

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

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

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

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

<sup>3</sup>*School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom*

## Abstract

We present the first measurement of the weak phase  $2\beta + \gamma$  obtained from a time-dependent (amplitude) analysis of  $B_s^0 \rightarrow D_s K \pi\pi$  decays using proton-proton collision data corresponding to an integrated luminosity of **xxx** fb<sup>-1</sup> recorded by the LHCb detector.



# 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 . . . . .	3
2.2.1	Form Factors and Resonance Lineshapes . . . . .	4
2.2.2	Spin Densities . . . . .	5
2.3	Validation . . . . .	8
<b>3</b>	<b>Selection</b>	<b>12</b>
3.1	Cut-based selection . . . . .	12
3.2	Multivariate stage . . . . .	13
<b>4</b>	<b>Fits to invariant mass distributions of signal and normalization channel</b>	<b>16</b>
4.1	Signal models for $m(D_s\pi\pi\pi)$ and $m(D_sK\pi\pi)$ . . . . .	16
4.2	Background models for $m(D_s\pi\pi\pi)$ . . . . .	17
4.3	Background models for $m(D_sK\pi\pi)$ . . . . .	17
4.4	Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates . . . . .	19
4.5	Fit to $B_s^0 \rightarrow D_sK\pi\pi$ candidates . . . . .	19
4.6	Extraction of signal weights . . . . .	19
<b>5</b>	<b>Flavour Tagging</b>	<b>21</b>
5.1	OS tagging calibration . . . . .	22
5.2	SS tagging calibration . . . . .	22
5.3	Tagging performance comparison between the signal and normalization channel . . . . .	22
5.4	Combination of OS and SS taggers . . . . .	23
<b>6</b>	<b>Acceptance</b>	<b>26</b>
6.1	MC corrections . . . . .	26
6.1.1	Truth matching of simulated candidates . . . . .	26
6.1.2	PID efficiencies . . . . .	27
6.1.3	BDT efficiencies . . . . .	28
6.1.4	Tracking efficiencies . . . . .	29
6.2	Decay-time acceptance . . . . .	30
6.2.1	Comparison of acceptance in subsamples . . . . .	31
6.2.2	Results . . . . .	33
6.3	Phasespace acceptance . . . . .	37
<b>7</b>	<b>Decay-time Resolution</b>	<b>38</b>
7.1	Calibration for Run-I data . . . . .	39
7.2	Calibration for Run-II data . . . . .	40
7.3	Cross-checks . . . . .	43
7.3.1	Kinematic dependence . . . . .	43
7.3.2	DTF constraints . . . . .	43

<b>8 Production and Detection Asymmetries</b>	<b>44</b>
8.1 $B_s$ Production Asymmetry . . . . .	44
8.2 $K^-\pi^+$ Detection Asymmetry . . . . .	45
<b>9 Time dependent fit</b>	<b>48</b>
9.1 sFit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data . . . . .	48
9.2 sFit to $B_s^0 \rightarrow D_sK\pi\pi$ data . . . . .	48
<b>10 Time dependent amplitude fit</b>	<b>50</b>
10.1 Signal Model Construction . . . . .	50
10.2 Results . . . . .	51
<b>A Details of multivariate classifier</b>	<b>52</b>
<b>B Detailed mass fits</b>	<b>53</b>
<b>C Decay-time Resolution fits</b>	<b>57</b>
<b>D Spin Amplitudes</b>	<b>61</b>
<b>E Considered Decay Chains</b>	<b>62</b>
<b>F MC corrections</b>	<b>63</b>
<b>G Data distributions</b>	<b>68</b>
G.1 Comparison of signal and calibration channel . . . . .	68
G.2 Comparison of Run-I and Run-II data . . . . .	70
G.3 Comparison of $D_s$ final states . . . . .	72
G.4 Comparison of trigger categories . . . . .	74
G.5 Comparison of $B_s$ and $B_d$ decays . . . . .	76
<b>References</b>	<b>77</b>

# 1 Introduction

- The weak phase  $\gamma$  is the least well known angle of the CKM unitary triangle. A key channel to measure  $\gamma$  is the time-dependent analysis of  $B_s^0 \rightarrow D_s K$  decays [1], [2].  
 The  $B_s^0 \rightarrow D_s K\pi\pi$  proceeds at tree level via the transitions shown in Fig. 1.1 a) and b).

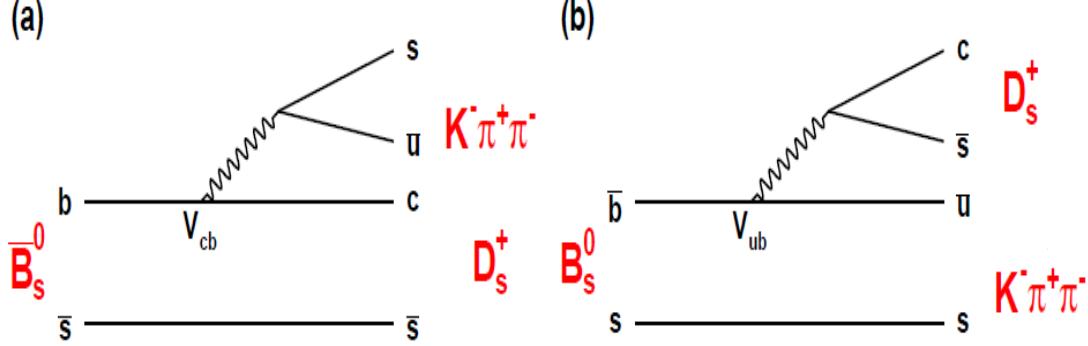


Figure 1.1: Feynman diagram of the  $B_s^0 \rightarrow D_s K\pi\pi$  decay, proceeding via a)  $b \rightarrow c$  transitions or b)  $b \rightarrow u$  transitions.

To measure the weak CKM phase  $\gamma \equiv \arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$ , a decay with interference between  $b \rightarrow c$  and  $b \rightarrow u$  transitions at tree level is needed [1]. As illustrated in Fig. 1.1, this is the case for the presented decay mode. A measurement of  $\gamma$  using  $B_s^0 \rightarrow D_s K\pi\pi$  decays, where the  $K\pi\pi$  subsystem is dominated by excited kaon states such as the  $K_1(1270)$  and  $K_1(1400)$  resonances, will succeed the branching ratio measurement presented in this note. It is complementary to the above mentioned analysis of  $B_s^0 \rightarrow D_s K$ , making use of a fully charged final state, where every track is detected in the vertex locator. To account for the non-constant strong phase across the Dalitz plot, one can either develop a time-dependent amplitude model or select a suitable phase-space region and introduce a coherence factor as additional hadronic parameter to the fit.  
 This analysis is based on the first observation of the  $B_s^0 \rightarrow D_s K\pi\pi$  decay presented in [3] and [4], where its branching ratio is measured relative to  $B_s^0 \rightarrow D_s \pi\pi\pi$ . The result obtained by the previous analysis is  $0.052 \pm 0.005 \pm 0.003$ , where the uncertainties are statistical and systematical, respectively. The branching ratio measurement is updated, exploiting the full Run 1 data sample, corresponding to  $3 \text{ fb}^{-1}$  of integrated luminosity.

## 20 2 Formalism

### 21 2.1 Decay rates and CP-observables

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

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

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

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

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

26 The full time-dependent amplitude pdf is given by:

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

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

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

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

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

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

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

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

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

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

<sup>32</sup> In the limit of only one contributing resonance  $\kappa \rightarrow 1$ .

<sup>33</sup>

## <sup>34</sup> 2.2 Amplitude model

<sup>35</sup> The differential decay rate of a  $B_s$  meson with mass,  $m_{B_s}$ , decaying into four pseudoscalar  
<sup>36</sup> particles with four-momenta  $p_i = (E_i, \vec{p}_i)$  ( $i = 1, 2, 3, 4$ ) is given by

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

<sup>37</sup> where the transition amplitude  $A(\mathbf{x})$ , describes the dynamics of the interaction,  $d\Phi_4$   
<sup>38</sup> is the four-body phase space element [5], and  $\mathbf{x}$  represents a unique set of kinematic  
<sup>39</sup> conditions within the phase space of the decay. Each final state particle contributes three  
<sup>40</sup> observables, manifesting in their three-momentum, summing up to twelve observables in  
<sup>41</sup> total. Four of them are redundant due to four-momentum conservation and the overall  
<sup>42</sup> orientation of the system can be integrated out. The remaining five independent degrees  
<sup>43</sup> of freedom unambiguously determine the kinematics of the decay. Convenient choices  
<sup>44</sup> for the kinematic observables include the invariant mass combinations of the final state  
<sup>45</sup> particles

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

<sup>46</sup> or acoplanarity and helicity angles. It is however important to take into account that,  
<sup>47</sup> while  $m_{12}^2, m_{23}^2$  are sufficient to fully describe a three-body decay, the obvious extension  
<sup>48</sup> to four-body decays with  $m_{ij}^2, m_{ijk}^2$  requires additional care, as these variables alone are  
<sup>49</sup> insufficient to describe the parity-odd moments possible in four-body kinematics.

<sup>50</sup> In practice, we do not need to choose a particular five-dimensional basis, but use the  
<sup>51</sup> full four-vectors of the decay in our analysis. The dimensionality is handled by the phase  
<sup>52</sup> space element which can be written in terms of any set of five independent kinematic  
<sup>53</sup> observables,  $\mathbf{x} = (x_1, \dots, x_5)$ , as

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

<sup>54</sup> 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,  
<sup>55</sup> the four-body phase space density function is not flat in the usual kinematic variables.  
<sup>56</sup> Therefore, an analytic expression for  $\phi_4$  is taken from Ref. [6].

<sup>57</sup> The total amplitude for the  $B_s \rightarrow h_1 h_2 h_3 h_4$  decay is given by the coherent sum  
<sup>58</sup> over all intermediate state amplitudes  $A_i(\mathbf{x})$ , each weighted by a complex coefficient  
<sup>59</sup>  $a_i = |a_i| e^{i\phi_i}$  to be measured from data,

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

<sup>60</sup> To construct  $A_i(\mathbf{x})$ , the isobar approach is used, which assumes that the decay process  
<sup>61</sup> can be factorized into subsequent two-body decay amplitudes [7–9]. This gives rise to  
<sup>62</sup> two different decay topologies; quasi two-body decays  $B_s \rightarrow (R_1 \rightarrow h_1 h_2) (R_2 \rightarrow h_3 h_4)$

63 or cascade decays  $B_s \rightarrow h_1$  [ $R_1 \rightarrow h_2$  ( $R_2 \rightarrow h_3 h_4$ )]. In either case, the intermediate state  
 64 amplitude is parameterized as a product of form factors  $B_L$ , included for each vertex  
 65 of the decay tree, Breit-Wigner propagators  $T_R$ , included for each resonance  $R$ , and an  
 66 overall angular distribution represented by a spin factor  $S$ ,

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

### 67 2.2.1 Form Factors and Resonance Lineshapes

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

### 119 2.2.2 Spin Densities

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

127 A spin- $S$  particle with four-momentum  $p$ , and spin projection  $\lambda$ , is represented by the  
 128 polarization tensor  $\epsilon_{(S)}(p, \lambda)$ , which is symmetric, traceless and orthogonal to  $p$ . These  
 129 so-called Rarita-Schwinger conditions reduce the a priori  $4^S$  elements of the rank- $S$  tensor

<sup>130</sup> to  $2S + 1$  independent elements in accordance with the number of degrees of freedom of a  
<sup>131</sup> spin- $S$  state [20, 24].

<sup>132</sup> The spin projection operator  $P_{(S)}^{\mu_1 \dots \mu_S \nu_1 \dots \nu_S}(p_R)$ , for a resonance  $R$ , with spin  $S =$   
<sup>133</sup>  $\{0, 1, 2\}$ , and four-momentum  $p_R$ , is given by [23]:

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

<sup>134</sup> where  $g^{\mu\nu}$  is the Minkowski metric. Contracted with an arbitrary tensor, the projection  
<sup>135</sup> operator selects the part of the tensor which satisfies the Rarita-Schwinger conditions.

<sup>136</sup> For a decay process  $R \rightarrow AB$ , with relative orbital angular momentum  $L$ , between  
<sup>137</sup> particle  $A$  and  $B$ , the angular momentum tensor is obtained by projecting the rank- $L$   
<sup>138</sup> 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- $L$   
<sup>139</sup> subspace,

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

<sup>140</sup> Their  $|\vec{q}_R|^L$  dependence accounts for the influence of the centrifugal barrier on the transition  
<sup>141</sup> amplitudes. For the sake of brevity, the following notation is introduced,

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

<sup>142</sup> Following the isobar approach, a four-body decay amplitude is described as a product  
<sup>143</sup> of two-body decay amplitudes. Each sequential two-body decay  $R \rightarrow A B$ , with relative  
<sup>144</sup> orbital angular momentum  $L_{AB}$ , and total intrinsic spin  $S_{AB}$ , contributes a term to the  
<sup>145</sup> overall spin factor given by

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

<sup>146</sup> where

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

<sup>147</sup> Here, a polarization vector is assigned to the decaying particle and the complex conjugate  
<sup>148</sup> vectors for each decay product. The spin and orbital angular momentum couplings are  
<sup>149</sup> described by the tensors  $P_{(S_{AB})}(R)$  and  $L_{(L_{AB})}(R)$ , respectively. Firstly, the two spins  $S_A$   
<sup>150</sup> and  $S_B$ , are coupled to a total spin- $S_{AB}$  state,  $\Phi(\mathbf{x}|S_{AB})$ , by projecting the corresponding  
<sup>151</sup> polarization vectors onto the spin- $S_{AB}$  subspace transverse to the momentum of the  
<sup>152</sup> decaying particle. Afterwards, the spin and orbital angular momentum tensors are  
<sup>153</sup> properly contracted with the polarization vector of the decaying particle to give a Lorentz  
<sup>154</sup> scalar. This requires in some cases to include the tensor  $\varepsilon_{\alpha\beta\gamma\delta} p_R^\delta$  via

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

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

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

160 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  
 161 pseudoscalar initial/final states are involved.

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

## 164 2.3 Validation

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

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

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

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

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

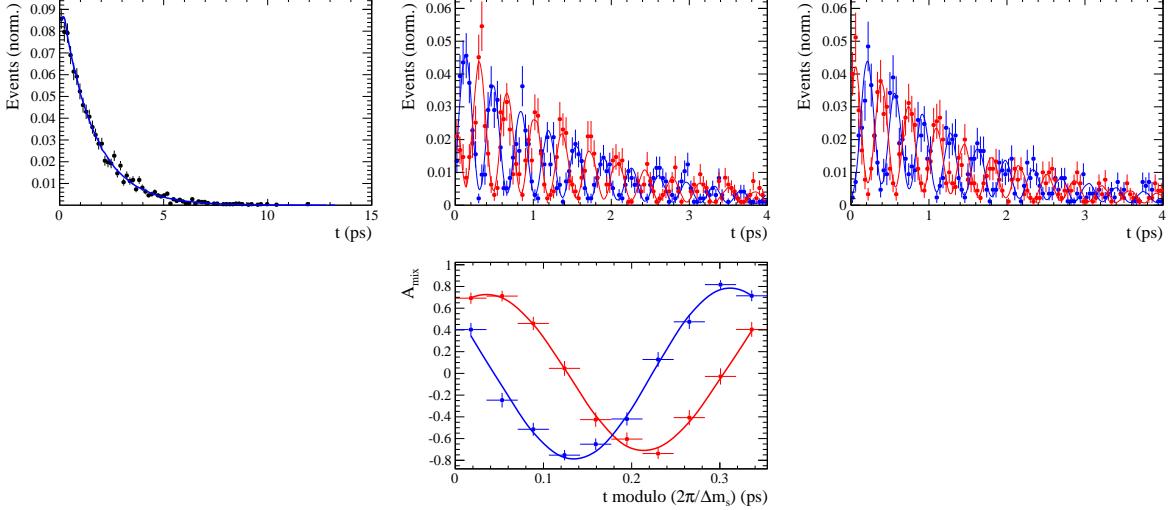


Figure 2.1: Time distribution of  $B_s \rightarrow D_s K \pi\pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

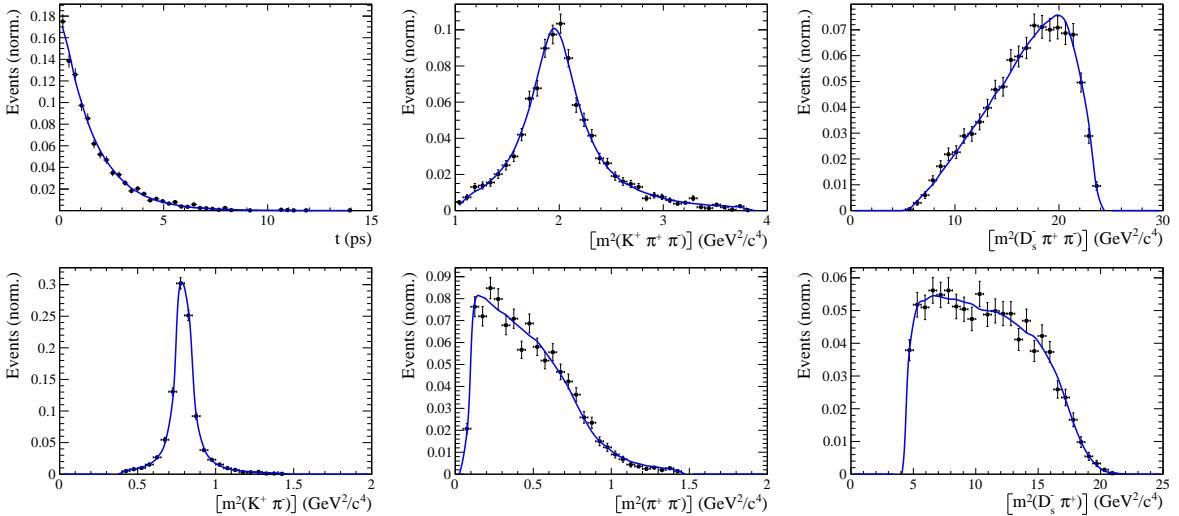


Figure 2.2: Time and invariant mass distributions of  $B_s \rightarrow D_s K \pi\pi$  events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

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

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

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

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

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

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

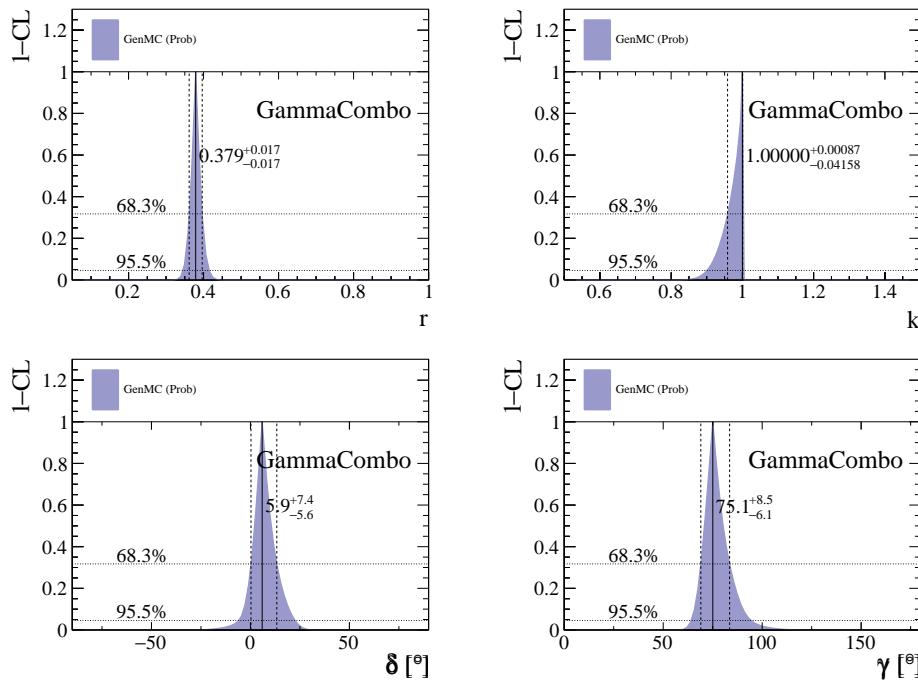


Figure 2.3: The 1-CL contours for the physical observable  $r, \kappa, \delta, \gamma$  obtained with the phasespace integrated fit to the EVTGEN toy sample.

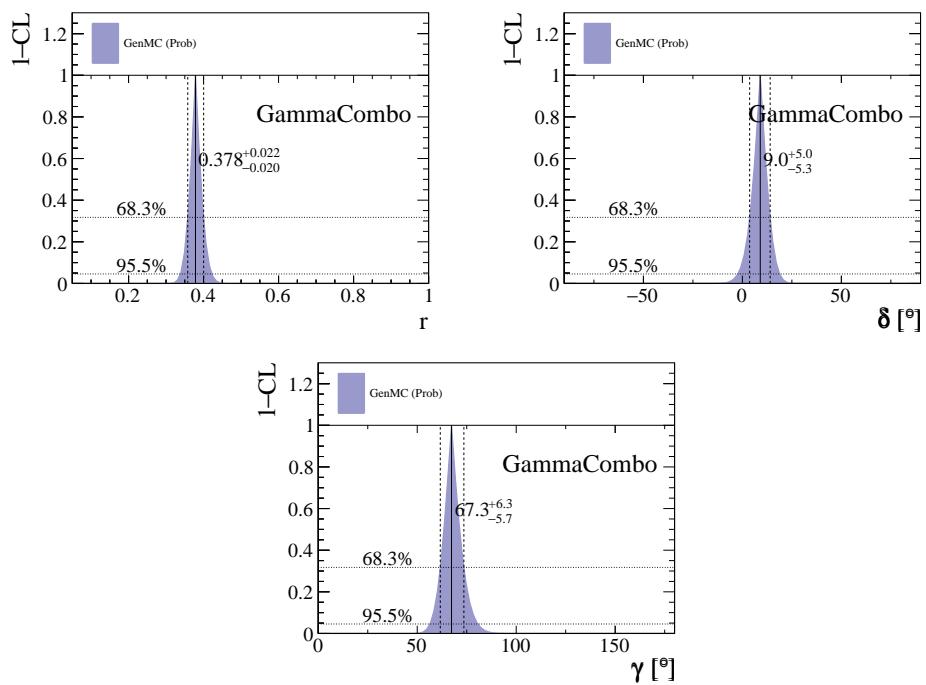


Figure 2.4: The 1-CL contours for the physical observable  $r, \delta, \gamma$  obtained with the time-dependent amplitude fit fit to the **EVTGEN** toy sample.

188 **3 Selection**

189 For the presented analysis, we reconstruct the  $B_s^0 \rightarrow D_s K\pi\pi$  decay through two different  
190 final states of the  $D_s$  meson,  $D_s \rightarrow KK\pi$  and  $D_s \rightarrow \pi\pi\pi$ . Of those two final states  
191  $D_s \rightarrow KK\pi$  is the most prominent one, while  $\mathcal{BR}(D_s \rightarrow \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$   
192 holds for the other one.

193 A two-fold approach is used to isolate the  $B_s^0 \rightarrow D_s K\pi\pi$  candidates from data passing  
194 the stripping line. First, further one-dimensional cuts are applied to reduce the level of  
195 combinatorial background and to veto some specific physical background. This stage is  
196 specific to the respective final state in which the  $D_s$  meson is reconstructed, since different  
197 physical backgrounds, depending on the respective final state, have to be taken into  
198 account. After that, a multivariate classifier is trained which combines the information  
199 of several input variables, including their correlation, into one powerful discriminator  
200 between signal and combinatorial background. For this stage, all possible  $D_s$  final states  
201 are treated equally.

202 **3.1 Cut-based selection**

203 In order to minimize the contribution of combinatorial background to our samples, we  
204 apply the following cuts to the b hadron:

- 205     • DIRA > 0.99994
- 206     • min IP  $\chi^2 < 20$  to any PV,
- 207     • FD  $\chi^2 > 100$  to any PV,
- 208     • Vertex  $\chi^2/\text{nDoF} < 8$ ,
- 209     •  $(Z_{D_s} - Z_{B_s^0}) > 0$ , where  $Z_M$  is the z-component of the position  $\vec{x}$  of the decay vertex  
210       for the  $B_s^0/D_s$  meson.

211 Additionally, we veto various physical backgrounds, which have either the same final  
212 state as our signal decay, or can contribute via a single misidentification of  $K \rightarrow \pi$  or  
213  $K \rightarrow p$ . In the following, the vetoes are ordered by the reconstructed  $D_s$  final state they  
214 apply to:

215 1. All:

- 216     (a)  $B_s^0 \rightarrow D_s^+ D_s^- : |M(K\pi\pi) - m_{D_s}| > 20 \text{ MeV}/c^2$ .
- 217     (b)  $B_s^0 \rightarrow D_s^- K^+ K^- \pi^+ : \text{possible with single missID of } K^- \rightarrow \pi^-$ , rejected by  
218       requiring  $\pi^-$  to fulfill  $\text{DLL}_{K\pi} < 5$ .

219 2.  $D_s \rightarrow KK\pi$

- 220     (a)  $B^0 \rightarrow D^+(\rightarrow K^+\pi^-\pi^+) K\pi\pi : \text{possible with single missID of } \pi^+ \rightarrow K^+$ , vetoed  
221       by changing particle hypothesis and recompute  $|M(K^+\pi^-\pi^+) - m_{D_p}| > 30$   
222       MeV/ $c^2$ , or the  $K^+$  has to fulfill  $\text{DLL}_{K\pi} > 10$ .

- 223 (b)  $\Lambda_b^0 \rightarrow \Lambda_c^+ (\rightarrow p K^- \pi^+) K \pi \pi$  : possible with single missID of  $p \rightarrow K^+$ , vetoed by  
 224 changing particle hypothesis and recompute  $M(p K^- \pi^+) - m_{\Lambda_c^+} > 30 \text{ MeV}/c^2$ ,  
 225 or the  $K^+$  has to fulfill  $(\text{DLL}_{K\pi} - \text{DLL}_{p\pi}) > 5$ .  
 226 (c)  $D^0 \rightarrow KK$  :  $D^0$  combined with a random  $\pi$  can fake a  $D_s \rightarrow KK\pi$  decay and  
 227 be a background to our signal, vetoed by requiring  $M(KK) < 1840 \text{ MeV}/c^2$ .

228 3.  $D_s \rightarrow \pi\pi\pi$

- 229 (a)  $D^0 \rightarrow \pi\pi$  : combined with a random  $\pi$  can fake a  $D_s \rightarrow \pi\pi\pi$  decay and be a  
 230 background to our signal, vetoed by requiring both possible combinations to  
 231 have  $M(\pi\pi) < 1700 \text{ MeV}/c^2$ .

232 The most prominent final state used in this analysis is  $B_s^0 \rightarrow D_s (\rightarrow KK\pi) K \pi \pi$ , where  
 233 the  $D_s$  decay can either proceed via the narrow  $\phi$  resonance, the broader  $K^{*0}$  resonance, or  
 234 non resonant. Depending on the decay process being resonant or not, we apply additional  
 235 PID requirements on this final state:

- 236 • resonant case:
  - 237 –  $D_s^+ \rightarrow \phi \pi^+$ , with  $|M(K^+ K^-) - m_\phi| < 20 \text{ MeV}/c^2$  : no additional requirements,  
   since  $\phi$  is narrow and almost pure  $K^+ K^-$ .
  - 239 –  $D_s^+ \rightarrow \bar{K}^{*0} K^+$ , with  $|M(K^- \pi^+) - m_{K^{*0}}| < 75 \text{ MeV}/c^2$  :  $\text{DLL}_{K\pi} > 0$  for kaons,  
   since this resonance is more than ten times broader than  $\phi$ .
- 241 • non resonant case:  $\text{DLL}_{K\pi} > 5$  for kaons, since the non resonant category has  
 242 significant charmless contributions.

243 For the  $D_s \rightarrow \pi\pi\pi$  final state, we apply global PID requirements:

- 244 •  $\text{DLL}_{K\pi} < 10$  for all pions.
- 245 •  $\text{DLL}_{p\pi} < 10$  for all pions.

246 3.2 Multivariate stage

247 We use TMVA [30] to train a multivariate discriminator, which is used to further improve  
 248 the signal to background ratio. The following variables are used for the training:

- 249 •  $\max(\text{ghostProb})$  over all tracks
- 250 •  $\text{cone}(p_T)$  asymmetry of every track, which is defined to be the difference between the  
    $p_T$  of the  $\pi/K$  and the sum of all other  $p_T$  in a cone of radius  $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}$   
    $< 1 \text{ rad}$  around the signal  $\pi/K$  track.
- 253 •  $\min(\text{IP}\chi^2)$  over the  $X_s$  daughters
- 254 •  $\max(\text{DOCA})$  over all pairs of  $X_s$  daughters
- 255 •  $\min(\text{IP}\chi^2)$  over the  $D_s$  daughters

- 256     •  $D_s$  and  $B_s^0$  DIRA  
 257     •  $D_s$  FD significance  
 258     •  $\max(\cos(D_s h_i))$ , where  $\cos(D_s h_i)$  is the cosine of the angle between the  $D_s$  and  
 259       another track i in the plane transverse to the beam  
 260     •  $B_s^0$  IP $\chi^2$ , FD $\chi^2$  and Vertex  $\chi^2$

261       Various classifiers were investigated in order to select the best performing discriminator.  
 262       Consequently, a boosted decision tree with gradient boost (BDTG) is chosen as nominal  
 263       classifier. We use truth-matched MC as signal input. Simulated signal candidates are  
 264       required to pass the same trigger, stripping and preselection requirements, that were  
 265       used to select the data samples. For the background we use events from the high mass  
 266       sideband ( $m_{B_s^0 \text{candidate}} > 5600 \text{ MeV}/c^2$ ) of our data samples. As shown in Fig. 3.1,  
 267       this mass region is sufficiently far away from signal structures and is expected to be  
 268       dominantly composed of combinatorial background. For completeness, the mass distribu-  
 269       tion of preselected  $D_s \rightarrow hh$  candidates (where  $h = \pi$  or  $h = K$ ) is also shown in Fig. 3.1.

270

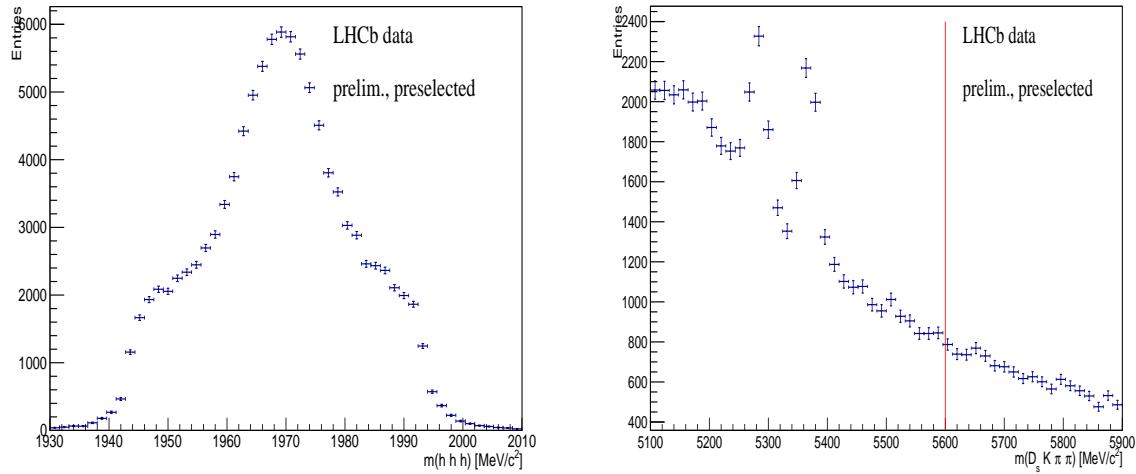


Figure 3.1: Invariant mass distribution of preselected (left)  $D_s \rightarrow hhh$  and (right)  $B_s^0 \rightarrow D_s K\pi\pi$  candidates. For the  $B_s^0 \rightarrow D_s K\pi\pi$  candidates, the region right from the red colored line with  $m_{B_s^0 \text{candidate}} > 5600 \text{ MeV}/c^2$  is used as background input for the boosted decision tree.

271       The distributions of the input variables for signal and background and the BDTG  
 272       output distribution are shown in the appendix.

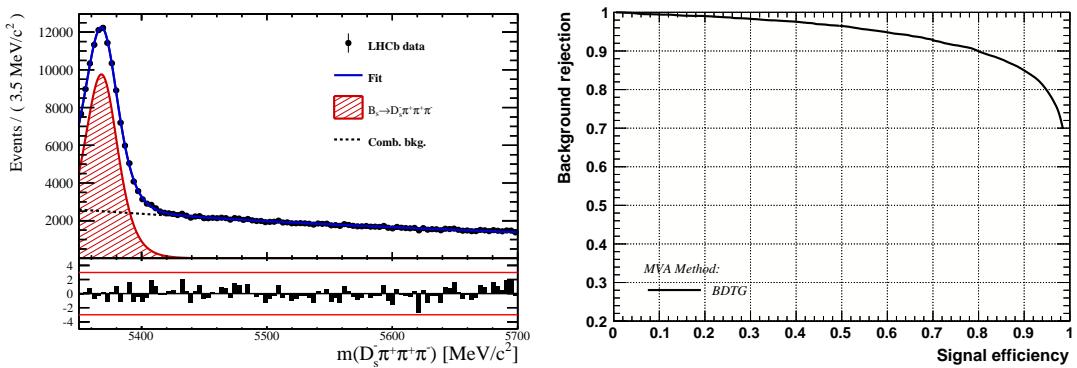


Figure 3.2

273 **4 Fits to invariant mass distributions of signal and**  
 274 **normalization channel**

275 In order to properly model the invariant mass distribution of  $B_s^0 \rightarrow D_s K\pi\pi$  and  $B_s^0 \rightarrow$   
 276  $D_s \pi\pi\pi$  candidates, the expected signal shape, as well as the expected shape for the  
 277 combinatorial and physical background has to be known. This model can then be used to  
 278 fit the distributions and obtain signal sWeights [31], which are employed to suppress the  
 279 residual background that is still left in the sample, for the time-dependent amplitude fit.

280 **4.1 Signal models for  $m(D_s \pi\pi\pi)$  and  $m(D_s K\pi\pi)$**

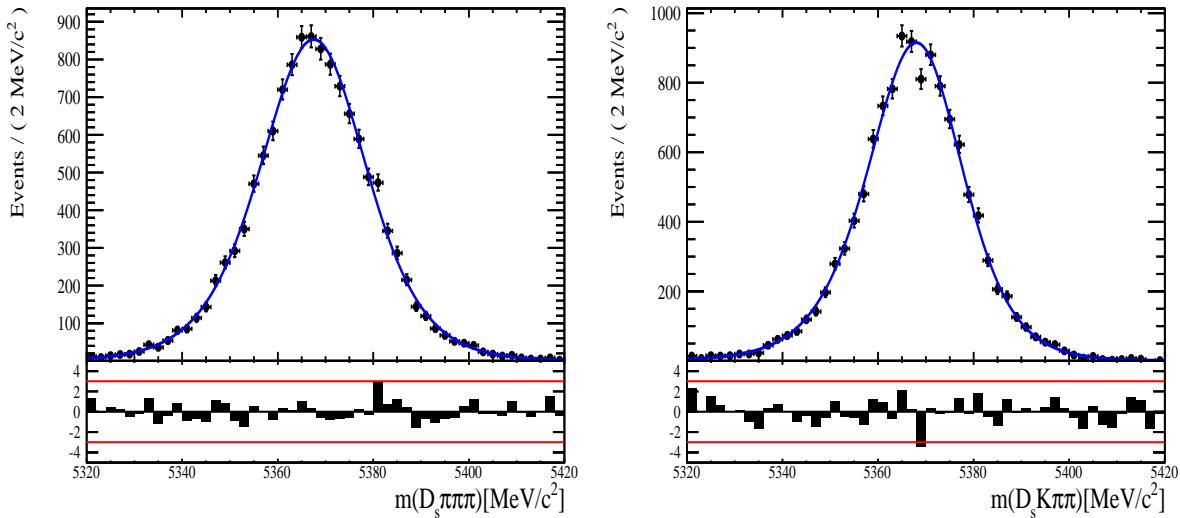


Figure 4.1: Invariant mass distributions of simulated (left)  $B_s^0 \rightarrow D_s \pi\pi\pi$  and (right)  $B_s^0 \rightarrow D_s K\pi\pi$  events. A fit of a RooJohnsonSU function to each distribution is overlaid.

281 The mass distribution of  $B_s^0 \rightarrow D_s K\pi\pi$  signals is modeled using a Johnson SU  
 282 function [32], which is a gaussian function with a Landau-like tail on one side,

$$J(m_{B_s^0}; \mu, \sigma, \gamma, \delta) = \frac{\delta}{\sigma 2\pi \sqrt{1 + (\frac{m_{B_s^0} - \mu}{\sigma})^2}} \exp\left(-\frac{1}{2}[\gamma + \delta \operatorname{Argsh}(\frac{m_{B_s^0} - \mu}{\sigma})]^2\right). \quad (4.1)$$

283 The sign of  $\gamma$  in Eq. 4.1 determines whether the tail is located at lower ( $\gamma > 0$ )  
 284 or higher ( $\gamma < 0$ ) invariant mass values than the mean  $\mu$  of the gaussian function and  
 285  $\delta$  describes the (a)symmetry of the fitted distribution. Higher values of  $\delta$  result in a  
 286 more symmetric, gaussian-like function. Another Johnson SU function function is used  
 287 to account for the contribution of the  $B^0 \rightarrow D_s K\pi\pi$  decay, which is also present in  
 288 the  $m(D_s K\pi\pi)$  spectrum. The width, as well as the tail parameters are fixed to values  
 289 obtained by a fit to the invariant mass distribution of simulated events shown in Fig 4.1.  
 290 A linear scaling factor for the mean  $\mu$  and width  $\sigma$  is floated in the fit to account for  
 291 possible differences between the simulation and real data.

292 The same approach is used to describe the invariant mass distribution of  $B_s^0 \rightarrow D_s \pi\pi\pi$

293 candidates. A Johnson SU function is used to model the signal, the parameters are  
294 determined by a fit to the invariant mass of simulated  $B_s^0 \rightarrow D_s\pi\pi\pi$  decays, shown in  
295 Fig 4.1. A scale factor for the width and the mean is floated to account for differences  
296 between data and MC.

## 297 4.2 Background models for $m(D_s\pi\pi\pi)$

298 Different background sources arise in the invariant mass spectrum of candidates in the  
299 normalization mode.

300 The following backgrounds have to be accounted for:

- 301 • Combinatorial background: This contribution arises from either a real  $D_s$ , which is  
302 paired with random tracks to form the  $B_s^0$  candidates, or via real  $X_d$ 's, which are  
303 combined with three tracks that fake a  $D_s$  candidate to form a fake  $B_s^0$ .
- 304 • Partially reconstructed  $B^0/B_s^0 \rightarrow D_s^*\pi\pi\pi$  decays, with  $D_s^* \rightarrow D_s\gamma$  or  $D_s^* \rightarrow D_s\pi^0$ ,  
305 where the  $\gamma/\pi^0$  is not reconstructed in the decay chain.

306 In both cases of combinatorial background, the distribution in the invariant mass of  
307  $B_s^0$  candidates is expected to be smooth and decrease with higher masses. Therefore, one  
308 exponential function is used to model these contributions.

309 The shape of the  $B_s^0 \rightarrow D_s^*\pi\pi\pi$  contribution is expected to be peaking in the  $m(D_s\pi\pi\pi)$   
310 spectrum, with large tails due to the missing momentum, which is carried away by the  $\pi^0$   
311 or  $\gamma$ . The pion or photon from  $D_s^* \rightarrow D_s(\gamma/\pi^0)$  is excluded from the reconstruction. We  
312 model the shape of this contribution using the sum of three bifurcated Gaussian functions.  
313 The shape parameters, as well as the yield of this contribution, are directly determined  
314 on data from a fit to the  $m(D_s\pi\pi\pi)$  invariant mass distribution.

## 315 4.3 Background models for $m(D_sK\pi\pi)$

316 For the signal channel, the following background sources have to be considered:

- 317 • Combinatorial background: same contributions as discussed in Sec. 4.2.
- 318 • Partially reconstructed  $B_s^0 \rightarrow D_s^*K\pi\pi$  decays, with  $D_s^* \rightarrow D_s\gamma$  or  $D_s^* \rightarrow D_s\pi^0$ ,  
319 where the  $\gamma/\pi^0$  is not reconstructed in the decay chain.
- 320 • Partially reconstructed  $B^0 \rightarrow D_s^*K\pi\pi$  decays, with  $D_s^* \rightarrow D_s\gamma$  or  $D_s^* \rightarrow D_s\pi^0$ ,  
321 where the  $\gamma/\pi^0$  is not reconstructed in the decay chain.
- 322 • Misidentified  $B_s^0 \rightarrow D_s\pi\pi\pi$  decays, where one of the pions is wrongly identified as a  
323 kaon  $\pi \rightarrow K$ .
- 324 • Misidentified, partially reconstructed  $B_s^0 \rightarrow D_s^*\pi\pi\pi$  decays, where one of the pions  
325 is wrongly identified as a kaon  $\pi \rightarrow K$  and the  $\gamma/\pi^0$  from  $D_s^* \rightarrow D_s\gamma/\pi^0$  is not  
326 reconstructed.

327 The combinatorial background is expected to be non-peaking in the spectrum of the  
328 invariant mass of  $B_s^0 \rightarrow D_sK\pi\pi$  candidates. An exponential function is used to model  
329 this contribution.

330 The shape of the partially reconstructed background without misID is taken from our  
 331 normalization channel, where it can be directly fitted by the sum of three bifurcated  
 332 Gaussian functions as described above. In the signal mass fit, all shape parameters for  
 333 the  $B_s^0 \rightarrow D_s^* K\pi\pi$  background are fixed to the input values from our normalization fit.

334 For the contribution of the  $B^0 \rightarrow D_s^* K\pi\pi$  background, the same shape is used but  
 335 the means  $\mu_i$  of the bifurcated gaussians are shifted down by  $m_{B_s^0} - m_{B^0}$  [?]. The yields  
 336 of both contributions are directly determined in the nominal fit.

337 To determine the shape of misidentified  $B_s^0 \rightarrow D_s \pi\pi\pi$  candidates in the  $m(D_s K\pi\pi)$   
 338 spectrum, we take a truth-matched signal MC sample of our normalization channel. We  
 339 then use the PIDCalib package to determine the  $\pi \rightarrow K$  fake rate. For every candidate  
 340 in our MC sample, a (momentum)  $p$  and (pseudorapidity)  $\eta$ -dependent event weight is  
 341 computed and assigned. We flip the particle hypothesis from pion to kaon for the  $\pi$  with  
 342 the biggest miss-ID weight for each event and recompute the invariant  $B_s^0$  mass. This  
 343 distribution is then modeled using two Crystal Ball functions. The distribution and the  
 344 fit are shown in Fig. 4.2(left).

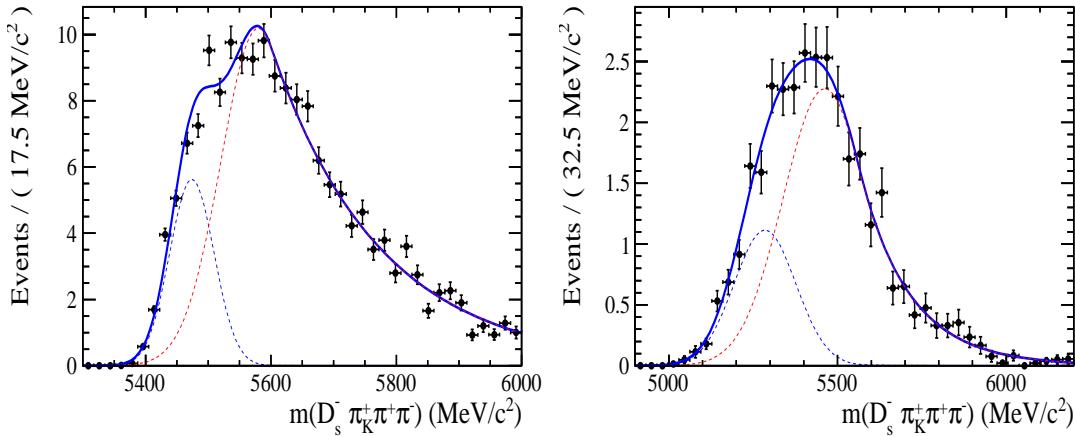


Figure 4.2: Invariant mass distribution of (left) simulated  $B_s^0 \rightarrow D_s \pi\pi\pi$  events, where one of the  $\pi$ 's is reconstructed as a  $K$  and the misID probability for each event is taken into account. The corresponding distribution for simulated  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  events, where the  $\gamma/\pi^0$  from the  $D_s^*$  is excluded from reconstruction, is shown on the right. The solid, black curve on each plot corresponds to the fit consisting of two Crystal Ball functions.

345 The expected yield of misidentified  $B_s^0 \rightarrow D_s \pi\pi\pi$  candidates in the  $m(D_s K\pi\pi)$  spec-  
 346 trum is computed by multiplying the fake probability of  $\propto 3.2\%$ , which is derived from  
 347 PIDCalib, by the yield of  $B_s^0 \rightarrow D_s \pi\pi\pi$  signal candidates, determined in the nominal  
 348 mass fit of our normalization channel.

349 In the same way as mentioned above, we can determine the rate of misidentified, partially  
 350 reconstructed  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  decays in our sample of  $B_s^0 \rightarrow D_s K\pi\pi$  decays using PIDCalib  
 351 and a MC sample of  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  events. The invariant mass distribution we obtain  
 352 when we exclude the  $\gamma/\pi^0$ , flip the the particle hypothesis  $\pi \rightarrow K$  and apply the event  
 353 weights given by the fake rate, is shown in Fig. 4.2 (right). The fit of two Crystal Ball  
 354 functions to this distribution is overlaid. The yield of this contribution is determined  
 355 from the yield of  $B_s^0 \rightarrow D_s^* \pi\pi\pi$  candidates in the nominal mass fit of our normalization  
 356 channel, multiplied by the misID probability of  $\propto 3.6\%$ .

357 **4.4 Fit to  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates**

358 An unbinned maximum likelihood fit is performed simultaneously to the invariant mass  
 359 distribution of  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates. As discussed in Sec. ??, the fit is given  
 360 as a Johnson SU signal model for the  $B_s^0$  and  $B^0$  signal, the sum of three bifurcated  
 361 Gaussian functions to model the partially reconstructed  $B_s^0 \rightarrow D_s^*\pi\pi\pi$  background and  
 362 an Exponential function to account for combinatorial background. The invariant mass  
 363 distribution and the fit is shown in Fig. 4.3. All simultaneously performed fits to the  
 364  $m(D_s\pi\pi\pi)$  distribution, ordered by the respective  $D_s$  final state, can be found in the  
 365 Appendix ???. The obtained yields are summarized in Table 4.1.

366 **4.5 Fit to  $B_s^0 \rightarrow D_sK\pi\pi$  candidates**

367 The shape of the invariant mass distribution of  $B_s^0 \rightarrow D_sK\pi\pi$  candidates is described by  
 368 Johnson SU functions for the  $B^0$  and  $B_s^0$  signal, two sums of three bifurcated Gaussians  
 369 for the  $B_s^0/B^0 \rightarrow D_s^*K\pi\pi$  partially reconstructed background contributions and two  
 370 sums of double Crystal Ball functions for the single misID  $B_s^0 \rightarrow D_s\pi\pi\pi$  and the partially  
 371 reconstructed, misidentified  $B_s^0 \rightarrow D_s^*\pi\pi\pi$  decays. A simultaneous unbinned maximum  
 372 likelihood fit is performed and the result is shown in Fig. 4.3. All simultaneously performed  
 373 fits to the  $m(D_sK\pi\pi)$  distribution, ordered by the respective  $D_s$  final state, can be found  
 374 in the Appendix ???. The obtained yields are summarized in Table 4.1.

375 **4.6 Extraction of signal weights**

376 The sPlot technique [31] is used to extract signal weights from the fits to the invariant  
 377 mass distributions of our signal and normalization channel. This statistical tool assigns  
 378 a weight to every event, according to its position in the respective mass distribution,  
 379 given the fitted signal and background models. The weights can then be used to suppress  
 380 the background components in every other observable distribution of interest. Figure ??  
 381 shows the distribution of weights across the invariant mass spectra of  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  
 382  $B_s^0 \rightarrow D_sK\pi\pi$  candidates.

fit component	yield 2011	yield 2012	yield 2015	yield 2016
$m(D_sK\pi\pi)$				
$B_s^0 \rightarrow D_sK\pi\pi$	$392 \pm 25$	$860 \pm 38$	$309 \pm 21$	$1984 \pm 55$
$B^0 \rightarrow D_sK\pi\pi$	$276 \pm 26$	$692 \pm 41$	$261 \pm 23$	$1385 \pm 58$
$B^0/B_s^0 \rightarrow D_s^*K\pi\pi$	$7 \pm 25$	$171 \pm 75$	$114 \pm 25$	$893 \pm 84$
$B_s^0 \rightarrow D_s^{(*)}\pi\pi\pi$	$63 \pm 0$	$158 \pm 0$	$53 \pm 0$	$314 \pm 0$
combinatorial	$1482 \pm 53$	$2884 \pm 100$	$605 \pm 43$	$4261 \pm 133$
$m(D_s\pi\pi\pi)$				
$B_s^0 \rightarrow D_s\pi\pi\pi$	$9183 \pm 105$	$22083 \pm 166$	$7574 \pm 95$	$43773 \pm 245$
$B^0 \rightarrow D_s\pi\pi\pi$	$289 \pm 58$	$716 \pm 95$	$229 \pm 54$	$968 \pm 147$
$B_s^0 \rightarrow D_s^*\pi\pi\pi$	$3640 \pm 130$	$9086 \pm 232$	$3047 \pm 110$	$17827 \pm 421$
combinatorial	$4991 \pm 154$	$11127 \pm 271$	$3728 \pm 126$	$24589 \pm 500$

Table 4.1: Summary of yields obtained from the fits to Run1 and Run2 data.

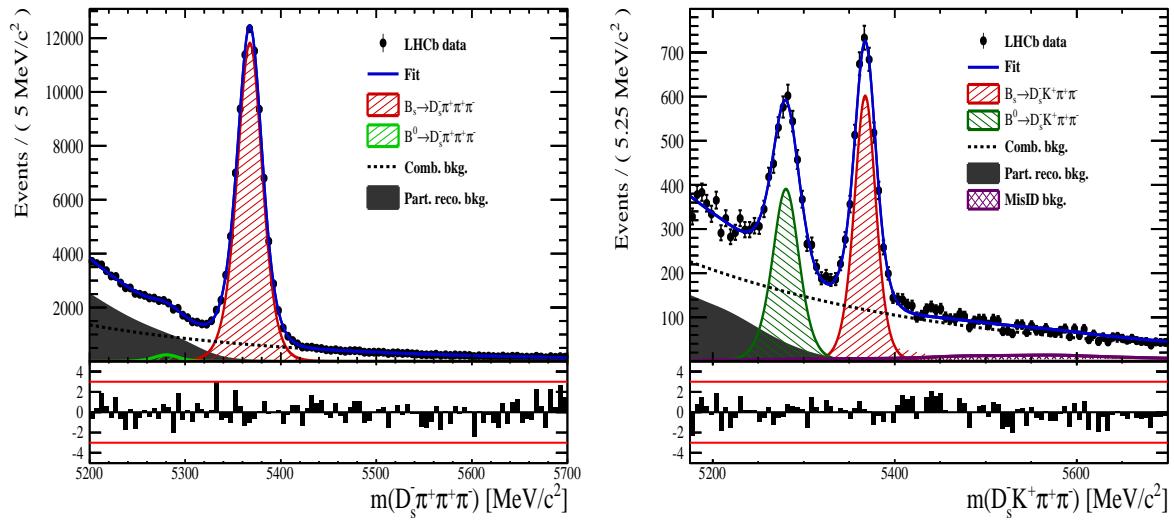


Figure 4.3: Invariant mass distribution of (left)  $B_s^0 \rightarrow D_s \pi \pi \pi$  and (right)  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for Run1 and Run2 data. The respective fit described in the text is overlaid.

## 383 5 Flavour Tagging

384 To successfully perform a time- and amplitude-dependent measurement of  $\gamma$ , the identifi-  
 385 cation of the initial state flavour of the  $B_s^0$  meson is crucial. In the presented analysis,  
 386 a number of flavour tagging algorithms are used that either determine the flavour of  
 387 the non-signal b-hadron produced in the event (opposite site, OS), or they use particles  
 388 produced in the fragmentation of the signal candidate  $B_s^0/\bar{B}_s^0$  (same side, SS).  
 389 For the same side, the algorithm searching for the charge of an additional kaon that accom-  
 390 panies the fragmentation of the signal candidate is used (SS-nnetKaon). For the opposite  
 391 site, four different taggers are chosen: The Two algorithms that use the charge of an  
 392 electron or a muon from semileptonic B decays (OS-  $e,\mu$ ), the tagger that uses the charge  
 393 of a kaon from a  $b \rightarrow c \rightarrow s$  decay chain (OS-nnetKaon) and the algorithm that determines  
 394 the  $B_s^0/\bar{B}_s^0$  candidate flavour from the charge of a secondary vertex, reconstructed from  
 395 the OS b decay product (OS-VtxCharge). All four taggers are then combined into a signal  
 396 OS tagger.  
 397 Every single tagging algorithm is prone to misidentify the signal candidate at a certain  
 398 mistag rate  $\omega = (\text{wrongtags})/(\text{alltags})$ . This might be caused by particle misidentifica-  
 399 tion, flavour oscillation of the neutral opposite site B-meson or by tracks that are wrongly  
 400 picked up from the underlying event. For every signal  $B_s^0/\bar{B}_s^0$  candidate, each tagging  
 401 algorithm predicts a mistag probability  $\eta$ , which is calculated using a combination of  
 402 inputs such as the kinematics of the tagging particles. The inputs are then combined  
 403 to a predicted mistag using neural networks. These are trained on simulated samples  
 404 of  $B_s^0 \rightarrow D_s^- \pi^+$  (SS algorithm) and  $B^+ \rightarrow J/\psi K^+$  (OS algorithms) decays. For the  
 405 presented analysis, the measurable CP-violating coefficients are damped by the tagging  
 406 dilution  $D$ , that depends on the mistag rate:

$$D = 1 - 2\omega. \quad (5.1)$$

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

$$\epsilon_{eff} = \epsilon_{tag}(1 - 2\omega)^2, \quad (5.2)$$

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

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

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - \langle \eta \rangle), \quad (5.3)$$

415 where the values of  $p_0$  and  $p_1$  are determined using the  $B_s^0 \rightarrow D_s \pi \pi \pi$  normalization  
 416 mode and  $\langle \eta \rangle$  is the average estimated mistag probability  $\langle \eta \rangle = \sum_{i=1}^{N_{cand}} (\eta_i)/N_{cand}$ .  
 417 Following this model, a perfectly calibrated tagger would lead to  $\omega(\eta) = \eta$  and one would  
 418 expect  $p_1 = 1$  and  $p_0 = \langle \eta \rangle$ . Due to the different interaction cross-sections of oppositely  
 419 charged particles, the tagging calibration parameters depend on the initial state flavour of  
 420 the  $B_s^0$ . Therefore, the flavour asymmetry parameters  $\Delta p_0$ ,  $\Delta p_1$  and  $\Delta \epsilon_{tag}$  are introduced.  
 421 For this analysis, the calibrated mistag is treated as per-event variable, giving a larger

weight to events that are less likely to have an incorrect tag. This adds one additional observable to the time- and amplitude-dependent fit.  
 The tagging calibration is determined using a time-dependent fit to the full  $B_s^0 \rightarrow D_s\pi\pi\pi$  sample, where the mixing frequency  $\Delta m_s$  is fixed to the nominal PDG value [33]. The calibration procedure for the OS tagging algorithms (Sec.5.1) and the SS kaon tagger (Sec.5.2) is applied on the full Run I and 2015 and 2016 Run II  $B_s^0 \rightarrow D_s\pi\pi\pi$  data sample, which is selected following the steps described in Sec. 3. The similar selection ensures as close as possible agreement between the  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$  samples in terms of the decay kinematics, which are crucial for the flavour tagging. Section 5.3 shows the compatibility of both samples. After applying the calibration, the response of the OS and SS taggers are combined, which is shown in Sec. 5.4.

## 5.1 OS tagging calibration

The responses of the OS electron, muon, neural net kaon and the secondary vertex charge taggers are combined for the mistag calibration. Figure ?? shows the distribution of the predicted OS mistag for signal candidates from  $B_s^0 \rightarrow D_s\pi\pi\pi$ . The extracted calibration parameters and tagging asymmetries are summarized in Table 5.1 and the measured tagging power for the OS combination is  $\epsilon_{eff,OS} = 4.81\%$ .

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

Table 5.1: Calibration parameters and tagging asymmetries of the OS tagger extracted from  $B_s^0 \rightarrow D_s\pi\pi\pi$  decays.

## 5.2 SS tagging calibration

The SS neural net kaon tagger can be calibrated using the flavour-specific  $B_s^0 \rightarrow D_s\pi\pi\pi$  decay. It's development, performance and calibration is described in detail in [34]. Figure ?? shows the distribution of the predicted mistag of the neural net kaon tagger. The extracted calibration parameters and tagging asymmetries are summarized in Table 5.2 and the measured tagging power for this algorithm is  $\epsilon_{eff,SS} = 3.22\%$ .

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

Table 5.2: Calibration parameters and tagging asymmetries of the SS tagger extracted from  $B_s^0 \rightarrow D_s\pi\pi\pi$  decays.

## 5.3 Tagging performance comparison between the signal and normalization channel

To justify the usage of the tagging calibration, obtained using the  $B_s^0 \rightarrow D_s\pi\pi\pi$  sample, for our signal decay, the performance of the taggers in the two decay channels needs to be compatible. This is verified using both, simulated signal samples of both decays and

450 sweighted data, to compare the similarity of the mistag probabilities, tagging decisions  
 451 and kinematic observables that are correlated with the tagging response, on simulation  
 452 and data.

453 The distributions of the predicted mistag probability  $\eta$  for the OS combination and the  
 454 SS kaon tagger are shown in Fig. ?? (simulation) and Fig. 5.1 (data).

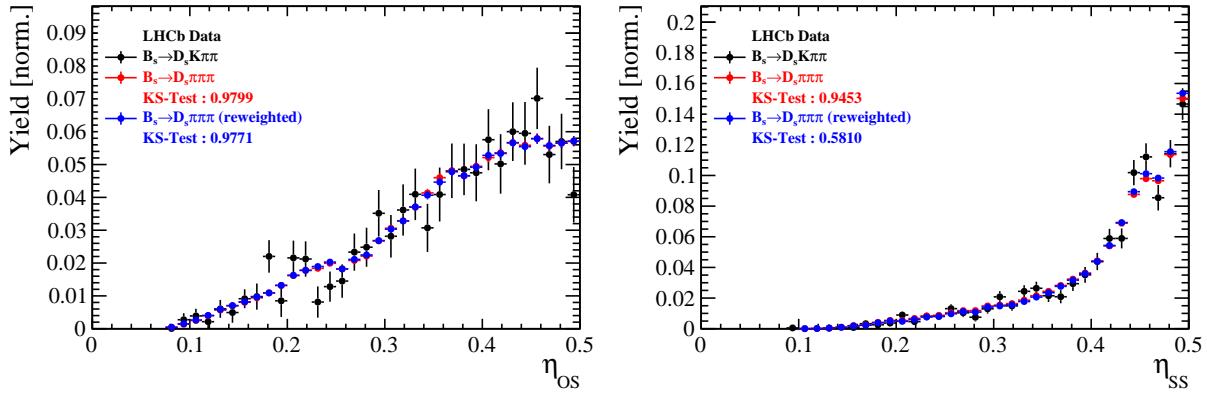


Figure 5.1: Distributions of the predicted mistag  $\eta$  for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the  $B_s^0 \rightarrow D_s K\pi\pi$  (black) and  $B_s^0 \rightarrow D_s \pi\pi\pi$  (red) data samples.

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

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

461 To justify the portability of the flavour tagging calibration obtained from  $B_s^0 \rightarrow D_s \pi\pi\pi$   
 462 to the  $B_s^0 \rightarrow D_s K\pi\pi$  channel, besides the good agreement of the distributions shown  
 463 above, the dependence of the measured mistag  $\omega$  on the predicted mistag  $\eta$  has to be  
 464 compatible in both channel. This dependence is shown in Fig. 5.2 for simulated signal  
 465 events of both channels, where good agreement is observed.

## 466 5.4 Combination of OS and SS taggers

467 In the time- and amplitude-dependent fit to  $B_s^0 \rightarrow D_s K\pi\pi$  data, the obtained tagging  
 468 responses of the OS and SS tagger will be combined after the calibration described in the  
 469 previous sections is applied. Events that aquire a mistag probability greater than 0.5 after  
 470 the calibration will have their tagging decision flipped. For events where only one of the  
 471 two taggers fired, the combination of the tagging decision is trivial. In those events where  
 472 both taggers made a decision, we use the standard combination of taggers [35] provided  
 473 by the flavour tagging group. In the nominal fit, the calibrated mistags  $\omega$  are combined  
 474 event by event for the OS and SS tager, thus adding one variable to observable to the  
 475 fit procedure. This ensures that the uncertainties of the OS and SS tagging calibration  
 476 parameters are propagated properly to the combined tagging response for each event.  
 477 The taggging performance for the combined tagger in the categories SS tagged only, OS  
 478 tagged only and SS+OS tagged, are shown in Tab. ?? for the signal and normalization

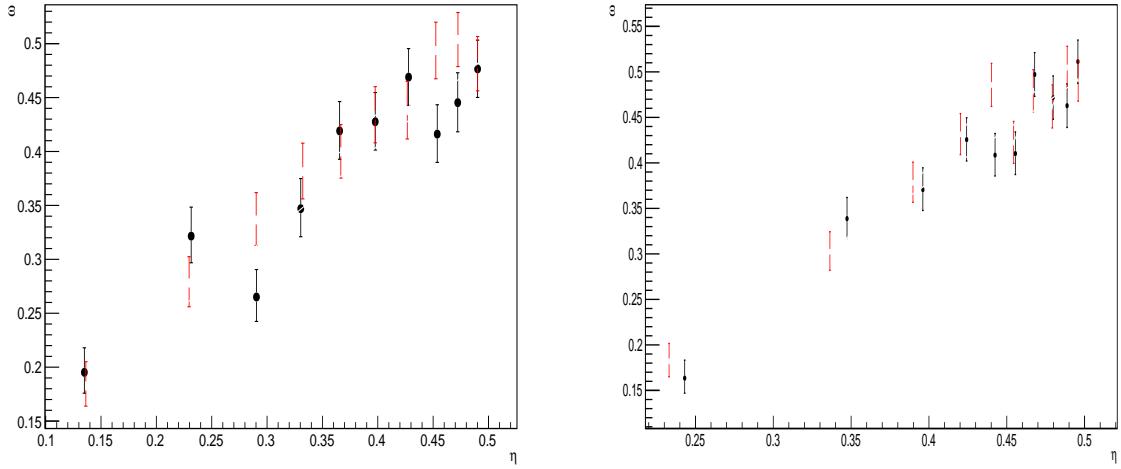


Figure 5.2: Dependence of the observed mistag  $\omega$  on the predicted mistag  $\eta$  for the (left) OS combination and the (right) SS kaon tagger, found in the simulated  $B_s^0 \rightarrow D_s K\pi\pi$  (black) and  $B_s^0 \rightarrow D_s \pi\pi\pi$  (red) signal samples.

479 channel. The distribution of the observed mistag  $\omega$  as a function of the combined mistag  
 480 probability  $\eta$  for  $B_s^0 \rightarrow D_s \pi\pi\pi$  decays is shown in Fig. 5.3.

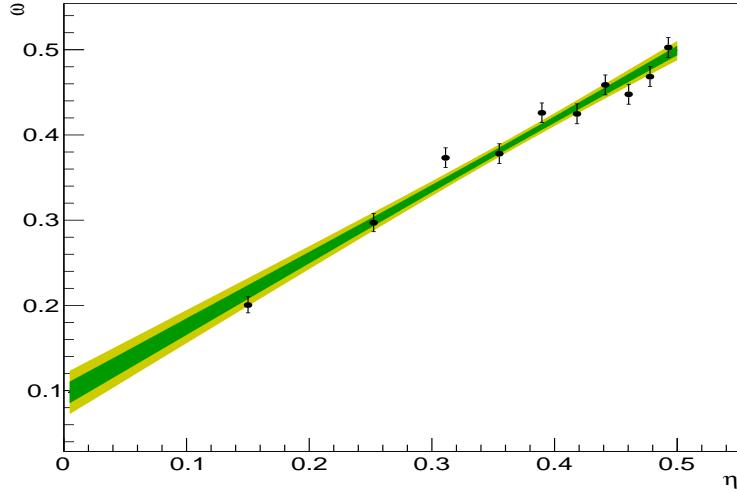


Figure 5.3: Distribution of the predicted combined mistag probability  $\eta$  versus the observed mistag  $\omega$  for  $B_s^0 \rightarrow D_s \pi\pi\pi$  signal candidates. The fit with a linear polynomial, used to determine  $p_0$  and  $p_1$  is overlaid.

$B_s^0 \rightarrow D_s\pi\pi\pi$		$\epsilon_{tag}$	$\epsilon_{eff}$
SS only		$(28.586 \pm 0.165)\%$	$(1.408 \pm 0.018(\text{stat}) \pm 0.082(\text{cal}))\%$
OS only		$(17.221 \pm 0.138)\%$	$(2.027 \pm 0.029(\text{stat}) \pm 0.100(\text{cal}))\%$
SS+OS		$(39.981 \pm 0.179)\%$	$(5.690 \pm 0.047(\text{stat}) \pm 0.196(\text{cal}))\%$
total			
$B_s^0 \rightarrow D_sK\pi\pi$		$\epsilon_{tag}$	$\epsilon_{eff}$
SS only		$(30.094 \pm 0.960)\%$	$(1.379 \pm 0.082(\text{stat}) \pm 0.085(\text{cal}))\%$
OS only		$(18.923 \pm 0.819)\%$	$(1.768 \pm 0.121(\text{stat}) \pm 0.099(\text{cal}))\%$
SS+OS		$(27.277 \pm 0.932)\%$	$(3.914 \pm 0.194(\text{stat}) \pm 0.220(\text{cal}))\%$
total			

Table 5.3: Flavour tagging performances for  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$  events which are only OS tagged, only SS tagged or tagged by both.

## 481 6 Acceptance

### 482 6.1 MC corrections

#### 483 6.1.1 Truth matching of simulated candidates

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

493 We therefore include `BKGCAT`= 60 and statistically subtract the ghost background by  
 494 using the `sPlot` technique. The `sWeights` are calculated from a fit to the reconstructed  $B_s$   
 495 mass. The signal contribution is modeled as described in Sec. 4.1 and the background  
 496 with a polynomial. The fit is performed simultaneously in two categories; the first includes  
 497 events with `BKGCAT` = 20 or 50 and the second events with `BKGCAT` = 60. To account  
 498 for the different mass resolution we use a different  $\sigma$  for each category, while the mean  
 499 and the tail parameters are shared between them. The background component is only  
 500 included for the second category.

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

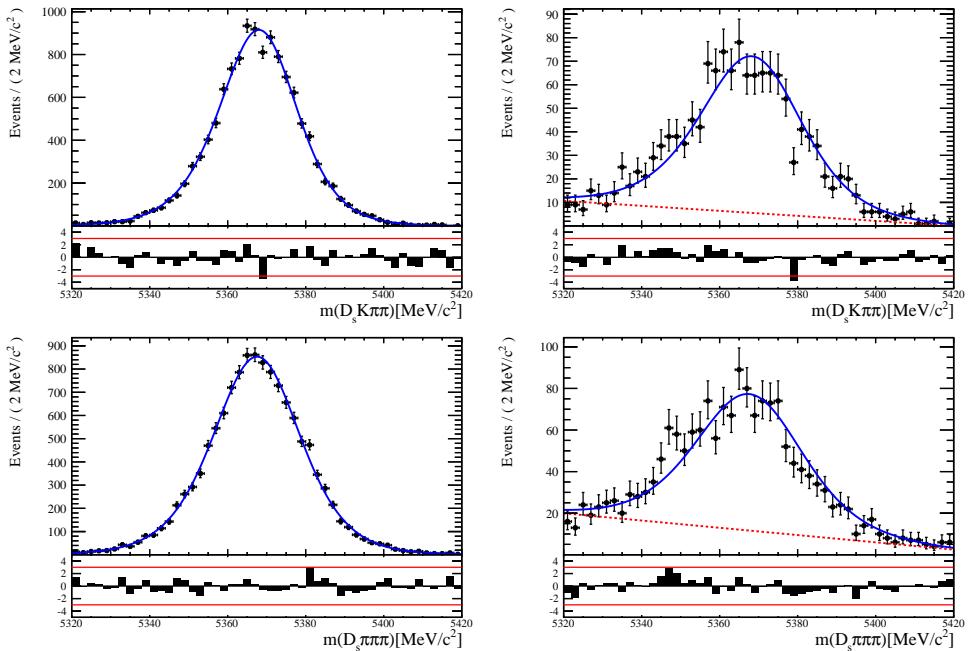


Figure 6.1: The reconstructed  $B_s$  mass distribution for simulated  $B_s \rightarrow D_s K\pi\pi$  (top) and  $B_s \rightarrow D_s \pi\pi\pi$  (bottom) decays classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

503 6.1.2 PID efficiencies

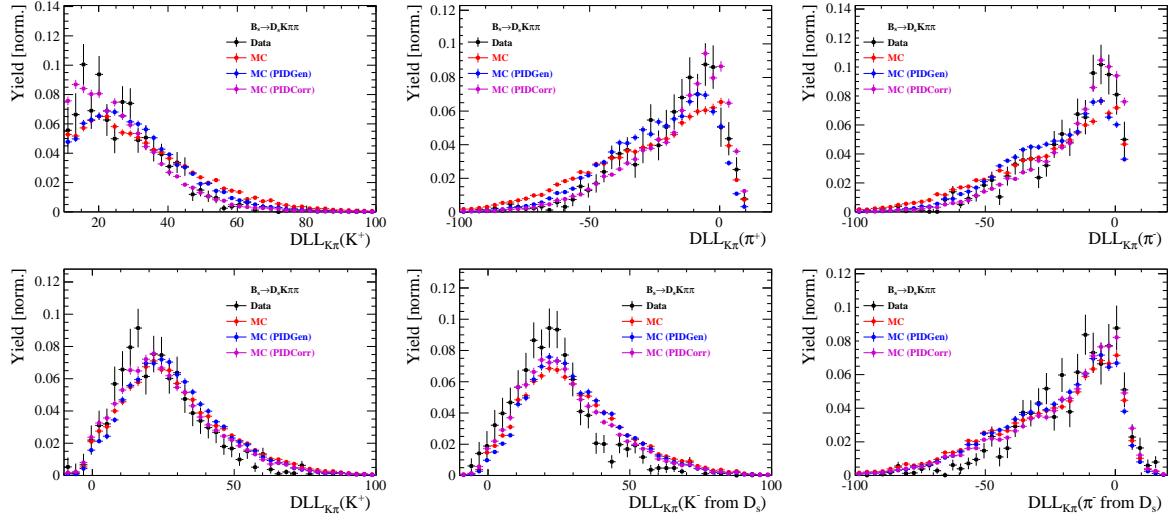


Figure 6.2

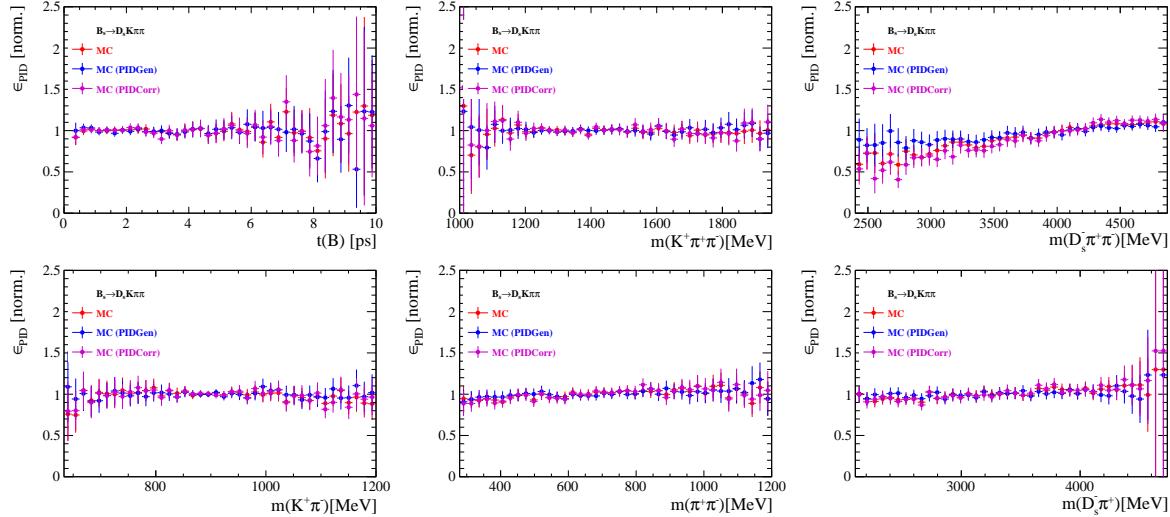


Figure 6.3

504 6.1.3 BDT efficiencies

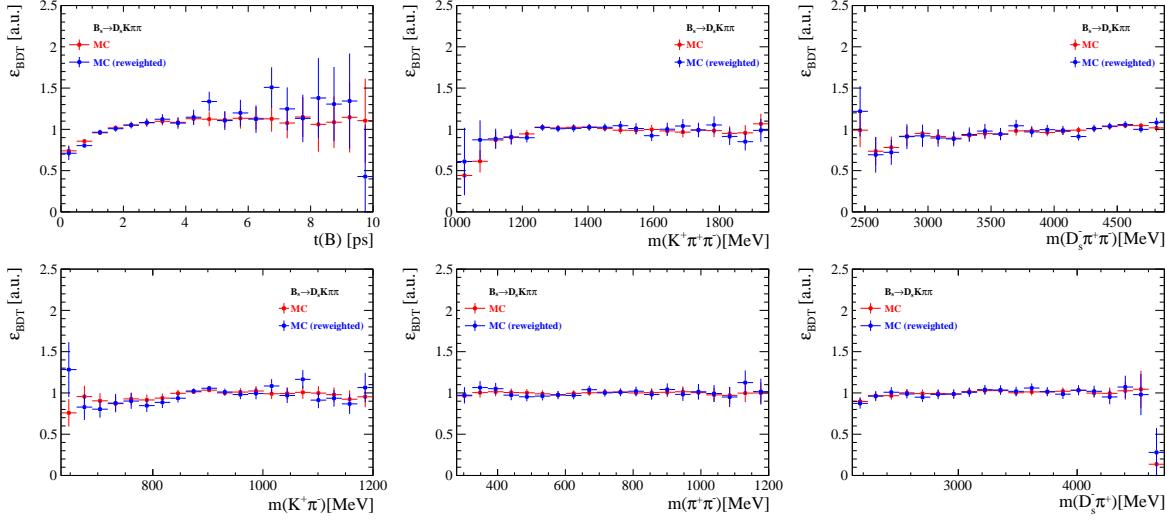


Figure 6.4

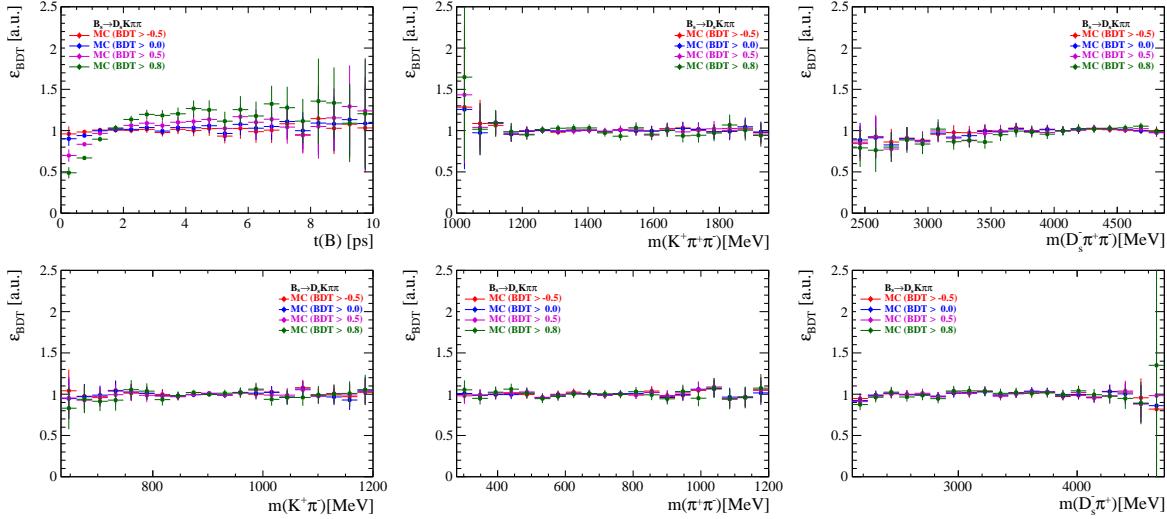


Figure 6.5

505    **6.1.4    Tracking efficiencies**

506 **6.2 Decay-time acceptance**

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

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

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

520 A PDF of the form

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

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

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

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

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

### 540 6.2.1 Comparison of acceptance in subsamples

541 It is possible that the decay-time dependent efficiency deviates in different subsamples of  
 542 our data. In particular, the acceptance could differentiate in subsamples with different  
 543 final state kinematics, such as the run I & run II sample, the various  $D_s$  final states and  
 544 the ways an event is triggered at the L0 stage. To investigate possible deviations, the  
 545 full selected  $B_s^0 \rightarrow D_s\pi\pi\pi$  sample is split into subsamples according to the categories  
 546 mentioned above (run,  $D_s$  state, L0 trigger). For each subsample, the fit procedure  
 547 described at the beginning of this chapter, using the pdf given by Eq. 6.2, is repeated  
 548 and the obtained values for the spline coefficients  $v_i$  are compared. Figure 6.6 shows the  
 549 comparison of the obtained spline coefficients for the different  $D_s$  final states.

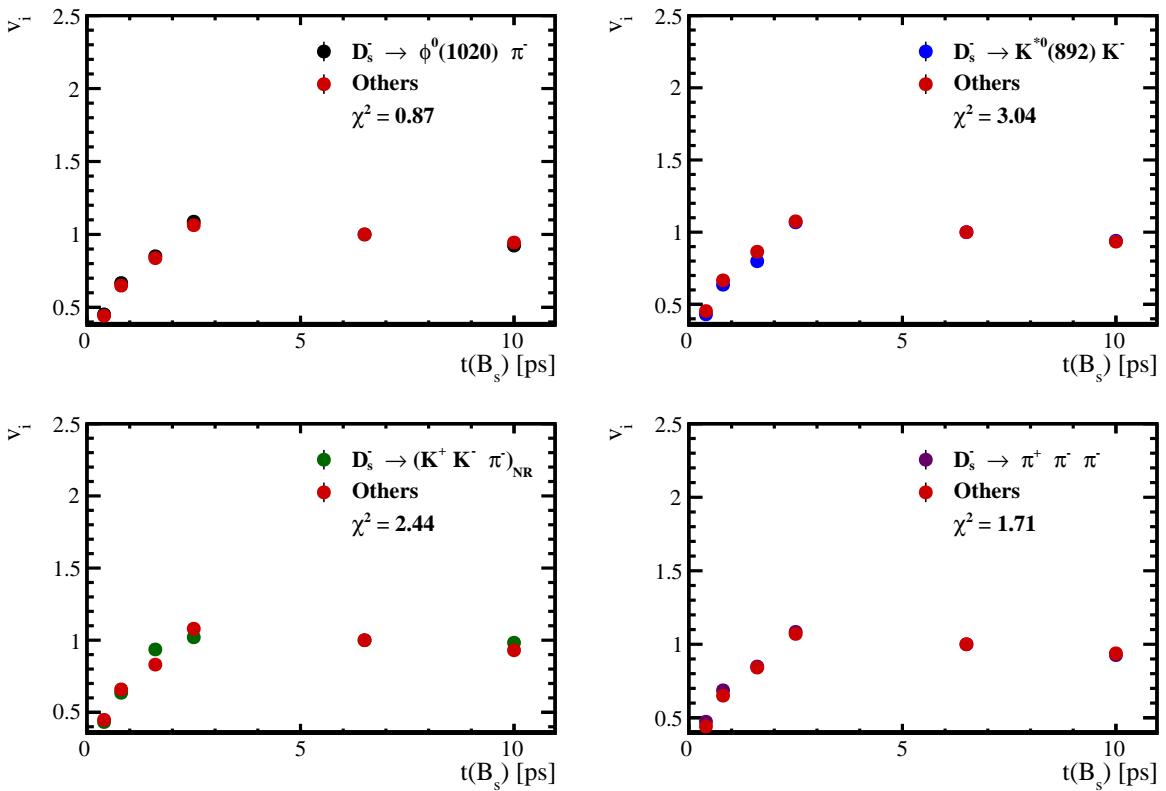


Figure 6.6: Comparison of the spline coefficients obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s\pi\pi\pi$  subsamples of different  $D_s$  final states. The comparison of one particular  $D_s$  state against all other states is shown.

550 Investigating the obtained spline coefficients from different  $D_s$  final states, good  
 551 agreement is observed between all four channels and no need to distinguish between  
 552 different final states in the time-dependent amplitude fit is found.

553 The comparison between spline coefficients for the different runs and L0 trigger categories  
 554 is shown in Figure 6.7.

555 Significant deviations between spline coefficients obtained from the two different runs  
 556 and L0 trigger categories can be observed. The deviations are most pronounced in the  
 557 (0 – 5) ps region, where the majority of statistics is found. Therefore, the time-dependent  
 558 efficiency has to be treated separately for the runs and L0 categories. This is achieved by

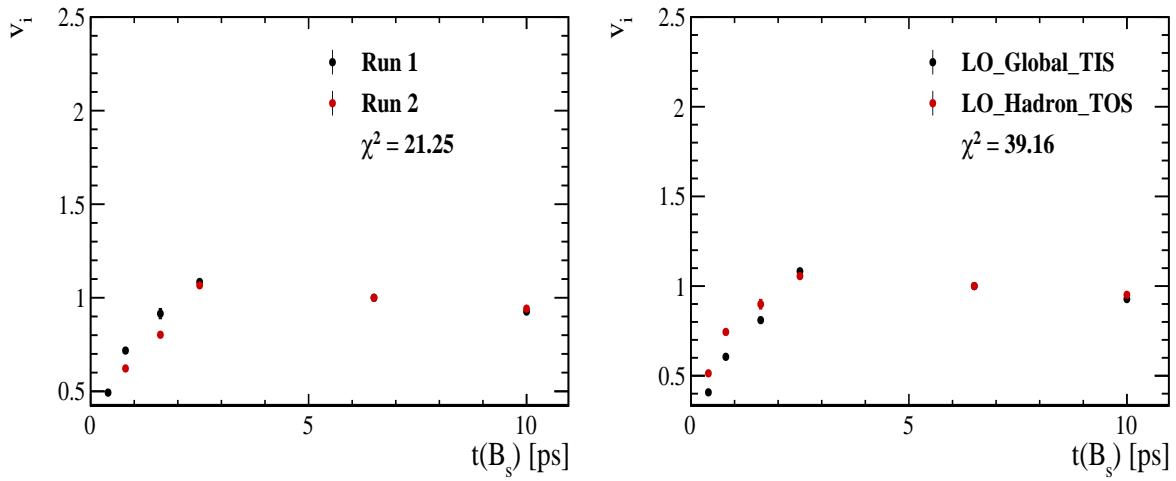


Figure 6.7: Comparison of the spline coefficients obtained from time-dependent fits to the  $B_s^0 \rightarrow D_s \pi \pi \pi$  subsamples of (left) the different runs and (right) L0 trigger categories.

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

561 **6.2.2 Results**

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

Table 6.1: Time acceptance parameters for events in category [Run-I,L0-TOS].

Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.561 \pm 0.038$	$0.546 \pm 0.022$	$0.953 \pm 0.060$
0.8	$v_1$	$0.826 \pm 0.059$	$0.785 \pm 0.034$	$0.910 \pm 0.066$
1.6	$v_2$	$0.843 \pm 0.087$	$0.905 \pm 0.056$	$1.055 \pm 0.095$
2.5	$v_3$	$1.154 \pm 0.036$	$1.118 \pm 0.028$	$0.930 \pm 0.045$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.866 (interpolated)	0.897 (interpolated)	1.061 (interpolated)

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

Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.368 \pm 0.031$	$0.412 \pm 0.020$	$0.955 \pm 0.077$
0.8	$v_1$	$0.583 \pm 0.050$	$0.648 \pm 0.033$	$0.910 \pm 0.074$
1.6	$v_2$	$0.939 \pm 0.101$	$0.953 \pm 0.061$	$0.947 \pm 0.096$
2.5	$v_3$	$1.052 \pm 0.054$	$1.077 \pm 0.035$	$1.003 \pm 0.051$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.954 (interpolated)	0.932 (interpolated)	0.998 (interpolated)

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

Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.486 \pm 0.009$	$0.482 \pm 0.009$	$1.000 \pm 0.000$
0.8	$v_1$	$0.691 \pm 0.014$	$0.707 \pm 0.015$	$1.000 \pm 0.000$
1.6	$v_2$	$0.851 \pm 0.024$	$0.926 \pm 0.026$	$1.000 \pm 0.000$
2.5	$v_3$	$1.061 \pm 0.017$	$1.086 \pm 0.018$	$1.000 \pm 0.000$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.946 (interpolated)	0.925 (interpolated)	1.000 (interpolated)

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

Knot position	Coefficient	$B_s^0 \rightarrow D_s K\pi\pi$ data	$B_s^0 \rightarrow D_s K\pi\pi$ MC	Ratio
0.4	$v_0$	$0.300 \pm 0.007$	$0.482 \pm 0.010$	$1.000 \pm 0.000$
0.8	$v_1$	$0.476 \pm 0.012$	$0.707 \pm 0.016$	$1.000 \pm 0.000$
1.6	$v_2$	$0.725 \pm 0.023$	$0.926 \pm 0.026$	$1.000 \pm 0.000$
2.5	$v_3$	$1.064 \pm 0.019$	$1.086 \pm 0.018$	$1.000 \pm 0.000$
6.5	$v_4$	1.0 (fixed)	1.0 (fixed)	1.0 (fixed)
10.0	$v_5$	0.944 (interpolated)	0.925 (interpolated)	1.000 (interpolated)

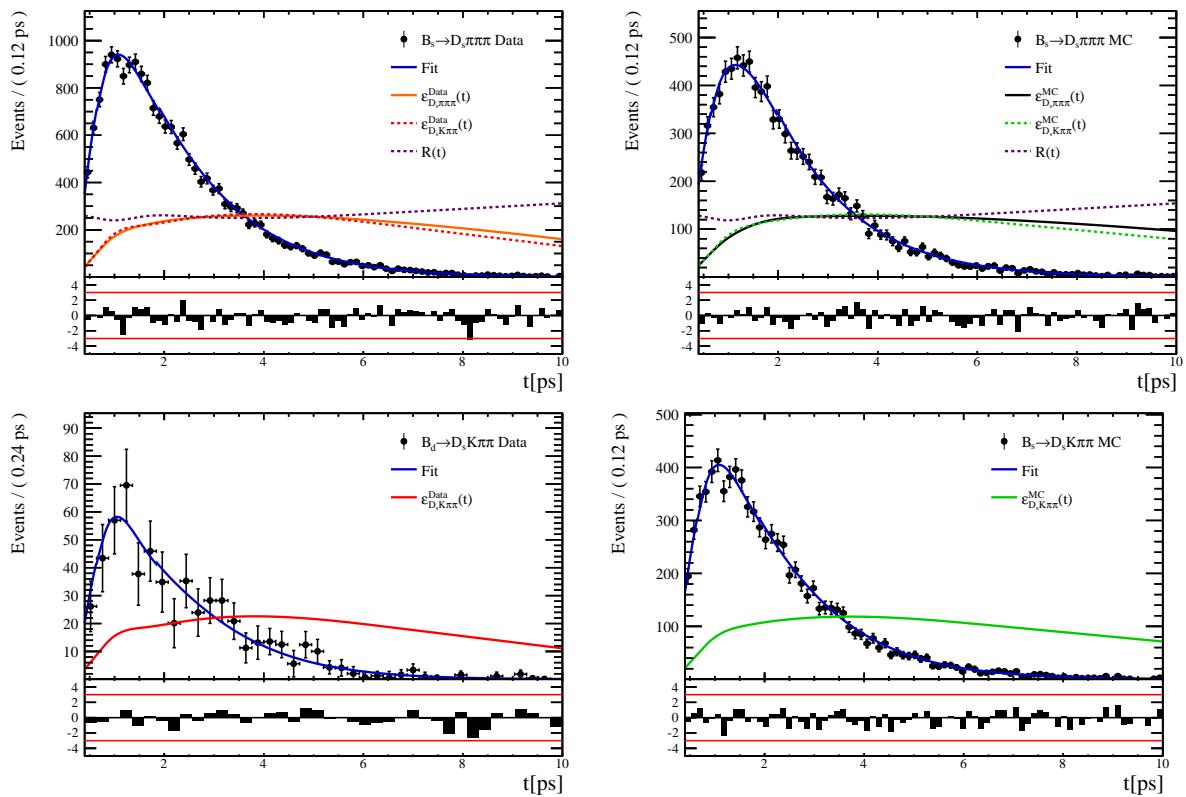


Figure 6.8

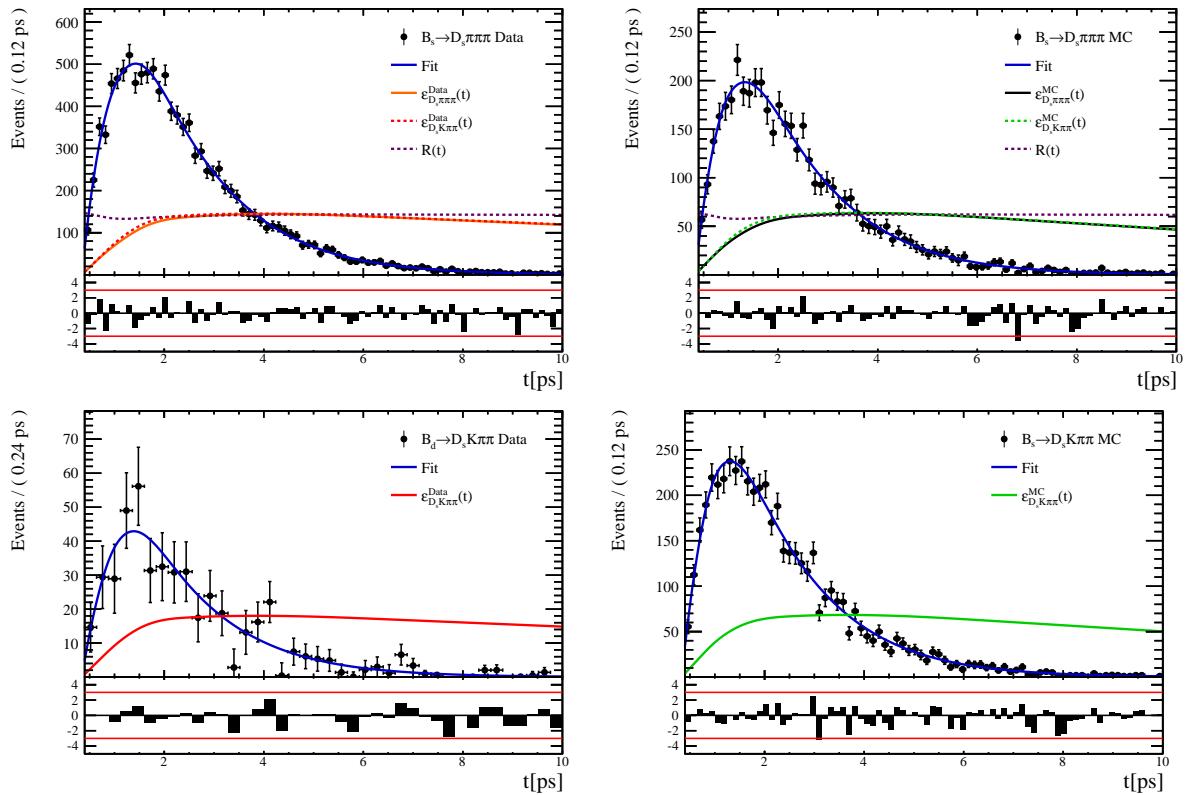


Figure 6.9

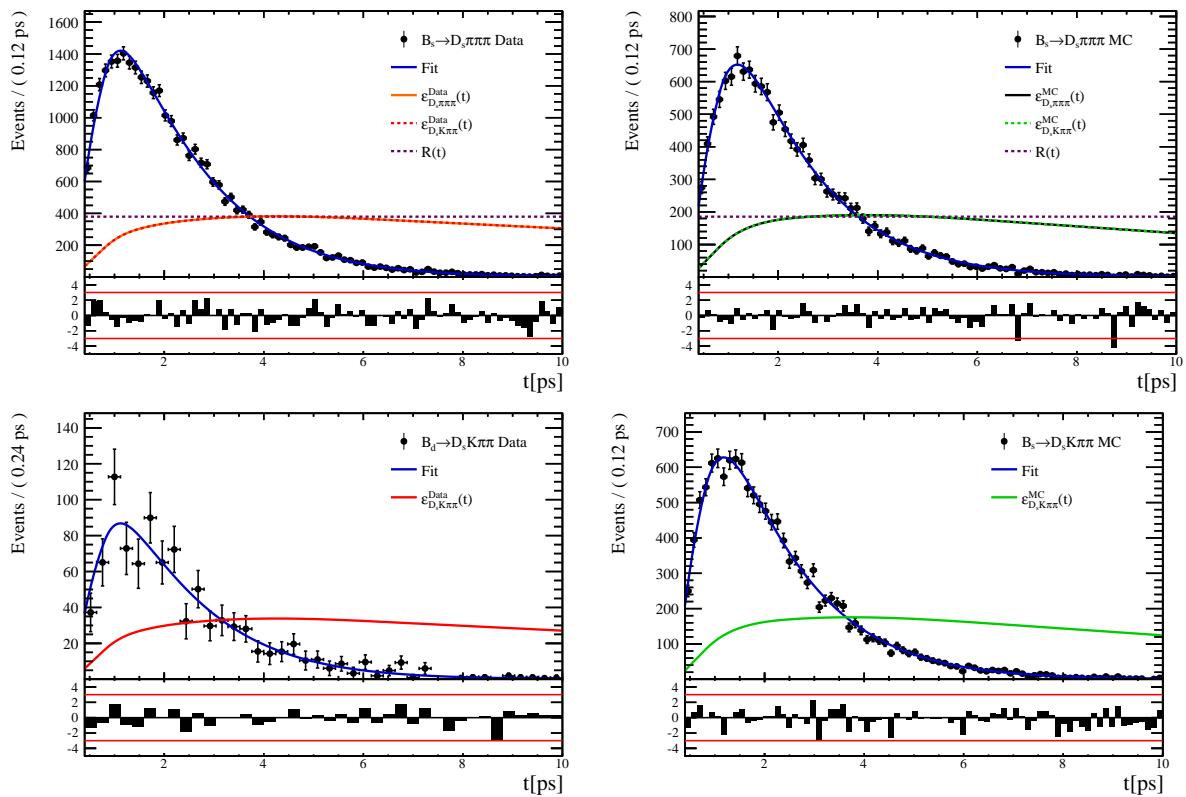


Figure 6.10

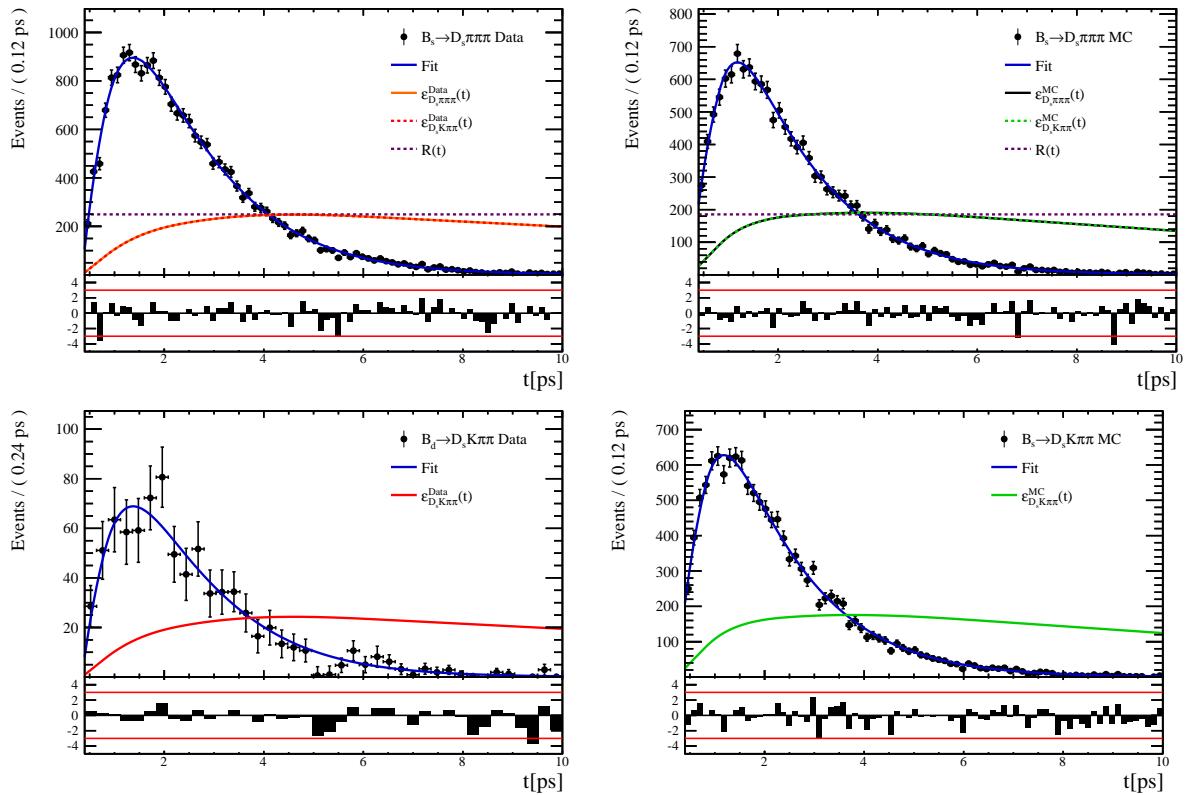


Figure 6.11

566 **6.3 Phasespace acceptance**

## 567 7 Decay-time Resolution

568 The observed oscillation of B mesons is prone to dilution, if the detector resolution is  
 569 of similar magnitude as the oscillation period. In the  $B_s^0$  system, considering that the  
 570 measured oscillation frequency of the  $B_s^0$  [33] and the average LHCb detector resolution [38]  
 571 are both  $\mathcal{O}(50 \text{ fs}^{-1})$ , this is the case. Therefore, it is crucial to correctly describe the  
 572 decay time resolution in order to avoid a bias on the measurement of time dependent CP  
 573 violation. Since the time resolution depends on the particular event, especially the decay  
 574 time itself, the sensitivity on  $\gamma$  can be significantly improved by using an event dependent  
 575 resolution model rather than an average resolution. For this purpose, we use the per-event  
 576 decay time error that is estimated based on the uncertainty obtained from the global  
 577 kinematic fit of the momenta and vertex positions (DTF) with additional constraints on  
 578 the PV position and the  $D_s$  mass. In order to apply it correctly, it has to be calibrated.  
 579 The raw decay time error distributions for  $B_s^0 \rightarrow D_s\pi\pi\pi$  signal candidates are shown in  
 580 Figure 7.1 for Run-I and Run-II data. Significant deviations between the two different  
 581 data taking periods are observed due to the increase in center of mass energy from Run-I  
 582 to Run-II, as well as (among others) new tunings in the pattern and vertex reconstruction.  
 583 The decay time error calibration is consequently performed separately for both data taking  
 584 periods.

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

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

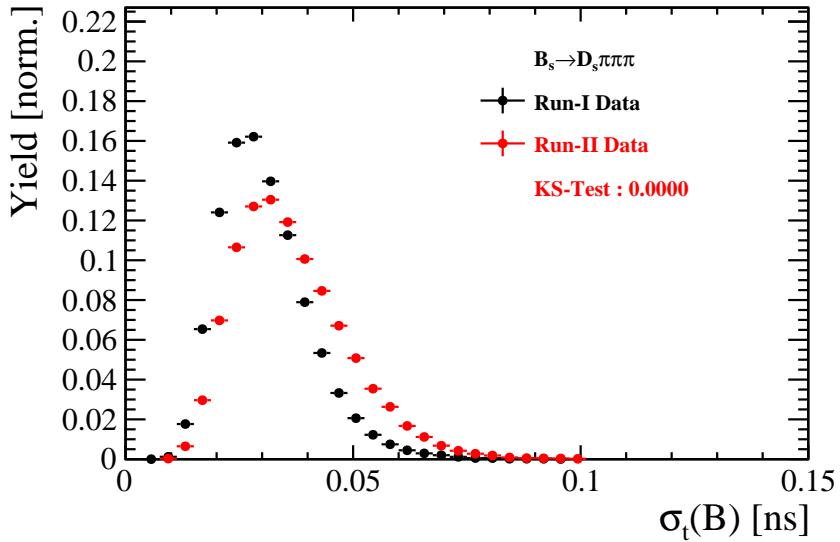


Figure 7.1: Distribution of the decay time error for  $B_s^0 \rightarrow D_s\pi\pi\pi$  signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

## 590 7.1 Calibration for Run-I data

591 For simulated  $B_s^0 \rightarrow D_s K \pi\pi$  events, the spread of the differences between reconstructed  
 592 decay time and true decay time,  $\Delta t = t - t_{true}$ , is a direct measure of the decay time  
 593 resolution. The sum of two Gaussian pdfs with common mean but different widths is used  
 594 to fit the distribution of the decay time difference  $\Delta t$  as shown in Fig. 7.2. The effective  
 595 damping of the CP amplitudes due to the finite time resolution is described by the dilution  
 596  $\mathcal{D}$ . In the case of infinite precision, there would be no damping and therefore  $\mathcal{D} = 1$  would  
 597 hold, while for a resolution that is much larger than the  $B_s^0$  oscillation frequency,  $\mathcal{D}$  would  
 598 approach 0. For a double-Gaussian resolution model, the dilution is given by [39]

$$\mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s^2 / 2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s^2 / 2}, \quad (7.1)$$

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

$$\sigma_{eff} = \sqrt{(-2/\Delta m_s^2) \ln \mathcal{D}}, \quad (7.2)$$

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

605 To analyze the relation between the per-event decay time error  $\delta_t$  and the actual  
 606 resolution  $\sigma_t$ , the simulated  $B_s^0 \rightarrow D_s K \pi\pi$  sample is divided into equal-statistics slices of  
 607  $\delta_t$ . For each slice, the effective resolution is determined as described above. Details of the  
 608 fit results in each slice are shown in appendix C. Figure 7.2 shows the obtained values  
 609 for  $\sigma_{eff}$  as a function of the per-event decay time error  $\sigma_t$ . To account for the variable  
 610 binning, the bin values are not placed at the bin center but at the weighted mean of the  
 611 respective per-event-error bin. A linear function is used to parametrize the distribution.  
 612 The obtained values are

$$\sigma_{eff}^{MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.232 \pm 0.010) \sigma_t \quad (7.3)$$

613 where the offset is fixed to 0. For comparison, the calibration function found for  $B_s^0 \rightarrow D_s K$   
 614 MC is also shown in Figure 7.2 [39]:

$$\sigma_{eff}^{D_s K, MC}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.201 \pm 0.013) \sigma_t. \quad (7.4)$$

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

$$\sigma_{eff}^{D_s K}(\sigma_t) = (10.26 \pm 1.52) \text{ fs} + (1.280 \pm 0.042) \sigma_t \quad (7.5)$$

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

$$\sigma_{eff}^{D_s K}(\sigma_t) = (-0.568 \pm 1.570) \text{ fs} + (1.243 \pm 0.044) \sigma_t \quad (7.6)$$

$$\sigma_{eff}^{D_s K}(\sigma_t) = (0.0 \pm 0.0) \text{ fs} + (1.772 \pm 0.012) \sigma_t \quad (7.7)$$

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

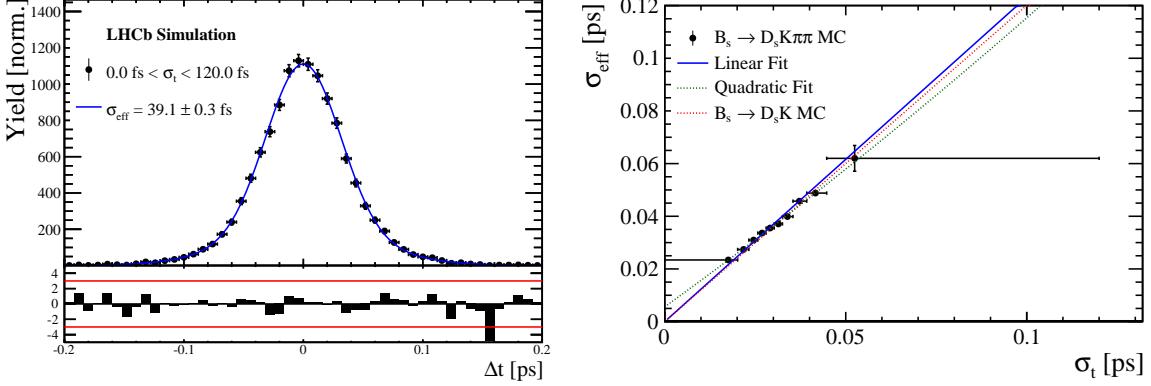


Figure 7.2: (Left) Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K\pi\pi$  MC candidates. (Right) The measured resolution  $\sigma_{eff}$  as function of the per-event decay time error estimate  $\sigma_t$  for  $B_s \rightarrow D_s K\pi\pi$  MC (Run-I). The fitted calibration curve is shown in blue.

## 7.2 Calibration for Run-II data

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

The offline selection of the fake  $B_s$  candidates is summarized in Tab. 7.1. The selection is similar than the one for real data with the important difference that the  $D_s$  candidate is required to come from the PV by cutting on the impact parameter significance. Even after the full selection, a significant number of multiple candidates is observed. These are predominantly fake  $B_s$  candidates that share the same  $D_s$  candidate combined with different random tracks from the PV. We select one of these multiple candidates randomly which retains approximately 20% of the total candidates. The invariant mass distribution of the selected  $D_s$  candidates is shown in Fig. 7.3. To separate true  $D_s$  candidates from random combinations, the `sPlot` method is used to statistically subtract combinatorial background from the sample.

Figure 7.4 shows the `sWeighted` decay-time distribution for fake  $B_s$  candidates. Similar as in the previous section, the decay-time distribution is fitted with a double-Gaussian resolution model in slices of the per-event decay time error. Since some  $D_s$  candidates might actually originate from true  $B_s$  decays, the decay-time distribution of the fake  $B_s$  candidates might show a bias towards positive decay times. Therefore, we determine the decay-time resolution from the negative decay-time distribution only. Details of the fit results in each slice are shown in appendix C. The resulting calibration function:

$$\sigma_{eff}^{Data}(\sigma_t) = (9.7 \pm 1.7) \text{ fs} + (0.915 \pm 0.040) \sigma_t \quad (7.8)$$

<sup>649</sup> is in good agreement with the one obtained for the  $B_s \rightarrow J/\psi\phi$  (Run-II) analysis.

$$\sigma_{eff}^{J/\psi\phi}(\sigma_t) = (12.25 \pm 0.33) \text{ fs} + (0.8721 \pm 0.0080) \sigma_t \quad (7.9)$$

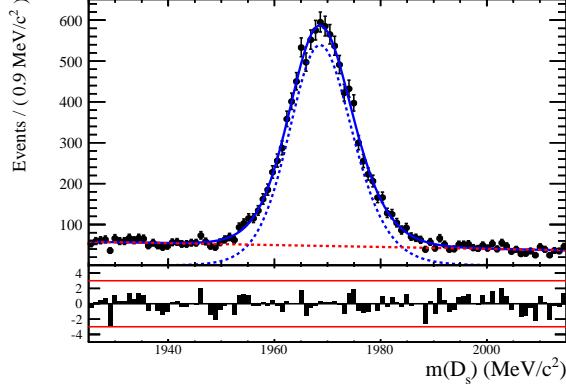


Figure 7.3: The invariant mass distribution for prompt  $D_s$  candidates.

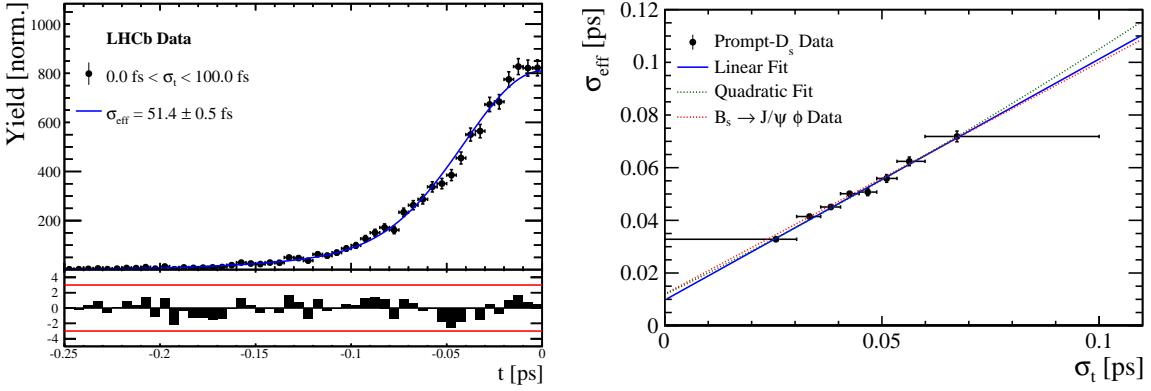


Figure 7.4: (Left) Decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks. (Right) The measured resolution  $\sigma_{eff}$  as function of the per-event decay time error estimate  $\sigma_t$  for fake  $B_s$  candidates (Run-II data). The fitted calibration curve is shown in blue.

Table 7.1: Offline selection requirements for fake  $B_s$  candidates from promptly produced  $D_s$  candidates combined with random prompt  $K\pi\pi$  bachelor tracks.

	Description	Requirement
$B_s \rightarrow D_s K\pi\pi$	$\chi^2_{vtx}/\text{ndof}$	< 8
	$\chi^2_{DTF}/\text{ndof}$	< 15
	$t$	< 0 ps
$D_s \rightarrow hhh$	$\chi^2_{vtx}/\text{ndof}$	< 5
	DIRA	> 0.99994
	$\chi^2_{FD}$	> 9
	$p_T$	> 1800 MeV
	$\chi^2_{IP}$	< 9
	$\chi^2_{IP}(h)$	> 5
	Wrong PV veto	$nPV = 1 \parallel \min(\Delta\chi^2_{IP}) > 20$
$D_s^- \rightarrow KK\pi^-$	$D^0$ veto	$m(KK) < 1840$ MeV
	$D^-$ veto	$m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 30$ MeV
	$\Lambda_c$ veto	$m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 30$ MeV
$D_s^- \rightarrow \phi\pi^-$	$m(KK)$	$= m_\phi \pm 20$ MeV
	PIDK( $K^+$ )	> -10
	PIDK( $K^-$ )	> -10
	PIDK( $\pi^-$ )	< 20
$D_s^- \rightarrow K^*(892)K^-$	$m(KK)$	$\neq m_\phi \pm 20$ MeV
	$m(K^+\pi^-)$	$= m_{K^*(892)} \pm 75$ MeV
	PIDK( $K^+$ )	> -10
	PIDK( $K^-$ )	> -5
	PIDK( $\pi^-$ )	< 20
$D_s^- \rightarrow (KK\pi^-)_{NR}$	$m(KK)$	$\neq m_\phi \pm 20$ MeV
	$m(K^+\pi^-)$	$\neq m_{K^*(892)} \pm 75$ MeV
	PIDK( $K^+$ )	> 5
	PIDK( $K^-$ )	> 5
	PIDK( $\pi^-$ )	< 10
$D_s \rightarrow \pi\pi\pi$	PIDK( $h$ )	< 10
	PIDp( $h$ )	< 10
	$D^0$ veto	$m(\pi^+\pi^-) < 1700$ MeV
$X_s \rightarrow K\pi\pi$	$\chi^2_{IP}(h)$	< 40
	PIDK( $K$ )	> 10
	PIDK( $\pi$ )	< 5
	isMuon( $h$ )	False
All tracks	$p_T$	> 500 MeV

650 7.3 Cross-checks

651 7.3.1 Kinematic dependence

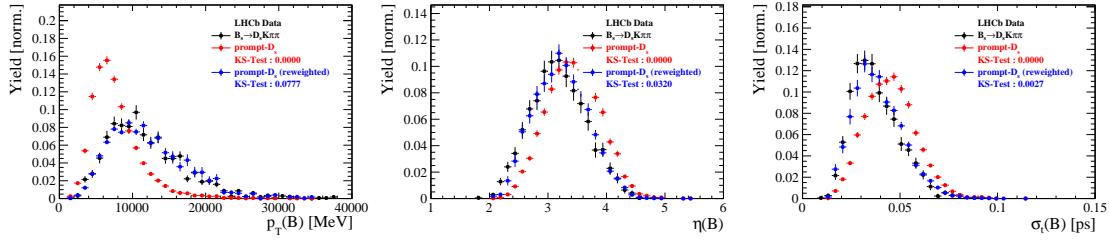


Figure 7.5

652 7.3.2 DTF constraints

## 653 8 Production and Detection Asymmetries

### 654 8.1 $B_s$ Production Asymmetry

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

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

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

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

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

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

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

Table 8.1:  $B_s$  production asymmetries in kinematic bins for 2011 and 2012 data. [40]

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

## 667 8.2 $K^-\pi^+$ Detection Asymmetry

668 The presented measurement of the CKM-angle  $\gamma$  using  $B_s^0 \rightarrow D_s K\pi\pi$  decays is sensitive  
 669 to a possible charge asymmetry of the kaon. This effect can be detector induced, because  
 670 kaons are known to have a nuclear cross-section which is asymmetrically dependent on  
 671 the sign of their charge. It is indispensable to determine the detector induced charge  
 672 asymmetry of the kaon, as fitting without taking this effect into account would introduce  
 673 a ‘fake’ CP violation. Instead of determining the single track detection asymmetry of a  
 674 kaon, it is found that the combined two track asymmetry of a kaon-pion pair is much  
 675 easier to access [41]. Therefore the two track asymmetry is used, which is defined as

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

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

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

678 where possible CP violation in the  $D^+ \rightarrow K_s^0\pi^+$  mode is predicted to be smaller than  
 679  $10^{-4}$  in the Standard Model [43]. The asymmetry in the neutral kaon system,  $A(K^0)$ , has  
 680 to be taken into account as a correction.

681 We use a dedicated LHCb tool to determine  $A^{det}(K^-\pi^+)$  for all data taking periods  
 682 used in this analysis. A detailed description can be found in [42]. The tool provides  
 683 large calibration samples of  $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$  and  $D^\pm \rightarrow K_s^0\pi^\pm$  decays, which are used to  
 684 determine the asymmetry following Eq. 8.6. Several weighting steps are performed to  
 685 match the kinematics of the calibration samples to our signal decay sample:

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

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

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

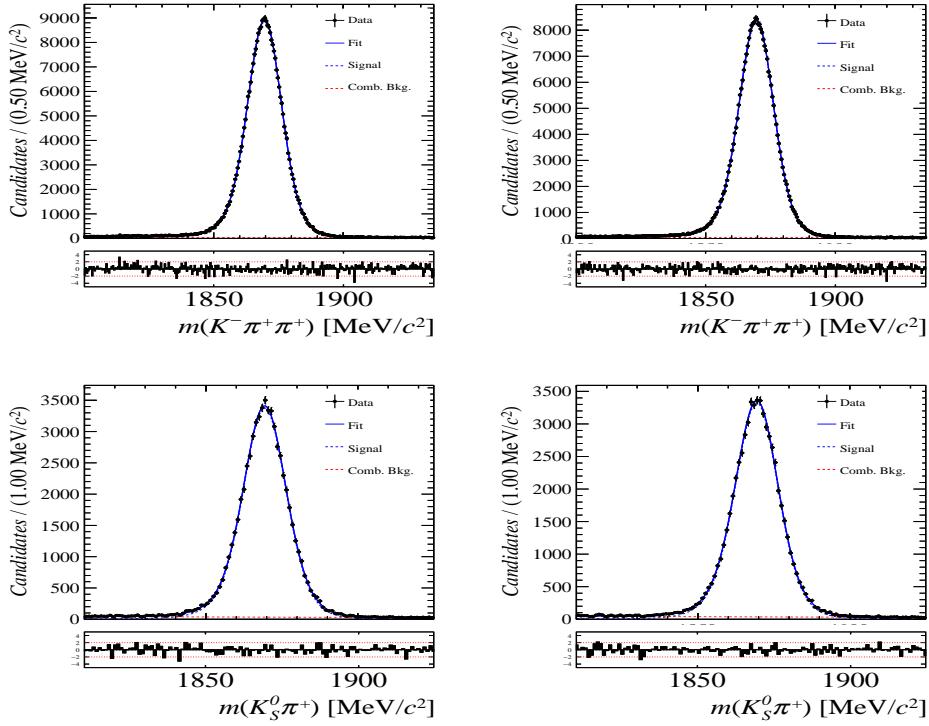


Figure 8.1: Distributions of the invariant mass of (top)  $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$  and (bottom)  $D^\pm \rightarrow K_S^0 \pi^\pm$  candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

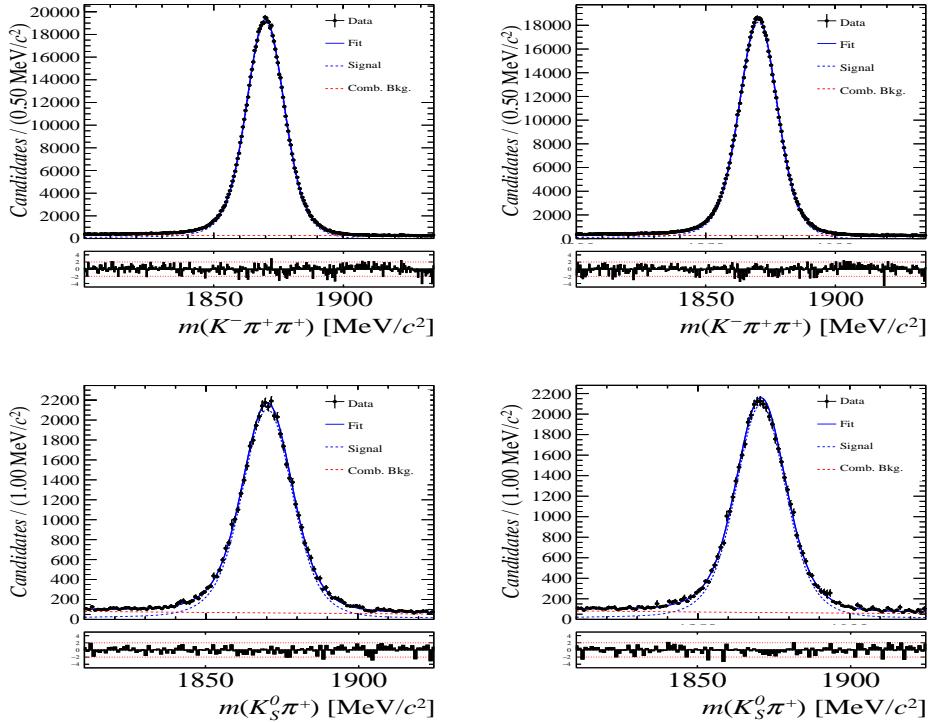


Figure 8.2: Distributions of the invariant mass of (top)  $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$  and (bottom)  $D^\pm \rightarrow K_S^0 \pi^\pm$  candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

Data sample	$A^{det}(K^-\pi^+) + A(K^0)$ [%]
Run-I	
2011, mag. up	-2.01 $\pm$ 0.32
2011, mag. down	-0.16 $\pm$ 0.28
2011, average	-1.09 $\pm$ 0.21
2012, mag. up	-0.90 $\pm$ 0.20
2012, mag. down	-1.01 $\pm$ 0.22
2012, average	-0.96 $\pm$ 0.15
Run-II	
2015, mag. up	-1.36 $\pm$ 0.36
2015, mag. down	-0.96 $\pm$ 0.24
2015, average	-1.16 $\pm$ 0.22
2016, mag. up	0.50 $\pm$ 0.88
2016, mag. down	1.23 $\pm$ 0.72
2016, average	0.87 $\pm$ 0.57

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

## 9 Time dependent fit

This section will cover the phasespace integrated, time dependent fit to  $B_s^0 \rightarrow D_s h\pi\pi$  data, which is described by the PDF formulated in Eq. 2.6. For the phasespace integrated fit to  $B_s^0 \rightarrow D_s K\pi\pi$  data, the sensitivity to the CKM phase  $\gamma$  will depend on the magnitude of the coherence factor  $\kappa$ , defined in Eq. 2.10, which is added as an additional fit parameter. In order to avoid any pollution of the final data samples by background events, both samples are weighted using the sWeights obtained by the fits to the invariant mass distributions, described in Sec. 4. This fit approach is commonly known as *sFit*. As additional input to the fit, the tagging information (Sec. 5), as well as the decay time acceptance (Sec. 6) and resolution (Sec. 7) is used and fixed to the values obtained by the dedicated studies. Taking all inputs into account, the final time dependent fit PDF is given by

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

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

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

The phasespace-integrated, time-dependent fit is performed to the full sWeighted sample of selected candidates from Run I and 2015+2016 Run II data, containing both possible magnet polarities and  $D_s$  final states. In the fit, the values of  $\Gamma_s$  and  $\Delta\Gamma_s$  are fixed to the latest PDG report. All tagging parameters are fixed to the central values found in the tagging calibration, described in Sec. 5. Due to the fact that the  $B_s^0 \rightarrow D_s \pi\pi\pi$  decay is flavour specific, the CP-coefficients can be fixed to  $C = 1$  and  $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$ , reducing Eq. 2.6 to

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

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

$$\Delta m_s = xx.xxx \pm 0.yyy. \quad (9.3)$$

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

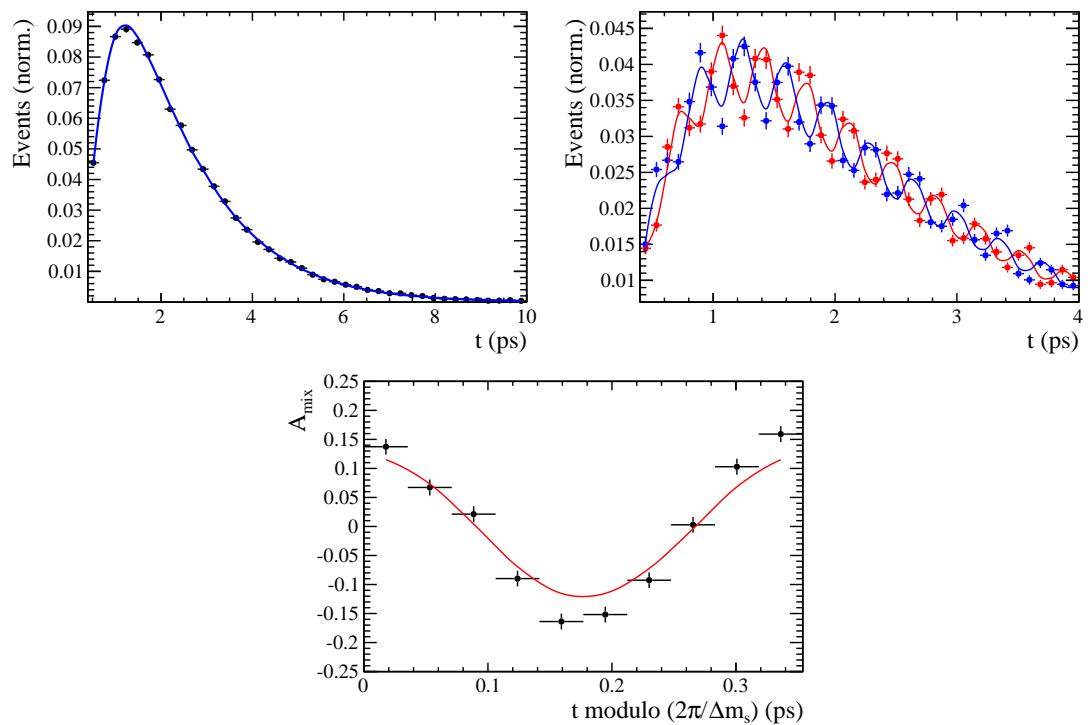


Figure 9.1

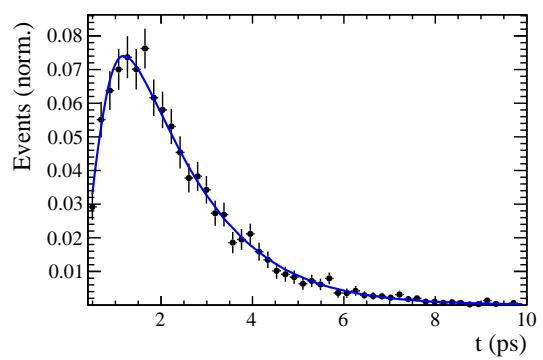


Figure 9.2

## 730 10 Time dependent amplitude fit

### 731 10.1 Signal Model Construction

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

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

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

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

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

748 where  $N_{\text{Sig}}$  is the number of signal events and  $r$  is the number of amplitudes with a decay  
749 fraction above a certain threshold. In this way, the optimal  $\lambda$  balances the fit quality  
750 ( $-2 \log \mathcal{L}$ ) against the model complexity. The LASSO penalty term is only used to select  
751 the model. Afterwards, this term must be discarded in the final amplitude fit with the  
752 selected model, otherwise the parameter uncertainties would be biased.

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

- 757 1. The fit fraction threshold for inclusion in the final model was varied within the  
758 interval [0.05, 5]%. The set of selected amplitudes is stable for thresholds between  
759 0.1% and 1%. Other choices result in marginally different models containing one  
760 component more or less.
  - 761 2. Instead of BIC, the Akaike information criteria ( $\text{AIC}(\lambda) = -2 \log \mathcal{L} + 2r$  [47]) was  
762 used to optimize  $\lambda$ . For a given threshold, the AIC method tends to prefer lower  
763  $\lambda$  values. However, the set of models obtained varying the threshold within the  
764 interval [0.05, 5]% is identical to the BIC method.
  - 765 3. The amplitudes selected under nominal conditions were excluded one-by-one from  
766 the set of all amplitudes considered.
- 767 From that we obtained a set of alternative models shown in Appendix ??.

768 **10.2 Results**

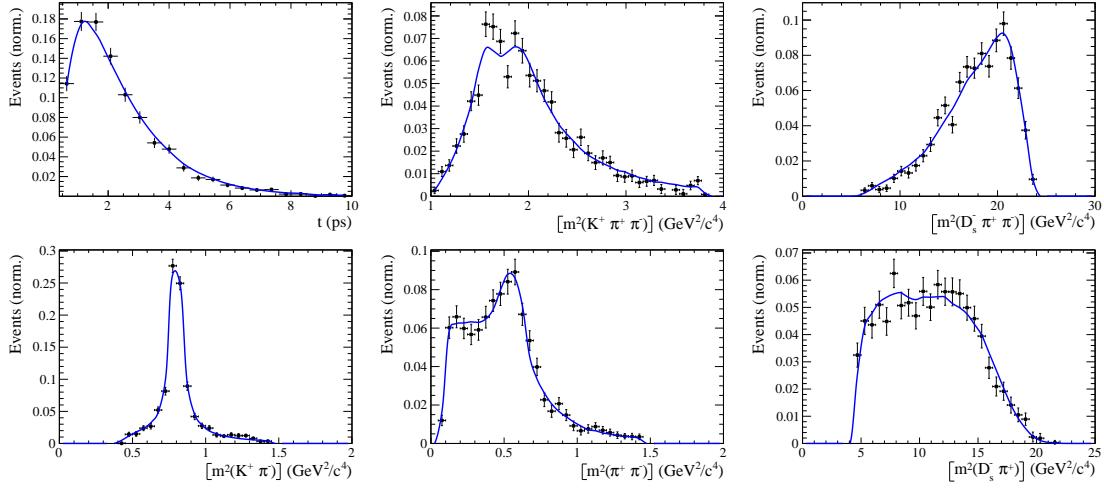


Figure 10.1

## A Details of multivariate classifier

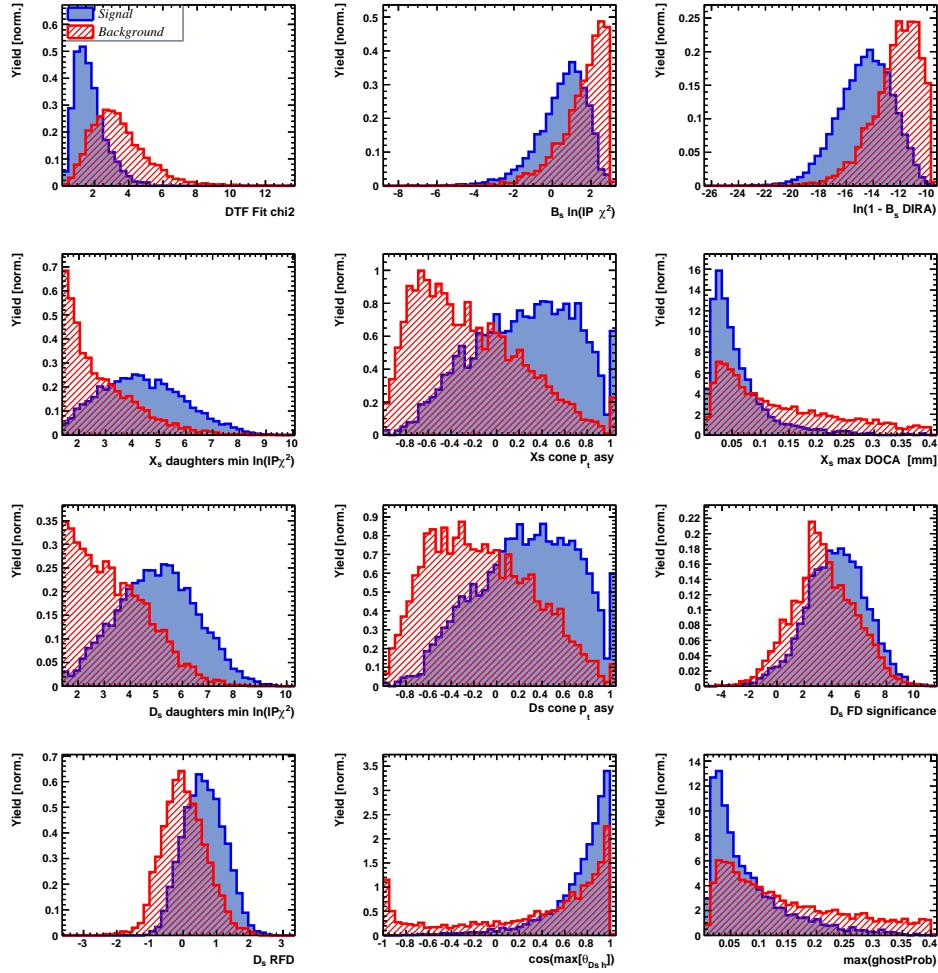


Figure A.1: Variables used to train the BDTG.

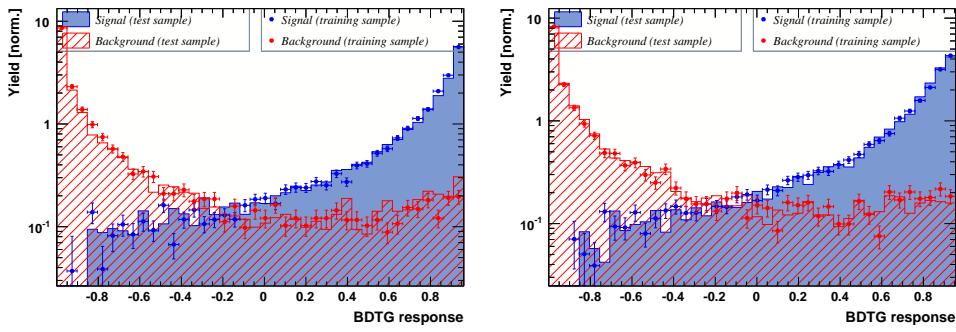


Figure A.2: Response of the classifier trained on the even (left) and odd (right) sample.

## 770 B Detailed mass fits

771 In this section, all fits to the mass distribution of  $B_s^0 \rightarrow D_s\pi\pi\pi$  and  $B_s^0 \rightarrow D_sK\pi\pi$   
 772 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-  
 773 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  
 774 L0 trigger category.

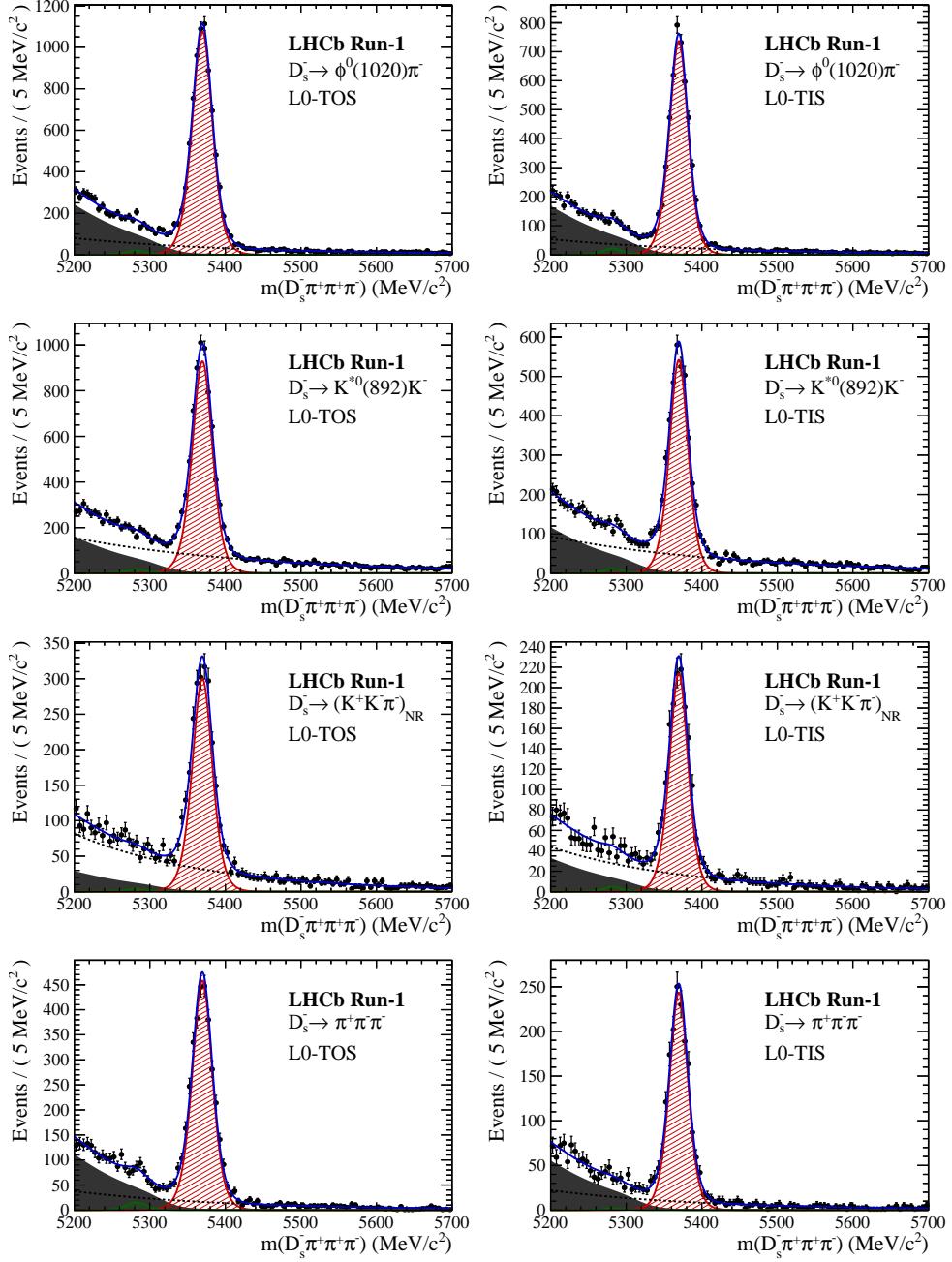


Figure B.1: Invariant mass distributions of  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates for Run-I data.

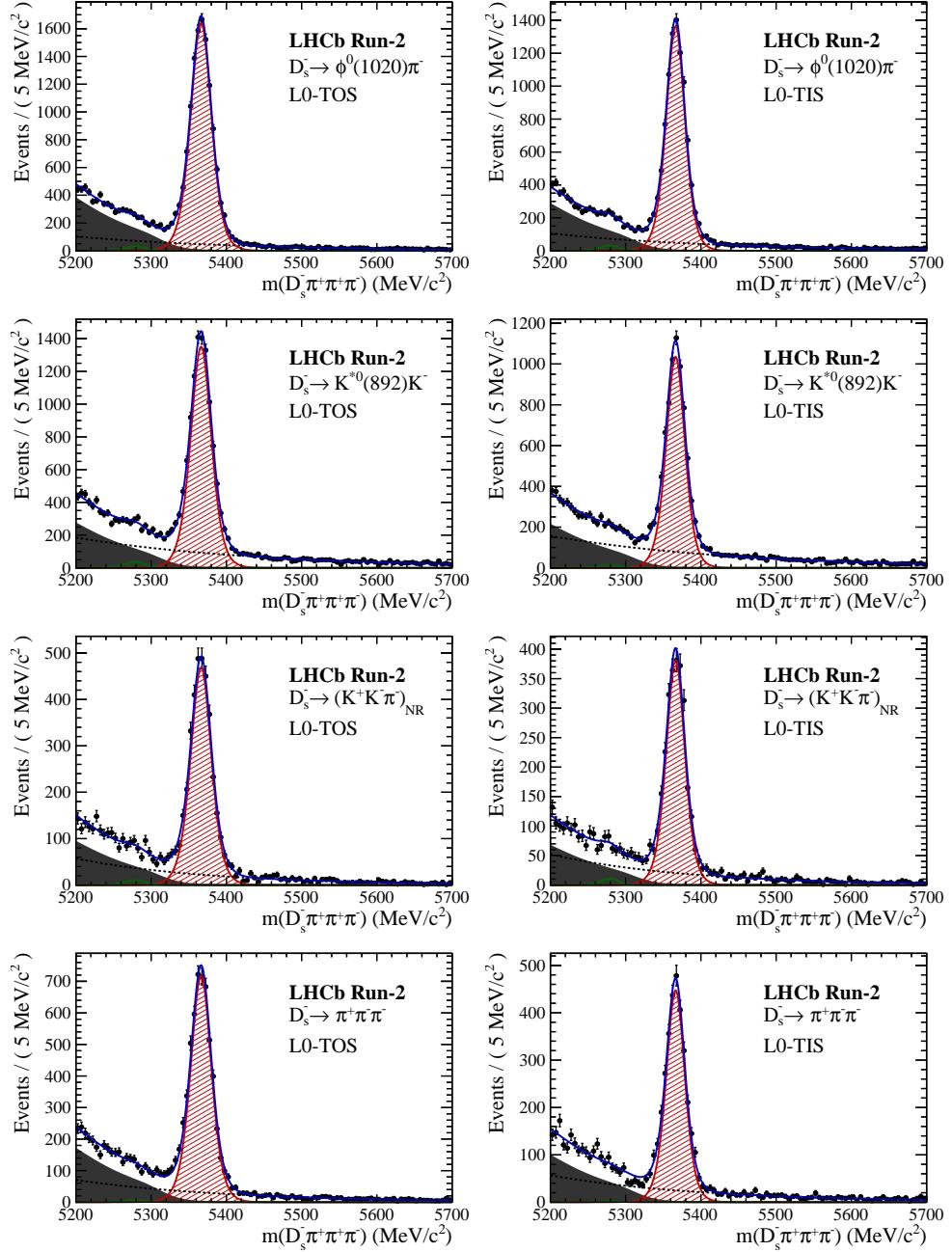


Figure B.2: Invariant mass distributions of  $B_s^0 \rightarrow D_s\pi\pi\pi$  candidates for Run-II data.

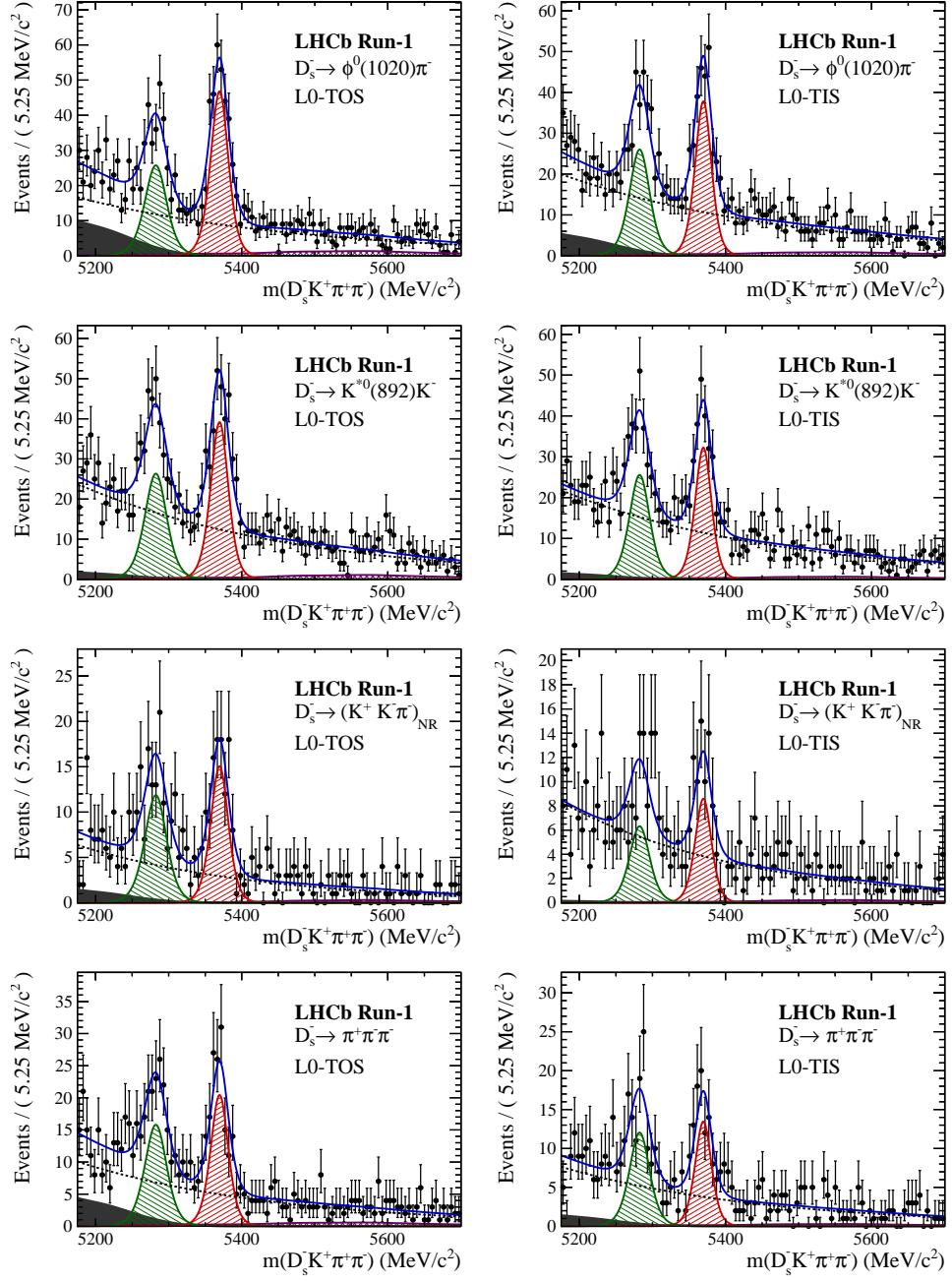


Figure B.3: Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for Run-I data.

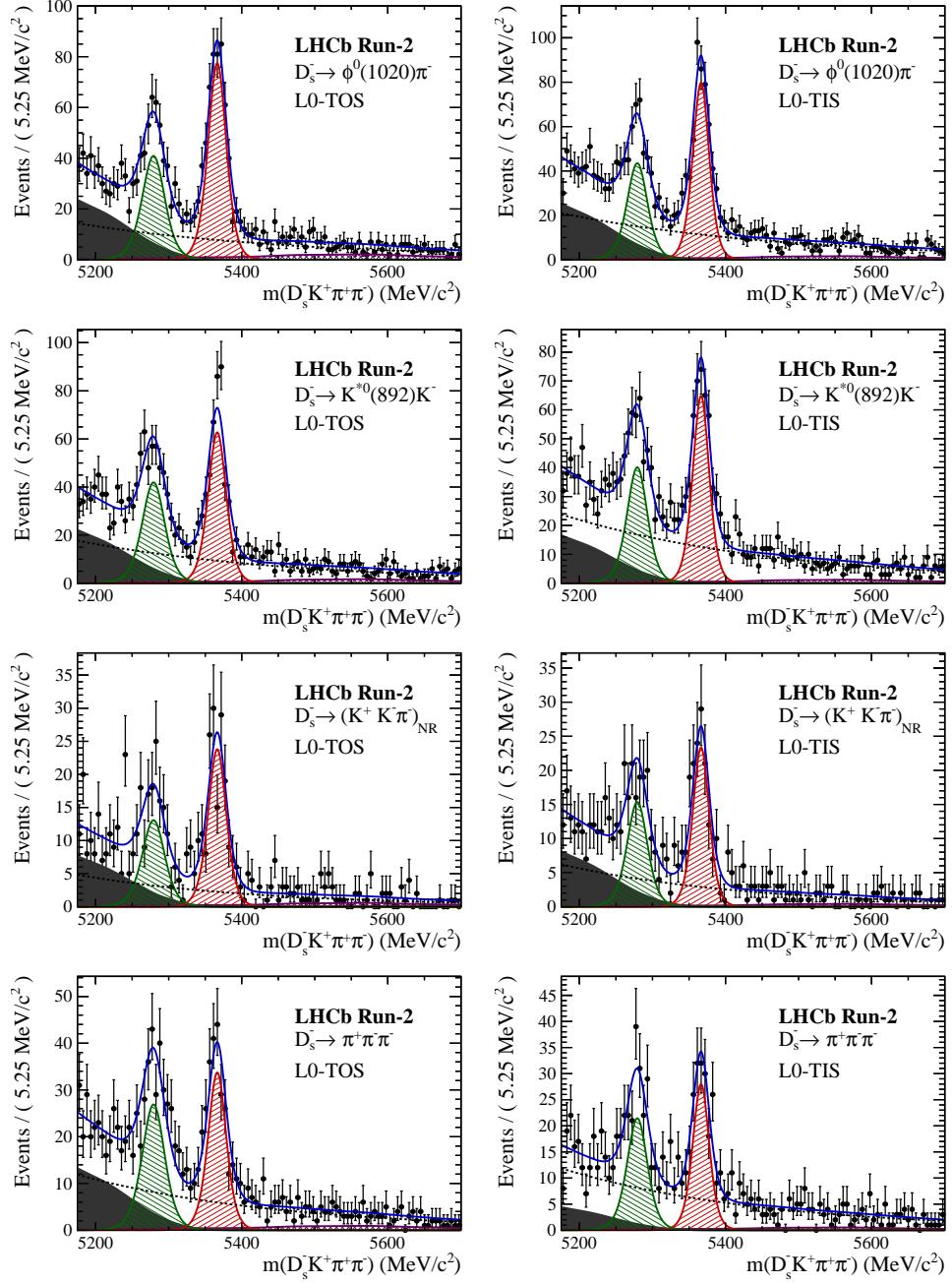


Figure B.4: Invariant mass distributions of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for Run-II data.

## 775 C Decay-time Resolution fits

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

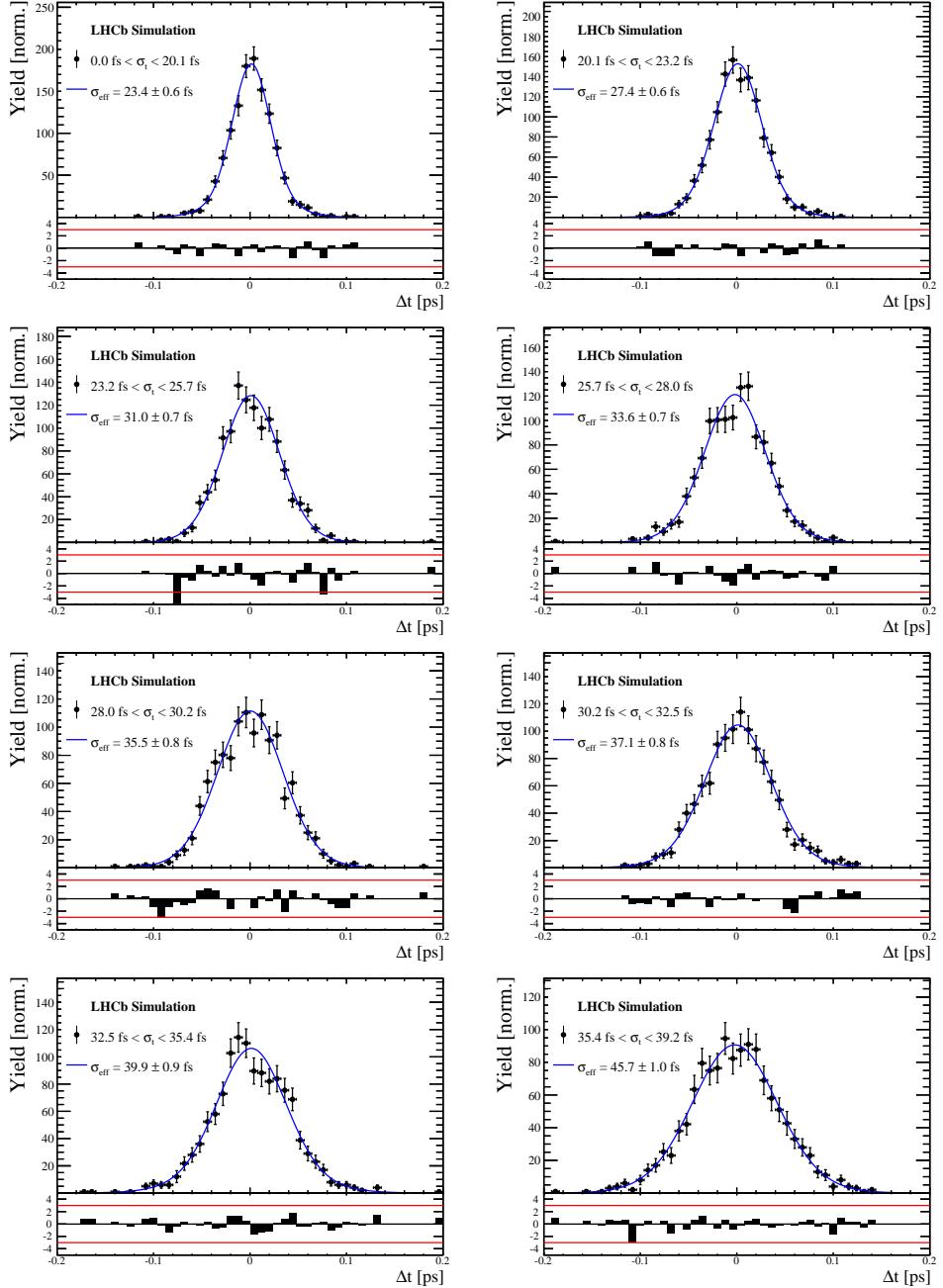


Figure C.1: Difference of the true and measured decay time of  $B_s^0 \rightarrow D_s K\pi\pi$  MC candidates in bins of the per-event decay time error estimate..

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

Table 3.1: Measured time resolution for  $B_s \rightarrow D_s K\pi\pi$  MC in bins of the per-event decay time error estimate.

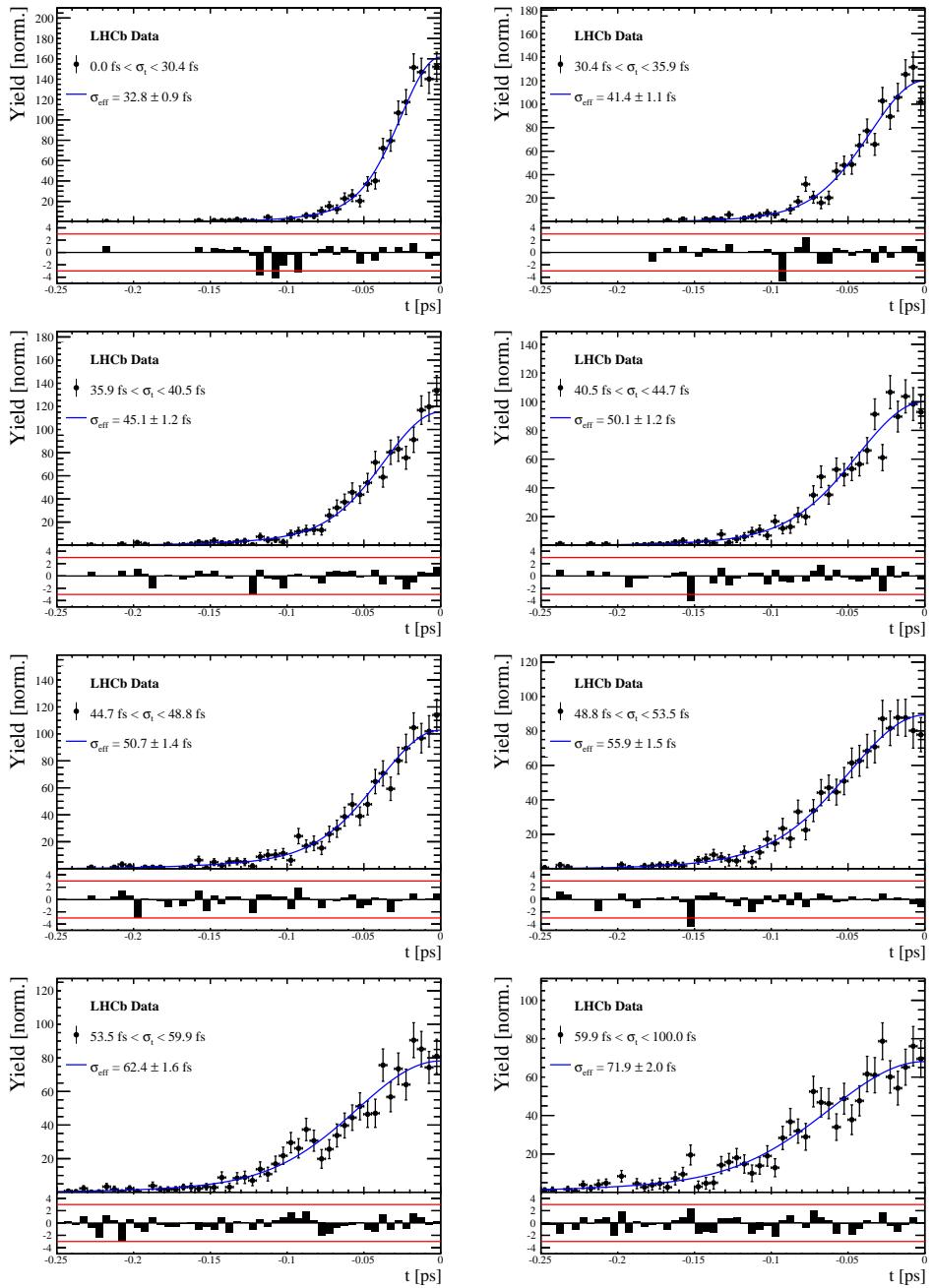


Figure C.2: Decay-time distribution for fake  $B_s$  candidates from promptly produced  $D_s$  candidates, combined with random prompt  $K\pi\pi$  bachelor tracks, for bins in the per-event decay time error estimate.

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

Table 3.2: Measured time resolution for prompt- $D_s$  data in bins of the per-event decay time error estimate.

## 781 D Spin Amplitudes

782 The spin factors used for  $B \rightarrow P_1 P_2 P_3 P_4$  decays are given in Table 4.1.

Table 4.1: 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	$B[D] \rightarrow (T_1 T_2)$ , $T_1 \rightarrow (P_1 P_2)$ , $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\gamma}(T_1) L_{(2)\gamma}^\beta(T_2)$

## 783 E Considered Decay Chains

784 The various decay channels considered in the model building are listed in Table 5.1.

Table 5.1: Decays considered in the LASSO model building.

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

785 F MC corrections

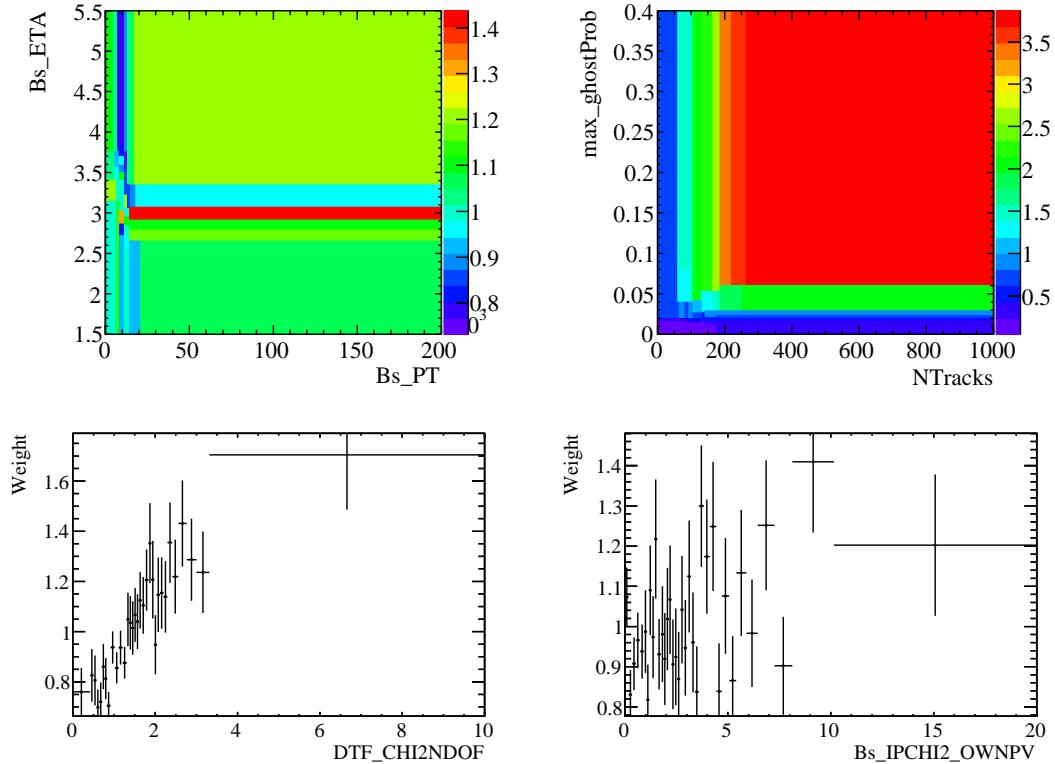


Figure C.1: Weights applied to correct for Data/MC differences.

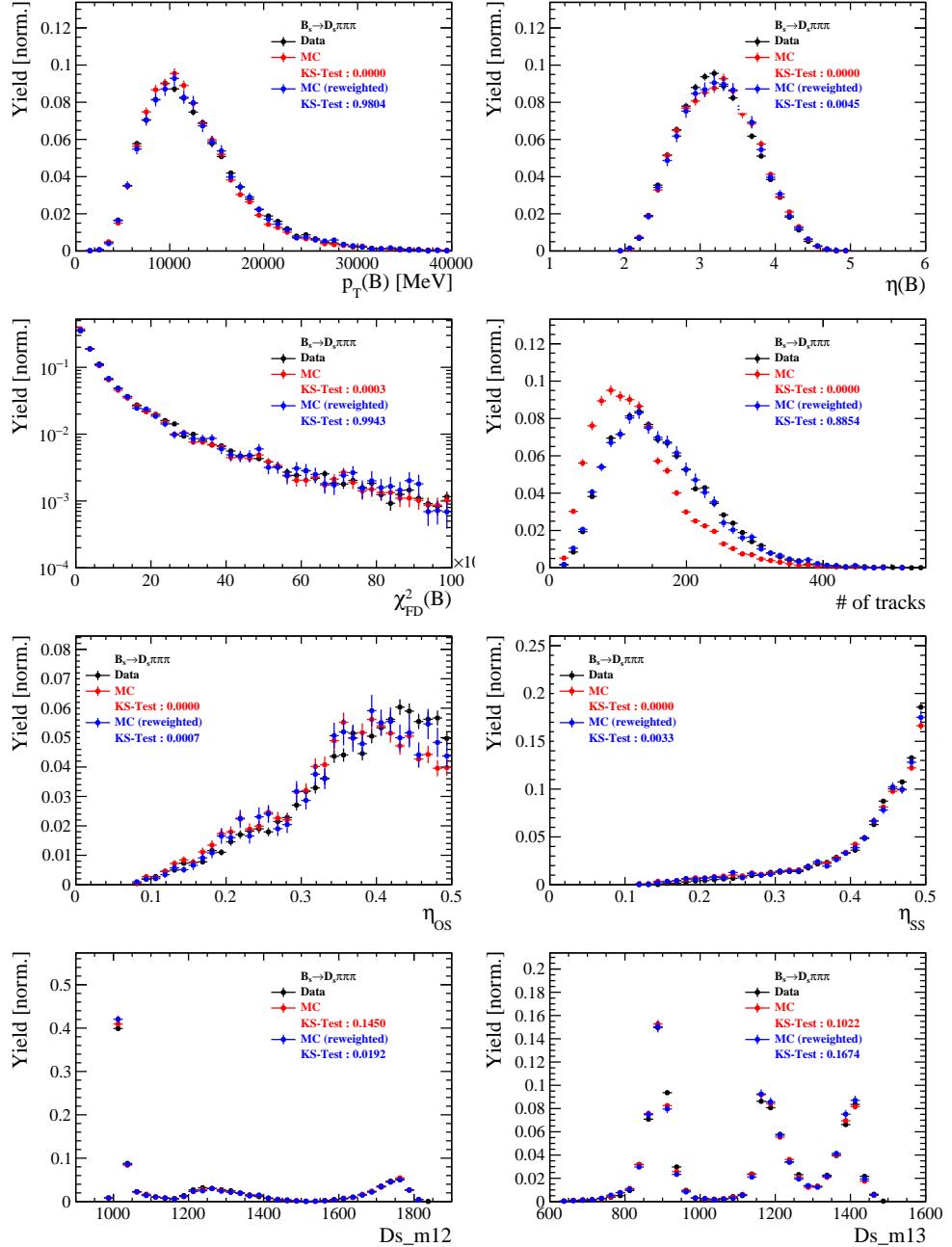


Figure C.2: Comparison of selected variables for  $B_s \rightarrow D_s \pi\pi\pi$  decays.

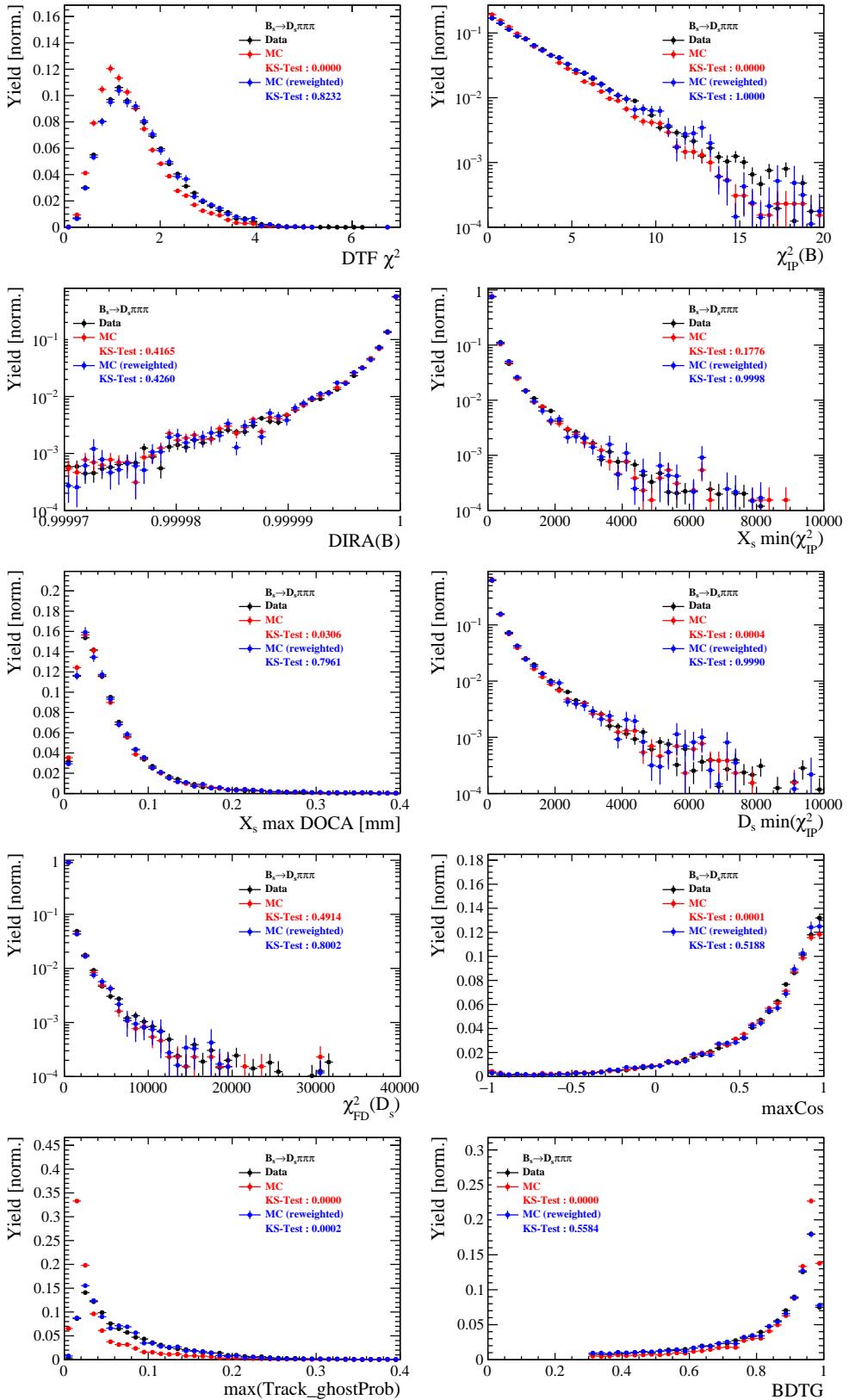


Figure C.3: Comparison of BDTG input variables and classifier response for  $B_s \rightarrow D_s \pi\pi\pi$  decays.

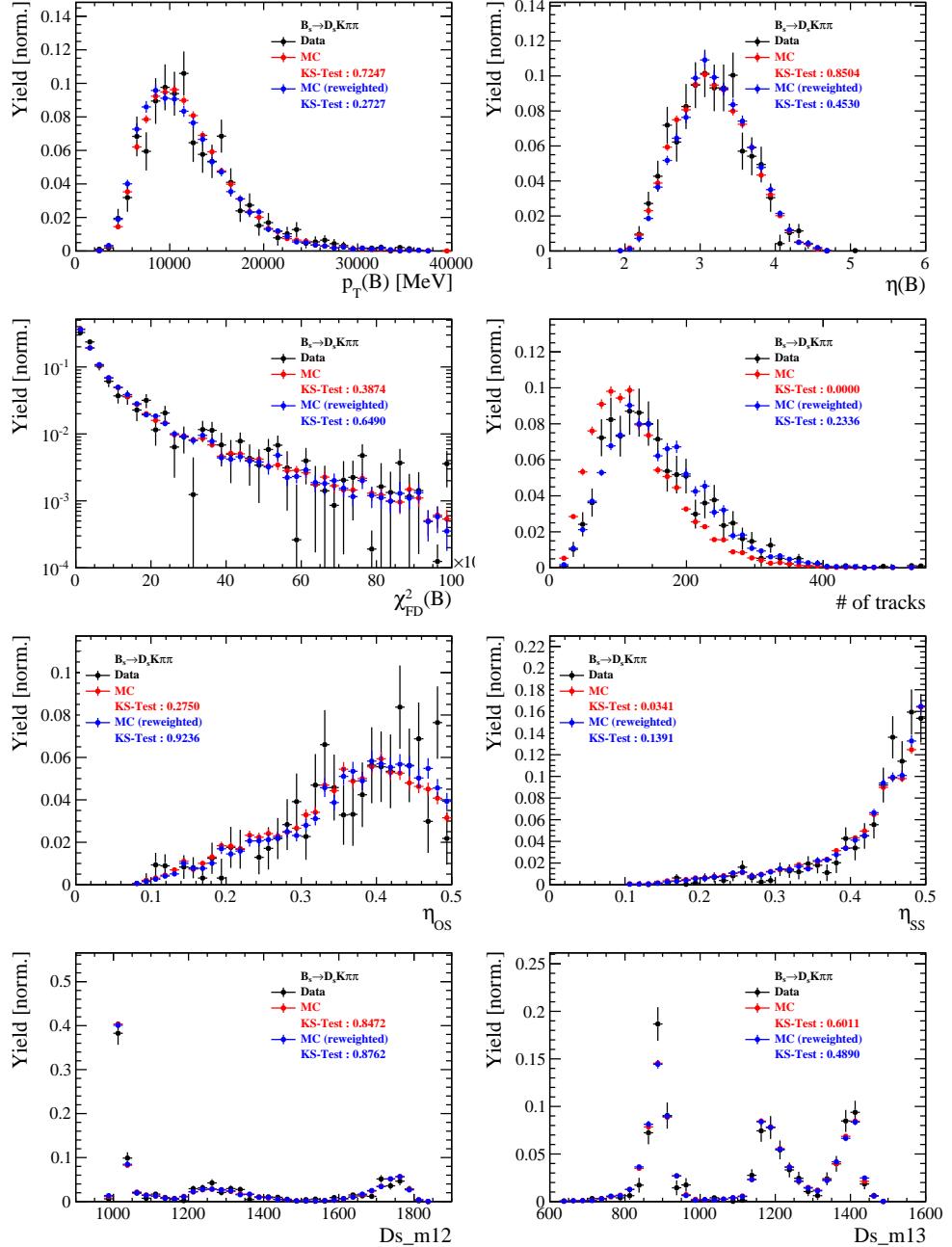


Figure C.4: Comparison of selected variables for  $B_s \rightarrow D_s K\pi\pi$  decays.

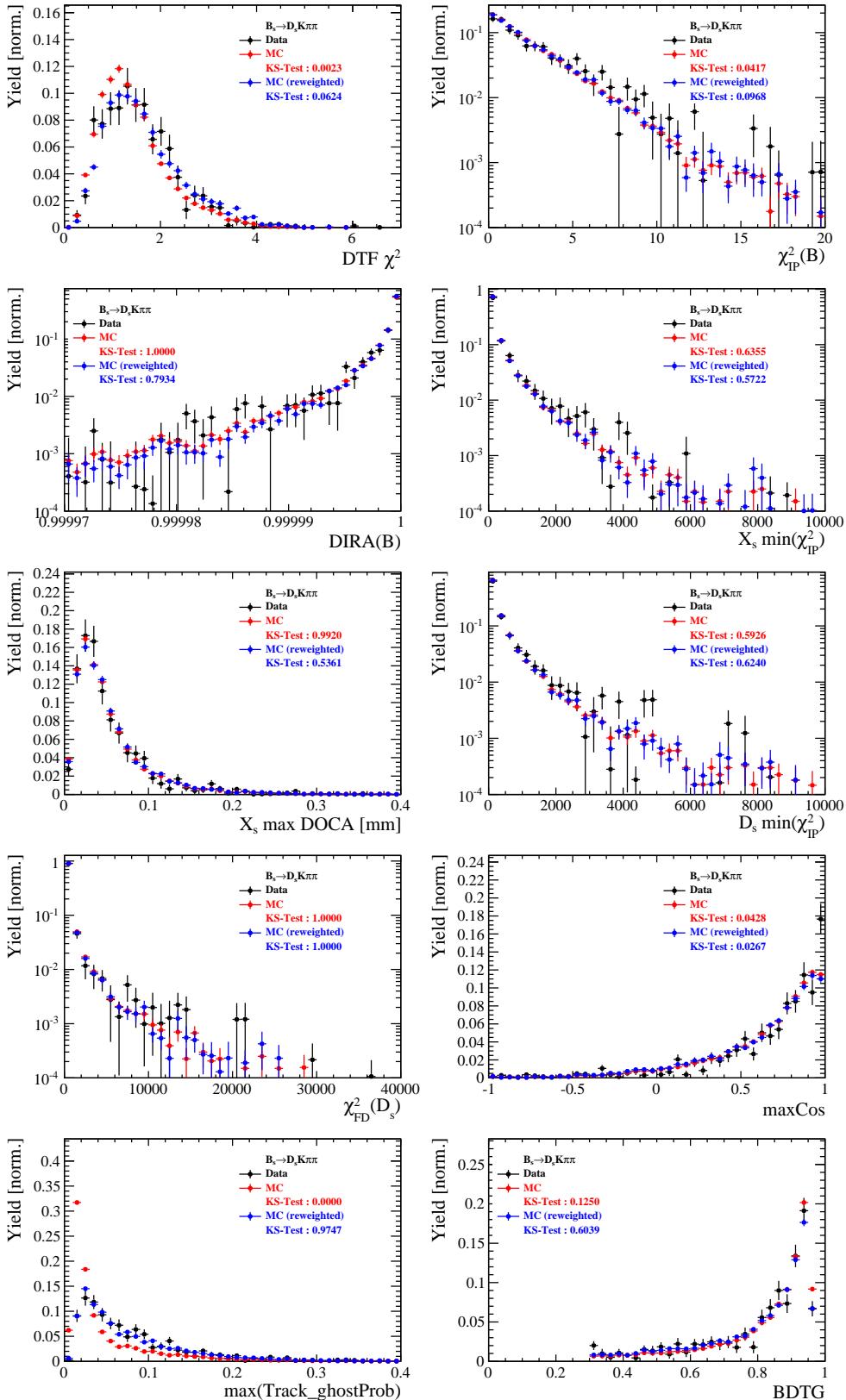


Figure C.5: Comparison of BDTG input variables and classifier response for  $B_s \rightarrow D_s K\pi\pi$  decays.

786 G Data distributions

787 G.1 Comparison of signal and calibration channel

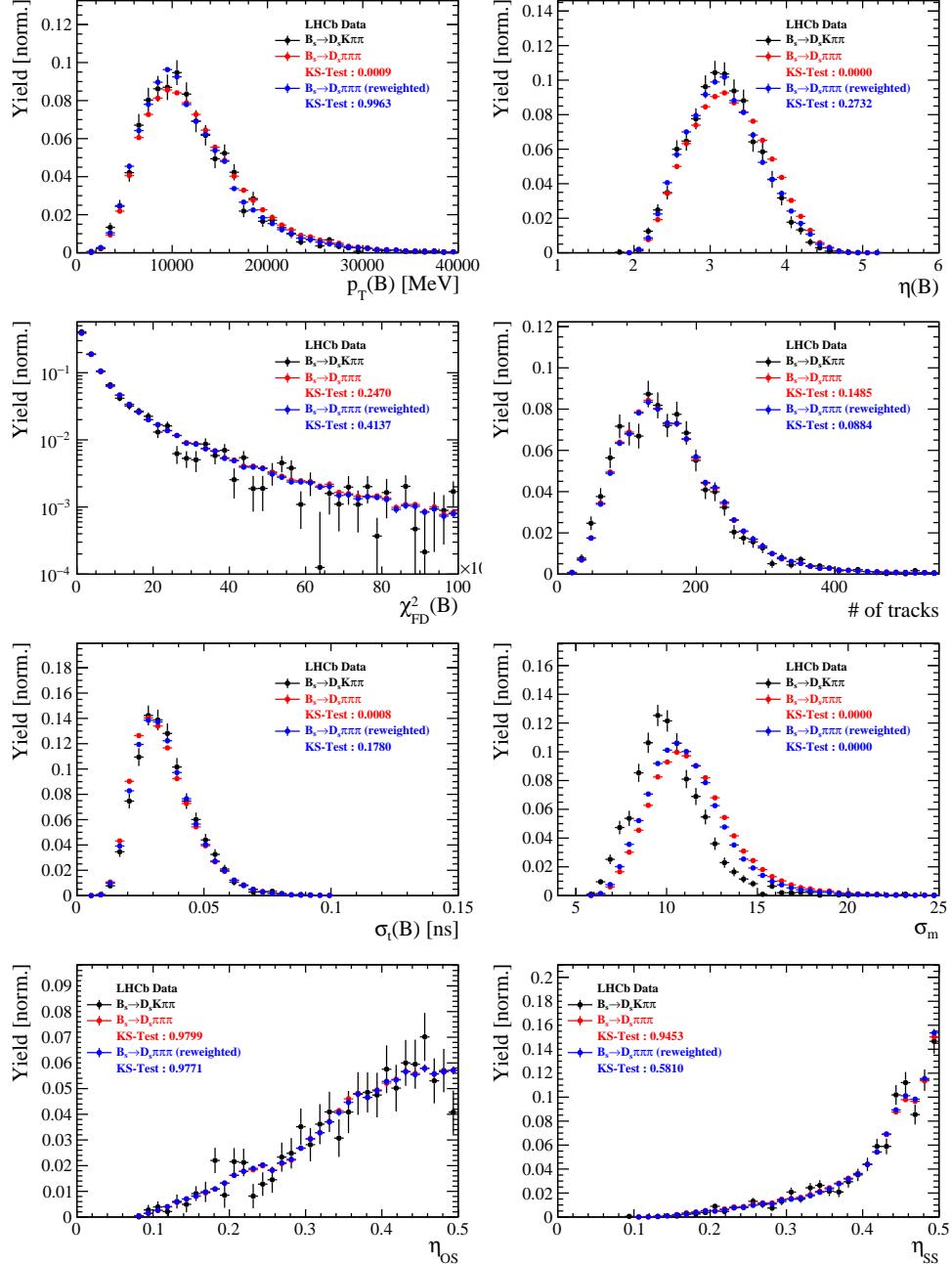


Figure C.1: Comparison of selected variables.

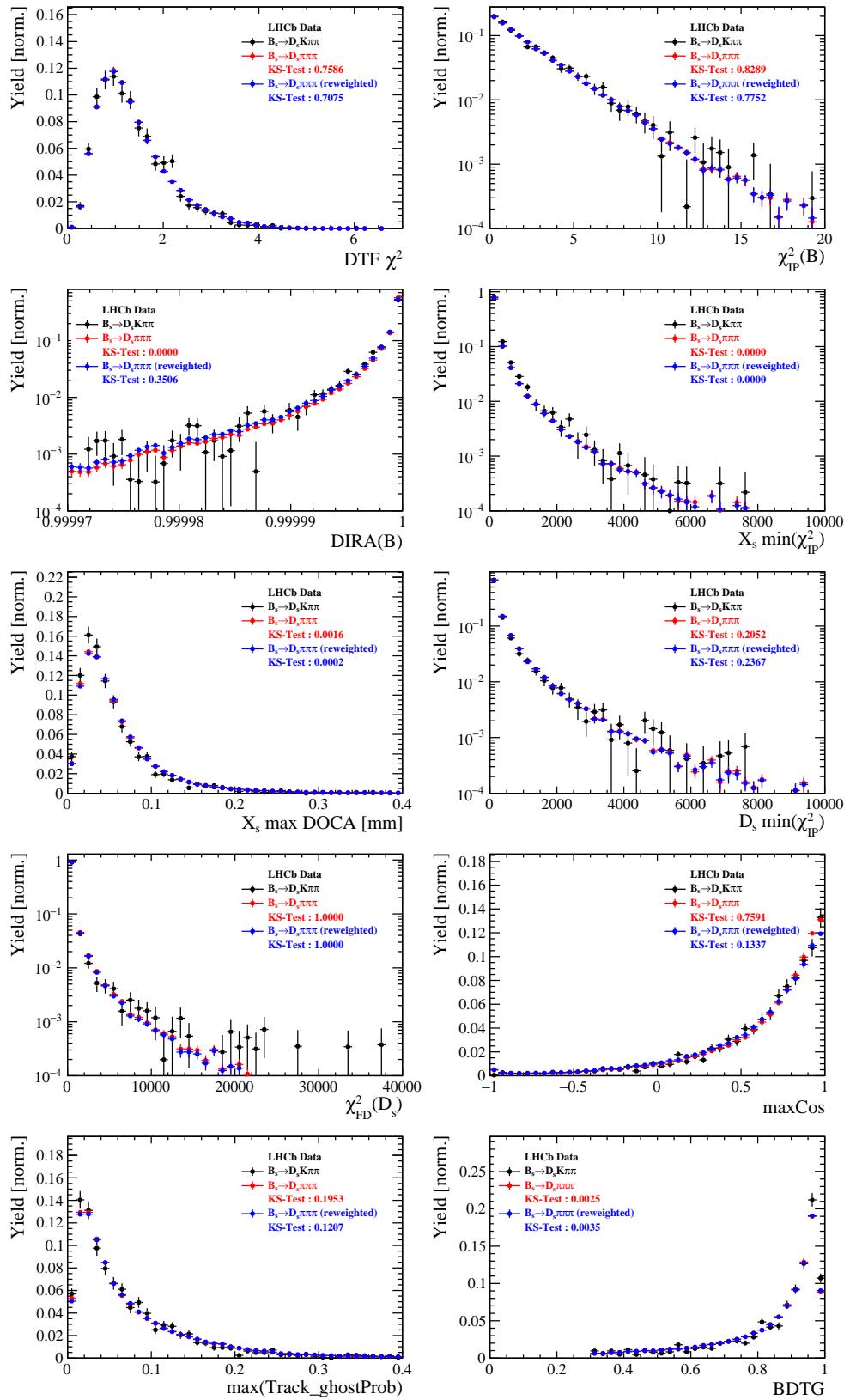


Figure C.2: Comparison of BDTG input variables and classifier response.

788 G.2 Comparison of Run-I and Run-II data

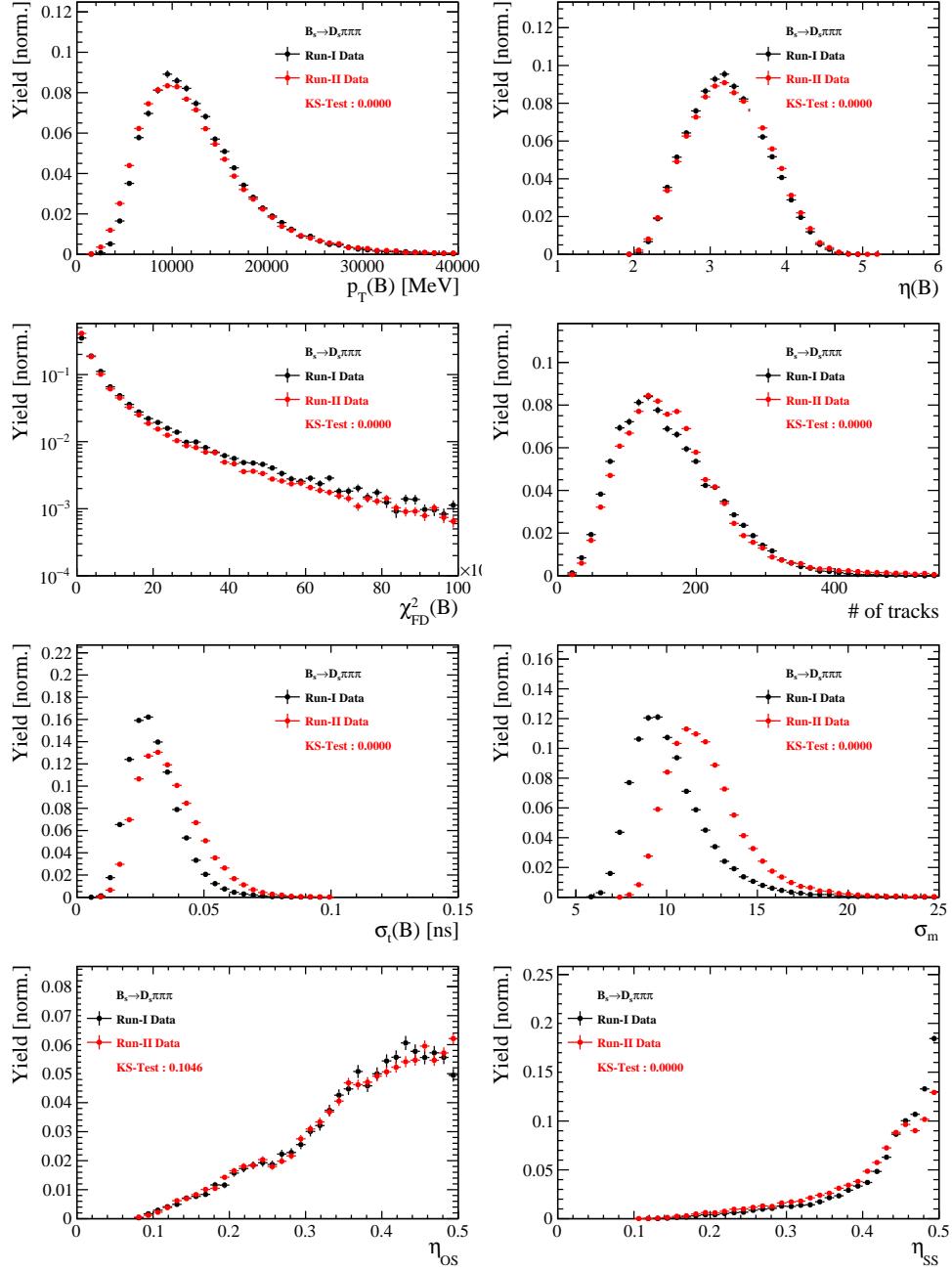


Figure C.3: Comparison of selected variables.

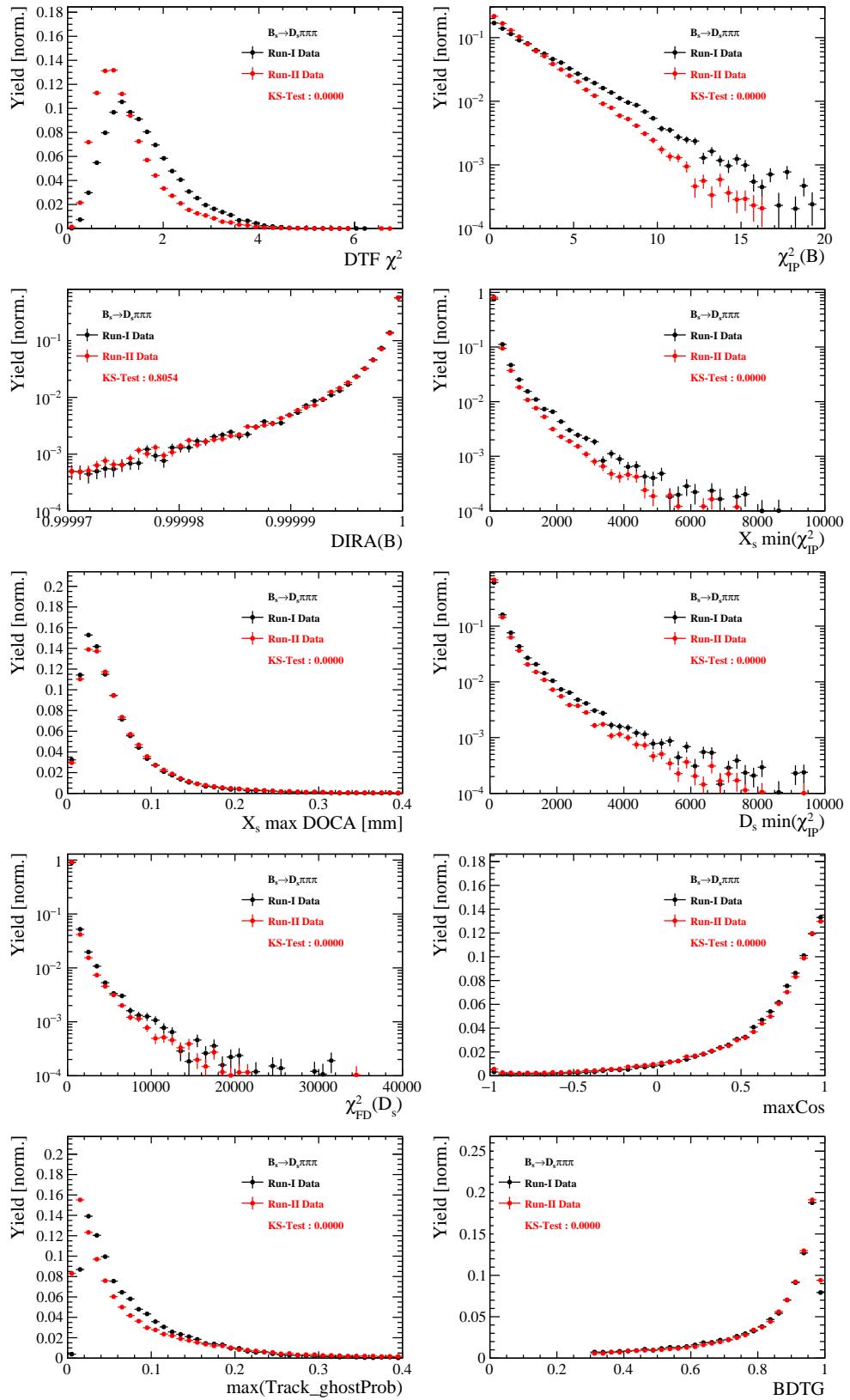


Figure C.4: Comparison of BDTG input variables and classifier response.

789 G.3 Comparison of  $D_s$  final states

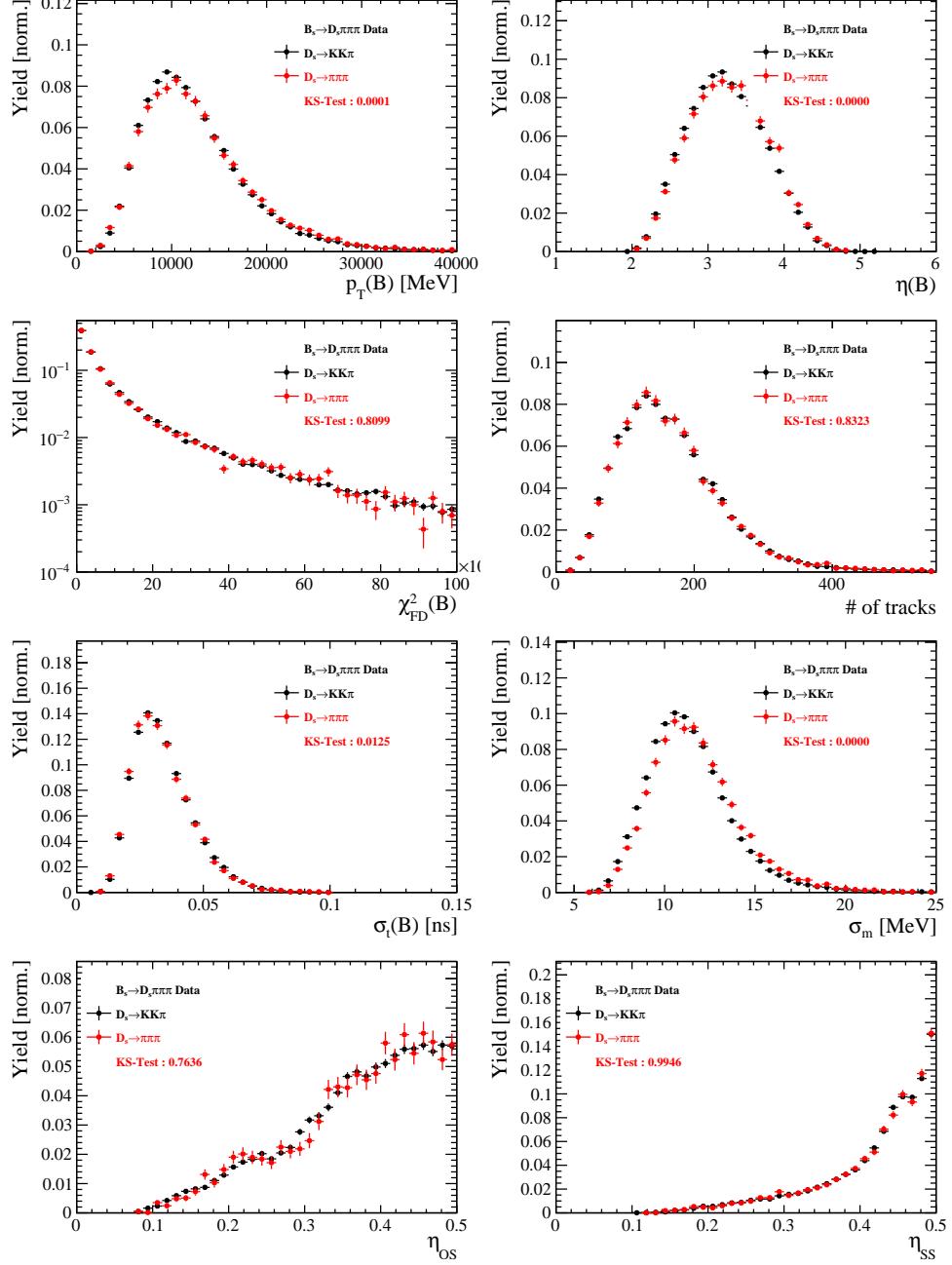


Figure C.5: Comparison of selected variables.

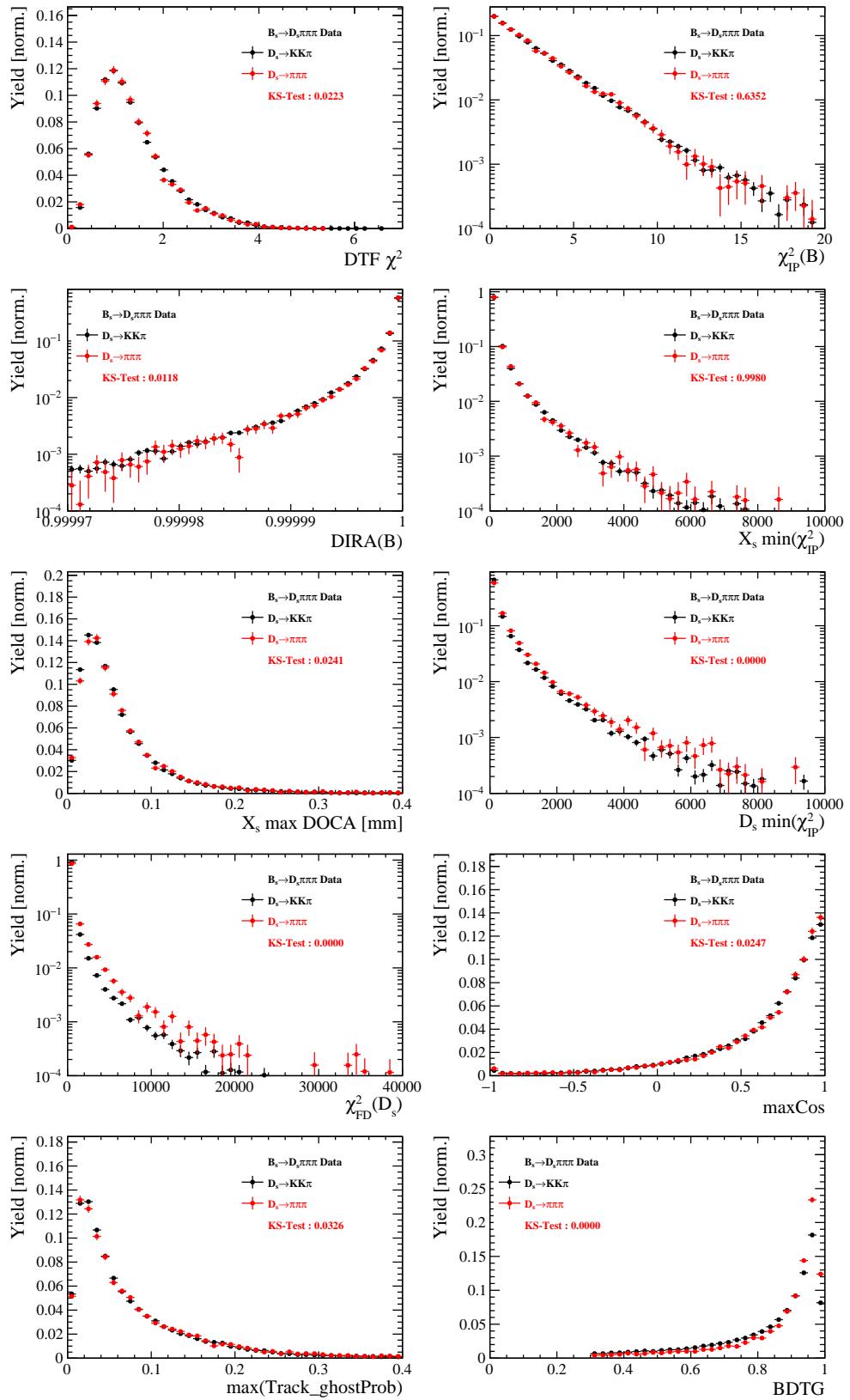


Figure C.6: Comparison of BDTG input variables and classifier response.

790 G.4 Comparison of trigger categories

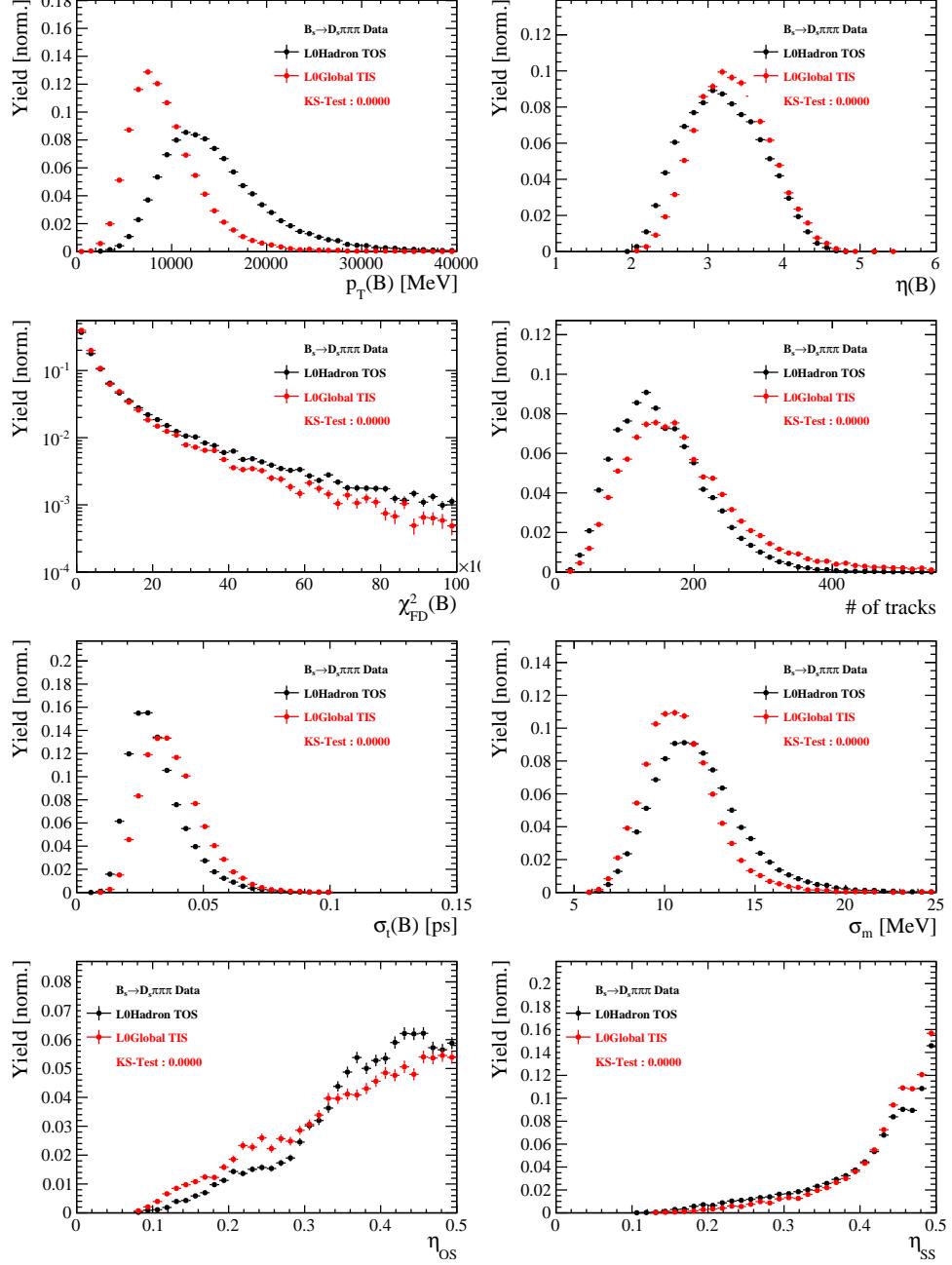


Figure C.7: Comparison of selected variables.

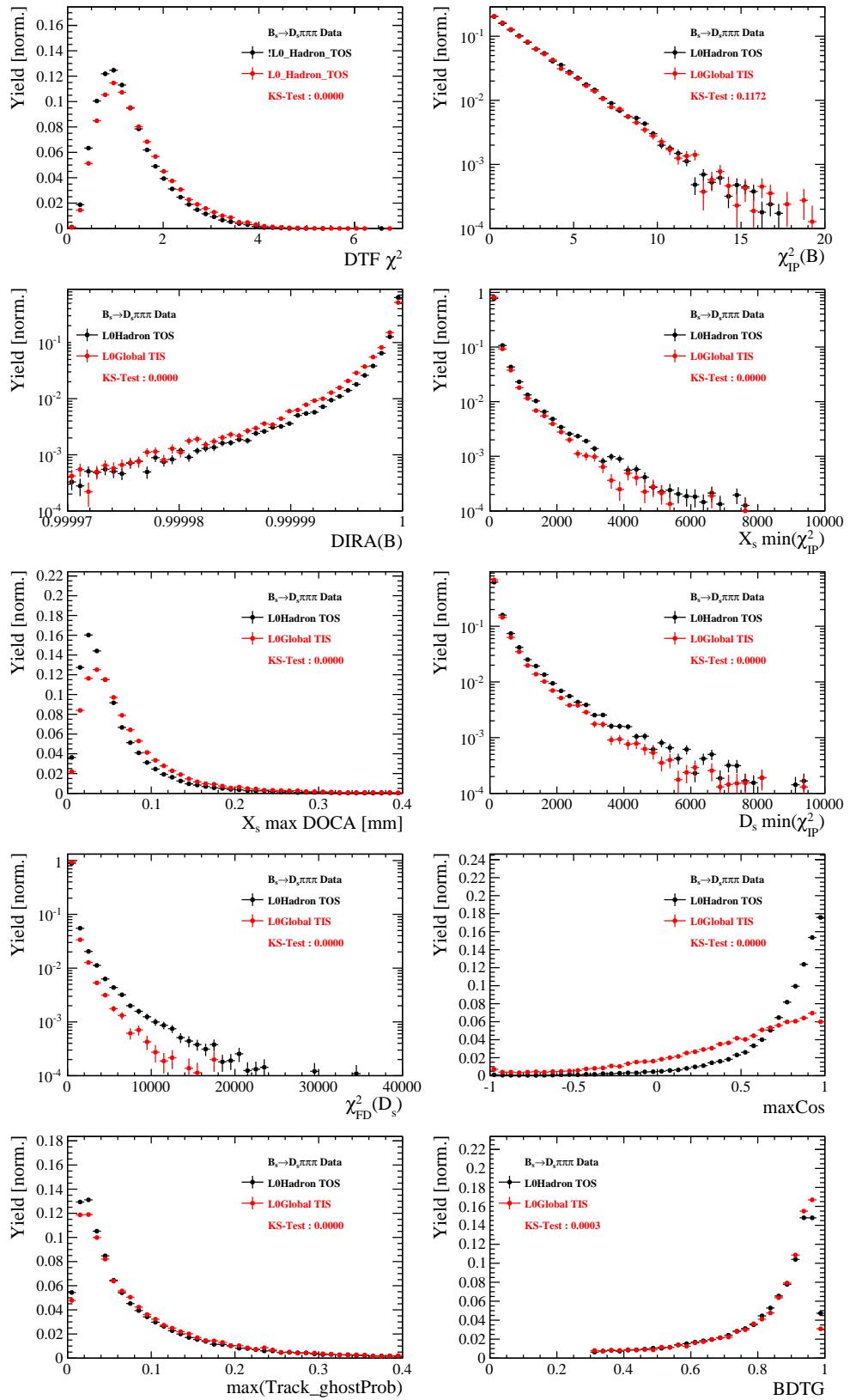


Figure C.8: Comparison of BDTG input variables and classifier response.

791 G.5 Comparison of  $B_s$  and  $B_d$  decays

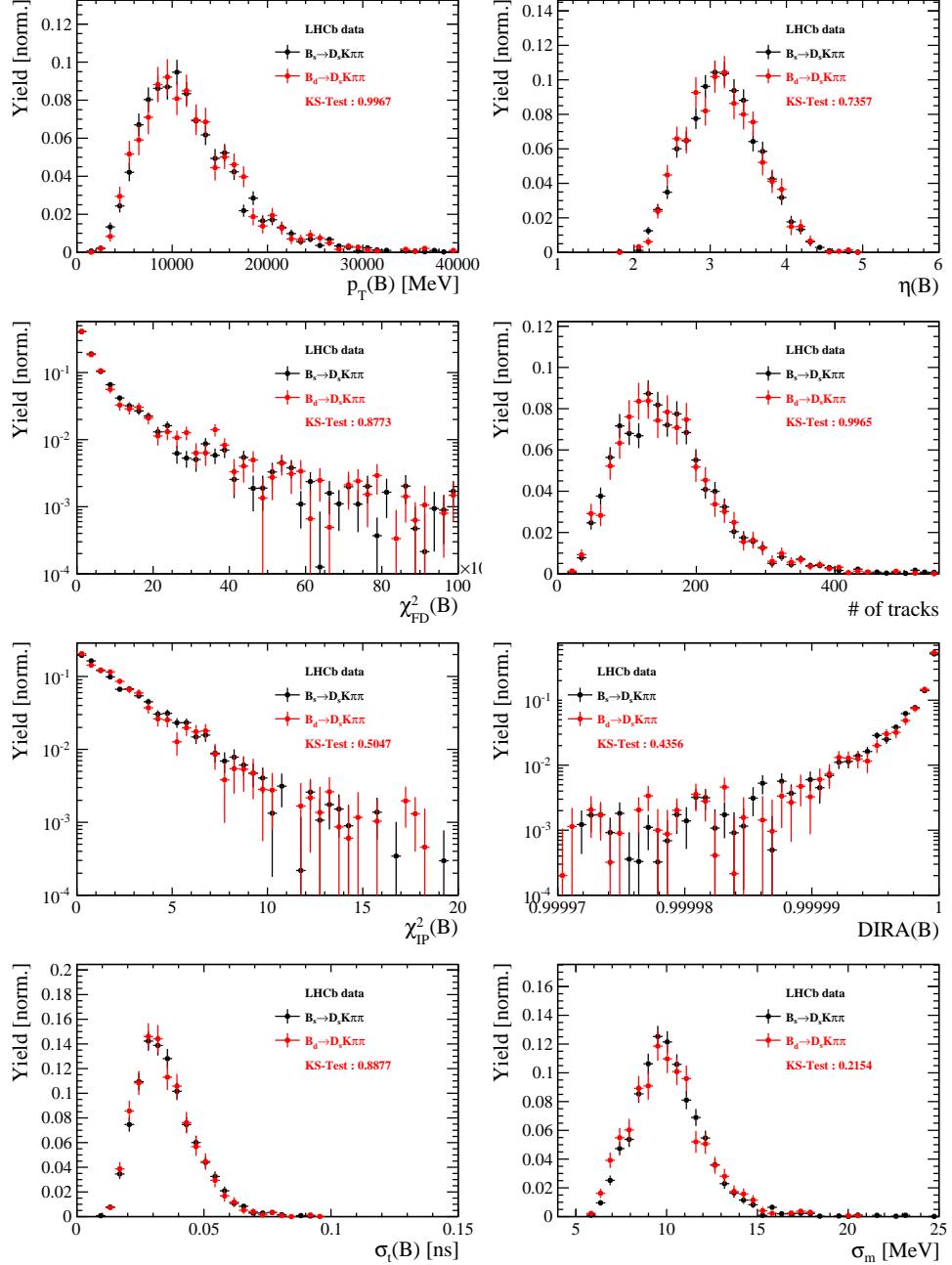


Figure C.9: Comparison of selected variables.

## 792 References

- 793 [1] R. Fleischer, *New strategies to obtain insights into CP violation through  $B(s) \rightarrow D(s) \rightarrow K \pi$ ,  $D(s)^* \rightarrow K \pi$ , ... and  $B(d) \rightarrow D \pi$ ,  $D^* \pi$ , ... decays*, Nucl.  
794 Phys. **B671** (2003) 459, arXiv:hep-ph/0304027.
- 795
- 796 [2] K. De Bruyn *et al.*, *Exploring  $B_s \rightarrow D_s^{(*)\pm} K^\mp$  Decays in the Presence of a Sizable*  
797 *Width Difference  $\Delta\Gamma_s$* , Nucl. Phys. **B868** (2013) 351, arXiv:1208.6463.
- 798 [3] S. Blusk, *First observations and measurements of the branching fractions for the*  
799 *decays  $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$  and  $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* .
- 800 [4] LHCb, S. Blusk, *Measurement of the CP observables in  $\bar{B}_s^0 \rightarrow D_s^+ K^-$  and first obser-*  
801 *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.
- 802 [5] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory* (*Frontiers in Physics*), Westview Press, 1995.
- 803
- 804 [6] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 805 [7] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation*  
806 *of production processes*, Annals of Physics **18** (1962), no. 2 198 .
- 807 [8] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*,  
808 Phys. Rev. D **11** (1975) 3165.
- 809 [9] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of  
810 Physics **108** (1977), no. 2 454 .
- 811 [10] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay*  
812 *widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.
- 813 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo  
814 Cimento Series 10 **34** (1964), no. 6 1644.
- 815 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys.  
816 **C40** (2016), no. 10 100001.
- 817 [13] D. V. Bugg, *The mass of the  $\sigma$  pole*, Journal of Physics G Nuclear Physics **34** (2007)  
818 151, arXiv:hep-ph/0608081.
- 819 [14] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-*  
820 *dominance prediction for  $\rho \rightarrow e^+ e^-$* , Phys. Rev. Lett. **21** (1968) 244.
- 821 [15] S. M. Flatté, *Coupled-channel analysis of the  $\pi\eta$  and  $KK$  systems near  $KK$  threshold*,  
822 Physics Letters B **63** (1976), no. 2 224 .
- 823 [16] BES Collaboration, M. Ablikim *et al.*, *Resonances in  $J/\psi \rightarrow \phi\pi^+\pi^-$  and  $\phi K^+K^-$* ,  
824 Phys. Lett. **B607** (2005) 243, arXiv:hep-ex/0411001.
- 825 [17] D. V. Bugg, *A study in depth of  $f_0(1370)$* , Eur. Phys. J. **C52** (2007) 55,  
826 arXiv:0706.1341.

- [827] [18] LHCb Collaboration, R. Aaij *et al.*, *Analysis of the resonant components in  $B_s \rightarrow J/\psi \pi^+ \pi^-$* , Phys. Rev. **D86** (2012) 052006, [arXiv:1204.5643](#).
- [829] [19] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- [830] [20] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys. Rev. **60** (1941) 61.
- [832] [21] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys. Rev. D **57** (1998) 431.
- [834] [22] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of  $\psi$  decay to mesons*, Eur. Phys. J. **A16** (2003) 537, [arXiv:hep-ph/0211457](#).
- [836] [23] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. **D51** (1995) 2247.
- [838] [24] J.-J. Zhu, *Explicit expressions of spin wave functions*, [arXiv:hep-ph/9906250](#).
- [839] [25] P. d'Argent *et al.*, *Amplitude Analyses of  $D^0 \rightarrow \pi^+ \pi^- \pi^+ \pi^-$  and  $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$  Decays*, JHEP **05** (2017) 143, [arXiv:1703.08505](#).
- [841] [26] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput. Phys. Commun. **180** (2009) 1847, [arXiv:0805.2956](#).
- [843] [27] LHCb, R. Aaij *et al.*, *Studies of the resonance structure in  $D^0 \rightarrow K^\mp \pi^\pm \pi^\pm \pi^\mp$  decays*, Submitted to: Eur. Phys. J. C (2017) [arXiv:1712.08609](#).
- [845] [28] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [847] [29] *Gammacombo package*, 2014.
- [848] [30] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, [arXiv:physics/0703039](#).
- [850] [31] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](#).
- [852] [32] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.
- [854] [33] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [856] [34] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of  $B_s^0$  mesons at LHCb*, JINST **11** (2016), no. 05 P05010, [arXiv:1602.07252](#).
- [858] [35] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of  $B$  mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, [arXiv:1202.4979](#).

- [36] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and  $\tau$ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots available at <http://www.slac.stanford.edu/xorg/hfag/>.
- [37] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, arXiv:1407.0748.
- [38] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.
- [39] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in  $B_s^0 \rightarrow D_s^\mp K^\pm$  decays*, Submitted to: JHEP (2017) arXiv:1712.07428.
- [40] LHCb, R. Aaij *et al.*, *Measurement of  $B^0$ ,  $B_s^0$ ,  $B^+$  and  $\Lambda_b^0$  production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- [41] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of the  $K\pi$  Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-2012-027, CERN, Geneva, Feb, 2013.
- [42] A. Davis *et al.*, *Measurement of the instrumental asymmetry for  $K^- \pi^+$ -pairs at LHCb in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN, Geneva, Mar, 2018.
- [43] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly forbidden transitions in  $D \rightarrow K(S)$ ,  $K(L) + \pi$ 's decays*, Phys. Lett. **B349** (1995) 363, arXiv:hep-ph/9502238.
- [44] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- [45] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal Statistical Society, Series B **58** (1994) 267.
- [46] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- [47] H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on Automatic Control **19** (1974) 716.