

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

P. d'Argent¹, E. Gersabeck², M. Kecke¹, M. Schiller³

¹*Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany*

²*School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom*

³*School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom*

Abstract

We present the first measurement of the weak phase $2\beta + \gamma$ obtained from a time-dependent (amplitude) analysis of $B_s^0 \rightarrow D_s K \pi\pi$ decays using proton-proton collision data corresponding to an integrated luminosity of 5 fb^{-1} recorded by the LHCb detector.

Contents

1	Introduction	1
2	Formalism	2
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
3	Data samples and event selection	12
3.1	Stripping and Trigger selection	12
3.2	Offline selection	12
3.2.1	Physics background vetoes	13
3.2.2	Phase space region	14
3.2.3	Training of multivariate classifier	15
3.2.4	Final selection	15
3.3	Simulation	16
4	Fits to invariant mass distributions of signal and normalization channel	18
4.1	Signal models for $m(D_s\pi\pi\pi)$ and $m(D_sK\pi\pi)$	18
4.2	Background models for $m(D_s\pi\pi\pi)$	19
4.3	Background models for $m(D_sK\pi\pi)$	19
4.4	Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates	21
4.5	Fit to $B_s^0 \rightarrow D_sK\pi\pi$ candidates	21
4.6	Extraction of signal weights	21
5	Flavour Tagging	23
5.1	OS tagging calibration	24
5.2	SS tagging calibration	24
5.3	Tagging performance comparison between the signal and normalization channel	24
5.4	Combination of OS and SS taggers	25
6	Acceptance	28
6.1	MC corrections	28
6.1.1	Truth matching of simulated candidates	28
6.1.2	PID efficiencies	29
6.1.3	BDT efficiencies	30
6.1.4	Tracking efficiencies	31
6.2	Decay-time acceptance	32
6.2.1	Comparison of acceptance in subsamples	33
6.2.2	Results	35
6.3	Phasespace acceptance	39

7 Decay-time Resolution	40
7.1 Calibration for Run-I data	41
7.2 Calibration for Run-II data	42
7.3 Cross-checks	45
7.3.1 Kinematic dependence	45
7.3.2 DTF constraints	45
8 Production and Detection Asymmetries	46
8.1 B_s Production Asymmetry	46
8.2 $K^-\pi^+$ Detection Asymmetry	47
9 Time dependent fit	50
9.1 sFit to $B_s^0 \rightarrow D_s \pi\pi\pi$ data	50
9.2 sFit to $B_s^0 \rightarrow D_s K\pi\pi$ data	53
10 Time dependent amplitude fit	54
10.1 Signal Model Construction	54
10.2 Results	55
A Stripping and Trigger cuts	58
B Details of multivariate classifier	60
C Detailed mass fits	61
D Decay-time Resolution fits	65
E Spin Amplitudes	69
F Considered Decay Chains	70
G MC corrections	71
H Data distributions	76
H.1 Comparison of signal and calibration channel	76
H.2 Comparison of Run-I and Run-II data	78
H.3 Comparison of D_s final states	80
H.4 Comparison of trigger categories	82
H.5 Comparison of B_s and B_d decays	84
References	85

1 Introduction

The weak phase γ is the least well known angle of the CKM unitary triangle. A key channel to measure γ is the time-dependent analysis of $B_s^0 \rightarrow D_s K$ decays [1, 2]. To measure the weak CKM phase $\gamma \equiv \arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$, a decay with interference between $b \rightarrow c$ and $b \rightarrow u$ transitions is needed [1]. This note present the first measurement of γ using $B_s^0 \rightarrow D_s K\pi\pi$ decays, where the $K\pi\pi$ subsystem is dominated by excited kaon states such as the $K_1(1270)$ and $K_1(1400)$ resonances. To account for the non-constant strong phase across the phasespace, one can either perform a time-dependent amplitude fit or select a suitable phase-space region and introduce a coherence factor as additional hadronic parameter to the fit. This analysis is based on the first observation of the $B_s^0 \rightarrow D_s K\pi\pi$ decay by LHCb [3, 4], where the branching ratio is measured relative to $B_s^0 \rightarrow D_s \pi\pi\pi$.

12 2 Formalism

13 2.1 Decay rates and CP-observables

14 In the following, we choose a convention in which $\Delta\Gamma_s = \Gamma_L - \Gamma_H < 0$ and $\Delta m_s =$
 15 $m_H - m_L > 0$, where the indices H and L refer to the heavy and light mass eigenstates
 16 of the B_s meson. We assume $|q/p| = 1$ for the complex coefficients p and q which relate
 17 the B_s meson mass eigenstates to the flavour eigenstates.

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

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

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

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

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

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

19 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
 20 $D_s^- K^+ \pi\pi$ ($D_s^+ K^- \pi\pi$) final states. Integrating over the phasespace, we get

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

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

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

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

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

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

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

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

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

²⁵

²⁶ 2.2 Amplitude model

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

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

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

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

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

⁴² In practice, we do not need to choose a particular five-dimensional basis, but use the
⁴³ full four-vectors of the decay in our analysis. The dimensionality is handled by the phase
⁴⁴ space element which can be written in terms of any set of five independent kinematic
⁴⁵ observables, $\mathbf{x} = (x_1, \dots, x_5)$, as

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

⁴⁶ where $\phi_4(\mathbf{x}) = \left| \frac{\partial \Phi_4}{\partial(x_1, \dots, x_5)} \right|$ is the phase space density. In contrast to three-body decays,
⁴⁷ the four-body phase space density function is not flat in the usual kinematic variables.
⁴⁸ Therefore, an analytic expression for ϕ_4 is taken from Ref. [6].

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

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

⁵² To construct $A_i(\mathbf{x})$, the isobar approach is used, which assumes that the decay process
⁵³ can be factorized into subsequent two-body decay amplitudes [7–9]. This gives rise to
⁵⁴ two different decay topologies; quasi two-body decays $B_s \rightarrow (R_1 \rightarrow h_1 h_2) (R_2 \rightarrow h_3 h_4)$

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

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

59 2.2.1 Form Factors and Resonance Lineshapes

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

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

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

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

67 featuring the energy-dependent mass $M(s)$ (defined below), and total width, $\Gamma(s)$. The
 68 latter is normalized to give the nominal width, Γ_0 , when evaluated at the nominal mass
 69 m_0 , *i.e.* $\Gamma_0 = \Gamma(s = m_0^2)$.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

106 Some particles may not originate from a resonance but are in a state of relative orbital
 107 angular momentum. We denote such non-resonant states by surrounding the particle
 108 system with brackets and indicate the partial wave state with an subscript; for example
 109 $(\pi\pi)_S$ refers to a non-resonant di-pion S -wave. The lineshape for non-resonant states is
 110 set to unity.

111 2.2.2 Spin Densities

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

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

122 to $2S + 1$ independent elements in accordance with the number of degrees of freedom of a
123 spin- S state [20, 24].

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

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

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

128 For a decay process $R \rightarrow AB$, with relative orbital angular momentum L , between
129 particle A and B , the angular momentum tensor is obtained by projecting the rank- L
130 tensor $q_R^{\nu_1} q_R^{\nu_2} \dots q_R^{\nu_L}$, constructed from the relative momenta $q_R = p_A - p_B$, onto the spin- L
131 subspace,

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

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

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

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

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

138 where

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

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

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

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

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

¹⁵² 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
¹⁵³ pseudoscalar initial/final states are involved.

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

156 2.3 Validation

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

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

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

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

τ	1.5 ps
$\Delta\Gamma$	-0.1 ps^{-1}
Δm_s	17.757 ps^{-1}
r	0.37
κ	1
δ	10.0°
γ	71.1°
β_s	0.0°

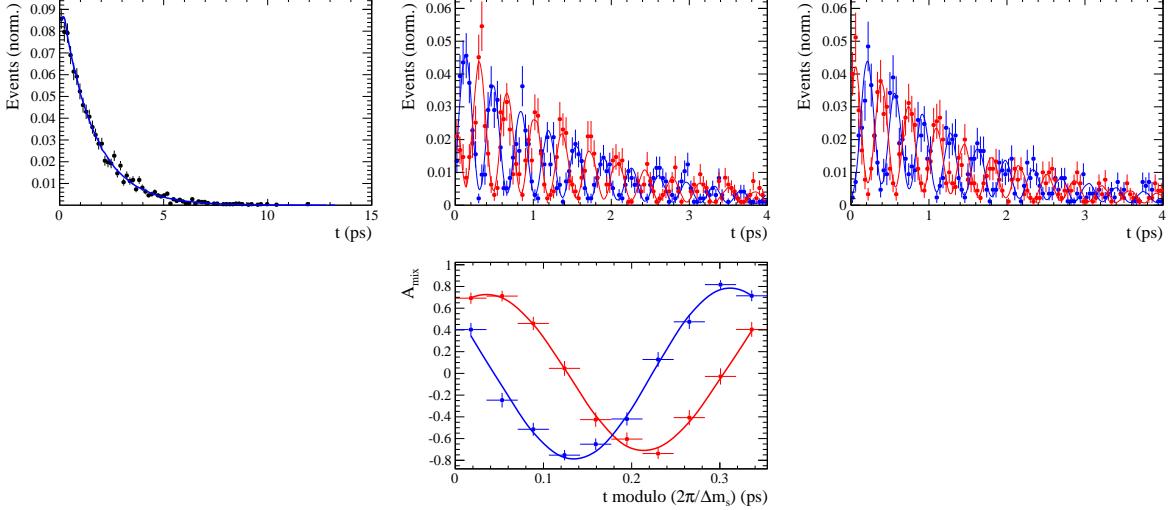


Figure 2.1: Time distribution of $B_s \rightarrow D_s K \pi\pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

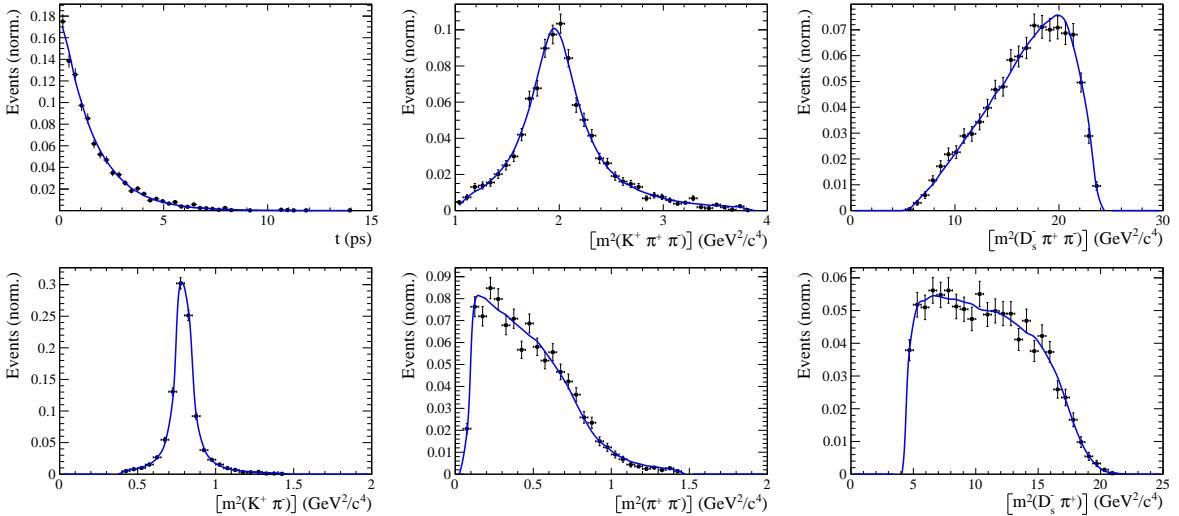


Figure 2.2: Time and invariant mass distributions of $B_s \rightarrow D_s K \pi\pi$ events generated with EVTGEN (points with error bars) and MINT2 fit projections (blue solid line).

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

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

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

	Generated	Fit result	Pull(σ)
x_-	0.179	0.135 ± 0.050	-0.9
y_-	-0.324	-0.307 ± 0.022	0.8
x_+	0.057	0.102 ± 0.065	0.6
y_+	0.366	0.394 ± 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 ± 0.021	0.379 ± 0.017
κ	1.0	1.0	1.000 ± 0.059
δ	10.0°	9.0 ± 5.1	5.9 ± 6.0
γ	71.1°	67.3 ± 5.9	75.1 ± 6.9

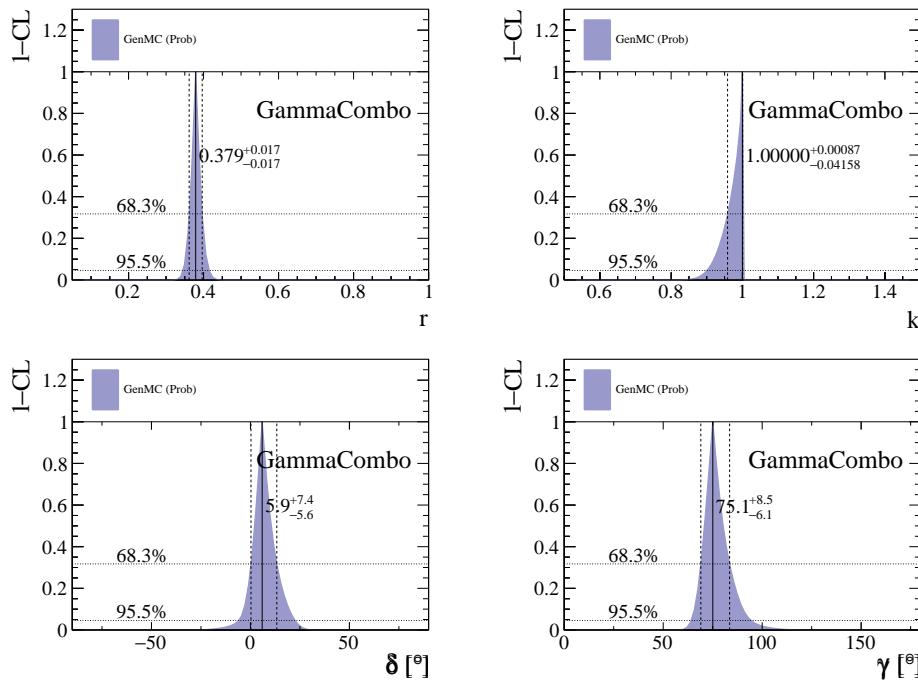


Figure 2.3: The 1-CL contours for the physical observable $r, \kappa, \delta, \gamma$ obtained with the phasespace integrated fit to the EVTGEN toy sample.

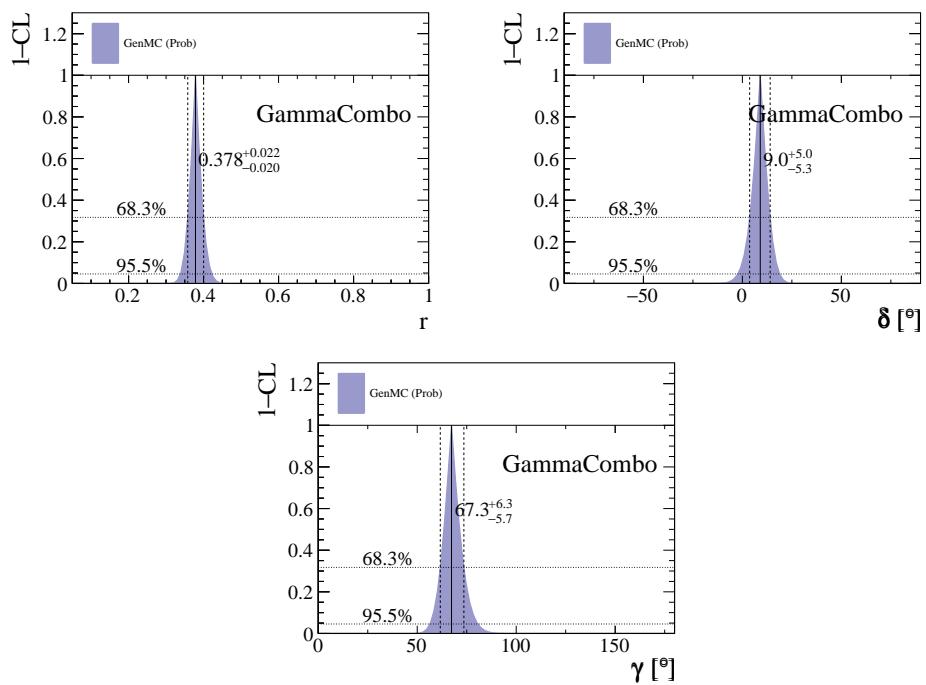


Figure 2.4: The 1-CL contours for the physical observable r, δ, γ obtained with the time-dependent amplitude fit fit to the **EVTGEN** toy sample.

180 3 Data samples and event selection

181 3.1 Stripping and Trigger selection

182 The dataset used for this analysis corresponds to 1 fb^{-1} of proton-proton collision data
183 collected in 2011 with a centre of mass energy $\sqrt{s} = 7 \text{ TeV}$, 2 fb^{-1} collected in 2012 with
184 $\sqrt{s} = 7 \text{ TeV}$ and 2 fb^{-1} collected in 2015/2016 with $\sqrt{s} = 13 \text{ TeV}$. Candidate $B_s^0 \rightarrow D_s K\pi\pi$
185 ($B_s^0 \rightarrow D_s \pi\pi\pi$) decays are reconstructed using the `B02DKPiPiD2HHHPIDBeauty2CharmLine`
186 (`B02DPiPiPiD2HHHPIDBeauty2CharmLine`) line of the `BHadronCompleteEvent` stream of
187 `Stripping21r1` (2011), `Stripping21` (2012), `Stripping24r1` (2015) and `Stripping28r1p1` (2016).
188 Both stripping lines employ the same selection cuts, listed in Appendix A, except for the
189 PID requirement on the bachelor kaon/pion.

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

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

202 3.2 Offline selection

203 The offline selection, in particular the requirements on the D_s hadron, are guided by
204 the previous analyses of $B_s \rightarrow D_s K/\pi$, $B_d \rightarrow D^0 \pi$ as well as the branching fraction
205 measurement of $B_s^0 \rightarrow D_s K\pi\pi$ decays.

206 In order to clean up the sample and to align the Run-I to the slightly tighter Run-2
207 stripping selection, we apply the following loose cuts to the b-hadron:

- 208 • DIRA > 0.99994
- 209 • min IP $\chi^2 < 20$ to any PV,
- 210 • FD $\chi^2 > 100$ to any PV,
- 211 • Vertex $\chi^2/\text{nDoF} < 8$.

212 We reconstruct the $B_s^0 \rightarrow D_s h\pi\pi$ decay through two different final states of the D_s meson,
213 $D_s \rightarrow KK\pi$ and $D_s \rightarrow \pi\pi\pi$. Of those, $D_s \rightarrow KK\pi$ is the most prominent one, while
214 $\mathcal{BR}(D_s \rightarrow \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$ holds for the other. For the $KK\pi$ final state we
215 make use of the well known resonance structure; the decay proceeds either via the narrow
216 ϕ resonance, the broader K^{*0} resonance or (predominantly) non-resonant. Within the ϕ
217 resonance region the sample is already sufficiently clean after the stripping so that we do
218 not impose additional criteria on the D_s daughters. For the K^{*0} and non-resonant regions
219 consecutively tighter requirements on the particle identification and the D_s flight-distance
220 are applied. We apply global requirements for the $D_s \rightarrow \pi\pi\pi$ final state.

221 **3.2.1 Physics background vetoes**

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

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

- 226 (a) $D^- \rightarrow K^+ \pi^- \pi^-$: possible with single missID of $\pi^- \rightarrow K^-$, vetoed by requiring
 227 $m(K^+ K_\pi^- \pi^-) \neq m(D^-) \pm 30$ MeV or the K^- has to fulfill more stringent PID
 228 criteria depending on the resonant region.
- 229 (b) $\Lambda_c^- \rightarrow K^+ \bar{p} \pi^-$: possible with single missID of $\bar{p} \rightarrow K^-$, vetoed by requiring
 230 $m(K^+ K_p^- \pi^-) \neq m(\Lambda_c) \pm 30$ MeV or the K^- has to fulfill more stringent PID
 231 criteria depending on the resonant region.
- 232 (c) $D^0 \rightarrow KK$: D^0 combined with a random π can fake a $D_s \rightarrow KK\pi$ decay,
 233 vetoed by requiring $m(KK) < 1840$ MeV.

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

- 235 (a) $D^- \rightarrow K^+ \pi^- \pi^-$:
- 236 (b) $D^0 \rightarrow \pi\pi$: combined with a random π can fake a $D_s \rightarrow \pi\pi\pi$ decay, vetoed by
 237 requiring both possible combinations to have $m(\pi\pi) < 1700$ MeV.

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

- 239 (a) $D^- \rightarrow \pi^- \pi^+ \pi^-$: $m(K_\pi^- \pi^+ \pi^-) \neq m(D^-) \pm 30$ MeV
- 240 (b) $\Lambda_c^- \rightarrow K^+ \bar{p} \pi^-$:
- 241 (c) $D^0 \rightarrow K\pi$: combined with a random π can fake a $D_s \rightarrow K\pi\pi$ decay, vetoed
 242 by requiring both possible combinations to have $m(K\pi) < 1700$ MeV.

243 1. $X_s \rightarrow K^+ \pi^+ \pi^-$:

- 244 (a) $B_s^0 \rightarrow D_s \pi\pi\pi$
- 245 (b) $B_s^0 \rightarrow D_s^- D_s^+$
- 246 (c) $B_s^0 \rightarrow D_s^- \mu\nu X$

247 2. $X_d \rightarrow \pi^+ \pi^+ \pi^-$:

- 248 (a) $B_s^0 \rightarrow D_s K\pi\pi$
- 249 (b) $B_s^0 \rightarrow D_s^- \mu\nu X$

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

257 **3.2.2 Phase space region**

258 Due to the comparable low masses of the final state particles with respect to the B_s
259 mass, there is, in contrast to for example b-hadron to charmonia or charm decays, a
260 huge phase-space available for the $B_s^0 \rightarrow D_s K\pi\pi$ decay. For the invariant mass of
261 the $K\pi\pi$ subsystem it extends up to 3.4 GeV. It has however been observed that the
262 decay proceeds predominantly through the low lying axial vector states $K(1270)$ and
263 $K(1400)$, while the combinatorial background is concentrated at high $K\pi\pi$ invariant
264 masses ($m(K\pi\pi) > 2000$ MeV). Moreover, the strange hadron spectrum above 2 GeV is
265 poorly understood experimentally such that an reliable extraction of the strong phase
266 motion in that region is not possible. We consequently choose the considered phase space
267 region to be $m(K\pi\pi) < 1950$ MeV, which is right below the charm-strange threshold
268 ($B_s^0 \rightarrow D_s^+ D_s^-$).

269 **3.2.3 Training of multivariate classifier**

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

- 272 • max(ghostProb) over all tracks
- 273 • cone(p_T) asymmetry of every track, which is defined to be the difference between the
 274 p_T of the π/K and the sum of all other p_T in a cone of radius $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}$
 275 < 1 rad around the signal π/K track.
- 276 • min(IP χ^2) over the X_s daughters
- 277 • max(DOCA) over all pairs of X_s daughters
- 278 • min(IP χ^2) over the D_s daughters
- 279 • D_s and B_s^0 DIRA
- 280 • D_s FD significance
- 281 • max($\cos(D_s h_i)$), where $\cos(D_s h_i)$ is the cosine of the angle between the D_s and
 282 another track i in the plane transverse to the beam
- 283 • B_s^0 IP χ^2 , FD χ^2 and Vertex χ^2

284 Various classifiers were investigated in order to select the best performing discriminator.
 285 Consequently, a boosted decision tree with gradient boost (BDTG) is chosen as nominal
 286 classifier. We use truth-matched MC as signal input. Simulated signal candidates are
 287 required to pass the same trigger, stripping and preselection requirements, that were used
 288 to select the data samples. For the background we use events from the high mass sideband
 289 ($m_{B_s^0 \text{ candidate}} > 5600$ MeV/ c^2) of our data samples.

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

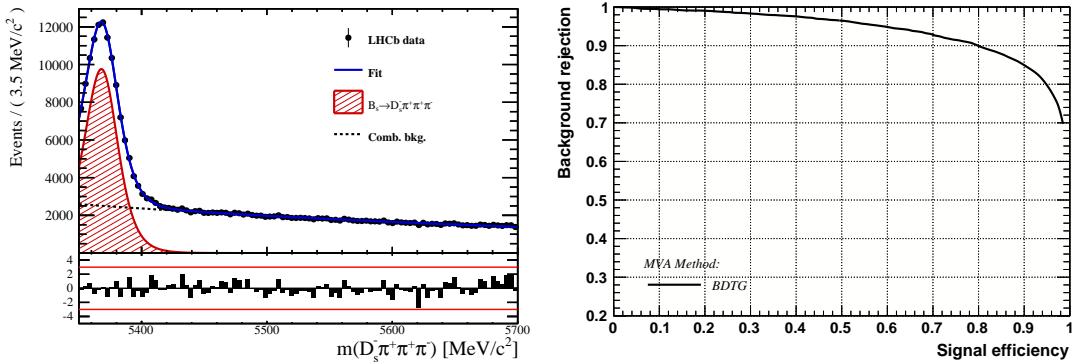


Figure 3.1

292 **3.2.4 Final selection**

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

	Description	Requirement
$D_s \rightarrow hh\bar{h}$	$m(hh\bar{h})$	$= m_{D_s} \pm 20$ MeV
$D_s^- \rightarrow KK\pi^-$	D^0 veto	$m(KK) < 1840$ MeV
$D_s^- \rightarrow \phi\pi^-$	$m(KK)$ PIDK(K^+) PIDK(K^-) PIDK(π^-) χ_{FD}^2 FD in z D^- veto Λ_c veto	$= m_\phi \pm 20$ MeV > -10 > -10 < 20 > 0 > -1 $m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 30$ MeV PIDK(K^-) > 0 $m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 30$ MeV PIDK(K^-) – PIDp(K^-) > 0
$D_s^- \rightarrow K^*(892)K^-$	$m(KK)$ $m(K^+\pi^-)$ PIDK(K^+) PIDK(K^-) PIDK(π^-) χ_{FD}^2 FD in z D^- veto Λ_c veto	$\neq m_\phi \pm 20$ MeV $= m_{K^*(892)} \pm 75$ MeV > -10 > -5 < 10 > 2 > 0 $m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 30$ MeV PIDK(K^-) > 5 $m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 30$ MeV PIDK(K^-) – PIDp(K^-) > 5
$D_s^- \rightarrow (KK\pi^-)_{NR}$	$m(KK)$ $m(K^+\pi^-)$ PIDK(K^+) PIDK(K^-) PIDK(π^-) χ_{FD}^2 FD in z D^- veto Λ_c veto	$\neq m_\phi \pm 20$ MeV $\neq m_{K^*(892)} \pm 75$ MeV > 5 > 5 < 10 > 5 > 0 $m(K^+K_\pi^-\pi^-) \neq m(D^-) \pm 30$ MeV PIDK(K^-) > 20 $m(K^+K_p^-\pi^-) \neq m(\Lambda_c) \pm 30$ MeV PIDK(K^-) – PIDp(K^-) > 5
$D_s \rightarrow \pi\pi\pi$	PIDK(π) PIDp(π) D^0 veto χ_{FD}^2 FD in z	< 10 < 10 $m(\pi^+\pi^-) < 1700$ MeV > 9 > 0
$D_s^- \rightarrow K^-\pi^+\pi^-$	PIDK(K) PIDK(π) PIDp(π) D^0 veto χ_{FD}^2 FD in z	> 10 < 5 < 10 $m(K^-\pi^+) < 1750$ MeV > 9 > 0

3.3 Simulation

293

Table 3.2: Offline selection requirements for $B_s \rightarrow D_s K\pi\pi(D_s\pi\pi\pi)$ candidates.

	Description	Requirement
$B_s \rightarrow D_s h\pi\pi$	$m(D_s h\pi\pi)$	$> 5200 \text{ MeV}$
	χ^2_{vtx}/ndof	< 8
	DIRA	> 0.99994
	χ^2_{FD}	> 100
	χ^2_{IP}	< 20
	χ^2_{DTF}/ndof	< 15
	t	$> 0.4 \text{ ps}$
	δt	$< 0.15 \text{ ps}$
	Phasespace region	$m(h\pi\pi) < 1.95 \text{ GeV}$ $m(h\pi) < 1.2 \text{ GeV}$ $m(\pi\pi) < 1.2 \text{ GeV}$
	Wrong PV veto	$\text{nPV} = 1 \parallel \min(\Delta\chi^2_{IP}) > 20$
$X_s^+ \rightarrow K^+\pi^+\pi^-$	PIDK(K)	> 10
	PIDK(π^+)	< 10
	PIDK(π^-)	< 5
	Semi.-lep. veto	$\text{isMuon}(K^+) = 0$
$X_s^+ \rightarrow \pi^+\pi^+\pi^-$	PIDK(π^+)	< 5
	PIDK(π^-)	< 10
	Semi.-lep. veto	$\text{isMuon}(\pi^+) = 0$

294 **4 Fits to invariant mass distributions of signal and**
 295 **normalization channel**

296 In order to properly model the invariant mass distribution of $B_s^0 \rightarrow D_s K\pi\pi$ and $B_s^0 \rightarrow$
 297 $D_s \pi\pi\pi$ candidates, the expected signal shape, as well as the expected shape for the
 298 combinatorial and physical background has to be known. This model can then be used to
 299 fit the distributions and obtain signal sWeights [31], which are employed to suppress the
 300 residual background that is still left in the sample, for the time-dependent amplitude fit.

301 **4.1 Signal models for $m(D_s \pi\pi\pi)$ and $m(D_s K\pi\pi)$**

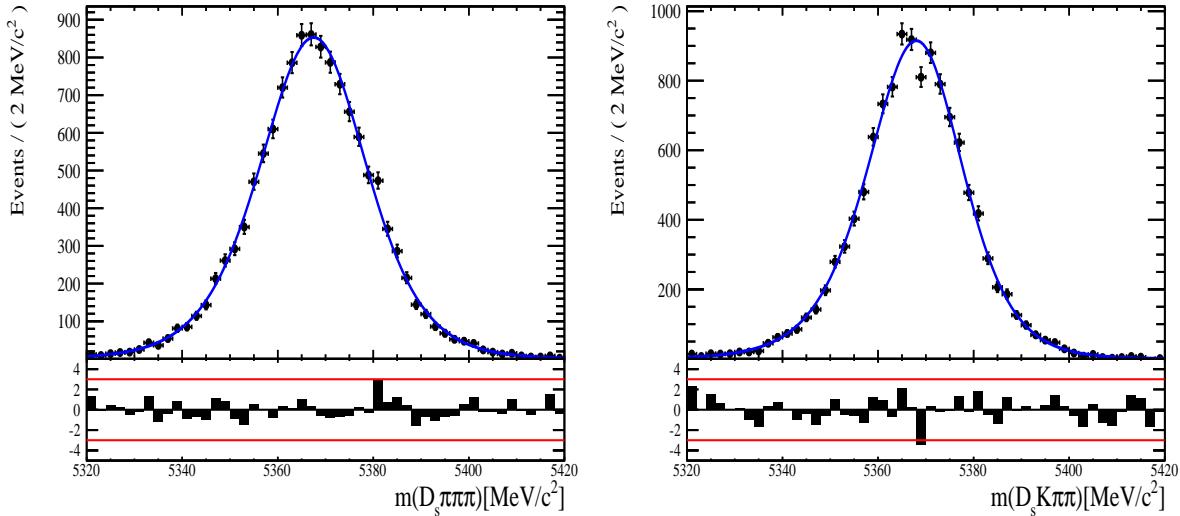


Figure 4.1: Invariant mass distributions of simulated (left) $B_s^0 \rightarrow D_s \pi\pi\pi$ and (right) $B_s^0 \rightarrow D_s K\pi\pi$ events. A fit of a RooJohnsonSU function to each distribution is overlaid.

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

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

304 The sign of γ in Eq. 4.1 determines whether the tail is located at lower ($\gamma > 0$)
 305 or higher ($\gamma < 0$) invariant mass values than the mean μ of the gaussian function and
 306 δ describes the (a)symmetry of the fitted distribution. Higher values of δ result in a
 307 more symmetric, gaussian-like function. Another Johnson SU function function is used
 308 to account for the contribution of the $B^0 \rightarrow D_s K\pi\pi$ decay, which is also present in
 309 the $m(D_s K\pi\pi)$ spectrum. The width, as well as the tail parameters are fixed to values
 310 obtained by a fit to the invariant mass distribution of simulated events shown in Fig 4.1.
 311 A linear scaling factor for the mean μ and width σ is floated in the fit to account for
 312 possible differences between the simulation and real data.

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

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

318 4.2 Background models for $m(D_s\pi\pi\pi)$

319 Different background sources arise in the invariant mass spectrum of candidates in the
320 normalization mode.

321 The following backgrounds have to be accounted for:

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

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

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

336 4.3 Background models for $m(D_sK\pi\pi)$

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

- 338 Combinatorial background: same contributions as discussed in Sec. 4.2.
- 339 Partially reconstructed $B_s^0 \rightarrow D_s^*K\pi\pi$ decays, with $D_s^* \rightarrow D_s\gamma$ or $D_s^* \rightarrow D_s\pi^0$, where the γ/π^0 is not reconstructed in the decay chain.
- 341 Partially reconstructed $B^0 \rightarrow D_s^*K\pi\pi$ decays, with $D_s^* \rightarrow D_s\gamma$ or $D_s^* \rightarrow D_s\pi^0$, where the γ/π^0 is not reconstructed in the decay chain.
- 343 Misidentified $B_s^0 \rightarrow D_s\pi\pi\pi$ decays, where one of the pions is wrongly identified as a kaon $\pi \rightarrow K$.
- 345 Misidentified, partially reconstructed $B_s^0 \rightarrow D_s^*\pi\pi\pi$ decays, where one of the pions is wrongly identified as a kaon $\pi \rightarrow K$ and the γ/π^0 from $D_s^* \rightarrow D_s\gamma/\pi^0$ is not reconstructed.

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

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

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

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

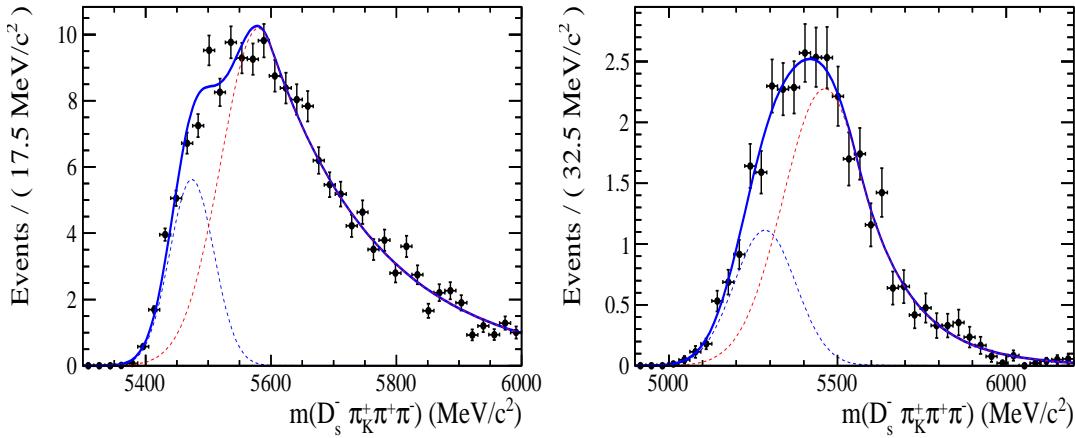


Figure 4.2: Invariant mass distribution of (left) simulated $B_s^0 \rightarrow D_s \pi\pi\pi$ events, where one of the π 's is reconstructed as a K and the misID probability for each event is taken into account. The corresponding distribution for simulated $B_s^0 \rightarrow D_s^* \pi\pi\pi$ events, where the γ/π^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.

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

370 In the same way as mentioned above, we can determine the rate of misidentified, partially
 371 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
 372 and a MC sample of $B_s^0 \rightarrow D_s^* \pi\pi\pi$ events. The invariant mass distribution we obtain
 373 when we exclude the γ/π^0 , flip the the particle hypothesis $\pi \rightarrow K$ and apply the event
 374 weights given by the fake rate, is shown in Fig. 4.2 (right). The fit of two Crystal Ball
 375 functions to this distribution is overlaid. The yield of this contribution is determined
 376 from the yield of $B_s^0 \rightarrow D_s^* \pi\pi\pi$ candidates in the nominal mass fit of our normalization
 377 channel, multiplied by the misID probability of $\propto 3.6\%$.

378 4.4 Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates

379 An unbinned maximum likelihood fit is performed simultaneously to the invariant mass
 380 distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates. As discussed in Sec. ??, the fit is given
 381 as a Johnson SU signal model for the B_s^0 and B^0 signal, the sum of three bifurcated
 382 Gaussian functions to model the partially reconstructed $B_s^0 \rightarrow D_s^*\pi\pi\pi$ background and
 383 an Exponential function to account for combinatorial background. The invariant mass
 384 distribution and the fit is shown in Fig. 4.3. All simultaneously performed fits to the
 385 $m(D_s\pi\pi\pi)$ distribution, ordered by the respective D_s final state, can be found in the
 386 Appendix ???. The obtained yields are summarized in Table 4.1.

387 4.5 Fit to $B_s^0 \rightarrow D_sK\pi\pi$ candidates

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

396 4.6 Extraction of signal weights

397 The sPlot technique [31] is used to extract signal weights from the fits to the invariant
 398 mass distributions of our signal and normalization channel. This statistical tool assigns
 399 a weight to every event, according to its position in the respective mass distribution,
 400 given the fitted signal and background models. The weights can then be used to suppress
 401 the background components in every other observable distribution of interest. Figure ??
 402 shows the distribution of weights across the invariant mass spectra of $B_s^0 \rightarrow D_s\pi\pi\pi$ and
 403 $B_s^0 \rightarrow D_sK\pi\pi$ 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 ± 25	860 ± 38	309 ± 21	1984 ± 55
$B^0 \rightarrow D_sK\pi\pi$	276 ± 26	692 ± 41	261 ± 23	1385 ± 58
$B^0/B_s^0 \rightarrow D_s^*K\pi\pi$	7 ± 25	171 ± 75	114 ± 25	893 ± 84
$B_s^0 \rightarrow D_s^{(*)}\pi\pi\pi$	63 ± 0	158 ± 0	53 ± 0	314 ± 0
combinatorial	1482 ± 53	2884 ± 100	605 ± 43	4261 ± 133
$m(D_s\pi\pi\pi)$				
$B_s^0 \rightarrow D_s\pi\pi\pi$	9183 ± 105	22083 ± 166	7574 ± 95	43773 ± 245
$B^0 \rightarrow D_s\pi\pi\pi$	289 ± 58	716 ± 95	229 ± 54	968 ± 147
$B_s^0 \rightarrow D_s^*\pi\pi\pi$	3640 ± 130	9086 ± 232	3047 ± 110	17827 ± 421
combinatorial	4991 ± 154	11127 ± 271	3728 ± 126	24589 ± 500

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

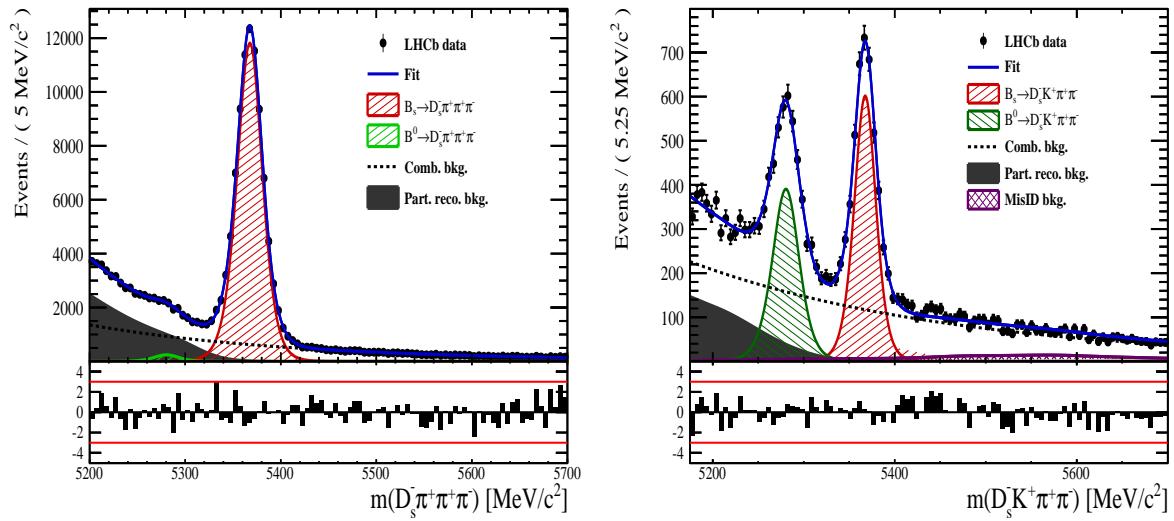


Figure 4.3: Invariant mass distribution of (left) $B_s^0 \rightarrow D_s \pi \pi \pi$ and (right) $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run1 and Run2 data. The respective fit described in the text is overlaid.

404 5 Flavour Tagging

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

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

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

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

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

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

429 where ϵ_{tag} is the fraction of events that have a tagging decision. The flavour
 430 tagging algorithms are optimized for highest ϵ_{eff} on data, using the $B_s^0 \rightarrow D_s^- \pi^+$ and
 431 $B^+ \rightarrow J/\psi K^+$ samples.

432 Utilizing flavour-specific final states, the predicted mistag η of each tagger has to be
 433 calibrated to match the observed mistag ω on the data sample. For the calibration, a
 434 linear model of the form

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

435 where the values of p_0 and p_1 are determined using the $B_s^0 \rightarrow D_s \pi \pi \pi$ normalization
 436 mode and $\langle \eta \rangle$ is the average estimated mistag probability $\langle \eta \rangle = \sum_{i=1}^{N_{cand}} (\eta_i)/N_{cand}$.
 437 Following this model, a perfectly calibrated tagger would lead to $\omega(\eta) = \eta$ and one would
 438 expect $p_1 = 1$ and $p_0 = \langle \eta \rangle$. Due to the different interaction cross-sections of oppositely
 439 charged particles, the tagging calibration parameters depend on the initial state flavour of
 440 the B_s^0 . Therefore, the flavour asymmetry parameters Δp_0 , Δp_1 and $\Delta \epsilon_{tag}$ are introduced.
 441 For this analysis, the calibrated mistag is treated as per-event variable, giving a larger
 442 weight to events that are less likely to have an incorrect tag. This adds one additional
 443 observable to the time- and amplitude-dependent fit.

444 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 Δm_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	$<\eta>$	ϵ_{tag}	Δp_o	Δp_1	$\epsilon_{eff} [\%]$
0.025 ± 0.005	0.944 ± 0.048	0.347	0.517 ± 0.002	0.028 ± 0.005	0.037 ± 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. Its 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	$<\eta>$	ϵ_{tag}	Δp_o	Δp_1	$\epsilon_{eff} [\%]$
0.008 ± 0.004	1.086 ± 0.059	0.381	0.571 ± 0.002	-0.017 ± 0.004	0.135 ± 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 weighted data, to compare the similarity of the mistag probabilities, tagging decisions and kinematic observables that are correlated with the tagging response, on simulation and data.

473 The distributions of the predicted mistag probability η for the OS combination and the
 474 SS kaon tagger are shown in Fig. ?? (simulation) and Fig. 5.1 (data).

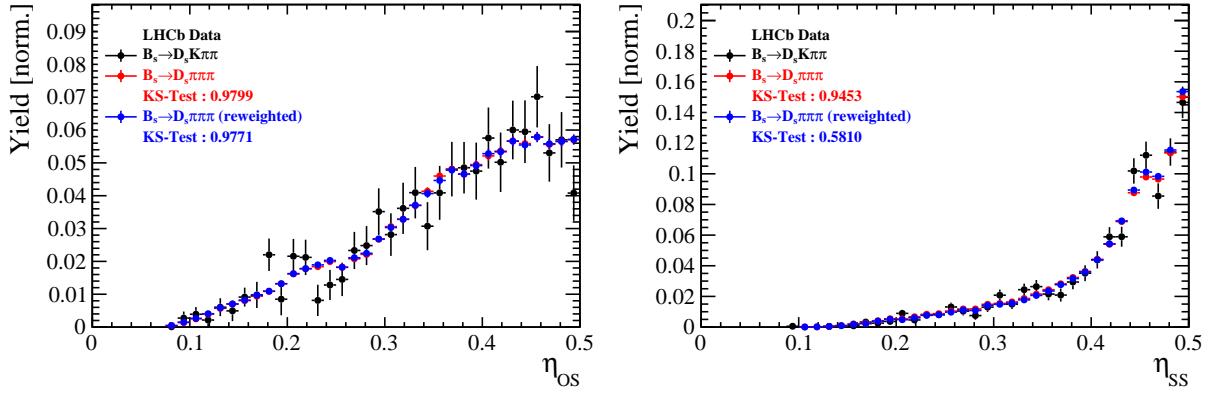


Figure 5.1: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \rightarrow D_s K\pi\pi$ (black) and $B_s^0 \rightarrow D_s \pi\pi\pi$ (red) data samples.

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

478 Fig. ?? shows the signal data distributions of the transverse B_s^0 momentum p_T , the
 479 pseudorapidity η of the signal candidate and the number of reconstructed tracks per event.
 480 Sufficient agreement is observed.

481 To justify the portability of the flavour tagging calibration obtained from $B_s^0 \rightarrow D_s \pi\pi\pi$
 482 to the $B_s^0 \rightarrow D_s K\pi\pi$ channel, besides the good agreement of the distributions shown
 483 above, the dependence of the measured mistag ω on the predicted mistag η has to be
 484 compatible in both channel. This dependence is shown in Fig. 5.2 for simulated signal
 485 events of both channels, where good agreement is observed.

486 5.4 Combination of OS and SS taggers

487 In the time- and amplitude-dependent fit to $B_s^0 \rightarrow D_s K\pi\pi$ data, the obtained tagging
 488 responses of the OS and SS tagger will be combined after the calibration described in the
 489 previous sections is applied. Events that aquire a mistag probability greater than 0.5 after
 490 the calibration will have their tagging decision flipped. For events where only one of the
 491 two taggers fired, the combination of the tagging decision is trivial. In those events where
 492 both taggers made a decision, we use the standard combination of taggers [35] provided
 493 by the flavour tagging group. In the nominal fit, the calibrated mistags ω are combined
 494 event by event for the OS and SS tager, thus adding one variable to observable to the
 495 fit procedure. This ensures that the uncertainties of the OS and SS tagging calibration
 496 parameters are propagated properly to the combined tagging response for each event.
 497 The taggging performance for the combined tagger in the categories SS tagged only, OS
 498 tagged only and SS+OS tagged, are shown in Tab. ?? for the signal and normalization
 499 channel. The distribution of the observed mistag ω as a function of the combined mistag
 500 probability η for $B_s^0 \rightarrow D_s \pi\pi\pi$ decays is shown in Fig. 5.3.

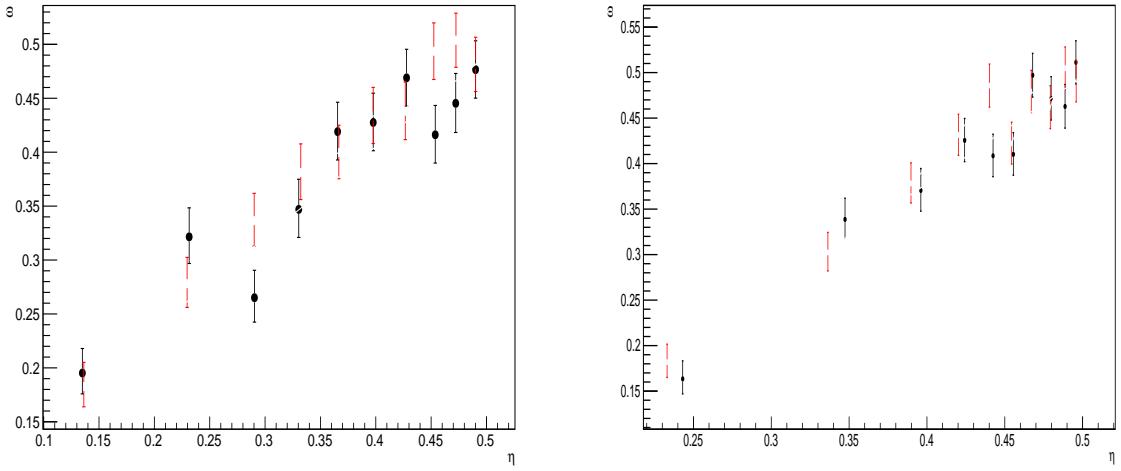


Figure 5.2: Dependence of the observed mistag ω on the predicted mistag η for the (left) OS combination and the (right) SS kaon tagger, found in the simulated $B_s^0 \rightarrow D_s K\pi\pi$ (black) and $B_s^0 \rightarrow D_s \pi\pi\pi$ (red) signal samples.

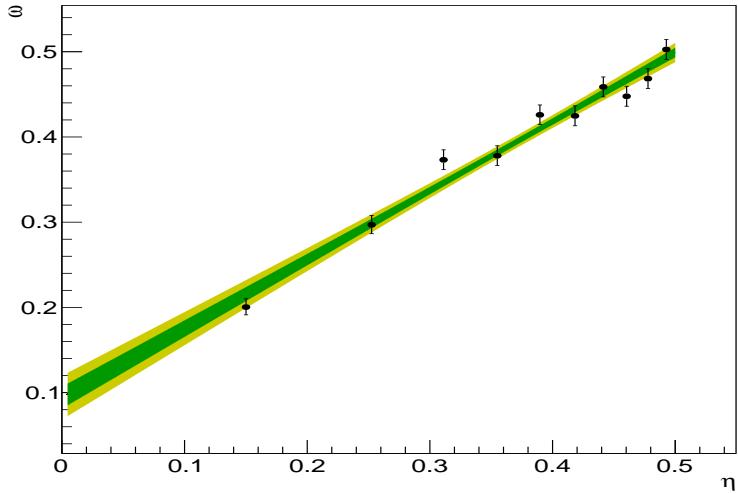


Figure 5.3: Distribution of the predicted combined mistag probability η versus the observed mistag ω for $B_s^0 \rightarrow D_s \pi\pi\pi$ signal candidates. The fit with a linear polynomial, used to determine p_0 and p_1 is overlaid.

Table 5.3: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-I data.

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

Table 5.4: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-II data.

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

501 6 Acceptance

502 6.1 MC corrections

503 6.1.1 Truth matching of simulated candidates

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

513 We therefore include `BKGCAT`= 60 and statistically subtract the ghost background by
 514 using the `sPlot` technique. The `sWeights` are calculated from a fit to the reconstructed B_s
 515 mass. The signal contribution is modeled as described in Sec. 4.1 and the background
 516 with a polynomial. The fit is performed simultaneously in two categories; the first includes
 517 events with `BKGCAT` = 20 or 50 and the second events with `BKGCAT` = 60. To account
 518 for the different mass resolution we use a different σ for each category, while the mean
 519 and the tail parameters are shared between them. The background component is only
 520 included for the second category.

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

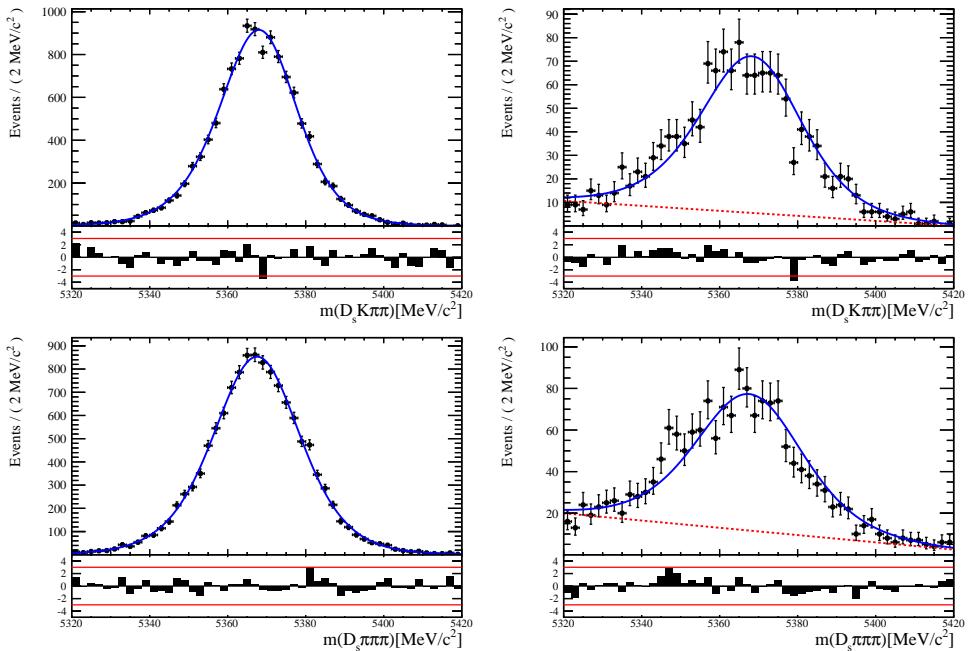


Figure 6.1: The reconstructed B_s mass distribution for simulated $B_s \rightarrow D_s K\pi\pi$ (top) and $B_s \rightarrow D_s \pi\pi\pi$ (bottom) decays classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

523 6.1.2 PID efficiencies

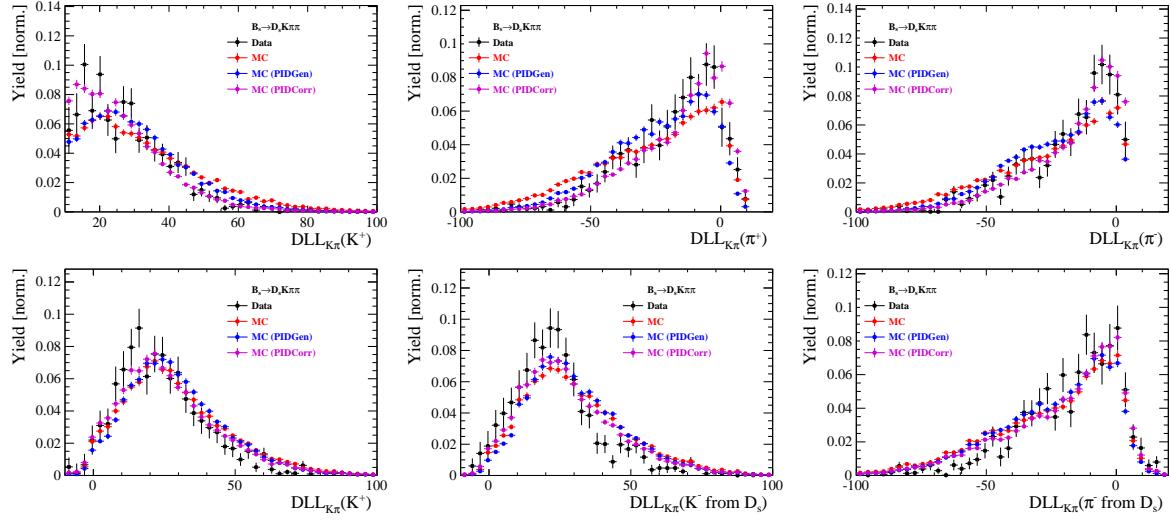


Figure 6.2

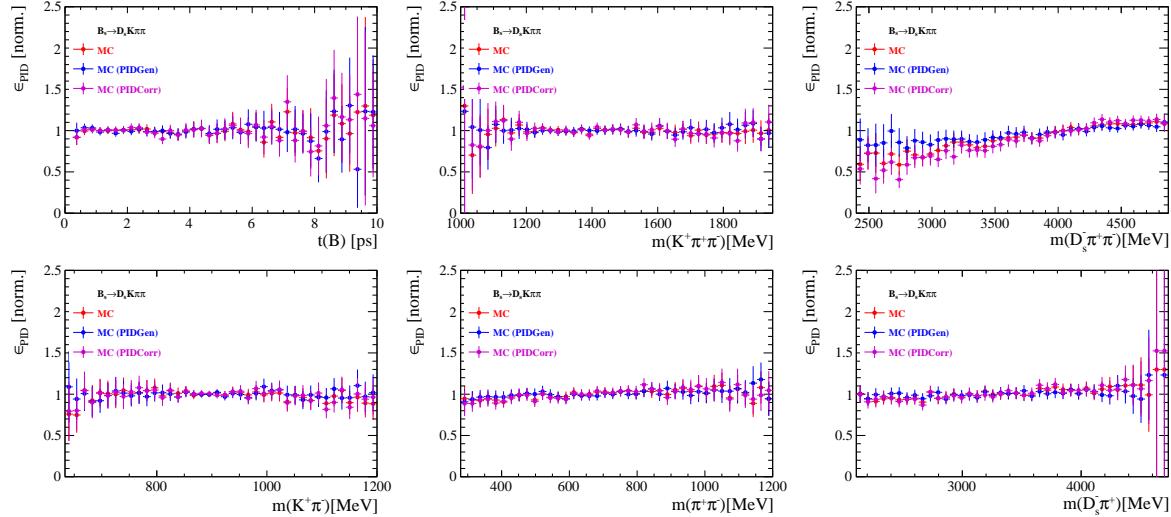


Figure 6.3

524 6.1.3 BDT efficiencies

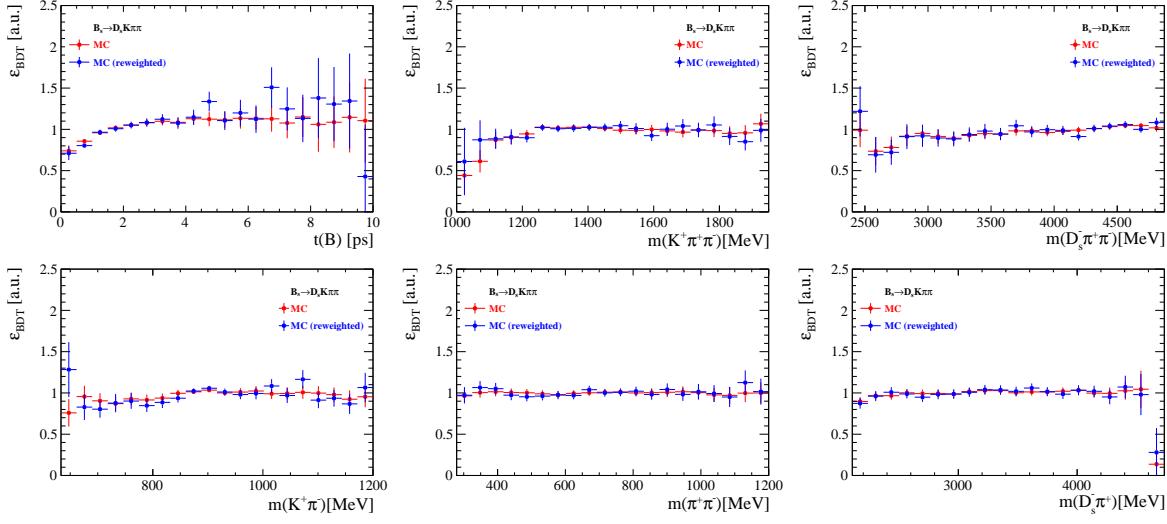


Figure 6.4

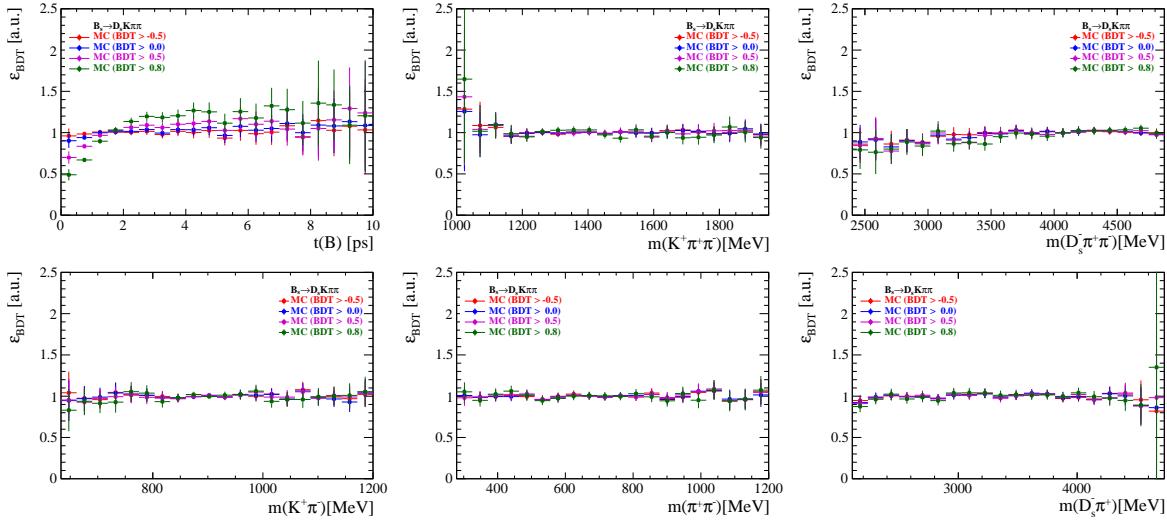


Figure 6.5

⁵²⁵ 6.1.4 Tracking efficiencies

526 **6.2 Decay-time acceptance**

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

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

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

540 A PDF of the form

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

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

550 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
 551 v_i , $i = [1, \dots, 8]$ are used. Since a rapid change of the decay time acceptance at low
 552 decay times due to the turn-on effect generated by the lifetime and other selection cuts is
 553 expected, more knots are placed in that regime. At higher decay times we expect linear
 554 behavior, with a possible small effect due to the VELO reconstruction. Therefore fewer
 555 knots are used. Furthermore, v_7 is fixed to 1 in order to normalize the overall acceptance
 556 function. To stabilise the last spline, v_8 is fixed by a linear extrapolation from the two
 558 previous splines:

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

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

560 6.2.1 Comparison of acceptance in subsamples

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

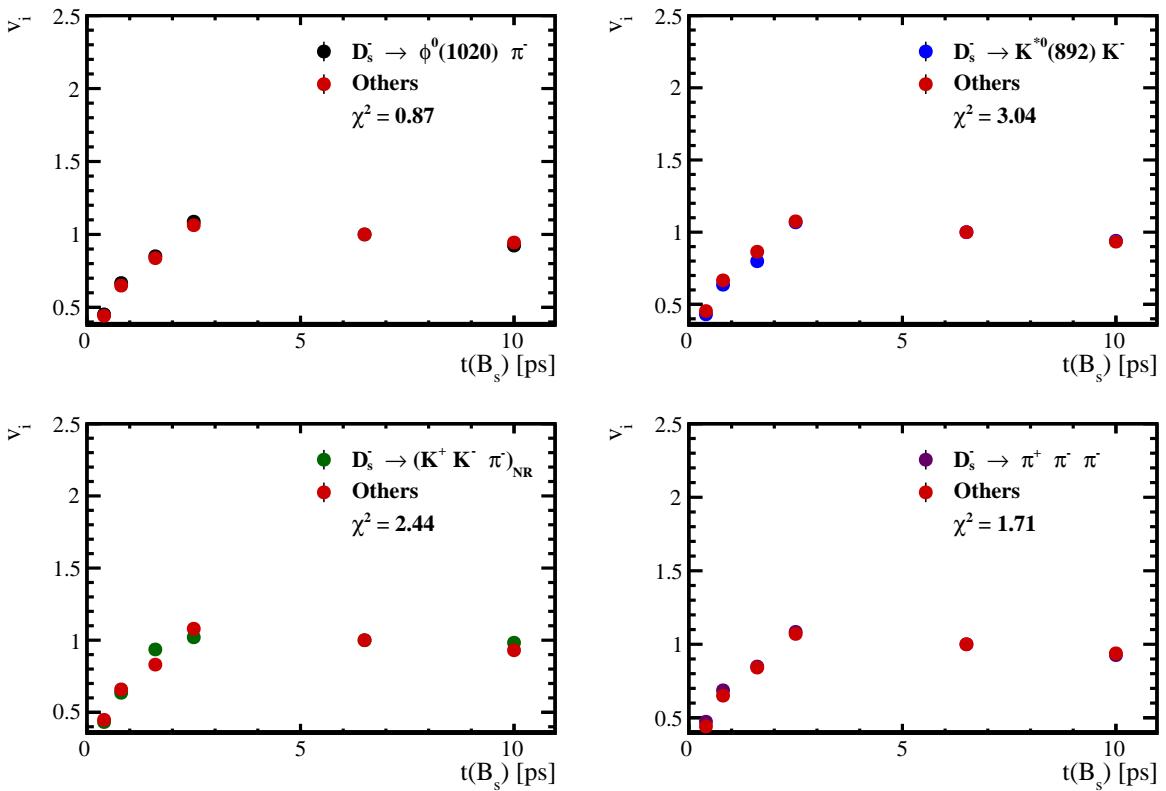


Figure 6.6: Comparison of the spline coefficients obtained from time-dependent fits to the $B_s^0 \rightarrow D_s\pi\pi\pi$ subsamples of different D_s final states. The comparison of one particular D_s state against all other states is shown.

570 Investigating the obtained spline coefficients from different D_s final states, good
 571 agreement is observed between all four channels and no need to distinguish between
 572 different final states in the time-dependent amplitude fit is found.
 573 The comparison between spline coefficients for the different runs and L0 trigger categories
 574 is shown in Figure 6.7.

575 Significant deviations between spline coefficients obtained from the two different runs
 576 and L0 trigger categories can be observed. The deviations are most pronounced in the
 577 (0 – 5) ps region, where the majority of statistics is found. Therefore, the time-dependent
 578 efficiency has to be treated separately for the runs and L0 categories. This is achieved by

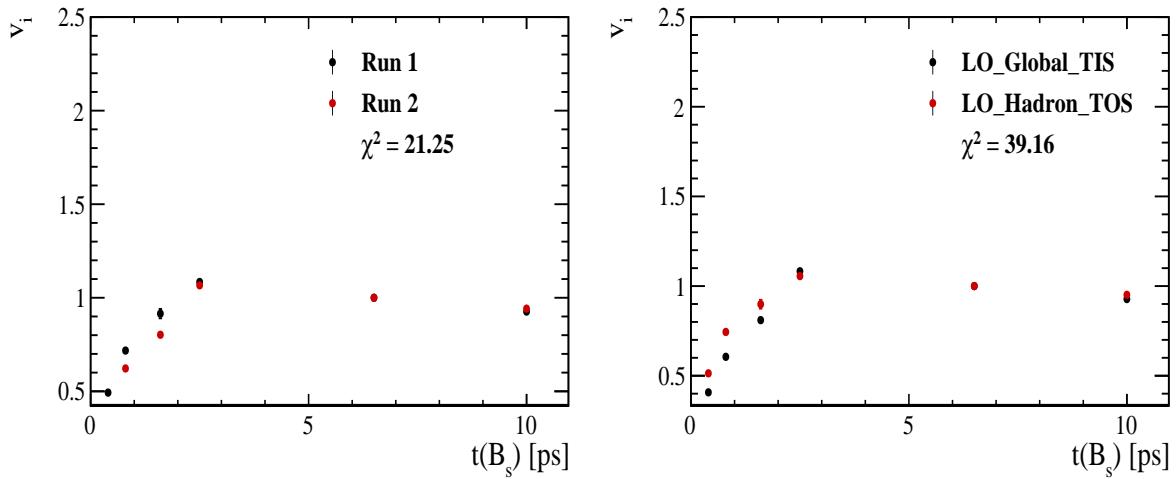


Figure 6.7: Comparison of the spline coefficients obtained from time-dependent fits to the $B_s^0 \rightarrow D_s \pi \pi \pi$ subsamples of (left) the different runs and (right) L0 trigger categories.

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

581 **6.2.2 Results**

582 The nominal fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data using this configuration is shown in Figure ??.
583 Note that the normalization of the splines in the following figures is not in scale. The fit
584 parameters obtained from the described fits to data and simulation are summarised in
585 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 ± 0.038	0.546 ± 0.022	0.953 ± 0.060
0.8	v_1	0.826 ± 0.059	0.785 ± 0.034	0.910 ± 0.066
1.6	v_2	0.843 ± 0.087	0.905 ± 0.056	1.055 ± 0.095
2.5	v_3	1.154 ± 0.036	1.118 ± 0.028	0.930 ± 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 ± 0.031	0.412 ± 0.020	0.955 ± 0.077
0.8	v_1	0.583 ± 0.050	0.648 ± 0.033	0.910 ± 0.074
1.6	v_2	0.939 ± 0.101	0.953 ± 0.061	0.947 ± 0.096
2.5	v_3	1.052 ± 0.054	1.077 ± 0.035	1.003 ± 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 ± 0.009	0.482 ± 0.009	1.000 ± 0.000
0.8	v_1	0.691 ± 0.014	0.707 ± 0.015	1.000 ± 0.000
1.6	v_2	0.851 ± 0.024	0.926 ± 0.026	1.000 ± 0.000
2.5	v_3	1.061 ± 0.017	1.086 ± 0.018	1.000 ± 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 ± 0.007	0.482 ± 0.010	1.000 ± 0.000
0.8	v_1	0.476 ± 0.012	0.707 ± 0.016	1.000 ± 0.000
1.6	v_2	0.725 ± 0.023	0.926 ± 0.026	1.000 ± 0.000
2.5	v_3	1.064 ± 0.019	1.086 ± 0.018	1.000 ± 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)

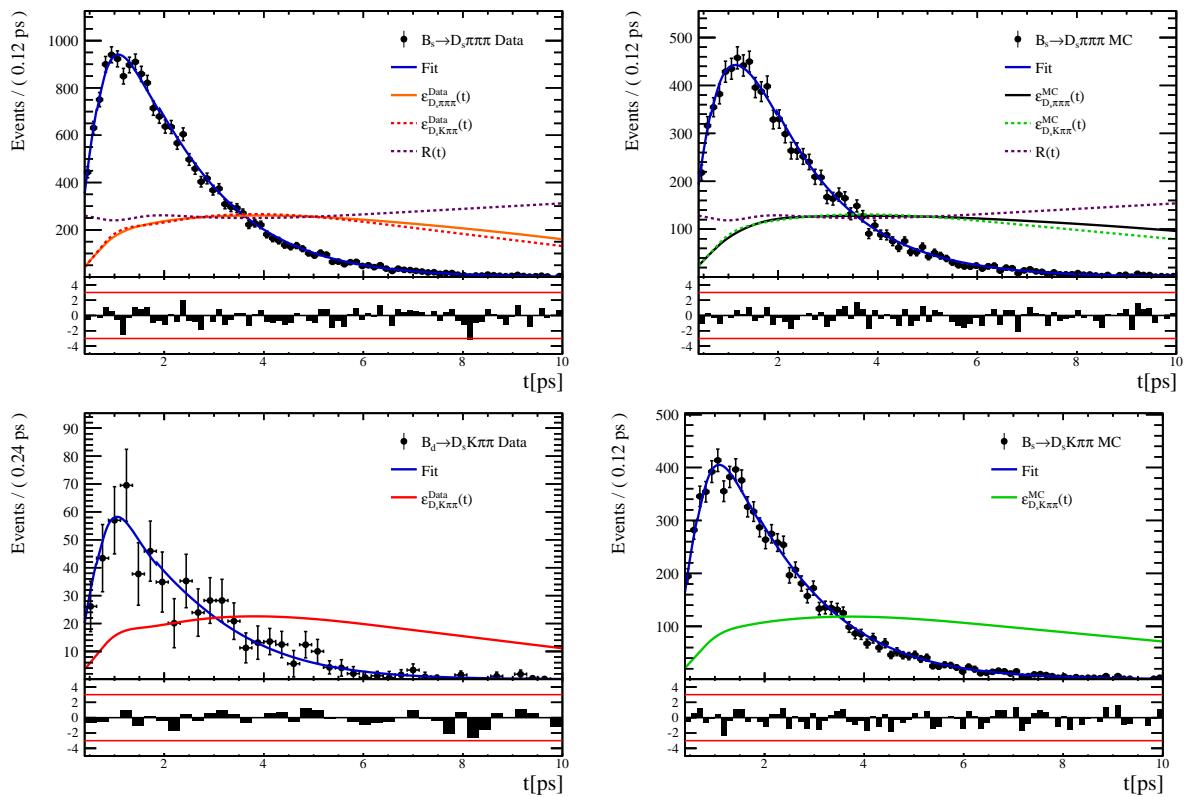


Figure 6.8

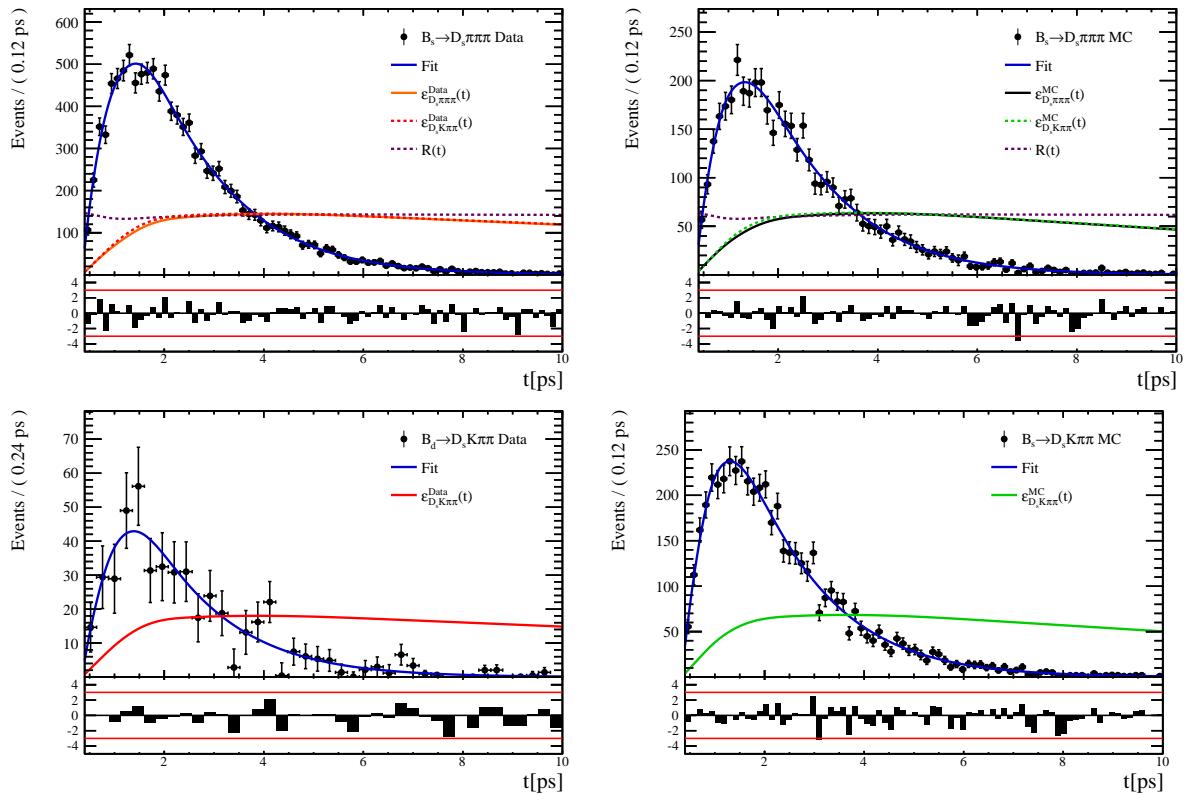


Figure 6.9

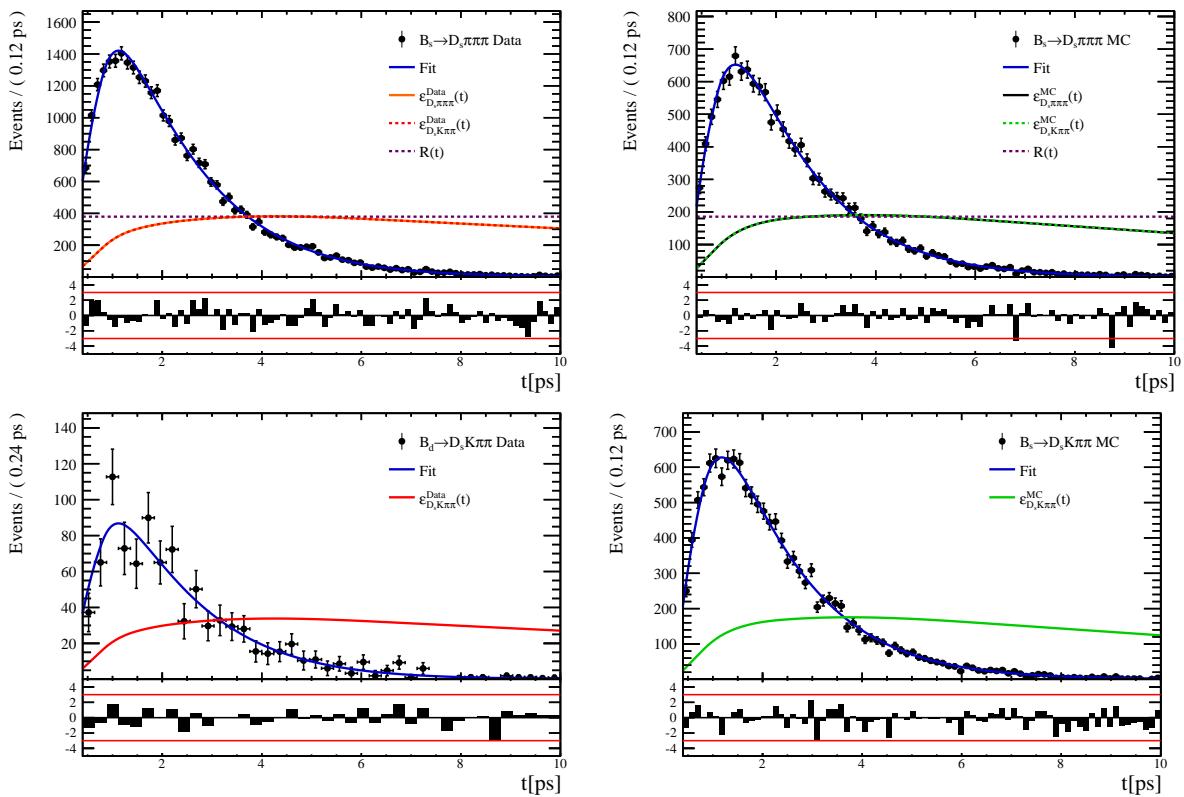


Figure 6.10

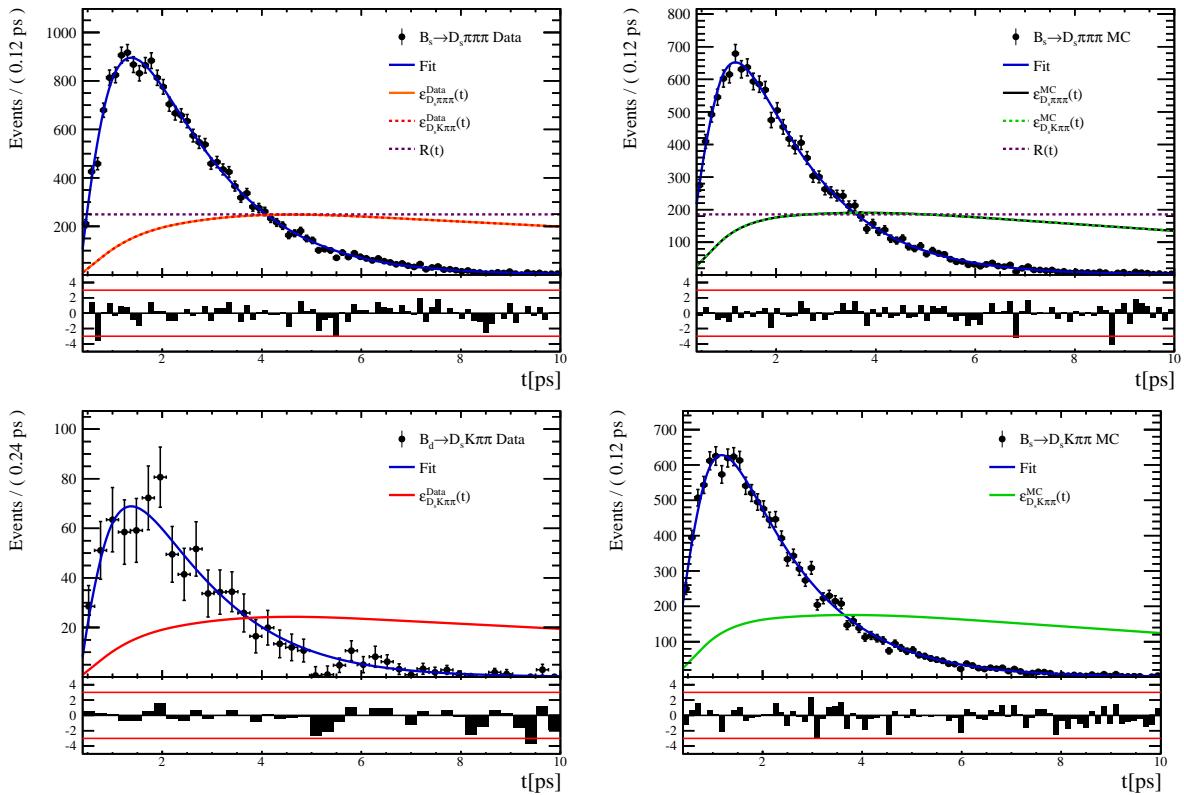


Figure 6.11

586 **6.3 Phasespace acceptance**

587 7 Decay-time Resolution

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

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

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

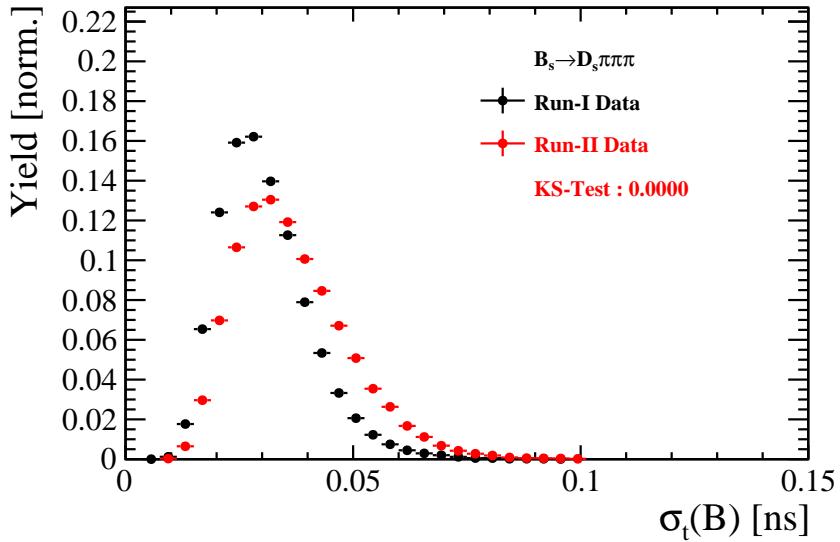


Figure 7.1: Distribution of the decay time error for $B_s^0 \rightarrow D_s\pi\pi\pi$ signal candidates for Run-I (black) and Run-II (red) data (sWeighted).

610 7.1 Calibration for Run-I data

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

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

619 where σ_1 and σ_2 are the widths of the Gaussians, f_1 is the relative fraction of events
 620 described by the first Gaussian relative to the second and Δm_s is the oscillation frequency
 621 of B_s^0 mesons. An effective single Gaussian width is calculated from the dilution as,

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

622 which converts the resolution into a single-Gaussian function with an effective resolution
 623 that causes the same damping effect on the magnitude of the B_s oscillation. For the Run-I
 624 $B_s^0 \rightarrow D_s K \pi\pi$ MC sample the effective average resolution is found to be $\sigma_{eff} = 39.1 \pm 0.3$ fs.

625 To analyze the relation between the per-event decay time error δ_t and the actual
 626 resolution σ_t , the simulated $B_s^0 \rightarrow D_s K \pi\pi$ sample is divided into equal-statistics slices of
 627 δ_t . For each slice, the effective resolution is determined as described above. Details of the
 628 fit results in each slice are shown in appendix D. Figure 7.2 shows the obtained values
 629 for σ_{eff} as a function of the per-event decay time error σ_t . To account for the variable
 630 binning, the bin values are not placed at the bin center but at the weighted mean of the
 631 respective per-event-error bin. A linear function is used to parametrize the distribution.
 632 The obtained values are

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

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

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

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

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

637 can be used for our analysis. The following calibration functions were used in the
 638 $B_s^0 \rightarrow D_s K$ analysis to estimate the systematic uncertainty on the decay-time resolution:

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

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

640 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
 641 assigned as additional systematic uncertainty.

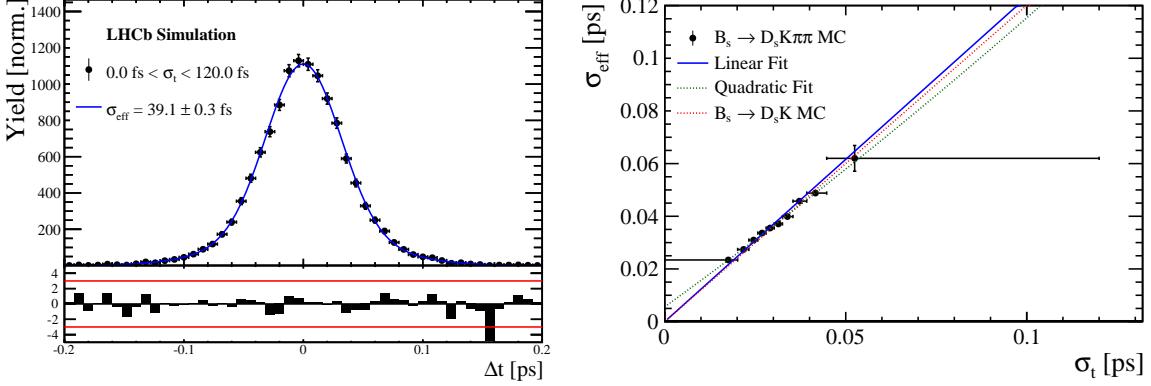


Figure 7.2: (Left) Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K\pi\pi$ MC candidates. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for $B_s \rightarrow D_s K\pi\pi$ MC (Run-I). The fitted calibration curve is shown in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

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

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

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

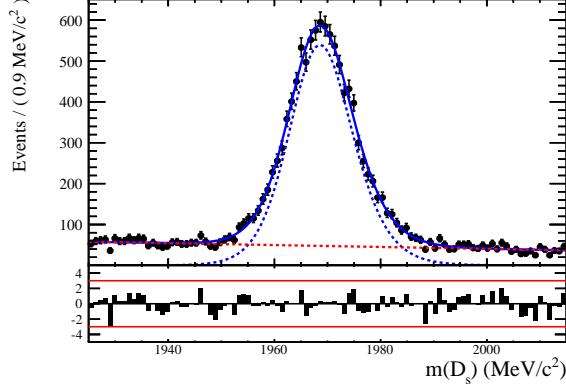


Figure 7.3: The invariant mass distribution for prompt D_s candidates.

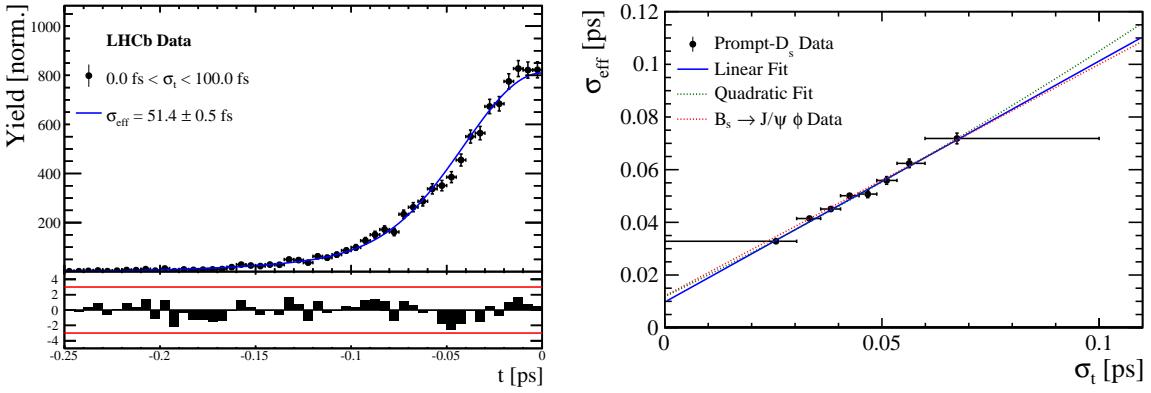


Figure 7.4: (Left) Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks. (Right) The measured resolution σ_{eff} as function of the per-event decay time error estimate σ_t for fake B_s candidates (Run-II data). The 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
$B_s \rightarrow D_s K\pi\pi$	χ^2_{vtx}/ndof	< 8
	χ^2_{DTF}/ndof	< 15
	t	< 0 ps
$D_s \rightarrow hhh$	χ^2_{vtx}/ndof	< 5
	DIRA	> 0.99994
	χ^2_{FD}	> 9
	p_T	> 1800 MeV
	χ^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
	Λ_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(π^-)	< 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(π^-)	< 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(π^-)	< 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(π)	< 5
	isMuon(h)	False
All tracks	p_T	> 500 MeV

670 7.3 Cross-checks

671 7.3.1 Kinematic dependence

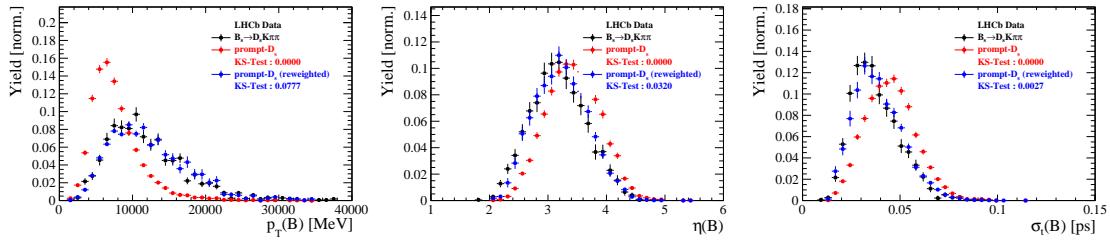


Figure 7.5

672 7.3.2 DTF constraints

673 8 Production and Detection Asymmetries

674 8.1 B_s Production Asymmetry

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

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

678 where σ are the corresponding production cross-section. This asymmetry was measured
 679 by LHCb in pp collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV by means of a time-dependent
 680 analysis of $B_s \rightarrow D_s^- \pi^+$ decays [40]. The results in bins of p_T and η of the B_s meson
 681 are shown in Table 8.1. To correct for the different kinematics of $B_s \rightarrow D_s^- \pi^+$ and
 682 $B_s^0 \rightarrow D_s K \pi \pi$ decays, the measured B_s production asymmetries $A_p(p_T, \eta)$ are folded with
 683 the sWeighted p_T, η distribution of our signal channel. The resulting effective production
 684 asymmetries are:

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

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

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

685 As for Run-II data no measurement is available yet, we determine the production asym-
 686 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]	η	$A_p(B_s^0)_{\sqrt{s}=7 \text{ TeV}}$	$A_p(B_s^0)_{\sqrt{s}=8 \text{ TeV}}$
(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)	$-0.0586 \pm 0.0648 \pm 0.0150$	$0.0537 \pm 0.0447 \pm 0.0124$

687 8.2 $K^-\pi^+$ Detection Asymmetry

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

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

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

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

698 where possible CP violation in the $D^+ \rightarrow K_s^0\pi^+$ mode is predicted to be smaller than
 699 10^{-4} in the Standard Model [43]. The asymmetry in the neutral kaon system, $A(K^0)$, has
 700 to be taken into account as a correction.

701 We use a dedicated LHCb tool to determine $A^{det}(K^-\pi^+)$ for all data taking periods
 702 used in this analysis. A detailed description can be found in [42]. The tool provides
 703 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
 704 determine the asymmetry following Eq. 8.6. Several weighting steps are performed to
 705 match the kinematics of the calibration samples to our signal decay sample:

706 First, weights are assigned to the K^\pm and π^\pm of the $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$ sample, using
 707 p, η of the K^\pm and p_T, η of the π^\pm from our $B_s^0 \rightarrow D_s K\pi\pi$ signal decay. Then, weights
 708 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
 709 the corresponding, weighted distributions of the $D^\pm \rightarrow K^\pm\pi^\pm\pi^\pm$ sample. In a last
 710 step, weights are assigned to match the bachelor pions ϕ distributions between the two
 711 calibration samples.

712 After the samples are weighted, fits are performed to the invariant
 713 $m(K^-\pi^+\pi^+)/m(K^+\pi^-\pi^-)$ and $m(K_s^0\pi^+)/m(K_s^0\pi^-)$ distributions to determine
 714 $A^{det}(K^-\pi^+)$. The PDFs used to describe the invariant mass distributions consist of
 715 gaussian functions for the signal component and exponentials describing the residual
 716 background.

717 The detection asymmetry is determined separately for every year and (since it is a
 718 charge asymmetry effect) magnet polarity. Serving as an example for Run-I and Run-
 719 II, the fits used to determine $N(D^+ \rightarrow K^-\pi^+\pi^+)/N(D^- \rightarrow K^+\pi^-\pi^-)$ and $N(D^+ \rightarrow$
 720 $K_s^0\pi^+)/N(D^- \rightarrow K_s^0\pi^-)$ for 2011, magnet up data and 2015, magnet up data are shown
 721 in Fig. 8.1 and 8.2 respectively. The obtained values of $A^{det}(K^-\pi^+) + A(K^0)$ for all years
 722 and polarities are shown in Table 8.2.

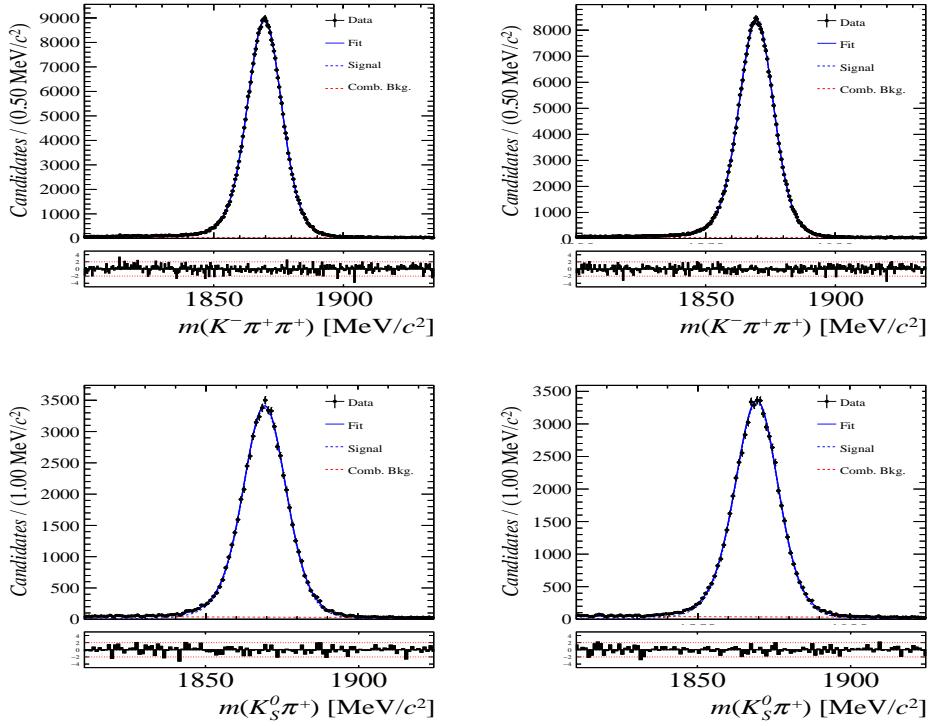


Figure 8.1: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_s^0 \pi^\pm$ candidates for Run-I data from the calibration samples. A fit described in the text is overlaid.

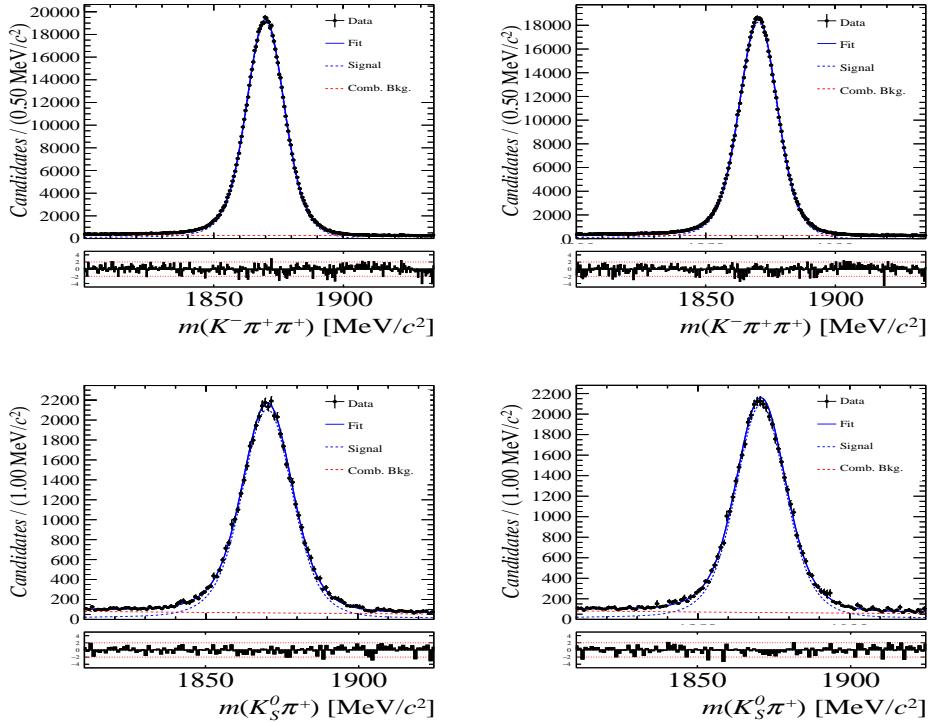


Figure 8.2: Distributions of the invariant mass of (top) $D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm$ and (bottom) $D^\pm \rightarrow K_s^0 \pi^\pm$ candidates for Run-II data from the calibration samples. A fit described in the text is overlaid.

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.

723 9 Time dependent fit

724 This section will cover the phasespace integrated, time dependent fit to $B_s^0 \rightarrow D_s h\pi\pi$ data,
 725 which is described by the PDF formulated in Eq. 2.6. For the phasespace integrated fit to
 726 $B_s^0 \rightarrow D_s K\pi\pi$ data, the sensitivity to the CKM phase γ will depend on the magnitude of
 727 the coherence factor κ , defined in Eq. 2.10, which is added as an additional fit parameter.
 728 In order to avoid any pollution of the final data samples by background events, both
 729 samples are weighted using the sWeights obtained by the fits to the invariant mass
 730 distributions, described in Sec. 4. This fit approach is commonly known as *sFit*. As
 731 additional input to the fit, the tagging information (Sec. 5), as well as the decay time
 732 acceptance (Sec. 6) and resolution (Sec. 7) is used and fixed to the values obtained by
 733 the dedicated studies. Taking all inputs into account, the final time dependent fit PDF is
 734 given by

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

735 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
 736 time acceptance effects and $\mathcal{R}(t - t')$ is the Gaussian time resolution function.

737 9.1 sFit to $B_s^0 \rightarrow D_s \pi\pi\pi$ data

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

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

745 Note that in this case, the dependence on the coherence factor κ is dropped and the
 746 same relation as found for $B_s^0 \rightarrow D_s \pi$ decays is recovered. Therefore, the only free fit
 747 parameter left is Δm_s . The data distribution with the overlaid fit is shown in Fig.

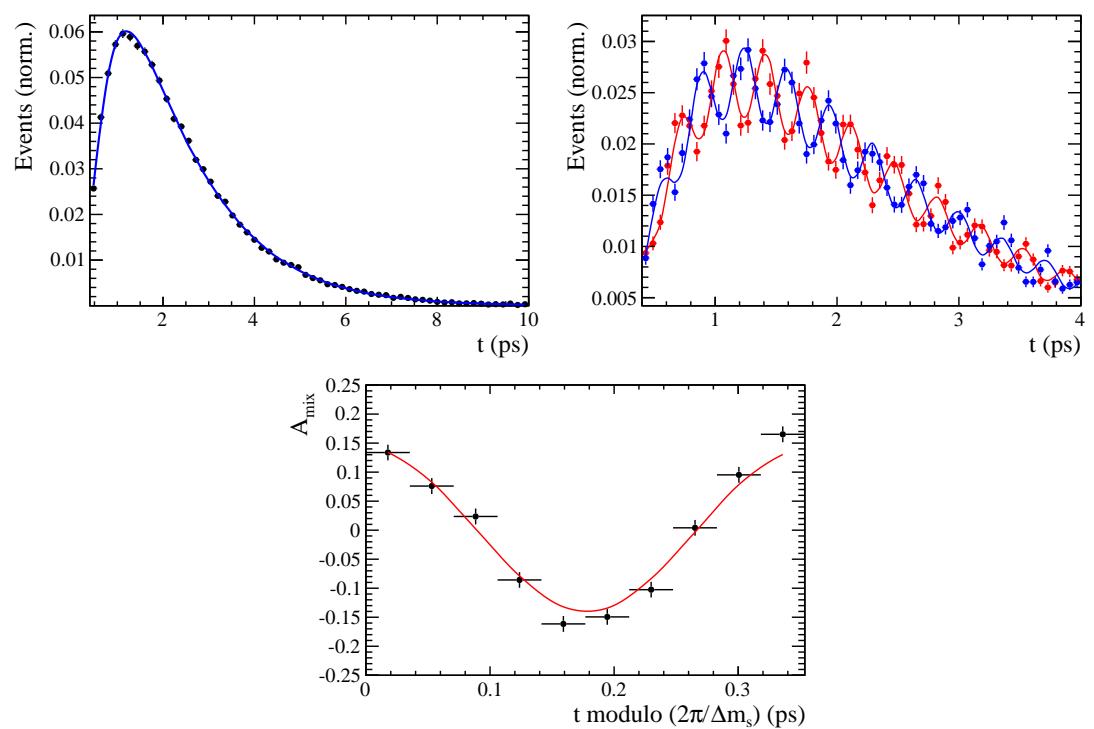


Figure 9.1

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

	Fit parameter	Value
Run-I	p_0^{OS}	0.3896 ± 0.0101
	p_1^{OS}	0.8883 ± 0.1074
	Δp_0^{OS}	0.0161 ± 0.0104
	Δp_1^{OS}	0.0005 ± 0.1095
	$\epsilon_{tag}^{\text{OS}}$	0.3851 ± 0.0031
	$\Delta \epsilon_{tag}^{\text{OS}}$	0.0069 ± 0.0123
	p_0^{SS}	0.4465 ± 0.0075
	p_1^{SS}	1.0748 ± 0.1012
	Δp_0^{SS}	-0.0190 ± 0.0076
	Δp_1^{SS}	0.1017 ± 0.1063
	$\epsilon_{tag}^{\text{SS}}$	0.6882 ± 0.0029
	$\Delta \epsilon_{tag}^{\text{SS}}$	-0.0076 ± 0.0117
	A_p	-0.0004 ± 0.0000
Run-II	p_0^{OS}	0.3669 ± 0.0074
	p_1^{OS}	0.9298 ± 0.0761
	Δp_0^{OS}	0.0118 ± 0.0085
	Δp_1^{OS}	0.0234 ± 0.0855
	$\epsilon_{tag}^{\text{OS}}$	0.3525 ± 0.0023
	$\Delta \epsilon_{tag}^{\text{OS}}$	0.0105 ± 0.0085
	p_0^{SS}	0.4532 ± 0.0055
	p_1^{SS}	0.9125 ± 0.0656
	Δp_0^{SS}	-0.0123 ± 0.0060
	Δp_1^{SS}	0.1374 ± 0.0757
	$\epsilon_{tag}^{\text{SS}}$	0.6804 ± 0.0023
	$\Delta \epsilon_{tag}^{\text{SS}}$	0.0076 ± 0.0083
	A_p	-0.0042 ± 0.0091
	Δm_s	$\text{xx.xx} \pm 0.0110$

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

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

Fit parameter	Value
C	xx.xx ± 0.170
D	xx.xx ± 0.390
\bar{D}	xx.xx ± 0.346
S	xx.xx ± 0.255
\bar{S}	xx.xx ± 0.221

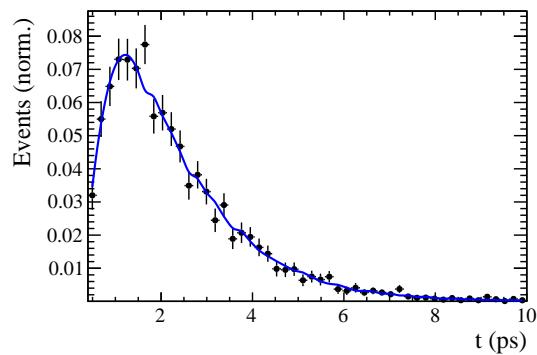


Figure 9.2

749 10 Time dependent amplitude fit

750 10.1 Signal Model Construction

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

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

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

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

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

767 where N_{Sig} is the number of signal events and r is the number of amplitudes with a decay
 768 fraction above a certain threshold. In this way, the optimal λ balances the fit quality
 769 ($-2 \log \mathcal{L}$) against the model complexity. The LASSO penalty term is only used to select
 770 the model. Afterwards, this term must be discarded in the final amplitude fit with the
 771 selected model, otherwise the parameter uncertainties would be biased.

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

- 776 1. The fit fraction threshold for inclusion in the final model was varied within the
 777 interval [0.05, 5]%. The set of selected amplitudes is stable for thresholds between
 778 0.1% and 1%. Other choices result in marginally different models containing one
 779 component more or less.
 - 780 2. Instead of BIC, the Akaike information criteria ($\text{AIC}(\lambda) = -2 \log \mathcal{L} + 2r$ [47]) was
 781 used to optimize λ . For a given threshold, the AIC method tends to prefer lower
 782 λ values. However, the set of models obtained varying the threshold within the
 783 interval [0.05, 5]% is identical to the BIC method.
 - 784 3. The amplitudes selected under nominal conditions were excluded one-by-one from
 785 the set of all amplitudes considered.
- 786 From that we obtained a set of alternative models shown in Appendix ??.

Table 10.1: Fit fractions for $B_s \rightarrow D_s K\pi\pi$ data.

Decay channel	Fraction [%]
$B_s \rightarrow K(1)(1400)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	34.70 ± 2.24
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K(0)^*(1430)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	6.85 ± 0.94
$B_s \rightarrow K(1)(1270)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	13.08 ± 1.70
$B_s \rightarrow K(1)(1270)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	9.25 ± 0.60
$B_s \rightarrow K(1460)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	0.99 ± 0.06
$BuggB_s \rightarrow K(1460)^+ (\rightarrow \sigma (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	3.42 ± 1.49
$B_s \rightarrow K^*(1410)^+ (\rightarrow K^*(892)^0 (\rightarrow K^+ \pi^-) \pi^+) D_s^-$	16.40 ± 1.06
$B_s \rightarrow K^*(1410)^+ (\rightarrow \rho(770)^0 (\rightarrow \pi^+ \pi^-) K^+) D_s^-$	4.88 ± 0.68
$B_s \rightarrow NonResS0 (\rightarrow D_s^- \pi^+) K^*(892)^0 (\rightarrow K^+ \pi^-)$	4.60 ± 1.44
$BuggB_s \rightarrow NonResS0 (\rightarrow D_s^- K^+) \sigma (\rightarrow \pi^+ \pi^-)$	4.96 ± 0.68
Sum	99.13 ± 5.87

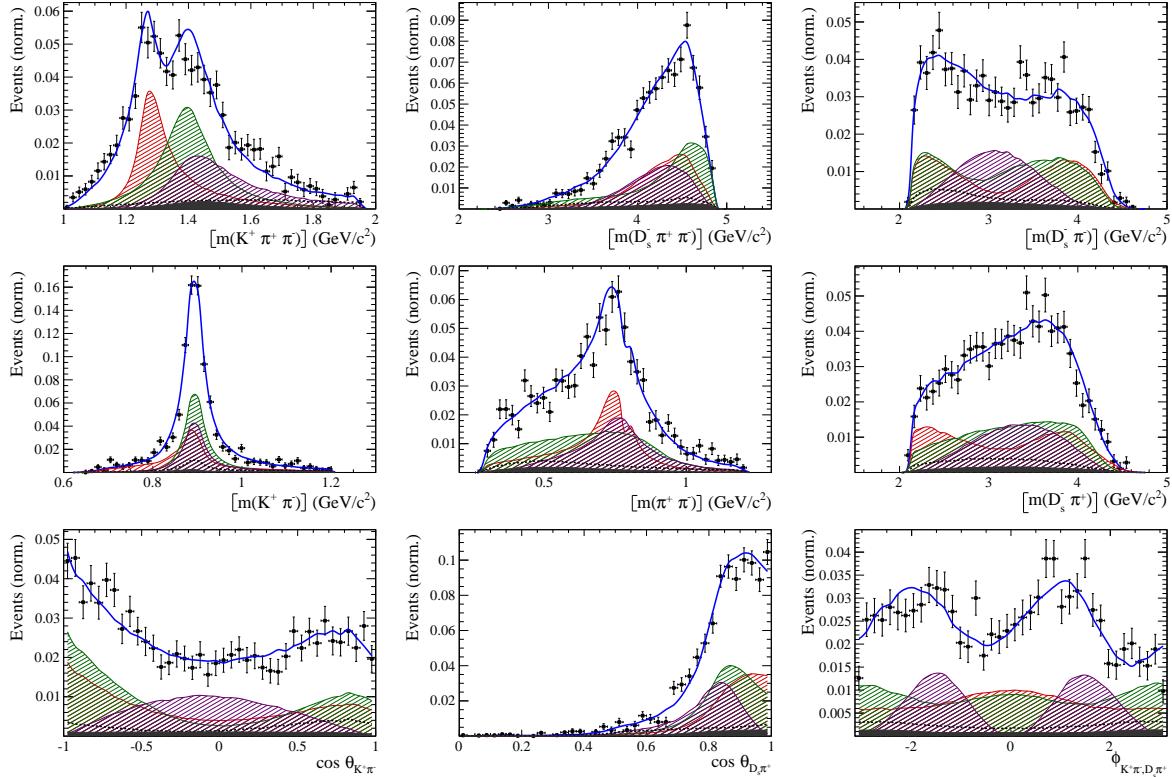


Figure 10.1

10.2 Results

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

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

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

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

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

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

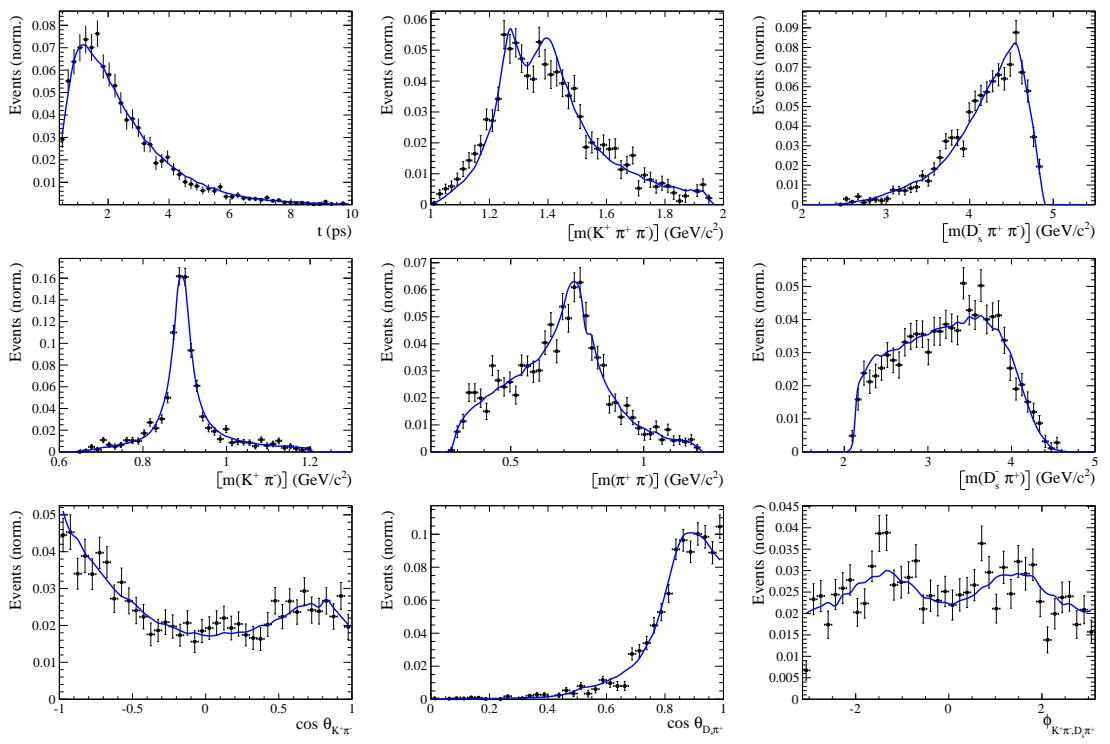


Figure 10.2

788 A Stripping and Trigger cuts

789 The following text describes variables which are used in Table 1.1 and might be ambiguous,
 790 or which benefits are not straight forward. Where noted, different cut values are applied
 791 for Run-I & Run-II data. In Table 1.1, DOCA is the abbreviation for distance of closest
 792 approach. This variable is used to ensure that all D_s and $X_{s,d}$ daughters originate from
 793 the same vertex. DIRA is the abbreviation for the cosine of the angle θ between the
 hadron's flight direction \vec{x} and it's corresponding momentum vector \vec{p} , $\cos \theta_{\vec{x}-\vec{p}}$.

Table 1.1: Summary of the stripping selections for $B_s^0 \rightarrow D_s K \pi \pi$ decays. The requirements for Run-II are only given if they have been changed.

Variable	Stripping Cut (Run-I)	Stripping Cut (Run-II)
Track χ^2/nDoF	< 3	
Track p	> 1000 MeV/ c	
Track p_T	> 100 MeV/ c	
Track IP χ^2	> 4	
Track ghost-prob.	< 0.4	
D_s mass	$m_{D_s} \pm 100$ MeV	$m_{D_s} \pm 80$ MeV
D_s Daughter p_T	$\sum_{i=1}^3 p_{t,i} > 1800$ MeV/ c	
D_s Daughter DOCA	< 0.5 mm	
D_s Vertex χ^2/nDoF	< 10	
D_s χ^2 -separation from PV	> 36	
D_s daughter PID(π)	< 20	
D_s daughter PID(K)	> -10	
$X_{s,d}$ mass	< 4000 MeV	< 3500 MeV
$X_{s,d}$ Daughter p	> 2 GeV/ c	
$X_{s,d}$ Daughter DOCA	< 0.4 mm	
$X_{s,d}$ Daughter p_T	$\sum_{i=1}^3 p_{t,i} > 1250$ MeV/ c	
$X_{s,d}$ Vertex χ^2/nDoF	< 8	
$X_{s,d}$ χ^2 -separation from PV	> 16	
$X_{s,d}$ DIRA	> 0.98	
$X_{s,d}$ $\Delta\rho$	> 0.1 mm	
$X_{s,d}$ Δz	> 2.0 mm	
$X_{s,d}$ daughter PID(π)	< 10	
X_s daughter PID(K)	> -2	> 4
B_s^0 mass	[4750, 7000] MeV/ c^2	[5000, 6000] MeV/ c^2
B_s^0 DIRA	> 0.98	> 0.99994
B_s^0 min IP χ^2	< 25	< 20
B_s^0 Vertex χ^2/nDoF	< 10	< 8
$B_s^0 \tau_{B_s^0}$	> 0.2 ps	

794

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

800 The Hlt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the
 801 b-hadron p_T , its flight distance χ^2 from the nearest PV and the sum of the B_s^0 and D_s
 802 vertex χ^2 divided by the sum of their number of degrees of freedom. Table 1.3 summarizes
 803 the cuts applied by the inclusive ϕ trigger, which requires that a $\phi \rightarrow KK$ candidate can
 be formed out of two tracks present in the event.

Table 1.2: Summary of the cuts applied by the Hlt1TrackAllL0 trigger for Run-I. At least one of the six decay particles must pass this requirements, in order for the event to be accepted.

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

Table 1.3: Summary of the cuts applied by the Hlt2 inclusive ϕ trigger. A $\phi \rightarrow KK$ candidate, formed by two tracks in the event, must pass this requirements in order for the event to be accepted.

Quantity	Hlt2IncPhi requirement
ϕ mass	$m_\phi \pm 12$ MeV/ c^2 of PDG value
ϕp_T	> 2.5 GeV/c
ϕ vertex χ^2/nDoF	< 20
ϕ IP χ^2 to any PV	> 5

B Details of multivariate classifier

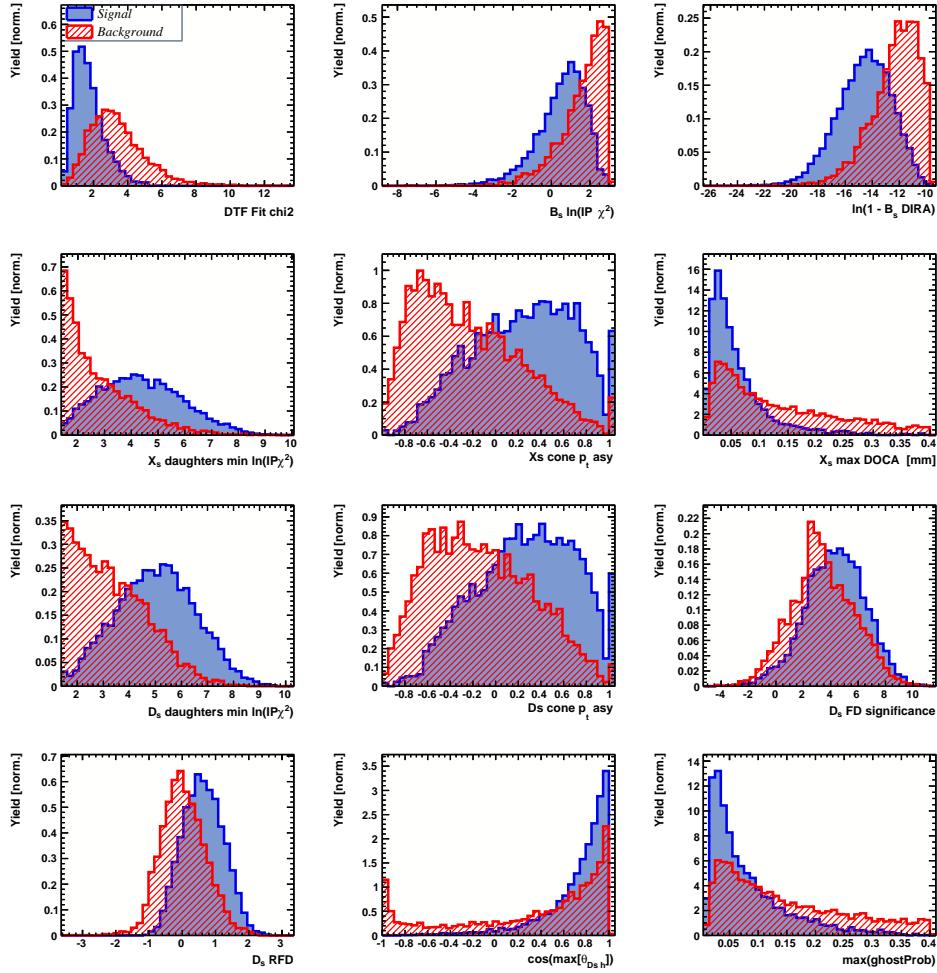


Figure A.1: Variables used to train the BDTG.

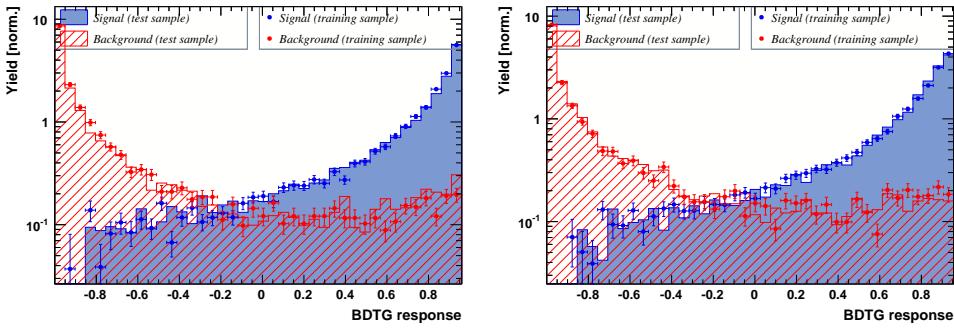


Figure A.2: Response of the classifier trained on the even (left) and odd (right) sample.

806 C Detailed mass fits

807 In this section, all fits to the mass distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ and $B_s^0 \rightarrow D_sK\pi\pi$
 808 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-
 809 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
 810 L0 trigger category.

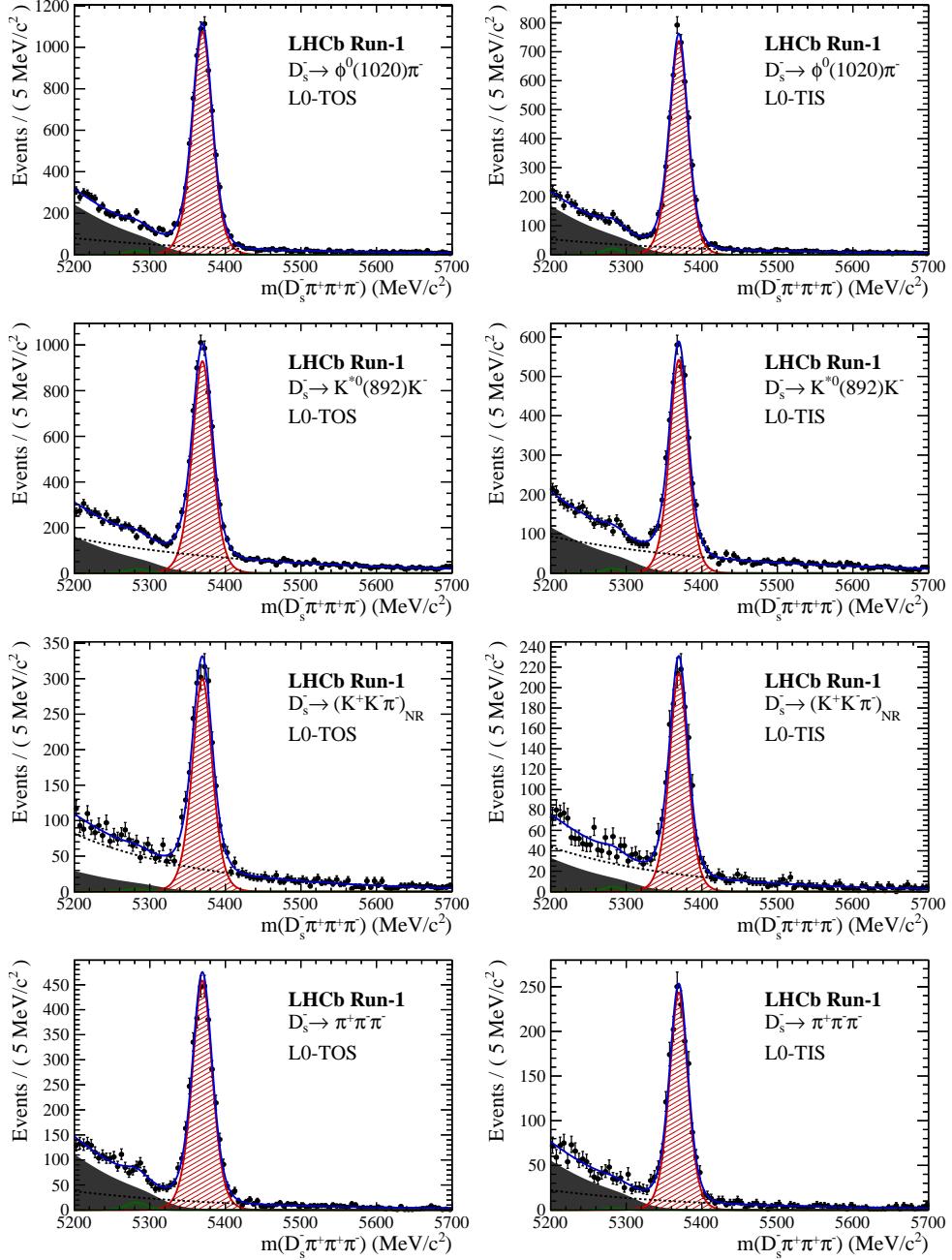


Figure B.1: Invariant mass distributions of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates for Run-I data.

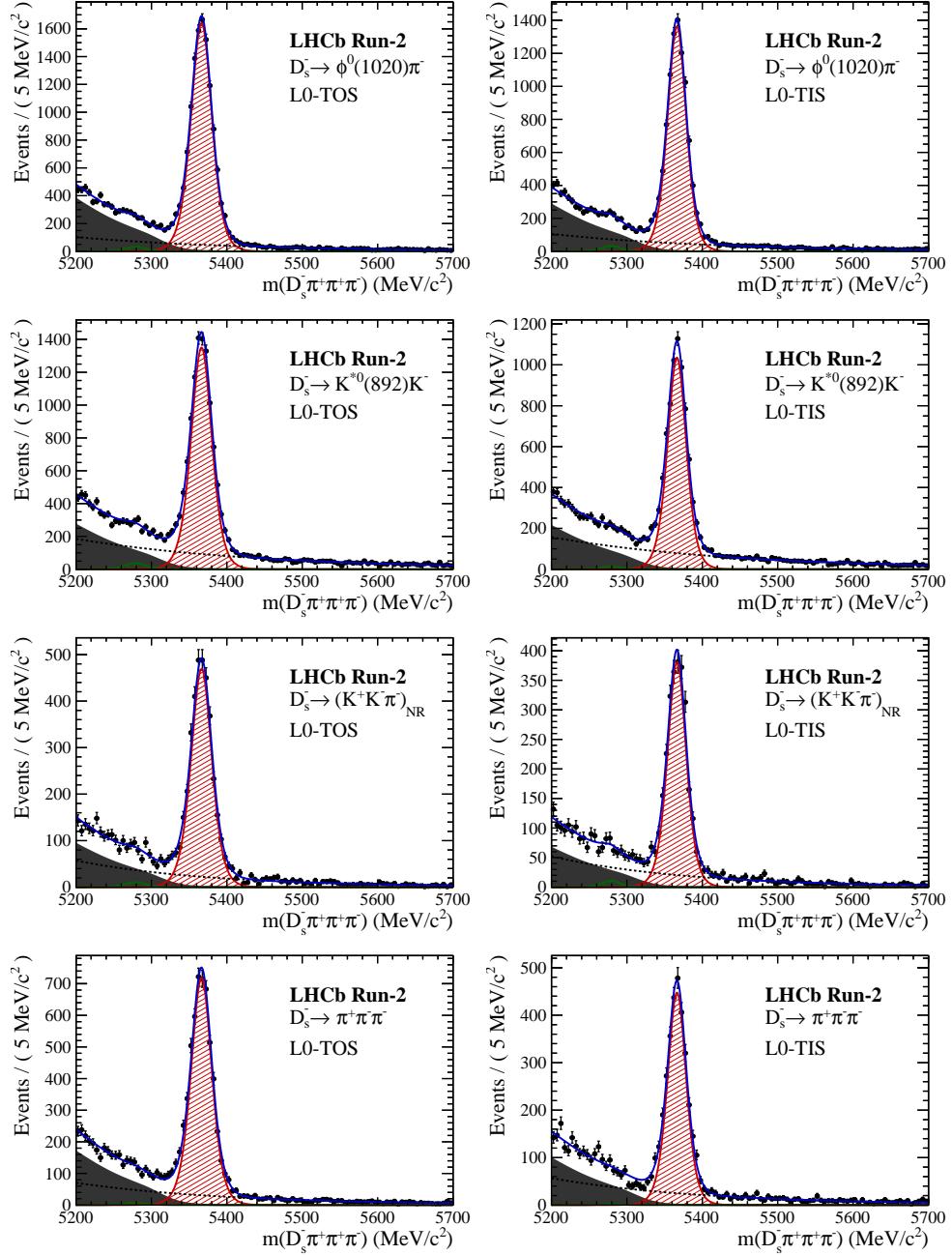


Figure B.2: Invariant mass distributions of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates for Run-II data.

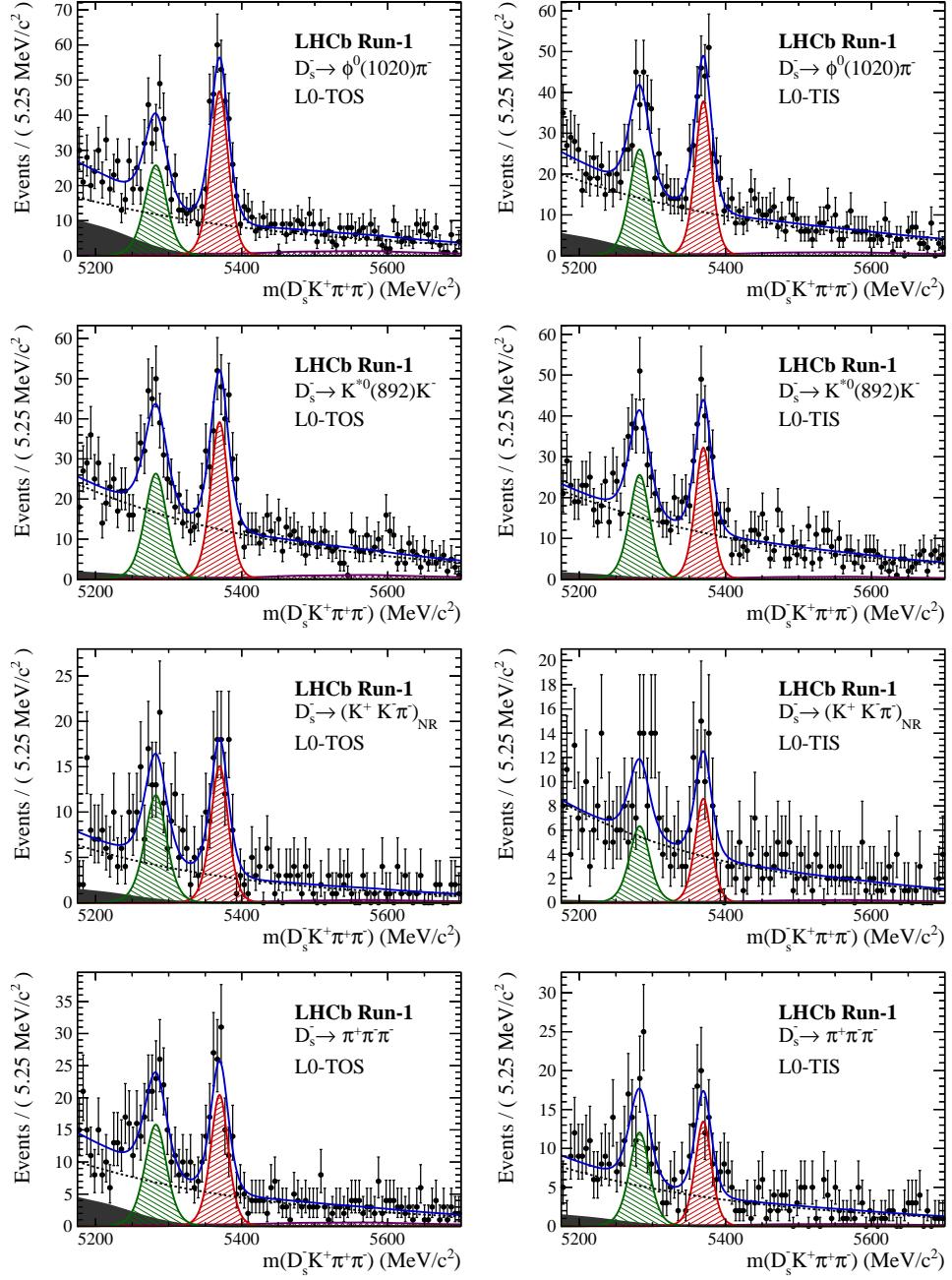


Figure B.3: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-I data.

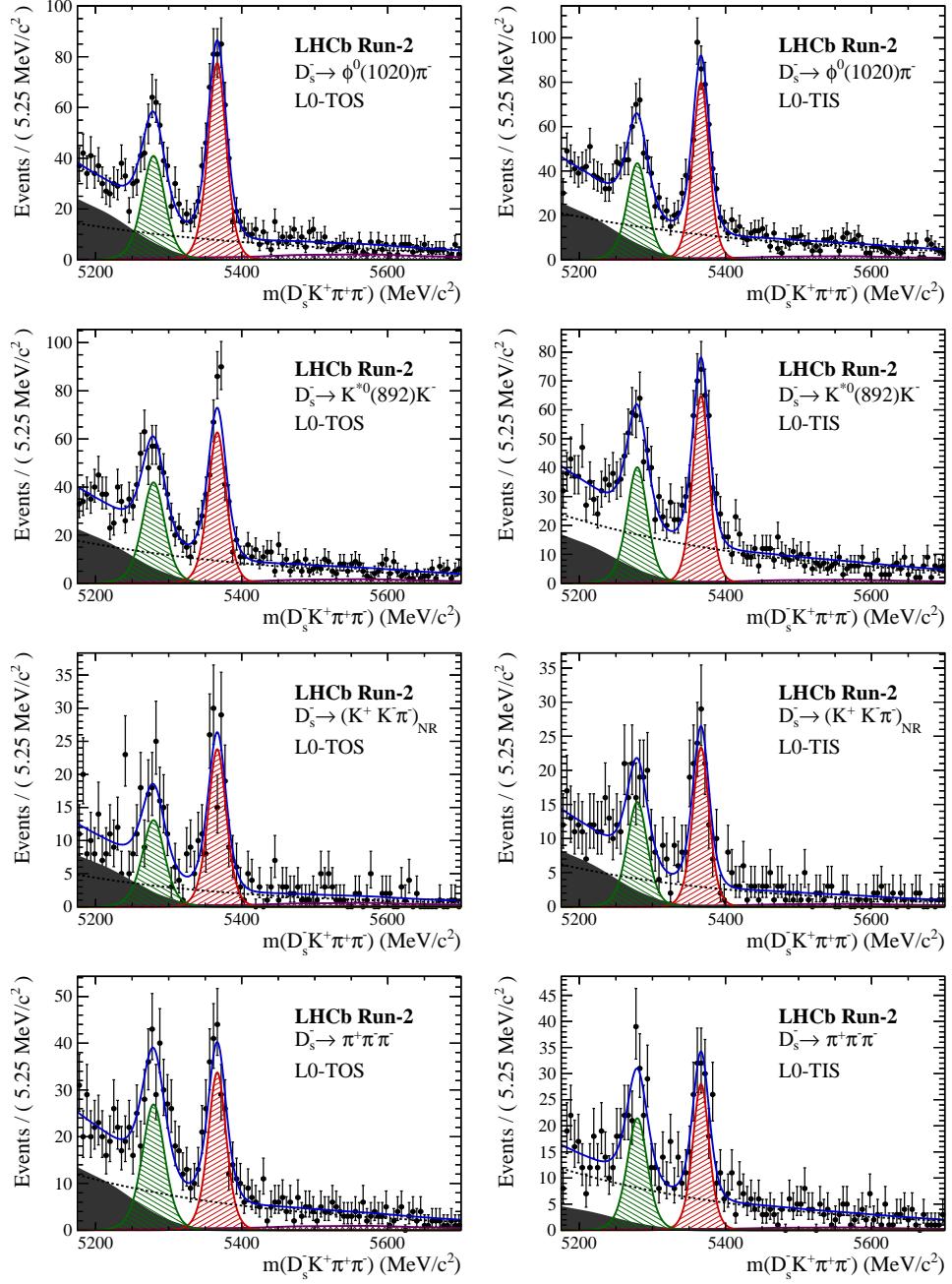


Figure B.4: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-II data.

811 D Decay-time Resolution fits

812 This section contains all fits to the distributions of the decay time difference Δt between
 813 the true and the reconstructed decay time of the truth-matched B_s^0 candidates on MC.
 814 The fits are performed in bins of the decay time error σ_t , where an adaptive binning
 815 scheme is used to ensure that approximately the same number of events are found in each
 816 bin.

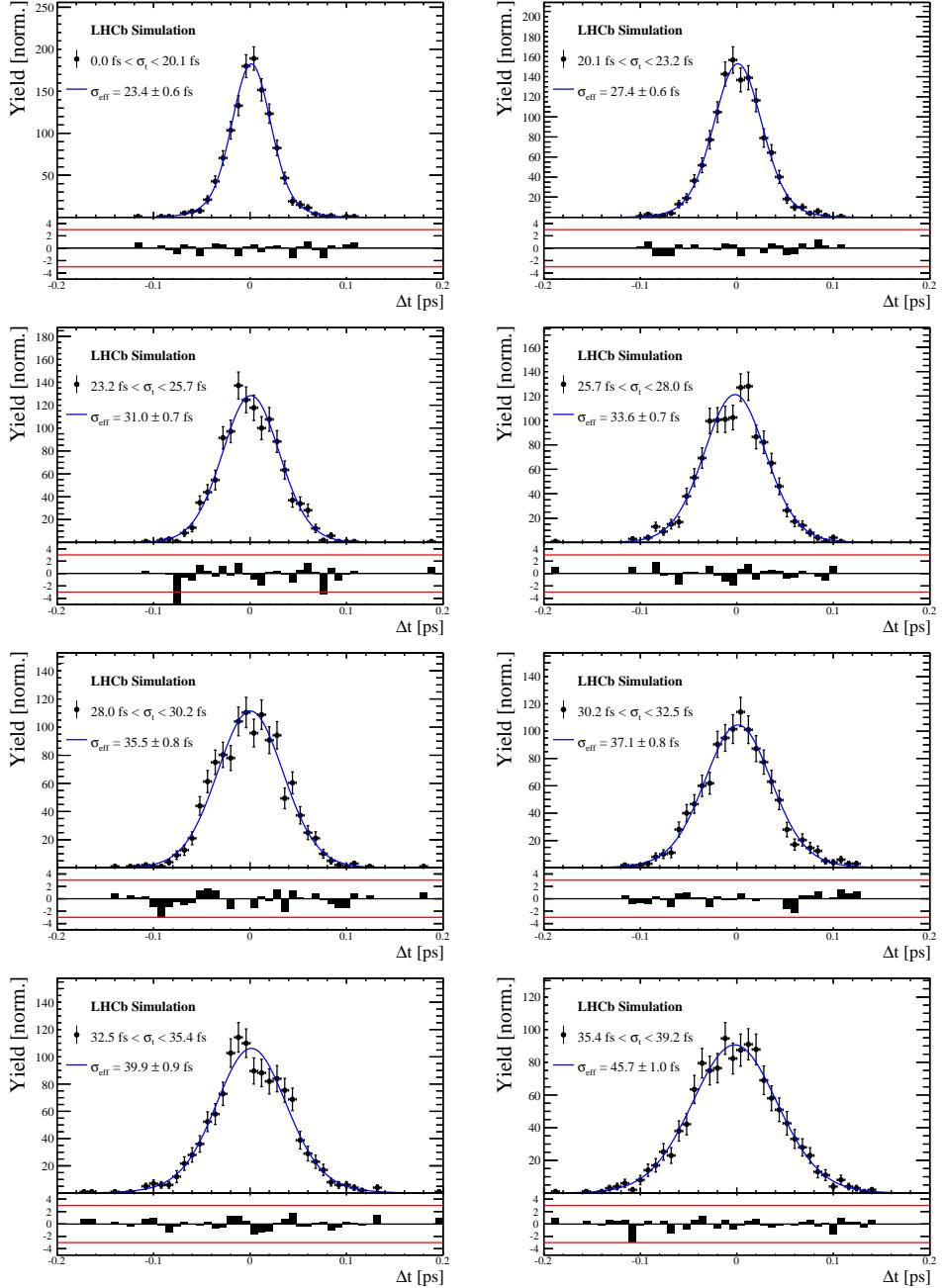


Figure C.1: Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K\pi\pi$ MC candidates in bins of the per-event decay time error estimate..

σ_t Bin [fs]	σ_1 [fs]	σ_2 [fs]	f_1	D	σ_{eff} [fs]
0.0 - 20.1	19 ± 0.675	33.8 ± 1.77	0.75 ± 0	0.917 ± 0.00406	23.4 ± 0.599
20.1 - 23.2	23.4 ± 0.86	37.4 ± 1.95	0.75 ± 0	0.888 ± 0.00477	27.4 ± 0.621
23.2 - 25.7	28.1 ± 1.02	38.7 ± 2.32	0.75 ± 0	0.86 ± 0.00563	31 ± 0.671
25.7 - 28.0	30.1 ± 1.12	43.2 ± 2.56	0.75 ± 0	0.837 ± 0.00651	33.6 ± 0.734
28.0 - 30.2	32.4 ± 1.12	44.2 ± 2.59	0.75 ± 0	0.819 ± 0.00694	35.5 ± 0.756
30.2 - 32.5	32.6 ± 1.38	49.2 ± 3.04	0.75 ± 0	0.805 ± 0.00792	37.1 ± 0.841
32.5 - 35.4	34.4 ± 1.19	54.7 ± 2.85	0.75 ± 0	0.778 ± 0.0086	39.9 ± 0.879
35.4 - 39.2	41.9 ± 1.8	56.9 ± 4.18	0.75 ± 0	0.719 ± 0.00997	45.7 ± 0.962
39.2 - 44.7	42.2 ± 1.56	68.1 ± 4.01	0.75 ± 0	0.687 ± 0.0114	48.8 ± 1.08
44.7 - 120.0	55.5 ± 2.59	83 ± 14.7	0.75 ± 0	0.546 ± 0.0521	62 ± 4.89

Table 4.1: Measured time resolution for $B_s \rightarrow D_s K\pi\pi$ MC in bins of the per-event decay time error estimate.

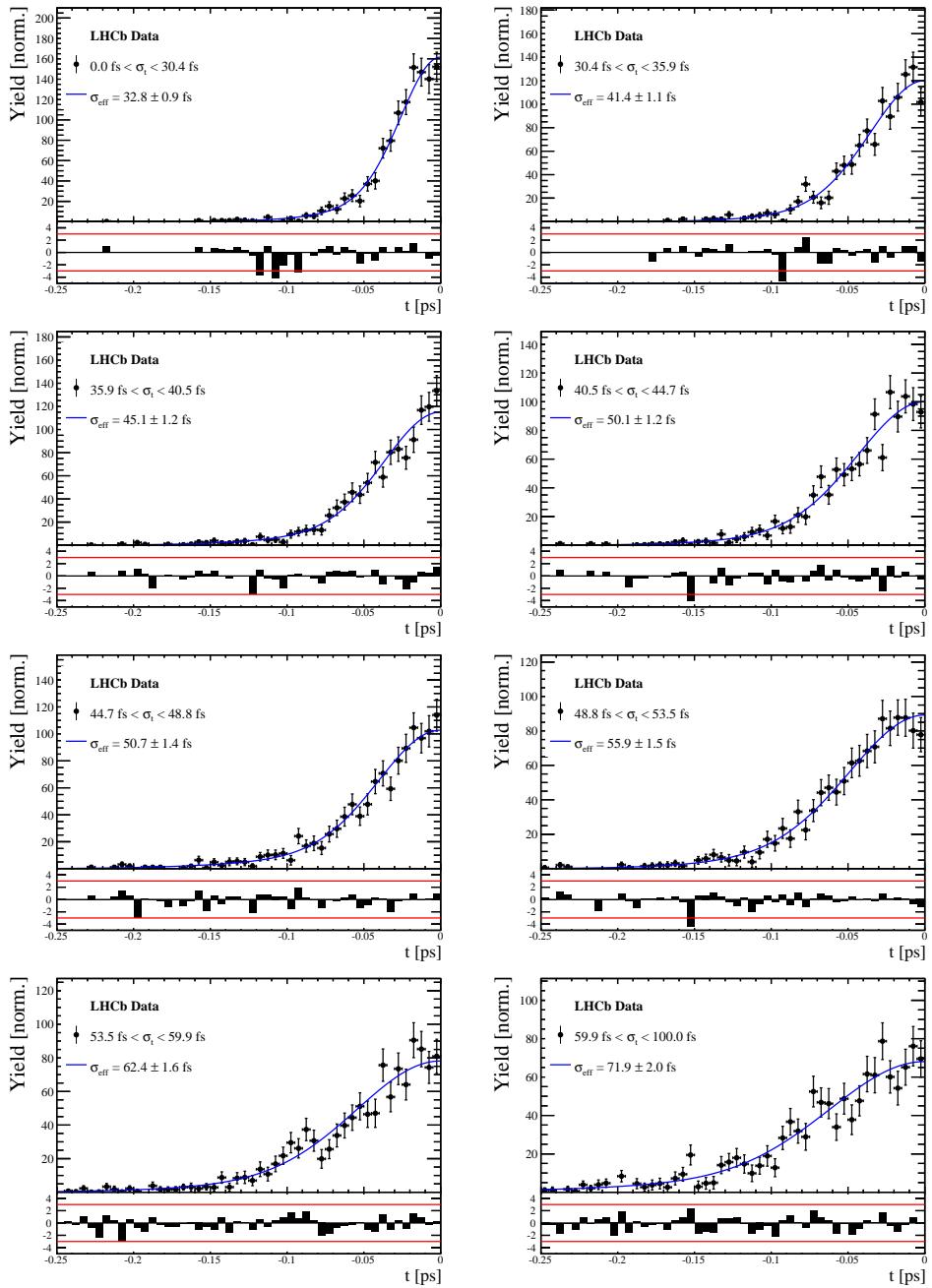


Figure C.2: Decay-time distribution for fake B_s candidates from promptly produced D_s candidates, combined with random prompt $K\pi\pi$ bachelor tracks, for bins in the per-event decay time error estimate.

σ_t Bin [fs]	σ_1 [fs]	σ_2 [fs]	f_1	D	σ_{eff} [fs]
0.0 - 30.4	25.4 ± 1.03	50.7 ± 2.77	0.75 ± 0	0.844 ± 0.00822	32.8 ± 0.942
30.4 - 35.9	34.5 ± 1.46	60.2 ± 3.48	0.75 ± 0	0.763 ± 0.0108	41.4 ± 1.08
35.9 - 40.5	35.6 ± 1.35	71.3 ± 3.84	0.75 ± 0	0.726 ± 0.0121	45.1 ± 1.18
40.5 - 44.7	42.3 ± 1.65	73.3 ± 4.21	0.75 ± 0	0.673 ± 0.0132	50.1 ± 1.24
44.7 - 48.8	39.6 ± 1.64	84.8 ± 5.07	0.75 ± 0	0.666 ± 0.0145	50.7 ± 1.36
48.8 - 53.5	47.6 ± 1.94	82.4 ± 5.48	0.75 ± 0	0.611 ± 0.0157	55.9 ± 1.46
53.5 - 59.9	53 ± 2.15	95.3 ± 6.84	0.75 ± 0	0.541 ± 0.0174	62.4 ± 1.63
59.9 - 100.0	60.5 ± 2.8	125 ± 14	0.75 ± 0	0.443 ± 0.0204	71.9 ± 2.03

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

817 E Spin Amplitudes

818 The spin factors used for $B \rightarrow P_1 P_2 P_3 P_4$ decays are given in Table 5.1.

Table 5.1: Spin factors for all topologies considered in this analysis. In the decay chains, S , P , V , A , T and PT stand for scalar, pseudoscalar, vector, axial vector, tensor and pseudotensor, respectively. If no angular momentum is specified, the lowest angular momentum state compatible with angular momentum conservation and, where appropriate, parity conservation, is used.

Number	Decay chain	Spin amplitude
1	$B \rightarrow (P P_1)$, $P \rightarrow (S P_2)$, $S \rightarrow (P_3 P_4)$	1
2	$B \rightarrow (P P_1)$, $P \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(P) L_{(1)}^\alpha(V)$
3	$B \rightarrow (A P_1)$, $A \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) P_{(1)}^{\alpha\beta}(A) L_{(1)\beta}(V)$
4	$B \rightarrow (A P_1)$, $A[D] \rightarrow (P_2 V)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(2)}^{\alpha\beta}(A) L_{(1)\beta}(V)$
5	$B \rightarrow (A P_1)$, $A \rightarrow (S P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(1)}^\alpha(A)$
6	$B \rightarrow (A P_1)$, $A \rightarrow (T P_2)$, $T \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(1)\beta}(A) L_{(2)}^{\alpha\beta}(T)$
7	$B \rightarrow (V_1 P_1)$, $V_1 \rightarrow (V_2 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$L_{(1)\mu}(B) P_{(1)}^{\mu\alpha}(V_1) \epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\beta(V_1) p_{V_1}^\gamma L_{(1)}^\delta(V_2)$
8	$B \rightarrow (PT P_1)$, $PT \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) P_{(2)}^{\alpha\beta\gamma\delta}(PT) L_{(1)\gamma}(PT) L_{(1)\delta}(V)$
9	$B \rightarrow (PT P_1)$, $PT \rightarrow (S P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}(PT)^{\alpha\beta}$
10	$B \rightarrow (PT P_1)$, $PT \rightarrow (T P_2)$, $T \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) P_{(2)}^{\alpha\beta\gamma\delta}(PT) L_{(2)\gamma\delta}(T)$
11	$B \rightarrow (T P_1)$, $T \rightarrow (V P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(2)\mu\nu}(B) P_{(2)}^{\mu\nu\rho\alpha}(T) \epsilon_{\alpha\beta\gamma\delta} L_{(2)\rho}^\beta(T) p_T^\gamma P_{(1)}^{\delta\sigma}(T) L_{(1)\sigma}(V)$
12	$B \rightarrow (T_1 P_1)$, $T_1 \rightarrow (T_2 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\mu\nu}(B) P_{(2)}^{\mu\nu\rho\alpha}(T_1) \epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\beta(T_1) p_{T_1}^\gamma L_{(2)\rho}^\delta(T_2)$
13	$B \rightarrow (S_1 S_2)$, $S_1 \rightarrow (P_1 P_2)$, $S_2 \rightarrow (P_3 P_4)$	1
14	$B \rightarrow (V S)$, $V \rightarrow (P_1 P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(1)}^\alpha(V)$
15	$B \rightarrow (V_1 V_2)$, $V_1 \rightarrow (P_1 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(V_1) L_{(1)}^\alpha(V_2)$
16	$B[P] \rightarrow (V_1 V_2)$, $V_1 \rightarrow (P_1 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$\epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\alpha(B) L_{(1)}^\beta(V_1) L_{(1)}^\gamma(V_2) p_B^\delta$
17	$B[D] \rightarrow (V_1 V_2)$, $V_1 \rightarrow (P_1 P_2)$, $V_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(1)}^\alpha(V_1) L_{(1)}^\beta(V_2)$
18	$B \rightarrow (T S)$, $T \rightarrow (P_1 P_2)$, $S \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\beta}(T)$
19	$B \rightarrow (V T)$, $T \rightarrow (P_1 P_2)$, $V \rightarrow (P_3 P_4)$	$L_{(1)\alpha}(B) L_{(2)}^{\alpha\beta}(T) L_{(1)\beta}(V)$
20	$B[D] \rightarrow (TV)$, $T \rightarrow (P_1 P_2)$, $V \rightarrow (P_3 P_4)$	$\epsilon_{\alpha\beta\delta\gamma} L_2^{\alpha\mu}(B) L_{2\mu}^\beta L_{(1)}^\gamma(V) p_B^\delta$
21	$B \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(T_1) L_{(2)}^{\alpha\beta}(T_2)$
22	$B[P] \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$\epsilon_{\alpha\beta\gamma\delta} L_{(1)}^\alpha(B) L_{(2)}^{\beta\mu}(T_1) L_{(2)\mu}^\gamma(T_2) p_B^\delta$
23	$B[D] \rightarrow (T_1 T_2)$, $T_1 \rightarrow (P_1 P_2)$, $T_2 \rightarrow (P_3 P_4)$	$L_{(2)\alpha\beta}(B) L_{(2)}^{\alpha\gamma}(T_1) L_{(2)\gamma}^\beta(T_2)$

819 F Considered Decay Chains

820 The various decay channels considered in the model building are listed in Table 6.1.

Table 6.1: Decays considered in the LASSO model building.

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

821 G MC corrections

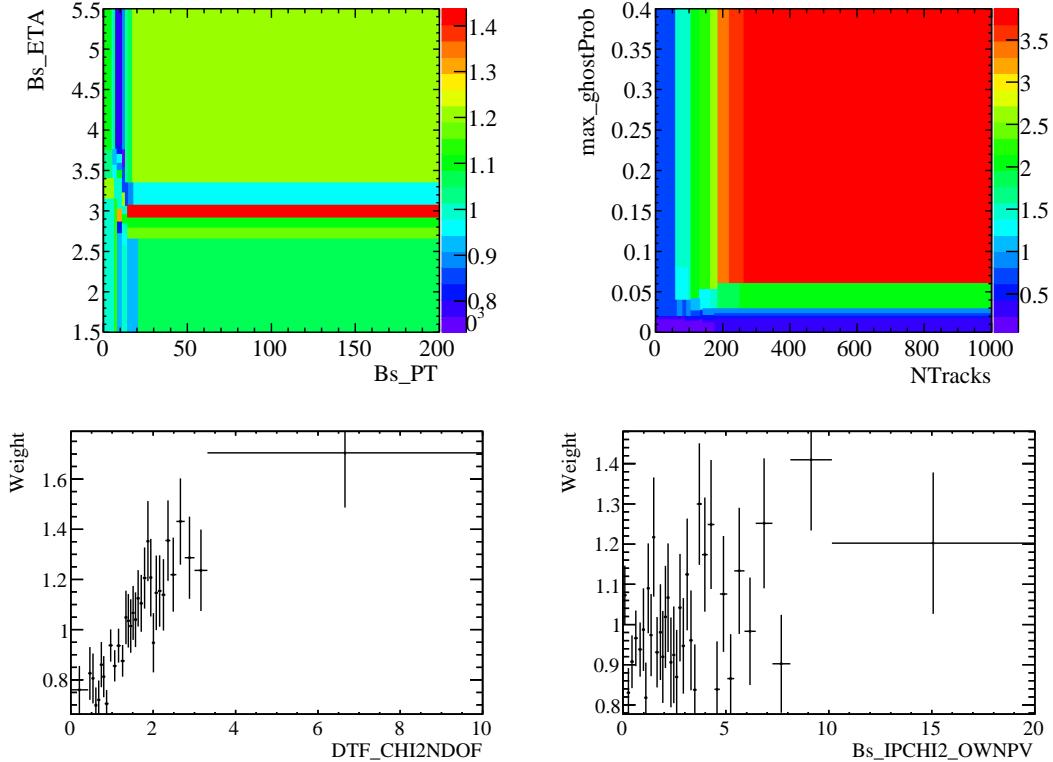


Figure C.1: Weights applied to correct for Data/MC differences.

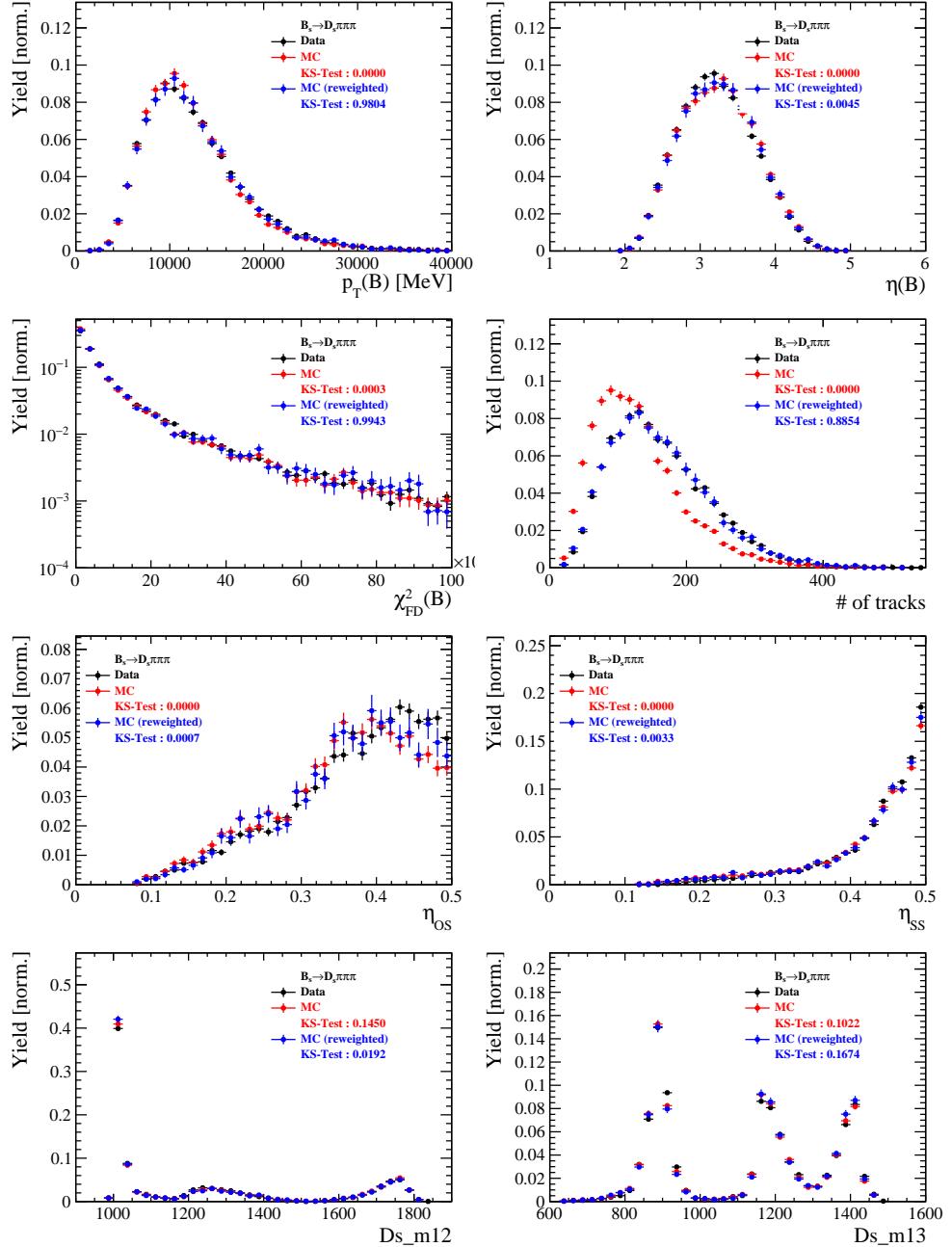


Figure C.2: Comparison of selected variables for $B_s \rightarrow D_s \pi\pi\pi$ decays.

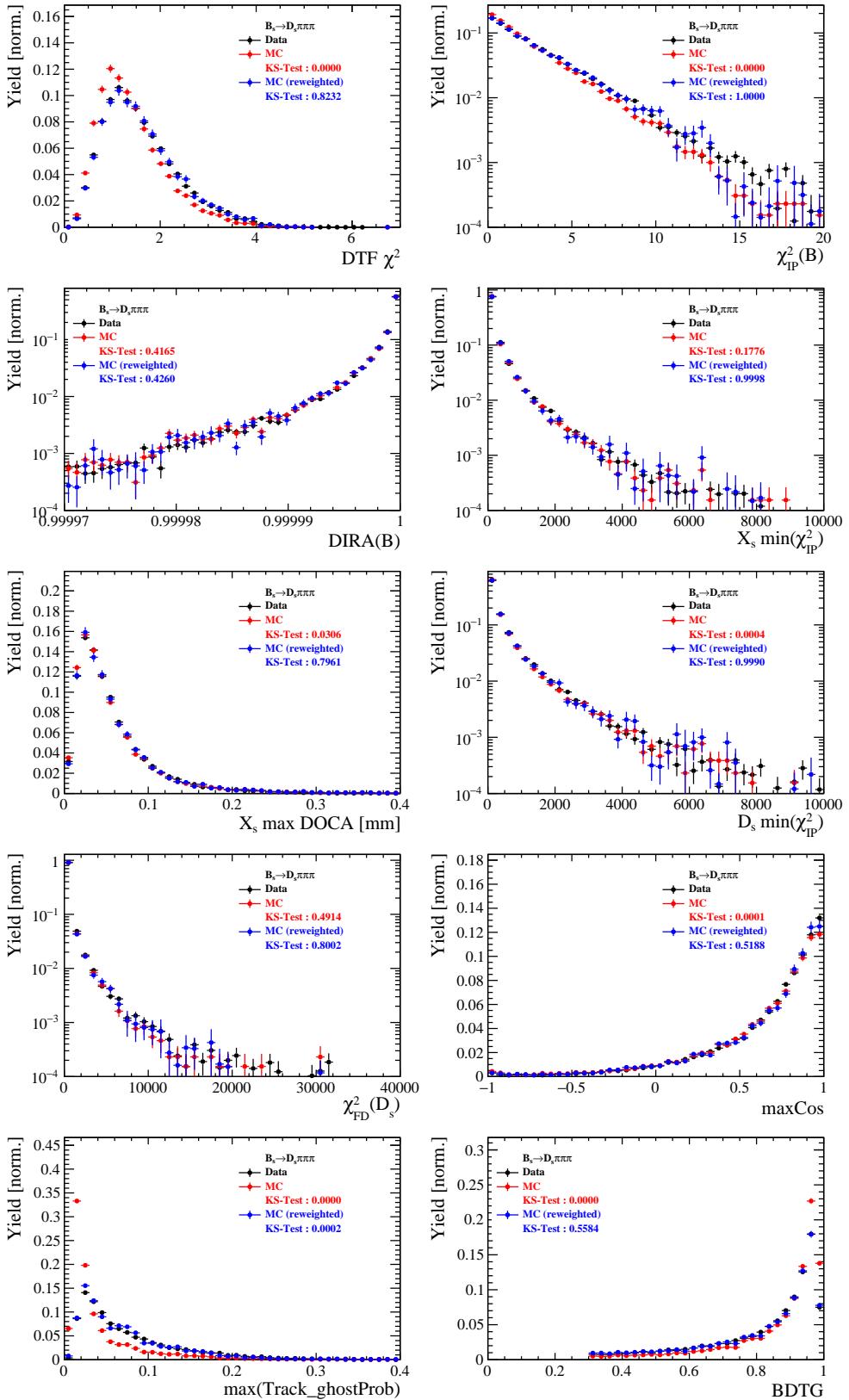


Figure C.3: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s\pi\pi\pi$ decays.

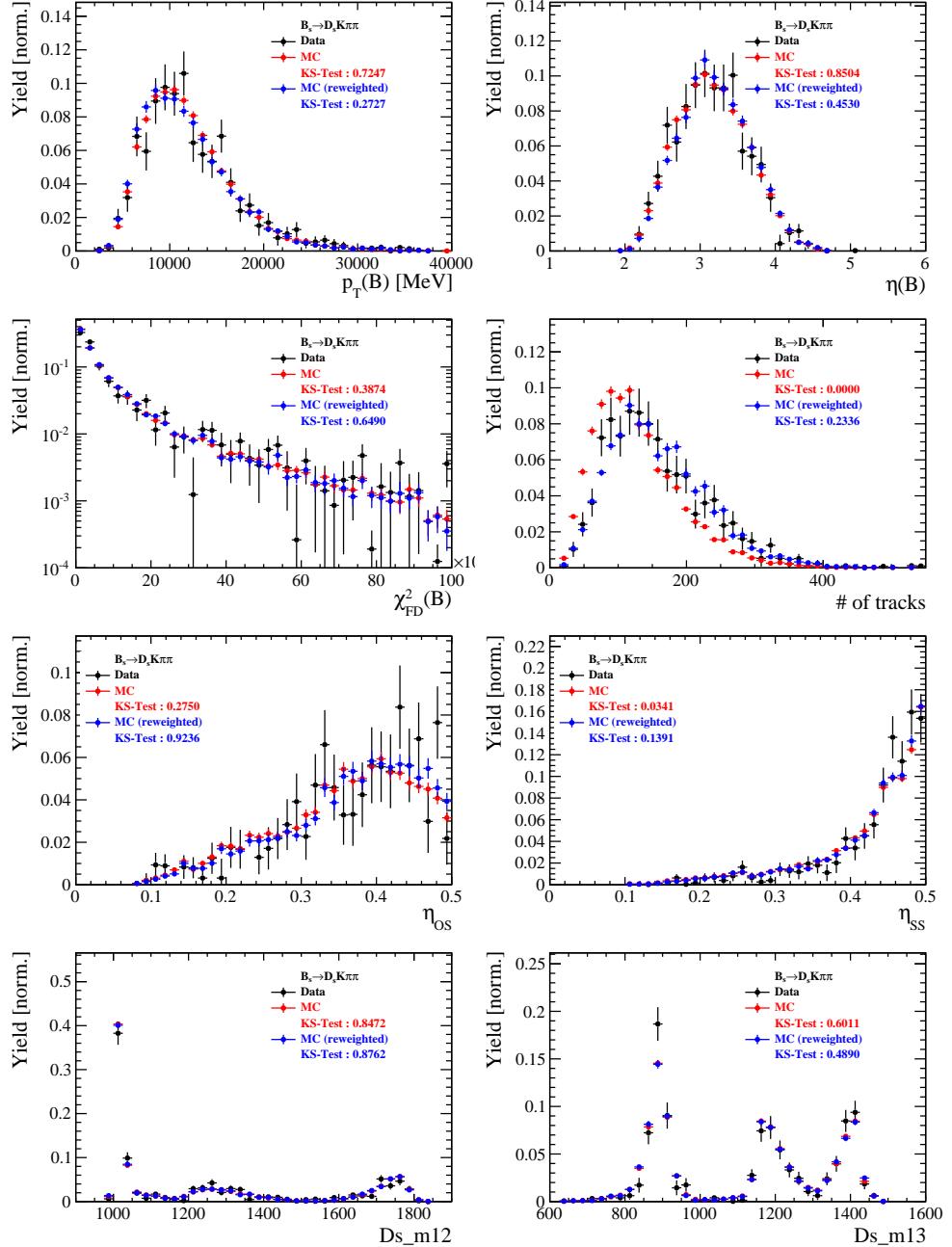


Figure C.4: Comparison of selected variables for $B_s \rightarrow D_s K\pi\pi$ decays.

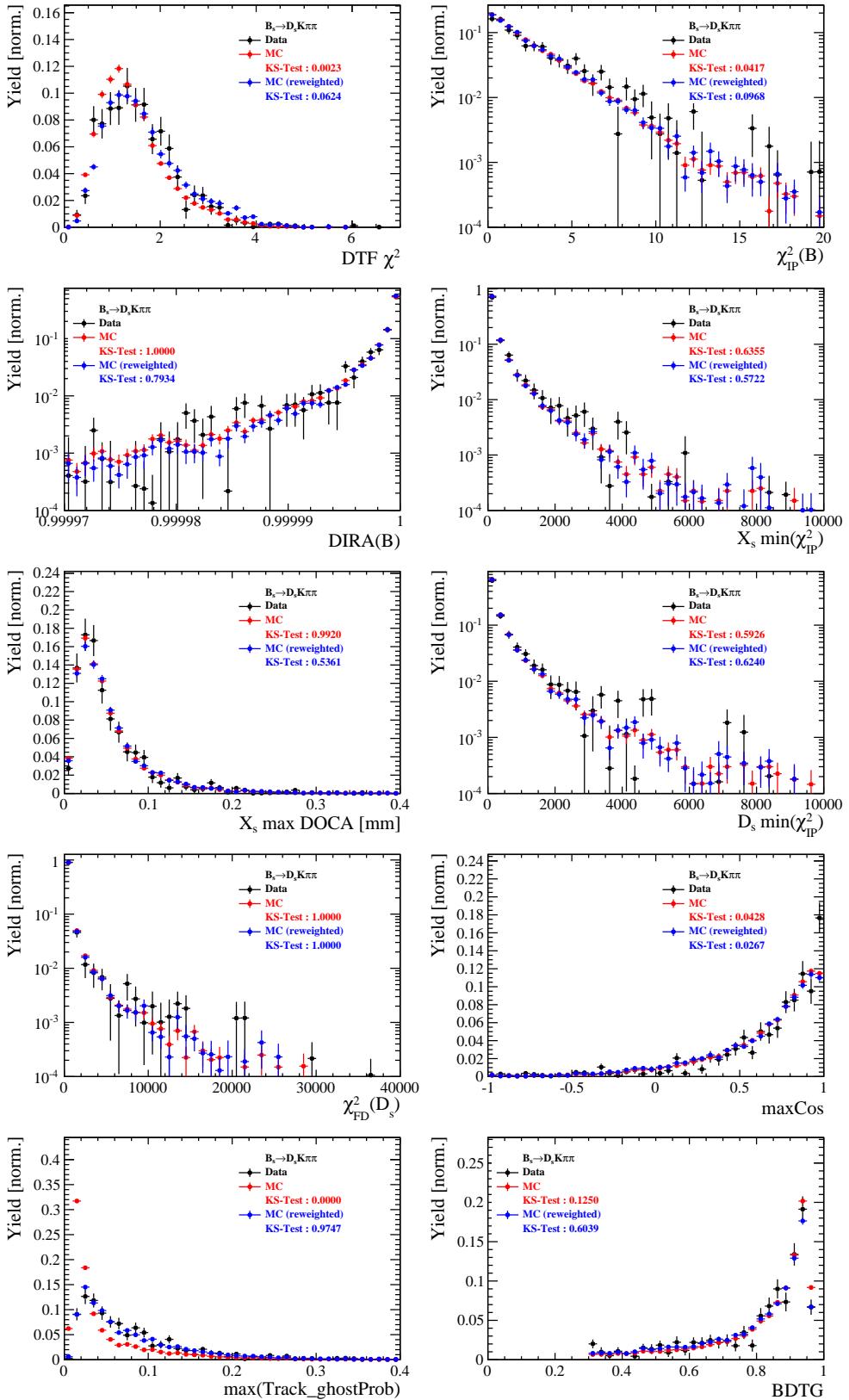


Figure C.5: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s K\pi\pi$ decays.

822 H Data distributions

823 H.1 Comparison of signal and calibration channel

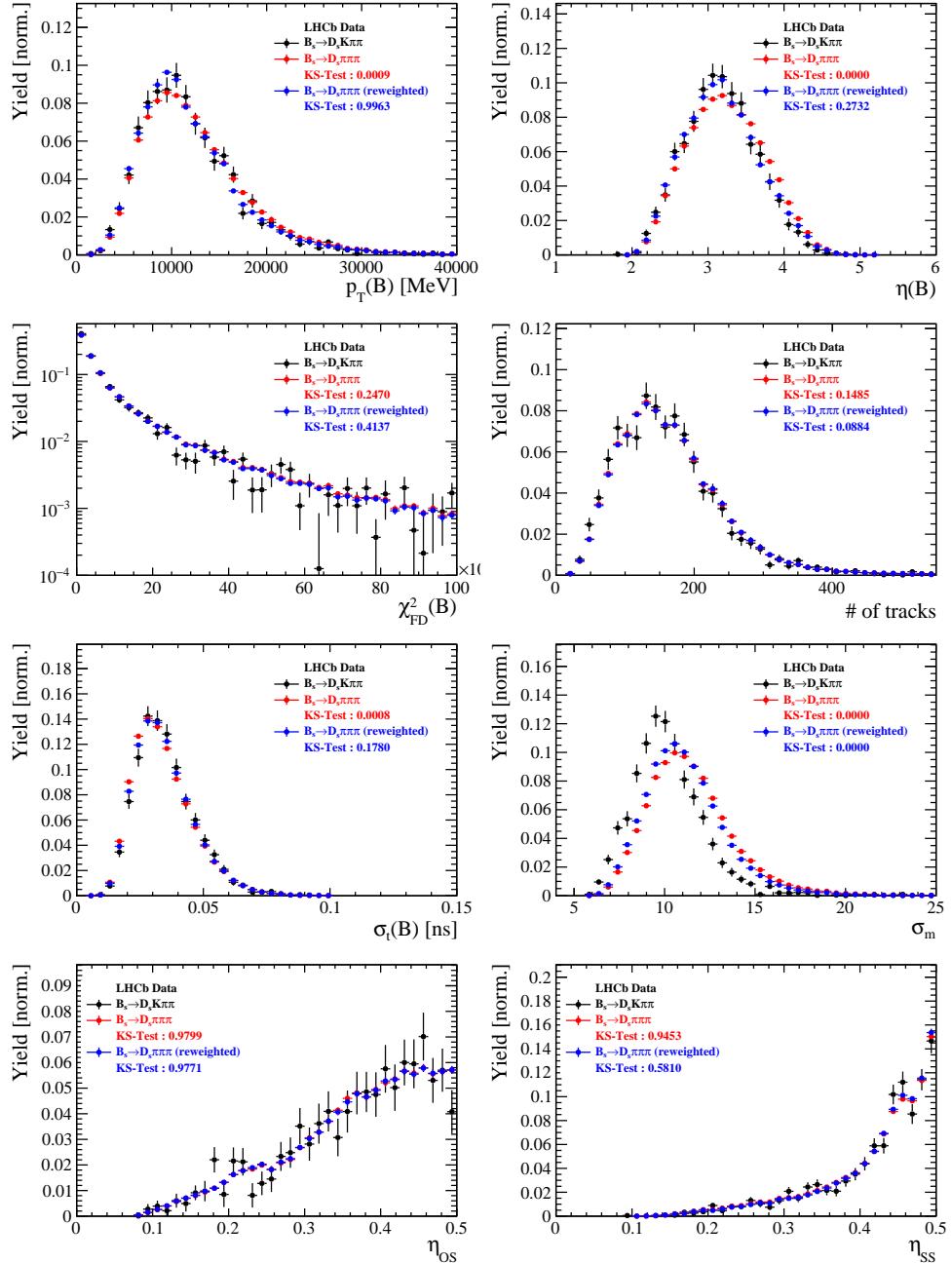


Figure C.1: Comparison of selected variables.

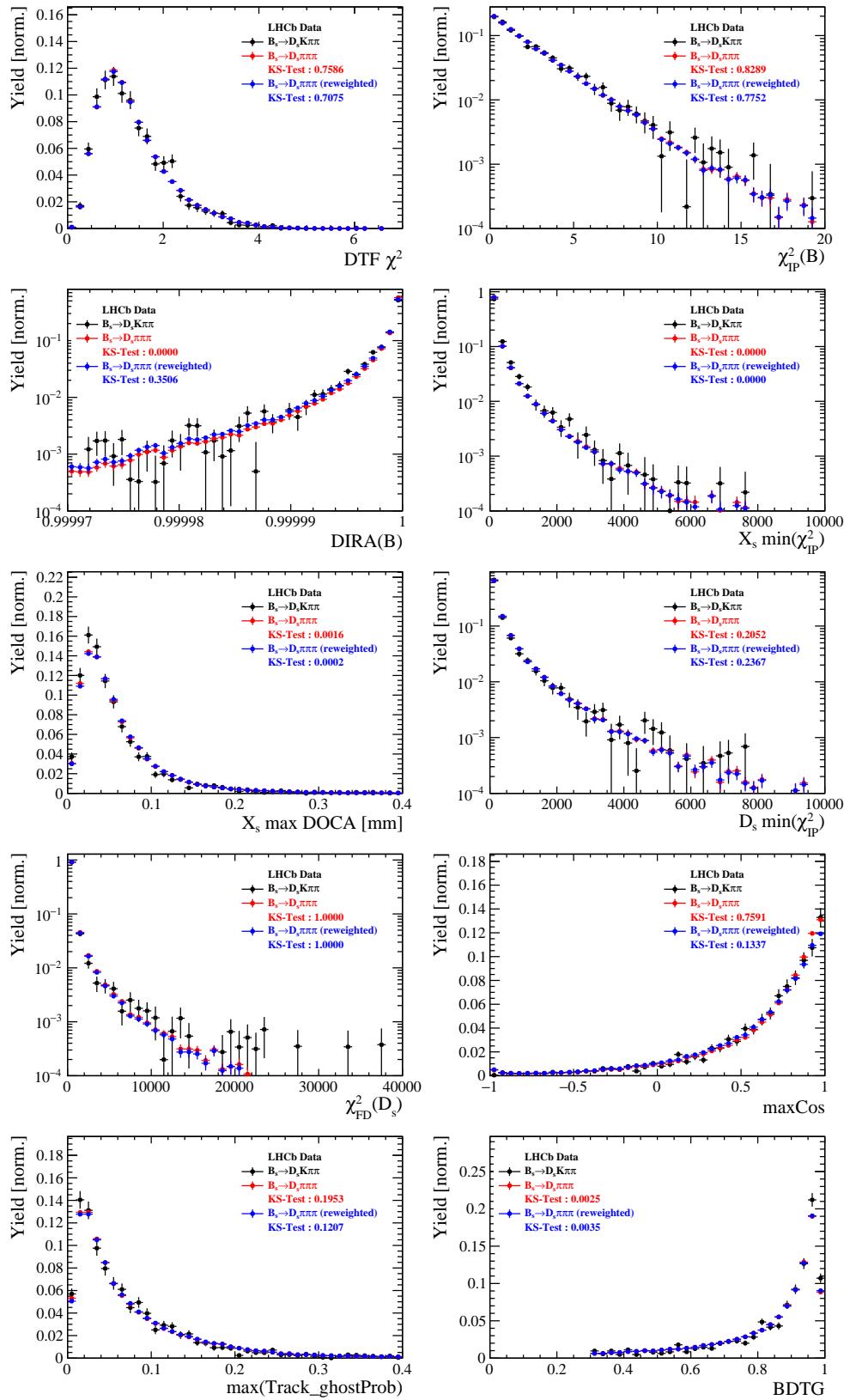


Figure C.2: Comparison of BDTG input variables and classifier response.

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

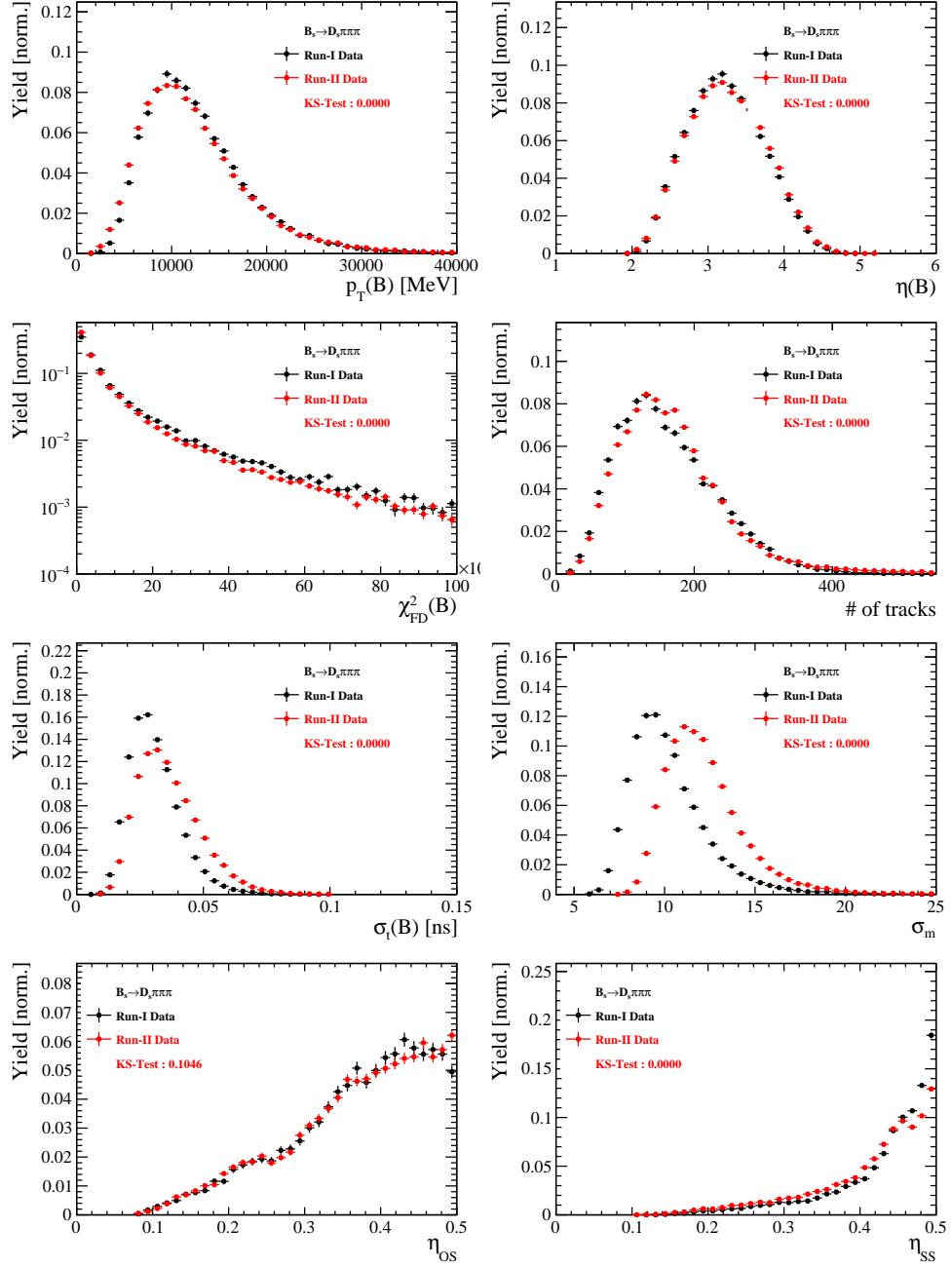


Figure C.3: Comparison of selected variables.

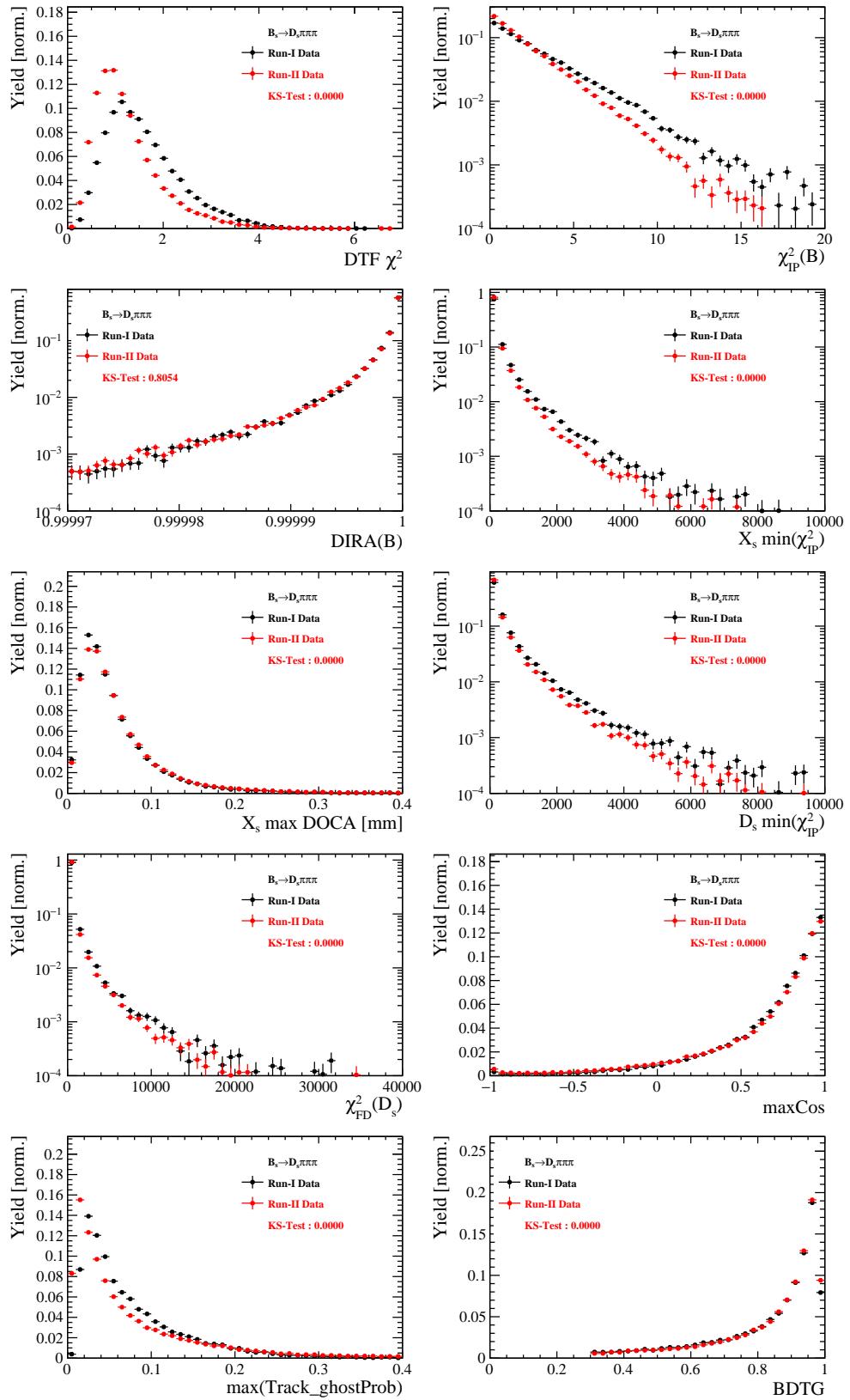


Figure C.4: Comparison of BDTG input variables and classifier response.

825 H.3 Comparison of D_s final states

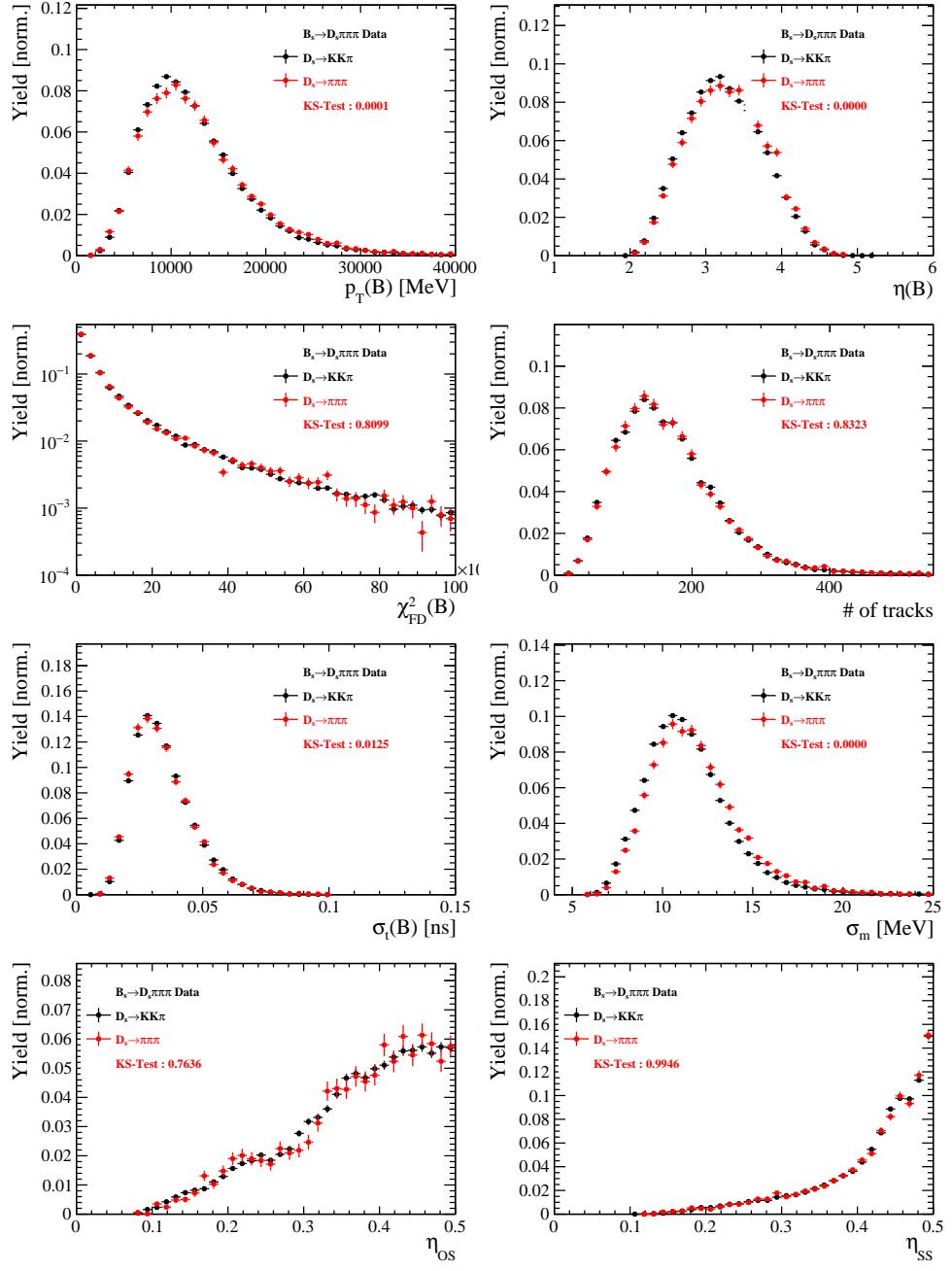


Figure C.5: Comparison of selected variables.

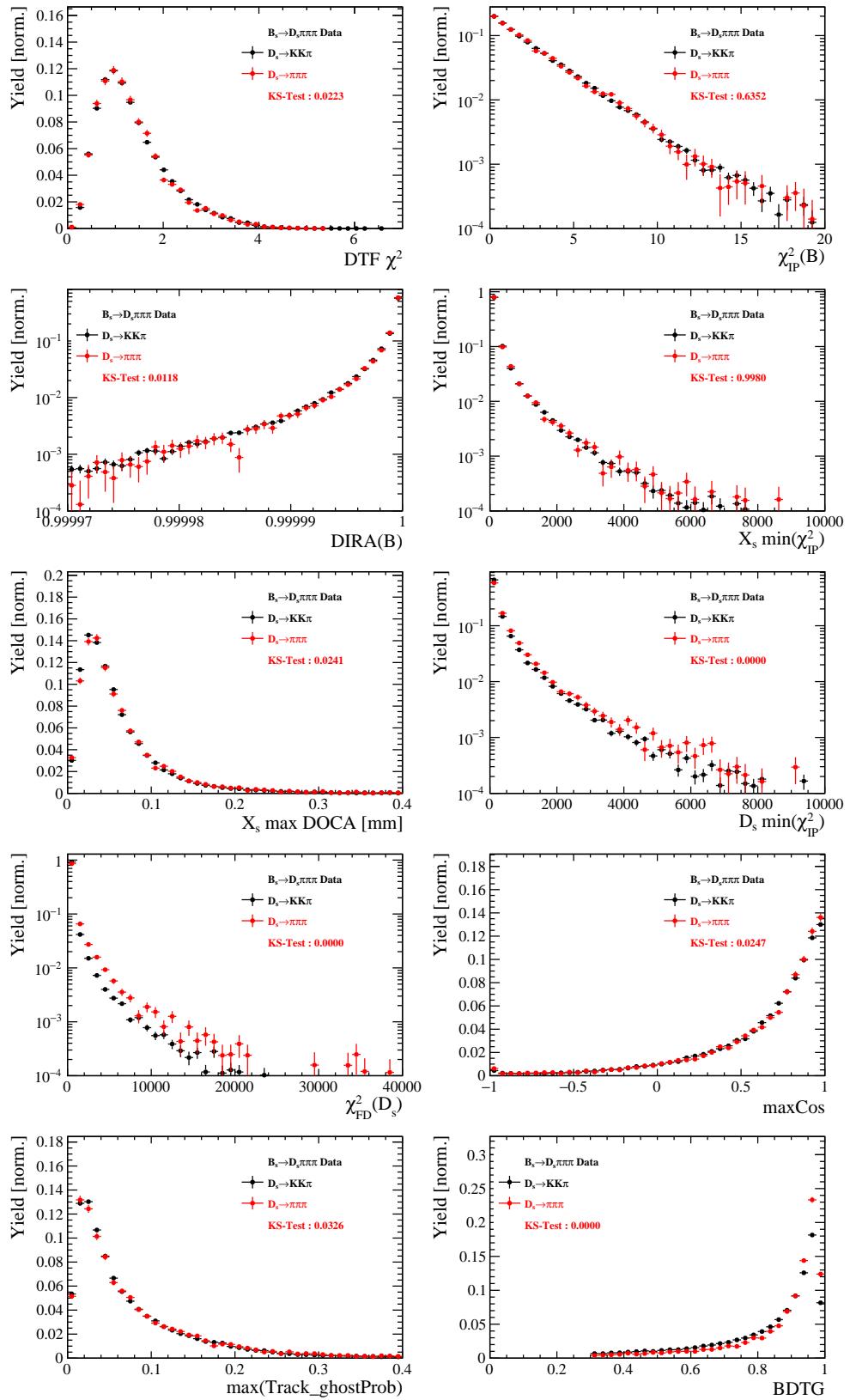


Figure C.6: Comparison of BDTG input variables and classifier response.

826 H.4 Comparison of trigger categories

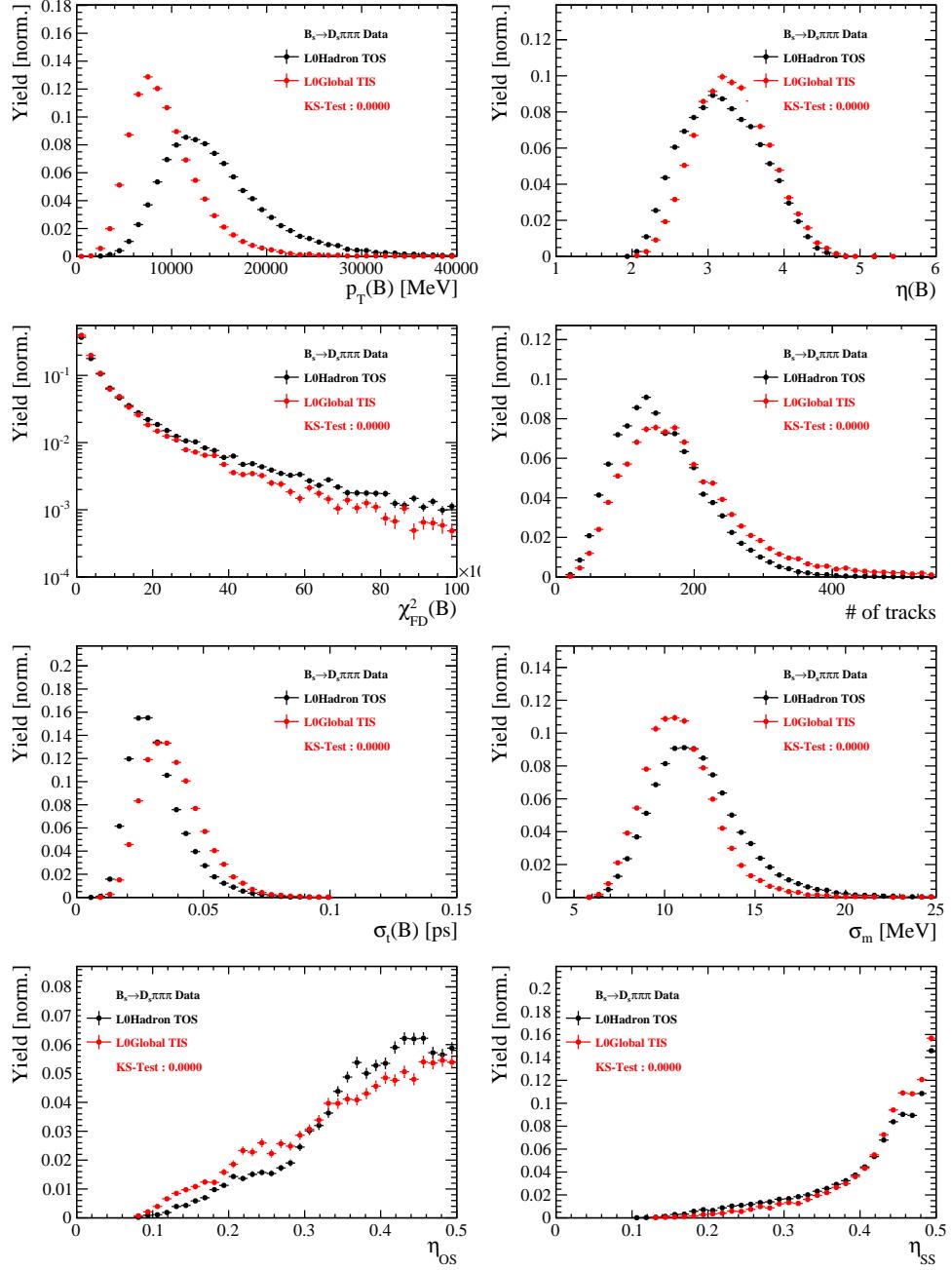


Figure C.7: Comparison of selected variables.

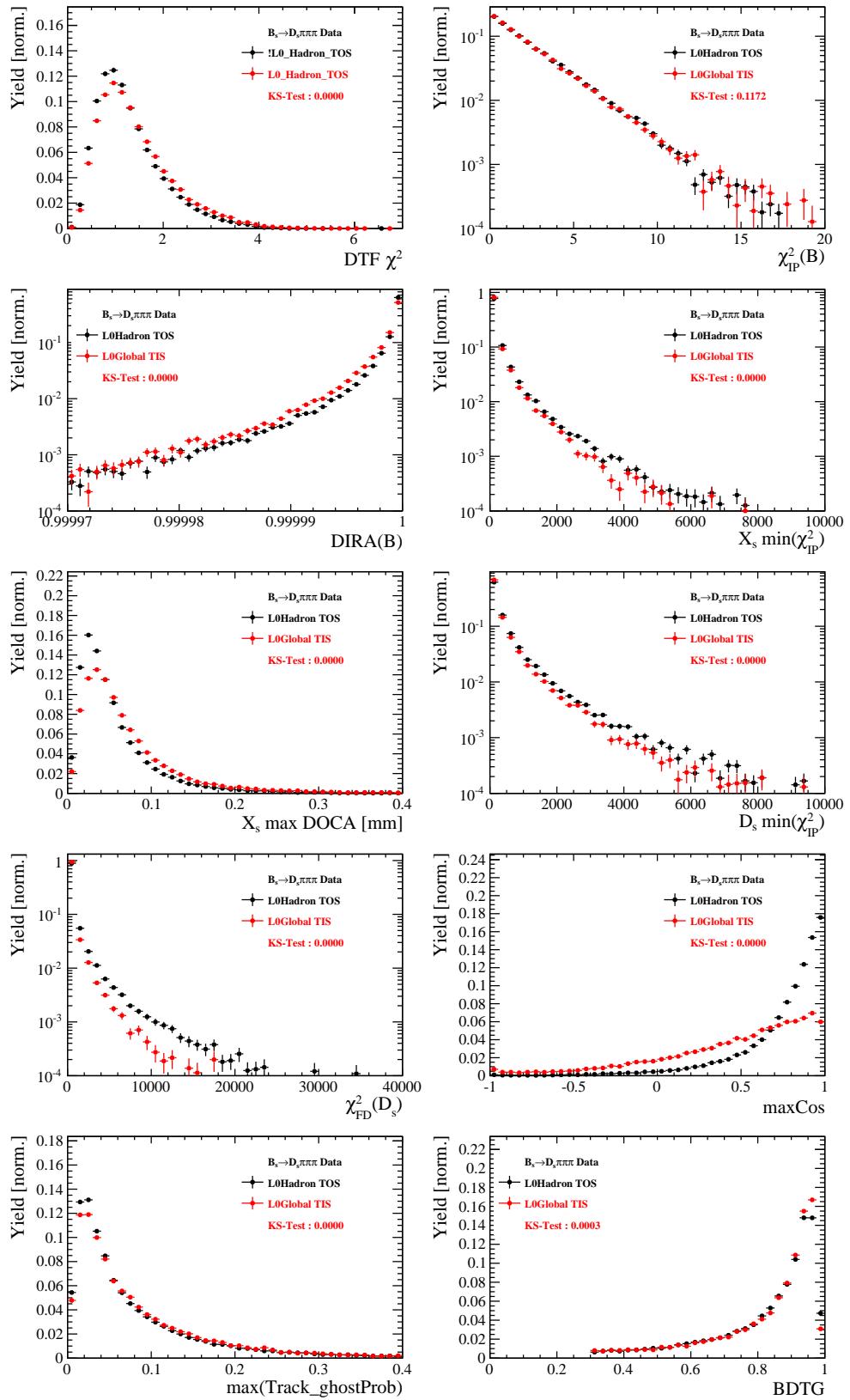


Figure C.8: Comparison of BDTG input variables and classifier response.

827 H.5 Comparison of B_s and B_d decays

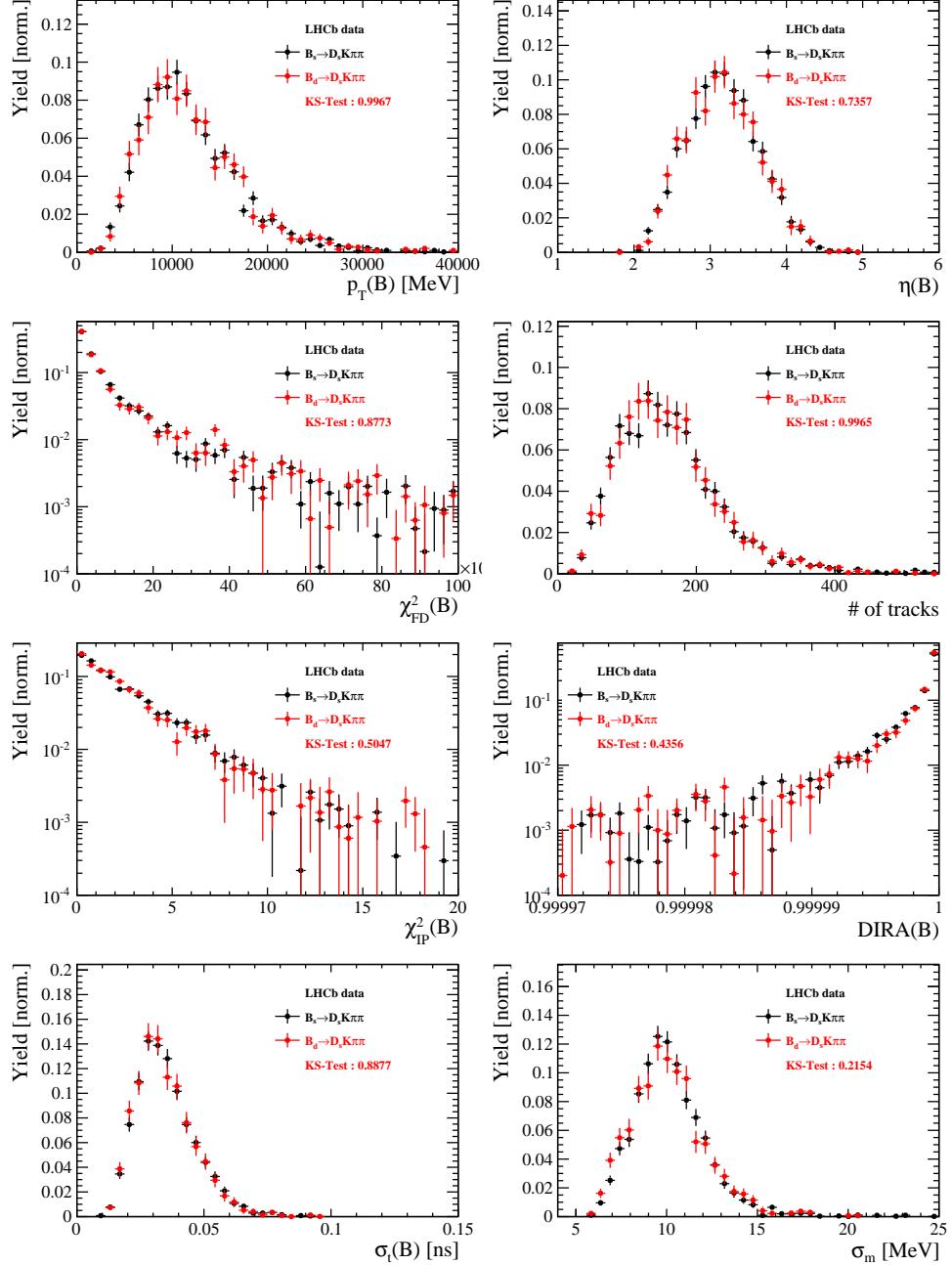


Figure C.9: Comparison of selected variables.

828 References

- 829 [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.
830 Phys. **B671** (2003) 459, arXiv:hep-ph/0304027.
- 832 [2] K. De Bruyn *et al.*, *Exploring $B_s \rightarrow D_s^{(*)\pm} K^\mp$ Decays in the Presence of a Sizable*
833 *Width Difference $\Delta\Gamma_s$* , Nucl. Phys. **B868** (2013) 351, arXiv:1208.6463.
- 834 [3] S. Blusk, *First observations and measurements of the branching fractions for the*
835 *decays $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* .
- 836 [4] LHCb, S. Blusk, *Measurement of the CP observables in $\bar{B}_s^0 \rightarrow D_s^+ K^-$ and first obser-*
837 *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.
- 838 [5] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory* (*Frontiers in Physics*), Westview Press, 1995.
- 840 [6] E. Byckling and K. Kajantie, *Particle Kinematics*, John Wiley & Sons, 1973.
- 841 [7] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, *Isobar approximation*
842 *of production processes*, Annals of Physics **18** (1962), no. 2 198 .
- 843 [8] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*,
844 Phys. Rev. D **11** (1975) 3165.
- 845 [9] J. J. Brehm, *Unitarity and the isobar model: Two-body discontinuities*, Annals of
846 Physics **108** (1977), no. 2 454 .
- 847 [10] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay*
848 *widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.
- 849 [11] J. D. Jackson, *Remarks on the phenomenological analysis of resonances*, Il Nuovo
850 Cimento Series 10 **34** (1964), no. 6 1644.
- 851 [12] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys.
852 **C40** (2016), no. 10 100001.
- 853 [13] D. V. Bugg, *The mass of the σ pole*, Journal of Physics G Nuclear Physics **34** (2007)
854 151, arXiv:hep-ph/0608081.
- 855 [14] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-*
856 *dominance prediction for $\rho \rightarrow e^+ e^-$* , Phys. Rev. Lett. **21** (1968) 244.
- 857 [15] S. M. Flatté, *Coupled-channel analysis of the $\pi\eta$ and KK systems near KK threshold*,
858 Physics Letters B **63** (1976), no. 2 224 .
- 859 [16] BES Collaboration, M. Ablikim *et al.*, *Resonances in $J/\psi \rightarrow \phi\pi^+\pi^-$ and ϕK^+K^-* ,
860 Phys. Lett. **B607** (2005) 243, arXiv:hep-ex/0411001.
- 861 [17] D. V. Bugg, *A study in depth of $f_0(1370)$* , Eur. Phys. J. **C52** (2007) 55,
862 arXiv:0706.1341.

- [863] [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](#).
- [864]
- [865] [19] C. Zemach, *Use of angular momentum tensors*, Phys. Rev. **140** (1965) B97.
- [866] [20] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, Phys. Rev. **60** (1941) 61.
- [867]
- [868] [21] S. U. Chung, *General formulation of covariant helicity-coupling amplitudes*, Phys. Rev. D **57** (1998) 431.
- [869]
- [870] [22] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of ψ decay to mesons*, Eur. Phys. J. **A16** (2003) 537, [arXiv:hep-ph/0211457](#).
- [871]
- [872] [23] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. **D51** (1995) 2247.
- [873]
- [874] [24] J.-J. Zhu, *Explicit expressions of spin wave functions*, [arXiv:hep-ph/9906250](#).
- [875] [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](#).
- [876]
- [877] [26] M. Williams, *Numerical Object Oriented Quantum Field Theory Calculations*, Comput. Phys. Commun. **180** (2009) 1847, [arXiv:0805.2956](#).
- [878]
- [879] [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](#).
- [880]
- [881] [28] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [882]
- [883] [29] *Gammacombo package*, 2014.
- [884] [30] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, [arXiv:physics/0703039](#).
- [885]
- [886] [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](#).
- [887]
- [888] [32] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.
- [889]
- [890] [33] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [891]
- [892] [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](#).
- [893]
- [894] [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](#).
- [895]

- 896 [36] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b -hadron, c -hadron, and*
 897 *τ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots
 898 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 899 [37] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*
 900 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 901 [38] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys.
 902 **A** **30** (2015) 1530022, arXiv:1412.6352.
- 903 [39] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*,
 904 Submitted to: JHEP (2017) arXiv:1712.07428.
- 905 [40] LHCb, R. Aaij *et al.*, *Measurement of B^0 , B_s^0 , B^+ and Λ_b^0 production asymmetries in 7*
 906 *and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.
- 907 [41] H. Gordon, R. W. Lambert, J. van Tilburg, and M. Vesterinen, *A Measurement of*
 908 *the $K\pi$ Detection Asymmetry*, Tech. Rep. LHCb-INT-2012-027. CERN-LHCb-INT-
 909 2012-027, CERN, Geneva, Feb, 2013.
- 910 [42] A. Davis *et al.*, *Measurement of the instrumental asymmetry for $K^- \pi^+$ -pairs at LHCb*
 911 *in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN,
 912 Geneva, Mar, 2018.
- 913 [43] I. I. Y. Bigi and H. Yamamoto, *Interference between Cabibbo allowed and doubly*
 914 *forbidden transitions in $D \rightarrow K(S)$, $K(L) + \pi$'s decays*, Phys. Lett. **B349** (1995)
 915 363, arXiv:hep-ph/9502238.
- 916 [44] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude*
 917 *analysis*, JINST **10** (2015), no. 09 P09002, arXiv:1505.05133.
- 918 [45] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, Journal of the Royal
 919 Statistical Society, Series B **58** (1994) 267.
- 920 [46] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.
- 921 [47] H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on
 922 Automatic Control **19** (1974) 716.