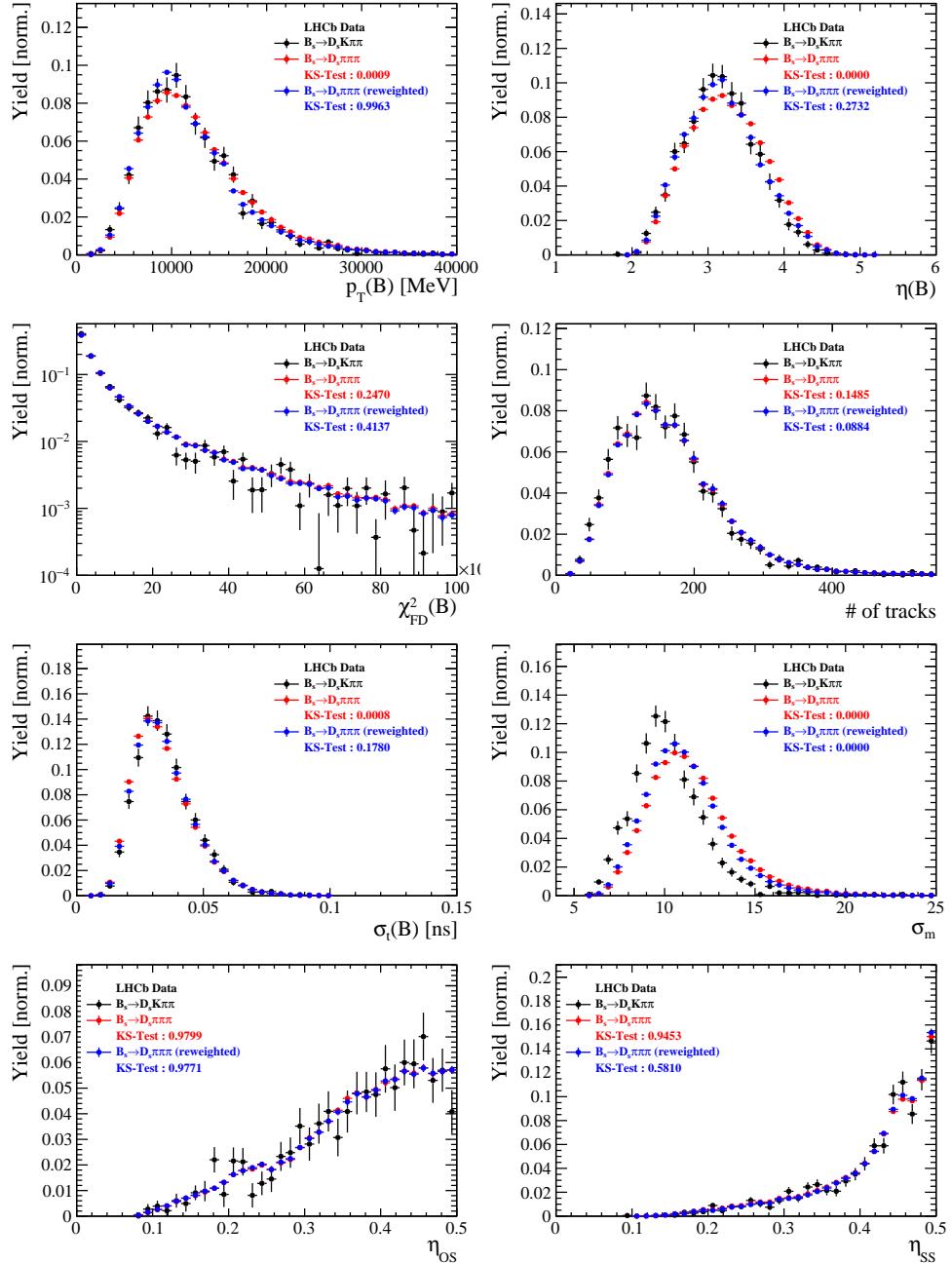


Figure C.1: Comparison of selected variables.

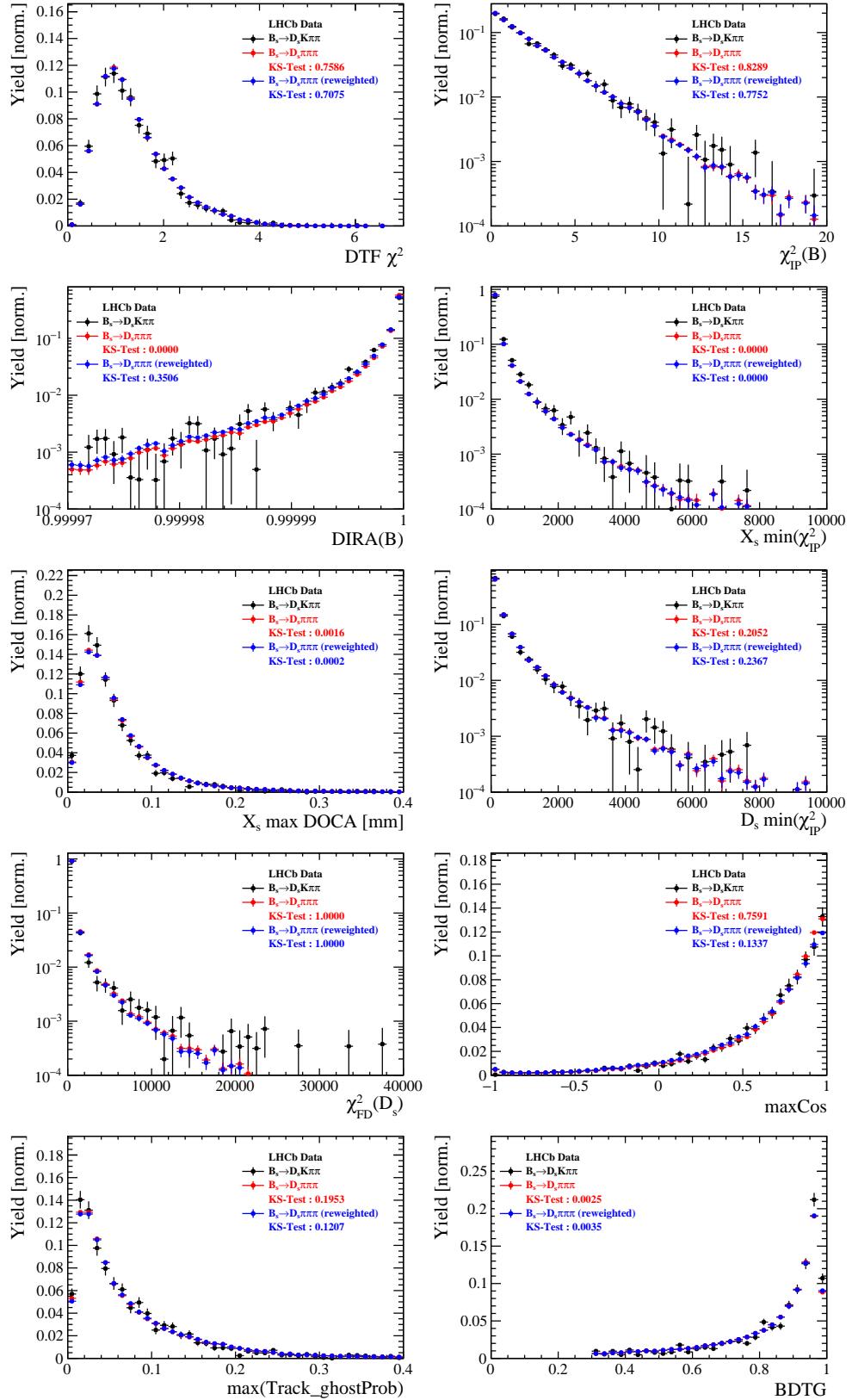


Figure C.2: Comparison of BDTG input variables and classifier response.

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

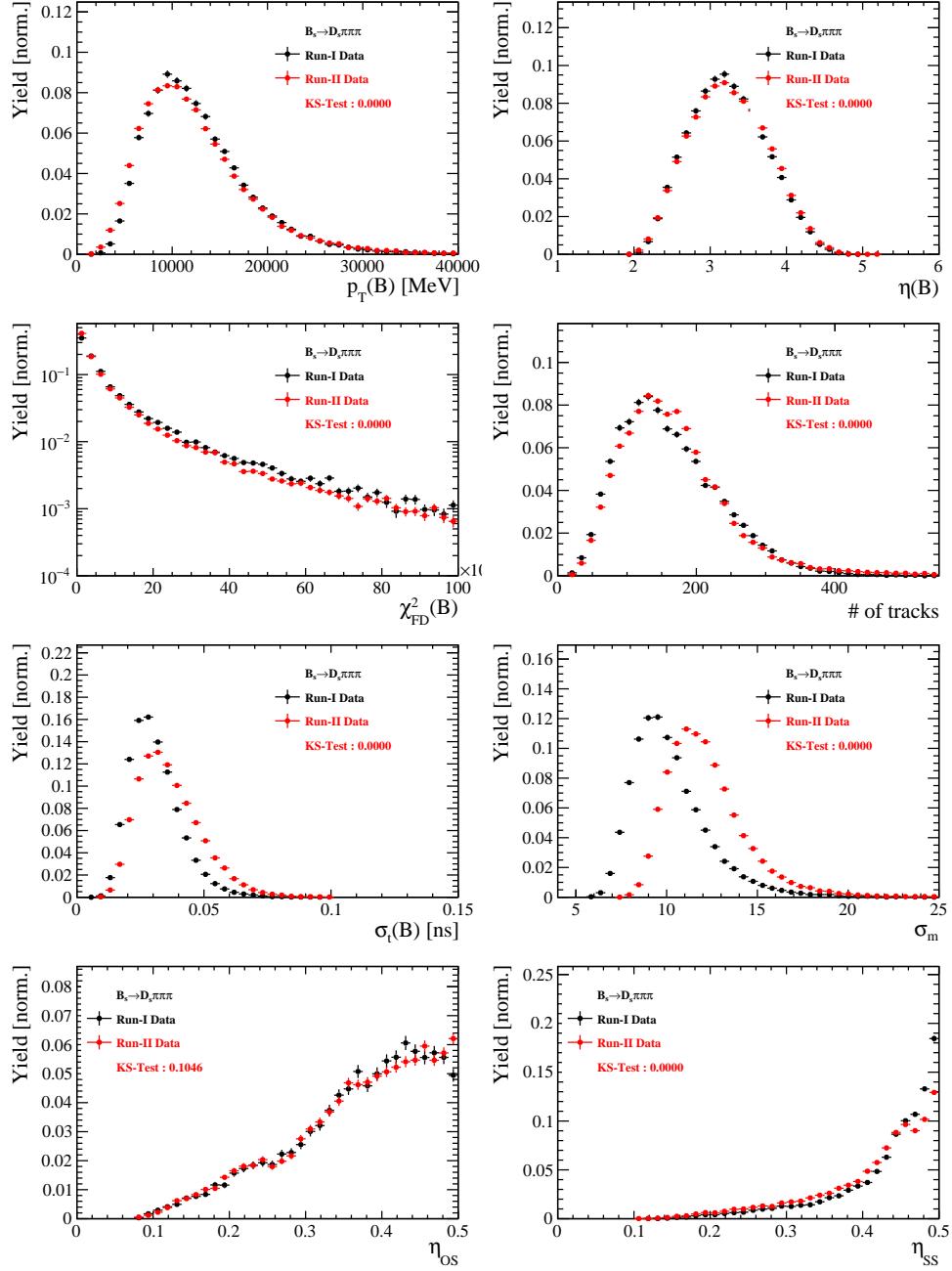


Figure C.3: Comparison of selected variables.

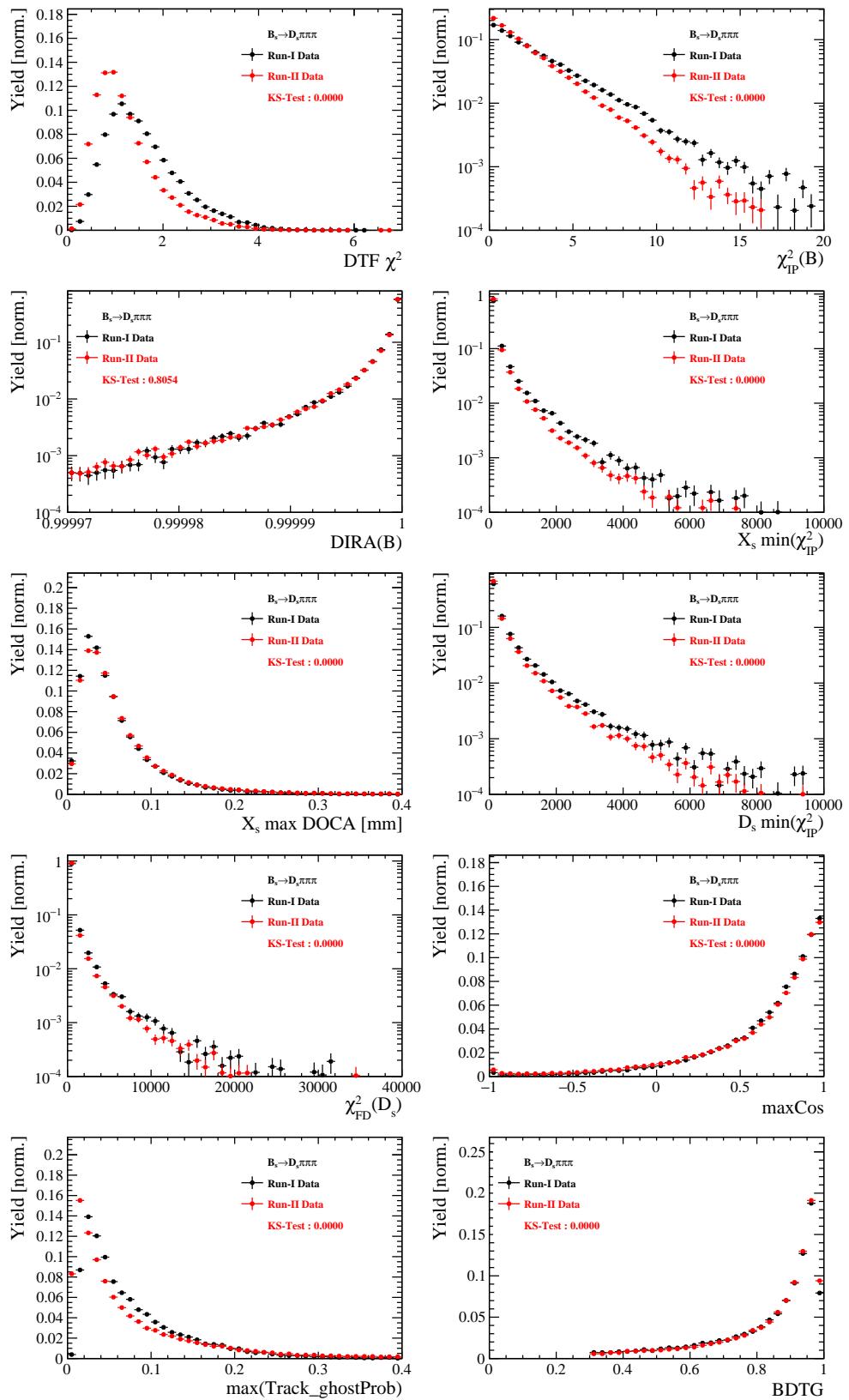


Figure C.4: Comparison of BDTG input variables and classifier response.

901 H.3 Comparison of  $D_s$  final states

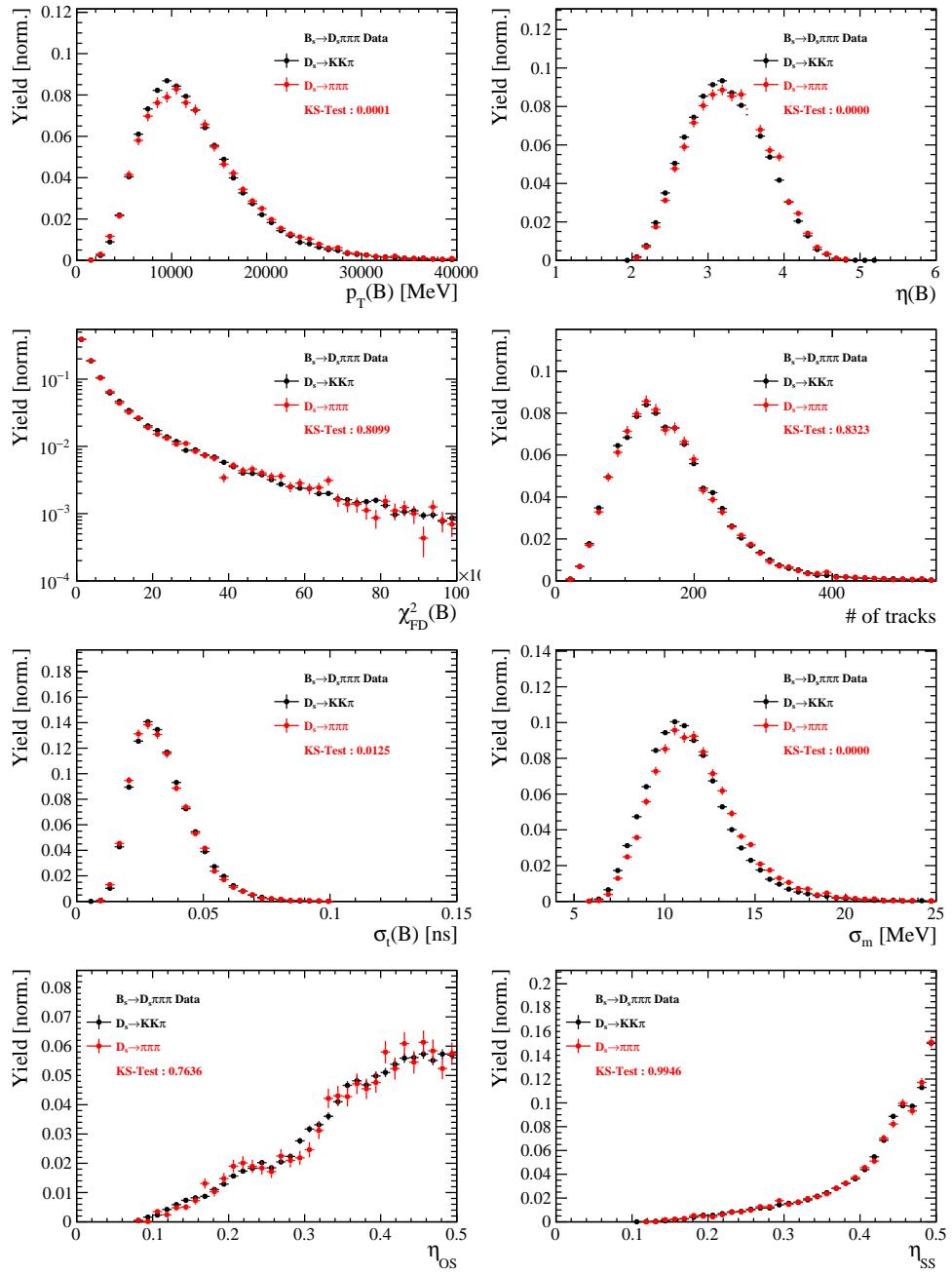


Figure C.5: Comparison of selected variables.

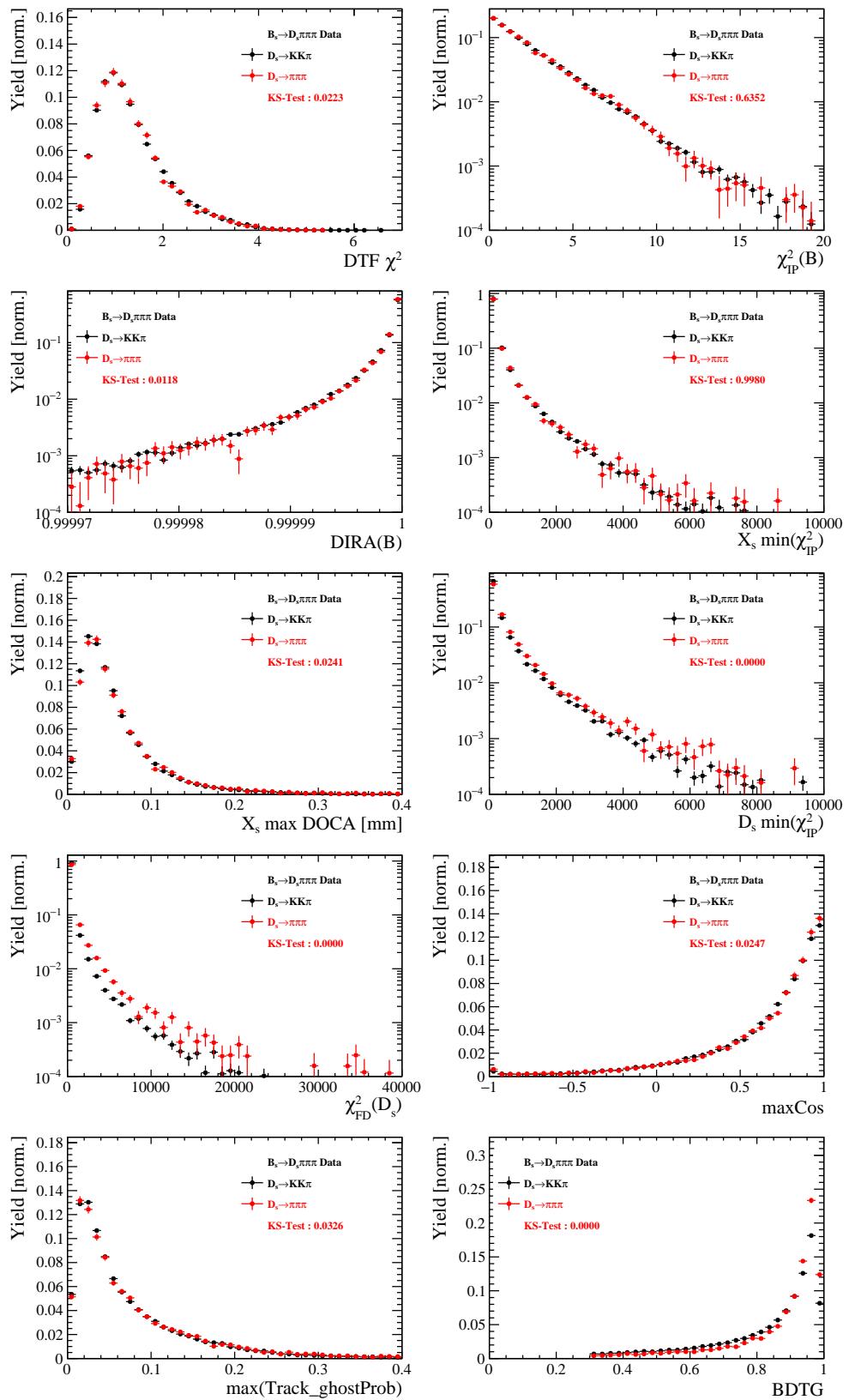


Figure C.6: Comparison of BDTG input variables and classifier response.

902 H.4 Comparison of trigger categories

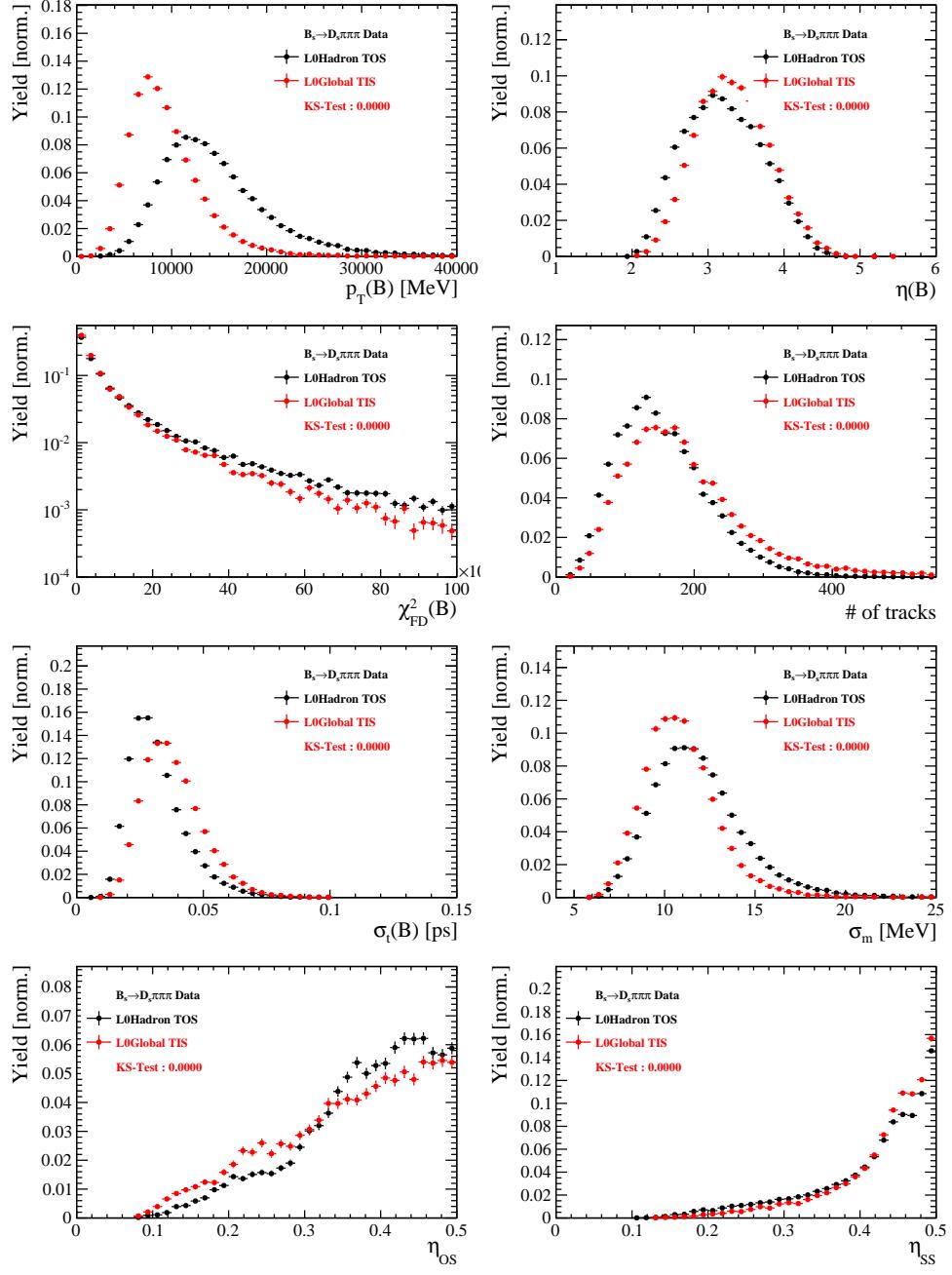


Figure C.7: Comparison of selected variables.

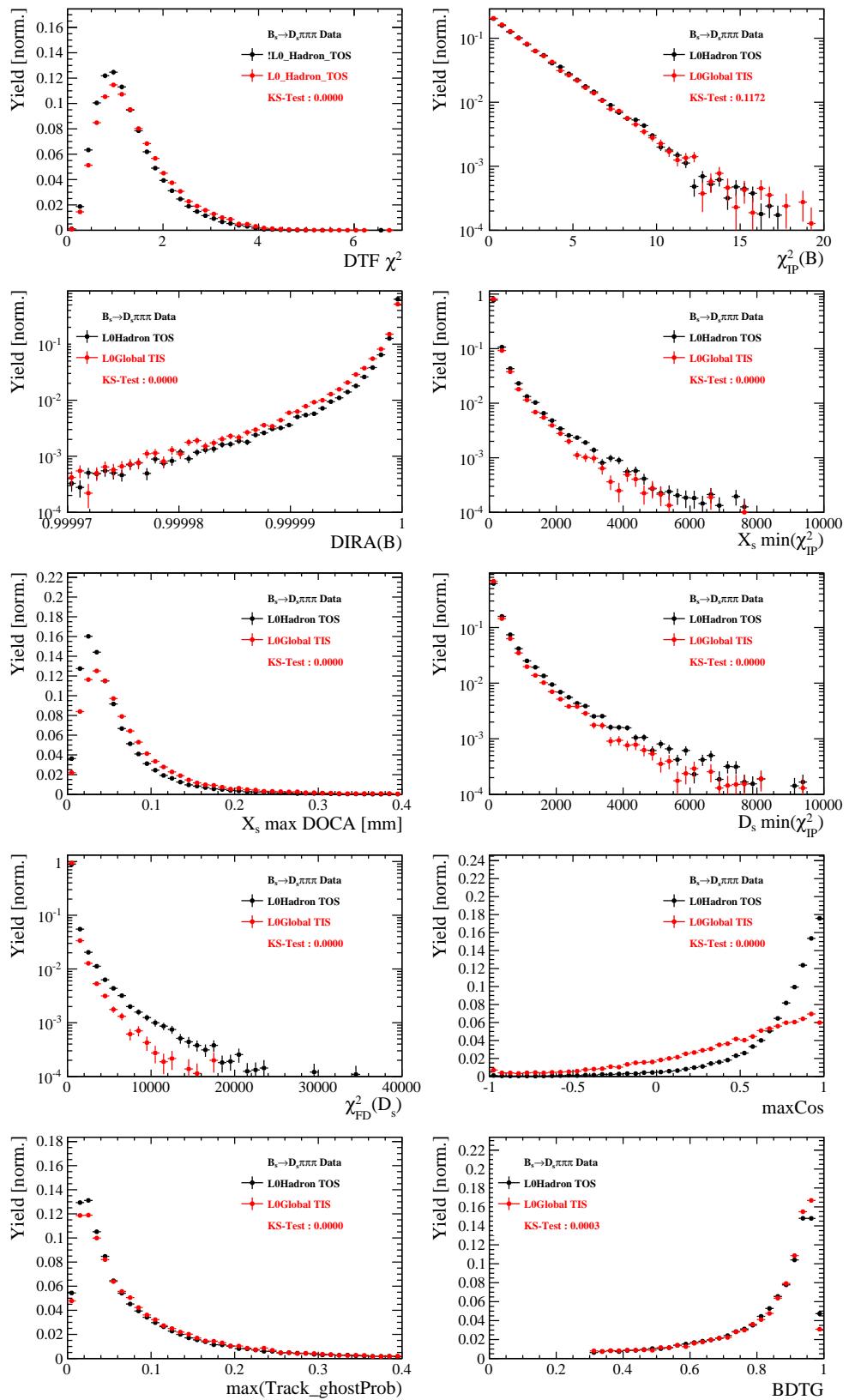


Figure C.8: Comparison of BDTG input variables and classifier response.

903 H.5 Comparison of  $B_s$  and  $B_d$  decays

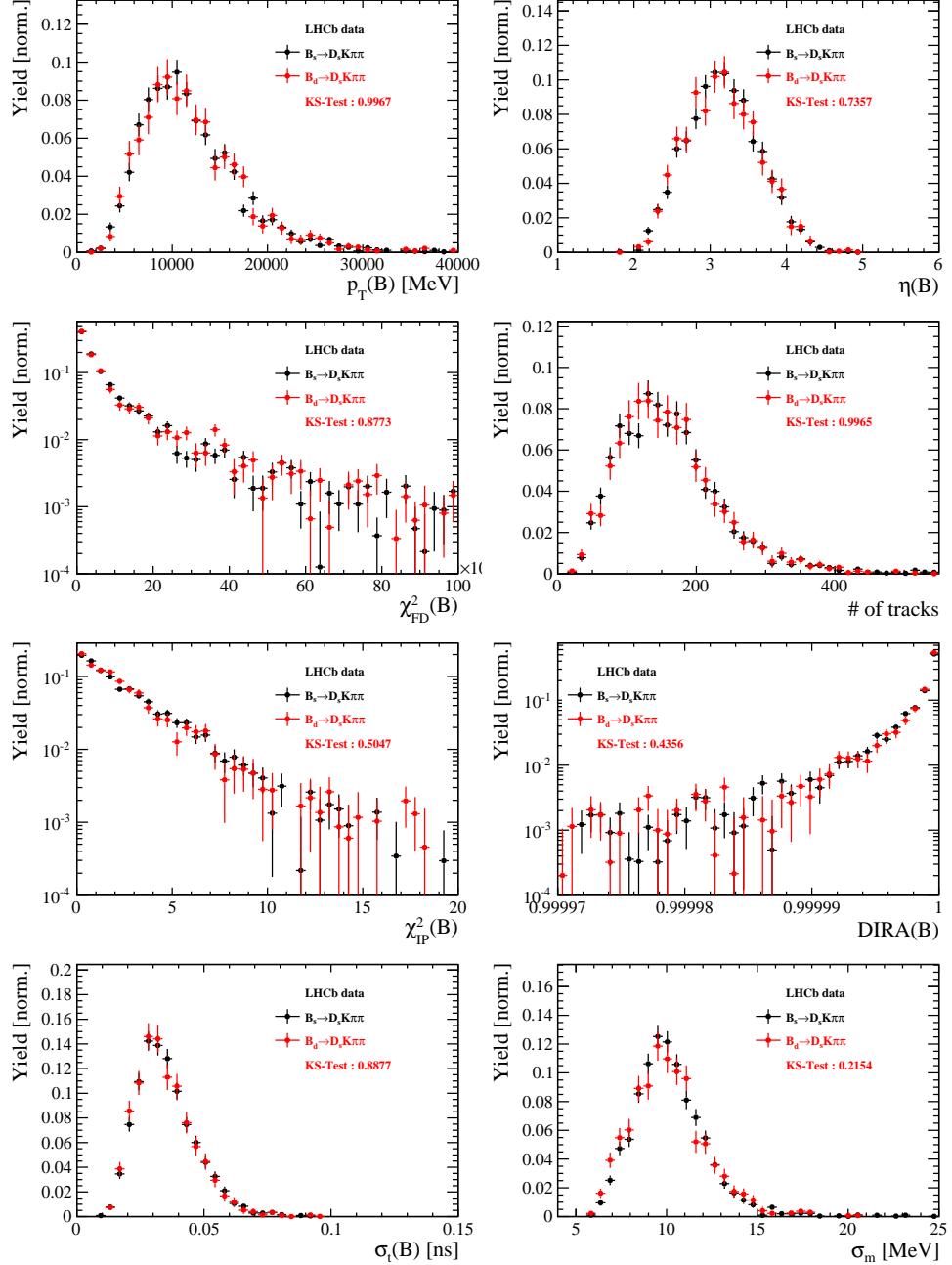


Figure C.9: Comparison of selected variables.

## 904 References

- 905 [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  $B(d) \rightarrow D \pi \pi$ ,  $D^* \pi \pi$ , ... decays*, Nucl.  
906 Phys. **B671** (2003) 459, [arXiv:hep-ph/0304027](#).
- 908 [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](#).
- 910 [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  $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* , .
- 912 [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](#).
- 914 [5] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory (Frontiers in Physics)*, Westview Press, 1995.
- 916 [6] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 918 [7] 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 .
- 920 [8] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*, Phys. Rev. D **11** (1975) 3165.
- 922 [9] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of Physics **108** (1977), no. 2 454 .
- 924 [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.
- 926 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo Cimento Series 10 **34** (1964), no. 6 1644.
- 928 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.
- 930 [13] D. V. Bugg, *The mass of the  $\sigma$  pole*, Journal of Physics G Nuclear Physics **34** (2007) 151, [arXiv:hep-ph/0608081](#).
- 932 [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.
- 934 [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 .
- 936 [16] BES Collaboration, M. Ablikim *et al.*, *Resonances in  $J/\psi \rightarrow \phi \pi^+ \pi^-$  and  $\phi K^+ K^-$* , Phys. Lett. **B607** (2005) 243, [arXiv:hep-ex/0411001](#).
- 938 [17] D. V. Bugg, *A study in depth of  $f_0(1370)$* , Eur. Phys. J. **C52** (2007) 55, [arXiv:0706.1341](#).

- [18] LHCb Collaboration, R. Aaij *et al.*, *Analysis of the resonant components in  $B_s \rightarrow J/\psi \pi^+ \pi^-$* , Phys. Rev. **D86** (2012) 052006, [arXiv:1204.5643](https://arxiv.org/abs/1204.5643).
- [19] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- [20] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys. Rev. **60** (1941) 61.
- [21] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys. Rev. D **57** (1998) 431.
- [22] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of  $\psi$  decay to mesons*, Eur. Phys. J. **A16** (2003) 537, [arXiv:hep-ph/0211457](https://arxiv.org/abs/hep-ph/0211457).
- [23] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. **D51** (1995) 2247.
- [24] J.-J. Zhu, *Explicit expressions of spin wave functions*, [arXiv:hep-ph/9906250](https://arxiv.org/abs/hep-ph/9906250).
- [25] 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](https://arxiv.org/abs/1703.08505).
- [26] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput. Phys. Commun. **180** (2009) 1847, [arXiv:0805.2956](https://arxiv.org/abs/0805.2956).
- [27] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in  $D^0 \rightarrow K^\mp \pi^\pm \pi^\pm \pi^\mp$  decays*, Submitted to: Eur. Phys. J. C (2017) [arXiv:1712.08609](https://arxiv.org/abs/1712.08609).
- [28] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [29] M. Karbach and M. Kenzie, *Gammacombo package*, <http://gammacombo.hepforge.org/web/HTML/index.html>, 2014.
- [30] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [31] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.
- [32] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [33] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](https://arxiv.org/abs/1412.6352).
- [34] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in  $B_s^0 \rightarrow D_s^\mp K^\pm$  decays*, Submitted to: JHEP (2017) [arXiv:1712.07428](https://arxiv.org/abs/1712.07428).
- [35] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and  $\tau$ -lepton properties as of summer 2014*, [arXiv:1412.7515](https://arxiv.org/abs/1412.7515), updated results and plots available at <http://www.slac.stanford.edu/xorg/hfag/>.

- [36] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, arXiv:1407.0748.
- [37] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of  $B_s^0$  mesons at LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- [38] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of  $B$  mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- [39] LHCb, R. Aaij *et al.*, *Measurement of  $B^0$ ,  $B_s^0$ ,  $B^+$  and  $\Lambda_b^0$  production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- [40] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of the  $K\pi$  Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-2012-027, CERN, Geneva, Feb, 2013.
- [41] A. Davis *et al.*, *Measurement of the instrumental asymmetry for  $K^- \pi^+$ -pairs at LHCb in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN, Geneva, Mar, 2018.
- [42] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly forbidden transitions in  $D \rightarrow K(S)$ ,  $K(L) + \pi$ 's decays*, Phys. Lett. **B349** (1995) 363, arXiv:hep-ph/9502238.
- [43] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- [44] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- [45] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal Statistical Society, Series B **58** (1994) 267.
- [46] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- [47] H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on Automatic Control **19** (1974) 716.
- [48] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.