



# 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 . . . . .	45
7.3	Tagging performance comparison between the signal and normalization channel . . . . .	46
7.4	Combination of OS and SS taggers . . . . .	47

<b>8 Production and Detection Asymmetries</b>	<b>49</b>
8.1 $B_s$ Production Asymmetry . . . . .	49
8.2 $K^-\pi^+$ Detection Asymmetry . . . . .	50
<b>9 Time dependent fit</b>	<b>53</b>
9.1 sFit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data . . . . .	53
9.2 sFit to $B_s^0 \rightarrow D_sK\pi\pi$ data . . . . .	56
9.3 sFit model validation using toy studies . . . . .	56
<b>10 Time dependent amplitude fit</b>	<b>59</b>
10.1 Signal Model Construction . . . . .	59
10.2 Results . . . . .	60
<b>11 Systematic uncertainties</b>	<b>63</b>
11.1 Models for $B_s^0$ mass distribution . . . . .	63
11.1.1 Signal model . . . . .	63
11.1.2 Background model . . . . .	63
11.1.3 Description of misidentified background . . . . .	64
11.1.4 Systematic effect on observables . . . . .	64
11.2 Decay-time acceptance . . . . .	65
11.2.1 Variation of knot positions . . . . .	65
11.2.2 Variation of spline coefficients . . . . .	65
11.3 Decay-time resolution . . . . .	66
11.4 Tagging calibration . . . . .	67
11.5 Summary of systematic uncertainties . . . . .	67
<b>A Stripping and Trigger cuts</b>	<b>68</b>
<b>B Details of multivariate classifier</b>	<b>70</b>
<b>C Detailed mass fits</b>	<b>74</b>
<b>D Decay-time Resolution fits</b>	<b>78</b>
<b>E Spin Amplitudes</b>	<b>82</b>
<b>F Considered Decay Chains</b>	<b>83</b>
<b>G MC corrections</b>	<b>84</b>
<b>H Data distributions</b>	<b>89</b>
H.1 Comparison of signal and calibration channel . . . . .	89
H.2 Comparison of Run-I and Run-II data . . . . .	91
H.3 Comparison of $D_s$ final states . . . . .	93
H.4 Comparison of trigger categories . . . . .	95
H.5 Comparison of $B_s$ and $B_d$ decays . . . . .	97
<b>References</b>	<b>98</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

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

<sup>22</sup> where the same convention for the CP coefficients as for the  $B_s \rightarrow D_s K$  analysis is used:

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

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

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

<sup>23</sup> The coherence factor  $\kappa$ , the strong phase difference  $\delta$  and the ratio of the suppressed  
<sup>24</sup> ( $b \rightarrow u$ ) over favored ( $b \rightarrow c$ ) decay mode are defined as:

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

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

<sup>25</sup> In the limit of only one contributing resonance  $\kappa \rightarrow 1$ .

<sup>26</sup>

## <sup>27</sup> 2.2 Amplitude model

<sup>28</sup> The differential decay rate of a  $B_s$  meson with mass,  $m_{B_s}$ , decaying into four pseudoscalar  
<sup>29</sup> particles with four-momenta  $p_i = (E_i, \vec{p}_i)$  ( $i = 1, 2, 3, 4$ ) is given by

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

<sup>30</sup> where the transition amplitude  $A(\mathbf{x})$ , describes the dynamics of the interaction,  $d\Phi_4$   
<sup>31</sup> is the four-body phase space element [5], and  $\mathbf{x}$  represents a unique set of kinematic  
<sup>32</sup> conditions within the phase space of the decay. Each final state particle contributes three  
<sup>33</sup> observables, manifesting in their three-momentum, summing up to twelve observables in  
<sup>34</sup> total. Four of them are redundant due to four-momentum conservation and the overall  
<sup>35</sup> orientation of the system can be integrated out. The remaining five independent degrees  
<sup>36</sup> of freedom unambiguously determine the kinematics of the decay. Convenient choices  
<sup>37</sup> for the kinematic observables include the invariant mass combinations of the final state  
<sup>38</sup> particles

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

<sup>39</sup> or acoplanarity and helicity angles. It is however important to take into account that,  
<sup>40</sup> while  $m_{12}^2, m_{23}^2$  are sufficient to fully describe a three-body decay, the obvious extension  
<sup>41</sup> to four-body decays with  $m_{ij}^2, m_{ijk}^2$  requires additional care, as these variables alone are  
<sup>42</sup> insufficient to describe the parity-odd moments possible in four-body kinematics.

<sup>43</sup> In practice, we do not need to choose a particular five-dimensional basis, but use the  
<sup>44</sup> full four-vectors of the decay in our analysis. The dimensionality is handled by the phase

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

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

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

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

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

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

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

### 60 2.2.1 Form Factors and Resonance Lineshapes

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

86 As experimental values are usually only available for the branching fractions, Eq. (2.22)  
 87 needs to be inverted to obtain values for the couplings. In practice, this is solved by  
 88 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-  
 89 hand side of Eq. (2.22).

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

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

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

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

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

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

## 112 2.2.2 Spin Densities

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

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

125 The spin projection operator  $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$ , for a resonance  $R$ , with spin  $S =$   
126  $\{0, 1, 2\}$ , and four-momentum  $p_R$ , is given by [23]:

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

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

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

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

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

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

135 Following the isobar approach, a four-body decay amplitude is described as a product  
136 of two-body decay amplitudes. Each sequential two-body decay  $R \rightarrow A B$ , with relative  
137 orbital angular momentum  $L_{AB}$ , and total intrinsic spin  $S_{AB}$ , contributes a term to the  
138 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>139</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>140</sup> Here, a polarization vector is assigned to the decaying particle and the complex conjugate  
<sup>141</sup> vectors for each decay product. The spin and orbital angular momentum couplings are  
<sup>142</sup> described by the tensors  $P_{(S_{AB})}(R)$  and  $L_{(L_{AB})}(R)$ , respectively. Firstly, the two spins  $S_A$   
<sup>143</sup> and  $S_B$ , are coupled to a total spin- $S_{AB}$  state,  $\Phi(\mathbf{x}|S_{AB})$ , by projecting the corresponding  
<sup>144</sup> polarization vectors onto the spin- $S_{AB}$  subspace transverse to the momentum of the  
<sup>145</sup> decaying particle. Afterwards, the spin and orbital angular momentum tensors are  
<sup>146</sup> properly contracted with the polarization vector of the decaying particle to give a Lorentz  
<sup>147</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>148</sup> where  $\varepsilon_{\alpha\beta\gamma\delta}$  is the Levi-Civita symbol and  $j$  refers to the arguments of  $X$  defined in  
<sup>149</sup> Eqs. 2.27 and 2.28. Its antisymmetric nature ensures the correct parity transformation  
<sup>150</sup> behavior of the amplitude. The spin factor for a whole decay chain, for example  $R \rightarrow$   
<sup>151</sup>  $(R_1 \rightarrow AB)(R_2 \rightarrow CD)$ , is obtained by combining the two-body terms and performing a  
<sup>152</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>153</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>154</sup> pseudoscalar initial/final states are involved.

<sup>155</sup> The spin factors for all decay topologies considered in this analysis are explicitly given  
<sup>156</sup> in Appendix E.

## 157 2.3 Validation

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

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

172 The CP coefficients  $C, D, \bar{D}, S, \bar{S}$  are the fit parameters in case of the phasespace-  
 173 integrated pdf, while the full pdf determines  $x_{\pm} = r \cos(\delta \pm (\gamma - 2\beta_s))$  and  $y_{\pm} = r \sin(\delta \pm$   
 174  $(\gamma - 2\beta_s))$ . The fit parameters are converted to the physical observables  $r, \kappa, \delta, \gamma$  using  
 175 the **GammaCombo** package [29]. As shown in Tab. 2.2, 2.3 and 2.4, the fit results are  
 176 in excellent agreement with the generated input values. The 1-CL contours are shown  
 177 in Figs. 2.1 and 2.2. The phasespace-integrated fit is, in addition, performed with the  
 178 **B2DX** fitter used for the  $B_s \rightarrow D_s K$  analysis yielding identical results. Note though that  
 179 some parts of the **B2DX** fitter have been taken over to our **MINT2** fitter, such that the  
 180 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
$\delta$	$10.0^\circ$
$\gamma$	$71.1^\circ$
$\beta_s$	$0.0^\circ$

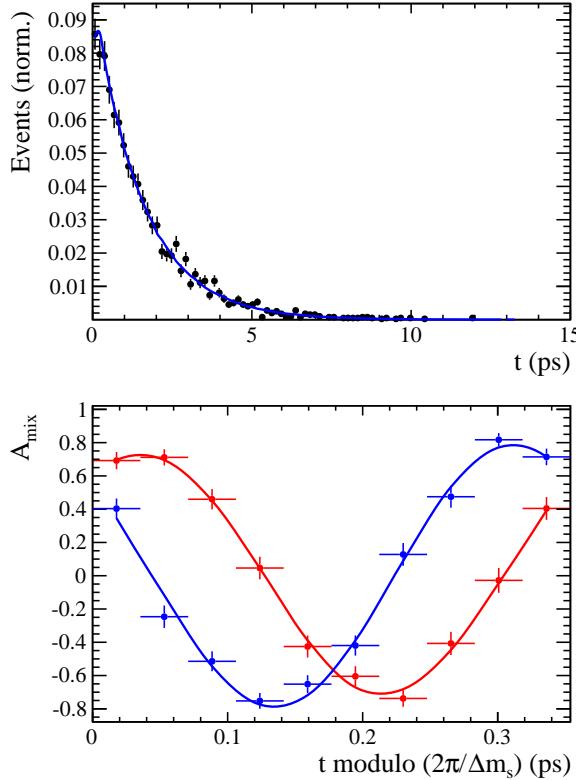


Figure 2.1: Left: Time distribution of  $B_s \rightarrow D_s K\pi\pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (solid line). Right: Time-dependent asymmetry between mixed and unmixed events over one mixing period  $2\pi/\Delta m_s$  for both final state charge configurations ( $D_s^\pm K^\mp\pi\pi$ ) in red & blue. The points are generated using EVTGEN and the MINT2 fit projections are overlaid.

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

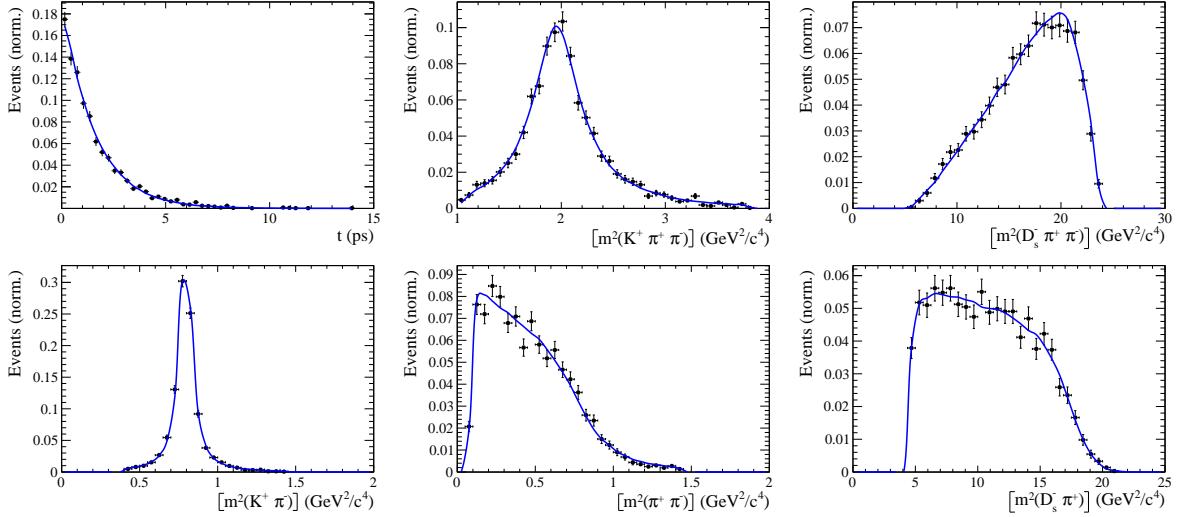


Figure 2.2: Time and invariant mass distributions of  $B_s \rightarrow D_s K\pi\pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

Table 2.3: 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	$1.000 \pm 0.059$
$\delta$	$10.0^\circ$	$9.0 \pm 5.1$	$5.9 \pm 6.0$
$\gamma$	$71.1^\circ$	$67.3 \pm 5.9$	$75.1 \pm 6.9$

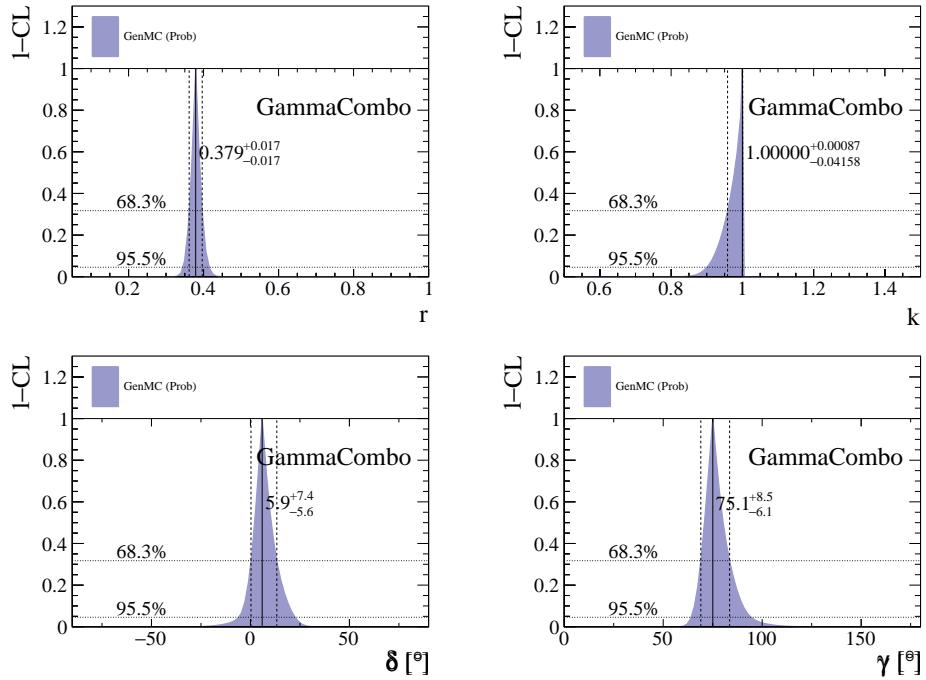


Figure 2.3: The 1-CL contours for the physical observable  $r, \kappa, \delta, \gamma$  obtained with the phasespace integrated fit to the EVTGEN toy sample.

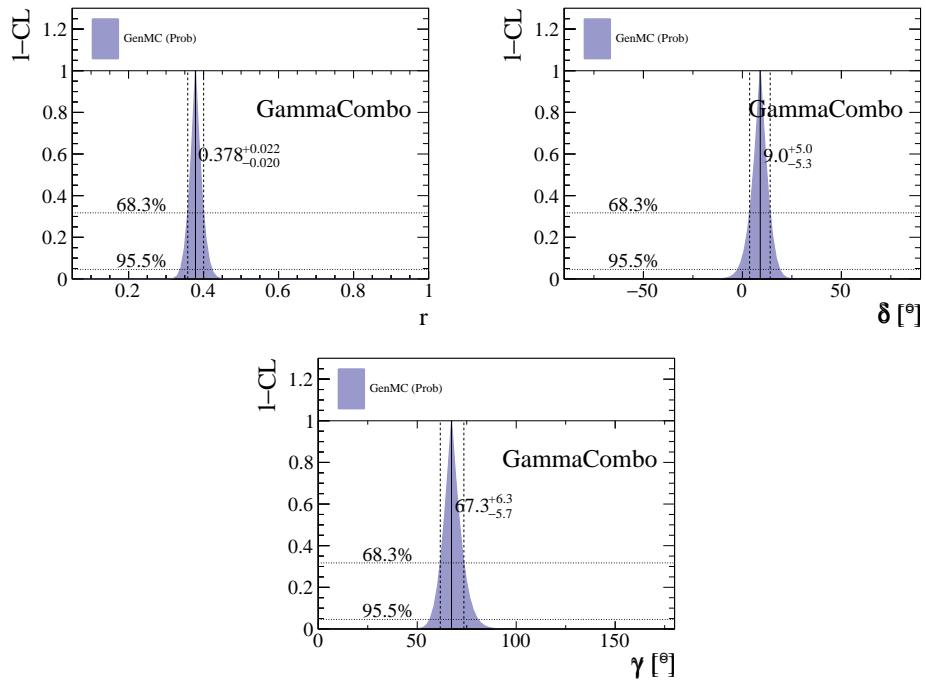


Figure 2.4: The 1-CL contours for the physical observable  $r, \delta, \gamma$  obtained with the time-dependent amplitude fit fit to the EVTGEN toy sample.

### 181 3 Data samples and event selection

182 Throughout the note, we abbreviate  $B_s^0 \rightarrow D_s X_s (\rightarrow K\pi\pi)$  and  $B_s^0 \rightarrow D_s X_d (\rightarrow \pi\pi\pi)$ ,  
183 identifying  $X_s \rightarrow K\pi\pi$  and  $X_d \rightarrow \pi\pi\pi$  as the various resonances through which the decays  
184 proceed.

#### 185 3.1 Stripping and Trigger selection

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

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

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

#### 207 3.2 Offline selection

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

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

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

- 221 • DIRA > 0.99994

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

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

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

### 3.2.1 Phase space region

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

254 **3.2.2 Physics background vetoes**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

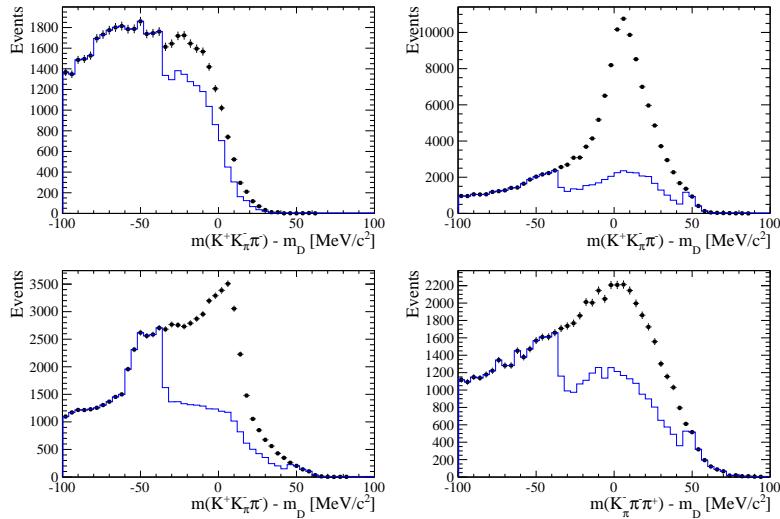


Figure 3.1: Background contributions from  $D^-$  decays where the  $\pi^-$  is misidentified as  $K^-$ . The  $D_s$  invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*(892)K$ ,  $D_s \rightarrow KK\pi$  (non-resonant)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

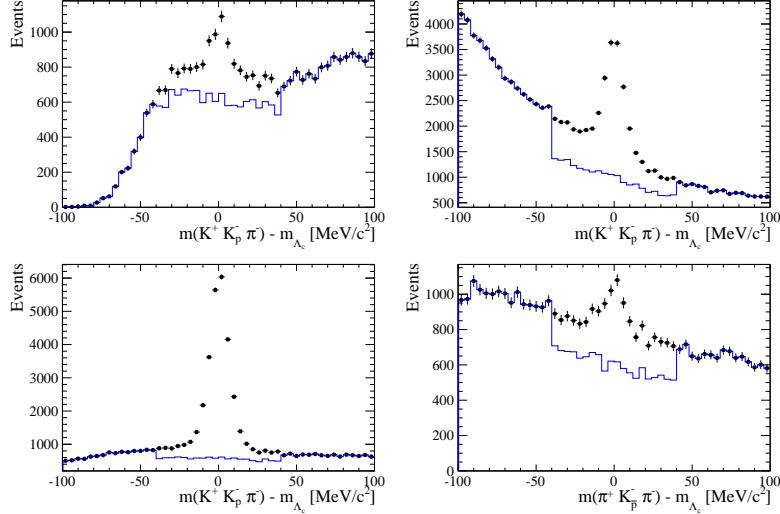


Figure 3.2: Background contributions from  $\Lambda_c$  decays where the  $\bar{p}$  is misidentified as  $K^-$ . The  $D_s$  invariant mass is recomputed applying the proton mass hypothesis to the kaon and shown for the  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*(892)K$ ,  $D_s \rightarrow KK\pi$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

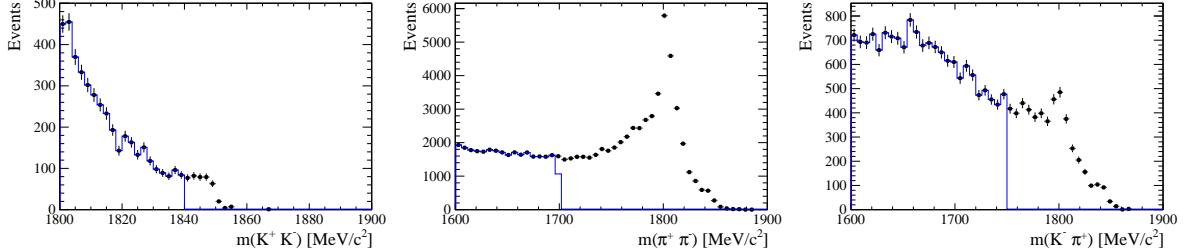


Figure 3.3: Background contributions to  $D_s \rightarrow KK\pi$  (left),  $D_s \rightarrow \pi\pi\pi$  (middle) and  $D_s \rightarrow K\pi\pi$  (right) from  $D^0 \rightarrow hh$  decays combined with a random pion.

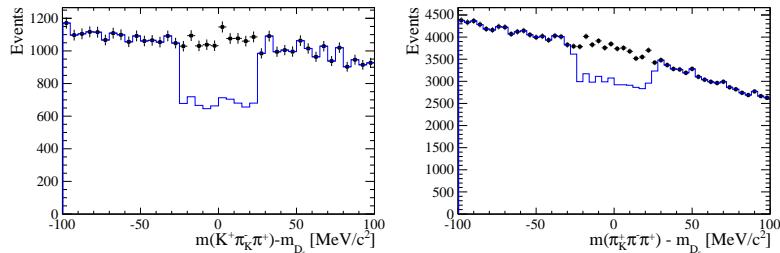


Figure 3.4: Background contributions to  $B_s \rightarrow D_s K\pi\pi$  (left) and  $B_s \rightarrow D_s \pi\pi\pi$  (right) from  $B_s \rightarrow D_s D_s$  decays where the kaon is misidentified as pion. The  $X_{s,d}$  invariant mass is recomputed applying the kaon mass hypothesis to the pion and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

308 **3.2.3 Training of multivariate classifier**

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

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

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

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

339 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  
340 to the training which are expected to be 100% signal efficient. Figure 3.6 shows the  
341 distributions of the input variables for signal and background. As shown in Appendix B,  
342 these distributions differ between data-taking period and trigger category. In particular  
343 variables depending on the  $B_s$  kinematics and the event multiplicity are affected (e.g.  
344  $\theta_{D_s h_i}$  or  $A_{pT}^{\text{cone}}$ ). The BDTG is consequently trained separately for these categories. The  
345 resulting classifier response is shown in Fig. 3.7 for each category (even and odd test  
346 samples combined) and in Appendix B for each training.

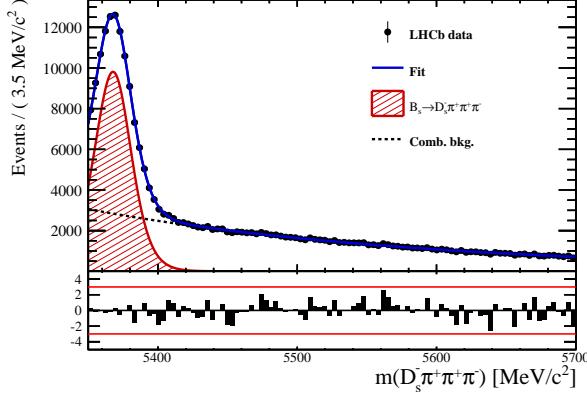


Figure 3.5: Reconstructed  $B_s$  mass for  $B_s \rightarrow D_s\pi\pi\pi$  events that pass the preselection. The fitted PDF is shown in blue, the signal component in red and the background component in black.

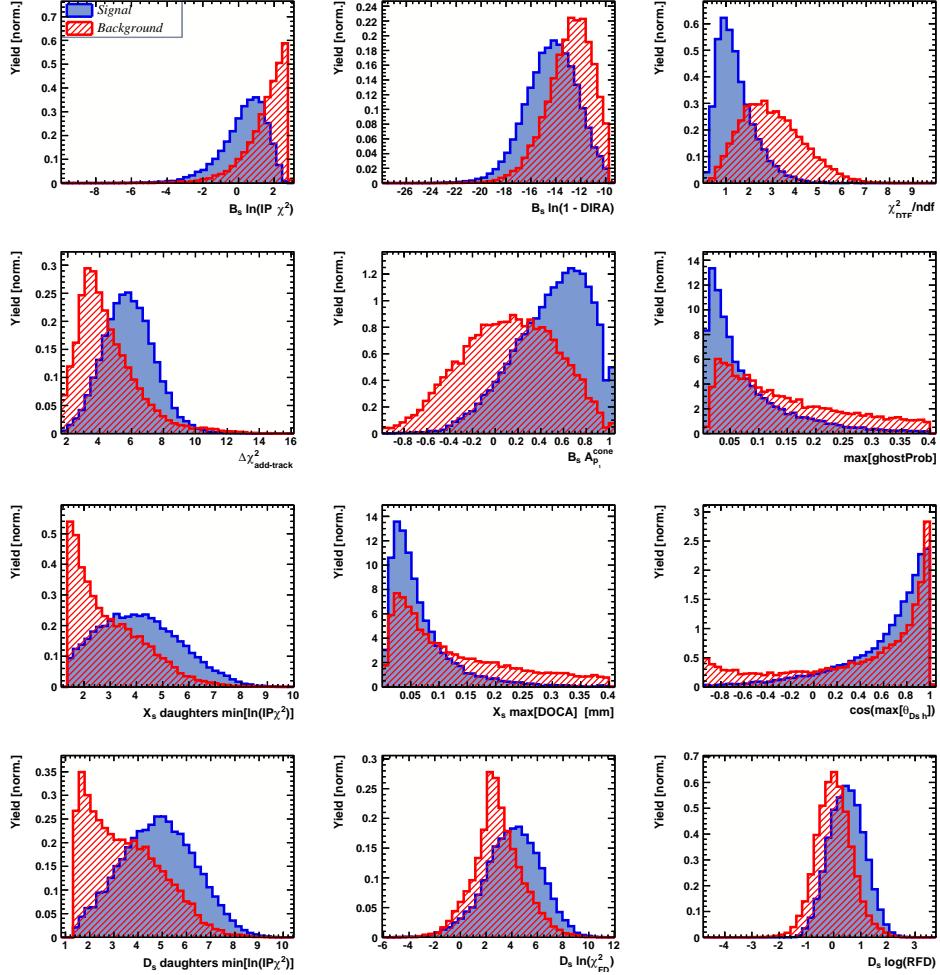


Figure 3.6: Discriminating variables used to train the BDTG for all data categories combined.

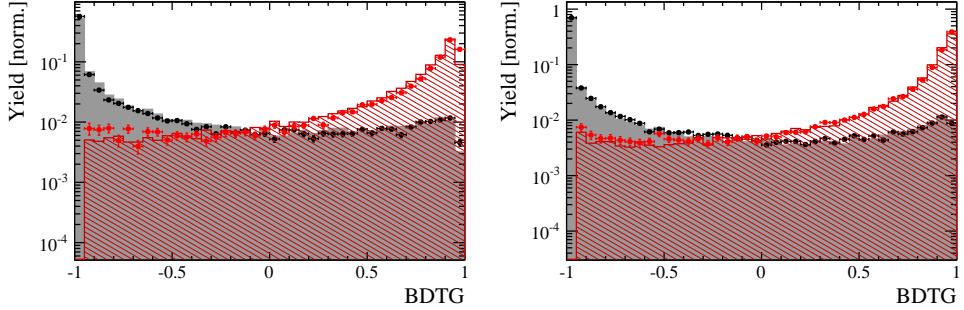


Figure 3.7: Signal (red) and background (black) distributions for the BDTG response for Run-I (left) and Run-II (right) data. Filled histograms (data points) show the BDTG response for the L0-TOS (L0-TIS) category. Even and odd test samples are combined.

### 3.2.4 Final selection

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

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

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

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

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

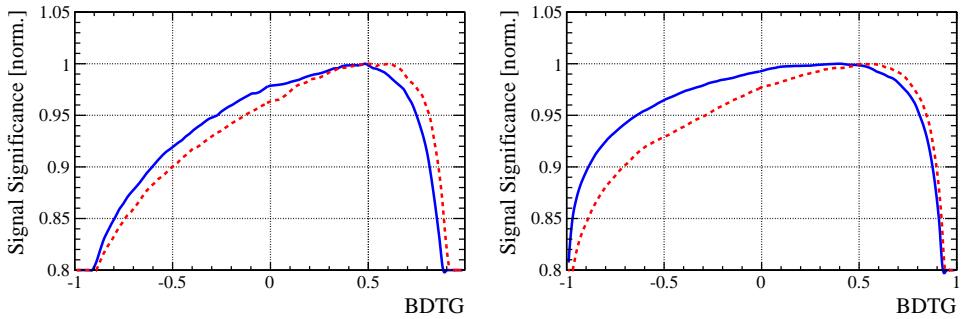


Figure 3.8: Signal significance as function of the applied cut on the BDTG response for Run-I (left) and Run-II (right) data. The scans for the L0-TOS (L0-TIS) category are shown in blue (red). The signal significance is normalized to be 1 at the optimal BDTG cut value.

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

	Description	Requirement
$D_s \rightarrow hh\bar{h}$	$m(hh\bar{h})$	$= m_{D_s} \pm 25$ MeV
$D_s^- \rightarrow KK\pi^-$	$D^0$ veto	$m(KK) < 1840$ MeV
$D_s^- \rightarrow \phi\pi^-$	$m(KK)$ PIDK( $K^+$ ) PIDK( $K^-$ ) PIDK( $\pi^-$ ) $\chi_{FD}^2$ 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	$= 1$

## 358 4 Yields determination

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

364 Due to different mass resolutions, we perform the invariant mass fits simultaneously  
 365 for each data-taking period and each trigger category. We further introduce four  $D_s$  final  
 366 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  
 367 different signal purities. The  $D_s \rightarrow Kh\pi$  category combines the two  $D_s$  decay channels  
 368 with the lowest statistics:  $D_s \rightarrow KK\pi$  (non-resonant) and  $D_s \rightarrow K\pi\pi$ . This amounts to  
 369 16 categories in total.

### 370 4.1 Signal model

371 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  
 372 using a Johnson's SU function [31], which results from a variable transformation of a  
 373 normal distribution to allow for asymmetric tails:

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

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

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

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

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

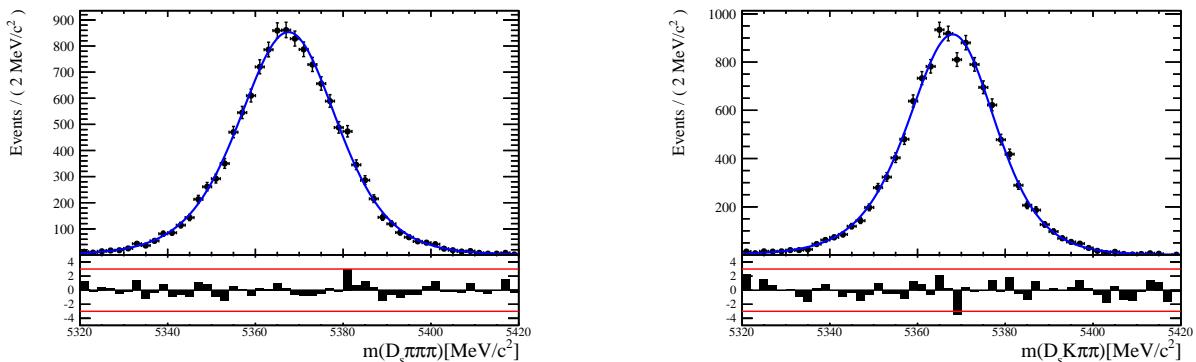


Figure 4.1: Invariant mass distributions of simulated (left)  $B_s^0 \rightarrow D_s \pi\pi\pi$  and (right)  $B_s^0 \rightarrow D_s K\pi\pi$  events. A fit with a Johnson's SU PDF is overlaid.

377 simulation and real data, linear scaling factors for the mean  $\mu$  and width  $\sigma$  are determined  
378 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  
379 factors are determined separately for each data-taking period and each trigger category.  
380

## 381 4.2 Background models

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

### 384 Combinatorial background

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

### 389 Peaking $B_d$ background

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

### 393 Partially reconstructed background

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

405 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  
406 shifted by  $m_{B_s^0} - m_{B^0}$ .

### 408 Misidentified background

409 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  
410 pions is misidentified as a kaon, contaminate the  $B_s^0 \rightarrow D_s K^+ \pi^+ \pi^-$  sample. To determine  
411 the corresponding background shapes, we use simulated events passing the nominal  
412 selection except for the PID cuts on the bachelor  $\pi^+$  tracks. The **PIDCalib** package  
413 is used to determine the  $p_T, \eta$ -dependent  $\pi^+ \rightarrow K^+$  misidentification probability for  
414 each pion. We change the particle hypothesis from pion to kaon for the pion with the  
415 higher misidentification probability and recompute the invariant  $B_s^0$  mass,  $m(D_s^- \pi_K^+ \pi^+ \pi^-)$ .  
416 Similarly, the invariant masses  $m(\pi_K^+ \pi^+ \pi^-)$  and  $m(\pi_K^+ \pi^-)$  are recomputed and required  
417 to be within the considered phasespace region. The background distributions are shown  
418 in Fig. 4.2 (middle,right) and modeled by the sum of two Crystal Ball functions. The  
419 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  
420  $B_s^0 \rightarrow D_s K\pi\pi$  sample is computed by multiplying the fake rate (within the considered  
421  $B_s^0 \rightarrow D_s K\pi\pi$  sample) with the total number of  $B_s^0 \rightarrow D_s K\pi\pi$  events.

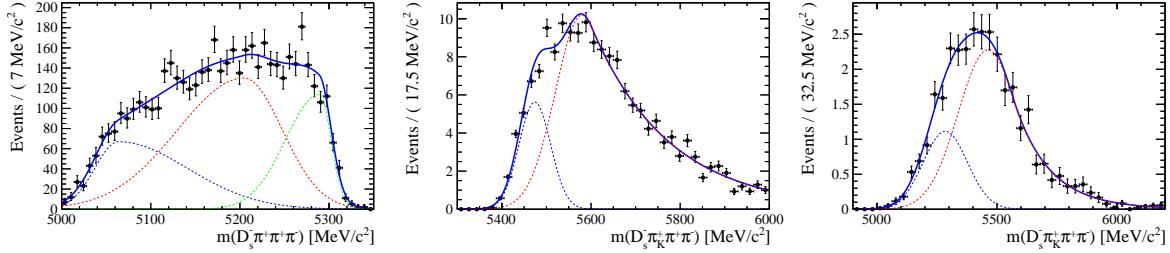


Figure 4.2: Left: Invariant mass distribution of simulated  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  events, where the  $\gamma/\pi^0$  is excluded from the reconstruction. Middle: Invariant mass distribution of simulated  $B_s^0 \rightarrow D_s \pi\pi\pi$  events, where one of the pions is reconstructed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

<sup>422</sup>  $B_s$  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  
<sup>423</sup> determined in the mass fit to the  $B_s^0 \rightarrow D_s \pi\pi\pi$  data sample which is corrected for the  
<sup>424</sup>  $\text{PID}(\pi^+) < 0$  requirement. The  $B_s^0 \rightarrow D_s^* \pi^+\pi^+\pi^-$  yield is additionally corrected for the  
<sup>425</sup> efficiency of the cut  $m(D_s K\pi\pi) > 5200$  MeV evaluated on MC. In the fit to data, the  
<sup>426</sup> misidentified background yields are fixed to the predicted ones.

<sup>427</sup> 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  
<sup>428</sup>  $B_s^0 \rightarrow D_s \pi\pi\pi$  data sample to be negligible due to the low branching fractions and the  
<sup>429</sup> tight PID cuts on the bachelor pions.

### 430 4.3 Results

<sup>431</sup> Figure 4.3 shows the invariant mass distribution for  $B_s^0 \rightarrow D_s \pi\pi\pi$  and  $B_s^0 \rightarrow D_s K\pi\pi$   
<sup>432</sup> candidates passing all selection criteria. The projections for all categories of the simula-  
<sup>433</sup> taneous fit are shown in Appendix C together with the results for all fitted parameters.  
<sup>434</sup> The integrated signal and background yields are listed in Tables 4.1 and 4.2.

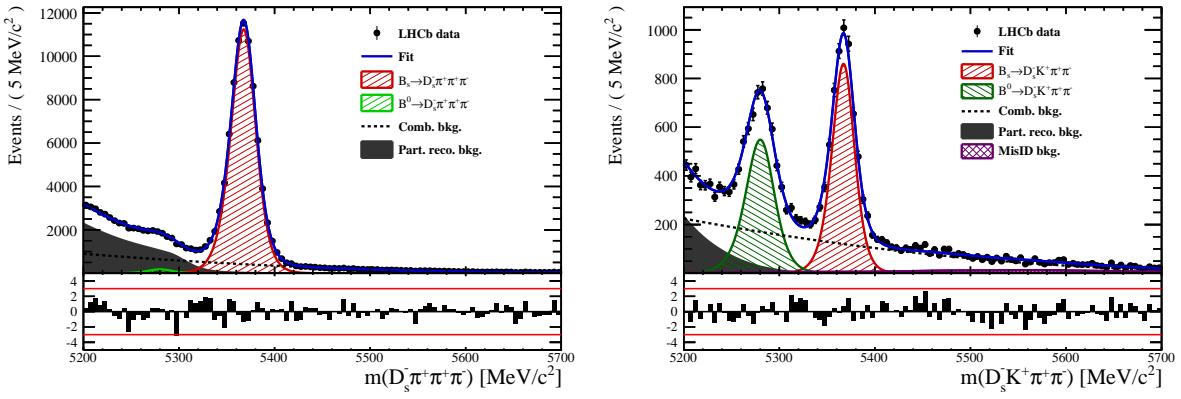


Figure 4.3: Invariant mass distribution of  $B_s^0 \rightarrow D_s \pi\pi\pi$  (left) and  $B_s^0 \rightarrow D_s K\pi\pi$  (right) candidates.

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.	$9376 \pm 177$		

## 435 5 Decay-time Resolution

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

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

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

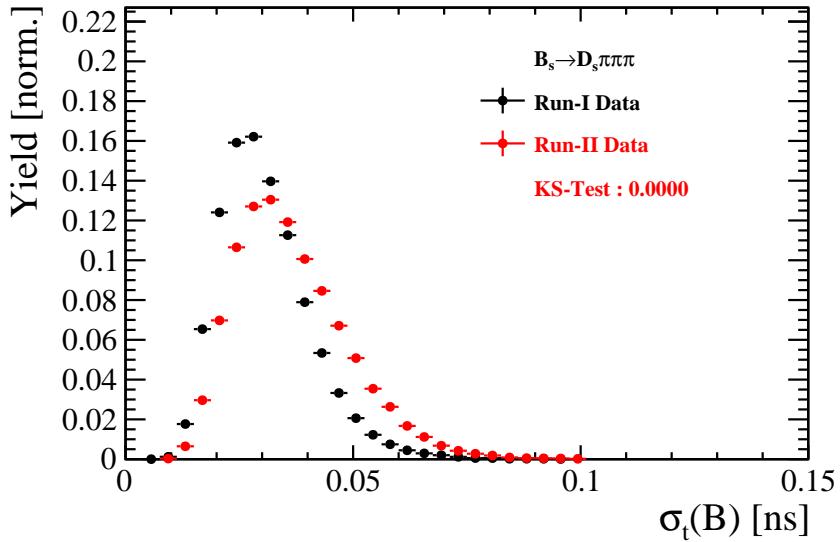


Figure 5.1: Distribution of the decay time error for  $B_s^0 \rightarrow D_s\pi\pi\pi$  signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

## 458 5.1 Calibration for Run-I data

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

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

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

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

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

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

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

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

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

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

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

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

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

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

488 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  
 489 assigned as additional systematic uncertainty.

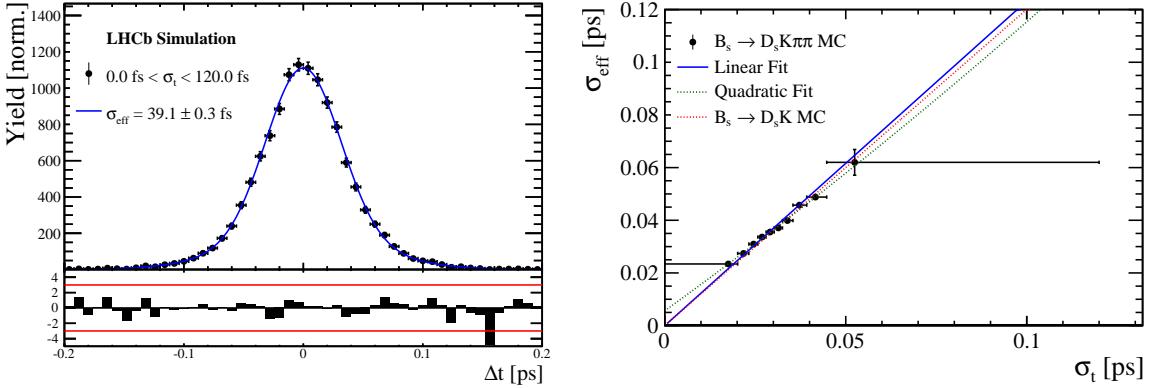


Figure 5.2: (Left) Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K\pi\pi$  MC candidates. (Right) The measured resolution  $\sigma_{\text{eff}}$  as function of the per-event decay time error estimate  $\sigma_t$  for  $B_s \rightarrow D_s K\pi\pi$  MC (Run-I). The fitted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

## 490 5.2 Calibration for Run-II data

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

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

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

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

<sup>517</sup> is in good agreement with the one obtained for the  $B_s \rightarrow J/\psi\phi$  (Run-II) analysis [35].

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

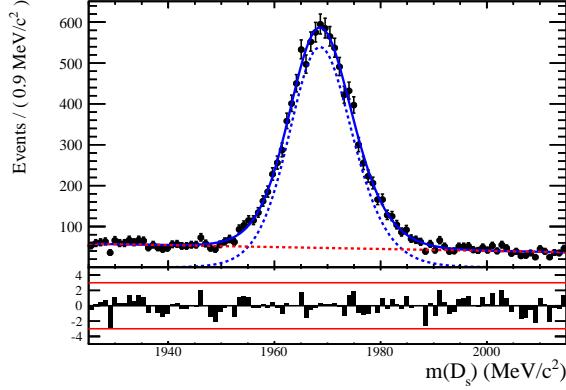


Figure 5.3: The invariant mass distribution for prompt  $D_s$  candidates.

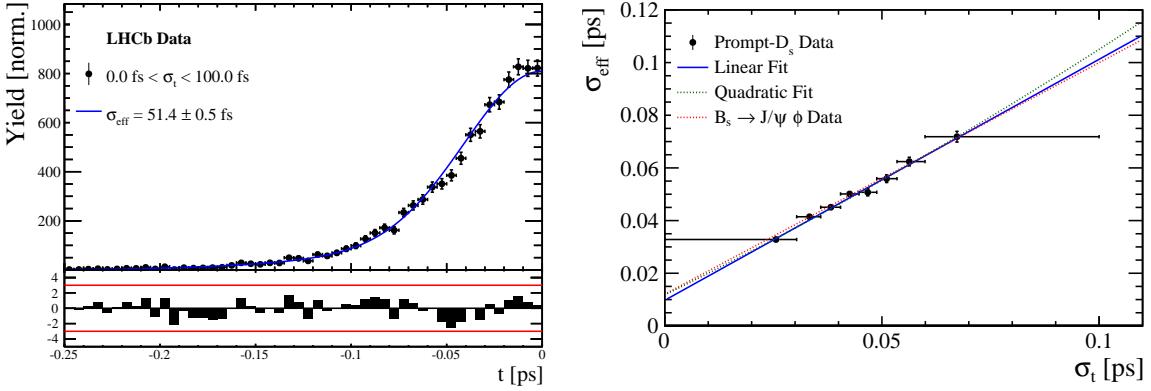


Figure 5.4: (Left) Decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks. (Right) The measured resolution  $\sigma_{eff}$  as function of the per-event decay time error estimate  $\sigma_t$  for fake  $B_s$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

518 **5.3 Cross-checks**

519 **5.3.1 Kinematic dependence**

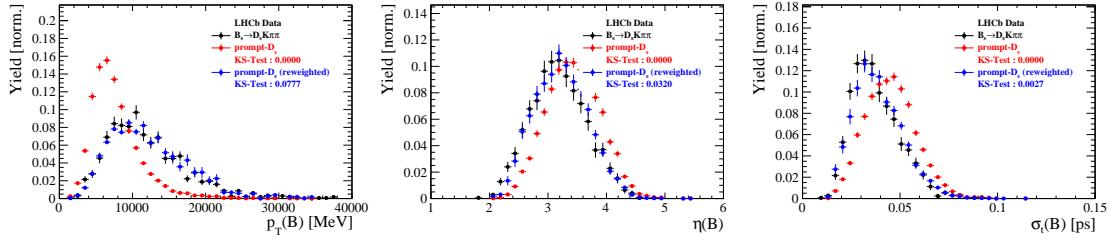


Figure 5.5

520 **5.3.2 DTF constraints**

## 521 6 Acceptance

### 522 6.1 MC corrections

#### 523 6.1.1 Truth matching of simulated candidates

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

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

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

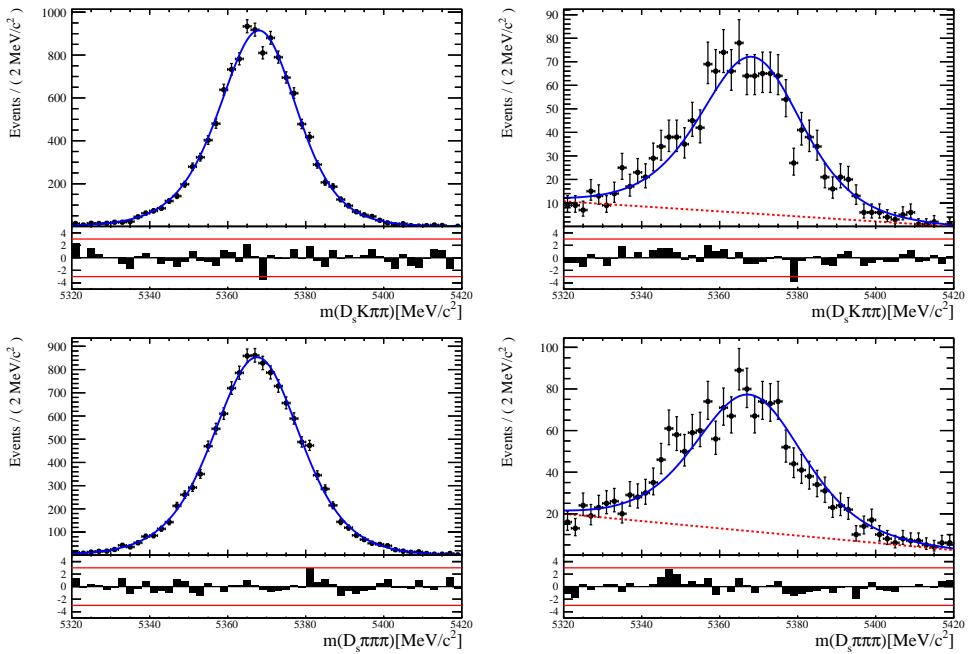


Figure 6.1: The reconstructed  $B_s$  mass distribution for simulated  $B_s \rightarrow D_s K\pi\pi$  (top) and  $B_s \rightarrow D_s \pi\pi\pi$  (bottom) decays classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

543 6.1.2 PID efficiencies

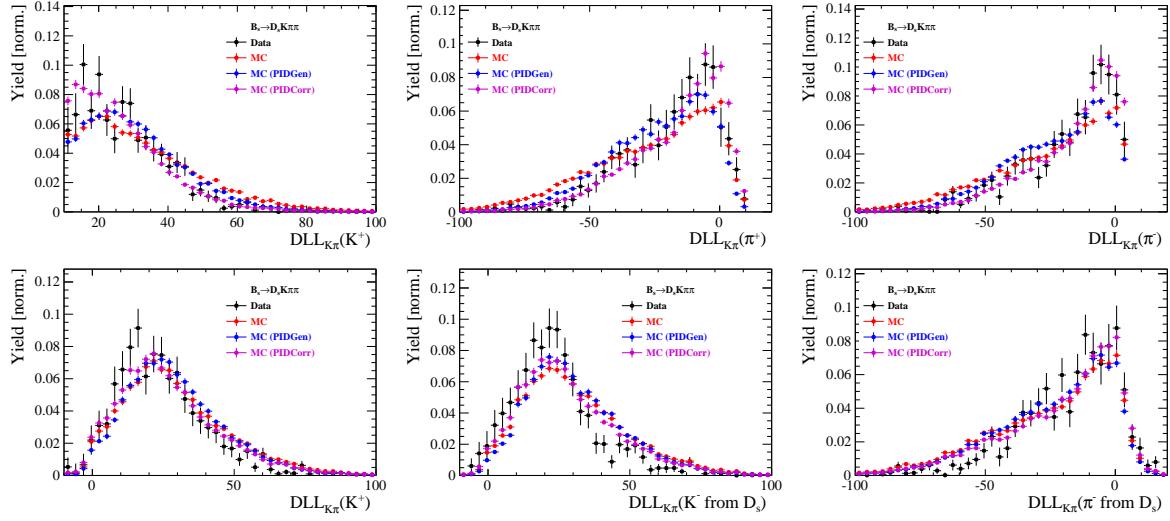


Figure 6.2

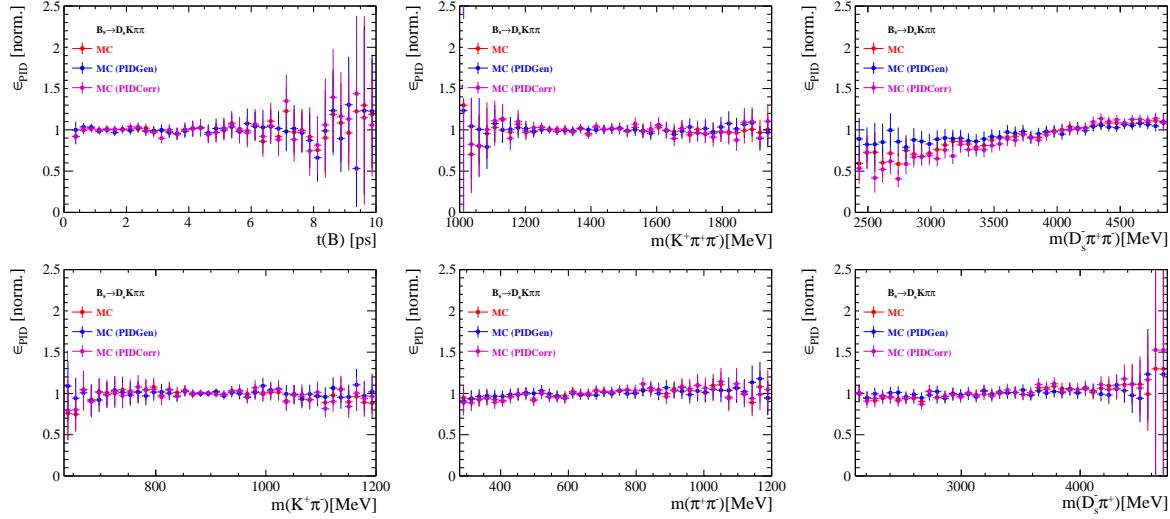


Figure 6.3

544 6.1.3 BDT efficiencies

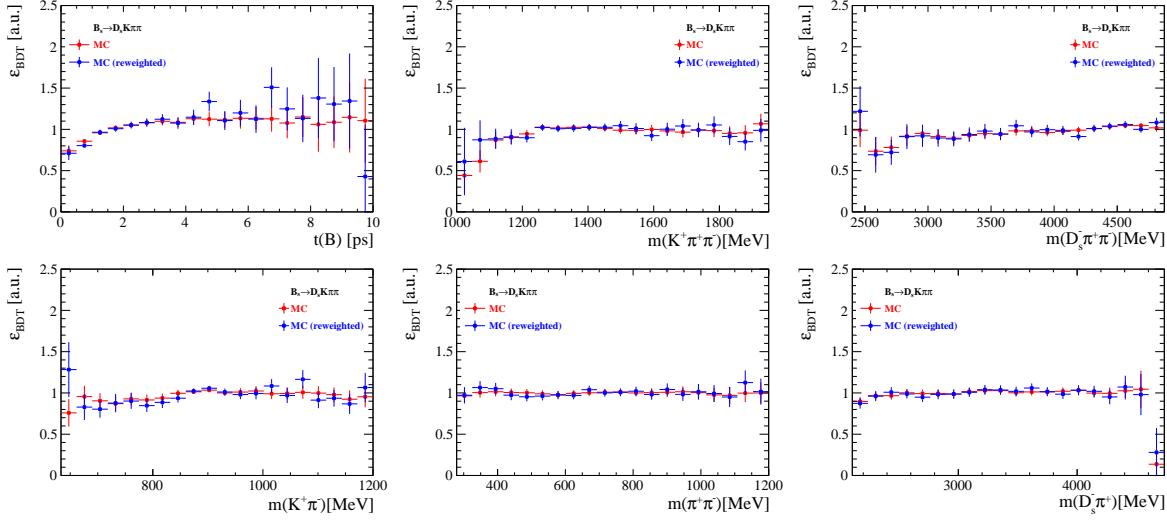


Figure 6.4

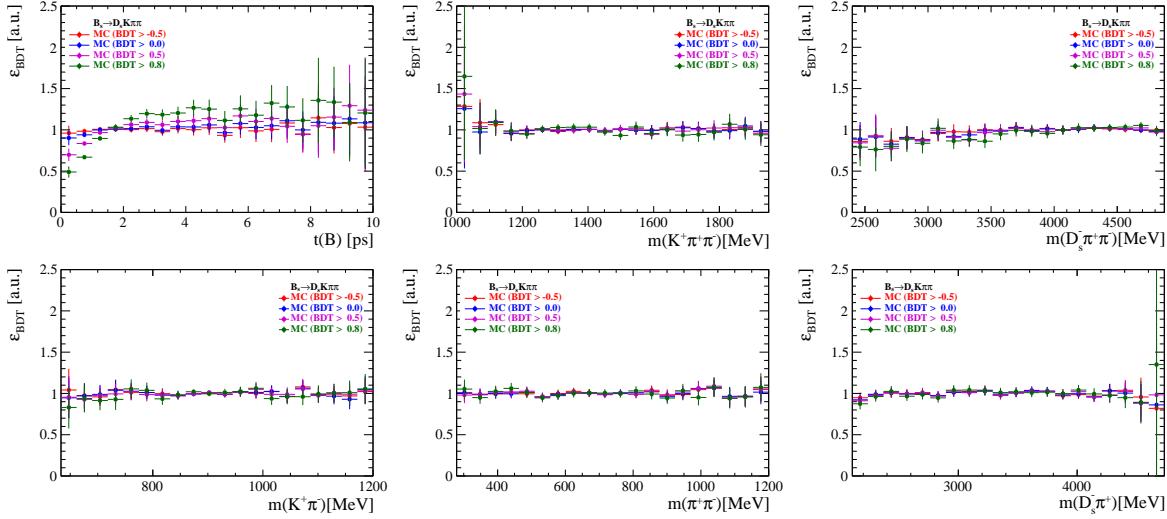


Figure 6.5

<sup>545</sup> 6.1.4 Tracking efficiencies

546 **6.2 Decay-time acceptance**

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

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

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

560 A PDF of the form

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

561 is fit to the decay time distribution of  $B_s^0 \rightarrow D_s \pi\pi\pi$  candidates in data. Since the  
 562 fit is performed untagged, the PDF shown in Eq. 6.2 contains no terms proportional  
 563 to  $\Delta m_s$ . The values for  $\Gamma_s$  and  $\Delta\Gamma_s$  are fixed to the latest HFAG results [36]. The  
 564 decay-time acceptance  $\epsilon(t', \vec{\lambda})$  is modelled using the sum of cubic polynomials  $v_i(t)$ , so  
 565 called Splines [37]. The polynomials are parametrised by so-called knots which determine  
 566 their boundaries. Knots can be set across the fitted distribution to account for local  
 567 changes in the acceptance shape. Using more knots is equivalent to using more base  
 568 splines which are defined on a smaller sub-range. In total,  $n + 2$  base splines  $v_i(t)$  are  
 569 needed to describe an acceptance shape which is parametrised using  $n$  knots.

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

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

578 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.

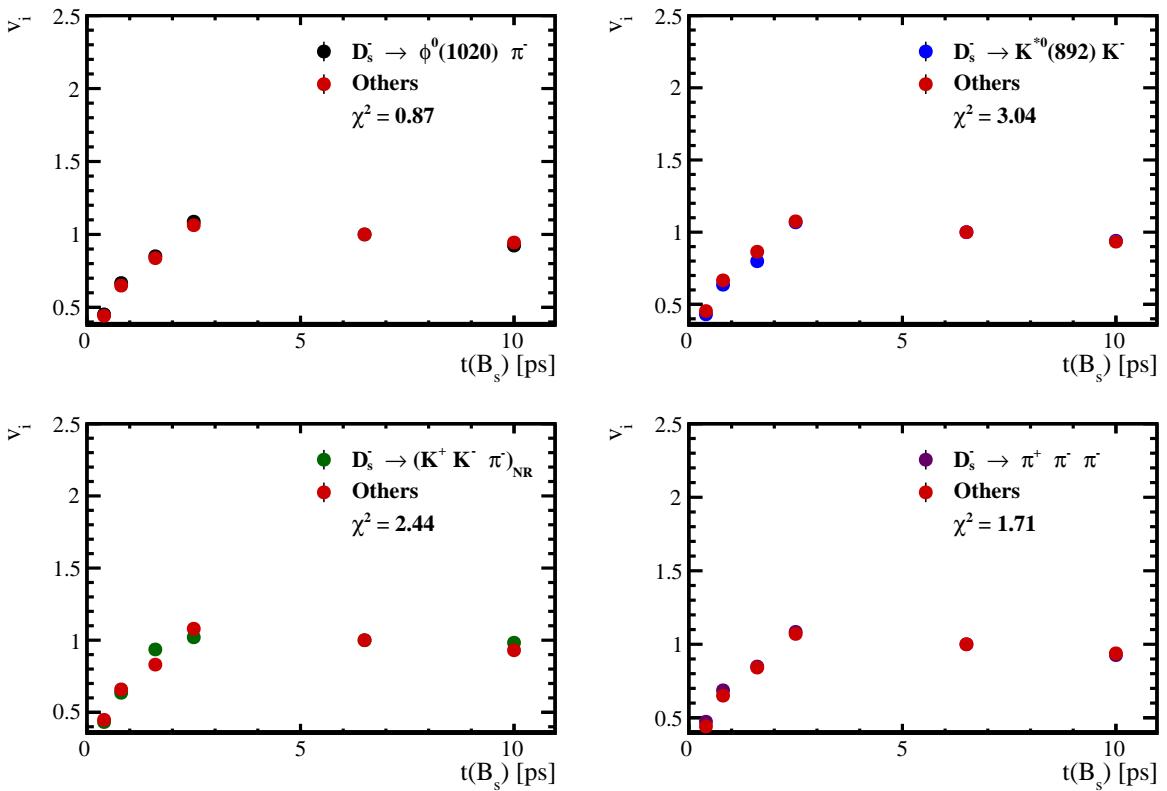


Figure 6.6: Comparison of the spline coefficients obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  subsamples of different  $D_s$  final states. The comparison of one particular  $D_s$  state against all other states is shown.

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

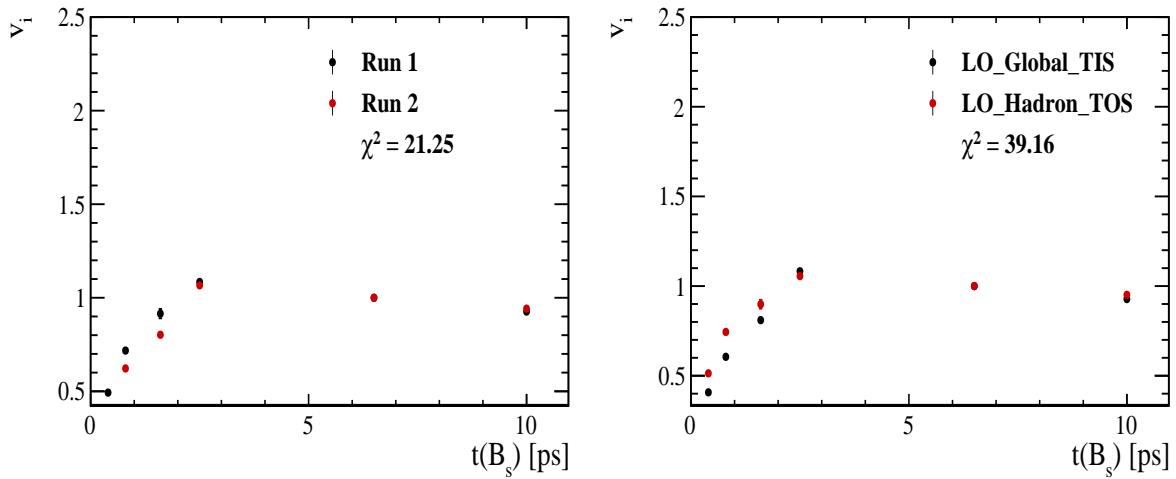


Figure 6.7: Comparison of the spline coefficients obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s \pi \pi \pi$  subsamples of (left) the different runs and (right) L0 trigger categories.

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

601 **6.2.2 Results**

602 The nominal fit to  $B_s^0 \rightarrow D_s\pi\pi\pi$  data using this configuration is shown in Figure ??.  
603 Note that the normalization of the splines in the following figures is not in scale. The fit  
604 parameters obtained from the described fits to data and simulation are summarised in  
605 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)

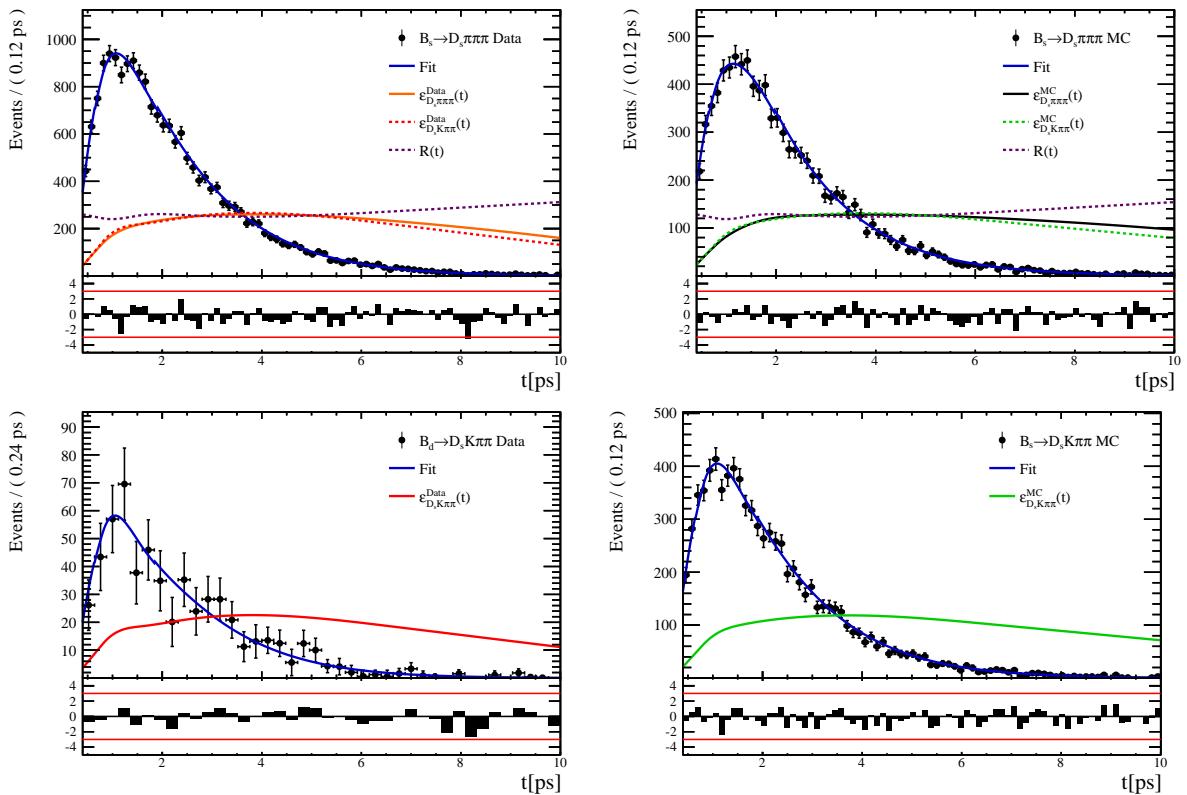


Figure 6.8

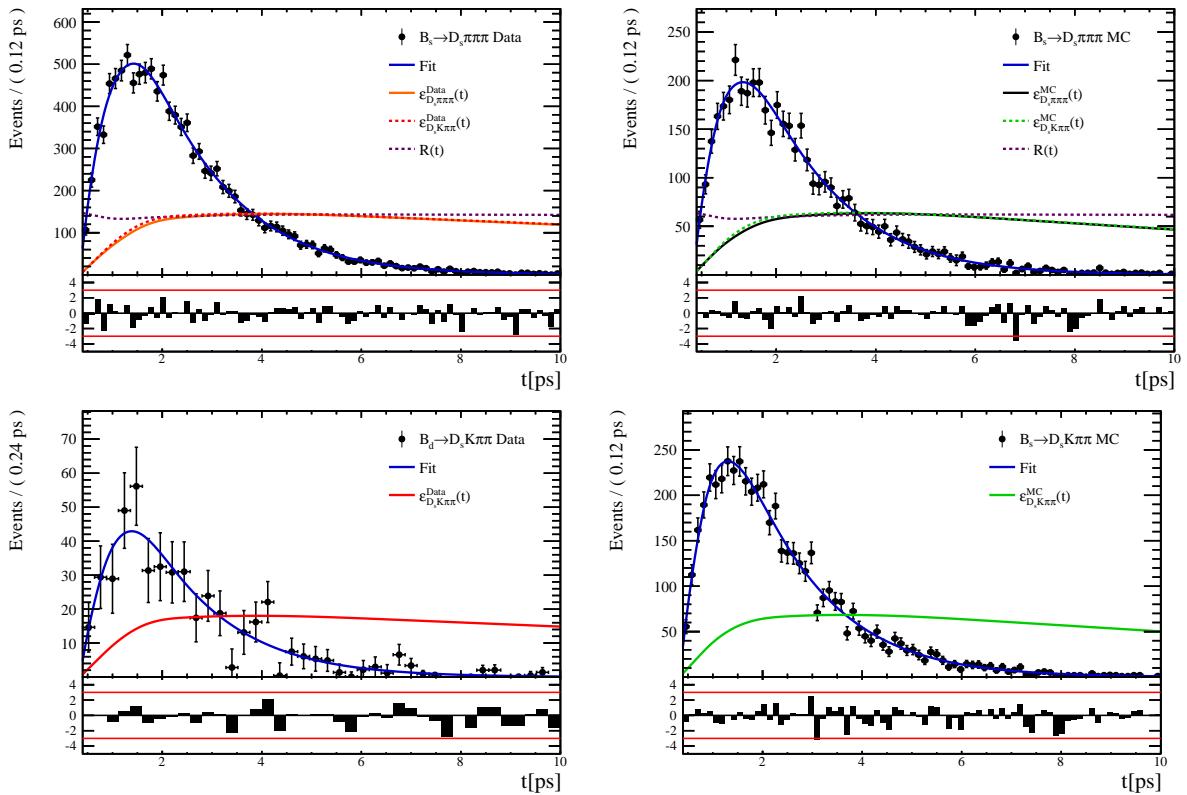


Figure 6.9:

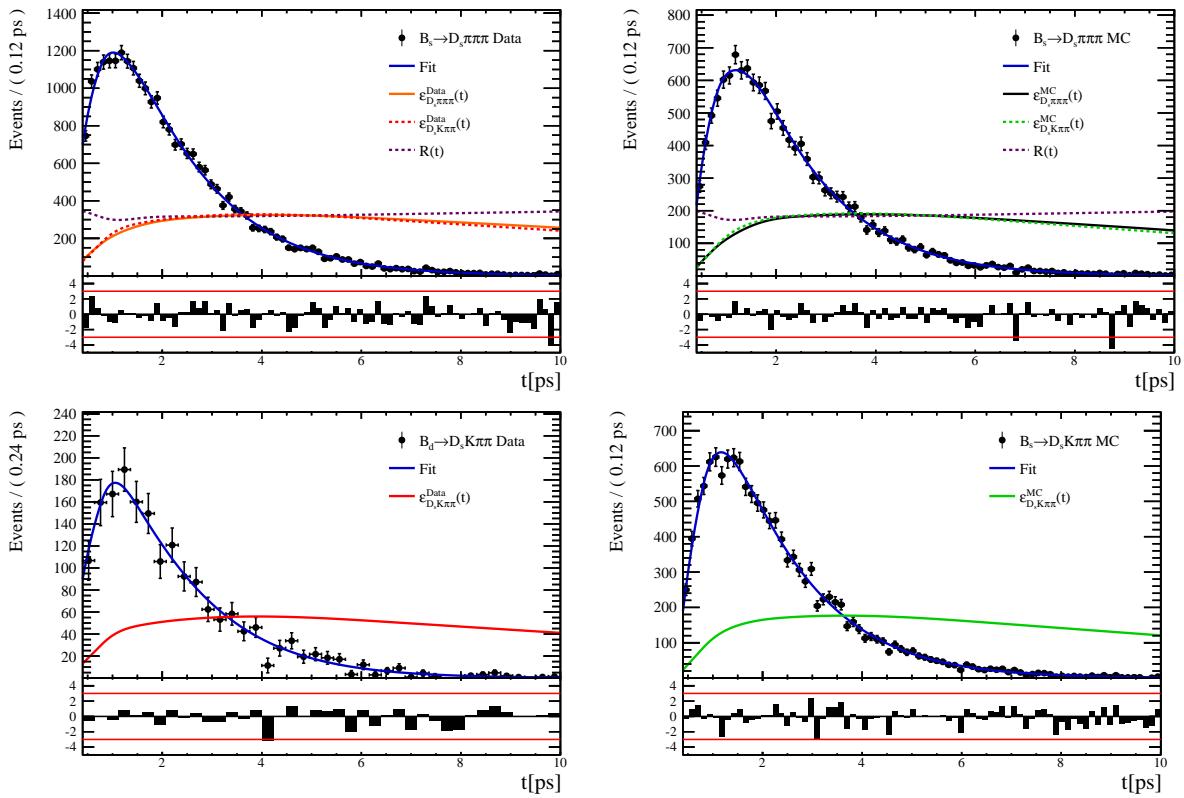


Figure 6.10

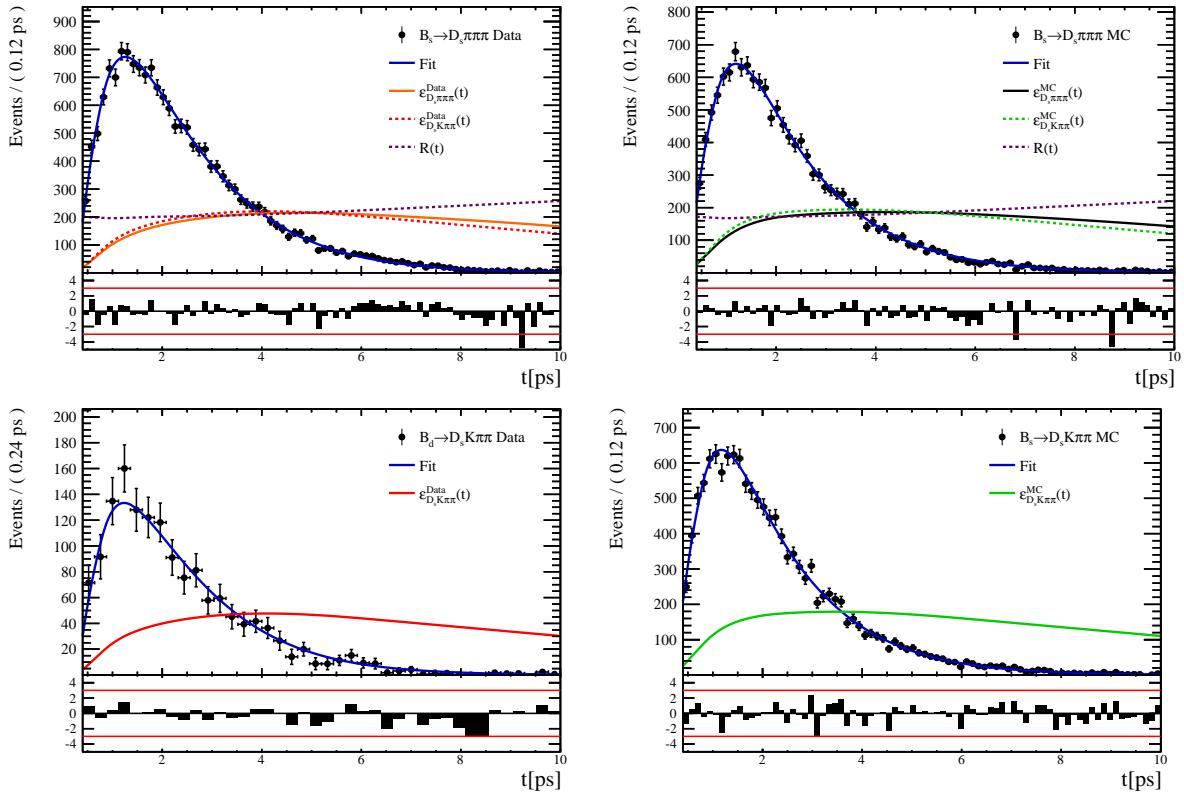


Figure 6.11:

606 **6.3 Phasespace acceptance**

## 607 7 Flavour Tagging

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

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

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

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

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

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

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

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

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

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

647 The tagging calibration is determined using a time-dependent fit to the full  $B_s^0 \rightarrow D_s\pi\pi\pi$   
 648 sample, where the mixing frequency  $\Delta m_s$  is fixed to the nominal PDG value [32]. The  
 649 calibration procedure for the OS tagging algorithms (Sec.7.1) and the SS kaon tagger  
 650 (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,  
 651 which is selected following the steps described in Sec. 3. The similar selection ensures  
 652 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  
 653 terms of the decay kinematics, which are crucial for the flavour tagging. Section 7.3 shows  
 654 the compatibility of both samples. After applying the calibration, the response of the OS  
 655 and SS taggers are combined, which is shown in Sec. 7.4.

## 656 7.1 OS tagging calibration

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

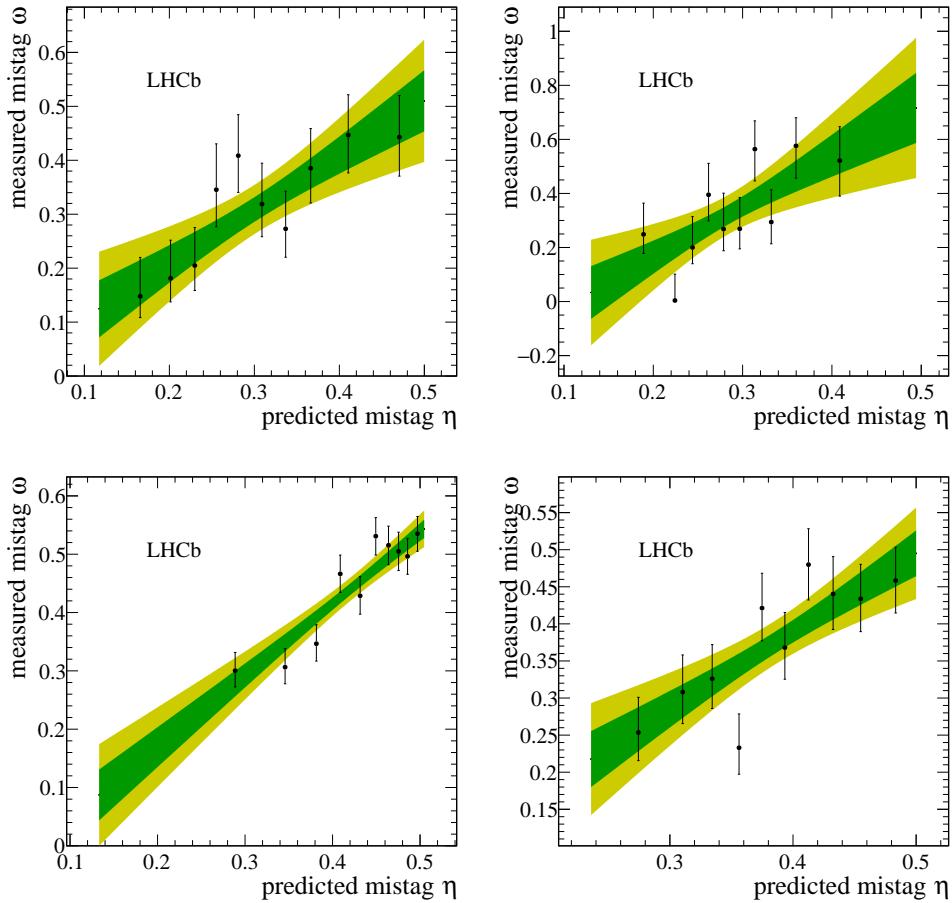


Figure 7.1: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run 1. A linear fit, including the  $1\sigma$  and  $2\sigma$  error bands is overlaid for each tagger.

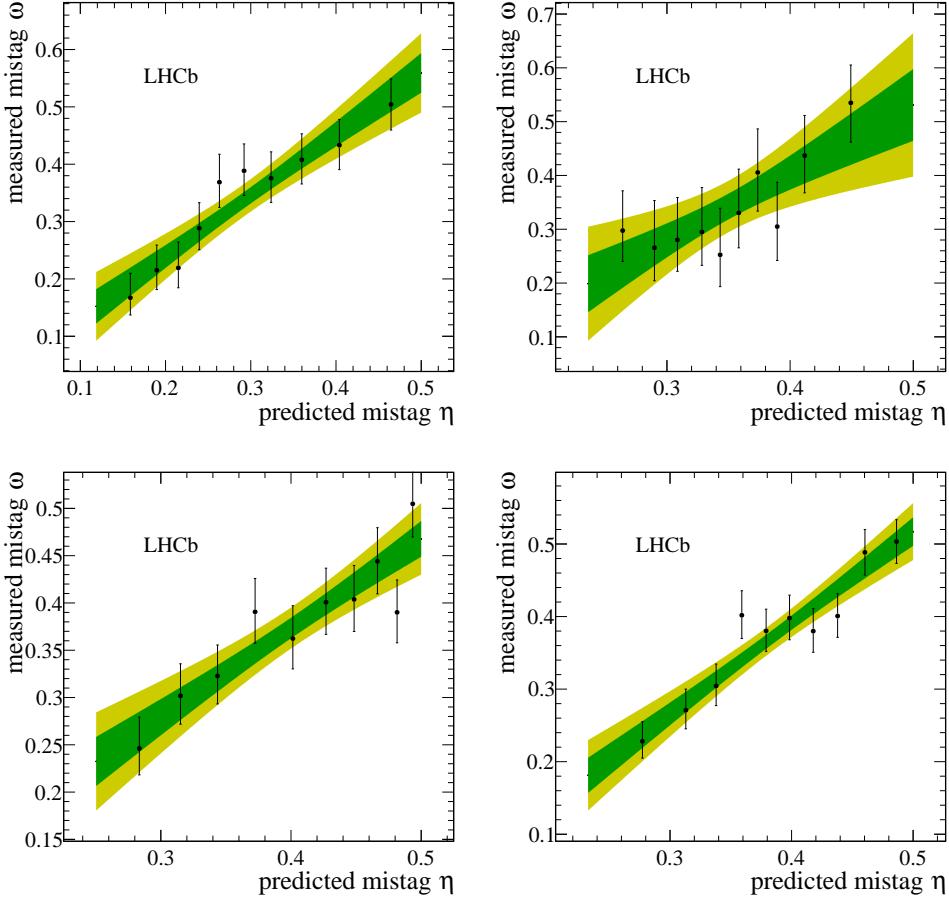


Figure 7.2: Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run 2. A linear fit, including the  $1\sigma$  and  $2\sigma$  error bands is overlaid for each tagger.

## 661 7.2 SS tagging calibration

662 The SS neural net kaon tagger can be calibrated using the flavour-specific  $B_s^0 \rightarrow D_s \pi \pi \pi$   
 663 decay. Its development, performance and calibration is described in detail in [38]. Figure  
 664 7.3 shows the distribution of the predicted mistag of the neural net kaon tagger. The  
 665 extracted calibration parameters and tagging asymmetries are summarized in Table 7.1  
 666 and the measured tagging power for this algorithm is  $\epsilon_{eff,SS} = 3.22\%$ .

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

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

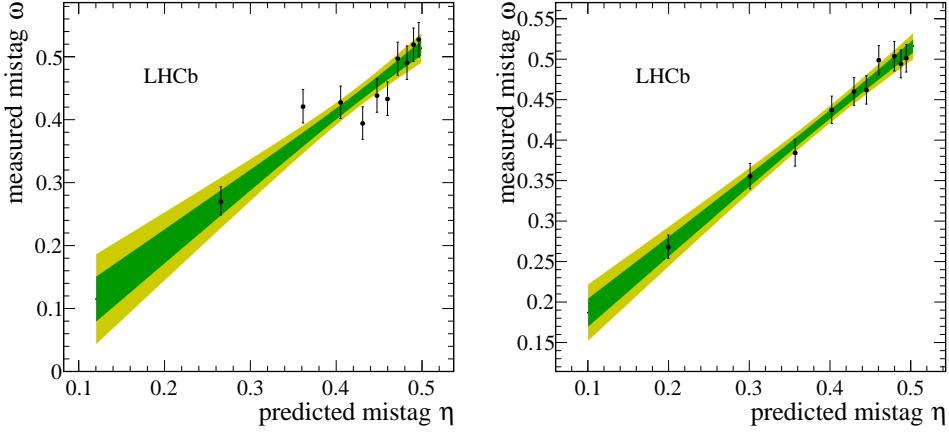


Figure 7.3: Predicted versus measured mistag probability for the SS neural net kaon tagger for (left) Run 1 and (right) Run 2. A linear fit, including the  $1\sigma$  and  $2\sigma$  error bands is overlaid for both distributions.

### 667 7.3 Tagging performance comparison between the signal and 668 normalization channel

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

675 The data distributions of the predicted mistag probability  $\eta$  for the OS combination and  
676 the SS kaon tagger are shown Fig. 7.4.

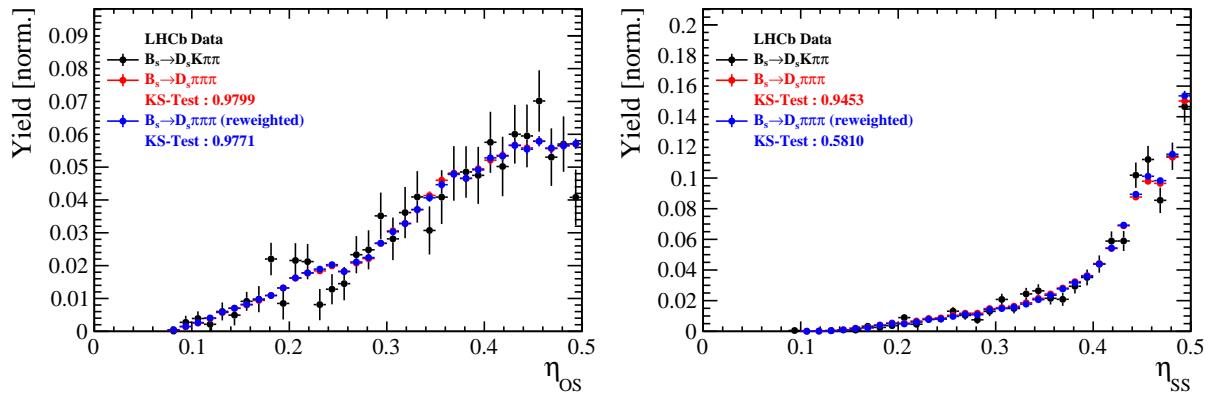


Figure 7.4: Distributions of the predicted mistag  $\eta$  for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the  $B_s^0 \rightarrow D_s K\pi\pi$  (black) and  $B_s^0 \rightarrow D_s \pi\pi\pi$  (red) data samples.

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

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

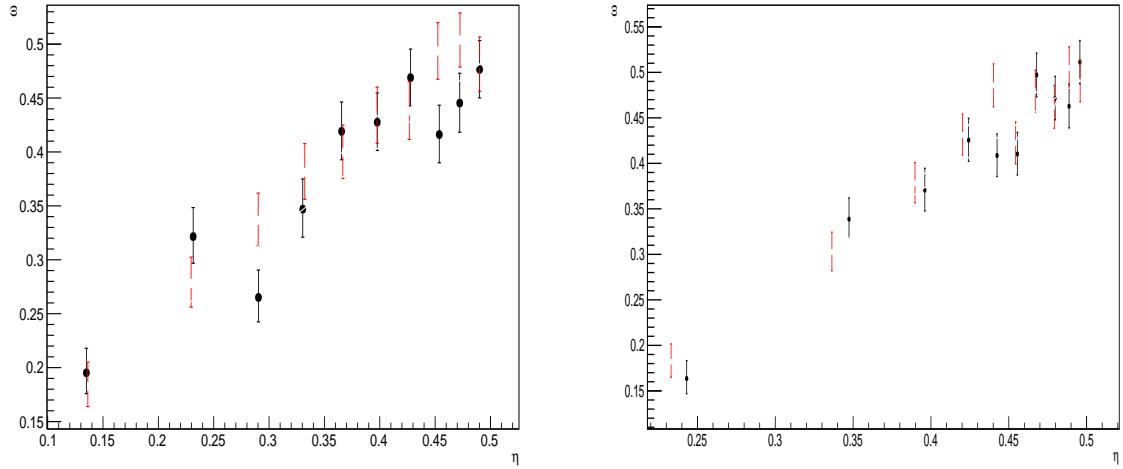


Figure 7.5: Dependence of the observed mistag  $\omega$  on the predicted mistag  $\eta$  for the (left) OS combination and the (right) SS kaon tagger, found in the simulated  $B_s^0 \rightarrow D_s K\pi\pi$  (black) and  $B_s^0 \rightarrow D_s\pi\pi\pi$  (red) signal samples.

## 7.4 Combination of OS and SS taggers

In the time- and amplitude-dependent fit to  $B_s^0 \rightarrow D_s K\pi\pi$  data, the obtained tagging responses of the OS and SS tagger will be combined after the calibration described in the previous sections is applied. Events that acquire a mistag probability greater than 0.5 after the calibration will have their tagging decision flipped. For events where only one of the two taggers fired, the combination of the tagging decision is trivial. In those events where both taggers made a decision, we use the standard combination of taggers [39] provided by the flavour tagging group. In the nominal fit, the calibrated mistags  $\omega$  are combined event by event for the OS and SS tagger, thus adding one observable to the fit procedure. This ensures that the uncertainties of the OS and SS tagging calibration parameters are propagated properly to the combined tagging response for each event.

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

Table 7.2: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-I data.

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
Only OS	$11.32 \pm 0.09$	$37.91 \pm 1.02$	$0.98 \pm 0.14$
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 7.3: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-II data.

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
Only OS	$10.51 \pm 0.07$	$35.32 \pm 0.77$	$1.25 \pm 0.11$
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	$78.55 \pm 0.24$	$39.65 \pm 0.55$	$6.02 \pm 0.30$

## 697 8 Production and Detection Asymmetries

### 698 8.1 $B_s$ Production Asymmetry

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

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

702 where  $\sigma$  are the corresponding production cross-section. This asymmetry was measured  
703 by LHCb in  $pp$  collisions at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV by means of a time-dependent  
704 analysis of  $B_s \rightarrow D_s^- \pi^+$  decays [40]. The results in bins of  $p_T$  and  $\eta$  of the  $B_s$  meson  
705 are shown in Table 8.1. To correct for the different kinematics of  $B_s \rightarrow D_s^- \pi^+$  and  
706  $B_s^0 \rightarrow D_s K \pi \pi$  decays, the measured  $B_s$  production asymmetries  $A_p(p_T, \eta)$  are folded with  
707 the sWeighted  $p_T, \eta$  distribution of our signal channel. The resulting effective production  
708 asymmetries are:

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

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

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

709 As for Run-II data no measurement is available yet, we determine the production asym-  
710 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. [40]

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

## 711 8.2 $K^-\pi^+$ Detection Asymmetry

712 The presented measurement of the CKM-angle  $\gamma$  using  $B_s^0 \rightarrow D_s K\pi\pi$  decays is sensitive to  
 713 a possible charge asymmetry of the kaon. Kaons are known to have a nuclear cross-section  
 714 which is asymmetrically dependent on the sign of their charge. It is indispensable to  
 715 determine the charge asymmetry of the kaon, as fitting without taking this effect into  
 716 account would introduce a 'fake' CP violation. Instead of determining the single track  
 717 detection asymmetry of a kaon, it is found that the combined two track asymmetry of a  
 718 kaon-pion pair is much easier to access [41]. Therefore, the two track asymmetry defined  
 719 as

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

720 is used.

721 This asymmetry can be measured from the difference in asymmetries in the  $D^+ \rightarrow$   
 722  $K^-\pi^+\pi^+$  and  $D^+ \rightarrow K_s^0\pi^+$  modes [42]:

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

723 where possible CP violation in the  $D^+ \rightarrow K_s^0\pi^+$  mode is predicted to be smaller than  
 724  $10^{-4}$  in the Standard Model [43]. The asymmetry in the neutral kaon system,  $A(K^0)$ , has  
 725 to be taken into account as a correction.

726 We use a dedicated LHCb tool to determine  $A^{det}(K^-\pi^+)$  for all data taking periods  
 727 used in this analysis. A detailed description can be found in [42]. The tool provides  
 728 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  
 729 determine the asymmetry following Eq. 8.6. Several weighting steps are performed to  
 730 match the kinematics of the calibration samples to our signal decay sample:

731 First, weights are assigned to the  $K^\pm$  and  $\pi^\pm$  of the  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  sample, using  
 732  $p, \eta$  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  
 733 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  
 734 the corresponding, weighted distributions of the  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  sample. In a last  
 735 step, weights are assigned to match the bachelor pions  $\phi$  distributions between the two  
 736 calibration samples.

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

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

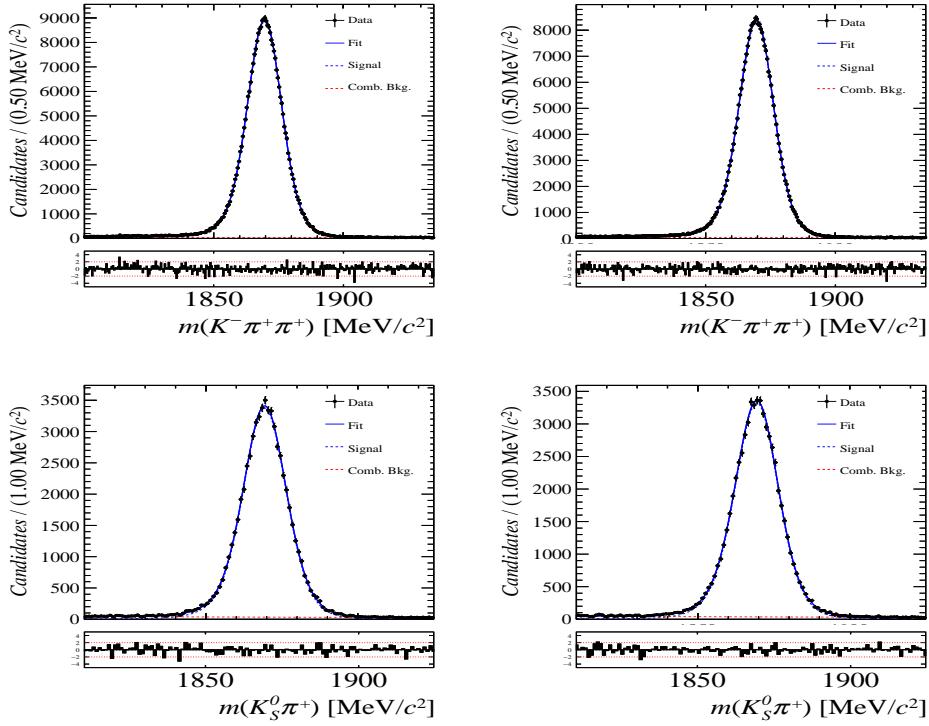


Figure 8.1: Distributions of the invariant mass of (top)  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  and (bottom)  $D^\pm \rightarrow K_S^0\pi^\pm$  candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

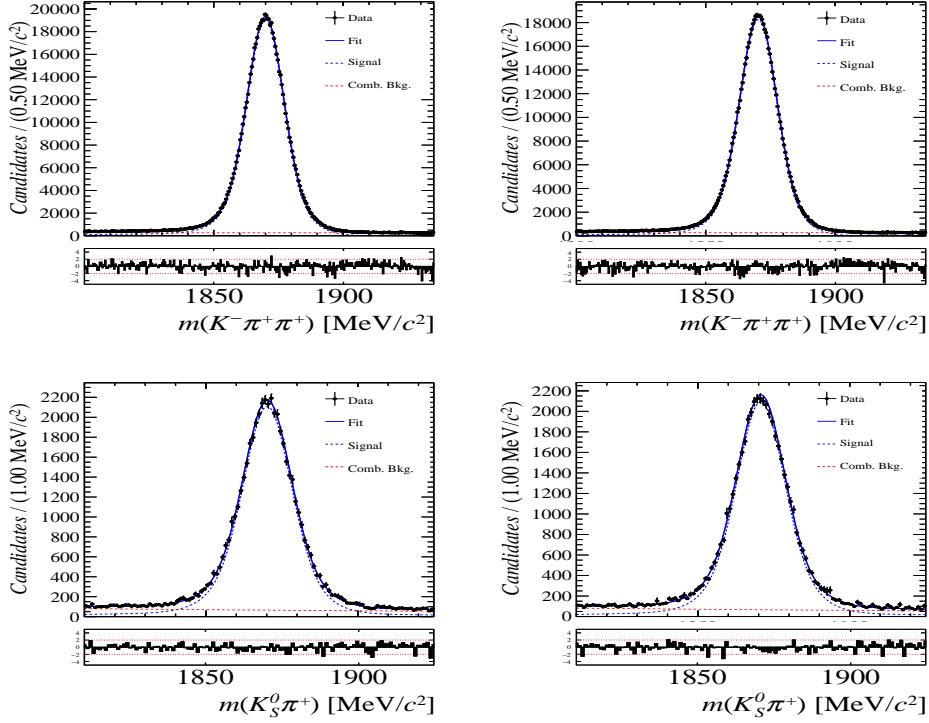


Figure 8.2: Distributions of the invariant mass of (top)  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  and (bottom)  $D^\pm \rightarrow K_S^0\pi^\pm$  candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

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.

## 748 9 Time dependent fit

749 This section covers the phasespace integrated, time-dependent fit to  $B_s^0 \rightarrow D_s h\pi\pi$  data.  
 750 We use the **sFit** technique [44] to statistically remove background from the decay time fit,  
 751 leaving only the signal PDF to describe the decay time. The **sWeights** are calculated based  
 752 on the fit to the reconstructed  $B_s$  mass distribution described in Sec. 4. As additional  
 753 input to the fit, the tagging information (Sec. 7), as well as the decay time acceptance  
 754 (Sec. 6) and resolution (Sec. 5) is used and fixed to the values obtained by the dedicated  
 755 studies. Taking all inputs into account, the final time dependent fit PDF is given by

$$756 \quad \mathcal{P}\mathcal{D}\mathcal{F}(t, \vec{\lambda}) = \left( \epsilon(t) \cdot \int P(x, t, q_t, q_f) dx \right) \times \mathcal{R}(t - t'), \quad (9.1)$$

756 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  
 757 acceptance effects and  $\mathcal{R}(t - t')$  is the Gaussian time resolution function.

### 758 9.1 sFit to $B_s^0 \rightarrow D_s \pi\pi\pi$ data

759 The phasespace-integrated, time-dependent fit is performed to the full sWeighted sample  
 760 of selected candidates from Run I and 2015+2016 Run II data, containing both possible  
 761 magnet polarities and all  $D_s$  final states. In the fit, the values of  $\Gamma_s$  and  $\Delta\Gamma_s$  are fixed to  
 762 the latest PDG report [12]. All tagging parameters are fixed to the central values found in  
 763 the tagging calibration, described in Sec. 7. Due to the fact that the  $B_s^0 \rightarrow D_s \pi\pi\pi$  decay  
 764 is flavour specific, the CP-coefficients can be fixed to  $C = 1$  and  $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$ ,  
 765 reducing Eq. 2.6 to

$$766 \quad \int P(x, t, q_t, q_f) dx = [\cosh\left(\frac{\Delta\Gamma}{2}t\right) + q_t q_f \cos(m_s t)] e^{-\Gamma t}. \quad (9.2)$$

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

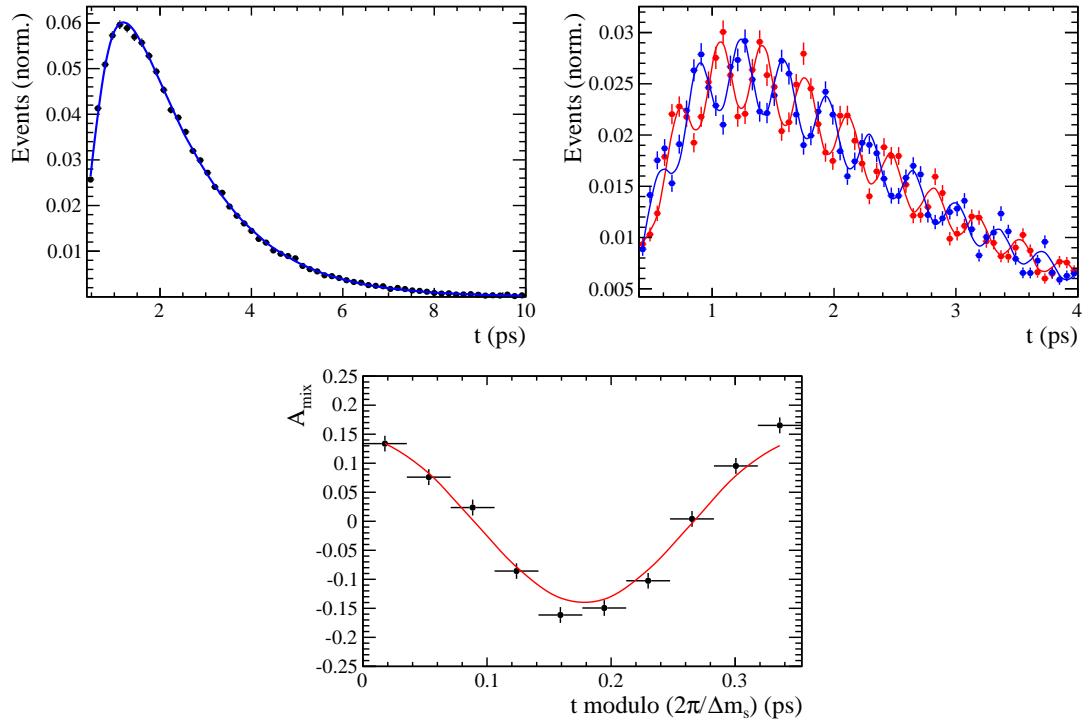


Figure 9.1: Top left: Flavour averaged decay time distribution of  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates with the averaged fit overlaid. Top right: Tagged decay time distribution of mixed (red) and unmixed (blue) signal candidates with the fit described in the text overlaid. Bottom: Time-dependent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

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	$p_0^{\text{OS}}$	$0.3669 \pm 0.0074$
	$p_1^{\text{OS}}$	$0.9298 \pm 0.0761$
	$\Delta p_0^{\text{OS}}$	$0.0118 \pm 0.0085$
	$\Delta p_1^{\text{OS}}$	$0.0234 \pm 0.0855$
	$\epsilon_{tag}^{\text{OS}}$	$0.3525 \pm 0.0023$
	$\Delta \epsilon_{tag}^{\text{OS}}$	$0.0105 \pm 0.0085$
	$p_0^{\text{SS}}$	$0.4532 \pm 0.0055$
	$p_1^{\text{SS}}$	$0.9125 \pm 0.0656$
	$\Delta p_0^{\text{SS}}$	$-0.0123 \pm 0.0060$
	$\Delta p_1^{\text{SS}}$	$0.1374 \pm 0.0757$
	$\epsilon_{tag}^{\text{SS}}$	$0.6804 \pm 0.0023$
	$\Delta \epsilon_{tag}^{\text{SS}}$	$0.0076 \pm 0.0083$
	$A_p$	$-0.0042 \pm 0.0091$
	$\Delta m_s$	$\text{xx.xx} \pm 0.0110$

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

The time-dependent fit to the sWeighted sample of  $B_s^0 \rightarrow D_s K\pi\pi$  signal candidates is performed simultaneously in the four bins defined in Sec. 6.2.1, splitting the data into Run I & II and trigger category 0 (L0Hadron TOS) & 1 (L0Hadron TIS). In these four bins, the respective description of the decay-time acceptance (Sec. 6) is used as an input. As further input the decay-time resolution scaling relation, found separately for Run I & II in Sec. 5, is used in the simultaneous fit. The full fit model is given in Eq. 9.1, where  $\int P(x, t, q_t, q_f)$  is:

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

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

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 $\pm$ 0.116
$D$	xx.xx $\pm$ 0.287
$\bar{D}$	xx.xx $\pm$ 0.269
$S$	xx.xx $\pm$ 0.162
$\bar{S}$	xx.xx $\pm$ 0.163

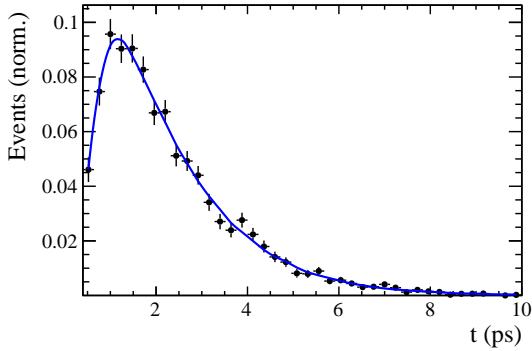


Figure 9.2: Tagged and sWeighted decay-time distribution of  $B_s^0 \rightarrow D_s K\pi\pi$  signal candidates. The fit described in the text is overlaid.

## 9.3 sFit model validation using toy studies

The fit model and procedure is validated using pseudo experiments. 1000 toys are generated using the model described in Eq. 9.1 and 2.6. Each pseudo experiment is generated with

785 the same amount of signal events found in the Run I + 2015/2016 data samples. Figure  
 786 9.3 shows the pull distributions for all CP coefficients, where every pull  $P$  of a parameter  
 787  $x$  is given as  $P = \frac{x_{gen} - x_{fit}}{\Delta x}$ .

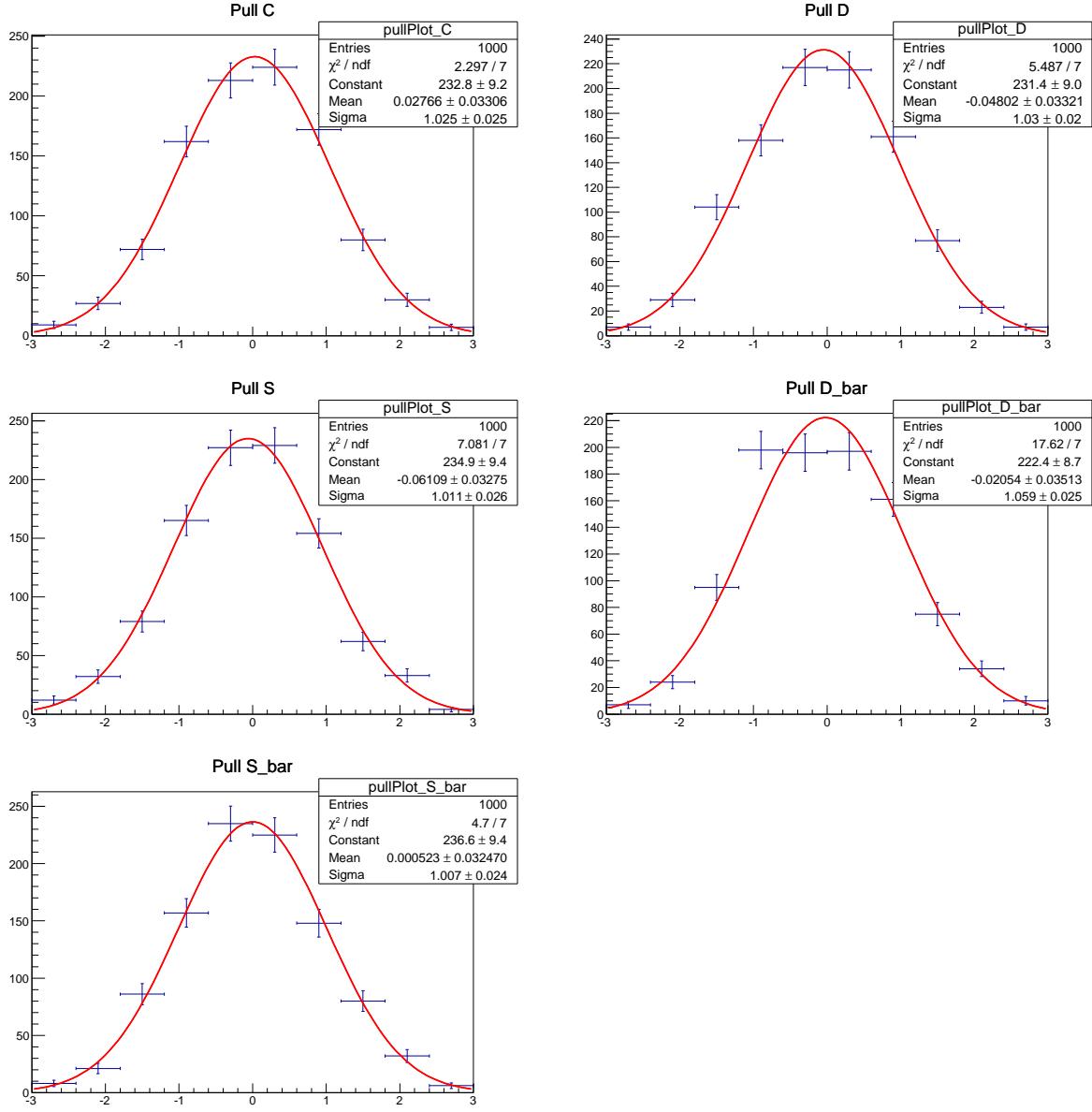


Figure 9.3: Pull distributions from toy studies for the time-dependent fit, done with 1000 pseudo experiments.

788 Table 9.3 summarizes the means  $\mu$  and widths  $\sigma$  of these pull distributions.

Table 9.3: Pull parameters for CP coefficients from the toy studies for the time-dependent fit.

Parameter	$\mu$ of pull distribution	$\sigma$ of pull distribution
C	$0.0276566 \pm 0.0330601$	$1.02492 \pm 0.0251927$
D	$-0.0480232 \pm 0.0332088$	$1.02998 \pm 0.0237533$
S	$-0.0610883 \pm 0.0327533$	$1.01134 \pm 0.0256215$
$\bar{D}$	$-0.0205393 \pm 0.0351271$	$1.05858 \pm 0.0252034$
$\bar{S}$	$0.000523013 \pm 0.0324702$	$1.00746 \pm 0.0244268$

## 789 10 Time dependent amplitude fit

### 790 10.1 Signal Model Construction

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

797 To build the amplitude model, one could successively add amplitudes on top of one  
798 another until a reasonable agreement between data and fit was achieved. However, this  
799 step-wise approach is not particularly suitable for amplitude analyses as discussed in  
800 Ref. [45]. Instead, we include the whole pool of amplitudes in the first instance and use  
801 the Least Absolute Shrinkage and Selection Operator [45, 46] (LASSO) approach to limit  
802 the model complexity. In this method, the event likelihood is extended by a penalty term

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

803 which shrinks the amplitude coefficients towards zero. The amount of shrinkage is  
804 controlled by the parameter  $\lambda$ , to be tuned on data. Higher values for  $\lambda$  encourage sparse  
805 models, *i.e.* models with only a few non-zero amplitude coefficients. The optimal value  
806 for  $\lambda$  is found by minimizing the Bayesian information criteria [47] (BIC),

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

807 where  $N_{\text{Sig}}$  is the number of signal events and  $r$  is the number of amplitudes with a decay  
808 fraction above a certain threshold. In this way, the optimal  $\lambda$  balances the fit quality  
809 ( $-2 \log \mathcal{L}$ ) against the model complexity. The LASSO penalty term is only used to select  
810 the model. Afterwards, this term must be discarded in the final amplitude fit with the  
811 selected model, otherwise the parameter uncertainties would be biased.

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

- 816 1. The fit fraction threshold for inclusion in the final model was varied within the  
817 interval [0.05, 5]%. The set of selected amplitudes is stable for thresholds between  
818 0.1% and 1%. Other choices result in marginally different models containing one  
819 component more or less.
  - 820 2. Instead of BIC, the Akaike information criteria ( $\text{AIC}(\lambda) = -2 \log \mathcal{L} + 2r$  [48]) was  
821 used to optimize  $\lambda$ . For a given threshold, the AIC method tends to prefer lower  
822  $\lambda$  values. However, the set of models obtained varying the threshold within the  
823 interval [0.05, 5]% is identical to the BIC method.
  - 824 3. The amplitudes selected under nominal conditions were excluded one-by-one from  
825 the set of all amplitudes considered.
- 826 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	$99.13 \pm 5.87$

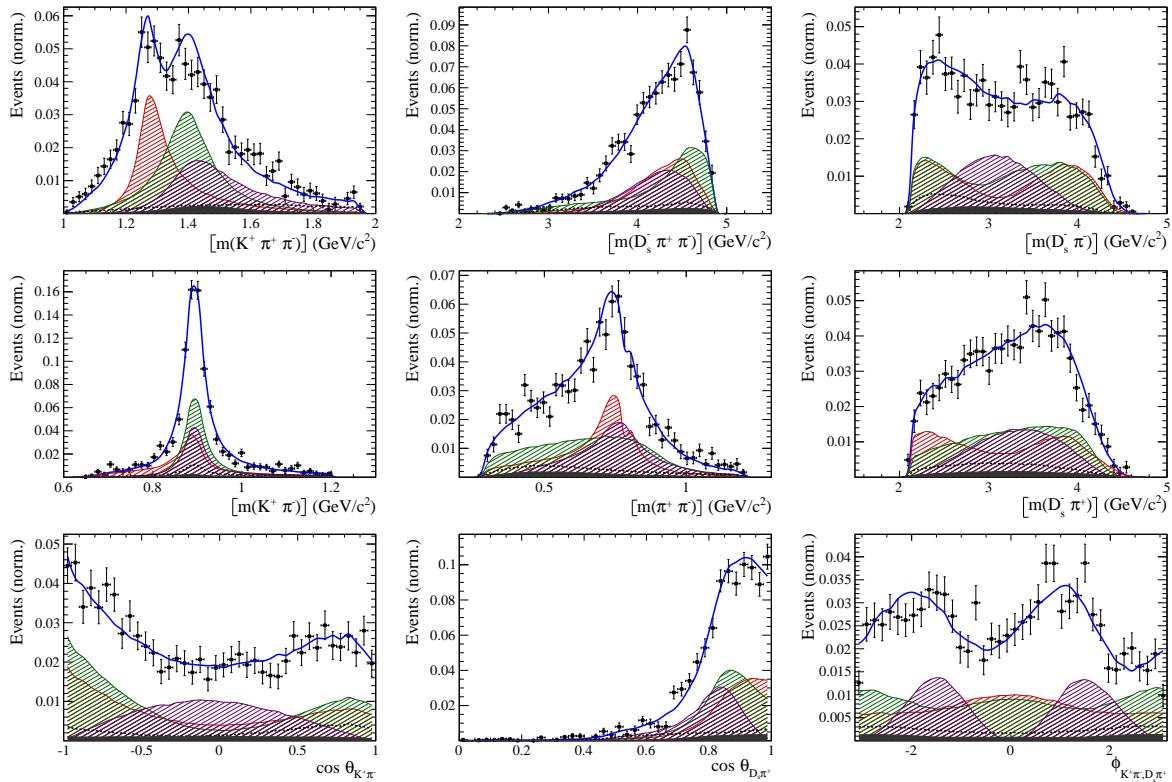


Figure 10.1

## 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

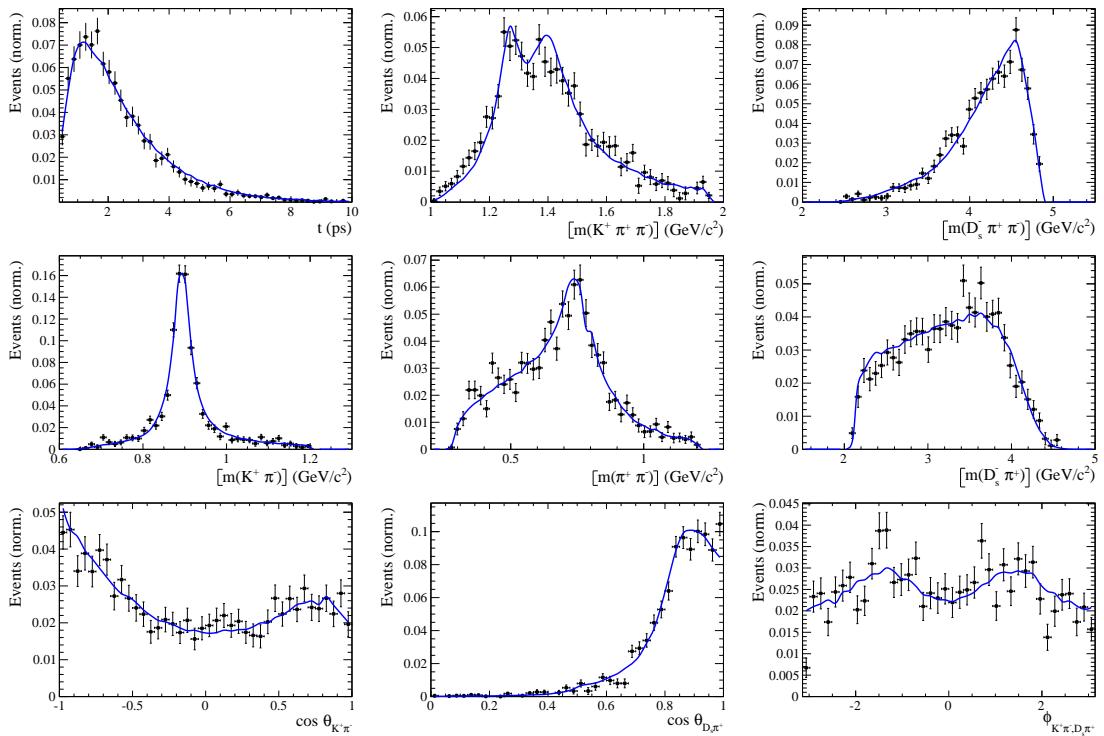


Figure 10.2

## 828 11 Systematic uncertainties

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

### 836 11.1 Models for $B_s^0$ mass distribution

837 The statistical subtraction of the residual background [44], left after the full selection,  
 838 relies on the correct description of the invariant  $B_s^0$  mass distribution. Since the choice  
 839 of signal and background models is not unique, alternative descriptions which lead to  
 840 slightly different yields for the signal and background components are available. The  
 841 difference in yields could result in shifted values for the measured observables and are  
 842 therefore treated as systematic uncertainty.

843

#### 844 11.1.1 Signal model

845 The Johnson's SU function which is used as nominal signal model is replaced by a double  
 846 Crystal Ball [49]. The crystal ball model is given by a gaussian core with an exponential  
 847 tail on one side. Choosing a double Crystal Ball allows for asymmetric tails in a slightly  
 848 different way compared to the Johnson's SU function.

#### 849 11.1.2 Background model

850 For the description of the partially reconstructed background, a combination of the  
 851 RooHILLdini and RooHORNsdini model [50] is used instead of the nominal model of  
 852 three bifurcated gaussians. The HORNsdini model is used to describe the  $B_s^0 \rightarrow D_s^* [\rightarrow$   
 853  $D_s(\pi^0)] X_{s/d}$  decay, where the brackets around the  $\pi^0$  indicate that it is missed in the  
 854 reconstruction. The  $D_s^* \rightarrow D_s\pi^0$  decay is a Vector  $\rightarrow$  Scalar-Scalar ( $1^- \rightarrow 0^-0^-$ )  
 855 transition. Using the helicity of the  $D_s$ , one can show that this results in a double-peak  
 856 structure in the reconstructed  $B_s^0$  mass. Therefore, the HORNsdini shape consists of a  
 857 gaussian-like double-peak structure:

$$858 HORN S(m_{B_s^0}) = \int_a^b dm_{B_s^0} \left( m_{B_s^0} - \frac{a+b}{2} \right)^2 \mathcal{D}\mathcal{G}(m_{B_s^0} | \mu, \sigma, f_G) \left( \frac{1-\zeta}{b-a} m_{B_s^0} + \frac{b\zeta - a}{b-a} \right), \quad (11.1)$$

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

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

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

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

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

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

### 11.1.3 Description of misidentified background

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

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

For the shape of the mis-ID background, the nominal approach is to use a simulated sample of  $B_s^0 \rightarrow D_s^- \pi^+ \pi^- \pi^+$  or  $B_s^0 \rightarrow D_s^{*-} \pi^+ \pi^- \pi^+$  decays and flip the mass hypothesis of the  $\pi^+$  with the higher misidentification probability (see Sec. 4). The resulting  $m(D_s^{(*)}\pi_K\pi\pi)$  distribution is then modelled and the shape obtained from the fit is used in the nominal mass fit to signal. This approach is modified as follows:

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

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

### 11.1.4 Systematic effect on observables

The shift of the central values of the observables in the full fit when using sWeights obtained from a combination of alternative models, as well as using only one alternative model for the signal/comb.background/part.reco.background and keeping the nominal model for the other parts, is shown in Table 11.1.

## 894 11.2 Decay-time acceptance

895 To investigate the systematic uncertainty related to the decay-time dependent efficiency,  
 896 we vary our parametrization of the acceptance using cubic splines. This is explicitly  
 897 done by choosing slightly different knot positions, varying the spline coefficients at the  
 898 nominal positions within their statistical uncertainties and adding/subtracting knots  
 899 in the range  $0.4 \text{ ps} < t < 11 \text{ ps}$ . Additionally, an adaptive binning scheme is used to  
 900 determine the knot positions in a way that roughly equal amounts of data is covered  
 901 between two knots. Strictly speaking, the variation of the spline coefficients within their  
 902 uncertainty gives the statistical uncertainty of the decay-time acceptance parametrization.  
 903 For the presented measurement, this is done using the Cholesky decomposition [51] of  
 904 the covariance matrix of coefficients  $c_i$ , generating toy splines with randomized coefficient  
 905 values  $c_{i,toy}$  from this decomposition and refitting using the toy spline. Furthermore, the  
 906 fit to the decay-time distribution of signal  $B_s^0 \rightarrow D_s \pi\pi\pi$  candidates, used to determine  
 907 the spline parametrization, is reiterated with varying fixed/constrained values for  $\Delta\Gamma_s$ .

### 908 11.2.1 Varition of knot positions

909 The nominal knot positions are changed to be:

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

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

### 912 11.2.2 Variation of spline coefficients

913 Due to the sizeable correlation of the spline coefficients  $c_i$  determined in Chapter 6.2, the  
 914 variations of the observables in the amplitude fit when changing one spline coefficient can  
 915 not be added up in quadrature for all coefficients. To simplify the problem, a Cholesky  
 916 decomposition [51] is used to generate a set of uncorrelated vectors from the covariance  
 917 matrix  $A_{cov}$ . It can be shown that every Hermitian positive-definite matrix, such as  $A_{cov}$ ,  
 918 has a unique Cholesky decomposition of the form:

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

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

921 Given the four free spline coefficients which are determined from the fit described in 6,  
 922  $A_{cov}$  is a  $4 \times 4$  matrix. Therefore, the lower triangular matrix  $L$  is of the form:

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

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

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

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

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

### 11.3 Decay-time resolution

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

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

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

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

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

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

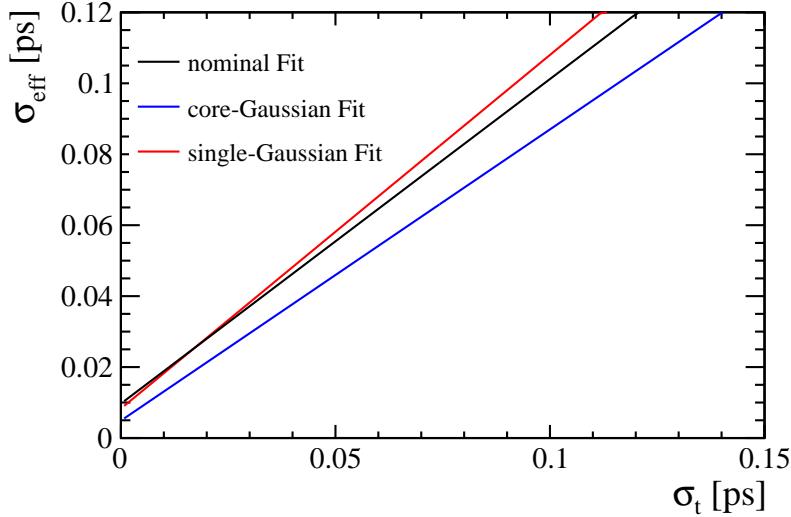


Figure 11.1: The measured resolution scaling function of the per-event decay time error estimate  $\sigma_t$  for fake  $B_s$  candidates (Run-II data) for (black line) the nominal scaling, (blue line) only using the narrow gaussian width of the double gaussian fit model or (red line) when determining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

## 960 11.4 Tagging calibration

961 Systematic uncertainties arise from the statistical precision of the tagging parameters  
 962 determined from the calibration, discussed in Sec. 7. These uncertainties are accounted  
 963 for by the inclusion of Gaussian constrains in the nominal fit. The width of the respective  
 964 constrain for the tagging parameter  $p_i$  is chosen to be  $\Delta p_i$ . In this way, the systematic  
 965 uncertainty due to the tagging calibration is included in the statistical uncertainty of the  
 966 time dependent fit.

## 967 11.5 Summary of systematic uncertainties

968 All contributing systematic uncertainties are summarized in Table 11.1. The individual  
 969 uncertainties are summed in quadrature to arrive at the total systematic uncertainty for the  
 970 respective CP observable. Their total magnitude ranges from (30-40)% of the statistical  
 971 uncertainty of the fit.

Table 11.1: Summary of systematic uncertainties for the time dependent fit.

	Fit bias	Acceptance	Resolution	$\Delta m_s$	Asymmetries	Background	Total	$\sigma_{syst}/\sigma_{stat}$
$C$	0.00	0.01	0.03	0.01	0.00	0.01	0.04	0.38
$D$	0.01	0.07	0.00	0.00	0.01	0.05	0.09	0.34
$\bar{D}$	0.01	0.07	0.01	0.00	0.01	0.04	0.08	0.32
$S$	0.00	0.00	0.02	0.04	0.01	0.03	0.05	0.40
$\bar{S}$	0.01	0.00	0.03	0.04	0.00	0.00	0.05	0.33

## 972 A Stripping and Trigger cuts

973 The following text describes variables which are used in Table 1.1 and might be ambiguous,  
 974 or which benefits are not straight forward. Where noted, different cut values are applied  
 975 for Run-I and Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest  
 976 approach. This variable is used to ensure that all  $D_s$  and  $X_{s,d}$  daughters originate from  
 977 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	

978

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

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

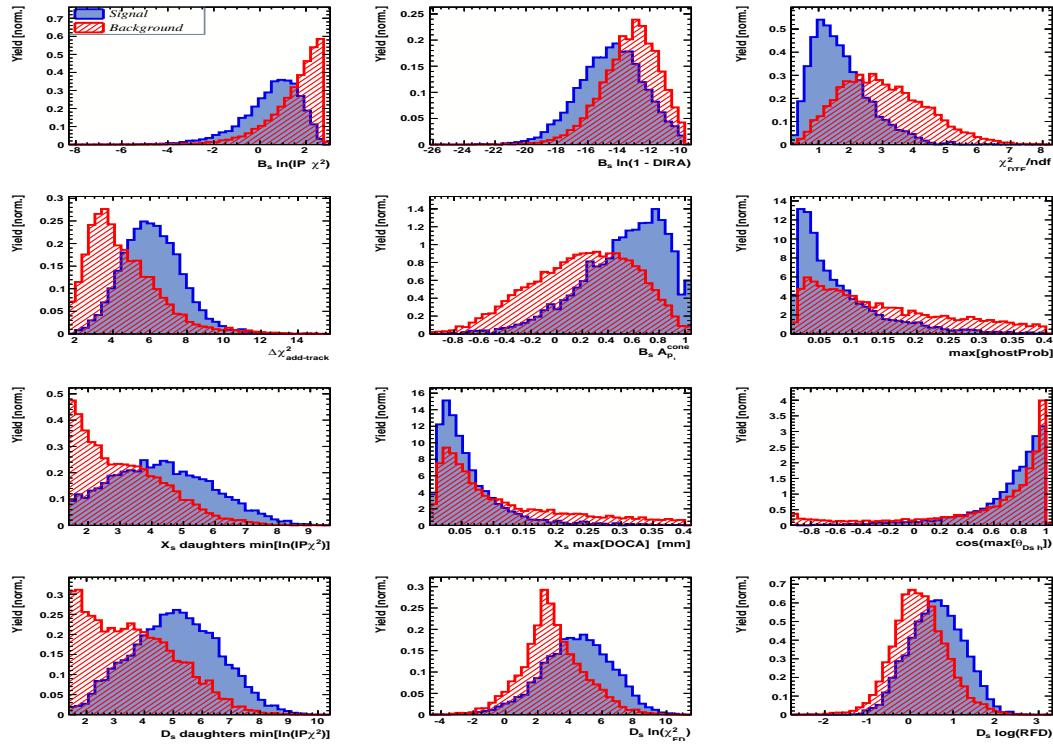


Figure A.1: Variables used to train the BDTG for category [Run-I,L0-TOS].

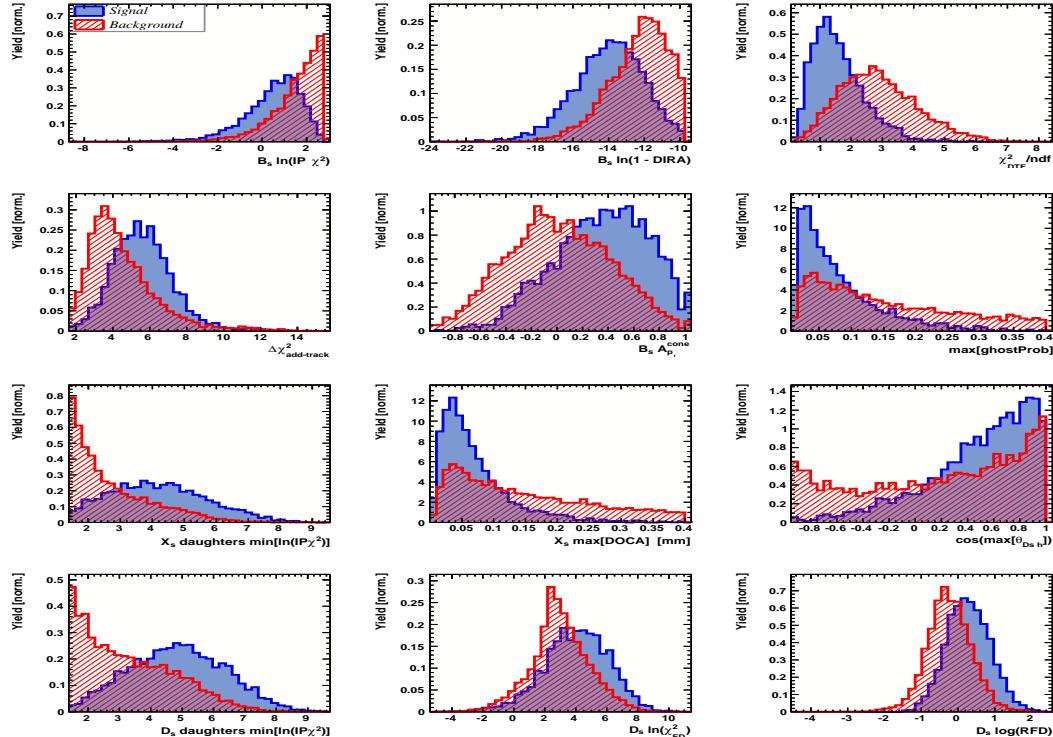


Figure A.2: Variables used to train the BDTG for category [Run-I,L0-TIS].

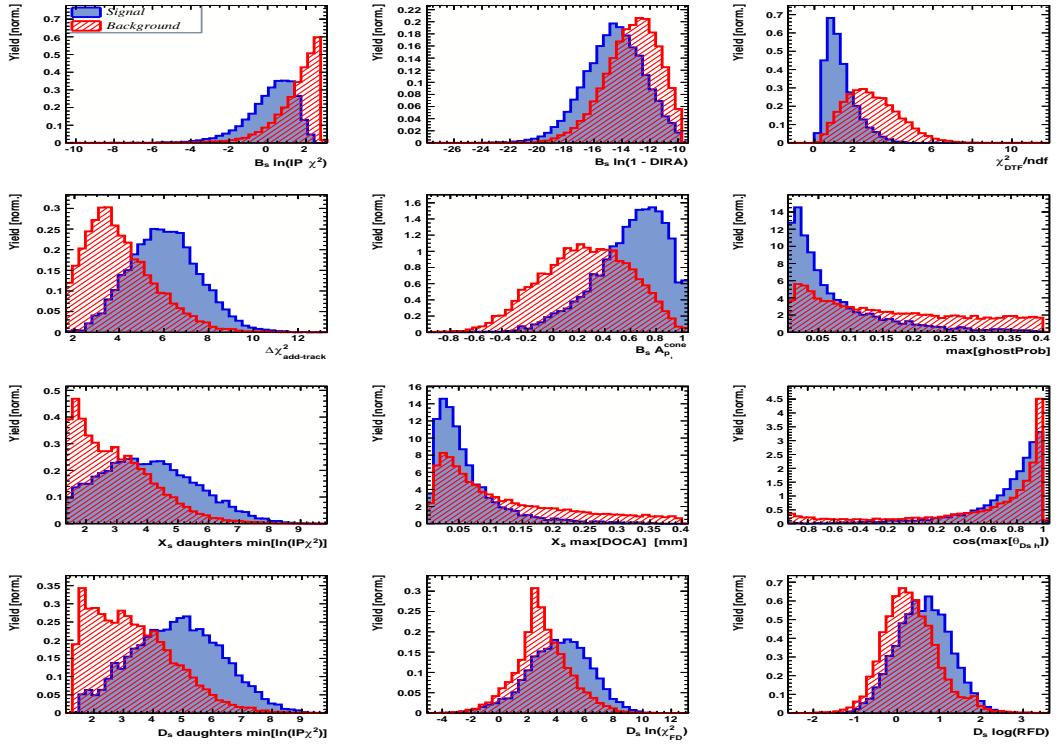


Figure A.3: Variables used to train the BDTG for category [Run-II,L0-TOS].

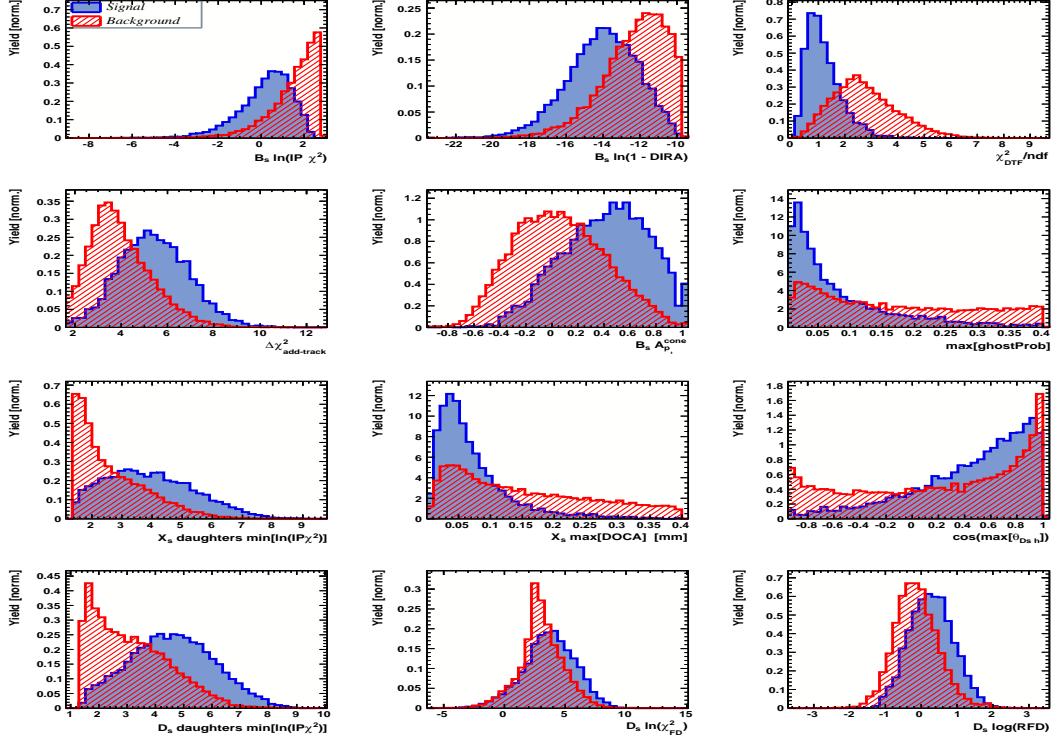


Figure A.4: Variables used to train the BDTG for category [Run-II,L0-TIS].

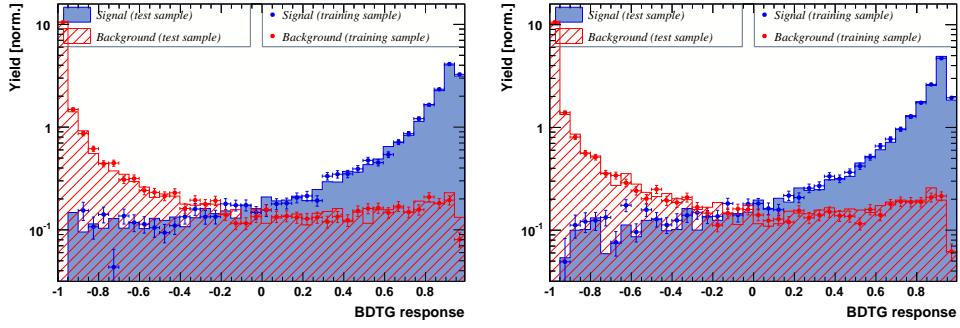


Figure A.5: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].

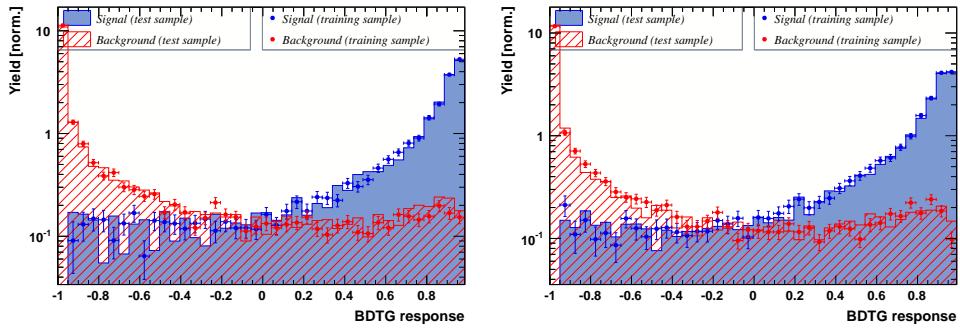


Figure A.6: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].

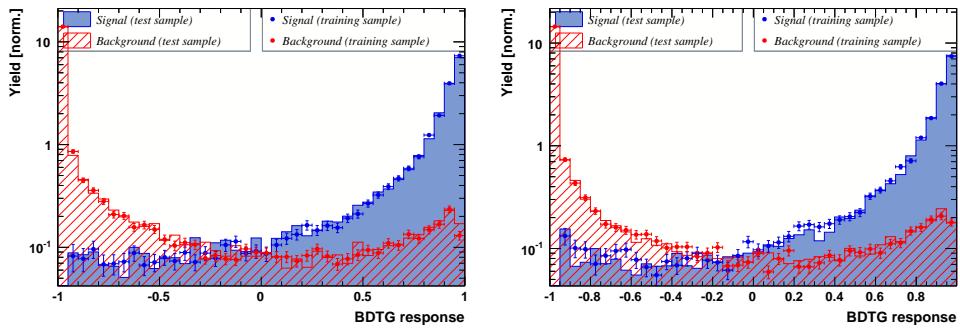


Figure A.7: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].

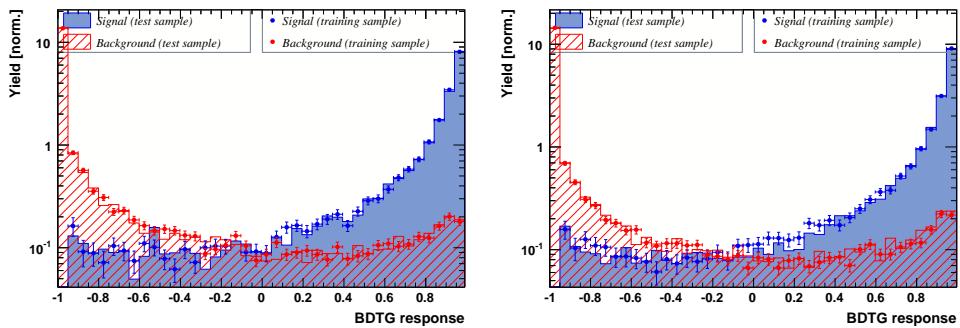


Figure A.8: Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].

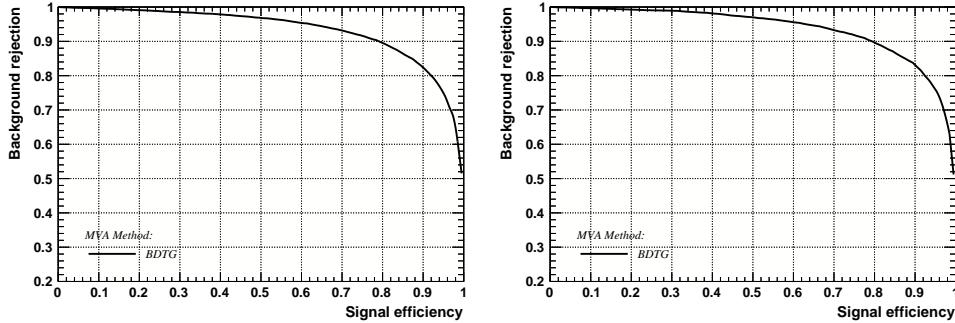


Figure A.9: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].

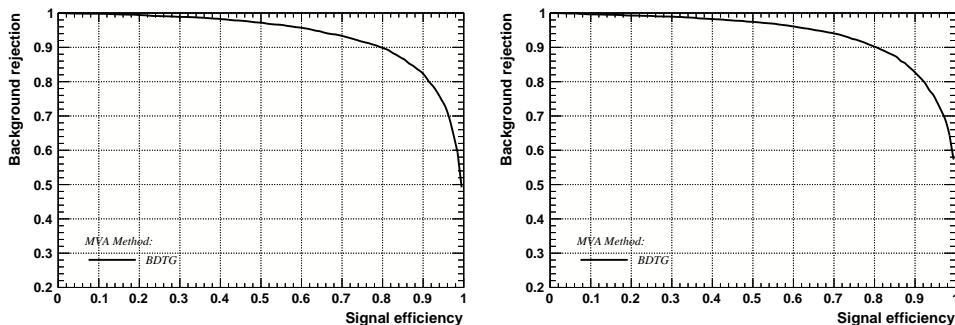


Figure A.10: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TIS].

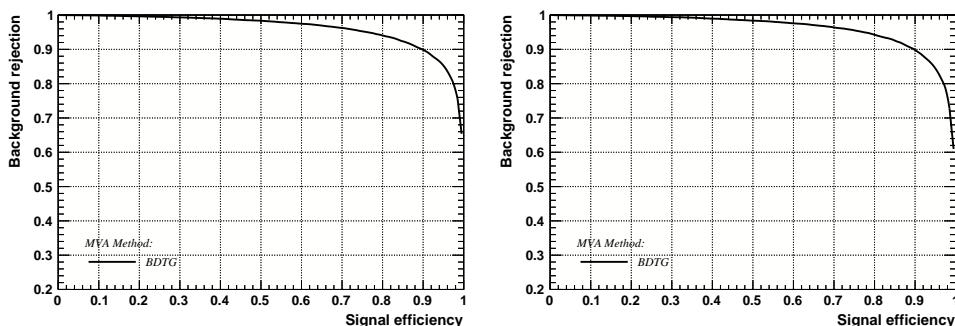


Figure A.11: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TOS].

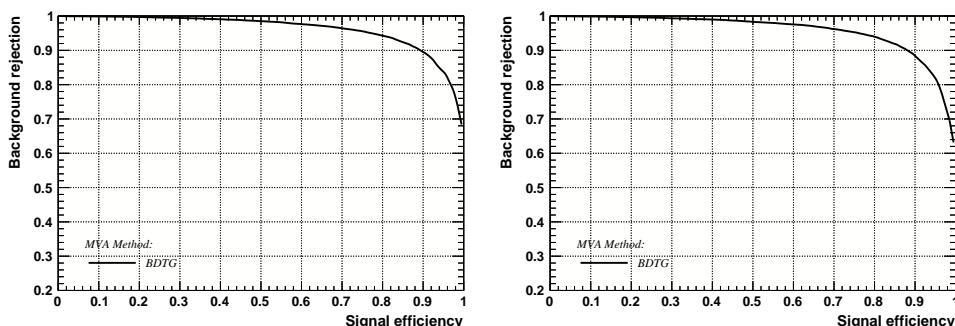


Figure A.12: Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TIS].

## 990 C Detailed mass fits

991 In this section, all fits to the mass distribution of  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$   
 992 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-  
 993 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  
 994 L0 trigger category.

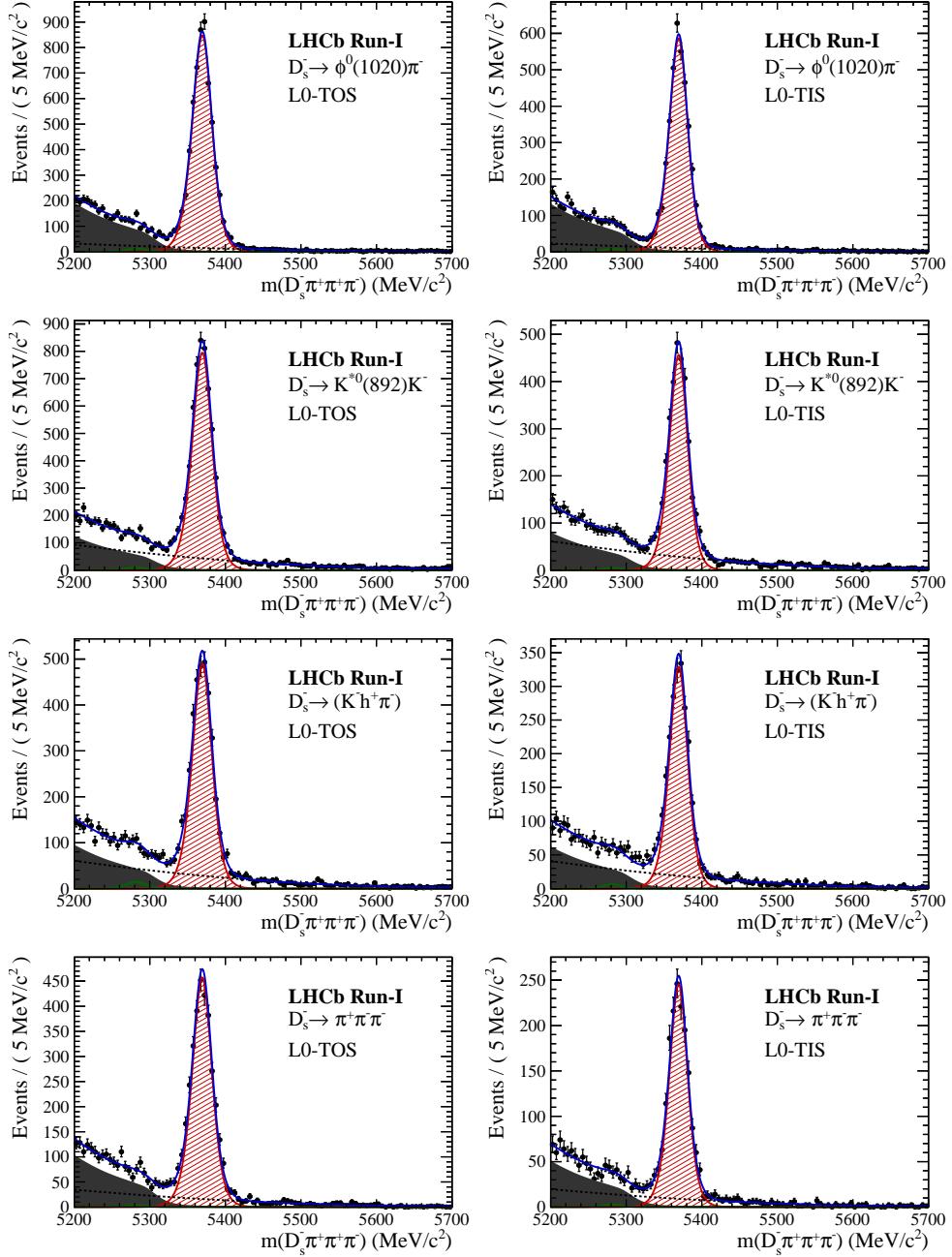


Figure B.1: Invariant mass distributions of  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates for Run-I data.

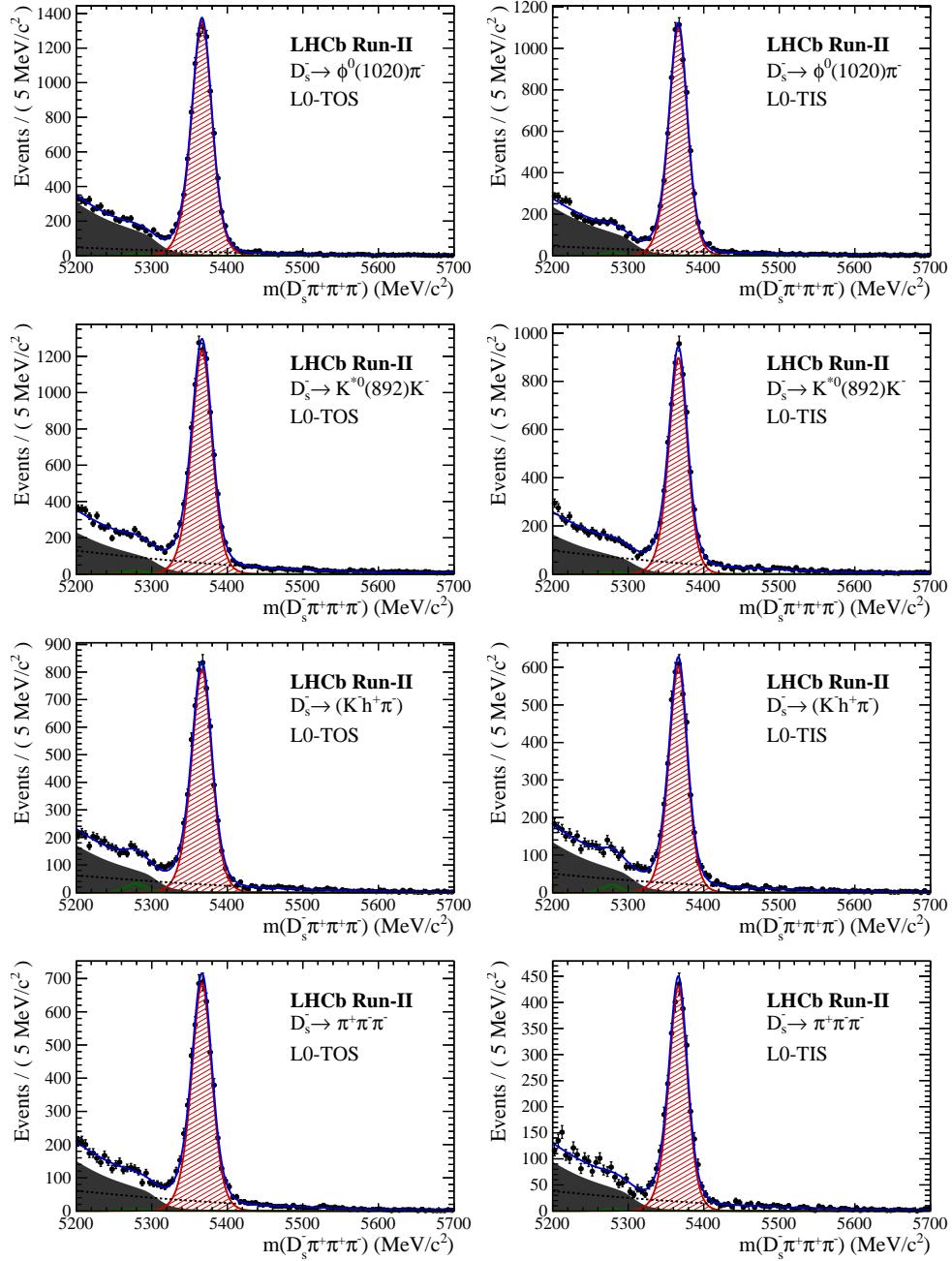


Figure B.2: Invariant mass distributions of  $B_s^0 \rightarrow D_s \pi^+ \pi^- \pi^+$  candidates for Run-II data.

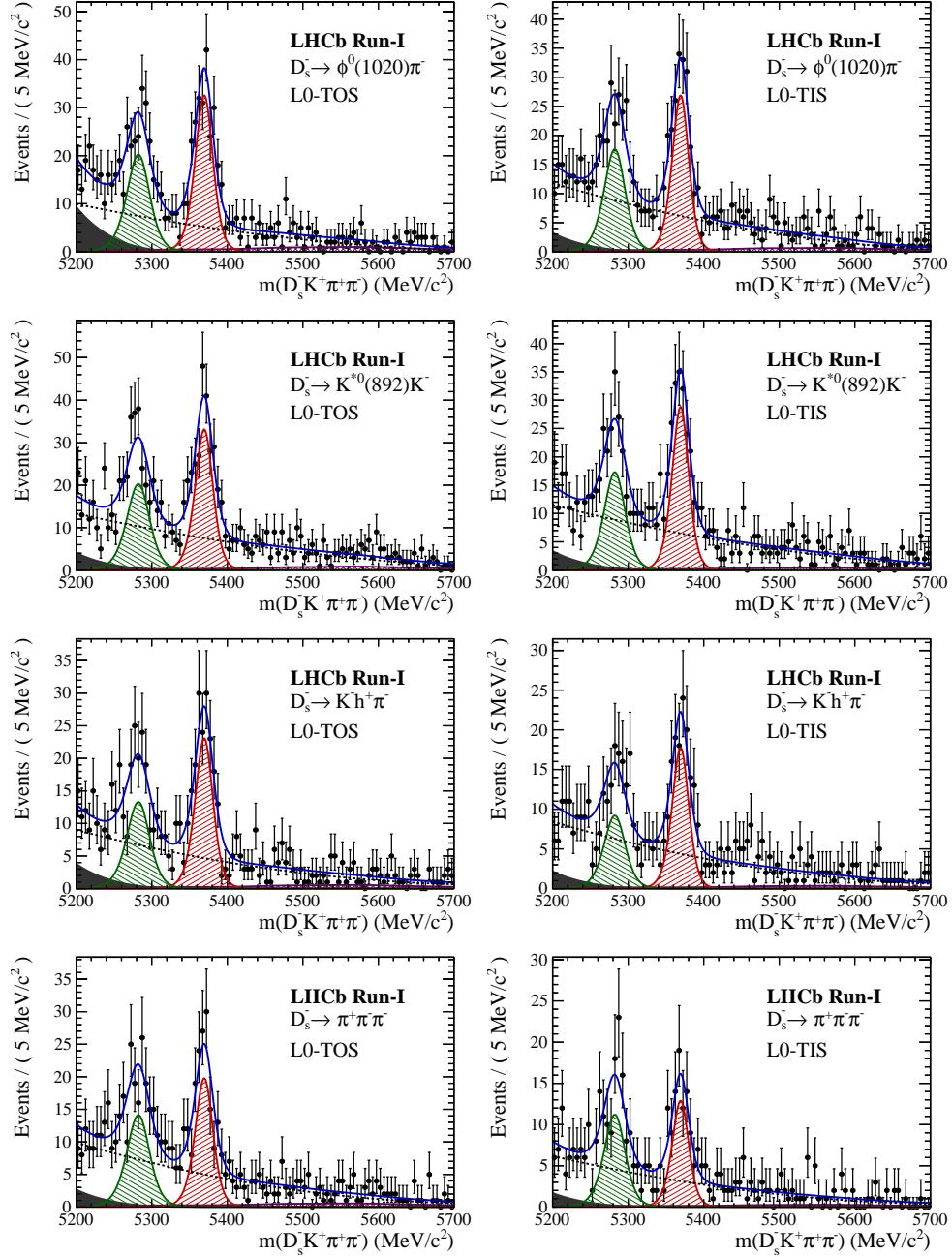


Figure B.3: Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for Run-I data.

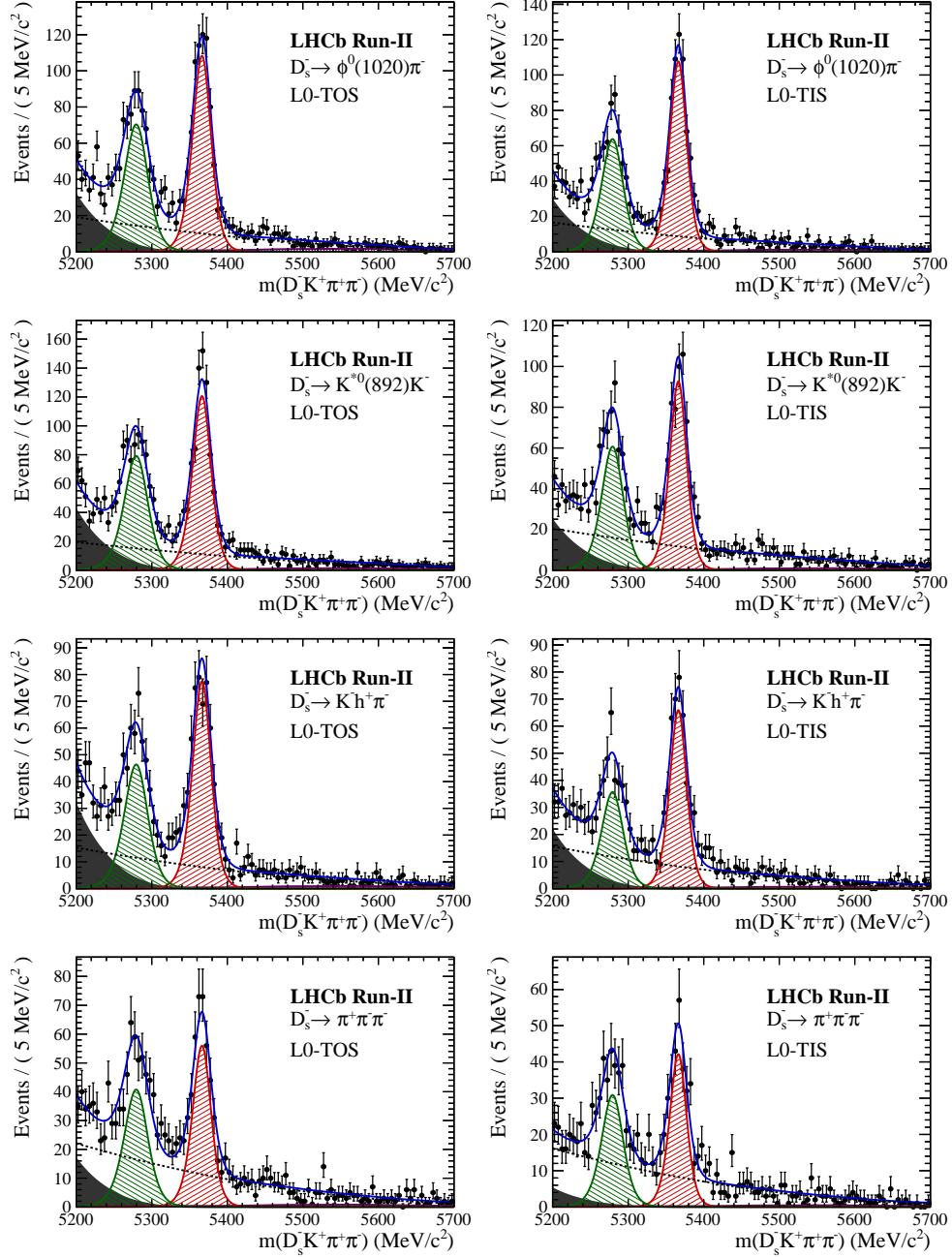


Figure B.4: Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi\pi$  candidates for Run-II data.

## 995 D Decay-time Resolution fits

996 This section contains all fits to the distributions of the decay time difference  $\Delta t$  between  
 997 the true and the reconstructed decay time of the truth-matched  $B_s^0$  candidates on MC.  
 998 The fits are performed in bins of the decay time error  $\sigma_t$ , where an adaptive binning  
 999 scheme is used to ensure that approximately the same number of events are found in each  
 1000 bin.

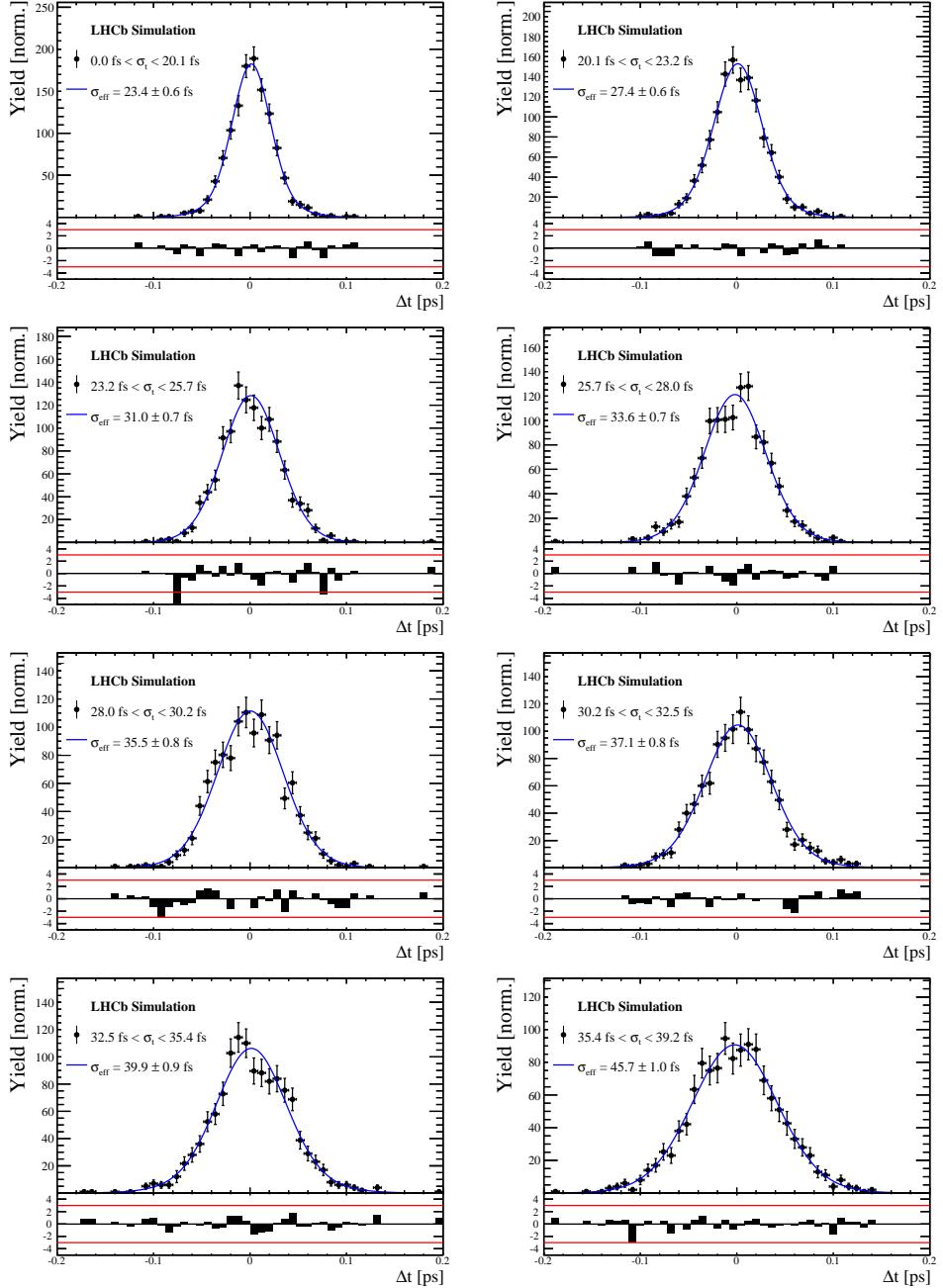


Figure C.1: Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K \pi\pi$  MC candidates in bins of the per-event decay time error estimate..

$\sigma_t$ 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.

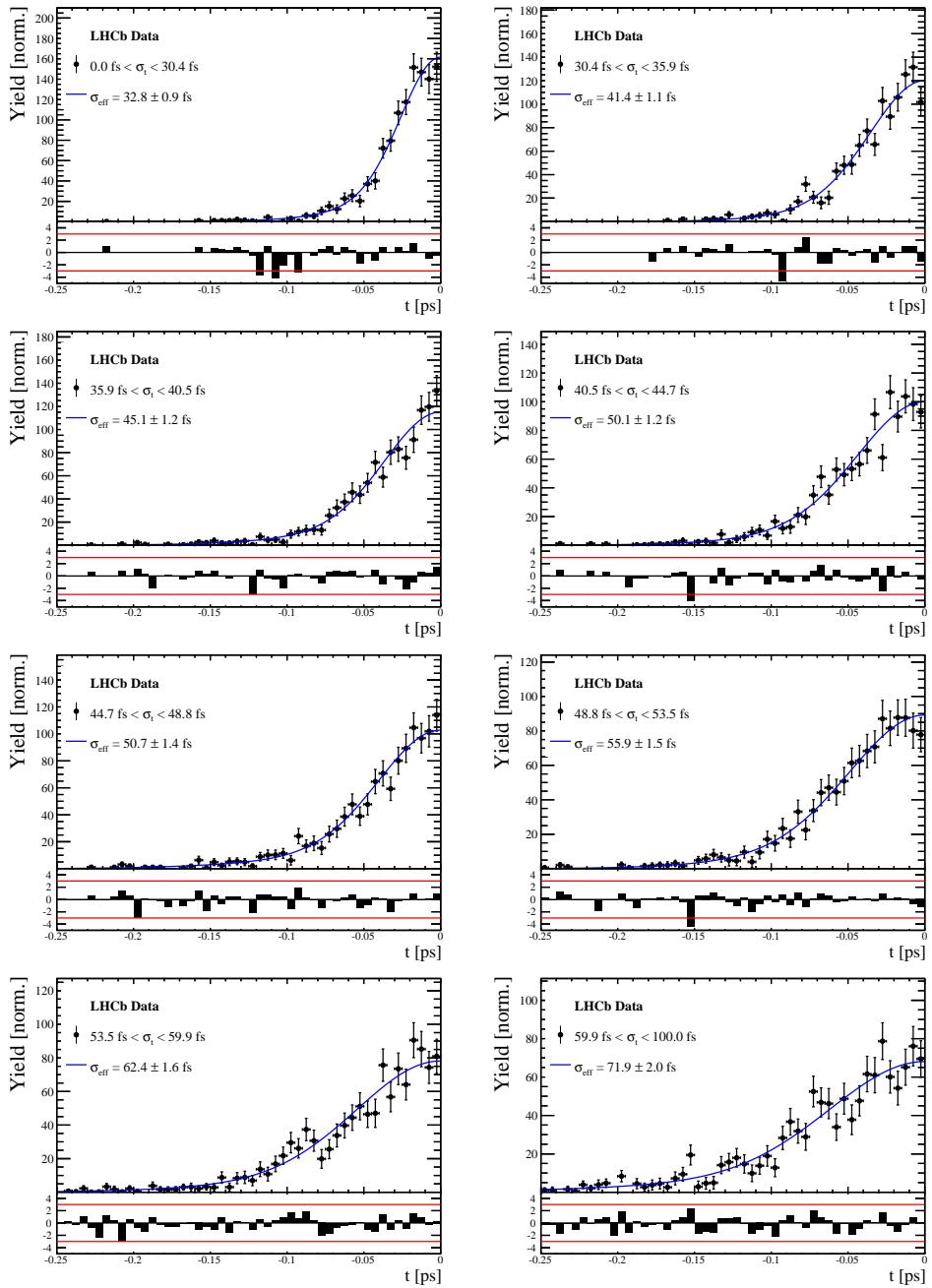


Figure C.2: Decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks, for bins in the per-event decay time error estimate.

$\sigma_t$ 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.

## 1001 E Spin Amplitudes

1002 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	$B[D] \rightarrow (T_1 T_2)$ , $T_1 \rightarrow (P_1 P_2)$ , $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\gamma}(T_1) L_{(2)\gamma}^\beta(T_2)$

## 1003 F Considered Decay Chains

1004 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
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ \rightarrow \pi^+ K^*(1430)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_1(1270)^+ [S, D] \rightarrow K^+ \omega(782)]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_1(1400)^+ [S, D] \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ \kappa]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \sigma]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K(1460)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1410)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2^*(1430)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K^*(1680)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow \pi^+ K^*(892)^0]$
$B_s \rightarrow D_s^- [K_2(1770)^+ \rightarrow K^+ \rho(770)^0]$
$B_s \rightarrow \sigma^0(D_s^- K^+_S)$
$B_s [S, P, D] \rightarrow \sigma^0(D_s^- K^+_V)$
$B_s \rightarrow \rho(770)^0 (D_s^- K^+_S)$
$B_s [S, P, D] \rightarrow \rho(770)^0 (D_s^- K^+_V)$
$B_s \rightarrow K^*(892)^0 (D_s^- \pi^+_S)$
$B_s [S, P, D] \rightarrow K^*(892)^0 (D_s^- \pi^+_V)$
$B_s \rightarrow (D_s^- K^+_S) (\pi^+ \pi^-)_S$

1005 G MC corrections

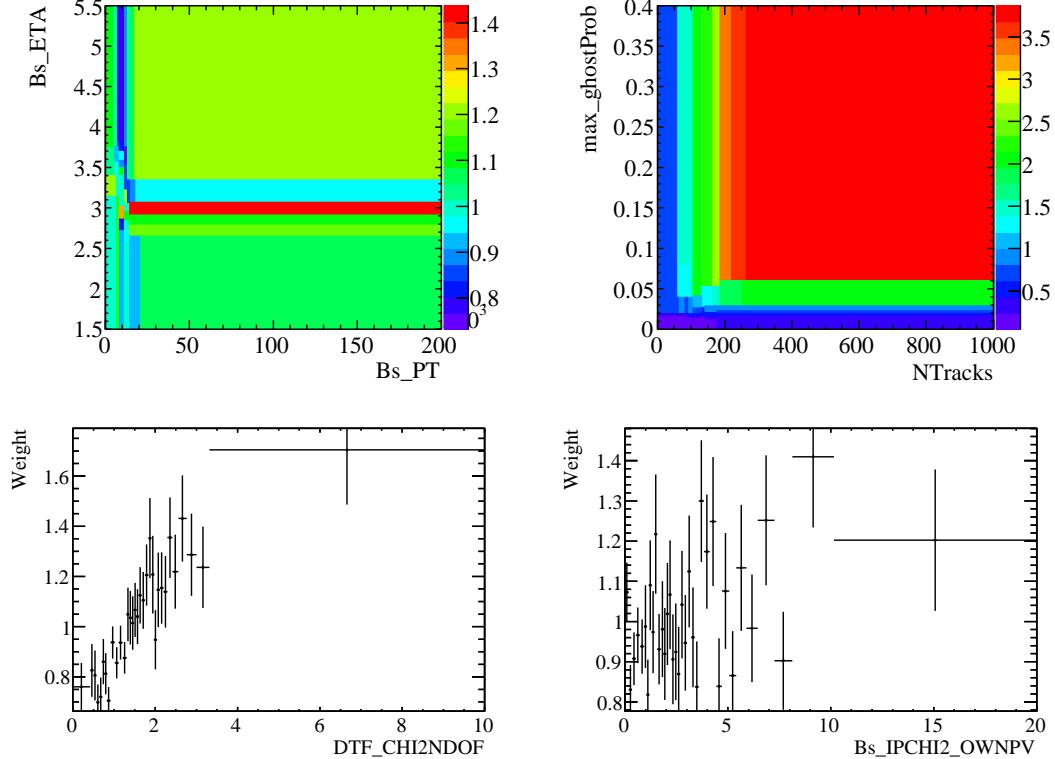


Figure C.1: Weights applied to correct for Data/MC differences.

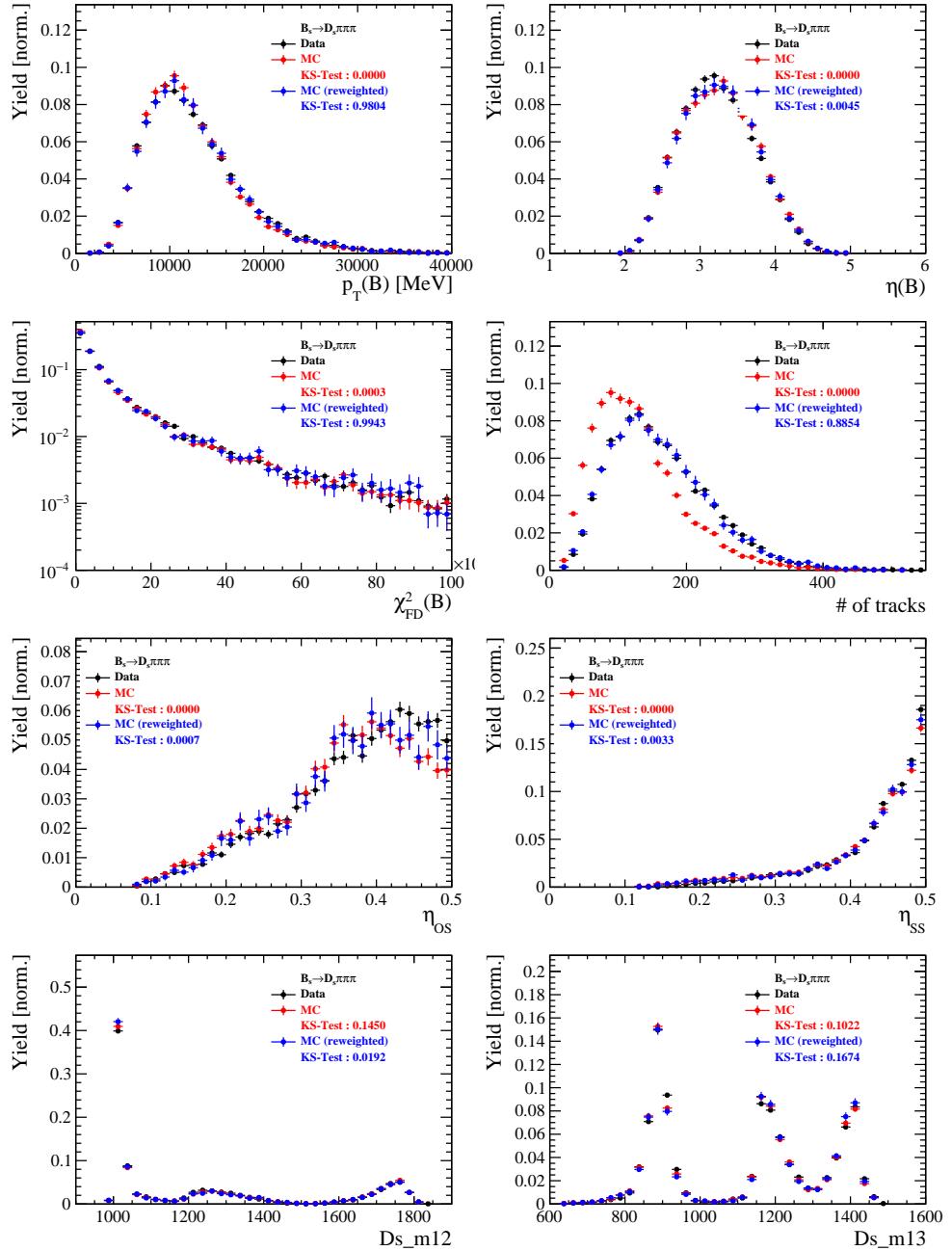


Figure C.2: Comparison of selected variables for  $B_s \rightarrow D_s\pi\pi\pi$  decays.

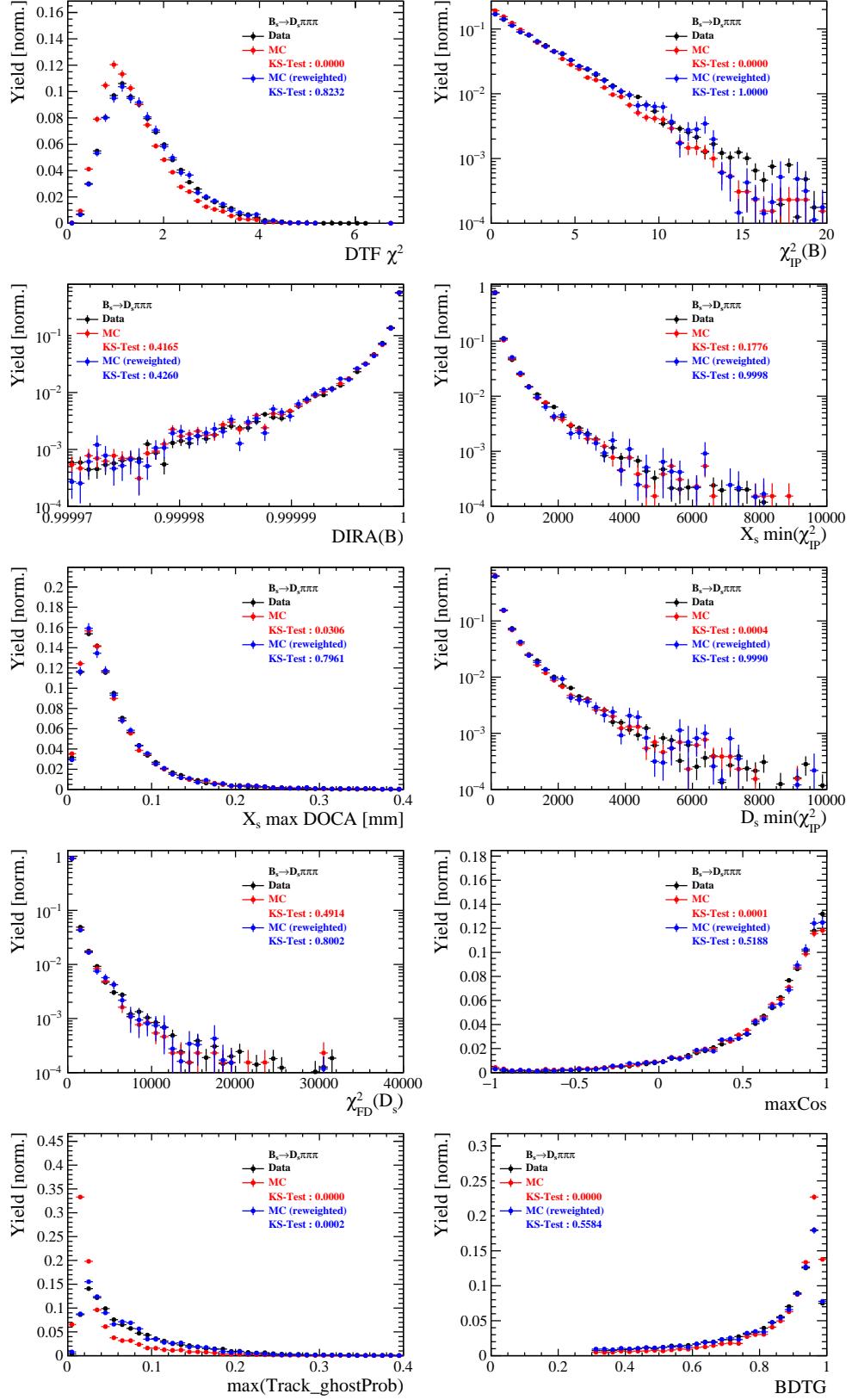


Figure C.3: Comparison of BDTG input variables and classifier response for  $B_s \rightarrow D_s\pi\pi\pi$  decays.

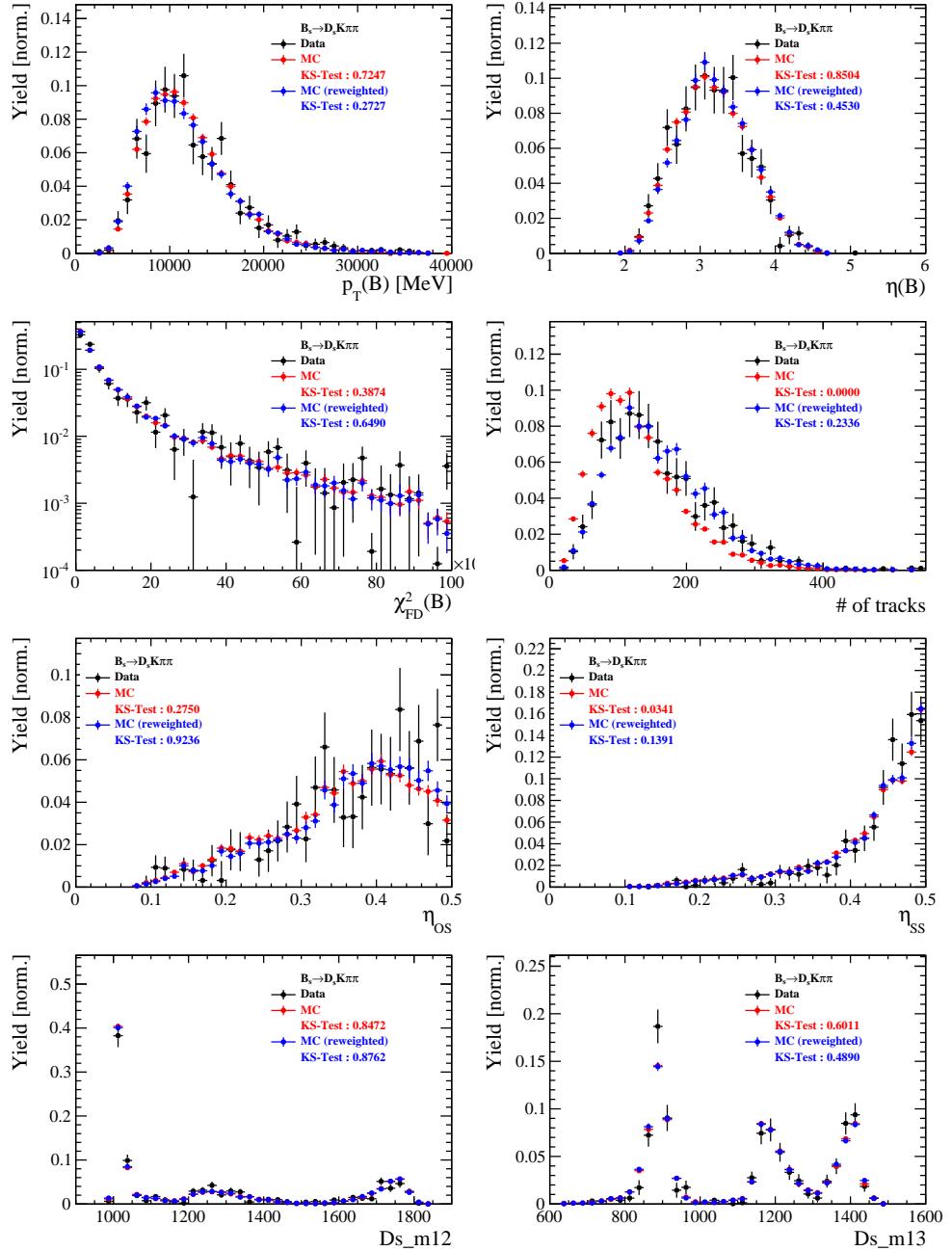


Figure C.4: Comparison of selected variables for  $B_s \rightarrow D_s K\pi\pi$  decays.

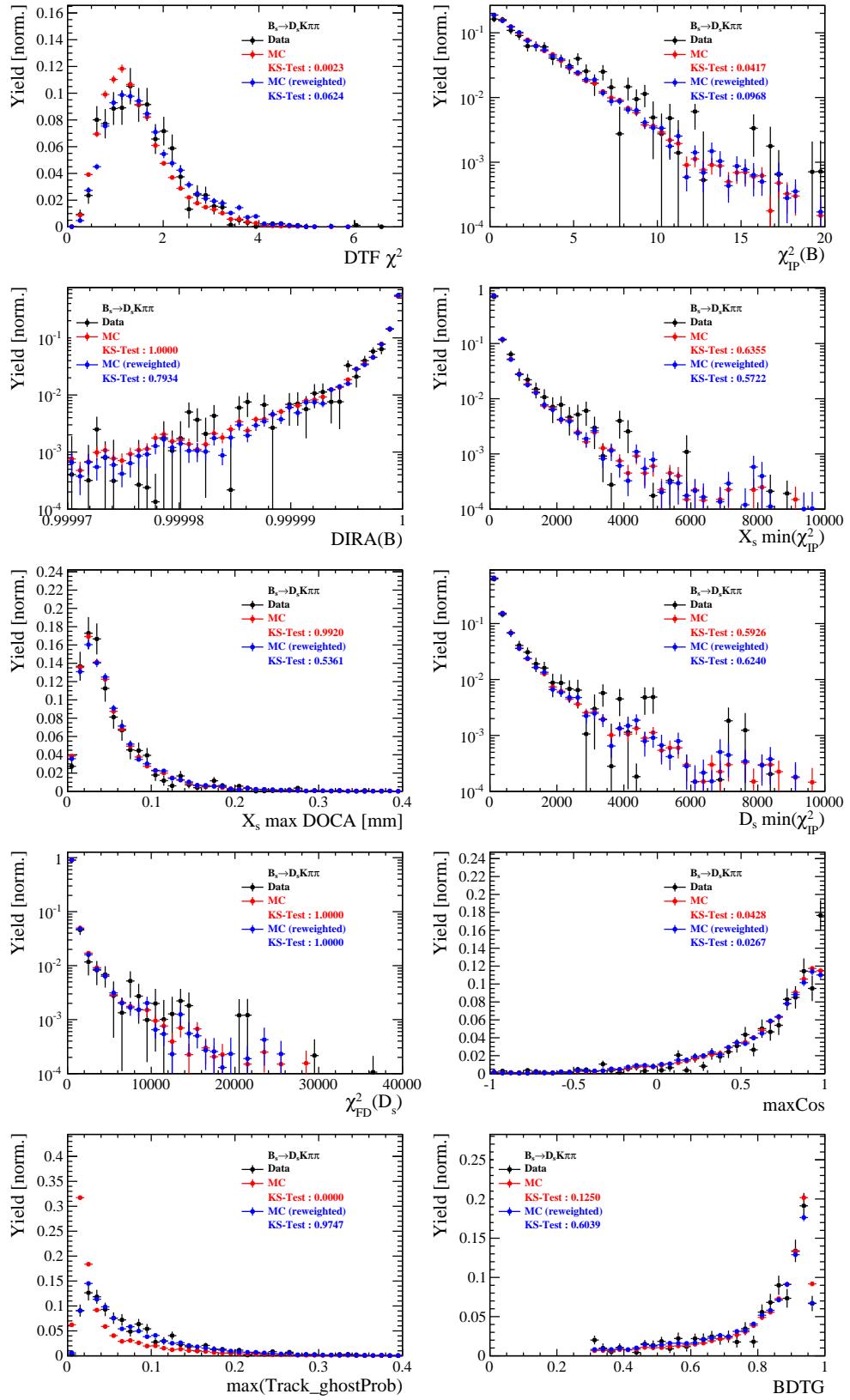


Figure C.5: Comparison of BDTG input variables and classifier response for  $B_s \rightarrow D_s K\pi\pi$  decays.

# 1006 H Data distributions

## 1007 H.1 Comparison of signal and calibration channel

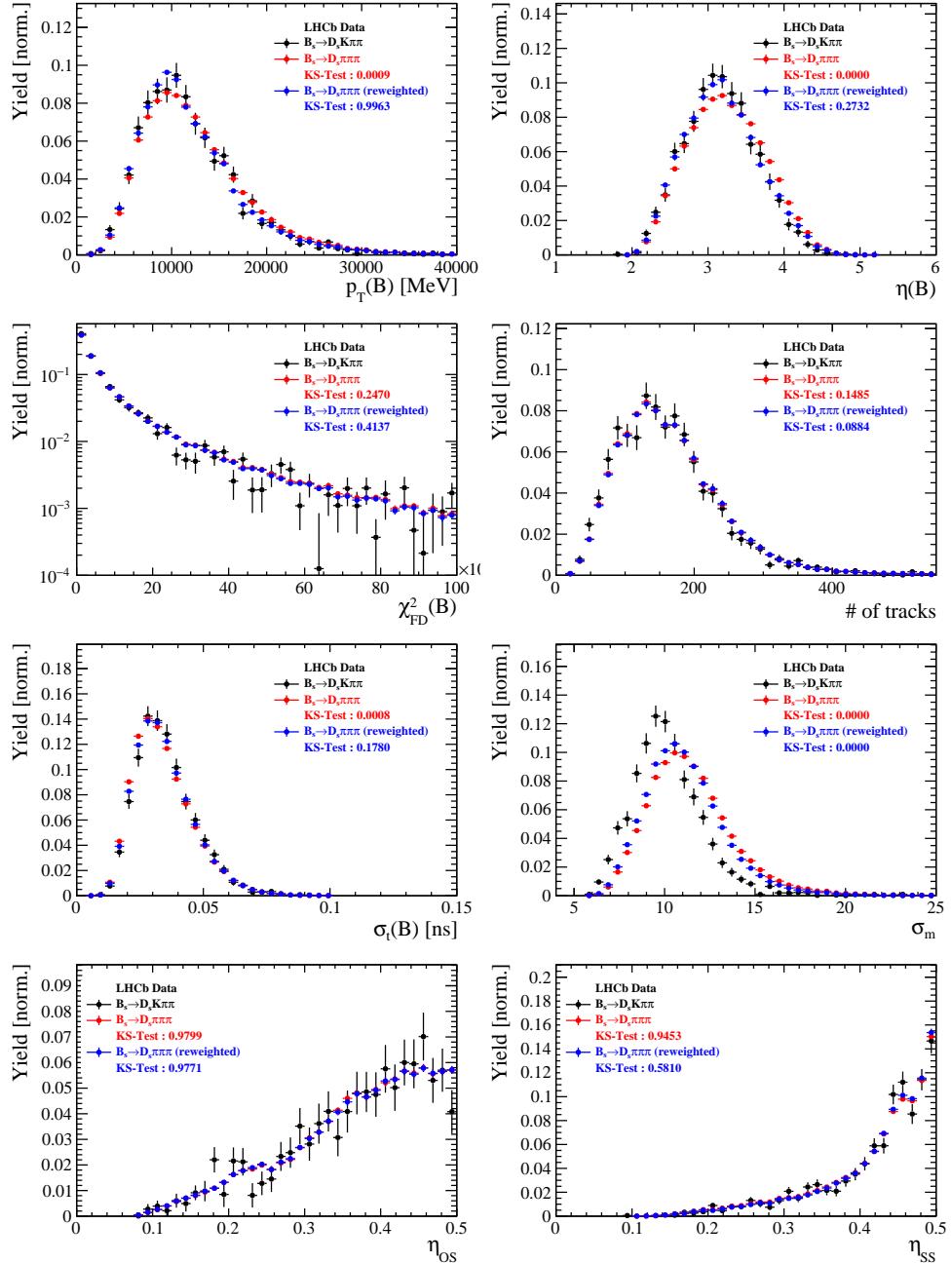


Figure C.1: Comparison of selected variables.

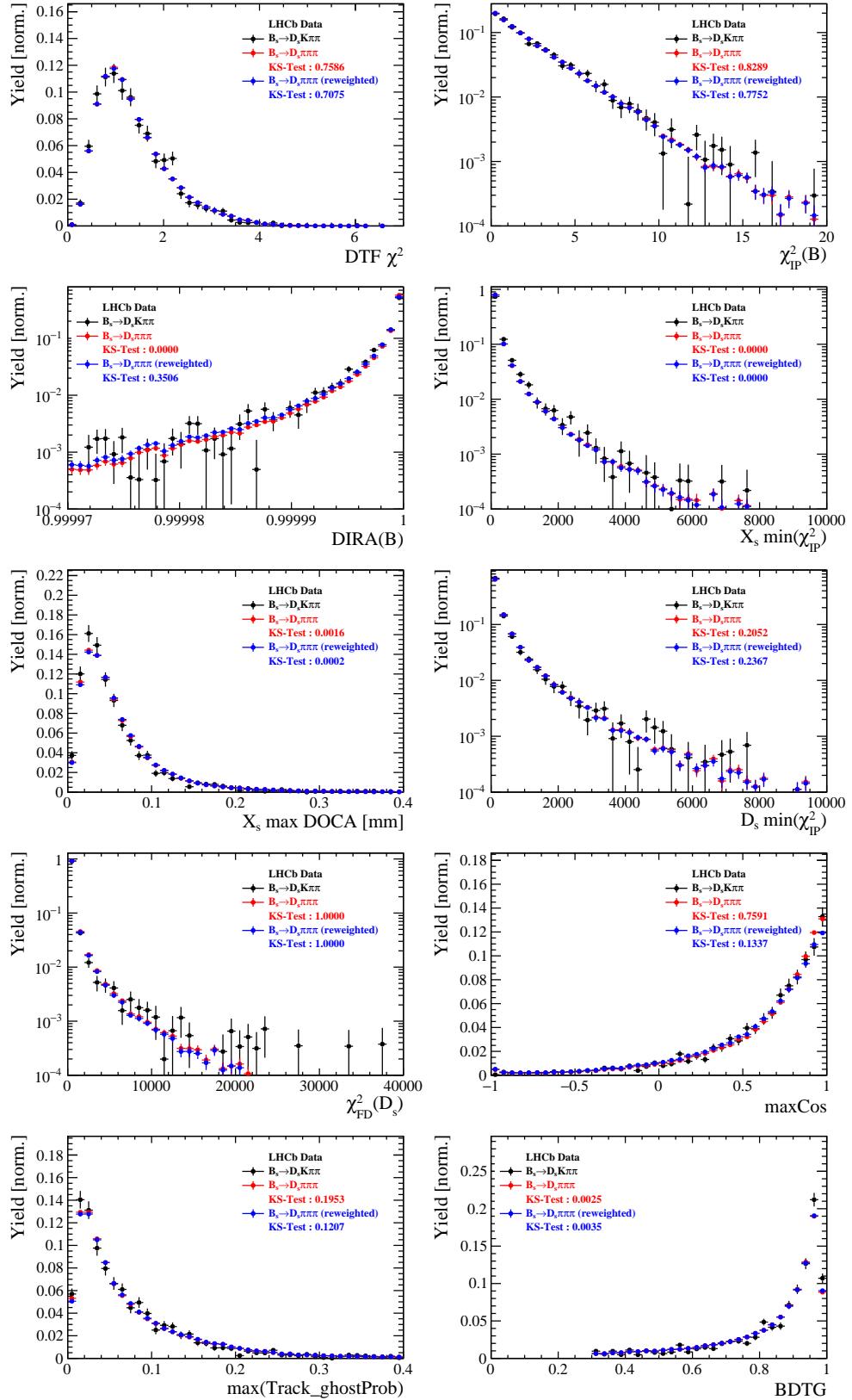


Figure C.2: Comparison of BDTG input variables and classifier response.

1008 H.2 Comparison of Run-I and Run-II data

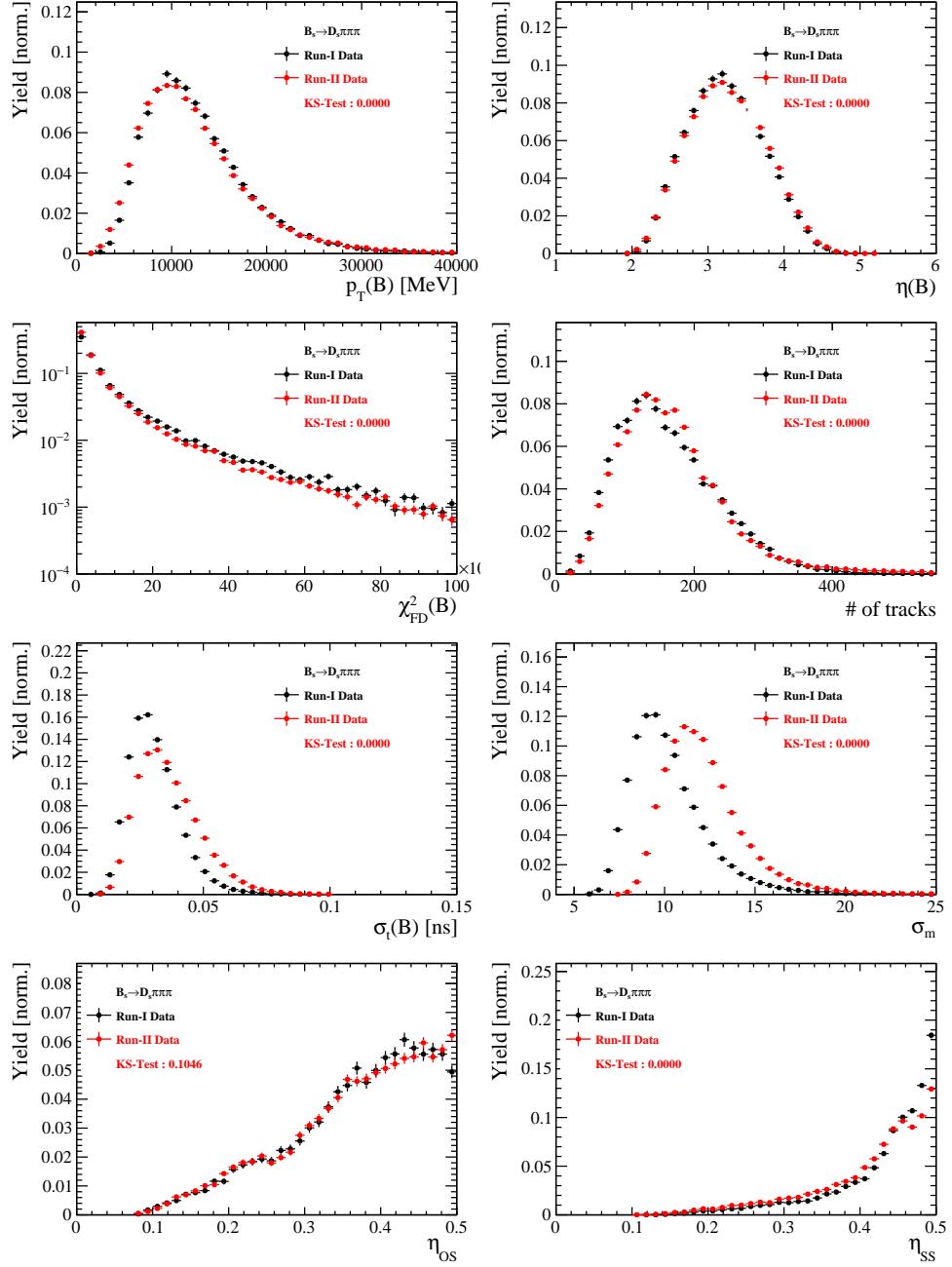


Figure C.3: Comparison of selected variables.

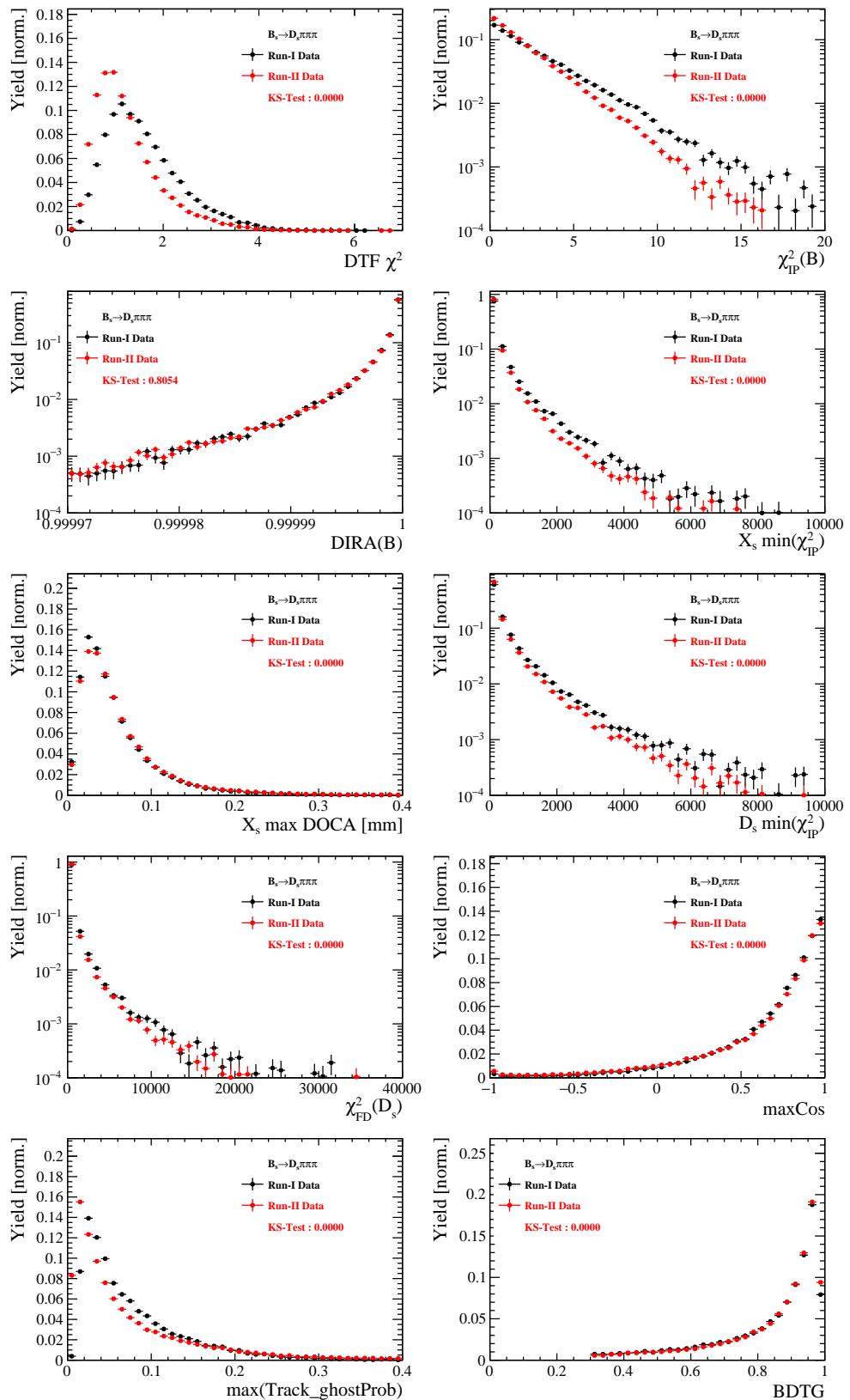


Figure C.4: Comparison of BDTG input variables and classifier response.

1009 H.3 Comparison of  $D_s$  final states

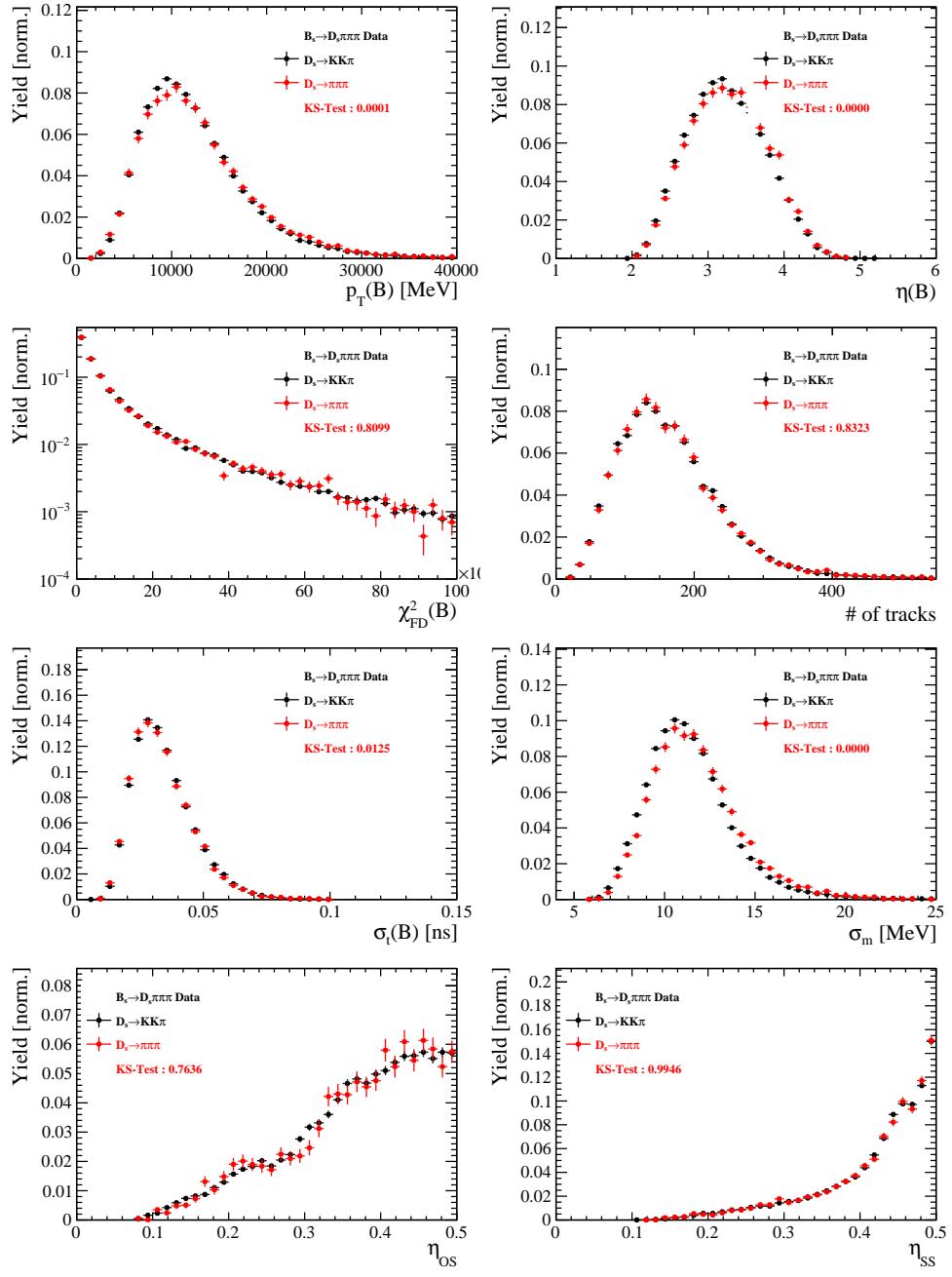


Figure C.5: Comparison of selected variables.

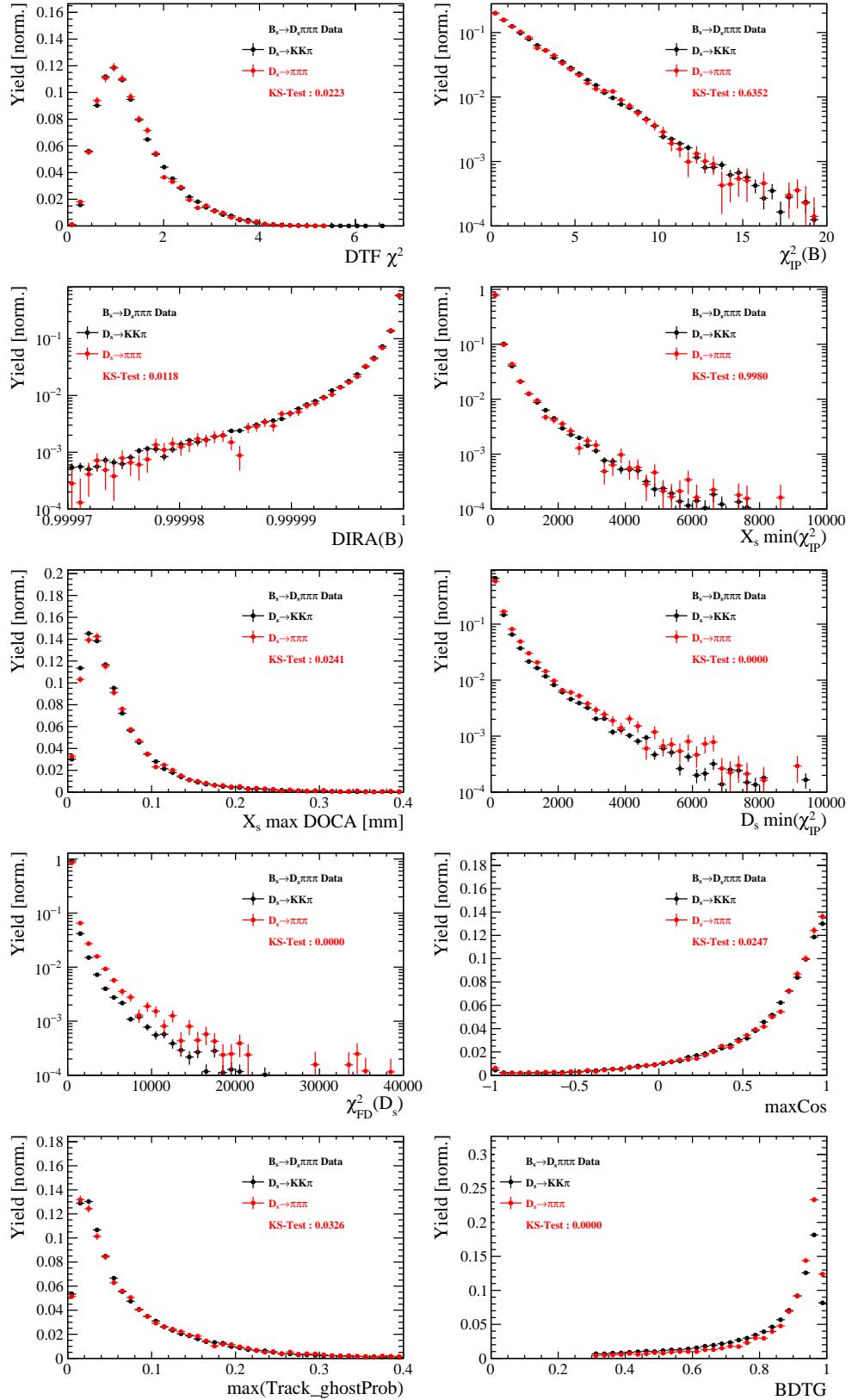


Figure C.6: Comparison of BDTG input variables and classifier response.

1010 H.4 Comparison of trigger categories

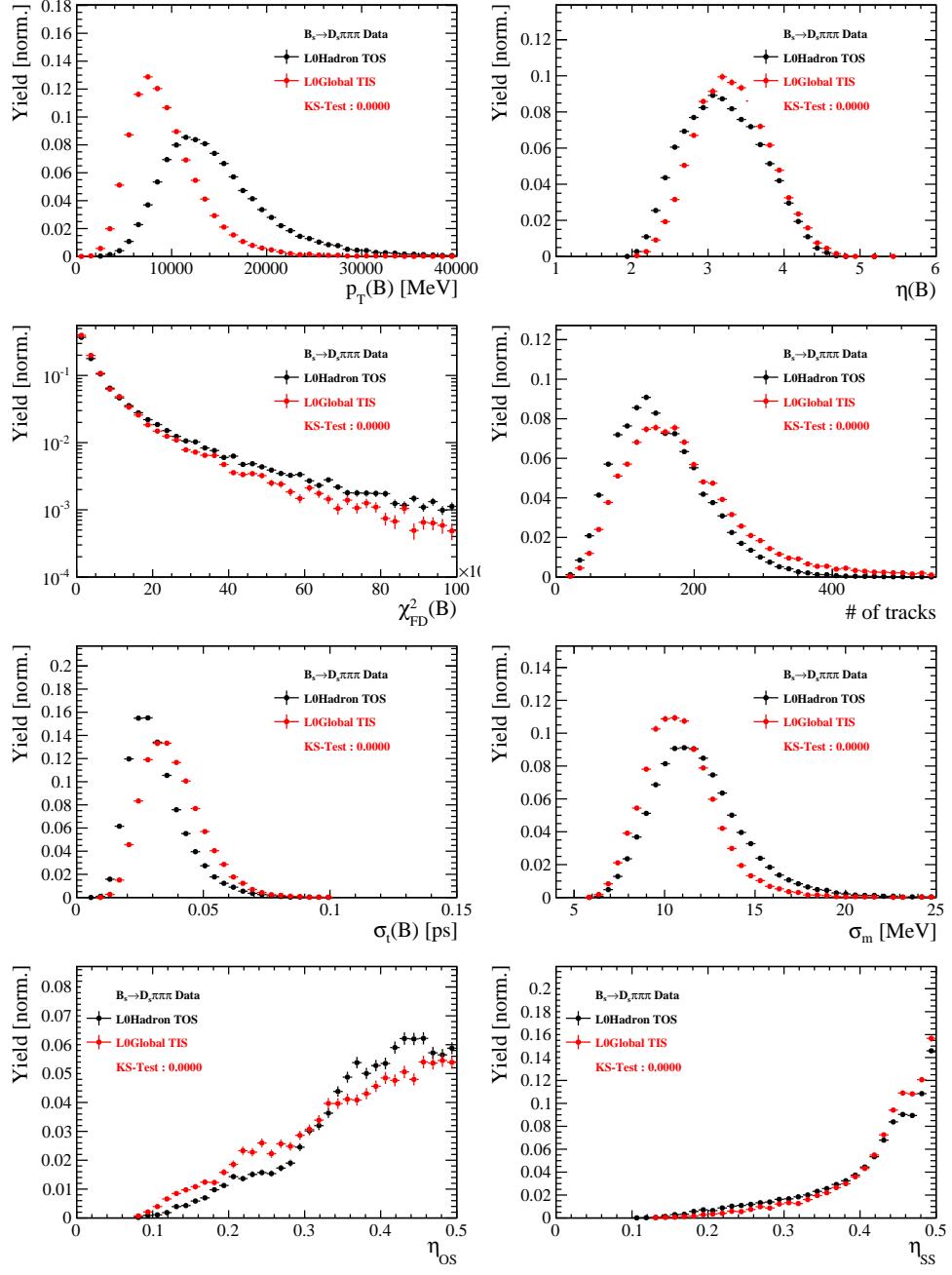


Figure C.7: Comparison of selected variables.

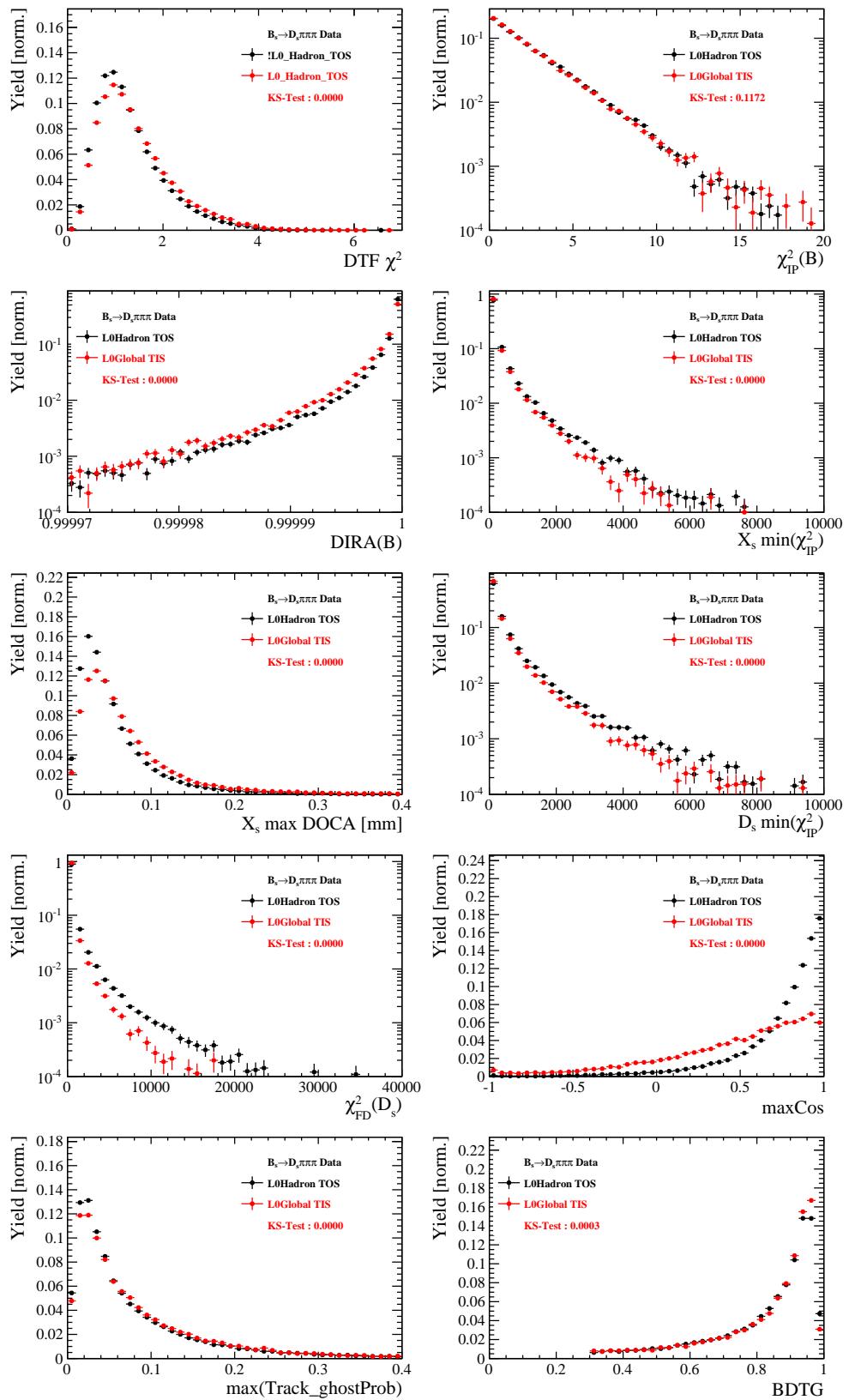


Figure C.8: Comparison of BDTG input variables and classifier response.

1011 H.5 Comparison of  $B_s$  and  $B_d$  decays

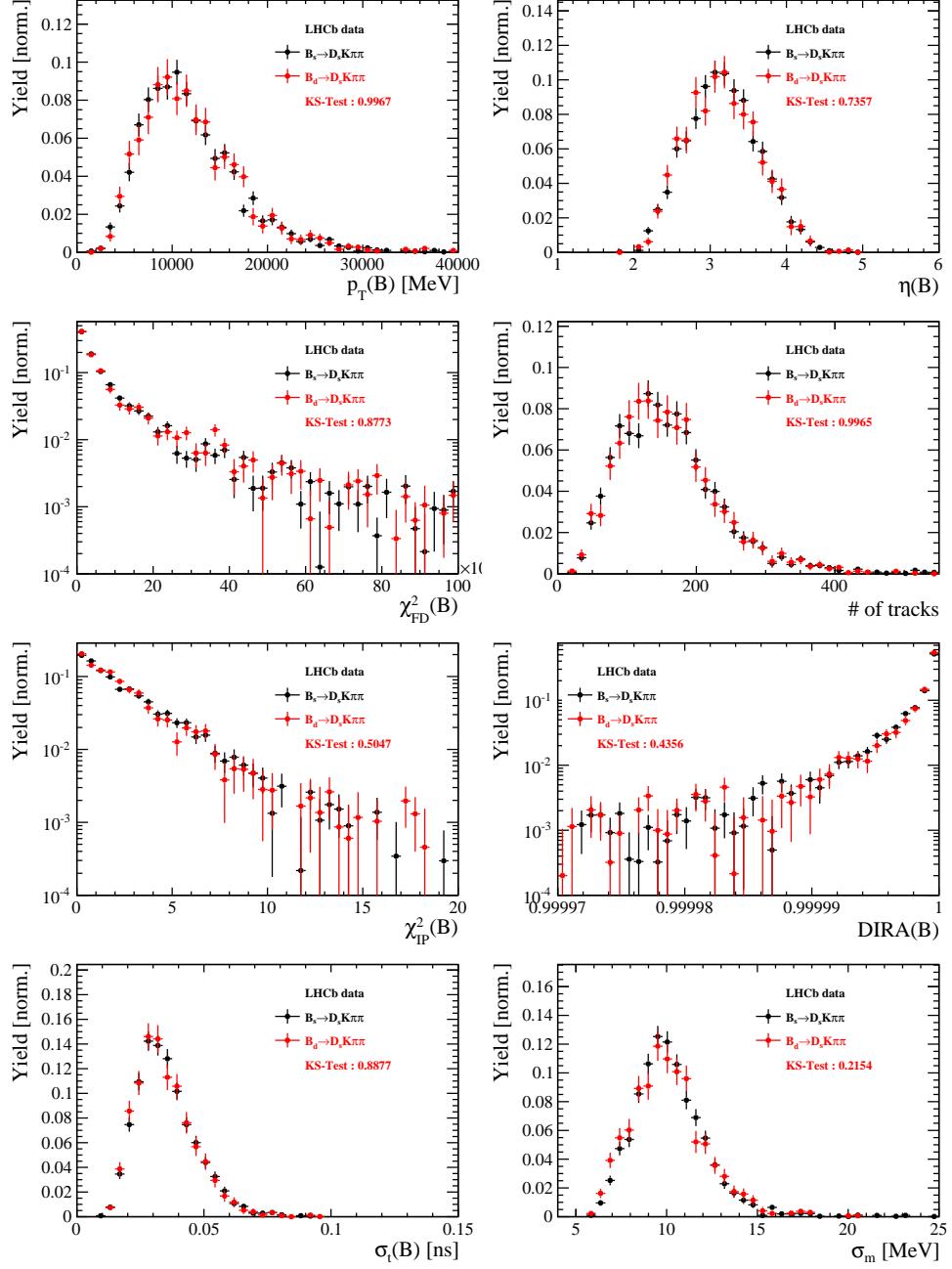


Figure C.9: Comparison of selected variables.

## 1012 References

- 1013 [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  $B(d) \rightarrow D \pi$ ,  $D^* \pi$ , ... decays*, Nucl.  
1014 Phys. **B671** (2003) 459, [arXiv:hep-ph/0304027](#).
- 1015
- 1016 [2] K. De Bruyn *et al.*, *Exploring  $B_s \rightarrow D_s^{(*)\pm} K^\mp$  Decays in the Presence of a Sizable*  
1017 *Width Difference  $\Delta\Gamma_s$* , Nucl. Phys. **B868** (2013) 351, [arXiv:1208.6463](#).
- 1018 [3] S. Blusk, *First observations and measurements of the branching fractions for the*  
1019 *decays  $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$  and  $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* .
- 1020 [4] LHCb, S. Blusk, *Measurement of the CP observables in  $\bar{B}_s^0 \rightarrow D_s^+ K^-$  and first obser-*  
1021 *vation of  $\bar{B}_{(s)}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$  and  $\bar{B}_s^0 \rightarrow D_{s1}(2536)^+ \pi^-$* , 2012. [arXiv:1212.4180](#).
- 1022 [5] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory* (*Frontiers in Physics*), Westview Press, 1995.
- 1023
- 1024 [6] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 1025 [7] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation of*  
1026 *production processes*, Annals of Physics **18** (1962), no. 2 198 .
- 1027 [8] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*,  
1028 Phys. Rev. D **11** (1975) 3165.
- 1029 [9] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of  
1030 Physics **108** (1977), no. 2 454 .
- 1031 [10] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay*  
1032 *widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.
- 1033 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo  
1034 Cimento Series 10 **34** (1964), no. 6 1644.
- 1035 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys.  
1036 **C40** (2016), no. 10 100001.
- 1037 [13] D. V. Bugg, *The mass of the  $\sigma$  pole*, Journal of Physics G Nuclear Physics **34** (2007)  
1038 151, [arXiv:hep-ph/0608081](#).
- 1039 [14] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-*  
1040 *dominance prediction for  $\rho \rightarrow e^+ e^-$* , Phys. Rev. Lett. **21** (1968) 244.
- 1041 [15] S. M. Flatté, *Coupled-channel analysis of the  $\pi\eta$  and  $KK$  systems near  $KK$  threshold*,  
1042 Physics Letters B **63** (1976), no. 2 224 .
- 1043 [16] BES Collaboration, M. Ablikim *et al.*, *Resonances in  $J/\psi \rightarrow \phi \pi^+ \pi^-$  and  $\phi K^+ K^-$* ,  
1044 Phys. Lett. **B607** (2005) 243, [arXiv:hep-ex/0411001](#).
- 1045 [17] D. V. Bugg, *A study in depth of  $f_0(1370)$* , Eur. Phys. J. **C52** (2007) 55,  
1046 [arXiv:0706.1341](#).

- 1047 [18] LHCb Collaboration, R. Aaij *et al.*, *Analysis of the resonant components in  $B_s \rightarrow$*
- 1048  $J/\psi\pi^+\pi^-$ , Phys. Rev. **D86** (2012) 052006, [arXiv:1204.5643](https://arxiv.org/abs/1204.5643).
- 1049 [19] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- 1050 [20] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys.
- 1051 Rev. **60** (1941) 61.
- 1052 [21] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys.
- 1053 Rev. D **57** (1998) 431.
- 1054 [22] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of*
- 1055  $\psi$  *decay to mesons*, Eur. Phys. J. **A16** (2003) 537, [arXiv:hep-ph/0211457](https://arxiv.org/abs/hep-ph/0211457).
- 1056 [23] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spec-*
- 1057 *troscopy*, Phys. Rev. **D51** (1995) 2247.
- 1058 [24] J.-J. Zhu, *Explicit expressions of spin wave functions*, [arXiv:hep-ph/9906250](https://arxiv.org/abs/hep-ph/9906250).
- 1059 [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).
- 1060 [26] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput.
- 1061 Phys. Commun. **180** (2009) 1847, [arXiv:0805.2956](https://arxiv.org/abs/0805.2956).
- 1062 [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).
- 1063 [28] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth.
- 1064 **A462** (2001) 152.
- 1065 [29] M. Karbach and M. Kenzie, *Gammacombo package*,  
<http://gammacombo.hepforge.org/web/HTML/index.html>, 2014.
- 1066 [30] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007)
- 1067 040, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- 1068 [31] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.
- 1069 [32] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38**
- 1070 (2014) 090001.
- 1071 [33] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys.
- 1072 **A30** (2015) 1530022, [arXiv:1412.6352](https://arxiv.org/abs/1412.6352).
- 1073 [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).
- 1074 [35] LHCb collaboration, L. Zhang, *Measurements of CP violation in  $B_s^0 \rightarrow J/\psi K^+K^-$  decays in the low  $K^+K^-$  mass range with 13 TeV data*, LHCb-ANA-2017-028.

- 1081 [36] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and*  
 1082  *$\tau$ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots  
 1083 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 1084 [37] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*  
 1085 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 1086 [38] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of  $B_s^0$  mesons at*  
 1087 *LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- 1088 [39] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of  $B$  mesons at the*  
 1089 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- 1090 [40] LHCb, R. Aaij *et al.*, *Measurement of  $B^0$ ,  $B_s^0$ ,  $B^+$  and  $\Lambda_b^0$  production asymmetries in 7*  
 1091 *and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- 1092 [41] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of*  
 1093 *the  $K\pi$  Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-  
 1094 2012-027, CERN, Geneva, Feb, 2013.
- 1095 [42] A. Davis *et al.*, *Measurement of the instrumental asymmetry for  $K^- \pi^+$ -pairs at LHCb*  
 1096 *in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN,  
 1097 Geneva, Mar, 2018.
- 1098 [43] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly*  
 1099 *forbidden transitions in  $D \rightarrow K(S)$ ,  $K(L) + \pi$ 's decays*, Phys. Lett. **B349** (1995)  
 1100 363, arXiv:hep-ph/9502238.
- 1101 [44] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*,  
 1102 Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- 1103 [45] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude*  
 1104 *analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- 1105 [46] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal  
 1106 Statistical Society, Series B **58** (1994) 267.
- 1107 [47] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- 1108 [48] H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on  
 1109 Automatic Control **19** (1974) 716.
- 1110 [49] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime*  
 1111 *and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986,  
 1112 DESY-F31-86-02.
- 1113 [50] D. Hill, M. John, and P. Gandini, *A study of partially reconstructed  $B^\pm \rightarrow D^{*0} h^\pm$*   
 1114 *decays using the  $D^0 \rightarrow K\pi$ ,  $KK$ ,  $\pi\pi$  final states*, .
- 1115 [51] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins  
 1116 University Press, Baltimore, MD, USA, 1996.