Measurement of the CKM angle γ and development of a novel, combined GGSZ analysis of $B \to D^{(*)}h^{(*)}$ decays at LHCb

Jackson William Smith
of Fitzwilliam College

CERN-THESIS-2019-185
//2019





UNIVERSITY OF CAMBRIDGE

Submitted March 2019
This thesis is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Jackson W Smith

Measurement of the CKM angle γ and development of a novel, combined GGSZ analysis of $B \to D^{(*)}h^{(*)}$ decays at LHCb

Jackson William Smith

The angle γ is a fundamental parameter of the Standard Model, within which it quantifies the degree to which $\mathcal{CP}$ violation is permitted. It is presently one of the least well-constrained parameters of the CKM sector, which embodies the description of quark interactions. This thesis details work undertaken by the author at the LHCb experiment with the aim of reducing the uncertainty in γ. It describes a measurement performed using a Dalitz analyis of $B^0 \to DK^{*0}$ decays, and a study which extends this work to a simultaneous Dalitz analysis of several B-meson decay modes of the form $B \to D^{(*)}K^{(*)}$. In each analysis, γ is extracted by studying the interference between $\mathcal{CP}$ eigenstates D^0 and $\overline{D^0}$ in the common final state $D \to K^0_S\pi^+\pi^-$, where D represents either a D^0 or $\overline{D^0}$ meson.

The measurement of γ in $B^0 \to DK^{*0}$ decays includes the full 3fb^{-1} Run 1 dataset gathered at LHCb, and uses a model-dependent approach to yield the 'Cartesian parameters' [1]

$$\begin{aligned} x_- &= -0.15 \pm 0.14 \pm 0.03 \pm 0.01, \\ y_- &= 0.25 \pm 0.15 \pm 0.06 \pm 0.01, \\ x_+ &= 0.05 \pm 0.24 \pm 0.04 \pm 0.01, \quad \text{and} \\ y_+ &= -0.65\ ^{+0.24}_{-0.23} \pm 0.08 \pm 0.01, \end{aligned}$$

where the first uncertainties are statistical, the second systematic and the third arise from the uncertainty on the $D \to K^0_S\pi^+\pi^-$ amplitude model used. These results imply (the relation between $\{x_\pm, y_\pm\}$, and angle γ, is described within) a value for γ of

$$\gamma = (80^{+21}_{-22})^\circ.$$

This measurement has been combined statistically with others performed at LHCb, to give a value for γ of $(74.0^{+5.0}_{-5.8})^\circ$ [2].

In the simultaneous analysis, both model-dependent and model-independent approaches to the determination of γ are studied. Using the full Run 1 LHCb dataset, the model-dependent approach yields preliminary uncertainties of

$$\begin{aligned} \sigma_{x_-} &= \pm 0.019 \pm 0.010 \pm 0.001, \\ \sigma_{y_-} &= \pm 0.013 \pm 0.010 \pm 0.005, \\ \sigma_{x_+} &= \pm 0.018 \pm 0.010 \pm 0.005, \quad \text{and} \\ \sigma_{y_+} &= \pm 0.018 \pm 0.010 \pm 0.010, \end{aligned}$$

where the first numbers are statistical, the second are estimated systematics arising from the experimental method used, and the third are estimated systematics arising from the uncertainty on the $D \to K^0_S\pi^+\pi^-$ amplitude model. The statistical covariances obtained for the Cartesian parameters

propagate to give an estimated statistical uncertainty of 12° on the value of γ. The model-independent approach yields preliminary uncertainties of

$$\begin{aligned} \sigma_{x_-} &= \pm 0.021 \pm 0.010 \pm 0.005, \\ \sigma_{y_-} &= \pm 0.022 \pm 0.005 \pm 0.010, \\ \sigma_{x_+} &= \pm 0.023 \pm 0.010 \pm 0.005, \quad \text{and} \\ \sigma_{y_+} &= \pm 0.029 \pm 0.005 \pm 0.010, \end{aligned}$$

where in this case the final numbers are estimated systematics arising from the uncertainty on the binned $D \to K^0_S \pi^+ \pi^-$ strong phase parameters. The statistical covariances obtained for the Cartesian parameters propagate to give an estimated statistical uncertainty of 13.5° on the value of γ.

In addition, work undertaken to ensure the continued performance of the RICH subdetectors of LHCb is described. These subdetectors form a crucial part of the particle-identification system of LHCb, whose accuracy allows the precise study of processes with hadronic final states, such as the decays mentioned above.

Acknowledgements

I must begin by expressing my gratitude to all those who have helped me over the course of this work; whether by providing technical advice, physical insights or simply a welcome break from the lab. You know who you are. Not to remain nameless is my supervisor Prof. Val Gibson, who has overseen this work from the very start.

This work could not have been completed without the immense technical and organisational efforts of the CERN staff, without whom the LHC would be worth only its weight in scrap metal, and the hundreds of collaborators within LHCb – both those whose ingenuity drives the development of ever more accurate and reliable instruments, and those who work with the LHCb detector data; converting binary zeroes and ones into physics results.

For Natalia and my parents, Lisa and Ali.

Preface

The Standard Model has now been the dominant theory of fundamental physics for over half a century. It is incredibly successful, giving rise to (amongst other things) one of the most accurately tested theoretical predictions within physics; the value of the electron's anomalous magnetic dipole moment [3]:

$$a = \frac{g_{\text{QED}} - 2}{2},$$

and

$$a_{\text{experiment}} = 0.00115965218076(27), \tag{1}$$

$$a_{\text{theory}} = 0.00115965218085(76), \tag{2}$$

where the theoretical calculation is performed by summing amplitudes for all Feynman diagrams with four loops or fewer [4].

However, there is good reason to believe that the Standard Model is not the last word. It provides no fundamental motivation for the existence of exactly *three* generations of quarks and leptons, and cannot account for non-massless neutrinos. Indeed, cosmological evidence of Dark Matter provide a strong indication that there remain unknown sources of mass within the Universe, which defy description within the Standard Model. Many theorists now believe that the Standard Model will eventually be revealed to be merely an effective low-energy approximation to some much more fundamental description of reality. Proposed candidates for this more fundamental description range from 'supersymmetric' extensions of the Standard Model [5] to theories based on entirely novel theoretical frameworks, such as 'string' theory [6] or 'M' theory [7].

Charge-Parity ($\mathcal{CP}$) violation, the asymmetry of matter and antimatter processes, provides an important arena in which to test the limits of the Standard Model. Evidence of $\mathcal{CP}$ violation is all around us; we live in an observable Universe which is completely dominated by matter. However, within the Standard Model, only very limited $\mathcal{CP}$ violation is allowed, and it is clear that this effect is insufficient to explain the observed discrepancy. By precisely measuring many parameters that are sensitive to $\mathcal{CP}$ violation, it should be possible to overconstrain the Standard Model, to quantify the precise extent of the $\mathcal{CP}$ violation it permits, and triangulate the processes by which any extra $\mathcal{CP}$ violation arises.

B-meson decays provide an excellent laboratory with which to study $\mathcal{CP}$ violation, since within the Standard Model the largest $\mathcal{CP}$ violating effects are expected here. The LHCb detector, based at CERN's Large Hadron Collider (LHC) is perfectly situated to contribute significant advances in the measurement of $\mathcal{CP}$ violation, since it has been purpose-built to make precision measurements of decays involving b quarks. In particular, a precise determination of the $\mathcal{CP}$ sensitive Standard Model parameter γ is a crucial aim of the LHCb experiment; being one of the 'flagship' measurements which the detector

was built to pursue. This parameter is currently one of the least well-constrained within the Standard Model.

This thesis reports work undertaken by the author towards improving the experimental precision with which the parameter γ is determined. The tasks described were undertaken between October 2012 and September 2016 as part of the LHCb collaboration. This work may be divided into three components, relating to

1. A measurement of γ using a model-dependent analysis of $B^0 \to DK^{*0}$ decays, where $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$, using the full 3fb^{-1} Run 1 LHCb dataset. This measurement has been published [1]. The author was one of the main contributors to this paper.

2. A study of γ using a simultaneous analysis of several decay modes of the form $B \to D^{(*)}h^{(*)}$, where h is used to represent a kaon or pion, using the full 3fb^{-1} Run 1 LHCb dataset. Specifically, the modes studied are:

   - $B^+ \to DK^+$, with $D \to K^0_S\pi^+\pi^-$;
   - $B^+ \to D\pi^+$, with $D \to K^0_S\pi^+\pi^-$, which acts as a control mode;
   - $B^0 \to DK^{*0}$, with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$;
   - $B^0_s \to DK^{*0}$, with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$, which acts as a control mode;
   - $B^+ \to D^*K^+$, with $D^* \to D\gamma/\pi^0$ and $D \to K^0_S\pi^+\pi^-$;
   - $B^+ \to D^*\pi^+$, with $D^* \to D\gamma/\pi^0$ and $D \to K^0_S\pi^+\pi^-$, which acts as a control mode; and
   - $B^+ \to DK^{*+}$, with $K^{*+} \to K^0_S\pi^+$ and $D \to K^0_S\pi^+\pi^-$,

   and their $\mathcal{CP}$ conjugate modes. Results are obtained using both model-dependent and model-independent approaches.

3. Work undertaken to test hardware designed for the upgrade of the RICH subdetector system of LHCb. Specifically, the author held sole responsibility for implementing the software used to configure and readout the digital circuit boards interfaced to the multi-anode photomultiplier tubes during beam tests. This work has also been published [8].

Finally, in addition to the work described herein, the author held collaboration wide responsibility for co-ordinating the production of a special class of 'filtered' simulation productions. These make use of selection requirements, imposed immediately after simulated events are reconstructed, to select only the best events to write to disk, and thus use storage space more efficiently.

Since the first analysis performed is encompassed and significantly extended by the second, it is the latter analysis which forms the main subject of this thesis. Throughout this document, 'the present analysis' will refer to the simultaneous $B \to D^{(*)}h^{(*)}$ analysis.

Chapter 1 begins by introducing the theory of $\mathcal{CP}$ violation, the significance of $\mathcal{CP}$ parameter γ, and the strategies that may be used to measure it in B-meson decays. Chapter 2 describes the resonance models used to extract γ in the phase space of $D^0 \to K^0_S\pi^+\pi^-$ decays, where the neutral D-meson

originates from a B-meson decay. Chapter 3 introduces the LHCb detector, and describes the features which make it well-suited to the study of B-meson decays. This chapter also describes the author's work towards the upgrade of the RICH subdetector system.

The remaining chapters describe the extraction of γ from B-meson decays recorded in LHCb's Run 1 dataset. Chapters 4–6 describe work undertaken by the author towards the simultaneous analysis. Chapter 4 defines the criteria used in the simulataneous study to select candidate events for each decay mode studied. Chapter 5 describes the fit performed on the mass spectrum of each selected sample in the simultaneus study, allowing the yields of signal and background components to be determined. Chapter 6 determines the variation of detection efficiency, over the phase space, for each component present in the simultaneous study. This is needed before the D-meson decay phase space can be fitted.

The measurement performed using only $B^0 \to DK^{*0}$ decays is described in Chapter 7. This work was performed by the present author in collaboration with others, and the author's responsibilties are detailed. It is included chiefly to illustrate the approach taken in the systematic studies, which have not yet been performed for the simultaneous analysis.

Chapter 8 returns to the simultaneous analysis, giving preliminary details of the model-dependent and model-independent simultaneous fits performed in the D-meson decay phase space. This work has been performed by collaborators of the author, using the mass fit and efficiency results of Chapters 5 and 6 as input.

Finally, Chapter 9 summarises the thesis.

Contents

1. $\mathcal{CP}$ violation in B-meson decays **1**

1.1. $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{T}$-symmetry . . . 2
1.2. Neutral meson systems . . . 2
1.3. $\mathcal{CP}$ violation . . . 7
1.4. $\mathcal{CP}$ violation within the Standard Model . . . 11
1.4.1. Experimental status of CKM parameters . . . 15
1.5. Measuring $\mathcal{CP}$ violation in $B \to D^{(*)}h^{(*)}$ decays . . . 18
1.6. Strategy of the simultaneous analysis . . . 25

2. Modelling $D^0 \to K_S\pi^+\pi^-$ decays **27**

2.1. Effective theory of resonances . . . 27
2.2. Isobar model . . . 28
2.2.1. Centrifugal barrier factors B_L . . . 29
2.2.2. Angular distribution Z_L . . . 30
2.2.3. Dynamical function T_r . . . 35
2.2.4. Limitations of the Isobar model . . . 37
2.3. The K-matrix model . . . 39
2.3.1. The S-matrix formalism . . . 39
2.3.2. The K-matrix formalism . . . 42
2.4. Decay model implemented in the present analysis . . . 51

3. The LHC and the LHCb detector **55**

3.1. The LHC . . . 55
3.2. The LHCb experiment . . . 58
3.2.1. Vertex Locator . . . 64
3.2.2. Tracking stations . . . 66
3.2.3. Dipole magnet . . . 70
3.2.4. Ring-Imaging Cherenkov detectors . . . 71
3.2.5. Calorimeters . . . 75
3.2.6. Muon systems . . . 77
3.2.7. The LHCb trigger . . . 78
3.2.8. Software and simulation . . . 81
3.3. The LHCb Upgrade . . . 83
3.3.1. Opto-electronic chain . . . 83
3.3.2. DAQ system . . . 84
3.3.3. Beam test setup . . . 87
3.3.4. Results . . . 89
3.4. Summary . . . 91

4. Selection of candidate events **93**
4.1. Stripping requirements . . . 95
4.2. Preselection requirements common to all decay modes . . . 97
4.2.1. Fake D/K^0_S candidate backgrounds . . . 99
4.3. Mode-specific requirements . . . 104
4.3.1. $B^\pm \to Dh^\pm$ . . . 104
4.3.2. $B^0_{(s)} \to DK^{*0}$ . . . 104
4.3.3. $B^\pm \to D^*h^\pm$ . . . 105
4.3.4. $B^\pm \to DK^{*\pm}$ . . . 110
4.4. MVA selection . . . 113
4.4.1. Training . . . 116
4.5. MVA and PID cut optimisation . . . 129
4.6. Selection summary . . . 139

5. Fit to reconstructed B-meson mass spectrum **141**
5.1. Fit to selected $B^\pm \to Dh^\pm$ candidates . . . 144
5.1.1. Constraints applied . . . 157
5.1.2. Mass fit result . . . 159
5.2. Fit to selected $B^0 \to DK^{*0}$ candidates . . . 170
5.2.1. Mass fit result . . . 177
5.3. Fit to selected $B^\pm \to DK^{*\pm}$ candidates . . . 182
5.3.1. Mass fit result . . . 187
5.4. Fit to selected $B^\pm \to D^*h^\pm$ candidates . . . 192
5.4.1. Constraints applied . . . 201
5.4.2. Mass fit result . . . 204
5.5. Estimation of fake D background yield . . . 217

6. Determination of the selection efficiency variation over the Dalitz plane **223**
6.1. Edge bin reweighting . . . 230
6.2. Trigger weights . . . 231
6.3. PID weights . . . 251
6.4. Total efficiency . . . 260

7. Measurement of the CKM angle γ using $B^0 \to DK^{*0}$ with $D \to K^0_S\pi^+\pi^-$ decays **263**
7.1. Publication . . . 264

8. Study of γ using the simultaneous fit to all $B \to D^{(*)}K^{(*)}$ decay modes **295**
8.1. $\mathcal{CP}$ fit model . . . 297
8.2. Preliminary results . . . 301
8.3. Expected systematic uncertainties . . . 301

9. Summary **305**

A. The relative weak phase for $B \to D^{(*)}K^{(*)}$ **309**

B. Useful relations of complex numbers **311**

C. Helicity eigenstates for a spin-1 resonance 313

D. Input variables used to train multivariate classifiers 317

Bibliography 335

"There was a fecklessness, a lack of symmetry and order in the clouds, as they thinned and thickened. Was it their own law, or no law, they obeyed?"

— Virginia Woolf

1. $\mathcal{CP}$ violation in B-meson decays

The Standard Model (SM) [9–11] of particle physics defines a set of fundamental particles, and the forces that control their interactions. All matter is described in terms of three generations of quarks and leptons, whilst forces are mediated by a set of bosons. These fundamental particles are summarised in Figure 1.1.

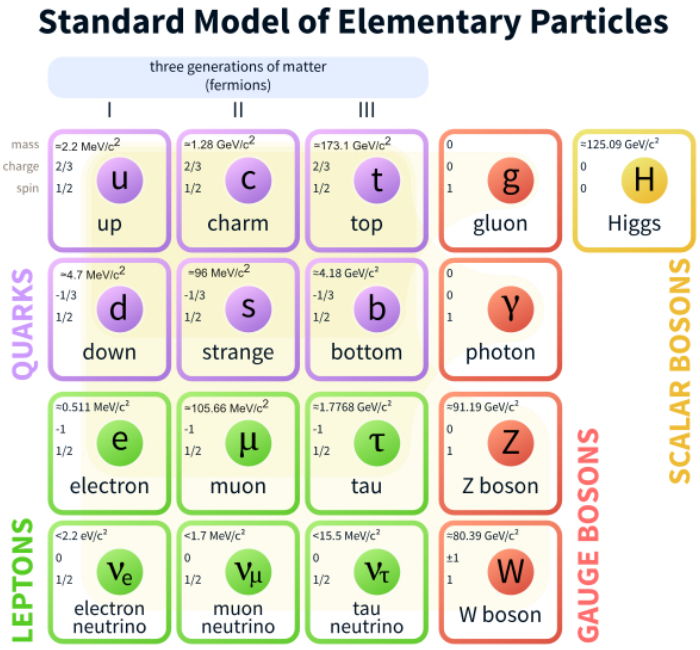


Figure 1.1.: Fundamental particles of the Standard Model [12].

The SM describes particle interactions in terms of symmetries, using the groups $SU(3) \times SU(2) \times U(1)$ to describe strong, weak and electromagnetic interactions in turn. In order to test the Standard Model it is essential to probe these underlying symmetries on which it is based.

This thesis pertains to the measurement of a quantity, γ, which is sensitive to the violation of $\mathcal{CP}$ symmetry. This chapter begins with a description of the symmetries that underly the Standard Model, including $\mathcal{CP}$ symmetry, before going on to focus on the discovery of $\mathcal{CP}$ symmetry violation in the weak sector, introduce the theoretical framework used within the Standard Model to accommodate $\mathcal{CP}$ violation, and to explain the significance of γ. Finally, the method that is used later in this thesis to measure γ will be outlined.

1.1. $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{T}$-symmetry

A physical process is said to have a symmetry if it occurs indistinguishably under some transformation. This statement may be made mathematically rigorous by describing transformations using quantum mechanical operators, $\hat{O}$, which act on quantum vectors describing the state of a system, $|\psi\rangle$. If a process is symmetric under $\hat{O}$, then the equations describing it are equally valid after the substitution $|\psi\rangle \to \hat{O}|\psi\rangle$.

The three discrete operators that lie at the heart of the Standard Model are:

- the charge conjugation operator, $\mathcal{C}$. This converts a particle into its antiparticle, by reversing the sign of all of its additive quantum numbers (or 'charges');
- the parity operator, $\mathcal{P}$. This performs a spatial inversion of a particle's coordinate system through the origin, such that $(t, \vec{x}) \to (t, -\vec{x})$; and
- the time reversal operator, $\mathcal{T}$. This reverses the time coordinate of a particle, such that $(t, \vec{x}) \to (-t, \vec{x})$, and has the effect of reversing a particle's momentum and spin vectors.

Each of these three operators appeals to the intuition as a potential symmetry of nature. The observation of charge and parity violation in weak interactions, made by Wu [13] in 1956, therefore came as something of a surprise to the physics community. It was found that the neutrinos produced in the decay $\pi^+ \to \mu^+\nu_\mu$ are always of left-handed parity, whilst the anti-neutrinos produced in $\pi^- \to \mu^-\overline{\nu}_\mu$ are always right-handed [14]. This case violates both $\mathcal{C}$ and $\mathcal{P}$ symmetry, since the application of either operator to one of these decays yields a forbidden decay. However, this 'maximal' violation of each symmetry suggested a simple solution; the combined operation, $\mathcal{CP}$, was still expected to be a symmetry of the system.

This state of affairs prevailed until in 1964 Cronin, Fitch, Turlay and Christenson [15] discovered that $\mathcal{CP}$ symmetry, too, is violated in weak interactions. In order to understand the experiment that first observed $\mathcal{CP}$ violation, and those that have measured it since, it is useful to study the neutral meson system, which neatly encapsulates the various means by which $\mathcal{CP}$ violation can arise. The combined symmetry, $\mathcal{CPT}$, is a key building block of the Standard Model, since it forms the basis for all local field theories [16–18].

1.2. Neutral meson systems

Neutral mesons provide a phenomenologically rich laboratory due to the non-conservation of flavour in the charged weak interaction. Since flavour is the only quantum number that differs between a neutral meson, X^0, and its conjugate anti-meson, $\overline{X^0}$, this means that it is possible for the two states to mix.

Observable eigenstates

Two sets of eigenstates provide natural bases to describe the neutral meson system. The flavour eigenstates, $\{|X^0\rangle, |\overline{X^0}\rangle\}$, are produced in flavour-conserving strong interactions. However, it is also useful to consider the $\mathcal{CP}$ eigenstates of the system, since these neutral mesons are often observed in final states of well-defined $\mathcal{CP}$, *e.g.* $\pi\pi$ ($\mathcal{CP}$-even) or $\pi\pi\pi$ ($\mathcal{CP}$-odd). The relevant eigenstates may be constructed by considering the action of the $\mathcal{CP}$ operator:

$$\mathcal{CP}|X^0\rangle = |\overline{X^0}\rangle, \tag{1.1}$$

$$\mathcal{CP}|\overline{X^0}\rangle = |X^0\rangle. \tag{1.2}$$

This implies that the superpositions

$$|X_\pm\rangle = \frac{|X^0\rangle \pm |\overline{X^0}\rangle}{\sqrt{2}} \tag{1.3}$$

are eigenstates of $\mathcal{CP}$.

Time evolution

Even before the quark picture of the Standard Model emerged, the weak force was known to violate flavour conservation, allowing recently-discovered 'strange' particles to decay to products without any trace of strangeness. It is therefore reasonable to expect some non-zero amplitude for $X^0 \longleftrightarrow \overline{X^0}$ conversion processes. The formalism that developed over time to describe this situation is covered in several monographs, for example References [19] and [20], or the shorter [21]. A brief exposition of the most relevant aspects is given here.

The Hamiltonian for a free particle, with non-zero probability of decay, is given by the sum of two terms $H = M - \frac{i}{2}\Gamma$. Under the application of the time-dependent Schrödinger equation (TDSE), each eigenstate gains the usual mass-dependent phase, e^{-iMt} plus an exponential attenuation of decay rate Γ. Working in the two-state flavour basis,

$$|\psi\rangle = a(t)|X^0\rangle + b(t)|\overline{X^0}\rangle, \tag{1.4}$$

mixing introduces off-diagonal elements to the Hamiltonian, giving the TDSE [22]

$$i\frac{\partial}{\partial t}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}. \tag{1.5}$$

Note that the individual M_{ij} and Γ_{ij} elements no longer correspond to physical observables, since they are not eigenvalues of the system. For example, M_{11} and Γ_{11} give the mass and decay rate that *would* be observed for an X^0 meson, in the absence of any flavour-violating mixing effect. Equivalently, they give the instantaneous mass and decay rate that would be observed for such a meson just after creation, precisely at $t = 0$, and are therefore not experimentally accessible.

$\mathcal{CPT}$ symmetry requires that the diagonal mass and lifetime elements are equal for both particle and antiparticle, giving $M_{11} = M_{22} \equiv M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. It also requires that the mass and lifetime matrices are separately Hermitian, such that $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. This gives the simplified TDSE

$$i\frac{\partial}{\partial t}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M - \frac{i}{2}\Gamma \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}. \tag{1.6}$$

It is useful to characterise this system by finding its conserved properties; the eigenvectors and their corresponding eigenvalues. In the flavour basis, the eigenstates of the system may be written as

$$|X_1\rangle = p|X^0\rangle + q|\overline{X^0}\rangle, \tag{1.7}$$
$$|X_2\rangle = p|X^0\rangle - q|\overline{X^0}\rangle, \tag{1.8}$$

where the two vectors share just one common pair of constant coefficients, (p, q), due to the restrictions imposed by $\mathcal{CPT}$ invariance [19], and normalisation implies the constraint $|p|^2 + |q|^2 = 1$.

These eigenstates satisfy the time-independent Schrödinger equation, $H|X_j\rangle = E_j|X_j\rangle$, which gives

$$(H - EI)X = 0 \tag{1.9}$$
$$\Rightarrow |H - EI| = 0. \tag{1.10}$$

Solving this characteristic equation allows the eigenvalues to be obtained:

$$(M - \frac{i}{2}\Gamma - E)(M - \frac{i}{2}\Gamma - E) - (M_{12} - \frac{i}{2}\Gamma_{12})(M_{12}^* - \frac{i}{2}\Gamma_{12}^*) = 0 \tag{1.11}$$

$$E_{1,2} = M - \frac{i}{2}\Gamma \mp \sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}. \tag{1.12}$$

Defining variables ΔM and $\Delta\Gamma$ to parameterize the mass and lifetime splitting between the two eigenstates,

$$\Delta M \equiv M_1 - M_2, \tag{1.13}$$
$$\Delta\Gamma \equiv \Gamma_2 - \Gamma_1, \tag{1.14}$$

these eigenvalues may be written as

$$E_{1,2} = M_{1,2} - \frac{i}{2}\Gamma_{1,2} \tag{1.15}$$
$$= \Big(M \mp \frac{\Delta M}{2}\Big) - \frac{i}{2}\Big(\Gamma \mp \frac{\Delta\Gamma}{2}\Big). \tag{1.16}$$

Note that the signs of ΔM and $\Delta\Gamma$ may differ; the question of whether the heavier eigenstate has a shorter or longer lifetime than the lighter eigenstate is a matter for empirical study in each neutral meson system.

The eigenvector equation,

$$(H - EI)\begin{pmatrix} p \\ \pm q \end{pmatrix} = 0, \tag{1.17}$$

allows the relation between p and q to be determined. It yields

$$(M - \frac{i}{2}\Gamma - E)p \pm (M_{12} - \frac{i}{2}\Gamma_{12})q = 0, \text{ and} \tag{1.18}$$

$$(M^*_{12} - \frac{i}{2}\Gamma^*_{12})p \pm (M - \frac{i}{2}\Gamma - E)q = 0, \tag{1.19}$$

which are degenerate, and give

$$\frac{q}{p} = \sqrt{\frac{M^*_{12} - \frac{i}{2}\Gamma^*_{12}}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{1.20}$$

Finally, this determines the time evolution for the eigenstates of the neutral meson system

$$|X_{1,2}(t)\rangle = e^{-i(M_{1,2} - \frac{i}{2}\Gamma_{1,2})t}|X_{1,2}\rangle. \tag{1.21}$$

Flavour oscillations and regeneration

This formalism allows the rate of mixing, for an initially pure X^0 or $\overline{X^0}$ beam to be determined. Equations 1.7 and 1.8 may be re-arranged to give

$$|X^0\rangle \propto |X_1\rangle + |X_2\rangle, \tag{1.22}$$

$$|\overline{X^0}\rangle \propto |X_1\rangle - |X_2\rangle \tag{1.23}$$

which evolve with time to give

$$|X^0(t)\rangle \propto e^{-i(M_1 - \frac{i}{2}\Gamma_1)t}|X_1\rangle + e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}|X_2\rangle, \tag{1.24}$$

$$|\overline{X^0}(t)\rangle \propto e^{-i(M_1 - \frac{i}{2}\Gamma_1)t}|X_1\rangle - e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}|X_2\rangle. \tag{1.25}$$

Substituting in from Equations 1.7 and 1.8 gives

$$|X^0(t)\rangle \propto f_+(t)|X^0\rangle + \frac{q}{p}f_-(t)|\overline{X^0}\rangle, \tag{1.26}$$

$$|\overline{X^0}(t)\rangle \propto f_-(t)|X^0\rangle + \frac{q}{p}f_+(t)|\overline{X^0}\rangle, \tag{1.27}$$

where

$$f_\pm(t) \equiv \frac{1}{2}\left[e^{-i(M_1 - \frac{i}{2}\Gamma_1)t} \pm e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}\right]. \tag{1.28}$$

The probabilities to measure each of the flavour eigenstates at time t after production (or equivalently at time t after a measurement in the flavour basis) are thus given by

$$P(X^0 \to X^0 : t) = |\langle X^0|X^0(t)\rangle|^2 = |f_+(t)|^2, \tag{1.29}$$
$$P(X^0 \to \overline{X^0} : t) = |\langle \overline{X^0}|X^0(t)\rangle|^2 = |\frac{q}{p} f_-(t)|^2, \tag{1.30}$$
$$P(\overline{X^0} \to X^0 : t) = |\langle X^0|\overline{X^0}(t)\rangle|^2 = |f_-(t)|^2, \tag{1.31}$$
$$P(\overline{X^0} \to \overline{X^0} : t) = |\langle \overline{X^0}|\overline{X^0}(t)\rangle|^2 = |\frac{q}{p} f_+(t)|^2, \tag{1.32}$$

where

$$|f_\pm(t)|^2 = \frac{1}{4}\left[e^{-\Gamma_1 t} + e^{-\Gamma_2 t} \pm 2e^{-\overline{\Gamma} t}\cos(\Delta M t)\right], \tag{1.33}$$

and $\overline{\Gamma} = \frac{(\Gamma_1+\Gamma_2)}{2}$ gives the mean decay width of the two eigenstates.

The cosine term of equation (1.33) describes the 'flavour oscillation' between the X^0 and $\overline{X^0}$ states. The likelihood of observing oscillation before the neutral state decays altogether is parameterized by the ratio of the period of the oscillations to the mean decay time, $x = \frac{|\Delta M|}{\overline{\Gamma}}$, known as the 'oscillation parameter'. For $x \ll 1$ it is unlikely for any oscillation to occur, whilst for $x \gg 1$ many oscillations are expected to occur before decay.

For the neutral kaon system, two free eigenstates of significantly different lifetime are observed [3]

$$\tau_S = (0.08954 \pm 0.00004)\,\text{ns}, \tag{1.34}$$
$$\tau_L = (51.16 \pm 0.21)\,\text{ns}. \tag{1.35}$$

The mass splitting between them is very small,

$$\Delta M_K = M_L - M_S = (52.93 \pm 0.09) \times 10^8 \hbar s^{-1} \tag{1.36}$$

giving very slow flavour oscillations relative to the decay time, with $x_K \sim 0.05$. In the neutral B-meson systems ($B^0_d - \overline{B^0_d}$ and $B^0_s - \overline{B^0_s}$), the lifetimes of the two free eigenstates are much closer, since almost all final states are common to the two. The mass splitting in the B^0_s system is significantly larger than in the B^0_d system, giving much faster oscillations in this case [3]

$$x_{B_d} = 0.771 \pm 0.008, \tag{1.37}$$
$$x_{B_s} = 26.2 \pm 0.5. \tag{1.38}$$

LHCb has contributed significantly to the precision of these values, by measuring the mass splitting for both neutral B-meson systems [23–25]

$$\Delta M_{B_d} = (0.5156 \pm 0.0051 \pm 0.0033)\,\text{ps}^{-1}, \tag{1.39}$$
$$\Delta M_{B_s} = (17.725 \pm 0.041 \pm 0.026)\,\text{ps}^{-1}, \tag{1.40}$$

where the statistical uncertainty appears before the systematic uncertainty in each case. The B_s^0 result represents the world's-best determination of the parameter, whilst the B_d^0 result agrees closely with previous measurements [26]. A first observation of mixing in the neutral D-meson system has been made by LHCb [27, 28].

In addition to oscillation, 'regeneration' has been studied for the neutral kaon system, where the lifetimes of the two free eigenstates vary significantly [29]. An initially pure K^0 beam is allowed to propagate over a distance $\mathcal{O}(10)$ m, over which time the shorter-lived free eigenstate, $|K_S\rangle = |K_1\rangle$ decays to a negligible level, leaving behind a pure $|K_L\rangle = |K_2\rangle$ beam. If the resulting beam is passed through a target of nuclear matter, the K^0 and $\overline{K^0}$ components interact differently.[1] This is equivalent to a measurement in the basis of flavour eigenstates and so each incoming state will be collapsed to either $|K^0\rangle$ or $|\overline{K^0}\rangle$. These contain the same relative amounts of $|K_S\rangle$ and $|K_L\rangle$ that would be present for a 'freshly-produced' kaon, and so the short-lived component has been regenerated.

1.3. $\mathcal{CP}$ violation

$\mathcal{CP}$ violation was first observed in the neutral kaon system, in 1964 [15]. At the time it was known that mixing must occur in this system, since two neutral kaons of very different lifetime had been observed [30]. However, due to the very low value of the oscillation parameter for this system, no direct time-dependent oscillation measurements had been made. Regeneration was studied by Piccione [31] and Adair [32], who observed an excess of $|K_S\rangle$ regeneration, implying that the free eigenstates of the system were not identical to the $\mathcal{CP}$ eigenstates as had been assumed.

In order to verify Adair's observations, Cronin, Fitch, Turlay and Christenson studied decays to $\mathcal{CP}$ eigenstates, for a beam of pure $|K_L\rangle$ obtained by propagating a beam of neutral kaons through a vacuum vessel approximately 20 m in length. From a total of 22700 decays, 45 $\mathcal{CP}$ even final states ($\pi\pi$) were detected, whilst the expected number would be < 1, if these decays were due only to pollution from surviving $|K_S\rangle$. This work firmly opened the door to $\mathcal{CP}$ violation, and ultimately won the Nobel Prize for its authors. Whilst $\mathcal{CP}$ violation may be parameterized within the Standard Model, it does not offer any explanation for the presence of $\mathcal{CP}$ violation in weak decays only, or the reason that the violation is so small, giving physical eigenstates that are very nearly identical to the $\mathcal{CP}$ eigenstates, but not quite.

The $\mathcal{CP}$ violation observed by Cronin *et al* arises from the fact that the eigenstates of the neutral meson system are not identical to the $\mathcal{CP}$ eigenstates ($|q/p| \neq 1$). However, there are two other ways in which $\mathcal{CP}$ violation can modify the amplitudes for various processes. To classify these different types of $\mathcal{CP}$ violation it is useful to consider the (experimentally-observable) rates for two $\mathcal{CP}$ conjugate neutral decays:

$$\Gamma(X^0(t) \to f) = |A_f|^2 \Big\{ |f_+(t)|^2 + |\lambda_f|^2 |f_-(t)|^2 + 2\mathrm{Re}(\lambda_f f_+^*(t) f_-(t)) \Big\}, \tag{1.41}$$

$$\Gamma(\overline{X^0}(t) \to \overline{f}) = |\overline{A}_{\overline{f}}|^2 \Big\{ |f_+(t)|^2 + |\overline{\lambda}_{\overline{f}}|^2 |f_-(t)|^2 + 2\mathrm{Re}(\overline{\lambda}_{\overline{f}} f_+^*(t) f_-(t)) \Big\}, \tag{1.42}$$

[1] K^0 can only scatter, whilst $\overline{K^0}$ may interact strongly with protons or neutrons to form strange baryons.

where the complex factors λ_f and $\overline{\lambda}_{\overline{f}}$ are given by

$$\lambda_f \equiv \frac{q\overline{A}_f}{pA_f}, \quad \overline{\lambda}_{\overline{f}} \equiv \frac{1}{\lambda_{\overline{f}}} \equiv \frac{pA_{\overline{f}}}{q\overline{A}_{\overline{f}}} \tag{1.43}$$

and the transition matrix elements are defined as

$$A_f \equiv \langle f|T|X^0\rangle, \qquad A_{\overline{f}} \equiv \langle \overline{f}|T|X^0\rangle, \tag{1.44}$$

$$\overline{A}_f \equiv \langle f|T|\overline{X^0}\rangle, \qquad \overline{A}_{\overline{f}} \equiv \langle \overline{f}|T|\overline{X^0}\rangle, \tag{1.45}$$

The first term within the braces corresponds to the situation in which the flavour state at decay matches that at $t = 0$, and so the behaviour matches what would be expected if no oscillations had occurred. The second term corresponds to the case in which an odd number of flavour oscillations occur between $t = 0$ and the decay, so that of the neutral meson is observed to have changed flavour, or 'mixed'. Finally, the third term describes the interference between the mixed and unmixed amplitudes.

Any difference between the two decay rates, Equations 1.41 and 1.42, indicates $\mathcal{CP}$-violation. There are three possible independent ways in which differences can arise:

- The magnitudes of the instantaneous decay amplitudes differ

$$|A_f| \neq |\overline{A}_{\overline{f}}|. \tag{1.46}$$

  This is known as 'direct' $\mathcal{CP}$ violation, or '$\mathcal{CP}$ violation in the decay amplitudes';

- The magnitudes of the conjugate λ-factors differ,

$$|\lambda_f| \neq |\overline{\lambda}_{\overline{f}}|, \tag{1.47}$$

  giving a discrepancy between the second terms of the two rate equations. This is known as '$\mathcal{CP}$ violation in the mixing'; and

- Finally, even if the magnitudes of the λ-factors are equal, their phases may differ,

$$\arg(\lambda_f) \neq \arg(\overline{\lambda}_{\overline{f}}), \text{ or equivalently} \tag{1.48}$$

$$\arg(\lambda_f) \neq -\arg(\lambda_{\overline{f}}) \tag{1.49}$$

  giving a discrepancy in the final terms of the two rate equations. This is known as '$\mathcal{CP}$ violation in the interference (between decay and mixing)'.

The observables used to study each of these cases are discussed in turn below.

$\mathcal{CP}$ violation in the decay

The conjugate instantaneous decay amplitudes may be written as summations over all contributing topologies (*i.e.* over all Feynman diagrams, within a local field theory), giving

$$A_f = \sum_m A_m e^{i\phi_m} e^{i\delta_m} \quad \text{and} \quad \overline{A}_{\overline{f}} = \sum_n A_n e^{-i\phi_n} e^{i\delta_n}, \tag{1.50}$$

where the associated phases have been split into their $\mathcal{CP}$-odd weak phase ($\phi_{m,n}$) and $\mathcal{CP}$-even strong phase ($\delta_{m,n}$) parts, and the magnitudes $A_{m,n}$ are real. The $\mathcal{CP}$-even phases include all contributions from strong processes, such as final state hadronization, since these are $\mathcal{CP}$ invariant.

$\mathcal{CP}$ violation in the decay occurs when

$$\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| = \left|\frac{\sum_n A_n e^{-i\phi_n} e^{i\delta_n}}{\sum_m A_m e^{i\phi_m} e^{i\delta_m}}\right| \neq 1, \tag{1.51}$$

and is thus possible whenever multiple amplitudes of different weak and strong phases contribute to either summation. $\mathcal{CP}$ violation in the decay is the only type that is possible in charged meson decays.

Defining

$$A_f \equiv \langle f|T|X^-\rangle, \tag{1.52}$$
$$\overline{A}_{\overline{f}} \equiv \langle \overline{f}|T|\overline{X^+}\rangle, \tag{1.53}$$

the charged $\mathcal{CP}$ asymmetry may be defined;

$$A_{\mathcal{CP}}^{\pm} \equiv \frac{\Gamma(X^- \to \overline{f}) - \Gamma(X^+ \to f)}{\Gamma(X^- \to \overline{f}) + \Gamma(X^+ \to f)} \tag{1.54}$$
$$= \frac{|\overline{A}_{\overline{f}}/A_f|^2 - 1}{|\overline{A}_{\overline{f}}/A_f|^2 + 1}. \tag{1.55}$$

The present thesis studies $\mathcal{CP}$ violation in the decay ('direct $\mathcal{CP}$ violation') for a variety of decay modes. However, instead of studying the charged $\mathcal{CP}$ asymmetry, where hadronic uncertainties in the included strong phases make it difficult to extract information on the desired weak phases, a more sophisticated approach is taken, which will be described in Section 1.5.

$\mathcal{CP}$ violation in the mixing

$\mathcal{CP}$ violation in the mixing requires $|\lambda_f| \neq |\overline{\lambda}_{\overline{f}}|$, which, assuming no direct $\mathcal{CP}$ violation, means that

$$\left|\frac{q}{p}\right| \neq 1, \tag{1.56}$$

and therefore

$$\left|M_{12} - \frac{i}{2}\Gamma_{12}\right| \neq \left|M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right|. \tag{1.57}$$

The effect of this form of $\mathcal{CP}$ violation is to 'bias' the flavour oscillations of the neutral system, so that over the course of many oscillations the population of one of the free eigenstates grows relative to the other. The corresponding $\mathcal{CP}$ asymmetry takes the form

$$A_{\mathcal{CP}}^{\text{mix}} \equiv \frac{\Gamma(X^0 \to f) - \Gamma(\overline{X^0} \to \overline{f})}{\Gamma(X^0 \to f) + \Gamma(\overline{X^0} \to \overline{f}))} \tag{1.58}$$

$$= \frac{1 - |q/p|^4}{1 + |q/p|^4}. \tag{1.59}$$

$\mathcal{CP}$ violation in the interference

$\mathcal{CP}$ violation in the interference is only possible when both decays proceed to a common final state, *i.e.* $\overline{f} = f$. If the two decays are chosen to be $\mathcal{CP}$ conjugates, as has been assumed thus far, then this common final state, f, must therefore be a $\mathcal{CP}$ eigenstate.

A time-dependent analysis of the conjugate rates may be performed, to extract a corresponding asymmetry. The conjugate rate equations may be written in the form [19]

$$\Gamma_f(t) \equiv \Gamma(X^0(t) \to f) = \frac{|A_f|^2}{2} e^{-\overline{\Gamma}t}\big[I_+(t) + I_-(t)\big], \tag{1.60}$$

$$\overline{\Gamma}_f(t) \equiv \Gamma(\overline{X^0}(t) \to f) = \frac{|\overline{A}_f|^2}{2|\lambda_f|^2} e^{-\overline{\Gamma}t}\big[I_+(t) - I_-(t)\big], \tag{1.61}$$

where the functions $I_\pm(t)$ are given by

$$I_+(t) \equiv (1 + |\lambda_f|^2)\cosh\left(\frac{\Delta\Gamma t}{2}\right) - 2\text{Re}\{\lambda_f\}\sinh\left(\frac{\Delta\Gamma t}{2}\right), \tag{1.62}$$

$$I_-(t) \equiv (1 - |\lambda_f|^2)\cos(\Delta M t) + 2\text{Im}\{\lambda_f\}\sin(\Delta M t). \tag{1.63}$$

Since this form of $\mathcal{CP}$ violation is most prevalent in the $B^0_{(s)}$-meson systems, it is usual to assume that $|q/p| \simeq 1$, since $\mathcal{CP}$ violation in the mixing is negligible for these systems [3]. Under this assumption, the time-dependent $\mathcal{CP}$ asymmetry, defined by

$$A_{\mathcal{CP}}(t) \equiv \frac{\Gamma_f - \overline{\Gamma}_f}{\Gamma_f + \overline{\Gamma}_f}, \tag{1.64}$$

becomes

$$A_{\mathcal{CP}}(t) = \frac{I_-(t)}{I_+(t)} = \frac{(1 - |\lambda_f|^2)\cos(\Delta M t) + 2\text{Im}\{\lambda_f\}\sin(\Delta M t)}{(1 + |\lambda_f|^2)\cosh\left(\frac{\Delta\Gamma t}{2}\right) - 2\text{Re}\{\lambda_f\}\sinh\left(\frac{\Delta\Gamma t}{2}\right)}. \tag{1.65}$$

For the B^0-system, where the mass splitting, ΔM is small, this expression reduces to

$$A_{\mathcal{CP}}(t) = \left(\frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}\right)\cos(\Delta M t) + \left(\frac{2\mathrm{Im}\{\lambda_f\}}{1+|\lambda_f|^2}\right)\sin(\Delta M t), \tag{1.66}$$

where the first term accounts for any direct $\mathcal{CP}$ violation that is present, and the second isolates the $\mathcal{CP}$ violation arising purely in the interference. By choosing a final state with only one contributing topology for each flavour eigenstate, X^0 and $\overline{X^0}$, it is possible to remove any direct $\mathcal{CP}$ violation (which may be affected by hadronic uncertainties) leaving an asymmetry depending on just the $\mathcal{CP}$-odd weak phase.

1.4. $\mathcal{CP}$ violation within the Standard Model

Within the Standard Model, violation of flavour symmetry by the charged weak interaction [33, 34] is accounted for by modifying the three generational quark picture of the model, such that the quark eigenstates involved are not identical to the mass eigenstates that are detected experimentally. These three generations are then defined by the doublets

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix} \text{ and } \begin{pmatrix} t \\ b' \end{pmatrix},$$

where no inter-generational weak couplings exist between quark states from different doublets. The d, s and b quarks that occur in bound states within matter are then superpositions of the primed states:

$$\begin{aligned} |d\rangle &= a_1|d'\rangle + a_2|s'\rangle + a_3|b'\rangle, \\ |s\rangle &= b_1|d'\rangle + b_2|s'\rangle + b_3|b'\rangle, \\ |b\rangle &= c_1|d'\rangle + c_2|s'\rangle + c_3|b'\rangle. \end{aligned} \tag{1.67}$$

These equations represent a change of basis from the set of primed states to the unprimed set.

As u quarks couple only to d' and likewise for the other two generations, a_1, for example, must represent the strength of the coupling between u and d quarks, whilst a_2 represents the coupling between c and d quarks and so on. Introducing the notation V_{ij} for the coupling of the i-th and j-th quark flavours, equation (1.67) can be rewritten in terms of the 'CKM matrix', V_{CKM}:

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}^{\mathbf{T}}}_{V_{\mathrm{CKM}}} \cdot \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}. \tag{1.68}$$

Parametrising the CKM matrix

The CKM matrix was first introduced by Kobayashi and Maskawa [35], as an extension of the Cabibbo matrix [36] for two-generational mixing. The Chau-Keung representation [37] of $V_{\rm CKM}$ emphasises this three-dimensional rotation between the weak eigenbasis and the physical quark eigenstates,

$$V_{\rm CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.69}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\mathrm{e}^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\mathrm{e}^{-i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}\mathrm{e}^{-i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\mathrm{e}^{-i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\mathrm{e}^{-i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.70}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. θ_{ij} are the three Euler angles [38] of the basis rotation ($\theta_{12} = \theta_c$ is the Cabibbo angle [36]), and δ is the only complex phase which enters the matrix. The angles θ_{ij} may be defined to lie in the first quadrant, ensuring all c_{ij} and s_{ij} are positive. It is known empirically that $\theta_{12} \sim 13^\circ$, $\theta_{13} \sim 0.2^\circ$ and $\theta_{23} \sim 2^\circ$ [3], and thus that $s_{13} \ll s_{23} \ll s_{12} \ll 1$. Within the Standard Model, then, the quark couplings of the weak charged interaction are governed by just four independent parameters.

Under $\mathcal{CP}$ conjugation, the elements of $V_{\rm CKM}$ are complex-conjugated. Thus the possibility of $\mathcal{CP}$ violation in the couplings, responsible for all three forms of $\mathcal{CP}$ violation discussed in the previous section, is allowed within the Standard Model, with its magnitude parameterized by the '$\mathcal{CP}$-violating phase', δ.

Another useful parameterization of the CKM matrix is given by taking a power expansion in terms of $s_{12} = \sin\theta_c$, which is known, from the ratio of u and d quark masses, to be less than one [39] (in fact, $s_{12} \sim 0.2$ [3]). This allows the relative sizes of the matrix elements to be estimated, and is thus useful in simplifying theoretical calculations, where high-order terms may be neglected. This parameterization, developed by Wolfenstein [40], uses the four parameters

$$\lambda \equiv s_{12}, \quad A \equiv \frac{s_{23}}{s_{12}^2}, \quad \rho \equiv \frac{s_{13}\cos\delta}{s_{12}s_{23}} \quad \text{and} \quad \eta \equiv \frac{s_{13}\sin\delta}{s_{12}s_{23}}, \tag{1.71}$$

and, to order λ^5, gives

$$V_{\rm CKM}^{(5)} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{1}{2}A^2\lambda^5(1 - 2(\rho + i\eta)) & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 \\ A\lambda^3(1 - \overline{\rho} - i\overline{\eta}) & -A\lambda^2 + A\lambda^4(\frac{1}{2} - \rho - i\eta) & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6), \tag{1.72}$$

where $\overline{\rho} \equiv \rho(1 - \frac{1}{2}\lambda^2)$ and $\overline{\eta} \equiv \eta(1 - \frac{1}{2}\lambda^2)$. Within this parameterization, all $\mathcal{CP}$ violation arises due to the imaginary term η, which appears only for terms of order λ^3 and higher, illustrating the small degree of $\mathcal{CP}$ violation permitted within the Standard Model.

One natural test of the Standard Model is to independently measure the $V_{\rm CKM}$ elements, to determine whether they are consistent with the parameterizations presented above. As more data are gathered by experiments, ever higher orders of the Wolfenstein parameterization could be tested, until a discrepancy is found between independent measurements of the four independent parameters, or even between independent measurements of the same matrix element. However, a more sophisticated, and more powerful, way to test the Standard Model involves combining different measurements under the constraint of unitarity.

Unitarity triangles and γ

The CKM matrix ought to be unitary, since it represents a change of basis only. Insisting on unitarity, $V^\dagger V = VV^\dagger = \mathbb{I}$, leads to 18 equations linking the elements of the CKM matrix

$$V^\dagger V = \mathbb{I} \quad \Rightarrow \quad (V^\dagger)_{ik} V_{kj} = \delta_{ij} \tag{1.73}$$

$$\Rightarrow \quad V^*_{ki} V_{kj} = \delta_{ij}, \tag{1.74}$$

and

$$VV^\dagger = \mathbb{I} \quad \Rightarrow \quad V_{ik} (V^\dagger)_{kj} = \delta_{ij}, \tag{1.75}$$

$$\Rightarrow \quad V_{ik} V^*_{jk} = \delta_{ij}. \tag{1.76}$$

The cases for which the right-hand side vanishes give vector sums of three complex numbers, and thus may be represented as triangles in the complex plane. The six unique triangles obtained involve the sums

$$\begin{aligned}
V^*_{ud} V_{cd} + V^*_{us} V_{cs} + V^*_{ub} V_{cb} &= 0, \quad (cu) \\
V^*_{ud} V_{td} + V^*_{us} V_{ts} + V^*_{ub} V_{tb} &= 0, \quad (tu) \\
V^*_{cd} V_{td} + V^*_{cs} V_{ts} + V^*_{cb} V_{tb} &= 0, \quad (tc) \\
V_{ud} V^*_{us} + V_{cd} V^*_{cs} + V_{td} V^*_{ts} &= 0, \quad (sd) \\
V_{ud} V^*_{ub} + V_{cd} V^*_{cb} + V_{td} V^*_{tb} &= 0, \quad (bd) \\
V_{us} V^*_{ub} + V_{cs} V^*_{cb} + V_{ts} V^*_{tb} &= 0, \quad (bs)
\end{aligned} \tag{1.77}$$

The fact that there is a single complex phase within the CKM matrix results in each triangle having the same area, given by $J_{\mathcal{CP}}/2$, where $J_{\mathcal{CP}}$ is the Jarlskog invariant [41], given in the Chau-Keung representation by

$$J_{\mathcal{CP}} = c_{12} c_{23} c^2_{13} s_{12} s_{23} s_{13} \sin\delta. \tag{1.78}$$

However, due to the relative sizes of the matrix elements, four of the triangles are extremely 'squashed', with one very small internal angle. The two remaining triangles, (bd) and (tu), have three sides of similar lengths, $\mathcal{O}\lambda^3$. Normalising the two equations by $V_{cd} V^*_{cb} = A\lambda^3$ gives two triangles which differ only at $\mathcal{O}\lambda^5$. For the (bd) case, the triangle obtained is known as *the* unitarity triangle (UT). The

normalisation gives

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0. \tag{1.79}$$

The (bd) and (tu) triangles are displayed in Figure 1.2.

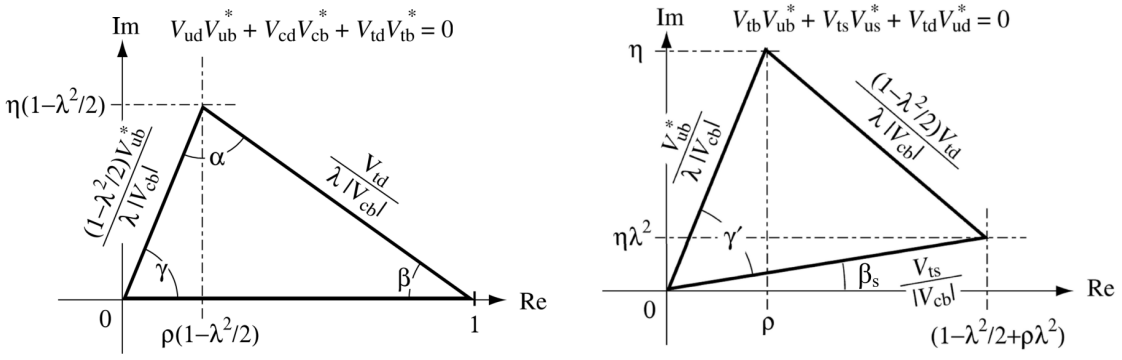


Figure 1.2.: The unitarity triangles, obtained from equation (1.77), for bd (left) and tu (right) [42].

The interior angles of the (bd) triangle are given by

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) = \arg\left(-\frac{1-\rho-i\eta}{\rho+i\eta}\right), \tag{1.80}$$

$$\beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) = \arg\left(-\frac{1}{1-\rho-i\eta}\right), \tag{1.81}$$

$$\gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) = \arg(\rho+i\eta) \tag{1.82}$$

whilst for the (tu) triangle the β and γ angles are shifted by $(O)\lambda^5$ terms, to give

$$\beta' \equiv \arg\left(-\frac{V_{ts}V_{us}^*}{V_{td}V_{ud}^*}\right) = \beta + \beta_s, \tag{1.83}$$

$$\gamma' \equiv \arg\left(-\frac{V_{tb}V_{ub}^*}{V_{ts}V_{us}^*}\right) = \gamma - \beta_s, \tag{1.84}$$

where

$$\beta_s \equiv \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \tag{1.85}$$

Overconstraining the Standard Model may now be achieved by overconstraining each of the unitarity triangles, by combining independent measurements of the different sides and internal angles. Since the normalisation fixes the lower side of the UT to lie on the real axis, with unit length, this task equates to overconstraining the position of the apex not lying on the real axis, $(1-\lambda^2/2)(\rho,\eta)$.

Relating γ to δ

Since they relate to elements of the CKM matrix, all parameters of the unitarity triangles are sensitive to $\mathcal{CP}$ violation. However, the interior angle γ is especially sensitive, owing to its near-equivalence to the $\mathcal{CP}$-violating phase δ.

Working in the Chau-Keung representation of the CKM matrix, the definition of γ yields

$$\gamma = \arg\left(\frac{-c_{12}c_{13}s_{13}e^{i\delta}}{\left(-s_{12}c_{23}-c_{12}s_{23}s_{13}e^{i\delta}\right)\left(s_{23}c_{13}\right)}\right), \tag{1.86}$$

which becomes

$$\gamma = \arg\left(-e^{i\delta}\right) - \arg\left(-s_{12}c_{23}-c_{12}s_{23}s_{13}e^{i\delta}\right). \tag{1.87}$$

Using the fact that $\arg(-1) = \pi$ gives

$$\gamma = \pi + \delta - \pi - \arg\left(s_{12}c_{23}+c_{12}s_{23}s_{13}e^{i\delta}\right). \tag{1.88}$$

Finally, noting the empirical relation above for the relative sizes of the s_{ij}, the second, complex, term within the parentheses is found to be significantly smaller than the first. Thus, to a very good approximation

$$\gamma \approx \delta. \tag{1.89}$$

As such, the precise determination of γ is not only crucial for the constraint it places on the apex of the unitarity triangle, but is relevant in its own right, for it quantifies the amount of CP violation permitted within the Standard Model.

1.4.1. Experimental status of CKM parameters

Measurements of CKM parameters are divided into two categories. 'Indirect' measurements give determinations of the side lengths of the unitarity triangles, *i.e.* by measuring the magnitudes of the different CKM matrix elements. By assuming a triangular construction, these measurements translate into a determination of the internal angles and the vertex postion; and thus the $\mathcal{CP}$-violating phase, δ.

'Direct' measurements are obtained from experimental determinations of the internal angles themselves. These give direct sensitivity to δ, and allow the assumptions of the unitarity triangles to be tested, by checking whether or not the triangles are closed (*i.e.* whether the interior angles sum to 180°). Any discrepancy between direct and indirect determinations of the UT parameters could indicate the presence of new physics contributions, which may enter as virtual particles within loop processes [43].

Two collaborations, UTFit [44] and CKMFitter [45], use statistical methods to combine independent measurements, in order to place constraints on the UTs. The current constraints provided by direct and indirect measurements, calculated by the CKMFitter group, are shown in Figure 1.3.

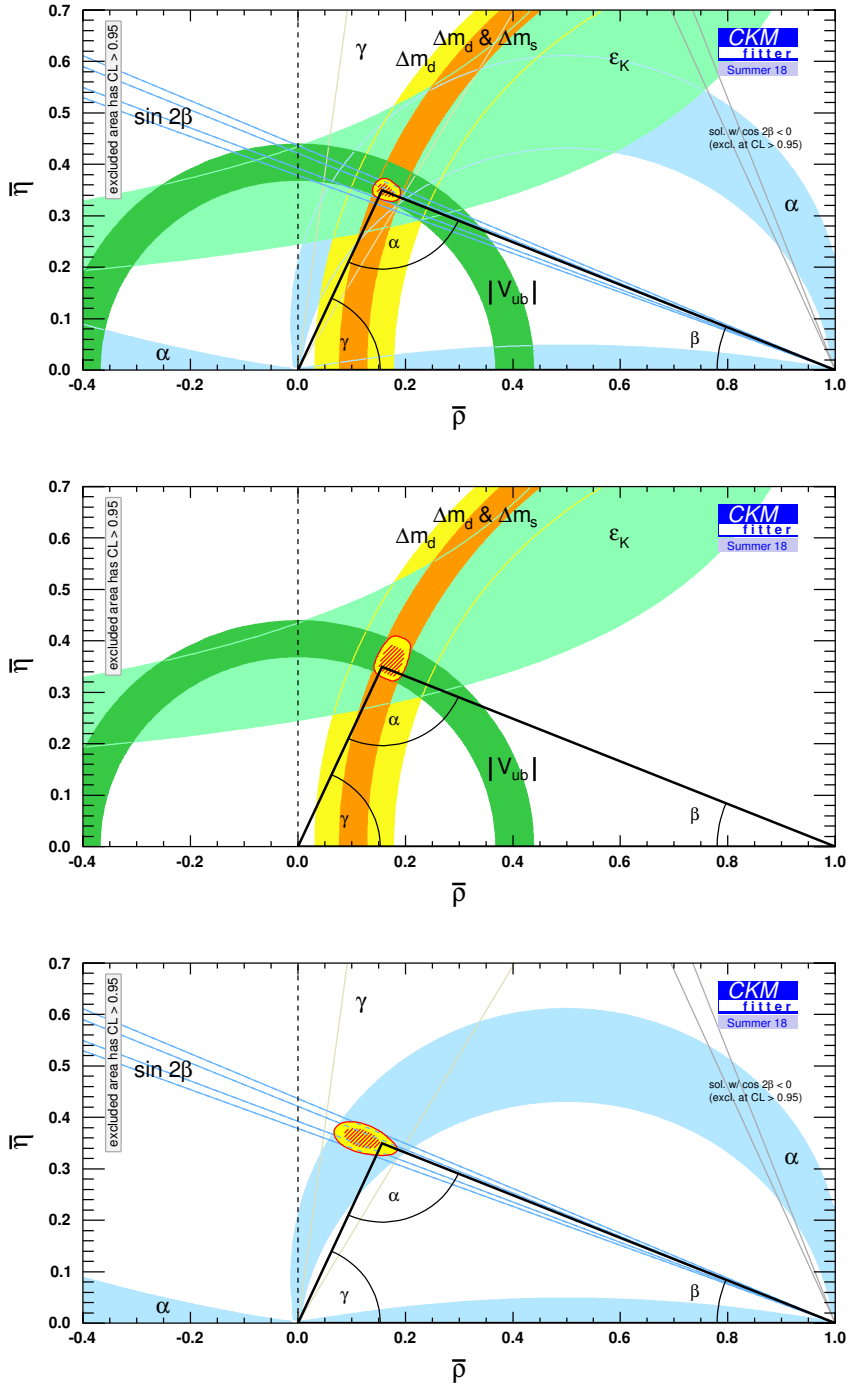


Figure 1.3.: Experimental constraints placed on the free apex of the unitarity triangle, calculated by the CKMFitter group [45]. The upper plot shows all constraints combined, whilst the middle (lower) plot shows just those constraints arising from indirect (direct) measurements. In all three cases, the triangle superimposed is the best-fit result for all constraints combined. The red hashed regions give 68% confidence intervals on the apex position, whilst the surrounding yellow regions give 95% confidence intervals.

The magnitudes of the CKM matrix elements $|V_{ud}|$, $|V_{us}|$, $|V_{ub}|$, $|V_{cd}|$, $|V_{cs}|$, $|V_{cb}|$ and $|V_{tb}|$ are determined from studying superallowed nuclear β-decays, and decays of pions, kaons, hyperons, D-mesons, B-mesons and top quarks (which decay too rapidly to hadronize into mesons). The other two magnitudes, $|V_{td}|$ and $|V_{ts}|$, are measured by their loop contributions in B-meson mixing. The global averages for these values, presented in reference [3], give

$$|V_{\text{CKM}}| = \begin{pmatrix} 0.97420 \pm 0.00021 & 0.2243 \pm 0.0005 & (3.94 \pm 0.36) \times 10^{-3} \\ 0.218 \pm 0.004 & 0.997 \pm 0.017 & (42.2 \pm 0.8) \times 10^{-3} \\ (8.1 \pm 0.5) \times 10^{-3} & (39.4 \pm 2.3) \times 10^{-3} & 1.019 \pm 0.025 \end{pmatrix} \tag{1.90}$$

Since the interior angles of the unitarity triangle give sensitivity to the $\mathcal{CP}$-violating phase, δ, it is no surprise that they can be measured in $\mathcal{CP}$-violating processes. Angle β may be directly extracted from the study of time-dependent $\mathcal{CP}$ asymmetries, for B^0 and $\overline{B^0}$ decaying to any common $\mathcal{CP}$ eigenstate. The current world average combining all such measurements gives [3]

$$\sin(2\beta) = 0.691 \pm 0.017. \tag{1.91}$$

For angle α, time-dependent asymmetries extracted from decays involving $b \to du\overline{u}$ processes may be used. However, such analyses are complicated by the fact that loop (specifically, $b \to d$ penguin [3]) diagrams contribute at the same order in λ as the desired tree-level decays. The current world average for this angle is [3]

$$\alpha = (84.5^{+5.9}_{-5.2})^\circ. \tag{1.92}$$

Angle γ represents a 'benchmark' Standard Model measurement, since it can be measured in tree-level processes with negligible loop contributions, and therefore little scope for new physics [43]. However, due to the CKM elements involved, decays sensitive to γ have small branching fractions ($\mathcal{O}(10^{-6})$ [3]), making it difficult to measure precisely. Prior to the contributions of LHCb, γ was the least well-constrained angle of the UT, with a 2012 world average of $\gamma = (68^{+10}_{-11})^\circ$ [46]. Since then, the uncertainty has more than halved, giving the current world average [3]

$$\gamma = (73.5^{+4.2}_{-5.1})^\circ. \tag{1.93}$$

Decays of the form $B \to D^{(*)}K^{(*)}$ give sensitivity to γ, since it parameterizes the relative weak phase between the contributing D^0 and $\overline{D^0}$ components. The 'GLW' [47–49] approach detects the D-meson in $\mathcal{CP}$ eigenstates, where these two contributions interfere, and extracts an asymmetry which is thus sensitive to γ. However, the degree of sensitivity is limited by the relative sizes of the two interfering components, being greatest when they are of comparable magnitude. For the mode with the highest branching fraction, $B^- \to DK^-$, the amplitude involving $\overline{D^0}$ is colour- and CKM-suppressed, whilst the amplitude involving $\overline{D^0}$ is colour- and CKM-favoured.

To increase sensitivity, the 'ADS' [50,51] method detects the D-meson in a chosen final state f (*e.g.* $K^\pm\pi^\mp$), such that the process $D^0 \to f$ is doubly-Cabibbo-suppressed, whilst $\overline{D^0} \to f$ is Cabibbo-

favoured. This gives two interfering amplitudes of comparable magnitude, yielding the maximum sensitivity possible. Once again asymmetries may be defined in order to extract γ. However, this method is limited by the small branching fractions of the D-meson decays involved.

The 'GGSZ' method [52], uses three body final states f that have large branching fractions for both D^0 and $\overline{D^0}$, *e.g.* $K^0_S\pi^+\pi^-$. The $\mathcal{CP}$ angle γ is extracted in an analysis of the full resonant structure of the interfering components in the phase space of f (the 'Dalitz plane'). An analysis of this type is the subject of the present thesis. The method is described more fully in the following section.

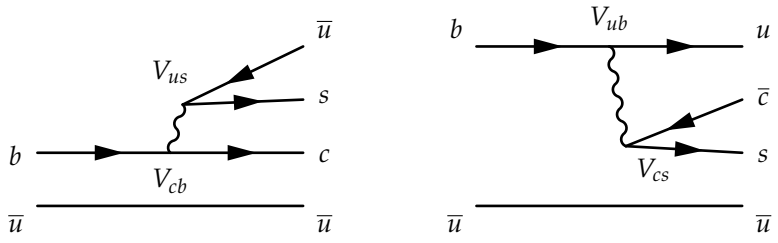
1.5. Measuring $\mathcal{CP}$ violation in $B \to D^{(*)}h^{(*)}$ decays

The present thesis concerns a combined GGSZ analysis to extract γ in the following modes:

- $B^- \to DK^-$, with $D \to K^0_S\pi^+\pi^-$;
- $B^0 \to DK^{*0}$, with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$;
- $B^- \to D^*K^-$, with $D^* \to D(\gamma/\pi^0)$ and $D \to K^0_S\pi^+\pi^-$; and
- $B^- \to DK^{*-}$, with $K^{*-} \to \overline{K^0_S}\pi^-$ and $D \to K^0_S\pi^+\pi^-$.

In every case, the neutral D-meson is detected in the final state ($K^0_S\pi^+\pi^-$). In addition, the corresponding $B^- \to D^{(*)}\pi^-$ modes provide control samples, used to reduce uncertainties in the mass spectrum fits performed (described in Chapter 5).

Feynman diagrams for the charged B-meson decays considered are shown in Figure 1.4, whilst those considered for the neutral B-meson are shown in Figure 1.5.



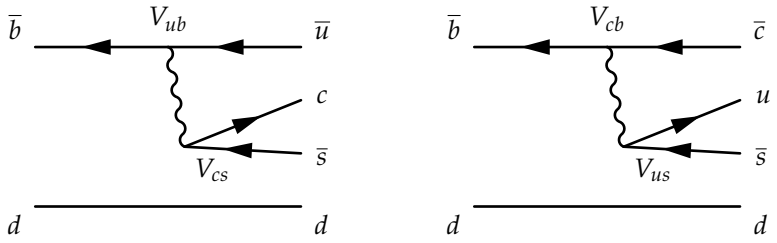
(a) B^- decay via D^0 **(b)** B^- decay via $\overline{D^0}$

Figure 1.4.: Decays considered for the B^- meson. The left-hand diagram corresponds to decay via a D^0-meson ($c\overline{u}$), whilst the right corresponds to decay via a $\overline{D^0}$-meson ($\overline{c}u$).

Contributing amplitudes

In each mode, m, considered, the two contributing amplitudes may be written as

$$A^m_{b\to c} = A^m_B e^{i\delta^m_c} \tag{1.94}$$



(a) B^0 decay via D^0 **(b)** B^0 decay via $\overline{D^0}$

Figure 1.5.: Decays considered for the B^0 meson. The left-hand diagram corresponds to decay via a D^0-meson ($c\bar{u}$), whilst the right corresponds to decay via a $\overline{D^0}$-meson ($\bar{c}u$).

for the diagram involving a $b \to c$ transition, and

$$A^m_{b\to u} = A^m_B r^m_B e^{\pm i\gamma} e^{i\delta^m_u} \tag{1.95}$$

for the diagram involving a $b \to u$ transition. The ratio of the amplitudes is then given by

$$\frac{A^m_{b\to u}}{A^m_{b\to c}} = r^m_B e^{i(\delta^m_B \pm \gamma)}. \tag{1.96}$$

The relative weak phase between the two processes is given to a very good approximation by γ (see Appendix A), due to the CKM matrix elements which appear in the two diagrams. The relative strong phase, δ^m_B, and magnitude r^m_B, between the interfering decays result from the different strong interactions that occur in the intermediate state to produce the hadronization into D^0 or $\overline{D^0}$. They take different values for the different modes considered, since the hadronization interactions will depend on the spin state, as well as the quark composition, of the mesons involved; so that modes $B^\pm \to DK^\pm$, $B^\pm \to DK^{*\pm}$ and $B^\pm \to D^*K^\pm$ are not equivalent. The $\mathcal{CP}$-violating weak phase γ, however, is common to all modes, thus motivating their combined study. It is useful to note that, for each mode studied, the flavour of the original B-meson may be distinguished (or 'tagged') using the charge of the kaon or of the daughter mesons in the decay of the K^*-meson.

For the decay modes $B^0 \to DK^{*0}$ and $B^\pm \to DK^{*\pm}$, the K^*-meson is detected in a two-body final state, by placing selection requirements on the mass reconstructed from its daughters. However, the K^*-meson has a large width. This means that any requirements encompassing a large fraction of the K^* resonance are also likely to select non-negligible contributions from other resonances. As such, the relative strong phase and magnitude observed are not those which would result from pure $B \to DK^*$ decays alone, but are instead given by the effective variables r^m_s and δ^m_s, obtained by integrating over the selected region of the $B^0 \to DK^+\pi^-$ or $B^\pm \to DK^0_S\pi^\pm$ phase space:

$$(r^m_s)^2 = \frac{\Gamma(B \to D^0 K\pi)}{\Gamma(B \to \overline{D^0} K\pi)} = \frac{\int dp (A^m_{b\to u}(p))^2}{\int dp (A^m_{b\to c}(p))^2}, \tag{1.97}$$

$$\kappa^m e^{i\delta^m_s} = \frac{\int dp A^m_{b\to c}(p) A^m_{b\to u}(p) e^{i\delta^m(p)}}{\sqrt{\int dp (A^m_{b\to c}(p))^2 \int dp (A^m_{b\to u}(p))^2}}, \tag{1.98}$$

where the variable p indicates the position in the decay phase space and the real parameter $0 < \kappa^m < 1$ parameterizes the fraction of the integrated phase space that is occupied by the K^* resonance. This same notation can be used to describe two-body B decays; where κ is set equal to unity, and r_s and δ_s simplify to become r_B and δ_B.

Interference in the neutral D-meson system

The eigenstates of the neutral D-meson system produced in the B-meson decay are then given, for each mode, by

$$\begin{aligned} |D^m_-\rangle &\sim \left|D^0\right\rangle + z^m_- \left|\overline{D^0}\right\rangle, \\ |D^m_+\rangle &\sim \left|\overline{D^0}\right\rangle + z^m_+ \left|D^0\right\rangle, \end{aligned} \tag{1.99}$$

where the subscript $(\pm)$ denotes the flavour eigenstate of the progenitor B-meson[2], and the relative phase and magnitude arising from the contributing B-meson decay diagrams are accounted for by the 'Cartesian parameters', $z^m_\pm$,

$$z^m_\pm = x^m_\pm + i\, y^m_\pm = r^m\, e^{i(\delta^m \pm \gamma)}. \tag{1.100}$$

The amplitude for the full $B \to D^{(*)}h^{(*)}, D \to f$ decay is then given by

$$A^m_- \sim A_f + z^m_- \overline{A}_f, \tag{1.101}$$
$$A^m_+ \sim \overline{A}_f + z^m_+ A_f, \tag{1.102}$$

where $A_f = \langle f\,|T|D^0\rangle$, $\overline{A}_f = \langle f|T|\overline{D^0}\rangle$ and $A^m_\pm = \langle f\,|T|D^m_\pm\rangle$.

Note that, in addition to the weak phase, γ, the strong phases δ^m and relative magnitudes r^m in each mode are also unknowns. Therefore, any extraction of γ must equally be an extraction of these parameters. However, a combined analysis of several modes can exploit the mode-invariance of γ to reduce the number of free parameters, giving a more accurate determination of γ.

Combined approach

As pointed out by Garra-Ticó [53], information from multiple modes may be combined efficiently by noticing that, for all modes, m,

$$\frac{z^m_+}{z^m_-} = e^{2i\gamma} \quad \Rightarrow \quad \gamma = \frac{1}{2}\arg\left(\frac{z^m_+}{z^m_-}\right). \tag{1.103}$$

This means that the z^m_+ and z^m_- coefficients of any mode may be multiplied by any complex factor, ξ, without altering the information on γ contained.

[2] B^+, B^- in the case of charged decays, $B^0, \overline{B^0}$ in the neutral case. These states may be differentiated ('tagged') by the charge of the K^* daughter in the B-decay.

In particular, it is useful to relate all coefficients back to those of a single mode, *e.g.* $B^\pm \to DK^\pm$:

$$z_\pm^m = \xi_m z_\pm^{\mathrm{DK}}, \tag{1.104}$$

where

$$\xi_m = \frac{r^m e^{i\delta^m}}{r^{\mathrm{DK}} e^{i\delta^{\mathrm{DK}}}}. \tag{1.105}$$

To give slightly neater notation, it is useful to abbreviate $z_\pm = z_\pm^{\mathrm{DK}}$, so that $z_\pm^m = \xi_m z_\pm$.

This substitution reduces the number of free parameters from $4N$ to $2(N+2)$, where N gives the number of modes included.[3] Therefore, this method is expected to give a more precise determination of γ than a statistical combination of separate results from each individual mode, when combining $N > 2$ modes.

Final state phase space

In the GGSZ approach [52], the interference parameterized by γ is studied in the phase space of the $D \to K_S^0\pi^+\pi^-$ decay. This phase space may be parameterized by the momentum configuration of the final state particles, (p_1, p_2, p_3), where

$$P^\mu(K_S^0) = P_1^\mu, \quad P^\mu(\pi^-) = P_2^\mu, \quad \text{and} \quad P^\mu(\pi^+) = P_3^\mu. \tag{1.106}$$

Following the usual definition for the combined squared invariant mass of two particles, it is useful to define

$$s_{ij} = |P_i^\mu + P_j^\mu|^2. \tag{1.107}$$

Any point in the phase space may then be uniquely specifed by any two of its combined squared mass coordinates, (s_i, s_j). Each (s_{ij}, s_{kj}) coordinate system represents a 'Dalitz plane', whilst a frequency histogram of observed events in such a plane is known as a 'Dalitz plot'. It is useful to work in this plane since, as noted by Richard Dalitz [54], for a three-body final state parameterized in these coordinates, the Lorentz-invariant phase space is uniform. This means that any non-uniformity observed in the Dalitz plot immediately indicates a Physics effect arising from the amplitudes involved. No unfolding is necessary to separate out Physics and phase-space effects. This makes the Dalitz coordinate system the natural basis for plotting and analysing three-body decays. A common feature of Dalitz plots is the presence of 'resonant substructure', indicating that the three-body final state arises from a series of two-body decays, with a short-lived resonance propagating between the two decay vertices. A typical Dalitz plot is shown in Figure 1.6.

[3] *I.e.* from $\{x_\pm^m, y_\pm^m\}$ to $\{x_\pm^{\mathrm{DK}}, y_\pm^{\mathrm{DK}}, \mathrm{Re}(\xi_m), \mathrm{Im}(\xi_m)\}$.

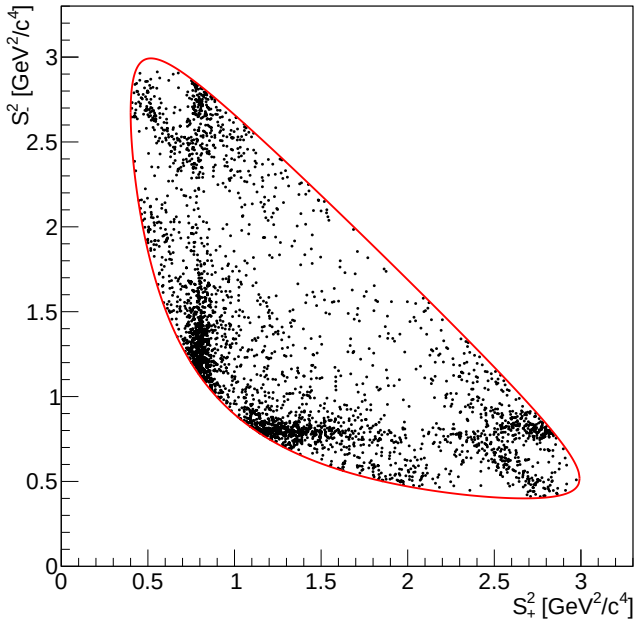


Figure 1.6.: Dalitz plot of $B^\pm \to DK^\pm, D \to K^0_S\pi^+\pi^-$ events selected from the full 2011 and 2012 combined LHCb dataset. Events are kinematically forbidden from occurring outside the closed red curve. Resonant structure is evident in the non-uniformity of the observed distribution, which would otherwise be flat, since the Lorentz-invariant phase space is uniform over the Dalitz plane [54].

Both the magnitude and phase of the D-decay amplitude may depend on position within the Dalitz plane. In this basis, the amplitude for a D^0 decay may be written as

$$A_f(s_{12}, s_{13}) \equiv A(D^0 \to K^0_S(p_1)\pi^-(p_2)\pi^+(p_3)), \tag{1.108}$$
$$= A_{12,13}e^{i\delta_{12,13}}. \tag{1.109}$$

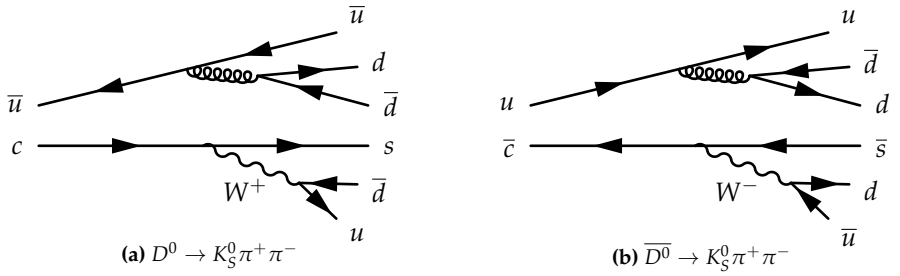
Assuming $\mathcal{CP}$ violation within the neutral D-meson system may be neglected compared to the direct $\mathcal{CP}$ violation in the B-decay, the amplitude for a $\overline{D^0}$ decay is then given by

$$\overline{A}_f(s_{12}, s_{13}) \equiv A(\overline{D^0} \to K^0_S(p_1)\pi^-(p_2)\pi^+(p_3)), \tag{1.110}$$
$$= A(D^0 \to K^0_S(p_1)\pi^-(p_3)\pi^+(p_2)), \tag{1.111}$$
$$= A_f(s_{13}, s_{12}), \tag{1.112}$$
$$= A_{13,12}e^{i\delta_{13,12}}, \tag{1.113}$$

and the equations 1.101 and 1.102 become

$$A^m_- \sim A_{12,13}e^{i\delta_{12,13}} + z^m_- A_{13,12}e^{i\delta_{13,12}}, \tag{1.114}$$
$$A^m_+ \sim A_{13,12}e^{i\delta_{13,12}} + z^m_+ A_{12,13}e^{i\delta_{12,13}}. \tag{1.115}$$

From inspection of the relevant Feynman diagrams (Figure 1.7), it is clear that there is no relative weak phase or magnitude between the decays $D^0 \to K^0_S\pi^+\pi^-$ and $\overline{D^0} \to K^0_S\pi^+\pi^-$, since the same CKM elements are involved in each case. As any global phase will give no effect on the decay rates observed (it disappears in the square modulus), and the absolute value of the rate is not important, only its variation over the Dalitz plane, then all weak effects can be neglected in these decays. In this case, the functions $A_{ij,kj}$ and $\delta_{ij,kj}$ give only the magnitudes and phases arising from the strong interaction.



(a) $D^0 \to K^0_S\pi^+\pi^-$ **(b)** $\overline{D^0} \to K^0_S\pi^+\pi^-$

Figure 1.7.: Feynman diagrams for decay of the neutral D^0 (left) and $\overline{D^0}$ (right) mesons to the final state mesons $K_S\ \pi^+\pi^-$. In each diagram, neighbouring pairs of quarks and antiquarks are hadronically bound.

Determination of $\mathcal{CP}$ parameters

To give expressions which may be brought to bear on data, allowing the extraction of the $\mathcal{CP}$ sensitive parameters, $z_\pm$ and ξ_m, Fermi's 'golden rule' for the differential rate is used

$$\mathrm{d}\Gamma = \int |A|^2\, \mathrm{d}\mathcal{P}. \tag{1.116}$$

Combining this with the useful relation $|z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2|z_1||z_2|\mathrm{Re}(z_1^* z_2)$ (derived in Appendix B), yields

$$\mathrm{d}\Gamma^m_- \sim \left|A_f\right|^2 + |z^m_-|^2 \left|\bar{A}_f\right|^2 + 2\kappa_m\, \mathrm{Re}\left(z^m_-\, A^\star_f \bar{A}_f\right), \tag{1.117}$$

$$\mathrm{d}\Gamma^m_+ \sim \left|\bar{A}_f\right|^2 + |z^m_+|^2 \left|A_f\right|^2 + 2\kappa_m\, \mathrm{Re}\left(z^m_+\, \bar{A}^\star_f A_f\right). \tag{1.118}$$

The authors of reference [52] present two ways of proceeding. The amplitude of the three-body D^0-meson decay, $D^0 \to K^0_S\pi^+\pi^-$ is presently beyond the limits of QCD calculation, meaning that there is no fundamental prescription for the functions $A_{12,13}$ and $\delta_{12,13}$. However, the empirical observation of resonant structure in the Dalitz plane for many decays, and the success of resonance-based phenomenology in describing scattering and decay processes over many decades, motivates the use of a resonance model to describe these functions. Any such model must of course be calibrated against a suitably clean data set before it can be used to extract new information from other decays. Thankfully, the BaBar collaboration has developed a sophisticated model of $D^0 \to K^0_S\pi^+\pi^-$ decays, calibrated against 468.5 fb^{-1} of data gathered in e^+e^- collisions [55]. The nominal approach taken within this thesis uses the BaBar model to fit the observed distribution of events in phase space, and thus to extract the $\mathcal{CP}$

violating parameters. Since the resonance model used depends on the choice of which resonances to include, and how to parameterize them, this approach is said to be **model-dependent**. The theory underlying resonance models, and the details of the specific model used within this thesis, are presented in Chapter 2.

An alternative, **model-independent** approach is also considered, which allows the magnitude and phase of $A_f(s_{12}, s_{13})$ to remain unknown. In this case, it is useful to re-express equations 1.117 and 1.118 in the form

$$\mathrm{d}\Gamma^m_- \sim |A_f|^2 + |z^m_-|^2|\overline{A}_f|^2 + 2\kappa_m\,|z^m_-||A_f||\overline{A}_f|\cos\Big(\arg(z^m_-) + \arg(\overline{A}_f) - \arg(A_f)\Big), \tag{1.119}$$

$$\mathrm{d}\Gamma^m_+ \sim |\overline{A}_f|^2 + |z^m_+|^2|A_f|^2 + 2\kappa_m\,|z^m_+||\overline{A}_f||A_f|\cos\Big(\arg(z^m_+) + \arg(A_f) - \arg(\overline{A}_f)\Big). \tag{1.120}$$

Substituting from equations 1.109 and 1.113, and using the relation $\cos(A \pm B) = \cos(A)\cos(B) \mp \sin(A)\sin(B)$, gives

$$\mathrm{d}\Gamma^m_- \sim A^2_{12,13} + |z^m_-|^2 A^2_{13,12} + 2\kappa_m\,|z^m_-|A_{12,13}A_{13,12} \Big\{\cos\Big(\arg(z^m_-)\Big)\cos\Big(\delta_{13,12} - \delta_{12,13}\Big) - \sin\Big(\arg(z^m_-)\Big)\sin\Big(\delta_{13,12} - \delta_{12,13}\Big)\Big\}, \tag{1.121}$$

$$\mathrm{d}\Gamma^m_+ \sim A^2_{13,12} + |z^m_+|^2 A^2_{12,13} + 2\kappa_m\,|z^m_+|A_{12,13}A_{13,12} \Big\{\cos\Big(\arg(z^m_+)\Big)\cos\Big(\delta_{12,13} - \delta_{13,12}\Big) - \sin\Big(\arg(z^m_+)\Big)\sin\Big(\delta_{12,13} - \delta_{13,12}\Big)\Big\}. \tag{1.122}$$

Integrating over some region, R of the Dalitz plane, it is useful to define the values

$$c_R \equiv \int_R \big[A_{12,13}A_{13,12}\cos(\delta_{12,13} - \delta_{13,12})\big]\,\mathrm{d}\mathcal{P}, \tag{1.123}$$

$$s_R \equiv \int_R \big[A_{12,13}A_{13,12}\sin(\delta_{12,13} - \delta_{13,12})\big]\,\mathrm{d}\mathcal{P}, \tag{1.124}$$

$$T_R \equiv \int_R \big[A^2_{12,13}\big]\,\mathrm{d}\mathcal{P}, \tag{1.125}$$

$$\tilde{T}_R \equiv \int_R \big[A^2_{13,12}\big]\,\mathrm{d}\mathcal{P}. \tag{1.126}$$

The symmetry of the terms c_R and s_R motivates the splitting of the Dalitz plane into bins arranged symmetrically about the line $s_{12} = s_{13}$. Let bin i lie above this symmetry axis, centred at the point (s_{12}, s_{13}), whilst bin $\tilde{i}$ is its conjugate below the symmetry axis, centred at the point (s_{13}, s_{12}). Going from bin i to bin $\tilde{i}$ (or vice-versa) therefore corresponds to interchanging the subscripted indices $12 \rightleftharpoons 13$.

This ensures the validity of the simplifying relations

$$c_{\tilde{i}} = c_i, \quad s_{\tilde{i}} = -s_i, \tag{1.127}$$

$$\tilde{T}_{\tilde{i}} = T_i, \quad \tilde{T}_i = T_{\tilde{i}}. \tag{1.128}$$

These give the following expressions for the partial widths for two conjugate bins in the Dalitz plane:

$$\Gamma^m_{-,i} \sim T_i + |z^m_-|^2 T_{\tilde{i}} + 2\kappa_m |z^m_-| \Big\{ c_i \cos\Big(\arg(z^m_-)\Big) - s_i \sin\Big(\arg(z^m_-)\Big) \Big\}, \tag{1.129}$$

$$\Gamma^m_{-,\tilde{i}} \sim T_{\tilde{i}} + |z^m_-|^2 T_i + 2\kappa_m |z^m_-| \Big\{ c_i \cos\Big(\arg(z^m_-)\Big) + s_i \sin\Big(\arg(z^m_-)\Big) \Big\}, \tag{1.130}$$

$$\Gamma^m_{+,i} \sim T_{\tilde{i}} + |z^m_+|^2 T_i + 2\kappa_m |z^m_+| \Big\{ c_i \cos\Big(\arg(z^m_+)\Big) + s_i \sin\Big(\arg(z^m_+)\Big) \Big\}, \tag{1.131}$$

$$\Gamma^m_{+,\tilde{i}} \sim T_i + |z^m_+|^2 T_{\tilde{i}} + 2\kappa_m |z^m_+| \Big\{ c_i \cos\Big(\arg(z^m_+)\Big) - s_i \sin\Big(\arg(z^m_+)\Big) \Big\}. \tag{1.132}$$

If k gives the number of bins above the line $s_{12} = s_{13}$, then the expressions (1.129), (1.130), (1.131) and (1.132) describe $4k$ independent equations. Thus by dividing the Dalitz plane into a sufficient number of bins it is possible to determine all of the $\mathcal{CP}$ parameters, $z_\pm$ and ξ_m. The values T_i and $T_{\tilde{i}}$ may be found from pure flavour-tagged samples of $D \to K^0_S \pi^+ \pi^-$, *e.g.* from a semi-leptonic mode such as $B^0 \to D^{*\pm} \mu^\mp \nu_\mu$, $\quad D^{*\pm} \to D\pi^\pm$, where the flavour of the neutral D-meson is tagged by the charge of the pion produced in the $D^{*\pm}$-meson decay.

Whilst this method benefits from model-independence, its performance is limited by the available statistics; in practise, the number of signal events, Γ_i, observed in each bin of the Dalitz plane is obtained by performing a fit to the distribution of the reconstructed B-meson mass (as described in Chapter 5). As such, there must be sufficient population within each bin to allow such a fit to be reliable. To reduce the resulting uncertainty it is typical to use c_i and s_i values (and the corresponding binning scheme) obtained from past CLEO measurements [56]. This approach is followed for all model-independent fits performed within this thesis.

1.6. Strategy of the simultaneous analysis

The simultaneous analysis combines several modes to maximise the precision with which $\mathcal{CP}$ parameter γ can ultimately be extracted. It uses a GGSZ approach, analysing the Dalitz plane of $D \to K^0_S \pi^+ \pi^-$ decays. Both a model-dependent and model-independent method are considered for the determination of the $\mathcal{CP}$ parameters from the observed distributions in the Dalitz plane. The stages of the analysis may be summarised as follows:

Selection: Candidate events are selected for each mode using similar criteria. In addition to placing requirements on individual variables, multivariate classifiers (MVAs) are trained to allow better discrimination between signal and background events. Separate MVAs are trained for each mode, though the input variables used in each case are the same.

Mass fitting: Even after selecting the most promising candidates for analysis, non-negligible background contributions remain. In order to determine the yields of selected signal and background events, the reconstructed B-meson mass spectrum is fitted, with separate fits performed for each

mode. These yields are needed as inputs for the Dalitz analysis, since the expected distributions within the Dalitz plane differ for different background components. Failing to account for this pollution would reduce the precision with which the $\mathcal{CP}$ parameters could be extracted.

Efficiency determination: Before the Dalitz plot can be analysed, the variation of the selection efficiency over the Dalitz plane must be determined. The extraction of the $\mathcal{CP}$ parameters relies on measuring asymmetries in the Dalitz plane, and so the effect of efficiency variations must be separated out in the final analysis. Separate selection efficiencies are obtained for each mode; one for each selection that is performed.

Combined $\mathcal{CP}$ fit: Candidates selected in all modes are combined in a single Dalitz analysis, along with their determined efficiencies. Different distributions are expected for events of each signal and background component, but in each case the expected distribution may be related back to the expressions presented in this chapter, equations 1.117 and 1.118. Both the model-dependent and model-independent method are considered.

Interpretation of results: Finally, the extraction of the $\mathcal{CP}$ phase γ, from measurements of the Cartesian parameters, $z_{\pm}^{m}$, makes use of a likelihood-maximising approach. The range of allowed (γ, r^m, δ^m) values are scanned to find those which maximise the likelihood of the observed Cartesians, before confidence intervals are extracted from the full likelihood profile obtained.

The simultaneous approach presents several benefits relative to measurements made using individual decay modes, in addition to the reduction in the number of fit parameters discussed in the previous section:

- The ability to correctly model signal decays that cross-feed between categories in the $\mathcal{CP}$ fit.[4] In a standalone analysis, such decays must be treated as a source of background, and reduce the precision of the measurement produced;
- The ability to make like-for-like comparisons between model-dependent and model-independent fits, since the exact same selected data samples, estimated efficiencies and mass fit results may be used as inputs in each case;
- The ability to extract correlations between fitted parameters coming from different decay modes. This information is inaccessible when standalone results are combined in a statistical meta-analysis;
- A coherent framework in which it is relatively simple to ensure that the different analysis stages are performed consistently between modes, including consistent treatment of systematic uncertainties.

[4] *I.e.* signal decays of one γ-sensitive mode that may satisfy the selection criteria of another.

2. Modelling $D^0 \to K_S\pi^+\pi^-$ decays

Section 1.5 outlined a model-dependent approach to extracting CKM angle γ from decays of the form $B \to D^{(*)}K^{(*)}$, with $D \to K^0_S\pi^+\pi^-$. The function $\mathcal{P}$ must be fitted to the D-decay phase space

$$\mathcal{P} = |A|^2 + |z|^2|\overline{A}|^2 + 2\kappa\text{Re}(zA^*\overline{A}),$$

where A is the amplitude for the decay $D^0 \to K^0_S\pi^+\pi^-$.

However, the theoretical calculation of three-body hadronic decay amplitudes, such as A, is currently out of reach (though progress is being made in this direction [57]). Instead, a simplified model is built and calibrated from a large, pure sample of $D^0 \to K^0_S\pi^+\pi^-$ decays. Section 2.1 introduces the effective theory of resonances; the theoretical framework within which such a model may be constructed. Section 2.2 and Section 2.3 apply this theory, with successively greater degrees of sophistication, to the problem of two-body resonant scattering, and show that this process is theoretically equivalent to a three-body decay. Finally, Section 2.4 gives details of the specific model implemented within the present analysis.

2.1. Effective theory of resonances

A model for the decay $D^0 \to K^0_S\pi^+\pi^-$ may be obtained under the assumption that it proceeds via a short-lived intermediate resonant state, r [58]

$$D \to rc, \tag{2.1}$$
$$r \to ab. \tag{2.2}$$

Equations 2.1 and 2.2 introduce the naming convention that will be followed for the rest of this chapter: a and b denote the final state particles produced by the resonance; and c labels the 'spectator' particle, produced directly by the D-meson decay. This picture may be described using an effective theory, which parameterizes the D and r decays as 'vertices', containing all short-range interactions, and a 'propagator' term to describe the resonant particle, as shown in Figure 2.1. Since this diagram is not a Feynman diagram - the interacting particles are mesons, not quarks - the couplings at the vertices are unknown, and hence any such resonance model will have parameters for which there is no *a priori* value. Therefore it is essential to fit the model to a pure $D^0 \to K^0_S\pi^+\pi^-$ data sample, to calibrate these parameters.

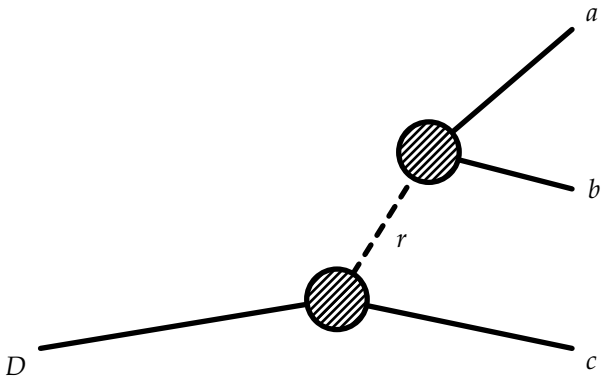


Figure 2.1.: Effective diagram for three-body D-meson decay.

2.2. Isobar model

The Isobar model [59–61] is the simplest type that may be constructed to describe the contributions of multiple resonances

$$\mathcal{M} = \sum_r a_r \mathcal{M}_r. \tag{2.3}$$

The different resonances sum at the amplitude level, with relative weights given by $|a_r|$ and relative phases $\delta_{ij} = \arg(a_i) - \arg(a_j)$.

Within this model, the amplitude for the decay proceeding via a given resonance r is

$$\mathcal{M}_r(L, m_{ab}, m_{ac}) = \sum_\lambda \langle ab|r_\lambda\rangle T_r(m_{ab}) \langle cr_\lambda|D^0\rangle, \tag{2.4}$$

where the operator for the interaction is defined as

$$\hat{\Delta}_{r,\lambda} = |r_\lambda\rangle T_r(m_{ab}) \langle cr_\lambda|. \tag{2.5}$$

L gives the spin of the resonance, and thus also the orbital angular momentum between r and c, and between a and b, since the progenitor D-meson and all the final state particles are spin-zero. The combined invariant mass of the 'resonant pair'[1] is given by m_{ab},

$$m_{ab}^2 \equiv |P_a^\mu + P_b^\mu|^2. \tag{2.6}$$

Equation 2.4 includes the usual sum over initial helicity states and average over final helicity states [62]. Since the resonance is the only particle in the decay chain which may have non-zero spin, this reduces to a sum over the helicity states, λ, of the resonance.

[1] *I.e.* the decay products of the resonance.

It is typical to rewrite equation (2.4), factorising it as

$$\mathcal{M}_r(L, m_{ab}, m_{ac}) = B_L^D(p, p_r)B_L^r(q, q_r)Z_L(m_{ab}^2, m_{ac}^2)T_r(m_{ab}), \tag{2.7}$$

where $\vec{p}$ represents the momentum of c (or r) in the rest frame of the progenitor D-meson, and $\vec{q}$ represents the momentum of a (or b) in the rest frame of the ab system.

One may then define

$$p \equiv |\vec{p}|, \quad q \equiv |\vec{q}|,$$

which can themselves be expressed as functions of m_{ab},

$$q = \frac{\lambda^{\frac{1}{2}}(m_{ab}^2, m_a^2, m_b^2)}{2m_{ab}} = \frac{\sqrt{[m_{ab}^2 - (m_a + m_b)^2][m_{ab}^2 - (m_a - m_b)^2]}}{2m_{ab}}, \tag{2.8}$$

$$p = \frac{\lambda^{\frac{1}{2}}(m_{ab}^2, m_D^2, m_c^2)}{2m_{ab}} = \frac{\sqrt{[m_{ab}^2 - (m_D + m_b)^2][m_{ab}^2 - (m_D - m_b)^2]}}{2m_{ab}}, \tag{2.9}$$

where

$$\lambda(x, y, z) \equiv x^2 + y^2 + z^2 - 2xy - 2xz - 2yz. \tag{2.10}$$

One may also define p_r and q_r, the values taken by p and q when $m_{ab} = m_r$ so that the resonance is produced on-shell.

The function Z_L parameterizes the angular distribution of the decay products, and will be seen to arise from the spin-sum rules for the effective theory defined in equation (2.4). B_L are centrifugal barrier factors for each 'vertex' of the decay. These account for the fact that angular momentum must be conserved. T_r is a dynamical function that acts as the 'propagator factor' for the resonance, quantifying how energetically favourable it is for a given resonance r to be produced. The following sections consider these functions in turn.

2.2.1. Centrifugal barrier factors B_L

Since D, K_S, and $\pi^\pm$ mesons have zero spin, the production and decay of any resonance with non-zero spin implies that some orbital angular momentum is required between the (c, r) pair, and also between the (a, b) pair, such that the total angular momentum of zero is conserved. This becomes harder to achieve the smaller the impact parameter between the orbiting pair, and the smaller their linear momenta. Barrier factors characterise this suppression with increasing resonance spin-values, and decreasing particle momenta.

Whilst the angular distribution Z_L arises naturally from the effective theory of resonances, barrier factors must be added 'by hand'; they parameterize knowledge that is not included within the effective theory, namely the non-zero size of the resonant particle. This size defines a characteristic impact

parameter, below which production of the resonance is likely to occur, and beyond which it is unlikely to occur.

Barrier factors have long been used to parameterize nuclear decays, where the characteristic impact parameter is naturally defined by the radius of the nucleus. The most commonly used barrier factors are those defined by Blatt and Weisskopf [63,64]. First, the Blatt-Weisskopf functions are defined, for the lowest three L-values, as:

$$F_0(q) = 1, \tag{2.11}$$

$$F_1(q) = \sqrt{\frac{2z^2}{z^2+1}}, \text{and} \tag{2.12}$$

$$F_2(q) = \sqrt{\frac{13z^4}{z^4+3z^2+9}}, \tag{2.13}$$

where $z = Rq$ and R is the Blatt-Weisskopf effective radius, the characteristic impact parameter for the interaction. The functions $F_L(q)$ are normalised such that they evaluate to unity when $z = 1$.

Finally, the barrier factors themselves are defined by

$$B_L(q, q_r) \equiv \frac{F_L(q)}{F_L(q_r)}, \tag{2.14}$$

which gives

$$B_0^2(q, q_r) = 1, \tag{2.15}$$

$$B_1^2(q, q_r) = \left(\frac{q}{q_r}\right)^2 \frac{(Rq_r)^2+1}{(Rq)^2+1}, \text{and} \tag{2.16}$$

$$B_2^2(q, q_r) = \left(\frac{q}{q_r}\right)^4 \frac{(Rq_r)^4+3(Rq_r)^3+9}{(Rq)^4+3(Rq)^3+9}. \tag{2.17}$$

For the present analysis, the effective radius is fixed to $R = 1.5\,\text{GeV}^{-1}$.

2.2.2. Angular distribution Z_L

Z_L originates from the two inner product terms of equation (2.4) (the 'vertex factors' of the effective theory defined in Figure 2.1),

$$Z_L = \sum_\lambda \langle ab|r_\lambda\rangle\langle cr_\lambda|D^0\rangle. \tag{2.18}$$

In the case of a scalar resonance ($L = 0$), no polarization states, λ, can be defined, and so the distribution of final state particles has no angular dependence

$$Z_0 = 1. \tag{2.19}$$

In the case of vector ($L = 1$) or tensor ($L = 2$) resonances, multiple helicity states are possible for the resonance, and so the sum must be performed. A quantum field-theoretical[2] approach is used to construct the relevant currents in each case, and to determine the resulting sum rules. A comprehensive guide to the spin formalisms used in the description of particle scattering is given by reference [66].

Vector resonances

We begin by considering the most general theory that can be used to describe Figure 2.1, guided simply by the fact that the matrix element (and thus Z_L) should be a Lorentz scalar quantity. In the most general case, Z will have some dependence on the helicity state of the resonance and on the 4-momenta of the final state particles, since these fully specify the process.

For a spin-1 resonance, three helicity eigenstates are possible ($\lambda = 0, \pm 1$). To construct a manifestly covariant expression for Z, the helicity state must be represented by a Lorentz scalar, Lorentz vector or Lorentz tensor. Three degrees of freedom are needed to represent the helicity state, since at any given time it will be a superposition of all possible helicity eigenstates:

$$|\psi\rangle = a_0|\psi(\lambda = 0)\rangle + a_1|\psi(\lambda = 1)\rangle + a_{-1}|\psi(\lambda = -1)\rangle, \tag{2.20}$$

and to fully specify the state $|\psi\rangle$ all three coefficients a_i must be known. The helicity state may thus be described using a Lorentz vector ϵ^μ_λ, where one additional constraint is required to reduce the degrees of freedom from four (the four elements of the vector) to three.

The constraint chosen is that of 'transversality'; it is required that the polarisation of the resonance is orthogonal to its 4-momentum. This also implies that the resonance is on-shell. This constraint can be expressed mathematically as

$$P_{\mu r}\epsilon^\mu_\lambda = 0, \qquad \text{(transversality)}$$
$$m_{ab} = m_r. \qquad \text{(on-shell)}$$

It turns out that requiring the helicity state to transform as a Lorentz vector, combined with the transversality condition above, is enough to determine the specific form that it must take. The derivation of this form is the subject of Appendix C, but for now it suffices to assume that ϵ^μ_λ is a legitimate Lorentz vector quantity.

We now consider the two 'vertex factors' in turn, constructing the most general expression that yields a Lorentz scalar. Beginning with the term arising from the D-decay vertex, the most general expression that can be written is

$$\langle cr_\lambda|D\rangle = g(m^2_{cr})f_\mu(P_{c\mu}, P_{r\mu}, P_{D\mu})\epsilon^\mu_\lambda, \tag{2.21}$$

where f_μ is some function of the 4-momenta of the three particles which interact at the vertex, and g is some function of the Lorentz scalar m^2_{cr}, the invariant squared mass of the combined decay products.

[2]Many introductory texts on QFT for particle physics are available, see for example reference [65].

Noting conservation of 4-momentum,

$$P_D^\mu = P_c^\mu + P_r^\mu, \tag{2.22}$$

the function f_μ can be re-written in terms of just two independent 4-momenta. Choosing to use the combinations $(P_D + P_c)_\mu$ and $(P_D - P_c)_\mu$, f_μ can be expressed, completely generally, as a polynomial in two variables, with unknown coefficients a_i,

$$\begin{aligned} f_\mu =& a_1(P_D + P_c)_\mu + a_2(P_D - P_c)_\mu + \\ & a_3(P_D + P_c)_\nu(P_D + P_c)^\nu(P_D + P_c)_\mu + a_4(P_D + P_c)_\nu(P_D - P_c)^\nu(P_D + P_c)_\mu + \\ & a_5(P_D + P_c)_\nu(P_D - P_c)^\nu(P_D - P_c)_\mu + a_6(P_D - P_c)_\nu(P_D - P_c)^\nu(P_D - P_c)_\mu + \cdots . \end{aligned} \tag{2.23}$$

This is the most general expression for a function of two variables which yields a single free Lorentz index μ.

Working in the low-momentum regime, keeping only the lowest order terms, and renaming $a_1 = f_+$ and $a_2 = f_-$ obtains

$$f_\mu = f_+(P_D + P_c)_\mu + f_-(P_D - P_c)_\mu. \tag{2.24}$$

Finally, this gives

$$\langle cr_\lambda | D\rangle = g(m_{cr}^2)[f_+(P_D + P_c)_\mu + f_-(P_D - P_c)_\mu]\epsilon_\lambda^\mu. \tag{2.25}$$

Likewise, an equivalent expression may be found for the vertex factor corresponding to the decay of the resonance,

$$\langle ab | r_\lambda\rangle = g'(m_{ab}^2)[f'_+(P_a + P_b)_\nu + f'_-(P_a - P_b)_\nu]\epsilon_\lambda^\nu. \tag{2.26}$$

Applying the transversality condition introduced earlier simplifies these expressions. Since

$$(P_D - P_c)_\mu = (P_r)_\mu \quad \text{and} \tag{2.27}$$

$$(P_a + P_b)_\nu = (P_r)_\nu, \tag{2.28}$$

transversality implies that terms involving these factors do not contribute to Z, so that effectively

$$f_- = f'_+ = 0. \tag{2.29}$$

In addition, since transversality implies the resonance is produced on-shell,

$$g'(m_{ab}^2) = g'(m_r^2), \tag{2.30}$$

which is constant, contributing no angular dependence. Likewise, assuming the progenitor D-meson is on-shell[3],

$$g(m^2_{cr}) = g(m^2_D). \tag{2.31}$$

Neglecting constant factors, one thus obtains

$$\langle cr_\lambda | D\rangle = (P_D + P_c)_\mu \epsilon^\mu_\lambda, \tag{2.32}$$
$$\langle ab | r_\lambda\rangle = (P_a - P_b)_\nu \epsilon^\nu_\lambda. \tag{2.33}$$

This allows Z_1 to be expressed as:

$$Z_1 = \sum_\lambda (P_D + P_c)_\mu \epsilon^\mu_\lambda (P_a - P_b)_\nu \epsilon^\nu_\lambda, \tag{2.34}$$
$$= \sum_\lambda (P_D + P_c)_\mu \epsilon^\mu_\lambda \epsilon^{*\nu}_\lambda (P_a - P_b)_\nu. \tag{2.35}$$

As mentioned earlier, the condition that ϵ^μ_λ behaves as a Lorentz vector is enough to explicitly determine its form. Doing so allows the evaluation of the 'spin sum',

$$\sum_\lambda \epsilon^\mu_\lambda \epsilon^{*\nu}_\lambda = -g^{\mu\nu} + \frac{p^\mu p^\nu}{m^2_r}, \tag{2.36}$$

where p^μ represents the 4-momentum of either of the D-meson decay products (c, r) in the D-meson rest frame. This is discussed in Appendix C. In the 'Zemach formalism' [67, 68], the transversality condition motivates the replacement $m_r \to m_{ab}$ within this expression. Alternatively, the 'helicity formalism' [69] does not make this substitution. For the present analysis, the Zemach formalism is used to obtain the nominal model, whilst a systematic uncertainty is determined by considering the model obtained in the helicity formalism as an alternative. Substituting the expression obtained for the spin sum into equation (2.35), and using the Zemach formalism,

$$Z_1 = (P_D + P_c)_\mu \big(-g^{\mu\nu} + p^\mu p^\nu / m^2_{ab}\big)(P_a - P_b)_\nu, \tag{2.37}$$
$$= (m^2_{bc} - m^2_{ac}) + (m^2_D - m^2_c)(m^2_a - m^2_b)/m^2_{ab}, \tag{2.38}$$
$$= -2\vec{p}\cdot\vec{q}. \tag{2.39}$$

Tensor resonances

For a spin-2 resonance, five helicity eigenstates are possible ($\lambda = 0, \pm 1, \pm 2$), and so a 4-component Lorentz vector is not sufficient to represent the helicity state in this case. Instead, a Lorentz tensor, $\epsilon^{\mu\nu}$, is used. However, since a Lorentz tensor has 16 degrees of freedom[4], and only five are needed, eleven

[3] A reasonable assumption, since the width of the D-meson is narrow.
[4] *I.e.* 16 components.

additional constraints are required. The constraints chosen are

$$\epsilon^{\mu\nu}P_\nu = 0, \qquad \text{(transversality)}$$
$$\epsilon^{\mu\nu} = \epsilon^{\nu\mu}, \text{and} \qquad \text{(symmetry)}$$
$$\epsilon^{\mu\mu} = 0, \qquad \text{(tracelessness)}$$

where for the final expression, the Einstein summation convention implies a sum over μ.

The transversality condition provides four constraints, since it holds for all values of the free index. As for a spin-1 resonance, this condition implies that the resonance is produced 'on-shell', so that $m_{ab} = m_r$. The symmetry condition provides six constraints, since it applies for all non-diagonal elements of the 4×4 matrix, $\epsilon^{\mu\nu}$. Tracelessness provides one constraint, since given the values of any three diagonal elements, it allows the fourth to be deduced.

Once again, we consider the two vertex factors in turn, in each case writing the most general expression which yields a Lorentz scalar. Beginning with the D-decay vertex,

$$\langle cr_\lambda|D\rangle = g(m^2_{cr})f_{\mu\nu}(P_{c\mu}, P_{r\mu}, P_{D\mu})\epsilon^{\mu\nu}_\lambda. \tag{2.40}$$

Again, choosing to work in terms of the two independent variables $(P_D + P_c)_\mu$ and $(P_D - P_c)_\mu$, we write the most general expression for f which yields two free Lorentz indices,

$$\begin{aligned} f_{\mu\nu} =& a_1(P_D + P_c)_\mu(P_D + P_c)_\nu + a_2(P_D + P_c)_\mu(P_D - P_c)_\nu + \\ & a_3(P_D - P_c)_\mu(P_D + P_c)_\nu + a_4(P_D - P_c)_\mu(P_D - P_c)_\nu + \cdots . \end{aligned} \tag{2.41}$$

An equivalent expression may be written for the vertex factor at the decay of the resonance,

$$\langle ab|r_\lambda\rangle = g'(m^2_{ab})f'_{\alpha\beta}(P_{a\alpha}, P_{b\alpha}, P_{r\alpha})\epsilon^{\alpha\beta}_\lambda, \tag{2.42}$$

where

$$\begin{aligned} f'_{\alpha\beta} =& a'_1(P_a + P_b)_\alpha(P_a + P_b)_\beta + a'_2(P_a + P_b)_\alpha(P_a - P_b)_\beta + \\ & a'_3(P_a - P_b)_\alpha(P_a + P_b)_\beta + a'_4(P_a - P_b)_\alpha(P_a - P_b)_\beta + \cdots . \end{aligned} \tag{2.43}$$

Working in the low-momentum limit, the higher order terms of $f_{\mu\nu}$ and $f'_{\alpha\beta}$, denoted by ellipsis, may be neglected.

Applying the transversality condition, terms of $f_{\mu\nu}$ containing any factors of $(P_D - P_c)_\nu = (P_r)_\nu$ will not contribute, nor will terms of $f'_{\alpha\beta}$ containing $(P_a + P_b)_\alpha$. Effectively then,

$$a_2 = a_3 = a_4 = 0, \tag{2.44}$$
$$a'_1 = a'_2 = a'_3 = 0. \tag{2.45}$$

As for the spin-1 case, assuming that both the D-meson and the resonance are on-shell (the latter is a consequence of the transversality condition), the functions $g(m^2_{cr})$ and $g'(m^2_{ab})$ become constant and thus contribute no angular dependence to Z_2.

Finally, the two vertex factors in the spin-2 case are given by

$$\langle cr_\lambda|D\rangle = (P_D + P_c)_\mu (P_D + P_c)_\nu \epsilon^{\mu\nu}_\lambda, \text{and} \tag{2.46}$$

$$\langle ab|r_\lambda\rangle = (P_a - P_b)_\alpha (P_a - P_b)_\beta \epsilon^{\alpha\beta}_\lambda, \tag{2.47}$$

and so Z_2 may be expressed as

$$Z_2 = \sum_\lambda (P_D + P_c)_\mu (P_D + P_c)_\nu \epsilon^{\mu\nu}_\lambda (P_a - P_b)_\alpha (P_a - P_b)_\beta \epsilon^{\alpha\beta}_\lambda, \tag{2.48}$$

$$= \sum_\lambda (P_D + P_c)_\mu (P_D + P_c)_\nu \epsilon^{\mu\nu}_\lambda \epsilon^{*\alpha\beta}_\lambda (P_a - P_b)_\alpha (P_a - P_b)_\beta. \tag{2.49}$$

As in the spin-1 case, the constraints placed on the helicity tensor allow its specific form to be deduced. This in turn allows the evaluation of the 'spin-sum', giving

$$\sum_\lambda \epsilon^{\mu\nu}_\lambda \epsilon^{*\alpha\beta}_\lambda = \frac{1}{2}\left(T^{\mu\alpha}T^{\nu\beta} + T^{\mu\beta}T^{\nu\alpha}\right) - \frac{1}{3}T^{\mu\nu}T^{\alpha\beta}, \tag{2.50}$$

where

$$T^{\mu\nu} = -g^{\mu\nu} + \frac{p^\mu p^\nu}{m^2_{ab}} \tag{2.51}$$

and the Zemach formalism has been used once more to replace m_r with m_{ab}. The derivation of this expression may be found in Sec. 8.3 of reference [70].

Substituting this expression into equation (2.49) yields

$$Z_2 = (P_D + P_c)_\mu (P_D + P_c)_\nu \left[\frac{1}{2}\left(T^{\mu\alpha}T^{\nu\beta} + T^{\mu\beta}T^{\nu\alpha}\right) - \frac{1}{3}T^{\mu\nu}T^{\alpha\beta}\right](P_a - P_b)_\alpha (P_a - P_b)_\beta, \tag{2.52}$$

$$= \left(m^2_{bc} - m^2_{ac} + \frac{(m^2_D - m^2_c)(m^2_a - m^2_b)}{m^2_{ab}}\right)^2 - \frac{1}{3}\left(m^2_{ab} - 2(m^2_D + m^2_c) + \frac{(m^2_D - m^2_c)^2}{m^2_{ac}}\right)\left(m^2_{ab} - 2(m^2_a + m^2_b) + \frac{(m^2_a - m^2_b)^2}{m^2_{ac}}\right), \tag{2.53}$$

$$= \frac{4}{3}\left[3(\vec{p}\cdot\vec{q})^2 - (|\vec{p}||\vec{q}|)^2\right]. \tag{2.54}$$

2.2.3. Dynamical function T_r

In deducing the angular distribution functions Z_L, it was assumed (following the Zemach formalism) that the resonance was created on-shell. However, as a short-lived quantum particle, the resonance has

a non-negligible amplitude to be created off-shell.[5] This is accounted for by the 'propagator factor' T_r, which models the energetical favourability for the decay to proceed via a given resonance.

Breit-Wigner propagator

The simplest propagator is the familiar Breit-Wigner function [71], which originates from the description of resonant scattering,

$$T_r^{\mathrm{BW}}(m_{ab}) = \frac{m_r\Gamma_r^0}{m_r^2 - m_{ab}^2 - im_r\Gamma_r(m_{ab})}, \tag{2.55}$$

where m_{ab} is the reconstructed combined invariant mass of the two daughters of the resonance, m_r is the mass of the resonance, Γ_r^0 is its total width, and the 'mass dependent width' [72] is given by

$$\Gamma_r(m_{ab}) = \Gamma_r\left(\frac{q_{ab}}{q_r}\right)^{2L+1}\left(\frac{m_r}{m_{ab}}\right)B_L(q_{ab},q_r)^2. \tag{2.56}$$

This corrects the 'raw' width of the resonance, Γ_r, to take into account the fact that, when angular momentum conservation requires non-zero orbital angular momentum between the resonance and spectator particle c, the energy required to achieve this reduces the mass-energy available to create the resonance itself.

The momentum of each resonance daughter (a,b) in the rest frame of the resonance is given by q_{ab}, whilst q_r gives the same momentum, calculated under the assumption that the resonance is produced on-shell. In the common case of a resonance decaying to two daughters of equal mass, m_a (*e.g.* $r \to \pi^+\pi^-$), these simplify to

$$q_{ab} = \sqrt{\frac{m_{ab}^2}{4} - m_a^2}, \text{and} \tag{2.57}$$

$$q_r = \sqrt{\frac{m_r^2}{4} - m_a^2}. \tag{2.58}$$

Gounaris-Sakurai propagator

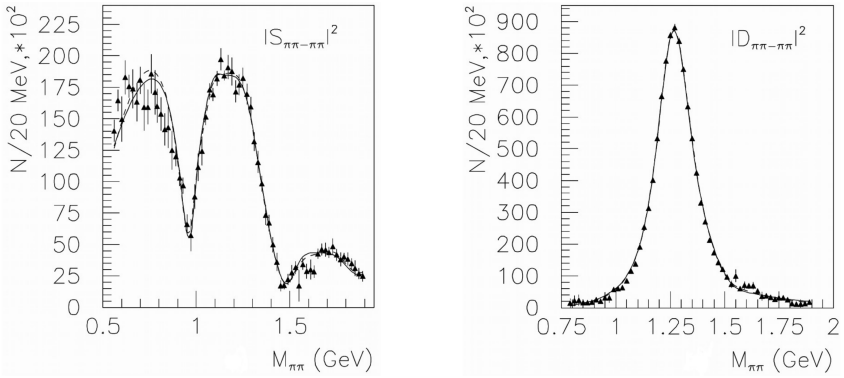
In some cases, an effective-range parameterization, known as the 'Gounaris-Sakurai' propagator [73] is preferred,

$$T_r^{\mathrm{GS}}(m_{ab}) = \frac{m_r^2 + lm_r\Gamma_r^0}{m_r^2 - m_{ab}^2 + f(m_{ab}^2) - im_{ab}\Gamma_r(m_{ab})}, \tag{2.59}$$

where

$$f(m_{ab}^2) = \frac{\Gamma_r^0 m_r^2}{q_r}\left\{\left(\frac{q_{ab}}{q_r}\right)^2\left[h(m_{ab}^2) - h(m_r^2)\right] + h'(m_r^2)\left(m_r^2 - m_{ab}^2\right)\right\}, \tag{2.60}$$

[5] *I.e.* such that $E^2 - p^2 \neq m^2$.



(a) $\pi\pi \to \pi\pi$ S-wave **(b)** $\pi\pi \to \pi\pi$ D-wave

Figure 2.2.: Observed $\pi\pi \to \pi\pi$ scattering rates proceeding via (a) S-wave (spin-zero) and (b) D-wave (spin-two) resonances. A relativistic Breit-Wigner lineshape is fitted to the isolated resonance in the D-wave component, whilst for the S-wave distribution a K-matrix model is used [74].

and

$$h(m^2) = \frac{2q_{ab}}{\pi m}\ln\left(\frac{m+2q_{ab}}{2m_a}\right), \tag{2.61}$$

$$h'(m_r^2) = \left.\frac{dh}{d\,(m^2)}\right|_{m^2=m_r^2} = h(m_r^2)\left(\frac{1}{8q_r^2}-\frac{1}{2m_r^2}\right)+\frac{1}{2\pi m_r^2}. \tag{2.62}$$

The range constant, l, is chosen such that $T_r^{\mathrm{GS}}(m_{ab}=0)=1$, and thus depends on the mass of the resonant particle,

$$l = \frac{3m_a^2}{\pi q_r^2}\ln\left(\frac{m_r+2q_r}{2m_a}\right)+\frac{m_r}{2\pi q_r}-\frac{m_a^2 m_r}{\pi q_r^3}. \tag{2.63}$$

2.2.4. Limitations of the Isobar model

Implicit in the derivation of the Breit-Wigner propagator is the assumption that, for the energy range over which a given resonance has non-negligible amplitude, it is the only available resonance, i.e. it is assumed that, for a given resonant pair, (ab), at any given value of m_{ab} there is only one possible resonance of spin-L that can couple to a and b. Hence, multiple resonances of the same spin can only be treated if they are well-separated in energy.[6]

However, when two resonances of the same spin have nearby central values in m_{ab}, this assumption breaks down. An isobar model built from two terms whose Breit-Wigner propagators overlap does not give a good match to the more complex structure observed in data. This is evident from Figure 2.2, which displays the spin-zero and spin-2 contributions to $\pi\pi$ scattering. A relativistic Breit-Wigner lineshape gives a good parameterization for the single isolated resonance in the D-wave component, but the distribution in the S-wave, where several nearby resonances contribute, must be fitted using a more sophisticated model. Furthermore, a model built in this way can violate unitarity.

[6] *I.e.* if their masses differ significantly.

For the decay mode considered by the present analysis,

$$D^0 \to K^0_S\pi^+\pi^-,$$

this problem arises for the spin-zero ('S-wave') component of the $(\pi^+\pi^-)$ system, which is seen to be made up of several broad overlapping resonances.

This problem cannot be solved by simply replacing the Breit-Wigner propagator with another function (e.g. the Gounaris-Sakurai propagator). Fundamentally, the isobar model is ill-equipped to handle the case of closely coupled resonances, since within this model, contributions from different resonances are combined at the amplitude level,

$$\mathcal{M} = \sum_r a_r \mathcal{M}_r(L, m_{ab}, m_{ac}), \tag{2.64}$$

$$= \sum_r a_r B^D_L(p, p_r) B^r_L(q, q_r) Z_L(m^2_{ab}, m^2_{ac}) T_r(m_{ab}). \tag{2.65}$$

Within this scheme, the calculation of each $\mathcal{M}_r$ in the summation depends only on the properties of the resonance r and the final state particles; there is no way for $\mathcal{M}_r$ to account for other nearby resonances. Only two degrees of freedom, $|a_r|$ and $\arg(a_r)$, are available to describe any possible interaction between resonances, and there is no scope for a more complex relation between any one pair of resonances than any other.

The K-matrix model [75, 76] is an alternative to the isobar model, which allows the treatment of many closely-coupled resonances. This model preserves unitarity by combining separate contributions at the level of T_r, the propagator. In doing so, it effectively treats all available resonances of a given spin as a single, closely coupled system, which carries a single phase relative to the other terms of the model. Since these resonances are now combined at the propagator level, interactions between them may be described, as quanta of energy can transfer between the different resonant modes. Within this model, the spin-zero resonances of the $(\pi^+\pi^-)$ system are no longer included in the summation of equation (2.64), but instead treated altogether in a single term, so that equation (2.64) becomes

$$\mathcal{M} = a_{\pi\pi}\mathcal{M}_{\pi\pi} + \sum_r a_r \mathcal{M}_r(L, m_{ab}, m_{ac}), \tag{2.66}$$

where r indexes only the remaining resonances.

The following section describes the K-matrix formalism used to obtain the amplitude, $\mathcal{M}_{\pi\pi}$, for the coupled resonance system. Within this formalism, it will be seen that the Breit-Wigner expression arises as the lowest order term of a series. For a single, isolated resonance, higher order terms are negligible, and the Breit-Wigner gives a good approximation. However, for a closely coupled system of resonances, higher order terms are non-negligible, and the Breit-Wigner formalism does not give the full amplitude, thus violating unitarity.

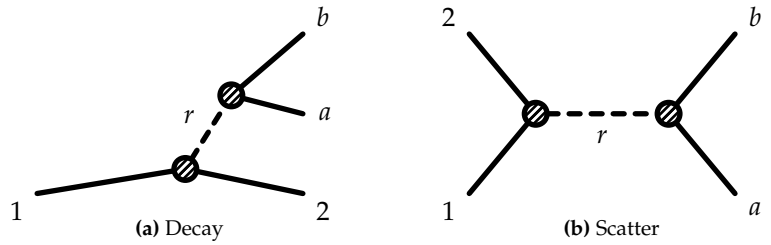


Figure 2.3.: Equivalence of diagrams for three-body decay and scattering processes.

2.3. The *K*-matrix model

The K-matrix formalism applies to any set of (potentially overlapping) resonances of the same spin, L, decaying to the same 'resonant pair' (ab). Spin is handled the same way as in the isobar model, using barrier factors B_L and an angular distribution term Z_L. For the present analysis, only spin-zero resonances are treated using the K-matrix formalism, and since

$$B_0 = 1, \text{and} \quad Z_0 = 1, \tag{2.67}$$

these terms are neglected for the rest of this section.

The K-matrix formalism has its roots in the 'S-matrix' formalism, which originated as a tool to describe the resonant scattering of hadrons in the 1950's and 1960's [77, 78]. The formalism represents an alternative to the second-quantisation field-theoretical approach most often encountered in phenomenology. It remains the *de facto* method for handling hadronic scattering processes.

As well as resonant scattering, the S-matrix formalism also applies to resonances produced in decays, since the diagrams for both cases are fundamentally equivalent. This is illustrated in Figure 2.3. The K-matrix formalism extends the S-matrix formalism to include a P-vector [76] description of the production process.

Before considering the S- and K-matrices and their mathematical properties, it is useful to note that they will be functions of a single variable, m_{ab}, since all spin and directional dependence is handled by B_L and Z_L. This variable defines the energy that is available for the creation of any resonance.

2.3.1. The *S*-matrix formalism

The S-matrix originated as a description for scattering experiments, which can be represented diagramatically by Figure 2.3b. It is defined in terms of the scattering operator, $\mathcal{S}$,

$$S_{fi} = \langle f|\mathcal{S}|i\rangle. \tag{2.68}$$

S_{fi} gives the amplitude for initial state $|i\rangle$ to scatter into final state $|f\rangle$, in some scattering process characterised by $\mathcal{S}$. In the context of Figure 2.3b, $|i\rangle$ denotes the 'channel' (12), whilst $|f\rangle$ denotes the

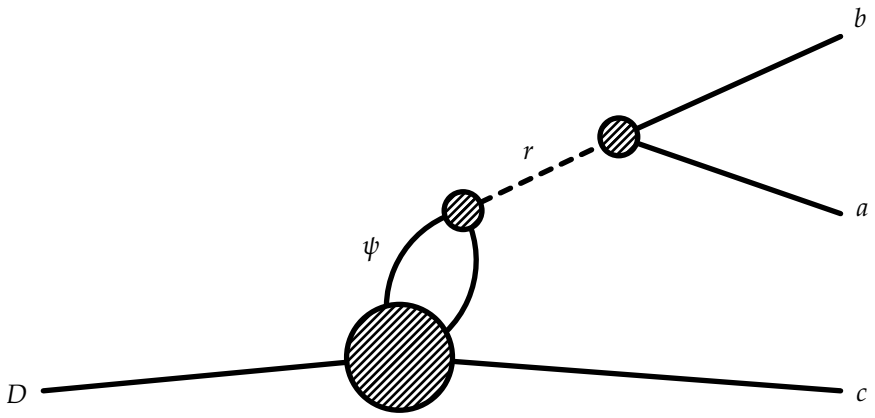


Figure 2.4.: Schematic diagram for resonant decay within the K-matrix model picture.

channel (ab). The amplitude for the scattering process $12 \to ab$ is thus

$$\mathcal{M}(12 \to ab) = \langle ab|\mathcal{S}|12\rangle. \tag{2.69}$$

S_{fi} is a square $n \times n$ matrix, where n is the dimension of the space of allowed states, the number of eigenvectors needed to completely span the space. In a scattering problem, n is given by the number of physically distinguishable channels that are possible; these are the eigenvectors of the S-matrix. Since S describes all interactions between the initial and final channels, it must encapsulate the behaviour of any resonances. The specific form taken by S will be discussed later in this chapter.

Accounting for the production environment

The treatment so far in this section has proceeded in terms of resonant scattering; going from some known initial channel (12), to some known final channel (ab), as displayed in Figure 2.3b. However, the present analysis concerns a resonant *decay*, in which the production mechanism of the resonance at the D-decay vertex is not known. To make the connection with the scattering theory outlined so far, the process is assumed to proceed as follows:

1. At the D-decay vertex, some state (ψ) is produced opposite the spectator hadron c;
2. This state is then propagated into the observed final channel (ab) by 're-scattering' via some resonance r.

This picture is illustrated in Figure 2.4.

The second stage in this process may be treated using the S-matrix formalism discussed so far, since it is equivalent to a typical scattering-type experiment. All that needs to be added is some description of the first stage of the process, the production mechanism at the D-decay vertex.

Since the initial state ψ is never observed, it needn't be a physical eigenstate. Instead, in general, it will be a superposition of all possible channels. The amplitude for the decay process thus becomes

$$\mathcal{M}(ab) = \langle ab|\mathcal{S}|\psi\rangle, \tag{2.70}$$

where $|\psi\rangle$ is given by a superposition in the eigenbasis of the $\mathcal{S}$ operator,

$$|\psi\rangle = a_1|x_1\rangle + a_2|x_2\rangle + a_3|x_3\rangle + \cdots . \tag{2.71}$$

Writing the identity operator in terms of the $\mathcal{S}$ operator eigenbasis - the set of well-defined channels,

$$\mathcal{I} = \sum_i |x_i\rangle\langle x_i| \tag{2.72}$$

and inserting this into equation (2.70), gives

$$\begin{aligned} \mathcal{M}(ab) &= \langle ab|\mathcal{S}\mathcal{I}|\psi\rangle, & (2.73)\\ &= \sum_i \langle ab|\mathcal{S}|x_i\rangle\langle x_i|\psi\rangle, & (2.74)\\ &= S_{fi}\Psi_i, & (2.75) \end{aligned}$$

where f denotes the specific final channel (ab), and the Einstein summation convention has been used in the final expression. Here, S_{fi} is the familiar S-matrix, accounting for the scattering part of the process. Ψ_i is a vector of amplitudes, which accounts for the production environment.

Equation (2.75) generalises simply to all possible final channels. In this case, one may define the final state vector ('F-vector'),

$$F_j = S_{ji}\Psi_i, \tag{2.76}$$

so that each element of F gives the amplitude for a different possible final channel.

Determining the form of the S-matrix and Ψ-vector

Equation (2.75) shows that the amplitude for a resonant decay can be treated in terms of a production vector Ψ_i and scattering matrix S_{fi}. However, the quantum mechanics leading to that expression does not indicate any way to determine, *a priori*, the specific form taken by either of these objects.

A way forward is indicated by considering the simplest case; a single possible channel, with a single possible resonance. In this case, the S-matrix must reduce to a single element which is identical to the Breit-Wigner propagator, since the assumptions made in its derivation are now satisfied. In this case, then,

$$S(m_{ab}) = \frac{m_r\Gamma_r^0}{m_r^2 - m_{ab}^2 - im_r\Gamma_r}, \tag{2.77}$$

where the mass-dependent width $\Gamma_{ab}(q)$ has reduced to Γ_r since only the $L = 0$ case is considered. Inspection of this expression reveals a pole at the point

$$m_{ab}^2 = m_r^2 - im_r\Gamma_r. \tag{2.78}$$

It is suggestive to note that all of the physical properties of the resonance, m_r and Γ_r, are described by the residue[7] at the pole, $m_r\Gamma_r^0$. It is also illuminating to consider adding an extra resonance, well-separated from the first. In this case, a sum of two Breit-Wigner propagators will give a good description, implying that the S-matrix should contain a pole for each resonance.

Considering now the Ψ-vector, the specific production mechanism that it describes is unknown (this is indicated by the large blob of Figure 2.4). It makes sense that it should include some resonant contribution. It is also reasonable to assume that the resonances that are available to take part in this production mechanism are the same as those that couple to channel ψ to allow the later scattering into channel (ab) - the process described by the S-matrix. Thus, the Ψ-vector ought to contain some resonant contribution, with poles at the same positions as those of the S-matrix. However, the possibility must remain open for a non-resonant contribution, since the production mechanism is unknown and could involve more complex interactions, without any resonances.

The 'K-matrix formalism' provides a way to satisfy all of the above criteria for the S-matrix and Ψ-vector, along with unitarity. It is useful to list these conditions for completeness:

1. Poles with correct residue values in the S-matrix elements;
2. Identically located poles in the Ψ-vector elements;
3. Inclusion of non-resonant contributions; and
4. Unitarity.

This formalism is described in the following section.

2.3.2. The K-matrix formalism

Definition of the K-matrix

Conservation of probability implies unitarity of the S-matrix, since it demands that S must preserve the norm of any amplitude vector on which it acts. Furthermore, time reversal invariance implies that S must be symmetric, since it requires that $\langle f|\mathcal{S}|i\rangle$ must equal $\langle i|\mathcal{S}|f\rangle$. This gives the constraint

$$SS^\dagger = S^\dagger S = I. \tag{2.79}$$

It is useful to define the 'transition matrix', T,

$$S = I + 2iT, \tag{2.80}$$

[7] From the theory of functions of a complex variable [79].

so that the S-matrix is decomposed into one part $(2iT)$, that describes the case of a transition between channels, and one part (I), that describes the case where no transition occurs. Applying condition (2.79) gives

$$T - T^\dagger = 2iT^\dagger T = 2iTT^\dagger, \tag{2.81}$$

which can be rewritten in terms of inverse operators as

$$(T^\dagger)^{-1} - T^{-1} = 2iI, \tag{2.82}$$

which further transforms into

$$(T^{-1} + iI)^\dagger = T^{-1} + iI. \tag{2.83}$$

The K operator is then defined as

$$K^{-1} \equiv T^{-1} + iI, \tag{2.84}$$

which is hermitian, due to equation (2.82). Furthermore, since time-reversal invariance implies that T is symmetric, K must be symmetric too. Combining this constraint with hermiticity, it is evident that the K-matrix can be chosen to be real. The definitions of the K- and T-matrices ensure that for any real, symmetric K-matrix, the corresponding S-matrix is guaranteed to satisfy unitarity, as desired. Multiplying equation (2.84) by K from the left and T from the right, and vice versa, gives

$$T = K + iKT = K + iTK, \tag{2.85}$$

which demonstrates that K and T commute. Solving for T gives

$$T = K(I - iK)^{-1} = (I - iK)^{-1}K, \tag{2.86}$$

and thus

$$S = (I + iK)(I - iK)^{-1} = (I - iK)^{-1}(I + iK). \tag{2.87}$$

Further useful relations may be obtained from equation (2.85). Since K is real, one finds

$$Re(T) = (I + K^2)^{-1}K = K(I + K^2)^{-1}, \text{and} \tag{2.88}$$

$$Im(T) = (I + K^2)^{-1}K^2 = K^2(I + K^2)^{-1}. \tag{2.89}$$

Combined with the unitarity condition (2.81), this gives

$$Im(T) = T^*T = TT^*. \tag{2.90}$$

In addition, from equation (2.84), one obtains

$$Im(T^{-1}) = -I. \tag{2.91}$$

Lorentz invariance

The T and K operators discussed so far are not manifestly Lorentz invariant. Invariance is achieved by applying relativistic normalisation factors for the phase space of the initial and final channels [80]. The diagonal phase space matrix is defined as

$$\rho_{jk} = \delta_j \rho_k, \tag{2.92}$$

so that

$$\rho = \begin{pmatrix} \rho_1 & & \\ & \rho_2 & \\ & & \ddots \end{pmatrix}, \tag{2.93}$$

where ρ_j is the density of states in channel j. For the final channel (ab), the density of states is

$$\rho_f(m_{ab}^2) = \sqrt{\left[1 - (\frac{m_a + m_b}{m_{ab}})^2\right]\left[1 - (\frac{m_a - m_b}{m_{ab}})^2\right]}, \tag{2.94}$$

normalised such that $\rho_f(m_{ab}^2) \to 1$ as $m_{ab}^2 \to \infty$.

The Lorentz invariant transition amplitude, $\hat{T}$, may then be defined as

$$T_{ij} = \rho_i^{\frac{1}{2}} \hat{T}_{ij} \rho_j^{\frac{1}{2}}, \tag{2.95}$$

or, in matrix notation

$$T = \{\rho\}^{\frac{1}{2}} \hat{T} \{\rho\}^{\frac{1}{2}}. \tag{2.96}$$

In terms of $\hat{T}$, equation (2.90) and (2.91) become

$$Im(\hat{T}) = \hat{T}^* \rho \hat{T} = \hat{T} \rho \hat{T}^*, \tag{2.97}$$

and

$$Im(\hat{T}^{-1}) = -\rho. \tag{2.98}$$

One may also define the invariant form of the K-matrix,

$$K = \{\rho\}^{\frac{1}{2}} \hat{K} \{\rho\}^{\frac{1}{2}}. \tag{2.99}$$

Combining this with equation (2.84) gives

$$\hat{K}^{-1} = \hat{T}^{-1} + i\rho. \tag{2.100}$$

Following the same steps which led to equation (2.85) gives

$$\hat{T} = \hat{K} + i\hat{K}\rho\hat{T} = \hat{K} + i\hat{T}\rho\hat{K}, \tag{2.101}$$

and thus

$$\hat{T}\rho\hat{K} = \hat{K}\rho\hat{T}. \tag{2.102}$$

Solving for $\hat{T}$ gives

$$\hat{T} = \hat{K}(I - i\rho\hat{K})^{-1} = (I - i\rho\hat{K})^{-1}\hat{K}, \tag{2.103}$$

and combining with equation (2.87) yields

$$S = (I + i\{\rho\}^{\frac{1}{2}}\hat{K}\{\rho\}^{\frac{1}{2}})(I - i\{\rho\}^{\frac{1}{2}}\hat{K}\{\rho\}^{\frac{1}{2}})^{-1} \tag{2.104}$$

$$= (I - i\{\rho\}^{\frac{1}{2}}\hat{K}\{\rho\}^{\frac{1}{2}})^{-1}(I + i\{\rho\}^{\frac{1}{2}}\hat{K}\{\rho\}^{\frac{1}{2}}). \tag{2.105}$$

Parameterisation of resonances

Within the K-matrix formalism, each resonance is described by a pole term, with independent coefficients to quantify the strength of its coupling to each possible channel. Typically, the simplest form that satisfies these conditions is used [75],

$$K_{ij}(m_{ab}^2) = \sum_\alpha \frac{g_{\alpha i}(m_{ab}^2)g_{\alpha j}(m_{ab}^2)}{m_\alpha^2 - m_{ab}^2}, \tag{2.106}$$

or, in Lorentz-invariant form,

$$\hat{K}_{ij}(m_{ab}^2) = \sum_\alpha \frac{g_{\alpha i}(m_{ab}^2)g_{\alpha j}(m_{ab}^2)}{(m_\alpha^2 - m_{ab}^2)\sqrt{\rho_i(m_{ab}^2)\rho_j(m_{ab}^2)}}, \tag{2.107}$$

where α indexes the possible resonances, and $g_{\alpha i}(m_{ab}^2)$ gives the coupling of resonance α to channel i. So long as the coupling factors are real, this parameterization gives a hermitian K-matrix, as required to preserve unitarity. The mathematical interpretation of the K-matrix is simple. For the case of a single resonance r, equation (2.106) may be written as

$$K_{ij}(m_{ab}^2) = \frac{g_{ri}(m_{ab}^2)}{(m_r + m_{ab})}\frac{g_{rj}(m_{ab}^2)}{(m_r - m_{ab})}. \tag{2.108}$$

It is seen that each element i, j of the K-matrix (corresponding to scattering between channels i and j) contains two poles, one for the initial and one for the final channel. The Laurent coefficient, or residue, at each pole gives the coupling strength g between the resonance and the corresponding channel. It is

useful to note that any hermitian matrix may be added to equation (2.106) and equation (2.107), without violating unitarity. This is the method used to describe any non-resonant contribution [76].

The coupling coefficients $g_{\alpha i}(m^2_{ab})$ may be written in terms of the 'mass' and 'width' of the corresponding resonance,

$$g^2_{\alpha i}(m^2_{ab}) = m_\alpha \Gamma_{\alpha i}(m^2_{ab}). \tag{2.109}$$

In the case of well-separated resonances, these parameters equate to the same mass, m_r, and width, Γ_r, which appear in the Breit-Wigner propagator. However, in general the interaction of nearby resonances produces complex values for these parameters, so that the couplings $g_{\alpha i}(m^2_{ab})$ are not necessarily positive.

It is typical to assume that phase space effects provide the dominant source of energy-dependence contributing to $\Gamma_{\alpha i}(m^2_{ab})$, so that one may write [61]

$$\Gamma_{\alpha i}(m^2_{ab}) = \Gamma^0_{\alpha i}\rho_i(m^2_{ab}), \tag{2.110}$$

where $\Gamma^0_{\alpha i}$ are taken to be constant.[8] Thus

$$g_{\alpha i}(m^2_{ab}) = g^0_{\alpha i}\sqrt{\rho_i(m^2_{ab})}, \tag{2.111}$$

where $g^0_{\alpha i}$ are constant. The K-matrix may then be expressed as

$$\hat{K}_{ij} = \sum_\alpha \frac{g^0_{\alpha i} g^0_{\alpha j}}{m^2_\alpha - m^2_{ab}}, \tag{2.112}$$

where the constant parameters $g^0_{\alpha i}$ must be determined from a fit to data.

It can be shown that the K-matrix, defined and parameterized above, gives rise to poles with correctly-valued residues in the corresponding S- and T-matrices. This is easily demonstrated for the simplest case, a single channel, with a single possible resonance. In this case, all matrices reduce to single elements, and equation (2.106) becomes

$$K(m^2_{ab}) = \frac{g^2_{ri}(m^2_{ab})}{(m^2_r - m^2_{ab})} = \frac{m_r\Gamma_r(m^2_{ab})}{(m^2_r - m^2_{ab})}, \tag{2.113}$$

where the second equality follows from equation (2.109), and the fact that the Breit-Wigner assumptions are satisfied. The definition of the K-matrix, equation (2.84), rearranges to give

$$T = (K^{-1} - iI)^{-1}, \tag{2.114}$$

[8] *I.e.* their energy-dependence is assumed to be negligible compared to the factors $\rho_i(m^2_{ab})$.

so substitution and simple algebra yields the final expression,

$$T = \frac{m_r\Gamma_r(m_{ab})}{m_r^2 - m_{ab}^2 - im_r\Gamma_r(m_{ab})}, \tag{2.115}$$

$$= \frac{m_r\Gamma_r^0}{m_r^2 - m_{ab}^2 - im_r\Gamma_r(m_{ab})}\rho(m_{ab}), \tag{2.116}$$

giving

$$\hat{T} = \frac{m_r\Gamma_r^0}{m_r^2 - m_{ab}^2 - im_r\Gamma_r(m_{ab})}, \tag{2.117}$$

which equates to the Breit-Wigner function as expected.

In order to illustrate how the K-matrix formalism generalises to more complex systems, it is useful to consider two further examples:

1. a single channel with two allowed resonances; and
2. two channels linked by a single resonance.

Single channel with two allowed resonances

Since there is only one possible channel, all matrices once again reduce to single elements. For two resonances, 1 and 2, the K-matrix becomes

$$K(m_{ab}^2) = \frac{m_1\Gamma_1(m_{ab}^2)}{(m_1^2 - m_{ab}^2)} + \frac{m_2\Gamma_2(m_{ab}^2)}{(m_2^2 - m_{ab}^2)} \tag{2.118}$$

$$= \frac{m_1\Gamma_1(m_2^2 - m_{ab}^2) + m_2\Gamma_2(m_1^2 - m_{ab}^2)}{(m_1^2 - m_{ab}^2)(m_2^2 - m_{ab}^2)}. \tag{2.119}$$

Using equation (2.114) for the T-matrix yields the expression

$$T = \frac{m_1\Gamma_1(m_2^2 - m_{ab}^2) + m_2\Gamma_2(m_1^2 - m_{ab}^2)}{(m_1^2 - m_{ab}^2)(m_2^2 - m_{ab}^2) - i\left[m_1\Gamma_1(m_2^2 - m_{ab}^2) + m_2\Gamma_2(m_1^2 - m_{ab}^2)\right]}. \tag{2.120}$$

In the limit $m_1 = m_2$, this expression reduces to the familiar Breit-Wigner form, with $m_r = m_1 = m_2$ and $\Gamma_r = \Gamma_1 + \Gamma_2$. The two resonances interfere purely constructively, so that the rates of scattering via the two resonances sum.

In the limit of two well-separated resonances, equation (2.120) reduces to the sum of two Breit-Wigners, as expected.[9] This is most easily demonstrated by constructing the sum of two Breit-Wigner functions,

$$T_1^{\text{BW}} + T_2^{\text{BW}} = \frac{m_1\Gamma_1(m_2^2 - m_{ab}^2 - im_2\Gamma_2) + m_2\Gamma_2(m_1^2 - m_{ab}^2 - im_1\Gamma_1)}{(m_2^2 - m_{ab}^2 - im_2\Gamma_2)(m_1^2 - m_{ab}^2 - im_1\Gamma_1)}, \tag{2.121}$$

[9] Since the assumptions of the Breit-Wigner formula are now valid.

which rearranges to

$$T_1^{\text{BW}} + T_2^{\text{BW}} = \frac{m_1\Gamma_1(m_2^2 - m_{ab}^2) + m_2\Gamma_2(m_1^2 - m_{ab}^2) - 2im_1m_2\Gamma_1\Gamma_2}{(m_1^2 - m_{ab}^2)(m_2^2 - m_{ab}^2) - i\left[m_1\Gamma_1(m_2^2 - m_{ab}^2) + m_2\Gamma_2(m_1^2 - m_{ab}^2)\right] - m_1m_2\Gamma_1\Gamma_2}. \quad (2.122)$$

Applying the condition that the two resonances be well separated,

$$m_2 - m_1 \gg \Gamma_1, \Gamma_2, \quad (2.123)$$

the final terms of the numerator and denominator become negligible, leaving equation (2.122) identical to equation (2.120).

Away from the limiting cases discussed above, equation (2.120) cannot be approximated by any Breit-Wigner-based approach.

Two channels with a single resonance

In this case, since two channels are possible, all matrices are 2×2. From equation (2.106), the K-matrix may be expressed as

$$K(m_{ab}^2) = \frac{1}{m_r^2 - m_{ab}^2}\begin{pmatrix} g_x^2 & g_xg_y \\ g_yg_x & g_y^2 \end{pmatrix}, \quad (2.124)$$

where subscripts x and y denote the two possible channels. The corresponding T-matrix may be most easily found by using equation (2.86). Letting

$$\eta = \frac{1}{m_r^2 - m_{ab}^2}, \quad (2.125)$$

one may write

$$I - iK = \begin{pmatrix} 1 - i\eta g_x^2 & -i\eta g_xg_y \\ -i\eta g_yg_x & 1 - i\eta g_y^2 \end{pmatrix}. \quad (2.126)$$

Inverting this expression yields

$$(I - iK)^{-1} = \frac{1}{(1 - i\eta g_x^2)(1 - i\eta g_y^2) - (-i\eta g_xg_y)^2}\begin{pmatrix} 1 - i\eta g_y^2 & -i\eta g_xg_y \\ -i\eta g_yg_x & 1 - i\eta g_x^2 \end{pmatrix}. \quad (2.127)$$

Finally, T may be obtained,

$$K(I - iK)^{-1} = \frac{\eta}{(1 - i\eta g_x^2)(1 - i\eta g_y^2) - (-i\eta g_xg_y)^2}\begin{pmatrix} g_x^2 & g_xg_y \\ g_yg_x & g_y^2 \end{pmatrix}\begin{pmatrix} 1 - i\eta g_y^2 & -i\eta g_xg_y \\ -i\eta g_yg_x & 1 - i\eta g_x^2 \end{pmatrix}. \quad (2.128)$$

Simple algebra reduces this to

$$T = \frac{1}{m_r^2 - m_{ab}^2 - i(g_x^2 + g_y^2)} \begin{pmatrix} g_x^2 - 2i\eta g_x^2 g_y^2 & g_x g_y - 2i\eta g_x^3 g_y \\ g_x g_y - 2i\eta g_x g_y^3 & g_y^2 - 2i\eta g_x^2 g_y^2 \end{pmatrix}. \tag{2.129}$$

Keeping only the lowest order terms in g yields the Flatté formula [81],

$$T^{\text{Flatté}} = \frac{1}{m_r^2 - m_{ab}^2 - i(g_x^2 + g_y^2)} \begin{pmatrix} g_x^2 & g_x g_y \\ g_y g_x & g_y^2 \end{pmatrix}, \tag{2.130}$$

originally derived empirically to describe the cusp in amplitude observed when the energy threshold of a new channel is reached.

The *P*-vector

The above discussion of the K-matrix applies to scattering between two channels. However, the process to be modelled in the present analysis,

$$D^0 \to K_S^0\pi^+\pi^-, \tag{2.131}$$

does not have a well-defined initial channel, but rather some unknown production mechanism at the D-decay vertex. Equation (2.76) anticipates the form of the full amplitude,

$$F_j = \hat{T}_{ji}\Psi_i, \tag{2.132}$$

where Ψ_i describes the superposition state produced at the D-decay vertex, in the basis of possible channels. The final state amplitude vector is given by F, whose j'th element gives the amplitude to produce channel j in the final state. No Standard Model description of the production mechanism is currently available, since it remains beyond the present limits of lattice QCD calculation. Instead, the elements of Ψ are treated as unknown constants, which must be determined by fitting experimental data.

Substituting the explicit form of $\hat{T}$ from equation (2.86) yields

$$F_j = [(I - i\hat{K}\rho)^{-1}\hat{K}]_{ji}\Psi_i, \tag{2.133}$$

$$= [(I - i\hat{K}\rho)^{-1}]_{jk}\hat{K}_{ki}\Psi_i, \tag{2.134}$$

which can be rewritten as

$$F = (I - i\hat{K}\rho)^{-1}P, \tag{2.135}$$

where the P-vector is defined as

$$P_k = \hat{K}_{ki}\Psi_i = \sum_\alpha \sum_i \frac{\Psi_i g_{\alpha i}^0 g_{\alpha k}^0}{m_\alpha^2 - m_{ab}^2}. \tag{2.136}$$

Defining the unknown parameters β^0_α,

$$\beta^0_\alpha = \sum_i \Psi_i g^0_{\alpha i}, \tag{2.137}$$

equation (2.136) becomes

$$P_k = \sum_\alpha \frac{\beta^0_\alpha g^0_{\alpha k}}{m^2_\alpha - m^2_{ab}}. \tag{2.138}$$

The poles of P precisely match those of the K-matrix, as desired.

LASS-like effective-range parameterization

The LASS experiment observed a broad, single resonance, centred at 1430 MeV, in the $K^-\pi^+$ S-wave component of $K^-p \to K^-\pi^+\eta$ scattering [82]. The Breit-Wigner line shape was found to give a poor description of the resonance, and so instead an effective range parameterization was used,

$$T_{\text{LASS}} = \sin\delta_R e^{i\delta_R} e^{2i\delta_B} + \sin\delta_B e^{i\delta_B} = \sin(\delta_R + \delta_B) e^{i(\delta_R+\delta_B)}, \tag{2.139}$$

where

$$\tan\delta_R = \frac{m_r \Gamma(m_{ab})}{m^2_r - m^2_{ab}}, \tag{2.140}$$

$$\cot\delta_B = \frac{1}{a q_{ab}} + \frac{r q_{ab}}{2}, \tag{2.141}$$

and the constants a and r give the scattering length and effective range of the coupling to the resonance. It is important to note that, despite this interpretation of the parameterization, neither a nor r are defined to be positive [83]. The operator T_{LASS} satisfies unitarity, and can be obtained in the K-matrix formalism by considering the special case of a two-body, single channel scattering, with the K-matrix set equal to $K = \tan(\delta_R + \delta_B)$.

This parameterization was originally introduced to describe elastic scattering, but may be generalized for the description of production experiments as

$$T_{\text{gLASS}} = R\sin\delta_R e^{i(\delta_R+\phi_R)} e^{2i(\delta_B+\phi_B)} + B\sin(\delta_B+\phi_B) e^{i(\delta_B+\phi_B)}, \tag{2.142}$$

where B, R, ϕ_B and ϕ_R are real parameters. This reduces to equation (2.139) in the case that $B = R = 1$ and $\phi_B = \phi_R = 0$.

Inserting the identities

$$e^{i\delta}\sin\delta = \frac{1}{\cot\delta - i}, \text{ and} \tag{2.143}$$

$$e^{2i\delta} = \frac{\cot\delta + i}{\cot\delta - i}, \tag{2.144}$$

allows the expression (2.142) to be rewritten as

$$\begin{aligned} T_{\text{gLASS}} =& Re^{i\phi_R}e^{2i\phi_B}\left(\frac{\cot\delta_B + i}{\cot\delta_B - i}\right)\frac{m_r\Gamma(m_{ab})}{m_r^2 - m_{ab}^2 - im_r\Gamma(m_{ab})} \\ &+ Be^{i\phi_B}\frac{\cos\phi_B + \sin\phi_B\cot\delta_B}{\cot\delta_B - i}, \end{aligned} \tag{2.145}$$

which therefore gives the Lorentz-invariant expression

$$\begin{aligned} \hat{T}_{\text{gLASS}} =& Re^{i\phi_R}e^{2i\phi_B}\left(\frac{q_{ab}\cot\delta_B + iq_{ab}}{q_{ab}\cot\delta_B - iq_{ab}}\right)\frac{m_r\Gamma_r^0}{m_r^2 - m_{ab}^2 - im_r\Gamma(m_{ab})} \\ &+ B\frac{m_{ab}}{2}e^{i\phi_B}\frac{\cos\phi_B + \sin\phi_B\cot\delta_B}{q_{ab}\cot\delta_B - iq_{ab}}. \end{aligned} \tag{2.146}$$

Factors of q_{ab} are not cancelled in this expression, since it is undesirable to work with $1/q_{ab}$ terms, which diverge for $q_{ab} \to 0$.

2.4. Decay model implemented in the present analysis

The present analysis implements the model introduced by BaBar in 2010 [55]. Model parameter values have been determined at BaBar using a fit to a flavour-tagged $D \to K_S^0\pi^+\pi^-$ sample obtained from 468.5 fb^{-1} of e^+e^- data, gathered at SLAC's PEP-II asymmetric-energy collider, for a centre-of-mass energy of $\sqrt{s} = 10.6$ GeV. The model may be summarised as

$$\mathcal{M} = \mathcal{M}_{\pi\pi} + a_{K\pi}e^{i\phi_{K\pi}}\mathcal{M}_{K\pi} + \sum_{\text{P,D}-\text{waves}} a_r e^{i\phi_r}\mathcal{M}_r, \tag{2.147}$$

where:

- $\mathcal{M}_{\pi\pi}$ gives the amplitude for the $\pi\pi$ S-wave, which is described using the K-matrix formalism with a P-vector approach;
- $\mathcal{M}_{K\pi}$ gives the amplitude for the $K\pi$ S-wave, which is described using a generalised LASS model; and
- $\mathcal{M}_r$ are the amplitudes included in the isobar model for the P- and D-wave components. The P-wave description includes the resonances $\rho(770)^0$, $\omega(782)$, $K^*(782)^\pm$ and $K^*(1680)^-$, whilst the D-wave includes the $K^*{}_2(1430)^\pm$ and $f_2(1270)$ resonances.

The coefficients of each model component, $a_{K\pi}e^{i\phi_{K\pi}}$ and $a_re^{i\phi_r}$, are included as model parameters, and have been calibrated experimentally by the BaBar collaboration [55].

P- and D-waves: An isobar model is used to describe the narrow, isolated resonances of the P- and D-wave components. Relativistic Breit-Wigner propagators are used, except for the $\rho(770)^0$ resonance, where a Gounaris-Sakurai propagator is preferred. The angular dependence of the decay amplitude is described using the Zemach formalism ($Z_l(m_{ab}^2, m_{ac}^2)$ factors), whilst spin suppression is accounted for using the Blatt-Weisskopf factors ($B_l(q, q_r)$), with the effective radius fixed to 1.5 GeV^{-1}. The

Lorentz-invariant amplitude for decay proceeding via a given resonance r is then expressed as

$$\mathcal{M}_r = Z_l(m_{ab}^2, m_{ac}^2) B_l(p, p_r) B_l(q, q_r) \Delta_r(m_{ab}). \tag{2.148}$$

***K*-matrix for $\pi\pi$ S-wave:** The $\pi\pi$ S-wave component is described using a K-matrix formalism with a P-vector description of the production environment. Five channels (denoted by α) are considered: $\pi^+\pi^-$; $K\bar{K}$; 4π; $\eta\,\eta$; and $\eta\,\eta'$.

The resulting amplitude for the $\pi\pi$ S-wave is given by

$$\mathcal{M}_{\pi\pi}(m_{\pi\pi}) = \sum_\alpha \left[I - i\hat{K}(m_{\pi\pi})\rho(m_{\pi\pi})\right]^{-1}_{\pi\pi,\alpha} \hat{P}_\alpha(m_{\pi\pi}), \tag{2.149}$$

where the K-matrix is parameterized in the form

$$\hat{K}_{ij}(m_{\pi\pi}) = \left(\sum_\alpha \frac{g^0_{\alpha i} g^0_{\alpha j}}{m_\alpha^2 - m_{\pi\pi}^2} + f_{ij}^{sc} \frac{1\,\text{GeV}^2 - s_0^{sc}}{m_{\pi\pi}^2 - s_0^A}\right) \left[\frac{1\,\text{GeV}^2 - s_0^A}{m_{\pi\pi}^2 - s_0^A}\left(m_{\pi\pi}^2 - \frac{s_A m_\pi^2}{2}\right)\right]. \tag{2.150}$$

The term in square brackets is known as the Adler zero [84, 85], and suppresses the false singularity that appears at the $\pi\pi$ production threshold. All of the parameters $g^0_{\alpha i}$, m_α, f_{ij}^{sc}, s_0^{sc}, s_0^A and s_A are fixed from scattering data [86], with their values given in Table 2.1.

Table 2.1.: K-matrix parameters, in units of GeV, obtained in the global analysis of $\pi\pi$ scattering data of reference [86].

m_α	$g^\alpha_{\pi\pi}$	$g^\alpha_{K\bar{K}}$	$g^\alpha_{4\pi}$	$g^\alpha_{\eta\eta}$	$g^\alpha_{\eta\eta'}$
0.65100	0.22889	-0.55377	0.	-0.39899	-0.34639
1.20360	0.94128	0.55095	0.	0.39065	0.31503
1.55817	0.36856	0.23888	0.55639	0.18340	0.18681
1.21000	0.33650	0.40907	0.85679	0.19906	-0.00984
1.82206	0.18171	-0.17558	-0.79658	-0.00355	0.22358

s_0^{sc}	$f^{sc}_{\pi\pi,\pi\pi}$	$f^{sc}_{\pi\pi,K\bar{K}}$	$f^{sc}_{\pi\pi,4\pi}$	$f^{sc}_{\pi\pi,\eta\eta}$	$f^{sc}_{\pi\pi,\eta\eta'}$
-3.92637	0.23399	0.15044	-0.20545	0.32825	0.35412

s_A^0	s_A
-0.15	1

The P-vector is parameterized as

$$\hat{P}_j(m_{\pi\pi}) = \sum_\alpha \frac{\beta^0_\alpha g^0_{\alpha j}}{m_\alpha^2 - m_{\pi\pi}^2} + f^{pr}_{\pi\pi,j} \frac{1\,\text{GeV}^2 - s_0^{pr}}{m_{\pi\pi}^2 - s_0^{pr}}, \tag{2.151}$$

where the (complex) resonance production parameters β^0_α and the non-resonant background parameters $f^{pr}_{\pi\pi,j}$ and s_0^{pr} are fixed to values measured by the BaBar collaboration [55]. These values are summarised in Table 2.2. Parameters $\beta_{\eta\eta'}$, $f^{pr}_{\pi\pi,\eta\eta}$ and $f^{pr}_{\pi\pi,\eta\eta'}$ are fixed to zero, since BaBar found these components to be negligible.

Table 2.2.: Parameters of the P-vector used to describe the production environment for the $\pi\pi$ S-wave component, with values taken from BaBar [55].

	Amplitude	Phase
$\beta_{\pi\pi}$	5.54 ± 0.06	-0.054 ± 0.007
$\beta_{K\bar{K}}$	15.64 ± 0.06	-3.125 ± 0.005
$\beta_{4\pi}$	44.6 ± 1.2	2.731 ± 0.015
$\beta_{\eta\eta}$	9.3 ± 0.2	2.30 ± 0.02
$f^{pr}_{\pi\pi,\pi\pi}$	11.43 ± 0.11	-0.005 ± 0.009
$f^{pr}_{\pi\pi,K\bar{K}}$	15.5 ± 0.4	-1.13 ± 0.02
$f^{pr}_{\pi\pi,4\pi}$	7.0 ± 0.7	0.99 ± 0.11

Generalised LASS-like parameterization:

A generalised-LASS model is used to describe the amplitude of the ${K^*}_0(1430)^{\pm}$ resonance in the $K^0_S\pi^{\pm}$ S-wave. This model may be written as

$$\mathcal{M}_{K\pi} = R\sin\delta_R e^{i(\delta_R+\phi_R)}e^{2i(\delta_B+\phi_B)} + B\sin(\delta_B+\phi_B)e^{i(\delta_B+\phi_B)}, \tag{2.152}$$

where

$$\tan\delta_R = \frac{m_r\Gamma(m_{ab})}{m_r^2 - m_{ab}^2}, \text{ and} \tag{2.153}$$

$$\cot\delta_B = \frac{1}{aq_{ab}} + \frac{rq_{ab}}{2}. \tag{2.154}$$

All of the shape parameters are fixed to the values obtained by BaBar [55], as summarized in Table 2.3.

Table 2.3.: Parameters of the generalised LASS line shape used to describe the $K\pi$ S-wave component, with all values taken from reference [55].

Parameter	Value
$M_{{K^*}_0(1430)}$	$1421.5 \pm 1.6\,\text{MeV}$
$\Gamma_{{K^*}_0(1430)}$	$247 \pm 3\,\text{MeV}$
B	0.62 ± 0.04
ϕ_B	$-0.100 \pm 0.010\,\text{rad}$
R	1
ϕ_R	$1.10 \pm 0.02\,\text{rad}$
a	$0.224 \pm 0.003\,\text{GeV}$
r	$-15.01 \pm 0.13\,\text{GeV}$

3. The LHC and the LHCb detector

This chapter describes CERN's Large Hadron Collider (LHC) and the LHCb detector, with a special focus on the features that make it an excellent laboratory for precision measurements of hadron decays.

3.1. The LHC

The Large Hadron Collider is designed to accelerate and collide protons with a nominal centre-of-mass energy of 14 TeV [87]. Built within the same circular underground tunnel used for its predecessor, LEP (the Large Electron-Positron collider [88]), it lies approximately 100 m below surface, on average, and has a circumference of 26.7 km.

Protons are first accelerated to successively higher energies in a complex of interconnected accelerators, shown in Fig. 3.1. Protons are stripped from hydrogen atoms by means of ionisation, and enter the first acceleration stage, LINAC2. This linear accelerator boosts the protons to an energy of 50 MeV, before acceleration through the Proton Synchrotron Booster yields an energy of 1.4 GeV. Protons then enter the Proton Synchrotron, where they reach an energy of 25 GeV. This synchrotron also uses radio-frequency (RF) cavities to regularise the spacing between bunches of protons. The bunches of protons are then circulated in the Super Proton Synchrotron, where they attain energies of 450 GeV before being injected into the main LHC ring. Here, two counter-rotating proton beams are accelerated in eight superconducting cavities, and finally collide.

The proton beams collide within four separate detectors lying on the LHC ring. ATLAS and CMS are General Purpose Detectors (GPDs), designed to detect a broad range of particle interactions. ALICE

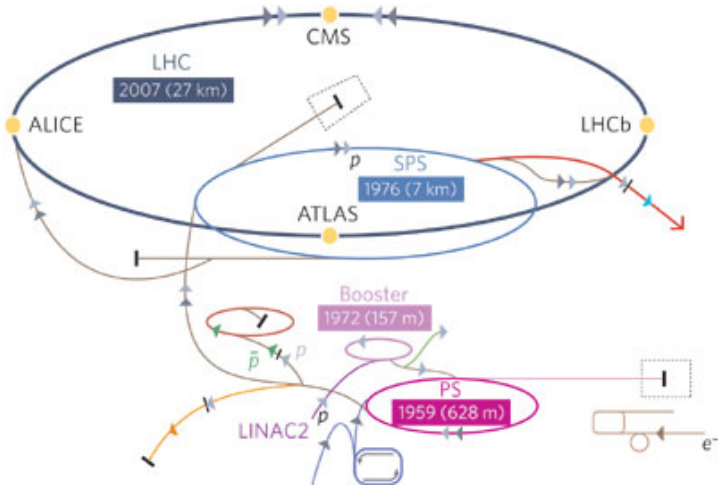


Figure 3.1.: Schematic of the CERN accelerator complex, which accelerates protons to be collided within one of the detectors [89].

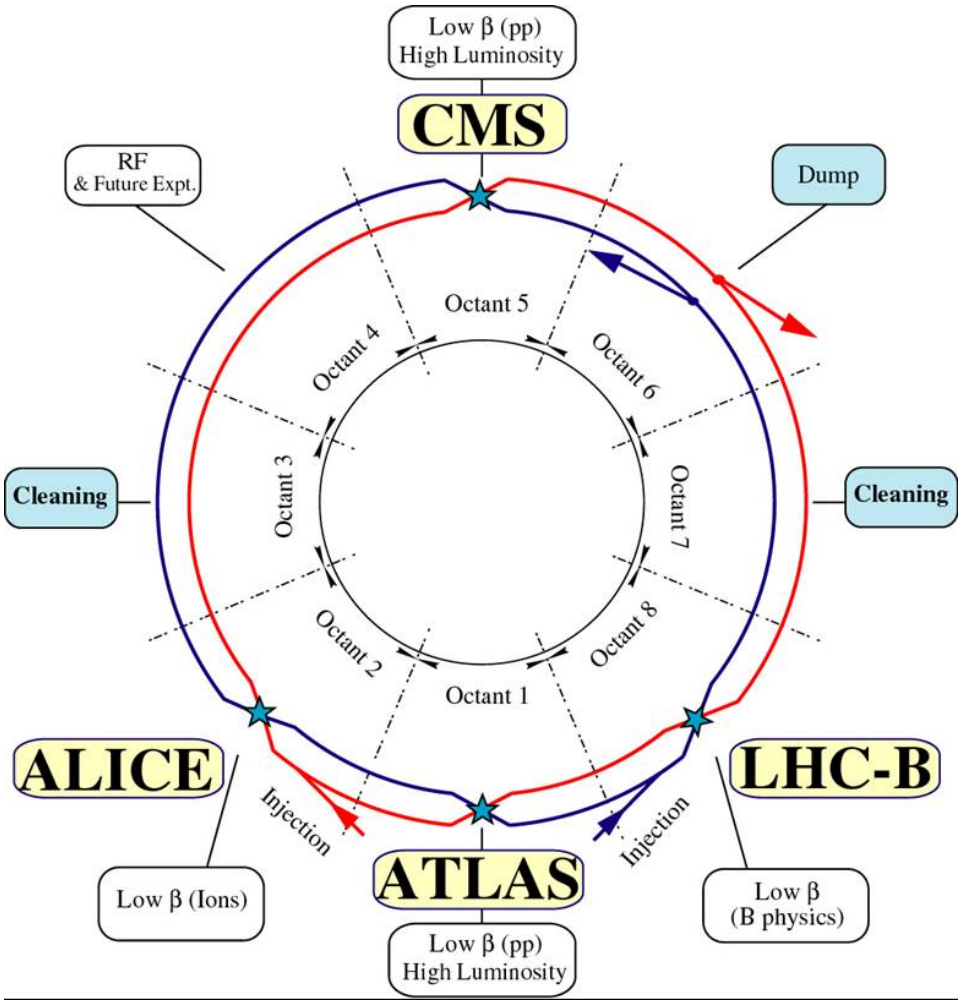


Figure 3.2.: Schematic diagram of the two counter-rotating beam paths within the LHC, displaying the interaction points for the four main detectors [87].

is designed primarily to study the collision of heavy ions, and LHCb is optimised to measure decays of b and c quarks.

The rate at which events of a given type occur within the LHC is given by

$$N = \mathcal{L}\sigma_t, \tag{3.1}$$

where σ_t gives the *cross-section* for the chosen event type, t, and $\mathcal{L}$ is the *luminosity*, equal to the flux of protons through unit area per unit time. The luminosity depends on a range of different factors, including the number of particles per bunch, the period of rotation of the circulating beam, and the cross-sectional area and crossing angle of the two beams at the collision point [87]. The GPDs are designed to operate at a luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, whilst LHCb is optimised to operate with a luminosity 50 times smaller. The number of interactions delivered to a given experiment may be quantified in terms of the *integrated luminosity*, given in units of inverse cross-section. During Run 1 of the LHC, approximately $30\,\mathrm{fb}^{-1}$ of pp collisions were produced for the GPDs, whilst for LHCb

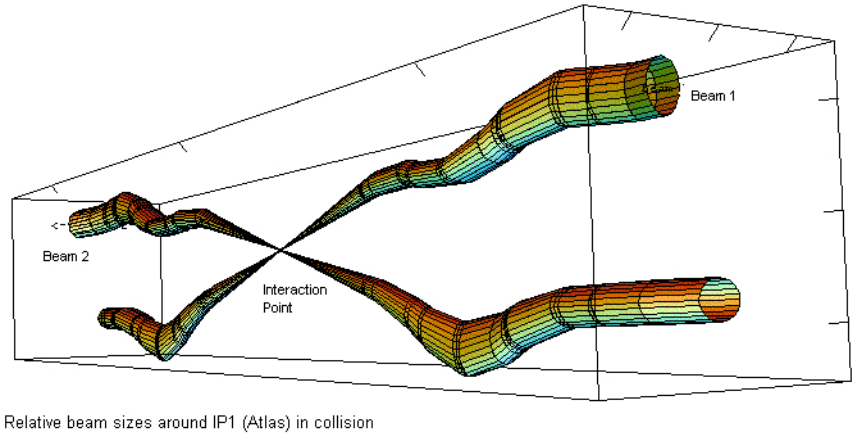


Figure 3.3.: Diagram to display the paths and relative sizes of the two colliding proton beams at the interaction point within the ATLAS detector [90].

approximately 3 fb^{-1} were produced. This is due to the fact that LHCb operates at a lower luminosity, with the colliding beams slightly de-centered. This gives fewer interactions per bunch crossing, allowing more precise measurements to be made in the region near the beam line.

The LHC's design collision energy of $\sqrt{s} = 14\,\text{TeV}$ corresponds to counter-rotating beams containing protons of energy 7 TeV, and thus travelling within approximately 3 ms^{-1} of the speed of light. The beams are steered and focused using 1232 helium-cooled superconducting magnets, which operate at temperatures below 2 K and produced fields above 8 T. The two counter-rotating beam paths are shown in Figure 3.2. Before colliding, the two beams are focused to a diameter of approximately 14 µm using quadrupole magnets. The machine is designed to circulate beams consisting of approximately 2800 bunches, each containing 1.1×10^{11} protons. These bunches circulate with a spacing of 25 ns, giving a 'bunch crossing rate' of 40 MHz. The beam paths at a typical interaction point are shown in Figure 3.3.

The LHC running schedule is divided into data taking runs of approximately four years, separated by 'long shutdown' periods of one or two years, along with more regular 'technical stops' of a few months. This is necessary to allow hardware to be maintained and upgraded, and also offers the benefit that systems can be regularly re-optimised, as their performance becomes better understood. During Run 1, the LHC operated well below its design energy. This was a consequence of an initial setback at the start-up of the machine, in September 2008, when a helium leak caused an explosion which damaged the superconducting magnet system. As a result of this incident the machine did not collide protons until March 2010, and then only at $\sqrt{s} = 7\,\text{TeV}$. In 2012, this energy was increased to $\sqrt{s} = 8\,\text{TeV}$. In addition to this reduced energy, the bunch spacing within the machine was doubled to 50 ns, halving the bunch crossing rate. To recoup some of the lost luminosity resulting from these changes, the number of protons per bunch was increased to 1.6×10^{11}. This yielded a peak luminosity approximately 77% of the design value, but increased the number of pp interactions per bunch crossing to around 30 for the GPDs; approximately twice the design value; making reconstruction of events more difficult.

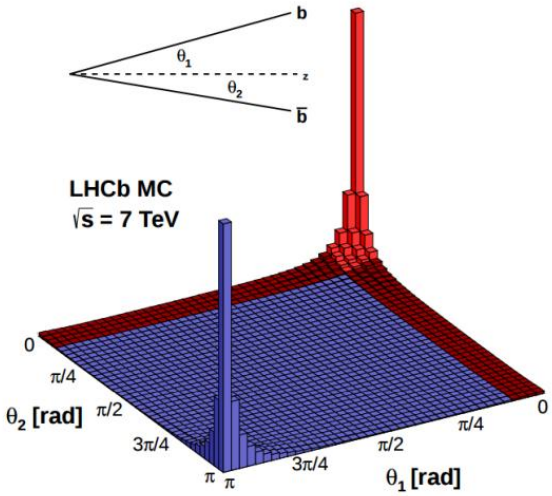


Figure 3.4.: Angular distribution of the $b\bar{b}$ production cross-section, $\sigma_{b\bar{b}}$ [92]. Angles are defined relative to the beamline. It is evident that $b\bar{b}$ pairs are preferentially produced with a boost into the forward ($\theta \approx 0$) or backward ($\theta \approx \pi$) region. The region lying within the acceptance of LHCb is coloured red, with light red indicating the region for which both beauty particles are within the acceptance.

Run 2 commenced in June 2015, with a push towards the original design energy of the machine. Five months of pp collision data were gathered, with the machine operating at $\sqrt{s} = 13\,\text{TeV}$, with a bunch spacing of 25 ns. At the time of writing the next milestone in the run schedule is the second long shutdown, LS2, planned to run from December 2018 to March 2021. The third run period, which follows this shutdown, will maintain similar run conditions, with the intention that it can be used to test upgraded hardware designed for the 'high luminosity era' of the LHC, due to commence in 2026 with the fourth run period of the machine.

3.2. The LHCb experiment

The LHCb detector lies at point 8 of the LHC ring within the cavern originally built to house the DELPHI experiment during the LEP era. It is optimised to make precision measurement of B- and D-meson decays, both to study $\mathcal{CP}$ violation in these systems and to search for the effects of New Physics in rare decays [91].

The GPDs aim to instrument as much of the solid angle surrounding their interaction point as possible, since they value the ability to measure missing energy and momentum for new particle searches. In contrast, LHCb is a single-arm forward spectrometer. Only the $\sim 4\%$ of solid angle nearest to the beam pipe, in a single direction, is instrumented, since this is the region where the $b\bar{b}$ production cross-section is largest. As shown in Figure 3.4, B-hadron pairs are predominantly produced with a significant boost in the same direction, either 'forwards' ($\theta \approx 0$) or 'backwards' ($\theta \approx \pi$).

The LHCb detector has a wedged structure, with subdetectors stacked in layers perpendicular to the beampipe in the forward region, as shown in Figure 3.5. The two proton beams collide at the 'interaction point', which lies within the VELO (Vertex Locator) at the left-hand-side of the diagram.

This subdetector is responsible for reconstructing the decay vertices of any b- or c-hadrons produced in the initial pp interaction. The particles originating from these decays then pass through a pair of Ring-Imaging Cherenkov (RICH) detectors and a series of tracking planes, located either side of a 4 Tm magnet. Finally, the particles deposit their energy in the electromagnetic and hadronic calorimeters, and the muon stations. A photograph of the LHCb detector, in situ within its cavern, is shown in Figure 3.6. Later subsections describe each of these subdetectors in turn.

The detector has a length of approximately 20 m, with a cross-section of 10×13 m at the muon stations. The coordinate system is defined such that the z-axis points along the beam pipe, from the VELO to the muon stations, the y-coordinate corresponds to the vertical dimension, and the x-axis points towards the centre of the LHC ring. The origin is defined as the centre of the pp interaction region, within the VELO subdetector. LHCb has an angular acceptance of 10-300 mrad in the bending plane of the magnet (x, z), and 10-250 mrad in the non-bending plane (y, z). The lack of acceptance in the central region is due to the beam pipe.

Angular acceptance is often quoted in terms of pseudorapidity, η, since this is the variable typically used to describe the angular dependence of particle probability distributions. It is useful to first define particle rapidity,

$$y \equiv \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{3.2}$$

where E gives the energy of the particle and p_z its momentum along the beam axis. Pseudorapidity is then defined as the rapidity obtained for a massless particle,

$$\eta = \frac{1}{2} \ln \left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z} \right) = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right), \tag{3.3}$$

where $\mathbf{p}$ is the particle's 3-momentum and θ its angle relative to the beam axis. Using this variable, the acceptance of the LHCb detector is $2 < \eta < 5$. This complements the GPDs, which are instrumented at lower η (for example, the inner detector of ATLAS has acceptance $\eta < 2.5$).

The LHCb beam pipe is assembled from four separate conical sections. The three sections nearest to the interaction point are fabricated from beryllium, due to its high efficiency of transmission for a range of particles. However, due to its toxicity, fragility and cost, it is not used for the fourth, which has the greatest diameter. Stainless steel is used for this section instead. The vacuum system of the LHCb beam pipe is isolated from that of the LHC, to enable it to be independently removed.

During Run 1 of the LHC, an integrated luminosity of 1.26 fb^{-1} was delivered to LHCb at a centre-of-mass energy of 7 TeV (2010 and 2011), and 2.21 fb^{-1} was delivered at a centre-of-mass energy of 8 TeV (2012). The cumulative growth of this integrated luminosity with time is displayed in Figure 3.7. LHCb recorded collision data with an average efficiency of 91% in 2011 and 94% in 2012, with dominant inefficiencies arising in the data acquisition system and the dead time of the readout chain [94].

By design, the integrated luminosity gathered by LHCb is significantly lower than that of the GPDs. This is because LHCb defocusses the interacting proton beams, and aligns them off-centre, to give a sub-maximal overlap of the beam cross-sections. This is done to avoid 'pile-up' - instances of multiple

pp interactions occurring during a single bunch crossing. Uniquely at the LHC, LHCb is optimised for precise reconstruction of individual hadron tracks close to the beam pipe, and multiple interactions serve to make this reconstruction very challenging. To combat pile-up, LHCb runs at a relatively low peak instantaneous luminosity, operating at $\mathcal{L} = 4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ during 2012. Since LHCb operates at sub-maximal luminosity, it is possible to keep the luminosity fairly constant for the duration of each fill. This luminosity levelling is achieved by gradually increasing the transverse overlap of the colliding beams, to counteract the decreasing population within their bunches. The luminosity during a typical fill is displayed for LHCb and the GPDs in Figure 3.8. The luminosity at LHCb is kept constant to within 5% for the first 15 hours of the fill. Since the peak luminosity experienced within the LHCb detector is less than that within the GPDs, the radiation damage to its components is correspondingly lower.

In order to conteract the 50 ns bunch spacing provided by the LHC during Run 1, LHCb was run at a luminosity more than double its design value. During 2012, the average number of interactions per bunch crossing was approximately 1.8, around 2.5 times greater than the design value of 0.7, as shown in Figure 3.9. Impressively, LHCb's performance was comparable to its design specification, despite the non-negligible pile-up.

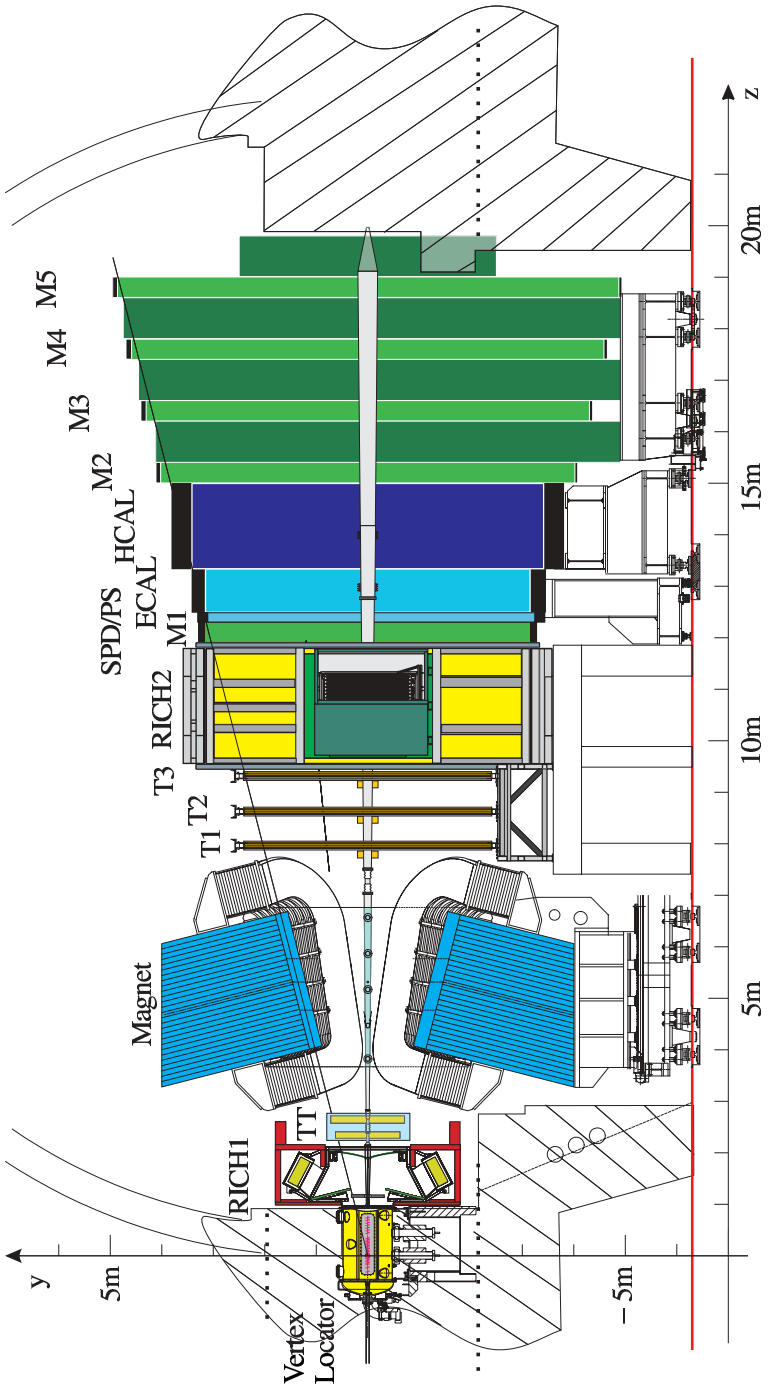


Figure 3.5.: Schematic diagram of a cross-section through the LHCb detector [93]. The colliding proton beams travel along the z-axis, and interact near the origin, within the Vertex Locator. The x-axis is defined to point towards the centre of the LHC ring, which for this schematic is into the page.

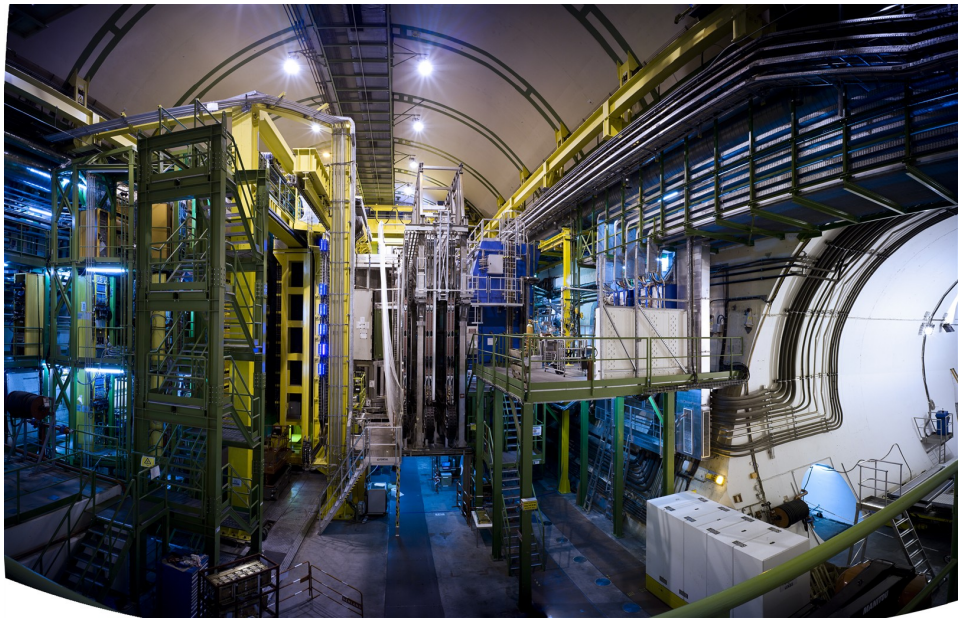


Figure 3.6.: Photograph of the LHCb detector in situ at point 8 of the LHC. Note that the z-axis runs from right to left in this photograph, so that the interaction point is located on the right-hand-side.

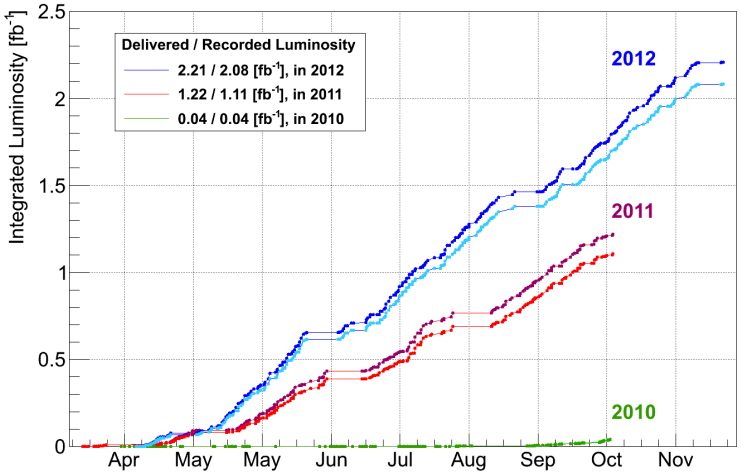


Figure 3.7.: Integrated luminosity delivered (darker shading) and recorded (lighter shading) for LHCb throughout Run 1 of the LHC [94].

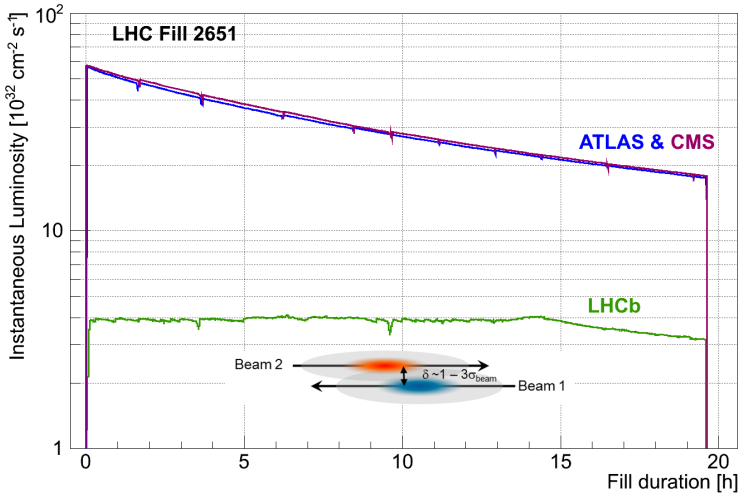


Figure 3.8.: Instantaneous luminosity profile over the duration of a typical fill of the LHC for LHCb and the GPDs. Luminosity levelling allows the luminosity at LHCb to be kept constant to within 5% for the first 15 hours of each fill [94].

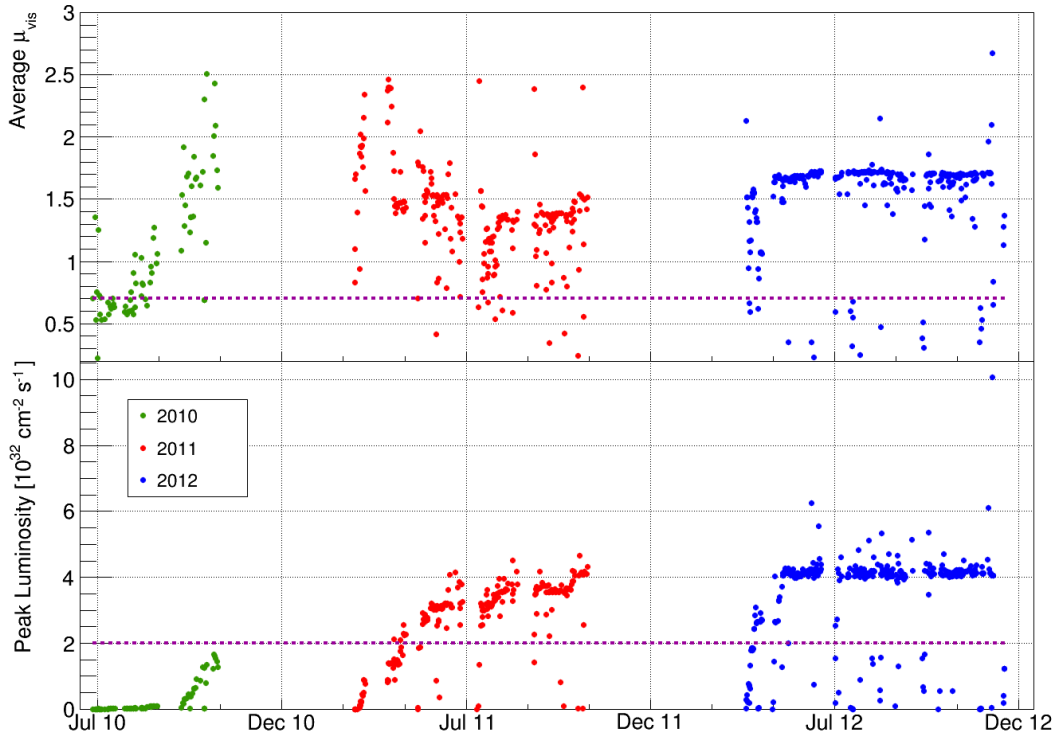


Figure 3.9.: Average pile-up (μ_{vis}) and peak instantaneous luminosity observed during each Run 1 fill at LHCb. Recorded values are consistently above the detector's design specifications, plotted as dotted lines [94].

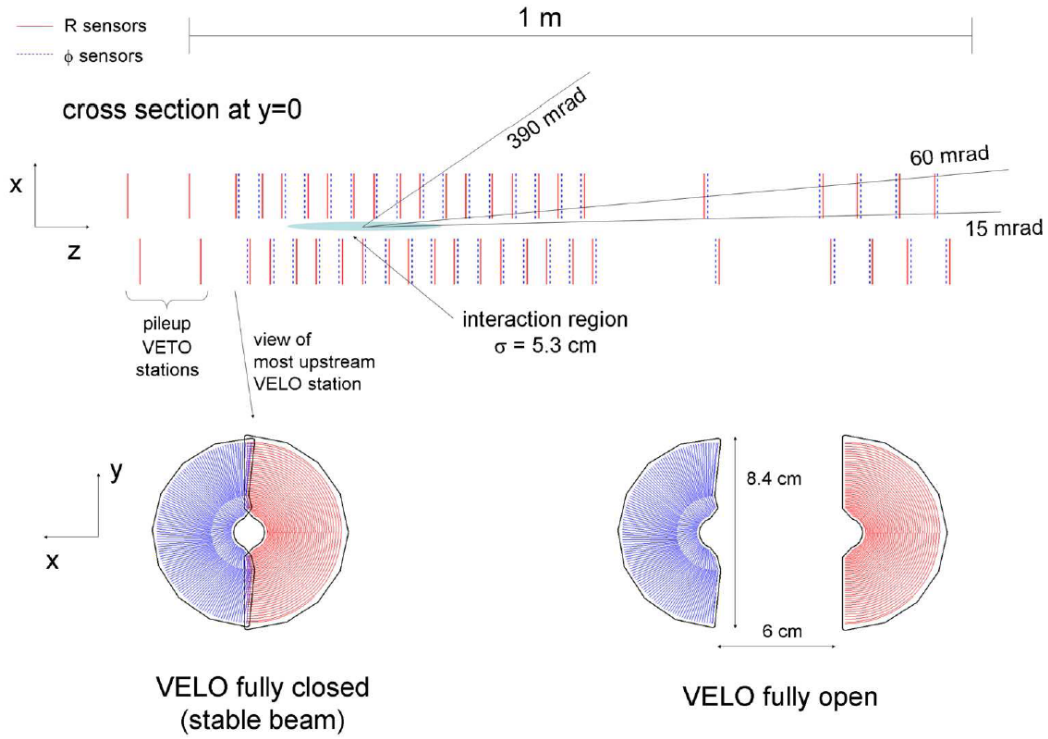


Figure 3.10.: Schematic diagram of the VELO subdetector, which surrounds the LHCb pp interaction point [93]. The upper diagram shows a top-down view of the distribution of VELO modules along the beam line. R-sensor modules are indicated in red, and ϕ-sensor modules in blue. The lower diagram shows a view of two modules, as seen when looking along the beam line, for both the closed and open VELO configurations.

3.2.1. Vertex Locator

The VELO is a silicon tracking detector that closely surrounds the LHCb pp interaction region. It is designed to allow for precise tracking of individual charged particles near the interaction region and the accurate and reliable reconstruction of the decay vertices of b- and c-hadrons. The VELO consists of 42 modules, paired into 21 tracking stations, distributed as shown in Figure 3.10. Each module is semi-circular in shape, and lies perpendicular to the beam on one side. A small semi-circular cutout at each module's centre allows the beam to pass through, as can be seen in Figure 3.11. The z-coordinates chosen for the modules ensure that the entire LHCb acceptance is covered, with each possible particle trajectory traversing at least three VELO stations. Two additional stations are installed upstream of the pp interaction region, to provide a veto against events with an unmanageable level of pile-up.

B-mesons have a mean lifetime of around 1.5 ps [3], and at LHC energies decay over a distance of $\mathcal{O}(1\,\text{cm})$. Since the role of the VELO is to track such mesons, it must be instrumented as close as possible to the pp interaction region. To achieve this, whilst avoiding subjecting the hardware to radiation damage during unstable beam periods, the VELO is able to move between two different configurations. In the 'closed' position, used during data-taking runs, the edges of the VELO modules lie just 5 mm from the beam axis, with the sensitive region beginning at a radius of 8 mm. In the 'open' position, used at the start of each LHC fill, when the beams are unstable, the modules are retracted by 3 cm. The

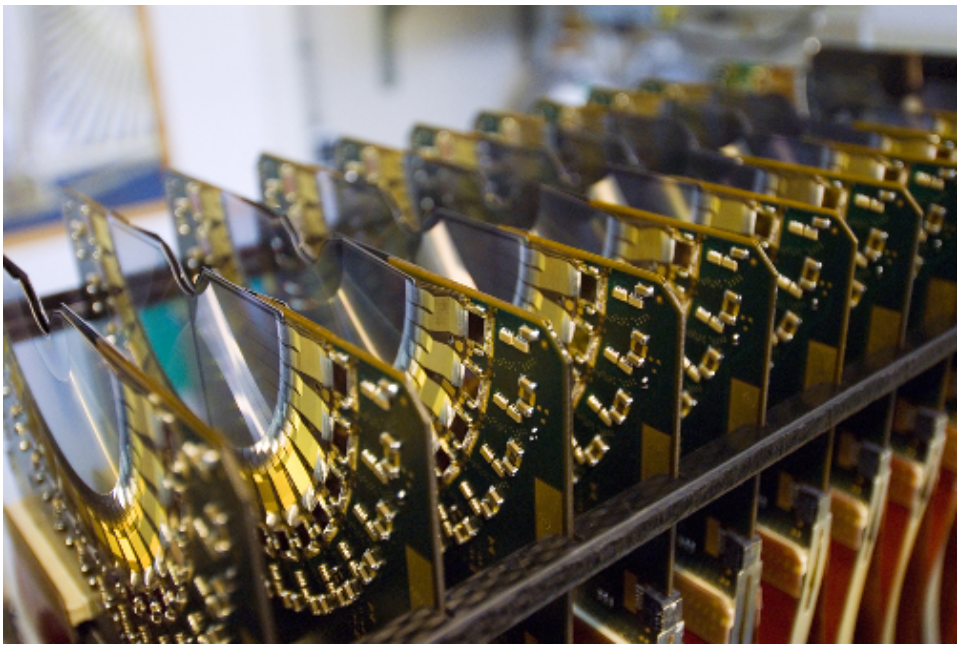


Figure 3.11.: Photograph of VELO modules at LHCb.

VELO modules overlap slightly in the closed position, to allow their relative alignment to be monitored and adjusted.

The VELO is operated in a vacuum in order to minimise the amount of material present in the detector acceptance. A vacuum vessel surrounds the entire subdetector, which is isolated from the LHC vacuum by aluminium RF-boxes, used to shield the VELO electronics from the electromagnetic fields produced by the LHC beams. The layout of the sensors within these boxes is shown in Figure 3.12.

Each module is comprised of one R-sensor and one ϕ-sensor. For each particle that leaves a 'hit' in the module, the ϕ-sensor records the azimuthal position of the hit and the R-sensor provides its radial distance relative to the beam axis, with the third coordinate provided by the module's position along the beam line. This coordinate system is chosen for the VELO since it is the one used by the fast vertex reconstruction algorithm used to trigger on events containing b-hadrons.

The geometry of the two sensor types is shown in Figure 3.13. Each R-sensor is divided into four sections of 45°, each containing 512 concentric silicon strips, centred on the beam line. The separation between the centres of adjacent strips, or *pitch*, ranges from 38 µm at the edge nearest the beam pipe, to 101.6 µm at the circumference, ensuring that similar occupancies are experienced by all strips, so that they each contribute with similar statistical weight to the impact parameter resolution of the VELO. Since the ϕ-sensors measure the azimuthal position of hits, a simple design could make use of radial strips lying orthogonal to those of the R-sensors. However, several sophistications are introduced to improve performance relative to this first approximation. Firstly, each ϕ-sensor is divided into two concentric regions, with an inner region extending from the inner sensitive radius (8 mm) to a radius of 17.25 mm, and an outer region extending from that radius to the outer circumference, 41.9 mm. The inner region contains 683 strips, with their pitch varying linearly from 37.7 µm at the inside to 78.3 µm at the outside. The outer region contains 1365 strips, with pitch varying linearly from 39.3 µm to 96.9 µm.

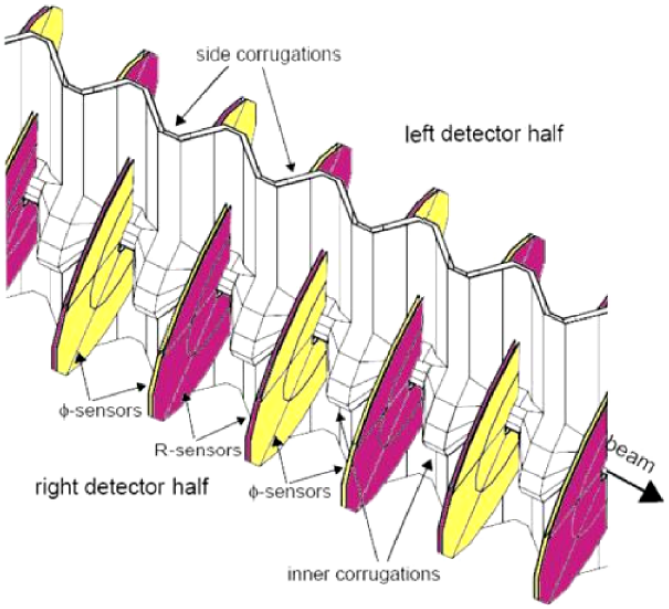


Figure 3.12.: Internal view of a VELO RF-box, showing the R- and ϕ-sensors and the RF-foil [93].

This design allows the occupancy in the inner strips to be reduced, whilst sensitivity is retained in the lower occupancy outer region. Pattern recognition is improved by skewing the strips by approximately 20° relative to the radial in the inner region, and by approximately 10° in the outer region, with the direction of the skew alternated in adjacent ϕ-sensors.

Vertex reconstruction in the VELO is achieved with a resolution of 10-25 μm in the directions transverse to the beam, and 50-100 μm parallel to the beam. The resolution varies with the number of tracks originating from the vertex, as shown in Figure 3.14 for the reconstruction of pp primary vertices (PVs). The VELO is also able to measure impact parameters of tracks relative to these vertices, with the resolution in 2012 data shown in Figure 3.15. A typical impact parameter resolution of 35 μm is obtained for a track with transverse momentum of $p_{\mathrm{T}} = 1\,\mathrm{GeV}/c$.

3.2.2. Tracking stations

The VELO forms the first part of the LHCb tracking system. In addition, there are four other silicon tracking stations within LHCb. The Tracker Turicensis (TT), is situated immediately downstream of the RICH1 subdetector, and three further tracking stations, T1-3, are located downstream of the magnet.

Tracks detected within LHCb are classified depending on which tracking stations are hit, as shown in Figure 3.16. For the present analysis it is important to distinguish between *long* and *downstream* tracks, since these will be treated separately when selecting candidate events. Long tracks leave hits in all of the tracking stations, including the VELO, and are thus more precisely measured than downstream tracks, for which no VELO hits are recorded. Downstream tracks can arise from daughter particles originating from a relatively long-lived meson, which decays downstream of the VELO.

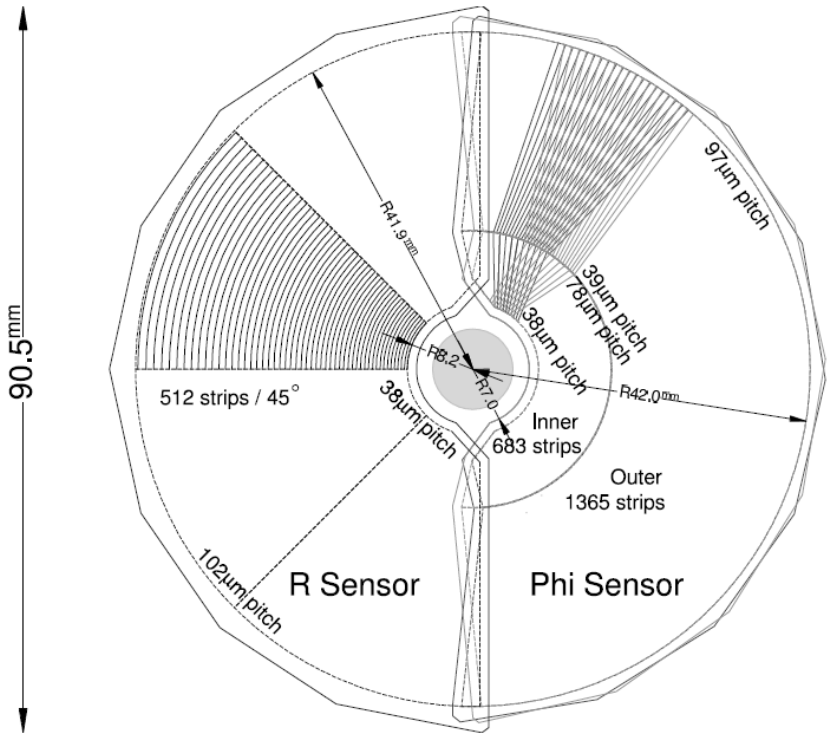
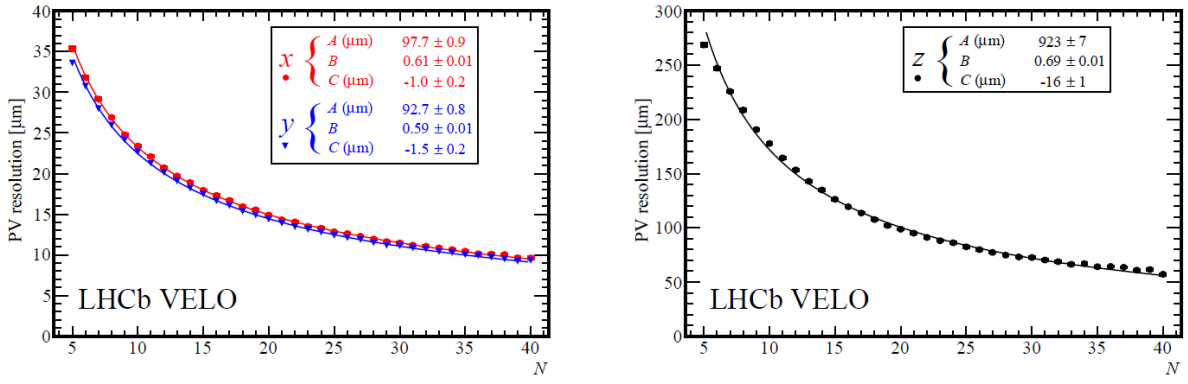


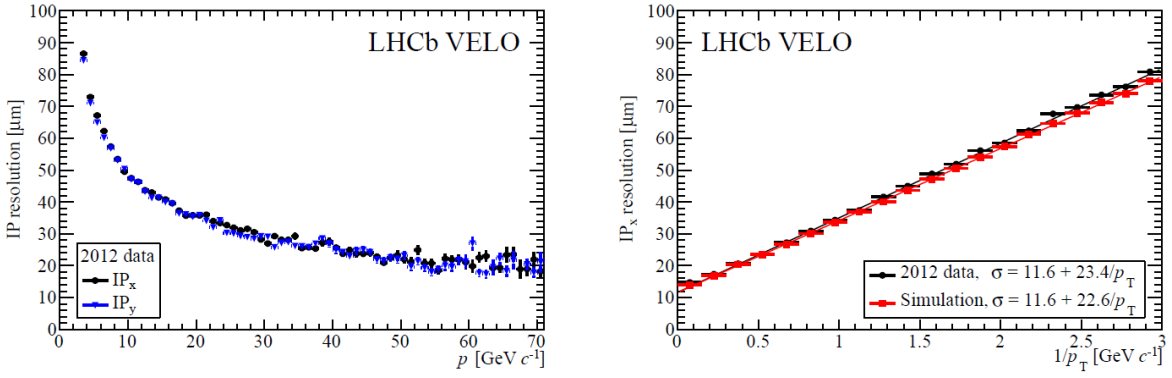
Figure 3.13.: Schematic diagram of the R- and ϕ-sensors used in the VELO subdetector [93]. The strips of two adjacent ϕ-sensors are displayed to indicate their alternating skew.

The TT uses four layers of silicon strips, with a pitch of approximately 200 µm, to cover the full LHCb acceptance, and is 150 cm wide and 130 cm tall, giving a total active area of about 8.4 m^2. The two middle layers are rotated by angles of $-5°$ and $+5°$ relative to the outermost layers, since the



(a) **(b)**

Figure 3.14.: Resolution obtained for primary vertex reconstruction in the VELO during 2011 [95]. The variation of resolution with the number of tracks, N, associated with the reconstructed vertex is shown for (a) the plane transverse to the beam axis and (b) the z direction. Functions of the form $A/N^B + C$ are fitted to the observations.



(a) (b)

Figure 3.15.: Resolution obtained for impact parameter reconstruction in the VELO during 2012 [95]. Resolution in the plane transverse to the beam axis is shown as a function of (a) track momentum, and (b) the inverse of the transverse track momentum.

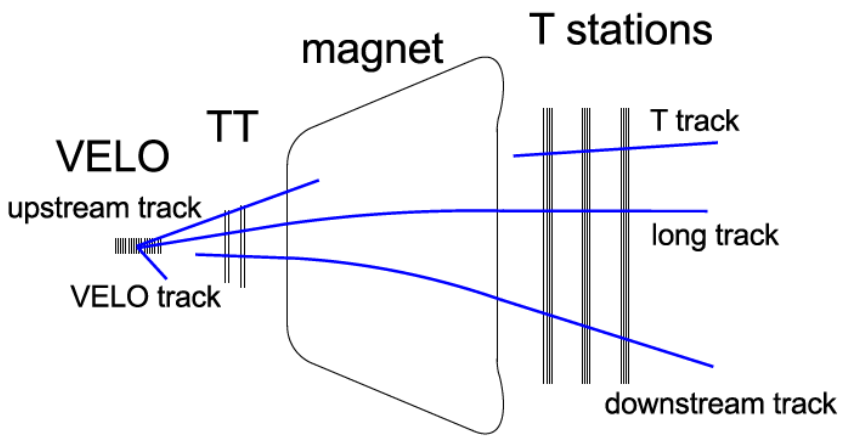


Figure 3.16.: Track categories defined at LHCb, depending on whether hits are detected in the VELO, TT and downstream tracking stations [96].

resulting stereo effect improves the performance of the track reconstruction algorithms. Particles with momenta below 1.5 GeV/c are bent out of the LHCb acceptance by the magnet between the TT and the downstream tracking stations, and so the TT is especially important in the reconstruction of these low momentum 'upstream' tracks.

Each downstream tracking station, T1-3, is divided into two parts, an Inner Tracker (IT) and an Outer Tracker (OT). As in the TT, these each use a four-layer structure with stereo-rotated detector planes. Each IT consists of four separate boxes, arranged in a cross configuration, as shown in Figure 3.17. In total, the four boxes occupy a 120×40 cm region around the beam pipe. The boxes are instrumented with planes of silicon strips, giving an active area of 4.0 m^2. The inner trackers cover only 2% of the total area of the T1-3 stations, but are intersected by approximately 20% of detected tracks. They provide a spatial resolution below 50 μm.

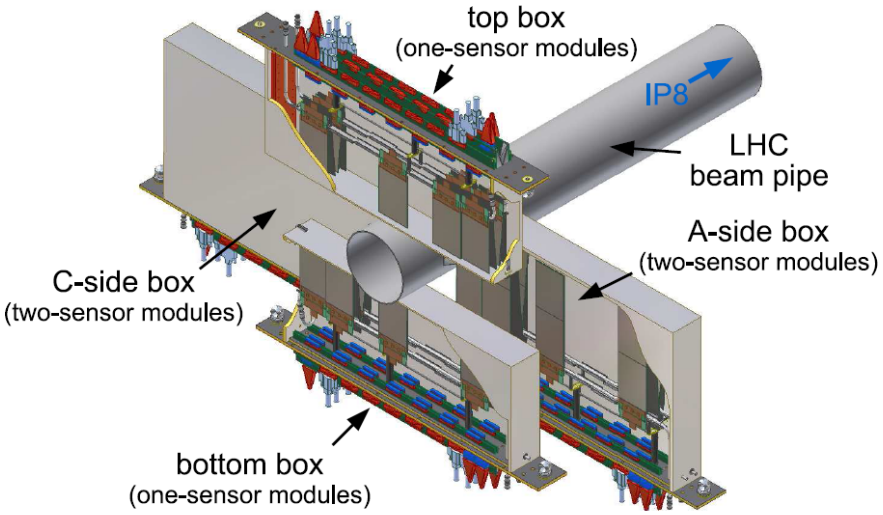


Figure 3.17.: An Inner Tracker subdetector at LHCb. The silicon strip detector planes are housed in four separate boxes, arranged in a cross configuration around the beam pipe [93].

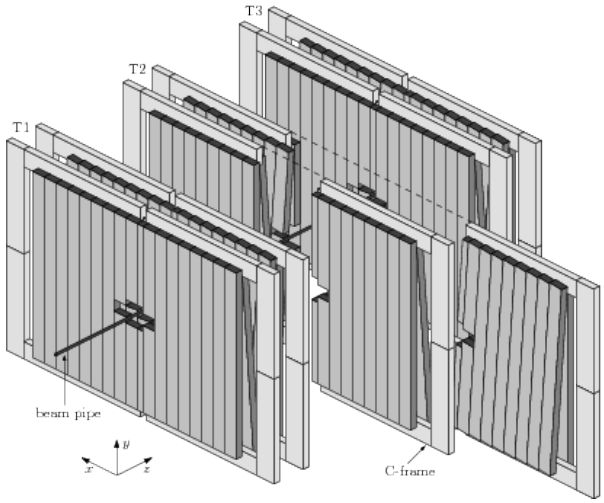


Figure 3.18.: Schematic diagram to show the configuration of the three Outer Trackers of LHCb. Each Outer tracker consists of 72 modules of double-monolayer straw tubes, divided between four independently moving support frames [97]. The region normally occupied by each Inner Tracker may be seen at the centre of each OT, nearest the beampipe.

The majority of the area covered by each tracking station T1-3 is occupied by its OT. In the interest of economy, silicon strips are eschewed here in favour of an array of straw-tube modules. Each module contains two layers of 2.4 m long straw tubes of inner diameter 4.9 mm. The straws are filled with a mixture of argon and carbon dioxide, and act as drift-time detectors. They are designed to provide a spatial drift-coordinate resolution of approximately 200 µm, thanks to their short drift time, which is below 50 ns. Each OT consists of 72 modules arranged in 4 layers of 18, supported by four independently moving frames, as shown in Figure 3.18. The total active area of all OTs combined is approximately 6×5 m, and during Run 1 a hit detection efficiency of 99.2% was obtained [97], with an average spatial resolution of 205 µm.

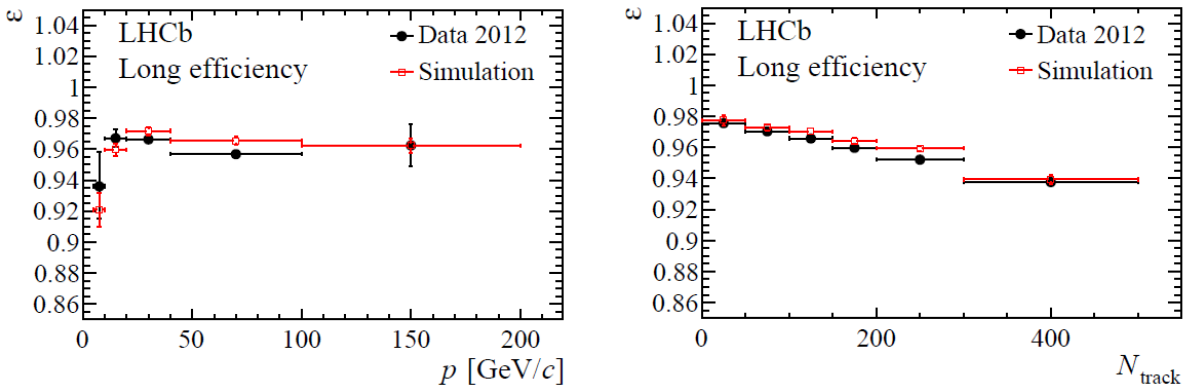


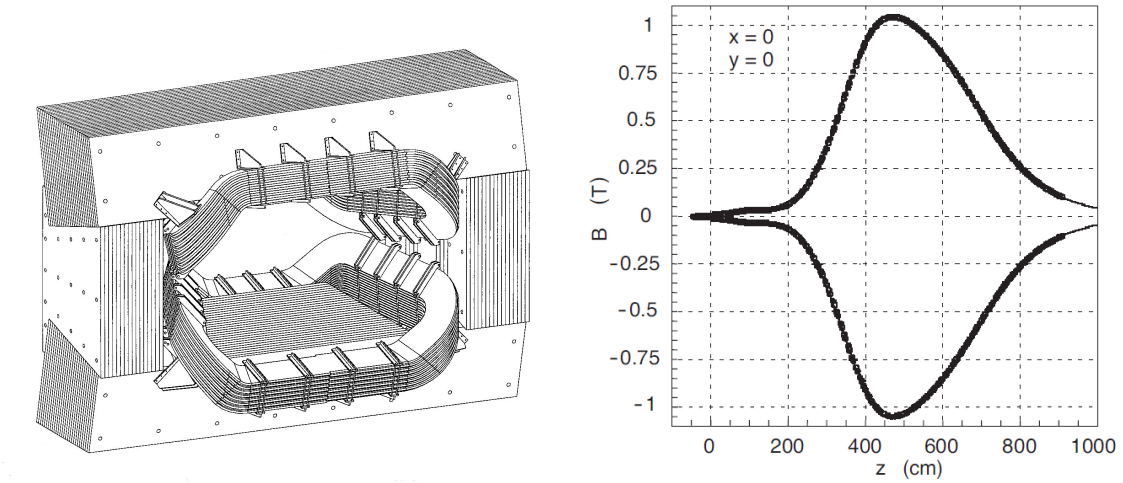
Figure 3.19.: Reconstruction efficiency for the full LHCb tracking system, obtained using 2012 data and simulation, for the case of long tracks, determined as a function of (a) track momentum, and (b) the number of tracks reconstructed in the event [96].

Data from the entire LHCb tracking system is taken together when reconstructing particle tracks through the detector. As demonstrated pictorially in Figure 3.16, the distance from the VELO to the final tracking station, T3, provides a significant lever-arm for track measurements. As a result, the reconstruction efficiency and momentum resolution when combining all subdetectors exceeds that of any single subdetector, especially for long tracks, which leave hits in both the VELO and the downstream tracking stations. In 2012, the reconstruction efficiency for long tracks was over 96% on average [96]. The variation of reconstruction efficiency with track momentum and the total number of tracks in an event is shown in Figure 3.19. The combined LHCb tracking system has a design momentum resolution of $\frac{\delta p}{p} < 0.4\%$, and this has been measured to lie in a range from 0.5% at low momentum, to 1.0% for tracks with $p = 200\,\text{GeV}/c$ [94].

3.2.3. Dipole magnet

LHCb uses a dipole magnet to bend the trajectories of charged particles passing through the detector, to allow their momenta to be inferred from the resulting curvature. It is a warm magnet built from two trapezoidal coils, each weighing 27 tonnes, situated inside a 1500-tonne iron yoke [98]. The magnet encloses the full LHCb acceptance, up to $\pm 300\,\text{mrad}$ horizontally and $\pm 250\,\text{mrad}$ vertically. The field is strongest in the region directly between the coils, which lie between the TT and T1 tracking stations. The peak field produced is 1.1 T, and for a track of length 10 m the integrated magnetic field is $\int B\text{dl} = 4\,\text{Tm}$. In order for momentum measurements to be correctly calibrated, the magnet properties must be accurately measured. An array of Hall probes is used to determine the value of $\int B\text{dl}$ to a precision of 10^{-4}, and the location of peak field strength to within a few millimetres [93]. The variation of the magnetic field magnitude, with z-position within the detector, is displayed in Figure 3.20, alongside a schematic diagram of the magnet assembly.

A crucial feature of the LHCb magnet is the ability to reverse its polarity. This allows systematic effects due to detector asymmetries to be studied and quantified, since in each field configuration



(a) The dipole magnet of LHCb

(b) Variation of the LHCb magnetic field with z-position within the detector

Figure 3.20.: Schematic diagram of LHCb's dipole magnet, and the variation of its magnetic field along the beam axis [93].

particles of a given charge are bent towards opposite sides of the detector. The two polarities are referred to as 'up' and 'down', such that 'up' corresponds to the magnetic field pointing in the positive y direction, bending positive particles towards negative x (away from the center of the LHC ring). Magnet polarity is alternated on a roughly fortnightly schedule. In 2011, the magnet was in the 'down' polarisation for approximately 56% of the data taken, whilst for 2012 it was 52%.

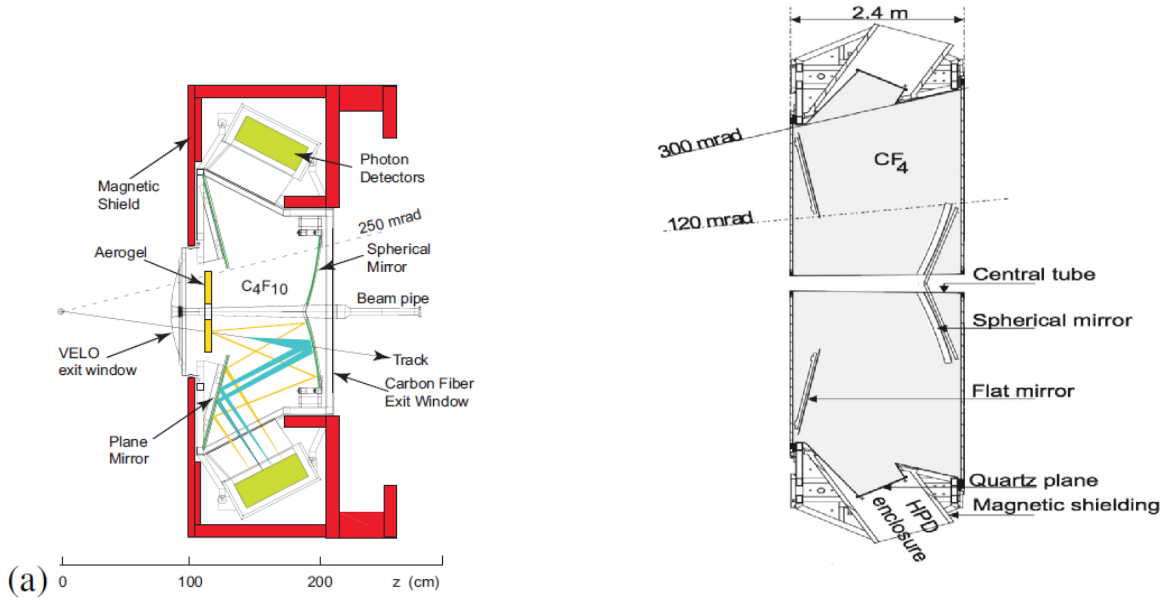
3.2.4. Ring-Imaging Cherenkov detectors

The Ring-Imaging Cherenkov (RICH) subsystem provides the primary means to distinguish between different species of charged hadrons at LHCb. It makes use of the phenomenon of Cherenkov radiation, which is emitted by a charged particle whenever it travels through a medium faster than the local speed of light. The radiation is emitted at a characteristic angle, θ, relative to the particle trajectory, and is a function of the particle's speed

$$\cos\theta = \frac{c}{vn}, \tag{3.4}$$

where c gives the speed of light in a vacuum, n is the refractive index of the medium, and v is the speed of the charged particle. By measuring the angular distribution of Cherenkov radiation emitted within a radiator of known refractive index, the particle's speed may be inferred. This can then be combined with a momentum measurement from the tracking system to determine the mass, and thus the identity, of the particle.

The RICH subsystem consists of two separate subdetectors. RICH1 is located between the VELO and the LHCb magnet; RICH2 lies downstream of the magnet and tracking stations. RICH1 uses a C_4F_{10} flourocarbon gas radiator, with a refractive index of 1.0014. The minimum momentum required for a



(a) RICH1 **(b)** RICH2

Figure 3.21.: Schematic diagrams of RICH1, viewed from the side, and RICH2, viewed from above [93].

kaon to produce Cherenkov emission in this medium is 9.3 GeV/c. For the duration of Run 1, a 50 mm thick block of silica aerogel was positioned at the entrance of RICH1. This material has a refractive index of 1.03, extending the momentum range for effective kaon identification down to roughly 2 GeV/c. RICH1 is also responsible for identifying very low-momentum particles (<1.5 GeV/c) before they are bent outside the LHCb acceptance by the magnet. RICH2 uses a CF_4 gas radiator, giving a refractive index of 1.0004, and is doped with 5% carbon dioxide to quench scintillations. This refractive index is optimised for the identification of particles with momenta between 15 and 100 GeV/c. RICH2 does not instrument the full range of the LHCb acceptance. Instead it covers the most central region, which is where high-momentum decay products are most likely to be found: 15-120 mrad in the horizontal (x, z) plane; and 15-100 mrad in the vertical plane.

In order to minimise the amount of material lying within the particle paths, both RICH subdetectors use lightweight spherical mirror assemblies to reflect emitted Cherenkov photons out of the LHCb acceptance. The photons are further reflected by a flat mirror, onto planes of photon detectors. The use of a spherical mirror transforms the solid disc of photons emitted by a charged particle continuously over its path, from entering the RICH subdetector to reaching the mirror surface, into a ring of photons in the detection plane. The Cherenkov angle during the emission may be determined from the radius of this ring. Each of the RICH subdetectors is divided into two halves around the beampipe, with RICH1 divided horizontally into upper and lower halves, and RICH2 vertically into A-side (the half lying nearest the centre of the LHC ring) and C-side. The RICH system is displayed schematically in Figure 3.21.

The RICH detection planes are instrumented with Hybrid Photon Detectors (HPDs), which are capable of detecting single Cherenkov photons of wavelength 200-600 nm. Each HPD consists of a

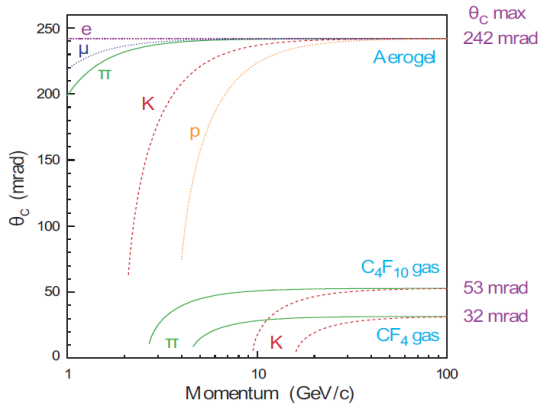


Figure 3.22.: Calculated Cherenkov angles expected for various particle types, as a function of particle momentum, for each of the radiator media used within the RICH subsystem [93]. An aerogel radiator is used to extend the range of effective identification for kaons down to $\sim 2\,\text{GeV}/c$.

vacuum tube of 75 mm diameter, with a 7 mm quartz window at one end and a silicon detector at the other. A multialkali film coats the inner surface of the entrance window, and is kept at a $-16\,\text{kV}$ potential difference relative to the silicon detector. Photons incident on the photocathode produce photoelectrons, which accelerate through this potential difference to be detected by the silicon detector, which is divided into a 32×32 array of $2.5 \times 2.5\,\text{mm}^2$ pixels. RICH1 is instrumented with 196 HPD tubes and RICH2 uses 288.

The HPD array making up each detector plane is enclosed within a shielding box to protect it from the field of the dipole magnet. The boxes are constructed from iron plates, 100 mm thick in RICH1 and 60 mm thick in RICH2. Due to this shielding, the design value for the magnetic field at the detector plane is 2.4 mT in RICH1 and 0.6 mT in RICH2, and the fields are measured during running so that they can be corrected for in the Cherenkov angle reconstruction.

The particle identification performance of the RICH subsystem is illustrated in Figure 3.22. This shows the Cherenkov angles expected for various particle types, as a function of momentum, for each of the radiator media used. The performance observed when measuring isolated tracks in 2012 data is shown in Figure 3.23.

Before data from the RICH subsystem can be used to calculate the Cherenkov angle for a particle, the direction of the particle's trajectory through the radiator must be known. This is achieved by combining RICH data with measurements from the tracking system. A further complication arises from the uncertainty of the emission point of each Cherenkov photon. As pointed out in the discussion above, a charged particle continuously emitting Cherenkov radiation would give rise to a disc of light incident on the spherical mirror. However, in practise a particle typically emits just $\mathcal{O}(10)$ photons during its traversal of each RICH detector. To simplify the calculation, all photons arising from a given particle are assumed to be emitted from the midpoint of its trajectory within the radiator medium. Using the known geometry of the mirror system and detector planes, photons detected in the HPDs are traced back and associated with charged tracks. The Cherenkov angle extracted has a resolution of approximately 1.6 mrad for C_4F_{10} and 0.7 mrad for CF_4 [99].

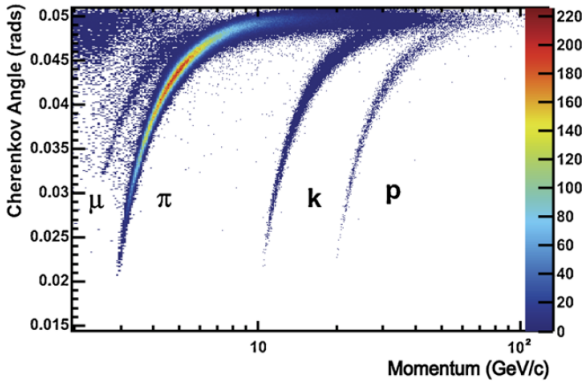


Figure 3.23.: Cherenkov angles measured in the RICH1 gas radiator, C_4F_{10}, for isolated tracks in 2012 data, as a function of particle momentum [99]. The particle identification power of the RICH subsystem is evident from the distinct bands observed for each particle species.

Since many charged hadrons are produced in each pp collision, the occupancy of the RICH1 subdetector is high, with many overlapping Cherenkov rings that must be distinguished. A likelihood-maximising algorithm is used to assign the identity of each particle track. The algorithm considers all tracks in the event together, and combines data from both RICH subdetectors. It proceeds as follows:

1. take as a starting point the hypothesis that every track in the event arise from a pion, and calculate total likelihood of the observed RICH and tracking hits under this assumption;
2. select one of the tracks, and consider each particle mass hypothesis in turn (e^-, μ^-, π^+, K^+, p), recalculating the likelihood in each case. All other track hypotheses are kept unchanged;
3. fix the hypothesis for this track to the one that gives the greatest increase in calculated likelihood; and
4. repeat this procedure until the optimal hypothesis has been reached for all tracks in the event, and no further improvement can be found.

The likelihoods calculated in this process are often used in analysis to select candidate events. By constructing the difference in log-likelihood between two separate mass hypotheses, variables known as delta-log-likelihoods (DLLs) may be calculated. For each decay mode studied in this thesis, the final state consists of just kaons and pions, so the variable

$$\mathrm{DLL}_{K\pi} = \ln\left(\frac{\text{Likelihood calculated with kaon hypothesis}}{\text{Likelihood calculated with pion hypothesis}}\right) \tag{3.5}$$

is an extremely useful discriminator. Positive values for this DLL indicate tracks that match the kaon hypothesis better than the pion hypothesis.

RICH performance may be evaluated using pure samples of protons, pions and kaons from data. The decay channels $D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$, $\Lambda \to p\pi^-$ and $K^0_S \to \pi^+\pi^-$ are used, with purely kinematic cuts used to purify the samples. Performance has been found to match the design specification, with the criterion $\mathrm{DLL}_{K\pi} > 0$ yielding a kaon identification efficiency of 95% in 2011 data, averaging over

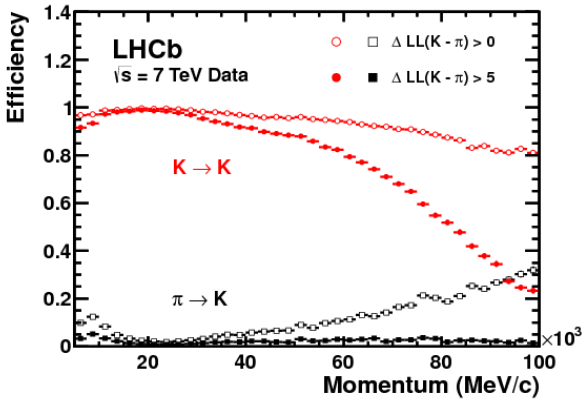


Figure 3.24.: Efficiency of kaon identification and probability of pion mis-identification, using information from the RICH subsystem for 2011 data. Two alternative criteria on the value of $\mathrm{DLL}_{K\pi}$ (here denoted '$\Delta LL(K-\pi)$') are considered [99].

the full accepted momentum range 2-100 GeV/c, with a pion mis-identification rate of 10%. Both of these quantities are studied as functions of particle momentum, for two alternative requirements on $\mathrm{DLL}_{K\pi}$, in Figure 3.24.

3.2.5. Calorimeters

The interactions in high energy collisions at the LHC occur directly between individual partons from within colliding protons. Since these partons may carry any fraction of the protons' energy and momentum, the total energy, E, and momentum, $\vec{p}$, present in each interaction is unknown. Instead, it is useful to define the *transverse* momentum, $\vec{p_T}$, as the momentum component in the plane perpendicular to the beam axis. This allows the transverse energy to be defined

$$E_T \equiv \sqrt{m^2 + p_T^2}. \tag{3.6}$$

These variables are often used in particle physics analyses. In particular, since the interacting partons must initially be travelling close to parallel with the beam axis, $\vec{p_T} \approx \vec{0}$. Conservation of momentum allows this to be used as a criterion within the reconstruction.

Particle energies at LHCb are measured using the calorimeter subsystem, which is comprised of four subdetectors: the Scintillating Pad Detector (SPD); Preshower Detector (PS); Electromagnetic Calorimeter (ECAL); and Hadron Calorimeter (HCAL). These serve to measure energy deposited by electrons, photons and hadrons, providing information used in the reconstruction. In addition, this subsystem is able to rapidly identify candidates with large transverse energies, on which criteria are placed in the first trigger stage, detailed fully in Section 3.2.7.

The SPD and PS subdetectors each consist of a 7.6×6.2 m plane of scintillating tiles, and are separated by 15 mm thick layer of lead. The SPD lies upstream of the lead layer whilst the PS lies downstream. Both of these subdetectors search for scintillations produced by charged particles. The thickness of the lead equates to 2.5 radiation lengths, meaning that photons traversing the layer are

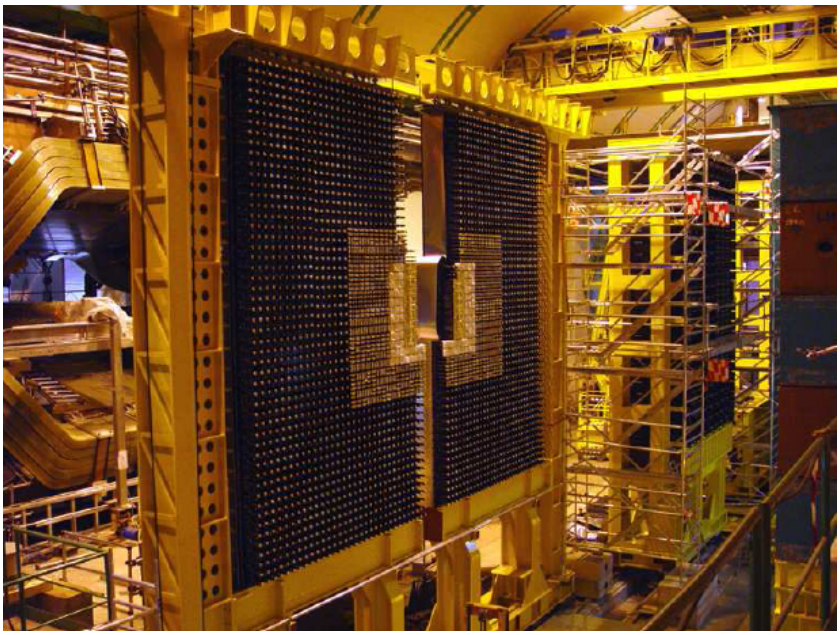


Figure 3.25.: Downstream view of the ECAL subdetector during installation.

likely to interact, producing electron-positron pairs. Thus the presence of a signal within the PS, but not within the SPD, gives a good indication of the presence of a neutral particle. This is useful in distinguishing high E_{T} neutral pions from electrons.

The ECAL uses 'shashlik' layering, with 2 mm thick lead absorber plates alternated with 4 mm thick scintillator tiles. Charged particles cause showers in the lead plates, giving rise to charged particles which interact within the scintillator tiles. The light produced by these scintillations is channelled through optical fibres to be read out. In order to accurately quantify the energy deposited, the ECAL must be long enough to contain the particle showers caused by high energy photons. To achieve this, its thickness is 25 radiation lengths. The resulting energy resolution has been measured using data from beam tests performed during the commissioning of the ECAL, and is found to be [100]

$$\frac{\sigma_E}{E} = \frac{(8.5\text{-}9.5)\%}{\sqrt{E}} \oplus 0.8\%,$$

for E in units of GeV. A photograph taken of the ECAL during its installation is displayed in Figure 3.25.

The HCAL uses scintillating tiles interspersed between iron plates, with a cross-section of $8.4 \times 6.8 \times 1.65\,\mathrm{m}$. Its thickness is limited by space constraints within the cavern housing LHCb, and is equivalent to 5.6 radiation lengths. The primary function of this subdetector is to provide measurements of hadron energies to be used in the first trigger stage, and so a low response time is more important than a precise measurement. The energy resolution of the HCAL is measured as [100]

$$\frac{\sigma_E}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%.$$

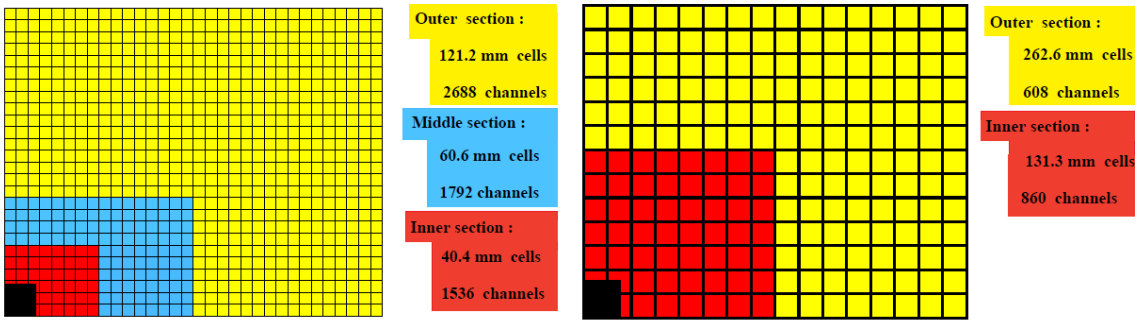


Figure 3.26.: Schematic to show the segmentation of the (left) SPD, PS and ECAL and (right) HCAL subdetectors, viewed in the plane transverse to the beam pipe. Only the upper right quadrant of the calorimeter is shown in each case, with the black shading indicating the non-instrumented region surrounding the beam pipe [93].

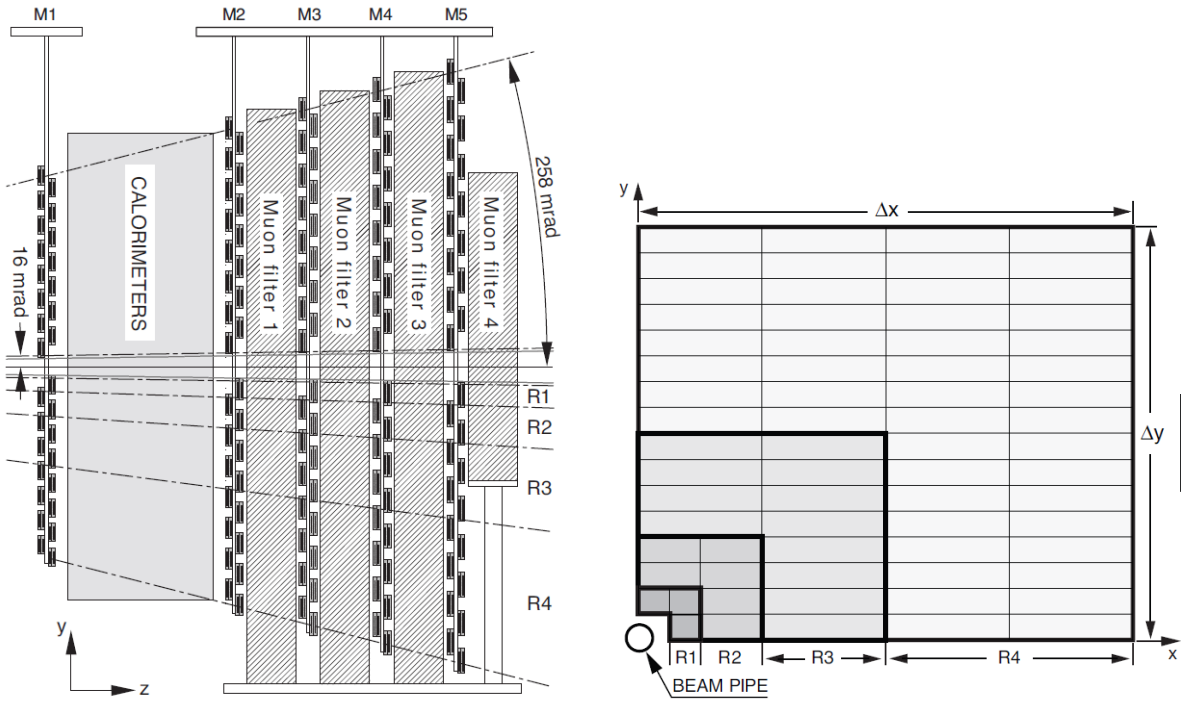
The LHCb calorimeter subsystem covers the entire acceptance, except for the most central region, 25×25 mrad, where radiation damage would outweigh the advantage of increased coverage. Since occupancy is higher nearer the beam pipe than farther away, each calorimeter uses segments of finer granularity near the center, with the HCAL segmented into two granularities, and each of the other calorimeters using three segments. This segmentation is illustrated in Figure 3.26.

3.2.6. Muon systems

Several decays of interest to LHCb produce final states containing muons. To measure these elusive particles, LHCb uses a subsystem of five muon stations, with the first (M1) situated immediately upstream of the calorimeters and the remaining four (M2-5) situated at the extreme downstream end of the detector. The full LHCb acceptance is covered, with angular ranges of 16-258 mrad in the non-bending plane (as shown in Figure 3.27) and 20-306 mrad in the bending plane of the magnet. Each muon station is subdivided into four separate regions, labelled R1-4, with a segmentation which increases with distance from the beam pipe in the ratio $1 : 2 : 4 : 16$. Roughly equal particle fluxes are expected in each of the four regions.

The most upstream of the muon stations, M1, is specialised to provide the fast response needed for its output to be used in the first stage of the LHCb trigger system; it is used to give a rough measurement of muon p_{T}. The downstream muon stations are alternated with absorber layers of 80 cm thick iron. These are used to identify and track high momentum muons, providing information useful for subsequent trigger stages and analysis. Only muons of energy greater than approximately 6 GeV/c can traverse all five muon stations.

The muon stations are predominantly instrumented with multi-wire proportional chambers (MWPCs). These consist of a gas enclosure containing a wire at a voltage of 2.8 kV. Incident muons cause ionisation of the gas, an $Ar/CO_2/CF_4$ mixture, with the ionisation products accelerated through the electric field of the wire, in which a corresponding current is induced. Each MWPC contains four separate 5 mm thick gas chambers, except for those used in M1, where only two are used to reduce



(a) Side view **(b)** Cross-sectional view

Figure 3.27.: (a) Schematic to show a side view of the full LHCb muon system. The subdivision of each muon station into four segmentation regions (R1-4) is shown. (b) Schematic diagram of the segmentation regions (R1-4) within the upper right quadrant of a muon station, viewed in the plane perpendicular to the beam axis [101].

the volume of absorber upstream of the calorimeters. In total, the muon subsystem of LHCb uses 1368 MWPCs.

The particle flux is highest in the innermost region of the upstream muon station, M1R1, and here MWPCs are eschewed in favour of more durable triple Gas Electron Multiplier (GEM) detectors. Instead of high tension wires, these detectors use three foils, separating an anode and cathode, with the gaps between layers filled with $Ar/CO_2/CF_4$ gas mixture.

The muon subsystem has a design efficiency of $> 95\%$, which requires each individual muon station to identify muons with an efficency greater than 99%. For data gathered during 2011, the total identification efficiency is measured to be over 98% for muons with $p > 3\,\text{GeV}/c$ and $p_\text{T} > 0.8\,\text{GeV}/c$, with hadrons of $p > 30\,\text{GeV}/c$ giving a mis-identification probability of less than 1%. This performance is shown in Figure 3.28.

3.2.7. The LHCb trigger

The LHC is designed to operate at a bunch crossing rate of 40 MHz, and the LHCb detector is designed to produce 0.7 pp interactions per crossing. This would imply that the detector ought to be read out at

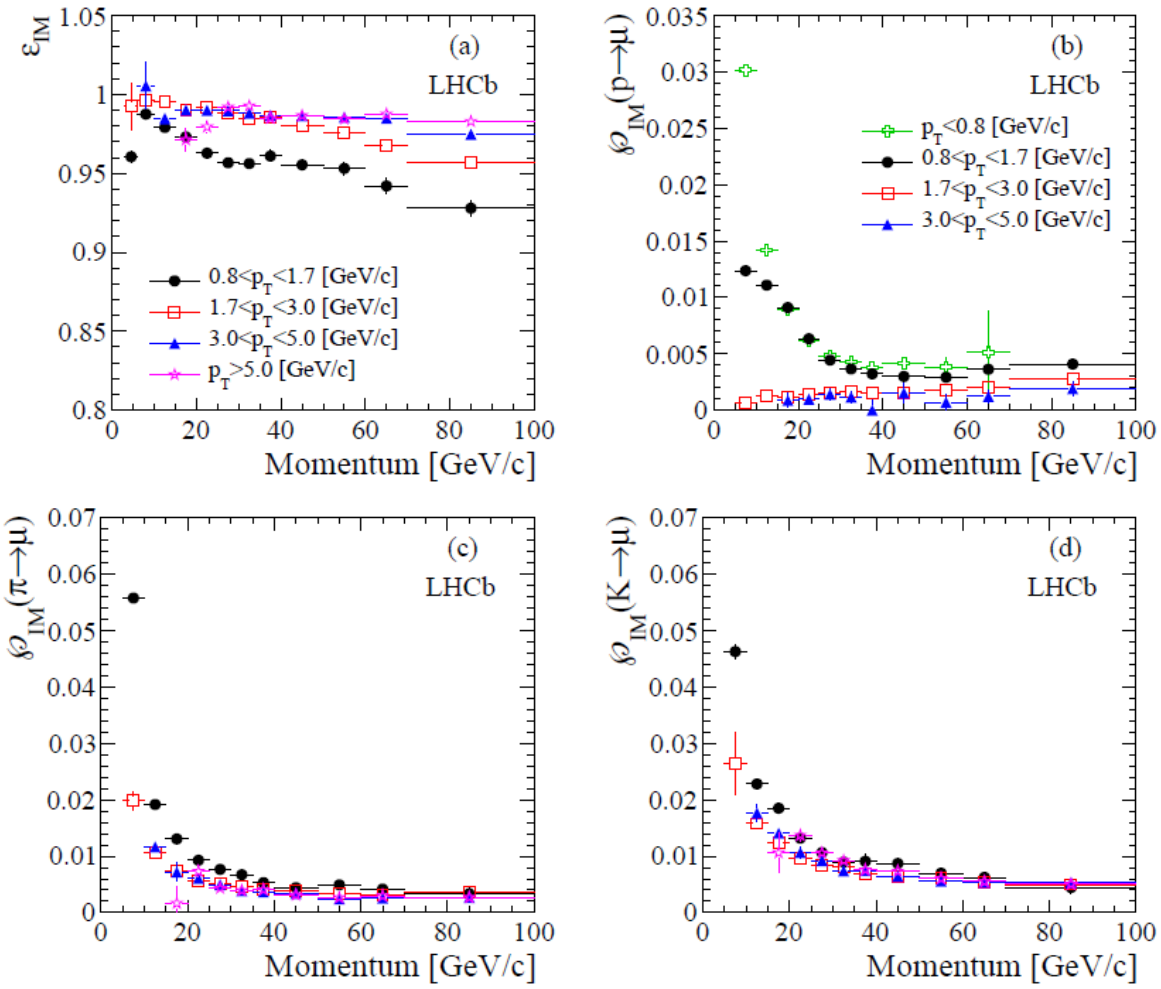


Figure 3.28.: (a) Total efficiency of muon identification, and probability for (b) protons, (c) pions and (d) kaons to be mis-identified as muons, given as functions of particle momentum, for different ranges of transverse momentum [102].

an average rate of 28 MHz, in order to gather data from each interaction.[1] However, detector data can only be stored at a rate of around 5 kHz. A trigger system is used to select only the most promising events for retention.

The LHCb trigger consists of an initial, fast-response, hardware stage, known as Level-0 (L0), which is synchronised with the bunch crossing rate of the collider. This stage uses limited information from the detector subsystems to discriminate between events, and outputs selected events at a rate of approximately 1 MHz. The data for these selected events are then passed into the second trigger stage, the software-based High Level Trigger (HLT). This stage runs on a dedicated on-site computing farm, and uses the full information from all detector subsystems to retain events at the final required rate.

[1] In actuality, for Run 1 the bunch crossing rate was half the design value, but the pile-up rate was approximately doubled, so the hypothetical required read-out rate is the same in each case.

Line	Requirements applied during 2011 (2012)
L0Electron	$n_{\text{SPD}} < 600$ and $E_{\text{T}} > 2.5\,(2.72)\,\text{GeV}$
L0Hadron	$n_{\text{SPD}} < 600$ and $E_{\text{T}} > 3.5\,(3.62)\,\text{GeV}$
L0Photon	$n_{\text{SPD}} < 600$ and $E_{\text{T}} > 2.5\,(2.72)\,\text{GeV}$
L0Muon	$n_{\text{SPD}} < 600$ and $p_{\text{T}} > 1.48\,(1.76)\,\text{GeV}/c$
L0DiMuon	$n_{\text{SPD}} < 900$ and $\sqrt{p_{\text{T}_1} \times p_{\text{T}_2}} > 1.296\,(1.6)\,\text{GeV}/c$

Table 3.1.: Criteria set by the LHCb L0 hardware trigger for data gathered during 2011 (2012) [104]. Several groups of criteria, or 'trigger lines' are defined, with events passing any line fed forward to the HLT software trigger. For the dimuon line, p_{T_1} and p_{T_2} refer to the two highest p_{T} tracks through the muon stations. Tighter requirements are used for 2012 data than for 2011, as the rate of events in the detector was greater due to the increased pp collision energy.

Hardware trigger: The L0 trigger has a latency of 4 µs, and uses a few key criteria to select events. The overall occupancy of the event is approximately quantified using the number of hits detected in the SPD subdetector, n_{SPD}, and events with a large number of hits are rejected. Such events would consume a disproportionate amount of CPU power in the full reconstruction, due to combinatoric effects. For 2011 data, approximately 9% of events containing b-hadrons were rejected by the requirement $n_{\text{SPD}} < 600$ [103]. In addition, the L0 trigger places criteria on the highest E_{T} clusters in the ECAL, the highest E_{T} cluster in the HCAL, and the two highest p_{T} tracks through the muon stations. Separate sets of criteria, or 'trigger lines' are used to select candidates for a range of decay types, as shown in Table 3.1. Events meeting any of these sets of criteria are sent to the HLT.

Software trigger: The HLT software trigger is divided into two stages. The first stage, HLT1, uses limited detector information from the VELO and tracking stations to reconstruct tracks within each event. Like the L0 trigger, a range of trigger lines are defined, each aiming to select some particular physically-interesting subset of events, with all events passing any of the lines sent to the next trigger stage. HLT1 produced an output rate of 40 kHz in 2011 and 80 kHz in 2012.

The HLT1 trigger line 'Hlt1TrackAllL0' is used for the analyses described in this thesis. This requires candidate events to contain at least one track reconstructed with transverse momentum $p_{\text{T}} > 1.6\,\text{GeV}/c$, and with good separation from the reconstructed pp interaction vertex, to indicate that it originated from the secondary decay of a B- or D-meson [105]. During 2012, around 40 such trigger lines were used to select events in HLT1.

The second software trigger stage, HLT2, fully reconstructs events using information read out from all of the detector subsystems. In order to produce an acceptable rate of throughput, only tracks meeting some minimum momentum requirements are fully reconstructed. For 2011 data, tracks were required to satisfy $p_{\text{T}} > 0.5\,\text{GeV}/c$ and $p > 5\,\text{GeV}/c$, giving a final output rate of 3 kHz, which increased to 5 kHz during 2012, with the loosening of the p_{T} requirement to $p_{\text{T}} > 0.3\,\text{GeV}/c$.

HLT2 trigger lines are tuned quite closely to the requirements of specific types of analyses, with over 200 lines presently defined. The analyses described in this thesis make use of a group of 'topological' triggers in the HLT2 stage, 'Hlt2Topo(2,3,4)BodyBDT'. These use good quality tracks detected in the

VELO subdetector to reconstruct candidate secondary vertices. Only tracks with a good separation from the PV are used, with adjacent sets of two, three and four tracks of separation below 0.2 mm combined to make 2-, 3- and 4-body objects - candidate b- or c-hadrons. The masses of these candidates are assigned using a corrected mass variable, to take into account any tracks missing from the reconstruction,

$$m_{\text{corr}} = \sqrt{m^2 + |p_{\text{T}}{}^{\text{miss}}|^2} + p_{\text{T}}{}^{\text{miss}}, \tag{3.7}$$

where m is the total reconstructed mass of the hadron candidate and $p_{\text{T}}{}^{\text{miss}}$ is the total missing transverse momentum in the decay. Finally, a multivariate classifier trained to select only the most physically interesting events, combines the corrected masses calculated for the event along with a range of other kinematic variables, to select a few percent of events for retention [103, 106].

Trigger categories: Once an event has been fully reconstructed, each trigger decision assigned to it may be classified as 'TOS' or 'TIS'. A trigger decision is considered TOS if it is 'triggered on signal', i.e. the trigger was fired by tracks coming from the reconstructed signal candidate. A decision is considered TIS if it is 'triggered independently of signal' - such that the same trigger decision would have been made if all tracks associated with the reconstructed signal candidate were removed from the event. The TIS category can be defined relative to a specific trigger line, or globally - allowing the event, after removal of the signal candidate, to be selected by any defined trigger line. Some events satisfy both TOS and TIS conditions, and thus provide a measure of the efficiency of the relevant trigger line [107].

3.2.8. Software and simulation

Data processing and simulation at LHCb is performed within a dedicated object-oriented framework implemented in C++, known as GAUDI [108]. Different tasks are handled by different software projects, each making use of the common tools and interfaces provided by the GAUDI framework. Since GAUDI sets up an environment within which all other LHCb software can be run, it allows effective use of distributed computing resources, in particular the LHC worldwide computing grid, with analysis code dispatched to run on nodes at data storage sites around the world, where the LHCb experimental dataset is stored. This distributed computing functionality is controlled using LHCb DIRAC [109, 110] software, with job configuration handled by the GANGA [111] project.

Simulated samples are used extensively in LHCb analyses. These events are generated using the GAUSS framework [112, 113], which brings together several different generation tools. First, PYTHIA [114, 115] is used to simulated pp interactions, before final state hadron decays are simulated using EVTGEN [116], with final state radiation added using PHOTOS [117]. Following this generation stage, the interaction of each simulated particle with the LHCb detector is treated using the GEANT4 toolkit [118, 119], with a description of the detector's material composition and geometry, before the digital signals arising from the generated particle trajectories are simulated by the BOOLE [120] package.

The next stage in the workflow proceeds identically for the simulation as it does for data. The BRUNEL [121] project reconstructs tracks in the detector, and gathers particle identification information from the RICH, calorimeter and muon subsystems. Finally, the dedicated physics analysis project, DAVINCI [122], is run to reconstruct and fully identify events. Track information from BRUNEL is used

to locate decay vertices, whilst particle identification information is used to assign mass hypotheses to observed tracks. These operations allow intermediate (i.e. not directly detected) particles to be reconstructed and identified. LHCb analysts typically configure DAVINCI to reconstruct only those events matching a given decay hypothesis, and specify a list of variables of relevance for the later stages of the analysis.

For data samples, selection requirements may be specified within the DAVINCI stage to determine which reconstructed candidate events are stored as output. A set of requirements designed to select the most promising candidates for a given analysis is known as a 'stripping' line. These requirements are not normally applied to simulation, so that the same generated samples can be used for multiple analyses.

The output of the DAVINCI stage is a ROOT [123] ntuple, storing the requested variable values for each selected event. These ntuples are small enough to be stored and manipulated locally by analysts, to perform the final 'offline' stages of the analysis. The analyses discussed in this thesis make extensive use of ROOT, which is the *de facto* standard tool used for statistical data analysis at LHCb. Fitting is performed in ROOFIT, a specialised modelling package within the ROOT framework, and in `cfit` [124], an object-oriented C++ library optimised specifically for fitting complex amplitude models.

Filtered simulation

The LHCb simulation framework may be configured in order to generate samples that are useful to specific analyses. Typically this is done by specifying the decay chain of interest as a 'decay descriptor', which is passed as input to the EVTGEN simulation loop. Only events matching this descriptor are then passed through the full simulation chain and written to disk.

However, in order to use simulated samples within an analysis, e.g. to examine mass PDFs or angular distributions expected to arise from certain decays, it is usually necessary to first apply the same requirements that will be used to select candidate events in data. In analyses for which these requirements must be made very stringent, due to the presence of background components with much higher branching fractions, an extremely large number of events must be simulated and stored in order to retain a statistically significant number of events once the requirements have been applied.

'Filtered simulation' addresses this problem by applying selection requirements that would normally be applied in the 'offline' analysis directly within the DAVINCI step of the simulation, reducing the fraction of simulated events written to disk. The process of writing and locally testing these selection requirements requires more expertise than a normal simulation, and simulations must be actively monitored to ensure that selection requirements are functioning correctly within the production environment. In addition, in order to evaluate the efficiency of each selection requirement used, retention statistics must be generated from the log files written during the DAVINCI step. The Filtered Production Liaison is the person responsible for overseeing these processes for all LHCb filtered productions. The author held this role from June 2013 to January 2016, and oversaw the filtered production of many simulated samples used within this thesis.

3.3. The LHCb Upgrade

The 'high luminosity era' of the LHC is set to begin in the third run period, commencing in May 2021. In this period, it is intended to increase LHCb's operating luminosity by a factor of five, to $\mathcal{L} = 2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. It is also intended to read out the detector at every bunch crossing, allowing more flexible software-based triggering to be performed using the signals recorded.

In order to gather data effectively in these new conditions, various upgraded components are in development for the LHCb detector. One area where significant changes are expected is in the photon detection arrays used within each RICH subdetector, with the existing Hybrid Photon Detectors and onboard readout electronics being replaced by an entirely new opto-electronic chain. The new system will use multi-anode photomultiplier tubes (MaPMTs), capable of recovering baseline voltage within the 25 ns window between bunch crossings, read out via a specially developed radiation-tolerant ASIC. The author took part in beam line testing of the candidate opto-electronic chain during 2014, holding responsibility for the data acquisition (DAQ) system used to configure the readout electronics and record data.

3.3.1. Opto-electronic chain

The opto-electronic chain used during 2014 beam tests is illustrated schematically in Figure 3.29 and consists of:

- Four Hamamatsu R11265 [8] MaPMTs of 8×8 pixels. Each MaPMT measures 26.2×26.2 mm^2, with a pixel size of 2.9×2.9 mm^2, giving an active area of 23×23 mm^2;
- A common baseboard (BB) providing high voltage connections to each MaPMT;
- Four front-end-boards (FEB) connecting to the baseboard, each equipped with eight eight-channel CLARO ASICs [125], used to digitize the MaPMT output;
- A common backboard (BkB), connecting to each front end board;
- Two digital boards (DB) connecting to the backboard, each housing a field-programmable gate array (FPGA) carrying firmware to interface with the DAQ PC via Gbit ethernet connection.

The block consisting of four MaPMTs and four FEBs, connected via a common baseboard and backboard, is known as an elementary cell (EC). This is intended to be the basic building block of the upgraded RICH detector planes, allowing a modular design to be used for mechanical support and cooling structures.

Each pixel of an MaPMT acts as an individual photomultiplier tube, with a high voltage photocathode followed by a series of dynodes. Photons incident on the photocathode stimulate the emission of photoelectrons, which accelerate through the potential difference and create cascades of electrons between the dynodes, generating a detectible current at the output. The CLARO ASIC ('chip') is responsible for converting this analogue output into a digital signal. Each CLARO has eight channels, so eight chips are needed to read out a single MaPMT. Within the CLARO, a separate binary discriminator with

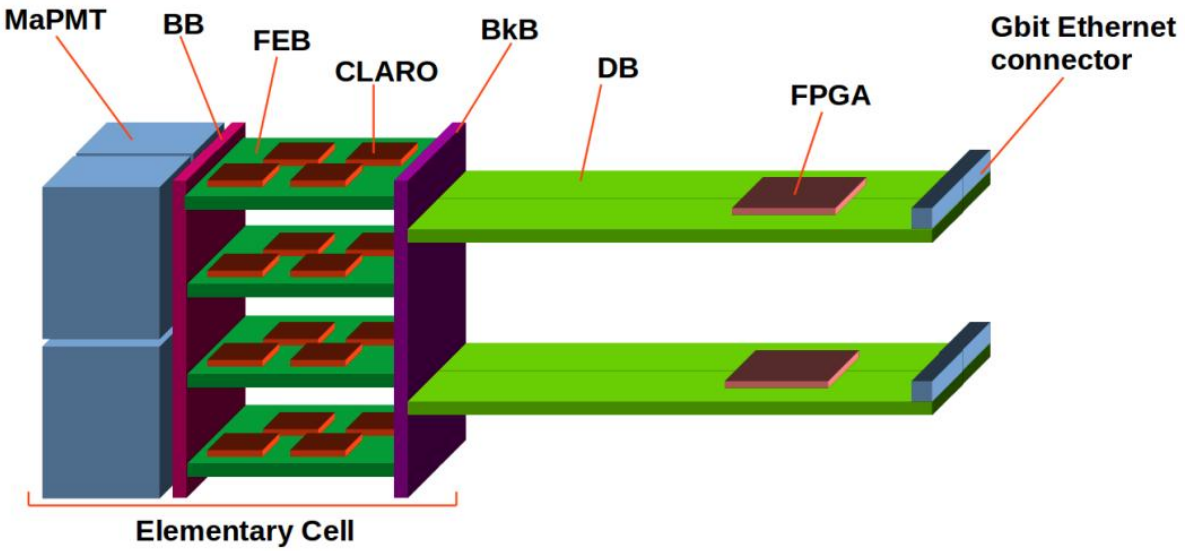


Figure 3.29.: Candidate RICH opto-electronic chain used during 2014 and 2015 beam tests.

variable threshold is used to digitise each channel. A discriminator converts an incoming analogue signal into a binary one by outputting a fixed 'high' voltage when the input signal exceeds the threshold voltage, and a fixed 'low' voltage otherwise. Each threshold is set using a 6-bit register, allowing 64 different possible threshold values. Different thresholds may be used for different channels to counter variations in the gain of each MaPMT pixel (the gain can differ between pixels by a factor of three [8]), or in order to suppress pixels that are particularly suceptible to cross-talk.

The digital signals output by sixteen CLARO chips are combined in each digital board. This uses a 40 MHz clock pulse, combined with trigger input (trigger signals may be generated either externally or using the inbuilt pulser of the digital board), to format the signals as 'events'. Events are then aggregated within a buffer on the digital board, from which they are read out as 'multi-event packets' via the gigabit ethernet link. Communication between the two digital boards and the DAQ system is routed via a network switch.

3.3.2. DAQ system

A JAVA based GUI, implemented using the `Swing` library, is used to configure the parameters of the opto-electronic chain and to send run control signals. Preferences which may be configured include:

- Network address of the DAQ PC and each of the connected digital boards. Using a recognised network protocol allows data packets to be routed efficiently by the network switch, since each packet is broadcast only to its intended recipients;
- The hexadecimal identifier ('DNA') of each digital board. These are used to generate unique identifiers for each connected CLARO chip. The resulting identifiers are interpreted by the firmware of the digital boards, and used to propagate the desired settings to each CLARO;

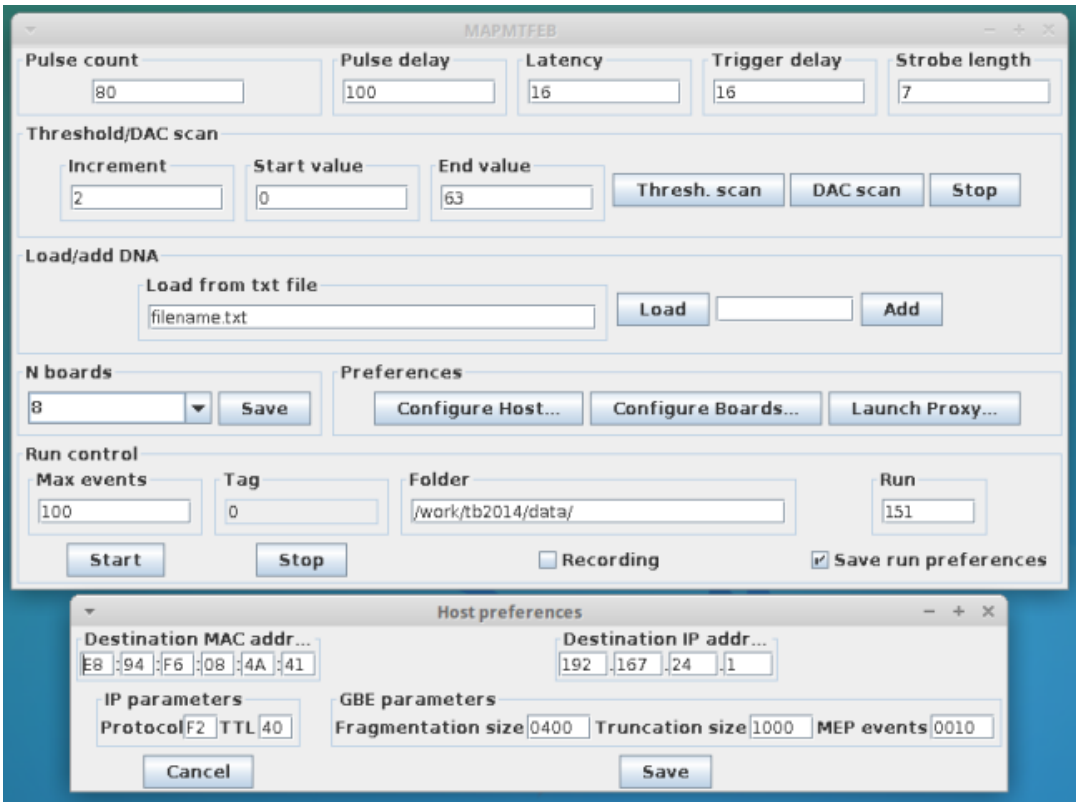


Figure 3.30.: Screen capture taken during operation of the DAQ GUI. The upper frame is the main window displayed each time the GUI is run. Clicking "Configure Host . . . " in this window generates the lower frame, where network addresses for the DAQ PC may be configured.

- The per-channel settings of each CLARO chip, including the threshold of the discriminator and the attenuation factor (used in pulse-shaping the incoming analogue signal).

The different windows of the GUI are shown in Figure 3.30 and Figure 3.31. Individual GUI windows are generated for the configuration of each digital board and each CLARO chip, allowing flexible configuration of any number of elementary cells. Configured preferences are stored in `.xml` format so that they can be remembered between runs.

Run control and data readout require the sending and receiving of data packets between the DAQ PC, where the GUI is running, and the digital boards. A simple application, implemented in the `C` language, is used to handle this interaction. The application is ported into JAVA as a shared library, so that it can be called from within the GUI.

At the start of each run, the stored preferences are gathered and formatted into a `C` struct, encoded in binary. This struct is then sent as a single data packet to the connected digital boards. The firmware implemented in the FPGA of each digital board interprets the encoded settings and uses them to configure each of the connected front-end boards, before commencing data-taking. During a data-taking run, multi-event packets are sent from the digital boards to the DAQ PC, where a waiting socket-reading process appends all incoming packets to a single file. To terminate the run, a final packet is sent from the DAQ PC to the boards.

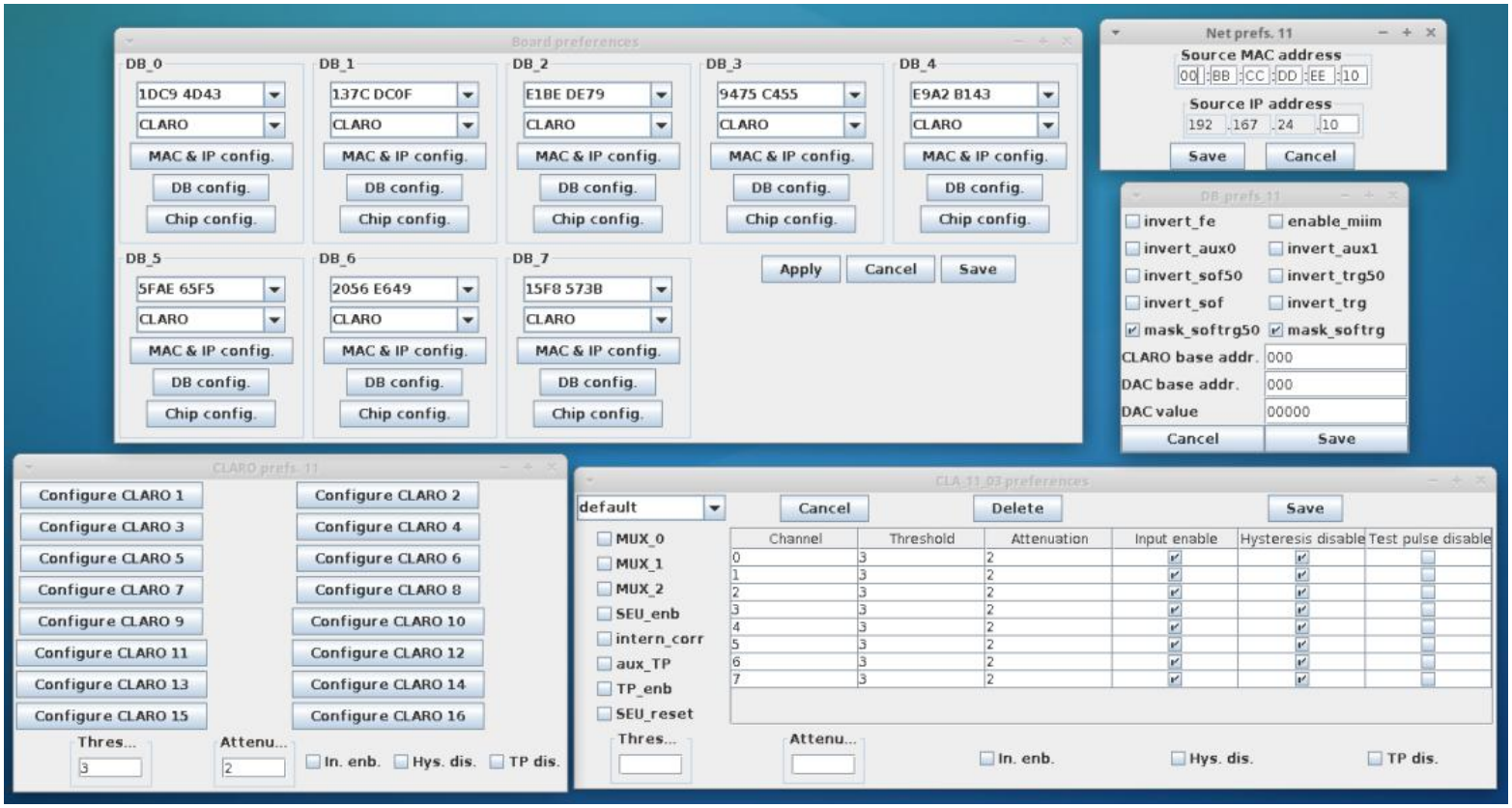


Figure 3.31.: Screen capture taken during operation of the DAQ GUI. The upper left frame is displayed by clicking "Configure Boards" in the main window of the GUI. Independent configuration of each connected digital board is possible. Clicking "Chip config." for a given digital board generates the lower left frame. Clicking on each of the buttons within this frame allows the user to configure the per-channel settings of each CLARO chip, via the lower right frame.

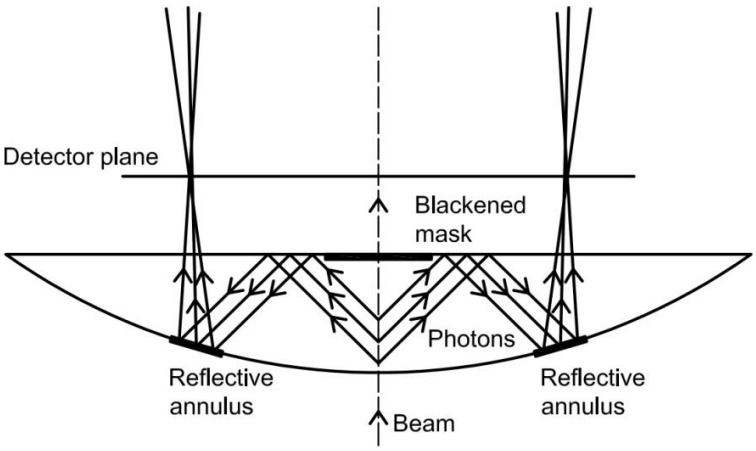


Figure 3.32.: Planoconvex lens used to generate and focus Cherenkov photons during beam tests of the candidate optoelectronic chain [8].

3.3.3. Beam test setup

To assess its suitability for use in the RICH subdetectors, beam tests focused on the candidate optoelectronic chain's ability to precisely detect individual Cherenkov photons. The tests were performed using a beam of momentum 180 GeV/c, consisting mainly of protons and pions, at the SPS facility at CERN.

A planoconvex borosilicate glass lens is used, both as a radiator medium and to focus the resulting Cherenkov photons. As shown in Figure 3.32, Cherenkov photons emitted within the lens are totally internally reflected at the flat surface, before being reflected once more from an aluminium annulus deposited on the convex surface. In order to reduce emission-point uncertainty, a circular portion at the centre of the flat surface is masked.

For these beam tests, sufficient components were available to fully instrument a single elementary cell only. In order to allow testing of the Cherenkov ring reconstruction, it was decided to instead half-instrument two elementary cells, placed at diamaterically opposing points of the expected ring position in the focal plane of the lens. This arrangement is shown in Figure 3.33.

The lens is housed within a light-tight box, with windows to accommodate the plastic housing of each elementary cell, as shown in Figure 3.34. The installation of the box within the SPS beam area is shown in Figure 3.35.

Readout is triggered by a dedicated circuit board, sending synchronous signals to each of the digital boards. Triggers are generated either in response to incident beam particles, or using the trigger board's internal pulser. Triggering on beam particles makes use of two beam scintillators, placed upstream and downstream of the box. The trigger signal toggles on the coincidence of the two scintillators. Adjustable dead time is used in each digital board so that signals originating from the same event reach the output of each board at the same time, along with the corresponding trigger pulse. The logic used to achieve this is shown in Figure 3.36, and the timing of each pulse is shown in Figure 3.37.

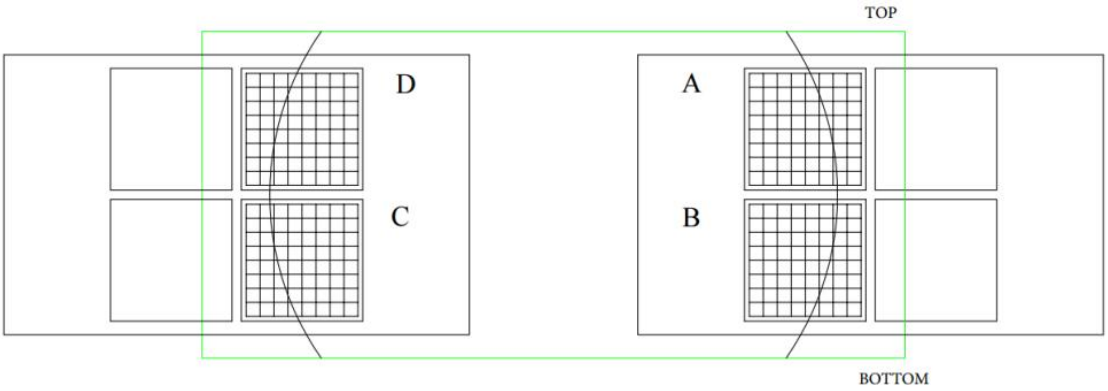


Figure 3.33.: Positioning of half-instrumented elementary cells relative to the expected Cherenkov ring during beam tests of the candidate optoelectronic chain [8].

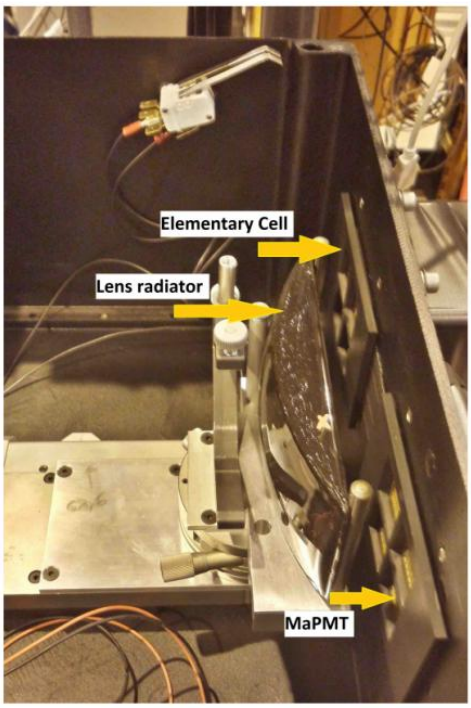


Figure 3.34.: Photograph of the lens used during beam tests of the candidate optoelectronic chain, *in situ* within the light-tight box. The entry windows of the elementary cells are visible [8].

To assist with the recontruction of particle trajectories, a dedicated 'tracking telescope' is used to track beam particles upstream of the light-tight box. This tracking information also allows events containing multiple beam particles to be identified. The telescope uses eight planes of silicon pixels detectors, of pixel size 55 μm $\times$ 55 μm, giving a position resolution of $\mathcal{O}(1)$ μm in the detector plane [8]. Telescope readout is triggerless, with continuous hit recording during each beam run. Each particle hit detected is associated with a timestamp of precision 1 ns. The trigger board is connected to the

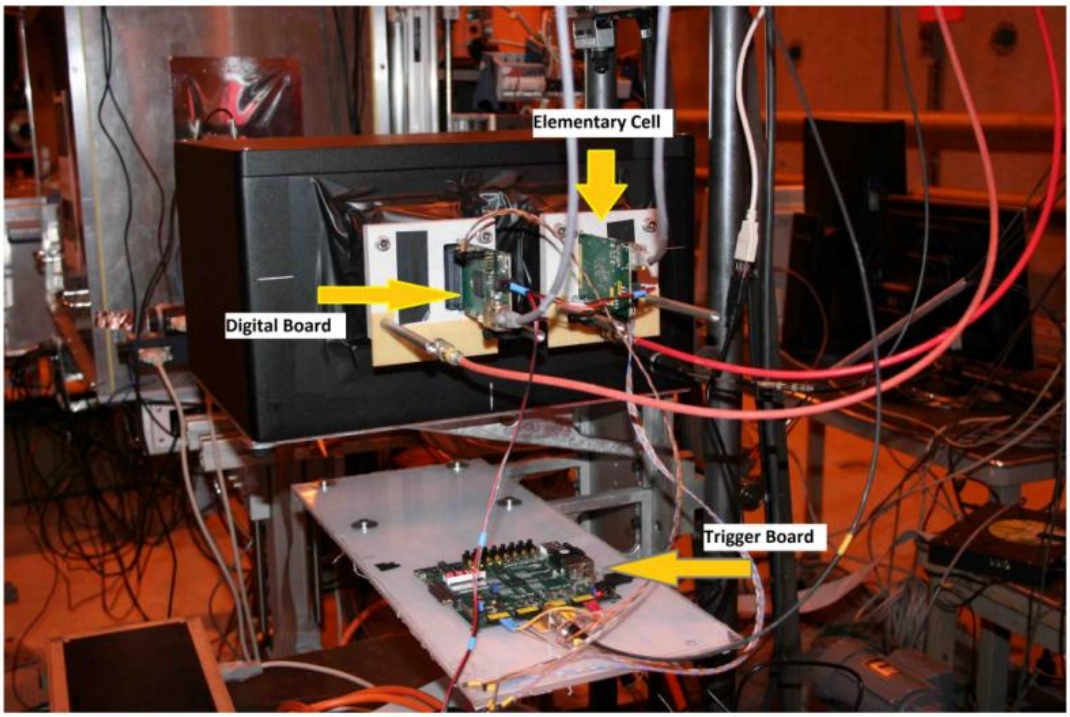


Figure 3.35.: Photograph of the experimental setup used for beam tests of the candidate optoelectronic chain. The light-tight box housing the lens is visible, with one digital board projecting from each connected elementary cell [8].

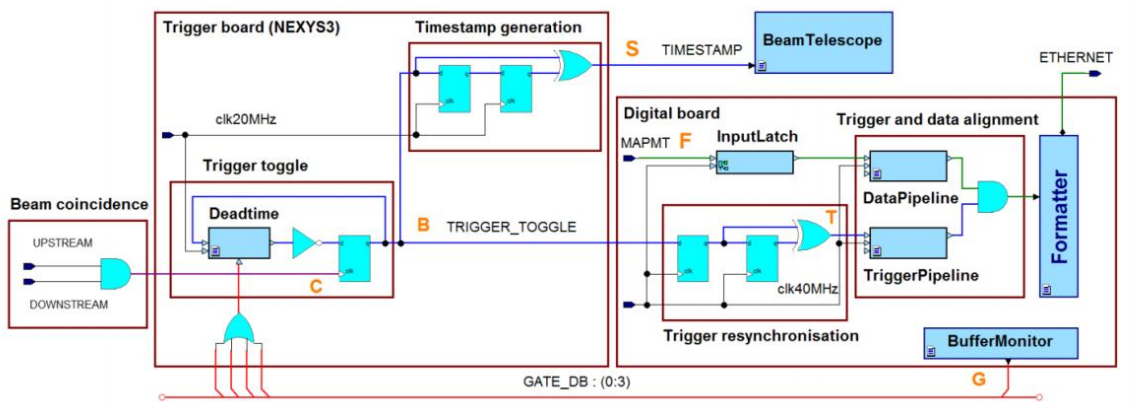


Figure 3.36.: Trigger logic used during beam tests of the candidate optoelectronic chain [8].

telescope, with a telescope timestamp recorded each time a trigger signal is sent. This allows matching of events between the telescope and digital boards.

3.3.4. Results

A hitmap obtained during a typical beam run is shown in Figure 3.38. The resolution of the Cherenkov ring is seen to be of the order of the pixel width, since few hits are observed for off-ring pixels.

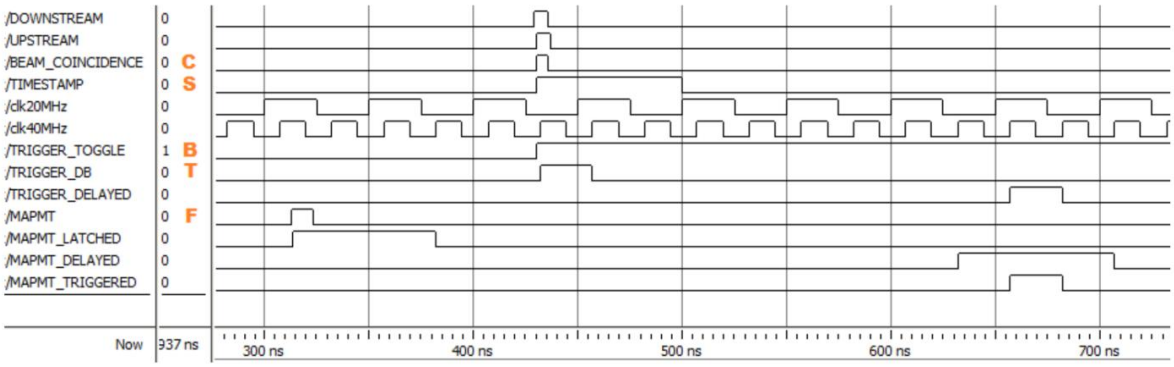


Figure 3.37.: Signal timing used during beam tests of the candidate optoelectronic chain [8].

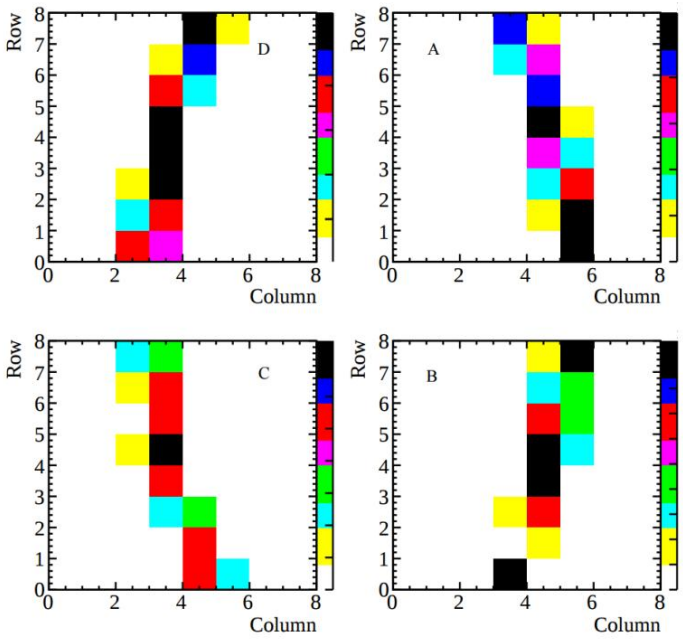


Figure 3.38.: Integrated hit map recorded during a typical beam test data-taking run [8].

Threshold scans are performed to characterise each detector pixel. Thresholds are varied between one and 63, in steps of two units, where for each CLARO channel one unit corresponds to a difference of approximately $35 \times 10^3\ e^-$ in the incoming pulse [8]. Integrated hits are plotted against threshold level, as shown in Figure 3.39, to give an 'S-curve' for each channel.

Numerical differentiation of this histogram gives the 'single photon spectrum' for each channel, as shown in Figure 3.40. This spectrum indicates what threshold value is most suitable for a given channel - somewhere in between the noise which dominates at low thresholds and the approximately Gaussian spectrum arising from a single incident Cherenkov photon.

Finally, the photon hits detected in each event may be used to reconstruct the full Cherenkov ring. A simple approach assumes the particle trajectory is perpendicular to the flat surface of the lens, giving

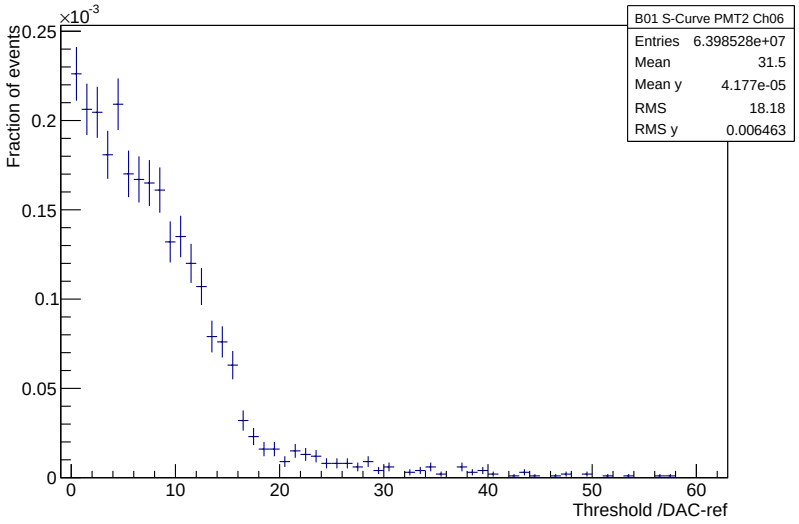


Figure 3.39.: Integrated hit spectrum ('S-curve') obtained during a threshold scan for a single CLARO channel during beam tests of the candidate optoelectronic chain.

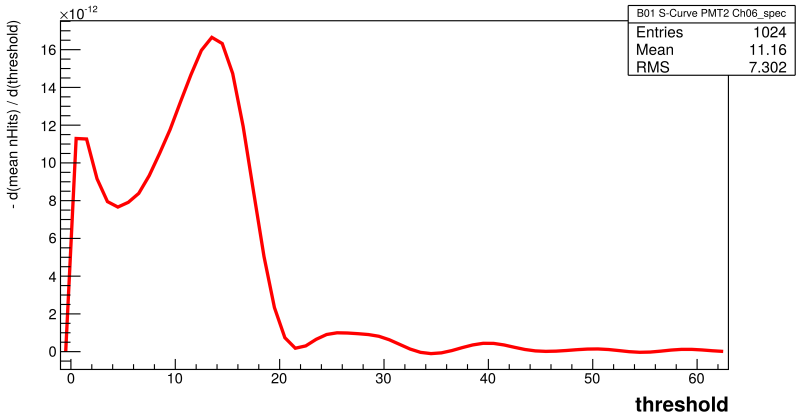


Figure 3.40.: Single photon spectrum obtained for a single CLARO channel during beam tests of the candidate optoelectronic chain.

a precisely circular ring in the detector plane. A least-squares fit may be then performed to find the ring centre, (x_0, y_0), and radius, r, which best describe the observed hits. The result of such a fit for a single typical event is shown in Figure 3.41.

3.4. Summary

The LHCb detector has been purpose built for the study of hadronic decays at the LHC. It is optimised for the measurement of interactions involving b- and c-quarks, thanks particularly to the fact it is instrumented very close to the pp interaction region; allowing secondary decay vertices to be located; and to the dedicated subsystems it uses to distinguish the identities of hadrons and muons.

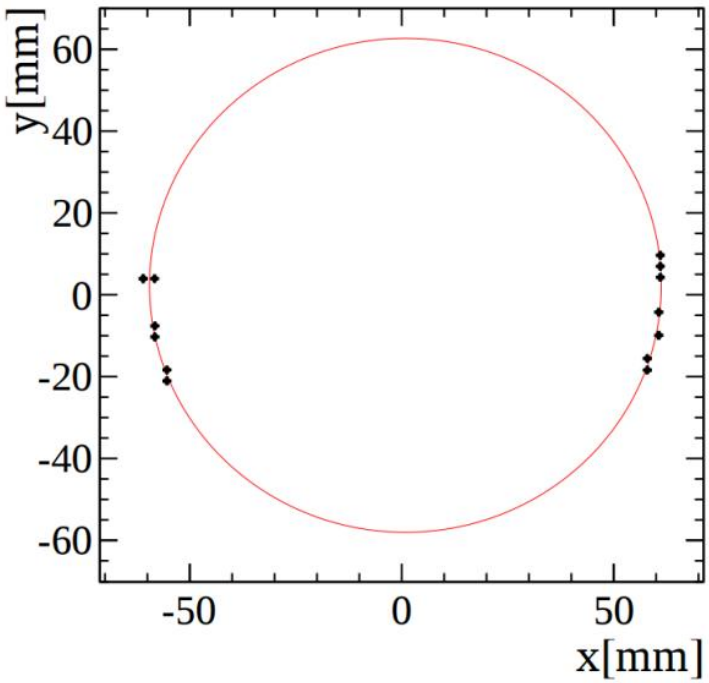


Figure 3.41.: Pixel hits and reconstructed Cherenkov ring obtained for a single beam test event [8].

All measurements presented in this thesis make use of the full 2011 and 2012 datasets gathered at the LHCb detector. A total integrated luminosity of $1.0\,\text{fb}^{-1}$ was collected during 2011 at centre-of-mass energy $\sqrt{s} = 7\,\text{TeV}$, of which 56% was obtained with the LHCb dipole magnet in the 'down' polarity. During 2012, a total integrated luminosity of $2.0\,\text{fb}^{-1}$ was collected at $\sqrt{s} = 8\,\text{TeV}$, of which 52% was obtained with the magnet in the 'down' polarity.

4. Selection of candidate events

Section 1.5 and Chapter 2 describe techniques that can be used to extract γ from decays of the form $B \to D^{(*)}K^{(*)}$. Before these techniques can be applied in practice, samples of candidate events must be isolated for each of the decay modes under study:

- $B^+ \to DK^+$, with $D \to K^0_S\pi^+\pi^-$;
- $B^0 \to DK^{*0}$, with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$;
- $B^+ \to D^*K^+$, with $D^* \to D(\gamma/\pi^0)$ and $D \to K^0_S\pi^+\pi^-$; and
- $B^+ \to DK^{*+}$, with $K^{*\pm} \to K^0_S\pi^\pm$ and $D \to K^0_S\pi^+\pi^-$.

The following control modes are also used:

- $B^+ \to D\pi^+$, with $D \to K^0_S\pi^+\pi^-$;
- $B^0_s \to DK^{*0}$, with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$; and
- $B^+ \to D^*\pi^+$, with $D^* \to D(\gamma/\pi^0)$ and $D \to K^0_S\pi^+\pi^-$.

Since it is often useful to describe a given signal mode and its corresponding control mode together, the label h is used to denote both π and K, so that e.g. $B^+ \to Dh^+$ denotes both $B^+ \to DK^+$ and $B^+ \to D\pi^+$.

Starting from the full LHCb Run 1 dataset, successive requirements ('cuts') are applied to select candidate events. In the language of statistical learning[1], each cut represents a binary classifier separating the dataset into two categories - 'signal' and 'background'.

Throughout this thesis the 2011 and 2012 datasets are merged and analysed together. As described in Chapter 3, changes were made to running conditions and detector configuration between the two years. However, cross-checks performed during the standalone $B^0 \to DK^{*0}$ analysis [128] did not find significant differences in the behaviour of the variables used in the analysis. Likewise, MC simulation generated using both 2011 and 2012 conditions behaves similarly, and is combined throughout the analysis. Throughout this thesis all samples are generated using 2012 conditions unless otherwise specified.

Selection stages

The first set of cuts are applied within the DaVinci stage of the event processing as stripping lines, as described in Section 3.2.8. The stripping lines used to select candidates in each decay mode are

[1] The theory of statistical learning and statistical inference has been comprehensively treated by Hastie *et al.* [126]. Practical details on implementing statistical learning are given by Géron [127]

discussed in Section 4.1. These requirements determine the events that are written out to ROOT ntuples for local ('offline') analysis.

The offline selection proceeds in two stages. First, a handful of physically-motivated simple cuts are used to reject obvious background from each sample. These 'preselection' cuts are described in Section 4.2 and Section 4.3. Section 4.2 describes a set of cuts common to all decay modes included, whilst Section 4.3 describes cuts specific to each mode. Finally, a multivariate binary classifier (or 'multivariate analysis', MVA) is applied for each decay mode. Various types of multivariate classifier are implemented, as discussed in Section 4.4.

Variables on which selection requirements are placed are introduced throughout this chapter. The first time each variable is mentioned, its name is written in *italics*, and a brief description is given.

Classification of K^0_S candidates

Each decay mode studied in this thesis includes a neutral D-meson decaying to the state $(K^0_S\pi^+\pi^-)$, with the K^0_S-meson reconstructed in the $(\pi^+\pi^-)$ final state. Throughout this work, a distinction will be made between two classes of K^0_S candidates. For *long-long* (LL) K^0_S candidates, the two pions leave *long* tracks in the LHCb detector, giving rise to hits both in the VELO and in the downstream tracking stations, and thus the decay vertex of the reconstructed K^0_S must lie within the VELO subdetector. For *down-down* (DD) K^0_S candidates, the two pions leave *downstream* tracks only, with no associated VELO hits. As such, the reconstructed K^0_S decay vertex lies downstream of the VELO region. It is useful to separate the two categories in the selection, since in each case the resolution and correlations of numerous variables relating to the K^0_S candidates differ.

For the decay mode $B^+ \to DK^{*+}$, with $D \to K^0_S\pi^+\pi^-$ and $K^{*+} \to K^0_S\pi^+$, an additional K^0_S candidate is reconstructed in the decay of the K^{*+}-meson. In principle, this could motivate the use of four separate categories in the selection for this mode, to treat each possible combination of K^0_S types. However, to keep the selection as similar as possible to the other modes, no criteria relating to the K^0_S originating from the $K^{*\pm}$ are used in the selection, except for a simple mass window. As such, no significant difference is expected whether this candidate is reconstructed from long or downstream pion tracks, and so only two categories are defined for this mode; separating events containing LL and DD K^0_S candidates originating from the D-meson decay.

Refitted intermediate mass variables

'Mass windows' are used at several points within the selection. These require that the reconstructed mass of a given particle lie within some range of its known mass, with the allowed range reflecting both the finite width, Γ, of the relativistic lineshape and the finite resolution of the LHCb detector and reconstruction algorithms.

Combinatoric candidates, reconstructed from tracks which do not actually originate from the same particle, can fall within these mass windows by chance. In order to reduce contamination from this source of background, the `DecayTreeFitter` package [129] is used to perform alternative

reconstructions of the full decay chain, to obtain 'refitted' mass variables for the intermediate particles in each case:

- in the 'KsFit' reconstruction, once a K^0_S-meson candidate has been identified and reconstructed from its daughter tracks, its mass is fixed to the nominal value, before it is used to reconstruct the other candidates above it in the decay chain;
- in the 'DFit' reconstruction, once a D-meson has been identified and reconstructed, its mass is fixed to the nominal value, before it is used to reconstruct the other candidates above it in the decay chain;
- the 'PVFit' reconstruction fixes the masses of both K^0_S- and D-meson candidates to their nominal values, before reconstructing the other candidates above them in the decay chain.

This combination of reconstructions allows mass windows to be placed on each intermediate hadron whilst constraining those further down the decay chain.

Events containing multiple candidates

In any case where multiple candidate B-meson decays are selected from a single event, the candidate returning the lowest `DecayTreeFitter` 'PVFit' χ^2 per degree of freedom (*i.e.* the highest quality of fit) is kept and the others are discarded.

Truth-matching Monte Carlo simulation

All MC samples used within this thesis are 'truth-matched'. This means that truth-level information produced in the simulation stage is used to match mother and daughter particles in the decay chain, and only events matching the desired decay hypothesis are retained.

4.1. Stripping requirements

Stripping requirements to select decay candidates of the form $B \to DX$ have been developed by the LHCb 'B decays to open charm' Working Group [130]. B represents any charged or neutral B-meson, D represents any charged or neutral D-meson and X represents a range of possible charmless states, including those obtained in the signal and control modes of the present analysis ($K^\pm$, $\pi^\pm$, K^{*0}, and $K^{*\pm}$). The requirements chosen are as follows:

- the number of long tracks present in each event is required to be less than 100. This removes events whose reconstruction requires disproportionate CPU usage;
- all charged particle tracks used in the reconstruction of the event are required to have a momentum greater than 2000 MeV, and transverse momentum greater than 250 MeV. For each track, the χ^2 returned by the track reconstruction algorithm, $(\chi^2)^{\text{track}}$, is required to be less than 5. This variable measures the goodness-of-fit achieved, with smaller values indicating better fit quality. In addition, the track's *impact parameter* χ^2 with respect to the primary vertex (PV), $(\chi^2)^{\text{IP,PV}}$, is required to

be greater than 4. This variable gives the change in χ^2 of the reconstruction, when the relevant track is associated directly with the PV, instead of its reconstructed creation vertex. As such, this variable is sensitive to the impact parameter (perpendicular separation) of the track relative to the PV. Tracks that really do originate at the PV, e.g. true B-meson candidates, give low values, whilst tracks that originate from displaced secondary vertices give higher values;

- K^*-meson candidates decaying to $K\pi$ are required to have a momentum greater than 1000 MeV, whilst the daughter kaon and pion used to reconstruct them are required to have transverse momenta greater than 300 MeV. Each K^*-meson candidate is also required to have a $(\chi^2)^{\text{IP,PV}}$ greater than 4. In addition, the candidate's *decay vertex* χ^2, $(\chi^2)^{\text{endvtx}}$, is required to be less than 12. This variable gives the χ^2 returned by the vertex reconstruction algorithm for the decay vertex of the candidate meson, and measures the goodness-of-fit achieved, with smaller values indicating better fit quality;

- D-meson candidates decaying to $K^0_S\pi^+\pi^-$ are required to have a mass, M_D, between 1764.8 and 1964.8 MeV, a transverse momentum greater than 1500 MeV, a $(\chi^2)^{\text{endvtx}}$ less than 12 and a $(\chi^2)^{\text{IP,PV}}$ greater than 100. The *distance of closest approach* between the two daughter pion tracks used to reconstruct the D-meson candidate, $l^{\text{DOCA}}_{\pi\pi}$, is required to be less than 1.5 mm. In addition, the *cosine of the direction angle* of the candidate, $\cos(\theta^{\text{DIRA}})$, is required to be greater than 0.9. This variable gives the cosine of the angle between a candidate's momentum vector, and the direction vector between its production and decay vertices. For a well-reconstructed particle this angle is small, giving cosine values which peak at 1;

- D-meson daughter K^0_S candidates are required to have a mass, $M_{K^0_S}$ between 467.6 and 527.6 MeV, and a $(\chi^2)^{\text{endvtx}}$ less than 16. In addition, the *flight distance* χ^2 of each K^0_S candidate relative to its own creation (or 'origin') vertex, $(\chi^2)^{\text{FD,orivtx}}$, is required to be greater than 4. For a given meson candidate, this variable gives the difference in χ^2 for the reconstruction when all of the meson's daughter candidates are associated directly with its creation vertex, *i.e.* as if the intermediate decayed instantly. This difference is larger for mesons that travel further between production and decay, and so longer-lived particles give higher values for this variable;

- B-meson candidates are required to have a mass, M_B, between 4750 and 7000 MeV, a transverse momentum greater than 5000 MeV, $(\chi^2)^{\text{IP,PV}}$ less than 25, $\cos(\theta^{\text{DIRA}})$ greater than 0.999, and a lifetime, τ, greater than 0.2 ps. In addition, the (χ^2) per degree of freedom in the vertex fit at the B-meson decay vertex, $(\chi^2/\text{ndof})^{\text{endvtx}}$ is required to be less than 10; and

- the output of a multivariate classifier [130], B2CMVA, is required to be greater than 0.05. The classifier is trained on three input variables:

  - the transverse momentum of the reconstructed B-meson candidate;
  - the B-meson $(\chi^2)^{\text{FD,orivtx}}_B$; and
  - the sum of the χ^2 per degrees of freedom over all reconstructed B- and D-meson decay vertices within the event.

These requirements are summarised in Table 4.1.

Candidate	Requirement
Each event	$N_{\text{long tracks}} < 100$
All charged tracks used in reconstruction	$(\chi^2)^{\text{track}} < 5$
	$p^T > 250$ MeV
	$p > 2000$ MeV
	$(\chi^2)^{\text{IP,PV}} > 4$
K^*-meson decaying to $K\pi$	$p^T > 1000$ MeV
	$(\chi^2)^{\text{endvtx}} < 12$
	$(\chi^2)^{\text{IP,PV}} > 4$
	$p^T_{\text{daughter}} > 300$ MeV
D-meson decaying to $K^0_S\pi^+\pi^-$	$1764.8 < M_D < 1964.8$ MeV
	$p^T > 1500$ MeV
	$(\chi^2)^{\text{endvtx}} < 12$
	$(\chi^2)^{\text{IP,PV}} > 100.0$
	$\cos(\theta^{\text{DIRA}}) > 0.9$
	$l^{\text{DOCA}}_{\pi\pi} < 1.5$ mm
D-meson daughter K^0_S	$467.6 < M_{K^0_S} < 527.6$ MeV
	$(\chi^2)^{\text{endvtx}} < 16$
	$(\chi^2)^{\text{FD,orivtx}} > 4$
B-meson	$4750 < M_B < 7000$ MeV
	$p^T > 5000$ MeV
	$(\chi^2/\text{ndof})^{\text{endvtx}} < 10$
	$(\chi^2)^{\text{IP,PV}} < 25$
	$\cos(\theta^{\text{DIRA}}) > 0.999$
	$\tau > 0.2$ ps
	B2CMVA > 0.05

Table 4.1.: Stripping selection requirements developed by the LHCb 'B decays to open charm' Working Group [130] for decays of the form $B \to DX$.

4.2. Preselection requirements common to all decay modes

Loose requirements are placed by eye to remove obvious background from samples before they proceed to the MVA stage of the selection. Requirements are placed on:

- all three stages of the LHCb trigger (as described in Section 3.2.7);
- the output of the `DecayTreeFitter` algorithm used to refit the reconstructed decay chain. To ensure that the refitted mass variables described in the previous section may safely be used, the `DecayTreeFitter` algorithm is required to have run successfully, with return code zero. In addition, well-reconstructed events are selected by requiring that the χ^2 per degree of freedom achieved by the refit, $(\chi^2/\text{ndof})_{\texttt{PVrefit}}$, is less than 90;

- the momenta, $p_{\pi_D^+}$ and $p_{\pi_D^-}$, of the pion daughters of the D-meson candidate, and the momentum of the 'bachelor' candidate, p_{bach}, which are all constrained to lie within the range $[2;100]$ GeV, for which the RICH subdetectors give reliable pion and kaon identification. Throughout this thesis, *bachelor* denotes the final state particle produced opposite the D-meson. For the $B^\pm \to DK^{*\pm}$ and $B^0 \to DK^{*0}$ decay modes, the bachelor requirement is placed on the charged daughter meson/s of the K^* candidate;

- the masses, $M_{K_S^0}$ and M_D, of the intermediate particles in the decay, which are each required to lie within a window surrounding their expected value, taken from the global averages of reference [3][2]. A 30 MeV window is used for $M_{K_S^0}$, and a 50 MeV window is used for M_D. These windows are shown in Figure 4.1 for a sample of $B^\pm \to D^0\pi^\pm$ MC, after the application of all other preselection requirements;

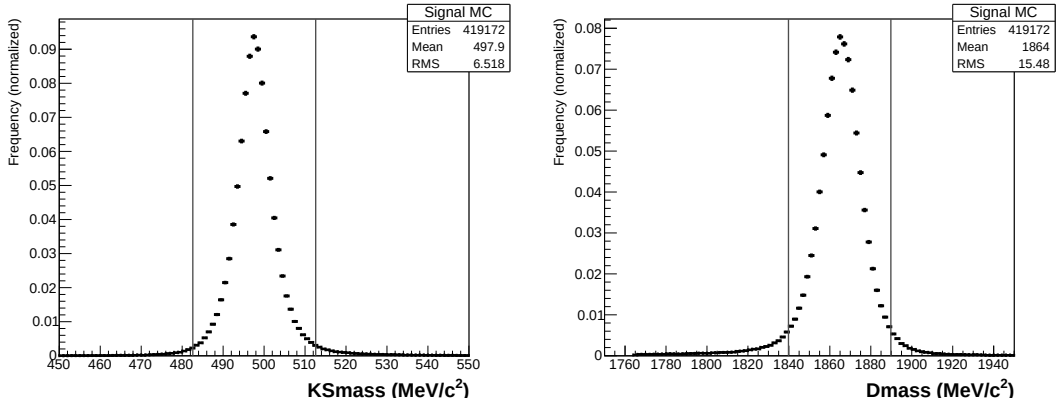


Figure 4.1.: An illustration of the mass window requirements placed on reconstructed K_S^0-meson candidates (left) and D-meson candidates (right): $482.6 < M_{K_S^0} < 512.6$ MeV and $1839.8 < M_D < 1889.8$ MeV. The $M_{K_S^0}$ and M_D distributions obtained from a sample of $B^\pm \to D^0\pi^\pm$ MC events are plotted, with vertical grey lines marking the cut positions chosen.

- the impact parameter χ^2 of the reconstructed B-meson candidate with respect to the primary vertex (PV), $(\chi^2)_B^{\text{IP,PV}}$, which is required to be less than 25;

- the difference in the *decay vertex z-coordinate* between the D- and B-meson candidates, $z_D^{\text{endvtx}} - z_B^{\text{endvtx}}$, which is required to be greater than zero. This requirement reduces contamination from events reconstructed with fake D-meson candidates, and is discussed in the following subsection; and

- the flight distance χ^2 of the reconstructed K_S^0-meson candidate relative to its origin vertex, $(\chi^2)_{K_S^0}^{\text{FD,orivtx}}$, which is required to be greater than 100. This requirement reduces contamination from events reconstructed with fake K_S^0-meson candidates, and is discussed in the following subsection.

These preselection requirements are summarized in Table 4.2.

[2]These windows are centred on different values to the corresponding windows defined in the stripping selection (Table 4.1), but given the looseness of both sets of windows, the difference is insignificant.

Preselection requirement	Purpose
`B_L0HadronDecision_TOS` = 1 `||` `B_L0Global_TIS` = 1 `B_Hlt1TrackAllL0Decision_TOS` = 1 `B_Hlt2Topo{2,3,4}BodyBBDTDecision_TOS` = 1	trigger selection - see Section 3.2.7
`DTF_PVrefit_status` = 0 $(\chi^2/\text{ndof})_{\texttt{PVrefit}} < 90$ $(\chi^2)^{\text{IP,PV}}_B < 25$	B-meson is well-reconstructed
$2 < p_{\text{bach}} < 100$ GeV $2 < p_{\pi^-_D} < 100$ GeV $2 < p_{\pi^+_D} < 100$ GeV	decay products in RICH momentum acceptance
$482.6 < M_{K^0_S} < 512.6$ MeV $1839.8 < M_D < 1889.8$ MeV	mass windows on intermediates
$z^{\text{endvtx}}_D - z^{\text{endvtx}}_B > 0$ $(\chi^2)^{\text{FD,orivtx}}_{K^0_S} > 100$	reduce combinatoric D candidates reduce combinatoric K^0_S candidates (LL K^0_S candidates only)

Table 4.2.: Preselection requirements common to all decay modes studied.

4.2.1. Fake D/K^0_S candidate backgrounds

The final two rows of Table 4.2 describe requirements that are used to discriminate against events reconstructed from 'fake' D- or K^0_S-meson candidates. These arise when the candidate event is a real B-meson decay, and all tracks are correctly associated (*i.e.* the candidate is not combinatoric background), but the true decay does not proceed via the hypothesised intermediate. For example, in the mode $B^\pm \to (K_S\pi\pi)_D K^\pm$:

1. fake D candidates may arise from $B^\pm \to K_S\pi\pi K^\pm$; and
2. fake K^0_S candidates may arise from $B^\pm \to (\pi\pi\pi\pi)_D K^\pm$. However, this is only possible for K^0_S candidates of track type LL, since for DD candidates the K^0_S candidate flies beyond the end of the VELO subdetector before decaying, so cannot be faked by two pion tracks originating within the VELO.

Fake D-meson candidates

For fake D-meson candidates, the $(K^0_S\pi^+\pi^-)$ state does not originate from a real D-meson, so these events cannot be used in the $\mathcal{CP}$ interference analysis. However, they do originate from the decay of a real B-meson, and therefore give a peak in the B-mass spectrum which coincides precisely with the real signal peak.

Figure 4.2 illustrates the effect of the requirement that $z^{\text{endvtx}}_D - z^{\text{endvtx}}_B > 0$, for a sample of fake D-meson candidates. First, in subfigure (a), events are selected from the upper sideband of the reconstructed D-mass spectrum for $B^\pm \to DK^\pm$ data candidates, after the application of all other

selection criteria except the D- and K^0_S-mass requirements. These criteria include a requirement on an MVA classifier trained to select this decay mode, as described in Section 4.4. However, in order to ensure that the MVA classifier used in this study does not bias the sample obtained, it is trained without any `DecayTreeFitter` mass variables used as inputs, since these variables apply constraints on the masses of the intermediates, and could indirectly teach the classifier to favour the constrained values.

To check that the resulting MVA is non-biasing, the distribution of the reconstructed D-meson mass is studied after the application of successively tighter MVA requirements, for a sample of combinatoric background. This sample is obtained from 'wrong-sign' $B^\pm \to DK^\pm$ data, for which the two charged pion daughters of the reconstructed D-meson candidate have the same sign charge, and thus cannot originate from a correctly-reconstructed decay. The resulting distributions are plotted in Figure 4.3, and show no sign of bias, since the distribution remains flat as the MVA classifier requirement is tightened. This indicates that the sample obtained from the sidebands of the reconstructed D-meson mass spectrum in Figure 4.2 (a) is indeed unbiased.

Next, the B-mass spectrum of the selected events is plotted, and fitted using an exponential shape and a Gaussian. The Gaussian peak describes the component originating from fake D-meson candidates, whilst the exponential shape describes the combinatoric component, containing D-meson candidates reconstructed from tracks which do not originate from the same B-decay.

Finally, the fit is repeated after the application of the chosen cut. The population within the Gaussian peak, and thus the yield of fake D-meson candidates, is seen to reduce. The amount of pollution from fake D-meson candidates remaining after this cut for each mode is studied in detail in Section 5.5.

The requirement $z_D^{\text{endvtx}} - z_B^{\text{endvtx}} > 0$ would be expected to retain half of the non-real D-meson candidates, since the tracks used to reconstruct the D-meson in this case actually do originate from the PV, and after smearing by resolution effects ought to lie upstream or downstream of the PV with equal probability. As such, background due to fake D-meson candidates is likely to be non-negligible, and should be considered as a source of systematic uncertainty.

Fake K^0_S-meson candidates

For fake K^0_S-meson candidates, the D meson does not decay to the desired $(K^0_S\pi^+\pi^-)$ state, and the weak phase involved gives no information on γ, so these events cannot be used in the $\mathcal{CP}$ interference analysis. However, they do originate from the decay of a real B-meson, thus giving a peak in the B-mass spectrum which coincides precisely with the real signal peak. Figure 4.4 illustrates the effect of the requirement that $(\chi^2)^{\text{FD,orivtx}}_{K^0_S} > 100$, for a sample of fake K^0_S-meson candidates. The study proceeds analogously to that shown in Figure 4.2 for fake D-meson candidates. The requirement used to reduce non-real K^0_S background is highly effective, retaining a negligible fraction of the background, and so the bias check of Figure 4.3 is not repeated.

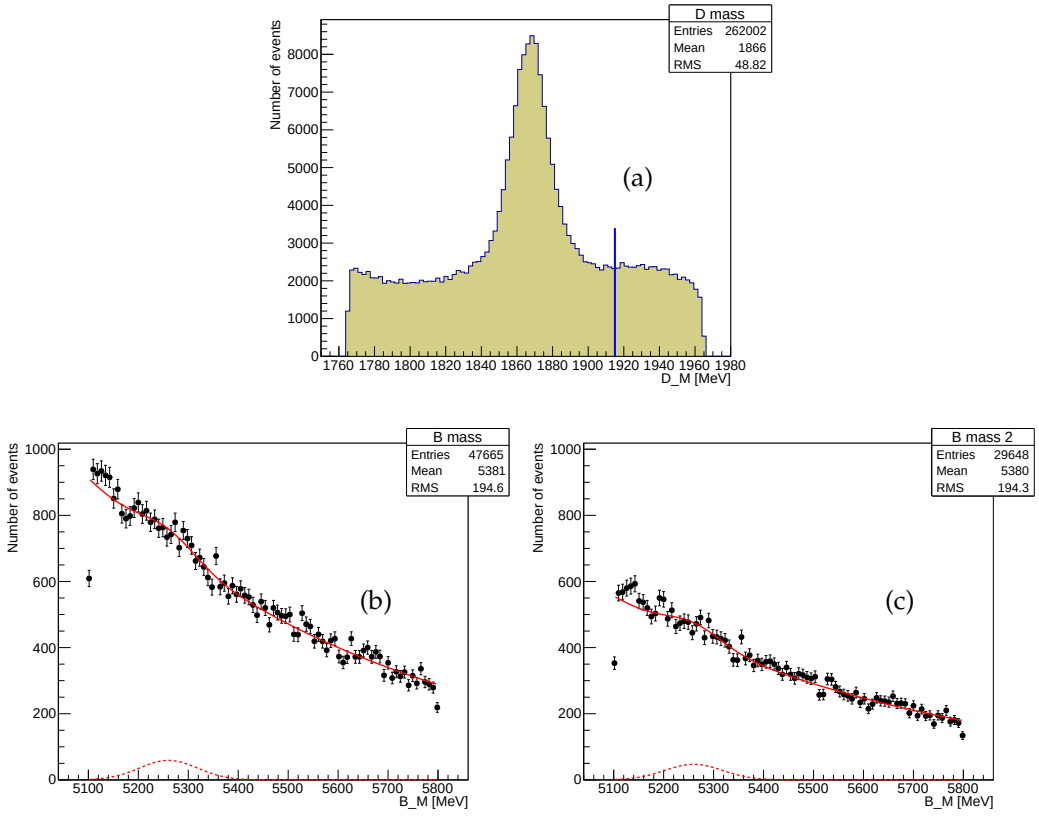


Figure 4.2.: A demonstration of the criterion used to reduce background from events containing fake D-meson candidates, $B \to K^0_S \pi^+ \pi^- h^{(*)}$, reconstructed as $B \to D^{(*)} h^{(*)}$. (a) illustrates the requirement used to select sideband data in the reconstructed D mass spectrum. (b) and (c) show the B-mass distribution of these sideband events before and after the application of the requirement that $z_D^{\text{endvtx}} - z_B^{\text{endvtx}} > 0$. In both (b) and (c), the solid red line shows the result of a fit to the observed distribution using the sum of an exponential and a Gaussian component, and the dashed red line shows the Gaussian component alone. In each plot, the x-axis displays reconstructed mass in MeV/c^2, and the y-axis displays the number of events per histogram bin. Identical binning schemes and axis ranges are used for plots (b) and (c), for ease of comparison.

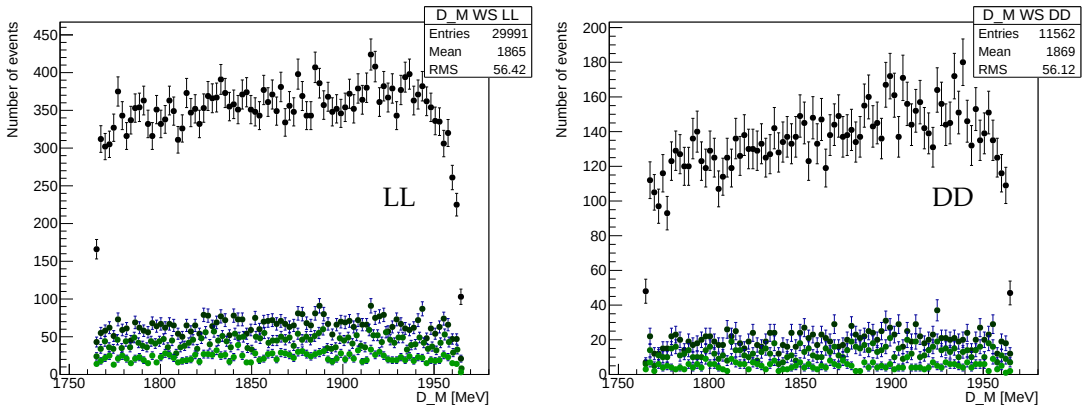


Figure 4.3.: Study performed to check that the MVA classifier used to select fake D-meson candidates does not bias the distribution of combinatoric background events in the reconstructed D mass spectrum. The distribution of reconstructed D-meson mass is shown for combinatoric D-meson candidates, reconstructed in events with LL (left) and DD (right) K^0_S-meson candidates. Distributions are obtained after successively tighter MVA requirements, shown in lighter shades. In both plots, the x-axis displays reconstructed mass in MeV/c^2, and the y-axis displays the number of events per histogram bin.

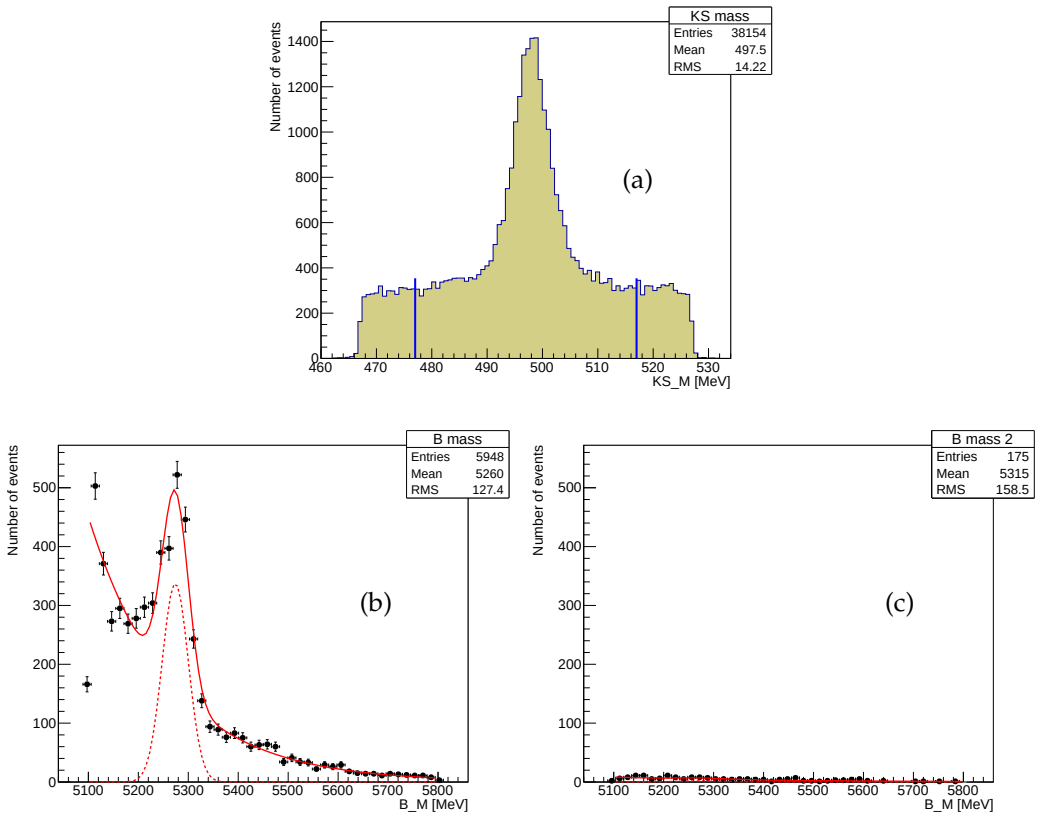


Figure 4.4.: A demonstration of the criterion used to reduce background from events containing fake K^0_S-meson candidates, $B \to (\pi^+\pi^-\pi^+\pi^-)_D h^{(*)}$, reconstructed as $B \to D^{(*)}h^{(*)}$. (a) illustrates the cuts used to select sideband data in the reconstructed K^0_S mass spectrum. (b) and (c) show the B-mass distribution of these sideband events before and after the application of the requirement that $(\chi^2)^{\mathrm{FD,orivtx}}_{K^0_S} > 100$. In both (b) and (c), the solid red line shows the result of a fit to the observed distribution using the sum of an exponential and a Gaussian component, and the dashed red line shows the Gaussian component alone. In each plot, the x-axis displays reconstructed mass in MeV/c^2, and the y-axis displays the number of events per histogram bin. Identical binning schemes and axis ranges are used for plots (b) and (c), for ease of comparison.

4.3. Mode-specific requirements

The mode-specific kinematic and PID (particle identification) requirements used to select candidates for further analysis in each decay mode are detailed in the following subsections, and summarised together in Table 4.4.

4.3.1. $B^{\pm} \to Dh^{\pm}$

Delta-log-likelihood (DLL) PID requirements are placed on the bachelor meson candidate to separate candidate events into the signal and control decay modes. Events for which $\mathrm{DLL}_{K\pi}(\mathrm{bach}) < 4$ are assigned to the control mode, $B^{\pm} \to D\pi^{\pm}$, whilst those for which $\mathrm{DLL}_{K\pi}(\mathrm{bach}) > 4$ are assigned to the signal mode, $B^{\pm} \to DK^{\pm}$. The PID variables used throughout this thesis are all DLLs, as explained in Section 3.2.4.

4.3.2. $B^0_{(s)} \to DK^{*0}$

For candidates reconstructed in the decay modes $B^0_{(s)} \to DK^{*0}$, the absolute value of the cosine of the *helicity angle* of the K^{*0}-meson candidate, $|\cos(\theta^{\mathrm{helicity}}_{K^*})|$, is used as a discriminating variable, since for signal events this angle peaks at 0 and π radians (spin aligned parallel or antiparallel to momentum), whilst for combinatoric K^{*0} candidates it is flat. The requirement $|\cos(\theta^{\mathrm{helicity}}_{K^*})| > 0.4$ is chosen, and is illustrated in Figure 4.5 for a sample of $B^0_s \to DK^{*0}$ signal MC, and a sample of combinatoric background from data, after the application of all other preselection requirements. The combinatoric sample is taken from 'upper sideband' data, $M_B > 5600$ MeV, since combinatoric background dominates peaking components in this region. This requirement matches a previous ADS/GLW analysis [131], ensuring that the same region of B-Dalitz phase space is selected. This allows meaningful comparison of efficiencies in this space.

The mass of the reconstructed K^{*0}-meson candidate, $M_{K^{*0}}$, is required to lie within a 100 MeV mass window centred on its known mass [3]. This window is illustrated in Figure 4.6 for a sample of $B^0_s \to DK^{*0}$ MC, and a sample of combinatoric background from upper sideband data, after the application of all other preselection requirements.

The PID requirements, $\mathrm{DLL}_{K\pi}(K) > 3$, and $\mathrm{DLL}_{K\pi}(\pi) < 3$, are applied to the K^{*0} daughter candidates, to eliminate double counting of events. It is possible for the same physical event to be reconstructed twice within the DaVinci stage, with the identities of the two K^{*0} daughter candidates reversed, to give two separate candidate events. Without this pair of PID requirements it would be possible for both candidate events to pass the selection.

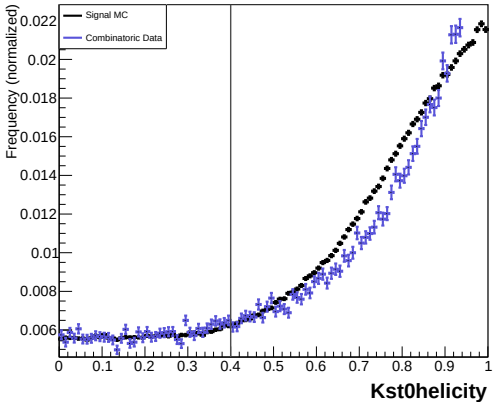


Figure 4.5.: An illustration of the requirement placed on the absolute value of the helicity angle cosine for K^{*0}-meson candidates reconstructed in modes $B^0_{(s)} \to DK^{*0}$. The $|\cos(\theta^{\text{helicity}}_{K^*})|$ distributions obtained in a sample of $B^0_s \to DK^{*0}$ MC and combinatoric background from data are plotted, with a vertical grey line marking the cut position chosen.

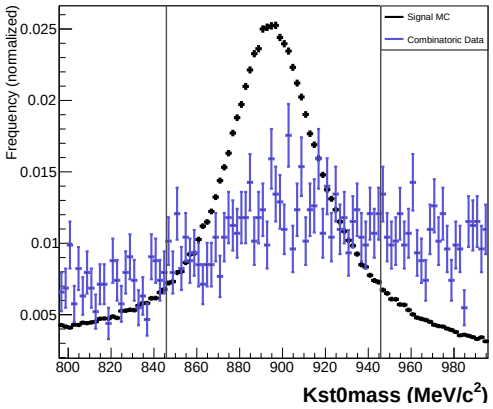


Figure 4.6.: An illustration of the mass window requirement placed on K^{*0}-meson candidates reconstructed in modes $B^0_{(s)} \to DK^{*0}$. The $M_{K^{*0}}$ distributions obtained in a sample of $B^0_s \to DK^{*0}$ MC and combinatoric background from data are plotted, with vertical grey lines marking the cut positions chosen.

4.3.3. $B^{\pm} \to D^* h^{\pm}$

Signal categories

Three distinct categories of signal are possible for these modes, which may be classified depending on the true neutral particles present in the decay, and the neutral candidates which are associated with the event in the reconstruction, as presented in Table 4.3. A real π^0 meson may be correctly reconstructed from either 'merged' or 'resolved' photon candidates, detected in the ECAL subdetector. In the 'merged' case, the two daughter photons interact in the same cell, giving a single energy deposit, which is correlated to the pion mass and momentum. In the 'resolved' case, the two daughter photons interact in separate cells of the ECAL, giving two separate energy deposits which must be combined in the

True decay	Reconstructed neutrals	Reconstructed hypothesis
$D^* \to D\gamma$	γ	$D^* \to D\gamma$
$D^* \to D\pi^0, \pi^0 \to \gamma\gamma$	γ, γ ('resolved') π^0 ('merged')	$D^* \to D\pi^0$
$D^* \to D\pi^0, \pi^0 \to \gamma\gamma$	γ ('partially reconstructed')	$D^* \to D\gamma$

Table 4.3.: Signal categories defined for decay mode $B^\pm \to D^* h^\pm$.

reconstruction to make the pion candidate. Reconstruction is more difficult in this case, and each of the photons may be incorrectly associated directly to the D^* decay vertex; these partially-reconstructed π^0 events thus provide a background to $D^* \to D\gamma$.

For all three cases, the $\mathcal{CP}$ parameters may be extracted, since the same weak phase difference is present in the B-meson decay in each case, and the B-meson flavour may be tagged by the bachelor charge. However, the three signal categories must be distinguished, since a strong phase shift of π radians is present between $D^* \to D\pi^0$ and $D^* \to D\gamma$ [43].

Missing mass variables

In order to achieve discrimination between these categories, it is useful to consider the 'missing mass variables', m_2 and m_3, which are defined, for a decay of the form $D^* \to DX^0$, by

$$m_2 = m(D^*) - m(D), \tag{4.1}$$

$$m_3 = m(D^*) - m(D) - m(X). \tag{4.2}$$

Distributions obtained for these variables, for simulated signal samples are displayed in Figure 4.7 (m_2) and Figure 4.8 (m_3). Separate distributions are shown for each of the three signal categories, which are identified by using truth-level information to determine the true decay chain, and thus the true provenance of the neutral particle reconstructed.

Each of the three signal components are observed to peak in both the m_2 and m_3 spectra. Three separate peaks are observed in m_3, whilst for m_2 the peaks arising from both of the correctly-reconstructed components coincide. This peaking structure suggests that requirements placed on these two variables can be used to efficiently separate the three components. In addition, any such requirements have the potential to remove contamination due to combinatoric candidates. Windows in the m_2 and m_3 spectra are considered, centred on each peak and with a width equal to the RMS of the component distribution found in truth-matched simulation. These windows are illustrated, for both simulation and combinatoric background samples, by shaded regions in Figure 4.9. In order to give the cleanest data samples possible, both the m_2 and m_3 requirements are applied in the final selection. In later analysis, m_3 requirements are also used to separate candidate events matching each of the three 'neutral hypotheses': $X^0 = \gamma$; $X^0 = \pi^0$; and $X^0 =$ partially-reconstructed π^0.

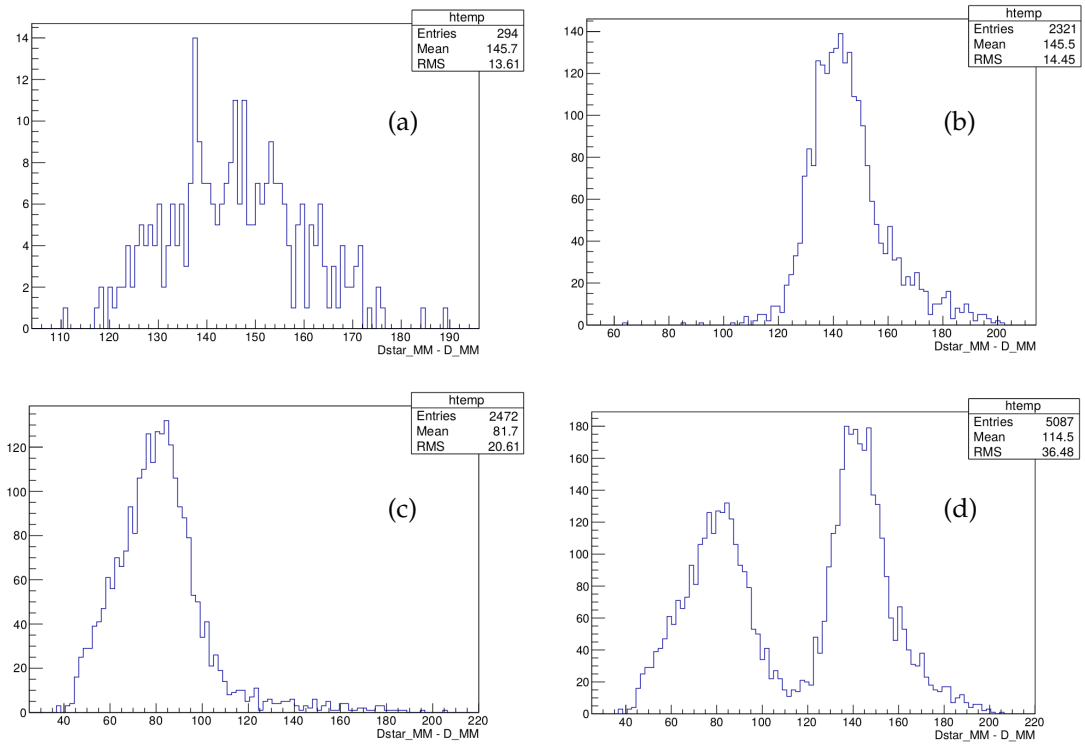


Figure 4.7.: Distributions obtained for missing mass variable $m_2 \equiv m(D^*) - m(D)$, for truth-matched $B^\pm \to D^*\pi^\pm$ Monte Carlo. Simulated events meeting truth-matching requirements for (a) $D^* \to D\pi^0$, (b) $D^* \to D\gamma$ and (c) partially-reconstructed $D^* \to D\pi^0, \pi^0 \to \gamma\gamma$ hypotheses are plotted separately, and are combined together in (d). It is seen that the peaks arising from events matching the hypotheses of (a) and (b) coincide for this variable. In each plot, the x-axis displays missing mass in units of MeV/c^2, and the y-axis displays the number of events per histogram bin.

Requirements

Finally then, the two requirements

- $(60 < m_2 < 100)||(130 < m_2 < 160)$ MeV, and
- $(2.5 < m_3 < 12.5)||(60 < m_3 < 100)||(130 < m_3 < 160)$ MeV

are used to discriminates against combinatoric background in these modes.

In addition, DLL PID requirements are placed on the bachelor meson candidate to separate candidate events into the signal and control decay modes. Events for which $\mathrm{DLL}_{K\pi}(\mathrm{bach}) < 4$ are assigned to the control mode, $B^\pm \to D^*\pi^\pm$, whilst those for which $\mathrm{DLL}_{K\pi}(\mathrm{bach}) > 4$ are assigned to the signal mode, $B^\pm \to D^*K^\pm$.

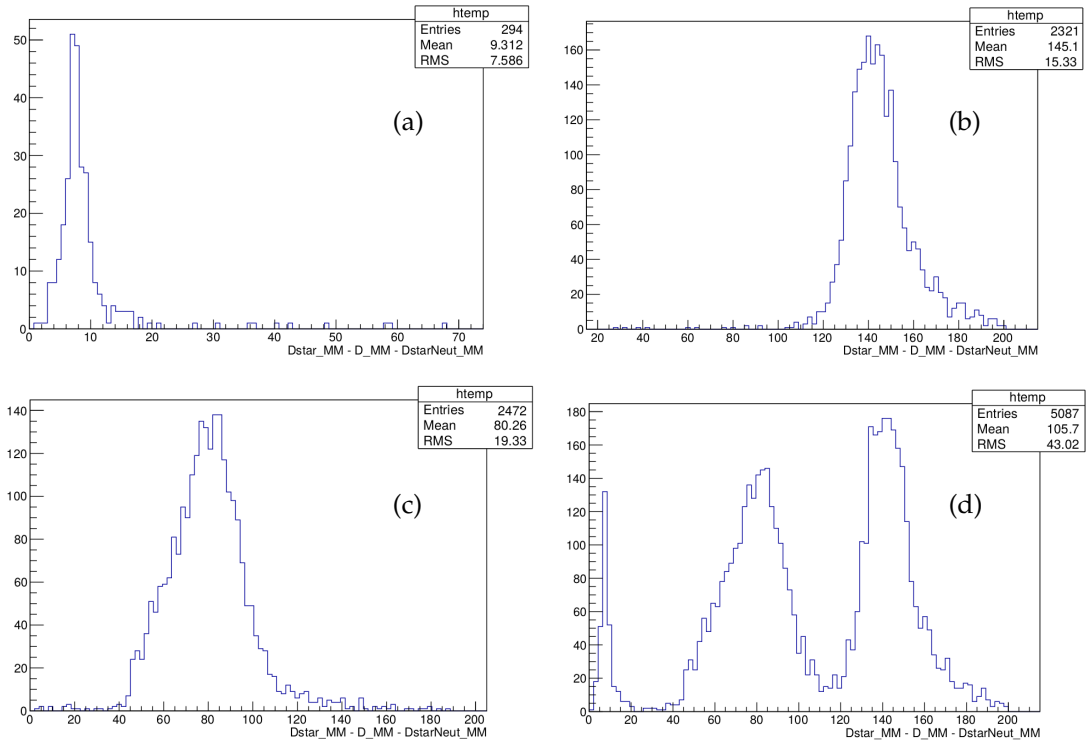


Figure 4.8.: Distributions obtained for missing mass variable $m_3 \equiv m(D^*) - m(D) - m(X)$, for truth-matched $B^\pm \to D^* \pi^\pm$ Monte Carlo. Simulated events meeting truth-matching requirements for (a) $D^* \to D\pi^0$, (b) $D^* \to D\gamma$ and (c) partially-reconstructed $D^* \to D\pi^0, \pi^0 \to \gamma\gamma$ hypotheses are plotted separately, and are combined together in (d). In each plot, the x-axis displays missing mass in units of MeV/c^2, and the y-axis displays the number of events per histogram bin.

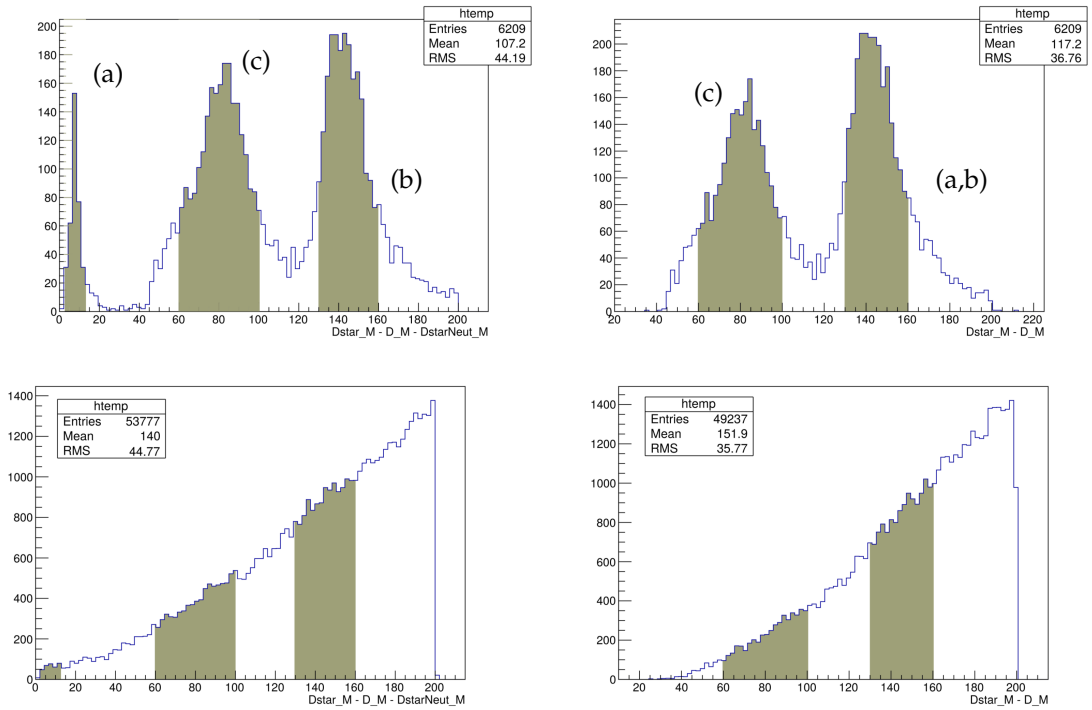


Figure 4.9.: Criteria considered to select data candidates matching (a) $D^* \to D\pi^0$, (b) $D^* \to D\gamma$ and (c) partially-reconstructed $D^* \to D\pi^0$, $\pi^0 \to \gamma\gamma$ hypotheses, using the variables (left) $m_3 \equiv m(D^*) - m(D) - m(X)$ and (right) $m_2 \equiv m(D^*) - m(D)$. The selected ranges considered are shown as shaded regions, for distributions taken from (top) truth-matched $B^\pm \to D^*\pi^\pm$ Monte Carlo simulation and (bottom) a sample of combinatoric $B^\pm \to D^*h^\pm$ candidates taken from the reconstructed B-mass data sideband ($> 5800\,\text{MeV}$). In each plot, the x-axis displays missing mass in units of MeV/c^2, and the y-axis displays the number of events per histogram bin.

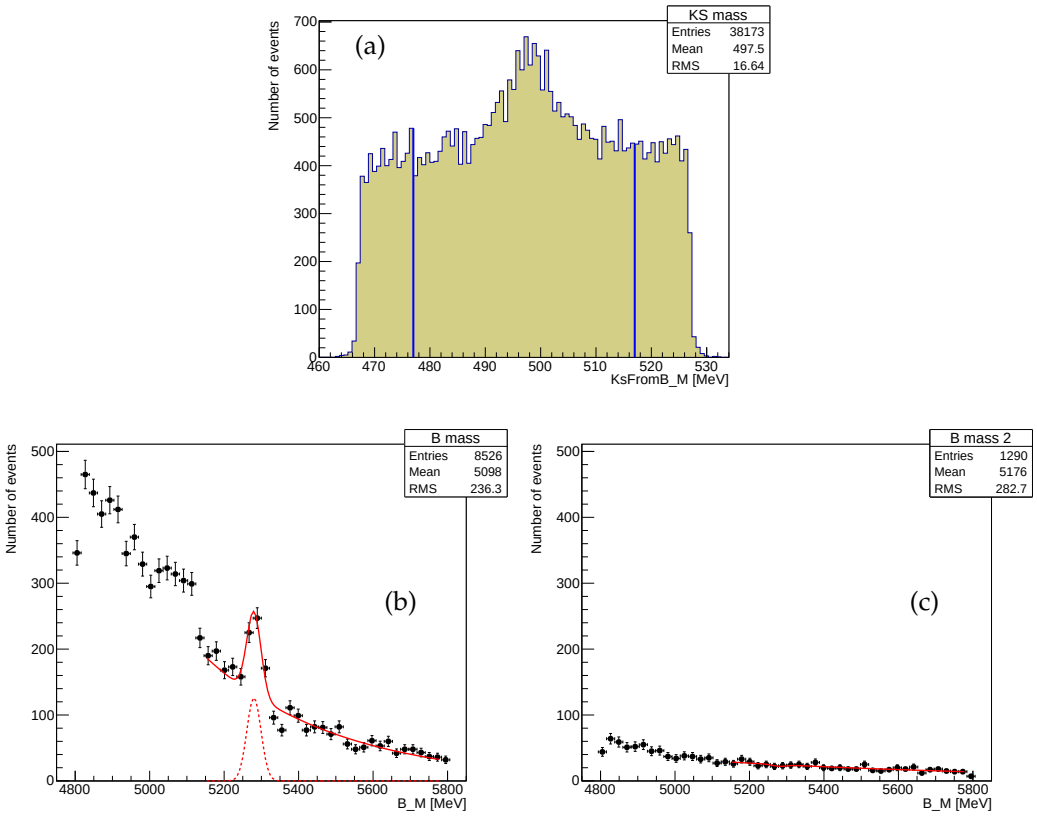


Figure 4.10.: A demonstration of the criterion used to reduce background from events containing fake K^0_S-meson candidates, $B^\pm \to D(\pi^+\pi^-\pi^\pm)_{K^{*\pm}}$, reconstructed as $B^\pm \to D(K^0_S\pi^\pm)_{K^{*\pm}}$. (a) illustrates the cuts used to select sideband data in the reconstructed K^0_S mass spectrum. (b) and (c) show the B-mass distribution of these sideband events before and after the application of the requirement that $(\chi^2)^{\mathrm{FD,orivtx}}_{K^0_S} > 25.0$. In both (b) and (c), the solid red line shows the result of a fit to the observed distribution using the sum of an exponential and a Gaussian component, and the dashed red line shows the Gaussian component alone. In each plot, the x-axis displays reconstructed mass in MeV/c^2, and the y-axis displays the number of events per histogram bin. Identical binning schemes and axis ranges are used for plots (b) and (c), for ease of comparison.

4.3.4. $B^\pm \to DK^{*\pm}$

For this mode, the $K^{*\pm}$-meson is reconstructed in the final state $K^0_S\pi^\pm$, meaning that an extra K^0_S-meson is present compared to the other modes considered. Background due to fake candidates is also possible for this meson, and is handled using the same approach detailed in Section 4.2.1. For these fake events, the $K^{*\pm}$ decays to the state $(\pi^+\pi^-\pi^\pm)$ instead of $(K^0_S\pi^\pm)$. The validation of the flight-distance χ^2 requirement used to reduced the contribution of these non-real events is presented in Figure 4.10. The requirement $(\chi^2)^{\mathrm{FD,oriVX}}_{K^0_S} > 25$ is observed to be highly effective at removing such events.

Mass window requirements are placed on $K^{*\pm}$-meson candidates, and on their K^0_S-meson daughter candidates. The $K^{*\pm}$-meson mass is required to lie within a 150 MeV window centred on its mass [3]. The requirement on the mass of the K^0_S-meson is identical to that placed on the K^0_S-meson arising from

the D-meson decay. These windows are shown in Figure 4.11 for a sample of $B^{\pm} \to DK^{*\pm}$ MC, after the application of all other preselection requirements.

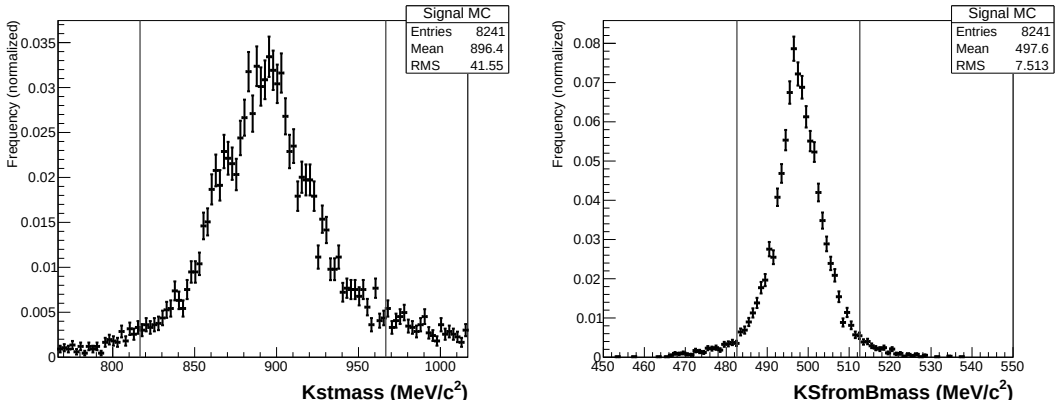


Figure 4.11.: An illustration of the mass window requirements placed on $K^{*\pm}$-meson candidates (left) and their K^0_S-meson daughter candidates (right), reconstructed in the $B^{\pm} \to DK^{*\pm}$ signal mode. The requirements $|M_{K^{*\pm}} - 891.66| < 75$ MeV and $482.6 < M_{K^0_S} < 512.6$ MeV are used. The $M_{K^{*\pm}}$ and $M_{K^0_S}$ distributions obtained from a sample of $B^{\pm} \to DK^{*\pm}$ MC events are plotted, with vertical grey lines marking the cut positions chosen.

As for the decay modes $B^0_{(s)} \to DK^{*0}$, the absolute value of the helicity angle cosine of the $K^{*\pm}$-meson candidate, $|\cos(\theta^{\text{helicity}}_{K^*})|$, is used as a discriminating variable. The requirement $|\cos(\theta^{\text{helicity}}_{K^*})| > 0.3$ is chosen, and is illustrated in Figure 4.12 for a sample of $B^{\pm} \to DK^{*\pm}$ signal MC, and a sample of combinatoric background from data, after the application of all other preselection requirements. The combinatoric sample is taken from upper sideband data, $M_B > 5600$ MeV.

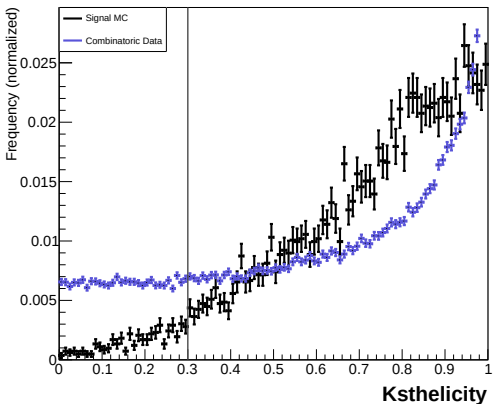


Figure 4.12.: An illustration of the requirement placed on the absolute value of the helicity angle cosine for $K^{*\pm}$-meson candidates reconstructed in the decay mode $B^{\pm} \to DK^{*\pm}$. The $|\cos(\theta^{\text{helicity}}_{K^*})|$ distributions obtained in a sample of $B^{\pm} \to DK^{*\pm}$ MC and combinatoric background from data are plotted, with a vertical grey line marking the cut position chosen.

Mode	Requirement	Applied to MVA training samples?	Description
$B^\pm \to D\pi^\pm$	$\mathrm{DLL}_{K\pi}(\mathrm{bach}) < 4$	No	PID criterion to remove mis-id bachelor candidates
$B^\pm \to DK^\pm$	$\mathrm{DLL}_{K\pi}(\mathrm{bach}) > 4$	No	PID criterion to remove mis-id bachelor candidates
$B^0 \to DK^{*0}$	$\mathrm{DLL}_{K\pi}(K) > 3 \,\&\&\, \mathrm{DLL}_{K\pi}(\pi) < 3$	Yes	Eliminates double counting due to PID-swapped K^{*0} daughter candidates
	$\lvert\cos(\theta^{\mathrm{helicity}}_{K^*})\rvert > 0.4$	Yes	Gives compatibility with previous ADS/GLW analysis
	$\lvert M_{K^{*0}} - 895.81\rvert < 50$ MeV	No	Selects K^{*0} resonance within B-decay Dalitz plane
$B^\pm \to D^*\pi^\pm$	$\mathrm{DLL}_{K\pi}(\mathrm{bach}) < 4$	No	PID criterion to remove mis-id bachelor candidates
$B^\pm \to D^*K^\pm$	$\mathrm{DLL}_{K\pi}(\mathrm{bach}) > 4$	No	PID criterion to remove mis-id bachelor candidates
$B^\pm \to D^*h^\pm$	$(60 < m_2 < 100)\lvert\lvert(130 < m_2 < 160)$ MeV	No	Discriminates against combinatoric background
	$(2.5 < m_3 < 12.5)\lvert\lvert(60 < m_3 < 100)\lvert\lvert(130 < m_3 < 160)$ MeV	No	Discriminates against combinatoric background
$B^\pm \to DK^{*\pm}$	$482.6 < M_{K^0_S} < 512.6$ MeV	Yes	Mass window on K^0_S from $K^{*\pm}$ decay
	$(\chi^2)^{\mathrm{FD,oriVX}}_{K^0_S} > 25$	Yes	Applied for LL K^0_S candidates from $K^{*\pm}$ decay only
	$\lvert M_{K^{*\pm}} - 891.66\rvert < 75$ MeV	No	Selects $K^{*\pm}$ resonance within B Dalitz plane
	$\lvert\cos(\theta^{\mathrm{helicity}}_{K^{*\pm}})\rvert > 0.3$	No	Selects $K^{*\pm}$ resonance within B Dalitz plane

Table 4.4.: Mode-specific requirements used to select candidates for further analysis in each mode. Only some of these requirements are applied to the samples used to train MVA classifiers, as described in Section 4.4.1.

4.4. MVA selection

Each cut described so far represents a simple univariate binary classifier - it takes a single variable as input, and separates data into two categories, signal and background. Combining several such cuts yields a multivariate *linear* classifier (or multivariate analysis, 'MVA'), since each input variable is treated independently.

However, in reality, it is to be expected that the optimal multivariate classifier may depend on a *non-linear* combination of many input variables. Essentially, one wishes to answer the question, 'what function $f(\vec{x})$ will best map the space of observables $\{x_j\}$ onto the classification space $\{\text{signal}, \text{background}\}$?'. This is an 'inverse problem', for which there is no general analytic solution. Instead, a range of algorithm-based approaches have been developed, collectively known as 'machine learning'.

One common approach to machine learning is known as 'supervised learning'. This method uses a trial classifier, which is optimised iteratively using a set of 'labelled' data, consisting of events whose category is known *a priori*. The parameters of the classifier are optimised to give the best possible match between the classifier output and the known categories of the training data, before the classifier is then applied on the experimental ('unlabelled') dataset. This optimisation process is known as 'training', and the labelled dataset used is known as 'training data'. Often, a subset of the training data is reserved for use as a validation dataset, to quantify the performance of the trained classifier.

Each multivariate classifier used within this thesis is implemented using the `TMVA` [132] module of the ROOT package, and is trained using supervised learning. Each classifier is set up so that for each event it outputs a single numerical value between zero and one, where values closer to one indicate greater confidence that the event belongs to the signal category, and values closer to zero indicate greater confidence that the event belongs to the background category. In each mode, two different types of classifier are considered - decision tree ensembles, and neural networks. These are described in the following subsections.

Decision trees

A decision tree is a multivariate classifier built using a set of sequentially-applied binary univariate cuts on the input observables. It uses these cuts to divide the space of input observables into regions which are dominated by each category. A simple example of a decision tree used in a biological setting is shown in Figure 4.13, for the case of a classifier trained to distinguish sick and healthy patients using two input observables $(X1, X2)$ [133]. Applying repeated cuts on the same input variable allows region boundaries of arbitrary complexity to be modelled. It is worth noting that the set of all preselection cuts defined so far in this chapter, taken together, form a decision tree, defining a decision boundary in N-dimensional space, where N is the number of variables for which preselection cuts are specified.

Decision trees may be trained (or 'built') using the CART (Classification And Regression Tree) algorithm [127]. Each successive cut of the tree is chosen by finding the input observable and cut value that divide the training data into the purest possible pair of subsamples. The algorithm ceases building

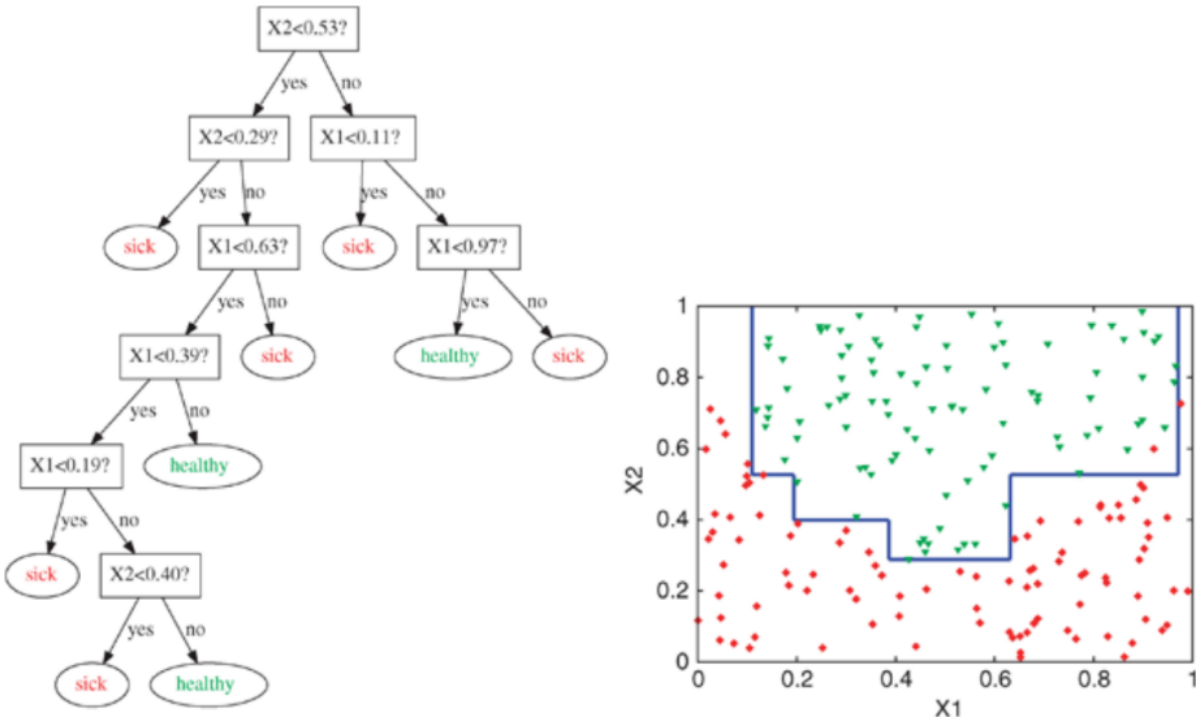


Figure 4.13.: An example decision tree (left) and the decision boundary it defines. The tree is trained to distinguish between two categories of data ('sick' and 'healthy'), using two input observables ($X1$ and $X2$) [133].

new layers of the tree once it has reached a user specified depth. Finally, those branch points which do not give statistically significant separation between categories are removed, in a process known as 'pruning'.

The CART algorithm is 'greedy' since at each branch point it chooses the optimal cut, with no consideration of the coming layers of the tree. As such, it is not guaranteed to find the globally optimal solution. Decision trees are also prone to overtraining, since the position of successive cuts can be sensitive to statistical fluctuations in the data, altering the structure of the tree created [126].

'Ensemble methods' are used to overcome these shortfalls. These work by combining many decision trees to improve accuracy whilst reducing the likelihood of overtraining. In particular, the 'AdaBoost' (or adaptive boosting) algorithm [134] is used to train ensembles of trees within this thesis. This method begins by training a single tree, and using it to classify the training data. The dataset is then reweighted to give a greater weight to misclassified events, and a second tree is trained on the reweighted data. This process is repeated iteratively to create an ensemble of trees. Finally, predictions on new data are made by combining the output of all trees in the ensemble, weighted by their accuracy.

Within this thesis, each AdaBoost ensemble trained contains 800 decision trees, each with a maximum depth of three. The 'learning rate', which determines how strongly the training data is reweighted at each step, is set to 0.3. Each ensemble of decision trees generated using the AdaBoost method will be referred to throughout this thesis as 'AdaBDT'.

Neural networks

A neural network is typically represented by a 'network diagram', consisting of an array of 'nodes', arranged in successive 'layers'. An example network diagram is shown in Figure 4.14. Each node of a given layer is connected to every node in the layer on its left, and every node in the layer on its right. A neural network set up using several layers of nodes is often referred to as a 'multi-layer perceptron' (MLP). The first layer of the network is the 'input layer', and the final layer is the 'output layer'. The layers in between are called 'hidden layers'.

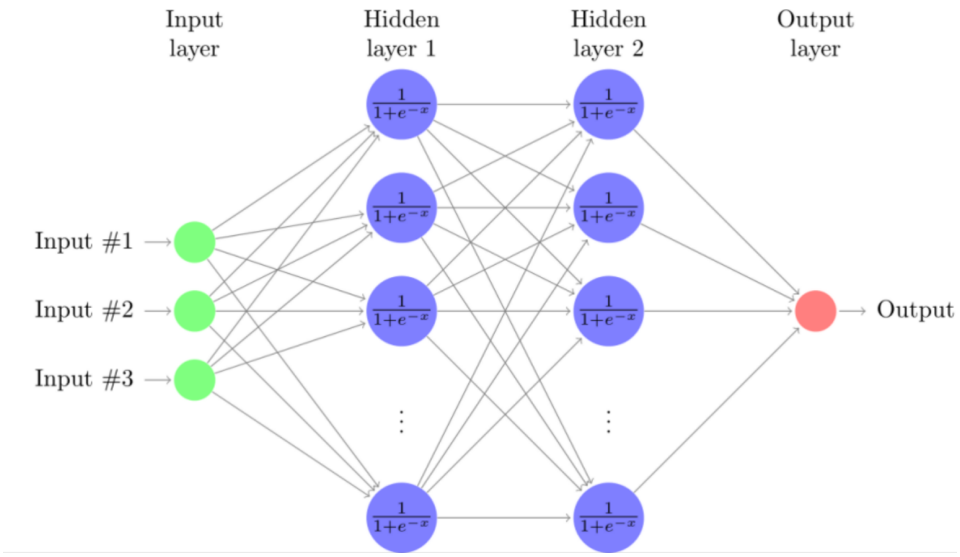


Figure 4.14.: Node diagram for a four-layer neural network [135].

An evaluation of the network output, for a single set of input values, proceeds as a single 'forward pass', with each node taking input values from the layer on its left, and passing its output to the layer on its right. The output values of each node in the l-th layer are calculated as

$$z_{lj} = \sigma_l\left(\alpha_{lj0} + \sum_{i=0}^{N_{l-1}} \alpha_{lji} z_{(l-1)i}\right) \quad , \tag{4.3}$$

where i indexes the nodes in the preceding layer, and j indexes the nodes in the current layer. σ_l denotes a univariate function, the 'activation function', which may be different for the nodes in each layer. The values α_{ljk} are constants, known as 'synapse weights', since they parameterize the degree to which each node is influenced by each of the preceding nodes.

This equation shows that at each node a weighted sum of input values from the preceding layer is formed, which is then fed to the activation function σ_l. A nonlinear function is used for the hidden layers (this is what allows a neural network to replicate nonlinear dependencies in the data), whilst the output layer may use a different function depending on the properties desired for the output variable. A commonly-used nonlinear function is the sigmoid,

$$\sigma_{\text{sigmoid}}(\nu) = \frac{1}{1+e^{-\nu}} \quad , \tag{4.4}$$

as plotted in Figure 4.15. The calculation performed at each node is simple, but it has been shown that a single neural network is mathematically sufficient to describe any multivariate function.[3]

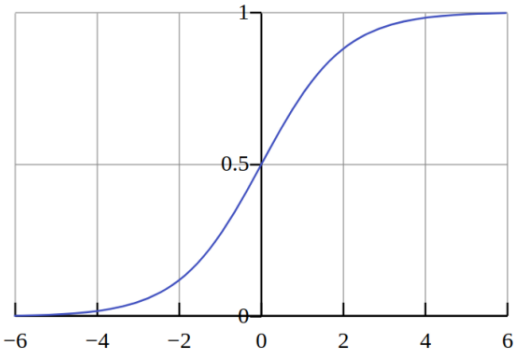


Figure 4.15.: The sigmoid activation function, $1/1 + e^{-\nu}$.

Neural networks are trained using 'backpropagation'. For each training datum, this algorithm calculates the network output, also recording the gradient of the activation function at each node. The error in the network output (i.e. the difference with respect to the training label) is propagated back through the nodes using the stored gradients, and used to update the weights, $\{\alpha_{l,j,k}\}$. This process is repeated for each datum in the training set, and the trained network is then specified by the final set of weights obtained.

Within this thesis, all MLP classifiers trained have a single hidden layer, with a number of nodes approximately 1.3 times that of the input layer. The activation function used is the hyperbolic tangent, $\tanh(x)$, since its gradient does not tend to zero as fast as the sigmoid function for $x \to \infty$ or $x \to -\infty$, reducing the chance of the backpropagation algorithm getting stuck in a region of near-zero gradient. The error function minimised by the backpropagation algorithm is chosen to be the 'cross-entropy' of the training data [126]. The number of training 'epochs' used is 750, meaning that the entire training dataset is iterated over 750 times. Each MLP classifier trained in this way will be referred to throughout this thesis as 'MLPBPCE'.

4.4.1. Training

Training samples

For all of the MVAs trained in this thesis, labelled background samples are taken from data lying in the upper sideband of the B-meson mass spectrum, since this region is dominated by combinatoric background. Labelled signal samples cannot be obtained from data, so instead truth-matched signal MC events are used. Separate classifiers are trained for each mode, and for each class of K^0_S candidates (LL and DD), since the distributions of the input kinematic variables are expected to differ in each case.

[3] More formally, it can be shown that a neural network is mathematically equivalent to a Fourier decomposition for any Borel-measurable function. Consequently, any such function can be represented to any desired degree of accuracy by including sufficiently many nodes in the network [136].

Mode	K_S type (from D)	Signal training sample description	Signal training sample no. events	Background training sample description	Background training sample no. events
$B^\pm \to Dh^\pm$	LL	$B^\pm \to D\pi^\pm$ MC	88509	$B^\pm \to D\pi^\pm$ data,	14929
	DD		271453	B_PVFit_M $>$ 5800	23336
$B^0 \to DK^{*0}$	LL	$B^0 \to DK^{*0}$ MC	22843	$B^0 \to DK^{*0}$ data,	4262
	DD		67584	B_PVFit_M $>$ 5800	8361
$B^\pm \to D^*h^\pm$	LL	$B^\pm \to D^*\pi^\pm$ MC	1548	$B^\pm \to D^*\pi^\pm$ data,	9414
	DD		5790	B_PVFit_M $>$ 5800	14000
$B^\pm \to DK^{*\pm}$	LL	$B^\pm \to DK^{*\pm}$ MC	1353	$B^\pm \to DK^{*\pm}$ data,	3354
	DD		4535	B_PVFit_M $>$ 5600	5952

Table 4.5.: Samples available for training each MVA classifier, after preselection cuts have been applied.

All common preselection cuts (as shown in Table 4.2) are applied before training of MVAs, along with some of the mode-specific cuts, as specified in Table 4.4. In particular, since PID variables are found to be imperfectly modelled in simulation [137], no PID requirements are placed on the training samples (as all signal samples are taken from simulation).

The numbers of candidates available for training each MVA classifier, after preselection cuts, are shown in Table 4.5. For each sample, 20% of these events are reserved for use as a validation sample, and are not used in the learning process. The limited statistics available for training, especially for modes $B^\pm \to D^*h^\pm$ and $B^\pm \to DK^{*\pm}$, motivate several choices later in the analysis. For now it is useful to note that no events from the signal and background samples are kept aside as a 'hold-back' dataset; all are used for either training or validation.

Input variables

The input variables to each MVA are chosen to be common to all decay modes. For those preselection requirements that are placed on the real data but not on on the MVA training samples, it is not efficient to include these variables as inputs to the MVA classifier, since the learning process may then rely on the information contained in these variables, which will not be present in the real data. However, for the requirements that are placed on both the real data and the MVA training samples, it is safe to include these same variables as inputs to the MVAs, and they may provide additional discriminating power.

A range of possible inputs are considered, and the following are chosen:

- the natural logarithm of the χ^2 per degree of freedom of the refit of the reconstructed B-meson candidate, $\ln((\chi^2/\text{ndof})_{\texttt{PVrefit}})$. Logarithms are applied to input variables that otherwise exhibit long upper tails, in order to create more symmetrical input distributions for training;
- the cosine of the direction angle, for reconstructed B-, D- and K^0_S-meson candidates: $\cos(\theta_B^{\text{DIRA}})$; $\cos(\theta_D^{\text{DIRA}})$; and $\cos(\theta_{K^0_S}^{\text{DIRA}})$. (Note that throughout this section K^0_S refers to the candidate originating in the D-meson decay. For decay mode $B^\pm \to DK^{*\pm}$, where an additional K^0_S-meson

candidate is reconstructed in the $K^{*\pm}$-meson decay, no variables pertaining to this candidate are included as MVA inputs);

- the natural logarithm of the momentum of the B-, D-meson candidates, $\ln(p_B)$ and $\ln(p_D)$;
- the impact parameter χ^2, with respect to the PV, for the reconstructed B-meson candidate, $(\chi^2)^{\text{IP,PV}}_B$. No logarithm is applied to this variable, since low values are expected, as the B-meson track should be consistent with originating at the PV;
- the natural logarithm of the impact parameter χ^2, with respect to the PV, for the reconstructed D-, K^0_S- and bachelor meson candidates: $\ln((\chi^2)^{\text{IP,PV}}_D)$; $\ln((\chi^2)^{\text{IP,PV}}_{K^0_S})$; and $\ln((\chi^2)^{\text{IP,PV}}_{\text{bach}})$. For the decay mode $B^0 \to DK^{*0}$, the charged kaon produced in the decay of the K^{*0} is treated as the bachelor, whilst for the decay mode $B^\pm \to DK^{*\pm}$, the charged pion produced in the decay of the $K^{*\pm}$ is used;
- the natural logarithm of the maximum and minimum impact parameter χ^2, with respect to the PV, taken across both pion daughters of the reconstructed D-meson candidate, $\ln(\max((\chi^2)^{\text{IP,PV}}_{\pi^+_D}, (\chi^2)^{\text{IP,PV}}_{\pi^-_D}))$ and $\ln(\min((\chi^2)^{\text{IP,PV}}_{\pi^+_D}, (\chi^2)^{\text{IP,PV}}_{\pi^-_D}))$. These variables are symmetrised in the two daughter pions, to avoid introducing a charge asymmetry;
- the natural logarithm of the maximum and minimum impact parameter χ^2, with respect to the PV, taken across both pion daughters of the reconstructed K^0_S-meson candidate, $\ln(\max((\chi^2)^{\text{IP,PV}}_{\pi^+_{K^0_S}}, (\chi^2)^{\text{IP,PV}}_{\pi^-_{K^0_S}}))$ and $\ln(\min((\chi^2)^{\text{IP,PV}}_{\pi^+_{K^0_S}}, (\chi^2)^{\text{IP,PV}}_{\pi^-_{K^0_S}}))$;
- the natural logarithm of the χ^2 per degree of freedom in the vertex fit at the decay vertex of the B- and D-meson candidates, $\ln((\chi^2/\text{ndof})^{\text{endvtx}}_B)$ and $\ln((\chi^2/\text{ndof})^{\text{endvtx}}_D)$;
- the natural logarithm of the flight distance χ^2 of the reconstructed B-, D- and K^0_S-meson candidates, with respect to the PV: $\ln((\chi^2)^{\text{FD,ownPV}}_B)$; $\ln((\chi^2)^{\text{FD,ownPV}}_D)$; and $\ln((\chi^2)^{\text{FD,ownPV}}_{K^0_S})$;
- the *significance* of the flight distance of the K^0_S-meson candidate with respect to its own creation vertex, constructed as

$$(z^{\text{endvtx}}_{K^0_S} - z^{\text{endvtx}}_D) \Big/ \sqrt{(\Delta z^{\text{endvtx}}_{K^0_S})^2 + (\Delta z^{\text{endvtx}}_D)^2}, \tag{4.5}$$

  where $\Delta z^{\text{endvtx}}_{K^0_S}$ and $\Delta z^{\text{endvtx}}_D$ give the uncertainty on the z-coordinate of the K^0_S- and D-meson decay vertices; and
- the *lifetime* of the D-meson candidate, τ_D.

A summary of the input variables chosen is shown in Table 4.6.

Table 4.6.: Input variables to the MVA classifier.

MVA Input variable	Description
$\cos(\theta_B^{\mathrm{DIRA}})$	cosine of the angle between the reconstructed B candidate momentum and direction vectors
$(\chi^2)_B^{\mathrm{IP,PV}}$	impact parameter χ^2 of the reconstructed B candidate with respect to the PV
$\ln((\chi^2)_B^{\mathrm{FD,ownPV}})$	log of the flight distance χ^2 of the reconstructed B candidate with respect to the PV
$\ln((\chi^2/\mathrm{ndof})_{\texttt{PVrefit}})$	log of the refit $\chi^2/ndof$ of the reconstructed B candidate
$\cos(\theta_D^{\mathrm{DIRA}})$	cosine of the angle between the reconstructed D candidate momentum and direction vectors
$\ln((\chi^2)_D^{\mathrm{IP,PV}})$	log of the impact parameter χ^2 of the reconstructed D candidate with respect to the PV
$\cos(\theta_{K_S^0}^{\mathrm{DIRA}})$	cosine of the angle between the reconstructed K_S^0 candidate momentum and direction vectors
$\ln((\chi^2)_{K_S^0}^{\mathrm{IP,PV}})$	log of the impact parameter χ^2 of the reconstructed K_S^0 candidate with respect to the PV
$\ln((\chi^2)_{\mathrm{bach}}^{\mathrm{IP,PV}})$	log of the impact parameter χ^2 of the bachelor with respect to the PV
$\ln(\max((\chi^2)_{\pi_D^+}^{\mathrm{IP,PV}}, (\chi^2)_{\pi_D^-}^{\mathrm{IP,PV}}))$	symmetrised log of the impact parameter χ^2 of the reconstructed D daughter pions with respect to the PV
$\ln(\min((\chi^2)_{\pi_D^+}^{\mathrm{IP,PV}}, (\chi^2)_{\pi_D^-}^{\mathrm{IP,PV}}))$	symmetrised log of the impact parameter χ^2 of the reconstructed D daughter pions with respect to the PV
$\ln(p_B)$	log of the reconstructed B candidate momentum
$\ln(p_D)$	log of the reconstructed D candidate momentum
$\ln(\max((\chi^2)_{\pi_{K_S^0}^+}^{\mathrm{IP,PV}}, (\chi^2)_{\pi_{K_S^0}^-}^{\mathrm{IP,PV}}))$	symmetrised log of the impact parameter χ^2 of the reconstructed K_S^0 daughter pions with respect to the PV
$\ln(\min((\chi^2)_{\pi_{K_S^0}^+}^{\mathrm{IP,PV}}, (\chi^2)_{\pi_{K_S^0}^-}^{\mathrm{IP,PV}}))$	symmetrised log of the impact parameter χ^2 of the reconstructed K_S^0 daughter pions with respect to the PV
$\ln((\chi^2/\mathrm{ndof})_B^{\mathrm{endvtx}})$	log of the B candidate decay vertex $\chi^2/ndof$
$\ln((\chi^2/\mathrm{ndof})_D^{\mathrm{endvtx}})$	log of the D candidate decay vertex $\chi^2/ndof$
$\ln((\chi^2)_D^{\mathrm{FD,ownPV}})$	log of the flight distance χ^2 of the reconstructed D candidate with respect to the PV
$\ln((\chi^2)_{K_S^0}^{\mathrm{FD,ownPV}})$	log of the flight distance χ^2 of the reconstructed K_S^0 candidate with respect to the PV
$(z_{K_S^0}^{\mathrm{endvtx}} - z_D^{\mathrm{endvtx}})/\sqrt{(\Delta z_{K_S^0}^{\mathrm{endvtx}})^2 + (\Delta z_D^{\mathrm{endvtx}})^2}$	flight distance significance of the K_S^0 candidate with respect to its origin vertex
τ_D	D candidate lifetime

In order to verify that the MVA input variables are well-modelled in the MC signal samples used for training[4], their distributions are compared for a sample of $B^{\pm} \to D\pi^{\pm}$ MC, a sample of real $B^{\pm} \to D\pi^{\pm}$ signal data, and a sample of combinatoric background taken from the upper sideband of the $B^{\pm} \to D\pi^{\pm}$ data. The preselection requirements of Table 4.2 are applied to each sample. In addition, for the signal data sample a requirement is placed on the reconstructed B-meson mass, which is constrained to lie within 30 MeV of its mean (as returned by the full mass spectrum fit, described in Chapter 5) $5252.0 < M_B < 5312.0$ MeV, to select the signal peak. For the combinatoric background sample, the reconstructed B-meson mass is required to lie above 5600 MeV. This gives 41647 signal data events, 63065 combinatoric data events, and 381987 MC events for comparison.

Distributions of each input variable, for each of the three samples described, are shown in Figures 4.16, 4.17, 4.18 and 4.19. For the flight distance significance of the K^0_S candidate and the symmetrised logarithms of the impact parameter χ^2 across the K^0_S daughter pions, separate comparisons are made for events containing LL and DD K^0_S meson candidates, since the distributions in each case differ significantly. In each case, the distributions in signal data and signal MC are very close, and quite different from the distribution in background data.[5] This indicates that training MVAs using signal MC and combinatoric background data should yield classifiers capable of separating signal data from combinatoric background data.

Distributions of all input variables and their correlations, for the signal and background training samples introduced in Table 4.5, are contained in Appendix D.

[4] If the input variables are not well-modelled, the MVA classifier may learn to rely on features which are not present in real data, giving poorer discrimination.

[5] The signal data and signal MC distributions are not expected to coincide, since the signal data sample retains some level of background contamination.

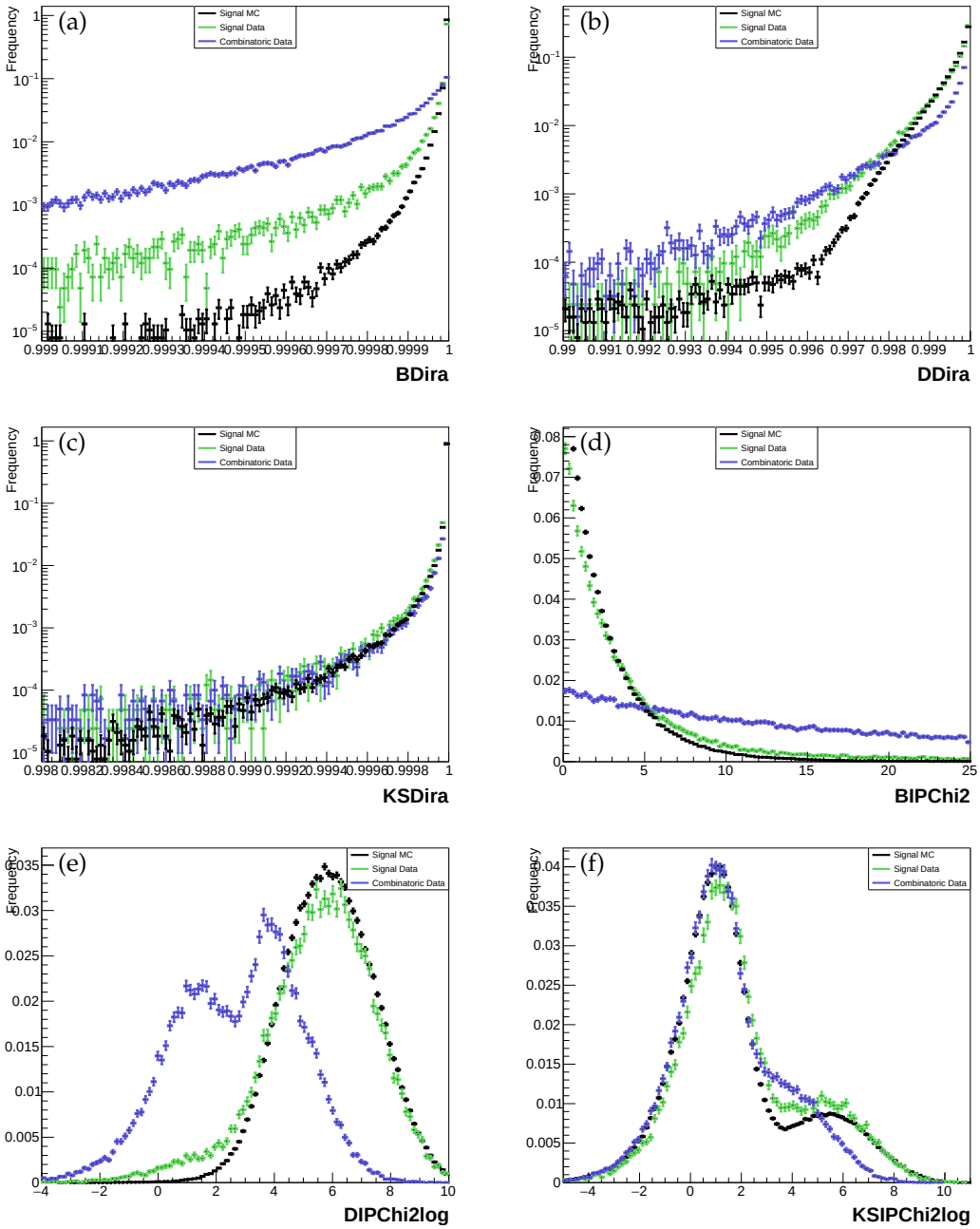


Figure 4.16.: Normalized distributions obtained for (a) $\cos(\theta_B^{\mathrm{DIRA}})$, (b) $\cos(\theta_D^{\mathrm{DIRA}})$, (c) $\cos(\theta_{K_S^0}^{\mathrm{DIRA}})$, (d) $(\chi^2)_B^{\mathrm{IP,PV}}$, (e) $\ln((\chi^2)_D^{\mathrm{IP,PV}})$, and (f) $\ln((\chi^2)_{K_S^0}^{\mathrm{IP,PV}})$, obtained in signal MC (black), signal data (green) and combinatoric background (blue).

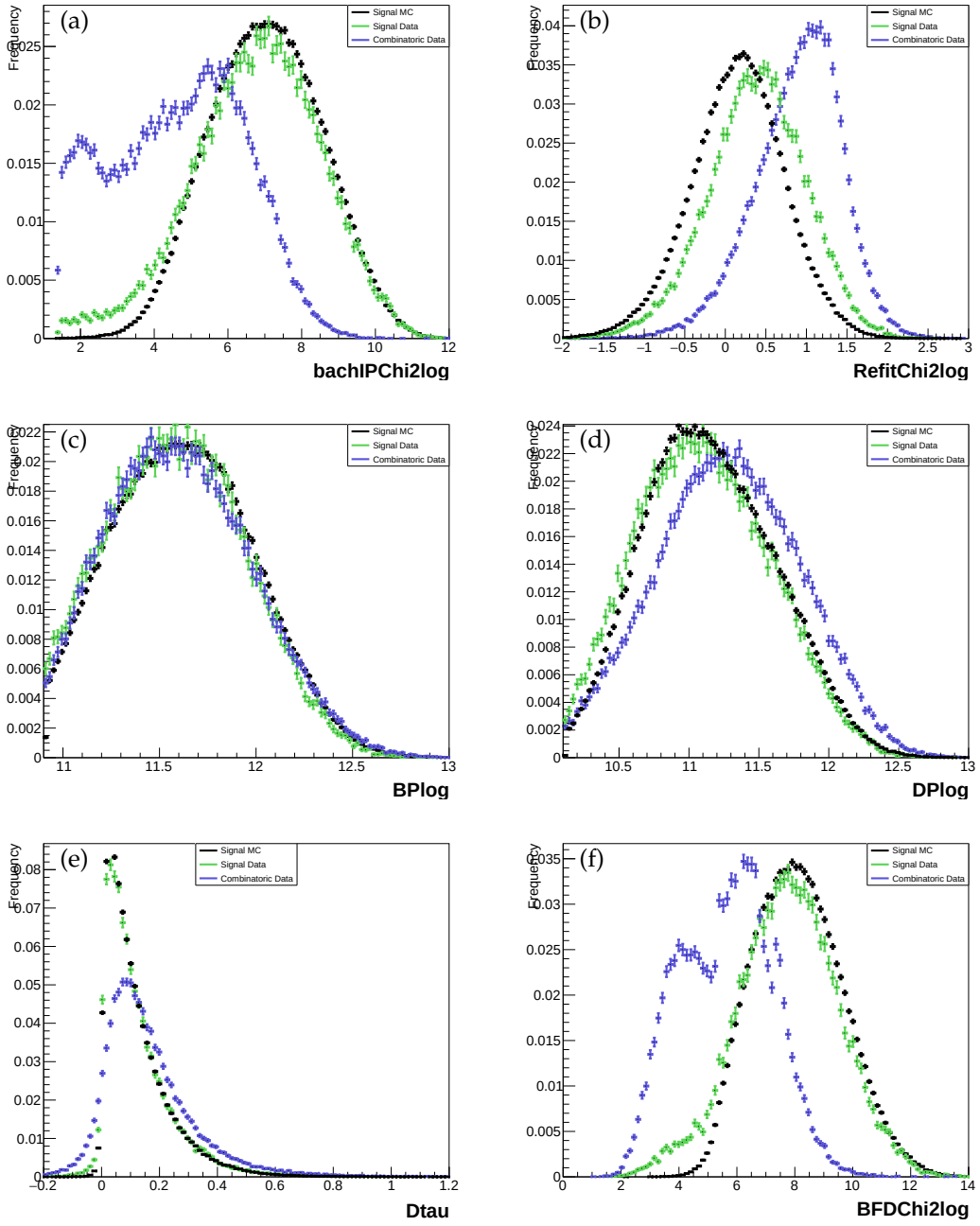


Figure 4.17.: Normalized distributions obtained for (a) $\ln((\chi^2)^{\mathrm{IP,PV}}_{\mathrm{bach}})$, (b) $\ln((\chi^2/\mathrm{ndof})_{\texttt{PVrefit}})$, (c) $\ln(p_B)$, (d) $\ln(p_D)$, (e) τ_D, and (f) $\ln((\chi^2)^{\mathrm{FD,ownPV}}_{B})$, obtained in signal MC (black), signal data (green) and combinatoric background (blue).

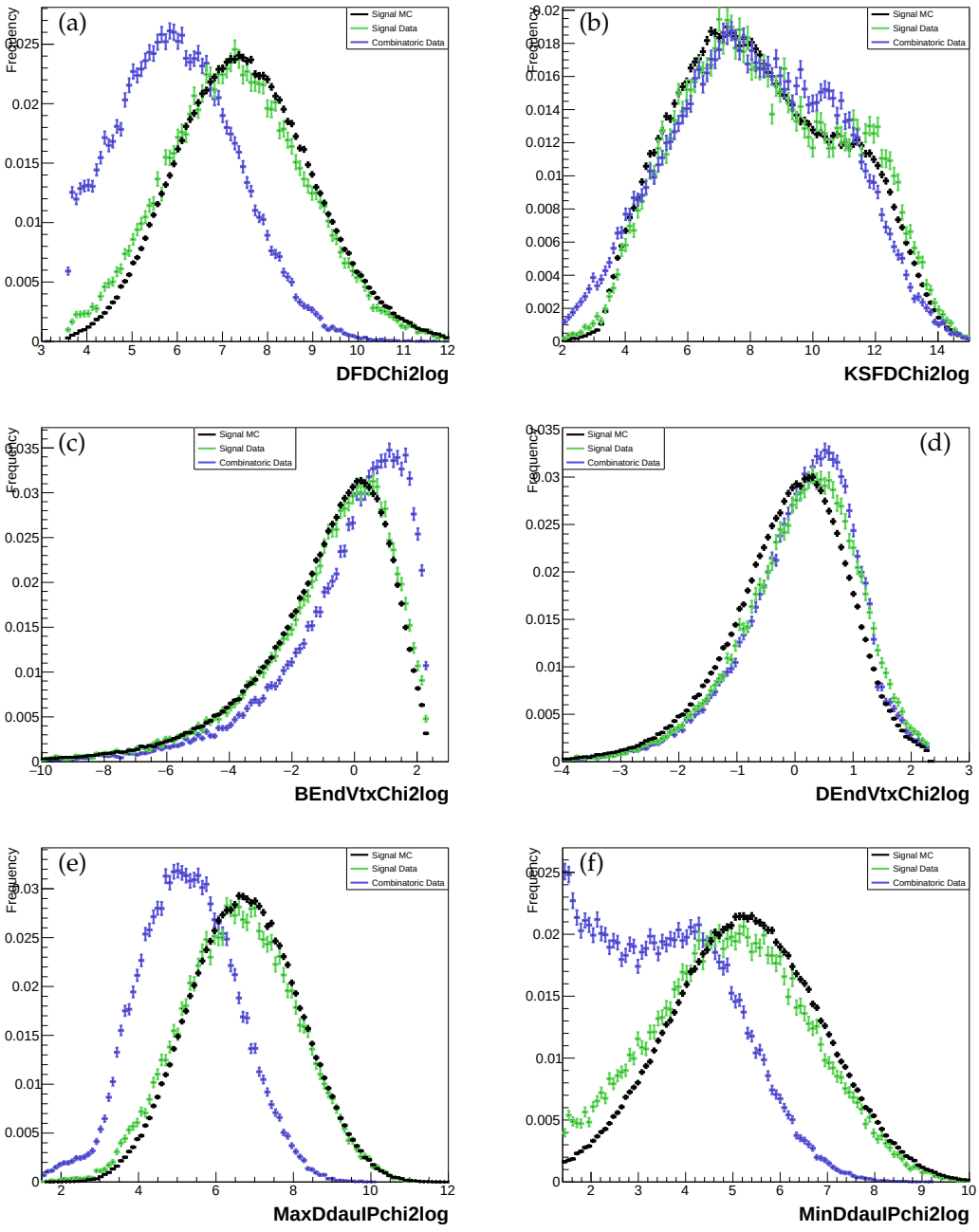


Figure 4.18.: Normalized distributions obtained for (a) $\ln((\chi^2)_D^{\mathrm{FD,ownPV}})$, (b) $\ln((\chi^2)_{K_S^0}^{\mathrm{FD,ownPV}})$, (c) $\ln((\chi^2/\mathrm{ndof})_B^{\mathrm{endvtx}})$, (d) $\ln((\chi^2/\mathrm{ndof})_D^{\mathrm{endvtx}})$, (e) $\ln(\max((\chi^2)_{\pi_D^+}^{\mathrm{IP,PV}}, (\chi^2)_{\pi_D^-}^{\mathrm{IP,PV}}))$, and (f) $\ln(\min((\chi^2)_{\pi_D^+}^{\mathrm{IP,PV}}, (\chi^2)_{\pi_D^-}^{\mathrm{IP,PV}}))$, obtained in signal MC (black), signal data (green) and combinatoric background (blue).

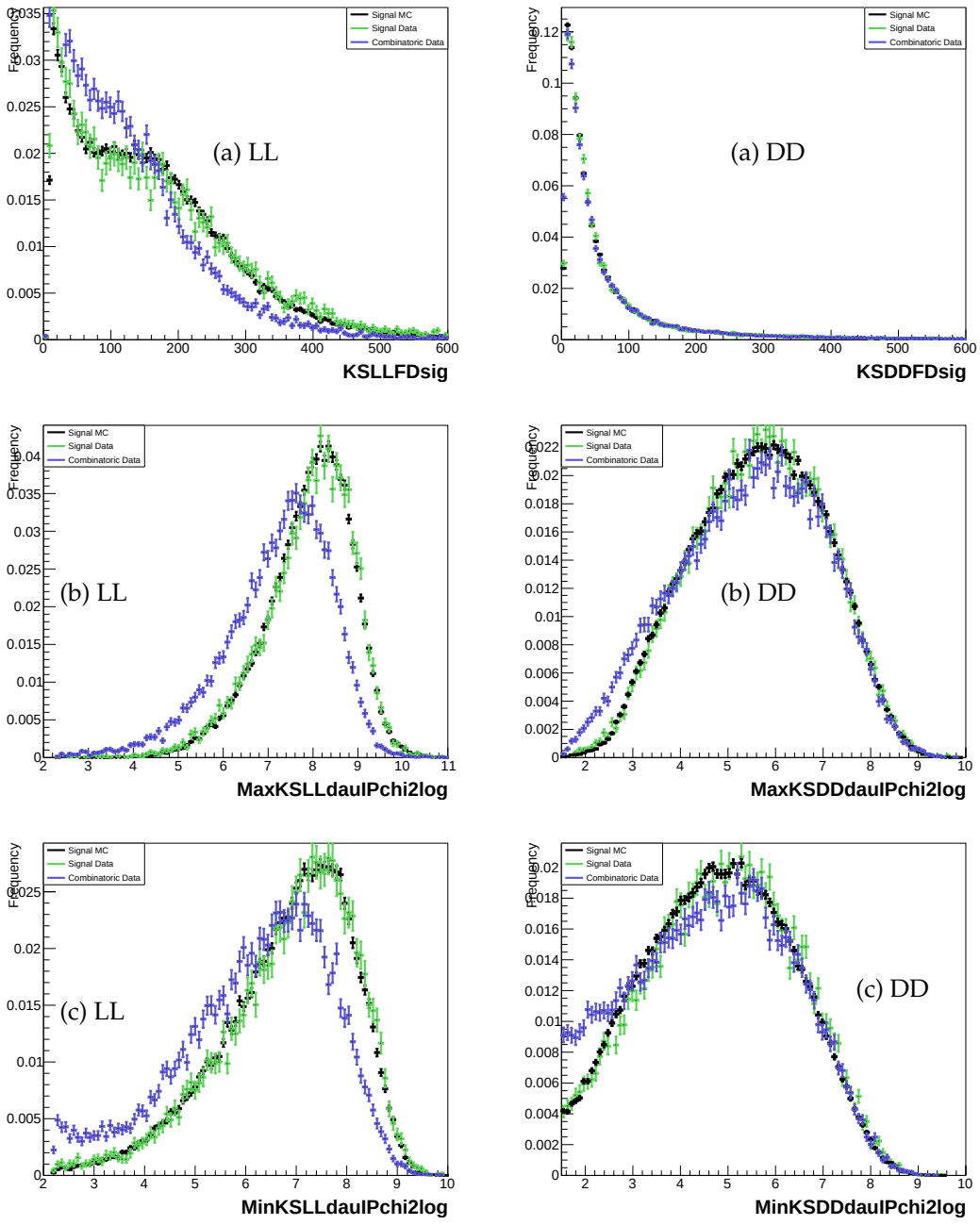


Figure 4.19.: Normalized distributions obtained for (a) $(z_{K_S^0}^{\text{endvtx}} - z_D^{\text{endvtx}})/\sqrt{(\Delta z_{K_S^0}^{\text{endvtx}})^2 + (\Delta z_D^{\text{endvtx}})^2}$, (b) ln($\max((\chi^2)^{\text{IP,PV}}_{\pi^+_{K_S^0}}, (\chi^2)^{\text{IP,PV}}_{\pi^-_{K_S^0}})$) and (c) ln($\min((\chi^2)^{\text{IP,PV}}_{\pi^+_{K_S^0}}, (\chi^2)^{\text{IP,PV}}_{\pi^-_{K_S^0}})$), obtained in signal MC (black), signal data (green) and combinatoric background (blue). Samples are split according to whether the K_S^0-meson candidate is of type LL (left) or DD (right).

Training results

A range of decision-tree and neural-network classifiers are trained and evaluated. The first check measures the performance of each classifier on its own training samples, by counting the number of signal and background events retained when a cut is placed on the classifier output. By scanning the cut value, a 'receiver operating characteristic' (ROC) curve may be constructed, plotting the fraction of signal events retained against the fraction of background events rejected. The ROC curves for the each MVA trained are shown in Figure 4.20. A perfect classifier would give a ROC curve which touches the point $(1,1)$, keeping all signal events whilst rejecting all background. The performance of most classifiers are very similar for each mode, whilst a few appear to give poor performance due to the relatively low training statistics available. To further compare the tree- and network-based classifiers, both AdaBDT and MLPBPCE classifiers (which each give good performance across all modes in Figure 4.20) are carried forward to the next check.

The next check performed involves comparing the normalized distributions of classifier output for both the data used to train the classifier, and the reserved validation set. Histograms of four distributions are plotted together: the classifier output for the signal training sample; background training sample; signal validation sample; and background validation sample. Plots of these distributions are shown for AdaBDT and MLPBPCE classifiers trained for each decay mode, in Figures 4.21, 4.22, 4.23 and 4.24. This check performs two fuctions. Firstly, it allows the separation between signal and background to be inspected. A perfect classifier would map signal and background events to completely separate regions of output space, so that a single cut on the output allows all background to be rejected. The smaller the overlap between the signal and background distributions, the greater the classifier's discriminatory power. Secondly, these plots give an indication of any potential 'overtraining', which occurs when a classifier gives better separation when acting on its training data than when acting on the validation data. This can indicate that the classifer has learned to rely on features that are only present within its training data, but not within the population more generally. This can occur when a classifier is trained using a very small training sample, since then the degrees of freedom in even a simple classifier may be sufficient for the optimisation algorithm to explore fine structure within parameter space, which may not generalise between samples.

The output of the MLPBPCE classifiers is observed to be slightly better separated than that of the AdaBDT classifiers and so these are chosen for the final selection.

For the two modes $B^{\pm} \to D^{*}h^{\pm}$ (Figure 4.23) and $B^{\pm} \to DK^{*\pm}$ (Figure 4.24), there is some difference between the histograms obtained for the training and validation samples. This may indicate slight overtraining of the classifiers, but in any case the trained classifiers are still shown to provide good discrimination between the signal and background validation samples. No overtraining is observed for decay modes $B^{\pm} \to Dh^{\pm}$ and $B^{0} \to DK^{*0}$, where more training and validation data is available.

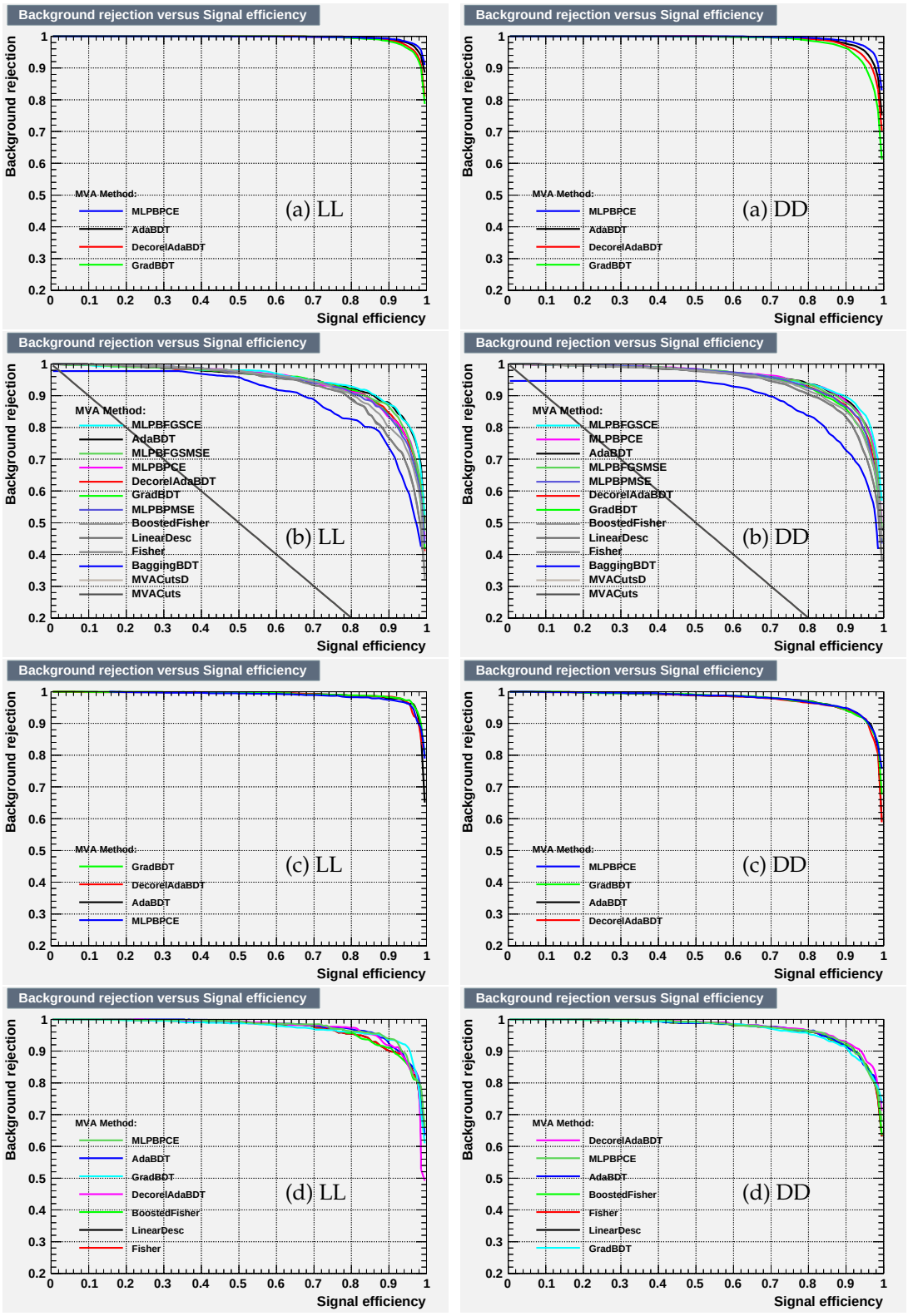


Figure 4.20.: ROC curves for MVAs trained to select LL and DD (a)$B^{\pm} \to Dh^{\pm}$, (b)$B^0 \to DK^{*0}$, (c)$B^{\pm} \to D^*h^{\pm}$, and (d)$B^{\pm} \to DK^{*\pm}$ candidates. MVAs trained to select LL K^0_S-meson candidates are shown on the left-hand side, whilst MVAs trained to select DD K^0_S-meson candidates are shown on the right-hand side.

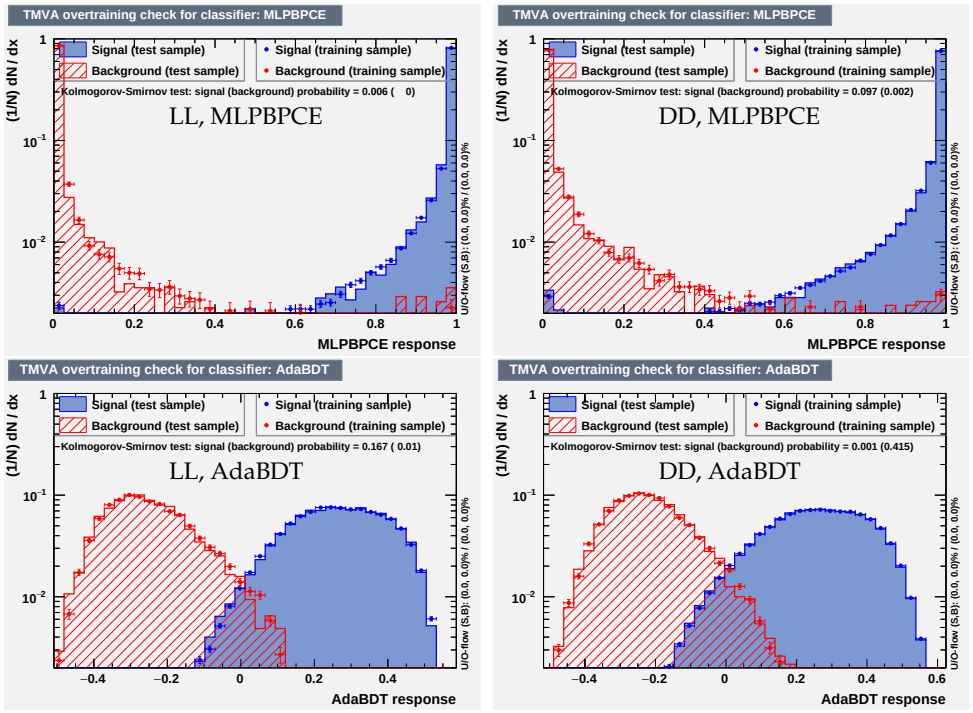


Figure 4.21.: Performance of classifiers trained to select $B^{\pm} \to Dh^{\pm}$ candidates.

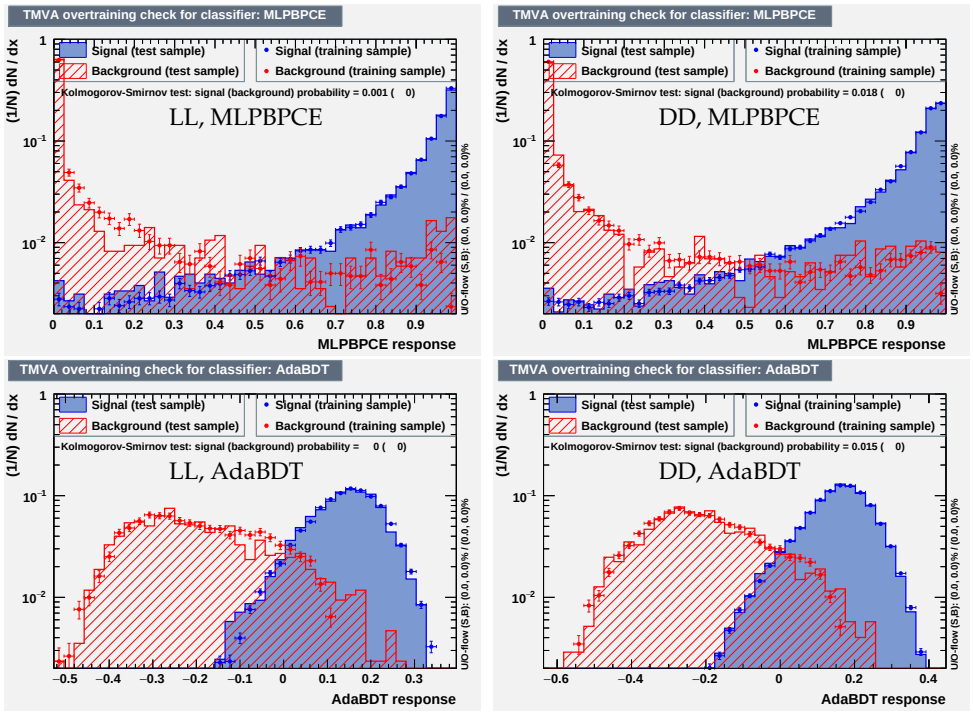


Figure 4.22.: Performance of classifiers trained to select $B^0 \to DK^{*0}$ candidates.

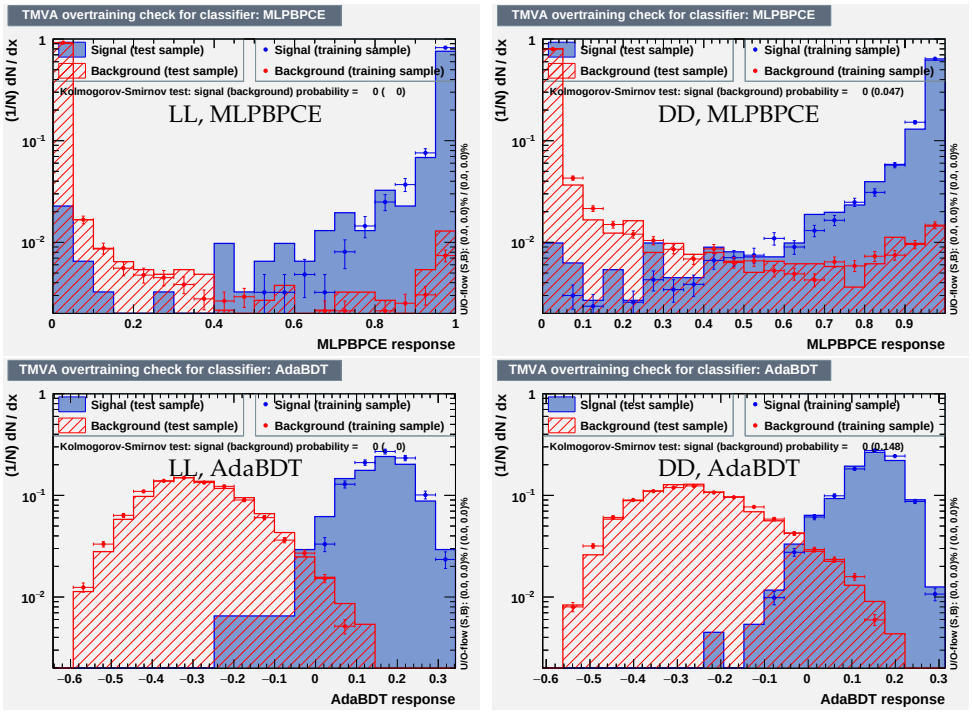


Figure 4.23.: Performance of classifiers trained to select $B^{\pm} \to D^{*}h^{\pm}$ candidates.

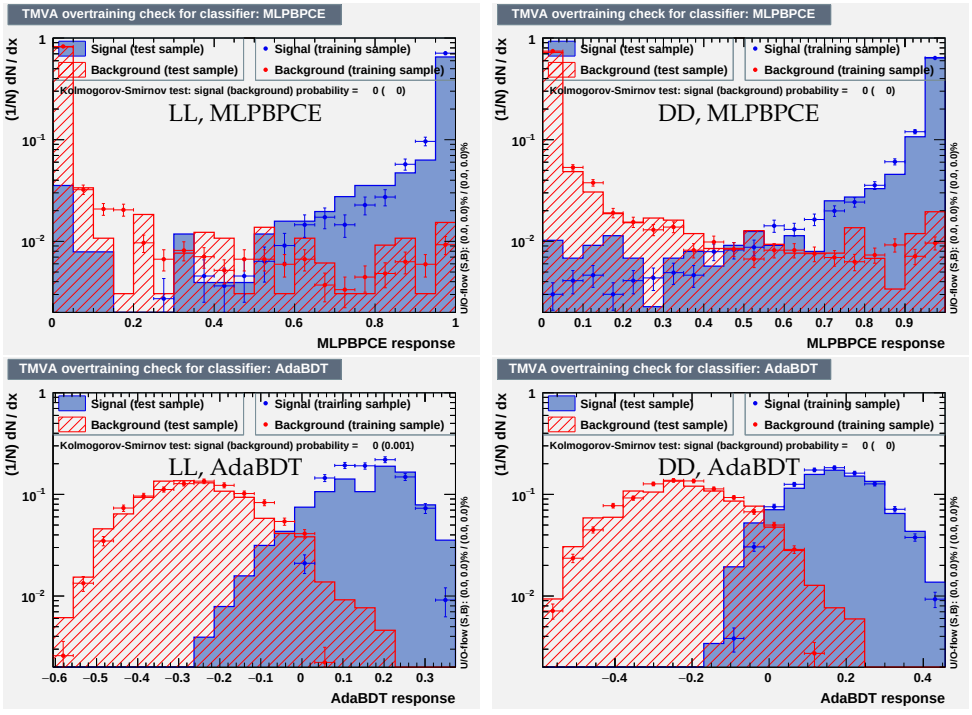


Figure 4.24.: Performance of classifiers trained to select $B^{\pm} \to DK^{*\pm}$ candidates.

4.5. MVA and PID cut optimisation

Having developed an MVA classifier for each mode, a cut of the form $(\text{MVA} > X_j)$ may be used to select candidate signal events for decay mode j. However, it remains to choose the most appropriate cut value, X_j, to be used for each decay mode.

In addition, a final set of PID requirements are introduced at this stage, placed on the bachelor candidate of decay modes $B^0 \to DK^{*0}$ and $B^\pm \to DK^{*\pm}$. For the decay mode $B^0 \to DK^{*0}$, a requirement is placed on the PID of the charged kaon originating from the K^{*0} decay of the form $(\text{DLL}_{K\pi}(K^\pm \text{ from } K^{*0}) > Y)$. For the decay mode $B^\pm \to DK^{*\pm}$, a requirement is placed on the PID of the charged pion originating from the $K^{*\pm}$ decay of the form $(\text{DLL}_{K\pi}(\pi^\pm \text{ from } K^{*\pm}) < Z)$. These variables cannot be included as MVA inputs, since they are not well-modelled in the training sample. Optimisation of these PID requirements is important, since they are crucial in reducing peaking background contributions.

The choice of cut value is made by defining a suitable figure of merit ('FOM') to maximise. Common FOMs include:

- purity $= \frac{S}{S+B}$;
- significance $= \frac{S}{\sqrt{S+B}}$; and
- hybrid $=$ purity $\cdot$ significance $= \frac{S^2}{(S+B)^{3/2}}$.

In this analysis, the hybrid FOM is used. It balances the need for high yield and high purity, and in a previous analysis it was found that this FOM was well-correlated with sensitivity to measured γ [128]. The MVA and PID cuts of decay modes $B^0 \to DK^{*0}$ and $B^\pm \to DK^{*\pm}$ are optimised simultaneously, since there is no guarantee that optimising each separately will yield the global maximum FOM.

This optimisation is performed on experimental data since, as mentioned above, low signal and background statistics mean that all available candidates are already used for training and validation of the MVA classifiers, and so cannot be used again here. In principle, it is possible that this approach may introduce bias, since the selection is optimised on precisely the same data it will be used to select. However, since the present analysis consists in extracting phase information within the Dalitz plane, rather than simply counting events, it is not easily susceptible to experimenter bias[6]. To guard against any possibility of such bias, no Dalitz fits are performed until the selection has been frozen.

The final optimisation is performed only after mass fits have been defined for each decay mode (described in Chapter 5), allowing accurate estimates of the signal and background populations to be produced for each trial cut value. The optimisation then proceeds as follows:

- apply initial trial MVA requirements, roughly tuned using simplified fits of control mode decays;

[6] For example, choosing cut values that give spuriously high signal purity in the mass fit, by increasing the mis-labelling of background candidates as signal, would not necessarily increase the apparent sensitivity to γ extracted in the Dalitz fit, since this depends on each selected candidate being modelled correctly in the Dalitz plane.

- develop full mass fits at this initial working point;
- use full fits of signal modes to optimise cut choice:
  - for each decay mode, scan a broad range of possible cut values. For decay modes $B^0 \to DK^{*0}$ and $B^\pm \to DK^{*\pm}$, a two-dimensional cut space is scanned, since the PID and MVA cuts are tuned simultaneously;
  - for each cut value, perform the full mass fit;
  - determine yields in the signal region and calculate the figures of merit; and
- choose cut values that maximise the hybrid FOM.

Figure 4.25 shows the corresponding FOM scans for $B^\pm \to DK^\pm$ decays with LL K^0_S candidates. A cut at MVA$>$ 0.92 is chosen as optimal in this case. Figure 4.26 shows the FOM scans for $B^\pm \to DK^\pm$ decays with DD K^0_S candidates. A cut at MVA$>$ 0.68 is chosen as optimal. The FOM scans for $B^\pm \to D^*K^\pm$ decays with LL K^0_S candidates are shown in Figure 4.27. A cut at MVA$>$ 0.96 is chosen as optimal. Figure 4.28 shows the FOM scans for $B^\pm \to D^*K^\pm$ decays with DD K^0_S candidates. A cut at MVA$>$ 0.72 is chosen as optimal. Figure 4.29 shows the FOM scans for $B^0 \to DK^{*0}$ decays with LL K^0_S candidates. Cut values of MVA$>$ 0.722 and $\mathrm{DLL}_{K\pi}(K^\pm \text{ from } K^{*0}) > 5.58$ are chosen as optimal in this case. Figure 4.30 shows the FOM scans for $B^0 \to DK^{*0}$ decays with DD K^0_S candidates. Cut values of MVA$>$ 0.781 and $\mathrm{DLL}_{K\pi}(K^\pm \text{ from } K^{*0}) > 4.68$ are chosen. The FOM scans for $B^\pm \to DK^{*\pm}$ decays with LL K^0_S candidates are shown in Figure 4.31. Cut values of MVA$>$ 0.871 and $\mathrm{DLL}_{K\pi}(\pi^\pm \text{ from } K^{*\pm}) < -2.08$ are chosen as optimal in this case. The FOM scans for $B^\pm \to DK^{*\pm}$ decays with DD K^0_S candidates are shown in Figure 4.32. Cut values of MVA$>$ 0.704 and $\mathrm{DLL}_{K\pi}(\pi^\pm \text{ from } K^{*\pm}) < 2.00$ are chosen as optimal. Table 4.7 summarises the optimised cut values obtained for the different decay modes.

Mode	Optimised MVA requirement	Optimised PID requirement
$B^\pm \to Dh^\pm$ LL	> 0.92	-
$B^\pm \to Dh^\pm$ DD	> 0.68	-
$B^\pm \to D^*h^\pm$ LL	> 0.96	-
$B^\pm \to D^*h^\pm$ DD	> 0.72	-
$B^0 \to DK^{*0}$ LL	> 0.722	$\mathrm{DLL}_{K\pi}(K^\pm \text{ from } K^{*0}) > 5.58$
$B^0 \to DK^{*0}$ DD	> 0.781	$\mathrm{DLL}_{K\pi}(K^\pm \text{ from } K^{*0}) > 4.68$
$B^\pm \to DK^{*\pm}$ LL	> 0.871	$\mathrm{DLL}_{K\pi}(\pi^\pm \text{ from } K^{*\pm}) < -2.08$
$B^\pm \to DK^{*\pm}$ DD	> 0.704	$\mathrm{DLL}_{K\pi}(\pi^\pm \text{ from } K^{*\pm}) < 2.00$

Table 4.7.: Optimised selection requirements applied to candidates of each mode.

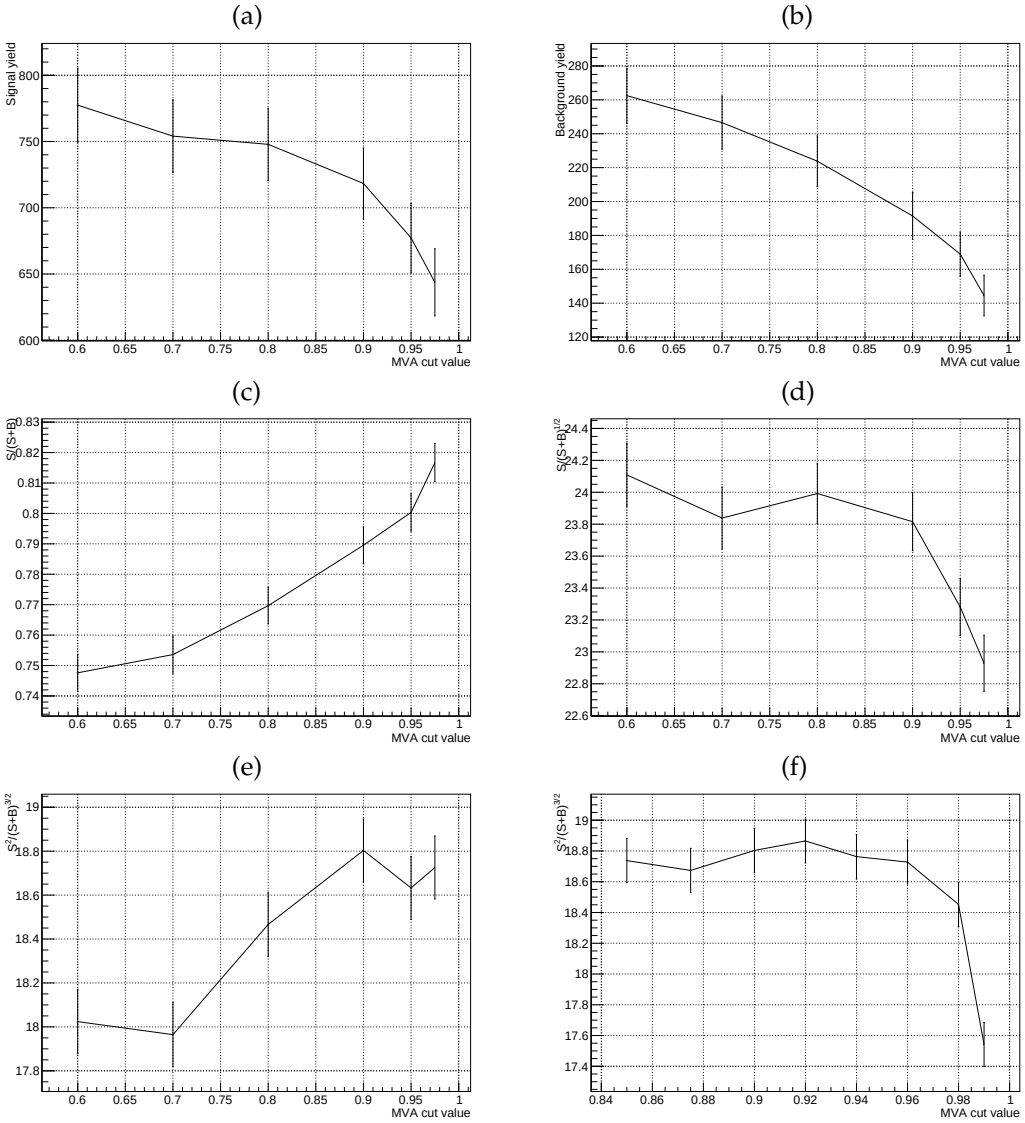


Figure 4.25.: Figures of merit extracted from fits to the B-mass spectrum of selected LL $B^{\pm} \to DK^{\pm}$ candidates, for a range of MVA cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f).

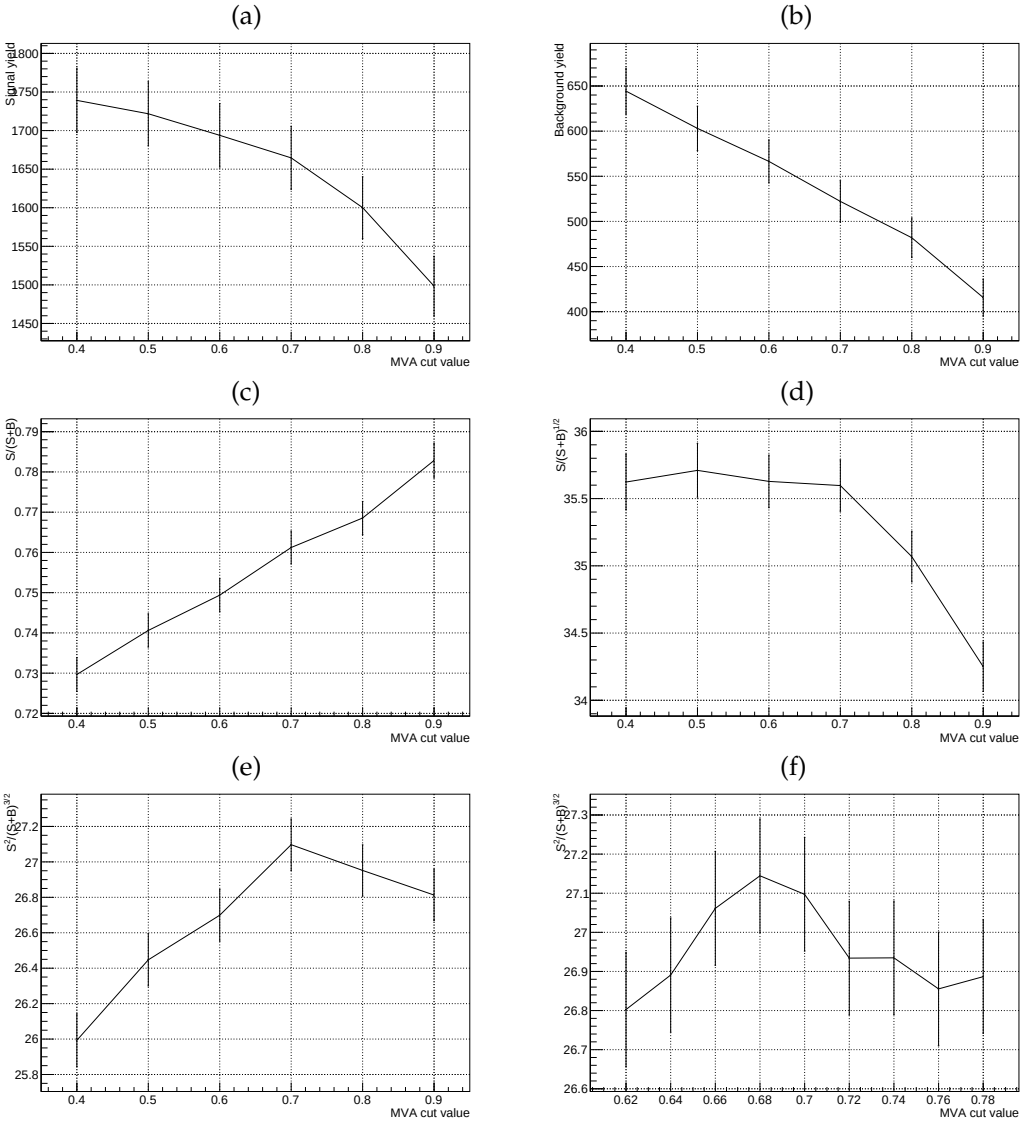


Figure 4.26.: Figures of merit extracted from fits to the B-mass spectrum of selected DD $B^{\pm} \to DK^{\pm}$ candidates, for a range of MVA cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f).

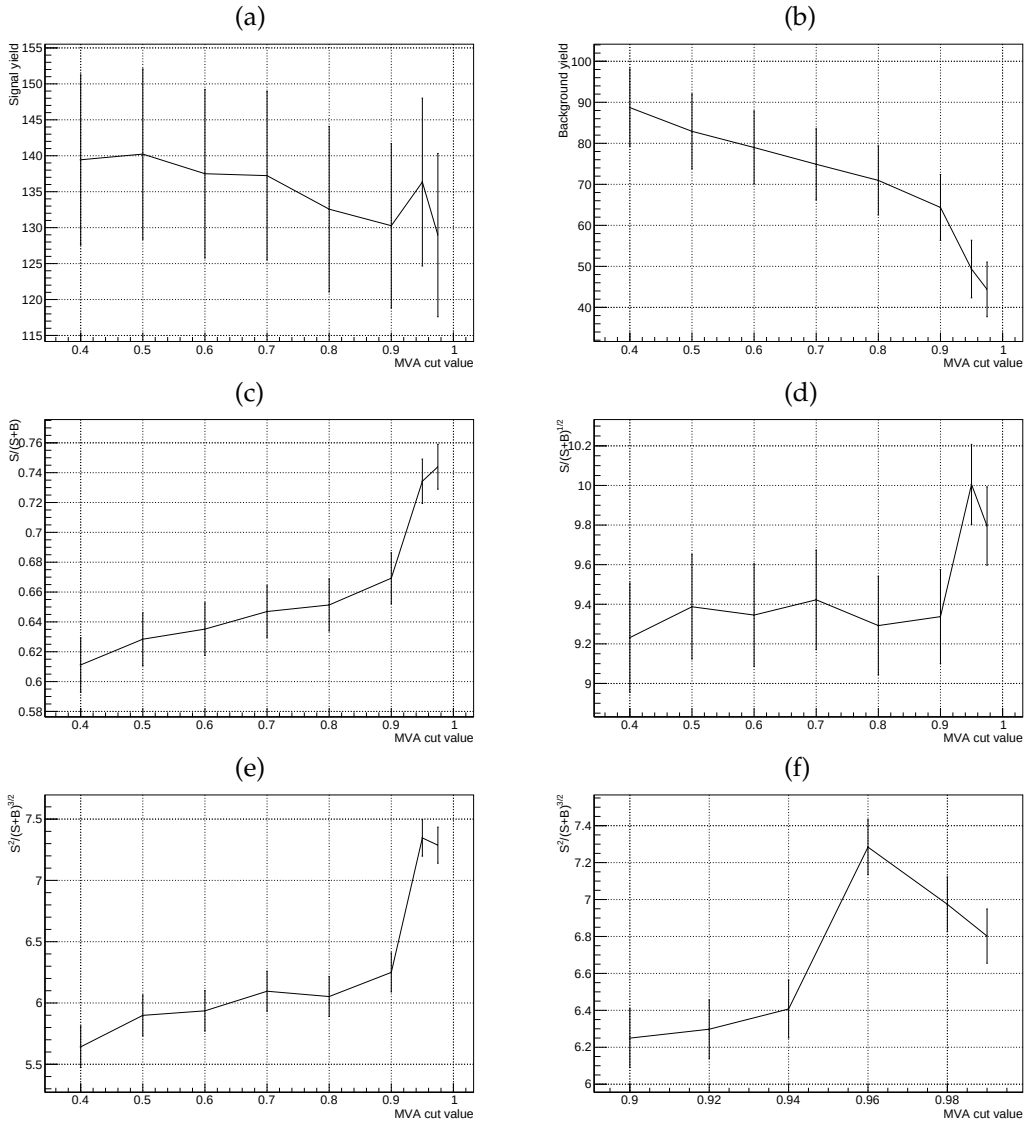


Figure 4.27.: Figures of merit extracted from fits to the B-mass spectrum of selected LL $B^{\pm} \to D^{*}K^{\pm}$ candidates, for a range of MVA cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f). The increase in signal yield (a) for high MVA cut values reflects the fact that these yields are extracted from the full mass fit, and statistical fluctuations in the fit result are possible as events are removed from the fitted sample.

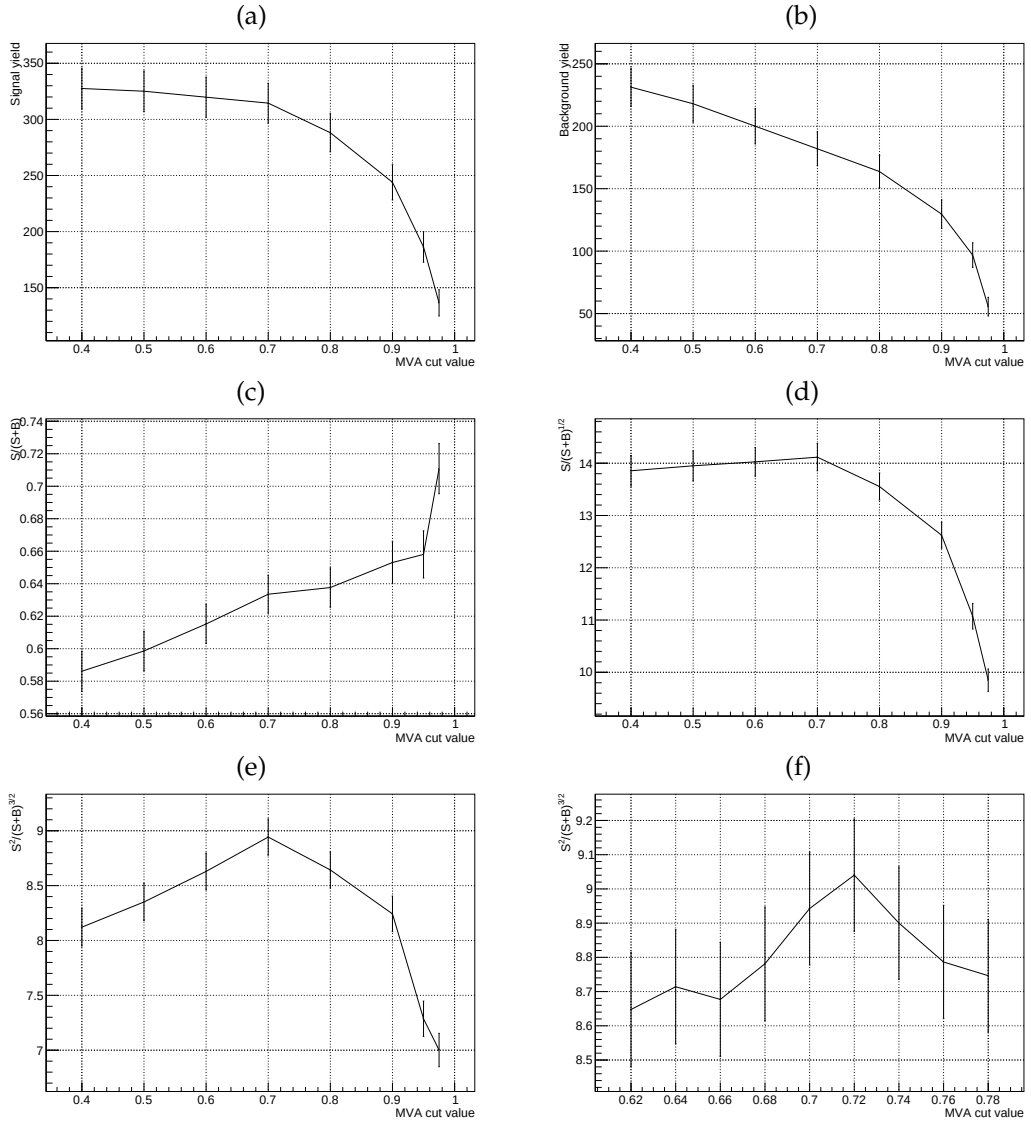


Figure 4.28.: Figures of merit extracted from fits to the B-mass spectrum of selected DD $B^{\pm} \to D^{*}K^{\pm}$ candidates, for a range of MVA cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f).

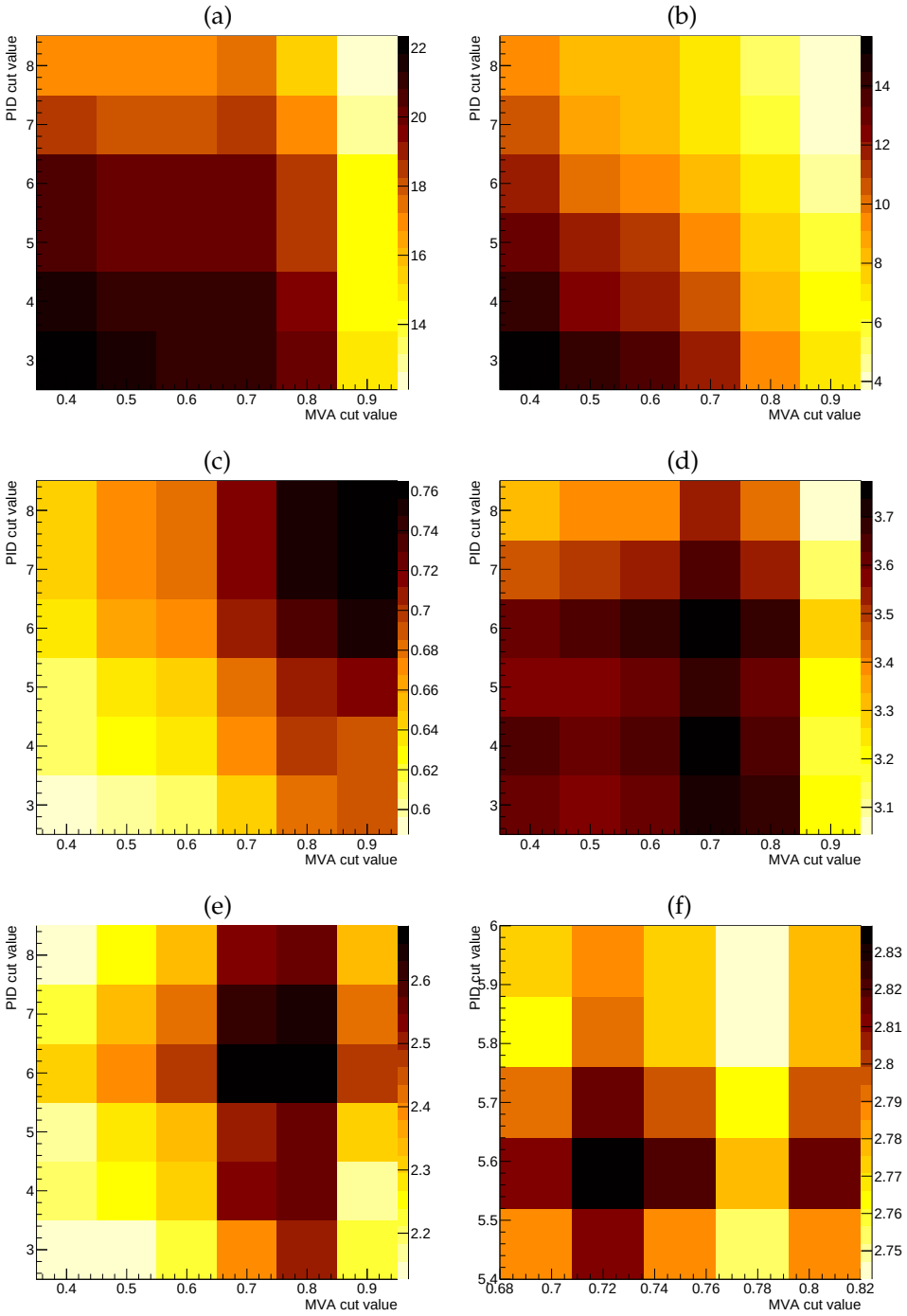


Figure 4.29.: Figures of merit extracted from fits to the B-mass spectrum of selected LL $B^0 \to DK^{*0}$ candidates, for a range of MVA and PID cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f).

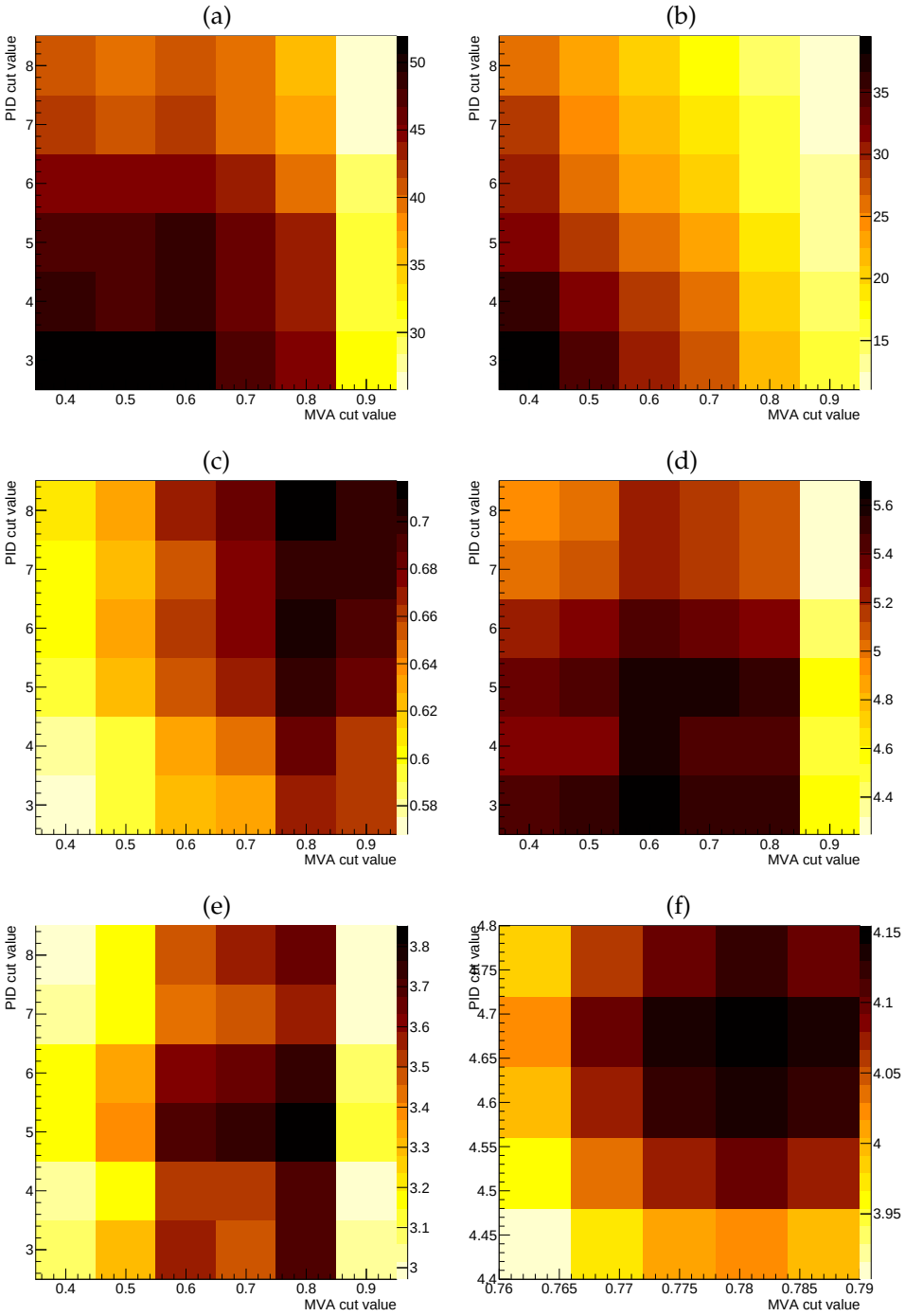


Figure 4.30.: Figures of merit extracted from fits to the B-mass spectrum of selected DD $B^0 \to DK^{*0}$ candidates, for a range of MVA and PID cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f).

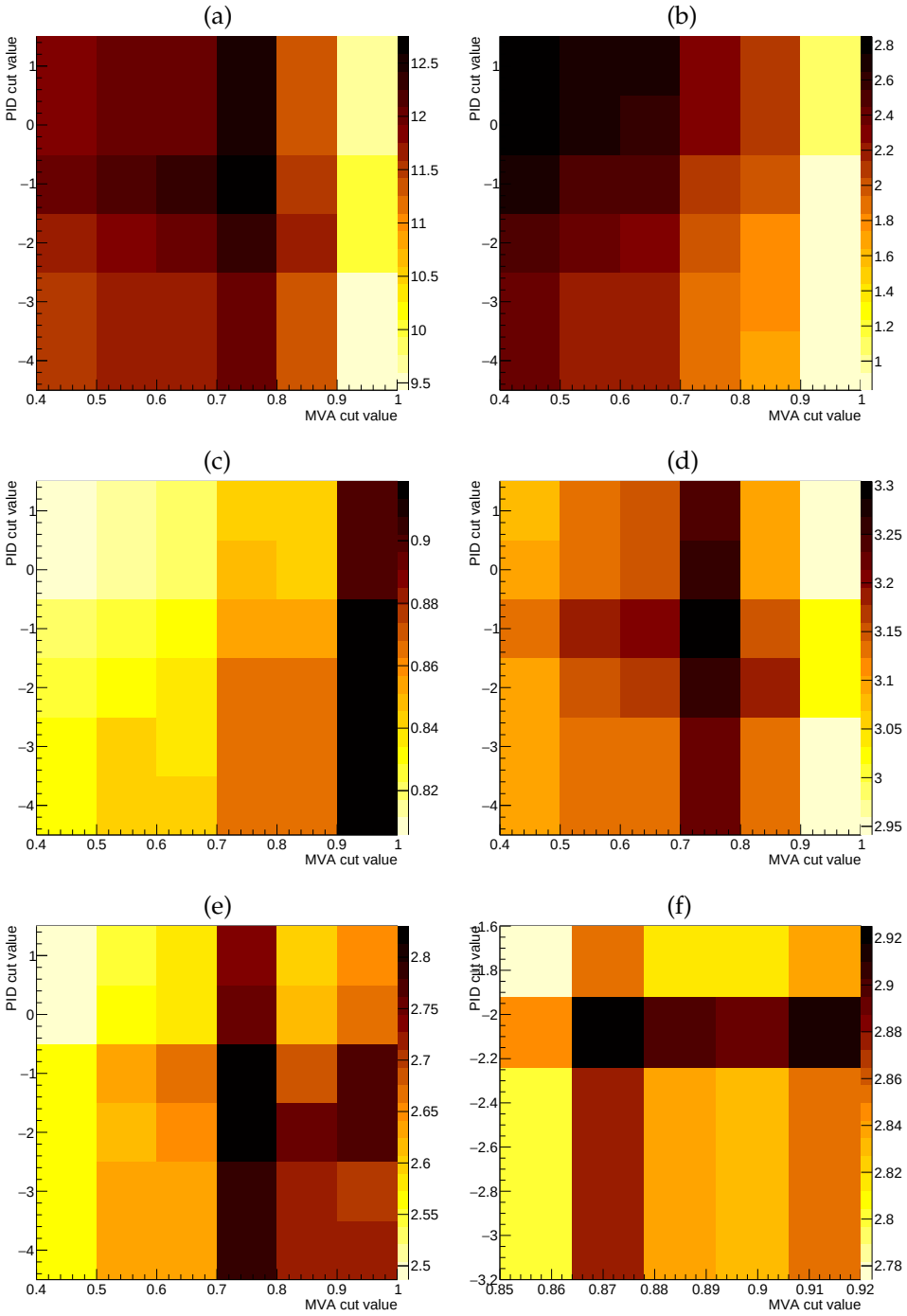


Figure 4.31.: Figures of merit extracted from fits to the B-mass spectrum of selected LL $B^{\pm} \to DK^{*\pm}$ candidates, for a range of MVA and PID cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f).

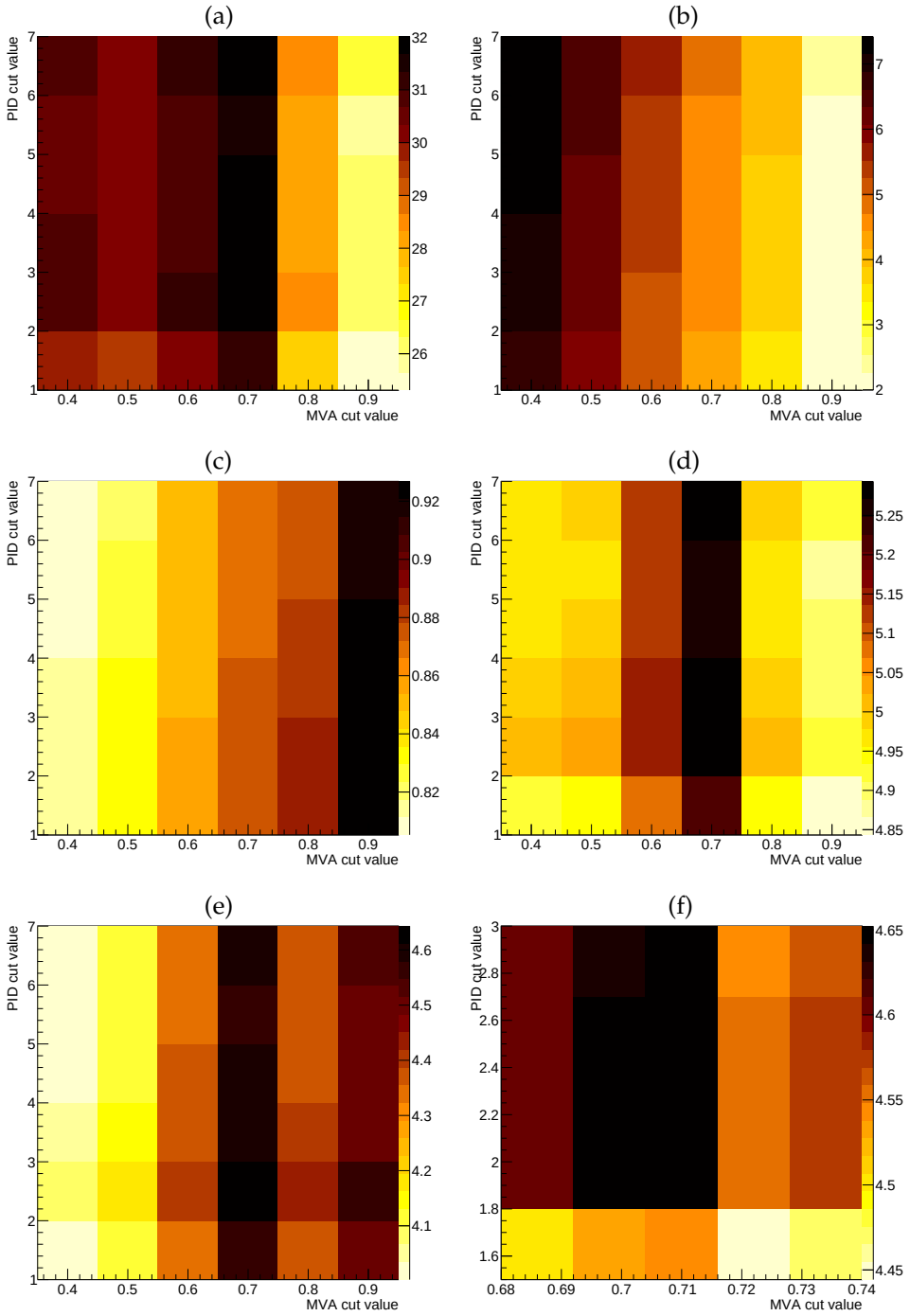


Figure 4.32.: Figures of merit extracted from fits to the B-mass spectrum of selected DD $B^{\pm} \to DK^{*\pm}$ candidates, for a range of MVA and PID cut values. Signal (a) and background (b) yields extracted within the signal window are plotted, along with the resulting purity (c), significance (d) and 'hybrid' (e) figure of merit, for a coarse sampling of cut values. Finally, a fine-grained scan is used to find the cut value which maximises the hybrid figure of merit (f).

4.6. Selection summary

In summary, the offline selection requirements placed on candidates of each mode are given in Tables 4.2, 4.4 and 4.7. MVA classifiers are trained, and the requirements placed on their output to select the final candidate samples are optimised using full fits to the reconstructed B-meson mass spectra of selected data.

5. Fit to reconstructed B-meson mass spectrum

After the application of the selection requirements described in Chapter 4, candidates remain that do not correspond to signal decay modes - non-negligible contributions are present for several background components. The yields of the signal and background components are determined using unbinned extended maximum likelihood fits (described below) to the reconstructed B-mass spectrum.

The spectrum fitted is that of the reconstructed B-meson mass. Where possible, analytic probability density functions (PDFs) are used to describe the expected distributions of the fitted components, with some shape parameters fixed to values extracted from fits to MC. Where this is unfeasible, non-analytic PDFs are extracted from MC directly, using the kernel estimation functionality provided within the `RooFit` software library [138]. In each case, simulated data provide an important input into the fit. This simulation is performed centrally within the LHCb collaboration, following the procedure outlined in Section 3.2.8. Before each simulated sample is used to extract a PDF, its full decay chain is first checked using truth-level information to match mother and daughter candidates at each decay vertex, allowing incorrectly reconstructed candidates to be rejected, before the full selection is finally applied.

Independent mass fits are performed for each signal mode. These are described sequentially in the following sections, following a brief interlude to describe the fitting process used.

Likelihood fitting

Likelihood fitting is a technique commonplace in statistics, often used when it is necessary to determine the unknown parameters of some theoretical model, by making a comparison[1] with experimental data.

The theoretical model takes the form of a probability density function, F, expressing the predicted conditional probability for the values of a set of observables, $\vec{x}$, given the values of a set of model parameters, $\vec{q}$;

$$F(\vec{x}\,|\,\vec{q}). \tag{5.1}$$

In order to make a comparison of the model expectation with some observed data point, $\vec{X} \in \vec{x}$, it is useful to consider the value of

$$F(\vec{X}\,|\,\vec{q}), \tag{5.2}$$

where greater values clearly indicate better agreement between the two, since the observables take values that are predicted by the model to be more likely. For a set of independent measurements of the

[1] Or, perhaps more appropriately, a 'calibration'.

observables, $\{\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n\}$, the conditional probability is given by

$$F(\{\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n\} \,|\, \vec{q}) = \prod_{i=1}^{n} F(\vec{X}_i \,|\, \vec{q}). \tag{5.3}$$

The likelihood formalism reinterprets the model as a function of *both* the observable values and the model parameters, since the above PDF can be evaluated for different parameter values in just the same way as it can for different observable values. The likelihood function, for a given set of observations, is thus defined as

$$\mathcal{L}(\vec{q};\, \vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n) = F(\{\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n\} \,|\, \vec{q}). \tag{5.4}$$

In order to calibrate the model, the likelihood function is maximised with respect to the model parameter values $\vec{q}$. For some simple models this maximisation may proceed analytically, but in more complex applications, typical of high energy physics analyses, a numerical approach is used. This approach relies on sampling points within the $\vec{q}$ parameter space, and calculating the likelihood for each choice, to find that which gives the greatest value. For computational efficiency, it is usual to consider the logarithm of the likelihood (or 'log-likelihood'), since this decomposes simply into a sum of terms corresponding to the set of observations

$$\ln \mathcal{L}(\vec{q};\, \vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n) = \sum_{i=1}^{n} \ln F(\vec{X}_i \,|\, \vec{q}). \tag{5.5}$$

In practice, all of the models considered within this thesis are composites, with separate terms to describe each signal and background component expected in data, weighted by their relative populations

$$F(\vec{X} \,|\, N, \vec{\theta}, \vec{q}) = \sum_j \theta_j F_j(\vec{X} \,|\, \vec{q}), \tag{5.6}$$

$$\theta_j \equiv \frac{N_j}{N}, \tag{5.7}$$

where $\vec{X}$ denotes the set of observed data points, $\vec{X} = \cup_i \vec{X}$, N denotes the total data yield predicted by the model, N_j gives the data yield predicted for the j-th model component, and the weights $\vec{\theta}$ thus sum to unity. Within this composite model, N and $\vec{\theta}$ are free parameters, and can thus be treated in the same way as the other parameters, $\vec{q}$, in the maximisation. However, additional information is available regarding the true value of N, since an estimator for N is given directly by the number of data points observed, n. To take account of this information, the likelihood function is multiplied by the Poissonian probability to observe n candidates, given the model expectation of N, to give an 'extended likelihood function' [139]:

$$\mathcal{L}(\vec{q}, N, \vec{\theta};\, \vec{X}) = \frac{N^n}{n} e^{-N} F(\vec{X} \,|\, N, \vec{\theta}, \vec{q}) = \frac{e^{-N}}{n} \prod_{i=1}^{n} \sum_j N_j F_j(\vec{X}_i \,|\, \vec{q}). \tag{5.8}$$

All likelihood maximisations performed in this thesis make use of the MINUIT [140] package, which uses a range of algorithms to minimise the negative logarithm of the likelihood (NLL). The minimum of the NLL is found using the MIGRAD algorithm, which estimates the gradient of the NLL function (in parameter space $\vec{q}$) after each evaluation and steps iteratively in the direction of steepest descent, until a minimum is reached.

Once the minimum has been located, the uncertainty on the estimation of each parameter must be determined. Keeping all other parameters fixed to their chosen values, a single parameter is varied at a time, to find an interval containing values within some chosen variation from the minimum NLL. The allowed variation is chosen such that the combined interval, including all parameters, corresponds to a maximum of one standard deviation from the minimum. Within MINUIT, symmetric likelihood intervals (*i.e.* intervals centred on the chosen values) are calculated by the HESSE module, under the assumption that the NLL function is well-approximated by a quadratic over the range of the interval considered.[2] Asymmetric likelihood intervals are calculated for the final fits, using the more sophisticated MINOS module. Uncertainties arising in this way are purely statistical, and are thus expected to reduce as more data are included.

Pulls

In order to visually assess the quality of a fit to a histogram, it is typical to include a plot to show the 'pull' p_k calculated for each bin, k. This is defined as the difference between the bin population observed, n_k^{obs}, and the bin population expected according to the fitted model, n_k^{expected}, in units of the standard deviation on the population that would be expected, σ_k,

$$p_k \equiv \frac{n_k^{\text{obs}} - n_k^{\text{expected}}}{\sigma_k}. \tag{5.9}$$

The event count in each bin is a discrete variable, and so the uncertainty σ_k is given by the Poissonian uncertainty on the bin population, $\sqrt{n_k^{\text{obs}}}$.

If the difference between observed and expected distributions are due purely to statistical fluctuations in the data, then the pulls should be Normally-distributed about zero, with unit variance. Any noticeable structure within the pulls therefore indicates imperfections in the model chosen to fit the data, and highlights the specific regions that are poorly modelled. Throughout this chapter, all fitted mass distributions are shown with their corresponding pull distribution directly below, for ease of comparison.

[2] As it will be for sufficiently narrow intervals around the minimum; where a Taylor expansion of the NLL is dominated by the quadratic term.

5.1. Fit to selected $B^\pm \to Dh^\pm$ candidates

Two sets of simultaneous fits are performed to the reconstructed B-meson mass of selected $B^\pm \to Dh^\pm$ candidates. Separate fits are performed depending on the track type (tt) of the reconstructed K^0_S candidate (LL or DD). For each fit, two data samples are fitted simultaneously; one containing $B^\pm \to DK^\pm$ candidates, selected by the requirement $\mathrm{DLL}_{K\pi}(\mathrm{bach}) > 4$, and one containing $B^\pm \to D\pi^\pm$ candidates, selected by the requirement $\mathrm{DLL}_{K\pi}(\mathrm{bach}) < 4$. For convenience, the shorthand $DK^\pm$ and $D\pi^\pm$ are adopted to describe these two samples.

For the fit to $B^\pm \to D\pi^\pm$ candidates, shapes are used to describe the contributions from:

- Signal candidates;
- 'Cross-feed' from $B^\pm \to DK^\pm$ candidates;
- 'Low-mass' background from partially-reconstructed decays:
  - $B \to D^*\pi$, where the B- and π-mesons may be charged or neutral;
  - $B \to D^*\rho$, where the B- and ρ-mesons may be charged or neutral;
  - $B \to D\rho$, where the B- and ρ-mesons may be charged or neutral; and
- Background from combinatoric candidates.

The fit to $B^\pm \to DK^\pm$ candidates includes shapes to describe all of the above contributions, plus additional 'low-mass' backgrounds:

- $B \to D^*K$, where the B- and K-mesons may be charged or neutral;
- $B^0 \to DK^{*0}$; and
- $B^0_s \to DK^{*0}$.

The PDF shapes used to describe each of these components are taken directly from a previous analysis [141][3], which was performed using the 2011 LHCb dataset only. The shapes are described here to allow a meaningful interpretation of the fit results obtained.

The shapes used to describe all components except combinatoric background were determined from fits to simulated samples in reference [141]. Each sample was reconstructed twice, once under the $B^\pm \to D\pi^\pm$ hypothesis, and once under the $B^\pm \to DK^\pm$ hypothesis, and separate shapes were extracted in each case, to be used in the separate categories of the simultaneous fit.

Most of the low-mass, partially-reconstructed background components include contributions from more than one mode, each expected to have a similar distribution in the reconstructed B-mass variable. For example, the PDF labelled '$B \to D^*\pi$' accounts for both $B^0 \to [D^{*\mp} \to D\pi^\mp]\pi^\pm$ and $B^\pm \to [D^{*0} \to D\gamma, \pi^0]\pi^\pm$ candidates. In such cases, reference [141] combined simulated samples

[3] This analysis resulted in the publication [142], but full details of the mass fit procedure are best found in the note [141].

corresponding to the different decays, according to their relative branching fractions and selection efficiencies obtained in MC, before the combined sample was fitted to calibrate the PDF parameters.

The choice to use the PDF shapes of the 2011 analysis [141] without recalibration is made to save time, since the present analysis has three other modes to consider, and carries more value as a proof of concept for the simultaneous approach than as an update of the 2011 analysis of this mode. This choice may lead to reduced mass fit quality, and thus poorer estimates of component yields than would be possible if each shape were calibrated to MC samples generated in 2012 conditions, on which the current selection had been applied. However, it will be seen that many of these shapes are negligible within the signal region, and so the impact of this choice on sensitivity to γ is expected to be slight.

The shapes used for each component are described below, summarized in Table 5.1, and are plotted in Figures 5.1–5.4.

$Dh^{\pm}$ signal

In each fit category, $D\pi^{\pm}$ and $DK^{\pm}$, signal candidates are described using the sum of a Gaussian with a Crystal Ball function [143], which is defined as

$$\mathcal{CB}(m;\mu,\sigma,\alpha,n) \equiv \begin{cases} \frac{(\frac{n}{|\alpha|})^n \cdot \exp(-\frac{1}{2}|\alpha|^2)}{(\frac{n}{|\alpha|}-|\alpha|-\frac{m-\mu}{\sigma})^n}, & \frac{m-\mu}{\sigma} < \alpha, \\ \exp\big(-\frac{1}{2}(\frac{m-\mu}{\sigma})^2\big), & \frac{m-\mu}{\sigma} \geq \alpha, \end{cases} \tag{5.10}$$

and consists of a Gaussian core with an added exponential tail on the low-mass side. Parameters μ and σ give the mean and standard deviation of the Gaussian core, whilst α defines the boundary point between the core and exponential tail. The slope of the tail is parameterized by parameter n.

The Crystal Ball and Gaussian shape share a common mean and standard deviation. The signal PDF used in the $D\pi^{\pm}$ fit category for each track type is given by

$$\begin{aligned} &F_{\text{sig}}^{Dh,tt}(m; N_{\text{sig}}^{Dh,tt}, \mu_{\text{sig}}^{tt}, \sigma_{\text{sig}}^{Dh,tt}) \equiv \\ &\quad N_{\text{sig}}^{Dh,tt} \cdot \big(f_{\text{sig}}^{Dh,tt} \cdot \mathcal{G}(m; \mu_{\text{sig}}^{tt}, \sigma_{\text{sig}}^{Dh,tt}) + (1 - f_{\text{sig}}^{Dh,tt}) \cdot \mathcal{CB}(m; \mu_{\text{sig}}^{tt}, \sigma_{\text{sig}}^{Dh,tt}, \alpha_{\text{sig}}^{Dh,tt}, n_{\text{sig}}^{Dh,tt})\big), \end{aligned} \tag{5.11}$$

where

- $\mathcal{G}(m;\mu,\sigma)$ denotes a Gaussian function of mean μ and standard deviation σ,
- $N_{\text{sig}}^{Dh,tt}$ denotes the signal yield,
- $f_{\text{sig}}^{Dh,tt}$ parameterizes the relative populations of the Gaussian and Crystal Ball components, and

Only the free parameters, $N_{\text{sig}}^{Dh,tt}$, μ_{sig}^{tt} and $\sigma_{\text{sig}}^{Dh,tt}$, are listed as arguments of the PDF function. This convention is used throughout the present chapter, to allow free parameters to be easily distinguished. All other parameters are fixed to their values from the previous analysis [141], where they were calibrated in fits to MC, shown in subplot (a) of Figures 5.1–5.4. Fixed parameter values used to describe

this PDF, and all others described in this section, are listed in Tables 5.2–5.3 for candidates of track type LL, and Tables 5.4–5.5 for candidates of track type DD.

A single mean, μ_{sig}^{tt}, is used to describe both $D\pi^{\pm}$ and $DK^{\pm}$ signal. The signal yield in the DK decay mode is parameterized in terms of that in the $D\pi$ decay mode, as

$$N_{\text{sig}}^{DK,tt} = r_N^{\text{tt}} \cdot N_{\text{sig}}^{D\pi,tt}, \tag{5.12}$$

where the ratio, r_N^{tt}, is a free parameter in the fit. Likewise, the width of the signal PDF in the DK decay mode is parameterized in terms of that in the $D\pi$ decay mode, as

$$\sigma_{\text{sig}}^{DK,tt} = r_\sigma^{\text{tt}} \cdot \sigma_{\text{sig}}^{D\pi,tt}, \tag{5.13}$$

where the ratio, r_σ^{tt}, is a free parameter in the fit.

Cross-feed

Cross-feed arises when candidates of one signal category are reconstructed under the opposite hypothesis, *i.e.* under a $(\pi \longleftrightarrow K)$ mis-identification. In each fit category, the contribution of such candidates is described using a two-tailed Crystal Ball function. This function modifies the Crystal Ball shape by the addition of an extra exponential tail on the high-mass side:

$$\mathcal{TCB}(m;\mu,\sigma,\alpha,n,\beta,l) \equiv \begin{cases} \frac{(\frac{n}{|\alpha|})^n \cdot \exp(-\frac{1}{2}|\alpha|^2)}{(\frac{n}{|\alpha|} - |\alpha| - \frac{m-\mu}{\sigma})^n}, & \frac{m-\mu}{\sigma} < \alpha, \\ \exp\big(-\frac{1}{2}(\frac{m-\mu}{\sigma})^2\big), & \alpha \leq \frac{m-\mu}{\sigma} < \beta, \\ \frac{(\frac{l}{|\beta|})^l \cdot \exp(-\frac{1}{2}|\beta|^2)}{(\frac{l}{|\beta|} - |\beta| + \frac{m-\mu}{\sigma})^l}, & \frac{m-\mu}{\sigma} \geq \beta. \end{cases} \tag{5.14}$$

This gives a shape with a Gaussian core and two independently parameterized exponential tails, and allows distributions with non-symmetric, non-Gaussian tails to be parameterized efficiently. The lower tail's boundary with the Gaussian core is parameterized by α, and its slope is parameterized by n. The upper tail's boundary with the Gaussian core is parameterized by β, and its slope is parameterized by l.

The cross feed contribution is given in the $D\pi$ decay mode by

$$F_{\text{xfeed}}^{D\pi,tt}(m; N_{\text{xfeed}}^{D\pi,tt}) = N_{\text{xfeed}}^{D\pi,tt} \cdot \mathcal{TCB}(m; \mu_{\text{xfeed}}^{D\pi,tt}, \sigma_{\text{xfeed}}^{D\pi,tt}, \alpha_{\text{xfeed}}^{D\pi,tt}, n_{\text{xfeed}}^{D\pi,tt}, \beta_{\text{xfeed}}^{D\pi,tt}, l_{\text{xfeed}}^{D\pi,tt}), \tag{5.15}$$

and in the DK decay mode by

$$F_{\text{xfeed}}^{DK,tt}(m; N_{\text{xfeed}}^{DK,tt}, \sigma_{\text{xfeed}}^{DK,tt}) = N_{\text{xfeed}}^{DK,tt} \cdot \mathcal{TCB}(m; \mu_{\text{xfeed}}^{DK,tt}, \sigma_{\text{xfeed}}^{DK,tt}, \alpha_{\text{xfeed}}^{DK,tt}, n_{\text{xfeed}}^{DK,tt}, \beta_{\text{xfeed}}^{DK,tt}, l_{\text{xfeed}}^{DK,tt}). \tag{5.16}$$

All of the shape parameters[4] except $\sigma_{\text{xfeed}}^{DK,tt}$ are fixed to their values from the previous analysis [141], where they were calibrated in fits to MC, as shown in subplot (b) of Figures 5.2–5.3.

[4] 'Shape parameters' is used to refer to all parameters describing a PDF, except for yields.

Background from $B \to D^*\pi$ decays

A single PDF shape is used to describe background arising from both $B^0 \to [D^{*\mp} \to D\pi^{\mp}]\pi^{\pm}$ and $B^{\pm} \to [D^{*0} \to D\gamma, \pi^0]\pi^{\pm}$ decays.

These decays can contribute to the selected $D\pi$ sample if the final state pion or photon emitted by the decaying D^*-meson is not correctly associated with the decay. With the addition of a $(\pi \longleftrightarrow K)$ mis-identification of the 'bachelor' pion[5], produced in the B-meson decay, these candidates can also enter as background in the DK sample. These backgrounds give a double-peaked shape in the reconstructed B-meson mass spectrum, due to the non-uniform angular distribution of the D^* daughters, which results from the helicity structure of the decay.

The decays $D^{*\mp} \to D\pi^{\mp}$ and $D^{*0} \to D\pi^0$ are all of the form Vector $\to$ Scalar Scalar (VSS). The helicity angle, θ, may be defined in each case, giving the angle between the pion momentum in the D^* rest frame, and the D^* momentum in the rest frame of the B-meson. The VSS structure of the decay favours the cases $\theta = 0$ and $\theta = \pi$, corresponding to a missing pion with momentum parallel or anti-parallel to its parent D^*. In the rest frame of the D^*-meson, the momentum carried by the pion is equal in either case, but accounting for the boost of the D^* in the lab frame separates the two cases, with more energy carried away by the pion when $\theta = \pi$, leading to a peak at lower reconstructed mass values, whilst the $\theta = 0$ gives a peak at higher mass values. The decay $D^{*0} \to D\gamma$ is of the form Vector $\to$ Scalar Vector (VSV). In this case the favoured values for the helicity angle of the missing photon are $\theta = \pi/2$ and $\theta = 3\pi/2$, and thus the two cases coincide to give a single peak.

In the fit to $B \to D\pi$ candidates, the sum of four Gaussian PDFs is used to describe this combination of background components, arranged in two pairs. Each pair consists of one wider and one narrower Gaussian PDF, sharing a common mean. This gives a double-peaked structure:

$$
\begin{aligned}
&F^{D\pi,tt}_{D^*\pi}(m; N^{D\pi,tt}_{D^*\pi\text{low}}, N^{D\pi,tt}_{D^*\pi\text{high}}, \mu^{D\pi,tt}_{D^*\pi\text{low}}, \mu^{D\pi,tt}_{D^*\pi\text{high}}, \sigma^{D\pi,tt}_{D^*\pi\text{low,narrow}}, \sigma^{D\pi,tt}_{D^*\pi\text{high,narrow}}) = \\
&N^{D\pi,tt}_{D^*\pi\text{low}} \cdot \left(f^{D\pi,tt}_{D^*\pi\text{low}} \cdot \mathcal{G}(m; \mu^{D\pi,tt}_{D^*\pi\text{low}}, \sigma^{D\pi,tt}_{D^*\pi\text{low,wide}}) + (1 - f^{D\pi,tt}_{D^*\pi\text{low}}) \cdot \mathcal{G}(m; \mu^{D\pi,tt}_{D^*\pi\text{low}}, \sigma^{D\pi,tt}_{D^*\pi\text{low,narrow}})\right) + \\
&N^{D\pi,tt}_{D^*\pi\text{high}} \cdot \left(f^{D\pi,tt}_{D^*\pi\text{high}} \cdot \mathcal{G}(m; \mu^{D\pi,tt}_{D^*\pi\text{high}}, \sigma^{D\pi,tt}_{D^*\pi\text{high,wide}}) + (1 - f^{D\pi,tt}_{D^*\pi\text{high}}) \cdot \mathcal{G}(m; \mu^{D\pi,tt}_{D^*\pi\text{high}}, \sigma^{D\pi,tt}_{D^*\pi\text{high,narrow}})\right)
\end{aligned}
\tag{5.17}
$$

where

- $f^{D\pi,tt}_{D^*\pi\text{low}}$ parameterizes the relative populations of the narrow and wide Gaussians constituting the lower mass peak,
- $f^{D\pi,tt}_{D^*\pi\text{high}}$ parameterizes the relative populations of the narrow and wide Gaussians constituting the higher mass peak, and

[5] Pions and kaons produced directly in the decay of the B-meson are typically referred to as 'bachelors', since unlike the pions produced in the decay of the D- or K^0_S-mesons they are not produced as one of a pair.

- The width of each broad peak is parameterized by its ratio relative to that of the corresponding narrow peak,

$$\sigma^{D\pi,tt}_{D^*\pi\text{low,wide}} = r^{D\pi,tt}_{D^*\pi,\sigma\text{low}} \cdot \sigma^{D\pi,tt}_{D^*\pi\text{low,narrow}}, \quad \text{and} \tag{5.18}$$

$$\sigma^{D\pi,tt}_{D^*\pi\text{high,wide}} = r^{D\pi,tt}_{D^*\pi,\sigma\text{high}} \cdot \sigma^{D\pi,tt}_{D^*\pi\text{high,narrow}}. \tag{5.19}$$

Calibration of this PDF to MC was performed in the previous analysis [141], as shown in subplot (c) of Figures 5.1 and 5.3. Previously all shape parameters were fixed from this calibration, but the present analysis allows four ($\mu^{D\pi,tt}_{D^*\pi\text{low}}$, $\mu^{D\pi,tt}_{D^*\pi\text{high}}$, $\sigma^{D\pi,tt}_{D^*\pi\text{low,narrow}}$, and $\sigma^{D\pi,tt}_{D^*\pi\text{high,narrow}}$) to float in the fit, subject to constraints discussed in the next section.

In the fit to $B \to DK$ candidates, two two-tailed Crystal Ball shapes are used, since the smearing caused by the mis-identification gives an asymmetry in the tails of the two peaks. These PDFs are described by

$$\begin{aligned} F^{DK,tt}_{D^*\pi}(m; N^{DK,tt}_{D^*\pi}) = N^{DK,tt}_{D^*\pi} \cdot \big(f^{DK,tt}_{D^*\pi} \cdot \mathcal{TCB}(m; \mu^{DK,tt}_{D^*\pi\text{low}}, \sigma^{DK,tt}_{D^*\pi\text{low}}, \alpha^{DK,tt}_{D^*\pi\text{low}}, n^{DK,tt}_{D^*\pi\text{low}}, \beta^{DK,tt}_{D^*\pi\text{low}}, l^{DK,tt}_{D^*\pi\text{low}}) + \\ (1 - f^{DK,tt}_{D^*\pi}) \cdot \mathcal{TCB}(m; \mu^{DK,tt}_{D^*\pi\text{high}}, \sigma^{DK,tt}_{D^*\pi\text{high}}, \alpha^{DK,tt}_{D^*\pi\text{high}}, n^{DK,tt}_{D^*\pi\text{high}}, \beta^{DK,tt}_{D^*\pi\text{high}}, l^{DK,tt}_{D^*\pi\text{high}})\big), \end{aligned} \tag{5.20}$$

where $f^{DK,tt}_{D^*\pi}$ parameterizes the relative populations of the lower and upper peaks. All of the shape parameters are fixed to their values from the previous analysis [141], where they were calibrated in fits to MC, as shown in subplot (f) of Figures 5.2 and 5.4.

Background from $B \to D^*\rho$ decays

A single PDF shape is used to describe background arising from both $B^0 \to [D^{*\mp} \to D\pi^{\mp}]\rho^{\pm}$ and $B^{\pm} \to [D^{*0} \to D\gamma, \pi^0]\rho^{\pm}$ decays. These decays can contribute background to the selected $B \to D\pi$ sample if, as above, the photon or pion emitted by the decaying D^*-meson is not correctly reconstructed. However, in this case it is necessary for the reconstruction to miss one of the ρ-meson's daughter pions as well. This additional missing particle smears the distributions obtained for these decays, such that in simulation no double-peaked structure is evident. Instead a single, broad Gaussian is used to describe the low-mass peak due to these decays. The mis-identification required for these decays to end up in the $B \to DK$ smears the distribution yet further and so a broad Gaussian is used to describe the shape in this case too.

These PDFs are described by

$$F^{Dh,tt}_{D^*\rho}(m; N^{Dh,tt}_{D^*\rho}) = N^{Dh,tt}_{D^*\rho} \cdot \mathcal{G}(\mu^{Dh,tt}_{D^*\rho}, \sigma^{Dh,tt}_{D^*\rho}), \tag{5.21}$$

where $\mu^{Dh,tt}_{D^*\rho}$ and $\sigma^{Dh,tt}_{D^*\rho}$ are fixed to their values from the previous analysis [141], where they were calibrated in fits to MC, as shown in subplot (d) of Figures 5.1 and 5.3, and subplot (g) of Figures 5.2 and 5.4.

Background from $B \to D\rho$ decays

A single PDF shape is used to describe background arising from both $B^0 \to D\rho^0$ and $B^\pm \to D\rho^\pm$ decays. In each case, these decays provide background to the selected $B \to D\pi$ sample if one of the ρ-meson's daughter pions is not associated with the decay in the reconstruction, whilst an additional mis-identification of the remaining daughter pion gives a background which can enter into the $B \to DK$ sample.

When reconstructed as $B \to D\pi$, this background gives a distribution with a sharp endpoint, and is described using the convolution of a Gaussian with an ARGUS function [144],

$$\text{ARGUS}(m; m_0, c, p) \equiv \begin{cases} m\left(1 - (\frac{m}{m_0})^2\right)^p \exp\left(c\left(1 - (\frac{m}{m_0})^2\right)\right), & m \leq m_0, \\ 0, & m > m_0, \end{cases} \tag{5.22}$$

where m_0 denotes the position of the high-mass endpoint, c parameterizes the curvature of the function, and p parameterizes the power of the low-mass slope. When reconstructed as $B \to DK$, the additional mis-identification required smears out any endpoint structure, and instead the sum of a Gaussian with a decreasing exponential is used to describe this component.

These PDFs are described by

$$F_{D\rho}^{D\pi,tt}(m; N_{D\rho}^{D\pi,tt}) = N_{D\rho}^{D\pi,tt} \cdot \left(\mathcal{G}(\mu_{D\rho}^{D\pi,tt}, \sigma_{D\rho}^{D\pi,tt}) \star \text{ARGUS}(m; (m_0)_{D\rho}^{D\pi,tt}, c_{D\rho}^{D\pi,tt}, p_{D\rho}^{D\pi,tt})\right), \tag{5.23}$$

and

$$F_{D\rho}^{DK,tt}(m; N_{D\rho}^{DK,tt}) = N_{D\rho}^{DK,tt} \cdot \left(f_{D\rho\text{gaus}}^{DK,tt} \cdot \mathcal{G}(\mu, \sigma) + (1 - f_{D\rho\text{gaus}}^{DK,tt}) \cdot \mathrm{e}^{\tau_{D\rho}^{DK,tt} m} |\tau_{D\rho}^{DK,tt}|\right), \tag{5.24}$$

where $f_{D\rho\text{gaus}}^{DK,tt}$ parameterizes the relative populations of the Gaussian and exponential components. All of the shape parameters are fixed to their values from the previous analysis [141], where they were calibrated in fits to MC, shown in subplot (e) of Figures 5.1 and 5.3, and subplot (h) of Figures 5.2 and 5.4.

Background from $B \to D^*K$ decays

This background is analogous to that arising from $B \to D^*\pi$ decays, and is considered only for the reconstructed $B \to DK$ sample. Due to its lower branching fraction, and the additional mis-identification required for these candidates to enter the $B \to D\pi$ sample, any contribution there is negligible. As for the background contribution from $B \to D^*\pi$ decays entering the $B \to D\pi$ sample, four Gaussian PDFs are used to describe this background component. Again, the Gaussian components are arranged in two pairs, each pair having a common mean, to give a double-peaked shape.

These PDFs are described by

$$F^{DK,tt}_{D^*K}(m; N^{DK,tt}_{D^*\text{Klow}}, N^{DK,tt}_{D^*\text{Khigh}}, \sigma^{DK,tt}_{D^*\text{Klow,narrow}}, \sigma^{DK,tt}_{D^*\text{Khigh,narrow}}) =$$
$$N^{DK,tt}_{D^*\text{Klow}} \cdot \left(f^{DK,tt}_{D^*\text{Klow}} \cdot \mathcal{G}(m; \mu^{DK,tt}_{D^*\text{Klow}}, \sigma^{DK,tt}_{D^*\text{Klow,wide}}) + (1 - f^{DK,tt}_{D^*\text{Klow}}) \cdot \mathcal{G}(m; \mu^{DK,tt}_{D^*\text{Klow}}, \sigma^{DK,tt}_{D^*\text{Klow,narrow}})\right) +$$
$$N^{DK,tt}_{D^*\text{Khigh}} \cdot \left(f^{DK,tt}_{D^*\text{Khigh}} \cdot \mathcal{G}(m; \mu^{DK,tt}_{D^*\text{Khigh}}, \sigma^{DK,tt}_{D^*\text{Khigh,wide}}) + (1 - f^{DK,tt}_{D^*\text{Khigh}}) \cdot \mathcal{G}(m; \mu^{DK,tt}_{D^*\text{Khigh}}, \sigma^{DK,tt}_{D^*\text{Khigh,narrow}})\right), \tag{5.25}$$

where

- $f^{DK,tt}_{D^*\text{Klow}}$ parameterizes the relative populations of the narrow and wide Gaussians constituting the lower mass peak,
- $f^{DK,tt}_{D^*\text{Khigh}}$ parameterizes the relative populations of the narrow and wide Gaussians constituting the higher mass peak, and
- The width of each broad peak is parameterized by its ratio relative to that of the corresponding narrow peak,

$$\sigma^{DK,tt}_{D^*\text{Klow,wide}} = r^{DK,tt}_{D^*K,\sigma\text{low}} \cdot \sigma^{DK,tt}_{D^*\text{Klow,narrow}}, \quad \text{and} \tag{5.26}$$
$$\sigma^{DK,tt}_{D^*\text{Khigh,wide}} = r^{DK,tt}_{D^*K,\sigma\text{high}} \cdot \sigma^{DK,tt}_{D^*\text{Khigh,narrow}}. \tag{5.27}$$

All of the shape parameters except $\sigma^{DK,tt}_{D^*\text{Klow,narrow}}$ and $\sigma^{DK,tt}_{D^*\text{Khigh,narrow}}$ are fixed to their values from the previous analysis [141], where they were calibrated in fits to MC, as shown in subplot (e) of Figures 5.2 and 5.4.

Background from $B^0_{(s)} \to DK^{*0}$ decays

Backgrounds arising from $B^0 \to DK^{*0}$ and $B^0_s \to DK^{*0}$ candidates are considered only for the reconstructed $B \to DK$ sample. Due to the mass shift between these two contributions caused by the difference in B^0 and B^0_s masses, separate shapes are used in each case, with each background parameterized using a Gaussian convolved with an ARGUS function.

These PDFs are described by

$$F^{DK,tt}_{B^0}(m; N^{DK,tt}_{B^0}) = N^{DK,tt}_{B^0} \cdot \left(\mathcal{G}(\mu^{DK,tt}_{B^0}, \sigma^{DK,tt}_{B^0}) \star \text{ARGUS}(m; (m_0)^{DK,tt}_{B^0}, c^{DK,tt}_{B^0}, p^{DK,tt}_{B^0})\right), \tag{5.28}$$

and

$$F^{DK,tt}_{B^0_s}(m; N^{DK,tt}_{B^0_s}) = N^{DK,tt}_{B^0_s} \cdot \left(\mathcal{G}(\mu^{DK,tt}_{B^0_s}, \sigma^{DK,tt}_{B^0_s}) \star \text{ARGUS}(m; (m_0)^{DK,tt}_{B^0_s}, c^{DK,tt}_{B^0_s}, p^{DK,tt}_{B^0_s})\right). \tag{5.29}$$

All of the shape parameters are fixed to their values from the previous analysis [141], where they were calibrated in fits to MC, as shown in subplots (c) and (d) of Figures 5.2 and 5.4.

Combinatoric background

Combinatoric candidates arise when tracks are incorrectly associated with an event in the reconstruction, giving one or more incorrectly-reconstructed vertices, and thus an incorrectly-reconstructed K^0_S-, D- or B-meson. The reconstructed B-meson mass in this case strongly depends on the energy of the incorrectly-associated track, and since there are many background processes which can give rise to low energy pions and kaons in LHC collisions, a decreasing exponential is used to describe this contribution. Independent shapes are used in the two fit categories.

These PDFs are described by

$$F^{Dh,tt}_{\text{comb}}(m; N^{Dh,tt}_{\text{comb}}, \tau^{Dh,tt}_{\text{comb}}) = N^{Dh,tt}_{\text{comb}} \cdot \mathrm{e}^{\tau^{Dh,tt}_{\text{comb}} m} |\tau^{Dh,tt}_{\text{comb}}|, \tag{5.30}$$

and the slope parameters, $\tau^{Dh,tt}_{\text{comb}}$, are left free in the final fit.

Component	Shape in $B^{\pm} \to D\pi^{\pm}$	Shape in $B^{\pm} \to DK^{\pm}$	Description
$B^{\pm} \to DK^{\pm}$	Two-tailed Crystal Ball	Gaussian + Crystal Ball	Signal mode
$B^{\pm} \to D\pi^{\pm}$	Gaussian + Crystal Ball	Two-tailed Crystal Ball	Control mode in fit
$B \to D^{*}\pi$	Two pairs of Gaussians	Two two-tailed Crystal Balls	One D^{*} daughter not reconstructed
$B \to D^{*}\rho$	Gaussian	Gaussian	One D^{*} daughter and one ρ daughter not reconstructed
$B \to D\rho$	Gaussian $\star$ ARGUS	Gaussian + exponential	One ρ daughter not reconstructed
$B \to D^{*}K$	–	Two pairs of Gaussians	One D^{*} daughter not reconstructed
$B^{0} \to DK^{*0}$	–	Gaussian $\star$ ARGUS	One K^{*0} daughter not reconstructed
$B_{s}^{0} \to DK^{*0}$	–	Gaussian $\star$ ARGUS	One K^{*0} daughter not reconstructed
Combinatoric	Decreasing exponential	Decreasing exponential	Candidate reconstructed using track/s from wrong event

Table 5.1.: Components included in the reconstructed B-mass fit to selected $B^{\pm} \to Dh^{\pm}$ candidates.

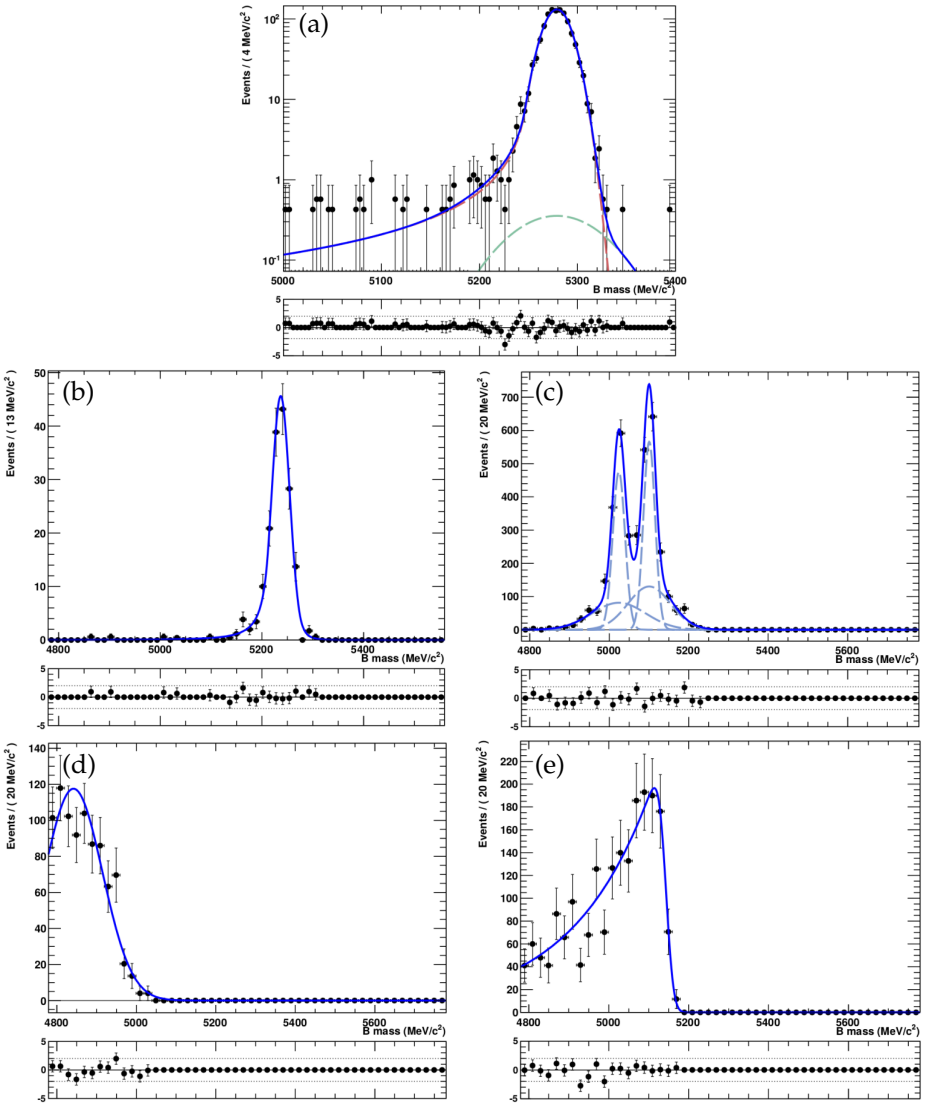


Figure 5.1.: B-mass PDF shapes used in the fit to LL $B^{\pm} \to D\pi^{\pm}$ candidates. These shapes are taken directly from a previous analysis [141]. Signal MC (a) is described using the sum of a Gaussian and a Crystal Ball function. Crossfeed from $B^{\pm} \to DK^{\pm}$ MC (b) is described using a two-tailed Crystal Ball function. $B \to D^*\pi$ MC (c) is described using the sum of four Gaussian shapes. $B \to D^*\rho$ MC (d) is described using a single Gaussian. $B \to D\rho$ MC (e) is described using the convolution of a Gaussian and an ARGUS shape. In each plot, histogram points show Monte Carlo simulation, and a solid blue line gives the PDF model fitted. Where the PDF model consists of a sum of multiple constituent shapes, these are indicated by dashed lines. In each plot, the lower canvas shows the pull.

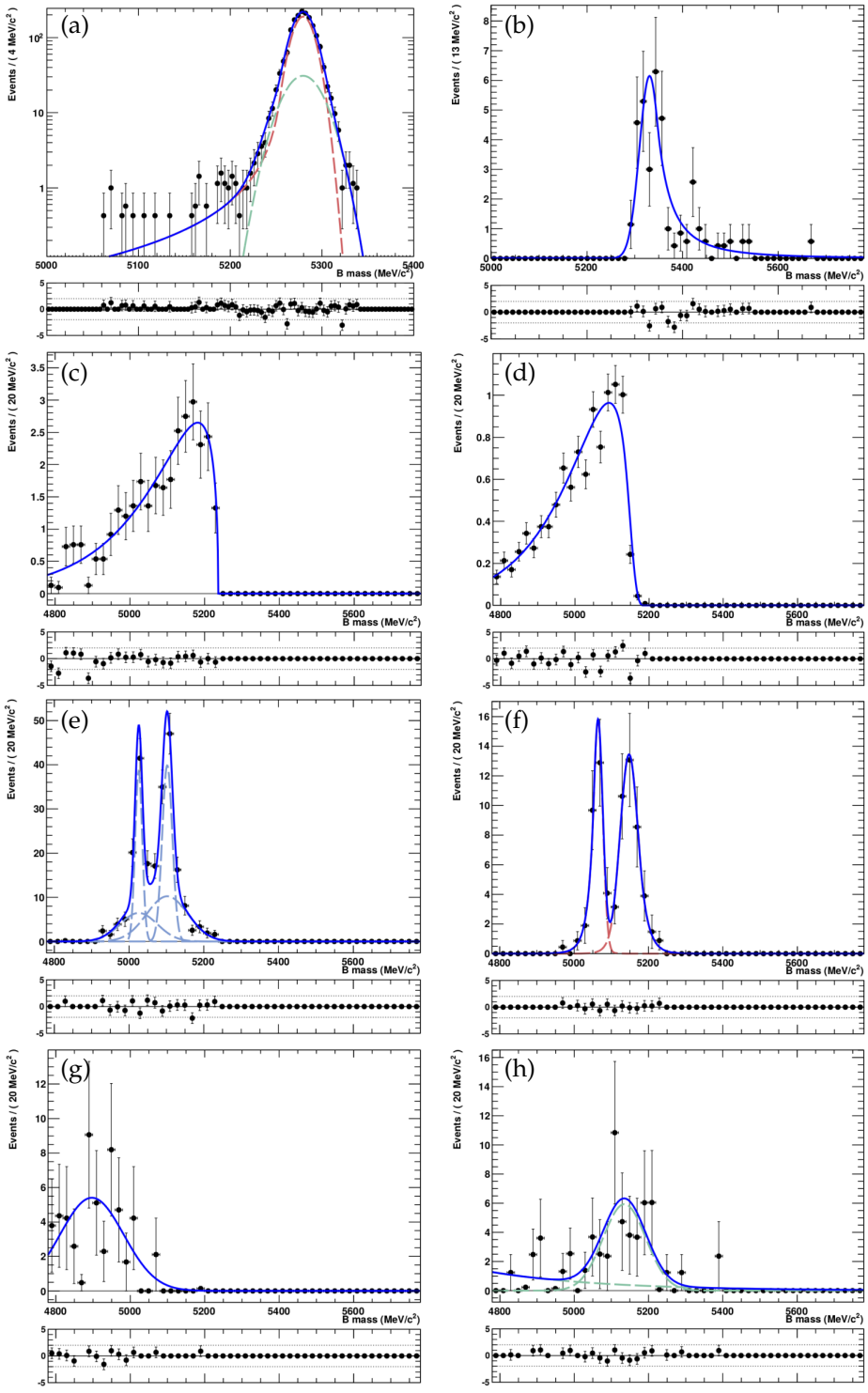


Figure 5.2.: B-mass PDF shapes used in the fit to LL $B^{\pm} \to DK^{\pm}$ candidates. These shapes are taken directly from a previous analysis [141]. Signal MC (a) is described using the sum of a Gaussian and a Crystal Ball function. Crossfeed from $B^{\pm} \to D\pi^{\pm}$ MC (b) is described using a two-tailed Crystal Ball function. $B_s^0 \to DK^{*0}$ MC (c) and $B_d^0 \to DK^{*0}$ MC (d) are each is described using the convolution of a Gaussian and an ARGUS shape. $B \to D^*K$ MC (e) is described using the sum of four Gaussian shapes. $B \to D^*\pi$ MC (f) is described using the sum of two two-tailed Crystal Ball functions. $B \to D^*\rho$ MC (g) is described using a single Gaussian. $B \to D\rho$ MC (h) is described using the sum of a Gaussian and a decreasing exponential function. In each plot, histogram points show Monte Carlo simulation, and a solid blue line gives the PDF model fitted. Where the PDF model consists of a sum of multiple constituent shapes, these are indicated by dashed lines. In each plot, the lower canvas shows the pull.

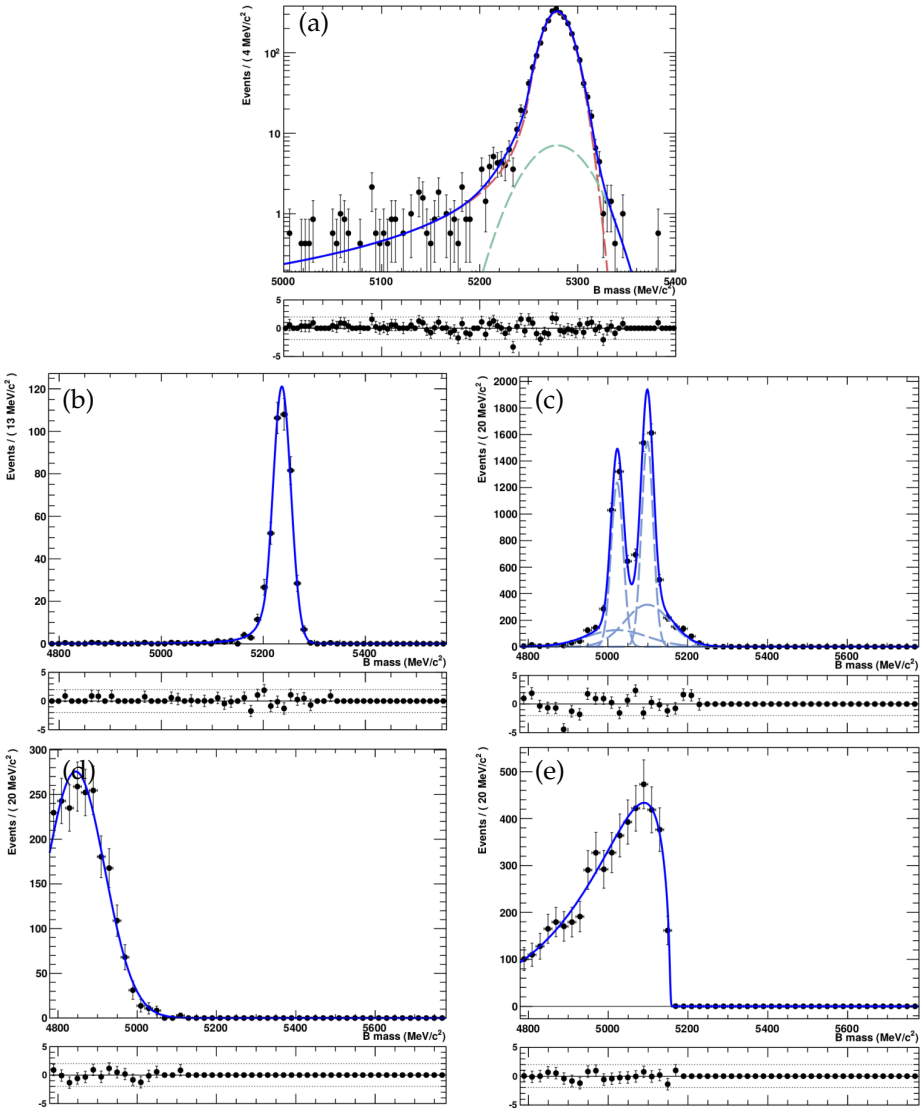


Figure 5.3.: *B*-mass PDF shapes used in the fit to DD $B^{\pm} \to D\pi^{\pm}$ candidates. These shapes are taken directly from a previous analysis [141]. Signal MC (a) is described using the sum of a Gaussian and a Crystal Ball function. Crossfeed from $B^{\pm} \to DK^{\pm}$ MC (b) is described using a two-tailed Crystal Ball function. $B \to D^*\pi$ MC (c) is described using the sum of four Gaussian shapes. $B \to D^*\rho$ MC (d) is described using a single Gaussian. $B \to D\rho$ MC (e) is described using the convolution of a Gaussian and an ARGUS shape. In each plot, histogram points show Monte Carlo simulation, and a solid blue line gives the PDF model fitted. Where the PDF model consists of a sum of multiple constituent shapes, these are indicated by dashed lines. In each plot, the lower canvas shows the pull.

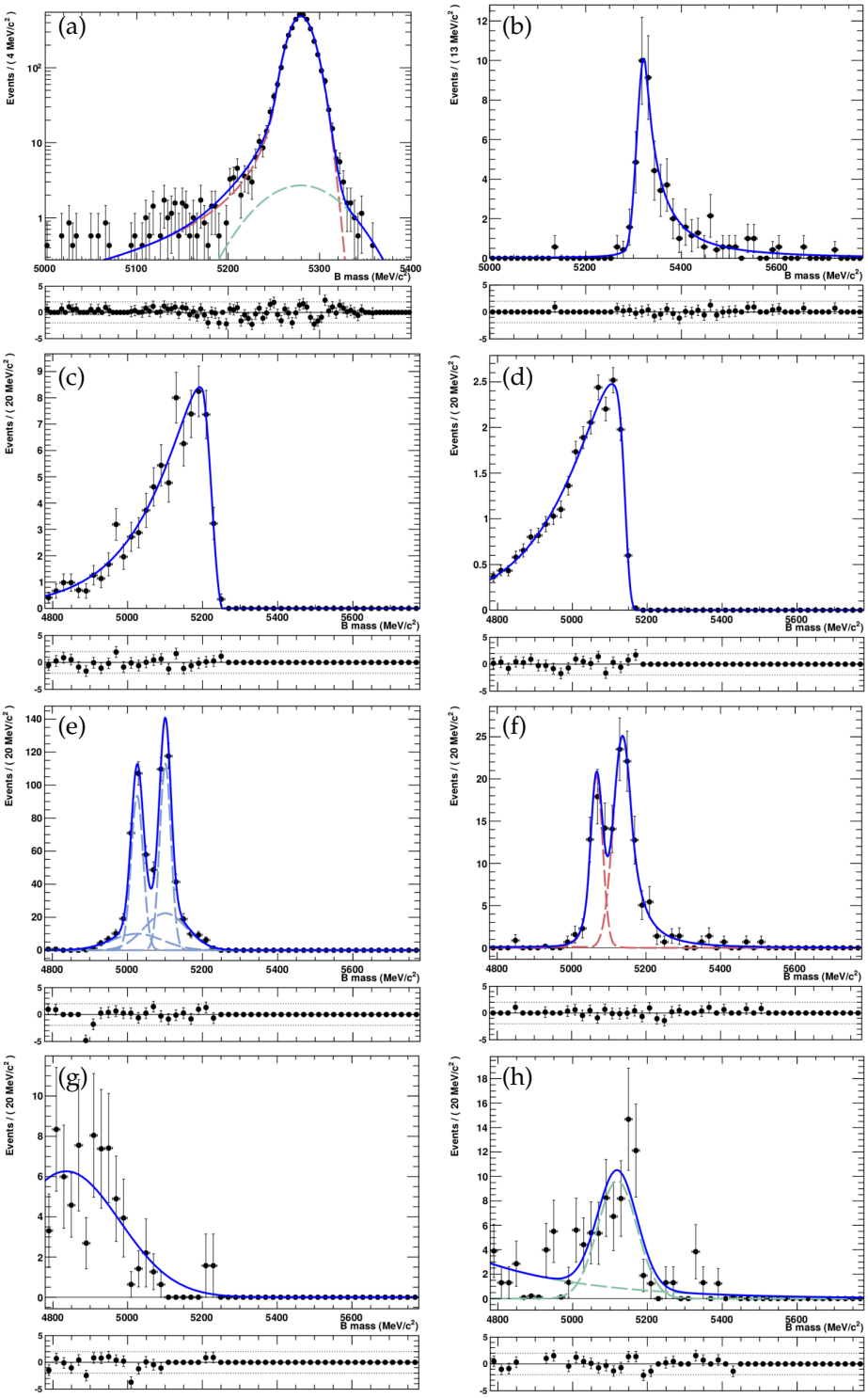


Figure 5.4.: *B*-mass PDF shapes used in the fit to DD $B^{\pm} \to DK^{\pm}$ candidates. These shapes are taken directly from a previous analysis [141]. Signal MC (a) is described using the sum of a Gaussian and a Crystal Ball function. Crossfeed from $B^{\pm} \to D\pi^{\pm}$ MC (b) is described using a two-tailed Crystal Ball function. $B_s^0 \to DK^{*0}$ MC (c) and $B_d^0 \to DK^{*0}$ MC (d) are each is described using the convolution of a Gaussian and an ARGUS shape. $B \to D^*K$ MC (e) is described using the sum of four Gaussian shapes. $B \to D^*\pi$ MC (f) is described using the sum of two two-tailed Crystal Ball functions. $B \to D^*\rho$ MC (g) is described using a single Gaussian. $B \to D\rho$ MC (h) is described using the sum of a Gaussian and a decreasing exponential function. In each plot, histogram points show Monte Carlo simulation, and a solid blue line gives the PDF model fitted. Where the PDF model consists of a sum of multiple constituent shapes, these are indicated by dashed lines. In each plot, the lower canvas shows the pull.

5.1.1. Constraints applied

The full PDF model described so far contains 29 degrees of freedom[6], and may be summarised as

$$\begin{aligned} F_{\text{tot}}^{D\pi,tt} \equiv\ & F_{\text{sig}}^{D\pi,tt}(m; N_{\text{sig}}^{D\pi,tt}, \mu_{\text{sig}}^{tt}, \sigma_{\text{sig}}^{D\pi,tt}) + F_{\text{xfeed}}^{D\pi,tt}(m; N_{\text{xfeed}}^{D\pi,tt}) \\ & + F_{D^*\pi}^{D\pi,tt}(m; N_{D^*\pi\text{low}}^{D\pi,tt}, N_{D^*\pi\text{high}}^{D\pi,tt}, \mu_{D^*\pi\text{low}}^{D\pi,tt}, \mu_{D^*\pi\text{high}}^{D\pi,tt}, \sigma_{D^*\pi\text{low,narrow}}^{D\pi,tt}, \sigma_{D^*\pi\text{high,narrow}}^{D\pi,tt}) \\ & + F_{D^*\rho}^{D\pi,tt}(m; N_{D^*\rho}^{D\pi,tt}) + F_{D\rho}^{D\pi,tt}(m; N_{D\rho}^{D\pi,tt}) + F_{\text{comb}}^{D\pi,tt}(m; N_{\text{comb}}^{D\pi,tt}, \tau_{\text{comb}}^{D\pi,tt}), \end{aligned} \tag{5.31}$$

and

$$\begin{aligned} F_{\text{tot}}^{DK,tt} \equiv\ & F_{\text{sig}}^{DK,tt}(m; r_N^{tt}, \mu_{\text{sig}}^{tt}, r_\sigma^{tt}) + F_{\text{xfeed}}^{DK,tt}(m; N_{\text{xfeed}}^{DK,tt}, \sigma_{\text{xfeed}}^{DK,tt}) + F_{D^*\pi}^{DK,tt}(m; N_{D^*\pi}^{DK,tt}) \\ & + F_{D^*\rho}^{DK,tt}(m; N_{D^*\rho}^{DK,tt}) + F_{D\rho}^{DK,tt}(m; N_{D\rho}^{DK,tt}) + F_{D^*K}^{DK,tt}(m; N_{D^*K\text{low}}^{DK,tt}, N_{D^*K\text{high}}^{DK,tt}, \sigma_{D^*K\text{low,narrow}}^{DK,tt}, \sigma_{D^*K\text{high,narrow}}^{DK,tt}) \\ & + F_{B^0}^{DK,tt}(m; N_{B^0}^{DK,tt}) + F_{B_s^0}^{DK,tt}(m; N_{B_s^0}^{DK,tt}) + F_{\text{comb}}^{DK,tt}(m; N_{\text{comb}}^{DK,tt}, \tau_{\text{comb}}^{DK,tt}). \end{aligned} \tag{5.32}$$

Constraints are applied to reduce the degrees of freedom and improve the fit quality. This is necessary due to the number of overlapping low-mass backgrounds present in both the $DK^\pm$ and $D\pi^\pm$ fits.

- The cross-feed yield of each fit category ($B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$) in the simultaneous fit is linked to the signal yield of the opposing category, using the efficiency of the PID criterion used to separate the two categories:

$$N_{\text{xfeed}}^{DK,tt} = N_{\text{sig}}^{D\pi,tt} \cdot \frac{\epsilon_{\text{DLL}_{K\pi}(\pi)>4}}{\epsilon_{\text{DLL}_{K\pi}(\pi)<4}}, \quad \text{and} \tag{5.33}$$

$$N_{\text{xfeed}}^{D\pi,tt} = N_{\text{sig}}^{DK,tt} \cdot \frac{\epsilon_{\text{DLL}_{K\pi}(K)<4}}{\epsilon_{\text{DLL}_{K\pi}(K)>4}}, \tag{5.34}$$

  where the efficiencies are fixed to their values determined in the previous analysis [141],

$$\epsilon_{\text{DLL}_{K\pi}(\pi)>4} / \epsilon_{\text{DLL}_{K\pi}(\pi)<4} = 0.039, \quad \text{and} \tag{5.35}$$

$$\epsilon_{\text{DLL}_{K\pi}(K)<4} / \epsilon_{\text{DLL}_{K\pi}(K)>4} = 0.185. \tag{5.36}$$

- Low mass background yields in the $B^\pm \to DK^\pm$ fit category are fixed relative to the low mass background with the largest yield, $B \to D^*\pi$,

$$N_{D\rho}^{DK,tt} = r_{D\rho,D^*\pi}^{DK} \cdot N_{D^*\pi}^{DK,tt}, \tag{5.37}$$

$$N_{B_s^0}^{DK,tt} = r_{B_s^0,D^*\pi}^{DK} \cdot N_{D^*\pi}^{DK,tt}, \quad \text{and} \tag{5.38}$$

$$N_{B^0}^{DK,tt} = r_{B^0,D^*\pi}^{DK} \cdot N_{D^*\pi}^{DK,tt}, \tag{5.39}$$

  where the ratios $r_{D\rho,D^*\pi}^{DK}$, $r_{B_s^0,D^*\pi}^{DK}$ and $r_{B^0,D^*\pi}^{DK}$ are fixed to the values extracted from MC by the previous analysis [141]:

[6] Note that the signal mean, μ_{sig}^{tt}, is common to both PDFs.

  - $r^{DK}_{D\rho,D^*\pi} = 0.61$;
  - $r^{DK}_{B^0_s,D^*\pi} = 0.74$; and
  - $r^{DK}_{B^0,D^*\pi} = 0.17$.

- The widths of the $B \to D^*K$ double-peaked shapes in the $B^\pm \to DK^\pm$ category are constrained relative to the $B \to D^*\pi$ widths in the $B^\pm \to D\pi^\pm$ category. Their ratio is constrained to match that between the $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ signal widths, since (if the small difference in bachelor mass is neglected) the kinematics and relative efficiency between $Dh^\pm$ signal and $D^*h^\pm$ background should be the same in each case:

$$\sigma^{DK,tt}_{D^*\text{Klow,narrow}} = r^{tt}_\sigma \cdot \sigma^{D\pi,tt}_{D^*\pi\text{low,narrow}}; \quad \text{and} \tag{5.40}$$

$$\sigma^{DK,tt}_{D^*\text{Khigh,narrow}} = r^{tt}_\sigma \cdot \sigma^{D\pi,tt}_{D^*\pi\text{high,narrow}}. \tag{5.41}$$

- Finally, for the low-mass background arising in the $B^\pm \to D\pi^\pm$ category due to $B \to D^*\pi$ candidates, constraints are implemented for the mean and width of each peak (each peak consists of the sum of two Gaussians, but their means are fixed equal and their widths are fixed in the ratio found from simulation, leaving a single mean and width parameter per peak). Since this background is the dominant low-mass contribution for the high-statistics control mode fit, the values of these parameters have the potential to strongly affect the fit quality if they are fixed to a slightly incorrect value. Thus, instead of fixing them to the central values obtained from fits to simulation, they are allowed to float subject to Gaussian constraints, where the constraining Gaussians have their widths set equal to the standard deviations returned from MC fits performed in the previous analysis [141].

  - $\mu^{D\pi,tt}_{D^*\pi\text{low}}$ is subject to a Gaussian constraint of mean 5027.4 MeV, and standard deviation 5 MeV;
  - $\mu^{D\pi,tt}_{D^*\pi\text{high}}$ is subject to a Gaussian constraint of mean 5103.8 MeV, and standard deviation 5 MeV;
  - $\sigma^{D\pi,tt}_{D^*\pi\text{low,narrow}}$ is subject to a Gaussian constraint of mean 15.4 MeV, and standard deviation 1.5 MeV; and
  - $\sigma^{D\pi,tt}_{D^*\pi\text{high,narrow}}$ is subject to a Gaussian constraint of mean 14.6 MeV, and standard deviation 1.5 MeV.

  Each constraining Gaussian is included as a multiplicative factor in the likelihood function of the fit, giving an additive penalty term in the NLL.

After the application of the above constraints, 22 free parameters remain in each simultaneous fit, including the four parameters that are linked by Gaussian constraints.

5.1.2. Mass fit result

Simulataneous fits are performed to the B-mass spectra of reconstructed $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ candidates. The full set of fixed parameters used in these fits is tabulated, for LL and DD candidate events, in Tables 5.2–5.3 and Tables 5.4–5.5. Some of the fixed values are very small, indicating that it may in future be worth considering simplifying the PDF shapes used to describe some components. For example, the parameter $p_{D\rho}^{D\pi,LL}$, obtained in the calibration to LL MC candidates performed in reference [141], is very small. Since p enters the ARGUS PDF only as the exponent in the factor $(1-(m/m_0)^2)^p$, this suggests that in the fit to LL candidates the factor could be replaced by unity. However, the parameter value obtained in the fit to DD candidates, $p_{D\rho}^{D\pi,DD}$, is not negligible, and here the full ARGUS PDF must be included. For simplicity the present analysis uses the full ARGUS PDF for both cases.

Fit results and corresponding pulls are plotted in Figure 5.5 and Figure 5.6 for candidates reconstructed with LL and DD K_S^0 candidates, respectively. The returned values for the free parameters of the fits to LL and DD categories are enumerated in Tables 5.6 and 5.7, and their correlations are illustrated in Figure 5.7.

The fits to reconstructed $B^\pm \to D\pi^\pm$ candidates (shown in the lower part of Figure 5.5 and Figure 5.6) both show pulls of relatively large magnitudes, indicating statistically significant deviations between the observed bin populations and those predicted by the fitted model. However, these fits are performed solely to provide constraints on background yields and shapes entering the γ-sensitive $B^\pm \to DK^\pm$ sample, and the fits to these candidates do not exhibit large pulls.

Only candidates in the $B^\pm \to DK^\pm$ sample falling within a $\pm 3\sigma$ window centred on the fitted signal mean are propagated forward to the full $\mathcal{CP}$ fit. This window is marked by vertical lines in Figures 5.5 and 5.6. For LL candidates, the window used is $[5237; 5327]$ MeV, and for DD candidates it is $[5237; 5328]$ MeV. These windows differ slightly due to the small differences in fitted signal mean, μ_{sig}^{tt}, and signal width ratio, r_N^{tt}, obtained in the fits to LL and DD candidates. Component yields extracted within these windows are enumerated in Table 5.8.

It may seem that the complex fit model developed for this mode is unnecessary, since most peaking background components are negligible within the signal region. However, the inclusion of these components allows a reliable fit in the full mass range from the signal region down to 4800 MeV, and thus improves the estimation of the slope of the combinatoric component, which extends across the full range and does give a non-negligible contribution in the signal region. A future analysis could perhaps consider excluding the region between approximately 4825 MeV and 5225 MeV from the fit. This approach may make it possible to remove almost all peaking background shapes from the fit model, whilst still reliably constraining the combinatoric slope using data from the regions below 4825 MeV (where combinatoric background appears to dominate) and above 5225 MeV.

Constant Parameter	Value	Description
$\alpha_{\text{sig}}^{D\pi,LL}$	2.4	Signal PDF parameters in $D\pi$ fit
$n_{\text{sig}}^{D\pi,LL}$	1.2	
$\sigma_{\text{sig}}^{D\pi,LL}$	45.0 MeV	
$\alpha_{\text{xfeed}}^{D\pi,LL}$	2.1	Cross-feed PDF parameters in $D\pi$ fit
$\beta_{\text{xfeed}}^{D\pi,LL}$	1.4	
$\mu_{\text{xfeed}}^{D\pi,LL}$	5239.3 MeV	
$\sigma_{\text{xfeed}}^{D\pi,LL}$	17.6 MeV	
$l_{\text{xfeed}}^{D\pi,LL}$	2.4	
$n_{\text{xfeed}}^{D\pi,LL}$	6.6	
$\mu_{D^*\pi\text{low}}^{D\pi,LL}$	5026.7 MeV	$B \to D^*\pi$ background PDF parameters in $D\pi$ fit
$\sigma_{D^*\pi\text{low,narrow}}^{D\pi,LL}$	15.4 MeV	
$\sigma_{D^*\pi\text{low,wide}}^{D\pi,LL}$	67.5 MeV	
$\mu_{D^*\pi\text{high}}^{D\pi,LL}$	5103.1 MeV	
$\sigma_{D^*\pi\text{high,narrow}}^{D\pi,LL}$	14.6 MeV	
$\sigma_{D^*\pi\text{high,wide}}^{D\pi,LL}$	53.3 MeV	
$\mu_{D^*\rho}^{D\pi,LL}$	4844.4 MeV	$B \to D^*\rho$ background PDF parameters in $D\pi$ fit
$\sigma_{D^*\rho}^{D\pi,LL}$	72.9 MeV	
$(m_0)_{D\rho}^{D\pi,LL}$	5145.0 MeV	$B \to D\rho$ background PDF parameters in $D\pi$ fit
$c_{D\rho}^{D\pi,LL}$	-12.5	
$p_{D\rho}^{D\pi,LL}$	1.4×10^{-5}	
$\sigma_{D\rho}^{D\pi,LL}$	14.9 MeV	

Table 5.2.: Constant parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected LL $B^{\pm} \to D\pi^{\pm}$ candidates. These values were calibrated in the previous analysis [141].

Constant Parameter	Value	Description
$\alpha_{\text{sig}}^{DK,LL}$	2.3	Signal PDF parameters in DK fit
$n_{\text{sig}}^{DK,LL}$	1.5	
$\alpha_{\text{xfeed}}^{DK,LL}$	0.8	Cross-feed PDF parameters in DK fit
$\beta_{\text{xfeed}}^{DK,LL}$	3.5	
$\mu_{\text{xfeed}}^{DK,LL}$	5333.3 MeV	
$l_{\text{xfeed}}^{DK,LL}$	1.1×10^{-4}	
$n_{\text{xfeed}}^{DK,LL}$	2.4	
$\mu_{D^*\text{Klow}}^{DK,LL}$	5028.3 MeV	$B \to D^*K$ background PDF parameters in DK fit
$\sigma_{D^*\text{Klow,narrow}}^{DK,LL}$	9.5 MeV	
$\sigma_{D^*\text{Klow,wide}}^{DK,LL}$	49.8 MeV	
$\mu_{D^*\text{Khigh}}^{DK,LL}$	5104.6 MeV	
$\sigma_{D^*\text{Khigh,narrow}}^{DK,LL}$	13.6 MeV	
$\sigma_{D^*\text{Khigh,wide}}^{DK,LL}$	55.2 MeV	
$\alpha_{D^*\pi\text{low}}^{DK,LL}$	2.9	$B \to D^*\pi$ background PDF parameters in DK fit
$\beta_{D^*\pi\text{low}}^{DK,LL}$	0.9	
$\mu_{D^*\pi\text{low}}^{DK,LL}$	5066.7 MeV	
$\sigma_{D^*\pi\text{low}}^{DK,LL}$	12.7 MeV	
$l_{D^*\pi\text{low}}^{DK,LL}$	13.5	
$n_{D^*\pi\text{low}}^{DK,LL}$	4.3	
$\alpha_{D^*\pi\text{high}}^{DK,LL}$	1.3	
$\beta_{D^*\pi\text{high}}^{DK,LL}$	1.9	
$\mu_{D^*\pi\text{high}}^{DK,LL}$	5150.2 MeV	
$\sigma_{D^*\pi\text{high}}^{DK,LL}$	24.3 MeV	
$l_{D^*\pi\text{high}}^{DK,LL}$	24.3	
$n_{D^*\pi\text{high}}^{DK,LL}$	109.9	
$\mu_{D^*\rho}^{DK,LL}$	4899.3 MeV	$B \to D^*\rho$ background PDF parameters in DK fit
$\sigma_{D^*\rho}^{DK,LL}$	85.3 MeV	
$\mu_{D\rho}^{DK,LL}$	5137.6 MeV	$B \to D\rho$ background PDF parameters in DK fit
$\sigma_{D\rho}^{DK,LL}$	57.6 MeV	
$\tau_{D\rho}^{DK,LL}$	-3.3×10^{-3}	
$(m_0)_{B_s^0}^{DK,LL}$	5238.8 MeV	$B_s^0 \to DK^{*0}$ background PDF parameters in DK fit
$c_{B_s^0}^{DK,LL}$	-21.0	
$p_{B_s^0}^{DK,LL}$	0.4	
$\sigma_{B_s^0}^{DK,LL}$	8.7×10^{-2} MeV	
$(m_0)_{B^0}^{DK,LL}$	5159.8 MeV	$B^0 \to DK^{*0}$ background PDF parameters in DK fit
$c_{B^0}^{DK,LL}$	-26.3	
$p_{B^0}^{DK,LL}$	0.7	
$\sigma_{B^0}^{DK,LL}$	13.0 MeV	

Table 5.3.: Constant parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected LL $B^\pm \to DK^\pm$ candidates. These values were calibrated in the previous analysis [141].

Constant Parameter	Value	Description
$\alpha_{\text{sig}}^{D\pi,DD}$	2.3	Signal PDF parameters in $D\pi$ fit
$n_{\text{sig}}^{D\pi,DD}$	1.3	
$\sigma_{\text{sig}}^{D\pi,DD}$	42.1 MeV	
$\alpha_{\text{xfeed}}^{D\pi,DD}$	3.2	Cross-feed PDF parameters in $D\pi$ fit
$\beta_{\text{xfeed}}^{D\pi,DD}$	1.6	
$\mu_{\text{xfeed}}^{D\pi,DD}$	5239.8 MeV	
$\sigma_{\text{xfeed}}^{D\pi,DD}$	17.6 MeV	
$l_{\text{xfeed}}^{D\pi,DD}$	2.3	
$n_{\text{xfeed}}^{D\pi,DD}$	1.8	
$\mu_{D^*\pi\text{low}}^{D\pi,DD}$	5026.6 MeV	$B \to D^*\pi$ background PDF parameters in $D\pi$ fit
$\sigma_{D^*\pi\text{low,narrow}}^{D\pi,DD}$	15.1 MeV	
$\sigma_{D^*\pi\text{low,wide}}^{D\pi,DD}$	76.1 MeV	
$\mu_{D^*\pi\text{high}}^{D\pi,DD}$	5102.4 MeV	
$\sigma_{D^*\pi\text{high,narrow}}^{D\pi,DD}$	14.7 MeV	
$\sigma_{D^*\pi\text{high,wide}}^{D\pi,DD}$	56.6 MeV	
$\mu_{D^*\rho}^{D\pi,DD}$	4848.3 MeV	$B \to D^*\rho$ background PDF parameters in $D\pi$ fit
$\sigma_{D^*\rho}^{D\pi,DD}$	73.3 MeV	
$(m_0)_{D\rho}^{D\pi,DD}$	5159.4 MeV	$B \to D\rho$ background PDF parameters in $D\pi$ fit
$c_{D\rho}^{D\pi,DD}$	-20.8	
$p_{D\rho}^{D\pi,DD}$	0.5	
$\sigma_{D\rho}^{D\pi,DD}$	1.6 MeV	

Table 5.4.: Constant parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected DD $B^\pm \to D\pi^\pm$ candidates. These values were calibrated in the previous analysis [141].

Constant Parameter	Value	Description
$\alpha_{\text{sig}}^{DK,DD}$	2.2	Signal PDF parameters in DK fit
$n_{\text{sig}}^{DK,DD}$	1.6	
$\alpha_{\text{xfeed}}^{DK,DD}$	0.5	Cross-feed PDF parameters in DK fit
$\beta_{\text{xfeed}}^{DK,DD}$	2.2	
$\mu_{\text{xfeed}}^{DK,DD}$	5324.8 MeV	
$l_{\text{xfeed}}^{DK,DD}$	0.9	
$n_{\text{xfeed}}^{DK,DD}$	2.1	
$\mu_{D^*K\text{low}}^{DK,DD}$	5029.3 MeV	$B \to D^*K$ background PDF parameters in DK fit
$\sigma_{D^*K\text{low,narrow}}^{DK,DD}$	16.0 MeV	
$\sigma_{D^*K\text{low,wide}}^{DK,DD}$	67.7 MeV	
$\mu_{D^*K\text{high}}^{DK,DD}$	5104.3 MeV	
$\sigma_{D^*K\text{high,narrow}}^{DK,DD}$	15.0 MeV	
$\sigma_{D^*K\text{high,wide}}^{DK,DD}$	56.0 MeV	
$\alpha_{D^*\pi\text{low}}^{DK,DD}$	2.4	$B \to D^*\pi$ background PDF parameters in DK fit
$\beta_{D^*\pi\text{low}}^{DK,DD}$	1.3	
$\mu_{D^*\pi\text{low}}^{DK,DD}$	5070.3 MeV	
$\sigma_{D^*\pi\text{low}}^{DK,DD}$	16.5 MeV	
$l_{D^*\pi\text{low}}^{DK,DD}$	5.0	
$n_{D^*\pi\text{low}}^{DK,DD}$	5.1	
$\alpha_{D^*\pi\text{high}}^{DK,DD}$	1.0	
$\beta_{D^*\pi\text{high}}^{DK,DD}$	3.3	
$\mu_{D^*\pi\text{high}}^{DK,DD}$	5140.2 MeV	
$\sigma_{D^*\pi\text{high}}^{DK,DD}$	24.9 MeV	
$l_{D^*\pi\text{high}}^{DK,DD}$	0.0018	
$n_{D^*\pi\text{high}}^{DK,DD}$	2.7	
$\mu_{D^*\rho}^{DK,DD}$	4838.8 MeV	$B \to D^*\rho$ background PDF parameters in DK fit
$\sigma_{D^*\rho}^{DK,DD}$	143.3 MeV	
$\mu_{D\rho}^{DK,DD}$	5122.8 MeV	$B \to D\rho$ background PDF parameters in DK fit
$\sigma_{D\rho}^{DK,DD}$	52.8 MeV	
$\tau_{D\rho}^{DK,DD}$	-3.5×10^{-3}	
$(m_0)_{B_s^0}^{DK,DD}$	5230.3 MeV	$B_s^0 \to DK^{*0}$ background PDF parameters in DK fit
$c_{B_s^0}^{DK,DD}$	-23.3	
$p_{B_s^0}^{DK,DD}$	0.3	
$\sigma_{B_s^0}^{DK,DD}$	16.3 MeV	
$(m_0)_{B^0}^{DK,DD}$	5147.5 MeV	$B^0 \to DK^{*0}$ background PDF parameters in DK fit
$c_{B^0}^{DK,DD}$	-21.6	
$p_{B^0}^{DK,DD}$	0.3	
$\sigma_{B^0}^{DK,DD}$	10.7 MeV	

Table 5.5.: Constant parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected DD $B^\pm \to DK^\pm$ candidates. These values were calibrated in the previous analysis [141].

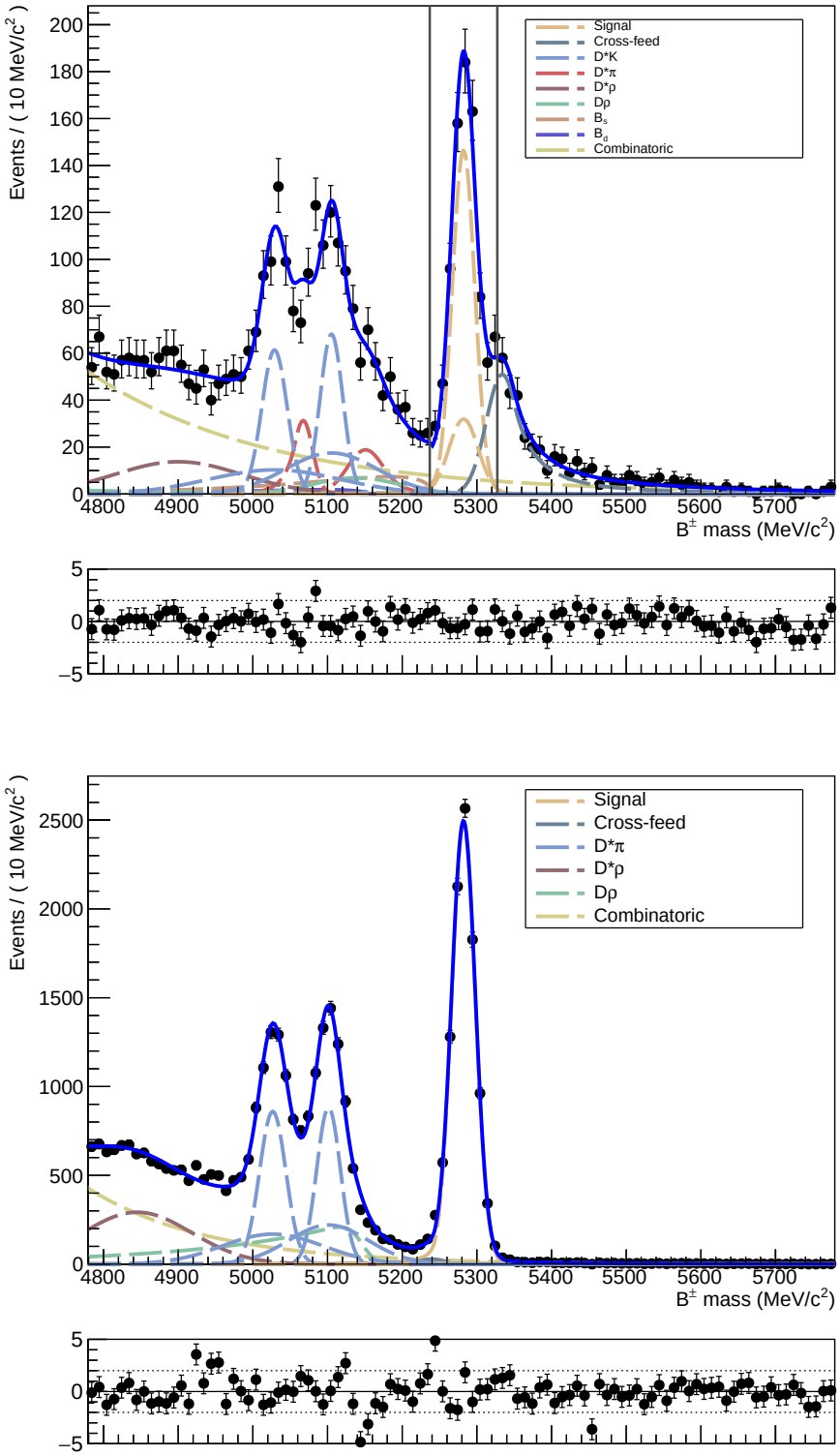


Figure 5.5.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for LL $B^\pm \to DK^\pm$ (top) and $B^\pm \to D\pi^\pm$ (bottom) candidates. In each plot, the lower canvas shows the pull.

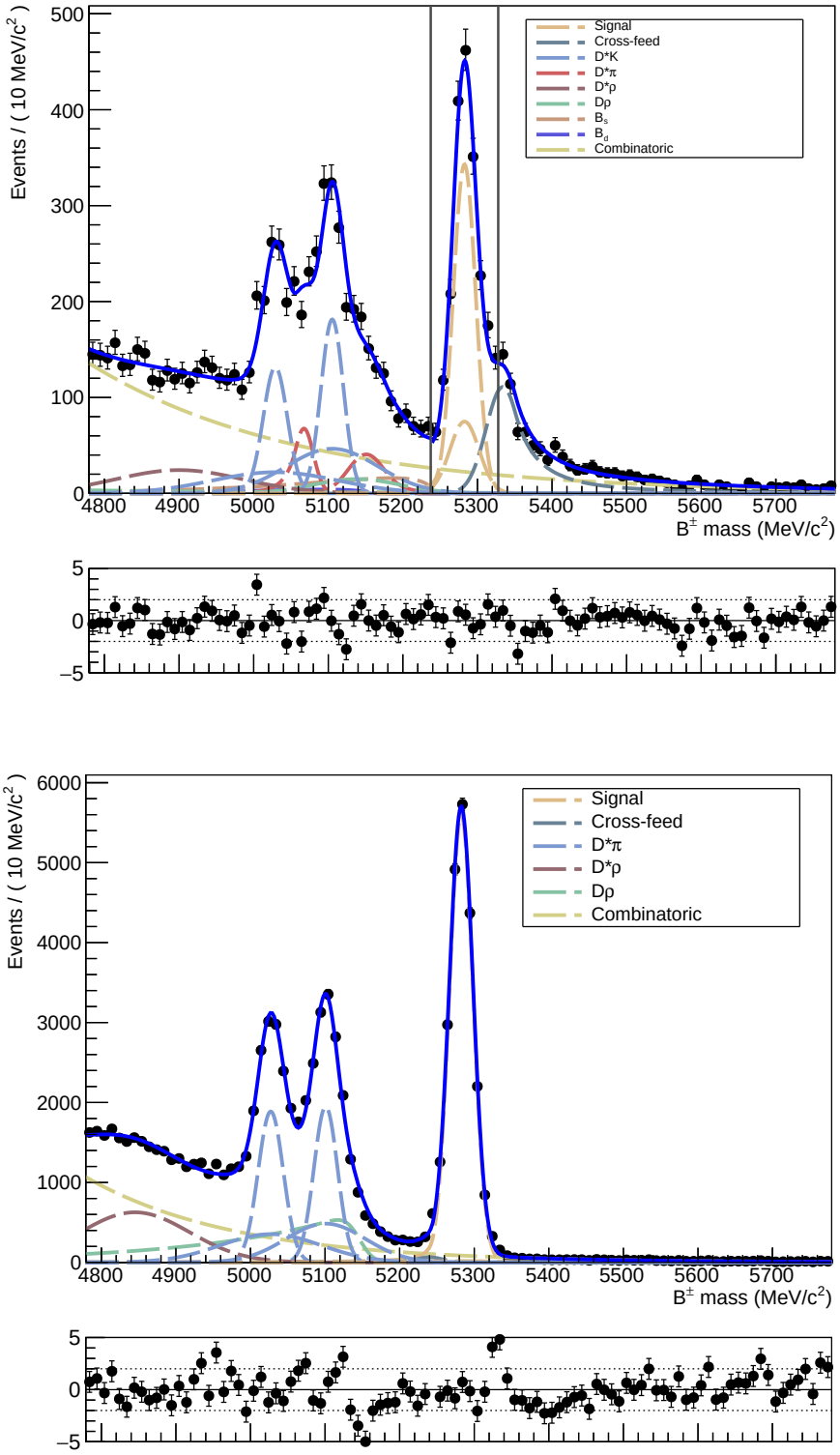


Figure 5.6.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for DD $B^\pm \to DK^\pm$ (top) and $B^\pm \to D\pi^\pm$ (bottom) candidates. In each plot, the lower canvas shows the pull.

Free parameter	Designation	Fitted Value	Description
$N_{\rm sig}^{D\pi,LL}$	`NSignalDpi`	10100^{+100}_{-100}	$B^\pm \to D\pi^\pm$ yield
r_N^{LL}	`Ratio`	$0.072^{+0.003}_{-0.003}$	Ratio of $(B^\pm \to DK^\pm)/(B^\pm \to D\pi^\pm)$ yields
$\mu_{\rm sig}^{LL}$	`GausSignalMeanDpi`	$5282.1^{+0.2}_{-0.2}$ MeV	Mean of signal peak
$\sigma_{\rm sig}^{D\pi,LL}$	`SignalCBsigmaDpi`	$15.8^{+0.1}_{-0.1}$ MeV	Width of signal peak in $B^\pm \to D\pi^\pm$ category
r_σ^{LL}	`widthratio_DK_DPi`	$0.953^{+0.049}_{-0.049}$ MeV	Ratio of $(B^\pm \to DK^\pm)/(B^\pm \to D\pi^\pm)$ widths
$\sigma_{\rm xfeed}^{DK,LL}$	`CBsigmaDK`	21^{+2}_{-2} MeV	Width of cross-feed peak in $B^\pm \to DK^\pm$ category
$N_{\rm comb}^{D\pi,LL}$	`Ncat1Dpi`	6900^{+400}_{-400}	Combinatoric yield in $B^\pm \to D\pi^\pm$ category
$\tau_{\rm comb}^{D\pi,LL}$	`Cat1ExpPowerDpi`	$-0.0062^{+0.0001}_{-0.0001}$	Combinatoric slope in $B^\pm \to D\pi^\pm$ category
$N_{\rm comb}^{DK,LL}$	`Ncat1DK`	1200^{+100}_{-100}	Combinatoric yield in $B^\pm \to DK^\pm$ category
$\tau_{\rm comb}^{DK,LL}$	`Cat1ExpPowerDK`	$-0.0040^{+0.0002}_{-0.0002}$	Combinatoric slope in $B^\pm \to DK^\pm$ category
$N_{D^*\pi,\rm low}^{D\pi,LL}$	`NDstpi1Dpi`	6700^{+200}_{-200}	Yield of lower-mass $B \to D^*\pi$ peak in $B^\pm \to D\pi^\pm$ category
$N_{D^*\pi,\rm high}^{D\pi,LL}$	`NDstpi2Dpi`	6400^{+200}_{-200}	Yield of higher-mass $B \to D^*\pi$ peak in $B^\pm \to D\pi^\pm$ category
$\mu_{D^*\pi,\rm low}^{D\pi,LL}$	`GausmeanDstpi1Dpi*`	$5026.7^{+0.5}_{-0.5}$ MeV	Mean of lower-mass $B \to D^*\pi$ peak in $B^\pm \to D\pi^\pm$ category
$\mu_{D^*\pi,\rm high}^{D\pi,LL}$	`GausmeanDstpi2Dpi*`	$5101.4^{+0.5}_{-0.5}$ MeV	Mean of higher-mass $B \to D^*\pi$ peak in $B^\pm \to D\pi^\pm$ category
$\sigma_{D^*\pi,\rm low,narrow}^{D\pi,LL}$	`GaussigmaDstpi1Dpi*`	$17.9^{+0.6}_{-0.6}$ MeV	Width of lower-mass $B \to D^*\pi$ peak in $B^\pm \to D\pi^\pm$ category
$\sigma_{D^*\pi,\rm high,narrow}^{D\pi,LL}$	`GaussigmaDstpi2Dpi*`	$15.7^{+0.5}_{-0.5}$ MeV	Width of higher-mass $B \to D^*\pi$ peak in $B^\pm \to D\pi^\pm$ category
$N_{D^*\pi}^{DK,LL}$	`NDstpiDK`	290^{+30}_{-20}	Yield of $B \to D^*\pi$ in $B^\pm \to DK^\pm$ category
$N_{D^*K,\rm low}^{DK,LL}$	`NDstK1DK`	360^{+30}_{-30}	Yield of lower-mass $B \to D^*K$ peak in $B^\pm \to DK^\pm$ category
$N_{D^*K,\rm high}^{DK,LL}$	`NDstK2DK`	480^{+40}_{-40}	Yield of higher-mass $B \to D^*K$ peak in $B^\pm \to DK^\pm$ category
$N_{D^*\rho}^{D\pi,LL}$	`NGausDstRhoDpi`	4300^{+400}_{-400}	Yield of $B \to D^*\rho$ in $B^\pm \to D\pi^\pm$ category
$N_{D^*\rho}^{DK,LL}$	`NGausDstRhoDK`	360^{+80}_{-80}	Yield of $B \to D^*\rho$ in $B^\pm \to DK^\pm$ category
$N_{D\rho}^{D\pi,LL}$	`NArgusDrhoDpi`	4200^{+500}_{-500}	Yield of $B \to D\rho$ in $B^\pm \to D\pi^\pm$ category

Table 5.6.: Free parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected LL $B^\pm \to Dh^\pm$ candidates. The designations assigned within the fitting program are listed, to allow the reader to interpret the corresponding plot of correlations, Figure 5.7. Variables marked with an asterisk (*) are subject to Gaussian constraints in the fit.

Free parameter	Designation	Fitted Value	Description
$N_{\text{sig}}^{D\pi,DD}$	`NSignalDpi`	23000^{+200}_{-200}	$B^{\pm} \to D\pi^{\pm}$ yield
r_N^{tt}	`Ratio`	$0.074^{+0.002}_{-0.002}$	Ratio of $(B^{\pm} \to DK^{\pm})/(B^{\pm} \to D\pi^{\pm})$ yields
μ_{sig}^{DD}	`GausSignalMeanDpi`	$5282.6^{+0.1}_{-0.1}$ MeV	Mean of signal peak
$\sigma_{\text{sig}}^{D\pi,DD}$	`SignalCBsigmaDpi`	$15.8^{+0.1}_{-0.1}$ MeV	Width of signal peak in $B^{\pm} \to D\pi^{\pm}$ category
r_{σ}^{DD}	`widthratio_DK_DPi`	$0.954^{+0.033}_{-0.032}$ MeV	Ratio of $(B^{\pm} \to DK^{\pm})/(B^{\pm} \to D\pi^{\pm})$ widths
$\sigma_{\text{xfeed}}^{DK,DD}$	`CBsigmaDK`	22^{+2}_{-2} MeV	Width of cross-feed peak in $B^{\pm} \to DK^{\pm}$ category
$N_{\text{comb}}^{D\pi,DD}$	`Ncat1Dpi`	21200^{+800}_{-700}	Combinatoric yield in $B^{\pm} \to D\pi^{\pm}$ category
$\tau_{\text{comb}}^{D\pi,DD}$	`Cat1ExpPowerDpi`	$-0.00502^{+0.00008}_{-0.00008}$	Combinatoric slope in $B^{\pm} \to D\pi^{\pm}$ category
$N_{\text{comb}}^{DK,DD}$	`Ncat1DK`	3600^{+200}_{-200}	Combinatoric yield in $B^{\pm} \to DK^{\pm}$ category
$\tau_{\text{comb}}^{DK,DD}$	`Cat1ExpPowerDK`	$-0.0035^{+0.0001}_{-0.0001}$	Combinatoric slope in $B^{\pm} \to DK^{\pm}$ category
$N_{D^*\pi\text{low}}^{D\pi,DD}$	`NDstpi1Dpi`	13900^{+300}_{-300}	Yield of lower-mass $B \to D^*\pi$ peak in $B^{\pm} \to D\pi^{\pm}$ category
$N_{D^*\pi\text{high}}^{D\pi,DD}$	`NDstpi2Dpi`	14100^{+400}_{-400}	Yield of higher-mass $B \to D^*\pi$ peak in $B^{\pm} \to D\pi^{\pm}$ category
$\mu_{D^*\pi\text{low}}^{D\pi,DD}$	`GausmeanDstpi1Dpi*`	$5027.2^{+0.3}_{-0.3}$ MeV	Mean of lower-mass $B \to D^*\pi$ peak in $B^{\pm} \to D\pi^{\pm}$ category
$\mu_{D^*\pi\text{high}}^{D\pi,DD}$	`GausmeanDstpi2Dpi*`	$5100.9^{+0.3}_{-0.3}$ MeV	Mean of higher-mass $B \to D^*\pi$ peak in $B^{\pm} \to D\pi^{\pm}$ category
$\sigma_{D^*\pi\text{low,narrow}}^{D\pi,DD}$	`GaussigmaDstpi1Dpi*`	$16.8^{+0.4}_{-0.4}$ MeV	Width of lower-mass $B \to D^*\pi$ peak in $B^{\pm} \to D\pi^{\pm}$ category
$\sigma_{D^*\pi\text{high,narrow}}^{D\pi,DD}$	`GaussigmaDstpi2Dpi*`	$15.7^{+0.3}_{-0.3}$ MeV	Width of higher-mass $B \to D^*\pi$ peak in $B^{\pm} \to D\pi^{\pm}$ category
$N_{D^*\pi}^{DK,DD}$	`NDstpiDK`	620^{+40}_{-40}	Yield of $B \to D^*\pi$ in $B^{\pm} \to DK^{\pm}$ category
$N_{D^*K\text{low}}^{DK,DD}$	`NDstK1DK`	720^{+50}_{-50}	Yield of lower-mass $B \to D^*K$ peak in $B^{\pm} \to DK^{\pm}$ category
$N_{D^*K\text{high}}^{DK,DD}$	`NDstK2DK`	1300^{+70}_{-70}	Yield of higher-mass $B \to D^*K$ peak in $B^{\pm} \to DK^{\pm}$ category
$N_{D^*\rho}^{D\pi,DD}$	`NGausDstRhoDpi`	9300^{+600}_{-600}	Yield of $B \to D^*\rho$ in $B^{\pm} \to D\pi^{\pm}$ category
$N_{D^*\rho}^{DK,DD}$	`NGausDstRhoDK`	700^{+100}_{-100}	Yield of $B \to D^*\rho$ in $B^{\pm} \to DK^{\pm}$ category
$N_{D\rho}^{D\pi,DD}$	`NArgusDrhoDpi`	10900^{+800}_{-800}	Yield of $B \to D\rho$ in $B^{\pm} \to D\pi^{\pm}$ category

Table 5.7.: Free parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected DD $B^{\pm} \to Dh^{\pm}$ candidates. The designations assigned within the fitting program are listed, to allow the reader to interpret the corresponding plot of correlations, Figure 5.7. Variables marked with an asterisk (*) are subject to Gaussian constraints in the fit.

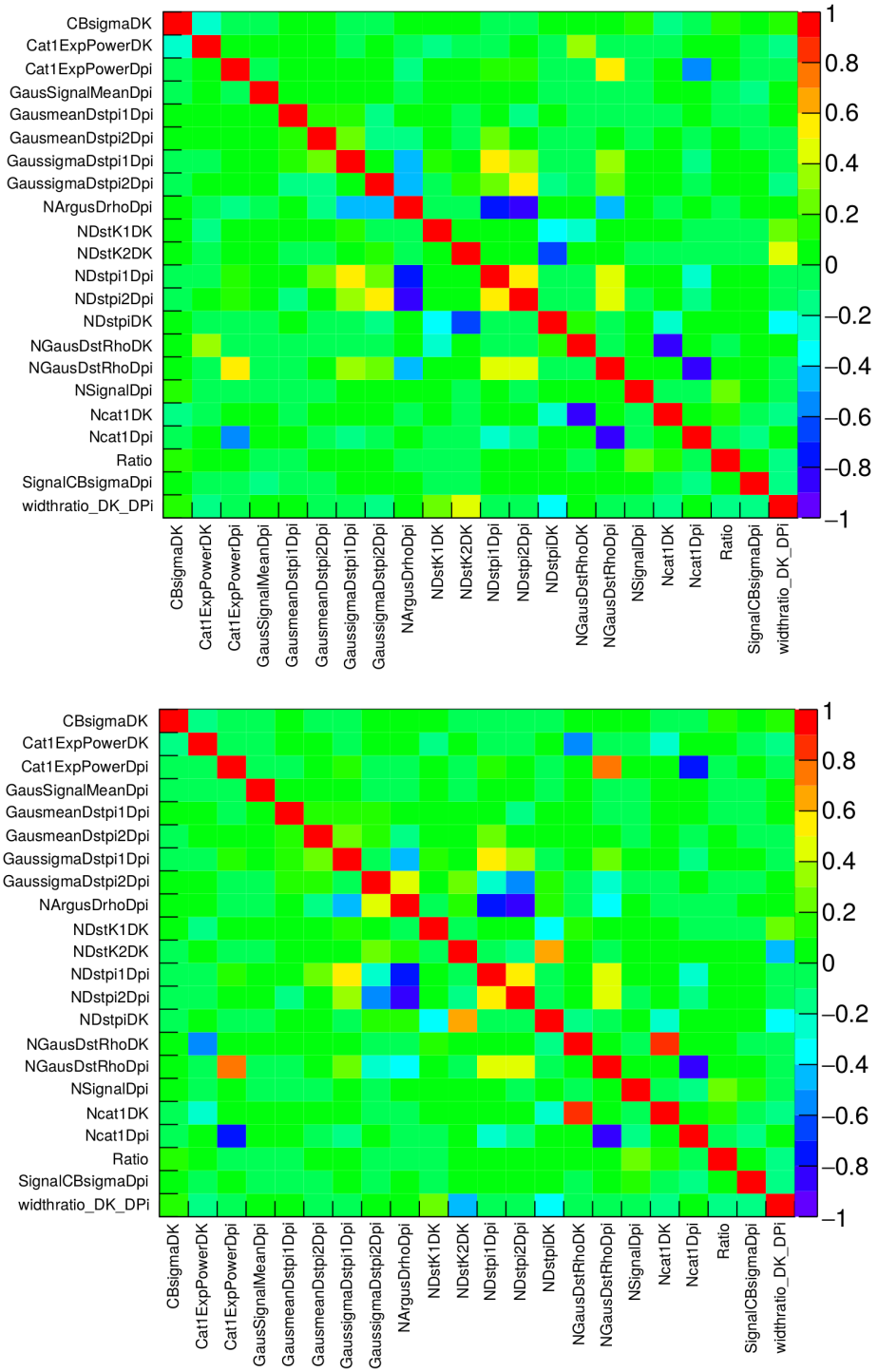


Figure 5.7.: Correlation matrices obtained for parameters in the simultaneous B-mass fit to LL (top) and DD (bottom) $B^{\pm} \to DK^{\pm}$ and $B^{\pm} \to D\pi^{\pm}$ candidates.

Component	Yield	
	LL	DD
signal	705 ± 7	1680 ± 10
cross-feed	114 ± 1	260 ± 2
D^*K	1.9 ± 0.4	7.6 ± 0.5
$D^*\pi$	0.5 ± 0.4	3.0 ± 0.5
$D^*\rho$	0.010 ± 0.006	0.027 ± 0.009
$D\rho$	3 ± 2	18 ± 3
B_s	1.2 ± 0.9	8 ± 1
B_d	negligible	negligible
combinatoric	60 ± 5	230 ± 10

Table 5.8.: Yields extracted in the signal region for mode $B^\pm \to DK^\pm$.

5.2. Fit to selected $B^0 \to DK^{*0}$ candidates

A single simultaneous fit is performed to selected $B^0 \to DK^{*0}$ candidates, which are split into two categories depending on the track type of the reconstructed K^0_S candidate. The PDFs used in this fit are described below. Several of these PDFs are obtained using Monte Carlo simulated samples, generated centrally by the LHCb collaboration following the procedure outlined in Section 3.2.8. The samples used for this fit are listed in Table 5.9.

Mode	No. of candidates generated
$B^0_s \to DK^{*0}$	25.0×10^6
$B^0_s \to D^*K^{*0}, D^* \to D\gamma$ (001)	33.3×10^6
$B^0_s \to D^*K^{*0}, D^* \to D\gamma$ (010)	33.3×10^6
$B^0_s \to D^*K^{*0}, D^* \to D\pi^0$ (001)	33.3×10^6
$B^0_s \to D^*K^{*0}, D^* \to D\pi^0$ (010)	33.3×10^6

Table 5.9.: Simulated Monte Carlo samples used to extract shapes for the reconstructed B-mass fit of selected $B^0 \to DK^{*0}$ candidates. All samples are generated in 2012 conditions using $\sqrt{s} = 8$ TeV. Labels (001) and (010) for the $B^0_s \to D^*K^{*0}$ samples refer to the simulated helicity combinations in the scalar to vector-vector decay, as described on page 172.

$B^0_{(s)}$ signal

'Double Crystal Ball' shapes are used to describe the signal from B^0 and B^0_s decays. Since the two decays are kinematically identical, the same PDF shape is used for both cases, with the yields of the two components floated separately in the fit. The 'double Crystal Ball' shape here refers to the sum of two Crystal Ball [143] functions with a single common mean, and with their width parameters, σ fixed to be equal and opposite. This means that one of the component crystal ball shapes has its exponential tail on the high-mass side, and the other has its exponential tail on the low mass side. The combined shape thus consists of a sum of two Gaussian cores, with independently parameterized exponential tails on either side.

The tail parameters of the two Crystal Ball shapes, along with the constant of proportionality between their relative magnitudes, are fixed to values extracted from a fit to simulated B^0_s signal candidates, as shown in Figure 5.8 (a) and Figure 5.9 (a) for LL and DD samples, respectively.

The full PDF used to describe $B^0_{(s)} \to D^*K^*$ candidates in the mass fit may then be expressed as

$$F^{tt}_{\text{sig}}(m; N^{tt}_{B^0_s}, r_{B^0}, \mu, k_\sigma) \equiv F^{tt}_{B^0_s}(m; N^{tt}_{B^0_s}, \mu, k_\sigma) + F^{tt}_{B^0}(m; r_{B^0}, N^{tt}_{B^0_s}, \mu, k_\sigma), \tag{5.42}$$

for

$$F^{tt}_{B^0_s}(m; N^{tt}_{B^0_s}, \mu, k_\sigma) \equiv N^{tt}_{B^0_s} \cdot \big(f^{tt} \cdot \mathcal{CB}(m; \mu, \sigma^{tt}, \alpha^{tt}_1, n^{tt}_1) + (1 - f^{tt}) \cdot \mathcal{CB}(m; \mu, -\sigma^{tt}, \alpha^{tt}_2, n^{tt}_2)\big), \quad \text{and} \tag{5.43}$$

$$F^{tt}_{B^0}(m; r_{B^0}, N^{tt}_{B^0_s}, \mu, k_\sigma) \equiv r_{B^0} N^{tt}_{B^0_s} \cdot \big(f^{tt} \cdot \mathcal{CB}(m; \mu - \delta m, \sigma^{tt}, \alpha^{tt}_1, n^{tt}_1) + (1 - f^{tt}) \cdot \mathcal{CB}(m; \mu - \delta m, -\sigma^{tt}, \alpha^{tt}_2, n^{tt}_2)\big), \tag{5.44}$$

where:

- $N^{tt}_{B^0_s}$ gives the yield of B^0_s candidates of track type tt;
- $r_{B^0} \equiv N^{tt}_{B^0} / N^{tt}_{B^0_s}$ gives the ratio of the yields of B^0 and B^0_s candidates, and is taken to be common for both track types;
- $\mathcal{CB}$ denotes a Crystal Ball function;
- μ is the mean of the B^0_s peak, and is taken to be common for both track types;
- δm is the mass difference between the B^0 and the B^0_s mesons. It is fixed to the PDG value $\delta m = 87.35\,\text{MeV}$ [3];
- σ^{tt} is the common width of the B^0 and B^0_s peaks. It is parameterized as

  $$\sigma^{tt} = k_\sigma \sigma^{tt}_{\text{MC}}, \tag{5.45}$$

  such that k_σ gives the smearing factor of the distribution observed in data relative to the value obtained from the signal MC fit, $\sigma^{\text{LL}}_{\text{MC}} = 13.5\,\text{MeV}$ and $\sigma^{\text{DD}}_{\text{MC}} = 13.6\,\text{MeV}$. A single smearing factor is used for both track types;
- $\alpha^{tt}_{\{1,2\}}$ parameterize the distance (in σ) from the mean of each Gaussian core at which the exponential tail begins, for each component Crystal Ball function. They are is fixed to values obtained from the signal MC fits; $\alpha^{\text{LL}}_1 = 1.44$, $\alpha^{\text{LL}}_2 = 1.07$, $\alpha^{\text{DD}}_1 = 0.565$ and $\alpha^{\text{DD}}_2 = 1.67$. The same values are used for B^0 and B^0_s;
- $n^{tt}_{\{1,2\}}$ denote the magnitudes of the exponential slopes of the two component Crystal Ball tails. They are fixed to values obtained from the signal MC fit; $n^{\text{LL}}_1 = 1.89$, $n^{\text{LL}}_2 = 3.30$, $n^{\text{DD}}_1 = 3.07$ and $n^{\text{DD}}_2 = 2.52$. The same values are used for B^0 and B^0_s; and
- f^{tt} is the fraction of the double Crystal Ball shape which is given by the Crystal Ball component with its exponential tail on the low mass side. It is fixed to the values obtained from the signal MC fits; $f^{\text{LL}} = 0.683$, and $f^{\text{DD}} = 0.192$. The same value is used for B^0 and B^0_s.

Partially-reconstructed $B^0_{(s)} \to D^* K^{*0}$ background

An important source of background arises from partially reconstructed $B^0_{(s)} \to D^* K^{*0}$ decays, with $D^* \to D X^0$, where $X^0 = \gamma, \pi^0$. Such candidates may pass the selection requirements for this mode if they are reconstructed without the associated neutral final state particle, X^0. The missing final state

particle results in a shape in the reconstructed B-mass spectrum that is shifted towards lower masses, and significantly broader than the $B^0_{(s)}$ peaks.

Since $B^0_{(s)} \to D^{*0}K^{*0}$ is a decay of a pseudo scalar meson to two vector mesons, conservation of angular momentum implies that the $B^0_{(s)}$ daughters are produced in one of three helicity configurations, each with an associated amplitude:

- $\mathcal{A}_{100}$, the amplitude for D^{*0} to be produced in helicity state $+1$, and the K^{*0} in helicity state -1;
- $\mathcal{A}_{010}$, the amplitude for D^{*0} to be produced in helicity state 0, and the K^{*0} in helicity state 0; and
- $\mathcal{A}_{001}$, the amplitude for D^{*0} to be produced in helicity state -1, and the K^{*0} in helicity state $+1$.

It is important to note that the relative proportions of these helicity components cannot be known a priori.

The B^0-candidate mass reconstructed by combining the D and K^{*0} candidates for this mode depends on the fraction of energy carried away by the neutral X^0, which may be parameterized by θ', the angle between the D^0 momentum in the D^{*0} rest frame and the D^{*0} momentum in the $B^0_{(s)}$ rest frame. The θ' distribution can be expressed for $D^{*0} \to D^0\pi^0$ decays as [145, 146]

$$I(\theta') \propto \frac{1}{2}\left(|\mathcal{A}_{001}|^2 + |\mathcal{A}_{100}|^2\right)\sin^2\theta' + |\mathcal{A}_{010}|^2\cos^2\theta', \tag{5.46}$$

and for $D^{*0} \to D^0\gamma$ decays as

$$I(\theta') \propto \frac{1}{2}\left(|\mathcal{A}_{001}|^2 + |\mathcal{A}_{100}|^2\right)(1+\cos^2\theta') + |\mathcal{A}_{010}|^2\sin^2\theta'. \tag{5.47}$$

Since in each case the 001 and 100 configurations follow the same angular dependence, no distinction can be made between these two contributions in the DK^{*0} mass spectrum. Therefore, only the 001 configuration is simulated to obtain the shape describing both 001 and 100 configurations. PDF shapes to describe the B-mass spectrum of these components are obtained directly from selected $B^0_s \to D^*K^*$ Monte Carlo, using the kernel estimation functionality provided by the `RooFit` software library. The shapes obtained are displayed in subplots (b), (c), (d) and (e) of Figure 5.8 and Figure 5.9 for LL and DD samples respectively.

To allow for the possibility of differences in the position and width of the kernel-estimated shapes obtained from simulation, relative to the true shapes that would be observed in data, a corrective Gaussian smearing and mass shift are applied to the simulated sample before the final kernel-estimated shapes are constructed. These are determined from the observed data-simulation differences in the width and central value of the B^0_s peak:

- The fits to simulated B^0_s decays yield values for the B^0_s mean of $\mu^{\text{LL}}_{\text{MC}} = 5367.3\,\text{MeV}$ and $\mu^{\text{DD}}_{\text{MC}} = 5367.2\,\text{MeV}$, whilst a previous fit to selected data returns the value $\mu_{\text{data}} = 5369.4\,\text{MeV}$. Therefore the $B^0_{(s)} \to D^*K^*$ shapes are shifted by $2.1(2.2)\,\text{MeV}$ in the final fit to the LL(DD) sample.

- The returned value for the smearing factor, k_σ, introduced in the signal PDF shape, parameterizes the relative difference in the signal PDF widths observed in data and simulation, and is taken as the appropriate width for the Gaussian smearing.

Finally, the different partially-reconstructed $B^0_{(s)} \to D^*K^*$ components are combined according to their known branching fractions [147], and their efficiencies in the full selection. In addition, the ratios of the yields of corresponding partially reconstructed B^0 and B^0_s components are fixed to match the ratio of the signal and control mode yields. The function describing this partially reconstructed background may then be expressed as

$$F^{tt}_{\text{part}}(m; N^{s\ tt}_{010}, r_{101}) = F^{s\ tt}_{\text{part}}(m; N^{s\ tt}_{010}, r_{101}) + F^{d\ tt}_{\text{part}}(m; N^{s\ tt}_{010}, r_{101}), \tag{5.48}$$

with

$$\begin{aligned} F^{s\ tt}_{\text{part}}(m; N^{s\ tt}_{010}, r_{101}) =& N^{s\ tt}_{010} \cdot \left(G^{tt}_{010} \cdot f^{tt}_{\gamma 010}(m) + (1 - G^{tt}_{010}) \cdot f^{tt}_{\pi^0 010}(m)\right) \\ &+ r_{101} N^{s\ tt}_{010} \cdot \left(G^{tt}_{101} \cdot f^{tt}_{\gamma 101}(m) + (1 - G^{tt}_{101}) \cdot f^{tt}_{\pi^0 101}(m)\right), \end{aligned} \tag{5.49}$$

$$\begin{aligned} F^{d\ tt}_{\text{part}}(m; N^{s\ tt}_{010}, r_{101}) =& r_{B^0} N^{s\ tt}_{010} \cdot \left(G^{tt}_{010} \cdot f^{tt}_{\gamma 010}(m - \delta m) + (1 - G^{tt}_{010}) \cdot f^{tt}_{\pi^0 010}(m - \delta m)\right) \\ &+ r_{101} r_{B^0} N^{s\ tt}_{010} \cdot \left(G^{tt}_{101} \cdot f^{tt}_{\gamma 101}(m - \delta m) + (1 - G^{tt}_{101}) \cdot f^{tt}_{\pi^0 101}(m - \delta m)\right), \end{aligned} \tag{5.50}$$

where:

- $N^{s\ tt}_{010}$ are the $B^0_s \to D^{*0}K^{*0}$ yields for helicity component 010, and r_{101} gives the relative yield for helicity component 101 as a fraction of this value, such that $r_{101} = N^{s\ tt}_{101}/N^{s\ tt}_{010}$ (here 101 designates the combination $100 + 001$, since these components are indistinguishable in the reconstructed B-mass spectrum). r_{101} is taken to be common for both track types;

- $f^{tt}_{\gamma H}$ and $f^{tt}_{\pi^0 H}$ denote the kernel-estimated PDFs for helicity component H, extracted from simulated $B^0_s \to D^{*0}K^{*0}$ samples; and

- G^{tt}_H give the proportionality of the $D^* \to D\gamma$ contribution relative to the $D^* \to D\pi^0$ contribution, for helicity component H. They are fixed from the measured branching fractions [147] and the relative efficiencies observed for the two components in Monte Carlo simulation, and are defined as

$$G^{tt}_H \equiv \frac{g^{tt}_H}{g^{tt}_H + p^{tt}_H}, \tag{5.51}$$

where

$$g^{tt}_H = \mathcal{B}(D\gamma) \cdot \varepsilon^{H\ tt}_{\text{gen}}(D\gamma) \cdot \varepsilon^{H\ tt}_{\text{strip}}(D\gamma) \cdot \varepsilon^{H\ tt}_{\text{sel}}(D\gamma), \quad \text{and} \tag{5.52}$$

$$p^{tt}_H = \mathcal{B}(D\pi^0) \cdot \varepsilon^{H\ tt}_{\text{gen}}(D\pi^0) \cdot \varepsilon^{H\ tt}_{\text{strip}}(D\pi^0) \cdot \varepsilon^{H\ tt}_{\text{sel}}(D\pi^0). \tag{5.53}$$

The Monte Carlo generator efficiency, ε_{gen}, includes the effects of all detector acceptance requirements applied. $\varepsilon_{\text{strip}}$ is the efficiency of the stripping selection, and ε_{sel} gives the efficiency of the full offline selection. The values of these parameters are summarised in Table 5.10.

Quantity	$D^{*0} \to D^0\pi^0$		$D^{*0} \to D^0\gamma$	
	LL	DD	LL	DD
Branching fraction	$(65.5 \pm 0.9)\%$ [147]		$(34.5 \pm 0.9)\%$ [147]	
$\varepsilon^{010}_{\text{gen}}$	$(6.32 \pm 0.06)\%$		$(6.36 \pm 0.02)\%$	
$\varepsilon^{101}_{\text{gen}}$	$(6.72 \pm 0.08)\%$		$(6.76 \pm 0.02)\%$	
$\varepsilon^{010}_{\text{strip}}$	$(1.177 \pm 0.002)\%$		$(1.177 \pm 0.002)\%$	
$\varepsilon^{101}_{\text{strip}}$	$(1.158 \pm 0.002)\%$		$(1.144 \pm 0.002)\%$	
$\varepsilon^{010}_{\text{sel}}$	$(2.37 \pm 0.01)\%$	$(6.87 \pm 0.02)\%$	$(2.28 \pm 0.01)\%$	$(6.58 \pm 0.02)\%$
$\varepsilon^{101}_{\text{sel}}$	$(1.18 \pm 0.01)\%$	$(3.35 \pm 0.02)\%$	$(1.12 \pm 0.01)\%$	$(3.20 \pm 0.02)\%$

Table 5.10.: Known branching fractions, and measured efficiencies obtained in Monte Carlo simulation, for $B^0_s \to D^{*0}K^{*0}$ partially reconstructed background components.

Combinatoric background

The combinatoric background component is modelled by a decreasing exponential distribution with slope parameter τ,

$$F^{tt}_{\text{comb}}(m; N^{tt}_{\text{comb}}, \tau^{tt}) = N^{tt}_{\text{comb}} \cdot \mathrm{e}^{\tau^{tt} m} |\tau^{tt}|. \tag{5.54}$$

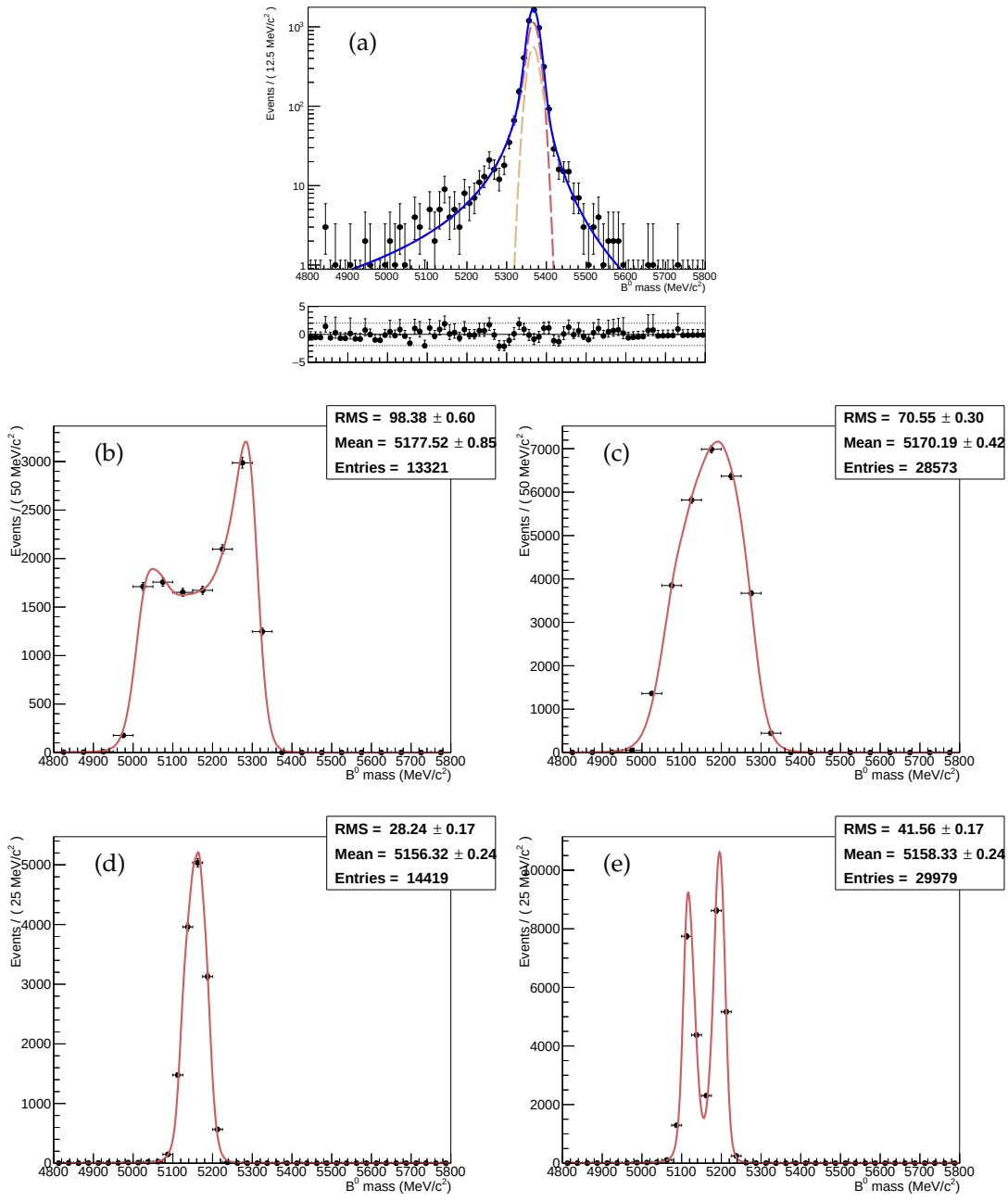


Figure 5.8.: *B*-mass PDF shapes used in the fit to LL $B^0 \to DK^{*0}$ candidates. A double Crystal Ball function is fitted to (a) B_s^0 signal Monte Carlo samples. Kernel-estimated shapes are obtained from $B_s^0 \to D^* K^{*0}$ Monte Carlo samples, for (b) $D^* \to D\gamma_{001}$, (c) $D^* \to D\gamma_{010}$, (d) $D^* \to D\pi^0_{001}$ and (e) $D^* \to D\pi^0_{010}$ candidates. For plot (a), histogram points show Monte Carlo simulation, a solid blue line gives the fitted PDF, and the two constituent Crystal Ball functions of the fitted PDF are shown as dashed red and yellow lines. The lower canvas of (a) shows the pull.

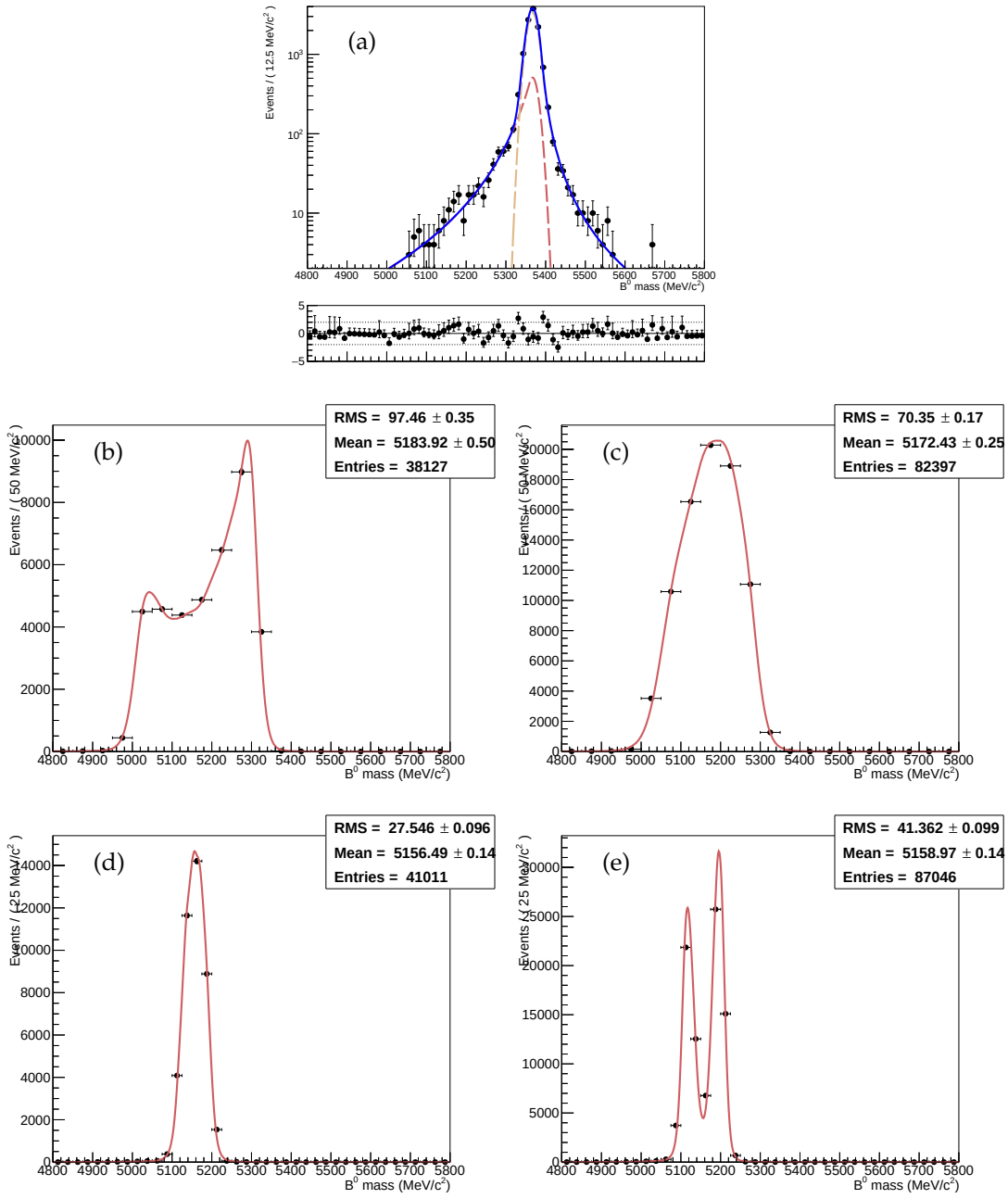


Figure 5.9.: B-mass PDF shapes used in the fit to DD $B^0 \to DK^{*0}$ candidates. A double Crystal Ball function is fitted to (a) B_s^0 signal Monte Carlo samples. Kernel-estimated shapes are obtained from $B_s^0 \to D^* K^{*0}$ Monte Carlo samples, for (b) $D^* \to D\gamma_{001}$, (c) $D^* \to D\gamma_{010}$, (d) $D^* \to D\pi^0_{001}$ and (e) $D^* \to D\pi^0_{010}$ candidates. For plot (a), histogram points show Monte Carlo simulation, a solid blue line gives the fitted PDF, and the two constituent Crystal Ball functions of the fitted PDF are shown as dashed red and yellow lines. The lower canvas of (a) shows the pull.

5.2.1. Mass fit result

In summary, the model used to fit the reconstructed B-mass spectrum for selected data of each track type is comprised of 11 components. Double Crystal Ball shapes are used to describe signal arising from B^0 and B_s^0 decays. For background arising from partially-reconstructed $B^0_{(s)} \to D^* K^{*0}$ decays, four kernel-estimated PDF shapes are used for B^0 decays and four for B_s^0, to describe the two distinguishable helicity distributions and the two possible decay modes $D^* \to D\gamma$ and $D^* \to D\pi^0$. The four PDF shapes obtained for $B^0 \to D^* K^{*0}$ background are combined according to their branching fractions and measured efficiencies in MC, and likewise for the four PDF shapes obtained for $B_s^0 \to D^* K^{*0}$ background. The combinatoric background component is described using an exponential.

The full PDF model used to fit each track type category may be written, in terms of its free parameters, as

$$\begin{aligned} F^{tt}_{\text{tot}} =& F^{tt}_{\text{sig}}(m; N^{tt}_{B^0_s}, r_{B^0}, \mu, k_\sigma) + \\ & F^{tt}_{\text{part}}(m; N^{s\ tt}_{010}, r_{101}) + F^{tt}_{\text{comb}}(m; N^{tt}_{\text{comb}}, \tau^{tt}). \end{aligned} \tag{5.55}$$

All parameters are taken to be independent for samples of different track type, except for:

- the central value of the B_s^0 peak, μ;
- the ratio of B^0 and B_s^0 yields, r_{B^0};
- the smearing factor for the Crystal Ball widths in data relative to simulation, k_σ; and
- the ratio of helicities in $B^0_{(s)} \to D^* K^{*0}$ background, r_{101}.

The constant parameters of the fit, and their fixed values, are summarized in Table 5.11. The full simultaneous fit to LL and DD candidates, performed on the 2011 and 2012 LHCb dataset, is displayed in Figure 5.10. The values returned for the free parameters are summarized in Table 5.12, and their obtained correlations are illustrated in Figure 5.11. Only candidates falling within a $\pm 3\sigma$ window centred on the fitted B^0 mean, $\mu - \delta m$, are propagated forward to the full $\mathcal{CP}$ fit.[7] This window is marked by vertical lines in Figure 5.10. For both LL and DD candidates, the window used is $[5245; 5319]$ MeV. Component yields extracted within this window are enumerated in Table 5.13.

[7] Note that, whilst kinematically identical to B^0 signal, B_s^0 candidates do not depend on the same CKM matrix elements, and give only very weak dependence on the $\mathcal{CP}$ phase γ.

Constant Parameter	Value	Description
δm	87.35 MeV	Mass shift between $B^0_{(s)}$ peaks
$\sigma^{\mathrm{LL}}_{\mathrm{MC}}$	13.5 MeV	Double Crystal Ball parameters for signal PDFs
$\sigma^{\mathrm{DD}}_{\mathrm{MC}}$	13.6 MeV	
α^{LL}_1	1.44	
α^{LL}_2	1.07	
α^{DD}_1	0.565	
α^{DD}_2	1.67	
n^{LL}_1	1.89	
n^{LL}_2	3.30	
n^{DD}_1	3.07	
n^{DD}_2	2.52	
f^{LL}	0.683	
f^{DD}	0.192	
$\delta m^{\mathrm{LL}}_{D\rho}$	44.1 MeV	Mass shift between $B^0 \to D\rho^0$ and B^0_s peaks
$\delta m^{\mathrm{DD}}_{D\rho}$	38.1 MeV	
G^{LL}_{010}	0.374	Proportionality constants for $B^0 \to D^* K^{*0}$ backgrounds
G^{DD}_{010}	0.372	
G^{LL}_{101}	0.367	
G^{DD}_{101}	0.369	

Table 5.11.: Constant parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected $B^0 \to DK^{*0}$ candidates.

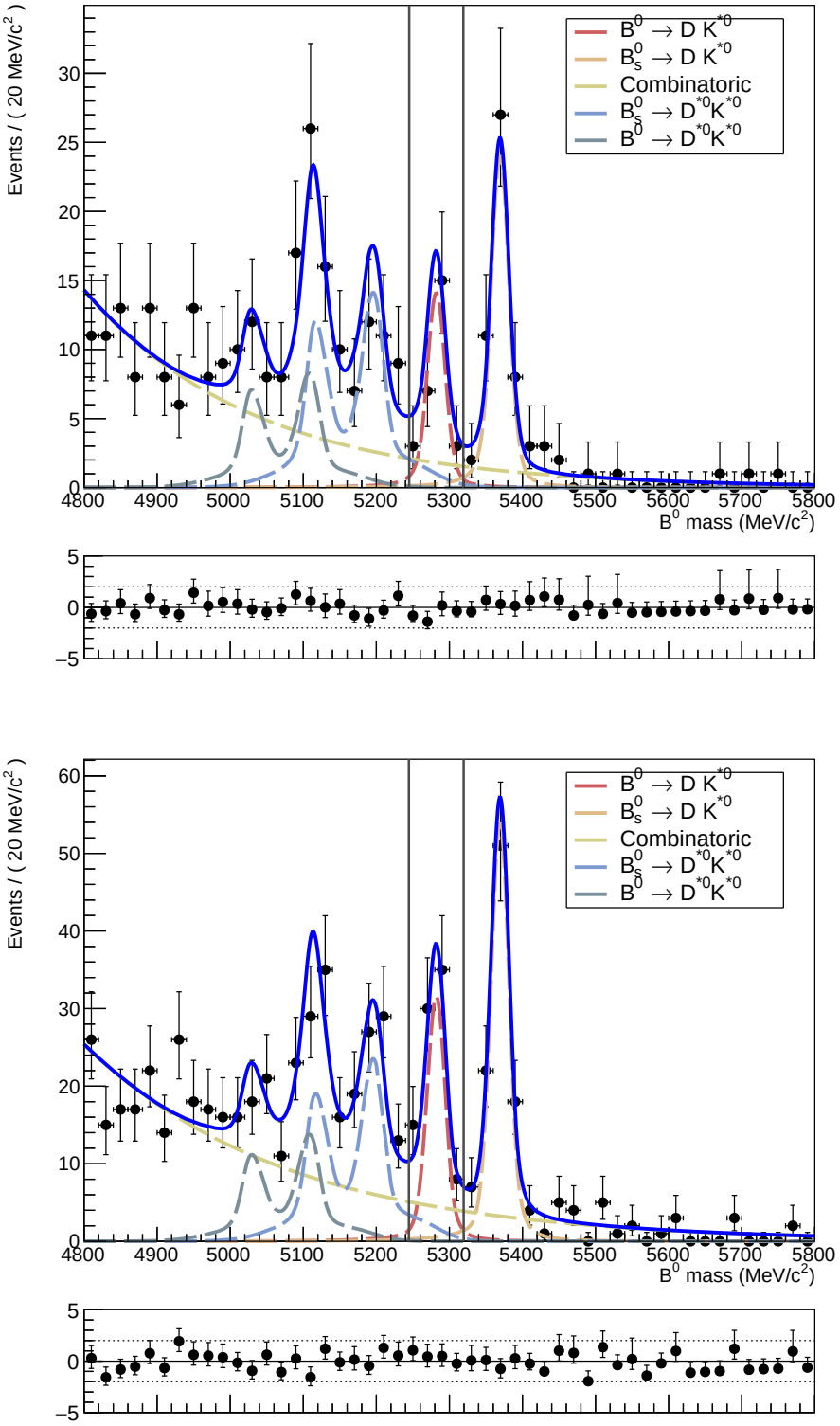


Figure 5.10.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for LL (top) and DD (bottom) $B^0 \to DK^{*0}$ candidates. In each plot, the lower canvas shows the pull.

Free Parameter	Designation	Fitted Value	Description
μ	`Bs_mean`	5369^{+1}_{-1}	Central value of B_s^0 peak
$N^{\mathrm{LL}}_{B_s^0}$	`LL_nBs`	43^{+7}_{-6}	$B_s^0 \to DK^{*0}$ yield (LL)
$N^{\mathrm{DD}}_{B_s^0}$	`DD_nBs`	93^{+11}_{-10}	$B_s^0 \to DK^{*0}$ yield (DD)
r_{B^0}	`BdBs_frac`	$0.585^{+0.095}_{-0.084}$	Ratio of B^0 and B_s^0 yields
k_σ	`smear_factor`	$0.907^{+0.093}_{-0.083}$	Ratio of Crystal Ball widths measured in data and simulation
$N^{s\ \mathrm{LL}}_{010}$	`LL_NDst_Bs_010`	65^{+12}_{-11}	010-helicity $B_s^0 \to D^{*0}K^{*0}$ yield (LL)
$N^{s\ \mathrm{DD}}_{010}$	`DD_NDst_Bs_010`	108^{+17}_{-16}	010-helicity $B_s^0 \to D^{*0}K^{*0}$ yield (DD)
r_{101}	`ratio_001_010`	$0.003^{+0.1}_{-0.1}$	Ratio of 101 and 010 helicity components in $B_s^0 \to D^{*0}K^{*0}$
$N^{\mathrm{LL}}_{\mathrm{comb}}$	`LL_nComb`	166^{+18}_{-17}	Combinatoric yield (LL)
$N^{\mathrm{DD}}_{\mathrm{comb}}$	`DD_nComb`	339^{+26}_{-25}	Combinatoric yield (DD)
τ^{LL}	`LL_Comb_tau`	$-0.00423^{+0.0005}_{-0.0005}$	Combinatoric slope (LL)
τ^{DD}	`DD_Comb_tau`	$-0.00364^{+0.0003}_{-0.0003}$	Combinatoric slope (DD)

Table 5.12.: Free parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected $B^0 \to DK^{*0}$ candidates. The designations assigned within the fitting program are listed, to allow the reader to interpret the corresponding plot of correlations, Figure 5.11.

Component	Yield	
	LL	DD
signal	21 ± 3	46 ± 5
B_s^0	1.0 ± 0.1	2.2 ± 0.2
$B_s^0 \to D^* K^{*0}$	2.4 ± 0.3	5 ± 2
$B^0 \to D^* K^{*0}$	0.026 ± 0.004	0.03 ± 0.01
combinatoric	4.5 ± 0.5	10.9 ± 0.8

Table 5.13.: Yields extracted in the signal region for mode $B^0 \to DK^{*0}$.

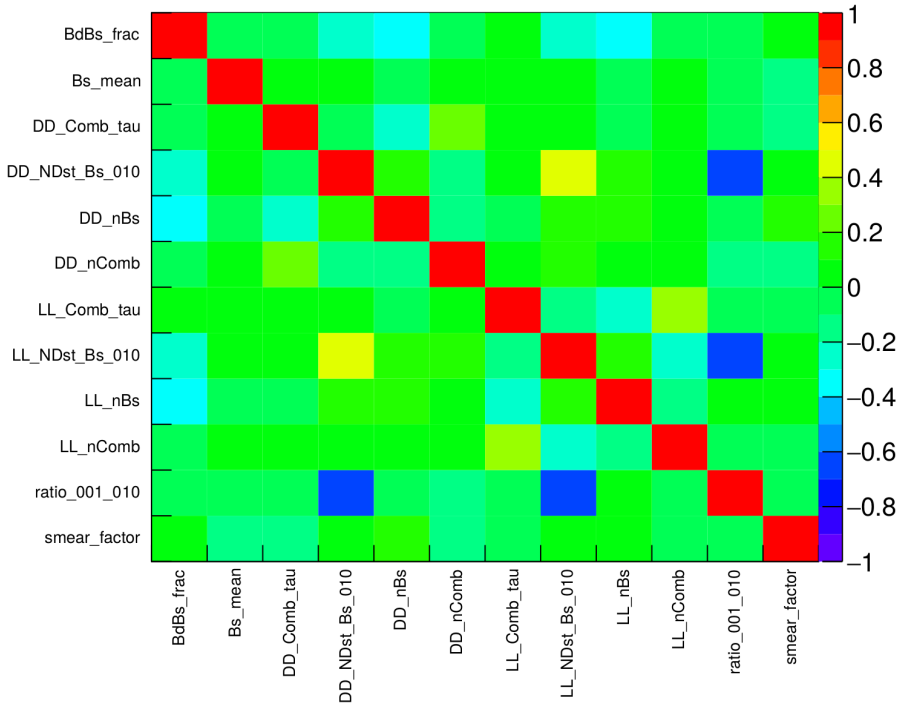


Figure 5.11.: Correlation matrix obtained for fitted parameters in the simultaneous B-mass fit to LL and DD $B^0 \to DK^{*0}$ candidates.

5.3. Fit to selected $B^{\pm} \to DK^{*\pm}$ candidates

A single simultaneous fit is performed to selected $B^{\pm} \to DK^{*\pm}$ candidates, which are split into two categories depending on the track type of the reconstructed K^0_S candidate.The PDFs used in this fit are described below, and the simulated samples used to obtain PDF shapes for this fit are listed in Table 5.14.

Mode	No. of candidates generated
$B^{\pm} \to DK^{*\pm}$	6.0×10^6
$B^{\pm} \to D^*K^{*\pm}, D^* \to D\gamma$ (001)	3.0×10^6
$B^{\pm} \to D^*K^{*\pm}, D^* \to D\gamma$ (010)	3.0×10^6
$B^{\pm} \to D^*K^{*\pm}, D^* \to D\gamma$ (100)	3.0×10^6
$B^{\pm} \to D^*K^{*\pm}, D^* \to D\pi^0$ (001)	3.0×10^6
$B^{\pm} \to D^*K^{*\pm}, D^* \to D\pi^0$ (010)	3.0×10^6
$B^{\pm} \to D^*K^{*\pm}, D^* \to D\pi^0$ (100)	3.0×10^6

Table 5.14.: Simulated Monte Carlo samples used to extract PDF shapes for the reconstructed B-mass fit of selected $B^{\pm} \to DK^{*\pm}$ candidates. Labels (001) and (010) for the $B^{\pm} \to D^*K^{*\pm}$ samples refer to the simulated helicity combinations in the scalar to vector-vector decay, as described in Section 5.2.

Signal

Double Crystal Ball shapes are used to describe the contribution from signal candidates. The tail parameters of the two component Crystal Ball shapes, along with the constant of proportionality between their relative magnitudes, are fixed to values extracted from fits to simulated signal candidates, as shown in Figure 5.12 (a) and Figure 5.13 (a) for LL and DD samples, respectively. The full selection is applied, with the exception of the MVA requirement, since these same simulation samples are used in its training.[8]

The full PDF used to describe $B^{\pm} \to D^*K^{*\pm}$ candidates in the mass fit may then be expressed as

$$F^{tt}_{\text{sig}}(m; N^{tt}_{\text{sig}}, \mu, \sigma^{tt}) \equiv N^{tt}_{\text{sig}} \cdot \left(f^{tt} \cdot \mathcal{CB}(m; \mu, \sigma^{tt}, \alpha^{tt}_1, n^{tt}_1) + (1 - f^{tt}) \cdot \mathcal{CB}(m; \mu, -\sigma^{tt}, \alpha^{tt}_2, n^{tt}_2) \right), \tag{5.56}$$

where the parameters fixed from fits to simulated signal candidates are given by:

- $\alpha^{\text{LL}}_1 = 1.68$ and $\alpha^{\text{DD}}_1 = 2.02$;
- $n^{\text{LL}}_1 = 2.56$ and $n^{\text{DD}}_1 = 1.75$;
- $\alpha^{\text{LL}}_2 = 1.72$ and $\alpha^{\text{DD}}_2 = 1.72$;
- $n^{\text{LL}}_2 = 3.26$ and $n^{\text{DD}}_2 = 3.26$; and

[8] Ideally, a subset of the Monte Carlo simulation would be excluded from the MVA training, so that it could safely be used here with the MVA requirement applied. However, insufficient simulated statistics are available to allow this in the present analysis, and the full simulated samples are used for training.

- $f^{\mathrm{LL}} = 0.869$ and $f^{\mathrm{DD}} = 0.806$.

Partially-reconstructed $B^{\pm} \to D^{*}K^{*\pm}$ background

Analogously with the fit to $B^0_{(s)} \to DK^{*0}$ candidates, an important source of background arises from partially reconstructed $B^{\pm} \to D^{*}K^{*\pm}$ decays, with $D^{*} \to DX^0$, where $X^0 = \gamma, \pi^0$ is missed in the reconstruction of the event. The helicity structure in these decays is the same as for $B^0_{(s)} \to D^{*}K^{*0}$, described in Section 5.2, since both are decays of a pseudo scalar meson to two vector mesons.

Once again, kernel-estimated shapes are used to describe these backgrounds, as displayed in subplots (b), (c), (d) and (e) of Figure 5.12 and Figure 5.13 for LL and DD samples respectively. The different partially-reconstructed $B^0_{(s)} \to D^{*}K^{*}$ components are again combined according to their known branching fractions [147], and their efficiencies in the full selection, and may be described mathematically as

$$\begin{aligned} F^{tt}_{\mathrm{part}}(m; N^{tt}_{010}, r_{101}) =& N^{tt}_{010} \cdot \left(G^{tt}_{010} \cdot f^{tt}_{\gamma 010}(m) + (1 - G^{tt}_{010}) \cdot f^{tt}_{\pi^0 010}(m)\right) \\ &+ r_{101} N^{tt}_{010} \cdot \left(G^{tt}_{101} \cdot f^{tt}_{\gamma 101}(m) + (1 - G^{tt}_{101}) \cdot f^{tt}_{\pi^0 101}(m)\right), \end{aligned} \tag{5.57}$$

where:

- N^{tt}_{010} are the $B^{\pm} \to D^{*}K^{*\pm}$ yields for helicity component 010, and r_{101} gives the relative yield for helicity component 101 as a fraction of this value, such that $r_{101} = N^{tt}_{101} / N^{tt}_{010}$ (101 again designates the combination $100 + 001$). r_{101} is taken to be common for both track types;
- $f^{tt}_{\gamma H}$ and $f^{tt}_{\pi^0 H}$ denote the kernel-estimated PDFs for helicity component H, extracted from simulated $B^{\pm} \to D^{*}K^{*\pm}$ samples; and
- G^{tt}_{H} give the proportionality of the $D^{*} \to D\gamma$ contribution relative to the $D^{*} \to D\pi^0$ contribution, for helicity state H. They are fixed from the measured branching fractions [147] and the relative efficiencies observed for the two components in Monte Carlo simulation, and are defined as

$$G^{tt}_{H} \equiv \frac{g^{tt}_{H}}{g^{tt}_{H} + p^{tt}_{H}}, \tag{5.58}$$

where

$$g^{tt}_{H} = \mathcal{B}(D\gamma) \cdot \varepsilon^{H\,tt}_{\mathrm{gen}}(D\gamma) \cdot \varepsilon^{H\,tt}_{\mathrm{strip}}(D\gamma) \cdot \varepsilon^{H\,tt}_{\mathrm{sel}}(D\gamma), \quad \text{and} \tag{5.59}$$

$$p^{tt}_{H} = \mathcal{B}(D\pi^0) \cdot \varepsilon^{H\,tt}_{\mathrm{gen}}(D\pi^0) \cdot \varepsilon^{H\,tt}_{\mathrm{strip}}(D\pi^0) \cdot \varepsilon^{H\,tt}_{\mathrm{sel}}(D\pi^0). \tag{5.60}$$

The Monte Carlo generator efficiency, $\varepsilon_{\mathrm{gen}}$, includes the effects of all detector acceptance requirements applied. $\varepsilon_{\mathrm{strip}}$ is the efficiency of the stripping selection, and $\varepsilon_{\mathrm{sel}}$ gives the efficiency of the full offline selection. The values of these parameters are summarised in Table 5.15.

Quantity	$D^{*0} \to D^0\pi^0$		$D^{*0} \to D^0\gamma$	
	LL	DD	LL	DD
Branching fraction	$(65.5 \pm 0.9)\%$ [147]		$(34.5 \pm 0.9)\%$ [147]	
$\varepsilon^{010}_{\text{gen}}$	$(15.9 \pm 0.02)\%$		$(18.4 \pm 0.03)\%$	
$\varepsilon^{101}_{\text{gen}}$	$(15.6 \pm 0.02)\%$		$(18.0 \pm 0.02)\%$	
$\varepsilon^{010}_{\text{strip}}$	$(0.106 \pm 0.002)\%$		$(0.0898 \pm 0.002)\%$	
$\varepsilon^{101}_{\text{strip}}$	$(0.106 \pm 0.001)\%$		$(0.0915 \pm 0.001)\%$	
$\varepsilon^{010}_{\text{sel}}$	$(8.8 \pm 0.5)\%$	$(27.7 \pm 1.0)\%$	$(8.9 \pm 0.6)\%$	$(28.2 \pm 1.1)\%$
$\varepsilon^{101}_{\text{sel}}$	$(9.8 \pm 0.4)\%$	$(29.8 \pm 0.7)\%$	$(9.9 \pm 0.4)\%$	$(29.2 \pm 0.8)\%$

Table 5.15.: Known branching fractions, and measured efficiencies obtained in Monte Carlo simulation, for partially-reconstructed $B^{\pm} \to D^* K^{*\pm}$ background components.

Combinatoric background

The combinatoric background component is modelled by a decreasing exponential distribution with slope parameter τ,

$$F^{tt}_{\text{comb}}(m; N^{tt}_{\text{comb}}, \tau^{tt}) = N^{tt}_{\text{comb}} \cdot \mathrm{e}^{\tau^{tt} m} |\tau^{tt}|. \tag{5.61}$$

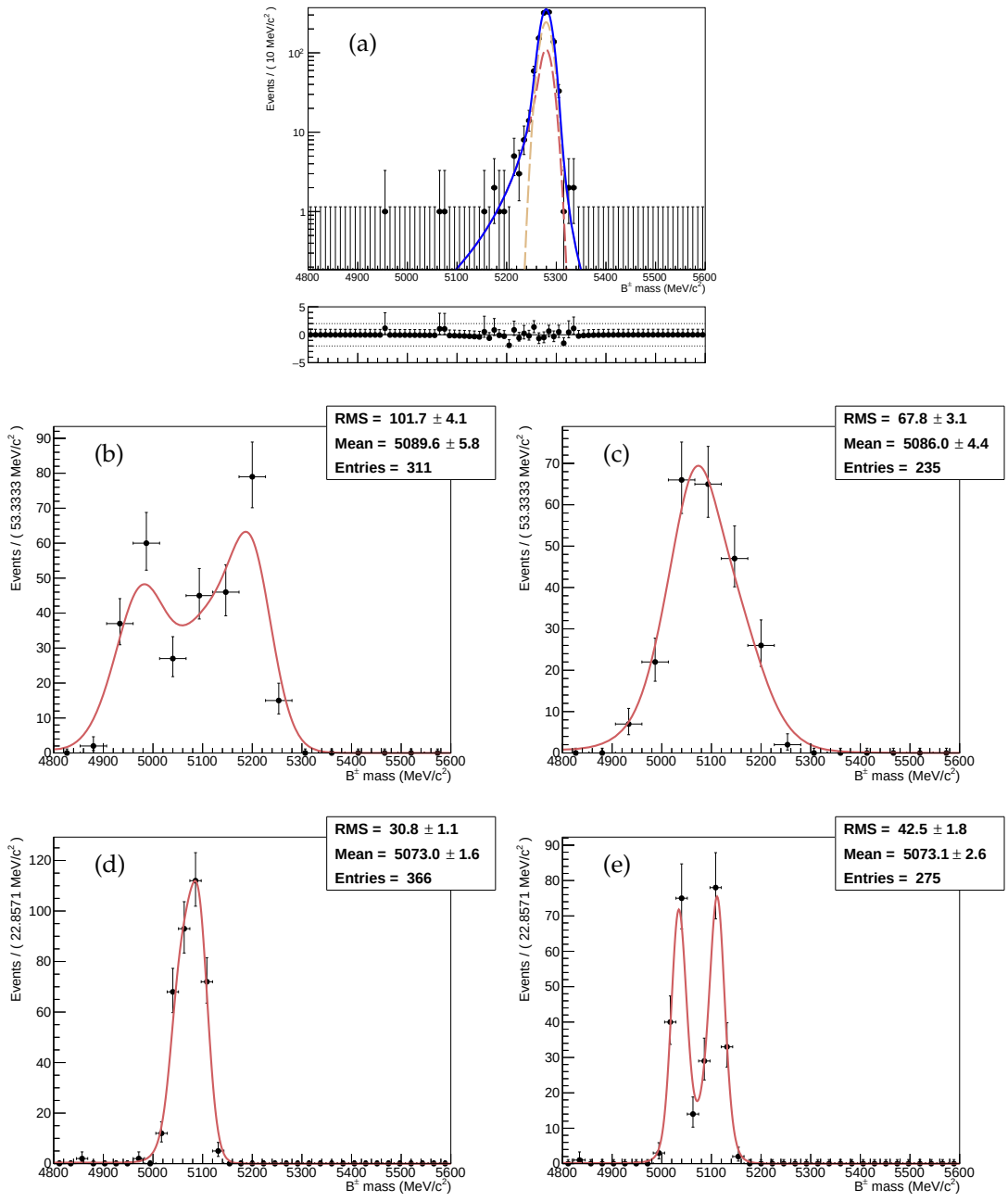


Figure 5.12.: B-mass PDF shapes used in the fit to LL $B^\pm \to DK^{*\pm}$ candidates. A Double Crystal Ball function is fitted to signal mode Monte Carlo (a). Kernel-estimated shapes are obtained from $B^\pm \to D^*K^{*\pm}$ Monte Carlo samples, for (b)$D^* \to D\gamma_{001}$, (c)$D^* \to D\gamma_{010}$, (d) $D^* \to D\pi^0_{001}$ and (e)$D^* \to D\pi^0_{010}$ candidates. For plot (a), histogram points show Monte Carlo simulation, a solid blue line gives the fitted PDF, and the two constituent Crystal Ball functions of the fitted PDF are shown as dashed red and yellow lines. The lower canvas of (a) shows the pull.

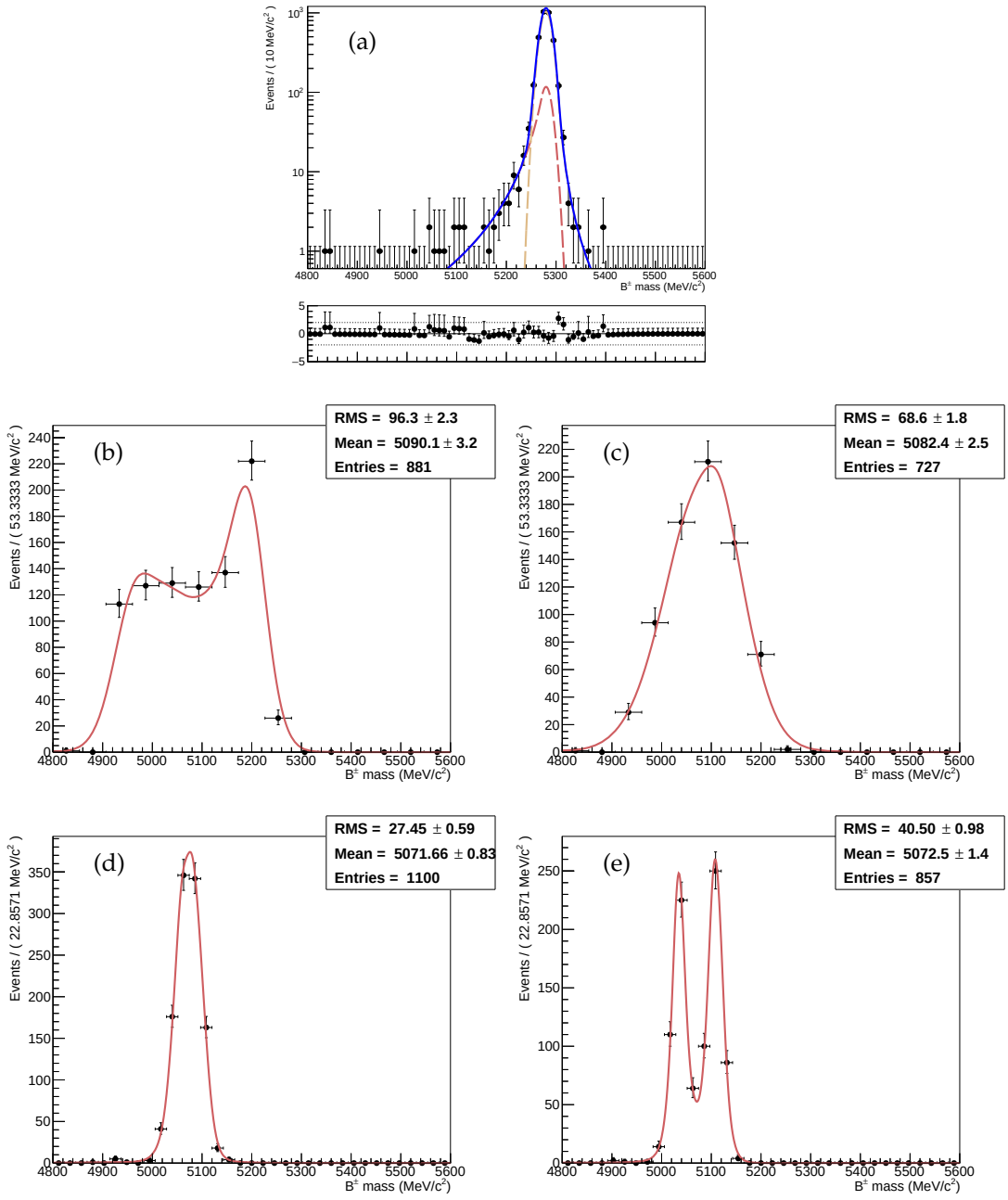


Figure 5.13.: *B*-mass PDF shapes used in the fit to DD $B^\pm \to DK^{*\pm}$ candidates. A Double Crystal Ball function is fitted to signal mode Monte Carlo (a). Kernel-estimated shapes are obtained from $B^\pm \to D^*K^{*\pm}$ Monte Carlo samples, for (b)$D^* \to D\gamma_{001}$, (c)$D^* \to D\gamma_{010}$, (d) $D^* \to D\pi^0_{001}$ and (e)$D^* \to D\pi^0_{010}$ candidates. For plot (a), histogram points show Monte Carlo simulation, a solid blue line gives the fitted PDF, and the two constituent Crystal Ball functions of the fitted PDF are shown as dashed red and yellow lines. The lower canvas of (a) shows the pull.

5.3.1. Mass fit result

In summary, the model used to fit the reconstructed B-mass spectrum for selected data of each track type is comprised of six components. Double Crystal Ball shapes are used to describe signal decays, whilst combinatoric background is described using an exponential. For background arising from partially-reconstructed $B^{\pm} \to D^* K^{*\pm}$ decays, four kernel-estimated PDF shapes are used to describe the two distinguishable helicity distributions and the two possible decay modes $D^* \to D\gamma$ and $D^* \to D\pi^0$. The four kernel-estimated PDFs are combined according to their branching fractions and measured efficiencies in MC.

The full PDF model used to fit each track type category may be written in terms of its free parameters as

$$\begin{aligned} F^{tt}_{\text{tot}} =& F^{tt}_{\text{sig}}(m; N^{tt}_{\text{sig}}, \mu, \sigma^{tt}) + F^{tt}_{\text{comb}}(m; N^{tt}_{\text{comb}}, \tau^{tt}) \\ &+ F^{tt}_{\text{part}}(m; N^{tt}_{010}, r_{101}). \end{aligned} \tag{5.62}$$

All parameters are taken to be independent for samples of different track type, except for:

- the central value of the signal peak, μ; and
- the ratio of helicities in $B^0_{(s)} \to D^* K^{*0}$ background, r_{101}.

The constant parameters of the fit, and their fixed values, are summarized in Table 5.16. The full simulataneous fit to LL and DD candidates, performed on the full 2011 and 2012 LHCb dataset, is displayed in Figure 5.14. The values returned for the free parameters are summarized in Table 5.17, and their obtained correlations are illustrated in Figure 5.15. Only candidates falling within a $\pm 3\sigma$ window centred on the fitted signal mean, μ, are propagated forward to the full $\mathcal{CP}$ fit. This window is marked by vertical lines in Figure 5.14. For LL candidates, the window used is $[5251; 5318]$ MeV, and for DD candidates it is $[5235; 5334]$ MeV. These windows differ in width due to the difference in the values obtained for σ^{LL} and σ^{DD} in the fit. Component yields extracted within these windows are enumerated in Table 5.18.

Constant Parameter	Value	Description
α_1^{LL}	1.68	Double Crystal Ball parameters for signal PDFs
α_2^{LL}	1.72	
α_1^{DD}	2.02	
α_2^{DD}	1.72	
n_1^{LL}	2.56	
n_2^{LL}	3.26	
n_1^{DD}	1.75	
n_2^{DD}	3.26	
f^{LL}	0.869	
f^{DD}	0.806	
G_{010}^{LL}	0.377	Proportionality constants for $B^{\pm} \to D^* K^{*\pm}$ backgrounds
G_{010}^{DD}	0.379	
G_{101}^{LL}	0.381	
G_{101}^{DD}	0.374	

Table 5.16.: Constant parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected $B^{\pm} \to DK^{*\pm}$ candidates.

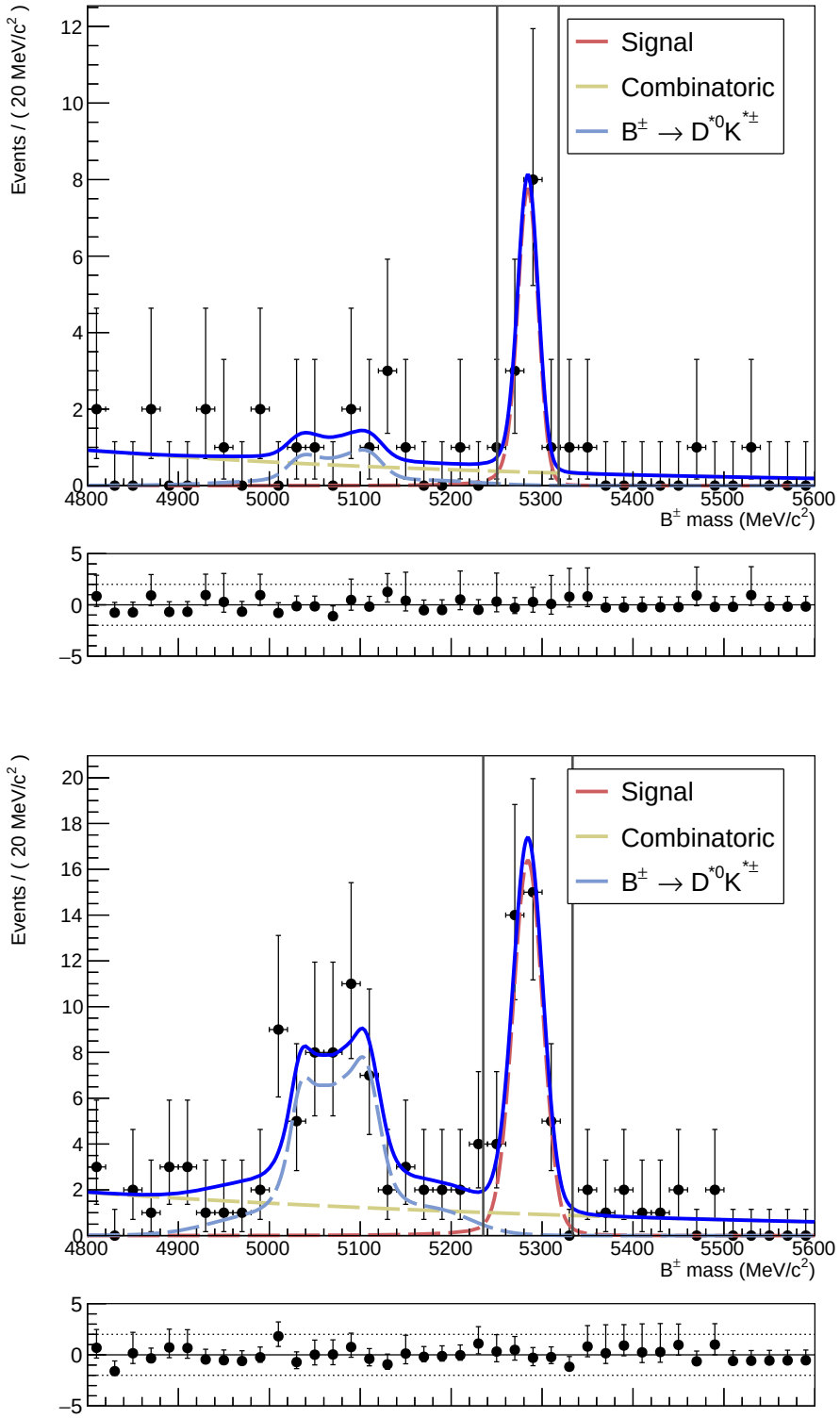


Figure 5.14.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for LL (top) and DD (bottom) $B^\pm \to DK^{*\pm}$ candidates. In each plot, the lower canvas shows the pull.

Free Parameter	Designation	Fitted Value	Description
μ	`mean`	5283^{+2}_{-2}	Central value of signal peak
$N_{\text{sig}}^{\text{LL}}$	`LL_nB`	12^{+4}_{-4}	$B^{\pm} \to DK^{*\pm}$ yield (LL)
$N_{\text{sig}}^{\text{DD}}$	`DD_nB`	38^{+8}_{-7}	$B^{\pm} \to DK^{*\pm}$ yield (DD)
σ^{LL}	`LL_sigmaCB1`	11^{+4}_{-3}	Width parameter of Crystal Ball shapes (LL)
σ^{DD}	`DD_sigmaCB1`	16^{+4}_{-4}	Width parameter of Crystal Ball shapes (DD)
N_{010}^{LL}	`LL_NDst_010`	4^{+4}_{-3}	Yield of 010-helicity $B^{\pm} \to D^{*}K^{*\pm}$ (LL)
N_{010}^{DD}	`DD_NDst_010`	36^{+15}_{-14}	Yield of 010-helicity $B^{\pm} \to D^{*}K^{*\pm}$ (DD)
r_{101}	`ratio_001_010`	$0.003^{+0.1}_{-0.1}$	Ratio of 101 and 010 helicity components in $B^{\pm} \to D^{*}K^{*\pm}$
$N_{\text{comb}}^{\text{LL}}$	`LL_nComb`	33^{+8}_{-7}	Combinatoric yield (LL)
$N_{\text{comb}}^{\text{DD}}$	`DD_nComb`	74^{+13}_{-11}	Combinatoric yield (DD)
τ^{LL}	`LL_Comb_tau`	$-0.0027^{+0.0009}_{-0.0009}$	Combinatoric slope (LL)
τ^{DD}	`DD_Comb_tau`	$-0.0013^{+0.0006}_{-0.0006}$	Combinatoric slope (DD)

Table 5.17.: Free parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected $B^{\pm} \to DK^{*\pm}$ candidates. The designations assigned within the fitting program are listed, to allow the reader to interpret the corresponding plot of correlations, Figure 5.15.

Component	Yield	
	LL	DD
signal	12 ± 4	32 ± 6
$B^{\pm} \to D^{*}K^{*\pm}$	0.02 ± 0.01	0.09 ± 0.03
combinatoric	1.3 ± 0.3	4.4 ± 0.6

Table 5.18.: Yields extracted in the signal region for mode $B^{\pm} \to DK^{*\pm}$

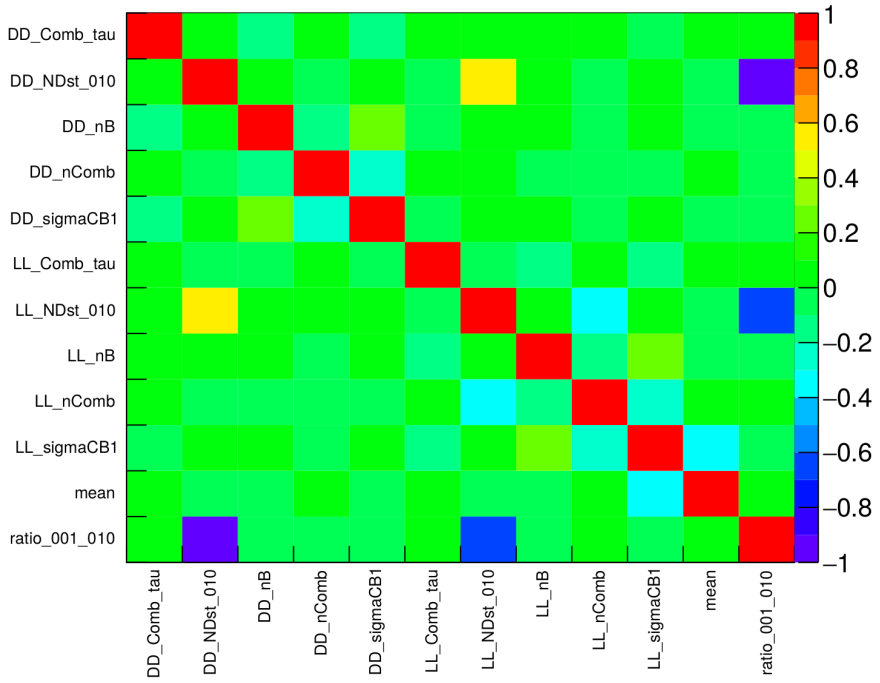


Figure 5.15.: Correlation matrix obtained for parameters in the simultaneous *B*-mass fit to LL and DD $B^{\pm} \to DK^{*\pm}$ candidates.

5.4. Fit to selected $B^\pm \to D^* h^\pm$ candidates

Two separate simultaneous fits are performed to selected $B^\pm \to D^* h^\pm$ candidates. Separate fits are performed for candidates containing K^0_S candidates of different track types. Within each fit, six separate categories are fitted simultaneously, split by the identification of the bachelor type (pion or kaon) and by the neutral hypothesis satisfied by the reconstructed D^*-meson decay, $D^* \to DX^0$ ($X^0 = \gamma, \pi^0$, or $\pi^0_{\text{part.reco.}}$), as described in Section 4.3.3. The neutral hypothesis $\pi^0_{\text{part.reco.}}$, or 'partially-reconstructed π^0' corresponds to the case when a real $D^* \to D\pi^0$ decay is reconstructed as $D^* \to D\gamma$, due to the loss of one photon from the subsequent $\pi^0 \to \gamma\gamma$ decay. The requirements used to separate the selected data samples into these categories are listed in Table 5.19. The missing mass variable m_3 is defined as $m_3 \equiv m(D^*) - m(D) - m(X^0)$, as discussed in Section 4.3.3.

Bachelor hypothesis	Neutral hypothesis		
	$X^0 = \gamma$	$X^0 = \pi^0$	$X^0 = \pi^0_{\text{part.reco.}}$
$B^\pm \to D^* K^\pm$	$\text{DLL}_{K\pi}(\text{bach}) > 4$	$\text{DLL}_{K\pi}(\text{bach}) > 4$	$\text{DLL}_{K\pi}(\text{bach}) > 4$
	$130 < m_3 < 160\,\text{MeV}$	$2 < m_3 < 12.5\,\text{MeV}$	$60 < m_3 < 100\,\text{MeV}$
$B^\pm \to D^* \pi^\pm$	$\text{DLL}_{K\pi}(\text{bach}) < 4$	$\text{DLL}_{K\pi}(\text{bach}) < 4$	$\text{DLL}_{K\pi}(\text{bach}) < 4$
	$130 < m_3 < 160\,\text{MeV}$	$2 < m_3 < 12.5\,\text{MeV}$	$60 < m_3 < 100\,\text{MeV}$

Table 5.19.: Requirements used to separate selected $B^\pm \to D^* h^\pm$ candidates into six simultaneous fit categories.

In what follows, each of the six categories defined will be referred to by the syntax '(b, n)', where b denotes the bachelor hypothesis (π or K) and n the neutral hypothesis (γ, π^0 or $\pi^0_{\text{part.reco.}}$). The PDFs used in this fit are described below. The simulated samples used to obtain PDF shapes for this fit are listed in Table 5.20.

Mode	Conditions	No. of candidates generated
$B^\pm \to D^* K^\pm$	2012	5.3×10^6
$B^\pm \to D^* \pi^\pm$	2012	6.0×10^6
$B^\pm \to D^* \rho^\pm, \rho^\pm \to \pi^\pm \pi^0$	2012	3.0×10^6
$B^\pm \to D^* K^{*\pm}, K^{*\pm} \to K^\pm \pi^0$	2012	3.0×10^6
$B^\pm \to DK^\pm$	2012	6.4×10^6
	2011	3.2×10^6
$B^\pm \to D\pi^\pm$	2012	6.4×10^6
	2011	3.2×10^6
$B^\pm \to D\rho^\pm, \rho^\pm \to \pi^\pm \pi^0$	2012	1.5×10^6
$B^0 \to D^{*\pm} \pi^\mp, D^{*\pm} \to \pi^\mp D^0$	2012	0.5×10^6
$B^0 \to D^{*\pm} K^\mp, D^{*\pm} \to \pi^\mp D^0$	2012	0.5×10^6

Table 5.20.: Simulated Monte Carlo samples used to extract shapes for the reconstructed B-mass fit of selected $B^\pm \to D^* h^\pm$ candidates.

It is important to note that all samples used in this mass fit, both data and MC, may be *reconstructed* under either hypothesis for the bachelor meson candidate, $h = \pi$ or $h = K$. These reconstructions differ in the masses assigned to the bachelor candidate, and thus in the mass reconstructed for the progenitor B-meson candidate. For each candidate B-meson decay, both reconstructions are performed, and candidates matching a given bachelor hypothesis ($b = \pi, K$) are selected by applying the requirements defined in Table 5.19 to the appropriately-reconstructed sample. This is safe because the PID requirements defined to separate bachelor categories in Table 5.19 are disjoint, ensuring that any candidate can enter one category only ($b = \pi$ or $b = K$).

The shapes used to describe all components except combinatoric background are determined from simulated samples. These samples are split using the same criteria listed in Table 5.19 to give six subsamples for each K^0_S track type. In addition, truth-level information is used to determine the true neutral hypothesis corresponding to each simulated event. Where the same background is expected to contribute in both categories of bachelor hypothesis (*i.e.* both $B^\pm \to D^*\pi^\pm$ and $B^\pm \to D^*K^\pm$), the simulated sample for this background is reconstructed separately under each hypothesis, and separate shapes are extracted in each case, after the application of the full selection.

Separate shapes are extracted for candidates satisfying the neutral hypothesis ($n = \pi^0_{\text{part.reco.}}$), known as the partially-reconstructed sample, and for those satisfying the neutral hypothesis ($(n = \gamma)||(n = \pi^0)$), known as the 'correctly-reconstructed' sample. Independent shapes are not extracted for candidates satisfying the neutral hypothesis ($n = \pi^0$) alone, since the yield of this signal category in both data and simulation is significantly lower than for the other categories.[9] Instead, common shapes are used for both of the categories with correctly-reconstructed neutral products.

Signal

Double Crystal Ball shapes are used to describe the contribution from signal candidates in each fit category. The tail parameters of the two component Crystal Ball shapes, along with the constant of proportionality between their relative magnitudes, are fixed to values extracted from fits to simulated signal candidates, $B^\pm \to D^*K^\pm$. In order for each calibration fit to be performed using a reasonably-sized simulation sample, candidates of both K^0_S track types are combined, and the selection applied to these samples does not include any requirement on the MVA output variable. This also ensures that the MVA is not applied to the same simulated samples on which it is trained, which could risk inducing a bias in the obtained PDF.[10] Corresponding fits are shown in Figures 5.16 (a) and 5.17 (a), for correctly-reconstructed and partially-reconstructed samples, respectively.

The signal PDF shapes may be expressed as

$$F^{b\,n}_{\text{sig}}(m; N^{b\,n}_{\text{sig}}, \mu^n, \sigma^{b\,n}) \equiv N^{b\,n}_{\text{sig}} \cdot \big(f^n \cdot \mathcal{CB}(m; \mu^n, \sigma^{b\,n}, \alpha^{b\,n}_1, n^{b\,n}_1) + (1 - f^n) \cdot \mathcal{CB}(m; \mu^n, -\sigma^{b\,n}, \alpha^n_2, n^n_2)\big), \tag{5.63}$$

[9] This indicates that a $D^* \to D\pi^0$ decay is more likely to be mis-reconstructed as $D^* \to D\gamma$ than it is to be correctly reconstructed. This is due to the fact that it is relatively common for the two photons emitted by the decaying π^0 to make a single cluster in the electromagnetic calorimeter, indistinguishable from a single photon.

[10] Ideally, a subset of the Monte Carlo simulation would be excluded from the MVA training, so that it could safely be used here with the MVA requirement applied. However, insufficient simulated statistics are available to allow this in the present analysis, and the full simulated samples are used for training.

where a common mean, μ^n, is used in the fit for both bachelor categories, and the parameter values fixed from fits to signal MC are:

- $\alpha_1^{\gamma||\pi^0} = 0.233$ and $\alpha_1^{\pi^0_{\text{part.reco.}}} = 0.670$;
- $\alpha_2^{\gamma||\pi^0} = 0.886$ and $\alpha_2^{\pi^0_{\text{part.reco.}}} = 1.03$;
- $n_1^{\gamma||\pi^0} = 112$ and $n_1^{\pi^0_{\text{part.reco.}}} = 22.4$;
- $n_2^{\gamma||\pi^0} = 131$ and $n_2^{\pi^0_{\text{part.reco.}}} = 3.26$; and
- $f^{\gamma||\pi^0} = 0.0536$ and $f^{\pi^0_{\text{part.reco.}}} = 0.278$.

Cross-feed between bachelor hypotheses

Cross-feed can occur when candidates are reconstructed with the wrong mass hypothesis $h = \pi^\pm, K^\pm$ for the bachelor track of the $B^\pm \to D^* h^\pm$ decay. Due to its much greater branching fraction, the cross-feed of true $h = \pi$ candidates into the selected $b = K$ samples is expected to be much greater than the opposite process. Crystal Ball PDFs are used to describe this component in each (K, n) fit category, with their tail parameters fixed from fits to simulation. Once again, candidates of both K^0_S track types are combined, and the selection applied to these samples does not include any requirement on the MVA output variable. Corresponding fits are shown in Figures 5.16 (b) and 5.17 (b), for correctly-reconstructed and partially-reconstructed samples respectively. These shape for signal cross=feeding into the $b = K$ category may be expressed as

$$F^{K\,n}_{\text{xfeed}}(m; N^{K\,n}_{\text{xfeed}}, \mu^{K\,n}_{\text{xfeed}}, \sigma^{K\,n}_{\text{xfeed}}) \equiv N^{K\,n}_{\text{xfeed}} \cdot \mathcal{CB}(m; \mu^{K\,n}_{\text{xfeed}}, \sigma^{K\,n}_{\text{xfeed}}, \alpha^{K\,n}_{\text{xfeed}}, n^{K\,n}_{\text{xfeed}}), \tag{5.64}$$

where the parameter values fixed from fits to MC are:

- $\alpha^{K\,\gamma||\pi^0}_{\text{xfeed}} = 0.460$ and $\alpha^{K\,\pi^0_{\text{part.reco.}}}_{\text{xfeed}} = 1.34$; and
- $n^{K\,\gamma||\pi^0}_{\text{xfeed}} = 126$ and $n^{K\,\pi^0_{\text{part.reco.}}}_{\text{xfeed}} = 144$.

Cross-feed of true $h = K$ candidates into the $b = \pi$ samples is described using a Gaussian PDF, since the yield of $h = K$ simulation which passes selection criteria for the $b = \pi$ hypothesis is too small to be reliably fitted using a PDF with more parameters. However, a precise determination of the tail parameters for this cross-feed is not crucial, since it is expected to have a very low yield in data, just as it does in the $B^\pm \to Dh^\pm$ decay modes (see Section 5.1). This PDF may be written as

$$F^{\pi\,n}_{\text{xfeed}}(m; N^{\pi\,n}_{\text{xfeed}}, \mu^{\pi\,n}_{\text{xfeed}}, \sigma^{\pi\,n}_{\text{xfeed}}) \equiv N^{\pi\,n}_{\text{xfeed}} \cdot \mathcal{G}(\mu^{\pi\,n}_{\text{xfeed}}, \sigma^{\pi\,n}_{\text{xfeed}}). \tag{5.65}$$

Cross-feed between neutral hypotheses

No PDFs are included to describe the contribution of cross-feed between signal corresponding to the three different neutral hypotheses. Simulated signal samples are studied using truth-level information

to determine the fraction of candidates that are assigned the wrong hypothesis under the selection criteria listed in Table 5.19 (see Figure 4.8, which illustrates the distributions of the missing mass variable, m_3, for truth-matched MC candidates generated under each hypothesis). In each fit category, the cross-feed yield in data is expected to be small, with the greatest fractional contribution cross-feeding from true $X^0 = \pi^0_{\text{part.reco.}}$ decays into the $n = \gamma$ fit category. For the fit to DD K^0_S candidates, this gives ~ 5 expected cross-feed candidates compared to a signal yield of ~ 135. At this level of statistics, the introduction of an extra PDF, almost coincident with the signal PDF, to describe this small contribution gives an unreliable fit. Instead, it is assumed that the signal PDF shape captures any small contribution from these cross-feed candidates. Since $D^* \to D\gamma$ and $D^* \to D\pi^0$ candidates require different phase models in the final $\mathcal{CP}$ fit, a systematic uncertainty may later be associated with the estimation of these cross-feed contributions.

Kernel-estimated background shapes

Kernel-estimated PDF shapes are used to describe the contributions from various background sources. For samples of bachelor hypothesis $b = K$, shapes are included to describe:

- $B^\pm \to D^* K^{*\pm}$, with $K^{*\pm} \to K^\pm \pi^0$;
- $B^\pm \to D K^\pm$;
- $B^\pm \to D \pi^\pm$; and
- $B^0 \to D^{*\pm} K^\mp$, $D^{*\pm} \to \pi^\mp D^0$.

$B^\pm \to D^* K^{*\pm}$ may provide a background in this sample if it is reconstructed without the inclusion of the neutral pion daughter of the $K^{*\pm}$-meson. $B^\pm \to D K^\pm$ may enter as background if in the reconstruction it is incorrectly combined with a neutral candidate from elsewhere in the event. The same is true for $B^\pm \to D\pi^\pm$, although in this case a $\pi \to K$ mis-identification must also occur.

For samples of bachelor hypothesis $b = \pi$, shapes are included to describe

- $B^\pm \to D^* \rho^\pm$, with $\rho^\pm \to \pi^\pm \pi^0$;
- $B^\pm \to D \pi^\pm$;
- $B^\pm \to D \rho^\pm$, with $\rho^\pm \to \pi^\pm \pi^0$; and
- $B^0 \to D^{*\pm} \pi^\mp$, with $D^{*\pm} \to \pi^\mp D^0$.

The backgrounds $B^\pm \to D K^\pm$ and $B^0 \to D^{*\pm} K^\mp$ are not included in this case, since their lower branching fractions and the additional $K \to \pi$ mis-identification required ensure that their yields are significantly lower than that of $B^\pm \to D\pi^\pm$, which is already expected to be small.

Kernel-estimated PDF shapes, extracted from simulated samples after the full selection, are shown in Figures 5.18, 5.19, 5.20 and 5.21, for correctly-reconstructed LL, partially-reconstructed LL, correctly-reconstructed DD and partially-reconstructed DD samples, respectively.

Combinatoric background

The combinatoric background component in each fit category is modelled by a decreasing exponential distribution,

$$F^{b\ n}_{\text{comb}}(m; N^{b\ n}_{\text{comb}}, \tau^{b\ n}) = N^{b\ n}_{\text{comb}} \cdot \mathrm{e}^{\tau^{b\ n} m |\tau^{b\ n}|}. \tag{5.66}$$

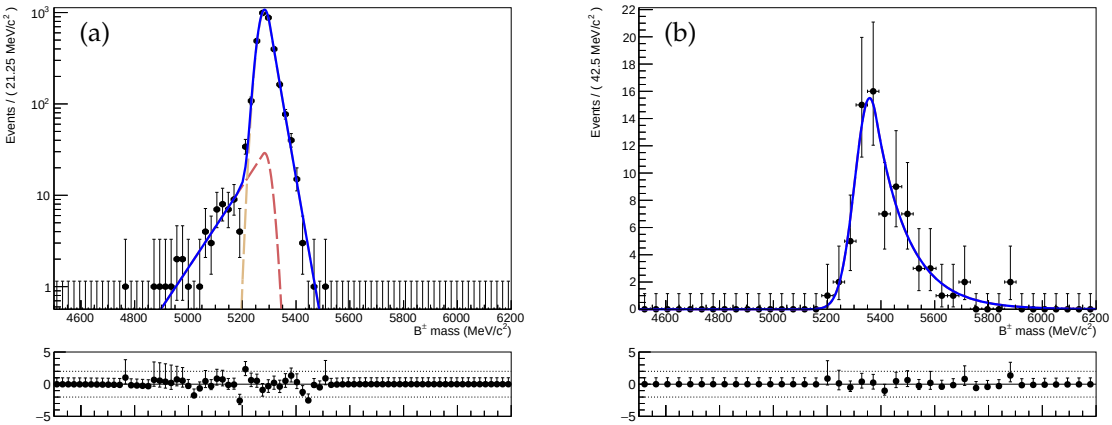


Figure 5.16.: Fitted B-mass PDF shapes used in the fit to both LL and DD $B^\pm \to D^* h^\pm$ correctly-reconstructed (*i.e.* $(n = \gamma || n = \pi^0)$) candidates. A Double Crystal Ball function is fitted to signal mode ($B^\pm \to D^* K^\pm$) MC (a). Cross-feed from control mode ($B^\pm \to D^* \pi^\pm$) MC, reconstructed in the signal mode, is fitted using a single Crystal Ball function (b). In each plot, histogram points show Monte Carlo simulation and a solid blue line gives the fitted PDF model. For plot (a), dashed red and yellow lines indicate the two constituent Crystal Ball shapes of the model. In each plot, the lower canvas shows the pull.

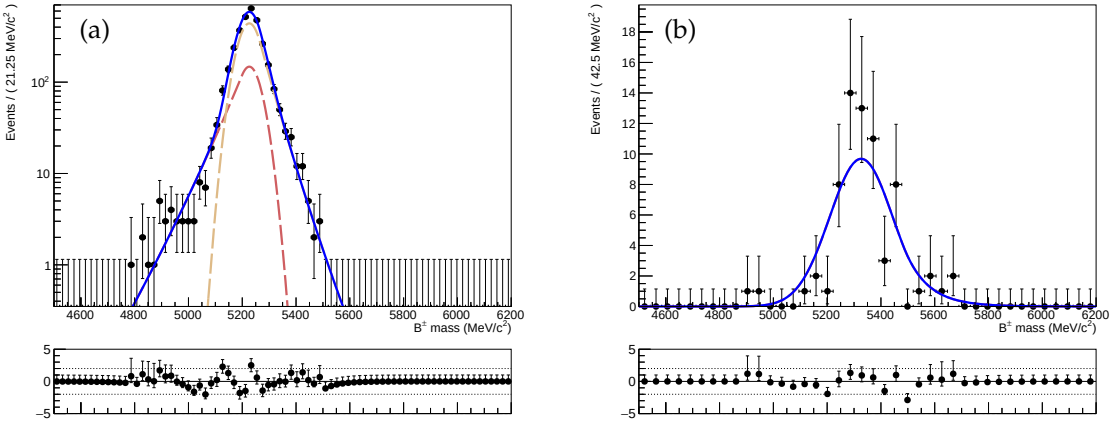


Figure 5.17.: Fitted B-mass PDF shapes used in the fit to both LL and DD $B^\pm \to D^* h^\pm$ partially-reconstructed (*i.e.* $n = \pi^0_{\text{part. reco.}}$) candidates. A Double Crystal Ball function is fitted to signal mode ($B^\pm \to D^* K^\pm$) MC (a). Cross-feed from control mode ($B^\pm \to D^* \pi^\pm$) MC, reconstructed in the signal mode, is fitted using a single Crystal Ball function (b). In each plot, histogram points show Monte Carlo simulation and a solid blue line gives the fitted PDF model. For plot (a), dashed red and yellow lines indicate the two constituent Crystal Ball shapes of the model. In each plot, the lower canvas shows the pull.

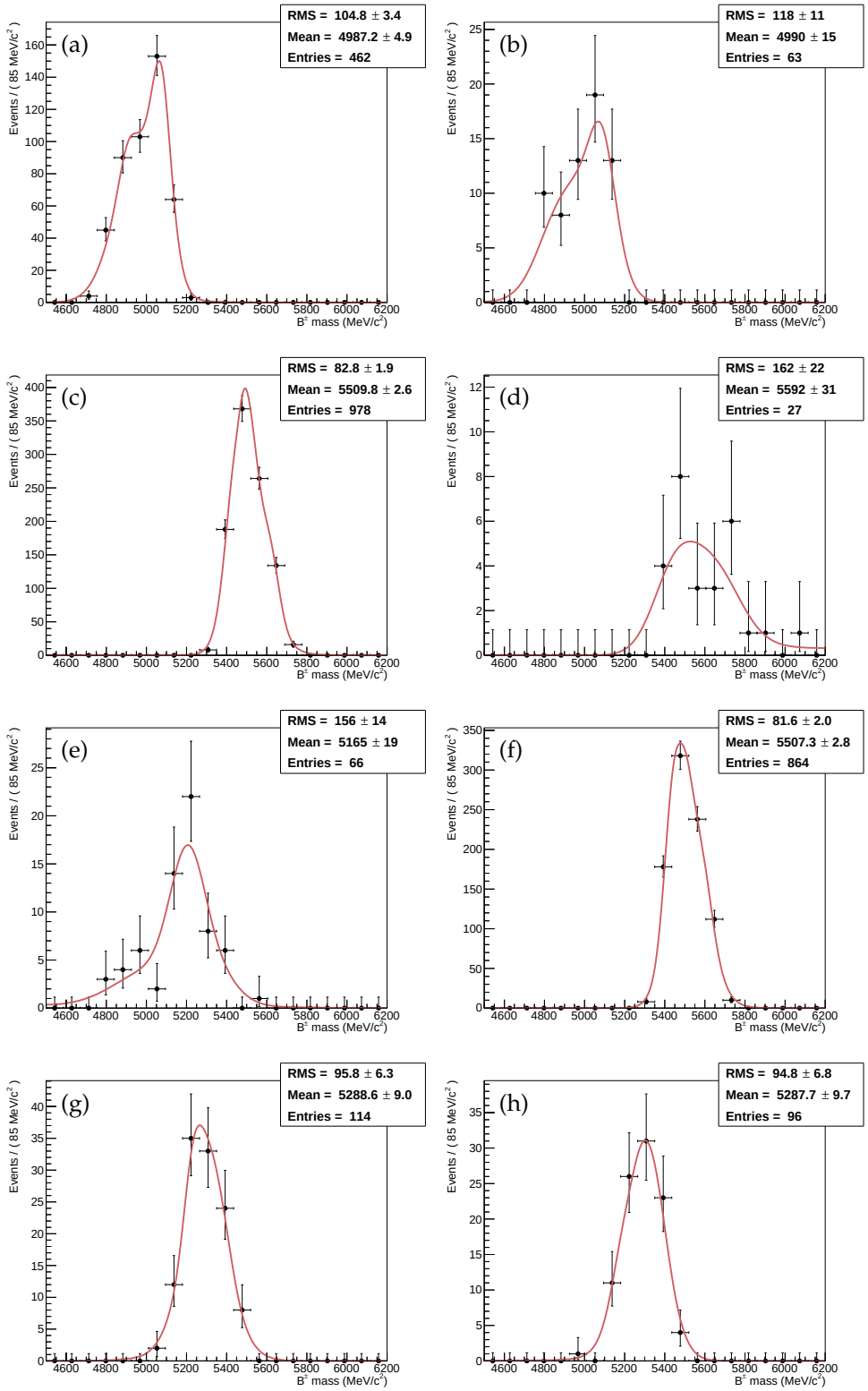


Figure 5.18.: Kernel-estimated B-mass PDF shapes used in the fit to LL $B^\pm \to D^* h^\pm$ correctly-reconstructed (*i.e.* $(n = \gamma || n = \pi^0)$) candidates. Shapes are obtained from: $B^\pm \to D^* \rho^\pm$ MC entering the $b = \pi$ category (a); $B^\pm \to D^* K^{*\pm}$ MC entering the $b = K$ category (b); $B^\pm \to D\pi^\pm$ MC entering the $b = \pi$ (c) and $b = K$ (d) categories; $B^\pm \to D\rho^\pm$ MC entering the $b = \pi$ category (e); $B^\pm \to DK^\pm$ MC entering the $b = K$ category (f); $B^0 \to D^{*\pm}\pi^\mp$ MC entering the $b = \pi$ category (g); and $B^0 \to D^{*\pm}K^\mp$ MC entering the $b = K$ category (h).

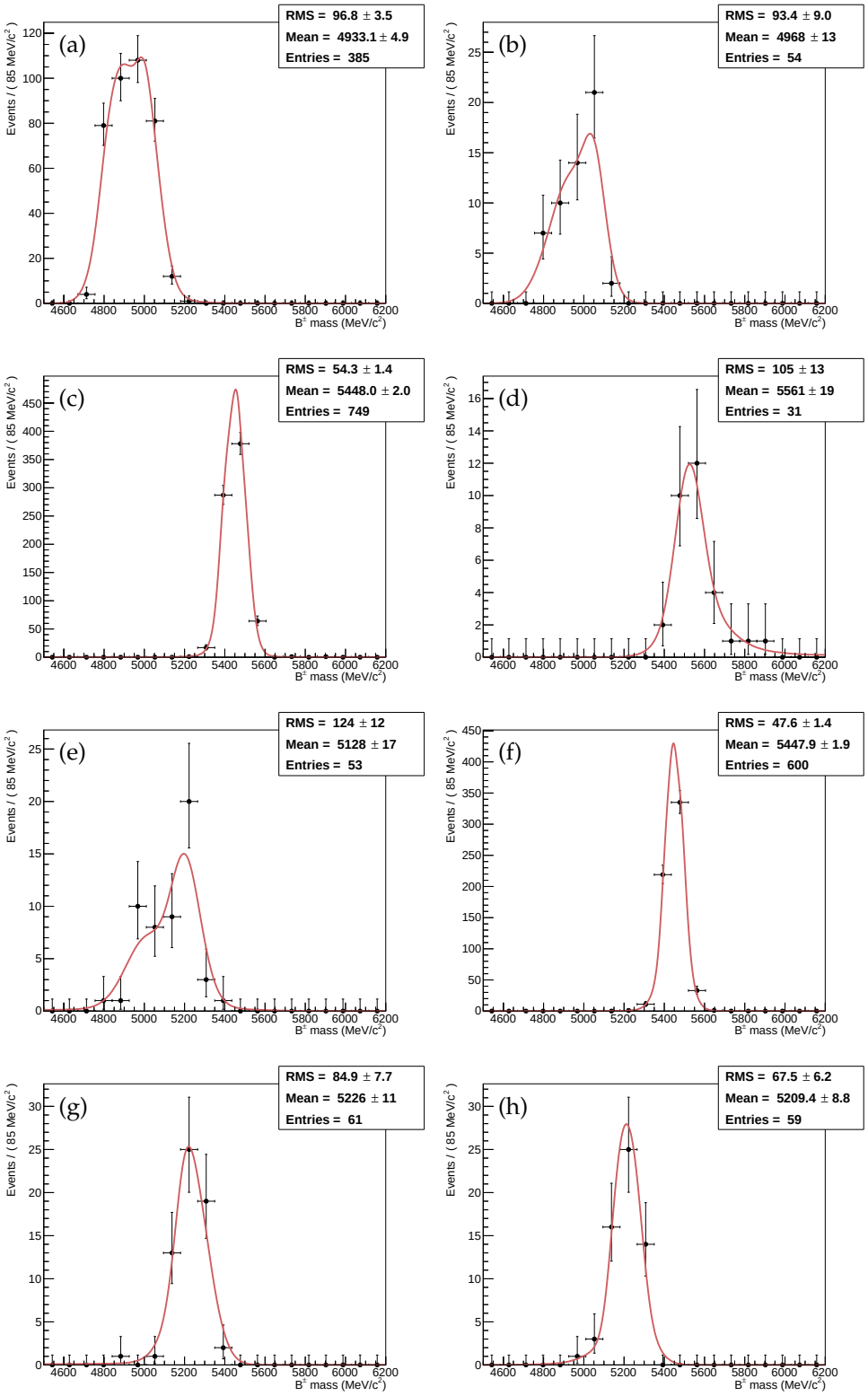


Figure 5.19.: Kernel-estimated B-mass PDF shapes used in the fit to LL $B^{\pm} \to D^{*}h^{\pm}$ partially-reconstructed (*i.e.* $n = \pi^{0}_{\text{part. reco.}}$) candidates. Shapes are obtained from: $B^{\pm} \to D^{*}\rho^{\pm}$ MC entering the $b = \pi$ category (a); $B^{\pm} \to D^{*}K^{*\pm}$ MC entering the $b = K$ category (b); $B^{\pm} \to D\pi^{\pm}$ MC entering the $b = \pi$ (c) and $b = K$ (d) categories; $B^{\pm} \to D\rho^{\pm}$ MC entering the $b = \pi$ category (e); $B^{\pm} \to DK^{\pm}$ MC entering the $b = K$ category (f); $B^{0} \to D^{*\pm}\pi^{\mp}$ MC entering the $b = \pi$ category (g); and $B^{0} \to D^{*\pm}K^{\mp}$ MC entering the $b = K$ category (h).

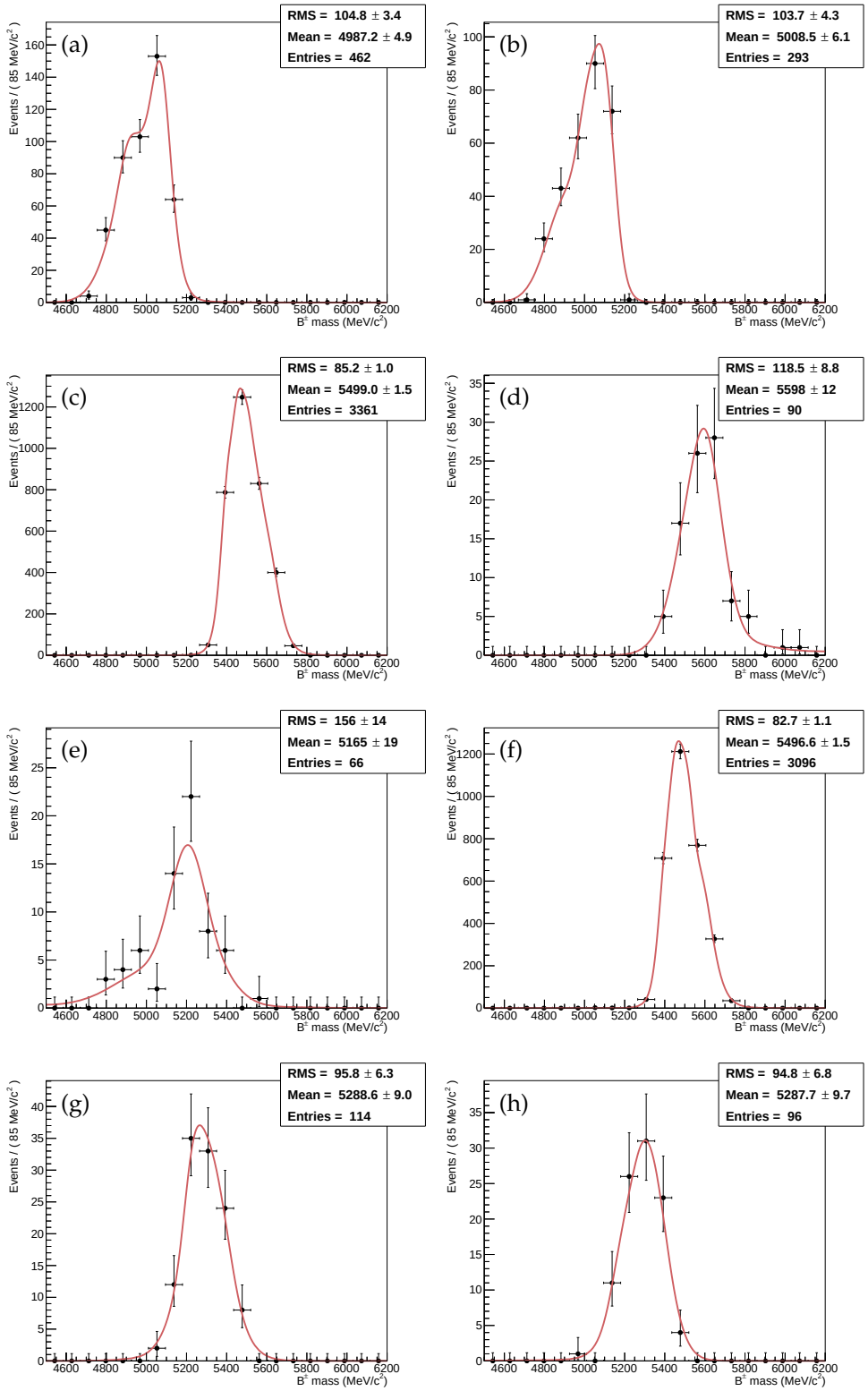


Figure 5.20.: Kernel-estimated B-mass PDF shapes used in the fit to DD $B^{\pm} \to D^{*}h^{\pm}$ correctly-reconstructed (*i.e.* $(n = \gamma || n = \pi^0)$) candidates. Shapes are obtained from: $B^{\pm} \to D^{*}\rho^{\pm}$ MC entering the $b = \pi$ category (a); $B^{\pm} \to D^{*}K^{*\pm}$ MC entering the $b = K$ category (b); $B^{\pm} \to D\pi^{\pm}$ MC entering the $b = \pi$ (c) and $b = K$ (d) categories; $B^{\pm} \to D\rho^{\pm}$ MC entering the $b = \pi$ category (e); $B^{\pm} \to DK^{\pm}$ MC entering the $b = K$ category (f); $B^0 \to D^{*\pm}\pi^{\mp}$ MC entering the $b = \pi$ category (g); and $B^0 \to D^{*\pm}K^{\mp}$ MC entering the $b = K$ category (h).

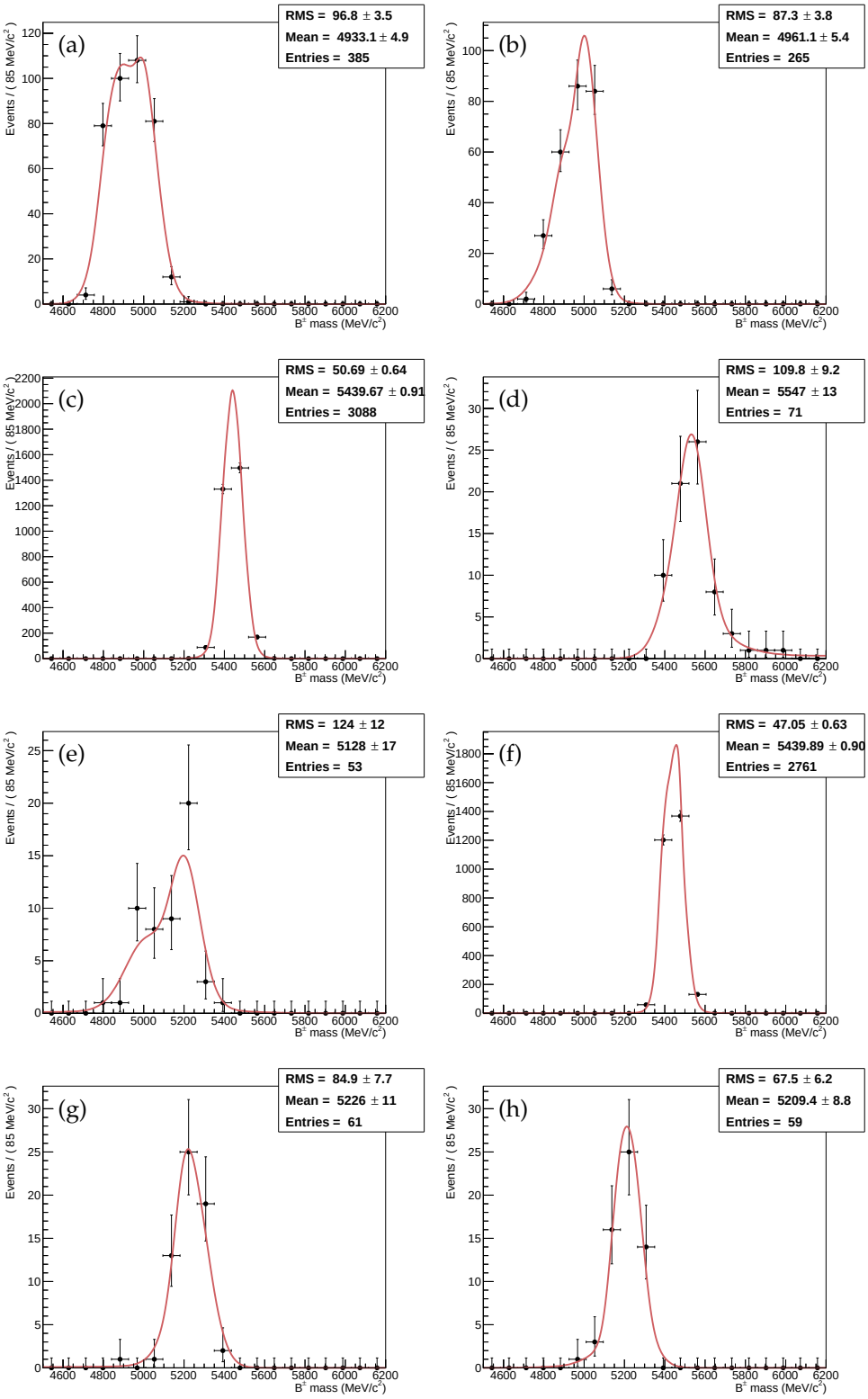


Figure 5.21.: Kernel-estimated B-mass PDF shapes used in the fit to DD $B^{\pm} \to D^{*}h^{\pm}$ partially-reconstructed (*i.e.* $n = \pi^{0}_{\text{part. reco.}}$) candidates. Shapes are obtained from: $B^{\pm} \to D^{*}\rho^{\pm}$ MC entering the $b = \pi$ category (a); $B^{\pm} \to D^{*}K^{*\pm}$ MC entering the $b = K$ category (b); $B^{\pm} \to D\pi^{\pm}$ MC entering the $b = \pi$ (c) and $b = K$ (d) categories; $B^{\pm} \to D\rho^{\pm}$ MC entering the $b = \pi$ category (e); $B^{\pm} \to DK^{\pm}$ MC entering the $b = K$ category (f); $B^{0} \to D^{*\pm}\pi^{\mp}$ MC entering the $b = \pi$ category (g); and $B^{0} \to D^{*\pm}K^{\mp}$ MC entering the $b = K$ category (h).

5.4.1. Constraints applied

The PDFs described so far combine to give the total model for the reconstructed B-meson mass, which may be written in terms of its free parameters as

$$\begin{aligned} F_{\text{tot}}^{\pi\,n} =& F_{\text{sig}}^{\pi\,n}(m; N_{\text{sig}}^{\pi\,n}, \mu^n, \sigma^{\pi\,n}) \\ &+ F_{\text{xfeed}}^{\pi\,n}(m; N_{\text{xfeed}}^{\pi\,n}, \mu_{\text{xfeed}}^{\pi\,n}, \sigma_{\text{xfeed}}^{\pi\,n}) \\ &+ N_{D^*\rho}^{\pi\,n} \cdot f_{D^*\rho}^{\pi\,n}(m) + N_{D\pi}^{\pi\,n} \cdot f_{D\pi}^{\pi\,n}(m) + N_{D\rho}^{\pi\,n} \cdot f_{D\rho}^{\pi\,n}(m) + N_{D^{*\pm}\pi}^{\pi\,n} \cdot f_{D^{*\pm}\pi}^{\pi\,n}(m) \\ &+ F_{\text{comb}}^{\pi\,n}(m; N_{\text{comb}}^{\pi\,n}, \tau^{\pi\,n}), \end{aligned} \tag{5.67}$$

for fit categories of bachelor hypothesis $b = \pi$, and

$$\begin{aligned} F_{\text{tot}}^{K\,n} =& F_{\text{sig}}^{K\,n}(m; N_{\text{sig}}^{K\,n}, \mu^n, \sigma^{K\,n}) \\ &+ F_{\text{xfeed}}^{K\,n}(m; N_{\text{xfeed}}^{K\,n}, \mu_{\text{xfeed}}^{K\,n}, \sigma_{\text{xfeed}}^{K\,n}) \\ &+ N_{D^*K^*}^{K\,n} \cdot f_{D^*K^*}^{K\,n}(m) + N_{D\pi}^{K\,n} \cdot f_{D\pi}^{K\,n}(m) + N_{DK}^{K\,n} \cdot f_{DK}^{K\,n}(m) + N_{D^{*\pm}K}^{K\,n} \cdot f_{D^{*\pm}K}^{K\,n}(m) \\ &+ F_{\text{comb}}^{K\,n}(m; N_{\text{comb}}^{K\,n}, \tau^{K\,n}), \end{aligned} \tag{5.68}$$

for fit categories of bachelor hypothesis $b = K$.

Given the three categories defined for the neutral hypothesis, n, this implies that the model contains 75 degrees of freedom. In order to obtain a reliable fit, constraints are applied, such that in the end only 20 free parameters are needed. The constraints applied are described in turn:

- a common mean, μ is used for all correctly-reconstructed signal components, and for the partially-reconstructed signal components the relative shift is fixed from simulation:

$$\mu^{\gamma} = \mu^{\pi^0} = \mu, \quad \text{and} \tag{5.69}$$

$$\mu^{\pi^0_{\text{part.reco.}}} = \mu + \delta m^{\text{part.reco.}}, \quad \text{where} \tag{5.70}$$

$$\delta m^{\text{part.reco.}} = (\mu_{\text{MC}}^{\pi^0_{\text{part.reco.}}} - \mu_{\text{MC}}^{\gamma||\pi^0}); \tag{5.71}$$

- for each $b = K$ category, the shift of the cross-feed shape, relative to the corresponding signal mean, is fixed from simulation. For the $b = \pi$ categories, where no fit is performed to simulated cross-feed, the shift is taken to be the same magnitude as the corresponding $b = K$ shift, but in the opposite direction:

$$\mu_{\text{xfeed}}^{K\,\gamma||\pi^0} = \mu + \delta m_{\text{xfeed}}^{\text{cor.reco.}}, \tag{5.72}$$

$$\mu_{\text{xfeed}}^{K\,\pi^0_{\text{part.reco.}}} = \mu + \delta m_{\text{xfeed}}^{\text{part.reco.}}, \tag{5.73}$$

$$\mu_{\text{xfeed}}^{\pi\,\gamma||\pi^0} = \mu - \delta m_{\text{xfeed}}^{\text{cor.reco.}}, \quad \text{and} \tag{5.74}$$

$$\mu_{\text{xfeed}}^{\pi\,\pi^0_{\text{part.reco.}}} = \mu - \delta m_{\text{xfeed}}^{\text{part.reco.}}, \quad \text{where} \tag{5.75}$$

$$\delta m^{\text{cor.reco.}}_{\text{xfeed}} = (\mu^{\gamma||\pi^0}_{\text{xfeed MC}} - \mu^{\gamma||\pi^0}_{\text{MC}}) \quad \text{and} \tag{5.76}$$

$$\delta m^{\text{part.reco.}}_{\text{xfeed}} = (\mu^{\pi^0_{\text{part.reco.}}}_{\text{xfeed MC}} - \mu^{\pi^0_{\text{part.reco.}}}_{\text{MC}}); \tag{5.77}$$

- common width parameters, σ^n, are used for the signal shapes of both bachelor categories in the fit:

$$\sigma^{K\,n} = \sigma^{\pi\,n} = \sigma^n; \tag{5.78}$$

- the ratios of the widths of the different signal shapes, and of the shapes describing cross-feed in the $b = K$ categories, are fixed to match the ratios found in the fits to simulation, by introducing a common 'smearing factor' parameter, k_σ:

$$\sigma^n = k_\sigma \sigma^{K\,n}_{\text{MC}}; \quad \text{and} \tag{5.79}$$

$$\sigma^{K\,n}_{\text{xfeed}} = k_\sigma \sigma^{K\,n}_{\text{xfeed MC}}; \tag{5.80}$$

- for the cross-feed shapes in the $b = \pi$ categories, where corresponding fits to simulation are not performed, the ratio of the width of the cross-feed shape to the signal shape is taken from the relative widths of the corresponding components in the $B^\pm \to D\pi^\pm$ mass fit (see Section 5.1). This ratio is taken to be common for all neutral hypotheses, and is parameterized by r^π_σ:

$$\sigma^{\pi\,n}_{\text{xfeed}} = r^\pi_\sigma \sigma^n; \tag{5.81}$$

- the ratio of yields of $b = K$ signal to $b = \pi$ signal is taken to be equal for each neutral category, since the m_3 requirements used to define each neutral hypothesis should not affect the distribution of the $\text{DLL}_{K\pi}(\text{bach})$ variable used to separate different bachelor hypotheses (as m_3 depends on properties of the D^* meson decay, whilst $\text{DLL}_{K\pi}(\text{bach})$ depends only on properties of the bachelor track). The common ratio is parameterized by $r_{K\pi}$:

$$N^{K\,n}_{\text{sig}} = r_{K\pi} N^{\pi\,n}_{\text{sig}}; \tag{5.82}$$

- the cross-feed yields, for samples of the same neutral hypothesis but different bachelor hypotheses, are linked to the corresponding signal yields using the known mis-indentification efficiency of the requirement used to separate the two bachelor hypotheses ($\text{DLL}_{K\pi}(\text{bach}) \gtrless 4$), calculated using the standard LHCb simulation calibration package, `PIDCalib` [148]:

$$N^{K\,n}_{\text{xfeed}} = \epsilon^{\text{PID}}_{\pi\to K} N^{\pi\,n}_{\text{sig}}; \quad \text{and} \tag{5.83}$$

$$N^{\pi\,n}_{\text{xfeed}} = \epsilon^{\text{PID}}_{K\to\pi} N^{K\,n}_{\text{sig}}; \tag{5.84}$$

- the ratio between the yield of $B^\pm \to D^*\rho^\pm$ background (included only for the $b = \pi$ categories) and the signal yield is taken to be equal for both $n = \pi^0$ and $n = \pi^0_{\text{part.reco.}}$ neutral categories, since the true signal decay is the same in each case, and equally susceptible to background containing

an additional π^0-meson. The common ratio is parameterized by $r_\rho^{\pi^0}$:

$$N_{D^*\rho}^{\pi\,\pi^0} = r_\rho^{\pi^0} N_{\text{sig}}^{\pi\,\pi^0}; \quad \text{and} \tag{5.85}$$

$$N_{D^*\rho}^{\pi\,\pi^0_{\text{part.reco.}}} = r_\rho^{\pi^0} N_{\text{sig}}^{\pi\,\pi^0_{\text{part.reco.}}}; \tag{5.86}$$

- the ratio of the $B^\pm \to D^* K^{*\pm}$ background (included only for the $b = K$ categories) and the signal yield is treated in an analogous way, but with a common ratio used for all three neutral hypotheses, parameterized by $r_{K^*}^{\pi^0}$:

$$N_{D^*K^*}^{K\,n} = r_{K^*}^{\pi^0} N_{\text{sig}}^{K\,n}; \tag{5.87}$$

- the yields of the $B^\pm \to D\pi^\pm$ background in the $b = K$ categories are constrained relative to those in the $b = \pi$ categories, using the known mis-identification efficiency of the PID requirement used to separate them:

$$N_{D\pi}^{K\,n} = \epsilon_{\pi\to K}^{\text{PID}} N_{D\pi}^{\pi\,n}; \tag{5.88}$$

- the yields of the $B^\pm \to DK^\pm$ background in the $b = K$ categories are constrained relative to the $B^\pm \to D\pi^\pm$ yields in the same categories, assuming their relative efficiencies match those of the signal ($B^\pm \to D^* K^\pm$) and cross-feed ($B^\pm \to D^*\pi^\pm$) modes:

$$N_{DK}^{K\,n} = N_{D\pi}^{K\,n} \cdot \frac{N_{\text{sig}}^{K\,n}}{N_{\text{xfeed}}^{K\,n}}; \tag{5.89}$$

- the yields of the small peaking backgrounds in the $b = \pi$ categories, $B^\pm \to D\rho^\pm$ and $B^0 \to D^{*\pm}\pi^\mp$, are fixed relative to the yield of the largest peaking background component, $B^\pm \to D^*\rho^\pm$, using their respective branching fractions and efficiencies, estimated from MC. These values are summarised in Table 5.21.

$$N_{D\rho}^{\pi\,n} = N_{D^*\rho}^{\pi\,n} \cdot \frac{\Gamma(B^\pm \to D\rho^\pm)\varepsilon_{D\rho}^{\pi\,n}}{\Gamma(B^\pm \to D^*\rho^\pm)\varepsilon_{D^*\rho}^{\pi\,n}}, \tag{5.90}$$

$$N_{D^{*\pm}\pi}^{\pi\,n} = N_{D^*\rho}^{\pi\,n} \cdot \frac{\sigma(B^0)\Gamma(B^0 \to D^{*\pm}\pi^\mp)\varepsilon_{D^{*\pm}\pi}^{\pi\,n}}{\sigma(B^\pm)\Gamma(B^\pm \to D^*\rho^\pm)\varepsilon_{D^*\rho}^{\pi\,n}}, \tag{5.91}$$

where $\sigma(B^0) = 38.1\ \mu b$ and $\sigma(B^\pm) = 38.9\ \mu b$ are the measured cross-sections for neutral and charged B-meson production at LHCb [149]. The same efficiencies are used for both of the correctly-reconstructed categories, $n = \gamma$ and $n = \pi^0$;

- the ratio of $B^0 \to D^{*\pm}K^\mp$ background to signal, in the $b = K$ categories, is assumed to match that between $B^0 \to D^{*\pm}\pi^\mp$ and π signal in the $b = \pi$:

$$N_{D^{*\pm}K}^{K\,n} = N_{\text{sig}}^{K\,n} \cdot \frac{N_{D^{*\pm}\pi}^{\pi\,n}}{N_{\text{sig}}^{\pi\,n}}; \quad \text{and} \tag{5.92}$$

- the yield of $B^\pm \to D\pi^\pm$ background in the $(b,n) = (\pi,\gamma)$ category is fixed relative to that in the $(b,n) = (\pi,\pi^0_{\text{part.reco.}})$ category, such that the ratio matches the value obtained in simulation. The fixed ratio is parameterized by $r^{\gamma\ \pi^0_{\text{part.reco.}}}_{D\pi\ \text{MC}}$:

$$N^{\pi\ \gamma}_{D\pi} = r^{\gamma\ \pi^0_{\text{part.reco.}}}_{D\pi\ \text{MC}} N^{\pi\ \pi^0_{\text{part.reco.}}}_{D\pi}. \tag{5.93}$$

  A single value for $r^{\gamma\ \pi^0_{\text{part.reco.}}}_{D\pi\ \text{MC}}$ is used in the separate fits to candidates with LL and DD K^0_S-meson candidates, since the value obtained by combining the corresponding simulated samples is consistent with each of those obtained separately, and gives a lower statistical uncertainty.

Mode	Quantity	correctly-reconstructed signal		partially-reconstructed signal	
		LL	DD	LL	DD
$B^\pm \to D^*\rho^\pm$	Γ	9.8×10^{-3}			
	ε_{gen}	$(14.8 \pm 0.03)\%$			
	$\varepsilon_{\text{strip}}$	$(0.408 \pm 0.004)\%$			
	ε_{sel}	$(0.60 \pm 0.08)\%$	$(3.1 \pm 0.2)\%$	$(0.47 \pm 0.07)\%$	$(2.7 \pm 0.2)\%$
$B^\pm \to D\rho^\pm$	Γ	13.4×10^{-3}			
	ε_{gen}	$(15.9 \pm 0.04)\%$			
	$\varepsilon_{\text{strip}}$	$(0.107 \pm 0.003)\%$			
	ε_{sel}	$(0.55 \pm 0.18)\%$	$(2.0 \pm 0.4)\%$	$(0.37 \pm 0.15)\%$	$(2.1 \pm 0.4)\%$
$B^0 \to D^{*\pm}\pi^\mp$	Γ	1.9×10^{-3}			
	ε_{gen}	$(17.0 \pm 0.06)\%$			
	$\varepsilon_{\text{strip}}$	$(0.257 \pm 0.007)\%$			
	ε_{sel}	$(1.5 \pm 0.3)\%$	$(4.8 \pm 0.6)\%$	$(0.42 \pm 0.17)\%$	$(3.0 \pm 0.5)\%$

Table 5.21.: Known branching fractions [3] and measured efficiencies obtained in Monte Carlo simulation, used to link $B^\pm \to D^{*0}\pi^\pm$ background yields.

5.4.2. Mass fit result

The constant parameters of the fit are all common to candidates containing both LL and DD K^0_S. Their fixed values are summarized in Table 5.22. The simultaneous invariant mass fit, described above, is performed separately to $B^\pm \to D^*h^\pm$ candidates reconstructed with LL and DD K^0_S track types. Fits to the full 2011 and 2012 LHCb dataset are displayed for each fit category in Figures 5.22, 5.24 and 5.26, for LL K^0_S candidates, and 5.23, 5.25 and 5.27, for DD K^0_S candidates. The free parameters of the fit and their returned values are given in Table 5.23 (LL) and Table 5.24 (DD). Correlations between the free parameters are illustrated in Figure 5.28.

Some of the fits to reconstructed $B^\pm \to D^*\pi^\pm$ candidates (shown in the lower part of Figures 5.22–5.27) show pulls of relatively large magnitudes, indicating statistically significant deviations between the observed bin populations and those predicted by the fitted model. However, these fits

are performed solely to provide constraints on background yields and shapes entering the γ-sensitive $B^{\pm} \to D^{*}K^{\pm}$ samples, and the fits to these samples do not exhibit large pulls.

Only candidates falling within $\pm 3\sigma$ windows centred on each $B^{\pm} \to D^{*}K^{\pm}$ signal peak are propagated forward to the full $\mathcal{CP}$ fit. These windows are marked by vertical lines in the plotted fits. For LL candidates in the correctly-reconstructed signal modes (*i.e.* $(n = \gamma || n = \pi^{0})$), the window used is $[5220; 5351]$ MeV. For DD candidates in the correctly-reconstructed signal modes, the window used is $[5220; 5352]$ MeV. For LL candidates in the partially-reconstructed signal mode (*i.e.* $n = \pi^{0}_{\text{part.reco.}}$), the window used is $[5108; 5349]$ MeV. For DD candidates in the partially-reconstructed signal mode, the window used is $[5107; 5351]$ MeV. Component yields extracted within these windows are enumerated in Table 5.25.

Constant Parameter	Value	Description
$\alpha_1^{\gamma\|\|\pi^0}$	0.233	Double Crystal Ball parameters for correctly-reconstructed signal PDF
$\alpha_2^{\gamma\|\|\pi^0}$	0.886	
$n_1^{\gamma\|\|\pi^0}$	112	
$n_2^{\gamma\|\|\pi^0}$	131	
$f^{\gamma\|\|\pi^0}$	0.0536	
$\sigma_{\text{MC}}^{\gamma\|\|\pi^0}$	22.2 MeV	
$\alpha_1^{\pi^0_{\text{part.reco.}}}$	0.670	Double Crystal Ball parameters for partially-reconstructed signal PDF
$\alpha_2^{\pi^0_{\text{part.reco.}}}$	1.03	
$n_1^{\pi^0_{\text{part.reco.}}}$	22.4	
$n_2^{\pi^0_{\text{part.reco.}}}$	3.26	
$f^{\pi^0_{\text{part.reco.}}}$	0.278	
$\sigma_{\text{MC}}^{\pi^0_{\text{part.reco.}}}$	40.9 MeV	
$\alpha_{\text{xfeed}}^{K\,\gamma\|\|\pi^0}$	0.460	Crystal Ball parameters for correctly-reconstructed crossfeed PDF
$n_{\text{xfeed}}^{K\,\gamma\|\|\pi^0}$	126	
$\sigma_{\text{xfeed MC}}^{K\,\gamma\|\|\pi^0}$	55.3 MeV	
$\alpha_{\text{xfeed}}^{K\,\pi^0_{\text{part.reco.}}}$	1.34	Crystal Ball parameters for partially-reconstructed crossfeed PDF
$n_{\text{xfeed}}^{K\,\pi^0_{\text{part.reco.}}}$	144	
$\sigma_{\text{xfeed MC}}^{K\,\pi^0_{\text{part.reco.}}}$	117 MeV	
$\delta m^{\text{part.reco.}}$	-56.9 MeV	Shift of partially-reconstructed signal PDF relative to correctly-reconstructed
$\delta m^{\text{cor.reco.}}_{\text{xfeed}}$	74.8 MeV	Shift of cross-feed PDF relative to signal, for correctly-reconstructed neutral candidates
$\delta m^{\text{part.reco.}}_{\text{xfeed}}$	100 MeV	Shift of cross-feed PDF relative to signal, for partially-reconstructed neutral candidates
$\epsilon^{\text{PID}}_{K\to\pi}$	0.185	Efficiency of $K \to \pi$ bachelor mis-identification
$\epsilon^{\text{PID}}_{\pi\to K}$	0.0389	Efficiency of $\pi \to K$ bachelor mis-identification
$r_{D\pi\ \text{MC}}^{\gamma\ \pi^0_{\text{part.reco.}}}$	1.15	Ratio of $B^\pm \to D\pi^\pm$ background yields in $X^0 = \gamma$ and $x^0 = \pi^0_{\text{part.reco.}}$ samples

Table 5.22.: Constant parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected $B^\pm \to D^* h^\pm$ candidates. All parameters listed are common to samples with both LL and DD K^0_S candidates.

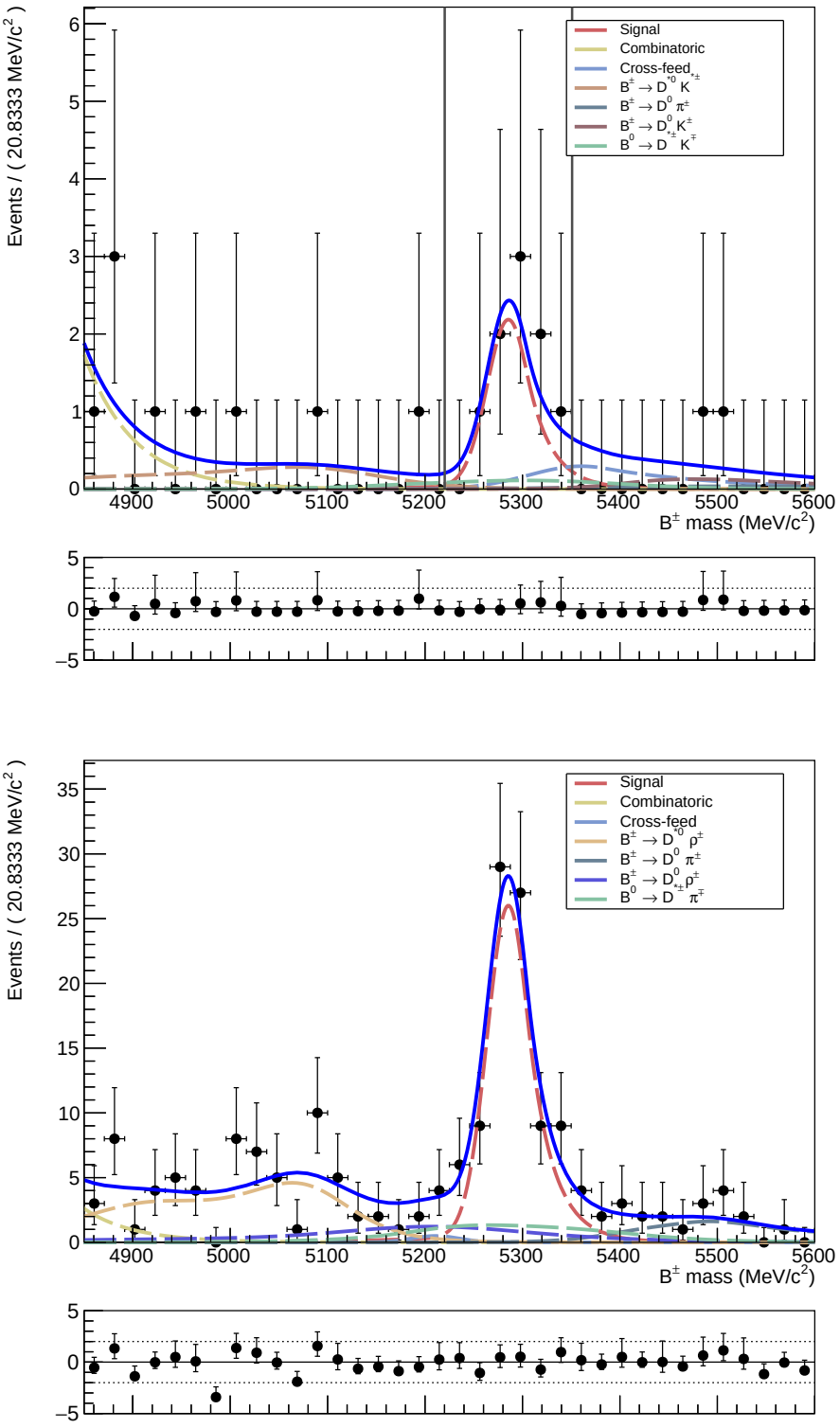


Figure 5.22.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for LL $B^{\pm} \to D^{*}K^{\pm}$ (top) and $B^{\pm} \to D^{*}\pi^{\pm}$ (bottom) candidates satisfying the $D^{*} \to D\pi^{0}$ (*i.e.* $n = \pi^{0}$) hypothesis. In each plot, the lower canvas shows the pull.

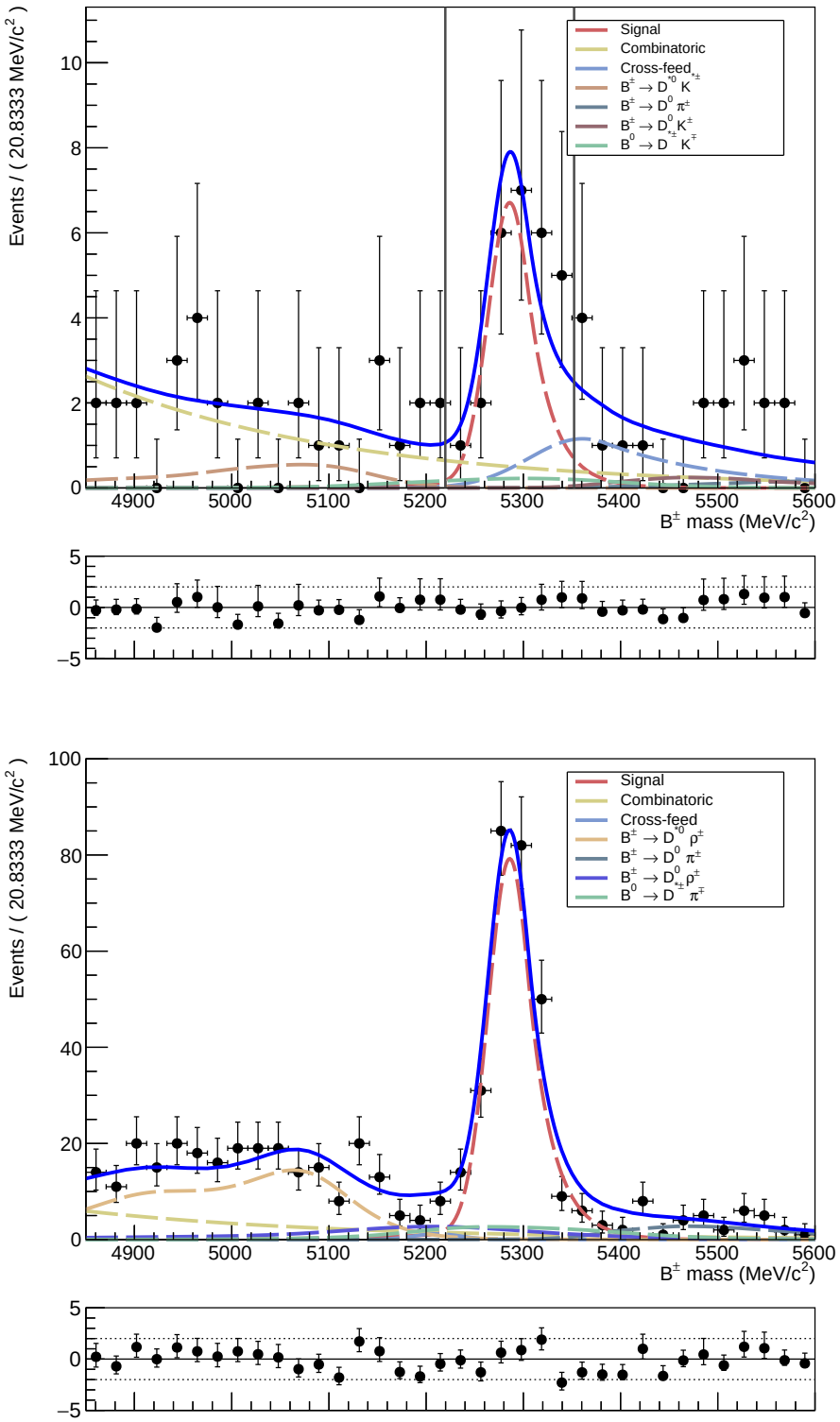


Figure 5.23.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for DD $B^{\pm} \to D^{*}K^{\pm}$ (top) and $B^{\pm} \to D^{*}\pi^{\pm}$ (bottom) candidates satisfying the $D^{*} \to D\pi^{0}$ (*i.e.* $n = \pi^{0}$) hypothesis. In each plot, the lower canvas shows the pull.

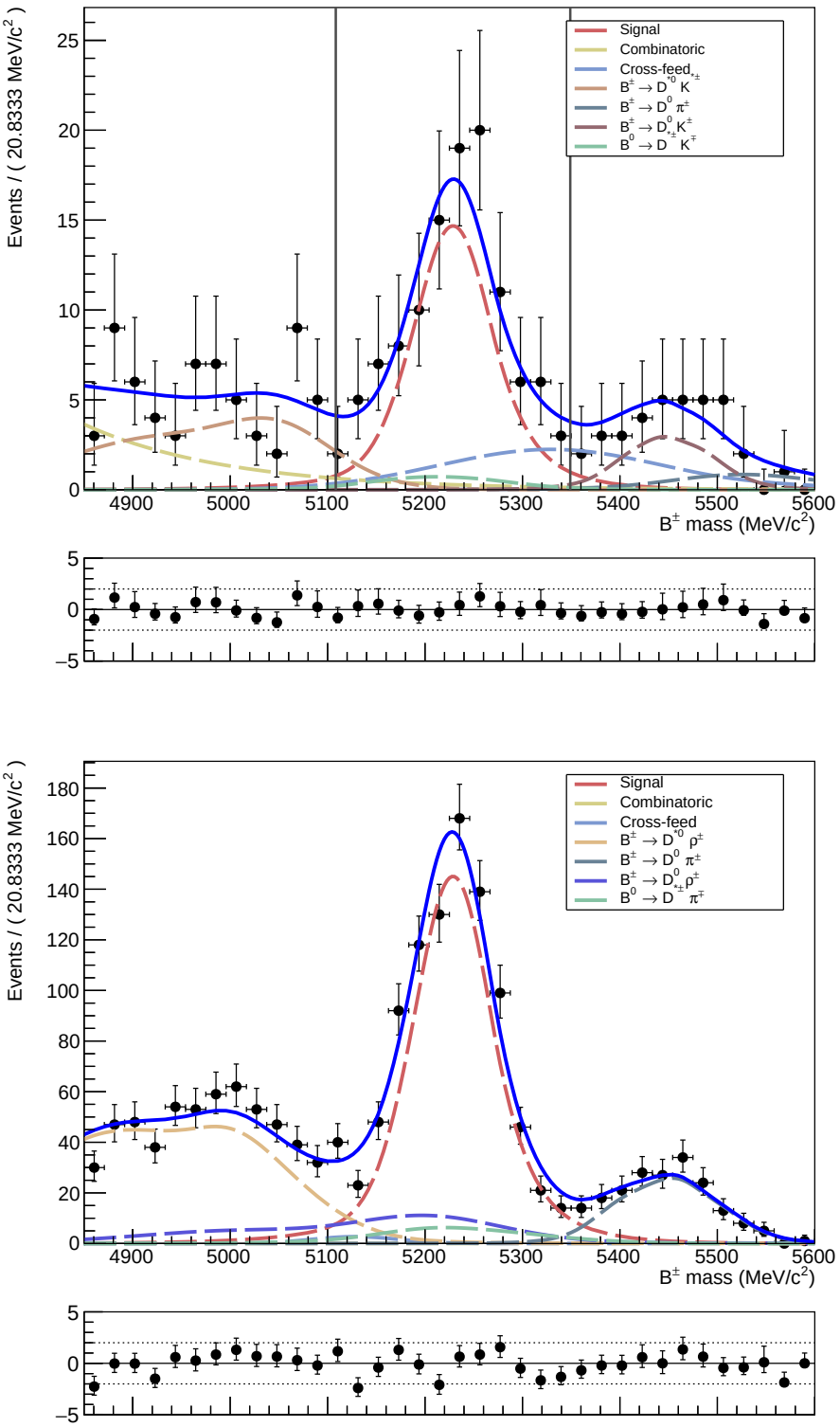


Figure 5.24.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for LL $B^{\pm} \to D^{*}K^{\pm}$ (top) and $B^{\pm} \to D^{*}\pi^{\pm}$ (bottom) candidates satisfying the partially-reconstructed $D^{*} \to D\pi^{0}, \pi^{0} \to \gamma\gamma$ (*i.e.* $n = \pi^{0}_{\text{part. reco.}}$) hypothesis. In each plot, the lower canvas shows the pull.

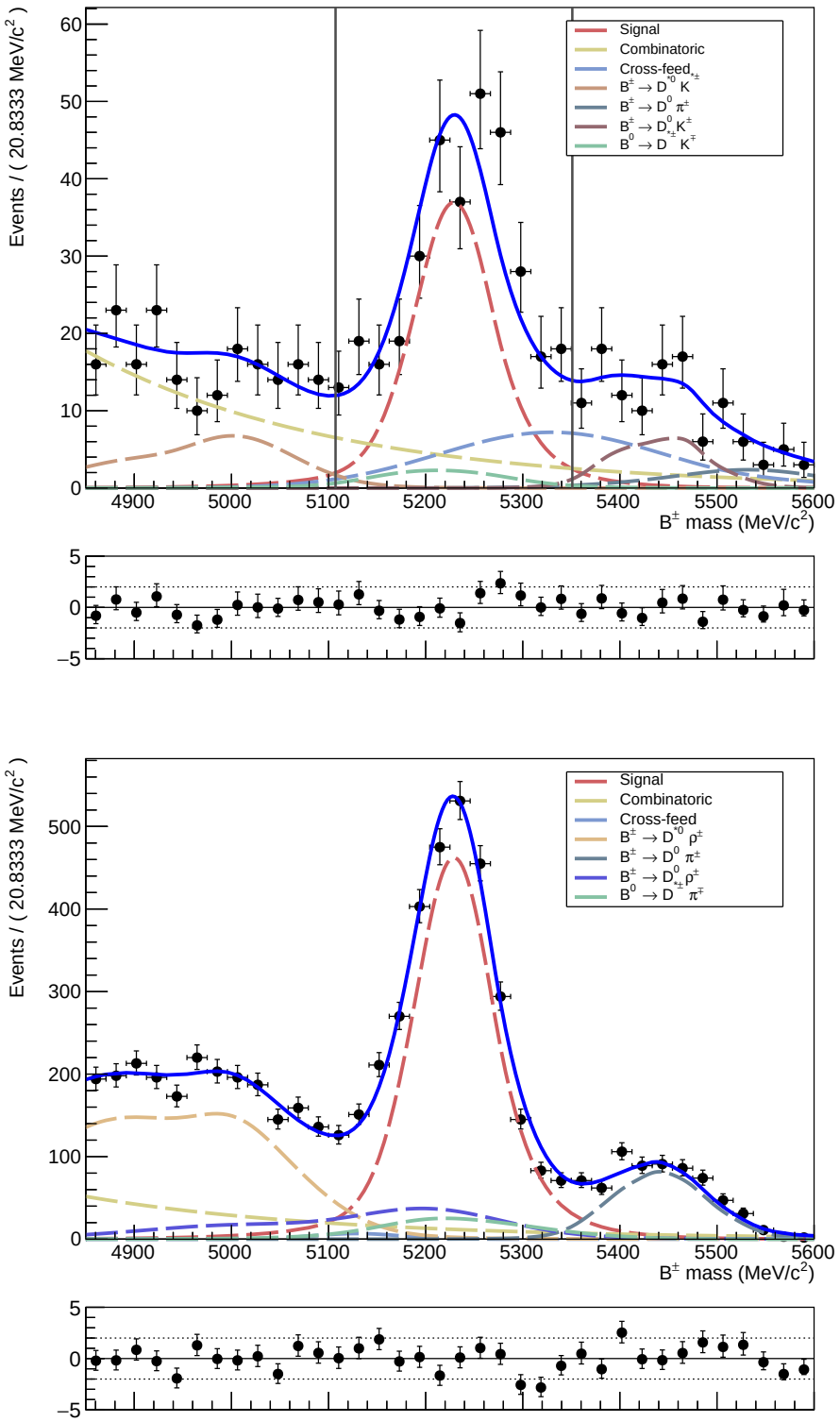


Figure 5.25.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for DD $B^\pm \to D^* K^\pm$ (top) and $B^\pm \to D^* \pi^\pm$ (bottom) candidates satisfying the partially-reconstructed $D^* \to D\pi^0, \pi^0 \to \gamma\gamma$ (*i.e.* $n = \pi^0_{\text{part. reco.}}$) hypothesis. In each plot, the lower canvas shows the pull.

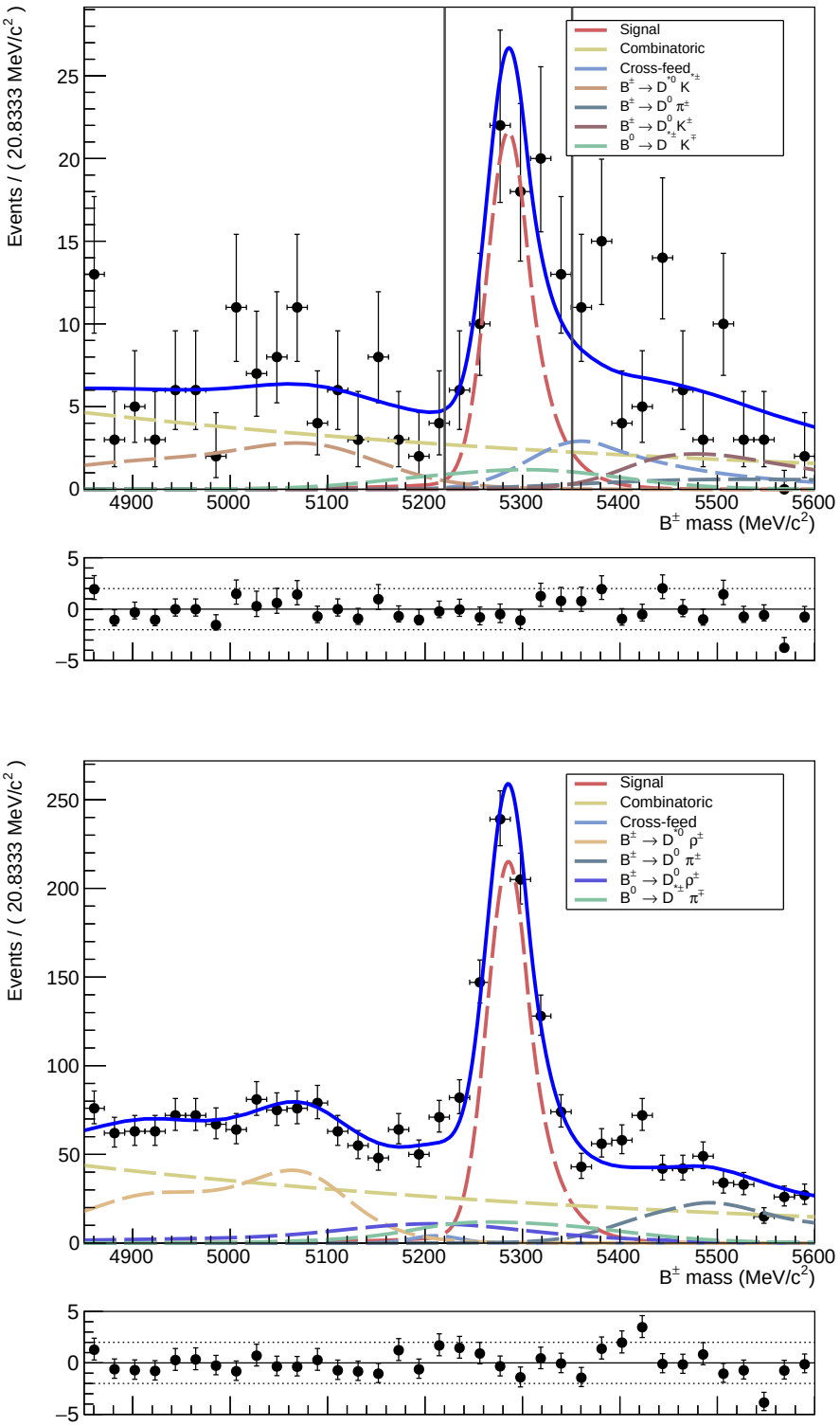


Figure 5.26.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for LL $B^{\pm} \to D^{*}K^{\pm}$ (top) and $B^{\pm} \to D^{*}\pi^{\pm}$ (bottom) candidates satisfying the $D^{*} \to D\gamma$ (*i.e.* $n = \gamma$) hypothesis. In each plot, the lower canvas shows the pull.

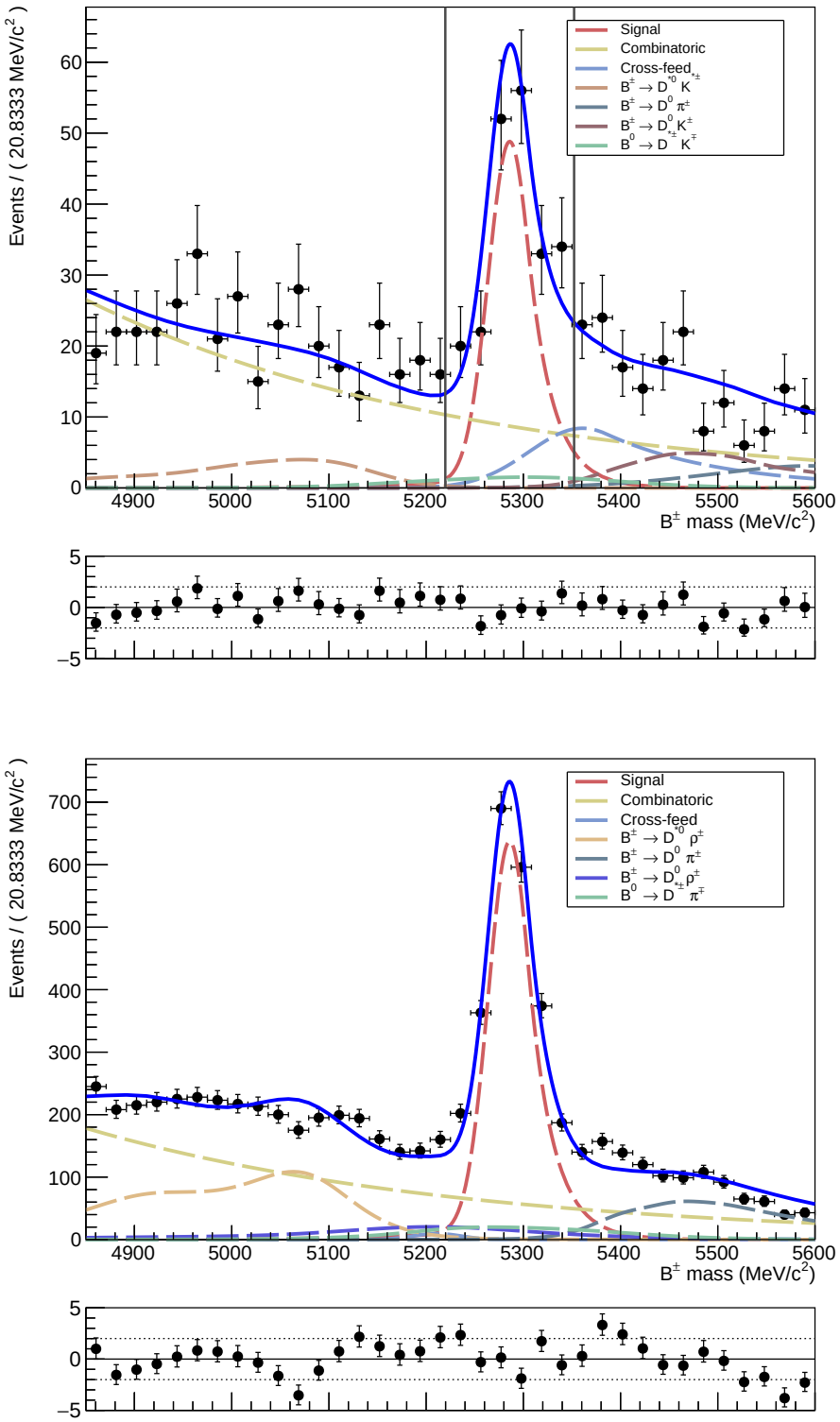


Figure 5.27.: Fit to reconstructed B-mass spectrum, and corresponding pulls, for DD $B^{\pm} \to D^{*}K^{\pm}$ (top) and $B^{\pm} \to D^{*}\pi^{\pm}$ (bottom) candidates satisfying the $D^{*} \to D\gamma$ (*i.e.* $n = \gamma$) hypothesis. In each plot, the lower canvas shows the pull.

Free Parameter	Designation	Fitted Value	Description
μ	`mean`	5286^{+1}_{-1} MeV	Central value of $X^0 = \pi^0, \gamma$ signal peaks
k_σ	`width_smear_factor`	$0.966^{+0.03}_{-0.03}$	Ratio of $\mathcal{CB}$ widths in data w.r.t. simulation
r_σ^π	`xfeedinDstPi_widthratio`	5^{+5}_{-3}	Ratio of cross-feed to signal widths, for $h = \pi$
$r_{K\pi}$	`Ka_Pi_ratio`	$0.10^{+0.01}_{-0.01}$	Ratio of $h = K$ and $h = \pi$ signal yields
$r_\rho^{\pi^0}$	`rho_wrt_sig_ratio`	$0.81^{+0.05}_{-0.05}$	Ratio of $D^*\rho^\pm$ to signal yields, for $X^0 = \pi^0, \pi^0_{\text{part.reco.}}$
$r_{K^*}^{\pi^0}$	`kst_wrt_sig_ratio`	$0.6^{+0.2}_{-0.2}$	Ratio of $D^*K^{*\pm}$ to signal yields, for $X^0 = \pi^0, \pi^0_{\text{part.reco.}}$
$N_{\text{sig}}^{\pi\ \gamma}$	`Pi_gam_sig_nB`	620^{+30}_{-30}	Signal yield for $h = \pi, X^0 = \gamma$ sample
$N_{\text{sig}}^{\pi\ \pi^0}$	`Pi_pi0_sig_nB`	76^{+10}_{-9}	Signal yield for $h = \pi, X^0 = \pi^0$ sample
$N_{\text{sig}}^{\pi\ \pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_sig_nB`	970^{+30}_{-30}	Signal yield for $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ sample
$N_{D^*\rho}^{\pi\ \gamma}$	`Pi_gam_rho_nB`	160^{+60}_{-60}	Yield of $D^*\rho^\pm$ in $h = \pi, X^0 = \gamma$ sample
$N_{D\pi}^{\pi\ \pi^0}$	`Pi_pi0_dpi_nB`	12^{+7}_{-7}	Yield of $D\pi^\pm$ in $h = \pi, X^0 = \pi^0$ sample
$N_{D\pi}^{\pi\ \pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_dpi_nB`	170^{+20}_{-10}	Yield of $D\pi^\pm$ in $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ sample
$\tau^{\pi\ \pi^0}$	`Pi_pi0_Comb_tau`	$-0.002^{+0.002}_{-0.004}$	Slope of $h = \pi, X^0 = \pi^0$ combinatoric
$N_{\text{comb}}^{\pi\ \pi^0}$	`Pi_pi0_nComb`	50^{+20}_{-20}	Yield of $h = \pi, X^0 = \pi^0$ combinatoric
$\tau^{K\ \pi^0}$	`Ka_pi0_Comb_tau`	$-0.011^{+0.005}_{-0.009}$	Slope of $h = K, X^0 = \pi^0$ combinatoric
$N_{\text{comb}}^{K\ \pi^0}$	`Ka_pi0_nComb`	7^{+4}_{-3}	Yield of $h = K, X^0 = \pi^0$ combinatoric
$\tau^{\pi\ \pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_Comb_tau`	$-0.05^{+0.03}_{-0.03}$	Slope of $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$N_{\text{comb}}^{\pi\ \pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_nComb`	30^{+20}_{-10}	Yield of $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$\tau^{K\ \pi^0_{\text{part.reco.}}}$	`Ka_pi0_miss_Comb_tau`	$-0.005^{+0.002}_{-0.007}$	Slope of $h = K, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$N_{\text{comb}}^{K\ \pi^0_{\text{part.reco.}}}$	`Ka_pi0_miss_nComb`	30^{+30}_{-20}	Yield of $h = K, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$\tau^{\pi\ \gamma}$	`Pi_gam_Comb_tau`	$-0.0008^{+0.0001}_{-0.0001}$	Slope of $h = \pi, X^0 = \gamma$ combinatoric
$N_{\text{comb}}^{\pi\ \gamma}$	`Pi_gam_nComb`	1940^{+80}_{-80}	Yield of $h = \pi, X^0 = \gamma$ combinatoric
$\tau^{K\ \gamma}$	`Ka_gam_Comb_tau`	$-0.0010^{+0.0004}_{-0.0004}$	Slope of $h = K, X^0 = \gamma$ combinatoric
$N_{\text{comb}}^{K\ \gamma}$	`Ka_gam_nComb`	210^{+30}_{-30}	Yield of $h = K, X^0 = \gamma$ combinatoric

Table 5.23.: Free parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected $B^\pm \to D^*h^\pm$ candidates, with K^0_S candidates of track type LL. The designations assigned within the fitting program are listed, to allow the reader to interpret the corresponding plot of correlations, Figure 5.28.

Free Parameter	Designation	Fitted Value	Description
μ	`mean`	$5284.8^{+0.6}_{-0.6}$ MeV	Central value of $X^0 = \pi^0, \gamma$ signal peaks
k_σ	`width_smear_factor`	$1.01^{+0.02}_{-0.02}$	Ratio of $\mathcal{CB}$ widths in data w.r.t. simulation
r_σ^π	`xfeedinDstPi_widthratio`	$0.32^{+0.1}_{-0.06}$	Ratio of cross-feed to signal widths, for $h = \pi$
$r_{K\pi}$	`Ka_Pi_ratio`	$0.073^{+0.005}_{-0.006}$	Ratio of $h = K$ and $h = \pi$ signal yields
$r_\rho^{\pi^0}$	`rho_wrt_sig_ratio`	$0.62^{+0.04}_{-0.04}$	Ratio of $D^*\rho^\pm$ to signal yields, for $X^0 = \pi^0, \pi^0_{\text{part.reco.}}$
$r_{K^*}^{\pi^0}$	`kst_wrt_sig_ratio`	$0.3^{+0.1}_{-0.1}$	Ratio of $D^*K^{*\pm}$ to signal yields, for $X^0 = \pi^0, \pi^0_{\text{part.reco.}}$
$N_{\text{sig}}^{\pi\,\gamma}$	`Pi_gam_sig_nB`	1850^{+60}_{-50}	Signal yield for $h = \pi, X^0 = \gamma$ sample
$N_{\text{sig}}^{\pi\,\pi^0}$	`Pi_pi0_sig_nB`	280^{+20}_{-20}	Signal yield for $h = \pi, X^0 = \pi^0$ sample
$N_{\text{sig}}^{\pi\,\pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_sig_nB`	3620^{+10}_{-10}	Signal yield for $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ sample
$N_{D^*\rho}^{\pi\,\gamma}$	`Pi_gam_rho_nB`	500^{+100}_{-100}	Yield of $D^*\rho^\pm$ in $h = \pi, X^0 = \gamma$ sample
$N_{D\pi}^{\pi\,\pi^0}$	`Pi_pi0_dpi_nB`	34^{+7}_{-7}	Yield of $D\pi^\pm$ in $h = \pi, X^0 = \pi^0$ sample
$N_{D\pi}^{\pi\,\pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_dpi_nB`	510^{+30}_{-30}	Yield of $D\pi^\pm$ in $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ sample
$\tau^{\pi\,\pi^0}$	`Pi_pi0_Comb_tau`	$-0.006^{+0.002}_{-0.002}$	Slope of $h = \pi, X^0 = \pi^0$ combinatoric
$N_{\text{comb}}^{\pi\,\pi^0}$	`Pi_pi0_nComb`	140^{+30}_{-30}	Yield of $h = \pi, X^0 = \pi^0$ combinatoric
$\tau^{K\,\pi^0}$	`Ka_pi0_Comb_tau`	$-0.011^{+0.0009}_{-0.0011}$	Slope of $h = K, X^0 = \pi^0$ combinatoric
$N_{\text{comb}}^{K\,\pi^0}$	`Ka_pi0_nComb`	40^{+9}_{-8}	Yield of $h = K, X^0 = \pi^0$ combinatoric
$\tau^{\pi\,\pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_Comb_tau`	$-0.0043^{+0.0004}_{-0.0005}$	Slope of $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$N_{\text{comb}}^{\pi\,\pi^0_{\text{part.reco.}}}$	`Pi_pi0_miss_nComb`	1300^{+200}_{-200}	Yield of $h = \pi, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$\tau^{K\,\pi^0_{\text{part.reco.}}}$	`Ka_pi0_miss_Comb_tau`	$-0.0035^{+0.0006}_{-0.0008}$	Slope of $h = K, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$N_{\text{comb}}^{K\,\pi^0_{\text{part.reco.}}}$	`Ka_pi0_miss_nComb`	280^{+50}_{-50}	Yield of $h = K, X^0 = \pi^0_{\text{part.reco.}}$ combinatoric
$\tau^{\pi\,\gamma}$	`Pi_gam_Comb_tau`	$-0.00144^{+0.00008}_{-0.00008}$	Slope of $h = \pi, X^0 = \gamma$ combinatoric
$N_{\text{comb}}^{\pi\,\gamma}$	`Pi_gam_nComb`	5300^{+100}_{-100}	Yield of $h = \pi, X^0 = \gamma$ combinatoric
$\tau^{K\,\gamma}$	`Ka_gam_Comb_tau`	$-0.0012^{+0.0002}_{-0.0002}$	Slope of $h = K, X^0 = \gamma$ combinatoric
$N_{\text{comb}}^{K\,\gamma}$	`Ka_gam_nComb`	570^{+50}_{-40}	Yield of $h = K, X^0 = \gamma$ combinatoric

Table 5.24.: Free parameters used in the simultaneous fit to the reconstructed B-mass spectrum for selected $B^\pm \to D^*h^\pm$ candidates, with K^0_S candidates of track type DD. The designations assigned within the fitting program are listed, to allow the reader to interpret the corresponding plot of correlations, Figure 5.28.

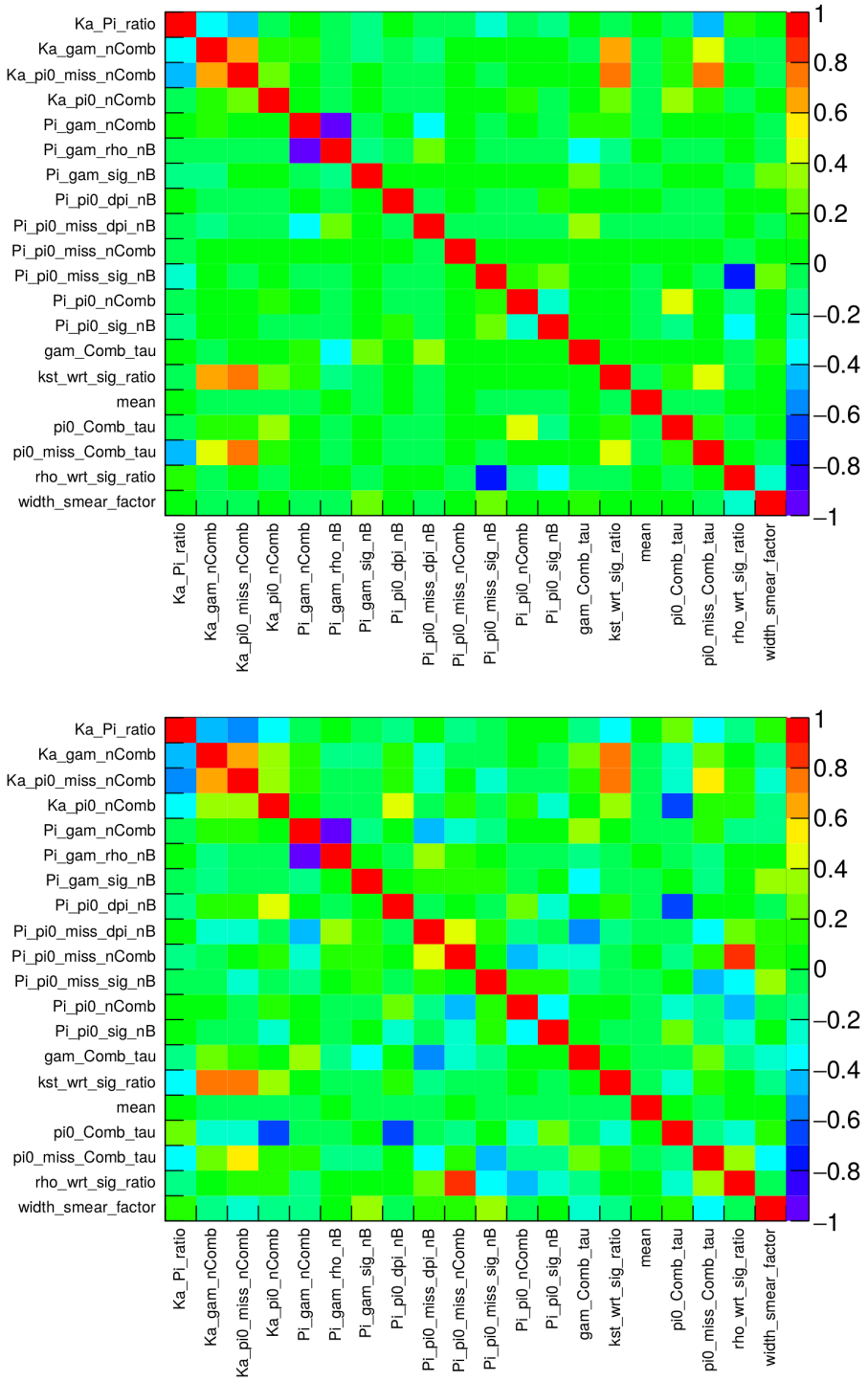


Figure 5.28.: Correlation matrices obtained for parameters in the simultaneous B-mass fits to LL (top) and DD (bottom) $B^{\pm} \to D^{*}h^{\pm}$ candidates.

Component	Yield					
	LL			DD		
	$D^* \to D\gamma$	$D^* \to D\pi^0$	$D^* \to D\pi^0$ part. reco.	$D^* \to D\gamma$	$D^* \to D\pi^0$	$D^* \to D\pi^0$ part. reco.
signal	53 ± 4	7.5 ± 0.8	70 ± 5	120 ± 10	18 ± 2	200 ± 17
$D^*\pi^\pm$ cross-feed	7.5 ± 0.3	1.07 ± 0.08	11.2 ± 0.4	22.3 ± 0.7	3.2 ± 0.2	42 ± 1
$B^\pm \to D^*K^{*\pm}$	0.8 ± 0.2	1.02 ± 0.03	1.5 ± 0.3	0.5 ± 0.1	0.07 ± 0.02	1.2 ± 0.3
$B^\pm \to DK^\pm$	negligible	0.034 ± 0.008	0.088 ± 0.009	0.14 ± 0.09	0.05 ± 0.01	0.087 ± 0.009
$B^\pm \to D\pi^\pm$	negligible	0.09 ± 0.02	0.094 ± 0.007	0.13 ± 0.09	0.05 ± 0.01	0.36 ± 0.02
combinatoric	21 ± 2	0.05 ± 0.02	5.5 ± 1.2	60 ± 4	4.1 ± 0.9	28 ± 5

Table 5.25.: Yields extracted in the signal region for mode $B^\pm \to D^*K^\pm$.

5.5. Estimation of fake D background yield

The combinatoric background in each of the modes studied consists of two components, arising from B-meson candidates containing fake and real D-meson candidates. It is important to distinguish between these two categories, since they give rise to different distributions over the Dalitz plane, which must be accounted for in the $\mathcal{CP}$ fit. In the 'real D' case, the combinatoric, incorrectly-reconstructed part of the B-meson candidate does not include the D-meson or any of its progeny, and so the D-meson candidate is correctly reconstructed. This ensures that the distribution of these candidates in the Dalitz plane must be an admixture of those for real D^0 and $\overline{D^0}$ decays. In the 'fake D' case, the D-meson candidate is reconstructed from incorrectly associated tracks, and so the kinematic variables of the decay are not correlated as they would be for a correctly-reconstructed candidate. Instead, these candidates are assumed to give a flat distribution over the Dalitz plane.[11]

In order to determine the separate fake and real D-meson background yields, fits to the reconstructed D-meson mass are performed for data events selected in each mode and lying within the signal window of the corresponding B-mass spectrum. Events containing real D candidates give peaking distributions, described using Gaussian shapes with floated mean and width parameters. Fake D candidates do not peak, and are instead modelled using a flat distribution.

Due to the low statistics entering these fits, common D-mass fits are performed across several categories:

- $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ events containing LL K^0_S candidates;
- $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ events containing DD K^0_S candidates;
- $B^0 \to DK^{*0}$ events containing both LL and DD K^0_S candidates;
- $B^\pm \to DK^{*\pm}$ events containing both LL and DD K^0_S candidates;
- $B^\pm \to D^*\pi^\pm$ and $B^\pm \to D^*K^\pm$ events, containing both LL and DD K^0_S candidates and satisfying the neutral hypothesis $D^* \to D\gamma$; and
- $B^\pm \to D^*\pi^\pm$ and $B^\pm \to D^*K^\pm$ events, containing both LL and DD K^0_S candidates and satisfying the neutral hypothesis $D^* \to D\pi^0$, where the neutral pion may be either fully or partially reconstructed.

The width of the D-meson mass window is doubled from its nominal value when performing these fits, to allow the tails of the Gaussian shapes to be fitted. To extract the yields of fake and real D-meson candidates in the final selected datasets, the fitted PDFs are then integrated over the nominal mass window. The resulting fits are shown in Figures 5.29–5.34, where in each case the nominal mass window is delimited by two vertical grey lines. The yields of real and fake D candidates extracted in these fits are shown in Table 5.26. The fake D yields obtained for modes $B^0 \to DK^{*0}$ and $B^\pm \to DK^{*\pm}$ are compatible with zero, suggesting it may be possible in future to constrain these values to zero in the nominal $\mathcal{CP}$ fit.

[11] A systematic uncertainty may later be evaluated considering an alternative distribution.

Mode	K^0_S type	n_{fakeD}	n_{realD}
$B^\pm \to D\pi^\pm$ & $B^\pm \to DK^\pm$	LL	500 ± 30	10800 ± 100
$B^\pm \to D\pi^\pm$ & $B^\pm \to DK^\pm$	DD	990 ± 60	26300 ± 200
$B^0 \to DK^{*0}$	LL & DD	2^{+4}_{-2}	120 ± 10
$B^\pm \to DK^{*\pm}$	LL & DD	0^{+1}_{-0}	50 ± 7
$B^\pm \to D^*\pi^\pm$ & $B^\pm \to D^*K^\pm$, $D^* \to D\gamma$	LL & DD	100 ± 20	3690 ± 60
$B^\pm \to D^*\pi^\pm$ & $B^\pm \to D^*K^\pm$, $D^* \to D\pi^0$	LL & DD	140 ± 20	5050 ± 70

Table 5.26.: Yields of real and fake D candidates extracted in fits to the reconstructed D-meson mass spectrum for each sample fitted.

The pulls arising in fits to selected $B^\pm \to Dh^\pm$ and $B^\pm \to D^*h^\pm$ candidates (Figures 5.29, 5.30, 5.33 and 5.34), where statistics are plentiful, reveal systematic deviations implying that the tails of the distribution of real D-meson candidates are not well modelled by a Gaussian shape. A systematic may be extracted by repeating these fits using alternative PDF shapes.

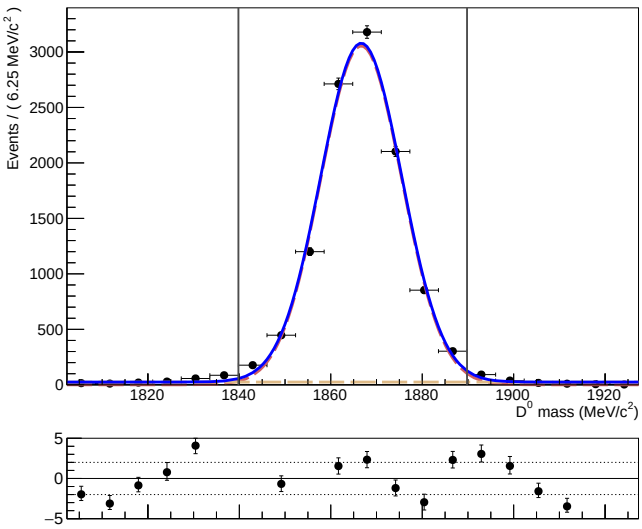


Figure 5.29.: Fit to the reconstructed D-meson mass spectrum for selected $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ data events lying within the B-meson mass signal window, with K^0_S candidates of track type LL. The lower canvas shows the pull.

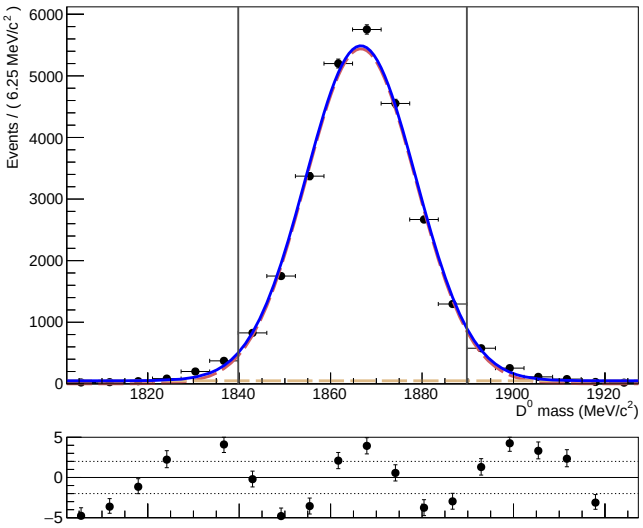


Figure 5.30.: Fit to the reconstructed D-meson mass spectrum for selected $B^{\pm} \to DK^{\pm}$ and $B^{\pm} \to D\pi^{\pm}$ data events lying within the B-meson mass signal window, with K^0_S candidates of track type DD. The lower canvas shows the pull.

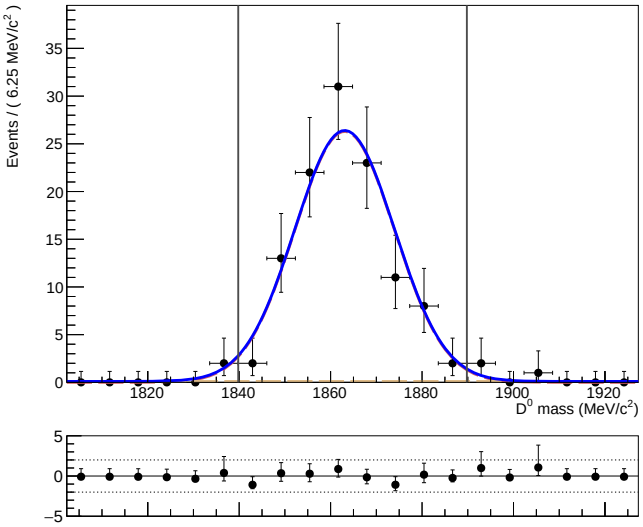


Figure 5.31.: Fit to the reconstructed D-meson mass spectrum for selected $B^0 \to DK^{*0}$ data events lying within the B-meson mass signal window, with K^0_S candidates of track type LL and DD combined. The lower canvas shows the pull.

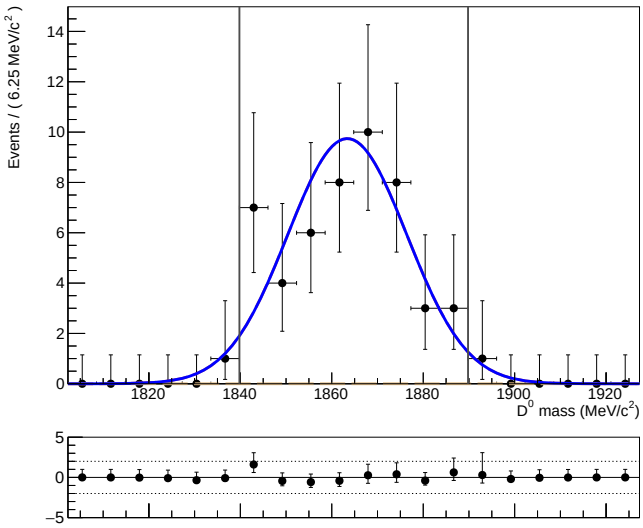


Figure 5.32.: Fit to the reconstructed D-meson mass spectrum for selected $B^{\pm} \to DK^{*\pm}$ data events lying within the B-meson mass signal window, with K^0_S candidates of track type LL and DD combined. The lower canvas shows the pull.

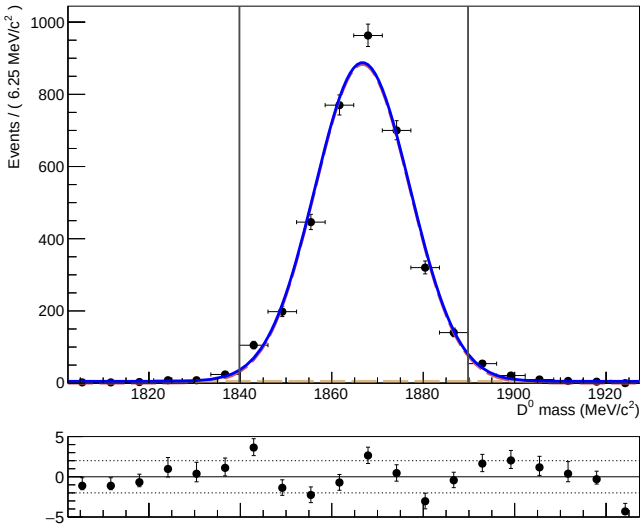


Figure 5.33.: Fit to the reconstructed D-meson mass spectrum for selected $B^{\pm} \to D^*K^{\pm}$ and $B^{\pm} \to D^*\pi^{\pm}$ data events lying within the B-meson mass signal window, identified under the neutral hypothesis $D^* \to D\gamma$. Events containing K^0_S candidates of track type LL and DD are combined. The lower canvas shows the pull.

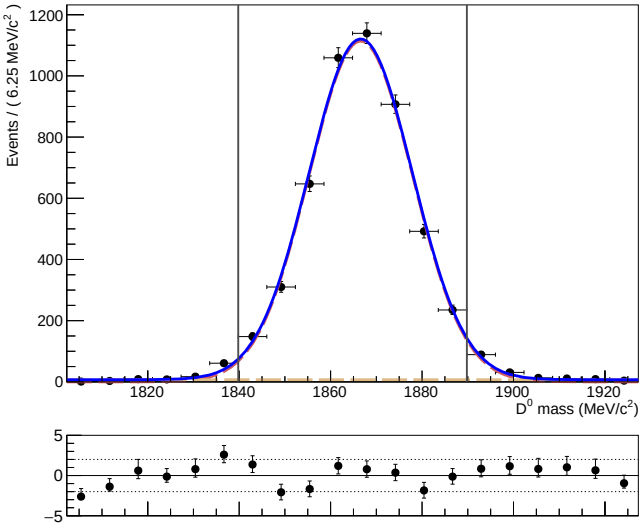


Figure 5.34.: Fit to the reconstructed D-meson mass spectrum for selected $B^{\pm} \to D^{*}K^{\pm}$ and $B^{\pm} \to D^{*}\pi^{\pm}$ data events lying within the B-meson mass signal window, identified under the neutral hypothesis $D^{*} \to D\pi^{0}$ (both correctly- and partially-reconstructed neutral hypotheses are combined). Events containing K^{0}_{S} candidates of track type LL and DD are combined. The lower canvas shows the pull.

6. Determination of the selection efficiency variation over the Dalitz plane

In order to extract the values of the $\mathcal{CP}$ parameters, the events selected in each mode are fitted in the Dalitz plane of the $D \to K^0_S\pi^+\pi^-$ decay, where non-uniformities, parameterized by γ, arise due to interference between $\mathcal{CP}$ conjugate D-meson eigenstates. It is therefore crucial to take into account any additional non-uniformities arising due to the selection efficiency.[1]

One approach is to include in the data fit a generalised smooth, symmetric efficiency function with floating parameters. In this way, the function is fitted to data alongside the model. This increases the number of free parameters in the final fit, and thus has the potential to reduce the statistical sensitivity with which the $\mathcal{CP}$ parameters can be extracted. However, this method is data-driven, and, where a high-statistics control mode is available that allows a reliable determination of the relative efficiency for the signal mode, the precision with which the $\mathcal{CP}$ parameters may be determined is negligibly affected.

Another approach is to use simulated events to extract a parameterization for the selection efficiency, which is then fixed in the final $\mathcal{CP}$ fit. Within this approach, care must be taken to account for any data/simulation differences to ensure that the obtained efficiency is accurate. In addition, any uncertainties in the efficiency function obtained must be propagated into the final $\mathcal{CP}$ fit result as systematic uncertainties.

The present chapter describes the extraction of efficiency histograms from simulation. Efficiency functions are extracted for all modes; the actual efficiency approach (data- or simulation-driven) chosen for each mode in the final $\mathcal{CP}$ fit is described alongside the fits, in Chapter 8.

The Monte Carlo simulated samples used are shown in Table 6.1. To prevent bias, these samples have no overlap with those used to train the MVA classifiers for each mode. All samples are truth-matched using full truth-level information to connect parent and daughter mesons within each event. For kinematically-equivalent modes, common efficiencies are used.

- For modes $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$, a common efficiency is determined using $B^\pm \to DK^\pm$ MC.
- For modes $B^0 \to DK^{*0}$ and $B^0_s \to DK^{*0}$, a common efficiency is determined from $B^0_s \to DK^{*0}$ MC.
- For modes $B^\pm \to D^*K^\pm$ and $B^\pm \to D^*\pi^\pm$, a common efficiency is obtained from $B^\pm \to D^*K^\pm$ MC, and the three signal categories ($D^* \to D\gamma$, $D^* \to D\pi^0$ and partially-reconstructed $D^* \to D\pi^0$) are combined.

In each mode, separate efficiencies are obtained for events containing LL and DD K^0_S candidates, since separate selections are applied in each case.

[1] Note that the absolute value of the efficiency is not needed, only the variations.

Mode	No. of events generated	No. of events in efficiency determination
$B^\pm \to DK^\pm$	10×10^6	82350 (LL)
		267597 (DD)
$B_s^0 \to DK^{*0}$	400×10^6	79817 (LL)
		244534 (DD)
$B^\pm \to DK^{*\pm}$	110×10^6	15514 (LL)
		56324 (DD)
$B^\pm \to D^*K^\pm$	110×10^6	29155 (LL)
		129847 (DD)

Table 6.1.: Simulated Monte Carlo samples used to extract efficiency variations over the Dalitz plane. The number of events generated is shown, alongside the number of events which pass the selection (not including trigger and PID requirements) and thus enter into the efficiency determination.

The Dalitz plane of the $D \to K_S^0\pi^+\pi^-$ decay may be parameterized using the combined invariant masses of each pair of decay products, denoted by

$$S_+^2 = |P^\mu_{K_S^0} + P^\mu_{\pi^+}|^2, \tag{6.1}$$

$$S_-^2 = |P^\mu_{K_S^0} + P^\mu_{\pi^-}|^2, \quad \text{and} \tag{6.2}$$

$$S_0^2 = |P^\mu_{\pi^+} + P^\mu_{\pi^-}|^2. \tag{6.3}$$

Any point in the phase space of the decay is fully specified by providing any two out of these three variables. Throughout this chapter, efficiency variations will be described as functions of S_+^2 and S_-^2.

The total efficiency for each selection may be decomposed into several component parts, concerning each set of requirements applied to the generated events:

$$\epsilon^{c,tt}(S_-^2, S_+^2) = \epsilon^c_{\text{acc}}\,\epsilon^c_{\text{rec}}\,\epsilon^{c,tt}_{\text{strip}}\,\epsilon^{c,tt}_{\text{sel}}\,\epsilon^{c,tt}_{\text{trig}}\,\epsilon^{c,tt}_{\text{PID}}, \tag{6.4}$$

where ϵ^c_{acc} is the efficiency of the acceptance requirements for mode c, ϵ^c_{rec} is the reconstruction efficiency, $\epsilon^{c,tt}_{\text{strip}}$ is the stripping efficiency for events in mode c with K_S^0 candidates of track type (tt = LL, DD), $\epsilon^{c,tt}_{\text{sel}}$ is the efficiency of the offline selection, excluding trigger and PID requirements, $\epsilon^{c,tt}_{\text{trig}}$ is the trigger efficiency, and $\epsilon^{c,tt}_{\text{PID}}$ is the PID efficiency.

Since each sample is generated uniformly over the allowed kinematic phase space of the decay, without any resonant structure included, one might naively assume that the distribution of selected events directly gives a histogram of the efficiency. However, the simulation is known to give a poor representation of the distributions of the Level-0 hadron trigger decision, and the PID variables, and so a separate weighting procedure is used to obtain the efficiencies ϵ_{L0} and ϵ_{PID}. Histograms, $\epsilon^{c,tt}_{\text{direct}}$, are obtained directly from simulated events passing the full selection, excluding trigger and PID requirements, so that

$$\epsilon^{c,tt}_{\text{direct}} = \epsilon^c_{\text{acc}}\,\epsilon^c_{\text{rec}}\,\epsilon^{c,tt}_{\text{strip}}\,\epsilon^{c,tt}_{\text{sel}}, \tag{6.5}$$

and thus

$$\epsilon^{c,tt}(S^2_-, S^2_+) = \epsilon^{c,tt}_{\text{direct}}\,\epsilon^{c,tt}_{\text{trig}}\,\epsilon^{c,tt}_{\text{PID}}. \tag{6.6}$$

The largest source of non-uniformity in the efficiencies arises due to the selection requirements preceding the offline selection; namely the acceptance, reconstruction and stripping selection criteria. The combined efficiencies of these stages, $\epsilon^{c}_{\text{acc}}\epsilon^{c}_{\text{rec}}\epsilon^{c,tt}_{\text{strip}}$, obtained from the distribution of MC events before any offline selection requirements are applied, are shown in Figures 6.1–6.4 for $B^\pm \to DK^\pm$, $B^0_s \to DK^{*0}$, $B^\pm \to DK^{*\pm}$ and $B^\pm \to D^*K^\pm$ MC. In each case, events containing LL and DD K^0_S candidates are shown together.

These efficiency distributions may be understood by considering the combined effects of detector acceptance, reconstruction methodology and stripping requirements. In general, reconstruction algorithms tend to be more effective when candidates have higher energies, and higher energy candidates are preferred by the stripping requirements defined in Chapter 4. However, the more high energy candidates within a decay, the more likely it becomes that one or more will have sufficient transverse momentum to leave the LHCb detector acceptance. Given a specific B-meson decay mode, then, the efficiency $\epsilon^{c}_{\text{acc}}\epsilon^{c}_{\text{rec}}\epsilon^{c,tt}_{\text{strip}}$ will be greatest when 4-momentum is partitioned amongst decay products such that all remain within detector acceptance, but with high enough energies to give reliable reconstruction, and to pass the selection requirements.

Figures 6.1–6.4 show efficiency distributions that decrease as either S^2_+ or S^2_- increase. The efficiency is low when S^2_+ is large and S^2_- is small, or *vice versa*, because in these circumstances the lower energy pion becomes difficult to distinguish from the low energy pions of the underlying event, and so the reconstructed D-meson is more likely to be a combinatoric candidate, and fail the selection. The efficiency is also low when *both* S^2_+ and S^2_- are large. In this case, the high energies of the D-meson decay products increase the likelihood that one or more candidates leave the LHCb acceptance. In adddition, even if all products remain within the acceptance, since now the 4-momentum partitioned amongst the rest of the candidates in the decay is lower, this leads to a greater likelihood of one of these candidates failing the selection.

Mode $B^\pm \to DK^{*\pm}$, Figure 6.3, appears to show steeper variation across the Dalitz plane than the other modes. This may be due to the presence of two K^0_S-meson candidates within the decay. When the D-meson decay products have high energies, resulting in large S^2_+ and S^2_-, the K^0_S candidate originating from the $K^{*\pm}$-meson decay tends to have lower energy, making it more difficult to reconstruct from its daughter pions and thus less likely to pass the selection.

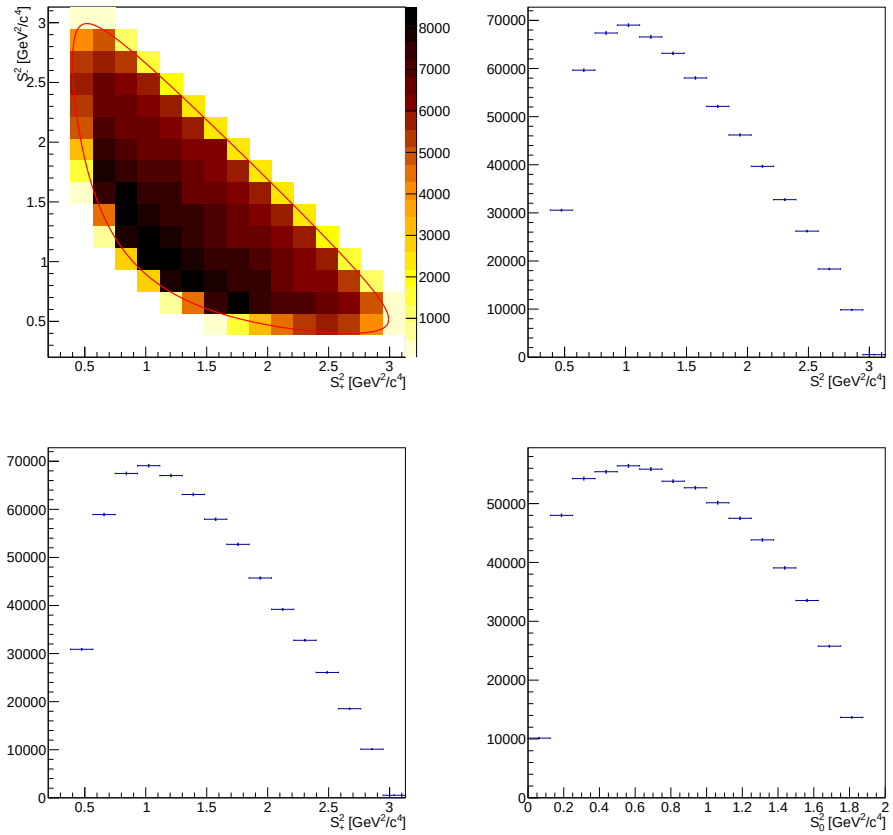


Figure 6.1.: Distributions of $B^\pm \to DK^\pm$ MC events after stripping selection. Both LL and DD K_S^0 candidates are shown.

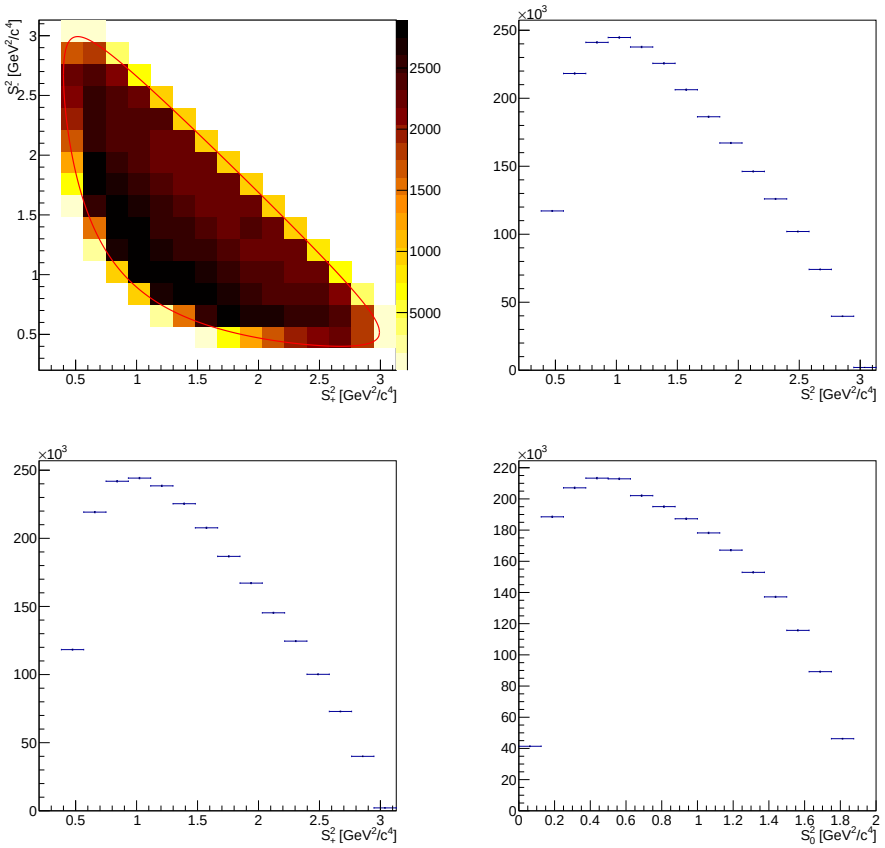


Figure 6.2.: Distributions of $B_s^0 \to DK^{*0}$ MC events after stripping selection. Both LL and DD K_S^0 candidates are shown.

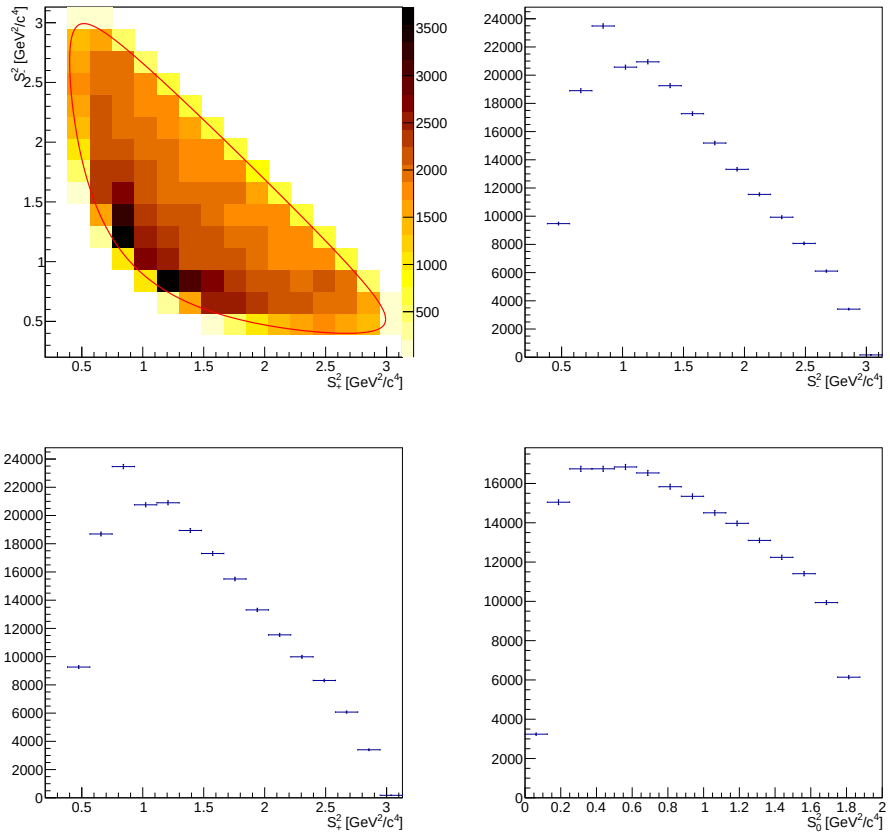


Figure 6.3.: Distributions of $B^{\pm} \to DK^{*\pm}$ MC events after stripping selection. Both LL and DD K_S^0 candidates are shown.

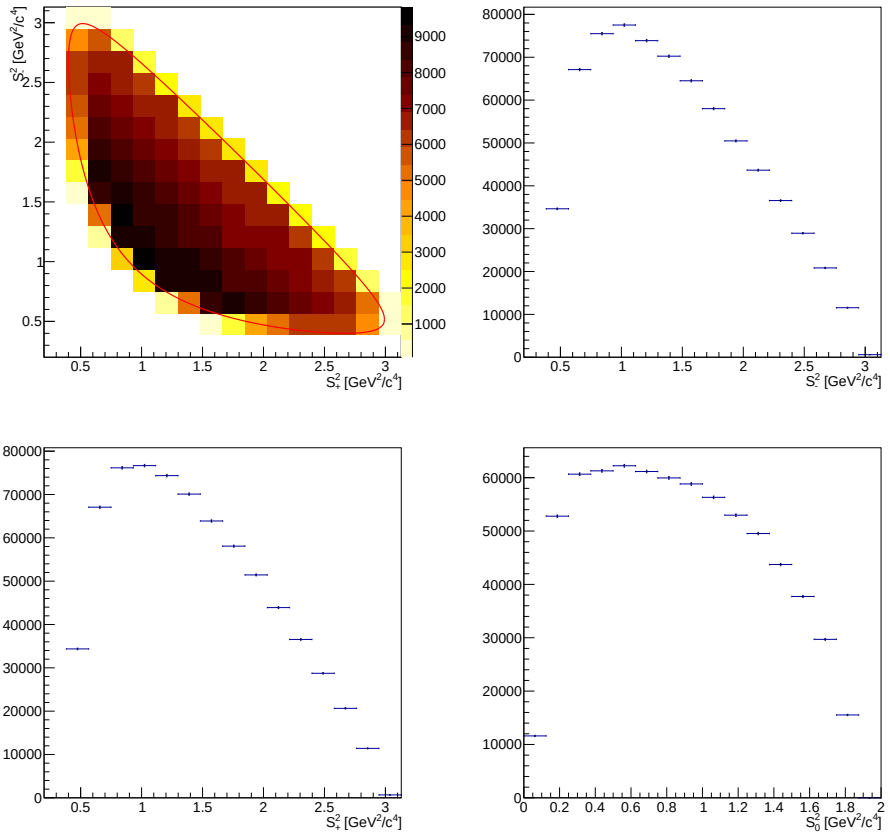


Figure 6.4.: Distributions of $B^{\pm} \to D^* K^{\pm}$ MC events after stripping selection. Both LL and DD K_S^0 candidates are shown.

6.1. Edge bin reweighting

In using a frequency histogram, $\epsilon^{c,tt}_{\text{direct}}$, to describe what is in reality a continuous function, the content of each bin represents the average of the true efficiency function over the area of the bin. However, for bins lying partially outside the kinematically-allowed region of the Dalitz plane, this averaging includes an unphysical, and thus depopulated region. This gives an underestimation of the function value; since it should be averaged only over the allowed part of the bin. This effect is corrected by re-weighting bins by a factor inversely proportional to the fraction of the bin lying within the allowed phase-space.

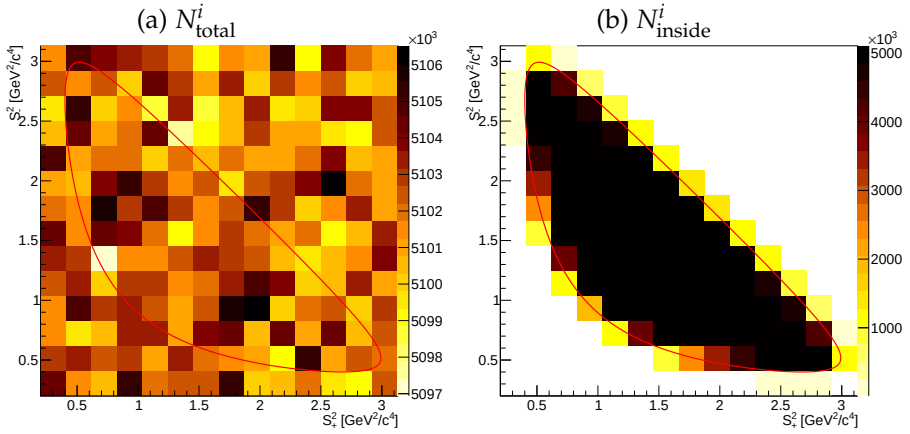


Figure 6.5.: Illustration of the method used to correct the populations of 'edge bins', lying partly outside the allowed region of phase space. Uniformly generated points (a), give the bin populations N^i_{total}, while the subsets of these points lying within the allowed region (b) give the bin populations N^i_{inside}.

The procedure used to obtain these correction weights is illustrated in Figure 6.5. Points are generated uniformly over the Dalitz plane, giving a frequency histogram with bin populations N^i_{total}. The subset of these points that lie within the allowed region is found, giving a histogram with bin populations N^i_{inside}. Correction weights ϵ^i_{edge} for each bin may then be extracted;

$$\epsilon^i_{\text{edge}} = N^i_{\text{total}} / N^i_{\text{inside}}. \tag{6.7}$$

Applying this correction, the expression for the total selection efficiency in mode c becomes

$$\epsilon^{c,tt}(S^2_-, S^2_+) = \epsilon^{c,tt}_{\text{direct}}\, \epsilon_{\text{edge}}\, \epsilon^{c,tt}_{\text{trig}}\, \epsilon^{c,tt}_{\text{PID}}. \tag{6.8}$$

6.2. Trigger weights

The L0 trigger is hardware-based, and thus more difficult to simulate than the selection variables which are generated in the reconstruction software. For all modes considered in the analysis, the L0 trigger requirements are the same:

$$(\texttt{B_L0HadronDecision_TOS} == 1 || \texttt{B_L0Global_TIS} == 1). \tag{6.9}$$

The L0 Hadron TOS trigger requires that the hardware trigger fired due to one or more tracks used to reconstruct the signal candidate, and thus it is important to ensure its efficiency is well described. The inclusion of the L0Global TIS trigger is crucial in allowing the trigger efficiency to be extracted, since this requires that the hardware trigger fired due to a track coming from the underlying event, and thus gives no bias on the kinematics, or on the Dalitz distribution, of selected candidates. It is useful to define two independent categories of events which may pass the above trigger requirements:

- Events are named 'TOS' if they fire the `B_L0HadronDecision_TOS` trigger;
- Events are named 'notTOS' if they do not fire the `B_L0HadronDecision_TOS` trigger, but do fire the `B_L0Global_TIS` trigger.

All events passing the trigger requirements must fall into one of these two categories.

Standard tables [150], calibrated from data, are available to describe the probability for different final state particle tracks (pions, kaons, electrons, etc.) to fire the `B_L0HadronDecision_TOS` trigger, binned in the transverse momentum of the track. Since the threshold of this trigger was adjusted between 2011 and 2012 data-taking, separate tables, $\epsilon^{2011}_{\text{TOS}}(p_T)$ and $\epsilon^{2012}_{\text{TOS}}(p_T)$, have been computed for each year.

Per-track weights, w^t_{TOS}, may be assigned to the simulated reference samples, by combining the values found in each table in the ratio of the luminosities measured in 2011 and 2012:

$$w^t_{\text{TOS}}(p_T) = \frac{\mathcal{L}_{2011}}{\mathcal{L}_{2011}+\mathcal{L}_{2012}}\epsilon^{2011}_{\text{TOS}}(p_T) + \frac{\mathcal{L}_{2012}}{\mathcal{L}_{2011}+\mathcal{L}_{2012}}\epsilon^{2012}_{\text{TOS}}(p_T). \tag{6.10}$$

A given event passes the TOS trigger if any one of the tracks it contains fires the trigger. Noting that

$$P(\text{event is TOS}) = P(\text{any track within event is TOS}), \tag{6.11}$$
$$= 1 - P(\text{no track within event is TOS}), \tag{6.12}$$

per-event weights, w^e_{TOS}, may then be calculated using

$$w^e_{\text{TOS}} = 1 - \prod_t (1 - w^t_{\text{TOS}}). \tag{6.13}$$

A histogram of the L0 selection efficiency for TOS events may then be constructed by taking the average of these weights within bins of the Dalitz plane, such that the value for the i-th bin is given by

$$\epsilon^i_{\text{TOS}} = \frac{\sum_{e=1}^{N_i} w^e_{\text{TOS}}}{N_i}. \tag{6.14}$$

Histograms obtained for each mode are displayed in Figures 6.6–6.13. Since a single high-p_T pion track is more likely to fire the hadron trigger than two low-p_T ones, the TOS efficiency is observed to be greatest in the high-S^2_+ and high-S^2_- corners of the allowed phase space.

Mode $B^\pm \to DK^{*\pm}$, Figures 6.10 and 6.11, appears to show less variation across the Dalitz plane than the other modes. This may be because when the D-meson decay products have lower energies, and are thus less likely to fire the hadron trigger, the K^0_S candidate, and its daughter pions, originating from the $K^{*\pm}$-meson decay tend to have higher energy, and thus these pions become more likely to fire the trigger.

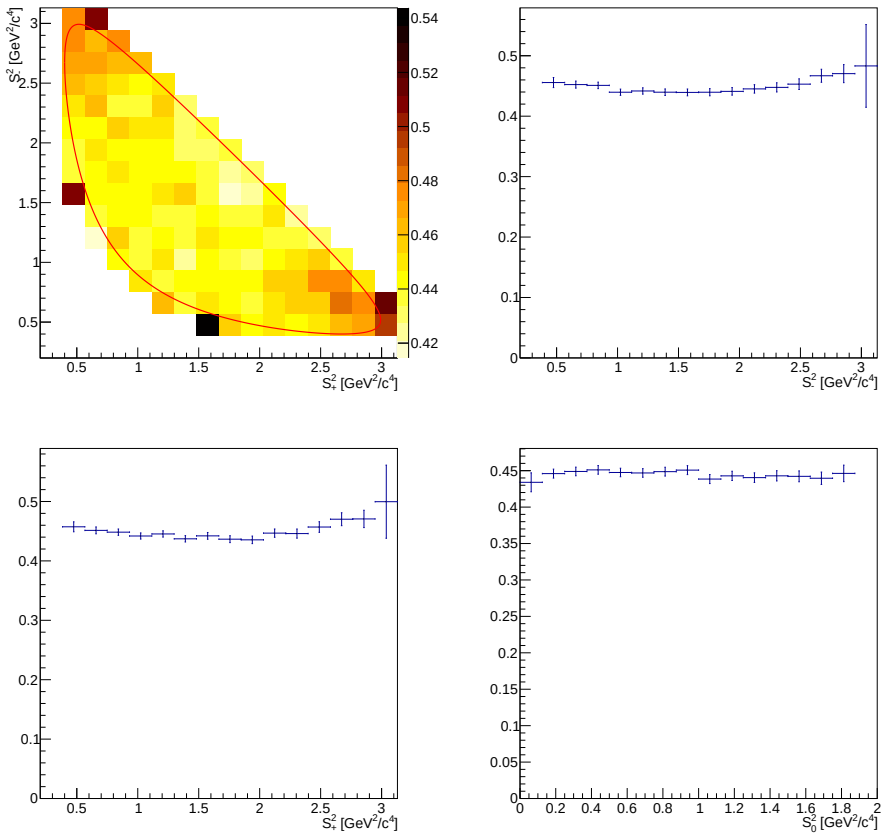


Figure 6.6.: L0 Hadron TOS efficiency weights obtained in selected $B^\pm \to DK^\pm$ MC, for events containing LL K^0_S candidates.

In order to construct the full L0 selection efficiency for each mode, the efficiency for events to be selected in the notTOS category must also be calculated. From the requirements that define this category,

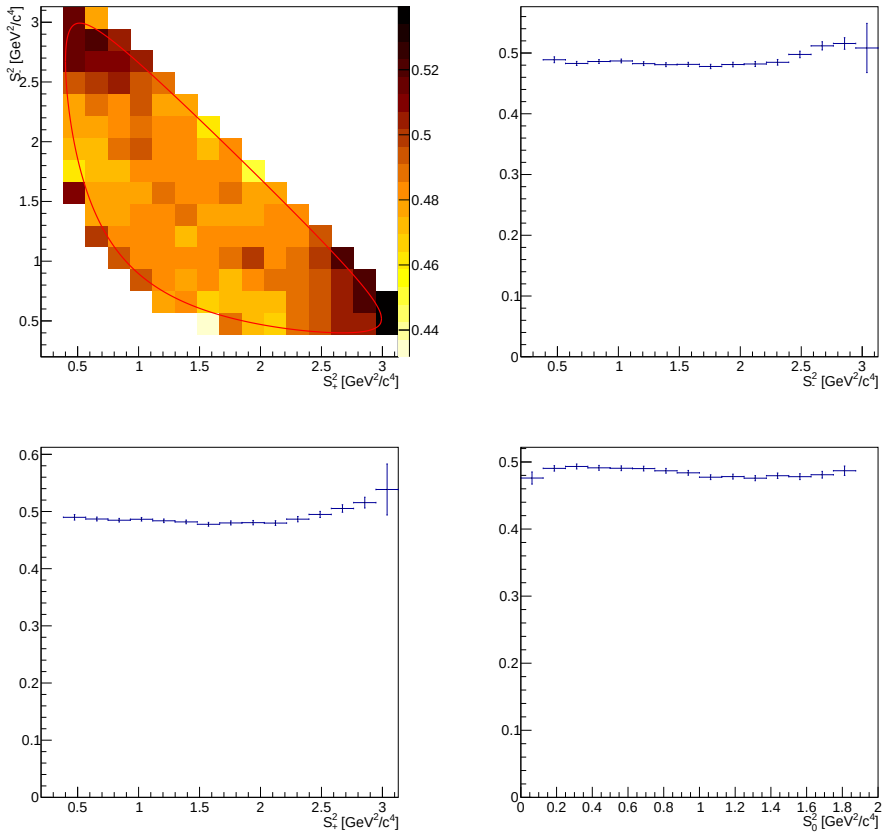


Figure 6.7.: L0 Hadron TOS efficiency weights obtained in selected $B^\pm \to DK^\pm$ MC, for events containing DD K^0_S candidates.

one may write

$$P(\text{event is notTOS}) = P(\text{event is not TOS}|\text{event is L0Global_TIS}) \cdot P(\text{event is L0Global_TIS}). \quad (6.15)$$

Since the Global_TIS trigger decision does not depend on any particle within the signal decay, the second factor is assumed to be uniform over the Dalitz plane. The notTOS efficiency can then be found similarly to the TOS efficiency, but considering only those reference events which pass the L0Global_TIS trigger, and using the probability that none of the tracks in a given event fires the Hadron trigger.

Considering only these events, per-event weights may be assigned as

$$w^e_{\text{notTOS}} = \prod_t (1 - w^t_{\text{TOS}}), \quad (6.16)$$

and a histogram, $\epsilon^i_{\text{notTOS}}$, may be constructed by averaging within bins of the Dalitz plane.

Histograms obtained for each mode are displayed in Figures 6.14–6.21. Since notTOS events are required not to fire the L0Hadron trigger, these histograms display distributions complementary to the

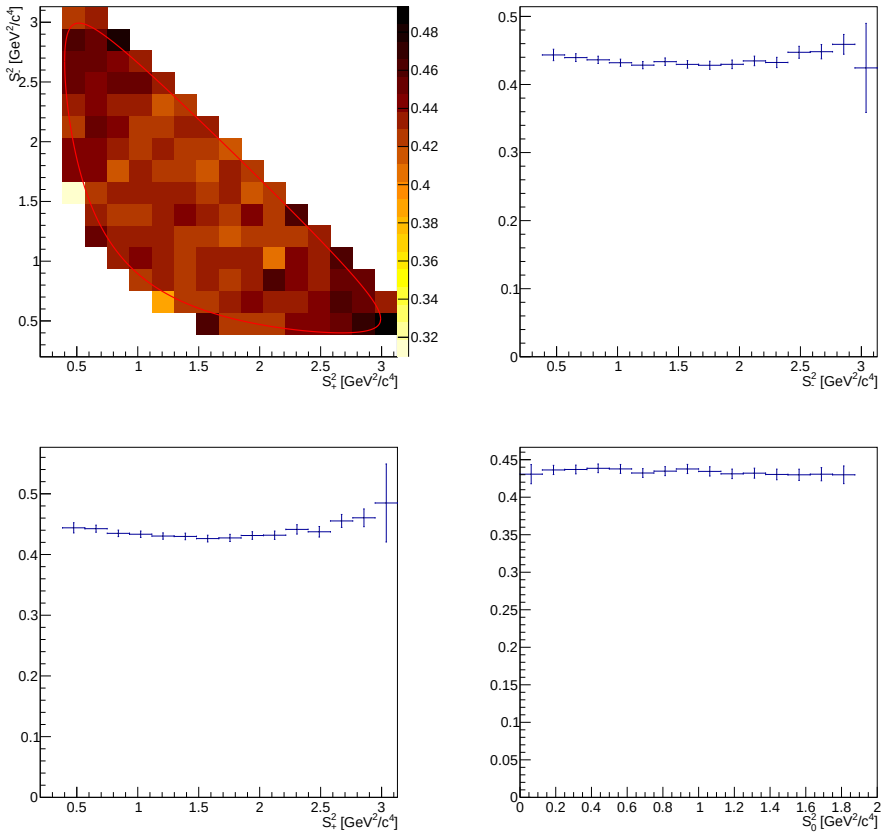


Figure 6.8.: L0 Hadron TOS efficiency weights obtained in selected $B_s^0 \to DK^{*0}$ MC, for events containing LL K_S^0 candidates.

TOS efficiencies, with efficiency decreasing in the high-S_+^2 and high-S_-^2 corners of the allowed phase space.

Finally, the full L0 trigger efficiency is obtained for each mode, by combining the corresponding TOS and notTOS histograms. The relative proportions are fixed to match the ratios of TOS and notTOS events found in the final data samples used in the $\mathcal{CP}$ fit. For the m-th mode, the final L0 trigger efficiency histogram is then given by

$$\epsilon^m_{\text{L0}} = f^m_{\text{TOS}}\epsilon^m_{\text{TOS}} + (1 - f^m_{\text{TOS}})\epsilon^m_{\text{notTOS}}. \tag{6.17}$$

The numbers of TOS and notTOS candidates selected for each data sample are shown in Table 6.2.

L0 efficiency histograms obtained for each mode are displayed in Figures 6.22–6.25.

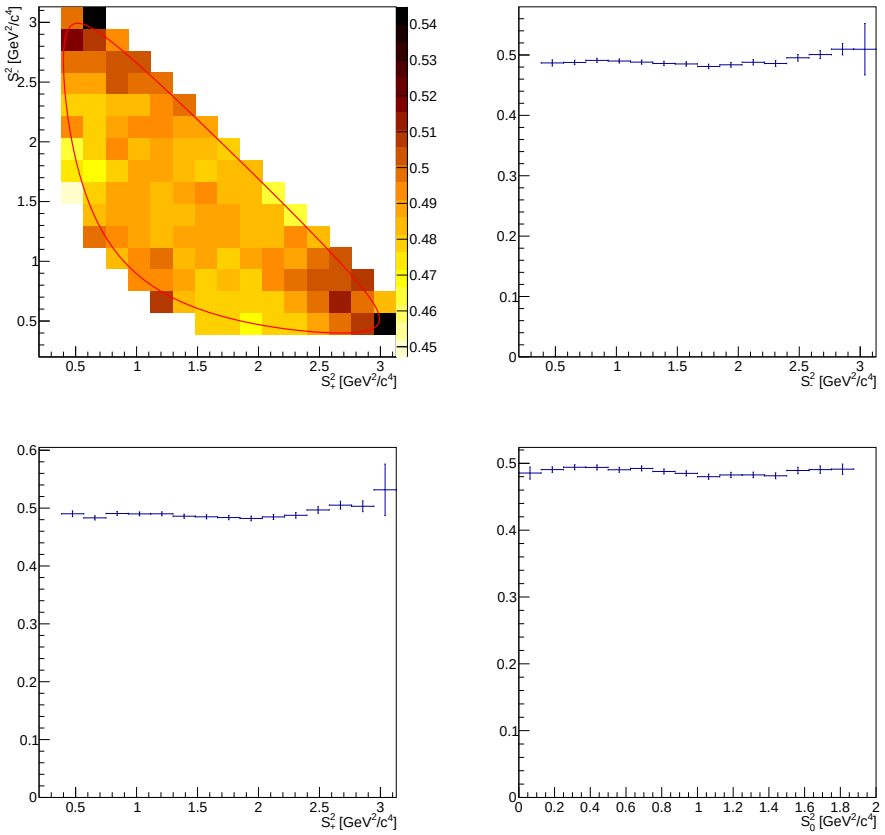


Figure 6.9.: L0 Hadron TOS efficiency weights obtained in selected $B^0_s \to DK^{*0}$ MC, for events containing DD K^0_S candidates.

Data sample	TOS yield	notTOS yield
$B^\pm \to DK^\pm$ (LL)	525	910
$B^\pm \to DK^\pm$ (DD)	1417	2225
$B^0 \to DK^{*0}$ (LL)	14	24
$B^0 \to DK^{*0}$ (DD)	43	70
$B^\pm \to DK^{*\pm}$ (LL)	11	13
$B^\pm \to DK^{*\pm}$ (DD)	23	35
$B^\pm \to D^*K^\pm$ (LL)	98	169
$B^\pm \to D^*K^\pm$ (DD)	309	446

Table 6.2.: TOS and notTOS yields observed in each data sample, after the application of all selection requirements.

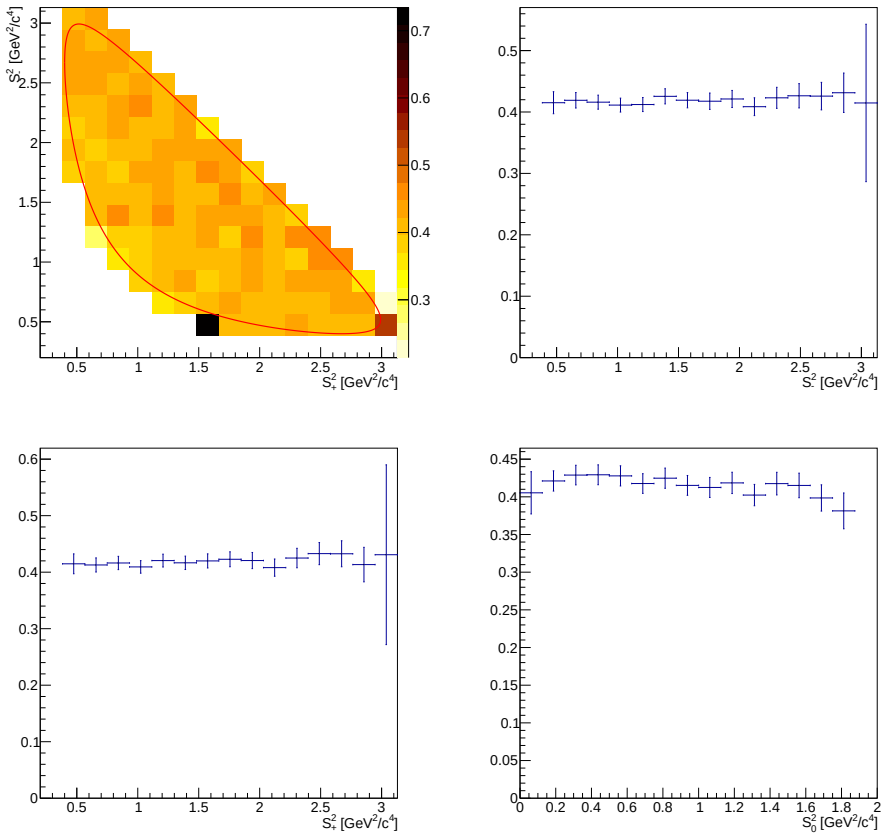


Figure 6.10.: L0 Hadron TOS efficiency weights obtained in selected $B^{\pm} \to DK^{*\pm}$ MC, for events containing LL K_S^0 candidates.

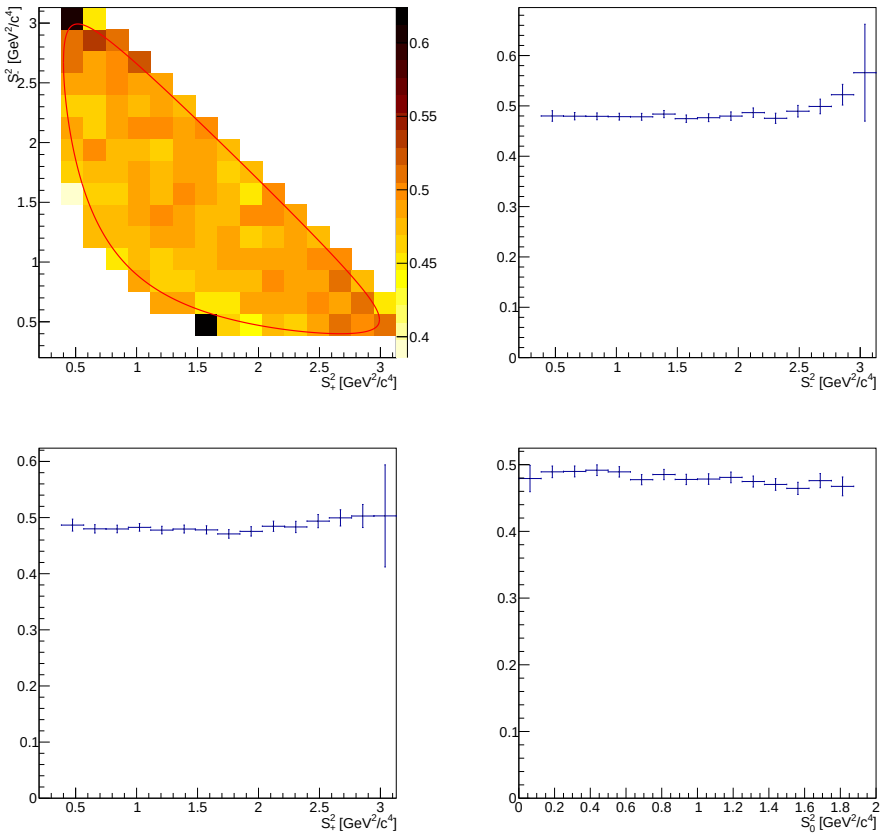


Figure 6.11.: L0 Hadron TOS efficiency weights obtained in selected $B^{\pm} \to DK^{*\pm}$ MC, for events containing DD K^0_S candidates.

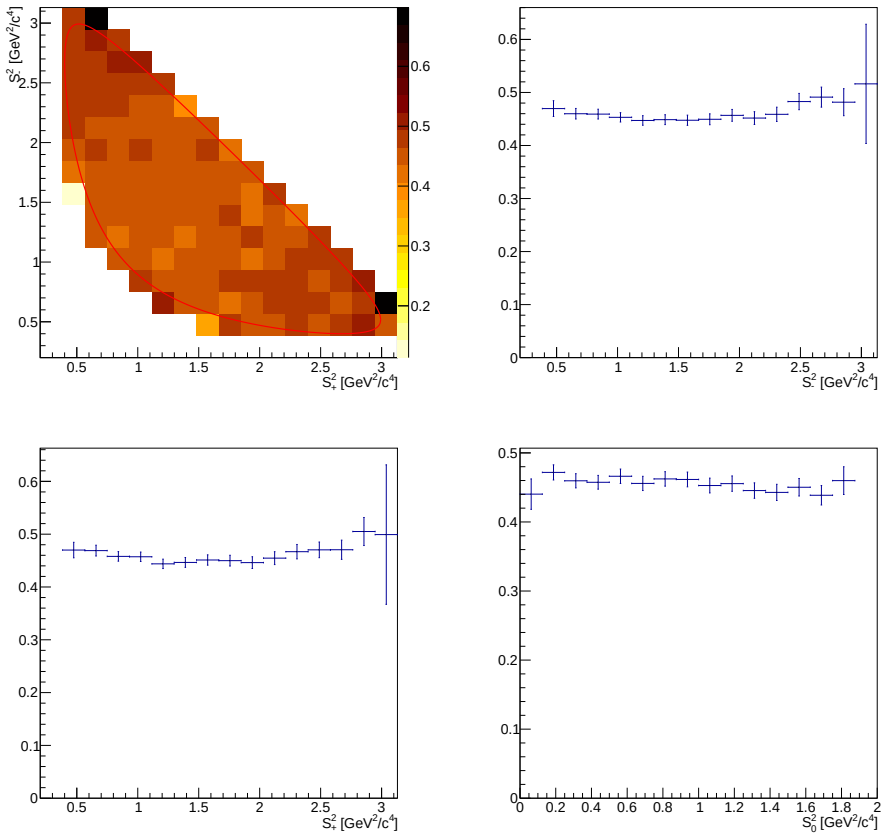


Figure 6.12.: L0 Hadron TOS efficiency weights obtained in selected $B^{\pm} \to D^* K^{\pm}$ MC, for events containing LL K_S^0 candidates.

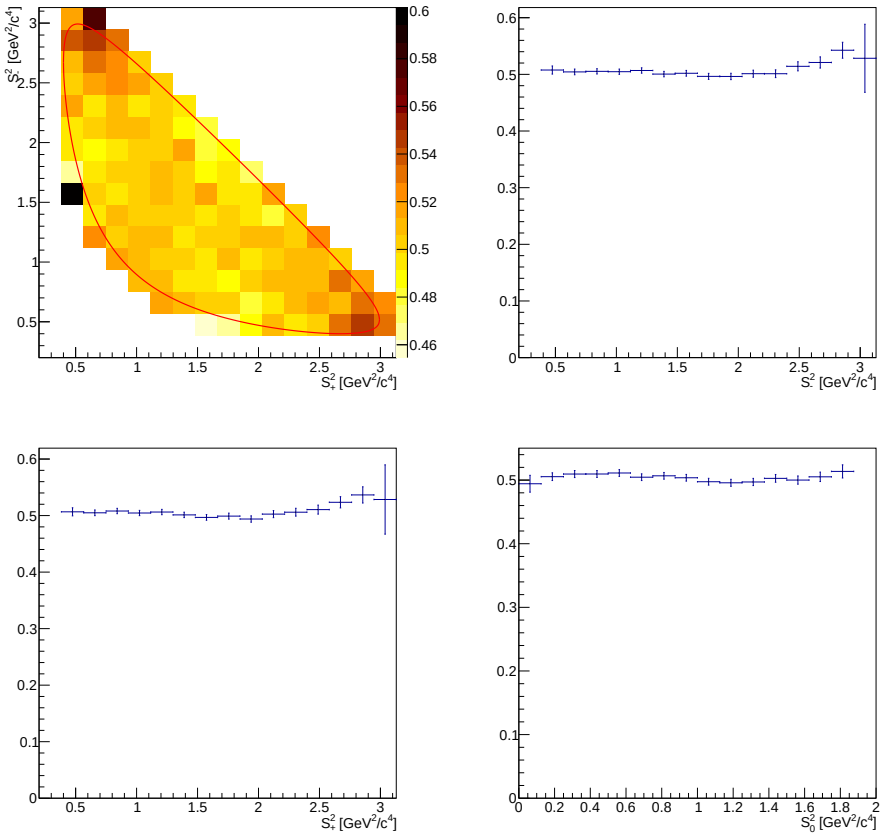


Figure 6.13.: L0 Hadron TOS efficiency weights obtained in selected $B^{\pm} \to D^* K^{\pm}$ MC, for events containing DD K_S^0 candidates.

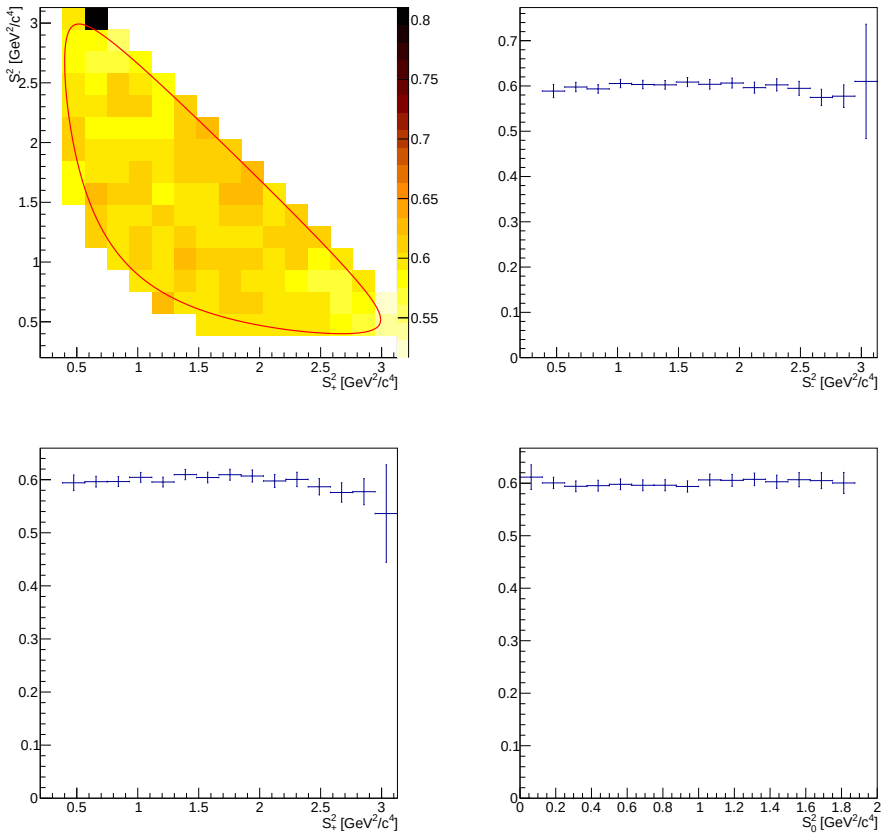


Figure 6.14.: L0 Hadron notTOS efficiency weights obtained in selected $B^{\pm} \to DK^{\pm}$ MC, for events containing LL K_S^0 candidates.

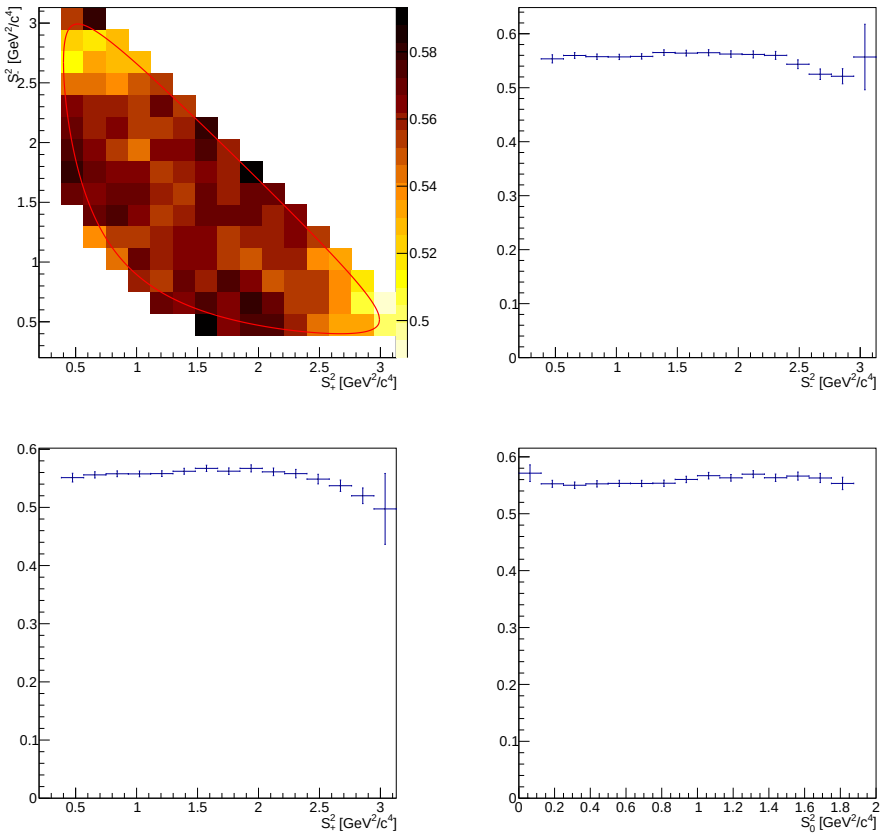


Figure 6.15.: L0 Hadron notTOS efficiency weights obtained in selected $B^{\pm} \to DK^{\pm}$ MC, for events containing DD K^0_S candidates.

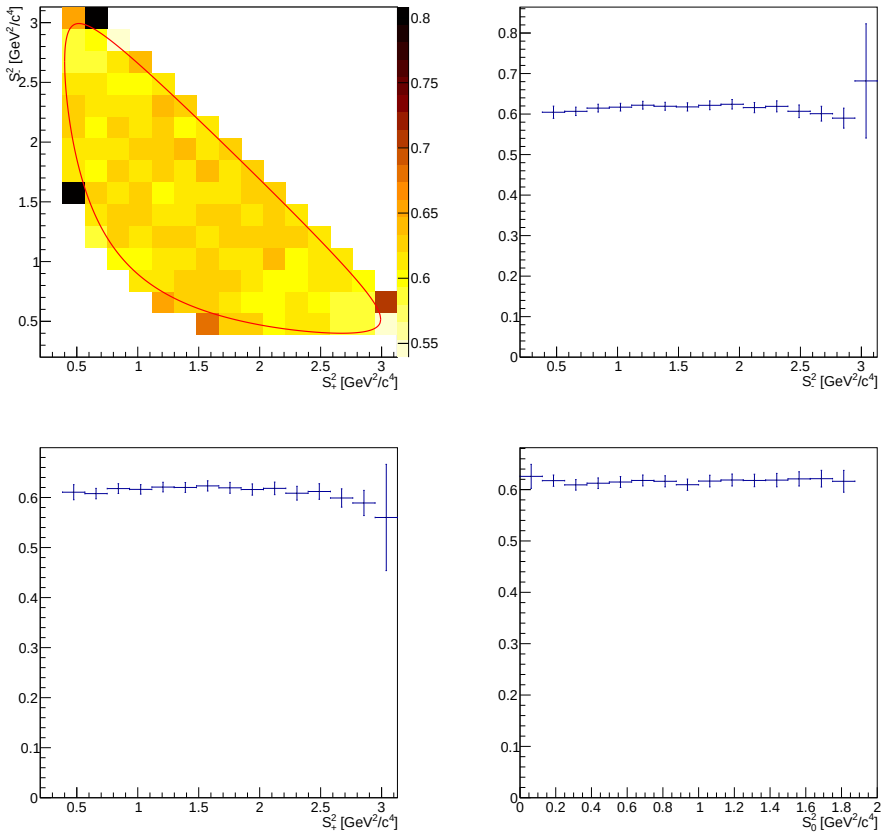


Figure 6.16.: L0 Hadron notTOS efficiency weights obtained in selected $B^0_s \to DK^{*0}$ MC, for events containing LL K^0_S candidates.

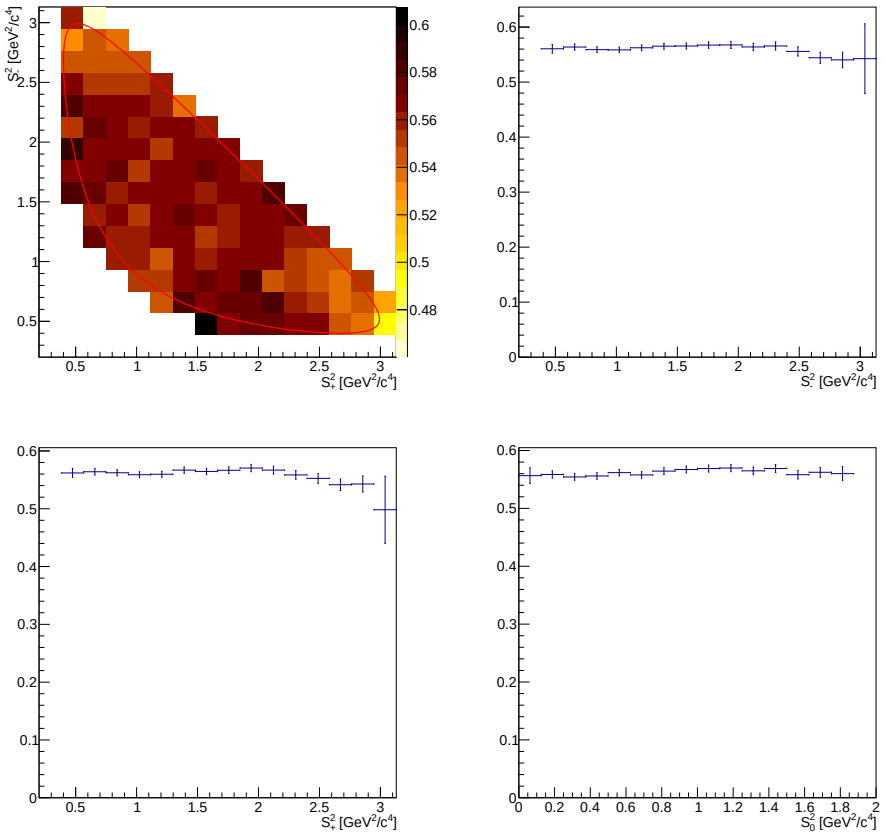


Figure 6.17.: L0 Hadron notTOS efficiency weights obtained in selected $B^0_s \to DK^{*0}$ MC, for events containing DD K^0_S candidates.

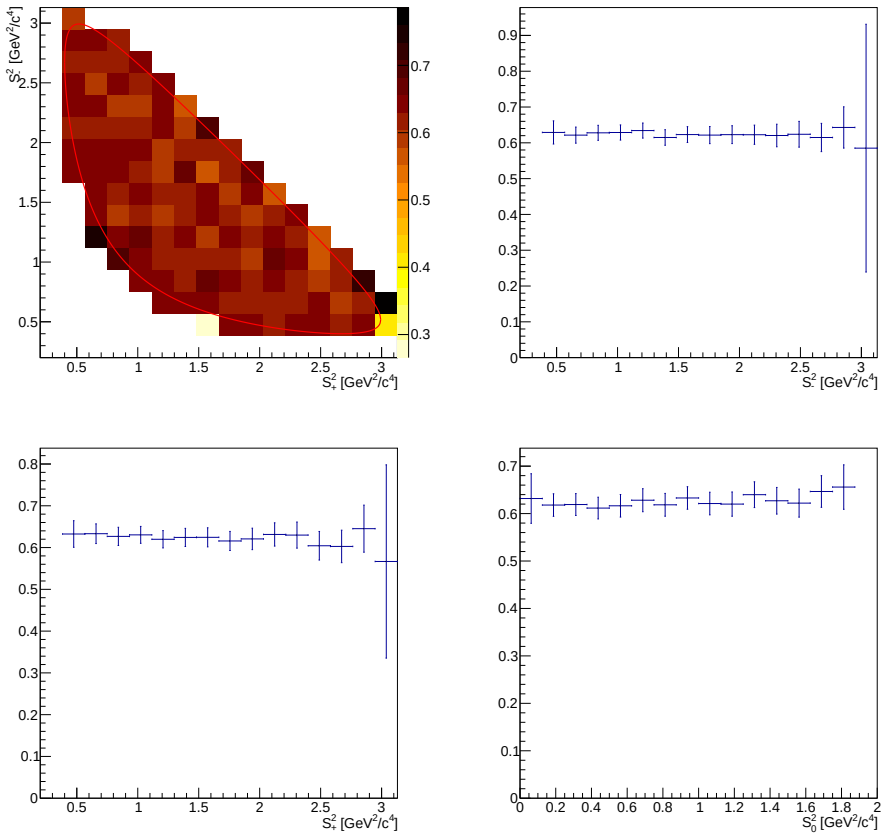


Figure 6.18.: L0 Hadron notTOS efficiency weights obtained in selected $B^{\pm} \to DK^{*\pm}$ MC, for events containing LL K_S^0 candidates.

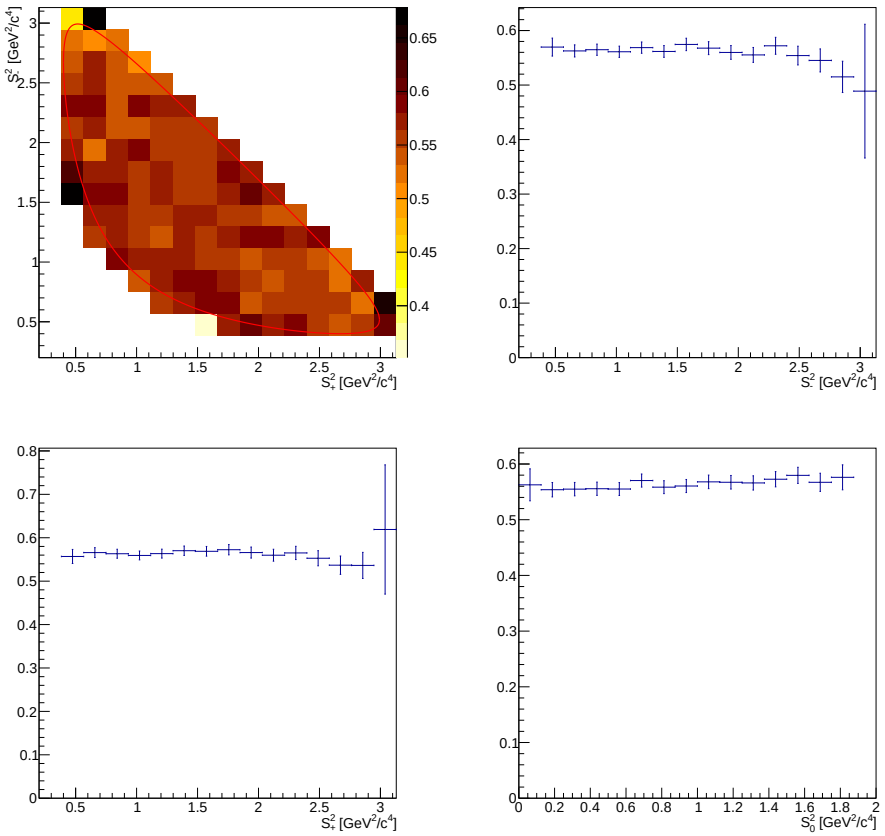


Figure 6.19.: L0 Hadron notTOS efficiency weights obtained in selected $B^{\pm} \to DK^{*\pm}$ MC, for events containing DD K^0_S candidates.

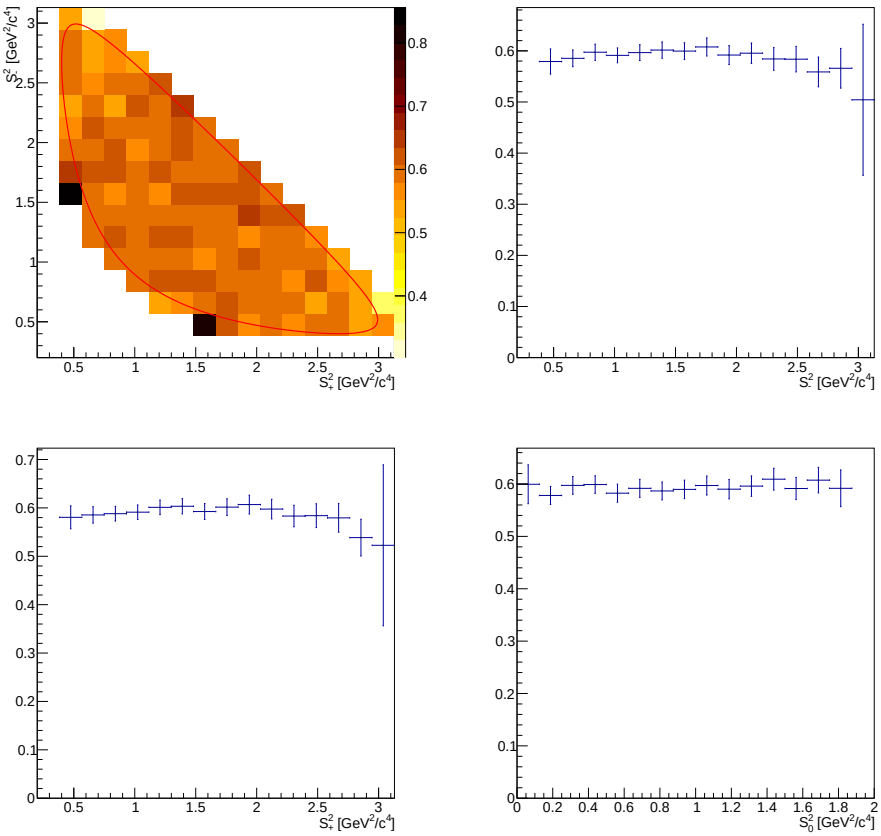


Figure 6.20.: L0 Hadron notTOS efficiency weights obtained in selected $B^{\pm} \to D^* K^{\pm}$ MC, for events containing LL K^0_S candidates.

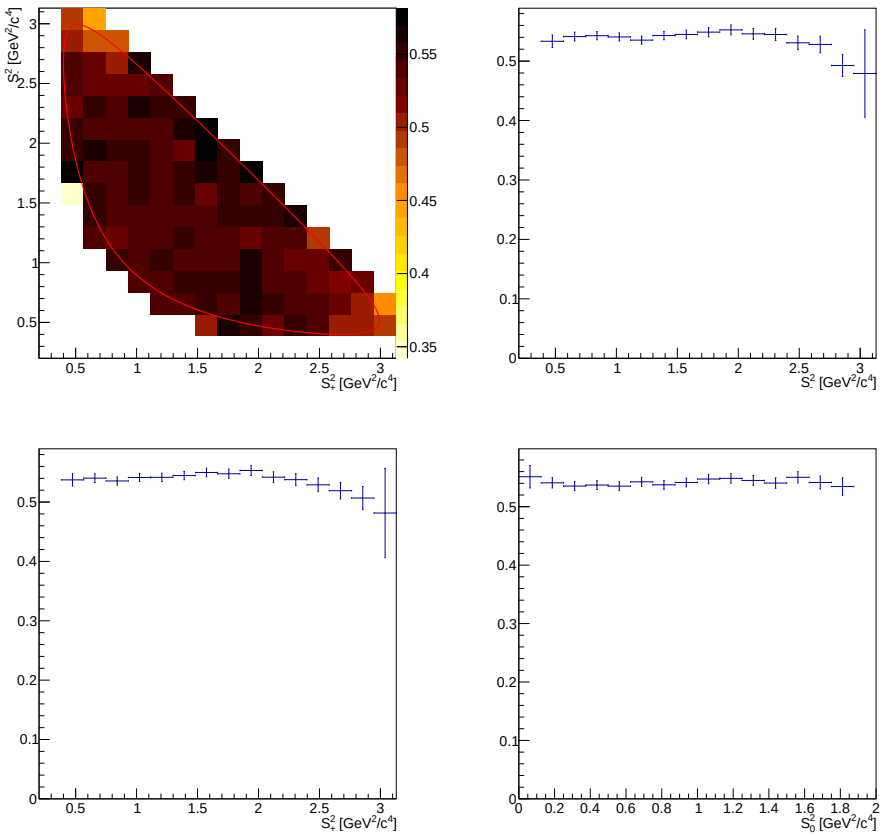


Figure 6.21.: L0 Hadron notTOS efficiency weights obtained in selected $B^{\pm} \to D^* K^{\pm}$ MC, for events containing DD K^0_S candidates.

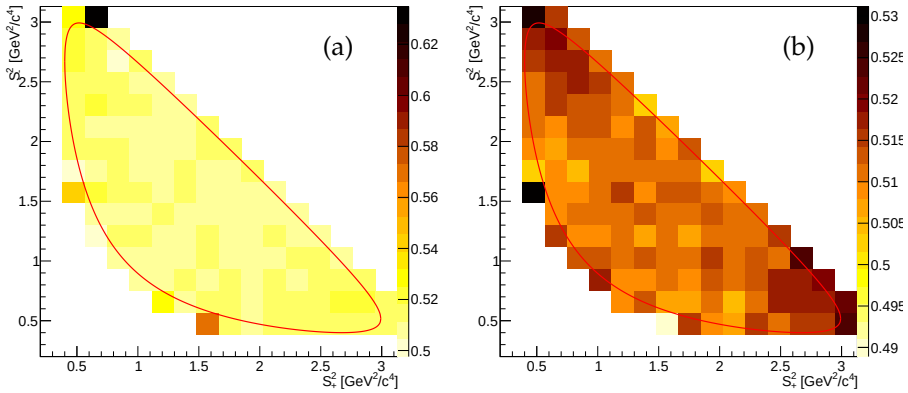


Figure 6.22.: Combined L0 Hadron TOS and notTOS efficiency weights obtained in selected $B^{\pm} \to D K^{\pm}$ MC, for events containing (a) LL and (b) DD K^0_S candidates.

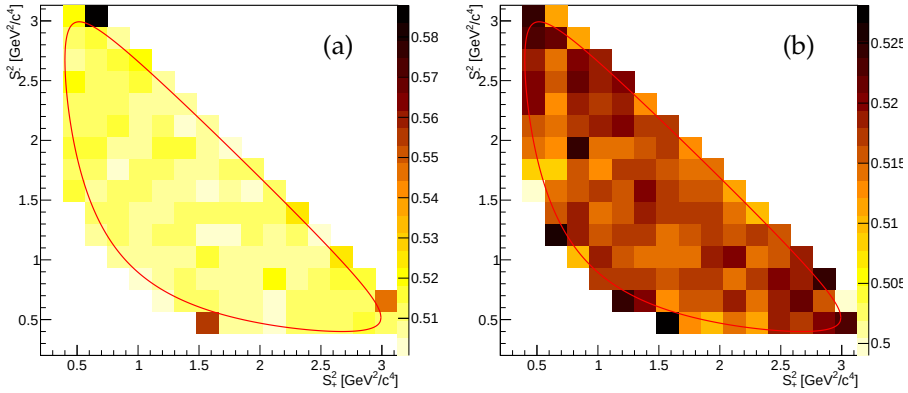


Figure 6.23.: Combined L0 Hadron TOS and notTOS efficiency weights obtained in selected $B_s^0 \to DK^{*0}$ MC, for events containing (a) LL and (b) DD K_S^0 candidates.

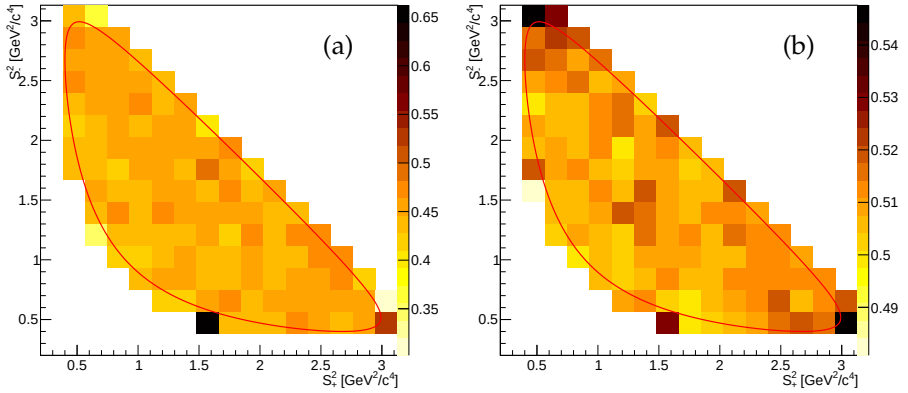


Figure 6.24.: Combined L0 Hadron TOS and notTOS efficiency weights obtained in selected $B^\pm \to DK^{*\pm}$ MC, for events containing (a) LL and (b) DD K_S^0 candidates.

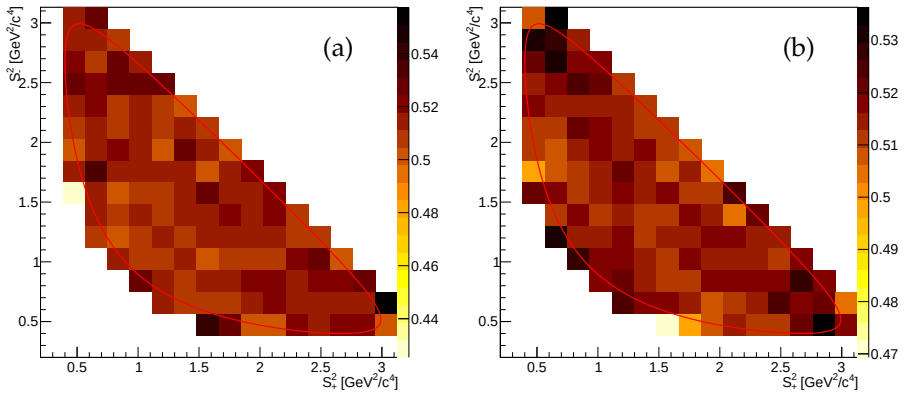


Figure 6.25.: Combined L0 Hadron TOS and notTOS efficiency weights obtained in selected $B^\pm \to D^*K^\pm$ MC, for events containing (a) LL and (b) DD K_S^0 candidates.

In order to obtain the L0 trigger weights as described above, HLT requirements are not applied on the samples used. The inclusion of these requirements gives a very small additional non-uniformity, which may be seen by taking the bin-by-bin ratio of the histograms obtained for events selected with and without the requirements applied, after all other selection criteria have been applied. The resulting HLT efficiencies for each mode are shown in Figures 6.26–6.29, HLT contributions are included in the final efficiency histograms for each mode.

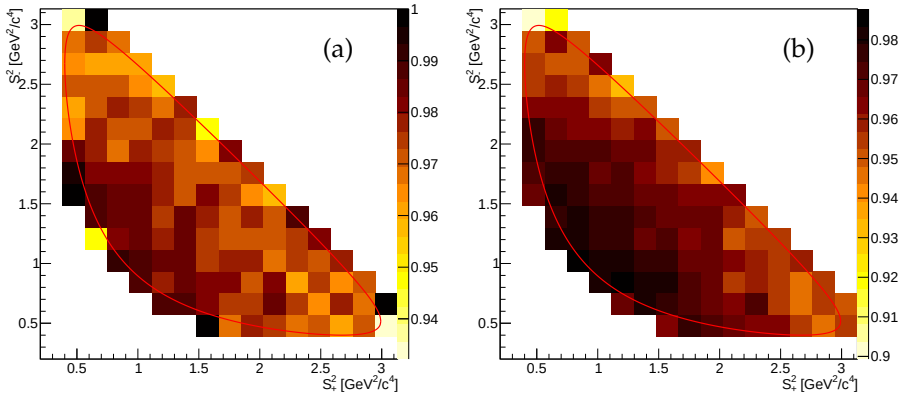


Figure 6.26.: HLT efficiency weights obtained in selected $B^{\pm} \to DK^{\pm}$ MC, for events containing (a) LL and (b) DD K^0_S candidates.

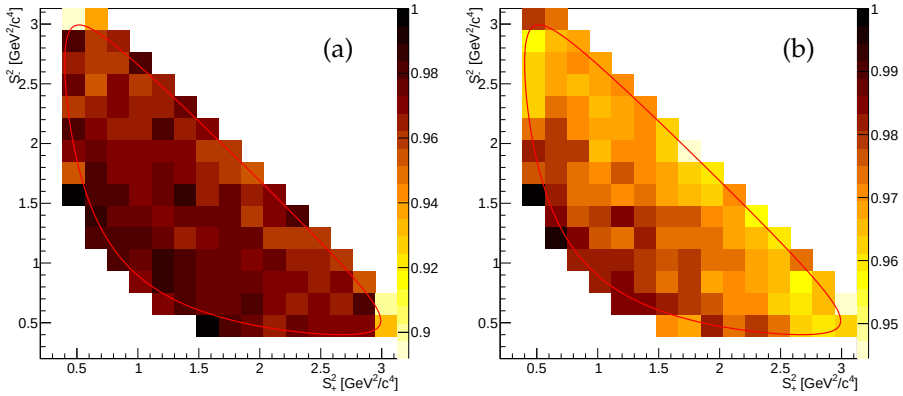


Figure 6.27.: HLT efficiency weights obtained in selected $B^0 \to DK^{*0}$ MC, for events containing (a) LL and (b) DD K^0_S candidates.

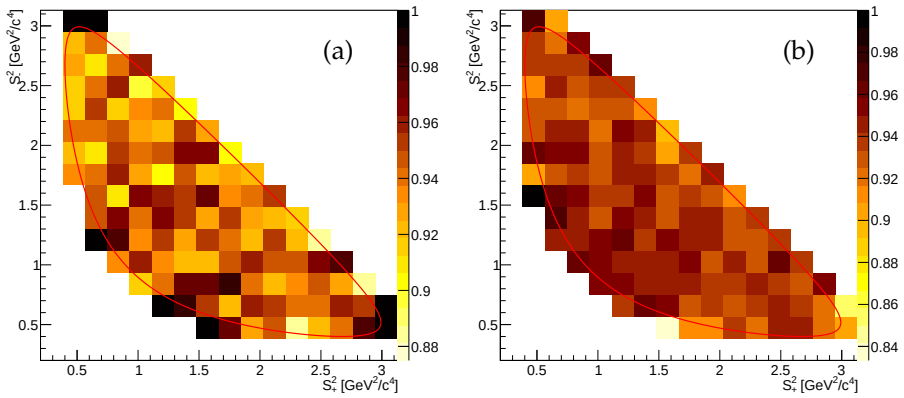


Figure 6.28.: HLT efficiency weights obtained in selected $B^{\pm} \to DK^{*\pm}$ MC, for events containing (a) LL and (b) DD K_S^0 candidates.

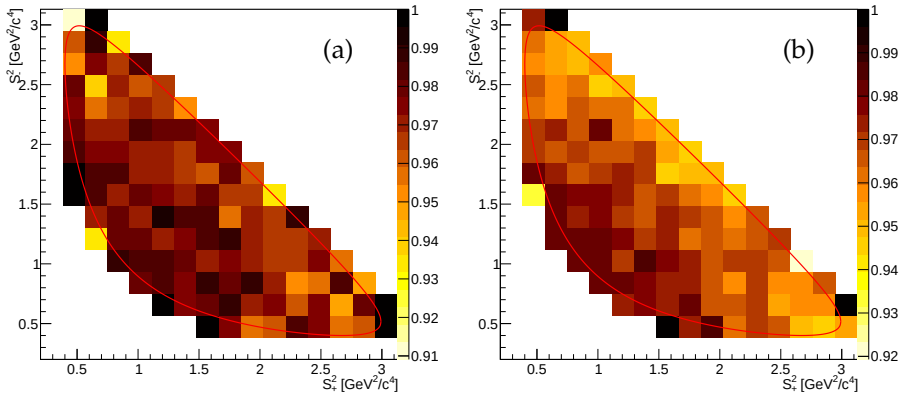


Figure 6.29.: HLT efficiency weights obtained in selected $B^{\pm} \to D^*K^{\pm}$ MC, for events containing (a) LL and (b) DD K_S^0 candidates.

6.3. PID weights

The distributions of PID variables are not well-described in simulation, and instead a standard tool, `PIDCalib` [148][2], has been developed by LHCb to provide PID efficiency histograms calibrated on data for different species of final state particles. For pion and kaon candidates, very pure calibration samples have been obtained at LHCb from decays of the form $D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$ [151].

For each PID requirement applied within the present analysis, as specified in Table 4.7, a two-dimensional calibration histogram is constructed, binned in momentum, p, and pseudorapidity, η. The same binning scheme is used in each case, with bin boundaries defined by

- p: $[2, 9.3, 15.6, 19] \cup 15$ uniform bins within $[19, 100] \cup [100, 1000]$ GeV; and
- η: $[1.5, 2.4, 3.15, 3.9, 5]$.

This binning scheme has been designed by the creators of the `PIDCalib` tool, in order to give approximately constant PID efficiency within each bin for the calibration samples. This is achieved by beginning with a fine-grained uniform binning scheme, and iteratively merging adjacent bins whose efficiencies are consistent to within some set tolerance [148].

These calibration histograms are used to assign weights, w^t_{PID}, to each track for which a PID requirement is placed in the selection. Per-event weights, w^e_{PID}, are constructed by taking the product of all such weights within each event,

$$w^e_{\text{PID}} = \prod_t w^t_{\text{PID}}. \tag{6.18}$$

Finally, histograms decribing the PID efficiency, ϵ^i_{PID}, are obtained by taking the average of these per-event weights within bins of the Dalitz plane. Histograms obtained for each reference sample are shown in Figures 6.30–6.37. In all cases the PID efficiency is very nearly flat over the Dalitz plane, since no PID requirements are placed on the D-meson or its daughters.

[2] Version v10r4.

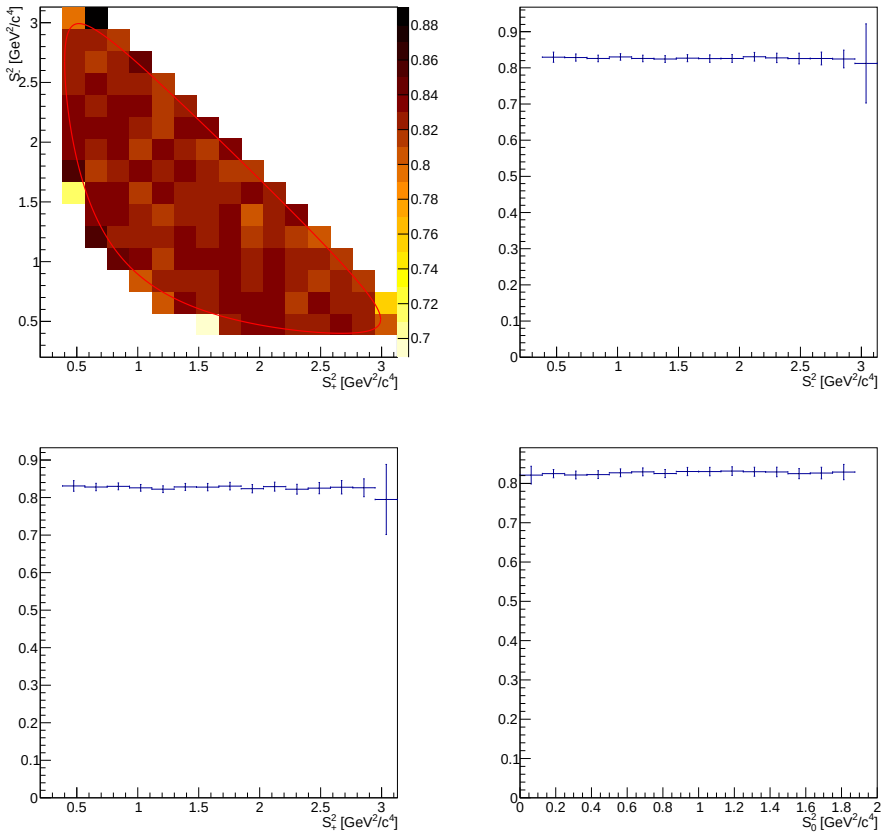


Figure 6.30.: PID efficiency weights obtained in selected $B^{\pm} \to DK^{\pm}$ MC, for events containing LL K^0_S candidates.

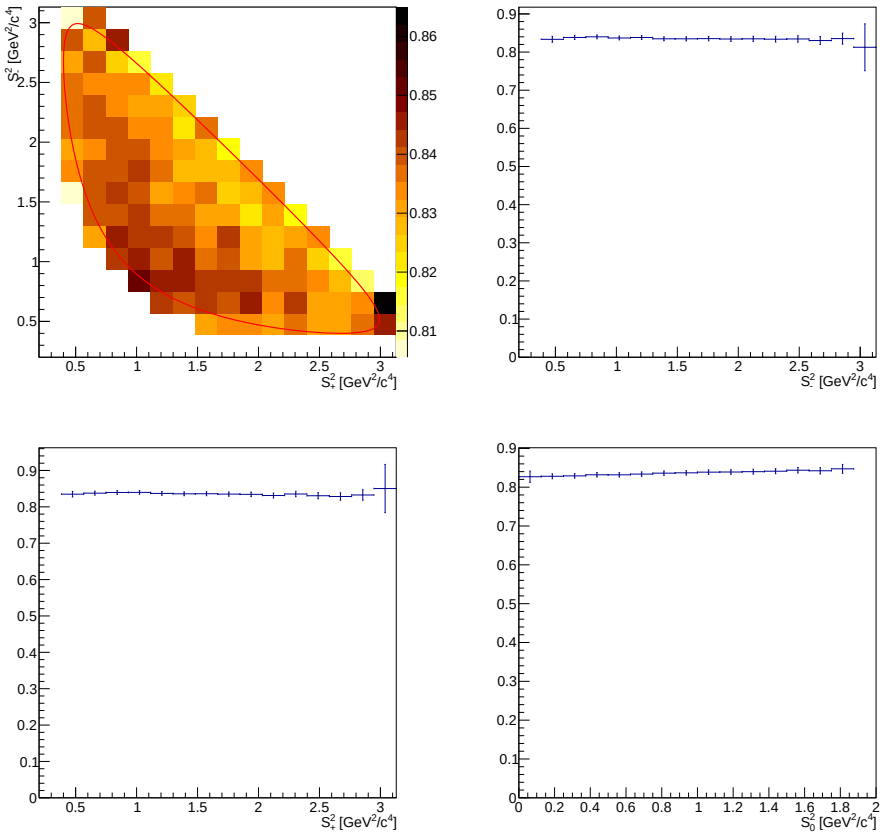


Figure 6.31.: PID efficiency weights obtained in selected $B^\pm \to DK^\pm$ MC, for events containing DD K^0_S candidates.

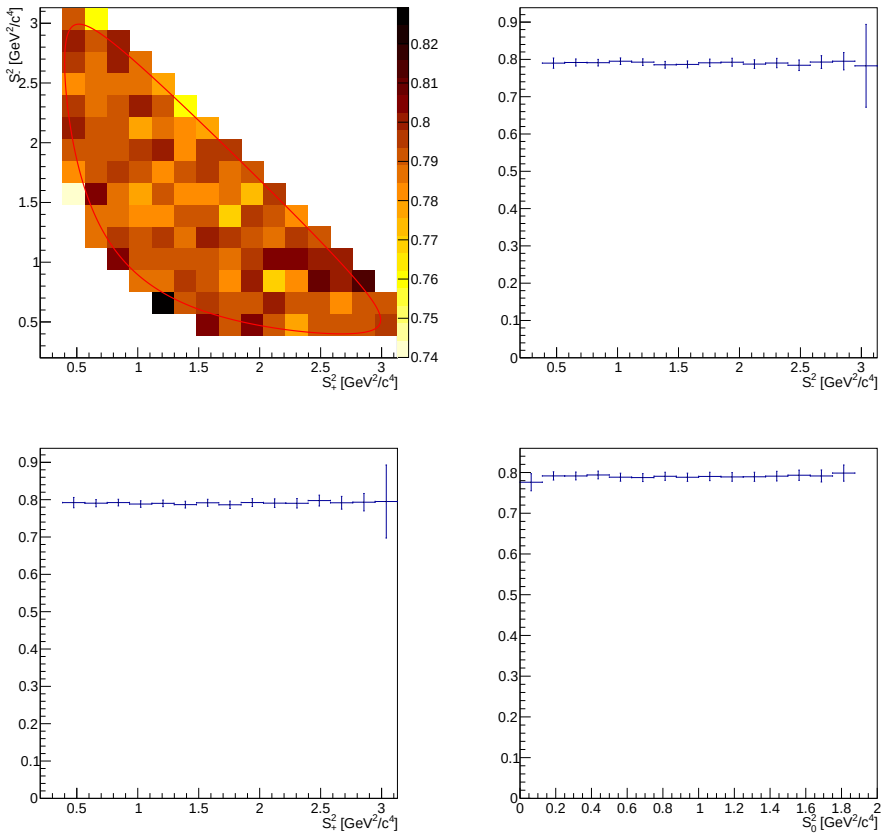


Figure 6.32.: PID efficiency weights obtained in selected $B_s^0 \to DK^{*0}$ MC, for events containing LL K_S^0 candidates.

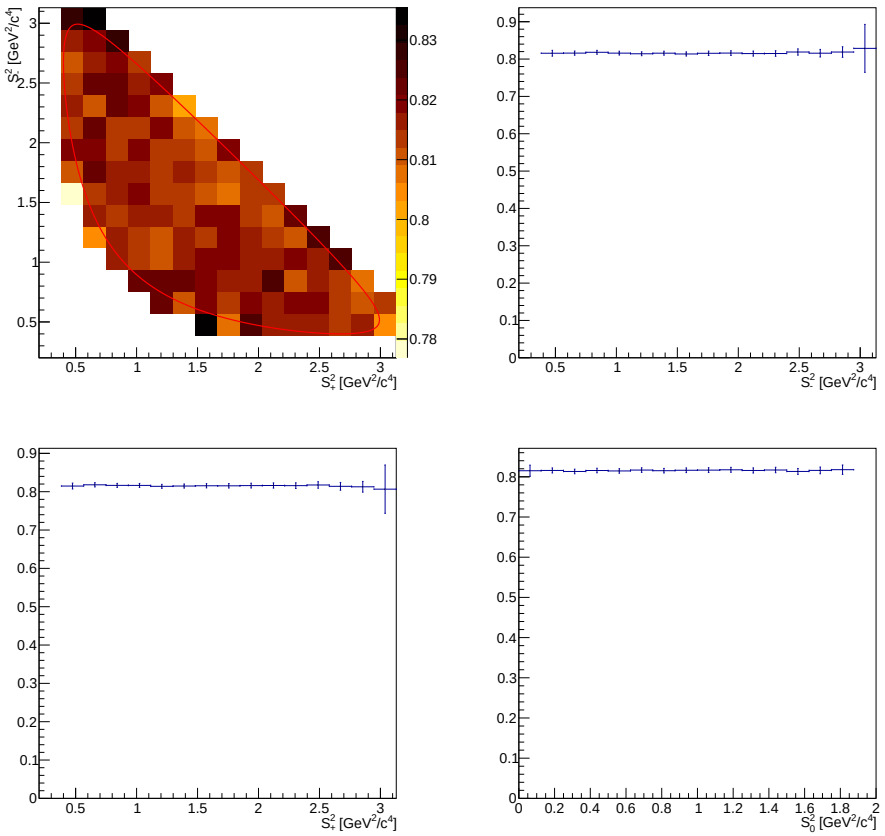


Figure 6.33.: PID efficiency weights obtained in selected $B_s^0 \to DK^{*0}$ MC, for events containing DD K_S^0 candidates.

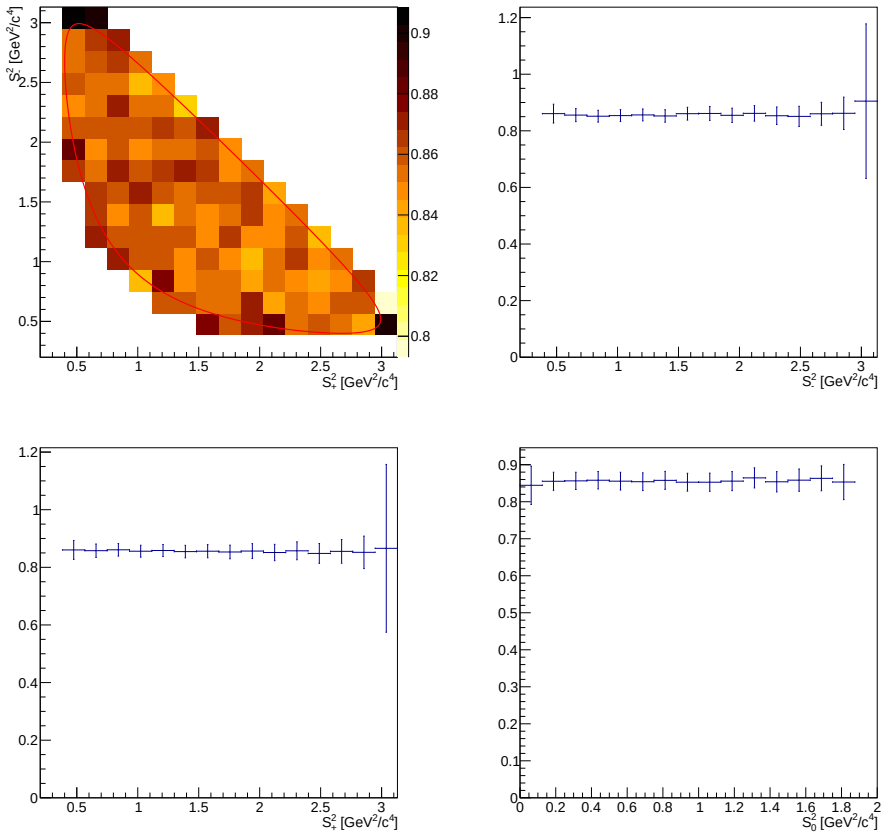


Figure 6.34.: PID efficiency weights obtained in selected $B^{\pm} \to DK^{*\pm}$ MC, for events containing LL K_S^0 candidates.

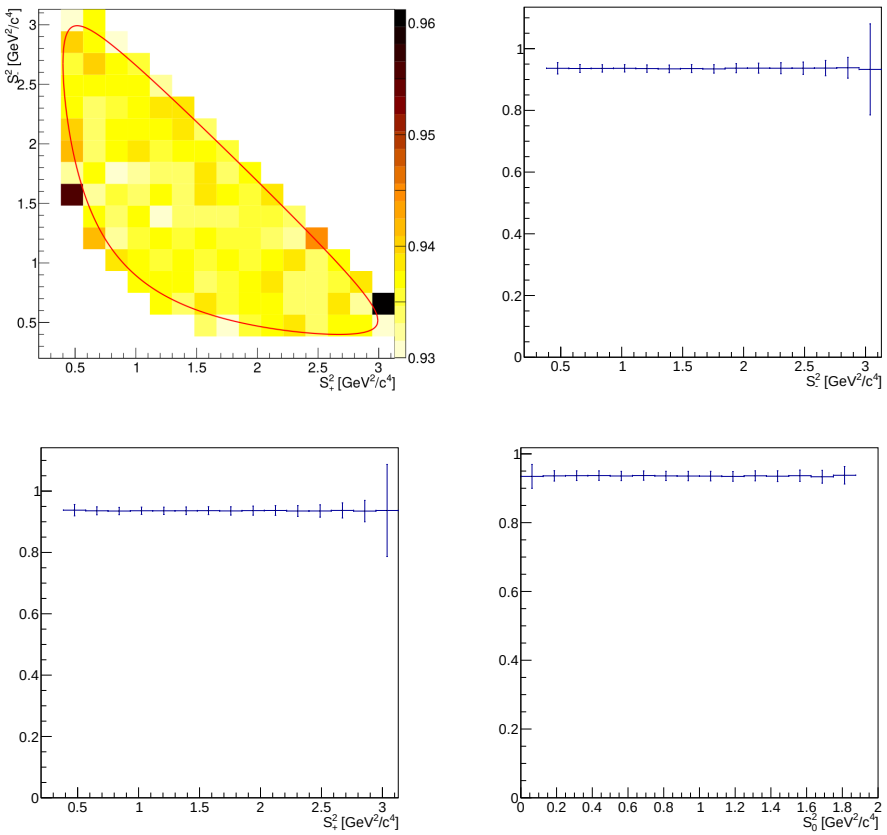


Figure 6.35.: PID efficiency weights obtained in selected $B^{\pm} \to DK^{*\pm}$ MC, for events containing DD K_S^0 candidates.

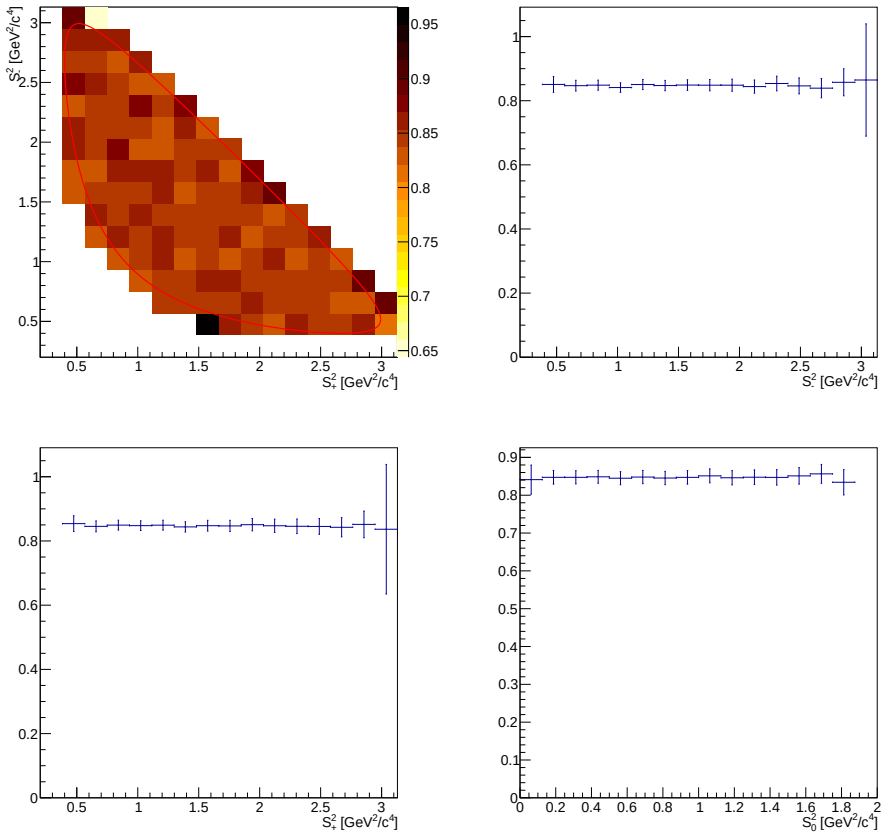


Figure 6.36.: PID efficiency weights obtained in selected $B^{\pm} \to D^* K^{\pm}$ MC, for events containing LL K_S^0 candidates.

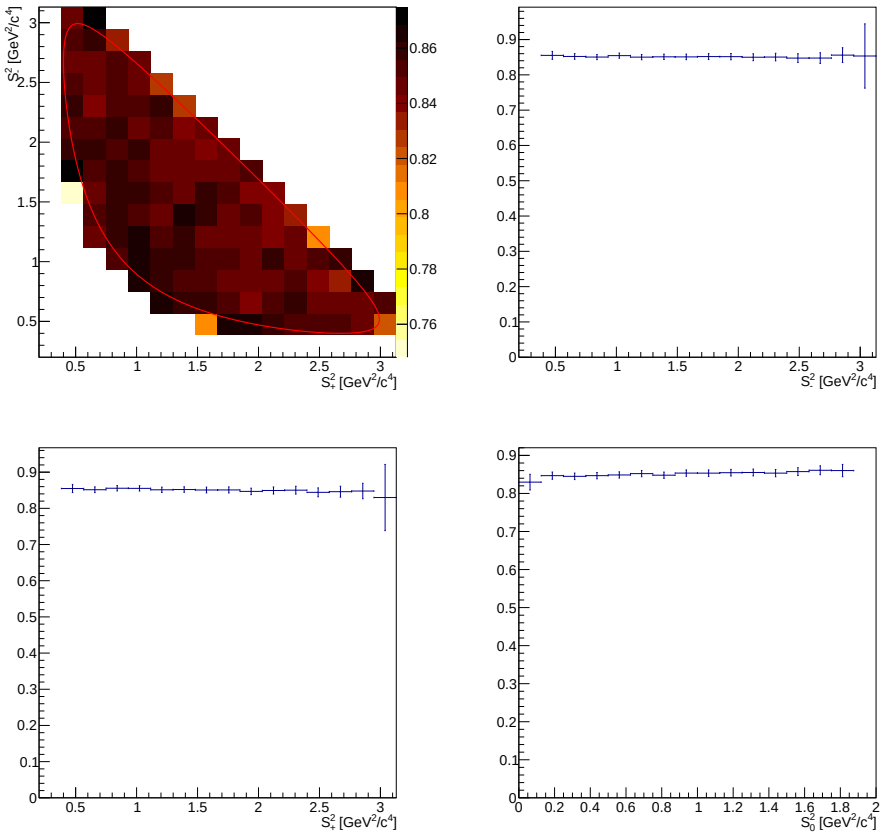


Figure 6.37.: PID efficiency weights obtained in selected $B^\pm \to D^* K^\pm$ MC, for events containing DD K^0_S candidates.

6.4. Total efficiency

To obtain the final efficiency histograms for each mode, the histograms obtained directly from selected MC events are multiplied bin-by-bin with the edge-bin, L0 and PID weights. The resulting histograms are symmetrised about the leading diagonal, since each selection applies identical requirements to the two pion daughters of the D-meson, and so no asymmetry is expected. The final efficiency histograms obtained for each mode are shown in Figures 6.38 and 6.39.

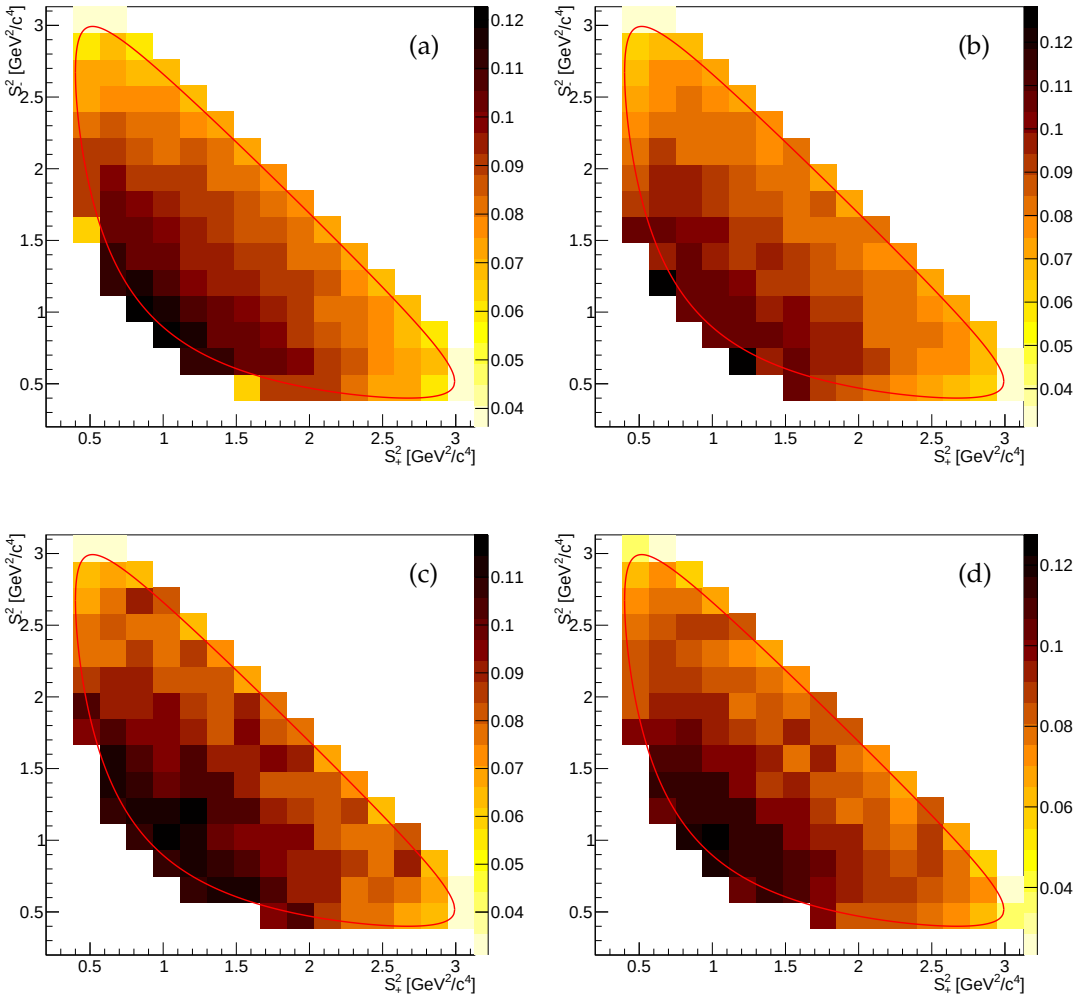


Figure 6.38.: Full selection efficiency histograms obtained for (a) $B^\pm \to DK^\pm$, (b) $B^0 \to DK^{*0}$, (c) $B^\pm \to DK^{*\pm}$, and (d) $B^\pm \to D^*K^\pm$ MC events containing LL K^0_S candidates. Edge-bin, trigger and PID weights are included, and each histogram is symmetrised about the leading diagonal.

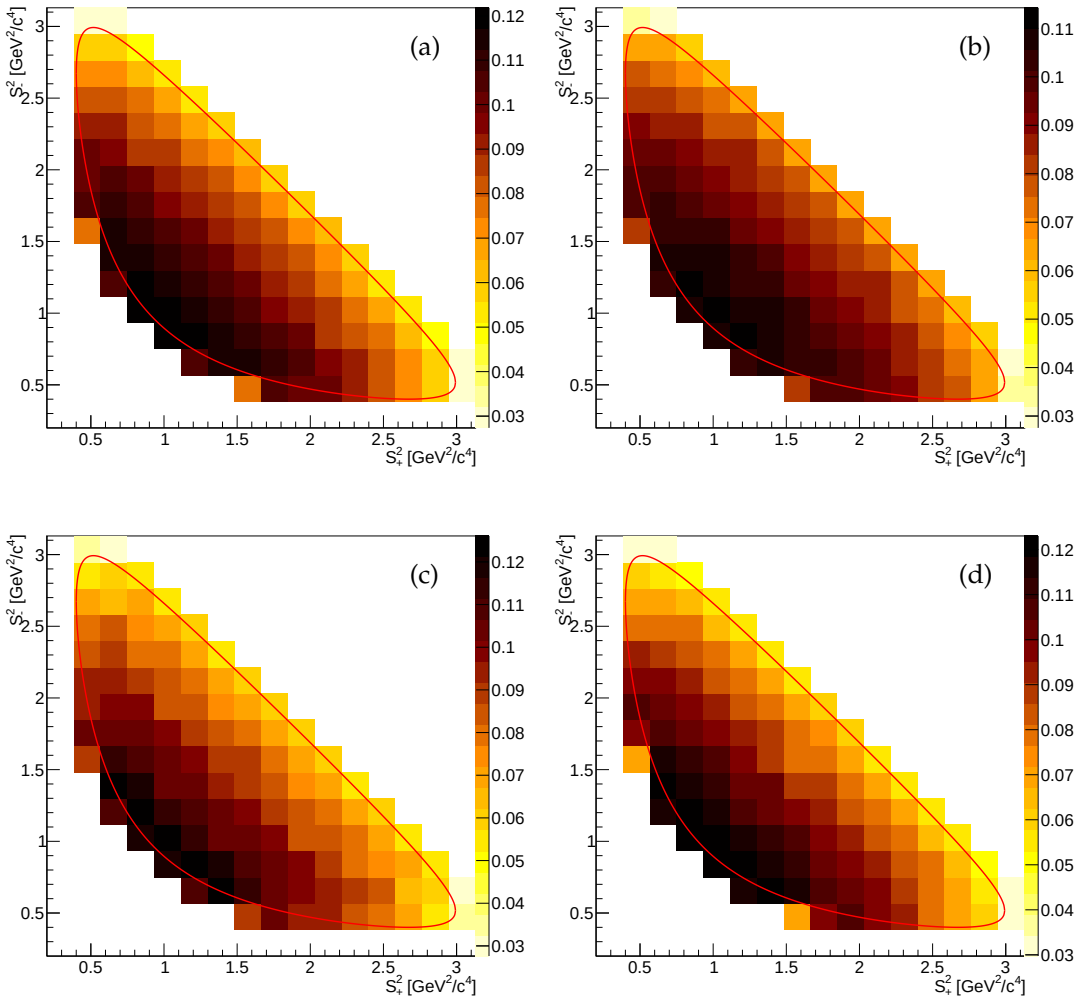


Figure 6.39.: Full selection efficiency histograms obtained for (a) $B^\pm \to DK^\pm$, (b) $B^0 \to DK^{*0}$, (c) $B^\pm \to DK^{*\pm}$, and (d) $B^\pm \to D^*K^\pm$ MC events containing DD K^0_S candidates. Edge-bin, trigger and PID weights are included, and each histogram is symmetrised about the leading diagonal.

7. Measurement of the CKM angle γ using $B^0 \to DK^{*0}$ with $D \to K^0_S\pi^+\pi^-$ decays

The work described so far in this thesis develops the foundations for a simultaneous Dalitz analysis of multiple decay modes of the form $B \to D^{(*)}h^{(*)}$. This approach represents a significant improvement on the method of statistically combining multiple single-mode analyses after the fact, since it reduces the number of degrees of freedom required to fit the distribution of candidates in the Dalitz plane, and allows cross-feed of signal decays between different decay modes to be modelled accurately. As such, it is expected to improve the sensitivity to γ that can be achieved using the available dataset, and the benefit it offers is expected to increase as more statistics become available.

However, systematic studies have not yet been performed for this simultaneous analysis. This task is left for future work, but it is natural to take as a starting point the studies performed for the previous single-mode analysis published in reference [1], to which the present author was a main contributor. Specifically, my contributions were focused on the selection procedure, mass fit and $\mathcal{CP}$ fit, and the evaluation of the systematic uncertainities arising from the amplitude model used in the $\mathcal{CP}$ fit. I was also responsible for managing the generation of several of the MC samples used. Another main contributor worked in parallel on the selection procedure and mass fit[1], and evaluated the systematic uncertainties arising from the PDF model used in the mass fit.

This publication is reproduced in full in the following section, but first it is useful to note some small differences relative to the present combined analysis, in the selection criteria used to choose candidate decays and the PDF model used in the fit to the B-meson mass spectrum. These are described below.

Selection of candidate decays

The mass requirements placed on D^0- and K^0_S-meson candidates differ slightly from those applied in the present analysis. The reconstructed D^0-meson mass is required to lie within 30 MeV of its expected value, instead of 25 MeV. The reconstructed K^0_S-meson mass is required to lie within 14.4(19.9) MeV of its expected value for LL(DD) candidates, whilst the present analysis uses 15 MeV in both cases.

The multivariate analysis stage of the selection makes use of boosted decision tree classifers instead of neural networks as in Chapter 4. Background training samples are taken from the upper sideband of the B-meson reconstructed mass spectrum beginning at 5500 MeV, instead of 5800 MeV. This region overlaps with the fitting region, so to prevent bias a 'crossed' architecture is used: the background training data is split into two subsamples, with one BDT trained using the first subsample and another trained using the second. Each BDT is then used to select data events from the subsample that it was not trained on. In total, four BDTs are trained, since separate classifiers are used to select LL and DD candidates.

[1] Two competing selection procedures were considered – one using a neural network classifier, and the other using a boosted decision tree classifier.

The cosine of the K^{*0}-meson helicity angle, $\cos(\theta^{\text{helicity}}_{K^{*0}})$ is included as an input to the BDT classifiers, whilst the D-meson lifetime, τ_D, is not. In addition, the inclusion of the impact parameter χ^2, with respect to the PV, for each final state candidate, is treated differently to the present analysis. Instead of constructing the six variables

- $\ln(\max((\chi^2)^{\text{IP,PV}}_{\pi^+_D}, (\chi^2)^{\text{IP,PV}}_{\pi^-_D}))$,
- $\ln(\min((\chi^2)^{\text{IP,PV}}_{\pi^+_D}, (\chi^2)^{\text{IP,PV}}_{\pi^-_D}))$,
- $\ln(\max((\chi^2)^{\text{IP,PV}}_{\pi^+_{K^0_S}}, (\chi^2)^{\text{IP,PV}}_{\pi^-_{K^0_S}}))$,
- $\ln(\min((\chi^2)^{\text{IP,PV}}_{\pi^+_{K^0_S}}, (\chi^2)^{\text{IP,PV}}_{\pi^-_{K^0_S}}))$,
- $\ln((\chi^2)^{\text{IP,PV}}_{\pi_{K^{*0}}})$, and
- $\ln((\chi^2)^{\text{IP,PV}}_{K_{K^{*0}}})$,

two variables are constructed, the first combining the impact parameter χ^2 of both daughter pions originating from the K^0_S-meson, and the second combining the impact parameter χ^2 of all other final state particles:

- $\ln((\chi^2)^{\text{IP,PV}}_{\pi^+_{K^0_S}} + (\chi^2)^{\text{IP,PV}}_{\pi^-_{K^0_S}})$; and
- $\ln((\chi^2)^{\text{IP,PV}}_{\pi^+_D} + (\chi^2)^{\text{IP,PV}}_{\pi^-_D} + (\chi^2)^{\text{IP,PV}}_{\pi_{K^{*0}}} + (\chi^2)^{\text{IP,PV}}_{K_{K^{*0}}})$.

Finally, a veto is used to exclude candidates which may arise from $B^\pm \to DK^\pm$ decays, by rejecting any candidates for which the invariant mass of the reconstructed D-meson and the final state kaon combined lies within 50 MeV of the $B^\pm$ mass.

Fit to reconstructed B-meson mass spectrum

The fit performed for this measurement includes one component not used in the fit described in Section 5.2, arising from misidentified $B^0 \to D\rho^0$ decays. These may satisfy the selection criteria for this mode if one of the ρ^0-meson's daughter pions (ρ^0 decays predominantly to $\pi^+\pi^-$ [3]) is mis-identified as a kaon. A kernel-estimated PDF shape for this component is obtained using selected MC decays.

7.1. Publication

Reference [1] is reproduced overleaf.

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)



CERN-EP-2016-089
LHCb-PAPER-2016-007
May 3, 2016

Measurement of the CKM angle γ using $B^0 \to DK^{*0}$ with $D \to K^0_S\pi^+\pi^-$ decays

The LHCb collaboration[†]

Abstract

A model-dependent amplitude analysis of the decay $B^0 \to D(K^0_S\pi^+\pi^-)K^{*0}$ is performed using proton-proton collision data corresponding to an integrated luminosity of $3.0\,\text{fb}^{-1}$, recorded at $\sqrt{s} = 7$ and $8\,\text{TeV}$ by the LHCb experiment. The *CP* violation observables $x_\pm$ and $y_\pm$, sensitive to the CKM angle γ, are measured to be

$$\begin{aligned} x_- &= -0.15 \pm 0.14 \pm 0.03 \pm 0.01, \\ y_- &= 0.25 \pm 0.15 \pm 0.06 \pm 0.01, \\ x_+ &= 0.05 \pm 0.24 \pm 0.04 \pm 0.01, \\ y_+ &= -0.65\,^{+0.24}_{-0.23} \pm 0.08 \pm 0.01, \end{aligned}$$

where the first uncertainties are statistical, the second systematic and the third arise from the uncertainty on the $D \to K^0_S\pi^+\pi^-$ amplitude model. These are the most precise measurements of these observables. They correspond to $\gamma = (80^{+21}_{-22})^\circ$ and $r_{B^0} = 0.39 \pm 0.13$, where r_{B^0} is the magnitude of the ratio of the suppressed and favoured $B^0 \to DK^+\pi^-$ decay amplitudes, in a $K\pi$ mass region of $\pm 50\,\text{MeV}$ around the $K^*(892)^0$ mass and for an absolute value of the cosine of the K^{*0} decay angle larger than 0.4.

© CERN on behalf of the LHCb collaboration, licence CC-BY-4.0.

[†]Authors are listed at the end of this paper.

Contents

1 Introduction **266**

2 The LHCb detector **268**

3 Candidate selection and background sources **269**

4 Efficiency across the phase space **271**

5 Analysis strategy and fit results **272**
5.1 Invariant mass fit of $B^0 \to DK^{*0}$ candidates 272
5.2 *CP* fit 273

6 Systematic uncertainties **276**

7 Determination of the parameters γ, r_{B^0} and δ_{B^0} **281**

8 Conclusion **283**

References **285**

1 Introduction

The Standard Model can be tested by checking the consistency of the Cabibbo-Kobayashi-Maskawa (CKM) mechanism [1,2], which describes the mixing between weak and mass eigenstates of the quarks. The CKM phase γ can be expressed in terms of the elements of the complex unitary CKM matrix, as $\gamma \equiv \arg\left[-V_{ud}V_{ub}{}^*/V_{cd}V_{cb}{}^*\right]$. Since γ is also the angle of the unitarity triangle least constrained by direct measurements, its precise determination is of considerable interest. Its value can be measured in tree-level processes such as $B^\pm \to DK^\pm$ and $B^0 \to DK^{*0}$, where D is a superposition of the D^0 and $\overline{D}{}^0$ flavour eigenstates, and K^{*0} is the $K^*(892)^0$ meson. Since loop corrections to these processes are of higher order, the associated theoretical uncertainty on γ is negligible [3]. As such, measurements of γ in tree-level decays provide a reference value, allowing searches for potential deviations due to physics beyond the Standard Model in other processes.

The combination of measurements by the BaBar [4] and Belle [5] collaborations gives $\gamma = (67 \pm 11)^\circ$ [6], whilst an average value of LHCb determinations in 2014 gave $\gamma = \left(73^{+9}_{-10}\right)^\circ$ [7]. Global fits of all current CKM measurements by the CKMfitter [8,9] and UTfit [10] collaborations yield indirect estimates of γ with an uncertainty of 2°. Some of the CKM measurements included in these combinations can be affected by new physics contributions.

Since the phase difference between V_{ub} and V_{cb} depends on γ, the determination of γ in tree-level decays relies on the interference between $b \to c$ and $b \to u$ transitions. The strategy of using $B^\pm \to DK^\pm$ decays to determine γ from an amplitude analysis of D-meson decays to the three-body final state $K^0_S\pi^+\pi^-$ was first proposed in Refs. [11, 12]. The method requires

knowledge of the $D \to K^0_S\pi^+\pi^-$ decay amplitude across the phase space, and in particular the variation of its strong phase. This may be obtained either by using a model to describe the D-meson decay amplitude in phase space (model-dependent approach), or by using measurements of the phase behaviour of the amplitude (model-independent approach). The model-independent strategy, used by Belle [13] and LHCb [14,15], incorporates measurements from CLEO [16] of the D decay strong phase in bins across the phase space. The present paper reports a new unbinned model-dependent measurement, following the method used by the BaBar [17–19], Belle [20–22] and LHCb [23] collaborations in their analyses of $B^\pm \to D^{(*)}K^{(*)\pm}$ decays. This method allows the statistical power of the data to be fully exploited.

The sensitivity to γ depends both on the yield of the sample analysed and on the magnitude of the ratio r_B of the suppressed and favoured decay amplitudes in the relevant region of phase space. Due to colour suppression, the branching fraction $\mathcal{B}(B^0 \to \overline{D}^0 K^{*0}) = (4.2 \pm 0.6) \times 10^{-5}$ is an order of magnitude smaller than that of the corresponding charged B-meson decay mode, $\mathcal{B}(B^+ \to \overline{D}^0 K^+) = (3.70 \pm 0.17) \times 10^{-4}$ [24]. However, this is partially compensated by an enhancement in r_{B^0}, which was measured to be $r_{B^0} = 0.240^{+0.055}_{-0.048}$ in $B^0 \to DK^{*0}$ decays in which the D is reconstructed in two-body final states [25]; the charged decays have an average value of $r_B = 0.097 \pm 0.006$ [8,9]. Model-dependent and independent determinations of γ using $B^0 \to D(K^0_S\pi^+\pi^-)K^{*0}$ decays have already been performed by the BaBar [26] and Belle [27] collaborations, respectively. The model-independent approach has also been employed recently by LHCb [28]. For these decays a time-independent CP analysis is performed, as the K^{*0} is reconstructed in the self-tagging mode $K^+\pi^-$, where the charge of the kaon provides the flavour of the decaying neutral B meson.

The K^{*0} meson is one of several possible states of the $(K^+\pi^-)$ system. Letting X^0_s represent any such state, the B-meson decay amplitude to $DK^+\pi^-$ may be expressed as a superposition of favoured $b \to c$ and suppressed $b \to u$ contributions:

$$\begin{aligned} \mathcal{A}(\overline{B}^0 \to D\overline{X}^0_s) &\propto |A_c|A_f + |A_u|e^{i(\delta_{B^0}-\gamma)}\bar{A}_f, \\ \mathcal{A}(B^0 \to DX^0_s) &\propto |A_c|\bar{A}_f + |A_u|e^{i(\delta_{B^0}+\gamma)}A_f, \end{aligned} \tag{1}$$

where $|A_{c,u}|$ are the magnitudes of the favoured and suppressed B-meson decay amplitudes, δ_{B^0} is the strong phase difference between them, and γ is the CP-violating weak phase. The quantities $A_{c,u}$ and δ_{B^0} depend on the position in the $B^0 \to DK^+\pi^-$ phase space. The amplitudes of the D^0 and $\overline{D}^0$ mesons decaying into the common final state f, $A_f \equiv \langle f | \mathcal{H} | D^0 \rangle$ and $\bar{A}_f \equiv \langle f | \mathcal{H} | \overline{D}^0 \rangle$, are functions of the $K^0_S\pi^+\pi^-$ final state, which can be completely specified by two squared invariant masses of pairs of the three final-state particles, chosen to be $m^2_+ \equiv m^2_{K^0_S\pi^+}$ and $m^2_- \equiv m^2_{K^0_S\pi^-}$. The other squared invariant mass is $m^2_0 \equiv m^2_{\pi^+\pi^-}$. Making the assumption of no CP violation in the D-meson decay, the amplitudes A_f and $\bar{A}_f$ are related by $\bar{A}_f(m^2_+, m^2_-) = A_f(m^2_-, m^2_+)$.

The amplitudes in Eq. 1 give rise to distributions of the form

$$\begin{aligned} \mathrm{d}\Gamma_{\overline{B}^0} &\propto |A_c|^2|A_f|^2 + |A_u|^2|\bar{A}_f|^2 + 2|A_c||A_u|\ \mathcal{R}e\left[A^\star_f \bar{A}_f\, e^{i(\delta_{B^0}-\gamma)}\right], \\ \mathrm{d}\Gamma_{B^0} &\propto |A_c|^2|\bar{A}_f|^2 + |A_u|^2|A_f|^2 + 2|A_c||A_u|\ \mathcal{R}e\left[A_f \bar{A}^\star_f\, e^{i(\delta_{B^0}+\gamma)}\right], \end{aligned} \tag{2}$$

which are functions of the position in the $B^0 \to DK^+\pi^-$ phase space. Integrating only over the region $\phi_{K^{*0}}$ of the $B^0 \to DK^+\pi^-$ phase space in which the K^{*0} resonance is dominant,

$$r^2_{B^0} \equiv \frac{\int_{\phi_{K^{*0}}} \mathrm{d}\phi \, |A_u|^2}{\int_{\phi_{K^{*0}}} \mathrm{d}\phi \, |A_c|^2}. \tag{3}$$

The functional

$$\mathcal{P}(A, z, \kappa) = \left|A\right|^2 + |z|^2 \left|\bar{A}\right|^2 + 2\kappa\mathcal{R}e\left[zA^\star\bar{A}\right], \tag{4}$$

describes the distribution within the phase space of the D-meson decay,

$$\begin{aligned} \mathcal{P}_{\overline{B}^0}(m^2_-, m^2_+) &\propto \mathcal{P}(A_f, z_-, \kappa), \\ \mathcal{P}_{B^0}(m^2_-, m^2_+) &\propto \mathcal{P}(\bar{A}_f, z_+, \kappa), \end{aligned} \tag{5}$$

where the coherence factor κ is a real constant ($0 \leq \kappa \leq 1$) [29] measured in Ref. [30], parameterising the fraction of the region $\phi_{K^{*0}}$ that is occupied by the K^{*0} resonance, and the complex parameters $z_\pm$ are

$$z_\pm = r_{B^0}\, e^{i(\delta_{B^0} \pm \gamma)}. \tag{6}$$

A direct determination of r_{B^0}, δ_{B^0} and γ can lead to bias, when r_{B^0} gets close to zero [17]. The Cartesian *CP* violation observables, $x_\pm = \mathcal{R}e(z_\pm)$ and $y_\pm = \mathcal{I}m(z_\pm)$, are therefore used instead.

This paper reports model-dependent Cartesian measurements of $z_\pm$ made using $B^0 \to D(K^0_S\pi^+\pi^-)K^{*0}$ decays selected from pp collision data, corresponding to an integrated luminosity of $3\,\mathrm{fb}^{-1}$, recorded by LHCb at centre-of-mass energies of 7 TeV in 2011 and 8 TeV in 2012. The measured values of $z_\pm$ place constraints on the CKM angle γ. Throughout the paper, inclusion of charge conjugate processes is implied, unless specified otherwise.

Section 2 describes the LHCb detector used to record the data, and the methods used to produce a realistic simulation of the data. Section 3 outlines the procedure used to select candidate $B^0 \to D(K^0_S\pi^+\pi^-)K^{*0}$ decays, and Sec. 4 describes the determination of the selection efficiency across the phase space of the D-meson decay. Section 5 details the fitting procedure used to determine the values of the Cartesian *CP* violation observables and Sec. 6 describes the systematic uncertainties on these results. Section 7 presents the interpretation of the measured Cartesian *CP* violation observables in terms of central values and confidence intervals for r_{B^0}, δ_{B^0} and γ, before Sec. 8 concludes with a summary of the results obtained.

2 The LHCb detector

The LHCb detector [31, 32] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing b or c quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the pp interaction region, a large-area silicon-strip detector located upstream of a dipole magnet of reversible polarity with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides

a measurement of the momentum p of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV. The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of $(15 + 29/p_{\rm T})\,\mu$m, where $p_{\rm T}$ is the component of the momentum transverse to the beam, in GeV. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

The trigger consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, in which all charged particles with $p_{\rm T} > 500\,(300)\,$MeV are reconstructed for 2011 (2012) data. The software trigger requires a two-, three- or four-track secondary vertex with a large sum of the transverse momentum, $p_{\rm T}$, of the tracks and a significant displacement from the primary pp interaction vertices. At least one track should have $p_{\rm T} > 1.7\,$GeV and $\chi^2_{\rm IP}$ with respect to any primary interaction greater than 16, where $\chi^2_{\rm IP}$ is defined as the difference in χ^2 of a given PV reconstructed with and without the considered track. A multivariate algorithm [33] is used for the identification of secondary vertices consistent with the decay of a b hadron. In the offline selection, trigger signals are associated with reconstructed particles. Selection requirements can therefore be made on the trigger selection itself and on whether the decision was due to the signal candidate, other particles produced in the pp collision, or a combination of both.

Decays of $K^0_{\rm S} \to \pi^+\pi^-$ are reconstructed in two different categories: the first involving $K^0_{\rm S}$ mesons that decay early enough for the daughter pions to be reconstructed in the vertex detector, and the second containing $K^0_{\rm S}$ that decay later such that track segments of the pions cannot be formed in the vertex detector. These categories are referred to as *long* and *downstream*, respectively. The long category has better mass, momentum and vertex resolution than the downstream category.

Large samples of simulated $B^0_{(s)} \to D\,\overline{(K)}^{*0}$ decays and various background decays are used in this study. In the simulation, pp collisions are generated using Pythia [34] with a specific LHCb configuration [35]. Decays of hadronic particles are described by EvtGen [36], in which final-state radiation is generated using Photos [37]. The interaction of the generated particles with the detector, and its response, are implemented using the Geant4 toolkit [38], as described in Ref. [39].

3 Candidate selection and background sources

In addition to the hardware and software trigger requirements, after a kinematic fit [40] to constrain the B^0 candidate to point towards the PV and the D candidate to have its nominal mass, the invariant mass of the $K^0_{\rm S}$ candidates must lie within $\pm 14.4\,$MeV ($\pm 19.9\,$MeV) of the known value [24] for long (downstream) categories. Likewise, after a kinematic fit to constrain the B^0 candidate to point towards the PV and the $K^0_{\rm S}$ candidate to have the $K^0_{\rm S}$ mass, the

reconstructed D-meson candidate must lie within $\pm 30\,\text{MeV}$ of the D^0 mass. To reconstruct the B^0 mass, a third kinematic fit of the whole decay chain is used, constraining the B^0 candidate to point towards the PV and the D and K^0_S to have their nominal masses. The χ^2 of this fit is used in the multivariate classifier described below. This fit improves the resolution of the $m^2_\pm$ invariant masses and ensures that the reconstructed D candidates are constrained to lie within the kinematic boundaries of the phase space. The K^{*0} candidate must have a mass within $\pm 50\,\text{MeV}$ of the world average value and $|\cos\theta^*| > 0.4$, where the decay angle θ^* is defined in the K^{*0} rest frame as the angle between the momentum of the kaon daughter of the K^{*0}, and the direction opposite to the B^0 momentum. The criteria placed on the K^{*0} candidate are identical to those used in the analysis of $B^0 \to DK^{*0}$ with two-body D decays [25].

A multivariate classifier is then used to improve the signal purity. A boosted decision tree (BDT) [41, 42] is trained on simulated signal events and background candidates lying in the high B^0 mass sideband $[5500, 6000]\,\text{MeV}$ in data. This mass range partially overlaps with the range of the invariant mass fit described below. To avoid a potential fit bias, the candidates are randomly split into two disjoint subsamples, A and B, and two independent BDTs (BDTA and BDTB) are trained with them. These classifiers are then applied to the complementary samples. The BDTs are based on 16 discriminating variables: the B^0 meson χ^2_{IP}, the sum of the χ^2_{IP} of the K^0_S daughter pions, the sum of the χ^2_{IP} of the final state particles except the K^0_S daughters, the B^0 and D decay vertex χ^2, the values of the flight distance significance with respect to the PV for the B^0, D and K^0_S mesons, the D (K^0_S) flight distance significance with respect to the B^0 (D) decay vertex, the transverse momenta of the B^0, D and K^{*0}, the cosine of the angle between the momentum direction of the B^0 and the displacement vector from the PV to the B^0 decay vertex, the decay angle of the K^{*0} and the χ^2 of the kinematic fit of the whole decay chain. Since some of the variables have different distributions for long or downstream candidates, the two event categories have separate BDTs, giving a total of four independent BDTs. The optimal cut value of each BDT classifier is chosen from pseudoexperiments to minimise the uncertainties on $z_\pm$.

Particle identification (PID) requirements are applied to the daughters of the K^{*0} to select kaon-pion pairs and reduce background coming from $B^0 \to D\rho^0$ decays. A specific veto is also applied to remove contributions from $B^\pm \to DK^\pm$ decays: $B^0 \to DK^{*0}$ candidates with a DK invariant mass lying in a $\pm 50\,\text{MeV}$ window around the $B^\pm$-meson mass are removed. To reject background from $D^0 \to \pi\pi\pi\pi$ decays, the decay vertex of each long K^0_S candidate is required to be significantly displaced from the D decay vertex along the beam direction.

The decay $B^0_s \to D\overline{K}^{*0}$ has a similar topology to $B^0 \to DK^{*0}$, but exhibits much less CP violation [30], since the decay $B^0_s \to D^0\overline{K}^{*0}$ is doubly-Cabbibo suppressed compared to $B^0_s \to \overline{D}^0\overline{K}^{*0}$. These decays are used as a control channel in the invariant mass fit. Background from partially reconstructed $B^0_{(s)} \to D^* {}^{(}\overline{K}{}^{)*0}$ decays, where D^* stands for either D^{*0} or $\overline{D}^{*0}$, are difficult to exclude since they have a topology very similar to the signal. The $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$ decays where the photon or the neutral pion is not reconstructed lead to $B^0_{(s)} \to D\,{}^{(}\overline{K}{}^{)*0}$ candidates with a lower invariant mass than the $B^0_{(s)}$ mass.

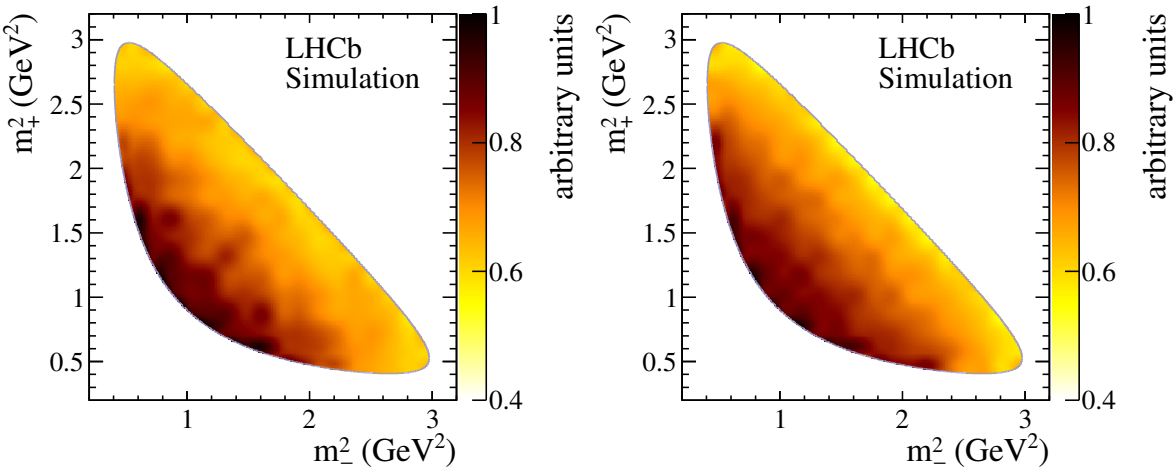


Figure 1: Variation of signal efficiency across the phase space for (left) long and (right) downstream candidates.

4 Efficiency across the phase space

The variation of the detection efficiency across the phase space is due to detector acceptance, trigger and selection criteria and PID effects. To evaluate this variation, a simulated sample generated uniformly over the $D \to K^0_S\pi^+\pi^-$ phase space is used, after applying corrections for known differences between data and simulation that arise for the hardware trigger and PID requirements.

The trigger corrections are determined separately for two independent event categories. In the first category, events have at least one energy deposit in the hadronic calorimeter, associated with the signal decay, which passes the hardware trigger. In the second category, events are triggered only by particles present in the rest of the event, excluding the signal decay. The probability that a given energy deposit in the hadronic calorimeter passes the hardware trigger is evaluated with calibration samples, which are produced for kaons and pions separately, and give the trigger efficiency as a function of the dipole magnet polarity, the transverse energy and the hit position in the calorimeter. The efficiency functions obtained for the two categories are combined according to their proportions in data.

The PID corrections are calculated with calibration samples of $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$ decays. After background subtraction, the PID efficiencies for kaon and pion candidates are obtained as functions of momentum and pseudorapidity. The product of the kaon and pion efficiencies, taking into account their correlation, gives the total PID efficiency.

The various efficiency functions are combined to make two separate global efficiency functions, one for long candidates and one for downstream candidates, which are used as inputs to the fit to obtain the Cartesian observables $z_\pm$. To smooth out statistical fluctuations, an interpolation with a two-dimensional cubic spline function is performed to give a continuous description of the efficiency $\varepsilon(m^2_+, m^2_-)$, as shown in Fig. 1.

5 Analysis strategy and fit results

To determine the *CP* observables $z_\pm$ defined in Eq. 6, an unbinned extended maximum likelihood fit is performed in three variables: the B^0 candidate reconstructed invariant mass m_{B^0} and the Dalitz variables m^2_+ and m^2_-. This fit is performed in two steps. First, the signal and background yields and some parameters of the invariant mass PDFs are determined with a fit to the reconstructed B^0 invariant mass distribution, described in Sec. 5.1. An amplitude fit over the phase space of the D-meson decay is then performed to measure $z_\pm$, using only candidates lying in a $\pm 25\,\text{MeV}$ window around the fitted B^0 mass, and taking the results of the invariant mass fit as inputs, as explained in Sec. 5.2. The `cfit` [43] library has been used to perform these fits. Candidate events are divided into four subsamples, according to K^0_S type (long or downstream), and whether the candidate is identified as a B^0 or $\overline{B}{}^0$-meson decay. In the B-candidate invariant mass fit, the B^0 and $\overline{B}{}^0$ samples are combined, since identical distributions are expected for this variable, whilst in the *CP* violation observables fit (*CP* fit) they are kept separate.

5.1 Invariant mass fit of $B^0 \to DK^{*0}$ candidates

An unbinned extended maximum likelihood fit to the reconstructed invariant mass distributions of the B^0 candidates in the range $[4900, 5800]\,\text{MeV}$ determines the signal and background yields. The long and downstream subsamples are fitted simultaneously. The total PDF includes several components: the $B^0 \to DK^{*0}$ signal PDF, background PDFs for $B^0_s \to D\overline{K}{}^{*0}$ decays, combinatorial background, partially reconstructed $B^0_{(s)} \to D^* \overline{(K)}{}^{*0}$ decays and misidentified $B^0 \to D\rho^0$ decays, as illustrated in Fig. 2.

The fit model is similar to that used in the analysis of $B^0 \to DK^{*0}$ decays with D-meson decays to two-body final states [25]. The $B^0 \to DK^{*0}$ and $B^0_s \to D\overline{K}{}^{*0}$ components are each described as the sum of two Crystal Ball functions [44] sharing the same central value, with the relative yields of the two functions and the tail parameters fixed from simulation. The separation between the central values of the $B^0 \to DK^{*0}$ and $B^0_s \to D\overline{K}{}^{*0}$ PDFs is fixed to the known B^0-B^0_s mass difference. The ratio of the $B^0 \to DK^{*0}$ and $B^0_s \to D\overline{K}{}^{*0}$ yields is constrained to be the same in both the long and downstream subsamples. The combinatorial background is described with an exponential PDF. Partially reconstructed $B^0_{(s)} \to D^* \overline{(K)}{}^{*0}$ decays are described with non-parametric functions obtained by applying kernel density estimation [45] to distributions of simulated events. These distributions depend on the helicity state of the D^{*0} meson. Due to parity conservation in $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$ decays, two of the three helicity amplitudes have the same invariant mass distribution. The $B^0_s \to D^*\overline{K}{}^{*0}$ PDF is therefore a linear combination of two non-parametric functions, with the fraction of the longitudinal polarisation in the $B^0_s \to D^*\overline{K}{}^{*0}$ decays unknown and accounted for with a free parameter in the fit. Each of the two functions describing the different helicity states is a weighted sum of non-parametric functions obtained from simulated $B^0_s \to D^*(D^0\gamma)\overline{K}{}^{*0}$ and $B^0_s \to D^*(D^0\pi^0)\overline{K}{}^{*0}$ decays, taking into account the known $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$ branching fractions [46] and the appropriate efficiencies. The PDF for $B^0 \to D^*K^{*0}$ decays is obtained from that for $B^0_s \to D^*\overline{K}{}^{*0}$ decays, by applying a shift corresponding to the known B^0-B^0_s mass difference. In the nominal fit, the

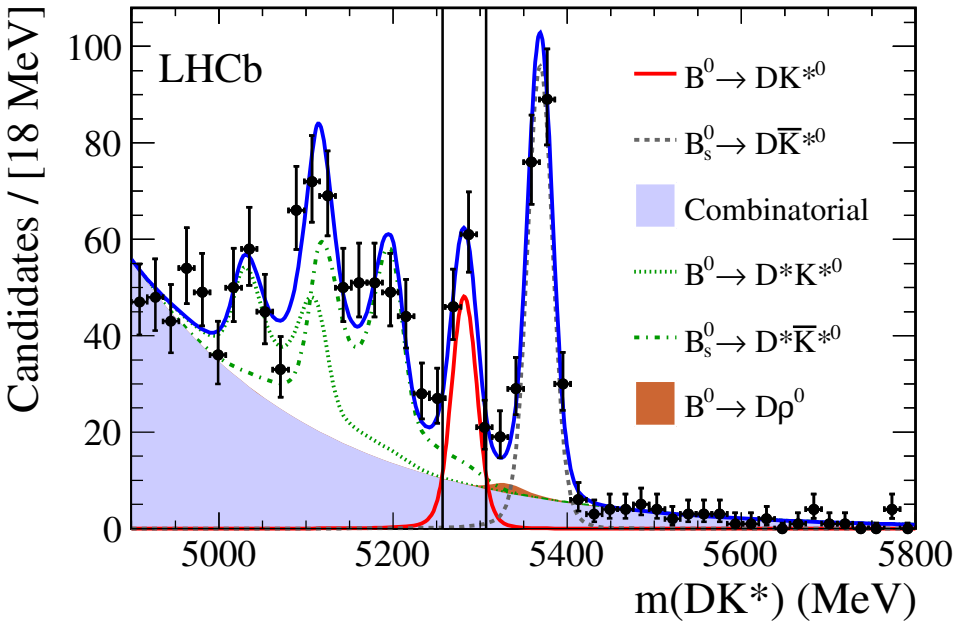


Figure 2: Invariant mass distribution for $B^0 \to DK^{*0}$ long and downstream candidates. The fit result, including signal and background components, is superimposed (solid blue). The points are data, and the different fit components are given in the legend. The two vertical lines represent the signal region in which the *CP* fit is performed.

polarisation fraction is assumed to be the same for $B^0 \to D^* K^{*0}$ and $B_s^0 \to D^* \overline{K}^{*0}$ decays. The effect of this assumption is taken into account in the systematic uncertainties. The $B^0 \to D\rho^0$ component is also described with a non-parametric function obtained from the simulation, using a data-driven calibration to describe the pion-kaon misidentification efficiency. This component has a very low yield and, to improve the stability of the fit, a Gaussian constraint is applied, requiring the ratio of yields of $B^0 \to D\rho^0$ and $B_s^0 \to D\overline{K}^{*0}$ to be consistent with its expected value.

The fitted distribution is shown in Fig 2. The resulting signal and background yields in a ± 25 MeV range around the B^0 mass are given in Table 1. This range corresponds to the signal region over which the *CP* fit is performed.

5.2 *CP* fit

A simultaneous unbinned maximum likelihood fit to the four subsamples is performed to determine the *CP* violation observables $z_\pm$. The value of the coherence factor is fixed to the central value of $\kappa = 0.958^{+0.005+0.002}_{-0.010-0.045}$, as measured in the recent LHCb amplitude analysis of

Table 1: Signal and background yields in the signal region, ±25 MeV around the B^0 mass, obtained from the invariant mass fit. Total yields, as well as separate yields for long and downstream candidates, are given.

Component	Yield		
	Long	Downstream	Total
$B^0 \to DK^{*0}$	29 ± 5	60 ± 8	89 ± 11
$B^0_s \to D\overline{K}^{*0}$	0.59 ± 0.12	1.21 ± 0.23	1.8 ± 0.3
Combinatorial	9.6 ± 1.0	16.1 ± 1.4	25.7 ± 1.7
$B^0 \to D^*K^{*0}$	0.06 ± 0.02	0.06 ± 0.02	0.12 ± 0.03
$B^0_s \to D^*\overline{K}^{*0}$	4.1 ± 0.8	7.9 ± 1.3	11.9 ± 1.7
$B^0 \to D\rho^0$	0.20 ± 0.05	0.37 ± 0.09	0.57 ± 0.11
Total background	14.5 ± 1.3	25.6 ± 1.8	40.1 ± 2.4

$B^0 \to DK^+\pi^-$ decays [30]. The negative logarithm of the likelihood,

$$\begin{aligned} -\ln\mathcal{L} = &-\sum_{B^0\text{cand.}} \ln\left(\sum_c N_c f_c^{\text{mass}}(m_B;\vec{q}_c^{\,\text{mass}}) f_c^{B^0\ \text{model}}(m_+^2, m_-^2; z_\pm, \kappa, \vec{q}_c^{\,\text{model}})\right) \\ &-\sum_{\overline{B}^0\text{cand.}} \ln\left(\sum_c N_c f_c^{\text{mass}}(m_B;\vec{q}_c^{\,\text{mass}}) f_c^{\overline{B}^0\ \text{model}}(m_+^2, m_-^2; z_\pm, \kappa, \vec{q}_c^{\,\text{model}})\right) \\ &+\sum_c N_c, \end{aligned} \tag{7}$$

is minimised, where c indexes the different signal and background components, N_c is the yield for each category, f_c^{mass} is the invariant mass PDF determined in the previous section, $\vec{q}_c^{\,\text{mass}}$ are the mass PDF parameters, $f_c^{B\ \text{model}}$ is the amplitude PDF and $\vec{q}_c^{\,\text{model}}$ are its parameters other than $z_\pm$ and κ, which have been included explicitly.

The non-uniformity of the selection efficiency over the $D \to K^0_S\pi^+\pi^-$ phase space is accounted for by including the function $\varepsilon(m_+^2, m_-^2)$, introduced in Sec. 4, within the $f_c^{B\ \text{model}}$ PDF:

$$f_c^{B\ \text{model}}(m_+^2, m_-^2; z_\pm, \kappa, \vec{q}_c^{\,\text{model}}) = \mathcal{F}_c(m_+^2, m_-^2; z_\pm, \kappa, \vec{q}_c^{\,\text{model}})\ \varepsilon(m_+^2, m_-^2), \tag{8}$$

where $\mathcal{F}_c$ is the PDF of the amplitude model.

The model describing the amplitude of the $D \to K^0_S\pi^+\pi^-$ decay over the phase space, $A_f\left(m_+^2, m_-^2\right)$, is identical to that used previously by the BaBar [19,47] and LHCb [23] collaborations. An isobar model is used to describe P-wave (including $\rho(770)^0$, $\omega(782)$, Cabibbo-allowed and doubly Cabibbo-suppressed $K^*(892)^\pm$ and $K^*(1680)^-$) and D-wave (including $f_2(1270)$ and $K_2^*(1430)^\pm$) contributions. The $K\pi$ S-wave contribution ($K_0^*(1430)^\pm$) is described using a generalised LASS amplitude [48], whilst the $\pi\pi$ S-wave contribution is treated using a P-vector approach within the K-matrix formalism. All parameters of the model are fixed in the fit to the

values determined in Ref. [47].[1]

All components included in the fit of the B-meson mass spectrum are included in the fit for the CP violation observables, with the exception of the $B^0 \to D^* K^{*0}$ background, because its yield within the signal region is negligible (Table 1). CP violation is neglected for $B^0_s \to D\overline{K}^{*0}$ and $B^0_s \to D^*\overline{K}^{*0}$ decays, since their Cabbibo-suppressed contributions are negligible. The relevant PDFs are therefore $\mathcal{F}_{B^0_s \to D^{(*)}\overline{K}^{*0}} = \mathcal{P}(\bar{A}_f, 0, 0)$ and $\mathcal{F}_{\overline{B}^0_s \to D^{(*)}K^{*0}} = \mathcal{P}(A_f, 0, 0)$, where $\mathcal{P}$ is defined in Eq. 4. For background arising from misidentified $B^0 \to D\rho^0$ events, the B flavour state cannot be determined, resulting in an incoherent sum of D^0 and $\overline{D}^0$ contributions: $\mathcal{F}_{B^0 \to D\rho^0} = (|A_f|^2 + |\bar{A}_f|^2)/2$.

The combinatorial background is composed of two contributions: one from non-D candidates, and the other from real D mesons combined with random tracks. Combinatorial D candidates arise from random combinations of four charged tracks, incorrectly reconstructed as a $D \to K^0_S\pi^+\pi^-$ decay, and this contribution is assumed to be distributed uniformly over phase space, $\mathcal{F}_{\text{Comb, non}-D} = 1$, consistent with what is seen in the data. Background from real D candidates arises when the $K^*(892)^0$ candidate is reconstructed from random tracks. Consequently, the B-meson flavour is unknown, resulting in an incoherent sum, $\mathcal{F}_{\text{Comb, real } D} = (|A_f|^2 + |\bar{A}_f|^2)/2$. The relative proportions of non-D and real D meson backgrounds ($\mathcal{O}(30\%)$) are fixed using the results of a fit to the reconstructed invariant mass of the D candidates in the signal B mass region. Figures 3 and 4 show the Dalitz plot and its projections, with the fit result superimposed, for B^0 and $\overline{B}^0$ candidates, respectively. A blinding procedure was used to obscure the values of the CP parameters until all aspects of the analysis were finalised. The measured values are

$$\begin{aligned} x_- &= -0.15 \pm 0.14, \\ y_- &= 0.25 \pm 0.15, \\ x_+ &= 0.05 \pm 0.24, \\ y_+ &= -0.65\ ^{+0.24}_{-0.23}, \end{aligned}$$

where the uncertainty is statistical only. The correlation matrix is

$$\begin{array}{c} \begin{array}{cccc} x_- & y_- & x_+ & y_+ \end{array} \\ \begin{pmatrix} 1 & 0.14 & 0 & 0 \\ 0.14 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.14 \\ 0 & 0 & 0.14 & 1 \end{pmatrix} \end{array},$$

and the corresponding likelihood contours for $z_\pm$ are shown in Fig. 5.

[1] As previously noted in Ref. [23], the model implemented by BaBar [47] differs from the formulation described therein. One of the two Blatt-Weisskopf coefficients was set to unity, and the imaginary part of the denominator of the Gounaris-Sakurai propagator used the mass of the resonant pair, instead of the mass associated with the resonance. The model used herein replicates these features without modification. It has been verified that changing the model to use an additional centrifugal barrier term and a modified Gounaris-Sakurai propagator has a negligible effect on the measurements.

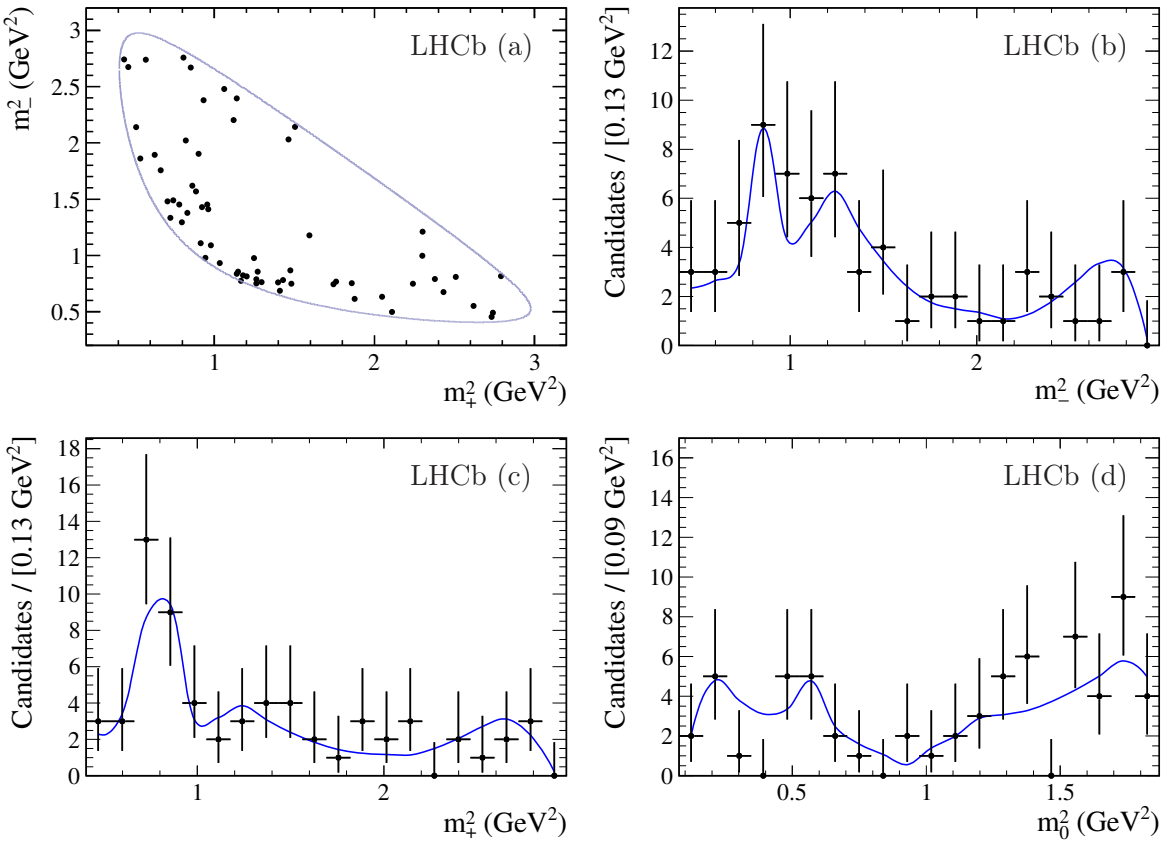


Figure 3: Selected $B^0 \to DK^{*0}$ candidates, shown as (a) the Dalitz plot, and its projections on (b) m^2_-, (c) m^2_+ and (d) m^2_0. The line superimposed on the projections corresponds to the fit result and the points are data.

6 Systematic uncertainties

Several sources of systematic uncertainty on the evaluation of $z_\pm$ are considered, and are summarised in Table 2. Unless otherwise stated, for each source considered, the *CP* fit is repeated and the differences in the $z_\pm$ values compared to the nominal results are taken as the systematic uncertainties.

The uncertainty on the description of the efficiency variation across the D-meson decay phase space arises from several sources. Statistical uncertainties arise due to the limited sizes of the simulated samples used to determine the nominal efficiency function and of the calibration samples used to obtain the data-driven corrections to the PID and hardware trigger efficiencies. Large numbers of alternative efficiency functions are created by smearing these quantities according to their uncertainties. For each fitted *CP* parameter, the residual for a given alternative efficiency function is defined as the difference between its value obtained using this function, and that obtained in the nominal fit. The width of the obtained distribution of residuals is taken as the corresponding systematic uncertainty. Additionally, since the nominal fit is perfomed using an efficiency function obtained using only BDTA, the fit is repeated using an alternative efficiency function obtained using BDTB, and an uncertainty extracted. The fit is also performed with

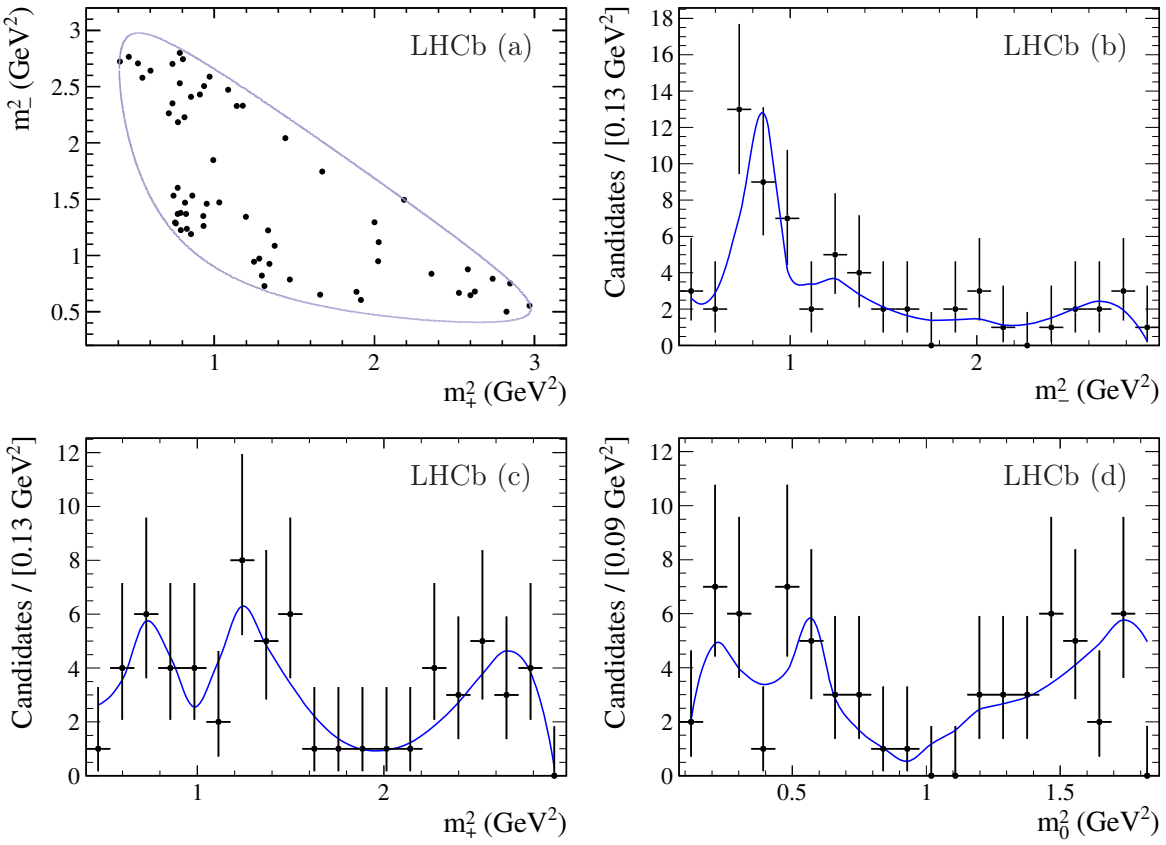


Figure 4: Selected $\overline{B}^0 \to D\overline{K}^{*0}$ candidates, shown as (a) the Dalitz plot, and its projections on (b) m^2_-, (c) m^2_+ and (d) m^2_0. The line superimposed on the projections corresponds to the fit result and the points are data.

alternative efficiency functions obtained by varying the fraction of candidates triggered by at least one product of the signal decay chain. Finally, for a few variables used in the BDT, a small difference is observed between the simulation and the background-subtracted data sample. To account for this difference, the simulated events are reweighted to match the data, and the fit is repeated with the resulting efficiency function.

The B-meson invariant mass fit result is used to fix the fractions of signal and background and the parameters of the B^0 mass PDF shapes in the *CP* fit. A large number of pseudoexperiments is generated, in which the free parameters of the invariant mass fit are varied within their uncertainties, taking into account their correlations. The *CP* fit is repeated for each variation. For each *CP* parameter, the width from a Gaussian fit to the resulting residual distribution is taken as the associated systematic uncertainty. This is the dominant contribution to the invariant mass fit systematic uncertainty quoted in Table 2. Other uncertainties due to assumptions in the invariant mass fit are evaluated by allowing the $B^0 \to DK^{*0}/B^0_s \to D\overline{K}^{*0}$ yield ratio to be different for long and downstream categories, by varying the $B^0 \to D\rho^0/B^0_s \to D\overline{K}^{*0}$ yield ratio, by varying the Crystal Ball PDF parameters within their uncertainties and by testing alternatives to the Crystal Ball PDFs. The proportions of $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$ in the $B^0_{(s)} \to D^* \overline{(K)}^{*0}$ background description are also varied, and the effect of neglecting the

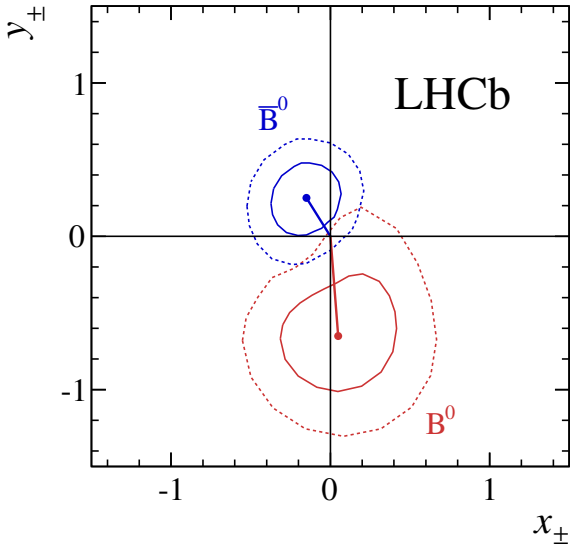


Figure 5: Likelihood contours at 68.3% and 95.5% confidence level for (x_+, y_+) (red) and (x_-, y_-) (blue), obtained from the *CP* fit.

Table 2: Summary of the systematic uncertainties on $z_\pm$, in units of (10^{-3}). The total experimental and total model-related uncertainties are also given as percentages of the statistical uncertainties.

Source of uncertainty	δx_-	δy_-	δx_+	δy_+
Efficiency	5.4	1.1	11	1.8
Invariant mass fit	12	21	15	48
Migration over the phase space	5.3	1.8	6.2	3.0
Misreconstructed signal	7.7	6.6	10	7.1
Background description				
Non-D background	20	15	28	47
Real D background	0.1	0.4	0.2	1.0
CP violation in $B^0_s \to D^* \overline{K}^{*0}$	1.5	0.8	4.0	1.6
$B^+ \to \overline{D}^0 \pi^+ \pi^+ \pi^-$ contribution	0.6	1.4	0.8	2.3
$\Lambda^0_b \to D^0 p \pi^-$ contribution	0.1	0.7	0.5	1.6
K^* coherence factor (κ)	4.8	2.4	8.5	2.6
CP fit bias	5	49	11	40
Total experimental	26 (19%)	56 (37%)	39 (16%)	78 (33%)
Total model-related (see Table 3)	8 (5%)	7 (5%)	10 (4%)	5 (2%)

$B^0 \to D^* K^{*0}$ component in the *CP* fit is evaluated.

The systematic uncertainty due to the finite resolution in $m^2_\pm$ is evaluated with a large number of pseudoexperiments. One nominal pseudodata sample is generated, with $z_\pm$ fixed to the values obtained from data. A large number of alternative samples are generated from the nominal one

by smearing the $m^2_\pm$ coordinates of each event according to the resolution found in simulation and taking correlations into account. For each *CP* parameter, the width of the residual distribution is taken as the systematic uncertainty.

The misreconstruction of $B^0 \to DK^{*0}$ signal events is also studied. This can occur *e.g.* when the wrong final state pions of a real signal event are combined in the reconstruction of the D-meson candidate, leading to migration of this event within the D-decay phase space. The uncertainty corresponding to this effect is evaluated using pseudoexperiments. The effect of signal misreconstruction due to K^{*0}–$\overline{K}^{*0}$ misidentification, corresponding to a $(K^\pm\pi^\mp) \to (\pi^\pm K^\mp)$ misidentification, is found to be negligible thanks to the PID requirements placed on the K^{*0} daughters.

The uncertainty arising from the background description is evaluated for several sources. The *CP* fit is repeated with the fractions of the two categories of combinatorial background (non-D and real D candidates) varied within their uncertainties from the fit to the D invariant mass distribution. Additionally, since in the nominal fit the non-D candidates are assumed to be uniformly distributed over the phase space of the $D \to K^0_S\pi^+\pi^-$ decay, the fit is repeated changing this contribution to the sum of a uniform distribution and a $K^*(892)^\pm$ resonance. The relative proportions of the two components are fixed based on the $m^2_\pm$ distributions found in data. The fit is also repeated with the D-meson decay model for the non-D component set to the distribution of data in the D mass sidebands. The uncertainty arising from the poorly-known fraction of non-D and real D background is the dominant systematic uncertainty for the $x_\pm$ parameters.

The description of the real D combinatorial background assumes that the probabilities of a D^0 or a $\overline{D}^0$ being present in an event are equal. The *CP* violation observables fit is repeated with the decay model for this background changed to include a D^0–$\overline{D}^0$ production asymmetry, whose value is set to the measured $D^\pm$ asymmetry $(-1.0 \pm 0.3) \times 10^{-2}$ [49].

CP violation is neglected in the $B^0_s \to D^*\overline{K}^{*0}$ decay nominal description. The *CP* fit is repeated with the inclusion of a small component describing the suppressed decay amplitude of $B^0_s \to D^{*0}\overline{K}^{*0}$, with *CP* violation parameters for this component fixed to $\gamma = 73.2°$, $r_{B^0_s} = 0.02$ and $\delta_{B^0_s} = \{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}$. The model used to describe $B^0_s \to D^*\overline{K}^{*0}$ decays consists of an incoherent sum of $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$ contributions. Between the $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$ decays, there is an effective strong phase shift of π that is taken into account [50].

The systematic uncertainties arising from the inclusion of background from misreconstructed $B^+ \to \overline{D}^0\pi^+\pi^+\pi^-$ and $\Lambda^0_b \to D^0p\pi^-$ decays are evaluated, by adding these components into the fit model. The *CP* fit is also repeated with the $K^*(892)^0$ coherence factor κ varied within its uncertainty [30].

The *CP* fit is verified using one thousand data-sized pseudoexperiments. In each experiment, the signal and background yields, as well as the distributions used in the generation, are fixed to those found in data. The fitted values of $z_\pm$ show biases smaller than the statistical uncertainties, and are included as systematic uncertainties. These biases are due to the current limited statistics and are found to reduce in pseudoexperiments generated with a larger sample size.

To evaluate the systematic uncertainty due to the choice of amplitude model for $D \to K^0_S\pi^+\pi^-$, one million $B^0 \to DK^{*0}$ and one million $B^0_s \to D\overline{K}^{*0}$ decays are simulated according to the nominal decay model, with the Cartesian observables fixed to the nominal fit result. These simulated decays are fitted with alternative models, each of which includes a single modification with respect to the nominal model, as described in the next paragraph. Each of these alternative models is first used to fit the simulated $B^0_s \to D\overline{K}^{*0}$ decays to determine values for the resonance coefficients of the model. Those coefficients are then fixed in a second fit, to the simulated $B^0 \to DK^{*0}$ decays, to obtain $z_\pm$. The systematic uncertainties are taken to be the signed differences in the values of $z_\pm$ from the nominal results.

The following changes, labelled (a)-(u), are applied in the alternative models, leading to the uncertainties shown in Table 3:

- $\pi\pi$ S-wave: The F-vector model is changed to use two other solutions of the K-matrix (from a total of three) determined from fits to scattering data [51] (a), (b). The slowly varying part of the nonresonant term of the P-vector is removed (c).
- $K\pi$ S-wave: The generalised LASS parametrisation used to describe the $K^*_0(1430)^\pm$ resonance, is replaced by a relativistic Breit–Wigner propagator with parameters taken from Ref. [52] (d).
- $\pi\pi$ P-wave: The Gounaris–Sakurai propagator is replaced by a relativistic Breit–Wigner propagator [19, 47] (e).
- $K\pi$ P-wave: The mass and width of the $K^*(1680)^-$ resonance are varied by their uncertainties from Ref. [48] (f)−(i).
- $\pi\pi$ D-wave: The mass and width of the $f_2(1270)$ resonance are varied by their uncertainties from Ref. [24] (j)−(m).
- $K\pi$ D-wave: The mass and width of the $K^*_2(1430)^\pm$ resonance are varied by their uncertainties from Ref. [53] (n)−(q).
- The radius of the Blatt–Weisskopf centrifugal barrier factors, $r_{\rm BW}$, is changed from 1.5 GeV^{-1} to 0.0 GeV^{-1} (r) and 3.0 GeV^{-1} (s).
- Two further resonances, $K^*(1410)^0$ and $\rho(1450)$, parametrised with relativistic Breit–Wigner propagators, are included in the model [19, 47] (t).
- The Zemach formalism used for the angular distribution of the decay products is replaced by the helicity formalism [19, 47] (u).

It results in total systematic uncertainties arising from the choice of amplitude model of

$$\begin{aligned}\delta x_- &= 8 \times 10^{-3},\\ \delta y_- &= 7 \times 10^{-3},\\ \delta x_+ &= 10 \times 10^{-3},\\ \delta y_+ &= 5 \times 10^{-3}.\end{aligned}$$

The different systematic uncertainties are combined, assuming that they are independent to obtain the total experimental uncertainties. Depending on the $(x_\pm, y_\pm)$ parameters, the leading systematic uncertainties arise from the invariant mass fit, the description of the non-

Table 3: Model related systematic uncertainties for each alternative model, in units of (10^{-3}). The relative signs indicate full correlation or anti-correlation.

	Description		δx_-	δy_-	δx_+	δy_+
(a)	K-matrix 1st solution		-2	0.9	2	1
(b)	K-matrix 2nd solution		0.3	0.3	0.0	-0.5
(c)	Remove slowly varying part in P-vector		-0.7	0.2	0.5	0.6
(d)	Generalised LASS $\rightarrow$ relativistic Breit–Wigner		2	3	-1	3
(e)	Gounaris-Sakurai $\rightarrow$ relativistic Breit–Wigner		0.7	0.0	-0.1	0.8
(f)	$K^*(1680)$	$m + \delta m$	-0.0	0.6	0.1	0.5
(g)		$m - \delta m$	-0.2	-0.5	0.2	-0.9
(h)		$\Gamma + \delta\Gamma$	-0.2	0.2	0.0	-0.2
(i)		$\Gamma - \delta\Gamma$	0.2	-0.1	0.5	-0.2
(j)	$f_2(1270)$	$m + \delta m$	-0.1	0.0	0.3	-0.2
(k)		$m - \delta m$	-0.0	0.1	0.2	-0.2
(l)		$\Gamma + \delta\Gamma$	-0.0	0.0	0.2	-0.2
(m)		$\Gamma - \delta\Gamma$	-0.1	0.0	0.2	-0.2
(n)	$K_2^*(1430)$	$m + \delta m$	0.3	0.2	0.2	-0.2
(o)		$m - \delta m$	-0.4	-0.2	0.3	-0.1
(p)		$\Gamma + \delta\Gamma$	-0.2	0.2	0.1	-0.2
(q)		$\Gamma - \delta\Gamma$	0.1	-0.1	0.3	-0.2
(r)	$r_{\rm BW} = 0.0\,{\rm GeV}^{-1}$		-2	0.7	-1	-0.3
(s)	$r_{\rm BW} = 3.0\,{\rm GeV}^{-1}$		4	-2	4	2
(t)	Add $K^*(1410)$ and $\rho(1450)$		-0.2	-0.2	0.3	-0.3
(u)	Helicity formalism		-6	6	-8	2
	Total model related		8	7	10	5

D background and the fit biases. A larger data sample is expected to reduce all three of these uncertainties. Whilst not intrinsically statistical in nature, the systematic uncertainty due to the description of the non-D background is presently evaluated using a conservative approach due to lack of statistics. The total systematic uncertainties, including the model-related uncertainties, are significantly smaller than the statistical uncertainties.

7 Determination of the parameters γ, r_{B^0} and δ_{B^0}

To determine the physics parameters r_{B^0}, δ_{B^0} and γ from the fitted Cartesian observables $z_\pm$, the relations

$$\begin{aligned} x_\pm &= r_{B^0}\cos(\delta_{B^0} \pm \gamma), \\ y_\pm &= r_{B^0}\sin(\delta_{B^0} \pm \gamma), \end{aligned} \tag{9}$$

must be inverted. This is done using the `GammaCombo` package, originally developed for the frequentist combination of γ measurements by the LHCb collaboration [7,54]. A global likelihood function is built, which gives the probability of observing a set of $z_\pm$ values given the true values $(r_{B^0}, \delta_{B^0}, \gamma)$,

$$\mathcal{L}(x_-, y_-, x_+, y_+ | r_{B^0}, \delta_{B^0}, \gamma). \tag{10}$$

All statistical and systematic uncertainties on $z_\pm$ are accounted for, as well as the statistical correlation between $z_\pm$. Since the precision of the measurement is statistics dominated, correlations between the systematic uncertainties are ignored. Central values for $(r_{B^0}, \delta_{B^0}, \gamma)$ are obtained by performing a scan of these parameters, to find the values that maximise $\mathcal{L}(x_-^{\rm obs}, y_-^{\rm obs}, x_+^{\rm obs}, y_+^{\rm obs} | r_{B^0}, \delta_{B^0}, \gamma)$, where $z_\pm^{\rm obs}$ are the measured values of the Cartesian observables. Associated confidence intervals may be obtained either from a simple profile-likelihood method, or using the Feldman-Cousins approach [55] combined with a "plugin" method [56]. Confidence level curves for $(r_{B^0}, \delta_{B^0}, \gamma)$ obtained using the latter method are shown in Figs. 6, 7 and 8. The measured values of $z_\pm$ are found to correspond to

$$\begin{aligned} \gamma &= \left(80^{+21}_{-22}\right)^\circ, \\ r_{B^0} &= 0.39 \pm 0.13, \\ \delta_{B^0} &= \left(197^{+24}_{-20}\right)^\circ. \end{aligned}$$

Intrinsic to the method used in this analysis [12], there is a two-fold ambiguity in the solution; the Standard Model solution $(0 < \gamma < 180)^\circ$ is chosen. Two-dimensional confidence level curves obtained using the profile-likelihood method are shown in Figs. 9 and 10.

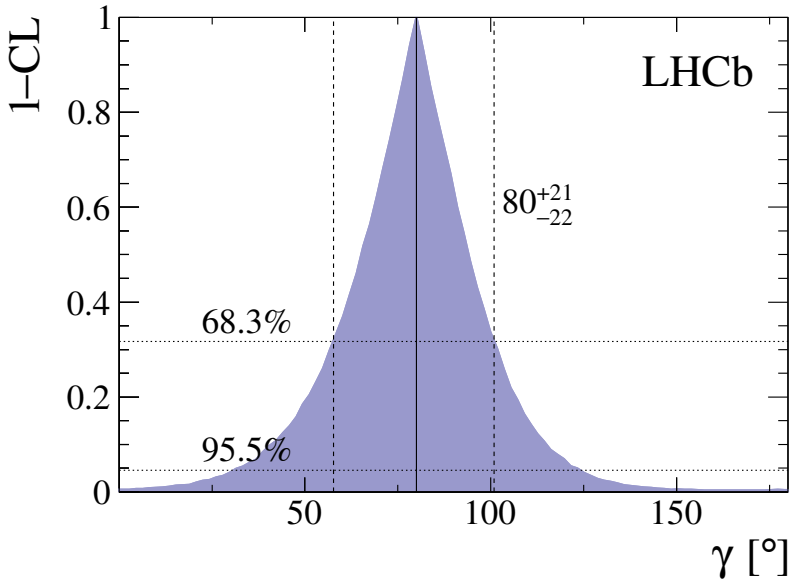


Figure 6: Confidence level curve on γ, obtained using the "plugin" method [56].

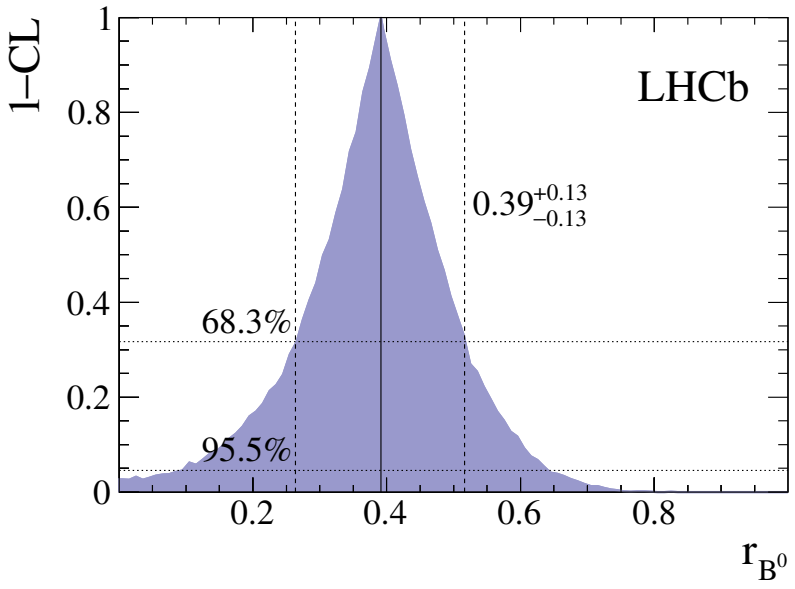


Figure 7: Confidence level curve on r_{B^0}, obtained using the "plugin" method [56].

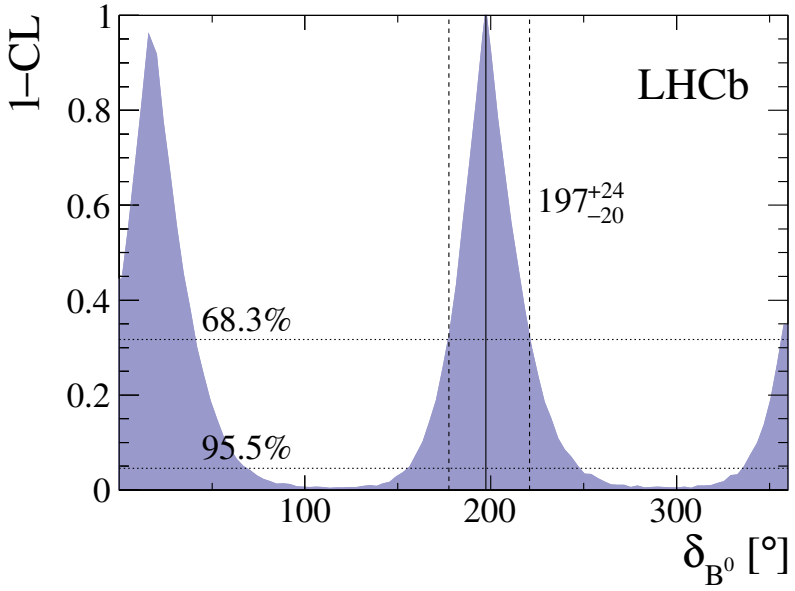


Figure 8: Confidence level curve on δ_{B^0}, obtained using the "plugin" method [56]. Only the δ_{B^0} solution corresponding to $0 < \gamma < 180°$ is highlighted; the other maximum is due to the $(\delta_{B^0}, \gamma) \to (\delta_{B^0} + \pi, \gamma + \pi)$ ambiguity.

8 Conclusion

An amplitude analysis of $B^0 \to DK^{*0}$ decays, employing a model description of the $D \to K^0_S\pi^+\pi^-$ decay, has been performed using data corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$, recorded by LHCb at a centre-of-mass energy of 7 TeV in 2011 and 8 TeV in 2012. The measured values of the CP violation observables $x_\pm = r_{B^0}\cos(\delta_{B^0} \pm \gamma)$ and $y_\pm = r_{B^0}\sin(\delta_{B^0} \pm \gamma)$

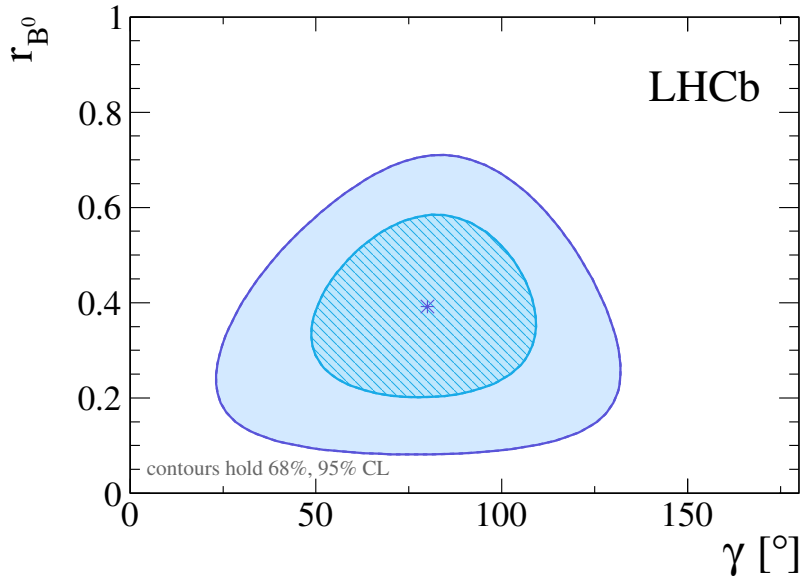


Figure 9: Two-dimensional confidence level curves in the (γ, r_{B^0}) plane, obtained using the profile-likelihood method.

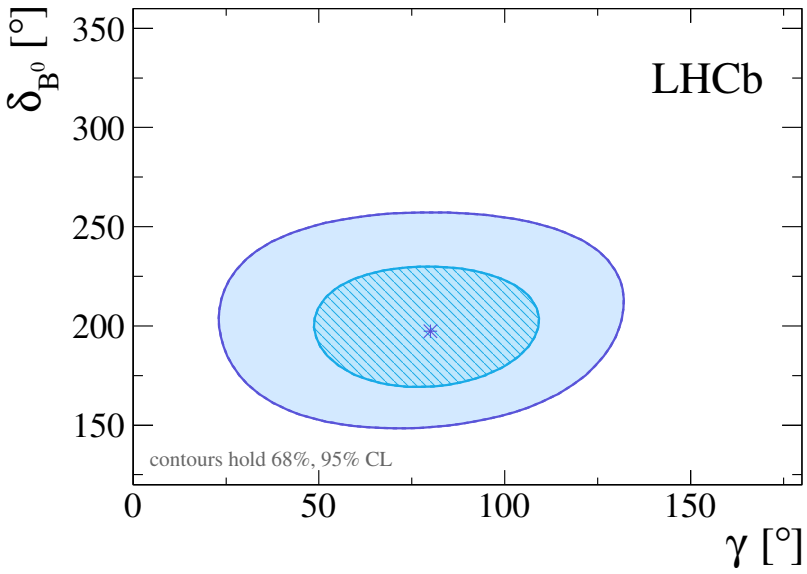


Figure 10: Two-dimensional confidence level curves in the (γ, δ_{B^0}) plane, obtained using the profile-likelihood method.

are

$$
\begin{aligned}
x_- &= -0.15 \pm 0.14 \pm 0.03 \pm 0.01, \\
y_- &= 0.25 \pm 0.15 \pm 0.06 \pm 0.01, \\
x_+ &= 0.05 \pm 0.24 \pm 0.04 \pm 0.01, \\
y_+ &= -0.65\,^{+0.24}_{-0.23} \pm 0.08 \pm 0.01,
\end{aligned}
$$

where the first uncertainties are statistical, the second are systematic and the third are due to the choice of amplitude model used to describe the $D \to K^0_S\pi^+\pi^-$ decay. These are the most precise measurements of these observables related to the neutral channel $B^0 \to DK^{*0}$. They place constraints on the magnitude of the ratio of the interfering B-meson decay amplitudes, the strong phase difference between them and the CKM angle γ, giving the values

$$\begin{aligned} \gamma &= \left(80^{+21}_{-22}\right)^\circ, \\ r_{B^0} &= 0.39 \pm 0.13, \\ \delta_{B^0} &= \left(197^{+24}_{-20}\right)^\circ. \end{aligned}$$

Here, r_{B^0} and δ_{B^0} are defined for a $K\pi$ mass region of $\pm 50\,\text{MeV}$ around the $K^*(892)^0$ mass and for an absolute value of the cosine of the K^{*0} decay angle greater than 0.4. These results are consistent with those reported in Ref. [28], where a different analysis method is used. The two results are based on the same data set and cannot be combined.

Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FANO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (USA). We are indebted to the communities behind the multiple open source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).

References

[1] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[2] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[3] J. Brod and J. Zupan, *The ultimate theoretical error on γ from $B \to DK$ decays*, JHEP **14** (2014) 051, arXiv:1308.5663.

[4] BaBar collaboration, J. P. Lees *et al.*, *Observation of direct CP violation in the measurement of the Cabibbo-Kobayashi-Maskawa angle γ with $B^\pm \to D^{(*)}K^{(*)\pm}$ decays*, Phys. Rev. **D87** (2013) 052015, arXiv:1301.1029.

[5] Belle collaboration, K. Trabelsi, *Study of direct CP in charmed B decays and measurement of the CKM angle γ at Belle*, in *7th Workshop on the CKM Unitarity Triangle (CKM 2012) Cincinnati, Ohio, USA, September 28-October 2, 2012*, 2013. arXiv:1301.2033.

[6] Belle, BaBar collaborations, A. J. Bevan *et al.*, *The physics of the B factories*, Eur. Phys. J. **C74** (2014) 3026, arXiv:1406.6311.

[7] LHCb collaboration, *Improved constraints on γ: CKM2014 update*, LHCb-CONF-2014-004.

[8] CKMfitter group, 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.

[9] CKMfitter group, J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91** (2015) 073007, arXiv:1501.05013.

[10] UTfit collaboration, M. Bona *et al.*, *The 2004 UTfit collaboration report on the status of the unitarity triangle in the standard model*, JHEP **07** (2005) 028, arXiv:hep-ph/0501199, updated results and plots available at: http://www.utfit.org/UTfit/.

[11] A. Bondar in *Proceedings of BINP special analysis meeting on Dalitz, unpublished*, 2002.

[12] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, *Determining γ using $B^\pm \to DK^\pm$ with multibody D decays*, Phys. Rev. **D68** (2003) 054018.

[13] Belle collaboration, H. Aihara *et al.*, *First measurement of ϕ_3 with a model-independent Dalitz plot analysis of $B^\pm \to DK^\pm$, $D \to K^0_S\pi^+\pi^-$ decay*, Phys. Rev. **D85** (2012) 112014, arXiv:1204.6561.

[14] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ using $B^\pm \to DK^\pm$ with $D \to K^0_S\pi^+\pi^-$, $K^0_SK^+K^-$ decays*, JHEP **10** (2014) 097, arXiv:1408.2748.

[15] LHCb collaboration, R. Aaij *et al.*, *A model-independent Dalitz plot analysis of $B^\pm \to DK^\pm$ with $D \to K^0_Sh^+h^-$ ($h = \pi, K$) decays and constraints on the CKM angle γ*, Phys. Lett. **B718** (2012) 43, arXiv:1209.5869.

[16] CLEO collaboration, J. Libby *et al.*, *Model-independent determination of the strong-phase difference between D^0 and $\bar{D}^0 \to K^0_{S,L}h^+h^-$ ($h = \pi, K$) and its impact on the measurement of the CKM angle γ/ϕ_3*, Phys. Rev. **D82** (2010) 112006, arXiv:1010.2817.

[17] BaBar collaboration, B. Aubert *et al.*, *Measurement of the Cabibbo-Kobayashi-Maskawa angle γ in $B^\mp \to D^{(*)}K^\mp$ decays with a Dalitz analysis of $D \to K^0_S\pi^-\pi^+$*, Phys. Rev. Lett. **95** (2005) 121802, arXiv:hep-ex/0504039.

[18] BaBar collaboration, B. Aubert *et al.*, *Improved measurement of the CKM angle γ in $B^{\mp} \to D^{(*)}K^{(*)\mp}$ decays with a Dalitz plot analysis of D decays to $K^0_S\pi^+\pi^-$ and $K^0_SK^+K^-$*, Phys. Rev. **D78** (2008) 034023, arXiv:0804.2089.

[19] BaBar collaboration, P. del Amo Sanchez *et al.*, *Evidence for direct CP violation in the measurement of the Cabibbo-Kobayashi-Maskawa angle γ with $B^{\mp} \to D^{(*)}K^{(*)\mp}$ decays*, Phys. Rev. Lett. **105** (2010) 121801, arXiv:1005.1096.

[20] Belle collaboration, A. Poluektov *et al.*, *Measurement of ϕ_3 with Dalitz plot analysis of $B^{\pm} \to D^{(*)}K^{\pm}$ decay*, Phys. Rev. **D70** (2004) 072003, arXiv:hep-ex/0406067.

[21] Belle collaboration, A. Poluektov *et al.*, *Measurement of ϕ_3 with a Dalitz plot analysis of $B^+ \to D^{(*)}K^{(*)+}$ decay*, Phys. Rev. **D73** (2006) 112009, arXiv:hep-ex/0604054.

[22] Belle collaboration, A. Poluektov *et al.*, *Evidence for direct CP violation in the decay $B^{\pm} \to D^{(*)}K^{\pm}$, $D \to K^0_S\pi^+\pi^-$ and measurement of the CKM phase ϕ_3*, Phys. Rev. **D81** (2010) 112002, arXiv:1003.3360.

[23] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and constraints on the CKM angle γ in $B^{\pm} \to DK^{\pm}$ with $D \to K^0_S\pi^+\pi^-$ decays*, Nucl. Phys. **B888** (2014) 169, arXiv:1407.6211.

[24] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.

[25] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation parameters in $B^0 \to DK^{*0}$ decays*, Phys. Rev. **D90** (2014) 112002, arXiv:1407.8136.

[26] BaBar Collaboration, B. Aubert *et al.*, *Constraints on the CKM angle γ in $B^0 \to \overline{D}^0K^{*0}$ and $B^0 \to D^0K^{*0}$ from a Dalitz analysis of D^0 and $\overline{D}^0$ decays to $K_S\pi^+\pi^-$*, Phys. Rev. D **79** (2009) 072003.

[27] Belle collaboration, K. Negishi *et al.*, *First model-independent Dalitz analysis of $B^0 \to DK^{*0}$, $D \to K^0_S\pi^+\pi^-$ decay*, arXiv:1509.01098.

[28] LHCb collaboration, R. Aaij *et al.*, *Model-independent measurement of the CKM angle γ using $B^0 \to DK^{*0}$ decays with $D \to K^0_S\pi^+\pi^-$ and $K^0_SK^+K^-$*, arXiv:1604.01525, submitted to JHEP.

[29] M. Gronau, *Improving bounds on γ in $B^{\pm} \to DK^{\pm}$ and $B^{\pm,0} \to DX^{\pm,0}_s$*, Phys. Lett. **B557** (2003) 198, arXiv:hep-ph/0211282.

[30] LHCb collaboration, R. Aaij *et al.*, *Constraints on the unitarity triangle angle γ from Dalitz plot analysis of $B^0 \to DK^+\pi^-$ decays*, arXiv:1602.03455, submitted to Phys. Rev. D.

[31] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[32] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[33] 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.

[34] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820; T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[35] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[36] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[37] 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.

[38] 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.

[39] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[40] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[41] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[42] 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.

[43] J. Garra Tico, *The `cfit` fitting package*, http://www.github.com/cfit.

[44] 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.

[45] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198, arXiv:hep-ex/0011057.

[46] BESIII collaboration, M. Ablikim *et al.*, *Precision measurement of the D^{*0} decay branching fractions*, Phys. Rev. **D91** (2015) 031101, arXiv:1412.4566.

[47] BaBar collaboration, P. del Amo Sanchez *et al.*, *Measurement of $D^0 - \bar{D}^0$ mixing parameters using $D^0 \to K^0_S\pi^+\pi^-$ and $D^0 \to K^0_S K^+ K^-$ decays*, Phys. Rev. Lett. **105** (2010) 081803, arXiv:1004.5053.

[48] D. Aston *et al.*, *A study of $K^-\pi^+$ scattering in the reaction $Kp \to K^-\pi^+ n$ at 11 GeV/c*, Nuclear Physics B **296** (1988) 493 .

[49] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $D^\pm$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B718** (2013) 902, arXiv:1210.4112.

[50] A. Bondar and T. Gershon, *On ϕ_3 measurements using $B^- \to D^* K^-$ decays*, Phys. Rev. **D70** (2004) 091503.

[51] V. V. Anisovich and A. V. Sarantsev, *K-matrix analysis of the ($IJ^{(PC)} = 00^{++}$)-wave in the mass region below 1900 MeV*, Eur. Phys. J. **A16** (2003) 229, arXiv:hep-ph/0204328.

[52] E791 collaboration, E. M. Aitala *et al.*, *Dalitz plot analysis of the decay $D^+ \to K^-\pi^+\pi^+$ and indication of a low-mass scalar $K\pi$ resonance*, Phys. Rev. Lett. **89** (2002) 121801, arXiv:hep-ex/0204018.

[53] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001, and 2013 partial update for the 2014 edition.

[54] LHCb collaboration, R. Aaij *et al.*, *A measurement of the CKM angle γ from a combination of $B^\pm \to Dh^\pm$ analyses*, Phys. Lett. **B726** (2013) 151, arXiv:1305.2050.

[55] G. J. Feldman and R. D. Cousins, *A unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873, arXiv:physics/9711021.

[56] S. Bodhisattva, M. Walker, and M. Woodroofe, *On the unified method with nuisance parameters*, Statist. Sinica **19** (2009) 301, http://www3.stat.sinica.edu.tw/statistica/j19n1/J19N116/J19N116.html.

LHCb collaboration

C. Abellán Beteta[41], P. Alvarez Cartelle[54], A.A. Alves Jr[58], O. Aquines Gutierrez[11], R. Aaij[39], B. Adeva[38], M. Adinolfi[47], Z. Ajaltouni[5], S. Akar[6], J. Albrecht[10], F. Alessio[39], M. Alexander[52], S. Ali[42], G. Alkhazov[31], S. Amato[2], S. Amerio[23], Y. Amhis[7], L. An[40], L. Anderlini[18], G. Andreassi[40], M. Andreotti[17,g], J.E. Andrews[59], R.B. Appleby[55], F. Archilli[39], A. Artamonov[36], M. Artuso[60], E. Aslanides[6], G. Auriemma[26,s], M. van Beuzekom[42], M. Baalouch[5], S. Bachmann[12], J.J. Back[49], A. Badalov[37], C. Baesso[61], W. Baldini[17], R.J. Barlow[55], C. Barschel[39], S. Barsuk[7], W. Barter[39], V. Batozskaya[29], V. Battista[40], A. Bay[40], L. Beaucourt[4], J. Beddow[52], F. Bedeschi[24], I. Bediaga[1], L.J. Bel[42], V. Bellee[40], N. Belloli[21,i], I. Belyaev[32], E. Ben-Haim[8], G. Bencivenni[19], S. Benson[39], J. Benton[47], A. Berezhnoy[33], R. Bernet[41], A. Bertolin[23], M.-O. Bettler[39], S. Bifani[46], P. Billoir[8], T. Bird[55], A. Birnkraut[10], A. Bitadze[55], A. Bizzeti[18,u], T. Blake[49], F. Blanc[40], J. Blouw[11], S. Blusk[60], V. Bocci[26], A. Bondar[35], N. Bondar[31,39], W. Bonivento[16], S. Borghi[55], M. Borsato[38], M. Boubdir[9], T.J.V. Bowcock[53], E. Bowen[41], C. Bozzi[17,39], S. Braun[12], M. Britsch[12], T. Britton[60], J. Brodzicka[55], E. Buchanan[47], C. Burr[55], A. Bursche[2], J. Buytaert[39], M. Calvo Gomez[37,m], D. Campora Perez[39], K. Carvalho Akiba[2], L. Castillo Garcia[40], X. Cid Vidal[39], M. Cruz Torres[61], S. Cadeddu[16], R. Calabrese[17,g], M. Calvi[21,i], P. Campana[19], L. Capriotti[55], A. Carbone[15,e], G. Carboni[25,j], R. Cardinale[20,h], A. Cardini[16], P. Carniti[21,i], L. Carson[51], G. Casse[53], L. Cassina[21,i], M. Cattaneo[39], Ch. Cauet[10], G. Cavallero[20], R. Cenci[24,t], M. Charles[8], Ph. Charpentier[39], M. Chefdeville[4], S. Chen[55], S.-F. Cheung[56], V. Chobanova[38], M. Chrzaszcz[41,27], G. Ciezarek[42], P.E.L. Clarke[51], M. Clemencic[39], H.V. Cliff[48], J. Closier[39], V. Coco[58], J. Cogan[6], E. Cogneras[5], V. Cogoni[16,f], L. Cojocariu[30], G. Collazuol[23,o], P. Collins[39], A. Comerma-Montells[12], A. Contu[39], A. Cook[47], S. Coquereau[8], G. Corti[39], M. Corvo[17,g], B. Couturier[39], G.A. Cowan[51], D.C. Craik[51], A. Crocombe[49], S. Cunliffe[54], R. Currie[54], C. D'Ambrosio[39], O. De Aguiar Francisco[2], K. De Bruyn[6], S. De Capua[55], M. De Cian[12], J.M. De Miranda[1], L. De Paula[2], P. De Simone[19], L. Del Buono[8], A. Di Canto[39], A. Dosil Suárez[38], E. Dall'Occo[42], J. Dalseno[47], P.N.Y. David[42], A. Davis[58], C.-T. Dean[52], D. Decamp[4], M. Deckenhoff[10], M. Demmer[10], D. Derkach[67], O. Deschamps[5], F. Dettori[39], B. Dey[22], H. Dijkstra[39], F. Dordei[39], M. Dorigo[40], A. Dovbnya[44], K. Dreimanis[53], L. Dufour[42], G. Dujany[55], P. Durante[39], R. Dzhelyadin[36], A. Dziurda[39], A. Dzyuba[31], N. Déléage[4], I. El Rifai[5], S. Easo[50,39], U. Egede[54], V. Egorychev[32], S. Eidelman[35], S. Eisenhardt[51], U. Eitschberger[10], R. Ekelhof[10], L. Eklund[52], Ch. Elsasser[41], S. Ely[60], S. Esen[12], H.M. Evans[48], T. Evans[56], V. Fernandez Albor[38], F. Ferreira Rodrigues[1], A. Falabella[15], N. Farley[46], S. Farry[53], R. Fay[53], D. Ferguson[51], F. Ferrari[15,39], M. Ferro-Luzzi[39], S. Filippov[34], M. Fiore[17,g], M. Fiorini[17,g], M. Firlej[28], C. Fitzpatrick[40], T. Fiutowski[28], F. Fleuret[7,b], K. Fohl[39], M. Fontana[16], F. Fontanelli[20,h], D.C. Forshaw[60], R. Forty[39], M. Frank[39], C. Frei[39], M. Frosini[18], J. Fu[22], E. Furfaro[25,j], C. Färber[39], A. Gallas Torreira[38], J. García Pardiñas[38], J. Garra Tico[48], M. Grabalosa Gándara[5], R. Graciani Diaz[37], L.A. Granado Cardoso[39], D. Galli[15,e], S. Gallorini[23], S. Gambetta[51], M. Gandelman[2], P. Gandini[56], Y. Gao[3], L. Garrido[37], P.J. Garsed[48], D. Gascon[37], C. Gaspar[39], L. Gavardi[10], G. Gazzoni[5], D. Gerick[12], E. Gersabeck[12], M. Gersabeck[55], T. Gershon[49], Ph. Ghez[4], S. Gianì[40], V. Gibson[48], O.G. Girard[40], L. Giubega[30], V.V. Gligorov[8], D. Golubkov[32], A. Golutvin[54,39], A. Gomes[1,a], C. Gotti[21,i], E. Graugés[37], E. Graverini[41], G. Graziani[18], A. Grecu[30], P. Griffith[46], L. Grillo[12], O. Grünberg[65], E. Gushchin[34], Yu. Guz[36,39], T. Gys[39], C. Göbel[61], J.A. Hernando Morata[38], E. van Herwijnen[39], T. Hadavizadeh[56], C. Hadjivasiliou[60], G. Haefeli[40], C. Haen[39], S.C. Haines[48], S. Hall[54], B. Hamilton[59], X. Han[12], S. Hansmann-Menzemer[12], N. Harnew[56], S.T. Harnew[47], J. Harrison[55], J. He[39], T. Head[40], A. Heister[9], K. Hennessy[53], P. Henrard[5], L. Henry[8], M. Heß[65], A. Hicheur[2], D. Hill[56], M. Hoballah[5], C. Hombach[55], W. Hulsbergen[42], T. Humair[54], N. Hussain[56], D. Hutchcroft[53], M. Idzik[28], P. Ilten[57], R. Jacobsson[39], A. Jaeger[12], J. Jalocha[56], E. Jans[42], A. Jawahery[59], M. John[56], D. Johnson[39], C.R. Jones[48], C. Joram[39], B. Jost[39], N. Jurik[60],

S. Kandybei[44], W. Kanso[6], M. Karacson[39], T.M. Karbach[39,†], S. Karodia[52], M. Kecke[12], M. Kelsey[60], I.R. Kenyon[46], M. Kenzie[39], T. Ketel[43], E. Khairullin[67], B. Khanji[21,39,i], C. Khurewathanakul[40], T. Kirn[9], S. Klaver[55], K. Klimaszewski[29], M. Kolpin[12], I. Komarov[40], R.F. Koopman[43], P. Koppenburg[42], M. Kozeiha[5], L. Kravchuk[34], K. Kreplin[12], M. Kreps[49], P. Krokovny[35], F. Kruse[10], W. Krzemien[29], W. Kucewicz[27,l], M. Kucharczyk[27], V. Kudryavtsev[35], A.K. Kuonen[40], K. Kurek[29], T. Kvaratskheliya[32], R. Le Gac[6], J. van Leerdam[42], E. Lemos Cid[38], M. Lucio Martinez[38], D. Lacarrere[39], G. Lafferty[55,39], A. Lai[16], D. Lambert[51], G. Lanfranchi[19], C. Langenbruch[49], B. Langhans[39], T. Latham[49], C. Lazzeroni[46], J.-P. Lees[4], A. Leflat[33,39], J. Lefrançois[7], R. Lefèvre[5], O. Leroy[6], T. Lesiak[27], B. Leverington[12], Y. Li[7], T. Likhomanenko[67,66], R. Lindner[39], C. Linn[39], F. Lionetto[41], B. Liu[16], X. Liu[3], D. Loh[49], I. Longstaff[52], J.H. Lopes[2], D. Lucchesi[23,o], H. Luo[51], A. Lupato[23], E. Luppi[17,g], O. Lupton[56], A. Lusiani[24], X. Lyu[62], C. Marin Benito[37], D. Martinez Santos[38], F. Martinez Vidal[68], D. Martins Tostes[2], J. Molina Rodriguez[61], F. Machefert[7], F. Maciuc[30], O. Maev[31], K. Maguire[55], S. Malde[56], A. Malinin[66], G. Manca[7], G. Mancinelli[6], P. Manning[60], A. Mapelli[39], J. Maratas[5], J.F. Marchand[4], U. Marconi[15], P. Marino[24,t], J. Marks[12], G. Martellotti[26], M. Martin[6], M. Martinelli[40], L.M. Massacrier[7], A. Massafferri[1], R. Matev[39], A. Mathad[49], Z. Mathe[39], C. Matteuzzi[21], A. Mauri[41], B. Maurin[40], A. Mazurov[46], M. McCann[54], J. McCarthy[46], A. McNab[55], R. McNulty[13], B. Meadows[58], F. Meier[10], M. Meissner[12], D. Melnychuk[29], M. Merk[42], E Michielin[23], D.A. Milanes[64], M.-N. Minard[4], D.S. Mitzel[12], I.A. Monroy[64], S. Monteil[5], M. Morandin[23], P. Morawski[28], A. Mordà[6], M.J. Morello[24,t], J. Moron[28], A.B. Morris[51], R. Mountain[60], F. Muheim[51], M. Mussini[15], B. Muster[40], D. Müller[55], J. Müller[10], K. Müller[41], V. Müller[10], P. Naik[47], T. Nakada[40], R. Nandakumar[50], A. Nandi[56], I. Nasteva[2], M. Needham[51], N. Neri[22], S. Neubert[12], N. Neufeld[39], M. Neuner[12], A.D. Nguyen[40], C. Nguyen-Mau[40,n], V. Niess[5], S. Nieswand[9], R. Niet[10], N. Nikitin[33], T. Nikodem[12], A. Novoselov[36], D.P. O'Hanlon[49], B. Osorio Rodrigues[1], J.M. Otalora Goicochea[2], A. Oblakowska-Mucha[28], V. Obraztsov[36], S. Ogilvy[19], O. Okhrimenko[45], R. Oldeman[48], C.J.G. Onderwater[69], A. Otto[39], P. Owen[54], A. Oyanguren[68], M. Pepe Altarelli[39], E. Picatoste Olloqui[37], M. Plo Casasus[38], A. Puig Navarro[40], A. Palano[14,d], F. Palombo[22,q], M. Palutan[19], J. Panman[39], A. Papanestis[50], M. Pappagallo[52], L.L. Pappalardo[17,g], C. Pappenheimer[58], W. Parker[59], C. Parkes[55], G. Passaleva[18], G.D. Patel[53], M. Patel[54], C. Patrignani[15,e], A. Pearce[55,50], A. Pellegrino[42], G. Penso[26,k], S. Perazzini[39], P. Perret[5], L. Pescatore[46], K. Petridis[47], A. Petrolini[20,h], M. Petruzzo[22], B. Pietrzyk[4], D. Pinci[26], A. Pistone[20], A. Piucci[12], S. Playfer[51], T. Poikela[39], F. Polci[8], A. Poluektov[49,35], I. Polyakov[32], E. Polycarpo[2], A. Popov[36], D. Popov[11,39], B. Popovici[30], C. Potterat[2], E. Price[47], J.D. Price[53], J. Prisciandaro[38], A. Pritchard[53], C. Prouve[47], V. Pugatch[45], G. Punzi[24,p], W. Qian[56], R. Quagliani[7,47], M. Ramos Pernas[38], A.C. dos Reis[1], V. Rives Molina[37], J.A. Rodriguez Lopez[64], P. Rodriguez Perez[55], A. Romero Vidal[38], P. Ruiz Valls[68], B. Rachwal[27], J.H. Rademacker[47], M. Rama[24], M.S. Rangel[2], I. Raniuk[44], G. Raven[43], F. Redi[54], S. Reichert[10], V. Renaudin[7], S. Ricciardi[50], S. Richards[47], M. Rihl[39], K. Rinnert[53,39], P. Robbe[7], A.B. Rodrigues[1], E. Rodrigues[55], A. Rogozhnikov[67], S. Roiser[39], V. Romanovskiy[36], J.W. Ronayne[13], M. Rotondo[23], T. Ruf[39], J.J. Saborido Silva[38], V. Salustino Guimaraes[2], C. Sanchez Mayordomo[68], B. Sanmartin Sedes[38], C. Santamarina Rios[38], R. Silva Coutinho[41], L. Silva de Oliveira[2], B. Souza De Paula[2], N. Sagidova[31], B. Saitta[16,f], R. Santacesaria[26], M. Santimaria[19], E. Santovetti[25,j], A. Sarti[19,k], C. Satriano[26,s], A. Satta[25], D.M. Saunders[47], D. Savrina[32,33], S. Schael[9], M. Schiller[39], H. Schindler[39], M. Schlupp[10], M. Schmelling[11], T. Schmelzer[10], B. Schmidt[39], O. Schneider[40], A. Schopper[39], M. Schubiger[40], M.-H. Schune[7], R. Schwemmer[39], B. Sciascia[19], A. Sciubba[26,k], A. Semennikov[32], A. Sergi[46], N. Serra[41], J. Serrano[6], L. Sestini[23], P. Seyfert[21], M. Shapkin[36], I. Shapoval[17,44,g], Y. Shcheglov[31], T. Shears[53], L. Shekhtman[35], V. Shevchenko[66], A. Shires[10], B.G. Siddi[17], G. Simi[23,o], M. Sirendi[48], N. Skidmore[47], T. Skwarnicki[60], E. Smith[54], I.T. Smith[51], J. Smith[48], M. Smith[55], H. Snoek[42], M.D. Sokoloff[58], F.J.P. Soler[52], F. Soomro[40], D. Souza[47],

B. Spaan[10], P. Spradlin[52], S. Sridharan[39], F. Stagni[39], M. Stahl[12], S. Stahl[39], S. Stefkova[54], O. Steinkamp[41], O. Stenyakin[36], S. Stevenson[56], S. Stoica[30], S. Stone[60], B. Storaci[41], S. Stracka[24,t], M. Straticiuc[30], U. Straumann[41], L. Sun[58], W. Sutcliffe[54], K. Swientek[28], S. Swientek[10], V. Syropoulos[43], M. Szczekowski[29], T. Szumlak[28], S. T'Jampens[4], J. van Tilburg[42], A. Tayduganov[6], T. Tekampe[10], G. Tellarini[17,g], F. Teubert[39], C. Thomas[56], E. Thomas[39], V. Tisserand[4], M. Tobin[40], S. Tolk[43], L. Tomassetti[17,g], D. Tonelli[39], S. Topp-Joergensen[56], E. Tournefier[4], S. Tourneur[40], K. Trabelsi[40], M.T. Tran[40], M. Tresch[41], A. Trisovic[39], A. Tsaregorodtsev[6], P. Tsopelas[42], N. Tuning[42,39], A. Ukleja[29], A. Ustyuzhanin[67,66], U. Uwer[12], R. Vazquez Gomez[19], P. Vazquez Regueiro[38], M. van Veghel[42], M. Vieites Diaz[38], J.A. de Vries[42], C. Vacca[16,39,f], V. Vagnoni[15,39], S. Valat[39], G. Valenti[15], A. Vallier[7], S. Vecchi[17], J.J. Velthuis[47], M. Veltri[18,r], G. Veneziano[40], M. Vesterinen[12], B. Viaud[7], D. Vieira[2], X. Vilasis-Cardona[37,m], V. Volkov[33], A. Vollhardt[41], D. Voong[47], A. Vorobyev[31], V. Vorobyev[35], C. Voß[65], C. Vázquez Sierra[38], R. Waldi[65], C. Wallace[49], R. Wallace[13], J. Walsh[24], J. Wang[60], D.R. Ward[48], N.K. Watson[46], D. Websdale[54], A. Weiden[41], M. Whitehead[39], J. Wicht[49], G. Wilkinson[56,39], M. Wilkinson[60], M. Williams[39], M.P. Williams[46], M. Williams[57], T. Williams[46], F.F. Wilson[50], J. Wimberley[59], J. Wishahi[10], W. Wislicki[29], M. Witek[27], G. Wormser[7], S.A. Wotton[48], K. Wraight[52], S. Wright[48], K. Wyllie[39], Y. Xie[63], Z. Xu[40], Z. Yang[3], H. Yin[63], J. Yu[63], X. Yuan[35], O. Yushchenko[36], M. Zangoli[15], M. Zavertyaev[11,c], L. Zhang[3], Y. Zhang[7], A. Zhelezov[12], Y. Zheng[62], A. Zhokhov[32], L. Zhong[3], V. Zhukov[9], S. Zucchelli[15], P. d'Argent[12].

[1] *Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil*
[2] *Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil*
[3] *Center for High Energy Physics, Tsinghua University, Beijing, China*
[4] *LAPP, Université Savoie Mont-Blanc, CNRS/IN2P3, Annecy-Le-Vieux, France*
[5] *Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, Clermont-Ferrand, France*
[6] *CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France*
[7] *LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France*
[8] *LPNHE, Université Pierre et Marie Curie, Université Paris Diderot, CNRS/IN2P3, Paris, France*
[9] *I. Physikalisches Institut, RWTH Aachen University, Aachen, Germany*
[10] *Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany*
[11] *Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany*
[12] *Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany*
[13] *School of Physics, University College Dublin, Dublin, Ireland*
[14] *Sezione INFN di Bari, Bari, Italy*
[15] *Sezione INFN di Bologna, Bologna, Italy*
[16] *Sezione INFN di Cagliari, Cagliari, Italy*
[17] *Sezione INFN di Ferrara, Ferrara, Italy*
[18] *Sezione INFN di Firenze, Firenze, Italy*
[19] *Laboratori Nazionali dell'INFN di Frascati, Frascati, Italy*
[20] *Sezione INFN di Genova, Genova, Italy*
[21] *Sezione INFN di Milano Bicocca, Milano, Italy*
[22] *Sezione INFN di Milano, Milano, Italy*
[23] *Sezione INFN di Padova, Padova, Italy*
[24] *Sezione INFN di Pisa, Pisa, Italy*
[25] *Sezione INFN di Roma Tor Vergata, Roma, Italy*
[26] *Sezione INFN di Roma La Sapienza, Roma, Italy*
[27] *Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland*
[28] *AGH - University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland*
[29] *National Center for Nuclear Research (NCBJ), Warsaw, Poland*

[30] Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
[31] Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
[32] Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia
[33] Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
[34] Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia
[35] Budker Institute of Nuclear Physics (SB RAS) and Novosibirsk State University, Novosibirsk, Russia
[36] Institute for High Energy Physics (IHEP), Protvino, Russia
[37] Universitat de Barcelona, Barcelona, Spain
[38] Universidad de Santiago de Compostela, Santiago de Compostela, Spain
[39] European Organization for Nuclear Research (CERN), Geneva, Switzerland
[40] Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[41] Physik-Institut, Universität Zürich, Zürich, Switzerland
[42] Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands
[43] Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, The Netherlands
[44] NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
[45] Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
[46] University of Birmingham, Birmingham, United Kingdom
[47] H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
[48] Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
[49] Department of Physics, University of Warwick, Coventry, United Kingdom
[50] STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
[51] School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
[52] School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
[53] Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
[54] Imperial College London, London, United Kingdom
[55] School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
[56] Department of Physics, University of Oxford, Oxford, United Kingdom
[57] Massachusetts Institute of Technology, Cambridge, MA, United States
[58] University of Cincinnati, Cincinnati, OH, United States
[59] University of Maryland, College Park, MD, United States
[60] Syracuse University, Syracuse, NY, United States
[61] Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to [2]
[62] University of Chinese Academy of Sciences, Beijing, China, associated to [3]
[63] Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China, associated to [3]
[64] Departamento de Fisica , Universidad Nacional de Colombia, Bogota, Colombia, associated to [8]
[65] Institut für Physik, Universität Rostock, Rostock, Germany, associated to [12]
[66] National Research Centre Kurchatov Institute, Moscow, Russia, associated to [32]
[67] Yandex School of Data Analysis, Moscow, Russia, associated to [32]
[68] Instituto de Fisica Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain, associated to [37]
[69] Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to [42]

[a] Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
[b] Laboratoire Leprince-Ringuet, Palaiseau, France
[c] P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
[d] Università di Bari, Bari, Italy
[e] Università di Bologna, Bologna, Italy
[f] Università di Cagliari, Cagliari, Italy
[g] Università di Ferrara, Ferrara, Italy
[h] Università di Genova, Genova, Italy
[i] Università di Milano Bicocca, Milano, Italy
[j] Università di Roma Tor Vergata, Roma, Italy

[k] *Università di Roma La Sapienza, Roma, Italy*
[l] *AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland*
[m] *LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain*
[n] *Hanoi University of Science, Hanoi, Viet Nam*
[o] *Università di Padova, Padova, Italy*
[p] *Università di Pisa, Pisa, Italy*
[q] *Università degli Studi di Milano, Milano, Italy*
[r] *Università di Urbino, Urbino, Italy*
[s] *Università della Basilicata, Potenza, Italy*
[t] *Scuola Normale Superiore, Pisa, Italy*
[u] *Università di Modena e Reggio Emilia, Modena, Italy*

[†] *Deceased*

8. Study of γ using the simultaneous fit to all $B \to D^{(*)}K^{(*)}$ decay modes

The mass fit results obtained in Chapter 5 act as inputs to the final $\mathcal{CP}$ fit, performed in the $D \to K^0_S\pi^+\pi^-$ Dalitz plane, as described in Section 1.5. Selected candidates lying within the signal region of each mass fit are analysed simultaneously, and both model-dependent and model-independent fits are performed. Moving from performing a model-dependent $\mathcal{CP}$ fit for a single decay mode, as described in Chapter 7, to performing *both* model-dependent and model-independent analyses for *multiple* decay modes simulataneously, presents a significant technical challenge. This work has been performed by a collaborator with significant expertise in $\mathcal{CP}$ analysis, and not by the present author. This chapter describes the fit models used, and presents the first results obtained, before discussing the estimation of systematic uncertainties.

It is useful to begin by introducing the relevant notation, which extends that of Section 1.5.

Model-dependent

The model-dependent approach begins with equations 1.117 and 1.118, which present the PDF for a signal B-meson decay of the form studied in this thesis:

$$\begin{aligned} \mathrm{d}\Gamma^m_- &\sim \left|A_f\right|^2 + |z^m_-|^2\left|\bar{A}_f\right|^2 + 2\kappa_m \,\mathrm{Re}\left(z^m_-\, A^\star_f \bar{A}_f\right); \quad \text{and} \\ \mathrm{d}\Gamma^m_+ &\sim \left|\bar{A}_f\right|^2 + |z^m_+|^2\left|A_f\right|^2 + 2\kappa_m \,\mathrm{Re}\left(z^m_+\, \bar{A}^\star_f A_f\right), \end{aligned}$$

where A_f gives the amplitude for the observed (K^0_S, π^+, π^-) momentum configuration to have arisen from a decay proceeding via a D^0-meson, and $\bar{A}_f$ gives the amplitude to have proceeded via a $\overline{D^0}$-meson. A resonance model is used to describe these amplitudes, as described in Chapter 2.[1] In order to model the observed distribution of events within the Dalitz plane, it is necessary to account for any efficiency variations, and so each of these expressions must be multiplied by an efficiency function before it can be fitted to data.

Defining the functional

$$\mathcal{P}_{\mathrm{MD}}(A, \xi z, \kappa, \epsilon) = \left(|A|^2 + |\xi z|^2|\bar{A}^2_f| + 2\kappa \mathrm{Re}(\xi z A^* \bar{A}_f)\right) \cdot \epsilon, \tag{8.1}$$

[1] Recall that A_f and $\bar{A}_f$ are linked by $\bar{A}_f(s_{12}, s_{13}) = A_f(s_{13}, s_{12})$.

the normalised PDF models for a given pair of conjugate decay modes are then given by

$$\mathrm{d}\Gamma^m_- = \frac{1}{N_m}\mathcal{P}_{\mathrm{MD}}(A_f, \xi_m z_-, \kappa_m, \epsilon_m), \quad \text{and} \tag{8.2}$$

$$\mathrm{d}\Gamma^m_+ = \frac{1}{N_m}\mathcal{P}_{\mathrm{MD}}(\bar{A}_f, \xi_m z_+, \kappa_m, \epsilon_m), \tag{8.3}$$

where N_m is the population of mode m within the model, and ϵ_m describes its efficiency variation over the Dalitz plane.

Note that the same functional also applies to decay modes for which the suppressed component, $\overline{A}$, is negligible, by setting $\xi_m = 0$. A flat PDF can also be obtained, by setting $A = 1$ and $\xi_m = 0$. This means that the same functional may be used to describe all of the signal and background components included in the $\mathcal{CP}$ fit.

Model-independent

As introduced in Section 1.5, the model-independent approach integrates the PDFs of equations 1.117 and 1.118 to model the bin populations of observed decays within the Dalitz plane, giving equations 1.129–1.132:

$$\begin{aligned}
\Gamma^m_{-,i} &\sim T_i + |z^m_-|^2 T_{\tilde{i}} + 2\kappa_m |z^m_-| \Big\{ c_i \cos\Big(\arg(z^m_-)\Big) - s_i \sin\Big(\arg(z^m_-)\Big) \Big\}; \\
\Gamma^m_{-,\tilde{i}} &\sim T_{\tilde{i}} + |z^m_-|^2 T_i + 2\kappa_m |z^m_-| \Big\{ c_i \cos\Big(\arg(z^m_-)\Big) + s_i \sin\Big(\arg(z^m_-)\Big) \Big\}; \\
\Gamma^m_{+,i} &\sim T_{\tilde{i}} + |z^m_+|^2 T_i + 2\kappa_m |z^m_+| \Big\{ c_i \cos\Big(\arg(z^m_+)\Big) + s_i \sin\Big(\arg(z^m_+)\Big) \Big\}; \quad \text{and} \\
\Gamma^m_{+,\tilde{i}} &\sim T_i + |z^m_+|^2 T_{\tilde{i}} + 2\kappa_m |z^m_+| \Big\{ c_i \cos\Big(\arg(z^m_+)\Big) - s_i \sin\Big(\arg(z^m_+)\Big) \Big\}.
\end{aligned}$$

Each of these expressions may be expressed using the functional

$$\mathcal{P}(T_i, T_{\tilde{i}}, c_i, s_i, \xi z, \kappa, \epsilon) = T_i + |\xi z|^2 T_{\tilde{i}} + 2\kappa |\xi z| \Big\{ c_i \cos\Big(\arg(\xi z)\Big) - s_i \sin\Big(\arg(\xi z)\Big) \Big\}. \tag{8.4}$$

Within the model-independent approach, the amplitudes A_f and $\bar{A}_f$ are never modelled directly, but all information relating to them is carried by the set of bin-integrated parameters $\{T_i, T_{\tilde{i}}, c_i, s_i\}$, so it is useful to denote this set of information as A_f. Making this substitution, and generalising to allow for non-uniform efficiency, the model-independent PDF may be described by a functional of the form,

$$\mathcal{P}_{\mathrm{MI}}(A, \xi z, \kappa, \epsilon). \tag{8.5}$$

This functional takes the same arguments as $\mathcal{P}_{\mathrm{MD}}$ and thus unifies the notation between the model-dependent and model-independent approaches.

name	channel	dominant amplitude	ξ_c	κ_c
`bdk-`	$B^- \to DK^-$	A_f	1	1
`bdpi-`	$B^- \to D\pi^-$	A_f	0	1
`b0dkst-`	$B^0 \to DK^{*0}$	A_f	$\xi_{DK^{*0}}$	$\kappa_{DK^{*0}}$
`bdkst-`	$B^- \to DK^{*-}$	A_f	$\xi_{DK^{*\pm}}$	$\kappa_{DK^{*\pm}}$
`bdstkg-`	$B^- \to D^*(D\gamma)K^-$	A_f	$-\xi_{D^*K}$	1
`bdstkr-`	$B^- \to D^*(D\pi^0)K^-$	A_f	ξ_{D^*K}	1
`bdstkm-`	$B^- \to D^*(D\pi^0_{\text{part.reco.}})K^-$	A_f	ξ_{D^*K}	1
`bdstpig-`	$B^- \to D^*(D\gamma)\pi^-$	A_f	0	1
`bdstpir-`	$B^- \to D^*(D\pi^0)\pi^-$	A_f	0	1
`bdstpim-`	$B^- \to D^*(D\pi^0_{\text{part.reco.}})\pi^-$	A_f	0	1
`d0+`	$D^0 \to K^0_S\pi^+\pi^-$	A_f	0	1
`d0-`	$\overline{D}^0 \to K^0_S\pi^+\pi^-$	$\bar{A}_f$	0	1
`flat`	flat	1	0	–

Table 8.1.: Phase space PDF models used for signal and background components included in the $\mathcal{CP}$ violation fit.

8.1. $\mathcal{CP}$ fit model

The phase space PDFs used to describe each signal and background component can be expressed in terms of the functionals defined above. The only differences between each of the components are the dominant amplitude, A, passed as the first argument of the functional, and the ξ_m and κ_m parameters. Table 8.1 summarises all of the phase space models that are defined. Note that, thanks to the shared notation, the same expressions give the PDFs used in both the model-dependent and model-independent fits. PDFs for the opposite charge are obtained by swapping $A_f \leftrightarrow \bar{A}_f$ and $z_- \leftrightarrow z_+$. The notation $B^- \to D^*(D\pi^0_{\text{part.reco.}})K^-$ indicates the $B^- \to D^*K^-$ signal mode with partially-reconstructed $D^* \to D\pi^0$, $\pi^0 \to \gamma\gamma$, as discussed in Section 4.3.3. The notation $B^- \to D^*(D\pi^0_{\text{part.reco.}})\pi^-$ indicates the corresponding control mode. The letters `r` and `m` in `bdstkr`, `bdstkm`, `bdstpir` and `bdstpim` stand for 'reconstructed' and 'misreconstructed', respectively, and refer to the π^0 candidate arising from the D^{*0} decay. An additional PDF model, `d0=`, is defined as the average of the `d0+` and `d0-` models. The value of the coherence factor for decay mode DK^{*0} is extracted from reference [152], and the value of the coherence factor for decay mode $DK^{*\pm}$ is calculated from the model described in reference [153].

Tables 8.2 – 8.6 indicate which of the defined phase space PDFs is used to model each signal and background component of each B-meson decay mode. Different PDFs are used to model combinatoric background reconstructed using real and fake D-meson candidates. For fake-D candidates, the final state (K^0_S, π^+, π^-) does not originate from a progenitor D-meson, and so the distribution in the Dalitz plane is assumed to be flat. For real-D candidates, the decay is assumed to proceed only via its dominant intermediate (D^0 or $\overline{D^0}$), without interference. Since the reconstructed decay chain contains combinatoric candidates, the flavour of the D-meson, whether D^0 or $\overline{D^0}$, cannot be tagged, and so an

	$B^- \to DK^-$	$B^- \to D\pi^-$
signal	`bdk-`	`bdpi-`
fake-D combinatoric	`flat`	`flat`
real-D combinatoric	`d0=`	`d0=`
crossfeed	`bdpi-`	`bdk-`
$B^- \to D^*K^-$	`bdstk`	–
$B^- \to D^*\pi^-$	`d0+`	`d0+`
$B^- \to D^*\rho^-$	`d0+`	`d0+`
$B^0 \to D^*\rho^0$	`d0+`	`d0+`
$B^- \to D\rho^-$	`d0+`	`d0+`
$B^0 \to D\rho^0$	`d0+`	`d0+`
$B_s^0 \to DK^{*0}$	`d0-`	–
$B^0 \to DK^{*0}$	`b0dkst-`	–

Table 8.2.: Phase space PDFs used to describe each component in the fit to $B^\pm \to Dh^\pm$ decay modes. Orange rows indicate decays which are signal in another category.

equal sum of the two models is used. The populations of fake-D and real-D combinatoric entering each sample are fixed from Section 5.5.

In addition to combining multiple decay modes into a single simultaneous fit, the PDF models defined may be used to perform standalone fits to individual decay modes. In this case, any cross-feed of signal decays from one mode into another is modelled using its dominant component only, instead of a correct PDF model with unknown $\mathcal{CP}$ parameters. This applies to *e.g.* background arising from $B^\pm \to DK^\pm$ candidates which enters the standalone fit to mode $B^\pm \to D^*K^\pm$. The decay model used for these components is `d0+` for models with a minus sign and `d0-` for models with a plus sign, i.e. the pure flavour models with the same dominant flavour amplitude. The correct PDF model for the crossfeed component is not used, since this introduces additional free $\mathcal{CP}$ parameters in the fit, and thus significantly increases the uncertainty on each parameter. The simultaneous fit is thus a unique tool, since it is the only proposed approach that assigns an accurate decay model for crossfeed components. Rows in Tables 8.2 – 8.6 describing such cross-feed components are highlighted in orange text.

The final fit PDF is obtained by combining the mass and phase space terms of each component, with their corresponding yield N_m,

$$\mathcal{P} = \sum_m N_m \, \mathcal{P}^m_{\text{mass}} \, \mathcal{P}^m_{\pm} \,. \tag{8.6}$$

All the yields and parameters of the mass PDF are fixed to the results obtained from the fits to B and D mass distributions in Chapter 5. This allows the likelihood to take into account the different mass distributions of the different signal and background categories.

	$B^0 \to DK^{*0}$	$B_s^0 \to DK^{*0}$
B^0 signal	`b0dkst-`	`b0dkst+`
B_s^0 signal	`d0-`	`d0+`
fake-D combinatoric	`flat`	`flat`
real-D combinatoric	`d0=`	`d0=`
$B_s^0 \to D^*(D\gamma)K^{*0}$	`d0-`	`d0+`
$B_s^0 \to D^*(D\pi^0)K^{*0}$	`d0-`	`d0+`
$B^0 \to D^*(D\gamma)K^{*0}$	`d0-`	`d0+`
$B^0 \to D^*(D\pi^0)K^{*0}$	`d0-`	`d0+`

Table 8.3.: Phase space PDFs used to describe each component in the fit to $B^0_{(s)} \to DK^{*0}$ decay modes.

	$B^- \to DK^{*-}$
signal	`bdkst-`
fake-D combinatoric	`flat`
real-D combinatoric	`d0=`
$B^- \to D^*(D\gamma)K^{*-}$	`d0+`
$B^- \to D^*(D\pi^0)K^{*-}$	`d0+`

Table 8.4.: Phase space PDFs used to describe each component in the fit to the $B^\pm \to DK^{*\pm}$ decay mode.

	$B^- \to D^*(D\gamma)K^-$	$B^- \to D^*(D\pi^0)K^-$	$B^- \to D^*(D\pi^0_{\text{part.reco.}})K^-$
signal	`bdstkg-`	`bdstkr-`	`bdstkm-`
fake-D combinatoric	`flat`	`flat`	`flat`
real-D combinatoric	`d0=`	`d0=`	`d0=`
crossfeed	`bdstpig-`	`bdstpir-`	`bdstpim-`
$B^- \to D^*K^{*-}$	`d0+`	`d0+`	`d0+`
$B^- \to D\pi^-$	`bdpi-`	`bdpi-`	`bdpi-`
$B^- \to DK^-$	`bdk-`	`bdk-`	`bdk-`
$B^0 \to D^{*+}K^-$	`d0+`	`d0+`	`d0+`

Table 8.5.: Phase space PDFs used to describe each component in the fit to the $B^\pm \to D^*K^\pm$ decay modes. Orange rows indicate decays which are signal in another category.

	$B^- \to D^*(D\gamma)\pi^-$	$B^- \to D^*(D\pi^0)\pi^-$	$B^- \to D^*(D\pi^0{}_{\text{part.reco.}})\pi^-$
signal	`d0+`	`d0+`	`d0+`
fake-D combinatoric	`flat`	`flat`	`flat`
real-D combinatoric	`d0=`	`d0=`	`d0=`
crossfeed	`bdstkg-`	`bdstkr-`	`bdstkm-`
$B^- \to D^*\rho^-$	`d0+`	`d0+`	`d0+`
$B^- \to D\pi^-$	`d0+`	`d0+`	`d0+`
$B^- \to D\rho^-$	`d0+`	`d0+`	`d0+`
$B^0 \to D^{*+}\pi^-$	`d0+`	`d0+`	`d0+`

Table 8.6.: Phase space PDFs used to describe each component in the fit to the $B^\pm \to D^*\pi^\pm$ decay modes.

Fit results are obtained using the `cfit` library [124] through an extended maximum likelihood fit, by minimising the function

$$-2\sum_j \ln \mathcal{P}_j + 2\sum_m N_m, \tag{8.7}$$

where $\mathcal{P}_j$ is the full PDF evaluated for datum j.

Efficiency non-uniformities over the D decay phase space are accounted for using a second order polynomial function, in the phase space variables S_+^2 and S_-^2,

$$\epsilon^m(x,y) = a + b(x+y) + cxy + d(x^2+y^2), \tag{8.8}$$
$$x = S_+^2 - m_0^2, \tag{8.9}$$
$$y = S_-^2 - m_0^2, \tag{8.10}$$

where the constant $m_0^2 = 1.9\,\text{GeV}^2$ improves the fit convergence by centering the polynomial variables within the allowed region of the Dalitz plane.

For $B^\pm \to Dh^\pm$ and $B^\pm \to D^*h^\pm$ fits, where candidate yields are high, the coefficients b, c and d of the efficiency functions float and are determined solely from data in the simultaneous fit. Coefficient a is fixed to be 1. For $B^0 \to DK^{*0}$ and $B^\pm \to DK^{*\pm}$ fits, the MC efficiency histograms described in Chapter 6 are used to introduce a Gaussian constraint on the efficiency parameters. The MC efficiencies take the form of a collection of bin contents ϵ_b with corresponding uncertainties σ_ϵ^b, using an arbitrary scale. The arbitrary scale is parameterized by introducing a parameter C_m to multiply the efficiency model ϵ^m for these modes. The Gaussian constraint is introduced by means of an additional χ^2 term in the fit function, which becomes

$$-2\sum_j \ln \mathcal{P}_j + 2\sum_m N_m + \sum_m \sum_b \left(\frac{\epsilon_b - C_m\,\epsilon^m}{\sigma_\epsilon^b}\right)^2. \tag{8.11}$$

channel	standalone					simultaneous		
$B^\pm \to DK^\pm$	x	x				x	x	x
$B^\pm \to D\pi^\pm$		x				x	x	x
$B^0 \to DK^{*0}$			x			x	x	x
$B_s^0 \to DK^{*0}$			x	x		x	x	x
$B^\pm \to DK^{*\pm}$				x			x	x
$B^\pm \to D^*h^\pm$					x			x

Table 8.7.: Decay mode configurations for which $\mathcal{CP}$ fits are performed.

8.2. Preliminary results

Fits are performed for a range of standalone and simultaneous decay mode combinations. The configurations used are summarised in Table 8.7. Standalone fit results for the Cartesian parameters are shown in Table 8.8. Simultaneous fit results for the Cartesian parameters are shown in Table 8.9. Only statistical uncertainties are shown, since all central values remain blinded whilst the analysis proceeds through peer review within the LHCb collaboration.

For each fit, the blinded Cartesian central values, and their statistical covariances, may be propagated to give an estimated statistical uncertainty on the value of γ, as described in Chapter 7. For the model-dependent and model-independent simultaneous fits including all decay modes studied, preliminary statistical uncertainties of 12° and 13.5° on γ are obtained, respectively.

8.3. Expected systematic uncertainties

Systematic uncertainties for the full simultaneous analysis have not yet been evaluated. However, their expected order of magnitude can be estimated by considering the previous standalone measurements that have been completed. The model-dependent approach has been used to extract Cartesian $\mathcal{CP}$ parameters in the decay modes

- $B^\pm \to D(\to K^0_S\pi^+\pi^-)K^\pm$, using data corresponding to an integrated luminosity of 1 fb^{-1} recorded at LHCb during 2011 [154], and
- $B^0 \to D(\to K^0_S\pi^+\pi^-)K^{*0}$, using the full LHCb Run 1 dataset [1], corresponding to an integrated luminosity of 3 fb^{-1}, as described in Chapter 7.

The model-independent approach has been used to extract Cartesian $\mathcal{CP}$ parameters in the decay modes

- $B^\pm \to D(\to K^0_S\pi^+\pi^-, K^0_SK^+K^-)K^\pm$, using the full LHCb Run 1 dataset [155], and
- $B^0 \to D(\to K^0_S\pi^+\pi^-, K^0_SK^+K^-)K^{*0}$, using the same dataset [156].

channels	parameter	$\sigma_{\text{stat}}^{\text{MD}}$	$\sigma_{\text{stat}}^{\text{MI}}$
$DK^{\pm}$	x_-	0.023	0.025
	y_-	0.028	0.034
	x_+	0.022	0.023
	y_+	0.026	0.031
$Dh^{\pm}$	x_-	0.022	0.025
	y_-	0.027	0.033
	x_+	0.021	0.023
	y_+	0.025	0.030
DK^{*0}	x_-	0.144	0.145
	y_-	0.164	0.175
	x_+	0.214	0.251
	y_+	0.234	0.251
$DK^{*\pm}$	x_-	0.144	0.196
	y_-	0.171	0.197
	x_+	0.353	0.277
	y_+	0.207	0.314
$D^*h^{\pm}$	x_-	0.055	0.060
	y_-	0.064	0.071
	x_+	0.059	0.065
	y_+	0.079	0.092

Table 8.8.: Comparison of standalone $\mathcal{CP}$ fit sensitivities using the model-dependent (MD) and model-independent (MI) approaches. All modes are included alongside their control mode, except for the first block ($DK^{\pm}$ without $D\pi^{\pm}$).

channels	parameter	$\sigma^{\text{MD}}_{\text{stat}}$	$\sigma^{\text{MI}}_{\text{stat}}$
	x_-	0.019	0.022
	y_-	0.015	0.021
$Dh^\pm$	x_+	0.018	0.023
DK^{*0}	y_+	0.018	0.029
	$\text{Re}(\xi_{DK^{*0}})$	3.224	3.250
	$\text{Im}(\xi_{DK^{*0}})$	2.908	3.229
	x_-	0.020	0.021
	y_-	0.014	0.021
	x_+	0.018	0.023
$Dh^\pm$	y_+	0.019	0.029
DK^{*0}	$\text{Re}(\xi_{DK^{*0}})$	3.211	3.201
$DK^{*\pm}$	$\text{Im}(\xi_{DK^{*0}})$	2.865	3.261
	$\text{Re}(\xi_{DK^{*\pm}})$	3.562	2.713
	$\text{Im}(\xi_{DK^{*\pm}})$	2.453	3.053
	x_-	0.019	0.021
	y_-	0.013	0.022
	x_+	0.018	0.023
$Dh^\pm$	y_+	0.018	0.029
DK^{*0}	$\text{Re}(\xi_{DK^{*0}})$	3.201	3.223
$DK^{*\pm}$	$\text{Im}(\xi_{DK^{*0}})$	2.791	3.302
$D^*h^\pm$	$\text{Re}(\xi_{DK^{*\pm}})$	3.530	2.743
	$\text{Im}(\xi_{DK^{*\pm}})$	2.420	3.051
	$\text{Re}(\xi_{D^*K^\pm})$	0.805	0.634
	$\text{Im}(\xi_{D^*K^\pm})$	0.781	0.809

Table 8.9.: Comparison of simultaneous $\mathcal{CP}$ fit sensitivities for different combinations of decay modes, using the model-dependent (MD) and model-independent (MI) approaches. All modes are included alongside their control mode.

An order of magnitude estimate for the systematic uncertainty arising in the present analysis may be found directly by taking the systematic uncertainty obtained in the existing standalone measurements of decay mode $B^{\pm} \to DK^{\pm}$. This mode is the highest statistics $\mathcal{CP}$-sensitive category within the present simultaneous fit. As such, the fitted values for the Cartesian $\mathcal{CP}$ parameters, and thus for γ, should change significantly more under systematic variations within this decay mode than they do under systematic variations within the other included modes.[2] The total systematic uncertainty should therefore be only slightly larger than that arising purely from $B^{\pm} \to DK^{\pm}$.

Taking the systematic uncertainties obtained in References [154] and [155] gives the expected orders of magnitude

$$\begin{aligned}
&\sigma_{\text{syst}}(x_-) \sim 1.0 \times 10^{-2}, \quad && \sigma_{\text{model}}(x_-) \sim 0.1 \times 10^{-2}, \\
&\sigma_{\text{syst}}(y_-) \sim 1.0 \times 10^{-2}, \quad && \sigma_{\text{model}}(y_-) \sim 0.5 \times 10^{-2}, \\
&\sigma_{\text{syst}}(x_+) \sim 1.0 \times 10^{-2}, \quad && \sigma_{\text{model}}(x_+) \sim 0.5 \times 10^{-2} \quad \text{and} \\
&\sigma_{\text{syst}}(y_+) \sim 1.0 \times 10^{-2}, \quad && \sigma_{\text{model}}(y_+) \sim 1.0 \times 10^{-2},
\end{aligned}$$

for the model-dependent simultaneous fit, and

$$\begin{aligned}
&\sigma_{\text{syst}}(x_-) \sim 1.0 \times 10^{-2}, \quad && \sigma_{\text{CLEO}}(x_-) \sim 0.5 \times 10^{-2}, \\
&\sigma_{\text{syst}}(y_-) \sim 0.5 \times 10^{-2}, \quad && \sigma_{\text{CLEO}}(y_-) \sim 1.0 \times 10^{-2}, \\
&\sigma_{\text{syst}}(x_+) \sim 1.0 \times 10^{-2}, \quad && \sigma_{\text{CLEO}}(x_+) \sim 0.5 \times 10^{-2} \quad \text{and} \\
&\sigma_{\text{syst}}(y_+) \sim 0.5 \times 10^{-2}, \quad && \sigma_{\text{CLEO}}(y_+) \sim 1.0 \times 10^{-2},
\end{aligned}$$

for the model-independent simultaneous fit. σ_{model} is the systematic uncertainty propagating from the calibration of the $D^0 \to K^0_S\pi^+\pi^-$ amplitude model by the BaBar collaboration. σ_{CLEO} is the systematic uncertainty propagating from the measurement of the bin phase parameters c_i and s_i by the CLEO collaboration. σ_{syst} includes all remaining systematics.

The actual systematic uncertainty obtained using the model-dependent approach may turn out to be lower than the above estimate, since the standalone measurement of reference [154] analysed only data obtained at LHCb during 2011, approximately one third of the full Run 1 dataset included in the present analysis. Several leading systematics are expected to reduce with increased statistics. In particular, uncertainties propagated from the results of the fits to reconstructed B-meson and D-meson mass spectra should reduce as statistics within those fits increase. Indeed, the systematic uncertainties obtained in the model-independent analysis of $B^{\pm} \to DK^{\pm}$ using the full Run 1 dataset [155] are lower than those obtained in the analysis using 2011 data only [157] by an average factor of 0.41.

[2] Of course, variations within other modes may give large systematic uncertainties on the corresponding ξ_m parameters, but these are not used in the extraction of γ.

9. Summary

This thesis describes two analyses to extract γ from B-meson decays of the form $B \to D^{(*)}h^{(*)}$ recorded at LHCb during Run 1. Direct $\mathcal{CP}$ violation, parameterized by γ, occurs between $\mathcal{CP}$-conjugate decay modes of this form. Both analyses follow the GGSZ methodology, studying the final state $D \to K^0_S\pi^+\pi^-$, which is accessible to both D^0- and $\overline{D^0}$-mesons, giving a B-meson decay amplitude consisting of two tree-level contributions. These contributions interfere in the $(K^0_S\pi^+\pi^-)$ Dalitz plane to generate an interference pattern that depends on their phase difference, which contains both weak and strong contributions. Performing a simultaneous fit to the interference patterns observed in two $\mathcal{CP}$-conjugate B-meson decay modes allows the weak and strong phases to be separated, and the γ-sensitive Cartesian parameters, $\{x_\pm, y_\pm\}$, to be extracted.

The first analysis studies the decay $B^0 \to DK^{*0}$ with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$, and its $\mathcal{CP}$ conjugate, using a model-dependent approach. This approach builds a resonance-based model to describe the amplitude for the decay $D^0 \to K^0_S\pi^+\pi^-$ in the Dalitz plane. The model used has been calibrated by the BaBar collaboration to a pure sample of $D^0 \to K^0_S\pi^+\pi^-$ decays [55]. This analysis yields the results

$$\begin{aligned} x_- &= -0.15 \pm 0.14 \pm 0.03 \pm 0.01, \\ y_- &= 0.25 \pm 0.15 \pm 0.06 \pm 0.01, \\ x_+ &= 0.05 \pm 0.24 \pm 0.04 \pm 0.01, \quad \text{and} \\ y_+ &= -0.65\ ^{+0.24}_{-0.23}\ \pm 0.08 \pm 0.01, \end{aligned}$$

where the first uncertainties are statistical, the second systematic and the third arise from the uncertainty on the $D^0 \to K^0_S\pi^+\pi^-$ amplitude model used. These results imply a value for γ of

$$\gamma = (80^{+21}_{-22})^\circ.$$

This measurement has been combined statistically with others performed at LHCb, to give a value for γ of $(74.0^{+5.0}_{-5.8})^\circ$ [2].

The second analysis combines the decay modes

- $B^+ \to DK^+$, with $D \to K^0_S\pi^+\pi^-$;
- $B^0 \to DK^{*0}$, with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$;
- $B^+ \to D^*K^+$, with $D^* \to D\gamma/\pi^0$ and $D \to K^0_S\pi^+\pi^-$; and
- $B^+ \to DK^{*+}$, with $K^{*+} \to K^0_S\pi^+$ and $D \to K^0_S\pi^+\pi^-$,

and their $\mathcal{CP}$ conjugate modes, in a single simultaneous $\mathcal{CP}$ fit. Candidate decays are selected using a combination of univariate and multivariate binary classifiers, trained using Monte Carlo simulated signal samples and combinatoric background. The reconstructed B- and D-meson mass spectra are

fitted to extract yields and PDFs for each signal and background component, which are taken as input in the final $\mathcal{CP}$ fit. Efficiency variations within the Dalitz plane are accounted for using a polynomial function whose parameters float in the $\mathcal{CP}$ fit, with Gaussian constraints applied using efficiencies obtained in MC.

The $\mathcal{CP}$ fit is performed both using the model-dependent approach, as in the first analysis, and using the model-independent approach. The model-independent approach allows the amplitude for the decay $D^0 \to K^0_S\pi^+\pi^-$ to remain unknown, and integrates the model PDF over bins within the Dalitz plane, to give a finite number of unknowns in the fit. To reduce the number of unknowns, the phase parameters of each bin, c_i and s_i, are fixed from previous CLEO measurements [56].

This analysis is still ongoing, but preliminary results yield the uncertainties

$$\begin{aligned} \sigma_{x_-} &= \pm 0.019 \pm 0.010 \pm 0.001, \\ \sigma_{y_-} &= \pm 0.013 \pm 0.010 \pm 0.005, \\ \sigma_{x_+} &= \pm 0.018 \pm 0.010 \pm 0.005, \quad \text{and} \\ \sigma_{y_+} &= \pm 0.018 \pm 0.010 \pm 0.010, \end{aligned}$$

using the model-dependent approach, where the first numbers are statistical, the second are estimated systematics arising from the experimental method used, and the third are estimated systematics arising from the calibration of the amplitude model by the BaBar collaboration. The model-independent approach yields the uncertainties

$$\begin{aligned} \sigma_{x_-} &= \pm 0.021 \pm 0.010 \pm 0.005, \\ \sigma_{y_-} &= \pm 0.022 \pm 0.005 \pm 0.010, \\ \sigma_{x_+} &= \pm 0.023 \pm 0.010 \pm 0.005, \quad \text{and} \\ \sigma_{y_+} &= \pm 0.029 \pm 0.005 \pm 0.010, \end{aligned}$$

where in this case the final uncertainties are estimated systematics arising from the measurement of the bin phase parameters, c_i and s_i, by the CLEO collaboration.

Unlike a standalone analysis, a simultaneous fit to multiple decay modes allows cross-feed of signal decays between samples to be modelled accurately in the $\mathcal{CP}$ fit. More generally, combining decay modes within a simultaneous fit naturally captures the correlations between parameters coming from different modes. These correlations are difficult to determine when standalone results are combined in a meta-analysis.

Both of these measurements rely heavily on LHCb's ability to reconstruct and identify individual mesons created during heavy-flavour decays. LHCb's RICH subdetectors provide discrimination between pion and kaon candidates, crucial to identifying the final states considered in this thesis, by measuring the Cherenkov 'rings' emitted as charged mesons traverse a radiator medium. In preparation for the high luminosity era, due to commence in 2021, the RICH photon detector planes and readout electronics will be upgraded. Beam tests have been performed using the candidate upgrade photon detectors and full readout chain.

A. The relative weak phase for $B \to D^{(*)}K^{(*)}$

We wish to find the weak phase of the process

$$B \to \overline{D^{(*)0}}K^{(*)}, \tag{A.1}$$

relative to the process

$$B^0 \to D^{(*)0}K^{(*)}. \tag{A.2}$$

Inspection of the relevant Feynman diagrams reveals that the CKM matrix elements involved are V_{ub} and V_{cs} for the $D^{(*)0}$ case, whilst for the $\overline{D^{(*)0}}$ case they are V_{cb} and V_{us}. The relative weak phase between the two is thus

$$\phi = \arg\left(\frac{V_{cb}V_{us}}{V_{ub}V_{cs}}\right). \tag{A.3}$$

It is useful here to note the definition of γ:

$$\gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right).$$

To bring ϕ into a form that can be compared more directly with γ, two relations that result from the theory of complex numbers are useful:

$$\arg(z^*) = -\arg(z),$$

and

$$\arg(z_1 z_2) = \arg(z_1) + \arg(z_2).$$

These allow us to write

$$\phi = \arg\left(\frac{V_{us}V_{cs}^*}{V_{ub}V_{cb}^*}\right). \tag{A.4}$$

Using the unitarity relation, $V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0$, to replace the numerator of the fraction, we obtain

$$\phi = \arg\left(-1 - \frac{V_{ud}V_{cd}^*}{V_{ub}V_{cb}^*}\right). \tag{A.5}$$

The final relation to be used is $\arg(z) = \tan^{-1}(\mathrm{Im}(z)/\mathrm{Re}(z))$, from which follows

$$\arg(z_1 + z_2) = \tan^{-1}\left(\frac{\mathrm{Im}(z_1 + z_2)}{\mathrm{Re}(z_1 + z_2)}\right) \tag{A.6}$$

$$= \tan^{-1}\left(\frac{|z_1|\sin(\theta_1) + |z_2|\sin(\theta_2)}{|z_1|\cos(\theta_1) + |z_2|\cos(\theta_2)}\right), \tag{A.7}$$

where $|z_i|$ is the modulus of complex number z_i, and θ_i is its argument. From inspection of equation (A.5), it is clear that ϕ takes the form $\arg(z_1 + z_2)$, with $|z_1| = 1$, $\theta_1 = \arg(-1) = \pi$, and $\theta_2 = \gamma$. Substituting these into equation (A.7), we find

$$\phi = \tan^{-1}\left(\frac{|z_2|\sin(\gamma)}{-1 + |z_2|\cos(\gamma)}\right). \tag{A.8}$$

$|z_2|$ involves just the moduli of the relevant CKM matrix elements, and can be found from the numerical values presented in [3]. Using these, one obtains $|z_2| \approx 1400$. Thus $\phi \approx \gamma$, so long as $|\cos(\gamma)| \gg 1/1400$.

B. Useful relations of complex numbers

A useful expression for the square-modulus of the sum of two complex numbers can be derived as follows. We begin with the straightforward relation

$$|z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2(z_1.z_2). \tag{B.1}$$

The term in parentheses may be readily re-expressed:

$$\begin{aligned} z_1.z_2 &= \mathrm{Re}(z_1)\mathrm{Re}(z_2) + \mathrm{Im}(z_1)\mathrm{Im}(z_2) && \text{(B.2)} \\ &= |z_1||z_2|\cos(\theta_1)\cos(\theta_2) + |z_1||z_2|\sin(\theta_1)\sin(\theta_2), && \text{(B.3)} \end{aligned}$$

where $\theta_i \equiv \arg(z_i)$. Using the trigonometric identity for $\cos(A \pm B)$, one finds

$$z_1.z_2 = |z_1||z_2|\cos(\theta_1 - \theta_2), \tag{B.4}$$

which finally gives

$$\begin{aligned} |z_1 + z_2|^2 &= |z_1|^2 + |z_2|^2 + 2|z_1||z_2|\cos(\theta_1 - \theta_2), && \text{(B.5)} \\ &= |z_1|^2 + |z_2|^2 + 2\mathrm{Re}(z_1 z_2^*). && \text{(B.6)} \end{aligned}$$

C. Helicity eigenstates for a spin-1 resonance

The aim of this appendix is to determine the value of the 'spin-sum' for a spin-1 resonance:

$$\sum_{\lambda} \epsilon_{\lambda}^{\mu} \epsilon_{\lambda}^{*\nu}, \tag{C.1}$$

where ϵ_{λ}^{μ} are the relativistic helicity[1] eigenvectors for a spin-1 particle, and λ gives the helicity eigenvalue:

$$\lambda = 0, \pm 1. \tag{C.2}$$

As such, only the properties of the helicity eigenstates that are relevant to the evaluation of this term are discussed. A comprehensive discussion of the treatment of helicity and spin within particle physics may be found in the very useful (and seemingly little-known) review reference [158].

In Section 2.2, two conditions on the states ϵ_{λ}^{μ} were introduced:

1. ϵ_{λ}^{μ} transform as Lorentz vectors.
2. $P_{\mu r} \epsilon_{\lambda}^{\mu} = 0$ ('transversality').

The first condition is essential, since it ensures that any amplitude constructed using ϵ_{λ}^{μ} will be a Lorentz scalar, as it must be.[2] The second condition was introduced to reduce the number of degrees of freedom of ϵ_{λ}^{μ} from four (the number of elements of a Lorentz vector) to three (the number of coefficients needed to fully specify a general superposition of three helicity eigenstates). These two conditions are enough to derive the specific form taken by ϵ_{λ}^{μ}, which in turn allows the evaluation of the 'spin-sum'. The derivation begins by formulating a set of spin eigenstates for a particle at rest, before applying a Lorentz transformation to find the general form for the helicity eigenstates, which applies for all momenta.

For a particle at rest,

$$P^{\mu} = \{m, \vec{0}\}. \tag{C.3}$$

Transversality thus implies that

$$\epsilon_{\lambda}^{\mu}(p = 0) = \{0, \vec{\epsilon}_{\lambda}\}. \tag{C.4}$$

for all λ.

[1] Helicity is the quantum number which corresponds to the component of a particle's spin which is measured parallel to its momentum.

[2] Otherwise, the cross-section for an interaction would depend on the reference frame from which it was viewed.

It is well known[3] that the angular momentum eigenstates for a spin-1 particle at rest are given by:

$$\vec{e}_{(m=\pm 1)} = \mp \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \\ 0 \end{pmatrix}, \quad \vec{e}_{(m=0)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{C.5}$$

where m is the quantum number corresponding to the operator J_z; which measures the z-component of the particle's angular momentum.

If we choose to set $\vec{\epsilon}_\lambda = \vec{e}_m$ for $\lambda = m$, this yields expressions for the Lorentz eigenvectors of the z-component of spin for a particle at rest. However, we want to obtain the *helicity* eigenstates for *general* momentum values. This can be achieved by applying the transformation:

$$\epsilon^\mu_\lambda(\vec{p}) = \big(R\Lambda_z(p)\big)^\mu_\nu \epsilon^\mu_\lambda(0), \tag{C.6}$$

where $\Lambda_z(p)$ represents a Lorentz boost along the z-axis, and R is the rotation which maps the z-axis onto the direction $\hat{\vec{p}}$. For the rest of this appendix, in the interests of simplicity, we let the momentum $\vec{p}$ define the z-axis; in this case, R reduces to the identity matrix.

Explicitly then, the three z-spin eigenstates for a stationary particle are given by

$$\epsilon^\mu_{\pm 1}(0) = \mp \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ \pm i \\ 0 \end{pmatrix}, \quad \epsilon^\mu_0(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{C.7}$$

The Lorentz boost that must be applied may be expressed as

$$\Lambda_z(p) = \begin{pmatrix} \gamma & & & \beta\gamma \\ & 1 & & \\ & & 1 & \\ \beta\gamma & & & \gamma \end{pmatrix}. \tag{C.8}$$

Since the boost is along the z-axis, it has no effect on the x and y components of the vectors it acts on, and so

$$\epsilon^\mu_{\pm 1}(p\hat{\vec{z}}) = \epsilon^\mu_{\pm 1}(0). \tag{C.9}$$

[3] See any introductory quantum mechanics text, e.g. Binney [159].

Transforming the $(\lambda = 0)$ eigenstate,

$$\epsilon_0^\mu(p\hat{z}) = \begin{pmatrix} \beta\gamma \\ 0 \\ 0 \\ \gamma \end{pmatrix} = \begin{pmatrix} p/m \\ 0 \\ 0 \\ E/m \end{pmatrix}. \tag{C.10}$$

Now that the helicity eigenstates have been determined explicitly, the 'sum-rule' may be evaluated;

$$\sum_\lambda \epsilon_\lambda^\mu \epsilon_\lambda^{*\nu} = \epsilon_{-1}^\mu(p)\epsilon_{-1}^{*\nu}(p) + \epsilon_1^\mu(p)\epsilon_1^{*\nu}(p) \tag{C.11}$$

$$+ \epsilon_0^\mu(p)\epsilon_0^{*\nu}(p). \tag{C.12}$$

The first two terms give the contributions

$$\epsilon_{-1}^\mu(p)\epsilon_{-1}^{*\nu}(p) = \frac{1}{2}\begin{pmatrix} 0 & & & \\ & 1 & -i & \\ & -i & 1 & \\ & & & 0 \end{pmatrix}, \tag{C.13}$$

$$\epsilon_1^\mu(p)\epsilon_1^{*\nu}(p) = \frac{1}{2}\begin{pmatrix} 0 & & & \\ & 1 & +i & \\ & +i & 1 & \\ & & & 0 \end{pmatrix}, \tag{C.14}$$

and the third term gives

$$\epsilon_0^\mu(p)\epsilon_0^{*\nu}(p) = \begin{pmatrix} p^2/m^2 & & pE/m^2 \\ & & \\ pE/m^2 & & E^2/m^2 \end{pmatrix} = \begin{pmatrix} (E^2/m^2 - 1) & & pE/m^2 \\ & & \\ pE/m^2 & & (p^2/m^2 + 1) \end{pmatrix}, \tag{C.15}$$

where the last expression follows from noting that $E^2 - p^2 = m^2$.

The spin-sum may thus be expressed as

$$\sum_\lambda \epsilon_\lambda^\mu \epsilon_\lambda^{*\nu} = \begin{pmatrix} E^2/m^2 & & pE/m^2 \\ & & \\ pE/m^2 & & p^2/m^2 \end{pmatrix} + \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \\ & & & 1 \end{pmatrix}. \tag{C.16}$$

The second term may be straightforwardly identified as $-g^{\mu\nu}$.

Noting that the particle momentum was chosen to define the z-axis, $\vec{p} = \hat{p}\hat{z}$, and thus $P^\mu = \{E, 0, 0, p\}$. Inspection of the first term therefore reveals that it may be written as $P^\mu P^\nu / m^2$, which generalises for arbitrary momentum. Finally, the 'spin-sum rule' is obtained

$$\sum_\lambda \epsilon^\mu_\lambda \epsilon^{*\nu}_\lambda = \frac{P^\mu P^\nu}{m^2} - g^{\mu\nu}. \tag{C.17}$$

D. Input variables used to train multivariate classifiers

MVA input variable distributions for signal and background training samples are shown in Figures D.1–D.8. Corresponding correlation matrices are shown in Figures D.9–D.16.

Figure D.1.: Distributions of input variables for signal and background samples, used to train MVA's to select LL $B^{\pm} \to Dh^{\pm}$ candidate events.

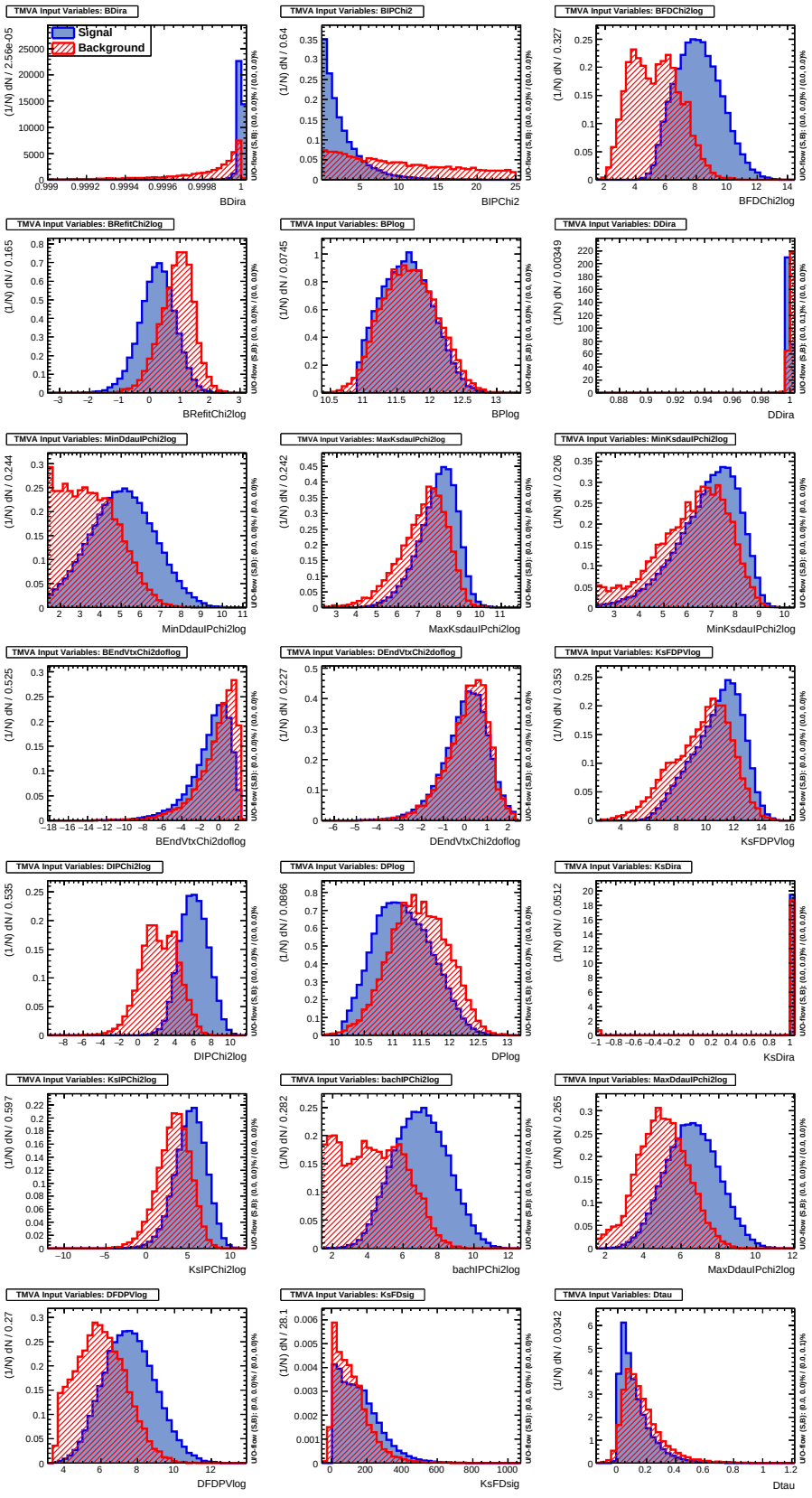


Figure D.2.: Distributions of input variables for signal and background samples, used to train MVA's to select DD $B^{\pm} \to Dh^{\pm}$ candidate events.

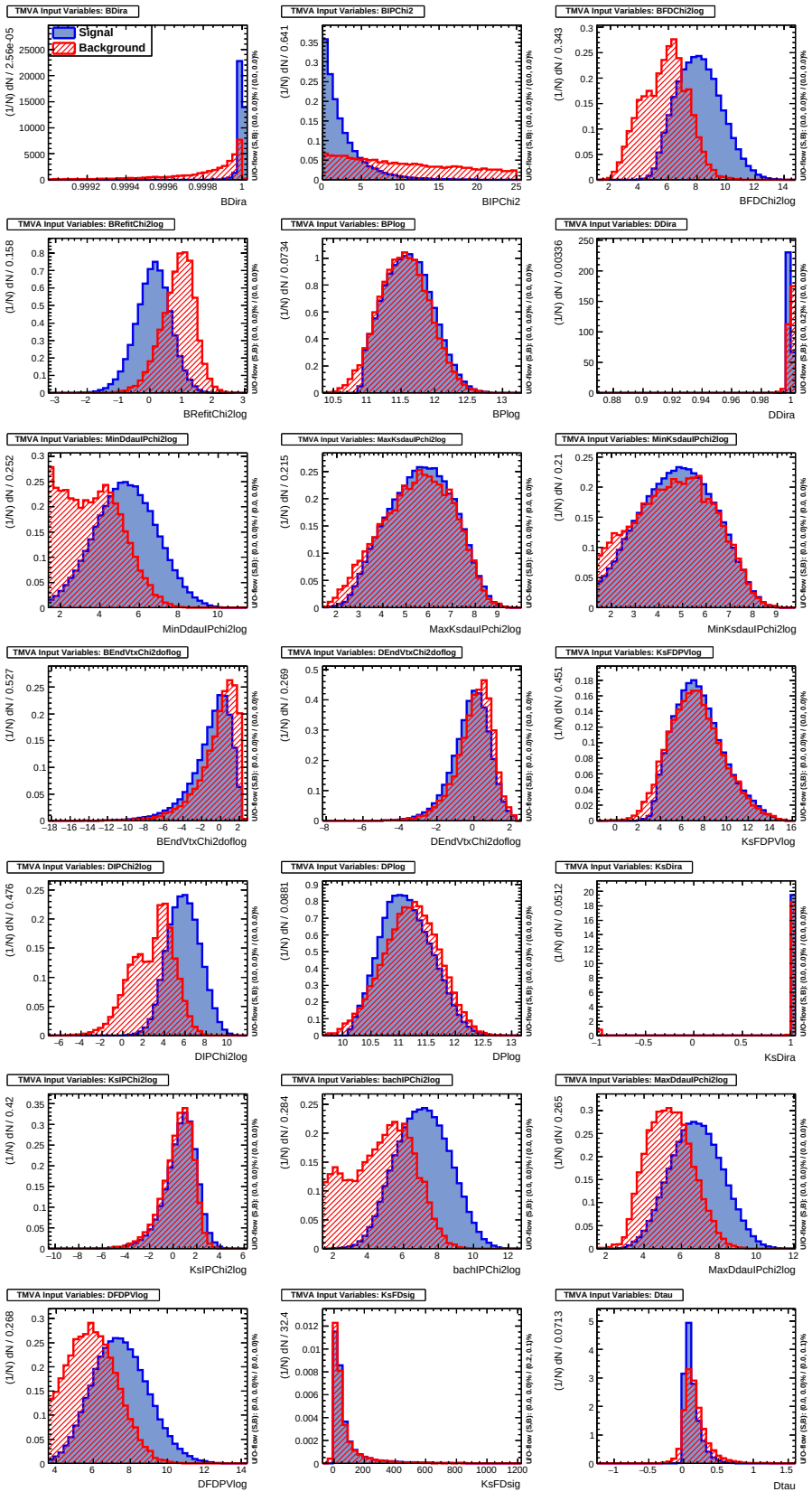


Figure D.3.: Distributions of input variables for signal and background samples, used to train MVA's to select LL $B^0 \to DK^{*0}$ candidate events.

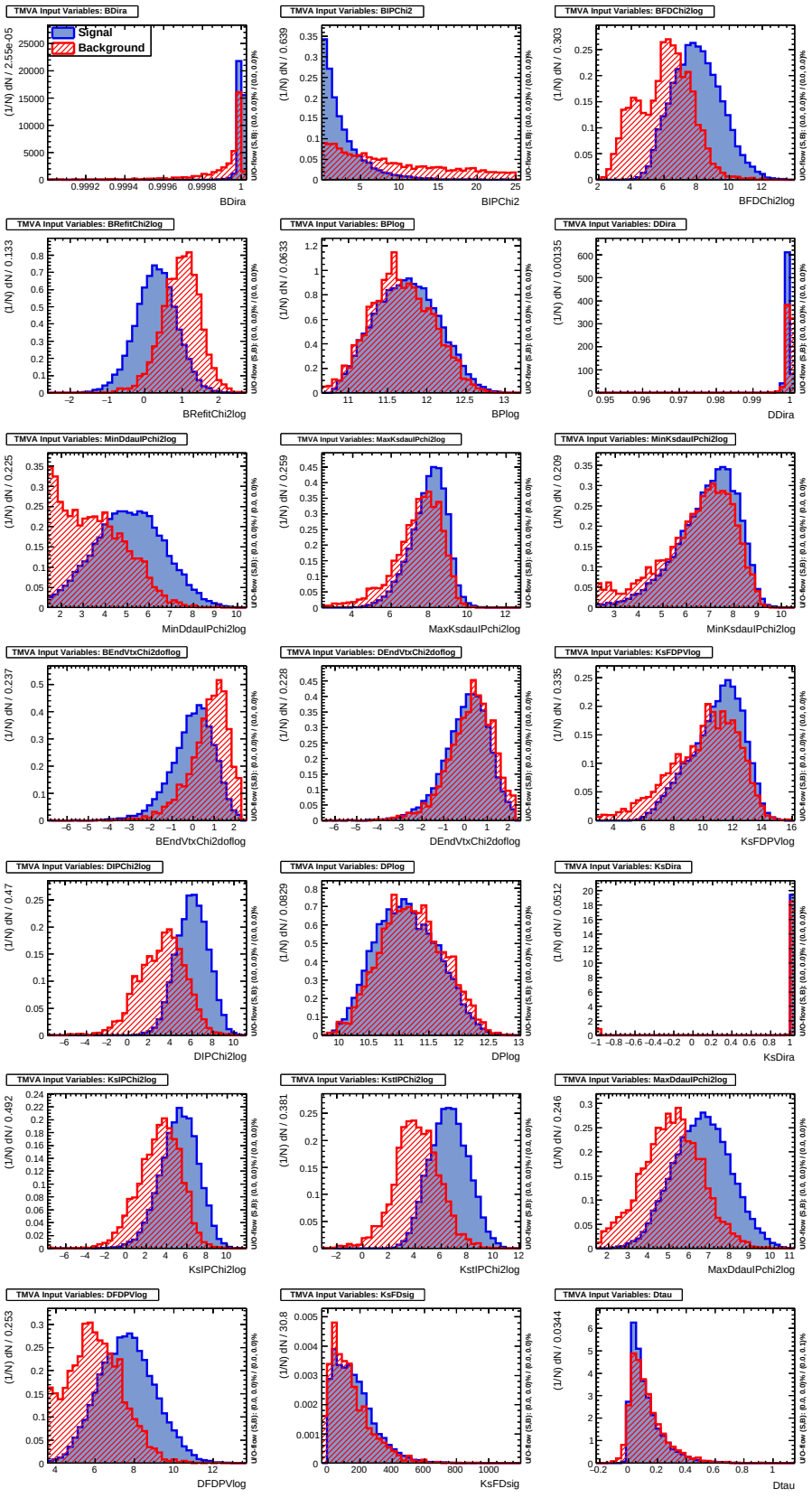


Figure D.4.: Distributions of input variables for signal and background samples, used to train MVA's to select DD $B^0 \to DK^{*0}$ candidate events.

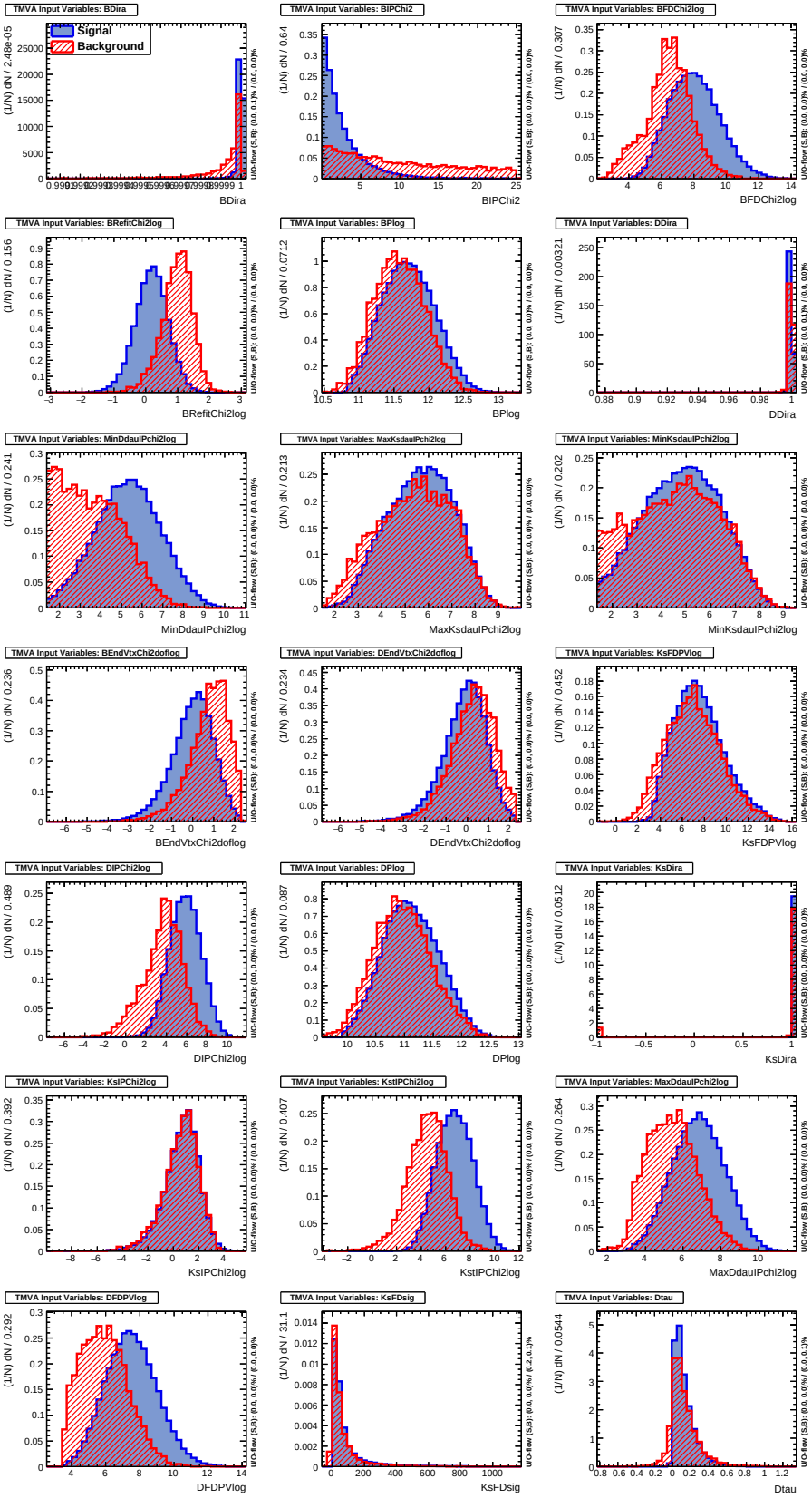


Figure D.5.: Distributions of input variables for signal and background samples, used to train MVA's to select LL $B^{\pm} \to DK^{*\pm}$ candidate events.

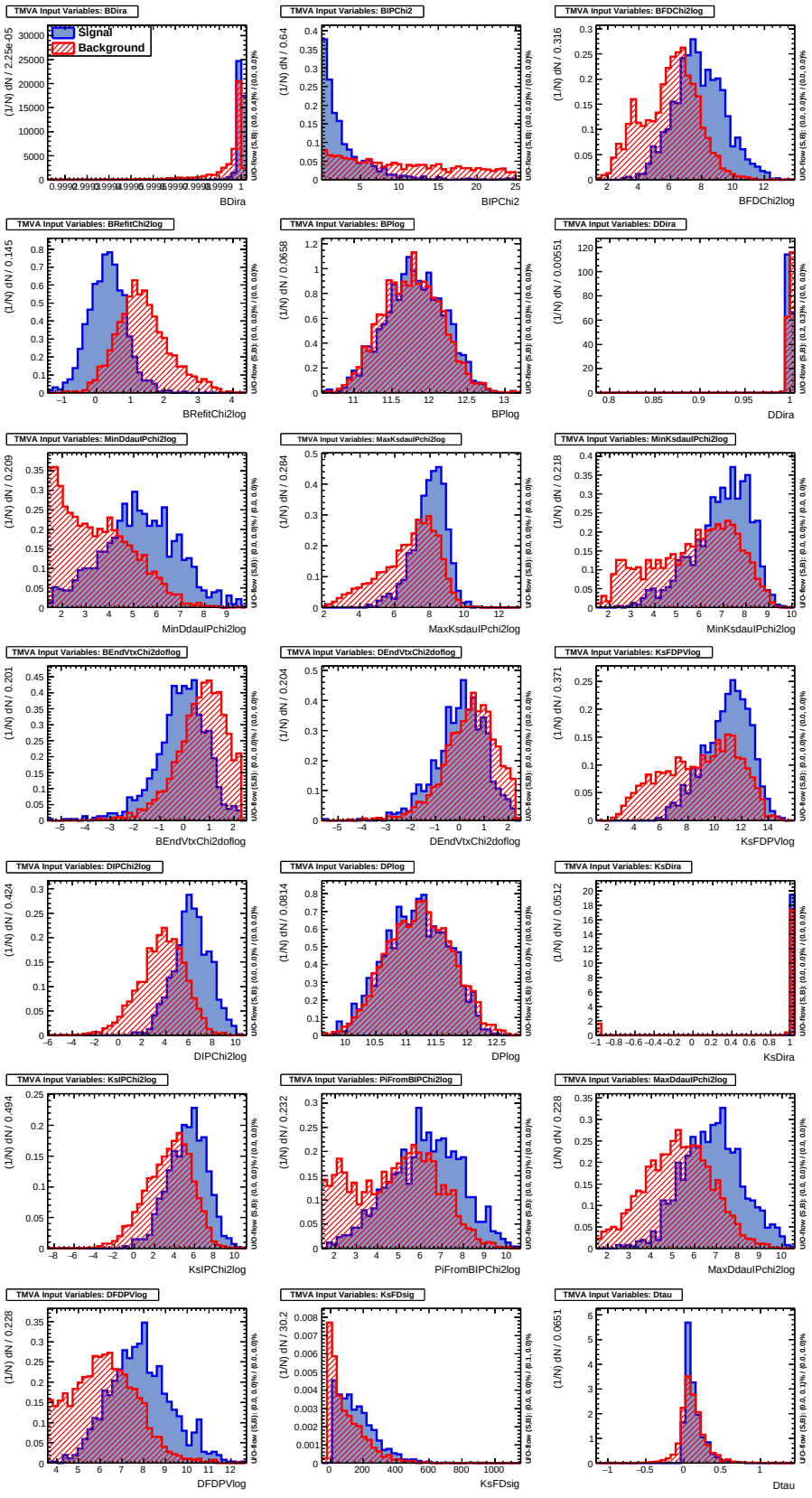


Figure D.6.: Distributions of input variables for signal and background samples, used to train MVA's to select DD $B^{\pm} \to DK^{*\pm}$ candidate events.

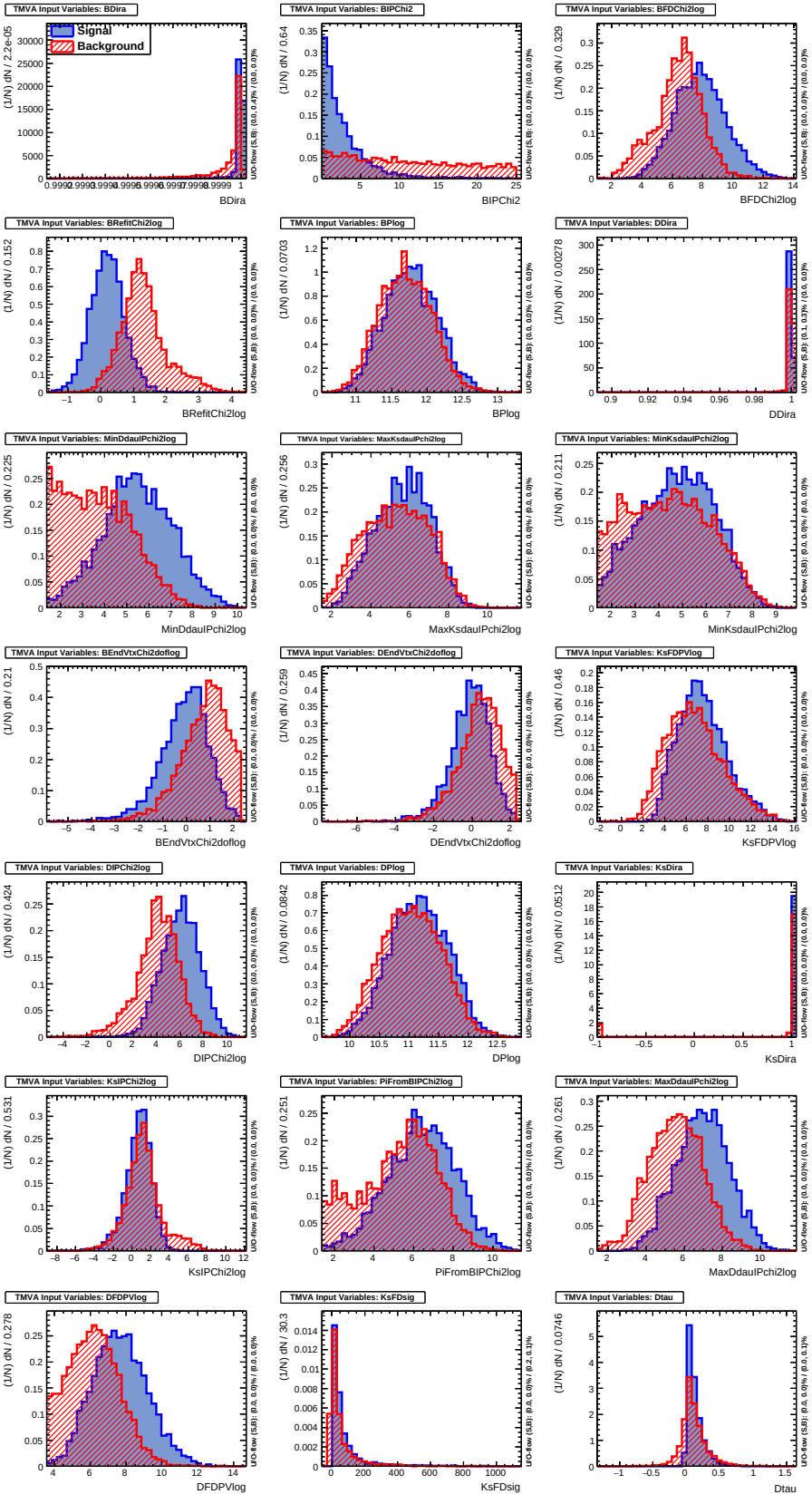


Figure D.7.: Distributions of input variables for signal and background samples, used to train MVA's to select LL $B^{\pm} \to D^{*}h^{\pm}$ candidate events.

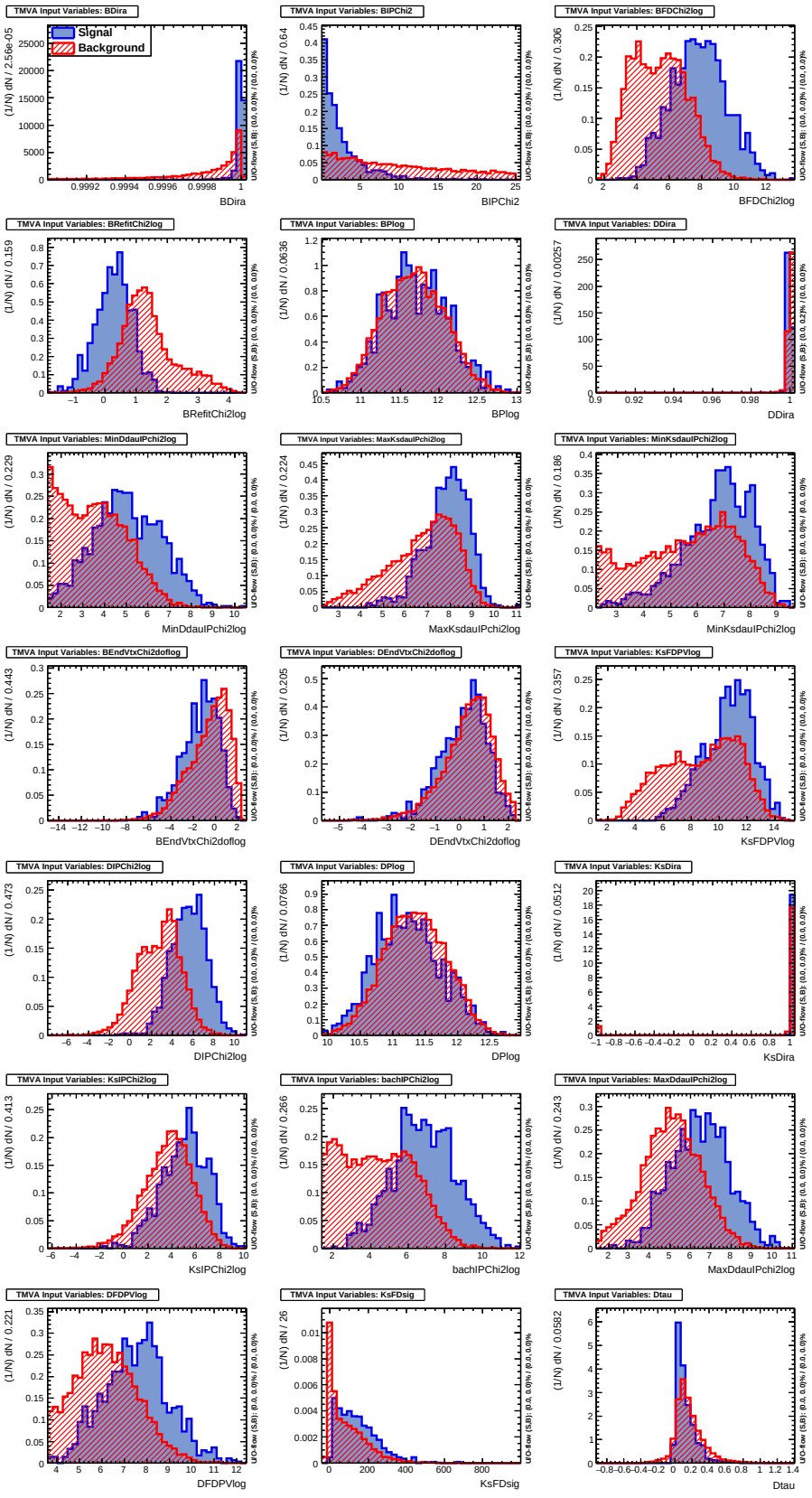
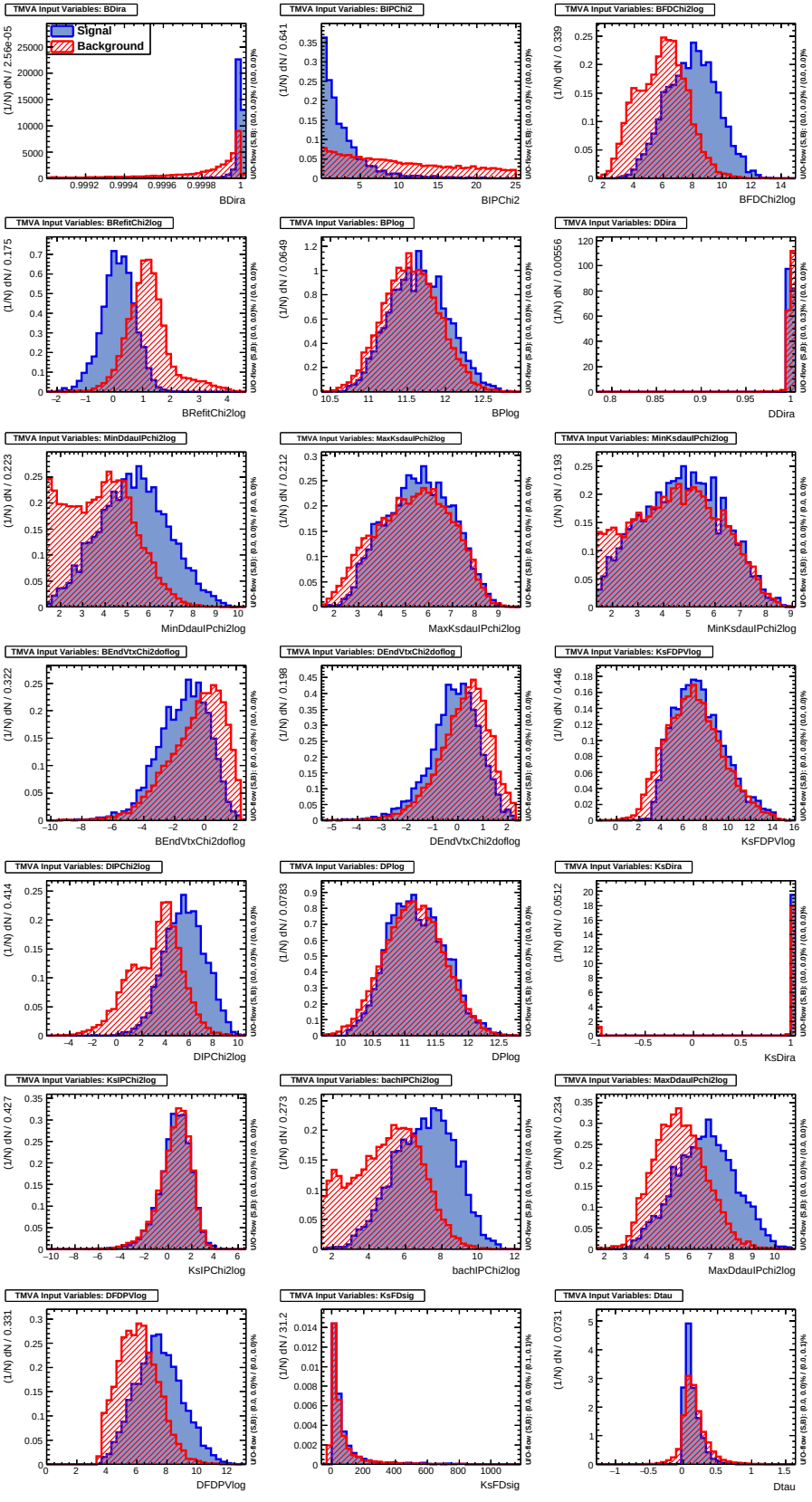


Figure D.8.: Distributions of input variables for signal and background samples, used to train MVA's to select DD $B^{\pm} \to D^{*}h^{\pm}$ candidate events.



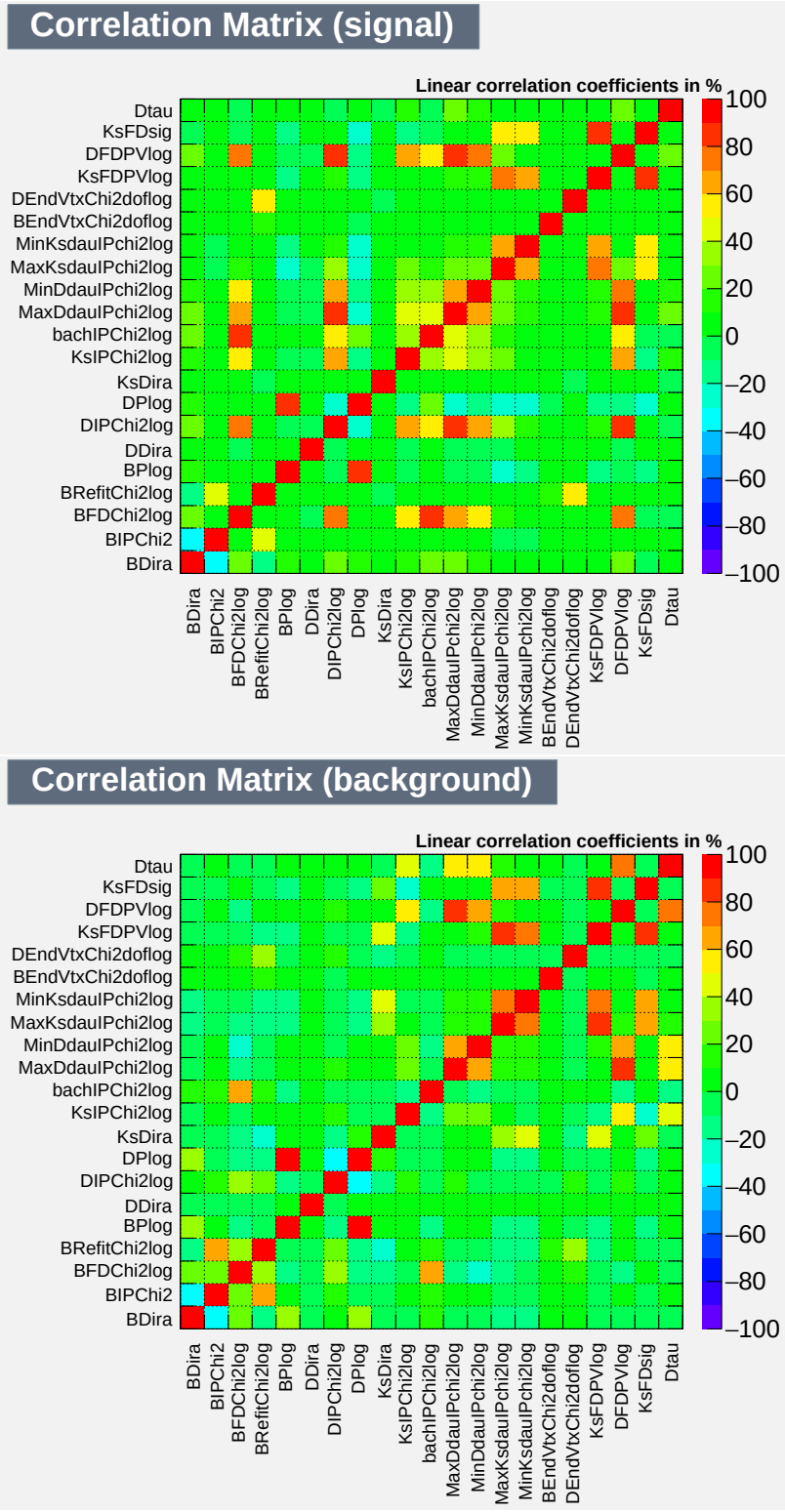


Figure D.9.: Correlation matrices for training variables used in MVA's to select LL $B^{\pm} \to Dh^{\pm}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

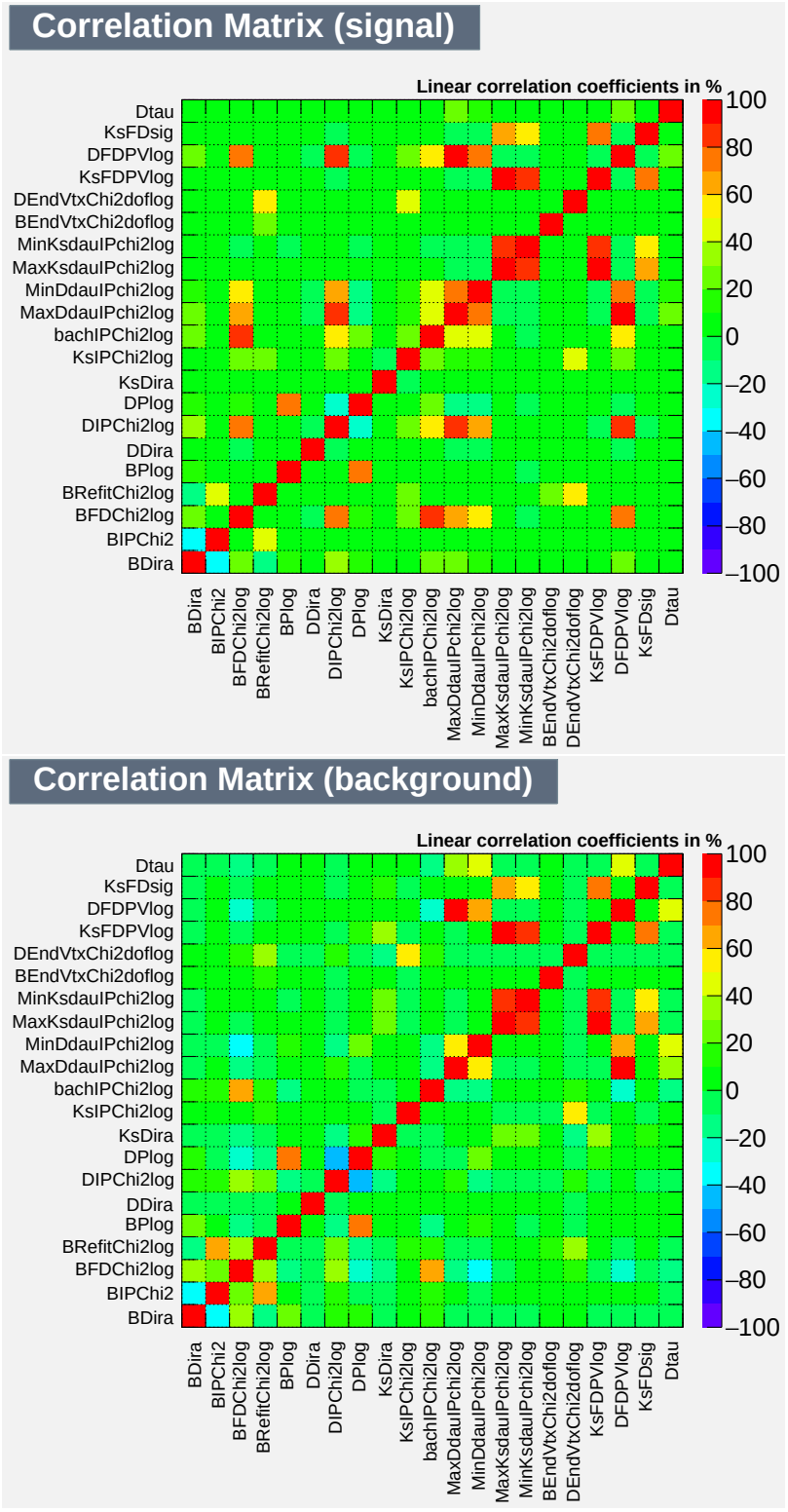


Figure D.10.: Correlation matrices for training variables used in MVA's to select DD $B^{\pm} \to Dh^{\pm}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

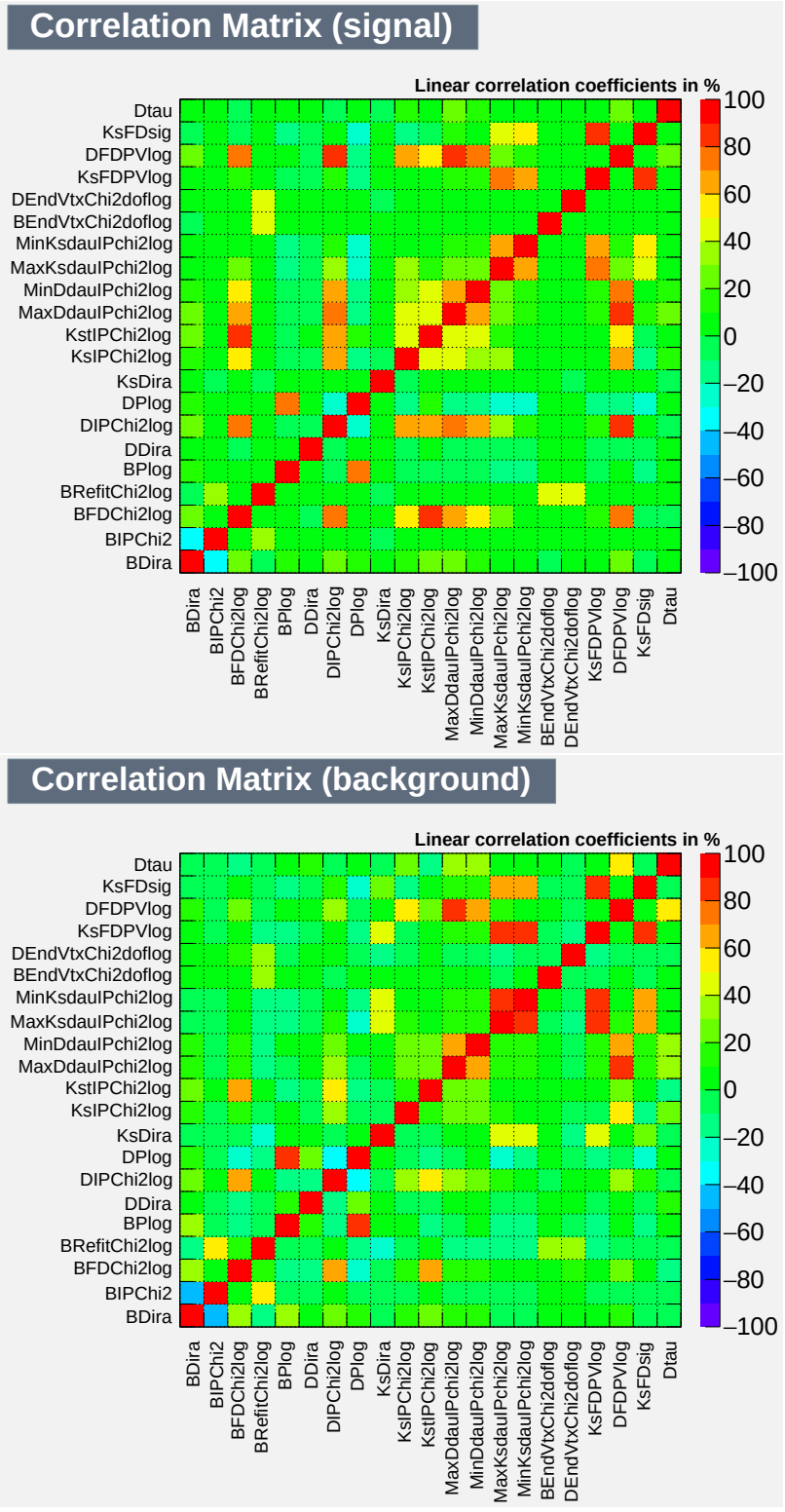


Figure D.11.: Correlation matrices for training variables used in MVA's to select LL $B^0 \to DK^{*0}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

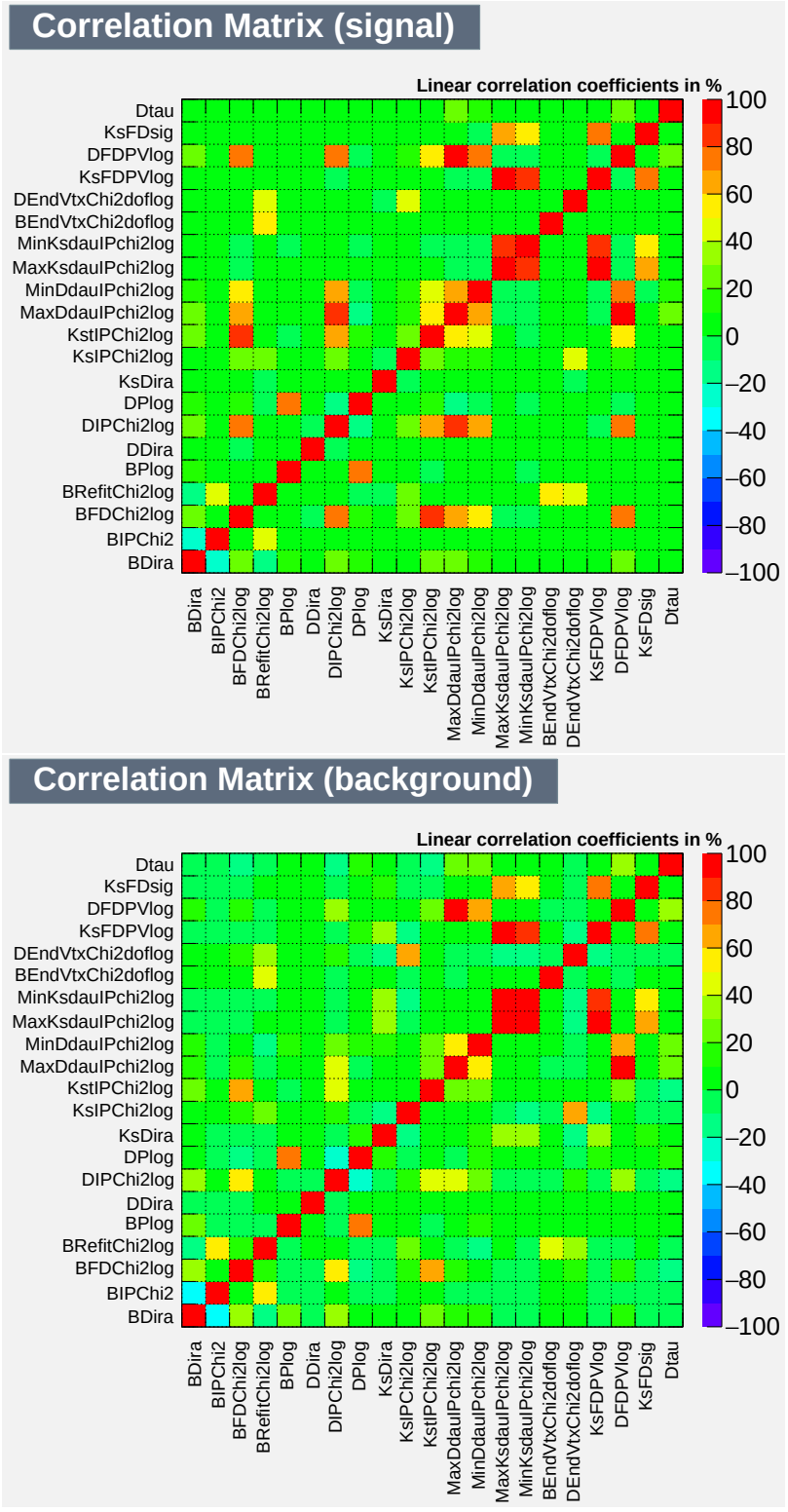


Figure D.12.: Correlation matrices for training variables used in MVA's to select DD $B^0 \to DK^{*0}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

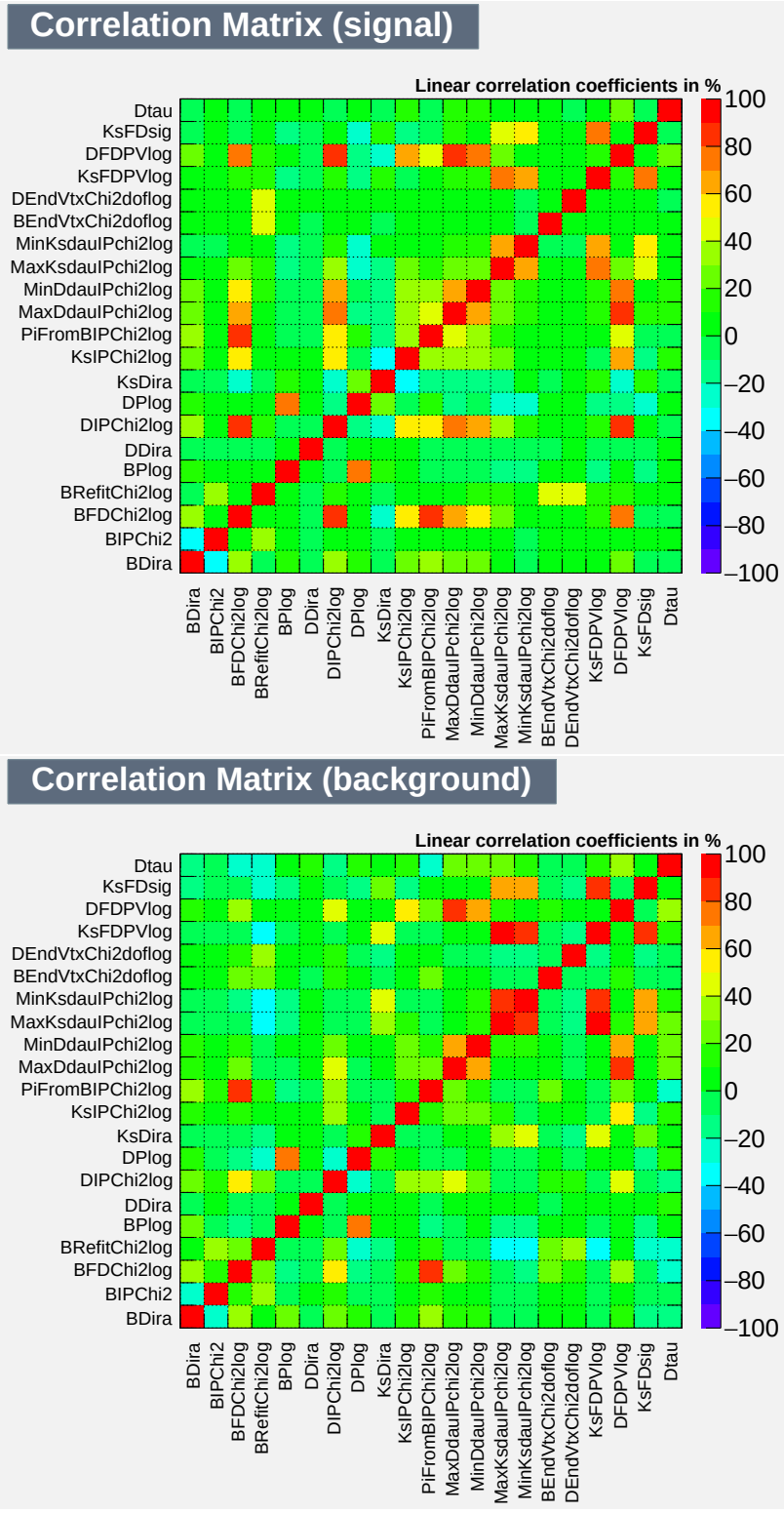


Figure D.13.: Correlation matrices for training variables used in MVA's to select LL $B^{\pm} \to DK^{*\pm}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

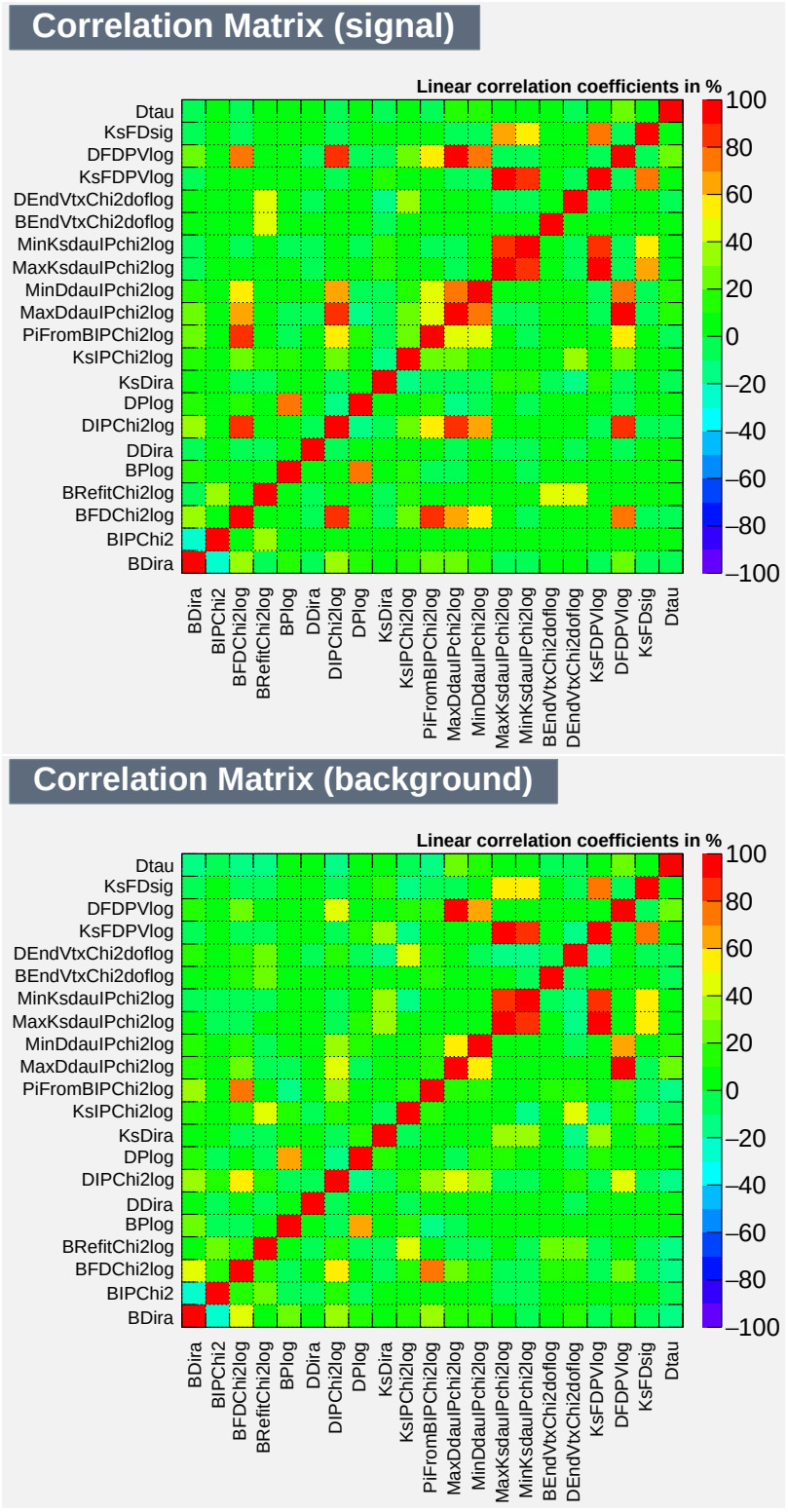


Figure D.14.: Correlation matrices for training variables used in MVA's to select DD $B^{\pm} \to DK^{*\pm}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

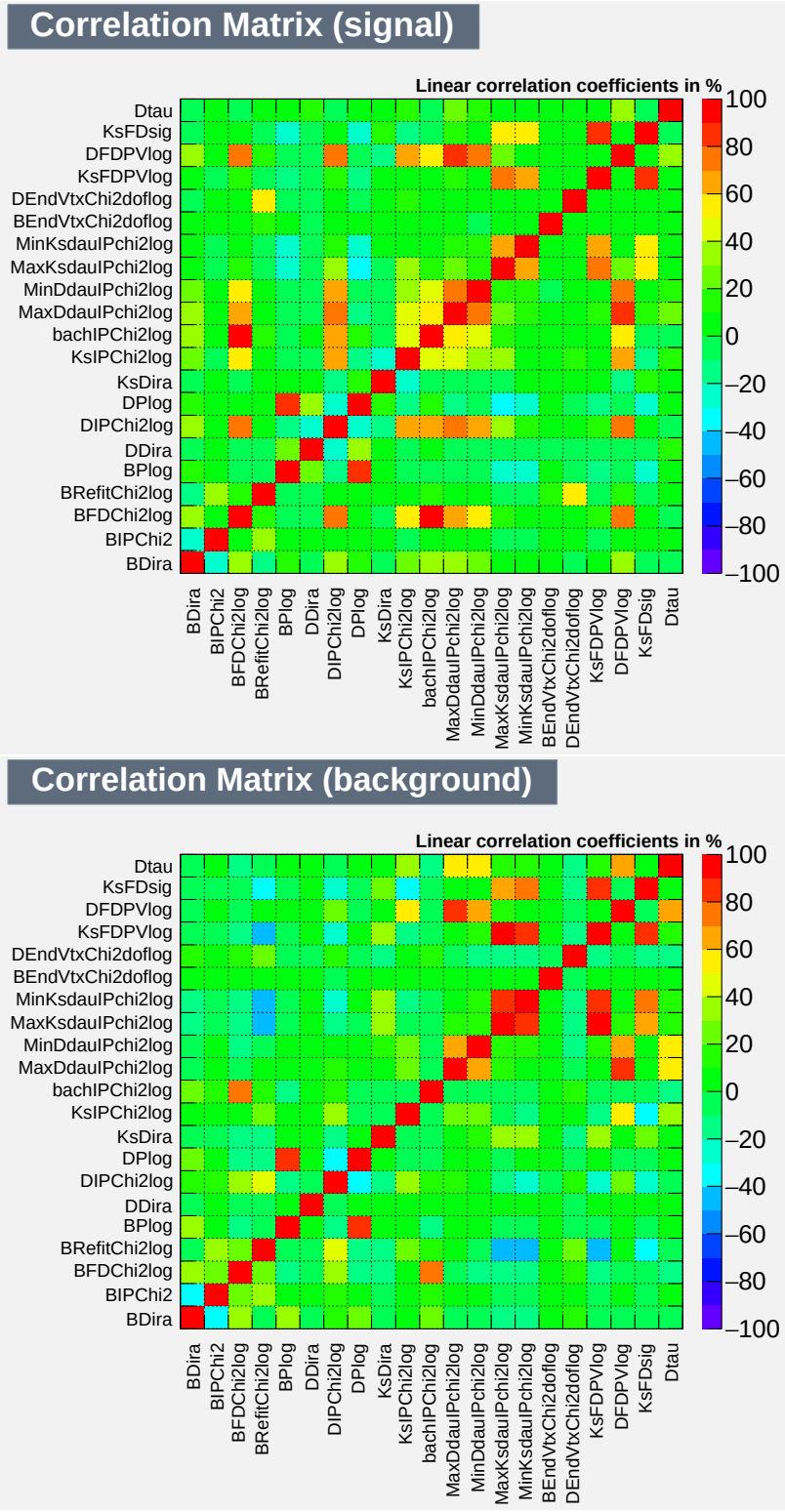


Figure D.15.: Correlation matrices for training variables used in MVA's to select LL $B^{\pm} \to D^{*}h^{\pm}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

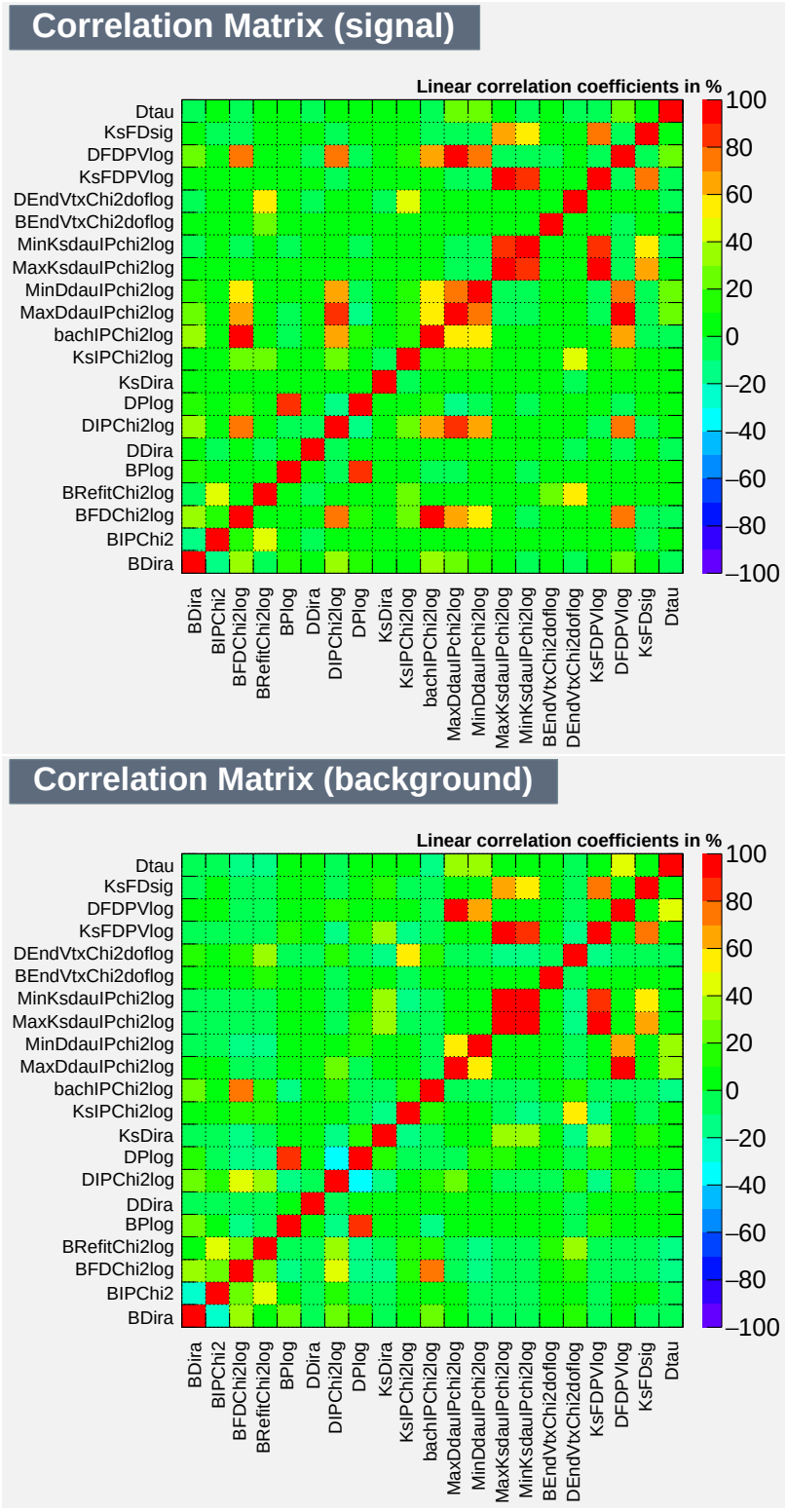


Figure D.16.: Correlation matrices for training variables used in MVA's to select DD $B^{\pm} \to D^{*}h^{\pm}$ candidate events. Correlations are shown for the signal (top) and background (bottom) training samples used.

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ using $B^0 \to DK^{*0}$ with $D \to K^0_S\pi^+\pi^-$ decays*, JHEP **08** (2016) 137, arXiv:1605.0108.

[2] LHCb Collaboration, *Update of the LHCb combination of the CKM angle γ*, Tech. Rep. LHCb-CONF-2018-002. CERN-LHCb-CONF-2018-002, CERN, Geneva, May, 2018. LHCb-CONF-2018-002.

[3] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. D **98** (2018) 030001.

[4] G. Gabrielse *et al.*, *New determination of the fine structure constant from the electron g value and QED*, Phys. Rev. Lett. **97** (2006) 030802.

[5] S. P. Martin, *A Supersymmetry primer*, arXiv:hep-ph/9709356, 1998.

[6] J. Bedford, *An Introduction to String Theory*, arXiv:1107.3967, 2011.

[7] M. Li, *Introduction to M theory*, in *Proceedings, Duality Superstring and M Theory: CCAST-WL workshop, Beijing, China*, 1998. arXiv:hep-th/9811019.

[8] M. K. Baszczyk *et al.*, *Test of the photon detection system for the LHCb RICH upgrade in a charged particle beam*, JINST **12** (2017), no. 01 P01012, arxiv:1610.02879.

[9] S. L. Glashow, *Partial-symmetries of weak interactions*, Selected Papers on Gauge Theory of Weak and Electromagnetic Interactions (1981) 171.

[10] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[11] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[12] *Standard model of elementary particles*, https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. Accessed: 12-08-2016.

[13] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[14] J. F. Donoghue, E. Golowich, and B. R. Holstein, *Dynamics of the standard model*, Cambridge University Press, 2014. doi: 10.1017/CBO9780511803512.

[15] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2 π decay of the K^0_2 meson*, Phys. Rev. Lett. **13** (1964) 138.

[16] J. Schwinger, *The theory of quantized fields. II*, Phys. Rev. **91** (1953) 713.

[17] J. Schwinger, *The theory of quantized fields. VI*, Phys. Rev. **94** (1954) 1362.

[18] G. Lüders, *Proof of the TCP theorem*, Annals of Physics **2** (1957), no. 1 1.

[19] I. I. Bigi and A. I. Sanda, *CP Violation*, Cambridge University Press, second ed., 2009. Cambridge Books Online.

[20] G. C. Branco, L. Lavoura, and J. P. Silva, *CP violation*, Oxford University Press, 1999.

[21] J. P. Silva, *Phenomenological aspects of CP violation*, in *Central European School in Particle Physics Prague, Czech Republic*, 2004. arXiv:hep-ph/0410351.

[22] P. K. Kabir, *The CP puzzle: strange decays of the neutral kaon*, Academic Press, 1968.

[23] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0 - \bar{B}^0$ oscillation frequency δm_d with the decays $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi K^{*0}$*, Physics Letters B **719** (2013), no. 4 318.

[24] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency Δm_s in $B_s^0 \to D_s^-(3)\pi$ decays*, Phys. Lett. **B709** (2012) 177, arXiv:1112.4311.

[25] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the B_s^0–oscillation frequency with the decay $B_s^0 \to D^- \pi^+$*, New Journal of Physics **15** (2013), no. 5 053021.

[26] Heavy Flavour Averaging Group, *Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012*, arXiv:1207.1158.

[27] LHCb collaboration, R. Aaij *et al.*, *Observation of D^0–$\overline{D^0}$ oscillations*, Physical Review Letters **110** (2013), no. 10 101802.

[28] LHCb collaboration, R. Aaij *et al.*, *Measurement of D^0–$\overline{D^0}$ mixing parameters and search for CP violation using $D \to K^+\pi^-$ decays*, Physical Review Letters **111** (2013), no. 25 251801.

[29] V. Gibson, *B physics*, in *Large Hadron Collider Phenomenology* (M. Krämer and F. J. P. Soler, eds.), p. 373. CRC Press, 2004.

[30] M. Bardon, K. Lande, L. Lederman, and W. Chinowsky, *Long-lived neutral K mesons*, Annals of physics **5** (1958), no. 2 156.

[31] R. H. Good *et al.*, *Regeneration of neutral K mesons and their mass difference*, Phys. Rev. **124** (1961) 1223.

[32] L. B. Leipuner *et al.*, *Anomalous regeneration of K_1^0 mesons from K_2^0 mesons*, Phys. Rev. **132** (1963) 2285.

[33] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, *The standard model of particle physics*, Rev. Mod. Phys. **71** (1999) S96.

[34] G. Altarelli, *The standard model of particle physics*, Tech. Rep. 206, CERN, Department of Physics, Theory Division, 2005.

[35] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Progress of Theoretical Physics **49** (1973), no. 2 652.

[36] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[37] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802.

[38] T. Kibble and F. Berkshire, *Classical mechanics*, World Scientific, 2004. doi: 10.1142/p310.

[39] H. Fritzsch, *Calculating the Cabibbo angle*, Physics Letters B **70** (1977), no. 4 436.

[40] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[41] 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.

[42] S. C. Haines and V. Gibson, *A study of charged $B \to DK$ and $B \to D\pi$ decays with the LHCb experiment*, PhD thesis, Cambridge University, Dec, 2011, Presented 28 Feb 2012.

[43] A. Bharucha *et al.*, *Implications of LHCb measurements and future prospects*, The European Physical Journal C **73** (2013), no. 4 1.

[44] UTfit collaboration, *The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of $\Delta m_{(s)}$ and BR($B \to \tau\nu(\tau)$)*, JHEP **0610** (2006) 081, arXiv:hep-ph/0606167, updated at http://www.utfit.org/UTfit/.

[45] CKMfitter Group, *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.

[46] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.

[47] M. Gronau and D. London, *How to determine all the angles of the unitarity triangle from $B^0_d \to DK^0_S$ and $B^0_s \to D\varphi$*, Physics Letters B **253** (1991), no. 3 483.

[48] M. Gronau and D. Wyler, *On determining a weak phase from charged B decay asymmetries*, Physics Letters B **265** (1991), no. 1 172.

[49] M. Gronau and D. London, *New physics in CP asymmetries and rare B decays*, Phys. Rev. D **55** (1997) 2845.

[50] D. Atwood, I. Dunietz, and A. Soni, *Enhanced CP violation with $B^\pm \to K^\pm D$ modes and extraction of the Cabibbo-Kobayashi-Maskawa angle γ*, Phys. Rev. Lett. **78** (1997) 3257.

[51] D. Atwood, I. Dunietz, and A. Soni, *Improved methods for observing CP violation in $B^\pm \to K^\pm D$ and measuring the CKM phase γ*, Phys. Rev. D **63** (2001) 036005.

[52] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, *Determining γ using $B^\pm \to DK^\pm$ with multibody D decays*, Phys. Rev. D **68** (2003) 054018.

[53] Garra-Ticó, Jordi. personal communication.

[54] R. H. Dalitz, *On the analysis of τ-meson data and the nature of the τ-meson*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **44** (1953), no. 357 1068.

[55] BaBar collaboration, *Measurement of $D^0 - \overline{D^0}$ mixing parameters using $D^0 \to K^0_S\pi^+\pi^-$ and $D^0 \to K^0_S K^+K^-$ decays*, Phys. Rev. Lett. **105** (2010) 081803, arXiv:1004.5053.

[56] CLEO Collaboration, *Model-independent determination of the strong-phase difference between D^0 and $\overline{D^0} \to K^0_{S,L}h^+h^-$, ($h = \pi, K$) and its impact on the measurement of the CKM angle γ/ϕ_3*, Phys. Rev. D **82** (2010) 112006.

[57] H.-S. Wang *et al.*, *The three body decays $B^+ \to \pi^+\pi^-\pi^+$ in perturbative QCD approach*, Nuclear Physics A **930** (2014) 117.

[58] M. Bauer, B. Stech, and M. Wirbel, *Exclusive non-leptonic decays of D-, Ds and B-mesons*, Zeitschrift fur Physik C Particles and Fields **34** (1987) 103.

[59] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*, Phys. Rev. D **11** (1975) 3165.

[60] Belle collaboration, *Measurement of ϕ_3 with Dalitz plot analysis of $B \to D(*)$ K decay at Belle*, in *3rd Conference on Flavor Physics and CP Violation, Daegu, Korea*, 2004. arXiv:hep-ex/0411049.

[61] BaBar collaboration, *Improved measurement of the CKM angle γ in $B^\mp \to D^{(*)}K^{(*)\mp}$ decays with a Dalitz plot analysis of D decays to $K^0_S\pi^+\pi^-$ and $K^0_S K^+K^-$*, Physical Review D **78** (2008), no. 3 034023, arXiv:0804.2089.

[62] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics, Volume I*, Taylor & Francis, 2012. doi: 10.1201/b13717.

[63] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics*, Dover, 2012. ISBN-13: 978-0486668277.

[64] F. von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay widths, shapes, and production amplitudes*, Phys. Rev. D **5** (1972) 624.

[65] M. Peskin and D. Schroeder, *An introduction to quantum field theory*, Taylor & Francis, 2018. doi: 10.1201/9780429503559.

[66] S. U. Chung, *Spin formalisms*, Tech. Rep. 71-8, Brookhaven Nat. Lab., 2008. BNL-QGS-02-0900.

[67] C. Zemach, *Use of angular-momentum tensors*, Phys. Rev. **140** (1965) B97.

[68] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. D **51** (1995) 2247.

[69] S. U. Chung, *Helicity-coupling amplitudes in tensor formalism*, Phys. Rev. D **48** (1993) 1225.

[70] S. U. Chung *et al.*, *Partial wave analysis in K-matrix formalism*, Annalen der Physik **507** (1995), no. 5 404.

[71] Many textbooks cover this material. I recommend B. Martin, "Nuclear and Particle Physics: An Introduction" (2009).

[72] H. M. Pilkuhn, *The interactions of hadrons*, North-Holland Pub. Co., 1967.

[73] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-dominance prediction for* $\rho \to e^+e^-$, Phys. Rev. Lett. **21** (1968) 244.

[74] V. Anisovich *et al.*, *Two-pion spectra for the reaction* $\pi^- p \to \pi^0\pi^0 n$ *at a pion momentum of 38 GeV/c and combined analysis of GAMS, crystal barrel, and BNL data*, Physics of Atomic Nuclei **63** (2000), no. 8 1410.

[75] I. J. R. Aitchison, *The K-matrix formalism for overlapping resonances*, Nuclear Physics A **189** (1972), no. 2 417.

[76] S. U. Chung *et al.*, *Partial wave analysis in K-matrix formalism*, Annalen der Physik **507** (1995), no. 5 404.

[77] N. N. Bogoliubov, D. V. Shirkov, and S. Chomet, *Introduction to the theory of quantized fields*, Interscience New York, 1959.

[78] C. J. Joachain, *Quantum collision theory*, Elsevier Science Ltd, 1975. ISBN 200720402948.

[79] G. F. Carrier, M. Krook, and C. E. Pearson, *Functions of a complex variable: theory and technique*, SIAM, 2005. doi: 10.1137/1.9780898719116.

[80] A. O. Barut, I. Muzinich, and D. N. Williams, *Construction of invariant scattering amplitudes for arbitrary spins and analytic continuation in total angular momentum*, Phys. Rev. **130** (1963) 442.

[81] S. M. Flatte, *Coupled-channel analysis of the* $\pi\eta$ *and KK systems near KK threshold*, Physics Letters B **63** (1976), no. 2 224, CERN-EP-PHYS-76-8.

[82] D. Aston *et al.*, *A study of* $K^-\pi^+$ *scattering in the reaction* $Kp \to K^-\pi^+ n$ *at 11 GeV/c* , Nuclear Physics B **296** (1988), no. 3 493 .

[83] H. A. Bethe, *Theory of the effective range in nuclear scattering*, Phys. Rev. **76** (1949) 38.

[84] S. L. Adler, *Consistency conditions on the strong interactions implied by a partially conserved axial-vector current*, Physical Review **137** (1965), no. 4B B1022.

[85] F. Arbab and J. Jackson, *Factorization, kinematic singularities, and conspiracies*, Physical Review **176** (1968), no. 5 1796, LBNL Report UCRL-18261.

[86] V. V. Anisovich and A. V. Sarantsev, *K-matrix analysis of the* $(IJPC = 00++)$*-wave in the mass region below 1900 MeV*, The European Physical Journal A - Hadrons and Nuclei **16** (2003), no. 2 229.

[87] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.

[88] G. Brianti, *The Large Hadron Collider project: Historical account*, Phys. Rept. **403-404** (2004) 349.

[89] O. S. Bruning and P. Collier, *Building a behemoth*, Nature **448** (2007) 285.

[90] *Collisions at the large hadron collider*, http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/collisions.htm. Accessed: 25-04-2016.

[91] LHCb collaboration, B. Adeva *et al.*, *Roadmap for selected key measurements of LHCb*, 2009.

arXiv:0912.4179.

[92] LHCb collaboration, S. Amato *et al.*, *LHCb technical proposal*, tech. rep., 1998. CERN-LHCC-98-4.

[93] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005, LHCb-DP-2008-001.

[94] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[95] LHCb collaboration, R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) 09007, arXiv:1405.7808.

[96] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, arXiv:1408.1251.

[97] LHCb collaboration, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014), no. 01 P01002, arXiv:1311.3893.

[98] LHCb collaboration, *LHCb magnet: Technical design report*, tech. rep., 2000. CERN-LHCC-2000-007.

[99] LHCb collaboration, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[100] P. Perret, X. Vilasís-Cardona, L. Collaboration *et al.*, *Performance of the LHCb calorimeters during the period 2010–2012*, in *Journal of Physics: Conference Series*, vol. 587, p. 012012, IOP Publishing, 2015. doi: 10.1088/1742-6596/587/1/012012.

[101] LHCb collaboration, A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[102] LHCb collaboration, F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[103] LHCb collaboration, R. Aaij *et al.*, *The LHCb Trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[104] A. Puig, *The LHCb trigger in 2011 and 2012*, Tech. Rep. LHCb-PUB-2014-046, CERN, Geneva, Nov, 2014.

[105] J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, arXiv:1310.8544.

[106] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, Tech. Rep. LHCb-INT-2011-030, CERN, Geneva, Sep, 2011.

[107] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039, CERN, Geneva, May, 2014.

[108] G. Barrand *et al.*, *GAUDI – a software architecture and framework for building HEP data processing applications*, Computer Physics Communications **140** (2001), no. 1-2 45.

[109] A. Tsaregorodtsev *et al.*, *DIRAC, the LHCb data production and distributed analysis system*, tech. rep., Feb, 2006. LHCb-TALK-2006-014.

[110] A. Tsaregorodtsev *et al.*, *DIRAC3–the new generation of the LHCb grid software*, in *Journal of Physics: Conference Series*, vol. 219, p. 062029, IOP Publishing, 2010. doi: 10.1088/1742-6596/219/6/062029.

[111] J. Mościcki *et al.*, *Ganga: a tool for computational-task management and easy access to grid resources*, Computer Physics Communications **180** (2009), no. 11 2303.

[112] LHCb collaboration, 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.

[113] LHCb collaboration, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[114] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[115] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[116] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[117] 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.

[118] Geant4 collaboration, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[119] Geant4 collaboration, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[120] LHCb Collaboration, *The BOOLE project*, http://lhcb-releasearea.web.cern.ch/LHCb-release-area/DOC/boole. Accessed: 12-08-2016.

[121] LHCb Collaboration, *The BRUNEL project*, http://lhcb-releasearea.web.cern.ch/LHCb-release-area/DOC/brunel. Accessed: 12-08-2016.

[122] LHCb Collaboration, *The DAVINCI project*, http://lhcb-releasearea.web.cern.ch/LHCb-release-area/DOC/davinci. Accessed: 12-08-2016.

[123] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[124] J. Garra Ticó, *cfit library*, https://github.com/cfit. Accessed: 12-08-2016.

[125] P. Carniti *et al.*, *The CLARO ASIC: Design and performance of prototype integrated circuits for fast single photon counting at low power*, in *Nuclear Science Symposium and Medical Imaging Conference*, pp. 1–4, IEEE, 2013.

[126] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning; Data Mining, Inference and Prediction*, Springer, New York, 2009. doi: 10.1007/978-0-387-84858-7.

[127] A. Géron, *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*, O'Reilly Media, 2017. ISBN: 9781492032632.

[128] LHCb collaboration, J. Garra Ticó *et al.*, *Measurement of CP violation observables and constraints on the CKM angle γ using the $B^0 \to D(K^0_S\pi^+\pi^-)K^{*0}$ decays*, Tech. Rep. LHCb-ANA-2014-105, 2014. Unpublished.

[129] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[130] LHCb collaboration, A. Cukierman, C. Dorothy, and M. Williams, *Baryonic double open charm decays*, Tech. Rep. LHCb-ANA-2013-078, Mar, 2014. Unpublished.

[131] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation parameters in $B^0 \to DK^{*0}$ decays*, Physical Review D **90** (2014), no. 11 112002.

[132] A. Hoecker *et al.*, *TMVA toolkit for multivariate data analysis*, arXiv:physics/0703039, 2007.

[133] P. Geurts, A. Irrthum, and L. Wehenkel, *Supervised learning with decision tree-based methods in computational and systems biology*, Molecular Biosystems **5** (2009), no. 12 1593.

[134] Y. Freund, R. Schapire, and N. Abe, *A short introduction to boosting*, Journal-Japanese Society For Artificial Intelligence **14** (1999), no. 771-780 1612.

[135] Custodio, C. , *How to add bias and weight to Neural Network diagram*, https://tex.stackexchange.com/questions/401681/how-to-add-bias-and-weight-to-neural-network-diagram. Accessed: 20-03-2019.

[136] K. Hornik, M. Stinchcombe, and H. White, *Multilayer feedforward networks are universal approximators*, Neural networks **2** (1989), no. 5 359.

[137] A. Powell *et al.*, *Particle identification at LHCb*, PoS **ICHEP2010** (2010) 020, LHCb-PROC-2011-008.

[138] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[139] G. Cowan, *Statistical Data Analysis*, Oxford Science Publications, 1998.

[140] F. James, *MINUIT function minimization and error analysis: Reference manual version 94.1*, 1994. CERN-D-506.

[141] S. C. Haines, J. Garra Ticó, V. Gibson, and C. R. Jones, *A model-dependent measurement of the CKM angle γ using $B^\pm \to D(K^0_S\pi\pi)K^\pm$ decays*, Tech. Rep. LHCb-ANA-2012-050, Apr, 2012. Unpublished.

[142] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and constraints on the CKM angle γ in $B^\pm \to DK^\pm$ with $D \to K^0_S\pi^+\pi^-$ decays*, Nucl. Phys. **B888** (2014) 169, arXiv:1407.6211.

[143] 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.

[144] ARGUS collaboration, H. Albrecht *et al.*, *A Partial wave analysis of the decay* $D^0 \to K^0_S\pi^+\pi^-$, Phys. Lett. **B308** (1993) 435.

[145] M. Jacob and G. C. Wick, *On the general theory of collisions for particles with spin*, Annals of Physics **7** (1959), no. 4 404 .

[146] J. D. Richman, *An experimenter's guide to the helicity formalism*, Tech. Rep. CALT-68-1148, Calif. Inst. Technol., Pasadena, CA, Jun, 1984. CALT-68-1148.

[147] BESIII collaboration, *Precision measurement of the* D^{*0} *decay branching fractions*, Phys. Rev. **D91** (2015), no. 3 031101, arXiv:1412.4566.

[148] LHCb collaboration, P. Hunt *et al.*, *PIDCalib packages*, https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage. Unpublished. Accessed: 12-08-2016.

[149] LHCb collaboration, R. Aaij *et al.*, *Measurement of B meson production cross-sections in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, JHEP **08** (2013) 117, arXiv:1306.3663.

[150] *L0 hadron trigger efficiencies*, https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/CalorimeterObjectsToolsGroupDOC. Unpublished. Accessed: 01-09-2016.

[151] LHCb collaboration, R. Aaij *et al.*, *Prompt charm production in pp collisions at* $\sqrt{s} = 7$ *TeV*, Nucl. Phys. **B871** (2013) 1, arXiv:1302.2864.

[152] LHCb collaboration, A. K. Nandi and S. S. Malde, *Analysis of* $B^\pm \to DK^{*\pm}$ *with* $D \to K\pi, KK, \pi\pi, \pi K, K\pi\pi\pi, \pi\pi\pi\pi, \pi K\pi\pi$ *decays*, Tech. Rep. LHCb-ANA-2017-005, 2017. Unpublished.

[153] LHCb collaboration, D. Craik *et al.*, *Measurement of* γ *with an amplitude analysis of* $B^0 \to DK^+\pi^-$ *decays*, Tech. Rep. LHCb-ANA-2015-024, 2015. Unpublished.

[154] LHCb collaboration, R. Aaij *et al.*, *Measurement of* $\mathcal{CP}$ *violation and constraints on the CKM angle* γ *in* $B^\pm \to DK^\pm$ *with* $D \to K^0_S\pi^+\pi^-$ *decays*, Nuclear Physics B **888** (2014), no. 0 169 .

[155] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle* γ *using* $B^\pm \to DK^\pm$ *with* $D \to K^0_S\pi^+\pi^-, K^0_S K^+K^-$ *decays*, JHEP **10** (2014) 097.

[156] LHCb collaboration, R. Aaij *et al.*, *Model-independent measurement of the CKM angle* γ *using* $B^0 \to DK^{*0}$ *with* $D \to K^0_S\pi^+\pi^-$ *and* $K^0_S K^+K^-$, JHEP **06** (2016) 131, arXiv:1604.0152.

[157] LHCb collaboration, R. Aaij *et al.*, *A model-independent Dalitz plot analysis of* $B^\pm \to DK^\pm$ *with* $D \to K^0_S h^+h^-$ $(h = \pi, K)$ *decays and constraints on the CKM angle* γ, Phys. Lett. **B718** (2012) 43, arXiv:1209.5869.

[158] S. U. Chung, *Spin Formalisms*, Tech. Rep. CERN-71-08, 1971.

[159] J. Binney and D. Skinner, *The Physics of Quantum Mechanics*, Oxford, 2013. ISBN: 9780199688579.