Determination of the CKM angle γ in $B^{\pm} \to DK^{\pm}, D\pi^{\pm}$ decays and strong phase determination of $D \to K^+K^-\pi^+\pi^-$ at BESIII

Martin Duy Tat

8th May 2021

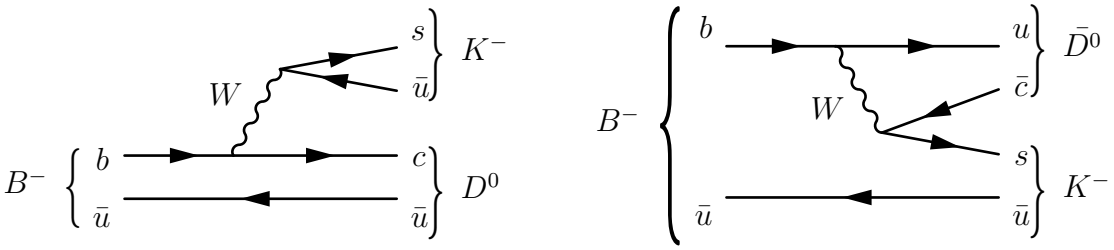
Abstract

Write abstract at the end

1 Introduction

In the Standard Model, CP-violation can be studied by by measuring the lengths and angles of the Unitary Triangle of the CKM matrix [1]. In particular, the angle $\gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)$ is the only angle that can be measured at tree level, with negligible theoretical uncertainties. Thus, a precise determination of γ is a good Standard Model benchmark which can be compared with indirect determinations from other CKM observables that may be sensitivty to new physics.

γ can be measured in processes with interference between $b \to c\bar{u}s$ and $b \to u\bar{c}s$ transitions, such as $B^{\pm} \to DK^{\pm}$ decays. D, a superposition of D^0 and $\bar{D}^0$, subsequently decays to a self-conjugate state. This is illustrated in Fig. 1. On the left, the colour favoured decay $B^- \to D^0K^-$ is shown, while on the right is the colour suppressed $B^- \to \bar{D}^0K^-$ decay. Interference is observed when D^0 and $\bar{D}^0$ decays to a common final state.



(a) $B^- \to D^0K^-$ **(b)** $B^- \to \bar{D}^0K^-$

Figure 1: Feynman diagrams of $B^- \to DK^-$ decays at tree level.

A wide range of subsequent D decays has been studied. Recently, the measurement $\gamma = (68.7^{+5.2}_{-5.1})^\circ$ from an analysis of $D \to K^0_S h^+h^-(h = \pi, K)$ was obtained [2], which is the single most precise γ measurement. In the following analysis, the subsequent decay $D \to K^+K^-\pi^+\pi^-$ is considered. An initial study [3] showed that a 14° precision is achievable with 1000 $B^{\pm} \to DK^{\pm}$ candidates. Considering similar channels, one expects that 2000 candidates can be reconstructed from the LHCb Run 1+2 dataset.

The challenge with the $D \to K^+K^-\pi^+\pi^-$ decay is the 5D four-body phase space, where the strong phase difference between the D^0 and $\bar{D}^0$ decays varies non-trivially. To

predict this strong phase, an amplitude model of the decay may be used. However, such a model introduces systematic uncertainties due to modelling.

In this analysis, a model-independent approach is chosen, in which strong phases are independently measured, in phase space bins, at the BESIII charm factory. The current BESIII 2010-11 dataset is insufficient, but significantly more data is expected from 2022.

A recently developed amplitude model for $D \to K^+K^-\pi^+\pi^-$, which is referred to as the LHCb model [4], will be used to develop an effective binning scheme. It is used to understand strong phase variations, but the final result will be model-independent. A poor binning scheme may decrease the statistical sensitivity, but will not bias the result. Thus, with a model-independent approach there is no modelling systematic uncertainty.

2 Formalism

2.1 γ sensitity through $B^\pm$ decays

The amplitude of $B^\pm \to DK^\pm$ is a coherent sum of the diagrams in Fig. 1,

$$\mathcal{A}(B^- \to DK^-) = \mathcal{A}_D + r_B^{DK} e^{i(\delta_B^{DK} - \gamma)} \mathcal{A}_{\bar{D}}, \tag{2.1}$$

$$\mathcal{A}(B^+ \to DK^+) = \mathcal{A}_{\bar{D}} + r_B^{DK} e^{i(\delta_B^{DK} + \gamma)} \mathcal{A}_D, \tag{2.2}$$

where r_B is the relative magnitude of the diagrams and $A_{D,\bar{D}}$ are D decay amplitudes as a function of phase space. δ_B^{DK} is the strong phase of the $B^\pm$ decay and is invariant under CP, while the weak phase γ swaps sign under CP.

The $B^\pm \to DK^\pm$ decay rate is considered in $2 \times N$ bins of phase space, labelled $i = -N, ..., N$, excluding zero. Bin i is related to $-i$ by a CP transformation. When integrating the square of Eqs. (2.1)-(2.2) over phase space Φ, the $B^- \to DK^-$ yield in bin i and $B^+ \to DK^+$ yield in bin $-i$ are

$$N_i^- = h_{B^-}\left[F_i + \left((x_-^{DK})^2 + (y_-^{DK})^2\right)\bar{F}_i + 2\sqrt{F_i\bar{F}_i}\left(x_-^{DK}c_i + y_-^{DK}s_i\right)\right], \tag{2.3}$$

$$N_{-i}^+ = h_{B^+}\left[F_i + \left((x_+^{DK})^2 + (y_+^{DK})^2\right)\bar{F}_i + 2\sqrt{F_i\bar{F}_i}\left(x_+^{DK}c_i + y_+^{DK}s_i\right)\right], \tag{2.4}$$

$$\begin{pmatrix} c_i \\ s_i \end{pmatrix} = \frac{\int_i \mathrm{d}\Phi\, |\mathcal{A}_D||\mathcal{A}_{\bar{D}}| \begin{pmatrix} \cos(\Delta\delta_D) \\ \sin(\Delta\delta_D) \end{pmatrix}}{\sqrt{\int_i \mathrm{d}\Phi\, |\mathcal{A}_D|^2 \int_i \mathrm{d}\Phi\, |\mathcal{A}_{\bar{D}}|^2}}, \quad F_i = \frac{\int_i \mathrm{d}\Phi\, |\mathcal{A}_D|^2}{\sum_i \int_i \mathrm{d}\Phi\, |\mathcal{A}_D|^2}, \quad \bar{F}_i = \frac{\int_i \mathrm{d}\Phi\, |\mathcal{A}_{\bar{D}}|^2}{\sum_i \int_i \mathrm{d}\Phi\, |\mathcal{A}_{\bar{D}}|^2}. \tag{2.5}$$

$h_{B^\pm}$ are a normalization constants. c_i (s_i) is the cosine (sine) of the strong phase difference $\Delta\delta_D$ between the D^0 and $\bar{D}^0$ decays, amplitude-averaged over bin i. F_i is the fractional yield of $B^- \to D^0K^-$ in bin i. Assuming CP conservation in D decays, $\bar{F}_i = F_{-i}$. Furthermore, the CP observables in Eqs. (2.1)-(2.2) are

$$x_\pm^{DK} = r_B^{DK}\cos\left(\delta_B^{DK} \pm \gamma\right), \quad y_\pm^{DK} = r_B^{DK}\sin\left(\delta_B^{DK} \pm \gamma\right). \tag{2.6}$$

By counting the number of $B^\pm \to DK^\pm$ decays in bins of phase space, one can do a Maximum Likelihood (ML) fit of Eqs. (2.3)-(2.4) to obtain the CP observables $x_\pm^{DK}$ and $y_\pm^{DK}$. These are then interpreted in terms of γ, δ_B^{DK} and r_B^{DK}. External inputs for c_i and s_i from BESIII are required for the ML fit.

In addition, F_i may either be obtained from flavour tagged D decays, or treated as free parameters in the fit. The second method is preferred to reduce systematic uncertainties. To constrain the F_i and improve the fit stability, the decay mode $B^\pm \to D\pi^\pm$, which has a similar selection, is included as a signal mode in the ML fit. This adds another set of Eqs. (2.1)-(2.2) with $K \to \pi$. The common topology means $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ will share the same F_i, but the second mode has much smaller CP-violation effects because $r_B^{D\pi}$ is much smaller than r_B^{DK}. The fit stability is improved by introducing the variables

$$x_\xi^{D\pi} = \mathrm{Re}(\xi^{D\pi}), \quad y_\xi^{D\pi} = \mathrm{Im}(\xi^{D\pi}), \quad \xi^{D\pi} = \frac{r_B^{D\pi}}{r_B^{DK}} e^{i(\delta_B^{D\pi} - \delta_B^{DK})}. \tag{2.7}$$

Therefore, the six CP observables in the ML fit are $x_\pm^{DK}$, $y_\pm^{DK}$, $x_\xi^{D\pi}$ and $y_\xi^{D\pi}$.

2.2 Strong phase and quantum correlations

c_i and s_i are measured in $\psi(3770) \to D^0\bar{D}^0$ decays using a double tag method. Since $\psi(3770)$ has charge conjugation $\mathcal{C} = -1$, the $D^0\bar{D}^0$ pair will be produced in a quantum correlated antisymmetric wavefunction.

The number of events where only one D meson is reconstructed as $D \to f$ is the single tag yield of f. If both D mesons are reconstructed in the signal mode $D \to K^+K^-\pi^+\pi^-$ and tag mode $D \to f$, one obtains the double tag yield.

If the tag is a flavour mode, such as $K^-\pi^+$, one can deduce the signal D flavour. The yield of flavour tagged $D^0 \to K^+K^-\pi^+\pi^-$ events in bin i is defined as K_i. The corresponding $\bar{D}^0$ decay belong in bin $-i$. c_i is measured in events where the tag mode is a CP eigenstate or a self-conjugate state with known CP-even fraction F_+. The yield of CP tagged $D \to K^+K^-\pi^+\pi^-$ events in bin i is given by

$$M_i^\pm = \frac{S_\pm}{2S_f}\Big[K_i - 2c_i(2F_+ - 1)\sqrt{K_iK_{-i}} + K_{-i}\Big], \tag{2.8}$$

where $S\pm$ and S_f are the single tag yields of the CP and flavour tag modes used, respectively. For CP even (odd) modes, $F_+ = 1$ (0). By measuring all the single and double tag yields, c_i is obtained from an ML fit to Eq. (2.8). To obtain s_i, the tag mode must be a self-conjugate mode, and an analogous ML fit is performed with the expression

$$M_{ij} = \frac{N_{D\bar{D}}}{2S_fS'_f}\Big[K_iK'_{-j} + 2\sqrt{K_iK'_{-j}K_{-i}K'_j}(c'_ic_j + s'_is_j) + K_{-i}K'_j\Big]. \tag{2.9}$$

$N_{D\bar{D}}$ is the number of $D^0\bar{D}^0$ pairs produced. Both f and f' are reconstructed in bins (i, j) of phase space. In the case $f = K^+K^-\pi^+\pi^-$ and $f' = K^0_S\pi^+\pi^-$, the $D \to K^0_S\pi^+\pi^-$ strong phases c'_i and s'_i are known [5]. If $f = f' = K^+K^-\pi^+\pi^-$, then $c'_i = c_i$ and $s'_i = s_i$.

Table 1 shows the tag modes in this analysis.

Table 1: Tag modes used in the BESIII double tag analysis. Subsequent decays are shown in parentheses. CP conjugates of flavour modes are implied.

Flavour	CP even	CP odd	Self conjugate
$K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$, $K^-e^+\nu_e$	K^+K^-, $\pi^+\pi^-$, $\pi^+\pi^-\pi^0$, $K^0_S\pi^0\pi^0$, $K^0_L\pi^0$, $K^0_L\eta(\gamma\gamma)$, $K^0_L\omega(\pi^+\pi^-\pi^0)$	$K^0_S\pi^0$, $K^0_S\phi$, $K^0_S\eta(\gamma\gamma, \pi^+\pi^-\pi^0)$, $K^0_S\omega(\pi^+\pi^-\pi^0)$, $K^0_S\eta'(\pi^+\pi^-\eta(\gamma\gamma), \pi^+\pi^-\gamma)$	$K^0_S\pi^+\pi^-$, $K^+K^-\pi^+\pi^-$

3 Detectors

3.1 LHCb

The LHCb [6] is a single arm forward spectrometer designed to study beauty and charm hadrons in pp collisions. The components important for this analysis are the high precision tracking system and the two Ring Imaging Cherencov counters (RICH1 and RICH2).

The tracking system includes the Vertex Locator (VELO). The VELO consists of silicon strip modules close to the interaction point, which provides high precision tracking and identification of displaced secondary vertices that are important for beauty and charm physics. A dipole magnet together with three tracking stations momentum of charged particles with an uncertainty of 0.5%-1.0%. The two RICH detectors, together with the tracking system and the calorimeter system, separates kaons from pions.

3.2 BESIII

BESIII [7] is a general purpose solenoidal detector. The main parts relevant to this analysis is the main drift chamber filled with Helium gas for measuring the momentum and $\frac{\mathrm{d}E}{\mathrm{d}x}$ of charged particles, a plastic scintillator based time of flight system for ID of charged particles and an electromagnetic calorimeter to measure neutral shower energies.

4 Binning scheme

In a three-body decay the 2D phase space is shown in a Dalitz plot. Ref. [2] separated the Dalitz space into bins of similar strong phase, such that the binning did not dilute the amplitude-averaged c_i and s_i. Bins with $i > 0$ and $i < 0$ were split such that Cabbibo favoured and suppressed resonances were in bins with opposite sign. This enhances the interference terms in Eqs. (2.3)-(2.4), and thus enhances the sensitivity to CP observables.

A binning scheme for the four-body decay $D \to K^+K^-\pi^+\pi^-$, where phase space is 5D, cannot easily be visualized. Instead, the LHCb model, implemented using the AmpGen [8], is used to predict the decay amplitude across phase space. It takes in the D daughter momenta and outputs the decay amplitude $\mathcal{A}(D)$. One can then define $\mathcal{A}(D^0)/\mathcal{A}(\bar{D^0}) \equiv r_D e^{i\Delta\delta_D}$, where $\Delta\delta_D$ and r_D are the strong phase difference and relative magnitude, respectively. A simple binning scheme is uniformly separated bins along $\Delta\delta_D$ such that decays with similar strong phases are in the same bins.

Under CP, $\Delta\delta_D \to -\Delta\delta_D$ and $\ln(r_D) \to -\ln(r_D)$. For $\ln(r_D) > 0$, the $\bar{D^0} \to K^+K^-\pi^+\pi^-$ decay is suppressed, relative to $D^0 \to K^+K^-\pi^+\pi^-$, while the converse is true for $\ln(r_D) < 0$. Therefore, the interference terms in Eqs. (2.3)-(2.4) may be enhanced if bins with $i > 0$ and $i < 0$ are split along $\ln(r_D) = 0$.

To optimize the binning, Q is defined as $x_\pm$ and $y_\pm$ sensitivity in a binned fit divided by that of an unbinned fit. It can be shown, with $N_i^\pm$ from Eqs. (2.3)-(2.4), that

$$Q^2 = \frac{1}{2}\left(Q_+^2 + Q_-^2\right), \quad Q_\pm^2 = 1 - \sum_i \frac{F_i F_{-i}\left(1 - c_i^2 - s_i^2\right)}{N_i^\pm}. \tag{4.1}$$

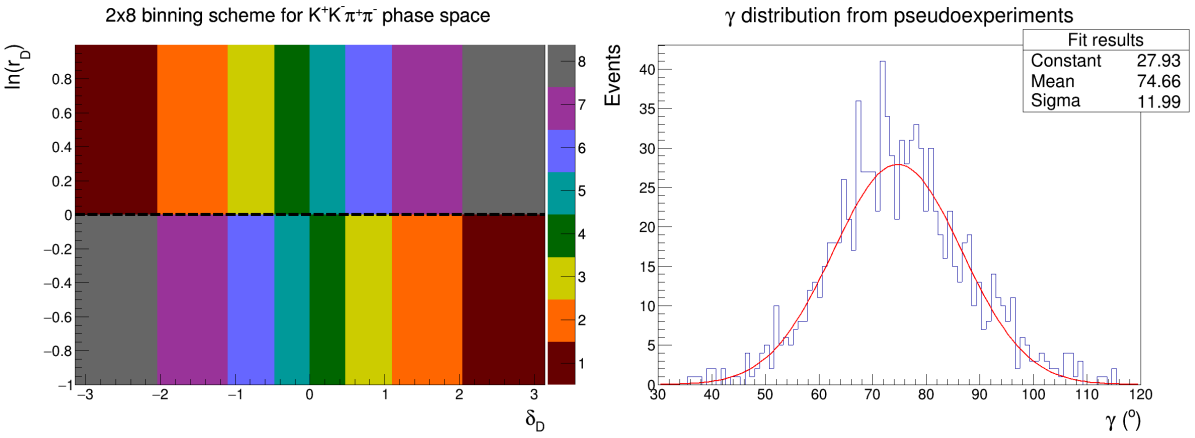
The bin boundaries are then moved symmetrically around $\Delta\delta_D = 0$ to maximize Q.

To assess the binning scheme efficiency, 1000 toy experiments, each with 2000 $B^\pm$ candidates, were generated with the LHCb model in AmpGen. The input parameters

used were $\gamma = 75°$, $\Delta_B = 130°$ and $r_B = 0.1$. An unbinned ML fit was performed to establish a benchmark for the γ precision, and the average precision of γ was $\Delta\gamma = 11°$.

With 2×8 bins, Fig. 2a shows the optimal binning scheme, with $Q = 0.90$, indicating that 10% sensitivity is lost due to binning. Using the binning in Fig. 2a, a ML fit was performed on each toy experiment using Eqs. (2.3)-(2.4) to extract $x_\pm^{DK}$ and $y_\pm^{DK}$. c_i, s_i and F_i were obtained from Monte Carlo integration of Eqs. (2.5) with $\mathcal{A}(D)$ from the LHCb model. Finally, γ, δ_B and r_B were obtained from $x_\pm^{DK}$ and $y_\pm^{DK}$ using Eqs. (2.6).

The pull distributions of $x_\pm^{DK}$, $y_\pm^{DK}$, γ, δ_B and r_B have mean and standard deviations consistent with zero and one, respectively, except for r_B, where the mean is $(1.72 \pm 0.32) \times 10^{-1}$. Similar behaviour has been found in similar analyses. Fig. 2b shows the γ distribution in the toy experiments. The binned fit γ precision achievable is $\Delta\gamma = (12.0 \pm 0.4)°$, which is consistent with the unbinned benchmark and $Q = 0.90$.



(a) Binning scheme definition for 2×8 bins **(b)** Distribution of γ from toy study

Figure 2

5 $B^\pm$ candidate selection

5.1 Signal candidate requirements

The $B^\pm \to Dh^\pm (h = \pi, K)$ candidates, with $D \to K^+K^-\pi^+\pi^-$, are reconstructed from five charged tracks. The standard track and trigger requirements follow Ref. [2]. To ensure that D is real, the $K^+K^-\pi^+\pi^-$ must be inside a 25 MeV mass window of the D^0 PDG mass. Then the tracks are refitted with their invariant mass constrained to the D^0 mass and their momenta pointing to the primary vertex. A cut on the χ^2 from this fit, at $\ln(\chi^2) < 3$, will remove most of the sidebands in the D invariant mass.

A mutually exclusive Particle Identification (PID) cut separates the $B^\pm \to D\pi^\pm$ and $DK^\pm$ candidates. Moreover, the $K^\pm$ daughters from D and the $h^\pm$ daughter from $B^\pm$ are required to have momentum $p < 100$ GeV, which is the optimal range for the RICH.

Sections 5.2-?? discuss more specialized selections to remove specific backgrounds.

5.2 Boosted Decision Tree

A Boosted Decision Tree (BDT) was used to remove the combinatorial background, which is the largest background. The BDT was trained using Monte Carlo (MC) $B^\pm$ samples as a

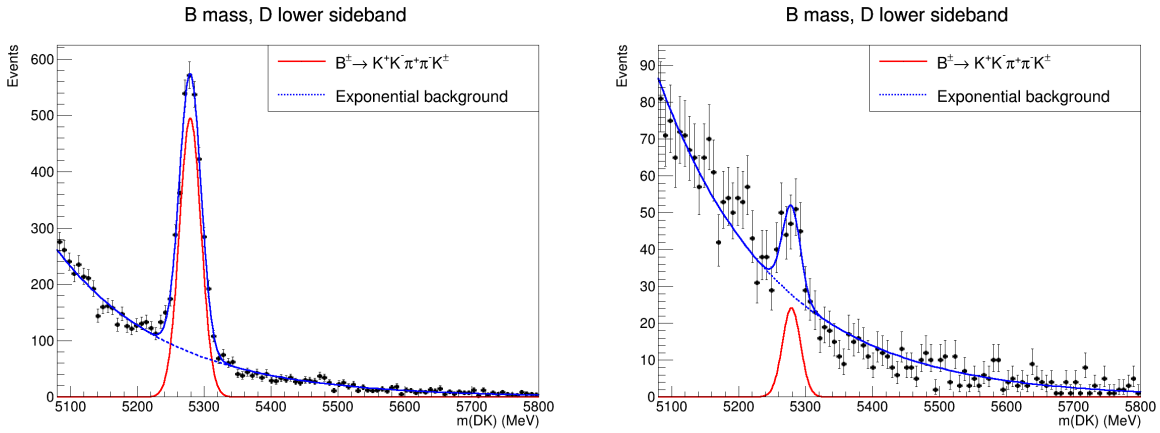
signal sample and the region $m(Dh^{\pm}) \in [5800, 7000]$MeV in data as a background sample. The training variables are similar to those in Ref. [2]. The input data was randomly split into equal training and test samples. After training, 99.4% of the background was removed and 93% of the signal remained in the test sample.

5.3 Mis-ID and charmless backgrounds

A significant mis-ID background are $B^{\pm} \to Dh^{\pm}$ candidates where $D \to K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}$ and a pion is assigned a kaon hypothesis. This is studied with MC samples. After reconstruction, the simulation indicate that 7.2% of $B^{\pm}$ candidates inside the signal region are $D \to K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}$. A tighter PID requirement reduces this fraction to 1.8% while keeping 93% of signal candidates.

Charmless backgrounds are $B^{\pm}$ candidates that decay into five charged particles without an intermediate D meson. These candidates form a signal in the $B^{\pm}$ mass spectrum, but does not peak in the D mass spectrum. The D mass window removes some of this, but a large fraction remains under the D peak.

To estimate the contamination, a selection was done without imposing a χ^2 cut on the refitted D daughters. This preserves the D mass sidebands. A mass window in the lower sideband, $m(K^+K^-\pi^+\pi^-) \in [1770\,\text{MeV}, 1820\,\text{MeV}]$ was selected. The $B^{\pm}$ mass spectrum is shown in Fig. 3a for $B^{\pm} \to DK^{\pm}$ candidates.



(a) The $B^{\pm}$ invariant mass in the D invariant mass lower sideband.

(b) The $B^{\pm}$ invariant mass in the D invariant mass lower sideband with a flight significance cut.

Figure 3

A fit to Fig. 3a with a Gaussian signal and an exponential background gives a yield of 2605 ± 57. A cut on the flight significance, defined as the flight distance divided by its error, reduces this yield to 110 ± 19, shown in Fig. 3b. No charmless peak was found in the corresponding $B^{\pm} \to D\pi^{\pm}$ invariant mass spectrum.

6 Fit to extract CP observables

6.1 Global fit and invariant mass spectra

A global maximum likelihood fit of all $B^{\pm}$ candidates was performed to determine the total yields and mass shape parameters in the $Dh^{\pm}$ invariant mass spectrum. The signal

and background parameterization are the same as those in Ref. [2], including which variables are floated in the fit and which are fixed. In addition, all fixed parameters are identical to the those in Ref. [2], and only minor adjustments are needed for the final analysis.

The combinatorial background is parameterized as an exponentially falling background. The signal is parameterized as a sum of a Gaussian $f_{\rm G}(m|m_B,\sigma)$ and a modified Gaussian that accounts for the radiate tail and wider resolution of signal events that are poorly reconstructed. The PDF has the form

$$f_{\rm MG}(m|m_B,\sigma,\alpha_L,\alpha_R,\beta) \propto \begin{cases} \exp\left(\frac{-\Delta m^2(1+\beta\Delta m^2)}{2\sigma^2+\alpha_L\Delta m^2}\right), \Delta m = m - m_B < 0 \\ \exp\left(\frac{-\Delta m^2(1+\beta\Delta m^2)}{2\sigma^2+\alpha_R\Delta m^2}\right), \Delta m = m - m_B > 0 \end{cases} . \quad (6.1)$$

To the left of the $B^\pm$ invariant mass peak are contributions from partially reconstructed backgrounds. These are mainly $B^\pm$ or B^0 decays to $D^*h^\pm$, and $D^* \to D\pi$ or $D^* \to D\gamma$, where the pion or photon are not reconstructed. The missing mass leads to peaking to the left of the signal. Furthermore, in the $B^\pm \to DK^\pm$, there is a component of $B^\pm \to D\pi^\pm$ candidates, where the $\pi^\pm$ daughter is wrongly assigned a kaon hypothesis. This is a small component to the right of the signal. Furthermore, this channel also has a component of $B_s \to DK^\pm\pi^\mp$ and the pion is not reconstructed, and a component of partially reconstructed background from $B^\pm$ and B^0 candidates where the $\pi^\pm$ daughter is mis-identified as a $K^\pm$. All partially reconstructed background components are fitted with PDF shapes described in Ref. [2].

The final global fits for $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ are shown in Fig. 4. The final yield of $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ candidates are 2290 ± 59 and $33\,113 \pm 211$, respectively.

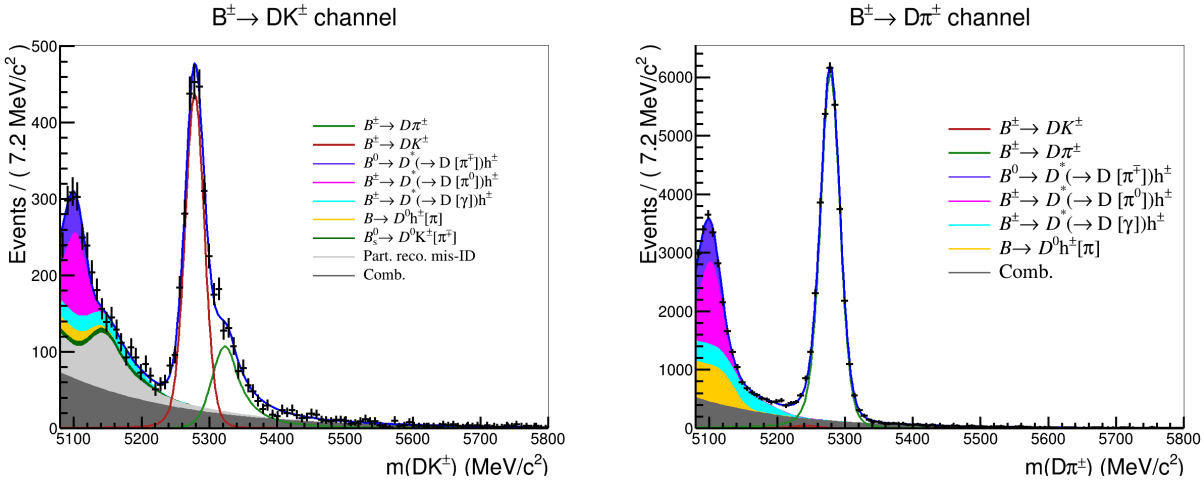


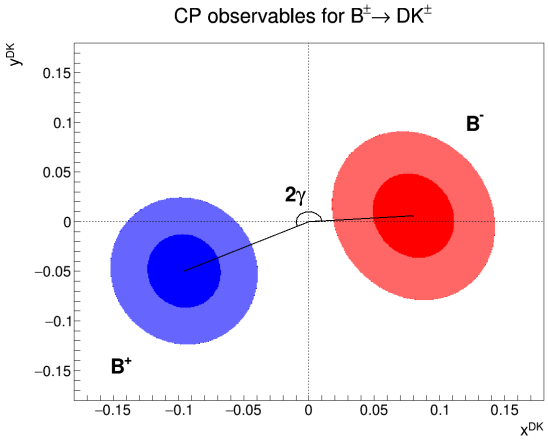
Figure 4: Global mass fit of $B^\pm \to DK^\pm$ (left) and $B^\pm \to D\pi^\pm$ (right).

6.2 Binned CP fit and CP observables

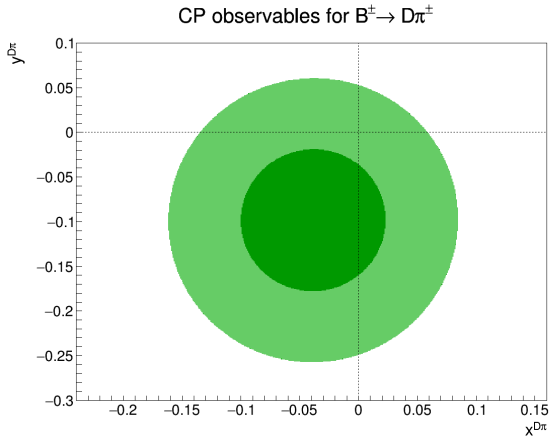
After the global fit described in Section 6.1, $B^\pm$ candidates are split by charge and phase space bins, using the binning scheme described in Section 4. A simultaneous maximum likelihood fit is performed for each charge and bin. All mass shape parameters are fixed

from the global fit while the yield of signal, total partially reconstructed background and combinatorial background in each bin are floated.

The final fitted parameters are shown in Figs. 5a-5b. Geometrically, the angle between the points (x_+^{DK}, y_+^{DK}) and (x_-^{DK}, y_-^{DK}) in Fig. 5a is 2γ, and these six CP observables may be interpreted in terms of γ, δ_B^{DK}, r_B^{DK}, $\delta^{D\pi}$ and $r_B^{D\pi}$. However, this step is only performed at the end to avoid any human bias in the final result.



(a) Confidence levels at 68.2% and 95.5% of (x_+^{DK}, y_+^{DK}) and (x_-^{DK}, y_-^{DK}).



(b) Confidence levels at 68.2% and 95.5% of $(x_\xi^{D\pi}, y_\xi^{D\pi})$.

6.3 Validation of fit procedure with toy studies

To ensure that the fit procedures are robust, for both the global and CP fit 1000 toy datasets are generated with the fitted parameters. Each toy datasets is then run through the same fitting procecure and the pull distibutions of each floated variable is studied closely. It was found that no CP observables were biased or had any understimated errors. Two nuisance parameters in the global fit had slightly underestimated errors which are easily corrected for.

7 External strong phase input from BESIII

Analysis of the BESIII data consists of e^+e^- collisions at the $\psi(3770)$ resonance. In addition, there exists inclusive simulated samples of each category of events. The relevant category for this analysis is $D^0\bar{D}^0$ events, which has 21.8 times more simulated events relative to data. Simulations of exclusive processes are generated when needed.

The most important aspects of this analysis is reconstructing single and double tagged event yields. Peaking backgrounds must be accounted for using the inclusive simulation samples. Finally, the yields are inputted into a maximum likelihood fit using Eqs. (2.8)-(2.9) to obtain c_i and s_i.

Since an e^+e^- collider is a very clean environment, charged and neutral particles mostly follow the same reconstruction strategy, described in Ref. [5]. Since the beams are symmetric, each D meson must have energy $E_{\rm beam}$. When the daughters of a D meson is reconstructed, the energy difference ΔE and beam-constrained mass $M_{\rm BC}$ may be defined as

$$\Delta E = \sum_i E_i - E_{\text{beam}}, \quad M_{\text{BC}} = \sqrt{E_{\text{beam}}^2 - \left|\sum_i \mathbf{p}_i\right|^2}.$$

For a perfectly reconstructed D meson, $\Delta E = 0$. However, due to detector resolution it is a peak of finite width. In addition, for tag modes with photon showers, the ΔE distribution has a left tail due to energy leakage through the calorimeter. The beam-constrained mass can be interpreted as the D invariant mass, but using the well measured beam energy to improve the resolution.

7.1 K_S^0 veto

The main peaking background in this analysis is $D \to K_S^0 K^+ K^-$, where $K_S^0 \to \pi^+\pi^-$. It has the same experimental signature as $D \to K^+K^-\pi^+\pi^-$ and therefore forms a signal peak in the M_{BC} distribution. Most of this background is removed with a flight significance cut at 2.0. The remaining background is from $K_S^0 K^+ K^-$ candidates where the K_S^0 decays close to the primary vertex, making it impossible to distinguish from $D \to K^+K^-\pi^+\pi^-$. This remaining background is therefore removed by imposing a veto on the $\pi^+\pi^-$ invariant mass. A sample of simulated $D \to K_S^0 K^+ K^-$ events was generated, but reconstructed as $D \to K^+K^-\pi^+\pi^-$ and the $\pi^+\pi^-$ invariant mass was studied. It was found that an asymmetric veto of $477\,\text{MeV} < m_{K_S^0} < 507\,\text{MeV}$ was more efficient in removing this peaking background. After the veto, a negligible amount of the peaking background remained.

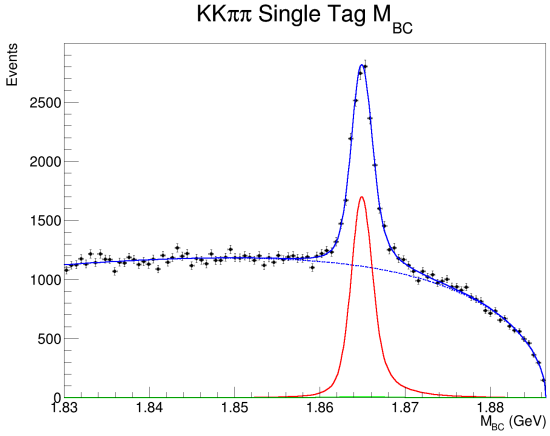
7.2 Single tag yield

The single tag yield is obtained by reconstructing events where one of the D mesons decay into a tag mode, while the other D meson is not reconstructed. A mode-dependent cut on ΔE is chosen by fitting a double Gaussian for the signal peak and a second order polynomial for the background in the ΔE distribution. A cut around $\pm 3\sigma$ is performed to remove combinatorial background. For modes with photon showers, the left cut is extended to -4σ.

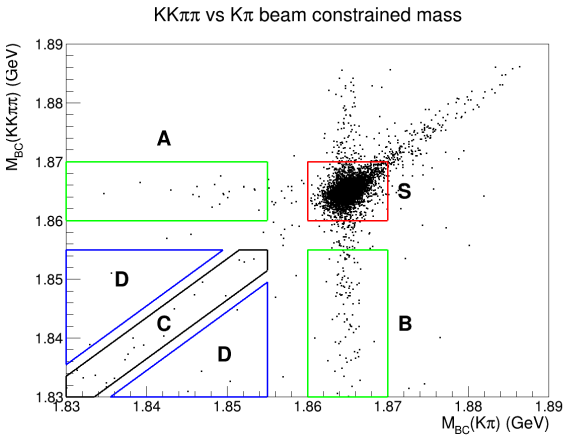
The single tag yield is then obtained by fitting the M_{BC} distribution. The continuous combinatorial background is modelled with an Argus shape. The signal shape is taken from a simulated sample of the signal, convolved with a Gaussian to account for detector resolution. Peaking backgrounds, if present, are accounted for with Gaussian shapes, with shape and yield fixed from the inclusive MC samples. An example of such a fit is shown in Fig. 6a.

7.3 Double tag yield

The double tagged yield is obtained by reconstructing both D mesons. The same cuts on ΔE used for single tagged events are applied, and the beam constrained mass for both D mesons are plotted in a 2D scatter plot, shown in Fig. 6b. Since the yields are expected to be too low for a maximum likelihood fit, a sideband subtraction technique is employed instead.



(a) Fit of single tagged $M_{\rm BC}$ distribution to obtain single tag yields.



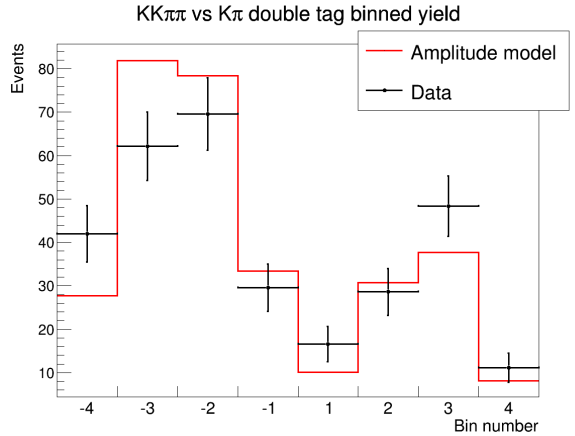
(b) Background subtraction technique of double tagged $M_{\rm BC}$ distributions to obtain double tag yields

In Fig. 6b, the data for double tagged $D \to K^+K^-\pi^+\pi^-$ versus $D \to K^\pm\pi^\mp$ events are shown. The $M_{\rm BC}$ region is split into 5 regions. Region S is the signal region, while A and B are events where one of the D mesons are correctly reconstructed. Region C are real events but daughter tracks between the D mesons have been swapped. Finally, region D contains purely combinatorial background. The total background is estimated using

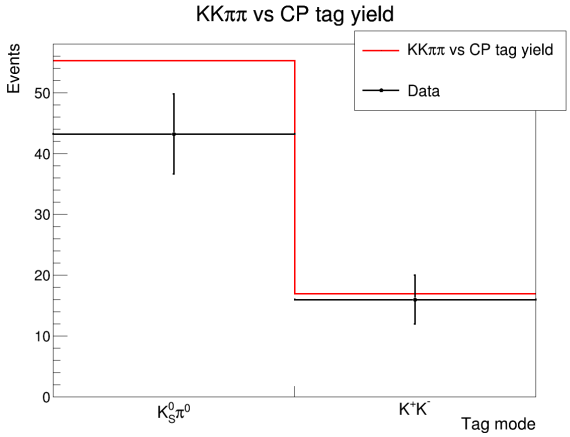
$$B = P + \frac{a_S}{a_D}Y_D + \sum_{i=A,B,C}\frac{a_S}{a_i}\left(Y_i - \frac{a_i}{A_D}Y_D\right),$$

where P is the peaking background found from inclusive MC, Y_i is the yield in region i and a_i is the corresponding area of the region.

Since the current dataset is insufficient for extracting c_i and s_i, it is more instructive to compare the double tag yields with the prediction from the LHCb amplitude model. In Fig. 7a, $D \to K^+K^-\pi^+\pi^-$ the yield of events tagged with the flavour tag $D \to K^\pm\pi^\mp$ are shown (data points), with the amplitude model prediction (solid line). The events have been binned using a 2×4 binning scheme as described in Section 4. In Fig. 7b, an equivalent plot for the $D \to K^+K^-\pi^+\pi^-$ events, tagged with a CP even tag mode $D \to K^+K^-$ and a CP odd tag mode $K^0_S\pi^0$, without binning. In both plots the yield is normalized by the single and double tag $K^\pm\pi^\mp$ yields.



(a) Binned double tag yield of $K^+K^-\pi^+\pi^-$ versus $K^\pm\pi^\mp$ events.



(b) Inclusive double tag yield of $K^+K^-\pi^+\pi^-$ versus K^+K^- and $K^0_S\pi^0$ events.

The LHCb model agrees reasonably well with the BESIII dataset. However, these are raw bin yields without any efficiency or bin migration corrections, and a perfect agreement is therefore not expected. In particular, $K^0_S\pi^0$ is expected to have a lower reconstruction efficiency than K^+K^- and $K^\pm\pi^\mp$ since it involves neutral particles, and therefore its double tag yield is lower than expected.

8 Discussion of future work

Discuss the future plans with Guy first!

References

[1] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49:652–657, 02 1973.

[2] R. Aaij et al. Measurement of the CKM angle γ in $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ decays with $D \to K^0_S h^+ h^-$. *JHEP*, 02:169, 2021.

[3] J. Rademacker and G. Wilkinson. Determining the unitarity triangle angle γ with a four-body amplitude analysis of $B^\pm \to (K^+K^-\pi^+\pi^-)_D K^\pm$ decays. *Physics Letters B*, 647(5):400–404, 2007.

[4] R. Aaij et al. Search for CP violation through an amplitude analysis of $D^0 \to K^+K^-\pi^+\pi^-$ decays. *JHEP*, 02:126, 2019.

[5] M. Ablikim et al. Improved model-independent determination of the strong-phase difference between D^0 and $\bar{D}^0 \to K^0_{S,L}K^+K^-$ decays. *Phys. Rev. D*, 102(5):052008, 2020.

[6] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[7] M. Ablikim et al. Future Physics Programme of BESIII. *Chin. Phys. C*, 44(4):040001, 2020.

[8] T. D. Evans. AmpGen. https://github.com/GooFit/AmpGen.

9 DPhil thesis plan

Discuss DPhil thesis plan with Guy first!