Chapter 6

Mass fit to $B^+ \to D_s^+ \phi$ candidates

Contents

In this chapter the methodology used to search for $B^+ \to D_s^+ \phi$ decays is described. The branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)$ is constructed by measuring the yield of $B^+ \to D_s^+ \phi$ decays relative to the normalisation channel $B^+ \to D_s^+ \overline{D}^0$. This ratio is corrected by the ratio of selection efficiencies for the two modes. The branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)$ is then determined by multiplying this corrected ratio by externally measured values for the branching fractions $\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)$ and $\mathcal{B}(\overline{D}^0 \to K^+ K^-)$, and dividing by $\mathcal{B}(\phi \to K^+ K^-)$.

6.1   Fit strategy

The strategy used to search for $B^+ \to D_s^+ \phi$ is more complicated than the method used to search for $B^+ \to D_s^+ K^+ K^-$ decays as outlined in Chapter 5. This is necessary to allow various different signal and background components to be distinguished. In particular, the majority of $B^+ \to D_s^+ K^+ K^-$ decays now act as a background; those not proceeding via the ϕ resonance must be distinguished from $B^+ \to D_s^+ \phi$ decays. The ratio of $B^+ \to D_s^+ \phi$ and $B^+ \to D_s^+ \overline{D}^0$ yields is determined using a simultaneous extended unbinned maximum likelihood fit. Three sets of categories are used, separating the candidates according to D_s^+ meson decay mode, invariant mass of the $K^+ K^-$ pair consiting the ϕ meson, $m(K^+ K^-)$, and the cosine of an angle $\cos \theta_K$. The details and definitions of these categories are listed in Sec 6.1.1. The total extended NLL for this fit is created from the sum of each NLL in each of the categories

$$-\log \mathcal{L}(n_0...n_j, \vec{p}) = \sum_\alpha \sum_\beta \sum_\gamma \left(-\log \mathcal{L}^{\alpha,\beta,\gamma}(n_0^{\alpha,\beta,\gamma}...n_j^{\alpha,\beta,\gamma}, \vec{p}) \right) \qquad (6.1)$$

where α, β and γ represent indexes over the D_s^+ mode, $m(K^+K^-)$ and $\cos\theta_K$ categories. The NLL for each category is defined as

$$-\log\mathcal{L}^{\alpha,\beta,\gamma}(n_0^{\alpha,\beta,\gamma}...n_j^{\alpha,\beta,\gamma}, \vec{p}) = -\sum_i^{N^{\alpha,\beta,\gamma}} \log\left(\sum_j n_j^{\alpha,\beta,\gamma} f_j^{\alpha,\beta,\gamma}(m=m_i, \vec{p})\right) + \sum_j n_j^{\alpha,\beta,\gamma}.$$
(6.2)

As before, j represents the index over each contribution to the fit model, and i represents each of $N^{\alpha,\beta,\gamma}$ entries in the data set for category α, β, γ. The composite extended NLL is minimised with respect to the parameters $\vec{p}$ to find the values for which the data is most likely.

The fit is implemented using the ROOFIT package within the ROOT framework. As the yields of candidates is likely to be small (especially in the D_s^+ decay modes with smaller branching fractions) the error on the branching fraction is determined asymmetrically using the MINOS procedure. This determines the error intervals by finding the positions either side of the minimum that the NLL value increases by one unit. As the NLL distribution may not be symmetric this can lead to different upper and lower uncertainties. However, this calculation is computationally expensive, especially in this large fit with many degrees of freedom.

6.1.1 Simultaneous categories

The data sample is split into categories primarily to aid the differentiation of different signal and background components that contribute in similar invariant mass ranges.

D_s^+ meson decay mode

The three D_s^+ decays modes used to reconstruct the signal and normalisation decays ($D_s^+ \to K^+K^-\pi^+$, $D_s^+ \to \pi^+\pi^-\pi^+$, and $D_s^+ \to K^+\pi^-\pi^+$) are fitted simultaneously in different categories. This allows the invariant mass distributions for the three modes to vary slightly in ways that could not be easily accounted for if the modes were combined in a single data set. In principle, the widths and resolutions of the B^+ meson mass distributions could vary for the three different modes as a result of the different numbers of pions and kaons in the final state. The background levels also differ

between the modes as a result of the smaller branching fractions for $D_s^+ \to \pi^+\pi^-\pi^+$ and $D_s^+ \to K^+\pi^-\pi^+$. This leads to the background from combinations of unrelated tracks having a larger relative contribution.

The $D_s^+ \to K^+K^-\pi^+$ decay mode is additionally split into two further categories; candidates consistent with $D_s^+ \to \phi\pi^+$ decays, and non-ϕ candidates. This exploits the high purity of $D_s^+ \to \phi\pi^+$ decays.

Invariant mass of K^+K^- pair, $m(K^+K^-)$

Three distinct ranges of $m(K^+K^-)$ invariant mass are used to split the candidates. The first of these corresponds to normalisation $B^+ \to D_s^+ \overline{D}^0$ candidates within the range $|m(K^+K^-) - m(\overline{D}^0)| < 25\,\mathrm{MeV}/c^2$. These are reconstructed separately to the signal decays, unlike in the search for $B^+ \to D_s^+ K^+K^-$ decays, as detailed in Chapter 4. The candidates reconstructed as $B^+ \to D_s^+ \phi$ decays are split into two ranges; those within $|m(K^+K^-) - m(\phi)| < 10\,\mathrm{MeV}/c^2$, referred to as the *inner ϕ mass category* and those candidates with $10 < |m(K^+K^-) - m(\phi)| < 40\,\mathrm{MeV}/c^2$, referred to as the *outer ϕ mass category*. These two categories for the signal mode allow contributions from decays that do not proceed via a ϕ meson to be distinguished from those that do. The *inner ϕ mass category* contains 88% of signal $B^+ \to D_s^+ \phi$ candidates, with the other 12% in the *outer ϕ mass category*. The $m(K^+K^-)$ invariant mass distribution for simulated $B^+ \to D_s^+ \phi$ is shown in Fig 6.1.

Helicity angle, $\cos\theta_K$

The $B^+ \to D_s^+ \phi$ decay involves the decay of a pseudoscalar particle to a pseudoscalar and vector particle. Therefore the ϕ vector meson ($J^P = 1^-$) must be produced longitudinally polarised. For a longitudinally polarised ϕ meson decaying to K^+K^-, the distribution of the angle θ_K, defined as the angle that the kaon meson forms with the B^+ momentum in the ϕ rest frame (Fig. 6.2), is proportional to $\cos^2\theta_K$. The distribution of $\cos\theta_K$ for $B^+ \to D_s^+ \phi$ as determined from simulated events is shown in Fig 6.1. Candidates are split into two categories; $|\cos\theta_K| > 0.4$ and $|\cos\theta_K| < 0.4$. These categories contain 93% and 7% of the signal respectively.
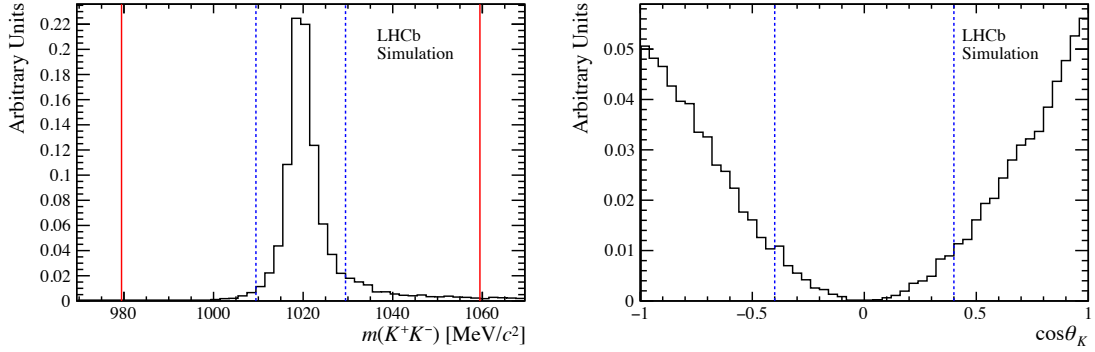
Figure 6.1: The distributions of $m(K^+K^-)$ (left) and $\cos\theta_K$ (right) in simulated $B^+ \to D_s^+ \phi$ decays. The vertical blue dashed lines represent the boundaries between categories defined in Sec. 6.1.1. The vertical red lines represent the mass window applied to candidates; those outside the red lines are not included in the data set.
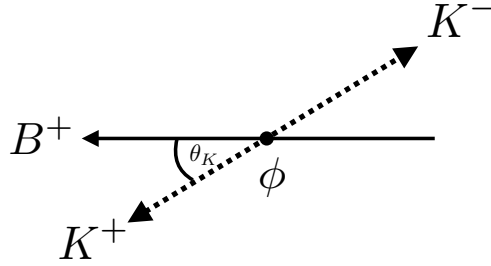
Figure 6.2: The angle θ_K (referred to as the helicity angle) is defined to be the angle that the kaon mesons forms with the B^+ meson momentum in the ϕ rest frame.

This helicity angle is constructed using the momentum of the decay products calculated after the whole decay chain has been refitted with a D_s^+ mass and B^+ direction constraint. This significantly increases the fraction of signal events expected in the first of the two categories.

6.1.2 $B^+ \to D_s^+ K^+ K^-$ model and assumptions

The search for $B^+ \to D_s^+ \phi$ decays includes a component for $B^+ \to D_s^+ K^+ K^-$ decays that didn't proceed via a ϕ meson. This is necessary as the search documented in Chapter 5 determined there is a non-zero contribution from these decays in the range of $m(K^+K^-)$ invariant mass considered here (Fig 5.18). To avoid overestimating $B^+ \to D_s^+ \phi$ signal yield, separate components are included in the fit model for the

$|m(K^+K^-) - m_\phi|$ (MeV/c^2)	Helicity Category	
	$|\cos\theta_K| > 0.4$	$|\cos\theta_K| < 0.4$
< 10	82%	6%
(10, 40)	11%	1%

Table 6.1: Fractions of $B^+ \to D_s^+ \phi$ candidates expected in the helicity and $m(K^+K^-)$ invariant mass categories of the simultaneous fit.

$B^+ \to D_s^+ \phi$ and $B^+ \to D_s^+ K^+ K^-$ decays. Although the invariant mass distributions of these contributions are identical, they can still be disentangled by exploiting the different fractions of these decays expected in each of the helicity angle and $m(K^+K^-)$ categories. The fractions for the $B^+ \to D_s^+ \phi$ signal decays as listed in Table 6.1 show the decays are concentrated in the *inner ϕ mass category* with $|\cos\theta_K| > 0.4$. To determine similar fractions for $B^+ \to D_s^+ K^+ K^-$ decays the LAURA++ package [40] is used to generate a number of simulation samples for different intermediate resonance models.

Only resonances in the K^+K^- system are considered as no significant structure is observed in the $m(D_s^+K^-)$ distribution in Fig. 5.18. As such, all resonances are neutral mesons. The models are generated separately, therefore the effect of interference between any combination of states has been entirely neglected. The generated samples are described in the following sections.

The $\phi(1020)$ resonance

Decays proceeding via a $\phi(1020)$ are produced as a crosscheck. As the simulations generated with LAURA++ have not been reconstructed with the full LHCb detector model, this sample is compared to the existing full simulation samples. The differences in between the fraction of the decays in the different $m(K^+K^-)$ and $\cos\theta_K$ categories in the two samples are taken as a proxy for the potential level of bias introduced by using these generator level samples instead of full simulation. The distribution of these simulated decays in $m(K^+K^-)$ and $\cos\theta_K$ are shown in Fig. 6.3. This figure also include the Dalitz plot distribution of the decays parametrised with the variables

$m^2(D_s^+ K^-)$ and $m^2(K^+ K^-)$. This resonance is generated with a Relativistic Breit-Wigner line shape [41].
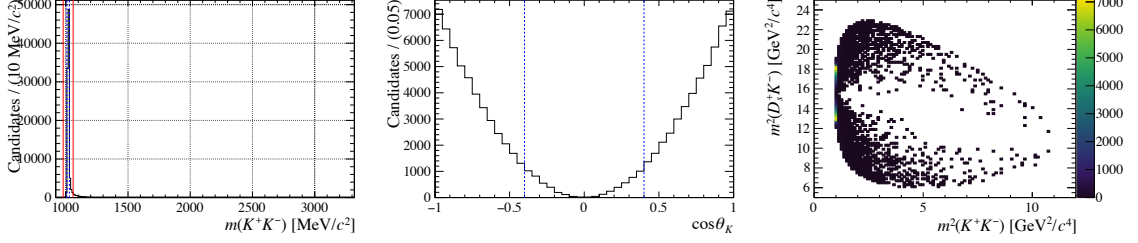

Figure 6.3: The distribution of $m(K^+ K^-)$ (left), Dalitz plot (middle) and the helicity angle $\cos\theta_K$ for generated for the $\phi(1020)$ resonance.

Non-resonant decays

In addition to $K^+ K^-$ resonances, a non-resonant model is considered. This model is defined to have a uniform amplitude across the allowed phase-space. The distribution in $m(K^+ K^-)$ of $B^+ \to D_s^+ K^+ K^-$ decays in Fig. 5.18 is not consistent with this model as there are no candidates above $m(K^+ K^-) \sim 1900\,\text{MeV}/c^2$. However, this component is included in this study for comparative purposes. The distributions of decays generated with this flat model are shown in Fig. 6.4.

Figure 6.4: The distribution of $m(K^+ K^-)$ (left), Dalitz plot (middle) and the helicity angle $\cos\theta_K$ for generated for non-resonant decays.

The $f_0^0(980)$ resonance

The $f_0^0(980)$ resonance is a light unflavoured $J^P = 0^+$ state with mass $990 \pm 20\,\text{MeV}/c^2$ and width 10–$100\,\text{MeV}/c^2$ [27]. It has been observed to decay to $K^+ K^-$ making it a suitable resonance to consider. Although it's mass is at the lower end of the range considered here, it's significant width allows it to contribute at higher invariant

masses. This component is modelled with the Flatté line shape [42] and the relevant distributions shown in Fig. 6.5.

Figure 6.5: The distribution of $m(K^+K^-)$ (left), Dalitz plot (middle) and the helicity angle $\cos\theta_K$ for generated for the $f_0^0(980)$ resonance.

The $a_0^0(980)$ resonance

The $a_0^0(980)$ resonance is a light unflavoured $J^P = 0^+$ state with mass $980 \pm 20 \, \text{MeV}/c^2$ and width 50–$100 \, \text{MeV}/c^2$ and has been observed to decay to $K\overline{K}$ final states [27]. This resonance is also modelled with the Flatté line shape and the relevant distributions are shown in Fig 6.6.
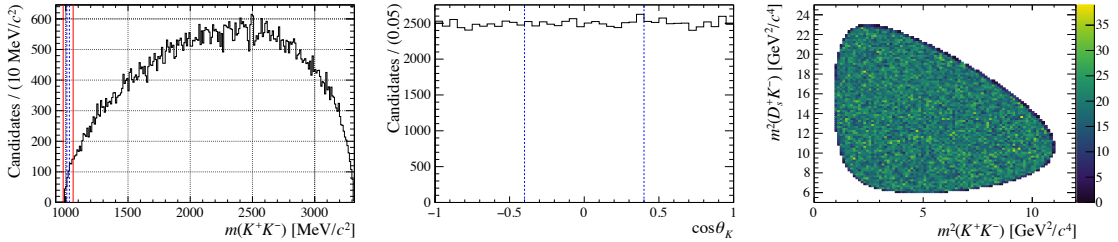

Figure 6.6: The distribution of $m(K^+K^-)$ (left), Dalitz plot (middle) and the helicity angle $\cos\theta_K$ for generated for the $a_0^0(980)$ resonance.

The $f_0^0(1370)$ resonance

The $f_0^0(1370)$ resonance is a light unflavoured $J^P = 0^+$ state with a mass in the range 1200–$1500 \, \text{MeV}/c^2$ and width in the range 200–$500 \, \text{MeV}/c^2$. It has been observed to decay to the $K\overline{K}$ final states. It is modelled with a Relativistic Breit-Wigner line shape and the relevant distributions are shown in Fig. 6.7.

Figure 6.7: The distribution of $m(K^+K^-)$ (left), Dalitz plot (middle) and the helicity angle $\cos\theta_K$ for generated for the $f_0^0(1370)$ resonance.

The $f_2^0(1270)$ resonance

The $f_2^0(1270)$ resonance is a $J^P = 2^+$ state with mass $1275.5 \pm 08\,\mathrm{MeV}/c^2$ and width $186.7^{+2.2}_{-2.5}\,\mathrm{MeV}/c^2$ that has been observed to decay to $K\overline{K}$ final states. This resonance is modelled with a Relativistic Breit-Wigner line shape as shown in Fig. 6.8.
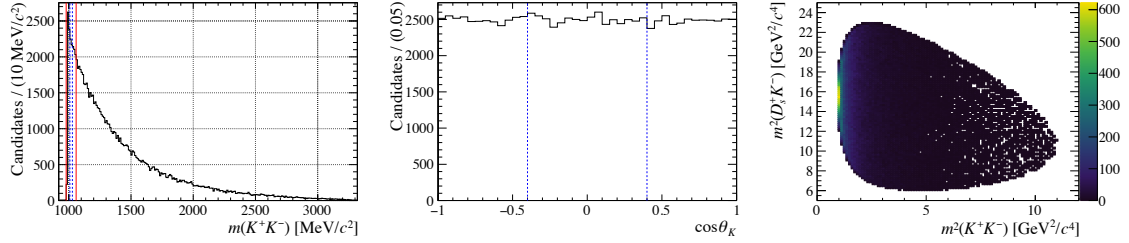

Figure 6.8: The distribution of $m(K^+K^-)$ (left), Dalitz plot (middle) and the helicity angle $\cos\theta_K$ for generated for the $f_2^0(1270)$ resonance.

The $a_2^0(1320)$ resonance

The $a_2^0(1320)$ resonance is a $J^P = 2^+$ state with a mass $1318.1 \pm 0.7\,\mathrm{MeV}/c^2$ and width $109.8 \pm 2.4\,\mathrm{MeV}/c^2$ (both measured in the $K\overline{K}$ mode) observed decaying to the $K\overline{K}$ final state. This resonance is modelled with a Relativistic Breit-Wigner line shape and shown in Fig. 6.9.

Summary of models

The fraction of decays expected in each $m(K^+K^-)$ and $\cos\theta_K$ category for the different models considered are tabulated in Table 6.2. These are calculated by counting the numbers of entries in the corresponding ranges delineated by the vertical lines in Figs. 6.3-6.9. For reference the $\phi(1020)$ fractions are included for both the fully
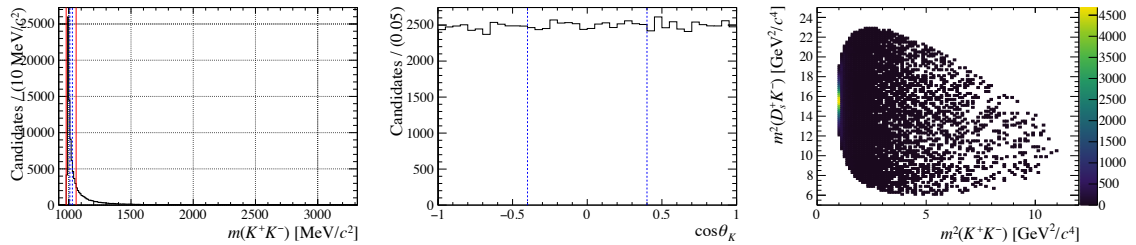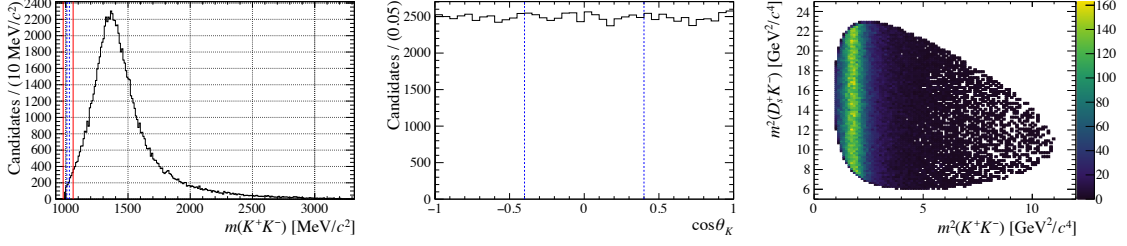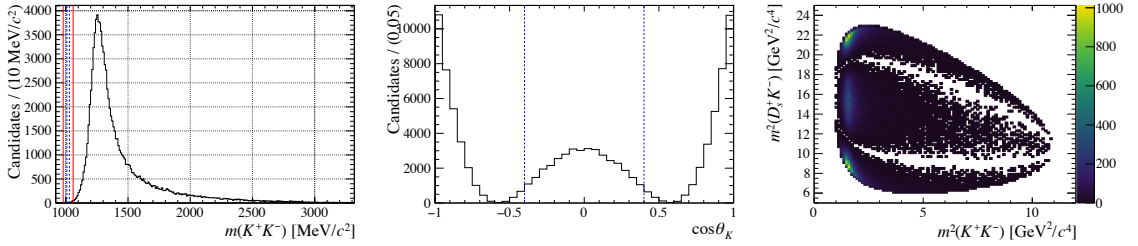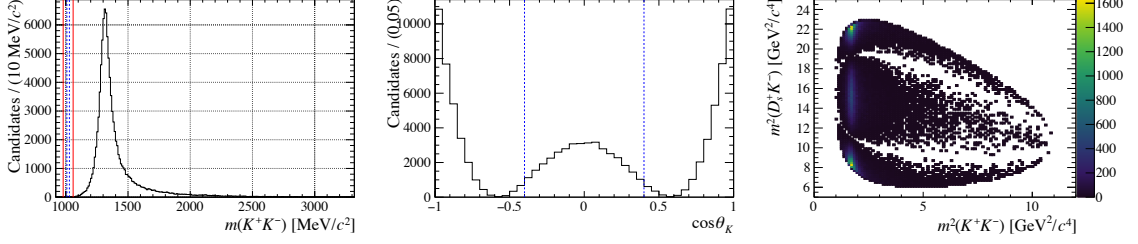
Figure 6.9: The distribution of $m(K^+K^-)$ (left), Dalitz plot (middle) and the helicity angle $\cos\theta_K$ for generated for the $a_2^0(1320)$ resonance.

simulated decays and those generated with LAURA++. The maximum difference between these fractions is included as a source of systematic uncertainty in Sec. 6.7. between For all of the models considered it is clear that the ϕ resonance has significantly different fractions to all of the models considered, allowing the component to be distinguished.

Model	$|\Delta m| < 10\,\mathrm{MeV}/c^2$		$10 < |\Delta m| < 40\,\mathrm{MeV}/c^2$	
	$|\cos\theta_K| > 0.4$	$|\cos\theta_K| < 0.4$	$|\cos\theta_K| > 0.4$	$|\cos\theta_K| < 0.4$
$\phi(1020)$ LAURA++	83.5	5.7	10.2	0.7
$\phi(1020)$ Full	82.4	5.9	10.9	0.8
Non-resonant	16.3	11.1	45.4	27.2
$f_0^0(980)$	16.5	11.2	43.3	29.0
$a_0^0(980)$	12.7	8.7	47.0	31.5
$f_0^0(1370)$	16.1	9.6	45.0	29.3
$f_2^0(1270)$	8.6	6.5	59.1	25.8
$a_2^0(1320)$	9.7	9.7	51.6	29.0
Chosen fractions	14.6 ± 1.9	10.0 ± 1.3	45.2 ± 1.9	30.3 ± 1.3

Table 6.2: Fractions of decays expected in each $m(K^+K^-)$ and $\cos\theta_K$ category for the various resonance models considered in Sec. 6.1.2 ($\Delta m = m(K^+K^-) - m(\phi)$).

In order to choose suitable fractions for the $B^+ \rightarrow D_s^+ K^+ K^-$ fit component a number of points are considered;

- Very few events are observed above $2000\,\mathrm{MeV}/c^2$ in the background-subtracted $m(K^+K^-)$ distribution, therefore the non-resonant model is neglected.

- No significant peaking structure is observed in the $m(K^+K^-)$ spectrum so on-shell resonances are neglected

- The helicity distribution shows no distinctive structure so spin zero states are favoured.

- As this is not a full amplitude analysis no attempt is made to include the effects of interference, either between the remaining off-shell resonances or between these and any possible $B^+ \rightarrow D_s^+ \phi$ decays.

These considerations leave the $f_0^0(980)$ and $a_0^0(980)$ resonances. The fractions of $B^+ \rightarrow D_s^+ K^+ K^-$ decays that have been used in the fit model are fixed to the average of these two, listed in the final row of Table 6.2. Uncertainties are assigned that correspond to half the difference of the two values. These uncertainties are propagated to the $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ branching fraction in Sec. 6.7.

6.2 Fit components

The yields of $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ \overline{D}^0$ decays are extracted from the invariant mass distributions of the data sets by representing each component by probability density functions. The components are broadly very similar to those considered in the search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays detailed in Sec. 5.2. There are a number of necessary differences:

- The B^+ invariant mass range considered for both the signal and normalisation channel is expanded to 4900–5900 MeV/c^2. This allows a more stable determination of the various backgrounds contributing in the vicinity of the signal decays. In particular it stabilises the fraction of decays assigned to the combinatorial and partially reconstructed backgrounds at low B^+ invariant mass.

- More components are included in the model. This include both the signal mode, $B^+ \rightarrow D_s^+ \phi$, and a closely related additional background mode $B^+ \rightarrow D_s^{*+} \phi$.

6.2.1   Signal and normalisation decays

The invariant mass distributions of $B^+ \to D_s^+ \phi$ and $B^+ \to D_s^+ \overline{D}^0$ decays are parametrised using the same DCB function as in the search for $B^+ \to D_s^+ K^+ K^-$ decays (Sec. 5.2.1)

$$\mathrm{DCB}(m|\mu, \sigma_1, \sigma_2, n, \alpha) = f_\sigma \times \mathrm{CB}(m|\mu, \sigma_1, n, \alpha) + (1 - f_\sigma) \times \mathrm{CB}(m|\mu, \sigma_2, n, \alpha), \quad (6.3)$$

where the CB function is defined as follows

$$\mathrm{CB}(m|\mu, \sigma, n, \alpha) = \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2} & \text{if } \left(\frac{m-\mu}{\sigma}\right) < -|\alpha| \\[2ex] \dfrac{\left(\frac{n}{|\alpha|}\right)^n \times e^{-\frac{1}{2}|\alpha|^2}}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{m-\mu}{\sigma}\right)^n} & \text{otherwise.} \end{cases} \quad (6.4)$$

Again, μ, σ, n, α and f_σ are adjustable parameters and m is the B meson invariant mass observable. The tail parameter α is fixed to values determined from maximum likelihood fits to simulated candidates for the signal and normalisation decays. The parameter n is fixed to unity in the fits to both simulation and data to increase the stability of the tails. The two CB function are allowed have different widths, σ_1 and σ_2, but the ratio σ_1/σ_2 is fixed from the fits to simulations, as is f_σ that determines the fractional contribution of the narrower CB function ($\sigma_1 < \sigma_2$). The values determined for each of these fixed parameters are tabulated in Table 6.3, along with the uncertainty obtained from the fits.

An extra constraint is added to the signal and normalisation DCB functions with respect to the configuration used in the search for $B^+ \to D_s^+ K^+ K^-$ decays. As the number of signal candidates is likely to be small, the relative width of the narrower signal and normalisation CB functions, $\sigma_1(D_s^+ \phi)/\sigma_1(D_s^+ \overline{D}^0)$, is also fixed to values obtained from the fits to simulations. All fixed parameters are determined separately for the different D_s^+ decay modes, and the results of the fits to simulated decays are shown in Fig. 6.10.

Parameter	Value		
	$D_s^+ \to K^+ K^- \pi^+$	$D_s^+ \to K^+ \pi^- \pi^+$	$D_s^+ \to \pi^+ \pi^- \pi^+$
$B^+ \to D_s^+ \phi$			
σ_1/σ_2	0.49 ± 0.01	0.47 ± 0.01	0.46 ± 0.01
f_σ	0.80 ± 0.01	0.84 ± 0.01	0.81 ± 0.01
α	2.76 ± 0.07	3.06 ± 0.16	3.71 ± 0.23
n	1 ± 0	1 ± 0	1 ± 0
$B^+ \to D_s^+ \overline{D}^0$			
σ_1/σ_2	0.43 ± 0.01	0.42 ± 0.01	0.40 ± 0.01
f_σ	0.88 ± 0.01	0.88 ± 0.01	0.88 ± 0.01
α	2.91 ± 0.06	3.36 ± 0.26	3.53 ± 0.25
n	1 ± 0	1 ± 0	1 ± 0
$\sigma_1(D_s^+ \phi)/\sigma_1(D_s^+ \overline{D}^0)$	1.27 ± 0.02	1.31 ± 0.02	1.26 ± 0.02

Table 6.3: Fixed values obtained in fits to MC used in the model for the signal pdf.

6.2.2 Partially reconstructed backgrounds

The accurate parametrisation of partially reconstructed backgrounds is particularly important in the search for $B^+ \to D_s^+ \phi$ decays as many different processes contribute to the low invariant mass range of the $m(D_s^+ \phi)$ spectrum. These processes involve decays of B_s^0, B^0 or B^+ mesons in which the five final state tracks reconstructed in the search for $B^+ \to D_s^+ \phi$ decays are only a subset of the background modes final state. Typically, processes in which a low momentum pion or photon has not been reconstructed are found closest in mass to the signal decays. Decays of B_s^0 mesons are particularly dominant as the B_s^0 meson has a larger mass than the B^+ meson.

Backgrounds to the normalisation channel

The modes $B^+ \to D_s^+ \overline{D}^{*0}$ and $B^+ \to D_s^{*+} \overline{D}^0$ can both contribute as partially reconstructed backgrounds to the $B^+ \to D_s^+ \overline{D}^0$ normalisation mode. These are parametrised using the same PDFs as in the search for $B^+ \to D_s K^+ K^-$ decays, detailed in Sec. 5.2.2.

The invariant mass fit range is wider than in the fit to $B^+ \to D_s^+ K^+ K^-$ candidates, therefore an extra contribution is included in the fit model to account for partially reconstructed $B^+ \to D_s^{*+} \overline{D}^{*0}$ decays at lower invariant masses.

(a) $D_s^+ \to K^+ K^- \pi^+$

(b) $D_s^+ \to \pi^+ \pi^- \pi^+$

(c) $D_s^+ \to K^+ \pi^- \pi^+$
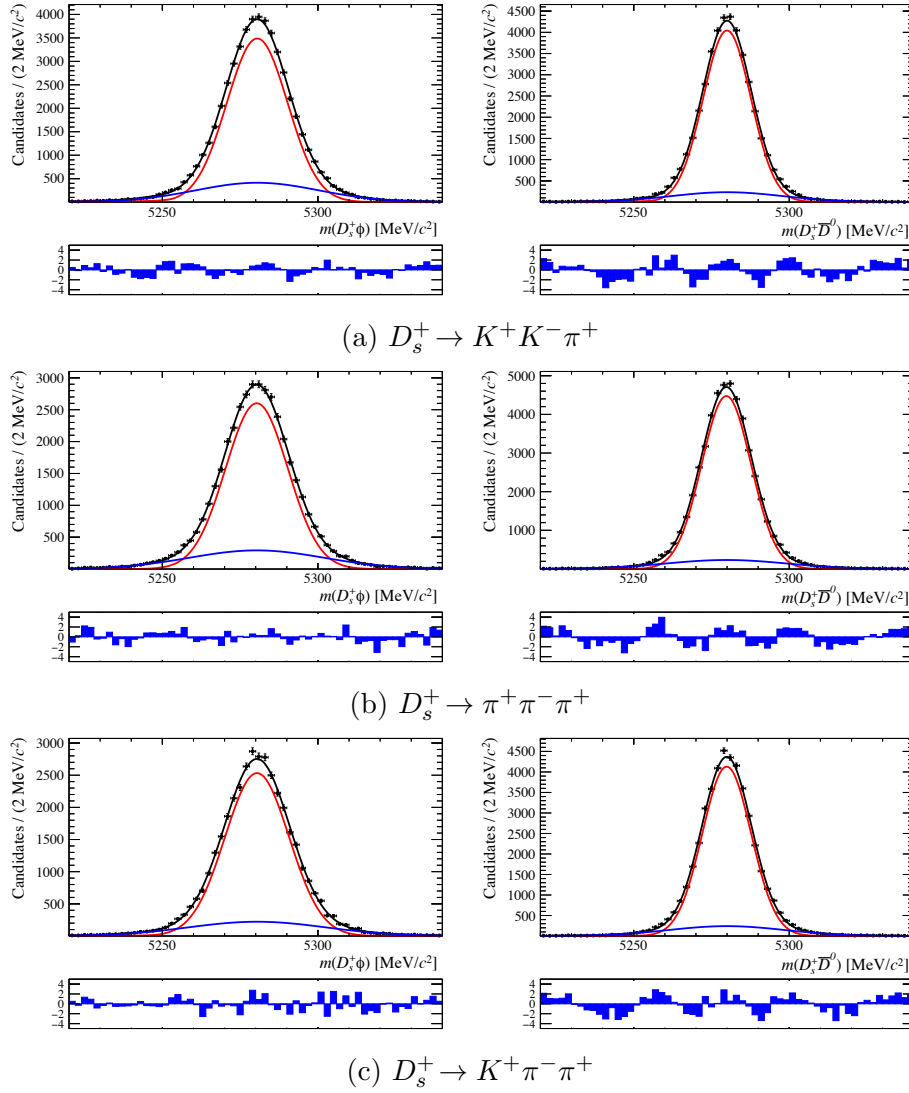
Figure 6.10: Invariant mass fits to simulated signal (left) and normalisation (right) decays. The results of maximum likelihood fits using the signal PDFs are over laid, with the total function in black and the two contributing CB shapes in red and blue.

$B^+ \to (D_s^{*+} \to D_s^+[\pi^0])\overline{D}^0$ **and** $B^+ \to D_s^+(\overline{D}^{*0} \to \overline{D}^0[\pi^0])$**:** these components are modelled by a parabola convolved with a resolution Gaussian. The parabola has a minimum in the centre and doesn't extend beyond endpoints a and b

$$f(m|a,b,\sigma,\xi,\delta) = \int_a^b \left(\mu - \frac{a+b}{2} \right)^2 \left(\frac{1-\xi}{b-a}\mu + \frac{b\xi - a}{b-a} \right) e^{-\frac{-(\mu-(m-\delta))^2}{2\sigma^2}} d\mu. \tag{6.5}$$

These components are shown by the black lines in Fig. 6.11.

$B^+ \to (D_s^{*+} \to D_s^+[\gamma])\overline{D}^0$ **and** $B^+ \to D_s^+(\overline{D}^{*0} \to \overline{D}^0[\gamma])$**:** these components are

modelled by a parabola convolved with a resolution Gaussian. The parabola has a maximum in the centre and doesn't extend beyond endpoints a and b

$$f(m|a,b,\sigma,\xi,\delta) = \int_a^b -(\mu-a)(\mu-b)\left(\frac{1-\xi}{b-a}\mu + \frac{b\xi-a}{b-a}\right)e^{-\frac{-(\mu-(m-\delta))^2}{2\sigma^2}}d\mu. \tag{6.6}$$

These components are shown by the blue lines in Fig. 6.11.
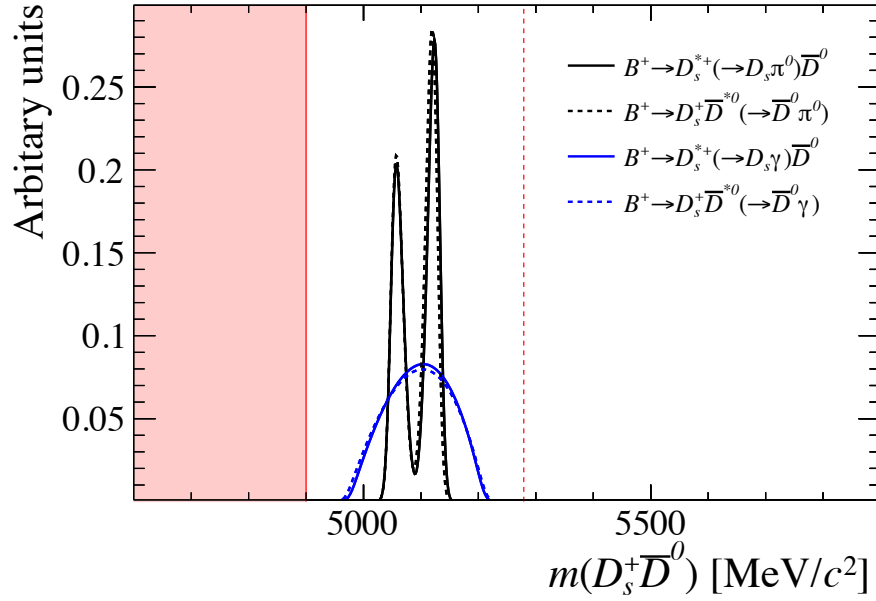
Figure 6.11: Partially reconstructed $B^+ \rightarrow D_s^+ \overline{D}{}^{*0}$ and $B^+ \rightarrow D_s^{*+} \overline{D}{}^0$ decay parametrisations. The B^+ mass is represented by a vertical dashed red line. Candidates in the area below $4900 \,\mathrm{MeV}/c^2$ are not included in the fit.

$B^+ \rightarrow D_s^{*+} \overline{D}{}^{*0}$: the lower invariant mass range used in this search necessitates including a PDF for $B^+ \rightarrow D_s^{*+} \overline{D}{}^{*0}$ decays in which two soft particles have been missed, one from each of the excited D meson decays. It is possible for either a π^0 or γ to be not reconstructed in the decays of both excited D mesons. Additionally, as this process involves a psecudo-scalar meson decaying to two vector mesons, there should be two distinguishable helicity combinations for each process. This leads to a total of eight PDFs necessary to fully parametrise this contribution. Instead, however, this component is parametrised using a single function of the form given in Eq. 6.6. The endpoints a and b are estimated

by combining the effects of missing two neutral particles. The resulting distribution is shown in Fig. 6.12. The choice of PDF for this component is found to have negligible effect on the determination of the $\mathcal{B}(B^+ \to D_s^+ \phi)$ branching fraction as detailed in Sec 6.7.
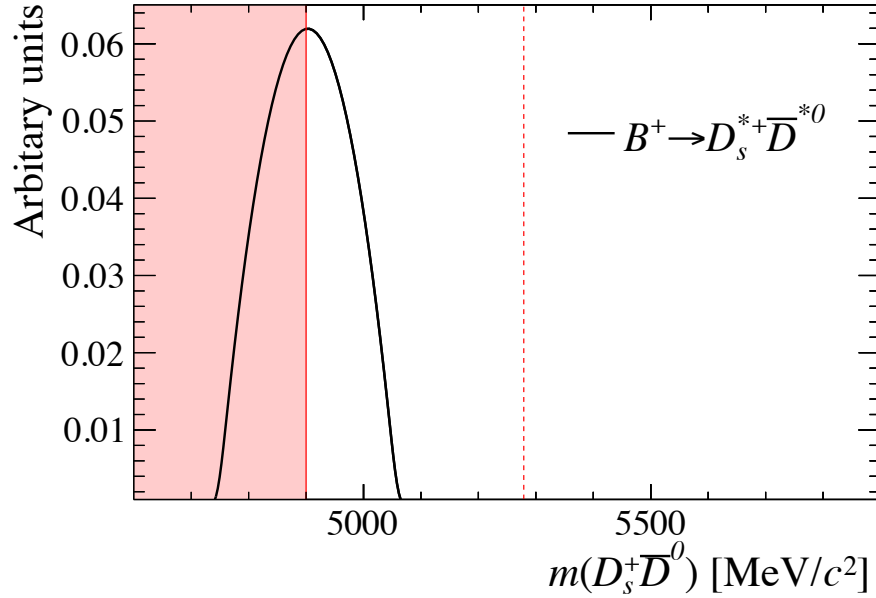
Figure 6.12: Partially reconstructed $B^+ \to D_s^{*+} \overline{D}^{*0}$ decay parametrisation. The PDF extends below the lower fit range at $4900 \, \mathrm{MeV}/c^2$.

Backgrounds to the signal channel

$B^+ \to (D_s^{*+} \to D_s^+[\gamma])\phi$ **and** $B^+ \to (D_s^{*+} \to D_s^+[\pi^0])\phi$: the decays of B^+ mesons to an excited D_s^+ meson and a ϕ meson could contribute to the $m(D_s^+ \phi)$ spectrum at low invariant masses when either a π^0 or γ is missed from the excited meson decay. The resulting invariant mass distribution depends on the mass and spin of the non-reconstructed particle, as well as the helicity state of the D_s^{*+} meson. This background involves the decay of a pseudo-scalar meson to two vector mesons, hence, as a result of angular momentum conservation, there are three helicity states of the D_s^{*+} meson to consider. These are labelled *001*, *010* and *100*. The two transversely polarised states, *100* and *001*, have identical invariant mass distributions and are therefore referred collectively as *101*. This leads to

a total of four contributions to consider for $B^+ \to D_s^{*+}\phi$ decays. These can be parametrised by parabolas convolved with resolution Gaussians in a similar way to the partially reconstructed backgrounds to the normalisation channel. These each share the same functional form

$$f(m|a,b,\sigma,\xi,\delta) = \int_a^b g(\mu,a,b)\left(\frac{1-\xi}{b-a}\mu + \frac{b\xi-a}{b-a}\right)e^{-\frac{-(\mu-(m-\delta))^2}{2\sigma^2}}d\mu. \quad (6.7)$$

where $g(\mu,a,b)$ represents the parabola for each of the four components, listed in Table 6.4. The resulting invariant mass distributions are shown in Fig.6.7.

This decay is unobserved, therefore the relative contribution to this decay in nature from the *101* and *010* helicty states is not known. The total PDF for this contribution is made by weighting the π^0 and γ contributions by their corresponding branching fractions, and by assuming that the *101* and *010* helicty states contribute with equal magnitudes. This assumption is varied and included as a source of systematic uncertainty in the $\mathcal{B}(B^+ \to D_s^+\phi)$ determination.

Missed particle	Helicity	$g(\mu,a,b)$	a (MeV/c^2)	b (MeV/c^2)
π^0	010	$\left(\mu - \frac{a+b}{2}\right)^2$	5026.8	5124.8
π^0	101	$-(\mu-a)(\mu-b)$	5026.8	5124.8
γ	010	$-(\mu-a)(\mu-b)$	4936.4	5220.6
γ	101	$\left(\mu - \frac{a+b}{2}\right)^2 + \left(\frac{a+b}{2}\right)^2$	4936.4	5220.6

Table 6.4: The parabolas contributing to the partially reconstructed $B^+ \to D_s^{*+}\phi$ decay parametrisation as defined in Eq. 6.7.

$\overline{B}_s^0 \to D_s^+ K^- K^{*0}$: this decay can form a background to $B^+ \to D_s^+\phi$ decays when the soft pion from the $K^{*0} \to K^+\pi^-$ decay is not reconstructed. The lower bound of the fit range is wide enough that a significant fraction of these decays are retained in the fitted data set. The K^-K^{*0} is modelled as originating from the $a_1(1260)$ resonance. This resonance has a width of $250-600$ MeV [27], allowing it to decay to K^-K^{*0} even though it's pole mass is below the K^-K^{*0} threshold. A PDF for this component is determined by reconstructing simulated $\overline{B}_s^0 \to$
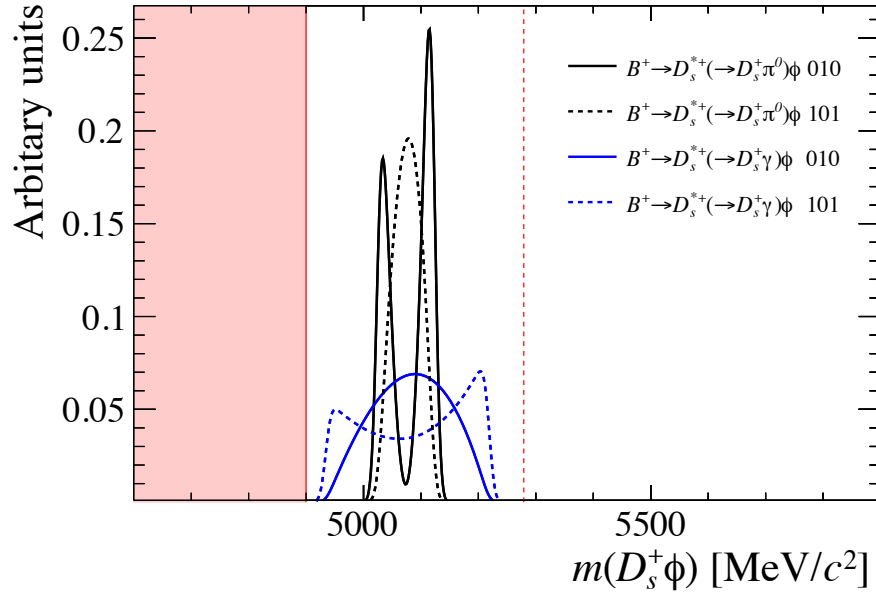
Figure 6.13: Partially reconstructed $D_s^+\phi$ shapes.

$D_s^+ K^- K^{*0}$ decays through the identical reconstruction and selection steps as the signal. The ROOFIT class `RooKeysPDF` is used to create a kernel estimation of the partially reconstructed B^+ mass distribution for the candidates passing the selection. This is shown in Fig. 6.15. The fraction of these decays expected in each of the four $m(K^+K^-)$ and $\cos\theta_K$ categories is determined from the simulations samples.

$\overline{B}_s^0 \to D_s^{*+} K^- K^{*0}$: this decay similarly forms a background to the signal mode when both a low momentum neutral particle (π^0 or γ) is not reconstructed in the decay of the D_s^{*+} meson, in addition to the low momentum pion from the K^{*0} decay. The PDF is similarly determined using a kernel estimation from simulated decays passing the full selection as shown in Fig. 6.15. The fraction of candidates expected in each of the $m(K^+K^-)$ and $\cos\theta_K$ categories are determined from these simulation samples.

$\overline{B}_s^0 \to D_s^+ D_s^-$, $\overline{B}^0 \to D_s^+ D^-$ **and** $\overline{B}_s^0 \to D_s^{*+} D_s^-$: the decays of neutral B mesons to two charged D mesons can form a background to the signal decay if a low
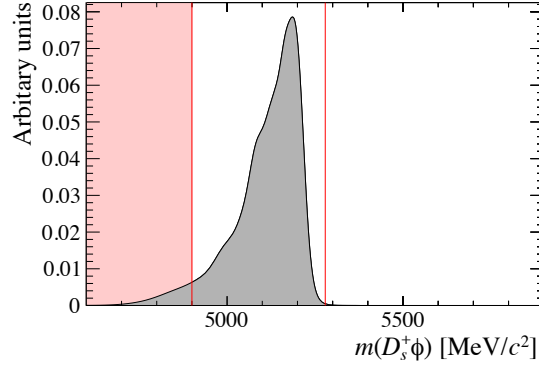
Figure 6.14: Partially reconstructed mass PDFs determined from samples of $\overline{B}_s^0 \to D_s^+ K^- K^{*0}$ simulations processed with the same reconstruction and selection as the signal decays. The B^+ meson mass is indicated by a vertical red line. The are below $4900 \, \text{MeV}/c^2$ is not included in the fit range, but included for reference. The PDF colours follow the same convention used in the final fit plots shown in Fig. 6.21.
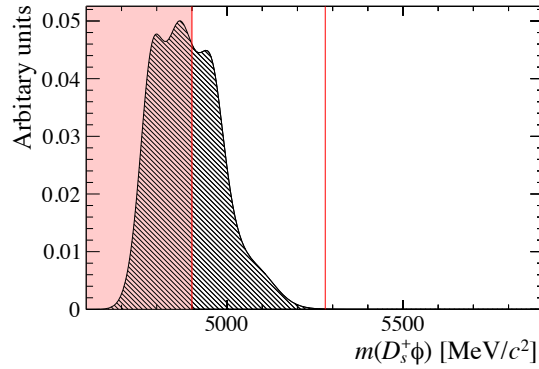
Figure 6.15: Partially reconstructed mass PDFs determined from samples of $\overline{B}_s^0 \to D_s^{*+} K^- K^{*0}$ simulations processed with the same reconstruction and selection as the signal decays. The B^+ meson mass is indicated by a vertical red line. The are below $4900 \, \text{MeV}/c^2$ is not included in the fit range, but included for reference. The PDF colours follow the same convention used in the final fit plots shown in Fig. 6.21.

momentum pion is not reconstructed in one of the charm meson decays. In the case of $\overline{B}_s^0 \to D_s^{*+} D_s^-$ decays an additional neutral particle is not reconstructed in the decay of the D_s^{*+} meson. The PDFs for these decays are determined from simulated decays that have been processed with the same reconstruction and selection steps as the signal. For $\overline{B}_s^0 \to D_s^+ D_s^-$ and $\overline{B}_s^0 \to D_s^{*+} D_s^-$ decays, the PDF is creating using a kernel estimation using the ROOFIT `RooKeysPDF` implementation, shown in Fig.6.16. Due to the similarities between $\overline{B}_s^0 \to D_s^+ D_s^-$

and $\overline{B}^0 \to D_s^+ D^-$ decays, the PDF for the latter is created by using the kernel estimation for $\overline{B}_s^0 \to D_s^+ D_s^-$ and shifting the shape down in mass by $40\,\mathrm{MeV}/c^2$ to account for the kinematic differences.

The branching fractions for these three decays are measured to be $\mathcal{B}(\overline{B}^0 \to D_s^+ D^-) = (7.2 \pm 0.8) \times 10^{-3}$, $\mathcal{B}(\overline{B}_s^0 \to D_s^+ D_s^-) = (4.4 \pm 0.5) \times 10^{-3}$ and $\mathcal{B}(\overline{B}_s^0 \to D_s^{*+} D_s^-) = (1.37 \pm 0.16)\%$. Therefore the relative contributions for these three processes are fixed using these branching fractions, estimates of the relative efficiencies for each mode and the production fraction of $\overline{B}_s^0$ mesons relative to $\overline{B}^0$ mesons, f_s/f_d. This helps to add stability to the fit.
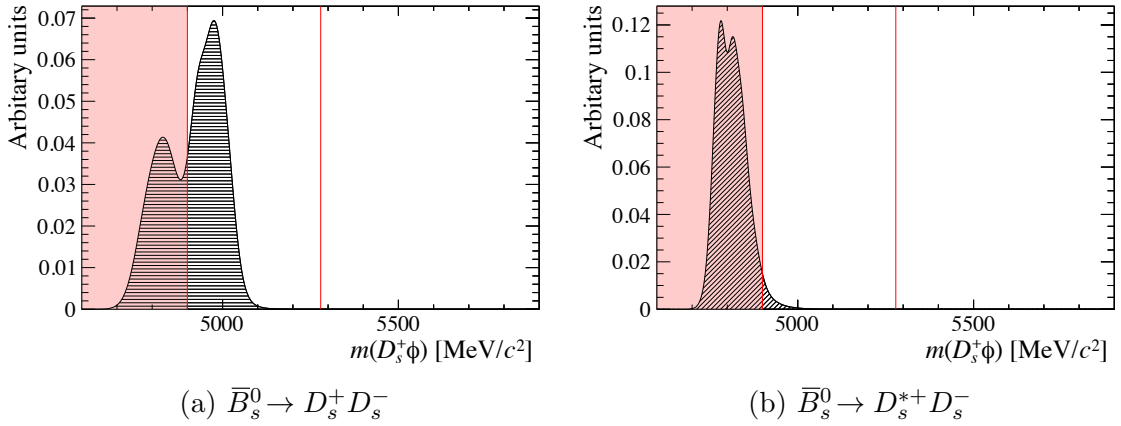
(a) $\overline{B}_s^0 \to D_s^+ D_s^-$ (b) $\overline{B}_s^0 \to D_s^{*+} D_s^-$

Figure 6.16: Partially reconstructed mass PDFs determined from samples of $\overline{B}_s^0 \to D_s^+ D_s^-$ and $\overline{B}_s^0 \to D_s^{*+} D_s^-$ simulations processed with the same reconstruction and selection as the signal decays. The B^+ meson mass is indicated by a vertical red line. The are below $4900\,\mathrm{MeV}/c^2$ is not included in the fit range, but included for reference. The PDF colours follow the same convention used in the final fit plots shown in Fig. 6.21.

To account for difference between the simulation and data samples all PDFs determined using the kernel estimation method are convolved with an additional Gaussian. The mean position of this Gaussian is given by δ, the same offset used for the analytically described partially decays. The width of the Gaussian is increased to account for the difference in resolution between simulation and data.

6.2.3 Combinatorial backgrounds

Combinations of unrelated tracks form a background to the signal decays and extend across the entire fitted B^+ meson mass range. This component is parametrised with a decaying exponential function

$$f(m|c) = e^{-m \times c}, \tag{6.8}$$

where the parameter c controls the slope and m is the observable B^+ meson invariant mass. The combinatorial background is the dominant background under the signal decays, therefore it is important to accurately extrapolate distribution from the high and low mass ranges to the signal range.

The yields of combinatorial background is allowed to vary freely in each of the separate simultaneous fit categories. To improve the stability of the fit, the single slope parameter, c, is shared between all of the categories. In categories with lower statistics, for example the $D_s^+ \to K^+\pi^-\pi^+$ decay mode, it would be possible for the partially reconstructed backgrounds to be incorrectly assigned to the combinatorial component. This can happen if the exponential slope for that mode gets to large, causing it to 'kick up' at low invariant masses. This may bias the signal yield in this mode. Fixing the slope parameter to be the same between the different D_s^+ modes and between the signal and normalisation channel allows the lower statistics modes to benefit from those with higher statistics.

6.3 Free and constrained parameters

The fit to $B^+ \to D_s^+\phi$ and $B^+ \to D_s^+\overline{D}^0$ decays contains a total of 54 free parameters. These are broadly divided into four different categories detailed here.

6.3.1 Parameter of interest

The parameter of interest (POI) is the branching fraction for $B^+ \to D_s^+\phi$ decays. This is determined directly in the fit to the signal and normalisation decays. The

same single branching fraction is used for all of the D_s^+ decays modes. The yield of signal candidates is calculated using the equation

$$N(B^+ \to D_s^+ \phi) = k_{D_s^+} \times \mathcal{B}(B^+ \to D_s^+ \phi) \times N(B^+ \to D_s^+ \overline{D}^0), \qquad (6.9)$$

where $k_{D_s^+}$ is a constant defined for each D_s^+ decay mode

$$k_{D_s^+} = \frac{\epsilon(B^+ \to D_s^+ \overline{D}^0)}{\epsilon(B^+ \to D_s^+ \phi)} \times \frac{\mathcal{B}(\phi \to K^+ K^-)}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)\mathcal{B}(\overline{D}^0 \to K^+ K^-)}. \qquad (6.10)$$

Here, the efficiencies ϵ are calculated for the specific D_s^+ decay mode in question and external measurements are used for the additional branching fractions. The yield $N(B^+ \to D_s^+ \phi)$ is the total signal yield in all four $m(K^+ K^-)$ and $\cos \theta_K$ categories. The yields in each category are constrained to be in the ratios previously listed in Table 6.1, *i.e.* the yield in the $|m(K^+ K^-) < 10 \, \mathrm{MeV}/c^2$ and $|\cos \theta_K| > 0.4$ category is $0.82 \times N(B^+ \to D_s^+ \phi)$.

6.3.2 Shape parameters

As detailed in Sec 6.2, the fit included eight parameters governing the shapes of various PDFs.

- The combinatorial background PDF is controlled by a single slope parameter c for all categories.

- The mean B^+ mass for the signal and normalisation mode is free to vary in the fit. The same value is used for all $B^+ \to D_s^+ \phi$ and $B^+ \to D_s^+ \overline{D}^0$ PDFs.

- The invariant mass offset δ shared between all partially reconstructed background PDFs is allowed to freely vary in the fit. The same value is used for the signal and normalisation modes.

- The relative heights of the two peaks in partially reconstructed $B^+ \to D_s^{*+} \phi$, $B^+ \to D_s^{*+} \overline{D}^0$ and $B^+ \to D_s^+ \overline{D}^{*0}$ decays, ξ, is allowed to vary. The same value is used for all modes.

- The smaller CB width parameter, σ_1, is allowed to vary for each of the D_s^+ decay modes, leading to four more free parameters.

6.3.3 Yields

The fit model contains a total of 36 free parameters that are yields.

- The total yield of normalisation decays in each D_s^+ decay mode is left floating in the fit. These variables are the sum of the yields in the two helicity categories, $|\cos\theta_K| < 0.4$ and $|\cos\theta_K| > 0.4$. This results in four free parameters.

- The yield of certain partially reconstructed decays in the normalisation channel is left floating for each D_s^+ decay mode independently. This yield is defined to be $N(B^+ \to D_s^+ \overline{D}^{*0}) + N(B^+ \to D_s^{*+} \overline{D}^0)$. The yield of the doubly excited $B^+ \to D_s^{*+} \overline{D}^0$ is not included in this total. This results in four free parameters.

- The yield of some partially reconstructed decays in the signal mode is left free in the fit. This is defined to be $N(D_s^{*+}\phi) + N(D_s^+ K^- K^{*0}) + N(D_s^{*+} K^- K^{*0})$. This leads to four free parameters.

- All yields of combinatorial backgrounds are left free in the fit. This results in 24 free parameters.

6.3.4 Fractions

The majority of the partially reconstructed backgrounds do not have unconstrained yields in each of the 24 simultaneous fit categories. Instead, a single partially reconstructed background yield per D_s^+ decay mode is left free in the fit. The relative contributions of other partially backgrounds are determined by floating fraction parameters multiplied by this yield. This allows the different background contributions to vary relative to one another, but keeps the ratios of the relative contributions the same across the different D_s^+ decay modes.

There are four free parameters controlling the backgrounds contributions in the signal mode:

1. Ratio of yields of $B^+ \to D_s^+ K^+ K^-$ decays to $B^+ \to D_s^+ \overline{D}^0$ decays

$$\frac{N(B^+ \to D_s^+ K^+ K^-)}{N(B^+ \to D_s^+ \overline{D}^0)}. \tag{6.11}$$

2. The fraction of $\overline{B}_s^0 \to D_s^+ K^- K^{*0}$ decays to the total of $\overline{B}_s^0 \to D_s^{(*)+} K^- K^{*0}$ decays

$$\frac{N(\overline{B}_s^0 \to D_s^+ K^- K^{*0})}{N(\overline{B}_s^0 \to D_s^{*+} K^- K^{*0}) + N(\overline{B}_s^0 \to D_s^+ K^- K^{*0})}. \tag{6.12}$$

3. The fraction of $B^+ \to D_s^{*+} \phi$ decays in the total of these and $\overline{B}_s^0 \to D_s^{(*)+} K^- K^{*0}$ decays

$$\frac{N(B^+ \to D_s^{*+} \phi)}{N(B^+ \to D_s^{*+} \phi) + N(\overline{B}_s^0 \to D_s^{*+} K^- K^{*0}) + N(\overline{B}_s^0 \to D_s^+ K^- K^{*0})} \tag{6.13}$$

4. The ratio of yields of the partially reconstructed $B_s^0 \to D_s^+ D_s^-$, $B_s^0 \to D_s^+ D_s^{*-}$ and $B^0 \to D_s^+ D^-$ decays to $B^+ \to D_s^{*+} \phi$ and $\overline{B}_s^0 \to D_s^{(*)+} K^- K^{*0}$ decays

$$\frac{N(B_s^0 \to D_s^+ D_s^-) + N(B_s^0 \to D_s^+ D_s^{*-}) + N(B^0 \to D_s^+ D^-)}{N(B^+ \to D_s^{*+} \phi) + N(\overline{B}_s^0 \to D_s^{*+} K^- K^{*0}) + N(\overline{B}_s^0 \to D_s^+ K^- K^{*0})} \tag{6.14}$$

Another five free parameters control the fractions of events in different categories for the normalisation mode:

1. Fraction of partially reconstructed $B^+ \to D_s^{*+} \overline{D}^0$ decays in the total of $B^+ \to D_s^{*+} \overline{D}^0$ and $B^+ \to D_s^+ \overline{D}^{*0}$ decays in the $|\cos\theta_K| > 0.4$ category

$$\left(\frac{N(B^+ \to D_s^{*+} \overline{D}^0)}{N(B^+ \to D_s^{*+} \overline{D}^0) + N(B^+ \to D_s^+ \overline{D}^{*0})}\right)_{|\cos\theta_K| > 0.4} \tag{6.15}$$

2. Fraction of partially reconstructed $B^+ \to D_s^{*+} \overline{D}^0$ decays in the total of $B^+ \to D_s^{*+} \overline{D}^0$ and $B^+ \to D_s^+ \overline{D}^{*0}$ decays in the alternative $|\cos\theta_K| < 0.4$ category

$$\left(\frac{N(B^+ \to D_s^{*+} \overline{D}^0)}{N(B^+ \to D_s^{*+} \overline{D}^0) + N(B^+ \to D_s^+ \overline{D}^{*0})}\right)_{|\cos\theta_K| < 0.4} \tag{6.16}$$

3. Fraction of partially reconstructed backgrounds in the $|\cos\theta_K| > 0.4$ category

$$\frac{(N(B^+ \to D_s^{*+} \overline{D}^0) + N(B^+ \to D_s^+ \overline{D}^{*0}))_{|\cos\theta_K| > 0.4}}{(N(B^+ \to D_s^{*+} \overline{D}^0) + N(B^+ \to D_s^+ \overline{D}^{*0}))_{\text{Total}}} \tag{6.17}$$

4. The ratio of $B^+ \to D_s^{*+} \overline{D}^{*0}$ decays to the total of $B^+ \to D_s^{*+} \overline{D}^0$ and $B^+ \to D_s^+ \overline{D}^{*0}$ decays

$$\frac{N(B^+ \to D_s^{*+} \overline{D}^{*0})}{N(B^+ \to D_s^{*+} \overline{D}^0) + N(B^+ \to D_s^+ \overline{D}^{*0})} \tag{6.18}$$

5. Fraction of fully reconstructed normalisation decays in the $|\cos\theta_K| > 0.4$ category

$$\frac{N(B^+ \to D_s^+ \overline{D}^0)_{|\cos\theta_K|>0.4}}{N(B^+ \to D_s^+ \overline{D}^0)_{\text{Total}}}. \tag{6.19}$$

This parameter is included as a cross check. The normalisation decays have a flat distribution in $\cos\theta_K$ so the paremeter would be expected to be around 0.6.

6.4   Fit validation

The simultaneous fitting framework is validated by generating pseudo-experiments (also referred to as *toys*). The total fit model PDF is randomly sampled to create a simulation sample with the same number of candidates as the nominal fit. These are then fitted using the same fit model, determining the best estimate and uncertainty of each parameter. The parameter values used to generate the pseudo-experiments are chosen to be the final parameter values as determined in a fit to data. This is known as the plug-in method [37]. The fitted value and associated error is used to determine the pull of each parameter of interest. As the errors are determined asymmetrically using Minos the is defined conditionally to incorporate the appropriate error

$$g_{\text{pull}} = \begin{cases} \dfrac{x_{\text{gen}} - x_{\text{fit}}}{\sigma_+}, & \text{if } x_{\text{fit}} < x_{\text{gen}} \\ \dfrac{x_{\text{fit}} - x_{\text{gen}}}{\sigma_-}, & \text{otherwise}, \end{cases} \tag{6.20}$$

where x_{fit} and x_{gen} are the fitted and generated values of the variable, and σ_+ and σ_- are the high and low asymmetric errors.

The distributions of the values, errors and pulls for the yields of the normalisation decay in each of the D_s^+ decays are shown in Fig. 6.17. The mean and widths are determined using simple fits to the pull distributions. The results and PDFs for these fits are overlaid on the distributions. The normalisation yield means and widths are found to be within 2σ of zero and one respectively.

Similarly, the distributions of the signal yield, error and pull for each of the different D_s^+ decay modes are shown in Fig. 6.19. The pull means and widths are all within 2σ of zero and one respectively.

(a) $D_s^+ \to \phi \pi^+$

(b) $D_s^+ \to K^+ K^- \pi^+$

(c) $D_s^+ \to \pi^+ \pi^- \pi^+$

(d) $D_s^+ \to K^+ \pi^- \pi^+$
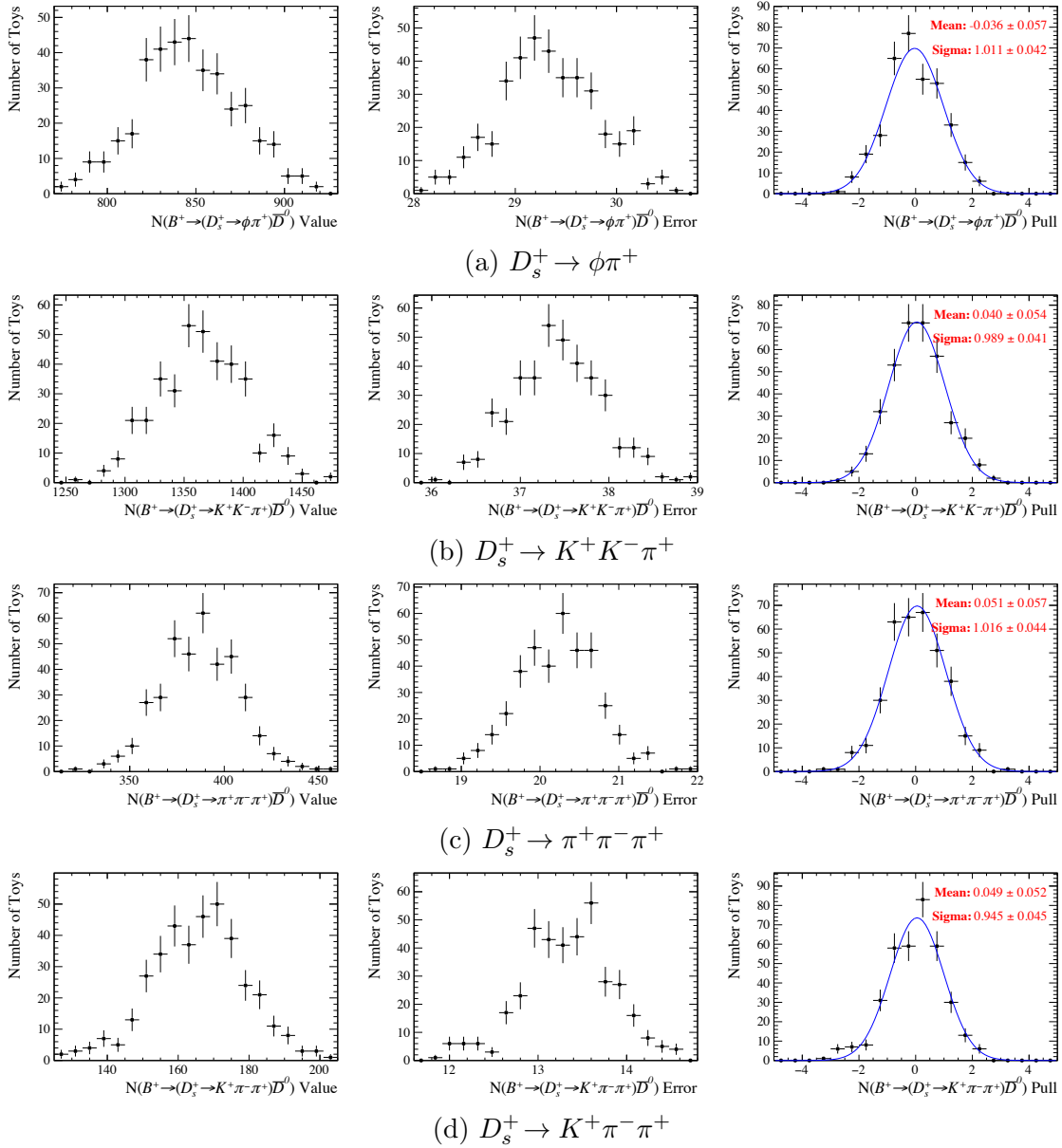
Figure 6.17: The yield, error and pull distributions for the normalisation channel.

These distributions were generated using pseudo-experiments in which the yields for each D_s^+ mode were free variables rather than being determined using a single free branching fraction parameter.

The model is also studied using a single branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)$ that determines the yields in each signal mode. This parameter is calculated directly in the fit so it can be assessed for any possible bias. The distribution of the measured values, uncertainties and pulls are shown in Fig. 6.19.

(a) $D_s^+ \to \phi \pi^+$

(b) $D_s^+ \to K^+ K^- \pi^+$

(c) $D_s^+ \to \pi^+ \pi^- \pi^+$

(d) $D_s^+ \to K^+ \pi^- \pi^+$
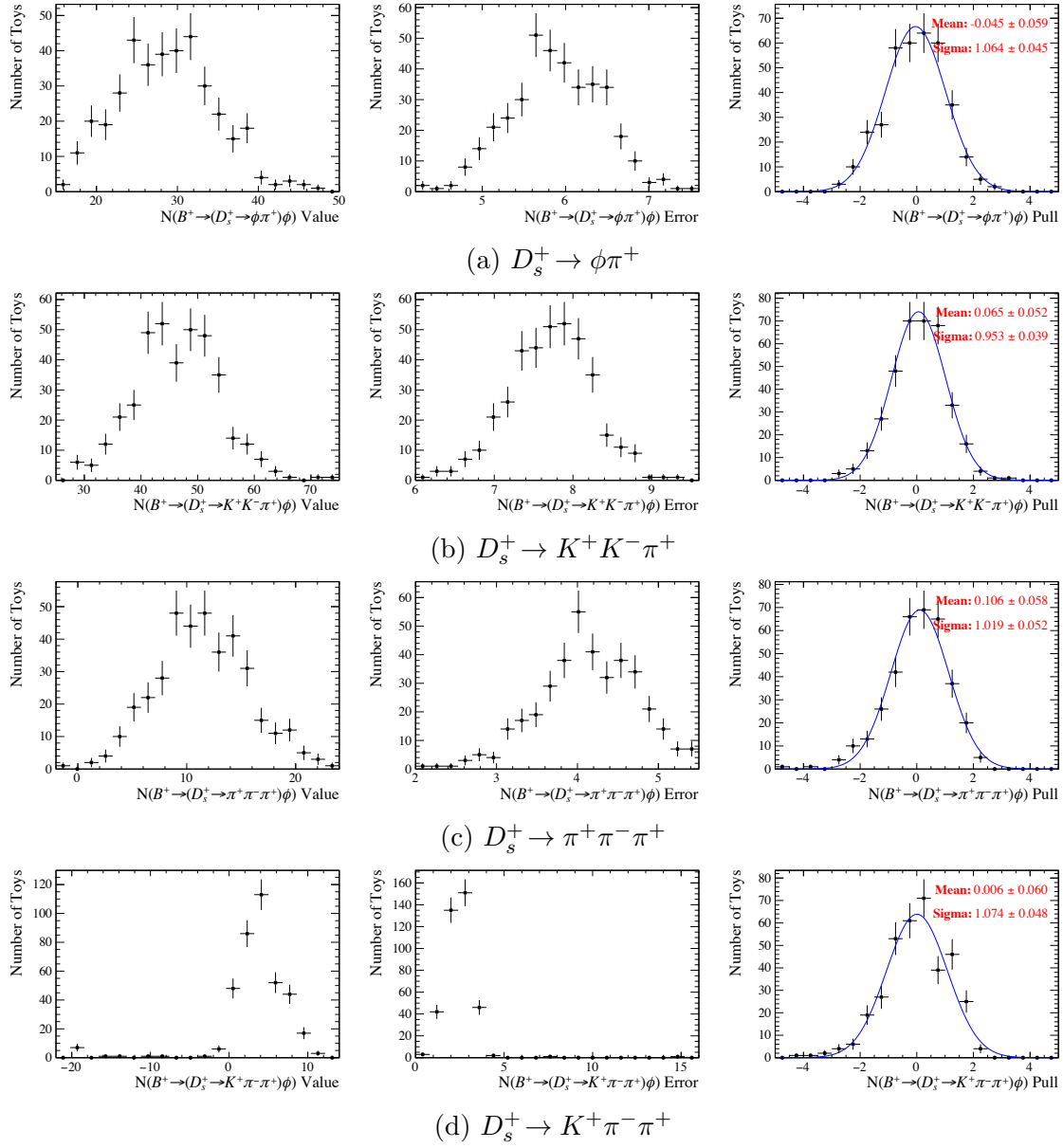
Figure 6.18: The yield, error and pull distributions for the signal channel.

The pull mean and width are both within 2σ of zero and one respectively. As no significant biases are observed in the pulls of any of the yields or branching fraction no corrections are applied to the values determined fit.
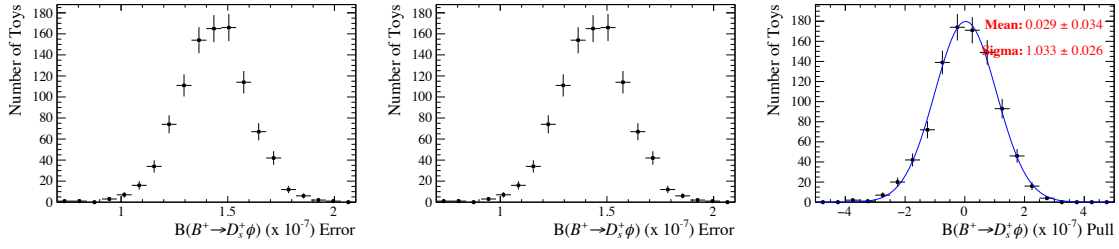
Figure 6.19: The branching fraction, error and pull distributions for the full simultaneous fit. The branching fraction is generated assuming the value obtained in the fit to data.

6.5 Normalisation and signal fits

The result of the simultaneous fit to the signal and normalisation decays is shown in Figs. 6.20 and 6.21. These figures show the distribution of candidates passing all selection requirements along with the total PDF model resulting from the minimisation of the log-likelihood. The four D_s^+ decay mode categories have been merged into a single distribution. The fits in each different D_s decay mode category can be found in Appendix A.
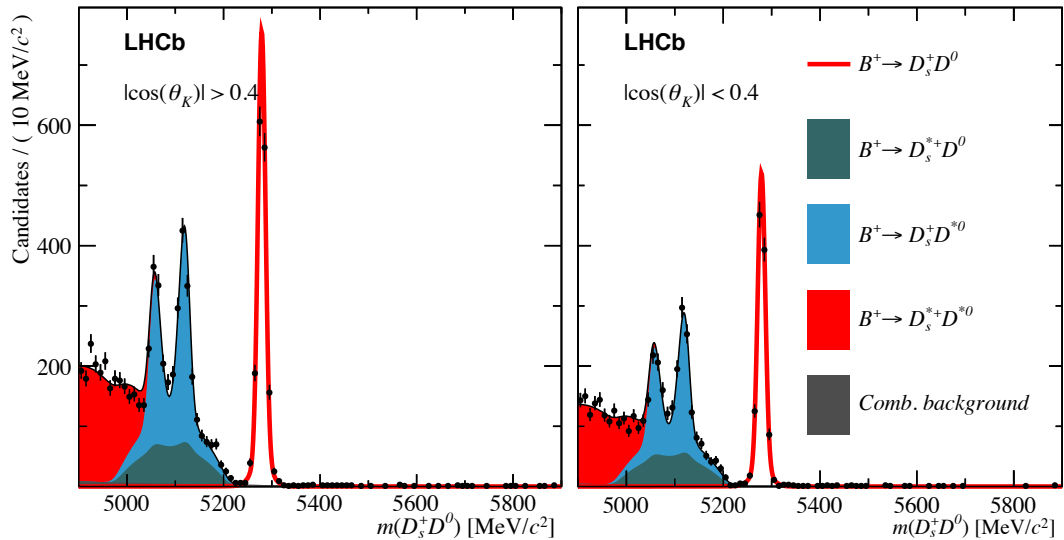
Figure 6.20: Invariant mass fits to $B^+ \to D_s^+ \overline{D}^0$ candidates

The high purity of the normalisation mode reconstruction can be seen in Fig. 6.20, as the contribution from the combinatorial background shape is very small. Additionally, the double-peaked structure of the partially reconstructed $B^+ \to D_s^{*+} \overline{D}^0$ and

$B^+ \to D_s^+ \overline{D}^{*0}$ decays is clearly visible.
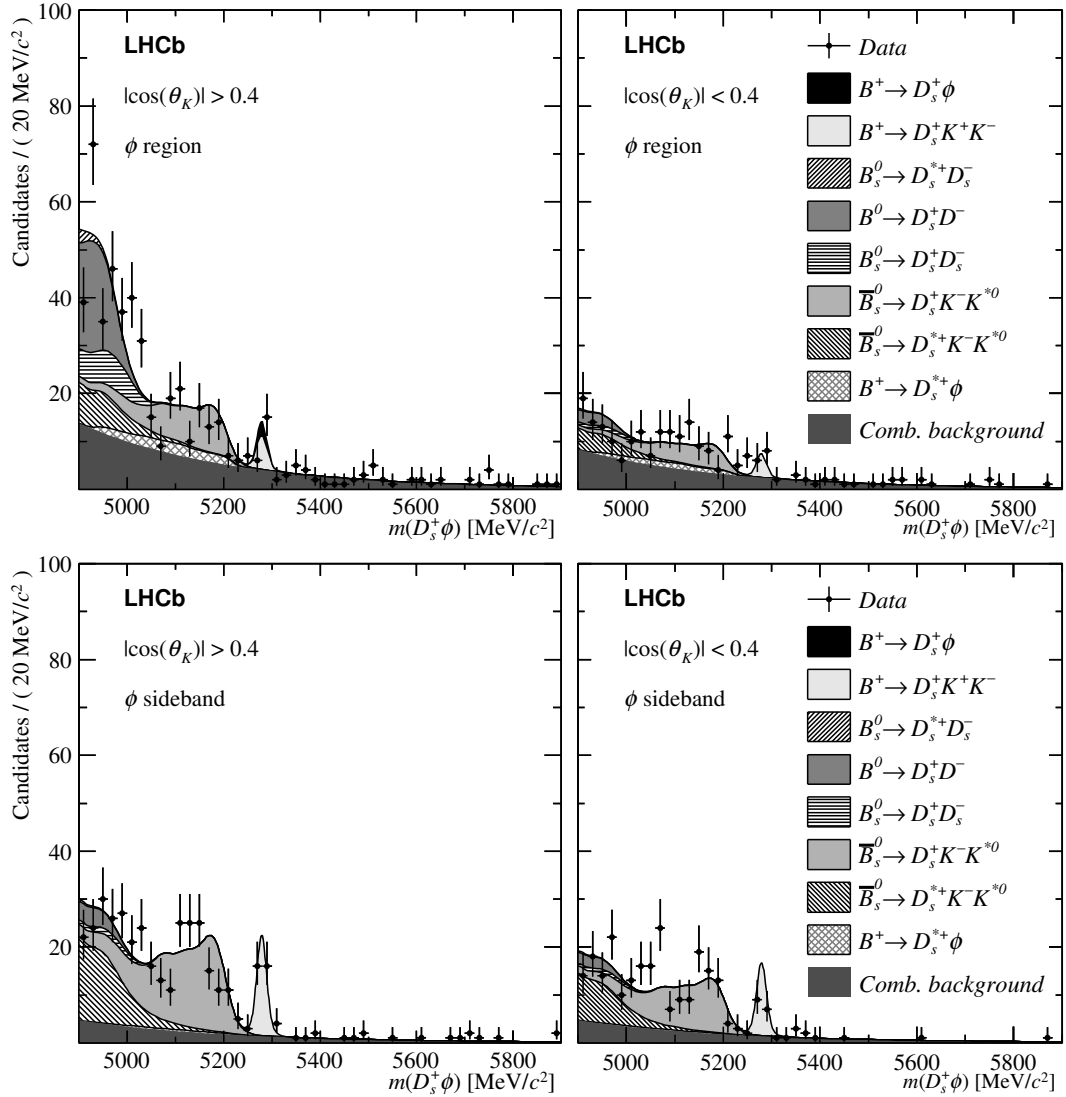
Figure 6.21: Invariant mass fits to $B^+ \to D_s^+ \phi$ candidates

The fit to the signal mode is shown in Fig. 6.21. This figure includes the $B^+ \to D_s^+ \phi$ candidates in both the $|m(K^+K^-)| < 10 \, \text{MeV}/c^2$ (ϕ-region) and $10 < |m(K^+K^-)| < 40 \, \text{MeV}/c^2$ (ϕ-sideband) categories. These are further split into the two $\cos\theta_K$ categories in an analogous way to the normalisation decays.

The final values of the parameters as determined by the fit are tabulated in Tab. 6.5, including the branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)$. This parameter includes a correction to account for the efficiencies of the signal and normalisation modes as discussed in Sec. 6.6.

Type	Parameter	Fit result
POI	Branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)(\times 10^{-7})$	$1.17^{+1.56}_{-1.37}$
Shape	Combinatorial slope c	$(-3.3^{+0.4}_{-0.4}) \times 10^{-3}$
	Mean B^+ mass (MeV/c^2)	$5279.10^{+0.17}_{-0.17}$
	Mass offset δ (MeV/c^2)	$-1.975^{+0.362}_{-0.362}$
	Relative heights ξ	$0.68^{+0.05}_{-0.05}$
	σ_1 for $D_s^+ \to K^+ K^- \pi^+$ (MeV/c^2)	$7.61^{+0.18}_{-0.18}$
	σ_1 for $D_s^+ \to K^+ \pi^- \pi^+$ (MeV/c^2)	$8.77^{+0.59}_{-0.55}$
	σ_1 for $D_s^+ \to \phi \pi^+$ (MeV/c^2)	$7.53^{+0.22}_{-0.21}$
	σ_1 for $D_s^+ \to \pi^+ \pi^- \pi^+$ (MeV/c^2)	$8.30^{+0.36}_{-0.34}$
Yields	$N(B^+ \to (D_s^+ \to K^+ K^- \pi^+)\overline{D}^0)$	1324^{+37}_{-36}
	$N(B^+ \to (D_s^+ \to K^+ \pi^- \pi^+)\overline{D}^0)$	182^{+14}_{-13}
	$N(B^+ \to (D_s^+ \to \phi \pi^+)\overline{D}^0)$	801^{+29}_{-28}
	$N(B^+ \to (D_s^+ \to \pi^+ \pi^- \pi^+)\overline{D}^0)$	369^{+20}_{-19}
	$N(D_s^{*+}\overline{D}^0 + D_s^+ \overline{D}^{*0})$ in $D_s^+ \to K^+ K^- \pi^+$	2827^{+52}_{-51}
	$N(D_s^{*+}\overline{D}^0 + D_s^+ \overline{D}^{*0})$ in $D_s^+ \to K^+ \pi^- \pi^+$	394^{+20}_{-20}
	$N(D_s^{*+}\overline{D}^0 + D_s^+ \overline{D}^{*0})$ in $D_s^+ \to \phi \pi^+$	1733^{+38}_{-37}
	$N(D_s^{*+}\overline{D}^0 + D_s^+ \overline{D}^{*0})$ in $D_s^+ \to \pi^+ \pi^- \pi^+$	804^{+25}_{-25}
	$N(D_s^{*+}\phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \to K^+ K^- \pi^+$	139^{+19}_{-20}
	$N(D_s^{*+}\phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \to K^+ \pi^- \pi^+$	7^{+4}_{-4}
	$N(D_s^{*+}\phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \to \phi \pi^+$	67^{+10}_{-10}
	$N(D_s^{*+}\phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \to K^+ \pi^- \pi^+$	25^{+7}_{-6}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to K^+ K^- \pi^+$ in H1	112^{+30}_{-25}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to K^+ K^- \pi^+$ in H2	50^{+17}_{-14}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to K^+ \pi^- \pi^+$ in H1	71^{+20}_{-17}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to K^+ \pi^- \pi^+$ in H2	41^{+12}_{-10}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to \phi \pi^+$ in H1	31^{+14}_{-11}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to \phi \pi^+$ in H2	12^{+9}_{-6}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to \pi^+ \pi^- \pi^+$ in H1	41^{+14}_{-11}
	$N_{\mathrm{comb}}(D_s^+ \overline{D}^0)\ D_s^+ \to \pi^+ \pi^- \pi^+$ in H2	35^{+14}_{-11}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to K^+ K^- \pi^+$ in H1	107^{+23}_{-20}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to K^+ K^- \pi^+$ in H2	52^{+12}_{-11}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to K^+ \pi^- \pi^+$ in H1	19^{+7}_{-6}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to K^+ \pi^- \pi^+$ in H2	18^{+5}_{-4}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to \phi \pi^+$ in H1	38^{+14}_{-12}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to \phi \pi^+$ in H2	27^{+8}_{-7}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to \pi^+ \pi^- \pi^+$ in H1	42^{+10}_{-9}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ D_s^+ \to \pi^+ \pi^- \pi^+$ in H2	25^{+7}_{-6}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ \phi\text{-sideband } D_s^+ \to K^+ K^- \pi^+$ in H1	33^{+13}_{-10}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ \phi\text{-sideband } D_s^+ \to K^+ K^- \pi^+$ in H2	26^{+11}_{-9}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ \phi\text{-sideband } D_s^+ \to K^+ \pi^- \pi^+$ in H1	18^{+7}_{-6}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ \phi\text{-sideband } D_s^+ \to K^+ \pi^- \pi^+$ in H2	14^{+5}_{-5}
	$N_{\mathrm{comb}}(D_s^+ \phi)\ \phi\text{-sideband } D_s^+ \to \phi \pi^+$ in H1	5^{+6}_{-4}

	$N_{\mathrm{comb}}(D_s^+\phi)$ ϕ-sideband $D_s^+ \to \phi\pi^+$ in H2	12^{+8}_{-7}
	$N_{\mathrm{comb}}(D_s^+\phi)$ ϕ-sideband $D_s^+ \to \pi^+\pi^-\pi^+$ in H1	16^{+8}_{-6}
	$N_{\mathrm{comb}}(D_s^+\phi)$ ϕ-sideband $D_s^+ \to \pi^+\pi^-\pi^+$ in H2	14^{+8}_{-8}
Fractions	Ratio of $D_s^+ K^+ K^-$ to $D_s^+ \overline{D}^0$	$0.024^{+0.004}_{-0.004}$
	Fraction of $D_s^+ K^- K^{*0}$ in $(D_s^{*+} K^- K^{*0} + D_s^+ K^- K^{*0})$	$0.664^{+0.046}_{-0.044}$
	Fraction of $D_s^{*+}\phi$ in $(D_s^{(*)+} K^- K^{*0} + D_s^{*+}\phi)$	$0.172^{+0.103}_{-0.135}$
	Ratio of $D_s^+ D_{(s)}^{(*)-}$ to $(D_s^{*+}\phi + D_s^{(*)+} K^- K^{*0})$	$0.567^{+0.129}_{-0.107}$
	Fraction of $D_s^{*+}\overline{D}^0$ in $(D_s^{*+}\overline{D}^0 + D_s^+\overline{D}^{*0})$ in H1	$0.302^{+0.031}_{-0.031}$
	Fraction of $D_s^{*+}\overline{D}^0$ in $(D_s^{*+}\overline{D}^0 + D_s^+\overline{D}^{*0})$ in H2	$0.342^{+0.037}_{-0.037}$
	Fraction of normalisation part reco in H1	$0.593^{+0.005}_{-0.005}$
	Fraction of normalisation peak in H1	$0.592^{+0.010}_{-0.010}$
	Ratio of $D_s^{*+}\overline{D}^{*0}$ to $(D_s^{*+}\overline{D}^0 + D_s^+\overline{D}^{*0})$	$0.607^{+0.016}_{-0.015}$

Table 6.5: The fit result with final values of all floating variables used in the fit model. Here H1 and H2 represent the $|\cos\theta_K| > 0.4$ and $|\cos\theta_K| < 0.4$ categories respectively.

6.6 Efficiency corrections

The branching fraction for $B^+ \to D_s^+\phi$ decays is determined by correcting the yields for the signal and normalisation channels by their respective efficiencies. In determining the efficiencies both the $B^+ \to D_s^+\phi$ and $B^+ \to D_s^+\overline{D}^0$ decays are assumed to be pseudo-two-body decays in which variations as a function of phase-space are negligible. This means the relative efficiency can be calculated as a simple ratio for each D_s^+ decay mode, simplifying the correction considerably.

The efficiencies for each stage of selection are either determined from the appropriate simulation samples for the signal and normalisation decays, or from dedicated calibration data samples for the efficiencies involving particle identification requirements. Each set of efficiencies are determined with respect to the previous one, such that the total efficiency is given by the product of each of them.

6.6.1 Efficiencies from simulation

The efficiencies for each step of selection are listed separately for the different D_s^+ decay modes in Table 6.6. These steps are closely related to the stages previously described in Chapter 4, however more specific details of what is included in each step

is described here. The efficiencies are calculated separately for the different years of data taking, however they are combined here, weighting according to the relative contributions to the total data set.

Requirement	D_s^+ mode	$\epsilon(B^+ \to D_s^+ \overline{D}^0)$	$\epsilon(B^+ \to D_s^+ \phi)$	Ratio
Acceptance	$K^+ K^- \pi^+$	17.42 ± 0.03	18.35 ± 0.04	0.950 ± 0.003
	$K^+ \pi^- \pi^+$	16.53 ± 0.04	17.36 ± 0.04	0.952 ± 0.003
	$\pi^+ \pi^- \pi^+$	15.91 ± 0.03	16.77 ± 0.04	0.949 ± 0.003
Reconstruction	$K^+ K^- \pi^+$	2.11 ± 0.02	1.96 ± 0.02	1.078 ± 0.013
	$K^+ \pi^- \pi^+$	2.27 ± 0.02	2.06 ± 0.02	1.095 ± 0.014
	$\pi^+ \pi^- \pi^+$	2.36 ± 0.02	2.13 ± 0.02	1.105 ± 0.014
Trigger	$K^+ K^- \pi^+$	93.3 ± 0.2	93.1 ± 0.2	1.003 ± 0.003
	$K^+ \pi^- \pi^+$	95.1 ± 0.2	93.5 ± 0.2	1.017 ± 0.003
	$\pi^+ \pi^- \pi^+$	95.3 ± 0.2	93.7 ± 0.2	1.017 ± 0.003
Mass window	$K^+ K^- \pi^+$	96.2 ± 0.2	94.1 ± 0.2	1.022 ± 0.003
	$K^+ \pi^- \pi^+$	94.9 ± 0.2	94.1 ± 0.2	1.008 ± 0.003
	$\pi^+ \pi^- \pi^+$	93.7 ± 0.2	92.4 ± 0.2	1.015 ± 0.004
Vetoes	$K^+ K^- \pi^+$	99.9 ± 0.0	95.2 ± 0.2	1.049 ± 0.002
	$K^+ \pi^- \pi^+$	99.9 ± 0.0	92.4 ± 0.3	1.080 ± 0.004
	$\pi^+ \pi^- \pi^+$	99.9 ± 0.0	90.3 ± 0.3	1.105 ± 0.004
Charmless	$K^+ K^- \pi^+$	72.6 ± 0.4	100.0 ± 0.0	0.726 ± 0.004
	$K^+ \pi^- \pi^+$	65.8 ± 0.4	63.0 ± 0.5	1.044 ± 0.011
	$\pi^+ \pi^- \pi^+$	66.0 ± 0.4	82.9 ± 0.4	0.796 ± 0.006
χ^2_{IP}	$K^+ K^- \pi^+$	100.0 ± 0.0	96.2 ± 0.2	1.039 ± 0.002
	$K^+ \pi^- \pi^+$	100.0 ± 0.0	95.7 ± 0.3	1.045 ± 0.003
	$\pi^+ \pi^- \pi^+$	100.0 ± 0.0	95.9 ± 0.3	1.043 ± 0.003

Table 6.6: Efficiencies (in %) determined from simulation samples for signal and normalisation decays and the ratio $\epsilon(B^+ \to D_s^+ \overline{D}^0)/\epsilon(B^+ \to D_s^+ \phi)$ for each D_s^+ decay mode. The errors are statistical.

Acceptance: this accounts for the fraction of generated decays in which all five final state tracks end up within the LHCb detector's acceptance. For all D_s^+ decay modes the signal decay has a slightly higher efficiency than the normalisation channel. This may because of the kinematics of the $\phi \to K^+ K^-$ decay in which the two kaon tracks are typically close to one another and therefore more likely to both be in the acceptance.

Reconstruction: this efficiency determines the fraction of decays in which all tracks are well reconstructed and combined into a suitable candidate. The candidate

must pass all requirements outlined in the *Stripping Line* for the specific decay. The reconstruction selections used for both the signal and normalisation decays explicitly require that the event in which the candidate was found had fired the trigger. Therefore this efficiency includes some, but not all, of the trigger efficiency. This results in small efficiencies around 2% for each mode. Here, the efficiencies are slightly larger for the normalisation channel than the signal. This may be because the kaons from the ϕ decay are typically lower momentum than those from the $\overline{D}^0$, and therefore less likely to be reconstructed.

Trigger: this efficiency effectively accounts for the likelihood that the candidate passed at least one of the requirements in Sec. 4.2, given a trigger fired in that event. These efficiencies have very high values around 94%.

Mass windows: this represents the efficiency for the candidates to be within the invariant mass windows for the D_s^+ and ϕ or $\overline{D}^0$ mesons. Again, this is slightly high for the normalisation than the signal. The ϕ meson invariant mass peak a fairly long tail extending to higher invariant masses that may be the cause.

Vetoes: the efficiency of the kinematic vetoes described in Sec.4.3.5 is included in this quantity. The misidentified D and Λ_c^+ hadron vetoes target the D_s^+ meson, present in both the signal and normalisation decays. The relative efficiency is assumed to be one for these specific vetoes therefore not included in this efficiency. The systematic uncertainty resulting from this assumption is discussed in Sec.6.7. Most of the kinematic vetoes are only applied to the signal mode, hence why the normalisation channel efficiencies are almost 100%.

Charmless: the requirements applied to the flight distance significance of the D_s^+ meson is tuned differently for each D_s^+ decay and signal and normalisation. As a result, the efficiencies have large variations between the different modes.

$\chi_{\mathbf{IP}}^2$: the efficiency of the χ_{IP}^2 requirements for the B^+ and D_s^+ candidates as detailed in Sec.4.3.8 are included in this quantity. These are only applied to the signal mode therefore the normalisation mode is 100% efficient.

6.6.2 Efficiencies requiring calibration samples

The efficiencies of the particle identification and MVA requirements are both determined using input from dedicated calibration samples as the distributions are known to be poorly represented in simulations. The method used here differs from that already described in Sec. 5.6.2 as the $B^+ \to D_s^+ \phi$ decay can be assumed to be a pseudo-two-body decay in which phase-space dependent efficiencies are not required.

Requirement	D_s^+ mode	$\epsilon(B^+ \to D_s^+ \overline{D}^0)$	$\epsilon(B^+ \to D_s^+ \phi)$	Ratio
PID	$K^+K^-\pi^+$	88.1 ± 0.1	89.9 ± 0.0	0.980 ± 0.001
	$K^+\pi^-\pi^+$	86.7 ± 0.2	88.8 ± 0.1	0.977 ± 0.002
	$\pi^+\pi^-\pi^+$	85.3 ± 0.1	87.0 ± 0.0	0.980 ± 0.001
MVA	$K^+K^-\pi^+$	53.2 ± 0.4	57.1 ± 0.4	0.932 ± 0.010
	$K^+\pi^-\pi^+$	42.9 ± 0.4	46.0 ± 0.5	0.932 ± 0.013
	$\pi^+\pi^-\pi^+$	46.3 ± 0.4	49.7 ± 0.4	0.933 ± 0.012

Table 6.7: Efficiencies (in %) determined from the relevant calibration and validation samples using input from simulation and the ratio $\epsilon(B^+ \to D_s^+ \overline{D}^0)/\epsilon(B^+ \to D_s^+ \phi)$ for each D_s^+ decay mode. The errors are statistical.

PID efficiency

The efficiency of the particle identification requirements described in Sec. 4.3.2 are determined using a package called PIDCALIB [38]. This uses calibrations samples for the different particle species to determine the fraction of candidates passing the various PID variable requirements. The samples are background-subtracted to isolate the distributions of the PID variables for the tracks of interest. The calibration samples for both K^+ and π^+ mesons are collected from a sample of $D^{*+} \to (D^0 \to K^+\pi^-)\pi^+$ decays, using the decay products of the D^0 decay.

Unlike in the search for $B^+ \to D_s^+ K^+ K^-$ decays, the PID variable distributions of the calibration samples are parametrised using a binned approach, with three characterising variables; the momentum p, pseudo-rapidity η, and total number of tracks n_{Track}. Input from the signal and normalisation simulation samples is used to determine per-candidate efficiencies. The characteristics $(p, \eta, n_{\text{Track}})$ of each simulated decay is used to find the corresponding calibration PID variable distribution. The

per-candidate efficiency is calculated by integrating this distribution above the PID requirement. The total efficiency is given by the average of the per-candidate efficiencies. The efficiencies of the PID requirements on the signal and normalisation for the different D_s^+ decay modes are listed in Table 6.7.

MVA efficiency

The efficiency of the MVA requirements is determined using the method outlined in Sec. 4.3.7. The validation samples of $\overline{B}_s^0 \to D_s^+ \pi^-$ and $B_s^0 \to J/\psi \phi$ decays are binned in transverse momentum p_T and flight distance significance χ^2_FD. The background-subtracted MVA responses are extracted for each bin. Samples of simulated decays are iterated though, finding the p_T and χ^2_FD values for each candidate and calculating the MVA cut efficiency from the corresponding validation samples. The total efficiency for each candidate is the product of the D_s^+ and ϕ MVA selection efficiencies. The total efficiency for the whole sample is then given by the sum of the per-candidate efficiencies as listed in Table 6.7.

6.6.3   Total efficiencies

The total relative efficiency between the signal and normalisation decays is determined as the product of each contributing relative efficiency

$$
\begin{aligned}
\epsilon^\mathrm{Tot.} = \epsilon^\mathrm{Accp.} \times \epsilon^\mathrm{Reco.|Accp.} \times \epsilon^\mathrm{Trig.|Reco.} \times \epsilon^\mathrm{Mass.|Trig.} \times \epsilon^\mathrm{Veto.|Mass.} \times \epsilon^\mathrm{FD|Veto.} \\
\times \epsilon^\mathrm{IP|FD} \times \epsilon^\mathrm{PID|IP} \times \epsilon^\mathrm{MVA|PID}, \quad (6.21)
\end{aligned}
$$

where each relative efficiency x is determined with respect to the previous selection step y as $\epsilon^{x|y}$. The total relative efficiencies are listed for each D^+ decay mode in Table 6.8. These values are used as an input in the simultaneous fit to the signal and normalisation modes to correct the yields ratios for each D_s^+ decay mode category.

D_s^+ decay mode	Ratio
$K^+K^-\pi^+$	0.757 ± 0.014
$K^+\pi^-\pi^+$	1.144 ± 0.027
$\pi^+\pi^-\pi^+$	0.907 ± 0.019

Table 6.8: The total relative efficiency $\epsilon(B^+ \to D_s^+ \overline{D}^0)/\epsilon(B^+ \to D_s^+ \phi)$ determined for each D_s^+ decay mode. The values for the individual years that contribute to the data set are weighted according to the size of their relative contribution. The errors are statistical.

6.7 Systematic uncertainties

The sources of systematic uncertainty are broadly similar to those considered in the search for $B^+ \to D_s^+ K^+ K^-$ decays. The more complex fit strategy and model requires most of these systematics to be reassessed and additional source included.

6.7.1 Relative efficiencies

The yields of signal and normalisation decays are corrected by their relative selection efficiencies, calculated separately for each D_s^+ decay mode. A number of sources contribute to the systematic uncertainty in these values.

Simulation statistics: limited simulation samples are used to determine some of the selection efficiencies. A systematic uncertainty of 2% is assigned to account for the bias this might incur.

Particle identification: the efficiency of the PID requirements are calculated using calibration samples within the PIDCALIB package. These are also of a limited size and possibly affected by the choice of binning scheme. To first order these should affect the signal and normalisation decays to the same extent as the same final state is used. However, differences in the kinematics could invalidate this assumption. Therefore a systematic uncertainty of 0.5% per track is assigned to the use of the PIDCALIB efficiencies. This is assumed to be uncorrelated for the five signal and five normalisation tracks, leading to a total uncertainty in the relative efficiency of 2.0%.

Veto efficiency: similar to the search for $B^+ \to D_s^+ K^+ K^-$ the efficiency of the misidentified D and Λ_c^+ hadron veto applied to the D_s^+ meson is assumed to be the same for the signal and normalisation channel. A similar systematic uncertainty of 1.4% is assigned to account for the uncertainty incurred if the relative efficiency were to be fully calculated.

MVA efficiency: unlike the search for $B^+ \to D_s^+ K^+ K^-$ decays, the efficiencies for the MVA requirements are just calculated using the $B_s^0 \to J/\psi\,\phi$ and $\overline{B}_s^0 \to D_s^+ \pi^-$ samples used in the validation of the MVA methods. There is no need account for the phase-space dependence, so the simulation samples have not been corrected. These data-driven efficiencies are affected by the same sources as previously discussed in 5.7.1. The validation samples are binned in four bins of both p_T and χ^2_FD. The choice of binning scheme is varied and the resulting variation in the efficiencies assigned as a systematic uncertainty. The yields of $B_s^0 \to J/\psi\,\phi$ and $\overline{B}_s^0 \to D_s^+ \pi^-$ are limited, which could lead to fluctuations in the efficiencies determined. The quantity $1/\sqrt{N}$ for the smallest sample is assigned as the systematic uncertainty.

Some differences are observed in the distributions of χ^2_IP between simulations and data for the normalisation mode (Fig. 4.18). The simulation samples are re-weighted to match the data distributions and the efficiencies recalculated. The resulting difference is included as a systematic uncertainty.

In training the MVA methods, the distribution of the MVA classifier shows some discrepancies between the training and validation samples at low classifier values. To quantify the effect this may have on the relative efficiencies the samples are swapped and the efficiencies recalculated. The resulting difference is assigned as the associated systematic uncertainty.

The total systematic uncertainty attributed to the relative MVA efficiencies is 6.2%.

6.7.2   Signal and normalisation PDFs

The PDFs used to describe the signal and normalisation channel invariant mass distributions are determined using input from fits to simulated decays. A similar approach to that already described in Sec. 5.7.1 is used to account for the systematic uncertainty that these fixed values contribute. The fit to data is repeated many times, each time the values of the fixed parameters are sampled from Gaussian distributions whose width is given by the associated statistical uncertainty from the fit to simulations. The resulting spread in the branching fraction obtained by re-sampling all fixed parameters simultaneously is 0.036×10^{-7}, assigned as the systematic uncertainty.

6.7.3   Background PDFs

Similarly the background PDFs require inputs determined from simulations. A large number of different properties of the PDFs are varied to determine the systematic uncertainty the fixed values incur. The resulting spread in the branching fraction is assigned as the systematic uncertainty.

The helicity fraction of $B^+ \to D_s^{*+} \phi$ decays is not measured and therefore assumed to be 0.5. This assumption is varied across all allowed values. Additionally, the PDF for both $B^+ \to D_s^{*+} \phi$ decay and the $B^+ \to D_s^{*+} \overline{D}^0$ and $B^+ \to D_s^+ \overline{D}^{*0}$ decays that contribute to the normalisation mode require the kinematic limits a and b to be defined. These values are varied.

The backgrounds to the signal mode require the fraction of decays expected in each $m(K^+ K^-)$ and $\cos \theta_K$ category to be determined from simulations. These are each have an associated statistical uncertainty due to the limited simulation samples, therefore these fractions are varied by sampling a Gaussian whose width is determined by the uncertainty.

In the fit to the normalisation mode the partially reconstructed background $B^+ \to D_s^{*+} \overline{D}^{*0}$ is approximated with a single PDF, rather than eight as required by the combinations of missed particles and helicity components. This function is replaced

with a DCB function with free widths mean and tail parameters. This results in a negligible change in the branching fraction.

The combinatorial background is parametrised with an exponential and constrained to have the same slope in all categories. To determine if these choices lead to systematic uncertainty the fit is repeated with the slope allowed to be different in the different D_s^+ decay mode categories. Additionally an exponential plus constant offset is tried as a parametrisation instead. The resulting changes in the branching fractions are included as systematic uncertainties.

The total systematic uncertainty associated to the background PDFs is determined to be 0.685×10^{-7}

6.7.4 Charmless contribution

A residual yield of charmless and single-charm backgrounds are expected be present in the final dataset as detailed in Table 4.9. The nominal measurement for the branching fraction neglects these contributions. The branching fraction is recalculated assuming the expected yields contribute and the difference assigned as the systematic uncertainty. This is likely to be an overestimation of the possible difference as the charmless candidates are likely to have a wider distribution than the signal decays. This is a result of the D_s^+ mass constraint applied to the candidates when calculating the fitted B^+ mass.

The charmless contributions lead to a systematic uncertainty of 2%.

6.7.5 $B^+ \to D_s^+ K^+ K^-$ model assumption

The search for $B^+ \to D_s^+ \phi$ candidates includes a component for $B^+ \to D_s^+ K^+ K^-$ decays that didn't proceed via a ϕ meson. As detailed in Sec. 6.1.2 various assumptions go into the choice of $B^+ \to D_s^+ K^+ K^-$ model used to determine the fraction of these decays expected in each $m(K^+ K^-)$ and $\cos \theta_K$ category. The final choice of fractions detailed in Table 6.2 are determined with associated uncertainties, calculated by taking the range of fractions in the models thought to be reasonable. The fit to data is performed many times changing these fixed fractions to values taken

from Gaussian distributions whose widths are given by these assigned uncertainties. The resulting spread in the measured branching fraction is used a a proxy for the systematic uncertainty resulting from this choice of model.

An additional systematic uncertainty is included because these fractions are taken directly from simulations produced by LAURA++, rather than full detector simulations. This is calculated by taking the difference in the fractions found for the $B^+ \to D_s^+ \phi$ decay generated using LAURA++ and the full LHCb simulation.

6.7.6 Total systematic uncertainty

The sources of systematic uncertainty are listed in Table 6.9 in decreasing order. The total is also included, calculated by summing the contributions in quadrature. Additionally the uncertainty arising from the externally measured branching fractions is included.

Source of Uncertainty	Systematic Uncertainty	
	Relative	Absolute ($\times 10^{-7}$)
Background PDF parametrisation	-	0.685
Choice of $B^+ \to D_s^+ K^+ K^-$ model	-	0.379
MVA relative efficiency	6.2%	0.072
Signal PDF parametrisation	-	0.036
Charmless contribution	2.0%	0.023
Simulation statistics	2.0%	0.023
PID relative efficiency	2.0%	0.023
Veto relative efficiency	1.4%	0.016
Using LAURA++ rather full sim.	1.1%	0.013
Total		0.788
Normalisation		0.1

Table 6.9: Contributions to the total systematic uncertainty in the search for $B^+ \to D_s^+ \phi$ decays. The contribution from the external measurements of the normalisation channel branching fraction is also included.

6.8 Results

The fit to $B^+ \to D_s^+ \phi$ candidates finds a total yield of $N(B^+ \to D_s^+ \phi) = 5.3 \pm 6.7$, summed across all categories and D_s^+ meson decay modes. A yield of $N(B^+ \to$

$D_s^+ K^- K^+) = 65 \pm 10$ is found, consistent with the yield obtained from the full $B^+ \to D_s^+ K^+ K^-$ measurement. The branching fraction for $B^+ \to D_s^+ \phi$ decays is calculated as

$$\mathcal{B}(B^+ \to D_s^+ \phi) = R \times \frac{\mathcal{B}(\overline{D}^0 \to K^+ K^-)}{\mathcal{B}(\phi \to K^+ K^-)} \times \mathcal{B}(B^+ \to D_s^+ \overline{D}^0), \qquad (6.22)$$

where the branching fraction $\mathcal{B}(\phi \to K^+ K^-) = 0.489 \pm 0.005$ has been used [27].

The free variable R is defined to be the ratio of the signal and normalisation yields, corrected for the selection efficiencies. The yield of signal candidates in each D_s^+ mode is constructed from R and the normalisation yield for the given D_s^+ decay mode, $N(B^+ \to D_s^+ \overline{D}^0)$. The product of these two quantities is corrected by the ratio of selection efficiencies

$$N(B^+ \to D_s^+ \phi) = R \times N(B^+ \to D_s^+ \overline{D}^0) \times \frac{\epsilon(B^+ \to D_s^+ \phi)}{\epsilon(B^+ \to D_s^+ \overline{D}^0)}. \qquad (6.23)$$

The simultaneous fit measures a single value of R for all D_s^+ decay mode categories. From an ensemble of pseudoexperiments, R is distributed normally. It can be written as the ratio of signal and normalisation branching fractions using Eq. 6.22. The value is determined to be

$$R = \frac{\mathcal{B}(B^+ \to D_s^+ \phi)}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} \times \frac{\mathcal{B}(\phi \to K^+ K^-)}{\mathcal{B}(\overline{D}^0 \to K^+ K^-)} = (1.6^{+2.2}_{-1.9} \pm 1.1) \times 10^{-3}, \qquad (6.24)$$

where the first uncertainty is statistical and the second systematic. This corresponds to a branching fraction for $B^+ \to D_s^+ \phi$ decays of

$$\mathcal{B}(B^+ \to D_s^+ \phi) = (1.2^{+1.6}_{-1.4} \pm 0.8 \pm 0.1) \times 10^{-7}, \qquad (6.25)$$

where the first uncertainty is statistical, the second systematic, and the third results from the uncertainty on the branching fractions $\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)$, $\mathcal{B}(\phi \to K^+ K^-)$ and $\mathcal{B}(\overline{D}^0 \to K^+ K^-)$. Considering only the statistical uncertainty, the significance of the $B^+ \to D_s^+ \phi$ signal is 0.8 standard deviations (σ).

Th branching fraction for $B^+ \to D_s^+ \phi$ decays is also determined separately for the different D_s^+ decays modes included in the search. These are found to be

$$
\begin{aligned}
D_s^+ \to \phi \pi^+ & \quad : \mathcal{B}(B^+ \to D_s^+ \phi) = & +2.7^{+2.9}_{-2.3} \times 10^{-7} \\
D_s^+ \to K^+ K^- \pi^+ & \quad : \mathcal{B}(B^+ \to D_s^+ \phi) = & +1.2^{+2.2}_{-1.8} \times 10^{-7} \\
D_s^+ \to \pi^+ \pi^- \pi^+ & \quad : \mathcal{B}(B^+ \to D_s^+ \phi) = & -9.4^{+3.6}_{-2.8} \times 10^{-7} \\
D_s^+ \to K^+ \pi^- \pi^+ & \quad : \mathcal{B}(B^+ \to D_s^+ \phi) = & +3.7^{+1.2}_{-7.6} \times 10^{-7},
\end{aligned}
\tag{6.26}
$$

where these results correspond to the following yields for each D_s^+ decay mode

$$
\begin{aligned}
D_s^+ \to \phi \pi^+ & \quad : N(B^+ \to D_s^+ \phi) = & +3.9^{+4.2}_{-3.3} \\
D_s^+ \to K^+ K^- \pi^+ & \quad : N(B^+ \to D_s^+ \phi) = & +2.7^{+5.0}_{-4.2} \\
D_s^+ \to \pi^+ \pi^- \pi^+ & \quad : N(B^+ \to D_s^+ \phi) = & -5.2^{+2.0}_{-1.6} \\
D_s^+ \to K^+ \pi^- \pi^+ & \quad : N(B^+ \to D_s^+ \phi) = & +0.8^{+2.6}_{-1.7}.
\end{aligned}
\tag{6.27}
$$

A visual representation of these measurements are shown in Fig. 6.22 along with the value determined using all modes simultaneously.
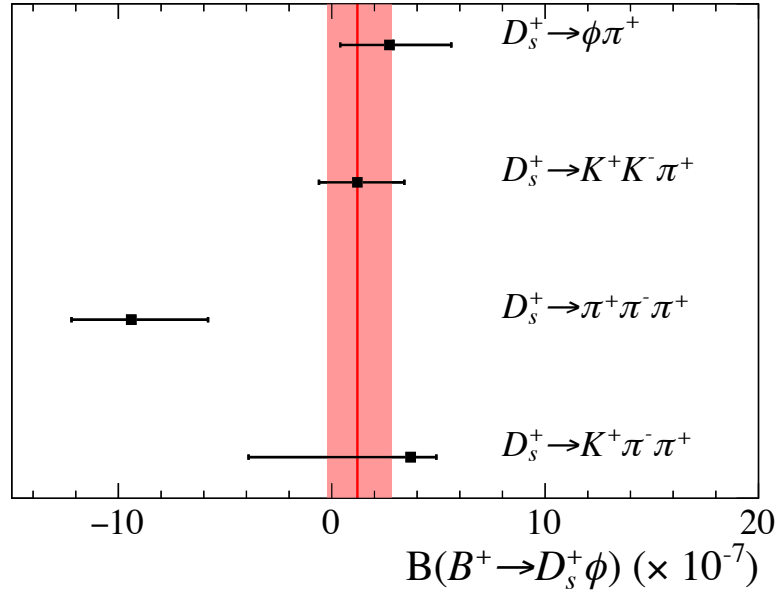

Figure 6.22: Results split for different D_s^+ decay modes.

6.8.1 Limit setting

The measured branching fraction, $\mathcal{B}(B^+ \to D_s^+ \phi) = (1.2^{+1.6}_{-1.4} \pm 0.8 \pm 0.1) \times 10^{-7}$, is not significant enough to constitute evidence or an observation for the $B^+ \to D_s^+ \phi$ decay

and is consistent with a branching fraction of zero. Whilst this measurement is useful in itself, for example it could provide constraints in combination with other results, it is also useful to set a limit on the branching fraction for a more straightforward comparison with theoretical predictions. Three different methods of limit estimation are attempted. These methods make different assumptions are therefore applicable in slightly different situations.

The CL_S method

The first method tried is the CL_S method, widely used in the high energy physics community. This method tests the p-value of a signal plus background hypothesis, $\text{CL}_{\text{S+B}}$, against a background only hypothesis, CL_B,

$$\text{CL}_\text{S} = \frac{\text{CL}_{\text{S+B}}}{\text{CL}_\text{B}}. \tag{6.28}$$

The free parameters in the fit other than the POI are considered nuisance parameters. This method is implemented using the RooStats package [43] within the Root framework.

The CL_S distribution is shown in Fig. 6.23 which includes the 1σ bands in green and 2σ bands in yellow. This limit is determined in the asymptotic limit [44]. The 95% upper limit is determined as the point where the CL_S value falls below 5% as illustrated by the red line. This corresponds to a 95% limit of

$$\mathcal{B}(B^+ \to D_s^+ \phi) < 4.2 \times 10^{-7}. \tag{6.29}$$

The profile likelihood method

An upper limit at 95% confidence limit is determined for the branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)$ using the profile likelihood method. This is calculated by determining the value of the branching fraction x_U that satisfies the equation

$$\frac{\int_0^{x_\text{U}} \mathcal{L}(x)dx}{\int_0^{\infty} \mathcal{L}(x)dx} = 0.95, \tag{6.30}$$
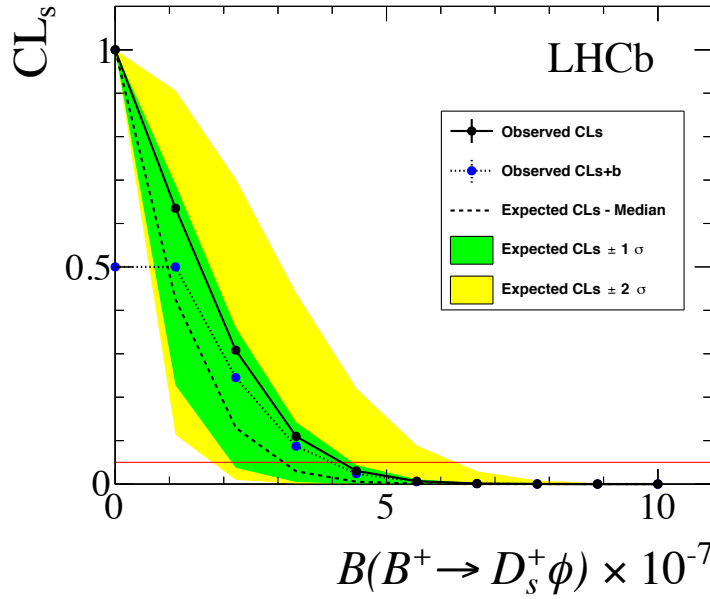
Figure 6.23: CL$_\mathrm{S}$ limit determination.

where $\mathcal{L}(x)$ is the profile likelihood as a function of the branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)$. This Bayesian method integrates the prior knowledge about the branching fraction, namely that the value must be greater than or equal to zero; the profile likelihood is integrated from zero upwards. The 95% CL limit determined when considering only statistical uncertainties is

$$\mathcal{B}(B^+ \to D_s^+ \phi) < 4.1 \times 10^{-7}. \tag{6.31}$$

To account for systematic uncertainty, the likelihood is convolved with a Gaussian distribution with a width given by the systematic uncertainty. The likelihood and difference in the log-likelihood are shown in Fig. 6.24, with and without the systematic uncertainty included. The limit at 95% CL including the systematic uncertainty is determined to be

$$\mathcal{B}(B^+ \to D_s^+ \phi) < 4.4 \times 10^{-7}. \tag{6.32}$$

The Feldman-Cousins method

Upper limits at 95% and 90% confidence levels (CL) are also determined using the Feldman-Cousins approach [45]. An ensemble of pseudo-experiments is generated for
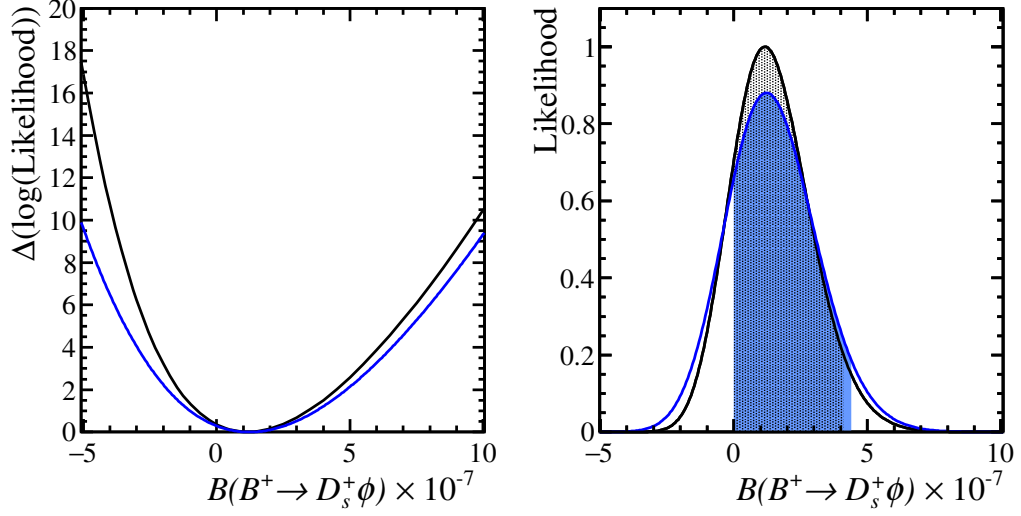
Figure 6.24: Bayesian profile likelihood limit determination: the (left) difference in the log-likelihood and (right) the likelihood as a function of the assumed $B^+ \to D_s^+ \phi$ branching fraction. The black distributions include only the statistical uncertainty, whilst the blue also include the systematic uncertainty. The shaded regions represent the areas integrated to determine the 95% CL limits.

different values of the branching fraction $\mathcal{B}(B^+ \to D_s^+ \phi)$. These generated pseudo-experiments are then fitted with the nominal fit model to calculate the fitted branching fraction and associated statistical uncertainty, σ_{stat}. This method constructs confidence bands based on a likelihood ratio method, calculating the probability of fitting a branching fraction for a given generated branching fraction. This probability is assumed to follow a Gaussian distribution with width $\sigma = \sqrt{\sigma_{\mathrm{stat}}^2 + \sigma_{\mathrm{syst}}^2}$, where σ_{stat} and σ_{syst} are the statistical and systematic uncertainties. The dominant source of systematic uncertainty in this measurement is from the background PDFs. As the size of this uncertainty is not expected to vary as a function of the generated branching fraction, σ_{syst} is assumed to be constant. Nuisance parameters are accounted for using the plug-in method [37]. The generated confidence bands are shown in Fig. 6.25, where the statistical-only 90% and 95% confidence level bands are shown, along with the 95% confidence level band with systematic uncertainty included. This corresponds to a statistical-only 95% (90%) confidence level of $\mathcal{B}(B^+ \to D_s^+ \phi) < 4.4 \times 10^{-7}$ (3.9×10^{-7}),

and a 95% (90%) confidence level including systematic uncertainties of

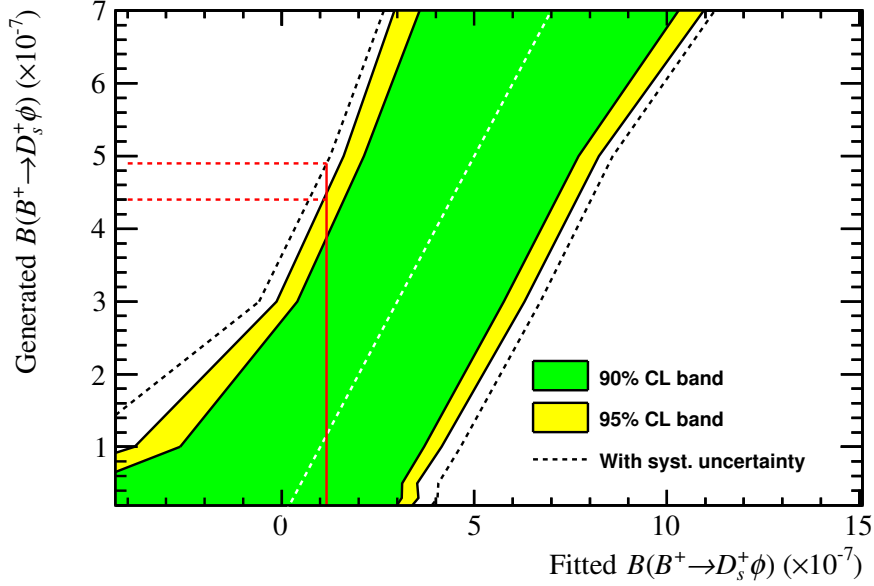$$\mathcal{B}(B^+ \to D_s^+ \phi) < 4.9 \times 10^{-7} \ (4.2 \times 10^{-7}). \tag{6.33}$$

Figure 6.25: Confidence bands produced using the Feldman-Cousins approach. The green and yellow bands represent the statistical-only 90% and 95% confidence level bands and the black dotted line represents the 95% limit including systematic uncertainties. The measured value of the branching fraction is shown by the vertical red line, and the corresponding 95% confidence levels, with and without systematic uncertainties, are represented by the dotted red lines.

6.8.2 Comparison to the previous measurement

The limit on the $B^+ \to D_s^+ \phi$ branching fraction presented here and in Ref [1] supersede the previous evidence reported by the LHCb collaboration in Ref. [46]. The updated analysis takes advantage of a much larger dataset now available at LHCb. This update determines that there is a sizeable contribution from $B^+ \to D_s^+ K^+ K^-$ decays that contribute within the ϕ-meson mass window that was previously not considered. The result is consistent with the prediction that rescattering contributions to $B^+ \to D_s^+ \phi$ decays are small.