Selection Study for $\mathrm{B}^{\pm} \to \left(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm}\right)_{\mathrm{D}} \mathrm{K}^{\pm}$ and $\mathrm{B}^{\pm} \to \left(\mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}\pi^{-}\right)_{\mathrm{D}} \mathrm{K}^{\pm}$

LHCb-PUB-2009-002
02/09/2009

Public Note

Issue: 1
Revision: 3

Reference: LHCb-PUB-2009-002
Created: 15th April 2009
Last modified: 28th August 2009

Prepared by: Philip Hunt[a], Malcolm John[a]
[a]University of Oxford, United Kingdom

Abstract

In this document, a selection of the decay modes $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ and $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ is presented. It has been designed to minimise background in this channel, whilst retaining as much signal as possible. This analysis will be used as part of LHCb's effort to measure the CKM angle γ using tree-level processes. The expected signal yield in one nominal year of LHC running is estimated to be 554^{+272}_{-179} events, with a total background of 1750^{+1970}_{-1100}. The selection has also been applied to the similar channel $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ for which the signal and background yields are estimated to be 1350 ± 116 and 1000 ± 510 respectively. The signal and background yields for the favoured mode $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ are expected to be 53000 ± 3800 and 1830^{+1000}_{-1100}.

Contents

1 Introduction

One of the key aims of the LHCb experiment is to make precise measurements of the amount of CP violation in B decays, and compare with the Standard Model prediction. In 1963, Nicola Cabibbo extended the concept of universality of the different generations of leptons interacting by the weak interaction to the quark sector by suggesting that the weak interaction acts on an admixture of the mass eigenstates of the strong interaction [1]. This is quantified by a 'rotation' angle known as the Cabibbo angle which accounts for the mixing of the mass eigenstates of the down and strange quarks. A similar mixing occurs between the up and charm quarks. Kobayashi and Maskawa noted that such

a 2×2 quark mixing matrix could not allow for CP violation in the weak interaction, extending Cabibbo's theory to a third generation of quarks quantified by a 3×3 mixing matrix called the CKM matrix [2]. By allowing arbitrary rotations and imposing unitarity, the number of degrees of freedom of the CKM matrix reduces to three, plus one complex phase. It is this phase, known as the angle γ that allows CP violation in the Standard Model. This angle is poorly measured, with a current uncertainty from direct measurements of $\gamma = \left(70^{+27}_{-29}\right)^{\circ}$ [3]. LHCb plans to reduce this uncertainty to a few degrees within an integrated luminosity of 10 fb^{-1}, which is the expected total luminosity that will be available during LHCb's lifetime.

The angle γ can be extracted by considering B mesons decaying through both loop processes and tree processes. Both types of process are crucial in measuring γ. Loop processes contain at least one virtual loop, and are sensitive to new physics through the exchange of the Standard Model virtual particles in the decay with non-Standard Model particles, such as supersymmetric particles. Tree decays, on the other hand, contain no loops, and so can be used to measure γ from purely Standard Model processes.

This study follows Ref. [4], but uses a recent simulation that implements a more realistic LHCb detector description and improved reconstruction. The study considers the the tree decay $\mathrm{B}^{\pm} \to \mathrm{DK}^{\pm}$, where the D meson decays into the four-body states $\mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}\pi^{-}$ or $\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm}$. The aim of the study is to produce a set of cuts that can be used to reject background events, whilst retaining as much signal as possible.

The remainder of this section outlines the formalism of measuring γ with the decays $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ and $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$. Section 2 details the samples used for this study, and the normalisation used when optimising these samples. Section 3 describes the event selection, explaining some of the variables used in optimising the selection, and how correlations between some of these variables were accounted for. In the final sections, an estimation of background and signal yields for the aforementioned decay modes within the 2 fb^{-1} LHCb nominal year is calculated.

1.1 Determining γ from $\mathrm{B}^{\pm} \to \mathrm{DK}^{\pm}$

The angle $-\gamma$ is defined as the phase between the CKM-matrix elements V_{ub} and V_{cb}. Like any phase, its extraction depends upon interference of two amplitudes that produce the same final state. If the two processes are of similar magnitude, then the interference is large, and extraction of the phase information becomes easier.

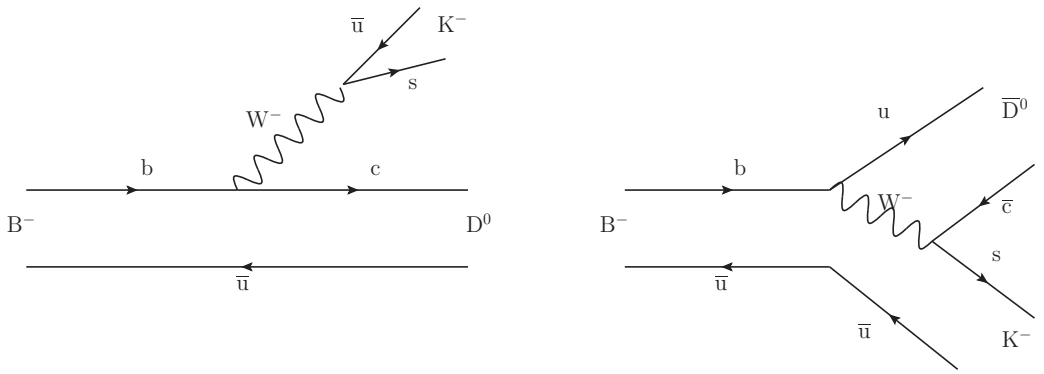

Figure 1 Feynman diagrams for the colour and CKM favoured / suppressed processes $\mathrm{B}^{-} \to \mathrm{D}^{0}\mathrm{K}^{-}$ and $\mathrm{B}^{-} \to \overline{\mathrm{D}}^{0}\mathrm{K}^{-}$.

Figure 1 shows the Feynman diagrams of a B^{-} decaying to $\mathrm{D}^{0}\mathrm{K}^{-}$ or $\overline{\mathrm{D}}^{0}\mathrm{K}^{-}$, which depend on V_{ub} and V_{cb} respectively. Thus, we will have interference which is sensitive to γ if the D^{0} and $\overline{\mathrm{D}}^{0}$ decay to the same final state. However, we can see from these diagrams that the latter has a suppressed branching ratio compared to the former, since the b quark decays more favourably to a c quark than a u quark, and is further suppressed by the allowed colour states of the quarks from the W^{-} decay. The decay amplitude of these processes may be written as follows:

$$A\left(\mathrm{B}^{-} \to \mathrm{D}^{0}\mathrm{K}^{-}\right) \equiv A_B, \tag{1}$$

and

$$A\left(\mathrm{B}^{-} \to \overline{\mathrm{D}}^{0}\mathrm{K}^{-}\right) \equiv A_B r_B e^{i(\delta_B - \gamma)}, \tag{2}$$

where r_B is the magnitude of the ratio of the suppressed amplitude to the favoured amplitude, and δ_B is a CP-conserving strong phase difference between these amplitudes. Equations 1 and 2 remain the same for a B^{+}, except for the sign of the weak phase transforming from $-\gamma \to \gamma$.

One method of extracting γ using this approach considers a D^0 or $\overline{\mathrm{D}}^0$ decaying to a CP-eigenstate, such as $\mathrm{K}^{+}\mathrm{K}^{-}$, $\pi^{+}\pi^{-}$ or $\mathrm{K}^0_{\mathrm{S}}\pi^0$. Since the D^0 and the $\overline{\mathrm{D}}^0$ will decay indistinguishably to one of these modes, the difference in the rates of B^{+} and B^{-} decaying to one of these final states, can be used to determine γ. Such a method is known as the Gronau, London and Wyler (GLW) method, after the original proposers [5, 6]. Equation 3 shows the rates for $\mathrm{B}^{-} \to \mathrm{DK}^{-}$ and $\mathrm{B}^{+} \to \mathrm{DK}^{+}$, where the D^0 or $\overline{\mathrm{D}}^0$ decay to a common CP-eigenstate f_{GLW}.

$$\Gamma\left(\mathrm{B}^{\pm} \to (f_{GLW})_{\mathrm{D}}\,\mathrm{K}^{\pm}\right) \propto 1 + r_B^2 + 2r_B\cos\left(\delta_B \pm \gamma\right) \tag{3}$$

This method is only efficient at extracting γ if the interference term is significant with respect to the other terms.

1.2 ADS Method

The ADS method, proposed by Atwood, Dunietz and Soni [7] considers processes in which the D^0 and $\overline{\mathrm{D}}^0$ decay to the same flavour-specific final state, such as $\mathrm{K}^{+}\pi^{-}$. Figure 2 shows the two neutral D decays resulting in the final state $\mathrm{K}^{-}\pi^{+}$; the second diagram is doubly Cabibbo suppressed compared to the first diagram by a factor λ^4, where λ is the sine of the Cabibbo angle.

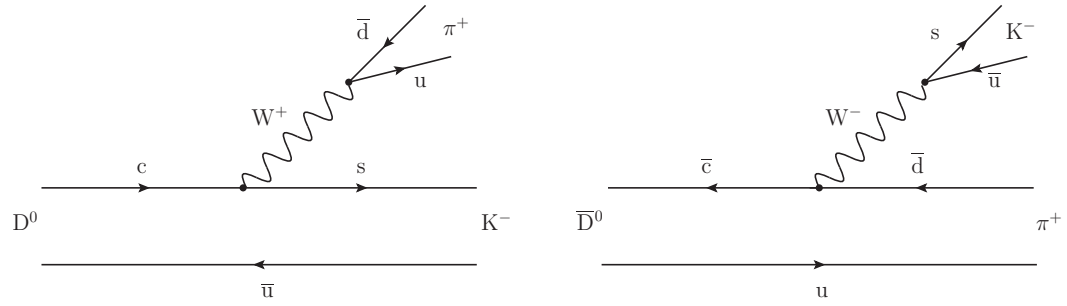

Figure 2 Feynman diagrams for the Cabibbo favoured / doubly Cabibbo suppressed processes $\mathrm{D}^0 \to \mathrm{K}^{-}\pi^{+}$ and $\overline{\mathrm{D}}^0 \to \mathrm{K}^{-}\pi^{+}$.

We can relate the amplitudes of these two decays by introducing the magnitude of the ratio between the suppressed and favoured amplitudes r_D , as well as a strong phase $\delta_D^{\mathrm{K}\pi}$.

$$A\left(\mathrm{D}^0 \to \mathrm{K}^{-}\pi^{+}\right) \equiv A_D^{\mathrm{K}\pi}, \tag{4}$$

and

$$A\left(\overline{\mathrm{D}}^0 \to \mathrm{K}^{-}\pi^{+}\right) \equiv A_D^{\mathrm{K}\pi} r_D^{\mathrm{K}\pi} e^{-i\delta_D^{\mathrm{K}\pi}} \tag{5}$$

The branching fractions for $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ and $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ are given by:

$$B\left(\mathrm{B}^{\pm} \to \left(\mathrm{K}^{\mp}\pi^{\pm}\right)_{\mathrm{D}}\mathrm{K}^{\pm}\right) \propto \left[r_B^2 + \left(r_D^{\mathrm{K}\pi}\right)^2 + 2r_B r_D^{\mathrm{K}\pi}\cos\left(\delta_B + \delta_D^{\mathrm{K}\pi} \pm \gamma\right)\right], \tag{6}$$

and

$$B\left(\mathrm{B}^{\pm} \to \left(\mathrm{K}^{\pm}\pi^{\pm}\right)_D \mathrm{K}^{\pm}\right) \propto \left[1 + \left(r_B r_D^{\mathrm{K}\pi}\right)^2 + 2 r_B r_D^{\mathrm{K}\pi} \cos\left(\delta_B - \delta_D^{\mathrm{K}\pi} \pm \gamma\right)\right], \tag{7}$$

Since $r_D^{\mathrm{K}\pi} = 0.0579 \pm 0.0007$ [8], which is comparable to $r_B = 0.087^{+0.022}_{-0.018}$ [3], the channel $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ is very sensitive to γ, and the rate can vary by $\sim 100\%$ depending on the value of γ and the strong phase differences. However, the branching ratio of $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\pm})_D \mathrm{K}^{\pm}$ is dominated by the favoured Feynman diagram, and so the interference in this channel is extremely small, albeit more abundant. Thus, the sensitivity to γ comes from the decays in which the kaons from the $\mathrm{B}^{\pm}$ and D decays have opposite signs.

1.3 Extending the ADS Method to Multi-Body D Decays

We can consider additional non-CP final states, such as $\mathrm{K}^{+}\pi^{+}\pi^{-}\pi^{+}$ to further constrain γ [9]. However, for multi-body modes, the situation is more complicated than the two-body case, since the D can decay through several resonant states, with a variety of amplitudes and strong phases. As with the two-body method, we consider the decays $\mathrm{B}^{-} \to \mathrm{DK}^{-}$ and $\mathrm{B}^{+} \to \mathrm{DK}^{+}$, where the D^0 or $\overline{\mathrm{D}}^0$ decay to the same final state f_D. The matrix element squared of the processes $\mathrm{B}^{-} \to f_{\mathrm{D}}\mathrm{K}^{-}$ and $\mathrm{B}^{+} \to f_{\mathrm{D}}\mathrm{K}^{+}$ are given by:

$$\mathcal{M}^2(\mathrm{B}^{-} \to f_{\mathrm{D}}\mathrm{K}^{-})(\mathbf{x}) = |A_B|^2 \left[|A_D(\mathbf{x})|^2 + r_B^2 |A_{\bar{D}}(\mathbf{x})|^2 + 2 r_B |A_D(\mathbf{x})|\,|A_{\bar{D}}(\mathbf{x})| \cos\left(\delta_B + \zeta(\mathbf{x}) - \gamma\right)\right], \tag{8}$$

and

$$\mathcal{M}^2(\mathrm{B}^{+} \to f_{D}\mathrm{K}^{+})(\mathbf{x}) = |A_B|^2 \left[r_B^2 |A_D(\mathbf{x})|^2 + |A_{\bar{D}}(\mathbf{x})|^2 + 2 r_B |A_D(\mathbf{x})|\,|A_{\bar{D}}(\mathbf{x})| \cos\left(\delta_B + \zeta(\mathbf{x}) + \gamma\right)\right], \tag{9}$$

where $\zeta(\mathbf{x}) = \arg\left(A^*_{\mathrm{D}^0}(\mathbf{x})\, A_{\overline{\mathrm{D}}0}(\mathbf{x})\right) = \arg\left(A_{\overline{\mathrm{D}}0}(\mathbf{x})\, A^*_{\mathrm{D}^0}(\mathbf{x})\right)$, A_B is the amplitude for the colour favoured decay $\mathrm{B}^{-} \to \mathrm{D}^0\mathrm{K}^{-}$, and $A_D(\mathbf{x})$ and $A_{\bar{D}}(\mathbf{x})$ are the amplitudes of D^0 and $\overline{\mathrm{D}}^0$ decaying to the state f_D respectively, at a given point $\mathbf{x}$ in the D decay phase space. The total rates for these processes can be found by integrating equations 8 and 9 over all points in the phase space. It can be shown that the results are simplified by defining a *coherence factor*, $0 \leq R_f \leq 1$, which quantifies the dilution of the total interference due to the presence of many competing resonances. The rates of the B^{-} and B^{+} decays are defined by:

$$\Gamma\left(\mathrm{B}^{-} \to f_{\mathrm{D}}\mathrm{K}^{-}\right) \propto A_f^2 + r_B^2 \bar{A}_f^2 + 2 r_B R_f A_f \bar{A}_{f_D} \cos\left(\delta_B + \delta_D^f - \gamma\right), \tag{10}$$

and

$$\Gamma\left(\mathrm{B}^{+} \to f_{\mathrm{D}}\mathrm{K}^{+}\right) \propto r_B^2 A_f^2 + \bar{A}_f^2 + 2 r_B R_f A_f \bar{A}_f \cos\left(\delta_B + \delta_D^f + \gamma\right), \tag{11}$$

where:

$$A_f^2 = \int \left|A_{\mathrm{D}^0}(\mathbf{x})\right|^2 d\mathbf{x}, \quad \bar{A}_f^2 = \int \left|A_{\overline{\mathrm{D}}0}(\mathbf{x})\right|^2 d\mathbf{x}, \tag{12}$$

are the partial widths for D^0 and $\overline{\mathrm{D}}^0$ decaying to the final state f_D respectively, and

$$R_f e^{i\delta_D^f} = \frac{\int \left|A_{\mathrm{D}^0}(\mathbf{x})\right| \left|A_{\overline{\mathrm{D}}0}(\mathbf{x})\right| e^{i\zeta(\mathbf{x})} d\mathbf{x}}{A_f \bar{A}_f}, \tag{13}$$

The phase difference δ_D^f can be considered as the average of the strong phase differences contributing to the final state over all phase space. If there is only one contribution to the decay $\mathrm{D} \to f_D$, then the phase term factors out of equation 13, the coherence factor becomes unity, and the equation simplifies to the two-body results in the previous section.

From equation 10, and 11, we can see that the rates of the suppressed four-body decays are:

$$\Gamma\left(\mathrm{B}^{\pm} \to \left(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm}\right)_{\mathrm{D}} \mathrm{K}^{\pm}\right) \propto \left[r_B^2 + \left(r_D^{\mathrm{K}3\pi}\right)^2 + 2r_B r_D^{\mathrm{K}3\pi} R_{\mathrm{K}3\pi} \cos\left(\delta_B + \delta_D^{\mathrm{K}3\pi} \pm \gamma\right)\right] \tag{14}$$

where $R_{\mathrm{K}3\pi}$ is the coherence factor for $\mathrm{D}^0 \to \mathrm{K}^-\pi^+\pi^-\pi^+$. Results from the CLEO-c experiment measure the four-body coherence factor as $R_{\mathrm{K}3\pi} = 0.33^{+0.20}_{-0.23}$ and the strong phase as $\delta_D^{\mathrm{K}3\pi} = \left(114^{+26}_{-23}\right)^\circ$ [10]. Using the values in Table 1 and their asymmetric errors as input parameters, we run a series of Monte Carlo simulations and estimate the branching fraction to be $\left(3.4^{+1.3}_{-1.1}\right) \times 10^{-7}$.

Input Parameter	Source	Value
γ	CKMFitter 2008 results [3]	$\left(70^{+27}_{-29}\right)^\circ$
r_B	CKMFitter 2008 results	$0.087^{+0.022}_{-0.018}$
δ_B	CKMFitter 2008 results	$\left(110^{+22}_{-27}\right)^\circ$
$r_D^{\mathrm{K}3\pi}$	CLEO-c [a]	0.0568 ± 0.0020
$\delta_D^{\mathrm{K}3\pi}$	CLEO-c [10]	$\left(114^{+26}_{-23}\right)^\circ$
$R_{\mathrm{K}3\pi}$	CLEO-c	$0.33^{+0.20}_{-0.23}$
$B\left(\mathrm{B}^- \to \mathrm{D}^0\mathrm{K}^-\right)$	Particle Data Group [12]	$(4.02 \pm 0.21) \times 10^{-4}$
$B\left(\mathrm{D}^0 \to \mathrm{K}^-\pi^+\pi^-\pi^+\right)$	Particle Data Group	$(8.10 \pm 0.20) \times 10^{-2}$

Table 1 Input parameters used for the determination of the branching fraction of $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})\,\mathrm{K}^{\pm}$.

[a] Value used as input by CLEO-c when calculating $R_{\mathrm{K}3\pi}$ and $\delta_D^{\mathrm{K}3\pi}$ [10, 11], and includes mixing parameters (see review of $\mathrm{D}^0 - \overline{\mathrm{D}}^0$ mixing [12]).

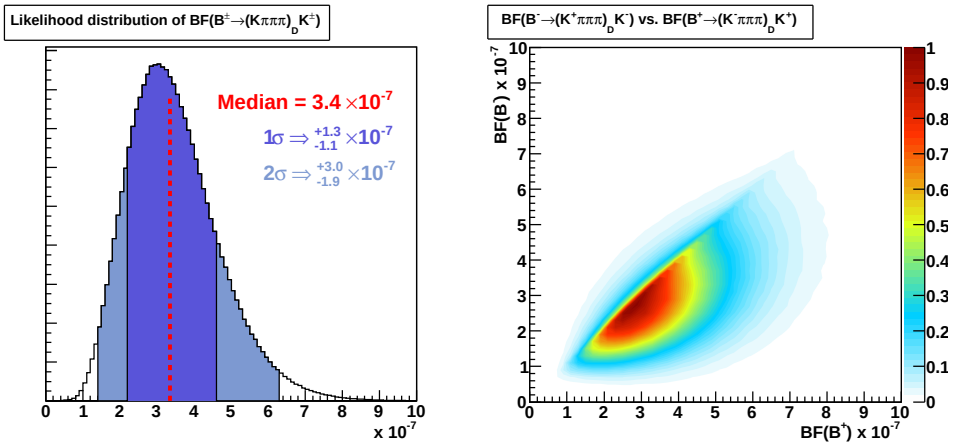

Figure 3 A frequentist estimation of the $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})\,\mathrm{K}^{\pm}$ branching fraction based on Equation 14 and the current world average values for the input parameters. On the left is the CP-average branching fraction including B^+ and B^- decays, and on the right is a 2D plot of B^- versus B^+ demonstrating the possibility of CP violation (away from the diagonal). This formulation neglects any correlations between variables.

1.4 Self-Conjugate, Multi-Body D Decays

Another method for extracting γ comes from the decay of the $\mathrm{D}^0/\overline{\mathrm{D}}^0$ to a multi-body self-conjugate state, such as $\mathrm{B}^{\pm} \to \left(\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-\right)_{\mathrm{D}} \mathrm{K}^{\pm}$ and $\mathrm{B}^{\pm} \to (\mathrm{K}^+\mathrm{K}^-\pi^+\pi^-)_{\mathrm{D}} \mathrm{K}^{\pm}$. This differs from the GLW approach in that as was the case for the multi-body extension to the ADS method, the final-states can occur via several possible resonances. Typically, γ is extracted from the decay $\mathrm{B}^{\pm} \to \left(\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-\right)_{\mathrm{D}} \mathrm{K}^{\pm}$ by fitting a Dalitz plot, assuming a particular model for the resonance structure.

In this technique, the amplitude of the decays $\mathrm{D} \to \mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ are assumed to have the following form

$$A_{\mathrm{D}\to\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-} = a_0 e^{i\delta_0} + \sum_n a_n e^{i\delta_n} A_n(m^2_+,\ m^2_-) \tag{15}$$

where $m^2_+ = \left(p_{\mathrm{K}^0_{\mathrm{S}}} + p_{\pi^+}\right)^2$ and $m^2_- = \left(p_{\mathrm{K}^0_{\mathrm{S}}} + p_{\pi^-}\right)^2$, a_0 is a constant non-resonant amplitude, with an associated phase δ_0, and a_n are the amplitudes of the resonances, with an associated phase δ_n and a Breit-Wigner function A_n. If there is no CP violation, then the Dalitz plot for $\mathrm{D}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ will be the same as $\overline{\mathrm{D}}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ under the transformation $m^2_+ \leftrightarrow m^2_-$.

Recently, a model-independent binned Dalitz fitting technique, first suggested by Giri et. al. [13], has been used to measure γ from multi-body D decays at CLEO-c [14]. This involves splitting a Dalitz plane into bins. The bins are constructed such that they are symmetric about the point $m^2_+ = m^2_-$. For each bin, we can construct a quantity similar to the coherence factor in equation 13. In this case, a double integration is performed over the bin contents. This represents the weighted average of $e^{i\Delta\delta_D}$ over the particular bin, where $\Delta\delta_D$ is the strong phase difference between the D^0 and the $\overline{\mathrm{D}}^0$ decays. We can then look for a difference between the bin content of the D^0 and $\overline{\mathrm{D}}^0$ plots as with the unbinned case. The precision on γ from a binned Dalitz technique has been investigated at LHCb [15], as well as for the model-dependent case [16].

These Dalitz techniques can also be extended to the four-body mode $\mathrm{B}^{\pm} \to (\mathrm{K}^+\mathrm{K}^-\pi^+\pi^-)_{\mathrm{D}} \mathrm{K}^{\pm}$, but this is more complicated than the three-body case, since we need five variables to fully describe the kinematics of the system rather than two. However, it does have the advantage that all of the final state particles are charged tracks. A model-dependent fitting technique has been considered for $\mathrm{B}^{\pm} \to (\mathrm{K}^+\mathrm{K}^-\pi^+\pi^-)_{\mathrm{D}} \mathrm{K}^{\pm}$ [17], and it has been estimated that LHCb can achieve an uncertainty on γ of 15° at LHCb.

For the final estimate of the signal yield for the decay modes $\mathrm{B}^{\pm} \to (\mathrm{K}^+\mathrm{K}^-\pi^+\pi^-)_{\mathrm{D}} \mathrm{K}^{\pm}$, no attempt has been made to estimate the separate branching fractions of the B^+ and B^-, and the total branching fraction is estimated by multiplying the average branching fractions of $\mathrm{B} \to \mathrm{DK}$ and $\mathrm{D} \to \mathrm{KK}\pi\pi$ reported in Ref. [12] , which gives $(9.8 \pm 0.7) \times 10^{-7}$.

2 Data Samples

This section describes the LHCb Monte Carlo data samples that are used to optimise the selection, and how these are normalised to a nominal year of LHC running.

2.1 Background Samples

The selection is optimised using ~22 million simulated inclusive $\mathrm{b\overline{b}}$ events. The LHCb generation and simulation software (Gauss v25r10) was used to generate $\mathrm{b\overline{b}}$ quarks produced in a proton-proton interaction at 14 TeV, requiring that at least one of the resulting hadrons has an angle of less than 400 mrad in the beam direction . The generator efficiency for this sample is $(43.7 \pm 0.1)\,\%$.

After the generation stage, the Gauss software uses GEANT4 [18, 19] to simulate the passage of the particles in the event through the detector, based on the current detector description. A digitisation step (using Boole v12r10) processes the output and simulates the detector response and readout electronics. Finally, the events are reconstructed using the offline reconstruction software (Brunel v30r17).

Of the $\sim 2.2 \times 10^7$ total $\mathrm{b\overline{b}}$ events available, 1.13×10^7 were used to optimise the selection, whilst the remaining 1.07×10^7 were reserved to make a subsequent unbiased estimate of the background.

2.2 Signal Samples

The signal samples were produced using a compatible version of the simulation software (Gauss v25r7). The generator cut, which requires the B daughters to be in the acceptance has an efficiency of $(17.5 \pm 0.1)\,\%$ for the $\mathrm{B} \to (\mathrm{K}\pi\pi\pi)_{\mathrm{D}}\,\mathrm{K}$ signal, and $(18.4 \pm 0.1)\,\%$ for $\mathrm{B} \to (\mathrm{KK}\pi\pi)_{\mathrm{D}}\,\mathrm{K}$ signal. After the generator cut, the samples available for this study were 50,676 $\mathrm{B} \to (\mathrm{K}\pi\pi\pi)\,\mathrm{K}$ events and 49,730 $\mathrm{B} \to (\mathrm{KK}\pi\pi)\,\mathrm{K}$ events.

2.3 Normalisation

LHCb is expected to receive an integrated luminosity of $2\ \mathrm{fb}^{-1}$ in one nominal year of running ($10^7\mathrm{s}$). Assuming a $\mathrm{b\overline{b}}$ cross-section of 500 µb, we expect $10^{12}\ \mathrm{b\overline{b}}$ events to be produced in a nominal year. Remembering the generator efficiency for $\mathrm{b\overline{b}}$ is 0.437, one estimates $4.37 \times 10^{11}\ \mathrm{b\overline{b}}$ events will occur within the LHCb acceptance. The optimisation performed uses 1.13×10^7 inclusive $\mathrm{b\overline{b}}$ Monte Carlo events, which represents one part in $4.37 \times 10^{11}/1.13 \times 10^7 = 38,700$. In terms of running time, the background sample represents 4.3 min. of LHCb operation. We refer to this figure as the background normalisation factor and it is used to estimate the background in a dataset containing one nominal year of events. A second sample of 1.07×10^7 is set aside for a subsequent unbiased estimate of the background, the normalisation factor for this sample is $40,800$.

Similarly, the signal events were normalised to 2 fb^{-1} when optimising the event selection. It is assumed that a $\mathrm{B}^{\pm}$ hadronises from each b and $\overline{\mathrm{b}}$ quark 40.5% of the time, so $8.1 \times 10^{11}\ \mathrm{B}^{\pm}$ would be produced in a dataset corresponding to an integrated luminosity of 2 fb^{-1}. In section 1.3, the branching fraction is estimated to be 3.4×10^{-7}, which leads to an estimate of $275,000\ \mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ events produced per nominal year. Given the generator efficiency of 0.175, one expects $48,200\ \mathrm{B}^{\pm}$ events within the LHCb acceptance each nominal year. The Monte Carlo sample contains $50,676$ events, so the normalisation factor is 0.95. For the channels $\mathrm{B}^{\pm} \to (\mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}\pi^{-})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ and $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$, the equivalent calculations give normalisation factors of 2.93 and 91 respectively.

3 Event Selection

The reconstructed signal and background samples were processed using the LHCb analysis software (DaVinci v20r3). The candidates passing a loose preselection are selected, and the event information is stored in ROOT ntuples [20]. The size of the background Monte Carlo sample is small compared to the real data. To partially compensate for this, an artificially large window in the reconstructed $\mathrm{B}^{\pm}$ mass of the background candidates is permitted. Nevertheless, not wanting to overestimate the background level, background events of category *low mass* are removed, as well as $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$ backgrounds, and *partially reconstructed physics background*, which are considered later. It is assumed that what remains is dominated by combinatoric background events, and the optimisation is run on this sample. In addition, the combinatoric background in the favoured and suppressed modes is expected to be topologically identical, so the background from both modes are allowed (any combination of kaon charge in the candidate), doubling the available statistics (and halving the normalisation factor).

The selection of $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ is designed to maximise the signal significance, given by $S/\sqrt{S+B}$, where S and B are the number of signal and background events respectively, normalised to $2\mathrm{fb}^{-1}$. The variables which appear to give the best signal-background separation are chosen. A full list of selection cuts is shown in Section 4.1. The variables used in the selection are explained in Section 3.1.

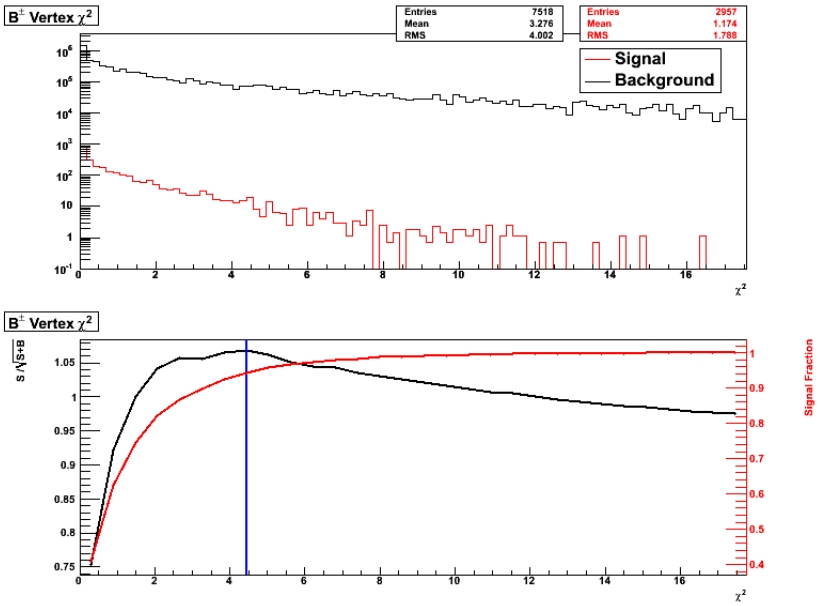

Figure 4 Optimisation plots of $\mathrm{B}^{\pm}$ vertex χ^2 . The first plot shows the distribution of signal and background, normalised to 2 fb^{-1}. The second plot shows the value of $S/\sqrt{S+B}$ for a given cut value, and the fraction of signal events remaining after the cut. The suggested cut $\chi^2 < 4.4$ is indicated by the vertical line.

Figure 4 shows an example of how the optimisation is performed. The top plot shows the binned distributions of background and signal for the vertex χ^2 of the $\mathrm{B}^{\pm}$ candidates. The histograms have been normalised to the estimated yield in a nominal year. The lower plot shows the signal significance normalised to 2 fb^{-1} (left scale), with the fraction of signal events remaining superimposed. The optimisation steps through values of the parameter and determines the value of $S/\sqrt{S+B}$ at each step. The maximum value is then calculated, which is indicated on the bottom plot by a vertical line. This particular variable has no significant correlations with the other variables used. The strategy used for dealing with correlations is described in Section 3.2.

3.1 Cut Variables

Mass Windows

Loose-mass windows of 500 MeV for the $\mathrm{B}^{\pm}$ and 50 MeV for the D reconstructed masses were applied in the preselection. Tight mass windows of 50 MeV for the reconstructed $\mathrm{B}^{\pm}$ and 22 MeV for the D were used to estimate the signal yield in the final offline selection.

Vertex χ^2

A cut on the maximum χ^2 of the reconstructed $\mathrm{B}^{\pm}$ and D vertices was used to reject events in which a B or D was reconstructed with tracks that did not come from a true decay vertex.

Impact Parameter

The impact parameter (IP) of a track with respect to a particular vertex is the distance of closest approach to the vertex. This will be the perpendicular distance between the track and the vertex. The impact parameter significance is defined as the IP divided by its error:

$$\mathrm{IPS} = \frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}. \tag{16}$$

The smallest Impact Parameter Significance (sIPS) is defined as the smallest IPS of all reconstructed PVs in a given event. Previously, the sIPS was used in selection studies. However, a new vertex fitter has been introduced in recent versions of the LHCb software, which calculates the IP χ^2 significance rather than the IPS. The two variables are highly correlated, and to leading order, one can assume the following relation between the two variables:

$$\text{IP } \chi^2 \text{ significance} \approxeq (\text{IPS})^2 . \tag{17}$$

For particles originating from the Primary Vertex (PV), the IP is expected to be very small, so a cut on the maximum IP χ^2 of the B meson with respect to the PV with the smallest IP χ^2, hereby referred to as the sIP χ^2, is implemented in the selection criteria. For particles from secondary vertices, in this case the D and bachelor kaon, a cut on the minimum sIP χ^2 is used in order to reject tracks used to reconstruct secondary vertices that originated from one of the PVs.

Flight Distance

The flight distance (FD) is the distance a particle travels before decaying. More explicitly, it is the magnitude of the vector difference between the position of the origin vertex and the decay vertex. The flight significance (FS) is defined as the FD divided by its error. The FS with respect to the PV is given by:

$$\text{FS} = \frac{|\mathbf{x_{decay}} - \mathbf{x_{PV}}|}{\sqrt{\sigma^2_{\text{decay}} + \sigma^2_{\text{PV}}}}, \tag{18}$$

where $\mathbf{x_{decay}}$ and $\mathbf{x_{PV}}$ are the 3-vector coordinates of the decay vertex and PV respectively, and σ_{decay} and σ_{PV} are their errors.

Similarly to the IP variable, the FD χ^2 significance is used in the current vertex fitter instead of the FS and is approximately equivalent to the square of the FS.

Pointing Angle

The pointing angle is defined as the angle between the flight direction of a reconstructed particle and its momentum. In the case of the $\text{B}^{\pm}$ meson, this is defined by:

$$\cos\theta = \frac{\mathbf{p}_{\text{B}^{\pm}} \cdot \left(\mathbf{x}_{\text{B}^{\pm}} - \mathbf{x}_{\text{PV}}\right)}{\left|\text{p}_{\text{B}^{\pm}}\right| \left|\mathbf{x}_{\text{B}^{\pm}} - \mathbf{x}_{\text{PV}}\right|}, \tag{19}$$

where θ is the pointing angle, $\mathbf{p}_{\text{B}^{\pm}}$ represents the $\text{B}^{\pm}$ 3-momentum, and $\mathbf{x}_{\text{B}^{\pm}}$ and $\mathbf{x}_{\text{PV}}$ are the 3-vector coordinates of the $\text{B}^{\pm}$ decay vertex and PV respectively.

For a perfectly reconstructed vertex, the momentum and flight directions will be identical, corresponding to $\cos\theta = 1$. Thus, a cut on the minimum value of $\cos\theta$ can be a powerful cut for discriminating between signal and background.

Particle Identification (PID)

LHCb has two Ring Imaging Cherenkov (RICH) detectors, which are situated either side of the dipole magnet. These RICH detectors are designed to discriminate between kaons and pions in a momentum range 2 GeV/c<p<100 GeV/c [21]. The PID algorithm uses the combined discrimination from the RICH detectors, electron and hadron calorimeters, and muon detectors, combining the results to obtain a likelihood of the particle satisfying a particular mass hypothesis. This information is most useful when we consider two different mass hypotheses; the difference between these two likelihoods is used by LHCb as a method of separating between these two mass hypothesis. For the final state particles in this study, the log likelihood difference between the kaon and pion mass hypotheses, hereby referred to as $\Delta \ln L\,(\text{K} - \pi)$ is the most powerful of these PID cuts. A particle with $\Delta \ln L\,(\text{K} - \pi) > 0$ is more in agreement with the kaon mass hypothesis than the pion hypothesis. The log likelihood difference between reconstructing the particle as a kaon and a proton, $\Delta \ln L\,(\text{K} - \text{p})$, also provides some discriminating power.

Maximum Bachelor Momentum

At high momentum, the difference between the Cherenkov angles of kaons and pions becomes is reduced, making discrimination difficult. LHCb predicts significant reduction in the ability of the RICH to separate kaons and pions at around 100 GeV/c. Thus, a maximum momentum cut of 100 GeV/c was applied to the bachelor kaon in order to reduce the number of $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$ reflection events, where the bachelor pion is misidentified as a kaon.

3.2 Correlations

The correlation between the cut variables was determined using the Kendall tau rank correlation statistic, τ, which runs between -1 (discordant) and +1 (concordant). This statistic was chosen because it makes no assumption about the distribution of the data (it is a non-parametric statistic), and it has a more straightforward interpretation than the Spearman rho rank statistic [a]. A correlation of $|\tau| > 0.5$ was assumed to be a strong correlation that merits investigation. We find that the correlations can be grouped; several topological variables are found to be correlated, as are the momentum measurements. Some variables, such as the vertex quality and the invariant mass distributions show no significant correlations.

Variables		**Kendall tau correlation**	
		Signal	*Background*
$\mathrm{B}^{\pm}$ FD χ^2	bach. sIP χ^2	0.76	0.51
D FD χ^2 from PV	D sIP χ^2	0.67	0.40
$\mathrm{B}^{\pm}$ FD χ^2	D sIP χ^2	0.58	0.07
$\mathrm{B}^{\pm}$ FD χ^2	D FD χ^2 from PV	0.58	0.14
$\mathrm{B}^{\pm}$ FD χ^2	$\mathrm{B}^{\pm} \cos\theta$	0.52	0.18

Table 2 Kendall tau correlation coefficients for $\mathrm{B}^{\pm} \to \mathrm{D}\,(\mathrm{K}\pi\pi\pi)\,\mathrm{K}^{\pm}$ topological variables. Only the pairings showing correlation of $|\tau| > 0.5$ are shown.

From Table 2, we can see that the strongest correlation in both the signal and background samples is between the $\mathrm{B}^{\pm}$ FD χ^2 and the sIP χ^2 of the bachelor $\mathrm{K}^{\pm}$; these variables are plotted in Figure 5. A clear diagonal structure in both the signal and background distributions is evident. To take account of this correlation, the dataset is rotated by an angle θ in the plane of these two variables, such that the two variables may be projected onto a transformed x-axis, $x' = x\cos\theta - y\sin\theta$. The optimisation then maximises $S/\sqrt{S+B}$ in the single dimension of x', as described in the previous section. In the case of the $\mathrm{B}^{\pm}$ FD χ^2 and the sIP χ^2 of the bachelor $\mathrm{K}^{\pm}$, an optimised value of $S/\sqrt{S+B}$ is found using two cuts, as indicated in Figure 5.

[a] The Kendall tau statistic represents the difference between the probability that the samples are in the same order compared to the probability they are differently ordered. For a comparison of the Spearman rho and Kendall tau rank statistics, see for example `http://rsscse.org.uk/ts/bts/noether/text.html`.

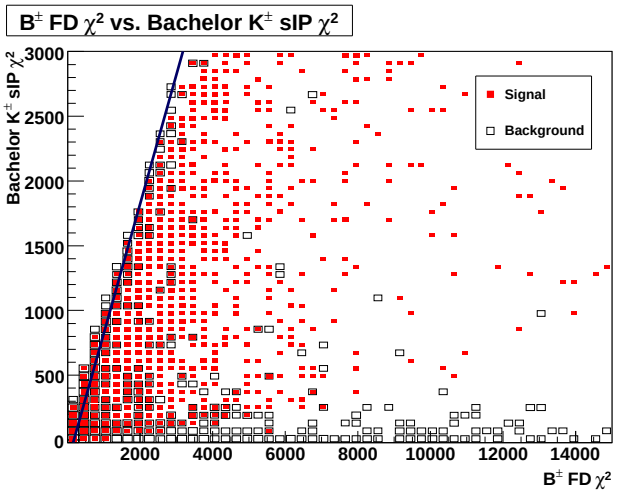

Figure 5 Plot of $\mathrm{B}^{\pm}$ FD χ^2 and bachelor $\mathrm{K}^{\pm}$ sIP χ^2 for the decay $\mathrm{B}^{\pm} \to \mathrm{D}\,(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})\,\mathrm{K}^{\pm}$, with tight-mass window cuts applied. The size of the boxes represent the number of signal and background events in each bin, normalised to 2 fb^{-1}. The area of the boxes is shown on a logarithmic scale to emphasise the features. The lines represent the final cut values found in the optimisation.

Variables		Kendall tau correlation	
		Signal	*Background*
D p_T	$\mathrm{K}^{\pm}$ from D p_T	0.57	0.41
D p_T	bach. p_T	-0.09	-0.49

Table 3 Kendall Tau Correlation Coefficients for $\mathrm{B}^{\pm} \to \mathrm{D}\,(\mathrm{K}\pi\pi\pi)$ Momentum Variables. Only the pairings showing correlation of $|\tau| \gtrapprox 0.5$ are shown.

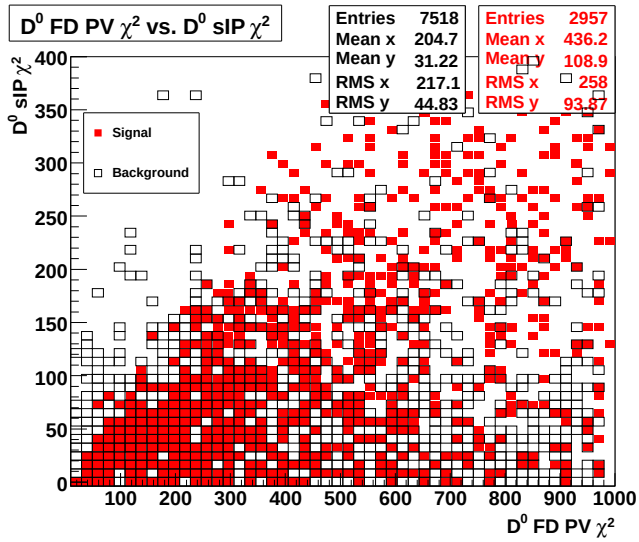

Figure 6 Plot of D FD χ^2 with respect to the PV and D sIP χ^2 for the decay $\mathrm{B}^{\pm} \to \mathrm{D}\,(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})\,\mathrm{K}^{\pm}$, with tight-mass window cuts applied.

The 2D plot of D FD χ^2 from the PV and D sIP χ^2 is shown in Figure 6. Once again, a clear structure can be seen in the signal distribution, though not as pronounced as the previous case, however, the optimisation suggests no great advantage in rotating the variables and applying a linear cut as opposed to applying a pair of one-dimensional cuts. The latter approach is therefore taken for simplicity. The remaining correlations in Table 2 are not investigated as they involve a variable that is already used in the aforementioned linear cut, and we chose not to go to higher dimensions because the Monte Carlo

statistics do not support such an optimisation. However, we note that correlations remain because the linear cut of $\mathrm{B}^{\pm}$ FD χ^2 and the sIP χ^2 of the bachelor $\mathrm{K}^{\pm}$ is observed to have less effect after applying other topological variables. After the other topological cuts are applied, the signal efficiency and background retention of the ($\mathrm{B}^{\pm}$ FD χ^2, bachelor sIP χ^2) cut are 93% and 25%. Prior to the application of the other topological cuts, the efficiency and retention are 80% and 12%.

The Kendall tau correlation coefficient for the three p_T variables are shown in Table 3. By inspection, we find it preferential to build a 2D linear cut between the D p_T and the p_T of the bachelor kaon. This cut is shown in Figure 7.

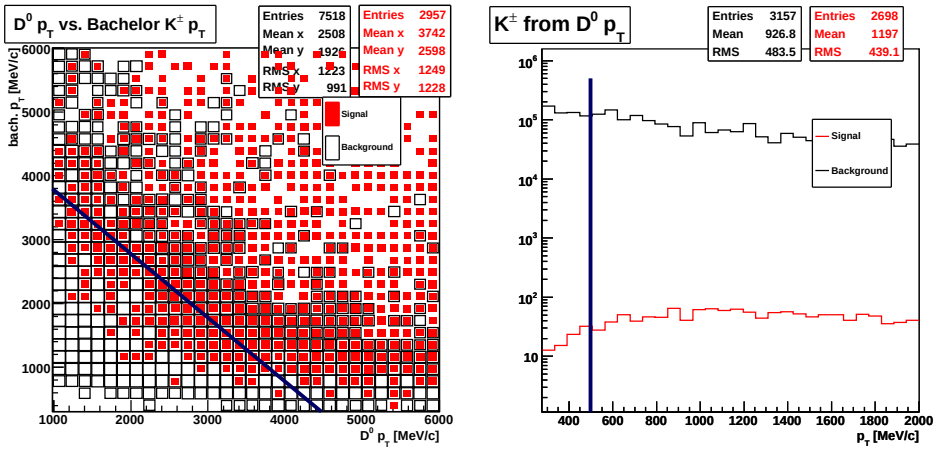

Figure 7 The left plot shows the 2-dimensional distribution of D p_T and bachelor $\mathrm{K}^{\pm}$ p_T, showing the advantage of using a linear cut for these variables. On the right is the 1D analysis plot of the p_T of the $\mathrm{K}^{\pm}$ from the D decay after the linear cut has been applied.

4 Background Estimation

Three background sources are identified: combinatoric background, which is formed from random combinations of tracks in the LHCb detector; the uncertainty on this source dominates the background estimation. We also consider partially reconstructed backgrounds separately, and backgrounds which suffer from a particle misidentification, hereafter referred to as *reflections*.

4.1 Combinatoric Background

The inclusive $\mathrm{b\overline{b}}$ sample is split into two: 11.3 million events are used to optimise the selection, whilst the remaining 10.7 million are kept in reserve to evaluate the cuts in an unbiased fashion. The result of the selection optimisation for both the $\mathrm{B}^{\pm} \to (\mathrm{K}\pi\pi\pi)_{\mathrm{D}}\,\mathrm{K}^{\pm}$ signal sample and the unbiased second half of the background sample is shown in Table 4.

	Signal		Bkg	
	Events	Efficiency	Events	Efficiency
Number of simulated LHCb events	50676		1.07×10^7	
LHCb offline reconstruction[a]	3761		6646	
Momentum Cuts				
bachelor $\mathrm{K}^{\pm}$ $p_T > -1 \times \mathrm{D}^0$ p_T+4780 MeV/c	3328	88.5%	3006	45.2%
$\mathrm{K}^{\pm}$ from D^0 p_T >500 MeV/c	3182	95.6%	2619	87.1%
$\pi^{\pm}$ from D^0 p_T >160 MeV/c	3008	94.5%	2218	84.7%
Topological Cuts				
bachelor $\mathrm{K}^{\pm}$ sIP $\chi^2 <$ $\mathrm{B}^{\pm}$ FD $\chi^2 - 178$	2303	76.6%	213	9.6%
bachelor $\mathrm{K}^{\pm}$ sIP $\chi^2 > 35$	2287	99.3%	152	71.4%
$\mathrm{B}^{\pm}$ $\cos\theta > 0.9999725$	2180	95.3%	74	48.7%
$\mathrm{B}^{\pm}$ sIP $\chi^2 < 11$	2061	94.5%	45	60.8%
D FD from PV $\chi^2 > 144$	2045	99.2%	43	95.6%
D sIP $\chi^2 > 45$	1998	97.7%	33	76.7%
Vertex Quality Cuts				
$\mathrm{B}^{\pm}$ vertex $\chi^2 < 4.4$	1897	94.9%	24	72.7%
D vertex $\chi^2 < 20$	1816	95.7%	16	66.7%
PID Requirements				
bachelor $\mathrm{K}^{\pm}$ $\Delta \ln L\,(\mathrm{K}-\pi) > -1.0$	1770	97.5%	10	62.5%
$\mathrm{K}^{\pm}$ from D $\Delta \ln L\,(\mathrm{K}-\pi) > -1.0$	1748	98.8%	8	80.0%
$\pi^{\pm}$ from D $\Delta \ln L\,(\mathrm{K}-\pi) < 5.0$	1641	93.9%	5	62.5%
bachelor $\mathrm{K}^{\pm}$ $\Delta \ln L\,(\mathrm{K}-\mathrm{p}) > -5.0$	1622	98.8%	5	100%
$\mathrm{K}^{\pm}$ from D $\Delta \ln L\,(\mathrm{K}-\mathrm{p}) > -5.0$	1588	87.9%	4	80%
bachelor $\mathrm{K}^{\pm}$ $p < 100\mathrm{GeV/c}$	1486	93.6%	4	100%
Mass Windows				
D mass window = $\pm$22 MeV/c^2	1414	95.2%	2	50%
$\mathrm{B}^{\pm}$ mass window = $\pm$50 MeV/c^2 [b]	1359	96.1%		
L0 Accepted[c]	583	42.9%		

Table 4 Signal and background efficiencies of optimised cuts for all signal events and unbiased sample of combinatoric background events for decay channel $\mathrm{B}^{\pm} \to \mathrm{D}\,(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})\,\mathrm{K}^{\pm}$.

[a] Including removal of non-combinatoric sources.
[b] $\mathrm{B}^{\pm}$ mass window of $50\mathrm{MeV/c^2}$ applied to signal sample. The background sample has a loose $500\mathrm{MeV/c^2}$ window applied.
[c] The background that remains after the selection is assumed to scale proportionally to the offline selected signal.

We see that the final selection retains 1,359 signal events from 50,676 simulated events, an offline reconstruction efficiency of 2.7%. This selection reduces 10.7 million combinatoric background events to just 2 events in the wide $\mathrm{B}^{\pm}$ mass window. The Level-0 hardware trigger emulator is also applied, and the signal is seen to reduce by another factor of 42.9%. It is assumed that the background events passing the offline selection would be accepted by the L0 trigger with the same efficiency ($\epsilon_{l0/sel}$) as the signal.

As described in Section 3, the combinatoric background statistics are effectively doubled by allowing wrong sign (opposite sign kaons) and right sign (same sign kaons) combinations to occur in the background sample. This leads to the normalisation factor calculated in section 2.3 to reduce by a factor 2. The background yield is therefore given by:

$$\begin{aligned} B_{comb} &= N_{sel} \times \frac{\mathrm{Norm.}}{2} \times \frac{M_{tight}}{M_{loose}} \times \epsilon_{L0/sel} \\ B_{comb} &= 2^{+2.25}_{-1.26} \times \frac{40800}{2} \times 0.1 \times 0.429 \\ B_{comb} &= 1750^{+1970}_{-1100} \end{aligned} \tag{20}$$

where M_{tight} and M_{loose} are the 50 MeV and 500 MeV windows respectively. The errors on the number of selected particles are calculated using the Feldman and Cousins approach [22], such that the 1σ confidence limit is given by $[547,\ 3719]$.

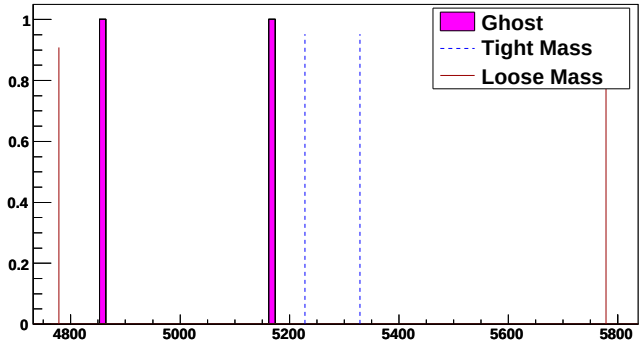

Figure 8 Combined WS and RS combinatoric candidates in the second half of the inclusive $\mathrm{b\overline{b}}$ background.

The same selection is used to select $\mathrm{B}^{\pm} \to (\mathrm{KK}\pi\pi)_{\mathrm{D}}\, \mathrm{K}^{\pm}$, except the D candidate is formed from two oppositely charged kaons and two oppositely charges pions. The PID requirements for these D daughters are the same as those listed in the table. The selection retains 1,084 events prior to the L0 and 481 after ($\epsilon_{L0/sel} = 42.5\%$). Three $\mathrm{B}^{\pm}$ candidates survived the selection cuts from the full sample of 2.2×10^7 $\mathrm{b\overline{b}}$ inclusive events. Two were classified as low mass backgrounds, and are expected to lie far outside the $\mathrm{B}^{\pm}$ mass signal window; the final remaining candidate is a 'ghost'. The background level in this singly Cabibbo suppressed mode is therefore estimated at 850^{+1480}_{-530}.

In both modes, the remaining backgrounds after the offline selection are classified as ghost events. No track quality cut were available to this analysis, but it is hoped that future iterations of this study may further reduce the level of this background using cuts on the χ^2 of the track reconstruction or other track quality variables.

4.2 $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$ Reflections

Background candidates in which one or more of the final-state particles in the decay have been misidentified are known as *reflections* by the background classification tool used by LHCb. The most prevalent source of reflection is $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$, where the D decays to the required final state. This background has a branching fraction approximately a factor 12 higher than the signal mode, and is expected to have a tail that extends into the tight $\mathrm{B}^{\pm}$ mass window. To limit the number of reflections, PID requirements of $DLL\,(\mathrm{K}-\pi) > -1$ were applied to the bachelor kaon as well as the $\mathrm{K}^{\pm}$ from the D decay.

In the $\mathrm{B}^{\pm} \to (\mathrm{K}\pi\pi\pi)_{\mathrm{D}}\, \mathrm{K}^{\pm}$ selection, only one reflection candidate passes the selection cuts, and lies outside the tight-mass window. This was identified to be a right sign candidate. Given that no right sign reflections enter the tight-mass window, we can place a limit on the number of WS events in the tight-mass window in $2\mathrm{fb}^{-1}$:

$$\begin{aligned} B_{refl} &= \left[0,\ N_{90}(0) \times \mathrm{Norm.} \times \epsilon_{L0/sel} \times \frac{BR(WS)}{BR(RS)}\right] \\ B_{refl} &= [0,\ 2.44 \times 20000 \times 0.429 \times 3.4 \times 10^{-3}] \\ B_{refl} &= [0,\ 67]\ (90\%\ \mathrm{CL}) \end{aligned} \tag{21}$$

where $BR(WS)/BR(RS) = (3.4 \pm 0.6) \times 10^{-3}$ is measured by Belle [23] for $\mathrm{B}^{\pm} \to (\mathrm{K}\pi)\,\pi^{\pm}$. As before, we have used the Feldman and Cousins approach when estimating the statistical background errors. Similarly, for the singly Cabibbo suppressed mode $\mathrm{B}^{\pm} \to (\mathrm{KK}\pi\pi)_{\mathrm{D}}\, \mathrm{K}^{\pm}$, no events were found in the tight-mass window and the congruent calculation gives an limit of [0, 20,600] (Feldman-Cousins 90% CL).

The signal extraction fit is expected to involve a fit to the spectrum of the invariant mass of the $\mathrm{B}^{\pm}$ meson taking into account the shape of the signal peak and the nearby $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$ reflection peak. As such, this background should be taken into account in the fitted PDF, and therefore is not included in the final irreducible background estimation in this document.

4.3 Biased Sample

In order to obtain a better estimate for the non-combinatoric background, a 'biased' sample of $\mathrm{b\overline{b}}$ background is analysed. This sample is focused on a particular region of phase space by placing some additional requirements on the proper time, psuedorapidity and p_T of the B meson [b]. To estimate the number of such events in 2 fb^{-1}, the number of favoured sign $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ signal events in the signal sample and biased sample are considered. Applying the same cuts used to generate the biased sample on the signal sample, 861 events are selected; 40 such events were found in the biased sample. We deduce the 2 fb^{-1} normalisation by identifying:

$$\mathrm{Norm.}_{\mathrm{biased}} \times n_{\mathrm{biased}} = \mathrm{Norm.}_{\mathrm{signal}} \times n_{\mathrm{signal}}$$

where the same selection criteria are applied to both samples.

Given that the normalisation for the signal sample is 91, the normalisation for the biased sample is estimated to be $91 \times (861/40) = 1,960$, approximately 1/10 of the total inclusive $\mathrm{b\overline{b}}$ sample. The different background contributions are considered below.

$B^{\pm} \to D\pi^{\pm}$ Reflections

As was the case in the inclusive $\mathrm{b\overline{b}}$ sample, no WS reflection events in the biased sample pass the selection, although several reflection events are selected in the RS mode. With no cuts applied, the number of RS $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$ reflection events in the biased sample is 125, of which 71 lie in the tight-mass window. This means that 56.8% of the reflection events are within the tight-mass window, corresponding to a 2 fb^{-1} normalisation of $1,960 \times 0.568 = 1,113$. After the selection cuts, 8 reflection events pass the L0 trigger in the tight-mass window. Thus, we estimate the RS reflection contribution to be 8900^{+3700}_{-3000}. Using the Belle estimate of the ratio of the WS and RS $\mathrm{B}^{\pm} \to (\mathrm{K}\pi)\,\pi^{\pm}$ branching ratios of $(3.4 \pm 0.6) \times 10^{-3}$, the WS contribution is estimated to be 30^{+12}_{-11}, which agrees with the estimated 90% Feldman-Cousin CL of $[0,\ 67]$ from the inclusive $\mathrm{b\overline{b}}$ sample. For the mode $\mathrm{B}^{\pm} \to (\mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}\pi^{-})_{\mathrm{D}}\,\mathrm{K}^{\pm}$, 5 reflection events are found before selection cuts, of which 2 are within the tight-mass window. Since there are so few reflection events in the tight-mass, we make the assumption that the distribution of reflection events will be similar to the RS mode. A single reflection event is triggered and selected. This event lies outside the tight-mass window, so the estimated background contribution from reflections is $[0,\ 1400]$ at the 68.27% CL, and $[0,\ 2700]$ at the 90% CL.

Partially Reconstructed Physics Background and Low Mass Background

As before, there are no low mass or partially reconstructed physics background events in the WS sample. The low mass background is expected to peak well below the tight-mass window. However, there is a small possibility that a few low mass events from the upper tail can extend into the tight-mass window. This is the case in the biased sample for $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$. One low mass event is found in the tight-mass window in the biased sample, both before and after the selection cuts. This low mass background is combined with the partially reconstructed physics background to obtain an overall estimate for these background types.

Before the selection cuts are applied, 431 partially reconstructed physics background events are found in the loose-mass window, of which 14 lie in the loose-mass window. Combining this with the 59 low mass events in the loose-mass window, the proportion of background in the tight-mass window is 3.1%, corresponding to a normalisation of 60. One low mass and one partially reconstructed event are selected and triggered within the tight-mass window, which correspond to a nominal background contribution of 120^{+135}_{-76}.

For the $\mathrm{B}^{\pm} \to (\mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}\pi^{-})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ mode, 31 partially reconstructed physics background are found, with 3 events within the tight-mass window before selection cuts. One L0 triggered event remains

[b] The events retained by the biased sample have at least one B hadron with $2.2 < \eta < 4.7$, p_T >8.4 GeV/c, $c\tau$ >0.16 mm and $p_T + 5.36 \times \eta > 26$ GeV/c.

after applying the selection. We therefore estimate the background contribution from partially reconstructed physics background to be 190^{+330}_{-120}.

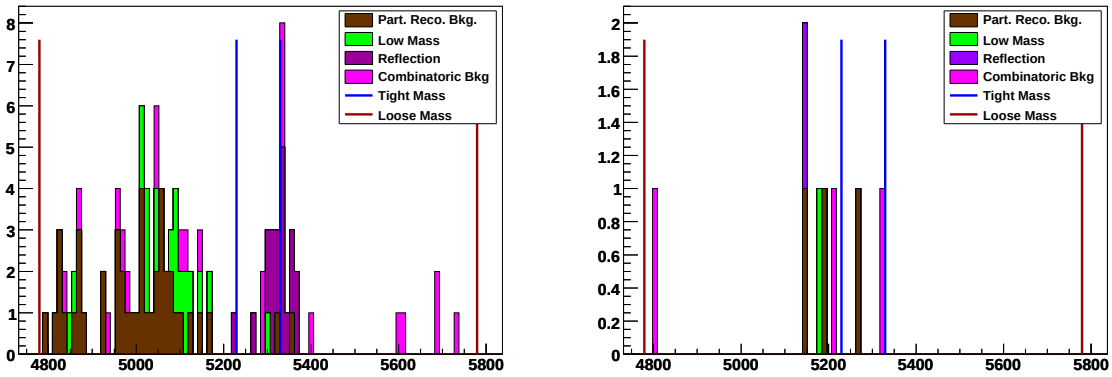

Table 5 Plots showing the background triggered and selected in the biased background sample for $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ (left plot) and $\mathrm{B}^{\pm} \to (\mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}\pi^{-})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ (right plot). No WS events were found in the biased sample.

5 Signal Yield and B/S Estimate

In section 4.1, the optimised event selection for $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ was presented. Including the emulation of the hardware trigger, 583 events remain from 50,676 fully simulated events. Using the normalisation factor calculated in Section 2.3, the annual 2 fb^{-1} yield of B^{+} and B^{-} combined is expected to be 554. Likewise, 1,350 $\mathrm{B}^{\pm} \to (\mathrm{K}\mathrm{K}\pi\pi)_{\mathrm{D}}\,\mathrm{K}^{\pm}$ and 53,000 $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$ events are anticipated. The concluding calculation is shown in Table 6.

2 fb^{-1} yields	$\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm}$	$\mathrm{K}^{\pm}\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}$	$\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}$
Combinatoric Background (inclusive $\mathrm{b\bar{b}}$)	1750^{+1970}_{-1100}	850^{+1480}_{-530}	1750^{+1970}_{-1100}
Part. reco. and low mass backgrounds (biased)	—	190^{+330}_{-120}	120^{+135}_{-76}
$\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$ reflections (biased)	30^{+12}_{-11}	$[0,\ 1400]$	8900^{+3700}_{-3000}
Total background	1780^{+1970}_{-1100}	< 4240 [a]	10700 ± 3200
Background estimate, excl. reflections [b]	1750^{+1970}_{-1100}	1000 ± 510	1830^{+1000}_{-1100}
Signal Yield	554^{+272}_{-179}	1350 ± 116	53000 ± 3800
B/S, incl. reflections	3 ± 2	< 3 [c]	$0.20^{+0.06}_{-0.06}$
B/S, excl. reflections	3 ± 2	0.7 ± 0.4	0.03 ± 0.02

Table 6 B/S ratios for $\mathrm{B}^{\pm} \to \mathrm{D}\,(\mathrm{K}\pi\pi\pi)\,\mathrm{K}^{\pm}$ and $\mathrm{K}^{\pm} \to \mathrm{D}\,(\mathrm{K}\mathrm{K}\pi\pi)\,\mathrm{K}^{\pm}$. The error on the suppressed signal yield is dominated by the error on the estimate of the branching fraction.

[a] The value quoted is the sum of the upper limits of the Feldman-Cousins 68.27% CLs of the contributing backgrounds.
[b] The estimated background yield, excluding $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$. We assume here that these events will be fitted with the signal.
[c] The value quoted is the ratio of the 68.27% Feldman-Cousins upper limit of the background and the $1\,\sigma$ lower limit of the signal yield, and is therefore a very conservative estimate.

These modes have been examined with an older version of the LHC simulation [4]. Table 7 shows the B/S estimates for the previous study, using the same estimates for the branching fractions used in this study, and the trigger efficiencies used above applied to the signal and background estimates. This way, a fair comparison can be made between the two studies. The results show that the signal yields in all modes have fallen. This is perhaps not unexpected, given the greater detail in the later simulation. The B/S estimations for the wrong sign mode are the same in both studies, and the results for $\mathrm{B}^{\pm} \to (\mathrm{K}^{+}\mathrm{K}^{-}\pi^{+}\pi^{-})\,\mathrm{K}^{\pm}$ agree within errors. The right sign result has improved, mainly due a much lower estimate of the partially reconstructed physics background from scaling the background found

in the tight mass window to the total partially reconstructed physics background. These results are encouraging, and demonstrate that both studies have broadly come to the same positive conclusion about the feasibility of this analysis.

$2\ \mathrm{fb}^{-1}$ yields	$\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm}$	$\mathrm{K}^{\pm}\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}$	$\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}$
Combinatoric Background	2300^{+1290}_{-1030}	918^{+1030}_{-578}	2300^{+1290}_{-1030}
Part. Reco. Background	—	287^{+323}_{-181}	4590^{+8030}_{-2890}
Background estimate, excl. reflections	2300^{+1290}_{-1030}	1210^{+1070}_{-1080}	6820^{+8190}_{-8050}
Signal Yield	815^{+311}_{-264}	2360 ± 169	78000 ± 4510
B/S, excl. reflections	3 ± 2	0.5 ± 0.5	0.09 ± 0.11

Table 7 B/S estimates for previous study [4]. The signal yields have been calculated using the same branching ratios as the current study. Reflections have not been included in the estimate of the background. Since the triggers used in the previous study were different to those used here, the trigger efficiencies of this study have been applied to the signal and background estimates.

6 References

[1] N. Cabibbo *Phys. Rev. Lett.* **10** (1963) 531–533.

[2] M. Kobayashi and T. Maskawa *Prog. Theor. Phys.* **49** (1973) 652–657.

[3] **CKMfitter Group** Collaboration, J. Charles *et al. Eur. Phys. J.* **C41** (2005) 1–131, arXiv:hep-ph/0406184.

[4] **LHCb** Collaboration, A. S. Powell. CERN-LHCb-2007-004.

[5] M. Gronau and D. London *Phys. Lett.* **B253** (1991) 483–488.

[6] M. Gronau and D. Wyler *Phys. Lett.* **B265** (1991) 172–176.

[7] D. Atwood, I. Dunietz, and A. Soni *Phys. Rev. Lett.* **78** (1997) 3257–3260, arXiv:hep-ph/9612433.

[8] **Heavy Flavor Averaging Group** Collaboration, E. Barberio *et al.* arXiv:0808.1297 [hep-ex].

[9] D. Atwood and A. Soni *Phys. Rev.* **D68** (2003) 033003, arXiv:hep-ph/0304085.

[10] **CLEO** Collaboration, N. Lowery arXiv:0903.4853 [hep-ex].

[11] **LHCb** Collaboration, K. Akiba *et al.* CERN-LHCb-2008-031.

[12] **Particle Data Group** Collaboration, C. Amsler *et al. Phys. Lett.* **B667** (2008) 1.

[13] A. Giri, Y. Grossman, A. Soffer, and J. Zupan *Phys. Rev.* **D68** (2003) 054018, arXiv:hep-ph/0303187.

[14] **CLEO** Collaboration, R. A. Briere *et al.* arXiv:0903.1681 [hep-ex].

[15] **LHCb** Collaboration, J. Libby. CERN-LHCB-2007-141.

[16] **LHCb** Collaboration, V. Gibson, C. Lazzeroni, and J. Libby. CERN-LHCB-2007-048.

[17] J. Rademacker and G. Wilkinson *Phys. Lett.* **B647** (2007) 400–404, arXiv:hep-ph/0611272.

[18] **GEANT4** Collaboration, S. Agostinelli *et al. Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[19] J. Allison *et al. IEEE Trans. Nucl. Sci.* **53** (2006) 270.

[20] R. Brun and F. Rademakers *Nucl. Instrum. Meth.* **A389** (1997) 81–86.

[21] **LHCb** Collaboration, "LHCb: RICH technical design report,". CERN-LHCC-2000-037.

[22] G. J. Feldman and R. D. Cousins *Phys. Rev. D* **57** no. 7, 3873.

[23] **Belle** Collaboration, Y. Horii *et al. Phys. Rev.* **D78** (2008) 071901, `arXiv:0804.2063 [hep-ex]`.

[24] T. Sjostrand *et al. Comput. Phys. Commun.* **135** (2001) 238–259, `arXiv:hep-ph/0010017`.

[25] **LHCb** Collaboration, V. Gligorov. CERN-LHCb-2007-044.

A Glossary

The following terminology is used throughout this paper.

- **Bachelor kaon** - The kaon produced from the $\mathrm{B}^{\pm}$ decay.
- **Wrong Sign (WS)** - The suppressed decay $\mathrm{B}^{\pm} \to (\mathrm{K}^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_{\mathrm{D}}\,\mathrm{K}^{\pm}$, where the two kaons are of opposite charge.
- **Right Sign (RS) -** In addition to the WS decay, we also refer to the kinematically identical Right Sign (RS) decay $\mathrm{B}^{\pm} \to (\mathrm{K}^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_{\mathrm{D}}\,\mathrm{K}^{\pm}$, where the kaons are the same sign. This mode is unsuppressed but is significantly less sensitive to γ.
- **Minimum bias** - This refers to a background sample that retains all proton-proton interaction events as generated by Pythia [24].
- **Inclusive $\mathrm{b\bar{b}}$** - A subset of a minimum bias sample that is required to contain a $\mathrm{b\bar{b}}$ pair.
- **Generator cut** - A cut applied at the generator stage. For the inclusive $\mathrm{b\bar{b}}$ background sample used, the generator cut requires that the B meson lies within the 400 mrad detector acceptance region. The signal samples used also have the requirement that at least of the the $\mathrm{B}^{\pm}$ mesons decays to the required signal mode.
- **Preselection** - A loose set of cuts that is used to reduce the combinatoric rate in the background datasets, reducing the amount of data that needs to be stored for analysis.
- **Background Category** - The background category tool used by LHCb [25] classifies an event type as signal, quasi-signal (signal decay that occurs via a resonant decay), or background, according to a signal decay descriptor. Background events are further classified by the type of background with respect to the signal decay. The background classifications mentioned in this note are described below.
- **Combinatoric Background** - Combinatoric background are events that are reconstructed from tracks that do not have a single origin. As such, these events can be found anywhere in the B mass window. We do not know what the true distribution of the combinatoric background will be, but we make the assumption that we can scale the number of events with reconstructed B mass found in the loose symmetric mass window about the true B rest mass to the tight-mass window by the relative size of the tight-mass window.
- **Ghost Background** - A ghost track has no associated Monte Carlo particle. The exact definition depends on the type of track. Generally, the track is classed as a ghost if it has <70% of hits associated to a single MC particle in the tracking stations it passed through, as well as the Vertex Locator (VELO) if it originated from the primary vertex. If the track also passes through the TT (the first tracking station, located between the inner RICH detector and the dipole magnet), it must also have no more than one incorrectly assigned TT hit. It is also possible for a track to be assigned as a ghost if the Monte Carlo associator fails. An event with one or more ghost tracks is classified as a ghost event by the background classification tool. This is the most common combinatoric background in the samples used in this study.
- **Partially Reconstructed Physics Background** - An event is classified as partially reconstructed background according to the signal decay descriptor if one or more particle in the event has not been reconstructed, and one or more final state particle has been misidentified. This type of background tends to peak below the tight-mass window, but may have an upper tail which enters the tight-mass window.

- **Low Mass Background** - This type of background is similar to the partially reconstructed background, but all the final state particles are correctly identified, and the true mass of the particle reconstructed as the B meson does not exceed the true $\mathrm{B}^{\pm}$ mass by more than 100 MeV/c^2.
- **Reflection Background** - Events classified as reflections have one or more misidentified final state particles (but no missing particles). The main type of reflection event for this study are decays of the form $\mathrm{B}^{\pm} \to \mathrm{D}\pi^{\pm}$, where the bachelor kaon has been misidentified as a pion, with the correct decay of the D meson. This is a potentially troublesome source of background, as it peaks just above the tight-mass window. It is expected that this background will inevitably be fit along with the signal.