LHCb-ANA-20XX-YYY
February 20, 2018

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

P. d'Argent[1], E. Gersabeck[2], M. Kecke[1], M. Schiller[3]

[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 $B_s^0 \to D_s K\pi\pi$ 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 Estimation of coherence factor . . . 3
2.3 Results . . . 5

3 Selection **8**
3.1 Cut-based selection . . . 8
3.2 Multivariate stage . . . 9

4 Fits to invariant mass distributions of signal and normalization channel **14**
4.1 Signal models for m($D_s\pi\pi\pi$) and $m(D_sK\pi\pi)$ . . . 14
4.2 Background models for m($D_s\pi\pi\pi$) . . . 15
4.3 Background models for $m(D_sK\pi\pi)$ . . . 15
4.4 Fit to $B_s^0 \to D_s\pi\pi\pi$ candidates . . . 17
4.5 Fit to $B_s^0 \to D_sK\pi\pi$ candidates . . . 17
4.6 Extraction of signal weights . . . 17

5 Flavour Tagging **19**
5.1 OS tagging calibration . . . 20
5.2 SS tagging calibration . . . 21
5.3 Tagging performance comparison between the signal and normalization channel . . . 21
5.4 Combination of OS and SS taggers . . . 23

6 Decay-time Acceptance **28**
6.1 Comparison of acceptance in subsamples . . . 30

7 Decay-time Resoution **32**
7.1 Formalism . . . 32
7.2 Fits to the decay time distributions . . . 32
7.3 Results . . . 33

8 Time dependent fit **36**
8.1 sFit to $B_s^0 \to D_s\pi\pi\pi$ data . . . 36
8.2 sFit to $B_s^0 \to D_sK\pi\pi$ data . . . 36
8.3 Results . . . 36

9 Time-dependent amplitude fit **37**

A Appendix **38**
A.1 Detailed mass fits . . . 38

References **43**

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 [1], [2]. The $B_s^0 \to D_s K\pi\pi$ proceeds at tree level via the transitions shown in Fig. 1.1 a) and b).

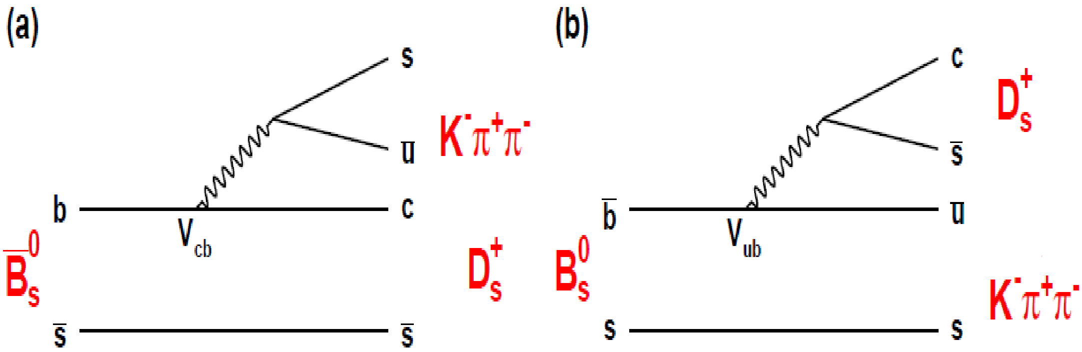


Figure 1.1: Feynman diagram of the $B_s^0 \to D_s K\pi\pi$ decay, proceeding via a) $b \to c$ transitions or b) $b \to u$ transitions.

To measure the weak CKM phase $\gamma \equiv arg[-(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)]$, a decay with interference between $b \to c$ and $b \to u$ 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 $B_s^0 \to D_s K\pi\pi$ decays, where the $K\pi\pi$ subsystem is dominated by excited kaon states such as the $K_1(1270)$ and $K_1(1400)$ resonances, is performed. 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 [3] and [4], where its branching ratio is measured relative to $B_s^0 \to D_s \pi\pi\pi$. The result obtained by the previous analysis is $0.052 \ \pm 0.005\ 0.003$, where the uncertainties are statistical and systematical, respectively. In this note, we present a measurement of γ, making use of the full phase space by using a 6 dimensional time- and amplitude-dependent fit.

2 Sensitivity studies

2.1 PDF

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

$$A(B_s^0 \to D_s^- K^+ \pi\pi) \equiv A(x) = \sum_i a_i\, A_i(x) \tag{2.1}$$

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

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

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

The full time-dependent amplitude pdf is given by:

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

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

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

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

$$\kappa\, e^{i\delta} \equiv \frac{\int A(x)^* \bar{A}(x)\mathrm{d}x}{\sqrt{\int |A(x)|^2 \mathrm{d}x}\sqrt{\int |\bar{A}(x)|^2 \mathrm{d}x}} \tag{2.7}$$

$$r \equiv \frac{\sqrt{\int |\bar{A}(x)|^2 \mathrm{d}x}}{\sqrt{\int |A(x)|^2 \mathrm{d}x}} \tag{2.8}$$

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

2.2 Estimation of coherence factor

To estimate the coherence factor we could generate many toys with random a_i and $\bar{a}_i$ values (see https://twiki.cern.ch/twiki/pub/LHCbPhysics/Bu2DKstar/LHCb-ANA-2017-005_v1.pdf) using the set of amplitudes show in our last talk. However with so many interfering amplitudes, I would be surprised if you couldn't generate every possible value for κ. In any case, this would give us a range where to expect possible values for κ. Worst case would be $0 \leq \kappa \leq 1$.

Assumptions:

- $A(x) = \sum_i a_i \, A_i(x)$
  $\bar{A}(x) = \sum_i \bar{a}_i \, \bar{A}_i(x)$
- Use amplitudes from flavor-averaged, time-integrated fit
- Draw random a_i and $\bar{a}_i$ values
- Constraints:
  $\int (|a_i A_i(x)|^2 + |\bar{a}_i \bar{A}_i(x)|^2) \, \mathrm{d}x / N = F_i^{eff}$
  $r \approx 0.4$ (ration of CKM elements)

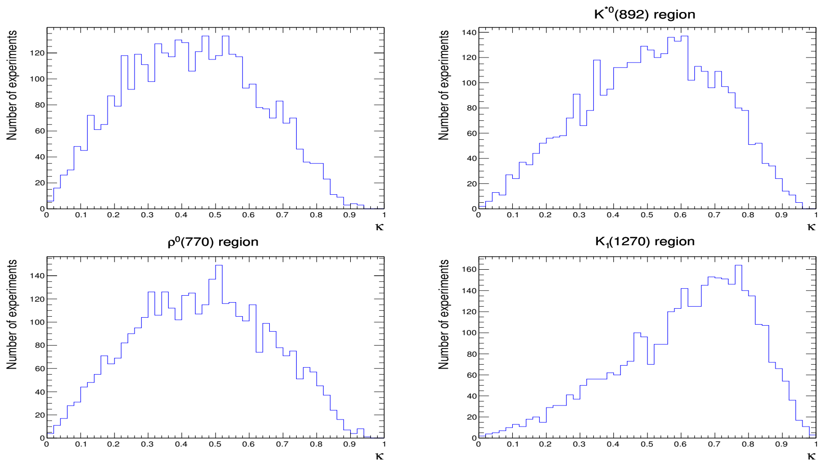


Figure 2.1

Table 2.1

Region	$< \kappa >$ (%)	Cut eff. (%)
Full	43	100
$K^*(892)$	51	43
$\rho^0(770)$	46	47
$K_1(1270)$	61	23

2.3 Results

Assumptions:

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

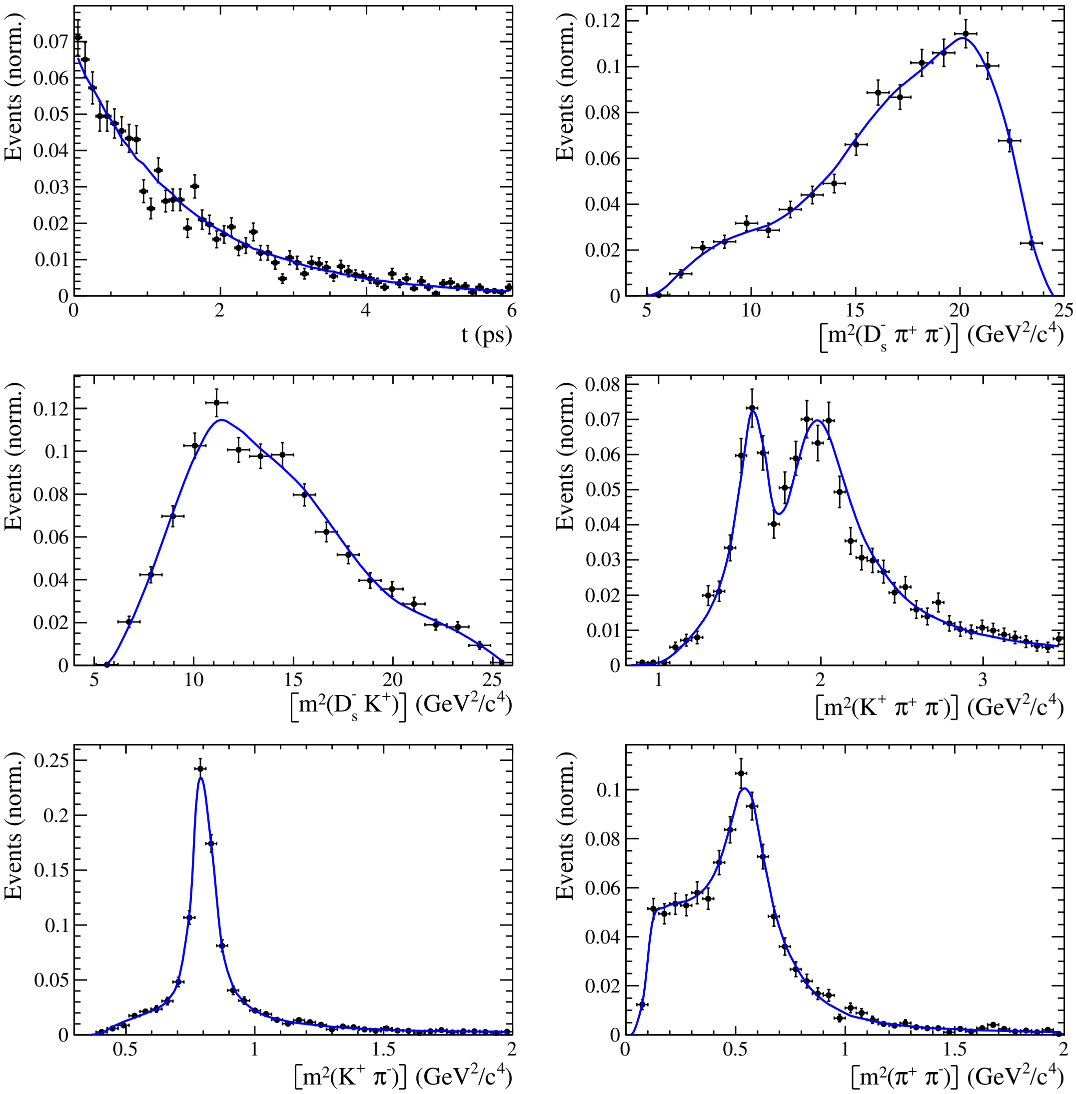
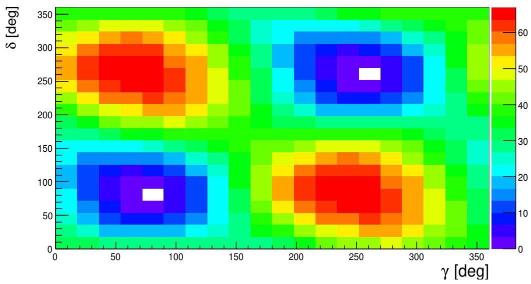


Figure 2.2: Example toy fit



Generated values:

$\gamma = 70°, \delta = 100°$

Fit result:

$\gamma = 74 \pm 15°, \delta = 84 \pm 15°$

$(\gamma = 254 \pm 15°, \delta = 264 \pm 15°)$

Figure 2.3: Likelihood scan

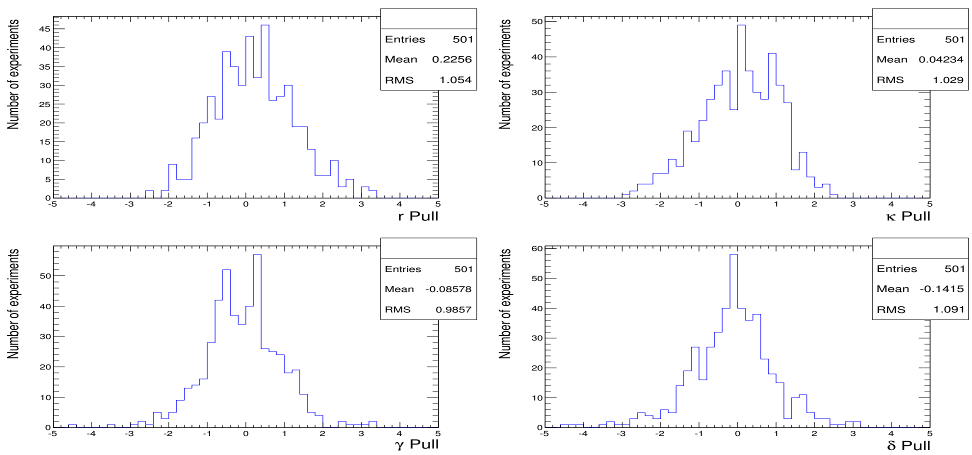


Figure 2.4: Pulls

Table 2.2

	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

3 Selection

For the presented analysis, we reconstruct the $B_s^0 \to D_s K\pi\pi$ decay through two different final states of the D_s meson, $D_s \to KK\pi$ and $D_s \to \pi\pi\pi$. Of those two final states $D_s \to KK\pi$ is the most prominent one, while $\mathcal{BR}(D_s \to \pi\pi\pi) \approx 0.2 \cdot \mathcal{BR}(D_s \to KK\pi)$ holds for the other one.
A two-fold 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. This stage is specific to the respective final state in which the D_s meson is reconstructed, since different physical backgrounds, depending on the respective final state, have to be taken into account. After that, a multivariate classifier is trained which combines the information of several input variables, including their correlation, into one powerful discriminator between signal and combinatorial background. For this stage, all possible D_s final states are treated equally.

3.1 Cut-based selection

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

- DIRA > 0.99994
- min IP $\chi^2 < 20$ to any PV,
- FD $\chi^2 > 100$ to any PV,
- Vertex χ^2/nDoF < 8,
- $(Z_{D_s} - Z_{B_s^0}) > 0$, 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 $K \to \pi$ or $K \to p$. In the following, the vetoes are ordered by the reconstructed D_s final state they apply to:

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

(b) $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)K\pi\pi$: possible with single missID of $p \to K^+$, vetoed by changing particle hypothesis and recompute $M(pK^-\pi^+) - m_{\Lambda_c^+} > 30$ MeV/c^2, or the K^+ has to fulfill (DLL$_{K\pi}$ - DLL$_{p\pi}$) > 5.

(c) $D^0 \to KK$: D^0 combined with a random π can fake a $D_s \to KK\pi$ decay and be a background to our signal, vetoed by requiring $M(KK) < 1840\,\text{MeV}/c^2$.

3. $D_s \to \pi\pi\pi$

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

The most prominent final state used in this analysis is $B_s^0 \to D_s(\to KK\pi)K\pi\pi$, where the D_s 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 on this final state:

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

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

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

3.2 Multivariate stage

We use TMVA [5] 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, which is defined to be the difference between the p_T of the π/K and the sum of all other p_T in a cone of radius $r = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}$ < 1 rad around the signal π/K 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 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 IPχ^2, FDχ^2 and Vertex χ^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 ($m_{B_s^0 candidate} > 5600$ 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.

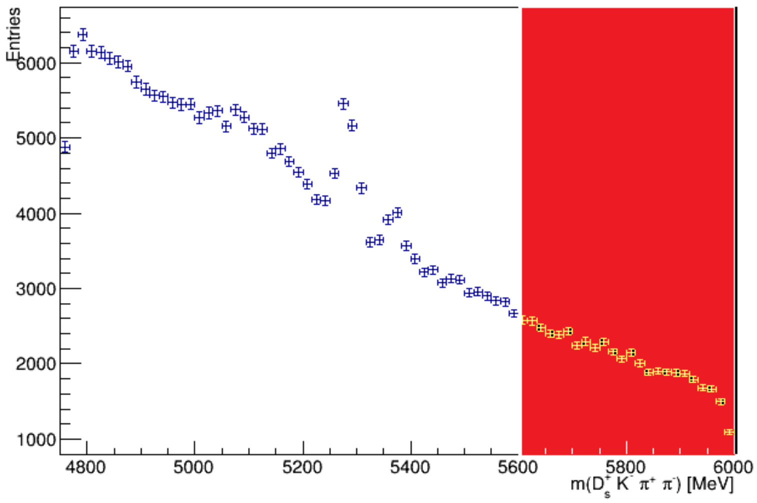


Figure 3.1: Invariant mass distribution of preselected $B_s^0 \to D_s K\pi\pi$ candidates. The red coloured region with $m_{B_s^0 candidate} > 5600$ MeV/c^2 is used as background input for the boosted decision tree.

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.

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,

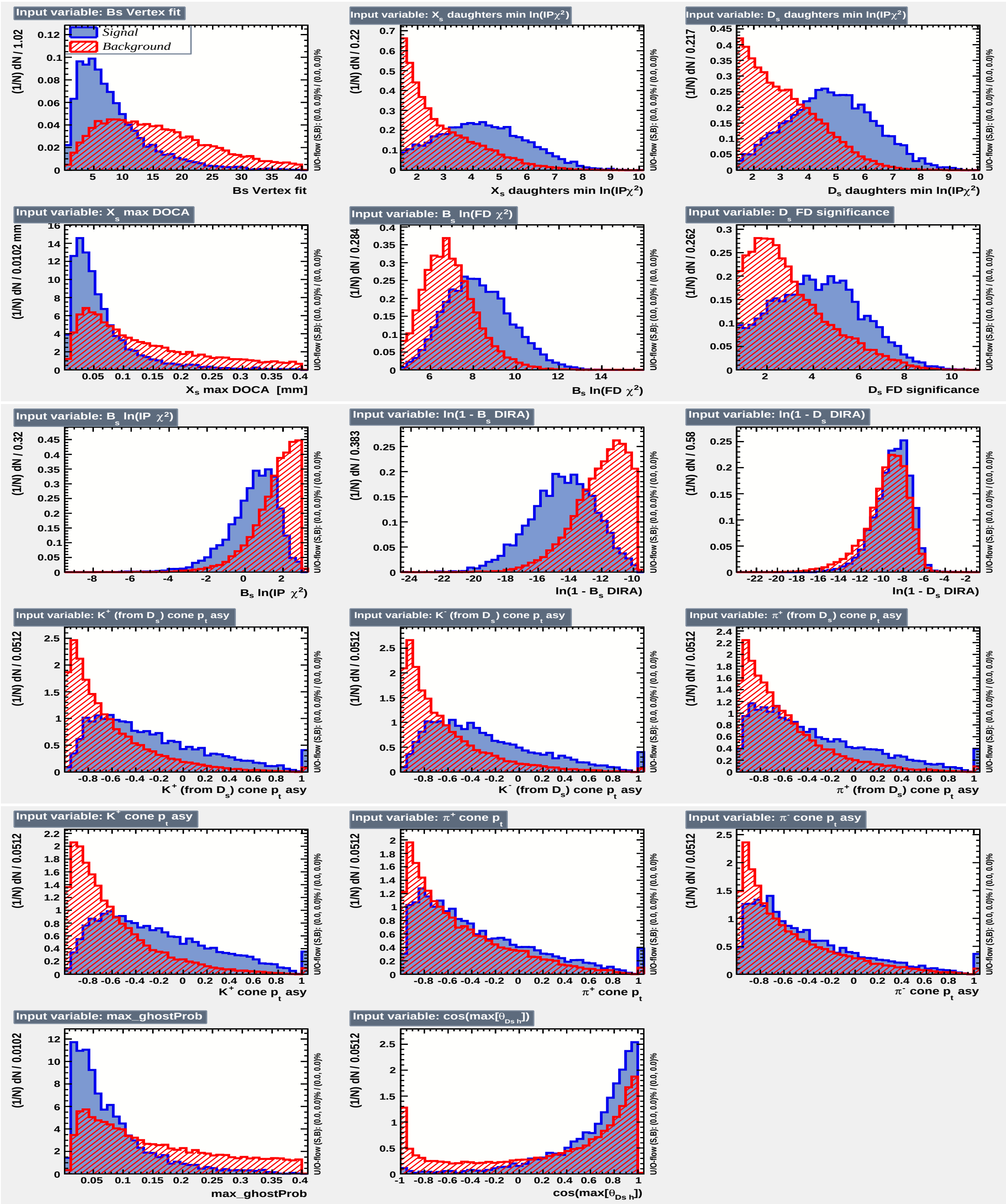


Figure 3.2: Distributions of the input variables used in the BDTG training. The background is shown as red hatched, while the signal is depicted solid blue.

stripping and additional selection criteria described in this and the previous chapter 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. **??**. 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

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.

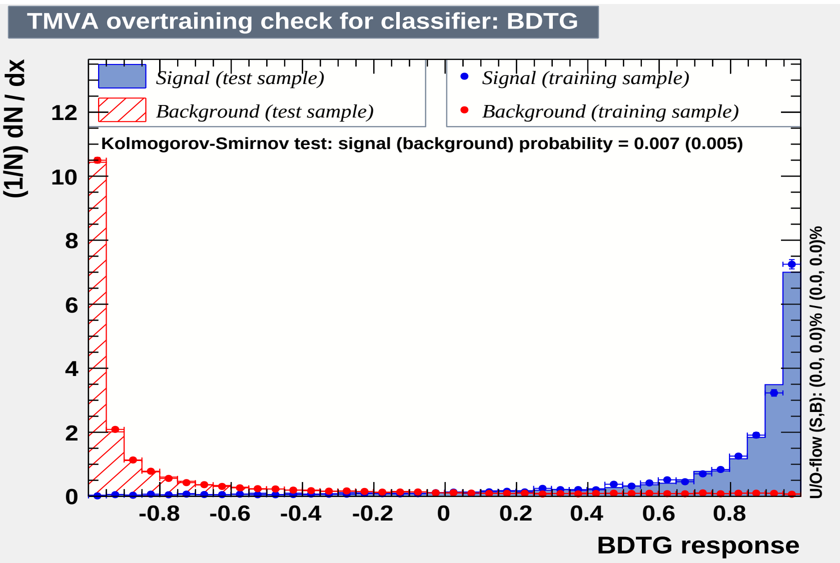


Figure 3.3: BDTG output classifier distribution for (blue) signal and (red) background. The response of an independent test sample (dots) is overlaid.

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

$$N_{bkg} = N_{all} - N_{sig}|_{m_{B_s^0} \pm 50\,\mathrm{MeV}/c^2}. \tag{3.1}$$

The efficiency curves as a function of the cut value are shown in Fig. 3.4. The optimal cut value is found to be 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 %.

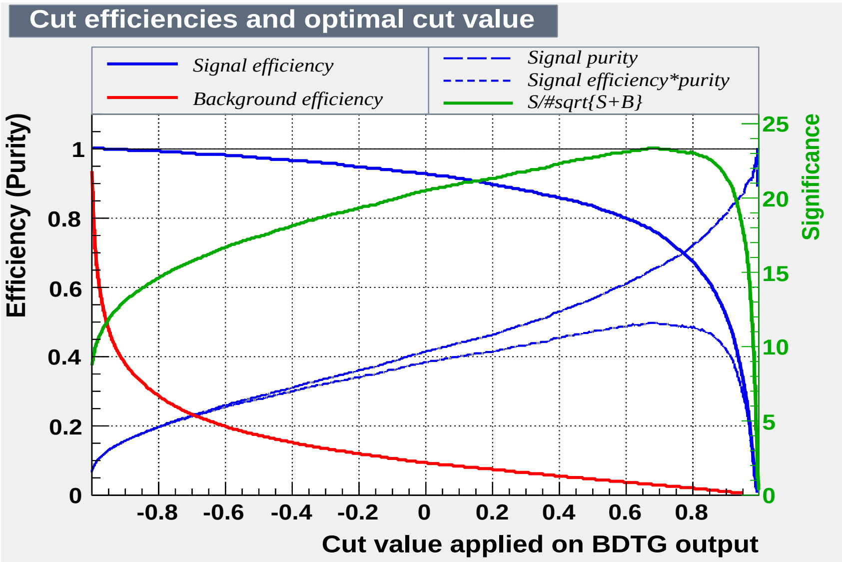


Figure 3.4: Efficiency and purity curves for (blue) signal, (red) background and the (green) FoM curve, as a function of the chosen cut value.

4 Fits to invariant mass distributions of signal and normalization channel

In order to properly model 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, the expected signal shape, as well as the expected shape for the combinatorial and physical background has to be known. This model can then be used to fit the distributions and obtain signal sWeights [6], which are employed to suppress the residual background that is still left in the sample, for the time-dependent amplitude fit.

4.1 Signal models for $\mathrm{m}(D_s\pi\pi\pi)$ and $m(D_s K\pi\pi)$

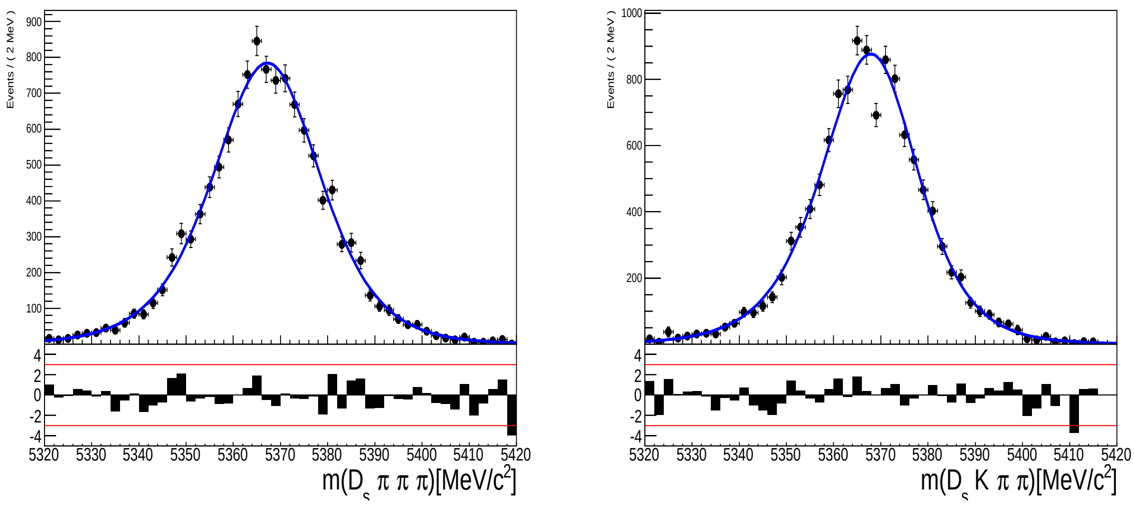


Figure 4.1: Invariant mass distributions of simulated (left) $B_s^0 \to D_s\pi\pi\pi$ and (right) $B_s^0 \to D_s K\pi\pi$ events. A fit of a RooJohnsonSU function to each distribution is overlaid.

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

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

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

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

4.2 Background models for $\mathrm{m}(D_s\pi\pi\pi)$

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 $B^0/B_s^0 \to {D_s}^*\pi\pi\pi$ 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 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. The pion or photon from ${D_s}^* \to D_s(\gamma/\pi^0)$ 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.

4.3 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. 4.2.
- 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.
- Misidentified $B_s^0 \to D_s\pi\pi\pi$ decays, where one of the pions is wrongly identified as a kaon $\pi \to K$.
- Misidentified, 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.

The combinatorial background is expected to be non-peaking 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 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 $B_s^0 \to D_s{}^* K\pi\pi$ 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 $m_{B_s^0} - m_{B^0}$ [?]. The yields of both contributions are directly determined in the nominal fit.

To determine the shape of misidentified $B_s^0 \to D_s\pi\pi\pi$ 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 (momentum) p and (pseudorapidity) 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 modeled using two Crystal Ball functions. The distribution and the fit are shown in Fig. 4.2(left).

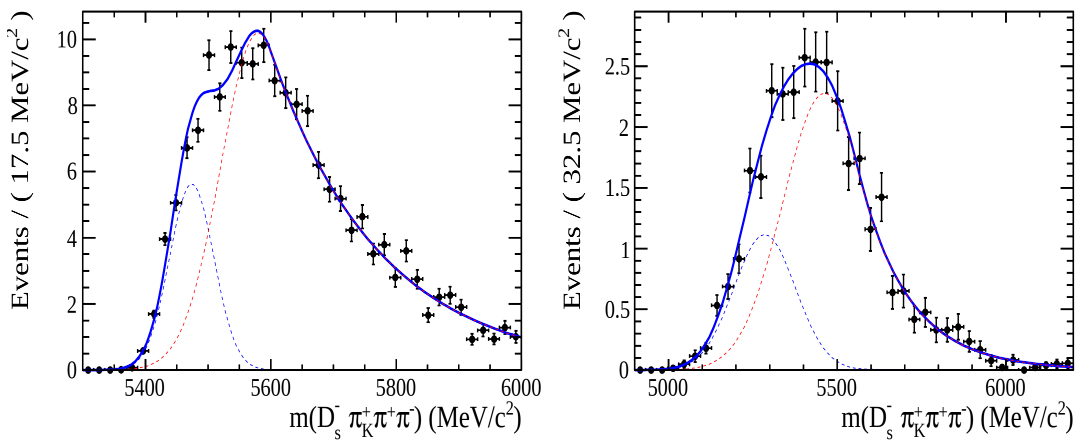


Figure 4.2: Invariant mass distribution of (left) simulated $B_s^0 \to D_s\pi\pi\pi$ events, where one of the π's is reconstructed as a K and the misID probability for each event is taken into account. The corresponding distribution for simulated $B_s^0 \to D_s{}^*\pi\pi\pi$ events, where the γ/π^0 from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

The expected yield of misidentified $B_s^0 \to D_s\pi\pi\pi$ 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 misidentified, 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 exclude 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. 4.2 (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 misID probability of $\propto 3.6\%$.

4.4 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. As discussed in Sec. 4.1, the fit is given as a Johnson SU signal model for the B_s^0 and B^0 signal, the sum of three bifurcated Gaussian functions to model the partially reconstructed $B_s^0 \to D_s{}^*\pi\pi\pi$ background and an Exponential function to account for combinatorial background. The invariant mass distribution and the fit is shown in Fig. 4.3. All simultaneously performed fits to the $m(D_s\pi\pi\pi)$ distribution, ordered by the respective D_s final state, can be found in the Appendix A.1. The obtained yields are summarized in Table 4.1.

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

The shape of the invariant mass distribution of$B_s^0 \to D_sK\pi\pi$ candidates is described by Johnson SU functions for the B^0 and B_s^0 signal, two sums of three bifurcated Gaussians for the $B_s^0/B^0 \to D_s{}^*K\pi\pi$ partially reconstructed background contributions and two sums of double Crystal Ball functions for the single misID $B_s^0 \to D_s\pi\pi\pi$ and the partially reconstructed, misidentified $B_s^0 \to D_s{}^*\pi\pi\pi$ decays. A simultaneous unbinned maximum likelihood fit is performed and the result is shown in Fig. 4.3. All simultaneously performed fits to the $m(D_sK\pi\pi)$ distribution, ordered by the respective D_s final state, can be found in the Appendix A.1. The obtained yields are summarized in Table 4.1.

4.6 Extraction of signal weights

The sPlot technique [6] is used to extract signal weights from the fits to the invariant mass distributions of our signal and normalizaton channel. This statistical tool assignes a weight to every event, according to it's position in the respective mass distribution, given the fitted signal and background models. The weights can then be used to suppress the background components in every other observable distribution of interest. Figure 4.4 shows the distribution of weights across the invariant mass spectra of $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_sK\pi\pi$ candidates.

fit component	yield 2011	yield 2012	yield 2015	yield 2016
$m(D_sK\pi\pi)$				
$B_s^0 \to D_sK\pi\pi$	392 ± 25	860 ± 38	309 ± 21	1984 ± 55
$B^0 \to D_sK\pi\pi$	276 ± 26	692 ± 41	261 ± 23	1385 ± 58
$B^0/B_s^0 \to D_s{}^*K\pi\pi$	7 ± 25	171 ± 75	114 ± 25	893 ± 84
$B_s^0 \to D_s{}^{(*)}\pi\pi\pi$	63 ± 0	158 ± 0	53 ± 0	314 ± 0
combinatorial	1482 ± 53	2884 ± 100	605 ± 43	4261 ± 133
$m(D_s\pi\pi\pi)$				
$B_s^0 \to D_s\pi\pi\pi$	9183 ± 105	22083 ± 166	7574 ± 95	43773 ± 245
$B^0 \to D_s\pi\pi\pi$	289 ± 58	716 ± 95	229 ± 54	968 ± 147
$B_s^0 \to D_s{}^*\pi\pi\pi$	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.

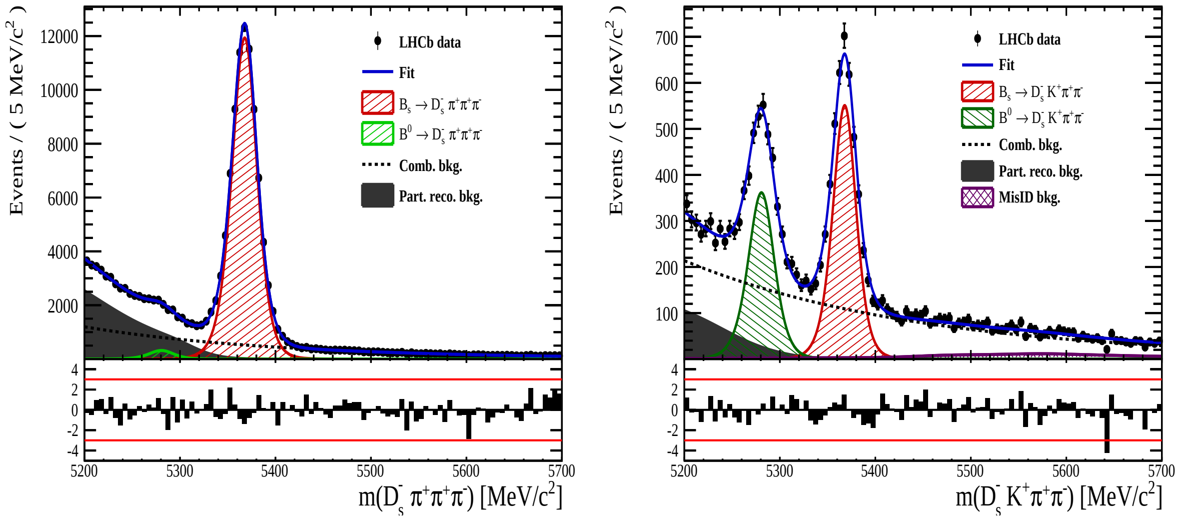


Figure 4.3: Invariant mass distribution of (left) $B_s^0 \to D_s\pi\pi\pi$ and (right) $B_s^0 \to D_sK\pi\pi$ candidates for Run1 and Run2 data. The respective fit described in the text is overlaid.

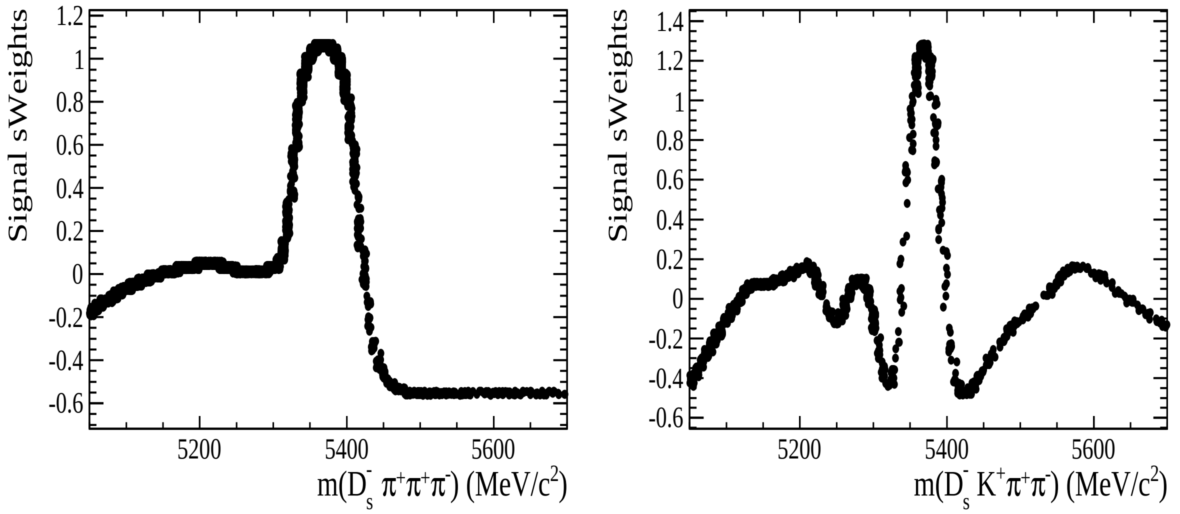


Figure 4.4: Distribution of sWeights across the invariant mass of (left) $B_s^0 \to D_s\pi\pi\pi$ and (right) $B_s^0 \to D_sK\pi\pi$ candidates for Run1 and Run2 data.

5 Flavour Tagging

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

$$D = 1 - 2\omega. \tag{5.1}$$

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

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

where $epsilon_{tag}$ is the fraction of events that have a tagging decision. The flavour tagging algorithms are optimised for highest ϵ_{eff} on data, using the $B_s^0 \rightarrow D_s^- \pi^+$ and $B^+ \rightarrow J/\psi K^+$ samples.
Utilizing flavour-specific final states, the predicted mistag η of each tagger has to be calibrated to match the observed mistag ω on the data sample. For the calibration, a linear model of the form

$$\omega(\eta) = p_o + p_1 \cdot (\eta - <\eta>), \tag{5.3}$$

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

The tagging calibration is determined using a time-dependent fit to the full $B_s^0 \to D_s\pi\pi\pi$ sample, where the mixing frequency Δm_s is fixed to the nominal PDG value [8]. The calibration procedure for the OS tagging algorithms (Sec.5.1) and the SS kaon tagger (Sec.5.2) is applied on the full Run I and 2015 and 2016 Run II $B_s^0 \to D_s\pi\pi\pi$ data sample, which is selected following the steps described in Sec. 3. The similar selection ensures as close as possible agreement between the $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_sK\pi\pi$ samples in terms of the decay kinematics, which are crucial for the flavour tagging. Section 5.3 shows the compatibility of both samples. After applying the calibration, the response of the OS and SS taggers are combined, which is shown in Sec. 5.4.

5.1 OS tagging calibration

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

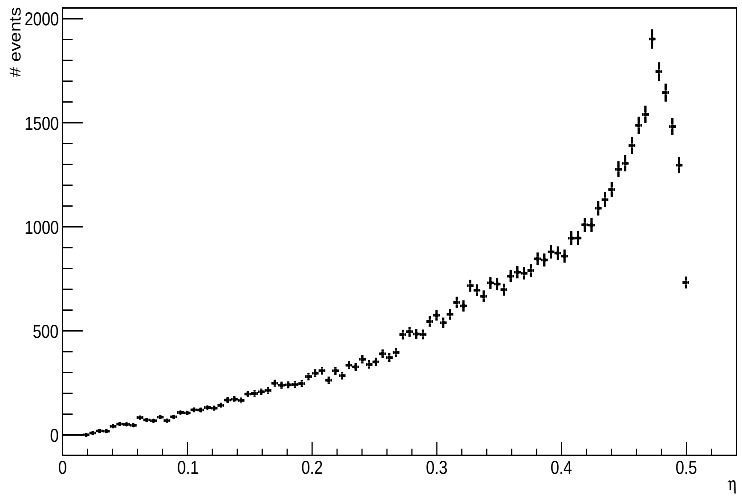


Figure 5.1: Distribution of the predicted OS combination mistag probablity for $B_s^0 \to D_s\pi\pi\pi$ signal candidates.

p_0	p_1	$<\eta>$	ϵ_{tag}	Δp_o	Δp_1	ϵ_{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 ± 0.04 (stat) ± 0.37 (cal)

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

5.2 SS tagging calibration

The SS neural net kaon tagger can be calibrated using the flavour-specific $B_s^0 \to D_s\pi\pi\pi$ decay. It's development, performance and calibration is described in detail in [9]. Figure 5.2 shows the distribution of the predicted mistag of the neural net kaon tagger. The extracted calibration parameters and tagging asymmetries are summarized in Table 5.2 and the measured tagging power for this algorithm is $\epsilon_{eff,SS} = 3.22\%$.

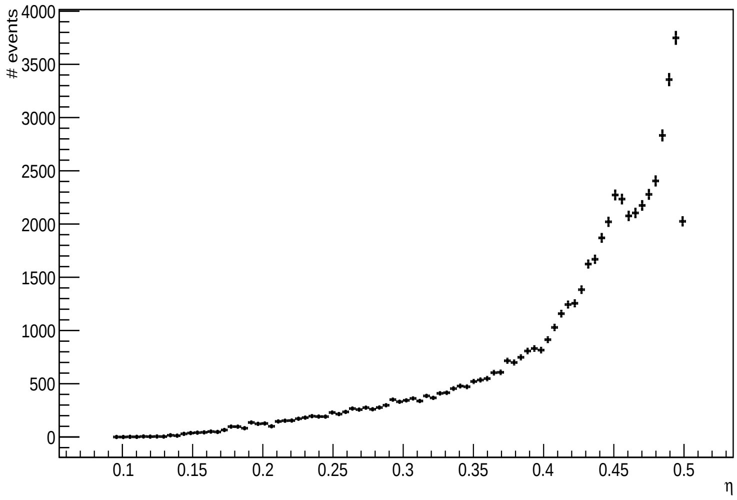


Figure 5.2: Distribution of the predicted SS neural net kaon tagger mistag probablity for $B_s^0 \to D_s\pi\pi\pi$ signal candidates.

p_0	p_1	$<\eta>$	ϵ_{tag}	Δp_o	Δp_1	ϵ_{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 ± 0.03 (stat) ± 0.26 (cal)

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

5.3 Tagging performance comparison between the signal and normalization channel

To justify the usage of the tagging calibration, obtained using the $B_s^0 \to D_s\pi\pi\pi$ sample, for our signal decay, the performance of the taggers in the two decay channels needs to be compatible. This is verified using both, simulated signal samples of both decays and 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. 5.3 (simulation) and Fig. 5.4 (data).

Both, data and simulated samples, show good agreement between the signal and normalization channel. Compatibility is also seen in Fig. 5.5, which shows the comparison of the tagging decision distributions of the OS and SS tagger for sweighted data.

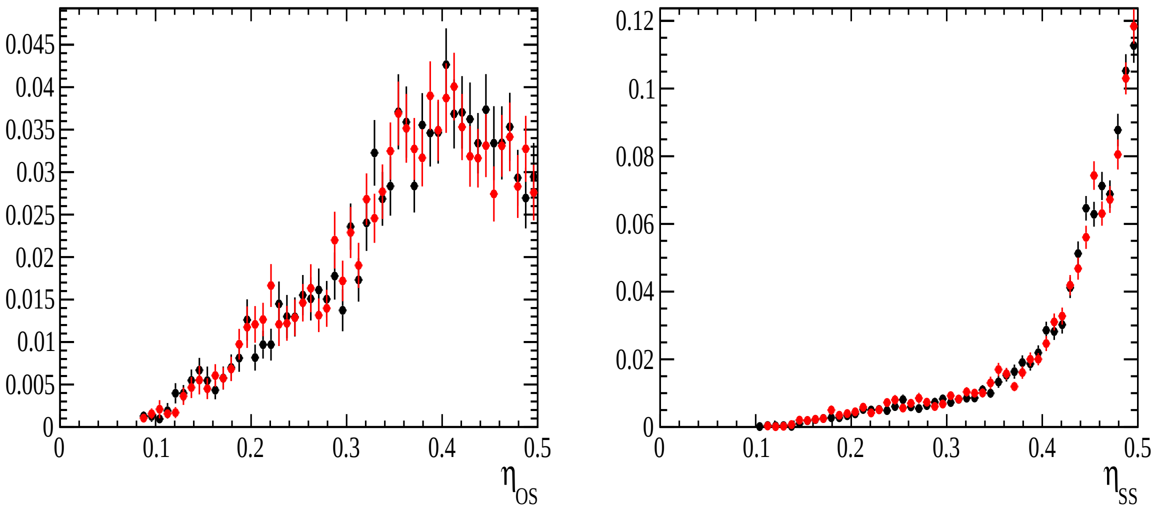


Figure 5.3: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) in simulated $B_s^0 \to D_s K\pi\pi$ (black) and $B_s^0 \to D_s\pi\pi\pi$ (red) signal.

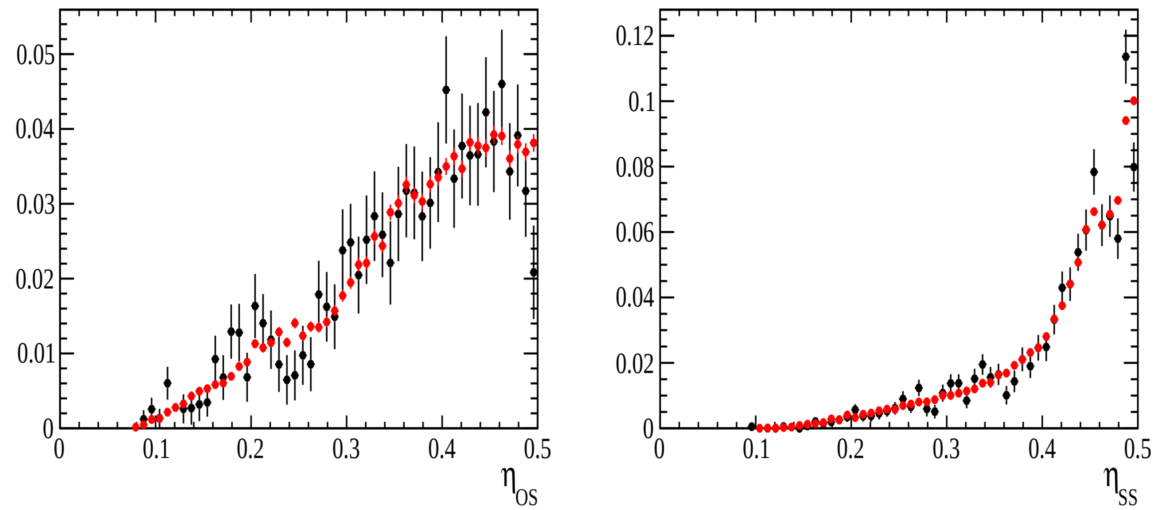


Figure 5.4: Distributions of the predicted mistag η for the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \to D_s K\pi\pi$ (black) and $B_s^0 \to D_s\pi\pi\pi$ (red) data samples. The signal distributions are obtained using sWeights, the procedure is described in Sec. 4.6.

Fig. 5.8 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 $B_s^0 \to D_s\pi\pi\pi$ to the $B_s^0 \to D_s K\pi\pi$ 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. 5.9 for simulated signal events of both channels, where good agreement is observed.

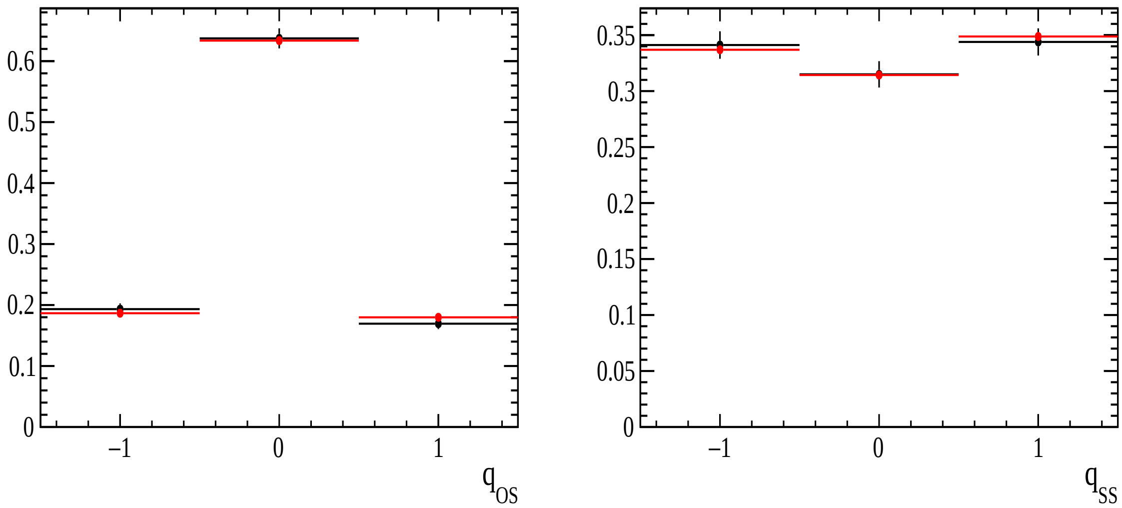


Figure 5.5: Distributions of the tagging decision from the OS combination (left) and the SS kaon tagger (right) for signal candidates in the $B_s^0 \to D_s K\pi\pi$ (black) and $B_s^0 \to D_s\pi\pi\pi$ (red) data samples. The signal distributions are obtained using sWeights, the procedure is described in Sec. 4.6.

5.4 Combination of OS and SS taggers

In the time- and amplitude-dependent fit to $B_s^0 \to D_s K\pi\pi$ 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 aquire 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 tager, 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 taggging performance for the combined tagger in the categories SS tagged only, OS tagged only and SS+OS tagged, are shown in Tab. 5.4 for the signal and normalization channel. The distribution of the observed mistag ω as a function of the combined mistag probability η for $B_s^0 \to D_s\pi\pi\pi$ decays is shown in Fig. 5.10.

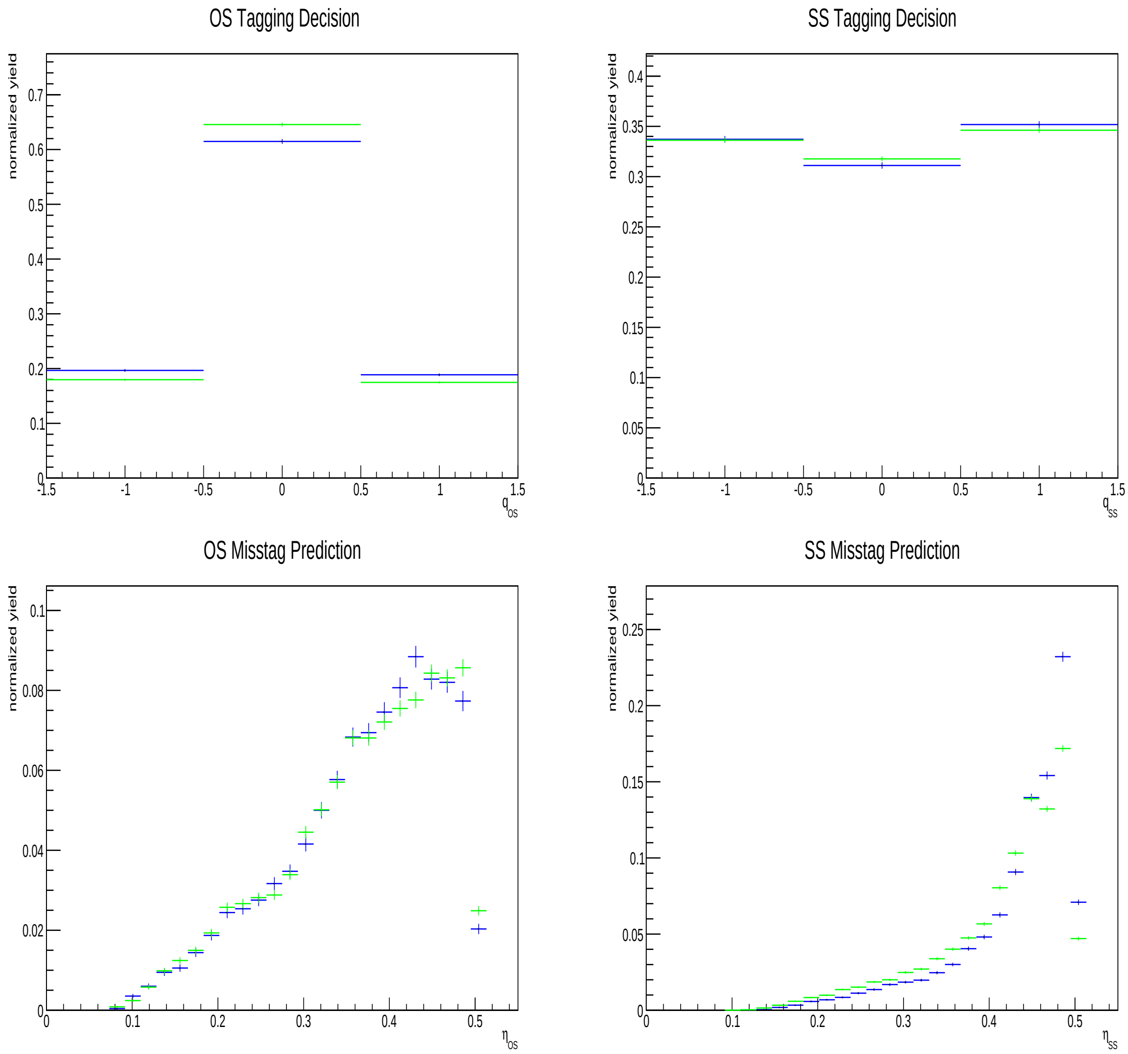


Figure 5.6: Distributions of the tagging decision from the OS combination (upper left) and the SS kaon tagger (upper right), as well as the predicted mistag η for OS (bottm left) and the SS (bottom right), for $B_s^0 \to D_s\pi\pi\pi$ candidates in the Run 1 (blue) and Run 2 (green) data samples. The signal distributions are obtained using sWeights, the procedure is described in Sec. 4.6.

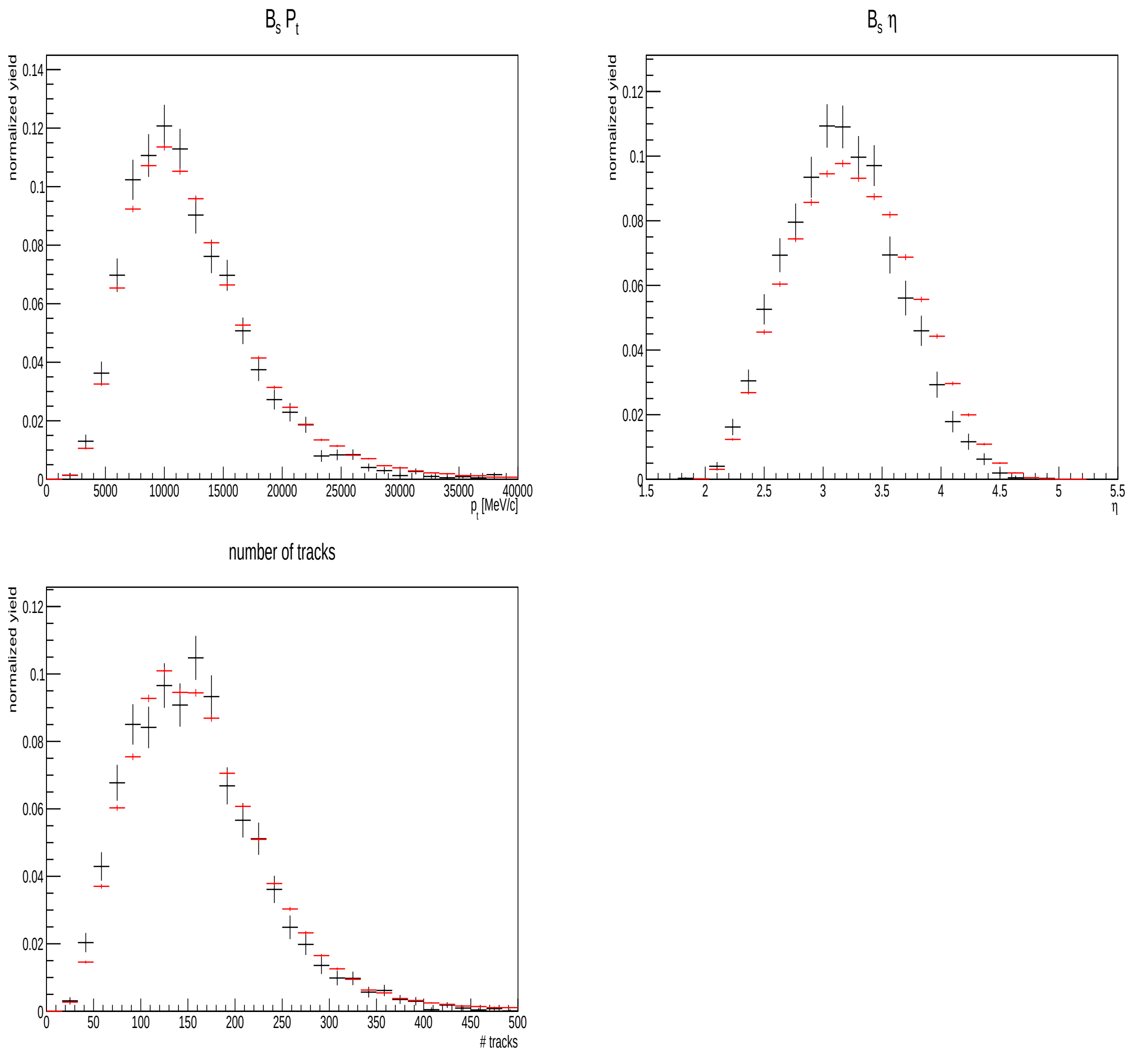


Figure 5.7: Distributions of the transverse momentum p_{T} (top left), the pseudorapidity η (top right) and the reconstructed number of tracks in the event (bottom left) for signal candidates in the $B_s^0 \to D_s K\pi\pi$ (black) and $B_s^0 \to D_s\pi\pi\pi$ (red) data samples. The signal distributions are obtained using sWeights, the procedure is described in Sec. 4.6.

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

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

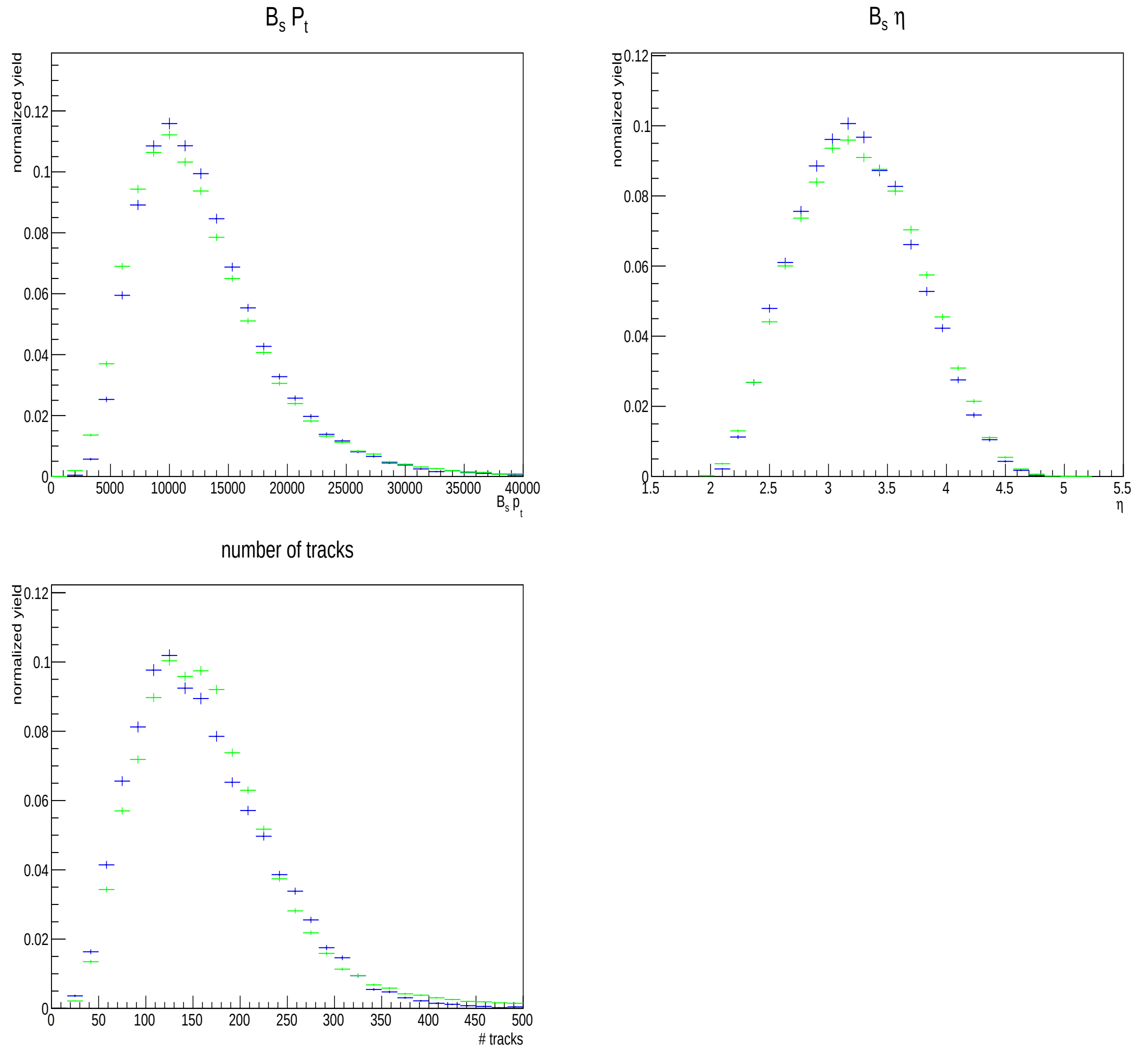


Figure 5.8: Distributions of the transverse momentum p_T (top left), the pseudorapidity η (top right) and the reconstructed number of tracks in the event (bottom left) for $B_s^0 \to D_s\pi\pi\pi$ candidates in the Run 1 (blue) and Run 2 (green) data samples. The signal distributions are obtained using sWeights, the procedure is described in Sec. 4.6.

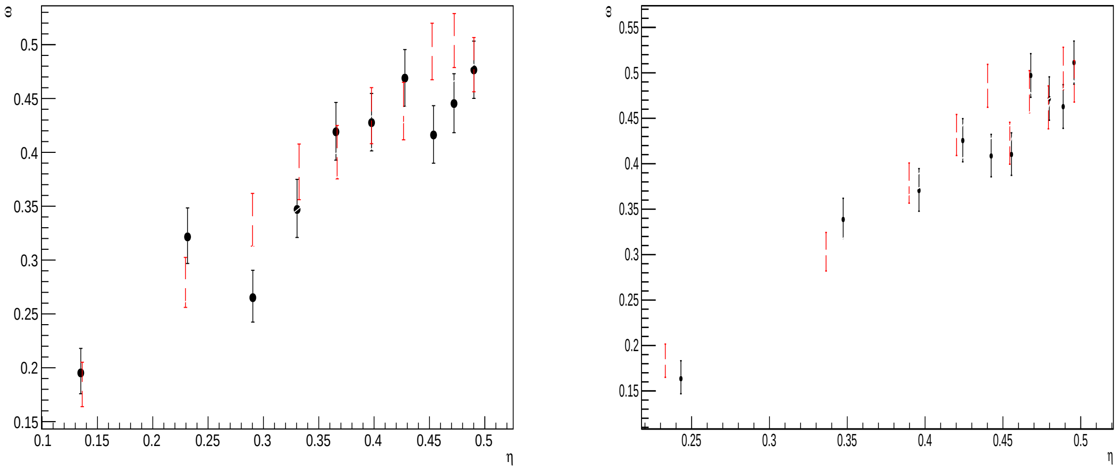


Figure 5.9: Dependence of the observed mistag ω on the predicted mistag η for the (left) OS combination and ther (right) SS kaon tagger, found in the simulated $B_s^0 \to D_s K\pi\pi$ (black) and $B_s^0 \to D_s\pi\pi\pi$ (red) signal samples.

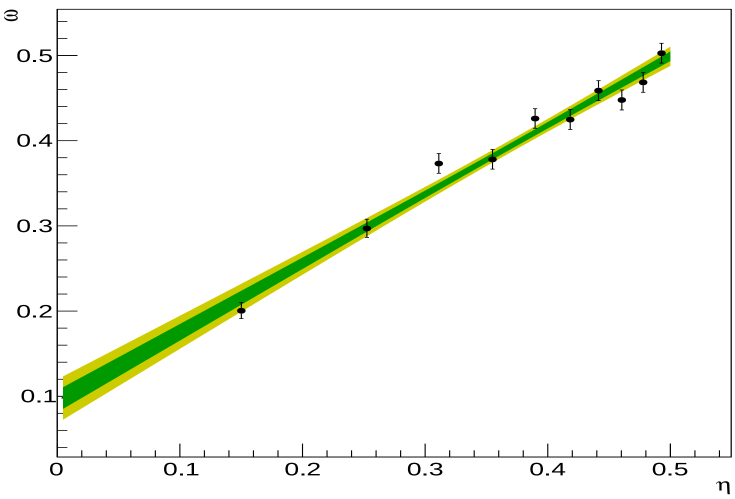


Figure 5.10: Distribution of the predicted combined mistag probablity η versus the observed mistag ω for $B_s^0 \to D_s\pi\pi\pi$ signal candidates. The fit with a linear polynomial, used to determine p_0 and p_1 is overlaid.

6 Decay-time Acceptance

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

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

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

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

is fit to the decay time distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates in data. Since the fit is performed untagged, the PDF shown in Eq. 6.2 contains no terms proportional to Δm_s. The values for Γ_s and $\Delta\Gamma_s$ are fixed to the latest HFAG results [11]. The decay-time acceptance $\epsilon(t', \vec{\lambda})$ is modelled using the sum of cubic polynomials $v_i(t)$, so called Splines [12]. The polynomials are parametrised by so-called knots which determine their boundaries. Knots can be set across the fitted distribution to account for local changes in the acceptance shape. Using more knots is equivalent to using more base splines which are defined on a smaller sub-range. In total, $n + 2$ base splines $v_i(t)$ are needed to describe an acceptance shape which is parametrised using n knots.
For fits shown in the following, the knots have been placed at $t = [0.5, 1.0, 1.5, 2.0, 3.0, 9.5] ps$. To accomodate these 6 knot positions, 8 basic splines v_i, $i = [1, ..., 8]$ are used. Since a rapid change of the decay time acceptance at low decay times due to the turn-on effect generated by the lifetime and other selection cuts is expected, more knots are placed in that regime. At higher decay times we expect linear behaviour, with a possible small effect due to the VELO reconstruction. Therefore fewer knots are used. Furthermore, v_7 is fixed to 1 in order to normalize the overall acceptance function. To stabilise the last spline, v_8 is fixed by a linear extrapolation from the two previous splines:

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

Here, $N = 8$ and t_{N-1} corresponds to the knot position associated with v_{N-1}. The nominal fit to $B_s^0 \to D_s\pi\pi\pi$ data using this configuration is shown in Figure 6.1. Note that the normalization of the splines in the following figures is not in scale.

The fits to $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_s K\pi\pi$ simulation are shown in Figure 6.2.

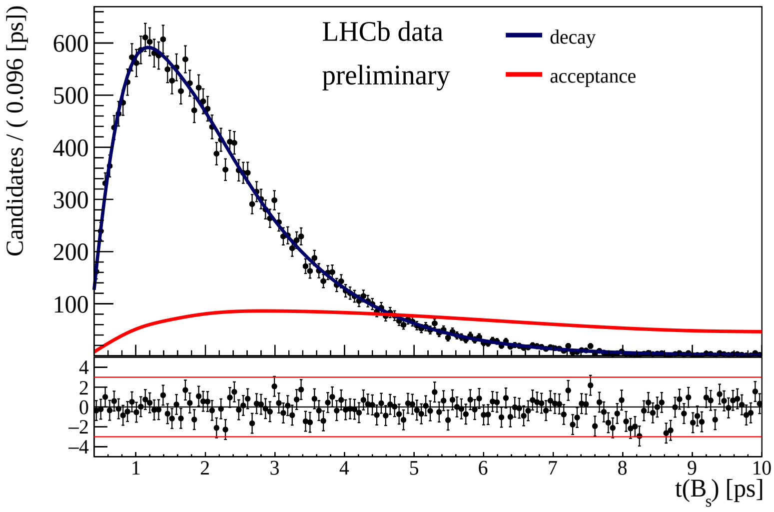


Figure 6.1: Decay-time distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates for the Run 1 data sample. The fit described in the text is overlaid. The red line shows the spline function describing the acceptance and the blue line depicts the total fit function.

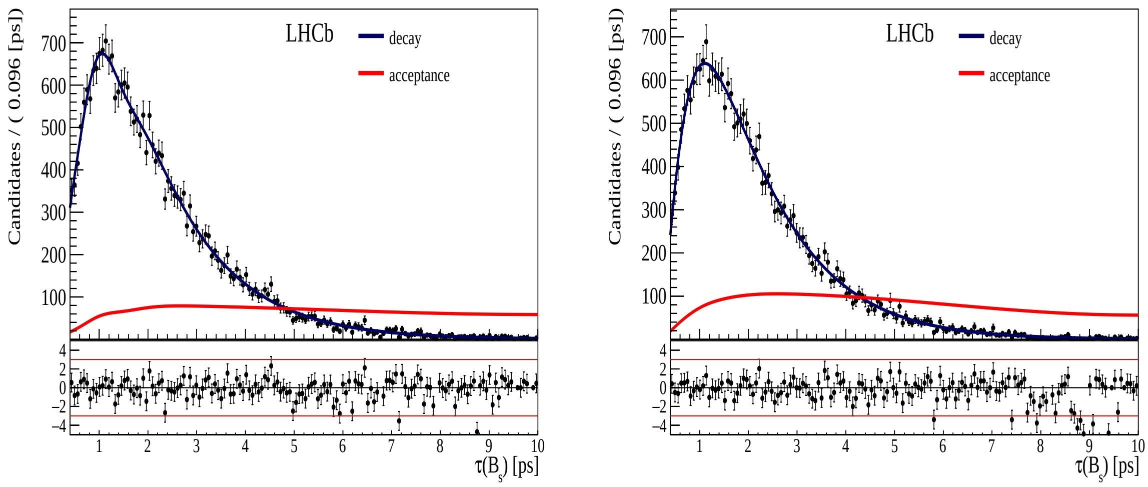


Figure 6.2: Decay-time distribution of (left) $B_s^0 \to D_s\pi\pi\pi$ and (right) $B_s^0 \to D_sK\pi\pi$ candidates in MC using truth information. The fit described in the text is overlaid. The red line shows the spline function describing the acceptance and the blue line depicts the total fit function.

The fit parameters obtained from the described fits to data and simulation are summarised in Table 6.1.

Parameter	Fit to $B_s^0 \to D_s\pi\pi\pi$ data	Fit to $B_s^0 \to D_s\pi\pi\pi$ MC	Fit to $B_s^0 \to D_sK\pi\pi$ MC
v_1			
v_2			
v_3			
v_4			
v_5			
v_6			
v_7	fixed	fixed	fixed

Table 6.1: Summary of the obtained parameters from the acceptance fits described above.

6.1 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 $B_s^0 \to D_s\pi\pi\pi$ 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 6.3 shows the comparison of the obtained spline coefficients for the different D_s final states.

Investigating the obtained spline coefficients from different D_s final states, good aggreement 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 6.4.

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 seperately for the runs and L0 categories. This is achieved by implementing a simultaneous fit, where the acceptance description is allowed to vary in the subsamples.

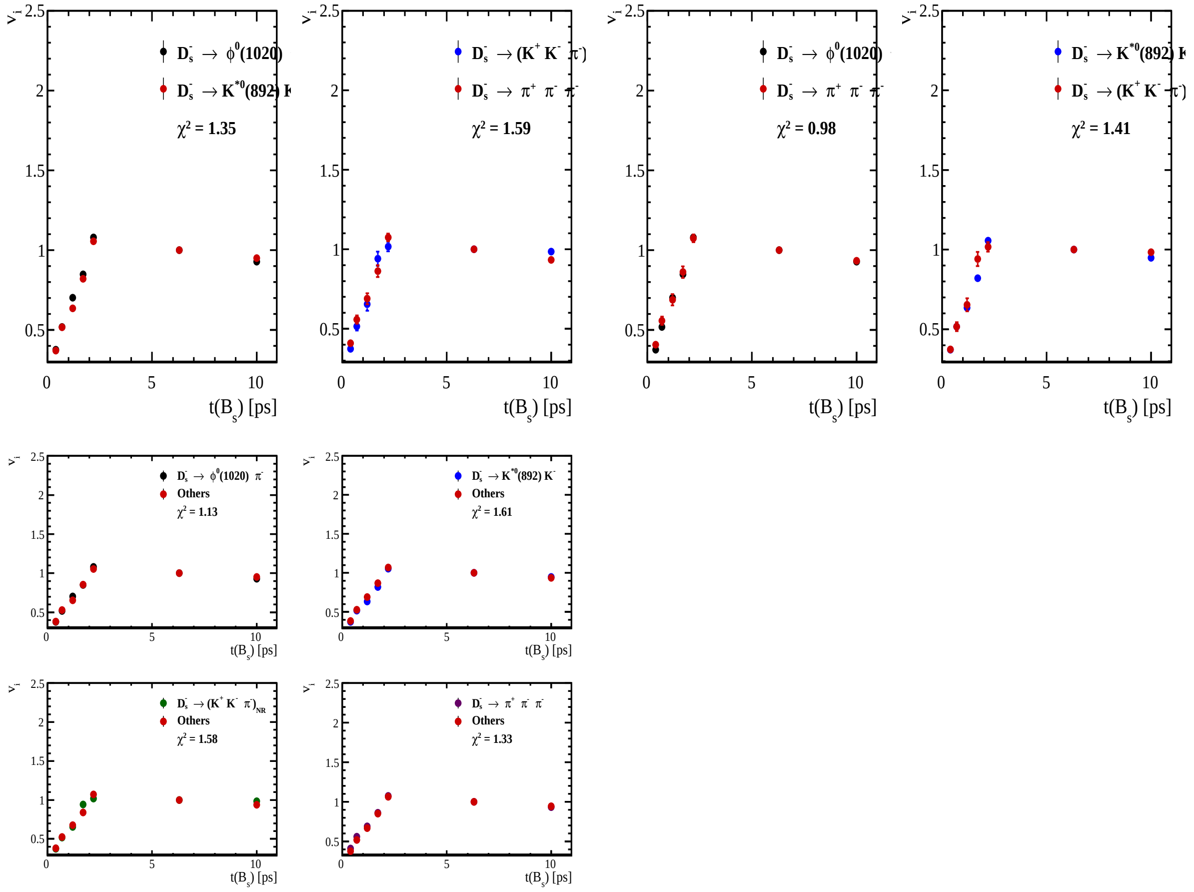


Figure 6.3: Comparison of the spline coefficients obtained from time-dependent fits to the $B_s^0 \to D_s\pi\pi\pi$ subsamples of different D_s final states. The top left and top right distributions show the direct comparison between different D_s final states, whereas the bottom plot shows the comparison of one particular D_s state against all other states.

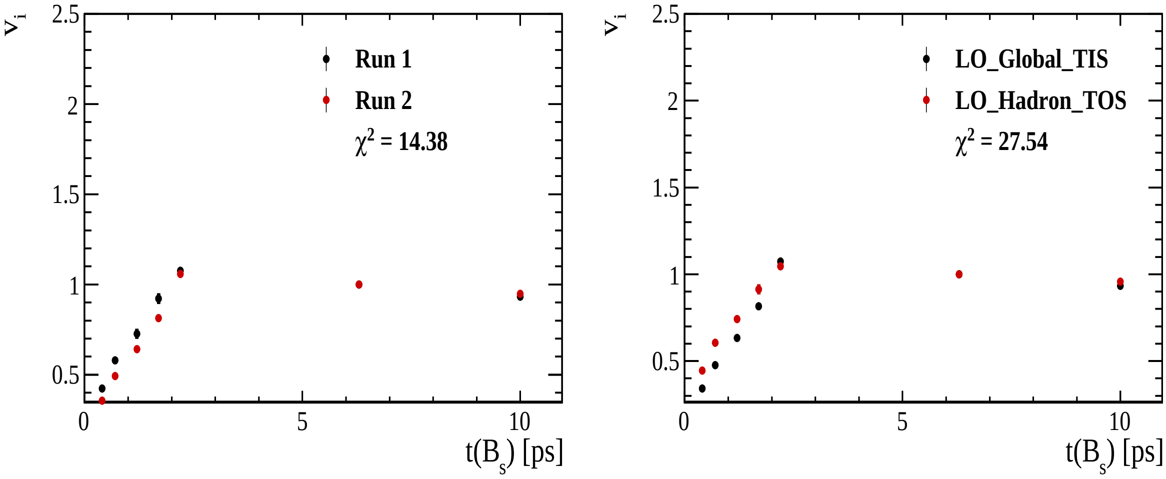


Figure 6.4: Comparison of the spline coefficients obtained from time-dependent fits to the $B_s^0 \to D_s\pi\pi\pi$ subsamples of (left) the different runs and (right) L0 trigger categories.

7 Decay-time Resoution

The observed oscillation of B mesons is prone to dilution, if the detector resolution is of similar magnitude as the oscillation period. In the B_s^0 system, considering that the measured oscillation frequency of the B_s^0 [8] and the average LHCb detector resolution [13] are both $\mathcal{O}(50\,\mathrm{fs}^{-1})$, this is the case. Therefore, it is crucial to correctly describe the decay time resolution in order to avoid a bias on the measurement of time dependent CP parameters.
In the presented analysis, we assume a gaussian resolution function with different widths for each event. This gives rise to a per-event decay time error σ_t, which is computed separately for every event along with the proper time t, by the decay time fitter. Furthermore, the per-event decay time error σ_t is usualy underestimated by the decay time fitter, making it necessary to derive a scaling function, which matches the per-event error to the actually measured decay time resolution.

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

7.1 Formalism

For simulated $B_s^0 \to D_s K\pi\pi$ events, the information on the true B_s^0 lifetime τ_{true} assigned at production of the event, as well as the measured decay time $\tau_{measured}$, which is determined after the interaction with the LHCb detector, is stored. In this analysis, the difference $\Delta t = \tau_{true} - \tau_{measured}$ is obtained for each simulated $B_s^0 \to D_s K\pi\pi$ candidate. The width of the distribution of Δt is a direct measure of the decay time resolution.
To analyse the relation between the per-event decay time eror σ_t and the actual resolution, the simulated sample is split into 8 bins of σ_t. A fit is then performed to the distribution of Δt in each bin to determine the decay time resolution in that bin. After that, the correlation between the binned per-event decay time error and the measured decay time resolution is analyzed to determine the scaling function.

7.2 Fits to the decay time distributions

The sum of two Gaussian functions is used to fit the distributions of the decay time difference Δt in each σ_t bin. One Gaussian function is relatively narrow and describes the decay time of the majority of candidates in each bin, while the other, broader Gaussian function describes candidates where the measure decay time differs considerably from τ_{true}. Those contributions are shifted to the tails of the distribution. From the two Gaussian functions, the combined, effective width σ_{eff} is quoted as decay time resolution in each bin. Figure 7.1 shows the double Gaussian fit to the distirbution of the decay time difference for events where $24\,\mathrm{ps} < \sigma_t < 29\,\mathrm{ps}$. All fits are shown in the Appendix xxY.

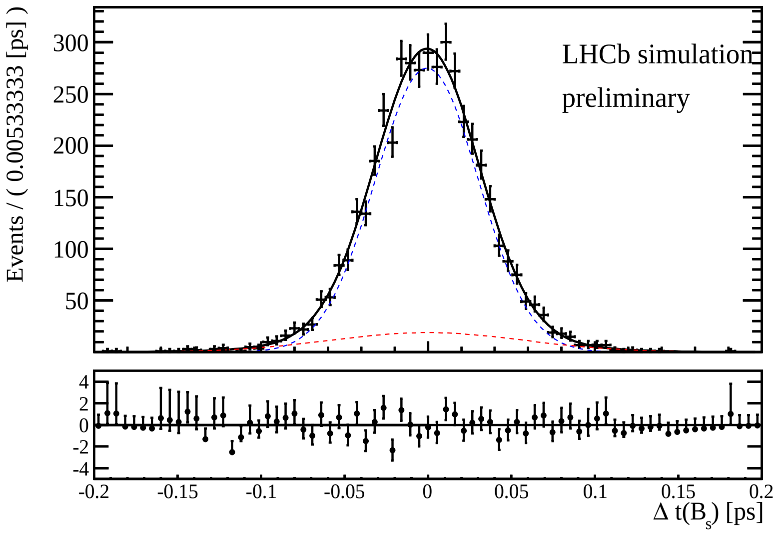


Figure 7.1: Difference of the true and measured decay time of $B_s^0 \to D_s K\pi\pi$ candidates from MC in the bin $24\,\mathrm{ps} < \sigma_t < 29\,\mathrm{ps}$. A fit of the sum of two Gaussian functions is overlaid.

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

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

where f_1 is the fraction of events described by the first Gaussian relative to the second and Δm_s is the oscillation period of the B_s^0 meson.
The dilution is computed in every bin of the per-event decay time error and can be converted into the effective resolution

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

7.3 Results

The fitted values for the Gaussian widths σ_1 and σ_2, the fraction of the first relative to the second Gaussian function f_1, as well as the effective resolution σ_{eff}, found in each bin σ_t, are shown in Tab. 7.1. Figure 7.2 shows the obtained values for σ_{eff} as a function of the per-event decay time error σ_t. A linear polynom of the form

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

is used to parametrise this distribution. The obtained values are

$$\sigma(\sigma_t)_{mc} = (xx.xxx \pm yy.yyy) + (z.zzz \pm a.aaa)\sigma_t. \tag{7.4}$$

For comparison, the linear scaling functions found for $\sigma(\sigma_t)$ in the $B_s^0 \to D_sK$ analysis [REF HERE] for data and MC are also shown in Figure 7.2. We introduce a correction factor to account for possible differences between the decay time resolution found in simulation and data. Motivated by the similarity between the $B_s^0 \to D_sK\pi\pi$ and $B_s^0 \to D_sK$ decay, we assume

$$\frac{\sigma(t)_{D_sK\pi\pi,data}}{\sigma(t)_{D_sK\pi\pi,mc}} \approx \frac{\sigma(t)_{D_sK,data}}{\sigma(t)_{D_sK,mc}}. \tag{7.5}$$

This leads to a correction factor

$$\sigma(t)_{D_sK\pi\pi,data} \approx \frac{\sigma(t)_{D_sK,data}}{\sigma(t)_{D_sK,mc}} \cdot \sigma(t)_{D_sK\pi\pi,mc}, \tag{7.6}$$

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

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

$$\sigma(t)_{D_sK\pi\pi,data} = xXx \pm yYy$$

,

which is the scaling factor used for the per-event decay time error in the nominal time- 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.

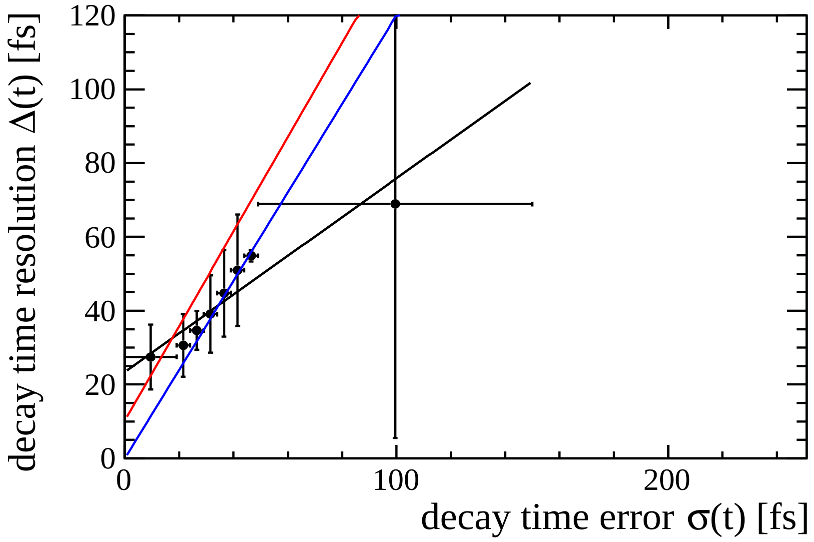


Figure 7.2: Decay-time resolution of $B_s^0 \to D_s K\pi\pi$ candidates from MC. The scaling functions found in for $B_s^0 \to D_s K$ (red) data and (blue) MC is also shown for comparison. The fit described in the text is overlaid.

8 Time dependent fit

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

$$\mathcal{PDF}(t, \vec{\lambda}) = \left(\epsilon(t) \cdot \int P(x, t, q_t, q_f)\mathrm{d}x\right) \times \mathcal{R}(t - t^{'}), \tag{8.1}$$

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

8.1 sFit to $B_s^0 \to D_s\pi\pi\pi$ data

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

$$\int P(x, t, q_t, q_f)\mathrm{d}x = [\cosh\left(\frac{\Delta\Gamma\, t}{2}\right) + q_t q_f \cos\left(\Delta m_s t\right)]e^{-\Gamma t}. \tag{8.2}$$

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

$$\Delta m_s = xx.xxx \pm 0.yyy. \tag{8.3}$$

8.2 sFit to $B_s^0 \to D_s K\pi\pi$ data

8.3 Results

9 Time-dependent amplitude fit

A Appendix

A.1 Detailed mass fits

In this section, all fits to the mass distribution of $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_sK\pi\pi$ candidates are shown. The fits are performed simultaneously for every year of datataking (2011, 2012, 2015 and 2016) and the D_s decay ($D_s \to KK\pi$ non-resonant, $D_s \to \phi\pi$, $D_s \to K^*K$, or $D_s \to \pi\pi\pi$) through which the final state is reached.

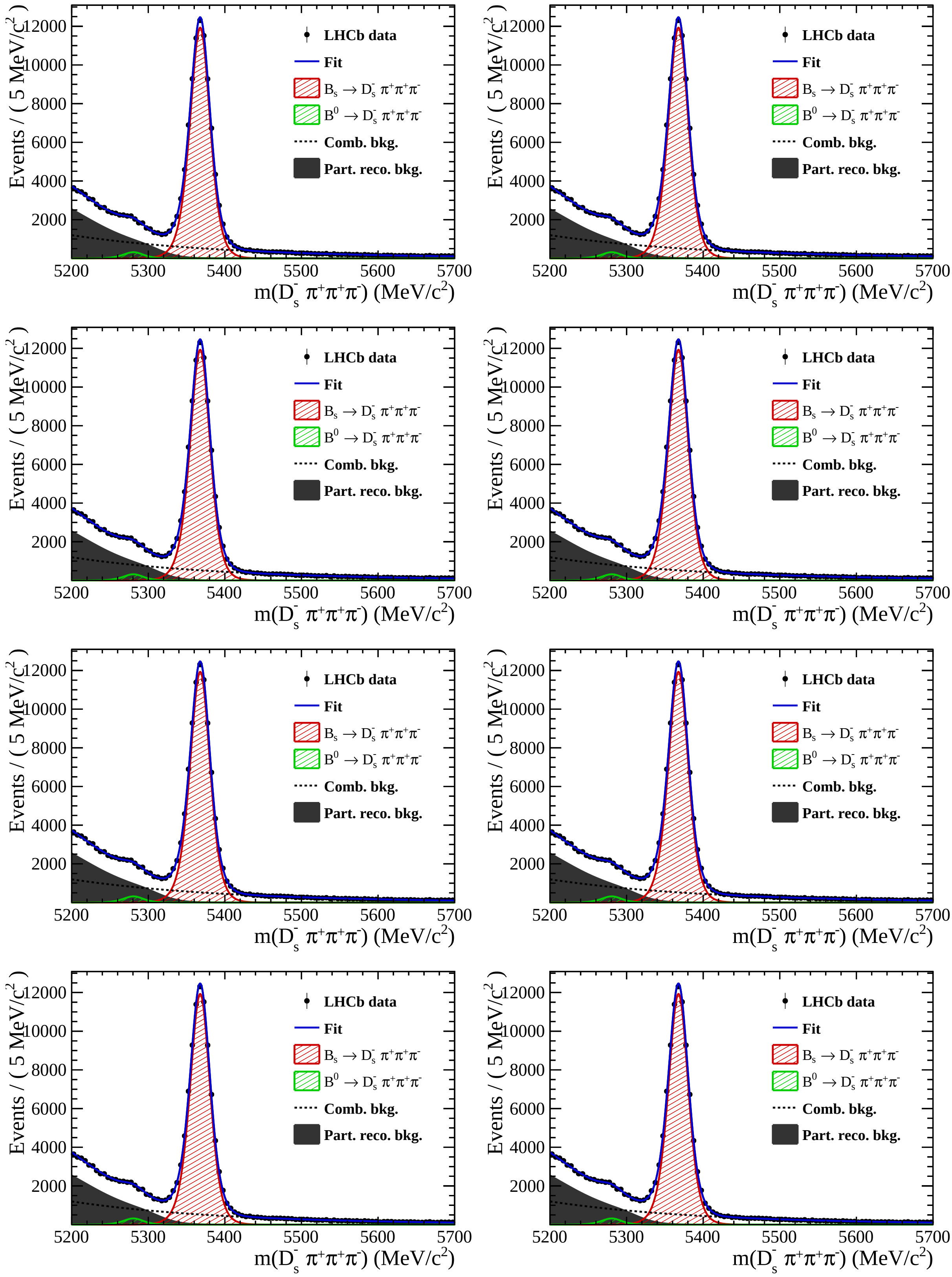


Figure 1.1: Invariant mass distributions of $B_s^0 \to D_s\pi\pi\pi$ candidates, ordered by D_s final state, for Run1 data. The fit described in 4.4 is overlaid.

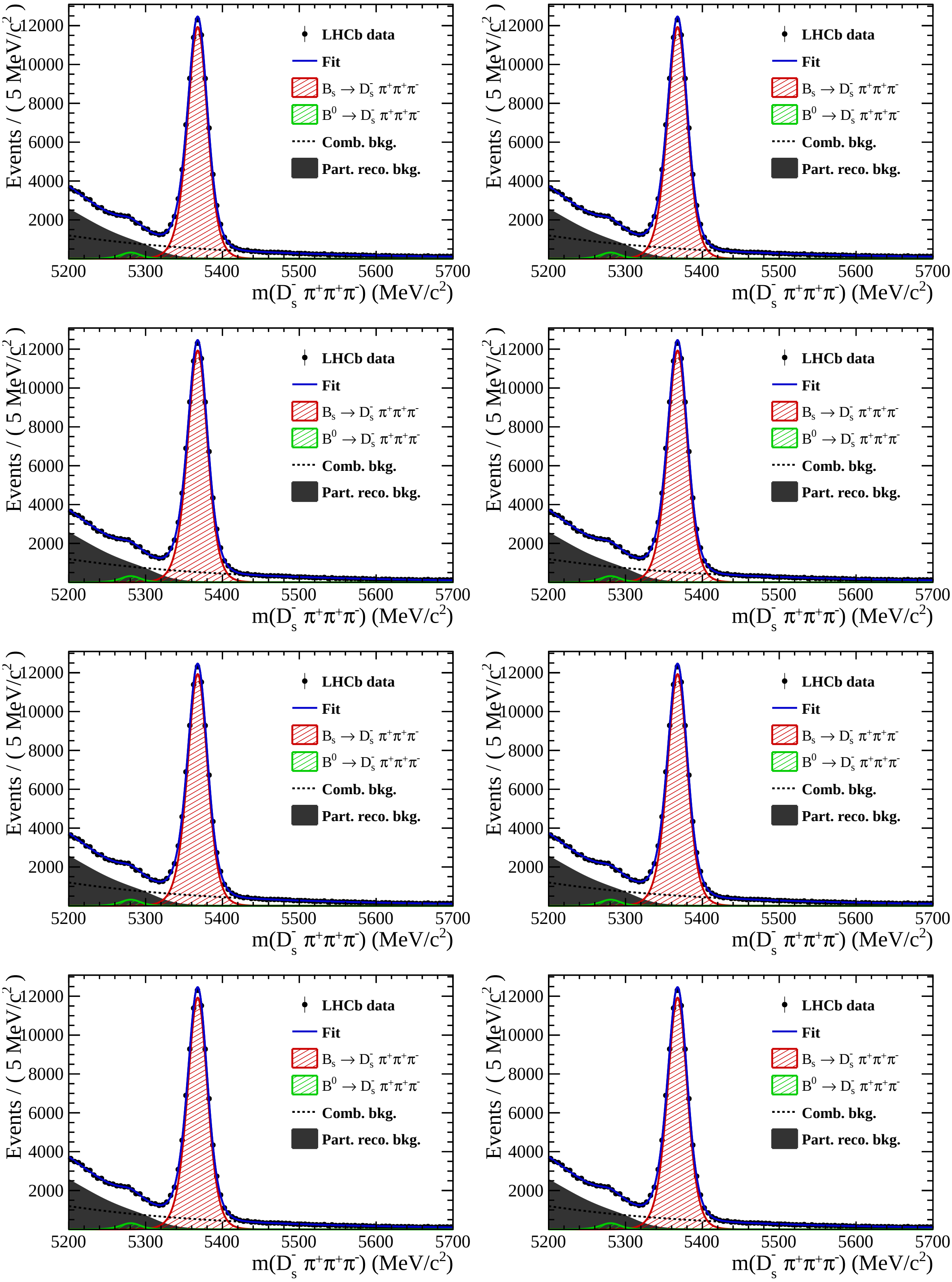


Figure 1.2: Invariant mass distributions of $B_s^0 \to D_s\pi\pi\pi$ candidates, ordered by D_s final state, for Run2 data. The fit described in 4.4 is overlaid.

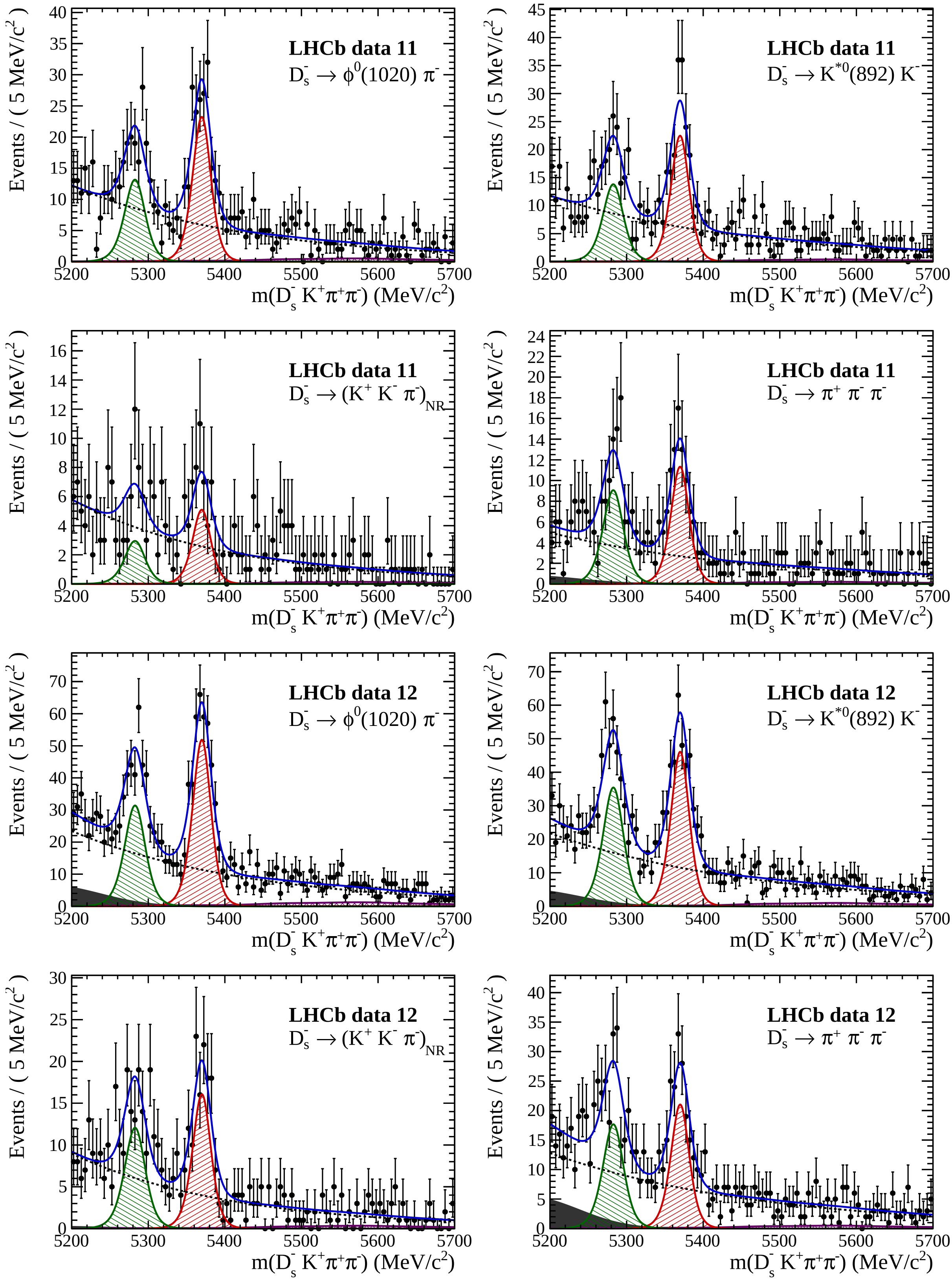


Figure 1.3: Invariant mass distributions of $B_s^0 \to D_s K\pi\pi$ candidates, ordered by D_s final state, for Run1 data. The fit described in 4.5 is overlaid.

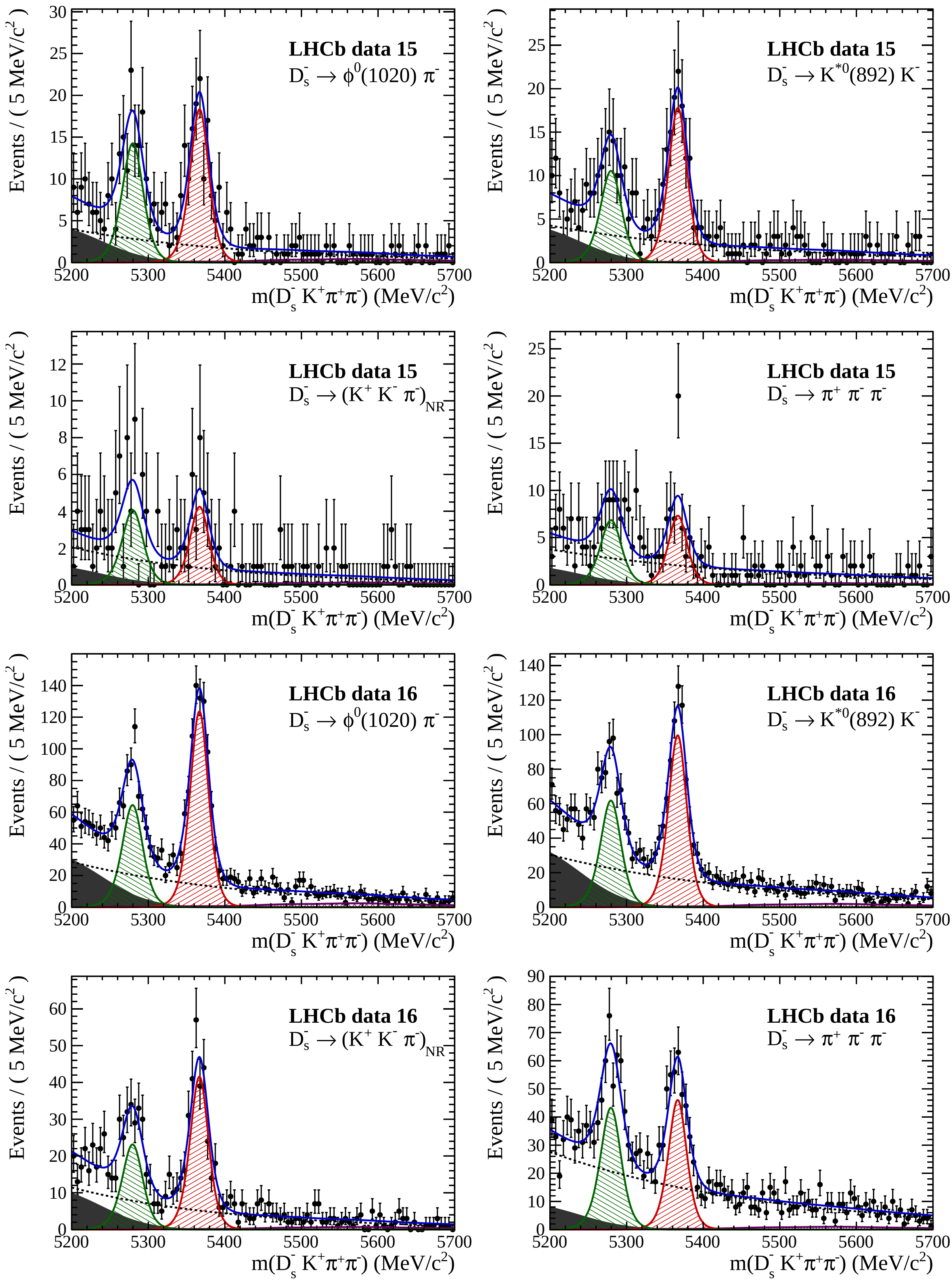


Figure 1.4: Invariant mass distributions of $B_s^0 \to D_s K\pi\pi$ candidates, ordered by D_s final state, for Ru2 data. The fit described in 4.5 is overlaid.

References

[1] R. Fleischer, *New strategies to obtain insights into CP violation through B(s) —¿ D(s)+- K-+, D(s)*+- K-+, ... and B(d) —¿ D+- pi-+, D*+- pi-+, ... decays*, Nucl. Phys. **B671** (2003) 459, `arXiv:hep-ph/0304027`.

[2] K. De Bruyn *et al.*, *Exploring $B_s \to D_s^{(*)\pm} K^{\mp}$ 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 $\bar{B}_s^0 \to D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}^0 \to D_s^+ K^- \pi^+ \pi^-$*, .

[4] LHCb, S. Blusk, *Measurement of the CP observables in $\bar{B}_s^0 \to D_s^+ K^-$ and first observation of $\bar{B}_{(s)}^0 \to D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}_s^0 \to D_{s1}(2536)^+ \pi^-$*, 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] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), no. 1/2 149.

[8] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.

[9] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016), no. 05 P05010, `arXiv:1602.07252`.

[10] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, `arXiv:1202.4979`.

[11] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, `arXiv:1412.7515`, updated results and plots available at `http://www.slac.stanford.edu/xorg/hfag/`.

[12] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, `arXiv:1407.0748`.

[13] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.