LHCb-ANA-20XX-YYY
July 6, 2016

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

P. d'Argent[1], E. Gersabeck[1], M. Kecke[1], M. Schiller[2]

[1] *Physikalisches Institut, Ruprecht-Karls-Universitt Heidelberg, Heidelberg, Germany*
[2] *European Organization for Nuclear Research (CERN), Geneva, Switzerland*

Abstract

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

$$\frac{\mathcal{B}(B_s^0 \to D_s K\pi\pi)}{\mathcal{B}(B_s^0 \to D_s\pi\pi\pi)} = 0.051 \pm 0.002 \pm 0.002$$

The $B_s^0 \to D_s K\pi\pi$ decay can be further used to measure the weak CKM phase γ 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

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 $B_s^0 \to D_s K$ decays [REF HERE]. The measurement of γ presented in this note uses $B_s^0 \to D_s K\pi\pi$ decays, where the $K\pi\pi$ subsystem is dominated by excited kaon states, such as the $K_1(1270)$ and $K_1(1400)$ resonances. It is complementary to the above mentioned analysis of $B_s^0 \to D_s K$, 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 $B_s^0 \to D_s K\pi\pi$ decay presented in [1] and [2], where its branching ratio is measured relative to $B_s^0 \to D_s\pi\pi\pi$. The branching ratio measurement is updated , exploiting the full Run 1 data sample, corresponding to 3 fb^{-1} of integrated luminosity.

2 Data samples

We use the full Run 1 sample from Stripping 21, consisting of 3 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 signal (TOS) or the L0 Global trigger independent of signal (TIS).
Events that pass the L0 stage are further required to pass the HLT1 TrackAllL0 trigger on signal (TOS).
All remaining candidates have to pass either the 2, 3 or 4-body topological trigger (TOS) of the HLT2 stage.
For the presented analysis the B02DKPiPiD2HHHPIDBeauty2CharmLine is used to preselect signal $B_s^0 \to D_s K\pi\pi$ candidates. A summary of the cuts employed by this stripping line can be found in Table 2.1. In this table and throughout the note, we abbreviate $B_s^0 \to D_s X_s(\to K\pi\pi)$ and $B_s^0 \to D_s X_d(\to \pi\pi\pi)$, identifying $X_s \to K\pi\pi$ and $X_d \to \pi\pi\pi$ as the various resonances through which the decays proceed.

3 Simulated samples

The simulated (MC) samples are generated using Pythia 8.
In order to use our MC samples during the BDT training, described in Chapter 4, and the calculation of efficiencies (Chapter 7), we have to make sure that the $B_s^0 \to D_s K\pi\pi$ decay is modelled 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 monte carlo samples. Signal distributions of observables in data are obtained using the sWeight technique [3]: We perform a fit of a

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

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

gaussian signal model and an exponential background to the invariant mass distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates (our normalization channel). Using the weights generated from this fit, we weight the distributions of data observables in $B_s^0 \to D_s K\pi\pi$ and obtain the corresponding signal distributions.
Figure 3.1 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 re-weight the MC samples using those two variables. All distributions of observables used in the BDT training, before and after the re-weighting procedure, are shown in the Appendix A.1.

4 Selection

A twofold approach is used to isolate the $B_s^0 \to D_s K\pi\pi$ candidates from data passing the stripping line. First, further one-dimensional cuts are applied to reduce the level of combinatorial background and to veto some specific physical background. After that, a

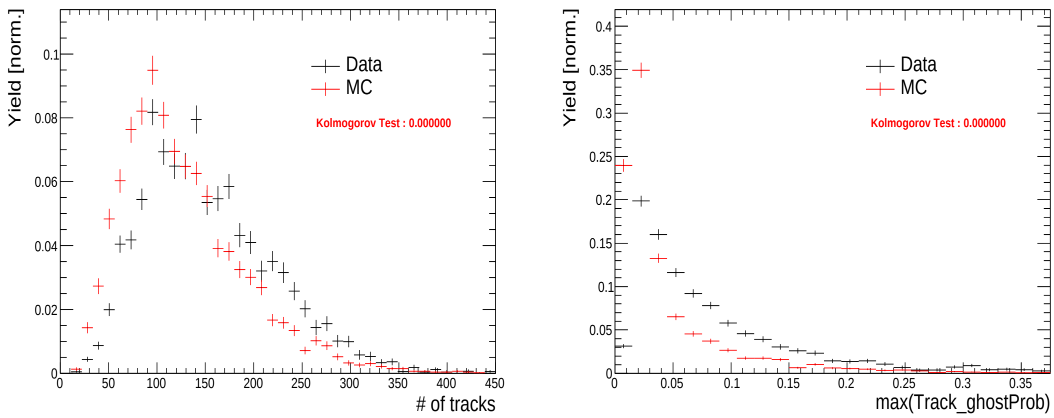

Figure 3.1: Comparison between the distribution of (left) the number of tracks and (right) the maximum ghost probability over all tracks, in (black) data and (red) simulation.

multivariate classifier is trained which combines the information of several input variables, including their correlation, into one powerful discriminator between signal and combinatorial background.

4.1 Cut-based selection

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

(i) DIRA > 0.99994

(ii) min IP $\chi^2 < 20$ to any PV

(iii) FD $\chi^2 > 100$ to any PV

(iv) Vertex $\chi^2/\mathrm{nDoF} < 8$

(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 miss-identification of $K \to \pi$ or $K \to p$:

- $B_s^0 \to D_s^+ D_s^-$: $|M(K\pi\pi) - m_{D_s}| > 20\,\mathrm{MeV}/c^2$
- $B_s^0 \to D_s K K \pi$: π^- $\mathrm{DLL}_{K\pi} < 5$ i

- $B^0 \to D^+(\to K^+\pi^-\pi^+)K\pi\pi$: possible with single miss-ID of $K^+ \to \pi^+$, vetoed by changing mass hypothesis and recompute $|M(K^+\pi^-\pi^+) - m_{Dp}| > 20$ MeV/c^2, or the K^+ has to fulfill $\mathrm{DLL}_{K\pi} > 10$

- $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)K\pi\pi$: possible with single miss-ID of $K^+ \to p$, vetoed by changing mass hypothesis and recompute $M(pK^-\pi^+) - m_{\Lambda_c^+} > 15$ MeV/c^2, or the K^+ has to fulfill $\mathrm{DLL}_{Kp} > 0$

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

1. resonant case:
   (a) $D_s^+ \to \phi\pi^+$, with $|M(K^+K^-) - m_\phi| < 20$ MeV/c^2 : no additional requirements
   (b) $D_s^+ \to \overline{K}^{*0}K^+$, with $|M(K^-\pi^+) - m_{K^{*0}} < 75$ MeV/c^2 : $\mathrm{DLL}_{K\pi} > 0$ for kaons
2. non-resonant case: $\mathrm{DLL}_{K\pi} > 5$ for kaons

4.2 Multivariate stage

We use TMVA [4] to train a multivariate discriminator, which is used to further improve the signal to background ratio. The 17 variables used for the training are:

- max(ghostProb) over all tracks
- cone(p_{T}) asymmetry of every track
- min(IPχ^2) over the X_s daughters
- max(DOCA) over all pairs of X_s daughters
- min(IPχ^2) over the D_s daughters
- D_s DIRA
- D_s FD significance
- max(cos(D_sh_i)), where cos(D_sh_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 IPχ^2, FDχ^2 and Vertex χ^2

Various classifiers were investigated in order to select the most efficient discriminator. As the result a boosted decision tree with gradient boost (BDTG) is chosen as nominal classifier. We use truth-matched Monte Carlo (MC) as signal input. Those 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 ($m_{B_s^0 candidate} > 5600$ MeV/c^2) of our data samples.
The distributions of the input variables for signal and background are shown in Fig. 4.1.

The relative importance of the input variables for the BDTG training is summarized in Table 4.1.

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

The BDTG output distribution for test and training samples is shown in Fig 4.2. No sign of overtraining is observed.

We determine the optimal cut value by maximizing the figure of merit $S/\sqrt{S+B}$ where S is the signal yield and B the background yield in the signal region, defined to be within ± 50 MeV/c^2 of the nominal B_s^0 mass. To avoid a bias in the determination of the branching fraction, we determine S and B using our normalization channel. All trigger, stripping and additional selections described in this and the previous chapters are applied to the $B_s^0 \to D_s \pi\pi\pi$ data samples. After that, we perform a simplified version of the fit to the invariant mass distribution of $B_s^0 \to D_s \pi\pi\pi$ candidates described in Sec. 6. Here, a gaussian signal model and an exponential function to model combinatorial background is used. From this fit we can estimate the number of signal events in our normalization channel. Multiplying that number with the PDG branching fraction of $\frac{\mathcal{B}(B_s^0 \to D_s K\pi\pi)}{\mathcal{B}(B_s^0 \to D_s \pi\pi\pi)}$ and

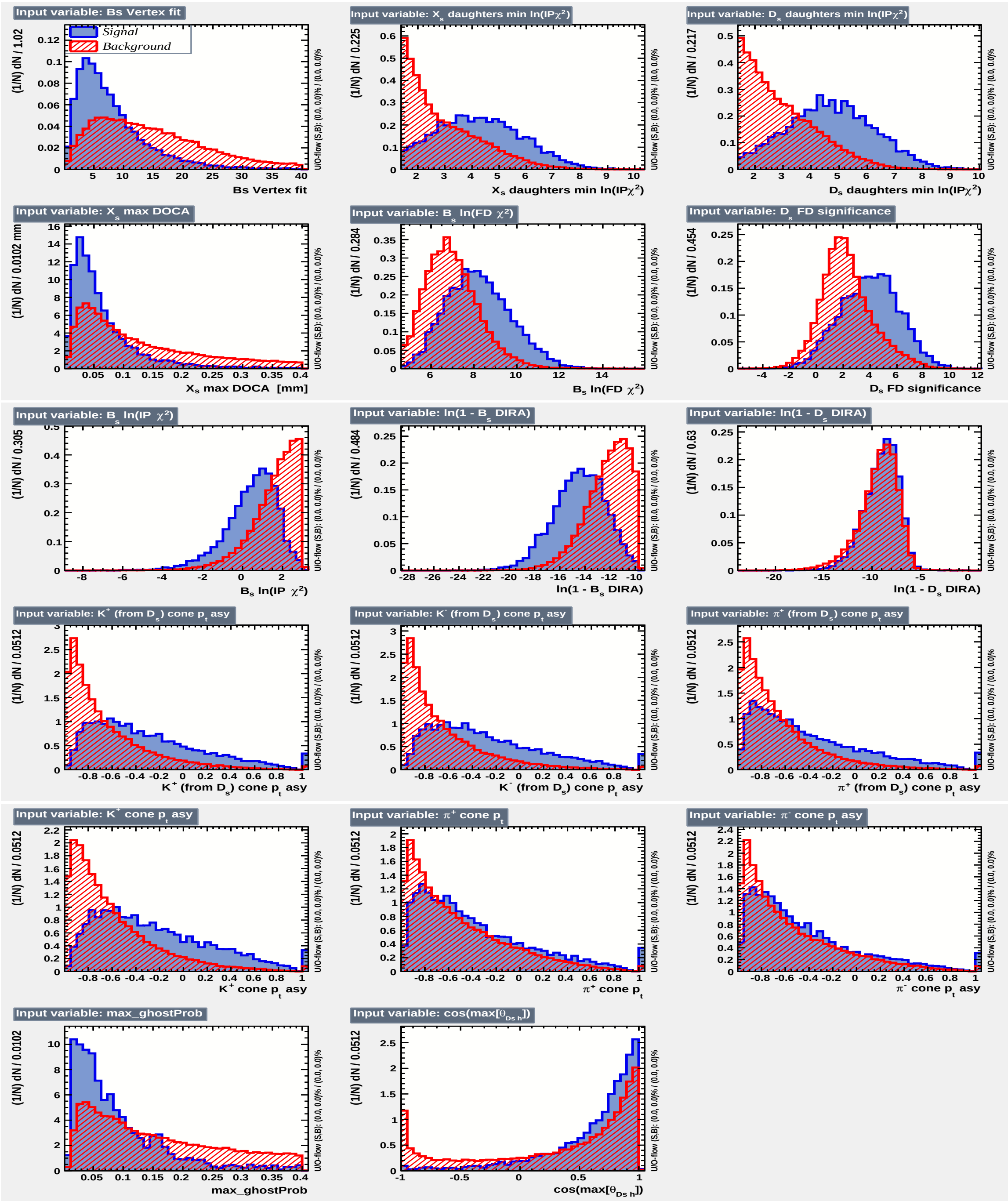

Figure 4.1: Distributions of the input variables used in the BDTG training. The background is shown as red hatched, while the signal is depicted solid blue.

the ratio of efficiencies discussed in Sec. 7 allows us to estimate the expected number of $B_s^0 \to D_s K\pi\pi$ signals. The number of background events can then be computed as

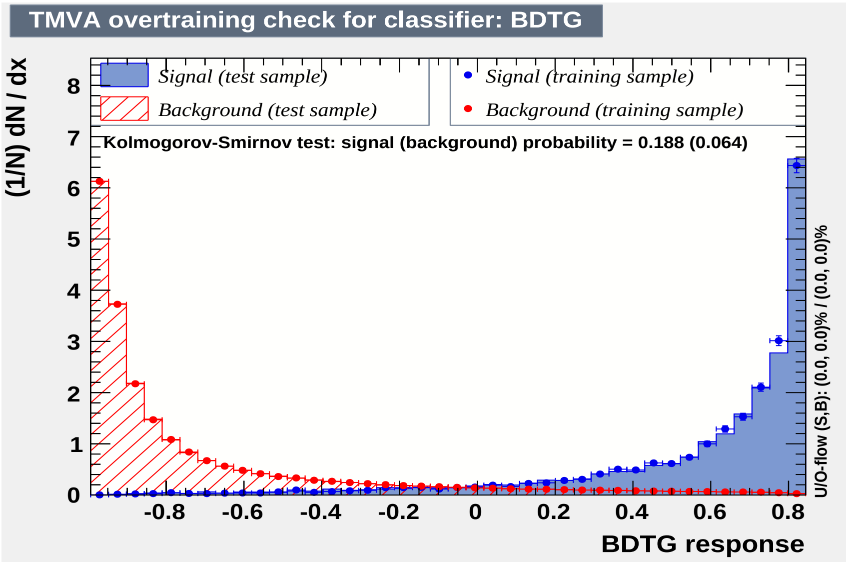

Figure 4.2: BDTG output classifier distribution for (blue) signal and (red) background. The response of an independent test sample is overlaid.

$$N_{bkg} = N_{all} - N_{sig}|_{m_{B_s^0}\pm 50\,\mathrm{MeV}/c^2}\,. \tag{4.1}$$

The efficiency curves as a function of the cut value are shown in Fig. 4.3.

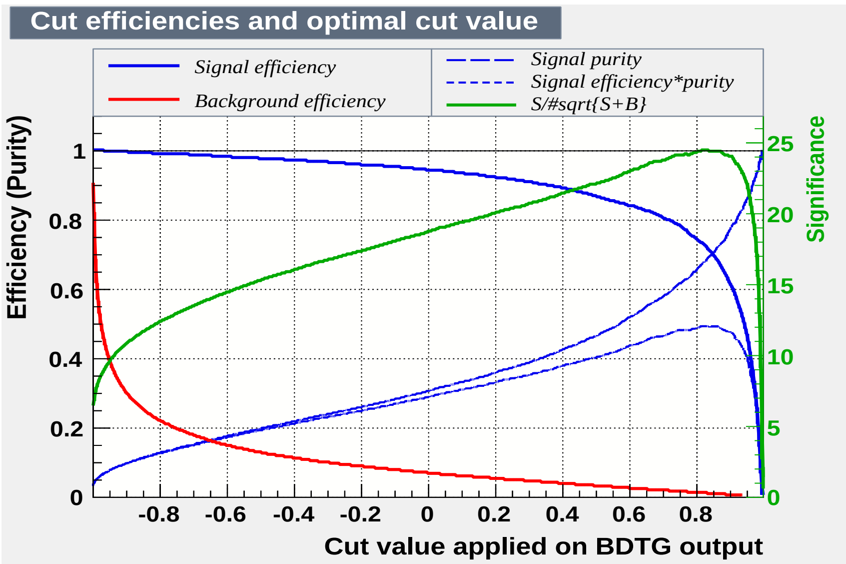

Figure 4.3: Efficiency and purity curves for (blue) signal, (red) background and the (green) FoM curve, as a function of the chosen cut value.

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 backgrounds have to be known in order to properly describe the invariant mass distribution of $B_s^0 \to D_s K\pi\pi$ and $B_s^0 \to D_s\pi\pi\pi$ candidates.

5.1 Signal model

The mass distribution of $B_s^0 \to D_s K\pi\pi$ signal is modeled using two gaussian functions, which share the same mean μ, but are allowed to have different widths σ_1 and σ_2. Another double gaussian is used to account for the contribution of $B^0 \to D_s K\pi\pi$ decays, which are also present in the $m(D_s K\pi\pi)$ spectrum. All parameters of both double gaussians except the core width σ_1 are allowed to float in the nominal fit. The core width is fixed to the value obtained from simulation in order to improve the stability of the fit.
The same approach is used to describe the invariant mass distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates. A double gaussian is used to model this signal shape, all parameters except the core width σ_1 are allowed to float.

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

Different background sources arise in the invariant mass spectrum of candidates for 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 $B_s^0 \to {D_s}^*\pi\pi\pi$ decays, with ${D_s}^* \to D_s\gamma$ or ${D_s}^* \to D_s\pi^0$, where the γ/π^0 is not reconstructed in the decay chain.

In both cases of combinatorial background, the distribution in the invariant mass spectrum 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 $B_s^0 \to {D_s}^*\pi\pi\pi$ 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 π^0 or γ. We rely on simulation to estimate the shape of this contribution.

Figure 5.1 shows the fit of the sum of three bifurcated gaussians to the invariant mass distribution of simulated $B_s^0 \to {D_s}^*\pi\pi\pi$ event. The pion or photon from ${D_s}^* \to D_s(\gamma/\pi^0)$ is excluded from the reconstruction. The obtained shape parameters are used as input values for the nominal $m(D_s\pi\pi\pi)$ mass fit. The yield of this contribution is directly determined in the nominal fit.

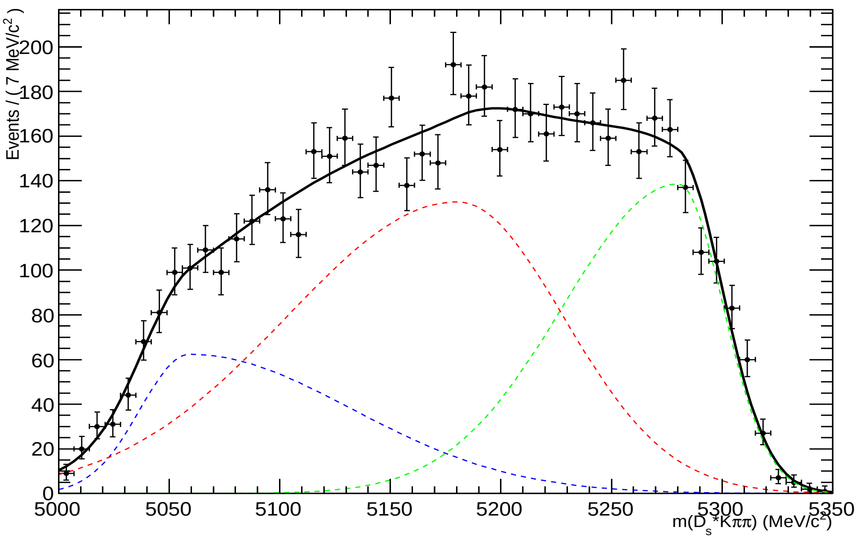

Figure 5.1: Invariant mass distribution of simulated $B_s^0 \to D_s{}^*\pi\pi\pi$ events, where the γ/π^0 is excluded from the reconstruction. A fit of the sum of three bifurcated gaussians to this distribution is overlaid.

5.3 Background models for $m(D_sK\pi\pi)$

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

- combinatorial background: Same contributions as discussed in Sec. 5.2.
- Partially reconstructed $B_s^0 \to D_s{}^*K\pi\pi$ decays, with $D_s{}^* \to D_s\gamma$ or $D_s{}^* \to D_s\pi^0$, where the γ/π^0 is not reconstructed in the decay chain.
- Partially reconstructed $B^0 \to D_s{}^*K\pi\pi$ decays, with $D_s{}^* \to D_s\gamma$ or $D_s{}^* \to D_s\pi^0$, where the γ/π^0 is not reconstructed in the decay chain.
- miss-identified $B_s^0 \to D_s\pi\pi\pi$ decays, where one of the pions is wrongly identified as a kaon $\pi \to K$.
- miss-identified, partially reconstructed $B_s^0 \to D_s{}^*\pi\pi\pi$ decays, where one of the pions is wrongly identified as a kaon $\pi \to K$ and the γ/π^0 from $D_s{}^* \to D_s\gamma/\pi^0$ is not reconstructed.

Again the combinatorial background is expected to be flat in the spectrum of the invariant mass of $B_s^0 \to D_sK\pi\pi$ candidates. An exponential function is used to model this contribution.
The shape of the partially reconstructed $B_s^0/B^0 \to D_s{}^*K\pi\pi$ background is taken from the normalization channel, where it can be directly fitted by the sum of three bifurcated

gaussians as described above. In the signal massfit, all shape parameters for the $B_s^0 \to D_s{}^*K\pi\pi$ background are fixed to the input values from our normalization fit.

For the contribution of the $B^0 \to D_s{}^*K\pi\pi$ background, the same shape is used, but the means μ_i of the bifurcated gaussians are shifted down by $m_{B_s^0} - m_{B^0}$ [5]. The yield of both contributions are directly determined in the nominal fit.

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

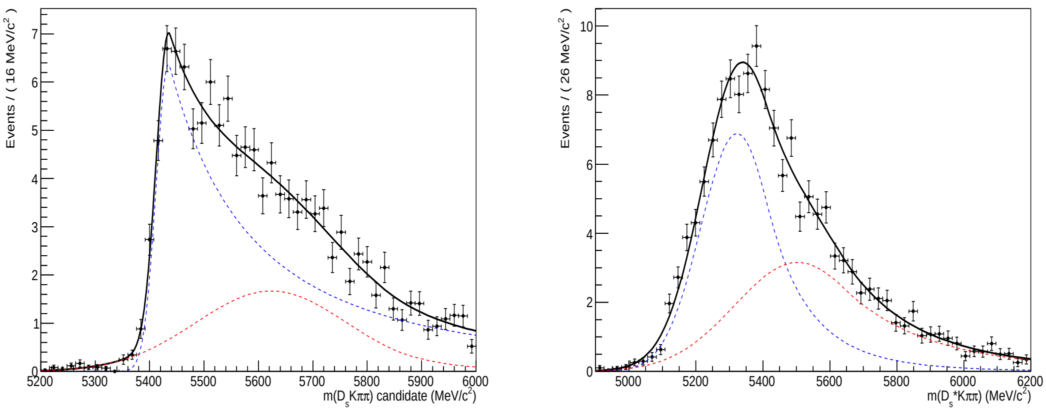

Figure 5.2: Invariant mass distribution of (left) simulated $B_s^0 \to D_s\pi\pi\pi$ events, where one of the π's is reconstructed as a K and the miss-ID probability for each event is taken into account. The corresponding distribution for simulated $B_s^0 \to D_s{}^*\pi\pi\pi$ events, where the γ/π^0 from the $D_s{}^*$ is excluded from reconstruction, is shown on the right. A fit of the sum of two crystal ball functions to each of these distributions is overlaid.

The expected yield of miss-identified $B_s^0 \to D_s\pi\pi\pi$ candidates in the $m(D_sK\pi\pi)$ spectrum is computed by multiplying the fake probability of $\propto 3.2\%$, which is derived from PIDCalib, by the yield of $B_s^0 \to D_s\pi\pi\pi$ 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 miss-identified, partially reconstructed $B_s^0 \to D_s{}^*\pi\pi\pi$ decays in our sample of $B_s^0 \to D_sK\pi\pi$ decays using PIDCalib and a MC sample of $B_s^0 \to D_s{}^*\pi\pi\pi$ events. The invariant mass distribution we obtain when we exlude the γ/π^0, flip the the particle hypothesis $\pi \to K$ 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 $B_s^0 \to D_s{}^*\pi\pi\pi$ candidates in the nominal mass fit of our normalization channel, multiplied by the miss-ID probability of $\propto 3.6\%$.

6 Massfits for signal and normalization channel

This section describes the nominal fits to the invariant mass distribution of $B_s^0 \to D_s K\pi\pi$ and $B_s^0 \to D_s\pi\pi\pi$ candidates after all selection steps, described in the previous Sections, are applied. The obtained yields are summarized in Tab. 6.1.

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

An unbinned maximum likelihood fit is performed simultaneously to the invariant mass distribution of $B_s^0 \to D_s\pi\pi\pi$ 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 gaussians to model the partially reconstructed $B_s^0 \to {D_s}^*\pi\pi\pi$ background, as well as an exponential to account for combinatorial background. The invariant mass distribution and the fit to it is shown in Fig. 6.1.

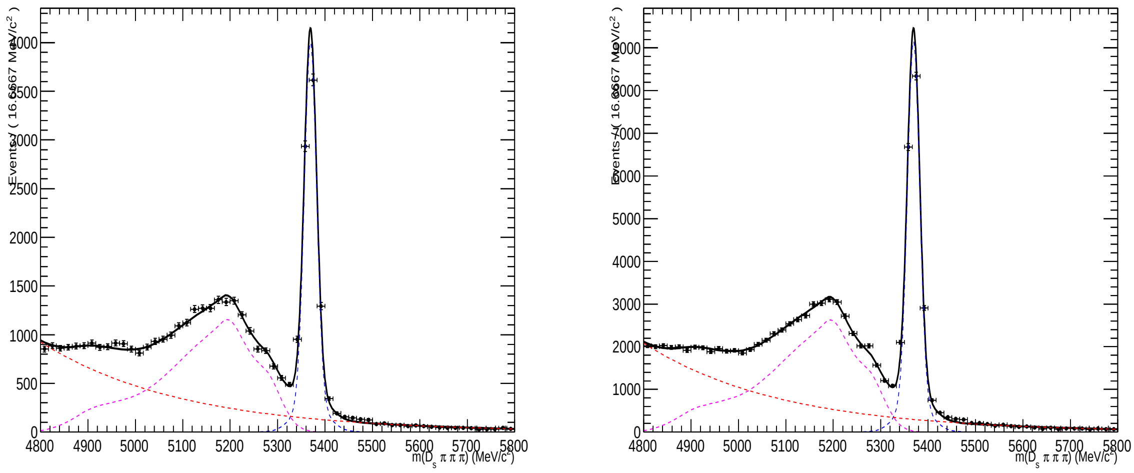

Figure 6.1: Invariant mass distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates for (left) 2011 and (right) 2012 data. A fit described in the text is overlaid. The dashed lines show the (green) partially reconstructed and (red) combinatorial background, as well as the (blue) signal component.

The determined number of $B_s^0 \to D_s\pi\pi\pi$ decays is 8496 ± 102 for 2011 data and 19410 ± 160 for 2012 data. The determined yield for the partially reconstructed $B_s^0 \to {D_s}^*\pi\pi\pi$ background is (2011) 16904 ± 299 and (2012) 38437 ± 589, while the yield for the combinatorial background is (2011) 16066 ± 304 and (2012) 35285 ± 596.

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

Fig. 6.2 shows the invariant mass distribution of $B_s^0 \to D_s K\pi\pi$ candidates. A simultaneous unbinned maximum likelihood fit is overlaid, which consists of two double gaussian models for the B^0 and B_s^0 signal, two sums of three bifurcated gaussians for the $B_s^0/B^0 \to {D_s}^* K\pi\pi$

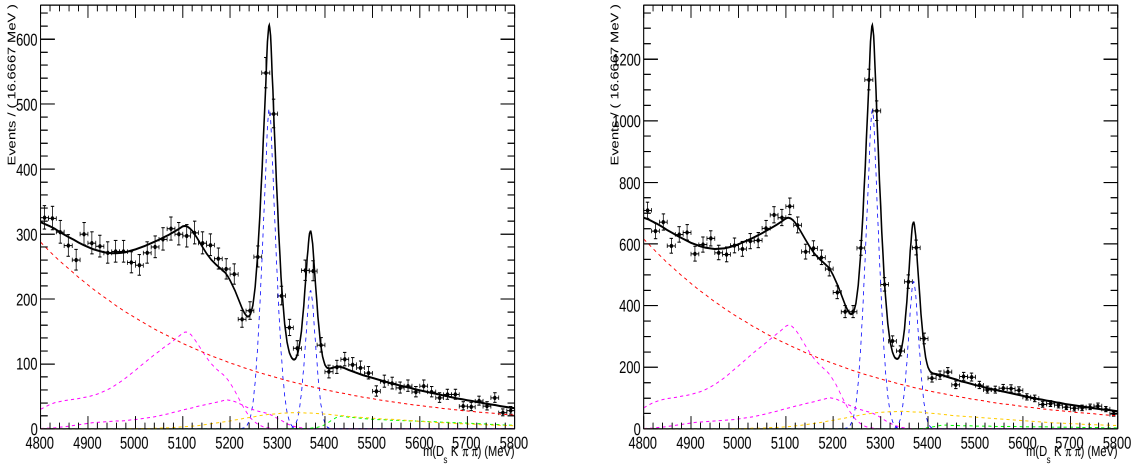

Figure 6.2: Invariant mass distribution of $B_s^0 \to D_s K\pi\pi$ candidates for (left) 2011 and (right) 2012 data. A fit described in the text is overlaid. The dashed lines show the (green) partially reconstructed and (red) combinatorial background, as well as the (blue) signal component. Additional, the dashed magenta line depicts the miss-ID background and the dashed yellow line shows the miss-identified, partially reconstructed background component.

partially reconstructed background contributions and two sums of double crystal ball functions for the single miss-ID $B_s^0 \to D_s\pi\pi\pi$ and the partially reconstructed, miss-identified $B_s^0 \to {D_s}^*\pi\pi\pi$ decays.
The extracted signal yields are (2011) 375 ± 26 and (2012) 866 ± 42.

Decay	yield 2011	yield 2012
$B_s^0 \to D_s K\pi\pi$	375 ± 26	866 ± 42
$B_s^0 \to D_s\pi\pi\pi$	8496 ± 102	19410 ± 160

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

7 Efficiency corrections

Several relative efficiency corrections are needed to measure the branching fraction of $B_s^0 \to D_s K\pi\pi$ with respect to $B_s^0 \to D_s\pi\pi\pi$. 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 γ and the branching ratio measurement.

7.1 Relative efficiency for BR measurement

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

$$\epsilon_{rel} = \epsilon_{rel}^{acc} \cdot \epsilon_{rel}^{sel} \cdot \epsilon_{rel}^{pid}, \tag{7.1}$$

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:

- ϵ_{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 different for the two channels.
- ϵ_{rel}^{sel}: The relative selection efficiency due to trigger and offline requirements.
- ϵ_{rel}^{pid}: The relative PID efficiency due to the identification likelihood requirements for tracks from both modes. This is evaluated using efficiencies from $D^{*+} \to D^0(K^-\pi^+)\pi^+$ calibration data, which is weighted by the expected momentum (p) distribution taken from simulation.

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

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

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 Table 7.1.

Efficiency (%)	$B_s^0 \to D_s K\pi\pi$	$B_s^0 \to D_s\pi\pi\pi$
2011 ϵ^{acc}	11.37 ± 0.02	10.66 ± 0.02
2012 ϵ^{acc}	11.63 ± 0.02	10.90 ± 0.02
2011 ϵ^{sel}	1.18 ± 0.01	1.21 ± 0.01
2012 ϵ^{sel}	1.06 ± 0.01	1.05 ± 0.01
2011 ϵ^{pid}	73.25 ± 0.88	88.50 ± 0.59
2012 ϵ^{pid}	71.96 ± 0.90	88.39 ± 0.59
2011 total ϵ	0.098 ± 0.002	0.114 ± 0.001
2012 total ϵ	0.089 ± 0.001	0.101 ± 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.

8 Systematic errors

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

- Particle identification
- Signal and background models
- Determination of the selection efficiency with MC
- MC statistics
- BDTG efficiency

The particle identification (PID) efficiency is determined using PIDcalib in bins of pseudorapidity η and transverse momentum p_{T} of each B_s^0 candidate. To estimate the systematic uncertainty, the binning scheme was changed to alternative η and p_{T} bins. 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 $B_s^0 \to D_s\pi\pi\pi$ or $B_s^0 \to {D_s}^*\pi\pi\pi$) and floating the other one during the fit is also considered. The variation in the yield of $B_s^0 \to D_s\pi\pi\pi$ candidates is found to be neglectable ($<< 1\%$), 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 $B_s^0 \to {D_s}^{(*)}\pi\pi\pi$ miss-ID 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 to be the systematical 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. Applying the weights, we observe a 0.9 % variation in the selection efficiency of $B_s^0 \to D_sK\pi\pi$ candidates, while no significant change can be found in the efficiency of the $B_s^0 \to D_s\pi\pi\pi$ channel.
The uncertainty on the BDTG efficiency is determined by a fit to the $B_s^0 \to D_s\pi\pi\pi$ invariant mass distribution with and without the BDTG cut. The maximum dissagrement is found to be 1.9 % and is assigned as the systematical uncertainty on the BDTG efficiency.
The uncertainty due to the limited MC statistic is 1.3 %.
All systematic uncertainties are summarized in Table 8.1. The quadratical sum of all contributions is 4.7 %.

Source	Uncertainty on $\frac{\mathcal{B}(B_s^0 \to D_s K\pi\pi)}{\mathcal{B}(B_s^0 \to D_s \pi\pi\pi)}$ [%]
PID	0.4 %
Mass fits	4.0 %
MC efficiency determination	0.9 %
BDTG efficiency	1.9 %
MC statistics	1.3 %
Total	4.7 %

Table 8.1: Summary of considered systematic uncertainties on the branching ratio determination.

9 Results and summary

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

$$\frac{\mathcal{B}(B_s^0 \to D_s K\pi\pi)}{\mathcal{B}(B_s^0 \to D_s \pi\pi\pi)} = 0.051 \pm 0.002 \pm 0.002, \tag{9.1}$$

where the uncertainties are statistical and systematical, respectively.
The results are in good agreement with the first observation and BR measurement of the $B_s^0 \to D_s K\pi\pi$ decay, done with 1 fb^{-1} of 2011 LHCb data [2]. The number of signal events already exceeds one thousand candidates, allthough only the $D_s \to KK\pi$ final state has been used for this analysis. Adding the $D_s \to \pi\pi\pi$ final state, which is expected to contribute roughly 20 % of signal on top, makes this channel a promising prospect for a time-dependent γ determination.

A Appendix

A.1 Re-weghted MC observables

Figure A.1 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 dissagement and were therefore chosen for the re-weighting procedure.

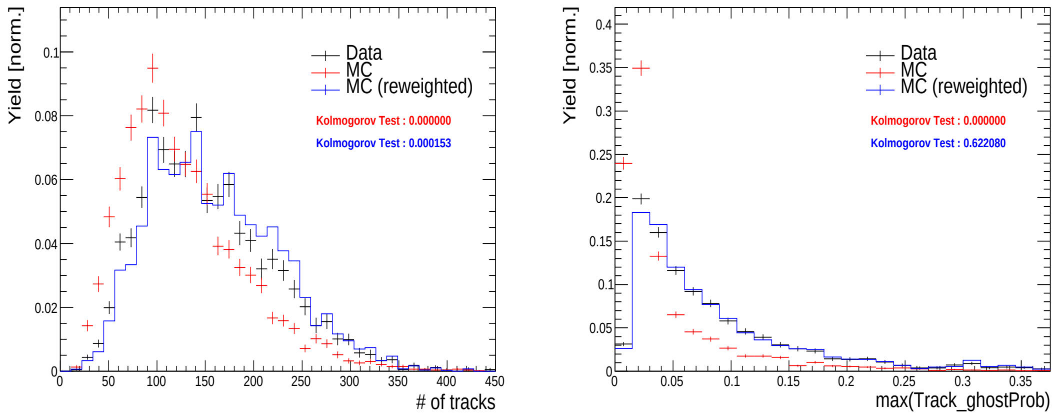

Figure 1.1: distributions of the (left) number of tracks and the (right) maximum ghost probability over all tracks for data (black), monte carlo (red) and re-weighted monte carlo (blue).

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

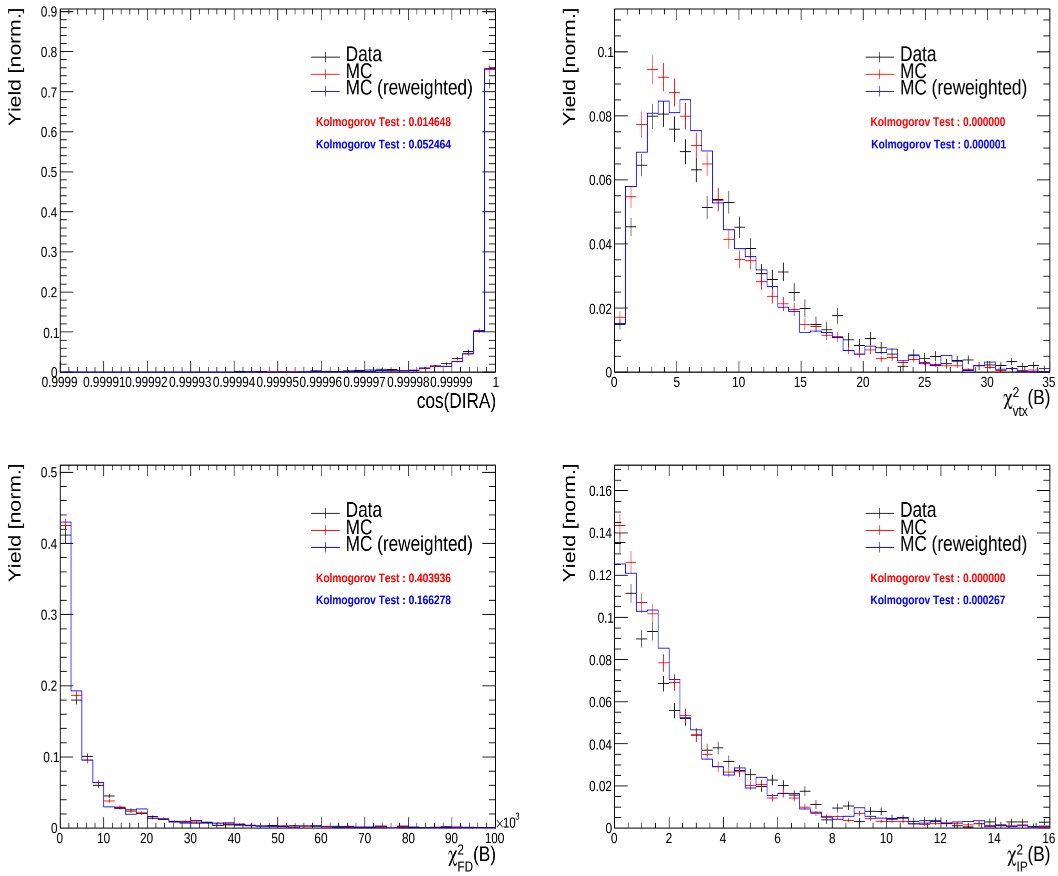

Figure 1.2: Comparison of data and simulated observables, before and after re-weighting 1.

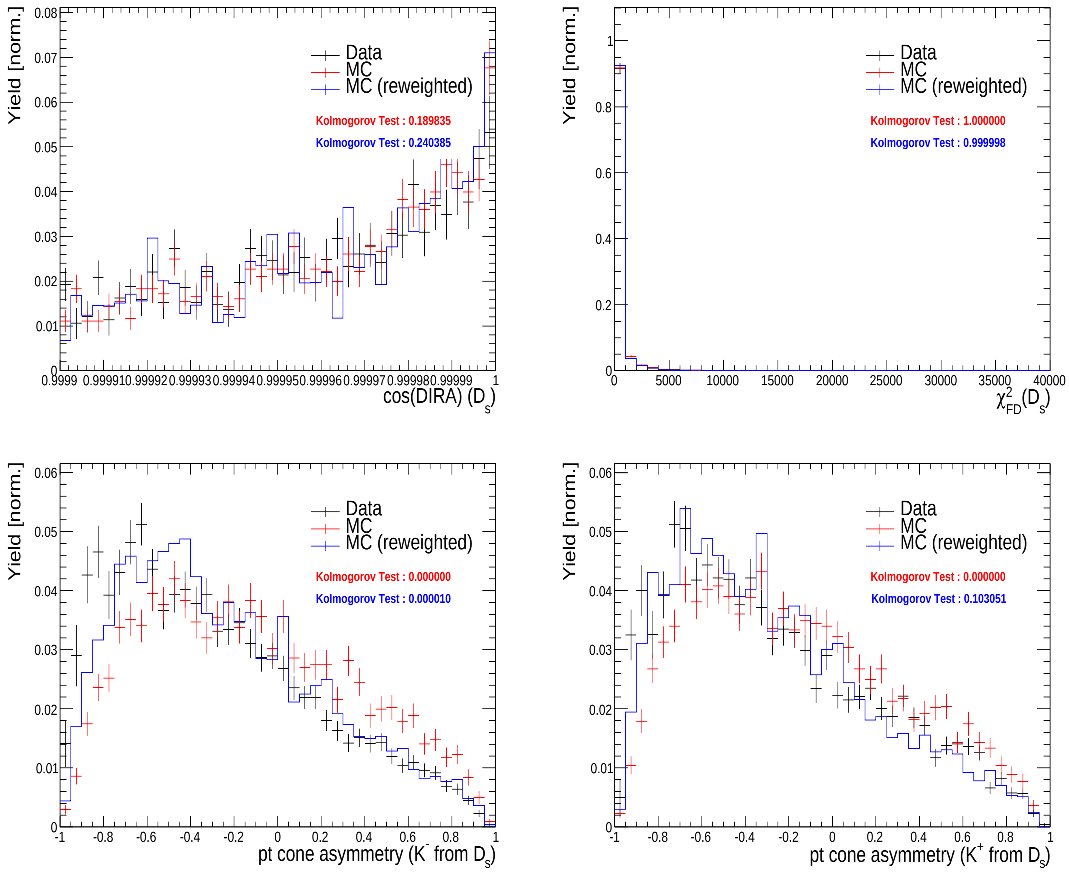

Figure 1.3: Comparison of data and simulated observables, before and after re-weighting 2.

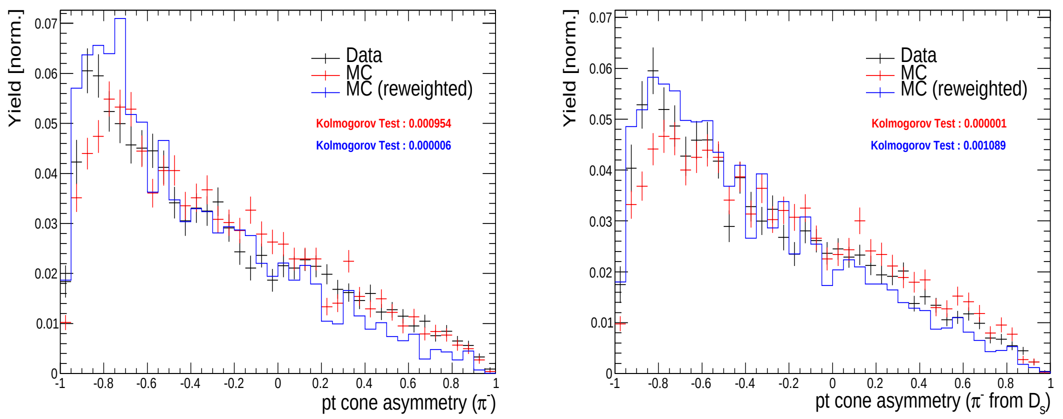

Figure 1.4: Comparison of data and simulated observables, before and after re-weighting 3.

References

[1] S. Blusk, *First observations and measurements of the branching fractions for the decays $\bar{B}^0_s \to D^+_s K^- \pi^+ \pi^-$ and $\bar{B}^0 \to D^+_s K^- \pi^+ \pi^-$*, .

[2] LHCb, S. Blusk, *Measurement of the CP observables in $\bar{B}^0_s \to D^+_s K^-$ and first observation of $\bar{B}^0_{(s)} \to D^+_s K^- \pi^+ \pi^-$ and $\bar{B}^0_s \to D_{s1}(2536)^+ \pi^-$*, 2012. `arXiv:1212.4180`.

[3] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[4] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[5] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.