

# 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
  - Changed  $D^0$  bkg vetoes (Sec. 4.2.2)
  - Revisited some systematics
    - \* Use uncertainty from fake rate instead of fixing misID yield to 0 or twice the value (Sec. 12.2)
    - \* Use GLASS model instead of RBW as alternative to LASS model for  $K\pi$  S-wave (Sec. 12.9)
    - \* 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. 12.9)
    - \* Calculate RMS of lineshape variations instead of summing them up in quadrature (Sec. 12.9)
    - \* Bug fix in evaluation of fixed lineshape parameters (mass, width, form-factor) systematic (Sec. 12.9)
  - Added tables with alternative amplitude models (Sec. 12.10)
  - Added comparison plots of 16 vs 17 data (Sec. M)
- Version 3.0:
  - Changed  $D^0$  bkg vetoes back to the ones from v1.0 to account for additional Kaon misID (Fig. 4.3)
  - Updated analysis with new MC samples (Sec. 4.3)
  - Studied time resolution separately for 16 and 17 data (Sec. 6.2)
  - Changed phasespace acceptance method from BDT to MC integration (Sec. 7.3)
  - Updated analysis with latest OS Tagger tuning (Sec. 8.1)
- Version 3.1: Minor changes, approved by Mark and Alessandro
- Version 3.2: Updated fit validation of phasespace-integrated and full fit using the same data set (Sec. 2.4)
- Version 4.0: Implemented first round of comments from Marta and Rafael
  - Added a small section summarizing the main analysis steps (Sec. 3)
  - Small improvements to the formalism (Sec. 2) and additional information on used lineshapes (Sec. A)
  - Additional information on fit stability and interference fractions (Sec. K)
  - Added additional systematic studies for the phase-space acceptance (Sec. 12.8)
  - Added rough estimate of sensitivity to physical observables from phase-space integrated fit (Sec. 10.2)
  - Added comparison of phase-space integrated and full time-dependent amplitude fit (Sec. 13)

- Version 4.1: Implemented second round of comments from Marta and Rafael
  - Added systematics for fit fractions (Table 11.3)
  - Added additional information to tagging calibration (Sec. 8.1) and detection asymmetry sections (Sec. 9.2)
- Version 5.0: Unblinded results
- Version 6.0:
  - Added studies of the decay-time bias (Sec. 6.3) and related systematic (Sec. 12.4)
  - Added stat. and syst. correlation matrix for  $CP$  coefficients (Tables 10.3,10.4)
  - New closure test for  $\Delta m_s$  (Sec. 12.1)
  - Added systematic study for ignoring correlation between reconstructed mass and decay-time (Sec. 12.2)
  - The assumption that time- and phasespace acceptance factorize is tested (Sec. 12.8)

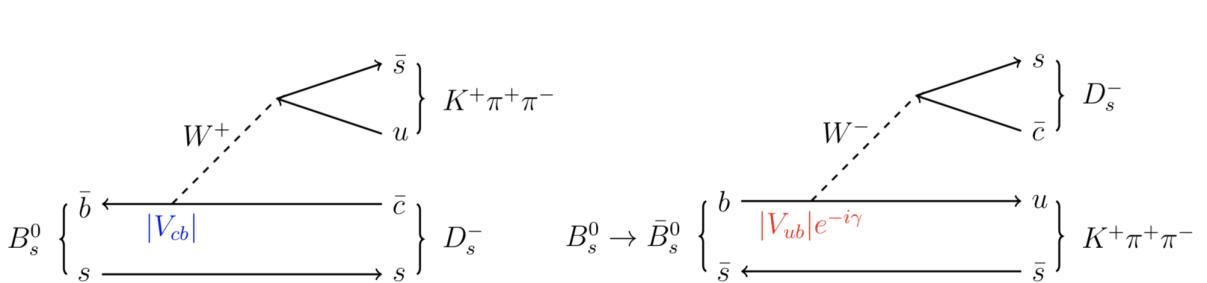
# Contents

<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 . . . . .	5
2.3	Fit implementation . . . . .	7
2.4	Validation . . . . .	8
<b>3</b>	<b>Analysis strategy</b>	<b>11</b>
<b>4</b>	<b>Data samples and event selection</b>	<b>11</b>
4.1	Stripping and Trigger selection . . . . .	11
4.2	Offline selection . . . . .	11
4.2.1	Phase space region . . . . .	12
4.2.2	Physics background vetoes . . . . .	13
4.2.3	Training of multivariate classifier . . . . .	16
4.2.4	Final selection . . . . .	17
4.3	Simulation . . . . .	22
<b>5</b>	<b>Yields determination</b>	<b>23</b>
5.1	Signal model . . . . .	23
5.2	Background models . . . . .	24
5.3	Results . . . . .	26
<b>6</b>	<b>Decay-time Resolution</b>	<b>27</b>
6.1	Calibration for Run-I data . . . . .	28
6.2	Calibration for Run-II data . . . . .	29
6.3	Decay-time bias . . . . .	32
<b>7</b>	<b>Acceptance</b>	<b>34</b>
7.1	MC corrections . . . . .	34
7.1.1	Truth matching of simulated candidates . . . . .	34
7.1.2	Correction of data-simulation differences . . . . .	34
7.2	Decay-time acceptance . . . . .	35
7.3	Phase space acceptance . . . . .	39
<b>8</b>	<b>Flavour Tagging</b>	<b>41</b>
8.1	OS tagger combination . . . . .	42
8.2	Tagging performance . . . . .	44
<b>9</b>	<b>Production and Detection Asymmetries</b>	<b>45</b>
9.1	$B_s$ Production Asymmetry . . . . .	45
9.2	$K^-\pi^+$ Detection Asymmetry . . . . .	46

<b>10 Decay-time fit</b>	<b>49</b>
10.1 Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data . . . . .	49
10.2 Fit to $B_s^0 \rightarrow D_sK\pi\pi$ data . . . . .	49
<b>11 Time-dependent amplitude fit</b>	<b>53</b>
11.1 Signal Model Construction . . . . .	53
11.2 Results . . . . .	57
<b>12 Systematic uncertainties</b>	<b>60</b>
12.1 Fit bias . . . . .	60
12.2 Background subtraction . . . . .	60
12.3 Decay-time acceptance . . . . .	61
12.4 Decay-time resolution and tagging . . . . .	62
12.5 Production, detection asymmetries and mixing frequency . . . . .	63
12.6 Multiple candidates . . . . .	63
12.7 Length and momentum scales . . . . .	63
12.8 Phase space acceptance . . . . .	63
12.9 Resonance description . . . . .	64
12.10 Alternative amplitude models . . . . .	65
<b>13 Summary</b>	<b>71</b>
<b>A Parametrization of Amplitude Lineshapes</b>	<b>72</b>
<b>B Stripping and Trigger cuts</b>	<b>75</b>
<b>C Details of multivariate classifier</b>	<b>77</b>
<b>D Detailed mass fits</b>	<b>82</b>
<b>E Decay-time Resolution fits</b>	<b>87</b>
<b>F Comparison of time-acceptance in subsamples</b>	<b>91</b>
<b>G Comparison of phase-space acceptance in subsamples</b>	<b>91</b>
<b>H OS tagger calibration parameters</b>	<b>94</b>
<b>I Spin Amplitudes</b>	<b>95</b>
<b>J Considered Decay Chains</b>	<b>96</b>
<b>K Additional information for the time-dependent amplitude fit</b>	<b>97</b>
<b>L Data-simulation comparisson</b>	<b>100</b>
<b>M Data distributions</b>	<b>101</b>
<b>References</b>	<b>107</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 [1–3]. 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 [4,5]. 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, in which case the procedure is very similar to the analysis of  $B_s^0 \rightarrow D_s K$  decays [6,7]. Both approaches are explored in this note. 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.

## 2 Formalism

### 2.1 Decay rates and $CP$ -observables

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 given by:

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

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

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

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

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

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

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

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, the four-body phase space density function is not flat in the usual kinematic variables. An analytic expression for  $\phi_4$  can be found in Ref. [8], but is not needed in practice, see Sec. 7.3.

49 Assuming there is no direct  $CP$  violation in the  $B_s$  decay implies for the  $CP$  conjugate  
50 transition amplitudes:

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

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

51 where the  $CP$ -conjugate phase space point  $\bar{\mathbf{x}}$  is defined such that it is mapped onto  
52  $\mathbf{x}$  by the interchange of final state charges, and the reversal of three-momenta. The  
53 phenomenological description of the intermediate state amplitudes is discussed in Sec. 2.2.

54

55 Rather than developing an amplitude model describing the whole phase space, it is also  
56 possible to determine the CKM angle  $\gamma$ , in a model-independent way, from the decay-time  
57 distribution only. The differential decay rate integrated over the phase space yields:

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

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

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

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

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

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

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

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

61 The five  $CP$  coefficients  $C, D \equiv D_f, \bar{D} \equiv D_{\bar{f}}, S \equiv S_f, \bar{S} \equiv S_{\bar{f}}$  can be extracted from  
62 a fit to the decay-time distribution which allows to infer the four physical observables  
63  $r, \delta, \gamma, \kappa$  as discussed in more detail in Sect. 2.4. This approach will be called phase-space  
64 integrated fit in the following. In contrast, the approach using Eq. 2.1 will be referred  
65 to as full time-dependent amplitude fit. In case of the phase-space integrated fit, the  
66 coherence factor dilutes the sensitivity to the weak phase  $\gamma$  due to the integration over  
67 the interfering amplitudes across the phase space. The value of  $\kappa$  is bounded between  
68 zero and unity. The latter corresponds to the limit of only one contributing intermediate  
69 state in which case the same sensitivity as in  $B_s \rightarrow D_s K$  decays is reached, while  $\kappa = 0$   
70 would result in no sensitivity to  $\gamma$  at all.

71 **2.2 Amplitude model**

72 To construct the intermediate state amplitudes  $A_i(\mathbf{x})$ , the isobar approach is used, which  
 73 assumes that the decay process can be factorized into subsequent two-body decay am-  
 74 plitudes [9–11]. This gives rise to two different decay topologies; quasi two-body decays  
 75  $B_s \rightarrow (R_1 \rightarrow h_1 h_2) (R_2 \rightarrow h_3 h_4)$  or cascade decays  $B_s \rightarrow h_1 [R_1 \rightarrow h_2 (R_2 \rightarrow h_3 h_4)]$ . In  
 76 either case, the intermediate state amplitude is parameterized as a product of orbital  
 77 angular momentum,  $L$ , dependent form factors  $B_L$ , included for each vertex of the decay  
 78 tree, Breit-Wigner propagators  $T_R$ , included for each resonance  $R$ , and an overall angular  
 79 distribution represented by a spin factor  $S$ ,

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

80 The following description of the individual components is adapted from Ref. [12] and  
 81 only included for completeness.

82 **2.2.1 Form Factors and Resonance Lineshapes**

83 To account for the finite size of the decaying resonances, the Blatt-Weisskopf penetration  
 84 factors, derived in Ref. [13] by assuming a square well interaction potential with radius  
 85  $r_{\text{BW}}$ , are used as form factors,  $B_L$ . They depend on the breakup momentum  $q$ , which is  
 86 defined as the magnitude of the three-momentum of one of the daughters in the rest frame  
 87 of the decaying resonance and the orbital angular momentum  $L$ , between the resonance  
 88 daughters. Their explicit expressions are

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

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

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

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

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

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

95 where  $q_0$  is the value of the breakup momentum at the resonance pole [14].

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

Equation (2.16) is used by default for resonances decaying into a two-body final state. For the  $K\pi$  and  $\pi\pi$  S-wave resonances as well as for the  $\rho(770)$  meson more sophisticated parameterizations are employed. Their explicit expressions are given in Appendix A. The nominal masses and widths of the resonances are taken from the PDG [15] with the exceptions described in Appendix K. The running width distributions for various 3-body resonances, calculated from Eq. 2.17, are shown in Appendix A.

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.

### 2.2.2 Spin Densities

The spin amplitudes are phenomenological descriptions of decay processes that are required to be Lorentz invariant, compatible with angular momentum conservation and, where appropriate, parity conservation. They are constructed in the covariant Zemach (Rarita-Schwinger) tensor formalism [16–18]. At this point, we briefly introduce the fundamental objects of the covariant tensor formalism which connect the particle’s four-momenta to the spin dynamics of the reaction and give a general recipe to calculate the spin factors for arbitrary decay trees. Further details can be found in Refs. [19, 20].

A spin- $S$  particle with four-momentum  $p$ , and spin projection  $\lambda$ , is represented by the polarization tensor  $\epsilon_{(S)}(p, \lambda)$ , which is symmetric, traceless and orthogonal to  $p$ . These so-called Rarita-Schwinger conditions reduce the a priori  $4^S$  elements of the rank- $S$  tensor to  $2S + 1$  independent elements in accordance with the number of degrees of freedom of a spin- $S$  state [17, 21].

The spin projection operator  $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$ , for a resonance  $R$ , with spin  $S = \{0, 1, 2\}$ , and four-momentum  $p_R$ , is given by [20]:

$$\begin{aligned} P_{(0)}^{\mu\nu}(p_R) &= 1 \\ P_{(1)}^{\mu\nu}(p_R) &= -g^{\mu\nu} + \frac{p_R^\mu p_R^\nu}{p_R^2} \\ P_{(2)}^{\mu\nu\alpha\beta}(p_R) &= \frac{1}{2} \left[ P_{(1)}^{\mu\alpha}(p_R) P_{(1)}^{\nu\beta}(p_R) + P_{(1)}^{\mu\beta}(p_R) P_{(1)}^{\nu\alpha}(p_R) \right] - \frac{1}{3} P_{(1)}^{\mu\nu}(p_R) P_{(1)}^{\alpha\beta}(p_R), \end{aligned} \quad (2.19)$$

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

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

$$L_{(L)\mu_1\dots\mu_L}(p_R, q_R) = (-1)^L P_{(L)\mu_1\dots\mu_L\nu_1\dots\nu_L}(p_R) q_R^{\nu_1} \dots q_R^{\nu_L}. \quad (2.20)$$

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

$$\begin{aligned} \varepsilon_{(S)}(R) &\equiv \varepsilon_{(S)}(p_R, \lambda_R), \\ P_{(S)}(R) &\equiv P_{(S)}(p_R), \\ L_{(L)}(R) &\equiv L_{(L)}(p_R, q_R). \end{aligned} \quad (2.21)$$

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

$$\begin{aligned} S_{R \rightarrow AB}(\mathbf{x}|L_{AB}, S_{AB}; \lambda_R, \lambda_A, \lambda_B) &= \varepsilon_{(S_R)}(R) X(S_R, L_{AB}, S_{AB}) L_{(L_{AB})}(R) \\ &\times \Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B), \end{aligned} \quad (2.22)$$

where

$$\Phi(\mathbf{x}|S_{AB}; \lambda_A, \lambda_B) = P_{(S_{AB})}(R) X(S_{AB}, S_A, S_B) \varepsilon_{(S_A)}^*(A) \varepsilon_{(S_B)}^*(B). \quad (2.23)$$

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

$$X(j_a, j_b, j_c) = \begin{cases} 1 & \text{if } j_a + j_b + j_c \text{ even} \\ \varepsilon_{\alpha\beta\gamma\delta} p_R^\delta & \text{if } j_a + j_b + j_c \text{ odd} \end{cases}, \quad (2.24)$$

where  $\varepsilon_{\alpha\beta\gamma\delta}$  is the Levi-Civita symbol and  $j$  refers to the arguments of  $X$  defined in Eqs. 2.22 and 2.23. Its antisymmetric nature ensures the correct parity transformation behavior of the amplitude. The spin factor for a whole decay chain, for example  $R \rightarrow (R_1 \rightarrow AB)(R_2 \rightarrow CD)$ , is obtained by combining the two-body terms and performing a sum over all unobservable, intermediary spin projections

$$\sum_{\lambda_{R_1}, \lambda_{R_2}} S_{R \rightarrow R_1 R_2}(\mathbf{x}|L_{R_1 R_2}; \lambda_{R_1}, \lambda_{R_2}) S_{R_1 \rightarrow AB}(\mathbf{x}|L_{AB}; \lambda_{R_1}) S_{R_2 \rightarrow CD}(\mathbf{x}|L_{CD}; \lambda_{R_2}), \quad (2.25)$$

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 pseudoscalar initial/final states are involved.

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

## <sup>164</sup> 2.3 Fit implementation

<sup>165</sup> The hadronic amplitudes are renormalized prior to the amplitude fit such that

$$\int |A_i(\mathbf{x})|^2 d\Phi_4 = 1. \quad (2.26)$$

<sup>166</sup> This allows us to set more intuitive starting values as the amplitude coefficients are all on  
<sup>167</sup> a comparable scale. Moreover, the total amplitudes  $\mathcal{A}_f^{c(u)}(\mathbf{x})$  are renormalized on-the-fly  
<sup>168</sup> (*i.e.* at each minimization step) such that

$$\begin{aligned} & \int |\mathcal{A}_f^{c(u)}(\mathbf{x})|^2 d\Phi_4 = 1 \\ & \arg \left( \int \mathcal{A}_f^c(\mathbf{x})^* \mathcal{A}_f^u(\mathbf{x}) d\Phi_4 \right) = 0. \end{aligned} \quad (2.27)$$

<sup>169</sup> As a result, the average amplitude ratio and strong phase difference between the  $b \rightarrow u$  and  
<sup>170</sup>  $b \rightarrow c$  transitions can be introduced as direct fit parameters instead of derived quantities  
<sup>171</sup> that have to be calculated from Equation 2.11 after the fit. For the differential decay rate  
<sup>172</sup> follows:

$$\begin{aligned} \frac{d\Gamma(\mathbf{x}, t, q, f)}{e^{-\Gamma_s t} dt d\Phi_4} \propto & (|\mathcal{A}_f^c(\mathbf{x})|^2 + r^2 |\mathcal{A}_f^u(\mathbf{x})|^2) \cosh \left( \frac{\Delta\Gamma_s t}{2} \right) \\ & + q_t q_f (|\mathcal{A}_f^c(\mathbf{x})|^2 - r^2 |\mathcal{A}_f^u(\mathbf{x})|^2) \cos (\Delta m_s t) \\ & - 2 r \text{Re} (\mathcal{A}_f^c(\mathbf{x})^* \mathcal{A}_f^u(\mathbf{x}) e^{i\delta - iq_f(\gamma - 2\beta_s)}) \sinh \left( \frac{\Delta\Gamma_s t}{2} \right) \\ & - 2 q_t q_f r \text{Im} (\mathcal{A}_f^c(\mathbf{x})^* \mathcal{A}_f^u(\mathbf{x}) e^{i\delta - iq_f(\gamma - 2\beta_s)}) \sin (\Delta m_s t) \end{aligned} \quad (2.28)$$

<sup>173</sup> This renormalization procedure was found to be crucial for the fit stability since it reduces  
<sup>174</sup> the correlation between the  $a_i^c$  and  $a_i^u$  amplitude coefficients significantly. Due to the  
<sup>175</sup> overall normalization, one of the complex amplitude coefficients  $a_i^c$  can be fixed to unity  
<sup>176</sup> and since  $r$  and  $\delta$  are included as fit parameters one of the complex amplitude coefficient  
<sup>177</sup>  $a_i^u$  can be additionally fixed to unity.

<sup>178</sup> We force strong decays in the cascade topology to have the same pattern in  $b \rightarrow c$   
<sup>179</sup> and  $b \rightarrow u$  transitions by the sharing of couplings between related quasi-two-body final  
<sup>180</sup> states. For example, given the two  $a_i^c$  parameters required for  $B_s \rightarrow D_s^- K_1(1270)^+$   
<sup>181</sup> with  $K_1(1270)^+ \rightarrow \rho(770)^0 \pi^+$  and  $K_1(1270)^+ \rightarrow K^*(892) \pi^+$ , the amplitude  $\bar{B}_s \rightarrow$   
<sup>182</sup>  $D_s^- K_1(1270)^+$  with  $K_1(1270)^+ \rightarrow \rho(770)^0 \pi^+$  and  $K_1(1270)^+ \rightarrow K^*(892) \pi^+$  only requires  
<sup>183</sup> one additional global complex parameter to represent the different production processes  
<sup>184</sup> 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  
<sup>185</sup> of  $K_1(1270)^+ \rightarrow \rho(770)^0 \pi^+$  and  $K_1(1270)^+ \rightarrow K^*(892) \pi^+$  are the same regardless of  
<sup>186</sup> the production mechanism. For this purpose, multiple decay amplitudes of a three-body  
<sup>187</sup> resonance are defined relative to a given reference channel.

## 188 2.4 Validation

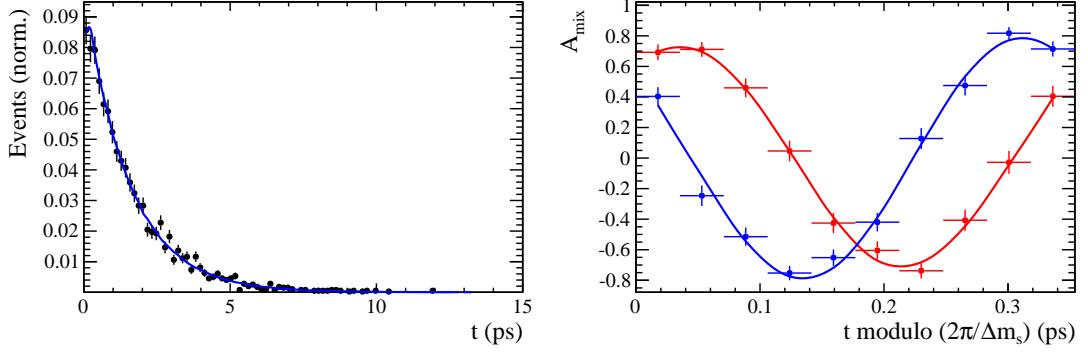
189 The amplitude fit is performed by using the amplitude analysis tool **MINT2**. It was  
 190 previously applied to analyze  $D^0 \rightarrow 4\pi$  and  $D^0 \rightarrow KK\pi\pi$  decays [12] which have an  
 191 identical general spin structure (*i.e.* scalar to four scalar decay) to  $B_s \rightarrow D_s K\pi\pi$  decays.  
 192 In the course of the  $D^0 \rightarrow hhhh$  analysis, the implementation of the amplitudes were  
 193 extensively cross-checked against other available tools such as **qft++** [22], **AmpGen** [23]  
 194 and where possible **EVTGEN** [24]. Since no additional line shapes or spin factors are  
 195 needed for this analysis, we consider the amplitude calculation as fully validated.

196 This does, however, not apply to the full time-dependent amplitude pdf which is  
 197 newly implemented for this analysis. To cross-check it, we use **EVTGEN** to generate  
 198 toy events with time-dependent  $CP$  violation according to the **SSD\_Cp** event model [24].  
 199 Since this event model does not allow for multiple interfering resonances, we generate  
 200 only the decay chain  $B_s \rightarrow D_s(K_1(1400) \rightarrow K^*(892)\pi)$ . Table 2.1 lists the generated  
 201 input parameters. The toy data set is fitted with our **MINT2** implementation of the full  
 202 time-dependent amplitude pdf and the phasespace-integrated pdf. Flat decay-time and  
 203 phase-space acceptances are used. The true ID of the  $B_s$  meson is used to tag the initial  
 204 flavor (*i.e.* perfect tagging). For technical reasons (the extreme case of perfect resolution  
 205 is currently not implemented in our fitter) we smeared the generated decay-times with a  
 206 Gaussian of width 45 fs. We use a Gaussian of the same width to convolute the pdf for  
 207 fitting. More details of the fit procedures are given in Secs. 10 and 11. The fit projections  
 208 are shown in Figs. 2.1 and 2.2.

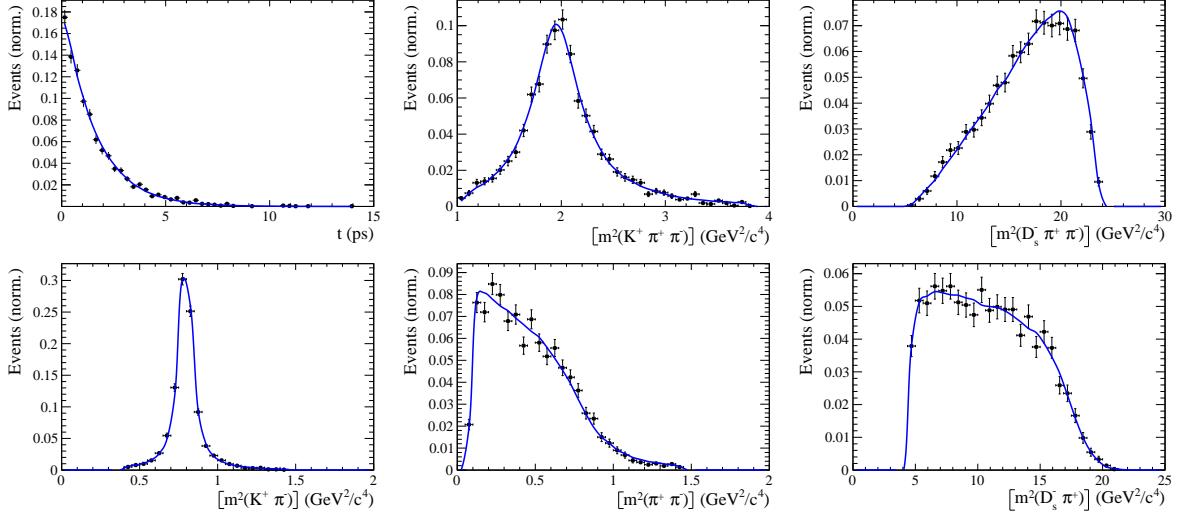
209 The  $CP$  coefficients  $C, D, \bar{D}, S, \bar{S}$  are the fit parameters in case of the phasespace-  
 210 integrated pdf, which are converted after to the fit to the physical observables  $r, \kappa, \delta$  and  $\gamma$   
 211 using the **GammaCombo** package [25]. The obtained 1-CL contours are shown in Fig. 2.3.  
 212 The full pdf determines  $r, \delta$  and  $\gamma$  directly. As shown in Tab. 2.2 and 2.3, the fit results  
 213 are in excellent agreement with the generated input values. The phasespace-integrated fit  
 214 is, in addition, performed with the **B2DX** fitter used for the  $B_s \rightarrow D_s K$  analysis yielding  
 215 identical results. Note though that some parts of the **B2DX** fitter have been taken over to  
 216 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
$\delta$	$10.0^\circ$
$\gamma$	$71.1^\circ$
$\beta_s$	$0.0^\circ$



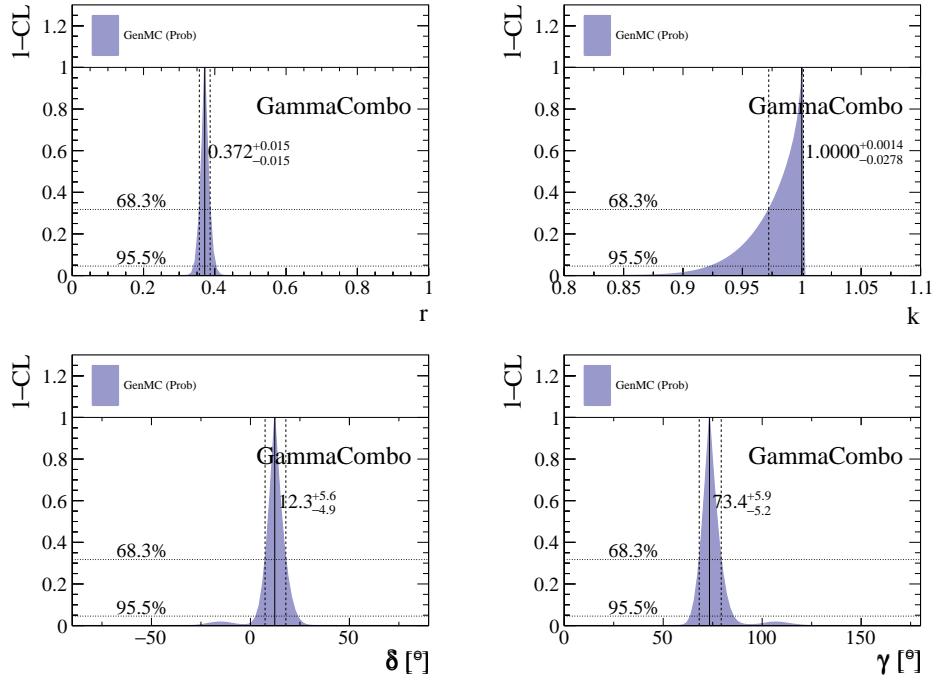
**Figure 2.1:** Left: Time distribution of  $B_s \rightarrow D_s K \pi \pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (solid line). Right: Time-dependent asymmetry between mixed and unmixed events folded into one oscillation period for  $D_s^- K^+ \pi \pi$  (red) 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	$0.767 \pm 0.023$	0.3
$D$	-0.314	$-0.194 \pm 0.205$	0.6
$\bar{D}$	-0.101	$-0.189 \pm 0.210$	-0.4
$S$	-0.570	$-0.556 \pm 0.033$	0.4
$\bar{S}$	-0.643	$-0.683 \pm 0.031$	-1.3



**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	$0.372 \pm 0.015$	$0.372 \pm 0.015$
$\kappa$	1.0	1.0	$1.000 \pm 0.035$
$\delta$	$10.0^\circ$	$12.2 \pm 5.1$	$12.3 \pm 5.1$
$\gamma$	$71.1^\circ$	$73.2 \pm 5.5$	$73.4 \pm 5.5$

### 217 3 Analysis strategy

218 The selection of the signal,  $B_s \rightarrow D_s K\pi\pi$ , and calibration channel,  $B_s \rightarrow D_s \pi\pi\pi$ , are  
219 outlined in Sect. 4, followed by the determination of the signal and background yields. The  
220 calibration channel is used to study the decay-time acceptance, see Sect. 7, and to calibrate  
221 the flavor tagging algorithms in Sect. 8. Moreover, the  $B_s$  mixing frequency is measured  
222 in Sect. 10.1. Afterwards, the CKM angle  $\gamma$  is extracted from  $B_s \rightarrow D_s K\pi\pi$  data using  
223 two different approaches: the results of the phase-space integrated fit are presented in  
224 Sect. 10.2, while Sect. 11 discusses the significantly more complicated full time-dependent  
225 amplitude fit. The systematic uncertainties of both methods are determined in Sect. 12,  
226 before we compare the results and conclude in Sect. 13.

## 227 4 Data samples and event selection

### 228 4.1 Stripping and Trigger selection

229 The dataset used for this analysis corresponds to  $1 \text{ fb}^{-1}$  of proton-proton collision data col-  
230 lected in 2011 with a centre of mass energy  $\sqrt{s} = 7 \text{ TeV}$ ,  $2 \text{ fb}^{-1}$  collected in 2012 with  $\sqrt{s} =$   
231  $7 \text{ TeV}$  and  $4 \text{ fb}^{-1}$  collected in 2015/2016/2017 with  $\sqrt{s} = 13 \text{ TeV}$ . Candidate  $B_s^0 \rightarrow D_s K\pi\pi$   
232 ( $B_s^0 \rightarrow D_s \pi\pi\pi$ ) decays are reconstructed using the `B02DKPiPiD2HHHPIDBeauty2CharmLine`  
233 (`B02DPiPiPiD2HHHPIDBeauty2CharmLine`) line of the `BHadronCompleteEvent` stream of  
234 `Stripping21r1` (2011), `Stripping21` (2012), `Stripping24r1` (2015) and `Stripping28r1p1` (2016)  
235 and `Stripping29r2` (2017). Both stripping lines employ the same selection cuts, listed in  
236 Appendix B, except for the PID requirement on the bachelor kaon/pion.

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

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

### 249 4.2 Offline selection

250 The offline selection, in particular the requirements on the  $D_s$  hadron, are guided by  
251 the previous analyses of  $B_s \rightarrow D_s K/\pi$ ,  $B_d \rightarrow D^0 \pi$  as well as the branching fraction  
252 measurement of  $B_s^0 \rightarrow D_s K\pi\pi$  decays. Tables 4.1 and 4.2 summarize all selection  
253 requirements which are described in the following. Throughout the note, we abbreviate  
254  $B_s^0 \rightarrow D_s X_s (\rightarrow K\pi\pi)$  and  $B_s^0 \rightarrow D_s X_d (\rightarrow \pi\pi\pi)$ .

255 Given the high number of  $pp$  interactions per bunch crossing, a large fraction of  
256 events have more than one reconstructed PV. We choose the 'best' PV to be the one  
257 to which the  $B_s$  candidate has the smallest  $\chi^2_{IP}$ . In instances where the association

of the  $B_s$  candidate to the best PV is wrong, the decay time of this candidate will be incorrect. These wrongly associated candidates are rejected by requiring that the  $B_s$   $\chi^2_{IP}$  with respect to any other PV is sufficiently higher than with respect to the best PV ( $\Delta\chi^2_{IP} = \chi^2_{IP,\text{SECONDBEST}} - \chi^2_{IP,\text{BEST}} > 20$ ). Events with only a single PV are not affected.

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

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

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

We reconstruct the  $B_s^0 \rightarrow D_s h \pi\pi$  decay through three different final states of the  $D_s$  meson:  $D_s \rightarrow KK\pi$ ,  $D_s \rightarrow \pi\pi\pi$  and  $D_s \rightarrow K\pi\pi$ . Of those,  $D_s \rightarrow KK\pi$  is the most prominent one, while  $\mathcal{BR}(D_s \rightarrow \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$  and  $\mathcal{BR}(D_s \rightarrow K\pi\pi) \approx 0.12 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$  holds for the others. For the  $KK\pi$  final state we make use of the well known resonance structure; the decay proceeds either via the narrow  $\phi$  resonance, the broader  $K^{*0}$  resonance or (predominantly) non-resonant. Within the  $\phi$  resonance region the sample is already sufficiently clean after the stripping so that we do not impose additional criteria on the  $D_s$  daughters. For the  $K^{*0}$  and non-resonant regions consecutively tighter requirements on the particle identification and the  $D_s$  flight-distance are applied. We apply global requirements (*i.e.* independent of the  $D_s$  Dalitz-plot position) for the other final states. All cuts are summarized in Table 4.1.

#### 4.2.1 Phase space region

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

297 **4.2.2 Physics background vetoes**

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

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

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

304 Possible with  $\pi^- \rightarrow K^-$  misidentification, vetoed by requiring  $m(K^+ K_\pi^- \pi^-) \neq$   
 305  $m(D^-) \pm 40$  MeV or the  $K^-$  has to fulfill more stringent PID criteria depending  
 306 on the resonant region (see Table 4.1).

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

308 Possible with  $\bar{p} \rightarrow K^-$  misidentification, vetoed by requiring  $m(K^+ K_p^- \pi^-) \neq$   
 309  $m(\Lambda_c^-) \pm 40$  MeV or the  $K^-$  has to fulfill more stringent PID criteria depending  
 310 on the resonant region (see Table 4.1).

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

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

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

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

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

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

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

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

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

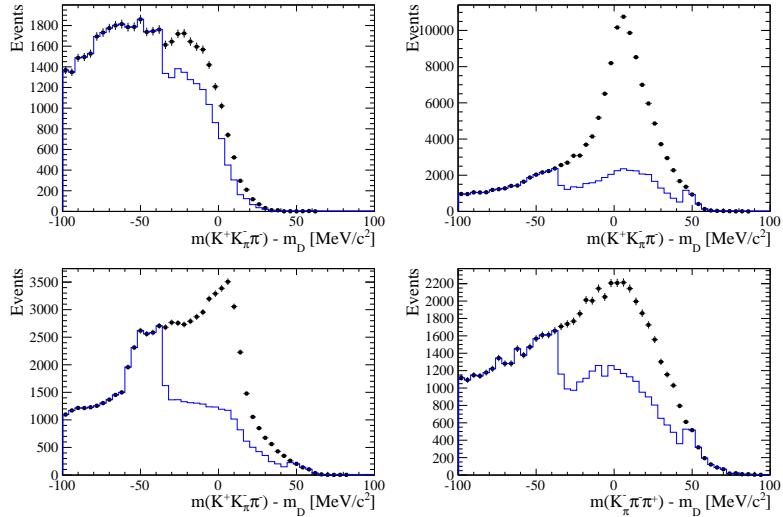
323 Possible with  $\bar{p} \rightarrow K^-$  misidentification, vetoed by requiring  $m(K_p^- \pi^+ \pi^-) \neq$   
 324  $m(\Lambda_c^-) \pm 40$  MeV or  $\text{PIDK}(K^-) - \text{PIDp}(K^-) > 5$ .

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

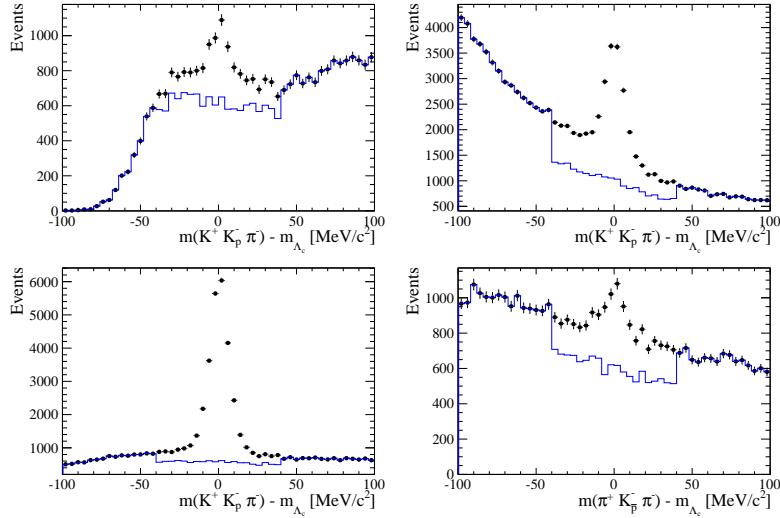
326  $D^0$  combined with a random  $\pi$  can fake a  $D_s \rightarrow K\pi\pi$  decay, vetoed by requiring  
 327 both possible combinations to have  $m(K\pi) < 1750$  MeV.

328 The effects of these veto cuts are illustrated in Figs. 4.1,4.2 and 4.3. To reduce cross-feed  
 329 from our calibration channel into the signal channel and vice-versa we require tight PID  
 330 cuts on the ambiguous bachelor track; for the signal channel we apply  $\text{PIDK}(K^+) > 10$   
 331 and for the calibration channel  $\text{PIDK}(\pi^+) < 0$ . In addition, we veto  $B_s^0 \rightarrow D_s^- D_s^+$  decays  
 332 which is illustrated in Fig. 4.4.

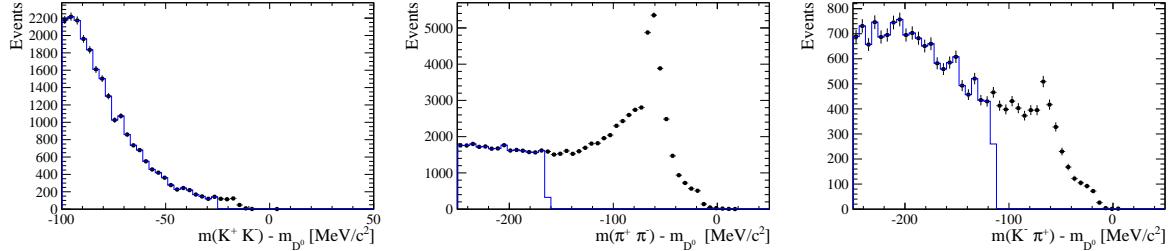
- 333 1.  $X_s^+ \rightarrow K^+\pi^+\pi^-$ :
- 334 (a)  $B_s^0 \rightarrow D_s\pi\pi\pi$ :  
 335 Possible with  $\pi^+ \rightarrow K^+$  misidentification, suppressed with  $\text{PIDK}(K^+) > 10$ .
- 336 (b)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^+\pi^-)$ :  
 337 Outside of considered phase-space region, see Sec. 4.2.1.
- 338 (c)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+K^-\pi^+)$ :  
 339 To suppress  $B_s^0 \rightarrow D_s^-K^+K^-\pi^+$  background, possible with  $K^- \rightarrow \pi^-$  misiden-  
 340 tification, we require  $\text{PIDK}(\pi^-) < 0$ . In case the invariant mass of the  $K^+\pi^+\pi^-$   
 341 system recomputed applying the kaon mass hypothesis to the pion is close to  
 342 the  $D_s$  mass ( $m(K^+\pi^+\pi_K^-) = m(D_s) \pm 20$  MeV), we further tighten the cut to  
 343  $\text{PIDK}(\pi^-) < -5$ .
- 344 2.  $X_d^+ \rightarrow \pi^+\pi^+\pi^-$ :
- 345 (a)  $B_s^0 \rightarrow D_sK\pi\pi$ :  
 346 Possible with single missID of  $K^+ \rightarrow \pi^+$ , suppressed with  $\text{PIDK}(\pi^+) < 0$ .
- 347 (b)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow \pi^+\pi^+\pi^-)$ :  
 348 Outside of considered phase-space region, see Sec. 4.2.1.
- 349 (c)  $B_s^0 \rightarrow D_s^-(D_s^+ \rightarrow K^+\pi^-\pi^+)$ :  
 350 Possible with single missID of  $K^+ \rightarrow \pi^+$ , vetoed by requiring  $m(\pi^+\pi_K^+\pi^-) \neq$   
 351  $m(D_s) \pm 20$  MeV or  $\text{PIDK}(\pi^+) < -5$  for both  $\pi^+$ .



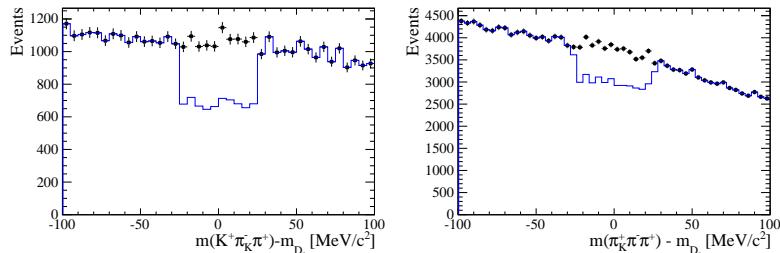
**Figure 4.1:** Background contributions from  $D^-$  decays where the  $\pi^-$  is misidentified as  $K^-$ . The  $D_s$  invariant mass is recomputed applying the pion mass hypothesis to the kaon and shown for the  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*(892)K$ ,  $D_s \rightarrow KK\pi$  (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 4.2:** 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 4.3:** 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 4.4:** 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.

352    **4.2.3 Training of multivariate classifier**

353    The Toolkit for Multivariate Analysis (TMVA [26]) is used to train a multivariate classifier  
 354    (BDT with gradient boosting, BDTG) discriminating signal and combinatorial background.  
 355    We use  $B_s \rightarrow D_s \pi\pi$  data that passes all previously mentioned selection steps, including  
 356    the trigger and stripping selection, offline cuts and physical background vetoes, as signal  
 357    proxy. The background is statistically subtracted by applying `sWeights` based on the  
 358    fit to the reconstructed  $B_s$  mass shown in Fig. 4.5(left). This is a simplified version  
 359    (performed in a reduced mass range) of the final mass fits described in Sec. 5. The  
 360    sideband  $B_s \rightarrow D_s K\pi\pi$  data ( $m(B_s) > 5500$  MeV) is used as background proxy.

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

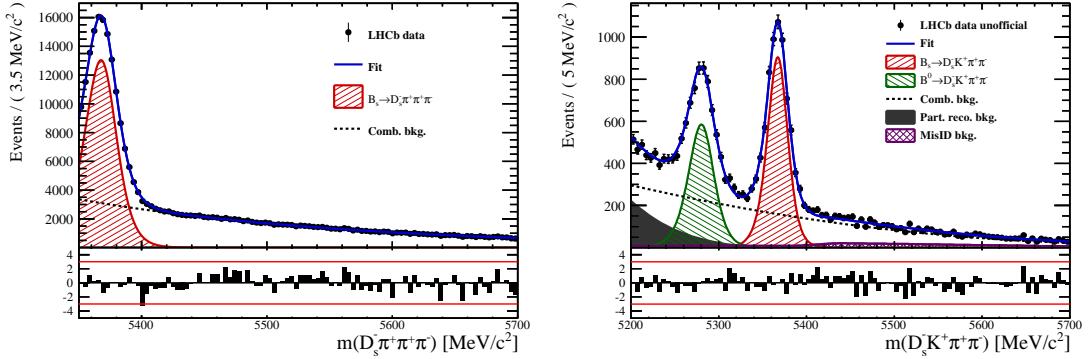
367    The following discriminating variables are used for the BDTG training<sup>1</sup>:

- 368    • logarithm of the  $B_s$  impact-parameter  $\chi^2$ ,  $B_s \log(\chi_{IP}^2)$
- 369    • logarithm of the cosine of the  $B_s$  direction angle,  $\log(\text{DIRA})$
- 370    • fit quality of the DTF with PV constrain,  $\chi_{DTF}^2/ndf$
- 371    • logarithm of the minimal  $B_s^0$  decay vertex quality difference for adding one extra  
 372    track,  $\log(\Delta\chi_{add-track}^2)$
- 373    • the difference between the transverse momentum of the  $B_s$ - candidate and the  
 374    transverse momentum of all the particles reconstructed with a cone of radius  
 375     $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2} < 1$  rad around the  $B_s$ - candidate normalized to the sum of  
 376    both,  $B_s A_{pT}^{\text{cone}}$
- 377    • largest ghost probability of all tracks,  $\max(\text{ghostProb})$
- 378    • logarithm of the the smallest  $X = X_d, X_s$  daughter impact-parameter  $\chi^2$ ,  $X \log(\min(\chi_{IP}^2))$
- 379    • largest distance of closest approach of the  $X = X_d, X_s$  daughters,  $\max(\text{DOCA})$
- 380    • cosine of the largest opening angle between the  $D_s$  and another bachelor track  $h_i$  in  
 381    the plane transverse to the beam,  $\cos(\max \theta_{D_s h_i})$
- 382    • logarithm of the the smallest  $D_s$  daughter impact-parameter  $\chi^2$ ,  $D_s \log(\min(\chi_{IP}^2))$
- 383    • logarithm of the  $D_s$  flight-distance significance,  $D_s \log(\chi_{FD}^2)$
- 384    • logarithm of the  $D_s$  radial flight-distance,  $D_s \log(RFD)$

---

<sup>1</sup> The following options are chosen: NTrees=500, MinNodeSize=2.5%, BoostType=Grad:Shrinkage=0.10, UseBaggedBoost:BaggedSampleFraction=0.5, nCuts=40, MaxDepth=3, NegWeightTreatment=Pray.

386 Loose cuts on the variables  $\chi^2_{DTG}/ndf$ ,  $\Delta\chi^2_{add-track}$  and  $\cos(\max\theta_{D_s h_i})$  are applied prior  
 387 to the training which are expected to be 100% signal efficient. Figure 4.6 shows the  
 388 distributions of the input variables for signal and background. As shown in Appendix C,  
 389 these distributions differ between data-taking period and trigger category. In particular  
 390 variables depending on the  $B_s$  kinematics and the event multiplicity are affected (*e.g.*  
 391  $\theta_{D_s h_i}$  or  $A_{pT}^{cone}$ ). The BDTG is consequently trained separately for these categories. The  
 392 resulting classifier response is shown in Fig. 4.7 for each category (even and odd test  
 393 samples combined) and in Appendix C for each training.



**Figure 4.5:** Left: Reconstructed  $B_s$  mass for  $B_s \rightarrow D_s \pi\pi\pi\pi$  events that pass the preselection (all categories combined). The fitted PDF is shown in blue, the signal component in red and the background component in black.

Right: Reconstructed  $B_s$  mass for  $B_s \rightarrow D_s K\pi\pi$  events that pass the  $BDTG > 0$  requirement (all categories combined).

#### 394 4.2.4 Final selection

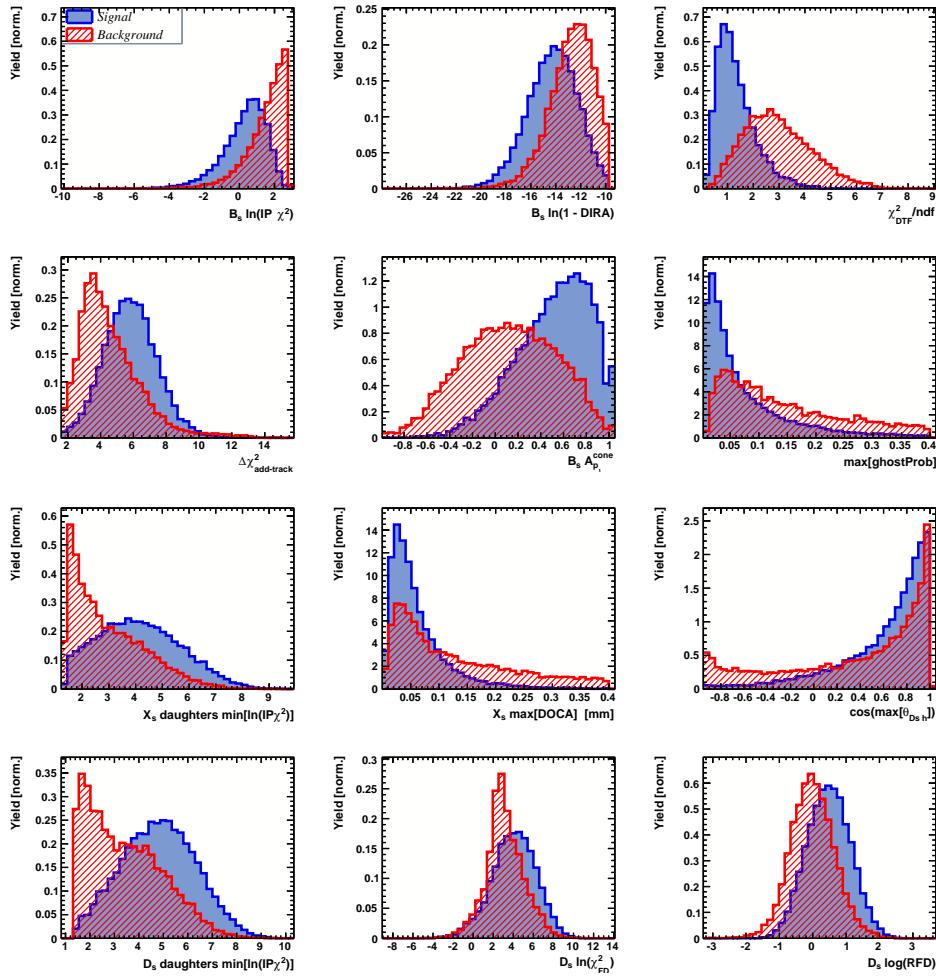
395 The cut on the BDTG output is optimized using the signal significance as figure of merit:

$$FOM(BDTG) = \frac{N_s(BDTG)}{\sqrt{N_s(BDTG) + N_b(BDTG)}} \quad (4.1)$$

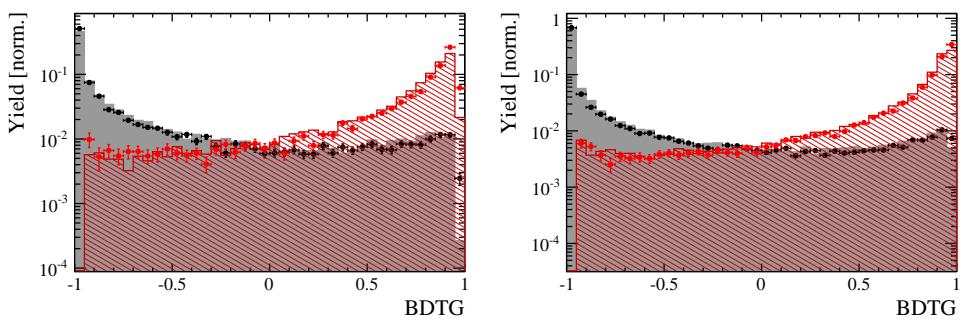
396 where  $N_s(BDTG)$  is the  $B_s \rightarrow D_s K\pi\pi$  signal yield for a given BDTG cut and  $N_b(BDTG)$   
 397 is the combinatorial background yield in the signal region ( $m(D_s K\pi\pi) = m_{B_s} \pm 40$  MeV).  
 398 To compute the yields as function of the BDTG cut, we use the BDTG efficiencies,  $\epsilon_{s,b}$ ,  
 399 evaluated on the corresponding test samples. To fix the overall scale, it is required to  
 400 know the yields at (at least) one point of the scanned range. We arbitrarily choose this  
 401 fix point to be  $BDTG > 0$  and perform a fit to the reconstructed  $B_s$  mass as described  
 402 in Sec. 5 to obtain the yields  $N_{s,b}(0)$ , see Fig. 4.5(right). These yields are then efficiency  
 403 corrected to calculate the yields for a given BDTG cut:

$$N_{s,b}(BDTG) = N_{s,b}(0) \cdot \frac{\epsilon_{s,b}(BDTG)}{\epsilon_{s,b}(0)}. \quad (4.2)$$

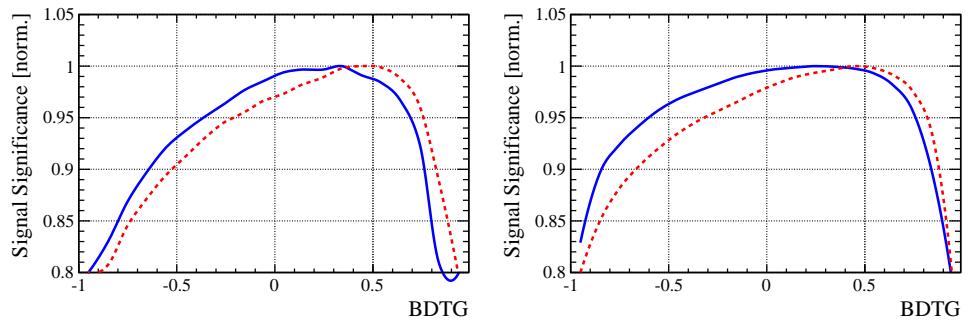
404 Figure 4.8 shows the resulting BDTG scans for each training category.



**Figure 4.6:** Discriminating variables used to train the BDTG for all data categories combined.



**Figure 4.7:** 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.



**Figure 4.8:** 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 4.1:** 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 4.2:** 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	$= 1$

### 405 4.3 Simulation

406 Several Monte Carlo (MC) samples are used in the analysis for acceptance and background  
 407 studies. A full list of them is given in Tab. 4.3. In each case, the decay model includes  
 408 a mixture of non-interfering resonances contributing to the bachelor system and a non-  
 409 resonant (phase-space) component. For  $B_s \rightarrow D_s X_s$  these are: 25%  $X_s \rightarrow \pi(K_1(1270) \rightarrow$   
 410  $K\rho(770))$ , 70%  $X_s \rightarrow \pi(K_1(1400) \rightarrow K^*(892)\pi)$  and 5%  $X_s \rightarrow K\pi\pi$  (non-resonant). And  
 411 similar for  $B_s \rightarrow D_s X_d$ : 85%  $X_d \rightarrow \pi(a_1(1260) \rightarrow \rho(770)\pi)$ , 10%  $X_d \rightarrow \pi(a_1(1260) \rightarrow \sigma\pi)$   
 412 and 5%  $X_d \rightarrow K\pi\pi$  (non-resonant). All MC samples are generated using Pythia8,  
 413 reconstructed using Reco14c, Reco15 and Reco16 for Run-I, 15 and 16 data and selected  
 414 using the same criteria as in data.

**Table 4.3:** 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

## 415 5 Yields determination

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

421 Due to different mass resolutions, we perform the invariant mass fits simultaneously  
 422 for each data-taking period and each trigger category. We further introduce four  $D_s$  final  
 423 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  
 424 different signal purities. The  $D_s \rightarrow Kh\pi$  category combines the two  $D_s$  decay channels  
 425 with the lowest statistics:  $D_s \rightarrow KK\pi$  (non-resonant) and  $D_s \rightarrow K\pi\pi$ . This amounts to  
 426 16 categories in total.

### 427 5.1 Signal model

428 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  
 429 using a Johnson's SU function [27], which results from a variable transformation of a  
 430 normal distribution to allow for asymmetric tails:

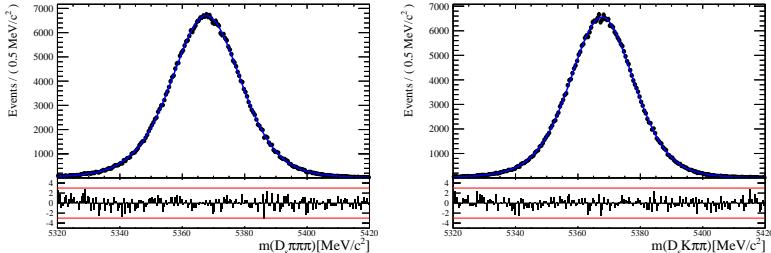
$$\mathcal{J}(x|\mu, \sigma, \nu, \tau) = \frac{e^{-\frac{1}{2}r^2}}{2\pi \cdot c \cdot \sigma \cdot \tau \cdot \sqrt{z^2 + 1}} \quad (5.1)$$

$$r = -\nu + \frac{\operatorname{asinh}(z)}{\tau} \quad (5.2)$$

$$z = \frac{x - (\mu - c \cdot \sigma \cdot e^\tau \sinh(\nu \cdot \tau))}{c \cdot \tau} \quad (5.3)$$

$$c = \frac{e^{\tau^2} - 1}{2\sqrt{e^{\tau^2} \cdot \cosh(2\nu \cdot \tau) + 1}}. \quad (5.4)$$

431 It is conveniently expressed in terms of the central moments up to order four: The mean  
 432 of the distribution  $\mu$ , the standard deviation  $\sigma$ , the skewness  $\nu$  and the kurtosis  $\tau$ . The  
 433 tail parameters  $\nu$  and  $\tau$  are fixed to the values obtained by a fit to the invariant mass  
 434 distribution of simulated events shown in Fig 5.1. To account for differences between  
 435 simulation and real data, linear scaling factors for the mean  $\mu$  and width  $\sigma$  are determined  
 436 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  
 437 factors are determined separately for each data-taking period and each trigger category.



**Figure 5.1:** 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.

438 **5.2 Background models**

439 After the full selection the following residual background components have to be accounted  
440 for:

441

442 **Combinatorial background**

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

446

447 **Peaking  $B_d$  background**

448 Decays of  $B_d$  mesons into the  $D_s h\pi\pi$  final state are described by the  $B_s$  signal PDF  
449 where the mean is shifted by the known mass difference  $m_{B_s} - m_{B_d}$  [15].

450

451 **Partially reconstructed background**

452 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  
453 to be peaking lower than signal in the  $m(D_s \pi\pi\pi)$  spectrum with large tails due to the  
454 momentum carried away by the not reconstructed  $\pi^0$  or  $\gamma$ . An empirical description for  
455 the shape of this contribution is derived from a  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  MC sample subject to  
456 the nominal  $B_s^0 \rightarrow D_s \pi\pi\pi$  selection. Figure 5.2 (top) shows the respective reconstructed  
457  $m(D_s \pi\pi\pi)$  distribution. A sum of three bifurcated Gaussian functions (*i.e.* Gaussian  
458 functions with different widths on the left and the right side of the maximum value) is used  
459 to describe it. In the fit to data, all parameters are fixed to the ones obtained from MC  
460 except for the parameter which describes the width of the right tail of the distribution to  
461 account for data-simulation differences in mass resolution. The equivalent  $B_s^0 \rightarrow D_s^* K\pi\pi$   
462 component contributing to the  $B_s^0 \rightarrow D_s K\pi\pi$  data sample is described by the same PDF  
463 with the right tail fixed to the  $B_s^0 \rightarrow D_s \pi\pi\pi$  result.

464 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  
465 shifted by  $m_{B_s^0} - m_{B^0}$ .

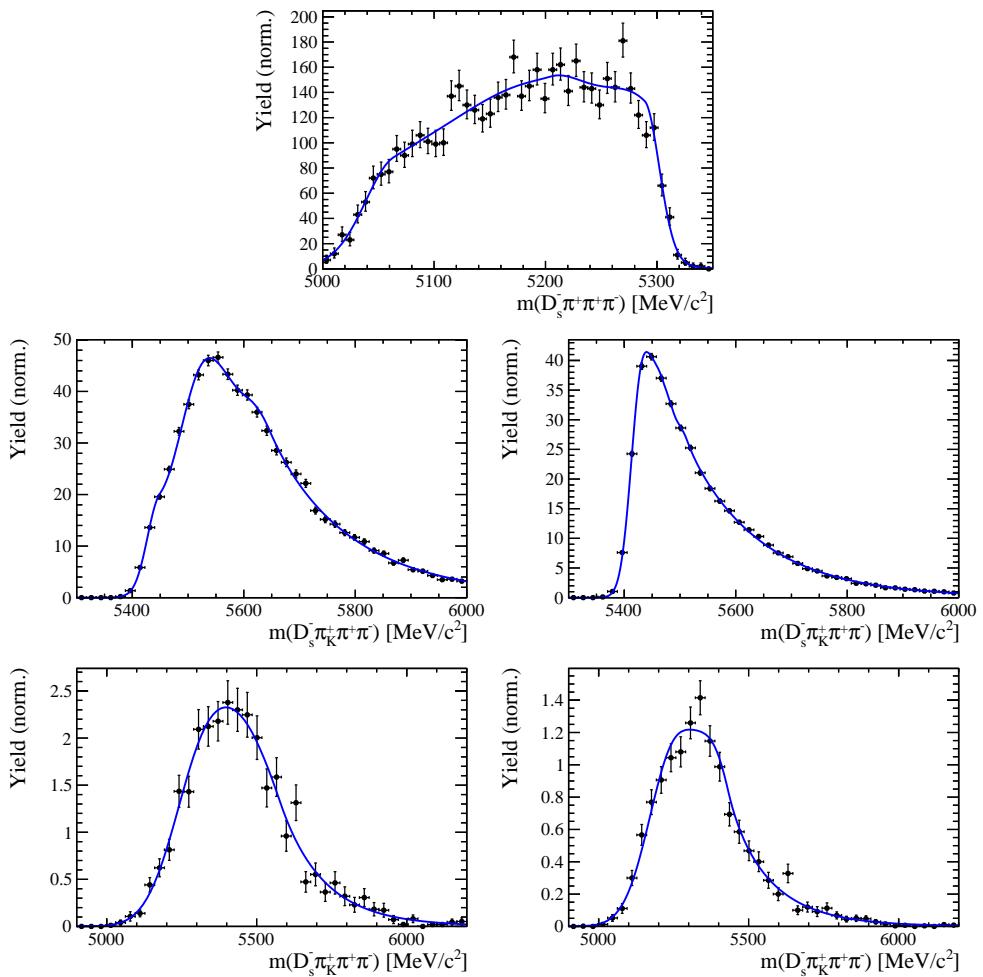
466

467 **Misidentified background**

468 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  
469 pions is misidentified as a kaon, contaminate the  $B_s^0 \rightarrow D_s K^+ \pi^+ \pi^-$  sample. To determine  
470 the corresponding background shapes, we use simulated events passing the nominal  
471 selection except for the PID cuts on the bachelor  $\pi^+$  tracks. The **PIDCalib** package  
472 is used to determine the  $p_T, \eta$ -dependent  $\pi^+ \rightarrow K^+$  misidentification probability for  
473 each pion. We change the particle hypothesis from pion to kaon for the pion with the  
474 higher misidentification probability and recompute the invariant  $B_s^0$  mass,  $m(D_s^- \pi_K^+ \pi^+ \pi^-)$ .  
475 Similarly, the invariant masses  $m(\pi_K^+ \pi^+ \pi^-)$  and  $m(\pi_K^+ \pi^-)$  are recomputed and required  
476 to be within the considered phasespace region. The background distributions are shown  
477 in Fig. 5.2 (middle, bottom) and modeled by the sum of three Crystal Ball functions.

478 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  
 479 the  $B_s^0 \rightarrow D_sK\pi\pi$  sample is computed by multiplying the fake rate (within the considered  
 480  $B_s$  mass range) of  $0.63 \pm 0.01\%$  ( $0.55 \pm 0.02\%$ ) for Run-I and  $0.33 \pm 0.01\%$  ( $0.24 \pm 0.01\%$ )  
 481 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  
 482 fit to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  data sample. The yields are corrected for the  $\text{PID}(\pi^+) < 0$   
 483 requirement which has an efficiency of  $77.1 \pm 0.1\%$  for Run-I and  $81.0 \pm 0.1\%$  for Run-II  
 484 data. The  $B_s^0 \rightarrow D_s^*\pi^+\pi^+\pi^-$  yield is additionally corrected for the efficiency of the cut  
 485  $m(D_sK\pi\pi) > 5200 \text{ MeV}$  evaluated on MC. In the fit to data, the misidentified background  
 486 yields are fixed to the predicted ones.

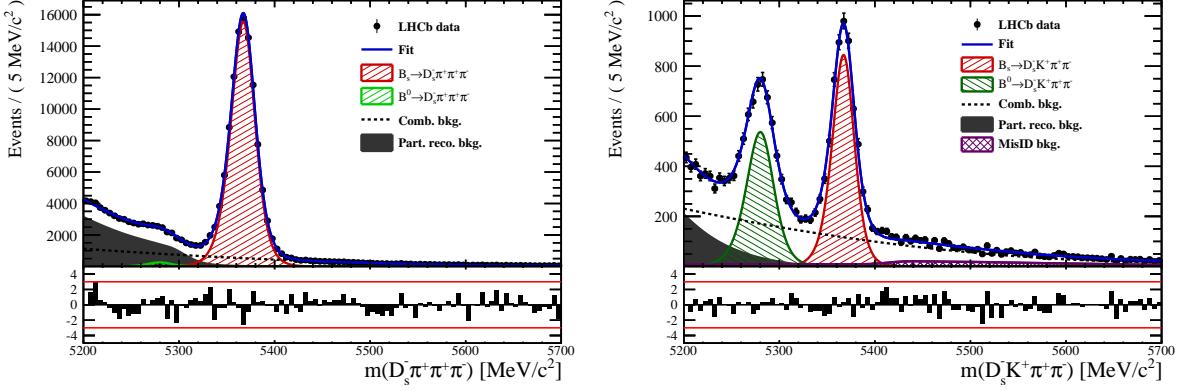
487 We consider the  $B_s^0 \rightarrow D_sK\pi\pi$  and  $B_s^0 \rightarrow D_s^*K\pi\pi$  components contributing to the  
 488  $B_s^0 \rightarrow D_s\pi\pi\pi$  data sample to be negligible due to the low branching fractions and the  
 489 tight PID cuts on the bachelor pions.



**Figure 5.2:** 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.

### 490 5.3 Results

491 Figure 5.3 shows the invariant mass distribution for  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$  can-  
 492 didates passing all selection criteria. The projections for all categories of the simultaneous  
 493 fit are shown in Appendix D. The integrated signal and background yields are listed in  
 494 Tables 5.1 and 5.2.



**Figure 5.3:** 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 integrated over all categories.

**Table 5.1:** 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 5.2:** 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.	$9172 \pm 221$

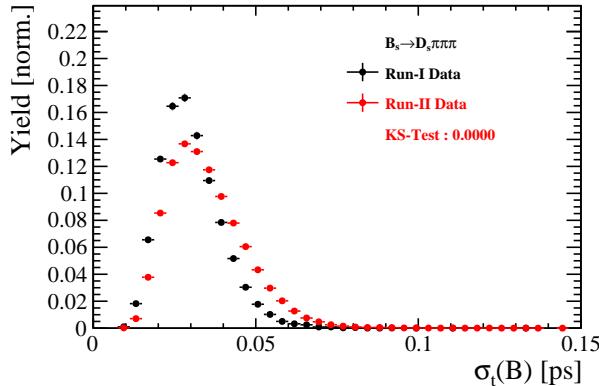
$D_s$ final state	Signal yield
$D_s^- \rightarrow \phi^0(1020)\pi^-$	$1637 \pm 47$
$D_s^- \rightarrow K^{*0}(892)K^-$	$1537 \pm 47$
$D_s^- \rightarrow (K^- h^+\pi^-)$	$1134 \pm 41$
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	$864 \pm 41$

## 495 6 Decay-time Resolution

496 The observed oscillation of  $B$  mesons is prone to dilution, if the detector resolution is  
 497 of similar magnitude as the oscillation period. In the  $B_s^0$  system, considering that the  
 498 measured oscillation frequency of the  $B_s^0$  [28] and the average LHCb detector resolution [29]  
 499 are both  $\mathcal{O}(50 \text{ fs}^{-1})$ , this is the case. Therefore, it is crucial to correctly describe the  
 500 decay time resolution in order to avoid a bias on the measurement of time dependent CP  
 501 violation. Since the time resolution depends on the particular event, especially the decay  
 502 time itself, the sensitivity on  $\gamma$  can be significantly improved by using an event dependent  
 503 resolution model rather than an average resolution. For this purpose, we use the per-event  
 504 decay time error that is estimated based on the uncertainty obtained from the global  
 505 kinematic fit of the momenta and vertex positions (DTF) with additional constraints on  
 506 the PV position and the  $D_s$  mass. In order to apply it correctly, it has to be calibrated.  
 507 The raw decay time error distributions for  $B_s^0 \rightarrow D_s\pi\pi\pi$  signal candidates are shown in  
 508 Figure 6.1 for Run-I and Run-II data. Significant deviations between the two different  
 509 data taking periods are observed due to the increase in center of mass energy from Run-I  
 510 to Run-II, as well as (among others) new tunings in the pattern and vertex reconstruction.  
 511 The decay time error calibration is consequently performed separately for both data taking  
 512 periods.

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

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



**Figure 6.1:** 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).

## 518 6.1 Calibration for Run-I data

519 For simulated  $B_s^0 \rightarrow D_s K \pi\pi$  events, the spread of the differences between reconstructed  
520 decay time and true decay time,  $\Delta t = t - t_{true}$ , is a direct measure of the decay time  
521 resolution. The sum of two Gaussian pdfs with common mean but different widths is used  
522 to fit the distribution of the decay time difference  $\Delta t$  as shown in Fig. 6.2. The effective  
523 damping of the CP amplitudes due to the finite time resolution is described by the dilution  
524  $\mathcal{D}$ . In the case of infinite precision, there would be no damping and therefore  $\mathcal{D} = 1$  would  
525 hold, while for a resolution that is much larger than the  $B_s^0$  oscillation frequency,  $\mathcal{D}$  would  
526 approach 0. For a double-Gaussian resolution model, the dilution is given by [7]

$$\mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s^2 / 2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s^2 / 2}, \quad (6.1)$$

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

$$\sigma_{eff} = \sqrt{(-2/\Delta m_s^2) \ln \mathcal{D}}, \quad (6.2)$$

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

533 To analyze the relation between the per-event decay time error  $\sigma_t$  and the actual  
534 resolution  $\sigma_{eff}$ , the simulated  $B_s^0 \rightarrow D_s K \pi\pi$  sample is divided into equal-statistics slices  
535 of  $\sigma_t$ . For each slice, the effective resolution is determined as described above. Details of  
536 the fit results in each slice are shown in appendix E. Figure 6.2 shows the obtained values  
537 for  $\sigma_{eff}$  as a function of the per-event decay time error  $\sigma_t$ . To account for the variable  
538 binning, the bin values are not placed at the bin center but at the weighted mean of the  
539 respective per-event-error bin. A linear function is used to parametrize the distribution.  
540 The obtained values are

$$\sigma_{eff}^{MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.232 \pm 0.010) \sigma_t \quad (6.3)$$

541 where the offset is fixed to 0. For comparison, the calibration function found for  $B_s^0 \rightarrow D_s K$   
542 MC is also shown in Figure 6.2 [7]:

$$\sigma_{eff}^{D_s K, MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.201 \pm 0.013) \sigma_t. \quad (6.4)$$

543 Due to the good agreement between the scale factors for  $B_s^0 \rightarrow D_s K$  and  $B_s^0 \rightarrow D_s K \pi\pi$   
544 MC, we conclude that the resolution calibration for  $B_s^0 \rightarrow D_s K$  data [7]:

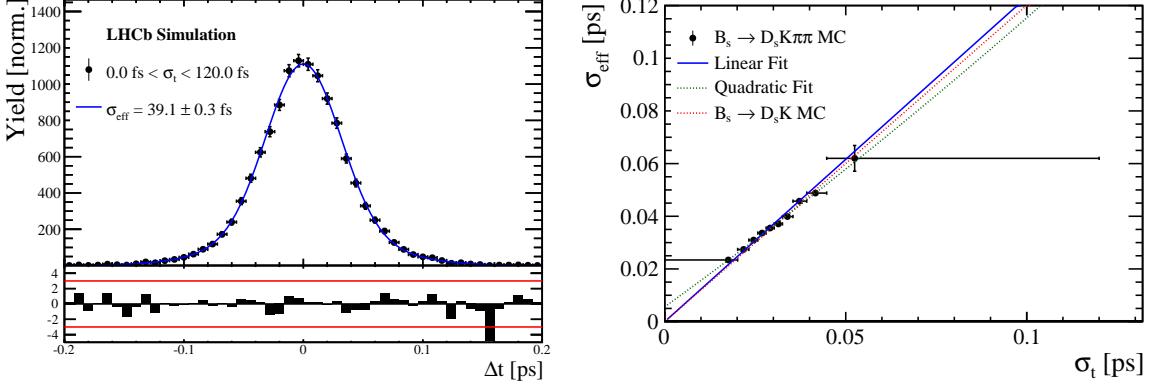
$$\sigma_{eff}^{D_s K}(\sigma_t) = (10.26 \pm 1.52) \text{ fs} + (1.280 \pm 0.042) \sigma_t \quad (6.5)$$

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

$$\sigma_{eff}^{D_s K}(\sigma_t) = (-0.568 \pm 1.570) \text{ fs} + (1.243 \pm 0.044) \sigma_t \quad (6.6)$$

$$\sigma_{eff}^{D_s K}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.772 \pm 0.012) \sigma_t \quad (6.7)$$

548 It was also observed in this analysis that the scale factor is largely independent of the  $B_s$   
549 kinematics and decay-time [6], consistent with other studies, for example in [30].



**Figure 6.2:** (Left) Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K\pi\pi$  MC candidates. (Right) The measured resolution  $\sigma_{eff}$  as function of the per-event decay time error estimate  $\sigma_t$  for  $B_s \rightarrow D_s K\pi\pi$  MC (Run-I). The fitted calibration curve is shown in blue.

## 550 6.2 Calibration for Run-II data

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

560 The offline selection of the fake  $B_s$  candidates is summarized in Tab. 6.1. The selection  
 561 is similar than the one for real data with the important difference that the  $D_s$  candidate  
 562 is required to come from the PV by cutting on the impact parameter significance. Even  
 563 after the full selection, a significant number of multiple candidates is observed. These  
 564 are predominantly fake  $B_s$  candidates that share the same  $D_s$  candidate combined with  
 565 different random tracks from the PV. We select one of these multiple candidates randomly  
 566 which retains approximately 20% of the total candidates. As can be seen in Figure K.3,  
 567 the shapes of the distributions of the unscaled decay time error  $\sigma_t$  for data taken in 2016  
 568 and 2017 are significantly different. Therefore, the scaling of the decay time error is  
 569 treated separately for 2015+2016 and 2017 data. The invariant mass distribution of the  
 570 selected  $D_s$  candidates is shown in Fig. 6.3. To separate true  $D_s$  candidates from random  
 571 combinations, the sPlot method is used to statistically subtract combinatorial background  
 572 from the sample.

573 Figure 6.4 and 6.5 show the sWeighted decay-time distributions for fake  $B_s$  candidates  
 574 from 2016 and 2017 data, respectively. Similar as in the previous section, the decay-time  
 575 distribution is fitted with a double-Gaussian resolution model in slices of the per-event  
 576 decay time error. Since some  $D_s$  candidates might actually originate from true  $B_s$  decays,  
 577 the decay-time distribution of the fake  $B_s$  candidates might show a bias towards positive  
 578 decay times. Therefore, we determine the decay-time resolution from the negative decay-  
 579 time distribution only. Details of the fit results in each slice are shown in appendix E.

580 The resulting calibration functions are:

$$\sigma_{eff}^{Data,16}(\sigma_t) = (11.6 \pm 1.6) \text{ fs} + (0.877 \pm 0.040) \sigma_t \quad (6.8)$$

581

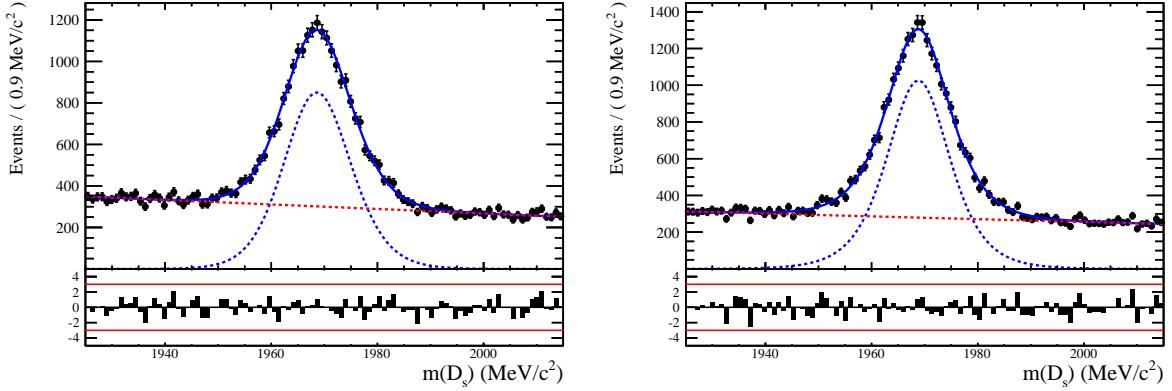
$$\sigma_{eff}^{Data,17}(\sigma_t) = (6.5 \pm 1.4) \text{ fs} + (0.961 \pm 0.036) \sigma_t \quad (6.9)$$

582 For 2016, the result is in good agreement with the one obtained for the  $B_s \rightarrow J/\psi\phi$   
583 (Run-II) analysis that uses 2016 data [31].

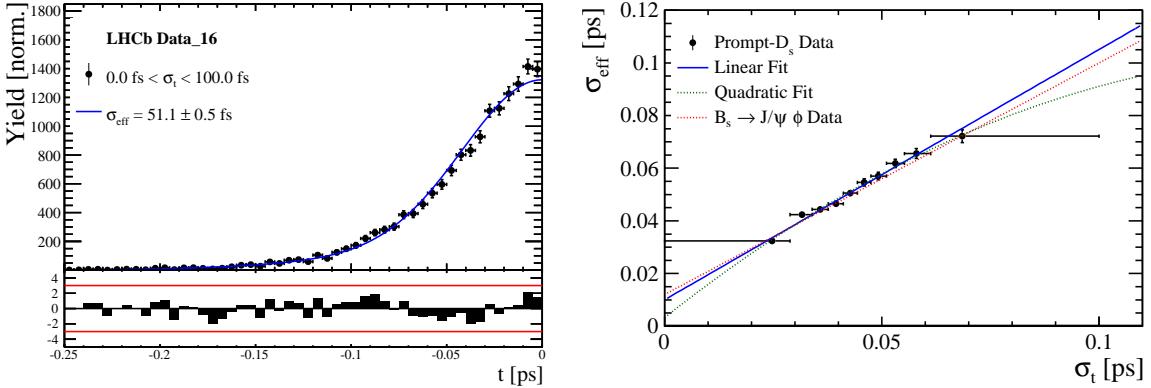
$$\sigma_{eff}^{J/\psi\phi}(\sigma_t) = (12.25 \pm 0.33) \text{ fs} + (0.8721 \pm 0.0080) \sigma_t \quad (6.10)$$

**Table 6.1:** 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. 4.1.

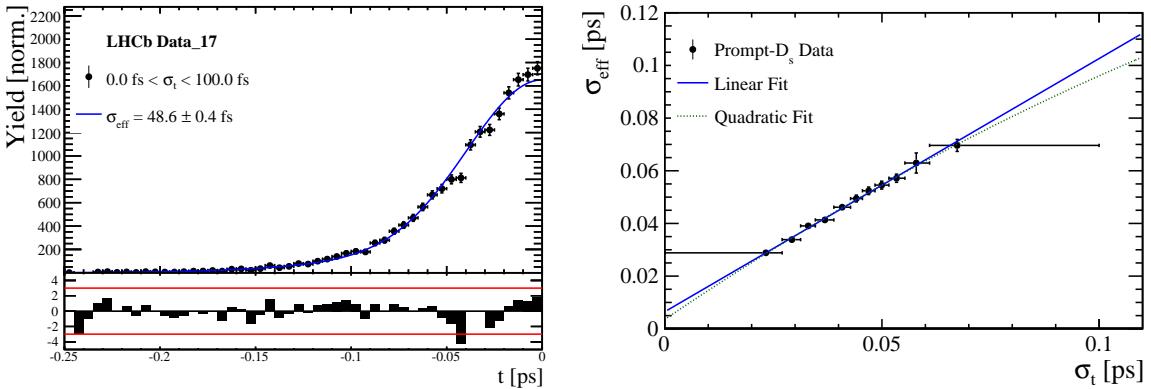
	Description	Requirement
$B_s \rightarrow D_s K\pi\pi$	$\chi_{vtx}^2/\text{ndof}$	< 8
	$\chi_{DTF}^2/\text{ndof}$	< 15
	$t$	< 0 ps
$D_s \rightarrow hh$	$\chi_{vtx}^2/\text{ndof}$	< 5
	DIRA	> 0.99994
	$\chi_{FD}^2$	> 9
	$p_T$	> 1800 MeV
	$\chi_{IP}^2$	< 9
	$\chi_{IP}^2(h)$	> 5
Wrong PV veto	nPV = 1   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



**Figure 6.3:** The invariant mass distribution for prompt  $D_s$  candidates for data taken from the LTUB stripping line in (left) 2016 and (right) 2017.



**Figure 6.4:** (Left) Decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks. (Right) The measured resolution  $\sigma_{eff}$  as function of the per-event decay time error estimate  $\sigma_t$  for fake  $B_s$  candidates (Run-II data). The fitted calibration curve is shown in blue. Data taken in 2016.



**Figure 6.5:** (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  $\sigma_{eff}$  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.

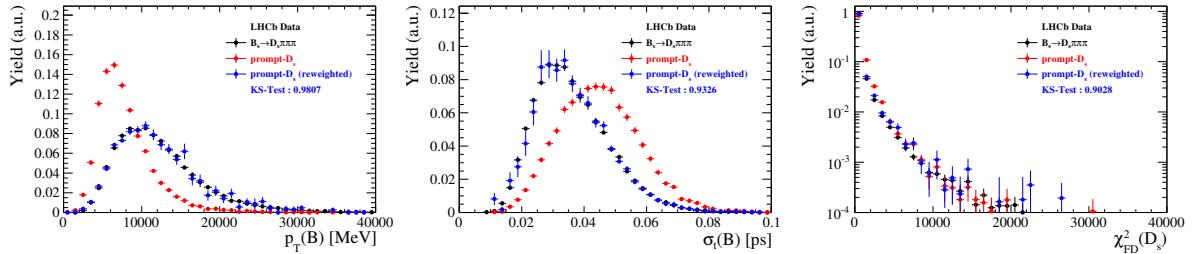
### 584 6.3 Decay-time bias

585 The analyses of  $B_s \rightarrow J/\psi\phi$  [32] and  $B_s \rightarrow D_s\pi$  [33] decays reported that the mean  
 586 of the resolution model is not centered around zero for Run-II data. This decay-time  
 587 bias is understood to be an effect of a misalignment between the VELO halves in the  
 588 order of  $8\text{ }\mu\text{m}$  in  $x$ -direction. To determine the decay-time bias from the prompt  $D_s K\pi\pi$   
 589 sample, the requirement  $t < 0$  used for the calibration of the decay-time error estimate  
 590 is removed. Since the decay-time bias shows a significant dependence on kinematic  
 591 and geometric properties of the decay, the prompt sample is reweighted to match the  
 592 three-dimensional  $p_T(B_s)$ ,  $\sigma_t$  and  $\chi_{FD}^2(D_s)$  distribution observed on  $B_s \rightarrow D_s\pi\pi\pi$  signal  
 593 data as shown in Fig. 6.6. The decay-time bias is extracted from a fit to the (reweighted)  
 594 decay-time distribution of the prompt candidates shown in Fig. 6.7. A systematic error  
 595 is estimated by utilizing different reweighting schemes: a 2D reweighting in  $p_T(B_s)$ ,  $\sigma_t$   
 596 and three consecutive 1D reweightings in  $p_T(B_s)$ ,  $\sigma_t$  and  $\chi_{FD}^2(D_s)$ . Additionally, a fit is  
 597 performed which includes an exponential decay-time distribution to account for a potential  
 598 long-lived background component from true  $B$  decays and the fit range is adjusted from  
 599  $[-0.25, 0.25]\text{ ps}$  to  $[-0.2, 0.05]\text{ ps}$ . The largest deviation from the nominal result is assigned  
 600 as systematic.

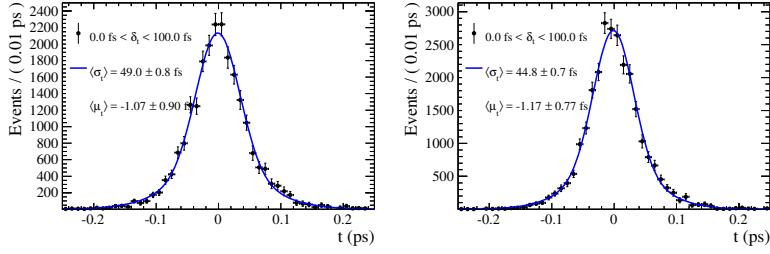
601 The procedure is validated using simulated  $B_s \rightarrow D_s\pi$  signal and prompt  $D_s\pi$  samples,  
 602 which are processed including a realistic VELO misalignment (as observed for data). As  
 603 shown in Fig. 6.8, the decay-time biases extracted from the two samples differ significantly  
 604 before the reweighting and are in a good agreement afterwards.

605 It was further observed that the offline selection, in particular the requirement on  
 606 the  $D_s$  flight distance significance, introduces an additional decay-time bias of the order  
 607  $-1\text{ fs}$  for signal decays, while prompt candidates are not affected. The decay-time bias  
 608 originating from the candidate selection is thus determined from the  $B_s \rightarrow D_s\pi\pi\pi$  MC  
 609 sample, see Fig. 6.9. Identical results are obtained with the  $B_s \rightarrow D_s K\pi\pi$  MC sample.

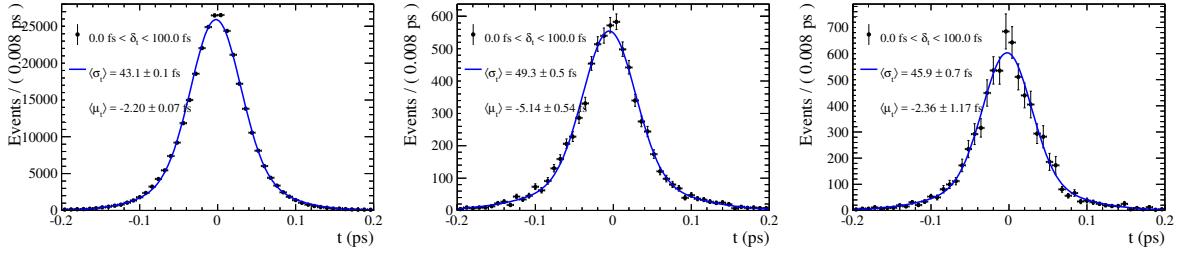
610 Table 6.2 summarizes the decay-time biases due to the selection and VELO misalign-  
 611 ment as well as the combined decay-time bias to be used in the resolution model for  
 612 different data-taking periods. It is verified on the  $B_s^0 \rightarrow D_s\pi$  MC sample (with VELO  
 613 misalignment) that the decay-time biases of these two sources add linearly.



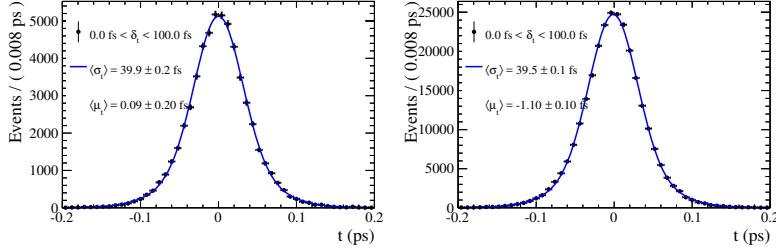
**Figure 6.6:** Comparison of selected variables for  $B_s \rightarrow D_s\pi\pi\pi$  signal data (black) and prompt  $D_s K\pi\pi$  data before (red) and after (blue) reweighting.



**Figure 6.7:** Reweighted decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks for 2016 (left) and 2017 (right) data.



**Figure 6.8:** Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s\pi$  MC candidates processed with a realistic VELO misalignment (left). Decay-time distribution of prompt  $D_s\pi$  MC candidates processed with a realistic VELO misalignment before (middle) and after (right) reweighting.



**Figure 6.9:** Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s\pi\pi\pi$  MC (Run-II) candidates after passing the stripping and trigger selection (left) and after full selection (right).

**Table 6.2:** Decay-time biases due to the selection and VELO misalignment as well as the total decay-time bias used in the resolution model.

	Decay-time bias [fs]		
	Selection	Alignment	Total
Run-1	$-1.20 \pm 0.11$	-	$-1.20 \pm 0.11$
15/16	$-1.10 \pm 0.10$	$-1.07 \pm 0.90 \pm 1.30$	$-2.17 \pm 1.68$
17	$-1.10 \pm 0.10$	$-1.17 \pm 0.77 \pm 1.58$	$-2.27 \pm 1.78$

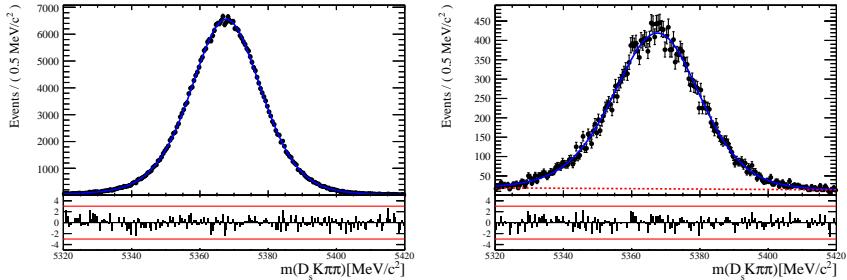
## 614 7 Acceptance

### 615 7.1 MC corrections

#### 616 7.1.1 Truth matching of simulated candidates

617 We use the `BackgroundCategory` tool to truth match our simulated candidates. Candidates  
 618 with background category (`BKGAT`) 20 or 50 are considered to be true signal. Background  
 619 category 60 is more peculiar since it contains both badly reconstructed signal candidates  
 620 and ghost background. This is due to the fact that the classification algorithms identifies  
 621 all tracks for which less than 70% of the reconstructed hits are matched to generated  
 622 truth-level quantities as ghosts. In particular for signal decays involving many tracks (as  
 623 in our case), this arbitrary cutoff is not 100% efficient. Moreover, the efficiency is expected  
 624 to depend on the kinematics which would lead to a biased acceptance determination if  
 625 candidates with `BKGAT`= 60 would be removed. We therefore include `BKGAT`= 60 and  
 626 statistically subtract the ghost background by using the `sPlot` technique. The `sWeights`  
 627 are calculated from a fit to the reconstructed  $B_s$  mass. The signal contribution is modeled  
 628 as described in Sec. 5.1 and the background with a polynomial. The fit is performed  
 629 simultaneously in two categories; the first includes events with `BKGAT` = 20 or 50 and  
 630 the second events with `BKGAT` = 60. To account for the different mass resolution we  
 631 use a different  $\sigma$  for each category, while the mean and the tail parameters are shared  
 632 between them. The background component is only included for the second category.

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



**Figure 7.1:** The reconstructed  $B_s$  mass distribution for simulated  $B_s \rightarrow D_s K\pi\pi$  decays classified with `BKGAT` = 20 or 50 (left) and `BKGAT` = 60 (right) after the full selection.

#### 635 7.1.2 Correction of data-simulation differences

636 For the evaluation of phase space efficiency and to a lesser extend also the decay-time  
 637 efficiency we rely on simulated data as discussed in the following sections. A number  
 638 of data-driven corrections are applied to the MC samples to account for known data-  
 639 simulation differences. The MC sample is reweighted to match the three-dimensional  
 640  $p_T$ ,  $\eta$  and track multiplicity distribution observed on real data. These corrections are  
 641 derived from the calibration channel  $B_s \rightarrow D_s \pi\pi\pi$  and applied to both the signal and  
 642 calibration channel MC samples. The distributions before and after reweighting are shown  
 643 in Appendix L. We use the `PIDCorr` tool to correct the simulated PID responses based on  
 644 PID calibration samples [34].

## 645 7.2 Decay-time acceptance

646 The decay-time distribution of the  $B_s^0$  mesons is sculpted due to the geometry of the LHCb  
 647 detector and the applied selection cuts, which are described in Section 4. In particular, any  
 648 requirement on the flight distance, the impact parameter or the direction angle (DIRA)  
 649 of the  $B_s^0$  mesons, as well as the direct cut on the proper-time, will lead to a decay-time  
 650 dependent efficiency  $\epsilon(t)$ .

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

$$661 \quad \mathcal{P}(t, \sigma_t) = \left[ e^{-\Gamma t} \cdot \cosh\left(\frac{\Delta\Gamma t'}{2}\right) \otimes \mathcal{R}(t - t', \sigma_t) \right] \cdot \epsilon(t), \quad (7.1)$$

662 is used to describe the decay-time distribution. For real data, the values for  $\Gamma_{s,d}$  and  
 663  $\Delta\Gamma_{s,d}$  are fixed to the latest HFLAV results [35], while for simulated data, the generated  
 664 values are used. A single Gaussian resolution function  $\mathcal{R}(t - t', \sigma_t)$  is used where the  
 665 decay-time error estimate is scaled with the respective calibration functions determined in  
 666 Sec. 6. Each decay-time acceptance  $\epsilon(t)$  is modeled using cubic splines, allowing for the  
 667 analytical computation of the decay-time integrals appearing in the PDF [36]. The splines  
 668 are parametrized by so-called knots  $(t_0, t_1, \dots, t_N)$  which determine their boundaries. Two  
 669 knots are located by default at the lower and upper edge of the interval allowed for the  
 670 decay time, the remaining ones are chosen such that there is an approximately equal  
 671 amount of data in-between two consecutive knots. In the basis of cubic b-splines,  $b_i(t)$ ,  
 the acceptance is then constructed as:

$$672 \quad \epsilon(t) = \sum_{i=0}^N v_i b_i(t) \quad (7.2)$$

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

$$675 \quad v_N = v_{N-1} + \frac{v_{N-2} - v_{N-1}}{t_{N-2} - t_{N-1}} \cdot (t_N - t_{N-1}). \quad (7.3)$$

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

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

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

682 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  
 683 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  
 684  $B_s^0 \rightarrow D_s K\pi\pi$  data is represented by  $\epsilon_{D_s K\pi\pi}^{Data}(t)$ .

685 The acceptances are determined separately for each data-taking period and each trigger  
 686 category as discussed in more detail in Appendix F. The fit results are shown in Figs. 7.2  
 687 to 7.5 and the fitted parameters are summarized in Tables 7.1 to 7.4.

**Table 7.1:** 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 7.2:** Time acceptance parameters for events in category [Run-I,L0-TIS].

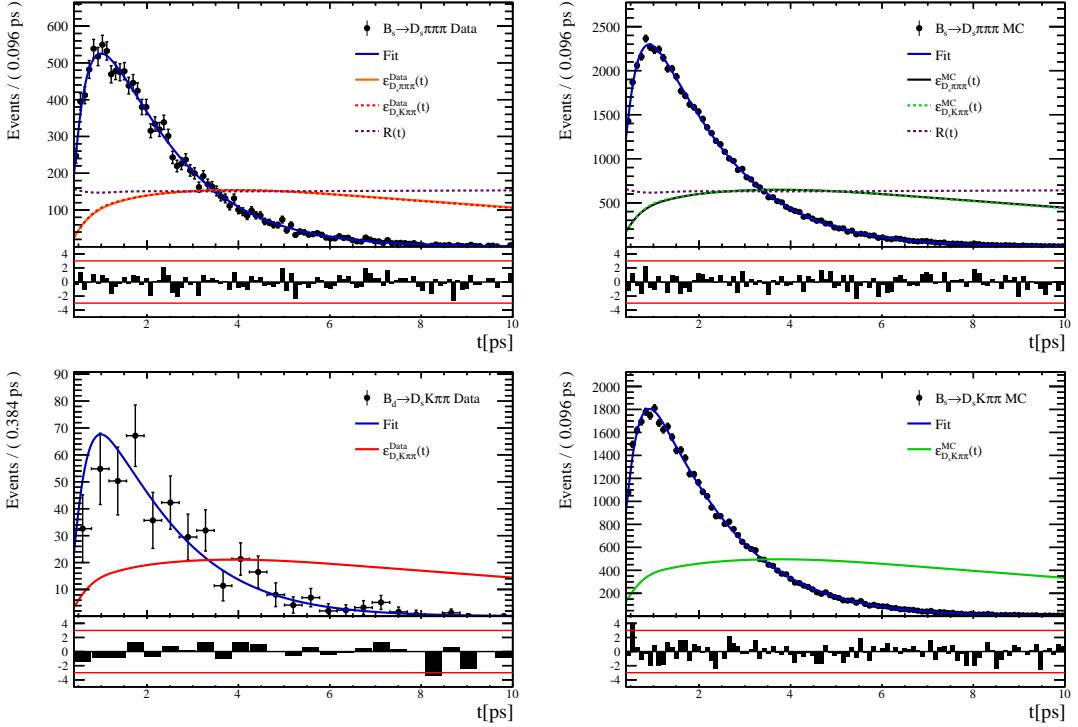
Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.158 \pm 0.014$	$0.216 \pm 0.005$	$0.986 \pm 0.040$
0.5	$v_1$	$0.422 \pm 0.029$	$0.524 \pm 0.010$	$0.965 \pm 0.029$
1.4	$v_2$	$0.802 \pm 0.047$	$0.860 \pm 0.017$	$0.982 \pm 0.029$
2.5	$v_3$	$1.099 \pm 0.034$	$1.098 \pm 0.011$	$1.002 \pm 0.019$
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 7.3:** Time acceptance parameters for events in category [Run-II,L0-TOS].

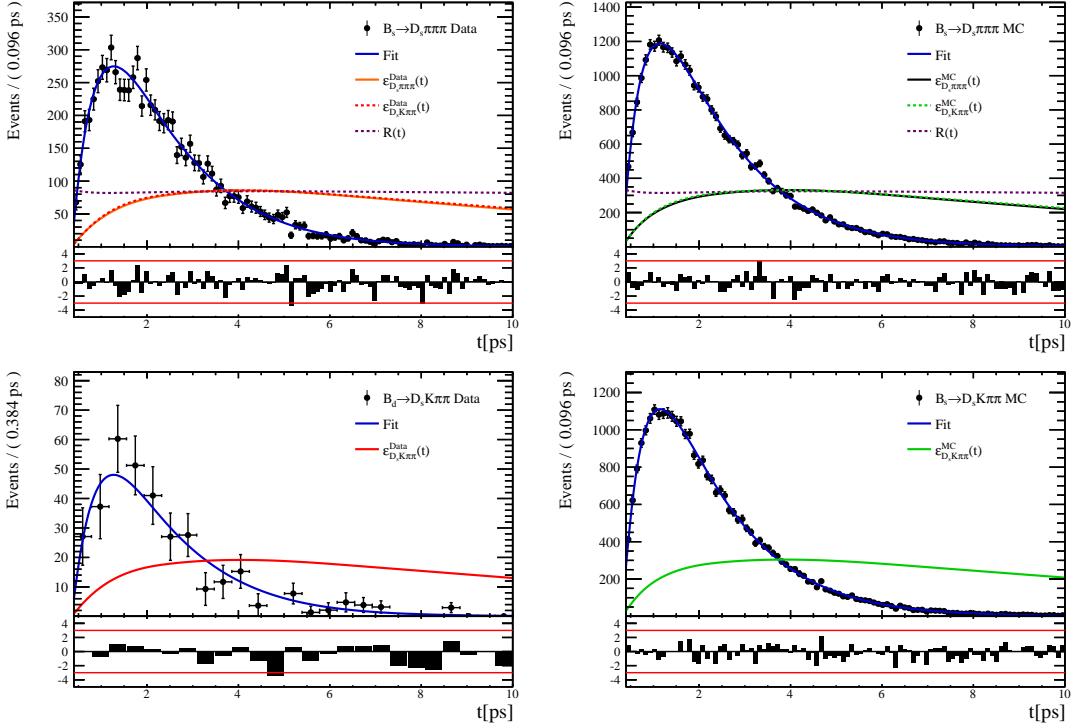
Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.285 \pm 0.009$	$0.368 \pm 0.005$	$1.023 \pm 0.020$
0.5	$v_1$	$0.663 \pm 0.017$	$0.749 \pm 0.009$	$0.911 \pm 0.016$
1.4	$v_2$	$0.856 \pm 0.025$	$0.893 \pm 0.012$	$1.016 \pm 0.019$
2.5	$v_3$	$1.060 \pm 0.017$	$1.071 \pm 0.008$	$0.996 \pm 0.013$
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 7.4:** Time acceptance parameters for events in category [Run-II,L0-TIS].

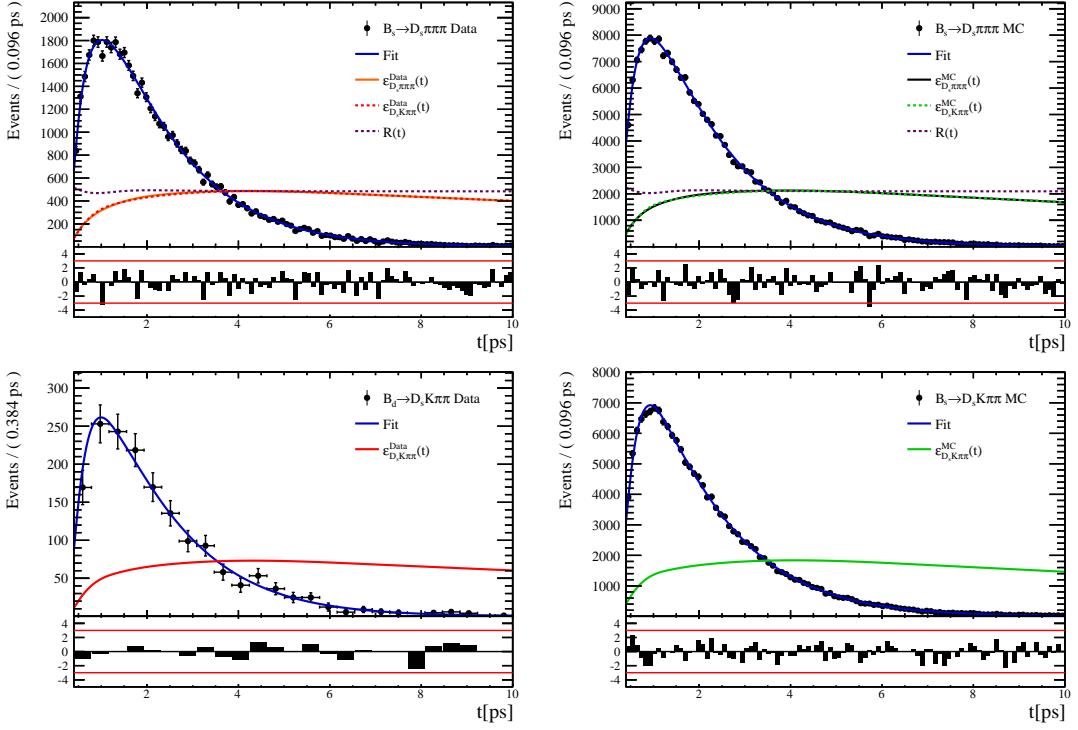
Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.117 \pm 0.008$	$0.171 \pm 0.003$	$0.965 \pm 0.034$
0.5	$v_1$	$0.422 \pm 0.019$	$0.474 \pm 0.008$	$0.952 \pm 0.024$
1.4	$v_2$	$0.733 \pm 0.027$	$0.777 \pm 0.013$	$0.973 \pm 0.025$
2.5	$v_3$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



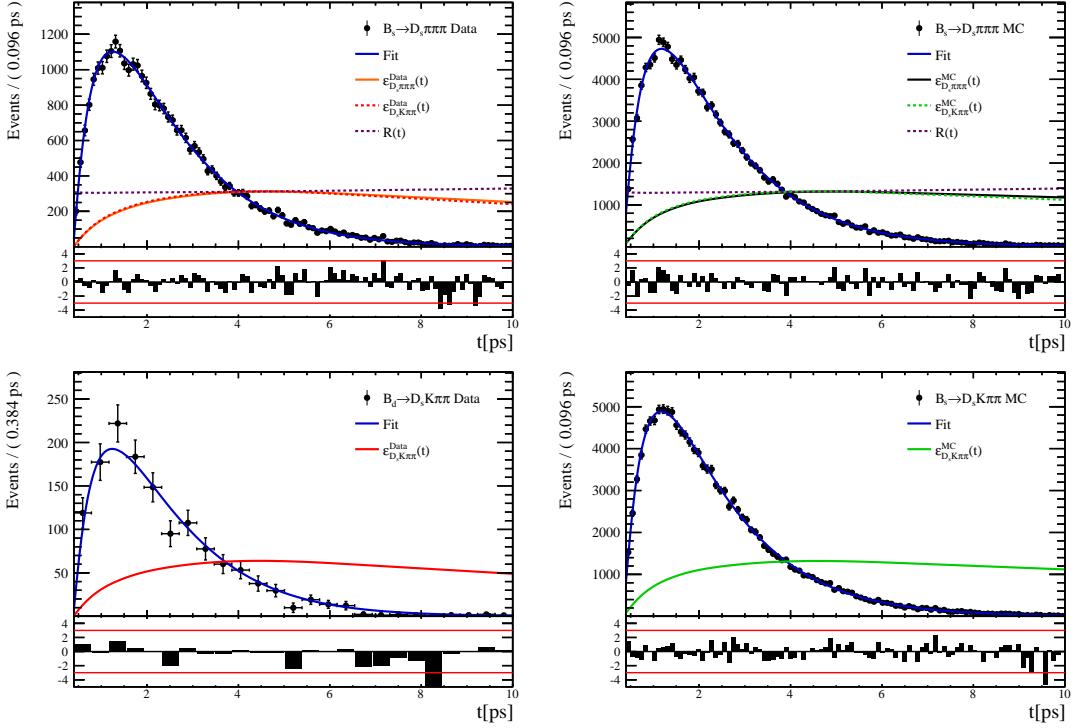
**Figure 7.2:** Decay-time fit projections for  $B_s^0 \rightarrow D_s\pi\pi\pi$  data (top-left),  $B_s^0 \rightarrow D_s\pi\pi\pi$  MC (top-right),  $B^0 \rightarrow D_s K\pi\pi$  data (bottom-left) and  $B_s^0 \rightarrow D_s K\pi\pi$  MC (bottom-right) in category [Run-I,L0-TOS]. The respective acceptance function is overlaid in an arbitrary scale.



**Figure 7.3:** 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 7.4:** 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 7.5:** 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.

### 688 7.3 Phase space acceptance

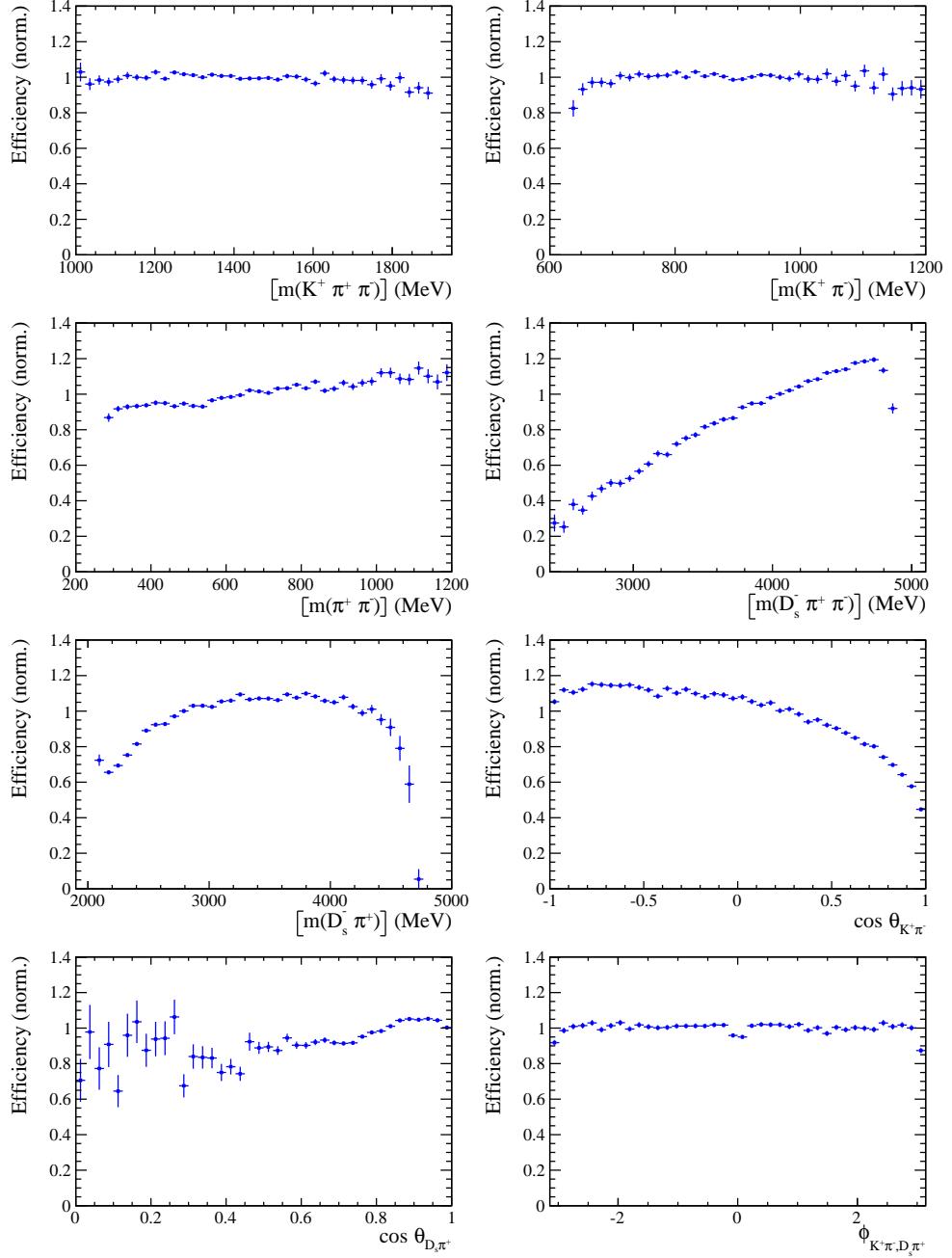
689 The signal PDF used for the full time-dependent amplitude fit can be written in terms of  
 690 the differential decay rate from Equation 2.28 as

$$\mathcal{P}(\mathbf{x}, t, g, f) = \frac{\left( \frac{d\Gamma(\mathbf{x}, t, q, f)}{dt d\Phi_4} \right) \cdot \epsilon(\mathbf{x}) \cdot \epsilon(t)}{\int \sum_{q,f} \left( \frac{d\Gamma(\mathbf{x}, t, q, f)}{dt d\Phi_4} \right) \cdot \epsilon(\mathbf{x}) \cdot \epsilon(t) dt d\Phi_4} \quad (7.4)$$

691 where  $\epsilon(\mathbf{x})$  is the phase-space efficiency. Note that the efficiency in the numerator appears  
 692 as an additive constant in the log  $\mathcal{L}$  that does not depend on any fit parameters such that it  
 693 can be ignored. However, the efficiency function still enters via the normalization integrals.  
 694 In contrast to the time integrals which can be performed analytically as discussed in  
 695 Sec. 7.2, the phase-space integrals are determined numerically. For this purpose, we use  
 696 simulated events generated with **EVTGEN**, pass them through the full detector simulation  
 697 and apply the same selection criteria as for data in order to perform the MC integrals. As  
 698 an example, the integral of the total  $b \rightarrow c$  amplitude squared can be approximated as

$$\int |\mathcal{A}_f^c(\mathbf{x})|^2 \epsilon(\mathbf{x}) d\Phi_4 \approx \frac{1}{N_{MC}} \sum_k^{N_{MC}} \frac{|\mathcal{A}_f^c(\mathbf{x}_k)|^2}{|A'(\mathbf{x}_k)|^2} \quad (7.5)$$

699 where  $A'$  labels the amplitude model used for the generation and  $x_k$  is the  $k$ -th MC  
 700 event. As a result, the phase-space efficiency can be included in the fit without explicitly  
 701 modeling it. The size of the fully selected MC sample ( $N_{MC} = 380k$ ) is more than 70 times  
 702 larger as the data sample which results in an integral precision smaller than 0.2%. The  
 703 efficiency projections are shown in Fig. 7.6 for visualization purposes only. As discussed  
 704 in Appendix G, the phase space efficiency differs significantly among L0-trigger categories  
 705 while the differences are small between the data-taking periods and negligible between  
 706 the  $D_s$  final states. To account for this, the MC events are scaled such that the relative  
 707 proportions of the four categories [Run-I,L0-TOS], [Run-I,L0-TIS], [Run-II,L0-TOS] and  
 708 [Run-II,L0-TIS] are the same as observed on the  $B_s \rightarrow D_s K\pi\pi$  data sample.



**Figure 7.6:** Efficiency variation as a function of the phase-space variables obtained from the ratio of selected and generated MC events.

## 709 8 Flavour Tagging

710 To identify the initial flavour state of the  $B_s^0$  meson, a number of flavour tagging algorithms  
 711 are used that either determine the flavour of the non-signal b-hadron produced in the  
 712 event (opposite site, OS [37]) or use particles produced in the fragmentation of the signal  
 713 candidate  $B_s^0/\bar{B}_s^0$  (same side, SS [38]). For the same side, the algorithm searching for the  
 714 charge of an additional kaon that accompanies the fragmentation of the signal candidate is  
 715 used (SS-Kaon). For the opposite site, five different taggers are chosen: the algorithms that  
 716 use the charge of an electron or a muon from semi-leptonic B decays (OS- $e,\mu$ ), the tagger  
 717 that uses the charge of a kaon from a  $b \rightarrow c \rightarrow s$  decay chain (OS-Kaon), the algorithm  
 718 which reconstructs opposite-side charm hadrons from a number of decay channels (OS-c)  
 719 and the algorithm that determines the  $B_s^0/\bar{B}_s^0$  candidate flavour from the charge of a  
 720 secondary vertex, reconstructed from the OS b decay product (OS-VtxCharge).

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

$$\epsilon_{eff} = \epsilon_{tag}(1 - 2\omega)^2, \quad (8.1)$$

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

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

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - \langle \eta \rangle), \quad (8.2)$$

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

## 741 8.1 OS tagger combination

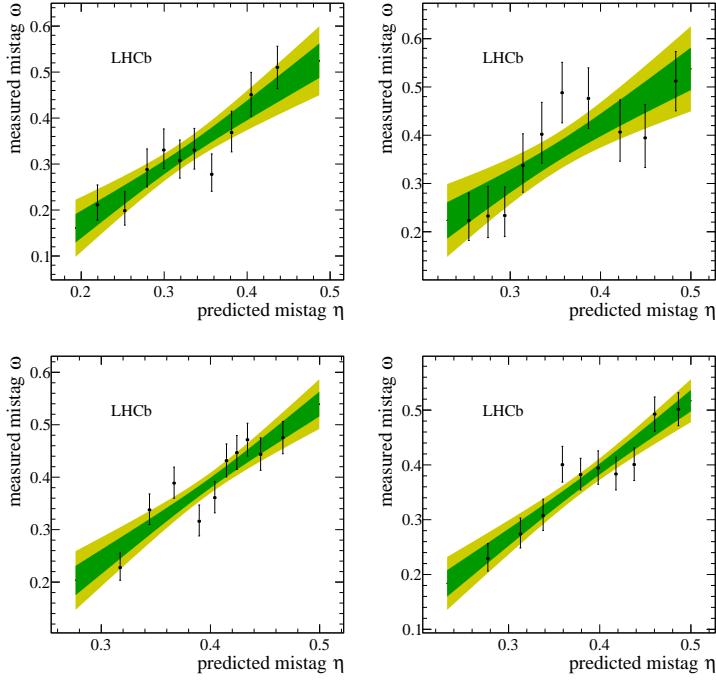
742 First, the OS electron, muon, kaon, charm and the secondary vertex charge taggers are  
 743 individually calibrated and then combined into a single OS-Combo tagger using the  
 744 **EspressoPerformanceMonitor** tool. We choose the flavour specific decay  $B_s \rightarrow D_s\pi\pi\pi$  as  
 745 calibration mode since it is very similar to the signal decay  $B_s \rightarrow D_s K\pi\pi$ . The calibration  
 746 is performed separately for Run-I and Run-II data. The OS-c tagger is not included for  
 747 Run-I data since the statistics is too low. Where available the latest Run-II tuning is used  
 748 for Run-II data, otherwise the Run-I tuning of the taggers is used. The calibration sample  
 749 is split in two subsamples. The single tagger calibration is performed for each and then  
 750 applied to the other sample. Figures 8.1 and 8.2 show the fitted calibration functions  
 751 and Tables 8.1 and 8.2 list the measured tagging performances (for both subsamples  
 752 combined). The tagging calibration parameters are listed in Appendix H.

**Table 8.1:** The flavour tagging performances for the used OS taggers for Run-I data.

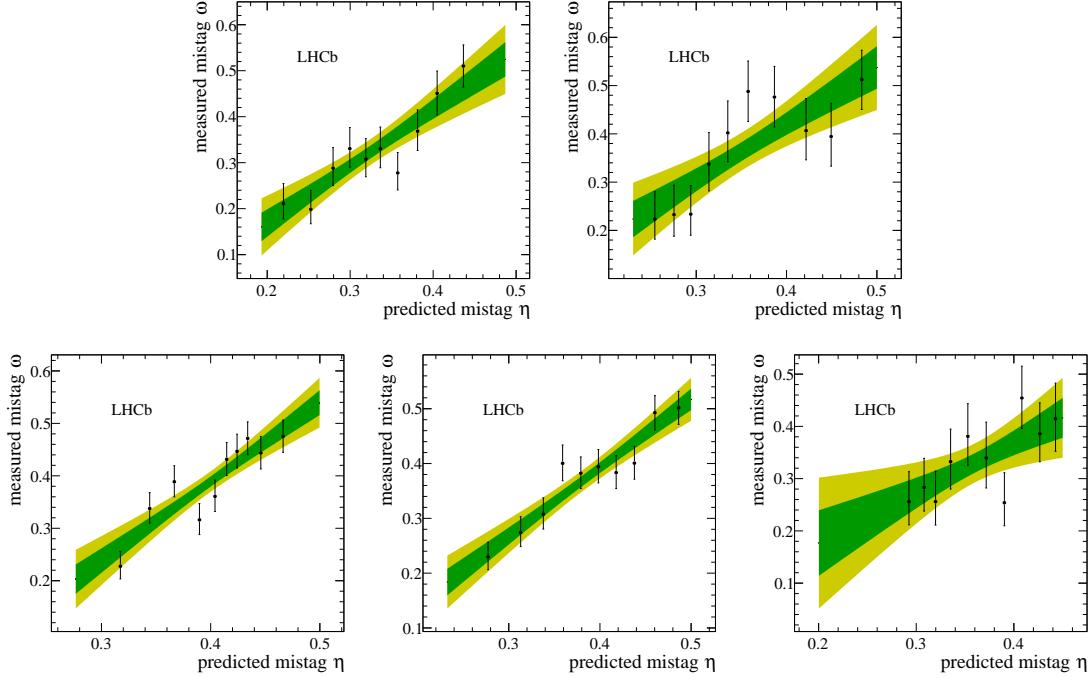
Tagger	$\epsilon$	$\omega$	$\epsilon\langle D^2 \rangle = \epsilon(1 - 2\omega)^2$
OS $\mu$	$(8.713 \pm 0.206)\%$	$(28.893 \pm 0.180(\text{stat}) \pm 2.291(\text{cal}))\%$	$(1.553 \pm 0.045(\text{stat}) \pm 0.337(\text{cal}))\%$
OS $e$	$(3.201 \pm 0.129)\%$	$(28.792 \pm 0.363(\text{stat}) \pm 3.611(\text{cal}))\%$	$(0.576 \pm 0.030(\text{stat}) \pm 0.196(\text{cal}))\%$
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 8.2:** 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 $c$	$(5.500 \pm 0.116)\%$	$(32.581 \pm 0.092(\text{stat}) \pm 1.848(\text{cal}))\%$	$(0.668 \pm 0.016(\text{stat}) \pm 0.142(\text{cal}))\%$



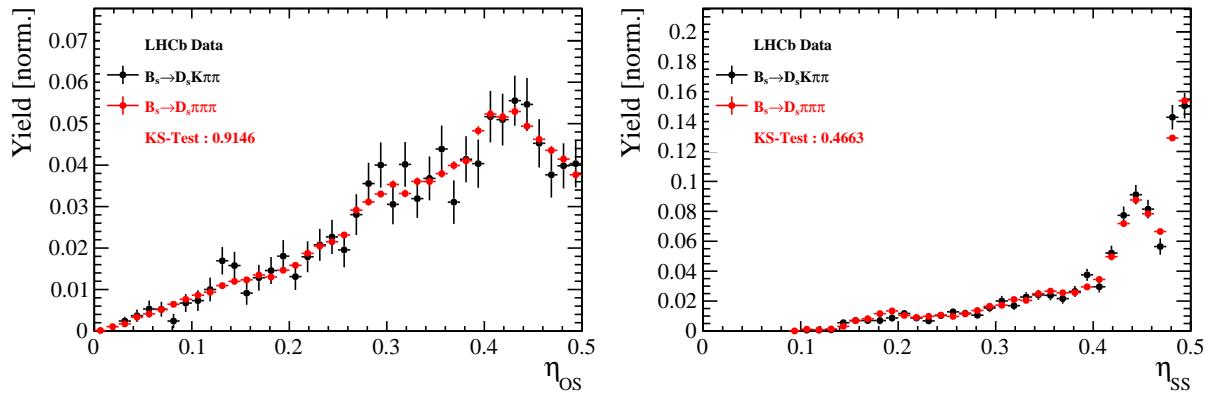
**Figure 8.1:** Predicted versus measured mistag probability for the (top left) OS muon, (top right) OS electron, (bottom left) OS kaon and (bottom right) OS vertex charge tagger for Run-I. A linear fit, including the  $1\sigma$  and  $2\sigma$  error bands is overlaid for each tagger.



**Figure 8.2:** 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.

## 753 8.2 Tagging performance

754 The OS-Combo and SS-Kaon taggers are calibrated simultaneously by fitting the  $B_s \rightarrow$   
 755  $D_s\pi\pi\pi$  decay-time distribution as discussed in Sec. 10. In this fit, the predicted mistag  
 756 probabilities  $\eta_{OS}$  and  $\eta_{SS}$ , shown Fig. 8.3 for  $B_s \rightarrow D_s\pi\pi\pi$  and  $B_s \rightarrow D_sK\pi\pi$  data, are  
 757 included as per-event observables, effectively giving a larger weight to the events that have  
 758 a lower mistag probability. The tagger responses are combined into a single response on  
 759 an event-by-event basis during the fit. Tables 8.3 and 8.4 report the tagging performances  
 760 for the OS and SS combination considering three mutually exclusive categories of tagged  
 761 events: OS only, SS only and both OS and SS. The tagging calibration parameters are  
 762 listed in Table 10.1.



**Figure 8.3:** 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 8.3:** 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 8.4:** 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	$81.68 \pm 0.19$	$39.28 \pm 0.40$	$6.57 \pm 0.21$

## 9 Production and Detection Asymmetries

### 9.1 $B_s$ Production Asymmetry

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

$$A_p(B_s^0) = \frac{\sigma(\bar{B}_s^0) - \sigma(B_s^0)}{\sigma(\bar{B}_s^0) + \sigma(B_s^0)} \quad (9.1)$$

where  $\sigma$  are the corresponding production cross-section. This asymmetry was measured by LHCb in  $pp$  collisions at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV by means of a time-dependent analysis of  $B_s \rightarrow D_s^- \pi^+$  decays [39]. The results in bins of  $p_T$  and  $\eta$  of the  $B_s$  meson are shown in Table 9.1. To correct for the different kinematics of  $B_s \rightarrow D_s^- \pi^+$  and  $B_s^0 \rightarrow D_s K\pi\pi$  decays, the measured  $B_s$  production asymmetries  $A_p(p_T, \eta)$  are folded with the sWeighted  $p_T, \eta$  distribution of our signal channel. The resulting effective production asymmetries are:

$$A_p(B_s^0)_{2011} = (-0.506 \pm 1.90)\% \quad (9.2)$$

$$A_p(B_s^0)_{2012} = (-0.164 \pm 1.30)\% \quad (9.3)$$

$$A_p(B_s^0)_{\text{Run-I}} = (-0.045 \pm 1.04)\%. \quad (9.4)$$

As for Run-II data no measurement is available yet, we determine the production asymmetry from  $B_s \rightarrow D_s \pi\pi\pi$  data together with the tagging parameters.

**Table 9.1:**  $B_s$  production asymmetries in kinematic bins for 2011 and 2012 data. [39]

$p_T$ [ GeV/c ]	$\eta$	$A_p(B_s^0)_{\sqrt{s}=7 \text{ TeV}}$	$A_p(B_s^0)_{\sqrt{s}=8 \text{ TeV}}$
(2.00, 7.00)	(2.10, 3.00)	$0.0166 \pm 0.0632 \pm 0.0125$	$0.0412 \pm 0.0416 \pm 0.0150$
(2.00, 7.00)	(3.00, 3.30)	$0.0311 \pm 0.0773 \pm 0.0151$	$-0.0241 \pm 0.0574 \pm 0.0079$
(2.00, 7.00)	(3.30, 4.50)	$-0.0833 \pm 0.0558 \pm 0.0132$	$0.0166 \pm 0.0391 \pm 0.0092$
(7.00, 9.50)	(2.10, 3.00)	$0.0364 \pm 0.0479 \pm 0.0068$	$0.0482 \pm 0.0320 \pm 0.0067$
(7.00, 9.50)	(3.00, 3.30)	$0.0206 \pm 0.0682 \pm 0.0127$	$0.0983 \pm 0.0470 \pm 0.0155$
(7.00, 9.50)	(3.30, 4.50)	$0.0058 \pm 0.0584 \pm 0.0089$	$-0.0430 \pm 0.0386 \pm 0.0079$
(9.50, 12.00)	(2.10, 3.00)	$-0.0039 \pm 0.0456 \pm 0.0121$	$0.0067 \pm 0.0303 \pm 0.0063$
(9.50, 12.00)	(3.00, 3.30)	$0.1095 \pm 0.0723 \pm 0.0179$	$-0.1283 \pm 0.0503 \pm 0.0171$
(9.50, 12.00)	(3.30, 4.50)	$0.1539 \pm 0.0722 \pm 0.0212$	$-0.0500 \pm 0.0460 \pm 0.0104$
(12.00, 30.00)	(2.10, 3.00)	$-0.0271 \pm 0.0336 \pm 0.0061$	$-0.0012 \pm 0.0222 \pm 0.0050$
(12.00, 30.00)	(3.00, 3.30)	$-0.0542 \pm 0.0612 \pm 0.0106$	$0.0421 \pm 0.0416 \pm 0.0162$
(12.00, 30.00)	(3.30, 4.50)	$-0.0586 \pm 0.0648 \pm 0.0150$	$0.0537 \pm 0.0447 \pm 0.0124$

## 777 9.2 $K^-\pi^+$ Detection Asymmetry

778 The presented measurement of the CKM-angle  $\gamma$  using  $B_s^0 \rightarrow D_s K\pi\pi$  decays is sensitive to  
 779 a possible charge asymmetry of the kaon. Kaons are known to have a nuclear cross-section  
 780 which is asymmetrically dependent on the sign of their charge. It is indispensable to  
 781 determine the charge asymmetry of the kaon, as fitting without taking this effect into  
 782 account would introduce a 'fake' CP violation. Instead of determining the single track  
 783 detection asymmetry of a kaon, it is found that the combined two track asymmetry of a  
 784 kaon-pion pair is much easier to access [40]. Therefore, the two track asymmetry defined  
 785 as

$$A^{det}(K^-\pi^+) = \frac{\epsilon^{det}(K^-\pi^+) - \epsilon^{det}(K^+\pi^-)}{\epsilon^{det}(K^-\pi^+) + \epsilon^{det}(K^+\pi^-)}, \quad (9.5)$$

786 is used. This asymmetry can be measured from the difference in asymmetries in the  
 787  $D^+ \rightarrow K^-\pi^+\pi^+$  and  $D^+ \rightarrow K_s^0\pi^+$  modes [41]:

$$\begin{aligned} A^{det}(K^-\pi^+) &= \frac{N(D^+ \rightarrow K^-\pi^+\pi^+) - N(D^- \rightarrow K^+\pi^-\pi^-)}{N(D^+ \rightarrow K^-\pi^+\pi^+) + N(D^- \rightarrow K^+\pi^-\pi^-)} \\ &\quad - \frac{N(D^+ \rightarrow K_s^0\pi^+) - N(D^- \rightarrow K_s^0\pi^-)}{N(D^+ \rightarrow K_s^0\pi^+) + N(D^- \rightarrow K_s^0\pi^-)} \\ &\quad - A(K^0), \end{aligned} \quad (9.6)$$

788 where possible CP violation in the  $D^+ \rightarrow K_s^0\pi^+$  mode is predicted to be smaller than  
 789  $10^{-4}$  in the Standard Model [42]. The small asymmetry in the neutral kaon system,  
 790  $A(\bar{K}^0) = -A(K^0) = (-0.054 \pm 0.014)\%$  [41], has to be taken into account as a correction.

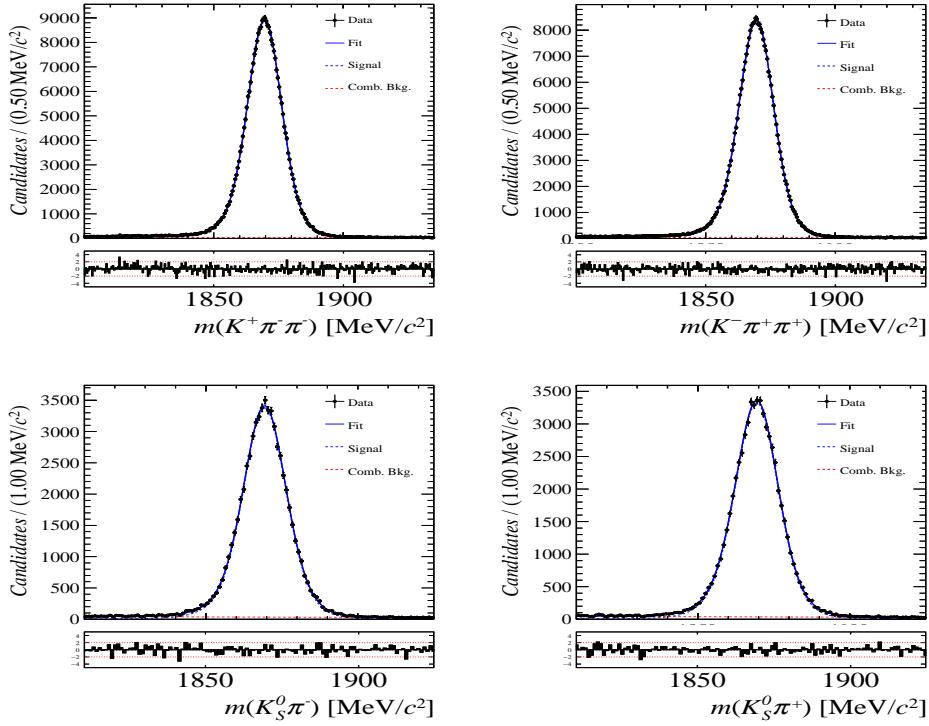
791 We use a dedicated LHCb tool to determine  $A^{det}(K^-\pi^+)$  for all data taking periods  
 792 used in this analysis. A detailed description can be found in [41]. The tool provides  
 793 large calibration samples of  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  and  $D^\pm \rightarrow K_s^0\pi^\pm$  decays, which are used  
 794 to determine the asymmetry following Eq. 9.6. Several weighting steps are performed  
 795 to match the kinematics of the calibration samples to our signal decay sample. First,  
 796 weights are assigned to the  $K^\pm$  and  $\pi^\pm$  of the  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  sample, using  $p, \eta$  of the  
 797  $K^\pm$  and  $p_T, \eta$  of the  $\pi^\pm$  from our  $B_s^0 \rightarrow D_s K\pi\pi$  signal decay. Then, weights are assigned  
 798 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 the corresponding,  
 799 weighted distributions of the  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  sample. In a last step, weights are assigned  
 800 to match the bachelor pions  $\phi$  distributions between the two calibration samples. After the  
 801 samples are weighted, fits are performed to the invariant  $m(K^-\pi^+\pi^+)/m(K^+\pi^-\pi^-)$  and  
 802  $m(K_s^0\pi^+)/m(K_s^0\pi^-)$  distributions to determine  $A^{det}(K^-\pi^+)$ . The PDFs used to describe  
 803 the invariant mass distributions consist of gaussian functions for the signal component  
 804 and exponentials describing the residual background.

805 The detection asymmetry is determined separately for every year and (since it is a  
 806 charge asymmetry effect) magnet polarity. Serving as an example for Run-I and Run-  
 807 II, the fits used to determine  $N(D^+ \rightarrow K^-\pi^+\pi^+)/N(D^- \rightarrow K^+\pi^-\pi^-)$  and  $N(D^+ \rightarrow$   
 $808 K_s^0\pi^+)/N(D^- \rightarrow K_s^0\pi^-)$  for 2011, magnet up data and 2015, magnet up data are shown  
 809 in Fig. 9.1 and 9.2 respectively. The obtained values of  $A^{det}(K^-\pi^+) + A(K^0)$  for all  
 810 years and polarities are shown in Table 9.2. Here, only the decay modes  $B_s^0 \rightarrow (D_s \rightarrow$   
 $811 KK\pi)K\pi\pi$  and  $B_s^0 \rightarrow (D_s \rightarrow \pi\pi\pi)K\pi\pi$  are considered. For  $B_s^0 \rightarrow (D_s \rightarrow K\pi\pi)K\pi\pi$ ,  
 812 the contributions of the two oppositely charged kaons cancel each other and a detection  
 813 asymmetry of zero with the same uncertainty as determined for  $D_s \rightarrow KK\pi$  and  $D_s \rightarrow \pi\pi\pi$

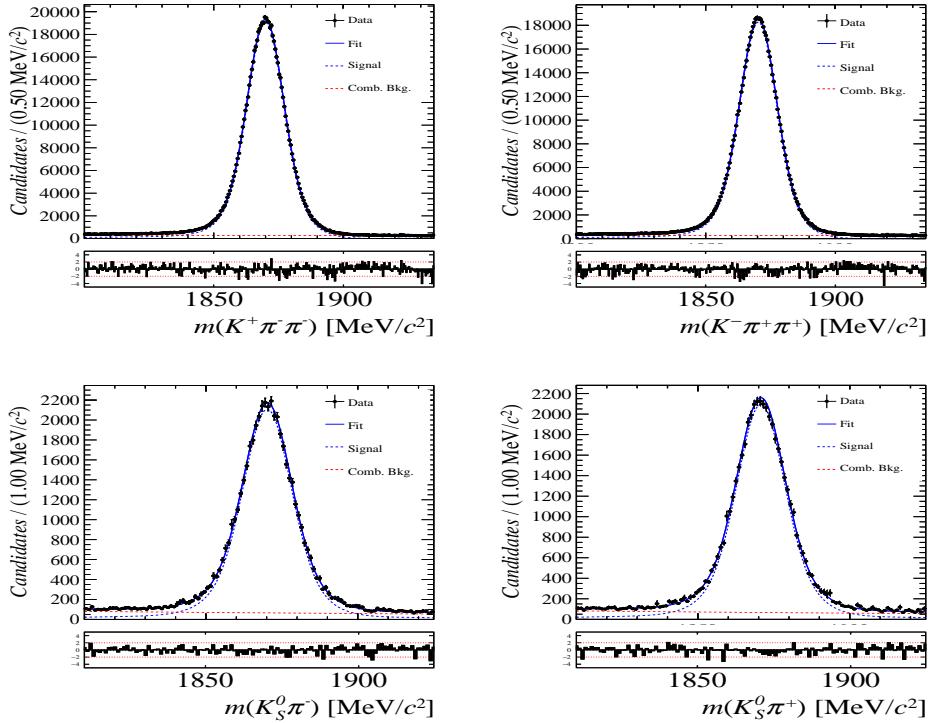
<sup>814</sup> is assumed. In contrast, only the decay mode  $D_s \rightarrow KK\pi$  is a source of detection  
<sup>815</sup> asymmetry for  $B_s^0 \rightarrow D_s\pi\pi\pi$  decays resulting in an average detection asymmetry of  
<sup>816</sup>  $A^{det}(K^-\pi^+) = -0.07 \pm 0.15\%$  (Run-I) and  $A^{det}(K^-\pi^+) = -0.08 \pm 0.21\%$  (Run-II).

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 9.2:** Summary of the  $K^-\pi^+$  detection asymmetries for  $B_s^0 \rightarrow D_s K\pi\pi$  decays obtained from the fits to the Run-I and Run-II calibration samples.



**Figure 9.1:** 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 9.2:** 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.

## 817 10 Decay-time fit

818 This section covers the (phase space integrated) decay-time fits to  $B_s^0 \rightarrow D_s h\pi\pi$  data. We  
 819 use the **sFit** technique [43] to statistically subtract the background, leaving only the signal  
 820 PDF to describe the decay-time. The **sWeights** are calculated based on the fit to the  
 821 reconstructed  $B_s$  mass distribution described in Sec. 5. The signal PDF is conditional on  
 822 the tagging decisions  $q_i$ , the mistag estimates  $\eta_i$  ( $i = \text{OS,SS}$ ) and the decay-time error  $\sigma_t$ :

$$\mathcal{P}(t|\sigma_t, q_{\text{OS}}, \eta_{\text{OS}}, q_{\text{SS}}, \eta_{\text{SS}}) \propto \left[ p(t'|q_{\text{OS}}, \eta_{\text{OS}}, q_{\text{SS}}, \eta_{\text{SS}}) \otimes \mathcal{R}(t - t', \sigma_t) \right] \cdot \epsilon(t) \quad (10.1)$$

823 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  
 824 account. The decay-time acceptance  $\epsilon(t)$  (Sec. 7) and the Gaussian time-resolution  
 825 function  $\mathcal{R}(t - t', \sigma_t)$  (Sec. 6) are fixed to the values obtained by the dedicated studies.  
 826 We fix the values of  $\Gamma_s$  and  $\Delta\Gamma_s$  to the latest HFLAV results [35].

827 The unbinned maximum likelihood fits are performed simultaneously in six categories:  
 828 [Run-I,L0-TOS], [Run-I,L0-TIS], [Year-15/16,L0-TOS], [Year-15/16,L0-TIS], [Year-17,L0-  
 829 TOS] and [Year-17,L0-TIS] to account for different time-acceptance shapes, time-resolution  
 830 and tagging calibrations.

### 831 10.1 Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data

832 Since the decay  $B_s^0 \rightarrow D_s\pi\pi\pi$  is flavour specific, the  $CP$  coefficients can be fixed to  $C = 1$   
 833 and  $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$ . The fit determines the calibration parameters for the  
 834 OS-Combo and SS-Kaon taggers, the  $B_s^0$  production asymmetry for Run-II data as well as  
 835 the mixing frequency  $\Delta m_s$ . Table 10.1 summarizes the fitted parameters. The **sWeighted**  
 836 decay-time distribution and the time-dependent asymmetry  $A_{mix}$  between mixed and  
 837 unmixed  $B_s^0$  candidates are shown in Fig. 10.1 along with the fit projections.

### 838 10.2 Fit to $B_s^0 \rightarrow D_sK\pi\pi$ data

839 The measured  $CP$  coefficients  $C, D_f, D_{\bar{f}}, S_f$  and  $S_{\bar{f}}$  extracted from a fit to the  $B_s \rightarrow$   
 840  $D_s K\pi\pi$  decay-time distribution are reported in Table 10.2. The fit projection is shown in  
 841 Fig. 10.2. We included a multi-dimensional Gaussian-constraint for the tagging calibration  
 842 parameters (including the tagging asymmetries) with the central values and covariance  
 843 matrix determined in Sec. 10.1.

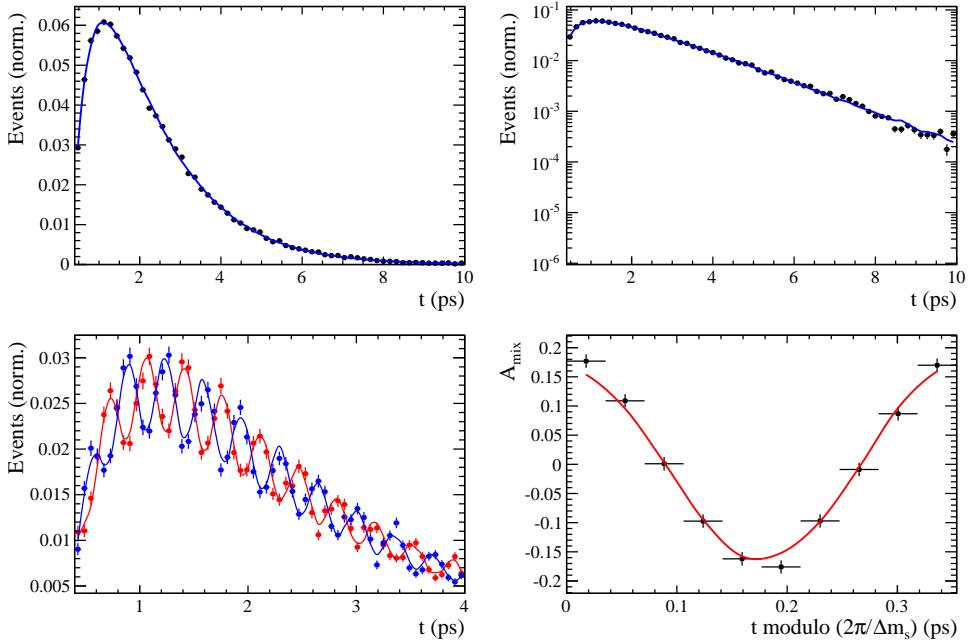
844 The  $CP$  coefficients are converted to the observables  $r, \kappa, \delta, \gamma$  using the GammaCombo  
 845 package. The corresponding confidence levels are shown in Fig 10.3.

**Table 10.1:** 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.

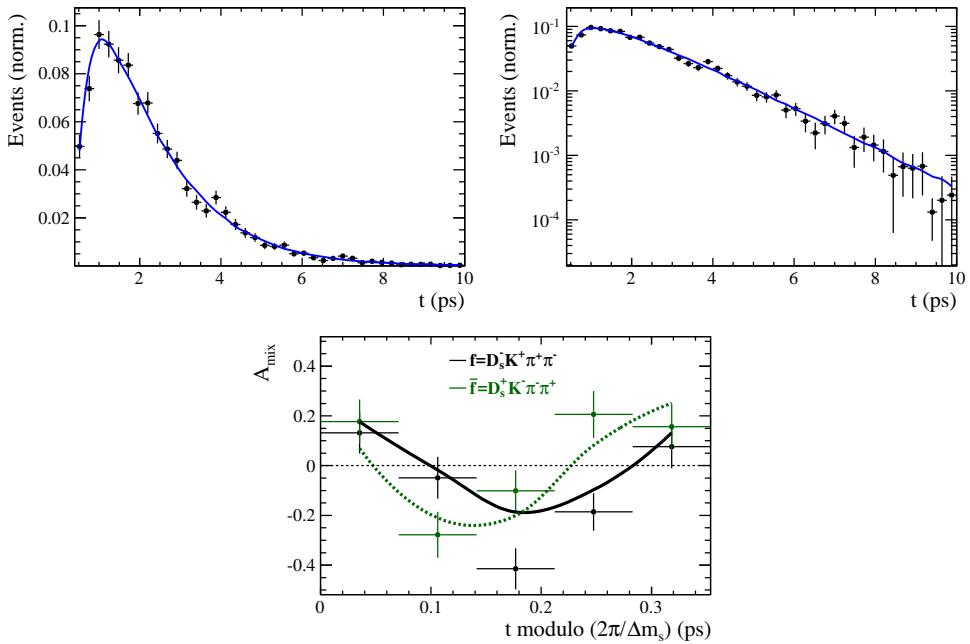
Fit Parameter	Run-I	Run-II
$p_0^{OS}$	$0.398 \pm 0.010 \pm 0.010$	$0.371 \pm 0.005 \pm 0.005$
$p_1^{OS}$	$0.894 \pm 0.085 \pm 0.089$	$0.784 \pm 0.043 \pm 0.026$
$\Delta p_0^{OS}$	$0.030 \pm 0.011 \pm 0.002$	$0.007 \pm 0.006 \pm 0.002$
$\Delta p_1^{OS}$	$0.011 \pm 0.095 \pm 0.023$	$0.065 \pm 0.053 \pm 0.009$
$\epsilon_{tag}^{OS} [\%]$	$47.775 \pm 0.365 \pm 0.080$	$40.413 \pm 0.182 \pm 0.036$
$\Delta \epsilon_{tag}^{OS} [\%]$	$-0.001 \pm 1.353 \pm 0.125$	$0.313 \pm 0.619 \pm 0.081$
$p_0^{SS}$	$0.444 \pm 0.008 \pm 0.005$	$0.428 \pm 0.004 \pm 0.002$
$p_1^{SS}$	$0.949 \pm 0.111 \pm 0.067$	$0.787 \pm 0.039 \pm 0.025$
$\Delta p_0^{SS}$	$-0.019 \pm 0.009 \pm 0.001$	$-0.017 \pm 0.004 \pm 0.000$
$\Delta p_1^{SS}$	$0.065 \pm 0.124 \pm 0.017$	$0.027 \pm 0.048 \pm 0.009$
$\epsilon_{tag}^{SS} [\%]$	$68.426 \pm 0.340 \pm 0.037$	$69.700 \pm 0.171 \pm 0.287$
$\Delta \epsilon_{tag}^{SS} [\%]$	$0.004 \pm 1.242 \pm 0.089$	$-0.374 \pm 0.578 \pm 0.121$
$A_P [\%]$	$-0.045$ (fixed)	$-0.256 \pm 0.644 \pm 0.168$
$\Delta m_s [\text{ps}^{-1}]$		$17.7537 \pm 0.0084 \pm 0.0092$

**Table 10.2:**  $CP$  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
$C$	$0.68 \pm 0.12 \pm 0.03$
$D$	$-0.00 \pm 0.32 \pm 0.12$
$\bar{D}$	$0.39 \pm 0.30 \pm 0.11$
$S$	$-0.15 \pm 0.17 \pm 0.03$
$\bar{S}$	$-0.53 \pm 0.17 \pm 0.03$



**Figure 10.1:** 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.



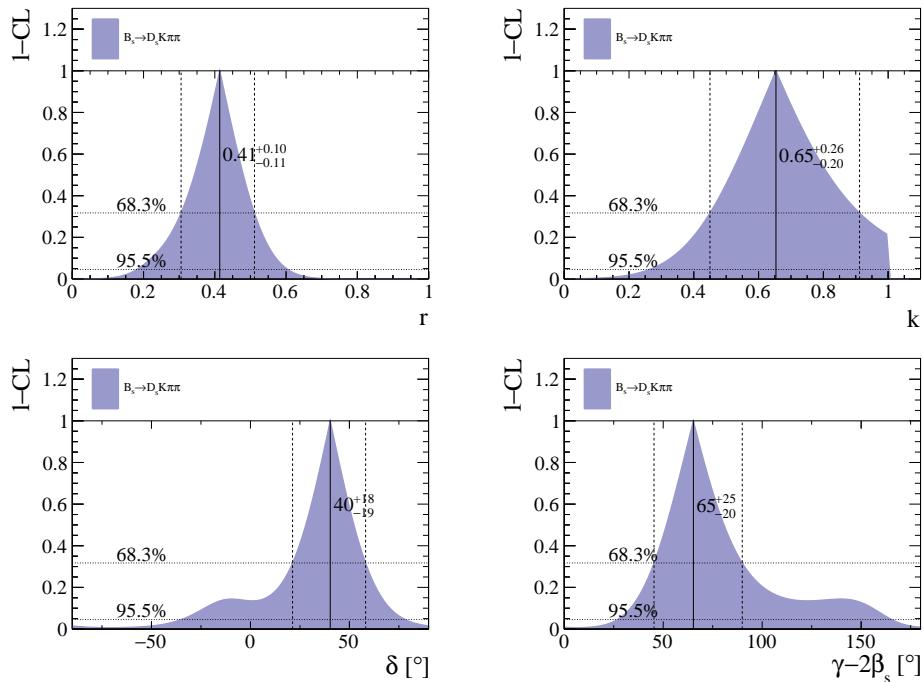
**Figure 10.2:** Decay-time distribution of  $B_s^0 \rightarrow D_s K\pi\pi$  signal candidates with the fit projection overlaid.

**Table 10.3:** Statistical correlation of the  $CP$  coefficients.

	$C$	$D$	$\bar{D}$	$S$	$\bar{S}$
$C$	1.000	-0.134	-0.181	0.048	0.074
$D$		1.000	0.542	0.019	0.060
$\bar{D}$			1.000	0.005	0.118
$S$				1.000	0.006
$\bar{S}$					1.000

**Table 10.4:** Systematic correlation of the  $CP$  coefficients.

	$C$	$D$	$\bar{D}$	$S$	$\bar{S}$
$C$	1.000	-0.057	-0.069	-0.109	0.102
$D$		1.000	0.206	0.022	0.040
$\bar{D}$			1.000	0.008	0.071
$S$				1.000	-0.342
$\bar{S}$					1.000



**Figure 10.3:** The 1-CL contours for the physical observable  $r$ ,  $\kappa$ ,  $\delta$  and  $\gamma$  obtained with the phasespace-integrated fit.

## 846 11 Time-dependent amplitude fit

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

$$\mathcal{P}(\mathbf{x}, t | \sigma_t, q_{os}, \eta_{os}, q_{ss}, \eta_{ss}) \propto [p(\mathbf{x}, t' | q_{os}, \eta_{os}, q_{ss}, \eta_{ss}) \otimes \mathcal{R}(t - t', \sigma_t)] \cdot \epsilon(t) \quad (11.1)$$

848 where  $p(\mathbf{x}, t | q_{os}, \eta_{os}, q_{ss}, \eta_{ss})$  is given the differential decay rate in Equation 2.28 taking  
849 the tagging dilution into account. The phase space efficiency  $\epsilon(\mathbf{x})$  is only included in the  
850 normalization of  $\mathcal{P}(\mathbf{x}, t | \sigma_t, q_{os}, \eta_{os}, q_{ss}, \eta_{ss})$  as discussed in Sec 7.3. The model selection  
851 of the amplitude components is described in the following Section. The remaining fitting  
852 strategy is exactly the same as for the decay-time fits, see Sec. 10.

### 853 11.1 Signal Model Construction

854 The light meson spectrum comprises multiple resonances which are expected to contribute  
855 to  $B_s \rightarrow D_s K\pi\pi$  decays as intermediate states. Apart from clear contributions coming  
856 from resonances such as  $K_1(1270)$ ,  $K_1(1400)$ ,  $\rho(770)$  and  $K^*(892)^0$ , the remaining structure  
857 is impossible to infer due to the cornucopia of broad, overlapping and interfering resonances  
858 within the phase space boundary. We follow the LASSO [44, 45] approach to limit the  
859 model complexity in two steps.

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

$$\mathcal{A}_f^{eff}(\mathbf{x}) = \sum_i a_i^{eff} A_i(\mathbf{x}) \quad (11.2)$$

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

$$|A_f^{eff}(\mathbf{x})|^2 = |A_f^c(\mathbf{x})|^2 + |A_f^u(\mathbf{x})|^2. \quad (11.3)$$

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

$$-2 \log \mathcal{L} \rightarrow -2 \log \mathcal{L} + \lambda \sum_i \sqrt{\int |a_i^{eff} A_i(\mathbf{x})|^2 d\Phi_4}, \quad (11.4)$$

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

$$BIC(\lambda) = -2 \log \mathcal{L} + r \log N_{Sig}, \quad (11.5)$$

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

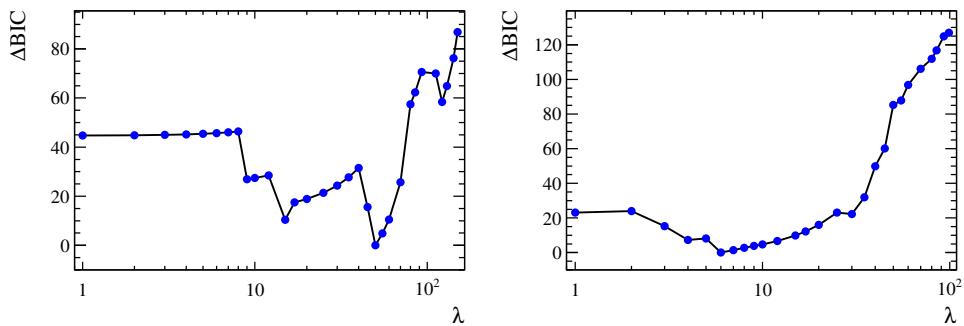
$$F_i \equiv \frac{\int |a_i^{eff} A_i(\mathbf{x})|^2 d\Phi_4}{\int |\mathcal{A}_f^{eff}(\mathbf{x})|^2 d\Phi_4}, \quad (11.6)$$

and are a measure of the relative strength between the different transitions. Figure 11.1(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 11.1. The fit projections are shown in Fig. 11.2. 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. 12.10.

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:

$$-2 \log \mathcal{L} \rightarrow -2 \log \mathcal{L} + \lambda \sum_i \sqrt{\int |a_i^c A_i(\mathbf{x})|^2 d\Phi_4} + \lambda \sum_i \sqrt{\int |a_i^u A_i(\mathbf{x})|^2 d\Phi_4} \quad (11.7)$$

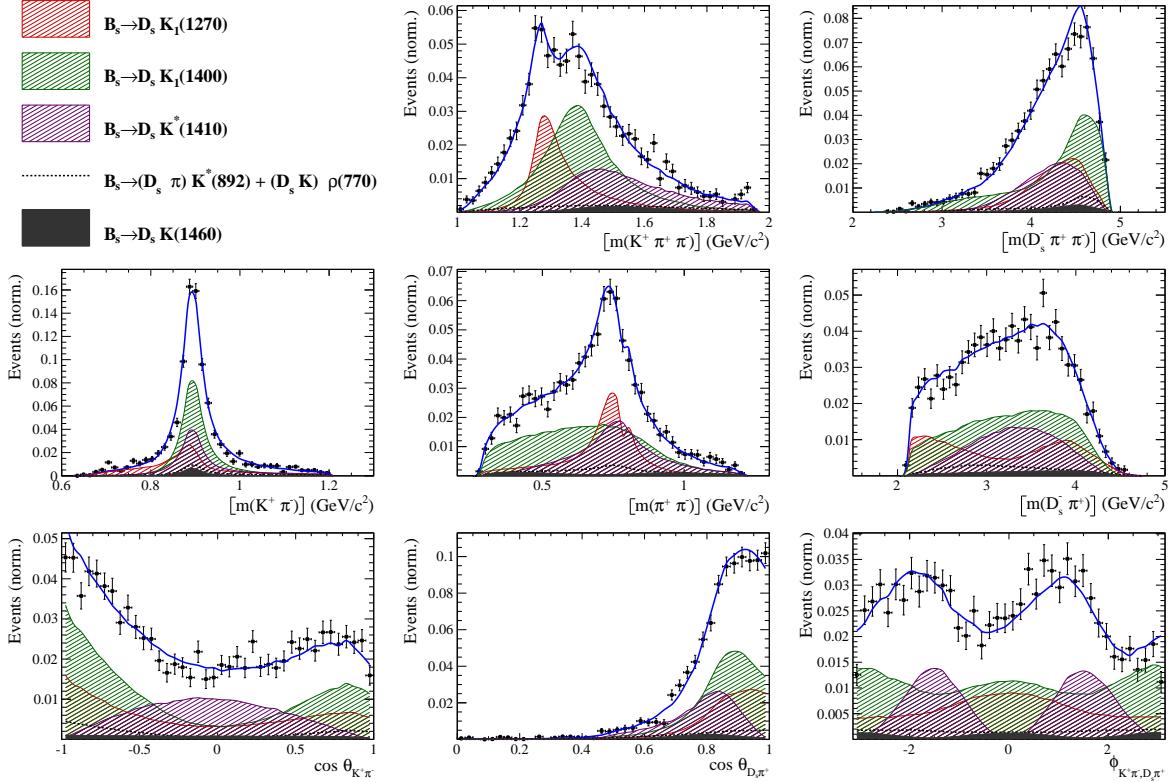
Figure 11.1(right) shows a plot of the complexity factor  $\lambda$ , against the resulting BIC values. The final set of  $b \rightarrow c$  and  $b \rightarrow u$  amplitudes is selected using the optimal value of  $\lambda = 6$ , and is henceforth called the LASSO model. The parameters of chosen resonances are summarized in Appendix K.



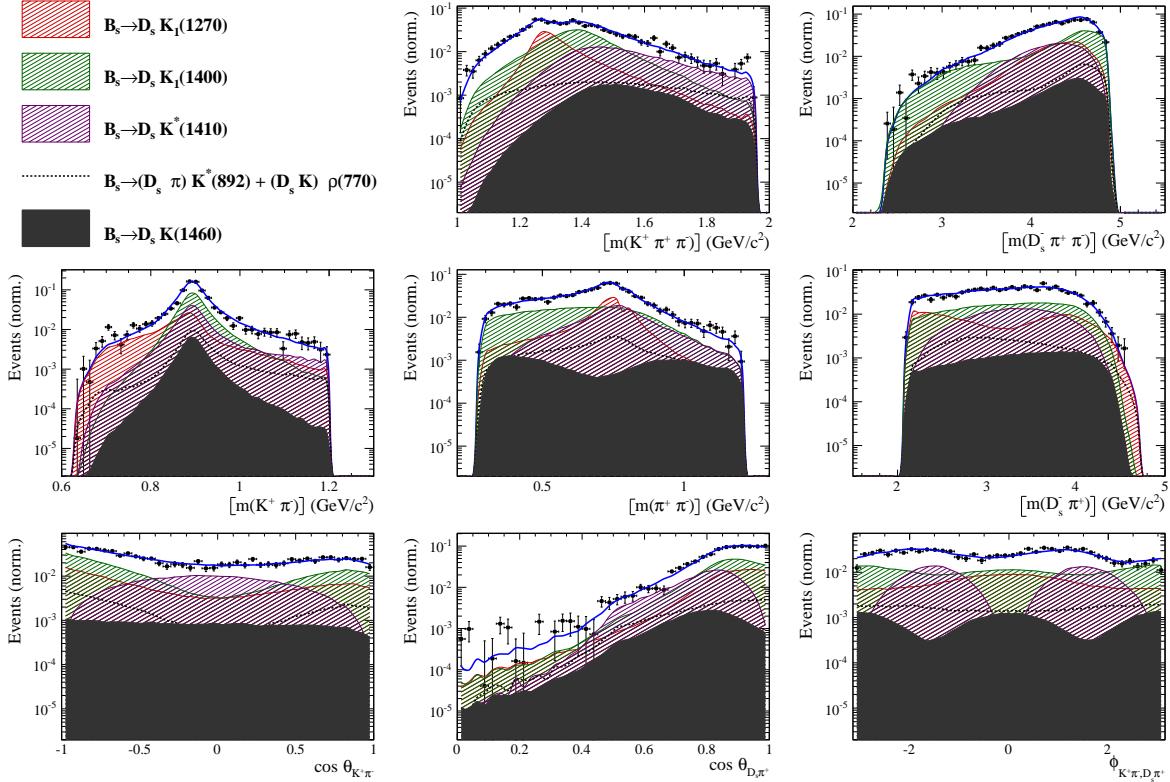
**Figure 11.1:** Difference in the BIC value from its minimum as function of the LASSO parameter  $\lambda$  for step 1 (left) and step 2 (right) of the model selection.

**Table 11.1:** 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 11.2:** 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. The incoherent contributions of the selected decay channels are plotted as well.



**Figure 11.3:** 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. The incoherent contributions of the selected decay channels are plotted as well.

## 888 11.2 Results

889 Table 11.2 lists the modulus and phases of the complex amplitude coefficients  $a_i^c$  and  $a_i^u$ ,  
 890 obtained by fitting the LASSO model to the data. The corresponding fit fractions for the  
 891  $b \rightarrow c$  and  $b \rightarrow u$  amplitudes are individually normalized

$$F_i^{c,u} \equiv \frac{\int |a_i^{c,u} A_i(\mathbf{x})|^2 d\Phi_4}{\int |\mathcal{A}_f^{c,u}(\mathbf{x})|^2 d\Phi_4} \quad (11.8)$$

892 and shown in Table 11.3. The statistical significances of the three smallest contributions  
 893 are:  $4.5\sigma$  for  $B_s \rightarrow (D_s K)_P \rho(770)$ ;  $6.2\sigma$  for  $B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$  and  
 894  $8.2\sigma$  for  $B_s \rightarrow (D_s \pi)_P K^*(892)$ . These are calculated from the likelihood difference  
 895  $\sqrt{\Delta(-2LL)}$  between a fit with and without the respective component assuming Wilk's  
 896 theorem. In addition to the amplitude coefficients, the amplitude ratio and the strong and  
 897 weak phase differences between the  $b \rightarrow c$  and  $b \rightarrow u$  decays are determined. Moreover,  
 898 the masses and widths of the  $K_1(1400)$  and  $K^*(1410)$  resonances are fitted.

899 Figure 11.4 shows the distributions of selected phase space observables, which demon-  
 900 strate reasonable agreement between data and the fit model. We also project into the  
 901 transversity basis to demonstrate good description of the overall angular structure (see  
 902 11.4 bottom row). The acoplanarity angle  $\chi$ , is the angle between the two decay planes  
 903 formed by the  $K^+ \pi^-$  system and the  $D_s^- \pi^+$  system in the  $B_s$  rest frame; boosting into the  
 904 rest frames of the two-body systems defining these decay planes, the two helicity variables  
 905 are defined as the cosine of the angle,  $\theta$ , of the  $K^+$  or  $D_S^-$  momentum with the  $B_s$  flight  
 906 direction.

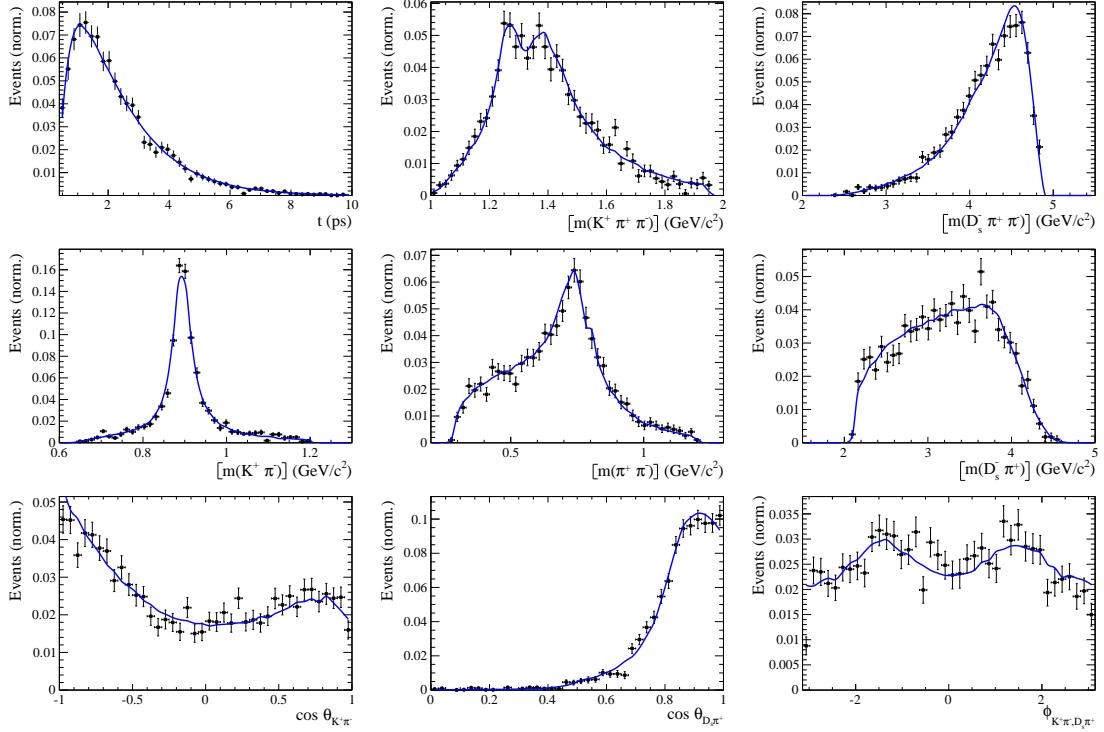
907 In order to quantify the quality of the fit in the five-dimensional phase space, a  $\chi^2$   
 908 value is determined by binning the data;

$$\chi^2 = \sum_{b=1}^{N_{\text{bins}}} \frac{(N_b - N_b^{\text{exp}})^2}{N_b^{\text{exp}}}, \quad (11.9)$$

909 where  $N_b$  is the number of data events in a given bin,  $N_b^{\text{exp}}$  is the event count predicted  
 910 by the fitted PDF and  $N_{\text{bins}}$  is the number of bins. An adaptive binning is used to ensure  
 911 sufficient statistics in each bin for a robust  $\chi^2$  calculation [47]. At least 25 events per  
 912 bin are required. The number of degrees of freedom  $\nu$ , in an unbinned fit is bounded by  
 913  $N_{\text{bins}} - 1$  and  $(N_{\text{bins}} - 1) - N_{\text{par}}$ , where  $N_{\text{par}}$  is the number of free fit parameters. We use  
 914 the  $\chi^2$  value divided by  $\nu = (N_{\text{bins}} - 1) - N_{\text{par}}$  as a conservative estimate. For the LASSO  
 915 model, this amounts to  $\chi^2/\nu = 1.40$  with  $\nu = 137$ , indicating a decent fit quality.

**Table 11.2:** 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
$K_1(1270) \rightarrow K^*(892) \pi$	$0.72 \pm 0.11 \pm 0.09$	$50.1 \pm 8.4 \pm 5.9$		
$K_1(1270) \rightarrow K_0^*(1430) \pi$	$0.52 \pm 0.04 \pm 0.06$	$128.9 \pm 5.0 \pm 18.0$		
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$1.98 \pm 0.08 \pm 0.19$	$10.5 \pm 9.3 \pm 6.2$	$0.74 \pm 0.20 \pm 0.16$	$-64.3 \pm 12.9 \pm 13.2$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$1.14 \pm 0.08 \pm 0.09$	$55.0 \pm 4.6 \pm 5.1$		
$K^*(1410) \rightarrow K \rho(770)$	$0.63 \pm 0.05 \pm 0.04$	$-164.1 \pm 5.3 \pm 3.2$		
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$			$0.87 \pm 0.10 \pm 0.10$	$-96.3 \pm 11.5 \pm 11.3$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$0.72 \pm 0.10 \pm 0.13$	$-17.2 \pm 12.4 \pm 13.0$	$1.13 \pm 0.12 \pm 0.15$	$-16.8 \pm 14.3 \pm 16.8$
$B_s \rightarrow (D_s K)_P \rho(770)$			$0.53 \pm 0.08 \pm 0.07$	$33.6 \pm 10.6 \pm 11.2$
Fit parameter	Value			
$m_{K_1(1400)}$ [MeV]	$1397 \pm 10 \pm 5 \pm 7$			
$\Gamma_{K_1(1400)}$ [MeV]	$205 \pm 15 \pm 11 \pm 8$			
$m_{K^*(1410)}$ [MeV]	$1432 \pm 10 \pm 17 \pm 8$			
$\Gamma_{K^*(1410)}$ [MeV]	$345 \pm 25 \pm 37 \pm 17$			
$r$	$0.50 \pm 0.05 \pm 0.03 \pm 0.02$			
$\delta [^\circ]$	$46 \pm 14 \pm 7 \pm 8$			
$\gamma - 2\beta_s [^\circ]$	$61 \pm 15 \pm 6 \pm 6$			



**Figure 11.4:** Projections of the full time-dependent amplitude fit.

**Table 11.3:** Fit fractions of the amplitudes contributing to  $b \rightarrow c$  and  $b \rightarrow u$  decays.

Decay Channel	$F_{b \rightarrow c} [\%]$	$F_{b \rightarrow u} [\%]$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$6.3 \pm 1.6 \pm 1.7 \pm 0.8$	$14.9 \pm 4.5 \pm 4.6 \pm 4.3$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$12.3 \pm 1.4 \pm 1.4 \pm 1.9$	$29.1 \pm 6.1 \pm 6.6 \pm 6.2$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$3.4 \pm 0.5 \pm 0.9 \pm 0.5$	$8.0 \pm 2.1 \pm 2.5 \pm 1.7$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$48.3 \pm 4.5 \pm 7.2 \pm 4.5$	$17.2 \pm 8.6 \pm 7.6 \pm 6.4$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$15.5 \pm 1.0 \pm 1.4 \pm 1.1$	
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$6.7 \pm 0.6 \pm 0.8 \pm 0.5$	
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$		$21.0 \pm 4.6 \pm 4.2 \pm 5.1$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$6.8 \pm 1.5 \pm 2.4 \pm 2.3$	$36.0 \pm 8.0 \pm 10.9 \pm 13.2$
$B_s \rightarrow (D_s K)_P \rho(770)$		$9.7 \pm 4.0 \pm 4.2 \pm 3.5$
Sum	$99.3 \pm 4.7 \pm 7.2 \pm 4.0$	$135.9 \pm 12.9 \pm 13.5 \pm 11.7$

## 916 12 Systematic uncertainties

917 The systematic uncertainties on the measured observables are summarized in Table 12.3 for  
918 the decay-time fit to  $B_s \rightarrow D_s \pi\pi\pi$ , in Table 12.4 for the decay-time fit to  $B_s \rightarrow D_s K\pi\pi$   
919 and in Table 12.5 for the full time-dependent amplitude fit to  $B_s \rightarrow D_s K\pi\pi$  decays. A  
920 description of each systematic effect is given in the following subsections starting with the  
921 ones common to all fits. Afterwards, systematic effect specific to the amplitude description  
922 are discussed. Systematic uncertainties estimated from pseudo-experiments (toys) are  
923 calculated relative to the statistical error.

### 924 12.1 Fit bias

925 Pseudo-experiments are performed, where a signal toy sample of the same size as the  
926 number of observed signal data events is generated according to the nominal fit model and  
927 subsequently fitted with the same model. The means of the pull distributions are taken  
928 as systematic uncertainties of the fit parameters. The width of the pull distributions are  
929 compatible with unity.

930 As the determination of the mixing frequency is a high precision measurement, a  
931 closure test is performed to further validate the fit procedure. A fit to the decay-time  
932 distribution of the high-statistics fully simulated signal  $B_s \rightarrow D_s \pi\pi\pi$  and  $B_s \rightarrow D_s K\pi\pi$   
933 candidates (using the TRUEID of the candidates for the tagging information) is performed  
934 to extract the mixing frequency resulting in  $\Delta m_s = 17.8027 \pm 0.0009 \text{ ps}^{-1}$  and  $\Delta m_s =$   
935  $17.8017 \pm 0.0009 \text{ ps}^{-1}$ , respectively. The larger shift with respect to the generated value of  
936  $\Delta m_s = 17.800 \text{ ps}^{-1}$  is assigned as additional systematic uncertainty (even though it could,  
937 at least to some extent, be explained by statistical fluctuations of the MC sample).

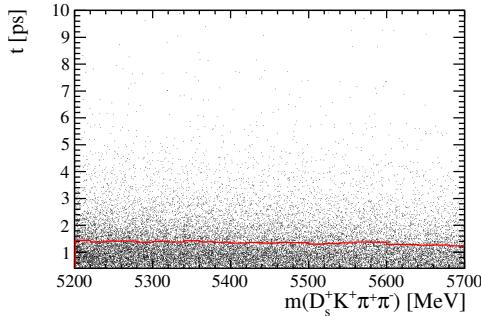
### 938 12.2 Background subtraction

939 The statistical subtraction of the residual background [43], left after the full selection,  
940 relies on the correct description of the invariant  $B_s^0$  mass distribution. Since the choice of  
941 signal and background models is not unique, alternative parameterizations are tested:

- 942 • The Johnson's SU function which is used as nominal signal model is replaced by the  
943 sum of two Crystal Ball functions [48].
- 944 • For the combinatorial background, the nominal second order polynomial is replaced  
945 by an exponential function.
- 946 • For the description of the partially reconstructed background, a combination of the  
947 RooHILLdini and RooHORNsdini model [49] is used instead of the nominal model of  
948 three bifurcated gaussians.
- 949 • For the shape of the mis-ID background, the nominal approach is to use a simulated  
950 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  
951 hypothesis of the  $\pi^+$  with the higher misidentification probability (see Sec. 5).  
952 Two alternative approaches are considered: we flip the mass hypothesis of the  $\pi^+$   
953 candidate with the lower probability of being misidentified; we randomly flip the  
954 mass hypothesis of a  $\pi^+$  candidate. To account for the systematic of the  $\pi \rightarrow K$   
955 fake rate, we vary the default PIDK cut ( $\text{PIDK} > 10$ ) by  $\pm 2$  when calculating the  
956 misID weights from PIDCalib.

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

The subtraction of the background via the sPlot technique relies on the assumption of  
 no correlation between the reconstructed mass and the observables that are used in the  
 final fit. To assess the impact of ignoring the small correlation between reconstructed mass  
 and decay-time ( $\rho \approx -0.04$ ) observed for the combinatorial background, we perform a toy  
 study including signal and background candidates. The two-dimensional reconstructed  
 mass vs. decay-time distribution for background candidates, shown in Fig. 12.1, is taken  
 from a (same-sign)  $D_s^+ K^+ \pi^+ \pi^-$  data sample. For simplicity, a flat phase-space distribution  
 is assumed for the background. The toy study including the small correlation in the  
 generation of the background candidates is contrasted to one where uncorrelated masses  
 and decay-times are generated. The difference of the pull means is assigned as systematic  
 for each fit parameter.



**Figure 12.1:** Scatter plot of reconstructed mass against decay-time for (same-sign)  $D_s^+ K^+ \pi^+ \pi^-$  candidates. The average decay-time in a given mass bin is overlaid in red.

### 12.3 Decay-time acceptance

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

To improve the coverage of the multi-dimensional parameter space, a Cholesky decom-  
 position [50] is used to generate a set of uncorrelated vectors from the covariance matrix  
 $\text{cov}(\lambda_i, \lambda_j)$ , where the vector  $\lambda$  includes the parameters  $\Gamma_s$ ,  $\Delta\Gamma_s$  and the  $N = 4$  spline  
 coefficients for each category of the simultaneous fit. The correlations between  $\Gamma_s$  ( $\Delta\Gamma_s$ )  
 and the spline coefficients are measured by rerunning the acceptance fits described in  
 Sec. 7.2 with the values of  $\Gamma_s$  ( $\Delta\Gamma_s$ ) varied by  $\pm 1\sigma$  and measuring the shift in the spline  
 coefficients as a fraction of their uncertainty. For the correlation between  $\Gamma_s$  and  $\Delta\Gamma_s$  we  
 use the HFLAV value [35].

## 990 12.4 Decay-time resolution and tagging

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

- 994 • A double Gaussian is fit to the decay-time distributions of fake  $B_s^0$  candidates, but  
 995 only the width of the core Gaussian is considered to represent the time resolution in  
 996 the respective bin. Therefore the resolution is slightly underestimated in this case.
- 997 • A single Gaussian is fit to the decay-time distributions of fake  $B_s^0$  candidates in a  
 998 wide range of  $[-3\sigma_t : 1.5\sigma_t]$ . Due to the tails of the distribution, which broaden the  
 999 width of the Gaussian function, this method slightly overestimates the decay-time  
 1000 resolution.

1001 For each case, a new scaling function is derived:

$$\sigma_{eff,Data,16}^{core-Gauss}(\sigma_t) = (6.8 \pm 1.9) \text{ fs} + (0.797 \pm 0.050) \sigma_t \quad (12.1)$$

$$\sigma_{eff,Data,16}^{single-Gauss}(\sigma_t) = (10.1 \pm 1.5) \text{ fs} + (0.950 \pm 0.039) \sigma_t \quad (12.2)$$

$$\sigma_{eff,Data,17}^{core-Gauss}(\sigma_t) = (0.1 \pm 1.5) \text{ fs} + (0.957 \pm 0.037) \sigma_t \quad (12.3)$$

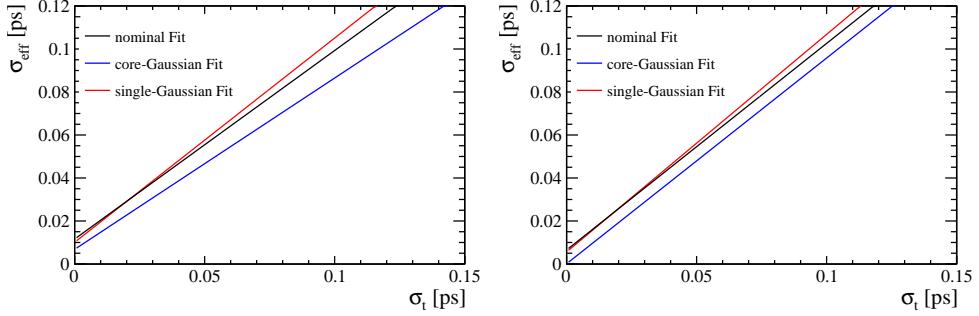
$$\sigma_{eff,Data,17}^{single-Gauss}(\sigma_t) = (5.6 \pm 1.2) \text{ fs} + (1.012 \pm 0.031) \sigma_t \quad (12.4)$$

1005 They are compared to the nominal result in Fig. 12.2.

1006 Due to the high correlation between the decay-time resolution and the tagging calibration,  
 1007 their systematic uncertainty has to be studied simultaneously. First, the decay-time  
 1008 fits to  $B_s \rightarrow D_s \pi \pi \pi$  data are repeated using the alternative decay-time error scaling  
 1009 functions. New tagging calibration parameters are obtained which are then used (together  
 1010 with the respective decay-time error scaling function) in the fits to  $B_s \rightarrow D_s K \pi \pi$  data to  
 1011 define the Gaussian-constraints as discussed in Sec. 10. For the width of the Gaussians  
 1012 only the statistical error of the tagging calibration parameters are used since systematic  
 1013 uncertainties (except the systematic arising from the decay-time resolution which is already  
 1014 included by the procedure described above) are found to be negligible, see Table 12.3.  
 1015 Finally, we take the biggest change in fit central value as the systematic for each parameter  
 1016 of the  $B_s \rightarrow D_s K \pi \pi$  fits.

1017 A systematic uncertainty due to the limited knowledge of the decay-time bias which is  
 1018 fixed in the fit is evaluated by means of a toy study, where the value is randomized within  
 1019 its uncertainty. This source of systematic uncertainty is only relevant for the measurement  
 1020 of  $\Delta m_s$ , while the impact on all parameters extracted from the fit to  $B_s \rightarrow D_s K \pi \pi$  data  
 1021 is negligible.

1022 The procedure of accounting for the VELO misalignment with a constant decay-time  
 1023 bias in the resolution model (*i.e.* ignoring the event-per-event dependency on the decay-  
 1024 time error estimate) is validated with a  $B_s^0 \rightarrow D_s \pi$  MC sample processed with a realistic  
 1025 VELO misalignment. A fit to the decay-time distribution (using the TRUEID of the  
 1026 candidates for the tagging information) is performed to extract  $\Delta m_s$ , where a constant  
 1027 decay-time bias is included in the resolution model. The result differs by  $0.003 \text{ ps}^{-1}$  from  
 1028 the result obtained by fitting simulated  $B_s^0 \rightarrow D_s \pi$  candidates processed without VELO  
 1029 misalignment. This difference is added as additional systematic uncertainty on  $\Delta m_s$  due  
 1030 to the decay-time bias correction method.



**Figure 12.2:** 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.

## 1031 12.5 Production, detection asymmetries and mixing frequency

1032 The systematic from the production, detection asymmetries and  $\Delta m_s$  (in case of  $B_s \rightarrow$   
 1033  $D_s K\pi\pi$  decays) which are fixed in the fit are evaluated by means of a toy study similar  
 1034 to the procedure performed for the time-acceptance. The parameters are assumed to be  
 1035 uncorrelated.

## 1036 12.6 Multiple candidates

1037 The fraction of events with multiple candidates has been found to be very small, it is  
 1038 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  
 1039 candidates, while a systematic uncertainty is assigned by repeating the fit randomly  
 1040 keeping only one candidate when multiple ones are founds. No shifts in the fit central  
 1041 values are observed.

## 1042 12.7 Length and momentum scales

1043 The uncertainty on the LHCb length scale is estimated to be at most 0.020% [51, 52],  
 1044 which translates directly in an uncertainty on  $\Delta m_s$  of 0.020% with other parameters  
 1045 being unaffected. The momentum scale uncertainty is at most 0.022%. In the fit to  
 1046  $B_s \rightarrow D_s K\pi\pi$  data, the scale uncertainties are then implicitly included in the systematic  
 1047 due to  $\Delta m_s$  which we vary within its uncertainties, see above.

## 1048 12.8 Phase space acceptance

1049 For the phase space acceptance we rely on simulated data. The integration error due to  
 1050 the limited size of the MC sample used to normalize the signal PDF is below 0.2% and  
 1051 thus negligible small.

1052 To assess the uncertainty due to possible data-simulation differences, we determine  
 1053 alternative phase space efficiencies by varying the selection requirements on quantities  
 1054 that are expected not to be well described by the simulation. In particular, we consider  
 1055 the following variations:

- No BDT cut is applied
- A tighter BDT requirement is used ( $\text{BDTG} > 0.6$ )
- No reweighting is applied
- Instead of the PID responses obtained from the `PIDCorr` tool, we use the `PIDGen` tool to resample the PID variables [34]
- The raw MC PID variables are used
- Candidates with `BKGCAT`= 60 are removed
- A tight cut on the  $B_s$  transverse momentum is applied ( $p_T > 10 \text{ GeV}$ )
- We require `L0Hadron-TOS` instead of (`L0Hadron-TOS` || `L0Global-TIS`)
- We require `L0Global-TIS` instead of (`L0Hadron-TOS` || `L0Global-TIS`)

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

It is assumed that the phase space acceptance is independent of the decay-time. This assumption is tested by using only simulated candidates within specific decay-time intervals to calculate the MC normalization integrals described in Sec. 7.3. Four approximately equally populated decay-time bins are chosen and the sample variance of the fitted values is assigned as systematic.

## 12.9 Resonance description

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

- The Lass description for the  $K\pi$   $S$ -wave is replaced by a more general model (Glass [53, 54])
- The Gounaris-Sakurai description for the  $\rho(770)$  is replaced by a relativistic Breit-Wigner propagator (Equation 2.16)
- The  $\omega$  contribution to the decay channel  $K_1(1270) \rightarrow K \rho(770)/\omega$  is set to zero
- For the decay channel  $K^*(1410) \rightarrow K \rho(770)$ , we include  $\rho(770) - \omega$  mixing with a relative magnitude and phase determined from data
- Instead of taking the energy-dependent widths of the three-body resonances from Refs. [12, 23], we derive them from Equation 2.17 taking only the dominant  $K\pi\pi$  decay mode into account

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

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

1093 **12.10 Alternative amplitude models**

1094 We tested several modifications of the LASSO model to assign an additional model  
1095 uncertainty to the measured observables  $r, \delta$  and  $\gamma - 2\beta_s$  as well as to the measured  
1096 masses and widths of the  $K_1(1400)$  and  $K^*(1410)$  resonances. The amplitude coefficients  
1097 are by definition parameters of a given model which is why we do not evaluate a model  
1098 uncertainty for them. The fit results of the following alternative models (Alt. 1 - Alt. 12)  
1099 are summarized in Tables 12.1 and 12.2.

- 1100 • All amplitudes selected by Stage 1 of the model selection are included for both  $b \rightarrow c$  and  
1101  $b \rightarrow u$  transitions (Alt. 1)
- 1102 • The decay channel  $K_1(1270)[D] \rightarrow K^*(892)\pi$ , where the  $K^*(892)\pi$  system is in relative a  
1103 D-wave state, is added (Alt. 2)
- 1104 • The decay channel  $K_1(1400) \rightarrow K\rho(770)$  is added (Alt. 12)
- 1105 • The decay channels  $K(1460) \rightarrow K\rho(770)$  and  $K(1460) \rightarrow K\sigma$  are added (Alt. 3)
- 1106 • The decay channels  $K^*(1680) \rightarrow K\rho(770)$  and  $K^*(1680) \rightarrow K^*(892)\pi$  are added (Alt. 9)
- 1107 • The decay channel  $K_2(1770) \rightarrow K^*(892)\pi$  is added (Alt. 10)
- 1108 • 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)
- 1109 • Higher orbital angular momentum states are added for the amplitudes:  $B_s \rightarrow$   
1110  $(D_s \pi)_P K^*(892)$  and  $B_s \rightarrow (D_s K)_P \rho(770)$  (Alt. 4 and Alt. 5)
- 1111 • 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  
1112  $B_s \rightarrow (D_s K)f_0(1370)$  are added in different combinations and angular momentum  
1113 configurations (Alt. 6, Alt. 7 and Alt. 8 )
- 1114 • The decay channels  $K_2^*(1430) \rightarrow K\rho(770)$  and  $K_2^*(1430) \rightarrow K^*(892)\pi$  are added (not  
1115 used)
- 1116 • 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  
1117 used)
- 1118 • A flat non-resonant  $B_s \rightarrow D_s K \pi\pi$  component is added (not used)

1119 In total 15 different sets of amplitudes are fitted. In some cases, the fit fractions of  
1120 additionally added amplitudes turn out to be exactly zero. These model are effectively  
1121 not distinguishable from the baseline LASSO model and are not considered further. From  
1122 the remaining 12 models, we compute the sample variance for each observable and take it  
1123 as model uncertainty.

**Table 12.1:** 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	6.5	6.7	5.2	5.4	6.1	6.0
$B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)$			1.3				
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	12.3	10.3	9.8	9.4	13.9	11.2	12.9
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))$							
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	3.3	2.5	2.3	2.6	3.5	2.8	3.5
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	47.9	54.3	53.5	58.8	47.0	53.7	43.7
$B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	15.9	15.0	15.6	15.7	17.6	15.7	16.1
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	6.5	5.5	6.2	6.3	6.5	6.5	6.6
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$		0.8					
$B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)$							
$B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))$							
$b \rightarrow c$							
$B_s \rightarrow (D_s \pi)_S K^*(892)$							
$B_s \rightarrow (D_s \pi)_P K^*(892)$		6.7	9.2	7.9	10.3	13.9	6.8
$B_s[P] \rightarrow (D_s \pi)_P K^*(892)$						0.1	
$B_s[D] \rightarrow (D_s \pi)_P K^*(892)$						0.9	
$B_s \rightarrow (D_s K)_S \sigma$							
$B_s \rightarrow (D_s K)_P \sigma$							
$B_s \rightarrow (D_s K)_S f_0(980)$							
$B_s \rightarrow (D_s K)_S f_2(1270)$							0.0
$B_s \rightarrow (D_s K)_P f_2(1270)$							
$B_s \rightarrow (D_s K)_S f_0(1370)$							
$B_s \rightarrow (D_s K)_S \rho(770)$							
$B_s \rightarrow (D_s K)_P \rho(770)$			0.7			0.1	
$B_s[P] \rightarrow (D_s K)_P \rho(770)$							
$B_s[D] \rightarrow (D_s K)_P \rho(770)$							
Sum	98.9	104.7	103.2	108.3	108.8	103.1	95.2
$b \rightarrow u$							
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	15.0	20.9	21.0	18.2	11.5	19.1	13.8
$B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)$			4.0				
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	29.5	33.3	30.8	33.0	29.8	35.2	29.6
$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))$							
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	8.0	8.2	7.2	9.2	7.5	8.9	8.2
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	15.5	29.5	21.2	22.1	23.5	15.4	19.6
$B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$		2.2					
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$		0.8					
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	22.3	19.1	12.9	17.2	15.1	21.7	21.9
$B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))$				3.5			
$B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)$				2.3			
$B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))$							
$B_s \rightarrow D_s (K_2(1770) \rightarrow K^*(892) \pi)$							
$B_s \rightarrow (D_s \pi)_S K^*(892)$							
$B_s \rightarrow (D_s \pi)_P K^*(892)$		37.6	16.3	23.1	17.3	61.2	29.3
$B_s[P] \rightarrow (D_s \pi)_P K^*(892)$						0.6	
$B_s[D] \rightarrow (D_s \pi)_P K^*(892)$						3.8	
$B_s \rightarrow (D_s K)_S \sigma$							
$B_s \rightarrow (D_s K)_P \sigma$							
$B_s \rightarrow (D_s K)_S f_0(980)$							
$B_s \rightarrow (D_s K)_S f_2(1270)$							0.2
$B_s \rightarrow (D_s K)_P f_2(1270)$							
$B_s \rightarrow (D_s K)_S f_0(1370)$							
$B_s \rightarrow (D_s K)_S \rho(770)$							
$B_s \rightarrow (D_s K)_P \rho(770)$	8.4	8.2	4.8	13.3	8.4	12.6	7.5
$B_s[P] \rightarrow (D_s K)_P \rho(770)$						0.9	
$B_s[D] \rightarrow (D_s K)_P \rho(770)$						0.8	
Sum	136.4	138.7	124.9	136.0	161.5	143.8	142.6
$m_{K_1(1400)} [\text{MeV}]$	1397	1389	1389	1394	1410	1392	1403
$\Gamma_{K_1(1400)} [\text{MeV}]$	205	210	207	220	222	207	205
$m_{K^*(1410)} [\text{MeV}]$	1432	1435	1436	1429	1435	1428	1432
$\Gamma_{K^*(1410)} [\text{MeV}]$	345	351	348	347	360	346	347
$r$	0.50	0.46	0.48	0.49	0.51	0.49	0.51
$\kappa$	0.52	0.61	0.54	0.30	0.53	0.51	0.51
$\delta [{}^\circ]$	46	58	57	46	42	55	45
$\gamma - 2\beta_s [{}^\circ]$	61	70	57	60	54	67	62

**Table 12.2:** 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
$b \rightarrow c$ <ul style="list-style-type: none"> <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)</math> 6.9 7.6 6.1 6.7 8.1 6.3</li> <li><math>B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)</math></li> <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))</math> 13.8 12.8 13.2 11.0 14.7 14.9</li> <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))</math></li> <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)</math> 3.8 3.1 3.3 3.3 4.0 3.3</li> <li><math>B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)</math> 47.8 45.7 49.8 52.6 46.4 49.9</li> <li><math>B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))</math></li> <li><math>B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))</math> 0.6</li> <li><math>B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)</math> 15.4 15.5 18.8 15.2 15.5 15.9</li> <li><math>B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))</math> 5.8 6.0 5.2 6.3 6.2 6.3</li> <li><math>B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)</math></li> <li><math>B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))</math></li> <li><math>B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)</math></li> <li><math>B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)</math> 0.8</li> <li><math>B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))</math> 0.9</li> <li><math>B_s \rightarrow D_s (K_2(1770) \rightarrow K^*(892) \pi)</math> 0.7</li> <li><math>B_s \rightarrow (D_s \pi)_S K^*(892)</math></li> <li><math>B_s \rightarrow (D_s \pi)_P K^*(892)</math> 7.8 5.7 7.7 7.3 7.1 6.5</li> <li><math>B_s [P] \rightarrow (D_s \pi)_P K^*(892)</math></li> <li><math>B_s [D] \rightarrow (D_s \pi)_P K^*(892)</math></li> <li><math>B_s \rightarrow (D_s K)_S \sigma</math> 1.6 0.4</li> <li><math>B_s \rightarrow (D_s K)_P \sigma</math> 2.9</li> <li><math>B_s \rightarrow (D_s K)_S f_0(980)</math> 0.6</li> <li><math>B_s \rightarrow (D_s K)_S f_2(1270)</math></li> <li><math>B_s \rightarrow (D_s K)_P f_2(1270)</math></li> <li><math>B_s \rightarrow (D_s K)_S f_0(1370)</math></li> <li><math>B_s \rightarrow (D_s K)_S \rho(770)</math></li> <li><math>B_s \rightarrow (D_s K)_P \rho(770)</math></li> <li><math>B_s [P] \rightarrow (D_s K)_P \rho(770)</math></li> <li><math>B_s [D] \rightarrow (D_s K)_P \rho(770)</math></li> <li>Sum 105.7 97.4 105.8 103.1 102.0 103.8</li>   <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)</math> 9.7 14.3 14.4 19.3 8.0 13.3</li> <li><math>B_s \rightarrow D_s (K_1(1270)[D] \rightarrow K^*(892) \pi)</math></li> <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))</math> 19.3 24.1 31.3 31.7 14.5 31.5</li> <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(1450))</math></li> <li><math>B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)</math> 5.3 5.9 7.8 9.5 4.0 7.1</li> <li><math>B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)</math> 6.8 14.3 12.0 18.7 8.9 15.7</li> <li><math>B_s \rightarrow D_s (K_1(1400) \rightarrow K \rho(770))</math> 0.2</li> <li><math>B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)</math></li> <li><math>B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))</math></li> <li><math>B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)</math> 29.4 27.0 23.1 22.6 28.9 22.8</li> <li><math>B_s \rightarrow D_s (K(1460) \rightarrow K \rho(770))</math></li> <li><math>B_s \rightarrow D_s (K(1460) \rightarrow K \sigma)</math></li> <li><math>B_s \rightarrow D_s (K^*(1680) \rightarrow K^*(892) \pi)</math></li> <li><math>B_s \rightarrow D_s (K^*(1680) \rightarrow K \rho(770))</math></li> <li><math>B_s \rightarrow D_s (K_2(1770) \rightarrow K^*(892) \pi)</math> 0.8</li> <li><math>B_s \rightarrow (D_s \pi)_S K^*(892)</math></li> <li><math>B_s \rightarrow (D_s \pi)_P K^*(892)</math> 45.5 40.9 39.3 27.0 46.6 35.3</li> <li><math>B_s [P] \rightarrow (D_s \pi)_P K^*(892)</math></li> <li><math>B_s [D] \rightarrow (D_s \pi)_P K^*(892)</math></li> <li><math>B_s \rightarrow (D_s K)_S \sigma</math> 0.3 0.7</li> <li><math>B_s \rightarrow (D_s K)_P \sigma</math> 0.6</li> <li><math>B_s \rightarrow (D_s K)_S f_0(980)</math> 0.9</li> <li><math>B_s \rightarrow (D_s K)_S f_2(1270)</math></li> <li><math>B_s \rightarrow (D_s K)_P f_2(1270)</math></li> <li><math>B_s \rightarrow (D_s K)_S f_0(1370)</math></li> <li><math>B_s \rightarrow (D_s K)_S \rho(770)</math></li> <li><math>B_s \rightarrow (D_s K)_P \rho(770)</math> 4.6 8.4 8.6 7.9 4.1 7.3</li> <li><math>B_s [P] \rightarrow (D_s K)_P \rho(770)</math></li> <li><math>B_s [D] \rightarrow (D_s K)_P \rho(770)</math></li> <li>Sum 121.6 136.5 136.5 137.5 115.0 133.1</li>   <li><math>m_{K_1(1400)}</math> [MeV] 1401 1393 1399 1394 1400 1393</li> <li><math>\Gamma_{K_1(1400)}</math> [MeV] 195 199 200 208 194 205</li> <li><math>m_{K^*(1410)}</math> [MeV] 1444 1438 1413 1432 1435 1433</li> <li><math>\Gamma_{K^*(1410)}</math> [MeV] 329 342 400 351 337 346</li> <li><math>r</math> 0.44 0.46 0.49 0.46 0.46 0.49</li> <li><math>\kappa</math> 0.46 0.51 0.51 0.54 0.56 0.44</li> <li><math>\delta [^\circ]</math> 34 49 47 49 34 44</li> <li><math>\gamma - 2\beta_s [^\circ]</math> 54 69 57 68 51 61</li> </ul>							

**Table 12.3:** 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	Closure	Bkg.	Correlations	Acceptance	Resolution	Decay-time bias	VELO-Misalign.	Asymmetries	Mon./z-Scale	Total
$p_0^{OS}$ Run-I	0.05	0.09	0.03	0.01	1.00	0.02		0.00			1.00
$p_1^{OS}$ Run-I	0.01	0.13	0.06	0.01	1.04	0.02		0.01			1.05
$\Delta p_0^{OS}$ Run-I	0.14	0.03	0.03	0.00	0.02	0.01		0.15			0.21
$\Delta p_1^{OS}$ Run-I	0.07	0.06	0.16	0.00	0.03	0.01		0.15			0.24
$\epsilon_{tag}^{OS}$ Run-I	0.06	0.17	0.12	0.00	0.00	0.00		0.01			0.22
$\Delta \epsilon_{tag}^{OS}$ Run-I	0.04	0.01	0.05	0.00	0.06	0.02		0.01			0.09
$p_0^{SS}$ Run-I	0.03	0.03	0.15	0.01	0.56	0.02		0.00			0.58
$p_1^{SS}$ Run-I	0.10	0.03	0.04	0.01	0.60	0.01		0.01			0.61
$\Delta p_0^{SS}$ Run-I	0.04	0.01	0.07	0.00	0.00	0.00		0.10			0.13
$\Delta p_1^{SS}$ Run-I	0.03	0.04	0.04	0.00	0.01	0.00		0.12			0.14
$\epsilon_{tag}^{SS}$ Run-I	0.02	0.02	0.10	0.00	0.00	0.00		0.01			0.11
$\Delta \epsilon_{tag}^{SS}$ Run-I	0.04	0.03	0.02	0.00	0.04	0.03		0.01			0.07
$p_0^{OS}$ Run-II	0.02	0.20	0.01	0.02	0.97	0.04		0.00			0.99
$p_1^{OS}$ Run-II	0.02	0.08	0.10	0.01	0.60	0.04		0.00			0.61
$\Delta p_0^{OS}$ Run-II	0.07	0.08	0.13	0.00	0.22	0.01		0.00			0.27
$\Delta p_1^{OS}$ Run-II	0.02	0.03	0.16	0.00	0.08	0.01		0.00			0.18
$\epsilon_{tag}^{OS}$ Run-II	0.01	0.16	0.04	0.00	0.11	0.00		0.00			0.20
$\Delta \epsilon_{tag}^{OS}$ Run-II	0.05	0.05	0.10	0.00	0.03	0.03		0.00			0.13
$p_0^{SS}$ Run-II	0.10	0.06	0.05	0.01	0.61	0.02		0.00			0.62
$p_1^{SS}$ Run-II	0.01	0.07	0.06	0.02	0.63	0.04		0.00			0.64
$\Delta p_0^{SS}$ Run-II	0.07	0.02	0.01	0.00	0.04	0.01		0.00			0.09
$\Delta p_1^{SS}$ Run-II	0.11	0.05	0.14	0.00	0.04	0.01		0.00			0.19
$\epsilon_{tag}^{SS}$ Run-II	0.03	0.03	0.09	0.00	1.68	0.00		0.00			1.68
$\Delta \epsilon_{tag}^{SS}$ Run-II	0.01	0.03	0.01	0.00	0.21	0.03		0.00			0.21
$A_P$ Run-II	0.04	0.02	0.17	0.00	0.19	0.03		0.01			0.26
$\Delta m_s$	0.01	0.32	0.11	0.02	0.00	0.13	0.69	0.36	0.02	0.67	1.09

**Table 12.4:** 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	Background	Correlations	Acceptance	Resolution	Decay-time bias	Asymmetries	$\Delta m_s$	Total
$C$	0.15	0.03	0.06	0.11	0.04	0.15	0.04	0.06	0.26
$D$	0.00	0.18	0.10	0.31	0.05	0.01	0.00	0.01	0.38
$\bar{D}$	0.05	0.20	0.13	0.28	0.04	0.02	0.00	0.01	0.37
$S$	0.05	0.01	0.06	0.06	0.04	0.07	0.07	0.10	0.18
$\bar{S}$	0.07	0.02	0.03	0.05	0.04	0.05	0.05	0.10	0.16

**Table 12.5:** 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	Background	Correlations	Time-Acc.	Resolution	Decay-time bias	Asymmetries	$\Delta m_s$	Phsp-Acc.	Acc. Factor.	Lineshapes	Resonances $m, \Gamma$	Form-Factors	Ampl. Model	Total
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892)\pi)$ Mag	0.10	0.12	0.01	0.03	0.01	0.00	0.01	0.07	0.44	0.42	0.21	0.46	0.46	0.82	
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892)\pi)$ Phase	0.07	0.08	0.09	0.01	0.04	0.00	0.01	0.25	0.29	0.33	0.30	0.36	0.36	0.70	
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430)\pi)$ Mag	0.04	0.25	0.26	0.01	0.01	0.00	0.00	0.06	0.36	0.11	0.11	0.18	0.18	1.25	
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430)\pi)$ Phase	0.04	0.15	0.17	0.00	0.02	0.01	0.00	0.01	0.46	3.53	0.15	0.50	0.50	3.60	
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow c$ )	0.13	0.51	0.05	0.04	0.25	0.04	0.03	0.09	0.48	1.00	1.67	0.55	1.18	2.46	
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow c$ )	0.14	0.14	0.13	0.01	0.08	0.02	0.01	0.02	0.21	0.25	0.36	0.21	0.32	0.67	
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow u$ )	0.10	0.28	0.00	0.04	0.04	0.02	0.03	0.10	0.22	0.07	0.29	0.18	0.62	0.82	
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow u$ )	0.02	0.10	0.28	0.05	0.04	0.03	0.03	0.11	0.29	0.47	0.68	0.24	0.35	1.02	
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow c$ )	0.08	0.11	0.18	0.02	0.05	0.00	0.02	0.05	0.63	0.67	0.34	0.16	0.47	1.12	
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow c$ )	0.35	0.15	0.00	0.02	0.07	0.02	0.01	0.12	0.30	0.57	0.22	0.80	0.80	1.12	
$B_s \rightarrow D_s (K^*(1410) \rightarrow K_\rho(770))$ Mag	0.35	0.13	0.00	0.01	0.01	0.00	0.00	0.01	0.07	0.41	0.44	0.09	0.26	0.76	
$B_s \rightarrow D_s (K^*(1410) \rightarrow K_\rho(770))$ Phase	0.18	0.24	0.19	0.00	0.01	0.01	0.00	0.01	0.15	0.25	0.32	0.09	0.21	0.61	
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892)\pi)$ Mag( $b \rightarrow u$ )	0.14	0.34	0.20	0.03	0.05	0.01	0.02	0.05	0.11	0.44	0.39	0.24	0.55	0.94	
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892)\pi)$ Phase( $b \rightarrow u$ )	0.13	0.16	0.30	0.03	0.08	0.02	0.02	0.05	0.42	0.60	0.38	0.36	0.36	0.98	
$B_s \rightarrow (D_s \pi)_P K^*(892)$ Mag( $b \rightarrow c$ )	0.03	0.29	0.26	0.02	0.07	0.02	0.01	0.03	0.23	0.44	0.98	0.14	0.67	1.35	
$B_s \rightarrow (D_s \pi)_P K^*(892)$ Phase( $b \rightarrow c$ )	0.20	0.34	0.25	0.01	0.09	0.02	0.01	0.01	0.34	0.47	0.64	0.12	0.35	1.05	
$B_s \rightarrow (D_s \pi)_P K^*(892)$ Mag( $b \rightarrow u$ )	0.14	0.14	0.28	0.04	0.09	0.02	0.02	0.08	0.37	0.27	0.88	0.29	0.66	1.28	
$B_s \rightarrow (D_s \pi)_P K^*(892)$ Phase( $b \rightarrow u$ )	0.24	0.27	0.10	0.03	0.11	0.02	0.02	0.03	0.28	0.51	0.87	0.16	0.34	1.18	
$B_s \rightarrow (D_s K)_P \rho(770)$ Mag( $b \rightarrow u$ )	0.35	0.22	0.00	0.03	0.02	0.02	0.02	0.05	0.21	0.16	0.50	0.26	0.51	0.91	
$B_s \rightarrow (D_s K)_P \rho(770)$ Phase( $b \rightarrow u$ )	0.12	0.58	0.00	0.02	0.04	0.01	0.02	0.05	0.11	0.40	0.39	0.33	0.57	1.06	
$m_{K_1(1400)}$	0.09	0.12	0.07	0.01	0.07	0.01	0.00	0.01	0.20	0.09	0.17	0.11	0.32	0.64	
$\Gamma_{K_1(1400)}$	0.01	0.12	0.22	0.01	0.01	0.01	0.01	0.16	0.27	0.42	0.12	0.40	0.57	0.91	
$m_{K^*(1410)}$	0.05	0.10	0.07	0.01	0.02	0.00	0.00	0.01	0.28	0.74	0.31	0.05	1.56	1.94	
$\Gamma_{K^*(1410)}$	0.25	0.13	0.18	0.00	0.02	0.00	0.00	0.01	0.31	0.15	0.04	1.38	0.69	1.62	
$r$	0.11	0.28	0.22	0.03	0.05	0.03	0.02	0.07	0.30	0.03	0.35	0.07	0.15	0.46	0.78
$\delta$	0.19	0.10	0.17	0.04	0.07	0.03	0.05	0.11	0.10	0.09	0.30	0.03	0.11	0.55	0.72
$\gamma - 2\beta_s$	0.10	0.13	0.11	0.07	0.13	0.01	0.02	0.07	0.04	0.02	0.28	0.03	0.11	0.42	0.59

## 1124 13 Summary

1125 The  $B_s^0$  oscillation frequency  $\Delta m_s$  is measured from the time-dependent fit to  $B_s^0 \rightarrow D_s\pi\pi\pi$   
 1126 data to be

$$\Delta m_s = 17.7537 \pm 0.0084 \pm 0.0092,$$

1127 where the errors are statistical and systematic, respectively. Table 13.1 summarizes the  
 1128 values for the ratio of the  $b \rightarrow u$  and  $b \rightarrow c$  transition amplitudes ( $r$ ) and their strong  
 1129 ( $\delta$ ) and weak phase ( $\gamma - 2\beta_s$ ) difference obtained from the phase-space integrated, as  
 1130 well as the time-dependent amplitude fit to  $B_s^0 \rightarrow D_s K\pi\pi$  data. Their comparison is  
 1131 not so straightforward as the results are obtained from the same data set. We define a  
 1132 measure of their compatibility as follows: For the statistical error we take the difference  
 1133 of the statistical uncertainties of both fits:  $\Delta\sigma_{stat} = \sigma_{stat}^{MI} - \sigma_{stat}^{Full}$ , where  $\sigma_{stat}^{MI}$  is the  
 1134 statistical uncertainty of the phase-space integrated (model-independent) fit and  $\sigma_{stat}^{Full}$   
 1135 is the statistical uncertainty of the full time-dependent amplitude fit. In the limit of  
 1136 equal statistical uncertainties,  $\Delta\sigma_{stat} = 0$ , and the results should agree perfectly. For  
 1137 the systematic error we assume that the model-dependent error (including resonance  
 1138 line shapes, form factors and alternative amplitude models) of the full time-dependent  
 1139 amplitude fit,  $\sigma_{model}^{Full}$ , is uncorrelated to the phase-space integrated results. The remaining  
 1140 systematics (including time-acceptance, resolution, tagging, etc.) are assumed to be 100%  
 1141 correlated such that their effect should cancel. Hence, the results should agree within  
 1142 a spread given by:  $\Delta\sigma = \sqrt{\Delta\sigma_{stat}^2 + (\sigma_{model}^{Full})^2}$ . The physical parameters agree within a  
 1143 range of 0.4 to 1.4  $\Delta\sigma$ .

**Table 13.1:** Parameters determined from the two fits performed to the  $B_s^0 \rightarrow D_s K\pi\pi$  data sample. Statistical and systematic uncertainties are combined.

Parameter	Phase-space integrated fit	Time-dependent amplitude fit	Difference [ $\Delta\sigma$ ]
$r$	$0.41 \pm 0.10$	$0.50 \pm 0.05$	1.4
$\kappa$	$0.65 \pm 0.21$	$0.52 \pm 0.11$	0.8
$\delta$ [°]	$40 \pm 18$	$46 \pm 18$	0.6
$\gamma - 2\beta_s$ [°]	$65 \pm 21$	$61 \pm 17$	0.4

## 1144 A Parametrization of Amplitude Lineshapes

### 1145 Bugg model for $\sigma$ resonance

1146 For the broad scalar resonance  $\sigma$ , the model from Bugg is used [55]:

$$T(s) = [M^2 - s - g_1^2 \frac{s - s_A}{M^2 - s_A} z(s) - iM\Gamma_{tot}(s)]^{-1}, \quad (1.1)$$

$$M\Gamma_1(s) = g_1^2 \frac{s - s_A}{M^2 - s_A} \rho_1(s), \quad (1.2)$$

$$g_1^2(s) = M(b_1 + b_2 s) \exp[-(s - M^2)A^{-1}], \quad (1.3)$$

$$j_1(s) = \frac{1}{\pi} [2 + \rho_1 \ln_e \frac{1 - \rho_1}{1 + \rho_1}], \quad (1.4)$$

$$z(s) = j_1(s) - j_1(M^2), \quad (1.5)$$

$$M\Gamma_2(s) = 0.6g_1^2(s) \frac{s}{M^2} \exp(-\alpha|s - 4m_K^2|) \rho_2(s), \quad (1.6)$$

$$M\Gamma_3(s) = 0.2g_1^2(s) \frac{s}{M^2} \exp(-\alpha|s - 4m_\eta^2|) \rho_3(s), \quad (1.7)$$

$$M\Gamma_4(s) = Mg_4 \frac{\rho_{4\pi}(s)}{\rho_{4\pi}(M^2)}, \quad (1.8)$$

$$\rho_2(s) = \sqrt{1 - 4m_\pi^2/s}, \quad (1.9)$$

$$\rho_{4\pi}(s) = 1.0[1 + \exp(7.082 - 2.845s)]^{-1}, \quad (1.10)$$

1147 where the numerical values for the parameters are [55]  $M = 0.953 \text{ GeV}$ ,  $b_1 = 1.302 \text{ GeV}$ ,  
 1148  $b_2 = 0.340 \text{ GeV}^{-1}$ ,  $A = 2.426 \text{ GeV}^2$ ,  $g_{4\pi} = 0.011 \text{ GeV}$  and  $s_A = 0.41m_\pi^2$ .

### 1149 Model for $K\pi$ -S-wave

1150 The LASS parameterization is used to model the  $K\pi$  S-wave contribution. It consists of  
 1151 the  $K_0^*(1430)$  resonance together with an effective range non-resonant component [56–58]:

$$T_{LASS}(s) = \frac{\sqrt{s}}{q \cot \delta_L - iq} + e^{2i\delta_L} \frac{m_0 \Gamma_0 \frac{m_0}{q_0}}{m_0^2 - s - i m_0 \Gamma_0 \frac{m_0}{\sqrt{s}} \frac{q}{q_0}} \quad (1.11)$$

1152 with  $\cot \delta_L = \frac{1}{aq} + \frac{1}{2}rq$ . We use the values for the scattering length  $a = 2.07 \pm 0.1 \text{ GeV}$   
 1153 and effective range parameter  $r = 3.32 \pm 0.34 \text{ GeV}$  from Ref. [56, 57].

1154 For systematic studies, the GLASS shape is used:

$$T_{GLASS}(s) = F \sin(\phi_B) e^{i\phi_B} + R \sin(\phi_R) e^{i\phi_R} e^{2i\phi_B} \quad (1.12)$$

$$\phi_B = \psi_B + \tan^{-1} \left( \frac{2aq(s)}{2 + arq^2(s)} \right) \quad (1.13)$$

$$\phi_R = \psi_R + \tan^{-1} \left( \frac{m\Gamma(s)}{m_0^2 - s} \right), \quad (1.14)$$

1155 with  $F = 0.62 \pm 0.04$ ,  $\phi_F = -0.100 \pm 0.010$ ,  $R = 1$ ,  $\phi_R = 1.10 \pm 0.02$ ,  $a = 0.224 \pm$   
 1156  $0.003 \text{ GeV}^{-1}$  and  $r = -15.01 \pm 0.13 \text{ GeV}^{-1}$ .

1157 **Model for  $\rho^0(770)$  resonance**

1158 We use the Gounaris-Sakurai parametrization for the  $\rho(770)^0 \rightarrow \pi\pi$  propagator [59]:

$$T_{GS}(m) = \frac{1 + f(0)/m_0^2}{m_0^2 + f(m) - m^2 - i m_0 \Gamma(m)}, \quad (1.15)$$

1159 where  $\Gamma(m)$  takes on the same form as in Eq. 2.16 and the function  $f(m)$  is defined as

$$f(m) = \Gamma_0 \frac{m_0^2}{q_0^3} \left[ q^2 (h(m) - h(m_0)) + (m^2 - m_0^2) q_0^2 \frac{dh}{dm} \Big|_{m_0} \right] \quad (1.16)$$

$$h(m) = \frac{2}{\pi} \frac{q}{m} \ln \left( \frac{m + 2q}{2m_\pi} \right). \quad (1.17)$$

1160 For the decay chain  $K_1(1270) \rightarrow \rho(770)K$ , we include  $\rho - \omega$  mixing [60]:

$$T(s) = T_{GS}(s) \cdot \left( 1 + \delta \frac{s}{m_\omega^2} T_\omega(s) \right) \quad (1.18)$$

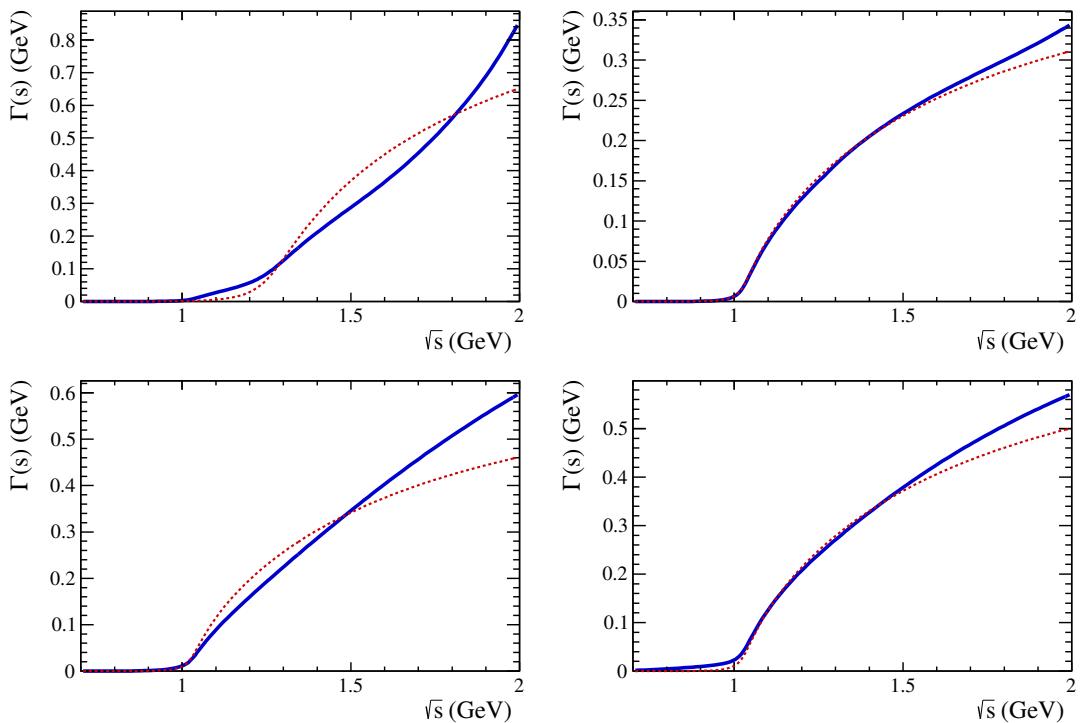
1161 where  $T_\omega$  is the relativistic Breit-Wigner propagator (Eq. 2.16) of the  $\omega$  and the relative  
1162 magnitude and phase between  $\rho$  and  $\omega$  fixed to the values determined in Ref. [61, 62]:  
1163  $|\delta| = 0.159 \pm 0.012 \pm 0.011$  and  $\arg(\delta) = 1.36 \pm 0.07 \pm 0.06$ .

## 1164 Running width distributions for 3-body resonances

1165 For the resonances  $K_1(1270)$  and  $K(1460)$ , the energy-dependent widths are reproduced  
1166 from Ref. [23]. We further use the energy-dependent widths for the  $K_1(1400)$ ,  $K^*(1410)$   
1167 and  $K^*(1680)$  mesons from Ref. [12]. For all other resonances decaying into a three-  
1168 body final state, an energy-dependent width distribution is derived from Equation 2.17  
1169 assuming an uniform phase space population. The running width distributions of the  
1170 3-body resonances included in the baseline model are shown in Fig. 1.1.

## 1171 Additional models

1172 Lineshape models for resonances which are not part of the nominal model, can be found  
1173 in the References. The energy-dependent width of the  $f_0(980)$  resonance is given by the  
1174 sum of the partial widths into the  $\pi\pi$  and  $KK$  channels (*i.e.* the Flatté lineshape [63]),  
1175 where the coupling constants as well as the mass and width are taken from a measurement  
1176 performed by the BES Collaboration [64]. For the  $f_2(1270)$  and the  $f_0(1370)$  mesons we  
1177 use the total decay widths calculated in Ref. [12] which take the channels  $\pi\pi$ ,  $KK$ ,  $\eta\eta$  and  
1178  $\pi\pi\pi\pi$  into account.



**Figure 1.1:** Running width distributions of the 3-body resonances included in the baseline model:  $K_1(1270)$  (top-left),  $K_1(1400)$  (top-right),  $K(1460)$  (bottom-left) and  $K^*(1410)$  (bottom-right). The nominal models are shown in blue, alternative models used for systematic studies in red.

## <sup>1179</sup> B Stripping and Trigger cuts

<sup>1180</sup> The following text describes variables which are used in Table A.1 and might be ambiguous,  
<sup>1181</sup> or which benefits are not straight forward. Where noted, different cut values are applied  
<sup>1182</sup> for Run-I and Run-II data. In Table A.1, DOCA is the abbreviation for distance of closest  
<sup>1183</sup> approach. This variable is used to ensure that all  $D_s$  and  $X_{s,d}$  daughters originate from  
<sup>1184</sup> 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 A.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 $\chi^2/\text{nDoF}$	$< 10$	$< 8$
$B_s^0 \tau_{B_s^0}$	$> 0.2 \text{ ps}$	

<sup>1185</sup>

1186 Table A.2 summarizes the trigger requirements imposed by the Hlt1 line used in this  
 1187 analysis for Run-I. At least one of the six decay particles must pass the listed requirements  
 1188 in order for the event to be stored for further analysis. For Run-II, this trigger line was  
 1189 updated and uses a multivariate classifier which takes the variables listed in Table A.2 as  
 1190 input, rather than directly cutting on them.

1191 The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the  
 1192 b-hadron  $p_T$ , its flight distance  $\chi^2$  from the nearest PV and the sum of the  $B_s^0$  and  $D_s$   
 1193 vertex  $\chi^2$  divided by the sum of their number of degrees of freedom. Table A.3 summarizes  
 1194 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 A.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 A.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

1195

## C Details of multivariate classifier

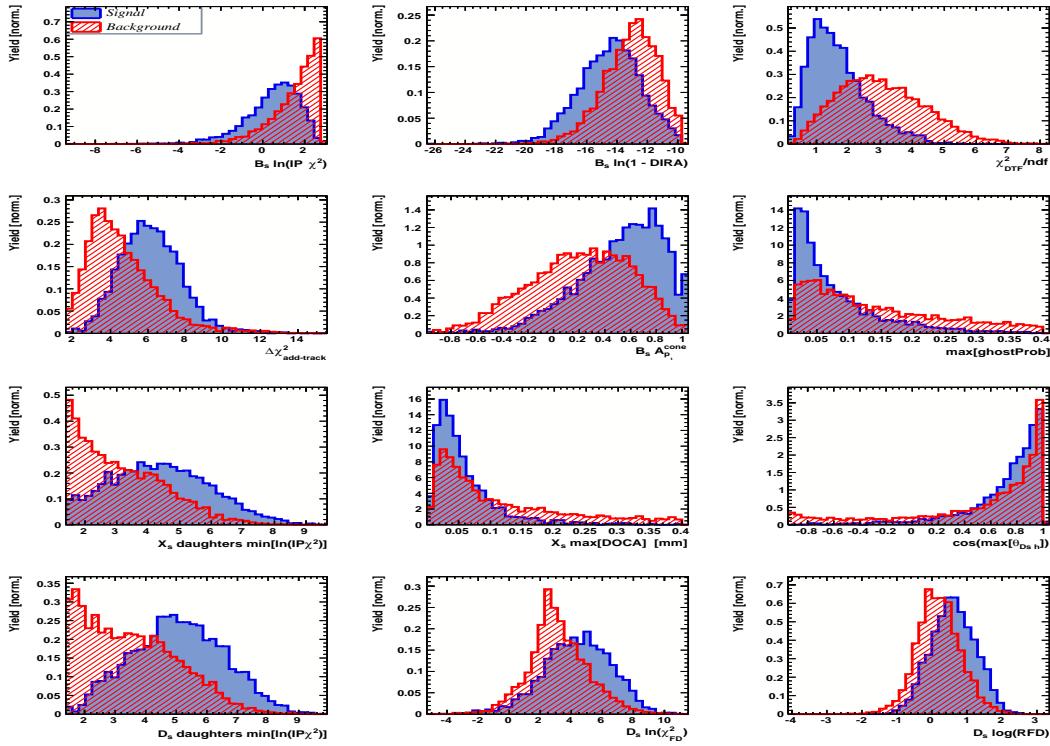


Figure B.1: Variables used to train the BDTG for category [Run-I,L0-TOS].

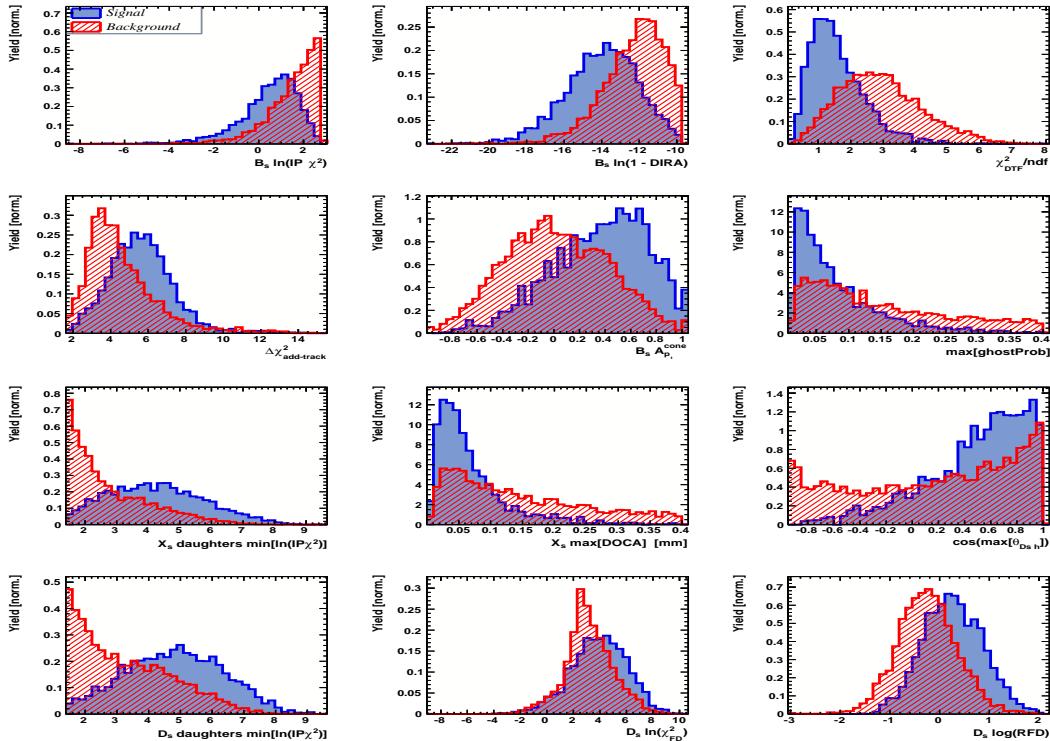
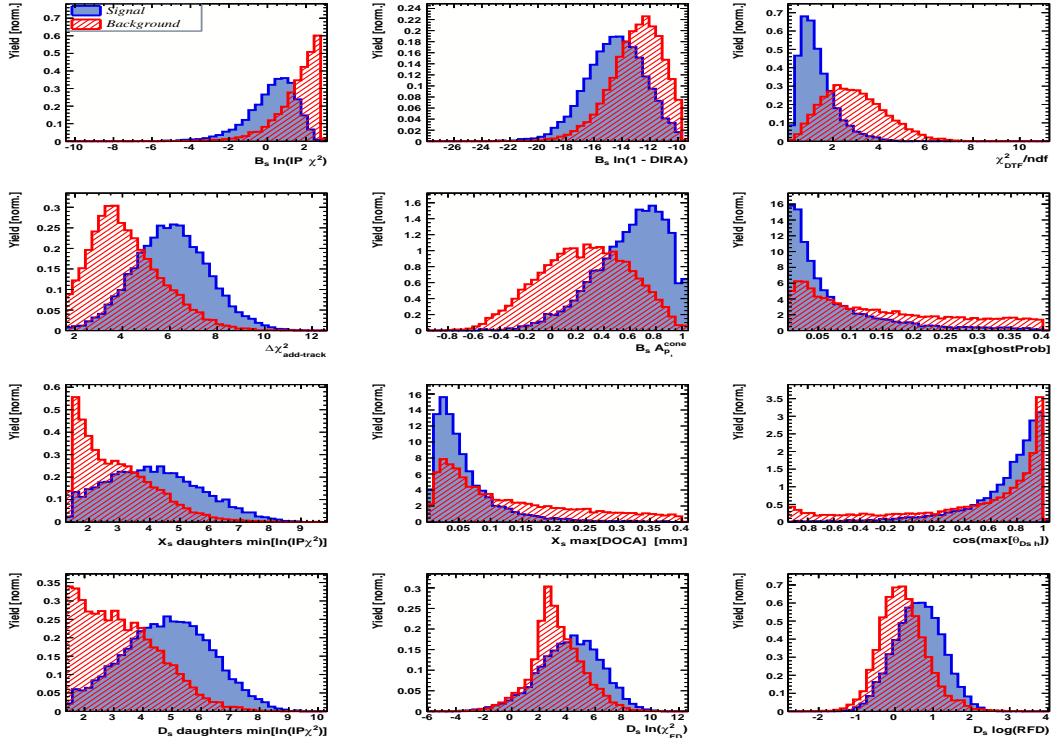
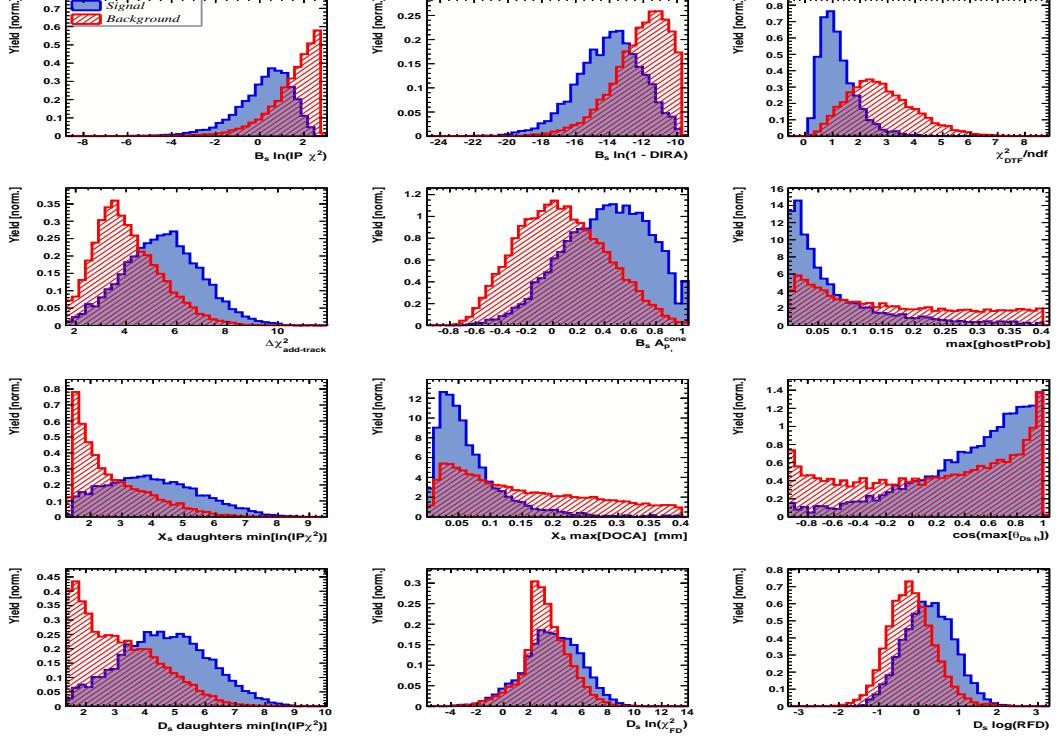


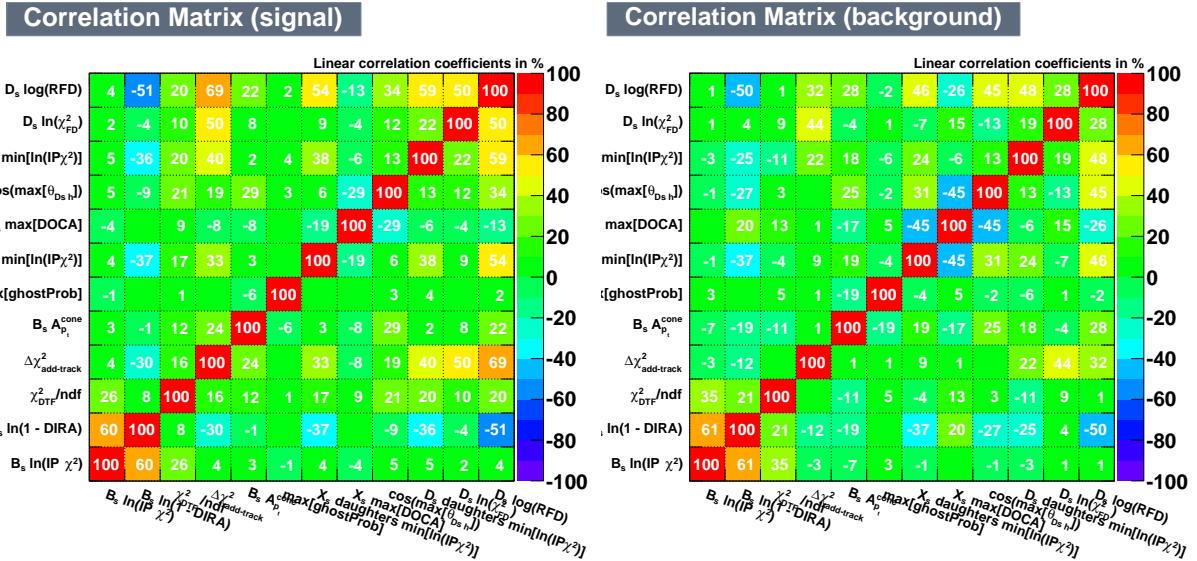
Figure B.2: Variables used to train the BDTG for category [Run-I,L0-TIS].



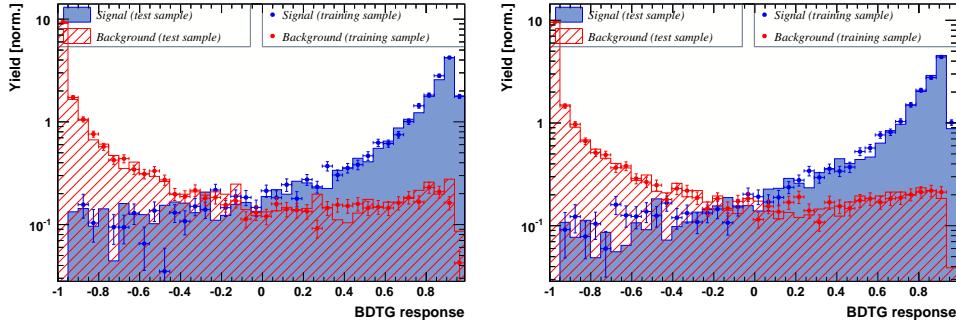
**Figure B.3:** Variables used to train the BDTG for category [Run-II,L0-TOS].



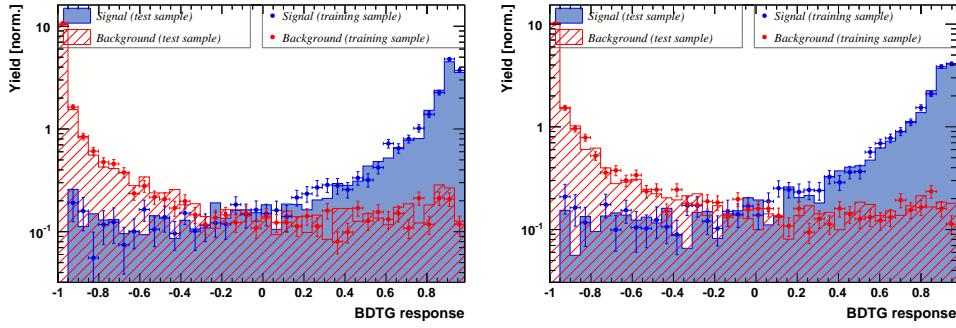
**Figure B.4:** Variables used to train the BDTG for category [Run-II,L0-TIS].



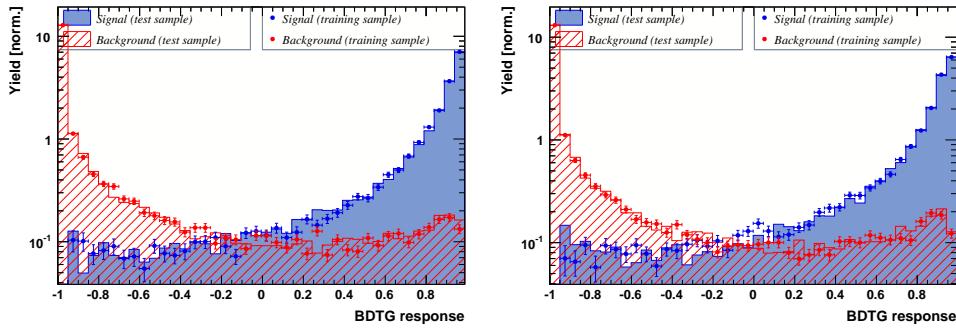
**Figure B.5:** Correlation matrix for the (left) signal and (right) background input distributions.



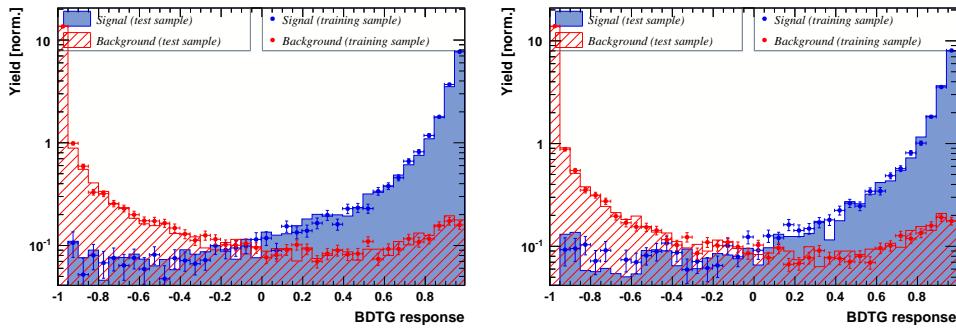
**Figure B.6:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TOS].



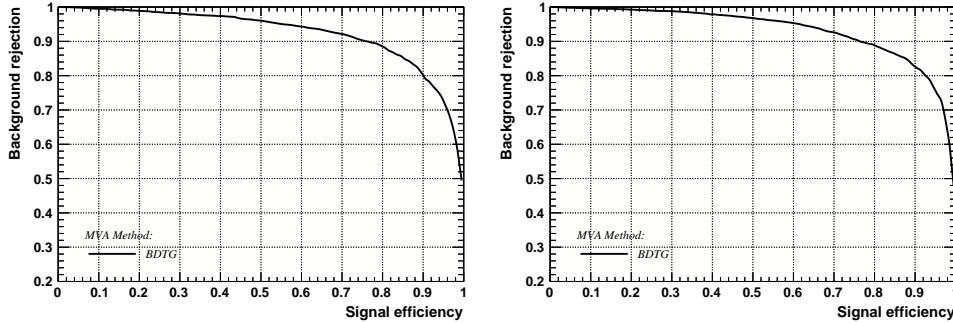
**Figure B.7:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-I,L0-TIS].



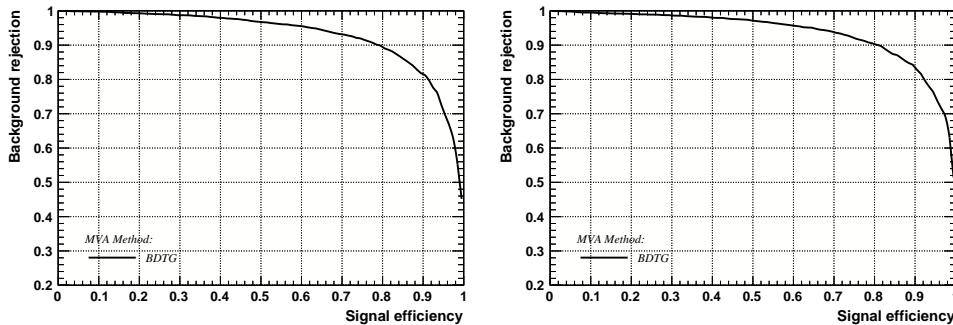
**Figure B.8:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TOS].



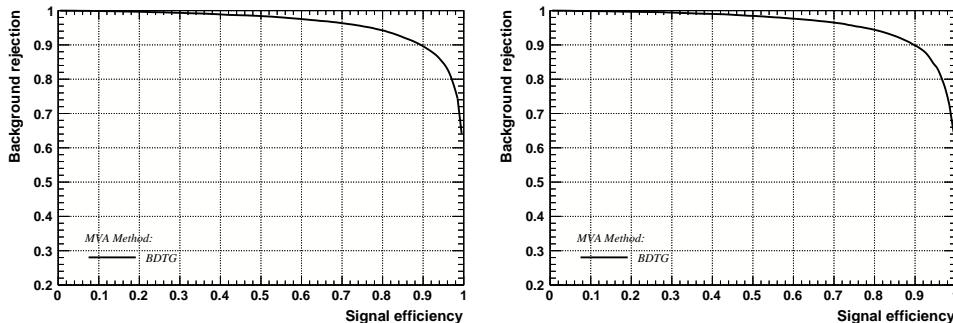
**Figure B.9:** Response of the classifier trained on the even and tested on the odd sample (left) and trained on the odd and tested on the even sample (right) for category [Run-II,L0-TIS].



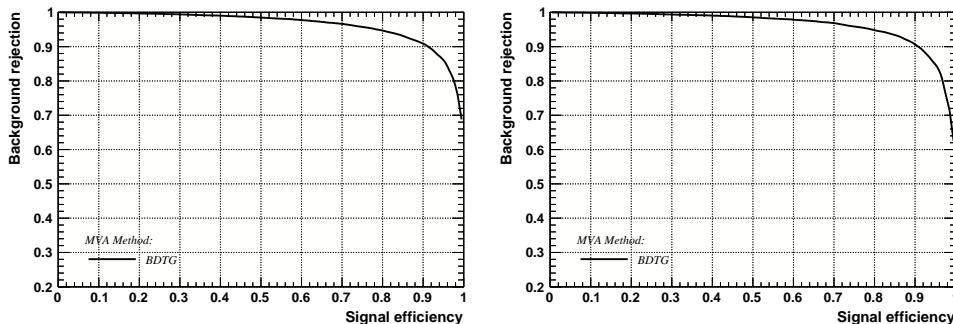
**Figure B.10:** Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-I,L0-TOS].



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



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



**Figure B.13:** Signal efficiency versus background rejection for the BDT trained on the even (left) and odd (right) sample for category [Run-II,L0-TIS].

## 1197 D Detailed mass fits

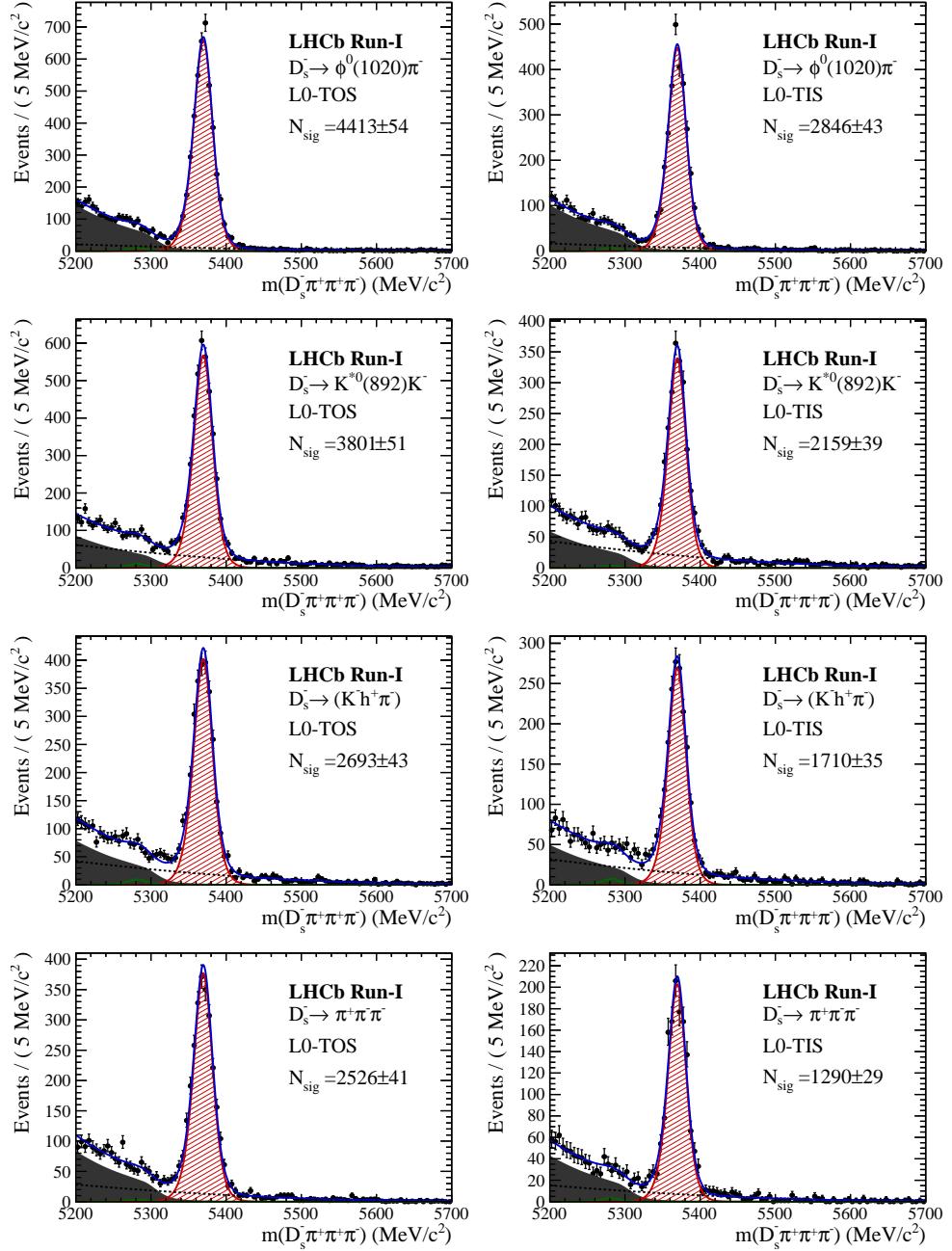
1198 In this section, all fits to the mass distribution of  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$   
 1199 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-  
 1200 II),  $D_s$  final state ( $D_s \rightarrow KK\pi$  non-resonant,  $D_s \rightarrow \phi\pi$ ,  $D_s \rightarrow K^*K$ , or  $D_s \rightarrow \pi\pi\pi$ ) and  
 1201 L0 trigger category.

**Table C.1:** Signal and background yields for the  $B_s \rightarrow D_s\pi\pi\pi$  sample split by data-taking period.

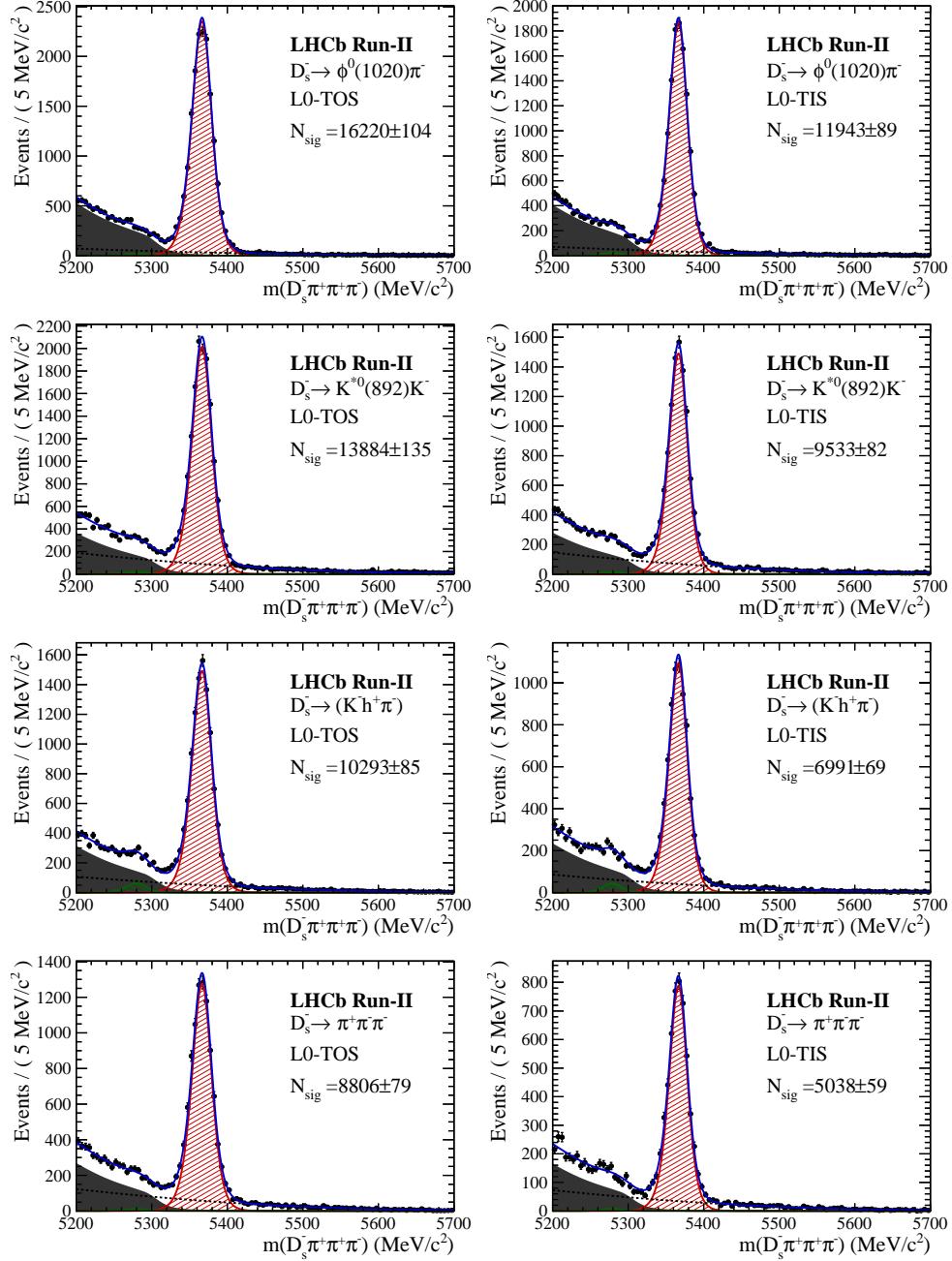
Component	Yield for Run I
$B_s \rightarrow D_s\pi\pi\pi$	$21443 \pm 122$
$B^0 \rightarrow D_s\pi\pi\pi$	$358 \pm 53$
Partially reconstructed bkg.	$8657 \pm 104$
Combinatorial bkg.	$9492 \pm 120$
Component	Yield for Run II
$B_s \rightarrow D_s\pi\pi\pi$	$82711 \pm 257$
$B^0 \rightarrow D_s\pi\pi\pi$	$1359 \pm 1326$
Partially reconstructed bkg.	$34471 \pm 937$
Combinatorial bkg.	$31574 \pm 680$

**Table C.2:** Signal and background yields for the  $B_s \rightarrow D_sK\pi\pi$  sample split by data-taking period.

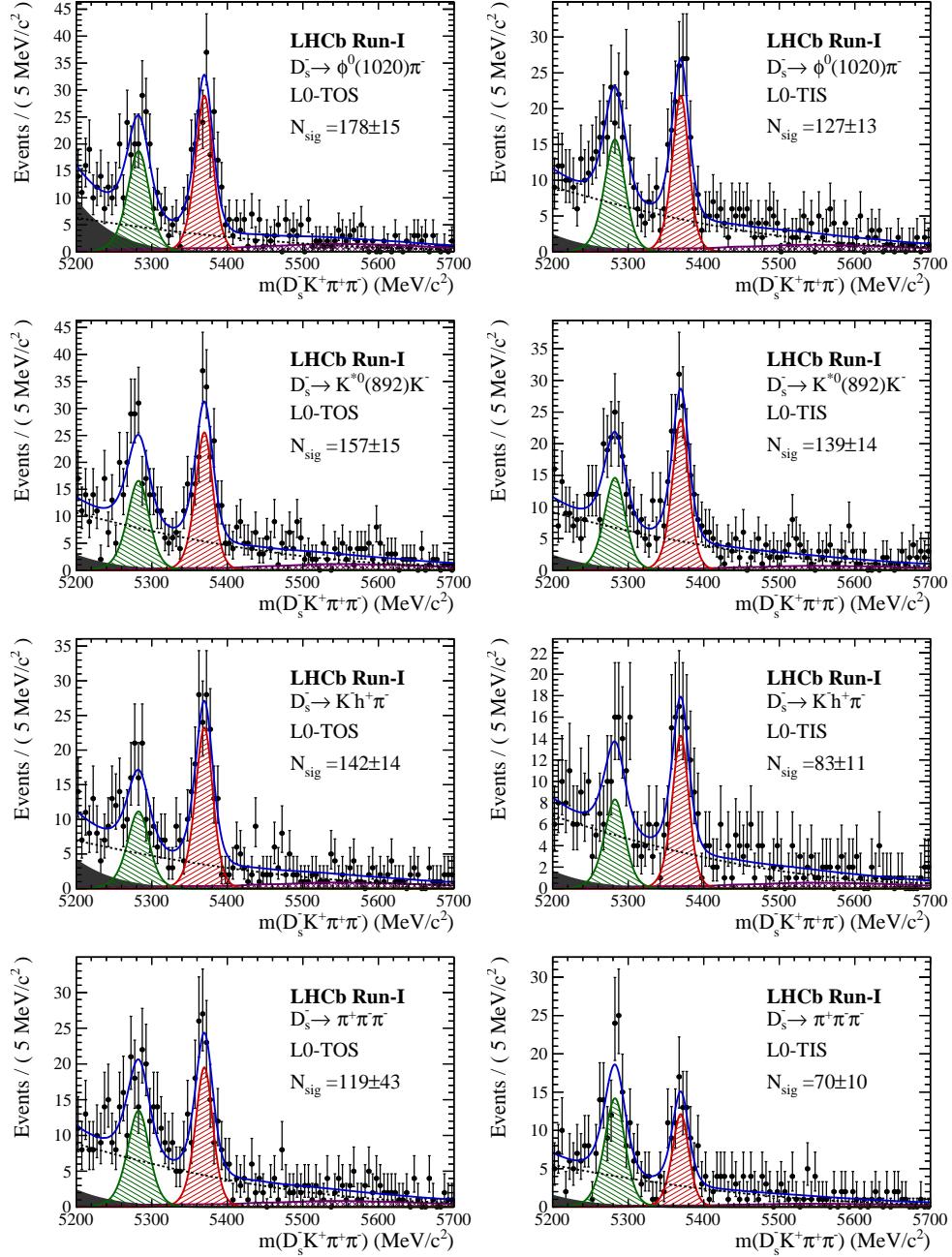
Component	Yield for Run I
$B_s \rightarrow D_sK\pi\pi$	$1018 \pm 56$
$B^0 \rightarrow D_sK\pi\pi$	$846 \pm 43$
Partially reconstructed bkg.	$232 \pm 137$
Misidentified bkg.	$426 \pm 0$
Combinatorial bkg.	$2520 \pm 324$
Component	Yield for Run II
$B_s \rightarrow D_sK\pi\pi$	$4153 \pm 76$
$B^0 \rightarrow D_sK\pi\pi$	$3264 \pm 86$
Partially reconstructed bkg.	$1592 \pm 312$
Misidentified bkg.	$760 \pm 0$
Combinatorial bkg.	$6653 \pm 186$



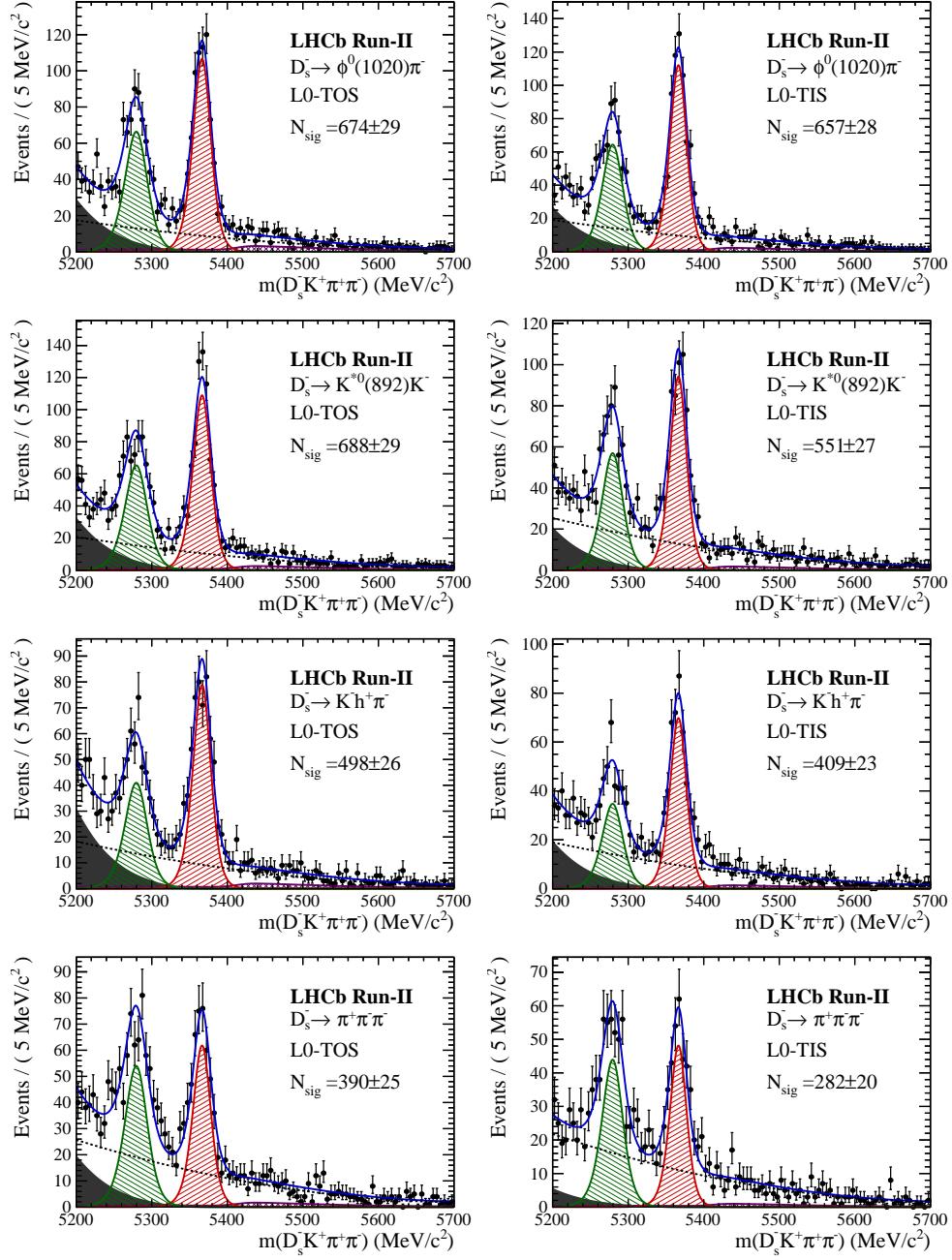
**Figure C.1:** Invariant mass distributions of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates for Run-I data.



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



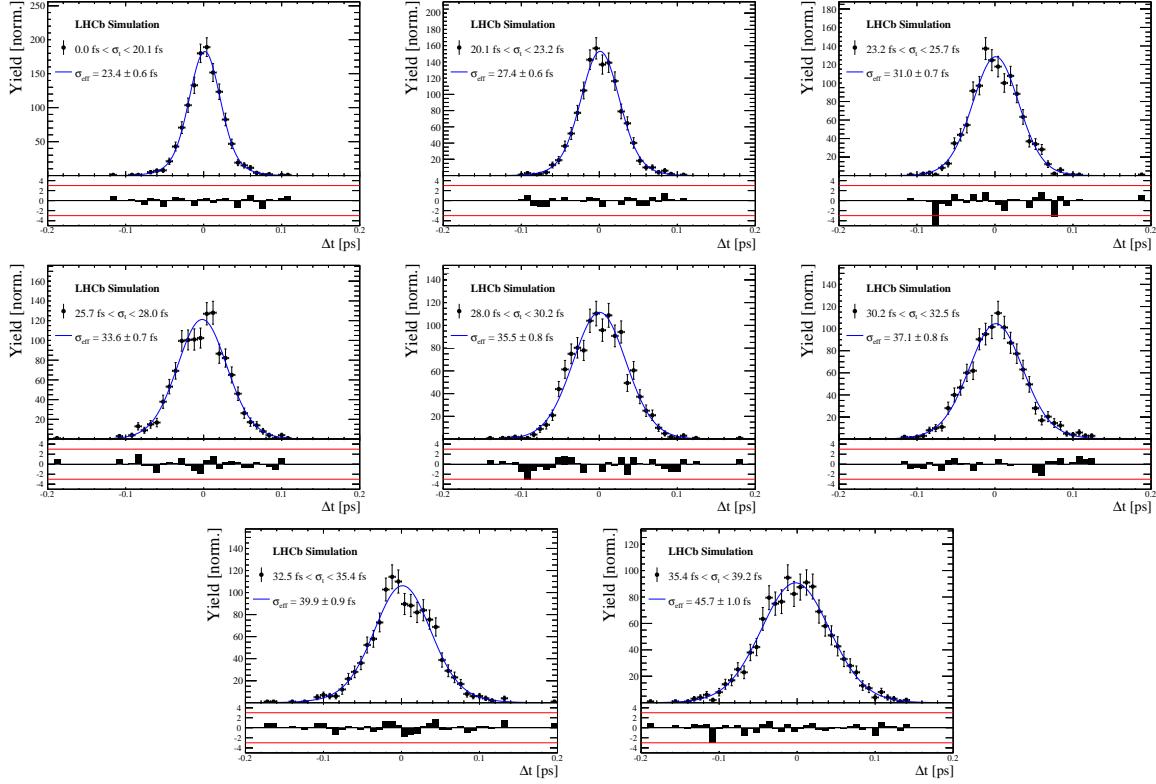
**Figure C.3:** Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for Run-I data.



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

## 1202 E Decay-time Resolution fits

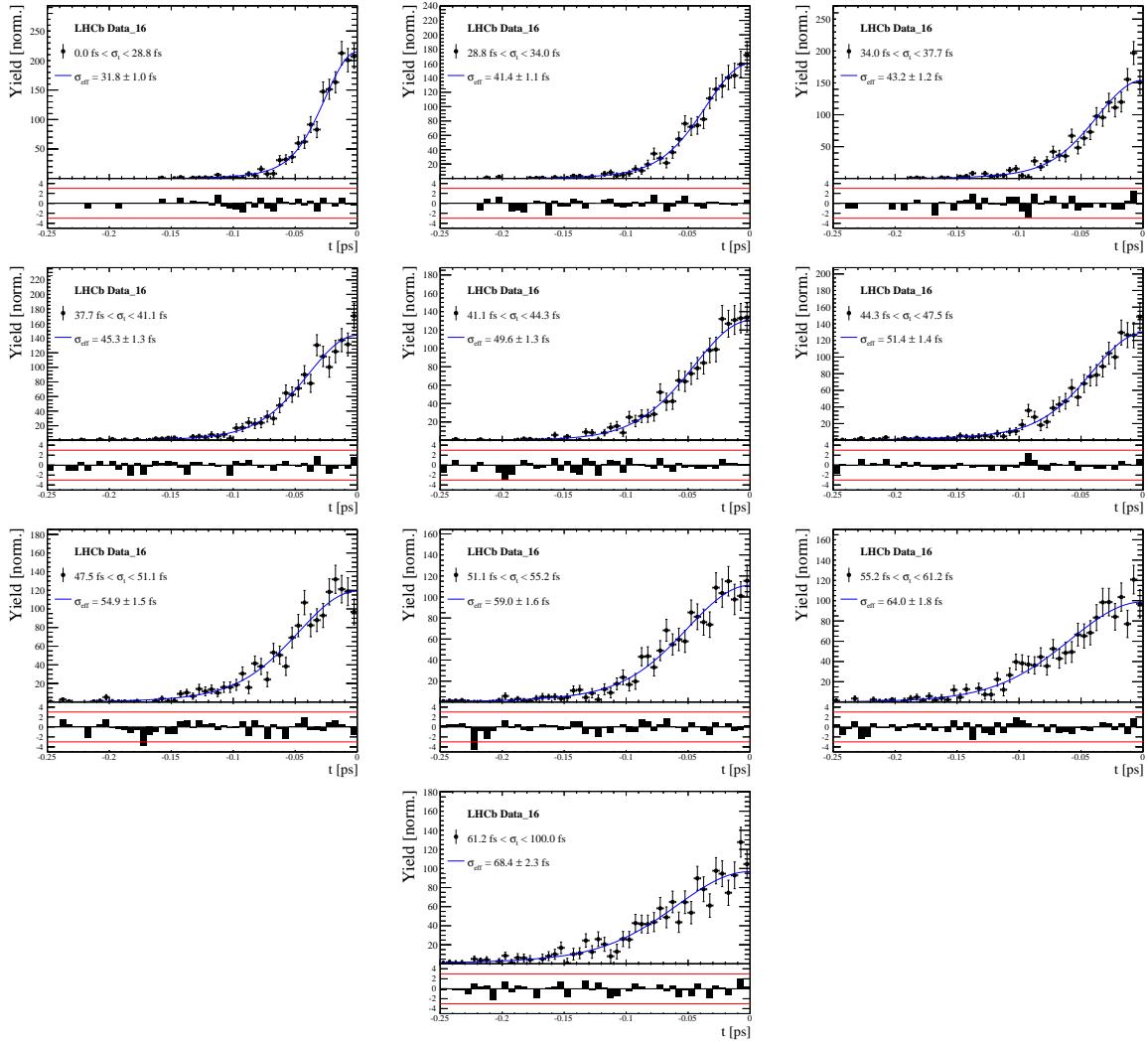
1203 This section contains all fits to the distributions of the decay time difference  $\Delta t$  between  
 1204 the true and the reconstructed decay time of the truth-matched  $B_s^0$  candidates on MC.  
 1205 The fits are performed in bins of the decay time error  $\sigma_t$ , where an adaptive binning  
 1206 scheme is used to ensure that approximately the same number of events are found in each  
 1207 bin.



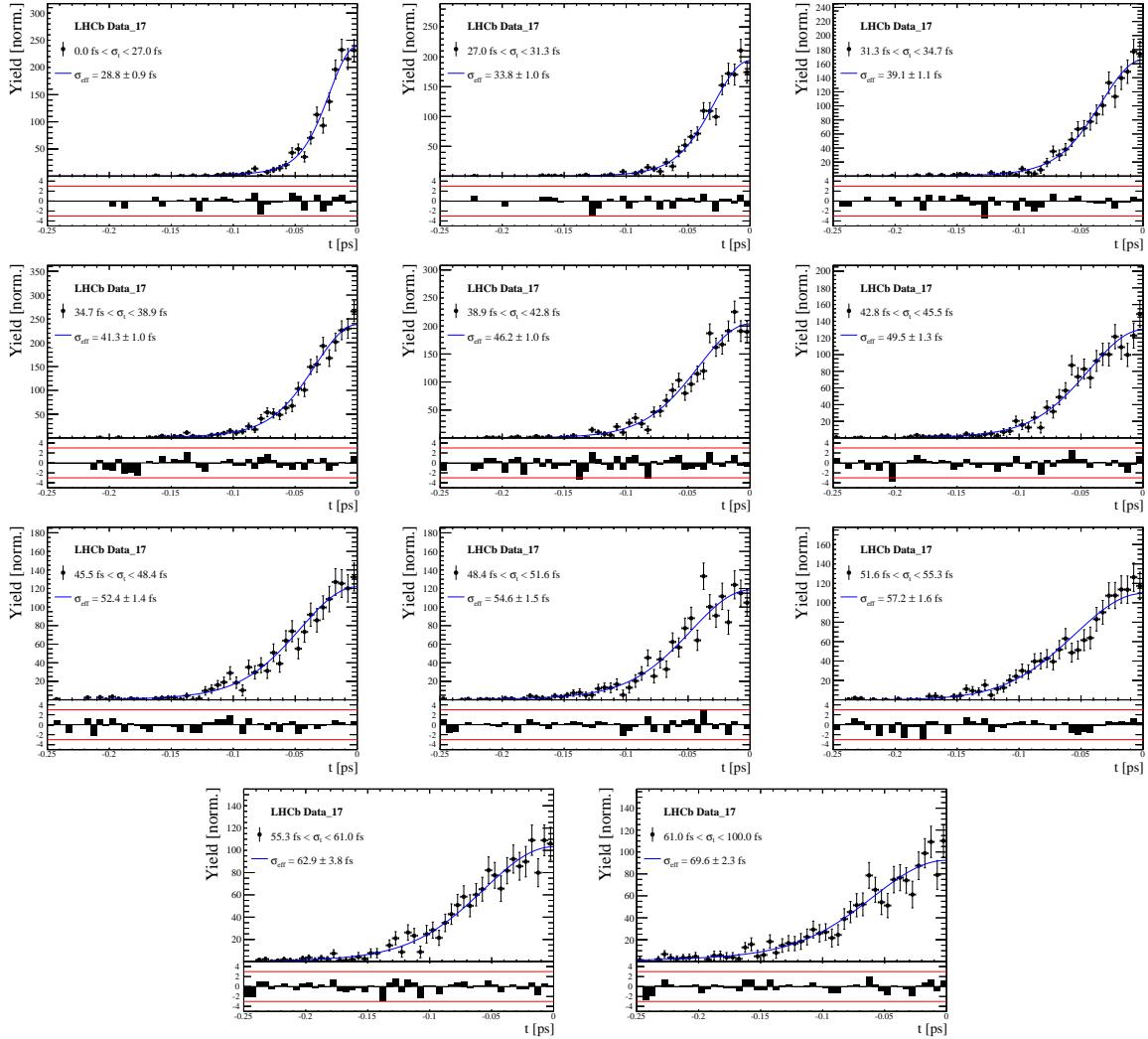
**Figure D.1:** Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K\pi\pi$  MC candidates in bins of the per-event decay time error estimate..

**Table D.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 D.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 D.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 D.2:** Measured time resolution for prompt- $D_s$  data in bins of the per-event decay time error estimate. Data taken in 2016.

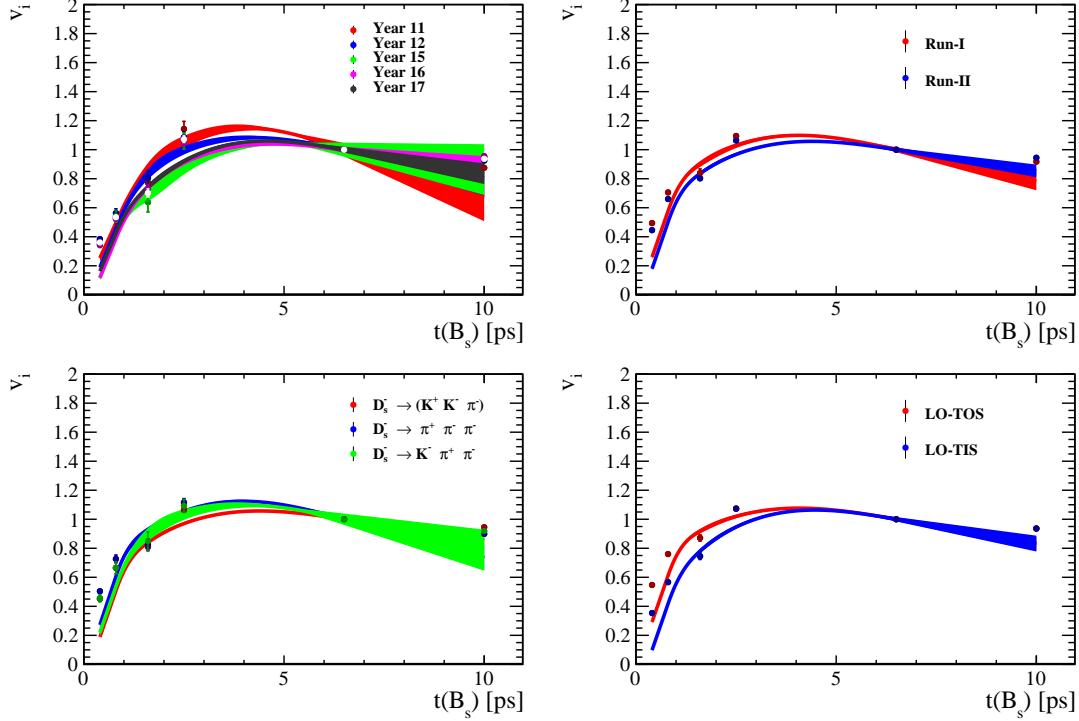
$\sigma_t$ Bin [fs]	$\sigma_1$ [fs]	$\sigma_2$ [fs]	$f_1$	D	$\sigma_{eff}$ [fs]
0.0 - 28.8	$25.3 \pm 1.05$	$47.8 \pm 2.82$	$0.75 \pm 0$	$0.853 \pm 0.00827$	$31.8 \pm 0.967$
28.8 - 34.0	$34.4 \pm 1.38$	$60.4 \pm 3.66$	$0.75 \pm 0$	$0.763 \pm 0.0114$	$41.4 \pm 1.14$
34.0 - 37.7	$35.9 \pm 1.7$	$63.1 \pm 4.05$	$0.75 \pm 0$	$0.745 \pm 0.0122$	$43.2 \pm 1.2$
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 D.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	$58.4 \pm 3.18$	$118 \pm 15.6$	$0.75 \pm 0$	$0.466 \pm 0.0233$	$69.6 \pm 2.28$

1208 **F Comparison of time-acceptance in subsamples**

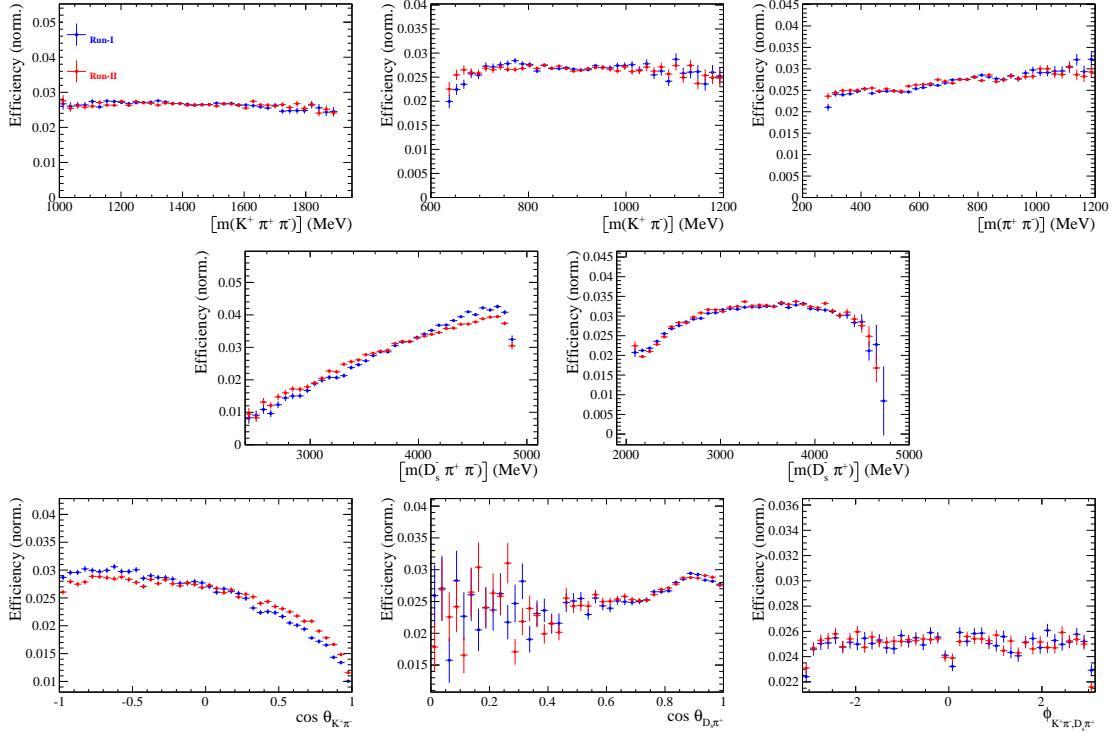
1209 Figure E.1 shows the spline coefficients obtained by fitting the decay-time distribution of  
 1210  $B_s^0 \rightarrow D_s\pi\pi\pi$  data candidates in different subsamples. Sufficient agreement is observed  
 1211 within a given data-taking period, while the acceptance shapes for Run-I and Run-II  
 1212 data differ significantly. The fitted splines for the different  $D_s$  final states are in a good  
 1213 agreement. The largest deviations are observed between the different L0 categories.



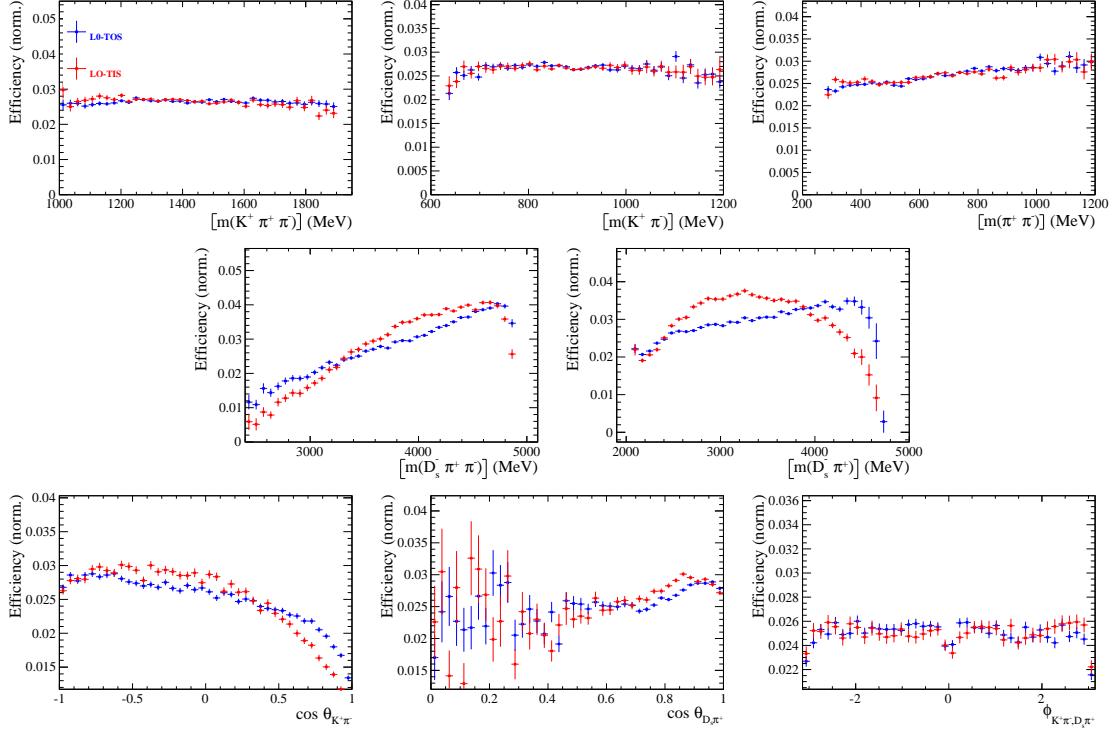
**Figure E.1:** Comparison of the spline coefficients (point with error bars) obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  decay-time for different subsamples: (top-left) different years of data-taking; (top-right) different data-taking periods; (bottom-left) different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

1214 **G Comparison of phase-space acceptance in subsam-  
 1215 ples**

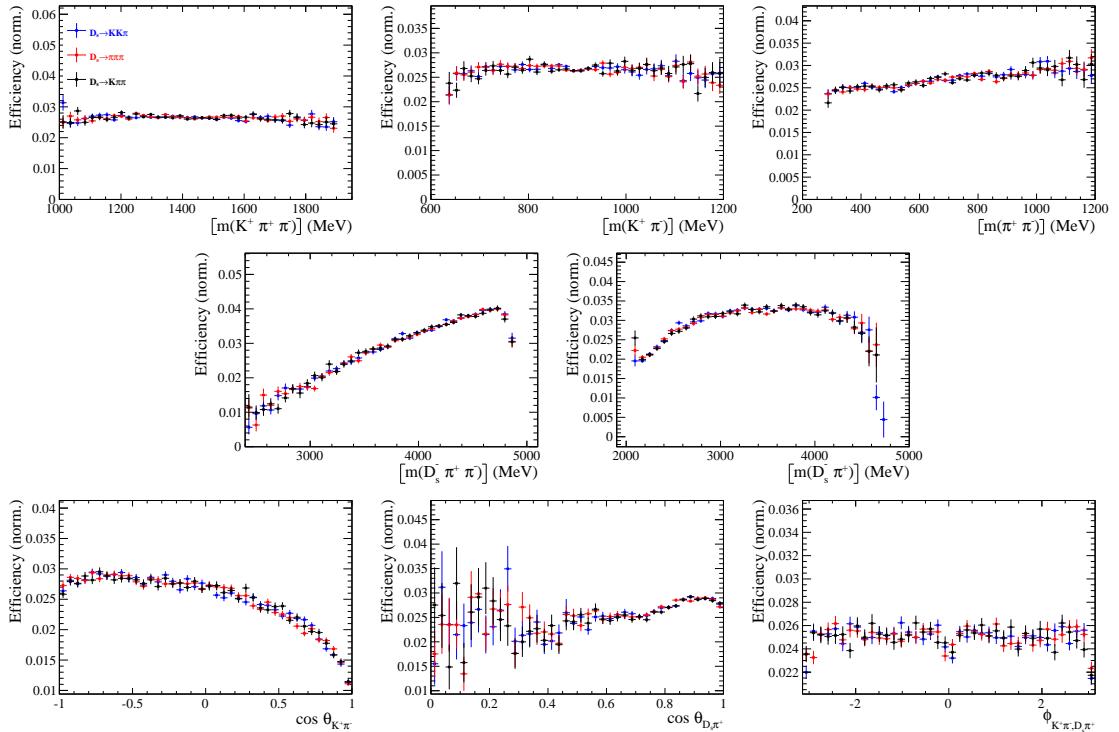
1216 Figures F.1, F.2 and F.3 compare the phase space-acceptance projections obtained from  
 1217  $B_s \rightarrow D_s K\pi\pi$  MC in different subsamples. Sufficient agreement is observed between  
 1218 different data-taking periods and  $D_s$  final states. The largest deviations are observed  
 1219 between the different L0 categories.



**Figure F.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 F.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 F.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  $D_s \rightarrow K\pi\pi$ .

<sub>1220</sub> **H 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 G.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 G.2:** Calibration parameters of the OS taggers for Run-II.

# I Spin Amplitudes

The spin factors used for  $B \rightarrow P_1 P_2 P_3 P_4$  decays are given in Table H.1.

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

## <sub>1223</sub> J Considered Decay Chains

<sub>1224</sub> The various decay channels considered in the model building are listed in Table I.1.

**Table I.1:** Decays considered in the LASSO model building.

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

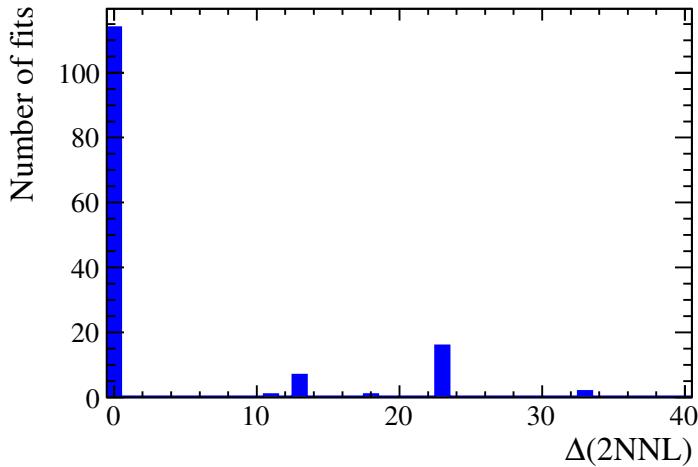
## K Additional information for the time-dependent amplitude fit

Table I.1 summarizes the fixed parameters for the resonances included in the baseline model. The parameters of the resonances  $K_1^+(1400)$  and  $K^{*+}(1410)$  are determined in the fit and their PDG values are only shown for comparison. Interference fractions are listed in Tables I.2 and I.3.

Figure F.1 shows the distribution of likelihood values for 150 time-dependent amplitude fits with randomized starting parameters. The majority of the fits (76%) converge to the global minimum. Two well isolated local minima are detected at  $\Delta(2NLL) = 13$  ( $3.3\sigma$ ) and  $\Delta(2NLL) = 23$  ( $4.8\sigma$ ) to which 5% and 11% of the fits converge, respectively. The fits at  $\Delta(2NLL) = 11$ ,  $\Delta(2NLL) = 18$  and  $\Delta(2NLL) = 33$  did not converge properly to one of the nearby local minima. The remaining fits (5%) have  $\Delta(2NLL) > 500$  and are not shown. This procedure has been repeatedly performed during the development of the analysis, for both LASSO stages and the final fit, to ensure we operate at the global minimum.

**Table I.1:** Parameters of the resonances chosen at the model selection stage.

Resonance	$m$ [ MeV ]	$\Gamma$ [ MeV ]	Source
$\rho(770)$	$775.26 \pm 0.25$	$149.1 \pm 0.8$	[15]
$K^{*0}(892)$	$895.55 \pm 0.20$	$47.3 \pm 0.5$	[15]
$K_1^+(1270)$	$1289.81 \pm 1.75$	$116.11 \pm 3.4$	[23]
$K^+(1460)$	$1482.4 \pm 15.64$	$335.6 \pm 10.64$	[23]
$K_1^+(1400)$	$1403 \pm 7$	$174 \pm 13$	[15]
$K^{*+}(1410)$	$1421 \pm 9$	$236 \pm 18$	[15]
$K^{0*}(1430)$	$1425 \pm 50$	$270 \pm 80$	[15]



**Figure F.1:** Likelihood differences with respect to the smallest value obtained from 150 fits with randomized starting parameters.

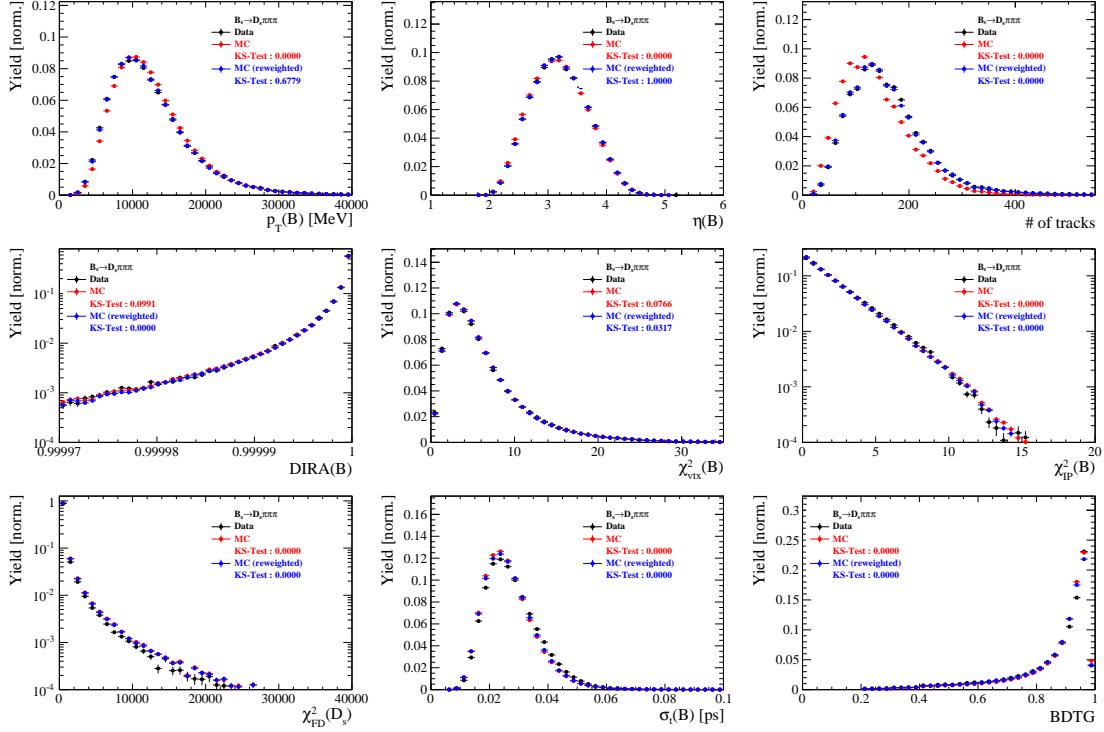
**Table I.2:** Interference fractions of the amplitudes contributing to  $b \rightarrow c$  decays.

Decay Channel $i$	Decay Channel $j$	$I_{ij}[\%]$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s \pi)_P K^*(892)$	$-10.1 \pm 4.2$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$8.8 \pm 4.1$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s \pi)_P K^*(892)$	$-5.6 \pm 1.3$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$5.3 \pm 0.5$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$5.1 \pm 0.3$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$-2.3 \pm 0.8$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$0.5 \pm 0.2$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$-0.2 \pm 0.4$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$-0.1 \pm 0.0$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$0.0 \pm 0.0$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s \pi)_P K^*(892)$	$-0.0 \pm 0.0$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$0.0 \pm 0.0$
$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K^*(1410) \rightarrow K^*(892) \pi)$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K^*(1410) \rightarrow K \rho(770))$	$0.0 \pm 0.0$

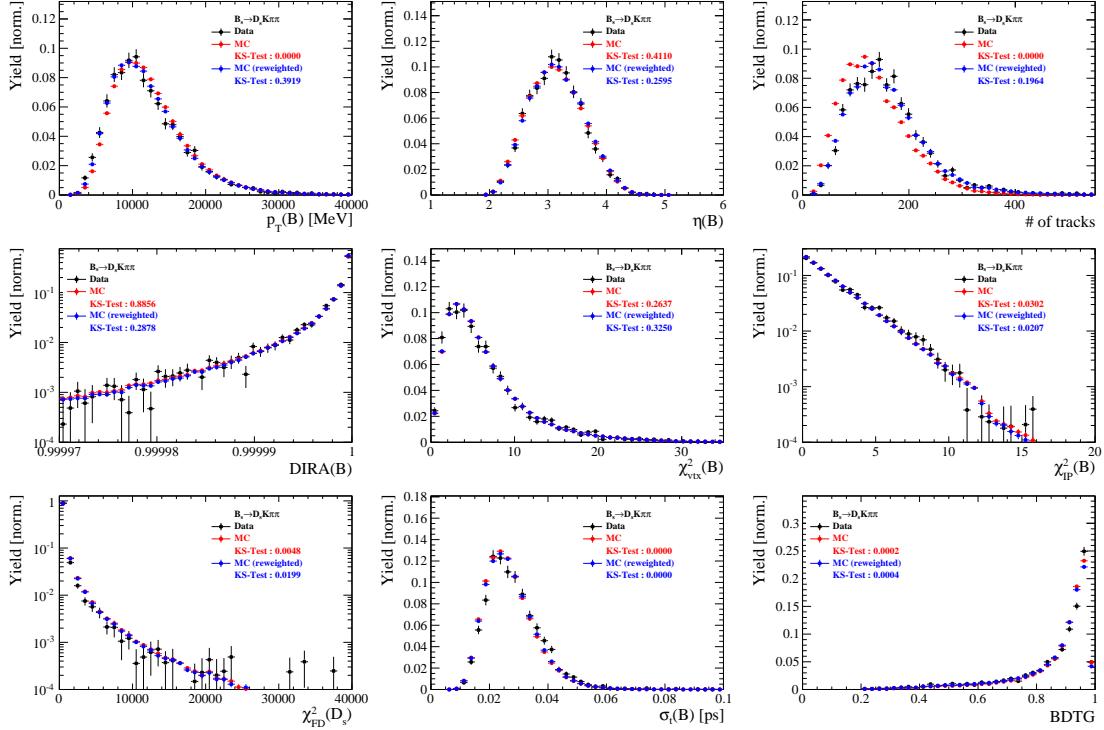
**Table I.3:** Interference fractions of the amplitudes contributing to  $b \rightarrow u$  decays.

Decay Channel $i$	Decay Channel $j$	$I_{ij}[\%]$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s \pi)_P K^*(892)$	$-20.7 \pm 4.9$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$-20.0 \pm 8.7$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s \pi)_P K^*(892)$	$18.9 \pm 9.7$
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s \pi)_P K^*(892)$	$-8.4 \pm 1.2$
$B_s \rightarrow (D_s K)_P \rho(770)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$-6.3 \pm 1.2$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$-5.5 \pm 1.7$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$B_s \rightarrow (D_s K)_P \rho(770)$	$3.1 \pm 0.6$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s K)_P \rho(770)$	$2.0 \pm 0.6$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$2.0 \pm 1.3$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$1.7 \pm 0.8$
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s K)_P \rho(770)$	$-1.5 \pm 0.3$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow (D_s K)_P \rho(770)$	$-1.3 \pm 0.4$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$-0.6 \pm 1.0$
$B_s \rightarrow (D_s \pi)_P K^*(892)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$0.1 \pm 0.0$
$B_s \rightarrow (D_s K)_P \rho(770)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$-0.1 \pm 2.9$
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K \rho(770))$	$0.1 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	$0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1270) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$-0.0 \pm 0.0$
$B_s \rightarrow D_s (K_1(1400) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$0.0 \pm 0.0$
$B_s \rightarrow D_s (K(1460) \rightarrow K^*(892) \pi)$	$B_s \rightarrow D_s (K_1(1270) \rightarrow K_0^*(1430) \pi)$	$0.0 \pm 0.0$

## L Data-simulation comparisson



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



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

1241 M Data distributions

1242 Comparison of signal and calibration channels

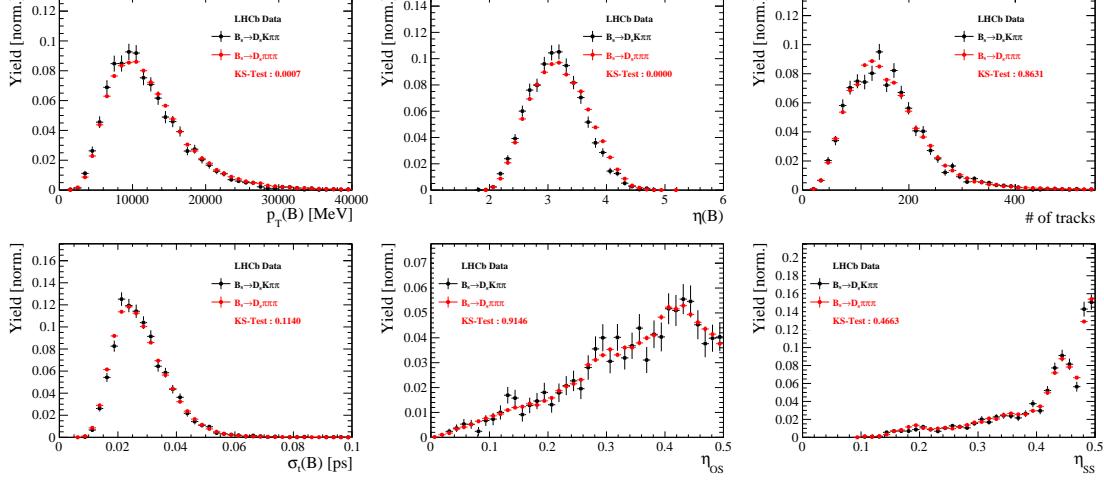


Figure K.1: Comparison between  $B_s \rightarrow D_s K\pi\pi$  and  $B_s \rightarrow D_s \pi\pi\pi$  decays for selected variables.

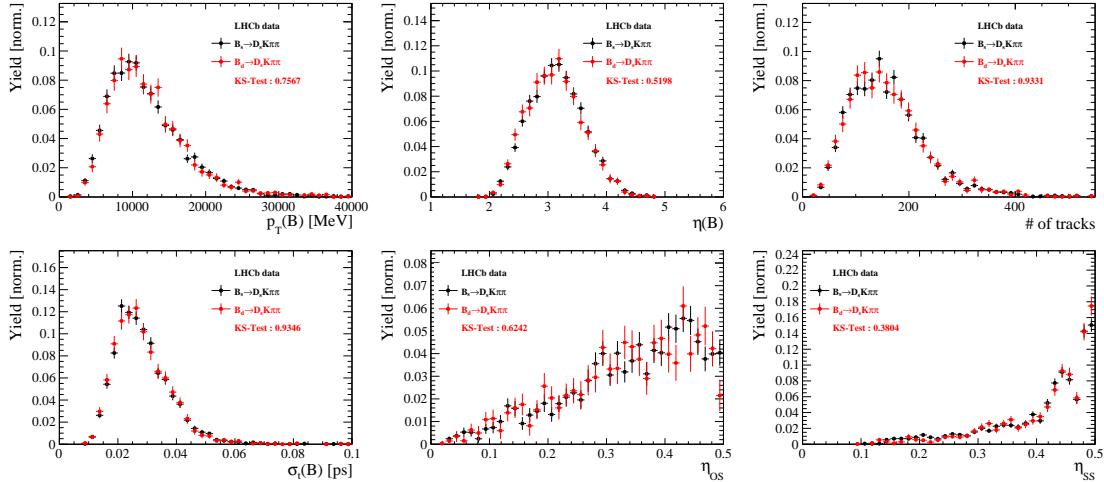


Figure K.2: Comparison between  $B_s \rightarrow D_s K\pi\pi$  and  $B_d \rightarrow D_s K\pi\pi$  decays for selected variables.

1243 Comparison of data taken in 2016 and 2017

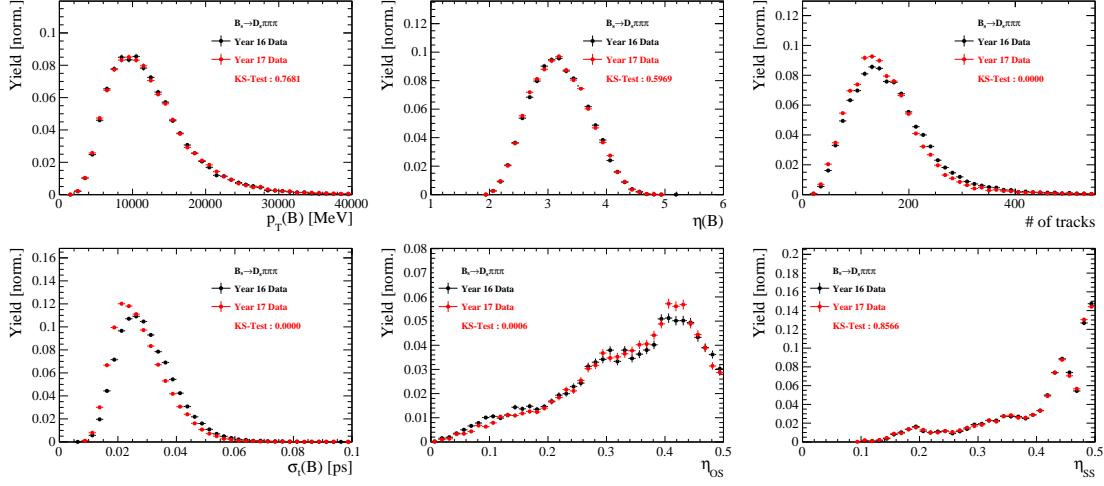


Figure K.3: Comparison of selected variables for  $B_s \rightarrow D_s \pi\pi\pi$  data taken in 2016 and 2017.

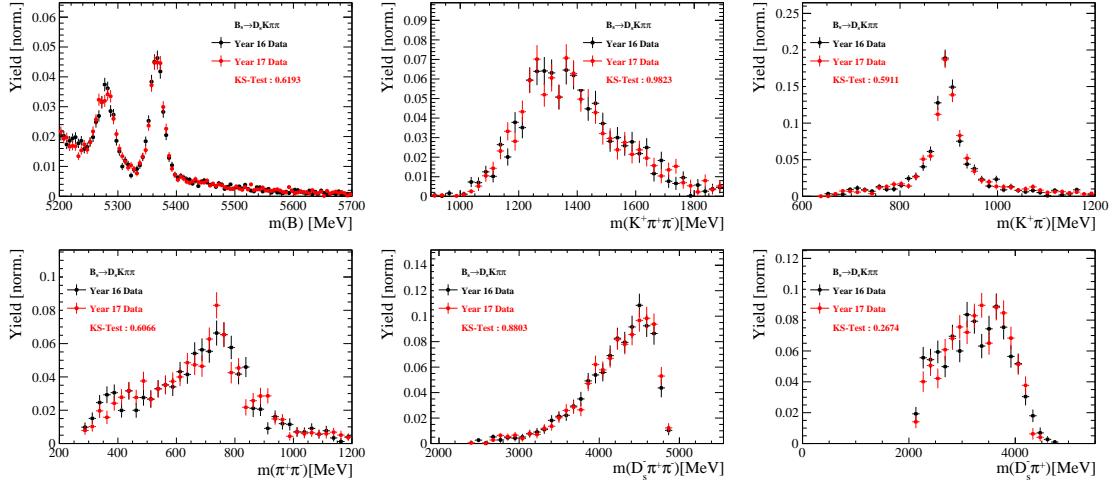
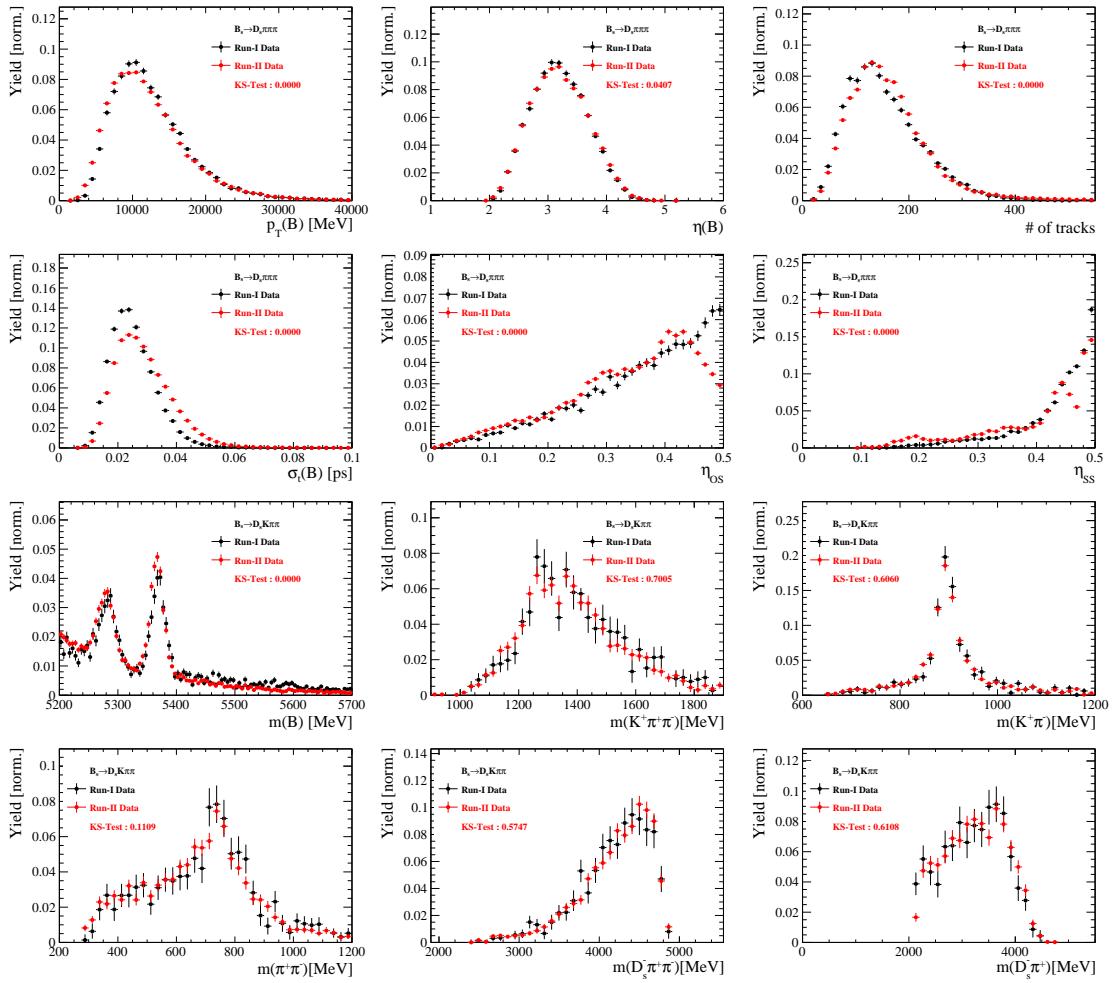


Figure K.4: Comparison of selected variables for  $B_s \rightarrow D_s K\pi\pi$  data taken in 2016 and 2017.

# Comparison of Run-I and Run-II data



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

1245 Comparison of  $D_s$  final states

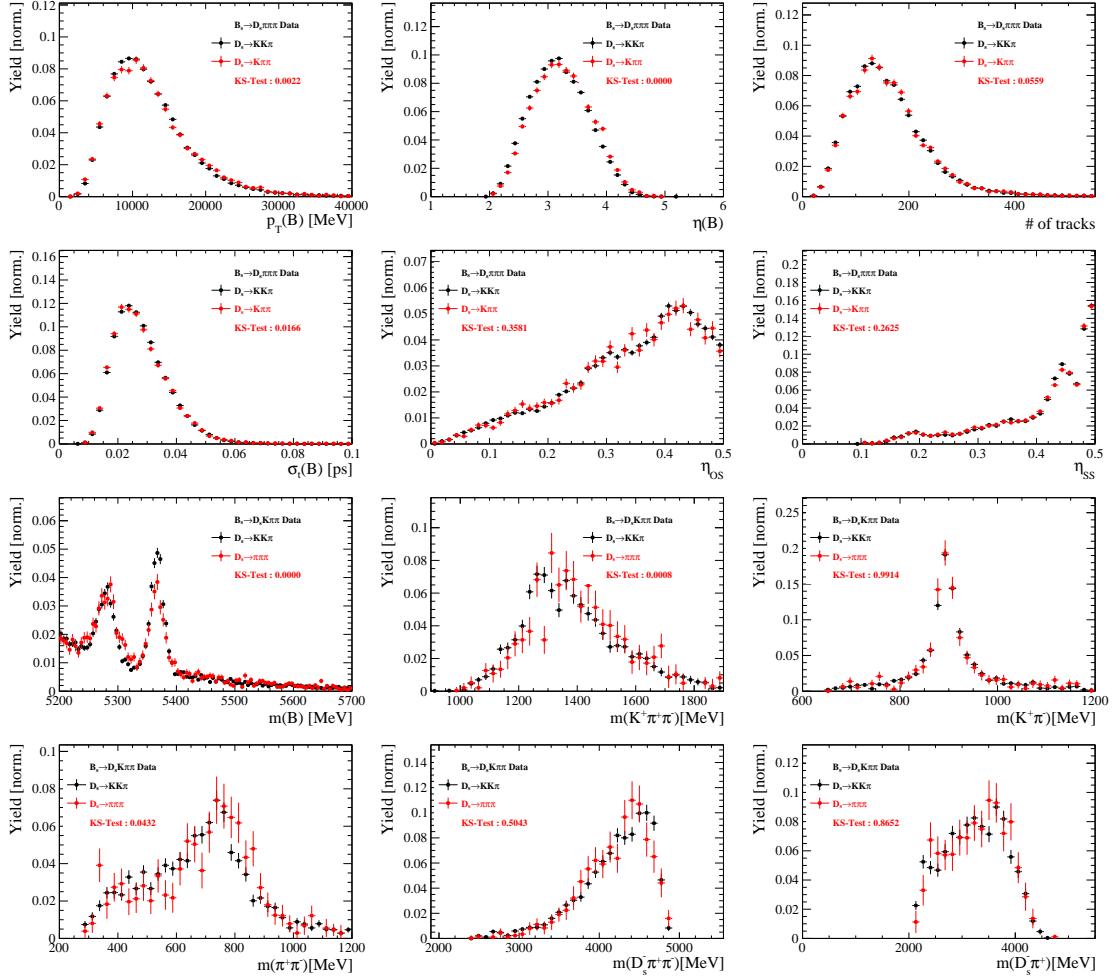
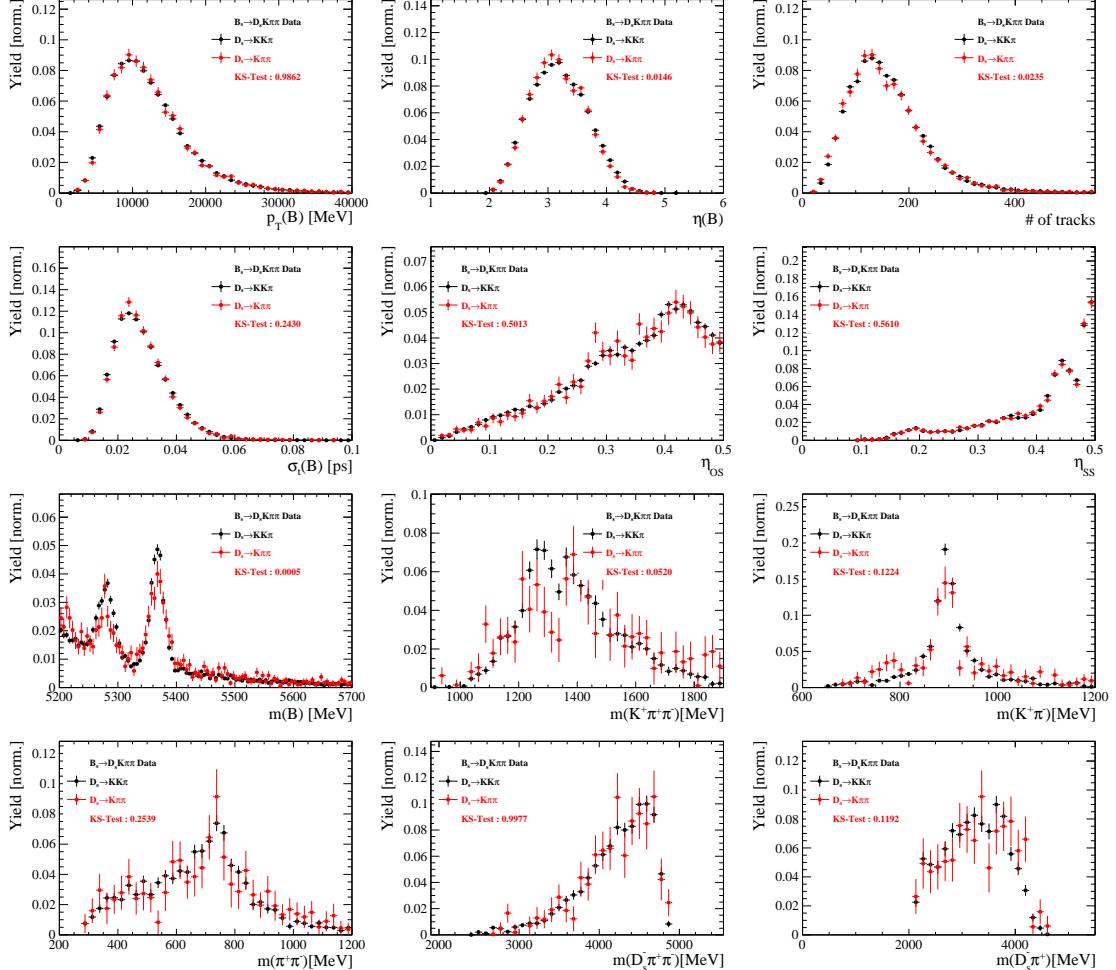


Figure K.6: Comparison of selected variables for different  $D_s$  final states.



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

1246 Comparison of trigger categories

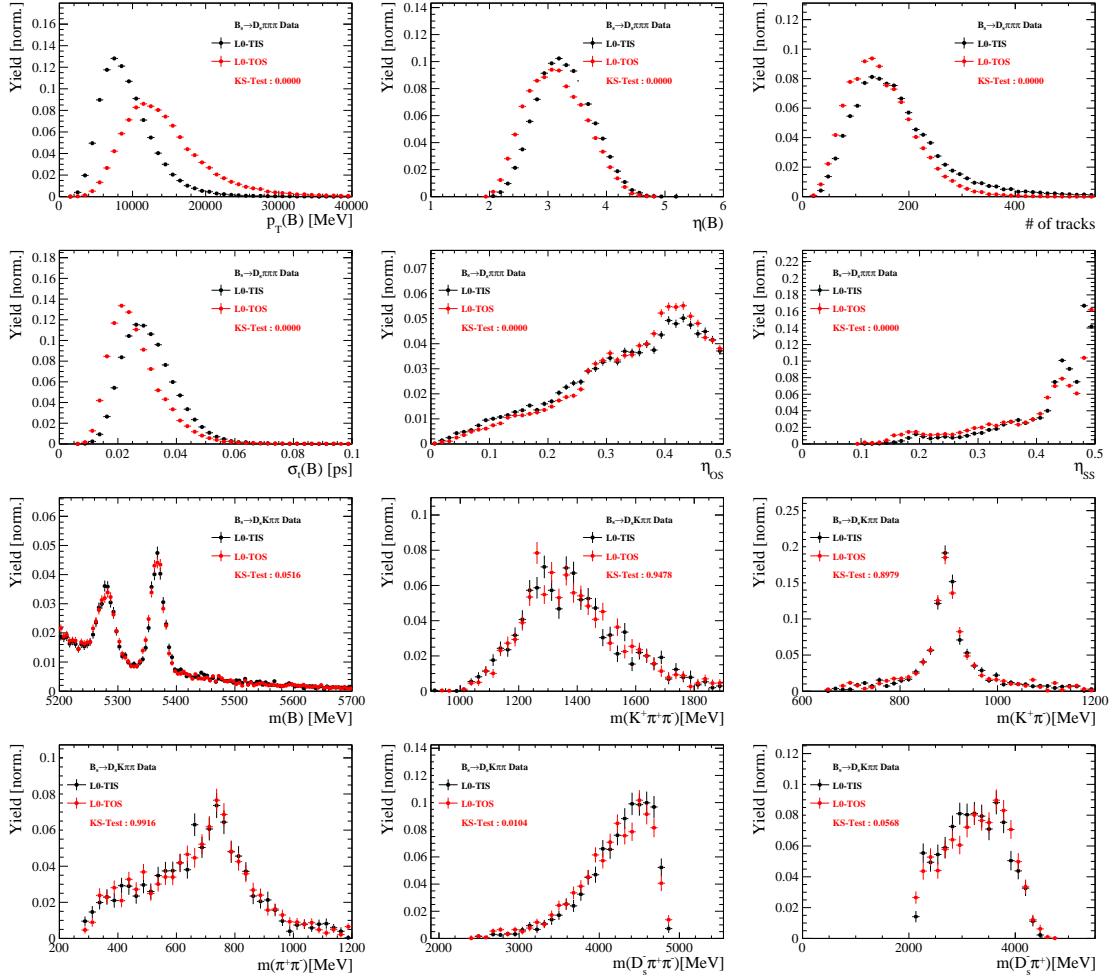


Figure K.8: Comparison of selected variables for different trigger categories.

1247 **References**

- 1248 [1] S. Blusk, *First observations and measurements of the branching fractions for the*  
1249 *decays  $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$  and  $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* , LHCb-ANA-2012-076.
- 1250 [2] LHCb, S. Blusk, *Measurement of the CP observables in  $\bar{B}_s^0 \rightarrow D_s^+ K^-$  and first obser-*  
1251 *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](#).
- 1252 [3] LHCb collaboration, R. Aaij *et al.*, *First observation of the decays  $\bar{B}_{(s)}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$*   
1253 *and  $\bar{B}_s^0 \rightarrow D_{s1}(2536)^+ \pi^-$* , Phys. Rev. **D86** (2012) 112005, [arXiv:1211.1541](#).
- 1254 [4] R. Fleischer, *New strategies to obtain insights into CP violation through  $B_s \rightarrow$*   
1255  *$D_s^\pm K^\mp, D_s^{*\pm} K^\mp, \dots$  and  $B_d \rightarrow D^\pm \pi^\mp, D^{*\pm} \pi^\mp, \dots$  decays*, Nucl. Phys. **B671** (2003) 459,  
1256 [arXiv:hep-ph/0304027](#).
- 1257 [5] K. De Bruyn *et al.*, *Exploring  $B_s \rightarrow D_s^{(*)\pm} K^\mp$  Decays in the Presence of a Sizable*  
1258 *Width Difference  $\Delta\Gamma_s$* , Nucl. Phys. **B868** (2013) 351, [arXiv:1208.6463](#).
- 1259 [6] L. collaboration, *Measurement of time-dependent CP-violation observables in  $B_s^0 \rightarrow$*   
1260  *$D_s^\mp K^\pm$* , LHCb-ANA-2016-052.
- 1261 [7] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in  $B_s^0 \rightarrow D_s^\mp K^\pm$  decays*, JHEP  
1262 **03** (2018) 059, [arXiv:1712.07428](#).
- 1263 [8] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 1264 [9] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation*  
1265 *of production processes*, Annals of Physics **18** (1962), no. 2 198 .
- 1266 [10] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*,  
1267 Phys. Rev. D **11** (1975) 3165.
- 1268 [11] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of  
1269 Physics **108** (1977), no. 2 454 .
- 1270 [12] P. d'Argent *et al.*, *Amplitude Analyses of  $D^0 \rightarrow \pi^+ \pi^- \pi^+ \pi^-$  and  $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$*   
1271 *Decays*, JHEP **05** (2017) 143, [arXiv:1703.08505](#).
- 1272 [13] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay*  
1273 *widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.
- 1274 [14] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo  
1275 Cimento Series 10 **34** (1964), no. 6 1644.
- 1276 [15] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys.  
1277 **C40** (2016), no. 10 100001.
- 1278 [16] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- 1279 [17] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys.  
1280 Rev. **60** (1941) 61.

- [1281] [18] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys.  
 1282 Rev. D **57** (1998) 431.
- [1283] [19] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of*  
 1284  $\psi$  *decay to mesons*, Eur. Phys. J. **A16** (2003) 537, arXiv:hep-ph/0211457.
- [1285] [20] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spec-*  
 1286 *troscopy*, Phys. Rev. **D51** (1995) 2247.
- [1287] [21] J.-J. Zhu, *Explicit expressions of spin wave functions*, arXiv:hep-ph/9906250.
- [1288] [22] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput.  
 1289 Phys. Commun. **180** (2009) 1847, arXiv:0805.2956.
- [1290] [23] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in  $D^0 \rightarrow K^\mp\pi^\pm\pi^\pm\pi^\mp$  decays*,  
 1291 Eur. Phys. J. **C78** (2018), no. 6 443, arXiv:1712.08609.
- [1292] [24] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth.  
 1293 **A462** (2001) 152.
- [1294] [25] M. Karbach and M. Kenzie, *Gammacombo package*,  
 1295 http://gammacombo.hepforge.org/web/HTML/index.html, 2014.
- [1296] [26] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007)  
 1297 040, arXiv:physics/0703039.
- [1298] [27] N. L. Johnson, *Systems of frequency curves generated by methods of translation*,  
 1299 Biometrika **36** (1949), no. 1/2 149.
- [1300] [28] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38**  
 1301 (2014) 090001.
- [1302] [29] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys.  
 1303 **A30** (2015) 1530022, arXiv:1412.6352.
- [1304] [30] M. Schiller, *Proper time resolution studies in  $B_s \rightarrow D_s\pi$* , Tech. Rep. LHCb-INT-  
 1305 2011-005. CERN-LHCb-INT-2011-005, CERN, Geneva, Feb, 2011.
- [1306] [31] LHCb collaboration, L. Zhang, *Measurements of CP violation in  $B_s^0 \rightarrow J/\psi K^+K^-$*   
 1307 *decays in the low  $K^+K^-$  mass range with 13 TeV data*, LHCb-ANA-2017-028.
- [1308] [32] LHCb, R. Aaij *et al.*, *Updated measurement of time-dependent CP-violating ob-*  
 1309 *servables in  $B_s^0 \rightarrow J/\psi K^+K^-$  decays*, Eur. Phys. J. **C79** (2019), no. 8 706,  
 1310 arXiv:1906.08356.
- [1311] [33] LHCb collaboration, *Precision measurement of the  $B_s - \bar{B}_s$  oscillation frequency with*  
 1312 *the decay  $B_s \rightarrow D_s\pi$* , LHCb-ANA-2019-XXX.
- [1313] [34] A. Poluektov, *Correction of simulated particle identification response in LHCb using*  
 1314 *kernel density estimation*, LHCb-INT-2017-007.

- 1315 [35] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of  $b$ -hadron,  $c$ -hadron, and*  
 1316  *$\tau$ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots  
 1317 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 1318 [36] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*  
 1319 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 1320 [37] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of  $B$  mesons at the*  
 1321 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- 1322 [38] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of  $B_s^0$  mesons at*  
 1323 *LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- 1324 [39] LHCb, R. Aaij *et al.*, *Measurement of  $B^0$ ,  $B_s^0$ ,  $B^+$  and  $\Lambda_b^0$  production asymmetries in 7*  
 1325 *and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- 1326 [40] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of*  
 1327 *the  $K\pi$  Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-  
 1328 2012-027, CERN, Geneva, Feb, 2013.
- 1329 [41] A. Davis *et al.*, *Measurement of the instrumental asymmetry for  $K^- \pi^+$ -pairs at LHCb*  
 1330 *in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN,  
 1331 Geneva, Mar, 2018.
- 1332 [42] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly*  
 1333 *forbidden transitions in  $D \rightarrow K(S)$ ,  $K(L) + \pi$ 's decays*, Phys. Lett. **B349** (1995)  
 1334 363, arXiv:hep-ph/9502238.
- 1335 [43] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*,  
 1336 Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- 1337 [44] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal  
 1338 Statistical Society, Series B **58** (1994) 267.
- 1339 [45] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude*  
 1340 *analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- 1341 [46] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- 1342 [47] CLEO Collaboration, M. Artuso *et al.*, *Amplitude analysis of  $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$* ,  
 1343 Phys. Rev. **D85** (2012) 122002, arXiv:1201.5716.
- 1344 [48] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime*  
 1345 *and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986,  
 1346 DESY-F31-86-02.
- 1347 [49] D. Hill, M. John, and P. Gandini, *A study of partially reconstructed  $B^\pm \rightarrow D^{*0} h^\pm$*   
 1348 *decays using the  $D^0 \rightarrow K\pi, KK, \pi\pi$  final states*, LHCb-ANA-2017-018.
- 1349 [50] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins  
 1350 University Press, Baltimore, MD, USA, 1996.

- 1351 [51] L. collaboration, *Precision measurement of the  $B_s$  oscillation frequency with the decay*  
 1352  $B_s \rightarrow D_s\pi$ , LHCb-ANA-2012-053.
- 1353 [52] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the  $B_s^0 - \bar{B}_s^0$  oscillation frequency in the decay  $B_s^0 \rightarrow D_s^-\pi^+$* , New J. Phys. **15** (2013) 053021,  
 1354 arXiv:1304.4741.
- 1356 [53] BaBar, P. del Amo Sanchez *et al.*, *Measurement of  $D0$ -anti $D0$  mixing parameters using  $D0 \rightarrow K(S)0$  pi+ pi- and  $D0 \rightarrow K(S)0$   $K+K-$  decays*, Phys. Rev. Lett. **105**  
 1357 (2010) 081803, arXiv:1004.5053.
- 1359 [54] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in  $D^0 \rightarrow K_S^0 K^\pm \pi^\mp$  decays*,  
 1360 Phys. Rev. **D93** (2016), no. 5 052018, arXiv:1509.06628.
- 1361 [55] D. V. Bugg, *The mass of the  $\sigma$  pole*, Journal of Physics G Nuclear Physics **34** (2007)  
 1362 151, arXiv:hep-ph/0608081.
- 1363 [56] W. Dunwoodie. Fits to  $K\pi$   $I = \frac{1}{2}$   $S$ -wave amplitude and phase data.
- 1364 [57] D. Aston *et al.*, *A Study of  $K^-\pi^+$  Scattering in the Reaction  $K^-p \rightarrow K^-\pi^+n$  at*  
 1365  $11\text{GeV}/c$ , Nucl. Phys. **B296** (1988) 493.
- 1366 [58] BaBar, B. Aubert *et al.*, *Dalitz-plot analysis of the decays  $B^\pm \rightarrow K^\pm \pi^\mp \pi^\pm$* ,  
 1367 Phys. Rev. **D72** (2005) 072003, arXiv:hep-ex/0507004, [Erratum: Phys.  
 1368 Rev.D74,099903(2006)].
- 1369 [59] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-dominance prediction for  $\rho \rightarrow e^+e^-$* , Phys. Rev. Lett. **21** (1968) 244.
- 1371 [60] CMD-2, R. R. Akhmetshin *et al.*, *Measurement of  $e^+e^- \rightarrow \pi^+\pi^-$  cross-section with*  
 1372 *CMD-2 around rho meson*, Phys. Lett. **B527** (2002) 161, arXiv:hep-ex/0112031.
- 1373 [61] M. S. et al. *Search for CP violation in the  $D^0 \rightarrow KK\pi\pi$  decay through a full*  
 1374 *amplitude analysis*, LHCb-ANA-2017-064.
- 1375 [62] LHCb Collaboration, R. Aaij, *Search for CP violation through an amplitude analysis*  
 1376 *of  $D^0 \rightarrow K^+K^-\pi^+\pi^-$  decays*, Tech. Rep. arXiv:1811.08304. LHCb-PAPER-2018-041,  
 1377 CERN, Geneva, Nov, 2018. \*Temporary entry\*.
- 1378 [63] S. M. Flatté, *Coupled-channel analysis of the  $\pi\eta$  and  $KK$  systems near  $KK$  threshold*,  
 1379 Physics Letters B **63** (1976), no. 2 224 .
- 1380 [64] BES Collaboration, M. Ablikim *et al.*, *Resonances in  $J/\psi \rightarrow \phi\pi^+\pi^-$  and  $\phi K^+K^-$* ,  
 1381 Phys. Lett. **B607** (2005) 243, arXiv:hep-ex/0411001.