

# Chapter 5

## Mass fit to $B^+ \rightarrow D_s^+ K^+ K^-$ candidates

### Contents

---

<b>5.1</b>	<b>Fit strategy . . . . .</b>	<b>88</b>
<b>5.2</b>	<b>Fit components . . . . .</b>	<b>89</b>
5.2.1	Signal and normalisation decays . . . . .	89
5.2.2	Partially reconstructed backgrounds . . . . .	90
5.2.3	Combinatorial background . . . . .	96
<b>5.3</b>	<b>Free and constrained parameters . . . . .</b>	<b>97</b>
5.3.1	Parameter of interest . . . . .	97
5.3.2	Shape parameters . . . . .	97
5.3.3	Yields . . . . .	97
<b>5.4</b>	<b>Fit validation . . . . .</b>	<b>98</b>
<b>5.5</b>	<b>Normalisation and signal fits . . . . .</b>	<b>100</b>
<b>5.6</b>	<b>Efficiency corrections . . . . .</b>	<b>103</b>
5.6.1	Efficiencies from simulations . . . . .	103
5.6.2	Efficiencies requiring calibration samples . . . . .	106
5.6.3	Total efficiency . . . . .	109
<b>5.7</b>	<b>Systematic uncertainties . . . . .</b>	<b>110</b>
5.7.1	Relative efficiencies . . . . .	110
5.7.2	Signal and normalisation PDFs . . . . .	112
5.7.3	Background PDFs . . . . .	113
5.7.4	Charmless contribution . . . . .	113
5.7.5	Total systematic uncertainty . . . . .	114
<b>5.8</b>	<b>Results . . . . .</b>	<b>114</b>
<b>5.9</b>	<b>Outlook . . . . .</b>	<b>117</b>

In this chapter the methodology used to search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is described. The branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)$  is determined by measuring the ratio of  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  yields. This ratio is corrected to account for the corresponding selection efficiencies of the two modes. Finally, the corrected yield ratio is multiplied by the externally measured branching fractions for the normalisation channel  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $\bar{D}^0 \rightarrow K^+ K^-$  decays to determine  $\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)$ .

The parametrisations used to model the signal and backgrounds components and extract the candidate yields are described in Section 5.2, the efficiency corrections are described in Section 5.6 and the resulting calculation of the branching fraction is in Section 5.8.

## 5.1 Fit strategy

The search for  $B^+ \rightarrow D_s^+ K^+ K^-$  involves two independent unbinned extended maximum likelihood fits for the signal and normalisation channels implemented with the RooFit package [34] within ROOT. The extended likelihoods for the two fits are constructed in a similar manner to those already detailed in Sec. 4.3.7. However, as a larger number of components are included there are correspondingly more contributions included in the likelihood

$$-\log \mathcal{L}(n_0 \dots n_j, \vec{p}) = -\sum_i^N \log \left( \sum_j n_j f_j(m = m_i, \vec{p}) \right) + \sum_j n_j, \quad (5.1)$$

where the index  $j$  represents each component of the model with yield  $n_j$  and PDF  $f_j$ . The constant  $\log N!$  term has been ignored. Separate likelihoods are created for the signal and normalisation fits.

The raw  $B^+ \rightarrow D_s^+ K^+ K^-$  yield is corrected on a per-candidate basis to account for the phase-space dependence of the signal efficiencies in this three-body decay. This is implemented using the *sPlot* technique [29] to determine a signal weight  $W_{s,i}$  for

each event  $i$  in the fitted data set. The weights for each component are constructed such that they sum to the fitted value of that components yield

$$n_{\text{sig}} = \sum_i^N W_{\text{sig},i}, \quad (5.2)$$

where  $n_{\text{sig}}$  is the fitted signal yield and  $N$  is the total number of entries in the data set. The efficiencies for the  $B^+ \rightarrow D_s^+ K^+ K^-$  signal decays are determined as a function of the kinematic properties of the decay. The weight for each entry  $i$  in the data set is corrected with the appropriate efficiency for its given kinematics

$$n_{\text{sig,corr}} = \sum_i^N \frac{W_{\text{sig},i}}{\epsilon_i}. \quad (5.3)$$

The propagation of the uncertainty in this corrected yields is described in Sec. 5.8.

The normalisation channel decay proceeds via a pseudo two-body process, therefore no kinematic dependent efficiency correction is required.

## 5.2 Fit components

In order to extract the yields of  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays the invariant mass distributions for the processes contributing within the invariant mass range are parametrised with probability density functions (PDFs). Both the signal and normalisation channels are considered within the same  $B^+$  meson invariant mass range 5100–5900 MeV/ $c^2$ . This is sufficiently wide to allow the contributions from different background components to be distinguished and accurately extrapolated into the signal region. The entire  $m(K^+ K^-)$  phase-space is included in the search for the signal decays, including the range in the vicinity of the  $\phi$  meson used later in Chapter 6.

### 5.2.1 Signal and normalisation decays

The invariant mass distributions of  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are parametrised as the sum of two Crystal Ball (CB) functions. The CB function consists

of a Gaussian function with a power-law tail and is typically used to parametrise losses due to radiative processes. This is defined as

$$\text{CB}(m|\mu, \sigma, n, \alpha) = \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2}, & \text{if } \left(\frac{m-\mu}{\sigma}\right) < -|\alpha| \\ \frac{\left(\frac{n}{|\alpha|}\right)^n \times e^{-\frac{1}{2}|\alpha|^2}}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{m-\mu}{\sigma}\right)^n}, & \text{otherwise} \end{cases} \quad (5.4)$$

where  $\mu$ ,  $\sigma$ ,  $n$  and  $\alpha$  are adjustable parameters and  $m$  is the  $B$  meson invariant mass observable. The sum of two CB functions is constructed with a variable fraction  $f_\sigma$  assigned to the CB function with the narrower width,

$$\text{DCB}(m|\mu, \sigma_1, \sigma_2, n, \alpha) = f_\sigma \times \text{CB}(m|\mu, \sigma_1, n, \alpha) + (1 - f_\sigma) \times \text{CB}(m|\mu, \sigma_2, n, \alpha), \quad (5.5)$$

where the same tail parameters,  $n$  and  $\alpha$  are used for both functions, but the widths,  $\sigma_1$  and  $\sigma_2$ , are allowed to be different (with  $\sigma_1 < \sigma_2$ ). As both CB shapes have the same parameter  $\alpha$ , the tails are constrained to be on the same side. Values for the adjustable parameters are determined from fits to simulated decays passing the selection requirements applied to the data. However, a number of parameters are not completely constrained from the simulations. The mean position  $\mu$  is allowed vary freely in the fit to data, as is the narrowest CB width of the normalisation and signal decays. The tail parameters  $n$  and  $\alpha$  are highly correlated, therefore the value of  $n$  is fixed to unity in both the fits to simulations and data. The values determined from simulations for  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are tabulated in Table 5.1 and the results of the corresponding fits are shown in Fig. 5.1.

### 5.2.2 Partially reconstructed backgrounds

Partially reconstructed decays are those in which the five final state particles combined in the signal mode are only a subset of a background mode's final state. Decays of  $b$ -hadrons can contribute at lower invariant masses below the signal peak when one or more the decay products have not been reconstructed. For decays to contribute within the fitted  $B^+$  invariant mass window, the particle or particles that have not been reconstructed must be fairly low-momentum (soft) such that the invariant mass

Parameter	Value
$B^+ \rightarrow D_s^+ K^+ K^-$	
$\sigma_1/\sigma_2$	$0.53 \pm 0.02$
$f_\sigma$	$0.87 \pm 0.02$
$\alpha$	$2.60 \pm 0.03$
$n$	$1 \pm 0$
$B^+ \rightarrow D_s^+ \bar{D}^0$	
$\sigma_1/\sigma_2$	$0.60 \pm 0.03$
$f_\sigma$	$0.66 \pm 0.12$
$\alpha$	$2.67 \pm 0.12$
$n$	$1 \pm 0$

Table 5.1: Fixed values obtained in fits to simulations used in the model for the signal and normalisation PDFs.

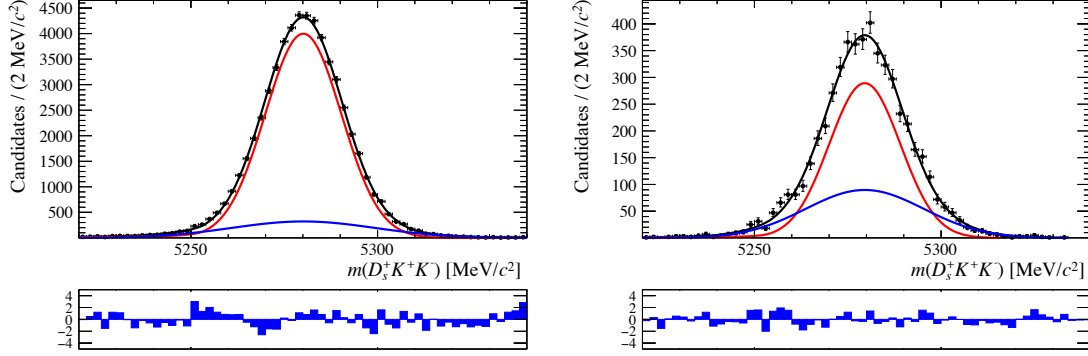


Figure 5.1: Invariant mass fits to signal (left) and normalisation (right) channel simulation samples.

of the remaining particles is large. Accurate parametrisation of these contributions is vital as many of the distributions extend close to or within the range of the signal distribution. Incorrectly attributing these decays to the signal component could lead to the incorrect branching fraction being measured.

### Backgrounds to the normalisation channel

The low invariant mass region of the  $D_s^+ \bar{D}^0$  spectrum is populated by decays of  $B^+$  mesons to combinations of  $D$  and excited  $D$  mesons. These  $\bar{D}^{*0}$  and  $D_s^{*+}$  mesons decay strongly to a ground state  $\bar{D}^0$  or  $D_s^+$  meson and a soft pion or photon. The branching fractions for these decays are listed in Table 5.2.

The excited charm mesons  $\bar{D}^{*0}$  and  $D_s^{*+}$  are vector ( $J^P = 1^-$ ) mesons. The

Decay	Branching fraction
$\bar{D}^{*0} \rightarrow \bar{D}^0 \gamma$	$(64.7 \pm 0.9)\%$
$\bar{D}^{*0} \rightarrow \bar{D}^0 \pi^0$	$(35.3 \pm 0.9)\%$
$D_s^{*+} \rightarrow D_s^+ \gamma$	$(93.5 \pm 0.7)\%$
$D_s^{*+} \rightarrow D_s^+ \pi^0$	$(5.8 \pm 0.7)\%$

Table 5.2: Branching fractions for excited charm mesons [27].

partially reconstructed invariant mass of the  $D_s^+$  and  $\bar{D}^0$  mesons,  $m(D_s^+ \bar{D}^0)$ , varies depending on the spin of the missed particle. Analytical PDFs are used to account for the spin and mass of the missing particle. These PDFs have been used to describe other partially reconstructed  $B \rightarrow DX$  decays investigated by the LHCb collaboration [35].

$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\pi^0])\bar{D}^0$  **and**  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\pi^0])$ : the  $\pi^0$  meson is a pseudo-scalar ( $J^P = 0^-$ ) particle with a mass  $m(\pi^0) = 134.9766 \pm 0.0006 \text{ MeV}/c^2$ . The partially reconstructed invariant mass distribution  $m(D_s^+ \bar{D}^0)$  can be described by a parabola convolved with a Gaussian resolution function. The parabola does not extend beyond kinematic endpoints defined by the parameters  $a$  and  $b$  and has a minimum in the centre

$$f(m|a, b, \sigma, \xi, \delta) = \int_a^b \left( \mu - \frac{a+b}{2} \right)^2 \left( \frac{1-\xi}{b-a} \mu + \frac{b\xi-a}{b-a} \right) e^{-\frac{(\mu - (m-\delta))^2}{2\sigma^2}} d\mu. \quad (5.6)$$

Here,  $\sigma$  is the width of the resolution Gaussian and  $\delta$  allows the function to be offset in invariant mass. The parameter  $\xi$  introduces the freedom for the two sides of the parabola to have difference heights, achieved by multiplying the parabola by a line whose slope is depends on  $\xi$ . The resulting function is a double peaked structure shown in Fig. 5.2. The parameters  $a$  and  $b$  are calculated from the kinematics of the  $B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ \pi^0)\bar{D}^0$  and  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0 \pi^0)$  decays respectively, and are listed in Table 5.3.

$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\gamma])\bar{D}^0$  **and**  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\gamma])$ : the  $\gamma$  boson is a mass-less vector ( $J^P = 1^-$ ) particle. The partially reconstructed  $m(D_s^+ \bar{D}^0)$  invariant

Mode	$a$ (MeV/ $c^2$ )	$b$ (MeV/ $c^2$ )
$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\pi^0])\bar{D}^0$	5051.4	5132.9
$B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\pi^0])$	5051.5	5128.6

Table 5.3: Kinematic endpoints for the partially reconstructed  $B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ \pi^0)\bar{D}^0$  and  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0 \pi^0)$  decays.

mass is also described by a parabola convolved with a Gaussian resolution function. The parabola does not extend beyond the endpoints  $a$  and  $b$ , and has a maximum in the centre

$$f(m|a, b, \sigma, \xi, \delta) = \int_a^b -(\mu - a)(\mu - b) \left( \frac{1 - \xi}{b - a} \mu + \frac{b\xi - a}{b - a} \right) e^{-\frac{(\mu - (m - \delta))^2}{2\sigma^2}} d\mu. \quad (5.7)$$

Again  $\sigma$ ,  $\delta$  and  $\xi$  control the width, offset and relative heights of two sides of the parabola. The resulting function is a broad single peak shown in Fig. 5.2.

Mode	$a$ (MeV/ $c^2$ )	$b$ (MeV/ $c^2$ )
$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\gamma])\bar{D}^0$	4976.7	5213.1
$B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\gamma])$	4970.1	5216.1

Table 5.4: Kinematic endpoints for the partially reconstructed  $B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ \gamma)\bar{D}^0$  and  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0 \gamma)$  decays.

## Backgrounds to the signal channel

The signal channel receives contributions at low invariant mass from a number of different decays. All modes considered involve a  $B_s^0$  or  $B^0$  meson decay in which one or more soft decay products have not been reconstructed.

$\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0}$ : this decay can form a background to the  $B^+ \rightarrow D_s^+ K^+ K^-$  signal when a soft pion from the  $K^{*0} \rightarrow K^+ \pi^-$  decay is not reconstructed. 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 its pole mass is below the  $K^- K^{*0}$  threshold. The PDF for this background is determined from a sample of fully reconstructed simulated events that have been processed using the same procedure as the signal decays. The PDF is created using a

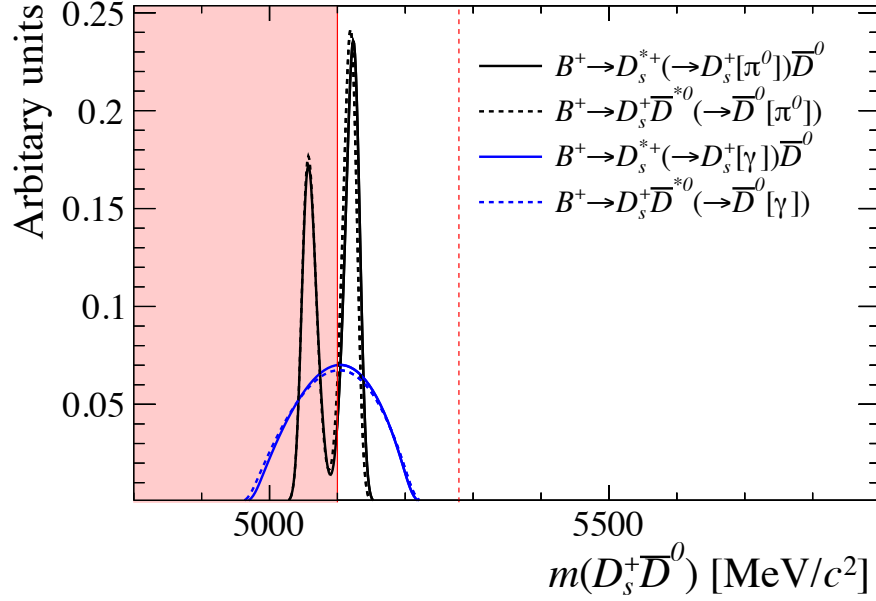


Figure 5.2: Partially reconstructed  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  PDFs as described in Sec. 5.2.2. The range below 5100 MeV/ $c^2$  (highlighted in red) is not included in the fit range, but included to shown the full PDF distributions. The  $B^+$  meson mass is represented by a vertical line.

kernel estimation technique [36] implemented in the RooKeysPDF class within RooFIT. The resulting PDF is shown in Fig. 5.3.

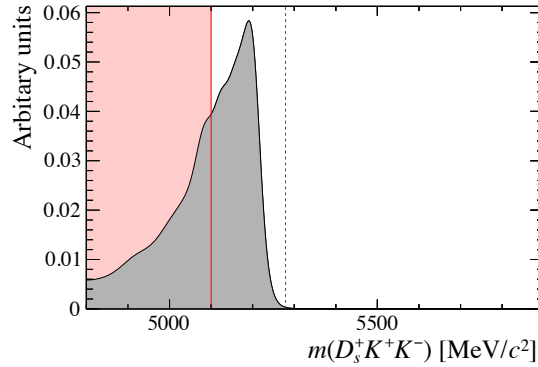


Figure 5.3: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0}$  mass shape. The  $B^+$  meson mass is represented by a vertical dashed line. The range below 5100 MeV/ $c^2$  (highlighted in red) is not included in the fit range.

$\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}$ : similarly this decay can form a background at low invariant mass when a soft neutral particle is not reconstructed in the  $D_s^{*+} \rightarrow D_s^+ X$  decay



in addition to the pion from the  $K^{*0}$ . This PDF is also determined using the RooKeysPDF class and shown in Fig. 5.4.

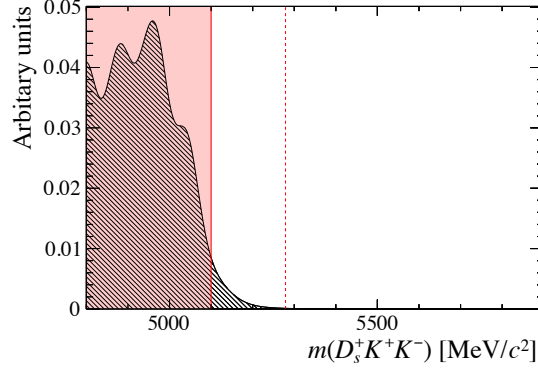


Figure 5.4: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}$  mass shape. The  $B^+$  meson mass is represented by a vertical dashed line. The range below 5100 MeV/ $c^2$  (highlighted in red) is not included in the fit range.

$\bar{B}_s^0 \rightarrow D_s^+ D_s^-$ : this decay can form a background to the signal when a pion is missed from either of the  $D_s^+$  decays. This requires both  $D_s^+$  mesons to decay to the  $D_s^+ \rightarrow K^+ K^- \pi^+$  final state. This background becomes significant at close to the  $B^+$  meson mass as shown in Fig. 5.5. This PDF is similarly determined by creating a kernel estimation of fully-reconstructed simulated events passing the signal selection.

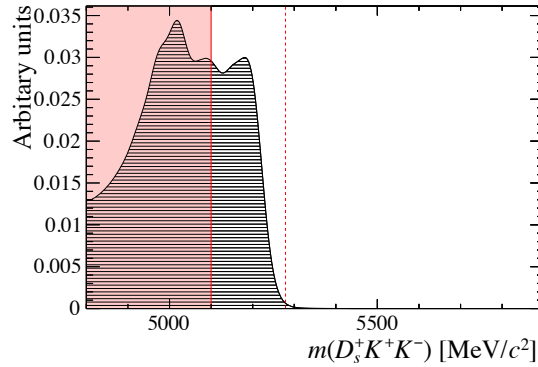


Figure 5.5: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  mass shape. The  $B^+$  meson mass is represented by a vertical dashed line. The range below 5100 MeV/ $c^2$  (highlighted in red) is not included in the fit range.

$\bar{B}^0 \rightarrow D_s^+ D^-$ : this decay can form a background when both the  $D_s^+$  and  $D^-$  mesons decay to the  $K^\pm K^\mp \pi^\pm$  final state. Due to the similar topology to the  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  decay, the same PDF determined from simulated events is used, however it is shifted down in mass by  $40 \text{ MeV}/c^2$  to account for the difference in kinematics.

$\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$ : similarly this decay can cause a background when both a soft neutral particle is missed from the  $D_s^{*+} \rightarrow D_s^+ X$  decay as well as a pion from either of the  $D_s^+$  mesons. As such this only has a small contribution within the fit range as shown in Fig. 5.6.

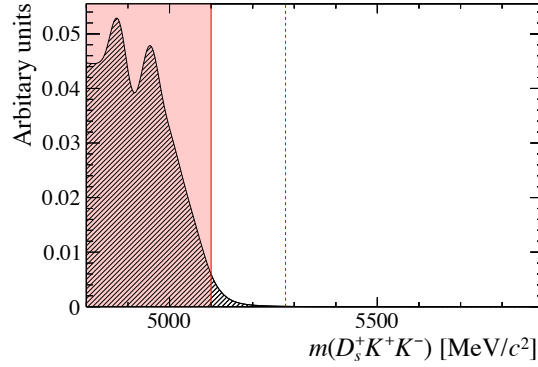


Figure 5.6: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$  mass shape.

All PDFs determined using kernel estimations from simulated events are convolved with a Gaussian distribution to account for difference between the simulations and data. The mean position of the Gaussian is given by a single free parameter  $\delta$ , allowing these PDF to move slightly high or lower in mass. The width of the Gaussian is increased to account for the difference in resolution between simulation and data.

### 5.2.3 Combinatorial background

The dominant source of background under the signal peak is due to combinations of unrelated tracks. This combinatorial background is modelled using an exponential function

$$f(m|c) = e^{-m \times c}, \quad (5.8)$$

where  $c$  is the single degree of freedom controlling the effective slope of the function and  $m$  is the observable  $B^+$  meson invariant mass. The separate fits to the signal and normalisation modes have the freedom to have different combinatorial slopes, motivated by the difference in background levels for the two decays.

### 5.3 Free and constrained parameters

The fit to signal and normalisation decays have eleven and nine free parameters respectively. These control the yields and shapes of the fit components.

#### 5.3.1 Parameter of interest

**Signal:** the parameter of interest in the signal decay is  $N(B^+ \rightarrow D_s^+ K^+ K^-)$ , the yield attributed to the signal PDF.

**Normalisation:** the parameter of interest in the signal decay is  $N(B^+ \rightarrow D_s^+ \bar{D}^0)$ , the yield attributed to the normalisation PDF.

#### 5.3.2 Shape parameters

**Signal:** The signal fit contains four free parameters that determine the shape of various PDFs. This corresponds to the mass offset  $\delta$ , signal mean value  $\mu$ , signal width  $\sigma_1$  and combinatorial slope  $c$ .

**Normalisation:** The normalisation fit has five free parameters that control the distributions of the signal and background PDFs. This includes the mean  $B^+$  mass  $\mu$ , signal width  $\sigma_1$ , the combinational slope  $c$ , the mass offset  $\delta$  and the relative heights of the double peaked partially reconstructed shapes  $\xi$ .

#### 5.3.3 Yields

**Signal:** the yields of each background component are left free in the signal fit. This corresponds to six free parameters for  $N_{\text{comb}}$ ,  $N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})$ ,  $N(\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0})$ ,  $N(B_s^0 \rightarrow D_s^+ D_s^-)$ ,  $N(B^0 \rightarrow D_s^+ D^-)$ , and  $N(B_s^0 \rightarrow D_s^{*+} D_s^-)$ .

**Normalisation:** the yields of the combinatorial and partially reconstructed backgrounds are left free in the fit to the normalisation channel. This corresponds to three parameters;  $N_{\text{comb}}$ ,  $N(B^+ \rightarrow D_s^{*+} \bar{D}^0)$  and  $N(B^+ \rightarrow D_s^+ \bar{D}^{*0})$ . The relative contributions of the  $D_s^{*+} \rightarrow D_s^+ \pi^0$  and  $D_s^{*+} \rightarrow D_s^+ \gamma$  decays are fixed to their ratio of branching fractions (similarly for  $\bar{D}^{*0}$ ). A factor is included to account for the fraction of each PDF that is within the fitted  $B^+$  invariant mass range.

## 5.4 Fit validation

The fitting framework is validated using large quantities of pseudo-experiments randomly generated using the same fit model. The free parameters are treated using the plug-in method [37]; the generated values of all free parameters are *plugged in* using the fitted values from the fit to data.

The fitted values and uncertainties are determined for each pseudo-experiment and a corresponding pull determined, defined as

$$g_{\text{pull}} = \frac{x_{\text{fit}} - x_{\text{gen}}}{\sigma} \quad (5.9)$$

where  $x_{\text{gen}}$  and  $x_{\text{fit}}$  are the generated and fitted values of the variable, and  $\sigma$  is the parameter's uncertainty. For an ideal unbiased fit model, the pull of each parameter of interest would be normally distributed with unit width and mean of zero.

The distributions of the yields, errors and pulls for the signal and normalisation pseudo-experiments are shown in Fig. 5.7. A fit is performed for each pull distributions using a Gaussian to determine the mean and width. For the signal yield the mean and width are within  $3\sigma$  of zero and one respectively. The normalisation yield shows a significant bias in the width. The bias implies the fit model is overestimating the uncertainty  $\sigma$  of the yield. If the normalisation yield uncertainty dominates the uncertainty in the branching fraction this could lead to an overestimation of the uncertainty on the final measured branching fraction.

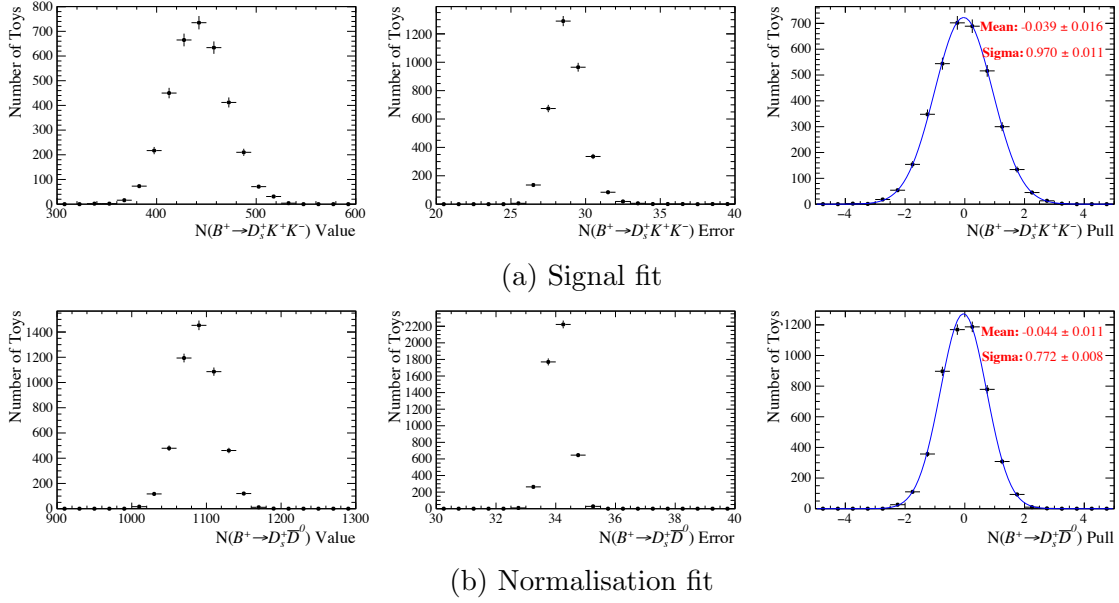


Figure 5.7: The distribution of the signal and normalisation yields, errors and pulls as determined from pseudo-experiments. The result of fit performed to the pull distributions is overlaid in blue, along with the numerical results in red.

To determine how the normalisation uncertainty propagates to the branching fraction another set of pseudo-experiments are produced. This set includes both signal and normalisation decays. To calculate the branching fraction the yield of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is corrected according to the signal efficiency as a function of the kinematics of a given candidate, as detailed previously in Eq. 5.3. The candidates are assumed to have a flat distribution in the two dimensional  $m^2(K^+ K^-)$  vs.  $m^2(D_s^+ K^-)$  space used to parametrise the efficiency. For each pseudo-experiment the branching fraction and uncertainty are produced. As no branching fraction is explicitly used to generate the pseudo-experiment (rather independent signal and normalisation yields), the pulls are redefined to be measured relative to the mean branching fraction

$$g_{\text{pull}} = \frac{x_{\text{fit}} - \bar{x}_{\text{fit}}}{\sigma} \quad (5.10)$$

where  $x_{\text{fit}}$  is the fitted value of the parameter,  $\bar{x}_{\text{fit}}$  is the average of all fitted values and  $\sigma$  is the uncertainty. The distribution of the branching fraction, uncertainty and pull are shown in Fig 5.8.

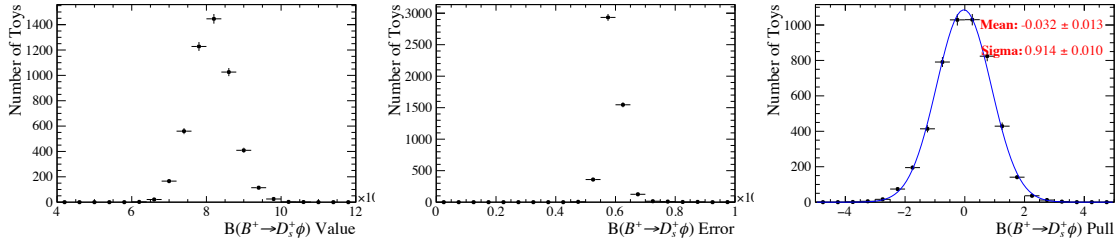


Figure 5.8: The distribution of the branching fraction, error and pull as determined from pseudo-experiments. The result of fit performed to the pull distributions is overlaid in blue, along with the numerical results in red.

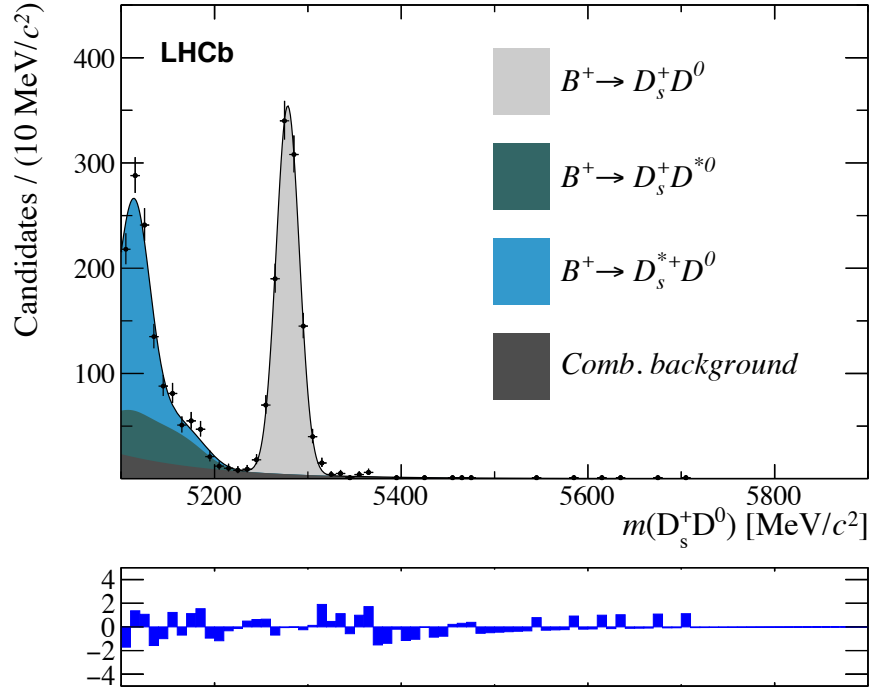
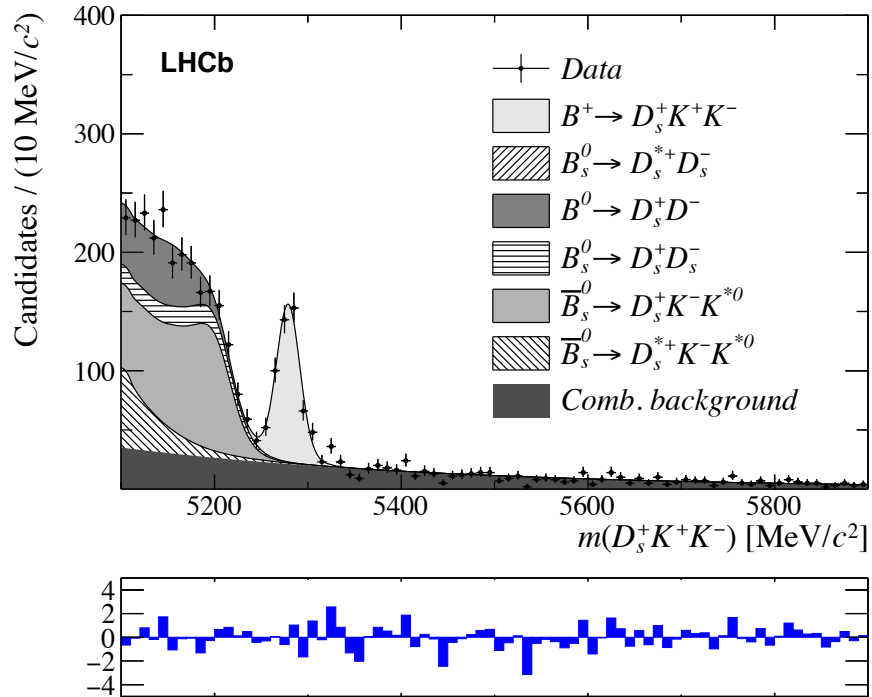
The pull in the branching fraction has a much smaller bias in the width than the normalisation pull. It would be possible to scale the uncertainty in the branching fraction as determined by the fit to account for the residual bias in the uncertainty. However, this would introduce a source of systematic uncertainty associated to scaling factor so the choice is made to not correct the slightly overestimated uncertainty.

## 5.5 Normalisation and signal fits

The signal and normalisation fits are performed using the model and dataset previously described. The results of the two fits are shown in Figs. 5.9 and 5.10. These figures shown the distribution of  $B^+$  candidates along with the total model PDF constructed with the values of the free parameters determined in the NLL minimisation process. The contributions from each different component in the model are superimposed, stacked upon one another, and detailed in the legends.

The distribution of  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates in Fig. 5.9 demonstrates the relatively high purity of the selection. Additionally, there is a good separation between the normalisation peak and the partially reconstructed background. Only the combinatorial background is found to have a contribution at the same invariant mass as the normalisation peak.

The distribution of  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates shown in Fig. 5.10 has a significant contribution from the signal decay. The background contribution under the signal PDF is larger than was observed for the normalisation channel. Although the

Figure 5.9: Invariant mass fit to  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates.Figure 5.10: Invariant mass fit to  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates.

dominant background below the signal is still the combinatorial background, the contribution from the partially reconstructed backgrounds is also larger; several of the background PDFs extend underneath the leftmost half of the signal distribution.

The numerical values of the free parameters determined in the fit to the signal and normalisation channels are listed in Tables 5.5 and 5.6.

Type	Parameter	Fit result
POI	$N(B^+ \rightarrow D_s^+ \bar{D}^0)$	$1091 \pm 34$
Shape	Mass shift $\delta$ (MeV/ $c^2$ )	$-10.6 \pm 1.3$
	Relative peak heights $\xi$	$0.0 \pm 0.2$
	Mean $B^+$ mass $\mu$ (MeV/ $c^2$ )	$5278.4 \pm 0.4$
	Signal width $\sigma_1$ (MeV/ $c^2$ )	$12.4 \pm 0.3$
	Combinatorial slope $c$	$(-9.4 \pm 1.1) \times 10^{-3}$
Yields	$N_{\text{comb}}$	$250 \pm 50$
	$N(B^+ \rightarrow D_s^+ \bar{D}^{*0})$	$330 \pm 70$
	$N(B^+ \rightarrow D_s^{*+} \bar{D}^0)$	$750 \pm 50$

Table 5.5: Normalisation fit result

Type	Parameter	Fit result
POI	$N(B^+ \rightarrow D_s^+ K^+ K^-)$	$442 \pm 29$
Shape	Mass shift $\delta$ (MeV/ $c^2$ )	$4 \pm 12$
	Mean $B^+$ mass $\mu$ (MeV/ $c^2$ )	$5278.9 \pm 0.9$
	Signal width $\sigma_1$ (MeV/ $c^2$ )	$12.3 \pm 0.9$
	Combinatorial slope $c$	$(-2.8 \pm 0.3) \times 10^{-3}$
Yields	$N_{\text{comb}}$	$1110 \pm 70$
	$N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})$	$1100 \pm 500$
	$N(\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0})$	$300 \pm 120$
	$N(B_s^0 \rightarrow D_s^+ D_s^-)$	$200 \pm 500$
	$N(B^0 \rightarrow D_s^+ D^-)$	$470 \pm 150$
	$N(B_s^0 \rightarrow D_s^{*+} D_s^-)$	$0 \pm 400$

Table 5.6: Signal fit result

Two of the common parameters between the two modes, namely the mean  $B^+$  mass and width, have consistent values between the two fits. The combinatorial slopes vary between the two, likely because of the different background levels in the two datasets. Similarly the mass shifts vary between the two.



## 5.6 Efficiency corrections

The branching fraction for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is determined by correcting the yields of signal and background decays by their corresponding selection efficiencies. These account for each step of the selection process and ultimately quantifies the extent to which the signal and normalisation channels are affected differently. As described in Sec. 5.1 the signal yield is corrected by the efficiency as a function of the candidate's kinematic properties. Therefore the relative signal efficiency is determined as a function of the two-dimensional Dalitz plot coordinates  $m^2(D_s^+ K^-)$  and  $m^2(K^+ K^-)$

$$\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-)) = \frac{\epsilon_{B^+ \rightarrow D_s^+ K^+ K^-}(m^2(D_s^+ K^-), m^2(K^+ K^-))}{\epsilon_{B^+ \rightarrow D_s^+ \bar{D}^0}}. \quad (5.11)$$

The normalisation decay efficiency  $\epsilon_{B^+ \rightarrow D_s^+ \bar{D}^0}$  is a pseudo-two-body decay and therefore the efficiency has no dependence on the phase-space coordinates. The kinematic phase-space is split into bins and the efficiencies determined in each.

The efficiencies are all determined using samples of simulated signal and normalisation decays. The PID and MVA efficiencies require additional input from external calibration samples to correct for the imperfect modelling of the PID variables in the simulations.

### 5.6.1 Efficiencies from simulations

**Acceptance:** this accounts for the likelihood for the five final state tracks to be within the LHCb detector's acceptance. The charged tracks are required to be in the range  $10 < \cos \theta < 400 \text{ mrad}$ , where  $\theta$  is the angle between the beam direction and the track. The relative efficiency is shown as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  Dalitz plot in Fig.5.11a. The relative efficiency is close to one for most of the phase-space.

**Reconstruction:** the reconstruction efficiency accounts for the fraction of decays in which all five final state tracks have been correctly reconstructed and combined into a  $B^+$  candidate that passes all selection requirements outlined in

the  $B^+ \rightarrow D_s^+ K^+ K^-$  *Stripping Line*. The *Stripping Line* reconstruction also explicitly requires that at least one trigger has fired for the event to be reconstructed, therefore this efficiency also accounts for part of the trigger efficiency. The distribution of the relative reconstruction efficiency between the signal and normalisation mode is shown in Fig. 5.11b. This is greater than one across the whole phase-space as a result of the choice to reconstruct the  $B^+ \rightarrow D_s^+ \bar{D}^0$  normalisation decays using the same selection as the  $B^+ \rightarrow D_s^+ K^+ K^-$  signal. The latter requires the  $D_s^+$ ,  $K^+$  and  $K^-$  candidates to form a well reconstructed vertex where the  $B^+$  meson decays. However, for the normalisation channel the  $\bar{D}^0$  meson travels away from this vertex, so some of the longer lived  $\bar{D}^0$  candidates do not pass the vertex quality requirement. This reduces the efficiency for the normalisation channel relative to the signal.

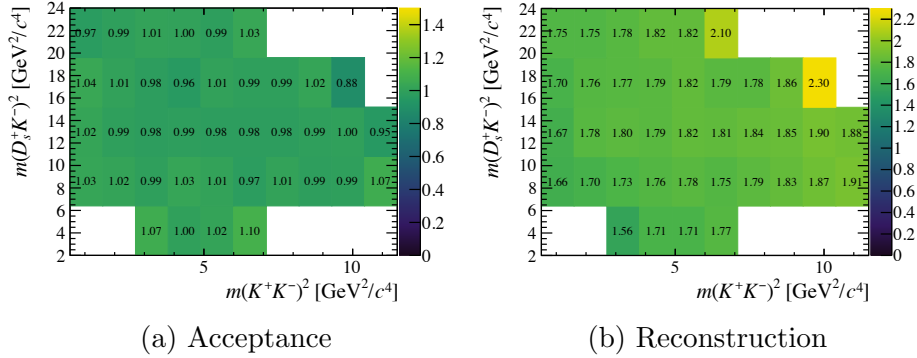


Figure 5.11: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

**Trigger:** the trigger efficiency accounts for the fraction of decays for which the signal candidate meets the TIS and TOS requirements outlined in Sec. 4.2, given that the trigger fired in that event. For the signal and normalisation simulation samples it is very likely that the signal candidate was the cause of the trigger if a trigger fired, therefore these are typically around 95%. The phase-space distribution of the relative efficiency is very uniform and shown in Fig. 5.12a.

**Veto:** this efficiency accounts for the fraction of decays passing the kinematic and normalisation vetoes detailed in Sec. 4.3.5 and 4.3.6. The vetoes to remove misidentified  $D$  and  $\Lambda_c^+$  hadrons are only applied to the  $D_s^+$  meson and therefore assumed to affect the signal and normalisation mode equally. The systematic uncertainty arising from this choice is discussed in Sections 5.7.1. Although the veto for misidentified normalisation decays proceeding via  $\bar{D}^0 \rightarrow K^+ \pi^-$  decays is included, the veto to remove correctly reconstructed normalisation decays with  $\bar{D}^0 \rightarrow K^+ K^-$  is not included. This is because the width of this veto is less than the binning width so it would therefore lead to the wrong efficiency for candidates within in the same bin but not in the window. This doesn't affect the final result, it simply means areas of the efficiency Dalitz plot are populated that will never be sampled. The distribution of the relative efficiency is shown in Fig. 5.12b. A vertical band is clear in the centre as a result of the veto for misidentified normalisation decays with  $\bar{D}^0 \rightarrow K^+ \pi^-$ .

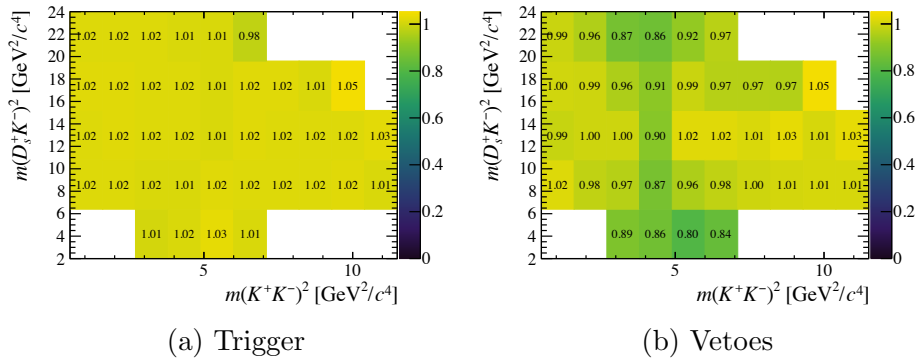


Figure 5.12: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

**Charmless:** this accounts for the fraction of decays failing the requirements on the  $D$  meson flight distance requirements aimed at removing charmless and single-charm backgrounds. The distribution shown in Fig. 5.13a shows a fairly flat distribution with the exception of some of the edge bins.

$\chi_{\text{IP}}^2$ : this accounts for the fraction of decays that pass the requirements on the  $B^+$  and  $D_s^+$  meson impact parameter significance. The relative efficiency is shown in Fig. 5.13b and shows a strong dependency on phase-space at high  $m^2(K^+ K^-)$ . The lower relative efficiencies at high  $m^2(K^+ K^-)$  may be because the  $D_s^+$  meson is produced almost at rest in this region. Therefore these  $D_s^+$  mesons are likely to have a lower transverse momentum and may fail the minimum impact parameter significance requirement.

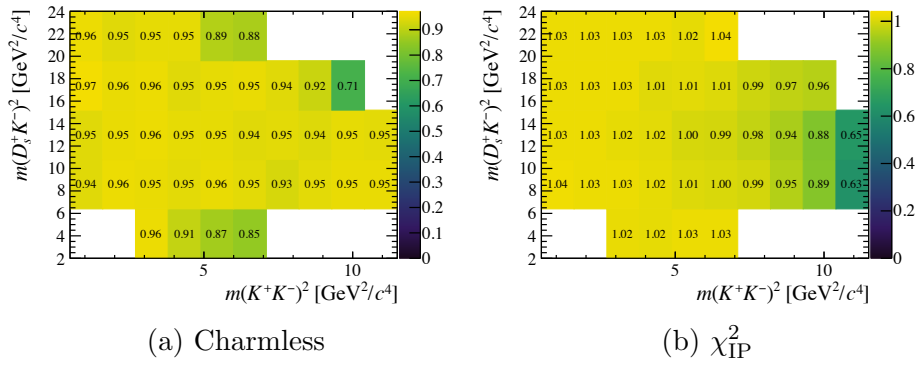


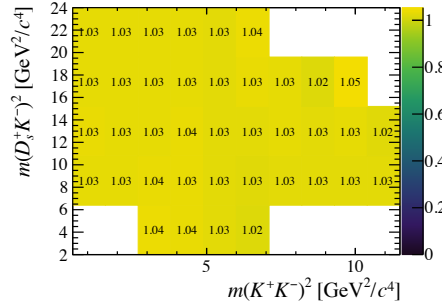
Figure 5.13: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

**Mass windows:** this accounts for the fraction of decays that pass the mass windows around the  $D_s^+$  and  $\bar{D}^0$  masses. No requirement is placed on the  $K^+ K^-$  pair, therefore the relative efficiency shown in Fig. 5.14a is slightly greater than one across the whole phase-space.

## 5.6.2 Efficiencies requiring calibration samples

### PID efficiency

The efficiencies for the particle identification requirements as described in Sec. 4.3.2 are calculated by correcting the PID variables in simulation using a package called PIDCALIB [38]. This uses calibration samples for the different particle species to determine the distribution of the PID variables in data. The calibration samples are background-subtracted to isolate the distributions of the PID variables for the tracks



(a) Mass windows

Figure 5.14: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

of interest. The calibration samples for both  $K^+$  and  $\pi^+$  mesons are collected from samples of  $D^{*+} \rightarrow (D^0 \rightarrow K^+ \pi^-) \pi^+$  decays, using the decay products of the  $D^0$  decay. Additionally, samples of protons are collected from  $\Lambda \rightarrow p \pi^-$  decays. The PID variable distributions depend on both the kinematics of the track in question and the occupancy of the detector as a whole. These can both affect the characteristics of the hits in the RICH sub-detectors and therefore result in different PID variable distributions. The calibration samples are characterised using three variables; the transverse momentum of the track  $p_T$ , the pseudo-rapidity  $\eta$ , and the number of tracks in an event  $n_{\text{Tracks}}$ . The calibration PID variables are parametrised using an unbinned approach, unlike in the search for  $B^+ \rightarrow D_s^+ \phi$  where a binned approach is used (Sec. 6.6.2). This unbinned approach creates four dimensional PDFs ( $\text{PID}, \eta, p_T, n_{\text{Tracks}}$ ) from the calibration samples using a kernel density estimation implemented with the MEERKAT package [39]. Similar four dimensional PDFs are created for simulation samples of the calibration modes. These two PDFs are used to calculate the transformation needed for each species as a function of  $(\text{PID}, \eta, p_T, n_{\text{Tracks}})$  to correct the distribution in simulation. This transformation preserves the correlations between the different PID variables and between the PID and kinematic properties. This correction is applied to both the DLL PID variables and ProbNNx variables used in the MVA selection. The phase-space distribution of the relative PID efficiency between the signal and normalisation channel is shown in Fig. 5.15a.

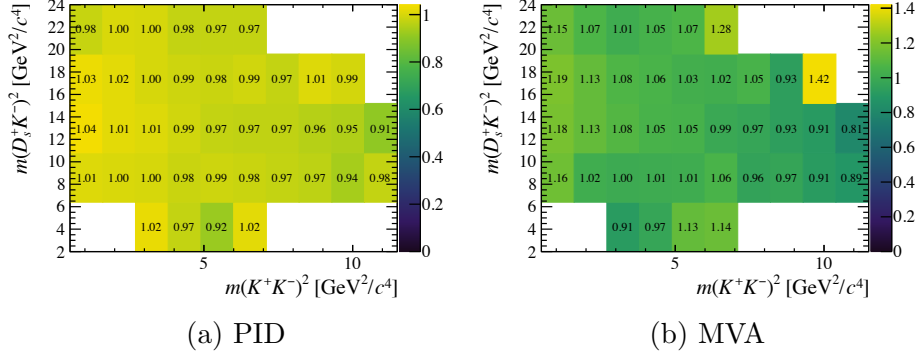


Figure 5.15: Efficiencies

### MVA efficiency

The MVA efficiency is determined from samples for signal and normalisation simulations in which the PID distributions have been corrected as described previously. This is different to the approach used in the search for  $B^+ \rightarrow D_s^+ \phi$  decays (Sec. 6.6.2) and is required to account for the variation in the efficiency across the phase-space. The corrected ProbNNx PID variables are used to determine the  $D_s^+$  and  $\phi$  MVA classifiers and the efficiency of the MVA requirements determined in different positions in the Dalitz plot as shown in Fig.5.15b. There is a strong dependence on the position, with decays at low  $m^2(K^+ K^-)$  values having a larger relative efficiency than those at higher values. This is likely to be because the MVA methods were trained using a sample of  $\phi \rightarrow K^+ K^-$  decays, so the selection favourably selects candidates with an invariant mass near the  $\phi$  meson mass. A cross check is performed to compare the MVA efficiencies determined directly using the classifier produced with the corrected PID variables and the method described in Sec. 4.3.7 that uses the MVA training samples to calculate the efficiency. The values of the efficiency determined for each mode are compared in Fig. 5.16 as a function of the  $m(K^+ K^-)$  mass. The MVA training modes only exist for two discrete values;  $m(K^+ K^-) = m(\phi)$  and  $m(K^+ K^-) = m(\bar{D}^0)$  represented by the red points. These are compared to the  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  efficiencies from the corrected simulations in black and blue respectively. A good agreement is found.

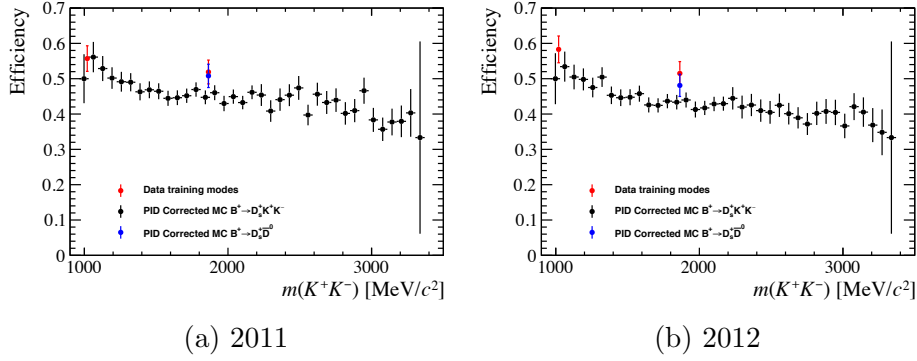


Figure 5.16: The PID efficiency variation as a function  $m(K^+ K^-)$  as determined from the MVA training modes (red) and the  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  simulation samples (black and blue respectively). A good agreement is found between the red and black point at  $m(K^+ K^-) = m(\phi)$ , and the red and blue point at  $m(K^+ K^-) = m(\bar{D}^0)$ .

### 5.6.3 Total efficiency

The total relative efficiency between the signal and normalisation decays is calculated from the products of each contributing efficiency

$$\epsilon^{\text{Tot.}} = \epsilon^{\text{Accp.}} \times \epsilon^{\text{Reco.}|\text{Accp.}} \times \epsilon^{\text{Trig.}|\text{Reco.}} \times \epsilon^{\text{Mass.}|\text{Trig.}} \times \epsilon^{\text{Veto.}|\text{Mass.}} \times \epsilon^{\text{FD}|\text{Veto.}} \\ \times \epsilon^{\text{IP}|\text{FD}} \times \epsilon^{\text{PID}|\text{IP}} \times \epsilon^{\text{MVA}|\text{PID}}, \quad (5.12)$$

where each relative efficiency  $x$  is defined relative to the previous selection step  $y$  as  $\epsilon^{x|y}$ . This is similarly determined as a function of the two-dimensional  $m^2(D_s^+ K^-)$  vs.  $m^2(K^+ K^-)$  space and shown in Fig. 5.17a. This two dimensional histogram has a rather coarse binning and is susceptible to variations from statistical fluctuation.

This could lead to biases when correcting the weights of signal candidates, especially if the candidates are concentrated within a one or a few bins. To reduce this the two dimensional histogram is made into a smoothly varying distribution by using a cubic spline interpolation as shown in Fig. 5.17b. The final distribution used to correct the yields of signal decays is the weighted sum of the splines for each year of data taking.

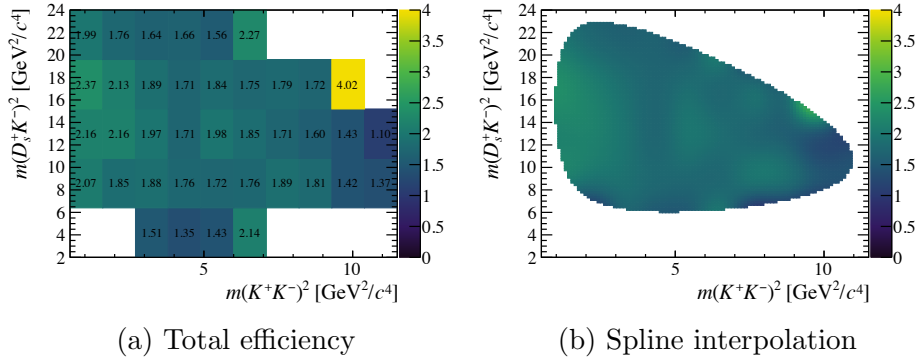


Figure 5.17: The total relative efficiency between the signal and normalisation mode as a function of the two-dimensional  $m^2(D_s^+ K^-)$  vs.  $m^2(K^+ K^-)$  (right) and a cubic spline interpolation of the same distribution.

## 5.7 Systematic uncertainties

A number of sources of systematic uncertainty are considered when determining the branching fraction for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays.

### 5.7.1 Relative efficiencies

The yields of signal and normalisation decays are corrected by the ratio of selection efficiencies. The limitations and assumptions surrounding this ratio contribute to the systematic uncertainty.

**Simulation statistics:** the relative efficiencies used to correct the signal and normalisation yields in this analysis are determined from samples of simulated decays. These are of a limited size, therefore the relative efficiencies have an associated statistical uncertainty. The size of the simulation samples results in a 2% systematic uncertainty.

**Particle identification:** the efficiencies for the particle identification requirements are determined from calibration samples using the PIDCALIB package. The procedure used to determine the efficiencies from these samples contains sources of systematic uncertainty, for example the choice of binning scheme used to match the calibration samples to the signal decays and the background-subtraction procedure used to isolate the distributions. The presence of the same particle



species in the signal and normalisation mode helps to reduce this affect in the ratio of efficiencies to some extent, but not completely. A relatively conservative estimate of the systematic uncertainty of 0.5% per track is assigned to account any residual differences. This is assumed to be uncorrelated for the five tracks contributing to the signal channel and five track contributing to the normalisation channel, resulting in a total relative uncertainty of 2.0%.

**Veto efficiency:** the efficiency of the misidentified  $D$  and  $\Lambda_c^+$  hadron veto applied to the  $D_s^+$  meson is assumed to be the same for the signal and normalisation channel. However, slight differences in the kinematics of the  $D_s^+$  mesons in each case might lead to this approximation not being valid when the particle identification requirements are tightened on the ambiguous track. The relative veto efficiency could be calculated by determining the relative PID efficiency for the sub-set of particles that would be affected by the tightening of the cuts. Rather than calculating this ratio, the systematic uncertainty that would be associated with it is assigned to account for any bias. This totals 1.4%, calculated in a similar way to the PID relative efficiency, except with three tracks instead of five.

**MVA efficiency:** the relative efficiency of the MVA selection receives contributions from a number of sources. Primarily, the relative efficiencies are determined from simulation samples in which the PID variable distributions have been corrected as described in Sec. 5.6.2. This method of correcting the PID distributions such that the simulation samples could be used directly was only available for Run I simulations at the time of analysis. The phase-space distributions of the Run II samples were assumed to be the same, but the overall scale adjusted using the MVA efficiencies derived from the  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decay modes used to train the MVAs. The systematic uncertainty arising from using the Run I variations in Run II are estimated by calculating the largest difference between the signal and normalisation efficiencies in

Run II and subtracting the smallest differences in Run I. This corresponds to a variation of 4.4%, assigned as the systematic uncertainty.

As they contribute to the scaling of the phase-space variations, all sourced of systematic uncertainty contributing to the data-driven efficiencies are also considered. The training modes are binned in four bins of  $p_T$  and  $\chi_{FD}^2$ . The choice of binning scheme is varied and the resulting variation in efficiency determined. The yields of the  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decays used to obtain the data-driven efficiencies are limited, therefore the quantity  $1/\sqrt{N}$  associated to the smallest of these yields is added as a systematic uncertainty.

Differences are observed in the distribution of  $\chi_{IP}^2$  in simulations and data for the normalisation channel (Fig. 4.18). To quantify the effect this may have on the MVA efficiencies, the simulations are re-weighted to match the data distributions and the efficiencies recalculated. The resulting difference is included as a systematic uncertainty.

When training the MVA methods some discrepancies are observed between the MVA classifier response for the training and validations samples (Fig. 4.15). To assess whether this affects the final relative MVA efficiencies the training and validation samples are swapped and the efficiencies recalculated. The differences for each year of data taking are weighted according to their contribution to the final data set and assigned as a systematic uncertainty.

The total systematic uncertainty associated to the relative MVA efficiency is therefore 7.6%.

### 5.7.2 Signal and normalisation PDFs

Some parameters in the signal and normalisation PDFs are fixed to values obtained from simulation. These include the tail parameters, relative widths, and fractional amounts of the two CB functions that make up the PDFs. The values obtained from simulation have associated uncertainties arising from the limited simulation sample sizes. The nominal fits are repeated with the fixed parameters modified to

values sampled from Gaussian distributions, with a width given by the parameter uncertainties. All parameters are changed simultaneously. The resulting variation of  $0.036 \times 10^{-6}$  is assigned as the associated systematic uncertainty.

### 5.7.3 Background PDFs

In the signal mode the PDFs for the partially reconstructed background modes are taken directly from simulated events using one-dimensional kernel estimations [36]. In the nominal fit, these are smeared to account for the differences in the mass resolution between data and simulation. To account for any systematic uncertainty arising from the choice of resolution difference, the fit is repeated, randomly varying the smearing resolution each time. The resulting variation in the branching fraction is assigned as a systematic uncertainty.

In the normalisation mode a number of choices are made about the partially reconstructed backgrounds, namely the kinematic limits of the  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$ , and the external branching fractions of  $D_s^{*+} \rightarrow D_s^+[\gamma/\pi^0]$ ,  $\bar{D}^{*0} \rightarrow \bar{D}^0[\gamma/\pi^0]$  decays. These assumptions are simultaneously varied within the relevant uncertainties, resulting in a change of  $0.015 \times 10^{-6}$  in the branching fraction.

### 5.7.4 Charmless contribution

The residual charmless and single charm contribution is only a significant fraction of the measured yield for the normalisation mode as detailed in Table 4.9. The residual yield of about eight candidates corresponds to a relative difference of 0.7% of the measured normalisation yield. This, however, would be a slight overestimation as the refitting of the decay chain with a constrained  $D_s^+$  mass would alter the resolution of the normalisation and charmless backgrounds differently. This has the effect of smearing the charmless candidates over a wider range. Rather than trying to include a PDF for this contribution or correcting the measured yield the expected size of the effect is used as the systematic uncertainty as it would propagate directly to the branching fraction.

### 5.7.5 Total systematic uncertainty

The systematic uncertainties contributed from each source discussed is detailed in Table 5.7. The total systematic uncertainty is also shown, along with the uncertainty arising from the externally measured normalisation channel branching fractions.

Source of Uncertainty	Systematic Uncertainty	
	Relative	Absolute ( $\times 10^{-6}$ )
MVA relative efficiency	7.6%	0.53
Simulation statistics	2.0%	0.14
PID relative efficiency	2.0%	0.14
Veto relative efficiency	1.4%	0.10
Charmless contribution	0.7%	0.05
Signal PDF parametrisation	-	0.036
Background PDF parametrisation	-	0.015
Total		0.59
Normalisation		0.70

Table 5.7: Contributions to the total systematic uncertainty of the  $B^+ \rightarrow D_s^+ K^+ K^-$  branching fraction measurement. The contributions are listed in descending order.

## 5.8 Results

The fit to  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates finds a total yield of  $N(B^+ \rightarrow D_s^+ K^+ K^-) = 443 \pm 29$  candidates. This constitutes the first observation of this decay mode. The branching fraction is calculated as

$$\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-) = \frac{N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)}{N(B^+ \rightarrow D_s^+ \bar{D}^0)} \times \mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0) \times \mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-) \quad (5.13)$$

where  $N(B^+ \rightarrow D_s^+ \bar{D}^0)$  is the yield of normalisation decays, and  $N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)$  is defined to be

$$N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-) = \sum_i \frac{W_i}{\epsilon_i^{\text{ratio}}}, \quad (5.14)$$

where  $W_i$  is the per-candidate weight, as determined by the *sPlot* technique for candidate  $i$ ; and  $\epsilon_i^{\text{ratio}}$  represents the relative efficiency of the signal and normalisation modes  $\epsilon_i(B^+ \rightarrow D_s^+ K^+ K^-)/\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)$  in the relevant bin of the  $B^+ \rightarrow D_s^+ K^+ K^-$  Dalitz plot as calculated in Sec. 5.6.

The uncertainty on the corrected yield,  $N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)$ , is in principle given by

$$\sigma(N_{\text{corr}}) = \sqrt{\sum_i \left( \frac{W_i}{\epsilon_i^{\text{ratio}}} \right)^2}. \quad (5.15)$$

However, the fit used to determine the  $sWeights$  only allows the yields to float, as opposed to the nominal fit that has additional floating parameters including the signal position and width. This means that this estimate of the corrected yield uncertainty could neglect the uncertainty due to the shape parameters: *i.e.* the uncertainty calculated from the weights,  $\sigma_{\text{yields only}}(N) = \sqrt{\sum W_i^2}$ , can be less than the uncertainty returned by the nominal fit,  $\sigma_{\text{fit}}(N)$ .

To correctly account for this possibility, the uncertainty from the shape parameters is separated from the the total uncertainty:  $\sigma_{\text{shape}}(N) = \sqrt{\sigma_{\text{fit}}(N)^2 - \sigma_{\text{yields only}}(N)^2}$ . This extra uncertainty is then scaled by the corrected yield to give the total uncertainty

$$\sigma_{\text{corr}}(N_{\text{corr}}) = \sqrt{\sigma(N_{\text{corr}})^2 + \left( \frac{N_{\text{corr}}}{N} \sigma_{\text{shape}}(N) \right)^2}. \quad (5.16)$$

For the nominal fit the uncertainties are summarised in Table 5.8.

Uncertainty	Value
$\sigma(N_{\text{corr}})$	13.0
$\sigma_{\text{fit}}(N)$	29.4
$\sigma_{\text{yields only}}(N)$	25.4
$\sigma_{\text{shape}}(N)$	14.7
$\sigma_{\text{corr}}(N_{\text{corr}})$	14.9

Table 5.8: The various uncertainties as detailed in Section 5.8 and their values in the nominal fit.

The corrected yield ratio can be expressed as the ratio of signal and normalisation branching fractions using Eq. 5.13. The value is measured to be

$$\frac{N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)}{N(B^+ \rightarrow D_s^+ \bar{D}^0)} = \frac{\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)}{\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0) \mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-)} = 0.197 \pm 0.015 \pm 0.017, \quad (5.17)$$

where the first uncertainty is statistical, and the second is systematic.

The branching fraction for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is determined to be

$$\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-) = (7.1 \pm 0.5 \pm 0.6 \pm 0.7) \times 10^{-6}, \quad (5.18)$$

where the first uncertainty is statistical, the second is systematic and the third from the branching fractions of  $\bar{D}^0 \rightarrow K^+ K^-$  and of the normalisation mode  $B^+ \rightarrow D_s^+ \bar{D}^0$ . The values used for the branching fractions are  $\mathcal{B}(D^0 \rightarrow K^+ K^-) = (4.01 \pm 0.07) \times 10^{-3}$  and  $\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0) = (9.0 \pm 0.9) \times 10^{-3}$  [27].

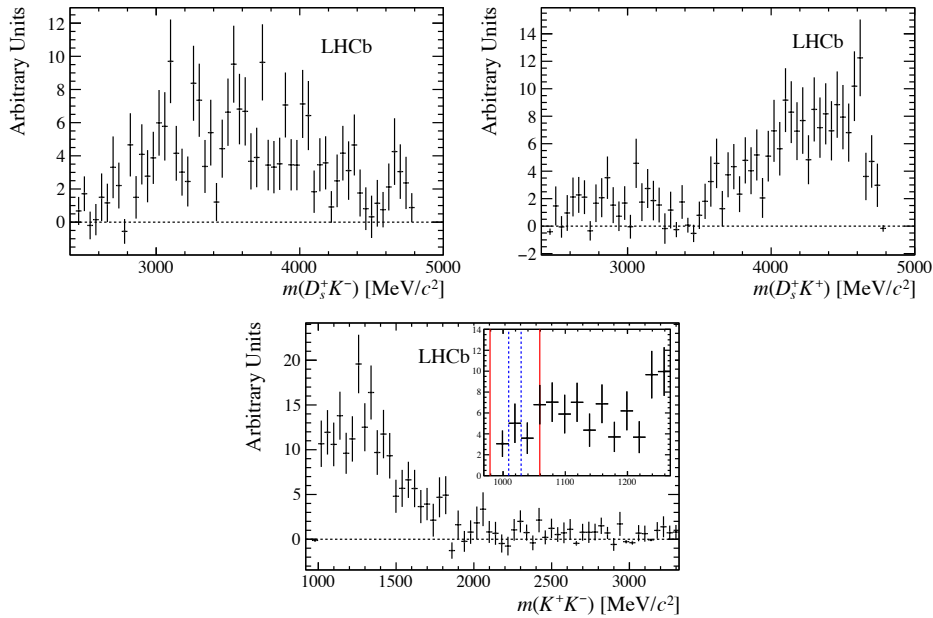


Figure 5.18: Two-body mass projections of the background-subtracted efficiency-corrected  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates. An inset has been included expanding the region around the  $\phi$  meson mass using the same ranges defined in the search for  $B^+ \rightarrow D_s^+ \phi$  decays in Fig. 6.1.

The two-body projections  $m(D_s^+ K^-)$  and  $m(K^+ K^-)$  are obtained for the signal component using the *sPlot* technique, shown in Fig. 5.18. No significant peak is observed in the  $\phi$  region of the  $m(K^+ K^-)$  plot; rather a broad distribution of candidates is found in the region up to  $m(K^+ K^-) \simeq 1900 \text{ MeV}/c^2$ . These plots have been corrected to account for the variation in the relative efficiencies as a function of the phase-space.

Additionally, the distribution of background-subtracted efficiency-corrected  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are shown as a function of the two-dimensional space  $m^2(D_s^+ K^-)$  vs.  $m^2(K^+ K^-)$  in Fig. 5.19. The candidates can be seen to be localised to a small area of the phase-space.

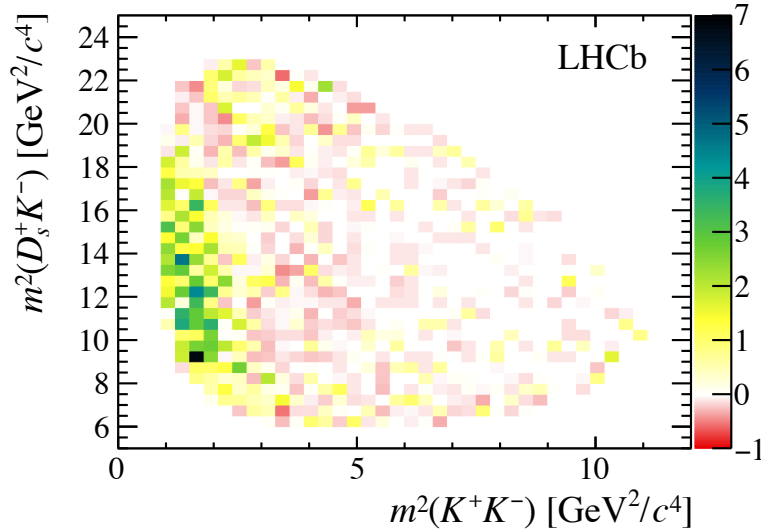


Figure 5.19: The Dalitz plot distribution of background-subtracted efficiency-corrected  $B^+ \rightarrow D_s^+ K^+ K^-$  decays.

The background-subtracted distribution of the angle,  $\cos\theta_K$ , defined to be the angle between the direction of the  $K^+$  meson and  $B^+$  momentum vector in the  $K^+ K^-$  rest frame is shown in Fig. 5.20. The distribution of this angle can change depending of the spin of the  $K^+ K^-$  system and therefore may help identify the contributing processes.

Additionally, the distribution of the angle  $\theta_K$  is plotted in bins of  $m(K^+ K^-)$  in Fig. 5.21. The distribution is observed to change as a function of  $m(K^+ K^-)$ , which may be a result of interference between competing processes.

## 5.9 Outlook

This analysis provides an initial insight to the possible processes that might contribute to the  $B^+ \rightarrow D_s^+ K^+ K^-$  decay. However, a full amplitude analysis would be needed to rigorously study the competing decays and account for any possible interference

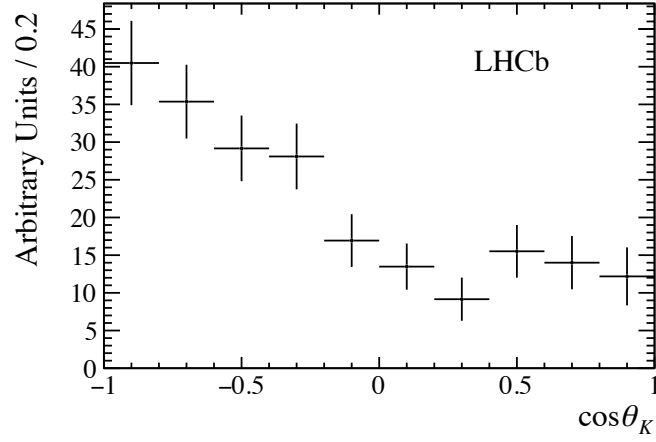


Figure 5.20: The background-subtracted efficiency-corrected distribution of  $\cos \theta_K$  for  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates.

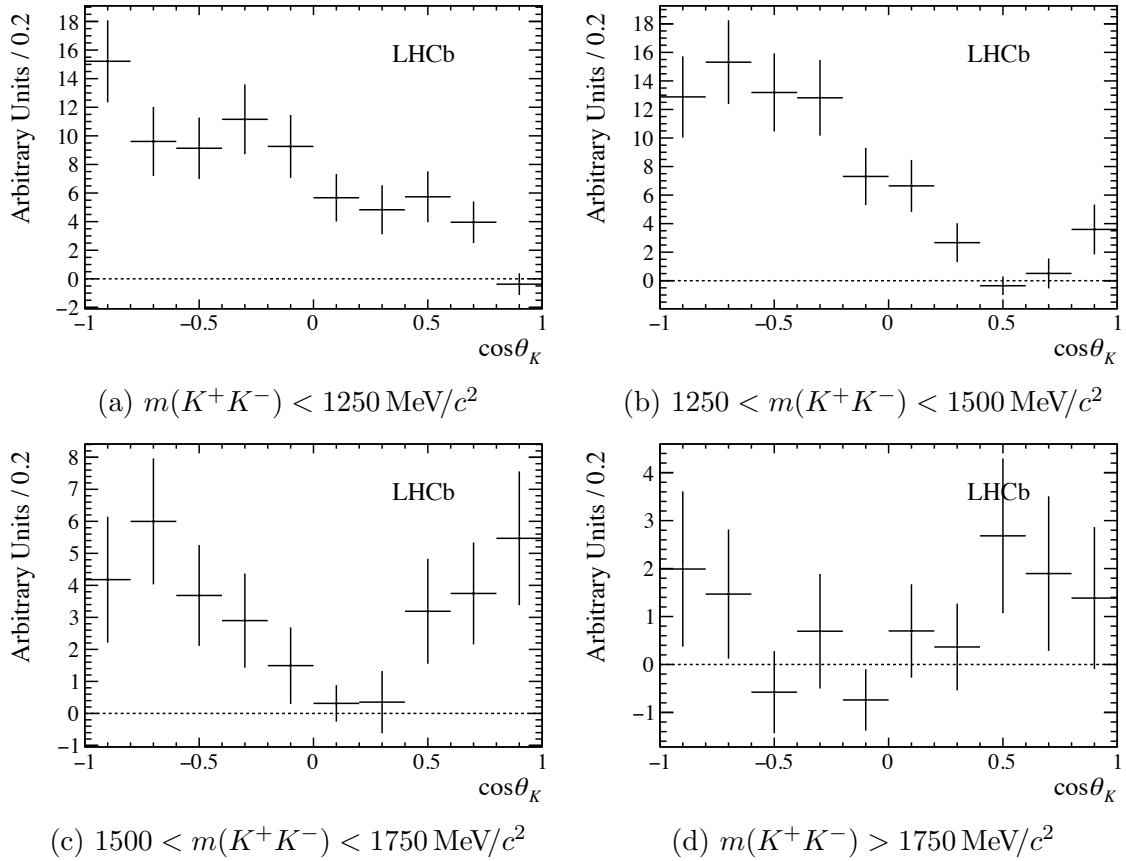


Figure 5.21: The distribution of  $\cos \theta_K$  in bins of  $m(K^+ K^-)$  mass.

between the amplitudes. This search contains only a fraction of the Run II data sample; the inclusion of the 2017 and 2018 samples would provide a statistically solid



base to further this line of investigation.