

# 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  $\gamma - 2\beta_s$  obtained from a time-dependent (amplitude) analysis of  $B_s^0 \rightarrow D_s K \pi\pi$  decays using proton-proton collision data corresponding to an integrated luminosity of  $7 \text{ fb}^{-1}$  recorded by the LHCb detector.



## Change-log

- Version 1.0: First note draft circulated to the WG
- Version 2.0: Implemented first round of comments from Mark, Alessandro, Agnieszka
  - Minor text changes throughout the note
  - Changed  $D^0$  bkg vetoes (Sec. 3.2.2)
  - Revisited some systematics
    - \* Use uncertainty from fake rate instead of fixing misID yield to 0 or twice the value (Sec. 11.2)
    - \* Use GLASS model instead of RBW as alternative to LASS model for  $K\pi$  S-wave (Sec. 11.9)
    - \* Instead of assuming pure phase space, recalculate running width of three-body resonances taking only the dominant  $K\pi\pi$  decay mode into account (Sec. 11.9)
    - \* Calculate RMS of lineshape variations instead of summing them up in quadrature (Sec. 11.9)
    - \* Bug fix in evaluation of fixed lineshape parameters (mass, width, form-factor) systematic (Sec. 11.9)
  - Added tables with alternative amplitude models (Sec. 11.10)
  - Added comparison plots of 16 vs 17 data (Sec. K.2)
- Version 3.0:
  - Changed  $D^0$  bkg vetoes back to the ones from v1.0 to account for additional Kaon misID (Fig. 3.3)
  - Updated analysis with new MC samples (Sec. 3.3)
  - Studied time resolution separately for 16 and 17 data (Sec. 5.2)
  - Changed phasespace acceptance method from BDT to MC integration (Sec. 6.3)
  - Updated analysis with latest OS Tagger tuning (Sec. 7.1)

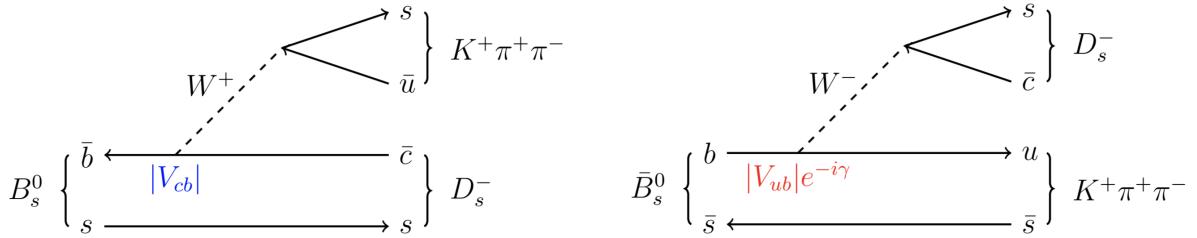
# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Formalism</b>	<b>2</b>
2.1	Decay rates and $CP$ -observables . . . . .	2
2.2	Amplitude model . . . . .	4
2.2.1	Form Factors and Resonance Lineshapes . . . . .	4
2.2.2	Spin Densities . . . . .	6
2.3	Fit implementation . . . . .	8
2.4	Validation . . . . .	9
<b>3</b>	<b>Data samples and event selection</b>	<b>12</b>
3.1	Stripping and Trigger selection . . . . .	12
3.2	Offline selection . . . . .	12
3.2.1	Phase space region . . . . .	13
3.2.2	Physics background vetoes . . . . .	14
3.2.3	Training of multivariate classifier . . . . .	17
3.2.4	Final selection . . . . .	19
3.3	Simulation . . . . .	22
<b>4</b>	<b>Yields determination</b>	<b>23</b>
4.1	Signal model . . . . .	23
4.2	Background models . . . . .	24
4.3	Results . . . . .	26
<b>5</b>	<b>Decay-time Resolution</b>	<b>27</b>
5.1	Calibration for Run-I data . . . . .	28
5.2	Calibration for Run-II data . . . . .	29
<b>6</b>	<b>Acceptance</b>	<b>32</b>
6.1	MC corrections . . . . .	32
6.1.1	Truth matching of simulated candidates . . . . .	32
6.1.2	Correction of data-simulation differences . . . . .	32
6.2	Decay-time acceptance . . . . .	33
6.3	Phase space acceptance . . . . .	37
<b>7</b>	<b>Flavour Tagging</b>	<b>39</b>
7.1	OS tagger combination . . . . .	40
7.2	Tagging performance . . . . .	42
<b>8</b>	<b>Production and Detection Asymmetries</b>	<b>43</b>
8.1	$B_s$ Production Asymmetry . . . . .	43
8.2	$K^-\pi^+$ Detection Asymmetry . . . . .	44
<b>9</b>	<b>Decay-time fit</b>	<b>47</b>
9.1	Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data . . . . .	47
9.2	Fit to $B_s^0 \rightarrow D_sK\pi\pi$ data . . . . .	47

<b>10 Time-dependent amplitude fit</b>	<b>50</b>
10.1 Signal Model Construction . . . . .	50
10.2 Results . . . . .	53
<b>11 Systematic uncertainties</b>	<b>56</b>
11.1 Fit bias . . . . .	56
11.2 Background subtraction . . . . .	56
11.3 Decay-time acceptance . . . . .	57
11.4 Decay-time resolution and tagging . . . . .	57
11.5 Production, detection asymmetries and mixing frequency . . . . .	58
11.6 Multiple candidates . . . . .	58
11.7 Length and momentum scales . . . . .	58
11.8 Phase space acceptance . . . . .	59
11.9 Resonance description . . . . .	59
11.10 Alternative amplitude models . . . . .	60
<b>A Stripping and Trigger cuts</b>	<b>66</b>
<b>B Details of multivariate classifier</b>	<b>68</b>
<b>C Detailed mass fits</b>	<b>72</b>
<b>D Decay-time Resolution fits</b>	<b>76</b>
<b>E Comparison of time-acceptance in subsamples</b>	<b>80</b>
<b>F Comparison of phase-space acceptance in subsamples</b>	<b>80</b>
<b>G OS tagger calibration parameters</b>	<b>83</b>
<b>H Spin Amplitudes</b>	<b>84</b>
<b>I Considered Decay Chains</b>	<b>85</b>
<b>J Data-simulation comparisson</b>	<b>86</b>
<b>K Data distributions</b>	<b>87</b>
K.1 Comparison of signal and calibration channels . . . . .	87
K.2 Comparison of data taken in 2016 and 2017 . . . . .	88
K.3 Comparison of Run-I and Run-II data . . . . .	89
K.4 Comparison of $D_s$ final states . . . . .	90
K.5 Comparison of trigger categories . . . . .	92
<b>References</b>	<b>93</b>

# 1 Introduction

This note presents the first measurement of the CKM angle  $\gamma \equiv \arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$  using  $B_s^0 \rightarrow D_s K\pi\pi$  decays, where the  $K\pi\pi$  subsystem is dominated by excited kaon states such as the  $K_1(1270)$  and  $K_1(1400)$  resonances [1, 2]. In these decays, sensitivity to the weak phase results from the interference between  $b \rightarrow c$  and  $b \rightarrow u$  transitions achieved through  $B_s^0 - \bar{B}_s^0$  mixing [3, 4]. The amplitudes for both processes are of the same order in the Wolfenstein parameters  $\lambda$ ,  $\mathcal{O}(\lambda^3)$ , so that interference effects are expected to be large. The corresponding Feynman diagrams are shown in Fig. 1.1. Due to the interference between mixing and decay amplitudes, the physical  $CP$  violating observables in these decays are functions of a combination of  $\gamma$  and the mixing phase  $\beta_s$ , namely  $\gamma - 2\beta_s$ . To account for the non-constant strong phase across the phase space, one can either perform a time-dependent amplitude fit or select a suitable phase-space region and introduce a coherence factor as additional hadronic parameter to the decay-time fit. The topologically very similar yet flavour specific decay  $B_s \rightarrow D_s \pi\pi\pi$  is used as calibration channel, not only to calibrate the tagging algorithms and determine the decay-time acceptance but also to constrain the  $B_s^0 - \bar{B}_s^0$  mixing frequency. The branching ration of these decay modes was measured by LHCb to be  $\frac{\mathcal{B}(B_s^0 \rightarrow D_s K\pi\pi)}{\mathcal{B}(B_s^0 \rightarrow D_s \pi\pi\pi)} = 0.052 \pm 0.005(\text{stat}) \pm 0.003(\text{syst})$  [1, 2].



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

## 18 2 Formalism

### 19 2.1 Decay rates and *CP*-observables

20 The differential decay rate of  $B_s^0$  or  $\bar{B}_s^0$  decays to the final state  $D_s^- K^+ \pi\pi$  or  $D_s^+ K^- \pi\pi$  is  
 21 given by:

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

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

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

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

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

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

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

42 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,  
 43 the four-body phase space density function is not flat in the usual kinematic variables.  
 44 Therefore, an analytic expression for  $\phi_4$  is taken from Ref. [5].

<sup>45</sup> Assuming there is no direct  $CP$  violation in the  $B_s$  decay implies for the  $CP$  conjugate  
<sup>46</sup> transition amplitudes:

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

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

<sup>47</sup> where the  $CP$ -conjugate phase space point  $\bar{\mathbf{x}}$  is defined such that it is mapped onto  $\mathbf{x}$  by  
<sup>48</sup> the interchange of final state charges, and the reversal of three-momenta.

<sup>49</sup> The phenomenological description of the intermediate state amplitudes is discussed  
<sup>50</sup> in Sec. 2.2. For a model-independent measurement, the differential decay rate can be  
<sup>51</sup> integrated over the phase space:

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

<sup>52</sup> where the same convention for the  $CP$  coefficients as for the  $B_s \rightarrow D_s K$  analysis is used:

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

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

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

<sup>53</sup> The coherence factor  $\kappa$ , the strong phase difference  $\delta$  and the ratio of the suppressed  
<sup>54</sup> ( $b \rightarrow u$ ) over favored ( $b \rightarrow c$ ) decay mode, averaged over the phase space, are defined as:

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

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

<sup>55</sup> The coherence factor dilutes the sensitivity to the weak phase  $\gamma$  due to the integration  
<sup>56</sup> over the interfering amplitudes across the phase space. The value of  $\kappa$  is bounded between  
<sup>57</sup> zero and unity. The latter corresponds to the limit of only one contributing intermediate  
<sup>58</sup> state in which case the same sensitivity as in  $B_s \rightarrow D_s K$  decays is reached, while  $\kappa = 0$   
<sup>59</sup> would result in no sensitivity to  $\gamma$  at all.

## 60 2.2 Amplitude model

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

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

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

### 71 2.2.1 Form Factors and Resonance Lineshapes

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

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

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

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

80 where the total width,  $\Gamma(s)$ , is normalized to give the nominal width,  $\Gamma_0$ , when evaluated  
 81 at the nominal mass  $m_0$ .

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

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

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

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

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

and therefore requires knowledge of the resonant substructure. The three-body amplitude  $A_{R \rightarrow ABC}$  can be parameterized similarly to the four-body amplitude in Eq. (2.13). In particular, it includes form factors and propagators of intermediate two-body resonances.

Both Eq. (2.16) and Eq. (2.17) give only the partial width for the decay into a specific channel. To obtain the total width, a sum over all possible decay channels has to be performed,

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

where the coupling strength to channel  $i$ , is given by  $g_i$ .

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

For the broad scalar resonance  $\sigma$ , the model from Bugg is used [13]. We use the Gounaris-Sakurai parametrization for the  $\rho(770)^0 \rightarrow \pi\pi$  propagator [14]. For the decay chain  $K_1(1270) \rightarrow \rho(770)K$ , we include  $\rho-\omega$  mixing with the relative magnitude and phase between  $\rho$  and  $\omega$  fixed to the values determined in Ref. [15]. The energy-dependent width of the  $f_0(980)$  resonance is given by the sum of the partial widths into the  $\pi\pi$  and  $KK$  channels (*i.e.* the Flatté lineshape [16]), where the coupling constants as well as the mass and width are taken from a measurement performed by the BES Collaboration [17]. For the  $f_2(1270)$  and the  $f_0(1370)$  mesons we use the total decay widths calculated in Ref. [9] which take the channels  $\pi\pi$ ,  $KK$ ,  $\eta\eta$  and  $\pi\pi\pi\pi$  into account. The Lass parameterization is used to model the  $K\pi$   $S$ -wave contribution. It consists of the  $K_0^*(1430)$  resonance together with an effective range non-resonant component [18–20]:

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

with  $\cot \delta_L = \frac{1}{aq} + \frac{1}{2}rq$ . We use the values for the scattering length  $a$  and effective range parameter  $r$  from Ref. [18, 19]. Equation (2.16) is used for all other resonances decaying into a two-body final state.

For the resonances  $K_1(1270)$  and  $K(1460)$ , the energy-dependent widths as well as the nominal mass and width are taken from Ref. [21]. We further use the energy-dependent widths for the  $K_1(1400)$ ,  $K^*(1410)$  and  $K^*(1680)$  mesons from Ref. [9]. For all other resonances decaying into a three-body final state, an energy-dependent width distribution is derived from Equation 2.17 assuming an uniform phase space population.

Some particles may not originate from a resonance but are in a state of relative orbital angular momentum. We denote such non-resonant states by surrounding the particle system with brackets and indicate the partial wave state with an subscript; for example  $(\pi\pi)_S$  refers to a non-resonant di-pion  $S$ -wave. The lineshape for non-resonant states is set to unity.

## 124 2.2.2 Spin Densities

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

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

137 The spin projection operator  $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$ , for a resonance  $R$ , with spin  $S =$   
 138  $\{0, 1, 2\}$ , and four-momentum  $p_R$ , is given by [26]:

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

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

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

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

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

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

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

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

151 where

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

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

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

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

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

165 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  
 166 pseudoscalar initial/final states are involved.

167 The spin factors for all decay topologies considered in this analysis are explicitly given  
 168 in Appendix H.

## 169 2.3 Fit implementation

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

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

171 This allows us to set more intuitive starting values as the amplitude coefficients are all on  
 172 a comparable scale. Moreover, the total amplitudes  $\mathcal{A}_f^{c(u)}(\mathbf{x})$  are renormalized on-the-fly  
 173 (*i.e.* at each minimization step) such that

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

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

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

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

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

## 193 2.4 Validation

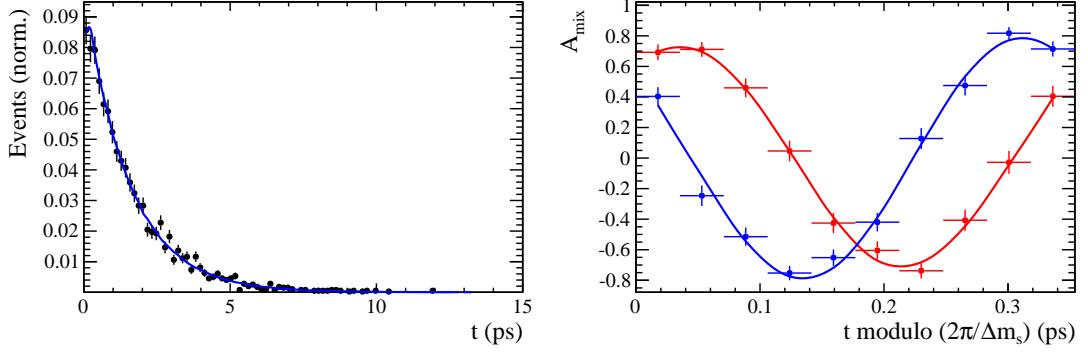
194 The amplitude fit is performed by using the amplitude analysis tool **MINT2**. It was  
 195 previously applied to analyze  $D^0 \rightarrow 4\pi$  and  $D^0 \rightarrow KK\pi\pi$  decays [9] which have an  
 196 identical general spin structure (*i.e.* scalar to four scalar decay) to  $B_s \rightarrow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

201 This does, however, not apply to the full time-dependent amplitude pdf which is  
 202 newly implemented for this analysis. To cross-check it, we use **EVTGEN** to generate  
 203 toy events with time-dependent  $CP$  violation according to the **SSD\_Cp** event model [29].  
 204 Since this event model does not allow for multiple interfering resonances, we generate  
 205 only the decay chain  $B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ . Table 2.1 lists the generated  
 206 input parameters. The toy data set is fitted with our **MINT2** implementation of the full  
 207 time-dependent amplitude pdf and the phasespace-integrated pdf. Flat decay-time and  
 208 phase-space acceptances are used. The true ID of the  $B_s$  meson is used to tag the initial  
 209 flavor (*i.e.* perfect tagging). For technical reasons (the extreme case of perfect resolution  
 210 is currently not implemented in our fitter) we smeared the generated decay-times with a  
 211 Gaussian of width 45 fs. We use a Gaussian of the same width to convolute the pdf for  
 212 fitting. More details of the fit procedures are given in Secs. 9 and 10. The fit projections  
 213 are shown in Figs. 2.1 and 2.2.

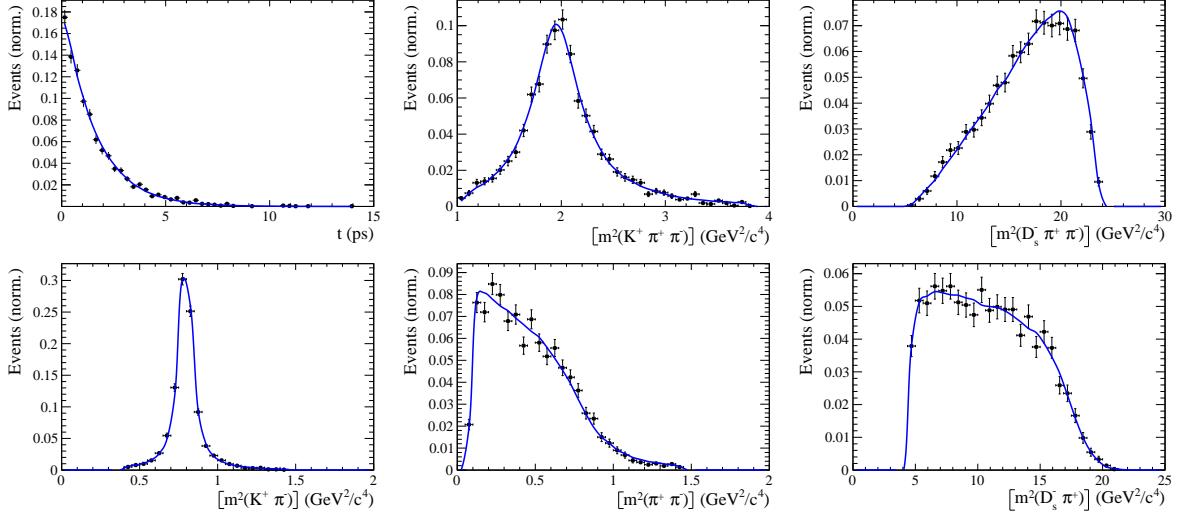
214 The  $CP$  coefficients  $C, D, \bar{D}, S, \bar{S}$  are the fit parameters in case of the phasespace-  
 215 integrated pdf, which are converted after to the fit to the physical observables  $r, \kappa, \delta$  and  $\gamma$   
 216 using the **GammaCombo** package [30]. The obtained 1-CL contours are shown in Fig. 2.3.  
 217 The full pdf determines  $r, \delta$  and  $\gamma$  directly. As shown in Tab. 2.2 and 2.3, the fit results  
 218 are in excellent agreement with the generated input values. The phasespace-integrated fit  
 219 is, in addition, performed with the **B2DX** fitter used for the  $B_s \rightarrow D_s K$  analysis yielding  
 220 identical results. Note though that some parts of the **B2DX** fitter have been taken over to  
 221 our **MINT2** fitter, such that the implementations are not fully independent.

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

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



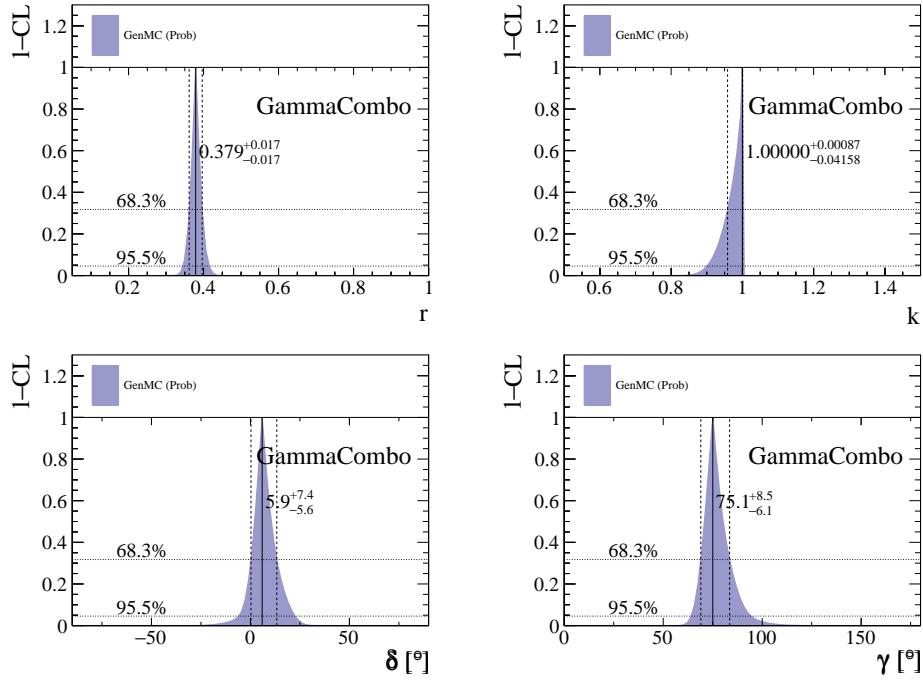
**Figure 2.1:** Left: Time distribution of  $B_s \rightarrow D_s K \pi \pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (solid line). Right: Time-dependent asymmetry between mixed and unmixed events folded into one oscillation period for  $D_s^- K^+ \pi \pi$  (red)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



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

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

	Generated	Fit result	Pull( $\sigma$ )
$C$	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 2.3:** The 1-CL contours for the physical observable  $r, \kappa, \delta$  and  $\gamma$  obtained with the phasespace-integrated fit to the EVTGEN toy sample.

**Table 2.3:** Results for the physical observables obtained from fits to EVTGEN toy events.

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

## 222 3 Data samples and event selection

### 223 3.1 Stripping and Trigger selection

224 The dataset used for this analysis corresponds to  $1 \text{ fb}^{-1}$  of proton-proton collision data col-  
225 lected in 2011 with a centre of mass energy  $\sqrt{s} = 7 \text{ TeV}$ ,  $2 \text{ fb}^{-1}$  collected in 2012 with  $\sqrt{s} =$   
226  $7 \text{ TeV}$  and  $4 \text{ fb}^{-1}$  collected in 2015/2016/2017 with  $\sqrt{s} = 13 \text{ TeV}$ . Candidate  $B_s^0 \rightarrow D_s K \pi \pi$   
227 ( $B_s^0 \rightarrow D_s \pi \pi \pi$ ) decays are reconstructed using the `B02DKPiPiD2HHHPIDBeauty2CharmLine`  
228 (`B02DPiPiD2HHHPIDBeauty2CharmLine`) line of the `BHadronCompleteEvent` stream of  
229 `Stripping21r1` (2011), `Stripping21` (2012), `Stripping24r1` (2015) and `Stripping28r1p1` (2016)  
230 and `Stripping29r2` (2017). Both stripping lines employ the same selection cuts, listed in  
231 Appendix A, except for the PID requirement on the bachelor kaon/pion.

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

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

### 244 3.2 Offline selection

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

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

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

- 259 • DIRA > 0.99994
- 260 • min IP  $\chi^2 < 16$  to the best PV,
- 261 • FD  $\chi^2 > 100$  to the best PV,
- 262 • Vertex  $\chi^2/\text{nDoF} < 8$ .

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

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

### 280 3.2.1 Phase space region

281 Due to the comparably low masses of the final state particles with respect to the  $B_s$   
282 mass, there is, in contrast to for example b-hadron to charmonia or charm decays, a  
283 huge phase-space available for the  $B_s^0 \rightarrow D_s K\pi\pi$  decay. For the invariant mass of  
284 the  $K\pi\pi$  subsystem it extends up to 3.4 GeV. It has however been observed that the  
285 decay proceeds predominantly through the low lying axial vector states  $K_1(1270)$  and  
286  $K_1(1400)$ , while the combinatorial background is concentrated at high  $K\pi\pi$  invariant  
287 masses ( $m(K\pi\pi) > 2000$  MeV). Moreover, the strange hadron spectrum above 2 GeV  
288 is poorly understood experimentally such that a reliable extraction of the strong phase  
289 motion in that region is not possible. We consequently choose to limit the considered  
290 phase space region to  $m(K\pi\pi) < 1900$  MeV, which is right below the charm-strange  
291 threshold ( $B_s^0 \rightarrow D_s^+ D_s^-$ ).

292 **3.2.2 Physics background vetoes**

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

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

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

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

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

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

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

307  $D^0$  combined with a random  $\pi$  can fake a  $D_s \rightarrow KK\pi$  decay,

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

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

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

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

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

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

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

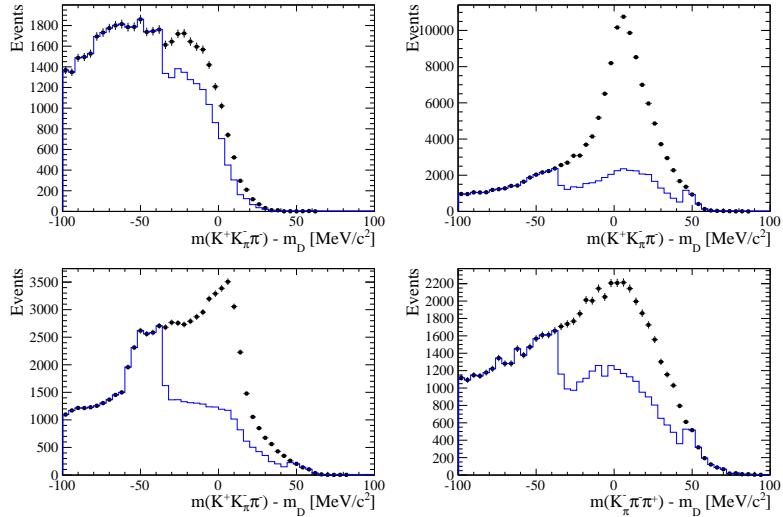
317 Possible with  $\bar{p} \rightarrow K^-$  misidentification, vetoed by requiring  $m(K_p^- \pi^+ \pi^-) \neq$   
 318  $m(\Lambda_c^-) \pm 40$  MeV or  $\text{PIDK}(K^-) - \text{PIDp}(K^-) > 5$ .

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

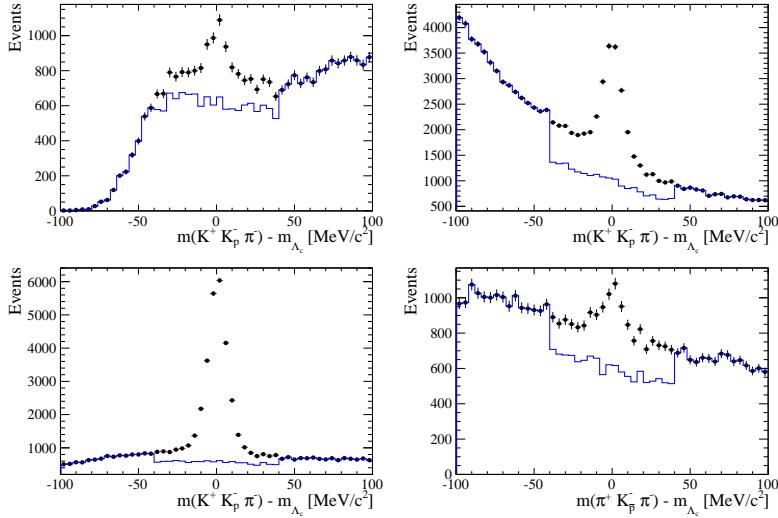
320  $D^0$  combined with a random  $\pi$  can fake a  $D_s \rightarrow K\pi\pi$  decay, vetoed by requiring  
 321 both possible combinations to have  $m(K\pi) < 1750$  MeV.

322 The effects of these veto cuts are illustrated in Figs. 3.1,3.2 and 3.3. To reduce cross-feed  
 323 from our calibration channel into the signal channel and vice-versa we require tight PID  
 324 cuts on the ambiguous bachelor track; for the signal channel we apply  $\text{PIDK}(K^+) > 10$   
 325 and for the calibration channel  $\text{PIDK}(\pi^+) < 0$ . In addition, we veto  $B_s^0 \rightarrow D_s^- D_s^+$  decays  
 326 which is illustrated in Fig. 3.4.

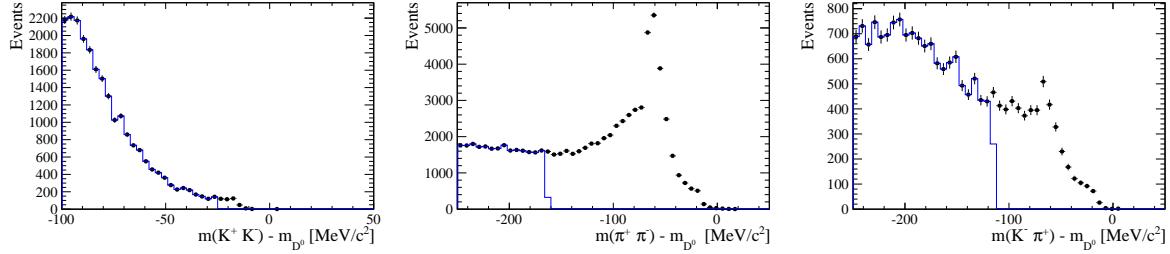
- 327 1.  $X_s^+ \rightarrow K^+\pi^+\pi^-$ :
- 328 (a)  $B_s^0 \rightarrow D_s\pi\pi\pi$ :  
329 Possible with  $\pi^+ \rightarrow K^+$  misidentification, suppressed with  $\text{PIDK}(K^+) > 10$ .
- 330 (b)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^+\pi^-)$ :  
331 Outside of considered phase-space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



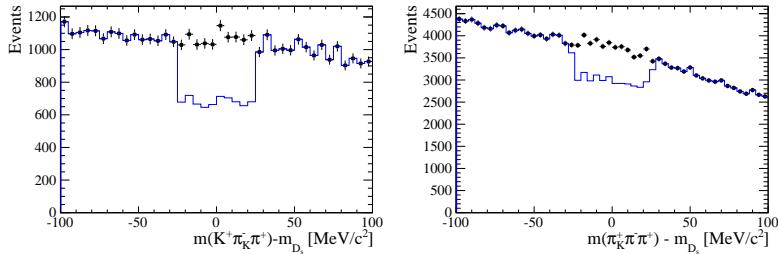
**Figure 3.1:** Background contributions from  $D^-$  decays where the  $\pi^-$  is misidentified as  $K^-$ . The  $D_s$  invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*(892)K$ ,  $D_s \rightarrow KK\pi$  (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. The peak at  $m(\pi\pi) - m(D^0) \approx -60$  MeV ( $m(K\pi) - m(D^0) \approx -60$  MeV) are due to  $D^0 \rightarrow K\pi$  ( $D^0 \rightarrow KK$ ) where a kaon is misidentified as 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.

346 **3.2.3 Training of multivariate classifier**

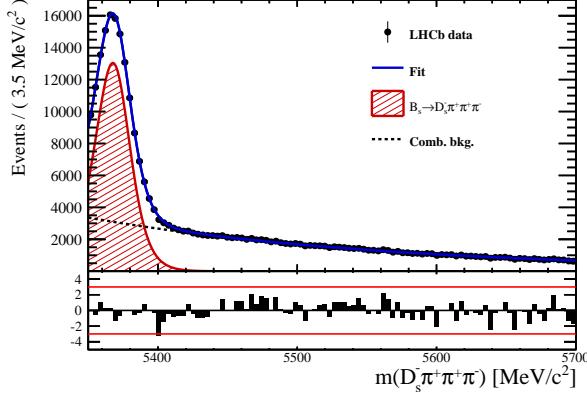
347 The Toolkit for Multivariate Analysis (TMVA [31]) is used to train a multivariate classifier  
348 (BDT with gradient boosting, BDTG) discriminating signal and combinatorial background.  
349 We use  $B_s \rightarrow D_s \pi\pi$  data that passes the preselection as signal proxy. The background 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

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

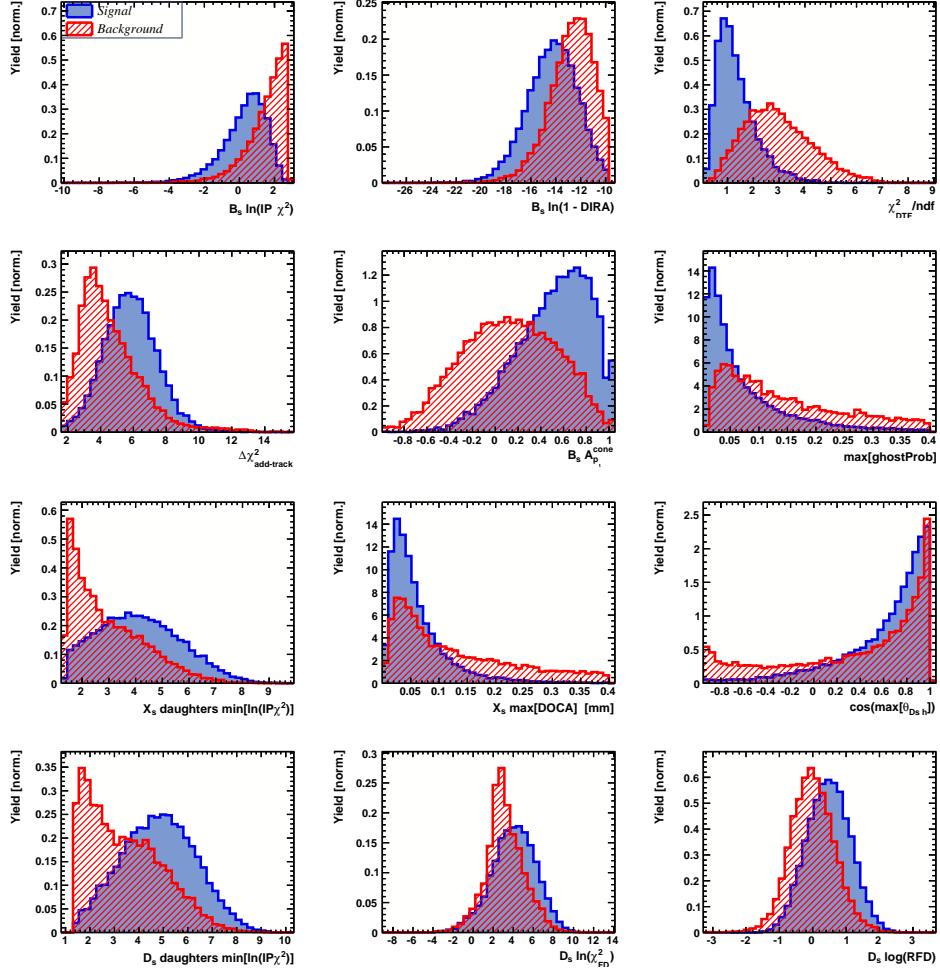
360 The following discriminating variables are used for the BDTG training:

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

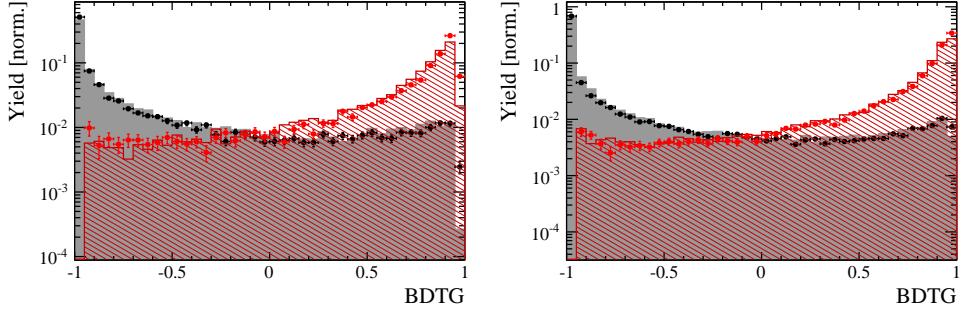
377 Loose cuts on the variables  $\chi_{DTF}^2/ndf$ ,  $\Delta\chi_{add-track}^2$  and  $\cos(\max \theta_{D_s h_i})$  are applied prior  
378 to the training which are expected to be 100% signal efficient. Figure 3.6 shows the  
379 distributions of the input variables for signal and background. As shown in Appendix B,  
380 these distributions differ between data-taking period and trigger category. In particular  
381 variables depending on the  $B_s$  kinematics and the event multiplicity are affected (e.g.  
382  $\theta_{D_s h_i}$  or  $A_{pT}^{\text{cone}}$ ). The BDTG is consequently trained separately for these categories. The  
383 resulting classifier response is shown in Fig. 3.7 for each category (even and odd test  
384 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

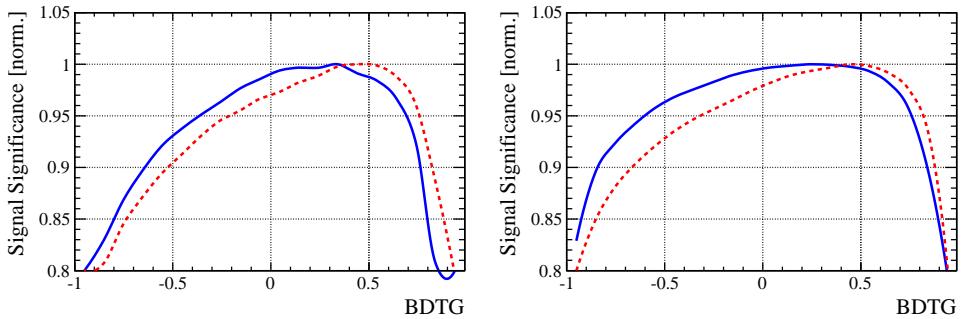
The cut on the BDTG output is optimized using the signal significance as figure of merit:

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

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 $> -10$ $> 0$ $< 10$ $> 0$ $> 0$ $m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 40$ 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		$> 0.35 \text{ [Run-I,L0-TOS]}$ $> 0.45 \text{ [Run-I,L0-TIS]}$ $> 0.25 \text{ [Run-II,L0-TOS]}$ $> 0.45 \text{ [Run-II,L0-TIS]}$
$X_s^+ \rightarrow K^+\pi^+\pi^-$	PIDK(K)	$> 10$
	PIDK( $\pi^+$ )	$< 10$
	PIDK( $\pi^-$ )	$< 0$
	$D_s$ veto	$m(K^+\pi^+\pi_K^-) \neq m(D_s) \pm 20 \text{ MeV} \parallel \text{PIDK}(\pi^-) < -5$
$X_d^+ \rightarrow \pi^+\pi^+\pi^-$	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	$= 1$

### 396 3.3 Simulation

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

**Table 3.3:** List of simulated samples used in the analysis.

Decay	Event Type	Sim	Statistics				Filter
			11	12	15	16	
$B_s \rightarrow (D_s \rightarrow KK\pi)K\pi\pi$	13266007	08i	1.2 M	1.2 M	-	-	Generator Level
$B_s \rightarrow (D_s \rightarrow KK\pi)K\pi\pi$	13266008	09c	70 k	80 k	50 k	100 k	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow K\pi\pi)K\pi\pi$	13266058	09c	70 k	80 k	50 k	100 k	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow \pi\pi\pi)K\pi\pi$	13266038	09c	70 k	80 k	50 k	100 k	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow KK\pi)\pi\pi\pi$	13266006	08i	1.2 M	1.2 M	-	-	Generator Level
$B_s \rightarrow (D_s \rightarrow KK\pi)\pi\pi\pi$	13266068	09c	70 k	80 k	50 k	100 k	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow K\pi\pi)\pi\pi\pi$	13266088	09c	70 k	80 k	50 k	100 k	Generator Level, Stripping
$B_s \rightarrow (D_s \rightarrow \pi\pi\pi)\pi\pi\pi$	13266078	09c	70 k	80 k	50 k	100 k	Generator Level, Stripping
$B_s \rightarrow D_s^*\pi\pi\pi, D_s \rightarrow KK\pi$	13266201	08i	1.2 M	1.2 M	-	-	Generator Level

## 4 Yields determination

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

Due to different mass resolutions, we perform the invariant mass fits simultaneously for each data-taking period and each trigger category. We further introduce four  $D_s$  final state categories:  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*(892)\pi$ ,  $D_s \rightarrow \pi\pi\pi$  and  $D_s \rightarrow Kh\pi$  to account for different signal purities. The  $D_s \rightarrow Kh\pi$  category combines the two  $D_s$  decay channels with the lowest statistics:  $D_s \rightarrow KK\pi$  (non-resonant) and  $D_s \rightarrow K\pi\pi$ . This amounts to 16 categories in total.

### 4.1 Signal model

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

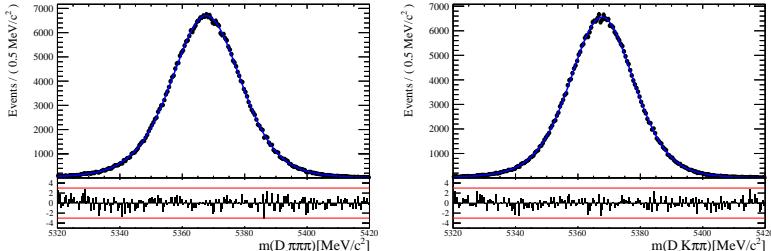
$$\mathcal{J}(x|\mu, \sigma, \nu, \tau) = \frac{e^{-\frac{1}{2}r^2}}{2\pi \cdot c \cdot \sigma \cdot \tau \cdot \sqrt{z^2 + 1}} \quad (4.1)$$

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

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

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

It is conveniently expressed in terms of the central moments up to order four: The mean of the distribution  $\mu$ , the standard deviation  $\sigma$ , the skewness  $\nu$  and the kurtosis  $\tau$ . The tail parameters  $\nu$  and  $\tau$  are fixed to the values obtained by a fit to the invariant mass distribution of simulated events shown in Fig 4.1. To account for differences between simulation and real data, linear scaling factors for the mean  $\mu$  and width  $\sigma$  are determined in the fit to  $B_s^0 \rightarrow D_s \pi\pi\pi$  data and later fixed in the fit to  $B_s^0 \rightarrow D_s K\pi\pi$  decays. The scale factors are determined separately for each data-taking period and each trigger category.



**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.

426 **4.2 Background models**

427 After the full selection the following residual background components have to be accounted  
428 for:

429

430 **Combinatorial background**

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

434

435 **Peaking  $B_d$  background**

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

438

439 **Partially reconstructed background**

440 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  
441 to be peaking lower than signal in the  $m(D_s \pi\pi\pi)$  spectrum with large tails due to the  
442 momentum carried away by the not reconstructed  $\pi^0$  or  $\gamma$ . An empirical description for  
443 the shape of this contribution is derived from a  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  MC sample subject to  
444 the nominal  $B_s^0 \rightarrow D_s \pi\pi\pi$  selection. Figure 4.2 (top) shows the respective reconstructed  
445  $m(D_s \pi\pi\pi)$  distribution. A sum of three bifurcated Gaussian functions (*i.e.* Gaussian  
446 functions with different widths on the left and the right side of the maximum value) is used  
447 to describe it. In the fit to data, all parameters are fixed to the ones obtained from MC  
448 except for the parameter which describes the width of the right tail of the distribution to  
449 account for data-simulation differences in mass resolution. The equivalent  $B_s^0 \rightarrow D_s^* K\pi\pi$   
450 component contributing to the  $B_s^0 \rightarrow D_s K\pi\pi$  data sample is described by the same PDF  
451 with the right tail fixed to the  $B_s^0 \rightarrow D_s \pi\pi\pi$  result.

452 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  
453 shifted by  $m_{B_s^0} - m_{B^0}$ .

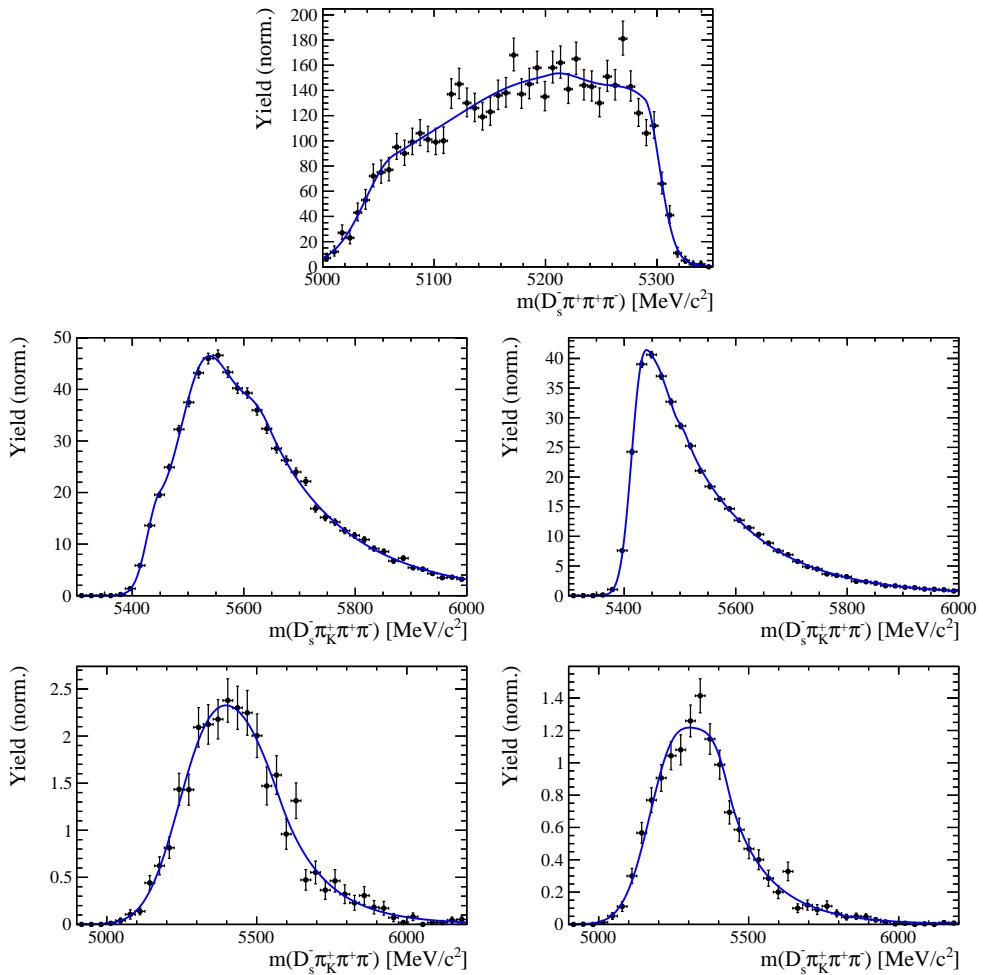
454

455 **Misidentified background**

456 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  
457 pions is misidentified as a kaon, contaminate the  $B_s^0 \rightarrow D_s K^+ \pi^+ \pi^-$  sample. To determine  
458 the corresponding background shapes, we use simulated events passing the nominal  
459 selection except for the PID cuts on the bachelor  $\pi^+$  tracks. The **PIDCalib** package  
460 is used to determine the  $p_T, \eta$ -dependent  $\pi^+ \rightarrow K^+$  misidentification probability for  
461 each pion. We change the particle hypothesis from pion to kaon for the pion with the  
462 higher misidentification probability and recompute the invariant  $B_s^0$  mass,  $m(D_s^- \pi_K^+ \pi^+ \pi^-)$ .  
463 Similarly, the invariant masses  $m(\pi_K^+ \pi^+ \pi^-)$  and  $m(\pi_K^+ \pi^-)$  are recomputed and required  
464 to be within the considered phasespace region. The background distributions are shown  
465 in Fig. 4.2 (middle, bottom) and modeled by the sum of three Crystal Ball functions.

466 The expected yield of misidentified  $B_s^0 \rightarrow D_s\pi\pi\pi$  ( $B_s^0 \rightarrow D_s^*\pi^+\pi^+\pi^-$ ) candidates in  
 467 the  $B_s^0 \rightarrow D_sK\pi\pi$  sample is computed by multiplying the fake rate (within the considered  
 468  $B_s$  mass range) of  $0.63 \pm 0.01\%$  ( $0.55 \pm 0.02\%$ ) for Run-I and  $0.33 \pm 0.01\%$  ( $0.24 \pm 0.01\%$ )  
 469 for Run-II by the  $B_s^0 \rightarrow D_s\pi\pi\pi$  ( $B_s^0 \rightarrow D_s^*\pi^+\pi^+\pi^-$ ) yield as determined in the mass  
 470 fit to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  data sample. The yields are corrected for the  $\text{PID}(\pi^+) < 0$   
 471 requirement which has an efficiency of  $77.1 \pm 0.1\%$  for Run-I and  $81.0 \pm 0.1\%$  for Run-II  
 472 data. The  $B_s^0 \rightarrow D_s^*\pi^+\pi^+\pi^-$  yield is additionally corrected for the efficiency of the cut  
 473  $m(D_sK\pi\pi) > 5200 \text{ MeV}$  evaluated on MC. In the fit to data, the misidentified background  
 474 yields are fixed to the predicted ones.

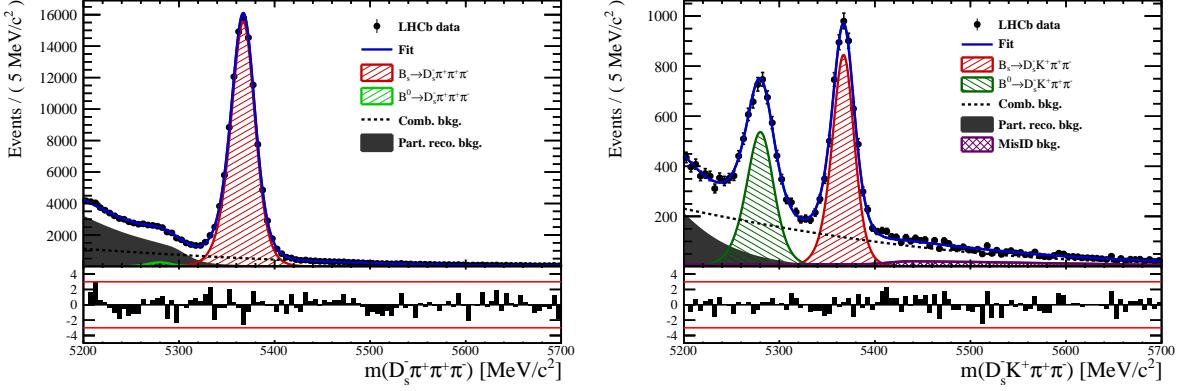
475 We consider the  $B_s^0 \rightarrow D_sK\pi\pi$  and  $B_s^0 \rightarrow D_s^*K\pi\pi$  components contributing to the  
 476  $B_s^0 \rightarrow D_s\pi\pi\pi$  data sample to be negligible due to the low branching fractions and the  
 477 tight PID cuts on the bachelor pions.



**Figure 4.2:** Top: 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 for Run-I (left) and Run-II (right), where one of the pions is reconstructed as a kaon taking the misidentification probability into account. Bottom: Invariant mass distribution for simulated  $B_s^0 \rightarrow D_s^*\pi\pi\pi$  events for Run-I (left) and Run-II (right), 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 PDFs are shown in blue.

### 4.3 Results

Figure 4.3 shows the invariant mass distribution for  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$  candidates passing all selection criteria. The projections for all categories of the simultaneous fit are shown in Appendix C. 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  $B_s^0 \rightarrow D_sK\pi\pi$  (right) candidates.

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

Component	Yield
$B_s \rightarrow D_s\pi\pi\pi$	$104176 \pm 356$
$B^0 \rightarrow D_s\pi\pi\pi$	$1742 \pm 363$
Partially reconstructed bkg.	$43157 \pm 407$
Combinatorial bkg.	$40992 \pm 455$

$D_s$ final state	Signal yield
$D_s^- \rightarrow \phi^0(1020)\pi^-$	$35429 \pm 202$
$D_s^- \rightarrow K^{*0}(892)K^-$	$29388 \pm 194$
$D_s^- \rightarrow (K^- h^+\pi^-)$	$21695 \pm 164$
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	$17665 \pm 148$

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

Component	Yield
$B_s \rightarrow D_sK\pi\pi$	$5172 \pm 88$
$B^0 \rightarrow D_sK\pi\pi$	$4109 \pm 100$
Partially reconstructed bkg.	$1825 \pm 204$
Misidentified bkg.	$1186 \pm 0$
Combinatorial bkg.	$9172 \pm 221$

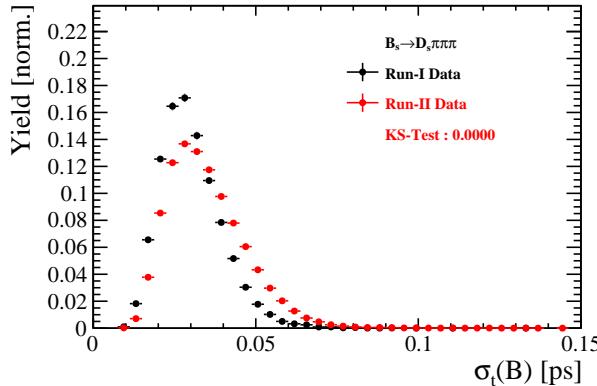
$D_s$ final state	Signal yield
$D_s^- \rightarrow \phi^0(1020)\pi^-$	$1637 \pm 47$
$D_s^- \rightarrow K^{*0}(892)K^-$	$1537 \pm 47$
$D_s^- \rightarrow (K^- h^+\pi^-)$	$1134 \pm 41$
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	$864 \pm 41$

## 483 5 Decay-time Resolution

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

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

504 For Run-II data, a new lifetime unbiased stripping line (LTUB) has been implemented  
 505 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).

## 506 5.1 Calibration for Run-I data

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

$$515 \quad \mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s^2 / 2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s^2 / 2}, \quad (5.1)$$

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

$$518 \quad \sigma_{eff} = \sqrt{(-2/\Delta m_s^2) \ln \mathcal{D}}, \quad (5.2)$$

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

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

$$529 \quad \sigma_{eff}^{MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.232 \pm 0.010) \sigma_t \quad (5.3)$$

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

$$531 \quad \sigma_{eff}^{D_s K, MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.201 \pm 0.013) \sigma_t. \quad (5.4)$$

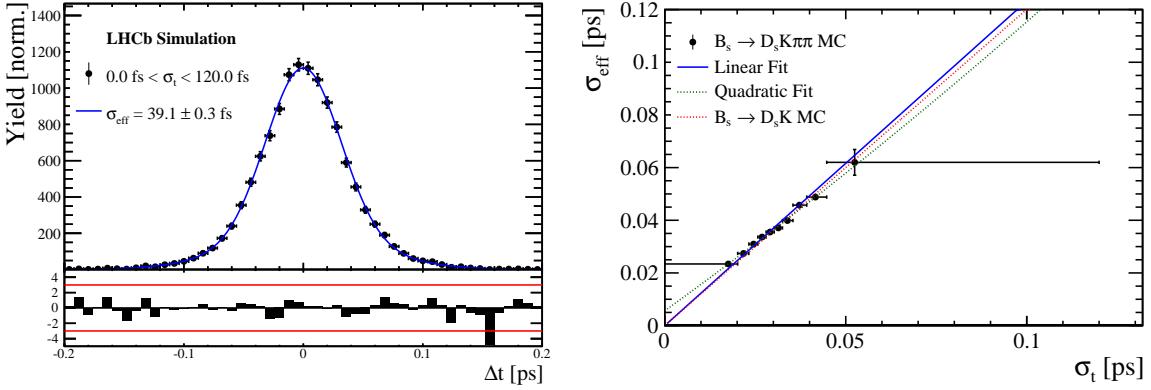
531 Due to the good agreement between the scale factors for  $B_s^0 \rightarrow D_s K$  and  $B_s^0 \rightarrow D_s K \pi\pi$   
 532 MC, we conclude that the resolution calibration for  $B_s^0 \rightarrow D_s K$  data:

$$533 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (10.26 \pm 1.52) \text{ fs} + (1.280 \pm 0.042) \sigma_t \quad (5.5)$$

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

$$535 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (-0.568 \pm 1.570) \text{ fs} + (1.243 \pm 0.044) \sigma_t \quad (5.6)$$

$$535 \quad \sigma_{eff}^{D_s K}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.772 \pm 0.012) \sigma_t \quad (5.7)$$



**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.

## 5.2 Calibration for Run-II data

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

The offline selection of the fake  $B_s$  candidates is summarized in Tab. 5.1. The selection is similar than the one for real data with the important difference that the  $D_s$  candidate is required to come from the PV by cutting on the impact parameter significance. Even after the full selection, a significant number of multiple candidates is observed. These are predominantly fake  $B_s$  candidates that share the same  $D_s$  candidate combined with different random tracks from the PV. We select one of these multiple candidates randomly which retains approximately 20% of the total candidates. As can be seen in Figure C.3, the shapes of the distributions of the unscaled decay time error  $\sigma_t$  for data taken in 2016 and 2017 are significantly different. Therefore, the scaling of the decay time error is treated separately for 2015+2016 and 2017 data. The invariant mass distribution of the selected  $D_s$  candidates is shown in Fig. 5.3. To separate true  $D_s$  candidates from random combinations, the sPlot method is used to statistically subtract combinatorial background from the sample.

Figure 5.4 and 5.5 show the sWeighted decay-time distributions for fake  $B_s$  candidates from 2016 and 2017 data, respectively. Similar as in the previous section, the decay-time distribution is fitted with a double-Gaussian resolution model in slices of the per-event decay time error. Since some  $D_s$  candidates might actually originate from true  $B_s$  decays, the decay-time distribution of the fake  $B_s$  candidates might show a bias towards positive decay times. Therefore, we determine the decay-time resolution from the negative decay-time distribution only. Details of the fit results in each slice are shown in appendix D.

<sup>566</sup> The resulting calibration functions are:

$$\sigma_{eff}^{Data,16}(\sigma_t) = (11.6 \pm 1.6) \text{ fs} + (0.877 \pm 0.040) \sigma_t \quad (5.8)$$

<sup>567</sup>

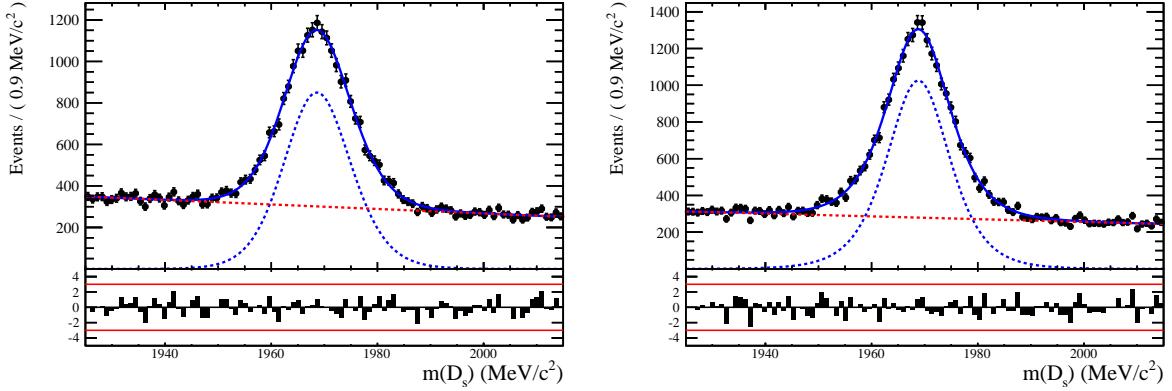
$$\sigma_{eff}^{Data,17}(\sigma_t) = (6.5 \pm 1.4) \text{ fs} + (0.961 \pm 0.036) \sigma_t \quad (5.9)$$

<sup>568</sup> For 2016, the result is in good agreement with the one obtained for the  $B_s \rightarrow J/\psi\phi$   
<sup>569</sup> (Run-II) analysis that uses 2016 data [36].

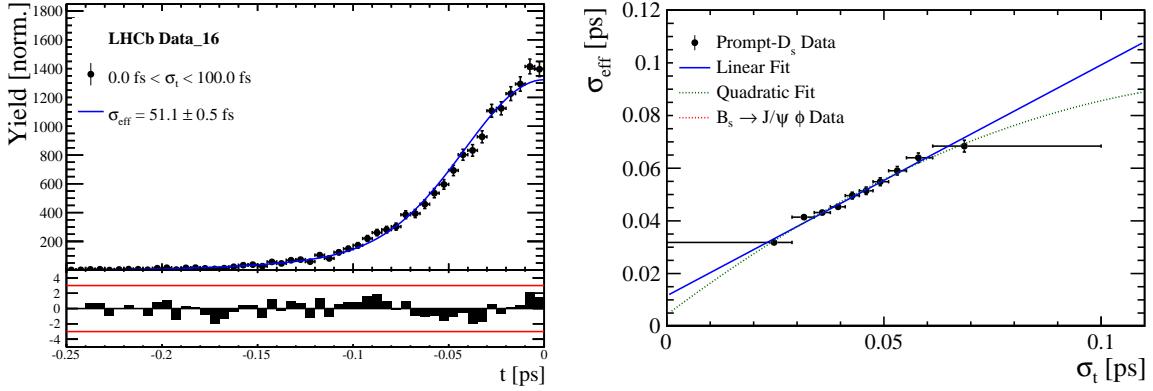
$$\sigma_{eff}^{J/\psi\phi}(\sigma_t) = (12.25 \pm 0.33) \text{ fs} + (0.8721 \pm 0.0080) \sigma_t \quad (5.10)$$

**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. The PID and veto cuts depending on the  $D_s$  final state and Dalitz plot position are the same as in Table. 3.1.

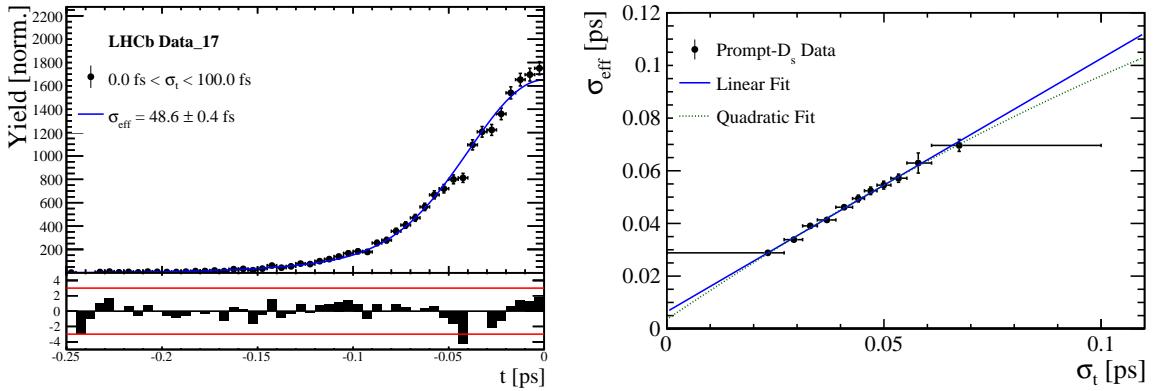
	Description	Requirement
$B_s \rightarrow D_s K\pi\pi$	$\chi^2_{vtx}/\text{ndof}$	< 8
	$\chi^2_{DTF}/\text{ndof}$	< 15
	$t$	< 0 ps
$D_s \rightarrow hh$	$\chi^2_{vtx}/\text{ndof}$	< 5
	DIRA	> 0.99994
	$\chi^2_{FD}$	> 9
	$p_T$	> 1800 MeV
	$\chi^2_{IP}$	< 9
	$\chi^2_{IP}(h)$	> 5
Wrong PV veto	nPV = 1    min( $\Delta\chi^2_{IP}$ ) > 20	
$X_s \rightarrow K\pi\pi$	$\chi^2_{IP}(h)$	< 40
	PIDK(K)	> 10
	PIDK( $\pi$ )	< 5
	isMuon(h)	False
All tracks	$p_T$	> 500 MeV



**Figure 5.3:** The invariant mass distribution for prompt  $D_s$  candidates for data taken from the LTUB stripping line in (left) 2016 and (right) 2017.



**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  $\sigma_{\text{eff}}$  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. Data taken in 2016.



**Figure 5.5:** (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  $\sigma_{\text{eff}}$  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. Data taken in 2017.

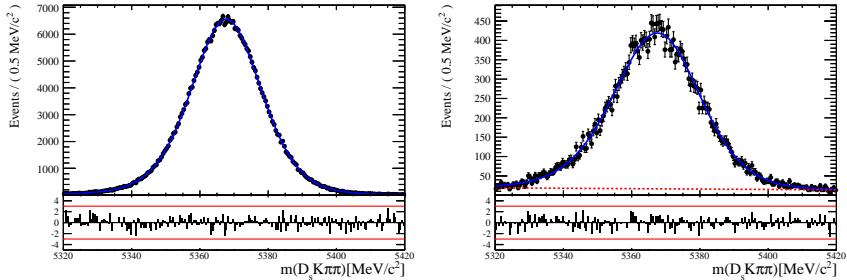
## 570 6 Acceptance

### 571 6.1 MC corrections

#### 572 6.1.1 Truth matching of simulated candidates

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

589 A significant fraction of 7% of the true signal candidates are classified as ghosts, while  
 590 only 20% of the events classified with `BKGCAT`= 60 are indeed genuine ghosts.



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

#### 591 6.1.2 Correction of data-simulation differences

592 For the evaluation of phase space efficiency and to a lesser extend also the decay-time  
 593 efficiency we rely on simulated data as discussed in the following sections. A number  
 594 of data-driven corrections are applied to the MC samples to account for known data-  
 595 simulation differences. The MC sample is reweighted to match the three-dimensional  
 596  $p_T$ ,  $\eta$  and track multiplicity distribution observed on real data. These corrections are  
 597 derived from the calibration channel  $B_s \rightarrow D_s \pi\pi\pi$  and applied to both the signal and  
 598 calibration channel MC samples. The distributions before and after reweighting are shown  
 599 in Appendix J. We use the `PIDCorr` tool to correct the simulated PID responses based on  
 600 PID calibration samples [37].

## 601 6.2 Decay-time acceptance

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

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

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

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

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

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

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

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

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

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

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

The acceptances are determined separately for each data-taking period and each trigger category as discussed in more detail in Appendix E. The fit results are shown in Figs. 6.2 to 6.5 and the fitted parameters are summarized in Tables 6.1 to 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	$v_0$	$0.309 \pm 0.018$	$0.410 \pm 0.007$	$1.007 \pm 0.029$
0.5	$v_1$	$0.694 \pm 0.031$	$0.776 \pm 0.011$	$0.936 \pm 0.021$
1.4	$v_2$	$0.858 \pm 0.043$	$0.896 \pm 0.015$	$1.004 \pm 0.024$
2.5	$v_3$	$1.090 \pm 0.028$	$1.099 \pm 0.009$	$0.992 \pm 0.015$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.921 (interpolated)	0.913 (interpolated)	1.007 (interpolated)

**Table 6.2:** Time acceptance parameters for events in category [Run-I,L0-TIS].

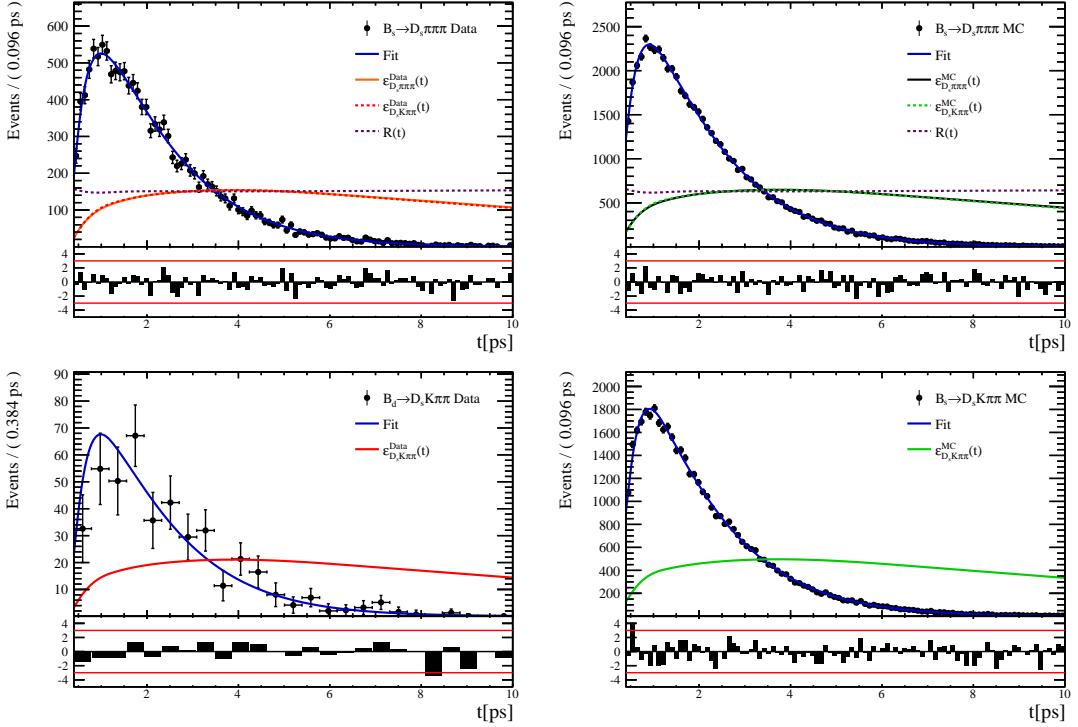
Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.158 \pm 0.014$	$0.216 \pm 0.005$	$0.986 \pm 0.040$
0.5	$v_1$	$0.422 \pm 0.029$	$0.524 \pm 0.010$	$0.965 \pm 0.029$
1.4	$v_2$	$0.802 \pm 0.047$	$0.860 \pm 0.017$	$0.982 \pm 0.029$
2.5	$v_3$	$1.099 \pm 0.034$	$1.098 \pm 0.011$	$1.002 \pm 0.019$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.913 (interpolated)	0.914 (interpolated)	0.998 (interpolated)

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

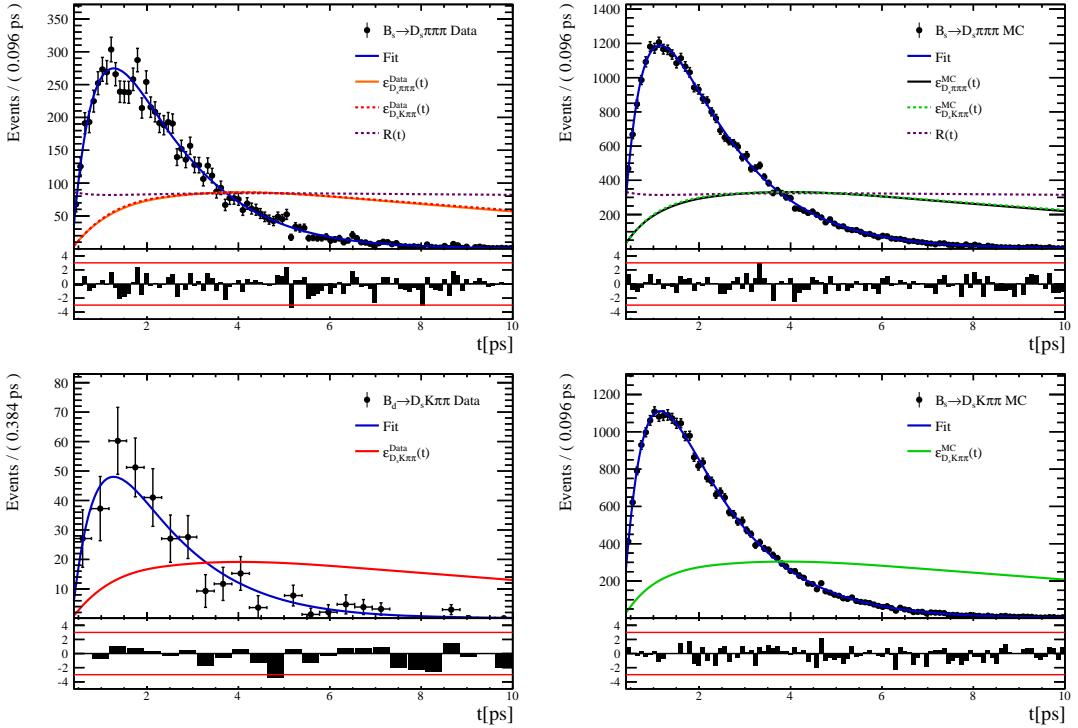
Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.285 \pm 0.009$	$0.368 \pm 0.005$	$1.023 \pm 0.020$
0.5	$v_1$	$0.663 \pm 0.017$	$0.749 \pm 0.009$	$0.911 \pm 0.016$
1.4	$v_2$	$0.856 \pm 0.025$	$0.893 \pm 0.012$	$1.016 \pm 0.019$
2.5	$v_3$	$1.060 \pm 0.017$	$1.071 \pm 0.008$	$0.996 \pm 0.013$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.948 (interpolated)	0.938 (interpolated)	1.004 (interpolated)

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

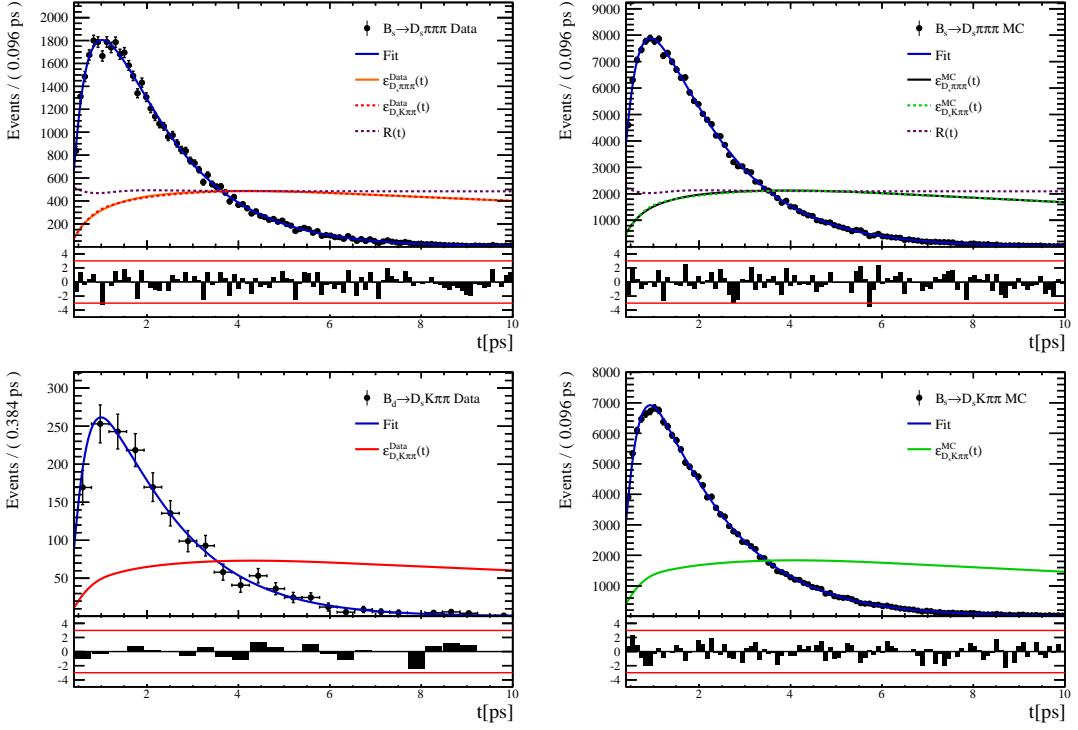
Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.117 \pm 0.008$	$0.171 \pm 0.003$	$0.965 \pm 0.034$
0.5	$v_1$	$0.422 \pm 0.019$	$0.474 \pm 0.008$	$0.952 \pm 0.024$
1.4	$v_2$	$0.733 \pm 0.027$	$0.777 \pm 0.013$	$0.973 \pm 0.025$
2.5	$v_3$	$1.071 \pm 0.020$	$1.046 \pm 0.010$	$0.989 \pm 0.015$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.938 (interpolated)	0.959 (interpolated)	1.009 (interpolated)



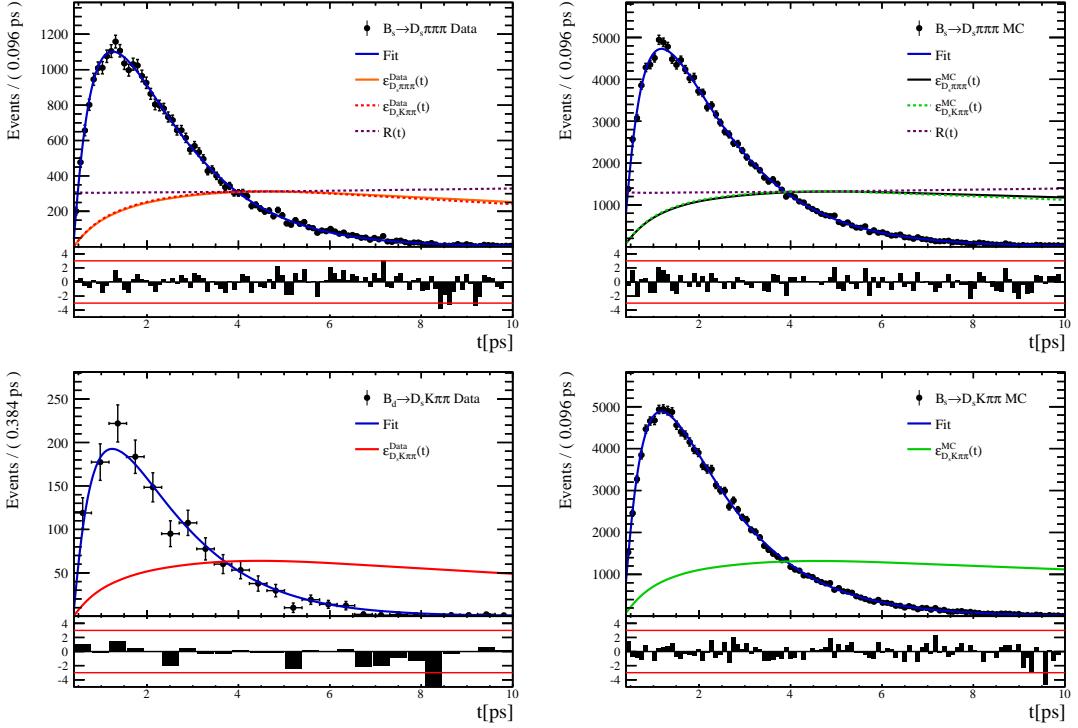
**Figure 6.2:** Decay-time fit projections for  $B_s^0 \rightarrow D_s\pi\pi\pi$  data (top-left),  $B_s^0 \rightarrow D_s\pi\pi\pi$  MC (top-right),  $B^0 \rightarrow D_s K\pi\pi$  data (bottom-left) and  $B_s^0 \rightarrow D_s K\pi\pi$  MC (bottom-right) in category [Run-I,L0-TOS]. The respective acceptance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



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



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



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

### 644 6.3 Phase space acceptance

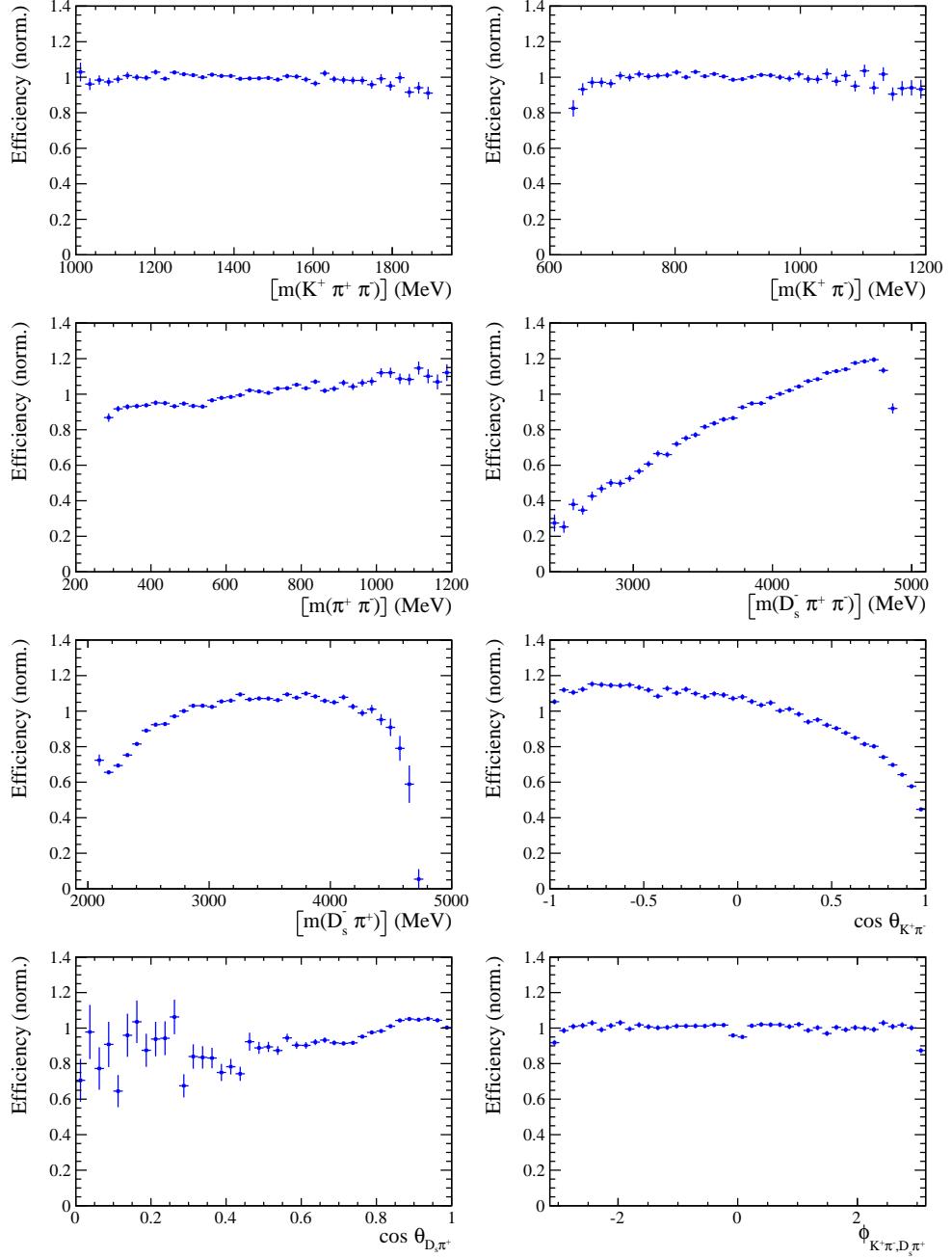
645 The signal PDF used for the full time-dependent amplitude fit can be written in terms of  
 646 the differential decay rate from Equation 2.29 as

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

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

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

655 where  $A'$  labels the amplitude model used for the generation and  $x_k$  is the  $k$ -th MC  
 656 event. As a result, the phase-space efficiency can be included in the fit without explicitly  
 657 modeling it. The size of the fully selected MC sample ( $N_{MC} = 380k$ ) is more than 70 times  
 658 larger as the data sample which results in an integral precision smaller than 0.2%. The  
 659 efficiency projections are shown in Fig. 6.6 for visualization purposes only. As discussed  
 660 in Appendix F, the phase space efficiency differs significantly among L0-trigger categories  
 661 while the differences are small between the data-taking periods and negligible between  
 662 the  $D_s$  final states. To account for this, the MC events are scaled such that the relative  
 663 proportions of the four categories [Run-I,L0-TOS], [Run-I,L0-TIS], [Run-II,L0-TOS] and  
 664 [Run-II,L0-TIS] are the same as observed on the  $B_s \rightarrow D_s K\pi\pi$  data sample.



**Figure 6.6:** Efficiency variation as a function of the phase-space variables obtained from the ratio of selected and generated MC events.

## 665 7 Flavour Tagging

666 To identify the initial flavour state of the  $B_s^0$  meson, a number of flavour tagging algorithms  
 667 are used that either determine the flavour of the non-signal b-hadron produced in the  
 668 event (opposite site, OS [40]) or use particles produced in the fragmentation of the signal  
 669 candidate  $B_s^0/\bar{B}_s^0$  (same side, SS [41]). For the same side, the algorithm searching for the  
 670 charge of an additional kaon that accompanies the fragmentation of the signal candidate is  
 671 used (SS-Kaon). For the opposite site, five different taggers are chosen: the algorithms that  
 672 use the charge of an electron or a muon from semi-leptonic B decays (OS- $e,\mu$ ), the tagger  
 673 that uses the charge of a kaon from a  $b \rightarrow c \rightarrow s$  decay chain (OS-Kaon), the algorithm  
 674 which reconstructs opposite-side charm hadrons from a number of decay channels (OS-c)  
 675 and the algorithm that determines the  $B_s^0/\bar{B}_s^0$  candidate flavour from the charge of a  
 676 secondary vertex, reconstructed from the OS b decay product (OS-VtxCharge).

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

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

684 where  $\epsilon_{tag}$  is the fraction of tagged candidates.

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

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

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

## 697 7.1 OS tagger combination

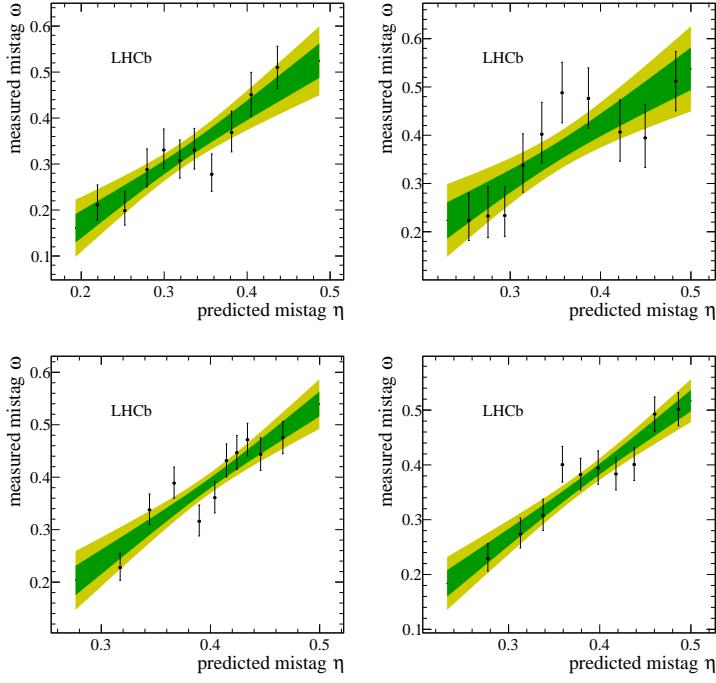
698 First, the OS electron, muon, kaon, charm and the secondary vertex charge taggers are  
 699 individually calibrated and then combined into a single OS-Combo tagger using the  
 700 **EspressoPerformanceMonitor** tool. We choose the flavour specific decay  $B_s \rightarrow D_s \pi\pi\pi$  as  
 701 calibration mode since it is very similar to the signal decay  $B_s \rightarrow D_s K\pi\pi$ . The calibration  
 702 is performed separately for Run-I and Run-II data. The OS-c tagger is not included for  
 703 Run-I data since the statistics is too low. Where available the latest Run-II tuning is  
 704 used for Run-II data, otherwise the Run-I tuning of the taggers is used. Figures 7.1 and  
 705 7.2 show the fitted calibration functions and Tables 7.1 and 7.2 list the measured tagging  
 706 performances. The tagging calibration parameters are listed in Appendix G.

**Table 7.1:** The flavour tagging performances for the used OS taggers for Run-I data.

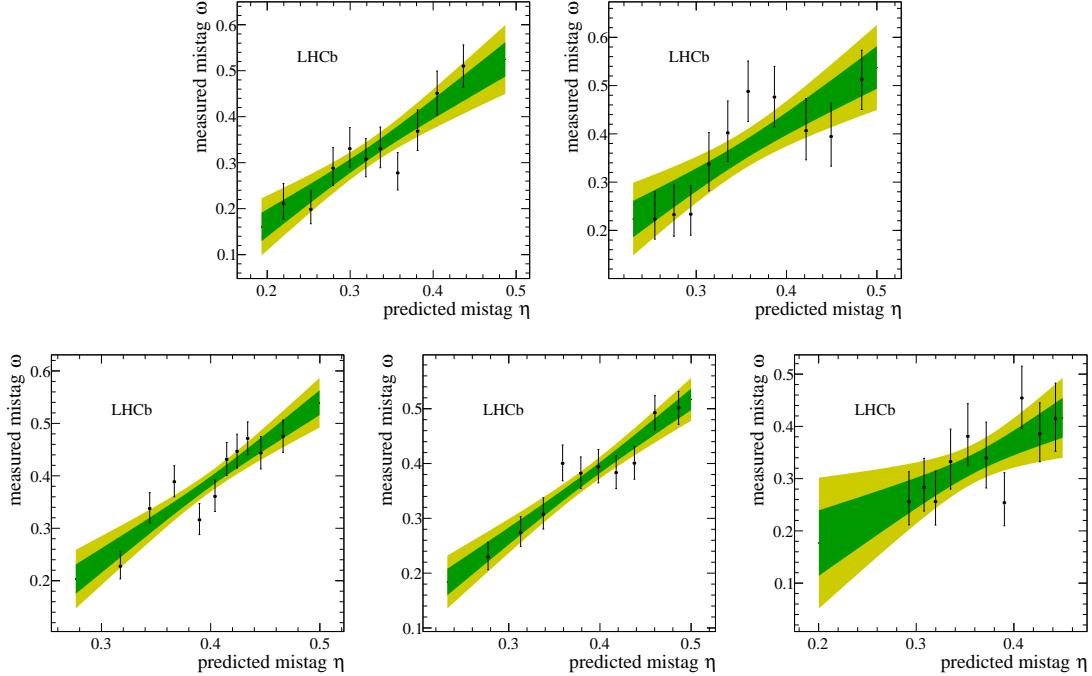
Tagger	$\epsilon$	$\omega$	$\epsilon\langle D^2 \rangle = \epsilon(1 - 2\omega)^2$
OS $\mu$	$(8.713 \pm 0.206)\%$	$(28.893 \pm 0.180(\text{stat}) \pm 2.291(\text{cal}))\%$	$(1.553 \pm 0.045(\text{stat}) \pm 0.337(\text{cal}))\%$
OS $e$	$(3.201 \pm 0.129)\%$	$(28.792 \pm 0.363(\text{stat}) \pm 3.611(\text{cal}))\%$	$(0.576 \pm 0.030(\text{stat}) \pm 0.196(\text{cal}))\%$
OS $K$	$(32.230 \pm 0.342)\%$	$(38.451 \pm 0.093(\text{stat}) \pm 1.145(\text{cal}))\%$	$(1.719 \pm 0.033(\text{stat}) \pm 0.341(\text{cal}))\%$
Vertex Charge	$(21.855 \pm 0.302)\%$	$(35.712 \pm 0.091(\text{stat}) \pm 1.474(\text{cal}))\%$	$(1.785 \pm 0.033(\text{stat}) \pm 0.368(\text{cal}))\%$

**Table 7.2:** The flavour tagging performances for the used OS taggers for Run-II data.

Tagger	$\epsilon$	$\omega$	$\epsilon\langle D^2 \rangle = \epsilon(1 - 2\omega)^2$
OS $\mu$	$(9.664 \pm 0.151)\%$	$(30.911 \pm 0.115(\text{stat}) \pm 1.369(\text{cal}))\%$	$(1.409 \pm 0.028(\text{stat}) \pm 0.202(\text{cal}))\%$
OS $e$	$(4.590 \pm 0.107)\%$	$(33.577 \pm 0.140(\text{stat}) \pm 2.007(\text{cal}))\%$	$(0.495 \pm 0.014(\text{stat}) \pm 0.121(\text{cal}))\%$
OS $K$	$(20.185 \pm 0.205)\%$	$(36.918 \pm 0.071(\text{stat}) \pm 0.969(\text{cal}))\%$	$(1.382 \pm 0.021(\text{stat}) \pm 0.205(\text{cal}))\%$
Vertex Charge	$(20.597 \pm 0.207)\%$	$(34.751 \pm 0.075(\text{stat}) \pm 0.961(\text{cal}))\%$	$(1.916 \pm 0.027(\text{stat}) \pm 0.242(\text{cal}))\%$
OS $c$	$(5.500 \pm 0.116)\%$	$(32.581 \pm 0.092(\text{stat}) \pm 1.848(\text{cal}))\%$	$(0.668 \pm 0.016(\text{stat}) \pm 0.142(\text{cal}))\%$



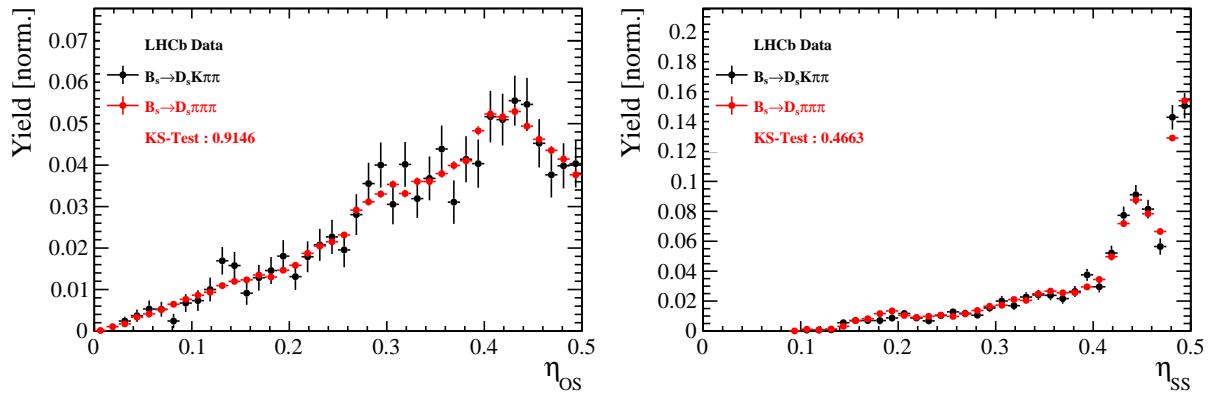
**Figure 7.1:** Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run-I. A linear fit, including the  $1\sigma$  and  $2\sigma$  error bands is overlaid for each tagger.



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

## 7.2 Tagging performance

The OS-Combo and SS-Kaon taggers are calibrated simultaneously by fitting the  $B_s \rightarrow D_s \pi\pi\pi$  decay-time distribution as discussed in Sec. 9. In this fit, the predicted mistag probabilities  $\eta_{OS}$  and  $\eta_{SS}$ , shown Fig. 7.3 for  $B_s \rightarrow D_s \pi\pi\pi$  and  $B_s \rightarrow D_s K\pi\pi$  data, are included as per-event observables, effectively giving a larger weight to the events that have a lower mistag probability. The tagger responses are combined into a single response on an event-by-event basis during the fit. Tables 7.3 and 7.4 report the tagging performances for the OS and SS combination considering three mutually exclusive categories of tagged events: OS only, SS only and both OS and SS. The tagging calibration parameters are listed in Table 9.1.



**Figure 7.3:** 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.

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

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
Only OS	$14.74 \pm 0.11$	$39.09 \pm 0.80$	$1.25 \pm 0.16$
Only SS	$35.38 \pm 0.18$	$44.26 \pm 0.62$	$1.05 \pm 0.18$
Both OS-SS	$33.04 \pm 0.30$	$37.33 \pm 0.73$	$3.41 \pm 0.33$
Combined	$83.16 \pm 0.37$	$40.59 \pm 0.70$	$5.71 \pm 0.40$

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

$B_s \rightarrow D_s \pi\pi\pi$	$\epsilon_{tag} [\%]$	$\langle \omega \rangle [\%]$	$\epsilon_{eff} [\%]$
Only OS	$11.78 \pm 0.05$	$37.01 \pm 0.51$	$1.15 \pm 0.07$
Only SS	$41.28 \pm 0.10$	$42.65 \pm 0.35$	$1.79 \pm 0.12$
Both OS-SS	$28.62 \pm 0.15$	$35.35 \pm 0.40$	$3.63 \pm 0.16$
Combined	$81.68 \pm 0.19$	$39.28 \pm 0.40$	$6.57 \pm 0.21$

## 717 8 Production and Detection Asymmetries

### 718 8.1 $B_s$ Production Asymmetry

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

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

722 where  $\sigma$  are the corresponding production cross-section. This asymmetry was measured  
723 by LHCb in  $pp$  collisions at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV by means of a time-dependent  
724 analysis of  $B_s \rightarrow D_s^- \pi^+$  decays [42]. The results in bins of  $p_T$  and  $\eta$  of the  $B_s$  meson  
725 are shown in Table 8.1. To correct for the different kinematics of  $B_s \rightarrow D_s^- \pi^+$  and  
726  $B_s^0 \rightarrow D_s K\pi\pi$  decays, the measured  $B_s$  production asymmetries  $A_p(p_T, \eta)$  are folded with  
727 the sWeighted  $p_T, \eta$  distribution of our signal channel. The resulting effective production  
728 asymmetries are:

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

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

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

729 As for Run-II data no measurement is available yet, we determine the production asym-  
730 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. [42]

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

## 731 8.2 $K^-\pi^+$ Detection Asymmetry

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

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

740 is used.

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

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

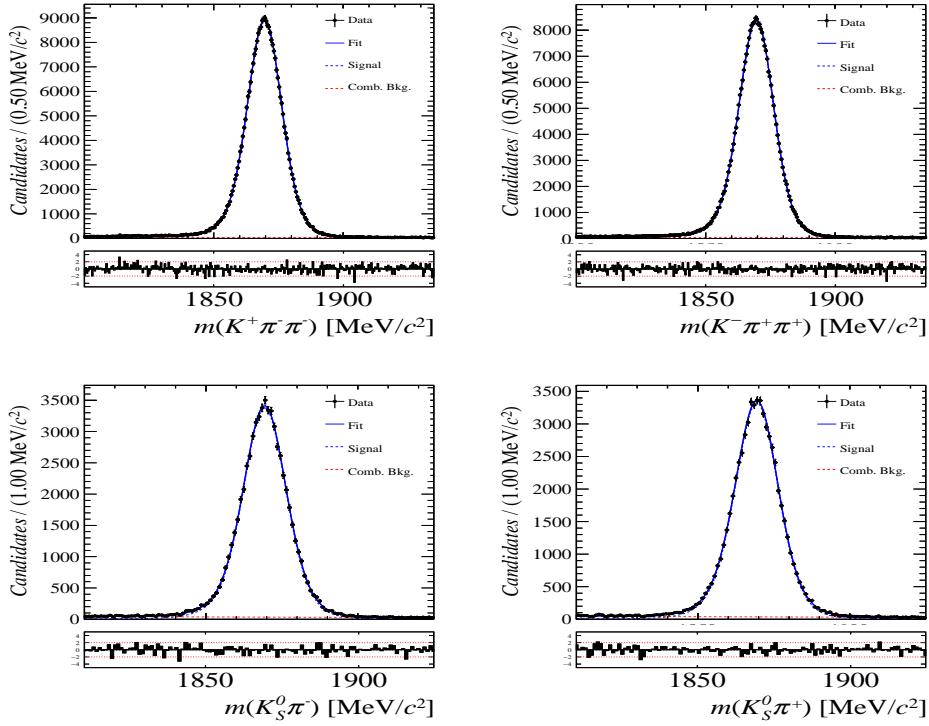
743 where possible CP violation in the  $D^+ \rightarrow K_s^0\pi^+$  mode is predicted to be smaller than  
 744  $10^{-4}$  in the Standard Model [45]. The asymmetry in the neutral kaon system,  $A(K^0)$ , has  
 745 to be taken into account as a correction.

746 We use a dedicated LHCb tool to determine  $A^{det}(K^-\pi^+)$  for all data taking periods  
 747 used in this analysis. A detailed description can be found in [44]. The tool provides  
 748 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  
 749 determine the asymmetry following Eq. 8.6. Several weighting steps are performed to  
 750 match the kinematics of the calibration samples to our signal decay sample:

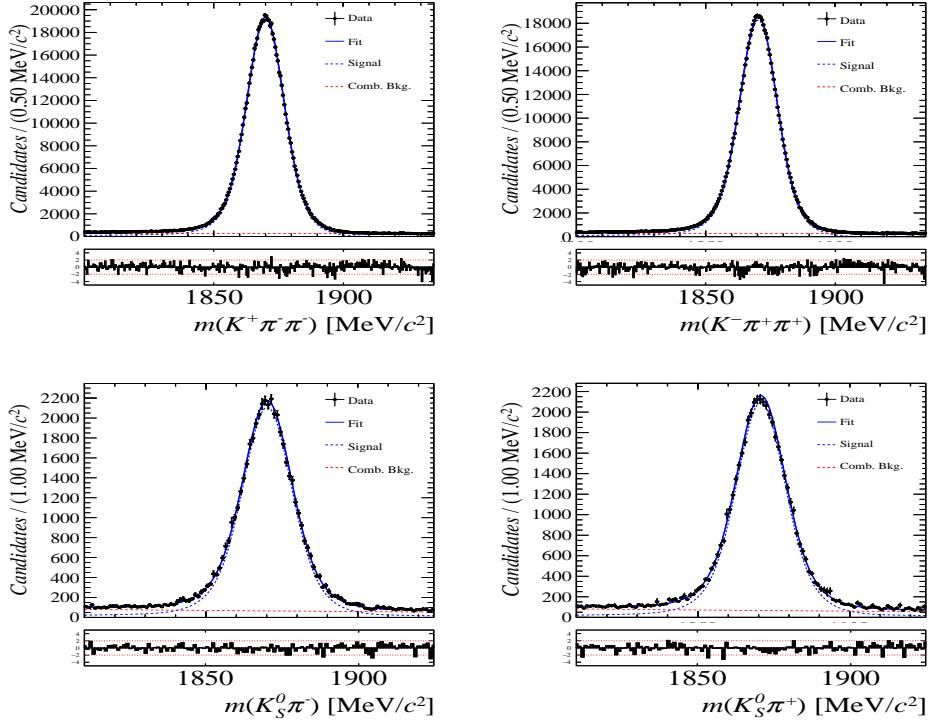
751 First, weights are assigned to the  $K^\pm$  and  $\pi^\pm$  of the  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  sample, using  
 752  $p, \eta$  of the  $K^\pm$  and  $p_T, \eta$  of the  $\pi^\pm$  from our  $B_s^0 \rightarrow D_s K\pi\pi$  signal decay. Then, weights  
 753 are assigned to the  $D^\pm$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

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

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



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



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

Data sample	$A^{det}(K^-\pi^+) + A(K^0)$ [%]
Run-I	
2011, mag. up	-2.01 $\pm$ 0.32
2011, mag. down	-0.16 $\pm$ 0.28
2011, average	-1.09 $\pm$ 0.21
2012, mag. up	-0.90 $\pm$ 0.20
2012, mag. down	-1.01 $\pm$ 0.22
2012, average	-0.96 $\pm$ 0.15
Run-II	
mag. up	-1.16 $\pm$ 0.34
mag. down	-0.65 $\pm$ 0.27
average	-0.91 $\pm$ 0.22

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

## 768 9 Decay-time fit

769 This section covers the (phase space integrated) decay-time fits to  $B_s^0 \rightarrow D_s h\pi\pi$  data. We  
 770 use the **sFit** technique [46] to statistically subtract the background, leaving only the signal  
 771 PDF to describe the decay-time. The **sWeights** are calculated based on the fit to the  
 772 reconstructed  $B_s$  mass distribution described in Sec. 4. The signal PDF is conditional on  
 773 the tagging decisions  $q_i$ , the mistag estimates  $\eta_i$  ( $i = \text{OS,SS}$ ) and the decay-time error  $\delta t$ :

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

774 where  $p(t|q_{\text{OS}}, \eta_{\text{OS}}, q_{\text{SS}}, \eta_{\text{SS}})$  is given by Equation 2.7 taking the tagging dilution into  
 775 account. The decay-time acceptance  $\epsilon(t)$  (Sec. 6) and the Gaussian time-resolution  
 776 function  $\mathcal{R}(t - t', \delta t)$  (Sec. 5) are fixed to the values obtained by the dedicated studies.  
 777 We fix the values of  $\Gamma_s$  and  $\Delta\Gamma_s$  to the latest HFAG results [38].

778 The unbinned maximum likelihood fits are performed simultaneously in six categories:  
 779 [Run-I,L0-TOS], [Run-I,L0-TIS], [Year-15/16,L0-TOS], [Year-15/16,L0-TIS], [Year-17,L0-  
 780 TOS] and [Year-17,L0-TIS].

### 781 9.1 Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data

782 Since the decay  $B_s^0 \rightarrow D_s\pi\pi\pi$  is flavour specific, the  $CP$  coefficients can be fixed to  $C = 1$   
 783 and  $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$ . The fit determines the calibration parameters for the  
 784 OS-Combo and SS-Kaon taggers, the  $B_s^0$  production asymmetry for Run-II data as well as  
 785 the mixing frequency  $\Delta m_s$ . Table 9.1 summarizes the fitted parameters. The **sWeighted**  
 786 decay-time distribution and the time-dependent asymmetry  $A_{mix}$  between mixed and  
 787 unmixed  $B_s^0$  candidates are shown in Fig. 9.1 along with the fit projections.

### 788 9.2 Fit to $B_s^0 \rightarrow D_s K\pi\pi$ data

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

793

794

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

796

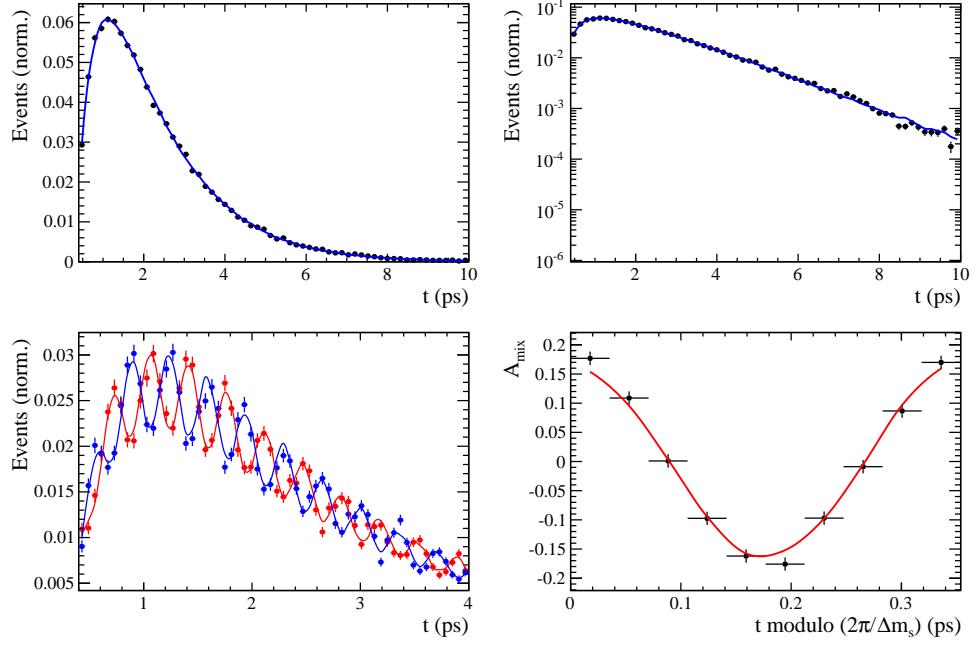
Currently the mixing frequency is fixed to the HFAG value. We intend to update  
 797 the fit after unblinding our result from the  $B_s^0 \rightarrow D_s\pi\pi\pi$  fit since our precision is  
 significantly higher.

**Table 9.1:** Parameters determined from a fit to the  $B_s \rightarrow D_s \pi\pi\pi$  decay-time distribution. The uncertainties are statistical and systematic, respectively.

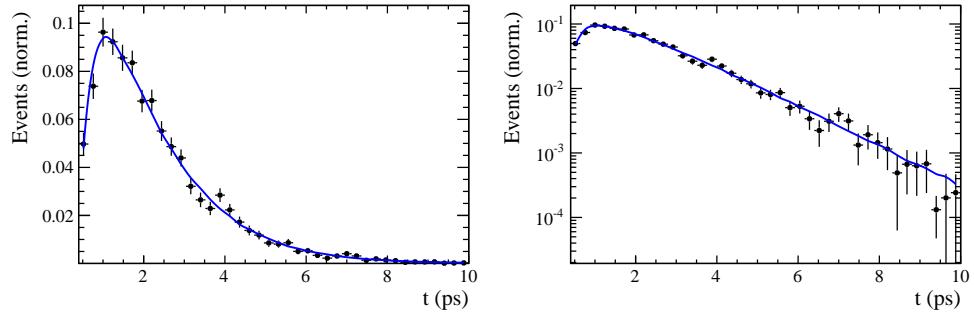
Fit Parameter	Run-I	Run-II
$p_0^{OS}$	$0.398 \pm 0.010 \pm 0.010$	$0.372 \pm 0.005 \pm 0.005$
$p_1^{OS}$	$0.895 \pm 0.085 \pm 0.090$	$0.788 \pm 0.043 \pm 0.030$
$\Delta p_0^{OS}$	$0.030 \pm 0.011 \pm 0.002$	$0.008 \pm 0.006 \pm 0.001$
$\Delta p_1^{OS}$	$0.011 \pm 0.095 \pm 0.017$	$0.067 \pm 0.052 \pm 0.002$
$\epsilon_{tag}^{OS} [\%]$	$47.775 \pm 0.365 \pm 0.067$	$40.399 \pm 0.182 \pm 0.029$
$\Delta \epsilon_{tag}^{OS} [\%]$	$0.016 \pm 1.353 \pm 0.097$	$0.316 \pm 0.618 \pm 0.046$
$p_0^{SS}$	$0.444 \pm 0.008 \pm 0.005$	$0.428 \pm 0.004 \pm 0.002$
$p_1^{SS}$	$0.949 \pm 0.111 \pm 0.067$	$0.787 \pm 0.039 \pm 0.025$
$\Delta p_0^{SS}$	$-0.019 \pm 0.009 \pm 0.001$	$-0.017 \pm 0.004 \pm 0.000$
$\Delta p_1^{SS}$	$0.064 \pm 0.124 \pm 0.017$	$0.028 \pm 0.048 \pm 0.006$
$\epsilon_{tag}^{SS} [\%]$	$68.426 \pm 0.340 \pm 0.013$	$69.903 \pm 0.170 \pm 0.007$
$\Delta \epsilon_{tag}^{SS} [\%]$	$-0.046 \pm 1.242 \pm 0.082$	$-0.319 \pm 0.575 \pm 0.062$
$A_P [\%]$	$-0.045$ (fixed)	
$\Delta m_s [\text{ps}^{-1}]$	$\text{xx.xx} \pm 0.0084 \pm 0.0058$	

**Table 9.2:**  $CP$  coefficients determined from a fit to the  $B_s \rightarrow D_s K\pi\pi$  decay-time distribution. The uncertainties are statistical and systematic, respectively.

Fit Parameter	Value
$C$	$\text{x.xx} \pm 0.12 \pm 0.02$
$D$	$\text{x.xx} \pm 0.32 \pm 0.08$
$\bar{D}$	$\text{x.xx} \pm 0.30 \pm 0.08$
$S$	$\text{x.xx} \pm 0.17 \pm 0.04$
$\bar{S}$	$\text{x.xx} \pm 0.17 \pm 0.04$



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



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

## 797 10 Time-dependent amplitude fit

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

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

799 where  $p(\mathbf{x}, t | q_{os}, \eta_{os}, q_{ss}, \eta_{ss})$  is given the differential decay rate in Equation 2.29 taking  
800 the tagging dilution into account. The phase space efficiency  $\epsilon(\mathbf{x})$  is only included in the  
801 normalization of  $\mathcal{P}(\mathbf{x}, t | \delta t, q_{os}, \eta_{os}, q_{ss}, \eta_{ss})$  as discussed in Sec 6.3. The model selection  
802 of the amplitude components is described in the following Section. The remaining fitting  
803 strategy is exactly the same as for the decay-time fits, see Sec. 9.

### 804 10.1 Signal Model Construction

805 The light meson spectrum comprises multiple resonances which are expected to contribute  
806 to  $B_s \rightarrow D_s K\pi\pi$  decays as intermediate states. Apart from clear contributions coming  
807 from resonances such as  $K_1(1270)$ ,  $K_1(1400)$ ,  $\rho(770)$  and  $K^*(892)^0$ , the remaining structure  
808 is impossible to infer due to the cornucopia of broad, overlapping and interfering resonances  
809 within the phase space boundary. We follow the LASSO [47, 48] approach to limit the  
810 model complexity in two steps.

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

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

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

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

814 This significantly simplifies the fitting procedure and allows us to include the whole pool  
815 of considered intermediate state amplitudes  $A_i$  which can be found in Appendix I. The  
816 LASSO penalty term added to the likelihood function

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

817 shrinks the amplitude coefficients towards zero. The amount of shrinkage is controlled by  
818 the parameter  $\lambda$ , to be tuned on data. Higher values for  $\lambda$  encourage sparse models, *i.e.*  
819 models with only a few non-zero amplitude coefficients. The optimal value for  $\lambda$  is found  
820 by minimizing the Bayesian information criteria [49] (BIC),

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

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

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

and are a measure of the relative strength between the different transitions. Figure 10.1(left) shows the distribution of BIC values obtained by scanning over  $\lambda$  where we choose the decay fraction threshold to be 0.5%. At the optimal value of  $\lambda = 50$ , the set of amplitudes with a decay fraction above the threshold are considered further for step two of the model selection. The selected amplitudes and their fractions are summarized in Table 10.1. The fit projections are shown in Fig. 10.2. The set of selected amplitudes is stable for thresholds between 0.1% and 1%.. Other choices result in marginally different models containing one component more or less. These are included in the set of alternative models used for the systematic studies presented in Sec. 11.10.

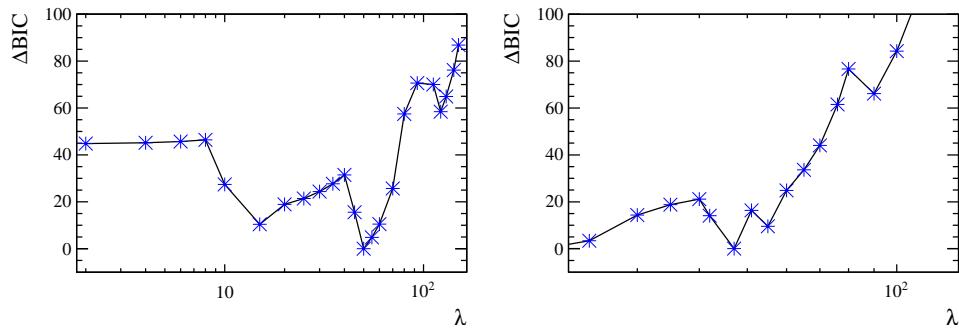
In Stage 2, the LASSO procedure is again performed by fitting the full time-dependent amplitude PDF. The components selected by Stage 1 are included for both  $b \rightarrow c$  and  $b \rightarrow u$  transitions and the likelihood is extended as follows:

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

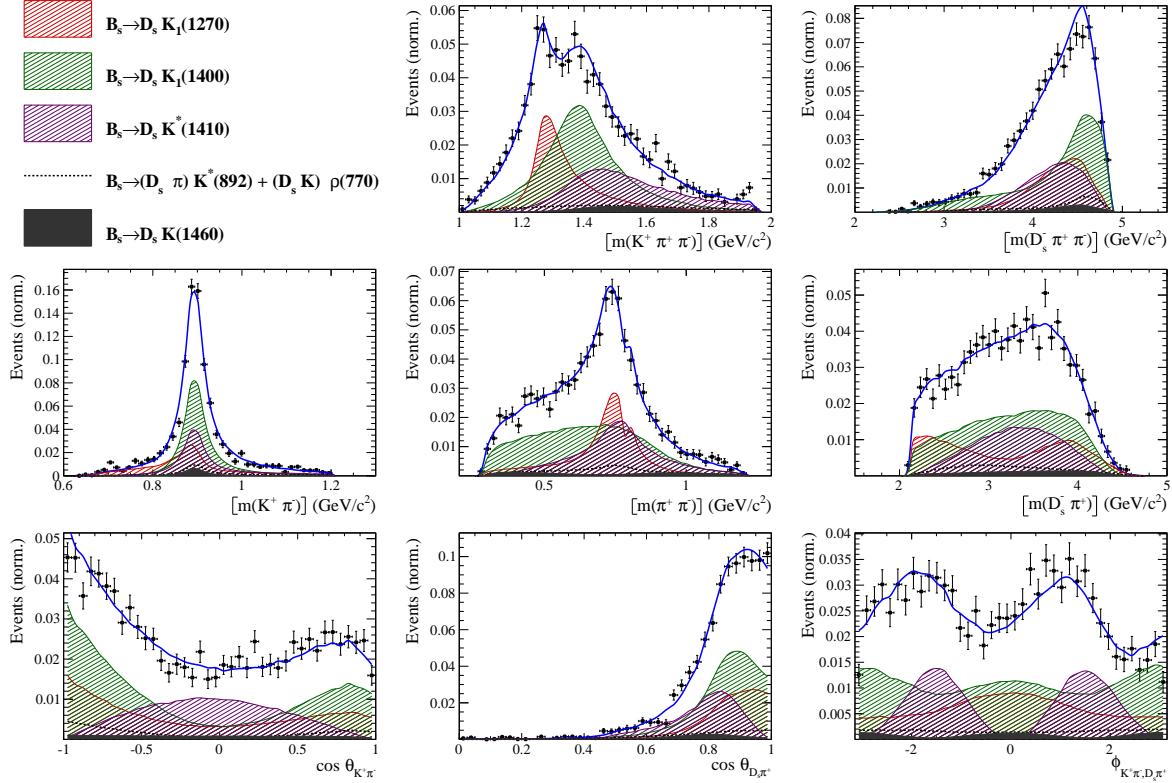
Figure 10.1(right) shows a plot of the complexity factor  $\lambda$ , against the resulting BIC values. The final set of  $b \rightarrow c$  and  $b \rightarrow u$  amplitudes is selected using the optimal value of  $\lambda = 28$ , and is henceforth called the LASSO model.

**Table 10.1:** Fit fractions of the amplitudes selected by Stage 1 of the model selection procedure.

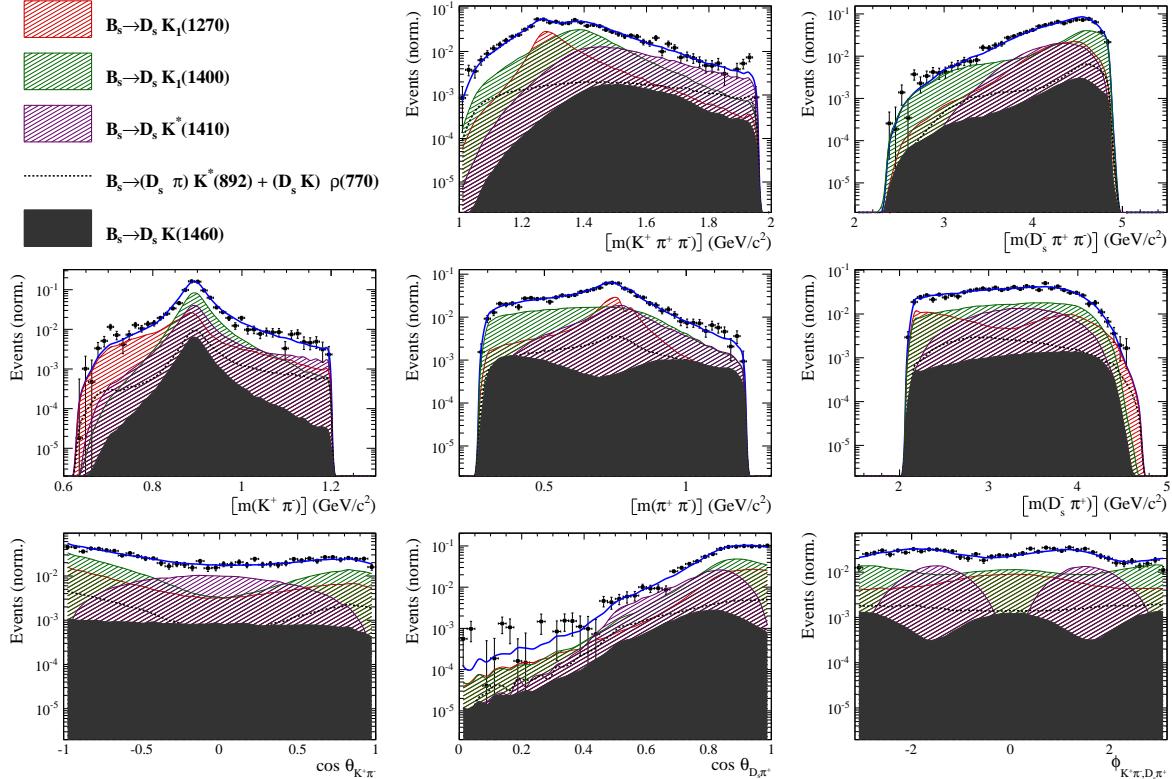
Decay channel	Fraction [%]
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	$8.56 \pm 1.43$
$B_s \rightarrow K(1)(1400)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	$43.72 \pm 2.80$
$B_s \rightarrow K(1460)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	$3.25 \pm 0.69$
$B_s \rightarrow K^*(1410)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	$15.33 \pm 1.13$
$B_s \rightarrow (D_s^- \pi^+)_P K^*(892)^0 (\rightarrow K^+ \pi^-)$	$4.63 \pm 0.69$
$B_s \rightarrow K^*(1410)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	$5.58 \pm 0.62$
$B_s \rightarrow (D_s^- K^+)_P \rho(770)^0 (\rightarrow \pi^+ \pi^-)$	$1.49 \pm 0.40$
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K(0)^*(1430)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	$4.72 \pm 0.54$
$B_s \rightarrow K(1)(1270)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	$14.20 \pm 1.56$
Sum	$101.47 \pm 3.86$



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



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



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

## 838 10.2 Results

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

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

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

846 Figure 10.4 shows the distributions of selected phase space observables, which demon-  
 847 strate reasonable agreement between data and the fit model. We also project into the  
 848 transversity basis to demonstrate good description of the overall angular structure (see  
 849 10.4 bottom row). The acoplanarity angle  $\chi$ , is the angle between the two decay planes  
 850 formed by the  $K^+\pi^-$  system and the  $D_s^-\pi^+$  system in the  $B_s$  rest frame; boosting into the  
 851 rest frames of the two-body systems defining these decay planes, the two helicity variables  
 852 are defined as the cosine of the angle,  $\theta$ , of the  $K^+$  or  $D_s^-$  momentum with the  $B_s$  flight  
 853 direction.

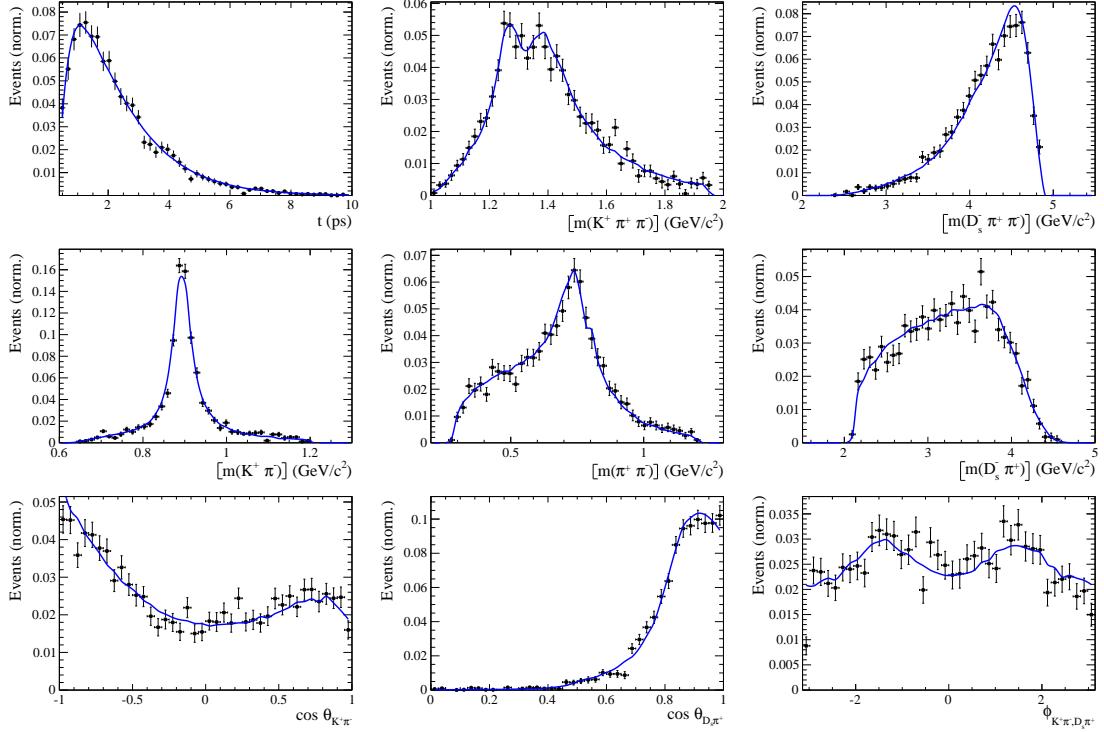
854 In order to quantify the quality of the fit in the five-dimensional phase space, a  $\chi^2$   
 855 value is determined by binning the data;

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

856 where  $N_b$  is the number of data events in a given bin,  $N_b^{\text{exp}}$  is the event count predicted  
 857 by the fitted PDF and  $N_{\text{bins}}$  is the number of bins. An adaptive binning is used to ensure  
 858 sufficient statistics in each bin for a robust  $\chi^2$  calculation [50]. At least 25 events per  
 859 bin are required. The number of degrees of freedom  $\nu$ , in an unbinned fit is bounded by  
 860  $N_{\text{bins}} - 1$  and  $(N_{\text{bins}} - 1) - N_{\text{par}}$ , where  $N_{\text{par}}$  is the number of free fit parameters. We use  
 861 the  $\chi^2$  value divided by  $\nu = (N_{\text{bins}} - 1) - N_{\text{par}}$  as a conservative estimate. For the LASSO  
 862 model, this amounts to  $\chi^2/\nu = 1.40$  with  $\nu = 137$ , indicating a decent fit quality.

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

Decay Channel	$A_{b \rightarrow c}$		$A_{b \rightarrow u}$	
	$ a_i $	$\arg(a_i)[^\circ]$	$ a_i $	$\arg(a_i)[^\circ]$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	1.0	0.0	1.0	0.0
$K_1(1270) \rightarrow K^*(892) \pi$	$0.71 \pm 0.10 \pm 0.08$	$49.5 \pm 8.9 \pm 5.2$		
$K_1(1270) \rightarrow K_0^*(1430) \pi$	$0.52 \pm 0.05 \pm 0.07$	$128.9 \pm 4.1 \pm 24.1$		
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$1.98 \pm 0.12 \pm 0.19$	$11.5 \pm 7.4 \pm 5.7$	$0.73 \pm 0.17 \pm 0.16$	$-65.8 \pm 14.7 \pm 13.2$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$1.14 \pm 0.05 \pm 0.05$	$55.3 \pm 5.4 \pm 5.4$		
$K^*(1410) \rightarrow K \rho(770)$	$0.63 \pm 0.04 \pm 0.03$	$-163.9 \pm 5.3 \pm 2.7$		
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$			$0.87 \pm 0.09 \pm 0.08$	$-97.3 \pm 8.9 \pm 9.9$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$0.74 \pm 0.11 \pm 0.13$	$-17.9 \pm 8.4 \pm 11.6$	$1.13 \pm 0.14 \pm 0.14$	$-17.7 \pm 8.2 \pm 14.5$
$B_s \rightarrow (D_s K)_P \rho(770)$			$0.54 \pm 0.07 \pm 0.08$	$33.3 \pm 9.1 \pm 10.4$
Fit parameter	Value			
$m_{K_1(1400)} [\text{MeV}]$	$1398 \pm 9 \pm 4 \pm 6$			
$\Gamma_{K_1(1400)} [\text{MeV}]$	$204 \pm 14 \pm 9 \pm 9$			
$m_{K^*(1410)} [\text{MeV}]$	$1432 \pm 12 \pm 15 \pm 8$			
$\Gamma_{K^*(1410)} [\text{MeV}]$	$344 \pm 25 \pm 35 \pm 18$			
$r$	$xx.xx \pm 0.03 \pm 0.03 \pm 0.02$			
$\delta [^\circ]$	$xx.xx \pm 15.0 \pm 6.0 \pm 7.8$			
$\gamma - 2\beta_s [^\circ]$	$xx.xx \pm 16.4 \pm 6.1 \pm 6.5$			



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

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

Decay Channel	$F_{b \rightarrow c} [\%]$	$F_{b \rightarrow u} [\%]$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$6.3 \pm 1.6$	$14.9 \pm 4.5$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$12.3 \pm 1.4$	$29.1 \pm 6.1$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$3.4 \pm 0.5$	$8.0 \pm 2.1$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$48.3 \pm 4.5$	$17.2 \pm 8.6$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$15.5 \pm 1.0$	
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$6.7 \pm 0.6$	
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$		$21.0 \pm 4.6$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$6.8 \pm 1.5$	$36.0 \pm 8.0$
$B_s \rightarrow (D_s K)_P \rho(770)$		$9.7 \pm 4.0$
Sum	$99.3 \pm 4.7$	$135.9 \pm 12.9$

## 863 11 Systematic uncertainties

864 The systematic uncertainties on the measured observables are summarized in Table 11.3 for  
865 the decay-time fit to  $B_s \rightarrow D_s \pi\pi\pi$ , in Table 11.4 for the decay-time fit to  $B_s \rightarrow D_s K\pi\pi$   
866 and in Table 11.5 for the full time-dependent amplitude fit to  $B_s \rightarrow D_s K\pi\pi$  decays. A  
867 description of each systematic effect is given in the following subsections starting with the  
868 ones common to all fits. Afterwards, systematic effect specific to the amplitude description  
869 are discussed.

### 870 11.1 Fit bias

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

### 875 11.2 Background subtraction

876 The statistical subtraction of the residual background [46], left after the full selection,  
877 relies on the correct description of the invariant  $B_s^0$  mass distribution. Since the choice of  
878 signal and background models is not unique, alternative parameterizations are tested:

- 879 • The Johnson's SU function which is used as nominal signal model is replaced by the  
880 sum of two Crystal Ball functions [51].
- 881 • For the combinatorial background, the nominal second order polynomial is replaced  
882 by an exponential function.
- 883 • For the description of the partially reconstructed background, a combination of the  
884 RooHILLdini and RooHORNsdini model [52] is used instead of the nominal model of  
885 three bifurcated gaussians.
- 886 • For the shape of the mis-ID background, the nominal approach is to use a simulated  
887 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  
888 hypothesis of the  $\pi^+$  with the higher misidentification probability (see Sec. 4).  
889 Two alternative approaches are considered: we flip the mass hypothesis of the  $\pi^+$   
890 candidate with the lower probability of being misidentified; we randomly flip the  
891 mass hypothesis of a  $\pi^+$  candidate. To account for the systematic of the  $\pi \rightarrow K$   
892 fake rate, we vary the default PIDK cut ( $\text{PIDK} > 10$ ) by  $\pm 2$  when calculating the  
893 misID weights from `PIDCalib`.

894 In total 10 (7) different combinations of the modifications discussed above are tested  
895 for the fit to the  $D_s K\pi\pi$  ( $D_s \pi\pi\pi$ ) mass distribution. For each case, new signal `sWeights`  
896 are calculated and the `sFits` to data are repeated. The sample variance of the obtained  
897 differences to the nominal fit value are used as systematic uncertainty due to the background  
898 subtraction.

### 899 11.3 Decay-time acceptance

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

908 To improve the coverage of the multi-dimensional parameter space, a Cholesky decom-  
 909 position [53] is used to generate a set of uncorrelated vectors from the covariance matrix  
 910  $\text{cov}(\lambda_i, \lambda_j)$ , where the vector  $\lambda$  includes the parameters  $\Gamma_s$ ,  $\Delta\Gamma_s$  and the  $N = 4$  spline  
 911 coefficients for each category of the simultaneous fit. The correlations between  $\Gamma_s$  ( $\Delta\Gamma_s$ )  
 912 and the spline coefficients are measured by rerunning the acceptance fits described in  
 913 Sec. 6.2 with the values of  $\Gamma_s$  ( $\Delta\Gamma_s$ ) varied by  $\pm 1\sigma$  and measuring the shift in the spline  
 914 coefficients as a fraction of their uncertainty. For the correlation between  $\Gamma_s$  and  $\Delta\Gamma_s$  we  
 915 use the HFAG value [38].

### 916 11.4 Decay-time resolution and tagging

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

- 920 • A double Gaussian is fit to the decay-time distributions of fake  $B_s^0$  candidates, but  
 921 only the width of the core Gaussian is considered to represent the time resolution in  
 922 the respective bin. Therefore the resolution is slightly underestimated in this case.
- 923 • A single Gaussian is fit to the decay-time distributions of fake  $B_s^0$  candidates in a  
 924 wide range of  $[-3\sigma_t : 1.5\sigma_t]$ . Due to the tails of the distribution, which broaden the  
 925 width of the Gaussian function, this method slightly overestimates the decay-time  
 926 resolution.

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

$$\sigma_{\text{eff},\text{Data},16}^{\text{core-Gauss}}(\sigma_t) = (6.8 \pm 1.9) \text{ fs} + (0.797 \pm 0.050) \sigma_t \quad (11.1)$$

$$\sigma_{\text{effData},16}^{\text{single-Gauss}}(\sigma_t) = (10.1 \pm 1.5) \text{ fs} + (0.950 \pm 0.039) \sigma_t \quad (11.2)$$

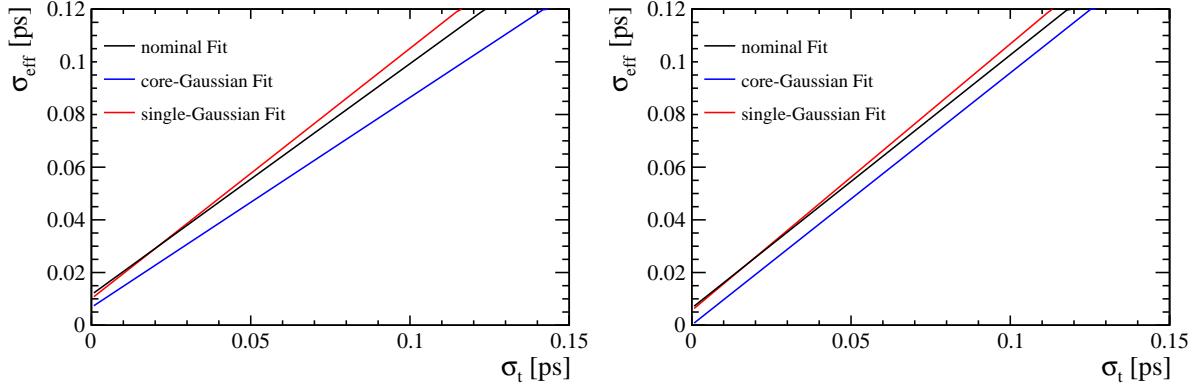
$$\sigma_{\text{eff},\text{Data},17}^{\text{core-Gauss}}(\sigma_t) = (0.1 \pm 1.5) \text{ fs} + (0.957 \pm 0.037) \sigma_t \quad (11.3)$$

$$\sigma_{\text{effData},17}^{\text{single-Gauss}}(\sigma_t) = (5.6 \pm 1.2) \text{ fs} + (1.012 \pm 0.031) \sigma_t \quad (11.4)$$

931 They are compared to the nominal result in Fig. 11.1.

932 Due to the high correlation between the decay-time resolution and the tagging calibra-  
 933 tion, their systematic uncertainty has to be studied simultaneously. First, the decay-time  
 934 fits to  $B_s \rightarrow D_s \pi\pi\pi$  data are repeated using the alternative decay-time error scaling  
 935 functions. New tagging calibration parameters are obtained which are then used (together  
 936 with the respective decay-time error scaling function) in the fits to  $B_s \rightarrow D_s K\pi\pi$  data

937 to define the Gaussian-constraints as discussed in Sec. 9. For the width of the Gaussians  
 938 only the statistical error of the tagging calibration parameters are used since systematic  
 939 uncertainties (except the systematic arising from the decay-time resolution which is already  
 940 included by the procedure described above) are found to be negligible, see Table 11.3.  
 941 Finally, we take the biggest change in fit central value as the systematic for each parameter  
 942 of the  $B_s \rightarrow D_s K\pi\pi$  fits.



**Figure 11.1:** The measured resolution scaling function of the per-event decay time error estimate  $\sigma_t$  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. Data taken in (left) 2016 and (right) 2017.

## 943 11.5 Production, detection asymmetries and mixing frequency

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

## 948 11.6 Multiple candidates

949 The fraction of events with multiple candidates has been found to be very small, it is  
 950 1.6% for  $D_s K\pi\pi$  and 1.5% for  $D_s \pi\pi\pi$ . Thus the nominal result is obtained keeping all  
 951 candidates, while a systematic uncertainty is assigned by repeating the fit randomly  
 952 keeping only one candidate when multiple ones are founds. No shifts in the fit central  
 953 values are observed.

## 954 11.7 Length and momentum scales

955 The uncertainty on the LHCb length scale is estimated to be at most 0.020% [54], which  
 956 translates directly in an uncertainty on  $\Delta m_s$  of 0.020% with other parameters being  
 957 unaffected. The momentum scale uncertainty is at most 0.022%.

## 958 11.8 Phase space acceptance

959 For the phase space acceptance we rely on simulated data. The integration error due to  
960 the limited size of the MC sample used to normalize the signal PDF is below 0.2% and  
961 thus negligible small.

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

- 966 • No BDT cut is applied
- 967 • A tighter BDT requirement is used ( $\text{BDTG} > 0.6$ )
- 968 • No reweighting is applied
- 969 • Instead of the PID responses obtained from the `PIDCorr` tool, we use the `PIDGen`  
970 tool to resample the PID variables [37]
- 971 • The raw MC PID variables are used
- 972 • Candidates with `BKGCAT`= 60 are removed

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

## 975 11.9 Resonance description

976 The following alternative line shape parameterizations are considered as part of the  
977 systematic studies:

- 978 • The Lass description for the  $K\pi$   $S$ -wave is replaced by a more general model  
979 (Glass [55, 56])
- 980 • The Gounaris-Sakurai description for the  $\rho(770)$  is replaced by a relativistic Breit-  
981 Wigner propagator (Equation 2.16)
- 982 • The  $\omega$  contribution to the decay channel  $K_1(1270) \rightarrow K \rho(770)/\omega$  is set to zero
- 983 • For the decay channel  $K^*(1410) \rightarrow K \rho(770)$ , we include  $\rho(770) - \omega$  mixing with a  
984 relative magnitude and phase determined from data
- 985 • Instead of taking the energy-dependent widths of the three-body resonances from  
986 Refs. [9, 21], we derive them from Equation 2.17 taking only the dominant  $K\pi\pi$   
987 decay mode into account

988 The data fits are repeated for each alternative model and the RMS of the central values  
989 are taken as systematic uncertainties.

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

## 995 11.10 Alternative amplitude models

996 We tested several modifications of the LASSO model to assign an additional model  
997 uncertainty to the measured observables  $r, \delta$  and  $\gamma - 2\beta_s$  as well as to the measured  
998 masses and widths of the  $K_1(1400)$  and  $K^*(1410)$  resonances. The amplitude coefficients  
999 are by definition parameters of a given model which is why we do not evaluate a model  
1000 uncertainty for them. The fit results of the following alternative models (Alt. 1 - Alt. 12)  
1001 are summarized in Tables 11.1 and 11.2.

- 1002 • All amplitudes selected by Stage 1 of the model selection are included for both  $b \rightarrow c$  and  
1003  $b \rightarrow u$  transitions (Alt. 1)
- 1004 • The decay channel  $K_1(1270)[D] \rightarrow K^*(892)\pi$ , where the  $K^*(892)\pi$  system is in relative a  
1005 D-wave state, is added (Alt. 2)
- 1006 • The decay channel  $K_1(1400) \rightarrow K\rho(770)$  is added (Alt. 12)
- 1007 • The decay channels  $K(1460) \rightarrow K\rho(770)$  and  $K(1460) \rightarrow K\sigma$  are added (Alt. 3)
- 1008 • The decay channels  $K^*(1680) \rightarrow K\rho(770)$  and  $K^*(1680) \rightarrow K^*(892)\pi$  are added (Alt. 9)
- 1009 • The decay channel  $K_2(1770) \rightarrow K^*(892)\pi$  is added (Alt. 10)
- 1010 • The amplitude  $B_s \rightarrow (D_s K)_P \rho(770)$  is replaced by  $B_s \rightarrow (D_s K)_S \rho(770)$  (Alt. 11)
- 1011 • Higher orbital angular momentum states are added for the amplitudes:  $B_s \rightarrow (D_s \pi)_P K^*(892)$  and  $B_s \rightarrow (D_s K)_P \rho(770)$  (Alt. 4 and Alt. 5)
- 1012 • The amplitudes  $B_s \rightarrow (D_s K)\sigma$ ,  $B_s \rightarrow (D_s K)f_0(980)$ ,  $B_s \rightarrow (D_s K)f_2(1270)$  and  
1013  $B_s \rightarrow (D_s K)f_0(1370)$  are added in different combinations and angular momentum  
1014 configurations (Alt. 6, Alt. 7 and Alt. 8 )
- 1015 • The decay channels  $K_2^*(1430) \rightarrow K\rho(770)$  and  $K_2^*(1430) \rightarrow K^*(892)\pi$  are added (not  
1016 used)
- 1017 • The amplitudes  $B_s \rightarrow (D_s \pi)_P K_0^*(1430)$  and  $B_s \rightarrow (D_s K)_S K_2^*(1430)$  are added (not  
1018 used)
- 1019 • A flat non-resonant  $B_s \rightarrow D_s K \pi\pi$  component is added (not used)

1021 In total 15 different sets of amplitudes are fitted. In some cases, the fit fractions of  
1022 additionally added amplitudes turn out to be exactly zero. These model are effectively  
1023 not distinguishable from the baseline LASSO model and are not considered further. From  
1024 the remaining 12 models, we compute the sample variance for each observable and take it  
1025 as model uncertainty.

**Table 11.1:** Fit fractions in percent for the baseline and several alternative amplitude models (Alt. 1 - Alt. 6). Resonance parameters and the physical observables  $r, \kappa, \delta, \gamma - 2\beta_s$  are also given. The values of the physical observables are given relative to the baseline result. The uncertainties are statistical only.

	Baseline	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5	Alt.6
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$6.2 \pm 2.1$	$7.0 \pm 0.0$	$7.2 \pm 0.0$	$5.2 \pm 1.8$	$5.5 \pm 0.0$	$6.1 \pm 1.9$	$6.3 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)$				$1.3 \pm 0.0$			
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$12.3 \pm 1.2$	$10.7 \pm 0.0$	$10.2 \pm 0.0$	$9.4 \pm 1.6$	$14.1 \pm 0.0$	$11.2 \pm 1.7$	$13.1 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))$							
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$3.3 \pm 0.5$	$2.7 \pm 0.0$	$2.4 \pm 0.0$	$2.6 \pm 0.4$	$3.6 \pm 0.0$	$2.8 \pm 0.6$	$3.7 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$47.9 \pm 3.0$	$53.8 \pm 0.0$	$54.3 \pm 0.0$	$58.8 \pm 6.8$	$47.2 \pm 0.0$	$53.7 \pm 9.4$	$43.3 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$15.9 \pm 0.9$	$14.5 \pm 0.0$	$15.2 \pm 0.0$	$15.7 \pm 1.1$	$17.5 \pm 0.0$	$15.7 \pm 1.0$	$15.8 \pm 0.0$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$6.5 \pm 0.6$	$5.7 \pm 0.0$	$6.3 \pm 0.0$	$6.3 \pm 0.8$	$6.6 \pm 0.0$	$6.5 \pm 0.8$	$6.9 \pm 0.0$
$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K \rho(770))$				$0.8 \pm 0.0$			
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)$							
$B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))$							
$b \rightarrow c$							
$B_s \rightarrow D_s (K_2(1770) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow (D_s \pi)_S K^*(892)$							
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$6.7 \pm 1.8$	$9.2 \pm 0.0$	$7.8 \pm 0.0$	$10.3 \pm 2.1$	$13.9 \pm 0.0$	$6.8 \pm 1.9$	$5.9 \pm 0.0$
$B_s [P] \rightarrow (D_s \pi)_P K^*(892)$					$0.1 \pm 0.0$		
$B_s [D] \rightarrow (D_s \pi)_P K^*(892)$					$0.9 \pm 0.0$		
$B_s \rightarrow (D_s K)_S \sigma$							
$B_s \rightarrow (D_s K)_P \sigma$							
$B_s \rightarrow (D_s K)_S f_0(980)$							
$B_s \rightarrow (D_s K)_S f_2(1270)$							$0.1 \pm 0.0$
$B_s \rightarrow (D_s K)_P f_2(1270)$							
$B_s \rightarrow (D_s K)_S f_0(1370)$							
$B_s \rightarrow (D_s K)_S \rho(770)$							
$B_s \rightarrow (D_s K)_P \rho(770)$		$0.8 \pm 0.0$					$0.1 \pm 0.0$
$B_s [P] \rightarrow (D_s K)_P \rho(770)$							
$B_s [D] \rightarrow (D_s K)_P \rho(770)$							
Sum	$98.9 \pm 3.8$	$105.3 \pm 0.0$	$104.9 \pm 0.0$	$108.3 \pm 6.7$	$109.5 \pm 0.0$	$103.1 \pm 9.6$	$95.1 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$15.0 \pm 4.9$	$21.3 \pm 0.0$	$21.7 \pm 0.0$	$18.2 \pm 5.2$	$11.6 \pm 0.0$	$19.1 \pm 6.7$	$14.1 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)$				$4.0 \pm 0.0$			
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$29.5 \pm 4.5$	$32.5 \pm 0.0$	$30.8 \pm 0.0$	$33.0 \pm 5.3$	$29.7 \pm 0.0$	$35.2 \pm 5.9$	$29.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))$							
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$8.0 \pm 2.0$	$8.3 \pm 0.0$	$7.3 \pm 0.0$	$9.2 \pm 1.9$	$7.5 \pm 0.0$	$8.9 \pm 1.9$	$8.2 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$15.5 \pm 6.3$	$31.7 \pm 0.0$	$21.4 \pm 0.0$	$22.1 \pm 12.7$	$22.1 \pm 0.0$	$15.4 \pm 5.4$	$19.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$			$2.4 \pm 0.0$				
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$			$0.9 \pm 0.0$				
$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	$22.3 \pm 3.2$	$19.1 \pm 0.0$	$13.3 \pm 0.0$	$17.2 \pm 3.9$	$15.5 \pm 0.0$	$21.7 \pm 3.4$	$21.4 \pm 0.0$
$B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))$				$3.5 \pm 1.1$			
$B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)$				$2.3 \pm 1.1$			
$b \rightarrow u$							
$B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K_2(1770) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow (D_s \pi)_S K^*(892)$							
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$37.6 \pm 6.3$	$15.7 \pm 0.0$	$23.4 \pm 0.0$	$17.3 \pm 10.3$	$59.7 \pm 0.0$	$29.3 \pm 9.2$	$40.3 \pm 0.0$
$B_s \rightarrow (D_s \pi)_D K^*(892)$					$0.6 \pm 0.0$		
$B_s \rightarrow (D_s \pi)_P K^*(892)$					$3.8 \pm 0.0$		
$B_s \rightarrow (D_s K)_S \sigma$							
$B_s \rightarrow (D_s K)_P \sigma$							
$B_s \rightarrow (D_s K)_S f_0(980)$							
$B_s \rightarrow (D_s K)_S f_2(1270)$							$0.2 \pm 0.0$
$B_s \rightarrow (D_s K)_P f_2(1270)$							
$B_s \rightarrow (D_s K)_S f_0(1370)$							
$B_s \rightarrow (D_s K)_S \rho(770)$							
$B_s \rightarrow (D_s K)_P \rho(770)$	$8.4 \pm 2.4$	$8.3 \pm 0.0$	$4.9 \pm 0.0$	$13.3 \pm 2.7$	$8.4 \pm 0.0$	$12.6 \pm 2.1$	$7.9 \pm 0.0$
$B_s [P] \rightarrow (D_s K)_P \rho(770)$						$0.9 \pm 1.2$	
$B_s [D] \rightarrow (D_s K)_P \rho(770)$						$0.8 \pm 0.1$	
Sum	$136.4 \pm 12.1$	$140.2 \pm 0.0$	$126.9 \pm 0.0$	$136.0 \pm 13.3$	$158.8 \pm 0.0$	$143.8 \pm 12.1$	$140.1 \pm 0.0$
$m_{K_1(1400)}$ [MeV]	$1398 \pm 9$	$1387 \pm 8$	$1386 \pm 10$	$1394 \pm 10$	$1409 \pm 10$	$1392 \pm 10$	$1400 \pm 10$
$\Gamma_{K_1(1400)}$ [MeV]	$204 \pm 14$	$207 \pm 15$	$205 \pm 16$	$220 \pm 18$	$221 \pm 17$	$207 \pm 18$	$203 \pm 13$
$m_{K^*(1410)}$ [MeV]	$1432 \pm 12$	$1436 \pm 11$	$1437 \pm 12$	$1429 \pm 12$	$1435 \pm 8$	$1428 \pm 11$	$1433 \pm 11$
$\Gamma_{K^*(1410)}$ [MeV]	$344 \pm 25$	$348 \pm 26$	$344 \pm 26$	$347 \pm 26$	$360 \pm 25$	$346 \pm 23$	$346 \pm 25$
$r$	$0.00 \pm 0.03$	$-0.04 \pm 0.08$	$-0.03 \pm 0.05$	$-0.01 \pm 0.04$	$0.01 \pm 0.04$	$-0.01 \pm 0.04$	$0.01 \pm 0.04$
$\kappa$	$0.00 \pm 0.05$	$0.09 \pm 0.00$	$0.01 \pm 0.00$	$-0.22 \pm 0.07$	$0.02 \pm 0.00$	$-0.01 \pm 0.06$	$0.00 \pm 0.00$
$\delta$ [ $^\circ$ ]	$0 \pm 15$	$12 \pm 15$	$12 \pm 15$	$-0 \pm 17$	$-4 \pm 15$	$9 \pm 16$	$0 \pm 15$
$\gamma - 2\beta_s$ [ $^\circ$ ]	$0 \pm 16$	$9 \pm 18$	$-4 \pm 16$	$-1 \pm 19$	$-8 \pm 14$	$6 \pm 16$	$2 \pm 16$

**Table 11.2:** Fit fractions in percent for several alternative amplitude models (Alt. 7 - Alt. 12). Resonance parameters and the physical observables  $r, \kappa, \delta, \gamma - 2\beta_s$  are also given. The values of the physical observables are given relative to the baseline result. The uncertainties are statistical only.

		Alt.7	Alt.8	Alt.9	Alt.10	Alt.11	Alt.12
$b \rightarrow c$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	6.9 ± 1.8	8.0 ± 0.0	6.4 ± 0.0	6.7 ± 1.2	8.4 ± 0.0	6.3 ± 2.3
	$B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)$						
	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	13.8 ± 2.0	13.0 ± 0.0	13.5 ± 0.0	11.0 ± 1.7	15.2 ± 0.0	14.9 ± 3.4
	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))$						
	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	3.8 ± 0.5	3.3 ± 0.0	3.6 ± 0.0	3.3 ± 0.7	4.2 ± 0.0	3.3 ± 0.8
	$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	47.8 ± 4.3	45.8 ± 0.0	50.5 ± 0.0	52.6 ± 9.1	45.9 ± 0.0	49.9 ± 8.4
	$B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))$						0.6 ± 0.5
	$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	15.4 ± 0.9	15.3 ± 0.0	18.3 ± 0.0	15.2 ± 1.3	15.2 ± 0.0	15.9 ± 1.1
	$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	5.8 ± 0.5	6.2 ± 0.0	5.7 ± 0.0	6.3 ± 0.6	6.5 ± 0.0	6.3 ± 0.7
	$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K^*(892) \pi)$						
	$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K \rho(770))$						
	$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$						
	$B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))$						
	$B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)$						
	$B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)$			0.8 ± 0.0			
	$B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))$			1.1 ± 0.0			
	$B_s \rightarrow D_s (K_2(1770) \rightarrow K^*(892) \pi)$				0.7 ± 0.3		
	$B_s \rightarrow (D_s \pi)_S K^*(892)$						
	$B_s \rightarrow (D_s \pi)_P K^*(892)$	7.8 ± 1.3	5.4 ± 0.0	7.7 ± 0.0	7.3 ± 2.0	6.8 ± 0.0	6.5 ± 2.5
	$B_s [P] \rightarrow (D_s \pi)_P K^*(892)$						
	$B_s [D] \rightarrow (D_s \pi)_P K^*(892)$						
	$B_s \rightarrow (D_s K)_S \sigma$	1.6 ± 0.9	0.4 ± 0.0				
	$B_s \rightarrow (D_s K)_P \sigma$	2.9 ± 1.2					
	$B_s \rightarrow (D_s K)_S f_0(980)$		0.5 ± 0.0				
	$B_s \rightarrow (D_s K)_S f_2(1270)$						
	$B_s \rightarrow (D_s K)_P f_2(1270)$						
	$B_s \rightarrow (D_s K)_S f_0(1370)$						
	$B_s \rightarrow (D_s K)_S \rho(770)$						
	$B_s \rightarrow (D_s K)_P \rho(770)$						
	$B_s [P] \rightarrow (D_s K)_P \rho(770)$						
	$B_s [D] \rightarrow (D_s K)_P \rho(770)$						
	Sum	105.7 ± 5.2	98.0 ± 0.0	107.7 ± 0.0	103.1 ± 8.7	102.2 ± 0.0	103.8 ± 9.2
$b \rightarrow u$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	9.7 ± 2.7	14.6 ± 0.0	14.4 ± 0.0	19.3 ± 5.8	7.7 ± 0.0	13.3 ± 5.6
	$B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)$						
	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	19.3 ± 3.7	23.8 ± 0.0	30.1 ± 0.0	31.7 ± 6.7	13.8 ± 0.0	31.5 ± 7.3
	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))$						
	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	5.3 ± 1.7	6.1 ± 0.0	7.9 ± 0.0	9.5 ± 1.8	3.8 ± 0.0	7.1 ± 1.8
	$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	6.8 ± 4.0	14.6 ± 0.0	10.7 ± 0.0	18.7 ± 8.6	9.1 ± 0.0	15.7 ± 10.4
	$B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))$						0.2 ± 0.3
	$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$						
	$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$						
	$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K^*(892) \pi)$						
	$B_s \rightarrow D_s (K_2^*(1430) \rightarrow K \rho(770))$						
	$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	29.4 ± 4.9	26.9 ± 0.0	23.5 ± 0.0	22.6 ± 4.6	28.9 ± 0.0	22.8 ± 4.4
	$B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))$						
	$B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)$						
	$B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)$						
	$B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))$						
	$B_s \rightarrow D_s (K_2(1770) \rightarrow K^*(892) \pi)$			0.8 ± 1.2			
	$B_s \rightarrow (D_s \pi)_S K^*(892)$						
	$B_s \rightarrow (D_s \pi)_P K^*(892)$	45.5 ± 7.5	39.9 ± 0.0	39.0 ± 0.0	27.0 ± 11.7	47.3 ± 0.0	35.3 ± 10.0
	$B_s [P] \rightarrow (D_s \pi)_P K^*(892)$						
	$B_s [D] \rightarrow (D_s \pi)_P K^*(892)$						
	$B_s \rightarrow (D_s K)_S \sigma$	0.3 ± 0.7	0.8 ± 0.0				
	$B_s \rightarrow (D_s K)_P \sigma$	0.6 ± 1.2					
	$B_s \rightarrow (D_s K)_S f_0(980)$		1.0 ± 0.0				
	$B_s \rightarrow (D_s K)_S f_2(1270)$						
	$B_s \rightarrow (D_s K)_P f_2(1270)$						
	$B_s \rightarrow (D_s K)_S f_0(1370)$						
	$B_s \rightarrow (D_s K)_S \rho(770)$					3.9 ± 0.0	
	$B_s \rightarrow (D_s K)_P \rho(770)$	4.6 ± 1.1	8.2 ± 0.0	8.6 ± 0.0	7.9 ± 3.2		7.3 ± 2.9
	$B_s [P] \rightarrow (D_s K)_P \rho(770)$						
	$B_s [D] \rightarrow (D_s K)_P \rho(770)$						
	Sum	121.6 ± 9.1	135.7 ± 0.0	134.3 ± 0.0	137.5 ± 13.2	114.6 ± 0.0	133.1 ± 15.8
$m_{K_1(1400)}$ [MeV]	$m_{K_1(1400)}$ [MeV]	1401 ± 9	1390 ± 10	1397 ± 10	1394 ± 9	1398 ± 8	1393 ± 11
	$\Gamma_{K_1(1400)}$ [MeV]	195 ± 17	199 ± 13	198 ± 16	208 ± 12	193 ± 14	205 ± 16
	$m_{K^*(1410)}$ [MeV]	1444 ± 11	1439 ± 11	1412 ± 17	1432 ± 11	1436 ± 10	1433 ± 10
	$\Gamma_{K^*(1410)}$ [MeV]	329 ± 24	339 ± 23	400 ± 20	351 ± 24	334 ± 24	346 ± 25
	$r$	-0.06 ± 0.04	-0.04 ± 0.04	-0.01 ± 0.04	-0.04 ± 0.05	-0.04 ± 0.04	-0.01 ± 0.05
	$\kappa$	-0.06 ± 0.04	-0.00 ± 0.00	-0.01 ± 0.00	0.02 ± 0.07	0.04 ± 0.00	-0.08 ± 0.05
	$\delta [^\circ]$	-12 ± 17	4 ± 14	2 ± 17	3 ± 14	-11 ± 17	-2 ± 16
	$\gamma - 2\beta_s [^\circ]$	-7 ± 17	8 ± 16	-4 ± 17	7 ± 14	-9 ± 17	-0 ± 17

**Table 11.3:** Systematic uncertainties on the fit parameters of the fit to  $B_s \rightarrow D_s\pi\pi$  data in units of statistical standard deviations.

Fit Parameter	Fit-bias	Acceptance	Resolution	Asymmetries	Background	Mom./z-Scale	Total
$p_0^{OS}$ Run-I	0.05	0.01	0.99	0.00	0.09		1.00
$p_1^{OS}$ Run-I	0.01	0.01	1.04	0.01	0.13		1.05
$\Delta p_0^{OS}$ Run-I	0.14	0.00	0.02	0.15	0.03		0.21
$\Delta p_1^{OS}$ Run-I	0.07	0.00	0.03	0.15	0.06		0.18
$\epsilon_{tag}^{OS}$ Run-I	0.06	0.00	0.00	0.01	0.17		0.18
$\Delta \epsilon_{tag}^{OS}$ Run-I	0.04	0.00	0.06	0.01	0.01		0.07
$p_0^{SS}$ Run-I	0.03	0.01	0.56	0.00	0.03		0.56
$p_1^{SS}$ Run-I	0.10	0.01	0.60	0.01	0.03		0.60
$\Delta p_0^{SS}$ Run-I	0.04	0.00	0.00	0.10	0.01		0.11
$\Delta p_1^{SS}$ Run-I	0.03	0.00	0.01	0.12	0.04		0.13
$\epsilon_{tag}^{SS}$ Run-I	0.02	0.00	0.00	0.01	0.02		0.04
$\Delta \epsilon_{tag}^{SS}$ Run-I	0.04	0.00	0.05	0.01	0.03		0.07
$p_0^{OS}$ Run-II	0.02	0.02	0.93	0.00	0.20		0.96
$p_1^{OS}$ Run-II	0.02	0.01	0.70	0.00	0.08		0.70
$\Delta p_0^{OS}$ Run-II	0.07	0.00	0.02	0.00	0.08		0.11
$\Delta p_1^{OS}$ Run-II	0.02	0.00	0.02	0.00	0.03		0.04
$\epsilon_{tag}^{OS}$ Run-II	0.01	0.00	0.00	0.00	0.16		0.16
$\Delta \epsilon_{tag}^{OS}$ Run-II	0.05	0.00	0.01	0.00	0.05		0.07
$p_0^{SS}$ Run-II	0.10	0.01	0.60	0.00	0.06		0.62
$p_1^{SS}$ Run-II	0.01	0.02	0.64	0.00	0.07		0.64
$\Delta p_0^{SS}$ Run-II	0.07	0.00	0.02	0.00	0.02		0.08
$\Delta p_1^{SS}$ Run-II	0.11	0.00	0.02	0.00	0.05		0.12
$\epsilon_{tag}^{SS}$ Run-II	0.03	0.00	0.00	0.00	0.03		0.04
$\Delta \epsilon_{tag}^{SS}$ Run-II	0.01	0.00	0.10	0.00	0.03		0.11
$A_P$ Run-II	0.04	0.00	0.06	0.01	0.02		0.08
$\Delta m_s$	0.01	0.00	0.16	0.02	0.11	0.67	0.69

**Table 11.4:** Systematic uncertainties on the fit parameters of the phase-space integrated fit to  $B_s \rightarrow D_s K\pi\pi$  data in units of statistical standard deviations.

Fit Parameter	Fit bias	Acceptance	Resolution	$\Delta m_s$	Asymmetries	Background	Total
$C$	0.06	0.04	0.15	0.06	0.02	0.06	0.19
$D$	0.02	0.22	0.01	0.02	0.04	0.10	0.25
$\bar{D}$	0.04	0.22	0.02	0.02	0.04	0.13	0.26
$S$	0.01	0.02	0.07	0.21	0.03	0.06	0.23
$\bar{S}$	0.07	0.02	0.05	0.22	0.02	0.03	0.24

**Table 11.5:** Systematic uncertainties on the fit parameters of the full time-dependent amplitude fit to  $B_s \rightarrow D_s K\pi\pi$  data in units of statistical standard deviations.

Fit Parameter	Fit bias	Time-Acc.	Resolution	$\Delta m_s$	Asymmetries	Background	Lineshapes	Resonances $m, \Gamma$	Form-Factors	Phsp-Acc.	Amp. Model	Total
$B_s \rightarrow D_s(K_1(1270) \rightarrow K^*(892)\pi)$ Mag	0.10	0.01	0.04	0.01	0.00	0.13	0.48	0.24	0.52	0.06	0.77	
$B_s \rightarrow D_s(K_1(1270) \rightarrow K^*(892)\pi)$ Phase	0.07	0.01	0.04	0.01	0.01	0.08	0.35	0.28	0.34	0.12	0.58	
$B_s \rightarrow D_s(K_1(1270) \rightarrow K_0^*(1430)\pi)$ Mag	0.04	0.01	0.01	0.00	0.00	0.24	1.44	0.11	0.17	0.04	1.47	
$B_s \rightarrow D_s(K_1(1270) \rightarrow K_0^*(1430)\pi)$ Phase	0.04	0.01	0.02	0.01	0.00	0.19	5.83	0.19	0.61	0.09	5.87	
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow c$ )	0.13	0.03	0.16	0.06	0.02	0.34	1.32	0.37	0.78	0.19	1.64	
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow c$ )	0.14	0.02	0.09	0.02	0.01	0.18	0.54	0.26	0.40	0.08	0.77	
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow u$ )	0.10	0.04	0.05	0.12	0.04	0.32	0.35	0.22	0.73	0.16	0.93	
$B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow u$ )	0.02	0.04	0.04	0.10	0.03	0.08	0.79	0.21	0.31	0.08	0.89	
$B_s \rightarrow D_s(K^*(1410) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow c$ )	0.08	0.03	0.08	0.08	0.03	0.18	0.61	0.25	0.75	0.28	1.06	
$B_s \rightarrow D_s(K^*(1410) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow c$ )	0.35	0.01	0.06	0.01	0.01	0.13	0.60	0.19	0.68	0.08	1.00	
$B_s \rightarrow D_s(K^*(1410) \rightarrow K\rho(770))$ Mag	0.35	0.01	0.02	0.01	0.00	0.18	0.59	0.12	0.34	0.06	0.79	
$B_s \rightarrow D_s(K^*(1410) \rightarrow K\rho(770))$ Phase	0.18	0.00	0.01	0.01	0.00	0.24	0.34	0.09	0.21	0.06	0.51	
$B_s \rightarrow D_s(K(1460) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow u$ )	0.14	0.03	0.05	0.05	0.02	0.37	0.43	0.27	0.60	0.12	0.89	
$B_s \rightarrow D_s(K(1460) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow u$ )	0.13	0.04	0.11	0.07	0.03	0.21	0.84	0.49	0.46	0.06	1.11	
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Mag( $b \rightarrow c$ )	0.03	0.02	0.06	0.02	0.01	0.24	0.95	0.11	0.55	0.13	1.14	
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Phase( $b \rightarrow c$ )	0.20	0.01	0.13	0.02	0.01	0.51	1.10	0.18	0.52	0.26	1.38	
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Mag( $b \rightarrow u$ )	0.14	0.04	0.07	0.06	0.02	0.11	0.78	0.24	0.54	0.17	1.01	
$B_s \rightarrow (D_s\pi)_P K^*(892)$ Phase( $b \rightarrow u$ )	0.24	0.05	0.19	0.06	0.03	0.47	1.54	0.28	0.59	0.17	1.77	
$B_s \rightarrow (D_sK)_P K^*(892)$ Phase( $b \rightarrow u$ )	0.35	0.04	0.02	0.05	0.02	0.25	0.75	0.31	0.60	0.06	1.10	
$B_s \rightarrow (D_sK)_P \rho(770)$ Mag( $b \rightarrow u$ )	0.12	0.03	0.05	0.06	0.02	0.68	0.50	0.38	0.66	0.08	1.14	
$m_{K_1(1400)}$	0.09	0.01	0.08	0.01	0.00	0.14	0.21	0.13	0.37	0.09	0.72	
$\Gamma_{K_1(1400)}$	0.01	0.01	0.01	0.02	0.01	0.14	0.46	0.13	0.44	0.10	0.62	
$m_{K^*(1410)}$	0.05	0.01	0.02	0.01	0.00	0.08	0.26	0.04	1.29	0.12	0.67	
$\Gamma_{K^*(1410)}$	0.25	0.00	0.02	0.01	0.00	0.14	0.15	0.04	1.40	0.07	1.61	
$r$	0.11	0.05	0.09	0.12	0.03	0.47	0.74	0.12	0.26	0.12	0.79	
$\delta$	0.19	0.04	0.07	0.10	0.05	0.10	0.29	0.03	0.11	0.02	0.52	
$\gamma - 2\beta_s$	0.10	0.06	0.12	0.06	0.02	0.12	0.27	0.03	0.10	0.03	0.39	
											0.53	

## 1026 A Stripping and Trigger cuts

1027 The following text describes variables which are used in Table 1.1 and might be ambiguous,  
 1028 or which benefits are not straight forward. Where noted, different cut values are applied  
 1029 for Run-I and Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest  
 1030 approach. This variable is used to ensure that all  $D_s$  and  $X_{s,d}$  daughters originate from  
 1031 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 $p$	$> 1000 \text{ MeV}/c$	
Track $p_T$	$> 100 \text{ MeV}/c$	
Track IP $\chi^2$	$> 4$	
Track ghost-prob.	$< 0.4$	
$D_s$ mass	$m_{D_s} \pm 100 \text{ MeV}$	$m_{D_s} \pm 80 \text{ MeV}$
$D_s$ Daughter $p_T$	$\sum_{i=1}^3 p_{t,i} > 1800 \text{ MeV}/c$	
$D_s$ Daughter DOCA	$< 0.5 \text{ mm}$	
$D_s$ Vertex $\chi^2/\text{nDoF}$	$< 10$	
$D_s$ $\chi^2$ -separation from PV	$> 36$	
$D_s$ daughter PID( $\pi$ )	$< 20$	
$D_s$ daughter PID(K)	$> -10$	
$X_{s,d}$ mass	$< 4000 \text{ MeV}$	$< 3500 \text{ MeV}$
$X_{s,d}$ Daughter $p$	$> 2 \text{ GeV}/c$	
$X_{s,d}$ Daughter DOCA	$< 0.4 \text{ mm}$	
$X_{s,d}$ Daughter $p_T$	$\sum_{i=1}^3 p_{t,i} > 1250 \text{ MeV}/c$	
$X_{s,d}$ Vertex $\chi^2/\text{nDoF}$	$< 8$	
$X_{s,d}$ $\chi^2$ -separation from PV	$> 16$	
$X_{s,d}$ DIRA	$> 0.98$	
$X_{s,d}$ $\Delta\rho$	$> 0.1 \text{ mm}$	
$X_{s,d}$ $\Delta z$	$> 2.0 \text{ mm}$	
$X_{s,d}$ daughter PID( $\pi$ )	$< 10$	
$X_s$ daughter PID(K)	$> -2$	$> 4$
$B_s^0$ mass	$[4750, 7000] \text{ MeV}/c^2$	$[5000, 6000] \text{ MeV}/c^2$
$B_s^0$ DIRA	$> 0.98$	$> 0.99994$
$B_s^0$ min IP $\chi^2$	$< 25$	$< 20$
$B_s^0$ Vertex $\chi^2/\text{nDoF}$	$< 10$	$< 8$
$B_s^0 \tau_{B_s^0}$	$> 0.2 \text{ ps}$	

1032

1033 Table 1.2 summarizes the trigger requirements imposed by the Hlt1 line used in this  
 1034 analysis for Run-I. At least one of the six decay particles must pass the listed requirements  
 1035 in order for the event to be stored for further analysis. For Run-II, this trigger line was  
 1036 updated and uses a multivariate classifier which takes the variables listed in Table 1.2 as  
 1037 input, rather than directly cutting on them.

1038 The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the  
 1039 b-hadron  $p_T$ , its flight distance  $\chi^2$  from the nearest PV and the sum of the  $B_s^0$  and  $D_s$   
 1040 vertex  $\chi^2$  divided by the sum of their number of degrees of freedom. Table 1.3 summarizes  
 1041 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.

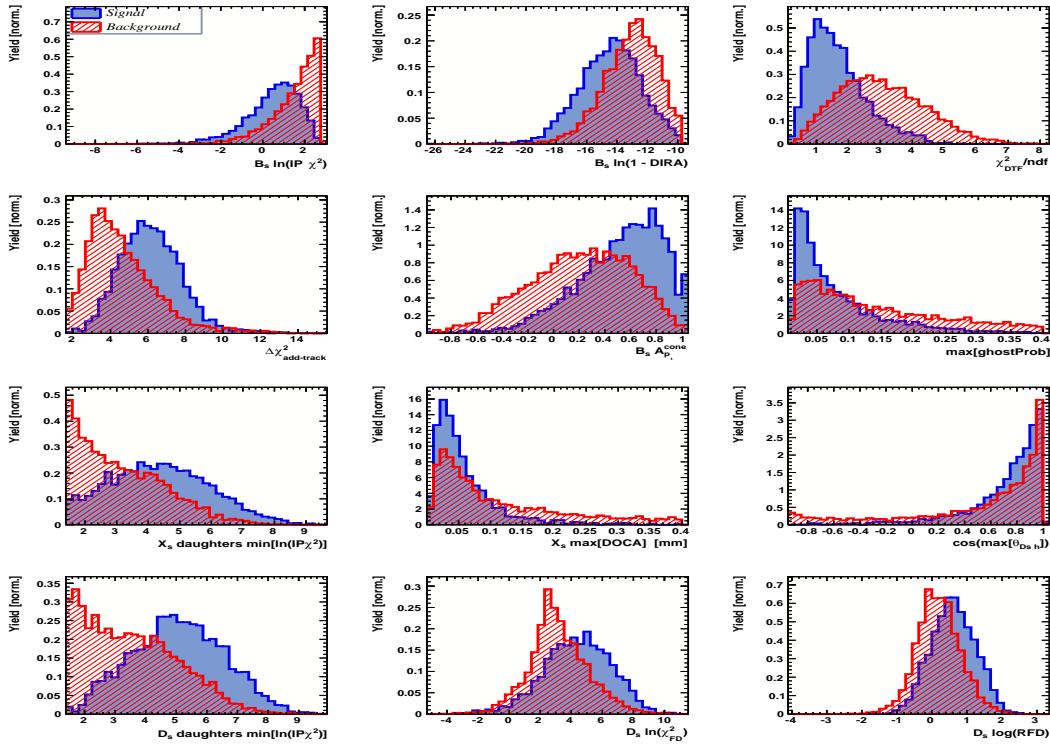
Quantity	Hlt1TrackAllL0 requirement
Track IP [mm]	> 0.1
Track IP $\chi^2$	> 16
Track $\chi^2/\text{nDoF}$	< 2.5
Track $p_T$	> 1.7 GeV/c
Track $p$	> 10 GeV/c
Number VELO hits/track	> 9
Number missed VELO hits/track	< 3
Number OT+IT $\times 2$ hits/track	> 16

**Table 1.3:** Summary of the cuts applied by the Hlt2 inclusive  $\phi$  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.

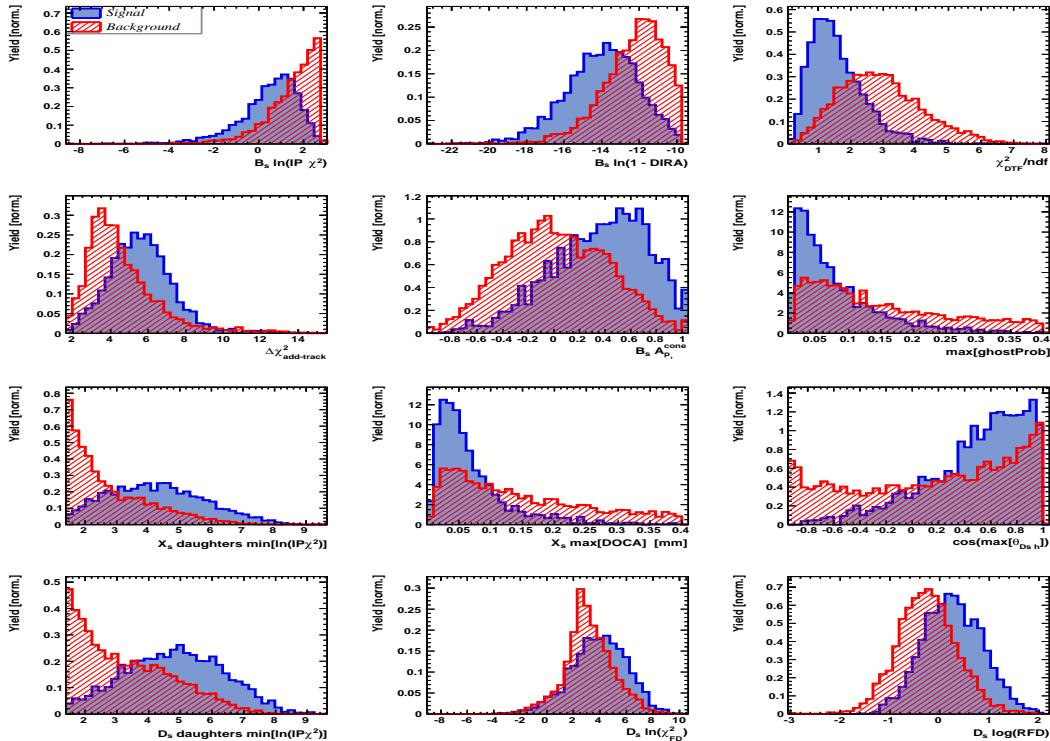
Quantity	Hlt2IncPhi requirement
$\phi$ mass	$m_\phi \pm 12$ MeV/ $c^2$ of PDG value
$\phi p_T$	> 2.5 GeV/c
$\phi$ vertex $\chi^2/\text{nDoF}$	< 20
$\phi$ IP $\chi^2$ to any PV	> 5

1042

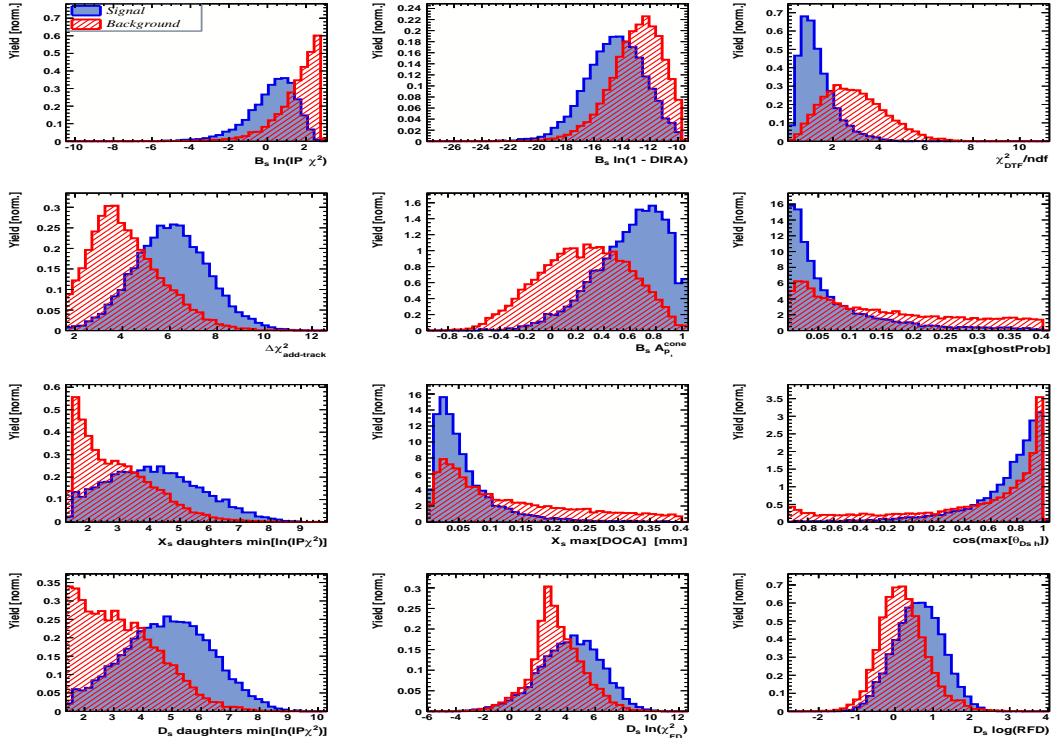
## B Details of multivariate classifier



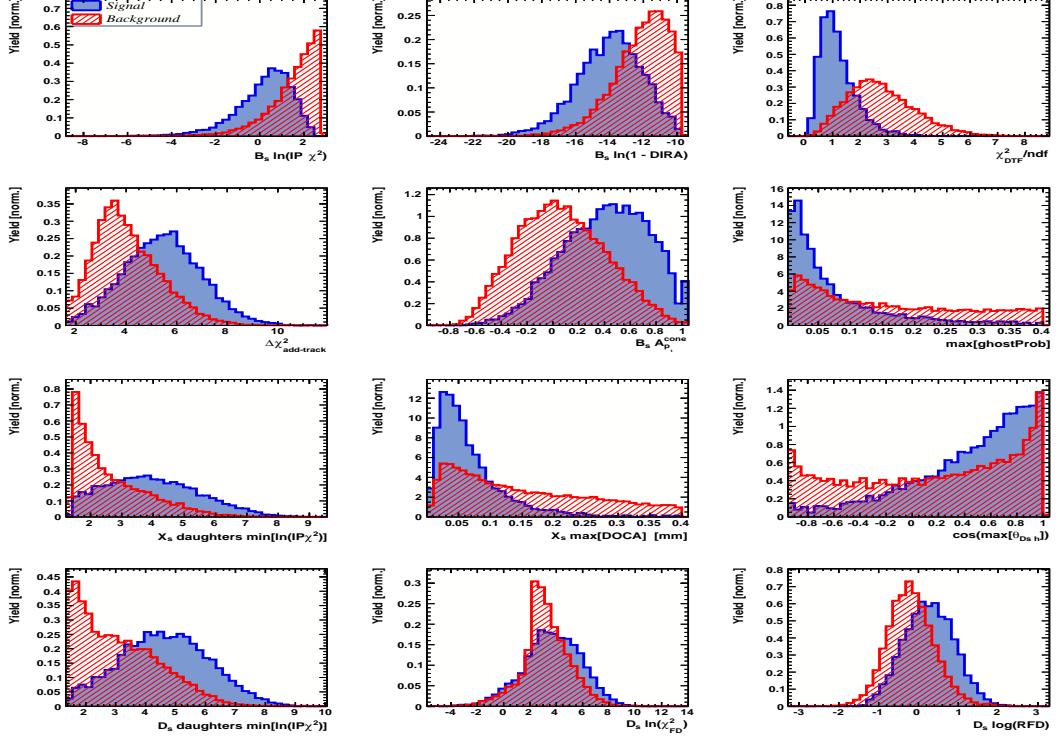
**Figure A.1:** Variables used to train the BDTG for category [Run-I,L0-TOS].



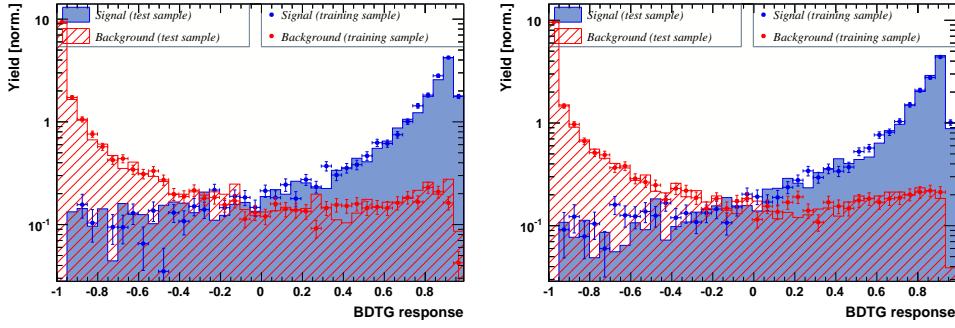
**Figure A.2:** Variables used to train the BDTG for category [Run-I,L0-TIS].



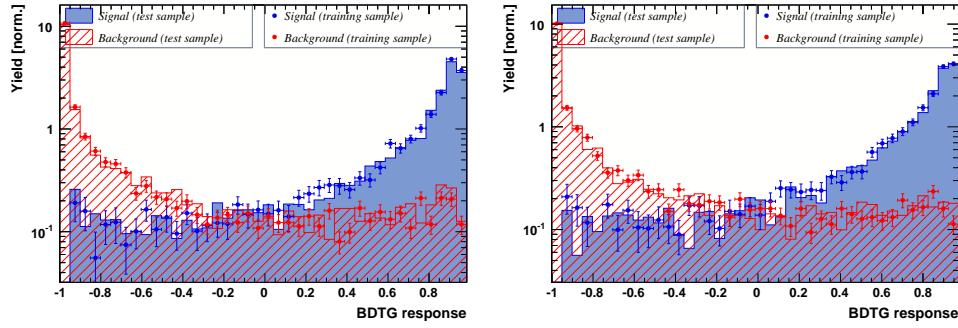
**Figure A.3:** Variables used to train the BDTG for category [Run-II,L0-TOS].



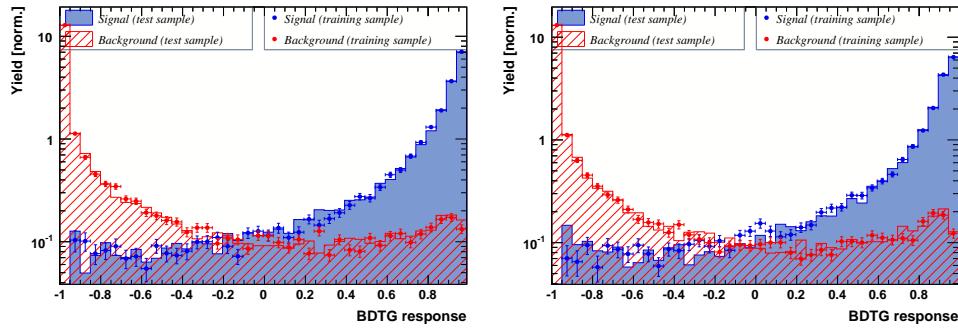
**Figure A.4:** Variables used to train the BDTG for category [Run-II,L0-TIS].



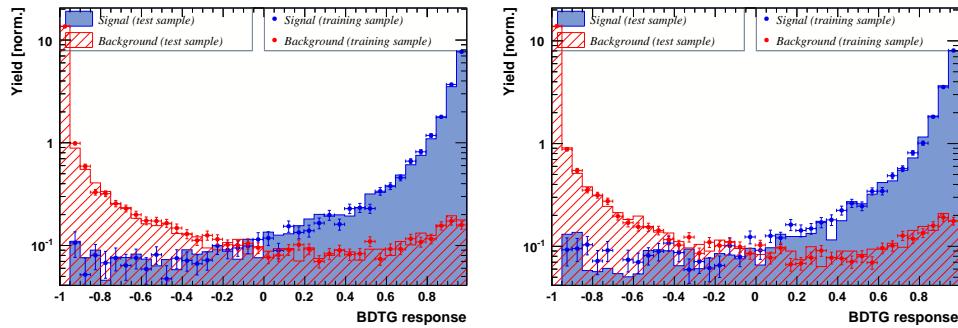
**Figure A.5:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].



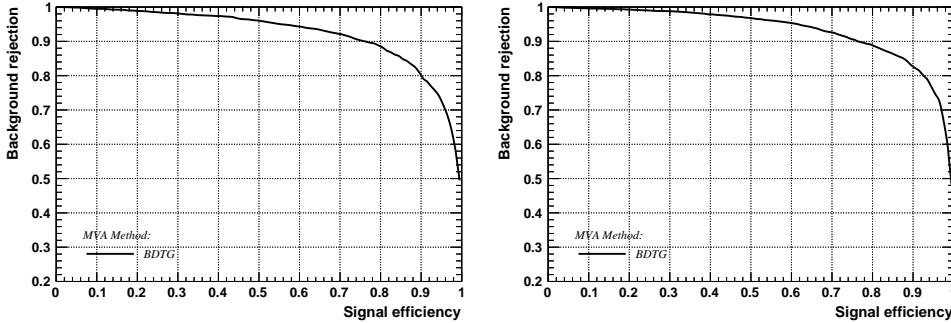
**Figure A.6:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].



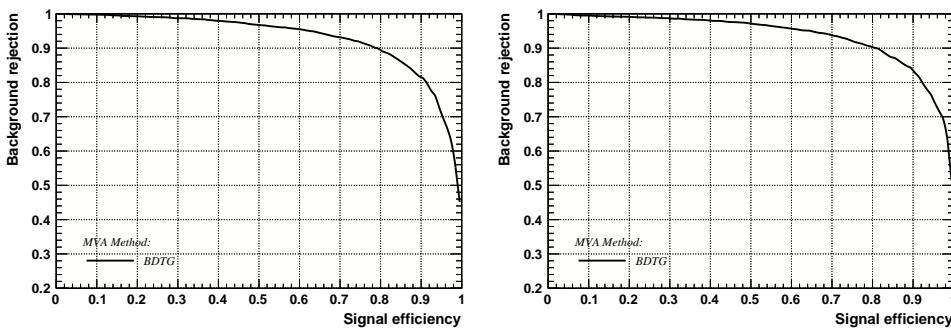
**Figure A.7:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].



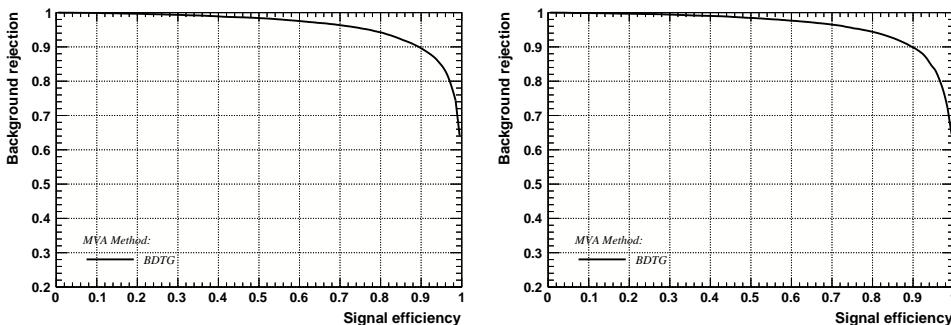
**Figure A.8:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].



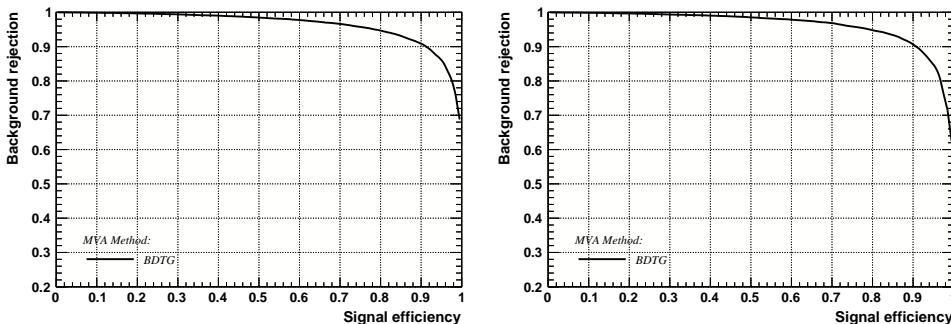
**Figure A.9:** Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].



**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].



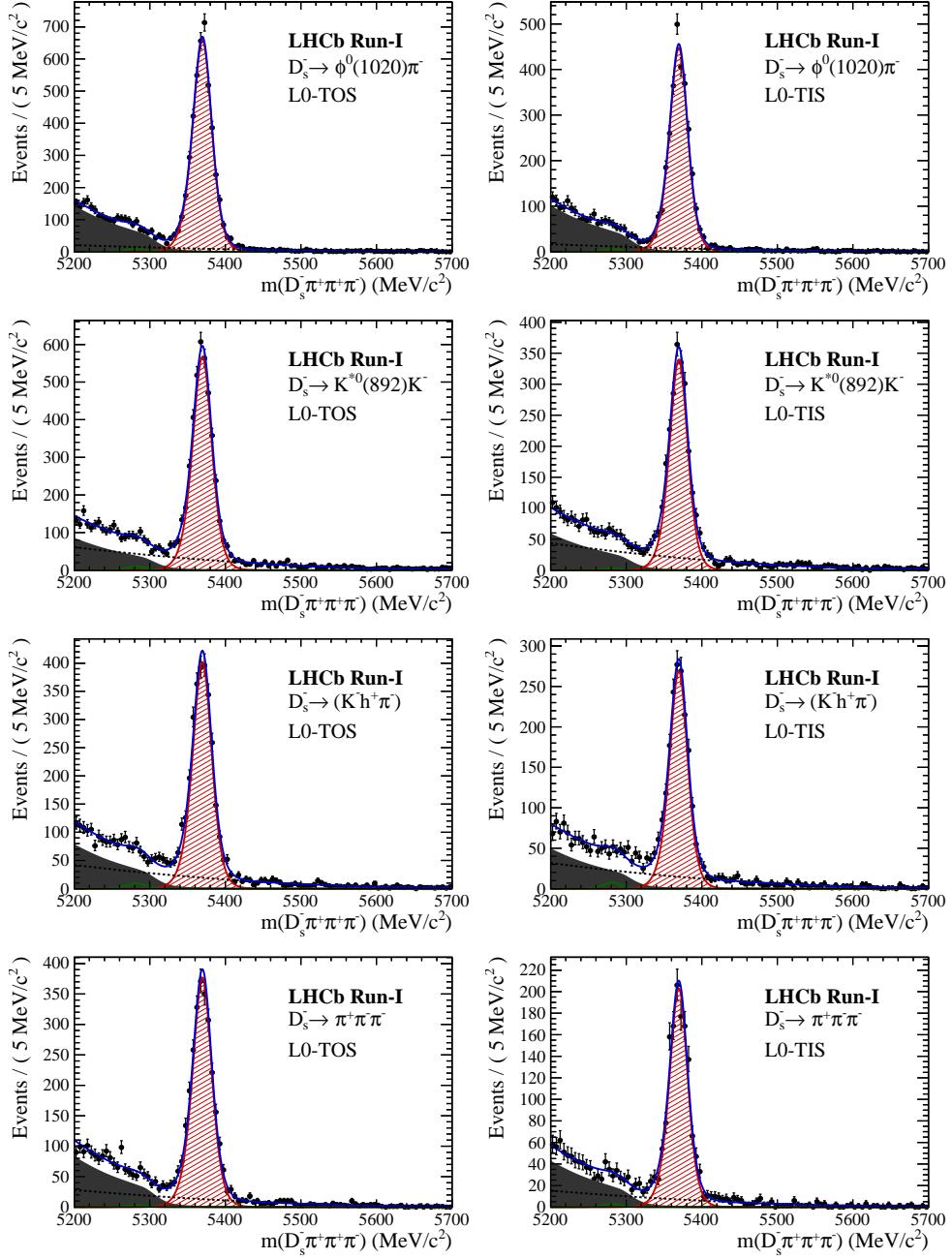
**Figure A.11:** Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TOS].



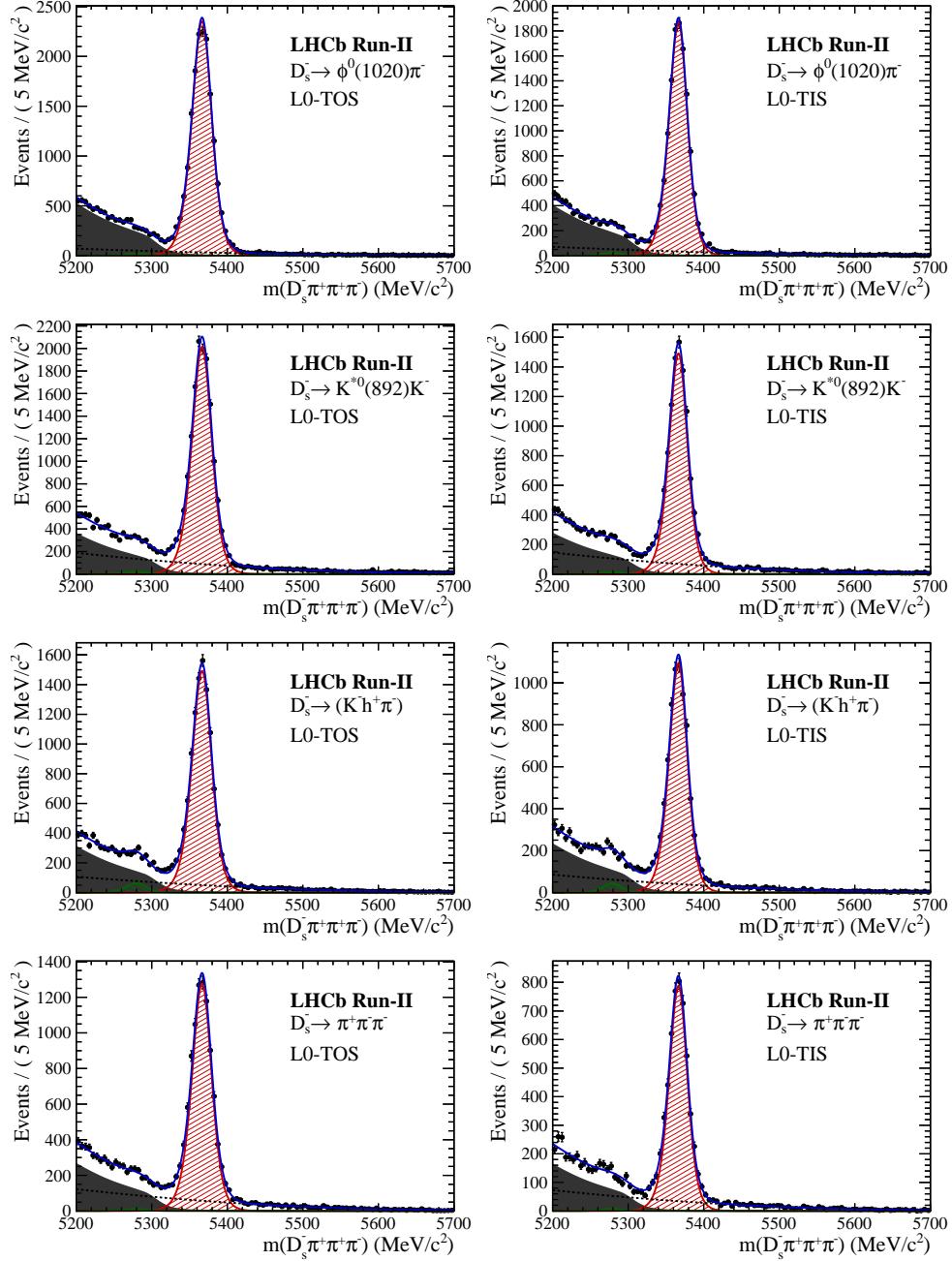
**Figure A.12:** Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TIS].

## 1044 C Detailed mass fits

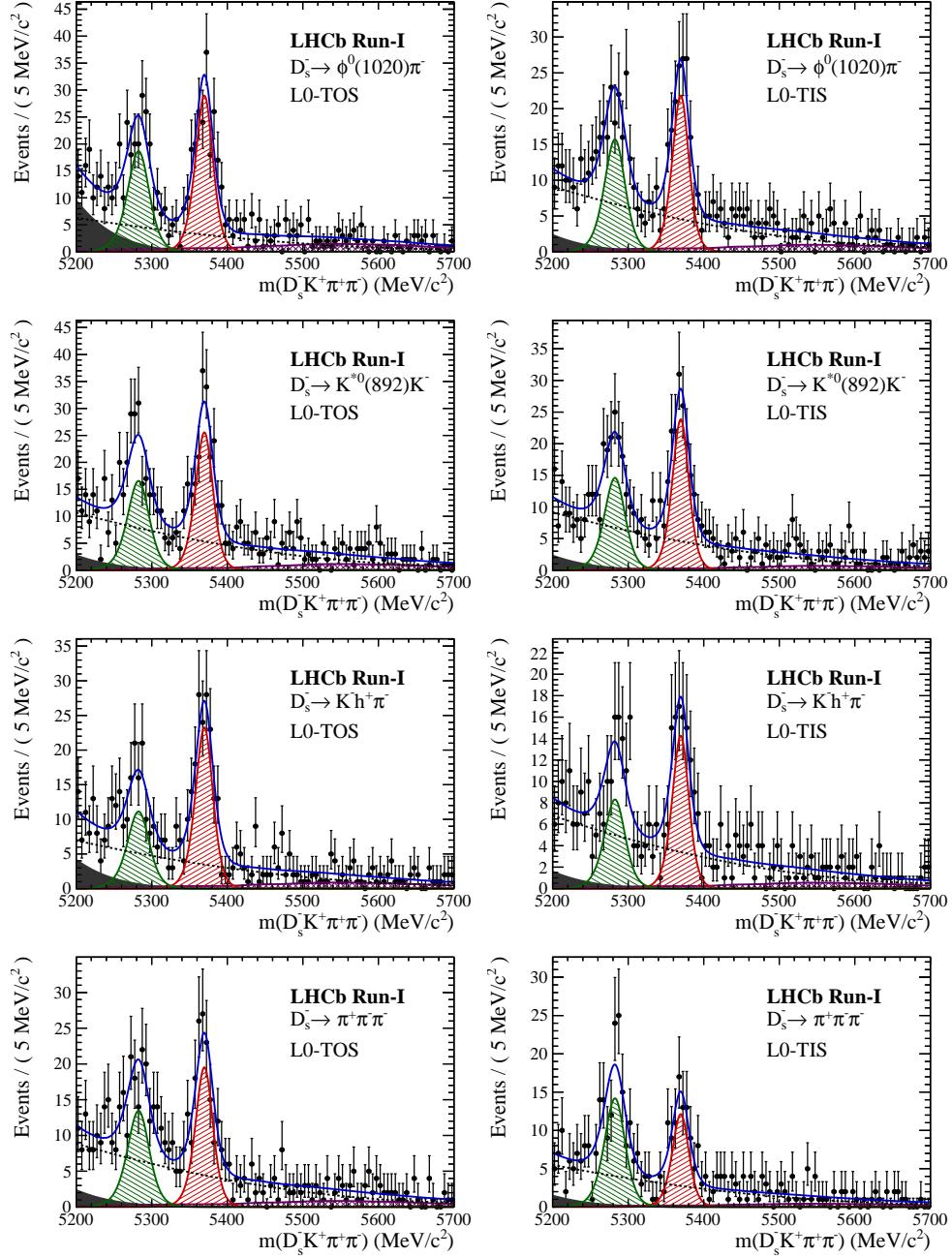
1045 In this section, all fits to the mass distribution of  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$   
 1046 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-  
 1047 II),  $D_s$  final state ( $D_s \rightarrow KK\pi$  non-resonant,  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*K$ , or  $D_s \rightarrow \pi\pi\pi$ ) and  
 1048 L0 trigger category.



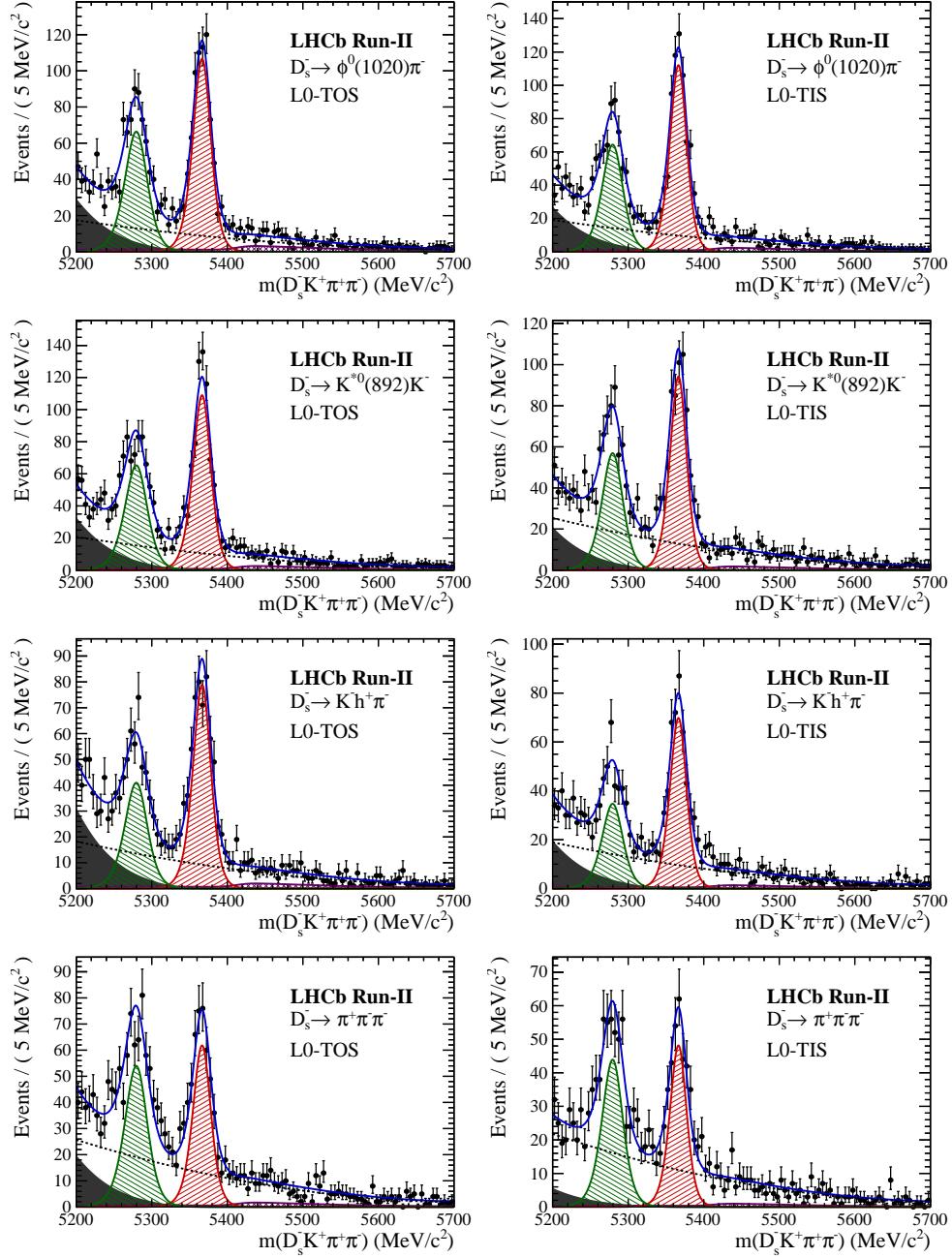
**Figure B.1:** Invariant mass distributions of  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates for Run-I data.



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



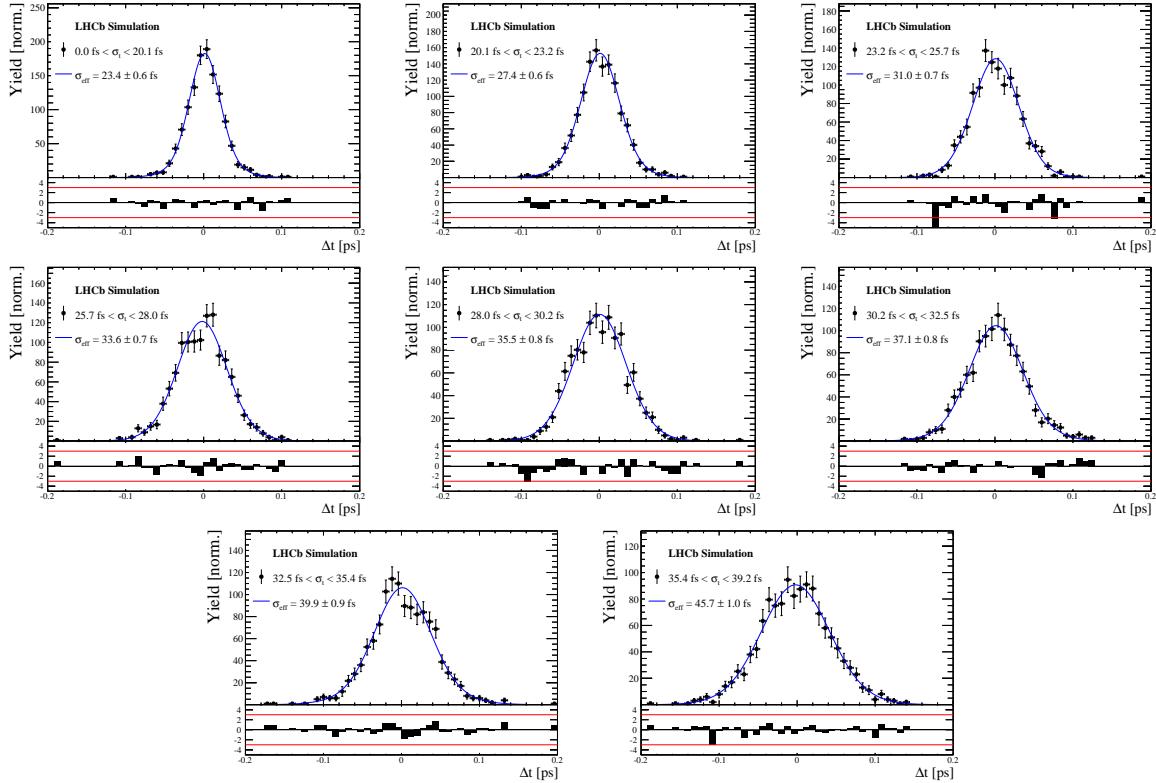
**Figure B.3:** Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for Run-I data.



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

## 1049 D Decay-time Resolution fits

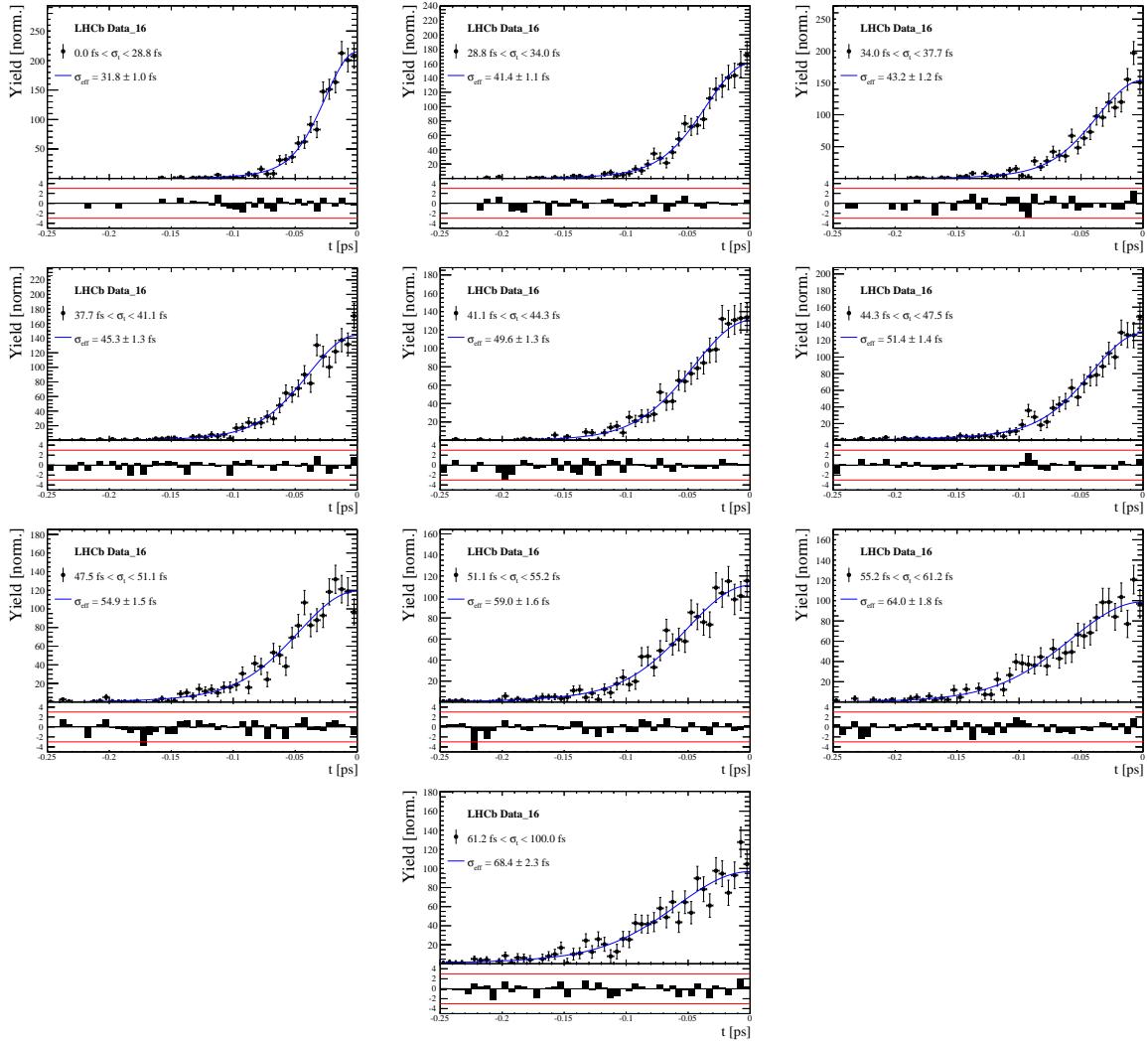
1050 This section contains all fits to the distributions of the decay time difference  $\Delta t$  between  
 1051 the true and the reconstructed decay time of the truth-matched  $B_s^0$  candidates on MC.  
 1052 The fits are performed in bins of the decay time error  $\sigma_t$ , where an adaptive binning  
 1053 scheme is used to ensure that approximately the same number of events are found in each  
 1054 bin.



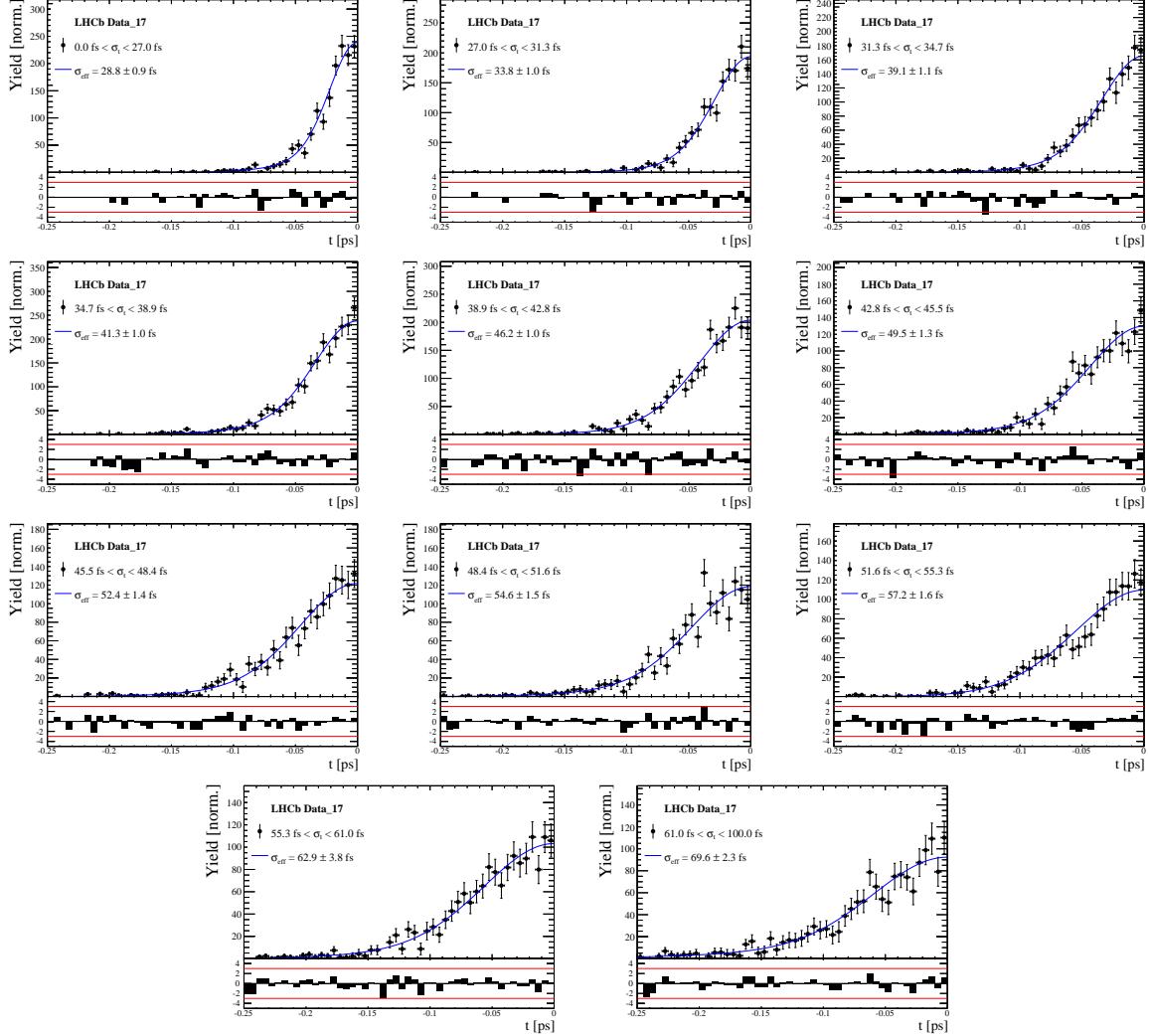
**Figure C.1:** Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K\pi\pi$  MC candidates in bins of the per-event decay time error estimate..

**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.

$\sigma_t$ 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	$55.5 \pm 2.59$	$83 \pm 14.7$	$0.75 \pm 0$	$0.546 \pm 0.0521$	$62 \pm 4.89$



**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. Data taken in 2016.



**Figure C.3:** 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. Data taken in 2017.

**Table 4.2:** Measured time resolution for prompt- $D_s$  data in bins of the per-event decay time error estimate. Data taken in 2016.

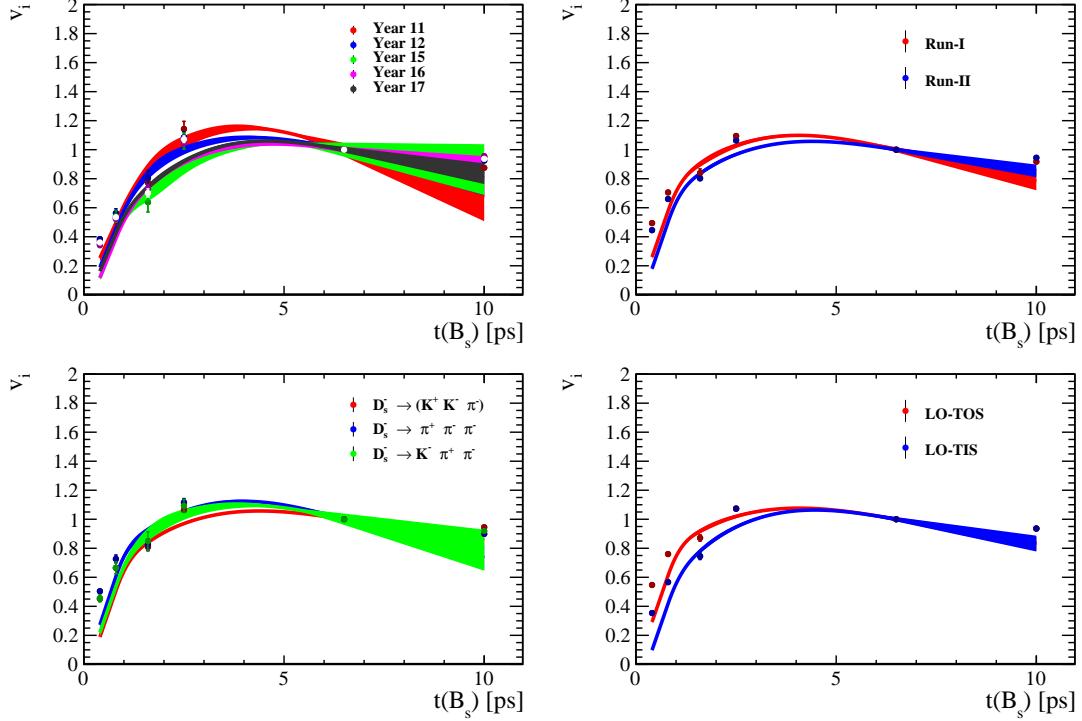
$\sigma_t$ Bin [fs]	$\sigma_1$ [fs]	$\sigma_2$ [fs]	$f_1$	D	$\sigma_{eff}$ [fs]
0.0 - 28.8	$25.3 \pm 1.05$	$47.8 \pm 2.82$	$0.75 \pm 0$	$0.853 \pm 0.00827$	$31.8 \pm 0.967$
28.8 - 34.0	$34.4 \pm 1.38$	$60.4 \pm 3.66$	$0.75 \pm 0$	$0.763 \pm 0.0114$	$41.4 \pm 1.14$
34.0 - 37.7	$35.9 \pm 1.7$	$63.1 \pm 4.05$	$0.75 \pm 0$	$0.745 \pm 0.0122$	$43.2 \pm 1.2$
37.7 - 41.1	$38.3 \pm 1.63$	$65.3 \pm 4.28$	$0.75 \pm 0$	$0.723 \pm 0.013$	$45.3 \pm 1.26$
41.1 - 44.3	$43.6 \pm 2.04$	$67.2 \pm 4.78$	$0.75 \pm 0$	$0.679 \pm 0.0137$	$49.6 \pm 1.3$
44.3 - 47.5	$40.4 \pm 1.64$	$85.4 \pm 5.28$	$0.75 \pm 0$	$0.659 \pm 0.0151$	$51.4 \pm 1.41$
47.5 - 51.1	$46.1 \pm 2.02$	$82.5 \pm 5.61$	$0.75 \pm 0$	$0.622 \pm 0.0161$	$54.9 \pm 1.5$
51.1 - 55.2	$49.2 \pm 2.1$	$92.3 \pm 6.57$	$0.75 \pm 0$	$0.577 \pm 0.0174$	$59 \pm 1.62$
55.2 - 61.2	$56.6 \pm 2.53$	$88.8 \pm 7.84$	$0.75 \pm 0$	$0.525 \pm 0.0192$	$64 \pm 1.82$
61.2 - 100.0	$57.6 \pm 3.09$	$112 \pm 14.5$	$0.75 \pm 0$	$0.478 \pm 0.0232$	$68.4 \pm 2.25$

**Table 4.3:** Measured time resolution for prompt- $D_s$  data in bins of the per-event decay time error estimate. Data taken in 2017.

$\sigma_t$ Bin [fs]	$\sigma_1$ [fs]	$\sigma_2$ [fs]	$f_1$	D	$\sigma_{eff}$ [fs]
0.0 - 27.0	$22.1 \pm 0.942$	$44.7 \pm 2.65$	$0.75 \pm 0$	$0.877 \pm 0.00729$	$28.8 \pm 0.914$
27.0 - 31.3	$28.3 \pm 1.12$	$48.2 \pm 2.98$	$0.75 \pm 0$	$0.835 \pm 0.00854$	$33.8 \pm 0.959$
31.3 - 34.7	$32.8 \pm 1.29$	$55.8 \pm 3.39$	$0.75 \pm 0$	$0.786 \pm 0.0105$	$39.1 \pm 1.08$
34.7 - 38.9	$33.4 \pm 1.2$	$62.5 \pm 3.04$	$0.75 \pm 0$	$0.764 \pm 0.00951$	$41.3 \pm 0.956$
38.9 - 42.8	$40.3 \pm 1.39$	$63 \pm 3.58$	$0.75 \pm 0$	$0.715 \pm 0.0104$	$46.2 \pm 1$
42.8 - 45.5	$42.2 \pm 1.63$	$70.7 \pm 4.35$	$0.75 \pm 0$	$0.68 \pm 0.0138$	$49.5 \pm 1.3$
45.5 - 48.4	$44.7 \pm 1.85$	$75.7 \pm 4.91$	$0.75 \pm 0$	$0.649 \pm 0.0152$	$52.4 \pm 1.42$
48.4 - 51.6	$46.3 \pm 2.05$	$80.3 \pm 5.66$	$0.75 \pm 0$	$0.626 \pm 0.0161$	$54.6 \pm 1.5$
51.6 - 55.3	$53.2 \pm 1.2$	$69.1 \pm 3.03$	$0.75 \pm 0$	$0.597 \pm 0.0167$	$57.2 \pm 1.55$
55.3 - 61.0	$54.8 \pm 2.29$	$90.9 \pm 12.2$	$0.75 \pm 0$	$0.535 \pm 0.0406$	$62.9 \pm 3.82$
61.0 - 100.0	$58.4 \pm 3.18$	$118 \pm 15.6$	$0.75 \pm 0$	$0.466 \pm 0.0233$	$69.6 \pm 2.28$

1055 **E Comparison of time-acceptance in subsamples**

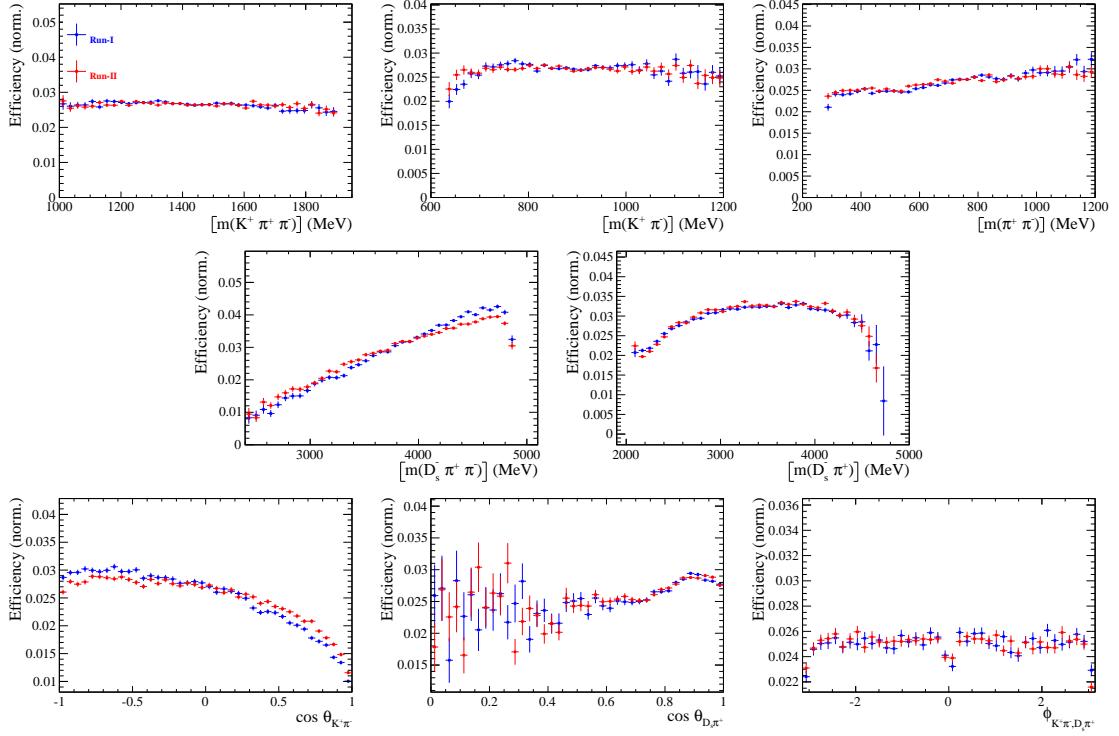
1056 Figure C.1 shows the spline coefficients obtained by fitting the decay-time distribution of  
 1057  $B_s^0 \rightarrow D_s\pi\pi\pi$  data candidates in different subsamples. Sufficient agreement is observed  
 1058 within a given data-taking period, while the acceptance shapes for Run-I and Run-II  
 1059 data differ significantly. The fitted splines for the different  $D_s$  final states are in a good  
 1060 agreement. The largest deviations are observed between the different L0 categories.



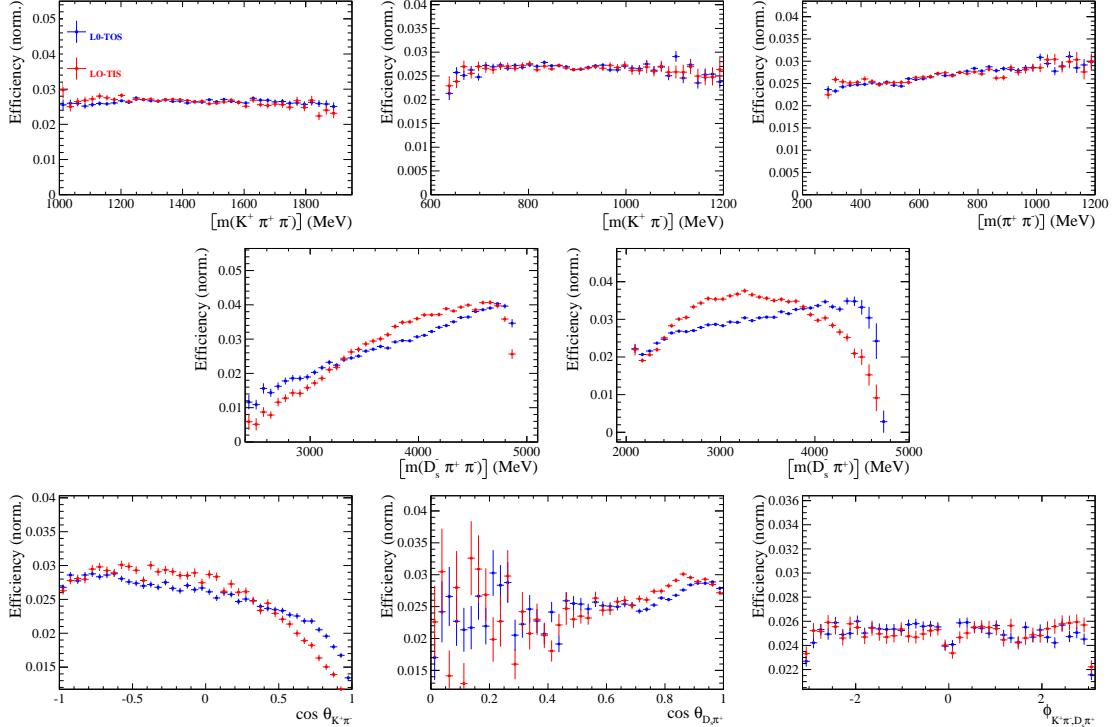
**Figure C.1:** Comparison of the spline coefficients (point with error bars) obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  decay-time for different subsamples: (top-left) different years of data-taking; (top-right) different data-taking periods; (bottom-left) different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

1061 **F Comparison of phase-space acceptance in subsam-  
 1062 ples**

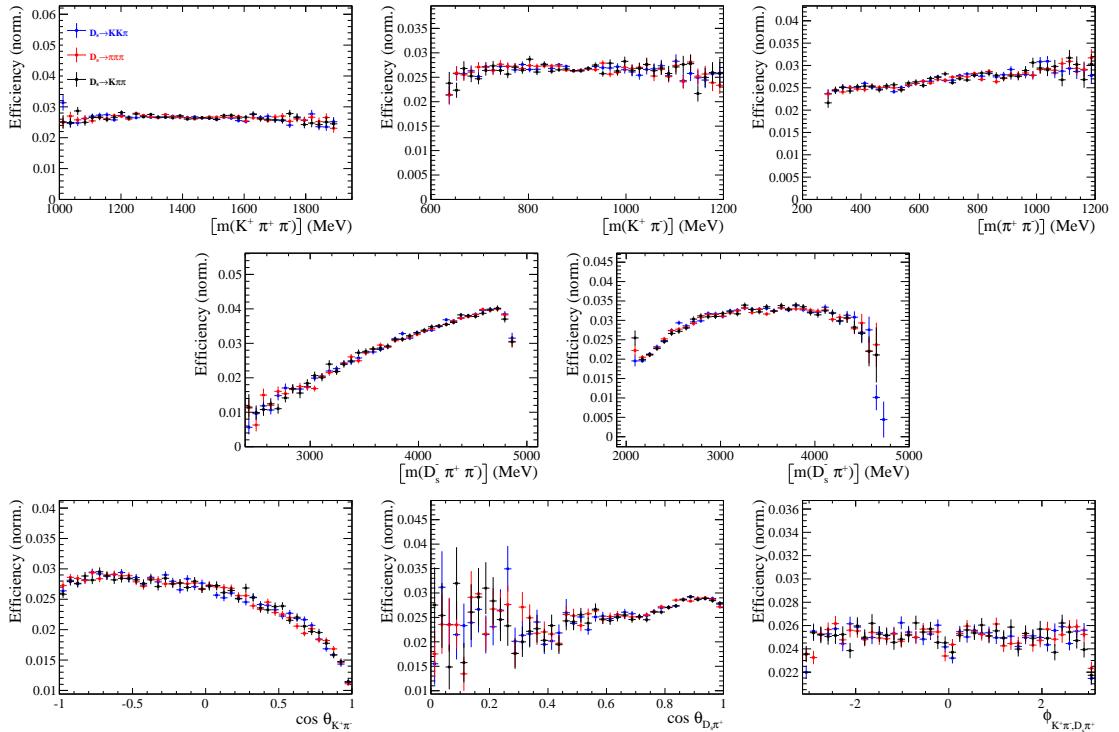
1063 Figures C.1, C.2 and C.3 compare the phase space-acceptance projections obtained from  
 1064  $B_s \rightarrow D_s K\pi\pi$  MC in different subsamples. Sufficient agreement is observed between  
 1065 different data-taking periods and  $D_s$  final states. The largest deviations are observed  
 1066 between the different L0 categories.



**Figure C.1:** Comparison of the phase space acceptance for different data-taking periods. A  $\chi^2$ -test between the samples yields  $\chi^2/\nu = 1.10$  (with  $\nu = 533$ ) using an adaptive 5D binning.



**Figure C.2:** Comparison of the phase space acceptance for different trigger categories. A  $\chi^2$ -test between the samples yields  $\chi^2/\nu = 1.62$  (with  $\nu = 1211$ ) using an adaptive 5D binning.



**Figure C.3:** Comparison of the phase space acceptance for different  $D_s$  final states. A  $\chi^2$ -test using an adaptive 5D binning between the  $D_s \rightarrow KK\pi$  and  $D_s \rightarrow K\pi\pi$  samples yields  $\chi^2/\nu = 1.01$  (with  $\nu = 728$ ),  $\chi^2/\nu = 0.96$  (with  $\nu = 988$ ) between  $D_s \rightarrow KK\pi$  and  $D_s \rightarrow \pi\pi\pi$  and  $\chi^2/\nu = 1.00$  (with  $\nu = 728$ ) between  $D_s \rightarrow \pi\pi\pi$  and  $D_s \rightarrow K\pi\pi$ .

## G OS tagger calibration parameters

tagger	$\langle \eta \rangle$	$p_0 - \langle \eta \rangle$	$p_1$	$\rho(p_0, p_1)$
OS $\mu$	0.30	$0.010 \pm 0.023$	$1.02 \pm 0.26$	0.03
OS $e$	0.29	$0.042 \pm 0.036$	$1.87 \pm 0.59$	0.08
OS $K$	0.42	$0.020 \pm 0.010$	$1.22 \pm 0.15$	0.03
OS Vtx charge	0.38	$-0.011 \pm 0.015$	$1.05 \pm 0.23$	-0.01

**Table 7.1:** Calibration parameters of the OS taggers for Run-I.

tagger	$\langle \eta \rangle$	$p_0 - \langle \eta \rangle$	$p_1$	$\rho(p_0, p_1)$
OS $\mu$	0.33	$-0.001 \pm 0.014$	$1.24 \pm 0.21$	0.06
OS $e$	0.36	$0.014 \pm 0.020$	$1.16 \pm 0.27$	0.06
OS $K$	0.40	$-0.011 \pm 0.010$	$1.51 \pm 0.21$	0.03
OS Vtx charge	0.39	$-0.011 \pm 0.010$	$1.25 \pm 0.15$	0.03
OS charm	0.36	$-0.030 \pm 0.019$	$0.96 \pm 0.37$	0.04

**Table 7.2:** Calibration parameters of the OS taggers for Run-II.

## 1068 H Spin Amplitudes

1069 The spin factors used for  $B \rightarrow P_1 P_2 P_3 P_4$  decays are given in Table 8.1.

**Table 8.1:** 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	$B[D] \rightarrow (T_1 T_2)$ , $T_1 \rightarrow (P_1 P_2)$ , $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\gamma}(T_1) L_{(2)\gamma}^\beta(T_2)$

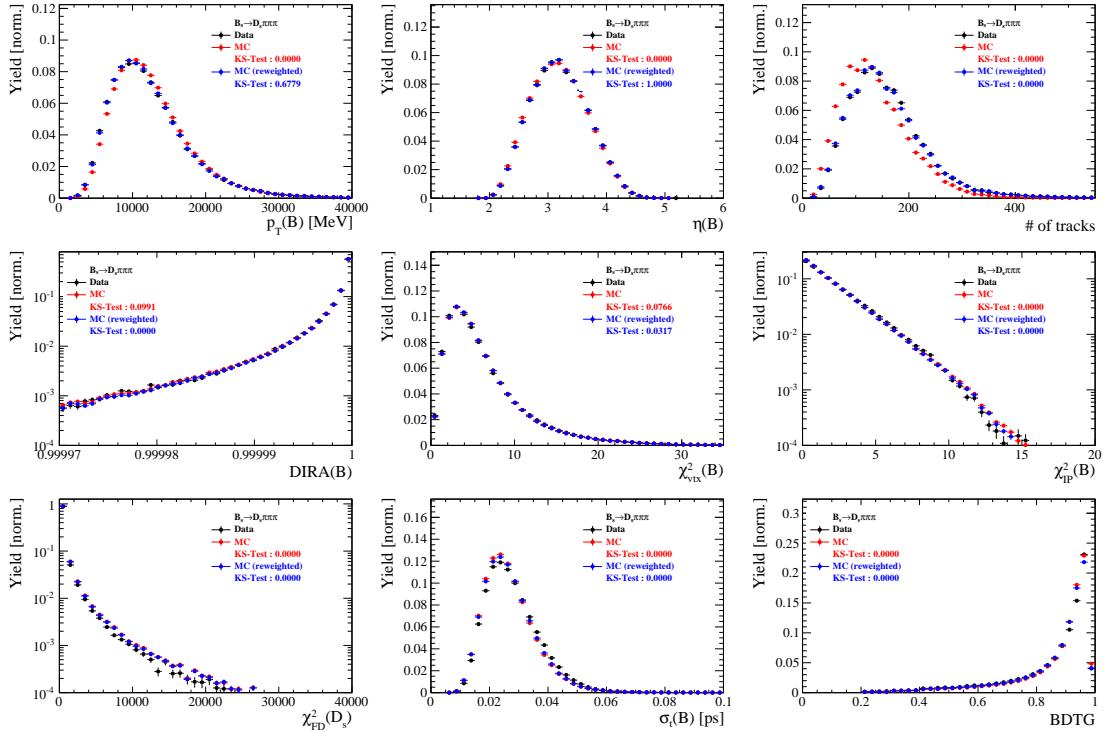
# I Considered Decay Chains

The various decay channels considered in the model building are listed in Table 9.1.

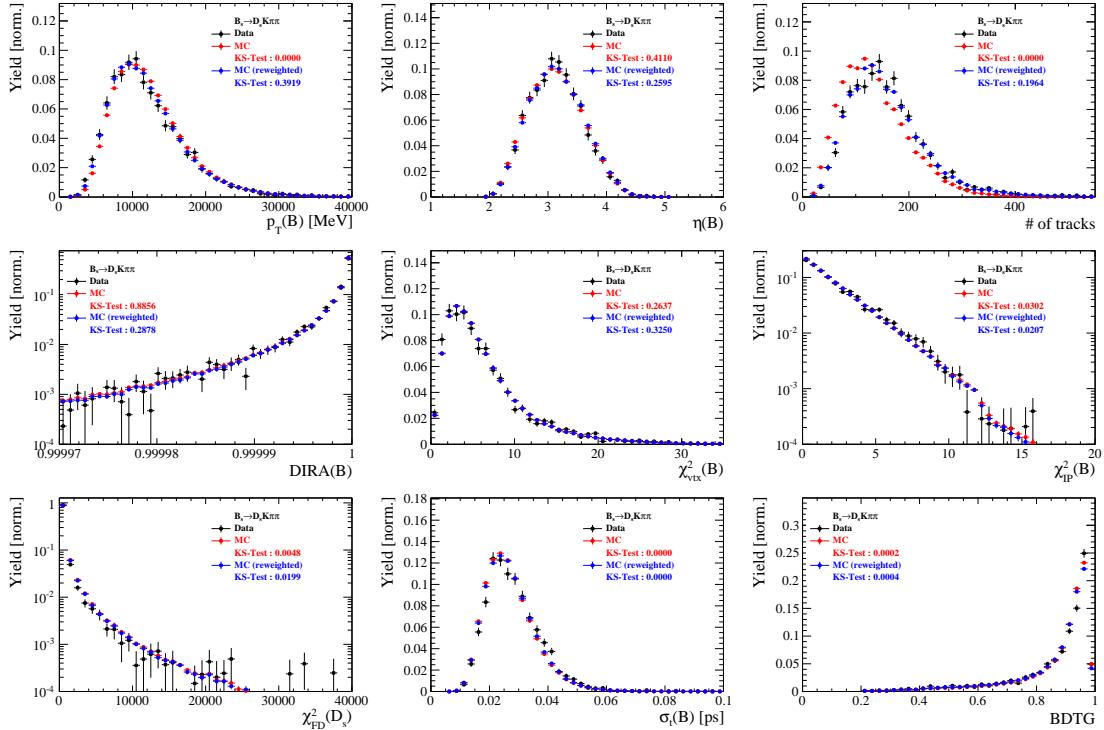
**Table 9.1:** Decays considered in the LASSO model building.

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

## J Data-simulation comparisson



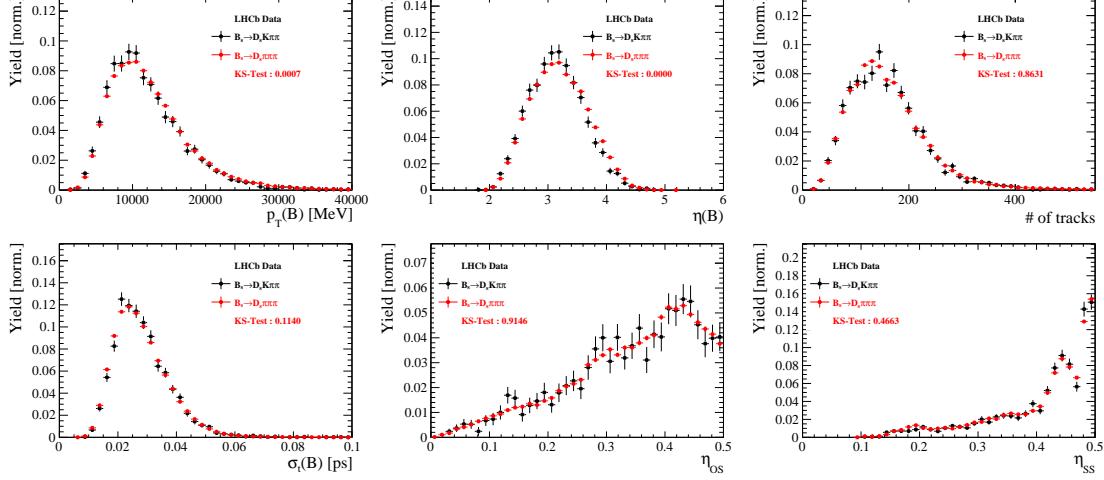
**Figure C.1:** Comparison between data and MC of selected variables for  $B_s \rightarrow D_s \pi\pi\pi$  decays.



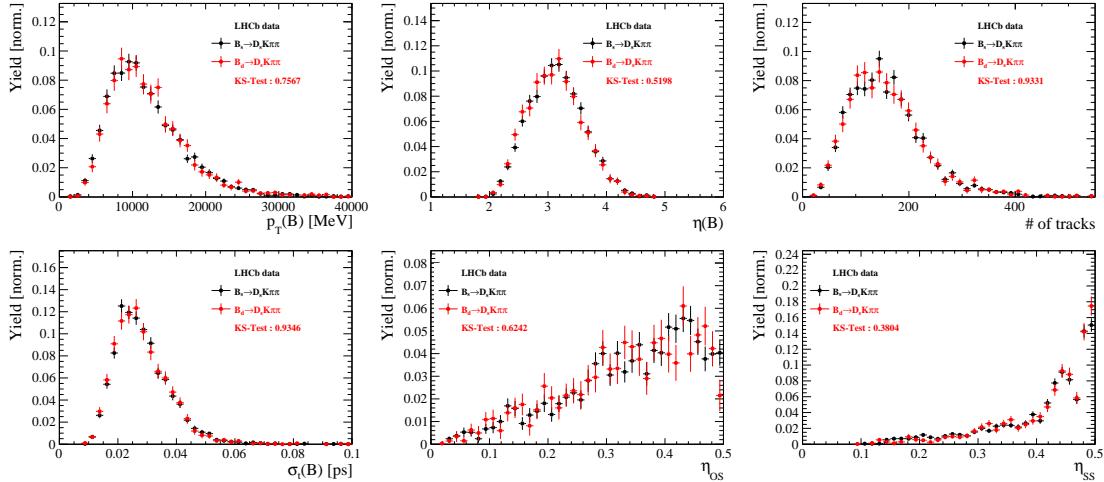
**Figure C.2:** Comparison between data and MC of selected variables for  $B_s \rightarrow D_s K\pi\pi$  decays.

1073 **K Data distributions**

1074 **K.1 Comparison of signal and calibration channels**

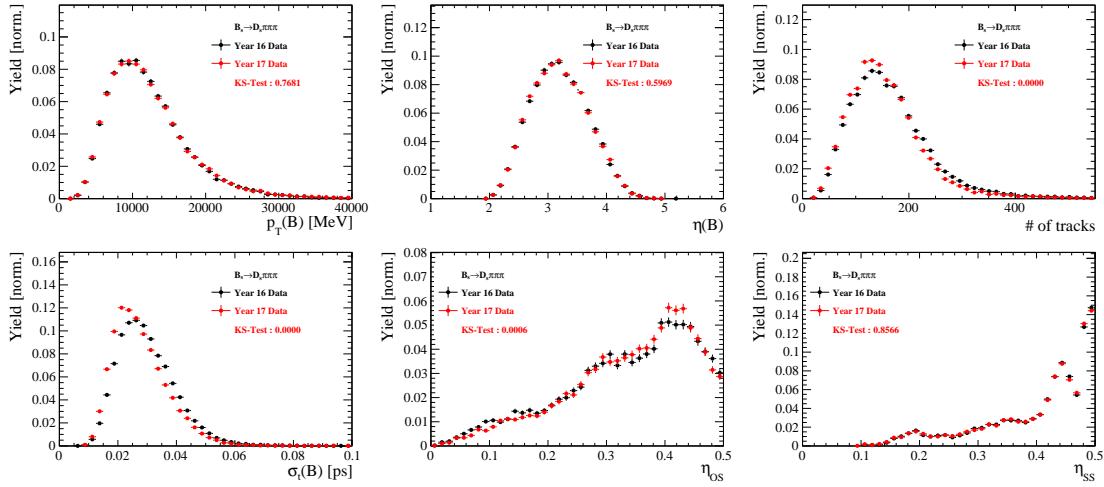


**Figure C.1:** Comparison between  $B_s \rightarrow D_s K\pi\pi$  and  $B_s \rightarrow D_s \pi\pi\pi$  decays for selected variables.

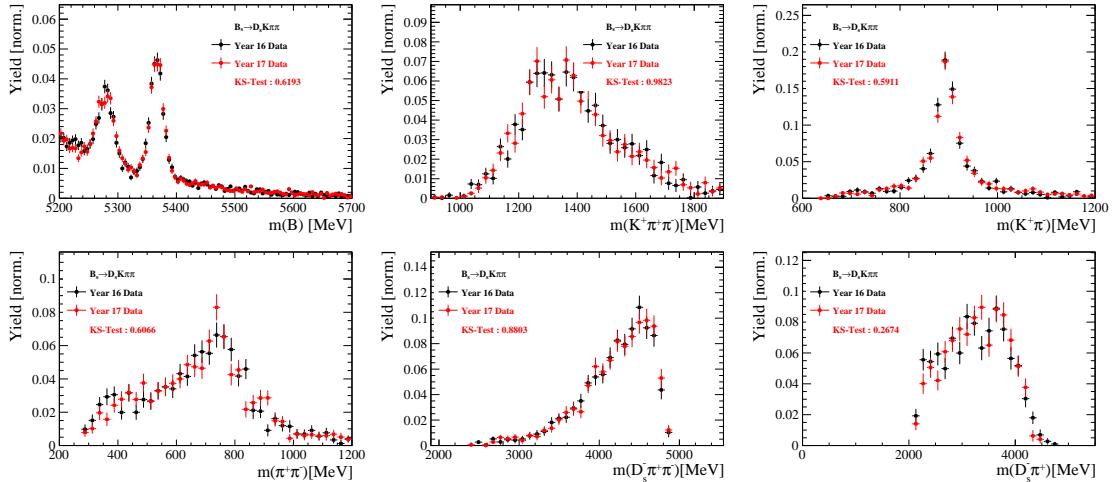


**Figure C.2:** Comparison between  $B_s \rightarrow D_s K\pi\pi$  and  $B_d \rightarrow D_s K\pi\pi$  decays for selected variables.

## K.2 Comparison of data taken in 2016 and 2017

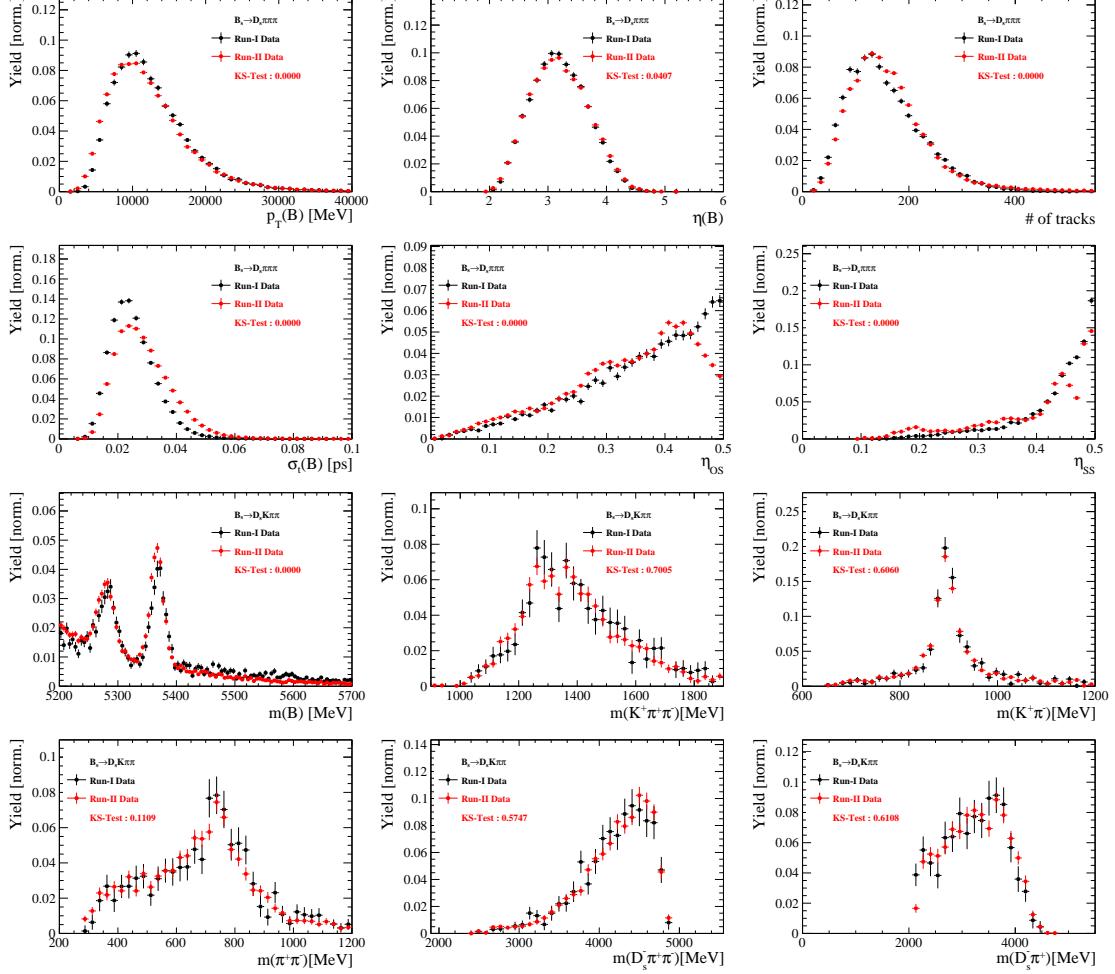


**Figure C.3:** Comparison of selected variables for  $B_s \rightarrow D_s \pi\pi\pi$  data taken in 2016 and 2017.



**Figure C.4:** Comparison of selected variables for  $B_s \rightarrow D_s K\pi\pi$  data taken in 2016 and 2017.

1076 K.3 Comparison of Run-I and Run-II data



**Figure C.5:** Comparison of selected variables for Run-I and Run-II data.

1077 K.4 Comparison of  $D_s$  final states

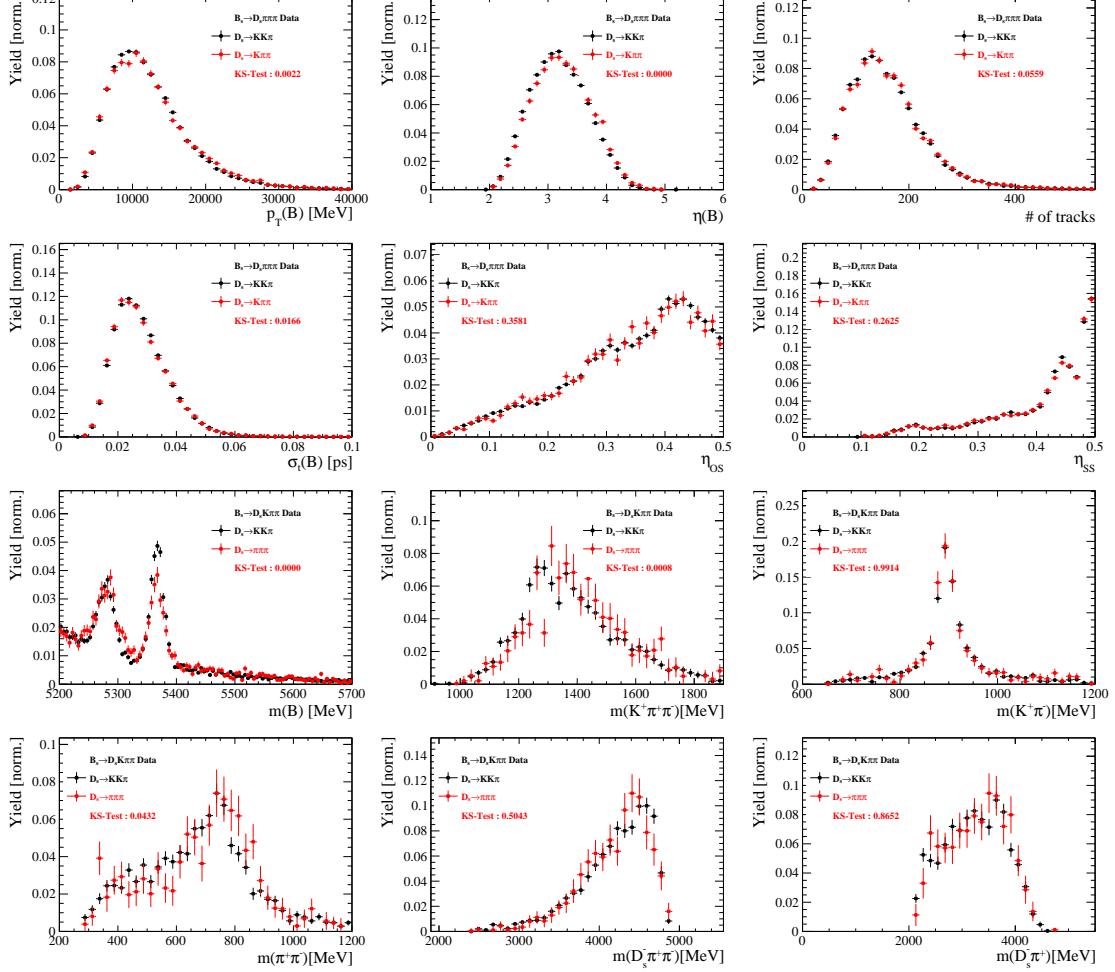
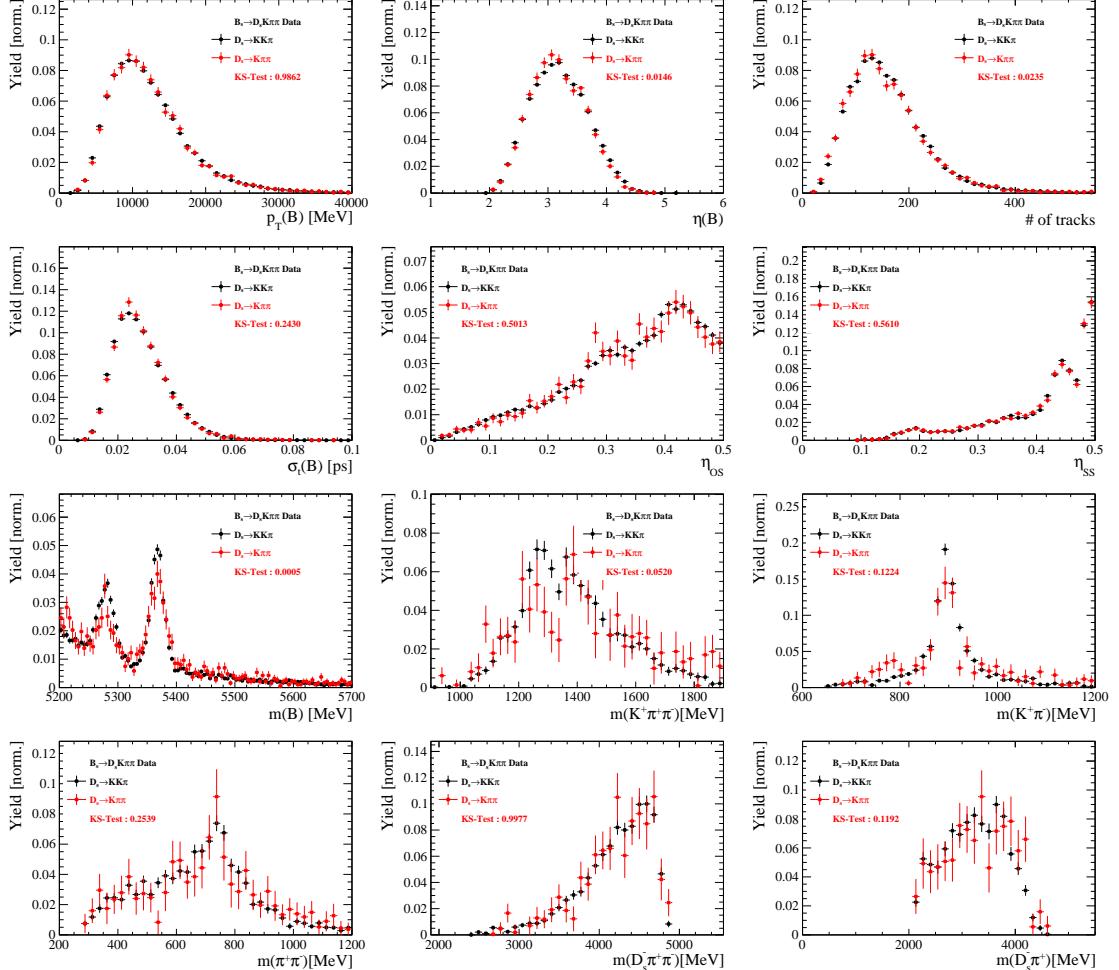
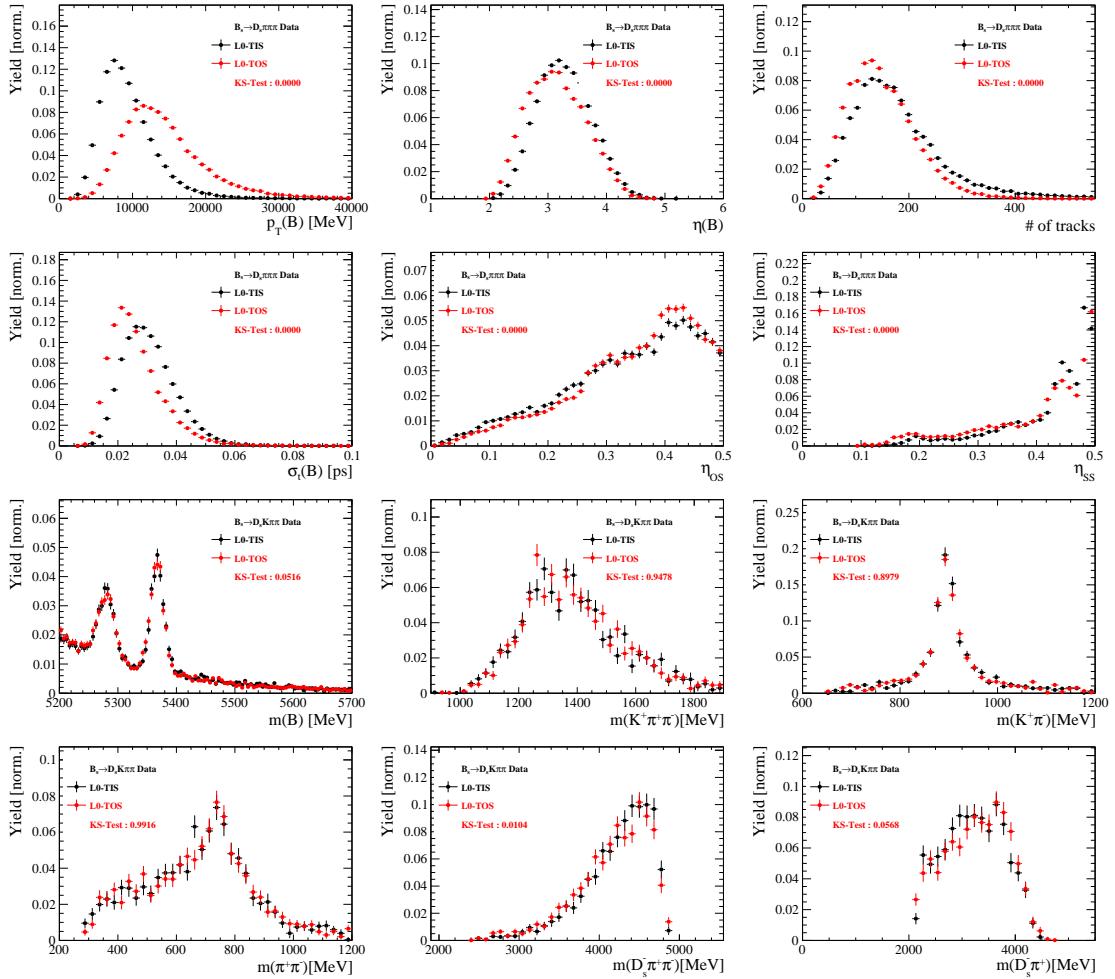


Figure C.6: Comparison of selected variables for different  $D_s$  final states.



**Figure C.7:** Comparison of selected variables for different  $D_s$  final states.

## K.5 Comparison of trigger categories



**Figure C.8:** Comparison of selected variables for different trigger categories.

1079 **References**

- 1080 [1] S. Blusk, *First observations and measurements of the branching fractions for the*  
1081 *decays  $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$  and  $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ ,* .
- 1082 [2] LHCb, S. Blusk, *Measurement of the CP observables in  $\bar{B}_s^0 \rightarrow D_s^+ K^-$  and first obser-*  
1083 *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](#).
- 1084 [3] R. Fleischer, *New strategies to obtain insights into CP violation through  $B_s \rightarrow$*   
1085  *$D_s^\pm K^\mp, D_s^{*\pm} K^\mp, \dots$  and  $B_d \rightarrow D^\pm \pi^\mp, D^{*\pm} \pi^\mp, \dots$  decays,* Nucl. Phys. **B671** (2003) 459,  
1086 [arXiv:hep-ph/0304027](#).
- 1087 [4] K. De Bruyn *et al.*, *Exploring  $B_s \rightarrow D_s^{(*)\pm} K^\mp$  Decays in the Presence of a Sizable*  
1088 *Width Difference  $\Delta\Gamma_s$ ,* Nucl. Phys. **B868** (2013) 351, [arXiv:1208.6463](#).
- 1089 [5] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 1090 [6] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation*  
1091 *of production processes*, Annals of Physics **18** (1962), no. 2 198 .
- 1092 [7] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*,  
1093 Phys. Rev. D **11** (1975) 3165.
- 1094 [8] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of  
1095 Physics **108** (1977), no. 2 454 .
- 1096 [9] P. d'Argent *et al.*, *Amplitude Analyses of  $D^0 \rightarrow \pi^+ \pi^- \pi^+ \pi^-$  and  $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$*   
1097 *Decays*, JHEP **05** (2017) 143, [arXiv:1703.08505](#).
- 1098 [10] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay*  
1099 *widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.
- 1100 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo  
1101 Cimento Series 10 **34** (1964), no. 6 1644.
- 1102 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys.  
1103 **C40** (2016), no. 10 100001.
- 1104 [13] D. V. Bugg, *The mass of the  $\sigma$  pole*, Journal of Physics G Nuclear Physics **34** (2007)  
1105 151, [arXiv:hep-ph/0608081](#).
- 1106 [14] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-*  
1107 *dominance prediction for  $\rho \rightarrow e^+ e^-$* , Phys. Rev. Lett. **21** (1968) 244.
- 1108 [15] M. S. et al. *Search for CP violation in the  $D^0 \rightarrow KK\pi\pi$  decay through a full*  
1109 *amplitude analysis*, LHCb-ANA-2017-064.
- 1110 [16] S. M. Flatté, *Coupled-channel analysis of the  $\pi\eta$  and  $KK$  systems near  $KK$  threshold*,  
1111 Physics Letters B **63** (1976), no. 2 224 .
- 1112 [17] BES Collaboration, M. Ablikim *et al.*, *Resonances in  $J/\psi \rightarrow \phi \pi^+ \pi^-$  and  $\phi K^+ K^-$* ,  
1113 Phys. Lett. **B607** (2005) 243, [arXiv:hep-ex/0411001](#).

- <sub>1114</sub> [18] W. Dunwoodie, Fits to  $K\pi$   $I = \frac{1}{2}$   $S$ -wave amplitude and phase data.
- <sub>1115</sub> [19] D. Aston *et al.*, *A Study of  $K^-\pi^+$  Scattering in the Reaction  $K^-p \rightarrow K^-\pi^+n$  at  $11\text{GeV}/c$* , Nucl. Phys. **B296** (1988) 493.
- <sub>1117</sub> [20] BaBar, B. Aubert *et al.*, *Dalitz-plot analysis of the decays  $B^\pm \rightarrow K^\pm\pi^\mp\pi^\pm$* , Phys. Rev. **D72** (2005) 072003, arXiv:hep-ex/0507004, [Erratum: Phys. Rev.D74,099903(2006)].
- <sub>1120</sub> [21] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in  $D^0 \rightarrow K^\mp\pi^\pm\pi^\pm\pi^\mp$  decays*, Eur. Phys. J. **C78** (2018), no. 6 443, arXiv:1712.08609.
- <sub>1122</sub> [22] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- <sub>1123</sub> [23] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys. Rev. **60** (1941) 61.
- <sub>1125</sub> [24] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys. Rev. D **57** (1998) 431.
- <sub>1127</sub> [25] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of  $\psi$  decay to mesons*, Eur. Phys. J. **A16** (2003) 537, arXiv:hep-ph/0211457.
- <sub>1129</sub> [26] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. **D51** (1995) 2247.
- <sub>1131</sub> [27] J.-J. Zhu, *Explicit expressions of spin wave functions*, arXiv:hep-ph/9906250.
- <sub>1132</sub> [28] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput. Phys. Commun. **180** (2009) 1847, arXiv:0805.2956.
- <sub>1134</sub> [29] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- <sub>1136</sub> [30] M. Karbach and M. Kenzie, *Gammacombo package*, <http://gammacombo.hepforge.org/web/HTML/index.html>, 2014.
- <sub>1138</sub> [31] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.
- <sub>1140</sub> [32] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.
- <sub>1142</sub> [33] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- <sub>1144</sub> [34] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.
- <sub>1146</sub> [35] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in  $B_s^0 \rightarrow D_s^\mp K^\pm$  decays*, Submitted to: JHEP (2017) arXiv:1712.07428.

- 1148 [36] LHCb collaboration, L. Zhang, *Measurements of CP violation in  $B_s^0 \rightarrow J/\psi K^+ K^-$*   
 1149 *decays in the low  $K^+ K^-$  mass range with 13 TeV data*, LHCb-ANA-2017-028.
- 1150 [37] A. Poluektov, *Correction of simulated particle identification response in LHCb using*  
 1151 *kernel density estimation*, LHCb-INT-2017-007.
- 1152 [38] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and*  
 1153  *$\tau$ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots  
 1154 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 1155 [39] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*  
 1156 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 1157 [40] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the*  
 1158 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- 1159 [41] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of  $B_s^0$  mesons at*  
 1160 *LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- 1161 [42] LHCb, R. Aaij *et al.*, *Measurement of  $B^0$ ,  $B_s^0$ ,  $B^+$  and  $\Lambda_b^0$  production asymmetries in 7*  
 1162 *and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- 1163 [43] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of*  
 1164 *the  $K\pi$  Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-  
 1165 2012-027, CERN, Geneva, Feb, 2013.
- 1166 [44] A. Davis *et al.*, *Measurement of the instrumental asymmetry for  $K^-\pi^+$ -pairs at LHCb*  
 1167 *in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN,  
 1168 Geneva, Mar, 2018.
- 1169 [45] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly*  
 1170 *forbidden transitions in  $D \rightarrow K(S)$ ,  $K(L) + \pi$ 's decays*, Phys. Lett. **B349** (1995)  
 1171 363, arXiv:hep-ph/9502238.
- 1172 [46] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*,  
 1173 Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- 1174 [47] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal  
 1175 Statistical Society, Series B **58** (1994) 267.
- 1176 [48] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude*  
 1177 *analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- 1178 [49] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- 1179 [50] CLEO Collaboration, M. Artuso *et al.*, *Amplitude analysis of  $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$* ,  
 1180 Phys. Rev. **D85** (2012) 122002, arXiv:1201.5716.
- 1181 [51] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime*  
 1182 *and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986,  
 1183 DESY-F31-86-02.

- <sub>1184</sub> [52] D. Hill, M. John, and P. Gandini, *A study of partially reconstructed  $B^\pm \rightarrow D^{*0} h^\pm$  decays using the  $D^0 \rightarrow K\pi, KK, \pi\pi$  final states*, .
- <sub>1186</sub> [53] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins University Press, Baltimore, MD, USA, 1996.
- <sub>1188</sub> [54] L. collaboration, *Precision measurement of the  $B_s$  oscillation frequency with the decay  $B_s \rightarrow D_s\pi$* , LHCb-ANA-2012-053.
- <sub>1190</sub> [55] BaBar, P. del Amo Sanchez *et al.*, *Measurement of  $D0$ -anti $D0$  mixing parameters using  $D0 \rightarrow K(S)0\pi^+\pi^-$  and  $D0 \rightarrow K(S)0K^+K^-$  decays*, Phys. Rev. Lett. **105** (2010) 081803, arXiv:1004.5053.
- <sub>1193</sub> [56] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in  $D^0 \rightarrow K_S^0 K^\pm \pi^\mp$  decays*, Phys. Rev. **D93** (2016), no. 5 052018, arXiv:1509.06628.