

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 **xxx** fb⁻¹ recorded by the LHCb detector.

Contents

1	Introduction	1
2	Sensitivity studies	2
2.1	PDF	2
2.2	Results	4
3	Selection	7
3.1	Cut-based selection	7
3.2	Multivariate stage	8
4	Fits to invariant mass distributions of signal and normalization channel	11
4.1	Signal models for $m(D_s\pi\pi\pi)$ and $m(D_sK\pi\pi)$	11
4.2	Background models for $m(D_s\pi\pi\pi)$	12
4.3	Background models for $m(D_sK\pi\pi)$	12
4.4	Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates	14
4.5	Fit to $B_s^0 \rightarrow D_sK\pi\pi$ candidates	14
4.6	Extraction of signal weights	14
5	Flavour Tagging	16
5.1	OS tagging calibration	17
5.2	SS tagging calibration	17
5.3	Tagging performance comparison between the signal and normalization channel	17
5.4	Combination of OS and SS taggers	18
6	Acceptance	21
6.1	MC corrections	21
6.1.1	Truth matching of simulated candidates	21
6.1.2	PID efficiencies	22
6.1.3	BDT efficiencies	22
6.1.4	Tracking efficiencies	22
6.2	Decay-time acceptance	23
6.2.1	Comparison of acceptance in subsamples	24
6.2.2	Results	26
6.3	Phasespace acceptance	30
7	Decay-time Resolution	32
7.1	Formalism	32
7.2	Decay-time Error in Run I & Run II	32
7.3	Fits to the decay time distributions	33
7.4	Results	34
8	B_s Production Asymmetry	36

9 Time dependent fit	37
9.1 sFit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data	37
9.2 sFit to $B_s^0 \rightarrow D_sK\pi\pi$ data	37
9.3 Results	37
A Details of multivariate classifier	38
B Detailed mass fits	39
C Decay-time Resolution fits	43
D MC corrections	44
E Data distributions	49
E.1 Comparison of signal and calibration channel	49
E.2 Comparison of Run-I and Run-II data	51
E.3 Comparison of D_s final states	53
E.4 Comparison of trigger categories	55
E.5 Comparison of B_s and B_d decays	57
References	58

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].
 The $B_s^0 \rightarrow D_s K\pi\pi$ proceeds at tree level via the transitions shown in Fig. 1.1 a) and b).

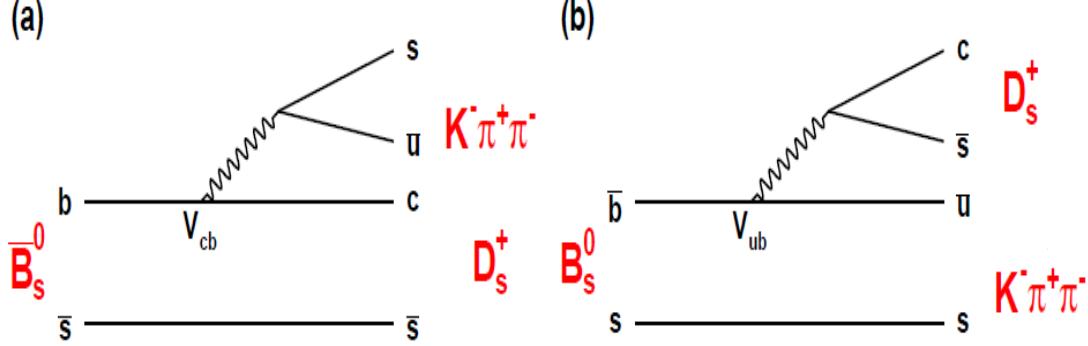


Figure 1.1: Feynman diagram of the $B_s^0 \rightarrow D_s K\pi\pi$ decay, proceeding via a) $b \rightarrow c$ transitions or b) $b \rightarrow u$ transitions.

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

2 Sensitivity studies

2.1 PDF

First, we define the purely hadronic amplitudes for a given phasespace point x . The weak phase dependence is written latter explicitly in the pdf.

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

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

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

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

The full time-dependent amplitude pdf is given by:

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

where $q_t = +1$ (-1) for a B_s^0 (\bar{B}_s^0) tag and $q_f = +1$ (-1) for $D_s^- K^+ \pi\pi$ ($D_s^+ K^- \pi\pi$) final states.

Integrating over the phasespace, we get

$$\begin{aligned} \int P(x, t, q_t, q_f) dx \propto & [\cosh\left(\frac{\Delta\Gamma t}{2}\right) \\ & + q_t q_f \left(\frac{1-r^2}{1+r^2}\right) \cos(m_s t) \\ & - 2 \left(\frac{\kappa r \cos(\delta - q_f(\gamma - 2\beta_s))}{1+r^2}\right) \sinh\left(\frac{\Delta\Gamma t}{2}\right) \\ & - 2q_t q_f \left(\frac{\kappa r \sin(\delta - q_f(\gamma - 2\beta_s))}{1+r^2}\right) \sin(m_s t)] e^{-\Gamma t} \\ = & [\cosh\left(\frac{\Delta\Gamma t}{2}\right) + q_t q_f C \cos(m_s t) - \kappa D_{q_f} \sinh\left(\frac{\Delta\Gamma t}{2}\right) - q_t \kappa S_{q_f} \sin(m_s t)] e^{-\Gamma t} \end{aligned} \quad (2.6)$$

where the C, D_{q_f}, S_{q_f} are defined exactly as for $D_s K$. The coherence factor is defined as :

$$\kappa e^{i\delta} \equiv \frac{\int A(x)^* \bar{A}(x) dx}{\sqrt{\int |A(x)|^2 dx} \sqrt{\int |\bar{A}(x)|^2 dx}} \quad (2.7)$$

$$r \equiv \frac{\sqrt{\int |\bar{A}(x)|^2 dx}}{\sqrt{\int |A(x)|^2 dx}} \quad (2.8)$$

³⁰ and appears in front of the D_{q_f}, S_{q_f} terms. This means one additional fit parameter for
³¹ the lifetime fit. In the limit of only one contributing resonance $\kappa \rightarrow 1$.

³²

33 2.2 Results

34 Assumptions:

- 35 • Use amplitudes from flavor-averaged, time-integrated fit
- 36 • $r = 0.4$ (ratio of CKM elements)
- 37 • PDG values for: $\tau, \Delta m_s, \Delta\Gamma, \beta_s$
- 38 • $\epsilon(x, t) = \text{const.}$, perfect resolution
- 39 • $\epsilon_{Tag} = 0.66, \langle \omega \rangle = 0.4$
- 40 • $N_{signal} = 3000$ (Run1+15/16 data)

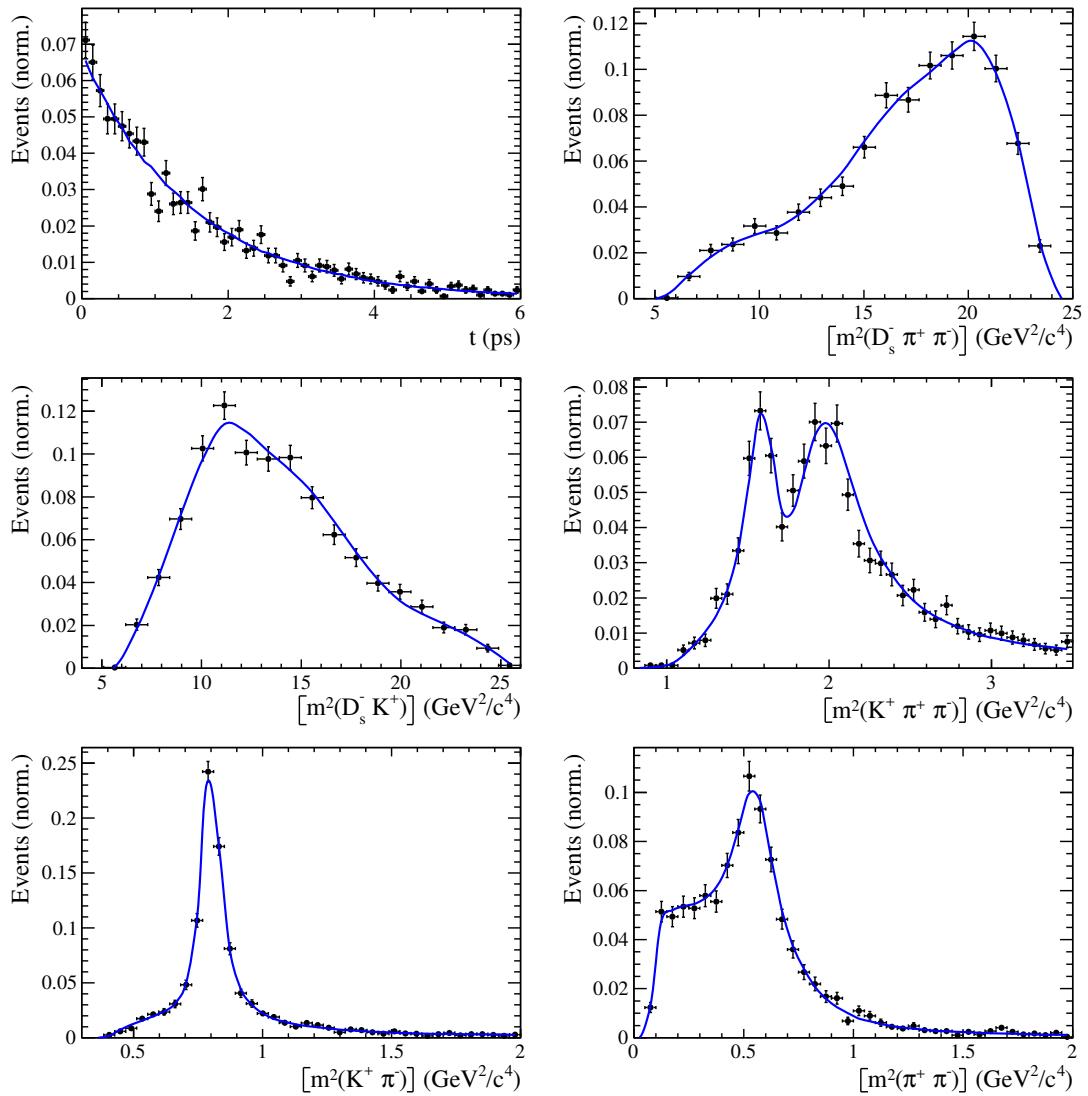
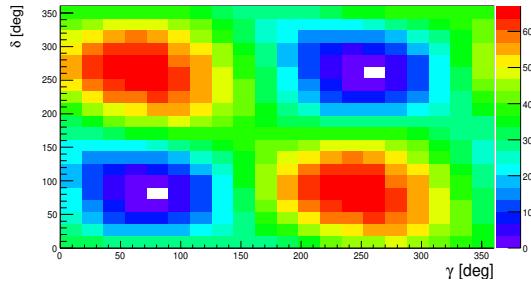


Figure 2.1: Example toy fit



Generated values:

$$\gamma = 70^\circ, \delta = 100^\circ$$

Fit result:

$$\begin{aligned} \gamma &= 74 \pm 15^\circ, \delta = 84 \pm 15^\circ \\ (\gamma &= 254 \pm 15^\circ, \delta = 264 \pm 15^\circ) \end{aligned}$$

Figure 2.2: Likelihood scan

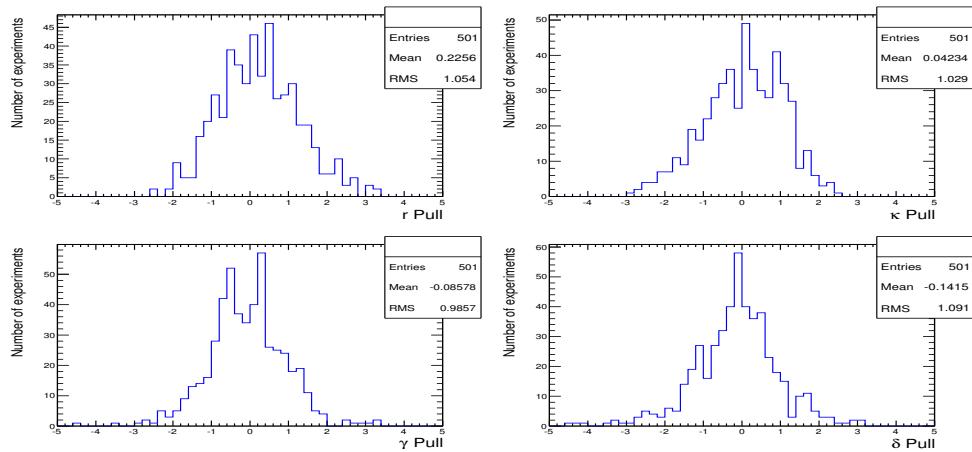


Figure 2.3: Pulls

Table 2.1

	Generated	Full PDF	Phasespace integrated
r	0.4	0.38 ± 0.06	unstable
κ	0.2	0.23 ± 0.13	0.2 (fixed)
δ	100	99 ± 22	unstable
γ	70	70 ± 17	unstable

	Generated	Full PDF	Phasespace integrated
r	0.4	0.44 ± 0.07	0.43 ± 0.11
κ	0.4	0.41 ± 0.14	0.4 (fixed)
δ	100	101 ± 19	95 ± 41
γ	70	69 ± 16	66 ± 40

	Generated	Full PDF	Phasespace integrated
r	0.4	0.41 ± 0.08	0.39 ± 0.11
κ	0.6	0.60 ± 0.13	0.6 (fixed)
δ	100	98 ± 17	92 ± 25
γ	70	68 ± 17	65 ± 28

	Generated	Full PDF	Phasespace integrated
r	0.4	0.42 ± 0.09	0.39 ± 0.09
κ	1.0	0.96 ± 0.03	1.0 (fixed)
δ	100	100 ± 17	100 ± 17
γ	70	66 ± 17	67 ± 17

41 **3 Selection**

42 For the presented analysis, we reconstruct the $B_s^0 \rightarrow D_s K\pi\pi$ decay through two different
 43 final states of the D_s meson, $D_s \rightarrow KK\pi$ and $D_s \rightarrow \pi\pi\pi$. Of those two final states
 44 $D_s \rightarrow KK\pi$ is the most prominent one, while $\mathcal{BR}(D_s \rightarrow \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$
 45 holds for the other one.

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

55 **3.1 Cut-based selection**

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

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

64 Additionally, we veto various physical backgrounds, which have either the same final
 65 state as our signal decay, or can contribute via a single misidentification of $K \rightarrow \pi$ or
 66 $K \rightarrow p$. In the following, the vetoes are ordered by the reconstructed D_s final state they
 67 apply to:

68 1. All:

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

72 2. $D_s \rightarrow KK\pi$

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

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

81 3. $D_s \rightarrow \pi\pi\pi$

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

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

- 89 • resonant case:

- 90 – $D_s^+ \rightarrow \phi\pi^+$, with $|M(K^+K^-) - m_\phi| < 20 \text{ MeV}/c^2$: no additional requirements,
 91 since ϕ is narrow and almost pure K^+K^- .
- 92 – $D_s^+ \rightarrow \bar{K}^{*0}K^+$, with $|M(K^-\pi^+) - m_{K^{*0}}| < 75 \text{ MeV}/c^2$: $\text{DLL}_{K\pi} > 0$ for kaons,
 93 since this resonance is more than ten times broader than ϕ .

- 94 • non resonant case: $\text{DLL}_{K\pi} > 5$ for kaons, since the non resonant category has
 95 significant charmless contributions.

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

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

99 3.2 Multivariate stage

100 We use TMVA [5] to train a multivariate discriminator, which is used to further improve
 101 the signal to background ratio. The following variables are used for the training:

- 102 • $\text{max}(\text{ghostProb})$ over all tracks
- 103 • $\text{cone}(p_T)$ asymmetry of every track, which is defined to be the difference between the
 104 p_T of the π/K and the sum of all other p_T in a cone of radius $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}$
 105 $< 1 \text{ rad}$ around the signal π/K track.
- 106 • $\text{min}(\text{IP}\chi^2)$ over the X_s daughters
- 107 • $\text{max}(\text{DOCA})$ over all pairs of X_s daughters
- 108 • $\text{min}(\text{IP}\chi^2)$ over the D_s daughters

- 109 • D_s and B_s^0 DIRA
 110 • D_s FD significance
 111 • $\max(\cos(D_s h_i))$, where $\cos(D_s h_i)$ is the cosine of the angle between the D_s and
 112 another track i in the plane transverse to the beam
 113 • B_s^0 IP χ^2 , FD χ^2 and Vertex χ^2

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

123

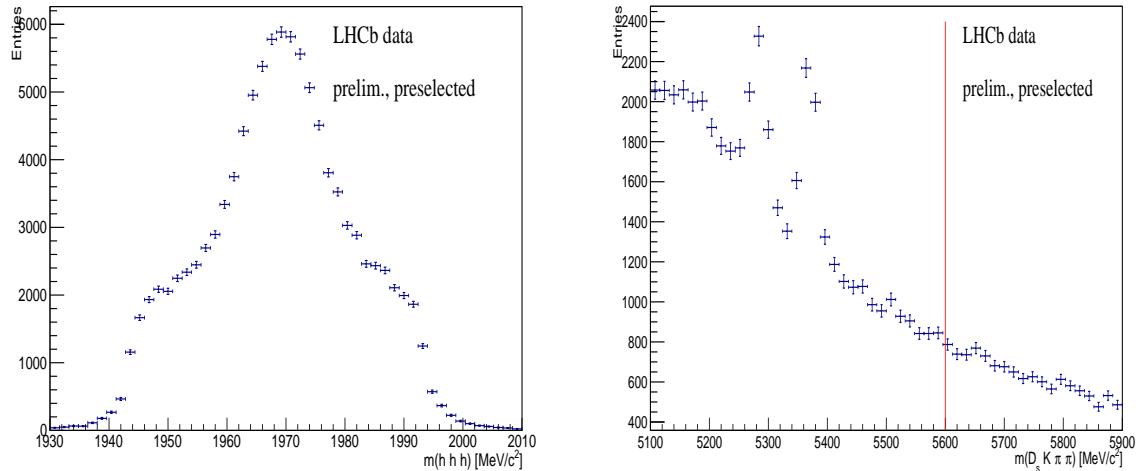


Figure 3.1: Invariant mass distribution of preselected (left) $D_s \rightarrow hhh$ and (right) $B_s^0 \rightarrow D_s K\pi\pi$ candidates. For the $B_s^0 \rightarrow D_s K\pi\pi$ candidates, the region right from the red colored line with $m_{B_s^0 \text{candidate}} > 5600 \text{ MeV}/c^2$ is used as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

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

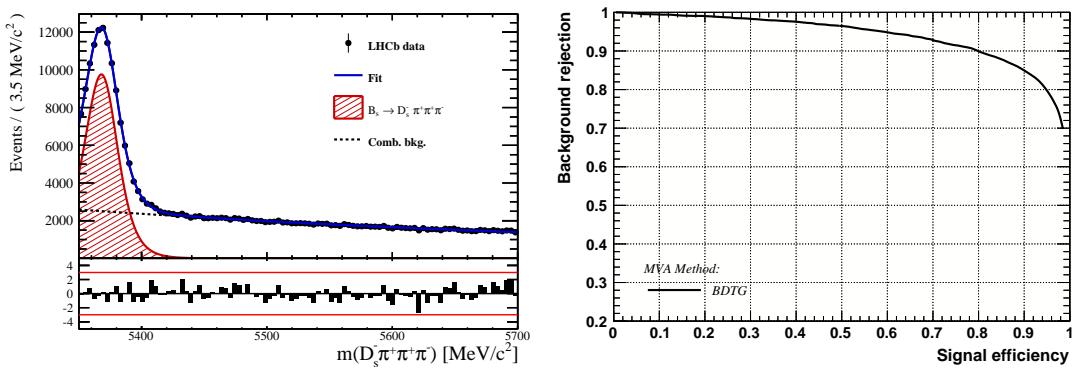


Figure 3.2

126 **4 Fits to invariant mass distributions of signal and**
 127 **normalization channel**

128 In order to properly model the invariant mass distribution of $B_s^0 \rightarrow D_s K\pi\pi$ and $B_s^0 \rightarrow$
 129 $D_s \pi\pi\pi$ candidates, the expected signal shape, as well as the expected shape for the
 130 combinatorial and physical background has to be known. This model can then be used to
 131 fit the distributions and obtain signal sWeights [6], which are employed to suppress the
 132 residual background that is still left in the sample, for the time-dependent amplitude fit.

133 **4.1 Signal models for $m(D_s \pi\pi\pi)$ and $m(D_s K\pi\pi)$**

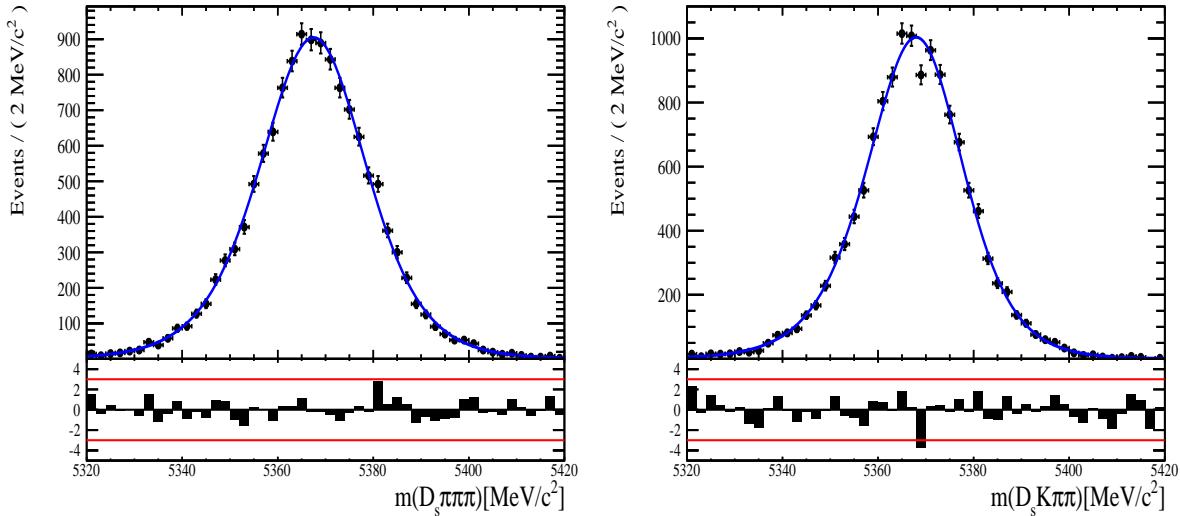


Figure 4.1: Invariant mass distributions of simulated (left) $B_s^0 \rightarrow D_s \pi\pi\pi$ and (right) $B_s^0 \rightarrow D_s K\pi\pi$ events. A fit of a RooJohnsonSU function to each distribution is overlaid.

134 The mass distribution of $B_s^0 \rightarrow D_s K\pi\pi$ signals is modeled using a Johnson SU
 135 function [7], which is a gaussian function with a Landau-like tail on one side,

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

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

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

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

150 4.2 Background models for $m(D_s\pi\pi\pi)$

151 Different background sources arise in the invariant mass spectrum of candidates in the
152 normalization mode.

153 The following backgrounds have to be accounted for:

- 154 • 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 .
- 157 • 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.

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

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

168 4.3 Background models for $m(D_sK\pi\pi)$

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

- 170 • Combinatorial background: same contributions as discussed in Sec. 4.2.
- 171 • 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.
- 173 • 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.
- 175 • Misidentified $B_s^0 \rightarrow D_s\pi\pi\pi$ decays, where one of the pions is wrongly identified as a kaon $\pi \rightarrow K$.
- 177 • 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.

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

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

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

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

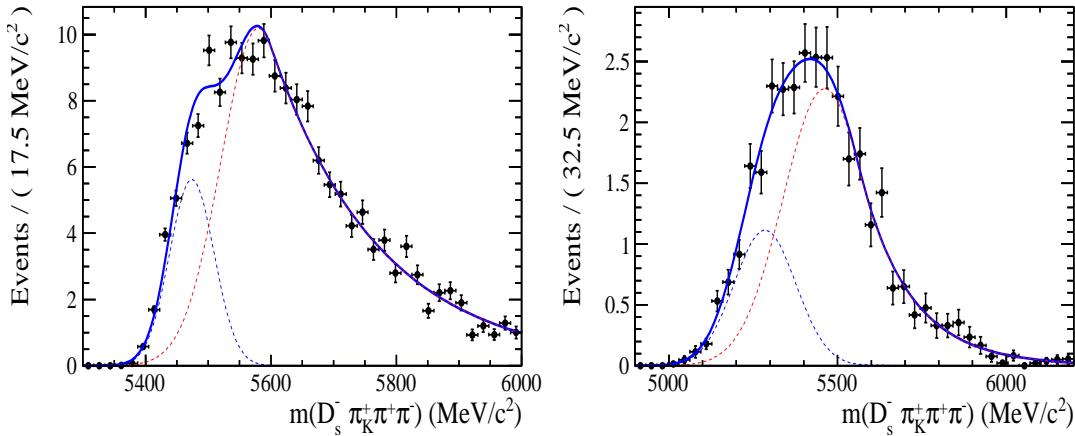


Figure 4.2: Invariant mass distribution of (left) simulated $B_s^0 \rightarrow D_s\pi\pi\pi$ events, where one of the π 's is reconstructed as a K and the misID probability for each event is taken into account. The corresponding distribution for simulated $B_s^0 \rightarrow D_s^*\pi\pi\pi$ events, where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

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

202 In the same way as mentioned above, we can determine the rate of misidentified, partially
 203 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
 204 and a MC sample of $B_s^0 \rightarrow D_s^*\pi\pi\pi$ events. The invariant mass distribution we obtain
 205 when we exclude the γ/π^0 , flip the the particle hypothesis $\pi \rightarrow K$ and apply the event
 206 weights given by the fake rate, is shown in Fig. 4.2 (right). The fit of two Crystal Ball
 207 functions to this distribution is overlaid. The yield of this contribution is determined
 208 from the yield of $B_s^0 \rightarrow D_s^*\pi\pi\pi$ candidates in the nominal mass fit of our normalization
 209 channel, multiplied by the misID probability of $\propto 3.6\%$.

210 4.4 Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates

211 An unbinned maximum likelihood fit is performed simultaneously to the invariant mass
 212 distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates. As discussed in Sec. ??, the fit is given
 213 as a Johnson SU signal model for the B_s^0 and B^0 signal, the sum of three bifurcated
 214 Gaussian functions to model the partially reconstructed $B_s^0 \rightarrow D_s^*\pi\pi\pi$ background and
 215 an Exponential function to account for combinatorial background. The invariant mass
 216 distribution and the fit is shown in Fig. 4.3. All simultaneously performed fits to the
 217 $m(D_s\pi\pi\pi)$ distribution, ordered by the respective D_s final state, can be found in the
 218 Appendix ???. The obtained yields are summarized in Table 4.1.

219 4.5 Fit to $B_s^0 \rightarrow D_sK\pi\pi$ candidates

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

228 4.6 Extraction of signal weights

229 The sPlot technique [6] is used to extract signal weights from the fits to the invariant
 230 mass distributions of our signal and normalization channel. This statistical tool assigns
 231 a weight to every event, according to its position in the respective mass distribution,
 232 given the fitted signal and background models. The weights can then be used to suppress
 233 the background components in every other observable distribution of interest. Figure ??
 234 shows the distribution of weights across the invariant mass spectra of $B_s^0 \rightarrow D_s\pi\pi\pi$ and
 235 $B_s^0 \rightarrow D_sK\pi\pi$ 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.

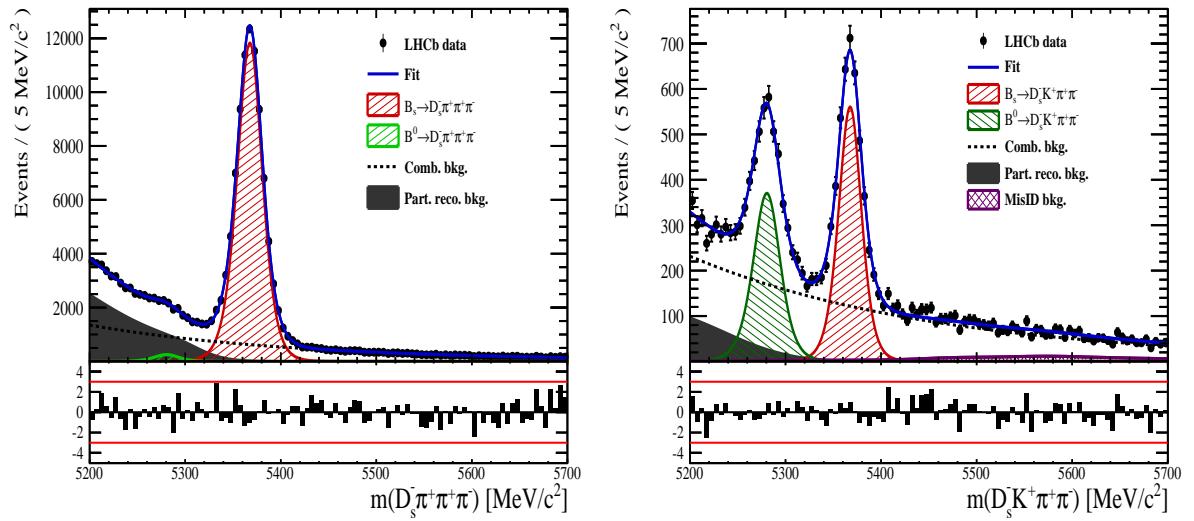


Figure 4.3: Invariant mass distribution of (left) $B_s^0 \rightarrow D_s\pi\pi\pi$ and (right) $B_s^0 \rightarrow D_sK\pi\pi$ candidates for Run1 and Run2 data. The respective fit described in the text is overlaid.

236 5 Flavour Tagging

237 To successfully perform a time- and amplitude-dependent measurement of γ , the identifi-
 238 cation of the initial state flavour of the B_s^0 meson is crucial. In the presented analysis,
 239 a number of flavour tagging algorithms are used that either determine the flavour of
 240 the non-signal b-hadron produced in the event (opposite site, OS), or they use particles
 241 produced in the fragmentation of the signal candidate B_s^0/\bar{B}_s^0 (same side, SS).
 242 For the same side, the algorithm searching for the charge of an additional kaon that accom-
 243 panies the fragmentation of the signal candidate is used (SS-nnetKaon). For the opposite
 244 site, four different taggers are chosen: The Two algorithms that use the charge of an
 245 electron or a muon from semileptonic B decays (OS- e,μ), the tagger that uses the charge
 246 of a kaon from a $b \rightarrow c \rightarrow s$ decay chain (OS-nnetKaon) and the algorithm that determines
 247 the B_s^0/\bar{B}_s^0 candidate flavour from the charge of a secondary vertex, reconstructed from
 248 the OS b decay product (OS-VtxCharge). All four taggers are then combined into a signal
 249 OS tagger.
 250 Every single tagging algorithm is prone to misidentify the signal candidate at a certain
 251 mistag rate $\omega = (\text{wrongtags})/(\text{alltags})$. This might be caused by particle misidentifica-
 252 tion, flavour oscillation of the neutral opposite site B-meson or by tracks that are wrongly
 253 picked up from the underlying event. For every signal B_s^0/\bar{B}_s^0 candidate, each tagging
 254 algorithm predicts a mistag probability η , which is calculated using a combination of
 255 inputs such as the kinematics of the tagging particles. The inputs are then combined
 256 to a predicted mistag using neural networks. These are trained on simulated samples
 257 of $B_s^0 \rightarrow D_s^- \pi^+$ (SS algorithm) and $B^+ \rightarrow J/\psi K^+$ (OS algorithms) decays. For the
 258 presented analysis, the measurable CP-violating coefficients are damped by the tagging
 259 dilution D , that depends on the mistag rate:

$$259 \quad D = 1 - 2\omega. \quad (5.1)$$

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

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

262 where ϵ_{tag} is the fraction of events that have a tagging decision. The flavour
 263 tagging algorithms are optimised for highest ϵ_{eff} on data, using the $B_s^0 \rightarrow D_s^- \pi^+$ and
 264 $B^+ \rightarrow J/\psi K^+$ samples.

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

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

268 where the values of p_0 and p_1 are determined using the $B_s^0 \rightarrow D_s \pi \pi \pi$ normalization
 269 mode and $\langle \eta \rangle$ is the average estimated mistag probability $\langle \eta \rangle = \sum_{i=1}^{N_{cand}} (\eta_i)/N_{cand}$.
 270 Following this model, a perfectly calibrated tagger would lead to $\omega(\eta) = \eta$ and one would
 271 expect $p_1 = 1$ and $p_0 = \langle \eta \rangle$. Due to the different interaction cross-sections of oppositely
 272 charged particles, the tagging calibration parameters depend on the initial state flavour of
 273 the B_s^0 . Therefore, the flavour asymmetry parameters Δp_0 , Δp_1 and $\Delta \epsilon_{tag}$ are introduced.
 274 For this analysis, the calibrated mistag is treated as per-event variable, giving a larger

275 weight to events that are less likely to have an incorrect tag. This adds one additional
 276 observable to the time- and amplitude-dependent fit.

277 The tagging calibration is determined using a time-dependent fit to the full $B_s^0 \rightarrow D_s\pi\pi\pi$
 278 sample, where the mixing frequency Δm_s is fixed to the nominal PDG value [8]. The
 279 calibration procedure for the OS tagging algorithms (Sec.5.1) and the SS kaon tagger
 280 (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,
 281 which is selected following the steps described in Sec. 3. The similar selection ensures
 282 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
 283 terms of the decay kinematics, which are crucial for the flavour tagging. Section 5.3 shows
 284 the compatibility of both samples. After applying the calibration, the response of the OS
 285 and SS taggers are combined, which is shown in Sec. 5.4.

286 5.1 OS tagging calibration

287 The responses of the OS electron, muon, neural net kaon and the secondary vertex charge
 288 taggers are combined for the mistag calibration. Figure ?? shows the distribution of the
 289 predicted OS mistag for signal candidates from $B_s^0 \rightarrow D_s\pi\pi\pi$. The extracted calibration
 290 parameters and tagging asymmetries are summarized in Table 5.1 and the measured
 291 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.

292 5.2 SS tagging calibration

293 The SS neural net kaon tagger can be calibrated using the flavour-specific $B_s^0 \rightarrow D_s\pi\pi\pi$
 294 decay. It's development, performance and calibration is described in detail in [9]. Figure
 295 ?? shows the distribution of the predicted mistag of the neural net kaon tagger. The
 296 extracted calibration parameters and tagging asymmetries are summarized in Table 5.2
 297 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.

298 5.3 Tagging performance comparison between the signal and 299 normalization channel

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

sweighted 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.

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

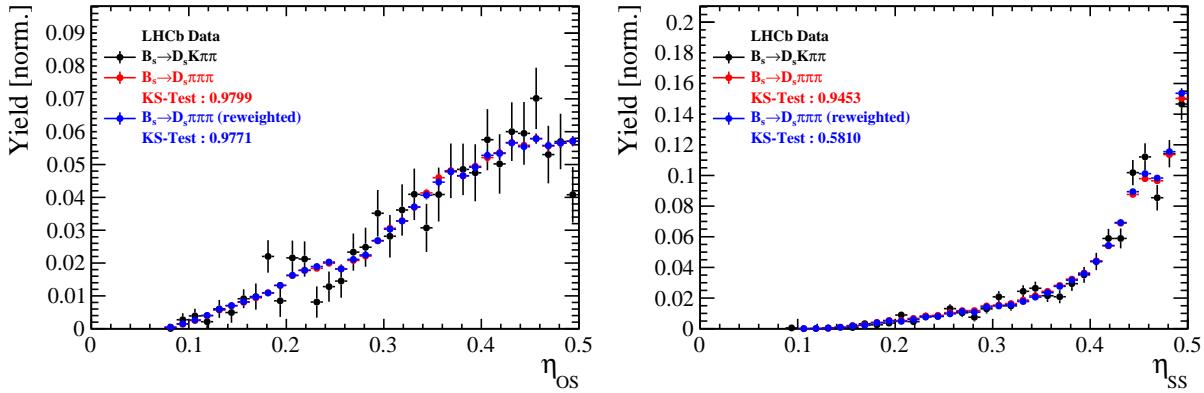


Figure 5.1: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \rightarrow D_s K\pi\pi$ (black) and $B_s^0 \rightarrow D_s \pi\pi\pi$ (red) data samples.

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

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

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

5.4 Combination of OS and SS taggers

In the time- and amplitude-dependent fit to $B_s^0 \rightarrow D_s K\pi\pi$ data, the obtained tagging responses of the OS and SS tagger will be combined after the calibration described in the previous sections is applied. Events that acquire a mistag probability greater than 0.5 after the calibration will have their tagging decision flipped. For events where only one of the two taggers fired, the combination of the tagging decision is trivial. In those events where both taggers made a decision, we use the standard combination of taggers [10] provided by the flavour tagging group. In the nominal fit, the calibrated mistags ω are combined event by event for the OS and SS tagger, thus adding one variable to observable to the fit procedure. This ensures that the uncertainties of the OS and SS tagging calibration parameters are propagated properly to the combined tagging response for each event. The tagging performance for the combined tagger in the categories SS tagged only, OS tagged only and SS+OS tagged, are shown in Tab. ?? for the signal and normalization

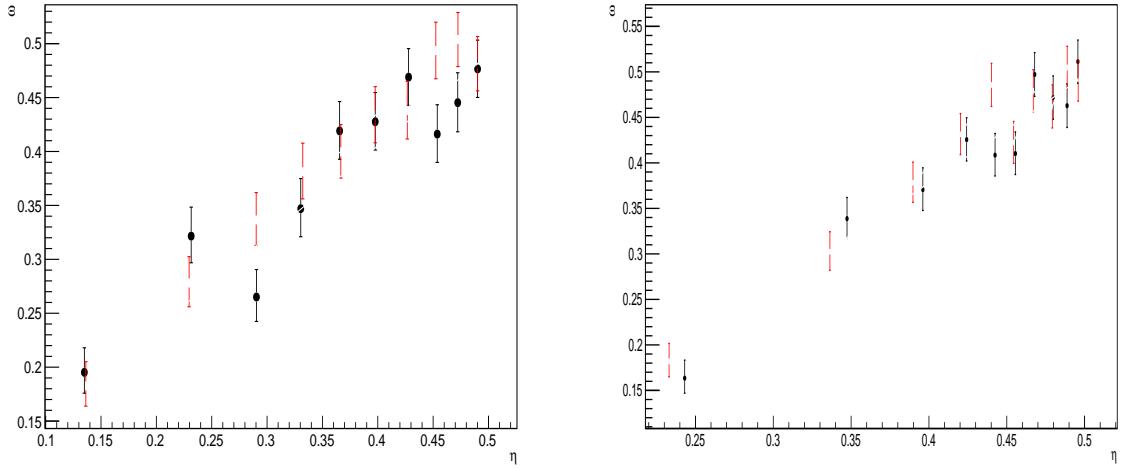


Figure 5.2: Dependence of the observed mistag ω on the predicted mistag η for the (left) OS combination and the (right) SS kaon tagger, found in the simulated $B_s^0 \rightarrow D_s K\pi\pi$ (black) and $B_s^0 \rightarrow D_s \pi\pi\pi$ (red) signal samples.

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

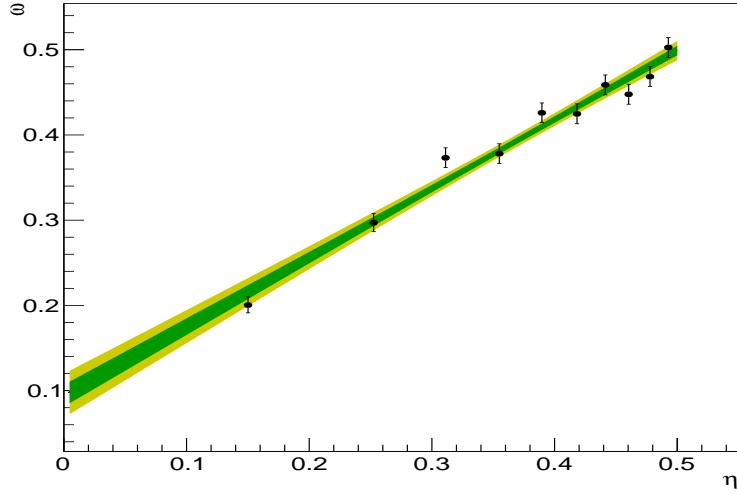


Figure 5.3: Distribution of the predicted combined mistag probability η versus the observed mistag ω for $B_s^0 \rightarrow D_s \pi\pi\pi$ signal candidates. The fit with a linear polynomial, used to determine p_0 and p_1 is overlaid.

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

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

334 6 Acceptance

335 6.1 MC corrections

336 6.1.1 Truth matching of simulated candidates

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

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

354 A significant fraction of 8% of the true signal candidates are classified as ghosts, while
 355 only $x.x\%$ of the events classified with `BKGCAT`= 60 are indeed genuine ghosts.

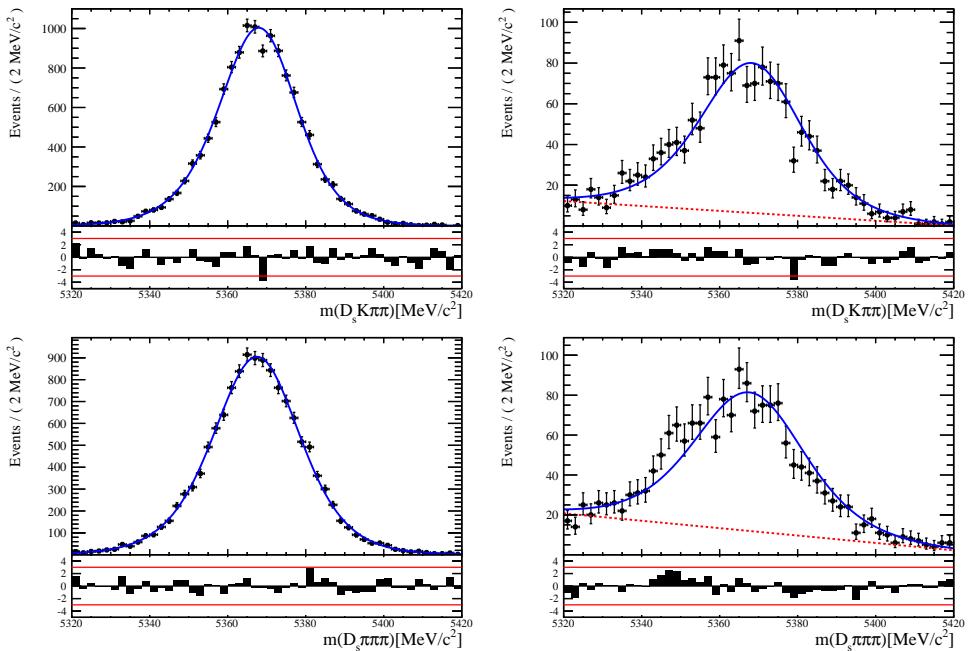


Figure 6.1: The reconstructed B_s mass distribution for simulated $B_s \rightarrow D_s K\pi\pi$ (top) and $B_s \rightarrow D_s \pi\pi\pi$ (bottom) decays classified with `BKGCAT` = 20 or 50 (left) and `BKGCAT` = 60 (right) after the full selection.

356 6.1.2 PID efficiencies

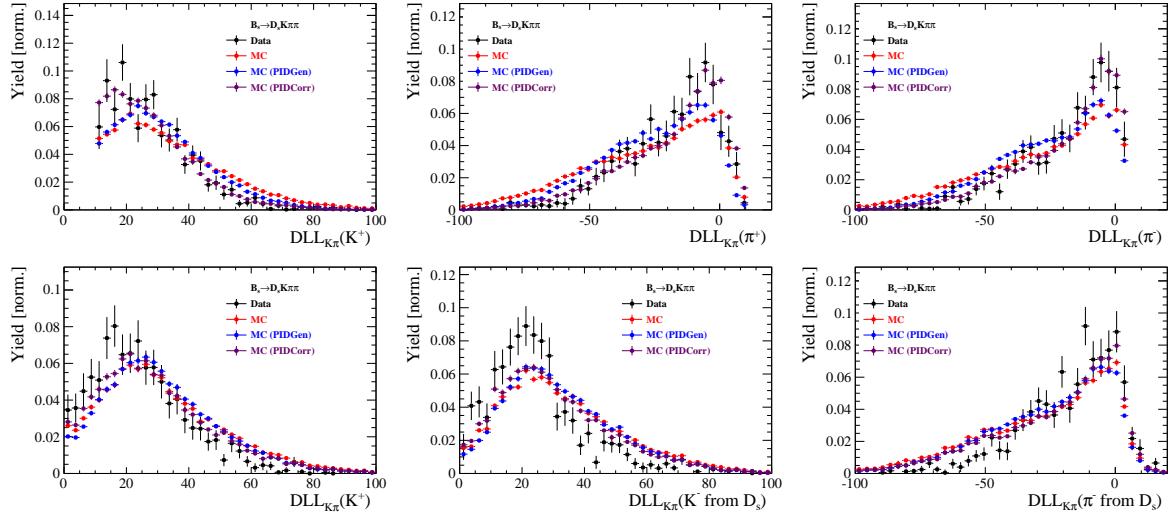


Figure 6.2

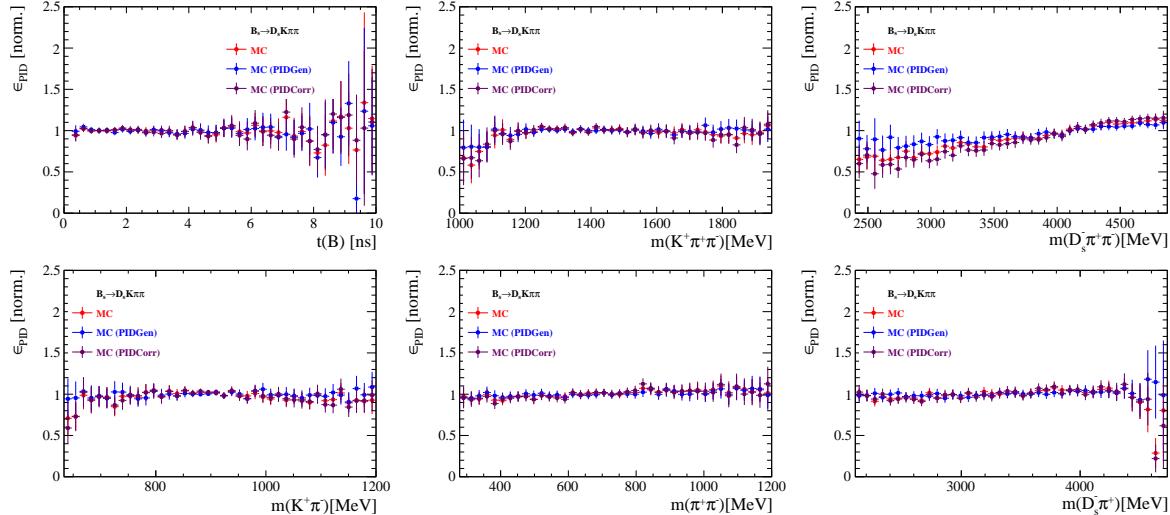


Figure 6.3

357 6.1.3 BDT efficiencies

358 6.1.4 Tracking efficiencies

359 **6.2 Decay-time acceptance**

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

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

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

373 A PDF of the form

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

374 is fit to the decay time distribution of $B_s^0 \rightarrow D_s \pi\pi\pi$ candidates in data. Since the
 375 fit is performed untagged, the PDF shown in Eq. 6.2 contains no terms proportional
 376 to Δm_s . The values for Γ_s and $\Delta\Gamma_s$ are fixed to the latest HFAG results [11]. The
 377 decay-time acceptance $\epsilon(t', \vec{\lambda})$ is modelled using the sum of cubic polynomials $v_i(t)$, so
 378 called Splines [12]. The polynomials are parametrised by so-called knots which determine
 379 their boundaries. Knots can be set across the fitted distribution to account for local
 380 changes in the acceptance shape. Using more knots is equivalent to using more base
 381 splines which are defined on a smaller sub-range. In total, $n + 2$ base splines $v_i(t)$ are
 382 needed to describe an acceptance shape which is parametrised using n knots.

383 For fits shown in the following, the knots have been placed at $t =$
 384 $[0.5, 1.0, 1.5, 2.0, 3.0, 9.5] ps$. To accommodate these 6 knot positions, 8 basic splines
 385 v_i , $i = [1, \dots, 8]$ are used. Since a rapid change of the decay time acceptance at low
 386 decay times due to the turn-on effect generated by the lifetime and other selection cuts is
 387 expected, more knots are placed in that regime. At higher decay times we expect linear
 388 behavior, with a possible small effect due to the VELO reconstruction. Therefore fewer
 389 knots are used. Furthermore, v_7 is fixed to 1 in order to normalize the overall acceptance
 390 function. To stabilise the last spline, v_8 is fixed by a linear extrapolation from the two
 391 previous splines:

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

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

6.2.1 Comparison of acceptance in subsamples

It is possible that the decay-time dependent efficiency deviates in different subsamples of our data. In particular, the acceptance could differentiate in subsamples with different final state kinematics, such as the run I & run II sample, the various D_s final states and the ways an event is triggered at the L0 stage. To investigate possible deviations, the full selected $B_s^0 \rightarrow D_s \pi\pi\pi$ sample is split into subsamples according to the categories mentioned above (run, D_s state, L0 trigger). For each subsample, the fit procedure described at the beginning of this chapter, using the pdf given by Eq. 6.2, is repeated and the obtained values for the spline coefficients v_i are compared. Figure 6.4 shows the comparison of the obtained spline coefficients for the different D_s final states.

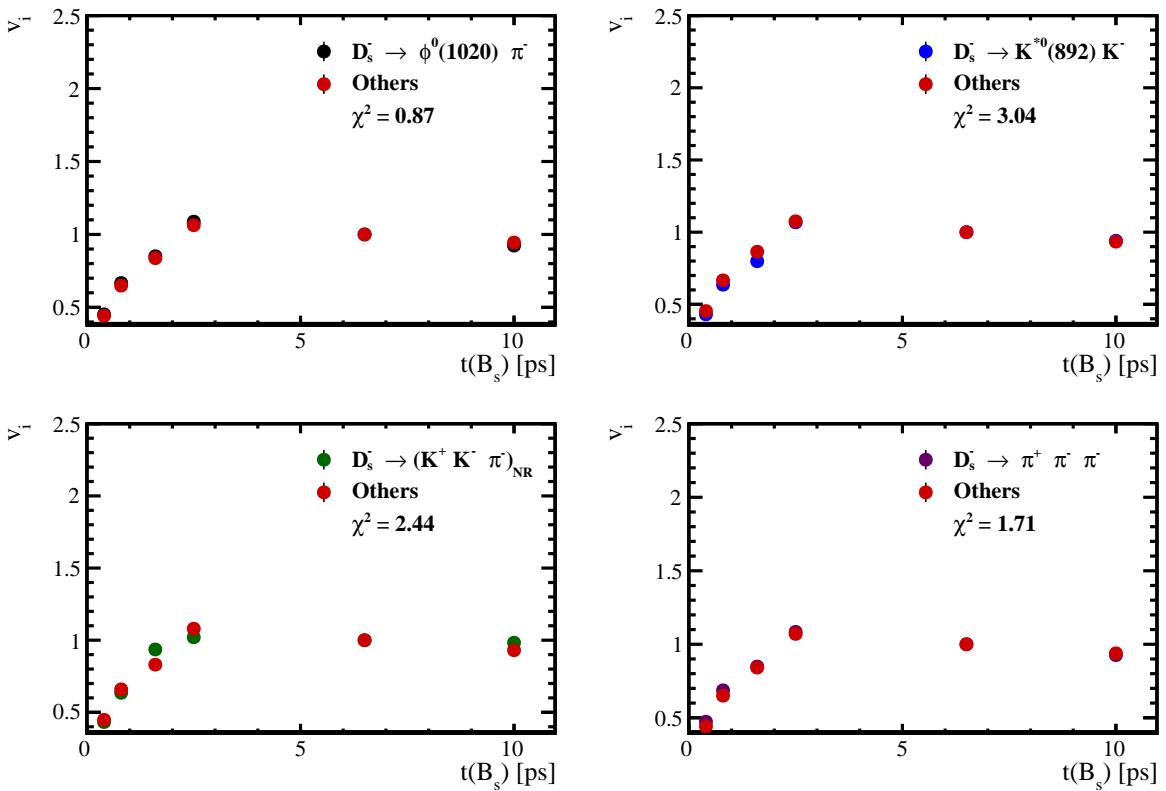


Figure 6.4: Comparison of the spline coefficients obtained from time-dependent fits to the $B_s^0 \rightarrow D_s \pi\pi\pi$ subsamples of different D_s final states. The comparison of one particular D_s state against all other states is shown.

Investigating the obtained spline coefficients from different D_s final states, good agreement is observed between all four channels and no need to distinguish between different final states in the time-dependent amplitude fit is found. The comparison between spline coefficients for the different runs and L0 trigger categories is shown in Figure 6.5.

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

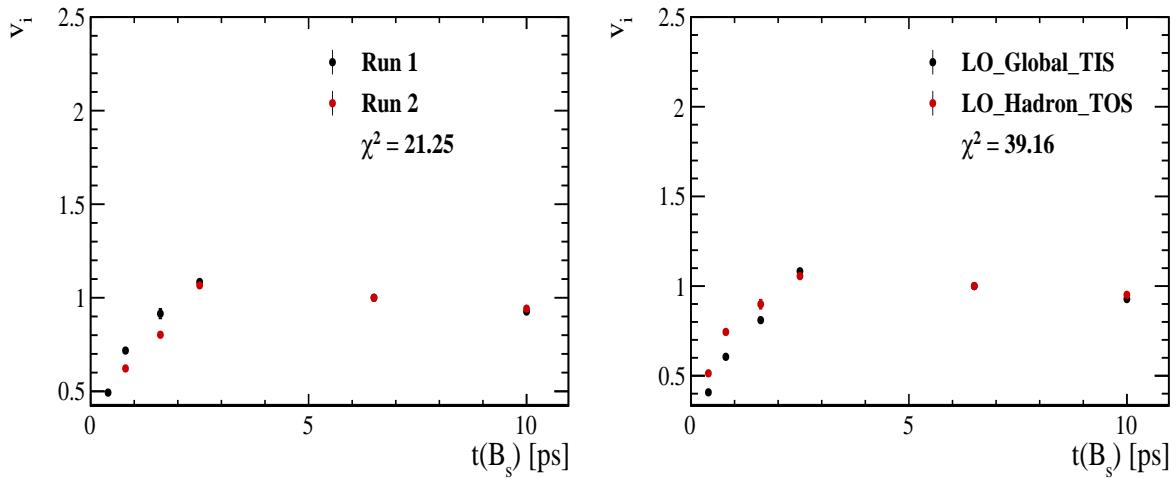


Figure 6.5: Comparison of the spline coefficients obtained from time-dependent fits to the $B_s^0 \rightarrow D_s \pi \pi \pi$ subsamples of (left) the different runs and (right) L0 trigger categories.

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

6.2.2 Results

The nominal fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ data using this configuration is shown in Figure 6.6. Note that the normalization of the splines in the following figures is not in scale. The fit parameters obtained from the described fits to data and simulation are summarised in Table 6.5.

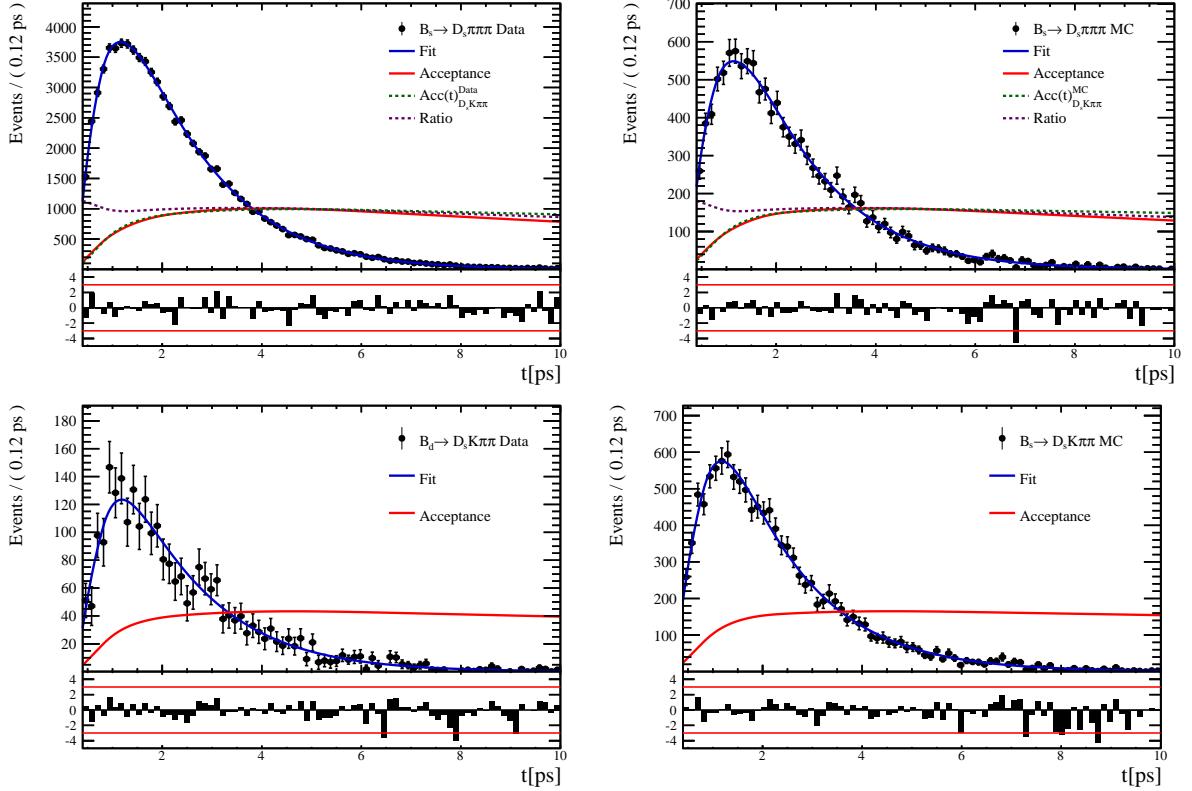


Figure 6.6: The red line shows the spline function describing the acceptance and the blue line depicts the total fit function.

Table 6.1: Summary of the obtained parameters from the acceptance fit

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	0.446 ± 0.034	0.479 ± 0.022	1.020 ± 0.078
0.8	v_1	0.673 ± 0.051	0.702 ± 0.034	0.943 ± 0.069
1.6	v_2	0.874 ± 0.076	0.915 ± 0.054	0.984 ± 0.084
2.5	v_3	1.028 ± 0.043	1.021 ± 0.038	1.042 ± 0.043
6.5	v_4	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
10.0	v_5	0.975 ± 0.000	0.982 ± 0.000	0.963 ± 0.000

Table 6.2: Summary of the obtained parameters from the acceptance fit

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	0.425 ± 0.039	0.456 ± 0.025	1.054 ± 0.091
0.8	v_1	0.626 ± 0.058	0.673 ± 0.039	1.028 ± 0.095
1.6	v_2	0.936 ± 0.105	0.887 ± 0.067	0.955 ± 0.102
2.5	v_3	1.081 ± 0.066	1.050 ± 0.049	1.019 ± 0.069
6.5	v_4	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
10.0	v_5	0.929 ± 0.000	0.956 ± 0.000	0.983 ± 0.000

Table 6.3: Summary of the obtained parameters from the acceptance fit

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	0.615 ± 0.080	0.525 ± 0.039	0.935 ± 0.106
0.8	v_1	0.922 ± 0.124	0.751 ± 0.060	0.842 ± 0.110
1.6	v_2	0.940 ± 0.123	0.959 ± 0.083	1.025 ± 0.126
2.5	v_3	0.940 ± 0.077	0.949 ± 0.062	1.111 ± 0.059
6.5	v_4	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
10.0	v_5	1.053 ± 0.000	1.045 ± 0.000	0.903 ± 0.000

Table 6.4: Summary of the obtained parameters from the acceptance fit

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	0.385 ± 0.019	0.478 ± 0.018	1.014 ± 0.056
0.8	v_1	0.584 ± 0.028	0.695 ± 0.029	0.975 ± 0.055
1.6	v_2	0.802 ± 0.067	0.919 ± 0.052	0.967 ± 0.078
2.5	v_3	1.032 ± 0.048	1.021 ± 0.038	1.045 ± 0.047
6.5	v_4	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
10.0	v_5	0.972 ± 0.000	0.982 ± 0.000	0.960 ± 0.000

Table 6.5: Summary of the obtained parameters from the acceptance fit

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	0.481 ± 0.032	0.479 ± 0.020	1.015 ± 0.061
0.8	v_1	0.711 ± 0.050	0.698 ± 0.031	0.962 ± 0.067
1.6	v_2	0.905 ± 0.079	0.918 ± 0.052	0.972 ± 0.079
2.5	v_3	1.002 ± 0.061	1.020 ± 0.041	1.045 ± 0.055
6.5	v_4	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
10.0	v_5	0.998 ± 0.000	0.982 ± 0.000	0.961 ± 0.000

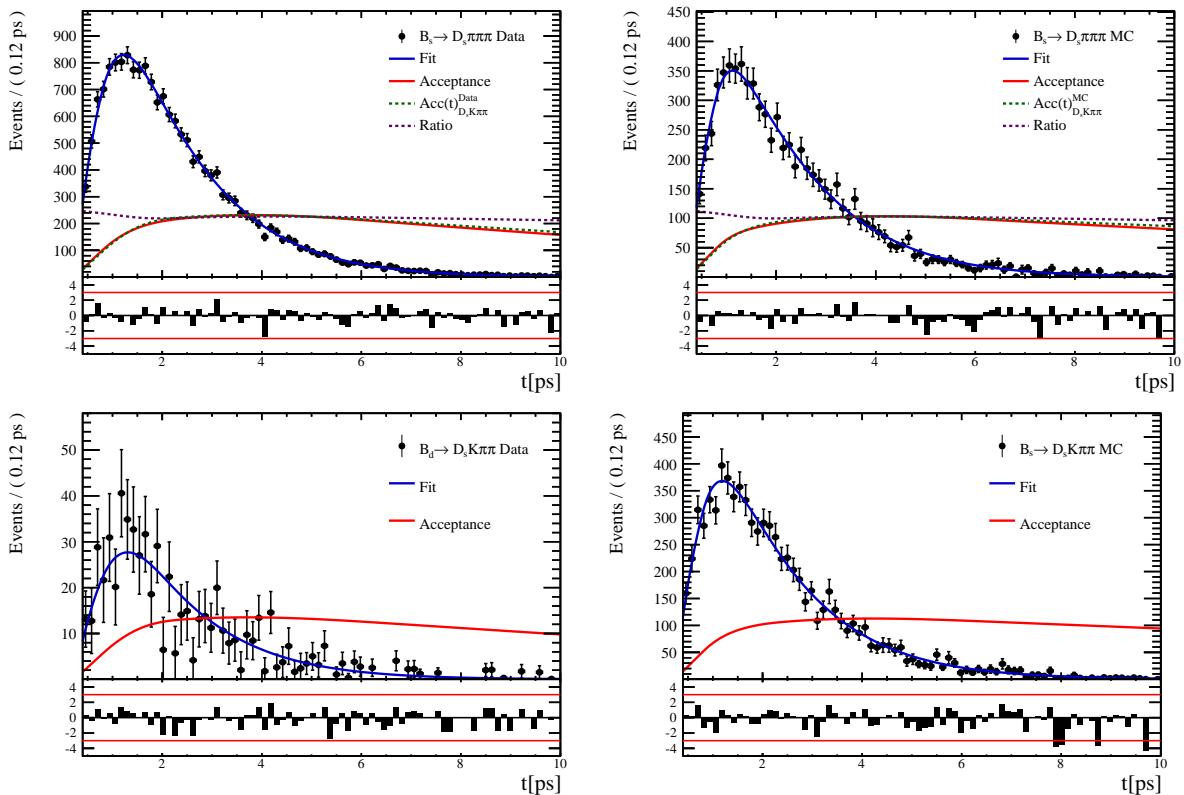


Figure 6.7

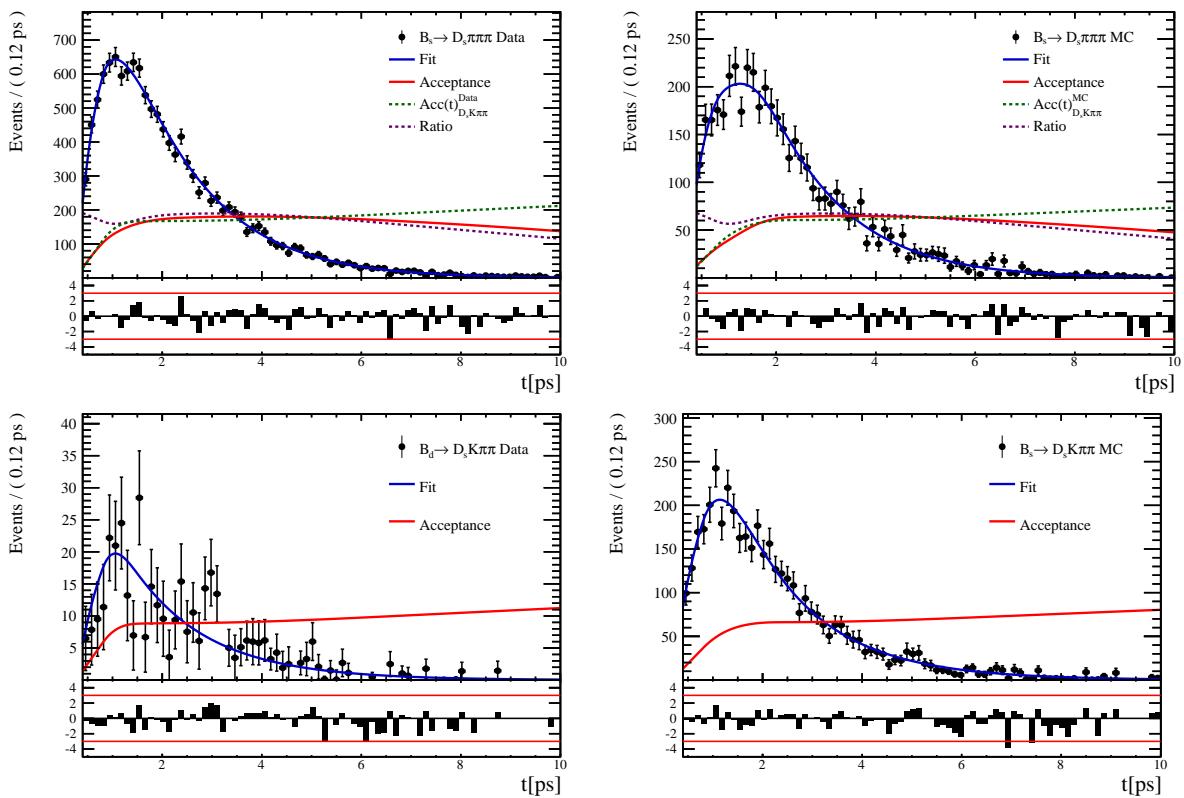


Figure 6.8

419 **6.3 Phasespace acceptance**

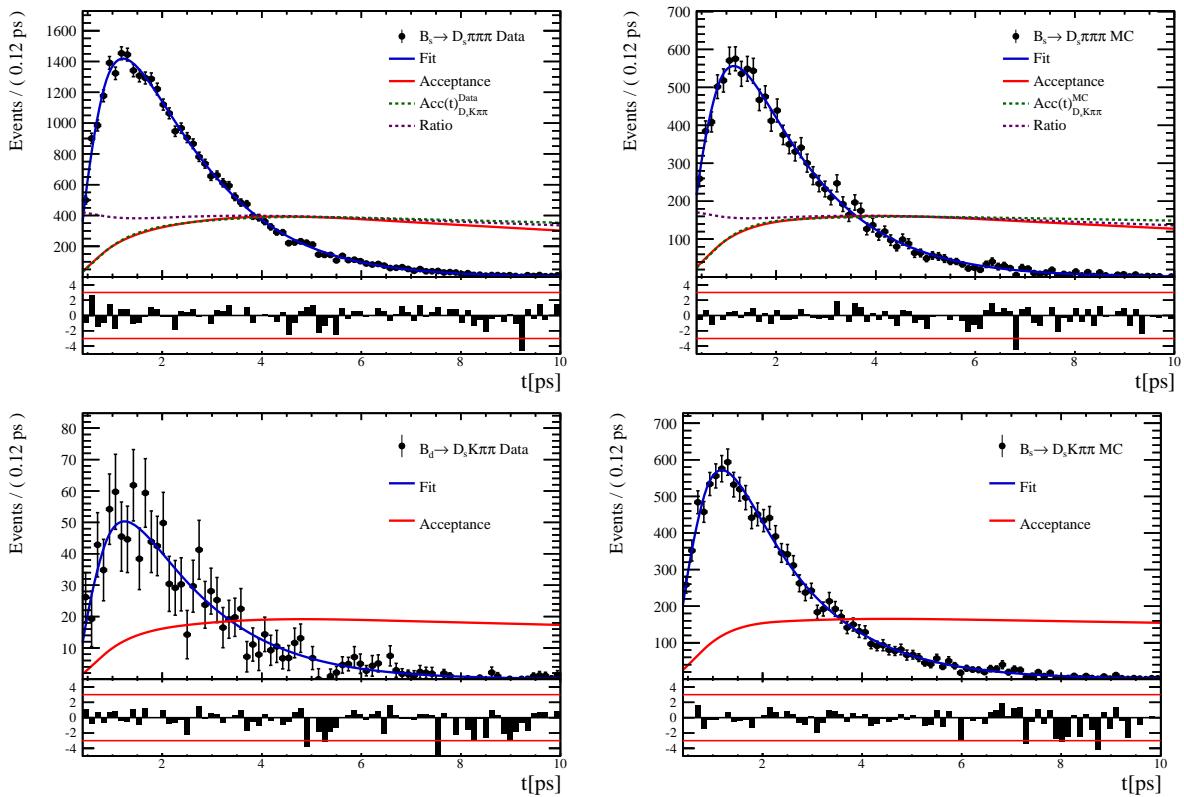


Figure 6.9

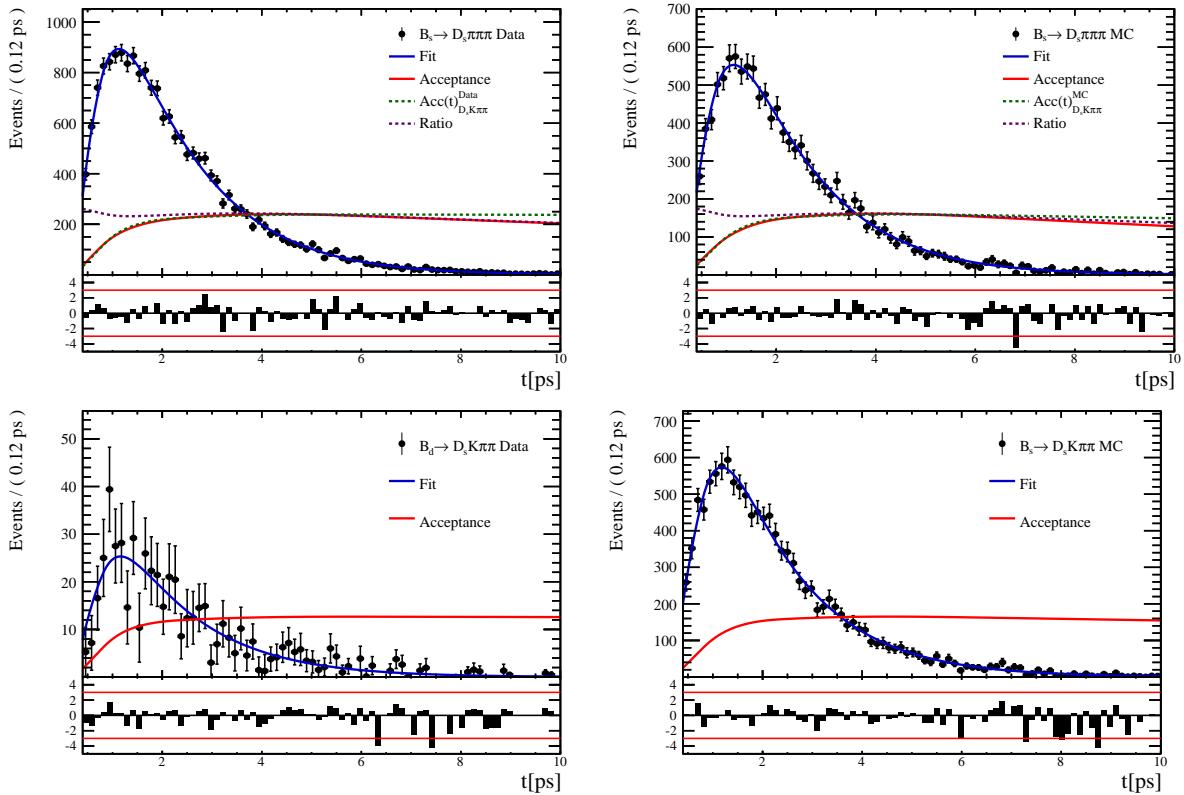


Figure 6.10

420 7 Decay-time Resoution

421 The observed oscillation of B mesons is prone to dilution, if the detector resolution is
422 of similar magnitude as the oscillation period. In the B_s^0 system, considering that the
423 measured oscillation frequency of the B_s^0 [8] and the average LHCb detector resolution [13]
424 are both $\mathcal{O}(50 \text{ fs}^{-1})$, this is the case. Therefore, it is crucial to correctly describe the
425 decay time resolution in order to avoid a bias on the measurement of time dependent CP
426 parameters.

427 In the presented analysis, we assume a gaussian resolution function with different
428 widths for each event. This gives rise to a per-event decay time error σ_t , which is
429 computed separately for every event along with the proper time t , by the decay time
430 fitter. Furthermore, the per-event decay time error σ_t is usually underestimated by the
431 decay time fitter, making it necessary to derive a scaling function, which matches the
432 per-event error to the actually measured decay time resolution.

433 Due to the lack of a decay time unbiased sample of real $B_s^0 \rightarrow D_s K \pi \pi$ decays, this
434 analysis relies on simulation to describe the time resolution. The obtained results will
435 be compared to those found in the closely related $B_s^0 \rightarrow D_s K$ analysis and systematic
436 uncertainties will be assigned conservatively. In the following, we investigate the Run1
437 and Run2 MC samples to find the proper decay time resolution in bins of the per-event
438 decay time errors and derive a scaling function from that.

440 7.1 Formalism

441 For simulated $B_s^0 \rightarrow D_s K \pi \pi$ events, the information on the true B_s^0 lifetime τ_{true} assigned
442 at production of the event, as well as the measured decay time $\tau_{measured}$, which is determined
443 after the interaction with the LHCb detector, is stored. In this analysis, the difference
444 $\Delta t = \tau_{true} - \tau_{measured}$ is obtained for each simulated $B_s^0 \rightarrow D_s K \pi \pi$ candidate. The width
445 of the distribution of Δt is a direct measure of the decay time resolution.

446 To analyse the relation between the per-event decay time error σ_t and the actual resolution,
447 the simulated sample is split into 8 bins of σ_t . Each bin width is chosen using an adaptive
448 binning scheme which ensures that approximately equal numbers of events are found
449 in every bin. A fit is then performed to the distribution of Δt in each of the bins to
450 determine the decay time resolution in the respective bin. After that, the correlation
451 between the binned per-event decay time error and the measured decay time resolution is
452 analyzed to determine the scaling function.

453 7.2 Decay-time Error in Run I & Run II

454 Due to the increase in center of mass energy from Run I to Run II, as well as (among
455 others) new tuning in the pattern and vertex reconstruction, the distributions of the raw
456 decay time error might not necessarily match each other between the two different runs.
457 Significant deviations can be observed in the shape and mean of those two distributions
458 for $B_s^0 \rightarrow D_s K \pi \pi$ signal candidates shown in Figure 7.1.

459 It can be observed that the decay time error distribution for signal candidates from Run
460 II is significantly broader and shifted to slightly higher values. Due to the discrepancies
461 between the distributions of the decay time error σ_r for Run I and Run II data, the time

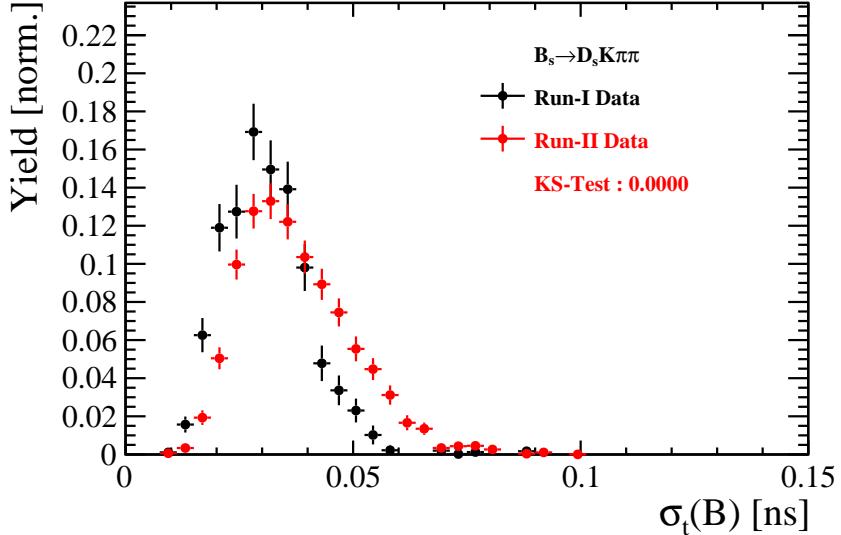


Figure 7.1: Distribution of the decay time error for $B_s^0 \rightarrow D_s K\pi\pi$ signal candidates on data for (black) Run I and (red) Run II. The signal distributions are obtained using the sWeight technique.

462 resolution studies have to be performed separately for both runs, which leads to two
463 different scaling functions to map $\sigma_t \rightarrow \sigma_{eff}$.

464 7.3 Fits to the decay time distributions

465 The sum of two Gaussian functions is used to fit the distributions of the decay time
466 difference Δt in each σ_t bin. One Gaussian function is relatively narrow and describes the
467 decay time of the majority of candidates in each bin, while the other, broader Gaussian
468 function describes candidates where the measured decay time differs considerably from
469 τ_{true} . Those contributions are shifted to the tails of the distribution. From the two
470 Gaussian functions, the combined, effective width σ_{eff} is quoted as decay time resolution
471 in each bin. Figure 7.2 shows the double Gaussian fit to the distribution of the decay time
472 difference for events where $20.7 \text{ ps} < \sigma_t < 24.3 \text{ ps}$. All fits are shown in the Appendix ???.
473

474 For the combination of the two separate widths σ_1 and σ_2 , a method which takes the
475 damping effect of the finite time resolution on the CP observables into account, is used.
476 The effective damping of the CP amplitudes is described by the dilution \mathcal{D} , which can
477 take values between 1 and 0. In the case of infinite precision, there would be no damping
478 and therefore $\mathcal{D} = 1$ would hold, while for a resolution that is much larger than the B_s^0
479 oscillation frequency, \mathcal{D} would approach 0. For two Gaussians describing the resolution,
480 the dilution can be defined as [14]

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

481 where f_1 is the fraction of events described by the first Gaussian relative to the second
482 and Δm_s is the oscillation period of the B_s^0 meson.

483 The dilution is computed in every bin of the per-event decay time error and can be
484 converted into the effective resolution

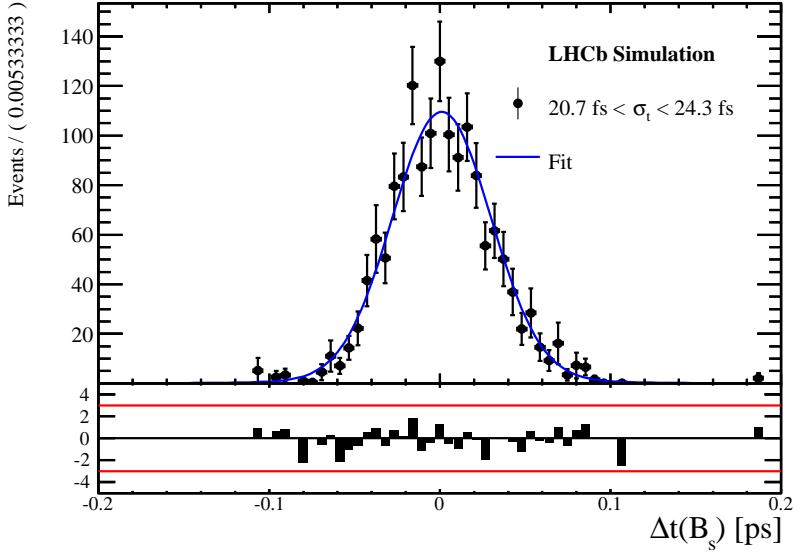


Figure 7.2: Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K\pi\pi$ candidates from MC in the bin $20.7 \text{ ps} < \sigma_t < 24.3 \text{ ps}$. A fit of the sum of two Gaussian functions is overlaid.

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

485 7.4 Results

486 The fitted values for the Gaussian widths σ_1 and σ_2 , the fraction of the first relative to
 487 the second Gaussian function f_1 , as well as the effective resolution σ_{eff} , found in each bin
 488 σ_t , are shown in Tab. 7.1. Figure 7.3 shows the obtained values for σ_{eff} as a function of
 489 the per-event decay time error σ_t . A linear polynom of the form

$$\sigma(\sigma_t)_{mc} = s_0 + s_1 \cdot \sigma_t \quad (7.3)$$

490 is used to parametrise this distribution. The obtained values are

$$\sigma(\sigma_t)_{mc} = 0 + (1.257 \pm 0.017)\sigma_t, \quad (7.4)$$

491 where s_0 is compatible with 0 in the fit and therefore is set to $s_0 = \sigma(\sigma_t = 0) = 0$. For
 492 comparison, the linear scaling functions found for $\sigma(\sigma_t)$ in the $B_s^0 \rightarrow D_s K$ analysis [14] for
 493 MC is also shown in Figure 7.3. Motivated by the similarity between the $B_s^0 \rightarrow D_s K\pi\pi$
 494 and $B_s^0 \rightarrow D_s K$ decay, we assume a comparable scaling relation for data,

$$\frac{\sigma(t)_{D_s K\pi\pi,data}}{\sigma(t)_{D_s K\pi\pi,mc}} \approx \frac{\sigma(t)_{D_s K,data}}{\sigma(t)_{D_s K,mc}}. \quad (7.5)$$

495 This leads to a correction factor

$$\sigma(t)_{D_s K\pi\pi,data} \approx \frac{\sigma(t)_{D_s K,data}}{\sigma(t)_{D_s K,mc}} \cdot \sigma(t)_{D_s K\pi\pi,mc}, \quad (7.6)$$

496 where all elements of the right side of the equation are known.

497

498 Taking the scaling function found in our simulation, as well as input from the $B_s^0 \rightarrow$
 499 $D_s K$ analysis for $\sigma(t)_{D_s K, mc/data}$, we find

$$\sigma(t)_{D_s K \pi\pi, data} = 10.3 fs + 1.28 \cdot t$$

500 ,
 501 which is the scaling factor used for the per-event decay time error in the nominal time-
 502 and amplitude-dependent fit.

σ_t Bin [fs]	σ_1 [fs]	σ_2 [fs]	f_1	D	σ_{eff} [fs]
0to19	22.57 ± 0.96	45.57 ± 4.061	0.827 ± 0.057	0.89 ± 0.067	27.46 ± 8.82
19to24	24.64 ± 1.03	46.65 ± 3.109	0.768 ± 0.061	0.86 ± 0.070	30.64 ± 8.48
24to29	30.96 ± 0.90	58.76 ± 5.684	0.884 ± 0.045	0.83 ± 0.05	34.66 ± 5.28
29to34	35.28 ± 1.54	57 ± 6.698	0.839 ± 0.098	0.79 ± 0.10	39.09 ± 10.47
34to39	37.05 ± 2.36	61.98 ± 5.769	0.707 ± 0.12	0.73 ± 0.12	44.76 ± 11.78
39to44	68.38 ± 8.33	42.15 ± 3.583	0.331 ± 0.18	0.66 ± 0.16	50.98 ± 15.11
44to49	199.9 ± 100.1	53.72 ± 1.419	0.020 ± 0.014	0.62 ± 0.02	54.89 ± 1.60
49to150	68.75 ± 165.3	68.92 ± 4.603	0.001 ± 0.97	0.47 ± 0.65	68.92 ± 63.42

Table 7.1: Summary of the obtained parameters from the resolution fits described above.

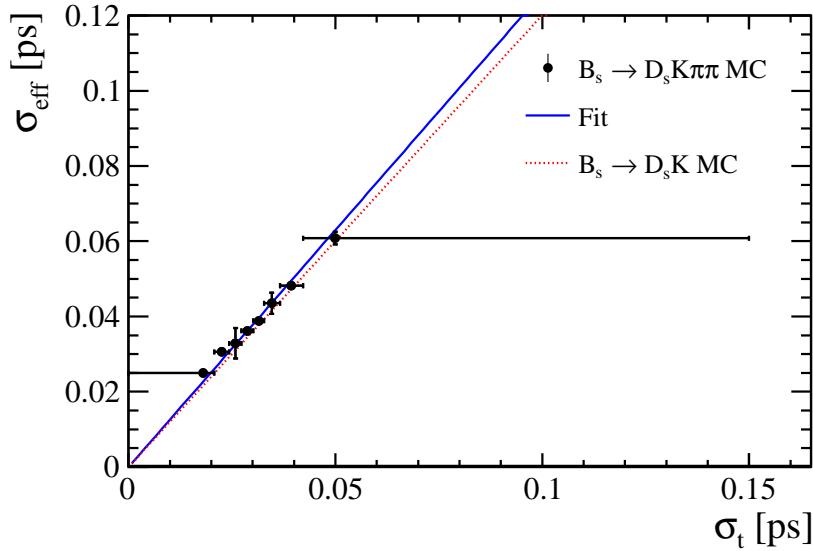


Figure 7.3: Decay-time resolution of $B_s^0 \rightarrow D_s K \pi\pi$ candidates from MC. The scaling functions found in $B_s^0 \rightarrow D_s K$ (dotted red line) MC is also shown for comparison. The fit described in the text is overlaid.

⁵⁰³ **8 B_s Production Asymmetry**

Table 8.1: B_s production asymmetry for 2011 data.

p_T [GeV/c]	y	$A_P(B_s^0)_{\sqrt{s}=7 \text{ TeV}}$
(2.00, 7.00)	(2.10, 3.00)	$0.0166 \pm 0.0632 \pm 0.0125$
(2.00, 7.00)	(3.00, 3.30)	$0.0311 \pm 0.0773 \pm 0.0151$
(2.00, 7.00)	(3.30, 4.50)	$-0.0833 \pm 0.0558 \pm 0.0132$
(7.00, 9.50)	(2.10, 3.00)	$0.0364 \pm 0.0479 \pm 0.0068$
(7.00, 9.50)	(3.00, 3.30)	$0.0206 \pm 0.0682 \pm 0.0127$
(7.00, 9.50)	(3.30, 4.50)	$0.0058 \pm 0.0584 \pm 0.0089$
(9.50, 12.00)	(2.10, 3.00)	$-0.0039 \pm 0.0456 \pm 0.0121$
(9.50, 12.00)	(3.00, 3.30)	$0.1095 \pm 0.0723 \pm 0.0179$
(9.50, 12.00)	(3.30, 4.50)	$0.1539 \pm 0.0722 \pm 0.0212$
(12.00, 30.00)	(2.10, 3.00)	$-0.0271 \pm 0.0336 \pm 0.0061$
(12.00, 30.00)	(3.00, 3.30)	$-0.0542 \pm 0.0612 \pm 0.0106$
(12.00, 30.00)	(3.30, 4.50)	$-0.0586 \pm 0.0648 \pm 0.0150$

Table 8.2: B_s production asymmetry for 2012 data.

p_T [GeV/c]	y	$A_P(B_s^0)_{\sqrt{s}=8 \text{ TeV}}$
(2.00, 7.00)	(2.10, 3.00)	$0.0412 \pm 0.0416 \pm 0.0150$
(2.00, 7.00)	(3.00, 3.30)	$-0.0241 \pm 0.0574 \pm 0.0079$
(2.00, 7.00)	(3.30, 4.50)	$0.0166 \pm 0.0391 \pm 0.0092$
(7.00, 9.50)	(2.10, 3.00)	$0.0482 \pm 0.0320 \pm 0.0067$
(7.00, 9.50)	(3.00, 3.30)	$0.0983 \pm 0.0470 \pm 0.0155$
(7.00, 9.50)	(3.30, 4.50)	$-0.0430 \pm 0.0386 \pm 0.0079$
(9.50, 12.00)	(2.10, 3.00)	$0.0067 \pm 0.0303 \pm 0.0063$
(9.50, 12.00)	(3.00, 3.30)	$-0.1283 \pm 0.0503 \pm 0.0171$
(9.50, 12.00)	(3.30, 4.50)	$-0.0500 \pm 0.0460 \pm 0.0104$
(12.00, 30.00)	(2.10, 3.00)	$-0.0012 \pm 0.0222 \pm 0.0050$
(12.00, 30.00)	(3.00, 3.30)	$0.0421 \pm 0.0416 \pm 0.0162$
(12.00, 30.00)	(3.30, 4.50)	$0.0537 \pm 0.0447 \pm 0.0124$

504 9 Time dependent fit

505 This section will cover the phasespace integrated, time dependent fit to $B_s^0 \rightarrow D_s h\pi\pi$ data,
 506 which is described by the PDF formulated in Eq. 2.6. For the phasespace integrated fit to
 507 $B_s^0 \rightarrow D_s K\pi\pi$ data, the sensitivity to the CKM phase γ will depend on the magnitude of
 508 the coherence factor κ , defined in Eq. 2.7, which is added as an additional fit parameter. In
 509 order to avoid any pollution of the final data samples by background events, both samples
 510 are weighted using the sWeights obtained by the fits to the invariant mass distributions,
 511 described in Sec. 4. This fit approach is commonly known as *sFit*. As additional input to
 512 the fit, the tagging information (Sec. 5), as well as the decay time acceptance (Sec. 6)
 513 and resolution (Sec. 7) is used and fixed to the values obtained by the dedicated studies.
 514 Taking all inputs into account, the final time dependent fit PDF is given by

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

515 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
 516 time acceptance effects and $\mathcal{R}(t - t')$ is the Gaussian time resolution function.

517 9.1 sFit to $B_s^0 \rightarrow D_s \pi\pi\pi$ data

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

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

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

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

529 9.2 sFit to $B_s^0 \rightarrow D_s K\pi\pi$ data

530 9.3 Results

A Details of multivariate classifier

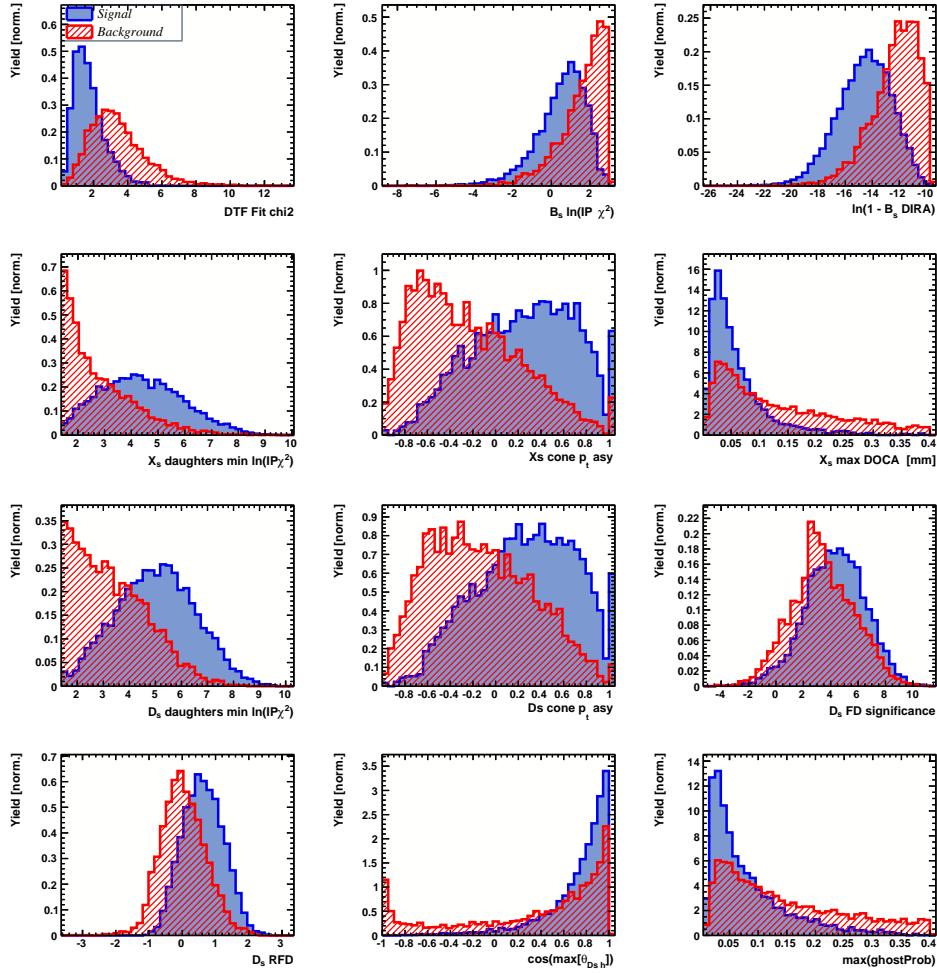


Figure 1.1: Variables used to train the BDTG.

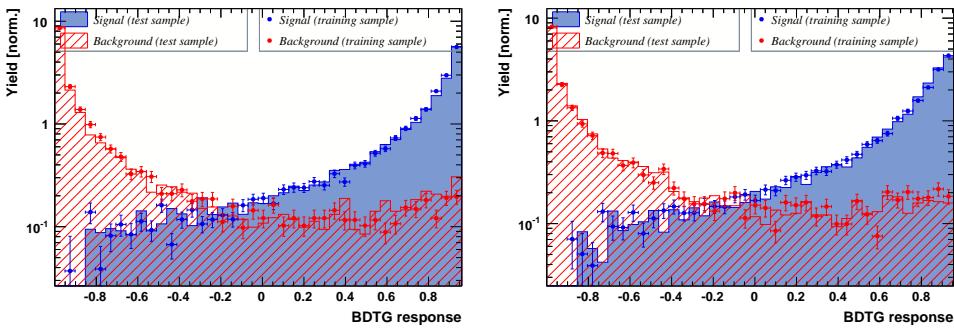


Figure 1.2: Response of the classifier trained on the even (left) and odd (right) sample.

532 B Detailed mass fits

533 In this section, all fits to the mass distribution of $B_s^0 \rightarrow D_s\pi\pi\pi$ and $B_s^0 \rightarrow D_sK\pi\pi$
 534 candidates are shown. The fits are performed simultaneously in run period (Run-I, Run-
 535 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
 536 L0 trigger category.

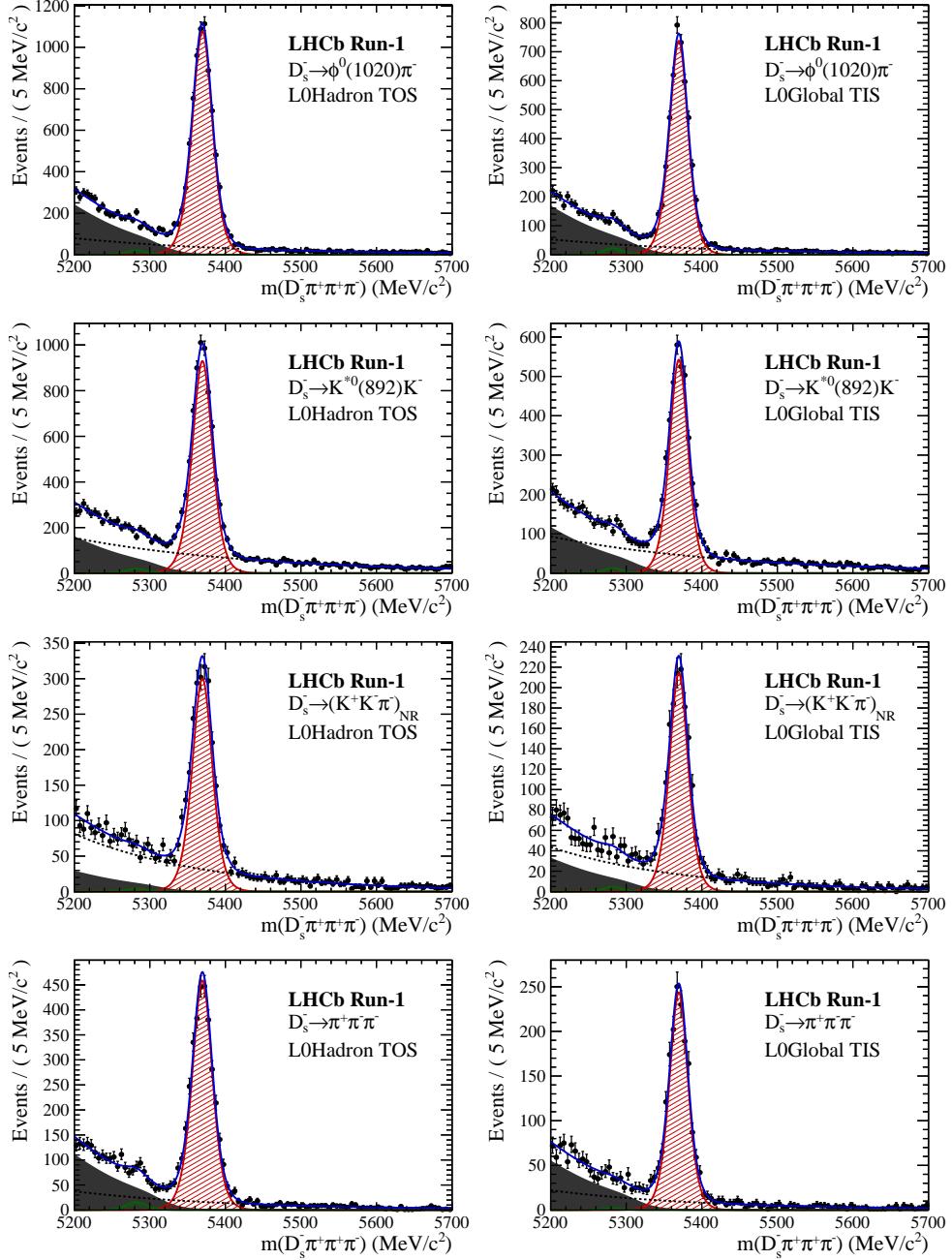


Figure 2.1: Invariant mass distributions of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates for Run-I data.

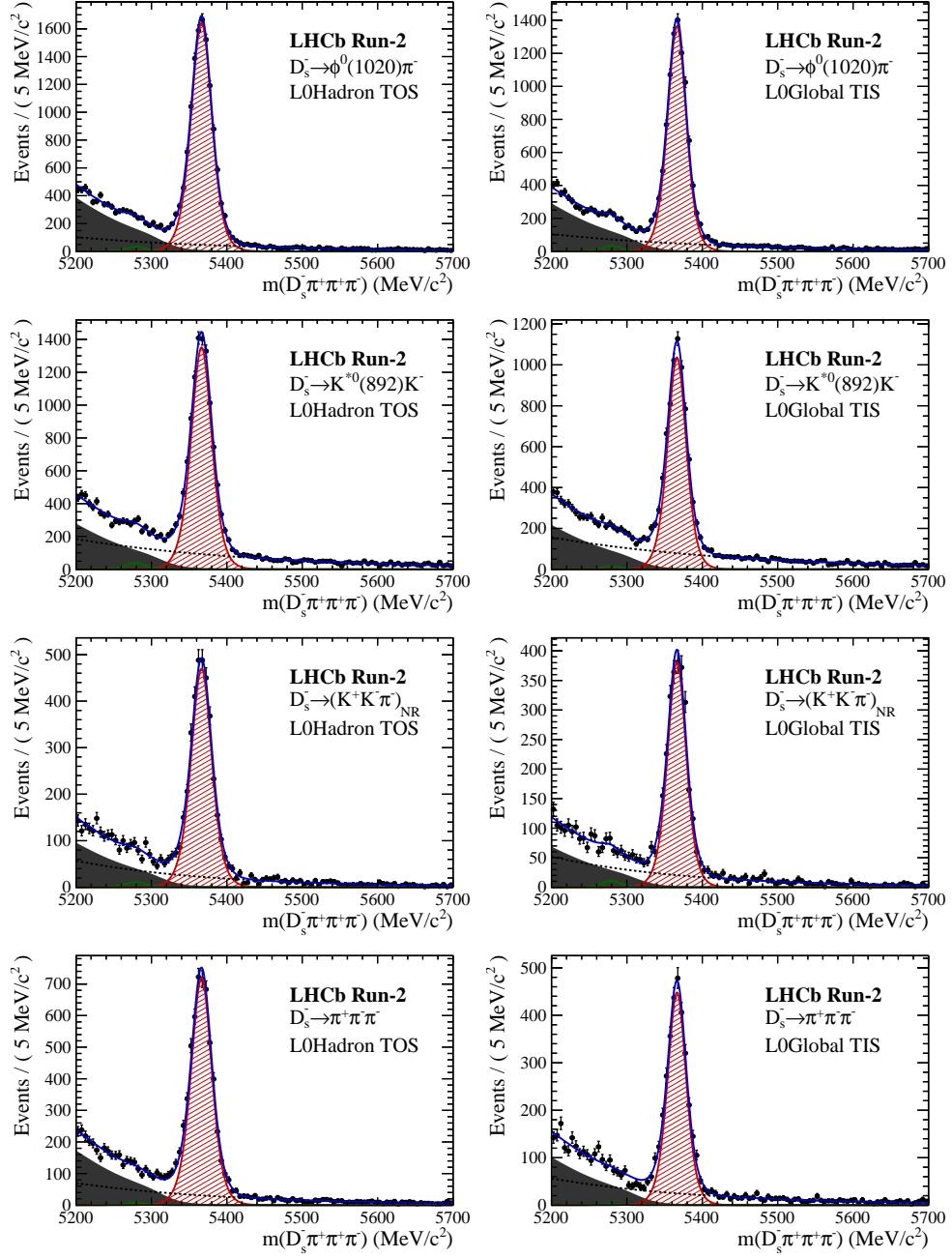


Figure 2.2: Invariant mass distributions of $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates for Run-II data.

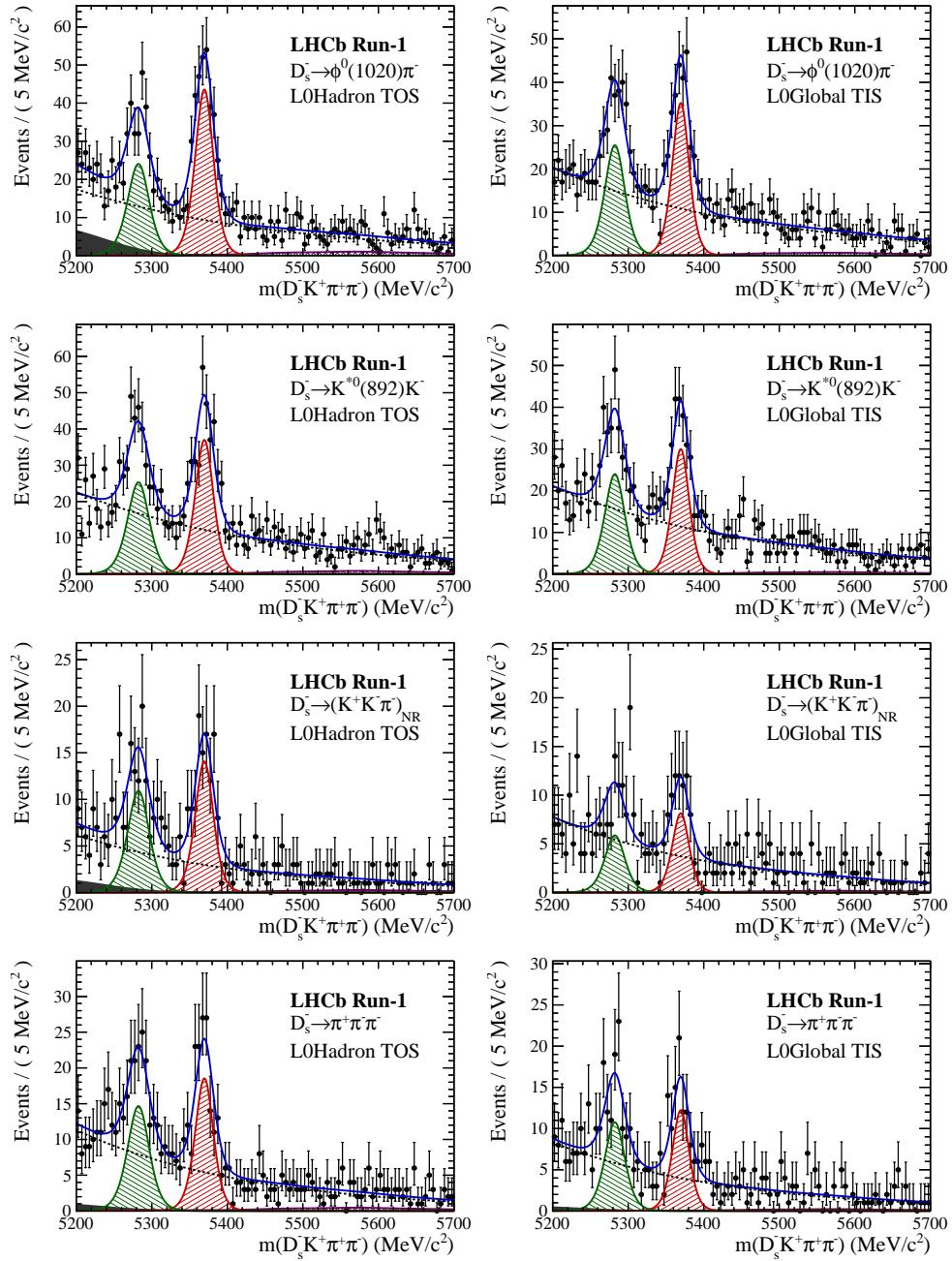


Figure 2.3: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-I data.

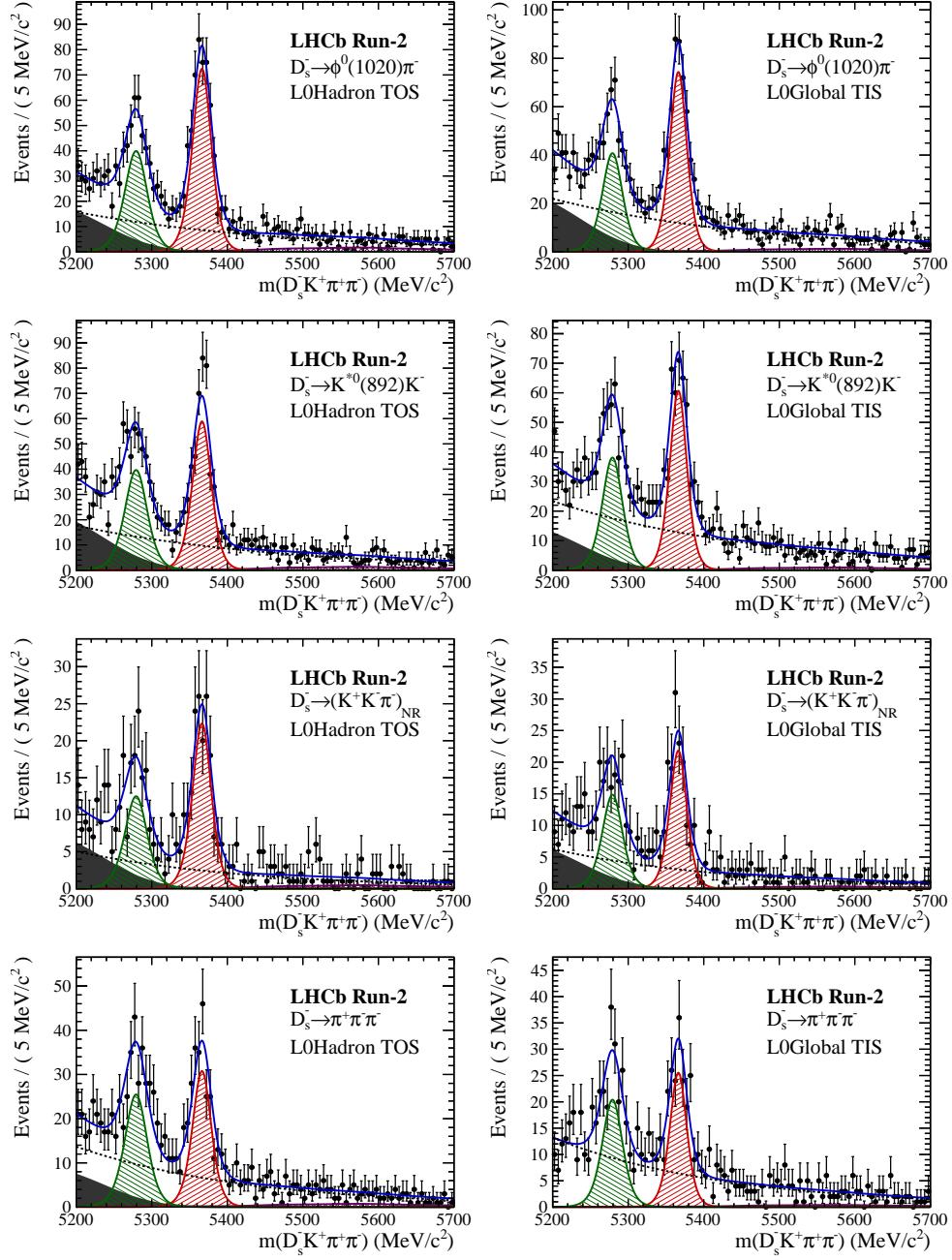


Figure 2.4: Invariant mass distributions of $B_s^0 \rightarrow D_s K \pi \pi$ candidates for Run-II data.

537 C Decay-time Resolution fits

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

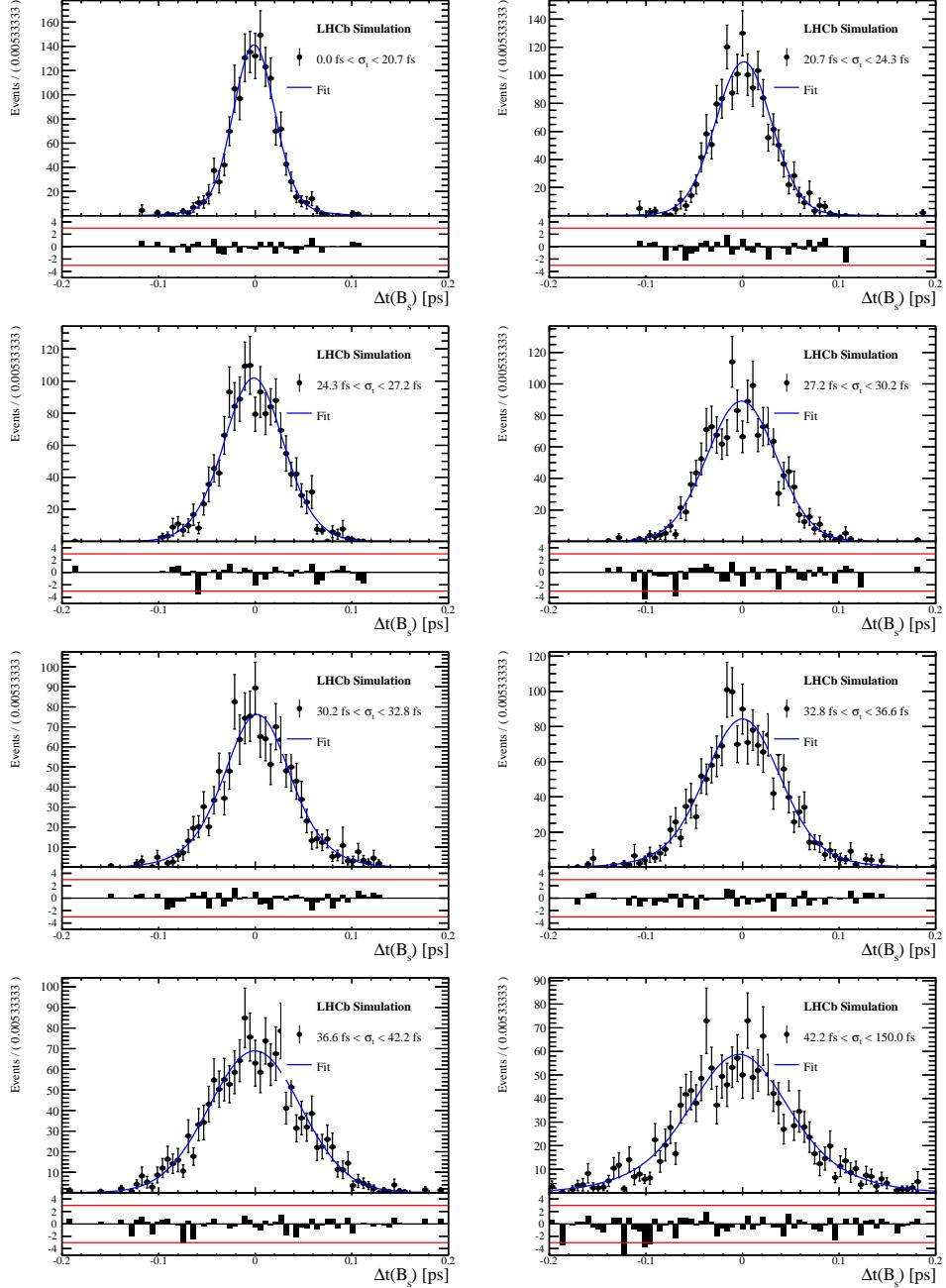


Figure 3.1: Difference of the true and measured decay time of $B_s^0 \rightarrow D_s K\pi\pi$ candidates from MC in all bins. The fit described in 7 is overlaid.

⁵⁴³ D MC corrections

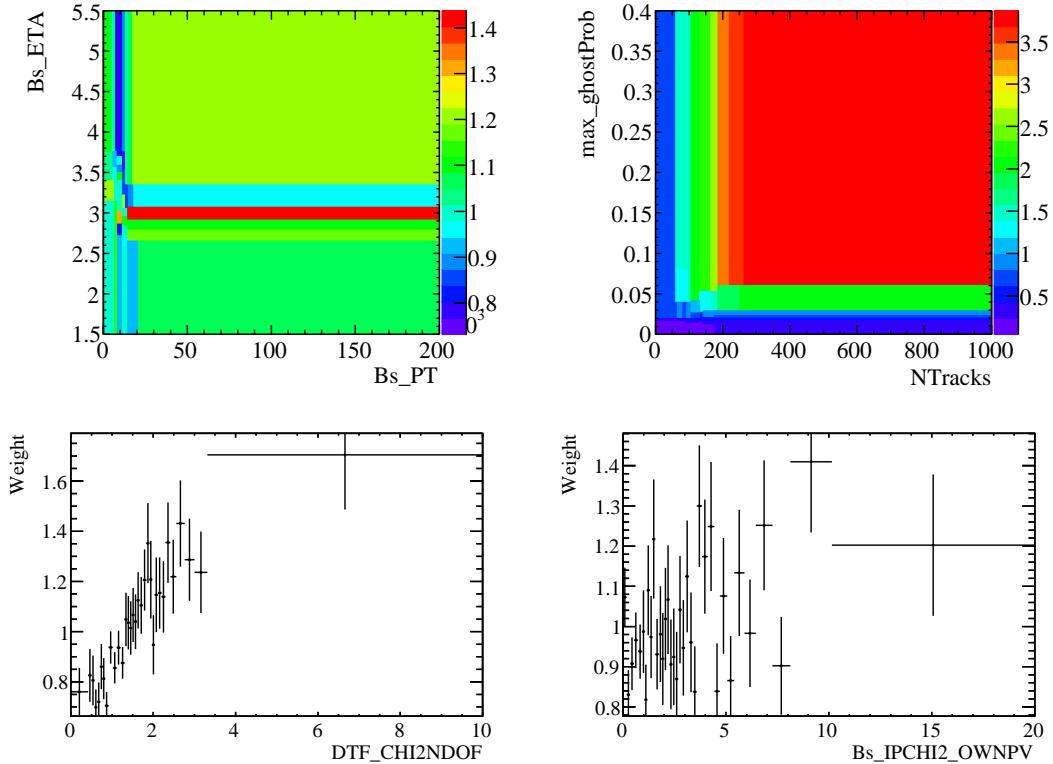


Figure 4.1: Weights applied to correct for Data/MC differences.

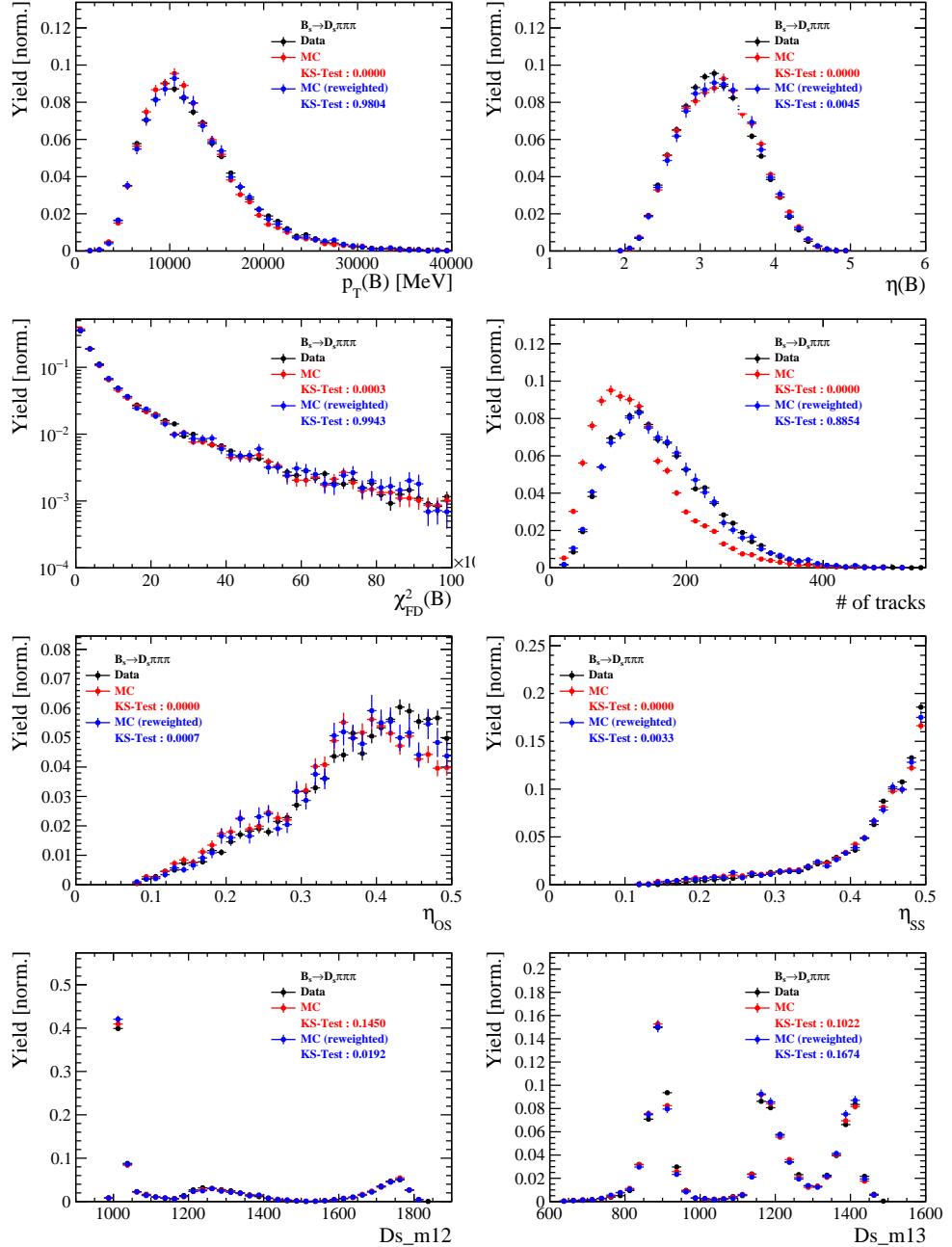


Figure 4.2: Comparison of selected variables for $B_s \rightarrow D_s \pi\pi\pi$ decays.

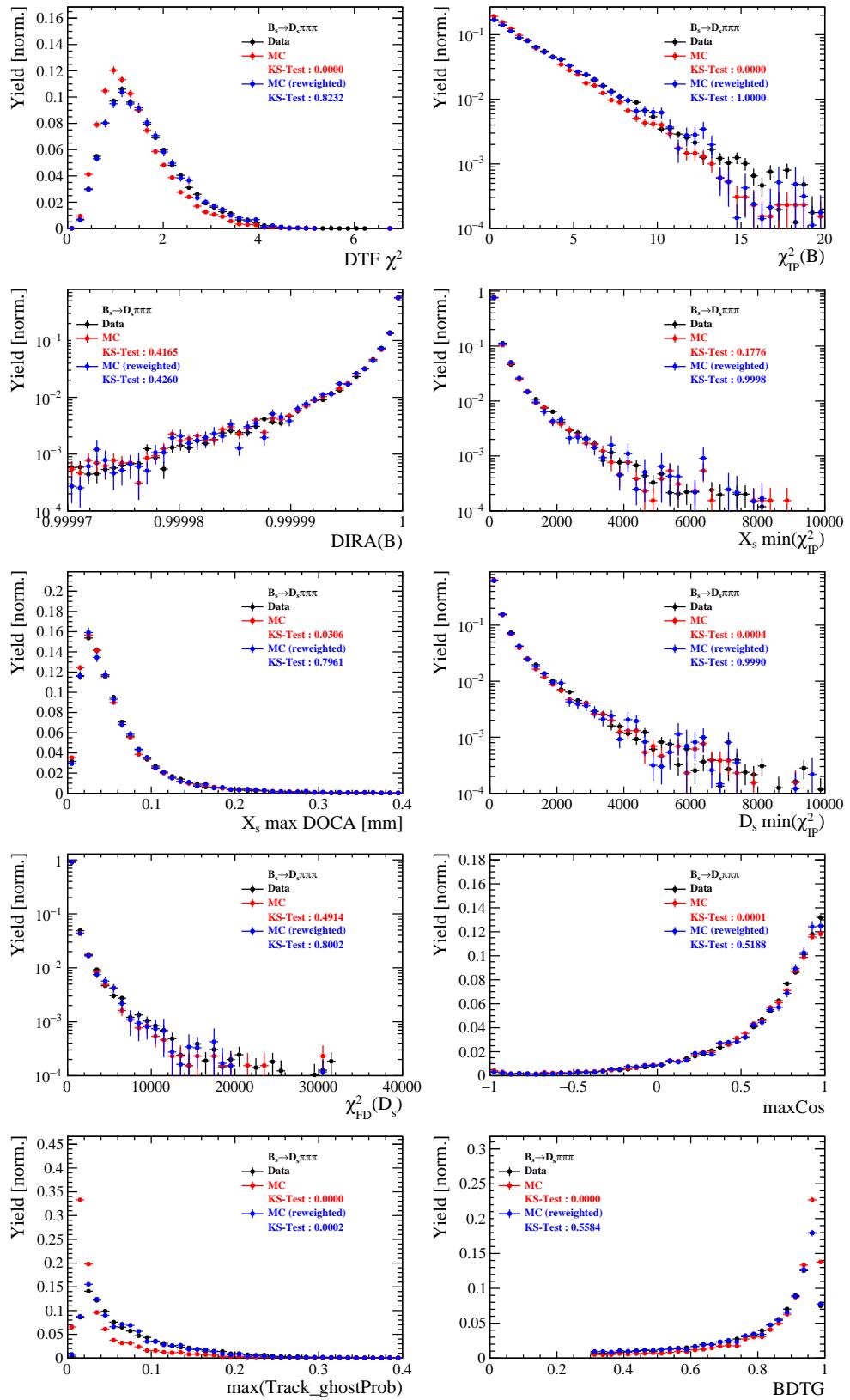


Figure 4.3: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s \pi\pi\pi$ decays.

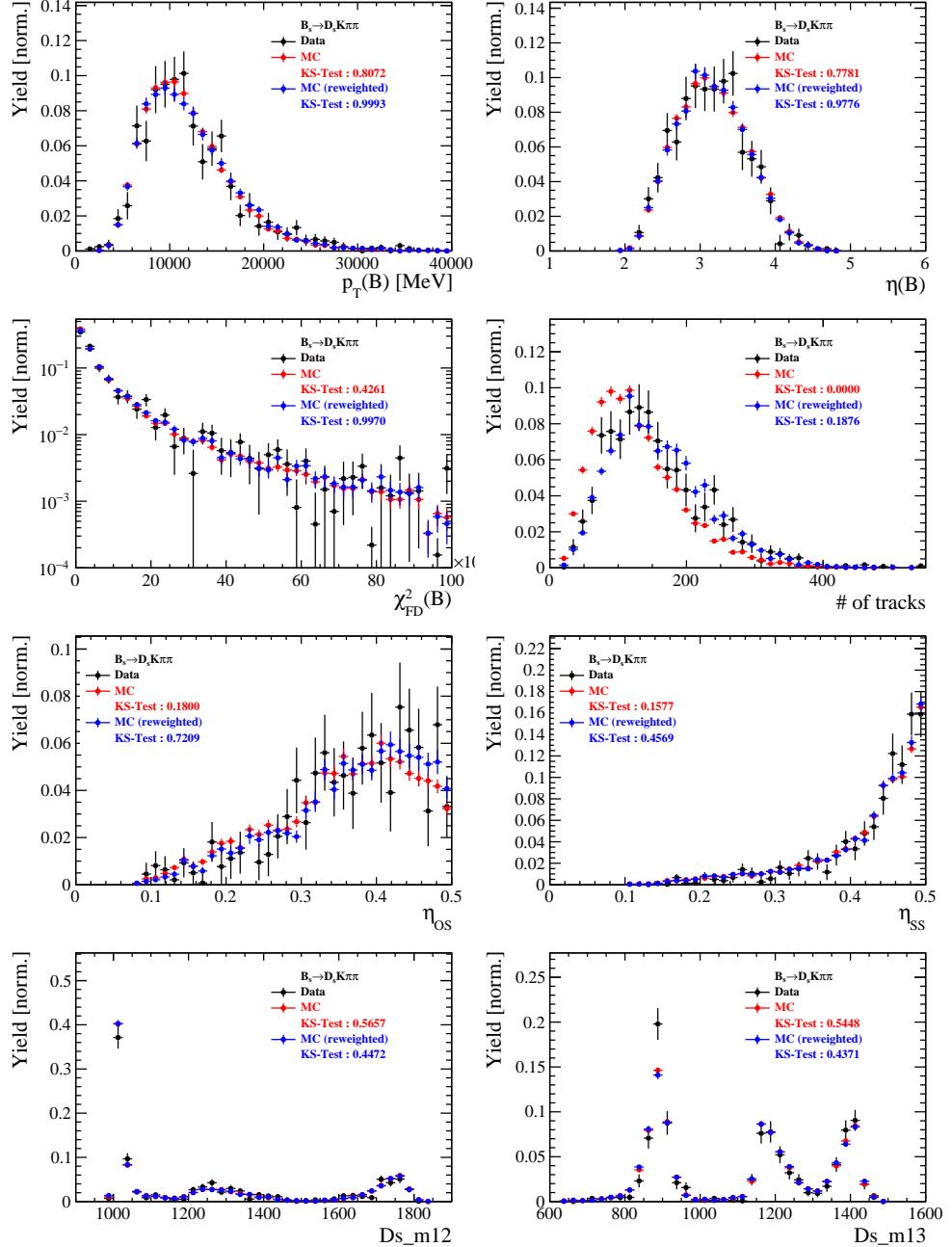


Figure 4.4: Comparison of selected variables for $B_s \rightarrow D_s K\pi\pi$ decays.

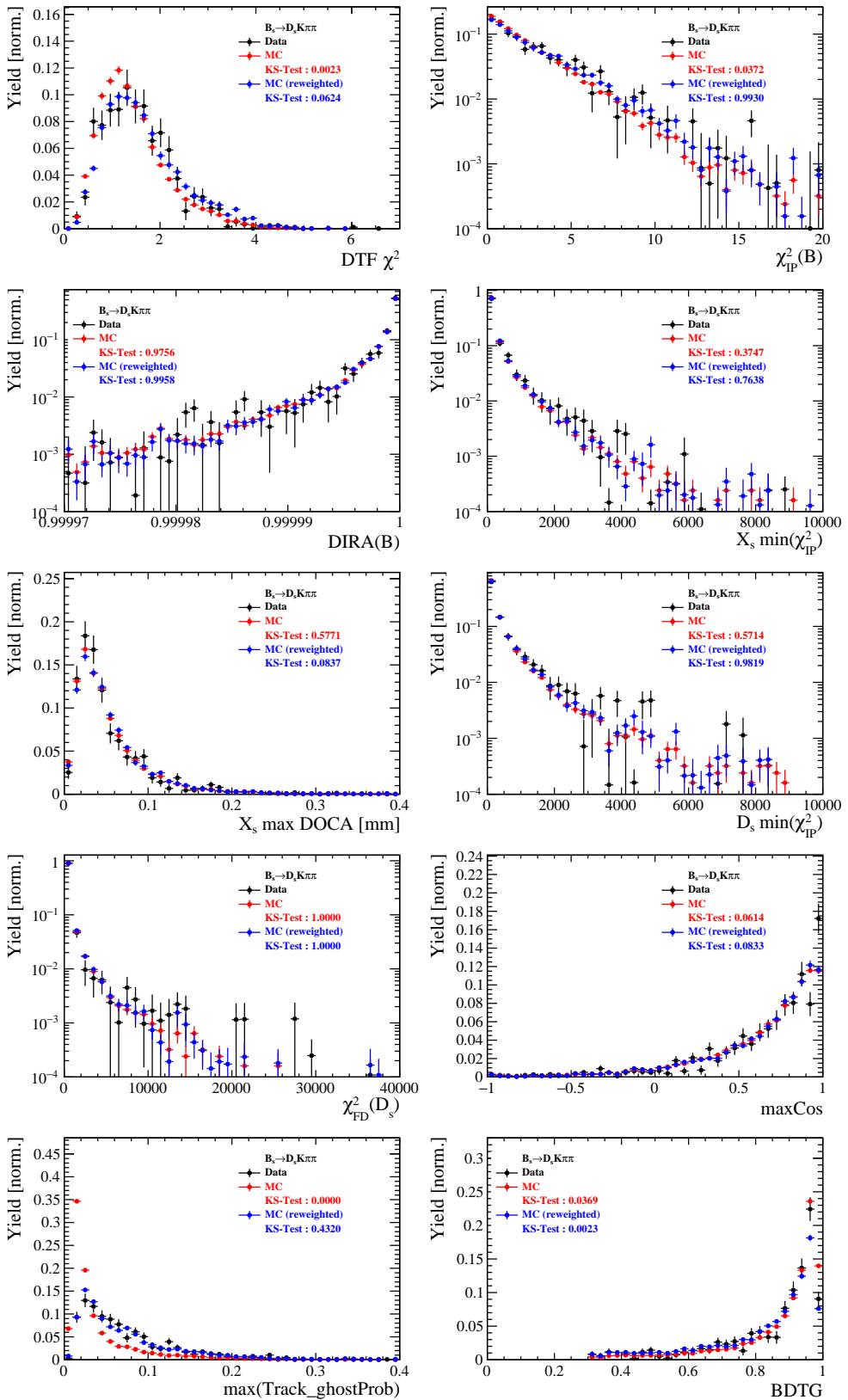


Figure 4.5: Comparison of BDTG input variables and classifier response for $B_s \rightarrow D_s K\pi\pi$ decays.

544 **E Data distributions**

545 **E.1 Comparison of signal and calibration channel**

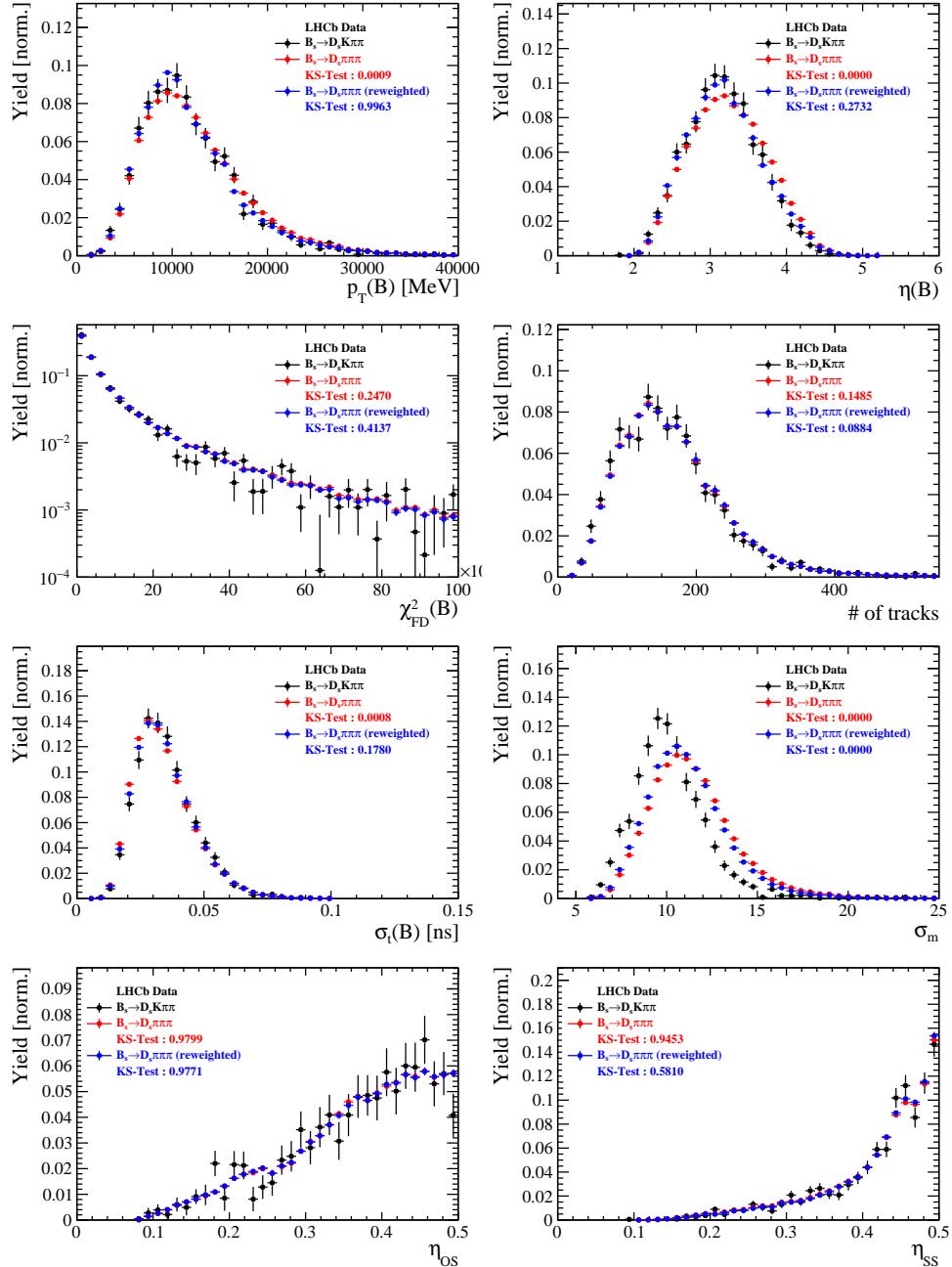


Figure 5.1: Comparison of selected variables.

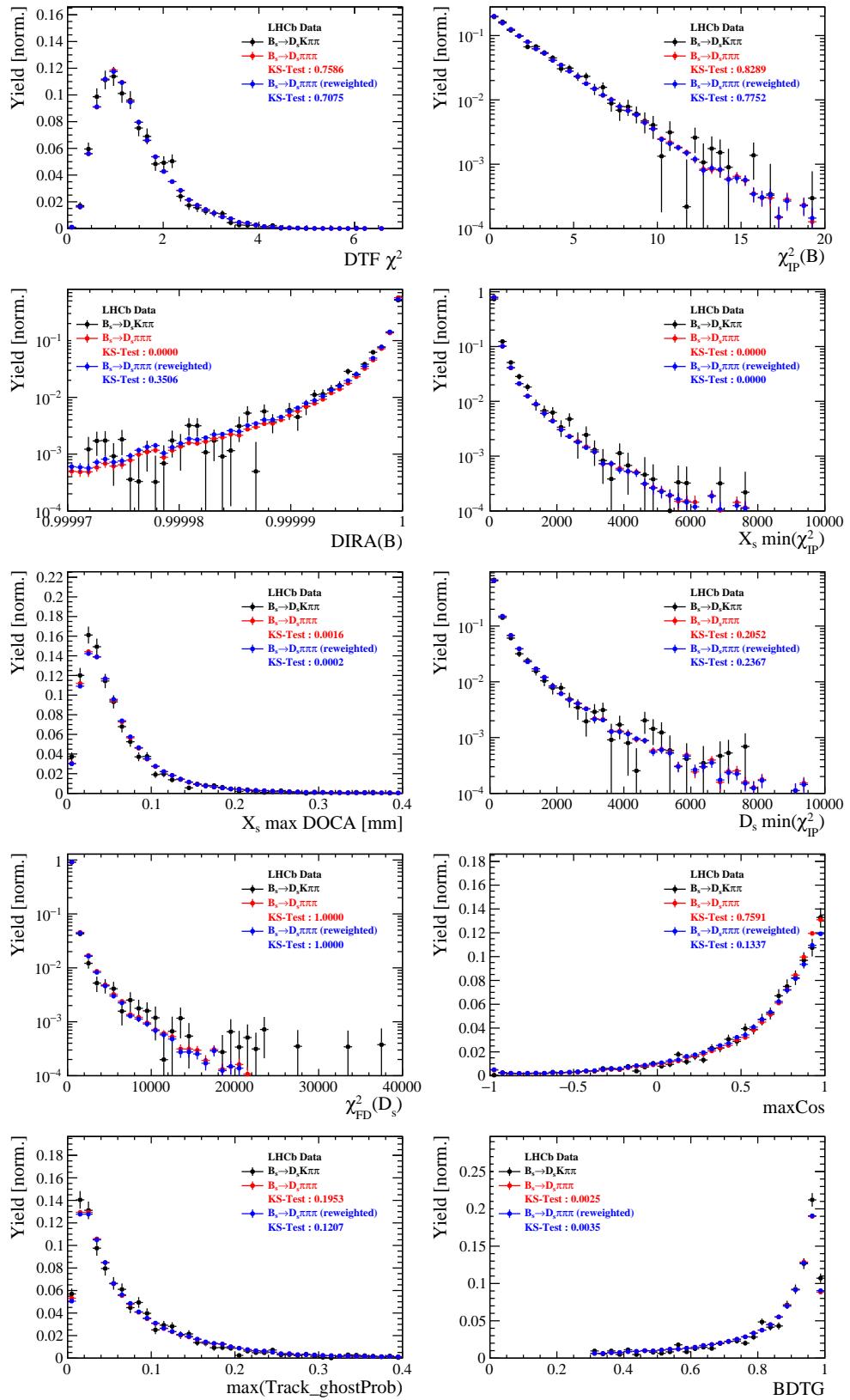


Figure 5.2: Comparison of BDTG input variables and classifier response.

546 E.2 Comparison of Run-I and Run-II data

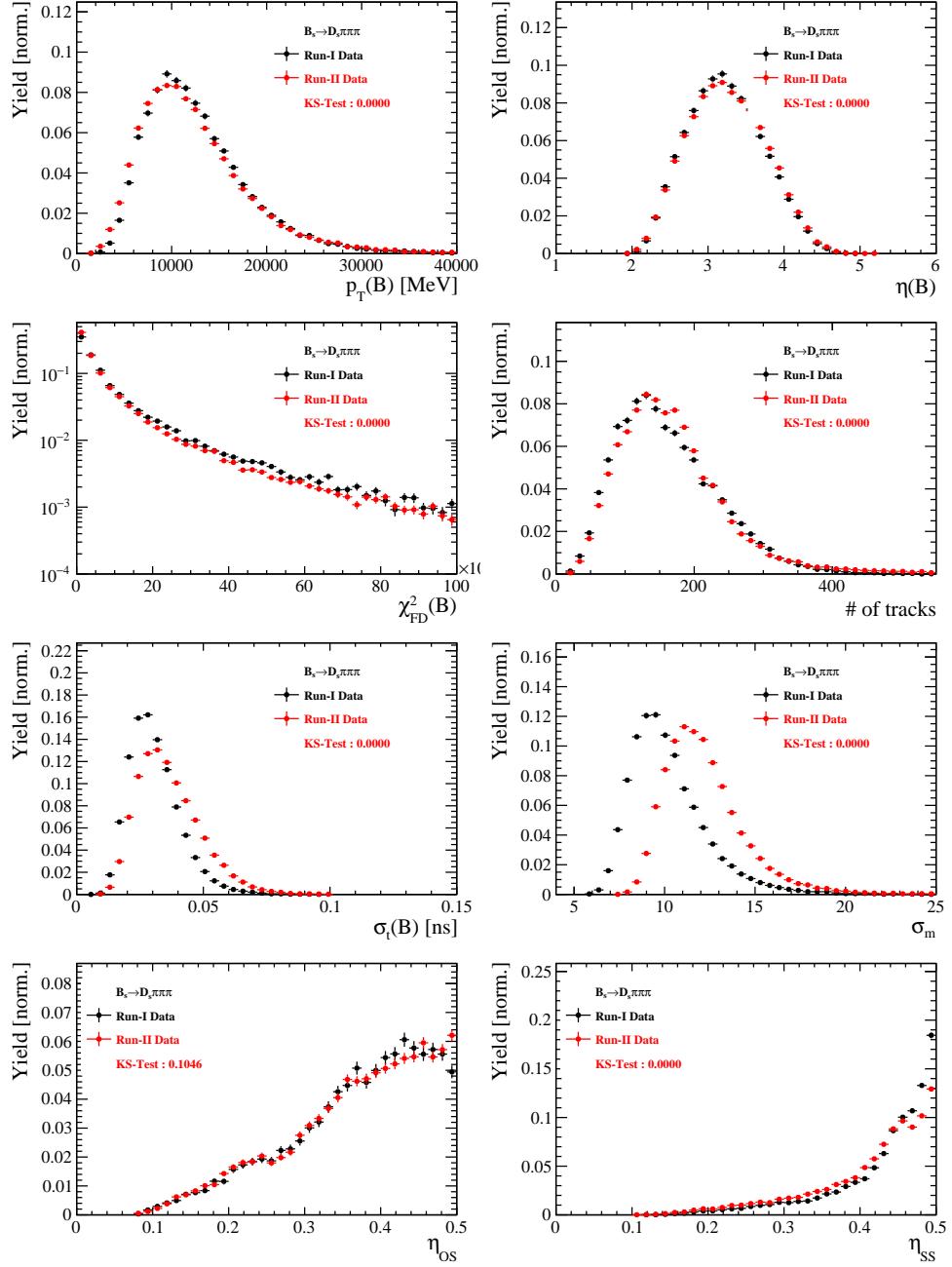


Figure 5.3: Comparison of selected variables.

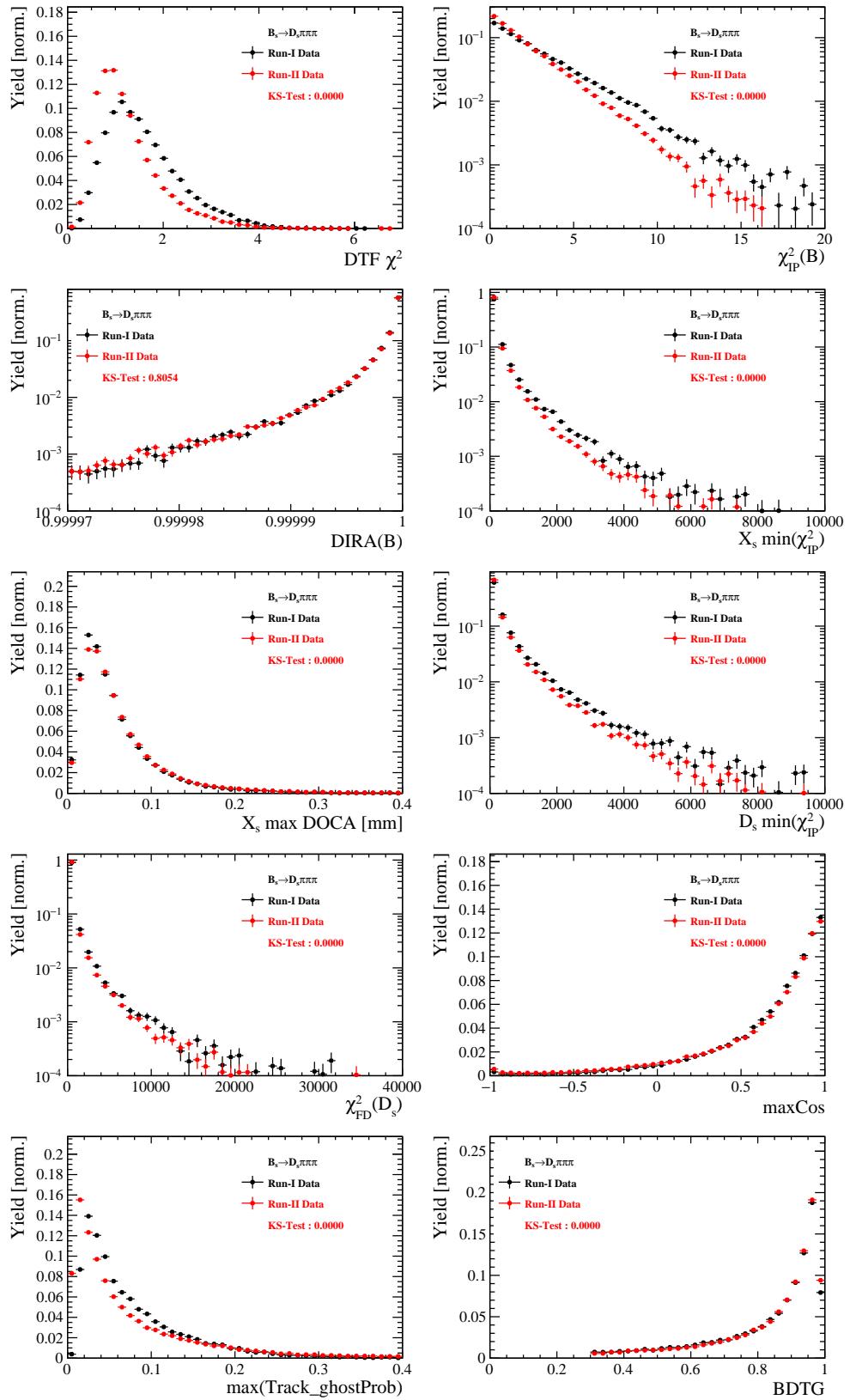


Figure 5.4: Comparison of BDTG input variables and classifier response.

547 E.3 Comparison of D_s final states

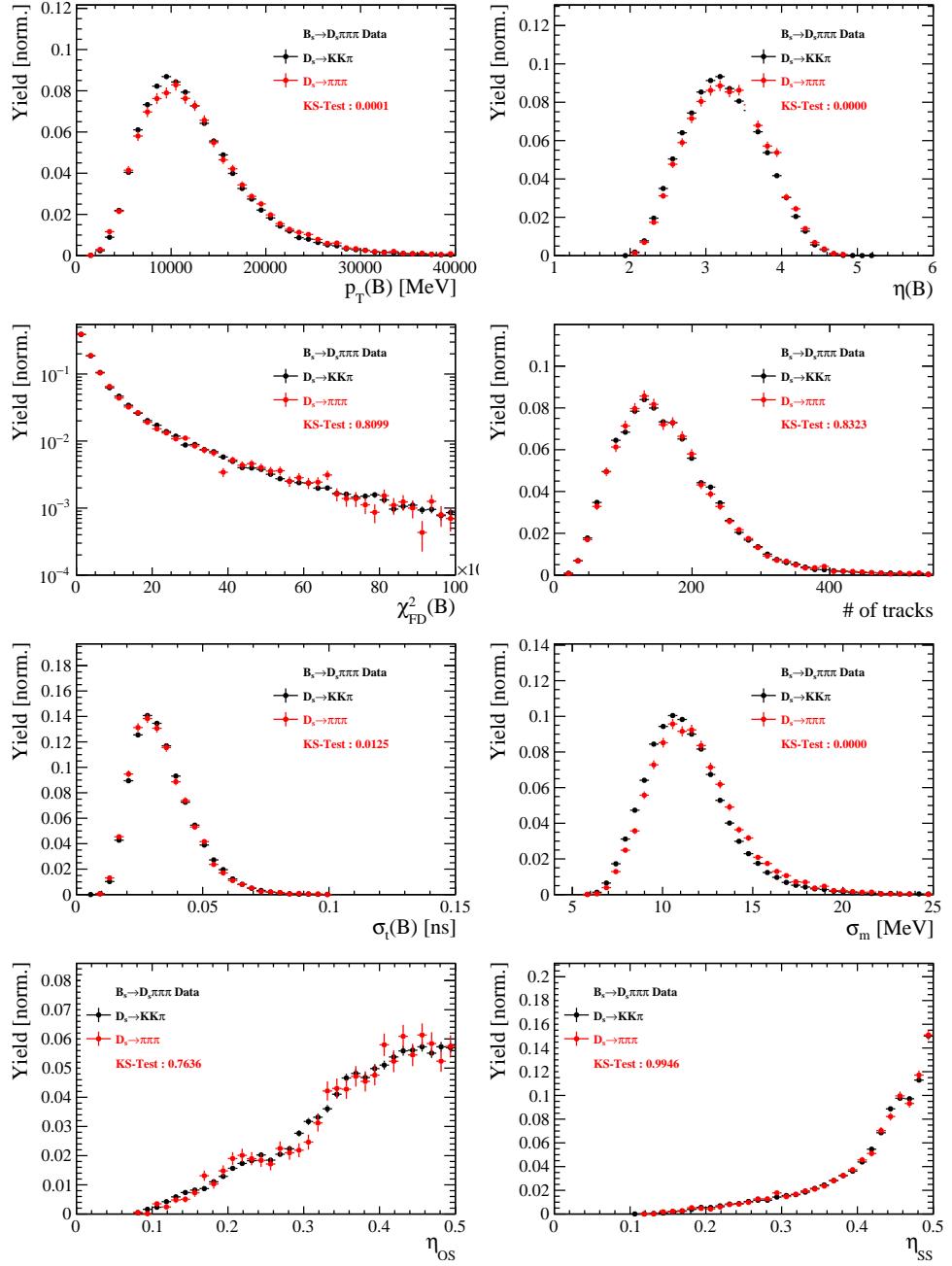


Figure 5.5: Comparison of selected variables.

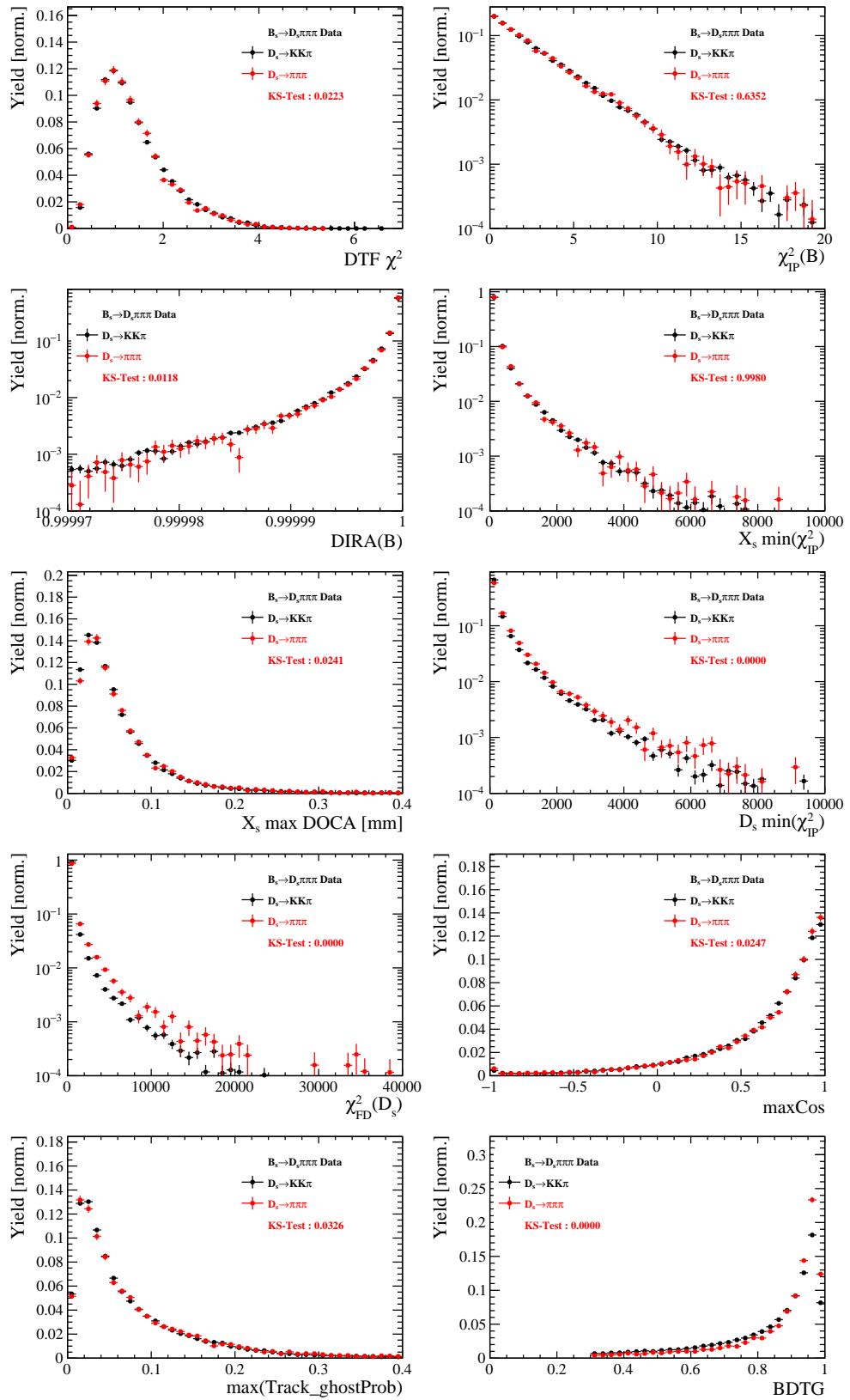


Figure 5.6: Comparison of BDTG input variables and classifier response.

548 E.4 Comparison of trigger categories

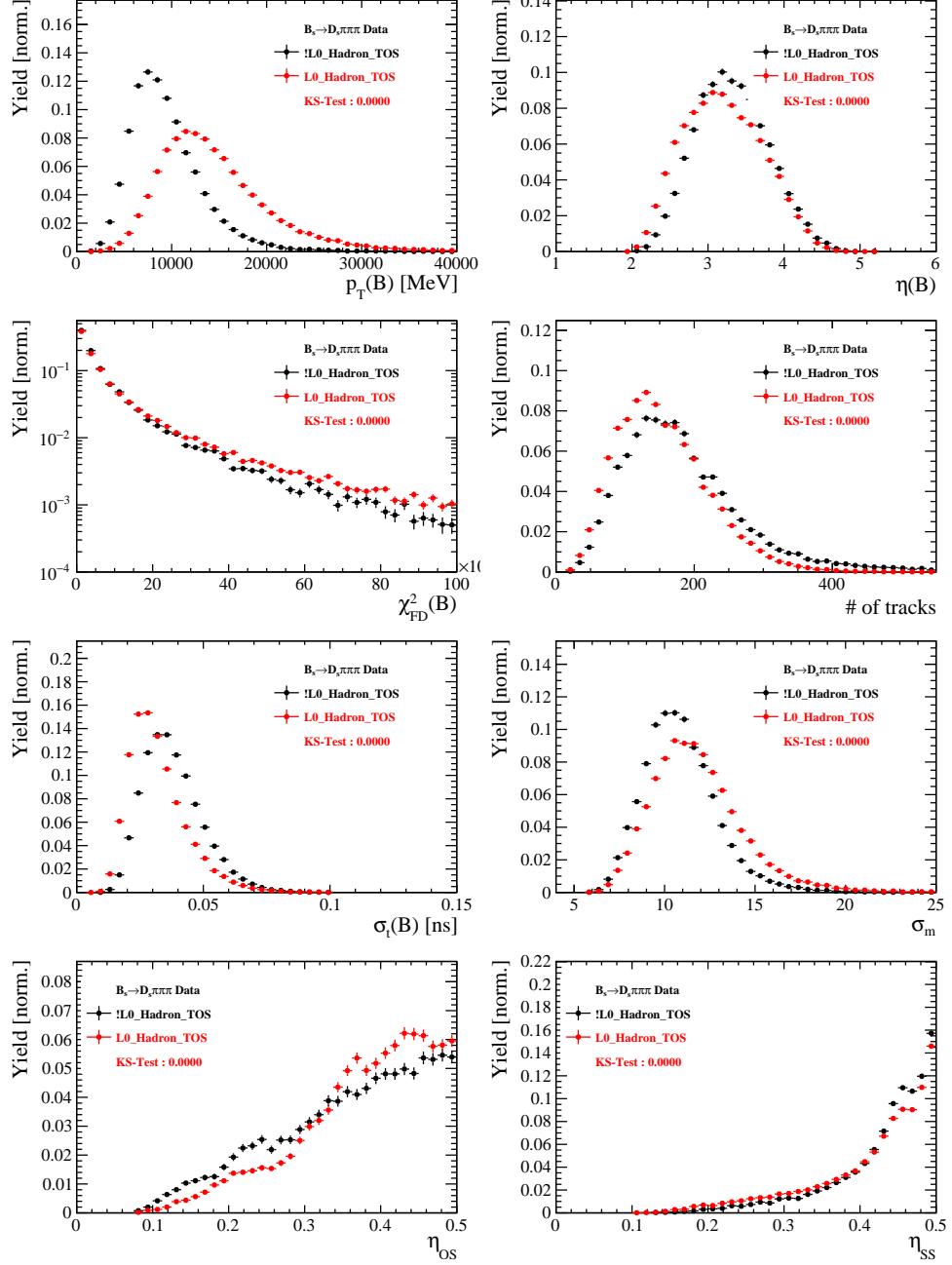


Figure 5.7: Comparison of selected variables.

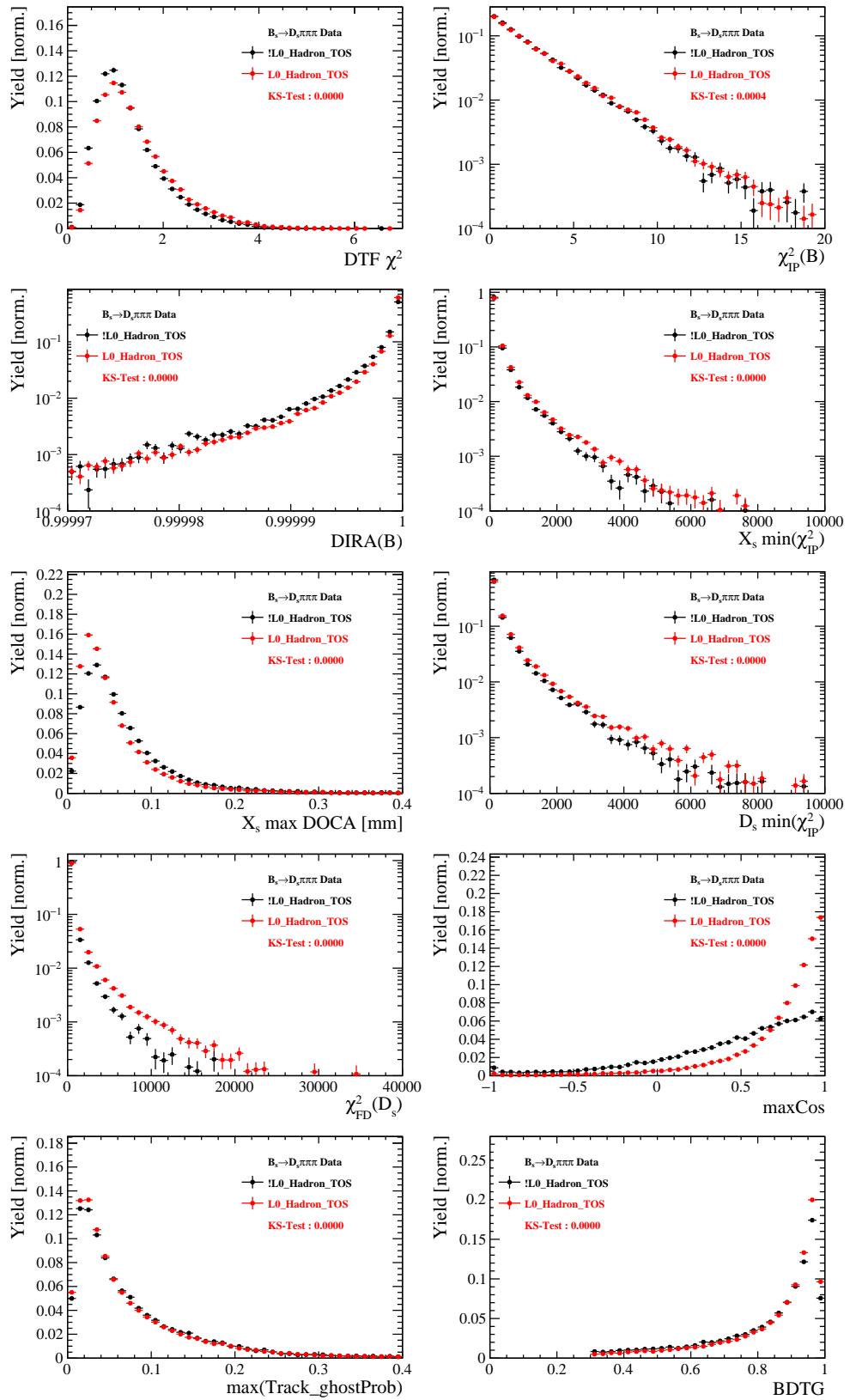


Figure 5.8: Comparison of BDTG input variables and classifier response.

549 E.5 Comparison of B_s and B_d decays

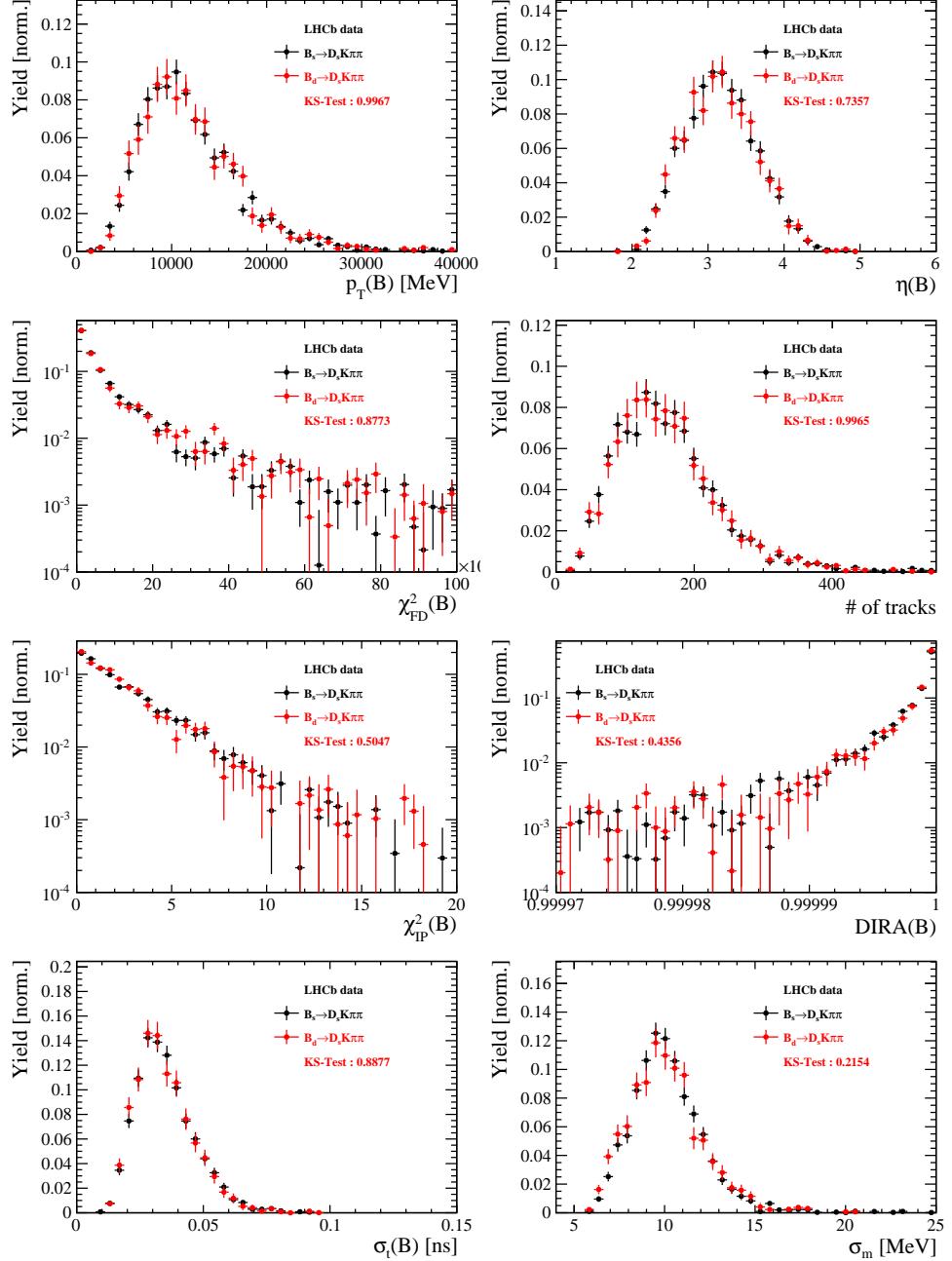


Figure 5.9: Comparison of selected variables.

550 **References**

- 551 [1] R. Fleischer, *New strategies to obtain insights into CP violation through $B(s) \rightarrow D(s)$, $D(s) \rightarrow K\pi$, $D(s) \rightarrow D\pi$, ... and $B(d) \rightarrow D\pi$, $D^* \rightarrow D\pi$, ... decays*, Nucl.
552 Phys. **B671** (2003) 459, arXiv:hep-ph/0304027.
- 553
- 554 [2] K. De Bruyn *et al.*, *Exploring $B_s \rightarrow D_s^{(*)\pm} K^\mp$ Decays in the Presence of a Sizable*
555 *Width Difference $\Delta\Gamma_s$* , Nucl. Phys. **B868** (2013) 351, arXiv:1208.6463.
- 556 [3] S. Blusk, *First observations and measurements of the branching fractions for the*
557 *decays $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* .
- 558 [4] LHCb, S. Blusk, *Measurement of the CP observables in $\bar{B}_s^0 \rightarrow D_s^+ K^-$ and first obser-*
559 *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.
- 560 [5] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007)
561 040, arXiv:physics/0703039.
- 562 [6] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*,
563 Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- 564 [7] N. L. Johnson, *Systems of frequency curves generated by methods of translation*,
565 Biometrika **36** (1949), no. 1/2 149.
- 566 [8] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38**
567 (2014) 090001, and 2015 update.
- 568 [9] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at*
569 *LHCb*, JINST **11** (2016), no. 05 P05010, arXiv:1602.07252.
- 570 [10] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the*
571 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- 572 [11] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and*
573 *τ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots
574 available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 575 [12] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson*
576 *mixing and their efficient evaluation*, arXiv:1407.0748.
- 577 [13] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys.
578 **A30** (2015) 1530022, arXiv:1412.6352.
- 579 [14] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*,
580 Submitted to: JHEP (2017) arXiv:1712.07428.