



# Measurement of the branching fraction for the decay $B_s^0 \rightarrow D_s K \pi \pi$ relative to the $B_s^0 \rightarrow D_s \pi \pi \pi$ decay

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

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

<sup>2</sup>*European Organization for Nuclear Research (CERN), Geneva, Switzerland*

## Abstract

We present the measurement of the branching fraction of decay  $B_s^0 \rightarrow D_s K \pi \pi$  using the complete  $3 \text{ fb}^{-1}$  of data, collected during Run 1 of the LHC. The branching fraction is measured relative to the decay  $B_s^0 \rightarrow D_s \pi \pi \pi$ , from which we obtain

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s K \pi \pi)}{\mathcal{B}(B_s^0 \rightarrow D_s \pi \pi \pi)} = 0.056 \pm 0.003 \pm 0.003$$

The  $B_s^0 \rightarrow D_s K \pi \pi$  decay can be further used to measure the weak CKM phase  $\gamma$  in a time-dependent analysis of the  $B_s^0$  and  $\bar{B}_s^0$  decay rates. This will be the final goal of the presented analysis.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data samples</b>	<b>1</b>
<b>3</b>	<b>Selection</b>	<b>2</b>
3.1	Cut-based selection . . . . .	2
3.2	Multivariate stage . . . . .	4
<b>4</b>	<b>Simulated samples</b>	<b>7</b>
<b>5</b>	<b>Models for signal and background components in invariant mass spectrum</b>	<b>9</b>
5.1	Signal models for $m(D_s\pi\pi\pi)$ and $m(D_sK\pi\pi)$ . . . . .	10
5.2	Background models for $m(D_s\pi\pi\pi)$ . . . . .	11
5.3	Background models for $m(D_sK\pi\pi)$ . . . . .	11
<b>6</b>	<b>Mass fits for signal and normalization</b>	<b>13</b>
6.1	Fit to $B_s^0 \rightarrow D_s\pi\pi\pi$ candidates . . . . .	13
6.2	Fit to $B_s^0 \rightarrow D_sK\pi\pi$ candidates . . . . .	13
<b>7</b>	<b>Efficiency corrections</b>	<b>13</b>
7.1	Relative efficiency for BR measurement . . . . .	14
<b>8</b>	<b>Systematic errors</b>	<b>15</b>
<b>9</b>	<b>Results and summary</b>	<b>19</b>
<b>A</b>	<b>Appendix</b>	<b>21</b>
A.1	High Level Trigger requirements . . . . .	21
A.2	Re-weghted MC observables . . . . .	21
	<b>References</b>	<b>24</b>

# 1 Introduction

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

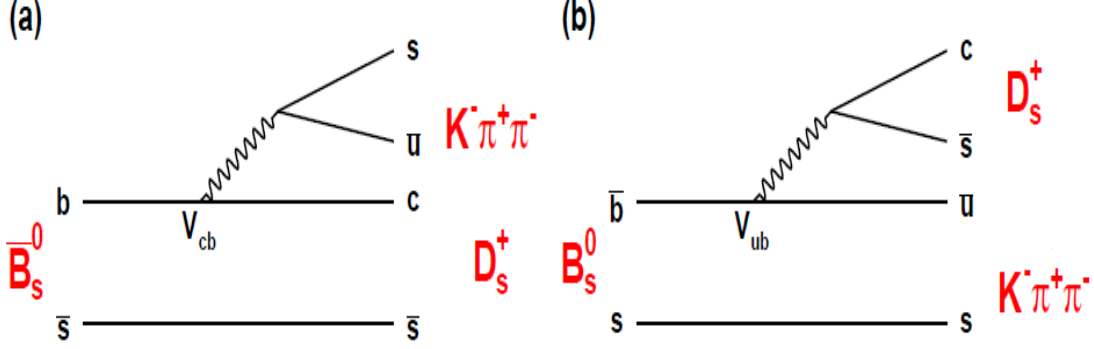


Figure 1.1: Feynman diagram of the  $B_s^0 \rightarrow D_s K \pi \pi$  decay, proceeding via a)  $b \rightarrow c$  transitions or b)  $b \rightarrow u$  transitions.

To measure the weak CKM phase  $\gamma \equiv \arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$ , a decay with interference between  $b \rightarrow c$  and  $b \rightarrow u$  transitions at tree level is needed [1]. As illustrated in Fig. 1.1, this is the case for the presented decay mode. A measurement of  $\gamma$  using  $B_s^0 \rightarrow D_s K \pi \pi$  decays, where the  $K \pi \pi$  subsystem is dominated by excited kaon states such as the  $K_1(1270)$  and  $K_1(1400)$  resonances, will succeed the branching ratio measurement presented in this note. It is complementary to the above mentioned analysis of  $B_s^0 \rightarrow D_s K$ , making use of a fully charged final state, where every track is detected in the vertex locator. To account for the non-constant strong phase across the Dalitz plot, one can either develop a time-dependent amplitude model or select a suitable phase-space region and introduce a coherence factor as additional hadronic parameter to the fit.

This analysis is based on the first observation of the  $B_s^0 \rightarrow D_s K \pi \pi$  decay presented in [3] and [4], where its branching ratio is measured relative to  $B_s^0 \rightarrow D_s \pi \pi \pi$ . The result obtained by the previous analysis is  $0.052 \pm 0.005 \pm 0.003$ , where the uncertainties are statistical and systematical, respectively. The branching ratio measurement is updated, exploiting the full Run 1 data sample, corresponding to  $3 \text{ fb}^{-1}$  of integrated luminosity.

## 2 Data samples

We use the full Run 1 sample from Stripping 21 with Reco14, consisting of  $3 \text{ fb}^{-1}$  of data, collected in the years 2011 and 2012 at a center of mass energies of 7 TeV and 8 TeV, respectively. The selected  $B_s^0$ -candidates are required to pass the L0 Hadron trigger on

24 signal (TOS) or the L0 Global trigger independent of signal (TIS).  
 25 Events that pass the L0 stage are further required to pass the HLT1 TrackAllL0 trigger  
 26 on signal (TOS).  
 27 All remaining candidates have to pass either the 2, 3 or 4-body topological trigger (TOS)  
 28 of the HLT2 stage. More information on the HLT lines used in this analysis is given in  
 29 Appendix A.1.  
 30 For the presented analysis the B02DKPiPiD2HHHPIDBeauty2CharmLine is used to pres-  
 31 elect signal  $B_s^0 \rightarrow D_s K \pi \pi$  candidates and the B02DPiPiPiD2HHHPIDBeauty2CharmLine  
 32 is used for the selection of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates. Both stripping lines employ the same  
 33 selection cuts for the  $B_s^0$ ,  $D_s$ ,  $K$  and  $\pi$  candidates. A summary of those cuts can be found in  
 34 Table 2.1. In this table and throughout the note, we abbreviate  $B_s^0 \rightarrow D_s X_s (\rightarrow K \pi \pi)$  and  
 35  $B_s^0 \rightarrow D_s X_d (\rightarrow \pi \pi \pi)$ , identifying  $X_s \rightarrow K \pi \pi$  and  $X_d \rightarrow \pi \pi \pi$  as the various resonances  
 36 through which the decays proceed. The following text describes variables that might be  
 37 ambiguous and/or cuts which benefits are not straight forward.  
 38 In Table 2.1, DOCA is the abbreviation for distance of closest approach. This variable is  
 39 used to ensure that all  $D_s$  and  $X_{s,d}$  daughters originate from the same vertex. The minimal  
 40 flight distance (FD)  $\chi^2$  is a measure on how likely a particle traveled some distance before  
 41 it decayed. A cut on this variable is employed to reject prompt background for  $D_s$  and  $X_{s,d}$   
 42 candidates. DIRA is the abbreviation for the cosine of the angle  $\theta$  between the hadron's  
 43 flight direction  $\vec{x}$  and it's corresponding momentum vector  $\vec{p}$ ,  $\cos \theta_{\vec{x}-\vec{p}}$ . For signal hadrons  
 44 this variable is expected to be very close to one, while it can be arbitrary distributed for  
 45 background.

## 46 **3 Selection**

47 A two-fold approach is used to isolate the  $B_s^0 \rightarrow D_s K \pi \pi$  candidates from data passing  
 48 the stripping line. First, further one-dimensional cuts are applied to reduce the level of  
 49 combinatorial background and to veto some specific physical background. After that, a  
 50 multivariate classifier is trained which combines the information of several input vari-  
 51 ables, including their correlation, into one powerful discriminator between signal and  
 52 combinatorial background.

### 53 **3.1 Cut-based selection**

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

- 56 (i)  $\text{DIRA} > 0.99994$
- 57 (ii)  $\min \text{IP } \chi^2 < 20$  to any PV
- 58 (iii)  $\text{FD } \chi^2 > 100$  to any PV
- 59 (iv)  $\text{Vertex } \chi^2/\text{nDoF} < 8$

Variable	Stripping Cut
Track $\chi^2/\text{nDoF}$	$< 3$
Track $p$	$> 1000 \text{ MeV}/c$
Track $p_T$	$> 100 \text{ MeV}/c$
Track IP $\chi^2$	$> 4$
$D_s$ Daughter $p_T$	$\Sigma_{i=1}^3 p_i > 1800 \text{ MeV}/c$
$D_s$ Daughter DOCA	$< 0.5 \text{ mm}$
$D_s$ mass $m_{D_s}$	within $\pm 40 \text{ MeV}/c^2$ of PDG value
$D_s$ Vertex $\chi^2/\text{nDoF}$	$< 10$
$D_s$ min FD $\chi^2$	$> 36$
$D_s$ FD $\chi^2 > 2$ to any PV	
$X_d$ Daughter $p_T$	$> 2 \text{ GeV}/c$
$X_{s,d}$ Daughter DOCA	$< 0.4 \text{ mm}$
$X_{s,d}$ Daughter $p_T$	$\Sigma_{i=1}^3 p_{t,i} > 1250 \text{ MeV}/c$
$X_{s,d}$ Vertex $\chi^2/\text{nDoF}$	$< 8$
$X_{s,d}$ min FD $\chi^2/\text{nDoF}$	$> 16$
$X_{s,d}$ DIRA	$> 0.98$
$X_{s,d}$ $\Delta\rho$ (vertex displacement perpendicular to z-axis)	$> 0.1 \text{ mm}$
$X_{s,d}$ $\Delta Z$ (vertex displacement along z-axis)	$> 2.0 \text{ mm}$
$B_s^0$ DIRA	$> 0.98$
$B_s^0$ min IP $\chi^2$	$< 25$
$B_s^0$ Vertex $\chi^2/\text{nDoF}$	$< 10$
$B_s^0$ $\tau_{B_s^0}$	$> 0.2 \text{ ps}$
$K$ DLL $_{K\pi}$	$> -5$
$\pi$ DLL $_{K\pi}$	$< 10$

Table 2.1: Summary of the stripping selections for  $B_s^0 \rightarrow D_s K \pi \pi$  decays.

(v)  $(Z_{D_s} - Z_{B_s^0}) > 0$ , where  $Z_M$  is the z-component of the position  $\vec{x}$  of the decay vertex for the  $B_s^0/D_s$  meson

Additionally, we veto various physical backgrounds, which have either the same final state as our signal decay, or can contribute via a single misidentification of  $K \rightarrow \pi$  or  $K \rightarrow p$ :

- $B_s^0 \rightarrow D_s^+ D_s^- : |M(K\pi\pi) - m_{D_s}| > 20 \text{ MeV}/c^2$
- $B_s^0 \rightarrow D_s^- K^+ K^- \pi^+ : \text{possible with single missID of } K^- \rightarrow \pi^-, \text{ rejected by requiring } \pi^- \text{ to fulfill } \text{DLL}_{K\pi} < 5$
- $B^0 \rightarrow D^+(\rightarrow K^+ \pi^- \pi^+) K \pi \pi : \text{possible with single missID of } \pi^+ \rightarrow K^+, \text{ vetoed by changing particle hypothesis and recompute } |M(K^+ \pi^- \pi^+) - m_{D_p}| > 30 \text{ MeV}/c^2, \text{ or the } K^+ \text{ has to fulfill } \text{DLL}_{K\pi} > 10$

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

76 All signal candidates used in the branching ratio measurement are reconstructed via the  
 77  $D_s \rightarrow K^+K^-\pi^+$  channel. This decay can either proceed via the narrow  $\phi$  resonance, the  
 78 broader  $K^{*0}$  resonance, or non resonant. Depending on the decay process being resonant  
 79 or not, we apply additional PID requirements:

80 1. resonant case:

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

85 2. non resonant case:  $\text{DLL}_{K\pi} > 5$  for kaons, since the non resonant category has  
 86 significant charmless contributions.

87 Since a measurement of the weak CKM phase  $\gamma$  in the  $B_s^0 \rightarrow D_s K\pi\pi$  channel will  
 88 almost certainly be statistically limited, we plan to add also the  $D_s \rightarrow \pi\pi\pi$  final state  
 89 with  $\mathcal{BR}(D_s \rightarrow \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \rightarrow KK\pi)$  to the  $\gamma$  analysis. Since the determination  
 90 of the branching ratio is expected to have the same level of statistical and systematical  
 91 uncertainties and due to the additional complexity (different selection, efficiencies, simulated  
 92 samples), we choose to only use the most prominent  $D_s \rightarrow KK\pi$  final state for the BR  
 93 determination.

## 94 3.2 Multivariate stage

95 We use TMVA [5] to train a multivariate discriminator, which is used to further improve  
 96 the signal to background ratio. The 17 variables used for the training are:

- 97 •  $\max(\text{ghostProb})$  over all tracks
- 98 •  $\text{cone}(p_T)$  asymmetry of every track, which is defined to be the difference between the  
 99  $p_T$  of the  $\pi/K$  and the sum of all other  $p_T$  in a cone of radius  $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}$   
 100  $< 1$  rad around the signal  $\pi/K$  track.
- 101 •  $\min(\text{IP}\chi^2)$  over the  $X_s$  daughters
- 102 •  $\max(\text{DOCA})$  over all pairs of  $X_s$  daughters
- 103 •  $\min(\text{IP}\chi^2)$  over the  $D_s$  daughters

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

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

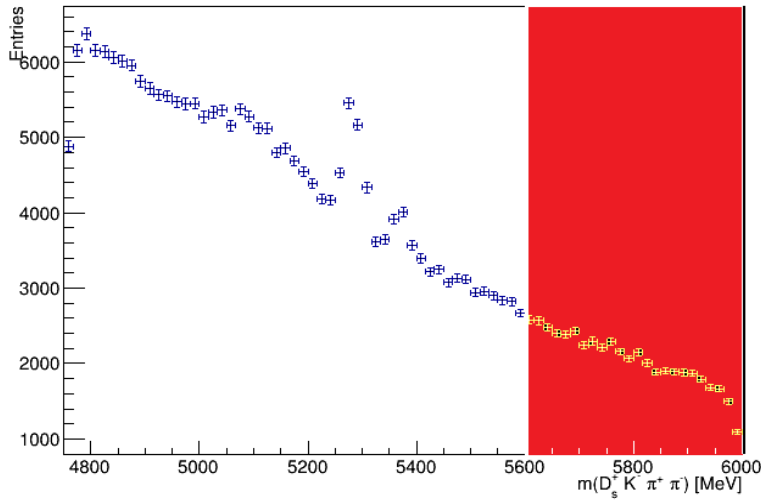


Figure 3.1: Invariant mass distribution of preselected  $B_s^0 \rightarrow D_s K \pi \pi$  candidates. The red coloured region with  $m_{B_s^0 \text{ candidate}} > 5600 \text{ MeV}/c^2$  is used as background input for the boosted decision tree.

The distributions of the input variables for signal and background are shown in Fig. 3.2.

The relative importance of the input variables for the BDTG training is summarized in Table 3.1.

The BDTG output distribution for test and training samples is shown in Fig 3.3. No sign of overtraining is observed.



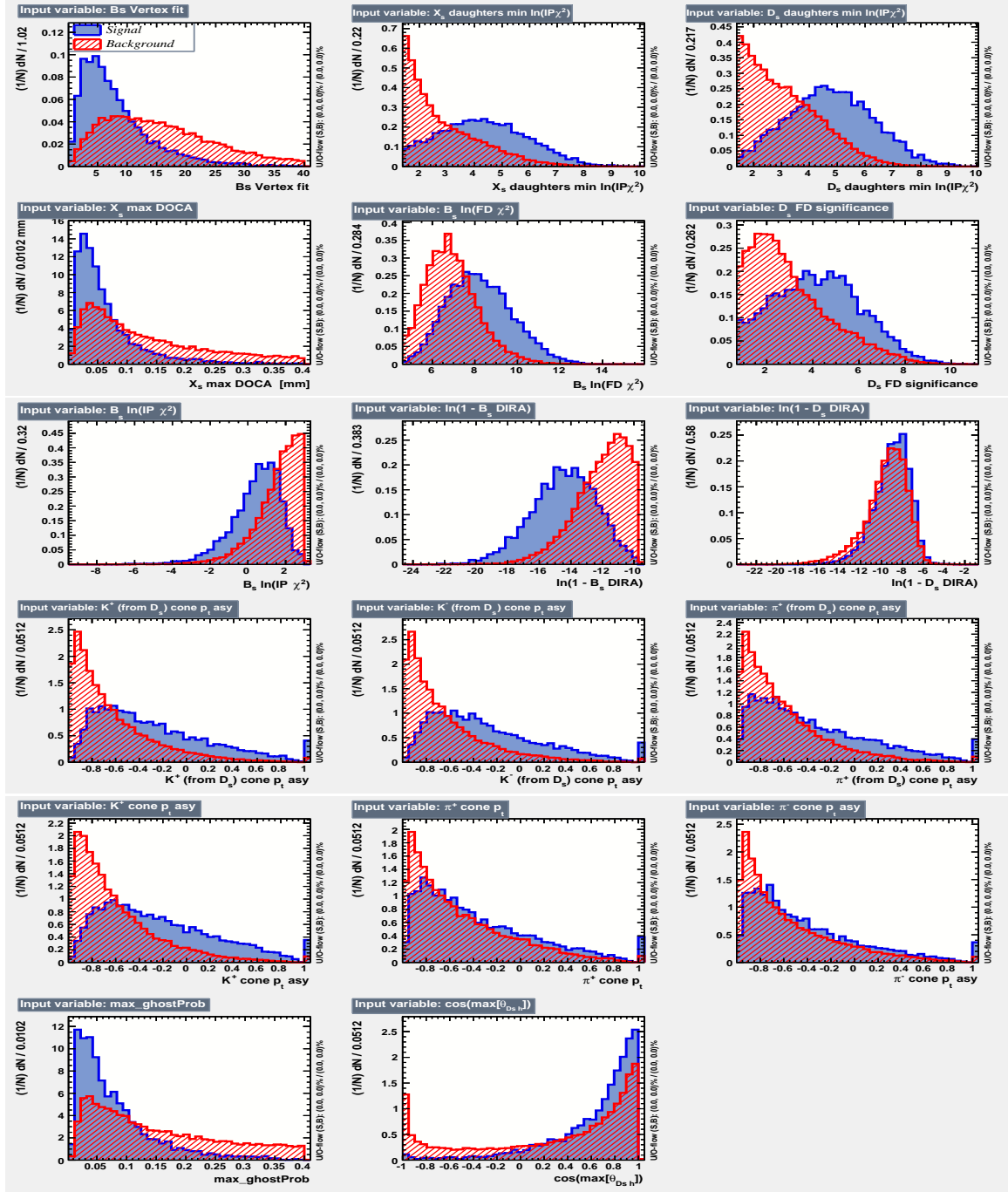


Figure 3.2: Distributions of the input variables used in the BDTG training. The background is shown as red hatched, while the signal is depicted solid blue.

124 We determine the optimal cut value by maximizing the figure of merit  $S/\sqrt{S+B}$   
 125 where  $S$  is the signal yield and  $B$  the background yield in the signal region, defined to be  
 126 within  $\pm 50$  MeV/ $c^2$  of the nominal  $B_s^0$  mass. To avoid a bias in the determination of the

Variable	relative importance [%]
pi_minus_ptasy_1.00	7.32
log_Ds_FDCHI2_ORIVX	7.23
K_plus_ptasy_1.00	7.17
log_Ds_DIRA	6.96
Bs_ENDVERTEX_CHI2	6.82
max_ghostProb	6.76
pi_plus_ptasy_1.00	6.57
log_DsDaughters_min_IPCHI2	6.21
log_Bs_DIRA	6.15
K_plus_fromDs_ptasy_1.00	6.10
log_XsDaughters_min_IPCHI2	5.87
K_minus_fromDs_ptasy_1.00	5.62
cos(Ds h)	5.58
log_Bs_IPCHI2_OWNPV	5.08
log_Bs_FDCHI2_OWNPV	4.04
Xs_max_DOCA	3.98
pi_minus_fromDs_ptasy_1.00	2.59

Table 3.1: Summary of the relative importance of each variable in the training of the BDTG.

branching fraction, we determine S and B using our normalization channel. All trigger, stripping and additional selection criteria described in this and the previous chapter are applied to the  $B_s^0 \rightarrow D_s \pi \pi \pi$  data samples. After that, we perform a simplified version of the fit to the invariant mass distribution of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates described in Sec. 6. Here, a Gaussian function to model the signal and an exponential function to model combinatorial background is used. From this fit we estimate the number of signal events in our normalization channel. Multiplying that number with the PDG branching fraction of  $\frac{\mathcal{B}(B_s^0 \rightarrow D_s K \pi \pi)}{\mathcal{B}(B_s^0 \rightarrow D_s \pi \pi \pi)}$  and the ratio of efficiencies discussed in Sec. 7 allows us to estimate the expected number of  $B_s^0 \rightarrow D_s K \pi \pi$  signal decays. The number of background events can then be computed as

$$N_{bkg} = N_{all} - N_{sig}|_{m_{B_s^0} \pm 50 \text{ MeV}/c^2}. \quad (3.1)$$

The efficiency curves as a function of the cut value are shown in Fig. 3.4. The optimal cut value is found to be  $\text{BDTG} > 0.7012$ . At this working point the signal efficiency is estimated to be 72.47 %, while the background rejection in the signal region is 97.38 %.

## 4 Simulated samples

The simulated (MC) samples are generated using Pythia 8 with Sim08 and Gauss v45r11. Samples are reconstructed with Brunel v43r2p11 and the DSTs were processed using

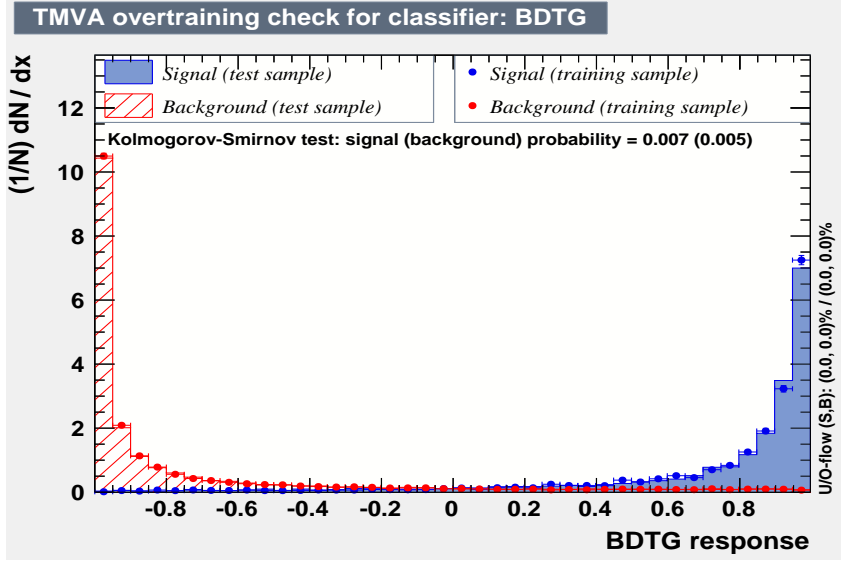


Figure 3.3: BDTG output classifier distribution for (blue) signal and (red) background. The response of an independent test sample (dots) is overlaid.

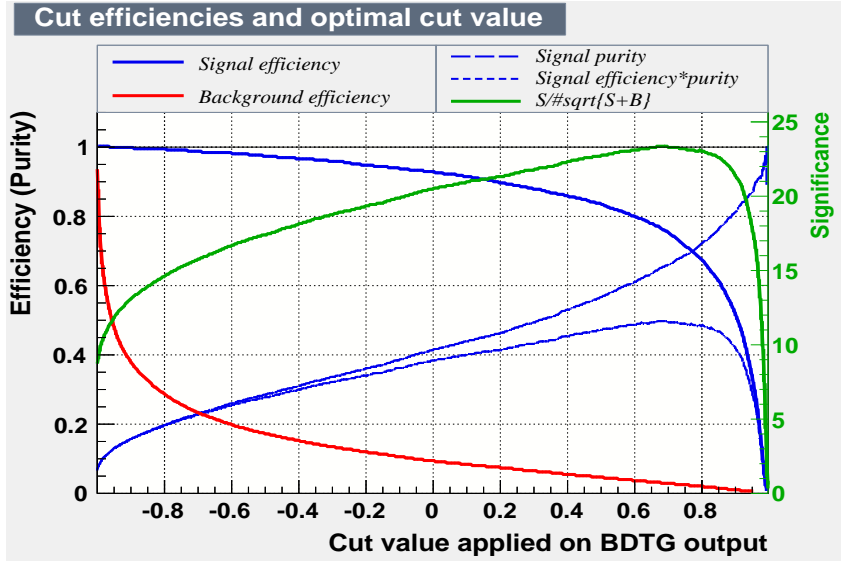


Figure 3.4: Efficiency and purity curves for (blue) signal, (red) background and the (green) FoM curve, as a function of the chosen cut value.

143 DaVinci v36r1p1. The event numbers for the simulated  $B_s^0 \rightarrow D_s K \pi \pi$  and  $B_s^0 \rightarrow D_s \pi \pi \pi$   
 144 samples, as well as the amount of generated and selected events, is shown in Tab. 4.1 and  
 145 4.2 for the different center of mass energies.  
 146 In order to use our MC samples during the BDT training, described in Chapter 3, and the  
 147 calculation of efficiencies (Chapter 7), we have to ensure that the  $B_s^0 \rightarrow D_s K \pi \pi$  decay is

Quantity	$B_s^0 \rightarrow D_s K \pi \pi$	$B_s^0 \rightarrow D_s \pi \pi \pi$
event type	1326607	1326606
events generated	1,131,662	1,188,549
events selected	14,480	15,793

Table 4.1: Generated and selected MC events for signal and normalization channel at  $\sqrt{s} = 7 \text{ TeV}$ .

Quantity	$B_s^0 \rightarrow D_s K \pi \pi$	$B_s^0 \rightarrow D_s \pi \pi \pi$
event type	1326607	1326606
events generated	1,257,244	1,167,428
events selected	14,422	13,423

Table 4.2: Generated and selected MC events for signal and normalization channel at  $\sqrt{s} = 8 \text{ TeV}$ .

modeled correctly by the simulation. To check this, we compare distributions of observables which we use during the multivariate selection stage, as well as some key event observables. The compared distributions need to be generated by signal decays only, therefore we truth-match all particles in the MC samples. Signal distributions of observables in data are obtained using the sWeight technique [6]. Due to the lack of a clear signal peak in the  $B_s^0 \rightarrow D_s K \pi \pi$  data after the stripping, trigger and cut-based preselection, we use preselected  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates to obtain signal sWeights: We perform a fit of a Gaussian signal model and an exponential background to the invariant mass distribution of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates. Using the weights generated from this fit, we weight the distributions of data observables in  $B_s^0 \rightarrow D_s K \pi \pi$  and obtain the corresponding signal distributions.

Figure 4.1 shows the distribution of the number of tracks per event and the distribution of the maximum ghost probability over all tracks in MC and data.

In both cases, the distributions differ significantly. Therefore, we reweight the MC samples using those two variables. The reweighting is done successively, e.g. we perform two 1D weighting processes, where we equalize the respective MC distribution to the observed distribution in data. All distributions of observables used in the BDT training, before and after the reweighting procedure, are shown in the Appendix A.2.

## 5 Models for signal and background components in invariant mass spectrum

The expected signal shape, as well as the expected shape for the combinatorial and physical background has to be known in order to properly model the invariant mass distribution of  $B_s^0 \rightarrow D_s K \pi \pi$  and  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates.

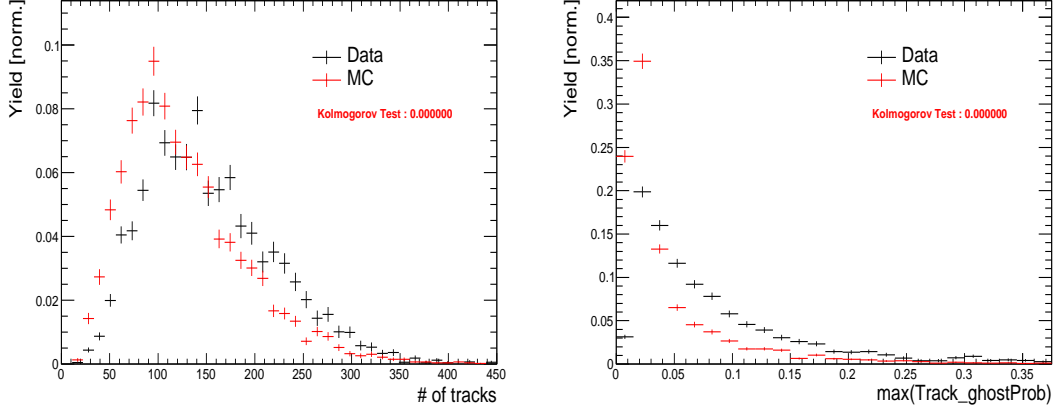


Figure 4.1: Comparison between the distribution of (left) the number of tracks and (right) the maximum ghost probability over all tracks in (black) data and (red) simulation.

## 171 5.1 Signal models for $m(D_s\pi\pi\pi)$ and $m(D_sK\pi\pi)$

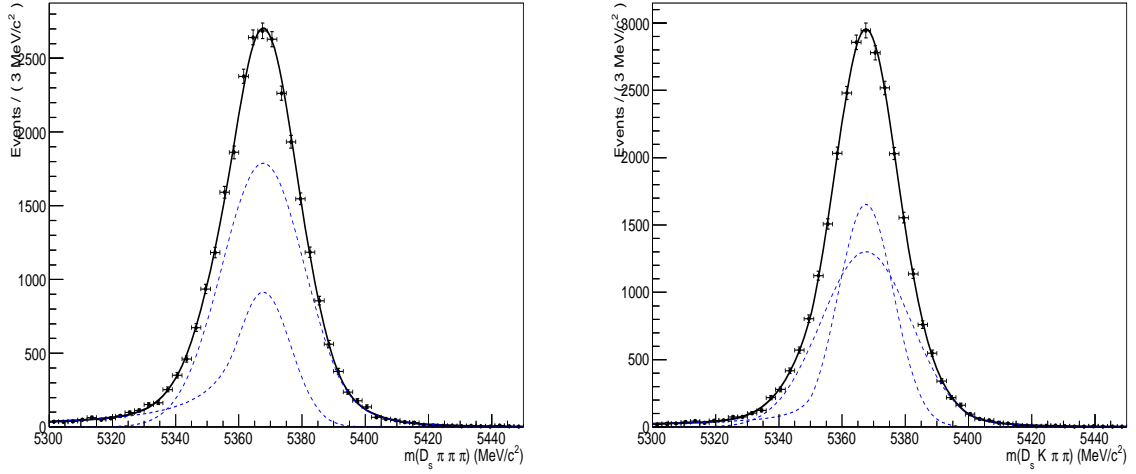


Figure 5.1: Invariant mass distributions of simulated (left)  $B_s^0 \rightarrow D_s\pi\pi\pi$  and (right)  $B_s^0 \rightarrow D_sK\pi\pi$  events. A fit of the sum of two Crystal Ball functions to each distribution is overlaid. The dotted lines represent the individual Crystal Ball functions.

172 The mass distribution of  $B_s^0 \rightarrow D_sK\pi\pi$  signals is modeled using two Crystal Ball  
 173 functions, which share the same mean  $\mu$ , but are allowed to have different widths  $\sigma_1$  and  
 174  $\sigma_2$ . Another double Crystal Ball function is used to account for the contribution of the  
 175  $B^0 \rightarrow D_sK\pi\pi$  decay, which is also present in the  $m(D_sK\pi\pi)$  spectrum. The core width,  
 176 as well as the tail parameters and the ratio of the two individual Crystal Ball functions

are fixed to values obtained by a fit to the invariant mass distribution of simulated events shown in Fig 5.1. The second width  $\sigma_2$  and the shared mean  $\mu$  are floated in the fit to account for possible differences between the simulation and real data. The same approach is used to describe the invariant mass distribution of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates. A double Crystal Ball function is used to model the signal, the parameters are determined by a fit to the invariant mass of simulated  $B_s^0 \rightarrow D_s \pi \pi \pi$  decays, shown in Fig 5.1. The second width and the shared mean are floated to account for differences between data and MC.

## 5.2 Background models for $m(D_s \pi \pi \pi)$

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

The following backgrounds have to be accounted for:

- 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  $B_s^0$ .
- Partially reconstructed  $B_s^0 \rightarrow D_s^* \pi \pi \pi$  decays, with  $D_s^* \rightarrow D_s \gamma$  or  $D_s^* \rightarrow D_s \pi^0$ , where the  $\gamma/\pi^0$  is not reconstructed in the decay chain.

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

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

## 5.3 Background models for $m(D_s K \pi \pi)$

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

- Combinatorial background: same contributions as discussed in Sec. 5.2.
- 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  $\gamma/\pi^0$  is not reconstructed in the decay chain.
- 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  $\gamma/\pi^0$  is not reconstructed in the decay chain.
- Misidentified  $B_s^0 \rightarrow D_s \pi \pi \pi$  decays, where one of the pions is wrongly identified as a kaon  $\pi \rightarrow K$ .

- 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  $\gamma/\pi^0$  from  $D_s^* \rightarrow D_s \gamma/\pi^0$  is not reconstructed.

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

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

For the contribution of the  $B^0 \rightarrow D_s^* K \pi \pi$  background, the same shape is used but the means  $\mu_i$  of the bifurcated gaussians are shifted down by  $m_{B_s^0} - m_{B^0}$  [7]. The yields of both contributions are directly determined in the nominal fit.

To determine the shape of misidentified  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates in the  $m(D_s K \pi \pi)$  spectrum, we take a truth-matched signal MC sample of our normalization channel. We then use the PIDCalib package to determine the  $\pi \rightarrow K$  fake rate. For every candidate in our MC sample, a (momentum)  $p$  and (pseudorapidity)  $\eta$ -dependent event weight is computed and assigned. We flip the particle hypothesis from pion to kaon for the  $\pi$  with the biggest miss-ID weight for each event and recompute the invariant  $B_s^0$  mass. This distribution is then modeled using two Crystal Ball functions. The distribution and the fit are shown in Fig. 5.2(left).

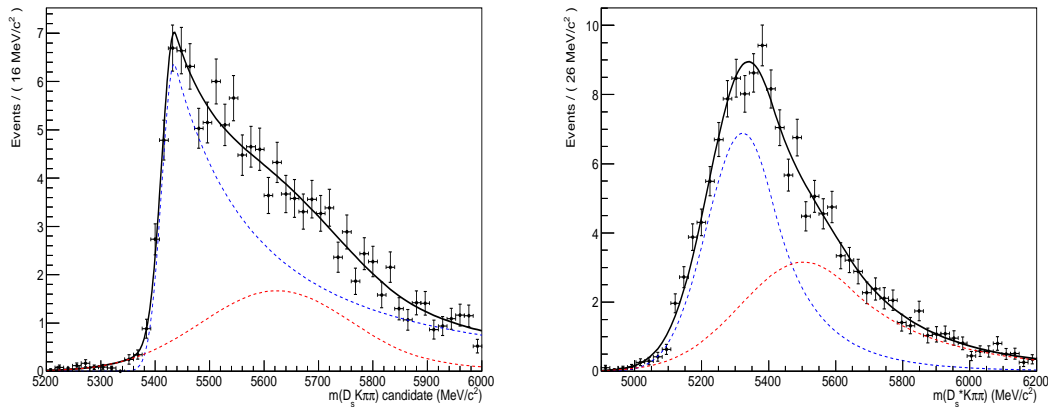


Figure 5.2: Invariant mass distribution of (left) simulated  $B_s^0 \rightarrow D_s \pi \pi \pi$  events, where one of the  $\pi$ 's is reconstructed as a  $K$  and the misID probability for each event is taken into account. The corresponding distribution for simulated  $B_s^0 \rightarrow D_s^* \pi \pi \pi$  events, where the  $\gamma/\pi^0$  from the  $D_s^*$  is excluded from reconstruction, is shown on the right. The solid, black curve on each plot corresponds to the fit consisting of two Crystal Ball functions.

The expected yield of misidentified  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates in the  $m(D_s K \pi \pi)$  spectrum is computed by multiplying the fake probability of  $\propto 3.2\%$ , which is derived from

PIDCalib, by the yield of  $B_s^0 \rightarrow D_s \pi \pi \pi$  signal candidates, determined in the nominal mass fit of our normalization channel.  
 In the same way as mentioned above, we can determine the rate of misidentified, partially 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 and a MC sample of  $B_s^0 \rightarrow D_s^* \pi \pi \pi$  events. The invariant mass distribution we obtain when we exclude the  $\gamma/\pi^0$ , flip the the particle hypothesis  $\pi \rightarrow K$  and apply the event weights given by the fake rate, is shown in Fig. 5.2 (right). The fit of two Crystal Ball functions to this distribution is overlaid. The yield of this contribution is determined from the yield of  $B_s^0 \rightarrow D_s^* \pi \pi \pi$  candidates in the nominal mass fit of our normalization channel, multiplied by the misID probability of  $\propto 3.6\%$ .

## 6 Mass fits for signal and normalization

This section describes the fits to the invariant mass distribution of  $B_s^0 \rightarrow D_s K \pi \pi$  and  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates after all selections are applied. The obtained yields are summarized in Tab. 6.1.

### 6.1 Fit to $B_s^0 \rightarrow D_s \pi \pi \pi$ candidates

An unbinned maximum likelihood fit is performed simultaneously to the invariant mass distribution of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates, for 7 and 8 TeV data. As discussed in Sec. 5.1, the fit is given as the sum of the double Gaussian signal model, the sum of three bifurcated Gaussian functions to model the partially reconstructed  $B_s^0 \rightarrow D_s^* \pi \pi \pi$  background and an Exponential function to account for combinatorial background. The invariant mass distribution and the fit is shown in Fig. 6.1.

### 6.2 Fit to $B_s^0 \rightarrow D_s K \pi \pi$ candidates

The shape of the invariant mass distribution of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates is described by the sum of two double Gaussian functions for the  $B^0$  and  $B_s^0$  signal, two sums of three bifurcated Gaussians for the  $B_s^0/B^0 \rightarrow D_s^* K \pi \pi$  partially reconstructed background contributions and two sums of double Crystal Ball functions for the single misID  $B_s^0 \rightarrow D_s \pi \pi \pi$  and the partially reconstructed, misidentified  $B_s^0 \rightarrow D_s^* \pi \pi \pi$  decays. A simultaneous unbinned maximum likelihood fit is performed and the result is shown in Fig. 6.2.

## 7 Efficiency corrections

Several relative efficiency corrections are needed to measure the branching fraction of the  $B_s^0 \rightarrow D_s K \pi \pi$  decay with respect to  $B_s^0 \rightarrow D_s \pi \pi \pi$ . Precise knowledge of the efficiency related to the detector acceptance, PID requirements, used trigger lines and offline selections are crucial for both, the determination of  $\gamma$  and the branching ratio measurement.



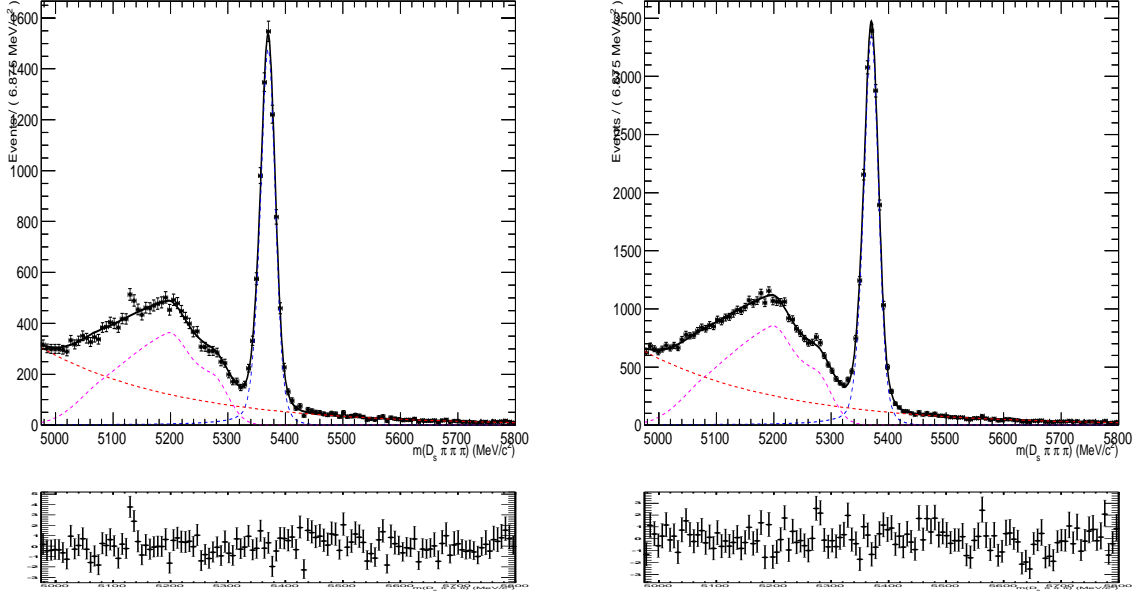


Figure 6.1: Invariant mass distribution of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates for (left) 2011 and (right) 2012 data. The fit described in the text is overlaid. The dashed lines show the (magenta) partially reconstructed and (red) combinatorial component, as well as the (blue) signal component. The pull distributions for the simultaneous fit are shown at the lower left and right.

## 7.1 Relative efficiency for BR measurement

For the branching ratio measurement, the relative efficiency is given by

$$\epsilon_{rel} = \epsilon_{rel}^{acc} \cdot \epsilon_{rel}^{sel} \cdot \epsilon_{rel}^{pid}, \quad (7.1)$$

where  $\epsilon = \frac{\epsilon_{Norm}}{\epsilon_{Sig}}$  is the ratio of the efficiency for the signal and normalization mode. To evaluate these efficiencies, we rely on simulation. The three efficiencies given in Eq. 7.1 are:

- $\epsilon_{rel}^{acc}$ : This is the relative efficiency due to the geometrical acceptance of the LHCb detector. All tracks are required to have a polar angle between 10 and 400 mrad and a minimal momentum of  $|p| > 1.6$  GeV/c in order to be recorded for further analysis. Since the particle species of one track differs between the signal and normalization mode, the efficiencies caused by the geometrical acceptance are expected to be slightly different for the two channels.
- $\epsilon_{rel}^{sel}$ : The relative selection efficiency due to trigger and offline requirements. This efficiency is evaluated using the simulated  $B_s^0 \rightarrow D_s K \pi \pi$  and  $B_s^0 \rightarrow D_s \pi \pi \pi$  samples, comparing the signal yield at generator level to the yield after the trigger and offline requirements are imposed. The respective yields used for the efficiency determination can be found in Tab. 4.1 and 4.2.

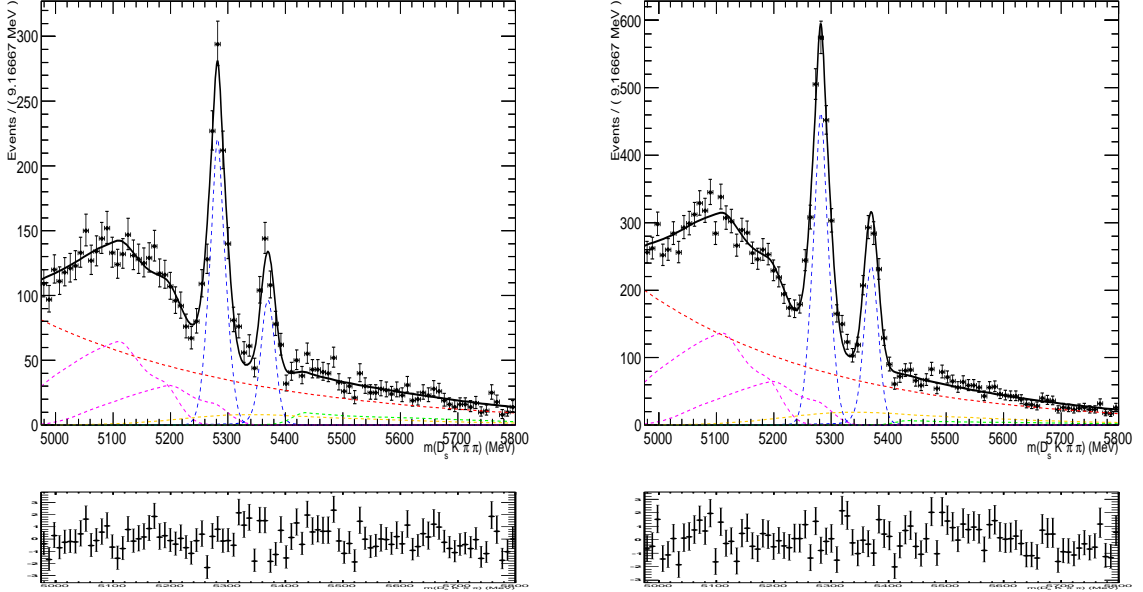


Figure 6.2: Invariant mass distribution of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates for (left) 2011 and (right) 2012 data. A fit described in the text is overlaid. The dashed lines show the (magenta) partially reconstructed and (red) combinatorial background, as well as the (blue) signal component. The dashed green line depicts the misID backgrounds and the dashed yellow line depicts the misidentified, partially reconstructed background component. The pull distributions for the simultaneous fit are shown at the lower left and right.

- $\epsilon_{rel}^{pid}$ : The relative PID efficiency due to the identification likelihood requirements for tracks from both modes. This is evaluated using efficiencies from  $D^{*+} \rightarrow D^0(K^-\pi^+)\pi^+$  calibration data, which is weighted by the expected momentum ( $p$ ) distribution taken from simulation. The PIDCalib tool with a  $p$ - and  $\eta$ -dependent binning is used to perform this evaluation.

Using the definition given in Eq. 7.1, the branching ratio can be expressed as

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s K \pi \pi)}{\mathcal{B}(B_s^0 \rightarrow D_s \pi \pi \pi)} = \frac{\mathcal{Y}(B_s^0 \rightarrow D_s K \pi \pi)}{\mathcal{Y}(B_s^0 \rightarrow D_s \pi \pi \pi)} \cdot \epsilon_{rel}, \quad (7.2)$$

where  $\mathcal{Y}(x)$  represents the yield of the respective channel.

The single efficiencies, as well as the total selection efficiency, for the signal and normalization channel, is given in Tab. 7.1.

## 8 Systematic errors

Several systematic errors contribute to the overall uncertainty on the branching fractions. We consider the most significant ones:

invariant mass spectrum/fit component	yield 2011	yield 2012
<hr/> $m(D_s K \pi \pi)$ <hr/>		
$B_s^0 \rightarrow D_s K \pi \pi$	$351 \pm 26$	$858 \pm 40$
$B^0 \rightarrow D_s K \pi \pi$	$821 \pm 41$	$1721 \pm 67$
$B_s^0 \rightarrow D_s^* K \pi \pi$	$629 \pm 68$	$1333 \pm 129$
$B^0 \rightarrow D_s^* K \pi \pi$	$1252 \pm 188$	$2653 \pm 400$
$B_s^0 \rightarrow D_s \pi \pi \pi$	257 (fixed)	582 (fixed)
$B_s^0 \rightarrow D_s^* \pi \pi \pi$	359 (fixed)	845 (fixed)
combinatorial	$2999 \pm 154$	$6689 \pm 240$
<hr/> $m(D_s \pi \pi \pi)$ <hr/>		
$B_s^0 \rightarrow D_s \pi \pi \pi$	$7671 \pm 96$	$17379 \pm 148$
$B_s^0 \rightarrow D_s^* \pi \pi \pi$	$9984 \pm 193$	$23479 \pm 357$
combinatorial	$10341 \pm 204$	$21737 \pm 373$

Table 6.1: Summary of yields from the fits to 2011 and 2012 data.

Efficiency (%)	$B_s^0 \rightarrow D_s K \pi \pi$	$B_s^0 \rightarrow D_s \pi \pi \pi$
2011 $\epsilon^{acc}$	$11.37 \pm 0.02$	$10.66 \pm 0.02$
2012 $\epsilon^{acc}$	$11.63 \pm 0.02$	$10.90 \pm 0.02$
2011 $\epsilon^{sel}$	$1.07 \pm 0.01$	$1.12 \pm 0.01$
2012 $\epsilon^{sel}$	$0.95 \pm 0.01$	$0.96 \pm 0.01$
2011 $\epsilon^{pid}$	$73.25 \pm 0.88$	$88.50 \pm 0.59$
2012 $\epsilon^{pid}$	$71.96 \pm 0.90$	$88.39 \pm 0.59$
2011 total $\epsilon$	$0.089 \pm 0.002$	$0.106 \pm 0.001$
2012 total $\epsilon$	$0.080 \pm 0.001$	$0.093 \pm 0.001$

Table 7.1: Efficiencies due to the detector acceptance, selection requirements and PID cuts for the signal and normalization mode. All values are obtained using simulated events.

- 296 • Particle identification
- 297 • Signal and background models
- 298 • Determination of the selection efficiency with MC
- 299 • The relative trigger efficiency
- 300 • The relative tracking efficiency
- 301 • MC statistics
- 302 • BDTG efficiency

303 The particle identification (PID) efficiency is determined using PIDcalib in bins of  
304 pseudorapidity  $\eta$  and transverse momentum  $p_T$  of each  $B_s^0$  candidate. To estimate the

systematic uncertainty, the baseline binning scheme was changed to alternative  $\eta$  and  $p_T$  bins. Tab. 8.1 summarizes the tested binning schemes and the observed effect on  $\epsilon^{pid}$ .

	baseline	alternative 1	alternative 2
$p$ bins [MeV/c]	3000-9300-15600-19000 -24400-29800-35200-40600 -46000-51400-56800-62200 -67600-73000-78400-83800 -89200-94600-100000	0-10000-20000-30000-40000 -50000-60000-70000-80000-90000 -100000	1000-8450-15900-23350 -30800-38250-45700-53150 -60600-68050-75500-82950 -90400-97850-105300-112750 -120200-127650-135100-142550 -150000
$\eta$ bins	1.5-2.375-3.25-4.125-5	2-2.3-2.6-2.9-3.2-3.5-3.8-4.1-4.4-4.7-5	1.5-2.2-3.1-4-5
$\Delta\epsilon^{pid}$	-	0.4 %	0.15 %

Table 8.1: Summary of considered binning schemes for the determination of the PID efficiency using PIDCalib.

The maximum change in the PID efficiency due to the binning scheme is observed to be 0.4 %.

The systematic uncertainty arising from the mass fits is introduced by the chosen fit model and the fixed peaking background yields in the signal channel. Those contributions to the overall uncertainty are estimated by varying the nominal fit model and changing the expected background yield within the uncertainties given by the PIDCalib tool. Fixing only one of the peaking background yields (either  $B_s^0 \rightarrow D_s \pi \pi \pi$  or  $B_s^0 \rightarrow D_s^* \pi \pi \pi$ ) and floating the other one during the fit is also considered. The variation in the yield of  $B_s^0 \rightarrow D_s \pi \pi \pi$  candidates is found to be negligible ( $< 1\%$ ), when a linear polynomial instead of an exponential is used to model the combinatorial background. Changing the signal component from a double Gaussian model to a Crystal Ball function has no significant effect on the signal yield either. In the signal channel, only a small change of the  $N_{B_s^0}$  yield is seen when a single Gaussian signal model is used instead of the nominal double gaussian. The most significant effect is observed when the yield of the  $B_s^0 \rightarrow D_s^{(*)} \pi \pi \pi$  misID background is directly determined in the fit. Depending on which component is floated, the signal yield increases or drops by 4%. Since this is the biggest observed effect, we quote it as the uncertainty of the mass fits.

The computed selection efficiency depends on how accurate the momentum spectrum of the final state particles is described by the simulation. To asses a potential systematic uncertainty due to the momentum modeling, we reweight the  $X_d/X_s$  mass spectrum in Monte Carlo to agree with our observed signal data. The spectra are shown in Figure 8.1.

Applying the weights, we observe a 0.9 % variation in the selection efficiency of  $B_s^0 \rightarrow D_s K \pi \pi$  candidates, while no significant change can be found in the efficiency of the  $B_s^0 \rightarrow D_s \pi \pi \pi$  channel.

We use the TISTOS method [8] to cross-check the simulated trigger efficiency on data. The efficiency of the used trigger lines on the signal decay (TOS) is calculated as  $\epsilon_{TOS} = N_{TIS\&TOS}/N_{TIS}$ , where  $N_{TIS}$  is the number of signal events which are triggered independent

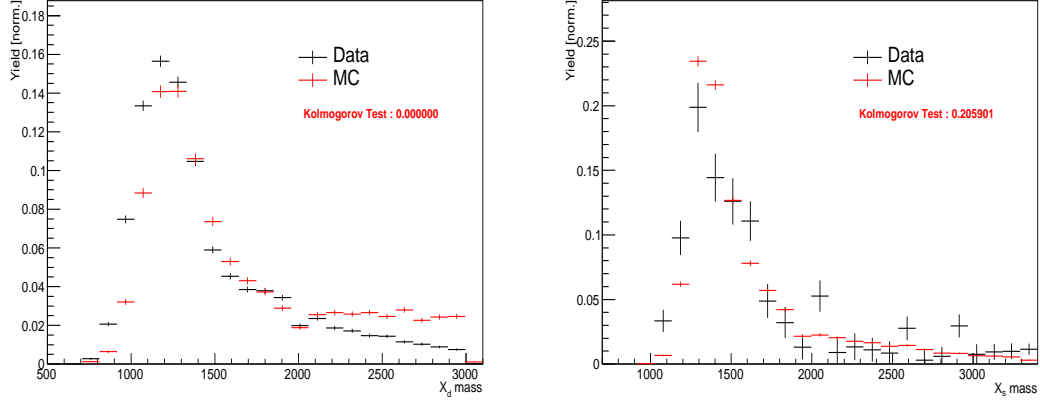


Figure 8.1: Comparison between (left) the  $X_d$  invariant mass spectrum and (right) the  $X_s$  invariant mass spectrum in (black) data and (red) simulation.

334 of the signal decay (TIS) and  $N_{TIS\&TOS}$  denotes the number of events which are both  
 335 triggered by signal decays and independently of signal decays. We perform this cross-check  
 336 using the full sample of  $B_s^0 \rightarrow D_s \pi \pi \pi$  decays and the corresponding MC. Figure 8.2 shows  
 337 the trigger efficiencies for data and MC obtained by the TISTOS method, depending on  
 338 the highest  $p_T$  of the six final state particles, for the L0 and HLT1 trigger stage.

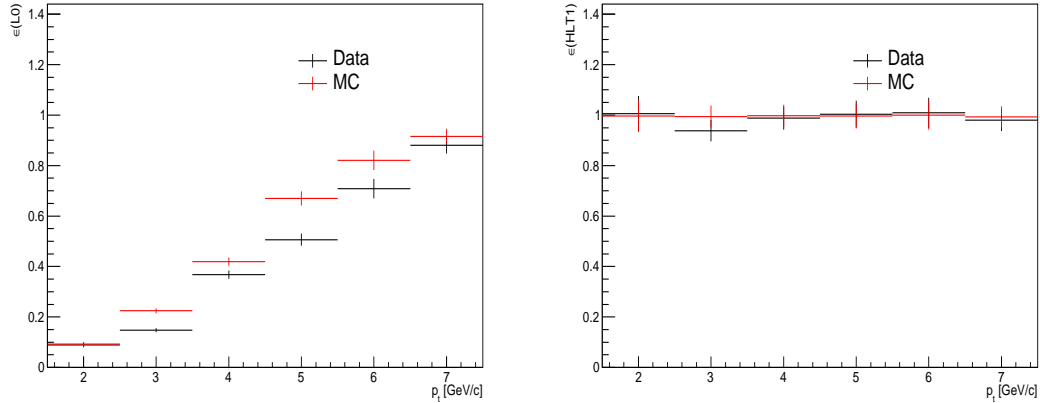


Figure 8.2: Trigger efficiencies of the (lef) L0 and (right) HLT1 trigger stage, determined by the TISTOS method on (black) data and (red) simulation.

339 Although a systematic deviation between the trigger efficiency of the L0 stage in data  
 340 and MC can be observed, this deviation appears in both the signal and normalization  
 341 channel and cancels almost completely in every bin of  $p_T$ . A small systematic uncertainty  
 342 of 0.1% remains. The trigger efficiencies for the signal and normalization channel with  
 343 respect to the HLT2 stage are compared on MC. An agreement, with deviations being

well below the percent level, is observed and thus no systematic uncertainty is assigned to the HLT2 stage.

Due to the different number of hadron types in the final state of the  $B_s^0 \rightarrow D_s K \pi \pi$  and  $B_s^0 \rightarrow D_s \pi \pi \pi$  decay, a systematic uncertainty related to the tracking efficiency for both channels can arise. To assess a potential difference in the tracking efficiency between the signal and normalization channel, we compare the ratio of the efficiency  $\epsilon_{track}(data)/\epsilon_{track}(MC)$  in bins of track momentum  $p$  and pseudorapidity  $\eta$  for both MC samples. Figure 8.3 shows the ratio and its  $p$ ,  $\eta$  dependence for 7 & 8 TeV, using  $J/\psi \rightarrow \mu^+ \mu^-$  decays. This table is used to weight the simulated samples according to the pseudorapidity and total momentum of each track.

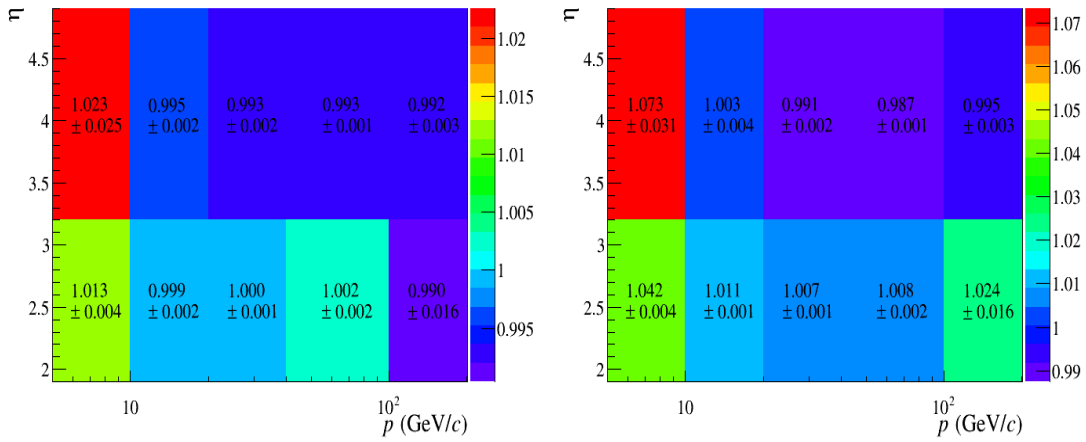


Figure 8.3: Tracking efficiency ratio for data over simulated events for (left)  $\sqrt{s} = 7 \text{ TeV}$  and (right)  $\sqrt{s} = 8 \text{ TeV}$ .

A maximum deviation of  $\Delta\epsilon_{track} = 1.5 \%$  can be observed between the signal and normalization channel. Thus, this value is quoted as the systematical uncertainty due to the tracking efficiency.

The uncertainty on the BDTG efficiency is determined by a fit to the  $B_s^0 \rightarrow D_s \pi \pi \pi$  invariant mass distribution with and without the BDTG cut. The maximum disagreement is found to be 1.9 % and is assigned as the uncertainty on the BDTG efficiency.

The uncertainty due to the limited MC statistic is 1.3 %.

All systematic uncertainties are summarized in Table 8.2. The quadratic sum of all contributions is 5.0 %.

## 9 Results and summary

Using the definition of the branching ratio given in Eq. 7.2, we compute from the measured yields and efficiencies:

Source	Uncertainty on $\frac{\mathcal{B}(B_s^0 \rightarrow D_s K \pi \pi)}{\mathcal{B}(B_s^0 \rightarrow D_s \pi \pi \pi)}$ [%]
PID	0.4 %
Mass fits	4.0 %
MC efficiency determination	0.9 %
Trigger efficiency	0.1 %
Tracking efficiency	1.5 %
BDTG efficiency	1.9 %
MC statistics	1.3 %
Total	5.0 %

Table 8.2: Summary of considered systematic uncertainties on the branching ratio determination.

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s K \pi \pi)}{\mathcal{B}(B_s^0 \rightarrow D_s \pi \pi \pi)} = 0.056 \pm 0.003 \pm 0.003, \quad (9.1)$$

where the uncertainties are statistical and systematical, respectively.  
The results are in good agreement with the first observation and BR measurement of the  $B_s^0 \rightarrow D_s K \pi \pi$  decay, done with  $1 \text{ fb}^{-1}$  of 2011 LHCb data [4]. The number of signal events already exceeds one thousand candidates, although only the  $D_s \rightarrow K K \pi$  final state has been used for this analysis. Adding the  $D_s \rightarrow \pi \pi \pi$  final state, which is expected to contribute roughly 20 % of signal on top, makes this channel a promising prospect for a time-dependent  $\gamma$  determination.

## A Appendix

### A.1 High Level Trigger requirements

The following table summarizes the trigger requirements imposed by the HLT1 line used in this analysis. At least one of the six decay particles must pass the requirements listed in Tab. 1.1 in order for the event to be stored for further analysis.

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

Table 1.1: Summary of the cuts applied by the Hlt1TrackAllL0 trigger. At least one of the six decay particles must pass this requirements, in order for the event to be accepted.

The HLt2 2, 3 and 4-body topological lines use a Boosted Decision Tree based on the b-hadron  $p_T$ , its flight distance  $\chi^2$  from the nearest PV and the sum of the  $B_s^0$  and  $D_s$  vertex  $\chi^2$  divided by the sum of their number of degrees of freedom.

### A.2 Re-weighted MC observables

Figure A.2 shows the distributions of the (left) number of tracks and the (right) maximum ghost probability over all tracks for data, monte carlo and re-weighted monte carlo. These two observables showed significant dissagrement and were therefore chosen for the re-weighting procedure.

The following figures show the comparison of all other observables, which were used during the multivariate selection stage.



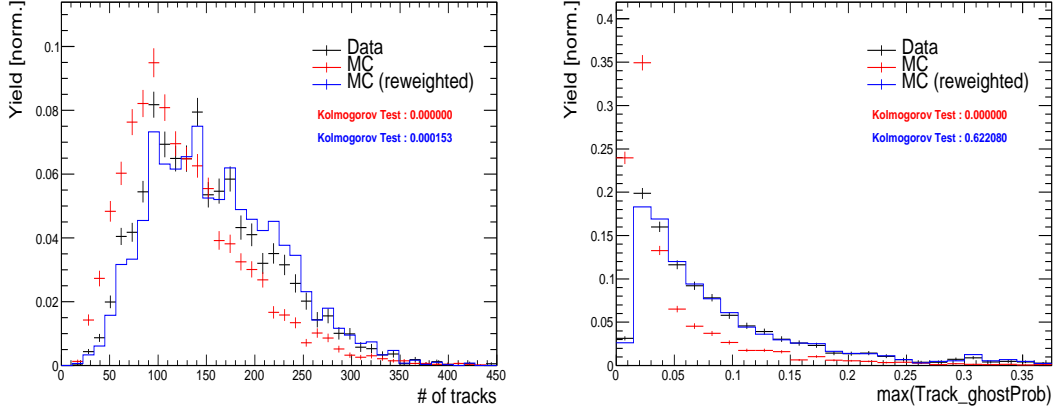


Figure 1.1: Distributions of the (left) number of tracks and the (right) maximum ghost probability over all tracks for data (black), MC (red) and reweighted MC (blue).

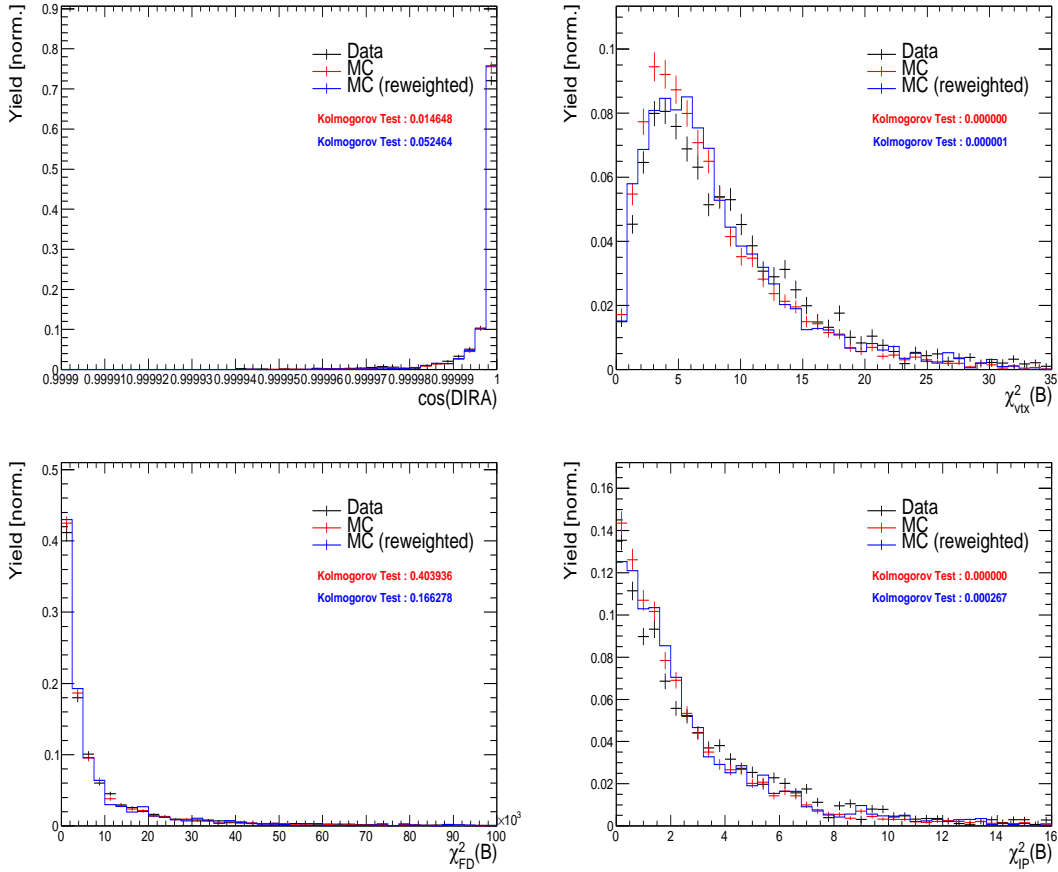


Figure 1.2: Comparison of data and MC observables, before and after reweighting 1.

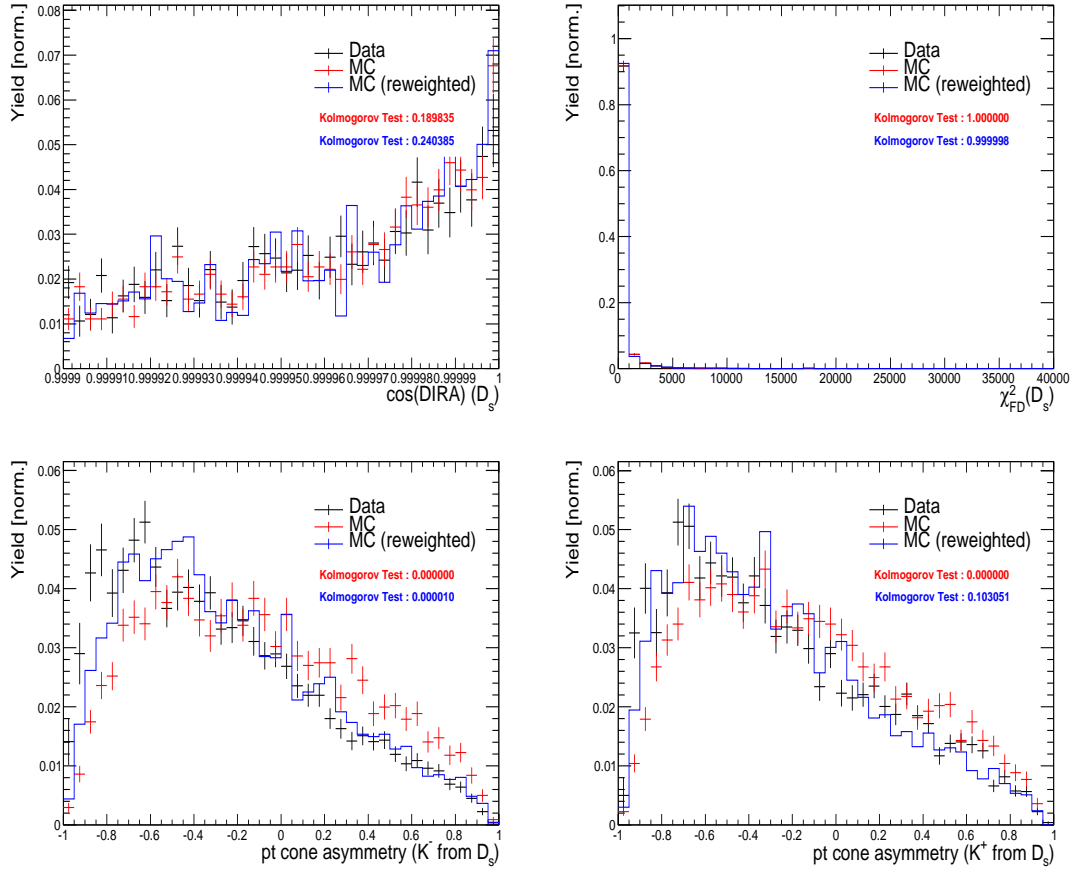


Figure 1.3: Comparison of data and MC observables, before and after reweighting 2.

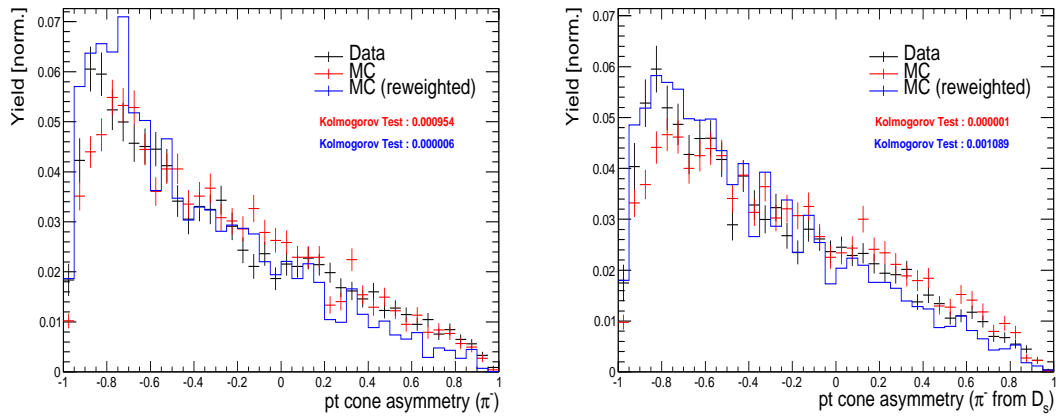


Figure 1.4: Comparison of data and MC observables, before and after reweighting 3.

## References

- [1] R. Fleischer, *New strategies to obtain insights into CP violation through  $B(s) \rightarrow D(s)^+ K^-$ ,  $D(s)^* K^-$ , ... and  $B(d) \rightarrow D^+ \pi^-$ ,  $D^* \pi^-$ , ... decays*, Nucl. Phys. **B671** (2003) 459, [arXiv:hep-ph/0304027](#).
- [2] K. De Bruyn *et al.*, *Exploring  $B_s \rightarrow D_s^{(*)\pm} K^\mp$  Decays in the Presence of a Sizable Width Difference  $\Delta\Gamma_s$* , Nucl. Phys. **B868** (2013) 351, [arXiv:1208.6463](#).
- [3] S. Blusk, *First observations and measurements of the branching fractions for the decays  $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$  and  $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$* , .
- [4] LHCb, S. Blusk, *Measurement of the CP observables in  $\bar{B}_s^0 \rightarrow D_s^+ K^-$  and first observation 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](#).
- [5] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, [arXiv:physics/0703039](#).
- [6] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](#).
- [7] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [8] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039, CERN, Geneva, May, 2014.