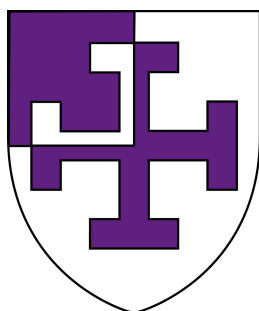


Measurements of the CKM angle γ at the LHCb experiment

Shu Faye Cheung
St Cross College, University of Oxford



Thesis submitted in partial fulfilment of the requirements for the degree
of Doctor of Philosophy at the University of Oxford

Trinity Term 2016



Measurements of the CKM angle γ at the LHCb experiment

Shu Faye Cheung

St Cross College, University of Oxford

Thesis submitted in partial fulfilment of the requirements for the degree
of Doctor of Philosophy at the University of Oxford

Trinity Term 2016

Abstract

Two measurements of the Cabibbo-Kobayashi-Maskawa angle γ using $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ decays are presented in this thesis. The subsequent D meson decays to the $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ final states are studied using a binned Dalitz plot analysis. The D strong-phase variation over the Dalitz plot is taken from measurements performed at the CLEO-c experiment, making the analysis independent of a model to describe the D decay amplitude. Both measurements are performed using proton-proton collision data collected by the Large Hadron Collider beauty (LHCb) experiment in 2011 and 2012, corresponding to an integrated luminosity of 3 fb^{-1} at centre-of-mass energies $\sqrt{s} = 7 \text{ TeV}$ and 8 TeV . The value $\gamma = (62_{-14}^{+15})^\circ$ is measured using $B \rightarrow DK$ decays and $\gamma = (71 \pm 20)^\circ$ is measured using $B^0 \rightarrow DK^{*0}$ decays, with a second solution for each value corresponding to $\gamma + 180^\circ$. The measurements are consistent with the world average value of γ .

*“There is an art, or rather, a knack to flying.
The knack lies in learning how to throw yourself at the ground and miss.
Pick a nice day and try it.”*

— Douglas Adams, *Life, the Universe and Everything*

Acknowledgements

I have many people to thank for an enjoyable and fruitful DPhil.

First and foremost, my utmost thanks go to my supervisors, Sneha Malde and Neville Harnew, for not only carefully reading this thesis but for their guidance and wisdom throughout the last four years. Sneha is responsible for never letting me get away with sub-standard work (this thesis included!) and I could not have asked for a better mentor and colleague. I have learnt an enormous amount from her and she has truly shaped the way I approach my work. Her passion and determination has also been a source of inspiration for whatever career path I choose next. I have Neville to thank for his constant encouragement, numerous teaching opportunities in Oxford and abroad, and for helping me to secure funding for conferences and my long term attachment at CERN.

I also thank many others in the Oxford LHCb group. Chris Thomas, with whom I worked closely on the $B \rightarrow DK$ analysis, utterly amazed me with the speed at which he worked, and I have been lucky to learn how to code by example from him. It goes without saying that Dan Johnson sowed the seeds for this thesis with his DPhil, and I have been grateful for his advice and friendship. I also owe many thanks to Donal Hill, Nazim Hussain, Scott Stevenson, Denis Derkach, Paolo Gandini, Malcolm John and Guy Wilkinson for their invaluable insights. Many thanks also go to Sue Geddes and Kim Proudfoot for getting me through the labyrinth of administrative paperwork.

I am lucky to have shared my time (coffees, lunchtimes, boardgames and DPhil-related griefs) in both Oxford and CERN with fellow DPhil students Will Kalderon, Mark Pickering, Will Fawcett, Alex Rollings, Barak Gruber and Neven Blaskovic, and also Katharina Behr and Jack Roberts, who were wonderful flatmates in St Genis. Special thanks go to Olli Lupton for teaching me that there is no such thing as too much coffee or too much cake. I also extend my thanks to Jackie Cheung, Philea Chim, Ella Fung, Dhivina Gagoomal, Jacob Kempster, Ben Kindler and Rebecca van der Hoeven for fun times shared in Oxford/London.

Thank you to the fantastic educators John Shearman, for setting me off on this path, and Owen Saxton, for keeping me on it.

My gratitude goes to the Croucher Foundation and Oxford University for funding this thesis.

Finally, I thank my mother, father and Thaddeus for their life-long encouragement, patience and love. I would not be where I am without them.

Contents

Preface	1
1 CP violation and measurements of the CKM angle γ	3
1.1 CP violation	4
1.2 The CKM matrix	5
1.3 The CKM unitarity triangle	8
1.4 The CKM angle γ	12
1.5 Tree-level determination of γ using $B \rightarrow DK$ decays	13
1.6 Measuring γ with the GGSZ method	16
1.6.1 Dalitz plots for three-body D decays	17
1.6.2 Measurement with $B \rightarrow DK$ decays	20
1.6.3 Determination of the strong-phase difference	23
1.6.4 Measurement with $B^0 \rightarrow DK^{*0}$ decays	30
1.7 Previous γ measurements using the GGSZ method	34
1.8 Analysis overview	35
2 The LHCb detector	38
2.1 The VELO	40
2.2 Tracking stations and magnet	42
2.3 The RICH detectors	46
2.4 Calorimeters	48
2.5 Muon system	50
2.6 Trigger system	51
2.6.1 Level-0 trigger	52
2.6.2 High Level Trigger	54
2.7 Reconstruction	55
2.7.1 Track reconstruction	55
2.7.2 Particle identification	57
2.7.3 Stripping	58
2.8 Simulation	58
3 Selection and parameterisation of $B \rightarrow DK$ decays	61
3.1 Selection of $B \rightarrow Dh, D \rightarrow K_s^0 h^+ h^-$ candidates	61
3.1.1 Reconstruction and trigger requirements	61
3.1.2 Boosted decision tree	63
3.1.3 Studies of peaking backgrounds	71

3.1.4	Particle identification requirements	74
3.1.5	Additional requirements	74
3.1.6	B mass distribution and Dalitz plots of selected candidates . . .	75
3.2	Mass fit	79
3.2.1	Signal	79
3.2.2	Combinatorial background	80
3.2.3	Misidentified decays	81
3.2.4	Partially reconstructed decays	85
3.2.5	Results	94
4	The flavour-tagged $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ control channel	99
4.1	Selection of semileptonic candidates	101
4.1.1	Reconstruction and trigger	101
4.1.2	Selection	102
4.1.3	Additional requirements	104
4.1.4	Dalitz plots of selected semileptonic candidates	104
4.2	Yield of $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ within the Dalitz plot	104
4.2.1	Fit to determine the shape parameters	105
4.2.2	Fit to measure the yield in each Dalitz plot bin	108
4.3	Dalitz plot efficiency correction	112
4.3.1	Efficiency differences in the semileptonic and hadronic modes .	112
4.3.2	Sources of efficiency difference	117
4.3.3	The efficiency correction	118
5	Measurement of the CP violation observables in $B^\pm \rightarrow DK^\pm$ decays	123
5.1	Fit configuration	123
5.1.1	Mass shapes and fit range	125
5.1.2	Yields of signal and background components	126
5.1.3	The likelihood function	127
5.1.4	The construction of r_B	129
5.1.5	Expected statistical sensitivity	130
5.1.6	Fitter bias	131
5.2	Results of the binned fit	133
5.2.1	Two-dimensional confidence level contours	140
5.2.2	Alternative fit to measure the yields	140
5.3	Systematic uncertainties	142
5.4	Summary of results	156
6	Measurement of the CP violation observables in $B^0 \rightarrow DK^{*0}$ decays	158
6.1	Selection of $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_s^0 h^+ h^-$ candidates	159
6.1.1	Boosted decision tree	159
6.1.2	Additional requirements	162
6.1.3	Dalitz plots of selected candidates	163
6.2	Parametrisation of the B mass distribution	163
6.2.1	Background studies and choice of fit range	165
6.2.2	Mass shapes	167

6.2.3	Constraints on background yields	170
6.2.4	Fit results	174
6.3	Calculation of the F_i parameters	177
6.3.1	Selection and parameterisation of $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates . . .	177
6.3.2	Dalitz plot efficiency correction	178
6.4	Fitting the CP violation observables	184
6.4.1	Fit configuration	185
6.4.2	Uncorrected fit results and fit bias	188
6.4.3	Fit results	189
6.4.4	Systematic uncertainties	190
6.5	Summary of results	202
7	Extraction of the CKM angle γ	204
7.1	Measurement of γ with $B \rightarrow DK$ decays	204
7.1.1	Extraction method and results	205
7.1.2	Comparison to previous model-independent measurements . . .	208
7.1.3	Comparison to model-dependent measurements	211
7.2	Measurement of γ with $B^0 \rightarrow DK^{*0}$ decays	212
7.2.1	Comparison to previous measurements	214
7.3	Combination of γ measurements	216
7.3.1	Combination of GGSZ MI measurements	217
7.3.2	Combination with other LHCb γ measurements	218
8	Conclusions	223
8.1	Summary of results	223
8.2	Outlook	224
	Bibliography	228

Preface

In the Standard Model of particle physics, the phenomenon of CP violation manifests itself through a single complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix. The CKM angle γ is a powerful parameterisation of the CP -violating phase, as it can be measured in tree-level processes without penguin pollution. Therefore, γ forms a vital benchmark for measuring Standard Model processes, which can be compared to values of γ obtained indirectly with loop-level measurements. A deviation between the direct and indirect values would be a sign of New Physics.

In this thesis, measurements of the CKM angle γ are performed using $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ decays, where D is a superposition of D^0 and $\bar{D}^0$ mesons decaying to either the $K_s^0 \pi^+ \pi^-$ or $K_s^0 K^+ K^-$ final state, collectively known as $K_s^0 h^+ h^-$. The hadronic parameters which describe the magnitude ratio and the strong-phase difference between the two interfering decay paths are also measured. The measurement involves an analysis of the $D \rightarrow K_s^0 h^+ h^-$ Dalitz plot, using a model-independent method to provide inputs for the D strong-phase variation, which have been measured at the CLEO experiment.

The data were collected at the Large Hadron Collider beauty (LHCb) experiment, which is a forward arm spectrometer situated at the Large Hadron Collider (LHC), built and operated by the European Organisation for Nuclear Research (CERN). The analysed data correspond to an integrated luminosity of 3 fb^{-1} of proton-proton collisions collected in 2011 and 2012 at centre-of-mass energies of $\sqrt{s} = 7 \text{ TeV}$ and 8 TeV .

The structure of this thesis is outlined, and places where this work has benefited from work performed by others are stated below. Chapter 1 provides an introduction to the angle γ and the method used to measure it. The general principles for the method were first proposed in Refs. [1] and [2]. Measurements of the D strong-phase variation performed at the CLEO experiment [3] are also described in this chapter. The LHC and the LHCb detector are described in Chapter 2. In Chapter 3 the selection and parameterisation of $B \rightarrow DK$ decays is described. Here, the Horns and Hill functions describing partially reconstructed backgrounds were first developed by Paolo Gandini. In Chapter 4 the $D^0 \rightarrow K_S^0 h^+ h^-$ Dalitz plot distribution of flavour-tagged D mesons using the semileptonic $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ control channel is determined. Chapter 5 outlines the measurement of the CP violation observables in $B \rightarrow DK$ decays. The work in Chapters 3–5 was undertaken in collaboration with Christopher Thomas. Chapter 6 summarises the measurement of the CP violation observables in $B^0 \rightarrow DK^{*0}$ decays. The multivariate selection of $B^0 \rightarrow DK^{*0}$ candidates was developed by Alexis Vallier. The extraction of the value of γ , together with the hadronic parameters, from the measured observables is described in Chapter 7. The code for performing the extraction was provided by Anton Poluektov, and the code for combining different γ measurements was written by Malcolm John.

The work in this thesis has been published as two articles in Journal of High Energy Physics **10** (2014) 097 and Journal of High Energy Physics **06** (2016) 131.

CHAPTER 1

CP violation and measurements of the CKM angle γ

The Standard Model (SM) of particle physics is a framework that describes all the known elementary particles and their interactions [4]. Its theoretical predictions have been successfully tested at increasingly rigorous levels of precision over the past decades. A recent significant milestone is the confirmation of the electroweak symmetry breaking mechanism with the discovery of the Higgs boson [5–8]. Despite its successes, on its own the SM appears to have a number of shortfalls, including its inability to incorporate gravitational interactions, provide a dark matter candidate or account for the large asymmetry in the abundance of matter and antimatter. These suggest that the current scheme of the SM is incomplete, hence the attempted confirmation of new physics (NP) theories that can account for these additional phenomena forms the driving force for experimental particle physics today.

Direct searches of new particles form one way in which NP processes can be identified, but this can only be achieved experimentally at energies above the threshold for their production. Of equal importance are precision measurements of parameters that have potential indirect contributions from NP. In general such measurements are able to place limits on NP theories at much higher energy scales than the direct searches. Furthermore, information about the structure of NP may be deduced if deviations from SM expectations are observed.

1.1 *CP* violation

The presence of our matter-dominated Universe with almost no antimatter content is a long-standing mystery in physics. The baryon-antibaryon asymmetry is typically parameterised in the ratio $\frac{N_B - N_{\bar{B}}}{N_\gamma} \approx \frac{N_B}{N_\gamma}$, where N_B , $N_{\bar{B}}$ and N_γ are the number densities of baryons, antibaryons and photons, respectively, and has been measured in astrophysical observations to be of order 10^{-10} [9]. The requirements for generating such an asymmetry from a symmetric initial state were first proposed by Andrei Sakharov in 1967 [10]. The three Sakharov Conditions for baryogenesis are baryon number violation, the violation of both charge (*C*) symmetry and the combined symmetry of charge and parity (*CP*), and a departure from thermal equilibrium. The *C* and *P* symmetries are defined by their operators: the charge-conjugation operator $\hat{C}$ replaces all particles by their antiparticles in the same state and the parity operator $\hat{P}$ reverses all of the spatial axes ($\mathbf{r} \rightarrow -\mathbf{r}$) used to describe the system. A system is said to be symmetric under *CP* if it remains unchanged under the combined operation of $\hat{C}\hat{P}$, and to violate *CP* if it does not.

The first evidence for *CP* violation was observed in the decays of neutral kaons in 1964 [11] and in the decays of *B* mesons by the BaBar and Belle collaborations in 2001 [12, 13]. To date, *CP* violation has only been observed in weak interactions in the quark sector. Experimental measurements of the amount of *CP* violation in the SM indicate that it only accounts for an asymmetry many orders of magnitude smaller than the observed 10^{-10} [14, 15]. Hence, NP theories must introduce new sources of *CP* violation and precise measurements of *CP* violation in the SM are crucial to further our understanding of the matter-antimatter problem.

1.2 The CKM matrix

The weak interaction of quarks can be described in terms of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. In a hadronic weak interaction, although it is the weak eigenstates of the quark that couple with the W boson, it is the flavour eigenstates (also known as the mass eigenstates) that form the observable quantities, because quarks hadronise and the observed hadrons are composed of quarks with a specific flavour (*e.g.* B^- meson $= b\bar{u}$). The misalignment between the flavour eigenstates (d, s, b) and weak eigenstates (d', s', b') is described by the CKM matrix, V_{CKM} , as follows:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.1)$$

The CKM matrix is complex, and each matrix element V_{ij} ($i = u, c, t; j = d, s, b$) describes the coupling of a $j \rightarrow i$ quark transition. Similarly the complex conjugate matrix element V_{ij}^* describes the coupling of a $\bar{j} \rightarrow \bar{i}$ antiquark transition.

The CKM matrix is unitary ($V_{\text{CKM}} V_{\text{CKM}}^\dagger = \mathbb{I}_3$) by construction, under the assumption that there are only three generations of up-type and down-type quarks with non-degenerate masses and that there are no flavour-changing neutral currents in the SM. Observing a violation of the unitary conditions would point to NP that opposes these assumptions. A complex 3×3 matrix has 18 parameters, which is reduced to 9 independent parameters due to the unitary condition. The phases of five quark fields can be redefined with the physics unchanged, leaving four independent parameters. A standard formalism is to write the CKM matrix in terms of 3 rotation angles ($\theta_{12}, \theta_{13}, \theta_{23}$) and a complex phase Δ :

$$V_{\text{CKM}} = \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\Delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\Delta} & 0 & c_{13} \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}, \quad (1.2)$$

where c_{ij} and s_{ij} are the cosine and sine of the rotation angle θ_{ij} . The overall complex phase Δ in V_{CKM} is responsible for all *CP*-violating phenomena in quark flavour-changing processes in the SM. The magnitudes of the matrix elements are [4]

$$|V_{\text{CKM}}| = \begin{pmatrix} 0.97425 \pm 0.00022 & 0.2253 \pm 0.0008 & 0.00413 \pm 0.00049 \\ 0.225 \pm 0.008 & 0.986 \pm 0.016 & 0.00411 \pm 0.00013 \\ 0.0084 \pm 0.0006 & 0.00400 \pm 0.00027 & 1.021 \pm 0.032 \end{pmatrix}. \quad (1.3)$$

The CKM elements are approximately unity along the diagonal, indicating a high probability of $t \rightarrow b$, $c \rightarrow s$ and $u \rightarrow d$ transitions, while the off-diagonal values have smaller values, indicating a lower probability of quark transitions amongst different generations. It is not currently understood why the CKM matrix possesses this apparent hierarchy. However, the relative sizes of the matrix elements can be written in a convenient form as an expansion of a parameter $\lambda = s_{12} \sim 0.2$. Every matrix element can then be closely approximated with its leading term in λ , which is known as the Wolfenstein parameterisation [16]. The leading order terms of the expansion are

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (1.4)$$

where λ , A , ρ and η are four real, independent parameters. For *CP* violation to occur, the presence of a complex phase corresponds to $\eta \neq 0$. As written, η only appears in V_{ub} and V_{td} up to $\mathcal{O}(\lambda^3)$.

Condition for direct *CP* violation

One mechanism by which *CP* violation manifests itself in particle decays is known as ‘direct *CP* violation’, where the magnitudes of the amplitudes of a decay and its charge-conjugate process differ. Consider a decay amplitude $\mathcal{A}$ which describes an initial particle i decaying to a final state f . The amplitude $\overline{\mathcal{A}}$ describes the charge-conjugate

process of antiparticle $\bar{i}$ decaying to a final state $\bar{f}$. *CP* violation is said to occur if the two processes occur at different rates ($|\mathcal{A}|^2 \neq |\bar{\mathcal{A}}|^2$).

Let an underlying $q' \rightarrow q$ quark transition be described by a decay amplitude $\mathcal{A}$ with an associated coupling $V_{qq'}$. The corresponding $\bar{q}' \rightarrow \bar{q}$ antiquark transition with the amplitude $\bar{\mathcal{A}}$ is associated with a coupling $V_{qq'}^*$. If the decay involves only a single flavour-changing quark transition then $|\mathcal{A}|^2 = |\bar{\mathcal{A}}|^2 \propto |V_{qq'}|^2$ and *CP* violation would not be observed, regardless of the value of Δ . Hence, although the overall complex phase Δ in V_{CKM} is necessary and sufficient for *CP* violation in the SM, it cannot be accessed experimentally in such a process.

In order to access the phase information, interfering amplitudes are required. Consider the case where there are two contributing Feynman diagrams to the amplitude $\mathcal{A}$. They proceed via different quark transitions and the two diagrams are written as amplitudes with magnitudes $A_{1,2}$, *CP*-violating phases $\phi_{1,2}$ and *CP*-conserving phases $\delta_{1,2}$, as illustrated in Fig. 1.1. The source of $\phi_{1,2}$ comes from the participating CKM factors and are related to Δ through Eq. (1.2). The strong phases $\delta_{1,2}$ come from the strong and electromagnetic interactions. $\mathcal{A}$ and $\bar{\mathcal{A}}$ can be written as:

$$\begin{aligned}\mathcal{A} &= A_1 e^{i(-\phi_1 + \delta_1)} + A_2 e^{i(-\phi_2 + \delta_2)}, \\ \bar{\mathcal{A}} &= A_1 e^{i(+\phi_1 + \delta_1)} + A_2 e^{i(+\phi_2 + \delta_2)},\end{aligned}\tag{1.5}$$

where the *CP*-violating phases $\phi_{1,2}$ have the opposite sign in the charge-conjugate decay amplitude. The difference between the charge-conjugate decay rates is

$$|\mathcal{A}|^2 - |\bar{\mathcal{A}}|^2 = \mathcal{A}^* \mathcal{A} - \bar{\mathcal{A}}^* \bar{\mathcal{A}} = 4A_1 A_2 \sin(\phi_2 - \phi_1) \sin(\delta_2 - \delta_1).\tag{1.6}$$

Equation (1.6) shows that $\phi_1 \neq \phi_2$ and $\delta_1 \neq \delta_2$ are both required for *CP* violation to be observed. This result can be generalised to show that *CP* violation will occur if two or more contributing amplitudes with different strong and weak phases interfere in the

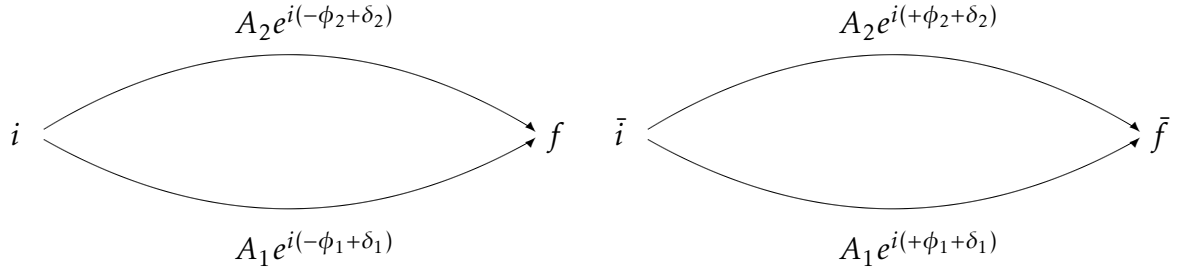


Figure 1.1. Illustration of two amplitudes with magnitude A , weak phase ϕ and strong phase δ contributing to the transition of an initial state i to a final state f . The corresponding amplitudes in the charge-conjugate transition of $\bar{i} \rightarrow \bar{f}$ are shown on the right.

decay of a given initial state i to final state f . As can be seen in Eq. (1.6), the phase differences, and therefore the relative phases between CKM matrix elements, can be accessed by studying asymmetries between the particle and antiparticle decay rates, which are called *CP* asymmetries.

1.3 The CKM unitarity triangle

An important constraint on the relative phases between the CKM matrix elements comes from the property that the CKM matrix is unitary, with $V_{\text{CKM}} V_{\text{CKM}}^\dagger = \mathbb{I}_3$. The unitary condition provides nine equations. There are three diagonal conditions, $\sum_j V_{ij} V_{kj}^* = \delta_{ik}$, which ensure that the total probability of an up-type quark transition to any down-type quark is unity. This is clearly seen when writing out the equation where $i = u$, $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1$. There are six off-diagonal conditions, $\sum_{i=u,c,t} V_{ij} V_{ik}^* = \delta_{jk}$, translating to six triangles in the complex plane. Writing out one of them explicitly,

$$V_{ud} V_{ub}^* + V_{cd} V_{cb}^* + V_{td} V_{tb}^* = 0. \quad (1.7)$$

This equation is known as the *unitarity triangle* (UT). This triangle is most commonly studied because it has sides which have comparable lengths. It is convenient to normalise one of the sides to unit length as shown in Fig. 1.2; in the Wolfenstein paramet-

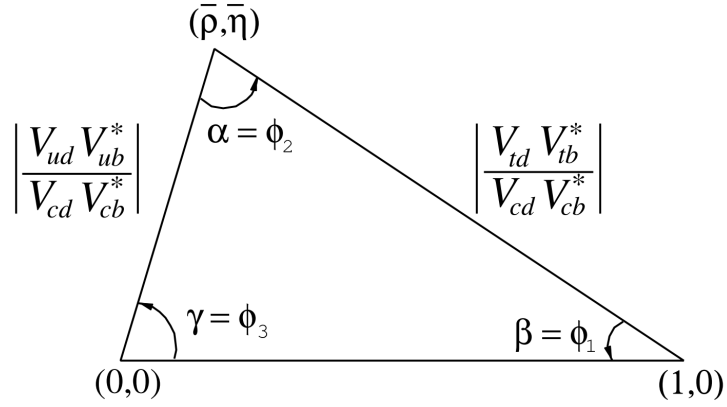


Figure 1.2. The CKM unitarity triangle with angles α , β and γ , and vertices $(0,0)$, $(1,0)$ and $(\bar{\rho}, \bar{\eta})$. Reproduced from [4].

terisation the apex of the triangle is then equal to

$$\bar{\rho} + i\bar{\eta} = \left(1 - \frac{\lambda^2}{2}\right)(\rho + i\eta). \quad (1.8)$$

The angles of the UT are denoted by α , β and γ , and have the values

$$\alpha \equiv \arg\left(-\frac{V_{td} V_{tb}^*}{V_{ud} V_{ub}^*}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*}\right). \quad (1.9)$$

The area of the unitarity triangle quantifies the amount of CP violation due to SM weak interactions in the quark sector [17]. In the Wolfenstein parameterisation, up to order λ^4 , all the CKM elements involved are real except V_{ub}^* and V_{td} , which also contain the leading order terms in η . In this approximation, the angles become

$$\alpha \approx \arg\left(-\frac{V_{td}}{V_{ub}^*}\right), \quad \beta \approx \arg(-V_{td}^*), \quad \gamma \approx \arg(-V_{ub}^*). \quad (1.10)$$

Precision measurements of each of the three angles and sides of the UT are necessary to test the unitary condition of Eq. (1.7) by over-constraining it. In other words, finding that $\alpha + \beta + \gamma \neq 180^\circ$, or that one of the sides is too long or short, would be a clear sign of NP. The angle α can be accessed through the interference between the tree

and penguin topologies of the $b \rightarrow u\bar{u}d$ transition. The precision on α is driven by measurements with $B^0 \rightarrow \rho^+\rho^-$ decays, where α is determined from the oscillation amplitude of the *CP* asymmetry as a function of the B^0 decay time. The angle β can be measured using $B^0 \rightarrow J/\psi K^0$ ($b \rightarrow c\bar{c}s$) decays, where there is interference between the amplitude for a direct decay, $B^0 \rightarrow J/\psi K^0$, and that where the B^0 meson has mixed before decay, $B^0 \rightarrow \bar{B}^0 \rightarrow J/\psi K^0$. As with measurements of α , the value of β determines the oscillation amplitude of the time-dependent *CP* asymmetry. The angle γ is the parameter studied in this thesis and various methods to measure it are discussed in Sect. 1.4.

In 2012, at the time of embarking on this research programme, the values of α , β and γ were [18]

$$\alpha = (88.7^{+4.6}_{-4.3})^\circ, \quad \beta = (21.38^{+0.79}_{-0.77})^\circ, \quad \gamma = (66 \pm 12)^\circ. \quad (1.11)$$

The current experimental status is

$$\alpha = (87.6^{+3.5}_{-3.3})^\circ, \quad \beta = (21.85^{+0.68}_{-0.67})^\circ, \quad \gamma = (73.2^{+6.3}_{-7.0})^\circ, \quad (1.12)$$

where the updated value of γ includes one of the measurements in this thesis. These values are used as inputs to a global fit for $(\bar{\rho}, \bar{\eta})$, the apex of the UT, using a frequentist technique by the CKMfitter group [18]. The measurements of α , β , γ are combined with measurements related to the sides of the triangle, which are $|V_{ub}|$ and $\Delta m_{d,s}$ (which describes mixing in $B_{d,s}$ mesons), and ε_K (which describes *CP* violation in the neutral kaon system). The constraints on these parameters are shown in Fig. 1.3. The resulting constraints on the apex position are $\bar{\rho} = 0.140^{+0.027}_{-0.026}$ and $\bar{\eta} = 0.343^{+0.015}_{-0.014}$ [18]. Consistent results are obtained using a Bayesian technique by the UTfit collaboration [19] and the results are also shown in Fig. 1.3. The compatibility between all of the different parameters shown indicates that the measurements are consistent with the unitary

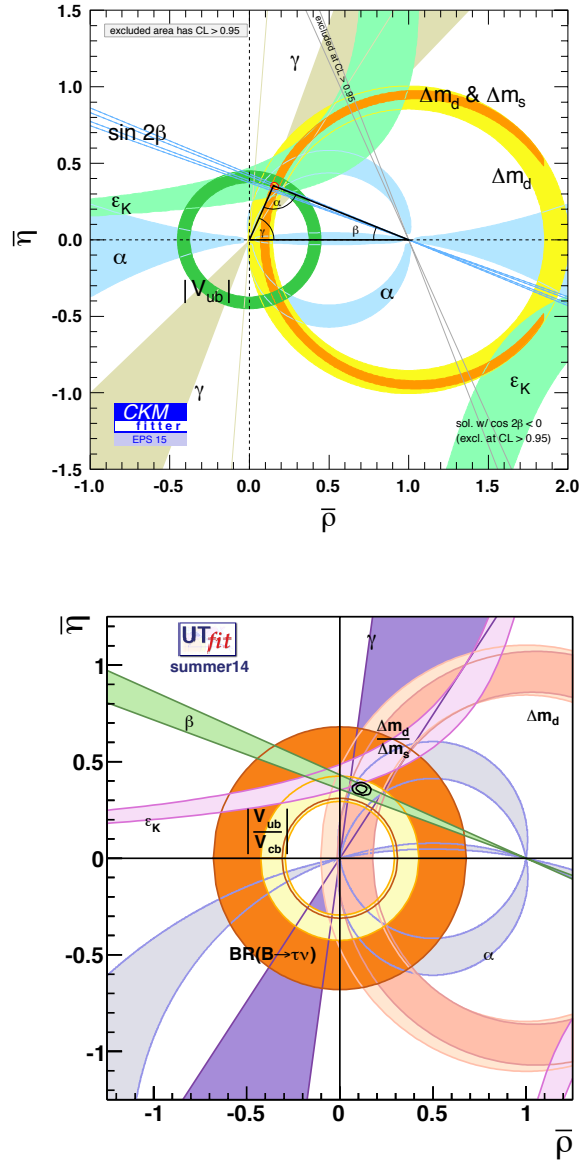


Figure 1.3. The most recent global fits of the unitarity triangle parameters from (top) the CKMfitter collaboration [18] and (bottom) the UTfit collaboration [19]. The various parameters are defined in the main text.

condition given by Eq. (1.7) and that *CP* violation is described by a single complex phase. Crucially however, several of the measurements, especially of the angle γ , are currently statistically limited.

1.4 The CKM angle γ

A striking feature of the results of Eqs. (1.11) and (1.12) is that the precision on γ is much poorer than that of α and β . Precision measurements of γ are of great interest and importance because it is the only angle out of the three that can be measured cleanly in tree-level processes with a negligible contribution from loop diagrams. Unlike α and β , the definition of γ does not contain any V_{ti} terms, which are accessed in loop diagrams of hadron processes as $m_t \gg m_b$. Since NP is expected to manifest through the loop contributions of new particles while leaving tree-level processes unchanged, the value of γ at tree-level is a ‘standard candle’ of the SM. Another attractive feature of direct measurements of γ is that the associated hadronic parameters can be determined experimentally and do not require additional theoretical inputs such as lattice calculations, which measurements of UT parameters depend on.

The benchmark value of γ can be compared to the indirect measurement of γ , which can be determined using a global fit from measurements of all the other UT parameters, excluding γ , assuming the unitary condition. Since the other parameters in the global fit are sensitive to loop-level processes, an inconsistency between the two values would point to the presence of NP. The value of γ from indirect measurements is $(66.85^{+0.94}_{-3.44})^\circ$ [18], which marks a 6° difference in central values compared to the direct measurement. Another way to determine γ indirectly from parameters sensitive to NP has recently been shown [20]. Assuming the unitarity condition, measurements of the angle β and the length $\left| \frac{V_{td} V_{tb}^*}{V_{cd} V_{cb}^*} \right|$ (the side opposite to γ , which is measured from the ratio of neutral B meson mixing parameters with input from lattice QCD [21]), are used to calculate a value of $\gamma = (62.7 \pm 2.1)^\circ$. Here, a 10° difference between the central values is present. However, the 7° uncertainty on the direct measurement means that the comparison will only become meaningful once the error has been reduced to the few-degree level.

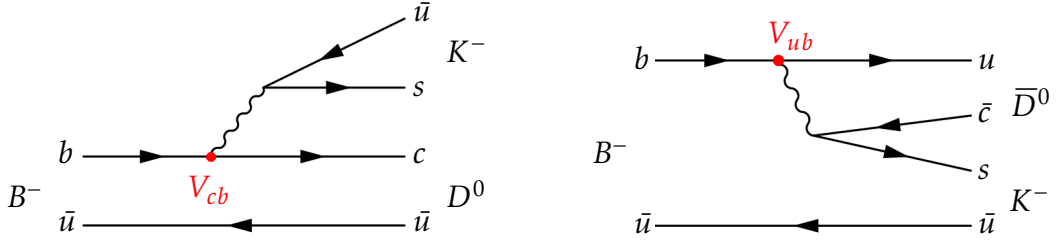


Figure 1.4. Feynman diagrams for the (left) $B^- \rightarrow D^0 K^-$ and (right) $B^- \rightarrow \bar{D}^0 K^-$ amplitudes.

1.5 Tree-level determination of γ using $B \rightarrow DK$ decays

From the definition of γ in Eq. (1.9), it is evident that it can be accessed in decays where there is interference between $b \rightarrow u$ and $b \rightarrow c$ quark transitions. Such interference occurs in $B^- \rightarrow DK^-$ decays, where D indicates a superposition of D^0 and $\bar{D}^0$ mesons that decay to a common final state f . The Feynman diagrams for the two contributing amplitudes, $B^- \rightarrow D^0 K^-$ and $B^- \rightarrow \bar{D}^0 K^-$, are shown in Fig. 1.4. It is noted that the CKM matrix elements involved in the decay are in fact $\frac{V_{cs} V_{ub}^*}{V_{us} V_{cb}^*}$, which differs from the definition of γ by the ratio [22]

$$\frac{V_{cd} V_{cs}}{V_{ud} V_{us}} = -1 + \frac{A\lambda^4}{2} - A^2 \lambda^5 \left(\rho + i\eta - \frac{1}{2} \right) + \mathcal{O}(\lambda^6). \quad (1.13)$$

Thus the leading order correction on γ is of the order $\lambda^5 \sim 10^{-4}$ which is small compared to the current precision of experimental measurements and can be ignored. Further, the theoretical uncertainty due to electroweak corrections is $\mathcal{O}(10^{-7})$ and can also be neglected [23].

The amplitudes contributing to the B^- decay can be written as

$$A(B^- \rightarrow D^0 K^-) = A_B, \quad (1.14)$$

$$A(B^- \rightarrow \bar{D}^0 K^-) = A_B r_B e^{i(\delta_B - \gamma)}, \quad (1.15)$$

where r_B is the ratio between the magnitude of the two amplitudes, $|A(B^- \rightarrow \bar{D}^0 K^-)|/|A(B^- \rightarrow D^0 K^-)|$, δ_B is the strong-phase difference and γ is the weak-phase difference. The amplitudes for the *CP*-conjugate B^+ decay can be obtained under the transformation $\gamma \rightarrow -\gamma$ and are

$$A(B^+ \rightarrow \bar{D}^0 K^+) = A_B, \quad (1.16)$$

$$A(B^+ \rightarrow D^0 K^+) = A_B r_B e^{i(\delta_B + \gamma)}. \quad (1.17)$$

The size of r_B dictates the size of the interference between the two amplitudes, and hence the sensitivity to γ . For $B^\pm \rightarrow DK^\pm$ decays, $r_B \sim 0.1$. This value is driven by the ratio of the participating CKM factors $\frac{|V_{cs} V_{ub}^*|}{|V_{us} V_{cb}^*|} \sim 0.4$, suppressed by a factor of three due to colour suppression in the $B^- \rightarrow \bar{D}^0 K^-$ amplitude, with corrections due to other strong interaction effects [24]. The amplitudes of the subsequent D^0 and $\bar{D}^0$ decays to a common final state f can be written

$$A(\bar{D}^0 \rightarrow f) = A_D, \quad (1.18)$$

$$A(D^0 \rightarrow f) = A_D r_D e^{i\delta_D}, \quad (1.19)$$

where $0 \leq r_D \leq 1$ and δ_D are the relative magnitude and strong-phase difference between the two decays, respectively. The level of direct *CP* violation in D decays is expected to be negligibly small [25], hence $A(D^0 \rightarrow f) \equiv A(\bar{D}^0 \rightarrow \bar{f})$ and $A(\bar{D}^0 \rightarrow f) \equiv A(D^0 \rightarrow \bar{f})$. Combining the above expressions for the B and D decays, as depicted in Fig. 1.5, the total amplitude for the B^- and B^+ decay is

$$\begin{aligned} A(B^- \rightarrow D(\rightarrow f)K^-) &\equiv A(B^- \rightarrow D^0 K^-)A(D^0 \rightarrow f) + A(B^- \rightarrow \bar{D}^0 K^-)A(\bar{D}^0 \rightarrow f) \\ &= A_B A_D r_D e^{i\delta_D} + A_B r_B e^{i(\delta_B - \gamma)} A_D, \end{aligned} \quad (1.20)$$

$$\begin{aligned} A(B^+ \rightarrow D(\rightarrow \bar{f})K^+) &\equiv A(B^+ \rightarrow \bar{D}^0 K^+)A(\bar{D}^0 \rightarrow \bar{f}) + A(B^+ \rightarrow D^0 K^+)A(D^0 \rightarrow \bar{f}) \\ &= A_B A_D + A_B r_B e^{i(\delta_B + \gamma)} A_D r_D e^{i\delta_D}. \end{aligned} \quad (1.21)$$

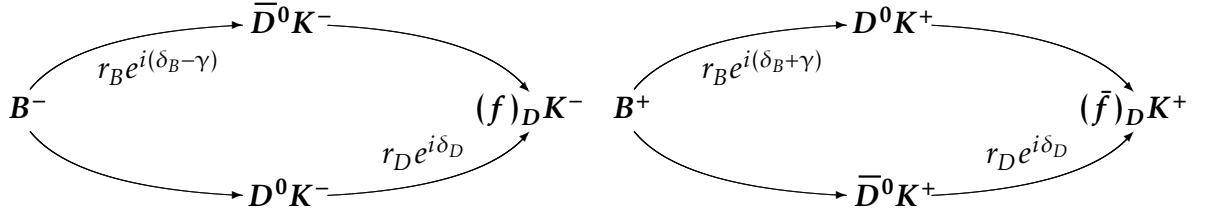


Figure 1.5. The interfering amplitudes in (left) a $B^- \rightarrow DK^-$ and (right) a $B^+ \rightarrow DK^+$ decay, where D is a superposition of a D^0 and $\bar{D}^0$ meson. The relative amplitude magnitude ratios (r_B , r_D), relative strong phases (δ_B , δ_D) and relative weak phase (γ) are indicated.

Therefore, the rates of $B^- \rightarrow DK^-$ and $B^+ \rightarrow DK^+$ decays are:

$$|A(B^- \rightarrow D(\rightarrow f)K^-)|^2 = |A_B|^2 |A_D|^2 [r_B^2 + r_D^2 + 2r_B r_D \cos(\delta_B - \gamma - \delta_D)], \quad (1.22)$$

$$|A(B^+ \rightarrow D(\rightarrow \bar{f})K^+)|^2 = |A_B|^2 |A_D|^2 [r_B^2 + r_D^2 + 2r_B r_D \cos(\delta_B + \gamma - \delta_D)]. \quad (1.23)$$

The expressions are consistent with the general result discussed in Sect. 1.2, where the interference of two amplitudes with different strong and weak phases result in direct CP violation.

So far the D final state f has not been specified and indeed several decay modes have been used to measure γ . It was first proposed to use CP eigenstates (the so-called ‘GLW’ method [26, 27], named after the proponents Gronau, London and Wyler), since $r_D = 1$ and $\delta_D = 0$ (π) are known for a CP -even (odd) eigenstate. Any CP asymmetries, defined as

$$A_{CP} = \frac{\Gamma(B^- \rightarrow D(\rightarrow f)K^-) - \Gamma(B^+ \rightarrow D(\rightarrow \bar{f})K^+)}{\Gamma(B^- \rightarrow D(\rightarrow f)K^-) + \Gamma(B^+ \rightarrow D(\rightarrow \bar{f})K^+)}, \quad (1.24)$$

would depend only on r_B , δ_B and γ , allowing them to be measured. However, the size of the interference, determined by $r_B \sim 0.1$, is relatively small and therefore sensitivity to γ is limited. A second method was proposed (the ‘ADS’ method, after Atwood, Dunitz and Soni [28–30]), where interference is enhanced in $B^- \rightarrow D(\rightarrow K^+\pi^-)K^-$ decays. The colour-favoured $B^- \rightarrow D^0 K^-$ decay is followed by a doubly Cabibbo-suppressed

$D^0 \rightarrow K^+\pi^-$ decay, while the colour-suppressed $B^- \rightarrow \bar{D}^0 K^-$ decay is followed by a Cabibbo-favoured $\bar{D}^0 \rightarrow K^+\pi^-$ decay, hence the amplitudes of the two decay paths are of similar size. The values of $r_D \sim 0.06$ and δ_D for $D \rightarrow K\pi$ are measured in a dedicated analysis and used as an input to the measurement. Although the ADS method provides large interference effects, it is limited by the small $\mathcal{O}(10^{-4})$ branching fraction of the suppressed $D^0 \rightarrow K^+\pi^-$ decay. A third way to measure γ , the ‘GGSZ’ method, (after Giri, Grossman, Soffer and Zupan [1]) makes use of Cabibbo-favoured $D \rightarrow K_S^0 \pi^+ \pi^-$ or $K_S^0 K^+ K^-$ decays, collectively known as $D \rightarrow K_S^0 h^+ h^-$. This is the method used in this thesis and the formalism and the advantages are discussed in more detail in Sect. 1.6.

A common feature to all the above methods is that the hadronic parameters for the B decays, r_B and δ_B , are determined experimentally, while external inputs for r_D and δ_D are provided where required. The measured values for the B hadronic parameters are summarised in Table 1.1. In addition to $B^\pm \rightarrow DK^\pm$ decays, similar analyses have been performed with $B^\pm \rightarrow DK^{*\pm}$, $B^\pm \rightarrow D^* K^\pm$ and $B^\pm \rightarrow D\pi^\pm$ using a similar formalism as that outlined above, but with different hadronic parameters for the different modes. Neutral B decays have also been used, including $B^0 \rightarrow DK^{*0}$, where the flavour of the B meson at decay is identified unambiguously using the charge of the K^{*0} daughter kaon. (A B^0 meson produces a K^{*0} which in turn produces a K^+ .) The $B^0 \rightarrow DK^{*0}$ decay mode is also studied in this thesis to perform a second measurement of γ using the GGSZ method. The associated hadronic parameters, r_{B^0} and δ_{B^0} , have been measured and are summarised in Table 1.1.

1.6 Measuring γ with the GGSZ method

In this section, the method used to measure γ with $B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-$ and $B^0 \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^{*0}$ decays is outlined. The key difference between Eqs. (1.20) and (1.21) and the total amplitude for these D final states is that, in the multibody decay $D \rightarrow$

Table 1.1. Summary of measurements of r_B , δ_B , r_{B^0} and δ_{B^0} from BaBar, Belle and LHCb [25]. The values of r_{B^0} have been measured in slightly different regions of the $B^0 \rightarrow DK^+\pi^-$ Dalitz plot at the three experiments.

Parameter	BaBar	Belle	LHCb
r_B	$0.092^{+0.013}_{-0.012}$	$0.112^{+0.014}_{-0.015}$	$0.1006^{+0.0059}_{-0.0060}$
δ_B [°]	105^{+16}_{-17}	116^{+18}_{-21}	$141.1^{+6.1}_{-7.7}$
r_{B^0}	0.26 ± 0.08	< 0.4 at 95% CL	$0.217^{+0.044}_{-0.048}$
δ_{B^0} [°]	-	-	189^{+24}_{-20}

$K_S^0 h^+ h^-$, the amplitude A_D is no longer fully specified by a single magnitude and a phase, which was assumed in the derivation in Sect. 1.5.

Consider the decay of the scalar D^0 meson to three scalar particles K_S^0 , π^+ and π^- . In the rest frame of the D^0 meson, the final state contains 12 degrees of freedom corresponding to the four-momenta of the three decay products. Conservation of energy and three-momentum imposes four constraints, and there are additionally three constraints due to the known invariant masses of the K_S^0 and two pions. A further three constraints are imposed if an integration is made over three Euler angles since the decaying particle and the final state particles are scalar and the spatial orientation is uniform. Hence, the final state contains two degrees of freedom and the D^0 decay amplitude can be fully specified as a function of two kinematic variables.

1.6.1 Dalitz plots for three-body D decays

The choice of the two kinematic variables to describe the decay is motivated by the use of a two-dimensional *Dalitz plot*, which provides a powerful parameterisation to visualise the $D \rightarrow K_S^0 h^+ h^-$ decay phase space [31]. Its axes are the Lorentz invariant squared masses of any pair of the final state particles defined by $m^2(AB) \equiv (p_A + p_B)^2$, where p_A and p_B are the four-momenta of particles A and B . For $D^0 \rightarrow K_S^0 h^+ h^-$ one can choose from any pair of $m^2(K_S^0 h^+)$, $m^2(K_S^0 h^-)$ and $m^2(h^+ h^-)$. For symmetry reasons

(further explained in Sect. 1.6.2) we choose the first two, and as a shorthand define $m_+^2 \equiv m^2(K_S^0 h^+)$ and $m_-^2 \equiv m^2(K_S^0 h^-)$. In this parameterisation, the partial decay rate provides direct information about the matrix element of the decay, $\mathcal{M}$. Using the general result for a three-body phase space, the $D \rightarrow K_S^0 h^+ h^-$ partial decay rate is

$$d\Gamma = \frac{1}{(2\pi)^3} \frac{1}{32m_{D^0}^3} |\mathcal{M}|^2 dm_-^2 dm_+^2, \quad (1.25)$$

where m_{D^0} is the D^0 invariant mass.

Figure 1.6 shows a sample of $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ and $D^0 \rightarrow K_S^0 K^+ K^-$ decays plotted in (m_-^2, m_+^2) space. The data are bound by the kinematically allowed region due to the constraint from m_{D^0} . Internal structure in the Dalitz plot is observed, which results from:

1. nonresonant contributions;
2. resonant contributions of the form $D^0 \rightarrow AR(\rightarrow BC)$, where R is a short-lived intermediate resonance, appearing as bands of increased density whose distributions depend on the mass, spin and decay products of R ;
3. constructive or destructive interference effects between resonances and/or the nonresonant component.

In the $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ Dalitz plot, a clear vertical band can be seen at $m^2(K_S^0 \pi^-) \sim 0.80 \text{ GeV}^2/c^4$ which is the spin-1 $K^*(892)^-$ resonance decaying to $K_S^0 \pi^-$. Similarly, the diagonal band close to the boundary in the $D^0 \rightarrow K_S^0 K^+ K^-$ Dalitz plot corresponds to the resonant $\phi(1020) \rightarrow K^+ K^-$ decay. In both Dalitz plots, there is more than one resonance contributing to the overall amplitude. Each resonance R contributes a magnitude $a_R(m_-^2, m_+^2)$ and phase $\delta_R(m_-^2, m_+^2)$ at the point (m_-^2, m_+^2) on the $D^0 \rightarrow K_S^0 h^+ h^-$ Dalitz plot. Since each resonance contributes to the same $K_S^0 h^+ h^-$ final state, they interfere such that the resulting amplitude $A_D(m_-^2, m_+^2)$ is a coherent sum of all the contributing amplitudes

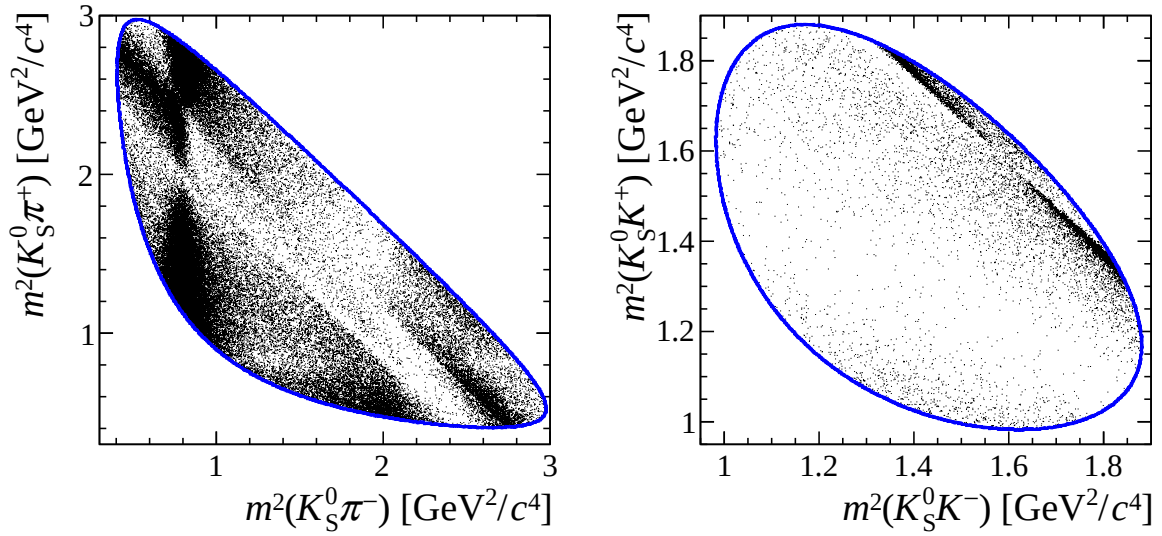


Figure 1.6. Example Dalitz plot of (left) $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ and (right) $D \rightarrow K_S^0 K^+ K^-$ decays from LHCb data. The blue line corresponds to the kinematic boundary imposed by the D^0 mass.

at the point (m_-^2, m_+^2) . Therefore,

$$\begin{aligned}
 A_D(m_-^2, m_+^2) &= a_{NR}(m_-^2, m_+^2) e^{i\delta_{NR}(m_-^2, m_+^2)} + \sum_R a_R(m_-^2, m_+^2) e^{i\delta_R(m_-^2, m_+^2)}, \\
 &= a(m_-^2, m_+^2) e^{i\delta(m_-^2, m_+^2)},
 \end{aligned} \tag{1.26}$$

where a_{NR} and δ_{NR} are the magnitude and phase of the nonresonant component, and a and δ are the summed magnitude and phase. There are a number of different conventions with which to model the various components [4]. The isobar formalism has been used to model $D \rightarrow K_S^0 h^+ h^-$ decays, where the resonances can be approximated with relativistic Breit-Wigner functions, with factors to take into account the spins of the particles and form factors to take into account the relative momenta of the decay products.

1.6.2 Measurement with $B \rightarrow DK$ decays

The following discussion summarises the derivation outlined in Ref. [1], where the GGSZ method was first proposed.

The expressions for the $B^\pm$ decay amplitudes are those written in Eqs. (1.14)–(1.17). Following the notation given in Eq. (1.26), the amplitudes for the CP -conjugate $D \rightarrow K_S^0 h^+ h^-$ decays are:

$$A(D^0 \rightarrow K_S^0 h^+ h^-) = A_D(m_-^2, m_+^2) = a(m_-^2, m_+^2) e^{i\delta(m_-^2, m_+^2)}, \quad (1.27a)$$

$$A(\bar{D}^0 \rightarrow K_S^0 h^+ h^-) = \bar{A}_D(m_-^2, m_+^2) = \bar{a}(m_-^2, m_+^2) e^{i\bar{\delta}(m_-^2, m_+^2)} = A_D(m_+^2, m_-^2). \quad (1.27b)$$

The equality $\bar{A}_D(m_-^2, m_+^2) = A_D(m_+^2, m_-^2)$ occurs because the final state has spin zero. The equality indicates that the expressions for the $\bar{D}^0$ decay amplitude can be obtained simply under the transformation $m_+^2 \leftrightarrow m_-^2$ and this symmetry is the reason for choosing the pair of variables m_-^2 and m_+^2 to specify the D^0 decay amplitude. Hence, the amplitude of the cascade decay $B^- \rightarrow DK^-$ with $D \rightarrow K_S^0 h^+ h^-$ at the coordinate (m_-^2, m_+^2) is

$$\begin{aligned} A(B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-) \\ &\equiv A(B^- \rightarrow D^0(\rightarrow K_S^0 h^+ h^-) K^-) + A(B^- \rightarrow \bar{D}^0(\rightarrow K_S^0 h^+ h^-) K^-) \\ &= A_B A_D(m_-^2, m_+^2) + A_B r_B e^{i(\delta_B - \gamma)} \bar{A}_D(m_-^2, m_+^2) \\ &\propto a(m_-^2, m_+^2) e^{i\delta(m_-^2, m_+^2)} + r_B e^{i(\delta_B - \gamma)} \bar{a}(m_-^2, m_+^2) e^{i\bar{\delta}(m_-^2, m_+^2)}. \end{aligned} \quad (1.28)$$

The partial width for the B^- decay is given by

$$\begin{aligned} \frac{d\Gamma(B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-)}{dm_-^2 dm_+^2} &= |A(B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-)|^2 \\ &\propto a^2(m_-^2, m_+^2) + r_B^2 \bar{a}^2(m_-^2, m_+^2) + \\ &\quad 2a(m_-^2, m_+^2) \bar{a}(m_-^2, m_+^2) r_B \cos[(\delta_B - \gamma) - (\delta(m_-^2, m_+^2) - \bar{\delta}(m_-^2, m_+^2))]. \end{aligned} \quad (1.29)$$

It is useful to define the strong-phase difference between the D^0 and $\bar{D}^0$ amplitudes:

$$\delta_D(m_-^2, m_+^2) = \delta(m_-^2, m_+^2) - \bar{\delta}(m_-^2, m_+^2) = \delta(m_-^2, m_+^2) - \delta(m_+^2, m_-^2), \quad (1.30)$$

such that, and henceforth implying the Dalitz plot coordinates (m_-^2, m_+^2) ,

$$\begin{aligned} \frac{d\Gamma(B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-)}{dm_-^2 dm_+^2} &\propto a^2 + r_B^2 \bar{a}^2 + 2a\bar{a}r_B \cos[(\delta_B - \gamma) - \delta_D], \\ &\propto a^2 + r_B^2 \bar{a}^2 + 2a\bar{a}r_B [\cos(\delta_B - \gamma) \cos \delta_D + \sin(\delta_B - \gamma) \sin \delta_D]. \end{aligned} \quad (1.31)$$

Similarly, the amplitude for the B^+ decay is

$$\begin{aligned} A(B^+ \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^+) &\equiv A(B^+ \rightarrow \bar{D}^0(\rightarrow K_S^0 h^+ h^-) K^+) + A(B^+ \rightarrow D^0(\rightarrow K_S^0 h^+ h^-) K^+) \\ &= A_B \bar{A}_D(m_-^2, m_+^2) + A_B r_B e^{i(\delta_B + \gamma)} A_D(m_-^2, m_+^2) \\ &\propto \bar{a}(m_-^2, m_+^2) e^{i\bar{\delta}(m_-^2, m_+^2)} + r_B e^{i(\delta_B + \gamma)} a(m_-^2, m_+^2) e^{i\delta(m_-^2, m_+^2)}, \end{aligned} \quad (1.32)$$

and the partial width is

$$\begin{aligned} \frac{d\Gamma(B^+ \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^+)}{dm_-^2 dm_+^2} &\propto \bar{a}^2 + r_B^2 a^2 + 2a\bar{a}r_B \cos[(\delta_B + \gamma) + \delta_D] \\ &\propto \bar{a}^2 + r_B^2 a^2 + 2a\bar{a}r_B [\cos(\delta_B + \gamma) \cos \delta_D - \sin(\delta_B + \gamma) \sin \delta_D]. \end{aligned} \quad (1.33)$$

Hence, in the presence of *CP* violation and when $r_B \neq 0$, the $D \rightarrow K_S^0 h^+ h^-$ Dalitz plot distribution will no longer be the same for $B^- \rightarrow DK^-$ and $B^+ \rightarrow DK^+$ decays under the transformation $m_+^2 \leftrightarrow m_-^2$, and the asymmetry will depend on the Dalitz plot coordinate.

One strategy is to measure the asymmetries in integrated regions of the Dalitz plot. The property that the interference term $A_D(m_-^2, m_+^2) A_D^*(m_+^2, m_-^2)$ is symmetric under the exchange $(m_-^2, m_+^2) \leftrightarrow (m_+^2, m_-^2)$ is exploited by defining bin regions that are symmetric about the line $m_-^2 = m_+^2$. An example of such a binning is shown in Fig. 1.7. For

convenience, the bins are labelled $i = \{1, 2, \dots, k\}$ in the region $m_-^2 \geq m_+^2$, and a further k bins are labelled $i = \{-1, -2, \dots, -k\}$ lying symmetrically opposite, such that there are $2k$ bins in total. In this arrangement of bins, the transformation $m_+^2 \leftrightarrow m_-^2$ corresponds to the exchange of bins labelled $i \leftrightarrow -i$. Defining integrals over the region of the Dalitz plot labelled bin i ,

$$T_i = \int_i dm_-^2 dm_+^2 a^2(m_-^2, m_+^2), \quad (1.34)$$

$$T_{-i} = \int_{-i} dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) = \int_{+i} dm_-^2 dm_+^2 \bar{a}^2(m_-^2, m_+^2), \quad (1.35)$$

$$c_i = \frac{1}{\sqrt{T_{+i} T_{-i}}} \int_i dm_-^2 dm_+^2 a(m_-^2, m_+^2) \bar{a}(m_-^2, m_+^2) \cos \delta_D(m_-^2, m_+^2), \quad (1.36)$$

$$s_i = \frac{1}{\sqrt{T_{+i} T_{-i}}} \int_i dm_-^2 dm_+^2 a(m_-^2, m_+^2) \bar{a}(m_-^2, m_+^2) \sin \delta_D(m_-^2, m_+^2). \quad (1.37)$$

The terms c_i and s_i contain $\cos \delta_D$ and $\sin \delta_D$ and are related by $c_i = c_{-i}$ and $s_i = -s_{-i}$ for the symmetric bin regions discussed. The physics parameters of interest, r_B , δ_B and γ , are translated into Cartesian parameters given by

$$x_{\pm} \equiv r_B \cos(\delta_B \pm \gamma), \quad (1.38)$$

$$y_{\pm} \equiv r_B \sin(\delta_B \pm \gamma). \quad (1.39)$$

The motivation for using these Cartesian parameters is given later. Replacing the terms in Eqs. (1.31) and (1.33) by these integrals, the equations for the rate of events in bins i and $-i$ of the Dalitz plots for B^- and B^+ decays become:

$$\Gamma_{+i}(B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-) \propto [T_{+i} + (x_-^2 + y_-^2) T_{-i} + 2\sqrt{T_{+i} T_{-i}}(x_- c_{+i} + y_- s_{+i})], \quad (1.40a)$$

$$\Gamma_{-i}(B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-) \propto [T_{-i} + (x_-^2 + y_-^2) T_{+i} + 2\sqrt{T_{+i} T_{-i}}(x_- c_{-i} + y_- s_{-i})], \quad (1.40b)$$

$$\Gamma_{+i}(B^+ \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^+) \propto [T_{-i} + (x_+^2 + y_+^2) T_{+i} + 2\sqrt{T_{+i} T_{-i}}(x_+ c_{+i} - y_+ s_{+i})], \quad (1.40c)$$

$$\Gamma_{-i}(B^+ \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^+) \propto [T_{+i} + (x_+^2 + y_+^2) T_{-i} + 2\sqrt{T_{+i} T_{-i}}(x_+ c_{-i} - y_+ s_{-i})]. \quad (1.40d)$$

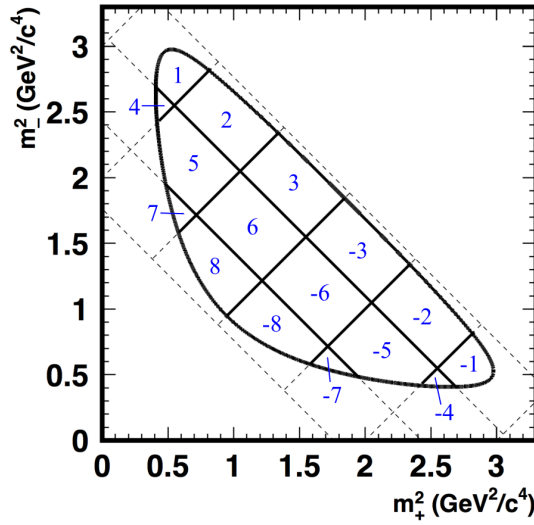


Figure 1.7. Example binning of the $D \rightarrow K_S^0 \pi^+ \pi^-$ Dalitz plot, where the bin regions are symmetric about the line $m_-^2 = m_+^2$. There are $k = 8$ bins in each half of the Dalitz plot, and the bins in the region $m_-^2 \leq m_+^2$ are labelled with negative signs. Reproduced from [2] and annotated.

For $i = \{1, 2, \dots, k\}$, there are $4 \times k$ equations in total. The parameters $T_{\pm i}$ can be determined by measuring the decay rates of flavour-tagged $D^0 \rightarrow K_S^0 h^+ h^-$ decays, *i.e.* where the D meson can be identified as D^0 or $\bar{D}^0$. By measuring the $B \rightarrow DK$ decay rates in each bin, there are $2k + 3$ unknowns in Eq. (1.40), which are c_i , s_i , r_B , δ_B and γ (noting the relations $c_i = c_{-i}$ and $s_i = -s_{-i}$). For $k \geq 2$ there are a greater number of equations than unknowns and γ can be determined. But since the goal is to obtain the best possible precision on γ , it is preferable to perform dedicated measurements of c_i and s_i and then use them as inputs to the analysis.

1.6.3 Determination of the strong-phase difference

Measurements of c_i and s_i for $D \rightarrow K_S^0 \pi^+ \pi^-$ and $D \rightarrow K_S^0 K^+ K^-$ decays have been performed at the CLEO experiment [3]. The results, including the optimisation of the Dalitz plot binning schemes, are described below.

The CLEO experiment was a charm factory that collected data from $e^+ e^-$ collisions at a centre-of-mass energy of 3.77 GeV, where $e^+ e^-$ annihilation produces the $c\bar{c}$ resonance,

$\psi(3770)$, at threshold. The $\psi(3770)$ has quantum numbers $J^{PC} = 1^{--}$ and subsequently decays to a $D\bar{D}$ pair. If no other particles are produced in the e^+e^- interaction, then in the CP -conserving strong decay $\psi(3770) \rightarrow D\bar{D}$ the final state has the properties $\hat{C}|D\bar{D}\rangle = (-1)|D\bar{D}\rangle$, $\hat{P}|D\bar{D}\rangle = (-1)|D\bar{D}\rangle$, $\hat{C}\hat{P}|D\bar{D}\rangle = (+1)|D\bar{D}\rangle$ [32], where $\hat{C}$ and $\hat{P}$ are the charge-conjugation and parity operators defined in Sect. 1.1. Selection rules apply to the subsequent decays of the D mesons and the $D\bar{D}$ pair is said to be *quantum correlated*. Consider the case where the two D mesons decay to final states f_a and f_b with CP eigenvalues η_a and η_b . It follows that $\hat{C}\hat{P}|f_a f_b\rangle = \eta_a \eta_b (-1)^L |f_a f_b\rangle$, where $L = 1$ is the relative orbital angular momentum between the two D mesons due to the spin of the $\psi(3770)$. CP conservation requires that $\eta_a \eta_b (-1)^L = 1$, hence $\eta_a/\eta_b = -1$. This means that if one D meson is reconstructed in a CP even (odd) eigenstate, the other D meson must be a CP odd (even) eigenstate. The CP eigenstates of the D meson are defined as follows. Beginning with the definition that a particle is transformed to its antiparticle by the $\hat{C}\hat{P}$ operator,

$$\hat{C}\hat{P}|D^0\rangle = |\bar{D}^0\rangle, \quad \hat{C}\hat{P}|\bar{D}^0\rangle = |D^0\rangle, \quad (1.41)$$

the CP -even (odd) eigenstates of the D meson, $D_{+(-)}$, are defined as

$$D_{\pm} = \frac{1}{\sqrt{2}}(D^0 \pm \bar{D}^0). \quad (1.42)$$

It can be shown using Eq. (1.41) that the $D_{+(-)}$ state has a CP eigenvalue of $+1$ (-1).

The measurements of c_i and s_i require that one of the D mesons decays to the $K_s^0 h^+ h^-$ final state and the other decays to some final state, X_D . If X_D is a CP even (odd) eigenstate, the D meson decaying to $K_s^0 h^+ h^-$ must be CP odd (even). The amplitude

and partial width of $D_{\pm} \rightarrow K_S^0 h^+ h^-$ at the Dalitz plot coordinate (m_-^2, m_+^2) is:

$$A(D_{\pm} \rightarrow K_S^0 h^+ h^-) = \frac{1}{\sqrt{2}} (A_D \pm \bar{A}_D), \quad (1.43)$$

$$\frac{d\Gamma(D_{\pm} \rightarrow K_S^0 h^+ h^-)}{dm_-^2 dm_+^2} = \frac{1}{2} (A_D^2 + \bar{A}_D^2) \pm A_D \bar{A}_D \cos \delta_D. \quad (1.44)$$

Using the integrals defined in Eqs. (1.34)–(1.37), the decay rate to bin i of the $D_{\pm}$ Dalitz plot is

$$\Gamma_i(D_{\pm} \rightarrow K_S^0 h^+ h^-) \propto \frac{1}{2} (T_i + T_{-i}) \pm \sqrt{T_i T_{-i}} c_i. \quad (1.45)$$

One can also consider the case where $X_D = K_S^0 h^+ h^-$. Since the $D\bar{D}$ system has odd parity, the two D mesons are produced in the antisymmetric state $\frac{1}{\sqrt{2}} (D\bar{D} - \bar{D}D)$. Using primed and unprimed symbols to indicate the coordinate positions of the two $D \rightarrow K_S^0 h^+ h^-$ Dalitz plots, the decay amplitude is

$$\begin{aligned} A_{D\bar{D}}(m_-^2, m_+^2, m_-'^2, m_+'^2) &= \frac{1}{\sqrt{2}} [\bar{A}_D(m_-^2, m_+^2) A_D(m_-'^2, m_+'^2) - \bar{A}_D(m_-'^2, m_+'^2) A_D(m_-^2, m_+^2)] \\ &= \frac{1}{\sqrt{2}} [A_D(m_+^2, m_-^2) A_D(m_-'^2, m_+'^2) - A_D(m_+'^2, m_-'^2) A_D(m_-^2, m_+^2)], \end{aligned} \quad (1.46)$$

where the amplitudes for the D^0 and $\bar{D}^0$ decays, A_D and $\bar{A}_D$, are as defined in Eq. (1.27). Then, using the integrals defined, the decay rate to bin i of one $D \rightarrow K_S^0 h^+ h^-$ Dalitz plot and bin j of the other is

$$\Gamma_{ij} \propto T_i T_{-j} + T_{-i} T_j - 2\sqrt{T_i T_{-i} T_j T_{-j}} (c_i c_j + s_i s_j). \quad (1.47)$$

The event rates in a *CP*-tagged $D_{\pm} \rightarrow K_S^0 h^+ h^-$ Dalitz plot are used to measure $\Gamma_i(D_{\pm})$ and those in a flavour-tagged $D^0 \rightarrow K_S^0 h^+ h^-$ Dalitz plot are used to measure T_i . Therefore, $D \rightarrow K_S^0 h^+ h^-$ decays are reconstructed with the companion $\bar{D}$ meson tagged in a variety of flavour eigenstates (*e.g.* $K^- \pi^+$) or *CP* eigenstates (*e.g.* *CP* even $\pi^+ \pi^-$ or *CP* odd $K_S^0 \pi^0$),

or with the $\psi(3770)$ decaying to a double $K_S^0 h^+ h^-$ final state. The values of c_i and s_i are determined using Eqs. (1.45) and (1.47).

Binning optimisation

Throughout this discussion neither the number of bins nor their shapes have been determined. The choice of binning scheme will affect the sensitivity to γ in three ways:

1. Inspecting Eq. (1.40), c_i and s_i are coupled to $x_{\pm}$ and $y_{\pm}$, hence partitioning the Dalitz plot to produce large, non-zero values of c_i and s_i will provide a better precision on γ .
2. The total number of bins will affect the uncertainties on c_i and s_i . A smaller number of bins will allow c_i and s_i to be determined with smaller statistical uncertainties.
3. The $B \rightarrow DK$ data must also be divided into bins to perform the measurement. Hence, there will be a loss in precision compared to an unbinned method.

The initial proposal in Ref. [1] was a rectangular binning similar to that in Fig. 1.7, but it was subsequently shown that better sensitivity to γ can be obtained by partitioning the Dalitz plot into regions over which the variation in $\delta_D(m_-^2, m_+^2)$ is small, which produces large values of c_i and s_i [2]. A model of the amplitude $A_D(m_-^2, m_+^2)$, determined using flavour-tagged $D^0 \rightarrow K_S^0 h^+ h^-$ decays, provides a prediction of $\delta_D(m_-^2, m_+^2)$ and can be used to determine a binning scheme. The measurements of c_i and s_i are not biased by the use of a specific amplitude model in defining the bin boundaries, which only affects this analysis to the extent that if the model gives a poor description of the underlying decay then there will be a reduction in the statistical sensitivity of the γ measurement. The optimisation procedure below summarises the results in Ref. [3].

An amplitude model for $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ developed by the BaBar collaboration [33] provides a prediction for the values of $\delta(m_-^2, m_+^2)$, and hence $\delta_D(m_-^2, m_+^2)$, where the Dalitz

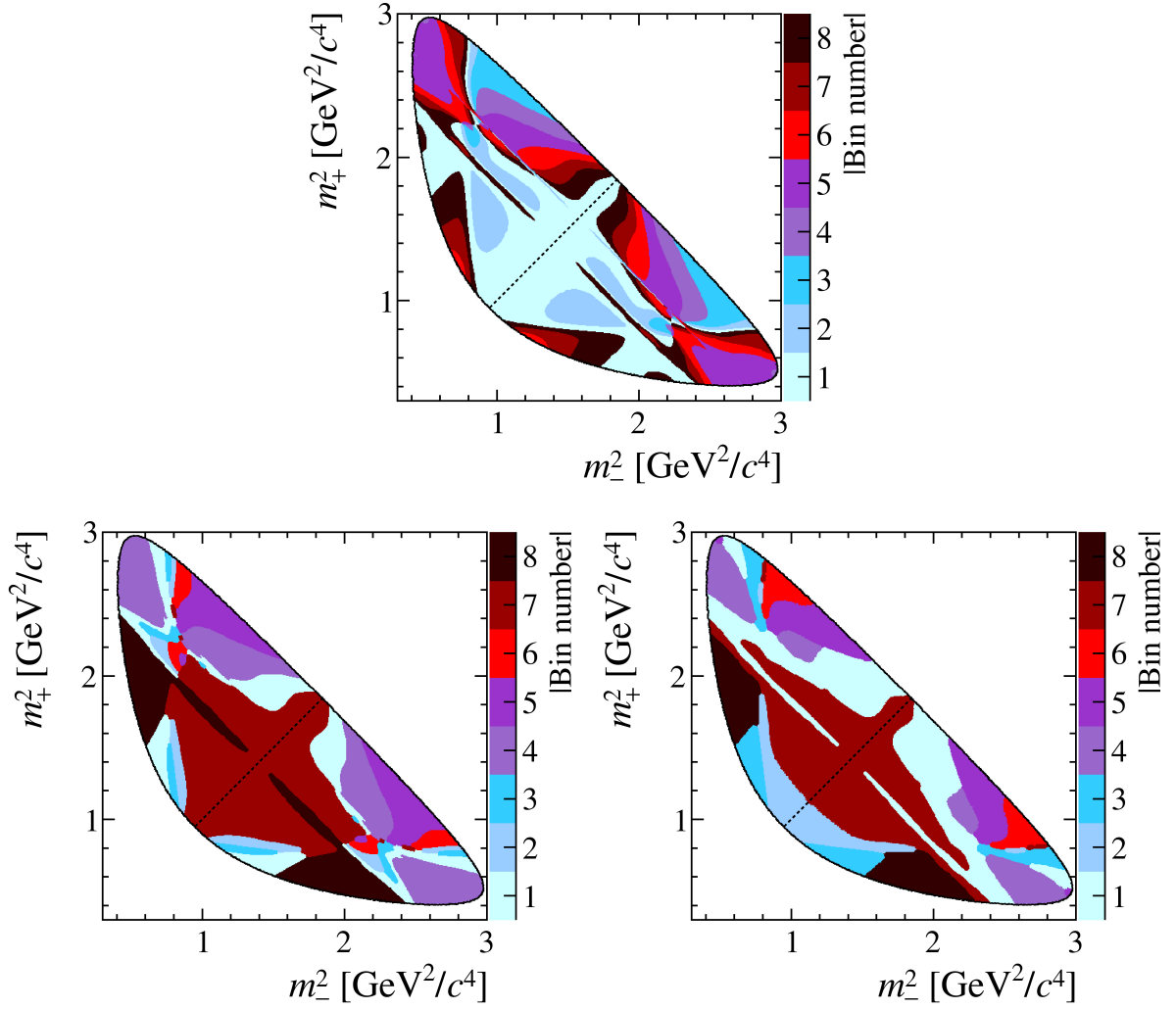


Figure 1.8. The (top) equal- δ_D , (bottom left) optimal and (bottom right) modified optimal binning schemes for $D \rightarrow K_s^0 \pi^+ \pi^-$. The dotted diagonal line separates the positive and negative bin numbers, where the negative bins are in the region $m_-^2 \leq m_+^2$.

plot has been finely segmented into pixels with width $0.0054 \text{ GeV}^2/c^4 \times 0.0054 \text{ GeV}^2/c^4$. The pixels are of similar size to the $m_\pm^2$ resolution found in simulated CLEO data. This amplitude model was used by CLEO to define and optimise the binning schemes [3]. In the so-called *equal- δ_D* binning scheme, the bin shapes are not regular but are chosen to give regions of approximately equal strong-phase difference, and the Dalitz plot region

labelled bin i contains pixels that satisfy

$$\frac{2\pi}{k} \left(i - \frac{3}{2} \right) < \delta_D(m_-^2, m_+^2) < \frac{2\pi}{k} \left(i + \frac{1}{2} \right), \quad (1.48)$$

where k is the number of bins in each half of the Dalitz plot, and bin $-i$ is defined in the symmetrically opposite region of the Dalitz plot. The equal- δ_D binnings for $K_S^0 \pi^+ \pi^-$ with $k = 8$ is shown in Fig. 1.8 (top). To determine whether the equal- δ_D binning will provide the best possible sensitivity, a metric is constructed to measure the statistical sensitivity to $x_{\pm}, y_{\pm}$ from different binning schemes relative to that of an unbinned method, or equivalently an infinite number of bins where $\cos \delta_D$ and $\sin \delta_D$ are measured at each point. The metric, $0 < Q < 1$, is defined as [3]

$$Q^2 = \frac{\sum_i \left[\left(\frac{1}{\sqrt{\Gamma_i}} \frac{d\Gamma_i}{dx_-} \right)^2 + \left(\frac{1}{\sqrt{\Gamma_i}} \frac{d\Gamma_i}{dy_-} \right)^2 \right]}{\int dm_-^2 dm_+^2 \left[\left(\frac{1}{\sqrt{|A_{B^-}|^2}} \frac{d|A_{B^-}|^2}{dx_-} \right)^2 + \left(\frac{1}{\sqrt{|A_{B^-}|^2}} \frac{d|A_{B^-}|^2}{dy_-} \right)^2 \right]}, \quad (1.49)$$

and is the ratio in the number of standard deviations by which the event rate Γ_i in bin i changes when the parameters x_- and y_- are varied, to the number of standard deviations by which the event rate, $|A_{B^-}|^2 \equiv |A(B^- \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^-)|^2$ (Eq. (1.31)), changes in the unbinned case. The metric is measured in terms of $x_{\pm}, y_{\pm}$ instead of γ because their sensitivities have weaker dependence on their central values. A low value of Q indicates a diluted sensitivity to γ and can occur in a scenario where a change in x or y induces local variations in the Dalitz plot density that average to only a small change in event rates in a certain bin.

For the equal- δ_D binning, $Q = 0.786$, indicating an approximate 20% loss in statistical precision compared to an unbinned method. Using the equal- δ_D as a starting point, the value of Q can be increased using an iterative optimisation algorithm to find the binning that produces the largest value of Q [2]. A smoothing procedure is also applied to remove regions of bins that are only a few pixels in size. The resulting *optimal*

$K_S^0 \pi^+ \pi^-$ binning with a maximal value of $Q = 0.887$ is shown in Fig. 1.8 (bottom left). This is a 13% relative improvement and the binning is significantly different in shape to the equal- δ_D binning. It is noted that an updated amplitude model for $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ is reported in Ref. [34] but using this does not lead to significantly better Q values.

An analogous metric Q' , which takes backgrounds into account, is used to optimise the binning in the presence of background that is distributed differently to the signal. Since the method involves measuring differences in yields between bins, the background, which is distributed differently to signal, will dilute these differences. The signal purity is assumed to be 41% and the background is assumed to be composed of 75% combinatorial background (uniform distribution) and a 25% equal mixture of D^0 and $\bar{D}^0$ resonant background. In this scenario, the equal- δ_D binning has $Q' = 0.882$ and the optimal binning has $Q' = 0.867$. The resulting *modified optimal* binning is shown in Fig. 1.8 (bottom right) and has a value $Q' = 0.904$. An additional advantage of the modified optimal binning is that it is more robust for analyses with low B yields in comparison to the alternatives, as no individual bin is very small.

An amplitude model for $D^0 \rightarrow K_S^0 K^+ K^-$ has been developed by BaBar [34], where the pixel width is $0.0018 \text{ GeV}^2/c^4 \times 0.0018 \text{ GeV}^2/c^4$. The equal- δ_D binning schemes for $K_S^0 K^+ K^-$, with $k = 2, 3$ and 4 , are shown in Fig. 1.9. The value $Q = 0.771$ for $k = 2$ is reported and the optimisation methods previously discussed did not produce any binning with a significantly higher Q value.

The measured CLEO values of c_i and s_i for $K_S^0 \pi^+ \pi^-$ are shown in Fig. 1.10 and are superimposed with predictions calculated from the BaBar model used to determine the binning. The level of agreement between the measured values and model expectation offers a *post hoc* justification that the BaBar model provides a reasonable guide for the binning optimisation. Some of the measured points lie outside the physical boundary $c_i^2 + s_i^2 = 1$ and using these central values will lead to a small bias in γ that is determined as part of the analysis. The uncertainties on s_i are larger than the ones on c_i and

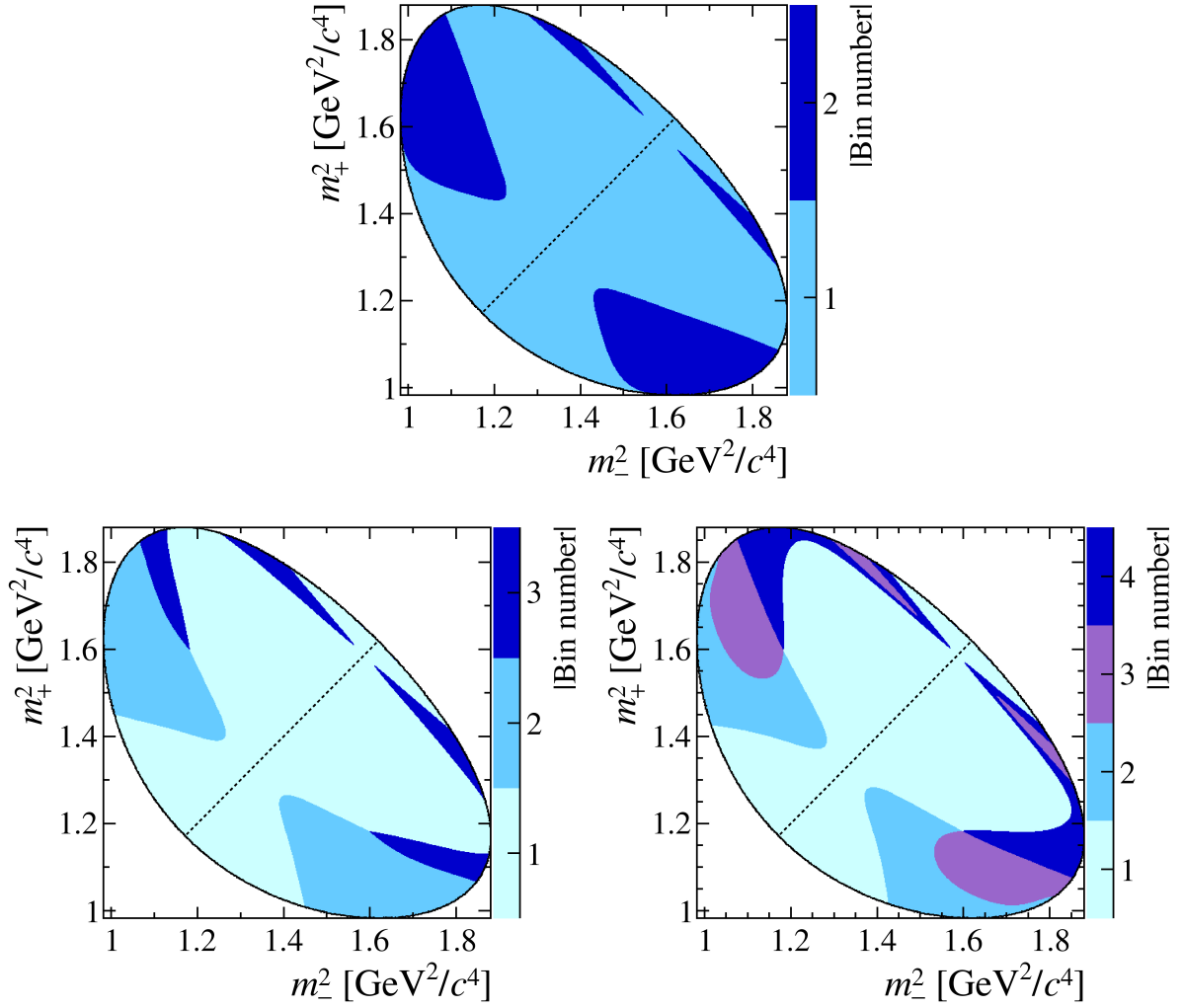


Figure 1.9. Binning of $D \rightarrow K_s^0 K^+ K^-$ Dalitz plot using the equal- δ_D scheme with (top, bottom left, bottom right) $k = 2, 3$ and 4 . The dotted diagonal line separates the positive and negative bin numbers, where the negative bins are in the region $m_-^2 \leq m_+^2$.

therefore the CLEO-related uncertainties in the γ measurement are expected to be larger in $y_{\pm}$ than $x_{\pm}$. Similar results are achieved with $K_s^0 K^+ K^-$.

1.6.4 Measurement with $B^0 \rightarrow DK^{*0}$ decays

Sensitivity to γ can also be obtained from $B^0 \rightarrow DK^{*0}$ decays, where K^{*0} refers to the $K^*(892)^0$; the charge of the kaon from the $K^*(892)^0 \rightarrow K^+ \pi^-$ decay unambiguously

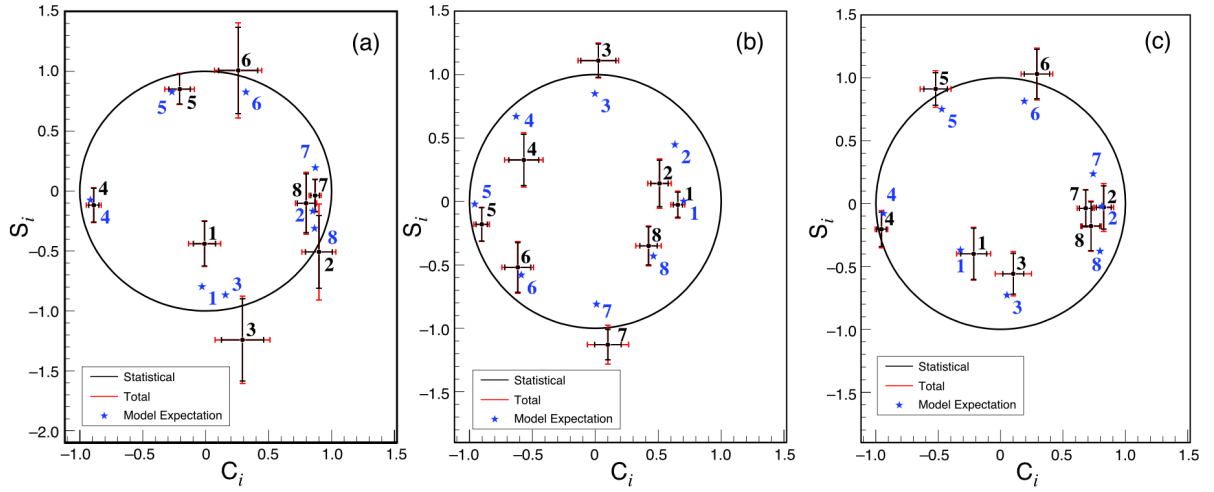


Figure 1.10. Measured values of c_i and s_i for $D \rightarrow K_S^0 \pi^+ \pi^-$ in the (left to right) optimal, equal- δ_D and modified optimal binnings. The stars show the model expectation calculated from Ref. [33]. The circle drawn indicates the physical region $c_i^2 + s_i^2 \leq 1$. Reproduced from [3]. A different phase convention used in [3] means that the values of s_i in this figure have the opposite sign to those used in this thesis.

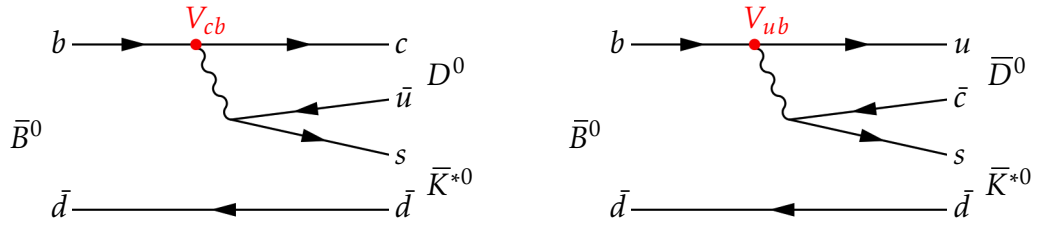


Figure 1.11. Feynman diagrams for the (left) $\bar{B}^0 \rightarrow D^0 \bar{K}^{*0}$ and (right) $\bar{B}^0 \rightarrow \bar{D}^0 \bar{K}^{*0}$ amplitudes.

identifies the flavour of the decaying B meson as B^0 or $\bar{B}^0$. The Feynman diagrams for the contributing amplitudes are shown in Fig. 1.11. Although the branching fraction of the $B^0 \rightarrow DK^{*0}$ decay is an order of magnitude smaller than that of the $B^- \rightarrow DK^-$ decay, it is expected to exhibit larger CP -violating effects since the two interfering amplitudes are both colour-suppressed and therefore comparable in magnitude. An estimate of the magnitude ratio between the suppressed and favoured amplitudes using the CKM factors gives $r_{B^0} \sim |V_{cs} V_{ub}^*| / |V_{us} V_{cb}^*| \sim 0.4$ [24].

The favoured and suppressed $\bar{B}^0$ decay amplitudes can be expressed as [35]

$$A(\bar{B}^0 \rightarrow D^0 \bar{K}_s^0; p) \equiv A_c(p) e^{i\delta_c(p)}, \quad A(\bar{B}^0 \rightarrow \bar{D}^0 \bar{K}_s^0; p) \equiv A_u(p) e^{i[\delta_u(p) - \gamma]}, \quad (1.50)$$

where p is the $(m^2(K\pi), m^2(D\pi))$ coordinate on the $\bar{B}^0 \rightarrow DK^-\pi^+$ Dalitz plot, $A_u(p)$ and $A_c(p)$ are the moduli of the $b \rightarrow u$ and $b \rightarrow c$ amplitudes, and $\delta_{c,u}(p)$ represent the strong phases of the relevant decay amplitudes. The symbol X_s^0 refers to a resonant or nonresonant $K^-\pi^+$ pair, which could be produced by the decay of the $\bar{K}^{*0}$ meson or by other contributions to the $\bar{B}^0 \rightarrow DK^-\pi^+$ final state. The expressions for the B^0 decay are

$$A(B^0 \rightarrow \bar{D}^0 X_s^0; p) \equiv A_c(p) e^{i\delta_c(p)}, \quad A(B^0 \rightarrow D^0 X_s^0; p) \equiv A_u(p) e^{i[\delta_u(p)+\gamma]}. \quad (1.51)$$

The natural width of the K^{*0} (approximately $50 \text{ MeV}/c^2$ [4]) must be considered when analysing these decays. In the region near the K^{*0} mass there is interference between the signal K^{*0} decay amplitude and amplitudes due to the other $B^0 \rightarrow DK^+\pi^-$ Dalitz plot contributions, such as higher mass $K\pi$ resonances and nonresonant $K\pi$ decays. The dilution in sensitivity to γ due to these interfering contributions is quantified by a coherence factor κ , where $0 \leq \kappa \leq 1$ [35]. Hence, r_{B^0} , the coherence factor κ , and the effective strong phase difference δ_{B^0} depend on the region of the B^0 Dalitz plot to be analysed. These are defined as

$$r_{B^0}^2 \equiv \frac{|A(B^0 \rightarrow D^0 K^{*0})|^2}{|A(B^0 \rightarrow \bar{D}^0 K^{*0})|^2} = \frac{\int_{K^{*0}} dp A_u^2(p)}{\int_{K^{*0}} dp A_c^2(p)}, \quad (1.52)$$

$$\kappa e^{i\delta_{B^0}} \equiv \frac{\int_{K^{*0}} dp A_c(p) A_u(p) e^{i[\delta_u(p)-\delta_c(p)]}}{\sqrt{\int_{K^{*0}} dp A_c^2(p)} \sqrt{\int_{K^{*0}} dp A_u^2(p)}}. \quad (1.53)$$

In order to maximise sensitivity to γ , an integration region $\int_{K^{*0}} dp$ which produces a large value of κ should be chosen. The region chosen in this thesis is defined later in Sect. 6.1.2. The region is based on a $K\pi$ invariant mass and a K^{*0} helicity angle window. The helicity angle is denoted by θ^* and is defined as the angle between the K^{*0} daughter kaon momentum vector and the direction opposite to the B^0 momentum vector in the K^{*0} rest frame.

Using $A_D(m_-^2, m_+^2)$ to describe the $D^0 \rightarrow K_S^0 h^+ h^-$ amplitude as before, the partial width for the $\bar{B}^0$ decays can be written as

$$\begin{aligned} \frac{d\Gamma(\bar{B}^0 \rightarrow D(\rightarrow K_S^0 h^+ h^-) \bar{K}_S^0)}{dp dm_-^2 dm_+^2} &\propto |A_c(p) e^{i\delta_c(p)} A_D(m_-^2, m_+^2) \\ &\quad + A_u(p) e^{i[\delta_u(p) - \gamma]} \bar{A}_D(m_-^2, m_+^2)|^2. \end{aligned} \quad (1.54)$$

Expanding and integrating over the defined K^{*0} region, one obtains

$$\begin{aligned} \frac{d\Gamma(\bar{B}^0 \rightarrow D(\rightarrow K_S^0 h^+ h^-) \bar{K}^{*0})}{dm_-^2 dm_+^2} &\propto |A_D(m_-^2, m_+^2)|^2 + r_{B^0}^2 |A_D(m_+^2, m_-^2)|^2 \\ &\quad + 2\kappa r_{B^0} \text{Re}[A_D(m_-^2, m_+^2) A_D^*(m_+^2, m_-^2) e^{-i(\delta_{B^0} - \gamma)}] \\ &\propto a^2 + r_{B^0}^2 \bar{a}^2 + 2\kappa r_{B^0} a \bar{a} \cos[(\delta_{B^0} - \gamma) - \delta_D]. \end{aligned} \quad (1.55)$$

In the last line, the coordinates (m_-^2, m_+^2) of a , $\bar{a}$ and δ_D are not explicitly written. The form of the $B^0 \rightarrow DK^{*0}$ partial width is identical to that for $B \rightarrow DK$ Eq. (1.31) with the addition of κ describing the dilution of the interference term. Integrating over the region defined by Dalitz plot bin i , the rate of $B^0 \rightarrow DK^{*0}$ decays in the bin are

$$\Gamma_{+i}(\bar{B}^0 \rightarrow D(\rightarrow K_S^0 h^+ h^-) \bar{K}^{*0}) \propto T_{+i} + (x_-^{*2} + y_-^{*2}) T_{-i} + 2\kappa \sqrt{T_{+i} T_{-i}} (x_-^* c_{+i} + y_-^* s_{+i}), \quad (1.56a)$$

$$\Gamma_{-i}(\bar{B}^0 \rightarrow D(\rightarrow K_S^0 h^+ h^-) \bar{K}^{*0}) \propto T_{-i} + (x_-^{*2} + y_-^{*2}) T_{+i} + 2\kappa \sqrt{T_{+i} T_{-i}} (x_-^* c_{-i} + y_-^* s_{-i}), \quad (1.56b)$$

$$\Gamma_{+i}(B^0 \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^{*0}) \propto T_{-i} + (x_+^{*2} + y_+^{*2}) T_{+i} + 2\kappa \sqrt{T_{+i} T_{-i}} (x_+^* c_{+i} - y_+^* s_{+i}), \quad (1.56c)$$

$$\Gamma_{-i}(B^0 \rightarrow D(\rightarrow K_S^0 h^+ h^-) K^{*0}) \propto T_{+i} + (x_+^{*2} + y_+^{*2}) T_{-i} + 2\kappa \sqrt{T_{+i} T_{-i}} (x_+^* c_{-i} - y_+^* s_{-i}), \quad (1.56d)$$

where the parameters T_i , c_i , s_i are as defined in Eqs. (1.34)–(1.37) and

$$x_{\pm}^* \equiv r_{B^0} \cos(\delta_{B^0} \pm \gamma), \quad (1.57)$$

$$y_{\pm}^* \equiv r_{B^0} \sin(\delta_{B^0} \pm \gamma). \quad (1.58)$$

1.7 Previous γ measurements using the GGSZ method

The measurement of γ using the GGSZ modes was first proposed in 2003 [1]. In addition to the method outlined above, an alternative *model-dependent* (MD) method which uses a D^0 amplitude model directly in Eqs. (1.31) and (1.33) was discussed. In the absence of measurements of c_i and s_i , the first GGSZ analyses by the B-factories, using charged B decays, employed the MD method, with BaBar and Belle each developing their own $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ amplitude model [36, 37]. In addition to the $B \rightarrow DK$ mode, the $B^- \rightarrow D^*(\rightarrow \{D\pi^0, D\gamma\})K^-$ and $B^- \rightarrow DK^{*-}(\rightarrow K_S^0 \pi^+)$ modes were also studied, since the same D^0 amplitude model and value of γ applies to all these decays. These measurements were later updated with larger datasets, and BaBar also included the $D \rightarrow K_S^0 K^+ K^-$ final state [33, 34, 38].

The $D^0 \rightarrow K_S^0 h^+ h^-$ models were instrumental in guiding the binnings in which c_i and s_i were measured by the CLEO experiment [3, 39]. Following the availability of c_i and s_i values in 2009–10, the first measurement of γ using the technique described in the previous sections, hereinafter known as the *model-independent* (MI) method, was performed by the Belle collaboration in $B \rightarrow DK$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decays [40] using the optimal binning scheme. An MI measurement using both the $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ final states was then performed by the LHCb collaboration [41]. This analysis also used the CLEO values of c_i and s_i and was performed using a subset of the data analysed in this thesis. The optimal binning scheme was used in $K_S^0 \pi^+ \pi^-$ and the $k = 2$ binning was used for $K_S^0 K^+ K^-$.

The first GGSZ analysis with neutral B decays was performed by BaBar [42] using the MD method to analyse $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decays. The 2008 BaBar amplitude model [33] was again used. At the time of starting this work, no MI measurement had been performed with neutral B decays. Since then, the first MI analysis of $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decays has been performed by Belle [43]. The LHCb collaboration has

also since performed two MD measurements of the $K_S^0 \pi^+ \pi^-$ final state with the 2010 BaBar amplitude model [34]; the first is a study of $B \rightarrow DK$ decays, using 1 fb^{-1} of data at $\sqrt{s} = 7 \text{ TeV}$ [44], and the second is a study of $B^0 \rightarrow DK^{*0}$ decays using the 3 fb^{-1} dataset collected in Run I of the LHC [45].

In the MD method, an amplitude model that predicts an incorrect δ_D distribution can bias the measurement of γ . Hence, these measurements have a model-related systematic uncertainty, which is evaluated by repeating the measurement using many alternative models. The BaBar and Belle models were developed at similar times using the most recent developments for amplitude models available; the model uncertainty is estimated to be 3° in the BaBar analysis [34] and 9° in the Belle analysis [38]. This large difference reflects the degree of subjectivity associated with the choice of alternative models. There is now a compelling reason for using the MI method over the MD method since CLEO measurements provide direct access to the δ_D distribution and the uncertainty related to model assumptions of the strong-phase difference distribution is removed. Since the ultimate goal is to determine γ as a SM benchmark, it is critical that model-related uncertainties are eliminated wherever possible. The results from the previous MD and MI measurements are compared together with the results obtained in this thesis in Sect. 7.1.3.

1.8 Analysis overview

In this thesis, the CKM angle γ and the hadronic parameters r_B , δ_B , r_{B^0} and δ_{B^0} are extracted using $B \rightarrow DK^{(*)}$ decays via the $D \rightarrow K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ final states. These measurements are performed with proton-proton collision data at $\sqrt{s} = 7 \text{ TeV}$ and 8 TeV , corresponding to an integrated luminosity of 3 fb^{-1} , collected by the LHCb experiment. The first measurement, with $B \rightarrow DK$ decays and published in [46],

supersedes the measurement in Ref. [41]. The second measurement, with $B^0 \rightarrow DK^{*0}$ decays, is published in Ref. [47].

In Chapter 3, data samples enriched in the $B \rightarrow DK$ signal are prepared using a multivariate selection. The optimisation of the selection and many specific background studies are performed using the $B \rightarrow D\pi$, $D \rightarrow K_S^0 h^+ h^-$ control channel because of the higher branching ratio and kinematic similarities shared with the signal channel. The B mass distribution of selected candidates is parameterised in preparation for the Dalitz plot fit to measure the physics parameters. Chapter 4 describes the use of $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ decays to determine the values of T_i , which are required to measure $x_\pm, y_\pm$ using Eq. (1.40). In this semileptonic decay, the flavour of the D^0 can be tagged because the strong decay of the D^{*-} ensures that the produced $\bar{D}^0$ meson flavour is determined by the charge of the associated pion. Here the $\bar{D}^0$ decays to either the $K_S^0 \pi^+ \pi^-$ or $K_S^0 K^+ K^-$ final state. The symbol X , henceforth omitted, indicates other particles which may be produced in the decay but are not reconstructed. The value of T_i is related to the fraction of $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates in bin i of the Dalitz plot. However, in any experimental measurement, the non-uniform selection (efficiency) of events in different parts of the Dalitz plot must be taken into account. Hence, the efficiency differences between the Dalitz plots of the hadronic and semileptonic modes are studied and corrected for.

In Chapter 5, the $B \rightarrow DK$ candidates are split into their respective Dalitz plot bins, where the optimal binning scheme is used for $K_S^0 \pi^+ \pi^-$ and the $k = 2$ binning scheme, also known as the 2×2 binning, is used for $K_S^0 K^+ K^-$. The yields in each bin are related to $x_\pm, y_\pm$ using Eq. (1.40). Instead of fitting for the physics parameters (r_B , δ_B and γ) directly, it is the values of $x_\pm$ and $y_\pm$ that are initially determined from the data. As defined in Eqs. (1.38) and (1.39), they are $x_\pm = r_B \cos(\delta_B \pm \gamma)$ and $y_\pm = r_B \sin(\delta_B \pm \gamma)$ and are better-behaved fit parameters in the limit that $(\delta_B \pm \gamma) = 0^\circ$ or 180° . The values of c_i and s_i for $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ are those obtained from the CLEO measurements [3]. The

choice of binning schemes from the options outlined in Sect. 1.6.3 is discussed in further detail in the relevant analysis chapters. Using the B mass parameterisation determined in Chapter 3, a simultaneous fit of the mass distribution in each bin determines the best fit values of $x_{\pm}, y_{\pm}$.

The $B^0 \rightarrow DK^{*0}$ decays are analysed in Chapter 6, where a similar procedure as that described for $B \rightarrow DK$ is used to measure $x_{\pm}^*$ and $y_{\pm}^*$. The modified optimal binning scheme is used for $K_S^0 \pi^+ \pi^-$ and the 2×2 scheme is used for $K_S^0 K^+ K^-$. The $B \rightarrow DK$ analysis was performed and published before work on the $B^0 \rightarrow DK^{*0}$ analysis began, hence areas where the two analyses differ reflect improvements that were only learnt after the publication of the $B \rightarrow DK$ analysis. The values of $r_B, \delta_B, r_{B^0}, \delta_{B^0}$ and γ are then extracted from the measured $x_{\pm}, y_{\pm}, x_{\pm}^*$ and $y_{\pm}^*$ in Chapter 7, where the results are presented together with those of other γ measurements made at LHCb. The conclusions are summarised in Chapter 8.

CHAPTER 2

The LHCb detector

The Large Hadron Collider (LHC) [48] is the world's largest and most powerful particle accelerator. It is situated at CERN (European Organisation for Nuclear Research) in an underground tunnel 27 km in circumference. The LHC circulates beams of protons or heavy ions and has achieved record collision energies of $\sqrt{s} = 13 \text{ TeV}$ in 2015. Hydrogen atoms are first ionised to form protons, which are accelerated in multiple stages in the CERN accelerator complex. A schematic diagram is shown in Fig. 2.1. Moving from the Proton Synchrotron to the Super Proton Synchrotron, the protons are accelerated to 26 GeV and 450 GeV, respectively, when they are injected into the LHC to be accelerated to the required collision energies. The proton beams are circulated in opposite directions and are crossed and focused for collision at four interaction points where the four major LHC experiments, including the Large Hadron Collider beauty (LHCb) experiment, are located. Situated at the other three interaction points are the ATLAS, CMS and ALICE experiments.

The first major data-taking period at the LHC, known as Run I, took place between 2010 and 2012. When proton beams were circulated, each proton bunch contained about 1.6×10^{11} protons, colliding at a frequency of 20 MHz. The instantaneous luminosity delivered at the ATLAS and CMS interaction points was approximately $5 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ [49], which decreased over the course of a fill as the number of circulat-

ing protons declined. A much lower instantaneous luminosity was delivered to the LHCb detector by providing a smaller number of pp interactions per bunch crossing. This choice is essential for the efficient reconstruction of B decays, which have a characteristic displaced secondary vertex from the primary vertex (PV) of $\mathcal{O}(10\text{mm})$ at the LHC. The lower luminosity also reduced the amount of radiation damage to the detector. The luminosity control is achieved in LHCb without reducing the luminosity delivered to the other experiments by colliding the beams at a transverse offset to reduce the overlapping cross-sectional area. The luminosity is then kept constant throughout each fill by gradually bringing the beams closer together. The levelled luminosity makes the data easier to analyse by allowing the same trigger configuration (Sect. 2.6) to be used in a fill and keeping the detector occupancy stable. The majority of data at LHCb was collected at a luminosity of $4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ [50]. In such conditions, there were typically 1.8 pp interactions per bunch crossing compared to an average of 20 in ATLAS at $\sqrt{s} = 8 \text{ TeV}$.

This thesis uses the full Run I dataset for pp collisions recorded by the LHCb experiment and corresponds to an integrated luminosity of 1 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$ and 2 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$, recorded in 2011 and 2012, respectively.

The LHCb detector is designed to study the decays of beauty and charm hadrons. At the LHC, the dominant production mechanism of b quarks is gluon fusion. Since it is stochastically favourable for the gluons to interact with asymmetric momenta in the pp centre-of-mass frame, the $b\bar{b}$ pair is boosted along the beamline [52] and both quarks tend to be produced in the same forward or backward cone, as seen in Fig. 2.2. Therefore, the LHCb detector is a single forward arm spectrometer covering the pseudorapidity range $2 < \eta < 5$, where 25% of produced $b\bar{b}$ quark pairs enter the detector acceptance. In this geometrical region, the b hadron has on average higher momentum, producing a larger displacement between the PV and decay vertex, which facilitates the identification of the b hadron. The detector is described with respect to a

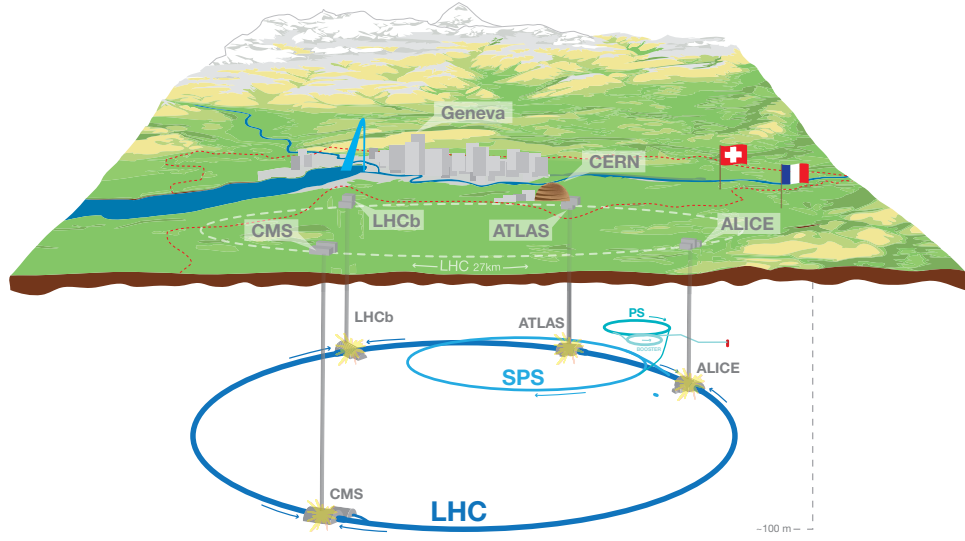


Figure 2.1. Schematic diagram of the LHC showing the approximate locations of the four major experiments: LHCb, ATLAS, CMS and ALICE. The booster rings, Super Proton Synchrotron (SPS) and Proton Synchrotron (PS), are also indicated. Reproduced from [51].

right-handed coordinate system, with the z -axis pointing along the beam, and x and y pointing in the horizontal and vertical directions. The detector is instrumented to a depth of $z = +20$ m, where $z = 0$ defines the primary interaction point, and is composed of many subdetectors, arranged as shown in Fig. 2.3. These are the VERtex LOcator (VELO), the Ring Imaging CHerenkov detectors (RICH1 and RICH2), the Tracker Turicensis (TT), the dipole magnet, the tracking stations (T1–T3), the calorimetry systems (SPD/PS, ECAL, HCAL), and the muon stations (M1–M5). The subdetectors are now discussed in turn, summarising the details described in Ref. [50, 52].

2.1 The VELO

The VERtex LOcator (VELO) is responsible for reconstructing primary and displaced secondary vertices, which are crucial for the identification of b -hadron decays. The reconstruction is achieved through precision tracking close to the interaction point. As

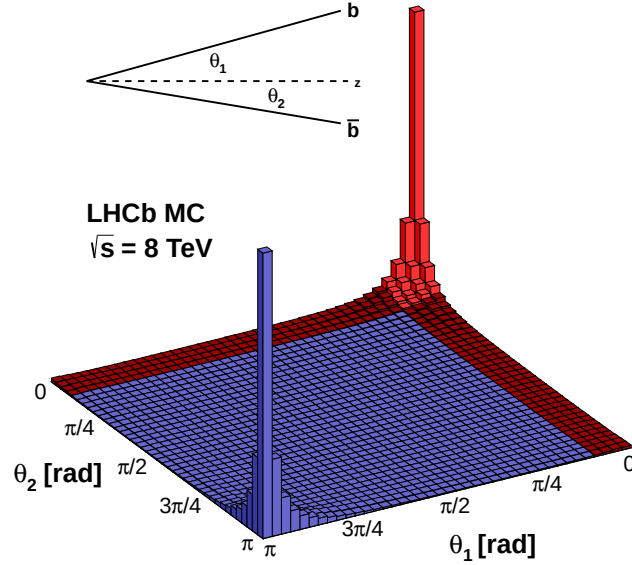


Figure 2.2. The cross-section of $b\bar{b}$ quark pair production in simulated pp collisions at $\sqrt{s} = 8 \text{ TeV}$ as a function of the polar angles, θ_1 and θ_2 , that each quark travels relative to the beam axis, labelled z . The red regions indicate the region instrumented by the LHCb detector. Reproduced from [53].

shown in Fig. 2.4, the VELO consists of 21 semicircular silicon modules placed along the z direction surrounding the pp interaction region. The modules contain single-sided silicon strip sensors with a thickness of $300 \mu\text{m}$ in either a concentric pattern to record the r position or a radial pattern to record the ϕ position of particle tracks, where r is the radial distance from the beam and ϕ is the azimuthal angle. The semicircular halves are retracted during beam injection and are only brought together once stable beams are achieved in order to prevent radiation damage. High precision of the vertex positions is attained by bringing the sensors to a distance of just 8 mm from the beam. The sensors are enclosed in a vacuum and separated from the beam vacuum by a corrugated aluminium sheet that shields the VELO modules from electromagnetic interference induced by the proton beams. The silicon sensors are kept at or below -5°C using liquid CO_2 to limit damage from the high radiation environment.

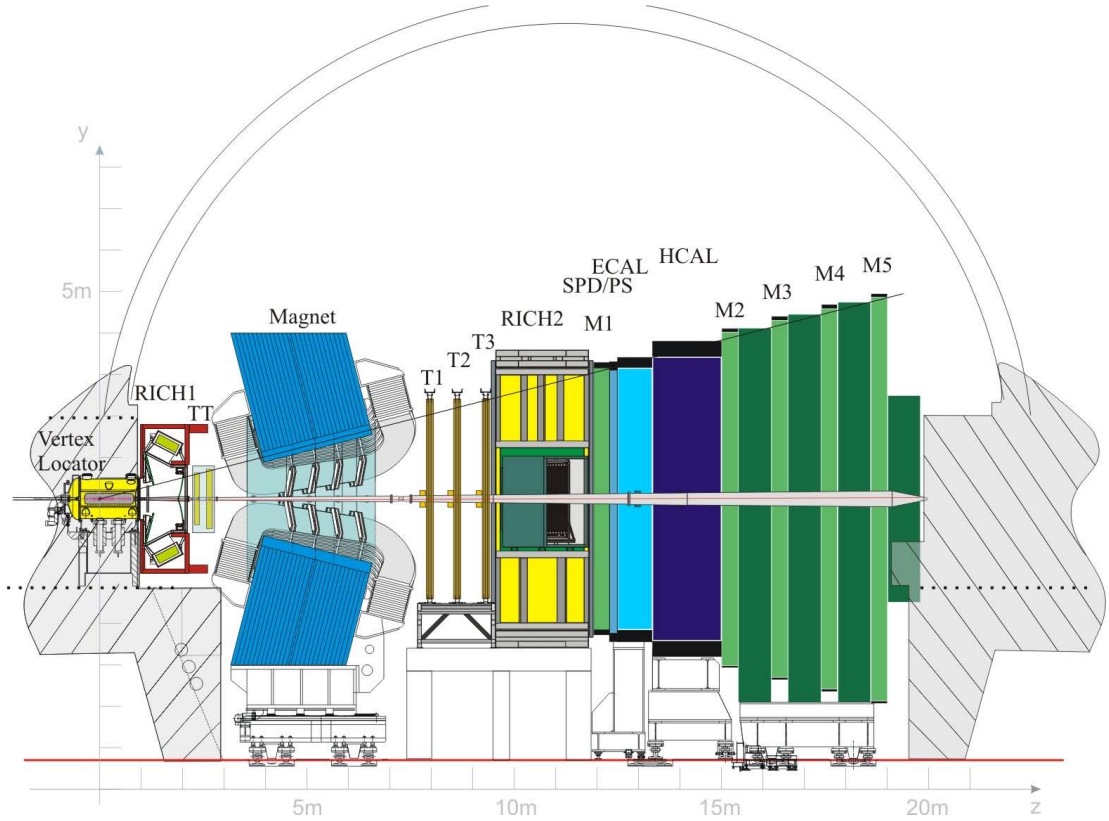


Figure 2.3. Schematic diagram of the LHCb detector with the subdetectors labelled and described in the main text. Reproduced from [52].

The performance of the VELO can be quantified by the resolution of the impact parameter (IP), which is defined as the distance of closest approach between a track and the PV associated with it. The IP resolution is $20\text{ }\mu\text{m}$ for a track with momentum $50\text{ GeV}/c$ or greater. The excellent resolution allows the decay length of most B mesons to be measured with an uncertainty of about 0.25 mm [54] compared to a typical flight distance of $\mathcal{O}(10)\text{ mm}$. The IP resolution in the x and y directions as a function of track momentum is shown in Fig. 2.5.

2.2 Tracking stations and magnet

Downstream of the VELO are the Tracker Turicensis (TT), a dipole magnet, followed by three tracking stations (T1–T3). The T1–T3 stations each consist of a cross-shaped

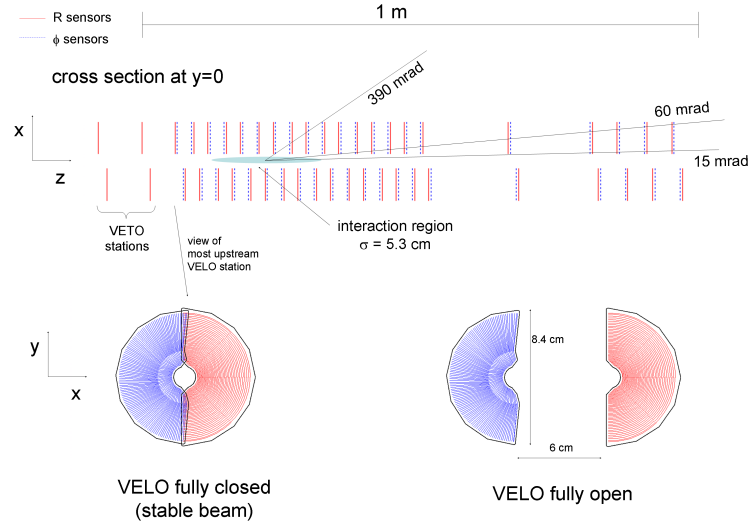


Figure 2.4. Plan of the VELO subdetector showing the positions of (top) the 21 silicon modules, and the arrangement of the semicircular halves in the (bottom left) closed and (bottom right) open positions. Reproduced from [52].

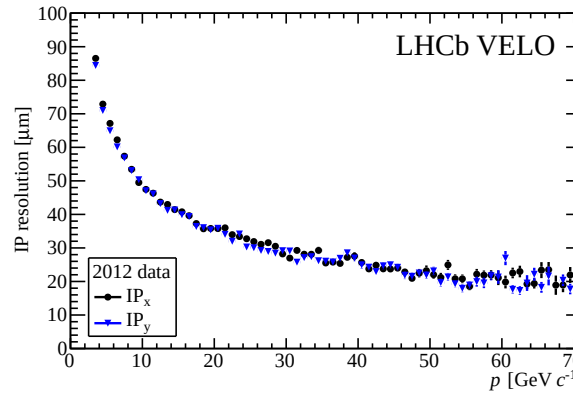


Figure 2.5. Impact parameter (IP) resolution of the VELO in 2012 conditions as a function of momentum in the (black) x and (blue) y directions. Reproduced from [54].

Inner Tracker (IT) and a straw tube Outer Tracker (OT). The TT's role is to provide space points for tracks about to enter the magnet and to allow the reconstruction of particle tracks with momentum below $1.5 \text{ GeV}/c$ whose trajectories are bent out of the detector acceptance after passing through the magnet.

Silicon microstrips with a pitch of $200 \mu\text{m}$ are used in both the TT and IT stations which provide a spatial resolution of $50 \mu\text{m}$ [55]. The layouts of one of the TT and one

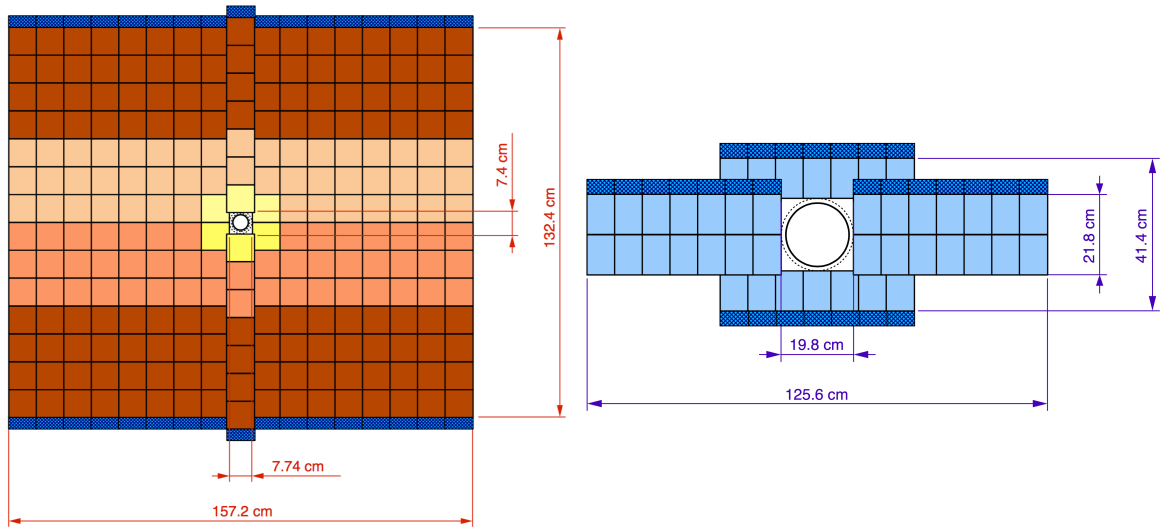


Figure 2.6. Layout of (left) one of the Tracker Turicensis (TT) layers and (right) one of the Inner Tracker (IT) in the x - y plane. Reproduced from [55].

of the IT layers are shown in Fig. 2.6. The TT is 150cm wide and 130cm high, and contains microstrips which are arranged into four layers along the beam axis, with the central two layers arranged at a -5° and $+5^\circ$ angle relative to the outer layers, covering an active area of 8.4m^2 . The IT is approximately 130cm wide and 40cm high, and occupies the region of T1–T3 closest to the beam pipe where the particle flux is highest.

Surrounding the IT subdetector is the OT, which is composed of an array of straw tubes with diameter 4.9mm, and circulated with a gas mixture of 70% argon, 28.5% carbon dioxide and 1.5% oxygen [56]. The drift-time of ionisation electrons, liberated in the gas by a traversing charged particle to a wire at the centre of each tube, is used to determine the particle trajectories. The OT design is driven by the momentum resolution required to achieve a B mass resolution of $20\text{MeV}/c^2$ for a fully charged hadronic final state.

In addition to the tracking instrumentation, momentum measurements rely on the dipole magnet, which curves charged tracks as they pass through the magnetic field. The magnet consists of a 1600 tonne iron yoke with water-cooled aluminium coils and provides a field integral of 4Tm for tracks 10m in length, which is the distance

from the VELO to the front of RICH2 (see Fig. 2.3). The variation in field strength along the z -axis is shown in Fig. 2.7. Beyond the volume of the magnetic yolk, the field strength is strongest around the tracking stations and is attenuated inside the VELO and RICH detectors which are surrounded by magnetic shielding. The low field strength in the VELO ensures that the tracks are straight to allow ease of triggering and precise vertexing, while the shielding around the RICH detectors is required to minimise image distortion in the photon detectors. The magnetic field is oriented along the vertical y -axis and bends particles in the x - z plane. Since positively and negatively charged particles will bend in opposite directions, a charge detection asymmetry can result if the left and right halves of the detector have different tracking efficiencies. This would affect CP violation studies which involve the measurements of charge asymmetries. Hence, to minimise systematics, the magnetic field direction is reversed regularly when taking data.

The excellent performance of the tracking system is demonstrated by the high tracking efficiency and good resolution of momentum measurements. These parameters are studied using $J/\psi \rightarrow \mu^+ \mu^-$ decays. The tracking efficiency is defined as the probability that a charged particle that has passed through the full tracking system is reconstructed [50]. (The track reconstruction algorithms are described in Sect. 2.7.1.) The efficiency is shown in Fig. 2.8 (left) as a function of the number of tracks in the event, N_{track} , and an efficiency of greater than 95% is achieved even at the highest track multiplicities. The tracking efficiency is also studied as a function of other kinematic and topological quantities, and is greater than 96% in the range $5 < p < 200 \text{ GeV}/c$, $2 < \eta < 5$ and $1 < N_{\text{PV}} < 5$, where N_{PV} is the number of reconstructed PVs. The momentum resolution is measured to be $\delta p/p \approx 0.5\%$ for tracks with a momentum of $20 \text{ GeV}/c$, increasing to $\delta p/p \approx 1\%$ for tracks with a momentum of $150 \text{ GeV}/c$.

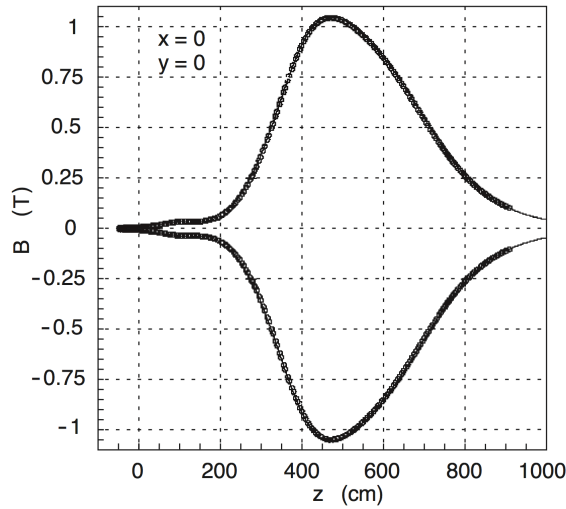


Figure 2.7. The variation in magnetic field strength, B , along the z -axis. The RICH1 and RICH2 detectors occupy the regions $99 < z < 217$ cm and $950 < z < 1183$ cm, respectively. Reproduced from [52].

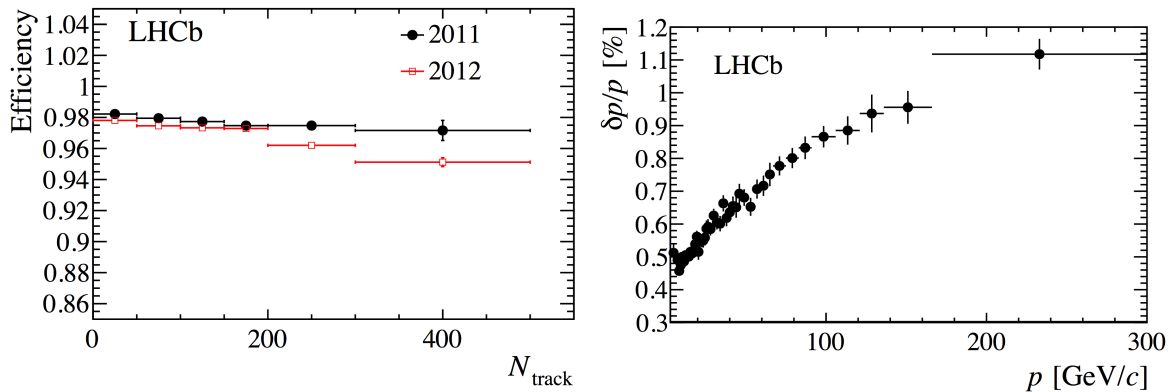


Figure 2.8. The performance of the tracking system measured using $J/\psi \rightarrow \mu^+ \mu^-$ decays: (left) track efficiency as a function of the number of tracks in the event, N_{track} , and (right) momentum resolution as a function of track momentum. Reproduced from [50].

2.3 The RICH detectors

Two Ring Imaging Cherenkov (RICH) detectors provide LHCb's particle identification (PID), and are paramount for distinguishing the various B and D final states studied in this thesis that contain combinations of kaons, pions and protons. This is vital for measurements of the CKM angle γ and the LHCb physics programme as a whole. The RICH detectors rely on the principle of Cherenkov radiation, in which a charged

particle passing through a dielectric material with refractive index n at a velocity v will emit a cone of light. The opening angle of the cone relative to the particle's trajectory, θ_c , is given by $\cos \theta_c = \frac{1}{n\beta}$, where $\beta = v/c$. Hence, a measurement of θ_c and a particle's momentum from the tracking stations allow its mass hypothesis to be determined.

Schematic diagrams of the RICH1 and RICH2 detectors are shown in Fig. 2.9. Three different radiators, each with different refractive indices, are used because they provide sensitivity in different momentum ranges. The RICH1 detector is located upstream of the magnet and is used to identify particles in the lower momentum range $2 < p < 40 \text{ GeV}/c$. It is composed of two radiators: aerogel [57] and C_4F_{10} . The RICH2 detector is located downstream of the tracking stations and identifies particles in the momentum region $15\text{--}100 \text{ GeV}/c$ using a CF_4 radiator. Figure 2.10 shows a two-dimensional distribution of θ_c and momentum of isolated tracks in C_4F_{10} , and clear bands for different particle species can be distinguished. The radiated cones of photons are reflected by spherical and flat mirrors towards an array of hybrid photon detectors (HPDs), where they are seen as rings. The value of θ_c can be determined from the diameter of the ring. The construction of a particle identification variable used in the analysis to separate pion and kaon candidates, $\text{DLL}_{K\pi}$, is discussed further in Sect. 2.7.2.

The performance of the RICH detectors depends on the angular deviation $\Delta\theta_c$, which is the difference between the reconstructed θ_c and the Cherenkov angle calculated from the momentum of the track and the refractive index of the radiator for a known particle species. The value of $\Delta\theta_c$ is evaluated for each photoelectron with respect to the measured track and found to have a spread of $1.618 \pm 0.002 \text{ mrad}$ for C_4F_{10} and $0.68 \pm 0.02 \text{ mrad}$ for CF_4 [58].

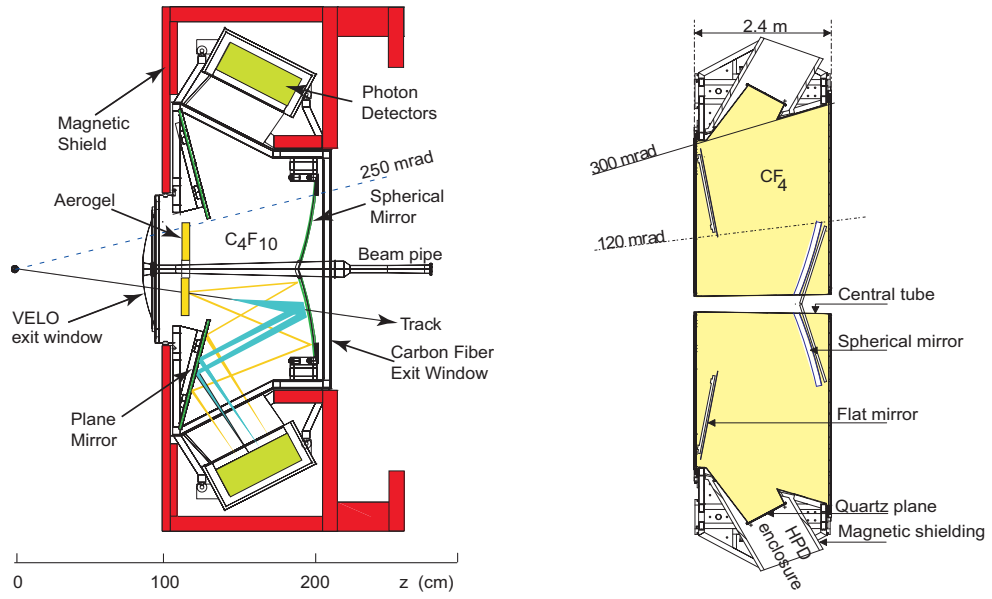


Figure 2.9. The RICH detectors: (left) side view schematic of RICH1 and (right) top view schematic of RICH2. Reproduced from [52].

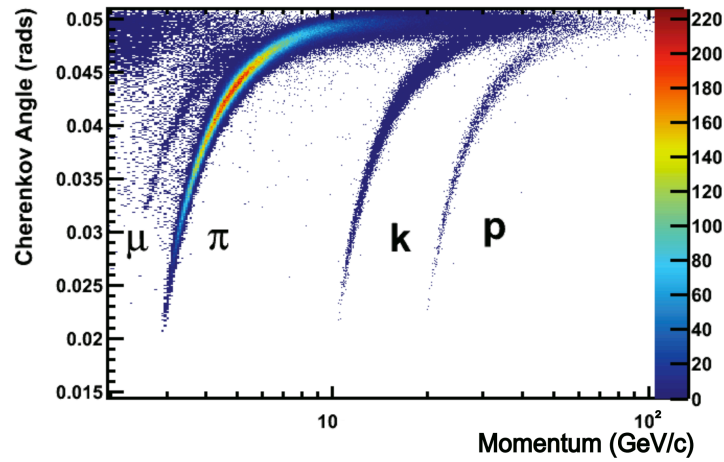


Figure 2.10. The reconstructed Cherenkov angle as a function of momentum showing bands for muons, pions, kaons and protons in a C₄F₁₀ radiator. Isolated tracks, whose Cherenkov rings do not overlap with any other rings, are used. Reproduced from [58].

2.4 Calorimeters

The calorimetry system of LHCb consists of four components: the scintillating pad detector (SPD), the preshower (PS), the electromagnetic calorimeter (ECAL), and the hadronic calorimeter (HCAL). Each of these subdetectors are constructed with organic

scintillators and the light produced is collected by photomultiplier tubes. The position and transverse energy (E_T) information collected by the calorimeter system plays a crucial role in the hardware trigger and in the reconstruction of electrons, prompt photons and π^0 mesons.

The SPD and PS consist of arrays of scintillator pads with a thin lead layer between them. The scintillator pads have a cell granularity of 4×4 cm closest to the beam line. This increases to 6×6 cm and 12×12 cm moving outwards to reflect the lower density of hits away from the beam line. Only charged particles deposit energy in the SPD, and both electrons and photons interact with the thin lead layer and deposit energy in the PS, with the shower continuing into the ECAL. This feature is used to distinguish electrons from photons in the hardware trigger. Photon candidates are identified by an ECAL cluster with hits in the corresponding PS cells upstream of it, while electron candidates are additionally required to have at least one SPD hit in front of the PS cells.

The ECAL has a Shashlik construction [59] with 66 layers of 2 mm thick lead as the absorber and 4 mm thick scintillator as the active material. The scintillator layers have a layout that is identical to the SPD/PS, where the cell granularity decreases away from the beam line. The thickness of the ECAL was chosen to be 25 radiation lengths in order to fully contain the electromagnetic shower from high energy photons. The energy resolution of the ECAL is given by $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\%$ (E in GeV), which describes the stochastic and constant terms.

Charged hadrons deposit only a small amount of energy in the calorimeter systems upstream of the HCAL. The HCAL has a similar Shashlik structure to the ECAL but with 1 cm thick iron absorbers. The layout of one HCAL module is shown in Fig. 2.11. In contrast to the ECAL, the layers alternate in the x -direction. Since the trigger does not require the full hadronic shower to be captured, the thickness of the HCAL is limited to 5 interaction lengths. As hadronic showers are typically larger than electromagnetic showers, the scintillating cell size is larger than that of the other calorimeter components:

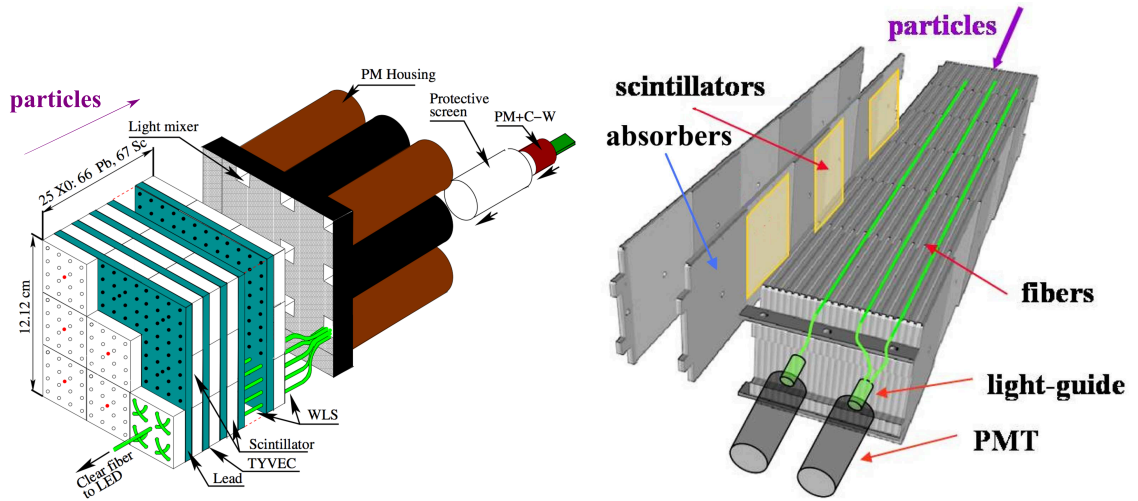


Figure 2.11. Exploded schematic of (left) an ECAL module and (right) an HCAL module, showing the alternating layers of absorber and scintillator, and the positions of (green) optical fibers leading to photomultiplier tubes (PMTs). The approximate direction of particle motion is indicated. Reproduced from [52, 60].

the inner section has a cell granularity of 13×13 cm and the outer section 26×26 cm.

The energy resolution of the HCAL is $\frac{\sigma_E}{E} = \frac{69\%}{\sqrt{E}} \oplus 9\%$ (E in GeV).

2.5 Muon system

The muon system consists of five stations labelled M1–M5. As shown in Fig. 2.12, the M1 station is situated upstream of the SPD; the four remaining stations are downstream of the HCAL and are separated by 80 cm thick iron absorbers which prevent hadrons from penetrating. Muons with momenta greater than $6 \text{ GeV}/c$ will typically traverse all five stations.

Each station consists of multiwire proportional chambers (MWPCs), each containing a gas mixture of 40% Ar, 55% CO_2 and 5% CF_4 . The inner region of M1 has ‘triple-GEM’ (gas electron multiplier) detectors which are used due to their higher radiation hardness. The ionising medium of the triple-GEMs is a gas mixture of 45% Ar, 15% CO_2 and 40%

CF_4 , which not only provides radiation hardness but a favourable time resolution of better than 3 ns.

The MWPCs are divided into pads that are read out by a single electronics channel. The granularity of the pads decreases away from the beam line in four regions labelled R1–R4 and indicated on Fig. 2.12, whose segments in the x, y -directions scale in the ratio 1:2:4:8 to give roughly equal particle flux and occupancy in each region. M1–M3 have a high spatial resolution along the x -axis, within the bending plane, allowing a stand-alone transverse momentum, p_T , to be measured with a resolution of 20%. This is crucial for providing fast information to the muonic hardware trigger which is based on p_T thresholds of muon tracks.

Muon candidates are identified in this thesis using a boolean selection, IsMuon , which considers the number of stations containing a hit within a field of interest defined around a track extrapolation [61]. As tracks with larger momenta more likely to reach M4 or M5, they require a greater number of stations to contain hits. The IsMuon requirement identifies muons with a $> 90\%$ efficiency and only a percent-level hadron misidentification rate across a large p and p_T range, as shown in Fig. 2.13.

2.6 Trigger system

The proton bunch crossing frequency that produces interactions visible to LHCb is approximately 10 MHz. (A visible interaction contains at least two charged particles that leave a sufficient number of hits in the detector for them to be reconstructed.) A dedicated trigger system [50] determines which events contain signatures that distinguish the physics of interest from the large background from QCD processes and writes them to storage. The trigger system is divided into two levels. Firstly a Level-0 (L0) trigger is implemented in hardware which reduces the rate to 1 MHz. This is the maximum rate at which the entire detector can be read out, imposed by the front-end

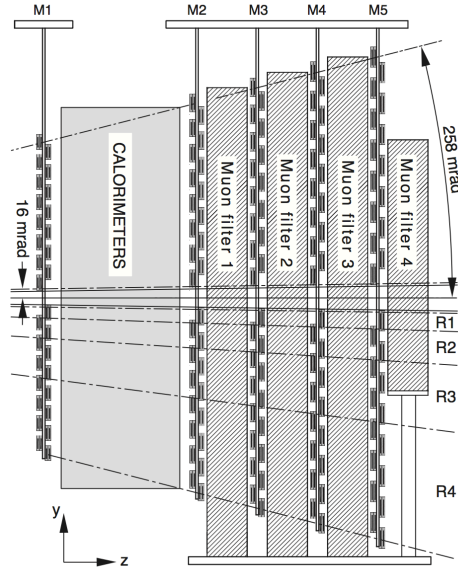


Figure 2.12. Side view of the muon stations, showing the relative positions of the five stations (M1–M5), the iron absorbers (‘filter’) and the calorimeters. The division of each muon station into four regions R1–R4 is denoted by the dash-dotted line. Reproduced from [52].

electronics. Secondly, a High Level Trigger (HLT) is implemented in software and this reduces the rate further to 2–5 kHz. The events written to storage are then subject to so-called *offline* processing, where a more accurate detector alignment and calibration is used in the reconstruction, and thence signal events are selected.

2.6.1 Level-0 trigger

The decay products of B mesons typically have high p_T due to the large B mass. Hence, the L0 trigger selects high E_T clusters in the calorimeters due to hadrons, photons and electrons, and high p_T muons in the muon systems. The trigger decision also requires a maximum number of SPD hits to prevent the HLT processing farms from saturating.

The L0 calorimeter trigger selects the 2×2 calorimeter cell cluster with the highest $E_T = \sum E \sin \theta$, where E is the energy deposited in a calorimeter cell, θ is its angle with respect to the z -axis and the sum is over the four cells in the cluster. The trigger creates L0Hadron, L0Photon or L0Electron candidates depending where energy has been

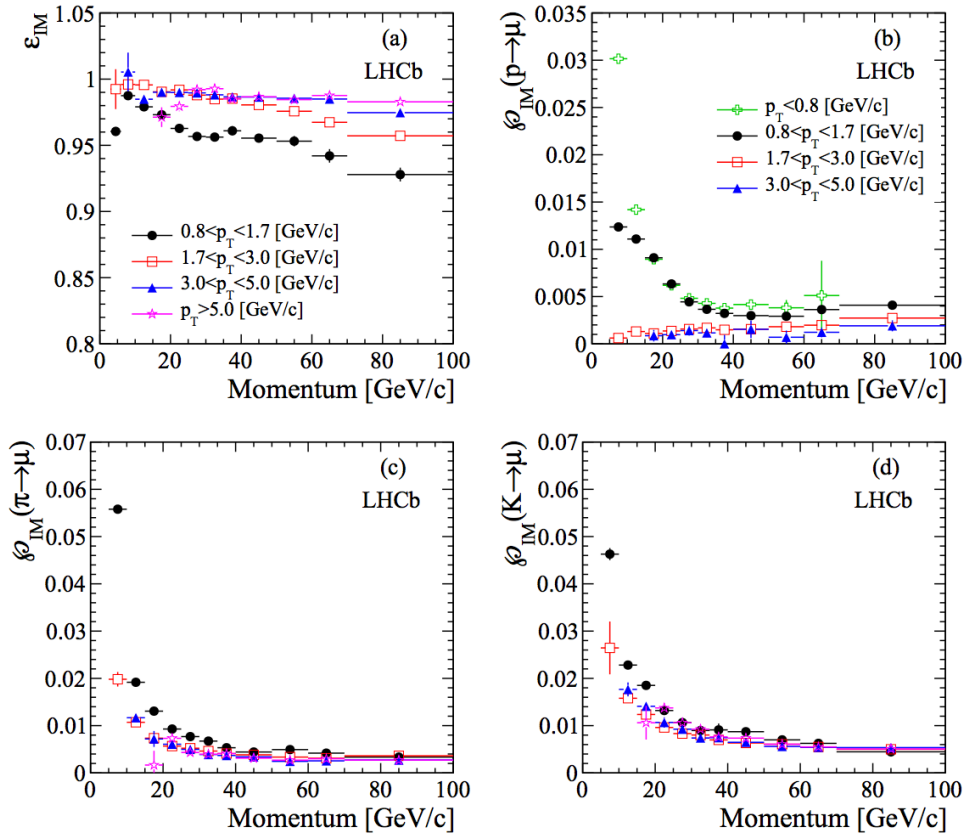


Figure 2.13. Efficiency of identifying muons, ϵ_{IM} , and probability of misidentifying p , π or K as muons, ρ_{IM} , of the IsMuon binary selection, as a function of momentum, in ranges of p_T . Reproduced from [61].

deposited in the calorimeter systems. Events containing at least one candidate with E_T above a fixed threshold are retained. For instance, an L0Hadron candidate is the highest E_T cluster in the HCAL that also contains the energy of the matching ECAL cluster, and has a threshold of approximately $E_T > 3.5$ GeV. The muon system is divided into quadrants and eight candidates based on the two muon tracks with the highest p_T from each quadrant are selected. The p_T is based on a standalone measurement of tracks with hits in all five muon stations. A threshold is set on either the largest p_T of the eight candidates (L0Muon), or on the product in p_T of the two tracks (L0DiMuon). Additionally, any events which pass any one of the above triggers are saved by the L0Global trigger.

2.6.2 High Level Trigger

The events accepted by the L0 trigger are placed in a buffer to be processed by the HLT trigger. The HLT is run on a farm of 2000 computing nodes and is divided into HLT1, where the event is only partially reconstructed, and HLT2, where the full event is reconstructed. Some of the reconstruction algorithms used by the HLT deviate from those used offline because of the limited time budget per event.

At HLT1, PVs are formed using VELO tracks (defined in Sect. 2.7.1) and a forward tracking algorithm matches the VELO tracks to hits in the tracking stations. An inclusive beauty and charm trigger that is highly efficient for decays with no lepton in the final state, `Hlt1TrackAllL0`, selects a track that has a p_T greater than $1.7 \text{ GeV}/c$ and a displacement from the PV of $IP > 0.1 \text{ mm}$. A similar trigger for tracks with matching hits in the muon chambers, *i.e.* significant displacement from all PVs and $p_T > 1.0 \text{ GeV}/c$, `Hlt1TrackMuon`, can be used to select semileptonic B decays [62].

After HLT1, the event rate is reduced to 40 kHz and is now low enough to perform forward tracking of all VELO tracks in HLT2. All tracks are required to have $p_T > 0.5 \text{ GeV}/c$ and $p > 5 \text{ GeV}/c$. The largest bandwidth in HLT2 is allocated to ‘topological’ triggers which select partially reconstructed b -hadron decays that have at least two charged decay products and a displaced secondary vertex [63]. Combinations of two, three and four-tracks, collectively known as n -body objects, are formed. Since the partially reconstructed candidates will have a mass below the nominal B mass, they are selected using a corrected mass variable, m_{corr} , defined as

$$m_{\text{corr}} = \sqrt{m^2 + |p_{T\text{miss}}|^2/c^2} + |p_{T\text{miss}}|/c, \quad (2.1)$$

where $p_{T\text{miss}}$ is the missing transverse momentum to the flight direction defined by the relative positions of the PV and the n -body vertex, and m is the invariant mass of

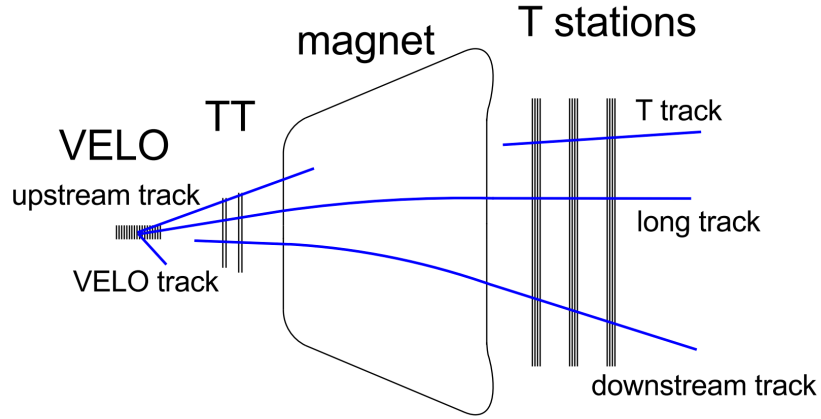


Figure 2.14. Schematic illustration of the VELO, TT, magnet and tracking stations (T1–T3), and (blue) the different classes of reconstructed tracks. The track types are called VELO, long, downstream and T. Reproduced from [65].

the observed n -body object. The value of m_{corr} is the least possible mass of the parent particle if there was a massless particle missing in its decay. The n -body objects are selected using a multivariate classifier [64] that takes m_{corr} and six other kinematic variables as input. These form three triggers called $\text{Hlt2Topo}N\text{BodyBBDT}$, where $N = \{2, 3, 4\}$. There are a further three triggers called $\text{Hlt2TopoMu}N\text{BodyBBDT}$ based on the same multivariate classifier with the added condition that one of the particles is identified as a muon.

2.7 Reconstruction

2.7.1 Track reconstruction

Track reconstruction algorithms combine the hits in different subdetector layers to form particle tracks from the VELO to the calorimeters. When tracks or track segments are reconstructed they are classed according to which subdetectors contain the hits. These are illustrated in Fig. 2.14 and are:

- *Long* tracks which traverse the VELO and all the tracking stations. They have the most precise momentum measurements because they experience the full magnetic field. Unless otherwise stated, the tracks used in the proceeding analyses are long tracks.
- *Downstream* tracks only leave hits in the TT and in the tracking stations. They typically result from decays of particles such as the K_S^0 meson that can live long enough to decay outside the VELO.
- *VELO* tracks are reconstructed using only hits in the VELO sensors. These tracks are straight lines due to the absence of a magnetic field.
- *T* tracks contain only hits in T1–T3.

The reconstruction of long tracks begins with the construction of VELO tracks. In the first adopted algorithm, the VELO tracks are associated with one hit in the T station to determine the particle momentum and trajectory. A search for additional hits along the particle's trajectory in T1–T3 is then performed to find the best combination of hits defining the long track. The second algorithm performs a separate track reconstruction to form T tracks, which are then combined with VELO tracks to form long tracks. Long tracks used in the analysis come from both algorithms with duplicates removed. Downstream tracks are reconstructed by first searching for T tracks and extrapolating them through the magnetic field upstream to the TT to find corresponding hits.

All of the tracks are fitted to obtain a best estimate of the track parameters [66], taking multiple scattering into account and correcting for energy loss due to ionisation. The χ^2 of this fit and a dedicated neural network classifier are used to reduce the number of fake tracks that are formed by wrongly associating track segments that do not correspond to the path of a genuine charged particle.

2.7.2 Particle identification

An essential component of reconstructing particle decays is the particle identification algorithm. A simultaneous global fit to RICH1 and RICH2 photon hits is used to determine the likelihood that a ring of Cherenkov photons associated with a given track matches the expectation from a π , K , p , e or μ mass hypothesis. The likelihood is labelled $\mathcal{L}_X$, where X is the hypothesised particle. Since the most abundant particles from a pp interaction are pions, the pion hypothesis is initially assumed. Then, for each track the log-likelihood difference,

$$\text{DLL}_{X\pi} = \log \mathcal{L}_X - \log \mathcal{L}_\pi, \quad (2.2)$$

is computed using the mass hypothesis of particle X . An overall event likelihood for all tracks in the three radiators in the entire event is constructed by combining the individual likelihoods for each track. It is necessary to consider an event likelihood because the Cherenkov rings from different tracks tend to overlap. The event likelihood is computed iteratively, and in each iteration, different mass hypotheses are considered for each track and the expected pattern of HPD hits due to Cherenkov photon emission is compared to the observed pattern. The tracks are then assigned the identities of the particles whose mass hypotheses provide the maximum event likelihood [67]. In the offline selection, cuts on the $\text{DLL}_{K\pi}$ variable are applied. The efficiency and misidentification probability for two $\text{DLL}_{K\pi}$ cuts as a function of track momentum is shown in Fig. 2.15. When averaged over a momentum range of 2–100 GeV/ c , the kaon identification efficiency is 95% and the misidentification fraction of a pion as a kaon is 10%. Although the response in simulation is shown to model the overall trend in data reasonably well, the absolute efficiencies are important in this thesis and a data-driven method of calculating them is described in Sect. 3.2.3.

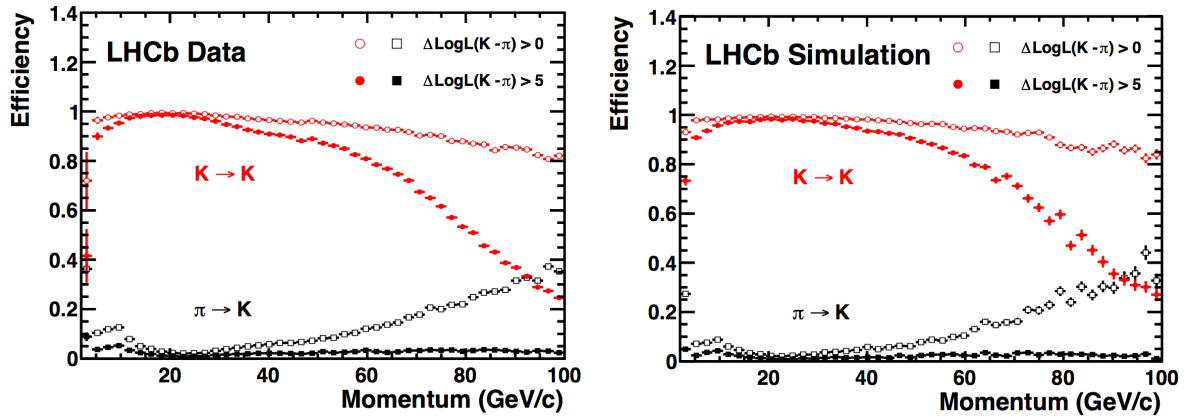


Figure 2.15. The efficiency of identifying kaons and misidentifying pions as kaons in (left) data and (right) simulation. The solid and open points refer to two different particle identification ($DLL_{K\pi}$) requirements. Reproduced from [50].

2.7.3 Stripping

After the HLT2 selection, the data collected are subsequently prepared offline with the latest alignment, calibration and PID information used to perform a full reconstruction of all tracks and vertices, in a process known as *stripping*. Hundreds of dedicated *stripping lines* are constructed, in which raw data are used to reconstruct individual decay channels required for analysis by building composite particles and imposing selection criteria. This computationally expensive task is processed centrally so that the analyst only needs to handle datasets of a manageable size. The stripping lines used in this thesis are discussed in Sects. 3.1, 4.1 and 6.1.

2.8 Simulation

An accurate Monte Carlo (MC) simulation of production processes, particle decays and their interaction with the detector is a vital tool for any physics analysis. MC samples of the signal modes, control channels and backgrounds studied in this thesis were generated as described below.

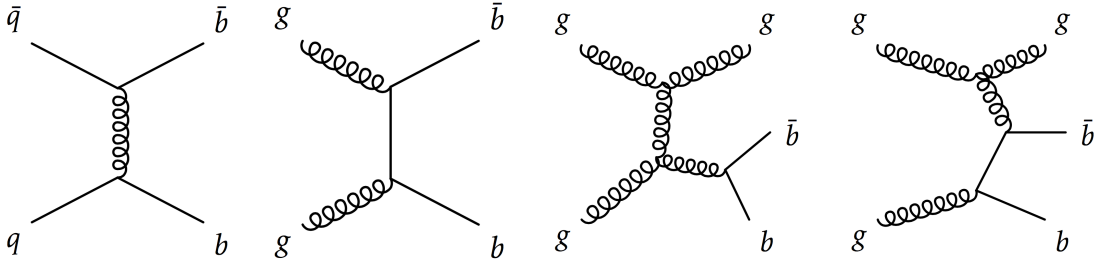


Figure 2.16. Feynman diagrams for leading order and higher order $b\bar{b}$ production processes in simulation [68].

The simulation of particle decays in the LHCb detector begins with pp collisions generated with a package called PYTHIA, either version 6.4 [69] or 8.1 [70] with a configuration of parton density functions given in Ref. [71]. At this stage, $b\bar{b}$ pairs are produced according to the underlying parton interactions scaled according to their cross-sections. Figure 2.16 shows the dominant production processes [68]. Both b quarks hadronise and their subsequent decays are handled by a package called EVTGEN [72], which specifies the properties of b -hadron decays including particle kinematics and angular correlations. One of the b quarks is chosen at random and forced to decay according to a user-specified decay mode and with all of the decay daughters within the LHCb detector geometry. In all of the MC samples, $D \rightarrow K_s^0 h^+ h^-$ decays are generated uniformly in the (m_-^2, m_+^2) Dalitz plot (*i.e.* no resonant structure) and no CP violation is present in any of the decays. The other b quark in the pair is required to decay according to a set list of all b -hadron decays scaled by their known branching ratios.

The GEANT4 toolkit [73] is used to simulate the passage of particles through the specified LHCb detector geometry and material. The physics processes for each particle (such as multiple scattering, bremsstrahlung [74], energy deposits and Cherenkov radiation) and the detector response is recorded, including hits in the active material. The simulated detector response is adjusted to reflect the resolution observed in test beams and calibration data. The response, including the L0 trigger information, is then digitised and converted into the same format as that read out by the real detector. From

this point, the simulated and collision data are reconstructed in the same way, except that the generator-level information from the simulated particles is retained.

CHAPTER 3

Selection and parameterisation of $B \rightarrow DK$ decays

3.1 Selection of $B \rightarrow Dh$, $D \rightarrow K_s^0 h^+ h^-$ candidates

In this section, the reconstruction and selection of candidates for the signal mode, $B \rightarrow DK$, and the control mode, $B \rightarrow D\pi$, in the $D \rightarrow K_s^0 \pi^+ \pi^-$ and $K_s^0 K^+ K^-$ final states is described. The $B \rightarrow DK$ and $B \rightarrow D\pi$ decays are collectively referred to as $B \rightarrow Dh$ and the kaon or pion produced directly in the B decay is called the *bachelor* particle.

3.1.1 Reconstruction and trigger requirements

As described in Sect. 2.7.3, candidate $B^\pm \rightarrow Dh^\pm$, $D \rightarrow K_s^0 h^+ h^-$ decays are formed offline and stored in a centrally-produced dataset during a processing stage known as stripping. The candidate decays are formed by reconstructing composite particles in the order: $K_s^0 \rightarrow \pi^+ \pi^-$, $D \rightarrow K_s^0 h^+ h^-$ and $B^\pm \rightarrow Dh^\pm$. The $K_s^0 \rightarrow \pi^+ \pi^-$ candidates are reconstructed in two categories. When two long pion tracks are combined, the K_s^0 candidate is categorised as LL, and if two downstream pion tracks are combined, the K_s^0 candidate is categorised as DD. The two K_s^0 types are analysed separately because the mass resolution of LL K_s^0 candidates is expected to be better than that of DD K_s^0 candidates (Sect. 2.7.1). The size of the stripping dataset is limited due to

constraints on computing infrastructure. This is achieved by placing loose requirements on the candidates to remove random combinations of particles, known as combinatorial background. These requirements include loose thresholds on the p , p_T and χ^2/ndf of each track. In addition, the K_S^0 (D) candidate is required to have a reconstructed mass within $30 \text{ MeV}/c^2$ ($100 \text{ MeV}/c^2$) of its known mass [75]. The D candidate is also required to have a good quality vertex that is displaced from the PV by 6 standard deviations in order to distinguish it from D mesons from prompt D^* decays which are also produced in abundance at LHCb. The B candidate is required to have a mass, $m(B)$, in the range $4750\text{--}7000 \text{ MeV}/c^2$ and $\chi_{\text{IP}}^2 < 25$, where χ_{IP}^2 is defined as the difference in χ^2 of the PV fitted with and without the B daughter tracks. All the composite particles are required to pass minimum thresholds of p and p_T .

The DecayTreeFitter (DTF) framework [76] is used to refit the entire decay chain of each B candidate passing the stripping requirements, with one or more constraints imposed. During the refit, the four-momenta of each particle are altered to their best fit values under the given constraints. The constraints used are labelled as follows. The labels `D0const` and `KS0const` indicate that the D and K_S^0 masses, respectively, are constrained to their known values. The label `PVconst` is used to indicate that the B momentum vector direction is constrained to be parallel to the vector joining the primary vertex (PV) to the B decay vertex. The $m(B)$ and the Dalitz coordinates (m_-^2, m_+^2) used in this analysis, unless otherwise specified, are obtained after the refit with `D0const`, `KS0const` and `PVconst`. The refit improves the $m(B)$ resolution. The refitted $m(B)$ distribution of $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$, $DD K_S^0$, candidates passing the stripping requirements is shown in Fig. 3.1. The refit also improves the resolution of the Dalitz coordinates, and the `D0const` constraint ensures that the candidates fall within the kinematically-allowed phase space for the $D \rightarrow K_S^0 h^+ h^-$ decay. The K_S^0 mass used is determined using a refit with `D0const` and `PVconst`. Candidates that fail either of the two refits are removed from the analysis.

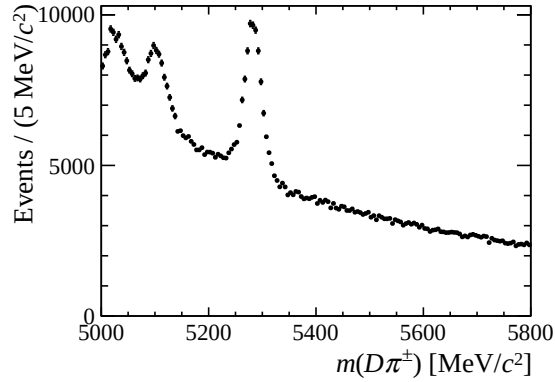


Figure 3.1. The refitted $m(D\pi^\pm)$ distribution for $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ (DD K_s^0) candidates after stripping from 2011 and 2012 data. The subsidiary peaks are described later in the chapter.

The trigger decision for each candidate is classified as T0S (Trigger On Signal) if the particles associated with the signal candidate triggered the event and as TIS (Trigger Independent of Signal) if other particles produced in the pp interaction, that are not associated with the signal candidate, triggered the event. At the hardware trigger level, the B candidates are required to be classified as L0Hadron T0S or L0Global TIS. The L0Global TIS requirement adds an additional 30% of candidates. Here, one track in the event that is not associated with the signal candidate satisfies any one of the hardware trigger requirements. At the software trigger level, the B candidates are required to be classified as Hlt1TrackAllL0 T0S and Hlt2TopoNBodyBBDT T0S, where $N = 2, 3$ or 4 . The triggers requirements were described in Sect. 2.6.

3.1.2 Boosted decision tree

A boosted decision tree [64] (BDT) is a multivariate event classifier that can be used to classify events as signal-like or background-like. The use of a BDT instead of a cut-based selection, in which individual requirements are placed on many variables, is particularly advantageous when there is a large degree of overlap in the distributions of the discriminating variables in signal and background. This is true of the signal $B \rightarrow Dh$, $D \rightarrow K_s^0 h^+ h^-$ decays and the combinatorial background in data, hence a BDT

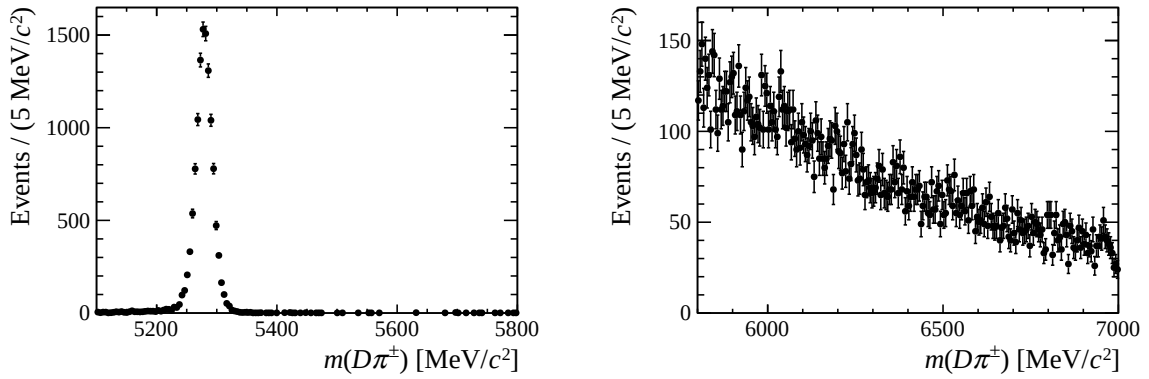


Figure 3.2. The $m(D\pi^\pm)$ distribution of $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ (DD K_s^0) candidates in the BDT (left) signal and (right) background training samples.

is used to suppress combinatorial background that passed the stripping requirements. The BDT is implemented using the `TOOLKIT FOR MULTIVARIATE ANALYSIS` [77]. The BDT is trained and tested using a sample of signal events and a sample of background events, as well as a list of *BDT variables* whose distributions are different in the signal and background samples. The signal and background samples are divided at random into training and testing samples of equal size, and the testing samples are used at a later stage. Many decision trees are constructed, and the BDT variables and cut values that provide the greatest separating power are found. A weight is assigned to each event and is updated on each iteration such that it increases (it is *boosted*) for misclassified events, where a signal event is incorrectly identified as a background event and vice versa. This allows misclassified events to have a higher chance of being correctly identified in subsequent trees. The gradient-boosting algorithm (BDTG), a specific method of updating the BDT weights at each iteration, is chosen for this analysis. A weighted average of all the decision trees is performed. The training results are used to analyse the BDT variables of each event in the data and an overall response value ranging from -1 (most background-like) to $+1$ (most signal-like) is assigned.

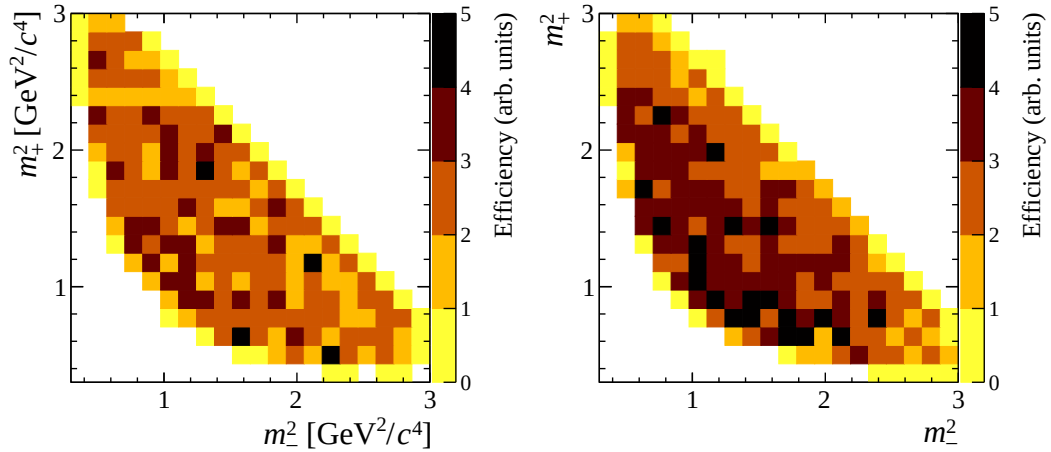


Figure 3.3. The Dalitz plot distributions of simulated $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ candidates in the BDT signal training samples for (left) LL and (right) DD K_s^0 .

Training samples

Both the signal and background samples must be independent of the data to be analysed. The signal training sample is obtained from $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ MC events in the range $5100 < m(B) < 5800 \text{ MeV}/c^2$, which is the approximate region in which the $m(B)$ variable will be fitted. The $m(B)$ distribution of these candidates is shown in Fig. 3.2 (left). The background training sample is obtained from $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ data candidates in the region $5800 < m(B) < 7000 \text{ MeV}/c^2$, known as the upper sideband. This mass range is sufficiently far away from the known B mass that candidates are unlikely to be real B candidates. As seen in Fig. 3.2 (right), no peaking structures are observed in this range. Candidates with $m(B) > 5800 \text{ MeV}/c^2$ are not subsequently used in the analysis and therefore this provides an independent sample for training. Although the training samples consist of only $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ candidates, this BDT training can be used to select all of the $B \rightarrow Dh$, $D \rightarrow K_s^0 h^+ h^-$ samples because the most discriminating BDT variables are topological and are hence insensitive to whether the bachelor is a pion or a kaon. A separate BDT is trained for LL and DD candidates as some of their distributions differ. The Dalitz plot distributions of the signal training samples are given in Fig. 3.3.

The candidates in the signal and background training samples are subject to initial requirements that eliminates candidates that would not pass the final selection. This is found to improve the BDT performance. The requirements are:

- The B candidate must satisfy the trigger requirements described.
- The bachelor must have information from the RICH detector and $p < 100 \text{ GeV}/c$.
- The reconstructed D mass lies within $25 \text{ MeV}/c^2$ of the known D mass.
- The reconstructed K_S^0 mass lies within $15 \text{ MeV}/c^2$ of the known K_S^0 mass.
- For LL candidates only, the separation of the K_S^0 vertex from the D vertex in units of χ^2 (FDCHI2) must be greater than 100. This removes events where the D decays to four pions. This requirement will be discussed in greater detail.

For the DD BDT, the signal training sample contains approximately 6000 events and the background training sample contains approximately 10 000 events. The training samples for the LL BDT are half this size.

The list of BDT variables used to classify events is informed by the selection requirements used in the previous LHCb analysis of $B \rightarrow Dh$, $D \rightarrow K_S^0 h^+ h^-$ decays [41]. Some of the variables are transformed using a logarithm function to increase their separation power. Variables found to be highly correlated with $m(B)$ are omitted from the BDT in order to avoid sculpting the $m(B)$ spectrum. Ten BDT variables are chosen for DD K_S^0 candidates: the χ^2 of the decay chain refit used to obtain $m(B)$, χ_{IP}^2 of the B , B momentum, the B isolation variable, D momentum, vertex separation χ^2 of the B decay vertex from the PV, $\log(\chi_{\text{IP}}^2)$ of the bachelor particle, $\log(\chi_{\text{IP}}^2)$ of the D and $\log(\chi_{\text{IP}}^2)$ of the two D daughter pions. The B isolation variable is a quantity designed to ensure the B candidate is well isolated from other tracks in the event, and is the asymmetry between the p_T of the signal candidate and the vector sum of the p_T of other tracks that lie within a distance of 1.5 rad in $\eta - \phi$ space, where ϕ is the azimuthal angle. For LL

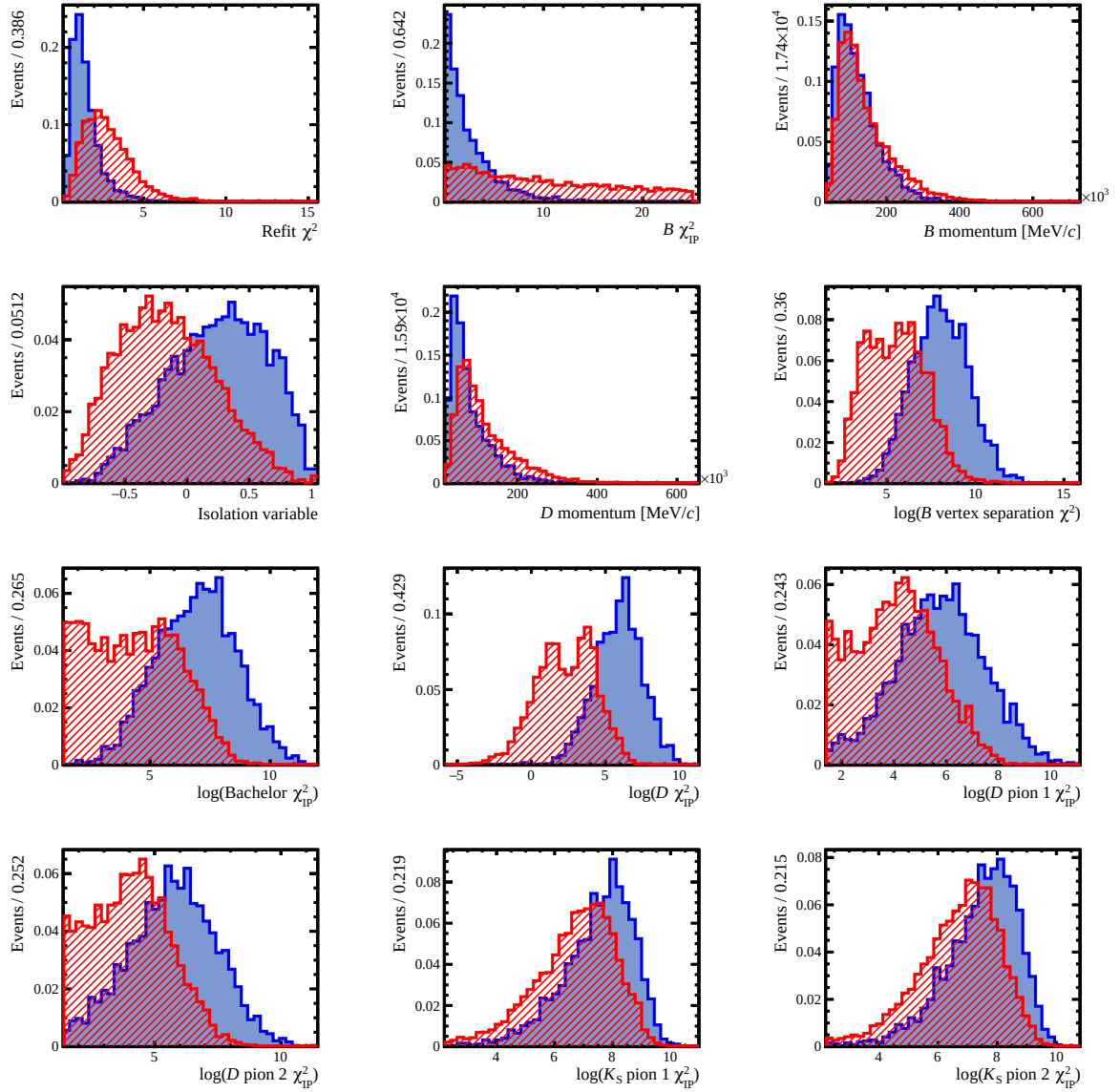


Figure 3.4. The 12 BDT training variables for the LL K_S^0 candidates in the (blue) signal training sample and in the (red) background training sample.

K_S^0 candidates, the same list of ten BDT variables as that used for DD K_S^0 is used, with the addition of the $\log(\chi_{IP}^2)$ of the two K_S^0 daughter pions. The distributions of these variables in the signal and background training samples are shown in Fig. 3.4 for LL K_S^0 and in Fig. 3.5 for DD K_S^0 .

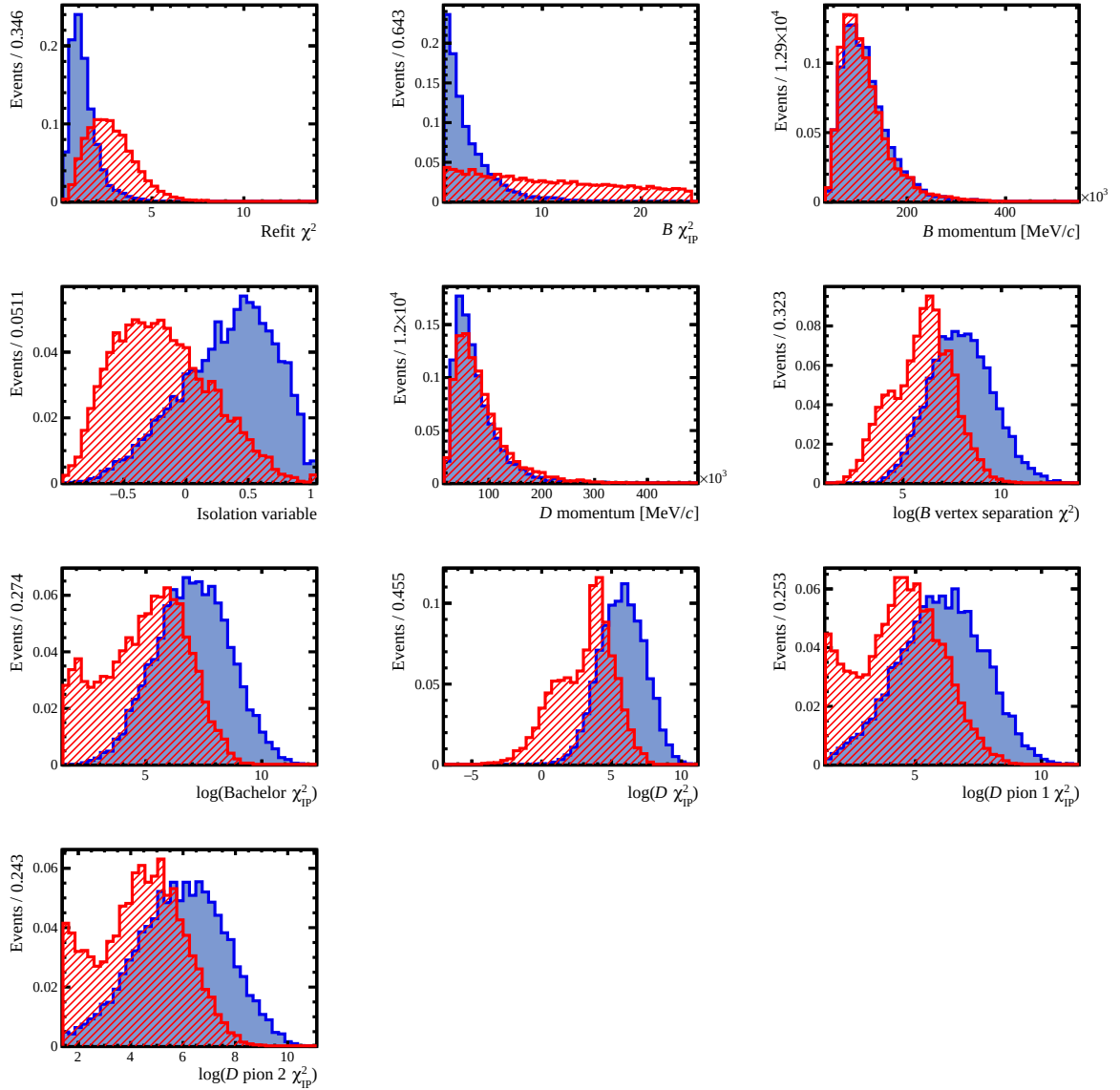


Figure 3.5. The 10 BDT training variables for the DD K_S^0 candidates in the (blue) signal training sample and in the (red) background training sample.

BDT performance and choice of working point

The results of the BDT training are applied back to the training samples. As shown in Fig. 3.6, the response values peak strongly at +1 (−1) for the signal (background) training samples. The training results are also applied to the testing samples to test for the presence of overtraining, which occurs when the classifier has learnt statistical fluctuations present in the given training sample. The results are also shown in Fig. 3.6,

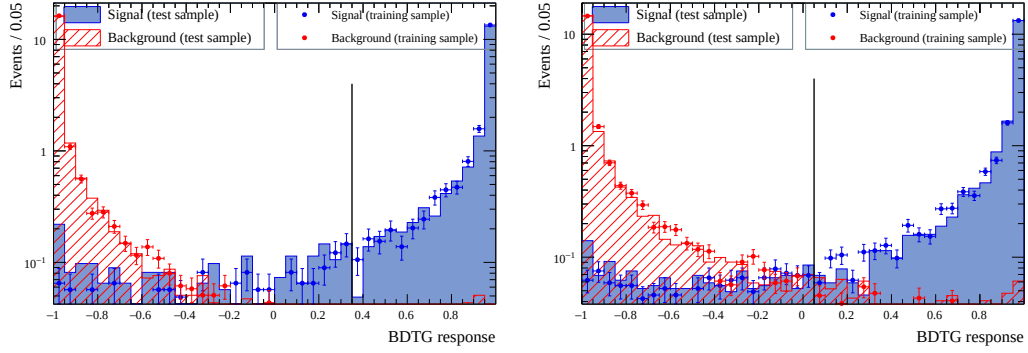


Figure 3.6. BDT response values of (left) LL and (right) DD K_S^0 candidates for the signal and background samples used in training and testing. The selection requirement of response value greater than 0.35 and 0.05 for LL and DD are indicated.

and shows that the distribution of response values in the training and testing samples are statistically compatible and that there is no evidence of overtraining.

The discriminating power of the BDT variables is ranked during the training stage and is given in Table 3.1. Although there are differences between the ranking of variables for the two K_S^0 types due to slightly different training variable distributions, the most powerful variables are the B isolation variable and the $B \chi_{\text{IP}}^2$. This is unsurprising, given that a B candidate containing tracks that do not come from the same B decay (*i.e.* combinatorial background) is unlikely to be well isolated and is expected to have a low $B \chi_{\text{IP}}^2$.

The signal yields and purities that are obtained after different requirements on the BDT response value are estimated using a simplified fit of the $m(B)$ distribution of $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ candidates. The yields and purities are used to generate realistic pseudoexperiments in order to evaluate the impact that different BDT requirements have on the sensitivities to the CP observables $x_{\pm}$ and $y_{\pm}$ (defined in Eqs. (1.38) and (1.39)). The pseudoexperiments are generated and fitted in a similar configuration to the one described later in Sect. 5.1.5 and the mean statistical uncertainties on $x_{\pm}$ and $y_{\pm}$ from the pseudoexperiments are measured. It is found that the sensitivity is

Table 3.1. BDT variables for LL and DD K_s^0 candidates, ranked according to their discriminating power.

Rank	LL variable	DD variable
1	$\log(D \chi_{\text{IP}}^2)$	B isolation variable
2	B isolation variable	$B \chi_{\text{IP}}^2$
3	$B \chi_{\text{IP}}^2$	D momentum
4	$\log(\text{bachelor } \chi_{\text{IP}}^2)$	$\log(D \chi_{\text{IP}}^2)$
5	Refit χ^2	$\log(D \text{ pion } 1 \chi_{\text{IP}}^2)$
6	$\log(K_s^0 \text{ pion } 1 \chi_{\text{IP}}^2)$	Refit χ^2
7	$\log(K_s^0 \text{ pion } 2 \chi_{\text{IP}}^2)$	$\log(D \text{ pion } 2 \chi_{\text{IP}}^2)$
8	D momentum	B momentum
9	B momentum	$\log(\text{bachelor } \chi_{\text{IP}}^2)$
10	$\log(D \text{ pion } 1 \chi_{\text{IP}}^2)$	$\log(B \text{ vertex separation } \chi^2)$
11	$\log(D \text{ pion } 2 \chi_{\text{IP}}^2)$	–
12	$\log(B \text{ vertex separation } \chi^2)$	–

very weakly dependent on the value of the BDT cut in the range -0.4 to 0.35 for LL K_s^0 and -0.55 to 0.05 for DD K_s^0 . The cut chosen is $\text{BDT} > 0.35$ for LL and $\text{BDT} > 0.05$ for DD candidates. When compared to the selection requirements in Ref. [41], this achieves a similar signal yield per pb^{-1} while reducing the combinatorial background in the signal region by 30–40%. Hence, it is also demonstrated that the BDT is able to provide a higher purity sample for a given signal yield compared to the selection previously developed. The cut is chosen on the basis that a lower level of combinatorial background can help reduce systematic uncertainties in the fit. When applied to the candidates in the testing samples, these requirements retain over 90% of simulated signal candidates and reject about 95% of candidates in the upper B mass sideband. The same requirements are applied to all $B \rightarrow Dh$, $D \rightarrow K_s^0 h^+ h^-$ samples.

3.1.3 Studies of peaking backgrounds

Dedicated studies are performed on charmless backgrounds and $D \rightarrow \pi\pi\pi\pi$ decays, which have the same final state particles as the signal. If every particle has been correctly identified, and all particles originate from the same B decay, these backgrounds will peak in the same place as the signal B mass and cannot be distinguished from the signal in the B mass distribution.

Charmless backgrounds occur when a B meson has decayed to a $K_S^0 h^+ h^- h^{(\prime)\pm}$ final state without decaying via a D meson, such as $B^\pm \rightarrow K^{*\pm}(\rightarrow K_S^0 \pi^\pm) \pi K$. The branching fraction of such decays is an order of magnitude less than $B \rightarrow DK$, so only a small contamination is expected. Since these backgrounds are expected to have a uniform D mass distribution, candidates in the D mass sideband are studied. However, the default selection suppresses events in the D mass sidebands because the refit χ^2 and the refitted K_S^0 mass both contain the $D0\text{const}$ constraint. Hence in order to estimate the potential contamination, $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ candidates are selected using the requirements in Ref. [41] and reproduced in Table 3.2, without the requirement on the refit χ^2 and with a requirement on the raw (no refit) K_S^0 mass.

This background is suppressed by requiring that the z -position of the D decay vertex $z(D)$ is downstream of the B decay vertex $z(B)$. In these decays, the true B and ‘ D ’ ($K_S^0 h^+ h^-$) vertices are at the same point, but reconstruction effects will push the reconstructed D vertex slightly in one z -direction with respect to the B , hence this cut will remove approximately half of any charmless backgrounds. To evaluate the effectiveness of this cut, the raw $m(B)$ distribution of candidates in the upper D mass sideband ($1925 < m(D) < 1955 \text{ MeV}/c^2$) is fitted before and after the application of the cut. The lower D mass sideband is not used as it may be populated by misidentified $K_S^0 K^+ K^-$ and $K_S^0 K^\pm \pi^\mp$ decays in the $K_S^0 \pi^+ \pi^-$ sample. The raw $m(B)$ distribution is fitted with an exponential function to model the combinatorial background and a Gaussian

Table 3.2. Selection requirements for $B \rightarrow D\pi$, $D \rightarrow K_S^0\pi^+\pi^-$ candidates in Ref. [41]. The variables are defined in the main text and DIRA is defined in Sect. 3.1.5. PDG refers to the values in Ref. [75].

Particle	Variable	LL K_S^0	DD K_S^0
K_S^0 daughter pions	χ_{IP}^2	> 16	–
	flight distance χ^2	> 100	–
	refitted – PDG mass	$\pm 15 \text{ MeV}/c^2$	$\pm 15 \text{ MeV}/c^2$
	DIRA	> 0.99	> 0.99
D daughter pions	χ_{IP}^2	> 9	> 16
	raw – PDG mass	$\pm 25 \text{ MeV}/c^2$	$\pm 25 \text{ MeV}/c^2$
	DIRA	> 0.99	> 0.99
	$z(D) - z(B)$	> 0	> 0
Bachelor	χ_{IP}^2	> 9	> 9
	RICH information	yes	yes
	p	$< 100 \text{ GeV}/c$	$< 100 \text{ GeV}/c$
	$\text{DLL}_{K\pi}$	< 4	< 4
B	χ_{IP}^2	> 25	> 25
	DIRA	> 0.9999	> 0.99995
	flight distance χ^2	> 169	> 169
	χ_{IP}^2	< 9	< 9
	refit χ^2	< 5	< 5

function, with a mean fixed to $5280 \text{ MeV}/c^2$, to model the charmless decays. The fit results are shown in Fig. 3.7. The red peaks indicate the potential contamination preferred by the fit, which is suppressed by the cut on $z(D) - z(B)$. The estimated contribution in the D signal region is no more than 5 background events in each of the four samples considered. Similar results are obtained with $B \rightarrow DK$ candidates.

In the case of LL $K_S^0\pi^+\pi^-$ candidates it is possible for two long pion tracks in the event to have a combined $m(\pi\pi)$ mass that falls within the K_S^0 mass window. To suppress these random combinations, the LL K_S^0 candidate is required to have a $\text{FDCHI2} > 100$. This ensures that all four pion tracks used to reconstruct the D meson do not originate from the same point. Since random combinations of pairs of pions in $D \rightarrow \pi\pi\pi\pi$ decays are expected to have a uniform K_S^0 mass distribution, candidates in the K_S^0 mass

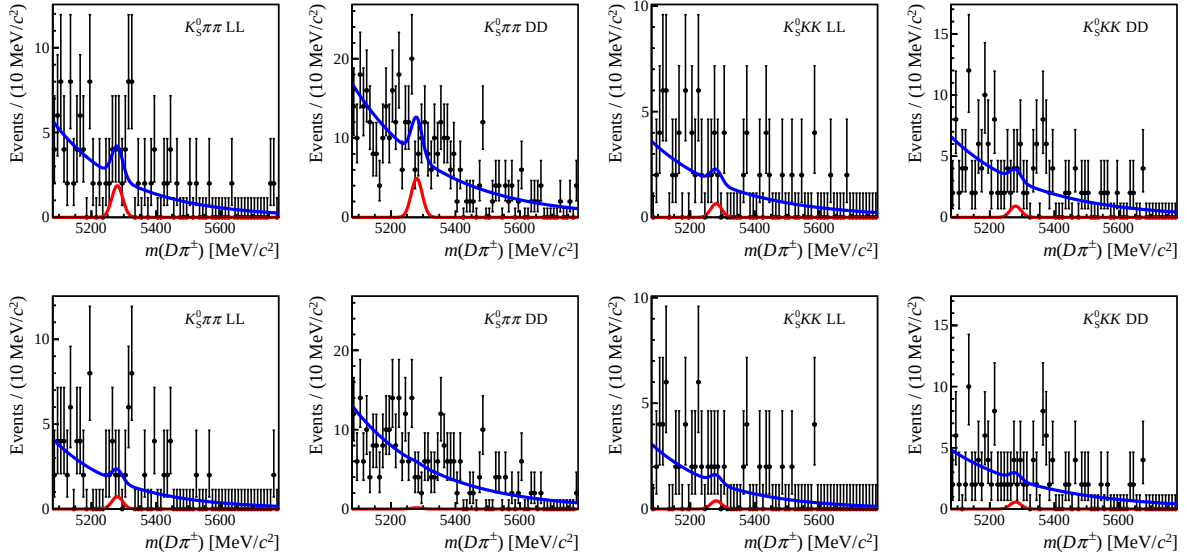


Figure 3.7. Raw $m(B)$ distribution of $B \rightarrow D\pi$ candidates in the upper D mass sideband in (left to right) the $K_S^0\pi^+\pi^-$ LL, $K_S^0\pi^+\pi^-$ DD, $K_S^0K^+K^-$ LL and $K_S^0K^+K^-$ DD final states. The data are shown (top) without and (bottom) with the cut to suppress charmless backgrounds. The fit results are superimposed, showing the (blue) total fit and (red) potential charmless background.

sideband are studied, defined as $m(K_S^0) < 477 \text{ MeV}/c^2$ or $m(K_S^0) > 517 \text{ MeV}/c^2$. Studies are performed using $B \rightarrow D\pi$ candidates satisfying the cuts in Table 3.2. The requirements on the K_S^0 FDCHI2 and the refit χ^2 are omitted. As before, the default selection cannot be used because the refit χ^2 and refitted mass variables contain KS0const and are therefore correlated with the quality of the reconstructed K_S^0 . The $m(\pi^+\pi^-\pi^+\pi^-\pi^\pm)$ spectrum of candidates in the K_S^0 sidebands before and after the proposed K_S^0 FDCHI2 cut is shown in Fig. 3.8. It can be seen that a large signal peak corresponding to $D \rightarrow \pi\pi\pi\pi$ decays is present at the known B mass, which is reduced to an $\mathcal{O}(0.1\%)$ contamination of the signal peak after the requirement on the K_S^0 FDCHI2 is included. 90% of simulated signal events passing the trigger and BDT requirements are retained by this cut.

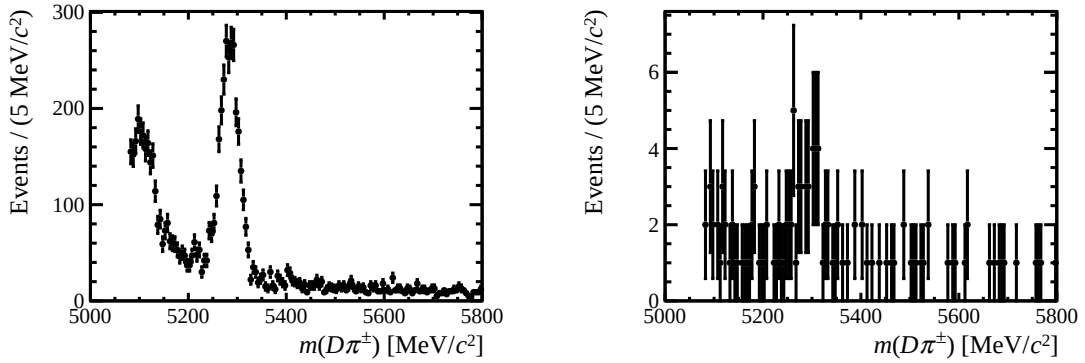


Figure 3.8. The raw B mass distribution for $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ candidates, where the reconstructed LL K_s^0 mass lies in the sidebands. The data (left) shows a peak corresponding to $D \rightarrow \pi\pi\pi\pi$ background and (right) after this background has been suppressed.

3.1.4 Particle identification requirements

PID requirements separate the different hadronic samples using the $DLL_{K\pi}$ variable defined in Sect. 2.7.2. For the $B \rightarrow D\pi$ and $B \rightarrow DK$ samples, the bachelor must have $DLL_{K\pi} < 4$ and $DLL_{K\pi} > 4$, respectively. For the $D \rightarrow K_s^0 K^+ K^-$ samples, the daughter kaon tracks must have $DLL_{K\pi} > -5$. In order for the $DLL_{K\pi}$ variable to be reliable, any track with a $DLL_{K\pi}$ requirement must also have information from the RICH detector and $p < 100 \text{ GeV}/c$.

3.1.5 Additional requirements

Some remaining combinatorial background is removed by placing cuts on the $DIRA$ of a particle, which is defined as the cosine of the angle between the momentum vector of the particle and the vector connecting the PV and its decay vertex. The $DIRA$ of the D , K_s^0 and B have the same requirements as those listed in Table 3.2. These cuts are highly efficient on signal.

In this analysis every B candidate is assigned a Dalitz plot bin number, which is obtained from the binning scheme provided in Ref. [3] and described in Sect. 1.6.3. The

exact position of the Dalitz plot kinematic boundary in the two-dimensional (m_-^2, m_+^2) phase space depends on the D mass. There is a $0.3 \text{ MeV}/c^2$ difference between the D mass in the 2008 BaBar amplitude model used to produce the binning [33] and that used by the implementation of DTF at the time of performing this work in 2012. In the DK signal region, 14 (5) candidates in the $K_S^0 \pi^+ \pi^-$ ($K_S^0 K^+ K^-$) samples are not assigned a Dalitz plot bin since they lie outside the defined binning scheme. Due to the small number of affected events, these events are simply excluded.

It is possible to reconstruct more than one B candidate per event. These candidates may have tracks in common and hence they are undesirable. Where more than one candidate is found, the one with the lower refit χ^2 is retained in the analysis and the other(s) discarded. After all of the above selection requirements have been imposed, the number of multiple candidates found is no larger than 0.6% in each sample.

3.1.6 B mass distribution and Dalitz plots of selected candidates

The $m(B)$ distributions of each sample passing the selection requirements are shown in Figs. 3.9 and 3.10. Dalitz plots of the candidates in the B mass signal window ($5247 < m(B) < 5317 \text{ MeV}/c^2$) are given in Figs. 3.11 and 3.12.

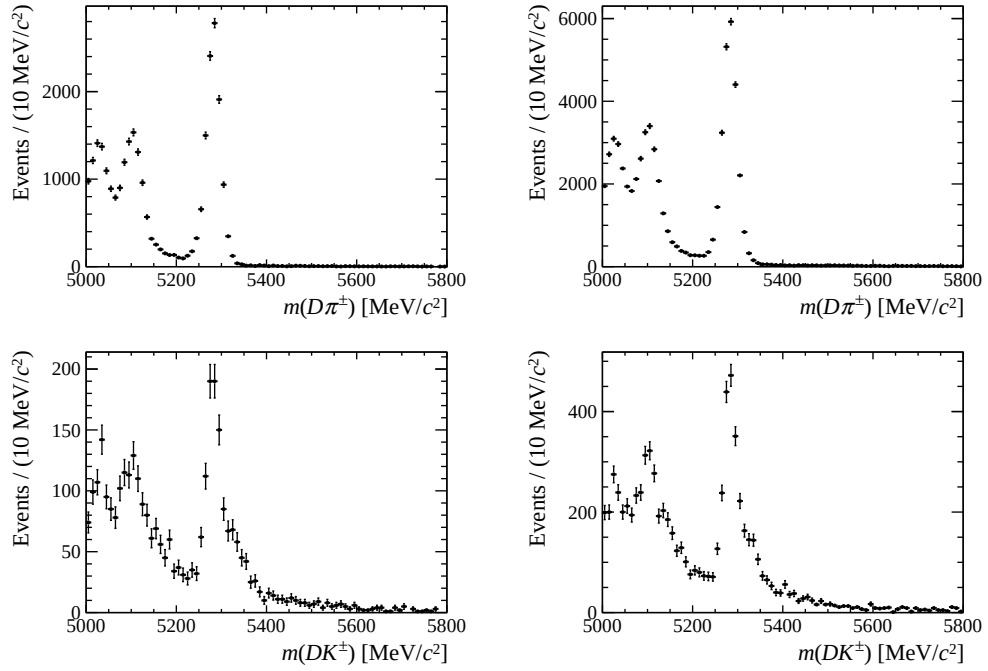


Figure 3.9. The B mass distribution of selected (top) $B \rightarrow D\pi$ and (bottom) $B \rightarrow DK$, $D \rightarrow K_s^0 \pi^+ \pi^-$ candidates for (left) LL and (right) DD K_s^0 .

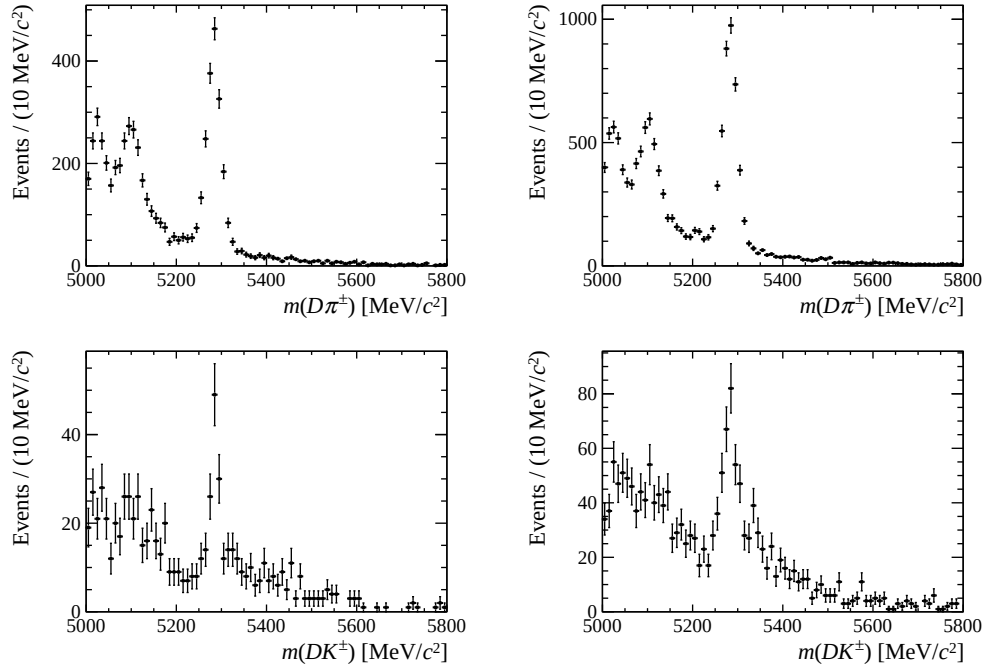


Figure 3.10. The B mass distribution of selected (top) $B \rightarrow D\pi$ and (bottom) $B \rightarrow DK$, $D \rightarrow K_s^0 K^+ K^-$ candidates for (left) LL and (right) DD K_s^0 .

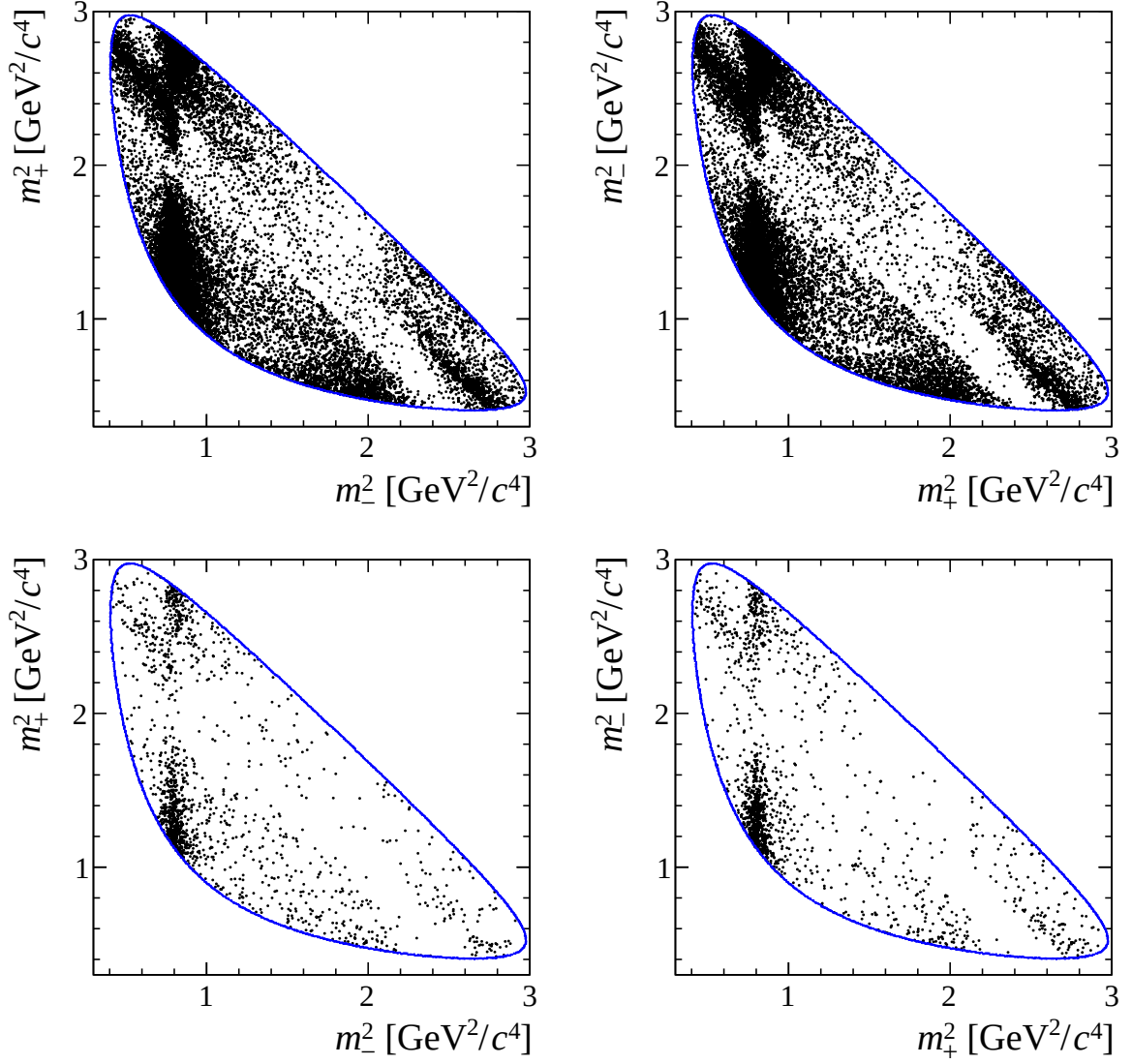


Figure 3.11. Dalitz plots for selected $B^\pm \rightarrow D(\rightarrow K_s^0 \pi^+ \pi^-) h^\pm$ candidates in the signal mass window ($5247 < m(B) < 5317 \text{ MeV}/c^2$). The LL and DD K_s^0 candidates are combined and the plots are separated by bachelor species, (top) $B \rightarrow D\pi$ or (bottom) $B \rightarrow DK$, and bachelor charge, (left) minus or (right) plus.

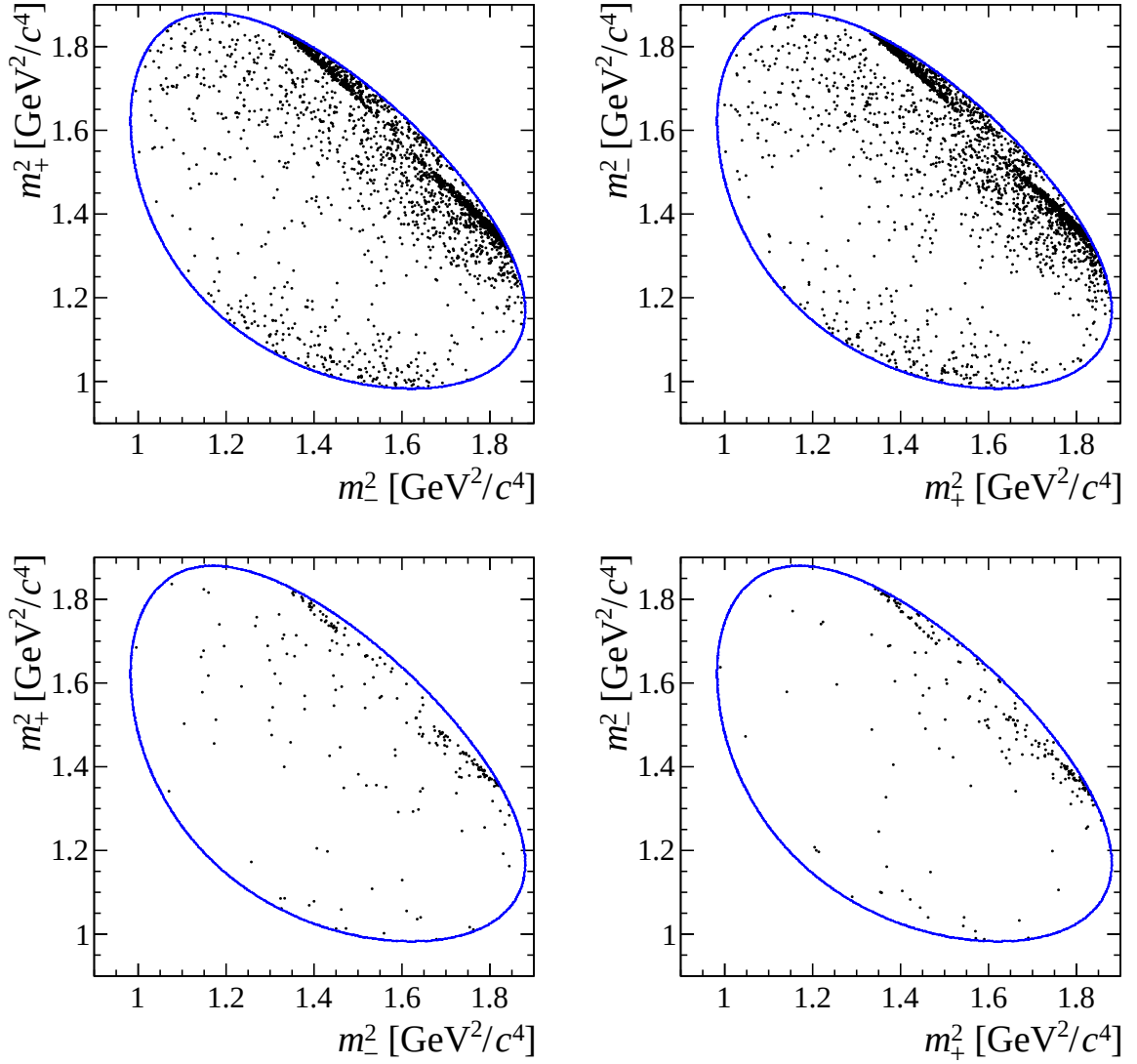


Figure 3.12. Dalitz plots for selected $B^\pm \rightarrow D(\rightarrow K_s^0 K^+ K^-) h^\pm$ candidates in the signal mass window ($5247 < m(B) < 5317 \text{ MeV}/c^2$). The LL and DD K_s^0 candidates are combined and the plots are separated by bachelor species, (top) $B \rightarrow D\pi$ or (bottom) $B \rightarrow DK$, and bachelor charge, (left) minus or (right) plus.

3.2 Mass fit

In this section the fit to the B mass distribution of selected $B \rightarrow DK$ and $B \rightarrow D\pi$ candidates is described. The purpose of this fit is to determine the parameterisation of the signal and backgrounds in preparation for the measurement of the CP violation observables in Chapter 5. The $B \rightarrow D\pi$ sample is fitted because the higher yield allows the shapes and relative yields of the various fit components to be better understood. In addition to the signal decay, three primary sources of background that the selection criteria have not fully eliminated are considered:

1. Combinatorial background, either containing a fake D meson, or when a real D meson is combined with a random bachelor track.
2. Misidentified (misID) decays, where $B \rightarrow D\pi$ candidates are reconstructed as $B \rightarrow DK$, and vice versa.
3. Partially reconstructed background, where one particle has been missed in the reconstruction, such as $B \rightarrow D^{*0}(\rightarrow D\pi^0)K$ decays when the π^0 is not reconstructed.

The shape of each fit component, and constraints on any background yields imposed, is discussed in the following sections. The data are divided into eight categories for the fit: $\{B \rightarrow D\pi, B \rightarrow DK\} \otimes \{K_S^0 \pi^+ \pi^-, K_S^0 K^+ K^-\} \otimes \{\text{LL } K_S^0, \text{DD } K_S^0\}$. The LL and DD K_S^0 categories are separated because of the different reconstructed $m(B)$ resolutions.

3.2.1 Signal

The signal can be described with a Gaussian function with asymmetric tails in order to account for radiative effects. The function is chosen by fitting simulated signal candidates to different modified Gaussian shapes that contain asymmetric tails and the Cruijff function is found to provide the best description. The Cruijff function is

defined as [78]:

$$\text{Cruiff}(m|\mu, \sigma, \alpha_L, \alpha_R) = \begin{cases} \exp\{-(m-\mu)^2/(2\sigma^2 + \alpha_L(m-\mu)^2)\}, & \text{if } m-\mu < 0, \\ \exp\{-(m-\mu)^2/(2\sigma^2 + \alpha_R(m-\mu)^2)\}, & \text{otherwise,} \end{cases} \quad (3.1)$$

where μ is the signal B mass, σ is the signal width, and α_L and α_R are the left-hand and right-hand tail parameters. The fit of a Cruiff function to simulated $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ candidates is shown in Fig. 3.13. The $m(B)$ distribution in MC exhibits small differences to that in data due to imperfect descriptions of the magnetic field, detector alignment and radiative effects, hence the fit parameters are all determined directly in the fit to data, while MC is used to provide information about which parameters can be shared among the different categories. The advantage of sharing parameters is that the $B \rightarrow D\pi$ sample can be used to constrain the shapes in the $B \rightarrow DK$ sample which has lower yield and lower purity. A similar strategy is employed for all the fit components.

It is found that the ratio of signal widths between $B \rightarrow D\pi$ and $B \rightarrow DK$ is 1.06 in MC. This reflects the larger Q -value for the $B \rightarrow D\pi$ decay, which means that the daughter particles have larger momenta, and hence a larger momentum resolution (Fig. 2.8) from the tracking systems. In order to validate the ratio of widths in data, the fit to data reports the signal width in $B \rightarrow DK$ σ_{DK} and the ratio of widths $\sigma_{D\pi}/\sigma_{DK}$ instead of the two signal widths independently. In MC the same $\sigma_{D\pi}/\sigma_{DK}$ ratio is observed in LL and DD K_S^0 candidates, hence a single ratio is measured across the categories.

3.2.2 Combinatorial background

The combinatorial background is primarily formed from random combinations of tracks in the event to form a fake D candidate. Two independent data samples are used to study their $m(B)$ distribution. The first sample contains ‘wrong sign’ candidates, where $D \rightarrow K_S^0 \pi^\pm \pi^\pm$ candidates are reconstructed and are subject to the same selection criteria as the

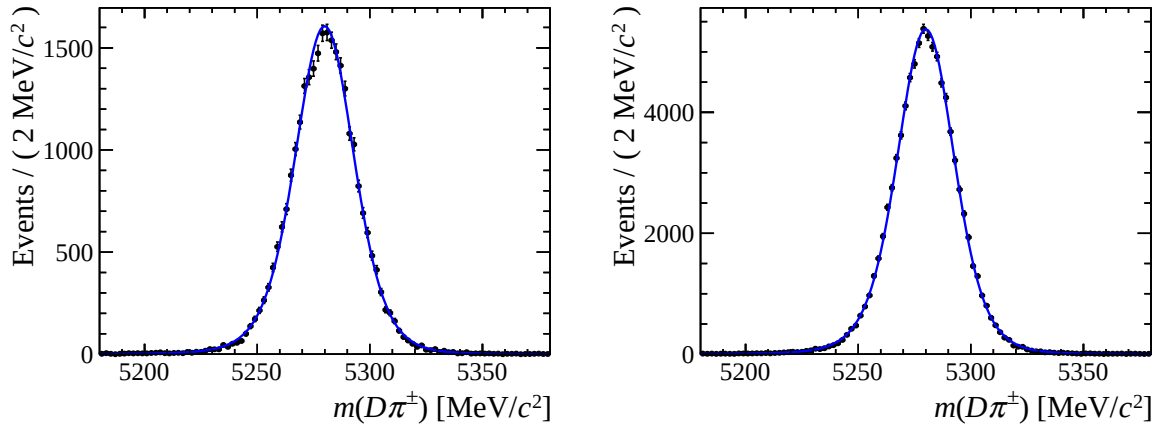


Figure 3.13. The B mass distribution of simulated $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ candidates containing (left) LL and (right) DD K_s^0 . The fit result of a Cruijff function is overlaid.

signal. The second sample consists of $D \rightarrow K_s^0 \pi^+ \pi^-$ events passing the full selection but are in the D and B sidebands ($m(K_s^0 \pi^+ \pi^-) < 1840 \text{ MeV}/c^2$ or $m(K_s^0 \pi^+ \pi^-) > 1890 \text{ MeV}/c^2$, and $m(DK) > 5350 \text{ MeV}/c^2$). Neither of these two data samples are contained in the fitted data. The distributions of both samples are shown in Fig. 3.14 with the results of a fit to an exponential function overlaid. The fitted decay constant τ of the exponential function is $(-3.6 \pm 0.5) \times 10^{-3} \text{ MeV}^{-1} c^2$ in the wrong sign sample and $(-3.7 \pm 0.5) \times 10^{-3} \text{ MeV}^{-1} c^2$ in the sideband sample. As the two measured values of τ are consistent, the first of the two values is used. The value of τ is implemented as a Gaussian constraint in the fit, where the mean and width of the Gaussian are the central value and uncertainty of τ , respectively. The potential contribution from combinations of real D mesons with a random bachelor track is studied by combining a D candidate from one event with a bachelor from another event. The resulting B mass distribution is uniform, and a fit to data with this component included prefers a negligibly small contribution.

3.2.3 Misidentified decays

The $B \rightarrow DK$ and $B \rightarrow D\pi$ categories are distinguished using $\text{DLL}_{K\pi}$ requirements in order to reduce cross-feed between the two decays. However, a small contamination

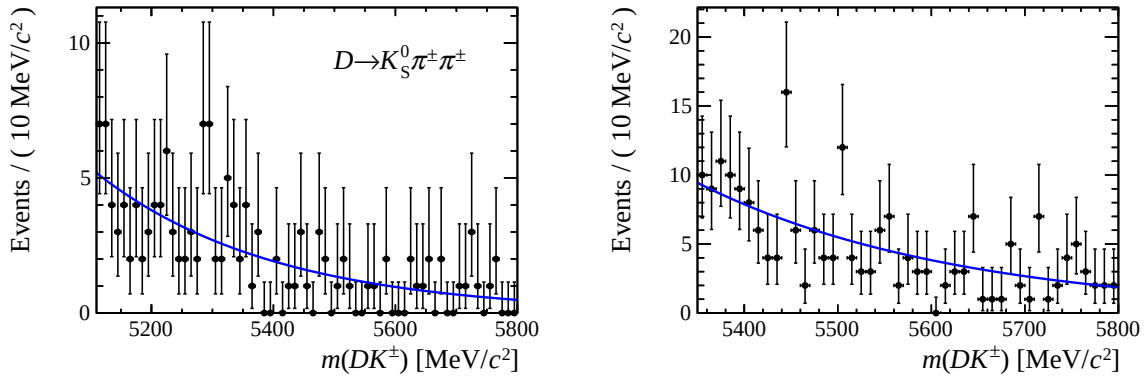


Figure 3.14. Fit of an exponential function to (left) $B \rightarrow DK$, $D \rightarrow K_S^0 \pi^\pm \pi^\pm$ candidates and (right) $B \rightarrow DK$, $D \rightarrow K_S^0 \pi^+ \pi^-$ candidates in the D and B sidebands. The (black) data points are shown with the (blue) fit result overlaid.

of pions misidentified as kaons remains in the $B \rightarrow DK$ sample because the branching fraction of $B \rightarrow D\pi$ is an order of magnitude larger than that of $B \rightarrow DK$. The PID efficiencies (ϵ_{PID}) with which pions are correctly identified or misidentified can be determined for the $\text{DLL}_{K\pi}$ requirements used. Then, the yield of misidentified $\pi \rightarrow K$ candidates can be calculated from the measured $B \rightarrow D\pi$ signal yield, corrected by the PID efficiencies.

The efficiencies cannot be directly determined with MC because the $\text{DLL}_{K\pi}$ variable is not sufficiently well simulated. The RICH detector conditions change over the course of a run and secondary effects such as the loss in photoelectron yields at high RICH detector occupancy because of multiple photons hitting the same HPD pixel are not simulated [58]. Hence, the efficiencies are determined using a calibration tool described in Ref. [79]. A calibration sample consisting of $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+) \pi^+$ decays is used to provide a large and pure sample of kaons and pions. The kaons and pions can be identified without using information from the RICH detector because the D^0 flavour is tagged by the charge of its accompanying pion in the strong D^{*+} decay. Hence, the true D^0 daughter species (kaon or pion) can be determined from its charge (minus or plus). The D^{*+} candidates are selected using kinematic and topological information only. The *sPlot* method [80] is used to perform background subtraction on a two-dimensional

distribution of $m(D^{*+})$ and the mass difference $m(D^{*+}) - m(D^0)$. This subtracts random combinations of D^0 candidates and pions which would cause the flavour of the D^0 to be tagged incorrectly. The efficiency reflects both the response of the RICH gas radiators and the particle identification algorithms. Hence, the performance of a $\text{DLL}_{K\pi}$ requirement depends on the momentum p and pseudorapidity η distributions of the particle to be identified, and the number of tracks in the event, N_{track} . The fraction of correctly identified pions ($\text{DLL}_{K\pi} < 4$) and kaons ($\text{DLL}_{K\pi} > 4$), and misidentified pions ($\text{DLL}_{K\pi} > 4$) and kaons ($\text{DLL}_{K\pi} < 4$) in the calibration sample can be determined in bins of p , η and N_{track} . However, there are differences between the p , η and N_{track} distributions between the D^0 daughters in the calibration sample, and the bachelor particle in the $B \rightarrow Dh$ samples. Therefore, the three variables in the calibration sample are weighted to match the distributions in $B \rightarrow D\pi$, $D \rightarrow K_S^0 h^+ h^-$ MC.

The results are listed in Table 3.3. 96% of bachelor pions are identified correctly using the $\text{DLL}_{K\pi} < 4$ requirement and 4% are misidentified as kaons with $\text{DLL}_{K\pi} > 4$. This means that the yield of $B \rightarrow D\pi$ decays in the $B \rightarrow DK$ sample is approximately $\frac{\epsilon_{\text{PID}}(\text{DLL}_{K\pi} > 4)}{\epsilon_{\text{PID}}(\text{DLL}_{K\pi} < 4)} = \frac{0.04}{0.96} = 4\%$ of the yield of $B \rightarrow D\pi$ signal decays in the $B \rightarrow D\pi$ sample. The $B \rightarrow D\pi$ contamination in the $B \rightarrow DK$ sample can clearly be seen as the enhanced slope to the right of the B signal peak in Fig. 3.9. Similarly, 84% of bachelor kaons are identified correctly using the $\text{DLL}_{K\pi} > 4$ requirement while 16% are misidentified as pions. The yield of $B \rightarrow DK$ decays in the $B \rightarrow D\pi$ sample is approximately $\frac{\epsilon_{\text{PID}}(\text{DLL}_{K\pi} < 4)}{\epsilon_{\text{PID}}(\text{DLL}_{K\pi} > 4)} = \frac{0.16}{0.84} = 20\%$ of the yield of $B \rightarrow DK$ signal decays in the $B \rightarrow DK$ sample. A crude estimation using the sizes of the signal peaks in Fig. 3.9 indicates that the actual level of $B \rightarrow DK$ decays in the $B \rightarrow D\pi$ sample is negligible because of the much smaller $B \rightarrow DK$ branching ratio.

The $m(B)$ distribution of $B \rightarrow D\pi$ candidates reconstructed as $B \rightarrow DK$ is now determined. A simple approach is to apply the $\text{DLL}_{K\pi} > 4$ requirement to simulated $B \rightarrow D\pi$ candidates, assign a kaon mass hypothesis to the bachelor and calculate the

Table 3.3. Particle identification efficiencies ϵ_{PID} for various $\text{DLL}_{K\pi}$ requirements on the bachelor pion and kaon in the $B \rightarrow Dh$ data samples.

Particle	PID requirement	D final state	ϵ_{PID} (%)	
			LL	DD
Kaon	$\text{DLL}_{K\pi} > 4$	$K_S^0 \pi^+ \pi^-$	83.34 ± 0.17	83.69 ± 0.12
		$K_S^0 K^+ K^-$	82.40 ± 0.49	82.12 ± 0.33
Pion	$\text{DLL}_{K\pi} < 4$	$K_S^0 \pi^+ \pi^-$	96.25 ± 0.01	96.21 ± 0.01
		$K_S^0 K^+ K^-$	96.27 ± 0.05	96.24 ± 0.03
Kaon	$\text{DLL}_{K\pi} < 4$	$K_S^0 \pi^+ \pi^-$	16.65 ± 0.17	16.30 ± 0.12
		$K_S^0 K^+ K^-$	17.59 ± 0.49	17.87 ± 0.33
Pion	$\text{DLL}_{K\pi} > 4$	$K_S^0 \pi^+ \pi^-$	3.74 ± 0.01	3.78 ± 0.01
		$K_S^0 K^+ K^-$	3.72 ± 0.05	3.75 ± 0.03

resulting B mass. However, the MC-data differences in the $\text{DLL}_{K\pi}$ variable means that an inaccurate shape will be obtained. Therefore, a data-driven method is used to determine the distribution and is outlined as follows:

1. The selection described in Sect. 3.1 is applied to $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decays, including bachelor $\text{DLL}_{K\pi} < 4$, to provide a sample of $B \rightarrow D\pi$ decays that contains a negligible contamination of $B \rightarrow DK$.
2. The B mass distribution of the $B \rightarrow D\pi$ sample is fitted using the shapes for the signal, combinatorial background and partially reconstructed background described in this section. Using the results of this fit, the *sPlot* method [80] is used to obtain per-event signal weights w_s .
3. Using the method and calibration sample previously described, the efficiencies of pions passing the $\text{DLL}_{K\pi} < 4$ and $\text{DLL}_{K\pi} > 4$ cuts in bins of p and η are obtained are used to obtain per-event signal weights w_π and w_K , respectively.
4. The $m(B)$ distribution of misidentified decays is determined by assigning a kaon mass hypothesis to the bachelor of each $B \rightarrow D\pi$ candidate in the sample.

5. The $m(B)$ distribution is weighted by an overall event weight equal to $\frac{w_s \times w_K}{w_\pi}$.

The weighting by w_s is equivalent to performing a background subtraction. The factor $1/w_\pi$ recovers the bachelor kinematic distributions before the $\text{DLL}_{K\pi} < 4$ requirement was applied, while the weights w_K are applied to obtain the bachelor kinematics of $B \rightarrow D\pi$ decays misidentified as $B \rightarrow DK$ in the selected data.

The resulting $m(DK^\pm)$ distribution is fitted with the sum of two Crystal Ball functions for LL K_s^0 candidates and the sum of a Crystal Ball and a Gaussian function for DD K_s^0 candidates. The Crystal Ball function is defined [81]:

$$\text{CB}(m|\mu, \sigma, \alpha, n) = \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2}, & \text{if } \frac{m-\mu}{\sigma} \geq -|\alpha|, \\ e^{-|\alpha|^2/2} \left(\frac{n}{|\alpha|}\right)^n \left(\frac{n}{|\alpha|} - |\alpha| - \frac{m-\mu}{\sigma}\right)^{-n}, & \text{otherwise,} \end{cases} \quad (3.2)$$

where μ and σ are mean and width parameters, and α and n are tail parameters. The Gaussian function is denoted $G(m|\mu, \sigma)$, where μ and σ are the mean and width. The fit results are shown in Fig. 3.15 and the fit parameters are given in Table 3.4. The shapes obtained using this procedure are subsequently fixed in the mass fit. This is necessary as there is a significant overlap between the signal shape and the left-hand tail, and between the combinatorial shape and the right-hand tail. The functions obtained if these background shapes are allowed to vary in the fit to $B \rightarrow Dh$ candidates is superimposed for comparison and a difference in the right-hand tail can be seen. This reinforces the need for an independent data-driven method to determine the shape of this distribution.

3.2.4 Partially reconstructed decays

Decays where one or more particles has failed to be reconstructed populate the mass region below the signal B mass. The B mass range in the fit is restricted to include only

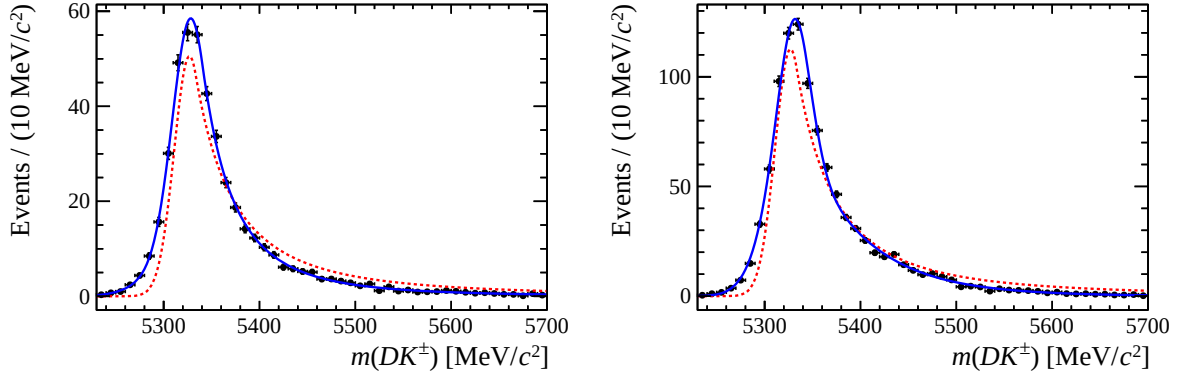


Figure 3.15. The $m(DK^\pm)$ distributions of $B \rightarrow D(\rightarrow K_s^0 \pi^+ \pi^-) \pi$ candidates, (left) LL and (right) DD K_s^0 , where the bachelor has been assigned the kaon mass hypothesis and the per-event weights $\frac{w_s \times w_K}{w_\pi}$ described in the main text have been applied. The (solid blue) fit results are overlaid. The (dashed red) distribution is obtained if the shapes were allowed to vary in the fit to $B \rightarrow Dh$ candidates.

Table 3.4. Shape parameters for $B \rightarrow D\pi$ decays misidentified as $B \rightarrow DK$. For LL K_s^0 the shape is a sum of two Crystal Ball functions given by $f_1 \times CB(m|\mu, \sigma_1, \alpha, n) + (1 - f_1) \times CB(m|\mu, \sigma_2, \alpha, n)$. For DD K_s^0 the shape is a sum of a Crystal Ball and Gaussian given by $f_1 \times CB(m|\mu, \sigma_1, \alpha, n) + (1 - f_1) \times G(m|\mu, \sigma_2)$.

Parameter	LL	DD
μ	5328.3	5329.6
σ_1	30.5	20.8
σ_2	18.0	20.6
α	-0.66	-0.33
n	3.45	28.1
f_1	0.37	0.77

the case where one particle has been lost. The decays to be considered and the analytic functions used to model them are based on a previous study of $B \rightarrow Dh$, $D \rightarrow h^+ h^-$ data, which also employs a similar simultaneous fit of $B \rightarrow D\pi$ and $B \rightarrow DK$ decays with over a factor of five larger yields [82]. It is assumed that the shape and composition of this background does not depend on the D final state, hence the same functions and shape parameters are used here.

The decays considered are listed below, where the particle that was not reconstructed is identified in square brackets. These decays have been identified in simulation to

be the dominant contributors to this background. For $B \rightarrow D\pi$ candidates, the decays considered are:

- $B^\pm \rightarrow D^{*0}(\rightarrow D^0[\pi^0])\pi^\pm,$
- $B^\pm \rightarrow D^{*0}(\rightarrow D^0[\gamma])\pi^\pm,$
- $B^0 \rightarrow D^{*\pm}(\rightarrow D^0[\pi^\pm])\pi^\mp,$
- $B^0 \rightarrow D^0\rho(\rightarrow [\pi^\pm]\pi^\mp).$

For $B \rightarrow DK$ candidates, these are:

- $B^\pm \rightarrow D^{*0}(\rightarrow D^0[\pi^0])K^\pm,$
- $B^\pm \rightarrow D^{*0}(\rightarrow D^0[\gamma])K^\pm,$
- $B^0 \rightarrow D^{*\pm}(\rightarrow D^0[\pi^\pm])K^\mp,$
- $B^\pm \rightarrow D^0K^{*\pm}(\rightarrow [\pi^0]K^\pm),$

as well as the four background decays considered for $B \rightarrow D\pi$, but where the bachelor pion has been misidentified as a kaon. The compatibility of the shape obtained from these backgrounds compared to a more comprehensive list is considered later in this section. The $B_s^0 \rightarrow D\bar{K}^{*0}$ decay is found to have a negligible fit fraction but is considered as a systematic uncertainty in Sect. 5.3.

The functions describing the backgrounds where the bachelor particle has not been misidentified as first considered. The simulated $m(D\pi)$ distributions of such decays are shown in Fig. 3.16. Two functions were developed by Paolo Gandini to parameterise these backgrounds and are also used in Refs. [82, 83]. The two functions are known as the *Horns* and *Hill* shapes, and are used when the missing particle has spin 0 ($\pi^0, \pi^\pm$) or spin 1 (γ), respectively.

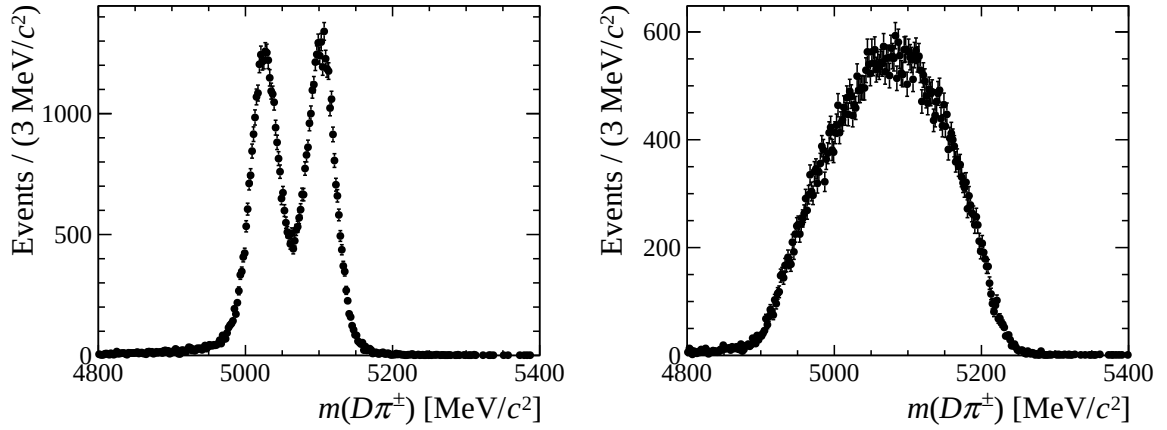


Figure 3.16. $B^\pm \rightarrow D^{*0}\pi^\pm$ MC events, with (left) $D^{*0} \rightarrow D^0\pi^0$ or (right) $D^{*0} \rightarrow D^0\gamma$, reconstructed as $B \rightarrow D\pi$ where the (left) π^0 or (right) γ is missed. The MC contains candidates in the $D \rightarrow h^+h^-$ final state.

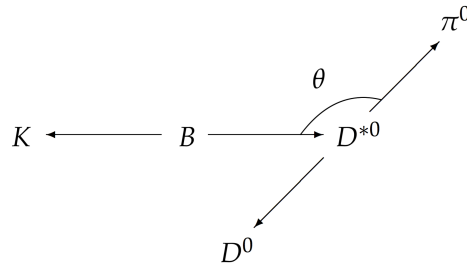


Figure 3.17. The $B \rightarrow D^{*0}(\rightarrow D^0\pi^0)K$ decay is illustrated, where the D^{*0} and K momentum vectors are drawn in the B rest frame, and the D^0 and π^0 momentum vectors are drawn in the D^{*0} rest frame. The angle θ is the helicity angle used in this section.

Horns function

Consider the decay of $B^- \rightarrow D^{*0}\pi^-$ with $D^{*0} \rightarrow D^0\pi^0$, where the π^0 helicity angle, θ , is defined as the angle between the π^0 momentum vector in the D^{*0} rest frame and the D^{*0} boost vector in the B rest frame, as illustrated in Fig. 3.17. The helicity of the D^{*0} and conservation of angular momentum mean that the spin-0 π^0 meson will be emitted predominantly along $\theta = 0^\circ$ or $\theta = 180^\circ$. When $\theta = 0^\circ$, the fraction of momentum carried by the π^0 in the B rest frame is at its smallest, leading to a larger reconstructed $m(D\pi^\pm)$. When $\theta = 180^\circ$ the converse occurs, and the momentum carried by the π^0 is

greatest, leading to a lower $m(D\pi^\pm)$. Thus the helicity angle distribution of the missing particle has a one-to-one relation to its momentum and therefore to the reconstructed B mass. This gives rise to the double-peak structure seen in Fig. 3.16. The kinematic endpoints of the distributions are fully defined by the particle masses in the decay, hence the resulting B mass distribution is a parabola $p(x)$ with kinematic endpoints a and b , where

$$p_{\text{HORNS}}(x) = \begin{cases} (x - \frac{a+b}{2})^2, & \text{if } a < x < b \\ 0, & \text{otherwise.} \end{cases} \quad (3.3)$$

The parabola is convolved with a sum of two Gaussian functions to account for resolution effects. For a Gaussian function $G(x|\mu, \sigma)$ with mean μ and width σ , the ‘double Gaussian’ is expressed as

$$\text{DG}(x) = f_G G(x|\mu, \sigma) + (1 - f_G) G(x|\mu, R_\sigma \sigma), \quad (3.4)$$

where σ is the width of the first Gaussian, f_G is the fraction contained by the first Gaussian and R_σ is the relative width between the two. Furthermore, selection effects can distort the Horns shape such that the right-hand peak is higher than the other, since the kinematics of candidates whose B mass is closer to the signal are favoured. This is taken into account by introducing a linear polynomial with a slope of $1 - \xi$, where $0 < \xi < 1$. As $\xi \rightarrow 0$, the left-hand peak decreases in size relative to the right hand peak. The resulting Horns function is

$$\text{Horns}(m) = \int_a^b dx \left(x - \frac{a+b}{2} \right)^2 \text{DG}(x|m, \sigma, f_G, R_\sigma) \left(\frac{1 - \xi_{\text{HORNS}}}{b - a} x + \frac{b \xi_{\text{HORNS}} - a}{b - a} \right), \quad (3.5)$$

where m is the mass variable to be fitted and x is the integration variable in the convolution.

Hill function

The Hill shape considers the case where a spin-1 particle is missed. Consider the case $B^- \rightarrow D^{*0}\pi^-$ with $D^{*0} \rightarrow D^0\gamma$, where θ is the photon helicity angle. The photon will decay preferentially in the $\theta = 90^\circ, 270^\circ$ directions, where the fraction of momentum carried by the photon is the same in these two directions and so a double-peak structure is not seen. The resulting $m(D\pi^\pm)$ distribution (Fig. 3.16) is a parabola with negative curvature and kinematic endpoints a, b described by

$$p_{\text{HILL}}(x) = \begin{cases} -(x-a)(x-b), & \text{if } a < x < b \\ 0, & \text{otherwise.} \end{cases} \quad (3.6)$$

As with the Horns function, p_{HILL} is convolved with a double Gaussian to describe the resolution and a linear polynomial to account for selection effects, giving

$$\text{Hill}(m) = \int_a^b dx [-(x-a)(x-b)] \text{DG}(x|m, \sigma, f_G, R_\sigma) \left(\frac{1 - \xi_{\text{HILL}}}{b-a} x + \frac{b\xi_{\text{HILL}} - a}{b-a} \right). \quad (3.7)$$

Misidentified decays

The partially reconstructed decays contributing to the $B \rightarrow D\pi$ sample can be reconstructed as $B \rightarrow DK$ when the bachelor pion is misidentified as a kaon. An approximate description for these decays is constructed in a similar fashion to the Horns and Hill functions. The shapes p_{HORNS} and p_{HILL} are each convolved with a sum of four Gaussians in order to model the asymmetric shape of misID decays seen in Sect. 3.2.3. A sum of four Gaussians is used to approximate the shape because the Crystal Ball function is not analytically integrable. The four Gaussians have different means ($\mu_1, \mu_2, \mu_3, \mu_4$) and widths ($\sigma_1, \sigma_2, \sigma_3, \sigma_4$) and are summed with recursive fractions (f_1, f_2, f_3). The parameters are fixed according to the results of MC studies and are listed later.

Implementation

The Horns function is used to describe $B^\pm \rightarrow (D^{*0} \rightarrow D^0[\pi^0])h^\pm$, $B^0 \rightarrow (D^{*\pm} \rightarrow D^0[\pi^\pm])h^\mp$, $B^0 \rightarrow D^0(\rho \rightarrow [\pi^\pm]\pi^\mp)$ and $B^\pm \rightarrow D^0(K^{*\pm} \rightarrow [\pi^0]K^\pm)$ decays. The Hill function is used to describe $B^\pm \rightarrow (D^{*0} \rightarrow D^0[\gamma])h^\pm$ decays.

The lower mass range of the fit is chosen to be $5080 \text{ MeV}/c^2$ so that the right-hand peak of the double-peaked structure is present, allowing the yield of the partially reconstructed background to be well estimated. Most of the parameters determining the shapes of the Horns and Hill functions are fixed to the results of fitting these shapes to $B^\pm \rightarrow Dh^\pm$, $D \rightarrow h^+h^-$ data in Ref. [82] and are listed in Table 3.5. These include the resolution parameters f_G and R_σ , which are common to all of the components, and ξ_{HILL} and ξ_{HORNS} which are common to all of the Hill and Horns shapes, respectively. Due to the width of the ρ meson in $B^0 \rightarrow D^0\rho^0$ decays, the left-hand peak extends far below the range in which these shapes are studied and ξ_{HORNS} is set to zero for this decay only. The endpoints a and b are calculated using the masses of the particles involved in each decay to obtain the minimum and maximum reconstructed B masses. The endpoints are fixed in the fit, and a further term δm is added to allow an overall translation of the full shape. This accommodates both residual imperfections in the parameterisation of this background and in the calibration of the momentum measurements. (The latter leads to a systematic shift in the B mass and is reflected in the fit results later in this section, which show that the mean signal $B^\pm$ mass is $\sim 3 \text{ MeV}/c^2$ higher than the known $B^\pm$ mass.) The width σ of the resolution function is also allowed to vary separately for K_S^0 LL and DD in $B \rightarrow D\pi$. The ratio of widths $\sigma_{D\pi}/\sigma_{DK}$ is fixed to that found in the fit to $D \rightarrow h^+h^-$ data, and is consistent with the expected value of 1.06 from $D \rightarrow K_S^0 h^+h^-$ MC.

The contributions from the different decay channels forming the partially reconstructed background in $B \rightarrow D\pi$ are summed with three recursive fractions $f_{D^*}^\pi, f_{D^{*0}}, f_\gamma$

as follows.

$$\begin{aligned}
 \mathcal{P}(B \rightarrow D^{*0}\pi) &= f_\gamma \mathcal{P}(B \rightarrow D^{*0}(\rightarrow D^0\gamma)\pi) & + (1 - f_\gamma) \mathcal{P}(B \rightarrow D^{*0}(\rightarrow D^0\pi^0)\pi), \\
 \mathcal{P}(B \rightarrow D^*\pi) &= f_{D^{*0}} \mathcal{P}(B \rightarrow D^{*0}\pi) & + (1 - f_{D^{*0}}) \mathcal{P}(B^0 \rightarrow D^{*\pm}\pi^\mp), \\
 \mathcal{P}_{\text{total}} &= f_{D^*}^\pi \mathcal{P}(B \rightarrow D^*\pi) & + (1 - f_{D^*}^\pi) \mathcal{P}(B \rightarrow D\rho),
 \end{aligned}$$

where $\mathcal{P}(B \rightarrow X)$ indicates the function describing the decay $B \rightarrow X$ and $\mathcal{P}_{\text{total}}$ is the overall shape. For $B \rightarrow DK$ a similar sum is constructed with three recursive fractions $f_{D^*}^K, f_{D^{*0}}, f_\gamma$, where

$$\mathcal{P}_{\text{total}} = f_{D^*}^K \mathcal{P}(B \rightarrow D^*K) + (1 - f_{D^*}^K) \mathcal{P}(B \rightarrow DK^*),$$

and the values of $f_{D^{*0}}$ and f_γ are the same as those for $B \rightarrow D\pi$. The values of the recursive fractions are obtained from the fit to $D \rightarrow h^+h^-$ data and are listed in Table 3.5. The overall fraction of partially reconstructed decays which are misidentified is $f_{\text{low,MISID}}$ and is allowed to vary.

Cross-check of the use of parameters obtained from the fit to $D \rightarrow h^+h^-$ data

The assumption that the same shape parameters as those used in the $D \rightarrow h^+h^-$ analysis can be used here is cross-checked. The following decays are generated with $D \rightarrow K_s^0\pi^+\pi^-$: $B^0 \rightarrow D^0\rho^0$, $B^0 \rightarrow D^{*-}K^+$, $B^0 \rightarrow D^{*-}\pi^+$, $B^0 \rightarrow D^{*-}\rho^+$, $B^+ \rightarrow D^0\rho^+$, $B^+ \rightarrow D^{*0}K^+$, $B^+ \rightarrow D^{*0}\pi^+$, $B^+ \rightarrow D^{*0}\rho^+$, $B_s^0 \rightarrow D^0K^{*0}$, $B_s^0 \rightarrow D^{*0}(D^0\pi^0)K^{*0}$, $B_s^0 \rightarrow D^{*0}(D^0\gamma)K^{*0}$, $\Lambda_b^0 \rightarrow D^0pK$, $\Lambda_b^0 \rightarrow D^0p\pi$ and $B^0 \rightarrow D^0K^{*0}$. Each of these decays is reconstructed as $B \rightarrow D\pi$ and $B \rightarrow DK$, and combined to form a cocktail according to their known production fractions and branching fractions, and the reconstruction and selection efficiencies determined using the simulated data. The $m(B)$ distribution of the cocktail is fitted in

Table 3.5. Fixed parameters for the partially reconstructed shapes which have been obtained from studies of $B^\pm \rightarrow Dh^\pm, D \rightarrow h^+h^-$ decays [82]. The parameters are defined in the main text. The parameters a and b refer to the lower and upper endpoints, respectively, of the Horns or Hill functions used for the decay specified.

Parameter	Value	Parameter	Value
$a(B^\pm \rightarrow D^0 K^{*\pm})$	3761.8	$\delta m(DK)$	0.70
$a(B^0 \rightarrow D^{*\pm}(\rightarrow D^0 \pi^\pm) K^\mp)$	5014.8	$\sigma_{D\pi}/\sigma_{DK}$	1.08
$a(B^0 \rightarrow D^{*\pm}(\rightarrow D^0 \pi^\pm) \pi^\mp)$	5012.6	$f_{1,\text{HILL}}$	0.63
$a(B^\pm \rightarrow D^0 \rho^0)$	3761.8	$f_{1,\text{HORNS}}$	0.50
$a(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \gamma) K^\pm)$	4909.2	$f_{2,\text{HILL}}$	0.20
$a(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \pi^0) K^\pm)$	5015.0	$f_{2,\text{HORNS}}$	0.36
$a(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \gamma) \pi^\pm)$	4905.6	$f_{3,\text{HILL}}$	0.032
$a(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \pi^0) \pi^\pm)$	5012.8	$f_{3,\text{HORNS}}$	0.14
$b(B^\pm \rightarrow D^0 K^{*\pm})$	5136.4	$\mu_{1,\text{HILL}}$	47.13
$b(B^0 \rightarrow D^{*\pm}(\rightarrow D^0 \pi^\pm) K^\mp)$	5105.8	$\mu_{1,\text{HORNS}}$	44.79
$b(B^0 \rightarrow D^{*\pm}(\rightarrow D^0 \pi^\pm) \pi^\mp)$	5104.8	$\mu_{2,\text{HILL}}$	103.08
$b(B^\pm \rightarrow D^0 \rho^0)$	5137.4	$\mu_{2,\text{HORNS}}$	88.92
$b(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \gamma) K^\pm)$	5226.2	$\mu_{3,\text{HILL}}$	322.36
$b(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \pi^0) K^\pm)$	5114.4	$\mu_{3,\text{HORNS}}$	169.69
$b(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \gamma) \pi^\pm)$	5226.7	$\mu_{4,\text{HILL}}$	159.85
$b(B^\pm \rightarrow D^{*0}(\rightarrow D^0 \pi^0) \pi^\pm)$	5108.5	$\mu_{4,\text{HORNS}}$	30.00
ξ_{HILL}	0.40	$\sigma_{1,\text{HILL}}$	12.29
ξ_{HORNS}	0.71	$\sigma_{1,\text{HORNS}}$	17.46
$\xi_{\text{HORNS},D\rho}$	0.00	$\sigma_{2,\text{HILL}}$	9.00
f_γ	0.37	$\sigma_{2,\text{HORNS}}$	39.92
$f_{D^{*0}}$	0.76	$\sigma_{3,\text{HILL}}$	70.82
$f_{D^*}^\pi$	0.76	$\sigma_{3,\text{HORNS}}$	126.53
$f_{D^*}^K$	1.00	$\sigma_{4,\text{HILL}}$	50.00
f_G	0.92	$\sigma_{4,\text{HORNS}}$	80.00
R_σ	2.44		

the range $5000 < m(B) < 5500 \text{ MeV}/c^2$ using the parameters fixed to the values listed in Table 3.5, apart from ξ_{HORNS} which is allowed to float. The results are shown in Fig. 3.18 and demonstrate that these parameters provide a good fit. The value of ξ_{HORNS} when fitted to the cocktail MC is 1.0 ± 0.2 , in comparison to a value of 0.71 (Table 3.5)

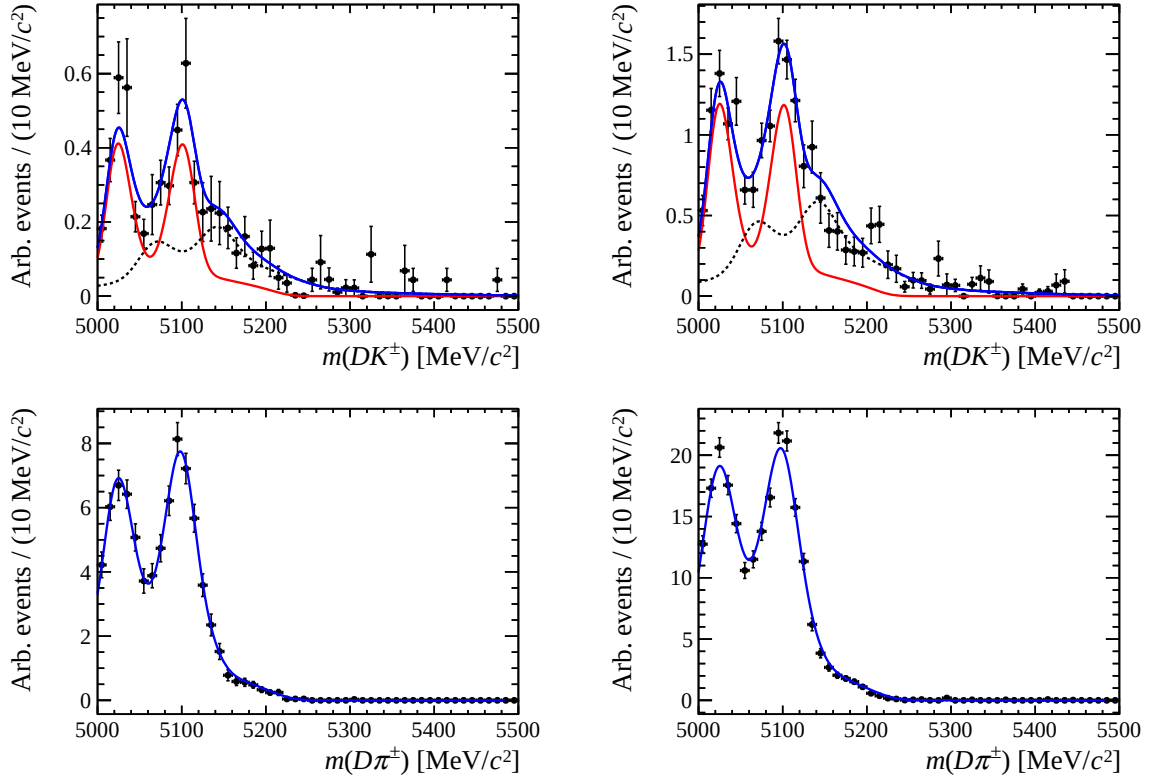


Figure 3.18. The $m(Dh^\pm)$ distribution of a cocktail of simulated background decays reconstructed as (top) $B \rightarrow DK$ and (bottom) $B \rightarrow D\pi$, (left) LL and (right) DD K_s^0 . The list of backgrounds is given in the main text. The y -axis scale is arbitrary and the error bars reflect the errors from the limited number of simulated events. The (blue) fit result is superimposed. The components in $B \rightarrow DK$ describing (red) real kaons and (dashed black) misidentified kaons are shown.

measured in the $D \rightarrow h^+h^-$ data. The difference in the two ξ_{HORNS} values can be ignored since it only affects the relative height of the left-hand peak, which is below the lower limit of $5080 \text{ MeV}/c^2$ used in the final fit.

3.2.5 Results

An extended maximum likelihood fit of the $B^\pm$ mass of the selected candidates is performed in the range 5080 to $5800 \text{ MeV}/c^2$. The values of the shape parameters are summarised in Table 3.6. The value of $\sigma_{D\pi}/\sigma_{DK}$ is measured to be 1.08 ± 0.03 which is consistent with the expectation of 1.06 from MC. Tables 3.7 and 3.8 list the fitted

yields of the signal and each background component integrated in the signal region ($5247 < m(B) < 5317 \text{ MeV}/c^2$). The uncertainties of the yields reflect the constraints on the signal and misID background yields through ϵ_{PID} (Sect. 3.2.3), and the fractional uncertainties of the yields are those measured in the full fit range. Approximately 2257 $K_s^0 \pi^+ \pi^-$ signal events and 324 $K_s^0 K^+ K^-$ signal events are found when the yields in the DD and LL K_s^0 categories are summed. The $m(B)$ distribution of the data and the results of the fit are shown in Figs. 3.19 – 3.22.

Table 3.6. Fit results for the $B^\pm$ mass fit to $B^\pm \rightarrow Dh^\pm$, $D \rightarrow K_s^0 h^+ h^-$ decays.

Parameter	Value	
	LL	DD
Signal (Cruijff)		
μ [MeV/ c^2]	5282.7 ± 0.1	
$\sigma_{D\pi}/\sigma_{DK}$	1.08 ± 0.03	
σ_{DK} [MeV/ c^2]	13.1 ± 0.3	13.3 ± 0.3
α_L	0.135 ± 0.008	0.109 ± 0.008
α_R	0.078 ± 0.009	0.084 ± 0.006
Combinatorial (Exponential)		
$\tau(B^\pm \rightarrow D(\rightarrow K_s^0 \pi^+ \pi^-) \pi^\pm)$ [GeV $^{-1} c^2$]	-4.0 ± 0.4	-3.5 ± 0.2
$\tau(B^\pm \rightarrow D(\rightarrow K_s^0 \pi^+ \pi^-) K^\pm)$ [GeV $^{-1} c^2$]	-3.6 ± 0.4	-2.9 ± 0.3
$\tau(B^\pm \rightarrow D(\rightarrow K_s^0 K^+ K^-) \pi^\pm)$ [GeV $^{-1} c^2$]	-4.3 ± 0.4	-4.4 ± 0.3
$\tau(B^\pm \rightarrow D(\rightarrow K_s^0 K^+ K^-) K^\pm)$ [GeV $^{-1} c^2$]	-3.7 ± 0.5	-4.1 ± 0.5
Partially reconstructed		
$f_{\text{low,MISID}}$	0.42 ± 0.02	
$\delta m_{D\pi}$ [MeV/ c^2]	5.2 ± 0.3	
$\sigma_{D\pi}$ [MeV/ c^2]	14.7 ± 0.5	15.1 ± 0.4

Table 3.7. The $D \rightarrow K_s^0 \pi^+ \pi^-$ yields of each fit component integrated in the signal region ($5247 < m(B) < 5317 \text{ MeV}/c^2$).

Fit component	$B^\pm \rightarrow DK^\pm$ selection		$B^\pm \rightarrow D\pi^\pm$ selection	
	LL	DD	LL	DD
$B^\pm \rightarrow DK^\pm$	702 ± 18	1555 ± 39	30 ± 5	64 ± 7
$B^\pm \rightarrow D\pi^\pm$	87 ± 9	164 ± 13	10338 ± 106	22779 ± 166
Combinatorial	59 ± 9	133 ± 14	103 ± 11	433 ± 25
Partially reconstructed	38 ± 2	82 ± 3	4.6 ± 0.1	14.2 ± 0.1

Table 3.8. The $D \rightarrow K_s^0 K^+ K^-$ yields of each fit component integrated in the signal region ($5247 < m(B) < 5317 \text{ MeV}/c^2$).

Fit component	$B^\pm \rightarrow DK^\pm$ selection		$B^\pm \rightarrow D\pi^\pm$ selection	
	LL	DD	LL	DD
$B^\pm \rightarrow DK^\pm$	101 ± 4	223 ± 7	4.5 ± 1.9	10.1 ± 2.9
$B^\pm \rightarrow D\pi^\pm$	13 ± 3	24 ± 5	1501 ± 38	3338 ± 57
Combinatorial	13 ± 3	30 ± 5	36 ± 5	78 ± 7
Partially reconstructed	4.6 ± 0.7	8.6 ± 1.2	0.60 ± 0.02	2.0 ± 0.1

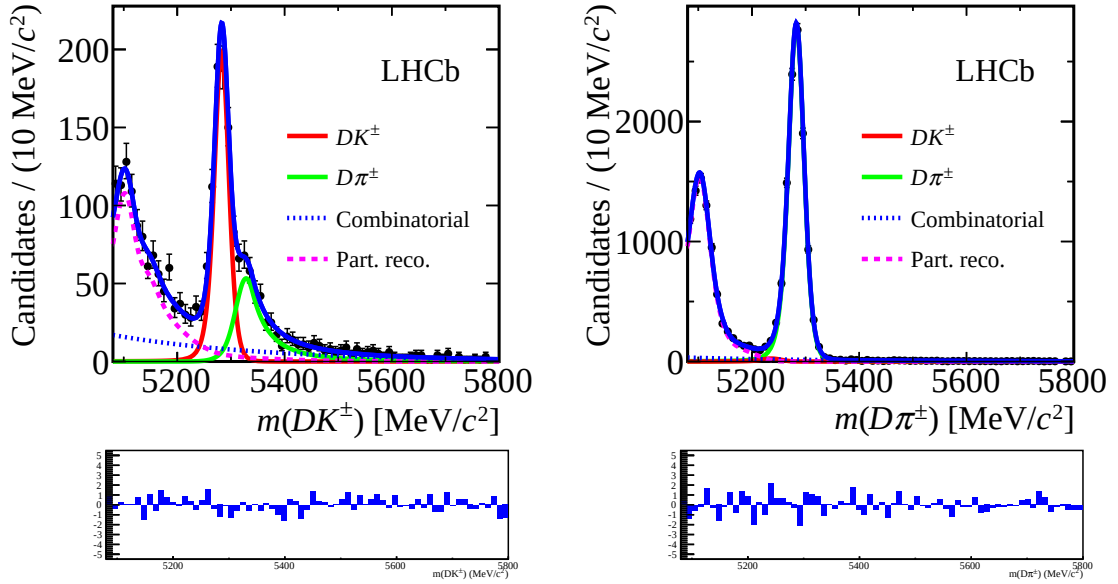


Figure 3.19. Invariant mass distributions of (left) $B^\pm \rightarrow DK^\pm$ and (right) $B^\pm \rightarrow D\pi^\pm$ candidates, with $D \rightarrow K_s^0 \pi^+ \pi^-$ in the LL K_s^0 category. Fit results, including the signal and background components, are superimposed and described in the legend. The associated pull distribution is shown below.

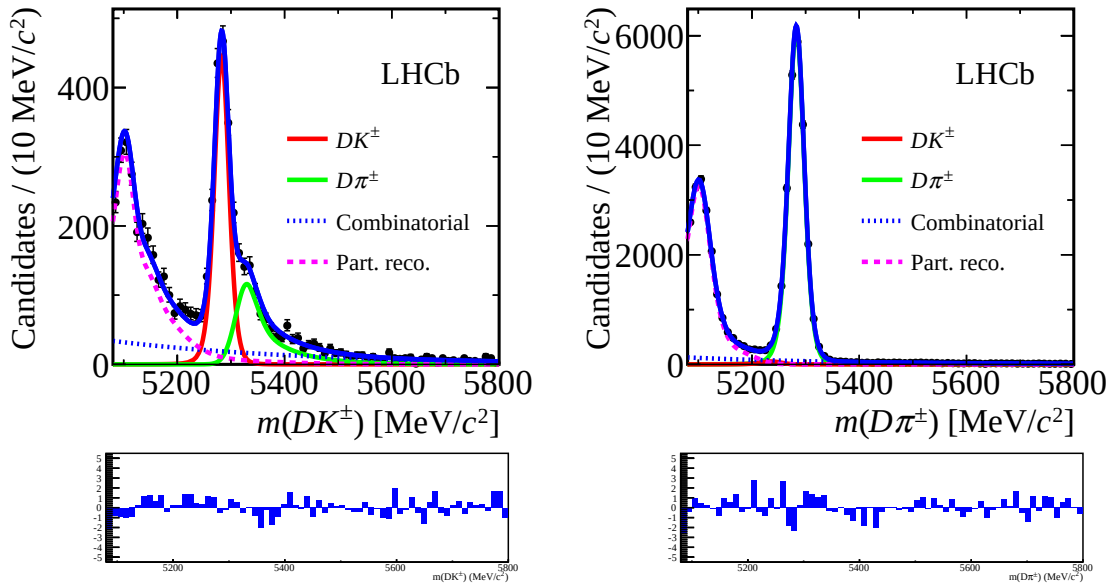


Figure 3.20. Invariant mass distributions of (left) $B^\pm \rightarrow DK^\pm$ and (right) $B^\pm \rightarrow D\pi^\pm$ candidates, with $D \rightarrow K_s^0 \pi^+ \pi^-$ in the DD K_s^0 category. Fit results, including the signal and background components, are superimposed and described in the legend. The associated pull distribution is shown below.

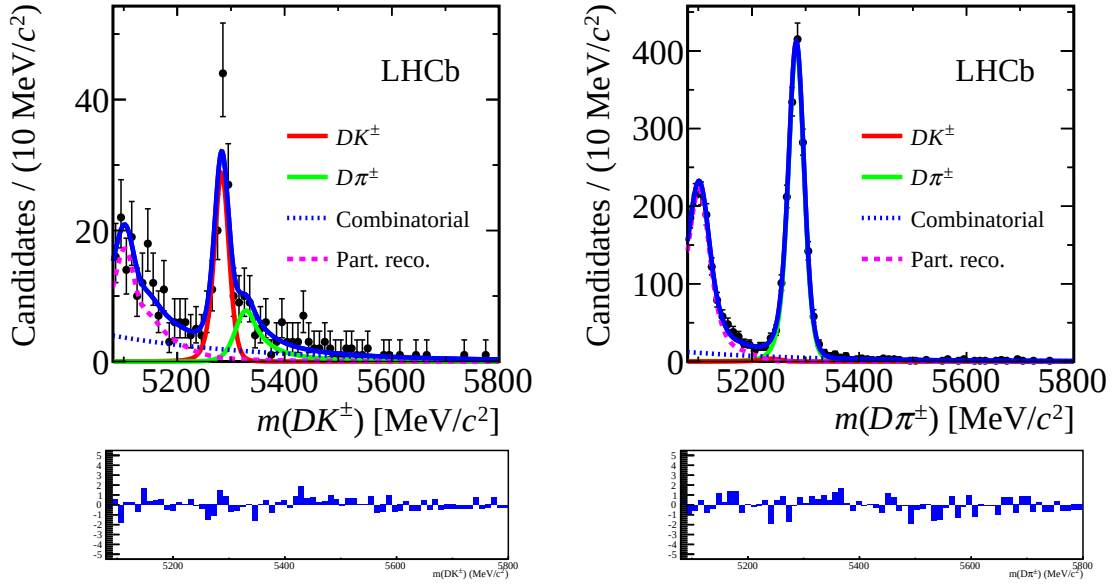


Figure 3.21. Invariant mass distributions of (left) $B^\pm \rightarrow DK^\pm$ and (right) $B^\pm \rightarrow D\pi^\pm$ candidates, with $D \rightarrow K_S^0 K^+ K^-$ in the LL K_S^0 category. Fit results, including the signal and background components, are superimposed and described in the legend. The associated pull distribution is shown below.

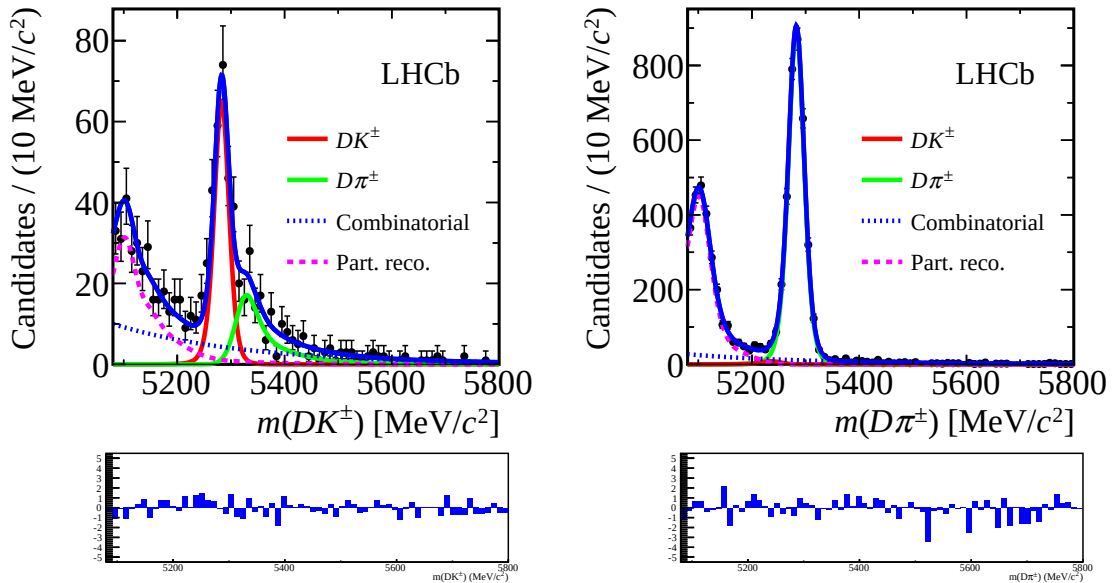


Figure 3.22. Invariant mass distributions of (left) $B^\pm \rightarrow DK^\pm$ and (right) $B^\pm \rightarrow D\pi^\pm$ candidates, with $D \rightarrow K_S^0 K^+ K^-$ in the DD K_S^0 category. Fit results, including the signal and background components, are superimposed and described in the legend. The associated pull distribution is shown below.

CHAPTER 4

The flavour-tagged $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ control channel

A flavour-tagged control channel is necessary to measure the quantities T_i in Eq. (1.40), which describe the decay intensity integrated over regions of the D^0 Dalitz plot. However, the non-uniform selection of events in different parts of the Dalitz plot must be considered. This is relevant for both the signal $B \rightarrow DK$ decay and for the flavour-tagged control channel because it is the number of events that are measured experimentally instead of the underlying decay rates. Only the relative efficiency from one point to another matters and not the absolute normalisation. Expressing the Dalitz plot efficiency distribution of the signal $B \rightarrow DK$ decays as $\eta(m_-^2, m_+^2)$ a new quantity is defined,

$$F_i \equiv \frac{\int_i dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta(m_-^2, m_+^2)}{\sum_j \int_j dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta(m_-^2, m_+^2)}. \quad (4.1)$$

which is the expected fraction of events in bin i of the $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ Dalitz plot if r_B were equal to zero. The sum in the denominator is over all the bins in the Dalitz plot. The Dalitz plot intensities $a^2(m_-^2, m_+^2)$ are as defined in Eq. (1.27). Writing the number of $B \rightarrow DK$ events as $N_{\pm i}^\pm$, where the subscript denotes the bin number and the superscript

charge refers to B^- or B^+ decays, the rate equations in Eq. (1.40) are modified to

$$N_{\pm i}^- = n_- \left[F_{\pm i} + (x_-^2 + y_-^2) F_{\mp i} + 2\sqrt{F_{+i} F_{-i}} (x_- c_{\pm i} + y_- s_{\pm i}) \right], \quad (4.2a)$$

$$N_{\pm i}^+ = n_+ \left[F_{\mp i} + (x_+^2 + y_+^2) F_{\pm i} + 2\sqrt{F_{+i} F_{-i}} (x_+ c_{\pm i} - y_+ s_{\pm i}) \right], \quad (4.2b)$$

where n_+ and n_- are normalisation constants for B^+ and B^- . The parameters c_i and s_i are as defined in Eqs. (1.36) and (1.37). The small impact of $\eta(m_-^2, m_+^2)$ on the values of c_i and s_i is neglected in the measurement, but is considered later as a systematic uncertainty. Using these equations, with knowledge of the values of F_i , c_i and s_i , the yields of $B \rightarrow DK$ in each bin can be used to determine the physics parameters of interest r_B , δ_B and γ . The optimal binning for $K_S^0 \pi^+ \pi^-$ is used in this analysis because it has been optimised for statistical sensitivity, and the 2×2 for $K_S^0 K^+ K^-$ is used because of the low yields expected.

The quantities F_i are determined using $B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu$ decays, where $\bar{D}^0$ decays to $K_S^0 h^+ h^-$. $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decays, as opposed to promptly produced D^{*-} mesons, are chosen because the triggers used at LHCb to select the D^{*-} candidates in $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decays are more similar to the triggers used to select the signal decay than those for promptly produced D^{*-} candidates. This produces more similar D Dalitz plot efficiencies between the signal and flavour-tagged channels. The compatibility in Dalitz plot efficiency between the signal and flavour-tagged decays is not strictly necessary, since efficiency differences can be corrected for, but preferable in order to limit the size of any efficiency corrections and the associated systematics. This decay mode also allows the D^0 meson flavour to be mistagged at a lower rate than the higher branching ratio $B^+ \rightarrow \bar{D}^0 \mu^+ \nu_\mu$ decay because of the ability to fully reconstruct the D^{*-} candidate.

The semileptonic candidates are selected and parameterised in Sects. 4.1 and 4.2. It is found that residual discrepancies in the D Dalitz plot efficiencies from $B \rightarrow Dh$

and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decays are present because of necessary differences in the selection requirements. The efficiency differences are documented in Sect. 4.3 and the efficiency correction and calculation of F_i values are described in Sect. 4.3.3.

4.1 Selection of semileptonic candidates

4.1.1 Reconstruction and trigger

$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates are formed in the stripping processing stage, where the composite particles $K_S^0 \rightarrow \pi^+ \pi^-$, $D \rightarrow K_S^0 h^+ h^-$ and $D^{*-} \rightarrow \bar{D}^0 \pi^-$ are formed, and loose requirements are placed to limit the size of the data samples. As with the signal mode, K_S^0 LL and DD candidates are treated differently. At this stage candidate B decays containing muons with both the same and opposite charge as the D^{*-} daughter pion, henceforth known as the *soft pion*, are created. Since the neutrino is not reconstructed by the LHCb detector, in practice the reconstructed candidates consist of $B \rightarrow D^{*\pm} \mu^+ X$ decays, where X corresponds to one or more particles that are missed.

The $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates stored in the stripping stage contain tracks and composite particles with higher p_T and p thresholds than in the $B \rightarrow Dh$ case because these thresholds are implemented for a different analysis. Additional stripping requirements are that the muon candidate have $DLL_{\mu\pi} > 0$ and that the reconstructed masses of the K_S^0 and D candidates lie within $30 \text{ MeV}/c^2$ and $100 \text{ MeV}/c^2$ of their known masses, respectively. The D^{*-} candidate is required to have a mass difference, $\Delta m = m(D^{*-}) - m(D^0)$, which lies between 0 and $170 \text{ MeV}/c^2$. True D^{*-} candidates are expected to peak close to $\Delta m = 145 \text{ MeV}/c^2$ [4]. The reconstructed B mass is required to fall within the range $2500\text{--}6000 \text{ MeV}/c^2$.

For each candidate the entire decay chain is refitted using DTF (see Sect. 3.1.1). The D^0 and D^{*-} masses used henceforth are obtained after a refit with KS0const, while

the Dalitz coordinates are obtained with D0const and KS0const. Candidates that fail either of these refits are removed. It should be noted that the Dalitz coordinates for $B \rightarrow Dh$ candidates are calculated with a refit that also contains PVconst, which cannot be used in the $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ case because of the missing neutrino. This results in a small relative shift in the position of the reconstructed Dalitz coordinates between the two modes and is considered as a systematic uncertainty.

At the hardware trigger level, at least one of two conditions must be satisfied. The muon candidate is required to be classified as L0Muon T0S or the D^0 candidate is required to be classified as L0Hadron T0S (defined in Sects. 2.6 and 3.1). Approximately 95% of muons in the final $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ sample are classified as L0Muon T0S and 5–15% of candidates satisfy both conditions. At HLT1, either the muon candidate is required to be classified as Hlt1TrackMuon T0S or the B candidate is required to be classified as Hlt1TrackAllL0 T0S. At HLT2, the B candidate is required to pass the topological trigger requirements of Hlt2TopoMuWBodyBBDT T0S ($N = 2, 3$ or 4) which selects multibody B decays where one track is consistent with a muon. An additional 5–10% of muon candidates are selected at HLT2 with a software trigger, which is designed to select muons from semileptonic b - and c -hadron decays, where the muon has a significant displacement from the PV and a high χ_{IP}^2 .

4.1.2 Selection

The selection criteria are chosen to minimise differences in Dalitz plot efficiency between the $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ and $B \rightarrow Dh$ modes. Hence, the selection requirements for $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates are based on the selection employed in Ref. [41], which are listed in Table 3.2. It is demonstrated in Fig. 4.1 that applying these cuts instead of the BDT requirement on $B \rightarrow Dh$, $D \rightarrow K_S^0 h^+ h^-$ candidates does not produce large relative differences in the Dalitz plot efficiency. (The proportion of candidates passing only one set of selection criteria is less than 10% of the number of candidates selected with either

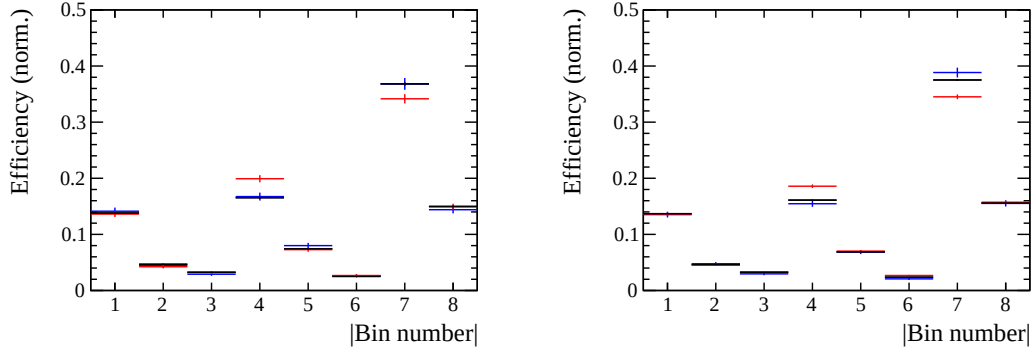


Figure 4.1. The distribution of simulated (left) LL and (right) DD $B \rightarrow D\pi$, $D \rightarrow K_s^0\pi^+\pi^-$ candidates in the optimal binning passing (red) only the BDT selection criteria, (blue) only the selection described in Table 3.2 and (black) both sets of selection criteria.

method.) Therefore, we choose to use the requirements of Ref. [41] for the D candidate and its daughters. The requirements to suppress $D \rightarrow \pi\pi\pi\pi$ decays and charmless background in Sect. 3.1.3, which have been demonstrated to be effective, are used here.

The selection criteria on the B candidate are different to those in Table 3.2 as it is not fully reconstructed. It is required to have DIRA (defined in Sect. 3.1.2) greater than 0.9999 and a flight distance χ^2 greater than 169. The reconstructed B mass is restricted to the range $m(D^{*-}\mu^+) < 5000 \text{ MeV}/c^2$, where approximately 98% of simulated $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays lie. The muon candidate is required to have χ_{IP}^2 greater than 25 as it should originate from a displaced vertex associated with a B meson decay, and its momentum is required to be less than $100 \text{ GeV}/c$ in order to match the momentum requirement on the bachelor in $B \rightarrow Dh$ decays. In addition, the muon is required to have opposite charge to the soft pion in order to suppress backgrounds when a random μ in the event is combined with a D^{*-} candidate. The charge requirement on the muon provides a double-tag that suppresses the rate of mistagging the D^0 flavour by a factor two. The fit to be discussed in Sect. 4.2 is designed to distinguish real D^{*-} candidates from D^0 candidates which have been combined with a random soft pion.

4.1.3 Additional requirements

The daughter kaons in $D \rightarrow K_S^0 K^+ K^-$ candidates are required to pass the same PID requirements as those used for the $B \rightarrow Dh$, $D \rightarrow K_S^0 K^+ K^-$ candidates. As in the $B \rightarrow Dh$ modes, candidates that fall outside the Dalitz plot kinematic boundary defined by the binning schemes are removed. Multiple candidates are removed with the criterion that the one retained has the lowest refit χ^2 from the refit containing D0const and KS0const.

4.1.4 Dalitz plots of selected semileptonic candidates

Dalitz plots of the candidates selected with the criteria outlined and in the signal mass windows are given in Figs. 4.2 and 4.3 for both D final states. The signal mass windows are defined as $|\Delta m - \Delta m_{\text{PDG}}| = 2.5 \text{ MeV}/c^2$ and $|m(D^0) - m(D^0)_{\text{PDG}}| = 25 \text{ (15) MeV}/c^2$ for $K_S^0 \pi^+ \pi^-$ ($K_S^0 K^+ K^-$), where the PDG subscript indicates the known value. The distributions are separated by muon charge (+ or -) and K_S^0 type (LL or DD).

4.2 Yield of $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ within the Dalitz plot

In this section the yield of semileptonic candidates in each Dalitz plot bin is measured. This is performed by first obtaining the shape parameters of the signal and background components with the full data sample, which are then fixed during a subsequent fit to measure the signal and background yields in each Dalitz plot bin.

A two-dimensional fit is performed in $m(D^0)$ and $\Delta m = m(D^{*-}) - m(D^0)$. There are three fit components: signal, combinatorial background that contains fake D candidates, and random slow pion (RSP) background, which occurs when real D candidates are combined with a random pion in the event. In this fit the RSP background peaks in $m(D^0)$ but forms a smooth background in Δm . Using the Δm distribution instead of

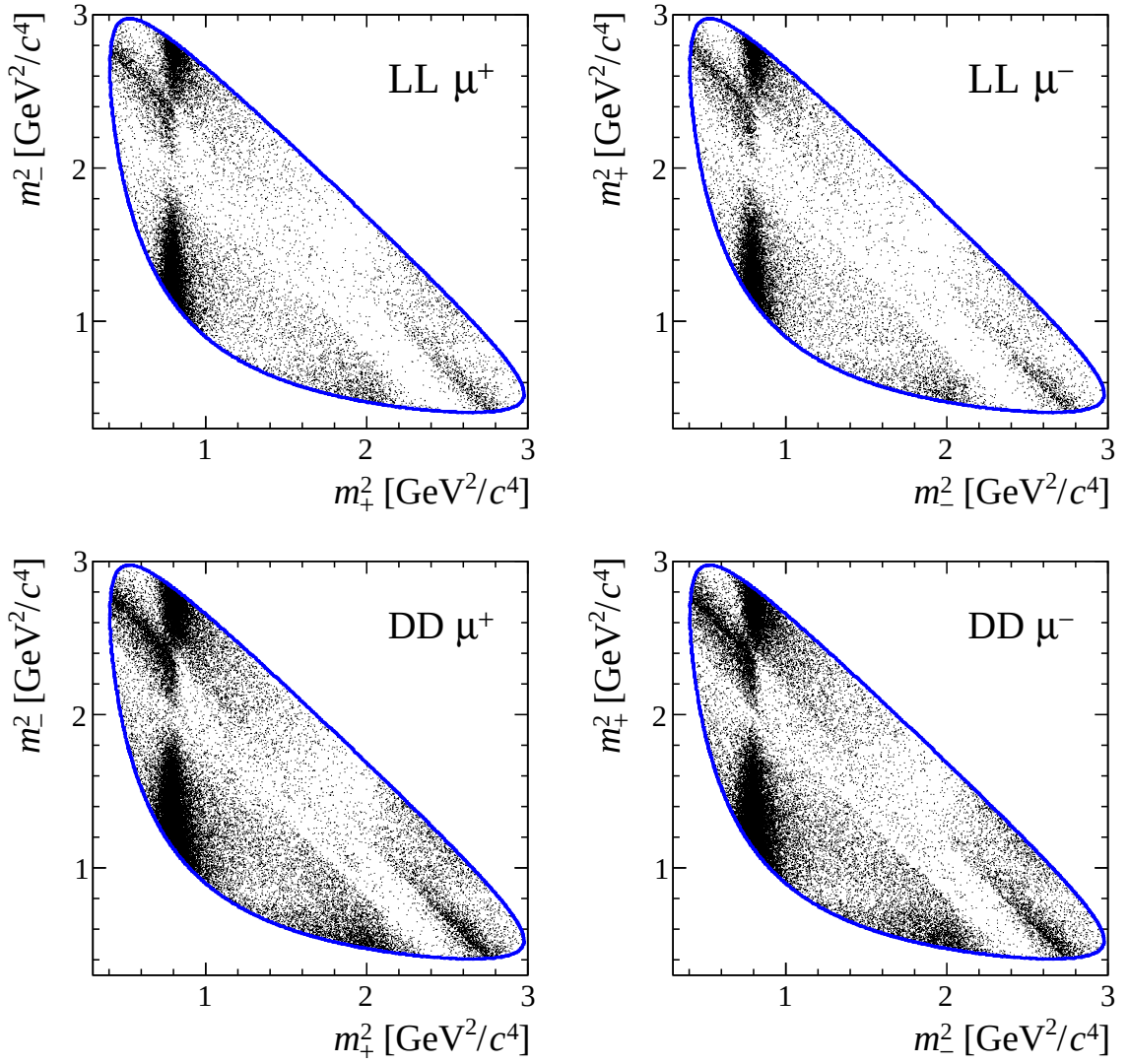


Figure 4.2. Dalitz plots for selected $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$, $D \rightarrow K_s^0\pi^+\pi^-$ candidates in the signal mass window. The data are divided by K_s^0 type, (top) LL and (bottom) DD, and muon charge, (left) positive and (right) negative.

the $m(D^{*-})$ distribution allows combinatorial D^{*-} candidates to be better distinguished, because true D^{*-} candidates form a narrow peak in Δm with a width of order $1 \text{ MeV}/c^2$.

4.2.1 Fit to determine the shape parameters

In the Δm projection, true D^{*-} decays are described using a sum of two Johnson functions. The Johnson function [84] is obtained by taking a general Gaussian function

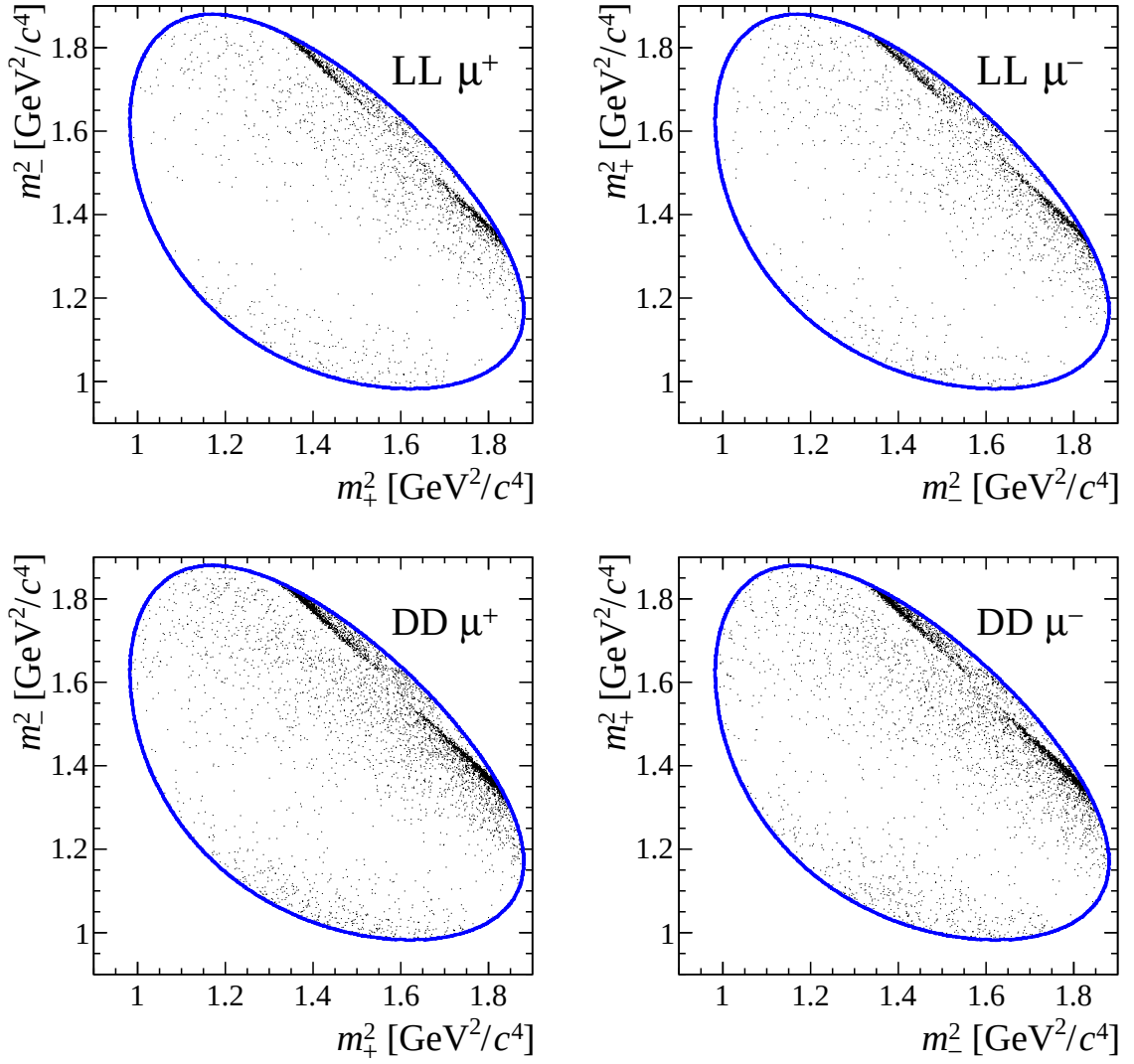


Figure 4.3. Dalitz plots for selected $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$, $D \rightarrow K_s^0 K^+ K^-$ candidates in the signal mass window. The data are divided by K_s^0 type, (top) LL and (bottom) DD, and muon charge, (left) positive and (right) negative.

$\frac{1}{\sqrt{2\pi}} \exp^{-Z^2/2}$ and performing the transformation $Z \rightarrow \beta + \delta \sinh^{-1} \left(\frac{x-\mu}{\sigma} \right)$, and is given by:

$$\text{Johnson}(\Delta m | \mu_m, \sigma, \beta, \delta) = \mathcal{N} \frac{\delta \exp \left(-\frac{1}{2} \left[\beta + \delta \sinh^{-1} \left(\frac{\Delta m - \mu_m}{\sigma} \right) \right]^2 \right)}{\sigma \sqrt{2\pi \left(1 + \left(\frac{\Delta m - \mu_m}{\sigma} \right)^2 \right)}}, \quad (4.3)$$

where μ_m is the mean parameter and σ is the width parameter. β is the asymmetry term, where $\beta > 0$ ($\beta < 0$) indicates a left-hand (right-hand) tail, and δ determines the steepness

of the tails. $\mathcal{N}$ is the normalisation factor with dimension MeV/c^2 . The two Johnson functions are summed $f_J \times \text{Johnson}(\Delta m | \mu_m, \sigma_1, \beta, \delta) + (1 - f_J) \times \text{Johnson}(\Delta m | \mu_m, \sigma_2, \beta, \delta)$, where f_J is the fraction of the first shape.

Combinatorial D^{*-} candidates are described using the *DstD0BG* function, which is defined in the region $\Delta m - \Delta m_0 > 0$ as

$$\text{DstD0BG}(\Delta m | \Delta m_0, x, p_1, p_2) = \left[1 - \exp\left(-\frac{\Delta m - \Delta m_0}{x}\right) \right] \left(\frac{\Delta m}{\Delta m_0} \right)^{p_1} + p_2 \left(\frac{\Delta m}{\Delta m_0} - 1 \right), \quad (4.4)$$

where Δm_0 , x , p_1 and p_2 are free parameters. The parameter Δm_0 describes the kinematic threshold, at $\Delta m_0 = m(\pi)$, below which it is unphysical to combine a D^0 meson with a pion, and therefore the function is set to zero in the region $\Delta m - \Delta m_0 \leq 0$.

In the $m(D^0)$ projection, true D candidates in the signal and RSP background are described using a Cruijff function, which is defined in Eq. (3.1). The combinatorial background is described with a second order Chebychev polynomial with linear coefficient, c_1 , and quadratic coefficient, c_2 . The combinatorial background is not expected to be uniform in the $m(K_s^0 h^+ h^-)$ variable, which is obtained after a refit with a K_s^0 mass constraint. While improving the D mass resolution, it also causes the combinatorial background to be sculpted towards the known D^0 mass, hence a quadratic with negative curvature ($c_2 < 0$) is expected.

An extended maximum likelihood fit to the $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates is performed to determine the shape parameters of each function. A simultaneous fit is performed with four data categories: $\{K_s^0 \pi^+ \pi^-, K_s^0 K^+ K^-\} \otimes \{\text{LL } K_s^0, \text{DD } K_s^0\}$. The fit range is $139.5 < \Delta m < 153.0 \text{ MeV}/c^2$ and $1830 < m(D^0) < 1910 \text{ MeV}/c^2$. The $m(D^0)$ range is chosen to be within a region where the Δm resolution does not vary significantly and the two variables are uncorrelated. Some categories have shape parameters in common, which are determined through an initial fit where all the parameters in every category

allowed to vary independently. The 2011 and 2012 data are fitted separately to allow studies of their Dalitz plot efficiency to be performed later.

Due to a high level of correlation between the parameters x , p_1 and p_2 in the DstD0BG function, the fit does not converge if all three are allowed to vary. Hence the values of p_1 and p_2 are fixed to the value returned in an initial fit where x , p_1 and p_2 vary freely. This has a negligible impact on the values and errors of both the parameter x and the yields, but allows the fit to converge with an accurate error matrix.

The results of the fit to 2012 data are shown in Fig. 4.4 and Fig. 4.5. The narrower width of the $m(K_s^0 K^+ K^-)$ spectrum reflects the smaller Q -value available in the $D \rightarrow K_s^0 K^+ K^-$ decay than in $D \rightarrow K_s^0 \pi^+ \pi^-$. The fitted shape parameters are given in Table 4.1. Similar results are found in the fit to 2011 data. The signal and background yields in the fits to 2011 and 2012 data are given in Table 4.2. The total signal yield of $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decays in all the categories is approximately 120 000, which ensures that the semileptonic data are not the limiting statistical error in measuring $x_\pm, y_\pm$. The purity is greater than 94% in all cases.

4.2.2 Fit to measure the yield in each Dalitz plot bin

A second fit is performed to measure the yield of semileptonic candidates in each Dalitz plot bin, separated by muon charge. The data categories are $\{\text{LL } K_s^0, \text{DD } K_s^0\} \otimes \{\mu^+, \mu^-\} \otimes \{16 \text{ bins } K_s^0 \pi^+ \pi^-, 4 \text{ bins } K_s^0 K^+ K^-\}$. As with the signal decay, the optimal binning is used for $K_s^0 \pi^+ \pi^-$ and the 2×2 binning is used for $K_s^0 K^+ K^-$ (Figs. 1.8 and 1.9). The shape parameters are fixed to the results obtained in the initial mass fit. The yield in each Dalitz plot bin is shown in Fig. 4.6 for $K_s^0 \pi^+ \pi^-$ and Fig. 4.7 for $K_s^0 K^+ K^-$. The yields from μ^+ and μ^- have been combined using the *effective* Dalitz plot bin numbers, which allow the data from B and $\bar{B}$ to be combined by adding the information in bin i of B decays with bin $-i$ of $\bar{B}$ decays.

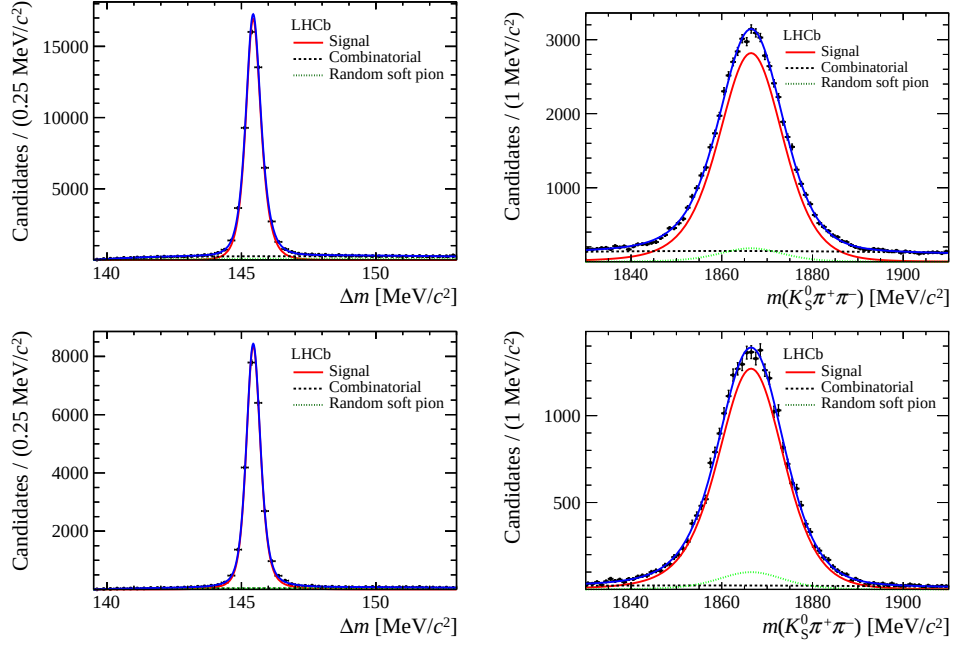


Figure 4.4. The (left) Δm and (right) $m(K_S^0 \pi^+ \pi^-)$ distributions of selected $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates, with (top) K_S^0 DD and (bottom) K_S^0 LL, from 2012 data are shown with the (blue) fit result superimposed.

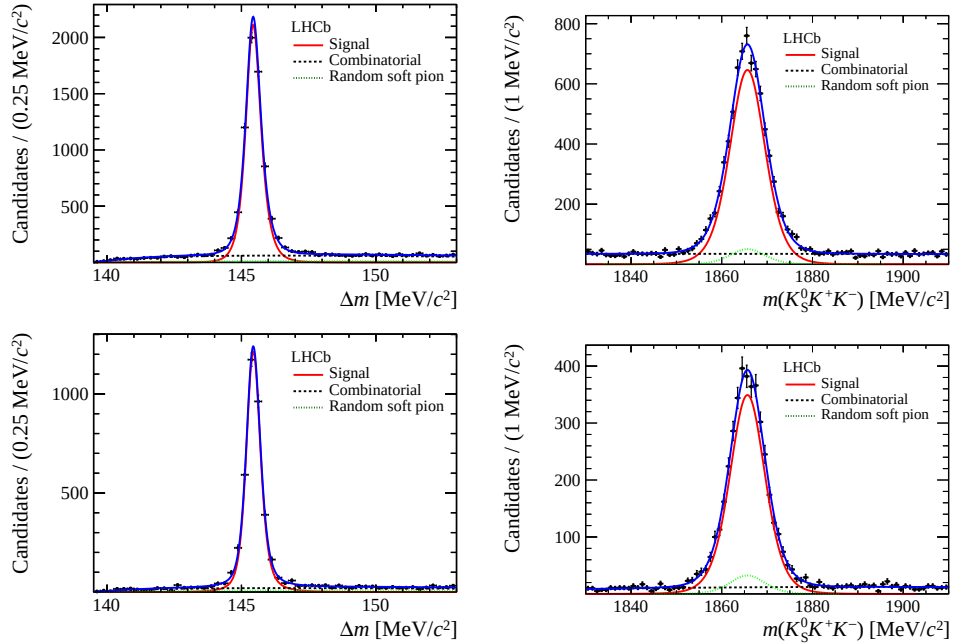


Figure 4.5. The (left) Δm and (right) $m(K_S^0 K^+ K^-)$ distributions of selected $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ candidates, with (top) K_S^0 DD and (bottom) K_S^0 LL, from 2012 data are shown with the (blue) fit result superimposed.

Table 4.1. Results of the fit to $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$, $D \rightarrow K_S^0 h^+ h^-$ candidates in 2012 data.

Parameter	$K_S^0 \pi^+ \pi^-$		$K_S^0 K^+ K^-$	
	LL	DD	LL	DD
Δm projection				
μ_m [MeV/ c^2]	145.410 ± 0.004		same as $K_S^0 \pi^+ \pi^-$	
σ_1 [MeV/ c^2]	1.66 ± 0.14	0.43 ± 0.03	same as $K_S^0 \pi^+ \pi^-$	
σ_2 [MeV/ c^2]	0.52 ± 0.04	0.43 ± 0.02	same as $K_S^0 \pi^+ \pi^-$	
β	-0.16 ± 0.02	-0.15 ± 0.01	same as $K_S^0 \pi^+ \pi^-$	
δ	1.92 ± 0.17	1.39 ± 0.02	same as $K_S^0 \pi^+ \pi^-$	
f_J	0.11 ± 0.02	0.39 ± 0.00	same as $K_S^0 \pi^+ \pi^-$	
Δm_0 [MeV/ c^2]	139.48 ± 0.03		same as $K_S^0 \pi^+ \pi^-$	
x (comb.) [MeV/ c^2]	3.52 ± 0.19		2.51 ± 0.19	
x (RSP) [MeV/ c^2]	9.9 ± 0.8		same as $K_S^0 \pi^+ \pi^-$	
$m(D^0)$ projection				
μ [MeV/ c^2]	1866.5 ± 0.03		1865.7 ± 0.05	
σ [MeV/ c^2]	6.91 ± 0.05	6.73 ± 0.04	3.73 ± 0.07	3.74 ± 0.06
α_L	0.138 ± 0.002		0.098 ± 0.008	
α_R	0.098 ± 0.003		0.095 ± 0.008	
c_1	-0.17 ± 0.05	-0.07 ± 0.02	0.16 ± 0.06	-0.007 ± 0.034
c_2	-0.11 ± 0.07	-0.07 ± 0.02	-0.07 ± 0.07	-0.006 ± 0.039

Table 4.2. The yields of each fit component and the signal purity of the $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ candidates measured in the fit, determined in the range $\pm 2.5 \text{ MeV}/c^2$ from the known Δm value and $\pm 25(15) \text{ MeV}/c^2$ from the known D mass for $K_S^0 \pi^+ \pi^-$ ($K_S^0 K^+ K^-$).

Mode	Signal	Combinatorial	RSP	Purity
2012				
$K_S^0 \pi^+ \pi^-$ LL	23840 ± 170	299 ± 16	499 ± 29	0.968 ± 0.001
$K_S^0 \pi^+ \pi^-$ DD	52020 ± 240	1924 ± 32	888 ± 44	0.949 ± 0.001
$K_S^0 K^+ K^-$ LL	3485 ± 61	90 ± 4	86 ± 9	0.952 ± 0.003
$K_S^0 K^+ K^-$ DD	6522 ± 84	271 ± 7	135 ± 12	0.941 ± 0.002
2011				
$K_S^0 \pi^+ \pi^-$ LL	11270 ± 120	120 ± 9	268 ± 20	0.967 ± 0.002
$K_S^0 \pi^+ \pi^-$ DD	21920 ± 160	703 ± 19	391 ± 28	0.952 ± 0.001
$K_S^0 K^+ K^-$ LL	1667 ± 43	42 ± 3	46 ± 6	0.950 ± 0.004
$K_S^0 K^+ K^-$ DD	2858 ± 56	117 ± 4	48 ± 8	0.945 ± 0.003

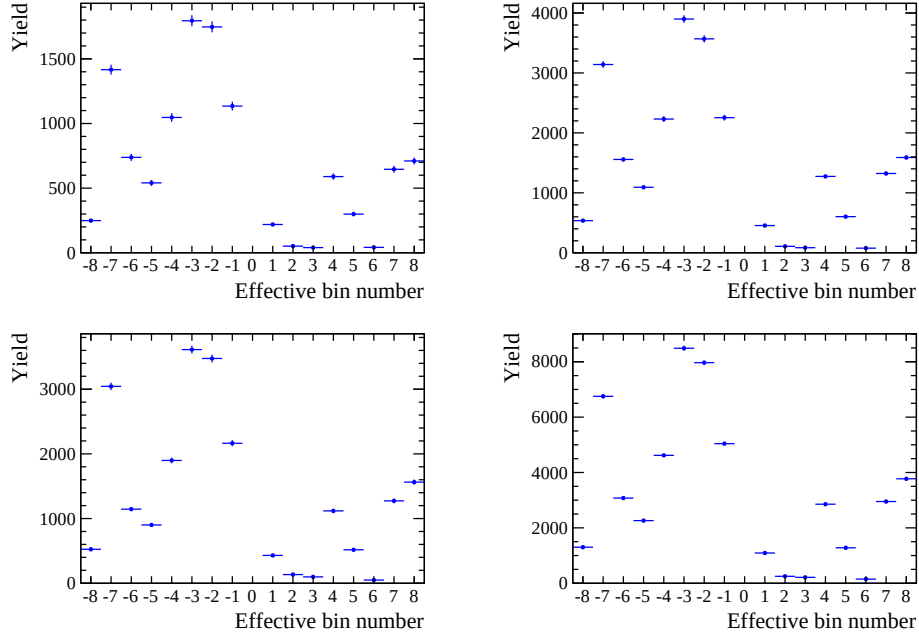


Figure 4.6. Yield of semileptonic $K_S^0 \pi^+ \pi^-$ candidates in each effective Dalitz plot bin of the optimal binning, separated by (top) LL and (bottom) DD K_S^0 , and (left) 2011 and (right) 2012 data. The results from both muon charges have been combined.

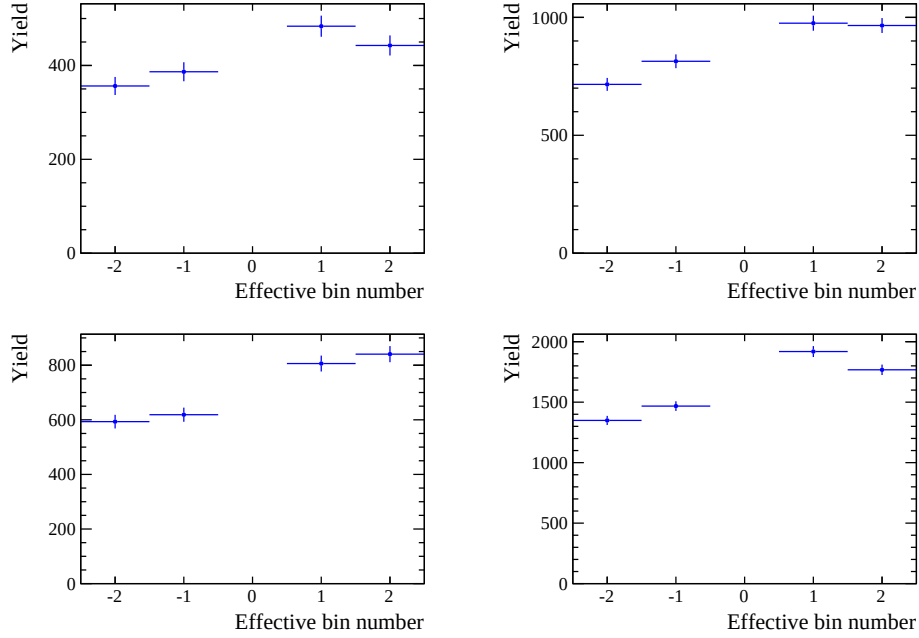


Figure 4.7. Yield of semileptonic $K_S^0 K^+ K^-$ candidates in each effective Dalitz plot bin of the 2×2 binning, separated by (top) LL and (bottom) DD K_S^0 , and (left) 2011 and (right) 2012 data. The results from both muon charges have been combined.

4.3 Dalitz plot efficiency correction

The measured yields of semileptonic candidates in each Dalitz plot bin in data are only equal to F_i if the $D^0 \rightarrow K_S^0 h^+ h^-$ Dalitz plot efficiency profiles of the semileptonic and $B \rightarrow DK$ modes are the same. The efficiency profiles of the two modes are compared by studying both data and MC in Sect. 4.3.1 and are found to be different. The sources of these differences and the reasons for correcting them using MC are discussed in Sect. 4.3.2. The efficiency correction method is described in Sect. 4.3.3.

4.3.1 Efficiency differences in the semileptonic and hadronic modes

The Dalitz plot efficiency of two samples of flavour-tagged $D^0 \rightarrow K_S^0 h^+ h^-$ candidates can be compared by measuring their relative yields in each Dalitz plot bin. The Dalitz plot distribution depends only on the $D^0 \rightarrow K_S^0 h^+ h^-$ decay amplitude and the efficiency profile, hence differences in the relative yields between the two sets of candidates can be attributed to differences in their efficiencies.

This technique cannot be used directly with $B \rightarrow DK$ data as CP violation changes the Dalitz plot distribution. Although this can be corrected using existing measurements of the CP parameters, it is preferable to use $B \rightarrow D\pi$ data, where the level of CP violation is quantified by $r_B^{D\pi} = \mathcal{O}(1\%)$ [85]. This level of CP violation can introduce large localised variations in the Dalitz plot, but only a few-percent change at most is incurred in the binned yields when the B and $\bar{B}$ data are combined, hence CP violation can be ignored in the qualitative investigation of efficiency differences in this section. Furthermore, studies of the $B \rightarrow DK$ and $B \rightarrow D\pi$ efficiencies in $D \rightarrow K_S^0 \pi^+ \pi^-$ MC, shown in Figs. 4.8 and 4.9, indicate that the Dalitz plot efficiencies of selected candidates from the two modes are almost identical. This is expected because the difference in the bachelor's mass and PID requirements between the two decays has negligible effects on the

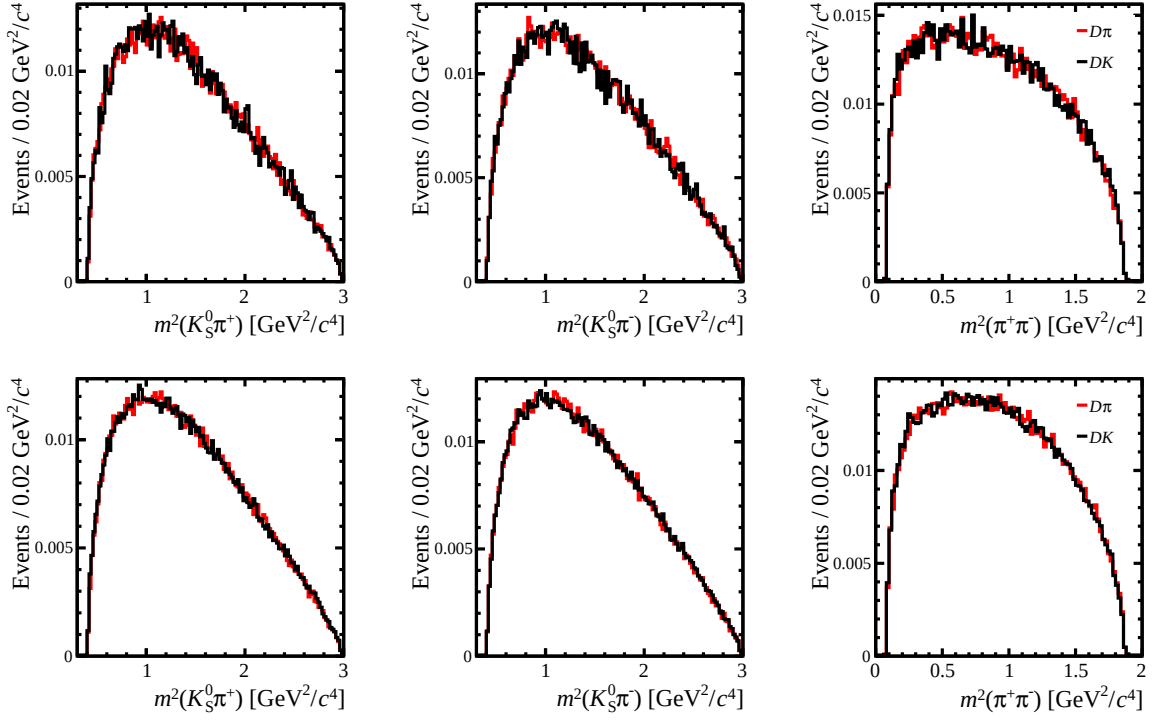


Figure 4.8. The Dalitz plot projections, (left to right) $m^2(K_s^0 \pi^+)$, $m^2(K_s^0 \pi^-)$ and $m^2(\pi^+ \pi^-)$, of simulated (red) $B \rightarrow D\pi$ and (black) $B \rightarrow DK$ candidates in the $D \rightarrow K_s^0 \pi^+ \pi^-$ final state reconstructed as (top) LL and (bottom) DD K_s^0 in 2012 conditions.

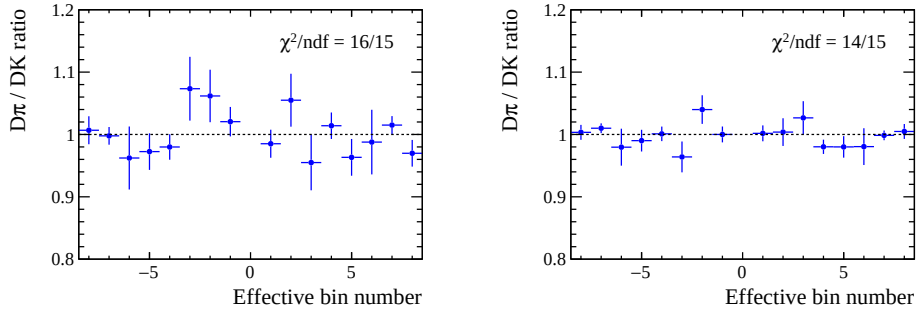


Figure 4.9. Ratio between the yields of simulated $B \rightarrow D\pi$ and $B \rightarrow DK$ candidates in each Dalitz plot bin for $K_s^0 \pi^+ \pi^-$ candidates in (left) LL and (right) DD K_s^0 from 2012 conditions. The dashed line at a ratio of one is drawn to guide the eye. The χ^2/ndf of fitting the ratios to a straight line is given.

kinematics of the D daughters. An additional advantage of using $B \rightarrow D\pi$ data is the factor of 15 larger yield than $B \rightarrow DK$ and the greater than 95% purity of the selected candidates, which reduces the uncertainties related to extracting the binned yields.

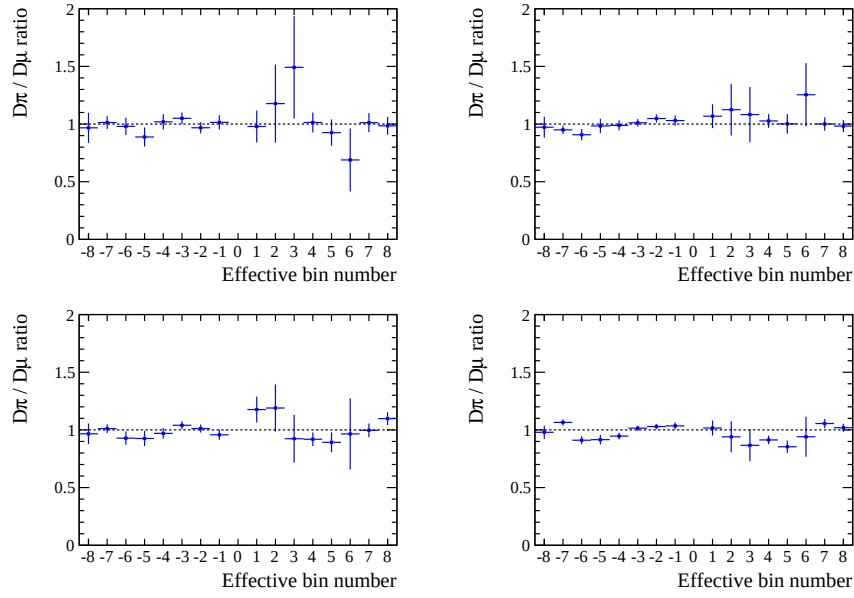


Figure 4.10. Ratio between the yields of $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ in each effective Dalitz plot bin for $K_S^0\pi^+\pi^-$ candidates in (top) LL and (bottom) DD K_S^0 from (left) 2011 and (right) 2012 data. The results from both charges have been combined. The dashed line at a ratio of one is drawn to guide the eye.

The data are summed and analysed in the effective bin numbers in order to reduce the statistical uncertainties.

The $B \rightarrow D\pi$ yields in data are measured using the fit configuration described in Sect. 3.2 and fixing the shape parameters to the results in Table 3.6. The fit is performed with $B \rightarrow D\pi$ candidates only, which are further divided according to their Dalitz plot bin number and bachelor charge. The $B \rightarrow D\pi$ signal yields are measured in each bin and are compared with the $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ yields measured in the previous section. The 2011 and 2012 data and the two K_S^0 types are studied separately as their efficiency profiles may differ.

The measured yields are normalised to an integral of one and then shown as a ratio between $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ for each category in Figs. 4.10 and 4.11. If the efficiency profiles are identical, the ratio is expected to be unity in each bin. However, in most bins, deviations of order 10% or less are observed. The discrepancies are most clearly observed in the DD $K_S^0\pi^+\pi^-$ 2012 sample, which contains the highest

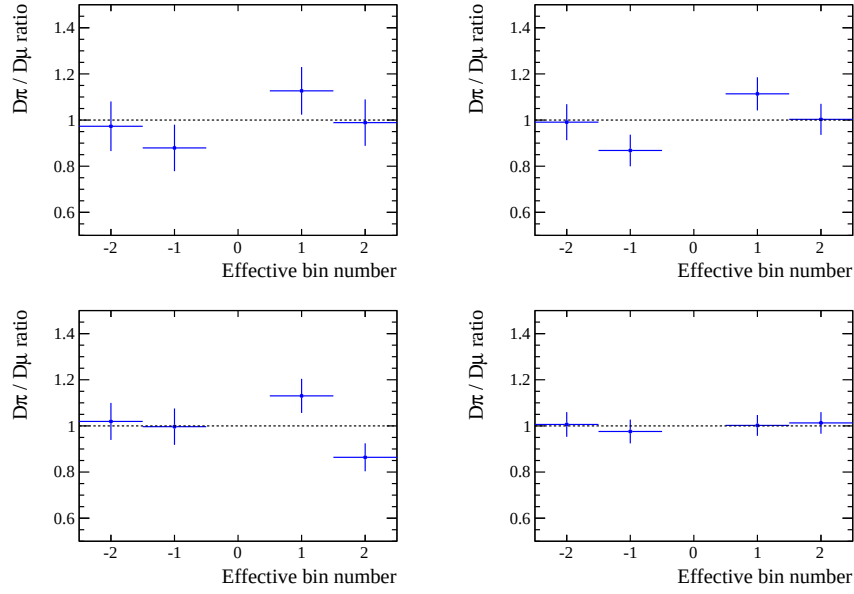


Figure 4.11. Ratio between the yields of $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ in each effective Dalitz plot bin for $K_s^0 K^+ K^-$ candidates in (top) LL and (bottom) DD K_s^0 from (left) 2011 and (right) 2012 data. The results from both charges have been combined. The dashed line at a ratio of one is drawn to guide the eye.

statistics, where fitting a straight line to the distribution results in $\chi^2/\text{ndf} = 40/15$. Deviations of this magnitude cannot be explained with statistical fluctuations or by taking into account the effect of CP violation in $B \rightarrow D\pi$. They also cannot be explained by candidates with mistagged D flavour in either sample. The $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ yields are measured with a fit to the Δm distribution which distinguishes true and fake D^{*-} candidates, and in the case of $B \rightarrow D\pi$, random pions are accounted for in the combinatorial background component in the fit to the $B^\pm$ mass distribution. Although the shifts are less evident in the $K_s^0 K^+ K^-$ candidates due to the lower statistics available, the ratios point to efficiency differences between the hadronic and semileptonic decays that need to be corrected for.

The presence of efficiency differences is reinforced by a visual inspection of the Dalitz plot distribution of $K_s^0 \pi^+ \pi^-$ candidates from simulated $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays in Fig. 4.12. Comparing the distributions for DD K_s^0 , in $B \rightarrow D\pi$ the efficiency approximately halves when moving along the bottom left to top right diagonal of the

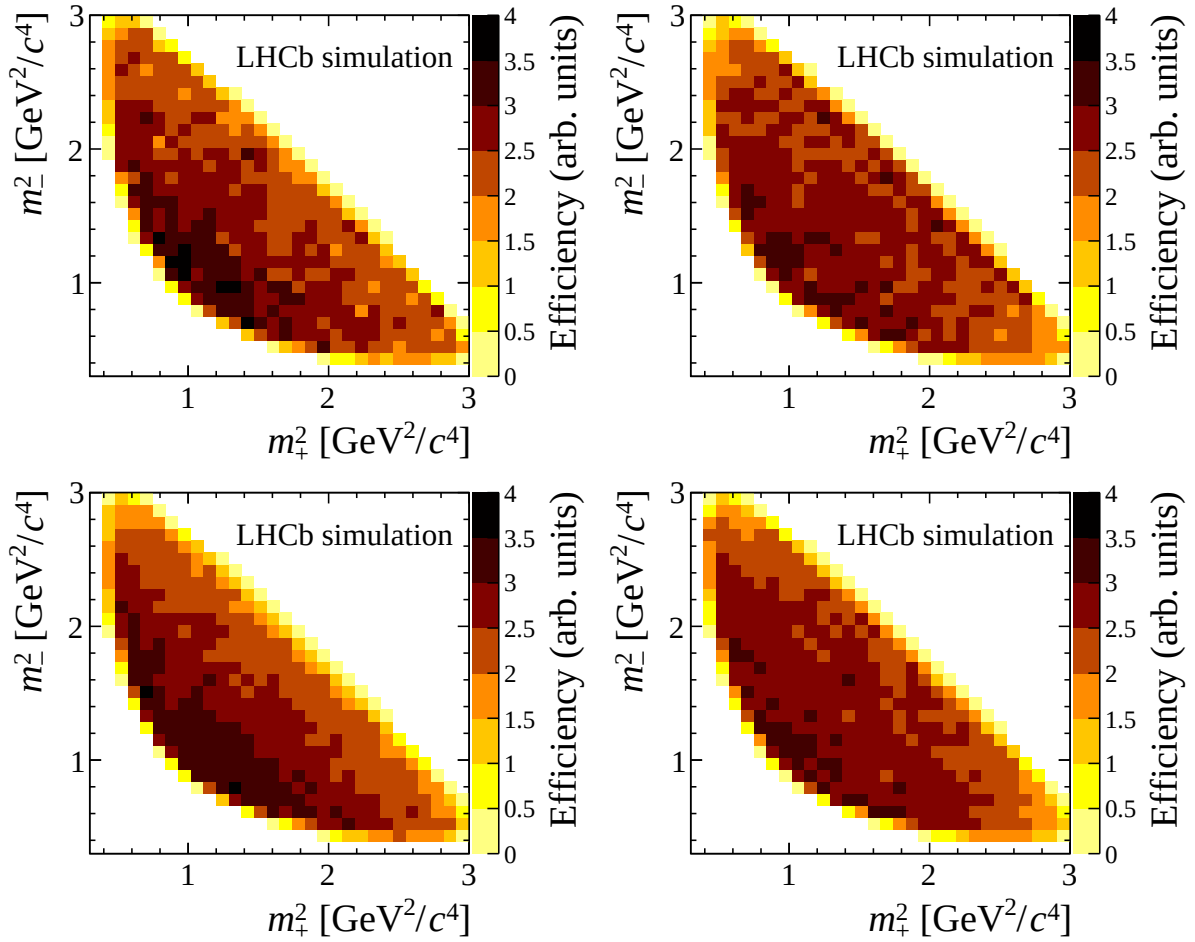


Figure 4.12. Dalitz plot distribution of (left) $B \rightarrow D\pi$ and (right) $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ candidates in the $D \rightarrow K_s^0\pi^+\pi^-$ final state reconstructed as (top) LL and (bottom) DD K_s^0 simulated in 2012 conditions and passing the respective selection criteria. The absolute scale on the z-axis is arbitrary.

Dalitz plot. In $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ the efficiency profile is flatter along this diagonal, with an additional decrease in efficiency in the top left and bottom right corners of the Dalitz plot, which corresponds to regions where one of the D daughter pions has a low momentum. The observed visual differences in the efficiency profiles are consistent with bin numbers four, five and six showing the largest deviations in the comparison of data yields in the $K_s^0\pi^+\pi^-$ DD 2012 sample. Similarly for LL K_s^0 , a flatter efficiency profile in $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ than in $B \rightarrow D\pi$ is observed.

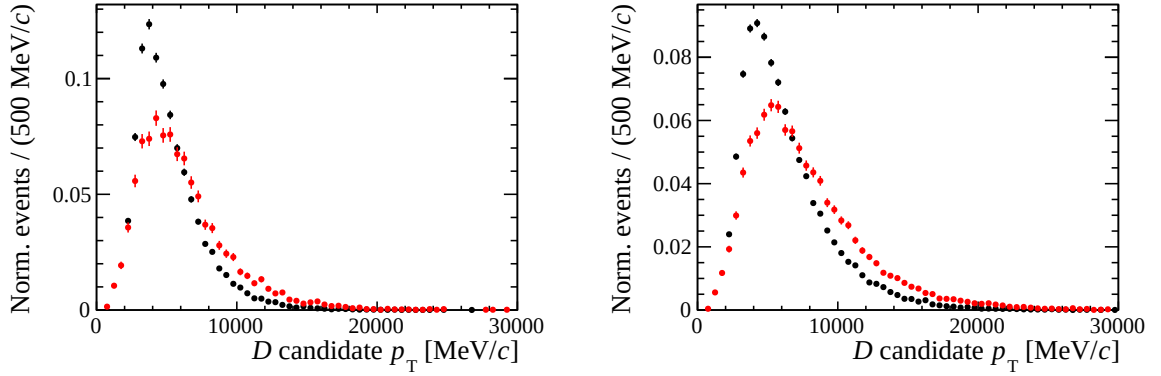


Figure 4.13. The p_T distribution of the D^0 candidate in selected (red) $B \rightarrow D\pi$ and (black) $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$, $D \rightarrow K_s^0\pi^+\pi^-$ candidates from 2012 data. D^0 candidates containing (left) LL K_s^0 and (right) DD K_s^0 candidates are shown. The total number of events in each sample has been normalised to unity.

4.3.2 Sources of efficiency difference

The disagreement in efficiency profiles can be attributed to differences in the decay product kinematics, and necessary differences in the selection of the semileptonic and hadronic modes. The presence of a muon in the $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decay means that the triggers used to select candidates necessarily differ from the $B \rightarrow Dh$ modes. Up to 95% of the selected semileptonic candidates pass the muonic hardware trigger, which has different p_T and p thresholds from the hadronic hardware trigger. In addition, the software topological triggers used are based on the application of a BDT, which is trained and optimised separately for the hadronic and muonic versions of the trigger (Hlt2TopoNBodyBBDT, Hlt2TopoMuNBodyBBDT). Hence, different trigger selections can result in different Dalitz plot efficiencies. Figure 4.13 compares the p_T distribution of the D^0 meson from selected $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ and $B \rightarrow D\pi$ candidates in data. In the $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decay, the D^0 meson is expected to have a smaller p_T on average because there are a greater number of particles in the final state, and this is consistent with the data shown. Differences in the p_T distribution between the two modes also come from different p_T thresholds imposed on the D^0 daughters during the stripping and offline selection.

Performing a re-weighting of the final state kinematics due to changes introduced at each stage of the selection is not practical because of the correlations that can exist between them. In addition, no single criterion in the selection is found to be the dominant cause of efficiency differences. Residual efficiency differences are still observed after re-weighting the semileptonic data using the $B \rightarrow D\pi$ kinematics and geometry. Therefore, although it was hoped that an efficiency correction could be achieved with data alone, the variety of contributing sources and inability to find an adequate solution with data motivated the use of MC to calculate and apply the efficiency correction instead.

4.3.3 The efficiency correction

The overall strategy to calculate the values of F_i , defined in Eq. (4.1), is to use the measurements of semileptonic yields in data and apply efficiency correction factors, which are calculated in each Dalitz plot bin using $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ MC. It is useful to define a number of efficiency profiles, denoted by $\eta \equiv \eta(m_-^2, m_+^2)$, where the profiles for $B \rightarrow D\pi$, $B \rightarrow DK$ and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ are $\eta_{D\pi}$, η_{DK} and $\eta_{D^*\mu}$, respectively. The fraction of semileptonic events measured in bin i is given by

$$M_i = \frac{\int_i dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta_{D^*\mu}(m_-^2, m_+^2)}{\sum_j \int_j dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta_{D^*\mu}(m_-^2, m_+^2)}. \quad (4.5)$$

Hence the required correction factor for bin i , ξ_i , is

$$\xi_i = \frac{\int_i dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta_{DK}(m_-^2, m_+^2)}{\int_i dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta_{D^*\mu}(m_-^2, m_+^2)}, \quad (4.6)$$

which relates F_i and M_i via

$$F_i = Z \xi_i M_i, \quad (4.7)$$

where Z is a normalisation factor such that $Z \sum_i \xi_i M_i = 1$. Assuming the same efficiency profile for the B^+ and B^- Dalitz plots gives $\xi_i(B^-) = \xi_{-i}(B^+)$.

The efficiency profiles are obtained from simulation. Throughout this discussion it is assumed that the efficiency profiles in MC describe the data perfectly, and any MC-data differences are evaluated as a systematic uncertainty. In practice $B \rightarrow D\pi$ MC is used to provide η_{DK} due to the larger sample of simulated decays available, and it has been demonstrated in Fig. 4.8 that $\eta_{DK} = \eta_{D\pi}$. Hence, Eq. (4.6) approximates to

$$\xi_i = \frac{\int_i dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta_{D\pi}(m_-^2, m_+^2)}{\int_i dm_-^2 dm_+^2 a^2(m_-^2, m_+^2) \eta_{D^*\mu}(m_-^2, m_+^2)}. \quad (4.8)$$

The amplitude models produced by BaBar for $D \rightarrow K_s^0 \pi^+ \pi^-$ in Ref. [33] and for $D \rightarrow K_s^0 K^+ K^-$ in Ref. [34] are used to provide the values of $a^2(m_-^2, m_+^2)$. The models provide a good description of the BaBar data and only a small systematic uncertainty is associated with the efficiency parameterisation. It should be noted that using intensities from an amplitude model does not affect the phase and thereby incur a model dependence. An amplitude model that is constructed with a different list of resonances, or that uses a different parameterisation, can still provide an intensity distribution that describes the data well. Further, any deficiencies in the model description will cancel to first order in the ratio in Eq. (4.8).

The Dalitz plot binning scheme and the values of $a^2(m_-^2, m_+^2)$ are both supplied on a fine pixel grid, hence the integral over bin i in Eq. (4.8) is performed by a summation over the pixels contained in bin i . The efficiency profile η is obtained by applying the same selection criteria as in data to the $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ MC and dividing the reconstructed events into the same fine pixel grid. Since the MC was generated uniformly in the Dalitz plot, the number of reconstructed events in a given pixel is proportional to the efficiency. In order to reduce the statistical uncertainty associated with the MC statistics available, it is assumed that the Dalitz plot efficiency is symmetric

under $m_+^2 \leftrightarrow m_-^2$ so that the reconstructed events on mirror-opposite pixels about the line of symmetry can be summed. This assumption is revisited as a systematic uncertainty in Sect. 5.3. Hence ξ_i is calculated using the sum

$$\xi_i = \frac{\sum_j (n_j^{D\pi} I_j)}{\sum_j (n_j^{D^*\mu} I_j)}, \quad (4.9)$$

where the index j runs over all the pixels that are contained in bin i . The parameter n_j is the sum of the number of MC events in pixel j and in its mirror-opposite pixel, and I_j is the value of $a^2(m_-^2, m_+^2)$ in pixel j of the BaBar amplitude model. The correction factors ξ_i are determined separately for events reconstructed as LL and DD K_S^0 , and separately for 2011 and 2012 conditions. The values of ξ_i are shown in Figs. 4.14 and 4.15.

The correction factors are applied to the semileptonic data using Eq. (4.7). To check the effect of the correction, the resulting F_i values are compared to the observed population as a function of the Dalitz plot in $B \rightarrow D\pi$ data. Figure 4.16 shows the ratio between the $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ fractional yields before and after the correction [46]. The ratios from the four samples (DD, LL, 2011, 2012) have been combined according to their proportions in data. The χ^2/ndf when considering the deviation from unity improves significantly from 48.1/20 to 29.5/20, demonstrating that the correction is effective.

The values of F_i are given in Table 4.3. The errors on F_i come from the uncertainties on M_i and ξ_i , which contribute equally. This method is thus data-driven, with MC used to provide a small correction, and is preferable to using only $B \rightarrow D\pi$ MC and an amplitude model to calculate the values of F_i , because any deficiencies in either the simulation or the amplitude model cancel to first order in the calculation of the correction factor.

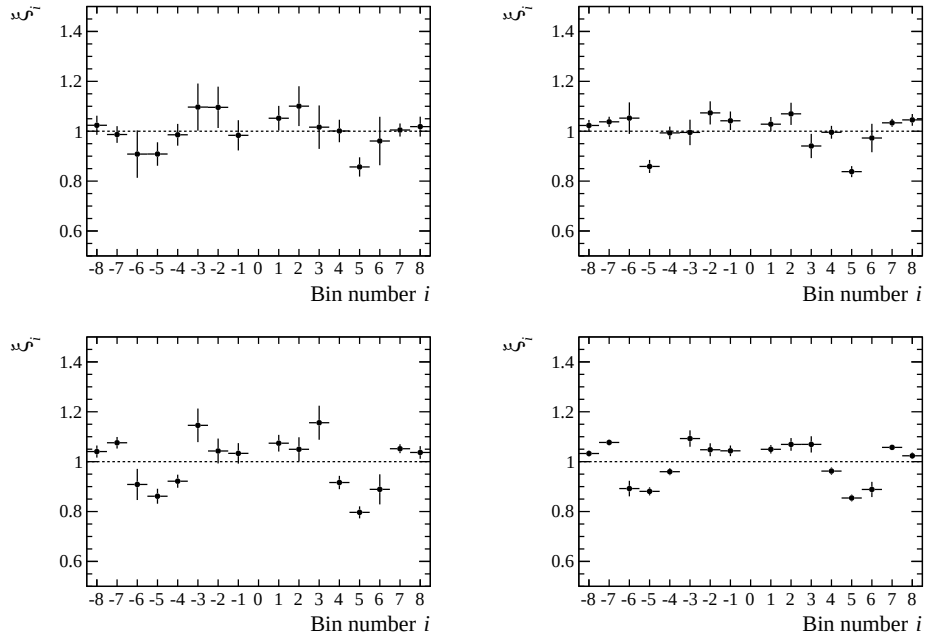


Figure 4.14. The efficiency correction factors ξ_i in each Dalitz plot bin of $K_s^0 \pi^+ \pi^-$ candidates in (top) LL and (bottom) DD K_s^0 from (left) 2011 and (right) 2012 data.

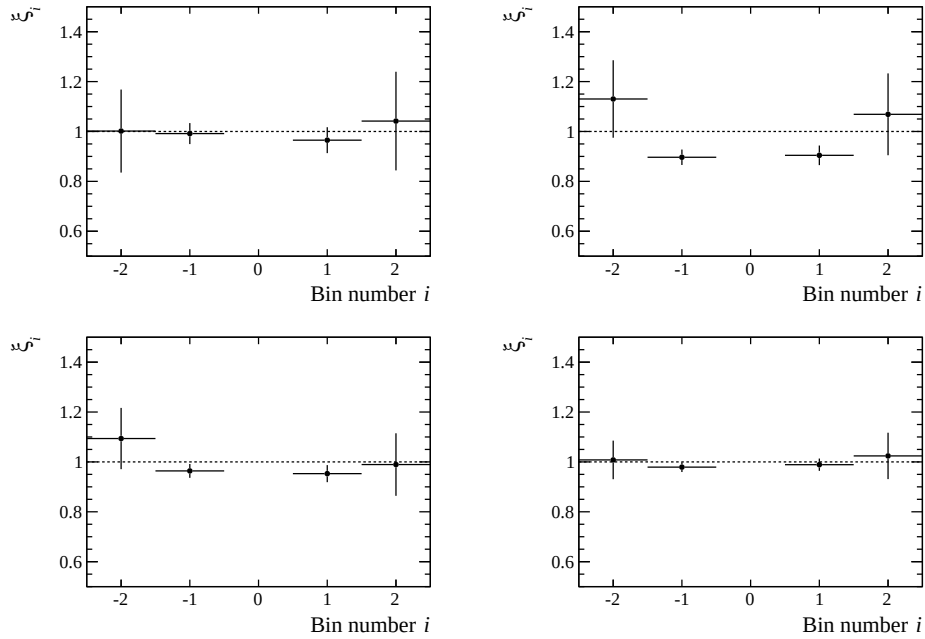


Figure 4.15. The efficiency correction factors ξ_i in each Dalitz plot bin of $K_s^0 K^+ K^-$ candidates in (top) LL and (bottom) DD K_s^0 from (left) 2011 and (right) 2012 data.

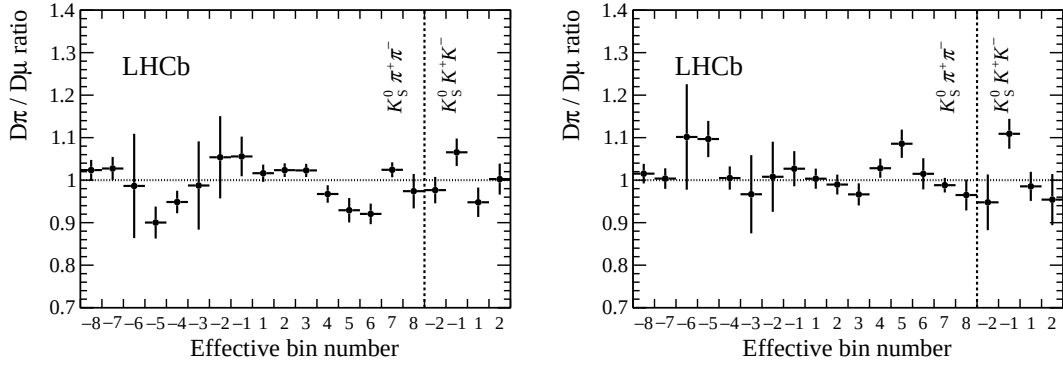


Figure 4.16. Fractional yield ratios of $B \rightarrow D\pi$ to $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decays for each effective Dalitz plot bin. The vertical dashed line separates the ratios for $K_S^0 \pi^+ \pi^-$ (bins -8 to $+8$) and $K_S^0 K^+ K^-$ (bins -2 to $+2$). The left (right) plot shows the ratios before (after) correcting for the efficiency differences. Also in [46].

Table 4.3. The central values and uncertainties of F_i calculated using the efficiency-corrected semileptonic yields.

i	F_i		i	F_i	
	LL	DD		LL	DD
$D \rightarrow K_S^0 \pi^+ \pi^-$ optimal binning			$D \rightarrow K_S^0 K^+ K^-$ 2×2 binning		
-8	0.0225 ± 0.0004	0.0249 ± 0.0003	-2	0.2276 ± 0.0246	0.2150 ± 0.0137
-7	0.1306 ± 0.0023	0.1391 ± 0.0014	-1	0.2175 ± 0.0060	0.2174 ± 0.0037
-6	0.0648 ± 0.0034	0.0499 ± 0.0016	+1	0.2636 ± 0.0092	0.2849 ± 0.0061
-5	0.0401 ± 0.0011	0.0365 ± 0.0006	+2	0.2914 ± 0.0353	0.2827 ± 0.0209
-4	0.0911 ± 0.0020	0.0816 ± 0.0011			
-3	0.1642 ± 0.0074	0.1769 ± 0.0049			
-2	0.1613 ± 0.0062	0.1579 ± 0.0036			
-1	0.0972 ± 0.0031	0.0989 ± 0.0018			
+1	0.0196 ± 0.0005	0.0212 ± 0.0003			
+2	0.0048 ± 0.0002	0.0054 ± 0.0001			
+3	0.0034 ± 0.0002	0.0045 ± 0.0001			
+4	0.0522 ± 0.0012	0.0498 ± 0.0007			
+5	0.0214 ± 0.0005	0.0199 ± 0.0003			
+6	0.0032 ± 0.0002	0.0023 ± 0.0001			
+7	0.0566 ± 0.0008	0.0588 ± 0.0005			
+8	0.0669 ± 0.0013	0.0723 ± 0.0008			

CHAPTER 5

Measurement of the CP violation observables in

$B^\pm \rightarrow DK^\pm$ decays

In this chapter the CP violation observables, $x_\pm$ and $y_\pm$, are measured using the relative distribution of $B^\pm \rightarrow DK^\pm$, $D \rightarrow K_S^0 h^+ h^-$ candidates in the D Dalitz plot. As outlined in Chapter 4, the number of signal events falling into each Dalitz plot bin, $N_{\pm i}^\pm$, is determined by F_i , c_i , s_i , $x_\pm$ and $y_\pm$ according to the relations in Eq. (4.2) reproduced below:

$$N_{\pm i}^- = n_- \left[F_{\pm i} + (x_-^2 + y_-^2) F_{\mp i} + 2\sqrt{F_{+i} F_{-i}} (x_- c_{\pm i} + y_- s_{\pm i}) \right], \quad (4.2a)$$

$$N_{\pm i}^+ = n_+ \left[F_{\mp i} + (x_+^2 + y_+^2) F_{\pm i} + 2\sqrt{F_{+i} F_{-i}} (x_+ c_{\pm i} - y_+ s_{\pm i}) \right]. \quad (4.2b)$$

The fit configuration and sensitivity studies are outlined in Sect. 5.1. The results of the fit to data are discussed in Sect. 5.2. The systematic uncertainties are evaluated in Sect. 5.3 and the results are summarised in Sect. 5.4.

5.1 Fit configuration

A simultaneous extended unbinned maximum likelihood fit to the invariant B mass distribution of candidates falling into each Dalitz plot bin is performed to measure

the values of the CP violation observables, $x_\pm$ and $y_\pm$. The optimal binning is used to divide the $K_S^0 \pi^+ \pi^-$ data, while the 2×2 binning scheme is used to divide the $K_S^0 K^+ K^-$ data. These binning schemes were shown in Figs. 1.8 and 1.9. The data are also separated according to the charge of the bachelor, as the $B^- \rightarrow DK^-$ data are used to determine x_- , y_- and the $B^+ \rightarrow DK^+$ data are used to determine x_+ , y_+ . As with the mass fit in Sect. 3.2, LL and DD K_S^0 candidates are separated. The $B \rightarrow D\pi$ data are used in this fit to determine the size of the misidentified $\pi \rightarrow K$ background in each bin, as described in Sect. 3.2.3. Hence the data are divided into the following 160 categories: $\{B \rightarrow D\pi, B \rightarrow DK\} \otimes \{\text{LL } K_S^0, \text{DD } K_S^0\} \otimes \{2 \times 8 K_S^0 \pi^+ \pi^- \text{ bins}, 2 \times 2 K_S^0 K^+ K^- \text{ bins}\} \otimes \{B^+, B^-\}$.

The fit configuration, which is motivated in the following sub-sections, is summarised as follows. The key concept in this analysis is to relate the relative signal $B \rightarrow DK$ yields in each Dalitz plot bin to the physics parameters, hence the CP observables $x_\pm, y_\pm$ are varied to alter the expected fractions of $B \rightarrow DK$ signal in each bin. The same values of $x_\pm, y_\pm$ determine the yields in each bin for both $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$. The yield of each background component is allowed to vary independently. The shape of each fit component is described in Sect. 3.2. All the shape parameters are fixed to the result of the mass fit, and it is assumed that the binned signal and background shapes are the same as the overall shapes. The values of F_i are Gaussian-constrained to the central values and errors that are listed in Table 4.3. The values of c_i and s_i are fixed in the fit to the CLEO measurements in Ref. [3] and are reproduced in Table 5.1. (The values for an additional $K_S^0 \pi^+ \pi^-$ binning scheme, to be used in Chapter 6, are also shown in Table 5.1 for convenience.) The fit is performed in the range $5150 < m(B) < 5800 \text{ GeV}/c^2$.

Table 5.1. The measurements of c_i and s_i by CLEO in the binning schemes used in this thesis. The first uncertainties are statistical and the second systematic. Reproduced from [3]. A different phase convention is used by CLEO and in this thesis, hence the values of s_i in this table have opposite sign to those in Ref. [3].

i	c_i	s_i
$D \rightarrow K_s^0 \pi^+ \pi^-$ optimal		
1	$-0.009 \pm 0.088 \pm 0.094$	$0.438 \pm 0.184 \pm 0.045$
2	$0.900 \pm 0.106 \pm 0.082$	$0.490 \pm 0.295 \pm 0.261$
3	$0.292 \pm 0.168 \pm 0.139$	$1.243 \pm 0.341 \pm 0.123$
4	$-0.890 \pm 0.041 \pm 0.044$	$0.119 \pm 0.141 \pm 0.038$
5	$-0.208 \pm 0.085 \pm 0.080$	$-0.853 \pm 0.123 \pm 0.035$
6	$0.258 \pm 0.155 \pm 0.108$	$-0.984 \pm 0.357 \pm 0.165$
7	$0.869 \pm 0.034 \pm 0.033$	$0.041 \pm 0.132 \pm 0.034$
8	$0.798 \pm 0.070 \pm 0.047$	$0.107 \pm 0.240 \pm 0.080$
$D \rightarrow K_s^0 \pi^+ \pi^-$ modified optimal		
1	$-0.216 \pm 0.104 \pm 0.088$	$0.399 \pm 0.204 \pm 0.049$
2	$0.827 \pm 0.060 \pm 0.053$	$0.031 \pm 0.172 \pm 0.084$
3	$0.101 \pm 0.085 \pm 0.118$	$0.558 \pm 0.161 \pm 0.070$
4	$-0.955 \pm 0.038 \pm 0.034$	$0.204 \pm 0.137 \pm 0.055$
5	$-0.522 \pm 0.095 \pm 0.079$	$-0.911 \pm 0.130 \pm 0.067$
6	$0.291 \pm 0.102 \pm 0.075$	$-1.030 \pm 0.196 \pm 0.065$
7	$0.682 \pm 0.051 \pm 0.047$	$0.037 \pm 0.146 \pm 0.029$
8	$0.724 \pm 0.071 \pm 0.044$	$0.180 \pm 0.194 \pm 0.05$
$D \rightarrow K_s^0 K^+ K^-$ 2 bins		
1	$0.818 \pm 0.107 \pm 0.037$	$-0.445 \pm 0.215 \pm 0.143$
2	$-0.746 \pm 0.083 \pm 0.035$	$-0.229 \pm 0.220 \pm 0.079$

5.1.1 Mass shapes and fit range

The signal and background shapes used are identical to those in the fit described in Sect. 3.2. Since there are insufficient data available to allow all of the shapes to vary in every Dalitz plot bin, the same shapes are used in each bin. The $m(B)$ distributions of candidates assigned to each bin in simulated $B \rightarrow D\pi$, $D \rightarrow K_s^0 \pi^+ \pi^-$ events are shown in Fig. 5.1 and the mean and rms between the bins is found to be consistent.

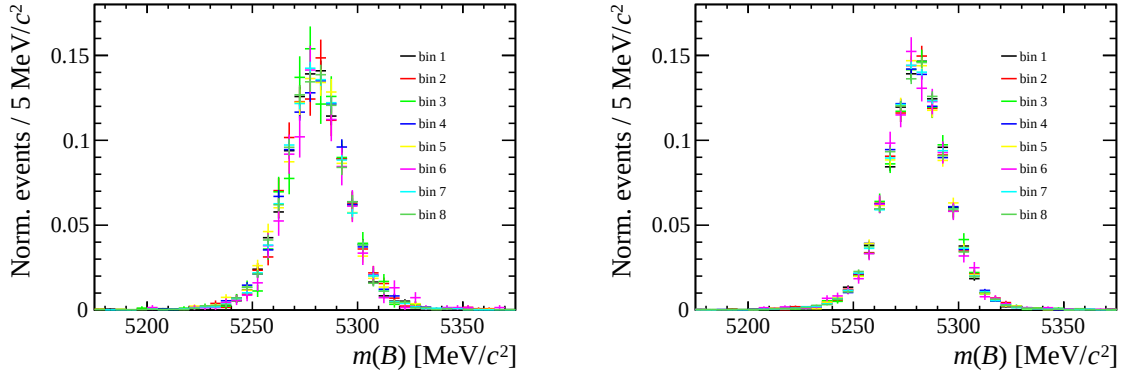


Figure 5.1. The $m(B)$ distributions in simulated $B \rightarrow D\pi$, $D \rightarrow K_S^0\pi^+\pi^-$ events, separated according to their effective bin numbers. The distributions are shown for (left) LL and (right) DD K_S^0 .

However, the partially reconstructed background is described with a combination of functions for different decays. The shape of the total function will vary between bins because the composition of the background changes. Furthermore, CP violation in these backgrounds is not well known. Hence, this fit is performed in a reduced range of $5150\text{--}5800\text{ GeV}/c^2$, where a stable fit can be achieved and only the upper tail of the function for partially reconstructed backgrounds participates in the fit. A systematic uncertainty from assuming that the shapes of the misID and partially reconstructed backgrounds are the same in every Dalitz plot bin is evaluated later.

5.1.2 Yields of signal and background components

The yield of each fit component in a given bin is determined by the relative composition of D^0 mesons, $\bar{D}^0$ mesons and fake D candidates, and is therefore allowed to vary. The total number of $B \rightarrow DK$ candidates in each Dalitz plot, separated by B charge, K_S^0 type and D final state, is measured and the yield of $B \rightarrow DK$ candidates in each bin is given by Eq. (4.2), where the CP observables $x_\pm, y_\pm$ are allowed to vary. The values of F_i are allowed to vary within the Gaussian-constraint given by the central values and errors on F_i given in Table 4.3. A sensitivity study to evaluate the impact of the errors on F_i is

performed in Sect. 5.1.5. The study shows that allowing F_i to vary within the calculated errors leads to only a few-percent increase in the errors of $x_\pm, y_\pm$.

The yields of $B \rightarrow D\pi$ signal candidates in each bin and category are varied independently. It is assumed that the PID efficiencies and misidentification fractions are identical in each bin and these are used to determine the yield of misidentified $B \rightarrow DK$ candidates in the corresponding bin. The combinatorial background yield and partially reconstructed background yield in each of the 160 data categories are free to vary independently.

In bins where the overall yield is very low, the fit may determine that the number of combinatorial or partially reconstructed background events is zero. When the fit prefers yields approaching the limit of zero, the fit may not be able to determine an accurate error matrix. Hence the fit is performed in two steps. Firstly, all of the parameters are allowed to vary, then any partially reconstructed or combinatorial yields less than 0.1 are fixed to zero. A second fit is performed with those yield parameters fixed to zero. When this fitting procedure is used, only about 1% of fits in pseudoexperiments fail to converge with an accurate error matrix. In the fit to data, this procedure affects 18 out of 320 background yield parameters, which typically have fit values of order 10^{-7} .

5.1.3 The likelihood function

The likelihood assigned to each candidate in a given Dalitz bin is constructed from each of the component shapes and yields. The log-likelihood is therefore a sum of the log-likelihoods for each of the 160 data categories and can be written as

$$\log \mathcal{L} = \sum_{B^+, B^-} \sum_{LL, DD} (\log \mathcal{L}^{D\pi} + \log \mathcal{L}^{DK}). \quad (5.1)$$

The terms that contribute to $\log \mathcal{L}^{D\pi}$ are as follows. The signal shape and yield is given by $\mathcal{S}^{D\pi}$ and $N_i^{D\pi}$, respectively. The subscript i denotes the bin number. The combinatorial and partially reconstructed backgrounds have shapes $\mathcal{S}_{\text{CO}}^{D\pi}$ and $\mathcal{S}_{\text{PR}}^{D\pi}$, and yields $N_{i,\text{CO}}^{D\pi}$ and $N_{i,\text{PR}}^{D\pi}$, respectively. Hence the log-likelihood for the $B \rightarrow D\pi$ invariant mass is given by

$$\log \mathcal{L}^{D\pi} = \sum_{\substack{K_S^0 \pi^+ \pi^-, \\ K_S^0 K^+ K^-}} \sum_i \sum_e^{N_{\text{events},i}} \log \left(N_i^{D\pi} \mathcal{S}^{D\pi} + N_{i,\text{CO}}^{D\pi} \mathcal{S}_{\text{CO}}^{D\pi} + N_{i,\text{PR}}^{D\pi} \mathcal{S}_{\text{PR}}^{D\pi} \right), \quad (5.2)$$

where the sum is over each event e in bin i , up to the number of events in the bin, $N_{\text{events},i}$, because the shape terms $\mathcal{S}$ depend on the B mass of event e . The sum is performed over each bin, labelled i , of the Dalitz plot for the two D final states.

The terms that contribute to $\log \mathcal{L}^{DK}$ are as follows. The signal has a shape $\mathcal{S}^{DK}$ and yields N_i^{DK} in each bin, which are not determined independently but are related to the total yield in that category N_{tot}^{DK} . The expected yields are given by expressions of the form $Y_i^+ = F_i + r_B^2 F_{-i} + 2\sqrt{F_i F_{-i}}(x_+ c_i - y_+ s_-)$, where Y_i^+ can be interpreted as the unnormalised fraction of $B \rightarrow DK$ signal in a particular bin. Hence the yield in the likelihood is found by normalising Y_i as follows:

$$N_i^{DK} = N_{\text{tot}}^{DK} \times \frac{Y_i}{\sum_j Y_j}. \quad (5.3)$$

The combinatorial, partially reconstructed and misidentified backgrounds have shapes denoted by $\mathcal{S}_{\text{CO}}^{DK}$, $\mathcal{S}_{\text{PR}}^{DK}$ and $\mathcal{S}_{\text{MIS}}^{DK}$ and yields $N_{i,\text{CO}}^{DK}$, $N_{i,\text{PR}}^{DK}$ and $N_{i,\text{MIS}}^{DK}$, respectively. The yield of misidentified background $N_{i,\text{MIS}}^{DK}$ is fixed to $\frac{\epsilon_{\text{PID}}(\text{DLL}_{K\pi} > 4)}{\epsilon_{\text{PID}}(\text{DLL}_{K\pi} < 4)} \times N_i^{D\pi}$, where ϵ_{PID} are

given in Table 3.3. Hence the log-likelihood for the $B \rightarrow DK$ invariant mass is given by

$$\log \mathcal{L}^{DK} = \sum_{\substack{K_S^0 \pi^+ \pi^-, \\ K_S^0 K^+ K^-}} \sum_i \sum_e^{N_{events,i}} \log(N_i^{DK} \mathcal{S}^{DK} + N_{i,CO}^{DK} \mathcal{S}_{CO}^{DK} + N_{i,PR}^{DK} \mathcal{S}_{PR}^{DK} + N_{i,MIS}^{DK} \mathcal{S}_{MIS}^{DK}), \quad (5.4)$$

where N_i^{DK} is given by Eq. (5.3). In each of Eqs. (5.2) and (5.4), there is a likelihood term corresponding to the Poisson distribution of the number of events in each data category (the ‘extended’ term) that has not been explicitly written.

5.1.4 The construction of r_B

The CP observables are defined $x_\pm = r_B \cos(\delta_B \pm \gamma)$ and $y_\pm = r_B \sin(\delta_B \pm \gamma)$, where the same value of r_B appears in both expressions. Fitting r_B , δ_B and γ directly can produce a bias towards higher values of r_B at low statistics, which is undesirable because the uncertainty $\sigma(\gamma) \propto 1/r_B$ (see Sect. 7.1). Hence, r_B is constructed such that x_+ and y_+ are determined with B^+ data and x_- and y_- are determined with B^- data. This leads to four parameters in the fit constructed from three underlying physics parameters (r_B , δ_B and γ). The four $x_\pm, y_\pm$ parameters are allowed to vary independently in the fit, which corresponds to two effective r_B parameters constructed by $r_{B^+} = \sqrt{x_+^2 + y_+^2}$ and $r_{B^-} = \sqrt{x_-^2 + y_-^2}$. Measuring $x_\pm, y_\pm$ directly produces smaller fitter biases and is preferred by the community for the combination of γ measurements. In addition, having access to two values for r_B provides a test of consistency, as some physics effects such as kaon CP violation and charm mixing would alter these values differently. Hence, a single value of r_B is later extracted from the combined values of $x_\pm, y_\pm$ in Chapter 7.

5.1.5 Expected statistical sensitivity

The expected statistical uncertainties on $x_\pm, y_\pm$ are investigated by generating pseudo-experiments. Realistic pseudoexperiments are generated assuming that the physics parameters are $r_B = 0.1$, $\gamma = 70^\circ$ and $\delta_B = 130^\circ$. The fraction of $B \rightarrow DK$ events per bin is calculated using Eq. (4.2). The fraction of $B \rightarrow D\pi$ events per bin is equal to F_i , which is a good approximation as the level of interference in this decay is very small. The fraction of partially reconstructed background events per bin is also equal to F_i because the decays contributing to the B^- sample are dominated by channels containing a D^0 . The misID background per bin is the $B \rightarrow D\pi$ yield in that bin multiplied by the PID efficiencies. The combinatorial background is distributed uniformly over the Dalitz plot. The extracted values of $x_\pm, y_\pm$ are similar for different assumptions of the Dalitz plot distribution of this background. The B mass distributions generated are based on the shapes in Sect. 3.2. Each pseudoexperiment is then fitted using the fit configuration described in this section.

The fit is performed using two methods to compare the expected sensitivity when the parameters F_i are either fixed or Gaussian-constrained. The first reproduces the exact method of the data fit, where the F_i are Gaussian-constrained to their determined central values and errors. The second method fixes the values of F_i in the fit. In both cases, 400 datasets are generated. The total signal and background yields generated in each category are obtained from the results of the fit in Sect. 3.2. With these yields, the error on the expected statistical uncertainty is approximately 0.0015. The expected statistical uncertainties for each type of fit are given in Table 5.2, which demonstrates that there is no significant difference in the expected sensitivity between the two fit scenarios. This is unsurprising since the errors on F_i are of order 2%. The study also indicates that reducing the uncertainties on F_i by increasing the size of simulation samples will not alter the sensitivity of the analysis and that the statistical uncertainty

Table 5.2. Results of the sensitivity study to $x_\pm, y_\pm$. The statistical uncertainties are shown when two different configurations of the fit are performed.

Parameter	F_i Gaussian-constrained	F_i fixed
$\sigma(x_+)$	0.026	0.026
$\sigma(x_-)$	0.025	0.024
$\sigma(y_+)$	0.032	0.032
$\sigma(y_-)$	0.030	0.029

is driven by the $B \rightarrow DK$ yields. It is preferable to allow F_i to vary under a Gaussian-constraint as it allows the $B \rightarrow Dh$ data to provide further information about residual differences in the model and efficiency. The distribution of central values, statistical errors and pulls on $x_\pm, y_\pm$ when F_i are fixed are shown in Fig. 5.2.

5.1.6 Fitter bias

A small bias in the fits to pseudoexperiments is observed in Fig. 5.2, where the mean fitted central values on $x_\pm, y_\pm$ differ from the input values used to generate the data. Defining the bias as the difference between the mean of the fitted central values and the input value of the pseudoexperiments, the biases on x_+, x_-, y_+ and y_- are +0.001, -0.001, -0.003 and -0.003, respectively. The bias occurs because of the low yields, causing signal or background in some bins to be consistent with zero. It has been demonstrated that these biases reduce in size with larger datasets [41, 86]. The quoted fit results of Sect. 5.2 are corrected by this offset. Half of the offset is taken as a systematic uncertainty (Sect. 5.3). An additional systematic uncertainty is considered due to the dependence of the correction on the input values of $x_\pm, y_\pm$.

In order to verify if the fit provides an adequate estimate of the statistical uncertainties, the pull distribution for each variable in the pseudoexperiments is shown in Fig. 5.2. The pull is defined as $(x_{\text{fit}} - x_{\text{gen}})/\sigma_{\text{fit}}$, where x_{fit} is the fitted value, x_{gen} is the input value and σ_{fit} is the fitted uncertainty. The pull distributions are fitted with a

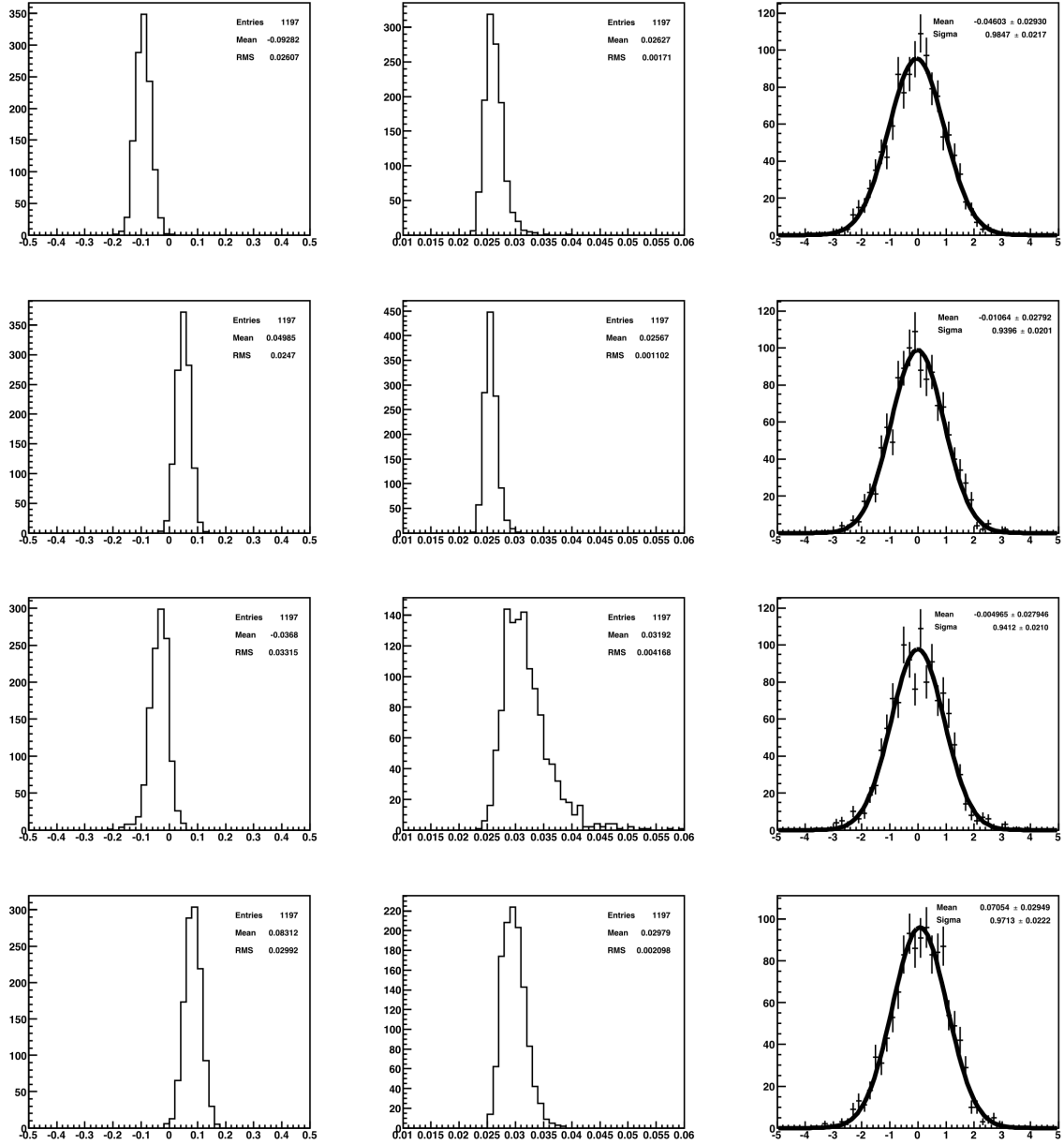


Figure 5.2. The result of performing the fit using pseudoexperiments generated with the CP values $x_+ = -0.094$, $x_- = 0.050$, $y_+ = -0.034$ and $y_- = 0.087$. The fitted (left) central values, (middle) statistical uncertainties and (right) pulls are shown for (top to bottom) x_+ , x_- , y_+ and y_- . The pull distributions are fitted to a Gaussian and the fit results are shown.

Gaussian function and it is found that their widths are consistent with unity, hence no correction is applied to the fitted statistical uncertainties.

5.2 Results of the binned fit

The fit is performed on data and the CP violation observables are:

$$\begin{aligned} x_+ &= -0.077 \pm 0.024, \quad y_+ = -0.022 \pm 0.025, \\ x_- &= 0.025 \pm 0.025, \quad y_- = 0.075 \pm 0.029. \end{aligned}$$

The central values have been corrected for the fitter bias discussed in Sect. 5.1.6. The fit projections for DD K_S^0 are shown in Figs. 5.3–5.6, where the corresponding bins in the B^- and B^+ Dalitz plots are shown side-by-side. The expected behaviour for a non-zero value of γ is that the yields of oppositely-signed bin numbers of B^+ and B^- data will differ, as given by Eq. (1.40). Differences in the $B \rightarrow DK$ yields can be observed between pairs of bins, for instance when comparing bin +4 of $B^- \rightarrow DK^-$ and bin -4 of $B^+ \rightarrow DK^+$. Similar results are observed in the fit projections for LL K_S^0 (Figs. 5.7–5.10).

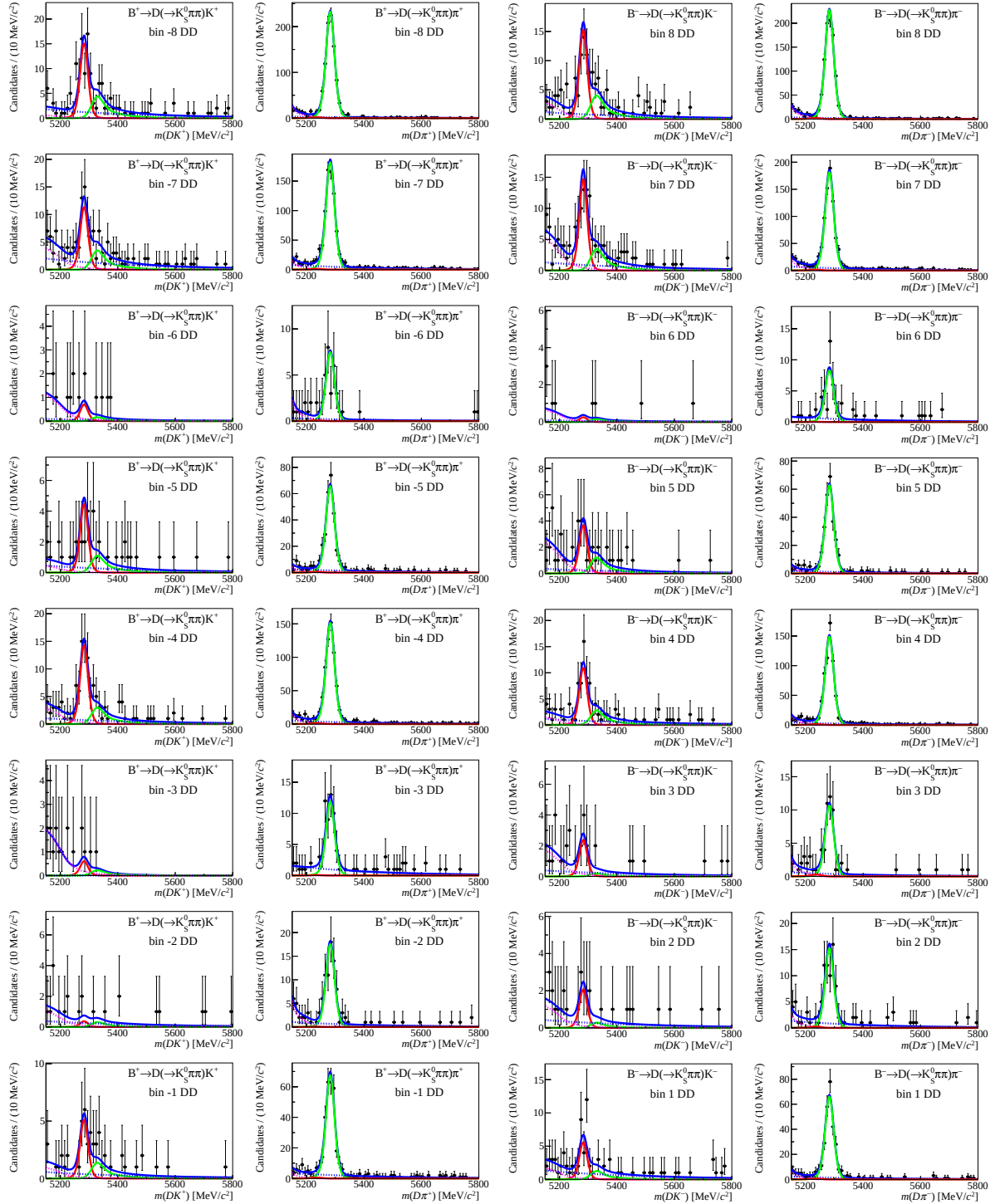


Figure 5.3. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ DD data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) -8 to -1 and the two right-hand columns show B^- data in bin numbers (top to bottom) +8 to +1. The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

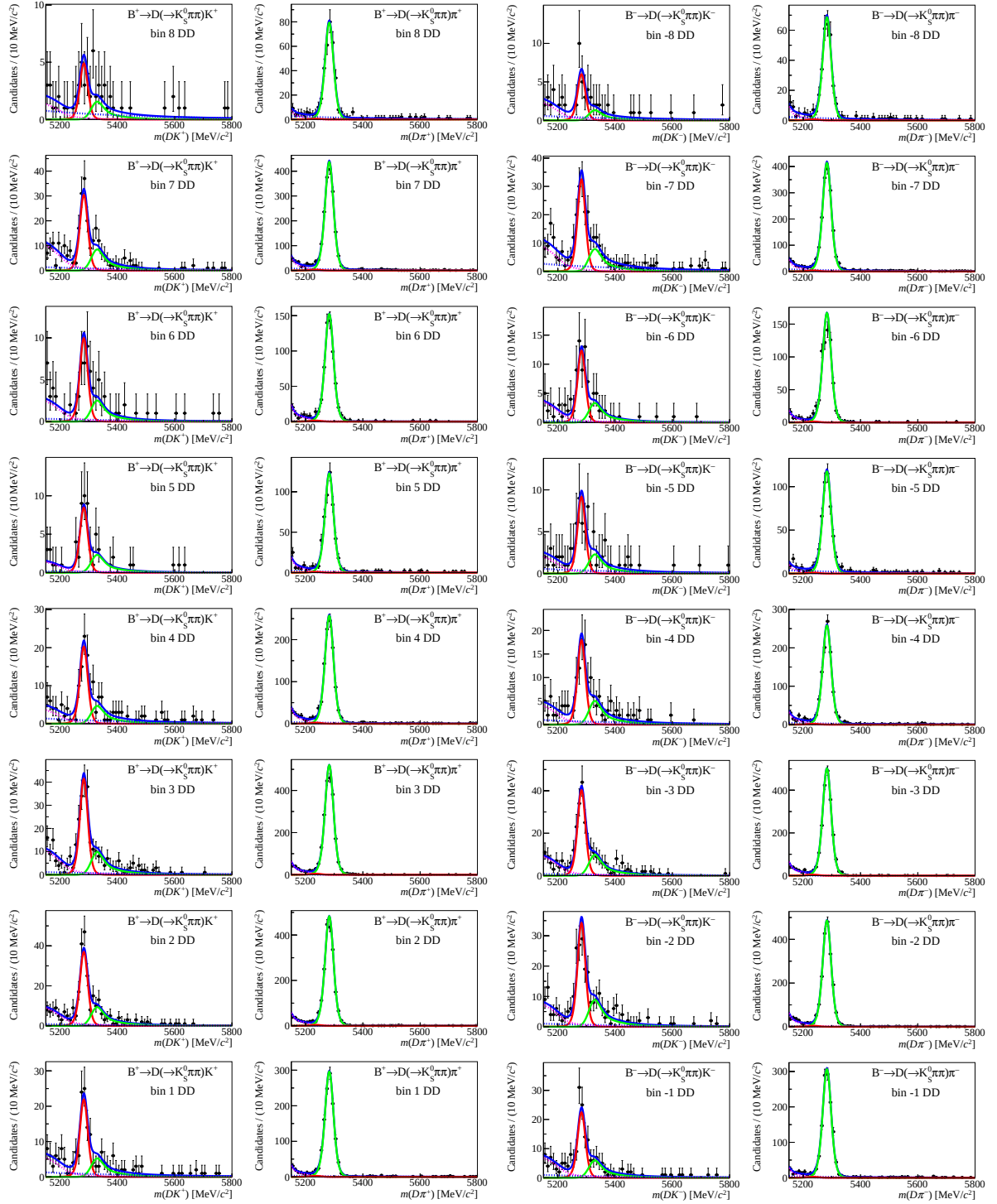


Figure 5.4. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ DD data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) +8 to +1 and the two right-hand columns show B^- data in bin numbers (top to bottom) -8 to -1. The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

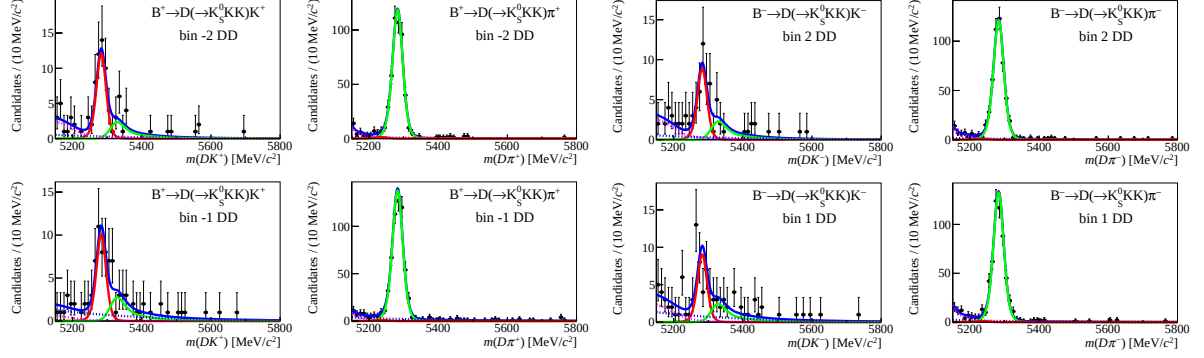


Figure 5.5. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 K^+ K^-$ DD data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) -2 to -1 and the two right-hand columns show B^- data in bin numbers (top to bottom) $+2$ to $+1$. The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

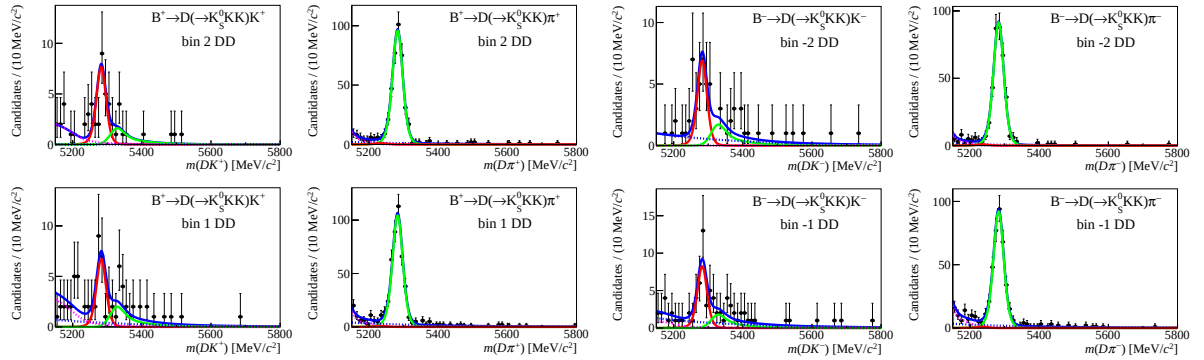


Figure 5.6. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 K^+ K^-$ DD data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) $+2$ to $+1$ and the two right-hand columns show B^- data in bin numbers (top to bottom) -2 to -1 . The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

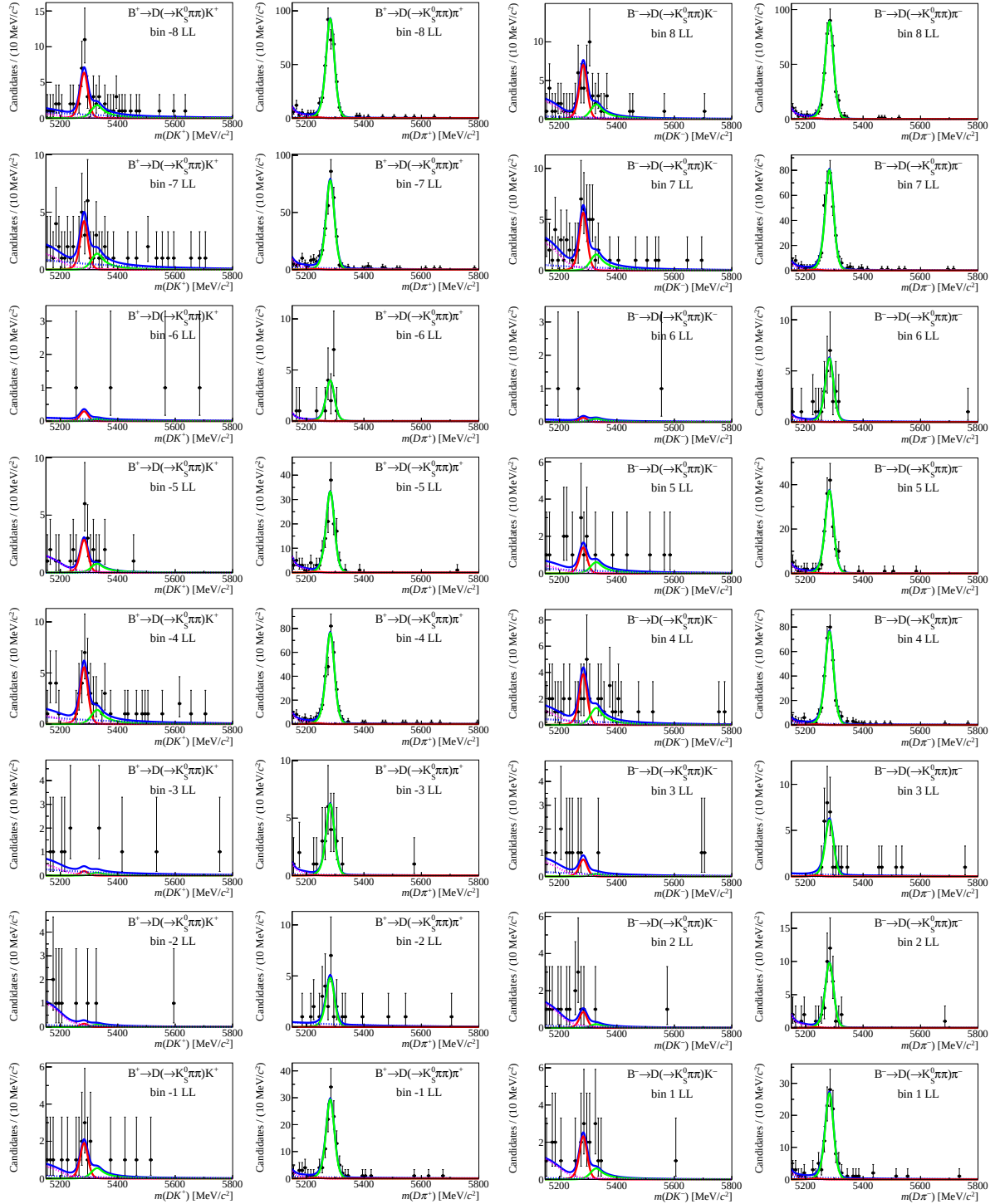


Figure 5.7. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ LL data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) -8 to -1 and the two right-hand columns show B^- data in bin numbers (top to bottom) +8 to +1. The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

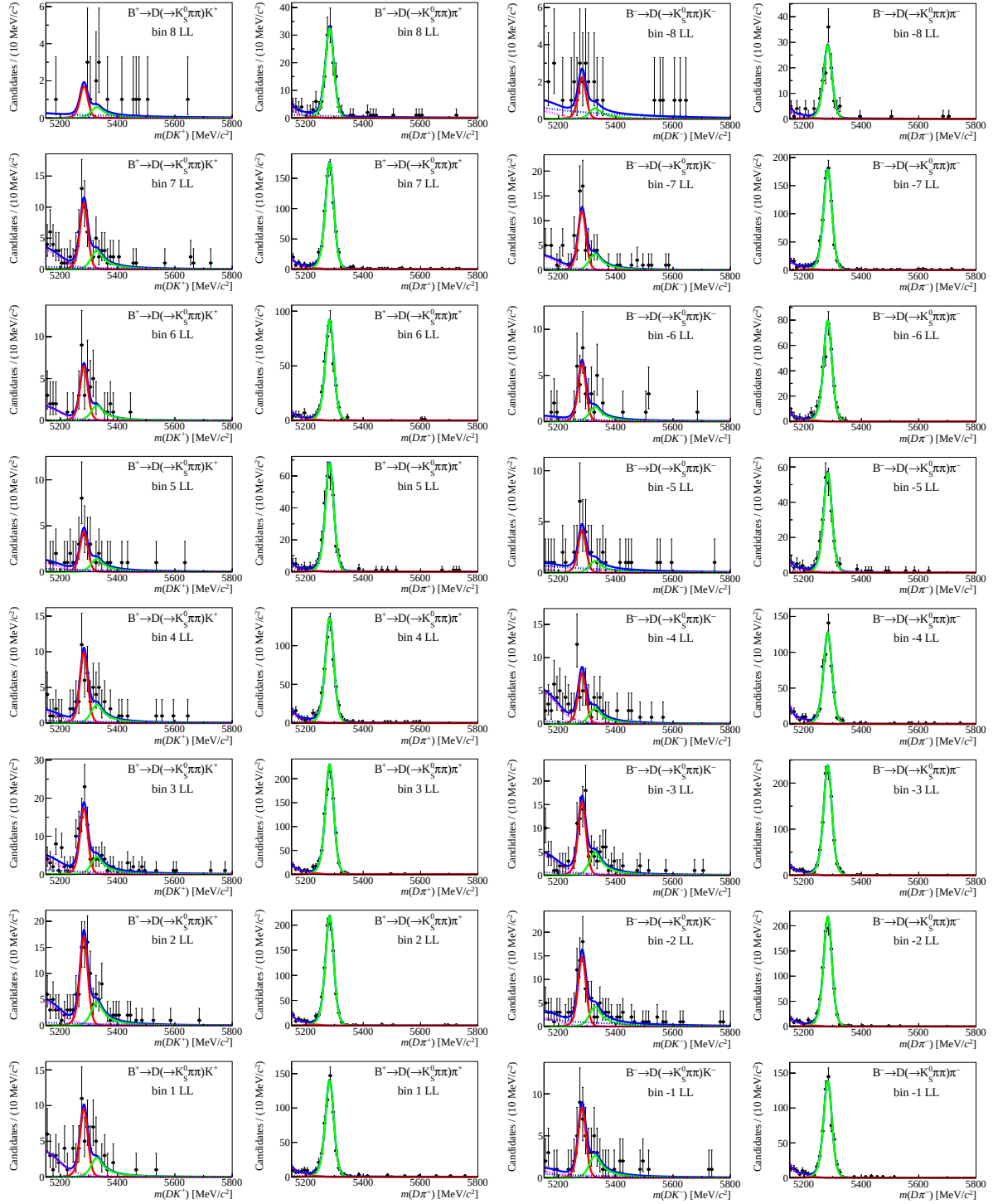


Figure 5.8. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 \pi^+ \pi^-$ LL data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) +8 to +1 and the two right-hand columns show B^- data in bin numbers (top to bottom) -8 to -1. The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

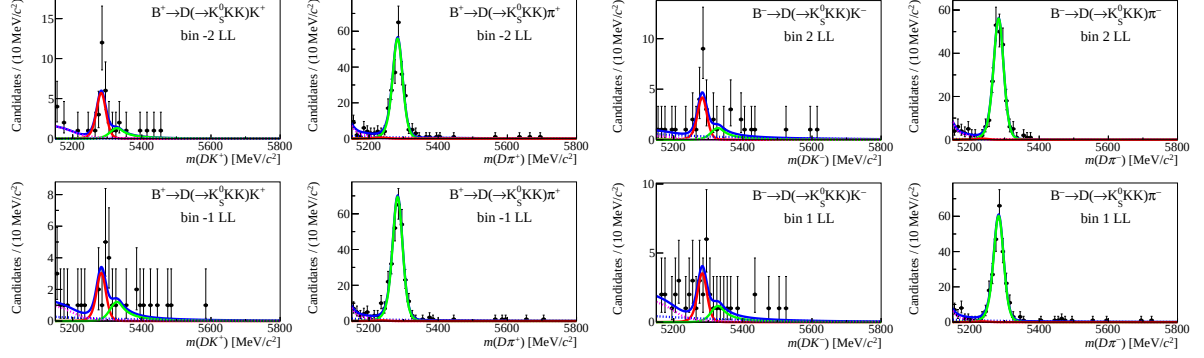


Figure 5.9. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 K^+ K^-$ LL data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) +2 to +1 and the two right-hand columns show B^- data in bin numbers (top to bottom) -2 to -1. The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

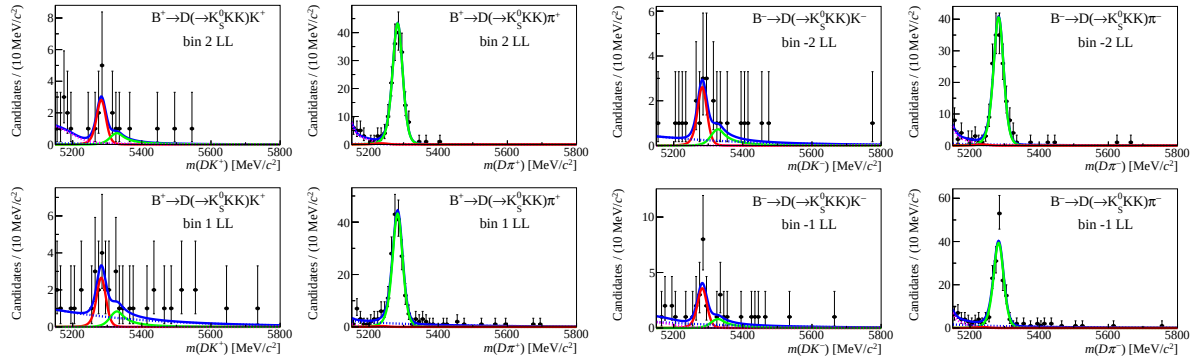


Figure 5.10. The fit projections for $B \rightarrow DK$ and $B \rightarrow D\pi$, $D \rightarrow K_S^0 K^+ K^-$ LL data are shown. The two left-hand columns show B^+ data in bin numbers (top to bottom) +2 to +1 and the two right-hand columns show B^- data in bin numbers (top to bottom) -2 to -1. The (blue) fit result is superimposed with the following fit components: (red) $B \rightarrow DK$, (green) $B \rightarrow D\pi$, (dashed blue) combinatorial background and (dashed magenta) partially reconstructed background.

5.2.1 Two-dimensional confidence level contours

A confidence level is computed for each pair of (x, y) parameters using the statistical uncertainties only. The confidence level is determined by evaluating the $\Delta \log \mathcal{L}$ at each point of $x_\pm, y_\pm$ -space, where $\Delta \log \mathcal{L}$ is the difference in log-likelihood between a given point and the minimum log-likelihood value. The $\Delta \log \mathcal{L}$ values for (x_+, y_+) are evaluated by running the fit with (x_+, y_+) fixed to a point of interest, with (x_-, y_-) fixed to their central values, and with all other fit parameters floating. The $\Delta \log \mathcal{L}$ values for (x_-, y_-) are evaluated similarly but with (x_+, y_+) fixed to their central values. The equal-confidence-level contours are drawn at $\Delta \log \mathcal{L} = 1.15, 3.09$ and 5.91 , corresponding to $\Delta \chi^2 = 1, 4$ and 9 for two degrees of freedom. The resulting plot is shown in Fig. 5.11. This procedure assumes that the pairs of (x_-, y_-) and (x_+, y_+) are uncorrelated, which is a good assumption given the correlation matrix between the parameters shown later in Sect. 5.4. In order to verify whether the contours for one (x, y) pair are affected by fixing the values of the opposite pair, an alternative scan is performed without fixing the opposite pair and it is found that the difference between the two scans is negligible. The expected signature for CP violation is that the vectors defined by the coordinates (x_+, y_+) and (x_-, y_-) are non-zero in magnitude and have a non-zero opening angle. This opening angle is equal to 2γ .

5.2.2 Alternative fit to measure the yields

The default fit measures only the total $B \rightarrow DK$ yield in each category and varies the values of $x_\pm, y_\pm$ and F_i to determine the distribution of yields in each Dalitz plot bin. As a cross-check, the fit is reconfigured to measure the yield of $B \rightarrow DK$ candidates in each Dalitz plot bin directly. The yields calculated from the measured $x_\pm, y_\pm$ are compared with the yields measured in the alternative fit. The results (LL and DD K_S^0 combined) are shown in Fig. 5.12 where the sum of signal yields in B^+ and B^- , $N_{B^+} + N_{B^-}$, and

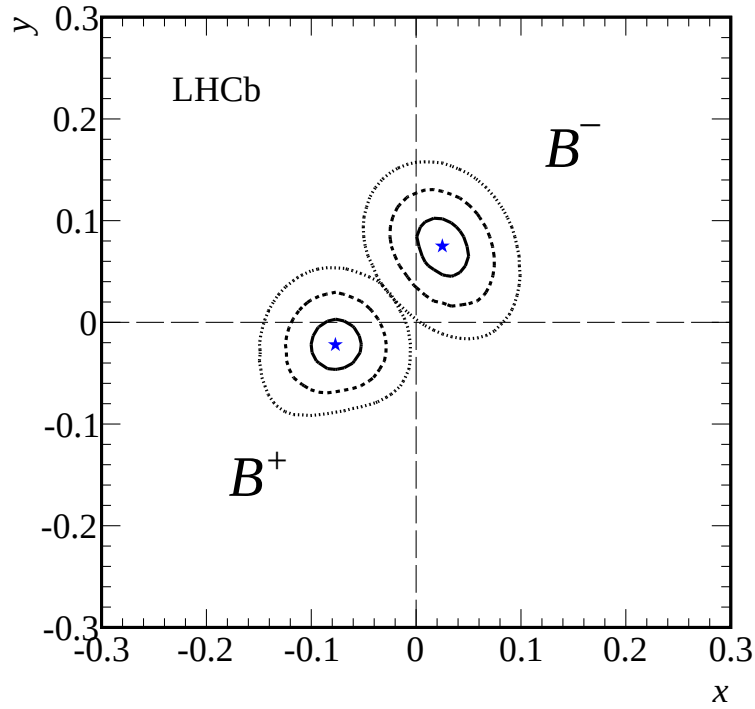


Figure 5.11. Confidence level contours for the statistical uncertainties on (x_+, y_+) and (x_-, y_-) . The blue stars indicate the best fit central values and the bands, extending outwards from the centre, correspond to the 1, 2 and 3σ intervals at 39.3%, 86.5% and 98.9% probability.

the difference in yields, $N_{B^+} - N_{B^-}$, are shown. In both cases good agreement between the calculated and measured yields are observed, which indicates that the fit results provide a good description of the data. The difference in B^+ and B^- yields is sensitive to CP violation and the expected signal yields assuming CP symmetry ($x_\pm = y_\pm = 0$) in this distribution is superimposed. It should be noted that the CP -symmetric distribution is not constant at $N_{B^+} - N_{B^-} = 0$ in every bin because the expected $N_{B^\pm}$ are calculated using the total B^+ and B^- yields, which are not identical. The data and fit expectations yield χ^2 probabilities (p -values) of 93% for $N_{B^+} + N_{B^-}$ and 80% for $N_{B^+} - N_{B^-}$. The results for the $N_{B^+} - N_{B^-}$ distribution are less compatible with the hypothesis of CP symmetry, which has a χ^2 probability of 4%.

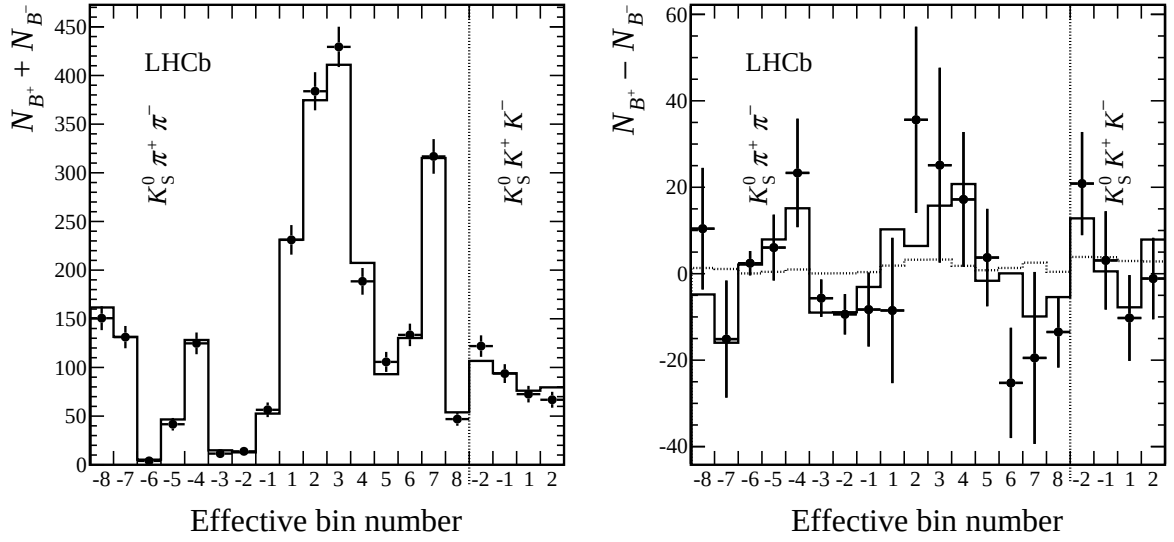


Figure 5.12. The signal yields (data points) measured in the alternative fit are compared with the predicted yields from the measured $x_\pm, y_\pm$ (solid histogram) for $D \rightarrow K_S^0 \pi^+ \pi^-$ and $D \rightarrow K_S^0 K^+ K^-$ decays. The left plot shows the sum of B^+ and B^- yields and the right plot shows the difference between B^+ and B^- yields. The dotted histogram on the right gives the prediction for $x_\pm = y_\pm = 0$.

5.3 Systematic uncertainties

Systematic uncertainties on $x_\pm$ and $y_\pm$ are evaluated in this section. Unless otherwise stated each systematic uncertainty is determined using pseudoexperiments, which are generated using an alternative configuration and fitted with the default fit configuration. The mean shift in the fitted values of $x_\pm$ and $y_\pm$ with respect to their input values is assigned as the systematic uncertainty. A summary is given at the end of the section.

Systematic from the CLEO uncertainties on c_i and s_i

In the default fit, the central values of the measured c_i and s_i at CLEO are used to measure $x_\pm$ and $y_\pm$. The systematic uncertainty associated with the CLEO uncertainties is evaluated by performing multiple fits to the data, where a different set of c_i and s_i are used in each fit. The alternative values of c_i and s_i are obtained by smearing the central values by their measured statistical and systematic uncertainties, taking into account

their correlations [3]. The systematic uncertainty associated with each parameter is assigned by taking the width of the measured values from 400 fits. For x_+ , x_- , y_+ and y_- this is found to be 0.004, 0.005, 0.010 and 0.014, respectively.

Effect of a non-uniform efficiency on the values of c_i and s_i

The measurements of c_i and s_i used in the fit correspond to the definitions in Eqs. (1.36) and (1.37), which do not take variations in the Dalitz plot efficiency within a bin into account. If the efficiency profile η_{DK} is considered, this modifies the definition of c_i to

$$c_i^{\text{corr}} \equiv \frac{\int_i dm_-^2 dm_+^2 \eta_{DK} a \bar{a} \cos \delta_D}{\sqrt{\int_i dm_-^2 dm_+^2 \eta_{DK} a^2} \sqrt{\int_i dm_-^2 dm_+^2 \eta_{DK} \bar{a}^2}}. \quad (5.5)$$

A similar definition for s_i^{corr} can be written by replacing $\cos \delta_D \rightarrow \sin \delta_D$. The measurements of $x_\pm, y_\pm$ are affected by any net shifts in the values of c_i and s_i caused by η_{DK} , which can be estimated using the BaBar amplitude models. The shifts are estimated by calculating $\Delta c_i = c_i^{\text{corr}} - c_i^{\text{model}}$ and $\Delta s_i = s_i^{\text{corr}} - s_i^{\text{model}}$. The values of c_i^{model} and s_i^{model} are calculated with Eqs. (1.36) and (1.37), using the intensity and phase values provided by BaBar. The values of c_i^{corr} and s_i^{corr} are calculated using simulated $B \rightarrow D\pi$ decays to provide the values of η_{DK} . The results are summarised in Table 5.3. The systematic uncertainty is estimated by fitting the data many times. The shifts Δc_i and Δs_i are taken as two-sided errors and assumed to be uncorrelated. In order to obtain a conservative estimate, the larger of Δc_i (DD) and Δc_i (LL) for bin i is used to smear c_i , and the same strategy is used to smear s_i . (The fits to data are performed similarly to the default, where the same values of c_i and s_i apply to the K_S^0 DD and LL data.) The systematic uncertainties on x_+ , x_- , y_+ and y_- due to the use of uncorrected c_i and s_i values are found to be very small, with values of 0.0002, 0.0002, 0.0008 and 0.0003, respectively.

Table 5.3. Values of c_i and s_i used to measure the systematic uncertainty from the effect of a non-uniform efficiency on c_i and s_i . The values of c_i^{model} , c_i^{corr} and Δc_i , and the equivalent s_i parameters are defined in the main text. The values and statistical uncertainties of c_i and s_i measured by CLEO are included for comparison.

i	CLEO c_i	c_i^{model}	c_i^{corr} (DD)	Δc_i (DD)	c_i^{corr} (LL)	Δc_i (LL)
$K_S^0 \pi^+ \pi^-$						
1	-0.009 ± 0.088	-0.029	-0.024	0.005	-0.029	0.000
2	0.900 ± 0.106	0.857	0.868	0.011	0.867	0.010
3	0.292 ± 0.168	0.182	0.166	-0.017	0.174	-0.009
4	-0.890 ± 0.041	-0.925	-0.921	0.003	-0.925	-0.000
5	-0.208 ± 0.085	-0.276	-0.294	-0.018	-0.275	0.001
6	0.258 ± 0.155	0.305	0.296	-0.009	0.309	0.005
7	0.869 ± 0.034	0.873	0.875	0.003	0.875	0.002
8	0.798 ± 0.070	0.864	0.862	-0.002	0.860	-0.004
$K_S^0 K^+ K^-$						
1	0.818 ± 0.107	0.739	0.733	-0.006	0.726	-0.013
2	-0.746 ± 0.083	-0.697	-0.704	-0.006	-0.682	0.016
i	CLEO s_i	s_i^{model}	s_i^{corr} (DD)	Δs_i (DD)	s_i^{corr} (LL)	Δs_i (LL)
$K_S^0 \pi^+ \pi^-$						
1	0.438 ± 0.184	0.816	0.816	0.000	0.815	-0.001
2	0.490 ± 0.295	0.141	0.153	0.012	0.149	0.008
3	1.243 ± 0.341	0.879	0.884	0.006	0.886	0.007
4	0.119 ± 0.141	0.080	0.087	0.006	0.082	0.001
5	-0.853 ± 0.123	-0.827	-0.823	0.004	-0.823	0.004
6	-0.984 ± 0.357	-0.829	-0.824	0.005	-0.826	0.002
7	0.041 ± 0.132	-0.197	-0.193	0.005	-0.196	0.001
8	0.107 ± 0.240	0.318	0.324	0.006	0.327	0.010
$K_S^0 K^+ K^-$						
1	-0.445 ± 0.215	-0.265	-0.262	0.003	-0.264	0.001
2	-0.229 ± 0.220	-0.332	-0.313	0.019	-0.309	0.023

Efficiency corrections in the calculation of F_i

The calculation of F_i includes an efficiency correction that is calculated with simulated $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays and $D^0 \rightarrow K_S^0 h^+ h^-$ amplitude models. There are three sources of bias that can lead to deviations from the true efficiency correction factors ξ_i (defined in Sect. 4.3.3). The first source is the imperfect description in the simulations, which can arise from the characterisation of the initial proton-proton

collision, the kinematic distributions of the signal decay or the reconstruction of particles in the LHCb detector. Furthermore, changes in trigger settings over the full data-taking period are not fully reproduced in simulations. Since the calculation of ξ_i involves a ratio between the results of simulated $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays, the differences between simulation and data should cancel to first order. The second source of bias in the relative efficiency profiles comes from kinematic criteria imposed on the generated candidates prior to reconstruction, which resulted in the removal of events that would otherwise pass the nominal selection. These were necessary to reduce CPU requirements when generating a large number of signal events. These criteria are found to affect the efficiencies of the two decay modes differently. The third source, due to deficiencies in the intensities from the amplitude models, is known to be small because of the large sample size of D^0 decays used in Ref. [33, 34], and the good fit quality of the models to the data.

The discrepancies between simulation and data come from many effects and it is not possible to determine each of these and their correlations exhaustively. Instead, the correction terms $\eta_C(m_-^2, m_+^2)_{D\pi}$ and $\eta_C(m_-^2, m_+^2)_{D^*\mu}$ are used to describe the net effect of all three sources of bias on $\eta(m_-^2, m_+^2)_{D\pi}$ and $\eta(m_-^2, m_+^2)_{D^*\mu}$, respectively. Hence, the true value of ξ_i is

$$\xi_i^{\text{true}} = \frac{\int_i dm_-^2 dm_+^2 \eta_C(m_-^2, m_+^2)_{D\pi} \eta_{MC}(m_-^2, m_+^2)_{D\pi} a^2(m_-^2, m_+^2)_{\text{Babar}}}{\int_i dm_-^2 dm_+^2 \eta_C(m_-^2, m_+^2)_{D^*\mu} \eta_{MC}(m_-^2, m_+^2)_{D^*\mu} a^2(m_-^2, m_+^2)_{\text{Babar}}}, \quad (5.6)$$

where $a^2(m_-^2, m_+^2)_{\text{Babar}}$ are the intensities from the BaBar amplitude model and η_{MC} are the efficiencies in MC. Although it is not possible to exactly determine $\eta_C(m_-^2, m_+^2)_{D\pi}$ or $\eta_C(m_-^2, m_+^2)_{D^*\mu}$ at each point on the Dalitz plot, it is possible to estimate their average values over a given region by comparing the yield directly measured in data, N_{data} , with the yield predicted by the simulated efficiency profile weighted by the appropriate intensity model, N_{pred} . This is performed in a rectangular binning scheme shown in Fig. 5.13, which allows regions of similar efficiency to be studied. In bin j of the

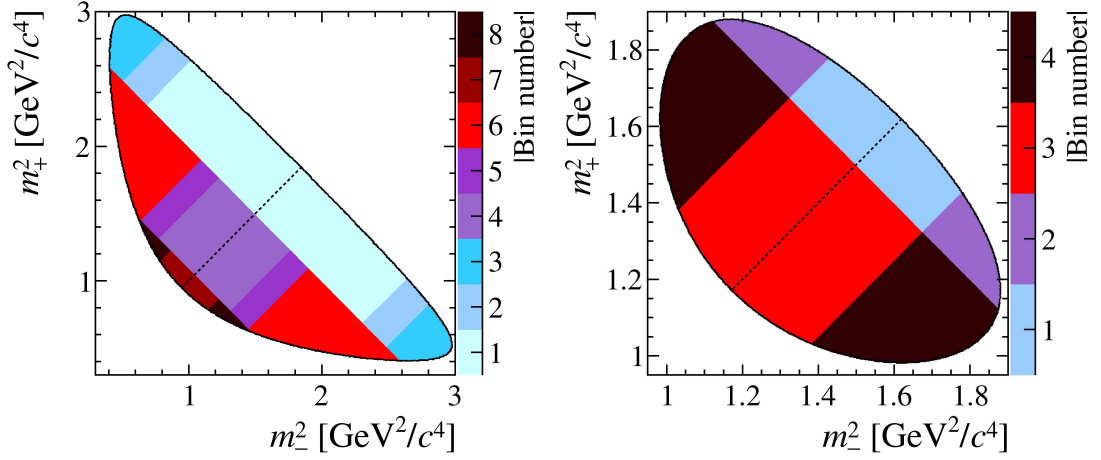


Figure 5.13. The rectangular binning schemes for the (left) $D \rightarrow K_s^0 \pi^+ \pi^-$ and (right) $D \rightarrow K_s^0 K^+ K^-$ decays. These schemes are used to evaluate systematic uncertainties. The dotted diagonal line separates the positive and negative bin numbers, where the negative bins are in the region $m_-^2 \leq m_+^2$.

rectangular binning scheme, the ratio $N_{\text{data}}/N_{\text{predicted}}$ is denoted by C_j . Figure 5.14 shows the ratio C_j for each bin of $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays for the $K_s^0 \pi^+ \pi^-$ final state. Values of C_j differing from unity by a maximum of $\pm 15\%$ are observed. Similar patterns are observed in all of the data categories, which show that the Dalitz plot variation along the diagonal defined by $m_+^2 = m_-^2$ has on average been overestimated in simulation. This trend also reflects the direction in which the Dalitz plot efficiency varies most significantly.

An alternative set of ξ_i values are calculated by modifying Eq. (4.9) to

$$\xi_i^{\text{alt}} = \frac{\sum_j (n_j^{D\pi} I_j C_j^{D\pi})}{\sum_j (n_j^{D^*\mu} I_j C_j^{D^*\mu})}. \quad (5.7)$$

Using ξ_i^{alt} , a second set of F_i is determined. The default and alternative F_i values are compared in Fig. 5.15. Only small differences remain because the data-simulation differences seen in Fig. 5.14 have been diluted by the irregular shapes of the bins in the optimal and equal-phase binning used for $K_s^0 \pi^+ \pi^-$ and $K_s^0 K^+ K^-$, respectively.

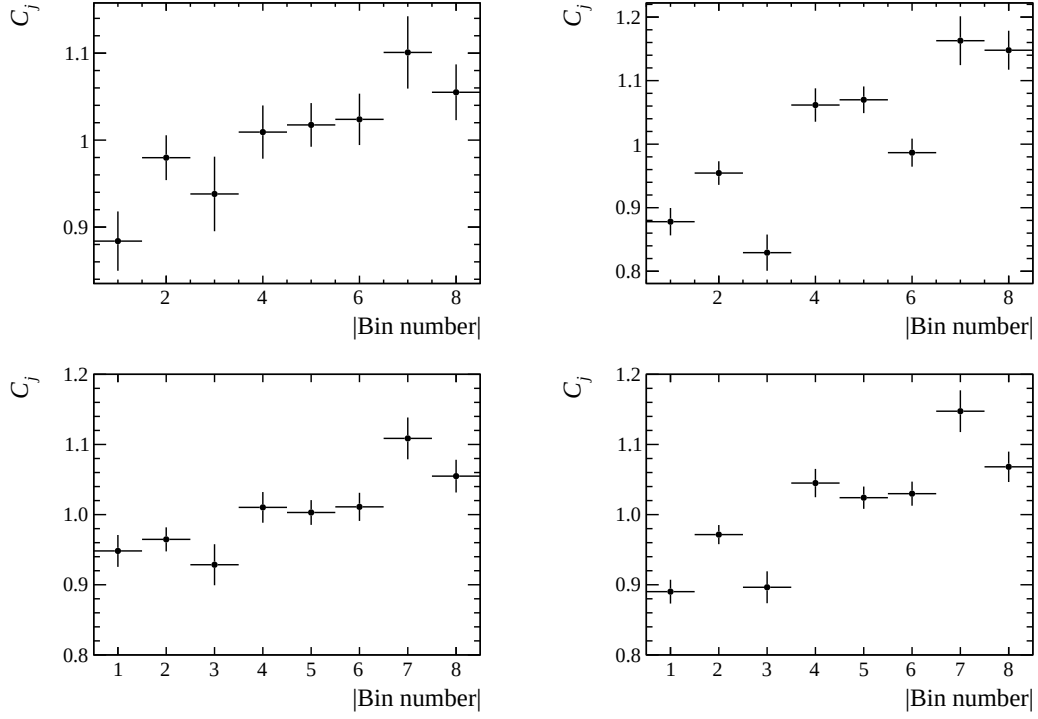


Figure 5.14. The values of C_j , which are the ratio between yields directly measured in data and yields predicted by simulation in $D \rightarrow K_s^0 \pi^+ \pi^-$ decays, DD K_s^0 , for (left) $B \rightarrow D\pi$ and (right) $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ in (top) 2011 and (bottom) 2012 conditions. The bin numbers refer to the rectangular binning.

When the statistical uncertainties on C_j are propagated, the variations in the alternative ξ_i are an order of magnitude smaller than $|\xi_i^{\text{alt}} - \xi_i|$. The systematic uncertainty is determined with pseudoexperiments generated with the alternative values of F_i and the uncertainties on x_+ , x_- , y_+ and y_- are 0.009, 0.009, 0.002 and 0.002, respectively.

Effect of assuming $\eta(m_-^2, m_+^2) = \eta(m_+^2, m_-^2)$

In the calculation of F_i and in the application of the same F_i values to $B^+ \rightarrow DK^+$ and $B^- \rightarrow DK^-$ data, it has been assumed that $\eta(m_-^2, m_+^2) = \eta(m_+^2, m_-^2)$, which can be broken in the presence of a charge detection asymmetry affecting the D daughters. The size of the possible asymmetry is estimated by comparing the fractional yields in bins numbered with opposite signs, which corresponds to events situated on mirror-opposite points of the Dalitz plot. The effect is studied in $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ data where a large yield with low

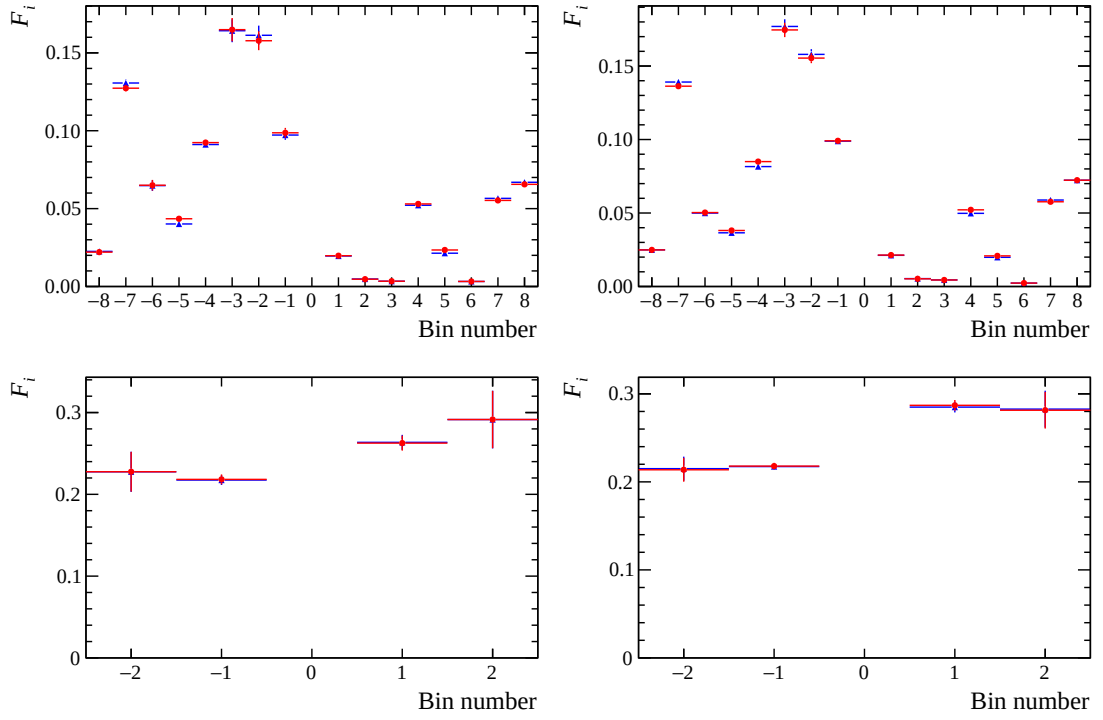


Figure 5.15. The (blue triangle) default values of F_i are compared with the (red circle) alternative values of F_i used to compute the systematic uncertainty associated with the efficiency correction. The values are shown for candidates in (left) LL and (right) DD K_S^0 from (top) $D \rightarrow K_S^0 \pi^+ \pi^-$ and (bottom) $D \rightarrow K_S^0 K^+ K^-$ decays.

background contamination is available. The fractional yields in oppositely-signed bins are consistent within the statistical uncertainties, and the same conclusion is drawn with simulation. Hence the presence of any charge asymmetry affecting F_i is much smaller than the statistical fluctuations in $B \rightarrow DK$ data that the measurement of $x_\pm, y_\pm$ is subject to. The assumption that $\eta(m_-^2, m_+^2) = \eta(m_+^2, m_-^2)$ is shown to be valid and a systematic uncertainty is not assigned.

Systematic uncertainty due to the fitter bias

A fitter bias has been evaluated in Sect. 5.1.6, where the low yields in the data sample result in the fitted signal yields in some bins to be consistent with zero. The fitter bias used to correct the measured central values was evaluated with $r_B = 0.1$, $\gamma = 70^\circ$ and $\delta_B = 130^\circ$. However, the size of the bias depends on the input values. Four alternate

inputs are considered, where $\gamma \in [60^\circ, 90^\circ]$, $r_B \in [0.08, 0.10]$ and $\delta_B \in [120^\circ, 130^\circ]$. The size of the bias on $x_\pm, y_\pm$ when different inputs are used varies by up to 0.0015–0.0020. The uncertainty on the bias because of the finite number of pseudoexperiments is 0.0007. The systematic uncertainty is calculated by adding in quadrature the spread of the biases, the uncertainty on the bias from the finite number of pseudoexperiments and half the size of the nominal fitter bias, and is found to be 0.002 for all four CP observables.

Systematic uncertainty from the B mass parameterisation

The systematic uncertainty from the B mass parameterisation of the signal and misID fit components is evaluated. Inaccurate descriptions of the signal or backgrounds can change the relative signal yields measured by the fit, and therefore the best fit values of $x_\pm, y_\pm$. This uncertainty is estimated by considering an alternative parameterisation for each fit component. An uncertainty for the combinatorial background shape is not considered because its slope is Gaussian constrained in the fit (see Sect. 3.2.2).

The alternative shape for the $B \rightarrow Dh$ signal is the sum of a Crystal Ball and a Gaussian function. The alternative signal shape obtained from fitting the data is compared to the default Cruijff function in Fig. 5.16. Although the alternative appears to prefer a larger width in DD K_S^0 , it does not result in a significant change in the fit χ^2 . The associated systematic uncertainties on x_+ , x_- , y_+ and y_- are 0.0003, 0.0001, 0.0009 and 0.0005, respectively. As a cross-check, the systematic uncertainty is also estimated by performing many data fits, where the Cruijff shape parameters are varied according to their uncertainties and correlations measured in Sect. 3.2. Systematic uncertainties of comparable size are observed using this method.

The shape for $B \rightarrow D\pi$ decays that have been misidentified as $B \rightarrow DK$ is fixed in the fit. Slight inaccuracies in the shape can come from imperfect background

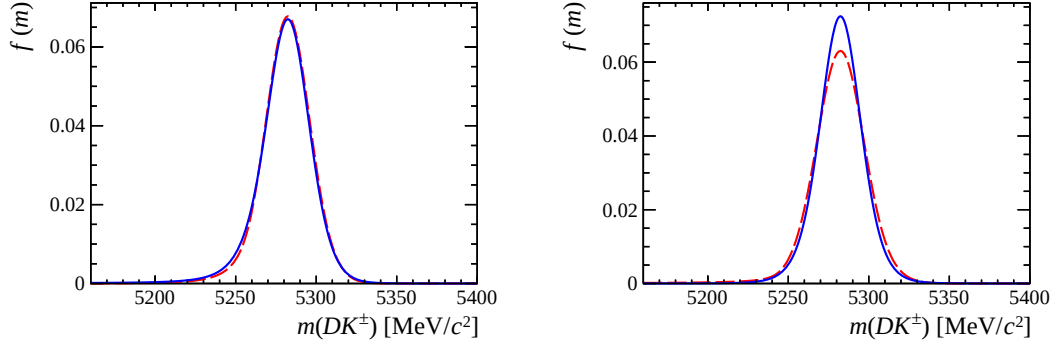


Figure 5.16. Comparison of (solid blue) default and (dashed red) alternative parameterisations of the signal $B \rightarrow DK$ shape for (left) LL (right) DD K_s^0 . The default parameterisation is a Cruijff function and the alternative is the sum of a Gaussian and a Crystal Ball function.

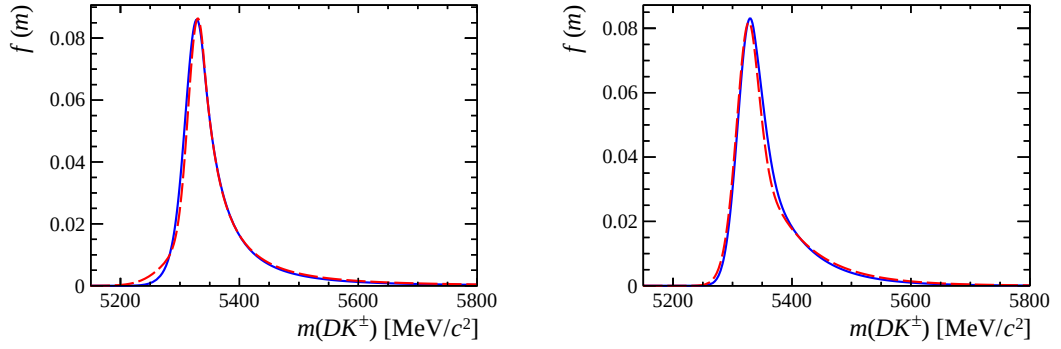


Figure 5.17. Comparison of (solid blue) default and (dashed red) alternative shapes to describe backgrounds where a bachelor pion has been misidentified as a kaon for (left) LL (right) DD K_s^0 .

subtraction in the $B \rightarrow D\pi$ sample used to derive the shape, and from discrepancies in the simulated bachelor kinematic distribution used to determine the PID efficiencies. Hence an alternative fit is performed where the means and widths of the misID shapes are allowed to vary. (This is not the default method due to the large overlap of the misID shape with the right-hand tail of the signal.) The default and alternative shapes are compared in Fig. 5.17. The systematic uncertainties on x_+ , x_- , y_+ and y_- when pseudoexperiments are generated with the alternative shape are 0.0016, 0.0000, 0.0011 and 0.0015, respectively.

Effect of using the same B mass shape in every Dalitz plot bin

It is assumed that each signal and background component can be described by the same shape in every Dalitz plot bin. In general this is a good approximation because the B mass has been calculated with a constraint on the D mass. It was shown in Fig. 5.1 that the B mass distribution is consistent within the statistical uncertainties among simulated signal decays lying in different bins. Hence, the assumption is revisited for misID, partially reconstructed and combinatorial backgrounds.

The B mass distribution of misID background can change if the D momentum spectrum varies between the Dalitz plot bins. The extent that such variations impact on the measurements of $x_\pm, y_\pm$ is estimated by producing a different shape for each bin using the method described in Sect. 3.2.3. When pseudoexperiments are generated with the individual shapes, the resulting systematic uncertainties are 0.0013, 0.0006, 0.0002 and 0.0005 for x_+ , x_- , y_+ and y_- , respectively.

The assumption that the B mass distribution of partially reconstructed backgrounds is revisited by considering physics effects in the participating backgrounds. The presence of CP violation is considered in the $B \rightarrow D^{*0}K$ decay channel. This mode has the largest known level of CP violation among the backgrounds, with values of $r_B(D^{*0}K) = 0.118$ and $\delta_B(D^{*0}K) = -51^\circ$ [19], and constitutes approximately 20% of the backgrounds partially reconstructed as $B \rightarrow DK$. Figure 5.18 shows how the B mass distribution changes in different bins when these effects are taken into account. The $B_s^0 \rightarrow D\bar{K}^{*0}$ background contribution is also considered here, because the D meson has the opposite flavour to the signal ($\bar{D}^0$) when reconstructed as $B^- \rightarrow DK^-$. A weighted MC cocktail, similar to the one in Sect. 3.2.4, is used to determine the background yield and shape in each Dalitz plot bin. It is estimated in simulation that the $B \rightarrow D^{*0}K$ and $B_s^0 \rightarrow D\bar{K}^{*0}$ backgrounds contribute 18% and 2% of the partially reconstructed background yield, respectively, while the remainder (e.g. $B \rightarrow D^{*0}\pi$, $B \rightarrow D^{*0}\rho$) con-

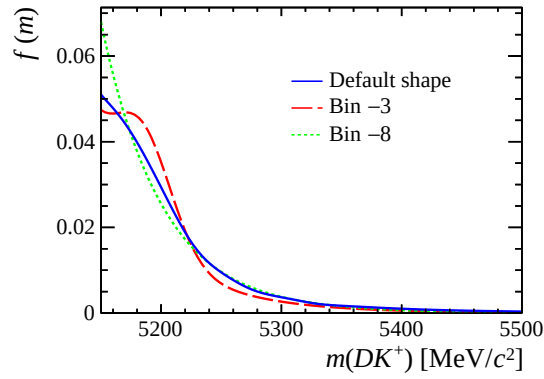


Figure 5.18. The (solid blue) default function in $m(DK^+)$ for partially reconstructed backgrounds is superimposed with separate shapes produced for two different Dalitz plot bins when CP violation in $B^+ \rightarrow D^{*0}K^+$ is taken into account.

tains D mesons that are of the same flavour as the signal. The associated systematic uncertainties are 0.0006, 0.0028, 0.0005 and 0.0023 for x_+ , x_- , y_+ and y_- , respectively.

The combinatorial background shape, described with an exponential function (Sect. 3.2.2), may also change from bin to bin. However, there is no practical way of estimating the extent to which the combinatorial slope might vary between bins using the selected data. The value of the exponential slope is primarily driven by the lower values of B mass, but this is dominated by the partially reconstructed background shape. Extrapolating individual slopes for each bin using only the upper values of the B mass, along with the relatively low combinatorial background level, would not be a reliable estimate. Instead, this is investigated by studying candidates where a real D meson and a fake bachelor particle are combined, where the largest variations in B mass distribution can be expected. The B mass distribution of such events is estimated by combining selected D candidates with a bachelor track from another randomly selected event, and is found to be uniform. When this flat shape is added to the default mass fit, the yield of such events is found to be consistent with zero. Hence, a systematic uncertainty associated with the combinatorial background shape is not considered further.

Systematic uncertainty due to migration on the Dalitz plot

The finite resolution of the tracking can cause the reconstructed values of m_+^2 and m_-^2 for a given event to differ from their true values. This can bias the CP observables if a net migration of events from one bin into another is induced, which can occur in the presence of resonances near bin boundaries. At first order this is taken into account by using semileptonic data to obtain the values of F_i , but the different resolution of m_-^2 and m_+^2 in $B \rightarrow D\pi$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays must be considered. This difference mostly comes from the use of different DTF constraints to calculate m_-^2 and m_+^2 in $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ and $B \rightarrow D\pi$. In the former case, the variables are refit with D0const and KS0const and in the latter case PVconst is also included. A non-zero value of r_B would also lead to different net migration in the $B \rightarrow DK$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ Dalitz plots.

The difference between the reconstructed and true values of m_+^2 and m_-^2 , Δm_+^2 and Δm_-^2 , are studied with simulation. Distributions of Δm_+^2 and Δm_-^2 and their correlations are shown in Fig. 5.19. Their correlations are negative because of the KS0const constraint. In both $B \rightarrow Dh$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$, the standard deviation of $\Delta m_\pm^2$ is $0.005 \text{ GeV}^2/c^4$ ($0.006 \text{ GeV}^2/c^4$) for LL (DD) $K_S^0 \pi^+ \pi^-$ events and $0.003 \text{ GeV}^2/c^4$ ($0.004 \text{ GeV}^2/c^4$) for LL (DD) $K_S^0 K^+ K^-$ events. The similarity between the two channels indicate that the D and K_S^0 mass constraints play the largest role in improving the resolution during the refit. Since the pixel width in the binning schemes (Sect. 1.6.3) is of similar size to the average $\Delta m_\pm^2$, most events will only move by one or two pixels, which will not result in a change in the assigned bin number.

To quantify the effect of the non-zero value of r_B in $B \rightarrow DK$, millions of pseudo-data events are generated using the BaBar amplitude models and efficiency profiles for $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ and $B \rightarrow D\pi$. To enable a fast generation, the efficiency profiles are parameterised with two-dimensional cubic polynomials. The $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ events are generated with no CP violation and the $B \rightarrow DK$ events are generated assuming $r_B = 0.1$,

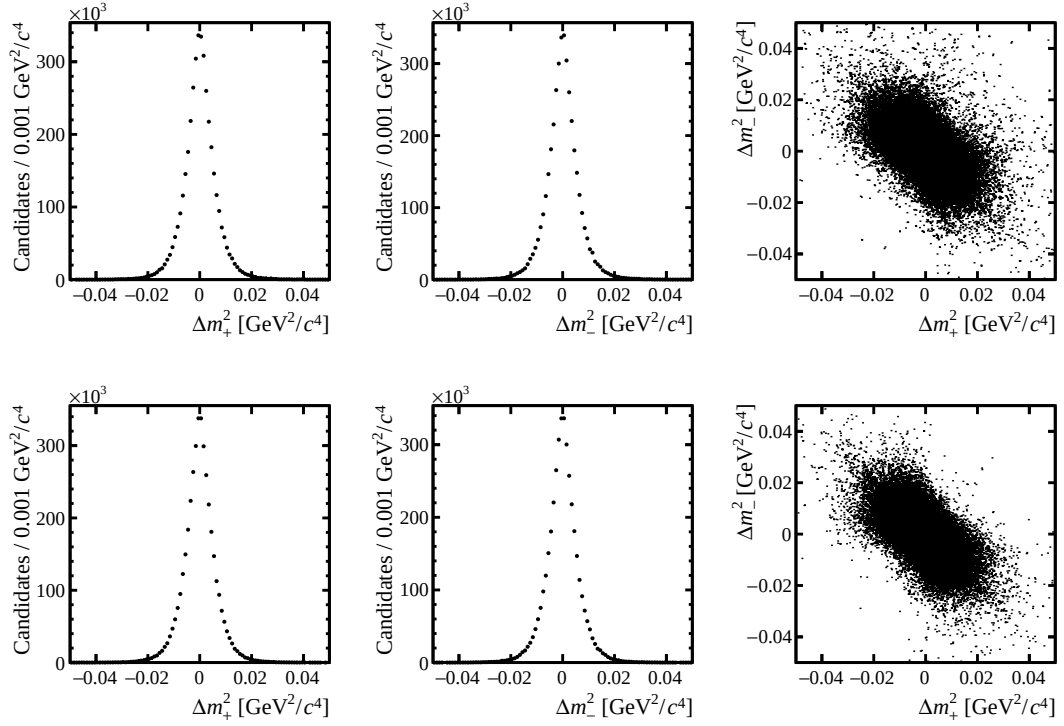


Figure 5.19. The difference between the true and reconstructed Dalitz variables, (left) Δm_+^2 and (centre) Δm_-^2 , in simulated (top) $B \rightarrow D\pi$ and (bottom) $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays, and (right) their correlation. The example shown involves DD $K_S^0\pi^+\pi^-$ candidates in 2012 conditions.

$\gamma = 70^\circ$ and $\delta_B = 130^\circ$. The Dalitz coordinates are smeared according to the values of Δm_-^2 , Δm_+^2 and their correlation observed in simulation. The net difference in migration between $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ and $B \rightarrow DK$ is found to be $< 1\%$ in nearly all the bins, except for a single bin in $D \rightarrow K_S^0\pi^+\pi^-$ where a relative change in bin population of up to 3% is observed. The associated systematic uncertainty is evaluated by correcting the F_i values to reflect the expected changes in bin populations due to migration effects. Using the altered F_i values the resulting systematic uncertainties are 0.0006, 0.0012, 0.0023 and 0.0015 for x_+ , x_- , y_+ and y_- , respectively.

The effect of charm mixing and neutral kaon detection asymmetries

The small effect of charm mixing is ignored in this analysis. The $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ and $B \rightarrow DK$ data are both time-integrated with similar D lifetime acceptances, hence any

effect is negligible and would be accounted for by using semileptonic data to provide the values of F_i . The measurements of c_i and s_i at CLEO are also insensitive to mixing at first order. The effect of ignoring charm mixing in the analysis is discussed in Ref. [87], which concludes that the maximum bias on γ is approximately 0.2° . This is negligible compared to the expected precision on γ .

The effect of ignoring kaon mixing and CP violation enters this analysis through the assumption that the K_S^0 is a pure CP -eigenstate and hence $A_D(m_-^2, m_+^2) = \bar{A}_D(m_+^2, m_-^2)$. The effect of this assumption is discussed in Ref. [22]. It is concluded that the resulting bias on γ is approximately $\Delta\gamma \sim |\epsilon_K|/r_B \sim \mathcal{O}(1^\circ)$, where $|\epsilon_K| = (2.228 \pm 0.011) \times 10^{-3}$ [4]. $\epsilon_K = A(K_L^0 \rightarrow \pi^+\pi^-)/A(K_S^0 \rightarrow \pi^+\pi^-)$ assuming that there is no direct CP violation in neutral kaon decays. This degree-level bias is ignored given the expected precision. The overall detection asymmetry between K^0 and $\bar{K}^0$ in $D \rightarrow K_S^0 h^+ h^-$ comes from both the mixing and CP violation effects above, and the different nuclear interaction cross-sections of K^0 and $\bar{K}^0$ flavour eigenstates with the detector material, which induces additional mixing between the states. The total kaon detection asymmetry is assessed in Ref. [88] using a data sample containing millions of $D^+ \rightarrow K_S^0 \pi^+$ decays. Using the expected trajectory of the kaon through the LHCb detector, it is found that the kaon detection asymmetry due to matter interactions is of similar magnitude to that of CP violation in the neutral kaon system. A cross-check is performed using the $B \rightarrow Dh$ and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ data samples here and it is found that the kaon asymmetries are compatible with zero within the statistical uncertainties. Hence detection asymmetries in the kaon system are not considered further.

Summary of systematic uncertainties

The systematic uncertainties are summarised in Table 5.4. All the uncertainties are taken as double-sided since in most cases only the magnitude of the effect is assessed and a preferred direction in which the systematic uncertainty acts is not demonstrated.

Table 5.4. The statistical and the various sources of systematic uncertainties for the parameters $x_\pm, y_\pm$. All values are given as multiples of 10^{-2} .

Source	$\sigma(x_+)$	$\sigma(x_-)$	$\sigma(y_+)$	$\sigma(y_-)$
Statistical	2.4	2.5	2.5	2.9
CLEO strong phase measurements	0.4	0.5	1.0	1.4
Efficiency corrections	0.9	0.9	0.2	0.2
Mass fit shapes	0.2	0.2	0.1	0.2
Shape of $D\pi$ misidentified as DK	0.1	0.1	0.0	0.1
Shape of partially reconstructed backgrounds	0.1	0.3	0.1	0.2
c_i, s_i bias due to efficiency	0.0	0.0	0.1	0.0
Bias correction	0.2	0.2	0.2	0.2
Migration	0.1	0.1	0.2	0.2
Total experimental systematic	1.0	1.0	0.4	0.5

The dominant experimental systematic uncertainties come from the efficiency correction in the measurements of F_i and the parameterisation of the B mass distribution. The systematic uncertainties from the CLEO measurements of c_i and s_i are reported separately.

5.4 Summary of results

Approximately 2580 $B \rightarrow DK$ decays with the final states $D \rightarrow K_S^0 \pi^+ \pi^-$ or $D \rightarrow K_S^0 K^+ K^-$ are used to determine the CP violation observables $x_\pm, y_\pm$, which are measured to be:

$$\begin{aligned}
 x_+ &= (-7.7 \pm 2.4 \pm 1.0 \pm 0.4) \times 10^{-2}, & y_+ &= (-2.2 \pm 2.5 \pm 0.4 \pm 1.0) \times 10^{-2}, \\
 x_- &= (2.5 \pm 2.5 \pm 1.0 \pm 0.5) \times 10^{-2}, & y_- &= (7.5 \pm 2.9 \pm 0.5 \pm 1.4) \times 10^{-2},
 \end{aligned}$$

where the first uncertainty is statistical, the second is systematic and the third is due to the uncertainty on the measurements of c_i and s_i . The correlations between $x_\pm$ and $y_\pm$ are given in Table 5.5. From Eq. (4.2), it can be seen that correlations in the

Table 5.5. Correlation matrix between the $x_\pm$ and $y_\pm$ parameters for (left) statistical uncertainties, (right) CLEO uncertainties and (bottom) sum in quadrature of all sources of uncertainties.

Statistical					CLEO				
	x_+	x_-	y_+	y_-		x_+	x_-	y_+	y_-
x_+	1.00	0.04	0.00	-0.01	x_+	1.00	-0.15	0.02	0.03
x_-		1.00	0.00	-0.25	x_-		1.00	0.17	0.01
y_+			1.00	0.01	y_+			1.00	-0.16
y_-				1.00	y_-				1.00

Total				
	x_+	x_-	y_+	y_-
x_+	1.00	0.03	0.00	-0.01
x_-		1.00	0.01	-0.20
y_+			1.00	-0.02
y_-				1.00

statistical uncertainties are expected between x_+ and y_+ , and between x_- and y_- , since they are measured using B^+ and B^- data, respectively. Further, correlations in the CLEO uncertainties are expected between x_+ and x_- , and between y_+ and y_- , since they are coupled to c_i and s_i , respectively. The statistical correlations are those measured in the fit to data and the CLEO correlations are determined using pseudoexperiments. The correlations are 25% or less and the expected trends are broadly observed. The experimental systematic uncertainties, arising from fitter and detector effects, are treated as uncorrelated. This is a reasonable assumption because the total systematic is dominated by the uncertainties from efficiency corrections, which are only weakly correlated. The correlations between the total uncertainties when they are summed in quadrature is also given in Table 5.5. The extraction of r_B , δ_B and γ from the values of $x_\pm, y_\pm$ are discussed in Sect. 7.1.

CHAPTER 6

Measurement of the CP violation observables in

$B^0 \rightarrow DK^{*0}$ decays

The measurement of the CP violation observables given in Eqs. 1.57 and 1.58, $x_{\pm}^* = r_{B^0} \cos(\delta_{B^0} \pm \gamma)$ and $y_{\pm}^* = r_{B^0} \sin(\delta_{B^0} \pm \gamma)$, in $B^0 \rightarrow DK^{*0}$ decays with $D \rightarrow K_s^0 \pi^+ \pi^-$ and $D \rightarrow K_s^0 K^+ K^-$ is described in this chapter. The selection of $B^0 \rightarrow DK^{*0}$ candidates is summarised in Sect. 6.1 and the B mass parameterisation of the selected candidates is given in Sect. 6.2. The relative decay rates in bins of the $D \rightarrow K_s^0 h^+ h^-$ Dalitz plots are related to $x_{\pm}^*$ and $y_{\pm}^*$ according to Eq. (1.56). A similar strategy to that used to measure $x_{\pm}$ and $y_{\pm}$ in $B \rightarrow DK$ decays is employed. Taking into account the Dalitz plot efficiency profile of $B^0 \rightarrow DK^{*0}$ by defining F_i as in Eq. (4.1), the rate equation in Eq. (1.56) is modified to

$$N_{\pm i}^- = n_-^* \left[F_{\pm i} + (x_-^{*2} + y_-^{*2}) F_{\mp i} + 2\kappa \sqrt{F_{+i} F_{-i}} (x_-^* c_{\pm i} + y_-^* s_{\pm i}) \right], \quad (6.1a)$$

$$N_{\pm i}^+ = n_+^* \left[F_{\mp i} + (x_+^{*2} + y_+^{*2}) F_{\pm i} + 2\kappa \sqrt{F_{+i} F_{-i}} (x_+^* c_{\pm i} - y_+^* s_{\pm i}) \right], \quad (6.1b)$$

where the superscript on N refers to the charge of the K^{*0} daughter kaon, hence $N_{\pm i}^-$ ($N_{\pm i}^+$) refers to the yield of $\bar{B}^0$ (B^0) decays in bin $\pm i$. $n_{\pm}^*$ are normalisation terms and κ is the coherence factor defined in Eq. (1.53). The F_i parameters are calculated from $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu}$ decays using the technique outlined in Chapter 4, with correction factors

to reflect the Dalitz plot efficiency of reconstructed $B^0 \rightarrow DK^{*0}$ decays. This is discussed in Sect. 6.3. The values of c_i and s_i are those provided by CLEO [3]. In this analysis, the modified optimal binning scheme is used for $D \rightarrow K_s^0 \pi^+ \pi^-$, as this minimises the fit biases observed in pseudoexperiments (see Sect. 6.4.2). For $D \rightarrow K_s^0 K^+ K^-$ candidates, the 2×2 binning scheme is used. The binning schemes are shown in Figs. 1.8 and 1.9. The CP observables $x_{\pm}^*, y_{\pm}^*$ are determined in Sect. 6.4 and the results are summarised in Sect. 6.5.

6.1 Selection of $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_s^0 h^+ h^-$ candidates

Events considered in the analysis fulfil the same hardware and software trigger requirements as the $B \rightarrow DK$ candidates, which are described in Sect. 3.1. At the stripping stage, composite particles are formed by combining $K_s^0 \rightarrow \pi^+ \pi^-$, $D \rightarrow K_s^0 h^+ h^-$, $\bar{K}^{*0} \rightarrow K^\pm \pi^\mp$, and $B^0 \rightarrow DK^{*0}$. The B candidate mass is required to be in the range 4750–6000 MeV/ c^2 . Loose requirements are placed on the quality of each track and vertex.

As with the reconstruction of $B \rightarrow DK$ candidates, the decay chain for each B candidate is refitted with constraints on $D0\text{const}$, $KS0\text{const}$ and $PV\text{const}$ (defined in Sect. 3.1) in order to improve the resolution on the B mass and the Dalitz plot coordinates. A refit is performed to improve the K_s^0 mass resolution using $D0\text{const}$ and $PV\text{const}$, and the D mass resolution is calculated after a refit with $KS0\text{const}$ and $PV\text{const}$.

6.1.1 Boosted decision tree

The motivation for and technical aspects of using a BDT classifier to select signal candidates have been described in Sect. 3.1.2. The BDT in this section was designed by Alexis Vallier for the model-dependent measurement of $x_{\pm}^*, y_{\pm}^*$ in $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_s^0 \pi^+ \pi^-$ decays [45, 89]. The main elements are summarised here.

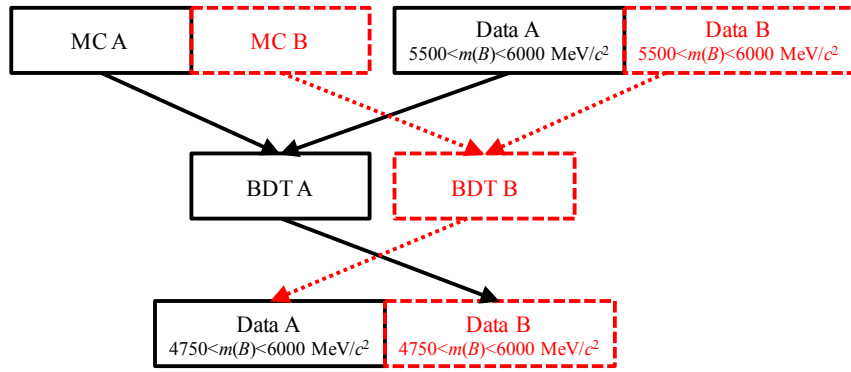


Figure 6.1. The procedure for performing a crossed BDT is summarised. The full data sample contains candidates in the range $4750 < m(B) < 6000 \text{ MeV}/c^2$, and the data samples used to train the BDT are in the range $5500 < m(B) < 6000 \text{ MeV}/c^2$.

The goal of the BDT is to suppress combinatorial background. The signal training sample is obtained from simulated $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_s^0 \pi^+ \pi^-$ decays, while the background training sample is obtained from candidates in data with a reconstructed B mass in the range $5500\text{--}6000 \text{ MeV}/c^2$. The background training sample is obtained from this mass range due to the presence of background $B_s^0 \rightarrow D\bar{K}^{*0}$ decays peaking at $m(B) \sim 5370 \text{ MeV}/c^2$ and the fact that only candidates with $m(B) < 6000 \text{ MeV}/c^2$ are retained after the stripping. The BDT must be trained with samples that are independent of the data to which the selection is subsequently applied. However, the B mass distribution is fitted at a later stage in the analysis, and a sufficiently wide range of the upper B mass sideband must be present in order to fit the combinatorial background reliably. This motivates the use of a *crossed* BDT, whose procedure is summarised in Fig. 6.1. All of the candidates in the MC and data samples are separated into two sets at random, which are labelled MC A, MC B, Data A and Data B. The candidates in MC A and Data A are used to train one BDT, denoted BDT A, while candidates in MC B and Data B are used to train a second BDT, denoted BDT B. The results of BDT A (B) are then applied to the candidates in Data B (A). Therefore the crossed BDT allows the data available to be fully exploited, and ensures that the BDT applied to one set of events is trained with a sample that is statistically independent to it.

The training samples are subject to loose criteria on the K^{*0} , D and K_S^0 candidate masses, and on the K^{*0} daughter $DLL_{K\pi}$, in order to remove candidates that are inconsistent with the final state. The BDTs are trained separately for candidates containing LL and DD K_S^0 candidates as their kinematic distributions differ. Hence there are four BDTs in total, and the same 16 variables are used to train all of them. The most discriminating variables in the set of DD BDTs are the refit χ^2 , the $K^{*0} p_T$ and the vertex separation χ^2 (defined in Sect. 3.1.2) of the B decay vertex from the PV. For the LL BDTs, two further variables are found to be strongly discriminating: the vertex separation χ^2 of the K_S^0 decay vertex from the PV and a ‘flight distance significance’ (FDSIG) variable characterising the separation of the K_S^0 and D meson vertices in the z -direction, which is defined as

$$K_S^0 \text{ FDSIG} = \frac{z(K_S^0) - z(D)}{\sqrt{\sigma_z^2(K_S^0) + \sigma_z^2(D)}}, \quad (6.2)$$

where z and σ_z are the value and error of the decay vertex position along the z -axis, respectively. The remaining variables in the BDT are the χ_{IP}^2 (defined in Sect. 3.1.1) of the B meson, the sum of the χ_{IP}^2 of the two K_S^0 daughter tracks, the sum of the χ_{IP}^2 of all the other tracks, the vertex quality of the B and D mesons, the vertex separation significance of the D meson from the PV, D FDSIG, defined as $\frac{z(D) - z(B)}{\sqrt{\sigma_z^2(D) + \sigma_z^2(B)}}$, the D and B meson p_T , the B DIRA (defined in Sect. 3.1.2), and the K^{*0} helicity angle θ^* . It is found that the performance of all four BDTs is consistent, with similar variable ranking and background rejection power. In all samples, the candidates are required to have a BDT response value greater than 0.75. This working point was determined by finding the value at which the mean statistical uncertainty on $x_{\pm}^*, y_{\pm}^*$ in pseudoexperiments is minimised. The same BDT is applied to candidates in the $D \rightarrow K_S^0 K^+ K^-$ final state.

6.1.2 Additional requirements

It is required that the LL (DD) refitted K_S^0 mass lies within 14.4 (19.9) MeV/c^2 of the known K_S^0 mass and the refitted D mass must lie within $30 \text{ MeV}/c^2$ of the known D mass.

Two backgrounds that peak in the B signal region are investigated further. For LL K_S^0 candidates, $D \rightarrow \pi^+\pi^-h^+h^-$ decays can contaminate the signal if two of the pions satisfy the requirements for a K_S^0 candidate. Studies in MC indicate that approximately 10% of events in the signal region selected by the BDT come from this background. Hence a requirement of $K_S^0 \text{ FDSIG} > 3$ is placed on the LL K_S^0 candidates, which is 99% efficient for signal and rejects 98% of the contamination. The efficacy of this criterion is validated in data by fitting the B mass distribution of LL K_S^0 candidates in the $D \rightarrow K_S^0\pi^+\pi^-$ final state and observing that the requirement on the $K_S^0 \text{ FDSIG}$ variable causes the $B_S^0 \rightarrow D\bar{K}^{*0}$ yield to change from 72 ± 10 to 63 ± 9 . This is consistent with 10% of the signal region being contaminated by $D \rightarrow \pi^+\pi^-\pi^+\pi^-$ decays, which is subsequently eliminated. Similar studies of the potential contribution of charmless decays finds that they are sufficiently suppressed by the BDT requirement, because the $D \text{ FDSIG}$ is one of the BDT variables.

Particle identification requirements are placed on the K^{*0} daughters in order to suppress the contribution from $B^0 \rightarrow D\rho^0$ decays, where either one of the pions from the $\rho^0 \rightarrow \pi^+\pi^-$ decay can be misidentified as a kaon. The requirements are that the K^{*0} daughter kaon has $\text{DLL}_{K\pi} > 5$ and the daughter pion has $\text{DLL}_{K\pi} < 3$, which retains 81% of the remaining signal events but suppresses 98% of the $B^0 \rightarrow D\rho^0$ background. This PID requirement also ensures that candidates caused by a double misID of the two K^{*0} daughters are removed. The D daughter kaons in $D \rightarrow K_S^0 K^+ K^-$ candidates are required to have $\text{DLL}_{K\pi} > -5$ which suppresses cross-feed from other $D \rightarrow K_S^0 h^+ h^-$ decays. (This requirement is identical to the $B \rightarrow DK$ analysis.)

Further requirements on the K^{*0} candidate mass and helicity angle, θ^* , are useful to suppress background due to random combinations of tracks faking a K^{*0} . However, these requirements also determine the values of the coherence factor κ , and the hadronic parameters r_{B^0} and δ_{B^0} by restricting the integrated region of the $B^0 \rightarrow DK^+\pi^-$ Dalitz plot, as shown in Eqs. (1.52) and (1.53). Choosing a $B^0 \rightarrow DK^+\pi^-$ Dalitz plot region to obtain a large value of κ is desirable as it increases the sensitivity to γ . Hence it is required that the K^{*0} candidate mass lies within $\pm 50 \text{ MeV}/c^2$ of the known K^{*0} mass [4] and that the absolute value of $\cos \theta^*$ is greater than 0.4. These requirements are identical in Ref. [90], where γ , r_{B^0} and δ_{B^0} are measured using $B^0 \rightarrow DK^{*0}$, $D \rightarrow h^+h^-$ decays at LHCb, and the value $\kappa = 0.95 \pm 0.03$ is determined using an amplitude model for the $B^0 \rightarrow DK^+\pi^-$ decay.

Multiple candidates per event are avoided by retaining only the candidate with the lowest refit χ^2 , which removes 23 out of 1977 selected candidates across all samples. A further two candidates are removed as they cannot be assigned Dalitz plot bin numbers due to a mismatch in the D mass used to produce the Dalitz plot binning scheme and to perform the refit.

6.1.3 Dalitz plots of selected candidates

Dalitz plots for selected $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_s^0 h^+ h^-$ candidates that lie within $25 \text{ MeV}/c^2$ of the known B^0 mass, separated by B^0 and $\bar{B}^0$ using the charge of the K^{*0} daughter kaon, are shown in Figs. 6.2 and 6.3.

6.2 Parametrisation of the B mass distribution

The B mass distribution of the selected candidates is fitted in order to determine the mass shapes and relative background yields, whose values are fixed in the subsequent

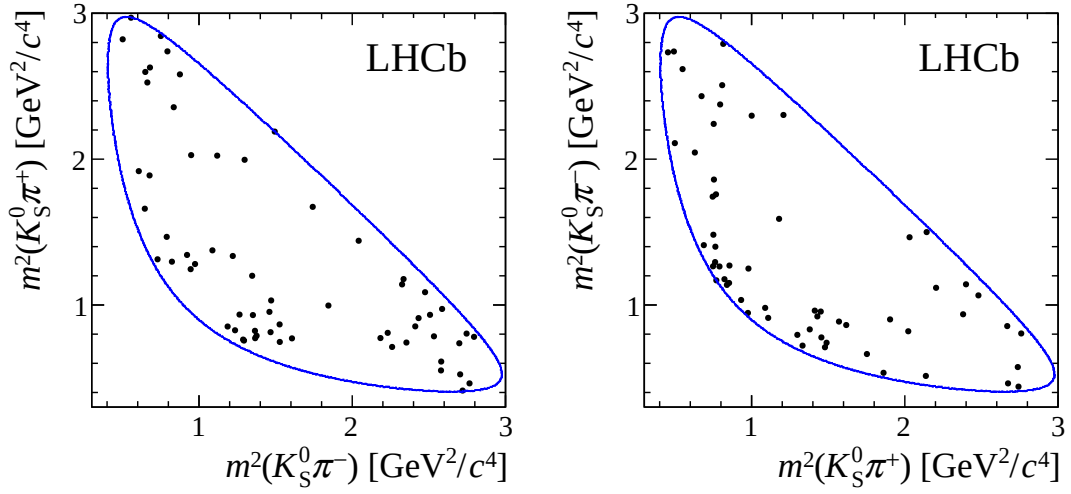


Figure 6.2. Dalitz plots of candidates in the signal region for $D \rightarrow K_S^0 \pi^+ \pi^-$ decays from (left) $\bar{B}^0 \rightarrow D \bar{K}^{*0}$ and (right) $B^0 \rightarrow D K^{*0}$ decays. The blue line indicates the kinematic boundary.

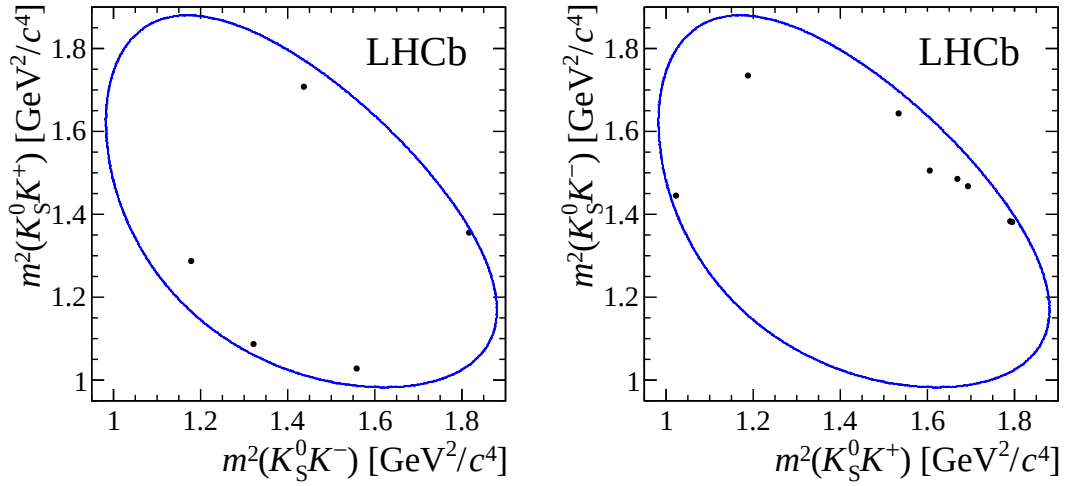


Figure 6.3. Dalitz plots of candidates in the signal region for $D \rightarrow K_S^0 K^+ K^-$ decays from (left) $\bar{B}^0 \rightarrow D \bar{K}^{*0}$ and (right) $B^0 \rightarrow D K^{*0}$ decays. The blue line indicates the kinematic boundary.

fit for $x_{\pm}^*, y_{\pm}^*$ in Sect. 6.4. Due to the small sample size, there is no sensitivity to the difference in mass resolution between LL and DD candidates and they are fitted together.

The fit strategy is driven by the small sample of approximately 1000 selected candidates, in which a robust knowledge of the mass shapes and the background yields must be obtained in order to minimise biases incurred in the Dalitz plot fit where the data are partitioned into many bins. This is particularly true of the function used to describe the combinatorial background, as its yield is measured independently in every bin.

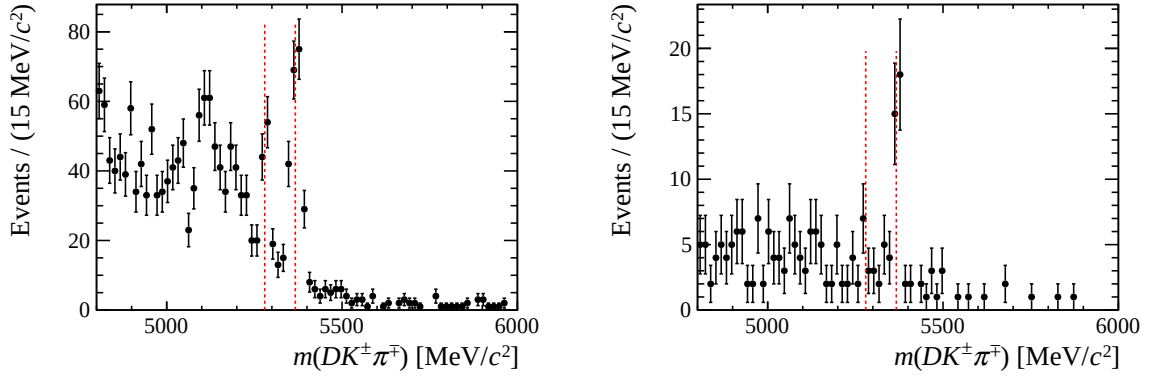


Figure 6.4. The $m(DK^\pm\pi^\mp)$ distribution of selected candidates in the range 4800–6000 MeV/ c^2 , in the (left) $D \rightarrow K_s^0\pi^+\pi^-$ and (right) $D \rightarrow K_s^0K^+K^-$ final states. The dashed red lines indicate the positions of the known B^0 and B_s^0 masses [4].

Table 6.1. Estimates of background yields passing the $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_s^0\pi^+\pi^-$ selection in two different $m(DK\pi)$ mass ranges: 4900–5800 MeV/ c^2 and 5200–5800 MeV/ c^2 . The yields are expressed as a ratio to the $B_s^0 \rightarrow D\bar{K}^{*0}$ yield, which is approximately 200 events.

Decay	Ratio to $N(B_s^0 \rightarrow D\bar{K}^{*0})$ (%)	
	4900–5800 MeV/ c^2	5200–5800 MeV/ c^2
$B^\pm \rightarrow DK^\pm\pi^+\pi^-$	18.9 ± 8.5	0.15 ± 0.08
$B^\pm \rightarrow D\pi^\pm\pi^+\pi^-$	10.9 ± 4.7	1.44 ± 0.63
$B \rightarrow D^{*0}K^\pm$	2.49 ± 0.98	1.38 ± 0.68
$B \rightarrow D^{*0}\pi^\pm$	1.25 ± 0.27	1.15 ± 0.25
$B^0 \rightarrow D^{*0}K^{*0}$	69.6 ± 37.1	2.13 ± 1.14
$B_s^0 \rightarrow D^{*0}K^{*0}$	142 ± 5	35.1 ± 1.4
$B^0 \rightarrow D\rho^0$	2.99 ± 0.82	2.94 ± 0.80
$B \rightarrow DK$	4.29 ± 0.97	4.24 ± 0.96
$\Lambda_b^0 \rightarrow Dp\pi$	0.76 ± 0.23	0.58 ± 0.18

6.2.1 Background studies and choice of fit range

The $m(DK^\pm\pi^\mp)$ distribution of the selected candidates is shown in Fig. 6.4. Peaking structures can be seen in the region below the B^0 and B_s^0 masses, which are partially reconstructed backgrounds. The contribution of various decays with reconstructed $m(DK\pi)$ in the region 4900–5800 MeV/ c^2 is estimated in order to determine which backgrounds need to be modelled in the fit. The backgrounds considered are: $B^0 \rightarrow D\rho^0$,

$B \rightarrow DK\pi\pi$, $B \rightarrow D\pi\pi\pi$, $\Lambda_b^0 \rightarrow Dp\pi$, $B \rightarrow DK$, $B \rightarrow D\pi$, $B \rightarrow D^{*0}K$, $B \rightarrow D^{*0}\pi$, $B^0 \rightarrow D^{*0}K^{*0}$, and $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$. A possible additional contribution from $B_s^0 \rightarrow DK^{*0}\pi^0$, where the π^0 is not reconstructed, is not considered in this analysis because its Dalitz plot structure is not known and therefore the $m(DK^{*0})$ distribution cannot be determined. Furthermore, an analysis of $B^0 \rightarrow DK^{*0}$, $D \rightarrow h^+h^-$ decays at LHCb [90] using an order of magnitude larger number of candidates than this work did not observe this background. Their reconstruction efficiencies are obtained from MC, and are weighted by the PID efficiencies of the $DLL_{K\pi}$ requirements on the two K^{*0} daughters using the technique described in Sect. 3.2.3. The total efficiencies in simulation are combined with the known values of the respective branching ratios [4] and fragmentation fractions [25] to estimate the relative contribution of each background. The expected yield of each background contribution is evaluated in terms of a ratio to the expected yield of background $B_s^0 \rightarrow D\bar{K}^{*0}$ decays, which is approximately 200 events. The $B_s^0 \rightarrow D\bar{K}^{*0}$ decay is used as a benchmark because its prominent peak in data (Fig. 6.4) makes a rough estimate of its yield, and therefore the size of the potential background contributions, possible. The ratios are given in Table 6.1. A data-driven method is used for the $B^0 \rightarrow D^{*0}K^{*0}$ and $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decays, as their branching fractions are only known as an upper limit in the former case and not measured in the latter. The information is taken from the higher statistics $B_s^0 \rightarrow D\bar{K}^{*0}$, $D \rightarrow K\pi$ decay that has been studied by LHCb [90], and the method is described in further detail in Sect. 6.2.3. The results are quoted here for ease of comparison and discussion. The ratios are also estimated in a limited mass region of 5200–5800 MeV/ c^2 . In this range the partially reconstructed $B^0 \rightarrow D^{*0}K^{*0}$ background, where either the π^0 or the photon from the D^{*0} decay is missed, is negligible. The CP parameters for this decay mode are not well known, and the relative contributions of $D^{*0} \rightarrow D^0\pi^0$ and $D^{*0} \rightarrow D^0\gamma$ decays vary in each Dalitz plot bin, hence with the current statistics it will be impossible to fit this component reliably in the Dalitz plot fit. One option is to fit the data in a wider mass range to establish the shape parameters, and then limit the fit range to 5200–5800 MeV/ c^2 when

performing the Dalitz plot fit. However, Table 6.1 indicates that there are approximately 60 candidates of partially constructed $B \rightarrow Dh\pi\pi$ decays, which lie primarily in the region $m(B) < 5200 \text{ MeV}/c^2$. (Only about 3 events are expected above this range.) While these decays can be modelled in the mass fit, there are other partially reconstructed decays that are likely to contribute in this region, such as the tails of the distributions due to decays where two particles are missed. Ignoring these backgrounds would result in these events being parameterised by the exponential function for the combinatorial background, biasing the fit. Alternatively, their mass shapes could be determined from simulation (with assumptions about potential resonant structure in the multibody B decay), but as many of these decays have no measured branching fraction their behaviour in the fit would be difficult to verify.

As a consequence of the above considerations, the lower range of the fit is chosen to be $5200 \text{ MeV}/c^2$, where the effect of partially reconstructed $B \rightarrow Dh\pi\pi$ decays in the low mass region and the tail contribution of 4–5 events of $B^0 \rightarrow D^{*0}K^{*0}$ decays can be ignored. It should be noted that performing the fit in this limited range results in a high level of correlation between the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ yield, the combinatorial background yield, and the combinatorial shape. Hence, the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ background yield will be constrained in a manner described in Sect. 6.2.3. Using the ratios in Table 6.1, the backgrounds included in the fit are $B^0 \rightarrow D\rho^0$, $B \rightarrow DK$ and $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decays. The $B^\pm \rightarrow Dh^\pm\pi^-\pi^+$, $B^0 \rightarrow D^{*0}K^{*0}$ and $B^\pm \rightarrow D^{*0}h^\pm$ backgrounds, which are expected to contribute 3–5 events, are assessed as a systematic uncertainty.

6.2.2 Mass shapes

With the exception of the combinatorial background, which is described using an exponential function whose slope is allowed to vary in the fit, each shape has some or all of its parameters determined using MC. In each case the invariant mass distributions are produced by simulating the background in question and reconstructing it as $B^0 \rightarrow DK^{*0}$,

$D \rightarrow K_s^0 \pi^+ \pi^-$. In order to represent the B mass resolution in data more accurately, each simulated B mass distribution is smeared with a Gaussian with mean $3.1 \text{ MeV}/c^2$ and width $5.0 \text{ MeV}/c^2$, which are determined by comparing the results of fitting the $B_s^0 \rightarrow D\bar{K}^{*0}$ signal peak in MC and data. The simulated distributions are also weighted using the method described in Sect. 3.2.3 in order to determine what effect the $DLL_{K\pi}$ requirements placed on the K^{*0} daughters have on the mass shapes.

The $B_s^0 \rightarrow D\bar{K}^{*0}$ background peak is modelled using a sum of two Crystal Ball functions [81], where the individual components have a common mean and width parameter that are determined in the fit to data. The tail parameters and the fraction between the two component shapes are fixed to the results of fitting the function to $B^0 \rightarrow DK^{*0}$ MC [89]. The $B^0 \rightarrow DK^{*0}$ signal is modelled using an identical shape, but with its mean shifted by the known $B_s^0 - B^0$ mass difference of $87.19 \text{ MeV}/c^2$ [4].

The distribution of $B^0 \rightarrow D\rho^0$ decays, where one of the pions from the $\rho^0 \rightarrow \pi^+ \pi^-$ decay has been misidentified as a kaon, is described by the sum of two Crystal Ball functions where the shape parameters are determined with a fit to simulated $B^0 \rightarrow D\rho^0$ decays. The mass shape for $B \rightarrow DK$ decays, where a random pion track in the event has been added, is determined in a similar fashion. A single Crystal Ball function provides an adequate description of this shape. The distributions of both backgrounds are shown in Fig. 6.5.

Partially reconstructed decays

The partially reconstructed decays of $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$, in which either a π^0 or a photon is missed from the D^{*0} decay, are modelled in a similar fashion to the $B \rightarrow D^{*0} h$ backgrounds in the $B \rightarrow Dh$ mass fit described in Sect. 3.2.4. The initial decay of the B_s^0 involves the decay of a pseudoscalar to two vector particles, hence due to angular momentum conservation there are three helicity amplitudes to consider, which

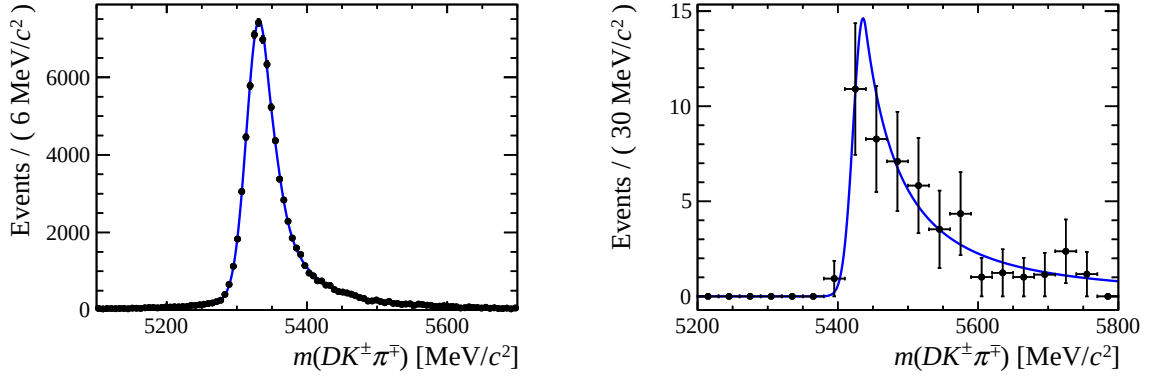


Figure 6.5. The $m(DK^\pm\pi^\mp)$ distributions of simulated (left) $B^0 \rightarrow D\rho^0$ and (right) $B \rightarrow DK$ decays reconstructed as $B^0 \rightarrow DK^{*0}$. The results of fitting (left) a sum of two Crystal Ball functions and (right) a Crystal Ball function to these decays are superimposed.

correspond to three D^{*0} helicity states, labelled $\lambda = -1, 0, +1$. In the subsequent parity-conserving decay $D^{*0} \rightarrow D^0\{\pi^0, \gamma\}$, the value of λ and the spin of the missing neutral particle determine the distribution of the missing particle's helicity angle θ , which is defined as the angle between the neutral particle's momentum vector in the D^{*0} rest frame and the D^{*0} boost vector in the B rest frame. The distribution of θ is identical for the cases where $\lambda = -1$ or $+1$, and due to the one-to-one correspondence of θ with the reconstructed B mass, the resulting B mass distributions for these two helicity states are indistinguishable and are therefore grouped together.

The functions describing the B mass distributions and the θ distribution of the different decays are summarised in Table 6.2. In addition to the Horns and Hill functions defined in Eq. (3.5) and (3.7), a function called LittleHorns is also required. It is described by a parabola $p_{\text{LITTLEHORNS}}(x)$ with kinematic endpoints a, b where

$$p_{\text{LITTLEHORNS}}(x) = \begin{cases} \left(m - \frac{a+b}{2}\right)^2 + \left(\frac{a-b}{2}\right)^2, & \text{if } a < x < b \\ 0, & \text{otherwise.} \end{cases} \quad (6.3)$$

The resulting LittleHorns function is obtained by convolving $p_{\text{LITTLEHORNS}}$ with a double Gaussian function, DG, defined in Eq. (3.4), to describe the experimental

Table 6.2. The functions used to describe the $m(DK\pi)$ distribution of partially reconstructed $B_s^0 \rightarrow (D^{*0} \rightarrow D\{\pi^0, \gamma\})\bar{K}^{*0}$ decays, depending on the species of the missed neutral particle and the D^{*0} helicity state. The distribution of the missing particle's helicity angle (θ) is also given.

Neutral	D^{*0} helicity state	θ -dependence	Function
γ	0	$\sin^2 \theta$	Hill, Eq. (3.7)
γ	± 1	$1 + \cos^2 \theta$	LittleHorns, Eq. (6.4)
π^0	0	$\cos^2 \theta$	Horns, Eq. (3.5)
π^0	± 1	$\sin^2 \theta$	Hill, Eq. (3.7)

resolution, and with a linear polynomial with slope $1 - \xi_{\text{LITTLEHORNS}}$ to account for selection effects, giving

$$\text{LittleHorns}(m) = \int_a^b dx \left\{ \left[\left(x - \frac{a+b}{2} \right)^2 + \left(\frac{a-b}{2} \right)^2 \right] \text{DG}(x|m, \sigma, f_G, R_\sigma) \right. \\ \left. \left(\frac{1 - \xi_{\text{LITTLEHORNS}}}{b-a} x + \frac{b\xi_{\text{LITTLEHORNS}} - a}{b-a} \right) \right\}. \quad (6.4)$$

The kinematic endpoints a and b are calculated. The remaining shape parameters, including the global shift δm (explained in Sect. 3.2.4), are obtained by performing a simultaneous fit to four samples of $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ MC, which simulate decays where the D^{*0} decays to either a π^0 or a photon and is produced in either the $\lambda = 0$ or $\lambda = \pm 1$ helicity state. The result of the simultaneous fit is shown in Fig. 6.6. The DG function parameters are required to be the same for all four samples, and each sample has an independent slope parameter to account for their different selection efficiencies.

6.2.3 Constraints on background yields

With the large number of overlapping signal and background contributions it is not possible to let all the background yields vary freely in the fit. The small expected number of $B^0 \rightarrow D\rho^0$ and $B \rightarrow DK$ decays, and their proximity to the $B_s^0 \rightarrow D\bar{K}^{*0}$ signal peak, results in a 100% uncertainty in their yields if they are floated as free

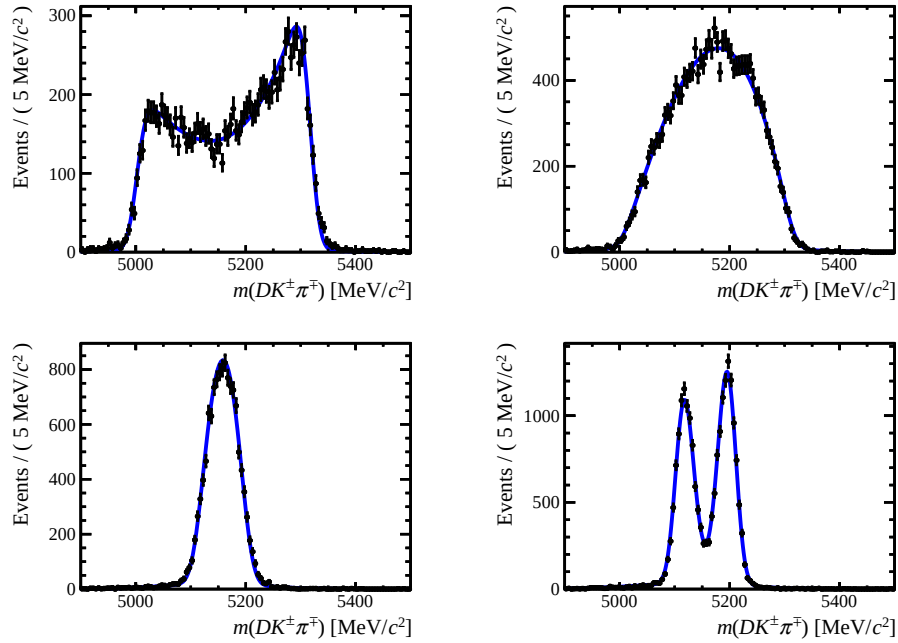


Figure 6.6. The $m(DK^\pm\pi^\mp)$ distribution of $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ MC reconstructed as $B_s^0 \rightarrow D\bar{K}^{*0}$ when the missing particle is (top) a photon or (bottom) a π^0 meson, and the D^{*0} helicity state λ is (left) ± 1 or (right) 0. The result of a fitting a (top left) LittleHorns, (top right, bottom left) Hill and (bottom right) Horns function is superimposed.

parameters. For $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decays, the lack of a clear peaking structure in the range $5200 < m(B) < 5800 \text{ MeV}/c^2$, and correlations with the combinatorial background shape and yield parameters, prevents both its total yield and the relative contribution from the different D^{*0} helicity states from being determined reliably. Hence the yields of the $B^0 \rightarrow D\rho^0$, $B \rightarrow DK$ and $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decays are implemented as Gaussian-constrained ratios relative to the $B_s^0 \rightarrow D\bar{K}^{*0}$ yield. The yield ratios are labelled R_ρ , R_{DK} and R_s , respectively, and are given in Table 6.1.

In the case of the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ background, neither its branching fraction nor the relative fraction of the D^{*0} helicity states have been measured. The total mass shape for the partially reconstructed $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decay is

$$\alpha \left[G_{\lambda=0} \mathcal{P}_{\lambda=0}^\gamma + (1 - G_{\lambda=0}) \mathcal{P}_{\lambda=0}^{\pi^0} \right] + (1 - \alpha) \left[G_{\lambda=\pm 1} \mathcal{P}_{\lambda=\pm 1}^\gamma + (1 - G_{\lambda=\pm 1}) \mathcal{P}_{\lambda=\pm 1}^{\pi^0} \right], \quad (6.5)$$

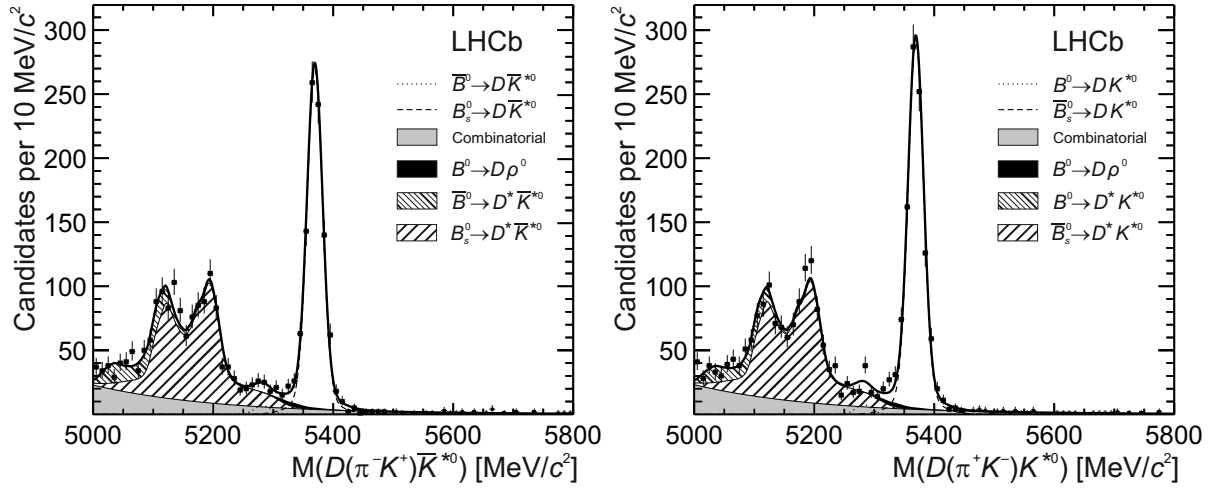


Figure 6.7. Distributions of (left) $m(D(\rightarrow \pi^- K^+) \bar{K}^{*0})$ and (right) $m(D(\rightarrow \pi^+ K^-) K^{*0})$. The data (black points) and the fitted model (thick solid line) are shown. The shapes corresponding to the different species are indicated in the legend. Figure reproduced and caption adapted from Fig. 3 in Ref. [90].

where $\mathcal{P}_\lambda^X$ are the Horns, Hill and LittleHorns functions specified in Table 6.2 for missing particle X and D^{*0} helicity state λ , and α is the fraction of D^{*0} mesons with $\lambda = 0$. The parameters $G_{\lambda=0} = 0.53 \pm 0.02$ and $G_{\lambda=\pm 1} = 0.80 \pm 0.05$ are the relative contributions of the $D^{*0} \rightarrow D^0 \gamma$ decay with $\lambda = 0$ and $\lambda = \pm 1$, respectively. $G_{\lambda=0}$ and $G_{\lambda=\pm 1}$ are calculated using the branching fractions and simulated reconstruction efficiencies of the $D^{*0} \rightarrow D^0 \gamma$ and $D^0 \pi^0$ final states. Their branching fractions are $\mathcal{B}(D^{*0} \rightarrow D^0 \gamma) = 0.381 \pm 0.029$ and $\mathcal{B}(D^{*0} \rightarrow D^0 \pi^0) = 0.619 \pm 0.029$ [4].

The parameter α of Eq. (6.5) is not known *a priori* but has previously been measured in the higher statistics analysis of $B^0 \rightarrow DK^{*0}$, $D \rightarrow h^+ h^-$ decays at LHCb [90]. Selected projections of the mass fit with the doubly Cabibbo-suppressed $D \rightarrow K\pi$ decay are reproduced in Fig. 6.7. In this projection, the suppression of the $B^0 \rightarrow DK^{*0}$ and $B^0 \rightarrow D^{*0} K^{*0}$ decays means that the $B_s^0 \rightarrow D \bar{K}^{*0}$ and $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$ decays dominate, allowing the yield and shape parameters of $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$ to be measured reliably in the range $5000 < m(B) < 5800 \text{ MeV}/c^2$. Since the underlying physics in the $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$ decay does not depend on the D final state, the measured value of $\alpha(D \rightarrow K\pi) = 0.73 \pm 0.04$ can be used to describe $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$, $D \rightarrow K_s^0 h^+ h^-$ decays. After taking into account

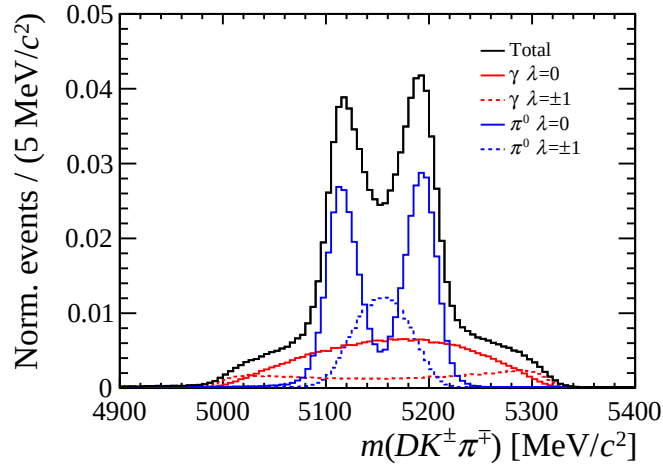


Figure 6.8. The $m(DK^\pm\pi^\mp)$ distribution of (solid black) $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$, $D \rightarrow K_s^0\pi^+\pi^-$ decays in MC, with the components where the D^{*0} decays to (blue) $D^0\pi^0$ or (red) $D^0\gamma$, and the D^{*0} helicity state is (solid) $\lambda = 0$ or (dashed) $\lambda = \pm 1$. The components have been added using the values of α , $G_{\lambda=0}$ and $G_{\lambda=\pm 1}$ given in the main text.

differences in selection and fit range, it is found that $\alpha = 0.72 \pm 0.13$ is the appropriate value for this analysis. The increase in uncertainty comes from $G_{\lambda=0,\pm 1}$ evaluated only in the region above $5200 \text{ MeV}/c^2$. The central value of α here is very similar to that in the $D \rightarrow K\pi$ analysis, as both contain the criterion that $|\cos\theta^*| > 0.4$ (where θ^* is the K^{*0} helicity angle). This suppresses the contribution from decays where the K^{*0} (and the D^{*0}) is produced in a $\lambda = \pm 1$ state, because the decay will have a $\sin^2\theta^*$ distribution. Hence, it is expected that the two analyses select the same fraction of partially reconstructed D^{*0} meson decays with $\lambda = 0$. The distributions of the four components in the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ background and their total distribution, when scaled according to the values of $G_{\lambda=0}$, $G_{\lambda=\pm 1}$ and α , are shown in Fig. 6.8. Since the branching ratio of $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ is unknown, the ratio R_s is also estimated from the fit to $D \rightarrow h^+h^-$ decays, using the distributions in Fig. 6.7. After correcting for differences in the selection and fit range, it is found that $R_s = 0.351 \pm 0.137$.

6.2.4 Fit results

The $B^0 \rightarrow DK^{*0}$ data is fitted and the fit projections are shown in Fig. 6.9. The two D final states, $K_s^0 \pi^+ \pi^-$ and $K_s^0 K^+ K^-$, are fitted simultaneously. Apart from the exponential slope parameter for the combinatorial background, the same shape parameters are used for the two D final states. The parameters α , R_{DK} , R_ρ and R_s are implemented as Gaussian constraints with means and widths corresponding to the central values and the errors given. The shape parameters obtained from MC are fixed in the fit and are listed in Table 6.3. The yields of $B^0 \rightarrow DK^{*0}$ signal, $B_s^0 \rightarrow D\bar{K}^{*0}$ background and combinatorial background are free parameters in the fit. The fit parameters are summarised in Table 6.4. The fitted values of α , R_{DK} , R_s and R_ρ deviate by less than 1σ from the Gaussian constraints imposed. Pseudoexperiments generated using these fit results show that only negligible biases are expected.

The number of $B^0 \rightarrow DK^{*0}$ signal events is 84_{-14}^{+15} in $D \rightarrow K_s^0 \pi^+ \pi^-$ and 7_{-4}^{+5} in $D \rightarrow K_s^0 K^+ K^-$. The ratio of yields between $D \rightarrow K_s^0 \pi^+ \pi^-$ and $D \rightarrow K_s^0 K^+ K^-$ is 12 ± 9 , which is consistent with the ratio of 7 observed in the fits to $B \rightarrow Dh$ data (Tables 3.7 and 3.8). The yield of $D \rightarrow K_s^0 K^+ K^-$ is smaller than expected, but is consistent with a downward statistical fluctuation. Pseudoexperiments generated prior to looking at the data, with 100 $K_s^0 \pi^+ \pi^-$ and 8 $K_s^0 K^+ K^-$ events, show that the addition of the $K_s^0 K^+ K^-$ events in the fit provides up to 5% improvement in the statistical precision on $x_\pm$ and a decrease in the fit bias. Hence, despite the small yield of $K_s^0 K^+ K^-$ observed, these events are retained in the analysis.

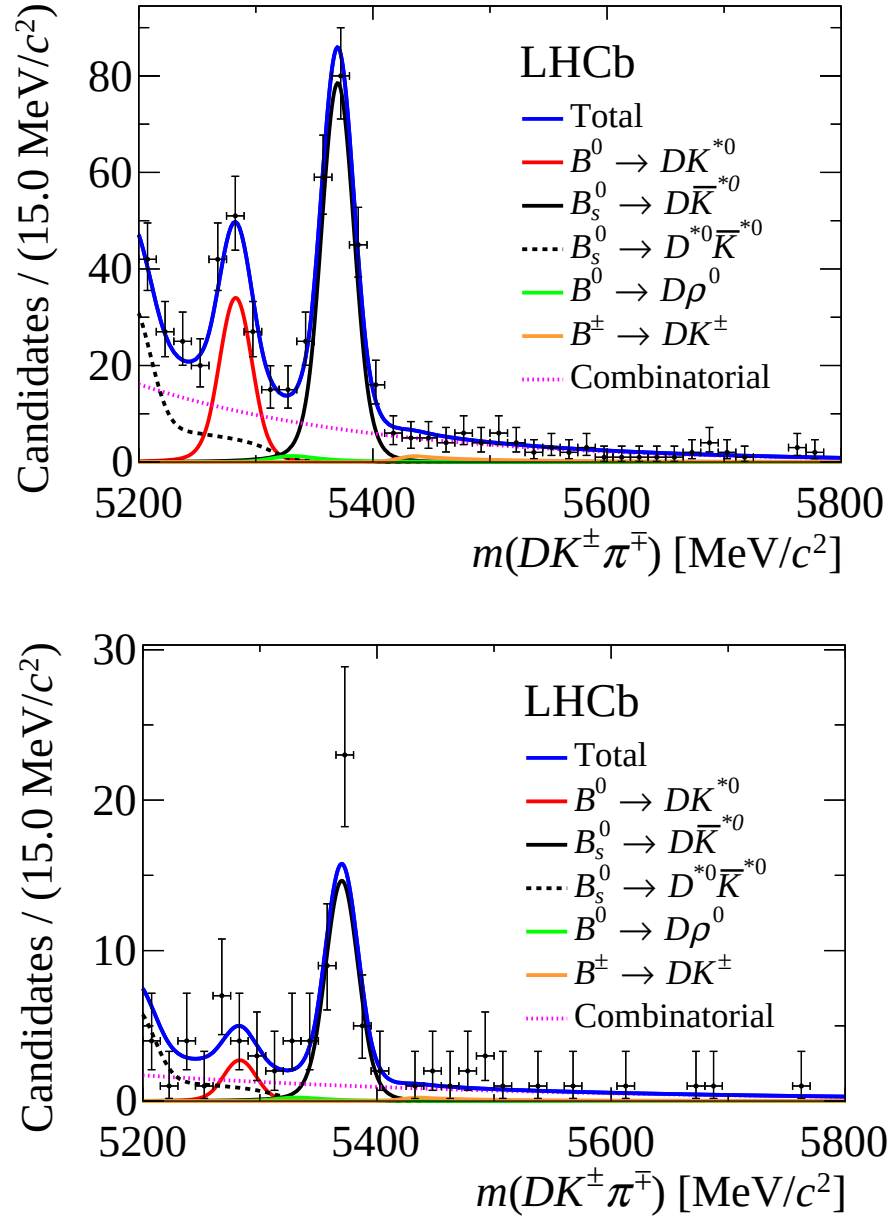


Figure 6.9. The $m(DK^\pm \pi^\mp)$ distributions of (red) $B^0 \rightarrow DK^{*0}$ candidates with (top) $D \rightarrow K_s^0 \pi^+ \pi^-$ and (bottom) $D \rightarrow K_s^0 K^+ K^-$. The fit results, including the signal and background components, are superimposed.

Table 6.3. Fixed parameters in the B mass fit to $B^0 \rightarrow DK^{*0}$ candidates. The signal and $B^0 \rightarrow D\rho^0$ shapes are the sum of two Crystal Balls. The Crystal Ball parameters μ and σ are the mean and width parameters, respectively, and have units of MeV/c^2 ; α^{CB} and n are tail parameters; and f_1 ($1 - f_1$) is the fraction of Crystal Ball 1 (2). The shape parameters for $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ are defined in the main text.

Fit component	Shape	Parameters
Signal	Crystal Ball 1	$\alpha_1^{\text{CB}} = 1.5, n_1 = 2.8, f_1 = 0.46$
	Crystal Ball 2	$\alpha_2^{\text{CB}} = 1.6, n_2 = 6.6$
$B^0 \rightarrow D\rho^0$	Crystal Ball 1	$\mu_1 = 5331.0, \sigma_1 = 17.46, \alpha_1^{\text{CB}} = 0.67, n_1 = 2.6, f_1 = 0.88.$
	Crystal Ball 2	$\mu_2 = 5347.1, \sigma_2 = 14.70, \alpha_2^{\text{CB}} = 0.41, n_2 = 2.6$
$B \rightarrow DK$	Crystal Ball	$\mu = 5436.0, \sigma = 14.1, \alpha^{\text{CB}} = 0.28, n = 1.8$
$B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$	Resolution (DG)	$\sigma = 13.93 \text{ MeV}/c^2, R_\sigma = 6.26, f_G = 0.93$
	Global shift	$\delta m = -35.1 \text{ MeV}/c^2$
	Slope parameter	$\xi_{\lambda=\pm 1}^\gamma = 0.57, \xi_{\lambda=0}^\gamma = 0.56,$ $\xi_{\lambda=\pm 1}^{\pi^0} = 0.82, \xi_{\lambda=0}^{\pi^0} = 0.83$

Table 6.4. Fit results for the fit to $B^0 \rightarrow DK^{*0}$ candidates with the $D \rightarrow K_S^0 \pi^+ \pi^-$ and $D \rightarrow K_S^0 K^+ K^-$.

Shape parameter	Value
B_s^0 mass [MeV/c^2]	$5369.2_{-1.0}^{+1.0}$
Signal width [MeV/c^2]	$13.3_{-0.9}^{+1.0}$
$K_S^0 K^+ K^-$ exponential slope [$(\text{MeV}/c^2)^{-1}$]	$(-3.4_{-1.4}^{+1.6}) \times 10^{-3}$
$K_S^0 \pi^+ \pi^-$ exponential slope [$(\text{MeV}/c^2)^{-1}$]	$(-5.4_{-0.8}^{+0.9}) \times 10^{-3}$
α	$0.74_{-0.13}^{+0.13}$
R_{DK}	$(4.3_{-1.0}^{+1.0}) \times 10^{-2}$
R_ρ	$(3.0_{-0.8}^{+0.8}) \times 10^{-2}$
R_s	$0.31_{-0.09}^{+0.09}$
Yield parameter	Value
$n(B^0 \rightarrow DK^{*0}, K_S^0 \pi^+ \pi^-)$	84_{-14}^{+15}
$n(B_s^0 \rightarrow D\bar{K}^{*0}, K_S^0 \pi^+ \pi^-)$	194_{-17}^{+18}
$n(\text{combinatorial}, K_S^0 \pi^+ \pi^-)$	207_{-35}^{+36}
$n(B^0 \rightarrow DK^{*0}, K_S^0 K^+ K^-)$	$6.7_{-4.2}^{+4.8}$
$n(B_s^0 \rightarrow D\bar{K}^{*0}, K_S^0 K^+ K^-)$	$36.3_{-6.4}^{+7.1}$
$n(\text{combinatorial}, K_S^0 K^+ K^-)$	$32.3_{-9.0}^{+10.0}$

6.3 Calculation of the F_i parameters

The parameters F_i are determined using the method described in Chapter 4. The yields of flavour-tagged semileptonic $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays are measured in each Dalitz plot bin. MC samples are then used to correct for Dalitz plot efficiency differences between the semileptonic and $B^0 \rightarrow DK^{*0}$ candidates.

6.3.1 Selection and parameterisation of $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ candidates

The selection of $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ candidates is identical to the one described in Sect. 4.1 for the $B \rightarrow DK$ analysis, with two changes found to better align the Dalitz plot efficiencies. Firstly, only events with a muon candidate that passes the L0Muon T0S trigger requirements are considered; previously events with a D candidate passing the L0Hadron T0S trigger requirements were also used. Secondly, the software topological trigger designed to select B candidates is identical to the one used for $B^0 \rightarrow DK^{*0}$ candidates (Hlt2TopoNBodyBBDT); an algorithm that also required one of the tracks to satisfy muon identification requirements was previously used (Hlt2TopoMuNBodyBBDT). Although 20% of the previous sample is removed, studies in MC show that better agreement in the Dalitz plot efficiency between $B^0 \rightarrow DK^{*0}$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ is achieved.

The selected $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ candidates are parameterised using the two-dimensional fit in $m(D)$ and Δm described in Sect. 4.2. An example fit projection is shown in Fig. 6.10, which indicates that the parameterisation provides a good description of this dataset. The signal and background yields are given in Table 6.5. The $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ signal yield is approximately 90 000 (12 000) for $D \rightarrow K_s^0\pi^+\pi^-$ ($D \rightarrow K_s^0K^+K^-$), which is three orders of magnitude larger than the $B^0 \rightarrow DK^{*0}$ yield. The results of a subsequent fit to measure the yield in each Dalitz plot bin are shown in Fig. 6.11. The modified optimal binning scheme is used for $K_s^0\pi^+\pi^-$ and the 2×2 binning is used for $K_s^0K^+K^-$.

Table 6.5. Yields of the signal, combinatorial background and random slow pion (RSP) backgrounds in the $B^0 \rightarrow D^{*-}(\rightarrow D^0\pi^-)\mu^+\nu_\mu$ control channel, evaluated in the signal mass window, which is defined as $|\Delta m - \Delta m_{\text{PDG}}| = 1.5 \text{ MeV}/c^2$ and $|m(D^0) - m(D^0)_{\text{PDG}}| = 25(15) \text{ MeV}/c^2$ for $K_S^0\pi^+\pi^-$ ($K_S^0K^+K^-$). The purity is also evaluated.

Mode	Signal	Combinatorial	RSP	Purity
2012				
$K_S^0\pi^+\pi^-$ LL	20001 ± 130	244 ± 61	391 ± 65	0.969 ± 0.004
$K_S^0\pi^+\pi^-$ DD	42929 ± 177	1542 ± 72	682 ± 101	0.950 ± 0.003
$K_S^0K^+K^-$ LL	2851 ± 44	64 ± 4	68 ± 9	0.956 ± 0.003
$K_S^0K^+K^-$ DD	5219 ± 61	204 ± 6	104 ± 11	0.944 ± 0.002
2011				
$K_S^0\pi^+\pi^-$ LL	9433 ± 105	92 ± 8	198 ± 16	0.970 ± 0.002
$K_S^0\pi^+\pi^-$ DD	17724 ± 139	553 ± 17	281 ± 23	0.955 ± 0.002
$K_S^0K^+K^-$ LL	1350 ± 38	32 ± 2	32 ± 5	0.955 ± 0.004
$K_S^0K^+K^-$ DD	2257 ± 49	88 ± 4	39 ± 7	0.947 ± 0.003

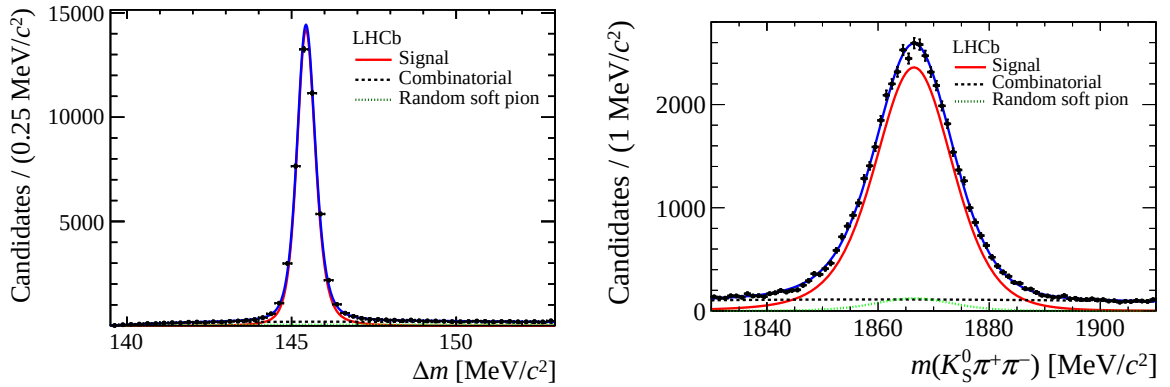


Figure 6.10. Result of the simultaneous fit to $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$, $D \rightarrow K_S^0\pi^+\pi^-$ decays with DD K_S^0 candidates, in 2012 data. A two-dimensional fit is performed in (left) Δm and (right) $m(K_S^0\pi^+\pi^-)$. The (blue) total fit result and the signal and background components are superimposed.

6.3.2 Dalitz plot efficiency correction

Efficiency maps of the signal $B^0 \rightarrow DK^{*0}$ and control $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ channels are obtained from MC in order to calculate the efficiency correction factors ξ_i defined Eq. (4.9). $B_S^0 \rightarrow D\bar{K}^{*0}$ MC is used to study the Dalitz plot efficiency in $B^0 \rightarrow DK^{*0}$ due to the larger sample available, and it is verified in simulation that the $87 \text{ MeV}/c^2$ mass

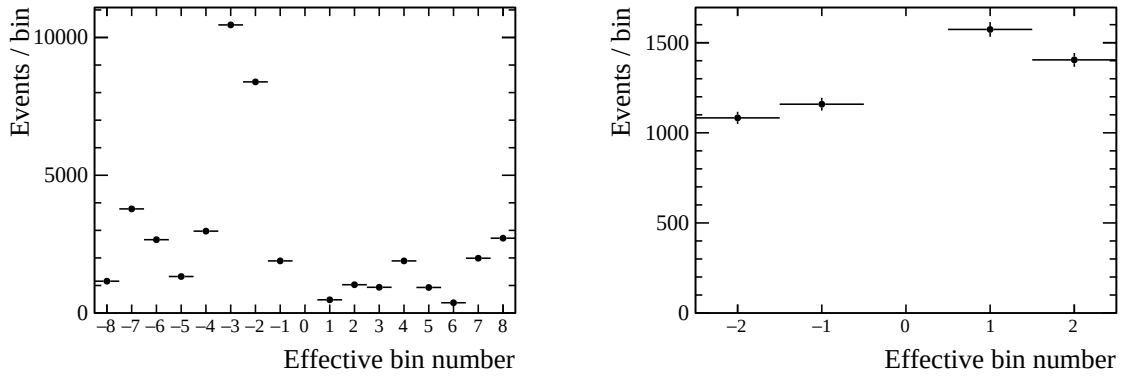


Figure 6.11. The binned yields of $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$, DD K_S^0 candidates from 2012 data. The (left) $K_S^0 \pi^+ \pi^-$ and (right) $K_S^0 K^+ K^-$ final states are divided using the modified optimal and 2×2 binning schemes, respectively.

difference between the B_s^0 and B^0 mesons has a negligible impact on the D kinematics or Dalitz plot efficiency. Separate efficiency maps from simulation are produced for 2011 and 2012, and for DD and LL K_S^0 . As before, it is assumed that the Dalitz plot efficiencies are symmetric about the diagonal defined by $m_-^2 = m_+^2$; this has been shown to be a good approximation (Sect. 5.3) and makes full statistical use of the MC sample.

The response of the L0Hadron trigger used to select $B^0 \rightarrow DK^{*0}$ candidates is not modelled accurately enough in the simulation and does not select candidates uniformly over the Dalitz plot, hence the trigger efficiencies of simulated $B^0 \rightarrow DK^{*0}$ candidates are calibrated using weights obtained from data [91]. The calibration is performed using $D^0 \rightarrow K^- \pi^+$ decays to provide a sample of kaons and pions, where the D^0 is produced in the $D^{*+} \rightarrow D^0 \pi^+$ decay. The efficiency of the L0Hadron trigger is measured in the calibration sample, using only tracks that are projected to fall within the calorimeter acceptance, as a function of the E_T deposited in the HCAL. Efficiencies are measured separately for tracks depositing energy in the inner or outer section of the HCAL (Sect. 2.4), 2011 or 2012 conditions, track charge, particle type and magnet polarity.

The simulated $B^0 \rightarrow DK^{*0}$ candidates are required to pass either the L0Hadron T0S or the L0Global TIS criteria defined in Sects. 2.6 and 3.1. Hence, the weights are assigned

in two categories. The first category is labelled ‘TOSall’ and contains candidates that pass only the L0Hadron TOS requirements or both the L0Hadron TOS and L0Global TIS requirements. The candidates are assigned a weight denoted ϵ_{TOS} , which is the probability that at least one out of the six B daughter tracks passes the L0Hadron TOS trigger requirements. The individual probabilities with which each B daughter track (with a given E_T deposit in the HCAL) will pass the L0Hadron trigger threshold are labelled p_i . For a two-body B decay, the probabilities p_1 and p_2 would be combined as $\epsilon_{\text{TOS}} = p_1(1 - p_2) + p_2(1 - p_1) + p_1p_2$. The expression for ϵ_{TOS} in the six-track decay of $B^0 \rightarrow D(\rightarrow K_S^0(\rightarrow \pi\pi)h^+h^-)K^{*0}(\rightarrow K\pi)$ is built analogously. The value of ϵ_{TOS} is calculated for all candidates passing the full selection (apart from the hardware trigger requirement). The second category contains candidates that only pass the L0Global TIS requirements and are labelled ‘TISonly’. These events are assigned weights equal to $1 - \epsilon_{\text{TOS}}$.

The average values of ϵ_{TOS} and $1 - \epsilon_{\text{TOS}}$ over the Dalitz plot are evaluated in a coarsely-binned map and example distributions of the averaged values are shown in Fig. 6.12. The mean values of ϵ_{TOS} are higher in the ‘corners’ of the Dalitz plot, as these correspond to regions where one of the D daughter pions has a high momentum and the other has a low momentum, hence there is a higher probability that one of the D daughter tracks will satisfy the trigger requirements. The fraction of candidates in data within a $\pm 25 \text{ MeV}/c^2$ window of the fitted B^0 mass that are classified as TOSall is 0.55 ± 0.11 (0.65 ± 0.06) for LL (DD) $D \rightarrow K_S^0 \pi^+ \pi^-$ and 0.50 ± 0.25 (0.67 ± 0.27) for LL (DD) for $D \rightarrow K_S^0 K^+ K^-$. For each data category, TOSall and TISonly Dalitz plot efficiency maps are combined as a weighted average, using the above fractions.

The PID requirements placed on the K^{*0} daughters do not bias the D Dalitz plot, hence the efficiencies in simulation are used. The selection efficiency for $B^0 \rightarrow DK^{*0}$ candidates, apart from the hardware trigger requirement, is determined using the number of simulated candidates passing the selection, which is proportional to the

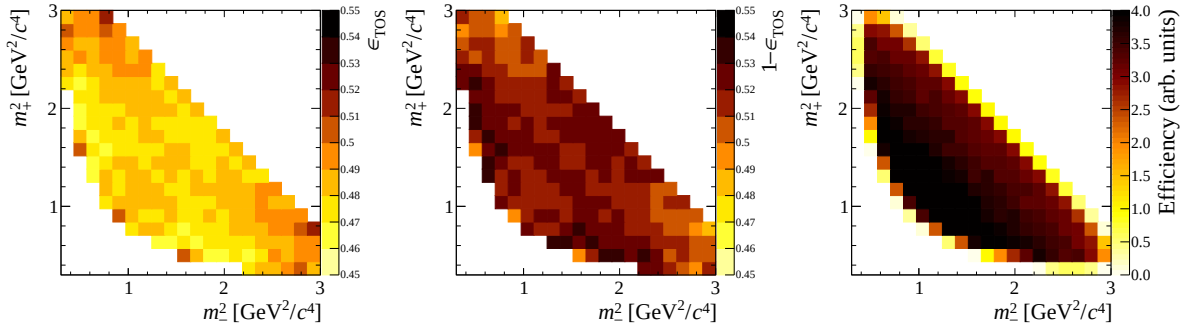


Figure 6.12. Dalitz plot efficiencies for simulated $B \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-) K^{*0}$, DD K_S^0 candidates in 2012 conditions. The (left) average values of ϵ_{TOS} , (centre) the average values of $1 - \epsilon_{\text{TOS}}$, and (right) the selection efficiencies excluding the hardware trigger are shown.

efficiency because the MC is generated uniformly in the Dalitz plot. An example Dalitz plot distribution of the selection efficiency is given in Fig. 6.12. The total efficiency is obtained by multiplying the hardware trigger and selection efficiencies. The Dalitz plot efficiencies of the $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decay are obtained purely from simulation, as the L0Muon trigger is modelled sufficiently accurately.

The above studies are performed with coarsely-binned efficiency maps for both channels, so that there is a sufficient number of events in each square bin. Square bins that cross the Dalitz plot kinematic boundary, which are less populated because only a fraction f_{bin} of the square bin is contained inside the Dalitz plot, are given a correction factor of $1/f_{\text{bin}}$. Square bins located outside but close to the Dalitz plot boundary that contain no candidates are assigned a value that is the average of neighbouring bins. They are then smeared using a cubic spline interpolation method. The correction for bins at the Dalitz plot edge ensures that the spline interpolation occurs smoothly at the border. Without the correction, the regions immediately outside the border would be treated as having values of zero. The final pixel maps are obtained by sampling the splines at intervals corresponding to the pixel width in the Dalitz plot binning schemes.

Efficiency maps for simulated $B_s^0 \rightarrow D \bar{K}^{*0}$ and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decays are shown in Fig. 6.13. In these example distributions, the Dalitz plot efficiencies for the different categories (LL, DD, 2011, 2012) have been combined. A variation of 50% between the

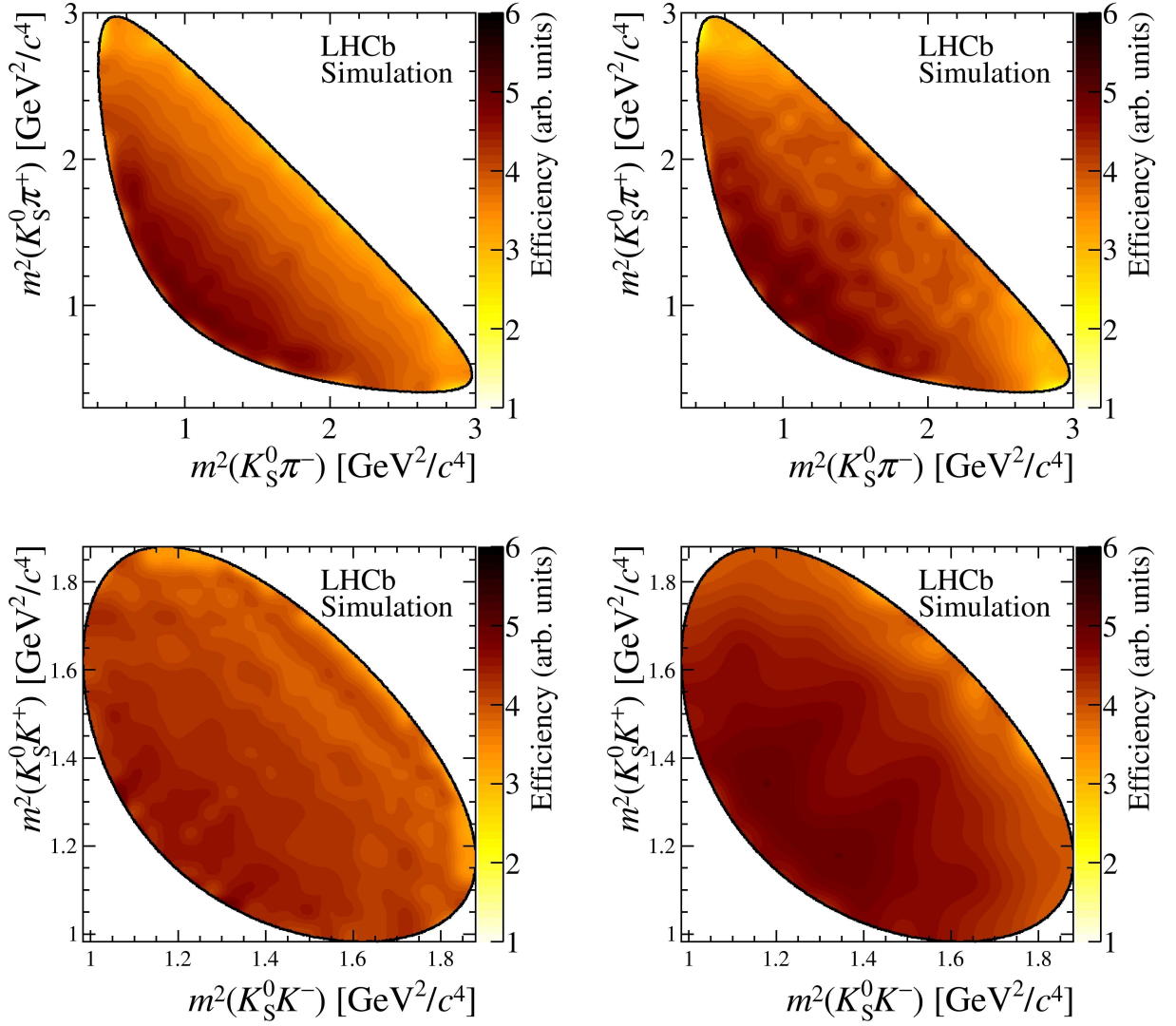


Figure 6.13. Efficiency profiles of (left) $B_s^0 \rightarrow D \bar{K}^{*0}$ and (right) $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ decays in simulation for the (top) $K_S^0 \pi^+ \pi^-$ and (bottom) $K_S^0 K^+ K^-$ final states, where a weighted average of the efficiency profiles for the different data categories has been performed.

highest and lowest efficiency regions is observed, although the efficiency changes within a Dalitz plot bin are not as large. The variation over the $D \rightarrow K_S^0 K^+ K^-$ Dalitz plot is smaller, at 35%. Similar behaviour is observed between the Dalitz plot efficiencies of the $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ sample here and those shown in Chapter 4, where the efficiency decreases in the corners of the Dalitz plot, however the decrease affects a smaller region here because of the changes in the selection.

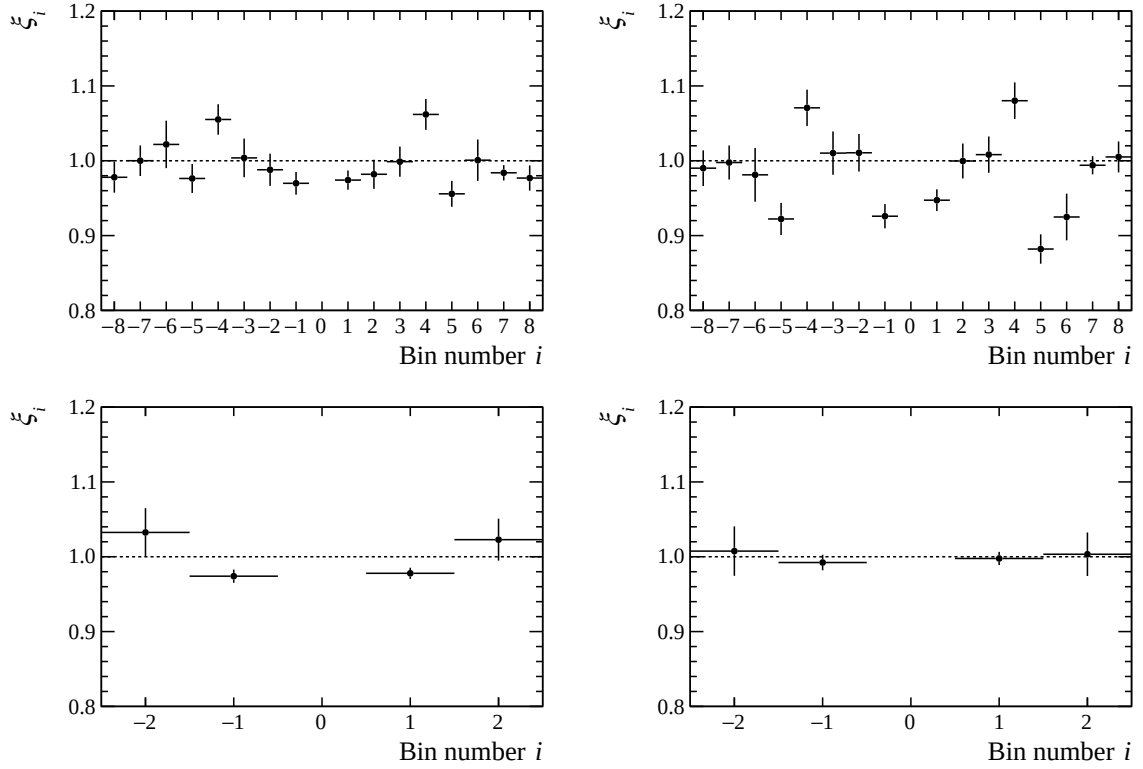


Figure 6.14. The efficiency correction factors ξ_i in each Dalitz plot bin of (top) $K_S^0 \pi^+ \pi^-$ and (bottom) $K_S^0 K^+ K^-$ candidates with (left) DD and (right) LL K_S^0 from 2012 data. The $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ bins are defined by the modified optimal and 2×2 binning schemes, respectively.

The efficiency correction factors ξ_i are determined from the $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D \bar{K}^{*0}$ efficiency maps and the BaBar amplitude models [33, 34]. The ξ_i values are calculated using Eq. (4.9), where the $B_s^0 \rightarrow D \bar{K}^{*0}$ efficiency profile replaces the $B \rightarrow D \pi$ profile, and the values of n_j are given by the efficiency profiles shown in Fig. 6.13. The values of ξ_i for 2012 data are given in Fig. 6.14. Similar values are obtained for 2011 data. As before, the values of ξ_i are within 10% of unity.

The binned semileptonic yields in Fig. 6.11 are combined with the ξ_i factors in Fig. 6.14 in a manner described in Sect. 4.3.3 to calculate the F_i parameters. The F_i parameters are determined for each year of data taking and K_S^0 category separately and are then combined in a weighted average according to the fractions in which they are observed in the $B^0 \rightarrow DK^{*0}$ signal region in data. The final F_i values are given in Table 6.6 and the fractions are shown later in Table 6.7. The total relative uncertainty

Table 6.6. The values of F_i for $B^0 \rightarrow DK^{*0}$ decays calculated using the efficiency-corrected semileptonic yields.

i	F_i	i	F_i
$D \rightarrow K_s^0 \pi^+ \pi^-$ (modified optimal)		$D \rightarrow K_s^0 K^+ K^-$ (2×2 binning)	
-8	0.0259 ± 0.0013	-2	0.2140 ± 0.0111
-7	0.0878 ± 0.0027	-1	0.2203 ± 0.0091
-6	0.0648 ± 0.0027	+1	0.2901 ± 0.0104
-5	0.0299 ± 0.0013	+2	0.2756 ± 0.0126
-4	0.0739 ± 0.0024		
-3	0.2440 ± 0.0073		
-2	0.1946 ± 0.0053		
-1	0.0425 ± 0.0016		
+1	0.0109 ± 0.0007		
+2	0.0229 ± 0.0012		
+3	0.0210 ± 0.0011		
+4	0.0461 ± 0.0018		
+5	0.0205 ± 0.0010		
+6	0.0085 ± 0.0007		
+7	0.0459 ± 0.0015		
+8	0.0607 ± 0.0020		

on F_i is 5% or less in all of the bins, and is a combination of the statistical uncertainty of the raw semileptonic yields and the uncertainty on ξ_i due to the limited size of the MC samples. The two contributions are comparable in magnitude.

6.4 Fitting the CP violation observables

The parameters $x_{\pm}^*, y_{\pm}^*$ are determined from a simultaneous extended maximum likelihood fit of the populations of B^0 and $\bar{B}^0$ decays in the Dalitz plot bins. The shape parameters for the signal and background invariant mass distributions are fixed to the values determined in the mass fit described in Sect. 6.2. It is assumed that with the small sample size there is no sensitivity to variation of the shapes in different bins.

6.4.1 Fit configuration

The $B^0 \rightarrow DK^{*0}$ data are split according the charge of the K^{*0} daughter kaon so that the Dalitz plots for B^0 and $\bar{B}^0$ data are analysed separately. As with the fit to determine the mass parameterisation, data from the two K_s^0 categories are merged. Hence, the data are divided into the following 40 bins: $\{2 \times 8 K_s^0 \pi^+ \pi^- \text{ bins}, 2 \times 2 K_s^0 K^+ K^- \text{ bins}\} \otimes \{B^0, \bar{B}^0\}$.

Signal and background yields

The $B^0 \rightarrow DK^{*0}$ yield in each bin is determined using Eq. (6.1) and the total yield of $B^0 \rightarrow DK^{*0}$, which is a free parameter in each of four categories ($\{B^0, \bar{B}^0\} \otimes \{K_s^0 \pi^+ \pi^-, K_s^0 K^+ K^-\}$). The parameters $x_{\pm}^*$ and $y_{\pm}^*$ are allowed to vary.

It is verified in simulation that the reconstruction efficiency over the D Dalitz plot does not depend on the parent B decay, and hence the same F_i values are used to describe the D Dalitz plot distribution of the $B_s^0 \rightarrow D\bar{K}^{*0}$, $B^0 \rightarrow D\rho^0$, $B \rightarrow DK$ and $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ backgrounds. The total yield of $B_s^0 \rightarrow D\bar{K}^{*0}$ decays integrated over the Dalitz plot for each category is a free parameter, with the fractions in each Dalitz plot bin assigned such that the yield of $B_s^0 \rightarrow D\bar{K}^{*0}$ decays in bin i is the total yield in that category multiplied by F_{-i} . In other words, it is assumed that CP violation in $B_s^0 \rightarrow D\bar{K}^{*0}$ is negligible [92], meaning the decay of the B_s^0 ($\bar{B}_s^0$) meson contains a $\bar{D}^0$ (D^0) meson. The total yields of the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$, $B^0 \rightarrow D\rho^0$ and $B \rightarrow DK$ backgrounds are determined by multiplying the total yield of $B_s^0 \rightarrow D\bar{K}^{*0}$ in that category by the fitted values of R_s , R_ρ and R_{DK} , respectively. Several assumptions are made about the Dalitz plot distributions of these backgrounds. The CP violation in $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decays is expected to be negligible as the underlying CKM factors are the same as that for $B_s^0 \rightarrow D\bar{K}^{*0}$ decays. Hence, the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decays are assumed to be distributed over the Dalitz plot in the same way as $B_s^0 \rightarrow D\bar{K}^{*0}$ decays, and the values of α , $G_{\lambda=0}$ and $G_{\lambda=\pm 1}$ are assumed to be the same in each bin. The D meson from $B^0 \rightarrow D\rho^0$ decays

is assumed to be an equal admixture of D^0 and $\bar{D}^0$ and hence the yield is distributed according to $(F_{+i} + F_{-i})/2$. In the case of the $B \rightarrow DK$ decay, CP violation is expected and it is assumed that the yield is distributed according to Eq. (4.2), where the values of $x_{\pm}$ and $y_{\pm}$ are those determined in Chapter 5. The yield of the combinatorial background in each bin is a free parameter, apart from the yields in bins in which an auxiliary fit determines it to be negligible. It is necessary to set these to zero to facilitate the calculation of the covariance matrix.

The parameters c_i , s_i , F_i and κ

The values of c_i and s_i are constrained to their measured values from CLEO [3], using a multivariate Gaussian constraint that takes into account statistical and systematic correlations. This strategy marks a departure from the $B \rightarrow DK$ analysis, where the c_i and s_i parameters are fixed in the fit and the systematic uncertainties are determined by performing many fits to data with smeared values of c_i and s_i . Due to the small dataset, the probability that the alternative fits can prefer secondary minima is increased, hence this method cannot be used to determine the associated systematic uncertainty. Instead, by constraining the c_i and s_i parameters, their uncertainties (primarily statistical) are included in the statistical uncertainties of $x_{\pm}^*$ and $y_{\pm}^*$.

The values of F_i are fixed to the central values listed in Table 6.6. Although the F_i parameters are implemented as constraints in the $B \rightarrow DK$ analysis, this is observed to reduce the fit stability and increase the fit biases on $x_{\pm}^*, y_{\pm}^*$. The impact of the F_i uncertainties on the fit is evaluated separately as a systematic uncertainty.

The coherence factor κ is measured in Ref. [92] and has a value $\kappa = 0.958^{+0.005+0.002}_{-0.010-0.045}$ in the region where $m(K\pi)$ is within $\pm 50 \text{ MeV}/c^2$ of the nominal K^{*0} mass and $|\cos \theta^*| > 0.4$. This region is chosen to obtain a large value of κ and to facilitate combination with γ measurements in Refs. [90, 92], which impose the same limits. The use of this

κ value requires that the efficiency within the selected K^{*0} region is uniform. The level of non-uniformities in the B Dalitz plot in this analysis are of similar size to that investigated in Ref. [90], where the impact of efficiency effects in this region on κ is found to be small. The κ variations are dominated by the uncertainties associated with the amplitude model and not the B Dalitz plot efficiency. Studies in simulation also show that the presence of θ^* in the BDT has no significant impact on the value of κ . The value of $\kappa = 0.958$ is fixed in the fit.

Fitter validation and choice of binning scheme

An ensemble of 10 000 pseudoexperiments is generated to validate the fit procedure and choose the Dalitz plot binning scheme. In each pseudoexperiment the yields and shapes of the signal and background candidates are generated according to the expected distributions in data. The combinatorial background is generated assuming a contribution of 50% D^0 and 50% $\bar{D}^0$ decays. Both $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ final states are generated. Each pseudoexperiment is generated with a unique set of c_i and s_i values, which have been smeared according to their measured errors and correlations. Three sets of $x_{\pm}^*, y_{\pm}^*$ values consistent with previous measurements are used ($r_{B^0} = 0.23$, $\delta_{B^0} = 340^\circ$, $\gamma = 73^\circ$; $r_{B^0} = 0.31$, $\delta_{B^0} = 62^\circ$, $\gamma = 73^\circ$; $r_{B^0} = 0.15$, $\delta_{B^0} = 62^\circ$, $\gamma = 73^\circ$) [25]. The results using the optimal and modified optimal binning schemes for $K_S^0 \pi^+ \pi^-$ are compared. The statistical sensitivity is found to be compatible for both binning schemes, but the modified optimal binning results in smaller fit biases and is therefore chosen. The biases in the central values (i.e. the mean values fitted in the pseudoexperiments are slightly offset from the input) have a magnitude of approximately 10% of the statistical uncertainty in the modified optimal binning and result from the low event yields, which can cause the fitted yields of signal or background to be consistent with zero in some bins. This also suggests why the modified optimal binning produces smaller biases, since no individual bin is very small.

6.4.2 Uncorrected fit results and fit bias

The raw central values and statistical uncertainties from the fit to data are:

$$\begin{aligned} x_+^* &= 0.042 \pm 0.345, \quad y_+^* = -0.729 \pm 0.277, \\ x_-^* &= -0.280 \pm 0.198, \quad y_-^* = 0.225 \pm 0.214, \end{aligned}$$

which correspond to central values of $r_{B^0} = 0.55$, $\delta_{B^0} = 204^\circ$ and $\gamma = 71^\circ$ (method explained in Chapter 7). The size of the fit bias depends on the values of the CP observables, hence the values of r_{B^0} , δ_{B^0} and γ obtained from the uncorrected fit result are used to estimate the bias correction. 10 000 pseudoexperiments are generated and the distribution of the fitted values, errors and pulls for the four $x_\pm^*, y_\pm^*$ parameters are shown in Fig. 6.15. The bias is determined using the difference between the input value to generate the pseudoexperiments and the mean fitted value of x_+^* , x_-^* , y_+^* and y_-^* to be +0.01, -0.03, -0.08, and +0.08, respectively. The quoted fit results in Sect. 6.5 are corrected by this offset.

From the above study (Fig. 6.15) we also note that the pull distributions on y_+^* and y_-^* are non-Gaussian and contain asymmetric tails. Additional pseudoexperiments are generated in which the yields are inflated by a factor of 100. The results are shown in Fig. 6.16 and the pull distributions here are Gaussian with unit width, which indicates that the non-Gaussian pulls observed in Fig. 6.15 are due to low statistics. A small fitter bias remains in the high statistics pseudoexperiments, most notably in y_+^* where the mean of the pull distribution is -0.10 ± 0.02 . This is the result of some of the s_i values lying outside the physical range of $c_i^2 + s_i^2 \leq 1$.

6.4.3 Fit results

After correction of fitter biases, the CP violation observables are:

$$\begin{aligned} x_+^* &= 0.05 \pm 0.35, \quad y_+^* = -0.81 \pm 0.28, \\ x_-^* &= -0.31 \pm 0.20, \quad y_-^* = 0.31 \pm 0.21, \end{aligned}$$

where the uncertainties are statistical only. The fit projections are shown in Figs. 6.17 and 6.18. The statistical uncertainties are compatible with those predicted by the simulated pseudoexperiments. Although the uncertainty on x_+^* is larger than the mean error in pseudoexperiments (Fig. 6.15), 16% of the generated pseudoexperiments have a larger uncertainty. The larger uncertainty on $x_{\pm}^*$ is also observed in the results using the model-dependent method [45], where identical data are analysed, hence this can be attributed to a statistical fluctuation. (The model-dependent results are discussed further in Sect. 7.2.1.)

The measured values of $x_{\pm}^*, y_{\pm}^*$ from the fit to data, with their likelihood contours corresponding to statistical uncertainties only, are displayed in Fig. 6.19. The expected signature for CP violation is that the vectors defined by the coordinates (x_+^*, y_+^*) and (x_-^*, y_-^*) are non-zero in magnitude and have a non-zero opening angle. This opening angle is equal to 2γ . No evidence for CP violation is observed. The fitted values of c_i and s_i are superimposed with the input values (the CLEO measurements) in Fig. 6.20. The fitted values are shifted by much less than 1σ from their inputs. This is expected given the low yields in this dataset. The fitted errors on c_i and s_i are nearly identical to the inputs.

To investigate whether the binned fit gives an adequate description of the distribution of events over the Dalitz plot, the signal yield in each bin is fitted directly as a cross-check. Figure 6.21 shows the results separately for the sum of B^0 and $\bar{B}^0$ yields,

and their difference. A comparison of these yields and those predicted by the fitted values of $x_{\pm}^*$ and $y_{\pm}^*$ shows good agreement. The difference between the B^0 and $\bar{B}^0$ yields includes the expected difference for $x_{\pm}^* = y_{\pm}^* = 0$. The expectation for this hypothesis is not exactly at zero in each bin as the difference is calculated using the total B^0 and $\bar{B}^0$ yields which are not identical. While these yields show consistency with the fitted values of $x_{\pm}^*$ and $y_{\pm}^*$, the uncertainties are large and there is also consistency with the hypothesis of CP symmetry.

6.4.4 Systematic uncertainties

Systematic uncertainties are evaluated on the measurements of $x_{\pm}^*, y_{\pm}^*$. Unless otherwise described, the systematic uncertainties are determined using pseudoexperiments, which are generated in various alternative configurations, and fitted with the default method described in Sect. 6.4.1. The mean shift in the fitted values of $x_{\pm}^*$ and $y_{\pm}^*$ in comparison to their input values is taken as the systematic uncertainty.

A separate systematic from the uncertainties on c_i and s_i is not evaluated, since this source of uncertainty is included within the statistical uncertainties. A systematic from the effects of a non-uniform Dalitz plot efficiency on the values of c_i and s_i is not determined, since the estimated shifts on c_i and s_i (calculated using the method outlined in Sect. 5.3 but with the $B_s^0 \rightarrow D\bar{K}^{*0}$ efficiency profile) are two orders of magnitude smaller than the errors of the Gaussian constraints.

Uncertainty due to the F_i parameters

A systematic uncertainty due to the efficiency correction in calculating F_i is evaluated by considering a conservative estimate of the level of MC-data differences. The uncertainty is estimated using only $B_s^0 \rightarrow D\bar{K}^{*0}$ MC and the BaBar $D \rightarrow K_s^0 h^+ h^-$ amplitude models to calculate alternative F_i parameters. The default and alternative values of F_i are shown

Table 6.7. Fractions of each category in data in the region $\pm 25 \text{ MeV}/c^2$ about the known B^0 and B_s^0 mass, used to average the F_i parameters from each category.

Year	K_s^0 type	B^0 window	B_s^0 window
2012	LL	0.17 ± 0.03	0.21 ± 0.03
2012	DD	0.51 ± 0.04	0.44 ± 0.04
2011	LL	0.13 ± 0.03	0.15 ± 0.03
2011	DD	0.19 ± 0.03	0.20 ± 0.03

in Fig. 6.22. No error bars are shown for the alternative F_i as the same $B_s^0 \rightarrow D\bar{K}^{*0}$ MC sample is used to calculate the default and alternative F_i and therefore their statistical errors are correlated. The deviations between the two sets of F_i are of order 1σ in each bin. The difference between the default and alternative F_i is of a similar size to or larger than the uncertainties on the default F_i . Hence, the alternative F_i are used in the generation of pseudoexperiments to determine the systematic uncertainty. The systematic uncertainty associated with the efficiency correction for x_+^* , x_-^* , y_+^* and y_-^* is 0.019, 0.034, 0.021 and 0.005, respectively.

A further uncertainty on the F_i parameters results from the combinations of the individual F_i parameters from the different data categories (year of data taking and K_s^0 type). A second alternate set of F_i are obtained using the fractions in which each data category is observed in the B_s^0 mass window instead of the B^0 mass window. The fractions are compared in Table 6.7 and are statistically consistent in the two mass windows. The associated uncertainty on x_+^* , x_-^* , y_+^* and y_-^* is 0.007, 0.001, 0.007 and 0.008, respectively.

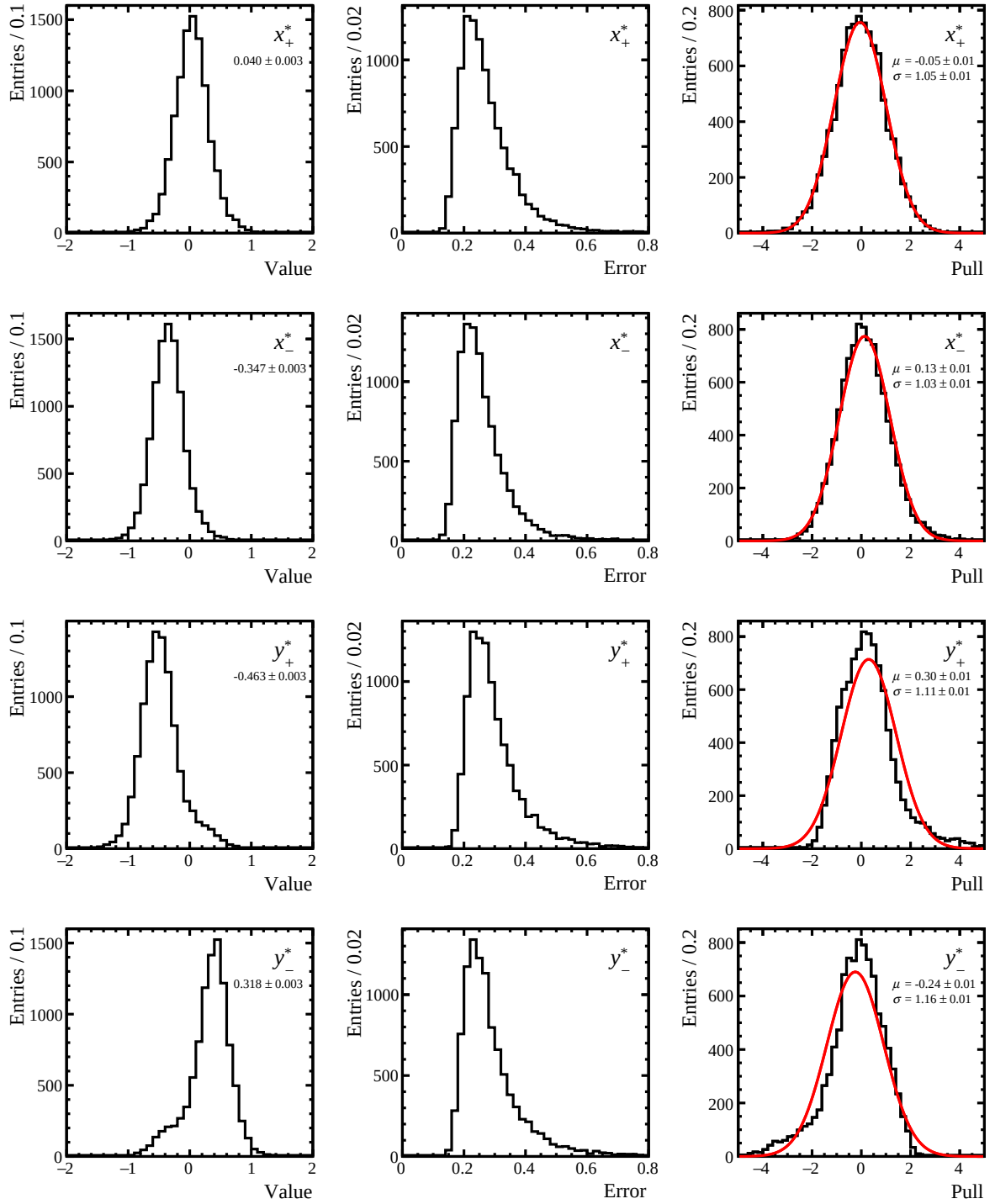


Figure 6.15. The results of fitting 10 000 pseudoexperiments generated with the parameters $\kappa = 0.958$, $r_{B^0} = 0.55$, $\delta_{B^0} = 204^\circ$ and $\gamma = 71^\circ$ ($x_+^* = 0.048$, $x_-^* = -0.375$, $y_+^* = -0.548$ and $y_-^* = 0.402$). The distributions of the (left to right) central values, errors and pulls for (top to bottom) x_+^* , x_-^* , y_+^* and y_-^* are shown. The mean central values are given on the left-hand plots and are used to determine the bias correction.

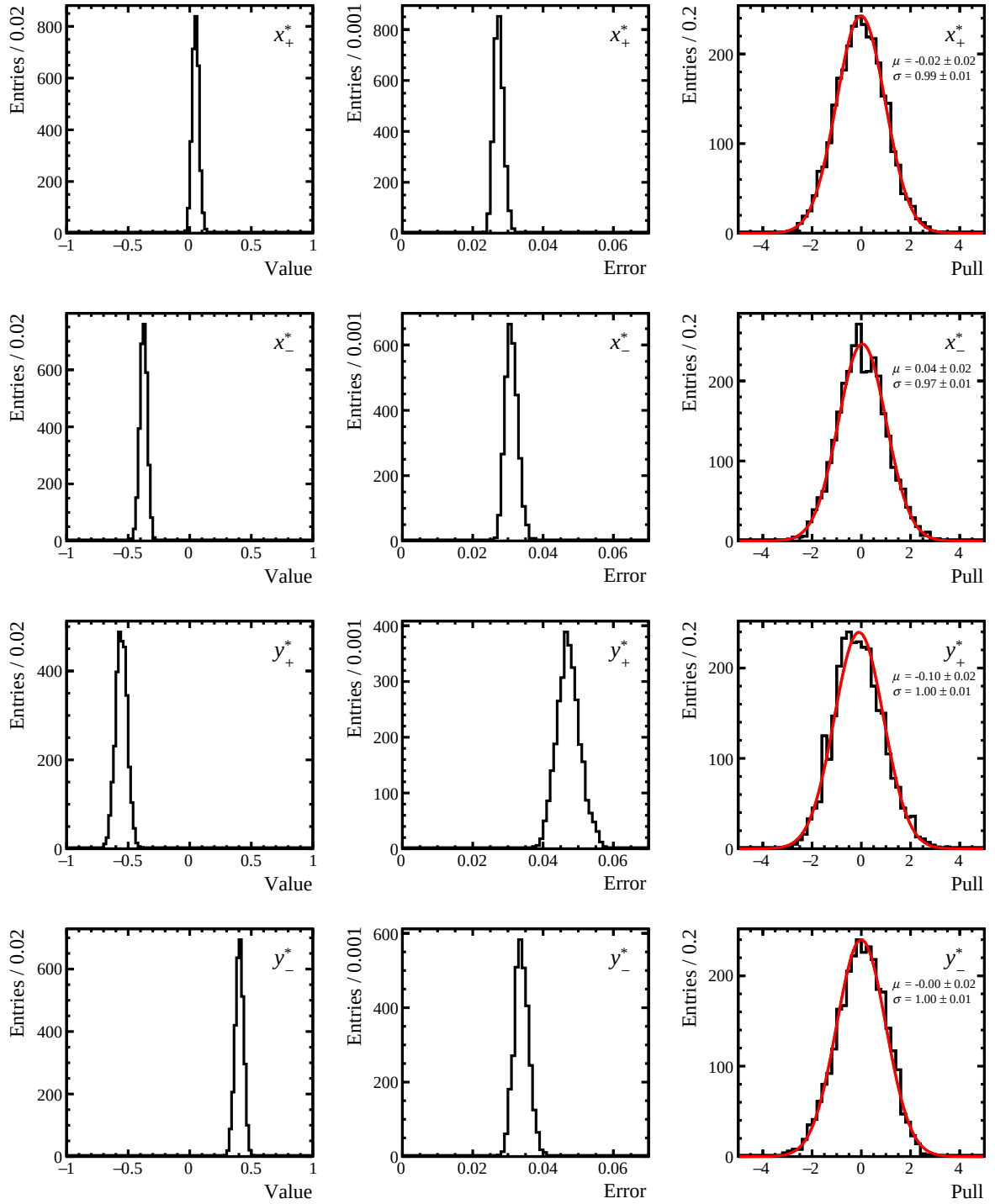


Figure 6.16. The results of fitting 10 000 pseudoexperiments generated with the parameters $\kappa = 0.958$, $r_{B^0} = 0.55$, $\delta_{B^0} = 204^\circ$ and $\gamma = 71^\circ$, with a factor of 100 higher statistics than the data. The distributions of the fit value, fit error and pull of (top to bottom) x_+^* , x_-^* , y_+^* and y_-^* are shown.

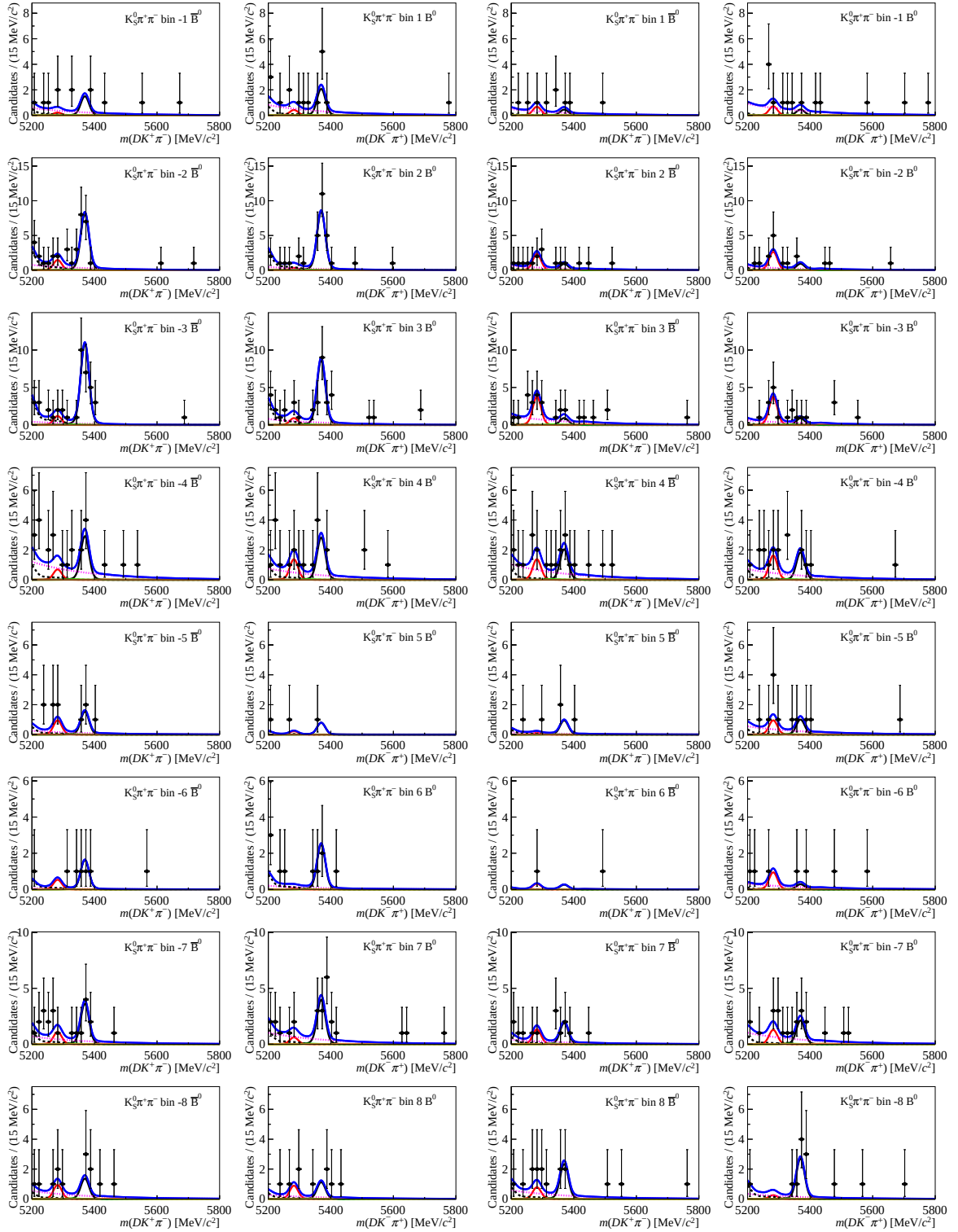


Figure 6.17. Fit projections for $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 \pi^+ \pi^-$ data. The first (third) column shows $\bar{B}^0$ data in Dalitz plot bin numbers -1 to -8 (1 to 8) and the corresponding bin in B^0 data is shown in the same row of the second (fourth) column. The (blue) fit result is superimposed with the following fit components: (red) $B^0 \rightarrow DK^{*0}$, (black) $B_s^0 \rightarrow D \bar{K}^{*0}$, (dashed black) $B_s^0 \rightarrow D^0 \bar{K}^{*0}$, (magenta) combinatorial, (green) $B^0 \rightarrow D \rho^0$ and (brown) $B \rightarrow DK$ decays.

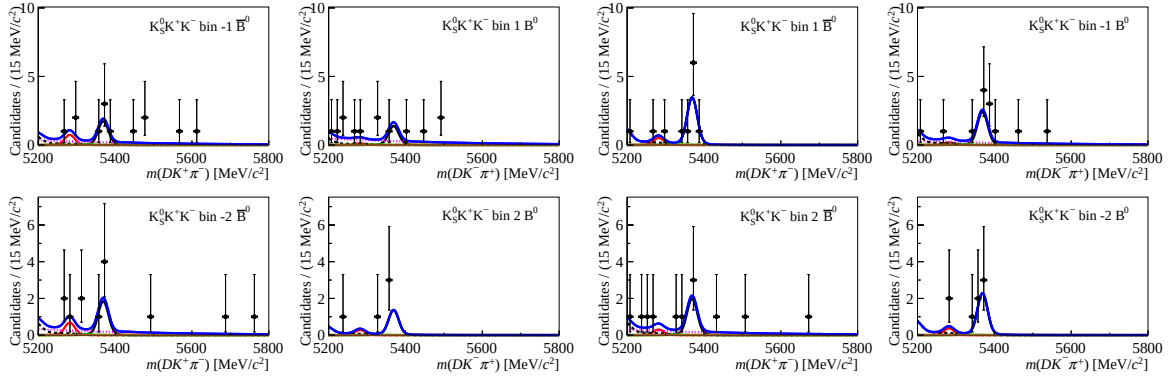


Figure 6.18. Fit projections for $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 K^+ K^-$ data. The first (third) column shows $\bar{B}^0$ data in Dalitz plot bin numbers -1 to -2 (1 to 2) and the corresponding bin in B^0 data is shown in the same row of the second (fourth) column. The (blue) fit result is superimposed with the following fit components: (red) $B^0 \rightarrow DK^{*0}$, (black) $B_S^0 \rightarrow D\bar{K}^{*0}$, (dashed black) $B_S^0 \rightarrow D^{*0}\bar{K}^{*0}$, (magenta) combinatorial, (green) $B^0 \rightarrow D\rho^0$ and (brown) $B \rightarrow DK$ decays.

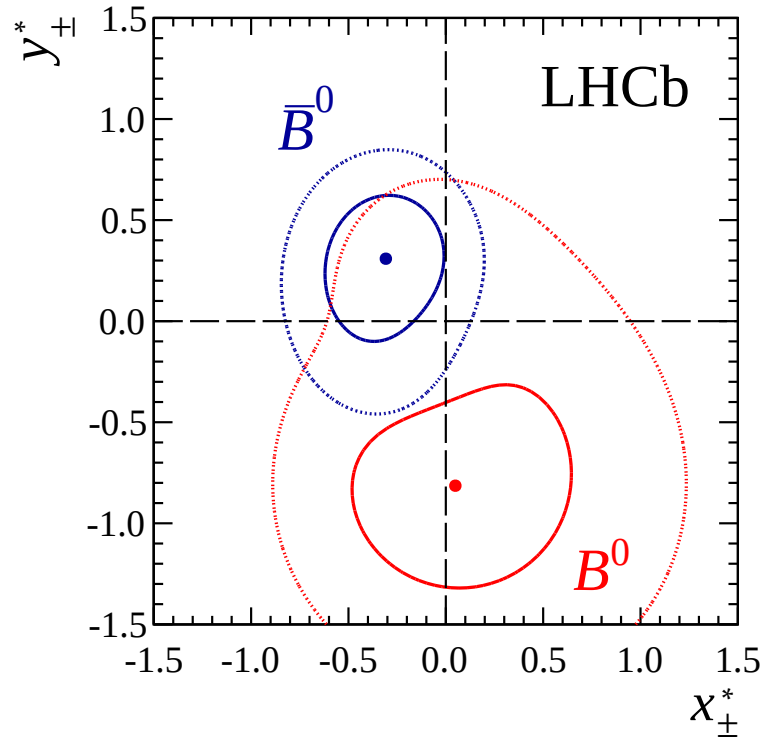


Figure 6.19. The two-dimensional confidence intervals on $x_{\pm}^*$, $y_{\pm}^*$ showing the likelihood contours for statistical uncertainties. The points indicate the central values and the bands correspond to the 1σ and 2σ intervals. The red bands correspond to the contours on (x_+^*, y_+^*) and the blue bands correspond to the contours on (x_-^*, y_-^*) .

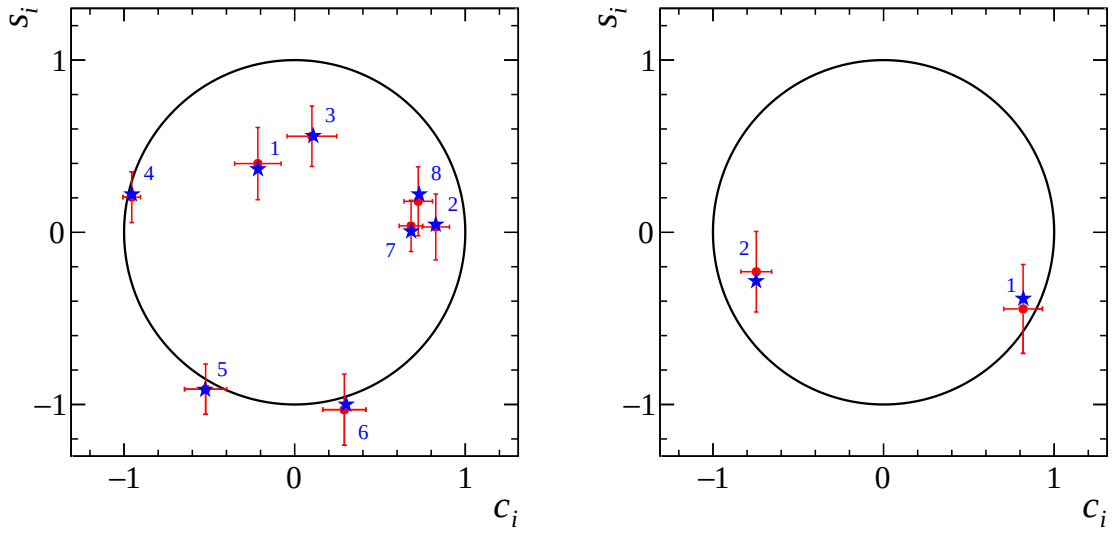


Figure 6.20. Fit results of the c_i and s_i parameters for (left) $D \rightarrow K_s^0 \pi^+ \pi^-$ and (right) $D \rightarrow K_s^0 K^+ K^-$, showing (red points) the CLEO central values and uncertainties that are input into the fit, and (blue stars) the fitted central values.

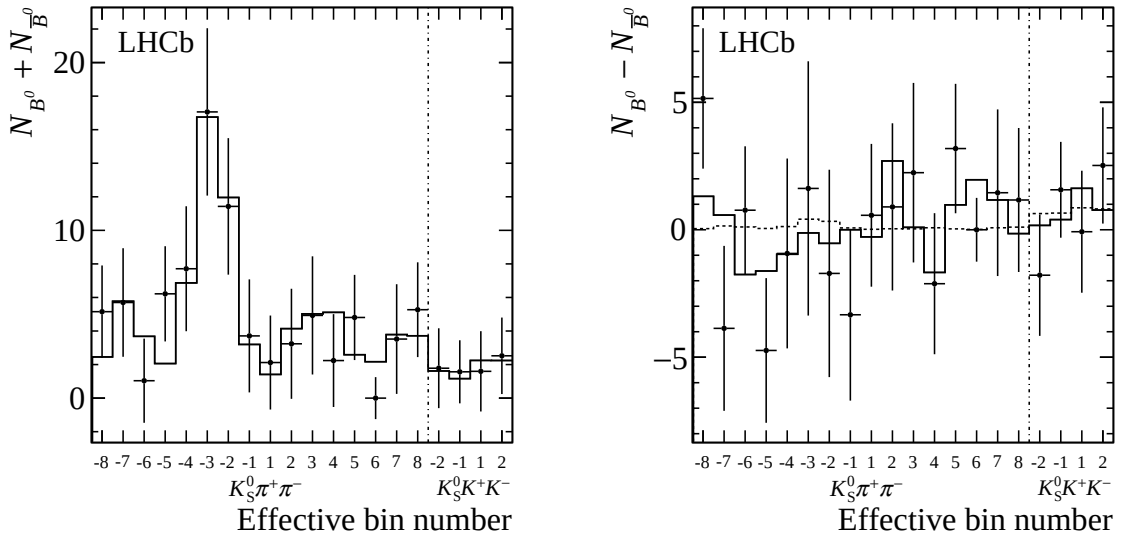


Figure 6.21. The values of (left) $N_{B^0} + N_{\bar{B}^0}$ and (right) $N_{B^0} - N_{\bar{B}^0}$ from two fit methods are shown. The data points show the result of fitting the B yields directly while the solid histogram indicates the expected yield calculated from the best fit values of $x_{\pm}^*, y_{\pm}^*$. On the right-hand plot, the dashed histogram shows the expectation for $x_{\pm}^* = y_{\pm}^* = 0$.

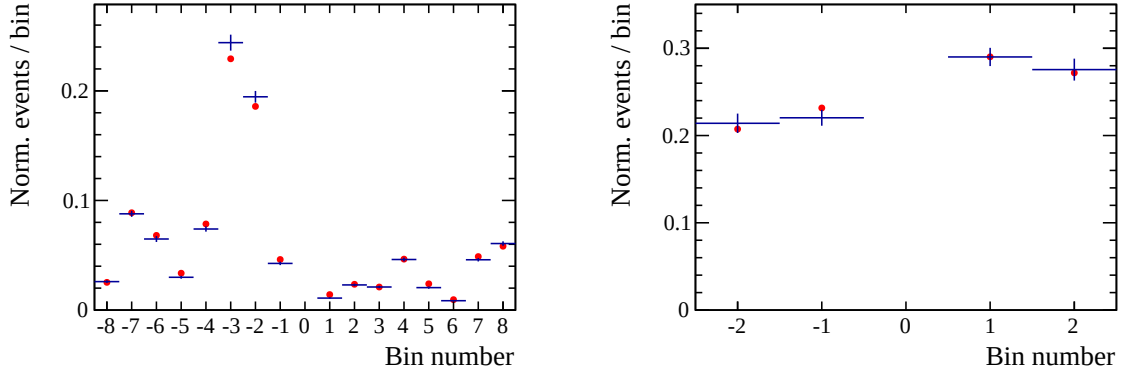


Figure 6.22. The (blue) default and (red) alternative values of F_i used to evaluate the systematic due to imperfect modelling in the simulation. The values are shown for the (left) $K_S^0 \pi^+ \pi^-$ and (right) $K_S^0 K^+ K^-$ final states.

Uncertainties due to the mass fit parameterisation

Several systematic uncertainties are associated with the parameterisation of the B mass distribution. These arise from uncertainties in the shape of the $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$ background, the size of the $B^0 \rightarrow D \rho^0$ background, CP violation in $B \rightarrow DK$ decays, the shape used to describe the signal peak and the presence of backgrounds that are neglected in the nominal fit due to their small yield.

The uncertainty in the shape of the $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$ background comes from the relative contribution of the different D^{*0} decay and helicity state components, each of which have a different $m(DK^\pm \pi^\mp)$ distribution. A different parameterisation of the data, with the fit range extending down to $4900 \text{ MeV}/c^2$, is used in the model-dependent measurement [45] and results in the value $\alpha = 0.9 \pm 0.1$. Adjusting for the difference in the mass range this becomes $\alpha = 0.91$, in comparison to the value of $\alpha = 0.74^{+0.13}_{-0.13}$ obtained in the nominal fit. The uncertainty associated with the alternate value of α is determined for x_+^* , x_-^* , y_+^* and y_-^* to be 0.002, 0.005, 0.021 and 0.020, respectively. A separate systematic uncertainty is evaluated for the relative fraction of $D^{*0} \rightarrow D^0 \pi^0$ and $D^{*0} \rightarrow D^0 \gamma$ decays in the $B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$ background. The uncertainties in the relative fractions come from uncertainties in the branching fractions of the D^{*0} decays and

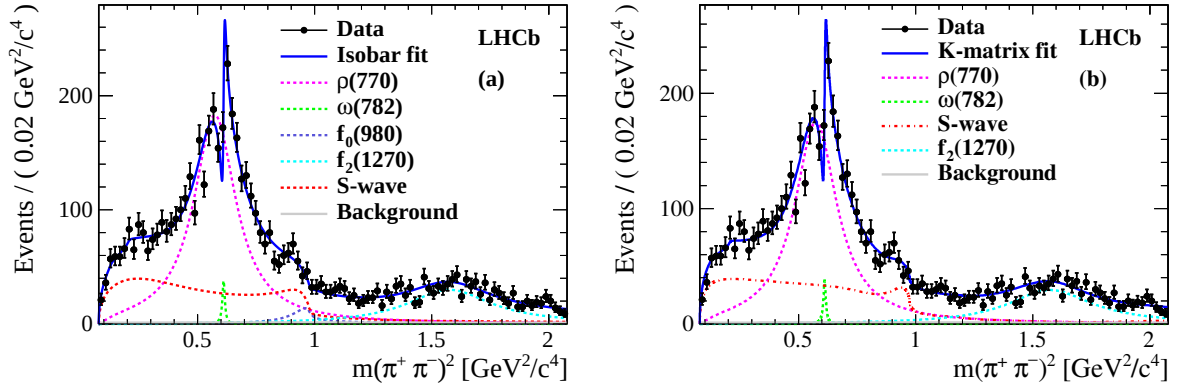


Figure 6.23. The distribution of $B^0 \rightarrow \bar{D}^0 \pi^+ \pi^-$ decays in the $m^2(\pi^+ \pi^-)$ projection near the $\rho(770)$ mass region is shown with the fit results from the (a) isobar and (b) K-matrix amplitude models superimposed. The (magenta) $\rho(770)$ and (red) S-wave contributions are shown. Reproduced from Ref. [93].

in the selection efficiencies determined from simulation. In this case the systematic uncertainty is estimated by fitting the data repeatedly with different fractions smeared around the central values. The resulting uncertainties on $x_{\pm}^*, y_{\pm}^*$ are no larger than 0.010.

The ratio R_{ρ} , which gives the size of the $B^0 \rightarrow D\rho^0$ background, is calculated ignoring the $B^0 \rightarrow D\pi^+\pi^-$ S-wave, which will contribute if the misidentified $K^+\pi^-$ mass falls within the K^{*0} mass window. The amplitude analysis of $B^0 \rightarrow D\pi^+\pi^-$ decays in Ref. [93] is used to determine the potential size of the S-wave contribution. The $m^2(\pi^+\pi^-)$ projection in the region close to the $\rho(770)$ from Ref. [93] is reproduced in Fig. 6.23. Assuming that the additional S-wave contribution will have the same $m(DK\pi)$ distribution, a rough estimate from these fits is that the S-wave would contribute an additional 50% of events to the ' $B^0 \rightarrow D\rho^0$ ' background. The systematic uncertainty is estimated by generating pseudoexperiments with R_{ρ} increased by 50%. The resulting uncertainties on $x_{\pm}^*, y_{\pm}^*$ are lower than 4×10^{-3} .

In the default fit the CP parameters of the $B \rightarrow DK$ background are fixed to the central values measured in Sect. 5.4. When the data are fitted repeatedly with multiple values of the $x_{\pm}, y_{\pm}$ parameters, smeared according to the measured uncertainties and correlations, it is found that the mean shifts in $x_{\pm}^*, y_{\pm}^*$ are less than 0.001.

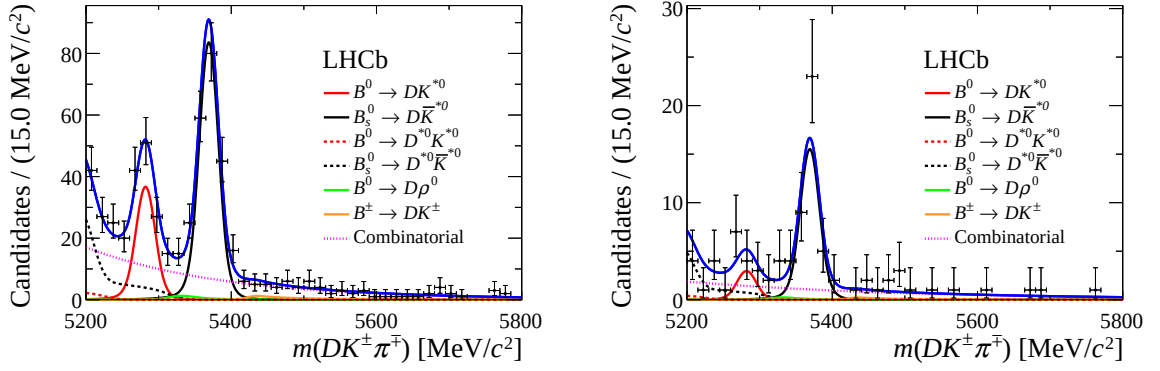


Figure 6.24. Results of the mass fit when the (dashed red) $B^0 \rightarrow D^{*0}K^{*0}$ background is included. The data for (left) $K_S^0\pi^+\pi^-$ and (right) $K_S^0K^+K^-$ are shown with the fit results superimposed.

The systematic uncertainty from the contributions of $B^0 \rightarrow D^{*0}K^{*0}$ decays, which are ignored in the default fit, is considered. The $B^0 \rightarrow D^{*0}K^{*0}$ decays are described with the same shape functions as the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ decays, but shifted by the $B_s^0 - B^0$ mass difference. The B mass fit described in Sect. 6.2 is performed with this background included, where the yield of $B^0 \rightarrow D^{*0}K^{*0}$ decays is constrained to be $(2.1 \pm 1.1)\%$ relative to that of the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$. The ratio is determined in a similar manner to R_s for the $B_s^0 \rightarrow D^{*0}\bar{K}^{*0}$ background. The results of including the $B^0 \rightarrow D^{*0}K^{*0}$ component in the mass fit are shown in Fig. 6.24, and it is observed that there is only a small shift in the mass fit parameters. However, its CP parameters are unknown, hence pseudoexperiments are generated with the $B^0 \rightarrow D^{*0}K^{*0}$ background decaying in three different CP violating hypotheses. The three hypotheses are: no CP violation (pure $\bar{D}^0$ associated with a K^{*0}), CP violation with the same parameters as $B^0 \rightarrow DK^{*0}$, and a $B_s^0 \rightarrow D\bar{K}^{*0}$ -like distribution (pure D^0 associated with a K^{*0}). The systematic uncertainty on each CP parameter is assigned using the largest shift from the three scenarios, which is found to be less than 0.01.

The $B^\pm \rightarrow D^{*0}\pi^\pm$, $B^\pm \rightarrow D^{*0}K^\pm$ and $B^\pm \rightarrow D\pi^\pm\pi^+\pi^-$ decays are ignored in the default fit as they are estimated to contribute approximately 2–4 events each. A systematic uncertainty from neglecting each of these decays is evaluated. Their shapes are determined from simulation, weighted for the PID efficiencies, and are shown in Fig. 6.25. The

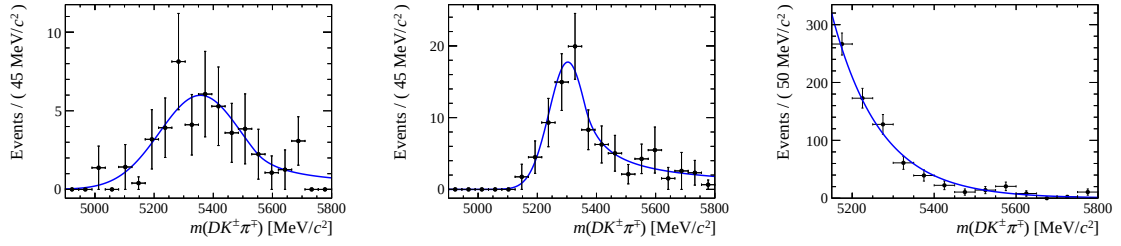


Figure 6.25. The $m(DK\pi)$ distribution of simulated (left) $B^\pm \rightarrow D^{*0}K^\pm$, (centre) $B^\pm \rightarrow D^{*0}\pi^\pm$ and (right) $B^\pm \rightarrow D\pi^\pm\pi^+\pi^-$ decays with their respective shape functions superimposed.

$B^\pm \rightarrow D^{*0}h^\pm$ decays are modelled with a Crystal Ball function, while the $B^\pm \rightarrow D\pi^\pm\pi^+\pi^-$ background is modelled with an exponential. Further pseudoexperiments are generated with each of these three decays added, and are fitted with the default configuration, demonstrating that the uncertainty due to ignoring these decays is 7×10^{-3} or less for all CP observables.

A final systematic uncertainty associated with the mass fit is the choice of shape function used to describe the signal. An imperfect description of the signal shape, due to the tail parameters which are fixed from simulation, can lead to biases in the signal yields. An alternative shape to describe the B^0 and B_s^0 signals is the sum of three Gaussian functions. The mean and width of the primary Gaussian is determined by performing a mass fit to data, with the relative means and widths of the two secondary Gaussians are taken from simulation. The associated systematic uncertainties on x_+^* , x_-^* , y_+^* and y_-^* are 0.005, 0.003, 0.003, and 0.002, respectively.

Uncertainty due to migration on the Dalitz plot

The systematic uncertainty from candidates being assigned the wrong Dalitz plot bin number, when reconstruction effects cause shifts in the measured values of m_+^2 and m_-^2 away from their true values, is considered in an identical manner to that described in Sect. 5.3. For both $B_s^0 \rightarrow D\bar{K}^{*0}$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ decays the width of the shift in $m_\pm^2$ from the true simulated values is $0.005 \text{ GeV}^2/c^4$ ($0.006 \text{ GeV}^2/c^4$) for candidates

with LL (DD) $K_S^0 \pi^+ \pi^-$ decays. For the $K_S^0 K^+ K^-$ final state the width is approximately $0.004 \text{ GeV}^2/c^4$ for both $B_S^0 \rightarrow D \bar{K}^{*0}$ and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$. This is small compared to the typical width of a Dalitz plot bin, but net migration can occur if the candidate lies close to a bin edge. To first order this effect is accounted for using the control channel, but residual effects enter because the non-zero value of r_{B^0} in the signal decay. The uncertainty due to these residual effects is determined via pseudoexperiments, in which alternative F_i values, altered from the default to reflect residual migration effects using the procedure in Sect. 5.3, are used as inputs. The alternate values of F_i deviate from the nominal values by 2% at most, and the uncertainties on x_+^* , x_-^* , y_+^* and y_-^* are 0.002, 0.006, 0.008, and 0.000, respectively.

Uncertainty due to the measurement of κ

The value of κ has an uncertainty associated with it, hence pseudoexperiments are generated assuming a value of $\kappa = 0.912$, which corresponds to the central value of κ lowered by one standard deviation. The mean shifts in $x_\pm^*$, $y_\pm^*$ are of order 0.01.

Uncertainty due to the fitter bias

As described in Sect. 6.4.2, the central values of the fit parameters $x_\pm^*$ and $y_\pm^*$ are corrected by a fitter bias that is determined with pseudoexperiments. The systematic uncertainty due to this is assigned using half of the size of the bias correction.

Summary of systematic uncertainties

The systematic uncertainties assessed are summarised in Table 6.8. The total experimental systematic uncertainties are calculated by adding all sources in quadrature and are 0.02 on x_+^* , 0.04 on x_-^* , 0.06 on y_+^* , and 0.05 on y_-^* . These uncertainties are dominated

Table 6.8. The statistical and the various sources of systematic uncertainties for $x_{\pm}^*, y_{\pm}^*$.

Source	$\sigma(x_+^*)$	$\sigma(x_-^*)$	$\sigma(y_+^*)$	$\sigma(y_-^*)$
Statistical	0.345	0.198	0.277	0.214
Efficiency corrections	0.019	0.034	0.021	0.005
Efficiency combination	0.007	0.001	0.007	0.008
Mass fit: α	0.002	0.005	0.021	0.020
$B_s^0 \rightarrow D^{*0} \bar{K}^{*0}$	0.002	0.002	0.010	0.005
$B^0 \rightarrow D \rho^0$	0.002	0.003	0.004	0.001
$B \rightarrow DK$	0.000	0.000	0.000	0.000
Signal shape	0.005	0.003	0.003	0.002
$B^0 \rightarrow D^{*0} K^{*0}$	0.006	0.007	0.008	0.004
$B \rightarrow D^{*0} h$	0.001	0.001	0.007	0.005
$B \rightarrow D \pi \pi \pi$	0.001	0.002	0.001	0.003
Dalitz plot migration	0.003	0.004	0.007	0.003
Value of κ	0.001	0.011	0.008	0.002
Fitter bias	0.004	0.014	0.042	0.042
Total systematic	0.022	0.040	0.056	0.048

by the efficiency corrections in F_i and the fitter bias. The systematic uncertainties are approximately 20% of the corresponding statistical uncertainties.

6.5 Summary of results

Approximately 84 $B^0 \rightarrow DK^{*0}$ decays with the final state $D \rightarrow K_S^0 \pi^+ \pi^-$ and 7 decays with $D \rightarrow K_S^0 K^+ K^-$ are used to determine the CP violation observables $x_{\pm}^*, y_{\pm}^*$. The observables are measured to be:

$$\begin{aligned}
 x_+^* &= 0.05 \pm 0.35 \pm 0.02, & y_+^* &= -0.81 \pm 0.28 \pm 0.06, \\
 x_-^* &= -0.31 \pm 0.20 \pm 0.04, & y_-^* &= 0.31 \pm 0.21 \pm 0.05,
 \end{aligned}$$

Table 6.9. Correlation matrix between the $x_{\pm}^*, y_{\pm}^*$ parameters for (left) statistical uncertainties only and (right) sum in quadrature of statistical and systematic uncertainties.

Statistical					Total				
	x_+^*	x_-^*	y_+^*	y_-^*		x_+^*	x_-^*	y_+^*	y_-^*
x_+^*	1.00	0.00	0.13	-0.01	x_+^*	1.00	0.00	0.13	-0.01
x_-^*		1.00	-0.01	0.15	x_-^*		1.00	-0.01	0.14
y_+^*			1.00	0.02	y_+^*			1.00	0.02
y_-^*				1.00	y_-^*				1.00

where the first uncertainties are statistical and the second are systematic. The correlation matrices are shown in Table 6.9. The systematic uncertainties are small in comparison to the statistical uncertainties and are only weakly correlated, so the systematic correlations are ignored. The extraction of r_{B^0} , δ_{B^0} and γ from these values is discussed in Sect. 7.2.

CHAPTER 7

Extraction of the CKM angle γ

The CP violation observables in $B \rightarrow DK$ decays, $x_{\pm}$ and $y_{\pm}$, and in $B^0 \rightarrow DK^{*0}$ decays, $x_{\pm}^*$ and $y_{\pm}^*$, have been measured in the $D \rightarrow K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ final states using a model-independent (MI) method. In this chapter, the CKM angle γ and the hadronic parameters r_B , δ_B , r_{B^0} and δ_{B^0} are determined from the observables using a frequentist method to extract their central values and confidence intervals. The results for $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ decays are given and are compared to measurements made using a model-dependent (MD) method. A combination of the two results and their combination with other LHCb γ measurements are shown.

7.1 Measurement of γ with $B \rightarrow DK$ decays

The CP observables $x_{\pm}, y_{\pm}$ are measured in $B \rightarrow DK$ decays to be

$$\begin{aligned} x_+ &= (-7.7 \pm 2.4 \pm 1.0 \pm 0.4) \times 10^{-2}, & y_+ &= (-2.2 \pm 2.5 \pm 0.4 \pm 1.0) \times 10^{-2}, \\ x_- &= (2.5 \pm 2.5 \pm 1.0 \pm 0.5) \times 10^{-2}, & y_- &= (7.5 \pm 2.9 \pm 0.5 \pm 1.4) \times 10^{-2}, \end{aligned}$$

where the first uncertainty is statistical, the second is systematic and the third comes from the uncertainty on the CLEO measurements.

7.1.1 Extraction method and results

The central values on r_B , δ_B and γ are extracted from the measured values of $x_{\pm}, y_{\pm}$ in $B \rightarrow DK$ decays. Equations (1.38) and (1.39) demonstrate that we have three unknowns (r_B, δ_B, γ) to determine with 4 observables ($x_{\pm}$ and $y_{\pm}$). Therefore the system is over-constrained and a multivariate Gaussian probability distribution function (PDF) is constructed to determine the most probable values of the underlying physics parameters. The PDF $\mathcal{P}(\vec{z}|\vec{z}')$ gives the probability of observing $x_{\pm}, y_{\pm}$ given *true* values $x'_{\pm}, y'_{\pm}$ and is defined:

$$\mathcal{P}(\vec{z}|\vec{z}') = \frac{1}{\sqrt{(2\pi)^4 |\mathbf{M}|}} \exp \left[-\frac{1}{2} (\vec{z} - \vec{z}')^{\top} \mathbf{M}^{-1} (\vec{z} - \vec{z}') \right], \quad (7.1)$$

where $\vec{z} = (x_+, x_-, y_+, y_-)$ is a vector of the measured values of $x_{\pm}, y_{\pm}$ and $\vec{z}' = (x'_+, x'_-, y'_+, y'_-)$ is a vector constructed from particular values of r_B, δ_B and γ using Eqs. (1.38) and (1.39). The matrix $\mathbf{M}$ is the covariance matrix between the four parameters $x_{\pm}$ and $y_{\pm}$. Additionally, $\vec{z}'$ is constructed such that $r_B = \sqrt{x_+'^2 + y_+'^2} = \sqrt{x_-'^2 + y_-'^2}$. Many values of r_B, δ_B and γ are used to calculate $\vec{z}'$ and hence $\mathcal{P}(\vec{z}|\vec{z}')$ using Eq. (7.1). The values of r_B, δ_B and γ which maximise $\mathcal{P}(\vec{z}|\vec{z}')$ for the measured $\vec{z}$ form the vector $\vec{\mu}_{\text{best}} = (r_B = 0.080, \delta_B = 134^\circ, \gamma = 62^\circ)$ and are assigned as the central values.

Neyman's construction [94] is employed to determine the confidence intervals on $\vec{\mu}_{\text{best}}$. The method is identical to that used by the Belle collaboration in the first GGSZ MI measurement of γ using $B \rightarrow DK, D \rightarrow K_S^0 \pi^+ \pi^-$ decays [40] and has been used in all subsequent GGSZ MI measurements. The confidence level α is defined for a given $\vec{\mu} = (r_B, \delta_B, \gamma)$ as

$$\alpha(\vec{\mu}) = \frac{\int_{\mathcal{D}(\vec{\mu})} d\vec{z} p(\vec{z}|\vec{\mu})}{\int_{\infty} d\vec{z} p(\vec{z}|\vec{\mu})}, \quad (7.2)$$

where $p(\vec{z}|\vec{\mu})$ is the PDF of obtaining the measured $\vec{z}$ given underlying physics parameters $\vec{\mu}$ and is defined to be identical to $\mathcal{P}(\vec{z}|\vec{z}')$ when $\vec{z}'$ is calculated from $\vec{\mu}_{\text{best}}$. The domain $\mathcal{D}(\vec{\mu})$ is the confidence interval constructed according to a Feldman-Cousins ordering [95], and $\int_{\infty} d\vec{z}$ integrates over the whole $x_{\pm}, y_{\pm}$ space. A likelihood-ratio, $\mathcal{R}$, is constructed by

$$\mathcal{R}(\vec{z}|\vec{\mu}) = \frac{p(\vec{z}|\vec{\mu})}{p(\vec{z}|\vec{\mu}_{\text{best}}(\vec{z}))}, \quad (7.3)$$

where $\vec{\mu}_{\text{best}}(\vec{z})$ is the $\vec{\mu}$ that maximises $p(\vec{z}|\vec{\mu})$ for the given $\vec{z}$. For a given $\vec{\mu}$, the value of $\mathcal{R}$ is calculated for many values of $\vec{z}$. The domain $\mathcal{D}(\vec{\mu})$ is then given by the condition $\mathcal{R}(\vec{z}|\vec{\mu}) > \mathcal{R}(\vec{z}_0|\vec{\mu})$, where $\vec{z}_0$ is the vector of $x_{\pm}, y_{\pm}$ measured in data. The confidence intervals for one, two and three standard deviations in $\vec{\mu}$ which enclose the observed $\vec{z}$ are determined at CL= 20%, 74% and 95%, respectively, corresponding to 1, 2 and 3 standard deviation volumes (*i.e.* $\Delta\chi^2 = 1, 4, 9$) in a three-dimensional Gaussian distribution¹.

The uncertainties on γ , δ_B and r_B are evaluated in two scenarios. First, only the statistical uncertainties and their correlations are considered. Second, the total uncertainties are considered, where the various sources of uncertainties are added in quadrature. The correlations between the parameters x_+ , x_- , y_+ and y_- are given in Table 5.5. Figure 7.1 shows the three-dimensional CL volume projected onto the (γ, r_B) and (γ, δ_B) planes. The contours in Fig. 7.1 are shown for the statistical uncertainties only, and when the combined statistical and systematic uncertainties are accounted for. The contours shown correspond to one, two and three standard deviations when projected into one dimension. For each of γ , r_B and δ_B , the uncertainties corresponding to one standard deviation are assigned. The solution for γ and δ_B has a two-fold ambiguity corresponding to the transformation $(\gamma, \delta_B) \rightarrow (\gamma + 180^\circ, \delta_B + 180^\circ)$. Quoting the results which correspond to the solution that satisfies $0 < \gamma < 180^\circ$ and accounting for

¹Calculated using `1-TMath::Prob(chi2={1,4,9},ndf=3)` in ROOT.

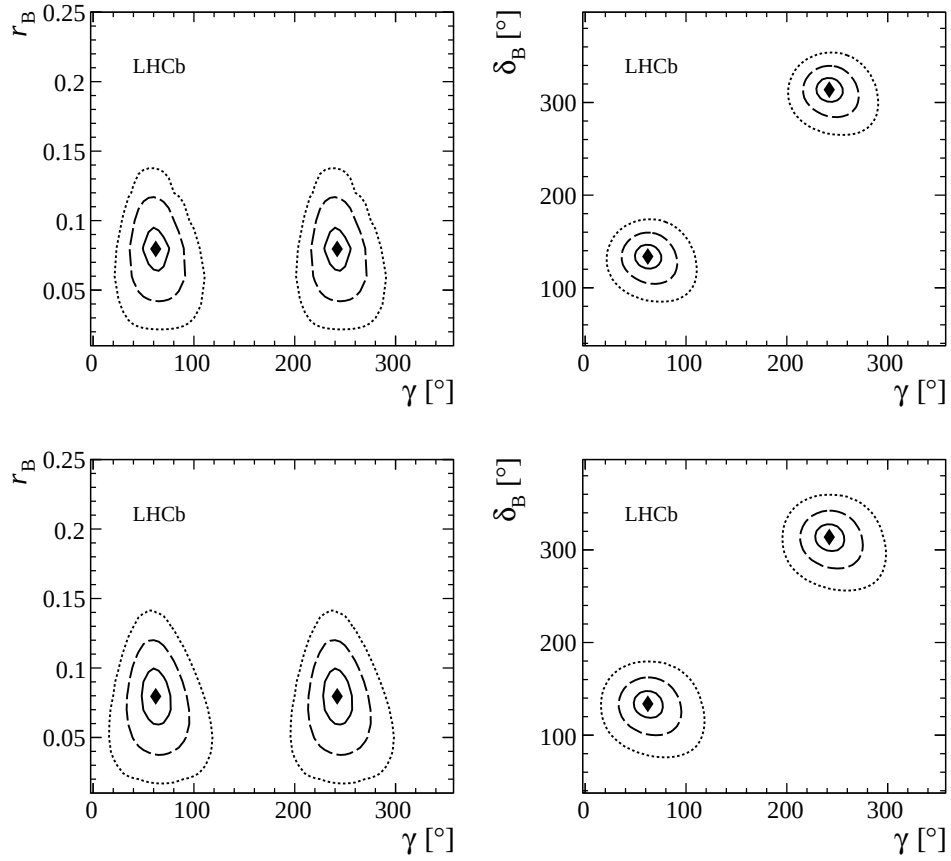


Figure 7.1. Confidence levels measured in $B \rightarrow DK$ decays projected onto the (left) (γ, r_B) and (right) (γ, δ_B) planes, accounting for (top) statistical uncertainties only and (bottom) all sources of uncertainty. The diamonds indicate the central values and the contours indicate the 1σ , 2σ and 3σ boundaries, corresponding to CL=20%, 74% and 95%, respectively.

statistical uncertainties only, the physics parameters are

$$\gamma = \left(62^{+14}_{-13}\right)^{\circ}, \quad r_B = 0.080^{+0.017}_{-0.018}, \quad \delta_B = \left(134^{+13}_{-14}\right)^{\circ}.$$

When the statistical and systematic uncertainties are considered, the parameters become

$$\gamma = \left(62^{+15}_{-14}\right)^{\circ}, \quad r_B = 0.080^{+0.019}_{-0.021}, \quad \delta_B = \left(134^{+14}_{-15}\right)^{\circ}.$$

7.1.2 Comparison to previous model-independent measurements

The results improve upon and supersede the MI measurement performed by the LHCb collaboration with a 1 fb^{-1} subset of data collected at $\sqrt{s} = 7 \text{ TeV}$ [41]. Approximately 650 signal candidates in the $K_s^0 \pi^+ \pi^-$ final state and 100 candidates in the $K_s^0 K^+ K^-$ final state were selected at $\sqrt{s} = 7 \text{ TeV}$. The CP observables measured with this dataset were $x_+ = (-10.3 \pm 4.5 \pm 1.8 \pm 1.4) \times 10^{-2}$, $y_+ = (-0.9 \pm 3.7 \pm 0.8 \pm 3.0) \times 10^{-2}$, $x_- = (0.0 \pm 4.3 \pm 1.5 \pm 0.6) \times 10^{-2}$, $y_- = (2.7 \pm 5.2 \pm 0.8 \pm 2.3) \times 10^{-2}$. The likelihood contours for the uncertainties on $x_{\pm}, y_{\pm}$ are compared with the contours for the analysis presented in this thesis in Fig. 7.2. The new measurement improves the statistical uncertainties on $x_{\pm}, y_{\pm}$ by the factor of $\sqrt{3}$ expected from the increase in integrated luminosity and signal candidates. A 30–40% reduction in the experimental systematic uncertainties is observed and the reasons that the dominant uncertainties have decreased are discussed below.

The largest systematic uncertainty in the 7 TeV analysis comes from the use of $B \rightarrow D\pi$ decays instead of $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu}$ to determine F_i . $B \rightarrow D\pi$ decays are not truly flavour-tagged and display small interference effects. The value of $r_B(B \rightarrow D\pi) = \mathcal{O}(1\%)$ is not well known and a conservative systematic uncertainty associated with CP violation in $B \rightarrow D\pi$ was assigned. In this thesis, this systematic uncertainty is replaced by a smaller uncertainty associated with efficiency corrections for the semileptonic mode. The second largest systematic uncertainty in the 7 TeV analysis comes from the choice of fit functions in the B mass parameterisation. This has been reduced through the use of data-driven methods to obtain their mass shapes and a candidate selection that rejects 30–40% more combinatorial background in the signal region.

The systematic uncertainty due to the CLEO inputs has decreased by 40–70%, even though the same binnings and c_i, s_i values were used in both analyses. This can be understood by considering the expression for the quantity $\frac{dx_+}{dc_i}$. The expected number of events per bin is given by $\bar{N}_i = h_+ \left[F_{+i} + (x_+^2 + y_+^2) F_{-i} + 2\sqrt{F_{+i} F_{-i}} (x_+ c_i - y_+ s_i) \right]$, where h_+ is

the normalisation term. The likelihood $\mathcal{L}$ between the observed yields N_i and expected yields $\bar{N}_i$ in each bin can be written as

$$\mathcal{L}(x_+, y_+) = \sum_i \frac{(N_i - \bar{N}_i)^2}{2N_i}. \quad (7.4)$$

The derivative $\frac{dx_+}{dc_i}$ can be related to the likelihood by

$$\frac{dx_+}{dc_i} \propto \frac{\partial^2 \mathcal{L}}{\partial x_+ \partial c_i} \left[\frac{\partial^2 \mathcal{L}}{\partial^2 x_+} \right]^{-1} \propto h_+ \sigma^2(x_+) \frac{N_i - \bar{N}_i}{N_i} + \mathcal{O}(x_+, y_+), \quad (7.5)$$

where $\sigma(x_+)$ is the statistical uncertainty on x_+ . The parameter h_+ is of order N and $\sigma(x_+)$ is of order $1/\sqrt{N}$, hence the N -dependence cancels in these two terms. The dependence of the systematic uncertainty of x_+ on the size of the B data sample arises from the term $\frac{(N_i - \bar{N}_i)}{N_i}$. This indicates the size of statistical fluctuations, which are expected to decrease for larger datasets, and in the high statistics limit, $\frac{(N_i - \bar{N}_i)}{N_i}$ will tend to a constant arising from differences between the CLEO and true values of c_i and s_i . Pseudoexperiments performed by CLEO indicate that the limiting systematic uncertainty on γ is 4° for the $K_S^0 \pi^+ \pi^-$ optimal binning [3].

The extracted physics parameters from the previous measurement are $\gamma = (45_{-38}^{+43})^\circ$, $r_B = 0.07 \pm 0.04$, $\delta_B = (137_{-46}^{+35})^\circ$. The precision on γ has improved by nearly a factor of three. This is driven primarily by the increase in the size of the dataset and improvements in systematic uncertainties. It is also driven by the slightly higher central value of $r_B = 0.08$. The reason is geometrical in nature and is illustrated in Fig. 7.3, which shows that two x, y measurements with the same uncertainties can provide a different γ precision depending on the value of r_B .

The Belle MI measurement [40] with 1200 $B \rightarrow DK$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decays obtained $x_+ = (-11.0 \pm 4.3 \pm 1.4 \pm 1.6) \times 10^{-2}$, $y_+ = (-5.0_{-5.5}^{+5.2} \pm 1.1 \pm 2.1) \times 10^{-2}$, $x_- = (9.5 \pm 4.5 \pm 1.4 \pm 1.7) \times 10^{-2}$, $y_- = (13.7_{-5.7}^{+5.3} \pm 1.9 \pm 2.9) \times 10^{-2}$, where the first uncertainty is statistical,

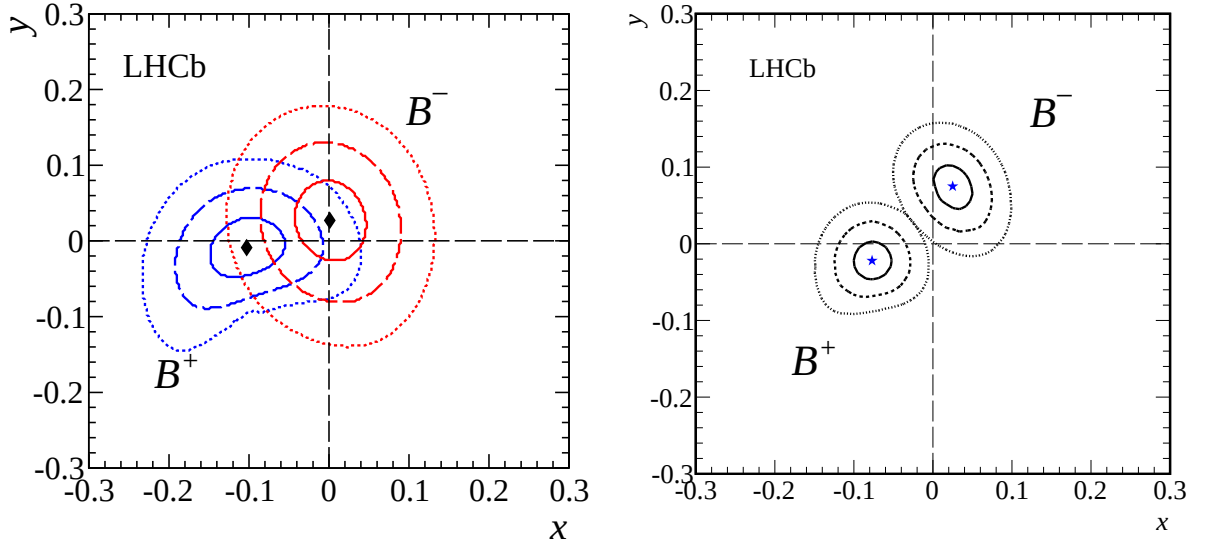


Figure 7.2. The one, two and three standard deviation contours for the model-independent measurements of $x_{\pm}, y_{\pm}$ (left) using 1 fb^{-1} of data collected at LHCb, reproduced from [41], and (right) presented in this thesis. The dotted lines at $x = 0$ and $y = 0$ are to guide the eye.

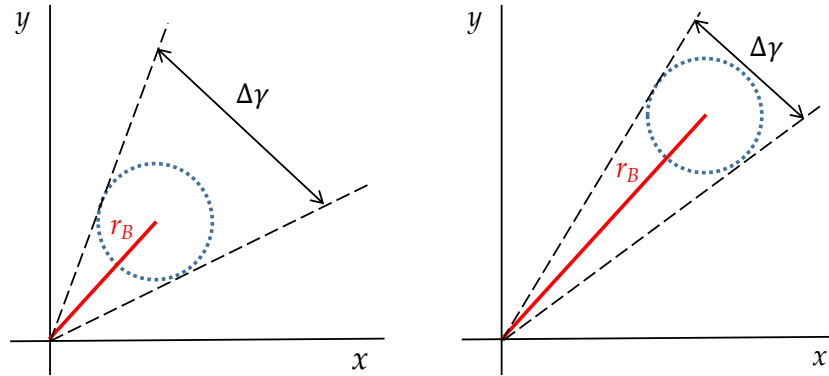


Figure 7.3. Illustration of the dependence of the precision on γ on the central value of r_B . The (short-dashed, blue) likelihood contours of two measurements of x and y of the same precision are drawn with (solid, red) the value of r_B and (long-dashed, black) the approximate uncertainty on γ , $\Delta\gamma$, superimposed. The (left) measurement with the smaller value of r_B provides a poorer precision on γ .

the second is systematic and the third is the CLEO systematic uncertainty (the optimal binning is also used). The value of γ is determined to be $(77 \pm 15 \pm 4 \pm 4)^\circ$, which is consistent with the measurement reported in this thesis. Despite the larger uncertainties on $x_{\pm}, y_{\pm}$, the precision on γ is similar to that obtained in this work because of the large central value of $r_B = 0.145 \pm 0.033$.

7.1.3 Comparison to model-dependent measurements

Model-dependent measurements of the parameters $x_{\pm}, y_{\pm}$ using $B \rightarrow DK$ decays have been performed at BaBar [34], Belle [38] and LHCb [44], where the δ_D variation (defined in Eq. 1.30) is provided using amplitude models of $D^0 \rightarrow K_s^0 h^+ h^-$ decays. The LHCb MD analysis uses a 1 fb^{-1} subset of data analysed in this thesis, collected at $\sqrt{s} = 7 \text{ TeV}$. The results of using the MI and MD methods are compared in this section.

At the B-factories, a combined analysis of three B^- decays is performed: $B \rightarrow DK$, $B \rightarrow D^*(\rightarrow \{D\pi^0, D\gamma\})K$, $B \rightarrow DK^*(\rightarrow K_s^0 \pi)$. The BaBar measurement uses both the $K_s^0 \pi^+ \pi^-$ and $K_s^0 K^+ K^-$ final states. The BaBar amplitude model is described in Ref. [33], in which $D^0 \rightarrow K_s^0 \pi^+ \pi^-$ and $D^0 \rightarrow K_s^0 K^+ K^-$ are each modelled by 8 resonant components. The Belle analysis uses $K_s^0 \pi^+ \pi^-$ only and a different amplitude model with 18 resonant components and an alternative parameterisation for the nonresonant contribution is used [38]. The LHCb measurement is performed with only $B^- \rightarrow DK^-$, $D \rightarrow K_s^0 \pi^+ \pi^-$ decays and the BaBar amplitude model is used.

The results are summarised with those of the MI measurements in Fig. 7.4; the ‘LHCb MI’ result is the one presented in this thesis. The statistical uncertainties from the MD measurements are larger due to the smaller datasets used. Of the three MD analyses, the BaBar dataset is the largest with 1050 ± 34 $B \rightarrow DK$ signal events, compared to over 2500 signal events in the LHCb MI dataset. In general, an MD measurement is expected to give smaller statistical uncertainties than an MI measurement performed with the same dataset, since the technique does not involve partitioning the data into bins of the Dalitz plot.

The model-related uncertainties in the MD analysis are evaluated by adding or removing resonances from the default amplitude models, by varying the model parameters or by using alternate parameterisations of particular resonances. When propagated to uncertainties on γ , the model uncertainties measured at BaBar and Belle are 3° and

9° , respectively. The substantial difference in the size of this systematic uncertainty is due to differing strategies taken in the construction of alternative models. This is because an amplitude model with an inaccurate strong phase distribution can still perfectly describe the D^0 Dalitz plot intensity distribution, and alternative amplitude models can produce a poorer fit χ^2 while not necessarily producing different strong phase behaviour on the Dalitz plot. The larger uncertainty estimated by Belle considers fits with alternate amplitude models where the strong phase distributions are different from the nominal, and this is reported to dominate the model uncertainty. These measurements demonstrate the difficulties in estimating the model uncertainty in a robust manner. By contrast, the uncertainties associated with the CLEO inputs in MI measurements are well-defined.

The differences between the BaBar model prediction and the measured values of c_i and s_i were shown in Fig. 1.10. The predictions and measurements show a $1-2\sigma$ difference in a few bins. The use of different δ_D distributions could result in a shift between the central values of $x_\pm, y_\pm$. Figure 7.4 shows that there is no significant systematic shift between the central values obtained in the MI and MD measurements, or between MD measurements using the Belle or BaBar model. However the uncertainties are heavily dominated by the statistical uncertainties, which are much larger than the differences in central values.

7.2 Measurement of γ with $B^0 \rightarrow DK^{*0}$ decays

The same procedure as that described in Sect. 7.1.1 is used to measure the parameters γ , r_{B^0} and δ_{B^0} , which are related to $x_\pm^*, y_\pm^*$ using Eqs. (1.57) and (1.58). The CP violation

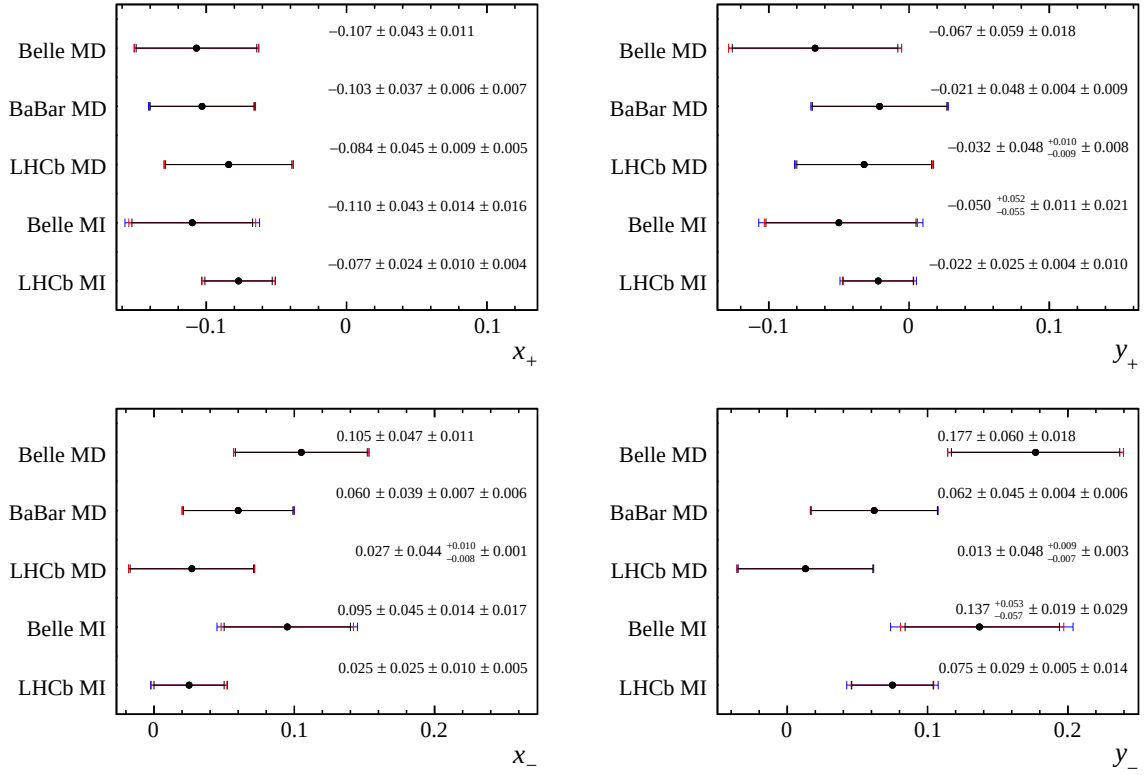


Figure 7.4. Summary of measurements of $x_{\pm}$ and $y_{\pm}$ using the model-independent (MI) and model-dependent (MD) variants of the GGSZ method by Belle, BaBar and LHCb. The first (black) uncertainty is statistical and the second (red) is systematic. The third (blue) uncertainty is due to the systematic from the CLEO inputs for the MI method and from the amplitude model for the MD method. In the case of Belle MD, the model uncertainty was not reported explicitly.

observables, $x_{\pm}^*$ and $y_{\pm}^*$, are measured in $B^0 \rightarrow DK^{*0}$ decays to be

$$\begin{aligned} x_+^* &= 0.05 \pm 0.35 \pm 0.02, & y_+^* &= -0.81 \pm 0.28 \pm 0.06, \\ x_-^* &= -0.31 \pm 0.20 \pm 0.04, & y_-^* &= 0.31 \pm 0.21 \pm 0.05, \end{aligned}$$

where the first uncertainty is statistical and the second systematic. The correlations between the uncertainties of x_+^* , x_-^* , y_+^* and y_-^* are listed in Table 6.9. As before, the confidence levels are evaluated using the statistical uncertainties and correlations only, and separately for the combined statistical and systematic uncertainties. The three-dimensional confidence levels are projected onto the (γ, r_{B^0}) and (γ, δ_{B^0}) planes in Fig. 7.5, and are projected onto one dimension for each of γ , r_{B^0} and δ_{B^0} in order to

determine their uncertainties. For statistical uncertainties only, the results for the solution consistent with $0 < \gamma < 180^\circ$ are

$$\gamma = (71 \pm 20)^\circ, \quad r_{B^0} = 0.56^{+0.16}_{-0.17}, \quad \delta_{B^0} = (204^{+21}_{-19})^\circ.$$

For the combined statistical and systematic uncertainties, the results are

$$\gamma = (71 \pm 20)^\circ, \quad r_{B^0} = 0.56 \pm 0.17, \quad \delta_{B^0} = (204^{+21}_{-20})^\circ.$$

The measured value of $r_{B^0} = 0.56 \pm 0.17$ is driven by the upward fluctuation in y_+ , and in turn this large value of r_{B^0} leads to a relatively small error on γ . The expected value of $r_{B^0} \sim |V_{ub}^* V_{cs}|/|V_{cb}^* V_{us}| \sim 0.3$ is consistent with this measurement at 2σ .

7.2.1 Comparison to previous measurements

The Belle collaboration has also performed a MI GGSZ measurement with $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decays [43] with 44^{+13}_{-12} signal events. The results obtained are $x_+^* = 0.1^{+0.7+0.0}_{-0.4-0.1} \pm 0.1$, $x_-^* = 0.4^{+1.0+0.0}_{-0.6-0.1} \pm 0.0$, $y_+^* = 0.3^{+0.5+0.0}_{-0.8-0.1} \pm 0.1$, $y_-^* = -0.6^{+0.8+0.1}_{-1.0-0.0} \pm 0.1$, where the first uncertainty is statistical, the second systematic and the third is the uncertainty associated with the CLEO inputs. The results are consistent with the measurements of this thesis, however the total uncertainties are two to four times larger in the Belle measurement. This is much larger than the expectation from the factor of two difference in number of signal events and is caused by the $K_S^0 \pi^+ \pi^-$ events being divided into 32 bins (8×2 bins per Dalitz plot, one Dalitz plot per B charge). With a small dataset the number of signal events in some bins is consistent with zero, which reduces the sensitivity to $x_\pm, y_\pm$. No limits on γ or δ_{B^0} are placed by Belle, and the parameter r_{B^0} is constrained to the region $r_{B^0} < 0.87$ at 68% CL.

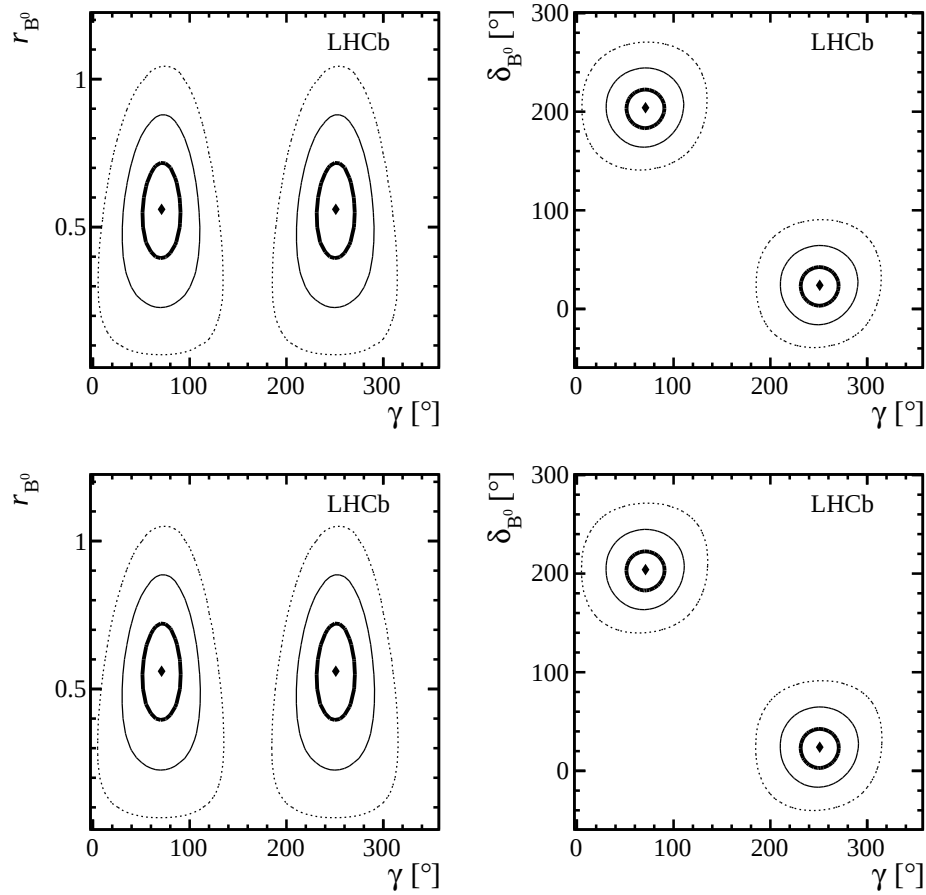


Figure 7.5. Confidence levels measured in $B^0 \rightarrow DK^{*0}$ decays projected onto the (left) (γ, r_{B^0}) and (right) (γ, δ_{B^0}) planes, accounting for (top) statistical uncertainties only and (bottom) all sources of uncertainty. The diamonds indicate the central values and the contours indicate the 1σ , 2σ and 3σ boundaries.

An MD measurement of $x_{\pm}^*, y_{\pm}^*$ has been performed by LHCb [45], where the $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 \pi^+ \pi^-$ dataset used is identical to the one described in this thesis. The BaBar 2010 amplitude model is used. The results are $x_+^* = 0.05_{-0.24}^{+0.24} \pm 0.04 \pm 0.01$, $y_+^* = -0.65_{-0.23}^{+0.24} \pm 0.08 \pm 0.01$, $x_-^* = -0.15_{-0.14}^{+0.14} \pm 0.03 \pm 0.01$, $y_-^* = 0.25_{-0.15}^{+0.15} \pm 0.06 \pm 0.01$, where the uncertainties are statistical, systematic and model-related. The statistical uncertainties are 17–45% smaller than for the MI measurement. This results from the use of the modified optimal binning instead of the optimal binning for $K_S^0 \pi^+ \pi^-$, which was necessary to mediate fit instabilities because of the low signal yield. With a larger data set, it is expected that the difference in statistical uncertainties between the MD and MI method will reduce to the predicted 10–15% loss (Sect. 1.6.3).

The BaBar MD measurement of $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decays is performed with only 39 ± 9 signal events, with a relatively large background. In addition to a different choice of K^{*0} region, no independent limits on the physics parameters are placed, and hence the result is not discussed further [42].

7.3 Combination of γ measurements

The current official LHCb combination of many γ measurements from $B \rightarrow DK$ -like decays is $\gamma = \left(70.9_{-8.5}^{+7.1}\right)^\circ$ [96]. This combination includes the $B \rightarrow DK$ GGSZ MI measurement. The $B^0 \rightarrow DK^{*0}$ GGSZ MD measurement [45] is included, but the numerical results of the combination do not change when the MI results are used instead. A frequentist treatment is used, although it has been demonstrated that a Bayesian combination produces compatible results.

The confidence intervals in this section are calculated using an independent implementation with respect to the official combination in Ref. [96]. The two combinations employ different approximations of the full frequentist treatment. Ref. [96] uses a method based on pseudoexperiments, which provides a more accurate approximation with better coverage. The approach in this section uses a profile likelihood method. It is much less computationally intensive and the accuracy provided is sufficient to demonstrate the impact of combining various γ measurements.

The effect of combining the two GGSZ MI measurements presented in this thesis is first considered. The results of combining the other LHCb γ measurements, excluding the GGSZ measurements, are then presented. Finally, the impact of adding the two GGSZ MI measurements to all other γ measurements is discussed.

7.3.1 Combination of GGSZ MI measurements

A combined value of γ is extracted from the CP violation observables measured with $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ decays described in this thesis. Correlations between the uncertainties on $x_{\pm}, y_{\pm}$ and $x_{\pm}^*, y_{\pm}^*$ can be expected, because the values of c_i and s_i in the two analyses were obtained using the same CLEO dataset. The uncertainties due to the CLEO inputs arise from two sources: the measured uncertainties from the CLEO measurements, and the size of the LHCb B data sample; only the former leads to correlations between different measurements. However, since two different $D \rightarrow K_S^0 \pi^+ \pi^-$ binning schemes are used, the resulting correlations between the CP observables cannot be quantified exactly without access to the CLEO data. An upper limit for the correlations is determined by generating pseudoexperiments assuming that the same binnings had been used for $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ decays. Large signal yields are generated so that the component of the uncertainty related to the B yields is negligible. Both the modes are generated with the same smeared sets of c_i and s_i . It is found that the correlations have a negligible effect on the combined results. Hence, the correlations between the uncertainties of the two measurements are neglected. A global fit incorporating the observables $x_{\pm}, y_{\pm}, x_{\pm}^*$ and $y_{\pm}^*$ is performed. The best fit value and the 1σ confidence interval of γ is

$$\gamma = (65 \pm 12)^\circ,$$

and the hadronic parameters are $r_B = 0.0790 \pm 0.020$, $\delta_B = (134_{-15}^{+14})^\circ$, $r_{B^0} = 0.55 \pm 0.17$, and $\delta_{B^0} = (204_{-20}^{+22})^\circ$. The resulting confidence levels in γ and two-dimensional projections of γ with each of the four hadronic parameters are shown in Fig. 7.6.

7.3.2 Combination with other LHCb γ measurements

Many other γ measurements have been performed using $B \rightarrow DK$ -like decays at LHCb. The decay modes and methods used are summarised in Table 7.1. The GLW and ADS methods have been outlined in Sect. 1.5. Two variants of the GLW method are used. The ‘qGLW’ method refers to the use of final states with only fractional CP -even content, such as $D \rightarrow \pi\pi\pi\pi$ or $\pi\pi\pi^0$. The ‘GLW-Dalitz’ method uses the two-body GLW modes, but an analysis of the $B^0 \rightarrow DK^+\pi^-$ and $\bar{B}^0 \rightarrow DK^-\pi^+$ Dalitz plots is used to identify the resonances contributing to the $b \rightarrow u$ or $b \rightarrow c$ amplitudes. In $B_s^0 \rightarrow D_s^+ K^-$ decays, access to γ is provided through the interference between direct $B_s^0 \rightarrow D_s^+ K^-$ and mixed $B_s^0 \rightarrow \bar{B}_s^0 \rightarrow D_s^+ K^-$ decays. Hence, γ is measured by studying the CP asymmetries of B_s^0 and $\bar{B}_s^0$ decays as a function of decay time. The result of combining all the LHCb γ measurements, excluding the two results presented in this thesis, is shown in Fig. 7.7. The combination also includes additional inputs for parameters describing the charm decay systems, coherence factors, and constraints on mixing and direct CP violation in the charm system. The full list can be found in Ref. [96]. The combined value of γ is $(106_{-7}^{+8})^\circ$ with a further three values which are allowed at 2σ . The different solutions of γ correspond to different values of δ_B and r_B , as shown with multiple ‘lobes’ in the two-dimensional scans in Fig. 7.7. The combination is driven by the measurement of γ in $B \rightarrow Dh$, two- and four-body D decays [83], where the observables with sensitivity to γ contain trigonometric relations of the form $\sin\gamma \sin\delta_B$ and $\cos\gamma \cos\delta_B$, resulting in a four-fold ambiguity in the value of $\gamma \in [0^\circ, 180^\circ]$.

The effect of adding the $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ GGSZ MI measurements to the other LHCb γ measurements is shown in Fig. 7.8 and complementarity is clearly demonstrated. The ability to measure the Cartesian parameters directly using the $D \rightarrow K_s^0 h^+ h^-$ modes provide only one solution in the range $0 < \gamma < 180^\circ$ and hence are able to resolve the trigonometric ambiguities presented by the other methods.

Table 7.1. List of decay modes and techniques used in LHCb γ measurements, whose results are combined in Fig. 7.7. The particle denoted h refers to a charged kaon or pion. The methods are described in the main text.

B mode	D mode	Method	Ref.
$B \rightarrow DK$	$D \rightarrow hh$	GLW/ADS	[83]
	$D \rightarrow h\pi\pi\pi$	qGLW/ADS	[83]
	$D \rightarrow hh\pi^0$	qGLW/ADS	[97]
	$D \rightarrow K_S^0 K\pi$	ADS	[98]
$B \rightarrow DK^{*0}$	$D \rightarrow hh$	ADS	[90]
$B \rightarrow DK\pi$	$D \rightarrow hh$	GLW-Dalitz	[92]
$B \rightarrow DK\pi\pi$	$D \rightarrow hh$	GLW/ADS	[99]
$B_S^0 \rightarrow D_S^\pm K^\mp$	$D_S^+ \rightarrow hhh$	Time-dependent	[100]

In turn, this places additional constraints in the parameter space for the hadronic parameters. It is observed that the addition of the GGSZ measurements causes the combination to prefer a different solution of γ and the best fit values of the hadronic parameters have shifted correspondingly. It is of note that the value of $r_{B^0} = 0.25$ from the GLW/ADS modes decreases slightly to 0.22 when the GGSZ modes are added to the combination, despite the high value of $r_{B^0} = 0.55$ measured in the $B^0 \rightarrow DK^{*0}$ GGSZ measurement. This is because the favoured solution of γ shifts from 106° to 70° . The result of this combination is $\gamma = (70.3_{-8.4}^{+7.0})^\circ$. The hadronic parameters are $r_B = 0.100 \pm 0.006$, $\delta_B = (140.7_{-7.7}^{+6.1})^\circ$, $r_{B^0} = 0.223_{-0.049}^{+0.042}$ and $\delta_{B^0} = (192_{-21}^{+26})^\circ$. These values are in good agreement with the results of combining the γ measurements performed by the B-factories. In particular, the γ combination exceeds the precision of the BaBar and Belle γ combinations, which are $\gamma = (69_{-16}^{+17})^\circ$ and $\gamma = (68_{-14}^{+15})^\circ$, respectively [18, 85].

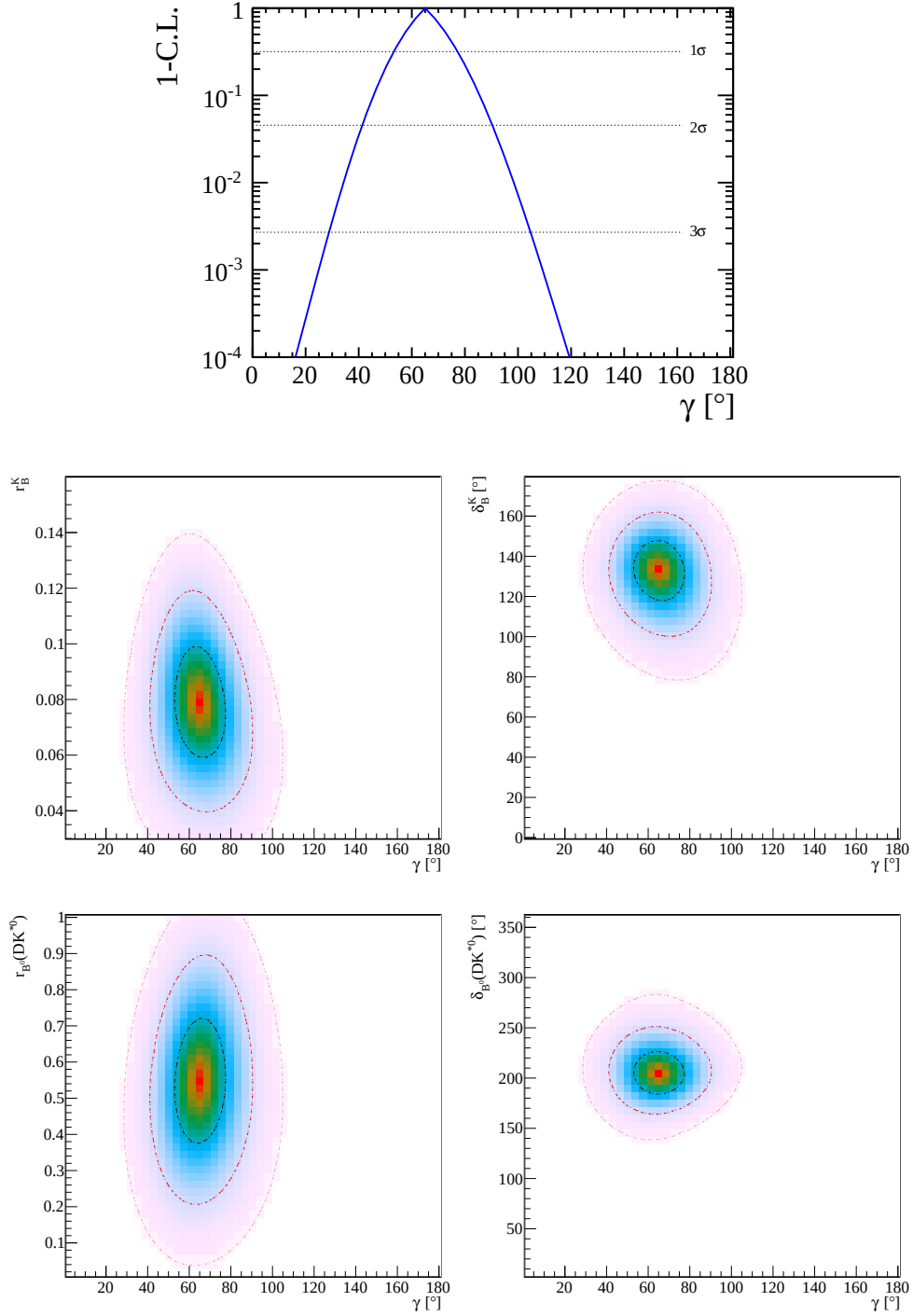


Figure 7.6. Combination of results from $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ decays with $D \rightarrow K_s^0 h^+ h^-$. Confidence levels on γ and two-dimensional projections on the (γ, r_B) , (γ, δ_B) , (γ, r_{B^0}) , (γ, δ_{B^0}) planes are shown. The dotted lines indicate the one, two and three standard deviation volumes.

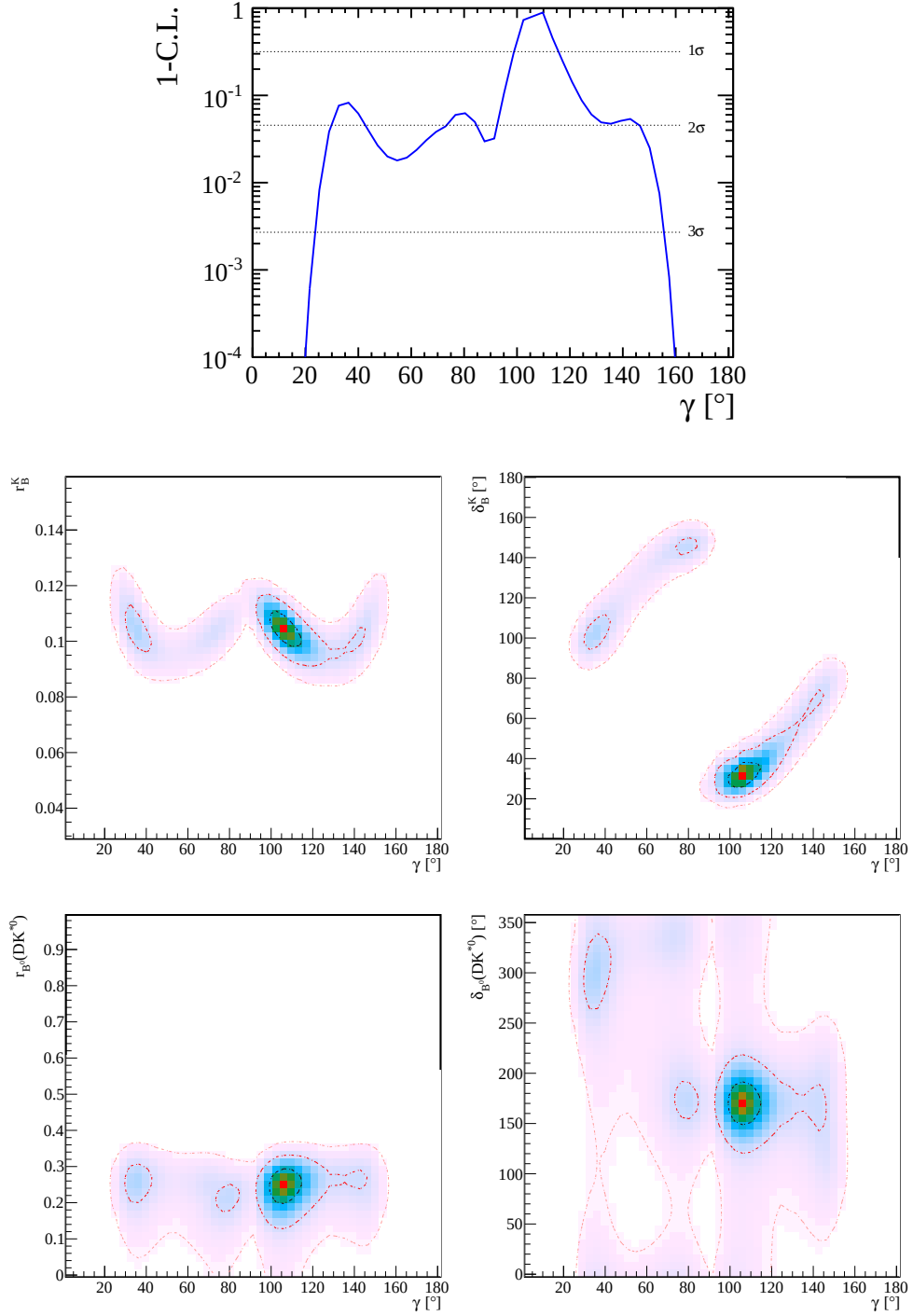


Figure 7.7. Combination of LHCb measurements listed in Table 7.1. Confidence levels on γ and two-dimensional projections on the (γ, r_B) , (γ, δ_B) , (γ, r_{B^0}) , (γ, δ_{B^0}) planes are shown. The dotted lines indicate the one, two and three standard deviation volumes.

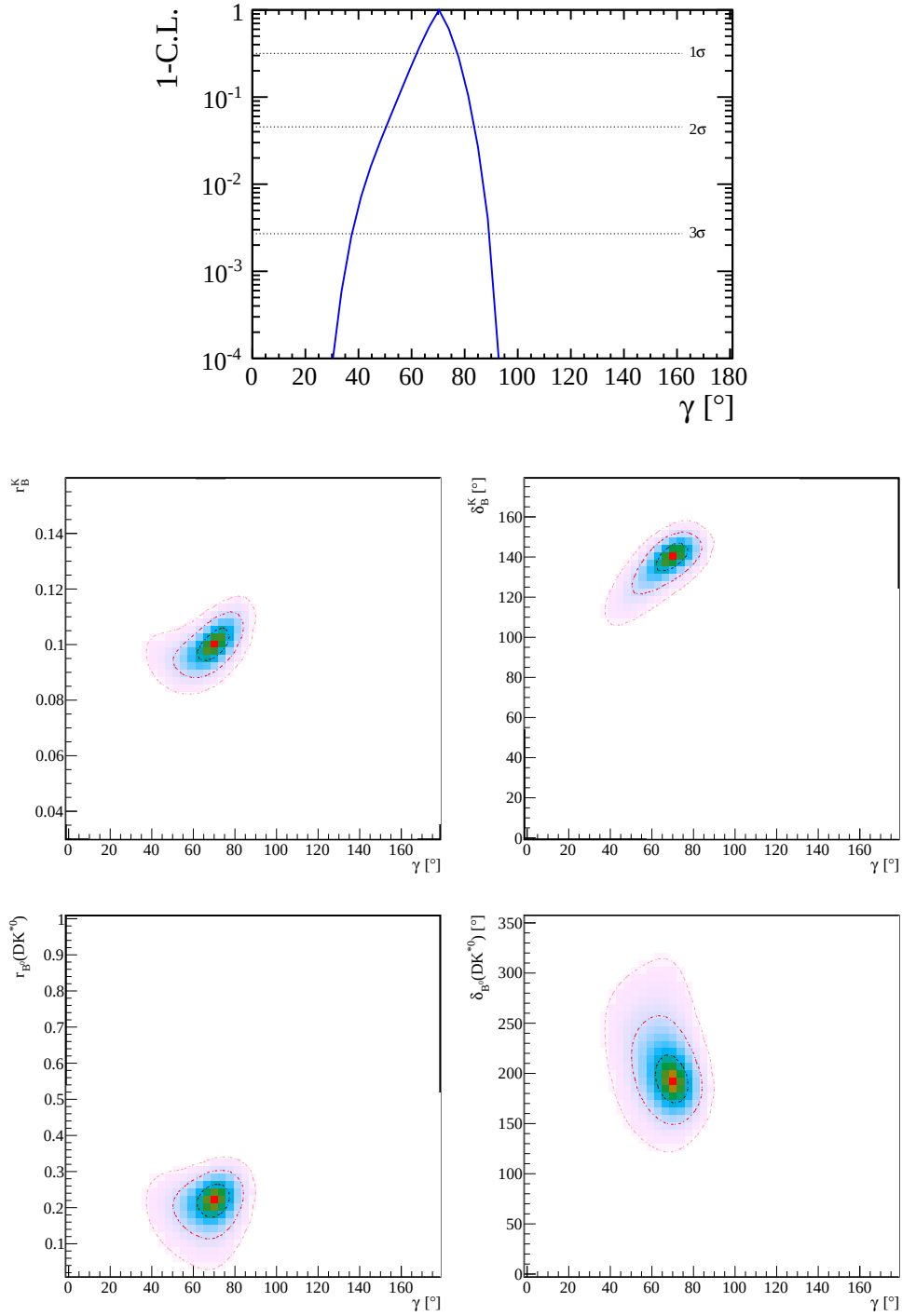


Figure 7.8. Combination of LHCb measurements listed in Table 7.1 and also the $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ GGSZ MI measurements presented in this thesis. Confidence levels on γ and two-dimensional projections on the (γ, r_B) , (γ, δ_B) , (γ, r_{B^0}) , (γ, δ_{B^0}) planes are shown. The dotted lines indicate the one, two and three standard deviation volumes.

CHAPTER 8

Conclusions

8.1 Summary of results

A study of CP violation in the decay modes $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ has been performed, where the D meson decays to the $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ final states. Proton-proton collision data collected at the LHCb detector corresponding to an integrated luminosity of 3 fb^{-1} at $\sqrt{s} = 7$ and 8 TeV were used to carry out a model-independent analysis of the $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$ Dalitz plots, where information about the D strong-phase difference distribution was taken from measurements made by the CLEO collaboration [3].

The world's most precise direct measurements of the CP violation observables in $B \rightarrow DK$ decays, $x_{\pm} = r_B \cos(\delta_B \pm \gamma)$ and $y_{\pm} = r_B \sin(\delta_B \pm \gamma)$, were measured using approximately 2580 signal events. The CP violation observables in $B^0 \rightarrow DK^{*0}$ decays, $x_{\pm}^* = r_{B^0} \cos(\delta_{B^0} \pm \gamma)$ and $y_{\pm}^* = r_{B^0} \sin(\delta_{B^0} \pm \gamma)$, were measured using approximately 90 signal events, and is the first time that the $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 K^+ K^-$ decay mode has been used in a measurement of γ . The numerical results of the two analyses are summarised in Table 8.1, showing that the measured values of γ are consistent with one another. When the two sets of CP violation observables are combined, a value of $\gamma = (65 \pm 12)^\circ$ is obtained. This single solution of γ in the range $[0^\circ, 180^\circ]$ currently

Table 8.1. Summary of measurements in this thesis using $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ decays, with $D \rightarrow K_s^0 \pi^+ \pi^-$ and $D \rightarrow K_s^0 K^+ K^-$ decays. The values of γ , δ_B and δ_{B^0} quoted correspond to the solution where $0^\circ < \gamma < 180^\circ$. The second solution is given by $\gamma \rightarrow \gamma + 180^\circ$.

$B \rightarrow DK$	$B^0 \rightarrow DK^{*0}$
$x_+ = (-7.7 \pm 2.4 \pm 1.0 \pm 0.4) \times 10^{-2}$	$x_+^* = 0.05 \pm 0.35 \pm 0.02$
$y_+ = (-2.2 \pm 2.5 \pm 0.4 \pm 1.0) \times 10^{-2}$	$y_+^* = -0.81 \pm 0.28 \pm 0.06$
$x_- = (2.5 \pm 2.5 \pm 1.0 \pm 0.5) \times 10^{-2}$	$x_-^* = -0.31 \pm 0.20 \pm 0.04$
$y_- = (7.5 \pm 2.9 \pm 0.5 \pm 1.4) \times 10^{-2}$	$y_-^* = 0.31 \pm 0.21 \pm 0.05$
$\gamma = (62_{-14}^{+15})^\circ$	$\gamma = (71 \pm 20)^\circ$
$r_B = 0.080_{-0.021}^{+0.019}$	$r_{B^0} = 0.56 \pm 0.17$
$\delta_B = (134_{-15}^{+14})^\circ$	$\delta_{B^0} = (204_{-20}^{+21})^\circ$

plays the crucial role of resolving the multiple solutions of γ obtained with the other D final states.

8.2 Outlook

Precision measurements of the CKM angle γ will continue to be among the physics priorities for the LHCb experiment and for flavour physics. This is because the direct determination of γ in tree-level processes serves as an excellent SM benchmark, which can be compared to the indirect determination of γ that is sensitive to NP processes. A naïve estimate of the current precision, obtained by averaging the BaBar, Belle and most recent LHCb combinations, is $(70 \pm 6)^\circ$. Assuming the unitarity condition of the UT, the world average indirect measurement of γ is $(66.85_{-3.44}^{+0.94})^\circ$ [18]. While these results are currently consistent, with reduced uncertainties, a difference if present, can be detected in the future.

All of the existing γ measurements are statistically limited, hence the constraints on γ should improve as LHCb not only collects more data, but exploits new B modes

and D final states. In Run II of the LHC, the LHCb detector collected 300 pb^{-1} of pp collision data in 2015 and has collected a similar amount so far in 2016, with a goal to collect 5 fb^{-1} by 2018. It has been observed in 2015 data that the signal yield per pb^{-1} of the relevant B decays increases by a factor of two compared to Run I data because of the increased cross-section for b production and improvements in the trigger efficiencies for multibody hadronic final states. The γ precision from a combination of LHCb measurements is expected to be 4° [101] when the Run II data have been analysed. A planned upgrade for the LHCb detector after Run II will enable the experiment to increase to higher luminosity conditions ($\sim 2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$) by using a fully software trigger to read out the entire detector at 40 MHz, the same rate as all visible pp interactions. This should allow 50 fb^{-1} of data to be collected by 2030, with signal events collected at an even higher efficiency. Using this dataset, the precision on γ is projected to reach 0.9° [101]. Competitive constraints on γ will also come from the Belle II experiment, which plans to collect e^+e^- collision data at the $\Upsilon(4S)$ resonance corresponding to an integrated luminosity of 50 ab^{-1} between 2018–2025. The expected precision on γ from Belle II is 1.5° [102].

Of the many methods proposed, the GGSZ method will continue to play an important role in constraining the value of γ as it provides direct access to the Cartesian parameters $x_{\pm}^{(*)}$ and $y_{\pm}^{(*)}$. Future updates of the $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ measurements with more data at LHCb, as well as measurements of the GGSZ modes in other B decays, will play a vital role in the experimental programme to increase the precision on γ and to determine its value. With 50 fb^{-1} of data, it is expected that the γ uncertainty from $B \rightarrow DK$ GGSZ modes will become systematically limited. In this scenario, the advantages offered by the model-independent method, which provides both direct access to the strong-phase differences and robustly defined systematic uncertainties, will far outweigh any loss in statistical power brought about by binning the Dalitz plot. In comparison, the difficulties faced in evaluating model-related uncertainties will persist

in the model-dependent method. Therefore, in the long run the model-independent strategy should be adopted.

The current measurements of the CP observables in $B \rightarrow DK^{(*)}$ decays are dominated by the statistical uncertainties. The major source of experimental systematic uncertainty in both the $B \rightarrow DK$ and $B^0 \rightarrow DK^{*0}$ analyses comes from the efficiency corrections applied to the semileptonic control channel. The magnitude of the efficiency corrections can be decreased by aligning the selection criteria of the signal and control channels, including those used at the central data processing stage. The systematic uncertainty associated with the simulation-derived correction can also be decreased with better understanding of the differences between data and simulation. Another large source of systematic uncertainty comes from the c_i and s_i inputs. This uncertainty should decrease with more data, and the limiting uncertainty due to the CLEO measurements is expected to be 4° when the optimal binning is used [3]. The analysis should also benefit from more precise measurements of c_i and s_i from the BESIII experiment, which has collected data at the $\psi(3770)$ resonance corresponding to an integrated luminosity of 2.9 fb^{-1} , or 3.5 times that of the CLEO dataset. The results of a preliminary measurement performed with $D \rightarrow K_s^0 \pi^+ \pi^-$ decays are consistent with the CLEO central values and with smaller statistical uncertainties. The limiting uncertainty on γ with input from BESIII data is anticipated to decrease to $1\text{--}2^\circ$ [103]. In the longer term, it will become necessary to take into account the effects of charm mixing and CP violation in the neutral kaon system, which are estimated to be of order 0.2° and 1° , respectively, in $B \rightarrow DK$ decays [22, 87]. These effects scale as $1/r_B$ and will therefore also eventually become relevant for $B^0 \rightarrow DK^{*0}$, where $r_{B^0} \sim 3r_B$.

The motivation for measuring γ also serves a wider goal of performing stringent tests of the Standard Model picture of CP violation in the quark sector in order to discover new sources of CP violation that can account for the observed matter-antimatter asymmetry. This will naturally require continued precision studies of all the parameters

participating in the CKM unitarity triangle and improvements in the theoretical (lattice QCD) inputs required to extract them. The current global fit of the CKM unitarity triangle parameters is consistent with the SM picture, but does not exclude the presence of sizeable NP contributions through flavour-changing neutral currents in the neutral B meson sector [104]. If new physics models or new particles are discovered during the LHC era, the CP -violating mechanisms and flavour structure of the models must be understood within the constraints imposed by the unitarity triangle measurements. Conversely, if new particles lie at a higher mass scale than observable by direct searches at the LHC energy of $\sqrt{s} = 14\text{TeV}$, precision measurements of the unitarity triangle parameters will be placed under the spotlight as the first conclusive evidence for NP could emerge here.

Bibliography

- [1] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, *Determining γ using $B^\pm \rightarrow DK^\pm$ with multibody D decays*, Phys. Rev. **D68** (2003) 054018.
- [2] A. Bondar and A. Poluektov, *Feasibility study of model-independent approach to ϕ_3 measurement using Dalitz plot analysis*, Eur. Phys. J. **C47** (2006) 347, arXiv:hep-ph/0510246; A. Bondar and A. Poluektov, *The use of quantum-correlated D^0 decays for ϕ_3 measurement*, Eur. Phys. J. **C55** (2008) 51, arXiv:0801.0840.
- [3] CLEO collaboration, J. Libby *et al.*, *Model-independent determination of the strong-phase difference between D^0 and $\bar{D}^0 \rightarrow K_{S,L}^0 h^+ h^-$ ($h = \pi, K$) and its impact on the measurement of the CKM angle γ/ϕ_3* , Phys. Rev. **D82** (2010) 112006.
- [4] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.
- [5] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [6] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [7] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.
- [8] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.
- [9] C. L. Bennett *et al.*, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, Astrophys. J. Suppl. **208** (2013) 20, arXiv:1212.5225.
- [10] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.
- [11] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the K_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.
- [12] BaBar collaboration, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson*

- system*, Phys. Rev. Lett. **87** (2001) 091801, arXiv:hep-ex/0107013.
- [13] Belle collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.
 - [14] E. Gardi, N. Glover, and A. Robson, eds., *LHC Phenomenology (Scottish Graduate Series)*, Springer, 2015 ed., 8, 2014.
 - [15] M. B. Gavela, M. Lozano, J. Orloff, and O. Pène, *Standard model CP-violation and baryon asymmetry (i). zero temperature*, Nuclear Physics B **430** (1994) 345, arXiv:9406288.
 - [16] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.
 - [17] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.
 - [18] J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, arXiv:hep-ph/0406184, updated results and plots available at: <http://ckmfitter.in2p3.fr>.
 - [19] UTfit collaboration, M. Bona *et al.*, *The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of Δm_s and $BR(B \rightarrow \tau \nu_\tau)$* , JHEP **10** (2006) 081, arXiv:hep-ph/0606167, updated results and plots available at: <http://www.utfit.org>.
 - [20] M. Blanke and A. J. Buras, *Universal Unitarity Triangle 2016 and the tension between $\Delta M_{s,d}$ and ε_K in CMFV models*, Eur. Phys. J. **C76** (2016), no. 4 197, arXiv:1602.04020.
 - [21] Fermilab Lattice, MILC, A. Bazavov *et al.*, *$B_{(s)}^0$ -mixing matrix elements from lattice QCD for the Standard Model and beyond*, Phys. Rev. **D93** (2016), no. 11 113016, arXiv:1602.03560.
 - [22] Y. Grossman and M. Savastio, *Effects of K^0 - $\bar{K}^0$ mixing on determining γ from $B^\pm \rightarrow DK^\pm$* , JHEP **03** (2014) 008, arXiv:1311.3575.
 - [23] J. Brod and J. Zupan, *The ultimate theoretical error on γ from $B \rightarrow DK$ decays*, JHEP **01** (2014) 051, arXiv:1308.5663.
 - [24] M. Gronau and J. L. Rosner, *Final state interaction effects on γ from $B \rightarrow DK$* , Phys. Lett. **B439** (1998) 171, arXiv:hep-ph/9807447.
 - [25] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of early 2012*, arXiv:1207.1158, updated results and plots available at <http://www.slac.stanford.edu/xorg/hfag/>.
 - [26] M. Gronau and D. London, *How to determine all the angles of the unitarity triangle from $B_d^0 \rightarrow DK_s^0$ and $B_s^0 \rightarrow D\phi$* , Phys. Lett. **B253** (1991) 483.

- [27] M. Gronau and D. Wyler, *On determining a weak phase from CP asymmetries in charged B decays*, Phys. Lett. **B265** (1991) 172.
- [28] D. Atwood, I. Dunietz, and A. Soni, *Enhanced CP violation with $B \rightarrow KD^0(\bar{D}^0)$ modes and extraction of the CKM angle γ* , Phys. Rev. Lett. **78** (1997) 3257, arXiv:hep-ph/9612433.
- [29] D. Atwood, I. Dunietz, and A. Soni, *Improved methods for observing CP violation in $B^\pm \rightarrow KD$ and measuring the CKM phase γ* , Phys. Rev. **D63** (2001) 036005, arXiv:hep-ph/0008090.
- [30] D. Atwood and A. Soni, *Role of charm factory in extracting CKM phase information via $B \rightarrow DK$* , Phys. Rev. **D68** (2003) 033003, arXiv:hep-ph/0304085.
- [31] R. H. Dalitz, *On the analysis of tau-meson data and the nature of the tau-meson*, Phil. Mag. **44** (1953) 1068.
- [32] M. Goldhaber and J. L. Rosner, *Mixing of Neutral Charmed Mesons and Tests for CP Violation in their Decays*, Phys. Rev. **D15** (1977) 1254.
- [33] BaBar collaboration, B. Aubert *et al.*, *Improved measurement of the CKM angle γ in $B^\mp \rightarrow D^{(*)}K^{(*)\mp}$ decays with a Dalitz plot analysis of D decays to $K_S^0\pi^+\pi^-$ and $K_S^0K^+K^-$* , Phys. Rev. **D78** (2008) 034023.
- [34] BaBar collaboration, P. del Amo Sanchez *et al.*, *Evidence for direct CP violation in the measurement of the CKM angle γ with $B^\mp \rightarrow D^{(*)}K^{(*)\mp}$ decays*, Phys. Rev. Lett. **105** (2010) 121801, arXiv:1005.1096.
- [35] M. Gronau, *Improving bounds on γ in $B^\pm \rightarrow DK^\pm$ and $B^{\pm,0} \rightarrow DX_s^{\pm,0}$* , Phys. Lett. **B557** (2003) 198, arXiv:hep-ph/0211282.
- [36] BaBar collaboration, B. Aubert *et al.*, *Measurement of γ in $B^\mp \rightarrow D^{(*)}K^\mp$ decays with a Dalitz analysis of $D \rightarrow K_S^0\pi^-\pi^+$* , Phys. Rev. Lett. **95** (2005) 121802, arXiv:hep-ex/0504039.
- [37] Belle collaboration, A. Poluektov *et al.*, *Measurement of ϕ_3 with Dalitz plot analysis of $B^\pm \rightarrow D^{(*)}K^\pm$ decay*, Phys. Rev. **D70** (2004) 072003, arXiv:hep-ex/0406067.
- [38] Belle collaboration, A. Poluektov *et al.*, *Evidence for direct CP violation in the decay $B^\pm \rightarrow D^{(*)}K^\pm$, $D \rightarrow K_S^0\pi^+\pi^-$ and measurement of the CKM phase ϕ_3* , Phys. Rev. **D81** (2010) 112002, arXiv:1003.3360.
- [39] CLEO collaboration, R. A. Briere *et al.*, *First model-independent determination of the relative strong phase between D^0 and $\bar{D}^0 \rightarrow K_S^0\pi^+\pi^-$ and its impact on the CKM angle γ/ϕ_3 measurement*, Phys. Rev. D **80** (2009) 032002.
- [40] Belle collaboration, H. Aihara *et al.*, *First measurement of ϕ_3 with a model-independent Dalitz plot analysis of $B \rightarrow DK$, $D \rightarrow K_S^0\pi\pi$ decay*, Phys. Rev. **D85** (2012) 112014, arXiv:1204.6561.
- [41] LHCb collaboration, R. Aaij *et al.*, *A model-independent Dalitz plot analysis of*

- $B^\pm \rightarrow DK^\pm$ with $D \rightarrow K_S^0 h^+ h^-$ ($h = \pi, K$) decays and constraints on the CKM angle γ , Phys. Lett. **B718** (2012) 43, arXiv:1209.5869.
- [42] BaBar collaboration, B. Aubert *et al.*, Constraints on the CKM angle $B^0 \rightarrow \bar{D}^0(D^0)K^{*0}$ from a Dalitz analysis of $D^0 \rightarrow K_S^0 \pi^+ \pi^-$, Phys. Rev. **D79** (2009) 072003, arXiv:0805.2001.
- [43] Belle collaboration, K. Negishi *et al.*, First model-independent Dalitz analysis of $B^0 \rightarrow DK^{*0}$, $D \rightarrow K_S^0 \pi^+ \pi^-$ decay, Prog. Theor. Exp. Phys. **2016** (2016) 043C01, arXiv:1509.01098.
- [44] LHCb collaboration, R. Aaij *et al.*, Measurement of CP violation and constraints on the CKM angle γ in $B^\pm \rightarrow DK^\pm$ with $D \rightarrow K_S^0 \pi^+ \pi^-$ decays, Nucl. Phys. **B888** (2014) 169, arXiv:1407.6211.
- [45] LHCb collaboration, R. Aaij *et al.*, Measurement of the CKM angle γ using $B^0 \rightarrow DK^{*0}$ with $D \rightarrow K_S^0 \pi^+ \pi^-$ decays, arXiv:1605.01082, submitted to JHEP.
- [46] LHCb collaboration, R. Aaij *et al.*, Measurement of the CKM angle γ using $B^\pm \rightarrow DK^\pm$ with $D \rightarrow K_S^0 \pi^+ \pi^-$, $K_S^0 K^+ K^-$ decays, JHEP **10** (2014) 097, arXiv:1408.2748.
- [47] LHCb collaboration, R. Aaij *et al.*, Model-independent measurement of the CKM angle γ using $B^0 \rightarrow DK^{*0}$ decays with $D \rightarrow K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$, JHEP **06** (2016) 131, arXiv:1604.01525.
- [48] L. Evans and P. Bryant, LHC Machine, Journal of Instrumentation **3** (2008), no. 08 S08001.
- [49] ATLAS collaboration, G. Aad *et al.*, Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC, Eur. Phys. J. **C73** (2013), no. 8 2518, arXiv:1302.4393, updated results and plots available at: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [50] LHCb collaboration, R. Aaij *et al.*, LHCb detector performance, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.
- [51] Service graphique CERN, Overall view of the LHC. Vue d'ensemble du LHC, 2014. General Photo, OPEN-PHO-CHART-2014-006.
- [52] LHCb collaboration, A. A. Alves Jr. *et al.*, The LHCb detector at the LHC, JINST **3** (2008) S08005.
- [53] C. Elsasser, $b\bar{b}$ production at the LHCb experiment, provided by the LHCb Speakers' Bureau.
- [54] R. Aaij *et al.*, Performance of the LHCb Vertex Locator, JINST **9** (2014) P09007, arXiv:1405.7808.
- [55] O. Steinkamp, Silicon strip detectors for the LHCb experiment, Nucl. Instrum. Meth. **A541** (2005) 83.
- [56] R. Arink *et al.*, Performance of the LHCb Outer Tracker, JINST **9** (2014) P01002,

arXiv:1311.3893.

- [57] E. Nappi, *Aerogel and its applications to RICH detectors*, Tech. Rep. SLAC-J-ICFA-17-3, Dec, 1998.
- [58] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.
- [59] G. S. Atoyan *et al.*, *Lead-scintillator electromagnetic calorimeter with wavelength shifting fiber readout*, Nucl. Instrum. Meth. **A320** (1992) 144.
- [60] LHCb collaboration, I. Machikhiliyan, *The LHCb electromagnetic calorimeter*, J. Phys. Conf. Ser. **160** (2009) 012047.
- [61] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.
- [62] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.
- [63] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.
- [64] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984; B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, arXiv:physics/0408124.
- [65] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, arXiv:1408.1251.
- [66] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, Tech. Rep. LHCb-2007-015, CERN, Geneva, May, 2007.
- [67] R. Forty, *Ring-imaging Cherenkov detectors for LHC-B*, Nucl. Inst. Meth. A **384** (1996) 167.
- [68] I. Belyaev *et al.*, *Handling of the generation of primary events in GAUSS, the LHCb simulation framework*, Nuclear Science Symposium Conference Record (NSS/MIC) **IEEE** (2010) 1155.
- [69] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.
- [70] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.
- [71] J. Pumplin *et al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195.
- [72] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

- [73] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [74] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.
- [75] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001, and 2013 partial update for the 2014 edition.
- [76] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.
- [77] A. Hoecker *et al.*, *Toolkit for multivariate data analysis*, arXiv:physics/0703039, <http://tmva.sourceforge.net>.
- [78] BaBar collaboration, P. del Amo Sanchez *et al.*, *Study of $B \rightarrow X\gamma$ decays and determination of $|V_{td}/V_{ts}|$* , Phys. Rev. **D82** (2010) 051101, arXiv:1005.4087.
- [79] A. Powell *et al.*, *Particle identification at LHCb*, PoS **ICHEP2010** (2010) 020, LHCb-PROC-2011-008.
- [80] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- [81] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
- [82] D. Hill, *Studies using $B^\pm \rightarrow Dh^\pm$ decays in the LHCb Run I data*, Chapter 3, PhD thesis, University of Oxford, 2015, CERN-THESIS-2015-187.
- [83] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP observables in $B^\pm \rightarrow DK^\pm$ and $B^\pm \rightarrow D\pi^\pm$ with two- and four-body D meson decays*, Phys. Lett. **B760** (2016) 117, arXiv:1603.08993.
- [84] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949) 149.
- [85] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots available at: <http://www.slac.stanford.edu/xorg/hfag/>.
- [86] LHCb collaboration, *Model-independent measurement of CP violation parameters in $B^\pm \rightarrow (K_S^0 h^+ h^-)_D K^\pm$ decays*, LHCb-CONF-2013-004.
- [87] A. Bondar, A. Poluektov, and V. Vorobiev, *Charm mixing in a model-independent analysis of correlated $D^0 \bar{D}^0$ decays*, Phys. Rev. D **82** (2010) 034033.
- [88] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \rightarrow K^- K^+$ and $D^0 \rightarrow \pi^- \pi^+$ decays*, JHEP **07** (2014) 041, arXiv:1405.2797.

- [89] A. Vallier, *Measurement of the CKM angle gamma in the $B^0 \rightarrow DK^{*0}$ decays using the Dalitz method in the LHCb experiment at CERN and photon reconstruction optimisation for the LHCb detector upgrade*, PhD thesis, Laboratoire de l'Accélérateur Linéaire, Université Paris-Sud, 2015.
- [90] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation parameters in $B^0 \rightarrow DK^{*0}$ decays*, Phys. Rev. **D90** (2014) 112002, arXiv:1407.8136.
- [91] A. Martín Sánchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 Calorimeter Trigger*, LHCb-PUB-2011-026.
- [92] LHCb collaboration, R. Aaij *et al.*, *Constraints on the unitarity triangle angle γ from Dalitz plot analysis of $B^0 \rightarrow DK^+\pi^-$ decays*, Phys. Rev. **D93** (2016), no. 11 112018, arXiv:1602.03455.
- [93] LHCb collaboration, R. Aaij *et al.*, *Dalitz plot analysis of $B^0 \rightarrow \bar{D}^0 \pi^+\pi^-$ decays*, Phys. Rev. **D92** (2015) 032002, arXiv:1505.01710.
- [94] J. Neyman, Phil. Trans. Royal Soc. London, Series A **236** (1937) 333, Reprinted in *A Selection of Early Statistical Papers on J. Neyman* (University of California Press, Berkeley, 1967), pp. 250–289.
- [95] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873.
- [96] LHCb collaboration, *LHCb γ combination update from $B \rightarrow DKX$ decays*, LHCb-CONF-2016-001.
- [97] LHCb collaboration, R. Aaij *et al.*, *A study of CP violation in $B^\mp \rightarrow Dh^\mp$ ($h = K, \pi$) with the modes $D \rightarrow K^\mp \pi^\pm \pi^0$, $D \rightarrow \pi^+ \pi^- \pi^0$ and $D \rightarrow K^+ K^- \pi^0$* , Phys. Rev. **D91** (2015) 112014, arXiv:1504.05442.
- [98] LHCb collaboration, R. Aaij *et al.*, *A study of CP violation in $B^\pm \rightarrow DK^\pm$ and $B^\pm \rightarrow D\pi^\pm$ decays with $D \rightarrow K_S^0 K^\pm \pi^\mp$ final states*, Phys. Lett. **B733** (2014) 36, arXiv:1402.2982.
- [99] LHCb collaboration, R. Aaij *et al.*, *Study of $B^- \rightarrow DK^- \pi^+ \pi^-$ and $B^- \rightarrow D\pi^- \pi^+ \pi^-$ decays and determination of the CKM angle γ* , Phys. Rev. **D92** (2015) 112005, arXiv:1505.07044.
- [100] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*, JHEP **11** (2014) 060, arXiv:1407.6127.
- [101] LHCb, R. Aaij *et al.*, *Implications of LHCb measurements and future prospects*, Eur. Phys. J. **C73** (2013), no. 4 2373, arXiv:1208.3355.
- [102] T. Aushev *et al.*, *Physics at Super B Factory*, arXiv:1002.5012.
- [103] BESIII collaboration, D. Ambrose, *Relative strong-phase difference between D^0 and $\bar{D}^0 \rightarrow K_S^0 \pi^+ \pi^-$* , in *American Physical Society April Meeting, Savannah, Georgia, USA, April 5 – April 8, 2014*.

- [104] J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91** (2015), no. 7 073007, arXiv:1501.05013.