CP violation measurements using fully reconstructed $B^{\pm} \to D^{*}h^{\pm}$ decays, $h \in (K, \pi)$

Alexandra Rollings

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2021

The more I see of the world, the more am I dissatisfied with it; and every day confirms my belief of the inconsistency of all human characters, and of the little dependence that can be placed on the appearance of merit or sense.

— *Jane Austen, Pride and Prejudice*

Abstract

Measurements of *CP* observables using fully reconstructed $B^{\pm} \to D^{*}h^{\pm}$ decays are made, $h \in (K, \pi)$, where the D^{*0} meson is reconstructed in the $D^{0}\pi^{0}$ and $D^{0}\gamma$ final states. The analysis is conducted using proton-proton collision data collected by the LHCb experiment, corresponding to a total integrated luminosity of $8.7\,\text{fb}^{-1}$: $3\,\text{fb}^{-1}$ taken at centre-of-mass energies of 7 TeV and 8 TeV (Run 1); and $5.7\,\text{fb}^{-1}$ at 13 TeV (Run 2). This work presents the first reconstruction of these decays at a hadron collider, and the first time D^{*0} mesons have been fully reconstructed for a *CP* violation analysis at a hadron collider. The 2-body D^{0} meson final states $K^{-}\pi^{+}$, $K^{-}K^{+}$, $\pi^{-}\pi^{+}$, and $\pi^{-}K^{+}$ are considered, facilitating the measurement of 16 *CP* observables in order to further constrain the CKM angle γ and the hadronic parameters $(r_{B}^{D^{*}K}, \delta_{B}^{D^{*}K}, r_{B}^{D^{*}\pi}, \delta_{B}^{D^{*}\pi})$. The first observation of the suppressed $B^{\pm} \to (D^{*} \to [\pi^{\mp}K^{\pm}]_{D}\gamma)K^{\pm}$ decay channel is also presented, colloquially known as the $D^{0}\gamma$ ADS mode.

Preface

This thesis describes the analysis of fully reconstructed $B^{\pm} \to D^{*}h^{\pm}$ decays, $h \in (K, \pi)$, at LHCb. Beyond common collaborative tasks of data collection, data processing and simulation production, I conducted this analysis alone. Therefore all the work presented in this thesis is my own, apart from the contributions from collaborators detailed below.

This work built on knowledge from the partially reconstructed (PR) analysis of such decays performed by Dónal Hill, a member of the LHCb collaboration and previous member of LHCb Oxford. The PR analysis is complementary to the fully reconstructed (FR) analysis in a number of ways, and the charged tracks considered are kinematically and topologically very similar. On a few occasions, therefore, where measurements made by the PR analysis were directly transferable to the FR technique, direct numbers were taken from his work and referenced in the thesis. The text under the title *Definition of asymmetry observables in terms of raw asymmetry and associated corrections* in Sec. 4.1.1 was also written by Dónal.

Malcolm John, a member of the LHCb Oxford group, ran the software required to perform the χ^2 minimisation procedure detailed in Chapter 5 for the interpretation of *CP* observables. The 2D confidence interval plots for the fundamental parameters given in Figs. 5.1 and 5.2 were therefore produced by him. I produced all other figures independently, unless explicitly referenced otherwise.

Acknowledgements

I would like to thank my supervisor, Malcolm John, for his guidance and advice over the course of this DPhil. His enthusiasm for everything physics is unrivalled. The work in this thesis would not have been possible without the insight and ideas of Dónal Hill, whose encouragement and motivational support was of immense help when the analysis hit (many) brick walls. I offer thanks to our collaborator, Vincent Tisserand, for his input and unfaltering enthusiasm. Sincere thanks to Sneha Malde for her help throughout this DPhil, particularly during the Oxford group meetings, where she was always on hand to offer useful suggestions. I am also grateful to Amanda Cooper-Sarkar for her support and advice during a challenging period.

I am grateful to the LHCb community at CERN, particularly the B Decays to Open Charm working group for their critical eye and invaluable input to the analysis. Thank you to Lorenzo Capriotti for the physics advice, many tasty dinners, and steadfast support.

I extend heartfelt thanks to Sue Geddes, who went above and beyond to support me and make my transition to and from CERN as smooth as possible during a very difficult time. Many thanks also to the rest of the administrative team in the particle physics sub-department, in particular to Kim Proudfoot.

The Science and Technology Facilities Council (STFC) made this DPhil possible by providing the funding for my studentship, and enabled me to spend one year on-site at CERN. For additional financial support, I would also like to thank New College, and the members of the Chancellor's Court of Benefactors who made donations to the Vice Chancellor's fund.

My time in the particle physics sub-department would not have been the same without the company of many inspiring colleagues. Thanks to all the fellow Oxford

LHCb students over the years, particularly Faye Cheung and Ollie Lupton, who helped me to get on my feet at the start of the DPhil (together, we made Christmas ever-present in the LHCb office). Thank you to the members of the SNO+ office next door, who offered a great escape out of the world of hadrons and into the world of neutrinos (and beer); your perspective and friendship has been invaluable. I also extend sincere thanks to the Scalable Parallel Computing Lab (SPCL) at ETH Zurich, who accepted me as an office refugee both during and after the COVID-19 pandemic.

Outside the world of physics, Maureen Freed and Jules Sealby showed me invaluable compassion and a generosity of time that I will never forget. Thank you to Harriet McAtee and the team at YogaQuota for offering me a sanctuary to escape to whenever I needed it.

It is safe to say that this DPhil has not been a smooth journey, and I am grateful to all the friends who have showed me patience, care, and empowered me to keep going; James and George must be named in particular, as I cannot imagine my time in Oxford without them. I would not be where I am today without the love and encouragement of my family: Mum, Dad, Josie, Matthew, Grandma, David, and Nanny; you can all relax now, it's finished! Last, but certainly not least, I am eternally grateful to Johannes, who never stopped believing in me and gave me the strength to make it to the end despite the many moments when that seemed like an impossible task.

Contents

1

CP violation and measurements of the CKM angle γ

This chapter places the work of the thesis in context of the wider aims of particle physics in the present day, and provides the theoretical background to understand the results presented. In Sec. 1.2 and Sec. 1.3, the concept of CP violation is explained and its origins in the Standard Model are presented. A specific strategy to make measurements of the CP-violating parameter γ, using $B^{\pm} \to D^{*}h^{\pm}$ decays, is then outlined in Sec. 1.4

1.1 Introduction

The Standard Model (SM) of particle physics is a quantum field theory which describes the electromagnetic, strong and weak interactions, and classifies all known elementary particles. Work over the last century has developed this successful theory, and many aspects of it have been validated by particle physics experiments around the world, a substantial contribution of which has been made by the CERN experiment in Geneva. The SM, however, is unable to explain a number of important phenomena. No mathematical unification of general relativity and the SM has been found, therefore the gravitational force is omitted. The matter that the SM describes makes up $< 5\%$ of the observable Universe, and no explanation has been found for *dark*

matter or *dark energy*, which make up the rest. Of particular importance to this thesis, the model fails to describe the matter-antimatter asymmetry of the observable, matter-dominated Universe [1, 2].

Studies of possible Beyond the Standard Model (BSM) effects to explain these (and further) short comings are wide-spread in particle physics. One way to search for BSM phenomena is to directly search for new, heavy particles in regions of (potentially infinite) phase space; another is to make precision measurements of low energy processes, and look for evidence of new particles by searching for discrepancies between these measurements and SM predictions. This is made possible by the fact that particles can influence interactions occurring at energies orders of magnitude smaller than would be required for their direct production. Within the field of *flavour physics*, which studies processes that distinguish different generations of quarks and leptons, this method has been historically successful; the existence of the charm and third-generation quarks were postulated before direct searches were possible, inferred by the absence of flavour changing neutral currents [3] and the observation of CP violation [4], respectively.

The LHCb experiment at CERN plays a significant role in making precision measurements within the field of flavour physics, including the work documented in this thesis. World-leading measurements using $B^{\pm} \to D^{*}\pi^{\pm}$ and $B^{\pm} \to D^{*}K^{\pm}$ decays are made, in order to constrain the CP-violating phase γ, a fundamental parameter of the SM which facilitates the small amount of matter-antimatter asymmetry that the model describes.

1.2 C, P and T symmetries

The mathematician Emmy Noether established that every continuous symmetry of a physical system has a corresponding conservation law [5]; the laws of physics are invariant under temporal, spatial and rotational translations, predicting the conservation of energy, momentum and angular momentum in physical systems, respectively. In elementary particle physics, the requirement that fundamental particle

fields are invariant under a local $U(1)_Y \times SU(2)_L \times SU(3)_C$ gauge symmetry forms the foundation of the Standard Model.

For this thesis, of particular relevance are the discreet symmetries of charge conjugation, parity transformation and time reversal; mathematically mediated by their respective operators C, P and T:

- C: converts particles to their corresponding antiparticles, achieved by changing the sign of all the quantum charges of a particle state.
- P: inverts the spatial dimensions of space time: $(t, \boldsymbol{x}) \to (t, -\boldsymbol{x})$.
- T inverts the temporal dimension of space time: $(t, \boldsymbol{x}) \to (-t, \boldsymbol{x})$.

The simultaneous transformation of a physical system under all three symmetries leaves the system unchanged. This is currently observed to be an exact symmetry of nature, and is bound up in the CPT theorem [6], however, any one of these symmetries can be broken individually.

Parity transformation symmetry was the first to be observed to be broken; in 1956, Chien-Shiung Wu studied decays of Co^{60} nuclei and determined that the weak interaction maximally violated parity [7]. Violation of C was also inferred from this discovery, leaving open the possibility that the two symmetries were conserved when combined, *i.e.* the laws of physics were *CP*-invariant. This was disproved in 1964 by Cronin and Fitch, who observed that long-lived kaons, which if *CP* was conserved would only decay to the *CP*-odd 3π final state, also decayed to the *CP*-even 2π final state [8].

Since then, *CP* violation has been looked for and observed in many other physical systems. Concerning B physics, the BaBar and Belle collaborations first observed *CP* violation in neutral B^0 mesons in the early 2000s [9, 10]. They later found evidence in charged $B^\pm$ mesons [11, 12], as did LHCb [13], and in 2013, LHCb observed *CP* violation in neutral B_s^0 systems [14]. The first observation of *CP* violation in D^0 decays was made by LHCb in 2019 [15]. In 2020, the first evidence of *CP* violation in the neutrino sector was found by the T2K experiment [16].

The observed *CP* violating effects, collectively responsible for these discoveries, can be grouped into three distinct categories:

1. *CP violation in decay* (direct *CP* violation) occurs when the decay amplitude for a process and its *CP*-conjugate are not equal: $\Gamma(X \to f) \neq \Gamma(\bar{X} \to \bar{f})$. This is the only type of *CP* violation accessible to charged initial states, and is therefore the focus of this thesis.

2. *CP violation in mixing* occurs in neutral meson systems when the oscillation rate from meson to anti-meson differs from the oscillation rate from anti-meson to meson: $\Gamma(X^0 \to \bar{X}^0) \neq \Gamma(\bar{X}^0 \to X^0)$.

3. *CP violation in the interference between mixing and decay* is also accessible to neutral meson systems. The interference between the amplitudes of the direct decay $\Gamma(X^0 \to f)$ and that of the decay after mixing $\Gamma(X^0 \to \bar{X}^0 \to f)$ can cause the *CP*-conjugate decay rates $\Gamma(X^0 \to f)$ and $\Gamma(\bar{X}^0 \to \bar{f})$ to differ, even in the absence of the prior two cases.

The study of *CP* violation remains important as it is the only known process that facilitates behavioural differences between matter and antimatter. Although the observed level of *CP* violation within the quark sector of the SM is not enough to explain the baryon anti-baryon asymmetry of the observable Universe [17], precise measurements of this level provide a benchmark against which to judge any matter-antimatter asymmetry effects arising from new physics.

1.3 *CP* violation in the Standard Model

In the Standard Model, the charged-current, weak interactions of quarks are described by the SM Lagrangian term given in Eq. (1.1), where *h.c.* stands for hermitian conjugate.

$$L_{CC} = -\frac{g}{\sqrt{2}} \overline{u}_{Li} \gamma^{\mu} W_{\mu}^{+} (V_{CKM})_{ij} d_{Lj} + h.c. \tag{1.1}$$

Here, g is the weak coupling constant; γ^u are the Dirac matrices; W_{μ}^{+} is the W boson field, which couples to the up-type and down-type left-handed quark triplets, $\overline{u}_{Li}$

and d_{Lj}, where $i, j = 1, 2, 3$ are the generation numbers; and V_{CKM} are elements of the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix [4].

The elements of the CKM matrix are coupling constants quantifying the strength of inter- and intra-generational quark mixing, and can be represented by complex numbers with arbitrary phases. For a 3×3 unitary matrix, $3^2 = 9$ real parameters must be specified. Our ability to absorb one phase into each quark field except for one overall absolute phase removes 5 of these parameters. 4 degrees of freedom therefore remain: 3 real rotations and 1 complex phase. It is the presence of this single, irreducible phase that generates CP violation in the quark sector.

In the Chau-Keung parameterisation of V_{CKM} [18], the three rotation angles are denoted θ_{12}, θ_{13} and θ_{23}, and the complex phase δ_{13}:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}c_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{1.2}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$; this is the standard representation, preferred by the *Particle Data Group* (PDG) [19]. The parameter θ_{12} is the well-known Cabibbo angle [20]. It is experimentally known that $s_{13} \ll s_{23} \ll s_{12} \ll 1$, therefore it is convenient to express the matrix in terms of an alternative parameterisation that naturally incorporates the hierarchy. The Wolfenstein parameterisation [21] formulates the elements in terms of a power series in $\lambda \equiv s_{12} \simeq 0.23$:

$$s_{12} \equiv \lambda \tag{1.3a}$$

$$s_{23} \equiv \lambda^2 A \tag{1.3b}$$

$$s_{13}e^{-i\delta_{13}} \equiv \lambda^3(\rho - i\eta), \tag{1.3c}$$

where A, ρ and η are all real parameters of order unity. To $\mathcal{O}(\lambda^4)$, the CKM matrix becomes:

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

It is worth noting that the CP violating term $(\rho - i\eta)$ is only present in the V_{ub} and V_{td} matrix elements, up to $\mathcal{O}(\lambda^3)$. Larger CP violating effects are therefore expected in B

systems compared to K or D systems, and *CP* violation in the interference between mixing and decay was observed in B^0 systems [9, 10] (whose mixing diagrams feature the V_{td} element) before B_s^0 systems [14] (whose mixing involves the V_{ts} element).

Verifying the unitarity of the CKM matrix is an important test of the SM [4]. Orthogonality relations can be derived between the rows and columns of the matrix:

$$\sum_{i\in\{u,c,t\}} V_{i\alpha}^* V_{i\beta} = \delta_{\alpha\beta},\ \ \alpha,\beta \in \{d,s,b\} \tag{1.5a}$$

$$\sum_{\alpha\in\{b,s,b\}} V_{\alpha i}^* V_{\alpha j} = \delta_{ij},\quad\ \ i,j \in \{u,c,t\} \tag{1.5b}$$

such that there are 6 off-diagonal conditions, each containing 3 complex numbers that sum to zero. These can be visualised as triangles in the complex plane, where the area of each triangle is proportional to the amount of *CP* violation in the SM quark sector; in the absence of any *CP* violation, they degenerate to lines along the real axis. Of particular interest is the equation:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.6}$$

where all terms are of $\mathcal{O}(\lambda^3)$. This condition forms the *Unitarity Triangle* (UT) in the complex plane when each side is divided by $V_{cd}V_{cb}^*$. An illustration of the UT is given in Fig. 1.1, where it can be seen that all sides and internal angles are of approximately equal size, allowing them to be accessed experimentally. By over-constraining the UT, making independent measurements using many different decay channels, it is possible to put CKM unitarity to the test.

The angles in this triangle, α, β, γ, also referred to in the literature as ϕ_2, ϕ_1 and ϕ_3, can be accessed experimentally via the study of B meson decays. For many years, the least well known angle was:

$$\gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*), \tag{1.7}$$

but recently its precision rivals that of α. γ is the Standard Candle of the UT, as it can be measured using decays that are described by *tree-level* amplitudes only, which benefit from very low theoretical uncertainty from higher-order diagrams,

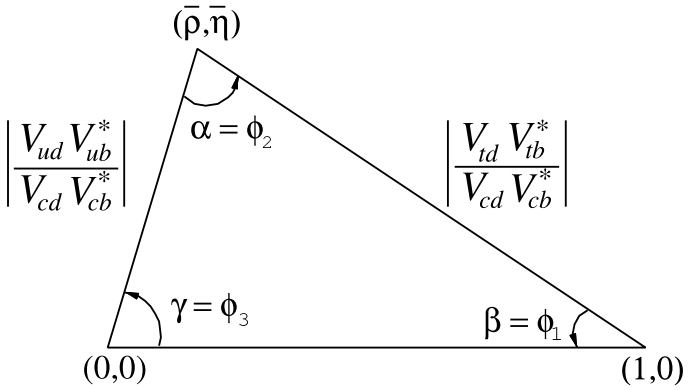


Figure 1.1: Illustration of the Unitary Triangle in the complex plane, where the upper apex is defined in terms of the Wolfenstein parameters $\bar{\rho} = \rho(1 - \lambda^2/2)$ and $\bar{\eta} = \eta(1 - \lambda^2/2)$, up to $\mathcal{O}(\lambda^3)$. The figure is reproduced from the *CKM Quark-Mixing Matrix* review of the PDG [19].

$\frac{\delta\gamma}{\gamma} \sim 10^{-7}$ [22]. Combining individual results from γ-sensitive decays, the latest published LHCb measurement of this angle is $\gamma = (65.4^{+3.8}_{-4.2})°$ [23]. This is the most precise determination of γ from a single experiment, and is consistent with previous results from BaBar and Belle.

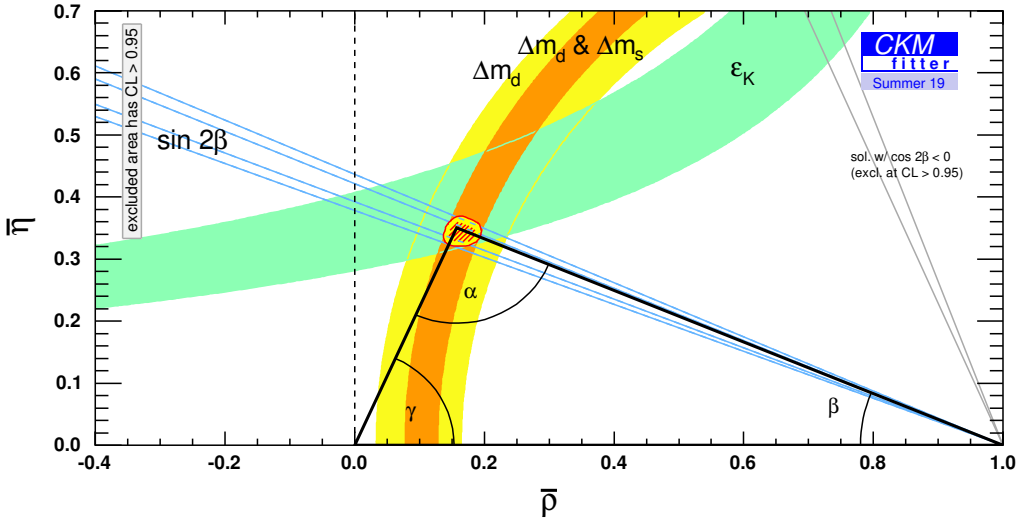


Figure 1.2: CKMfitter constraints in the $(\bar{\rho}, \bar{\eta})$ plane using only 'loop' quantities [24]. Each coloured region indicates an area of allowed parameter space for a particular class of *CP* violating measurements, constraining the sides and angles of the UT.

It is also possible to constrain γ indirectly, by excluding tree-level measurements and performing global fits to the UT whilst assuming unitarity. The CKMfitter group

currently estimates the *indirect* measurement to be $\gamma = (65.7^{+0.9}_{-2.65})^\circ$ [24]. This value is chiefly constrained by the angle β and the length of the side of the triangle opposite γ, as can be seen in Fig. 1.2; both rely on measurements from neutral B mixing processes. Over the past few years, flavour anomalies have appeared in lepton universality tests with $b \to sll$ decays [25–29]: rare flavour-changing neutral currents that are mediated by loop-level processes. These anomalies could be explained by the presence of new, heavy particles that also have the potential to enter neutral B meson mixing loops [30]. If new physics effects enter the mixing loops, but are not accounted for in the global fit, different values of γ could result from tree-level and loop-level constraints.

The latest LHCb combination does include *CP* violation measurements from the interference of mixing and decay in B^0 and B^0_s systems, but the precision is driven by time-integrated measurements of direct *CP* violation in $B^\pm \to D^{(*)}K^{(*)\pm}$ decays. It also contains many recent LHCb analyses that have not yet been included in an updated direct result by CKMfitter, including the full LHCb dataset analyses of $B^\pm \to D^{(*)}h^\pm$ decays using the ADS/GLW methods [31], and of $B^\pm \to Dh^\pm$ decays using the BPGGSZ method [32]. The GLW and ADS techniques will be described in Secs. 1.4.1 and 1.4.2; the BPGGSZ method [33–35] will not be covered in this thesis, but it differs from the former techniques as it considers 3-body D decays and exploits phase-space dependent D decay rates.

The direct and indirect measurements of γ currently show good agreement, but the direct measurement in particular possesses significant uncertainties. The uncertainties on the indirect measurement are currently limited by lattice QCD calculations; the uncertainties on the direct measurement are currently limited by experimental precision. It is therefore highly motivated to explore new ways to exploit LHCb data to make measurements of γ-sensitive decays using tree-level processes.

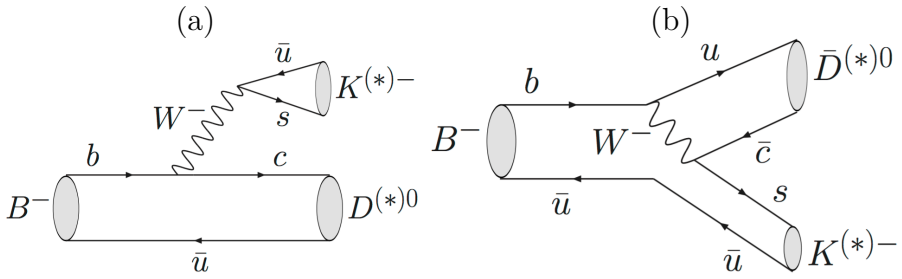


Figure 1.3: Feynman diagrams for $B^- \to D^{(*)}K^{(*)-}$ decays. The decay to $\overline{D}^{(*)0}$ is CKM $|V_{ub}V_{cs}^*/V_{cb}V_{us}^*| \approx 0.4$ and colour $F_{CS} \approx 1/3$ suppressed with respect to the $D^{(*)0}$ final state. The interference of these two diagrams is the common mechanism for accessing γ.

1.4 Tree-level determination of γ using $B^\pm \to D^*K^\pm$ decays

Decays of the kind $B^- \to D^{(*)}K^{(*)-}$, with intermediate states $\overline{D}^{0(*)}$ and $D^{0(*)}$, are mediated by $b \to u\bar{c}s$ and $b \to c\bar{u}s$ transitions, respectively. The Feynman diagrams for these decays are shown in Fig. 1.3. The amplitude ratio in terms of CKM elements is $V_{ub}V_{cs}^*/V_{cb}V_{us}^* \approx 0.4$, therefore the decay involving a $\overline{D}^{0(*)}$ is *Cabibbo suppressed*. This transition also suffers from colour suppression, as the colours of the $\bar{c}$ and s quarks are fixed from the pre-defined colours of the u and $\overline{u}$ quarks with which they form bound states. The relative weak phase between the suppressed and favoured decays, $\arg\left(V_{ub}V_{cs}^*/V_{cb}V_{us}^*\right)$, equates to the definition of the CKM angle γ given in Eq. (1.7), up to $\mathcal{O}(\lambda^4) \simeq 2.6\times10^{-3}$. This is well below the current level of experimental uncertainty, therefore γ measurements can be made from studying $B^\pm \to D^{(*)}K^{(*)\pm}$ decays. This thesis studies the variant $B^\pm \to D^*K^\pm$, where the neutral D^* meson represents D^{*0} or $\overline{D}^{*0}$ and is fully reconstructed in both the $D\pi^0$ and $D\gamma$ final states, for which the branching fractions are given below:

$$\mathcal{B}(D^{*0} \to D^0\pi^0) = (64.7 \pm 0.9) \times 10^{-2} \tag{1.8}$$

$$\mathcal{B}(D^{*0} \to D^0\gamma) = (35.3 \pm 0.9) \times 10^{-2}. \tag{1.9}$$

Defining the amplitude ratio $r_B^{D^*K} = \frac{|A(B^-\to\bar{D}^{*0}K^-)|}{|A(B^-\to D^{*0}K^-)|}$, and the relative strong phase $\delta_B^{D^*K}$, between $B^- \to \overline{D}^{*0}K^-$ and $B^- \to D^{*0}K^-$ decays, the amplitude for producing

a neutral D^* meson in the decay $B^- \to \tilde{D}^* K^-$ can be expressed as:

$$\tilde{D}^* = D^{*0} + r_B^{D^*K} e^{i(\delta_B^{D^*K} - \gamma)} \overline{D}^{*0}. \tag{1.10}$$

Here, the term *strong* phase refers to a CP conserving phase arising from possible intermediate states in the decay amplitude. Writing Eq. (1.10) in terms of odd and even CP eigenstates, where ${D^*}_+ = \frac{D^{*0} + \overline{D}^{*0}}{\sqrt{2}}$ and ${D^*}_- = \frac{D^{*0} - \overline{D}^{*0}}{\sqrt{2}}$, gives:

$$\tilde{D}^* = \frac{{D^*}_+ + {D^*}_-}{\sqrt{2}} + r_B^{D^*K} e^{i(\delta_B^{D^*K} - \gamma)} \frac{{D^*}_+ - {D^*}_-}{\sqrt{2}}. \tag{1.11}$$

The CP eigenvalue of the D^* state can be evaluated using the product:

$$\lambda_{D^*} = \lambda_D \times \lambda_{\pi^0/\gamma} \times (-1)^l, \tag{1.12}$$

where l is the orbital angular momentum quantum number. In the case of strong D^* ($J^P = 1^-$) decay via π^0 ($J^{PC} = 0^{-+}$) emission, $\lambda_{\pi^0} = -1$, and $l = 1$ by conservation of angular momentum, therefore $\lambda_{D^*} = +\lambda_D$ and ${D^*}_\pm \to D_\pm \pi^0$. In the case of γ ($J^{PC} = 1^{--}$) emission, $\lambda_\gamma = +1$; by conservation of momentum l could be zero, but this would violate parity, so $l = 1$.[1] This means that $\lambda_{D^*} = -\lambda_D$ and ${D^*}_\pm \to D_\mp \gamma$, introducing a phase shift of π ($e^{i\pi} = -1$) into Eq. (1.11) for the $D\gamma$ mode:

$$\tilde{D}^* \to \tilde{D}\pi^0 : \tilde{D} = D^0 + r_B^{D^*K} e^{i(\delta_B^{D^*K} - \gamma)} \overline{D}^0 \tag{1.13}$$

$$\tilde{D}^* \to \tilde{D}\gamma : \tilde{D} = D^0 + r_B^{D^*K} e^{i(\delta_B^{D^*K} + \pi - \gamma)} \overline{D}^0. \tag{1.14}$$

This logic was first shown in Ref. [36].

Labelling the amplitude of the diagram displayed in Fig. 1.3 (a) as $A_B = A(B^- \to D^{*0}K^-) = A(B^+ \to \overline{D}^{*0}K^+)$, we can construct 4 equations for the decay amplitudes of B^- mesons to final states $f(D)$, or their complex conjugates $\bar{f}(D)$:

$$A_{Bf}^{\pi^0} = A(B^- \to [f(D)\pi^0]_{D^*}K^-) = A_B A_{\pi^0}(A_D + \bar{A}_D e^{-i\delta_D} r_B^{D^*K} e^{i(\delta_B^{D^*K} - \gamma)}) \tag{1.15}$$

$$A_{B\bar{f}}^{\pi^0} = A(B^- \to [\bar{f}(D)\pi^0]_{D^*}K^-) = A_B A_{\pi^0}(\bar{A} e_D^{-i\delta_D} + A_D r_B^{D^*K} e^{i(\delta_B^{D^*K} - \gamma)}) \tag{1.16}$$

$$A_{Bf}^{\gamma} = A(B^- \to [f(D)\gamma]_{D^*}K^-) = A_B A_{\gamma}(A_D + \bar{A}_D e^{-i\delta_D} r_B^{D^*K} e^{i(\delta_B^{D^*K} + \pi - \gamma)}) \tag{1.17}$$

$$A_{B\bar{f}}^{\gamma} = A(B^- \to [\bar{f}(D)\gamma]_{D^*}K^-) = A_B A_{\gamma}(\bar{A}_D e^{-i\delta_D} + A_D r_B^{D^*K} e^{i(\delta_B^{D^*K} + \pi - \gamma)}). \tag{1.18}$$

[1] The $D^0\gamma$ final state must have the angular momentum projection quantum number $m = 0$ to conserve angular momentum despite having $l = 1$ total angular momentum.

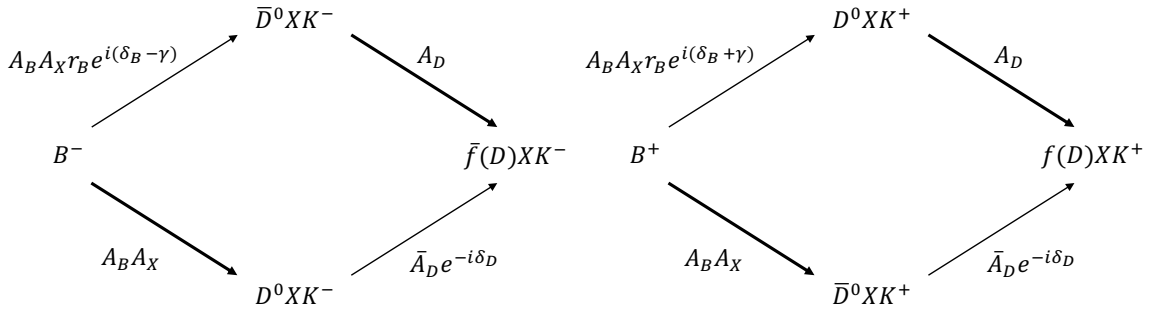


Figure 1.4: The decay diagram for $B^- \to D^*K^-$ transitions to a general D^* final state $\bar{f}(D)X$, $X = \pi^0/\gamma$, is depicted on the left. The complex conjugate decay for $B^+ \to D^*K^+$ transitions, to the $f(D)X$ final state, is given on the right. These decays can proceed via 2 interfering intermediate states containing D^0 or $\overline{D}^0$ mesons.

The D meson amplitudes are defined as: $\bar{A}_D = A(D^0 \to \bar{f}(D)) = A(\overline{D}^0 \to f(D))$ and $A_D = A(\overline{D}^0 \to \bar{f}(D)) = A(D^0 \to f(D))$; δ_D is the CP invariant relative strong phase difference between the D and $\overline{D}$ mesons decaying to the same final state. 4 analogous, charge conjugate equations can be written for B^+ mesons, where the only change is the sign of the CP violating phase γ in the exponent. Flow diagrams for B^+ and B^- decays are given in Fig. 1.4 to illustrate this.

It is also possible to access γ by studying $B^\pm \to D^*\pi^\pm$ decays, extracting the weak phase between $b \to c\bar{u}d$ and $b \to u\bar{d}s$ transitions. These decays have a branching fraction ~ 10 times larger than their $B^\pm \to D^*K^\pm$ counterparts:

$$\mathcal{B}(B^\pm \to D^*\pi^\pm) = (4.90 \pm 0.17) \times 10^{-3} \tag{1.19}$$

$$\mathcal{B}(B^\pm \to D^*K^\pm) = (3.97^{+0.31}_{-0.28}) \times 10^{-4}. \tag{1.20}$$

The suppressed to favoured amplitude ratio $r_B^{D^*\pi} = \frac{|A(B^- \to \bar{D}^{*0}\pi^-)|}{|A(B^- \to D^{*0}\pi^-)|}$, however, is smaller by a factor $|V_{us}/V_{cs}|^2 \approx 0.05$ than $r_B^{D^*K}$, making this mode considerably less sensitive to interference effects and therefore γ. Both types of decay, with the pion and kaon *companion* hadron, are studied in this thesis, as the former can also be used as 'control channels' to develop models and constrain nuisance parameters.

1.4.1 The Gronau-London-Wyler (GLW) method

When applying the GLW method [37], the D meson is reconstructed in the CP-even final states $D \to K^+K^-$ and $D \to \pi^+\pi^-$, therefore $\delta_D = 0$ and $A_D = \bar{A}_D$. Time

integrated decay rates are proportional to the squared magnitude of the amplitudes, therefore the following equations for $B^{\pm} \to D^* h^{\pm}$ decays can be constructed, where the companion particle h can take the form of a kaon or a pion:

$$\Gamma_{\pm}^{CP+}(B^{\pm} \to (D^* \to [hh]\pi^0)h^{\pm}) \propto |A_{Bf_{CP}}^{\pi^0}|^2 \propto 1 + (r_B^{D^*h})^2 + 2r_B^{D^*h}\cos(\delta_B^{D^*h} \pm \gamma) \tag{1.21}$$

$$\Gamma_{\pm}^{CP-}(B^{\pm} \to (D^* \to [[hh]\gamma)h^{\pm}) \propto |A_{Bf_{CP}}^{\gamma}|^2 \propto 1 + (r_B^{D^*h})^2 - 2r_B^{D^*h}\cos(\delta_B^{D^*h} \pm \gamma). \tag{1.22}$$

Here, the D decay particles hh stand for K^+K^- or $\pi^+\pi^-$. The strong phase shift of π from the $D^* \to D\gamma$ decay results in a $CP-$ (CP-odd) eigenstate.

1.4.2 The Atwood-Dunietz-Soni (ADS) method

When applying the ADS method [38], the D meson is reconstructed in the doubly-Cabibbo suppressed $D^0 \to K^+\pi^-$ and Cabibbo-favoured $D^0 \to K^-\pi^+$ final states, and their charge conjugates. These modes are given the respective labels SUP and FAV. The SUP mode is also known as the ADS mode, and is referred to with the symbols πK in this text; the favoured mode is labelled by the opposite symbol arrangement, $K\pi$.

Defining $A_D = A(D^0 \to K^-\pi^+) = A(\overline{D}^0 \to K^+\pi^-)$ and $\bar{A}_D = A(D^0 \to K^+\pi^-) = A(\overline{D}^0 \to K^-\pi^+) = A_D r_D$, where r_D is the amplitude ratio of the SUP with respect to the FAV D decay mode, the following decay rate equations for $B^{\pm} \to D^* h^{\pm}$ decays are built:

$$\Gamma_{\pm}^{SUP,\pi^0}(B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\pi^0)h^{\pm}) \propto (r_B^{D^*h})^2 + r_D^2 + 2r_B^{D^*h}r_D\cos(\delta_B^{D^*h} + \delta_D \pm \gamma) \tag{1.23}$$

$$\Gamma_{\pm}^{SUP,\gamma}(B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\gamma)h^{\pm}) \propto (r_B^{D^*h})^2 + r_D^2 - 2r_B^{D^*h}r_D\cos(\delta_B^{D^*h} + \delta_D \pm \gamma) \tag{1.24}$$

$$\Gamma_{\pm}^{FAV,\pi^0}(B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)h^{\pm}) \propto 1 + (r_B^{D^*h})^2 r_D^2 + 2r_B^{D^*h}r_D\cos(\delta_B^{D^*h} - \delta_D \pm \gamma) \tag{1.25}$$

$$\Gamma_{\pm}^{FAV,\gamma}(B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)h^{\pm}) \propto 1 + (r_B^{D^*h})^2 r_D^2 - 2r_B^{D^*h}r_D\cos(\delta_B^{D^*h} - \delta_D \pm \gamma). \tag{1.26}$$

The D^*K ADS mode possesses particularly high sensitivity to γ as the amplitude ratios r_D and $r_B^{D^*K}$ are of similar magnitudes; the size of the interference term is therefore relatively large.

It should be noted that when studying excited D^* intermediate states, double the number of γ-constraining equations are obtained compared to traditional analyses of $B^\pm \to Dh^\pm$ decays due to the strong phase difference between $D^* \to D\pi^0$ and $D^* \to D\gamma$ decays [36].

1.4.3 Including D-mixing

The effect of D-mixing on $B^\pm \to D^{(*)}h^{(*)\pm}$ observables was modelled mathematically by Matteo Rama in 2013 [39].[2] He showed that the GLW mode observables were unaffected due to the cancellation of D mixing terms in R^{CP} and A_h^{CP}, for which the D decay modes are common to both the numerator and denominator. The D-mixing correction to the ADS mode observables, however, has the potential to cause a significant bias in the $D^*\pi$ sample if left unaccounted for. The formulae below have therefore been calculated to include these corrections, and are used in the interpretation of the final results given in Chapter 5:

$$\begin{aligned}\Gamma_\pm^{SUP,\pi^0}(B^\pm \to D^*(\to [K^\mp\pi^\pm]_D\pi^0)h^\pm) \propto\ & r_D^2 + (r_B^{D^*h})^2 \\ & + 2r_D r_B^{D^*h}\cos(\delta_B^{D^*h} + \delta_D \mp \gamma) \\ & - \alpha y(1 + (r_B^{D^*h})^2) r_D \cos\delta_D \\ & - \alpha y(1 + r_D^2) r_B^{D^*h}\cos(\delta_B^{D^*h} \mp \gamma) \\ & + \alpha x(1 - (r_B^{D^*h})^2) r_D \sin\delta_D \\ & - \alpha x(1 - r_D^2) r_B^{D^*h}\sin(\delta_B^{D^*h} \mp \gamma)\end{aligned} \tag{1.27}$$

$$\begin{aligned}\Gamma_\pm^{SUP,\gamma}(B^\pm \to D^*(\to [K^\mp\pi^\pm]_D\gamma)h^\pm) \propto\ & r_D^2 + (r_B^{D^*h})^2 \\ & - 2r_D r_B^{D^*h}\cos(\delta_B^{D^*h} + \delta_D \mp \gamma) \\ & + \alpha y(1 + (r_B^{D^*h})^2) r_D \cos\delta_D \\ & - \alpha y(1 + r_D^2) r_B^{D^*h}\cos(\delta_B^{D^*h} \mp \gamma) \\ & - \alpha x(1 - (r_B^{D^*h})^2) r_D \sin\delta_D \\ & - \alpha x(1 - r_D^2) r_B^{D^*h}\sin(\delta_B^{D^*h} \mp \gamma).\end{aligned} \tag{1.28}$$

[2] In the calculations, mixing induced CP violation in D mesons was neglected.

Compared to the definitions given in Sec. 1.4.2, these are dependent on three additional variables:

- α: the analysis specific decay-time acceptance co-efficient. The decay time acceptance function for D mesons is needed when integrating the state of the system over time, in order to evaluate the time-integrated effects of D-mixing. The effect of the non-uniform function in this analysis is bound up in the multiplicative correction factor, $\alpha = 0.88 \pm 0.03$. A description of the evaluation of this parameter using fits to D decay times is given in Appendix C.

- x, y: the D mixing parameters are taken from the *Heavy Flavour Averaging Group* (HFLAV) 2020 [40] (CPV allowed):

$$x = \frac{\Delta M}{\Gamma} = (0.31^{+0.11}_{-0.12})\% \tag{1.29}$$

$$y = \frac{\Delta \Gamma}{2\Gamma} = (0.651^{+0.063}_{-0.0.069})\%, \tag{1.30}$$

  where ΔM and $\Delta\Gamma$ are the mass and width differences between the mass eigenstates of the D system, respectively, and Γ is their average width.

Referring to the flow diagram pictured in Fig. 1.4, D-mixing in the suppressed mode equates to a favoured $B^- \to D^0 X K^-$ decay, followed by the oscillation: $D^0 \to \overline{D}^0$, followed by a favoured $\overline{D}^0 \to K^+\pi^-$ decay (or the charge conjugate process). Favoured decays therefore infiltrate the suppressed sample at a rate $\sim |x|, |y|$, as the corrections in Eqs. (1.27) and (1.28) are linear in the terms αx and αy; the relative size of these terms compared to $r_B^{D^*h}$ determines the size of D mixing effects. In the D^*K ADS mode, $r_B^{D^*K} \sim \mathcal{O}(0.1)$, therefore D mixing effects are not visible in this channel at the current level of experimental uncertainty. In the $D^*\pi$ ADS mode, however, $r_B^{D^*\pi} \sim \mathcal{O}(0.005)$, therefore D mixing effects must be accounted for. A similar rate of D mixing occurs in the favoured mode, but seeing as this corresponds to the rate of suppressed B decays entering the favoured sample, the effect is negligible.

1.4.4 CP Observables

In order to extract γ and the hadronic parameters $r_B^{D^*K}$, $\delta_B^{D^*K}$, $r_B^{D^*\pi}$, and $\delta_B^{D^*\pi}$, ratios are constructed of topologically identical final states. In this way, the CP observables do not depend on absolute efficiencies, and many systematic uncertainties cancel in the final results. Descriptions of the observables and definitions in terms of the decay rates derived earlier in this section are given below:

1. The ratio of $B^\pm \to D^*K^\pm$ to $B^\pm \to D^*\pi^\pm$ decays for the favoured D decay mode:

$$R_{K/\pi}^{K\pi,\pi^0/\gamma} = \frac{\Gamma(B^- \to ([K^-\pi^+]_D\pi^0/\gamma)_{D^*}K^-) + \Gamma(B^+ \to ([K^+\pi^-]_D\pi^0/\gamma)_{D^*}K^+)}{\Gamma(B^- \to ([K^-\pi^+]_D\pi^0/\gamma)_{D^*}\pi^-) + \Gamma(B^+ \to ([K^+\pi^-]_D\pi^0/\gamma)_{D^*}\pi^+)}. \tag{1.31}$$

2. The CP asymmetry between B^- and B^+ mesons in the favoured D decay mode (for $B^\pm \to D^*K^\pm$ decays only):

$$A_K^{K\pi,\pi^0/\gamma} = \frac{\Gamma(B^- \to ([K^-\pi^+]_D\pi^0/\gamma)_{D^*}K^-) - \Gamma(B^+ \to ([K^+\pi^-]_D\pi^0/\gamma)_{D^*}K^+)}{\Gamma(B^- \to ([K^-\pi^+]_D\pi^0/\gamma)_{D^*}K^-) + \Gamma(B^+ \to ([K^+\pi^-]_D\pi^0/\gamma)_{D^*}K^+)}. \tag{1.32}$$

3. The CP asymmetries between B^- and B^+ mesons decaying to CP eigenstates, where $hh = \pi\pi/KK$:

$$A_\pi^{CP,\pi^0/\gamma} = \frac{\Gamma(B^- \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}\pi^-) - \Gamma(B^+ \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}\pi^+)}{\Gamma(B^- \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}\pi^-) + \Gamma(B^+ \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}\pi^+)} \tag{1.33}$$

$$A_K^{CP,\pi^0/\gamma} = \frac{\Gamma(B^- \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}K^-) - \Gamma(B^+ \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}K^+)}{\Gamma(B^- \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}K^-) + \Gamma(B^+ \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}K^+)}. \tag{1.34}$$

   A correction to account for the difference in CP asymmetry between $D \to K^+K^-$ and $D \to \pi^+\pi^-$ decays is considered in Sec. 4.1.2.

4. The ratio of $B^\pm \to D^*K^\pm$ to $B^\pm \to D^*\pi^\pm$ decays for $D \to hh$, where $hh =$

$\pi\pi/KK$, divided by $R_{K/\pi}^{K\pi,\pi^0/\gamma}$:

$$R^{CP,\pi^0/\gamma} = \frac{\Gamma(B^- \to ([h^-h^+]_D\pi^0/\gamma)_{D^*}K^-) + \Gamma(B^+ \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}K^+)}{\Gamma(B^- \to ([h^-h^+]_D\pi^0/\gamma)_{D^*}\pi^-) + \Gamma(B^+ \to ([h^+h^-]_D\pi^0/\gamma)_{D^*}\pi^+)} \times \frac{1}{R_{K/\pi}^{K\pi,\pi^0/\gamma}}. \quad (1.35)$$

Here, $R^{KK,\pi^0/\gamma} = R^{\pi\pi,\pi^0/\gamma} = R^{CP,\pi^0/\gamma}$, where direct CP violation in D decays is considered negligible.

5. The ratio of the ADS to the favoured D decay mode, for $B^\pm \to D^*h^\pm$ decays:

$$R_{\pi^\pm}^{\pi K,\pi^0/\gamma} = \frac{\Gamma(B^\pm \to [[K^\mp\pi^\pm]_D\pi^0/\gamma]_{D^*}K^\pm)}{\Gamma(B^\pm \to [[K^\pm\pi^\mp]_D\pi^0/\gamma]_{D^*}K^\pm)} \quad (1.36)$$

$$R_{K^\pm}^{\pi K,\pi^0/\gamma} = \frac{\Gamma(B^\pm \to [[K^\mp\pi^\pm]_D\pi^0/\gamma]_{D^*}\pi^\pm)}{\Gamma(B^\pm \to [[K^\pm\pi^\mp]_D\pi^0/\gamma]_{D^*}\pi^\pm)}. \quad (1.37)$$

The individual ratios for each B charge are measured, as opposed to the asymmetry $A_h^{\pi K,\pi^0/\gamma}$ and summed over charge ratio $R_h^{\pi K,\pi^0/\gamma}$, because the latter are statistically correlated; they are therefore calculated by hand using the formulae:

$$R_h^{\pi K,\pi^0/\gamma} = \frac{R_{h^-}^{\pi K\pi^0/\gamma} + R_{h^+}^{\pi K\pi^0/\gamma}}{2} \quad (1.38)$$

$$A_h^{\pi K,\pi^0/\gamma} = \frac{R_{h^-}^{\pi K\pi^0/\gamma} - R_{h^+}^{\pi K\pi^0/\gamma}}{R_{h^-}^{\pi K\pi^0/\gamma} + R_{h^+}^{\pi K\pi^0/\gamma}}. \quad (1.39)$$

1.4.5 Previous Measurements

Decays of the kind $B^\pm \to D^*h^\pm$ have been successfully studied at LHCb using the *partially reconstructed* method [31]. In this analysis, the neutral particle produced in the D^* decay was not reconstructed, therefore these events could be found at low mass in the fitting variable $m(Dh)$. The analysis was performed using the full LHCb dataset, and an example invariant-mass fit to data is shown in Fig. 1.5 for B^- candidates, where the D meson is reconstructed in the K^+K^- final state. The large peak sitting at the nominal B mass, $(5279.34 \pm 0.12)\,\text{MeV}/c^2$, is due to traditionally studied $B^- \to DK^-$ decays; the dark blue, double horned structure in the low mass region $[5000, 5150]\,\text{MeV}/c^2$ is due to $B^- \to (D^* \to D\pi^0)K^-$ decays; the light blue,

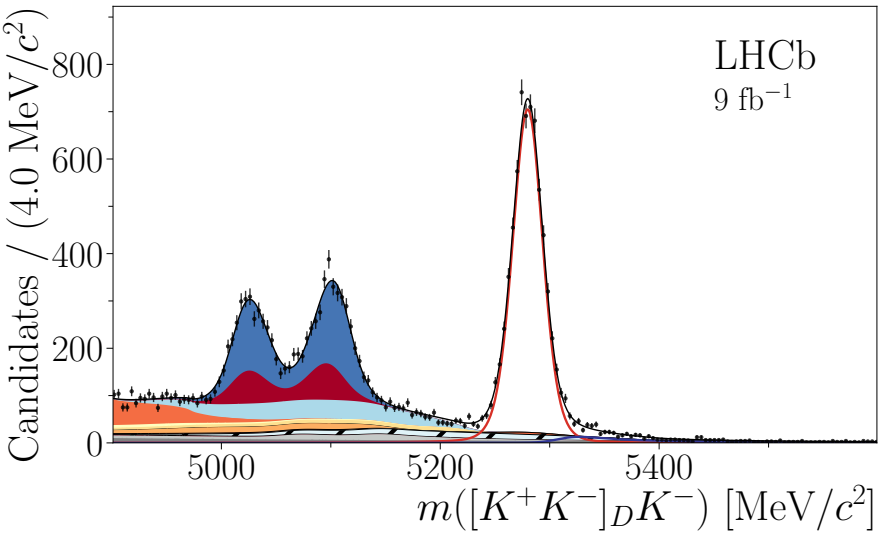


Figure 1.5: Invariant-mass fit to the $m(DK)$ distribution of $B^- \to D^{(*)}K^-$ data, where the D meson is reconstructed in the K^+K^- final state, performed by the partially reconstructed analysis [31].

broad peak sitting underneath this is from $B^- \to (D^* \to D\gamma)K^-$ decays. The characteristic *horns* and *hill* nature of the D^* modes, respectively, is due to the angular distributions of the decay products.

The partially reconstructed (PR) analysis benefited from high statistics, as it was unaffected by the low neutral reconstruction efficiencies at LHCb (which will be covered in Sec. 2.3.3). Measurements of *CP* observables for the $D^* \to D\pi^0$ mode have been made with competitive precision, however the $D^* \to D\gamma$ final state suffered from low purity due to the many overlapping backgrounds sitting underneath the broad signal peak.

This is where the work presented in this thesis comes in, which looks to measure the same observables using the *fully reconstructed* method, by including the neutral particle in the final state. The fully reconstructed (FR) technique compromises on statistics with the aim of achieving increased signal purity. The results of the PR analysis will be given and compared to the FR results of this thesis in Sec. 4.6. It is worth noting that selected FR candidates were removed from the PR data sample, so the analyses are statistically independent.

The most precise measurements of the hadronic parameters $r_B^{D^*h}$ and $\delta_B^{D^*h}$, given by the latest LHCb γ combination [23], come from the results of the PR analysis:

$$r_B^{D^*K} = (9.9^{+1.6}_{-1.9})\% \tag{1.40}$$

$$\delta_B^{D^*K} = (310^{+12}_{-23})^\circ \tag{1.41}$$

$$r_B^{D^*\pi} = (0.95^{+0.85}_{-0.61})\% \tag{1.42}$$

$$\delta_B^{D^*\pi} = (139^{+22}_{-86})^\circ. \tag{1.43}$$

GLW analyses of $B^\pm \to D^*K^\pm$ decays have also been successfully conducted by the Belle [41] and BaBar [42] collaborations. These results, combined with measurements using the BPGGSZ method [43], were used to produce the combination [44]:

$$r_B^{D^*K} = (10.6^{+1.9}_{-3.6})\% \tag{1.44}$$

$$\delta_B^{D^*K} = (294^{+21}_{-31})^\circ, \tag{1.45}$$

which shows good agreement with the latest LHCb measurements.

2

The LHCb experiment

The reconstruction of $B^{\pm} \to (D^* \to [h^{\pm}h^{\mp}]_D \pi^0/\gamma)h^{\pm}$ decays was performed by the LHCb detector [45], a single-arm forward spectrometer located on the ring of the Large Hadron Collider (LHC), a 27 km circular pp collider in Geneva. Bunches of $\mathcal{O}(10^{11})$ protons, obtained from ionising hydrogen atoms, are accelerated through a series of linear and circular accelerators, depicted in Fig. 2.1, before being injected into the LHC in opposite directions. During the years 2011 and 2012, know as Run 1, the beams were accelerated to centre of mass energies of $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$ respectively. During the years $2015 - 2018$, known as Run 2, the beams were accelerated to $\sqrt{s} = 13\,\text{TeV}$. The beams are focused to collide at four points, home to the four experiments located on the ring: ATLAS, CMS, ALICE and LHCb, with collisions occurring at a frequency of 20 (40) MHz in Run 1 (Run 2). The protons continue to circulate for $\sim$ 12 hours, known as a *fill*.

LHCb has been designed to study b and c quarks, which are predominantly produced from gluon-gluon fusions with asymmetric momenta in the laboratory frame, leading to $c\bar{c}$ or $b\bar{b}$ pairs with boosted centre of mass energies. For this reason, LHCb is a forward spectrometer covering the pseudorapidity region $1.9 < \eta < 4.9$, where $\eta = -\ln(\tan\theta/2)$. Here, θ is the angle to the beam axis. In terms of angular acceptance, this equates to 15 to 300 (250) mrad in the horizontal (vertical) plane,

capturing $24 - 25\%$ of all $b\bar{b}$ pairs produced at the interaction point [46].

In order to maximise the data collection efficiency whilst maintaining the necessary precision and purity of secondary vertex reconstruction of heavy mesons, the number of pp interactions is reduced at the interaction point of LHCb by reducing the transverse overlap of the LHC beams (a process known as *luminosity levelling*). The number of pp collisions was limited to 1.7 (1.1) interactions per bunch crossing during Run 1 (Run 2), resulting in an instantaneous luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The total integrated luminosity collected by the LHCb detector was $3\,\mathrm{fb}^{-1}$ in Run 1 and $5.7\,\mathrm{fb}^{-1}$ in Run 2, providing a total of $8.7\,\mathrm{fb}^{-1}$. The measurements presented in this thesis are based on the full LHCb dataset.

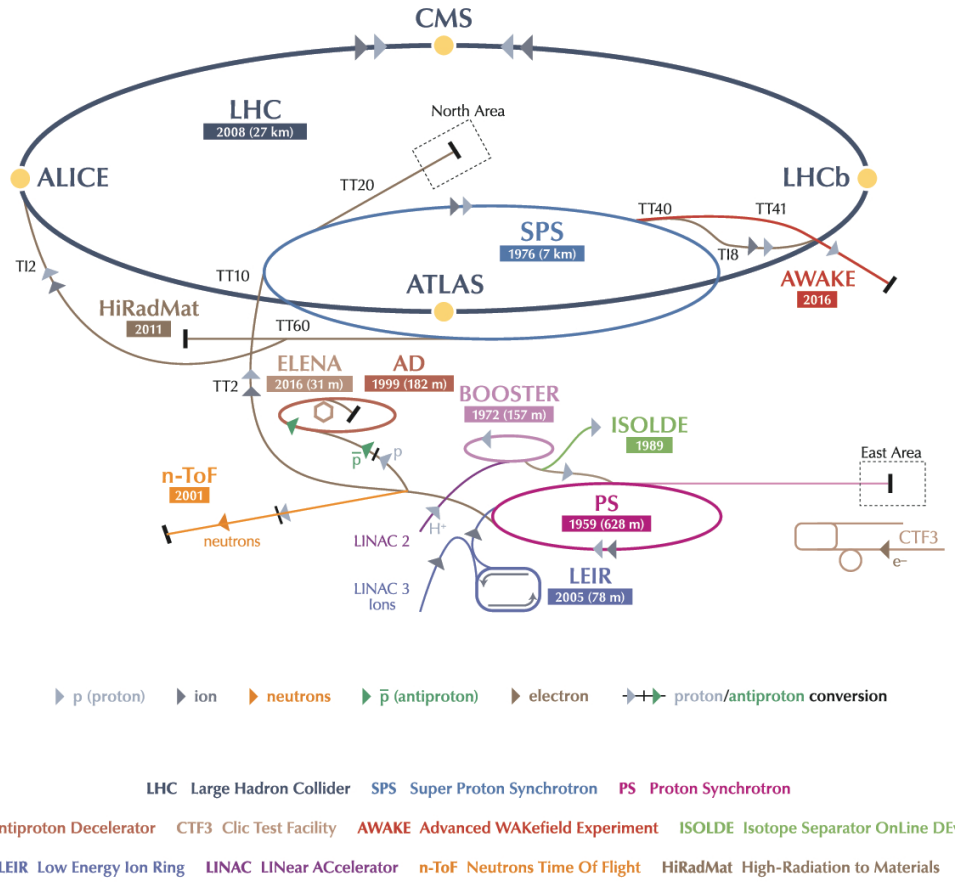


Figure 2.1: The CERN accelerator complex, including the construction year and length for a number of accelerators, reproduced from Ref. [47]. The accelerator chain used during pp operation is: LINAC 2 $\rightarrow$ BOOSTER $\rightarrow$ PS $\rightarrow$ SPS $\rightarrow$ LHC.

2.1 The LHCb sub-detectors

The coordinate system for detector description is such that the z-axis lies along the beamline, and the x-axis (y-axis) in the horizontal (vertical) direction in the plane perpendicular to this. The origin is located at the collision point. The LHCb sub-detectors, which will be described in the text below, are positioned along the beamline from the interaction point up to $z = 20\,\mathrm{m}$, as depicted in Fig. 2.2. For the rest of this chapter, *upstream* refers to the positive z direction, *downstream* to the negative z direction.

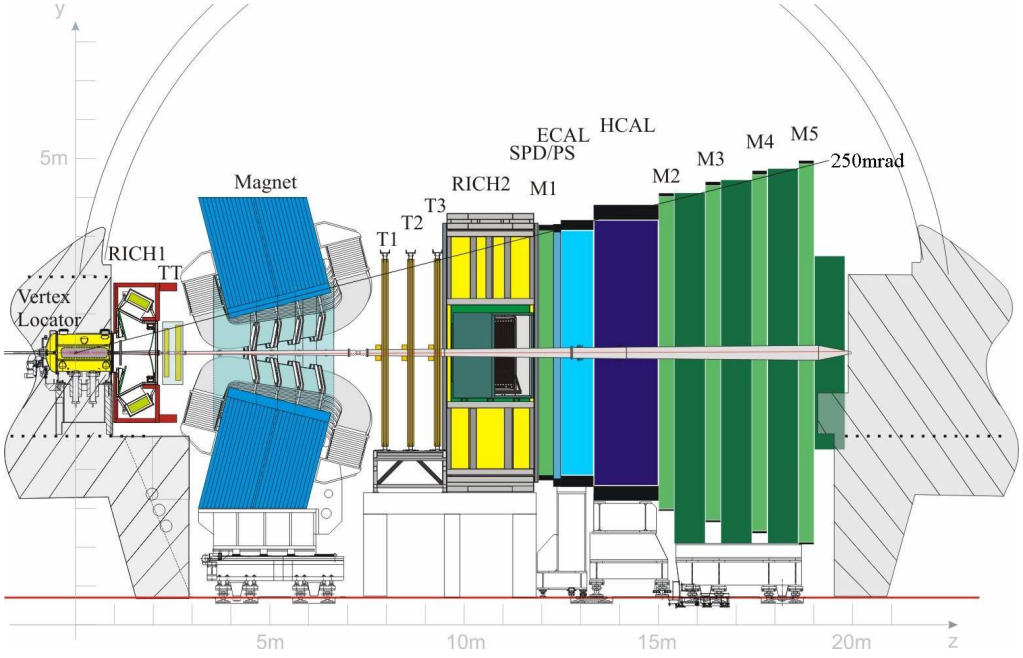


Figure 2.2: Overview of the LHCb detector, reproduced from Ref. [45].

2.1.1 VErtex LOcator (VELO)

The VELO is a silicon micro-strip detector located very close to the interaction region. It provides crucial coordinate information for decays of b and c hadrons, facilitating the reconstruction of primary (PV) and secondary vertices. Of importance is also the calculation of impact parameters (IP), the distance of closest approach between tracks and their PVs.

There are 21 VELO stations positioned along the beamline from $z = -18\,\text{cm}$ to $z = 88\,\text{cm}$, where the nominal beam interaction point is positioned at $z = 0 \pm 5\,\text{cm}$, as displayed in Fig. 2.3. Each station consists of 2 semi-circular modules mounted on either side of the beamline, and each module has 2 silicon strip detectors designed to measure the azimuthal angle of the track and the radial distance from the beam. The z coordinate is aligned with the beam direction and can be inferred by the fixed module position. The silicon in $\mathcal{R}$ sensors is arranged in concentric semi-circles; in ϕ sensors radial strips run from the inner to outer radius of the module, as illustrated in Fig. 2.4. The strip pitch varies between 40 µm and 100 µm depending on the distance from the beam line.

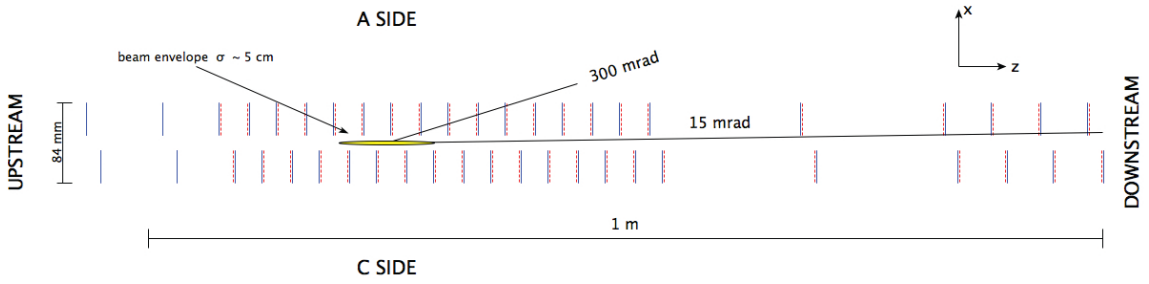


Figure 2.3: Cross-section in the $x - z$ plane at $y = 0$ of the VELO stations in their closed position. $\mathcal{R}$ (ϕ) sensors are shown with solid blue (dashed red) lines. The modules at positive (negative) x are known as the left or A-side (right or C-side). Reproduced from Ref. [48].

The sensitive area of the silicon sensors starts 8 mm from the LHC beams, therefore the modules are attached to two retractable halves that are opened during injection, to shield from radiation damage during periods of unstable beam, and closed for data-taking at the beginning of *stable beams*. To align with the other sub-detectors, the VELO has an angular acceptance of $15 - 300\,\text{mrad}$ downstream, so that tracks pass through at least 3 VELO stations within this coverage.

The PV resolution achieved by the VELO is typically $\sim 10\,\mu\text{m}$ in x/y and $\sim 50\,\mu\text{m}$ in z. The impact parameter resolution is $< 35\,\mu\text{m}$ for particles with $p_\text{T} > 1\,\text{GeV}/c$ [48], offering discriminatory power between particles produced from primary and secondary vertices.

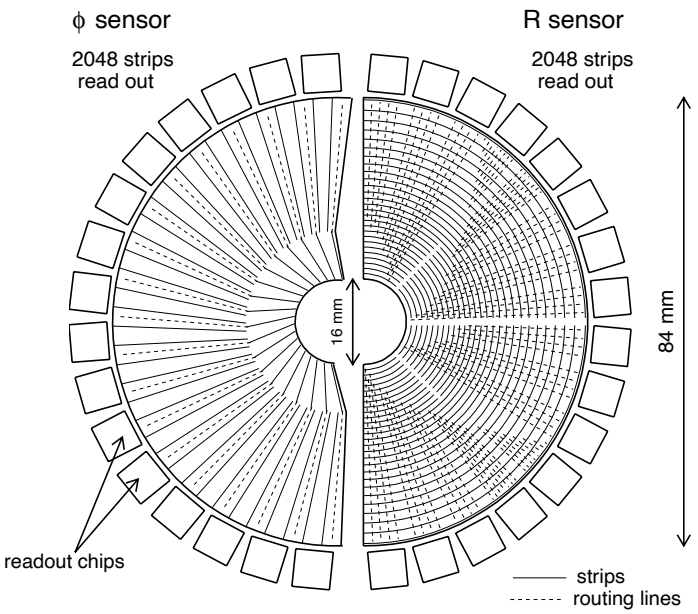


Figure 2.4: Illustration of the silicon strip layout for a VELO module, designed to measure (left) the azimuthal angle, ϕ, of a track and (right) the radial distance, r, from the beam. Reproduced from Ref. [49].

2.1.2 Dipole magnet

In order to measure the momentum of charged particles, a warm (non-superconducting) dipole magnet consisting of 2 sets of aluminium coils within an iron yoke is positioned downstream of the VELO and the first RICH detector, RICH1. The principle component of the magnetic field is aligned along the vertical (y) direction, deflecting particles in the horizontal ($x - z$) plane [50]. The magnetic field strength varies along the z direction, as shown in Fig. 2.13, reaching a maximum strength of 1 T. The field has been measured with a relative precision of $\delta B/B \approx 4 \times 10^{-4}$ and is uniform in the $x - y$ plane within 1% in the tracking volume [51]. The total bending power provided is 3.6 Tm over $z \in [2.5, 8]$ m.

Within each year of data taking, the magnetic field spends approximately half the time oriented along the positive y direction in the "Up" configuration, and half oriented along the negative y direction in the "Down" configuration. Positive and negatively charged particles bend in opposite directions in this field, therefore this procedure significantly reduces charge detection asymmetries from small differences

in the left and right parts of the detector.

2.1.3 Tracking stations

The LHCb tracking system is comprised of the VELO and 4 additional tracking stations: the Tracker Turicencis (TT) upstream of the magnet, and the tracking stations 1-3 (T1, T2, T3) downstream of the magnet.

The TT is a planar silicon microstrip detector with pitch $\sim 200\,\mu\text{m}$, giving a spatial resolution of $50\,\mu\text{m}$. It has 4 layers of detector material: the first and last have strip readouts with vertical orientation (x-layers); the second and third are oriented at $+5°$ and $-5°$ respectively (u- and v- layers). This $x-u-v-x$ geometry allows the transverse components of particle trajectories to be resolved. The front layer is depicted in Fig. 2.5. The TT makes it possible to reconstruct long-lived K^0_S mesons, which decay after the VELO. It also facilitates the reconstruction of low momentum tracks that are swept out of detector acceptance by the magnet before reaching the downstream tracking stations (illustrated in Fig. 2.13).

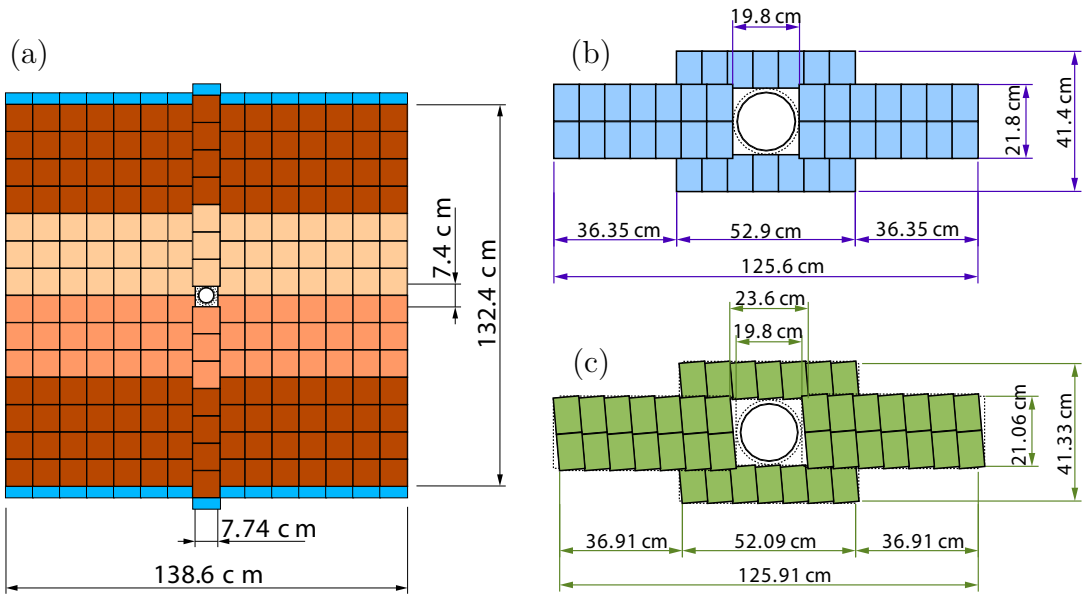


Figure 2.5: Schematic of (a) an x-layer module of the TT, (b) an x-layer module of the IT and (c) a v-layer module of the IT [52].

The tracking stations T1-T3 are split into two sections: a silicon based inner tracker (IT) and an outer tracker (OT) that employs drift tubes. The IT is made up

of 4 modules surrounding the central region, illustrated in Fig. 2.5, where particle fluxes are highest. It uses the same silicon microstrip technology as the TT, arranged in 4 layers oriented using the $x-u-v-x$ geometry. The IT and TT are collectively known as the Silicon Tracker (ST).

The OT covers the large area of LHCb acceptance not occupied by the IT. 4 layers of vertically oriented drift tubes, arranged in the same $x-u-v-x$ geometry, cover an area of approximately $5 \times 6\,\mathrm{m}^2$ starting $\sim 10\,\mathrm{cm}$ from the beamline. This set up is illustrated in Fig. 2.6. Each layer has 64 drift tubes of 4.9 mm in diameter, filled with an $Ar/CO_2/O_2$ (70/28.5/1.5) gaseous mixture, which ensures a drift time $< 50\,\mathrm{ns}$. The coordinate resolution of the OT was 205 μm (171 μm) during Run 1 (Run 2) [53].

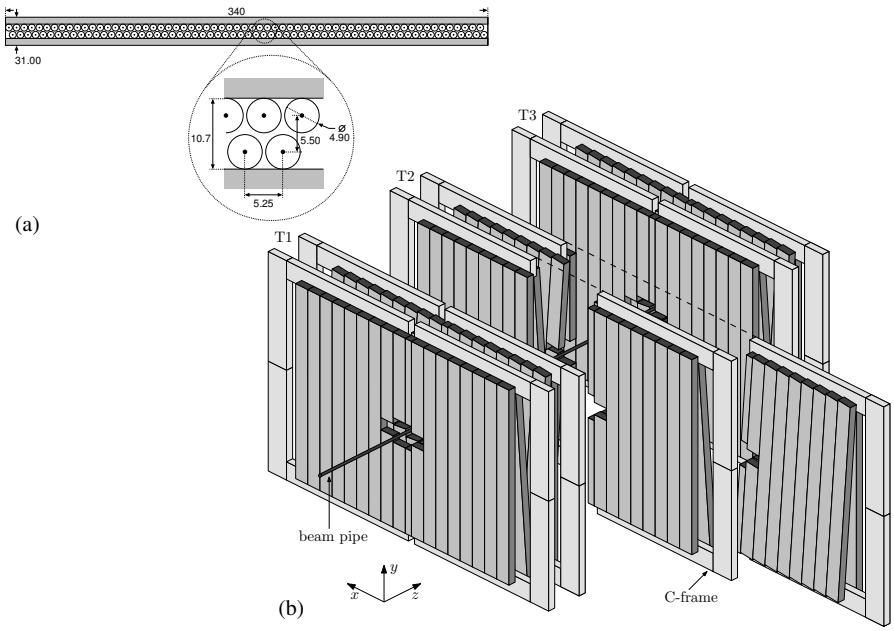


Figure 2.6: The cross section of an OT module is illustrated in (a); the arrangement of OT modules in T1-T3 is shown in (b). Reproduced from Ref. [54].

For charged tracks that pass through the entire tracking system (known as long tracks), the overall relative momentum resolution achieved is $< 1\%$ for $p < 200\,\mathrm{GeV}/c$. This was calculated using data from $J/\psi \to \mu^+\mu^-$ decays, for which the momentum resolution can be related to the invariant mass resolution of the J/ψ candidate. The resolution over the full momentum spectra is plotted in Fig. 2.7.

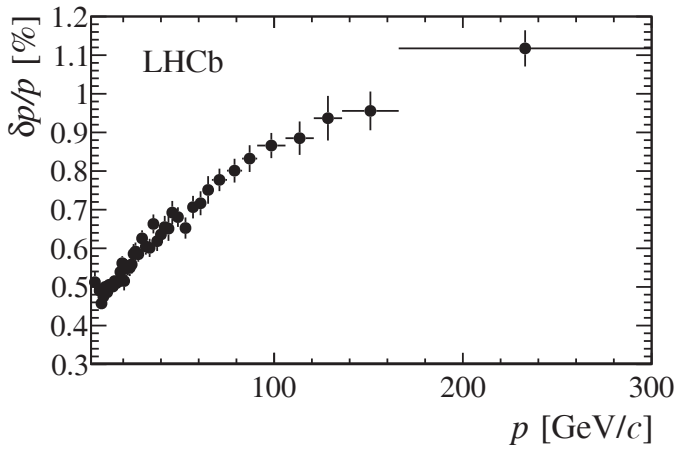


Figure 2.7: Relative momentum resolution plotted against momentum for charged particles that pass through the entire tracking system [51], measured using $J/\psi \to \mu^+\mu^-$ decays in data.

2.1.4 Ring Imaging CHerenkov detectors (RICH)

The identification of different particle species is an important component of flavour physics; for this thesis in particular, the ability to separate pions and kaons is crucial. At the LHCb experiment, this is made possible by the RICH detectors [55]. The RICH1 detector sits between the VELO and TT, targetting tracks in the momentum range $p \in [1, 60]\,\text{GeV}/c$; the RICH2 detector is positioned upstream of the tracking system and targets tracks in the momentum range $p \in [15, 100]\,\text{GeV}/c$. RICH1 covers the full LHCb acceptance; RICH2 only operates in the region where higher momentum tracks lie, $\theta < 120(100)\,\text{mrad}$ in the horizontal (vertical) plane. An overview of both detectors is given in Fig. 2.8.

Cherenkov radiation is emitted by a particle passing through a medium at a speed v, which is greater than the phase velocity of light in that medium. Light is emitted in a cone, with angle θ_c with respect to the particle trajectory, defined by Eq. (2.1):

$$\cos\theta_c = \frac{1}{n\beta}, \tag{2.1}$$

where $\beta = v/c$ and n is the refractive index of the RICH radiator medium. A measurement of β, along with information of the track momentum from the tracking system, allows the particle mass and therefore species to be determined.

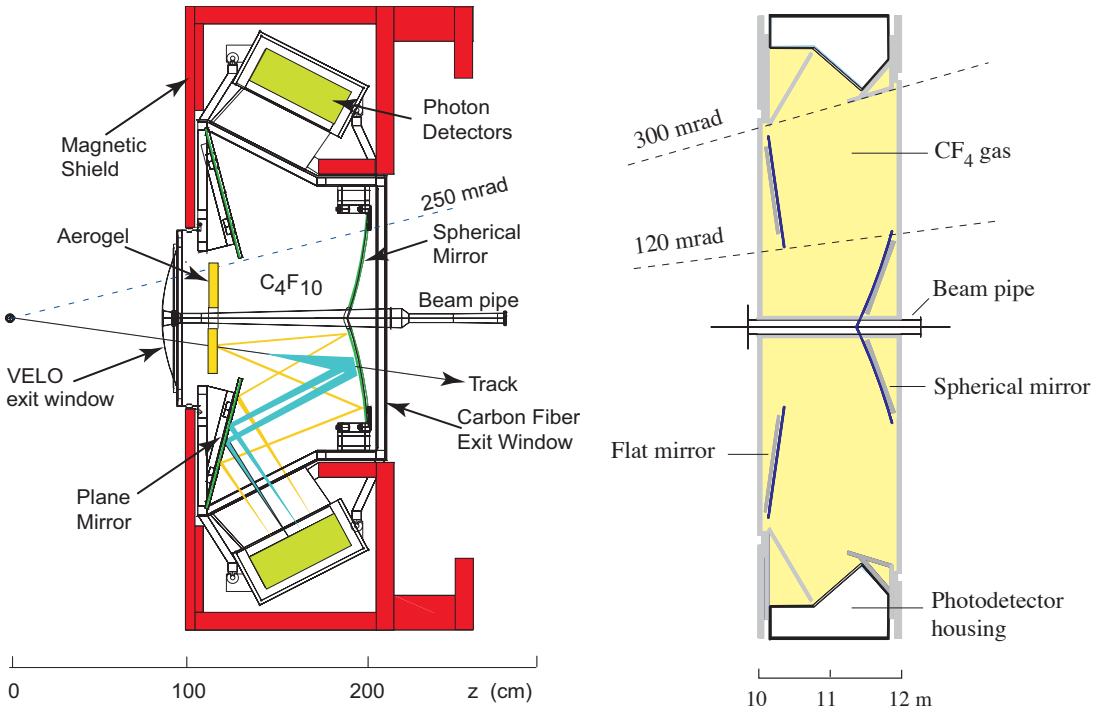


Figure 2.8: Illustration of the RICH 1 (left) and RICH 2 (right) detectors, reproduced from Refs. [45, 56].

The threshold for emission depends on n and the particle momentum. Sensitivity is lost at high momentum when $\beta \to 1$ as $\cos\theta_c \to 1/n$. Different radiator mediums are therefore chosen for RICH1 and RICH2 to provide coverage for all particle types within the designated momentum spectra. The first detector is filled with a C_4F_{10} radiator, $n = 1.0014$; the second with a CF_4 radiator, $n = 1.0005$. During Run 1, an Aerogel radiator with $n = 1.03$ was also installed to provide particle identification (PID) at very low momenta, however this was removed in Run 2 as it brought minimal performance benefits at the cost of Cherenkov photon yield [55, 57].

A series of curved and flat mirrors are designed so that the light from incident tracks is focused conically onto Hybrid Photon Detectors (HPDs), outside the LHCb acceptance. Incident light causes the emission of photoelectrons in the HPDs, which are focused onto a silicon chip using an electrostatic field. The final photoelectron positions on the chip are used to deduce the location of incident photons on the HPDs and therefore calculate θ_c.

The separation power of RICH1 can be seen in Fig. 2.9 for particle types (p, K, π, μ), where the Cherenkov angle is plotted against particle momentum. Details of the PID variables used in this analysis are given in Sec. 2.3.2.

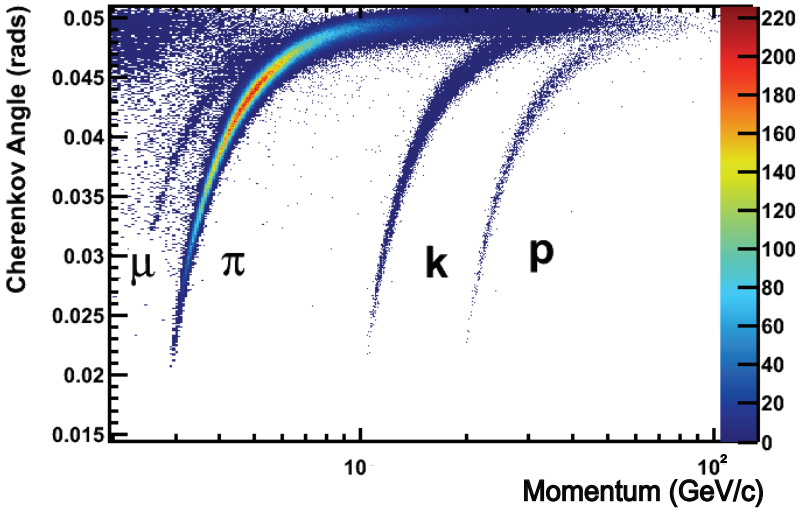


Figure 2.9: Cherenkov angle *vs.* track momentum for isolated tracks in the RICH1 C_4F_{10} radiator [57].

2.1.5 Calorimeters

The calorimeter system is involved in the L0 trigger, particle identification and, of significance for this analysis, the reconstruction of neutral particles (see Sec. 2.3.3). The electromagnetic (ECAL) and hadronic (HCAL) calorimeters are preceded by the Silicon Pad Detector (SPD) and Preshower (PS) detector, separated by a lead wall (Fig. 2.10) [45]. There is a consistent design across all calorimeter components: metal absorbers are interspaced with organic scintillator plates; scintillation light is collected by wavelength-shifting fibres (WLS) and transported to photomultiplier tubes (PMTs); the PMTs convert light into electrical signals, which are read out and digitised by Front End Boards (FEB).

The hit density varies by two orders of magnitude over the calorimeter surface: it is a steep function of distance from the beampipe. The SPD, PS and ECAL are segmented in the plane perpendicular to the beam axis to account for this variation, providing a one-to-one cell correspondence. The cell size is $4.04 \times 4.04\,\mathrm{cm}^2$ in the

inner region, $6.06 \times 6.06\,\mathrm{cm}^2$ in the middle region and $12.12 \times 12.12\,\mathrm{cm}^2$ in the outer region. The HCAL, on the other hand, has only an inner and an outer region with cells of size $13.13 \times 13.13\,\mathrm{cm}^2$ and $26.26 \times 26.26\,\mathrm{cm}^2$ respectively, due to the dimensions of hadronic showers [45].

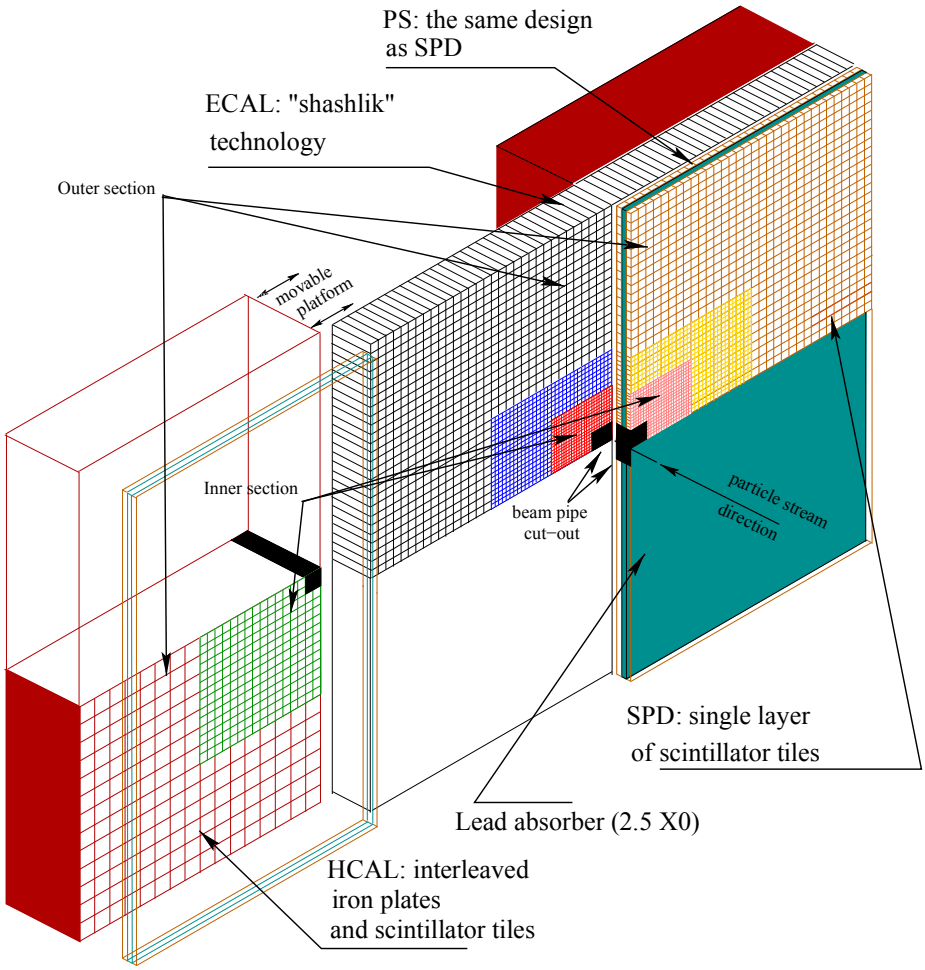


Figure 2.10: Layout of the calorimeter system, taken from Ref. [58].

Scintillating pad/preshower detector

The SPD/PS detector consists of two identical scintillating pads separated by a 15 mm lead absorber, corresponding to 2.5 radiation lengths (X_0). Each pad contains 6016 cells with granularity described above, where signal is read out by multi-anode PMTs [45].

The SPD delivers binary information per cell, dependent on whether the energy

deposited is over a certain threshold. This identifies whether incident particles are charged (signal) or neutral (no signal) [58]. The PS distinguishes between electromagnetic and hadronic particles by looking at the longitudinal segmentation of the electromagnetic shower after the lead absorber. Using this procedure, photons, electrons and pions can be identified by their energy deposition sequence. When evaluating the energy of electrons, photons and neutral pions, the deposits in the PS are combined with their projected partners in the ECAL.

Electromagnetic calorimeter

The ECAL has a shashlik structure with 66 alternating layers of 4 mm scintillator and 2 mm lead absorber. The Molière radius of the ECAL is 3.5 cm. The depth of the 42cm Pb/scintillator stack corresponds to 25 X_0, which ensures optimal energy resolution and full containment of high energy electromagnetic showers. The outer acceptance of the ECAL matches that of the tracking system: $\theta_x < 300\,\text{mrad}$, $\theta_y < 350\,\text{mrad}$. The inner acceptance, limited by radiation levels close to the beampipe, is $\theta_x, \theta_y < 25\,\text{mrad}$. There are a total of 6016 ECAL cells over the inner, middle and outer regions; each cell is read out by a single PMT [45], as shown in Fig. 2.11.

The ECAL energy resolution, measured using test-beam electrons [59], is parameterised by Eq. (2.2), where the particle energy E is measured in GeV and θ is the angle between the beam axis and the line from the interaction point to the cell centre [58]

$$\frac{\sigma(E)}{E} = \frac{(9.0 \pm 0.5\%)}{\sqrt{E}} \oplus (0.8 \pm 0.2)\% \oplus \frac{0.3}{E \sin\theta}. \tag{2.2}$$

The first term accounts for fluctuations in the number of signal generating processes, for example the number of photo-electrons produced by PMTs. The second, constant term results from imperfections in calorimeter construction, for example the loss of particle energy in dead material, mis-calibrations and non-uniform response across the detector. The third term is due to electronic noise. For more realistic LHCb operating conditions, considering event pile-up, material before the ECAL and non-isolated photons, the energy resolution is closer to $\sigma(E)/E = (9.0 \pm 0.5\%)/\sqrt{E} \oplus 4\%$.

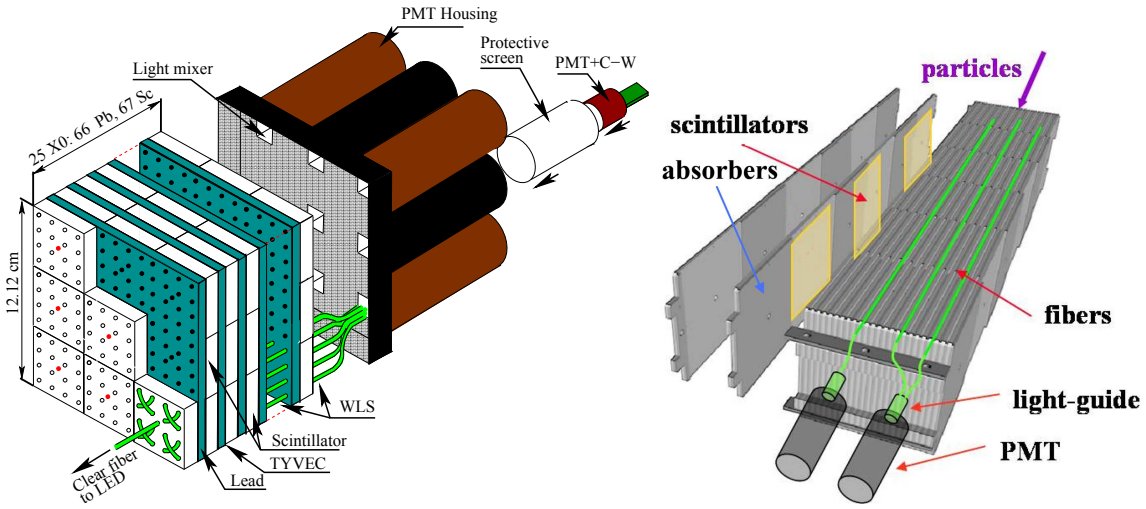


Figure 2.11: Illustration of an ECAL cell (left) and HCAL cell (right), reproduced from Refs. [45, 58].

Hadronic calorimeter

The HCAL has a sampling structure, with alternating 1 cm iron plates and scintillating tiles oriented parallel to the beamline, enhancing light collection. This layout is illustrated in Fig. 2.11. There are a total of 1488 cells in the HCAL, over the inner and outer regions, and each cell is read out by a single PMT [45].

Unlike the other three calorimeter components, which are involved in neutral particle identification and reconstruction, the HCAL is only used for the hardware L0 triggering (see Sec. 2.2) in this work. The L0 trigger does not require good hadron energy resolution, therefore HCAL thickness it limited to 5.6 interaction lengths due to space limitations. The energy resolution achieved is given by Eq. (2.3), where the particle energy E is measured in GeV [58].

$$\frac{\sigma(E)}{E} = \frac{(67 \pm 5\%)}{\sqrt{E}} \oplus (9 \pm 2)\%. \tag{2.3}$$

As in Sec. 2.1.5, the first term accounts for fluctuations in the number of signal generating processes and the second, constant term results from imperfections in calorimeter construction.

2.1.6 Muon stations

Muons are the only particles that penetrate the full calorimeter system (excluding neutrinos), therefore the muon detection system is the final downstream sub-detector. Five rectangular tracking stations, M1-M5, covering the full LHCb acceptance are dedicated to muon identification and triggering [60]. The first muon station, M1, is located upstream of the calorimeter system and provides p_{T} information for the Level 0 muon trigger. The final four stations, M2-M5, are used to identify and trace muons for online and offline analysis and are positioned downstream of the calorimeters. They are interspaced with 80 cm think iron absorbers to select only penetrating muons. This layout is depicted in Fig. 2.12.

All muon stations employ Multi-Wire Proportional Chambers (MWPCs) to track traversing particles [61]. The only exception is the central region of the M1 station, which utilises gas-electron multiplier detectors around the beamline due to higher track multiplicity and radiation levels in this region [62].

Events for which the companion particle of the signal decay triggers the `isMuon` line are removed from this analysis in order to target semi-leptonic backgrounds, as will be described in Sec. 3.3.8. However, events containing muons with p_{T} above a certain threshold, which by definition pass the global L0 trigger (described in Sec. 2.2.1), are included in the dataset which is analysed. This is because it is possible for an event to be considered if it is triggered *independent* of signal, *i.e.* due to other particles produced in the pp interaction.

2.2 The LHCb Trigger

During Run 1 (Run 2), the LHC operated with a collision frequency of approximately 20 MHz (40 MHz). At the interaction point, the transverse overlap of the 2 beams is reduced so that, on average, 1.6 (1.1) pp interactions occur per bunch crossing in Run 1 (Run 2). To reduce the rate of data uptake, online event selection consists of a hardware and software trigger. The hardware trigger uses information from

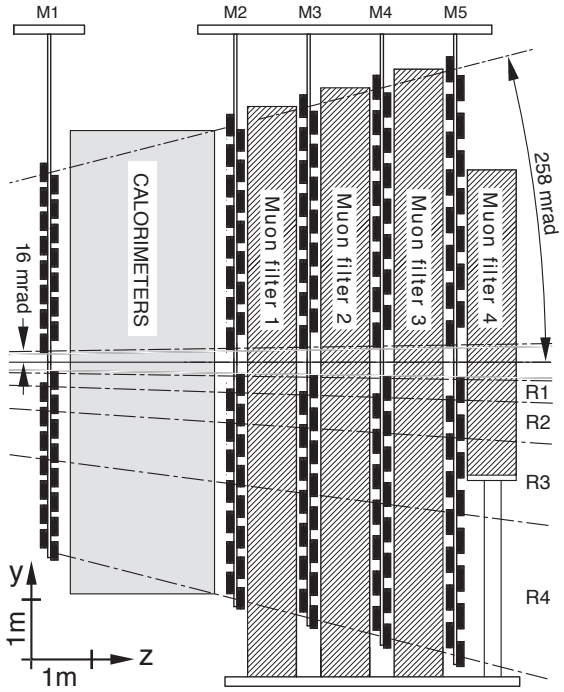


Figure 2.12: Layout of the muon chambers (black dashed lines) and iron absorbers (grey areas) [60]. M1 is placed before the calorimeter system to improve the p_{T} measurement, which is used in the Level 0 muon trigger.

the calorimeter and muon systems, and reduces the rate of information that needs to be stored down to 1 MHz, allowing latency for the full detector to be read out. This is required for the software trigger, which involves a full event reconstruction. Events that pass the software trigger are read out at a rate of 3 kHz (2011), 5 kHz (2012) and 12.5 kHz (Run 2).

The data considered in this thesis has to pass a set of *trigger lines*. Each line consists of a sequence of algorithms and thresholds, designed to select for specific decays, and returns an accept or reject decision. Events must pass at least one hardware and software line to be retained.

2.2.1 L0 trigger

There is insufficient processing power to store events at the nominal bunch crossing rate of 20 (40) MHz in Run 1 (Run 2), therefore the level-0 (L0) trigger is tasked with reducing this down to the more manageable level of 1 MHz. The large mass

of the B meson means that its decay products are typically of high p_{T} relative to the large background of soft QCD processes. The L0 trigger exploits this decay signature by using quick readout from the calorimeter and muon systems to select high p_{T} events. This selection is achieved by splitting the ECAL and HCAL into 2×2 cell clusters with transverse energy:

$$E_T = \sum_{i=1}^{4} E_i \sin\theta_i, \tag{2.4}$$

where θ_i is the cell angle with respect to the beam axis and the average collision point.

The trigger line `L0Hadron` is passed if the highest HCAL transverse energy deposit in the event is greater than 3.68 GeV. The lines `L0Electron` and `L0Photon` are triggered by the ECAL if the highest transverse energy deposit in the event meets the threshold $E_T > 3\,\mathrm{GeV}$. The identification of the deposit as an electron or photon depends on whether there is an associated SPD hit. The triggering of muons is based on whether an approximate straight line can be draw from the interaction point to hits in the muon stations. The minimum p_{T} for the line `L0Muon` is $p_{\mathrm{T}} > 1.76\,\mathrm{GeV}/c$; the p_{T} threshold for the product of two muons to trigger the line `L0DiMuon` is $p_{\mathrm{T}} > 1.76\,\mathrm{GeV}/c$. The `L0Global` line used in this thesis is triggered if an event contains at least one candidate with E_T above the described thresholds.

High multiplicity events that would take too long to process by the high-level trigger (HLT) are additionally discarded by rejecting events with more than 1000 SPD hits.

2.2.2 High-level trigger

The output of the L0 trigger is processed by the HLT on the Event Farm Filter (EFF): a farm consisting of 900 (1700) multiprocessor computing nodes during Run 1 (Run 2). The HLT consists of two stages: HLT1 and HLT2.

There is enough latency for the full detector to be read out at 1 MHz for a partial event reconstruction to be performed by HLT1. Tracks that pass through all elements of the tracking system (long tracks) and have a p_{T} greater than a certain threshold are built, and tracks in the VELO are used to determine the position of primary vertices (PVs). The trigger line `HLT1TrackAllL0` (re-optimised and renamed

`HLT1TrackMVA` in Run 2), an inclusive trigger that selects for high p_{T} tracks with significant displacement from the PV, is used to filter data in this thesis. When selecting Run 2 data, an additional inclusive trigger `HLTTwoTrackMVA` is employed: a multivariate classifier trained on track and vertex properties that is applied to two pronged vertices formed from high p_{T} tracks not originating from a PV. These lines are specific to this thesis, but other triggers designed to select muons, calibration data, low-multiplicity events and a number of exclusive lines exist. The overall effect of HLT1 is to reduce the event rate further, down to 40 (110) kHz in Run 1 (Run 2), so that a full event reconstruction can be performed by HLT2.

The full event reconstruction of HLT2 was also based on long tracks with high momentum during Run 1. During Run 2, a complete, fully aligned reconstruction was performed, made possible by the increased number of multiprocessor nodes on the EFF. The data considered in thesis was filtered by inclusive 'topological' HLT2 lines. Tracks that satisfy fit quality requirements, have high p_{T} and are displaced from the PV are combined one-by-one and identified as having either 2-, 3-, or 4-body topology depending on their distance of closest approach, which must be $< 0.2\,\mathrm{mm}$. A multivariate classifier, which discriminates using invariant mass combinations and event topology, is applied to the $n-$body object to see whether it should be accepted or rejected. These lines are denoted `HLT2Topo2,3,4BodyBBDT` (`HLT2Topo2,3,4Body`) for Run 1 (Run 2) data. Events that pass HLT2 lines are stored permanently, ready for offline analysis.

Improvements in and re-optimisation of the HLT before Run 2 lead to an increase in trigger efficiency. This effect, combined with the increase in $B^{\pm}$ production cross section at higher collision energies[1], increased the number of signal candidates per fb^{-1} for many analyses. The yields per fb^{-1} of LHCb data measured in this thesis are 2.5 times higher for Run 2 than for Run 1 data.

[1]The $B^{\pm}$ production cross section within LHCb acceptance is measured to be $43\,\mu\mathrm{b}$ at $\sqrt{s} = 7\,\mathrm{TeV}$ and $87\,\mu\mathrm{b}$ at $\sqrt{s} = 13\,\mathrm{TeV}$ [63].

2.2.3 Offline data filtering

For qualifying events, reconstruction is performed offline on the Worldwide LHC Computing Grid, and candidates that fulfill a set of pre-selection requirements are stored in a centrally produced dataset known as *stripping*. Each *stripping line* selects exclusively for a particular family of decays. Signal candidates are built and loose pre-selection requirements are chosen to help reduce combinatorial background, therefore reducing disk usage and CPU time required for offline analysis. For the purpose of this thesis, the combinatorial background targeted at this stage refers to the random combination of charged tracks in the event not originating from the same B meson.

2.3 Reconstruction

This section describes how charged and neutral particles are reconstructed and identified using the various LHCb sub-detectors.

2.3.1 Track reconstruction

Track reconstruction uses information from the VELO, TT and T1-T3 tracking stations to fit particle trajectories. The type of track is defined by which of these sub-detectors it traverses, where the different possibilities are illustrated in Fig. 2.13. Long tracks, which pass through the entire tracking system and have the best momentum resolution, are the only type used to form the B meson candidates considered in this thesis.

To construct long track candidates, hits are first combined in the VELO and T1-T3 tracking stations (known as T tracks) separately, as the low magnetic field within these detectors means that traversing particles have approximately straight trajectories. There are two methods that are then employed to match these segments:

1. *Forward tracking*, where VELO tracks are used as seed tracks and extrapolated to match T-tracks [64].

2. *Track matching*, where VELO tracks and T tracks are extrapolated into the bending region to determine whether they are compatible [65].

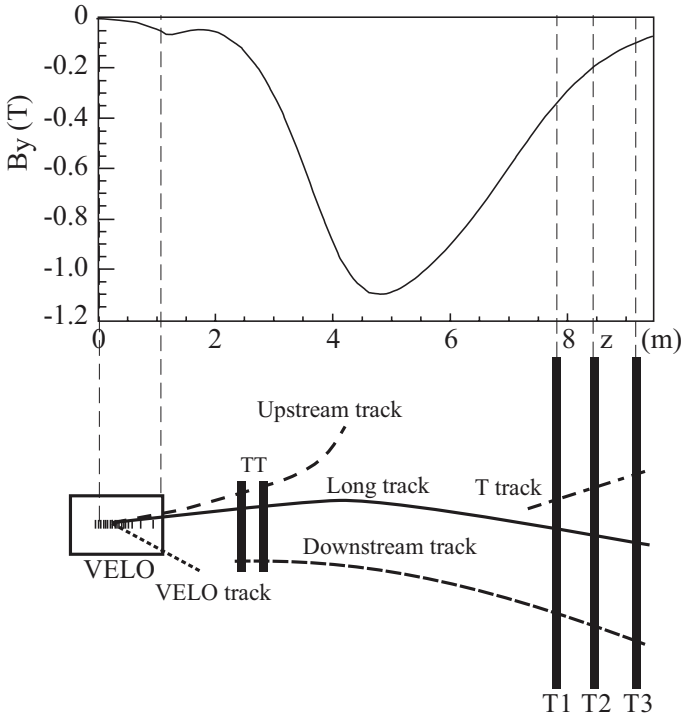


Figure 2.13: At the top of the illustration, the main B-field component (B_y) is plotted as a function of z. Below, the different types of tracks are depicted: long, upstream, downstream VELO and T tracks. This figure is taken from Ref. [51].

TT hits are added last. Tracks found using both methods are saved unless a track appears twice, then only the candidate with the best fit quality is kept. Lastly, the long tracks are refitted with a Kalman filter, which accounts for multiple scattering and energy losses due to ionisation as the particles pass through detector material [66, 67].

The particle decays considered in this thesis have 3 final state tracks and in many LHCb analyses this number is even higher, therefore it is important to have a single track reconstruction efficiency close to 100%. This is plotted in Fig. 2.14 as a function of track momentum, and as a function of the number of tracks in the event, for Run 1 data [68]. The lower efficiencies in 2012 compared to 2011 are partially due to the higher event multiplicities at higher centre-of-mass energy. Similar efficiencies have been achieved in Run 2.

The development of tools using multivariate analysis on the kinematic and track properties of ghost compared to normal tracks means that ghost tracks do not pose a problem in offline analyses and are not considered as a potential source of background

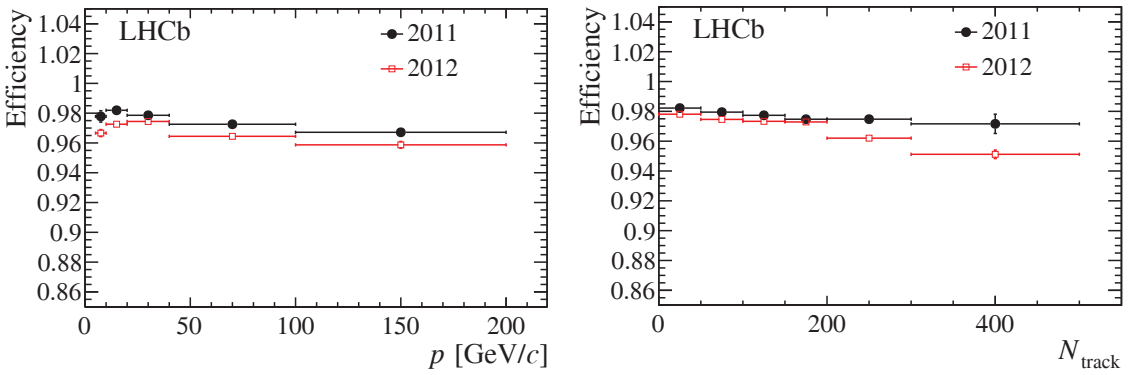


Figure 2.14: The single track reconstruction efficiency as a function of the momentum, p, (left) and the total number of tracks in the event, N_{track}, (right). This figure is reproduced from Ref. [51].

in this thesis. Here, ghost tracks refer to tracks that are not associated to any charged particle and are made up of uncorrelated real and noise hits.

2.3.2 Charged particle identification

Charged tracks for a given particle $x \in (K, p, \pi, \mu, e)$ are identified using information from the RICH detectors, calorimeters and muon stations. The absolute likelihood, L, for each particle type is constructed by multiplying the individual likelihoods from each sub-detector. For the analysis considered in this thesis it was essential to separate pions and kaons, therefore these likelihoods are considered here:

$$L(K) = L^{RICH}(K) \times L^{CALO}(\text{non-}e) \times L^{MUON}(\text{non-}\mu) \tag{2.5}$$

$$L(\pi) = L^{RICH}(\pi) \times L^{CALO}(\text{non-}e) \times L^{MUON}(\text{non-}\mu). \tag{2.6}$$

Due to differences in kinematics and underlying event multiplicity, the likelihood scale changes between events. ΔLL values with respect to a reference likelihood, chosen to be the pion hypothesis, are therefore constructed, and the global PID likelihood function is defined as:

$$\Delta LL_{x/\pi} = \ln L(x) - \ln L(\pi). \tag{2.7}$$

Substituting Eq. (2.5) and Eq. (2.6) into Eq. (2.7), it can be seen that pions and kaons are identified primarily using the RICH detectors (see Sec. 2.1.4).

The RICH likelihoods are constructed by comparing the observed pattern of hit pixels on the RICH photo detectors to the expected pattern, given all reconstructed tracks in an event under a given set of particle hypotheses [69]. All tracks in an event must be considered simultaneously because the Cherenkov rings of different tracks overlap. The likelihood function is maximised by varying the particle hypotheses for each track being a kaon, proton, pion, muon or electron. In the case of pion-kaon separation:

$$\Delta LL_{K/\pi}^{RICH} = \ln L_{max}^{RICH}(K|track) - \ln L_{max}^{RICH}(\pi|track), \tag{2.8}$$

where the maximum likelihood of the kaon hypothesis for a given track is taken relative to the maximum likelihood of the pion hypothesis, as per Eq. (2.7).

The RICH PID performance for K/π separation is demonstrated in Fig. 2.15 using data samples from Run 1 and Run 2. The kaon identification efficiency (red) and pion misidentification rate (black) are plotted as a function of track momentum. Two different $\Delta LL_{K/\pi}$ requirements have been imposed on the samples, resulting in the open and filled marker distributions for the looser and tighter threshold, respectively. As discussed in Sec. 2.1.4, at high momentum the Cherenkov angle saturates and at low momentum the tracks suffer from a low Cherenkov photon yield, degrading the RICH separation power at both ends of the spectrum. The majority of pion/kaon tracks considered in this thesis lie in the central, high performance momentum region.

As well as track momentum, the RICH performance is also known to depend on the particle rapidity and the track multiplicity of an event [70]; the exact calculation of PID efficiencies will be covered in Sec. 3.1.6.

2.3.3 Neutral particle identification and reconstruction

Of particular significance for this thesis is the reconstruction of neutral pions and photons from strong D^{*0} decays. This process begins by grouping together energy deposits in ECAL cells into clusters by applying a 3×3 cell pattern around the local energy deposition maxima [72]. If multiple clusters overlap in a single cell, the cell energy is redistributed between the clusters proportional to the total cluster

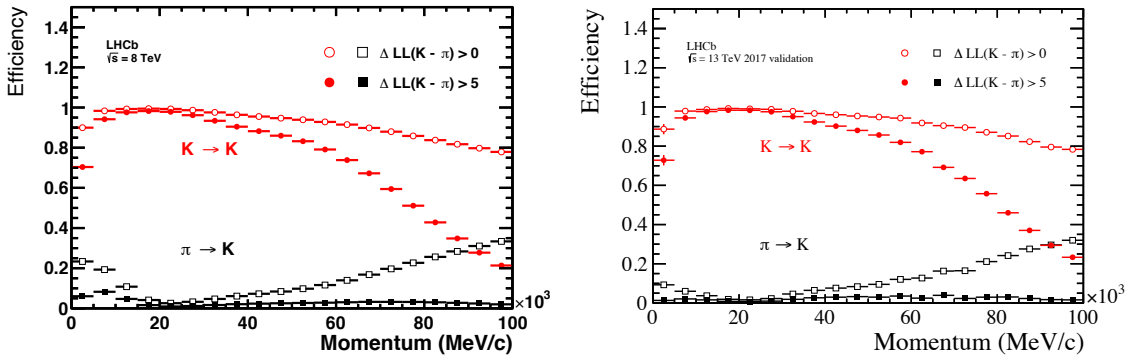


Figure 2.15: The RICH kaon identification efficiency (red) and pion misidentification rate (black) are plotted as a function of track momentum for 2012 data (left) and 2017 data (right) [71]. Two different $\Delta LL_{K/\pi}$ requirements have been imposed on the samples, resulting in the open and filled marker distributions for the looser and tighter threshold, respectively.

energy using an iterative procedure. The Molière radius of the ECAL is smaller than the size of each cell, therefore this converges quickly. The following cluster parameters are then evaluated, where for each cell i, E_i is the local energy deposit and (x_i, y_i) is the position of the cell centre:

- Energy:

$$E_{cl} = \sum_i E_i. \tag{2.9}$$

- Transverse barycentre:

$$\vec{r}_{cl} = (x_b, y_b) = \frac{1}{E_{cl}} \left(\sum_i E_i x_i, \sum_i E_i y_i \right). \tag{2.10}$$

- Transverse dispersion:

$$\mathcal{S}_{cl} = \frac{1}{E_{cl}} \begin{pmatrix} \sum_i E_i \left(x_i - x_b\right)^2 & \sum_i E_i \left(x_i - x_b\right)\left(y_i - y_b\right) \\ \sum_i E_i \left(x_i - x_b\right)\left(y_i - y_b\right) & \sum_i E_i \left(y_i - y_b\right)^2 \end{pmatrix}. \tag{2.11}$$

The efficiency of reconstructing a cluster out of an ECAL deposit depends on the transverse energy (E_T) of the cluster itself and is particularly affected by the overlap from other particles arriving in the ECAL. A neutral particle is only considered to be reconstructed if it contributes to at least 90% of the cluster's energy and the cluster contains 90% of the particle's energy. The efficiency of reconstructing ECAL clusters using simulated $B^0 \to K^{*0}\gamma$ decays has been studied in Run3 conditions [73] and is

shown in Fig. 2.16 as a function of E_T. For low E_T photons in the range $350-1000\,\text{MeV}$, as are considered in this thesis, the efficiency is between $40-60\%$ due to high levels of low energy combinatorial photons. The instantaneous luminosity at LHCb in Run 1 and Run 2 was $4\times10^{32}\,\text{cm}^{-2}\text{s}^{-1}$; in Run3, this will increase by a factor of ~ 10 to $2\times10^{33}\,\text{cm}^{-2}\text{s}^{-1}$. The cluster reconstruction efficiency will therefore be slightly higher than those shown in Fig. 2.16 for Run 1 and Run 2 due to less pile up, but this is still the dominant source of inefficiency when reconstructing neutral particles at low E_T.

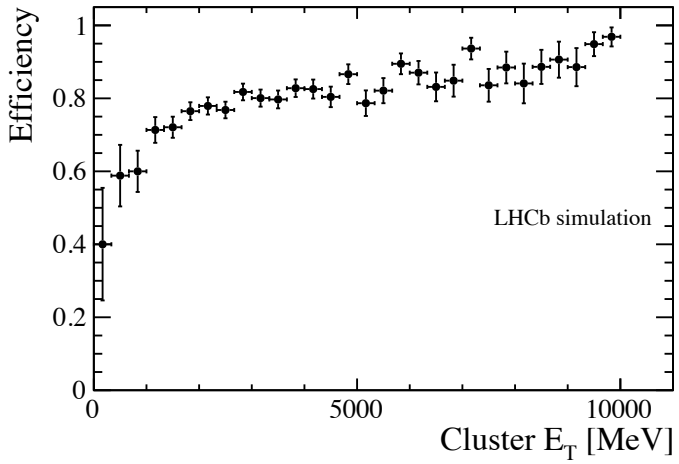


Figure 2.16: ECAL cluster reconstruction efficiency *vs.* transverse energy, E_T, in Run3 conditions, measured using $B^0 \to K^{*0}\gamma$ simulation [73].

Once an energy cluster has been reconstructed, photon candidates are identified as those without any associated extrapolated track. This is realised by performing a one-to-one matching between reconstructed tracks and clusters in the event. A 2-dimensional χ^2_{2D} is built for each pairing, representing the geometric distance between the cluster barycentre and the extrapolated track position to the ECAL reference plane:

$$\chi^2_{2D}(\vec{r}) = (\vec{r}_{tr}-\vec{r})^T\mathcal{C}^{-1}_{tr}(\vec{r}_{tr}-\vec{r}) + (\vec{r}_{cl}-\vec{r})^T\mathcal{S}^{-1}_{cl}(\vec{r}_{cl}-\vec{r}), \tag{2.12}$$

where $\vec{r}_{tr}$ and $\vec{r}_{cl}$ represent the 2D coordinates of extrapolated tracks and energy-weighted cluster centres, respectively. C_{tr} is the covariance matrix of the $\vec{r}_{tr}$ parameters and S_{cl} is the transverse energy dispersion, defined above. The χ^2_{2D} estimator is minimised with respect to $\vec{r}$ and energy clusters are identified as neutral clusters or photon candidates if none of the pairings have a minimum $\chi^2_{2D} < 4$. This cut was

chosen have to a 100% rejection rate of electrons clusters. It also significantly reduces clusters due to other charged particles whilst maintaining a high efficiency for photons ($\sim 90\%$). Apart from the χ^2_{2D} requirements, no further cuts are placed on photon clusters; higher efficiency is favoured over increased purity as the majority of photon backgrounds are from neutral combinatorial, not charged tracks [58].

Once a photon candidate has been established, its energy is derived from the total cluster energy in the ECAL and the reconstructed energy deposit in the PS. Corrections are applied to account for longitudinal leakage, where energy is lost in the lead absorber, and transverse leakage, which considers shower extension outside the cluster area and energy lost in the passive material in the transverse plane of the detector. The photon direction is taken along the vector connecting the primary vertex to the barycentre of the photon shower (x_c, y_c, z_c). The longitudinal barycentre position, z_c, is determined from the longitudinal position in the ECAL; the transverse barycentre position, (x_c, y_c), is evaluated from the energy-weighted barycentre (x_b, y_b).

There are three categories of reconstructed photons [72]:

1. Converted photons that pair-produce *before* the magnet/tracking stations. The electron-positron pair can be reconstructed from the 2 opposite sign tracks.

2. Converted photons that pair-produce *after* the magnet/tracking stations. These are recognised by single or double clusters in the ECAL with an associated SPD hit.

3. Unconverted photons, which reach the SPD/PS detector without pair-producing. These are recognised by a single cluster in the ECAL, with no associated SPD hit.

For photons within category 1, only electron/positron tracks with $p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$ are considered by the reconstruction algorithm: below this, electrons are swept out of detector acceptance by the dipole magnet. The majority of photons from strong D^{*0} decays have $p_{\mathrm{T}} < 1000\,\mathrm{MeV}/c$, therefore photons that have converted before

the magnet, accounting for $< 20\%$ of all photon candidates, are not considered in this thesis.

Neutral pions originating from D^{*0} decays are reconstructed as pairs of well-separated photons with $p_{\mathrm{T}} > 200\,\mathrm{MeV}/c$ to remove large combinatorial backgrounds from low-momentum photons. These resolved π^0 candidates are formed by looping over photon candidates, pairing them and comparing their invariant mass sum to the nominal π^0 mass ($134.98\,\mathrm{MeV}/c^2$). Candidates with reconstructed mass within the range $100\,\mathrm{MeV}/c^2 < m(\pi^0) < 200\,\mathrm{MeV}/c^2$ are kept. Merged π^0 candidates, where ECAL granularity fails to resolve the daughter photons, become relevant for the transverse momentum region $p_{\mathrm{T}} > 2\,\mathrm{GeV}/c$. These merged candidates are therefore not considered in this thesis. A summary of neutral candidates, classified by their reconstruction signature, is given in Tab. 2.1.

Neutral candidate	**Reconstruction**	**Considered**
Photon converted before magnet	Two opposite sign e^+e^- tracks	No
Photon converted after magnet	Single/double PS and ECAL cluster with SPD hit	Yes
Unconverted photon	PS and ECAL cluster with no SPD hit	Yes
Resolved π^0	Pair of well separate photon deposits in ECAL	Yes
Merged π^0	Unresolved photon deposits in ECAL	No

Table 2.1: Neutral candidates, classified by their method of reconstruction. Whether a particular candidate has been considered in this thesis is indicated in the right hand column.

A charged particle passing through the whole tracking system, with momentum within the range $5\,\mathrm{GeV}/c < p < 200\,\mathrm{GeV}/c$, has a 96% probability of being reconstructed. In contrast, photons and π^0 mesons have a reconstruction efficiency of around 20% and 3%, respectively. These neutral reconstruction efficiencies were evaluated by taking the ratio of the estimated number of $B^\pm \to (D^* \to [K^\pm\pi^\mp]_D\pi^0/\gamma)\pi^\pm$ events straight out of stripping for the fully and partially reconstructed analyses. The yields measured by the final invariant mass fits to data of the FR analysis (which will be

given in Sec. 4.2.1) are divided by offline selection efficiencies obtained from simulation (which will be given in Sec. 3.2.4) to estimate the number of events straight out of stripping. An equivalent calculation was performed to the yields measured by the partially reconstructed analysis of $B^{\pm} \to D^{(*)}h^{\pm}$ decays [31], where the neutral particle was not included in the final state. The number of D^{*0} candidates fully reconstructed with the neutral particle, compared to partially reconstructed without, gives an estimate of the neutral reconstruction efficiency at LHCb.

2.4 Simulation

Centralised LHCb Monte Carlo (MC) simulation samples are used in this thesis to determine selection efficiencies, efficiencies of PID variables and to model the invariant-mass distributions of signal and background decay modes.

To generate these samples, simulated $b\overline{b}$ pairs are produced from pp collisions using Pythia 8 [74], with an LHCb specific configuration [75]. One quark is then chosen at random to decay via a user-specified process and its decay is simulated using EvtGen [76], with Photos [77] modelling any final state radiation. The interaction of generated particles with the detector, and the detector response, is modelled using Geant4 [78]. The simulated samples are then processed through the trigger, reconstruction and stripping as for real data.

3

Selection and parameterisation of $B^{\pm} \to (D^* \to D\pi^0/\gamma)h^{\pm}$ decays

3.1 Candidate reconstruction and selection

In this chapter, the procedure for reconstructing and selecting $B^{\pm} \to (D^* \to D\pi^0/\gamma)h^{\pm}$ decays is described, where the D^* (D) symbol is inclusive of both D^{*0} and $\overline{D}^{*0}$ (D^0 and $\overline{D}^0$) intermediate states. The strong D^* decay product, either a π^0 meson or a photon, is denoted the *neutral*, and h is labelled the *companion* particle, $h \in (\pi, K)$. The D meson is reconstructed in the $K^{\pm}\pi^{\mp}, K^{\pm}K^{\mp}, \pi^{\pm}\pi^{\mp}, \pi^{\pm}K^{\mp}$ final states, corresponding to the favoured D decay mode, the 2 *CP* modes, and the suppressed mode, respectively. These modes will be referenced by the symbols $K\pi, KK, \pi\pi, \pi K$ in following text. The *CP* modes are easily identified at first glance; the identifiable difference between the favoured and suppressed modes is that for the former, the kaon D decay product has the same sign as the companion particle. In contrast, for the suppressed mode, the tracks have opposite charge.

The data presented in this thesis corresponds to an integrated luminosity of $3\,\text{fb}^{-1}$ collected at centre of mass energy $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$ during Run 1, and $5.7\,\text{fb}^{-1}$ collected at $\sqrt{s} = 13\,\text{TeV}$ during Run 2.

3.1.1 Reconstruction and trigger requirements

The data analysed in this thesis comes from two centrally produced *stripping lines* (see Sec. 2.2.3), which reconstruct and store decays of the kind $B \to [hh]_D K$ and $B \to [hh]_D \pi$. These lines discard high multiplicity events prior to the formation of any combinations by applying a Global Event Cut (GEC) on the number of tracks, which must be less than 500. Signal candidates are then built by selecting combinations of D candidates and charged tracks ($h^{\pm}$) within the invariant mass range $4750 < m(Dh) < 7000\,\text{MeV}/c^2$. The loose pre-selection requirements, described below, are designed to reject combinatorial backgrounds from charged tracks in the event not originating from the same B meson.

In order to select candidates that are displaced from the primary vertex (PV), the B proper lifetime must exceed 0.2 ps and the χ^2 of the distance between the decay vertex and the PV for D candidates must be greater than 36. For the companion particle and D decay products, the minimal χ^2 distance between each track and the PV is required to be greater than 4. To ensure the B meson originated close to the PV, the cosine of the direction angle (DIRA) between the B meson momentum vector and the line connecting the PV to its decay vertex must be > 0.999, and the minimal χ^2 distance between the B meson and the PV must be less than 25.

For both B and D candidates, the vertex quality $\chi^2_{\text{vtx}}/\text{ndf}$ (where ndf refers to the number of degrees of freedom in the vertex fit) is required to be below 10. The invariant mass of D candidates must lie within ± 100 MeV/c^2 of the PDG D^0 mass, $m(D^0) = 1864.84 \pm 0.05\,\text{MeV}/c^2$. Tracks must be well reconstructed, with track quality χ^2_{tr}/ndf less than 3 (4) for Run 1 (Run 2), and pions and kaons are subject to loose PID requirements. The distance of closest approach (DOCA) between D decay products must be less than 0.5 mm. B candidates must additionally satisfy a loose multivariate selection using a Boosted Decision Tree classifier (BDT) [79]. This takes as input the particle momentum and flight distance significance, as well as the sum of all χ^2_{vtx} values in the decay chain (*i.e.* $\chi^2_{\text{vtx}}(B) + \chi^2_{\text{vtx}}(D)$). Here, the

flight distance significance is defined as:

$$\frac{V_B - PV}{\sigma}, \tag{3.1}$$

where V_B represents the B decay vertex and $\sigma = \sqrt{\sigma^2_{V_B} + \sigma^2_{PV}}$ the flight distance uncertainty.

Once a $B^{\pm} \to Dh^{\pm}$ candidate has been loaded from the stripping, neutral particles from the underlying event are combined with the D meson to make $B^{\pm} \to D^{*}h^{\pm}$ candidates. When building the D^{*} candidates, the mass difference $m(D^{*}) - m(D)$ is required to lie in the range $[0, 500]\,\mathrm{MeV}/c^2$. In order to reduce combinatorial backgrounds from low energy photons in the event, transverse momentum requirements are placed on neutral particles. For the $D^{*} \to D\gamma$ final state, the transverse momentum of the photon is required to be greater than $350\,\mathrm{MeV}/c$. For the $D^{*} \to D\pi^{0}$ final state, the transverse momentum of the π^{0} meson is required to be greater than $350\,\mathrm{MeV}/c$, and the π^{0} secondary photons must have $p_{\mathrm{T}} > 200\,\mathrm{MeV}/c$. There are approximately 2(3) $D^{*}h$ candidates formed for every Dh candidate in the data when reconstructed in the $D\gamma$ ($D\pi^{0}$) final state. This is due to high levels of low E_{T} photons in the underlying event. Of these $D^{*}h$ combinations, only $\sim 11\%(\sim 1\%)$ are signal candidates, which we label *true* D^{*} mesons. D^{*} candidates formed using the *wrong* neutral are labelled *fake* D^{*} candidates. As a point of comparison, for the events considered straight out of stripping in the partially reconstructed $B^{\pm} \to D^{(*)}h^{\pm}$ analysis [31], $\sim 87\%$ of Dh combinations are true $B^{\pm} \to Dh^{\pm}$ signal candidates. This demonstrates the high level of combinatorial background from neutral particles in the fully reconstructed data. These numbers were calculated using MC truth information, which provides the user with the *true*, generated properties of each candidate in the event.

Selected $B^{\pm} \to D^{*}h^{\pm}$ candidates must then fulfil the hardware trigger, meaning that decay products of the signal candidate detected by the HCAL (`L0HadronDecision_TOS`, where TOS stands for triggered *on* signal), or events containing at least one candidate from elsewhere in the event (`L0GlobalDecision_TIS`, where TIS stands for triggered *independent* of signal), must lie above a fixed threshold in transverse energy. For signal candidates reconstructed in the $D\pi^{0}$ final state, 55%

are triggered on signal (TOS), 15% are triggered independent of signal (TIS) and 30% of events pass both triggers. The statistics are similar for signal candidates reconstructed in the $D\gamma$ final state, where 49% are triggered on signal (TOS), 25% are triggered independent of signal (TIS) and 26% of events pass both triggers. All candidates in the signal decay chain must additionally pass the software trigger. This places requirements on the quality of tracks (`HLT1Track`, `HLT1TwoTrack`), which are combined one-by-one and identified as having either 2- (`HLT2Topo2Body`), 3-(`HLT2Topo3Body`), or 4-body (`HLT2Topo4Body`) topology, depending on their distance of closest approach. A more thorough description of these trigger lines can be found in Sec. 2.2.

3.1.2 Multivariate analysis with Boosted Decision Trees (BDT)

The first stage of the optimised offline selection designed specifically for this analysis uses multivariate analysis (MVA) techniques. This allows correlations between variables to be exploited and in doing so achieves a higher level of purity in the data sample than would be possible by cutting on theses variables independently. There are two stages of MVA implemented to reduce combinatorial backgrounds:

- **BDT1: charged BDT** - discriminates against random combinations of charged tracks in the event.
- **BDT2: neutral BDT** - trained on neutral particle information to minimise fake D^* combinations.

BDTs are supervised machine learning algorithms, so must be provided with a sample of classified events. A set of variables from labelled signal and background samples are used as input, which are randomly split into training and testing sub-samples. The former is then used to train the BDT, classifying the events as *signal-like* or *background-like*, and the latter to check for any bias in the algorithm towards the training sample, known as *overtraining*. Both stages use Adaptive (Ada-)Boosted Decision Trees [79, 80], implemented by the Toolkit for Multivariate Analysis (TMVA) [81, 82].

A schematic of a simple Decision Tree (DT) is shown in Fig. 3.1. Each level of the tree looks for the best combination of variable and value to split the training data into two subsets. These splitting points are known as nodes, and in this binary implementation, each node i holds S_i signal events and B_i background events. The separation is performed in order to maximise the separation gain, defined by:

$$\text{Gain} = N_{ParentNode} \times G_{ParentNode} - N_{LeftNode} \times G_{LeftNode} - N_{RightNode} \times G_{RightNode}, \tag{3.2}$$

where $N_i = S_i + B_i$. This particular parameterisation uses the Gini Index $G_i = p_i \times (1 - p_i)$, where the purity of a node $p_i = S_i / (S_i + B_i)$. A straight-forward DT repeats this process, optimising the variable and value choice for each node, until the separation gain is no longer increased by further splitting, or a minimum number of events is reached. The number of decisions determines the depth of the tree, and the final node is known as a leaf. The leaves are labelled as either signal-like or background-like, according to the class that the majority of events belong to. Background events on a signal leaf or signal events on a background leaf are deemed mis-classified.

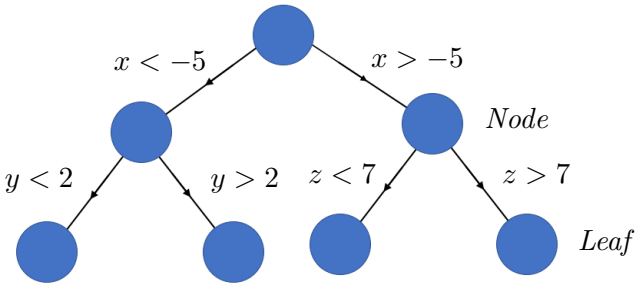


Figure 3.1: Schematic of a decision tree, of depth 2, where x, y and z represent variable names.

A boosting algorithm uses a collection of classifiers of the same type trained sequentially on a training sample that is re-weighted each time. The final response is a weighted average of the individual classifiers, which increases the stability of the classifier to statistical fluctuations in the training data. An Ada-Boosted Decision Tree algorithm uses a set of DTs with shallow depth (≤ 3). All events in the training sample are initially given equal weights. Following this, before training the $(j+1)^{th}$

tree, mis-classified events in the j^{th} tree are weighted by $\alpha_j = (1 - f_j)/f_j \geq 1$, where $f_j < 0.5$ is the fraction of mis-classified events. The data sample is then re-normalised so that the sum of weights remains constant. In this way, mis-classified events are given a higher weight in the training of the next tree. Each event $\mathbf{x}$, represented by a tuple of training variables, put into a single tree j, will either give a response $h_j(\mathbf{x}) = +1$ for signal or $h_j(\mathbf{x}) = -1$ for background. The boosted event classification is then given by the weighted average over all trees in the set, as defined in Eq. (3.3), favouring those that correctly classify events:

$$Response = \frac{1}{N_{trees}} \sum_{j=1}^{N_{trees}} \ln(\alpha_j)\, h_j(\mathbf{x}). \tag{3.3}$$

This response can range from -1 (background-like) to +1 (signal-like).

Charged BDT

The strategy used for the first stage BDT builds on work published in previous ADS/GLW analyses of $B^{\pm} \to [h^{\pm}h^{\mp}]_D h^{\pm}$ decays [31, 83]. The purpose of this BDT is to target fake D candidates, formed from the combination of two random tracks, and fake B candidates, comprised of a true D mesons plus a random track.

A single BDT classifier is trained to encompass data collected across both Run 1 and Run 2, and both D^{*} decay modes. Data side-bands for the favoured D decay modes $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_D\pi^{0})h^{\pm}$ and $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_D\gamma)h^{\pm}$, $h = \pi/K$, are provided as the background sample, where side-bands refers to data that falls in the range $5800\,\mathrm{MeV}/c^2 < m(Dh) < 6800\,\mathrm{MeV}/c^2$, as depicted in the 2D plots of Fig. 3.2 for the pion companion. The Δm variable, on the vertical axes, is defined as the raw mass difference:

$$\Delta m = \begin{cases} m(D^{*}) - m(D), & \text{for } D^{*} \to D\gamma \\ m(D^{*}) - m(D) - m(\pi^{0}) + m(\pi^{0})_{\mathrm{PDG}}, & \text{for } D^{*} \to D\pi^{0}, \end{cases} \tag{3.4}$$

where the PDG subscript refers to the meson mass listed on the PDG [19], $m(\pi^{0})_{\mathrm{PDG}} = 134.98\,\mathrm{MeV}/c^2$. The projection of this 2D data onto the $m(D\pi)$ dimension is shown in Fig. 3.3.

The partially reconstructed mass variable $m(Dh)$ is used to determine the background sample due to the presence of over-reconstructed $B^{\pm} \to Dh^{\pm}$ decays in the data, where a neutral particle from the underlying event has been added to the D meson. The inclusion of the neutral shifts the reconstructed B upwards in mass and into the $m(D^*h)$ upper sideband. It would reduce the discriminating power of the classifier if these candidates contaminated the background sample, as they contain true D and B mesons. The presence of this background can be clearly seen in Fig. 3.3; the largest peak depicted on the right hand side of the $m(D\pi)$ distributions is due to $B^{\pm} \to D\pi^{\pm}$ decays. The D^* modes are partially reconstructed to the left hand side of this peak; the $D\pi^0$ decay channel candidates can be found in the double horned structure, and the $D\gamma$ candidates are distributed broadly underneath this (as for the PR analysis described in Sec. 1.4.5).

In this thesis, all measured CP observables are constructed from ratios of common D decay products, or a double ratio thereof, to minimise systematic uncertainty from the relative reconstruction efficiency of different D decay modes. It is therefore considered sufficient to analyse one set of simulation samples for this work; the favoured mode $D \to K\pi$ mode is chosen. Simulated $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0/\gamma)\pi^{\pm}/K^{\pm}$ signal events are therefore provided as a combined signal sample to the BDT, where MC candidates are required to satisfy the same stripping and trigger requirements as data. The described signal and background samples are split into training and test subsets using an 80 : 20 split, respectively.

The set of training variables, listed by rank in Tab. 3.1, are chosen to exploit the decay topology of tracks, and represent a sub-sample of those used in previous $B^- \to [h^-h^+]_D h^-$ ADS/GLW analyses [31, 83]. A description of each variable is given in the table, where some are transformed using a logarithm function to increase their separation power. The quantity labelled `Bu_ptasy_1.50`, defined by Eq. (3.5), quantifies the isolation of the B candidate by calculating the p_{T} asymmetry between the B meson and other tracks from the same PV:

$$A_{p_{\mathrm{T}}} = \frac{p_T^B - p_T^{cone}}{p_T^B + p_T^{cone}}. \tag{3.5}$$

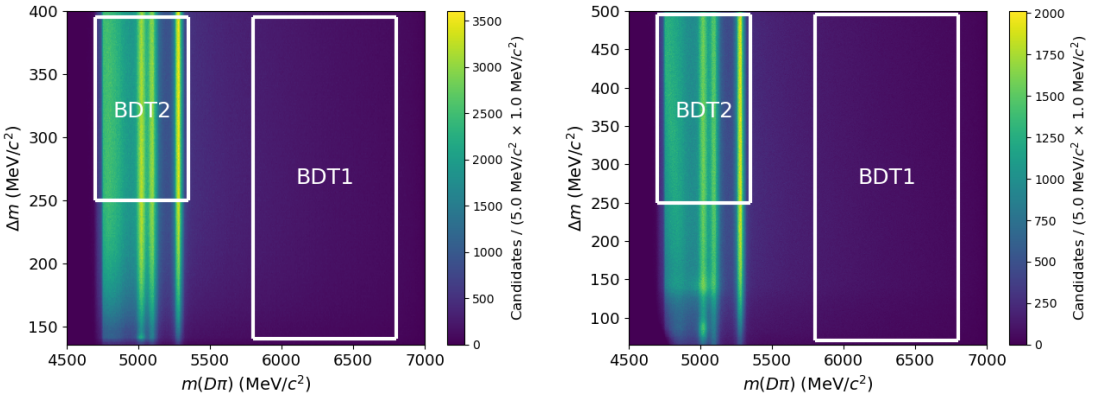


Figure 3.2: 2D histogram of Δm plotted against $m(D\pi)$ for 2011-2018 data. The background samples for BDT training are taken from the regions indicated in the labelled boxes for $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ (left) and $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ (right) data.

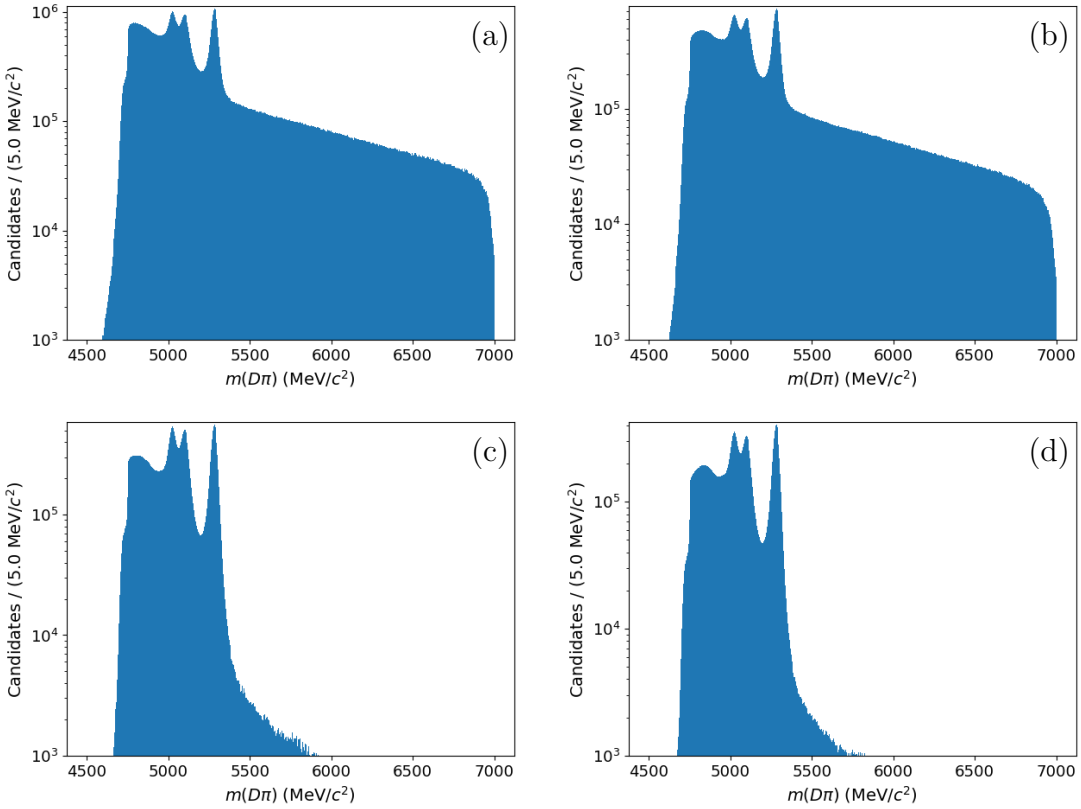


Figure 3.3: 1D histograms of $m(D\pi)$ for 2011-2018 data with a logarithmic y-axis scale. $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ data before and after the application of the charged BDT1 is shown in (a) and (c), respectively. $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ data before and after the application of the charged BDT1 is shown in (b) and (d), respectively. The large peaks to the right hand side, above $m(D\pi) \gtrsim 200\,\text{MeV}/c^2$, are due to $B^{\pm} \to D\pi^{\pm}$ decays; $B^{\pm} \to (D^{*} \to D\pi^{0})\pi^{\pm}$ decays are part of the double horned structures in the range $5000\,\text{MeV}/c^2 \gtrsim m(D\pi) \gtrsim 5200\,\text{MeV}/c^2$, whilst $B^{\pm} \to (D^{*} \to D\gamma)\pi^{\pm}$ candidates form broad resonances in the same region.

Here, p_T^B is the p_{T} of the B candidate and p_T^{cone} is the p_{T} of all other tracks in a cone around the B candidate of radius 1.50 radians. The relative discriminating power of each variable is evaluated on the training set by counting how often it is used to split decision nodes, weighting each split by the square of the separation gain, defined by Eq. (3.2), and the number of events considered on the node. This can be interpreted as a measure of the ease of telling the signal and background distributions apart. A visual illustration of the separation power of each variable is given in Fig. 3.4.

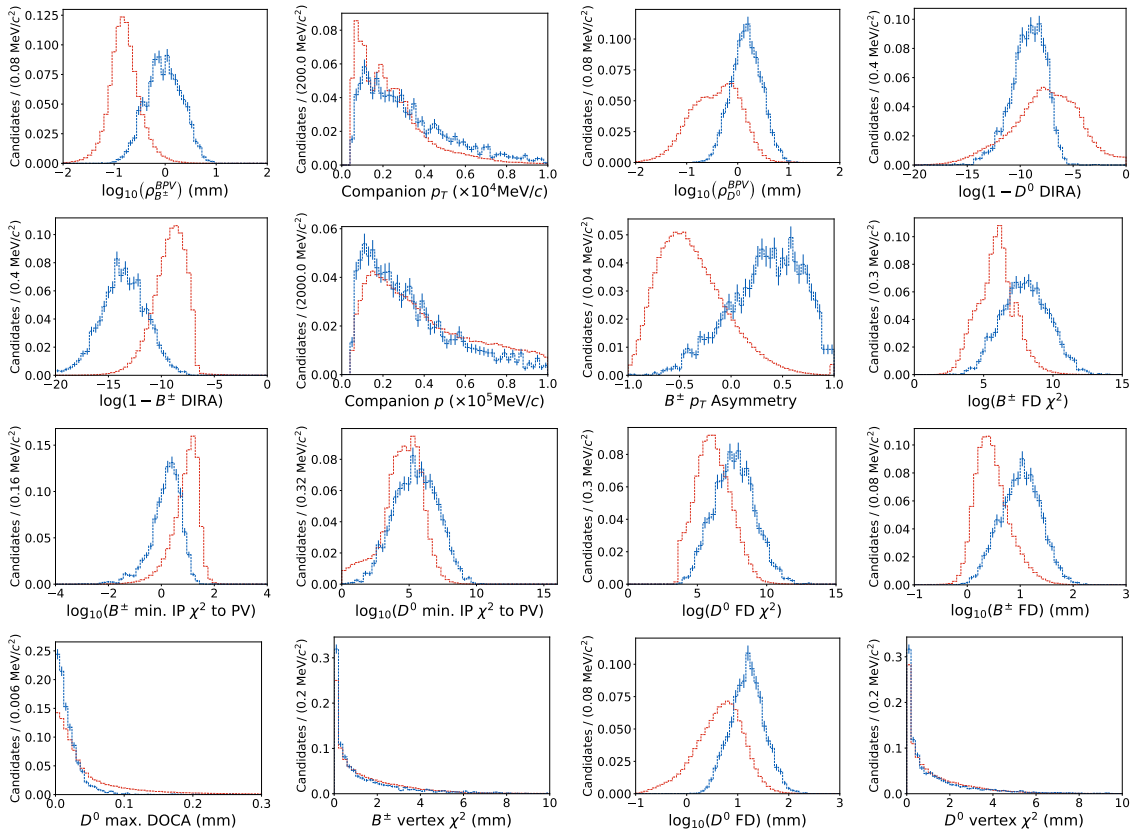


Figure 3.4: Normalised training variable distributions for the signal (blue) and background (red) samples that feed into the charged BDT.

The following BDT hyperparameters are optimised using a grid search: the number of trees, the maximum depth of a single tree, and the minimum fraction of events on a single leaf. Due to limited simulation sample sizes, this is done using 5-fold cross-validation. During this process, the training data is split into five subsets; 4 of these subsets are used to train the model, and the final 'hold-out' subset is used to evaluate model performance. In this case, performance is evaluated using the area under the

Receiver Operating Characteristic (ROC) curve; the larger the area under the ROC curve, the more powerful the classifier. This process is repeated 5 times, and for each iteration (fold) a different subset acts as the hold-out sample, whilst the remaining four are combined to make the training sample. The mean performance over all iterations is reported and used to select optimum values for the hyperparameters: 850 trees, a maximum tree depth of 3, and a minimum of 2.5% of events per leaf. The BDT is then retrained on the entire training dataset with these hyperparameter values specified.

The performance of the charged BDT is evaluated using the test dataset and is summarised in Fig. 3.5. The distribution of the output classifier, defined by Eq. (3.3), is shown for signal candidates (blue) and background candidates (red). The response for training and testing data are overlaid, displaying a high degree of overlap, therefore no evidence of overtraining is found. The ROC curve, in this case signal efficiency *vs.* background rejection, for the BDT applied to the test data is also plotted. In the ideal case, the BDT would keep 100% of signal and reject 100% of background, corresponding to the apex in the top right of plot (b), Fig. 3.5; the curve shown closely approaches this apex, indicating a high level of performance.

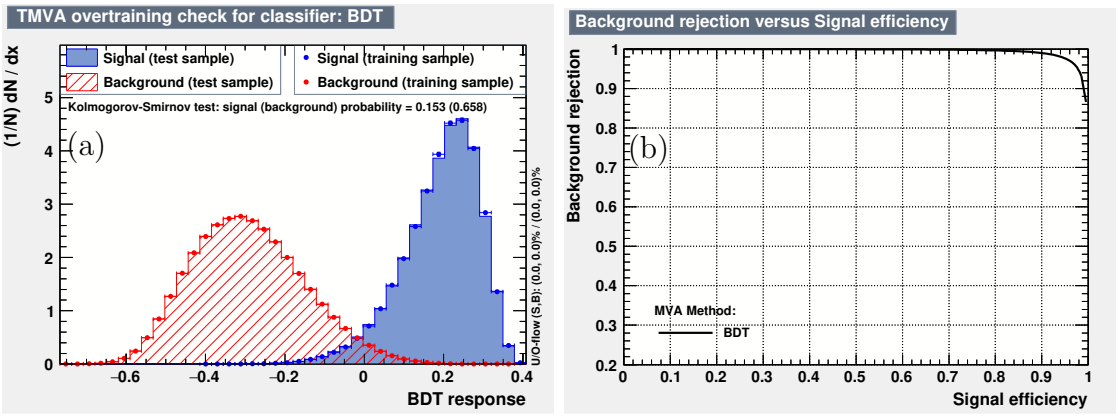


Figure 3.5: Image (a) shows the output classifier response of BDT1 for signal (red) and background (blue). The response of the testing data is overlaid on top of the training data response. The corresponding signal efficiency *v.s.* background rejection curve is shown in (b).

Once the BDT has been trained, the classifier is used to assign *response values* to data that passes the trigger and stripping requirements. Comparisons of the

$m(Dh)$ distributions of this data before and after application of the charged BDT is given in Fig. 3.3, where the optimisation procedure for selecting the applied cut is described in Sec. 3.1.4.

Neutral BDT

The second stage, neutral BDT is designed to select for true D^* candidates. Separate classifiers are trained for the $D\pi^0$ and $D\gamma$ final states. The classifiers are supplied with a set of variable distributions for signal and background samples. These variables, listed in Tab. 3.2, are a mixture of neutral particle momenta and photon identification parameters. Comparisons of the normalised signal and background sample distributions are given in Fig. 3.6.

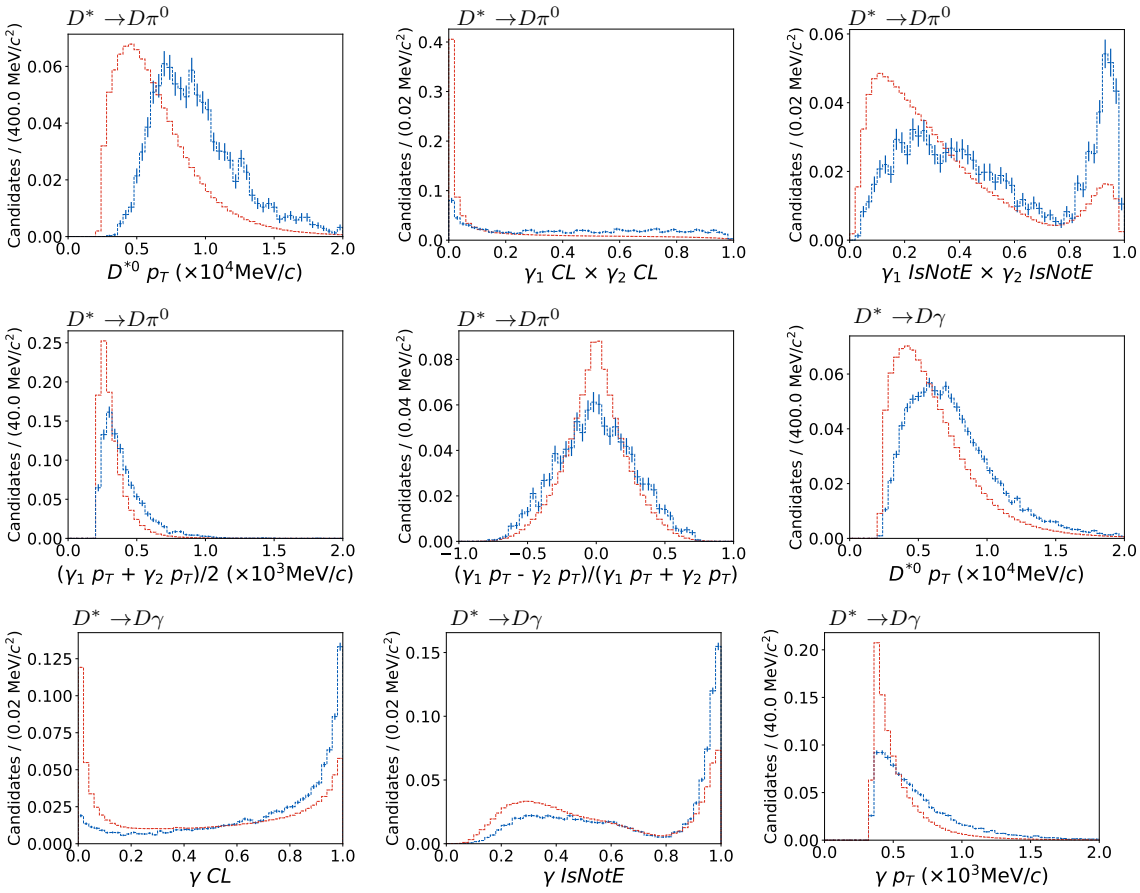


Figure 3.6: Normalised training variable distributions for the signal and background samples that feed into the $D^* \to D\pi^0$ and $D^* \to D\gamma$ neutral BDTs, where the D^* final state is indicated on the top left of each plot.

Variable	Description	BDT Importance ($\times 10^{-2}$)
$\log_{10}$(Bu_RHO_BPV)	$\Delta(\rho)$ (cylindrical coordinates) between the end vertex of the B and the best primary vertex	8.92
bach_PT	Companion transverse momentum	8.03
$\log_{10}$(D0_RHO_BPV)	$\Delta(\rho)$ (clindrical coordinates) between the end vertex of the D and the best primary vertex	7.56
log(1-D0_DIRA_BPV)	DIRA is the cosine of the angle between the D momentum and the vector from the primary to the secondary vertex	7.35
log(1-Bu_DIRA_BPV)	DIRA is the cosine of the angle between the B momentum and the vector from the primary to the secondary vertex	7.34
bach_P	Companion momentum	6.80
Bu_ptasy_1.50	p_{T} asymmetry of the B candidate using a cone radius of 1.50 radians	6.75
log(Bu_FDCHI2_OWNPV)	Flight distance χ^2 of the B with respect to the primary vertex	6.33
$\log_{10}$(Bu_MIPCHI2_PV)	Minimal χ^2 for the impact parameter of the B with respect to the primary vertex	6.00
log(D0_MIPCHI2_PV)	Minimal χ^2 for the impact parameter of the D with respect to the primary vertex	5.71
log(D0_FDCHI2_OWNPV)	Flight distance χ^2 of the D with respect to the primary vertex	5.48
$\log_{10}$(Bu_FD_BPV)	Flight distance of the B with respect to the primary vertex	5.43
D0_AMAXDOCA	Distance of closest approach of the D decat products	5.20
Bu_VTXCHI2DOF	Vertex χ^2 per degree of freedom of the B	5.07
$\log_{10}$(D0_FD_OWNPV)	Flight distance of the D with respect to the primary vertex	4.33
D0_VTXCHI2DOF	Vertex χ^2 per degree of freedom of the D	3.81

Table 3.1: Variables used in the training of the charged BDT, BDT1, ranked by importance. The label Bu refers to a B candidate, bach to the companion particle, and D0 to the D candidate.

Variable		Description	BDT Importance($\times 10^{-2}$)	
$D^* \to D\pi^0$	$D^* \to D\gamma$		$D^* \to D\pi^0$	$D^* \to D\gamma$
D^* p_{T}	D^* p_{T}	Transverse momentum of D^*	2.21	2.18
$\gamma^1_{\mathtt{CL}} \times \gamma^2_{\mathtt{CL}}$	$\gamma_{\mathtt{CL}}$	Distinguishes photons from non-electromagnetic deposits	2.17	2.79
$\gamma^1_{\mathtt{IsNotE}} \times \gamma^2_{\mathtt{IsNotE}}$	$\gamma_{\mathtt{IsNotE}}$	Separates photons and electrons	2.15	2.62
-	$\gamma_{p_{\mathrm{T}}}$	Transverse energy of photon	-	2.41
$(\gamma^1_{p_{\mathrm{T}}} + \gamma^2_{p_{\mathrm{T}}})/2$	-	Average transverse energy of π^0 decay photons	1.76	-
$(\gamma^1_{p_{\mathrm{T}}} - \gamma^2_{p_{\mathrm{T}}})/(\gamma^1_{p_{\mathrm{T}}} + \gamma^2_{p_{\mathrm{T}}})$	-	E_{T} asymmetry of π^0 decay photons	1.71	-

Table 3.2: Variables used in the training of the neutral BDT, BDT2. γ^1 and γ^2 refer to the two decay photons of the π^0 meson.

The neutral PID variables, `CL` and `IsNotE`, are multi-layered perceptrons implemented with the TMVA tool [81, 82]. These supervised algorithms are trained on properties of photon, electron and hadron clusters in the PS and ECAL, for example the χ^2_{2D} cluster isolation variable defined in Eq. (2.12), the transverse energy dispersion given by Eq. (2.11), and ratios with corresponding energy deposits in the HCAL. Using these classifiers as inputs to the neutral BDTs is possible for this analysis as precise knowledge of BDT efficiencies are not needed; all measured ratios have the same D^* decay in the numerator and denominator, so any efficiencies associated with the neutral selection cancel.

The signal samples provided to the $D\pi^0$ and $D\gamma$ neutral BDTs are made up of $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ and $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ MC samples, respectively, which have passed the stripping and trigger requirements. Data side-bands for the favoured D decay modes $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)h^{\pm}$ and $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)h^{\pm}$, $h = \pi/K$, are provided as background samples. For the neutral BDT, side-bands refers to data falling in the Δm upper region $250\,\text{MeV}/c^2 < \Delta m < 500(400)\,\text{MeV}/c^2$ for the $D\gamma$ ($D\pi^0$) BDT, but $m(Dh)$ signal region. This 2D data selection is shown in Fig. 3.2, and the 1D projection onto the Δm sample is shown in Fig. 3.7. The described signal and background samples are split into training and test subsets using an 80 : 20 split, respectively.

A loose cut on the charged BDT, BDT1 > -0.2, is also applied to all samples before training the second stage BDTs. This requirement removes obvious backgrounds from track combinatorial whilst retaining 100% of signal, therefore preserving simulation statistics. Hyperparameter optimisation was also performed for the neutral BDTs, and the same values were found as for the charged BDT: 850 trees, a maximum tree depth of 3, and a minimum of 2.5% of events per leaf.

The performance of the neutral BDTs are evaluated on the test datasets and are summarised in Fig. 3.8. The output classifier distributions are shown for signal candidates (blue) and background candidates (red). The classifier trained to distinguish $D\pi^0$ candidates is depicted on the left, and $D\gamma$ on the right. By comparing the ROC

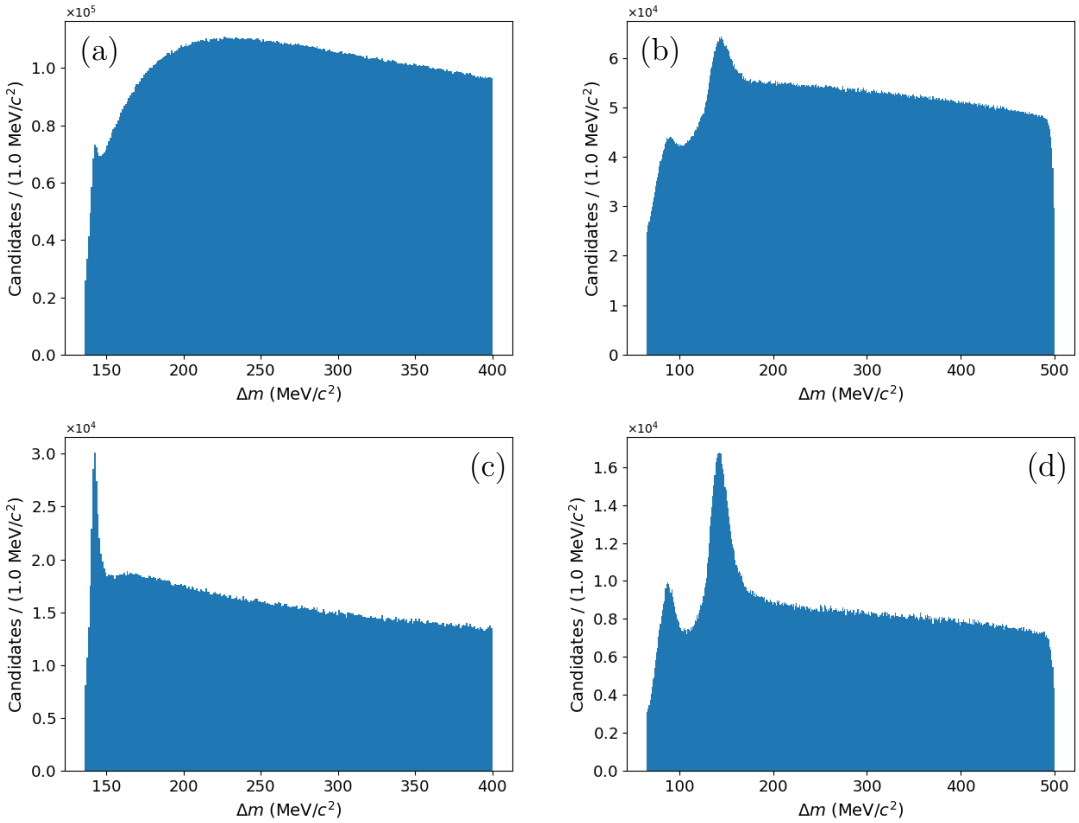


Figure 3.7: 1D histograms of Δm for 2011-2018 data. $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ data before and after the application of the neutral BDT is shown in (a) and (c), respectively. $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ data before and after the application of the neutral BDT2 is shown in (b) and (d), respectively. The first stage, charged BDT has been cut on in all distributions.

curves in (c) and (d), it can be seen that the $D\pi^0$ classifier achieves a higher level of performance. The quality of any BDT is largely dependant on its ability to distinguish signal and background with the input variables provided. For the $D\pi^0$ final state, information from the π^0 decay photons is also used in the training, increasing the classifier's performance compared to the $D\gamma$ final state. The responses for training and testing data are overlaid, displaying a high degree of overlap, therefore no evidence of overtraining is found.

The trained classifiers are used to assign *response values*, calculated using Eq. (3.3), to the data samples that pass the trigger, stripping and BDT1 > -0.2 requirements. Comparisons of the Δm distributions before and after application of the neutral BDTs are given in Fig. 3.7, where the optimisation procedure to select the applied BDT

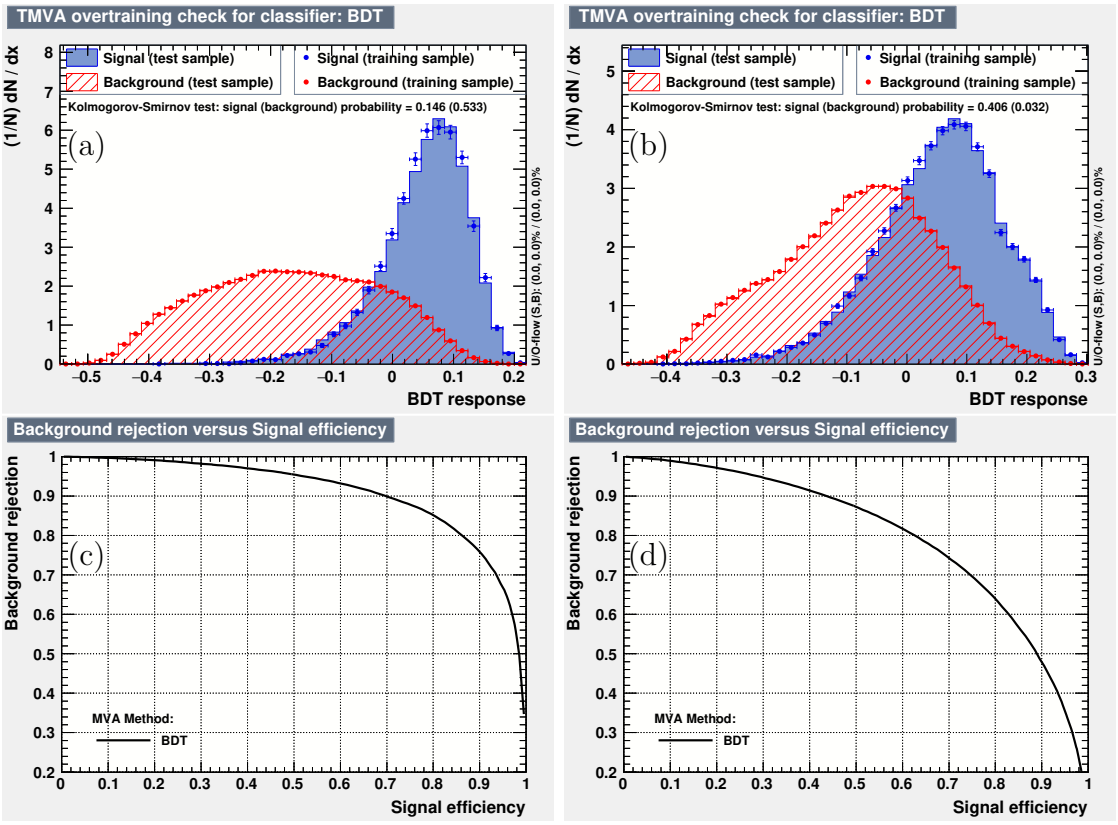


Figure 3.8: For $D^{*} \to D\pi^{0}$ and $D^{*} \to D\gamma$ candidates, respectively, images (a), (b) show the neutral BDT classifier response for signal (red) and background (blue), with testing and training data superimposed. The corresponding signal efficiency *v.s.* background rejection curves are shown in (c) and (d).

cuts is described in Sec. 3.1.4. Two peaks are visible in the $D\gamma$ mode distributions, taking plot (d) in Fig. 3.7 as an example. The peak on the right hand side is from fully reconstructed $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ decays; the peak on the left hand side is due to partially reconstructed $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ decays, where only one of the photons from the neutral pion decay has been reconstructed. In order to exploit the extra information provided by this resonance, observables from $D^{*} \to D\pi^{0}$ decays reconstructed in the $D\gamma$ data tuples are also measured.

3.1.3 Rectangular cuts

Pre-optimisation boundary cuts are placed on the BDT response values, BDT1 > -0.1 and BDT2 > -0.1, which have a combined signal efficiency of 95% (92%) for data reconstructed in the $D\pi^{0}$ ($D\gamma$) final state, whilst rejecting obvious background. The

L-shaped cuts illustrated by the 2D BDT1 *vs.* BDT2 distributions in Fig. 3.9 depict this. A series of rectangular cuts, described in this sub-section, to improve the signal purity of the data samples are applied before the BDTs are optimised.

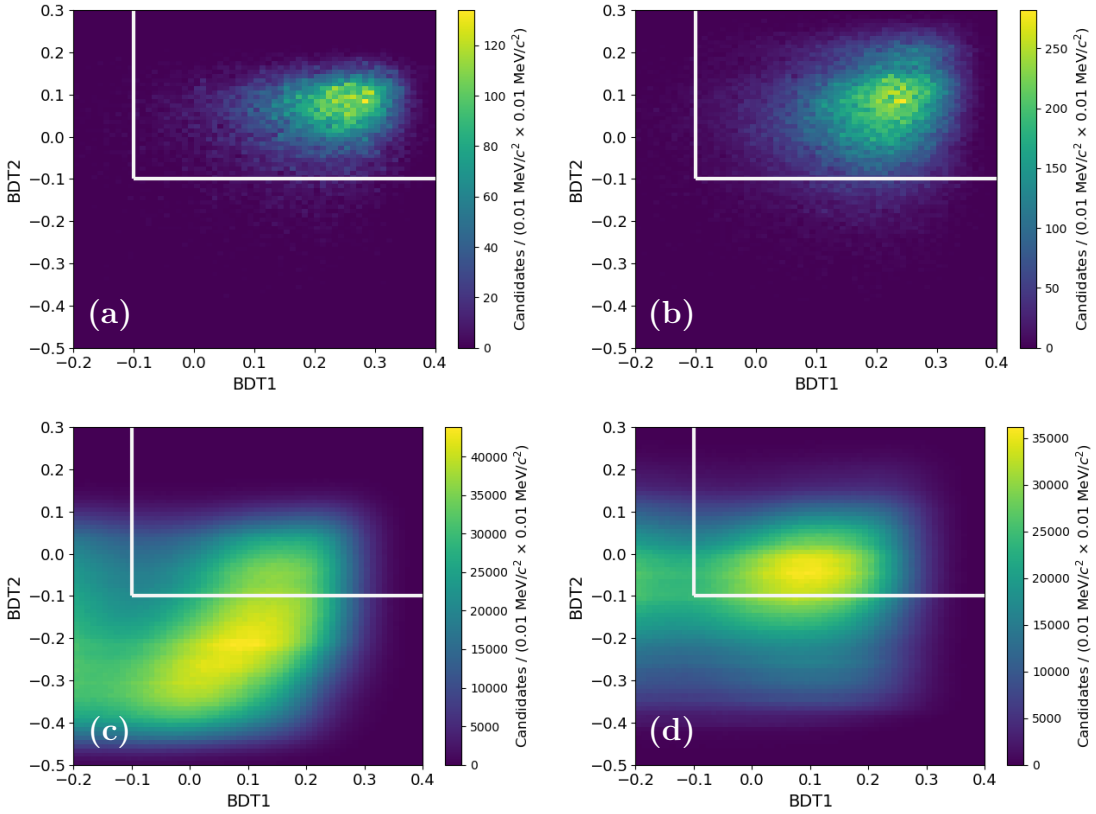


Figure 3.9: 2D plots of the response values of BDT1 *vs.* BDT2. Images (a) and (b) are of $D^{*} \to D\pi^{0}$ and $D^{*} \to D\gamma$ simulation signal candidates, respectively. Images (c) and (d) are of $D^{*} \to D\pi^{0}$ and $D^{*} \to D\gamma$ data samples, respectively. The L-shaped cuts applied to the data before optimisation are depicted in white.

The following mass cuts have no effect on the results of the analysis, as the same requirements are placed on all modes, and all observables are ratios of common D and D^{*} final states. The D meson is required to be within $\pm 25\,\mathrm{MeV}/c^{2}$ of the known D^{0} mass, which corresponds to approximately three times the mass resolution and is 97% efficient on signal; the D invariant mass distributions for signal $B^{\pm} \to D^{*}\pi^{\pm}$ simulation samples are shown in Fig. 3.10. When a neutral pion is present, it is required to lie within the invariant mass range $125\,\mathrm{MeV}/c^{2} < m(\pi^{0}) < 165\,\mathrm{MeV}/c^{2}$, which is within 1.5σ of the mean π^{0} mass and has a signal efficiency of 73%; a

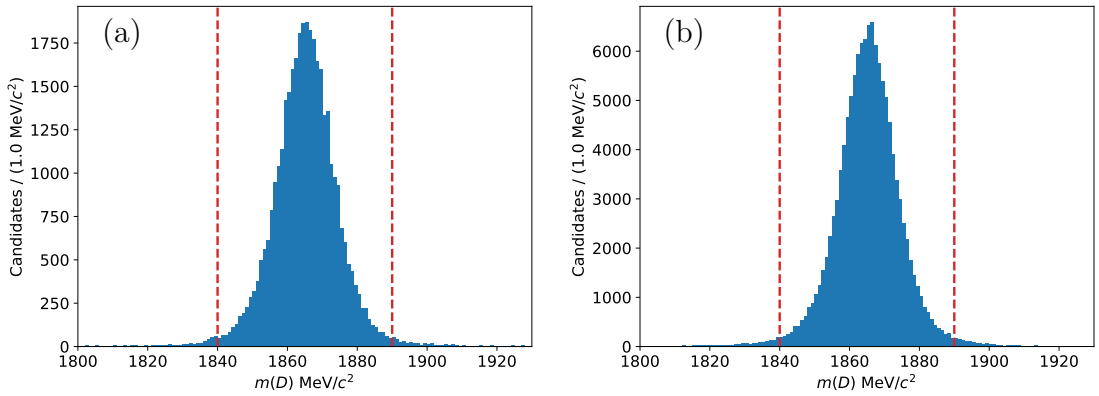


Figure 3.10: $m(D)$ distributions for $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ simulation on the left (a) and $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ simulation on the right (b). The selection requirements $1840\,\text{MeV}/c^2 < m(D) < 1890\,\text{MeV}/c^2$ are indicated by the dashed red lines.

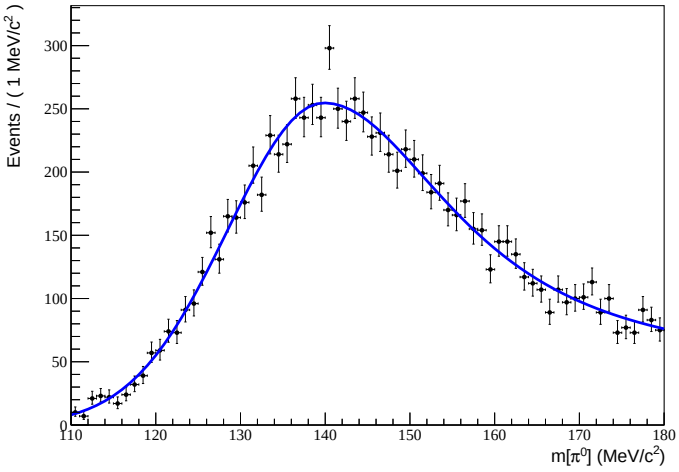
fit to the $m(\pi^0)$ distribution for simulated $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ decays is shown in Fig. 3.11. The partially reconstructed $m(Dh)$ mass must lie in the range $4900\,\text{MeV}/c^2 < m(Dh) < 5200\,\text{MeV}/c^2$, retaining all signal candidates whilst reducing the rate of partially-reconstructed backgrounds and removing over-reconstructed backgrounds. Partially-reconstructed backgrounds refer to decays of the kind $B \to (D^* \to D\pi/\gamma)h^{\pm}\pi$, where two particles in the decay chain have been missed and a neutral particle added. These candidates sit low in $m(D^*h)$ and $m(Dh)$ mass, and the $m(Dh) > 4900\,\text{MeV}/c^2$ requirement is 73% (80%) efficient at removing backgrounds with a photon (pion) associated with the D^* meson. Over-reconstructed decays refer to those of the kind $B^{\pm} \to Dh^{\pm}$, for which the full decay chain has been reconstructed and a neutral particle from the underlying event combined with the D meson candidate to create a fake D^* candidate. The $m(Dh) < 5200\,\text{MeV}/c^2$ requirement is 100% efficient at removing these events, as can be seen by referring back to the $m(Dh)$ plots in Fig. 3.3.

The flight distance significance of the D candidate in the z direction is required to be greater than 2. This variable is defined as:

$$\mathcal{S}_z = \frac{z^{\text{Decay}} - z^{\text{Origin}}}{\sqrt{\left(\sigma_{z^{\text{Decay}}}\right)^2 + \left(\sigma_{z^{\text{Origin}}}\right)^2}}, \tag{3.6}$$

where z^{Origin} and $\sigma_{z^{\text{Origin}}}$ (z^{Decay} and $\sigma_{z^{\text{Decay}}}$) mark the z positions and uncertainties of the D origin (decay) vertex. This selects D candidates that have flown a significant

distance from the B decay vertex, and is traditionally put in place to remove *charmless* backgrounds, comprising decays of the form $B \to hhhX$. Due to the small section of D^{*} phase space that this analysis is conducted in, charmless backgrounds do not pose a problem, but the D flight distance cut is necessary to reduce the rate of mis-reconstruction between signal decay categories. This cut is 89% (87%) efficient on $D\pi^{0}$ ($D\gamma$) signal MC.



$\mu = (139.85 \pm 0.57)\,\mathrm{MeV}/c^2$
$\sigma_L = (11.34 \pm 0.37)\,\mathrm{MeV}/c^2$
$\sigma_R = (15.05 \pm 1.12)\,\mathrm{MeV}/c^2$

Figure 3.11: Fit to the $m(\pi^0)$ distribution of signal $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ simulation using an asymmetric Gaussian function. The values found for the mean (μ), left-hand width (σ_L) and right-hand width (σ_R) are given to the right hand side.

Companion particles (h_B) and D decay products (h_D) with the same sign have the potential to be swapped. Reconstructing $B^{\pm} \to (D^{*} \to [h^{\mp}h_{D}^{\pm}]_{D}\pi^{0}/\gamma)h_{B}^{\pm}$ decays as $B^{\pm} \to (D^{*} \to [h^{\mp}h_{B}^{\pm}]_{D}\pi^{0}/\gamma)h_{D}^{\pm}$ decays results in mis-reconstructed D mesons. In this work, the resulting backgrounds are from favoured $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0}/\gamma)\pi^{\pm}$ decays mis-reconstructed as *CP* mode $B^{\pm} \to (D^{*} \to [\pi^{\pm}\pi^{\mp}]_{D}\pi^{0}/\gamma)K^{\pm}$ decays, and $B^{\pm} \to (D^{*} \to [K^{\pm}K^{\mp}]_{D}\pi^{0}/\gamma)\pi^{\pm}$ decays mis-reconstructed as suppressed mode $B^{\pm} \to (D^{*} \to [\pi^{\pm}K^{\mp}]_{D}\pi^{0}/\gamma)K^{\pm}$ decays. The danger of this cross-mode contamination arises from the varying branching fractions ($\mathcal{B}$) of the different D decay modes, which

are given in Eqs. (3.7) to (3.10):

$$\mathcal{B}(D^0 \to K^-\pi^+) = (3.946 \pm 0.030)\% \tag{3.7}$$

$$\mathcal{B}(D^0 \to K^-K^+) = (0.408 \pm 0.006)\% \tag{3.8}$$

$$\mathcal{B}(D^0 \to \pi^-\pi^+) = (0.1453 \pm 0.0024)\% \tag{3.9}$$

$$\mathcal{B}(D^0 \to \pi^-K^+) = (0.00150 \pm 0.0007)\%. \tag{3.10}$$

Without removing these cross-mode reconstructions, decay channels with larger branching fractions, and therefore higher statistics samples, will cause significant contamination of the lower statistics modes. Consequently, in addition to the D flight distance significance cut, the invariant mass combination $m\left(h^{\pm}h_B^{\mp}\right)$ is formed for all samples, and those candidates that lie within 25 MeV/c^2 of the nominal D mass are removed; this veto is found to be $> 99\%$ efficient on signal simulation. Although several decay modes do not suffer from this background, application of the veto across all samples ensures that CP observables are unaffected.

For the final states $D \to K\pi, KK, \pi\pi$, the amount of mis-reconstruction is negligible between the D decay products themselves because the invariant mass of the mis-reconstructed D candidate falls outside the selection window. The D meson mass of candidates with pions misidentified as kaons will be shifted upwards; the D meson mass of candidates with kaons misidentified as pions will be shifted downwards. There are also particle identification requirements placed on final state hadrons, which will be detailed in Sec. 3.1.6. When both D secondaries are mis-reconstructed, however, there is some contamination of Cabibbo-favoured $D^0 \to K^-\pi^+$ decays in the doubly-Cabibbo suppressed $D^0 \to K^+\pi^-$ sample (herein referred to as *crossfeed*), as the distribution of the double-misidentification crossfeed remains centred at the D mass, albeit much wider. To suppress this crossfeed, a veto is applied to D final states containing a pion and kaon. The D candidate is reconstructed with the mass hypothesis of the daughters swapped, *i.e.* the kaon is reconstructed as a pion and the pion is reconstructed as a kaon. These mass-swapped D candidates that lie within 15 MeV/c^2 of the nominal D mass are removed. The crossfeed veto is 93% efficient

on signal and the residual crossfeed rate is discussed in Sec. 3.3.7.

At this stage, mass cuts are also placed on the fitting variables: $4950\,\mathrm{MeV}/c^2 < m(D^*h) < 5650\,\mathrm{MeV}/c^2$ and $60\,\mathrm{MeV}/c^2 < \Delta m < 190\,\mathrm{MeV}/c^2$.

3.1.4 BDT optimisation

In order to optimise the first stage, charged BDT, invariant mass fits in Δm are performed to 2011-2018 data consisting of high statistics $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0/\gamma)\pi^{\pm}$ decays. For data reconstructed in the $D\gamma$ final state, the fitting region chosen to perform the optimisation selects for $D^* \to D\gamma$ signal candidates only. Events considered are required to pass the selection requirements described earlier in this section. For both D^* decay modes, the signal resonance is modelled using the sum of two Crystal Ball functions, and the combinatorial background is described using a probability density function (PDF) that specifically models backgrounds for $m(D^*) - m(D)$ invariant-mass difference distributions. Both of these functions are described in more detail in Sec.3.3. Unbinned maximum likelihood fits are performed across a wide range of possible BDT cuts, $-0.1 <$ BDT1 $< 0.35(0.3)$, for the $D\pi^0(D\gamma)$ mode. The signal (S) and background (B) yields $\pm 3\sigma$ around the measured signal mean are extracted at each cut value, and the significance figure of merit, defined by Eq. (3.11), is calculated:

$$FOM = S/\sqrt{S+B}. \tag{3.11}$$

The results are depicted in Fig. 3.12, and a cut of BDT1 > 0.05 is chosen for both D^* decay modes. This choice provides the highest purity samples before the signal significance begins to fall. The efficiency of this cut is 94% (95%) on $(D\pi^0)\pi$ ($(D\pi^0)K$) signal MC and 91% (92%) on $(D\gamma)\pi$ ($(D\gamma)K$) signal MC. The invariant mass fits at the optimised points are provided in Fig. 3.13.

Before optimising the second stage, neutral BDT, candidates are required to pass the first stage, BDT1 > 0.05. Simultaneous fits to Δm distributions are constructed across data categories spanning both companion particles, π/K, and the favoured and suppressed D decay modes, $D^0 \to K^-\pi^+$ and $D^0 \to K^+\pi^-$. For data reconstructed

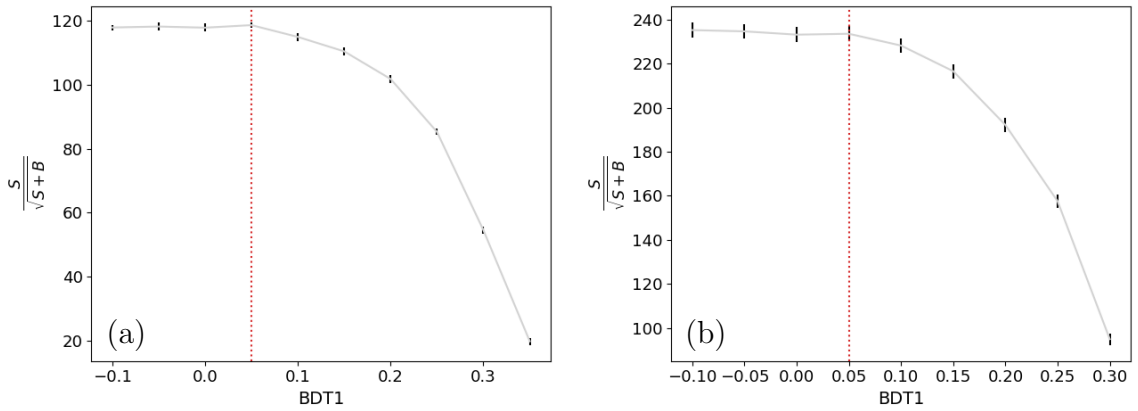


Figure 3.12: The signal yield significance *vs.* BDT1 response for $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ and $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ data is shown (a) and (b), respectively. The chosen cut, BDT1 > 0.05, is identified using the red dotted lines.

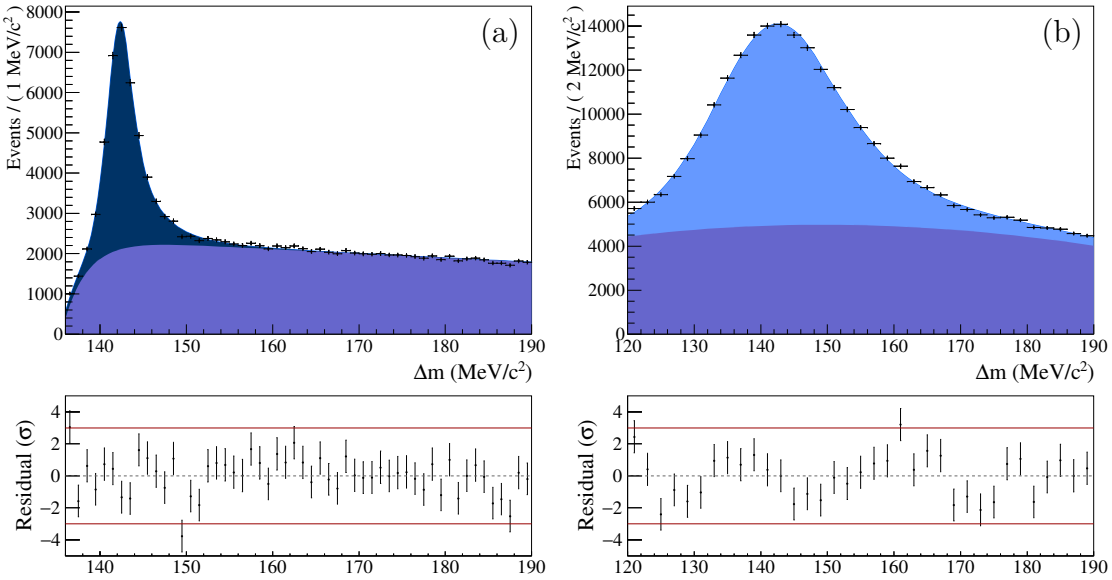


Figure 3.13: Fits to Δm at the optimised point BDT1 > 0.05 for $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ and $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ data is shown (a) and (b), respectively.

in the $D\gamma$ final state, the fitting region chosen to perform the optimisation selects for $D^{*} \to D\gamma$ signal candidates only. Once again, the sum of two Crystal Balls plus the $m(D^{*}) - m(D)$ mass difference PDF are employed to model the data. The parameter describing the curvature of the $m(D^{*}) - m(D)$ PDF is floated independently in all categories. The fits to the suppressed D decay modes are blind, so as not to bias the analysis during the optimisation procedure.

Simultaneous fits are performed across a range of neutral BDT values: $-0.1 <$

BDT2 $< 0.15(0.2)$ for the $D\pi^0$ ($D\gamma$) mode. A high purity sample of signal candidates is then selected by cutting on the BDT2 response that minimises the percentage error on the yield ratio of $B^{\pm} \to (D^* \to [SUP]_D\pi^0/\gamma)K^{\pm}$ to $B^{\pm} \to (D^* \to [FAV]_D\pi^0/\gamma)K^{\pm}$ decays (typically called $R_K^{\pi K,\pi^0/\gamma}$); the measured values of $R_K^{\pi K,\pi^0/\gamma}$ plotted against BDT cut are shown in Fig. 3.14. The chosen cut in the $D\pi^0$ mode is BDT2 > 0, which is 80% efficient on signal MC in both the $D^*\pi$ and D^*K sub-samples. The chosen cut in the $D\gamma$ mode is BDT2 > 0.05, which is 56% efficient on $D\gamma$ signal MC in both the $D^*\pi$ and D^*K sub-samples. The invariant-mass fits performed at the optimised points are provided in Fig. 3.15 for the $D\pi^0$ mode and Fig. 3.16 for the $D\gamma$ mode.

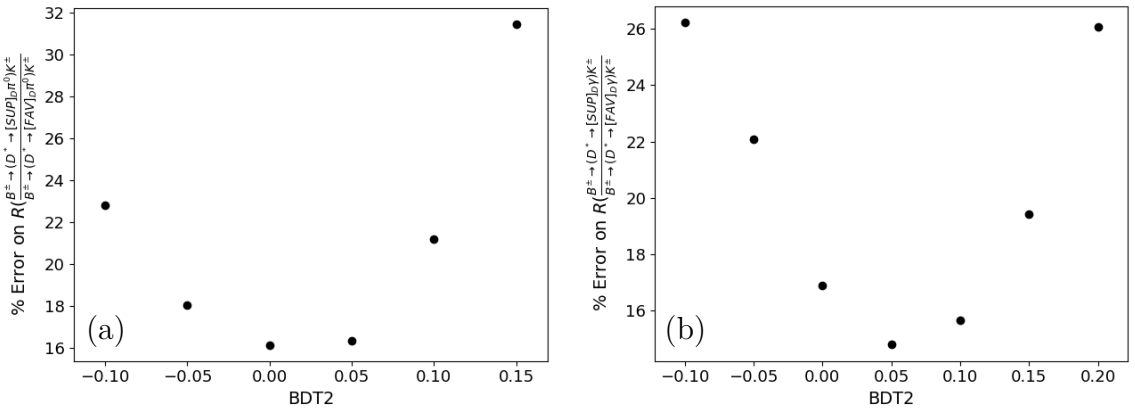


Figure 3.14: The percentage error on $R_K^{\pi K,\pi^0/\gamma}$ plotted against the selection cut on the neutral BDT (BDT2) is shown (a) for the $D\pi^0$ mode. The equivalent plot for the $D\gamma$ mode is shown in (b).

3.1.5 Multiple candidate removal

Events reconstructed in both D^* decay modes are kept in the $D\pi^0$ data sample and removed from the $D\gamma$ data sample. This choice is motivated by the lower statistics of the $D\pi^0$ final state and the desire for increased purity of signal data in the $D\gamma$ final state. By construction, this choice does not bias the analysis.

It is possible for multiple candidates per event to be present in the final data sample that passes all selection requirements; only one is kept, chosen at random. There are found to be ~ 1.20 candidates per event in the $D\pi^0$ mode and ~ 1.07 candidates per event in the $D\gamma$ mode. This reduced rate in the $D\gamma$ mode is due to

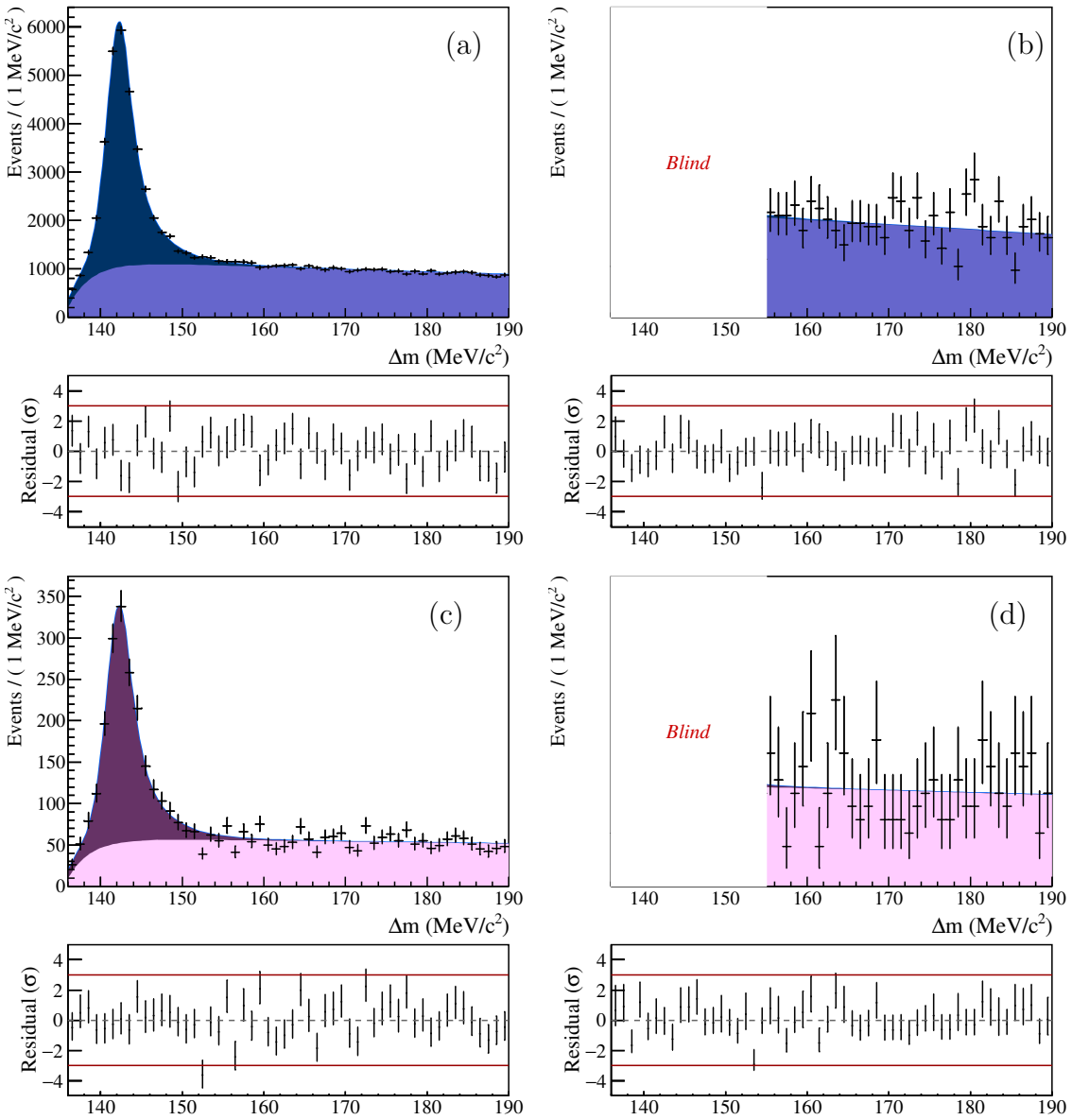


Figure 3.15: Simultaneous fit to Δm at the optimised point BDT2 > 0. Decays of the type $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ are depicted in (a); $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\pi^0)\pi^{\pm}$ in (b) (blind); $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)K^{\pm}$ in (c); and $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\pi^0)K^{\pm}$ in (d) (blind).

the tighter photon transverse momentum cut, removing substantial combinatorial backgrounds from low energy photons.

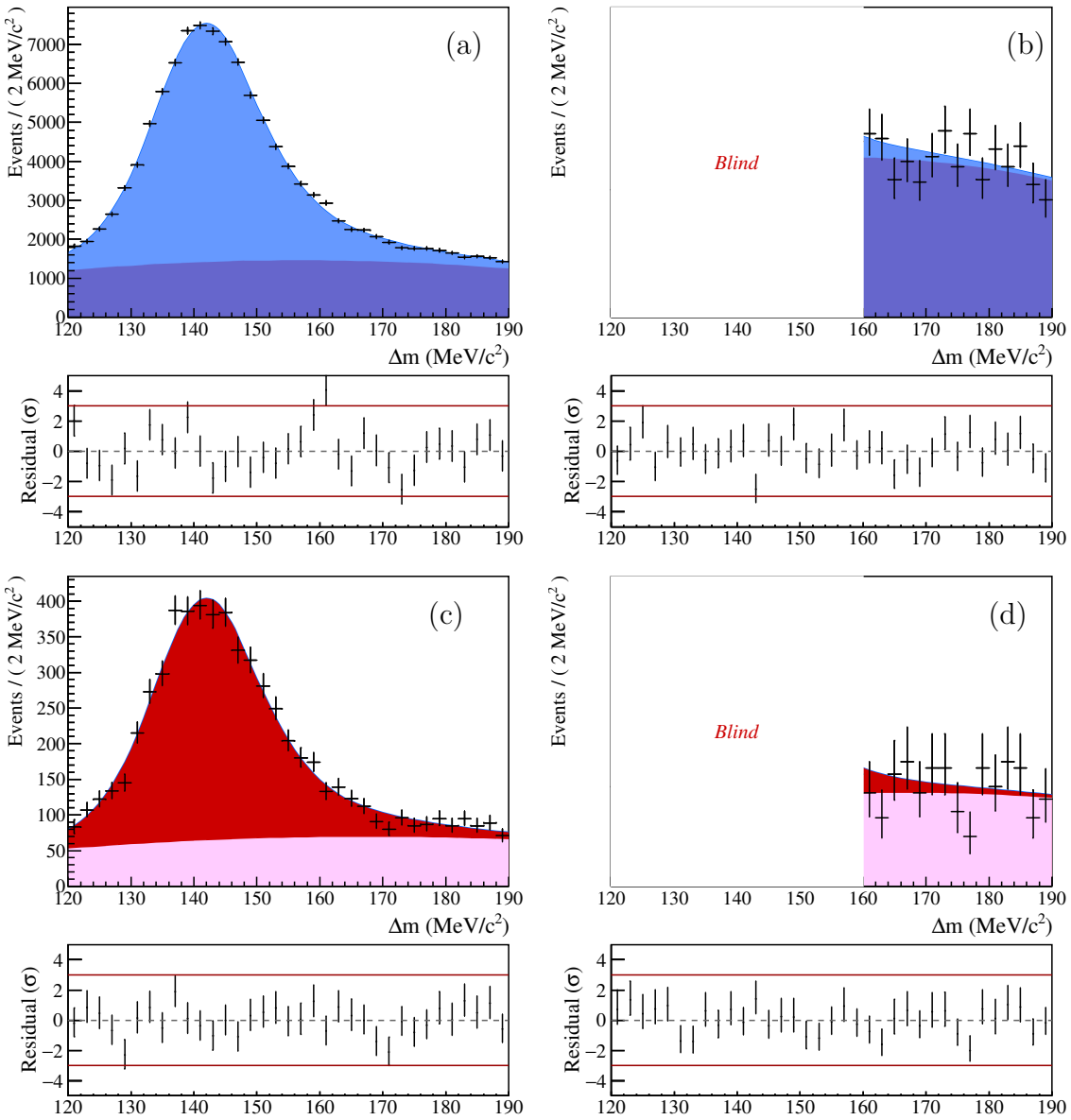


Figure 3.16: Simultaneous fit to Δm at the optimised point BDT2 > 0.05. Decays of the type $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ are depicted in (a); $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\gamma)\pi^{\pm}$ in (b) (blind); $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)K^{\pm}$ in (c); and $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\gamma)K^{\pm}$ in (d) (blind).

3.1.6 Particle identification requirements

The ability to distinguish between kaons and pions is crucial to this analysis, in order to separate Cabibbo-suppressed $B^{\pm} \to D^*K^{\pm}$ decays from Cabibbo-favoured $B^{\pm} \to D^*\pi^{\pm}$ decays. As well as isolating $B^{\pm} \to D^*K^{\pm}$ decays from their high statistic counterpart, this separation enables the measurement of CP observables in both decay modes.

Particle identification (PID) information is obtained from the RICH detector, described in Sec. 2.3.2. The charged pion and kaon from the $B^{\pm}$ decay vertex are required to occupy the momentum regions $5 < p < 100$ GeV/c and $0.5 < p_T < 10$ GeV/c, to lie within the RICH's kinematic range. A binary cut on the difference in log-likelihood, $\Delta LL_{K/\pi}$, between the kaon and pion mass hypotheses of 12 is then applied to these companion hadrons. Candidates with $\Delta LL_{K/\pi} > 12$ are placed in the $D^{*}K$ data samples, reconstructed under the kaon mass hypothesis; candidates with $\Delta LL_{K/\pi} < 12$ are placed in the $D^{*}\pi$ data samples, reconstructed under the pion mass hypothesis. Any true $D^{*}K$ ($D^{*}\pi$) candidates that fail this identification requirement end up in the $D^{*}\pi$ ($D^{*}K$) sample. This ensures that all candidates are accounted for, but no candidate appears simultaneously in both samples.

The simultaneous fit used to extract observables, described in Sec. 4.1, employs the strategy of fitting $B \to D^{*}K^{\pm}$ decays and $B^{\pm} \to D^{*}\pi^{\pm}$ decays together, where the yield of each mode is determined by summing the total number of correctly identified and mis-identified decays. The efficiency of the $\Delta LL_{K/\pi}$ cut on both companion types must therefore be known, so that the proportion of correctly identified decays in each sample can be accurately determined. The efficiency of the $\Delta LL_{K/\pi} > 12$ cut applied to the kaon companions in the $B^{\pm} \to D^{*}K^{\pm}$ sample is denoted ϵ^{K}_{PID}; the efficiency of the $\Delta LL_{K/\pi} < 12$ cut applied to the pion companions in the $B^{\pm} \to D^{*}\pi^{\pm}$ sample is denoted ϵ^{π}_{PID}.

The $\Delta LL_{K/\pi}$ response in simulation is not sufficiently well modelled, therefore a high statistics sample of $D^{*\pm} \to [K^{\mp}\pi^{\pm}]_{D^0}\pi^{\pm}$ decays is used to calibrate signal MC. Decays of this kind are chosen as both $D^{*\pm}$ decay products can be determined without the any RICH PID information; the track types of the D^{0} decay products can be unambiguously identified using their relative charge compared to that of the companion pion. This provides a high purity sample of pion and kaon tracks that can be used to calibrate the RICH detector performance. This depends on several factors, most notably the momentum, p, pseudorapidity, η, and the total number of tracks in the event, N_{Tr}.

The PID efficiencies for pions and kaons in the calibration sample are mapped in bins of (p, η), so that per-event efficiencies can be assigned to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma/\pi^0)\pi^{\pm}$ MC samples by identifying the corresponding bin that the companion falls into. Information from the N_{Tr} variable is not used as the event multiplicity is not well modelled in simulation (this will be discussed in more detail later in the sub-section). The average efficiency over all companion tracks is then taken to represent the overall PID efficiency, which is calculated separately for each D^* final state, data taking year, and magnetic polarity. The bachelor kinematics are highly similar across the $D^*\pi$ and D^*K samples, therefore both efficiencies can be calculated from the same $D^*\pi$ MC reference sample.

The resulting values for ϵ^{π}_{PID} and ϵ^{K}_{PID} are displayed in Tabs. 3.3 and 3.4 for the $D^* \to D\gamma$ and $D^* \to D\pi^0$ decay modes, respectively. The total efficiencies evaluated in the final row are fixed in the signal extraction fits and are shared across all D decay modes. These are luminosity-scaled averages over all data taking years and magnetic polarities, and are varied within the uncertainties quoted in the table when evaluating systematic uncertainties (see Sec. 4.5). These quoted uncertainties account for:

- The use of binned efficiencies: an alternative binning scheme is used to evaluate the average PID efficiencies and the difference between these numbers and those calculated using the default binning scheme is taking as a systematic [31].

- The finite size of the reference $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma/\pi^0)\pi^{\pm}$ simulation samples.

- A global 0.2% systematic for underlying methodological assumptions in `PIDCalib`, the LHCb software used to perform the calibration.

- Potential differences in PID efficiencies between decays with D meson final states $K\pi$, KK and $\pi\pi$, which can arise due to correlations between PID performance and other tracks in the event. The range of PID efficiencies calculated for each D final state separately in the partially reconstructed analysis was found to be 0.67%. This additional source of uncertainty is therefore added in quadrature

to the other sources described when calculating the error on ϵ^K_{PID}. Concerning ϵ^{π}_{PID}, this effect is found to be negligible.

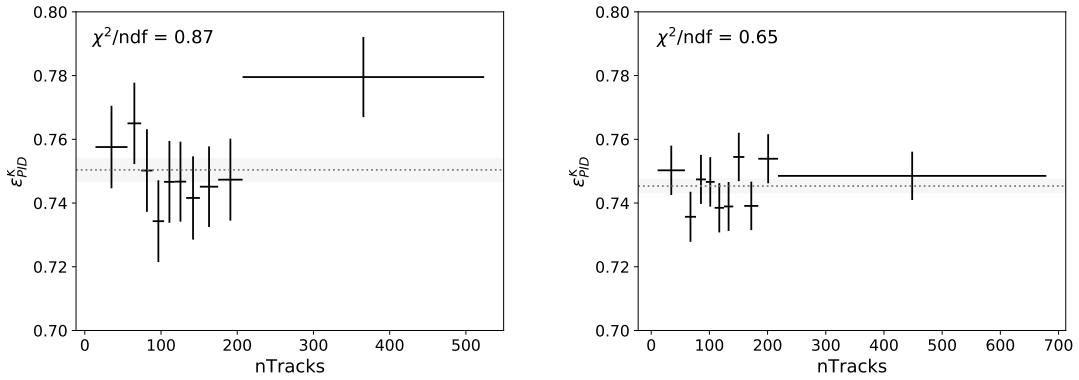


Figure 3.17: The kaon PID efficiency, luminosity averaged across years and magnetic polarities, is plotted in bins of equal occupancy of event multiplicity (N_{Tr}) for the $D\pi^0$ ($D\gamma$) final state on the left (right). ϵ^K_{PID} is represented by the grey dotted line, and the shaded region represents the 1σ statistical error on this value. All efficiencies were calculated using reference samples of $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0/\gamma)\pi^{\pm}$ MC. The χ^2/ndf values, calculated using Eq. (3.12), are also given.

The contamination of signal data with neutral combinatorial in the underlying event makes it difficult to obtain high purity signal samples using data. This motivated the use of simulation samples when performing the PID efficiency calculation, and a study was conducted to validate this choice at the expense of binning PID efficiencies in the N_{Tr} variable. The variation of ϵ^K_{PID} in 10 bins of N_{Tr} for the reference samples are shown in Fig. 3.17. The total kaon PID efficiencies $\pm 1\sigma$ are also depicted in grey. The χ^2/ndf was calculated for both D^* final states:

$$\chi^2/\text{ndf} = \frac{1}{N_{bins}} \sum_{i=1}^{N_{bins}} \frac{(\epsilon^i_{PID} - \epsilon^K_{PID})^2}{\sigma(\epsilon^i_{PID})^2 + \sigma(\epsilon^K_{PID})^2}, \tag{3.12}$$

where N_{bins} represents the total number of N_{Tr} bins, ϵ^i_{PID} is the total PID efficiency in bin i, ϵ^K_{PID} is the kaon PID efficiency for the given final state (displayed at the bottom of Tabs. 3.3 and 3.4) and $\sigma(\epsilon^i_{PID})$ and $\sigma(\epsilon^K_{PID})$ are their respective statistical errors. The χ^2/ndf was calculated to be 0.87 (0.65) for the $D\pi^0$ ($D\gamma$) final state. It was therefore concluded that there is no evidence of PID efficiency variation over N_{Tr}, justifying the use of simulation to evaluate PID efficiencies.

PID requirements are also applied to the D decay products. Kaons are required to have $\Delta LL_{K/\pi} > 2$ in Run 1 and $\Delta LL_{K/\pi} > 1$ in Run 2; pions are required to have $\Delta LL_{K/\pi} < -2$ in Run 1 and $\Delta LL_{K/\pi} < -1$ in Run 2. This reduces the occurrence of mis-identified D decays across the different modes, which would otherwise pose a significant background in the doubly Cabibbo-suppressed mode. The threshold is lowered for Run 2 as the PID performance at a given cut value improved compared to Run 1, primarily due to the removal of the aerogel radiator as discussed in Sec. 2.1.4. For this work, it is not necessary to evaluate the efficiencies of the D decay product PID cuts as all measurements are constructed as ratios, where the D final state is common to both the numerator and denominator.

	ϵ^{π}_{PID}		ϵ^{K}_{PID}	
Year	MagUp	MagDown	MagUp	MagDown
2011	99.36 ± 0.31	99.33 ± 0.32	72.00 ± 2.06	73.33 ± 2.24
2012	99.45 ± 0.26	99.48 ± 0.26	71.20 ± 1.65	70.77 ± 1.62
2015	99.64 ± 0.30	99.66 ± 0.29	74.70 ± 2.33	74.98 ± 2.24
2016	99.75 ± 0.22	99.64 ± 0.22	75.24 ± 1.21	74.79 ± 1.22
2017	99.75 ± 0.22	99.64 ± 0.22	75.24 ± 1.21	74.79 ± 1.22
2018	99.75 ± 0.22	99.64 ± 0.22	75.24 ± 1.21	74.79 ± 1.22
Total	99.64 ± 0.07		74.53 ± 0.76	

Table 3.3: PID efficiencies ϵ^{π}_{PID} and ϵ^{K}_{PID} calculated using $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ signal MC, where pions are used as a proxy for kaons. All efficiencies are given in percentages.

	ϵ^{π}_{PID}		ϵ^{K}_{PID}	
Year	MagUp	MagDown	MagUp	MagDown
2011	99.20 ± 0.64	99.37 ± 0.56	76.75 ± 4.10	71.06 ± 4.15
2012	99.40 ± 0.44	99.46 ± 0.42	71.50 ± 3.39	68.68 ± 3.33
2015	99.62 ± 0.37	99.66 ± 0.38	73.72 ± 3.42	77.19 ± 3.55
2016	99.75 ± 0.24	99.62 ± 0.26	75.74 ± 1.81	76.08 ± 1.80
2017	99.75 ± 0.24	99.62 ± 0.26	75.74 ± 1.81	76.08 ± 1.80
2018	99.75 ± 0.24	99.62 ± 0.26	75.74 ± 1.81	76.08 ± 1.80
Total	99.62 ± 0.10		75.04 ± 0.91	

Table 3.4: PID efficiencies ϵ^{π}_{PID} and ϵ^{K}_{PID} calculated using $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ signal MC, where pions are used as a proxy for kaons. All efficiencies are given in percentages.

3.1.7 Invariant-mass distributions

The $m(D^{*}\pi)$ and Δm distributions for $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ and $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ data, following the application of all selections described in this section (stripping, trigger, BDT, rectangular cuts and PID requirements), are shown in Fig. 3.18.

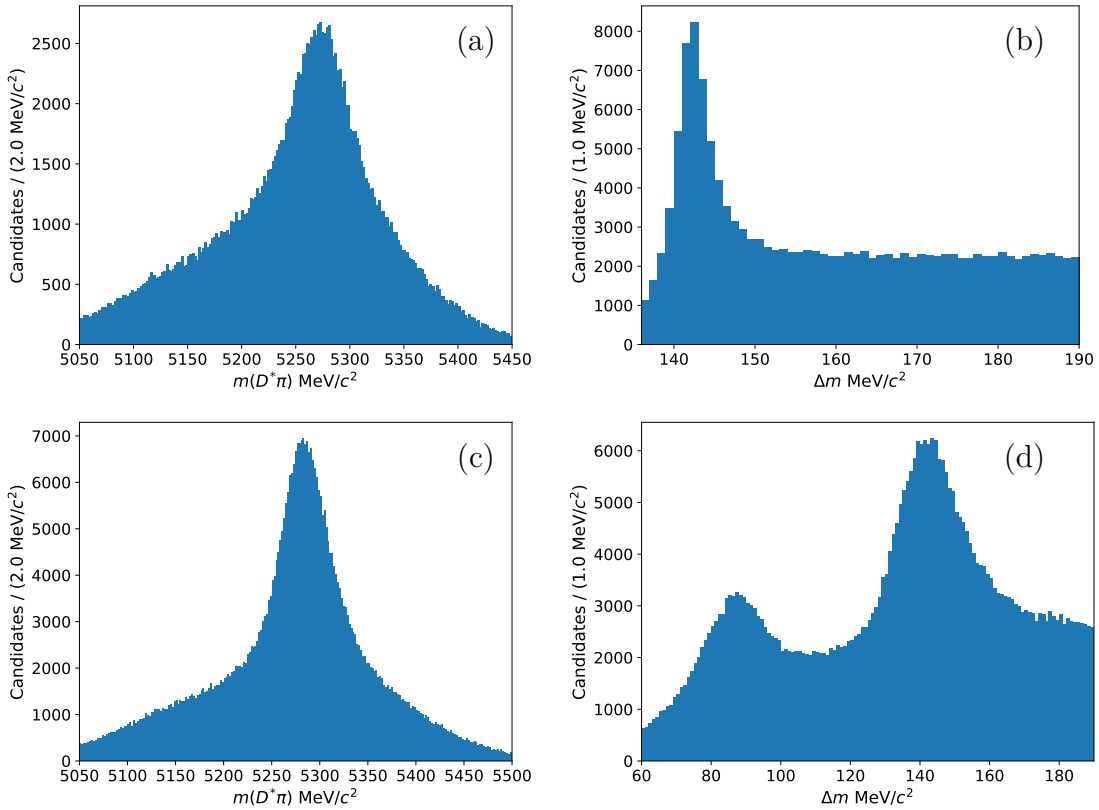


Figure 3.18: 1D histograms of $m(D^{*}\pi)$ (left) and Δm (right) for 2011-2018 favoured mode data post-stripping, trigger, BDT, rectangular cuts and PID requirements. The distribution ranges correspond to those used in the invariant-mass fits to data in Sec. 3.2. $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ data are shown in the top two plots, (a) and (b). $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ data are shown in the bottom two plots, (c) and (d).

3.2 Set up of double-1D (D1D) signal extraction fits

Events that pass the entire selection procedure described in Sec. 3.1 are fed into binned, extended maximum likelihood fits to determine the best model parameters

that describe the data. This work uses two separate invariant-mass fits to determine the observables listed in Sec. 1.4.4, one for data reconstructed in the $D^* \to D\pi^0$ final state, and one for data reconstructed in the $D^* \to D\gamma$ final state. The novel double-1D fitting technique developed for this analysis will be introduced in this section.

3.2.1 Binned, extended maximum likelihood

A given invariant mass distribution is described by N independent data points, $\boldsymbol{m} = \{m_1, m_2, ..., m_N\}$, where each data point follows the *probability density function* (PDF), $f(m; \boldsymbol{\theta})$. Here, $\boldsymbol{\theta}$ are the set of unknown variables that parameterise the shape of the distribution. The likelihood function:

$$\mathcal{L} = \prod_i^N f(m_i; \boldsymbol{\theta}) \tag{3.13}$$

represents the probability, for given values of $\boldsymbol{\theta}$, of observing the dataset $\boldsymbol{m}$. The maximum likelihood estimates (MLE) are the values $\boldsymbol{\theta}$ for which $\mathcal{L}$ has its global maximum. In practice, it is more convenient to search for the minimum of the negative log-likelihood function:

$$-\ln \mathcal{L} = -\sum_i^N \ln f(m_i; \boldsymbol{\theta}). \tag{3.14}$$

The final favoured mode datasets considered in this thesis contain $\mathcal{O}(10^5)$ events, therefore, in order to improve the computation efficiency of the likelihood function, binned datasets are used. This choice degrades the statistical errors on the observables in only their third significant figure. The likelihood function therefore takes the form:

$$\mathcal{L} = N! \prod_{i=1}^{n_B} \frac{P_i(\boldsymbol{\theta})^{N_i}}{N_i!}, \tag{3.15}$$

where N_i are the number of entries in bin i, for a total of N events over n_B bins. The variable $P_i(\boldsymbol{\theta})$ represents the expected probability for an event to lie in a particular bin. The numerical methods implemented by the `RooFit` package [84], which is used to construct and minimise the negative log-likelihood functions employed in this work, calculates this probability by multiplying the PDF evaluated at the bin

centre, m_i^c, by the bin width, δm_i:

$$P_i(\boldsymbol{\theta}) = \delta m_i \times f(m_i^c; \boldsymbol{\theta}). \tag{3.16}$$

When the number of events described by a particular PDF also needs to be measured, the extended maximum likelihood method is used [85]. Eq. (3.15) is then multiplied by a Poisson term for N observed events, when $\mu(\boldsymbol{\theta})$ events are expected:

$$\mathcal{L} = e^{-\mu(\boldsymbol{\theta})}\frac{\mu(\boldsymbol{\theta})^N}{N!}N!\prod_{i=1}^{n_B}\frac{P_i(\boldsymbol{\theta})^{N_i}}{N_i!} = \prod_{i=1}^{n_B} e^{-\mu_i(\boldsymbol{\theta})}\frac{\mu_i(\boldsymbol{\theta})^{N_i}}{N_i!}. \tag{3.17}$$

To get from the first to the second equation, the definitions $N = \sum_i N_i$ and $\mu_i(\boldsymbol{\theta}) = P_i(\boldsymbol{\theta})\mu(\boldsymbol{\theta})$ have been used. From the second equation, it can be seen that the binned extended maximum likelihood function can be interpreted as the product over all bins, $i \in [1, n_B]$, of the Poisson probabilities of observing N_i events, when $\mu_i(\boldsymbol{\theta})$ events are expected. Removing constant terms, the log-likelihood function to be minimised is therefore:

$$-\ln\mathcal{L}(\mu, \boldsymbol{\theta}; \boldsymbol{N}) = -\sum_i^{n_B} N_i \ln\mu_i(\mu, \boldsymbol{\theta}) + \mu. \tag{3.18}$$

3.2.2 D1D fitting method

The most challenging aspect of the fully reconstructed $B^{\pm} \to D^* h^{\pm}$ analysis is distinguishing signal from the many contaminating neutral combinatorial backgrounds. The strategy that has been developed trains a dedicated neutral BDT for each D^* decay mode to remove the majority of combinatorial events (see Sec. 3.1.2), then employs a simultaneous extended maximum likelihood fit for each D^* final state across multiple invariant-mass samples to pick out signal candidates.

The idea of using a simultaneous fit across multiple variables was first presented in the measurement of $\mathcal{R}(D^{*-}) \equiv \mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)/\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ [86], where backgrounds from double-charmed candidates were constrained using simultaneous fits across up to 4 mass samples. This was the inspiration for the double-1D technique developed in this work; the difference being that the $\mathcal{R}(D^{*-})$ analysis did not consider the impact on statistical errors from including candidates in more than one sample.

In this thesis, this effect is studied and accounted for, as will be described in this section and expanded on in Sec. 4.4.[1]

The two discriminating variables used in the fit are:

$$m(B) = m(D^*h) - m(D^*) + m(D^*)_{\text{PDG}} \tag{3.19}$$

$$\Delta m = \begin{cases} m(D^*) - m(D), & \text{for } D^* \to D\gamma \\ m(D^*) - m(D) - m(\pi^0) + m(\pi^0)_{\text{PDG}}, & \text{for } D^* \to D\pi^0, \end{cases} \tag{3.20}$$

where the PDG subscript refers to the meson masses listed on the PDG [19], $m(D^*)_{\text{PDG}} = 2006.85\,\text{MeV}/c^2$ and $m(\pi^0)_{\text{PDG}} = 134.98\,\text{MeV}/c^2$. The key benefit of using both variables is that *only* the signal modes peak in both distributions, allowing their yields to be constrained. The implementation of these simultaneous fits are outlined below.

As data reconstructed in the $D\pi^0$ or $D\gamma$ final state is read into the associated fit, it is split into the different mass samples depending on the location of the event in two-dimensional $m(B) - \Delta m$ space. For $m(B)$ fits, a cut is placed on the Δm values of that sample; for Δm fits, a cut is placed on the $m(B)$ values of that sample. These final, sample-defining requirements are referred to as *box cuts*, and are chosen to increase the signal purity of the 1D mass samples, as illustrated in Fig. 3.19.

The total yield within $m(B) - \Delta m$ space considered in the fit, for each component, is denoted N_{box}. Visually, this represents the number of candidates for a given decay mode that lie within the white boxes of Fig. 3.19. The transformation of N_{box} into the 1D yields of each mass sample is performed using the following equations:

$$N_{m(B)} = N_{box} \times \epsilon_{m(B)} \tag{3.21}$$

$$N_{\Delta m} = N_{box} \times \epsilon_{\Delta m}, \tag{3.22}$$

where the *box efficiencies*, $\epsilon_{m(B)}$ and $\epsilon_{\Delta m}$, which are measured using MC and fixed in the invariant-mass fits, are covered in more detail in Sec. 3.2.3. The log-likelihood

[1] Ongoing analyses of $B \to D^* l\nu$ decays in the Oxford LHCb group are now implementing the method developed in this work to correct statistical uncertainties for events that are present in more then one sample. Their initial findings suggest that the corrections are less significant for analyses of this kind, which are of much lower purity than the decays studied in this thesis.

function to be minimised therefore becomes:

$$\begin{aligned} -\ln \mathcal{L}(N_{box}, \boldsymbol{\theta}; \epsilon_{m(B)}, \epsilon_{\Delta m}, \boldsymbol{N}) = &- \sum_{i}^{n_B^{m(B)}} N_i^{m(B)} \ln \mu_i^{m(B)}(\epsilon_{m(B)}, N_{box}, \boldsymbol{\theta}) \\ &- \sum_{j}^{n_B^{\Delta m}} N_j^{\Delta m} \ln \mu_j^{\Delta m}(\epsilon_{\Delta m}, N_{box}, \boldsymbol{\theta}) \\ &+ N_{box}, \end{aligned} \tag{3.23}$$

where $N_i^{m(B)}$ ($N_j^{\Delta m}$) are the observed yields in the i^{th} (j^{th}) bin of the $m(B)$ (Δm) sample, $\mu_i^{m(B)}$ and $\mu_j^{\Delta m}$ are the expected yields per bin, and $n_B^{m(B)}$ and $n_B^{\Delta m}$ are the number of bins in the respective distributions.

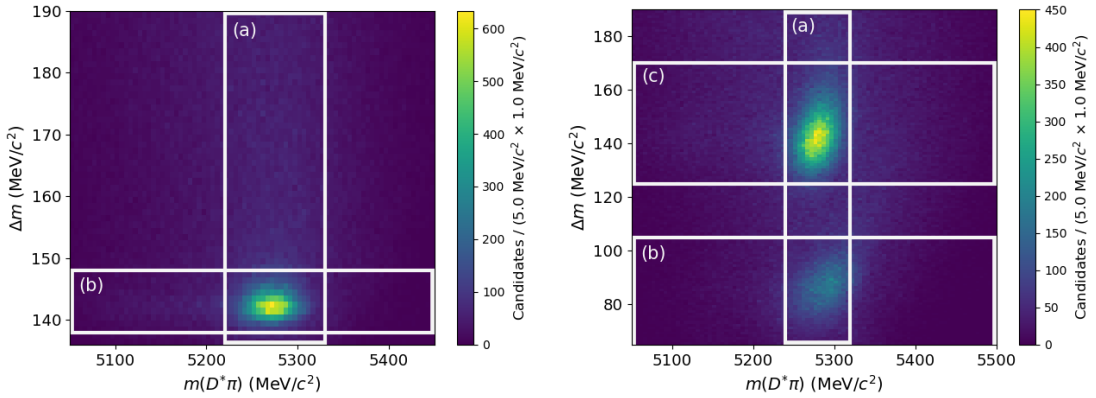


Figure 3.19: 2D histogram of Δm plotted against $m(D^*\pi)$ for 2011-2018 data, post-selections, with the $D^* \to D\pi^0$ mode on the left and the $D^* \to D\gamma$ mode on the right. The phase space shown is the region spanned by the invariant-mass fits, and the box cuts are indicated by the white boxes. Region (a) indicates the data considered in the Δm sample; region (b) indicates data considered in the $m(B)$ sample selecting $D\pi^0$ signal; region (c) indicates data considered in the $m(B)$ sample selecting $D\gamma$ signal.

The presence of $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ signal in the $D\gamma$ data tuples should be noted in the 2D plot depicted in Fig. 3.19: this is the partially-reconstructed $D\pi^0$ signal peak, where only one of the photons from the neutral pion decay has been reconstructed. The presence of this partially-reconstructed $D\pi^0$ peak motivates a simultaneous fit across 3 mass samples to model the $D\gamma$ datasets, one in Δm and two in $m(B)$. The box cuts in the Δm variable applied to each of the B mass samples are chosen to select for $D\gamma$ candidates and partially-reconstructed $D\pi^0$ candidates separately.

It can also be seen from the 2D distributions that a subset of events are included in two mass samples. This requires the correction of statistical uncertainties for the

double-counting of events where boxes overlap in $m(B) - \Delta m$ space; the method developed to handle this correction is described in Sec. 4.4. In the $D\gamma$ mode, candidates only appear in one of the $m(B)$ distributions, due to the non-overlapping box cuts in Δm, therefore no 'triple-counting' of events occurs.

For the $B^{\pm} \to (D^{*} \to D\pi^{0})h^{\pm}$ data sample, the box cuts are:

- Data included in the $m(B)$ sample is required to fall within the region $138\,\mathrm{MeV}/c^2 < \Delta m < 148\,\mathrm{MeV}/c^2$.
- Data included in the Δm sample is required to fall within the region $5220\,\mathrm{MeV}/c^2 < m(B) < 5330\,\mathrm{MeV}/c^2$.

For the $B^{\pm} \to (D^{*} \to D\gamma)h^{\pm}$ data sample, the box cuts are:

- Data included in the fully reconstructed $D\gamma$ $m(B)$ sample is required to fall within the region $125\,\mathrm{MeV}/c^2 < \Delta m < 170\,\mathrm{MeV}/c^2$.
- Data included in the partially reconstructed $D\pi^{0}$ $m(B)$ sample is required to fall within the region $60\,\mathrm{MeV}/c^2 < \Delta m < 105\,\mathrm{MeV}/c^2$.
- Data included in the Δm sample is required to fall within the region $5240\,\mathrm{MeV}/c^2 < m(B) < 5320\,\mathrm{MeV}/c^2$.

The $D\pi^{0}$ ($D\gamma$) data displayed in Fig. 3.18 is therefore split into 2 (3) samples. The final distributions that are fed into the invariant mass fits, with box cuts applied, are shown in Fig. 3.20 (Fig. 3.21). At this stage, the total number of candidates in the $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ mode, *i.e.* the number of candidates depicted in Fig. 3.20, is $\sim$ 180 000; the total number of candidates in the $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ mode, *i.e.* the number of candidates depicted in Fig. 3.21, is $\sim$ 470 000.

3.2.3 Box efficiencies

The invariant-mass fits used in this analysis are simultaneous across 2 or 3 mass slices. The total yield for each component is transformed into 1D yields within each slice using

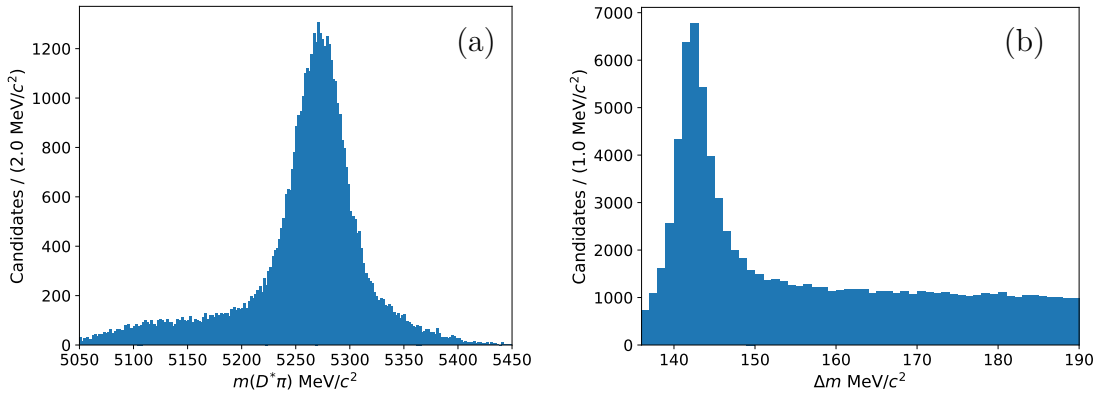


Figure 3.20: 1D histograms depicting 2011-2018 $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ data that is fed into the $D\pi^{0}$ signal-extraction fit. The $m(B)$ distribution with Δm mass window cut is shown in (a); the Δm distribution with $m(B)$ mass window cut is shown in (b).

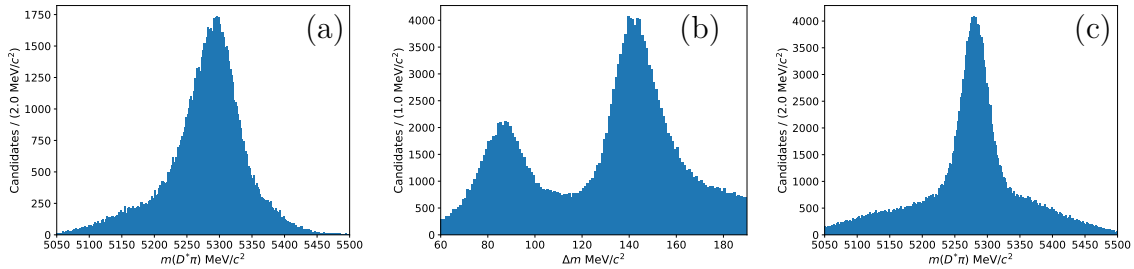


Figure 3.21: 1D histograms depicting 2011-2018 $B^{\pm} \to (D^{*0} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ data that is fed into the $D\gamma$ signal-extraction fit. The $m(B)$ distribution with the Δm mass window cut selecting for partially reconstructed $D\pi^{0}$ signal is shown in (a); the Δm distribution with $m(B)$ mass window cut is shown in (b); the $m(B)$ distribution with the Δm mass window cut selecting for fully reconstructed $D\gamma$ signal is shown in (c).

fixed box efficiencies, as given by Eq. (3.22). For the fit to data reconstructed in the $D\pi^{0}$ final state, there are two mass slices and therefore two box efficiencies to consider:

- $\epsilon^{FR}_{m(B)}$ is the proportion of events from the total yield that end up in the $m(B)$ sample of the fit. In terms of number of candidates, N:

$$\epsilon^{FR}_{m(B)} = \frac{N \text{ in } m(B)_{FR}}{N \text{ in } m(B)_{FR} \text{ OR } \Delta m}. \tag{3.24}$$

- $\epsilon_{\Delta m}$: the proportion of events from the total yield that end up in the Δm sample of the fit:

$$\epsilon_{\Delta m} = \frac{N \text{ in } \Delta m}{N \text{ in } m(B)_{FR} \text{ OR } \Delta m}. \tag{3.25}$$

For the fit to data reconstructed in the $D\gamma$ final state, there are three mass slices and therefore three box efficiencies to consider:

- $\epsilon^{FR}_{m(B)}$: the proportion of events from the total yield that end up in the $m(B)$ sample of the fit that selects for fully reconstructed $D^{*} \to D\gamma$ candidates:

$$\epsilon^{FR}_{m(B)} = \frac{N \text{ in } m(B)_{FR}}{N \text{ in } m(B)_{FR} \text{ OR } m(B)_{PR} \text{ OR } \Delta m}. \tag{3.26}$$

- $\epsilon^{PR}_{m(B)}$: the proportion of events from the total yield that end up in the $m(B)$ sample of the fit that selects for partially reconstructed $D^{*} \to D\pi^{0}$ candidates:

$$\epsilon^{PR}_{m(B)} = \frac{N \text{ in } m(B)_{PR}}{N \text{ in } m(B)_{FR} \text{ OR } m(B)_{PR} \text{ OR } \Delta m}. \tag{3.27}$$

- $\epsilon_{\Delta m}$: the proportion of events from the total yield that end up in the Δm sample of the fit:

$$\epsilon_{\Delta m} = \frac{N \text{ in } \Delta m}{N \text{ in } m(B)_{FR} \text{ OR } m(B)_{PR} \text{ OR } \Delta m}. \tag{3.28}$$

These 1D box efficiencies are calculated for each decay mode considered in the invariant-mass fits using simulation samples. Their values are given in Tab. 3.5, averaged across years and magnetic polarities, for events reconstructed in the following categories: $(D\pi^{0})K$, $(D\pi^{0})\pi$, $(D\gamma)K$ and $(D\gamma)\pi$. The uncertainties given in the table are due to the finite size of the simulation samples; systematic uncertainties are handled in Sec. 4.5, where the errors associated with each efficiency are increased in order to account for possible differences between data and simulation.

3.2.4 Selection efficiencies

Most reconstruction and selection efficiencies fully cancel in this analysis, as all observables are ratios of topologically identical final states. However, corrections must be applied to ratios with differing companion hadrons in the denominator and the numerator to account for any efficiency difference between $B^{\pm} \to D^{*}\pi^{\pm}$ and $B^{\pm} \to D^{*}K^{\pm}$ decays. The invariant-mass fits to data also read in selection efficiencies for the following use-cases:

Mode	Reco. as	$\epsilon_{\Delta m}$	$\epsilon^{FR}_{m(B)}$	$\epsilon^{PR}_{m(B)}$
$B^{\pm} \to (D^* \to D\pi^0)\pi^{\pm}$	$(D\pi^0)\pi$	97.37 ± 0.13	90.49 ± 0.24	-
	$(D\pi^0)K$	35.99 ± 0.43	94.52 ± 0.21	-
	$(D\gamma)\pi$	69.10 ± 0.22	6.96 ± 0.12	87.41 ± 0.16
	$(D\gamma)K$	18.31 ± 0.22	7.31 ± 0.14	91.32 ± 0.16
$B^{\pm} \to (D^* \to D\pi^0)K^{\pm}$	$(D\pi^0)\pi$	43.56 ± 0.92	92.23 ± 0.49	-
	$(D\pi^0)K$	98.06 ± 0.12	90.15 ± 0.25	-
	$(D\gamma)\pi$	42.37 ± 0.55	7.95 ± 0.30	85.17 ± 0.40
	$(D\gamma)K$	69.88 ± 0.23	7.04 ± 0.13	87.44 ± 0.17
$B^{\pm} \to (D^* \to D\gamma)\pi^{\pm}$	$(D\gamma)\pi$	92.78 ± 0.12	92.60 ± 0.13	0.15 ± 0.02
	$(D\gamma)K$	18.46 ± 0.22	98.50 ± 0.07	0.16 ± 0.02
$B^{\pm} \to (D^* \to D\gamma)K^{\pm}$	$(D\gamma)\pi$	35.29 ± 0.53	95.79 ± 0.22	0.12 ± 0.04
	$(D\gamma)K$	94.04 ± 0.12	92.45 ± 0.13	0.14 ± 0.02
Mis-reco. $B^{\pm} \to (D^* \to D\pi^0)\pi^{\pm}$	$(D\pi^0)\pi$	91.14 ± 0.22	27.57 ± 0.35	-
	$(D\pi^0)K$	79.94 ± 0.40	34.78 ± 0.48	-
	$(D\gamma)\pi$	46.48 ± 0.34	63.54 ± 0.32	19.39 ± 0.27
	$(D\gamma)K$	30.97 ± 0.36	68.30 ± 0.36	19.88 ± 0.31
Mis-reco. $B^{\pm} \to (D^* \to D\pi^0)K^{\pm}$	$(D\pi^0)\pi$	82.82 ± 0.72	33.00 ± 0.90	-
	$(D\pi^0)K$	91.21 ± 0.23	26.98 ± 0.36	-
	$(D\gamma)\pi$	40.53 ± 0.69	65.70 ± 0.67	17.86 ± 0.54
	$(D\gamma)K$	46.09 ± 0.35	63.95 ± 0.34	19.18 ± 0.28
Mis-reco. $B^{\pm} \to (D^* \to D\gamma)\pi^{\pm}$	$(D\pi^0)\pi$	86.44 ± 0.66	25.87 ± 0.85	-
	$(D\pi^0)K$	85.90 ± 0.84	25.66 ± 1.05	-
	$(D\gamma)\pi$	41.46 ± 0.46	62.51 ± 0.45	21.99 ± 0.39
	$(D\gamma)K$	35.22 ± 0.53	64.17 ± 0.53	22.43 ± 0.46
Mis-reco. $B^{\pm} \to (D^* \to D\gamma)K^{\pm}$	$(D\pi^0)\pi$	84.95 ± 1.51	28.67 ± 1.91	-
	$(D\pi^0)K$	86.36 ± 0.70	25.80 ± 0.89	-
	$(D\gamma)\pi$	36.41 ± 0.95	64.57 ± 0.95	21.89 ± 0.82
	$(D\gamma)K$	39.99 ± 0.48	63.77 ± 0.47	21.96 ± 0.40
$B^0 \to (D^{*\mp} \to D\pi^{\mp})\pi^{\pm}$	$(D\pi^0)\pi$	92.42 ± 0.48	24.83 ± 0.78	-
	$(D\pi^0)K$	84.20 ± 0.79	28.17 ± 0.97	-
	$(D\gamma)\pi$	45.99 ± 0.43	62.77 ± 0.42	21.15 ± 0.36
	$(D\gamma)K$	31.37 ± 0.45	66.56 ± 0.46	21.72 ± 0.40
$B^0 \to (D^{*\mp} \to D\pi^{\mp})K^{\pm}$	$(D\pi^0)\pi$	81.95 ± 1.65	32.60 ± 2.01	-
	$(D\pi^0)K$	92.31 ± 0.49	23.89 ± 0.79	-
	$(D\gamma)\pi$	40.86 ± 0.92	66.01 ± 0.89	19.85 ± 0.75
	$(D\gamma)K$	46.34 ± 0.44	63.10 ± 0.43	20.06 ± 0.36
$B^{\pm} \to D\rho^{\pm}$	$(D\pi^0)\pi$	84.25 ± 0.93	28.49 ± 1.15	-
	$(D\pi^0)K$	82.68 ± 1.02	28.84 ± 1.23	-
	$(D\gamma)\pi$	37.06 ± 0.58	65.56 ± 0.57	22.00 ± 0.50
	$(D\gamma)K$	37.29 ± 0.58	64.52 ± 0.58	21.39 ± 0.49
$B^{\pm} \to DK^{*\pm}$	$(D\pi^0)\pi$	81.75 ± 2.29	32.63 ± 2.78	-
	$(D\pi^0)K$	86.91 ± 0.86	27.41 ± 1.14	-
	$(D\gamma)\pi$	32.97 ± 1.22	66.76 ± 1.23	22.39 ± 1.09
	$(D\gamma)K$	39.20 ± 0.58	62.91 ± 0.58	22.81 ± 0.50
$B^{\pm} \to (D^* \to D\gamma)\rho^{\pm}$	$(D\gamma)\pi$	10.16 ± 0.69	88.95 ± 0.72	7.53 ± 0.61
	$(D\gamma)K$	22.85 ± 0.56	89.75 ± 0.41	5.55 ± 0.31
$B^{\pm} \to (D^* \to D\pi^0)\rho^{\pm}$	$(D\pi^0)\pi$	15.96 ± 1.89	84.84 ± 1.85	-
	$(D\pi^0)K$	44.93 ± 1.17	71.85 ± 1.06	-
	$(D\gamma)\pi$	12.27 ± 1.08	48.10 ± 1.65	47.67 ± 1.65
	$(D\gamma)K$	26.84 ± 0.72	34.17 ± 0.77	59.28 ± 0.80
$B^{\pm} \to (D^* \to D\gamma)K^{*\pm}$	$(D\gamma)K$	9.60 ± 0.61	88.52 ± 0.66	7.85 ± 0.56
$B^{\pm} \to (D^* \to D\pi^0)K^{*\pm}$	$(D\pi^0)K$	12.39 ± 1.54	88.04 ± 1.51	-
	$(D\gamma)K$	12.20 ± 0.92	47.29 ± 1.41	48.41 ± 1.41
$\Lambda_b^0 \to \Sigma_c^{\mp}\pi^{\pm}$	$(D\pi^0)\pi$	84.08 ± 2.15	28.03 ± 2.64	-
	$(D\gamma)\pi$	43.48 ± 1.95	47.36 ± 1.97	41.15 ± 1.94
$B_s^0 \to DK^{*0}$	$(D\pi^0)K$	79.00 ± 1.24	30.80 ± 1.40	-
	$(D\gamma)K$	35.39 ± 0.68	62.22 ± 0.69	23.75 ± 0.61
$B_s^0 \to (D^* \to D\pi^0)K^{\mp}\pi^{\pm}$	$(D\pi^0)K$	14.13 ± 0.89	94.47 ± 0.60	-
	$(D\gamma)K$	16.73 ± 0.64	7.94 ± 0.48	88.11 ± 0.57
$B_s^0 \to (D^* \to D\gamma)K^{\mp}\pi^{\pm}$	$(D\gamma)K$	5.60 ± 0.32	98.89 ± 0.15	0.11 ± 0.05
Mis-reco. $B_s^0 \to (D^* \to D\pi^0)K^{*0}$	$(D\pi^0)K$	60.16 ± 2.21	48.58 ± 2.25	-
	$(D\gamma)K$	29.49 ± 1.33	68.73 ± 1.35	20.59 ± 1.18
Mis-reco. $B_s^0 \to (D^* \to D\gamma)K^{*0}$	$(D\pi^0)K$	76.19 ± 2.94	37.62 ± 3.34	-
	$(D\gamma)K$	33.86 ± 1.37	68.31 ± 1.35	19.82 ± 1.15

Table 3.5: 1D box efficiencies for all fit components. All efficiency values are given in percentages. The total efficiencies over all years are luminosity averaged.

- Branching ratios, which must be corrected for by selection efficiencies, are used to constrain certain backgrounds.

- Mis-identified decays, fixed in yield relative to correctly identified decays, must be corrected for by selection efficiencies due to B mass displacements caused by assigning the companion particle the wrong mass hypothesis. This displacement has a noticeable effect on the efficiency for a candidate to pass the box cuts outlined in the previous sub-section.

These corrections will be covered in more detail in Sec. 4.1.4 and Sec. 3.3, respectively, but the efficiencies used to make them are given in Tabs. 3.6 and 3.7 for signal and background decay modes. In these tables, the total selection efficiency is labelled ϵ_{tot}, which represents the product of the detector acceptance efficiency, ϵ_{acc}, with the efficiency of the combined reconstruction, stripping, trigger and offline selection, ϵ_{sel}, and the box efficiency, ϵ_{box}. Note that PID efficiencies are not included in ϵ_{sel}: these are handled separately when parameterising D^*K and $D^*\pi$ yields in Sec. 3.3.5.

The box efficiency for each mode is defined as the proportion of events that end up in at least one mass sample:

$$\epsilon_{box} = \begin{cases} \frac{N \text{ in } m(B)_{FR} \text{ OR } \Delta m}{N \text{ in 2D}} & \text{for } D^* \to D\pi^0 \text{ candidates} \\ \frac{N \text{ in } m(B)_{FR} \text{ OR } m(B)_{PR} \text{ OR } \Delta m}{N \text{ in 2D}} & \text{for } D^* \to D\gamma \text{ candidates,} \end{cases} \tag{3.29}$$

where 'N in 2D' represents the number of candidates, N, in the 2D $m(B) - \Delta m$ phase space spanning the fitting region, depicted in Fig. 3.19.

Mode	Reco. as	ϵ_{acc}	ϵ_{sel}	ϵ_{box}	ϵ_{tot}
$B^{\pm} \to (D^* \to D\pi^0)\pi^{\pm}$	$(D\pi^0)\pi$	15.7574 ± 0.0229	0.1069 ± 0.0011	99.0787 ± 0.0934	0.0167 ± 0.0002
	$(D\pi^0)K$	15.7574 ± 0.0229	0.0924 ± 0.0010	90.2896 ± 0.3173	0.0131 ± 0.0001
	$(D\gamma)\pi$	15.7574 ± 0.0229	0.3191 ± 0.0018	93.9380 ± 0.1426	0.0472 ± 0.0003
	$(D\gamma)K$	15.7574 ± 0.0229	0.2464 ± 0.0016	90.0140 ± 0.2018	0.0349 ± 0.0002
$B^{\pm} \to (D^* \to D\pi^0)K^{\pm}$	$(D\pi^0)\pi$	16.0077 ± 0.0233	0.0909 ± 0.0010	89.1048 ± 0.3504	0.0130 ± 0.0002
	$(D\pi^0)K$	16.0077 ± 0.0233	0.1003 ± 0.0011	99.2512 ± 0.1046	0.0159 ± 0.0002
	$(D\gamma)\pi$	16.0077 ± 0.0233	0.2257 ± 0.0016	93.5707 ± 0.1760	0.0338 ± 0.0003
	$(D\gamma)K$	16.0077 ± 0.0233	0.2945 ± 0.0018	93.6890 ± 0.1542	0.0442 ± 0.0003
$B^{\pm} \to (D^* \to D\gamma)\pi^{\pm}$	$(D\gamma)\pi$	17.6739 ± 0.0257	0.5274 ± 0.0031	97.8897 ± 0.0878	0.0912 ± 0.0006
	$(D\gamma)K$	17.6739 ± 0.0257	0.4205 ± 0.0028	92.8401 ± 0.1662	0.0690 ± 0.0005
$B^{\pm} \to (D^* \to D\gamma)K^{\pm}$	$(D\gamma)\pi$	18.0172 ± 0.0262	0.3710 ± 0.0026	93.2694 ± 0.1725	0.0623 ± 0.0005
	$(D\gamma)K$	18.0172 ± 0.0262	0.4995 ± 0.0030	98.0959 ± 0.0839	0.0883 ± 0.0006

Table 3.6: Total selection efficiencies, ϵ_{tot}, for the signal modes, where the quoted uncertainties are derived from the finite MC sample sizes. ϵ_{acc} represents the detector acceptance efficiency; ϵ_{sel} represents the combined reconstruction, stripping, trigger and offline selection efficiency; ϵ_{box} represents the box efficiency. All efficiencies are given in percentages.

Mode	Reco. as	ϵ_{acc}	ϵ_{sel}	ϵ_{box}	ϵ_{tot}
Mis-reco. $B^{\pm} \to (D^* \to D\pi^0)\pi^{\pm}$	$(D\pi^0)\pi$	15.7574 ± 0.0229	0.1769 ± 0.0014	64.7146 ± 0.3792	0.0180 ± 0.0002
	$(D\gamma)\pi$	15.7574 ± 0.0229	0.2072 ± 0.0015	75.1508 ± 0.3119	0.0245 ± 0.0002
Mis-reco. $B^{\pm} \to (D^* \to D\gamma)\pi^{\pm}$	$(D\pi^0)\pi$	17.6739 ± 0.0257	0.0587 ± 0.0010	55.5585 ± 0.8653	0.0058 ± 0.0001
	$(D\gamma)\pi$	17.6739 ± 0.0257	0.1853 ± 0.0018	73.7572 ± 0.4270	0.0242 ± 0.0003
$B^{\pm} \to D\rho^{\pm}$	$(D\pi^0)\pi$	17.0000 ± 0.0225	0.0148 ± 0.0004	51.3922 ± 1.4674	0.0013 ± 0.0001
	$(D\gamma)\pi$	17.0000 ± 0.0225	0.0476 ± 0.0008	70.1801 ± 0.7444	0.0057 ± 0.0001
$B^0 \to (D^{*\mp} \to D\pi^{\mp})\pi^{\pm}$	$(D\pi^0)\pi$	15.7574 ± 0.0229	0.0589 ± 0.0012	58.9851 ± 1.0002	0.0055 ± 0.0001
	$(D\gamma)\pi$	15.7574 ± 0.0229	0.2078 ± 0.0022	74.7002 ± 0.4611	0.0245 ± 0.0003
$B^{\pm} \to (D^* \to D\pi^0)\rho^{\pm}$	$(D\pi^0)\pi$	17.0000 ± 0.0225	0.0099 ± 0.0005	39.3217 ± 2.2057	0.0007 ± 0.0000
	$(D\gamma)\pi$	17.0000 ± 0.0225	0.0120 ± 0.0005	77.9278 ± 1.7406	0.0016 ± 0.0001
$B^{\pm} \to (D^* \to D\gamma)\rho^{\pm}$	$(D\gamma)\pi$	17.0000 ± 0.0225	0.0246 ± 0.0008	80.1022 ± 1.1797	0.0033 ± 0.0001
$B_s^0 \to (D^* \to D\pi^0)K^{\mp}\pi^{\pm}$	$(D\pi^0)K$	17.0000 ± 0.0225	0.0080 ± 0.0004	82.1002 ± 1.6251	0.0011 ± 0.0001
	$(D\gamma)K$	17.0000 ± 0.0225	0.0170 ± 0.0005	91.9845 ± 0.7460	0.0027 ± 0.0001
$B_s^0 \to (D^* \to D\gamma)K^{\mp}\pi^{\pm}$	$(D\gamma)K$	17.0000 ± 0.0225	0.0267 ± 0.0006	91.1923 ± 0.6686	0.0041 ± 0.0001
Mis-reco. $B_s^0 \to (D^* \to D\pi^0)K^{*0}$	$(D\pi^0)K$	17.0000 ± 0.0225	0.0111 ± 0.0004	34.6833 ± 1.7099	0.0007 ± 0.0000
	$(D\gamma)K$	17.0000 ± 0.0225	0.0131 ± 0.0005	70.5882 ± 1.5175	0.0016 ± 0.0001
Mis-reco. $B_s^0 \to (D^* \to D\gamma)K^{*0}$	$(D\pi^0)K$	17.0000 ± 0.0225	0.0040 ± 0.0002	31.2869 ± 2.5274	0.0002 ± 0.0000
	$(D\gamma)K$	17.0000 ± 0.0225	0.0126 ± 0.0004	75.4436 ± 1.4182	0.0016 ± 0.0001

Table 3.7: Total selection efficiencies, ϵ_{tot}, used to constrain select background modes in the invariant-mass fits, where the quoted uncertainties are derived from the finite MC sample sizes. ϵ_{acc} represents the detector acceptance efficiency; ϵ_{sel} represents the combined reconstruction, stripping, trigger and offline selection efficiency; ϵ_{box} represents the box efficiency. All efficiencies are given in percentages. The detector acceptance efficiency for many backgrounds is approximated at 17% as the data fits are not sensitive to the exact values.

3.3 Decay modes considered and their parameterisation

A number of background decay modes cannot be separated from signal during the selection procedure, therefore dedicated PDFs are needed in the signal extraction fits to model these contributions. An overview of all decay modes and descriptions of the PDFs used to model signal and background components are given in this section.

3.3.1 Overview

The decays considered in this analysis can be grouped into six categories. In the following, $h = \pi/K$:

1. **Signal modes**: fully reconstructed $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0/\gamma)h^{\pm}$ decays and partially reconstructed $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)h^{\pm}$ decays.

2. **Mis-reconstructed backgrounds**: a charged pion or neutral particle from the true decay has been missed and a neutral particle from the underlying event added to the D meson to form a D^{*0} candidate. These decays form broad peaks under the signal channels in $m(B)$, but are flat in Δm phase space. In the favoured D decay mode fit category, the associated components are mis-reconstructed $B^{\pm} \to (D^* \to D\pi^0/\gamma)h^{\pm}$ decays, $B \to Dh^{\pm}\pi$ decays and $\overset{(-)}{B^0} \to (D^{*\pm} \to D\pi^{\pm})h^{\mp}$ decays. Additional backgrounds infiltrate the invariant mass spectrum in the lower statistics modes: in the $D \to KK$ category, $\Lambda_b^0 \to (\Sigma_c^+ \to (\Lambda_c^+ \to pK^-\pi^+))h^-$ decays must be considered; in the doubly-Cabibbo suppressed mode, $B_s^0 \to D^0K^+\pi^-$ must be accounted for.

3. **Partially reconstructed backgrounds**: contain a true D^{*0} or charged $D^{*\mp}$ candidate. A charged pion or neutral particle, either the decay product of the D^* or an associated companion particle, has been missed in the final state. These decays peak low in B mass and form a combination of peaking (events for which a true D^{*0} is reconstructed) and flat (events for which a neutral from the underlying event is combined with the D^0 meson) distributions in Δm.

Partially-reconstructed decays are of the form $B \to D^* h^{\pm}\pi$. In the doubly-Cabibbo suppressed mode, $B_s^0 \to D^{*0}K^+\pi^-$ decays must also be considered.

4. **Mis-identified decays**: the companion hadron is assigned the wrong mass hypothesis. These decays are shifted upwards/downwards in B mass due to the incorrect mass assignment; in Δm, they have similar shapes to their correctly identified counterparts.

5. **Combinatorial background**: the correct neutral D^* meson is reconstructed and combined with a random track, or a D meson is combined with a neutral particle from the underlying event and a random track. These events are flat in $m(B)$ and have peaking (true D^*) and flat (fake D^*) components in Δm.

6. **Crossfeed**: the decay products of Cabibbo-favoured D decays are both mis-identified, therefore infiltrating the doubly-Cabibbo suppressed sample. Crossfeed components are assigned for all signal components, and peak directly under the signal in B mass and Δm. The rate of these events is therefore fixed in the signal extraction fits.

Approximately 15 (10) million simulated events are generated within the LHCb detector acceptance for the $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0(\gamma))K^{\pm}$ and $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0(\gamma))\pi^{\pm}$ signal modes using the procedure described in Sec. 2.4. The generation is performed to reproduce the conditions for each data taking year separately, where the relative size of each per-year sample is representative of the integrated luminosity of LHCb data collected during that period. Such a large sample size is needed in order to evaluate selection efficiencies and model signal mode line shapes due to the low reconstruction efficiency of neutral particles, as discussed Sec. 2.3.3. Similar sample sizes are also generated for the background modes listed above, in order to determine the component PDFs any any relevant efficiencies.

In order to obtain PDFs for each component, MC events reconstructed in the $D\pi^0$ final state are fit to with unbinned maximum likelihood functions simultaneously across Δm and $m(B)$ distributions. Box cuts are applied (see Sec. 3.2.2): events

considered in Δm are required to fall in the mass window $5220\,\mathrm{MeV}/c^2 < m(B) < 5330\,\mathrm{MeV}/c^2$; events considered in $m(B)$ are required to fall in the mass window $138\,\mathrm{MeV}/c^2 < \Delta m < 148\,\mathrm{MeV}/c^2$. For MC events reconstructed in the $D\gamma$ final state, unbinned maximum likelihood fits are performed simultaneously to the Δm distribution and 2 slices of $m(B)$. Events considered in the Δm sample are required to fall in the mass window $5240\,\mathrm{MeV}/c^2 < m(B) < 5320\,\mathrm{MeV}/c^2$; events considered in the $m(B)$ samples are required to fall in the mass windows $125\,\mathrm{MeV}/c^2 < \Delta m < 170\,\mathrm{MeV}/c^2$ for fully reconstructed $D\gamma$ candidates, or $60\,\mathrm{MeV}/c^2 < \Delta m < 105\,\mathrm{MeV}/c^2$ for partially reconstructed $D\pi^0$ candidates. In the text and figures below, the $m(B)$ slice around the $D\pi^0$ peak is referred to as the PR (partially-reconstructed) $m(B)$ distribution, and the $m(B)$ slice around the $D\gamma$ peak is referred to as the FR (fully reconstructed) $m(B)$ distribution.

3.3.2 Signal modes

The signal modes are predominantly parameterised by the sum of two Crystal Ball (CB) functions, given by Eqs. (3.30) and (3.31) as a function of reconstructed mass, m:

$$f_{CB}\left(m;\alpha,n,\mu,\sigma\right) \propto \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2}, & \text{if } \frac{m-\mu}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}} \left(\frac{n}{|\alpha|} - |\alpha| - \frac{m-\mu}{\sigma}\right)^{-n}, & \text{if } \frac{m-\mu}{\sigma} \leq -\alpha. \end{cases} \tag{3.30}$$

Crystal Balls are Gaussian-like functions with additional tail parameters that facilitate the modelling of resonances for which some fraction of the particles' energy and momenta have been lost during detection. The parameters μ and σ are the mean and width of the Gaussian, respectively; the power-law tail of the function is parameterised by n and starts $\alpha\sigma$ away from the peak position. The tail parameters have the typical ranges: $\alpha \in [-5, 5]$ and $n \in [0, 10]$; they are highly correlated, and it is often necessary to fix one of them to ensure the minimisation procedure converges. The double Crystal Ball function (DCB) has an additional floating parameter, k, which determines the relative size of the individual CBs:

$$\begin{aligned} f_{DCB}(m;k,\alpha_1,n_1,\mu_1,\sigma_1,\alpha_2,n_2,\mu_2,\sigma_2) = k \times f_{CB,1}(m;\alpha_1,n_1,\mu_1,\sigma_1) \\ + (1-k) \times f_{CB,2}(m;\alpha_2,n_2,\mu_2,\sigma_2). \end{aligned} \tag{3.31}$$

For the signal modes, $f_{CB,1}$ has a radiative tail extending to low invariant mass and $f_{CB,2}$ has a tail extending to high invariant mass in order to describe the reconstruction and resolution effects of neutral particles.

Decay	Reconstructed as $D\pi^0$		Reconstructed as $D\gamma$		
	$m(B)$	Δm	$m(B)_{\mathrm{PR}}$	Δm	$m(B)_{\mathrm{FR}}$
$B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$	f_{DCB}	f_{DCB}	$f_{DCB} + f_G$	f_{DCB}	f_{CB}
$B^{\pm} \to (D^* \to D\gamma)h^{\pm}$	-	-	f_{CB}	f_{DCB}	f_{DCB}

Table 3.8: Summary of PDFs used to parameterise signal decay modes, where f_{CB} represents the Crystal Ball function defined in Eq. (3.30), f_{DCB} represents the double Crystal Ball function defined in Eq. (3.31) and f_G represents a Gaussian function. Mis-reconstructed $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays that end up in the $D\pi^0$ data tuples are not considered signal decays.

A summary of the PDFs used to describe the signal modes is given in Tab. 3.8; more details are given below. In the fit to data, the PDF means are shared across $D^*\pi$ and D^*K and are allowed to float for all signal modes. The widths in the $D^*\pi$ mode also vary freely; the widths in D^*K, defined respectively to these, are floated then fixed in the data fits to allow for differences in resolution between the two B meson decay modes due to their different Q values. The α_1 and α_2 tail parameters are also floated then fixed in the fit to data, to account for any mis-modelling of the distribution tails in simulation. These parameters are varied within the uncertainties measured by the data fit to determine the associated systematic uncertainty. All other shape parameters are fixed to the values found in simulation, and varied within the uncertainties measured by the MC fits when evaluating systematics.

$B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays reconstructed in the $D\pi^0$ final state

To determine the signal PDFs for $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays reconstructed in the $D\pi^0$ final state, a simultaneous fit across $m(B)$ and Δm is performed to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ MC. The 2D MC distribution depicted in Fig. 3.22 is considered in

the fit. A total floating yield is assigned to the MC falling within the region selected by the box cuts, outlined by white rectangles. The total number of events is then split into each mass sample by multiplying the total yield by the relevant box efficiency.

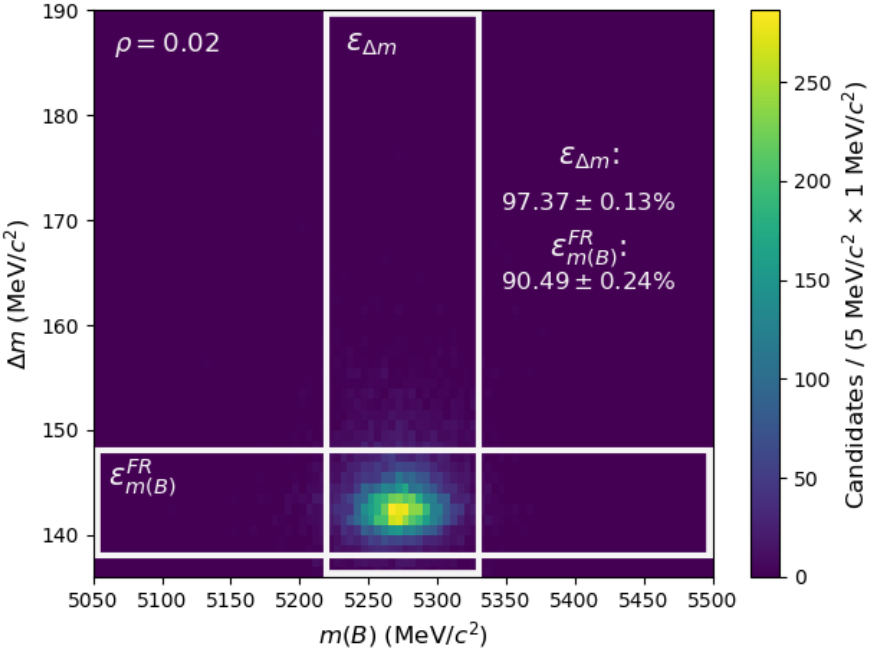


Figure 3.22: 2D distribution, $m(B)$ vs. Δm, of $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ signal MC. The vertical (horizontal) rectangle encompasses the MC included in the Δm ($m(B)$) sample, where the box efficiencies are printed as a reference. The efficiencies match up to those list in Tab. 3.5. The Pearson correlation coefficients, ρ, of the 2D mass distributions are also given.

The simultaneous fit is shown in Fig. 3.23. The residuals displayed below the fits are defined as the difference between the data and the PDF projection at the bin centre, divided by the Poisson error on the number of candidates in that bin. The n parameters are fixed to values of 10 to aid stability. The values of α are allowed to vary freely, settling on small positive values for α_1 and small negative values for α_2 in both mass samples. The relative contribution of the two Crystal Ball functions, k, floats in both fits. The means and widths vary freely but are shared across both Crystal Balls within each sample: $\mu_1 = \mu_2, \sigma_1 = \sigma_2$. In $m(B)$, $\mu(B) \sim 5274\,\text{MeV}/c^2$ and $\sigma(B) \sim 22\,\text{MeV}/c^2$; in Δm, $\mu(\Delta) \sim 142\,\text{MeV}/c^2$ and $\sigma(\Delta) \sim 2\,\text{MeV}/c^2$.

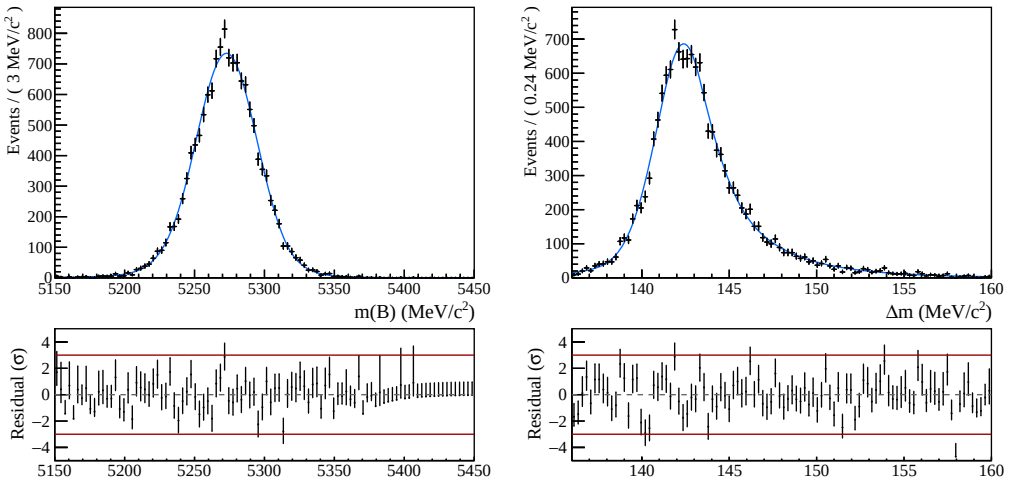


Figure 3.23: Simultaneous fit to $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ signal MC in $m(B)$ and Δm. 2011-2018 samples are used, where all years are considered together.

$B^{\pm} \to (D^{*} \to D\pi^{0})h^{\pm}$ decays reconstructed in the $D\gamma$ final state

To determine the signal PDFs for $B^{\pm} \to (D^{*} \to D\pi^{0})h^{\pm}$ decays reconstructed in the $D\gamma$ final state, a simultaneous fit across Δm and two slices of $m(B)$ is performed to $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ MC. The 2D MC distribution depicted in Fig. 3.24 is considered in the fit.

The simultaneous fit is shown in Fig. 3.23. The MC distribution in Δm is parameterised by the sum of two CBs, which share a single mean and width: $\mu_1 = \mu_2, \sigma_1 = \sigma_2$. These obtain the values $\mu(\Delta) \sim 85\,\text{MeV}/c^2$ and $\sigma(\Delta) \sim 9\,\text{MeV}/c^2$. The PR $m(B)$ slice is modelled by the sum of two CBs and a Gaussian function. The two CBs share a mean, found to be $\mu(B)_{PR} \sim 5279\,\text{MeV}/c^2$, but the Gaussian mean is offset by $\sim 27\,\text{MeV}/c^2$ above this to better describe the asymmetry of the resonance. All three functions are assigned different widths, and these find the values $\sigma_1(B)_{PR} \sim 26\,\text{MeV}/c^2$, $\sigma_2(B)_{PR} \sim 45\,\text{MeV}/c^2$ and $\sigma_3(B)_{PR} \sim 18\,\text{MeV}/c^2$ for the two CBs and the Gaussian, respectively. The n parameters of the CBs in both mass samples are fixed to values of 10 to aid stability. The values of α are allowed to vary freely, settling on small positive values for α_1 and small negative values for α_2. There is a low efficiency for partially reconstructed $D\pi^0$ MC in the $D\gamma$ $m(B)$ slice, as can be seen in Fig. 3.24. These events are parameterised using a CB function, where

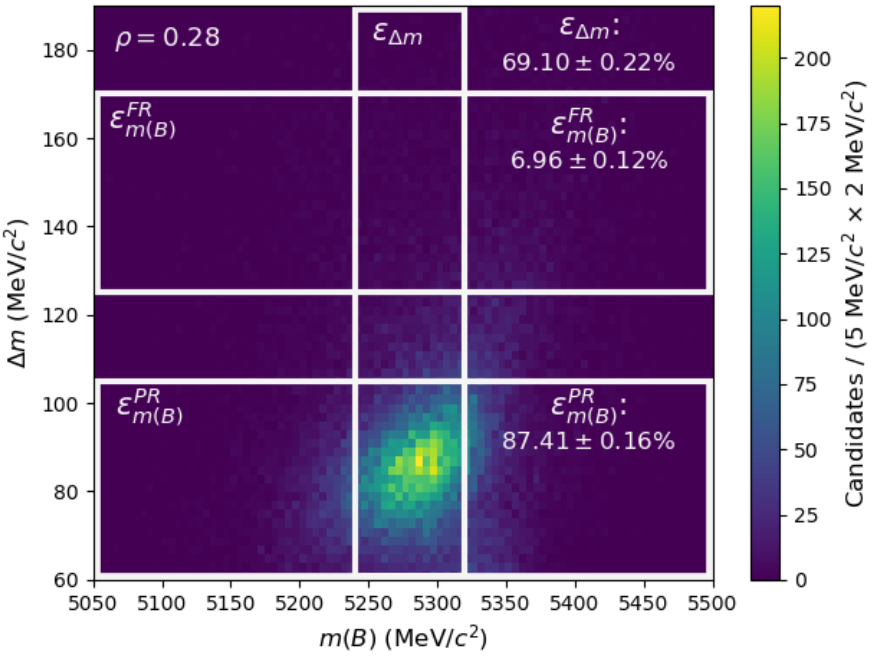


Figure 3.24: 2D distribution, $m(B)$ vs. Δm, of $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ signal MC reconstructed as $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$. The vertical rectangle encompasses the MC included in the Δm sample; the top (bottom) horizontal rectangle encompasses the MC included in the $m(B)$ sample selected around the $D\gamma$ ($D\pi^{0}$) peak. The box efficiencies are printed as a reference and match up to those list in Tab. 3.5. The Pearson correlation coefficients, ρ, of the 2D mass distributions are also given.

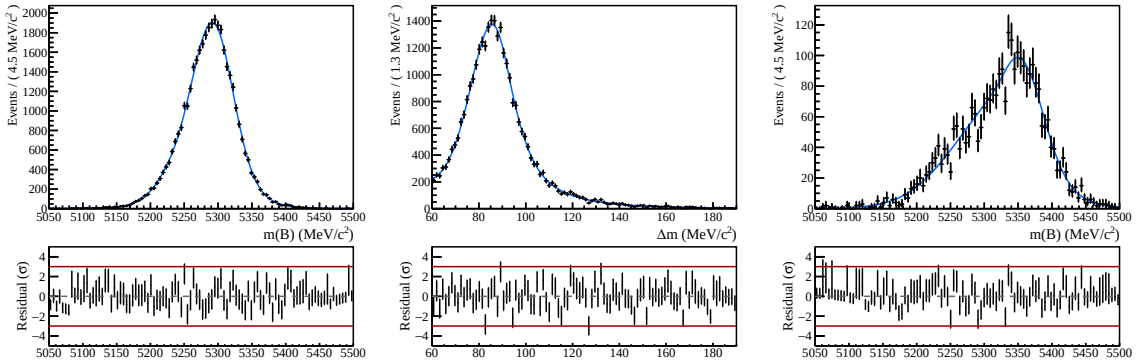


Figure 3.25: Simultaneous fit to $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ signal MC in Δm and two slices of $m(B)$: the partially reconstructed $D\pi^{0}$ slice on the left, and the $D\gamma$ slice on the right. 2011-2018 samples are used, where all years are considered together.

the tail parameter n has been fixed to 10. The mean and width in this sample are found to be $\mu(B)_{FR} \sim 5337\,\text{MeV}/c^2$ and $\sigma(B)_{FR} \sim 49\,\text{MeV}/c^2$.

$B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays reconstructed in the $D\gamma$ final state

To determine the signal PDFs for $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays reconstructed in the $D\gamma$ final state, a simultaneous fit across Δm and two slices of $m(B)$ is performed to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ MC. The 2D MC distribution depicted in Fig. 3.26 is considered in the fit.

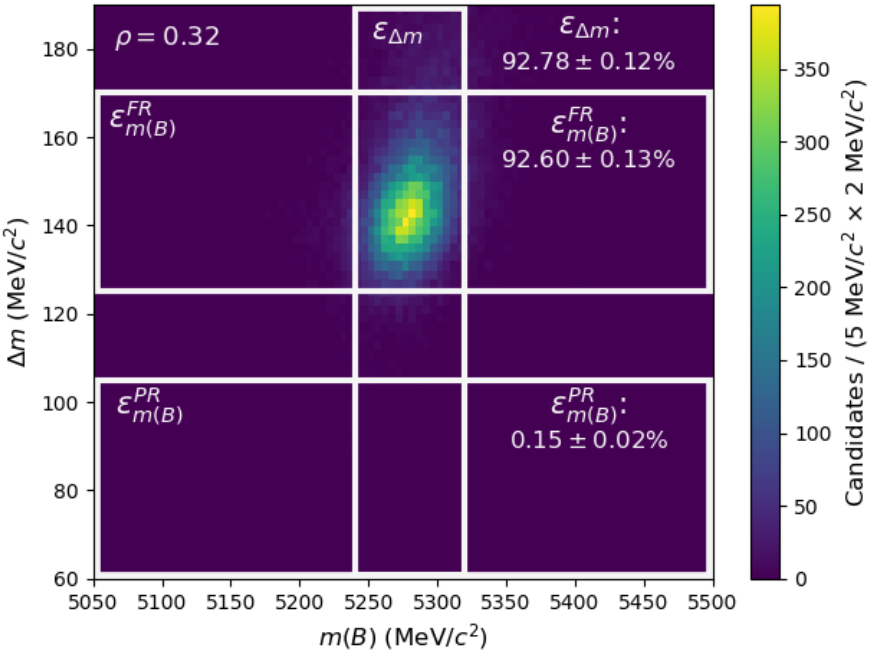


Figure 3.26: 2D distribution, $m(B)$ vs. Δm, of $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ signal MC reconstructed as $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$. The vertical rectangle encompasses the MC included in the Δm sample; the top (bottom) horizontal rectangle encompasses the MC included in the $m(B)$ sample selected around the $D\gamma$ ($D\pi^0$) peak. The box efficiencies are printed as a reference and match up to those list in Tab. 3.5. The Pearson correlation coefficients, ρ, of the 2D mass distributions are also given.

The simultaneous fit is shown in Fig. 3.27. The MC distributions in Δm and the FR slice of $m(B)$ are parameterised by the sum of two CBs. The n tail parameters are floated then fixed in the fit to aid stability. The values of α are allowed to vary freely, settling on small positive values for α_1 and small negative values for α_2 in both mass samples. The relative contribution of the two Crystal Ball functions, k, floats in the both fits. The means and widths vary freely but are shared across both Crystal Balls within each sample: $\mu_1 = \mu_2, \sigma_1 = \sigma_2$. In $m(B)$, $\mu(B)_{FR} \sim 5281\,\text{MeV}/c^2$

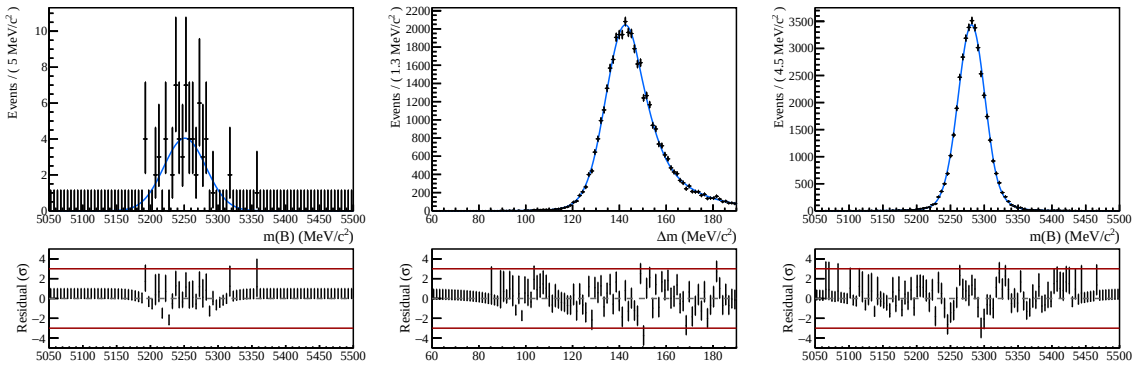


Figure 3.27: Simultaneous fit to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\gamma)\pi^{\pm}$ signal MC in Δm and two slices of $m(B)$: the partially reconstructed $D\pi^0$ slice on the left, and the $D\gamma$ slice on the right. 2011-2018 samples are used, where all years are considered together.

and $\sigma(B)_{FR} \sim 20\,\text{MeV}/c^2$; in Δm, $\mu(\Delta) \sim 143\,\text{MeV}/c^2$ and $\sigma(\Delta) \sim 8\,\text{MeV}/c^2$. The component reconstructed in the PR slice of $m(B)$ is small due to the low Δm region the partially reconstructed π^0 slice selects for ($< 105\,\text{MeV}/c^2$). The chosen PDF is a single CB function. The mean and width settle on the values of $\mu(B)_{PR} \sim 5250\,\text{MeV}/c^2$ and $\sigma(B)_{PR} \sim 31\,\text{MeV}/c^2$.

3.3.3 Mis-reconstructed backgrounds

Mis-reconstructed decays are formed when a charged or neutral pion from the true decay has been missed and a neutral particle from the underlying event added to the D meson to form a D^{*0} candidate. The mis-reconstructed decay modes considered in this work and their associated PDFs are summarised in Tab. 3.9, and are discussed in more detail below.

In order to parameterise the Δm distributions of decays involving neutral combinatorial, where fake D^* mesons are built, a PDF shape that models $m(D^*)-m(D)$ mass differences is used:

$$f_{D^*-D}\left(\Delta m; \Delta m_0, A, B, C\right) = \left(1 - e^{\left(-\frac{\Delta m - \Delta m_0}{C}\right)}\right) \times \left(\frac{\Delta m}{\Delta m_0}\right)^A + B \times \left(\frac{\Delta m}{\Delta m_0} - 1\right). \tag{3.32}$$

Here, Δm_0 is the mass difference threshold; the shape parameter A governs the power law rate with which the right hand tail increases/decreases; B also effects the right hand tail, but describes a linear rate of increase/decrease and $C > 0$

determines the dominance of the exponential decay function. Large correlations between the function variables exist therefore it is often necessary to fix $1-3$ parameters to obtain a stable fit.

Decay	Reconstructed as $D\pi^0$		Reconstructed as $D\gamma$		
	$m(B)$	Δm	$m(B)_{\rm PR}$	Δm	$m(B)_{\rm FR}$
Mis-reco. $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$	f_{CB}	f_{D^*-D}	$f_{DCB} + f_G$	f_{D^*-D}	f_{DCB}
Mis-reco. $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$	f_{CB}	f_{D^*-D}	f_G	f_{D^*-D}	$f_{DCB} + f_G$
$B^0 \to (D^{*\mp} \to D\pi^{\mp})h^{\pm}$	f_G	f_{D^*-D}	f_{DCB}	f_{D^*-D}	f_{DCB}
$B \to Dh^{\pm}\pi$	f_G	f_{D^*-D}	f_G	f_{D^*-D}	$f_{DCB} + f_G$
$\Lambda_b^0 \to (\Sigma_c^+ \to (\Lambda_c^+ \to pK^-\pi^+)\pi^0)\pi^-$	f_G	f_{D^*-D}	f_{CB}	f_{D^*-D}	f_{CB}
$B_s^0 \to D^0K^-\pi^+$	f_G	f_{D^*-D}	$f_{CB} + f_G$	f_{D^*-D}	f_{CB}

Table 3.9: Summary of PDFs used to parameterise mis-reconstructed decay modes, where f_{D^*-D} represents the $m(D^*) - m(D)$ mass difference function defined in Eq. (3.32), f_{CB} represents the Crystal Ball function defined in Eq. (3.30), f_{DCB} represents the double Crystal Ball function defined in Eq. (3.31) and f_G represents a Gaussian function.

Mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays

Mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays occur when the charged section of the decay is reconstructed correctly but a fake D^* meson is built by combining the D meson with a neutral particle from the underlying event. These are mis-reconstructed 'signal' decays, and therefore can be constrained in the invariant-mass fit to data by the peaking, correctly reconstructed, signal modes; this parameterisation will be described in more detail in Sec. 4.1.5.

Four categories of mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays exist. The PDFs used to parameterise the invariant-mass distributions for each are described below:

- $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays mis-reconstructed in the $D\pi^0$ final state, where the *wrong* π^0 meson has been used to build a fake D^* candidate. Decays of this kind are parameterised with a single CB shape in $m(B)$ and the $m(D^*) - m(D)$

mass difference function in Δm.

- $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays mis-reconstructed in the $D\gamma$ final state, where a photon from the underlying event has been used to build a fake D^* candidate. Decays of this kind are parameterised with a double CB combined with a Gaussian in the PR slice of $m(B)$, the $m(D^*) - m(D)$ mass difference function in Δm and a double CB in the FR slice of $m(B)$.

- $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays mis-reconstructed in the $D\pi^0$ final state, where a π^0 meson from the underlying event has been used to build a fake D^* candidate. Decays of this kind are parameterised with a single CB shape in $m(B)$ and the $m(D^*) - m(D)$ mass difference function in Δm.

- $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays mis-reconstructed in the $D\gamma$ final state, where the *wrong* photon has been used to build a fake D^* candidate. Decays of this kind are parameterised with a Gaussian function in the PR slice of $m(B)$, the $m(D^*) - m(D)$ mass difference function in Δm and a double CB and a Gaussian in the FR slice of $m(B)$.

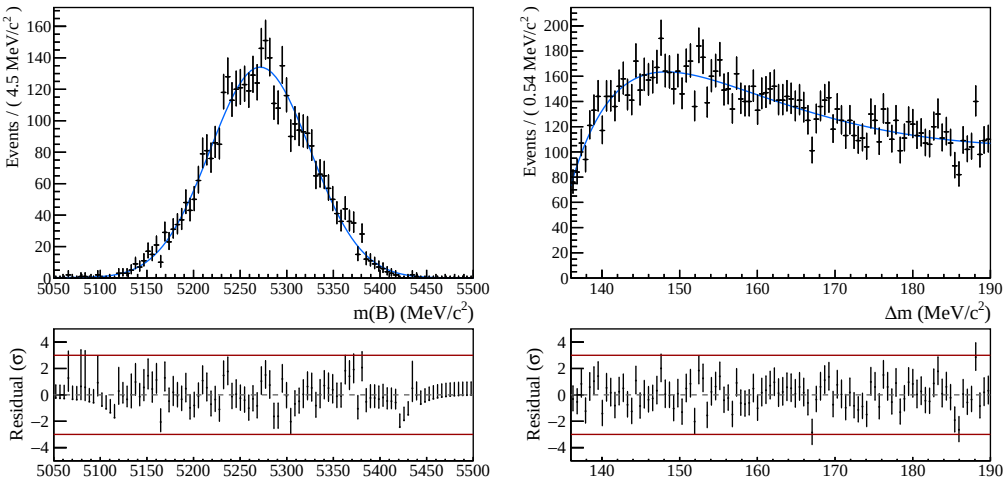


Figure 3.28: Simultaneous fit to mis-reconstructed $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ MC in $m(B)$ and Δm. 2011-2018 samples are used, where all years are considered together.

A summary of these PDFs is given in Tab. 3.9. The distributions of mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays are very similar for those that were generated as true $D^* \to D\gamma$

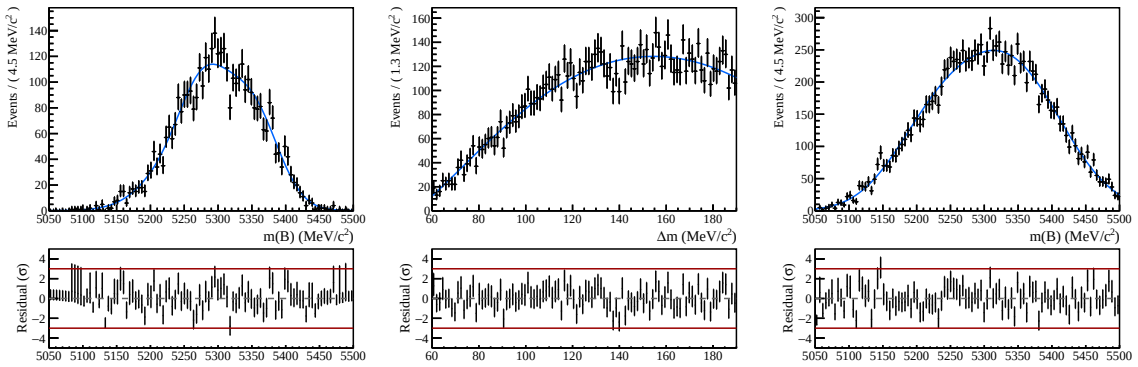


Figure 3.29: Simultaneous fit to mis-reconstructed $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ MC in Δm and two slices of $m(B)$: the partially reconstructed $D\pi^{0}$ slice on the left, and the $D\gamma$ slice on the right. 2011-2018 samples are used, where all years are considered together.

decays or as true $D^{*} \to D\pi^{0}$ decays. Fits to simulation when reconstructed in the $D\pi^{0}$ and $D\gamma$ final states are therefore only shown for $B^{\pm} \to (D^{*} \to D\pi^{0})\pi^{\pm}$ decays in Figs. 3.28 and 3.29, respectively. The shape parameters measured by these simultaneous fits enter as fixed terms in the invariant-mass fits to data, and are shared across all decay categories.

$B^{0} \to (D^{*\mp} \to D\pi^{\mp})h^{\pm}$ decays

Decays of the kind $B^{0} \to (D^{*\mp} \to D\pi^{\mp})h^{\pm}$ end up in the $D\pi^{0}$ ($D\gamma$) data tuples when the charged pion from the $D^{*\mp}$ is not reconstructed, and a neutral pion (photon) from the underlying event is combined with the D meson to make a fake D^{*} candidate.

In the $D\pi^{0}$ final state, the MC distribution in $m(B)$ is parameterised by a single Gaussian; the Δm sample is described using the $m(D^{*}) - m(D^{0})$ mass difference PDF. The simultaneous fit is shown in Fig. 3.30. For reconstruction in the $D\gamma$ final state, the MC distribution in Δm is parameterised using the $m(D^{*}) - m(D^{0})$ mass difference function; the FR and PR slices of $m(B)$ are parameterised by the sum of two Crystal Balls. The simultaneous fit is shown in Fig. 3.31.

$B \to Dh^{\pm}\pi$ decays

Decays of the type $B \to Dh^{\pm}\pi$ end up in the $D\pi^{0}$ ($D\gamma$) data tuples when the pion produced alongside the companion, h, is not reconstructed, but a neutral pion (photon)

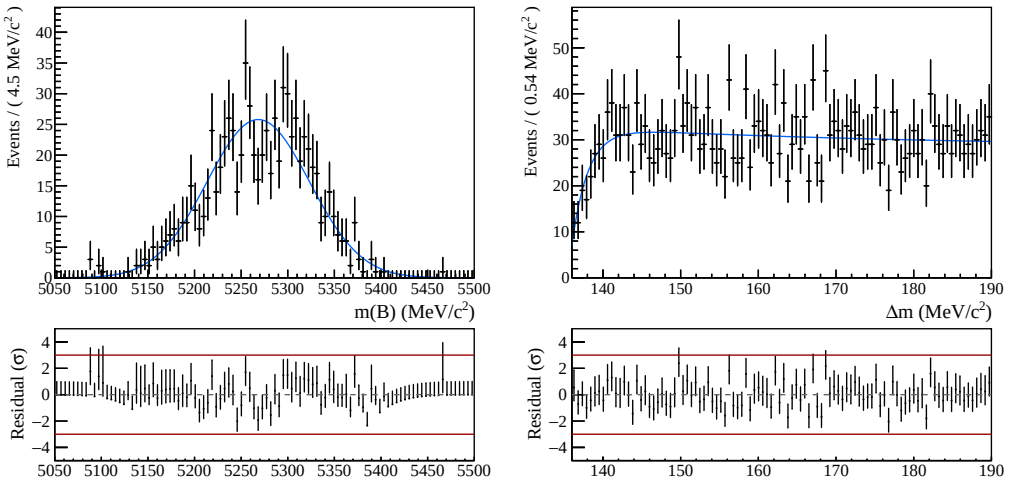


Figure 3.30: Simultaneous fit to $B^0 \to (D^{*\mp} \to [K^{\pm}\pi^{\mp}]_D\pi^{\pm})\pi^{\pm}$ MC in $m(B)$ and Δm. 2011-2018 samples are used, where all years are considered together.

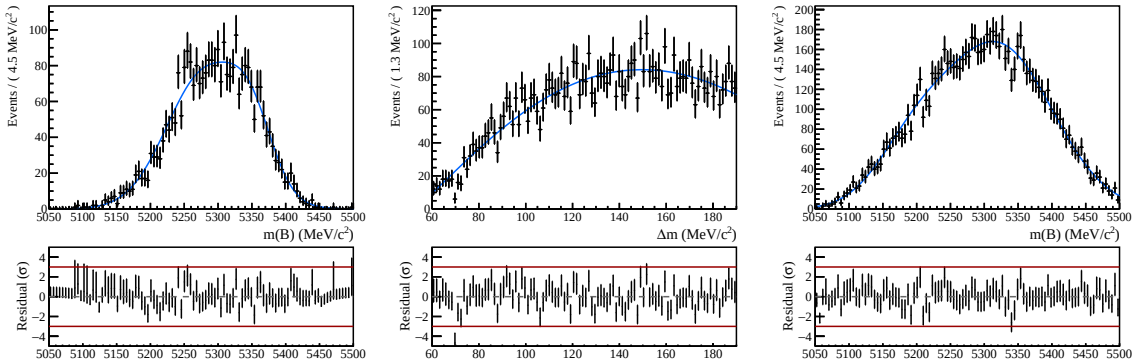


Figure 3.31: Simultaneous fit to $B^0 \to (D^{*\mp} \to [K^{\pm}\pi^{\mp}]_D\pi^{\pm})\pi^{\pm}$ MC in Δm and two slices of $m(B)$: the partially reconstructed $D\pi^0$ slice on the left, and the $D\gamma$ slice on the right. 2011-2018 samples are used, where all years are considered together.

from the underlying event is combined with the D meson to make a fake D^* candidate. In the $D^*\pi$ sub-sample, two types of decay fall into this category: $B^{\pm} \to D\pi^{\pm}\pi^0$ and $\overline{(}B^0\overline{)} \to D\pi^{\pm}\pi^{\mp}$ decays, whose branching fractions are:

$$\mathcal{B}(B^{\pm} \to D\pi^{\pm}\pi^0) = (1.34 \pm 0.18)\% \tag{3.33}$$

$$\mathcal{B}(\overline{(}B^0\overline{)} \to D\pi^{\pm}\pi^{\mp}) = (8.8 \pm 0.5) \times 10^{-4}. \tag{3.34}$$

Despite having a much smaller branching fraction, the latter features prominently in the ADS mode as Cabibbo-favoured $B^0 \to [K^+\pi^-]_D\pi^+\pi^-$ decays are disguised as suppressed decays when the π^- is reconstructed as the companion.

In the D^*K sub-sample, $B^{\pm} \to DK^{\pm}\pi^0$ and $\overline{(}B^0\overline{)} \to DK^{\pm}\pi^{\mp}$ decays contribute,

where the latter have a larger value of r_B as Cabibbo suppressed as $B^0 \to D^0 K^+\pi^-$ decays, involving a V_{ub}^* matrix element, are not additionally colour suppressed with respect to favoured $B^0 \to \overline{D}^0 K^+\pi^-$ decays; from the latest LHCb γ combination, $r_{B^\pm}^{DK^{*\pm}} = 0.106^{+0.087}_{-0.122}$ and $r_{B^0}^{DK^{*0}} = 0.25^{+0.23}_{-0.27}$ [23]. These channels are therefore more prominent in the ADS and GLW modes despite having a branching fraction $\sim 17\%$ the size of the former:

$$\mathcal{B}(B^{\pm} \to DK^{\pm}\pi^{0}) = (5.3 \pm 0.4) \times 10^{-4} \tag{3.35}$$

$$\mathcal{B}(\overline{B^{0}} \to DK^{\pm}\pi^{\mp}) = (0.88 \pm 0.17) \times 10^{-4}. \tag{3.36}$$

Full simulation MC samples of $B^{\pm} \to D(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ decays, re-weighted to describe the full $\pi^{\pm}\pi^{0}$ phase space, are used to model all $B \to Dh^{\pm}\pi$ decay modes. The re-weighting method is as follows:

- The normalised $m(D\pi)$ distribution of $B^0 \to D\pi^{\pm}\pi^{\mp}$ samples, generated using the full LHCb amplitude model [87], is divided by that of full simulation $B^{\pm} \to D(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ decays.
- Bin-by-bin ratios in the $m(D\pi)$ variable are obtained.
- These are then used to re-weight the $m(D^{*}\pi)$ and Δm distributions of $B^{\pm} \to D(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ decays by matching the $m(D\pi)$ value of each event to the associated weight.

The samples used to obtain the weights were generated for the partially reconstructed $B^{\pm} \to D^{(*)}h^{\pm}$ analysis [31]. The $\rho^{\pm}$ resonance dominates in the region of phase space considered in this analysis, therefore the re-weighting procedure has a minor effect.

In the $D\pi^0$ final state, the MC distribution of re-weighted $B^{\pm} \to D(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ decays in $m(B)$ is parameterised by a single Crystal Ball; the Δm sample is described using the $m(D^{*}) - m(D^{0})$ mass difference PDF. The simultaneous fit is shown in Fig. 3.32. For reconstruction in the $D\gamma$ final state, the MC distribution in the PR slice of $m(B)$ is described using a Gaussian function; the Δm sample is parameterised using the $m(D^{*}) - m(D^{0})$ mass difference function; the FR slice of

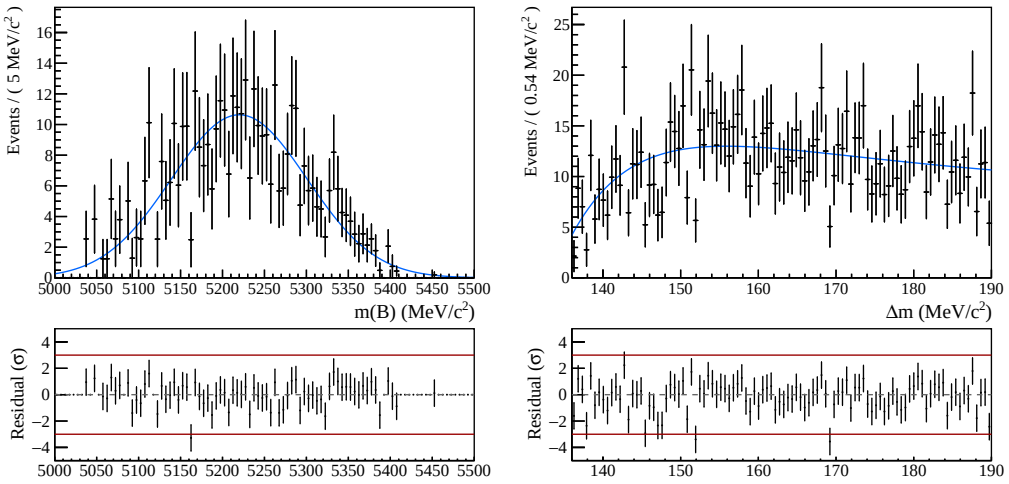


Figure 3.32: Simultaneous fit to $B^{\pm} \to D^{0}(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ MC in $m(B)$ and Δm. 2011-2018 samples are used, where all years are considered together.

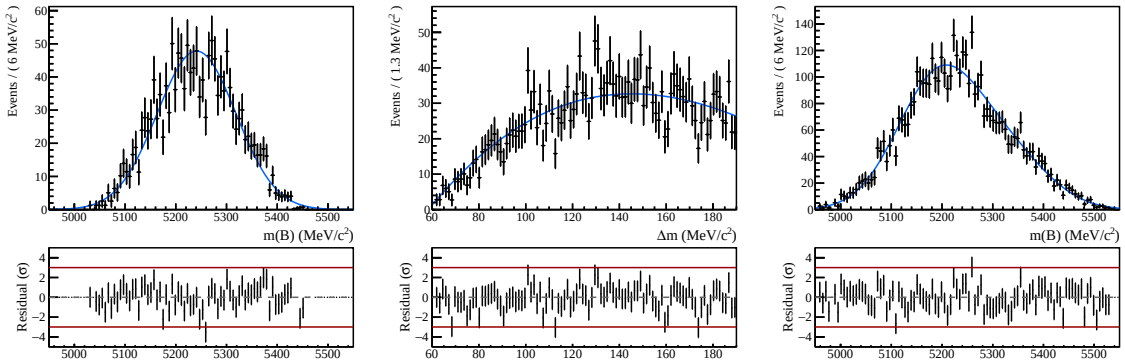


Figure 3.33: Simultaneous fit to $B^{\pm} \to D^{0}(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ MC in Δm and two slices of $m(B)$: the partially reconstructed $D\pi^{0}$ slice on the left, and the $D\gamma$ slice on the right. 2011-2018 samples are used, where all years are considered together.

$m(B)$ is parameterised by the sum of two Crystal Balls plus a Gaussian function. The simultaneous fit is shown in Fig. 3.33.

The shape parameters measured by these simultaneous fits enter as fixed terms in the invariant-mass fits to data, and are shared across all decay categories apart from the $D^{*}\pi$ ADS modes. In this category, freedom is given in both data fits to allow for any phase-space differences due to the dominance of mis-reconstructed $\overline{B}^{0} \to D\pi^{\pm}\pi^{\mp}$ decays; the means and widths of the resonances in B mass and the A parameters describing the mass difference function curvature in Δm are floated then fixed in the data fits to their final values $\pm$ uncertainties.

$\Lambda_b^0 \to \Sigma_c^+ h^-$ decays

$\Lambda_b^0 \to (\Sigma_c^+ \to (\Lambda_c^+ \to pK^-\pi^+)\pi^0)h^-$ decays end up in the $D\pi^0$ ($D\gamma$) data tuples when the proton is mis-identified as a kaon, the charged pion is not reconstructed, and a neutral pion (photon) from the underlying event is combined with the mis-reconstructed K^+K^- pair to make a D^* candidate. These decays therefore form a prominent background in the $D \to KK$ category.

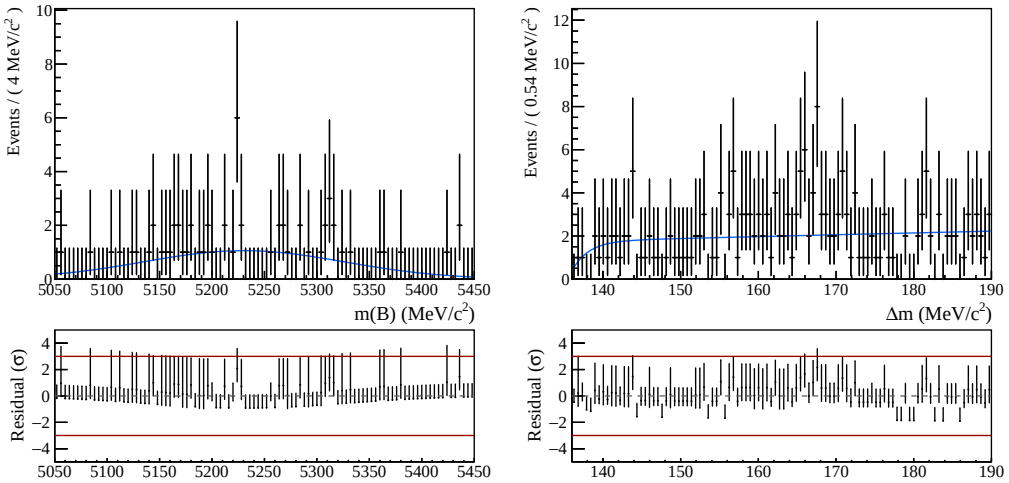


Figure 3.34: Simultaneous fit to $\Lambda_b^0 \to (\Sigma_c^+ \to (\Lambda_c^+ \to pK^-\pi^+)\pi^0)\pi^-$ MC in $m(B)$ and Δm. 2012-2016 samples are used, where all years are considered together.

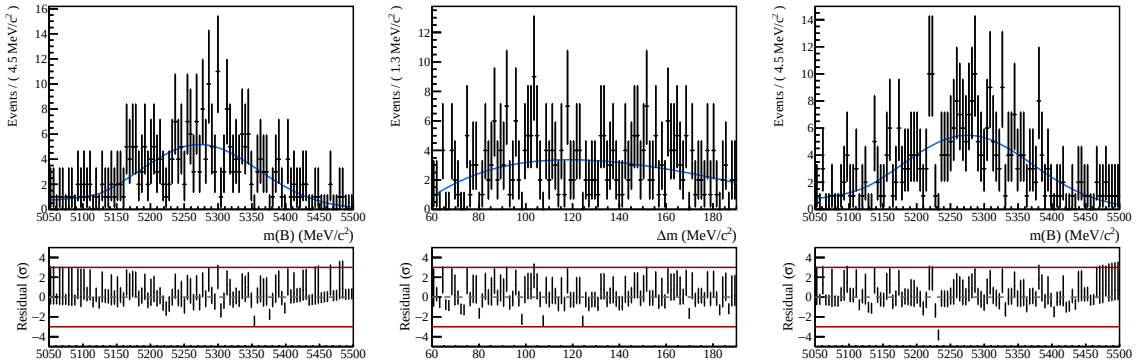


Figure 3.35: Simultaneous fit to $\Lambda_b^0 \to (\Sigma_c^+ \to (\Lambda_c^+ \to pK^-\pi^+)\pi^0)\pi^-$ MC in Δm and two slices of $m(B)$: the partially reconstructed $D\pi^0$ slice on the left, and the $D\gamma$ slice on the right. 2011-2016 samples are used, where all years are considered together.

In the $D\pi^0$ final state, the MC distribution in $m(B)$ is parameterised by a Gaussian function and the Δm sample is described using the $m(D^*) - m(D^0)$ mass difference

PDF; the simultaneous fit is shown in Fig. 3.34. For reconstruction in the $D\gamma$ final state, the MC distribution in Δm is parameterised using the $m(D^*) - m(D^0)$ mass difference function and both slices of $m(B)$ are parameterised using a single Crystal Ball; the simultaneous fit is shown in Fig. 3.35. The shape parameters measured by these simultaneous fits enter as fixed terms in the invariant-mass fits to data, and are shared across both $D^*\pi$ and D^*K sub-samples.

$B_s^0 \to DK^-\pi^+$ decays

$B_s^0 \to DK^-\pi^+$ decays form a prominent background in the D^*K ADS category because Cabibbo-favoured $D^0 \to K^-\pi^+$ decays are produced with an oppositely-charged kaon companion, giving them the same signature as suppressed mode signal candidates. These decays are also considered in the D^*K GLW modes, with reduced contributions in line with to the lower D branching fractions; the contribution to the favoured mode is considered negligible. This background is not included in the $D^*\pi$ sub-sample as the reduction in B mass when the companion kaon is not reconstructed pushes the contribution below the invariant mass range considered.

Full simulation MC samples of $B_s^0 \to D(K^{*0}(892) \to K^-\pi^+)$ decays, re-weighted to describe the full $K^-\pi^+$ phase space, are used to determine the PDFs for this contribution. The re-weighting procedure is the same as for $B^{\pm} \to D\rho^{\pm}$ decays, described earlier in the section:

- The normalised $m(DK)$ distribution of $B_s^0 \to DK^-\pi^+$ samples, generated using the full LHCb amplitude model [88], is divided by that of full simulation $B_s^0 \to D(K^{*0}(892) \to K^-\pi^+)$ decays.
- Bin-by-bin ratios in the $m(DK)$ variable are obtained.
- These are then used to re-weight the $m(D^*K)$ and Δm distributions of $B_s^0 \to D(K^{*0}(892) \to K^-\pi^+)$ decays by matching the $m(DK)$ value of each event to the associated weight.

The samples used to obtain the weights were generated for the partially reconstructed

$B^{\pm} \to D^{(*)}h^{\pm}$ analysis [31]. The $K^{*0}(892)$ resonance dominates in the region of phase space considered in this analysis, therefore this process has a minor effect.

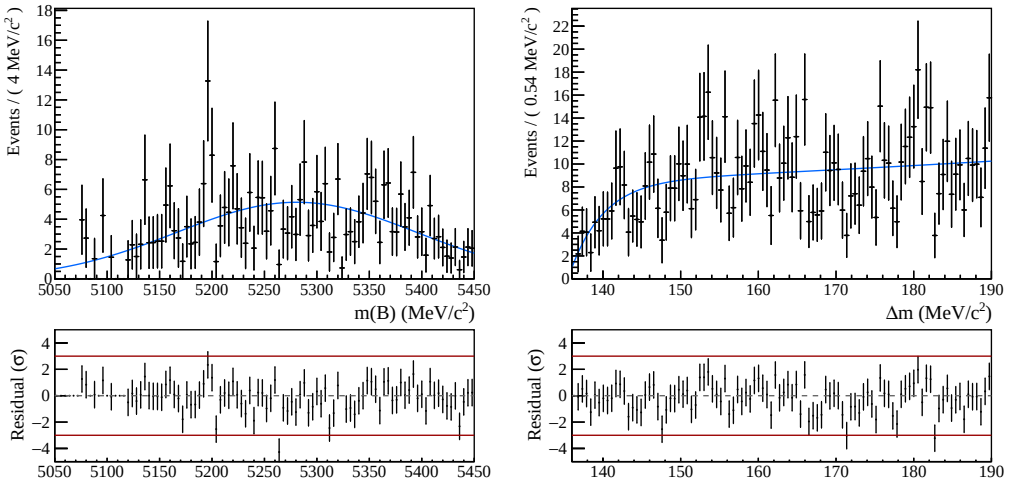


Figure 3.36: Simultaneous fit to weighted $B_s^0 \to D(K^{*0} \to K^-\pi^+)$ MC in $m(B)$ and Δm. 2011-2018 samples are used, where all years are considered together.

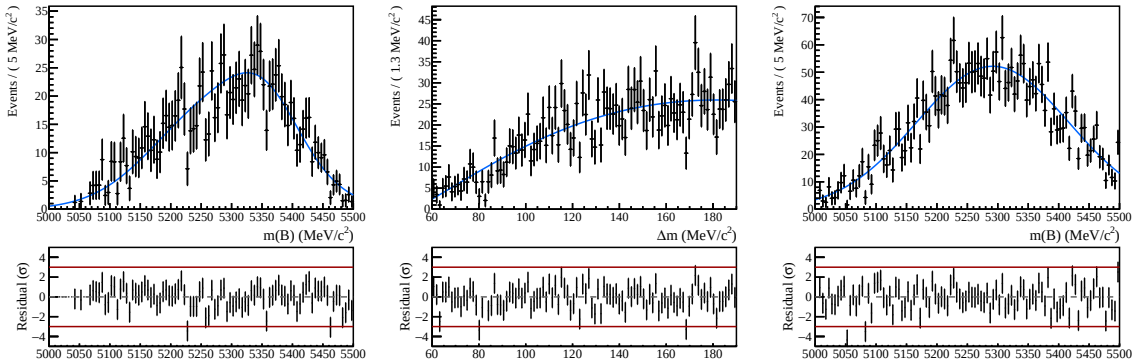


Figure 3.37: Simultaneous fit to weighted $B_s^0 \to D(K^{*0} \to K^-\pi^+)$ MC in Δm and two slices of $m(B)$: the partially reconstructed $D\pi^0$ slice on the left, and the $D\gamma$ slice on the right. 2011-2018 samples are used, where all years are considered together.

In the $D\pi^0$ final state, the MC distribution of re-weighted $B_s^0 \to D(K^{*0}(892) \to K^-\pi^+)$ decays in $m(B)$ is parameterised by a Gaussian function; the Δm sample is described using the $m(D^*) - m(D)$ mass difference PDF. The simultaneous fit is shown in Fig. 3.36. For reconstruction in the $D\gamma$ final state, the MC distribution in the PR slice of $m(B)$ is described using the sum of a Crystal Ball and a Gaussian function; the Δm sample is parameterised using the $m(D^*) - m(D)$ mass difference function; the FR slice of $m(B)$ is parameterised by a Crystal Ball. The simultaneous

fit is shown in Fig. 3.37. The shape parameters measured on simulation enter as fixed terms in the invariant-mass fits to data, and are shared across the GLW and ADS modes in the D^*K sub-sample.

3.3.4 Partially reconstructed backgrounds

Partially reconstructed decays contain a true D^{*0} or charged $D^{*\mp}$ candidate. A charged or neutral pion, either the decay product of the D^* or an associated companion particle, has been missed in the final state. The decays considered in this category and their associated PDFs are summarised in Tab. 3.10; more details are given below.

Decay	Reconstructed as $D\pi^0$		Reconstructed as $D\gamma$		
	$m(B)$	Δm	$m(B)_{\mathrm{PR}}$	Δm	$m(B)_{\mathrm{FR}}$
$B \to (D^* \to D\pi^0)h^{\pm}\pi$	f_G	f_G	f_G	f_G	f_G
$B \to (D^* \to D\gamma)h^{\pm}\pi$	-	-	f_G	f_G	f_{DCB}
$B_s^0 \to (D^{*0} \to D\pi^0)K^-\pi^+$	$f_{CB} + f_G$	f_{DCB}	f_{CB}	f_{DCB}	-
$B_s^0 \to (D^{*0} \to D\gamma)K^-\pi^+$	-	-	-	f_{D^*-D}	f_{DCB}
Mis-reco. $B_s^0 \to (D^{*0} \to D\pi^0)K^-\pi^+$	f_G	f_{DCB}	f_G	f_{D^*-D}	f_G
Mis-reco. $B_s^0 \to (D^{*0} \to D\gamma)K^-\pi^+$	f_G	f_{DCB}	f_G	f_{D^*-D}	f_G

Table 3.10: Summary of PDFs used to parameterise partially-reconstructed decay modes, where f_{D^*-D} represents the $m(D^*) - m(D)$ mass difference function defined in Eq. (3.32), f_{CB} represents the Crystal Ball function defined in Eq. (3.30), f_{DCB} represents the double Crystal Ball function defined in Eq. (3.31) and f_G represents a Gaussian function.

$B \to D^* h^{\pm}\pi$ decays

Decays of the kind $B \to D^*h^{\pm}\pi$ end up in the $D\pi^0$ ($D\gamma$) data tuples when the pion produced along with the companion hadron, h is not reconstructed. If the decay contains a neutral D^* meson, the D^* candidate is either fully reconstructed or the neutral particle is missed, in which case a neutral pion (photon) from the underlying event is combined with the D to make a fake D^* candidate. If the decay

contains a $D^{*\mp}$ meson, the charged pion in the $D^{*\mp}$ decay is not reconstructed and a neutral pion (photon) from the underlying event is combined with the D meson to make a fake D^{*} candidate.

There are several decays that contribute to the low-mass $B \to D^{*}h^{\pm}\pi$ component: $B^{\pm} \to (D^{*} \to D\pi^{0})h^{\pm}\pi^{0}$, $B^{\pm} \to (D^{*} \to D\gamma)h^{\pm}\pi^{0}$, $\overline{(}B^{0}\overline{)} \to (D^{*\mp} \to D\pi^{\mp})h^{\pm}\pi^{0}$, $\overline{(}B^{0}\overline{)} \to (D^{*} \to D\pi^{0})h^{\pm}\pi^{\mp}$ and $\overline{(}B^{0}\overline{)} \to (D^{*} \to D\gamma)h^{\pm}\pi^{\mp}$ decays. The corresponding branching fractions for decays containing the $\rho(770)$ and $K^{*}(892)$ resonances (which dominate the phase space considered) are given below:

$$\mathcal{B}(B^{\pm} \to D^{*0}\rho^{\pm}) = (9.8 \pm 1.7) \times 10^{-3} \tag{3.37}$$

$$\mathcal{B}(\overline{(}B^{0}\overline{)} \to D^{*\mp}\rho^{\pm}) = (6.8 \pm 0.9) \times 10^{-3} \tag{3.38}$$

$$\mathcal{B}(\overline{(}B^{0}\overline{)} \to D^{*0}\rho^{0}) < 5.1 \times 10^{-4} \tag{3.39}$$

$$\mathcal{B}(B^{\pm} \to D^{*0}K^{*\pm}) = (8.1 \pm 1.4) \times 10^{-4} \tag{3.40}$$

$$\mathcal{B}(\overline{(}B^{0}\overline{)} \to D^{*\mp}K^{*\pm}) = (3.3 \pm 0.6) \times 10^{-4} \tag{3.41}$$

$$\mathcal{B}(\overline{(}B^{0}\overline{)} \to D^{*0}K^{*0}) < 6.9 \times 10^{-5}, \tag{3.42}$$

where the hadronic parameters for these decay modes have not yet been measured. Neutral D^{*0} mesons decay to the via $D\pi^{0} \sim 65\%$ of the time and via $D\gamma \sim 35\%$ of the time; the branching fraction for charged $D^{*\pm}$ decays is of a similar size:

$$\mathcal{B}(D^{*\mp} \to D^{0}\pi^{\mp}) = (67.7 \pm 0.5) \times 10^{-2}. \tag{3.43}$$

As for the mis-reconstructed $B \to Dh^{\pm}\pi$ component, decays originating from a B^{0} meson dominate in the $D^{*}\pi$ ADS mode when Cabibbo-favoured $B^{0} \to (D^{*\pm} \to [K^{+}\pi^{-}]_{D}\pi^{\pm})\pi^{+}\pi^{-}$ decays are reconstructed with the π^{-} as the companion.

To obtain PDFs for this component, decays where the D^{*} meson decays to a D meson and a charged or neutral pion are modelled using $B^{\pm} \to (D^{*} \to D\pi^{0})(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ MC; decays where the D^{*} meson decays to a D meson and a photon are modelled using $B^{\pm} \to (D^{*} \to D\gamma)(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ MC. The PDF parameters measured from fits to MC (documented below) are shared across all decay modes in the invariant-mass fits to data, except in the $D^{*}\pi$ ADS mode. In this sub-sample, freedom is given to the

means and widths of the distributions to accommodate any phase-space differences from the dominance of mis-reconstructed favoured mode B^0 decays.

For events reconstructed in the $D\pi^0$ final state, the selection efficiency of $B^{\pm} \to (D^{*} \to D\gamma)\rho^{\pm}$ decays is ~ 10 times smaller than that of $B^{\pm} \to (D^{*} \to D\pi^{0})\rho^{\pm}$ decays, therefore the $B \to D^{*}h^{\pm}\pi$ component is modelled entirely on the latter. Gaussian PDFs parameterise both the $m(B)$ and Δm distributions for $B^{\pm} \to (D^{*} \to D\pi^{0})\rho^{\pm}$ decays, as depicted in Fig. 3.38. There is no peaking component in the Δm distribution due to the box cut on the $m(B)$ variable, as can be seen in the 2D distribution of Fig. 3.40; the selection window cuts out the resonance from true D^{*0} candidates.

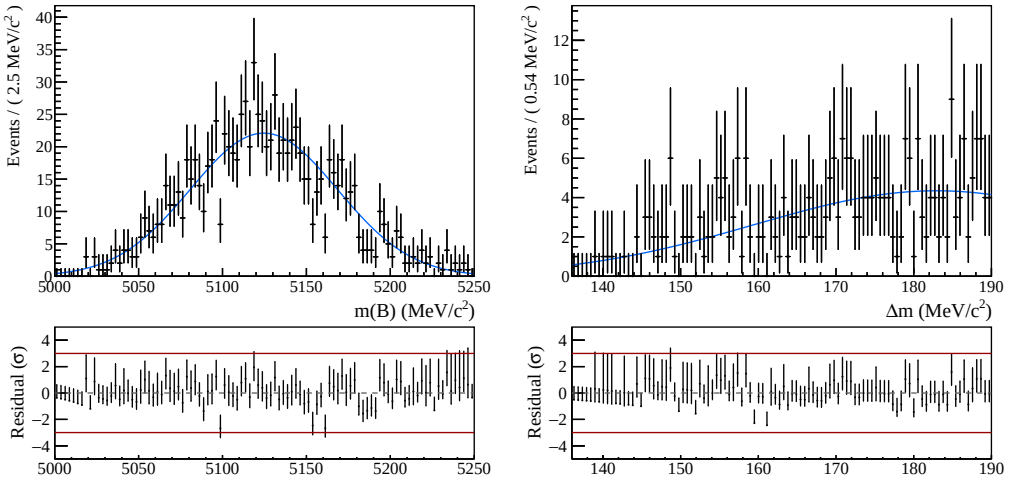


Figure 3.38: Simultaneous fit to $B^{\pm} \to (D^{*} \to D\pi^{0})(\rho^{\pm} \to \pi^{\pm}\pi^{0})$ MC in $m(B)$ and Δm. 2011-2018 samples are used, where all years are considered together.

For events reconstructed in the $D\gamma$ final state, two PDFs are modelled: decays in which the D^* meson decays to a D meson and a pion, and decays in which the D^* meson decays to a D meson and a photon. For the former component, the MC distributions of $B^{\pm} \to (D^{*} \to D\pi^{0})\rho^{\pm}$ decays in Δm and both slices of $m(B)$ are parameterised by Gaussian functions. For the latter component, the MC distributions of $B^{\pm} \to (D^{*} \to D\gamma)\rho^{\pm}$ decays in Δm and the PR slice of $m(B)$ are parameterised by Gaussian functions; the FR slice of $m(B)$ is modelled using the sum of two Crystal Balls. Again, Fig. 3.40 illustrates how the box cuts remove any peaking component due to true D^* mesons from the Δm samples. The simultaneous fits are shown in

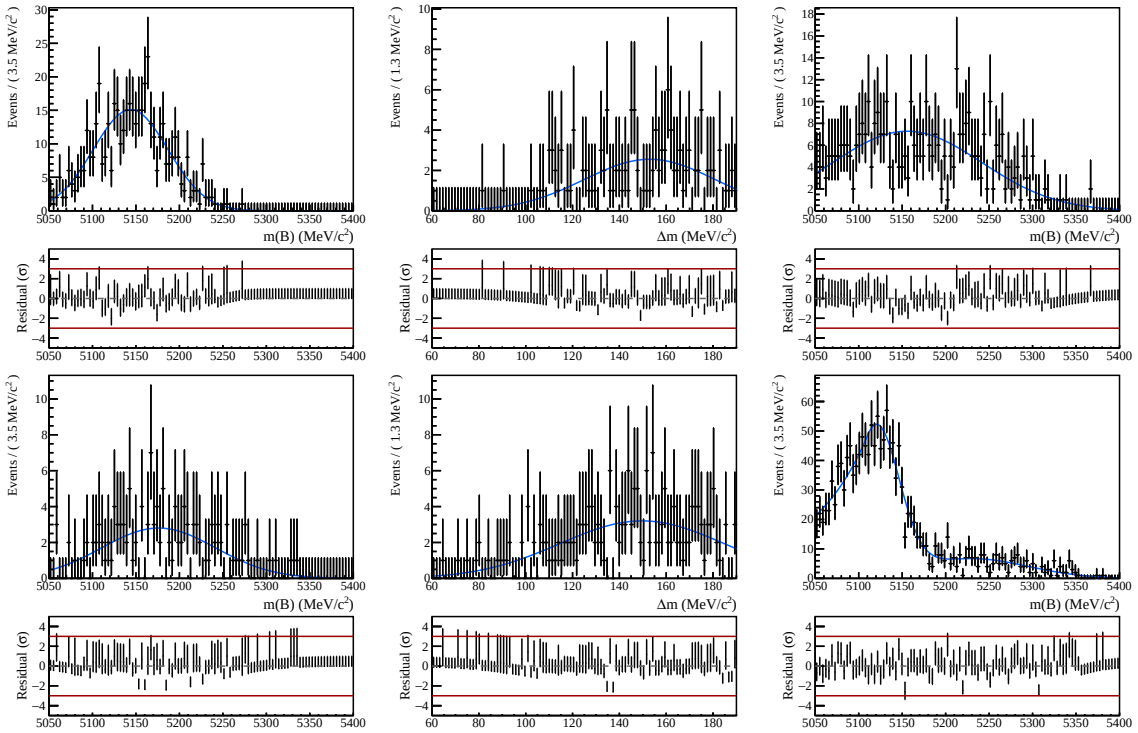


Figure 3.39: Simultaneous fits to $B^{\pm} \to (D^* \to D\pi^0)(\rho^{\pm} \to \pi^{\pm}\pi^0)$ MC at the top and $B^{\pm} \to (D^* \to D\gamma)(\rho^{\pm} \to \pi^{\pm}\pi^0)$ MC at the bottom, using the Δm distributions and two slices of $m(B)$: the partially reconstructed $D\pi^0$ slices are depicted on the left, and the $D\gamma$ slices are on the right. 2011-2018 samples are used, where all years are considered together.

Fig. 3.39. In the fit data, these two components are combined into a single PDF by fixing the fraction of each using the relative selection efficiencies, given in Tab. 3.7, and the branching fractions of the contributing decay modes.

$B_s^0 \to D^{*0}K^-\pi^+$ decays

$B_s^0 \to D^{*0}K^-\pi^+$ decays form a prominent background in the D^*K ADS category because Cabibbo-favoured $D^0 \to K^-\pi^+$ decays are produced with an oppositely-charged kaon companion, giving them the same signature as suppressed mode signal candidates. These decays are also considered in the D^*K GLW modes, with reduced contributions in line with the lower D branching fractions; the contribution to the favoured mode is considered negligible. This background is not included in the $D^*\pi$ sub-sample as the reduction in B mass when the companion kaon is not reconstructed pushes the contribution below the invariant mass range considered.

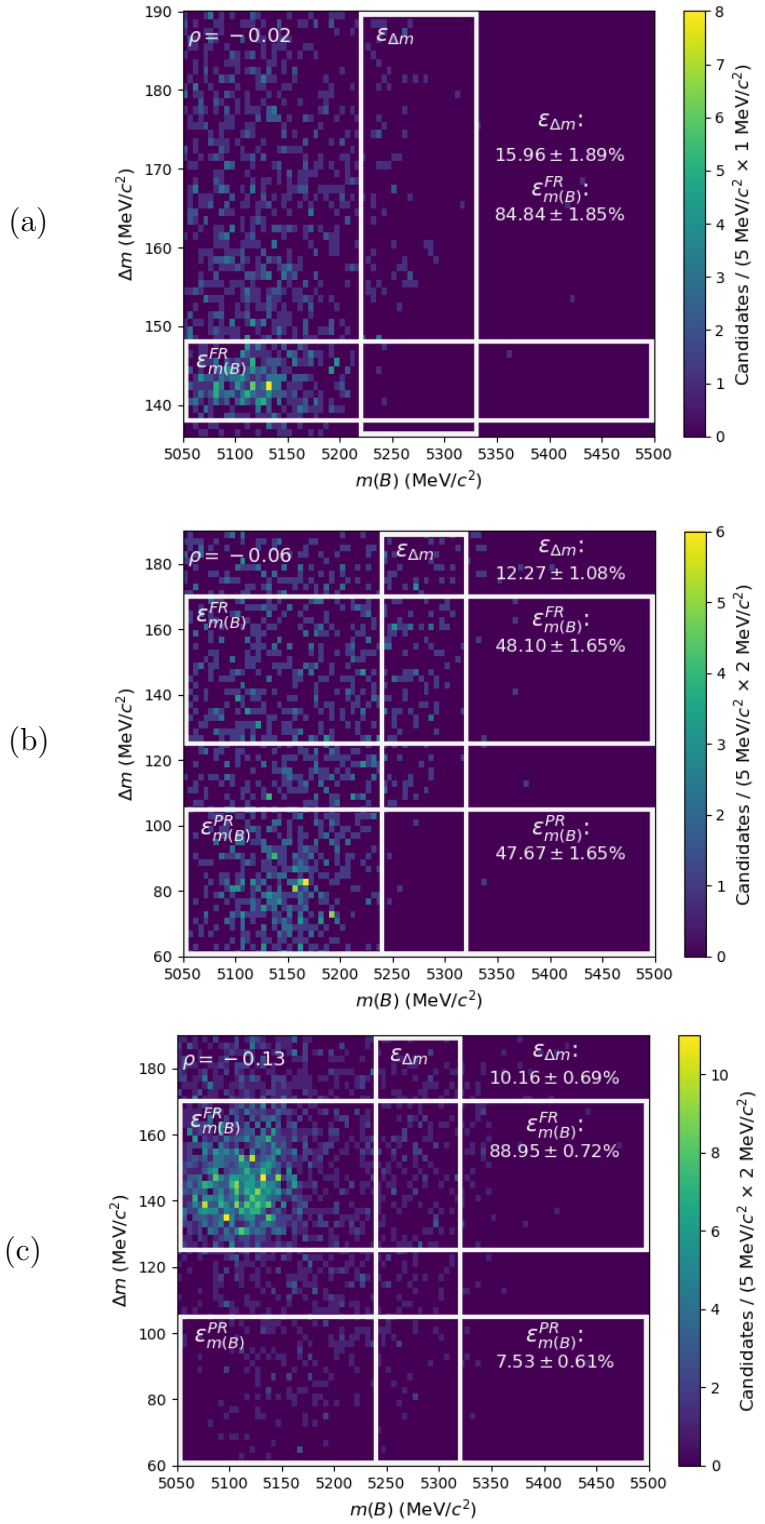


Figure 3.40: 2D distributions, $m(B)$ vs. Δm, of $B^{\pm} \to D^{*}(\rho^{\pm} \to \pi^{\pm}\pi^{0})$: (*a*) displays generated $D^{*} \to D\pi^{0}$ decays reconstructed in the $D\pi^{0}$ final state; (*b*) displays generated $D^{*} \to D\pi^{0}$ decays reconstructed in the $D\gamma$ final state; (*c*) displays generated $D^{*} \to D\gamma$ decays reconstructed in the $D\gamma$ final state. The vertical (horizontal) rectangles encompass the MC included in the Δm ($m(B)$) samples, where the box efficiencies are printed as a reference. The Pearson correlation coefficients, ρ, of the 2D mass distributions are also given.

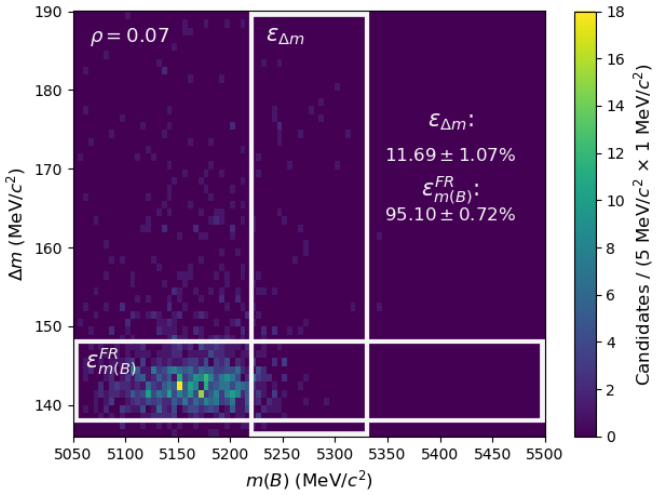
$B_s^0 \to D^{*0}K^-\pi^+$ decays are split into the five separate D^* reconstruction categories listed below:

1. Fully reconstructed $B_s^0 \to (D^{*0} \to D\gamma)K^-\pi^+$ decays.
2. Fully reconstructed $B_s^0 \to (D^{*0} \to D\pi^0)K^-\pi^+$ decays.
3. Partially reconstructed $B_s^0 \to (D^{*0} \to D\pi^0)K^-\pi^+$ decays, where only one photon from the π^0 decay has been detected.
4. Mis-reconstructed $B_s^0 \to (D^{*0} \to D\gamma)K^-\pi^+$ decays, where the D meson has been combined with a neutral particle from the underlying event.
5. Mis-reconstructed $B_s^0 \to (D^{*0} \to D\pi^0)K^-\pi^+$ decays, where the D meson has been combined with a neutral particle from the underlying event.

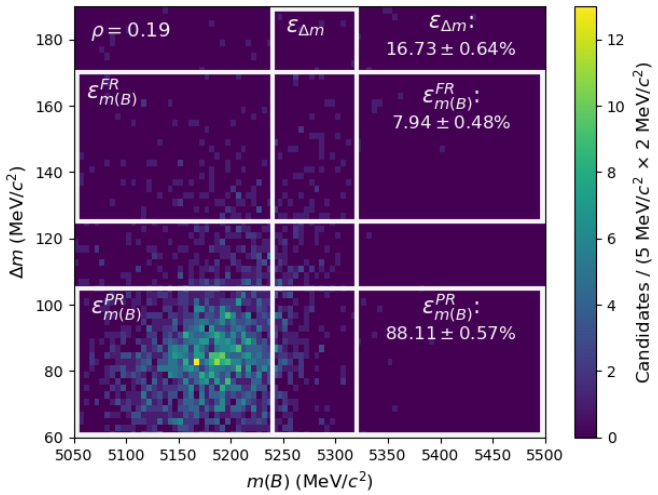
Categories 2, 4 & 5 are considered in the invariant-mass fit to data reconstructed in the $D\pi^0$ final state. Categories 1, 3, 4 & 5 are considered in the invariant-mass fit to data reconstructed in the $D\gamma$ final state. In this way, modelling doesn't rely on simulation samples having the correct proportion of mis- and correctly reconstructed D^* mesons, and the B_s^0 decay modes are handled identically to fully, partially and mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays (*i.e.* signal).

The majority of true D^* mesons in categories 1, 2, & 3 are removed from the Δm sample due to the $m(B)$ mass window cut, as for $B \to D^*h^{\pm}\pi$ decays. However, due to the higher B_s^0 mass, $(5366.88 \pm 0.14)\,\mathrm{MeV}/c^2$, some true D^* candidates do end up in the Δm sample, as can be seen in the 2D distributions of Fig. 3.41. Therefore, unlike partially reconstructed $B \to D^*h^{\pm}\pi$ decays, $B_s^0 \to D^{*0}K^-\pi^+$ decays have peaking components in Δm. The shape of each distribution in this mass sample is determined by D^* decay kinematics, which are shared between the corresponding categories of signal and B_s^0 decays. The Δm PDF for each D^* decay category is therefore shared with the corresponding $B^{\pm} \to D^*h^{\pm}$ component, as these MC samples have significantly higher statistics. The power of the D1D method is evident here, as these peaking B_s^0 components sit dangerously under the signal peaks in

(a)



(b)



(c)

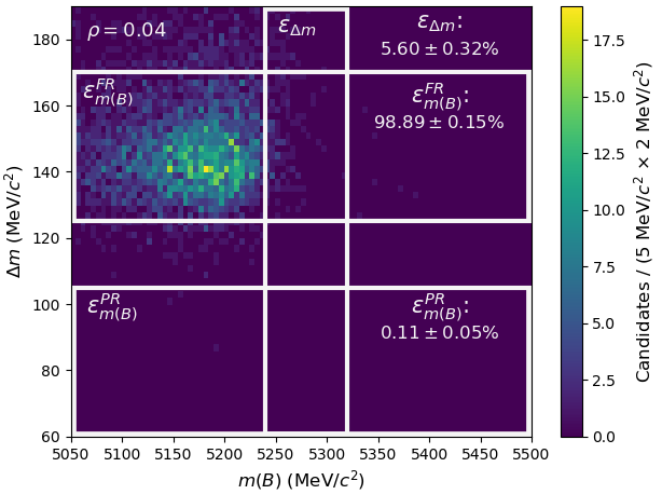


Figure 3.41: 2D distributions, $m(B)$ vs. Δm, of $B^0_s \to D^* K^- \pi^+$ simulation: (a) displays generated $D^* \to D\pi^0$ decays reconstructed in the $D\pi^0$ final state; (b) displays generated $D^* \to D\pi^0$ decays reconstructed in the $D\gamma$ final state; (c) displays generated $D^* \to D\gamma$ decays reconstructed in the $D\gamma$ final state. The vertical (horizontal) rectangles encompass the MC included in the Δm ($m(B)$) samples, where the box efficiencies are printed as a reference. The Pearson correlation coefficients, ρ, of the 2D mass distributions are also given.

Δm. Their yields in the invariant-mass fits to data, however, are constrained by the corresponding B mass distributions, which are displaced to low mass due to their partially reconstructed nature.

Simulation samples are needed to model the B mass distributions of $B_s^0 \to D^* K^- \pi^+$ decays. Taking the ratio of resonance contributions to the $K\pi$ phase space from the amplitude analysis of $B_s^0 \to D^0 K^- \pi^+$ decays [88], 60% of the phase space is occupied by the $K^{*0}(892)$ resonance and 40% by the $K^{*0}(1430)$ resonance. The phase space considered in this analysis sits in the far-right corner of the Dalitz plane, excluding $\sim 50\%$ of the $K^{*0}(1430)$. The $B_s^0 \to D^{*0} K^- \pi^+$ samples used to model the B mass distributions for categories 1, 2 & 3 are therefore hybrid simulation samples, consisting of $(80 \pm 10)\%$ $B_s^0 \to D^{*0}(K^{*0}(892) \to K^- \pi^+)$ decays and $(20 \pm 10)\%$ $B_s^0 \to D^{*0} K^- \pi^+$ decays generated over a flat, square Dalitz space. The $K^{*0}(1430)$ resonance forms a broad peak, with a natural width of $270\,\mathrm{MeV}/c^2$, so its contribution to the distribution in the far corner of the Dalitz plot is approximated with the phase space generation. The box efficiencies given in Tab. 3.5, used to partition the total component yield between the simultaneous mass samples, are also calculated using this hybrid dataset. The $\pm 10\%$ uncertainties assigned to the resonance contributions are used when evaluating the systematic error associated with this component in Sec. 4.5.1.

Mis-reconstructed decays, described by categories 4 & 5, are modelled using $B_s^0 \to D^{*0}(K^{*0}(892) \to K^- \pi^+)$ MC; the phase space MC is generated with $D^* \to D\gamma$ and $D^* \to D\pi^0$ decays in the same sample, so that mis-reconstructed decays of each kind are indistinguishable from each other. Seeing as these mis-reconstructed samples have lower selection efficiencies and do not peak in D^* mass, this does not pose a problem for the analysis. A systematic uncertainty to account for this is described in Sec. 4.5.1.

For events reconstructed in the $D\pi^0$ final state, the B mass PDFs are determined by performing invariant-mass fits to simulation samples of fully and mis- reconstructed $B_s^0 \to (D^{*0} \to D^0\pi^0)K^-\pi^+$ decays, alongside mis-reconstructed $B_s^0 \to (D^{*0} \to D^0\gamma)K^-\pi^+$ decays. The fully reconstructed sample is modelled using the sum of a Crystal Ball and an offset Gaussian, where as both the mis-reconstructed $D\pi^0$ and

$D\gamma$ samples are modelled using Gaussian functions; the fits are shown in Fig. 3.42.

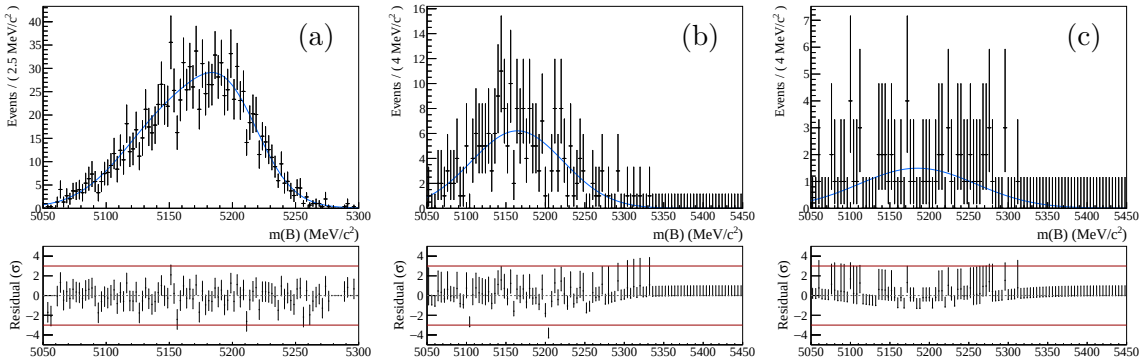


Figure 3.42: Fits to the $m(B)$ distributions of candidates reconstructed in the $D\pi^{0}$ final state for fully reconstructed $B_s^0 \to (D^{*0} \to D^0\pi^0)K^-\pi^+$ MC in (a), mis-reconstructed $B_s^0 \to (D^{*0} \to D^0\pi^0)K^-\pi^+$ MC in (b) and mis-reconstructed $B_s^0 \to (D^{*0} \to D^0\gamma)K^-\pi^+$ MC in (c).

The line shapes for the three D^* decay categories considered in the $D\pi^0$ final state are then combined into a single $B_s^0 \to D^{*0}K^-\pi^+$ component. The two mis-reconstructed PDFs are first combined, where the relative fraction of each is fixed using selection efficiencies, given in Tab. 3.7, and D^* branching fractions, given by Eqs. (1.8) and (1.9). The total mis-reconstructed component is then combined with the fully reconstructed $D^* \to D\pi^0$ PDF; the fraction of each, *i.e.* the amount of mis- to fully reconstructed D^* decays, is shared with the relative amount of mis- to fully reconstructed $B^{\pm} \to D^*h^{\pm}$ decays in the favoured mode. Relative efficiency differences, dominated by differences in box efficiencies, are accounted for.

For B_s^0 events reconstructed in the $D\gamma$ final state, the B mass PDFs are determined by performing invariant-mass fits to simulation samples of fully and mis-reconstructed $B_s^0 \to (D^{*0} \to D^0\gamma)K^-\pi^+$ decays, and partially and mis-reconstructed $B_s^0 \to (D^{*0} \to D^0\pi^0)K^-\pi^+$ decays. Fully reconstructed $B_s^0 \to (D^{*0} \to D^0\gamma)K^-\pi^+$ decays have a negligible efficiency when reconstructed in the PR slice of $m(B)$, therefore a 1D fit using the sum of two Crystal Balls is performed only to the FR slice of $m(B)$ for this component. Similarly, partially reconstructed $B_s^0 \to (D^{*0} \to D^0\pi^0)K^-\pi^+$ decays have a negligible efficiency when reconstructed in the FR slice of $m(B)$, therefore a 1D fit using a Crystal Ball function is performed only to the PR slice of $m(B)$ for

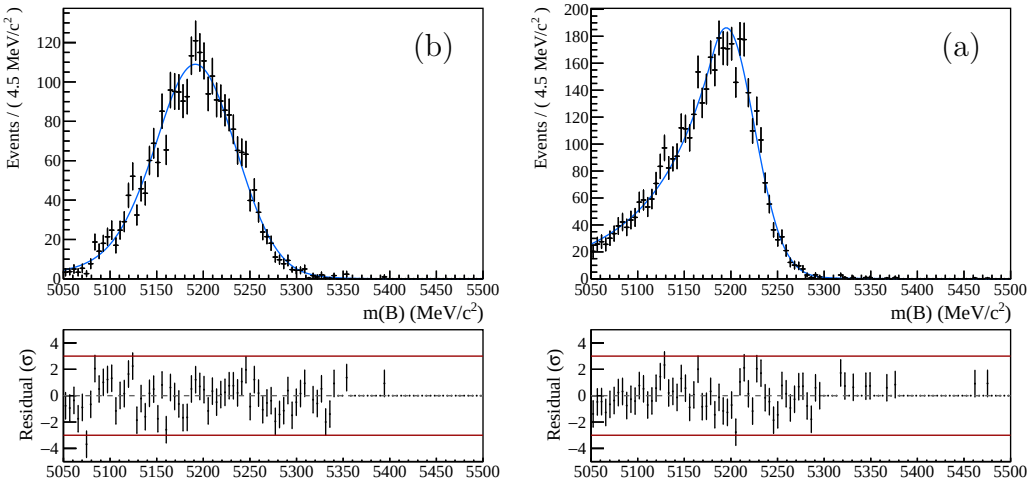


Figure 3.43: Fits to candidates reconstructed in the $D\gamma$ final state. Partially reconstructed $B_s^0 \to (D^{*0} \to D^0\pi^0)K^-\pi^+$ MC in the PR slice of $m(B)$ is shown in (a); fully reconstructed $B_s^0 \to (D^{*0} \to D^0\gamma)K^-\pi^+$ MC in the FR slice of $m(B)$ is shown in (b).

this component. These fits are shown in Fig. 3.43. For the mis-reconstructed samples, simultaneous fits are performed to the FR and PR slices of $m(B)$ simultaneously; the distributions are modelled using Gaussian functions. The mis-reconstructed $D^* \to D\pi^0$ and $D^* \to D\gamma$ distributions are very similar, therefore a single example is given for mis-reconstructed $B_s^0 \to (D^{*0} \to D^0\pi^0)K^-\pi^+$ decays in Fig. 3.44.

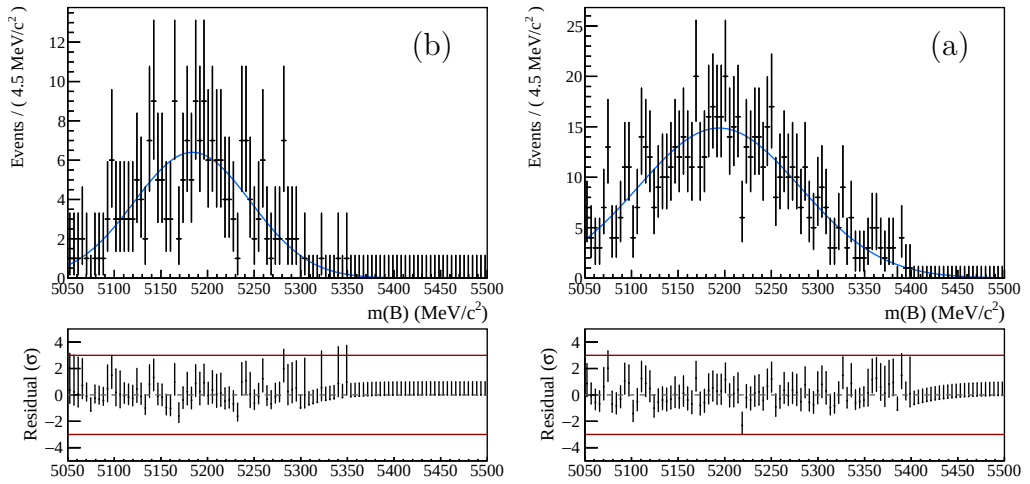


Figure 3.44: Simultaneous fits to the PR slice of $m(B)$, (a), and the FR slice of $m(B)$, (b), for $B_s^0 \to (D^{*0} \to D^0\pi^0)K^{*0}$ MC mis-reconstructed in the $D\gamma$ final state.

The PDFs of the four D^* decay categories considered in the $D\gamma$ final state are then combined into a single $B_s^0 \to D^{*0}K^-\pi^+$ component. The fully reconstructed $D\gamma$

and partially-reconstructed $D\pi^0$ PDFs are first combined, where the relative fraction of each is shared with the relative amount of fully to partially reconstructed signal decays in the favoured mode. The two mis-reconstructed PDFs are also combined, where the relative fraction of each is fixed using selection efficiencies and D^* branching fractions. The two composite PDFs are then used to build a total $B^0_s \to D^{*0}K^-\pi^+$ PDF; the proportion of each, *i.e.* the fraction of mis- to correctly reconstructed D^* decays, is shared with the relative amount of mis- to correctly reconstructed $B^{\pm} \to D^*h^{\pm}$ decays. Relative efficiency differences, again dominated by differences in box efficiencies, are accounted for.

3.3.5 Mis-identified decays

Events with a pion companion end up in the $B^{\pm} \to D^*K^{\pm}$ sub-sample at a rate $(1 - \epsilon^{\pi}_{PID})$, and events with a kaon companion end up in the $B^{\pm} \to D^*\pi$ sub-sample at a rate $(1 - \epsilon^{K}_{PID})$, where the PID efficiencies are given in Sec. 3.1.6. Mis-identified decays must therefore be accounted for and are modelled for each component in the invariant-mass fits using simulation samples that have been assigned the incorrect companion mass hypothesis. When pions are mis-identified as kaons, the reconstructed B invariant mass is shifted upwards; when kaons are mis-identified as pions, the reconstructed B invariant mass is shifted downwards. This effect is illustrated in Fig. 3.45, which shows the 2D distributions of mis-identified $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ and $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)K^{\pm}$ simulation samples. This means that box efficiencies, ϵ_{box}, and consequently ϵ_{tot}, are very different for correctly identified and mis-identified decays. The yields of mis-identified components are therefore fixed to be $(1-\epsilon^{h}_{PID}) \times (\epsilon^{\text{Mis ID}}_{tot}/\epsilon^{\text{Correct ID}}_{tot})$ the size of the associated correctly identified yields, where $h \in (\pi, K)$.

Mis-identified components are included in the invariant-mass fits to data for the signal modes and all favoured mode backgrounds. A summary of the PDFs used to parameterise each decay mode are given in Tabs. 3.11 and 3.12. Notably, partially reconstructed $B \to D^*\pi^{\pm}\pi$ decays reconstructed in the D^*K sub-sample

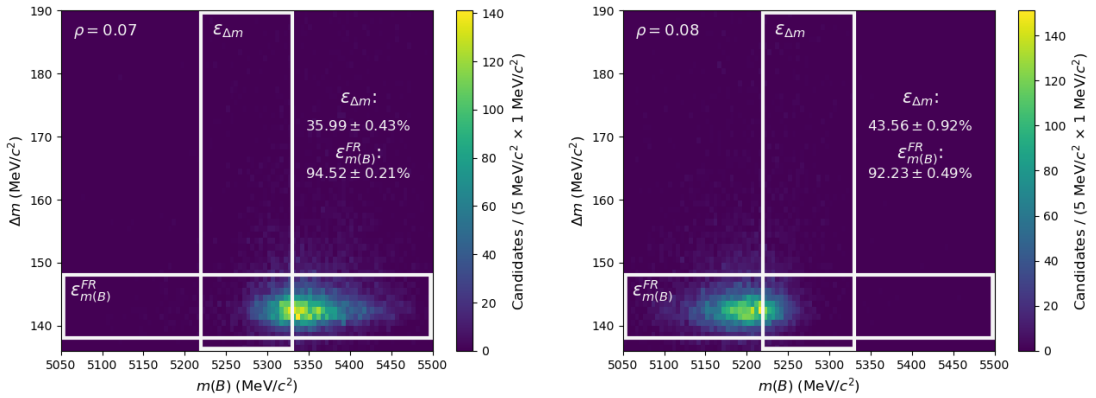


Figure 3.45: The 2D distribution, $m(B)$ vs. Δm, of $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ reconstructed in the D^*K sub-sample is depicted on the LHS; $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)K^{\pm}$ decays reconstructed in the $D^*\pi$ sub-sample are shown on the RHS. The vertical (horizontal) rectangles encompass events included in the Δm ($m(B)$) sample, where the box efficiencies are printed as a reference. The efficiencies match up to those list in Tab. 3.5. The Pearson correlation coefficients, ρ, of the 2D mass distributions are also given.

have a peaking component in Δm, modelled by the sum of two CBs, as the upwards shift in B mass pushes true D^* mesons into the $m(B)$ box window. On the other hand, partially reconstructed $B \to D^*K^{\pm}\pi$ decays are shifted below the invariant mass range of the fit when the kaon is mis-identified as a pion, therefore these components are not modelled.

Decay	Reconstructed as $D\pi^0$		Reconstructed as $D\gamma$		
	$m(B)$	Δm	$m(B)_{\mathrm{PR}}$	Δm	$m(B)_{\mathrm{FR}}$
$B^{\pm} \to (D^* \to D\pi^0)\pi^{\pm}$ as D^*K	f_{DCB}	f_{DCB}	f_{DCB}	f_{DCB}	f_{CB}
$B^{\pm} \to (D^* \to D\gamma)\pi^{\pm}$ as D^*K	-	-	-	f_{DCB}	f_{DCB}
Mis-reco. $B^{\pm} \to (D^* \to D\pi^0)\pi^{\pm}$ as D^*K	f_{CB}	f_{D^*-D}	f_{CB}	f_{D^*-D}	$f_G + f_G + f_{CB}$
Mis-reco. $B^{\pm} \to (D^* \to D\gamma)\pi^{\pm}$ as D^*K	f_G	f_G	f_{CB}	f_{D^*-D}	f_{CB}
$B^0 \to (D^{*\mp} \to D\pi^{\mp})\pi^{\pm}$ as D^*K	f_G	f_{D^*-D}	f_{CB}	f_{D^*-D}	f_{CB}
$B \to D\pi^{\pm}\pi$ as D^*K	f_G	f_{D^*-D}	f_{CB}	f_{D^*-D}	f_{CB}
$B \to D^*\pi^{\pm}\pi$ as D^*K	f_{CB}	$f_{D^*-D} + f_{DCB}$	f_G	$f_{D^*-D} + f_{DCB}$	$f_{CB} + f_G$

Table 3.11: Summary of PDFs used to parameterise decay modes with a companion pion mis-identified as a kaon.

Simultaneous fits to mis-identified $D\pi^0$ signal simulation are given as an example

Decay	Reconstructed as $D\pi^0$		Reconstructed as $D\gamma$		
	$m(B)$	Δm	$m(B)_{\rm PR}$	Δm	$m(B)_{\rm FR}$
$B^{\pm} \to (D^* \to D\pi^0)K^{\pm}$ as $D^*\pi$	$f_{CB} + f_G$	f_{DCB}	f_{DCB}	f_{DCB}	f_{CB}
$B^{\pm} \to (D^* \to D\gamma)K^{\pm}$ as $D^*\pi$	-	-	-	f_{DCB}	f_{DCB}
Mis-reco. $B^{\pm} \to (D^* \to D\pi^0)K^{\pm}$ as $D^*\pi$	f_{CB}	f_{D^*-D}	f_{CB}	f_{D^*-D}	$f_G + f_G + f_{CB}$
Mis-reco. $B^{\pm} \to (D^* \to D\gamma)K^{\pm}$ as $D^*\pi$	f_G	f_{D^*-D}	f_G	f_{D^*-D}	f_{CB}
$B^0 \to (D^{*\mp} \to D\pi^{\mp})K^{\pm}$ as $D^*\pi$	f_G	f_{D^*-D}	f_{CB}	f_{CB}	f_G
$B \to DK^{\pm}\pi$ as $D^*\pi$	f_G	f_{D^*-D}	f_{CB}	f_{CB}	f_{CB}
$B \to D^*K^{\pm}\pi$ as $D^*\pi$	-	-	-	-	-

Table 3.12: Summary of PDFs used to parameterise decay modes with a companion kaon mis-identified as a pion.

in Fig. 3.46; $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ decays reconstructed in the $(D\pi^0)K$ sub-sample are shown on the top, and $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)K^{\pm}$ decays reconstructed in the $(D\pi^0)\pi$ sub-sample are shown on the bottom.

3.3.6 Combinatorial background

There are three distinct sources of combinatorial background that have been considered in this thesis:

1. Charmless backgrounds, of the form $B \to hhhX$. Due to the narrow region of Δm phase space that this analysis is conducted in, there are no peaking charmless backgrounds. The small number of charmless events that are present fall harmlessly into the background PDFs already defined.

2. Neutral combinatorial, where correctly reconstructed Dh decays are combined with a neutral particle from the underlying event. Events of this kind are labelled mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays and have already been accounted for in Sec. 3.3.3.

3. Charged combinatorial, where the correct neutral D^* meson is reconstructed

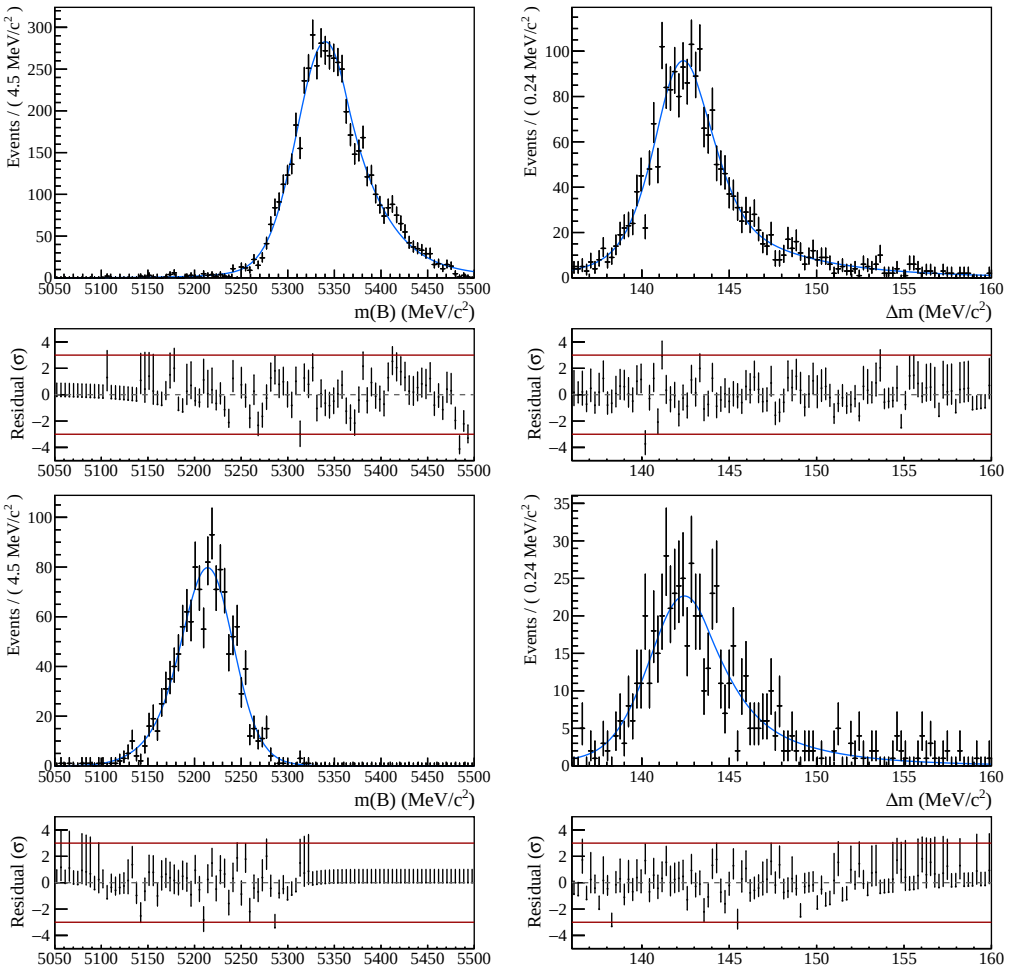


Figure 3.46: The simultaneous fit to the $m(B)$ and Δm distributions of $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ decays reconstructed in the $(D\pi^{0})K$ sub-sample is depicted in the top two plots. The simultaneous fit for $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})K^{\pm}$ decays reconstructed in the $(D\pi^{0})\pi$ sub-sample is shown in the bottom two plots.

and combined with a random track, or a D meson is combined with a neutral particle from the underlying event and a random track. Events of this kind will be considered here.

By nature, combinatorial backgrounds formed from the combination of a true D^{*} meson, or a D meson combined with a neutral particle from the underlying event, and a random track are non peaking in B mass. Their shape will also not be exponential due to the removal of events far out to the left and right hand sides of the fit (with the $4900\,\mathrm{MeV}/c^{2} < m(Dh) < 5200\,\mathrm{MeV}/c^{2}$ mass window cut). These events are therefore modelled with a wide Gaussian in $m(B)$, with mean $5200 \pm 100\,\mathrm{MeV}/c^{2}$ and width

$300 \pm 100\,\mathrm{MeV}/c^2$. These parameters are fixed in the invariant-mass fits to data and varied within the given errors as a source of systematic uncertainty.

In the Δm sample, combinatorial events share peaking and flat backgrounds with fully and mis-reconstructed $B^\pm \to D^* h^\pm$ decays, respectively representing events with true and fake D^* candidates; the proportion of these PDFs is then constrained by the same ratio in signal decays. For the fit to data reconstructed in the $D\gamma$ final state, the proportion of combinatorial events sitting under the partially-reconstructed $D\pi^0$ peak, with respect to the fully-reconstructed $D\gamma$ peak, is also constrained by the relative size of these components in signal decays. The justification for these shared fractions is that the well constrained ratios in the signal modes represent the proportion of D^* mesons that are reconstructed as either $D\pi^0$ or $D\gamma$ candidates, or are mis-reconstructed, and therefore can be shared with any events containing D^* candidates.

The associated box efficiencies, used to split the total combinatorial yield in each fit category into the different mass samples, are estimated using the PDF shapes described. In B mass, this amounts to integrating the wide Gaussian over the box cut range. In Δm, the efficiencies are calculated from a normalised sum of the efficiencies for fully and mis-reconstructed $B^\pm \to D^* h^\pm$ decays, where the relative proportion of each component is taken into account.

3.3.7 Crossfeed

Favoured $B^\pm \to (D^* \to [K^\pm\pi^\mp]_D\pi^0/\gamma)h^\pm$ decays occur at a much higher rate than their suppressed ADS mode counterparts due to the suppression factors $r_B^{D^*h}$ and r_D, as can be seen from the rate equations given by Eqs. (1.23) to (1.26). This is most prominent in the $B^\pm \to D^*\pi^\pm$ modes, which have a smaller amplitude ratio $r_B^{D^*\pi}$. When both D decay products are mis-identified ($\pi \to K$ and $K \to \pi$), favoured mode candidates end up in the ADS mode samples. This effect would dwarf the ADS mode signal if it weren't for three specific selection requirements implemented to reduce the rate to the percentage level:

- The $\pm 25\,\mathrm{MeV}/c^2$ D mass window cut, which has a 97% efficiency when applied

to correctly identified decays and a lower efficiency for doubly misidentified decays, as the latter form wider resonances.

- The D meson mass calculated when both the kaon and pion hypotheses are swapped must lie outside a ± 15 MeV/c^2 mass window. This efficiently removes favoured decays reconstructed in the ADS sample, which form the true D mass distribution when reconstructed in this way; 93% of the ADS signal is retained.

- The PID requirements placed on the D decay products; pions must have a $\Delta LL_{K/\pi}$ value less than -2 (-1) in Run 1 (Run 2); kaons must have a $\Delta LL_{K/\pi}$ value greater than 2 (1) in Run 1 (Run 2).

These cuts are applied to both favoured and suppressed mode samples to ensure that CP observables are unaffected; it is assumed that their effect on both samples, when correctly identified, is the same.

The efficiency of these requirements on correctly and doubly-misidentified D decays was calculated for the 2-body $B^{\pm} \to D^{(*)}h^{\pm}$ analysis [31] using simulation samples of $B^{\pm} \to [K^{\pm}\pi^{\mp}]_D\pi^{\mp}$ decays, reconstructed with the double swapped D mass hypothesis. The PID efficiencies were evaluated using the LHCb software `PIDCalib`, and the efficiencies of the D mass window cuts were determined using smeared simulation samples to ensure accurate representation of data. The rate of favoured mode crossfeed in the ADS sample, ϵ_{CF}, was then calculated by taking the efficiency ratio of these requirements on doubly-misidentified events relative to correctly identified events. Due to the different PID cuts applied to D decay products in Run 1 and Run 2 data, separate measurements were made for each run. The values quoted below are used directly in this work:

$$\epsilon_{CF} = \begin{cases} (5.34 \pm 0.27) \times 10^{-5}, & \text{in Run 1} \\ (10.16 \pm 0.37) \times 10^{-5}, & \text{in Run 2.} \end{cases} \tag{3.44}$$

In the invariant-mass fits to data, these numbers are averaged using the relative luminosity of each run. The yield of the favoured mode crossfeed in the ADS mode is then fixed from the correctly identified favoured mode yield using this average rate, in both the D^*K and $D^*\pi$ sub-samples.

Decay	Reconstructed as $D\pi^0$		Reconstructed as $D\gamma$		
	$m(B)$	Δm	$m(B)_{\mathrm{PR}}$	Δm	$m(B)_{\mathrm{FR}}$
$B^{\pm} \to (D^* \to D_{SWAP}\pi^0)h^{\pm}$	f_G	f_{DCB}	f_G	f_{DCB}	f_G
$B^{\pm} \to (D^* \to D_{SWAP}\gamma)h^{\pm}$	-	-	-	f_{DCB}	f_G

Table 3.13: Summary of PDFs used to parameterise favoured to ADS mode crossfeed, where f_{CB} represents the Crystal Ball function defined in Eq. (3.30), f_{DCB} represents the double Crystal Ball function defined in Eq. (3.31) and f_G represents a Gaussian function.

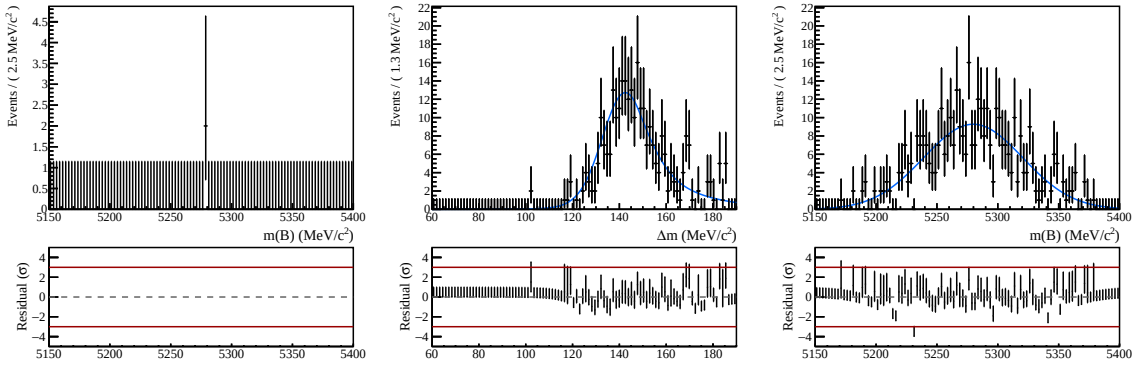


Figure 3.47: Simultaneous fit in Δm and the FR slice of $m(B)$ to $B^{\pm} \to (D^* \to D_{SWAP}\gamma)h^{\pm}$ MC, where the D meson is reconstructed as if it were an ADS mode candidate with both the kaon and pion mass hypothesis swapped. Both companion final states, $h \in (\pi, K)$ are included to increase statistics; 2011-2018 samples are used, where all years are considered together. It can be seen that the efficiency for reconstructing mass-swapped $D\gamma$ signal in the PR slice of $m(B)$ is negligible by the low numbers of events in the left hand plot.

Double mis-identification of the D decay products results in wider invariant mass distributions, therefore simultaneous fits to simulated favoured decays where both D decay products are given the incorrect mass hypothesis are used to determine the crossfeed PDFs. All selections are applied to the samples except D decay product PID cuts, and the BDT cuts are loosened to increase statistics. $D^*\pi$ and D^*K samples are combined and all years considered together to ensure maximum statistics. An associated crossfeed component is included for every peaking signal component: the PDFs used to parameterise the mass samples are given in Tab. 3.13. As an example, the simultaneous fit to $B^{\pm} \to (D^* \to D_{SWAP}\gamma)h^{\pm}$ decays reconstructed in the $D\gamma$ final state is shown in Fig. 3.47.

3.3.8 Semi-leptonic backgrounds

Semi-leptonic backgrounds, for example $B^{\pm} \to D^{(*)}\mu^{\pm}\nu_{\mu}$, have the potential to infiltrate the $D^{*}\pi$ sub-sample when the neutrino is not reconstructed and the muon is mis-identified as a pion. This mis-identification occurs due to the similar masses of the two particles, making them harder to distinguish in the RICH detector (as is evident in Fig. 2.9). Decays of this kind are not explicitly modelled in the signal-extraction fits, therefore events that are consistent with a muon hypothesis are removed from the final data sample by requiring that the companion particle does not trigger the muon line (`isMuon==0`). This binary decision tests whether the extrapolation of a long or downstream track through the muon stations is consistent with a muon hypothesis.

In the $D^{*}\pi$ sub-sample, $\sim 3\%$ of background events sitting at low B mass are removed by this cut. Decays in flight, when the companion hadron undergoes the decay $\pi^{\pm} \to \mu^{\pm}\nu_{\mu}$ or $K^{\pm} \to \mu^{\pm}\nu_{\mu}$ in the detector, are also removed. The rate of decay in flight to muonic final states for pions and kaons is not equal, therefore the efficiency of the `isMuon` cut differs for data in the $D^{*}\pi$ and $D^{*}K$ sub-samples. The total, luminosity-averaged efficiencies for each signal decay mode are given below:

$$\epsilon_{isMuon}(B^{\pm} \to (D^{*} \to D\pi^{0})\pi^{\pm}) = (99.21 \pm 0.16)\%$$
$$\epsilon_{isMuon}(B^{\pm} \to (D^{*} \to D\pi^{0})K^{\pm}) = (98.05 \pm 0.26)\%$$
$$\epsilon_{isMuon}(B^{\pm} \to (D^{*} \to D\gamma)\pi^{\pm}) = (99.45 \pm 0.06)\%$$
$$\epsilon_{isMuon}(B^{\pm} \to (D^{*} \to D\gamma)K^{\pm}) = (99.10 \pm 0.08)\%,$$

where the quoted uncertainties are due to finite simulation sample sizes. Efficiency differences between the two neutral final states are due to selection effects.

The efficiency of the `isMuon==0` requirement for each decay mode has been included when evaluating the offline selection efficiencies, ϵ_{sel}, given in Sec. 3.2.4. This ensures that the ratios $R_{K/\pi}^{K\pi,\pi^{0}/\gamma}$ are calculated accurately. All other observables are ratios with the companion particle common to the numerator and denominator and are therefore unbiased by the cut (across all D decay modes, the same proportion of signal events will be removed from each category of the fit).

4

Measurement of *CP* observables

In this chapter, measurements of *CP* observables are made using the signal extraction fits to data reconstructed in the $D\gamma$ and $D\pi^0$ final states. This chapter builds on the D1D method detailed in Sec. 3.2 of the previous chapter; the full setup of the *CP* fits are detailed in Sec. 4.1. The raw fit results are given in Sec. 4.2; fit stability is demonstrated in Sec. 4.3; the observables are corrected for the double counting of events in Sec. 4.4; systematic uncertainties are evaluated in Sec. 4.5. The final results are presented in Sec. 4.6 and validated using comparisons with previous measurements.

4.1 *CP* fit setup

In order to measure the CP observables given by Eqs. (1.31) to (1.37), binned extended maximum-likelihood fits are performed. The fit to data reconstructed in the $D\pi^0$ final state is simultaneous across Δm and $m(B)$; the fit to data reconstructed in the $D\gamma$ final state is simultaneous across Δm and 2 slices of $m(B)$. For details on how the datasets are split into these separate mass samples, please refer to Sec. 3.2.2. The invariant-mass fits are also simultaneous across the following 16 categories: $(D^*\pi, D^*K) \times (B^+, B^-) \times (K\pi, KK, \pi\pi, \pi K)$. Candidates considered in the D^*K and $D^*\pi$ sub-samples are separated using a binary PID cut to ensure statistical independence. Candidates whose companion hadron has a $\Delta LL_{K/\pi}$ value

less than 12 are reconstructed as $D^*\pi$, whilst those with a $\Delta LL_{K/\pi}$ value greater than 12 are reconstructed as D^*K.

A total simultaneous PDF describing the data distribution in each category is formed by combining the individual line shapes described in Sec. 3.3. Signal yields are expressed in terms of the CP parameters of interest (Sec. 1.4.4), and are allowed to vary freely. The log-likelihood estimator to be maximised is therefore:

$$-\ln\mathcal{L} = \sum_c -\ln\mathcal{L}_c(\boldsymbol{N}^c_{\boldsymbol{box}}, \boldsymbol{\theta}; \boldsymbol{\epsilon}_{\boldsymbol{m(B)}}, \boldsymbol{\epsilon}_{\boldsymbol{\Delta m}}, \boldsymbol{N}^c), \tag{4.1}$$

which is a sum over all categories, c, of the D1D log-likelihood function, $\mathcal{L}_c$, defined by Eq. (3.23). Here, $\boldsymbol{N}^c_{\boldsymbol{box}}$ are the set of measured yields within a given category; $\boldsymbol{\theta}$ are the set of variables that parameterise the distribution, including the desired observables; $\boldsymbol{\epsilon}_{\boldsymbol{m(B)}}$ and $\boldsymbol{\epsilon}_{\boldsymbol{\Delta m}}$ are sets of box efficiencies and $\boldsymbol{N}^c$ are the observed per-bin yields within a given category.

Data samples spanning all years (2011-2018), summed over magnet polarity, are included. In the $D\pi^0$ case, a bin width of 10 (1) MeV/c^2 is used over the range $5050-5450$ $(136-190)$ MeV/c^2 in $m(B)$ (Δm), resulting in $40+54=94$ bins per category, and $640+848=1504$ bins in total. In the $D\gamma$ case, a bin width of 10 (2) MeV/c^2 is used over the range 5050−5500 (60−190) MeV/c^2 in $m(B)$ (Δm), resulting in $45+75+45=165$ bins per category, and $720+1200+720=2640$ bins in total.

4.1.1 Production and detection asymmetry corrections

Raw charge asymmetries, given by Eq. (4.2), must be corrected for sources of experimental asymmetry in order to measure CP asymmetries:

$$A_{raw} = \frac{N^- - N^+}{N^- + N^+}. \tag{4.2}$$

Here, N^+ and N^- are the B^+ and B^- yields, respectively.

$B^\pm$ production asymmetry

Effective values for the $B^\pm$ production asymmetry, A_{B^+}, are measured in the $D\pi^0$ and $D\gamma$ invariant mass fits. This is achieved by fixing the favoured mode asymmetries,

$A_\pi^{K\pi,\pi^0/\gamma}$, thereby releasing a degree of freedom for A_{B^+} to float. $A_\pi^{K\pi,\pi^0}$ is fixed to zero with asymmetric uncertainties $0^{+0.03\%}_{-0.08\%}$; $A_\pi^{K\pi,\gamma}$ is fixed to zero with asymmetric uncertainties $0^{+0.08\%}_{-0.03\%}$. The given uncertainties cover the possible range of CP asymmetry, calculated using the latest LHCb measurement of γ [23] and World-averaged values of hadronic parameters [40]. The effective production asymmetries are measured to be:

- $A_{B^+} = (-0.008 \pm 0.005)\%$ using the $D\pi^0$ invariant-mass fit.
- $A_{B^+} = (+0.005 \pm 0.003)\%$ using the $D\gamma$ invariant-mass fit.

where the quoted uncertainties are statistical. These measurements cannot be considered main results of the analysis as the data samples include other species of B hadron. This is also the reason for measuring the asymmetry separately for the two D^* decay samples: differences in data selection requirements result in different rates of each B hadron species in the invariant mass spectrum. The measurements are, however, consistent within 2σ.

Kaon and pion detection asymmetries

The detection asymmetry for a $K^+\pi^-$ pair was reported in the LHCb paper dedicated to measuring the $B^\pm$ production asymmetry [89]. The measurement was made in bins of $B^\pm$ momentum and pseudorapidity using Run 1 data. Averaging these results over 2011 and 2012 data, and across all kinematic bins, results in the detection asymmetry:

$$A_{K\pi} = (-0.960 \pm 0.080)\%, \tag{4.3}$$

where the quoted uncertainty combines both statistical and systematic uncertainties. The 2-body $B^\pm \to D^{(*)}h^\pm$ analysis corrected this uncertainty for an extension of the measurement to Run 2 data [31]. This was achieved by comparing the kinematic distributions of the companion hadron and D decay products for data collected during Run 1 and Run 2. Conservatively, assuming the observed deviation in kinematics translated to the same change in $A_{K\pi}$, the correction was evaluated to be:

$$A_{K\pi} = (-0.960 \pm 0.134)\%. \tag{4.4}$$

This value is used directly in this work and enters as a fixed parameter in the invariant-mass fits to correct the raw asymmetries of modes containing a $K^+\pi^-$ pair.

For the detection asymmetry of kaons in isolation, $A_{K\pi}$ must itself be corrected for the pion detection asymmetry, A_π. This quantity was also measured for Run 1 data, in bins of (p_{T}, η), in Ref. [89]. Averaging the results over kinematic bins yields:

$$A_\pi = (-0.064 \pm 0.018)\%, \tag{4.5}$$

where the quoted uncertainty combines both statistical and systematic uncertainties. The pion asymmetry is also assumed to apply to Run 2, assigning a larger uncertainty of 0.019% to account for kinematic differences of the D decay products and the companion hadron [31].

Modes containing a net non-zero number of kaons in isolation are correct by a factor of $A_K = A_{K\pi} + A_\pi$. Modes containing a net non-zero number of pions in isolation are corrected by a factor of A_π. All corrections are varied within the given uncertainties in order to determine the associated systematic uncertainties.

L0 hadron trigger asymmetry

An additional kaon asymmetry for candidates that are required to trigger `L0Hadron` must also be considered. K^- mesons have a larger nuclear interaction cross section than K^+ mesons, therefore are preferentially absorbed by detector material. This means that a higher proportion of positive kaons may reach the hadronic calorimeter and trigger the `L0Hadron` line.

This effect was quantified in the 2-body $B^\pm \to D^{(*)}h^\pm$ analysis [31]. Fits to selected candidates were performed in two categories: candidates passing the `L0_HadronDecision_TOS` trigger line, and candidates passing the `L0_Global_TIS` line but not `L0_HadronDecision_TOS`. The measured difference in $B^\pm$ production asymmetry between these two data samples was measured to be:

$$\delta A_{B^+} = A_{B^+}(\mathrm{TOS}) - A_{B^+}(\mathrm{TIS}) = (-2.54 \pm 1.19) \times 10^{-3},$$

which represents the L0 hadron trigger asymmetry in a sample of purely `L0Hadron` candidates. This choice is justified by the fact that A_{B^+} was determined by the

asymmetry in the favoured $B^{\pm} \to [K^{\pm}\pi^{\mp}]_D\pi^{\pm}$ mode, which contains a net non-zero number of kaons but net zero pions.

The percentage of `L0Hadron` candidates in the final datasets that enter the invariant-mass fits of this analysis are 79.4% for data reconstructed in the $D\pi^0$ final state, and 73.5% for data reconstructed in the $D\gamma$ final state. The effective additional kaon asymmetries due to the `L0Hadron` trigger are therefore $(-2.01 \pm 0.94) \times 10^{-3}$ and $(-1.87 \pm 0.88) \times 10^{-3}$, respectively. These values are added to the value of $A_{K\pi}$ given above, and varied when evaluating systematic uncertainties.

Definition of asymmetry observables in terms of raw asymmetry and associated corrections

The raw asymmetries in Eq. (4.2) must be corrected for the different sources of detection and production asymmetries listed above. The formulae used in the fit to convert raw asymmetries to the measured observables are therefore:

$$A_{B^+} = A_{\rm raw}(B^+ \to [[K^+\pi^-]_D\pi^0/\gamma]_{D^*}\pi^+) \; - A_{\pi}^{K\pi} - A_{K\pi} - A_{\pi} \tag{4.6}$$

$$A_{K}^{K\pi} = A_{\rm raw}(B^+ \to [[K^+\pi^-]_D\pi^0/\gamma]_{D^*}K^+) - A_{B^+} - A_{K\pi} - (A_{K\pi} + A_{\pi}) \tag{4.7}$$

$$A_{\pi}^{KK} = A_{\rm raw}(B^+ \to [[KK]_D\pi^0/\gamma]_{D^*}\pi^+) \quad - A_{B^+} \qquad - A_{\pi} \tag{4.8}$$

$$A_{K}^{KK} = A_{\rm raw}(B^+ \to [[KK]_D\pi^0/\gamma]_{D^*}K^+) \quad - A_{B^+} \qquad - (A_{K\pi} + A_{\pi}) \tag{4.9}$$

$$A_{\pi}^{\pi\pi} = A_{\rm raw}(B^+ \to [[\pi\pi]_D\pi^0/\gamma]_{D^*}\pi^+) \quad - A_{B^+} \qquad - A_{\pi} \tag{4.10}$$

$$A_{K}^{\pi\pi} = A_{\rm raw}(B^+ \to [[\pi\pi]_D\pi^0/\gamma]_{D^*}K^+) \quad - A_{B^+} \qquad - (A_{K\pi} + A_{\pi}) \tag{4.11}$$

$$A_{\pi}^{\pi K} = A_{\rm raw}(B^+ \to [[\pi^+K^-]_D\pi^0/\gamma]_{D^*}\pi^+) \; - A_{B^+} + A_{K\pi} - A_{\pi} \tag{4.12}$$

$$A_{K}^{\pi K} = A_{\rm raw}(B^+ \to [[\pi^+K^-]_D\pi^0/\gamma]_{D^*}K^+) - A_{B^+} + A_{K\pi} - (A_{K\pi} + A_{\pi}). \tag{4.13}$$

Here, the $-A_{K\pi}$ factors correct for the detection asymmetry of a $K^+\pi^-$ pair, and $+A_{K\pi}$ for a $K^-\pi^+$ pair. Factors of $-A_{\pi}$ correct for the pion detection asymmetry, and the $-(A_{K\pi} + A_{\pi}) = -A_K$ factors correct for the kaon detection asymmetry.

Equivalent expressions for B^- decays are obtained by changing the sign of the A_{B^+}, $A_{K\pi}$, A_π, and $(A_{K\pi} + A_\pi)$ terms.

Different factors are assigned for each final state, with the following reasoning:

- A_{B^+}: $K^-\pi^+$ pair from D is assigned a factor $-A_{K\pi}$, and the companion π^- a factor $-A_\pi$.
- $A_K^{K\pi}$: $K^-\pi^+$ pair from D is assigned a factor $-A_{K\pi}$, and the companion K^- a factor $-(A_{K\pi} + A_\pi)$, where A_π is added in order to isolate the kaon detection asymmetry A_K.
- A_π^{KK} and $A_\pi^{\pi\pi}$: the D decay has zero net kaons and pions and thus is assigned no factor. A factor $-A_\pi$ is assigned to the π^- companion hadron.
- A_K^{KK} and $A_K^{\pi\pi}$: the D decay has zero net kaons and pions and thus is assigned no factor. A factor $-(A_{K\pi} + A_\pi)$ is assigned to the K^- companion hadron.
- $A_\pi^{\pi K}$: $K^+\pi^-$ pair from D is assigned a factor $+A_{K\pi}$, and the companion π^- a factor $-A_\pi$.
- $A_K^{\pi K}$: $K^+\pi^-$ pair from D is assigned a factor $+A_{K\pi}$, and the companion K^- a factor $-(A_{K\pi} + A_\pi)$.

4.1.2 Difference in CP asymmetry between KK and $\pi\pi$ decays

CP violation in charm decays was observed by LHCb in 2019 [15]; in this work, asymmetry observables measured across the KK and $\pi\pi$ final states are corrected for differences in CP violating effects using these results. The difference in CP asymmetry between $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays was found to be:

$$\Delta A_{CP} = A_{CP}(D^0 \to KK) - A_{CP}(D^0 \to \pi\pi) = (-15.4 \pm 2.9) \times 10^{-4}. \quad (4.14)$$

CP asymmetries for the KK and $\pi\pi$ final states are therefore defined separately, in

terms of the desired observable $A_h^{CP,\pi^0/\gamma}$ and ΔA_{CP}:

$$A_h^{CP,\pi^0/\gamma}(KK) = A_h^{CP,\pi^0/\gamma} + \frac{\Delta A_{CP}}{2} \tag{4.15}$$

$$A_h^{CP,\pi^0/\gamma}(\pi\pi) = A_h^{CP,\pi^0/\gamma} - \frac{\Delta A_{CP}}{2}. \tag{4.16}$$

This makes no assumption about the values of the individual CP asymmetries, $A_{CP}(D^0 \to KK)$ or $A_{CP}(D^0 \to \pi\pi)$, by splitting the ΔA_{CP} correction equally across both modes. A systematic uncertainty is assigned to this correction by varying it within the measured uncertainty.

4.1.3 Yield parameterisation

The component yields that pass the box cuts within each category, $\boldsymbol{N^c_{box}}$, do not all enter the fit as floating parameters. Instead, there are inter-category relations between yields in order to extract the parameters of interest; this parameterisation is described below.

The total yield of each component, for a given combination of companion particle, D decay mode and D^* final state, is defined as the sum of the B^+ yield (N^+) and the B^- yield (N^-):

$$N_{tot} = N^+ + N^-. \tag{4.17}$$

In practice, this is achieved by defining the split by charge yields in terms of the total yield, summed over B charge, and the specific decay mode asymmetry, $A^{h_B}_{h_D h'_D}$, given by Eqs. (4.7) to (4.13). Here, $h_D h'_D$ represent the D decay products and h_B the companion particle:

$$N^- = \frac{N_{tot}}{2} \times (1 + A^{h_B}_{h_D h'_D}) \tag{4.18}$$

$$N^+ = \frac{N_{tot}}{2} \times (1 - A^{h_B}_{h_D h'_D}). \tag{4.19}$$

The B^+/B^- yields are themselves calculated by summing correctly identified candidates ($N^\pm_{\text{trueID}}$) and mis-identified candidates ($N^\pm_{\text{misID}}$):

$$N^\pm = N^\pm_{\text{trueID}} + N^\pm_{\text{misID}}, \tag{4.20}$$

which is implemented by parameterising the correctly and mis-identified yields in the following way:

$$N^{\pm}_{\text{trueID}}(B^{\pm} \to D^{*}\pi^{\pm}) = \epsilon^{\pi}_{PID} \times N^{\pm}(B^{\pm} \to D^{*}\pi^{\pm}) \tag{4.21}$$

$$\begin{aligned} N^{\pm}_{\text{misID}}(B^{\pm} \to D^{*}\pi^{\pm}) &= (1 - \epsilon^{\pi}_{PID}) \times N^{\pm}(B^{\pm} \to D^{*}\pi^{\pm}) \\ &\quad \times \epsilon^{\text{misID/trueID}}_{tot}(B^{\pm} \to D^{*}\pi^{\pm}) \end{aligned} \tag{4.22}$$

$$N^{\pm}_{\text{trueID}}(B^{\pm} \to D^{*}K^{\pm}) = \epsilon^{K}_{PID} \times N^{\pm}(B^{\pm} \to D^{*}K^{\pm}) \tag{4.23}$$

$$\begin{aligned} N^{\pm}_{\text{misID}}(B^{\pm} \to D^{*}K^{\pm}) &= (1 - \epsilon^{K}_{PID}) \times N^{\pm}(B^{\pm} \to D^{*}K^{\pm}) \\ &\quad \times \epsilon^{\text{misID/trueID}}_{tot}(B^{\pm} \to D^{*}K^{\pm}), \end{aligned} \tag{4.24}$$

where $\epsilon^{\text{misID/trueID}}_{tot}(B^{\pm} \to D^{*}h^{\pm})$ is the ratio of selection efficiencies for the mis-identified $B^{\pm} \to D^{*}h^{\pm}$ component, compared to the correctly identified $B^{\pm} \to D^{*}h^{\pm}$ component.

Further relations are dependent on the D decay mode. The yields $N_{tot}(B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\pm}]_{D}\pi^{0}/\gamma)\pi^{\pm})$, $N_{tot}(B^{\pm} \to (D^{*} \to [K^{\pm}K^{\pm}]_{D}\pi^{0}/\gamma)\pi^{\pm})$ and $N_{tot}(B^{\pm} \to (D^{*} \to [\pi^{\pm}\pi^{\pm}]_{D}\pi^{0}/\gamma)\pi^{\pm})$ float freely in both invariant-mass fits. All other yields are defined in terms of these, and the ratios of interest:

- The summed over charged favoured mode $D^{*}K$ yields are defined in terms of the summed over charge favoured mode $D^{*}\pi$ yields in order to measure the floating ratio $R^{K\pi,\pi^{0}/\gamma}_{K/\pi}$, where the efficiency correction $R^{K\pi,\pi^{0}/\gamma}_{\epsilon(K/\pi)}$ will be explained in Sec. 3.2.4:

$$\begin{aligned} &N_{tot}(B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\pm}]_{D}\pi^{0}/\gamma)K^{\pm}) = \\ &\quad R^{K\pi,\pi^{0}/\gamma}_{K/\pi} \times R^{K\pi,\pi^{0}/\gamma}_{\epsilon(K/\pi)} \times N_{tot}(B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\pm}]_{D}\pi^{0}/\gamma)\pi^{\pm}). \end{aligned} \tag{4.25}$$

- The summer over charge $D^{*}K$ yields in the CP modes are defined in terms of the equivalent summed over charge $D^{*}\pi$ yields and the observables $R^{CP,\pi^{0}/\gamma}$:

$$\begin{aligned} &N_{tot}(B^{\pm} \to (D^{*} \to [K^{\pm}K^{\pm}]_{D}\pi^{0}/\gamma)K^{\pm}) = \\ &\quad R^{CP,\pi^{0}/\gamma} \times R^{K\pi,\pi^{0}/\gamma}_{K/\pi} \times N_{tot}(B^{\pm} \to (D^{*} \to [K^{\pm}K^{\pm}]_{D}\pi^{0}/\gamma)\pi^{\pm}) \\ &N_{tot}(B^{\pm} \to (D^{*} \to [\pi^{\pm}\pi^{\pm}]_{D}\pi^{0}/\gamma)K^{\pm}) = \\ &\quad R^{CP,\pi^{0}/\gamma} \times R^{K\pi,\pi^{0}/\gamma}_{K/\pi} \times N_{tot}(B^{\pm} \to (D^{*} \to [\pi^{\pm}\pi^{\pm}]_{D}\pi^{0}/\gamma)\pi^{\pm}). \end{aligned} \tag{4.26}$$

- The suppressed mode yields are defined split by charge so that the yield in the $D^*h^\pm$ category, $h \in (\pi, K)$, is parameterised in terms of the corresponding favoured mode yield and the observable $R_{h^\pm}^{\pi K,\pi^0/\gamma}$:

$$N^\pm(B^\pm \to (D^* \to [\pi^\pm K^\pm]_D\pi^0/\gamma)h^\pm) = R_{h^\pm}^{\pi K,\pi^0/\gamma} \times N^\pm(B^\pm \to (D^* \to [K^\pm\pi^\pm]_D\pi^0/\gamma)h^\pm). \tag{4.27}$$

4.1.4 Efficiency corrections

The ratios $R_{K/\pi}^{K\pi,\pi^0}$ and $R_{K/\pi}^{K\pi,\gamma}$ must be corrected to account for any efficiency difference between $B^\pm \to D^*\pi^\pm$ and $B^\pm \to D^*K^\pm$ decays. Corrections of the form:

$$R_{\epsilon(K/\pi)}^{K\pi,\pi^0/\gamma} = \frac{\epsilon_{tot}(B^\pm \to (D^* \to [K^\pm\pi^\mp]_D\pi^0/\gamma)K^\pm)}{\epsilon_{tot}(B^\pm \to (D^* \to [K^\pm\pi^\mp]_D\pi^0/\gamma)\pi^\pm)} \tag{4.28}$$

are therefore applied to $B^\pm \to D^*K^\pm$ yields, defined with respect to $B^\pm \to D^*\pi^\pm$, in the data fit. Here, ϵ_{tot} represents the signal mode selection efficiencies, given in Tab. 3.6.

4.1.5 Summary of fit components

Component yields in each decay mode are parameterised in term of CP parameters, as described in Sec. 4.1.3. The potential of each decay to violate CP is discussed here but in the invariant-mass fits to data any decays involving $B^\pm$ mesons are given the freedom to do so. A summary of the parameter values assigned for the various physics decay modes and details of what is fixed/floated in the signal extraction fits are given in Tab. 4.1; for details on the fit PDFs, please refer to Sec. 3.3.

In the following text, if the companion particle is represented as h, it is inclusive of both kaon and pion decay channels ($h = \pi/K$).

Signal modes

The signal modes comprise of $B^\pm \to D^*h^\pm$ decays, reconstructed in the following categories:

- $B^\pm \to (D^* \to D\pi^0)h^\pm$ decays fully reconstructed in the $D\pi^0$ final state or partially reconstructed in the $D\gamma$ final state.

Decay	Reco. as	$R^{K\pi}_{K/\pi}$	$A^{K\pi}_{K}$	R^{CP}	A^{CP}_{π}	A^{CP}_{K}	$R^{\pi K}_{\pi}$	$A^{\pi K}_{\pi}$	$R^{\pi K}_{K}$	$A^{\pi K}_{K}$
$\overline{(}\overline{B^0}) \to D^{*\pm}h^{\mp}$	$D\pi^0$	$(7.74 \pm 0.66)\%$	0.0 ± 0.0	1.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	$(3.80 \pm 0.18) \times 10^{-3}$	0.0 ± 0.0	$(3.80 \pm 0.18) \times 10^{-3}$	0.0 ± 0.0
	$D\gamma$	$(7.74 \pm 0.66)\%$	0.0 ± 0.0	1.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	$(3.80 \pm 0.18) \times 10^{-3}$	0.0 ± 0.0	$(3.80 \pm 0.18) \times 10^{-3}$	0.0 ± 0.0
$B \to Dh^{\pm}\pi$	$D\pi^0$	*free*	$(0.0 \pm 0.9)\%$	*free*	$(6.0 \pm 9.0)\%$	*free*	*free*	$(0.00 \pm 0.08)\%$	0.016 ± 0.012	-0.69 ± 0.50
	$D\gamma$	*free*	$(0.0 \pm 0.9)\%$	1.04 ± 0.06	$(6.0 \pm 9.0)\%$	*free*	*free*	*free*	0.016 ± 0.012	-0.69 ± 0.50
$B \to D^{*}h^{\pm}\pi$	$D\pi^0$	*free*	$(0 \pm 2)\%$	*free*	*free*	*free*	*free*	$(0.0 \pm 0.1)\%$	0.015 ± 0.010	0.0 ± 0.5
	$D\gamma$	*free*	*free*	*free*	*free*	*free*	*free*	*free*	0.015 ± 0.010	0.0 ± 0.5
$\Lambda^0_b \to \Sigma^+_c h^-$	$D\pi^0$	$7.3 \pm 3.7\%$	-	-	*free*	*free*	-	-	-	-
	$D\gamma$	*free*	-	-	*free*	*free*	-	-	-	-
$B^0_s \to D^{(*)}K^{\mp}\pi^{\pm}$	$D\pi^0$	-	-	-	-	-	-	-	-	0.0 ± 0.0
	$D\gamma$	-	-	-	-	-	-	-	-	0.0 ± 0.0

Table 4.1: The background yields in each decay mode are parameterised in term of *CP* parameters as described in Sec. 4.1.3. This table indicates which variables are given freedom in the invariant-mass fits (*free*) and provides the values of the parameters that are fixed. The constraints are given separately for the $D\pi^0$ and $D\gamma$ signal extraction fits, where any differences are to ensure fit stability. The suppressed mode parameters $R^{\pi K}_{h}$ and $A^{\pi K}_{h}$, defined in Eqs. (1.38) and (1.39), are given instead of $R^{\pi K}_{h^{\pm}}$ as measurements of the prior are reported in the reference papers and on HFLAV [40]. Neutral B meson decays are assumed to not exhibit any *CP* violation. The $B^{\pm} \to D^{*}h^{\pm}$ modes are not documented in this table as all signal observables are allowed to float.

- $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays fully reconstructed in the $D\gamma$ final state.

In the invariant-mass fits to data, the signal yields are parameterised as described in Sec. 4.1.3, with freedom given to all *CP* observables.

Mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays

Mis-reconstructed $B^{\pm} \to D^*h^{\pm}$ decays form broad resonances under the signal peaks B mass and flat distributions in Δm. They are constrained by their associated peaking signal component in order to be distinguished from the other mis-reconstructed backgrounds:

- In the $D\pi^0$ invariant-mass fit, the yields of mis-reconstructed $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays are constrained the yields of fully reconstructed $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays.

- In the $D\gamma$ invariant-mass fit, the yields of mis-reconstructed $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays are constrained the yields of fully reconstructed $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays.

- In the $D\gamma$ invariant-mass fit, the yields of mis-reconstructed $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays are constrained the yields of partially reconstructed $B^{\pm} \to (D^* \to D\pi^0)h^{\pm}$ decays.

In each fit category, the yield of a given mis-reconstructed component is related to its associated signal component via the selection efficiency ratio $\epsilon_{tot}^{\text{mis. reco.}}/\epsilon_{tot}^{\text{signal}}$ (given in Tabs. 3.6 and 3.7), multiplied by a freely varying parameter. This floating parameter is shared across all categories of the fit and accounts for any difference in multiplicity between simulation and data, representing the relative probability of combining a D^0 meson from the decay chain with a neutral particle from the underlying event.

Mis-reconstructed $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ decays that end up in the $D\pi^0$ data tuples, however, must be handled differently due to the lack of an associated peaking component. Instead, the *CP* parameters of these decays are fixed using the results of the $D\gamma$ invariant-mass fit and are varied as a source of systematic uncertainty.

$\overline{(}B^0\overline{)} \to (D^{*\pm} \to D\pi^{\pm})h^{\mp}$ decays

This analysis is time-independent, therefore it is assumed that decays involving $\overline{(}B^0\overline{)}$ mesons do not exhibit any *CP* violation. This results in the fixed *CP* parameters displayed in Tab. 4.1, where asymmetries have been set to zero, R_{CP} to one, and the ratios of ADS to favoured mode candidates are fixed using the well-known D branching fractions given in Eqs. (3.7) and (3.10). The branching ratio $\mathcal{B}(B^0 \to D^*K^{\pm})/\mathcal{B}(B^0 \to D^{*\mp}\pi^{\pm}) = (7.74 \pm 0.66)\%$ also enters both invariant-mass fits as a fixed parameter, taken from the PDG [19].

$B \to Dh^{\pm}\pi$ decays

As explain in Sec. 3.3.3, for the $B \to Dh^{\pm}\pi$ component, decays of the type $B^{\pm} \to Dh^{\pm}\pi^0$ and $B^0 \to Dh^{\pm}\pi^{\mp}$ are both considered. These decays are allowed to *CP* violate, but it is necessary to fix certain *CP* parameters in order to improve fit stability. These fixed parameters are summarised Tab. 4.1, and are taken either from previous measurements [90, 91] or calculated using World-averaged values of fundamental hadronic parameters [40] and the recent LHCb measurement of γ [23].

$B \to D^*h^{\pm}\pi$ decays

As described in Sec. 3.3.4, $B^{\pm} \to D^{*0}h^{\pm}\pi^0$, $\overline{(}B^0\overline{)} \to D^{*0}h^{\pm}\pi^{\mp}$ and $\overline{(}B^0\overline{)} \to D^{*\mp}h^{\pm}\pi^{\mp}$ decays are all considered. The partially reconstructed $B \to D^*h^{\pm}\pi^-$ component representing these channels is mostly given the freedom to vary freely, as can be seen in Tab. 4.1.

Hadronic parameters for these decays are unknown. The yield in the ADS D^*K subsample is therefore fixed to be $(1.5 \pm 1.0)\%$ of that in the D^*K favoured mode, and the *CP* asymmetry is fixed to 0 with an uncertainty of ± 0.5. The motivation behind fixing these parameters is to aid fit stability. The central values have been chosen to align with similar decays of the type $B^{\pm} \to D^{(*)}K^{\pm(*)}$; the uncertainties assigned are conservative. This component sits at low B mass, removed from the signal PDF, and with a negligible contribution in the Δm sample due to the applied box cuts. It therefore is

of little consequence to the signal yields and *CP* observables of this analysis, which is demonstrated by the small systematic effect of the fixed ADS parameters in Sec. 4.5.

$\Lambda_b^0 \to (\Sigma_c^+ \to (\Lambda_c^+ \to pK^-\pi^+))h^-$ decays

A PDF component for Λ_b^0 decays is necessary in the $D \to KK$ category when the pion from the $\Lambda_c^+ \to pK^-\pi^+$ decay is not reconstructed and the proton is mis-identified as a kaon. These decays are mostly given the freedom to vary freely, as can be seen in Tab. 4.1. In the $D\pi^0$ invariant-mass fit, for stability, the rate of $\Lambda_b^0 \to \Sigma_c^+\pi^-$ decays is fixed with respect to $\Lambda_b^0 \to \Sigma_c^+ K^-$ decays using the known branching ratio $\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+\pi^-)/\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ K^-) = 7.3 \pm 3.7\%$ [19], where a 50% uncertainty has been assigned.

$B_s^0 \to D^{(*)}K^+\pi^-$ decays

$B_s^0 \to D^{(*)}K^\mp\pi^\pm$ decays are not considered in the $D^*\pi$ sub-samples, as to miss a kaon would shift the invariant mass of the candidates below the lower $m(B)$ range of both data fits. The yields in the favoured mode D^*K sub-samples are also assumed to be zero, as this would correspond to suppressed D decays in the B_s^0 case. The rate of both $B_s^0 \to DK^\mp\pi^\pm$ and $B_s^0 \to D^*K^\mp\pi^\pm$ decays float freely in the D^*K ADS mode sub-samples. The GLW mode yields are fixed relative to these using the well known D branching fractions, given by Eqs. (3.7) to (3.9). Again, due to the time integrated nature of the analysis, B_s^0 decays are assumed to exhibit no CP violation.

Mis-identified decays

Companion misidentification is handled in the fit using mis-identified PDFs, where the relative size of these compared to correctly identified components are fixed using measured PID and selection efficiencies.

Combinatorial

The combinatorial component is assumed to exhibit zero *CP* asymmetry, which is handled as a source of systematic uncertainty in Sec. 4.5. Combinatorial yields

float in the $D^*\pi$ favoured mode, then are fixed relative to this in all other modes using the measured yields in the partially reconstructed $B^\pm \to D^{(*)}h^\pm$ analysis [31]. These fixed ratios are varied as a source of systematic uncertainty, described in Sec. 4.5. The assumption made here is that the fake companion background level across decay modes is similar for the partially and fully reconstructed analyses. The favoured mode combinatorial yield in the $D\pi^0$ ($D\gamma$) invariant-mass fit is found to be 75 ± 18 (76 ± 26) candidates.

Crossfeed

Crossfeed from favoured $B^\pm \to (D^* \to [K^\pm\pi^\mp]_D\pi^0/\gamma)h^\pm$ decays enter the ADS $D^*\pi$ and D^*K sub-samples when both D decay products are misidentified. All signal modes are assigned a crossfeed component. The crossfeed rate is taken from the partially reconstructed analysis [31] and is fixed in the invariant-mass fits.

4.2 Raw fit results split by charge

The results of the invariant-mass fits to 2011-2018 data, split by B charge, will now be given. Both the $D\pi^0$ and $D\gamma$ fits converge to stable minima with well behaved covariance matrices (in `ROOFIT` language, this equates to the flags: `MIGRAD` $= 0$ and `HESSE` $= 0$, and estimated distance to minimum values: `EDM` $= 6.32 \times 10^{-4}$ and 4.70×10^{-4} , respectively). The results are illustrated in Figs. 4.2 to 4.17; each D decay mode is displayed separately, and a legend is provided in Fig. 4.1.

The *raw* numerical `RooFit` results of the split by charge $D\pi^0$ fit are given below. Here, raw implies statistical uncertainties taken straight from `RooFit`:

$$
\begin{aligned}
R_{K^-}^{\pi K,\pi^0} &= 0.0215 \pm 0.0055 \\
R_{K^+}^{\pi K,\pi^0} &= 0.0077 \pm 0.0046 \\
R_{\pi^-}^{\pi K,\pi^0} &= 0.0039 \pm 0.0006 \\
R_{\pi^+}^{\pi K,\pi^0} &= 0.0044 \pm 0.0006 \\
R^{CP,\pi^0} &= 1.0692 \pm 0.0665
\end{aligned}
$$

$$
\begin{aligned}
R_{K/\pi}^{K\pi,\pi^0} &= \quad 0.0732 \pm 0.0018 \\
A_K^{K\pi,\pi^0} &= -0.0031 \pm 0.0240 \\
A_K^{CP,\pi^0} &= -0.1804 \pm 0.0497 \\
A_\pi^{CP,\pi^0} &= \quad 0.0195 \pm 0.0143.
\end{aligned}
$$

The raw `RooFit` results of the split by charge $D\gamma$ fit are as follows:

$$
\begin{aligned}
R_{K^-}^{\pi K,\gamma} &= \quad 0.0086 \pm 0.0039 \\
R_{K^+}^{\pi K,\gamma} &= \quad 0.0414 \pm 0.0049 \\
R_{\pi^-}^{\pi K,\gamma} &= \quad 0.0056 \pm 0.0005 \\
R_{\pi^+}^{\pi K,\gamma} &= \quad 0.0032 \pm 0.0005 \\
R_{K^-}^{\pi K,\pi^0} &= \quad 0.0333 \pm 0.0059 \\
R_{K^+}^{\pi K,\pi^0} &= -0.0001 \pm 0.0041 \\
R_{\pi^-}^{\pi K,\pi^0} &= \quad 0.0033 \pm 0.0006 \\
R_{\pi^+}^{\pi K,\pi^0} &= \quad 0.0051 \pm 0.0007 \\
R^{CP,\gamma} &= \quad 0.9357 \pm 0.0358 \\
R^{CP,\pi^0} &= \quad 1.0586 \pm 0.0548 \\
R_{K/\pi}^{K\pi,\gamma} &= \quad 0.0708 \pm 0.0010 \\
R_{K/\pi}^{K\pi,\pi^0} &= \quad 0.0723 \pm 0.0013 \\
A_K^{K\pi,\gamma} &= -0.0354 \pm 0.0130 \\
A_K^{K\pi,\pi^0} &= \quad 0.0563 \pm 0.0161 \\
A_K^{CP,\gamma} &= \quad 0.1525 \pm 0.0327 \\
A_\pi^{CP,\gamma} &= -0.0126 \pm 0.0085 \\
A_K^{CP,\pi^0} &= -0.1922 \pm 0.0387 \\
A_\pi^{CP,\pi^0} &= \quad 0.0083 \pm 0.0120.
\end{aligned}
$$

The quoted uncertainties are statistical, and will be corrected for double-counting using the procedure described in Sec. 4.4. Systematic uncertainties are evaluated in Sec. 4.5, and full numerical results given in Sec. 4.6.

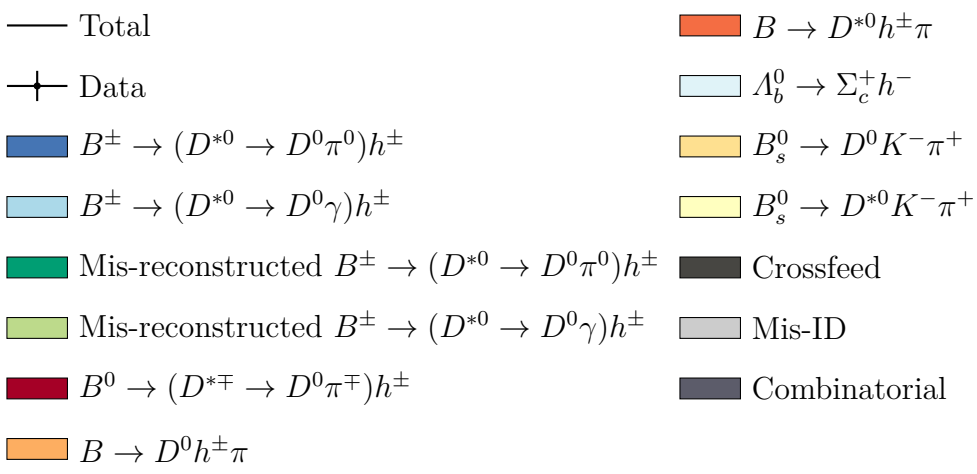


Figure 4.1: Invariant-mass fit legend, linking the fit components to their associated colours.

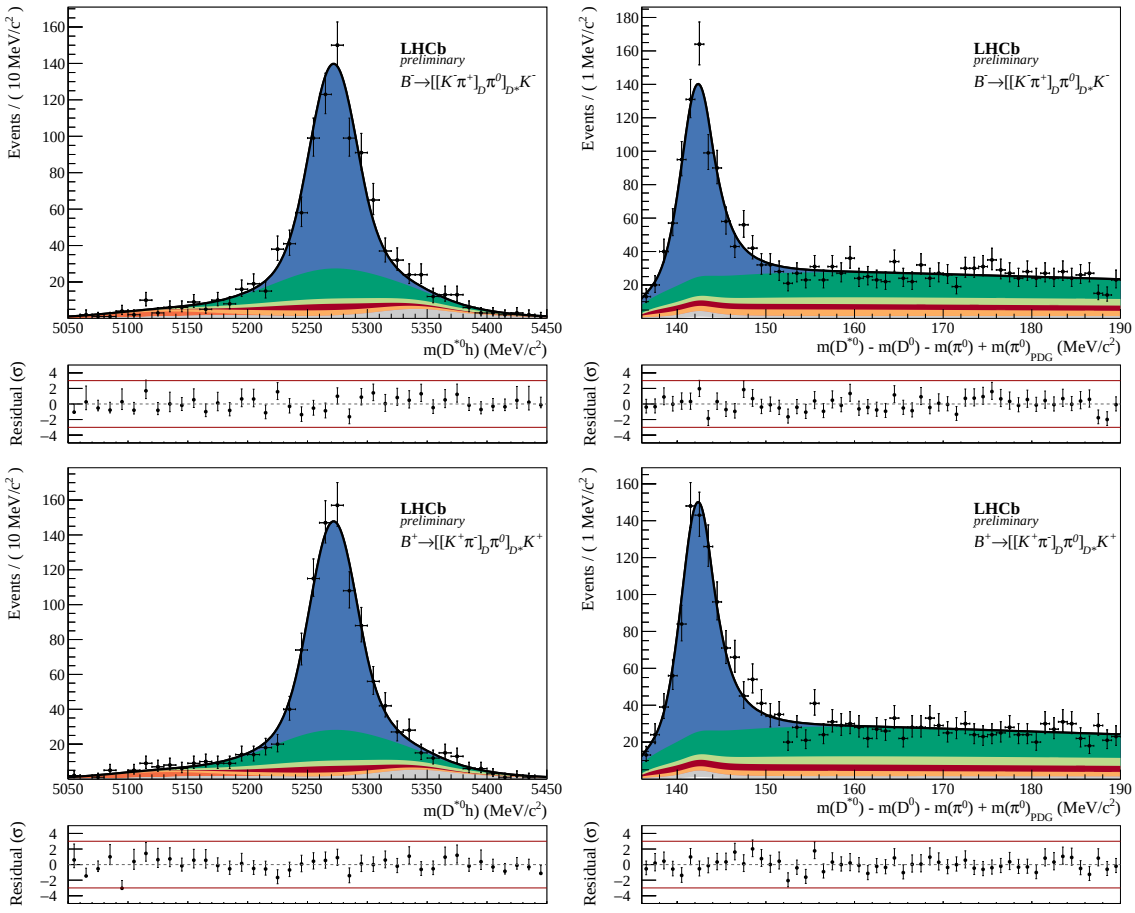


Figure 4.2: Invariant mass fit to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)K^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

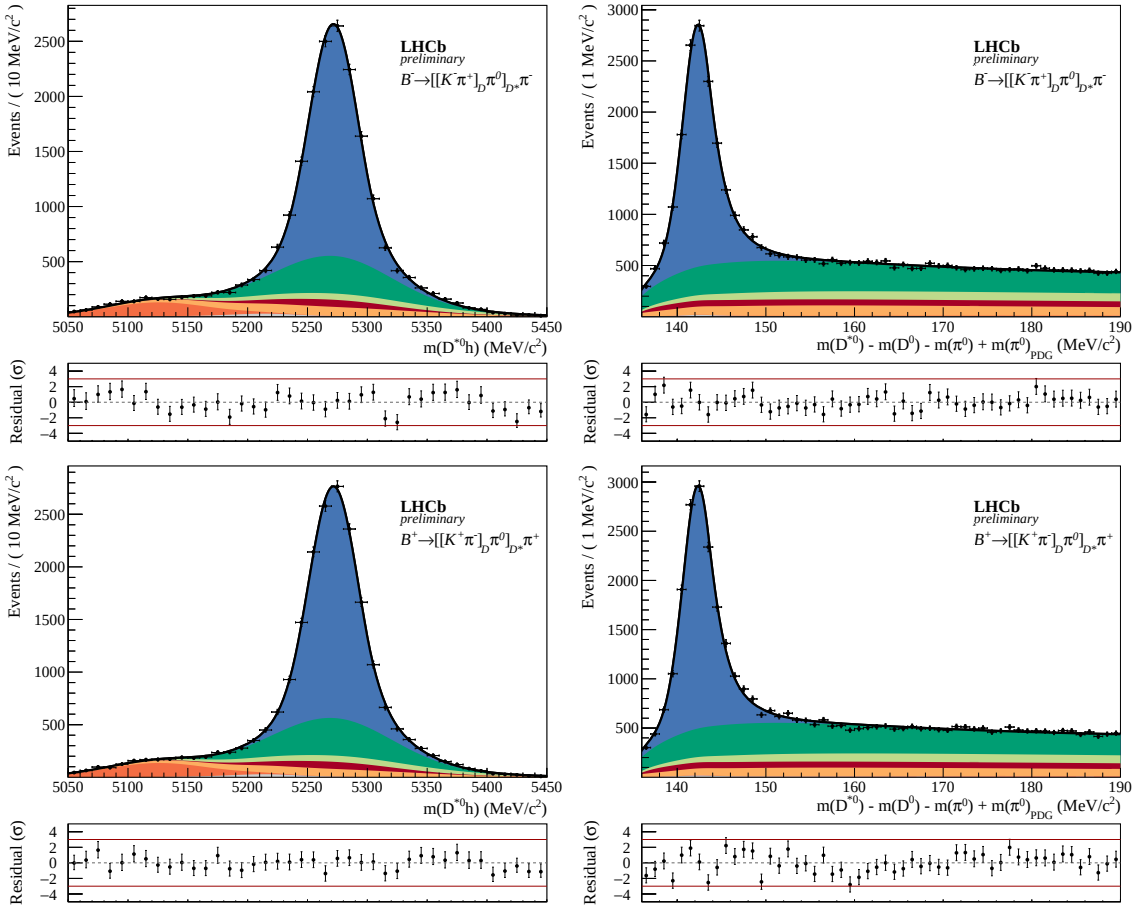


Figure 4.3: Invariant mass fit to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

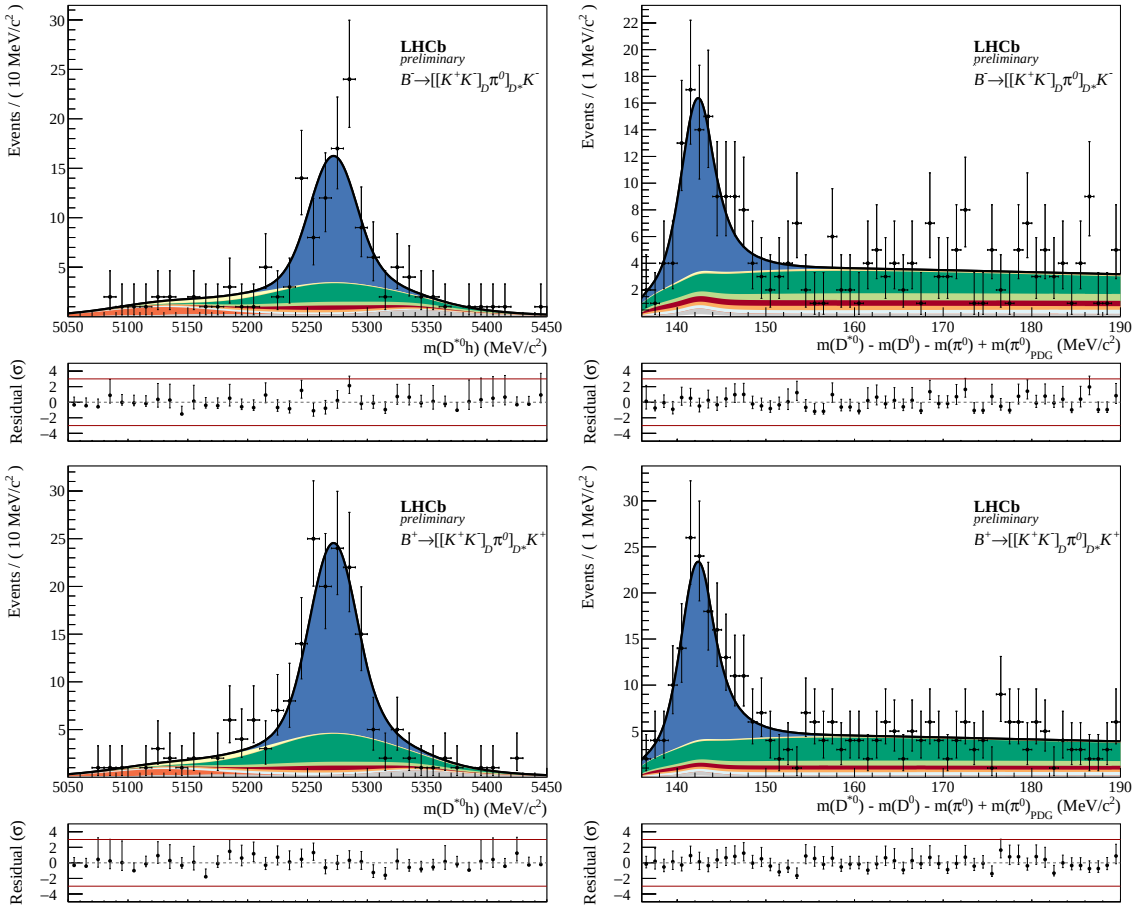


Figure 4.4: Invariant mass fit to $B^{\pm} \to (D^* \to [K^{\pm}K^{\mp}]_D\pi^0)K^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

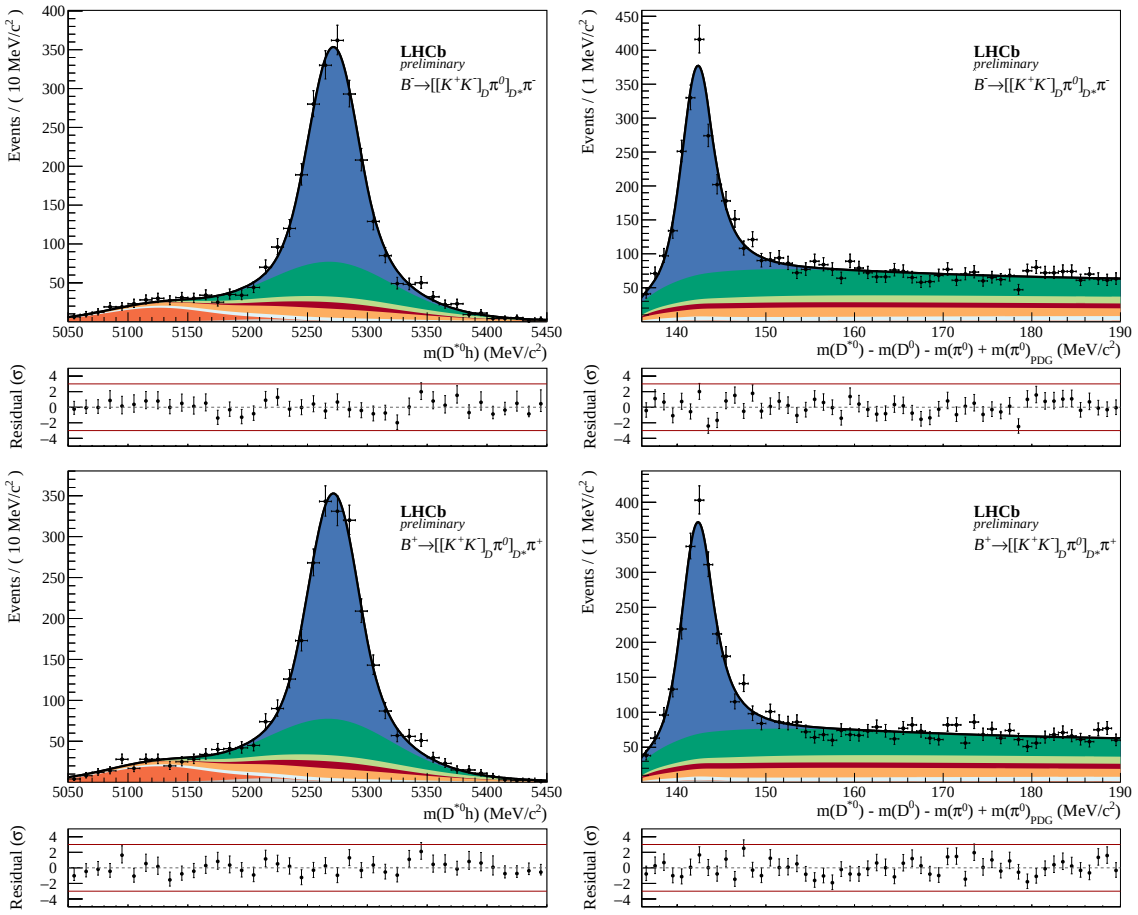


Figure 4.5: Invariant mass fit to $B^{\pm} \to (D^* \to [K^{\pm}K^{\mp}]_D\pi^0)\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

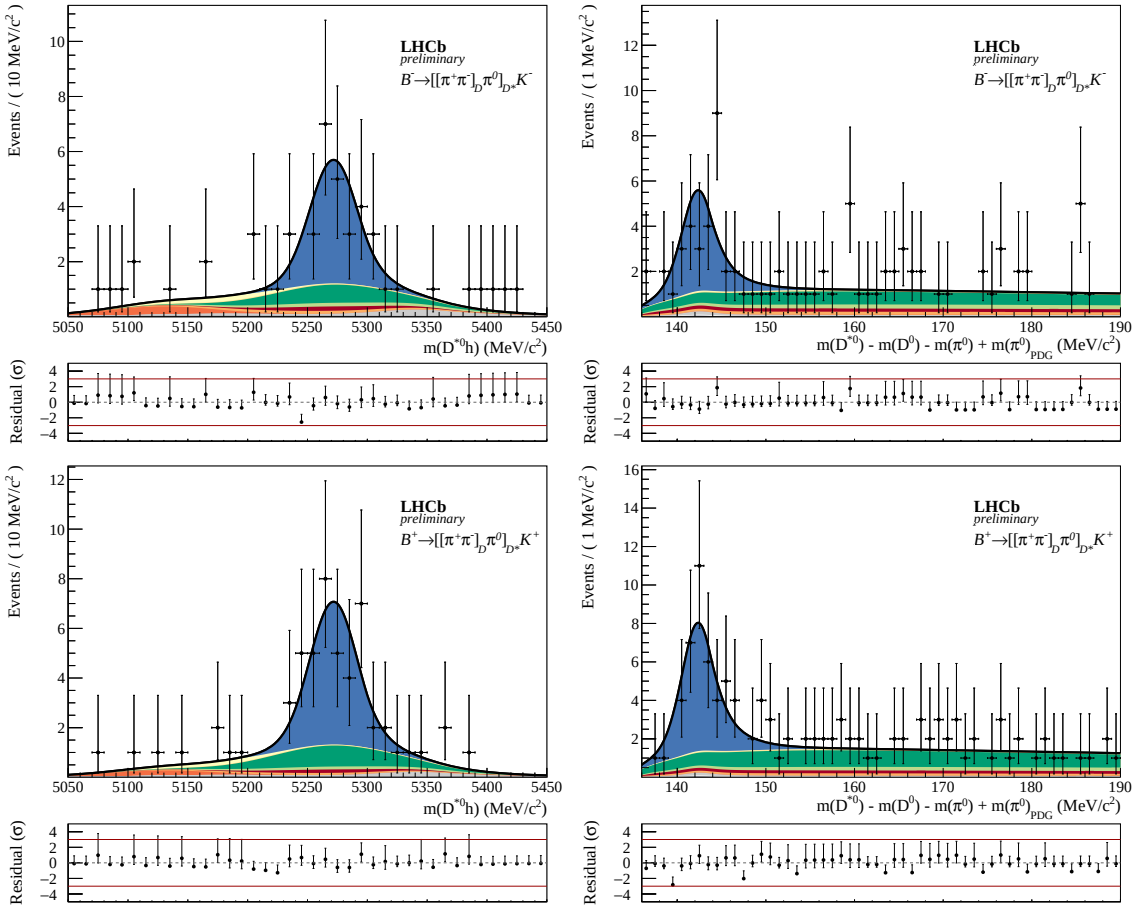


Figure 4.6: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}\pi^{\mp}]_D\pi^0)K^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

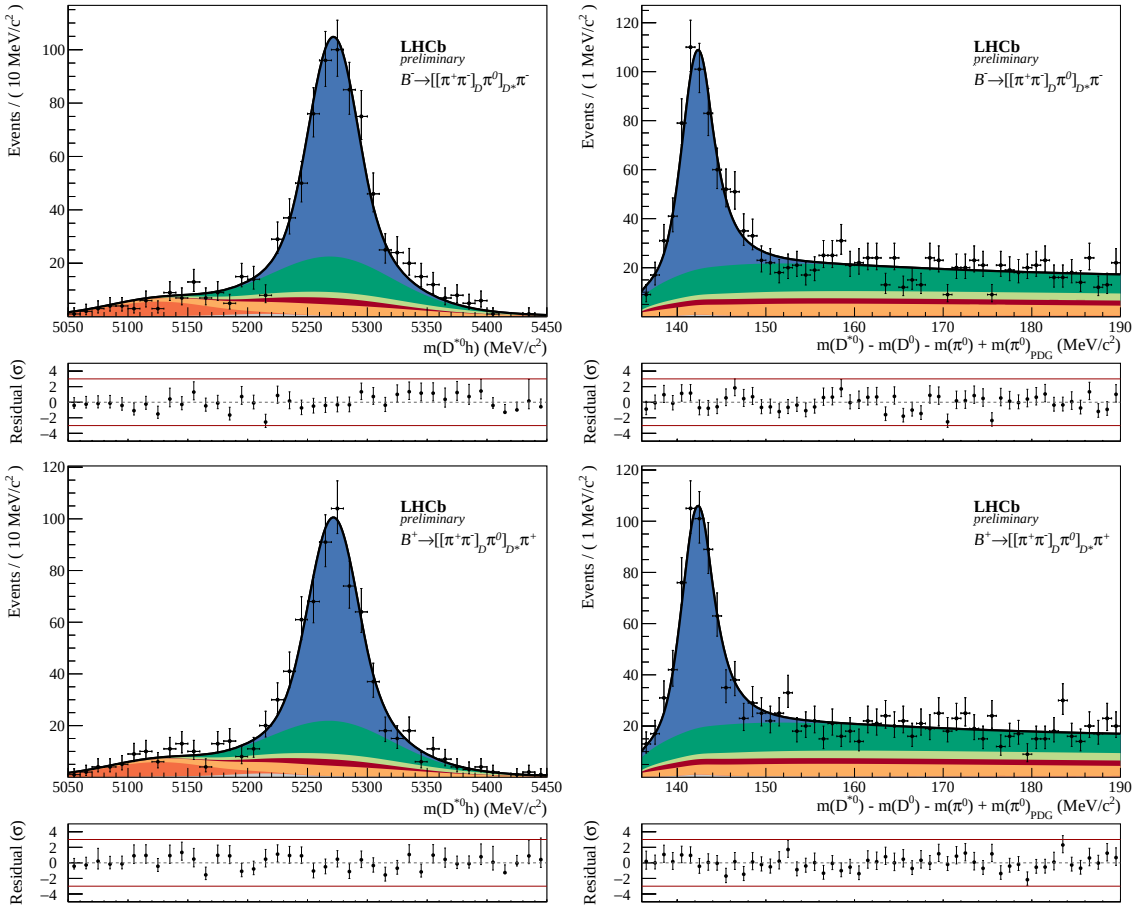


Figure 4.7: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}\pi^{\mp}]_D\pi^0)\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

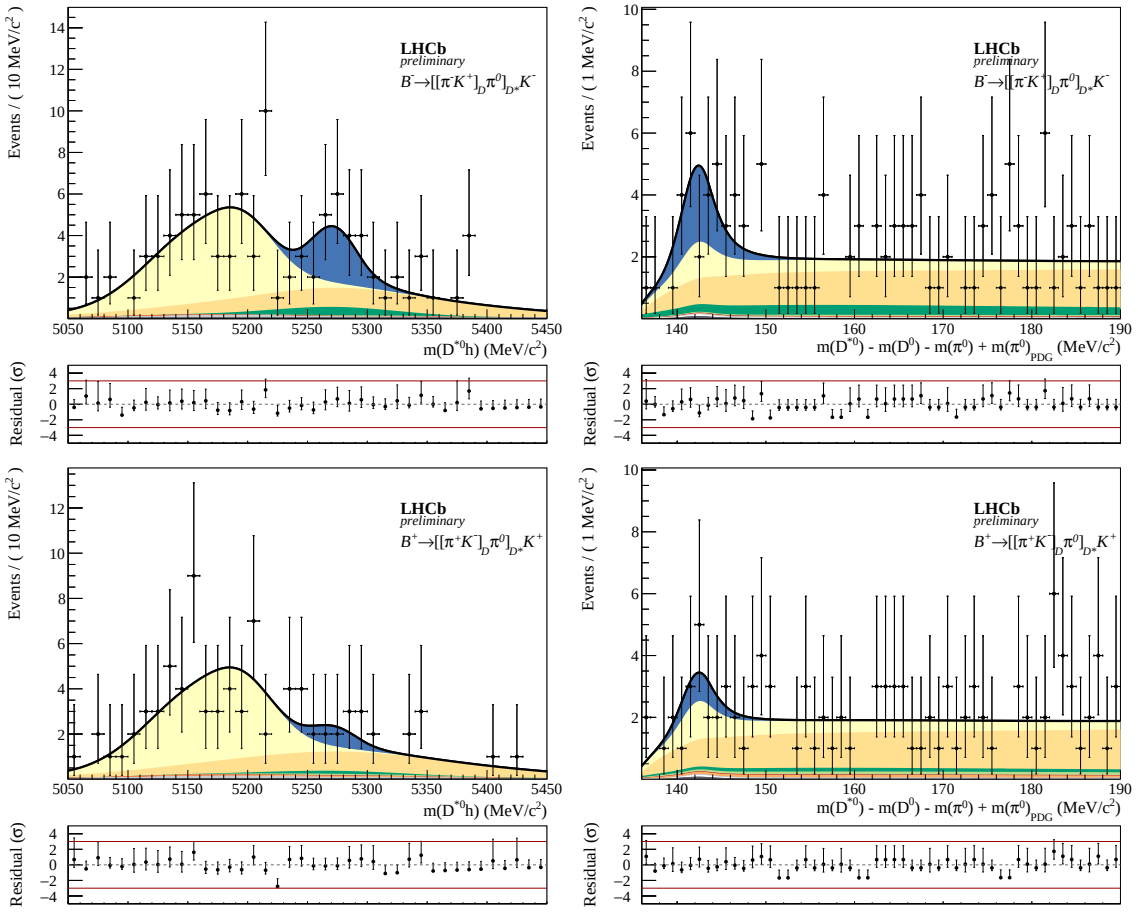


Figure 4.8: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\pi^0)K^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

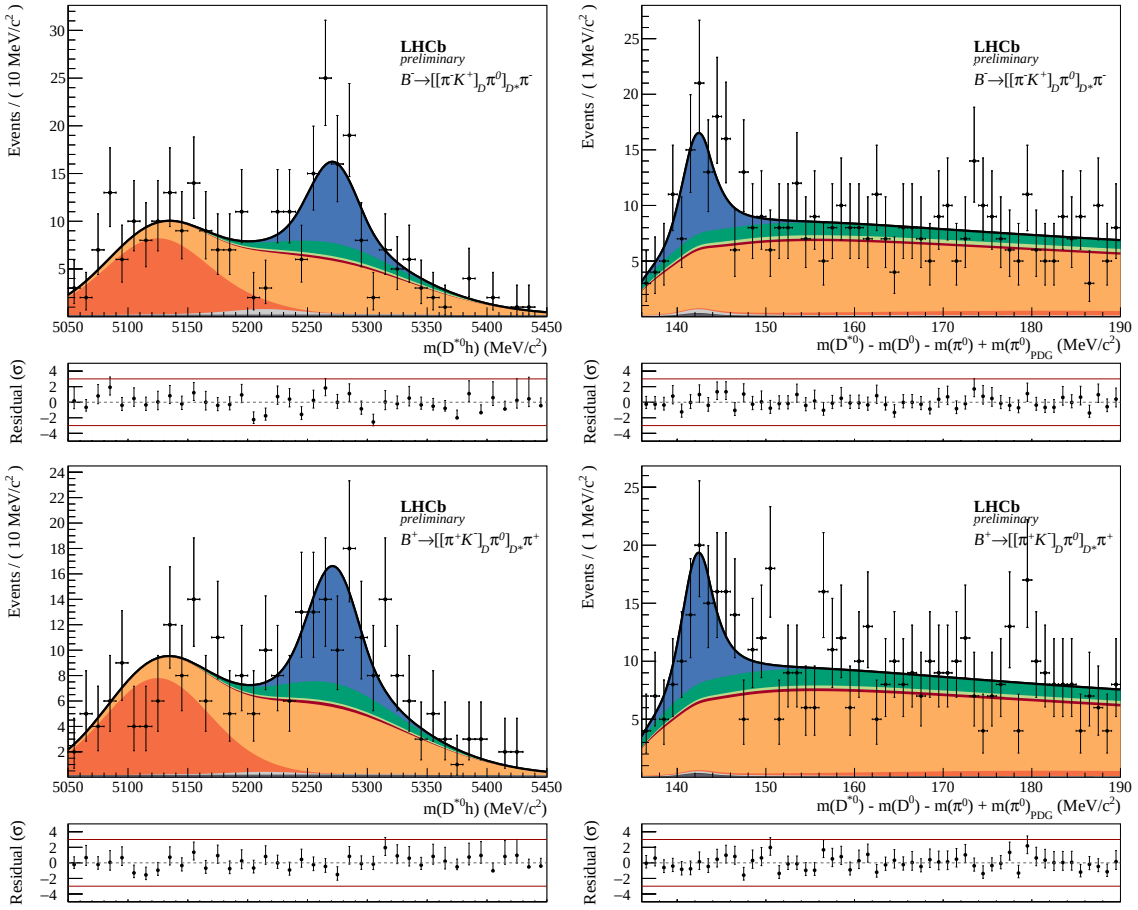


Figure 4.9: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_D\pi^0)\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across $m(B)$ (left) and Δm (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

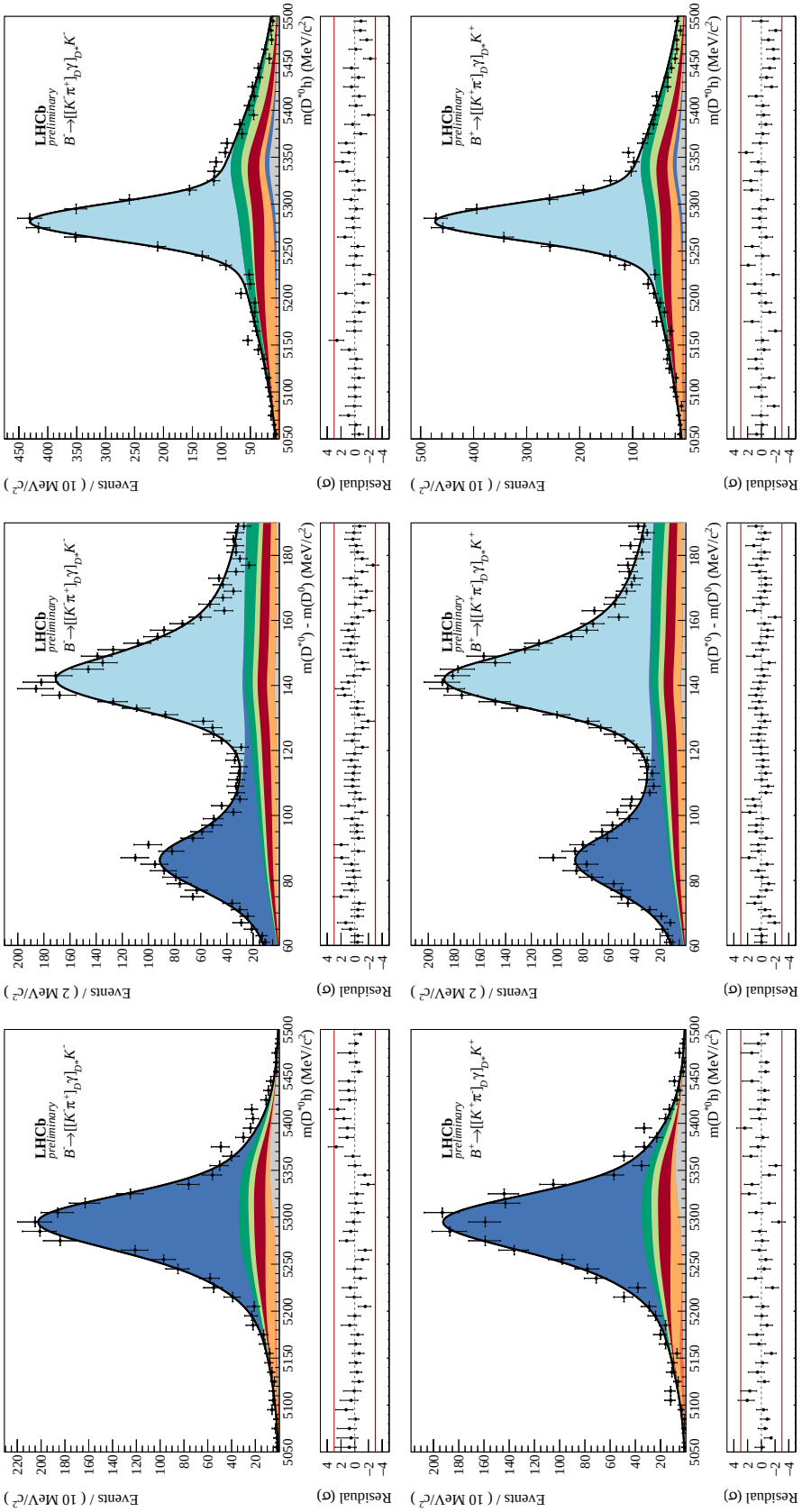


Figure 4.10: Invariant mass fit to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_{D\gamma})K^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

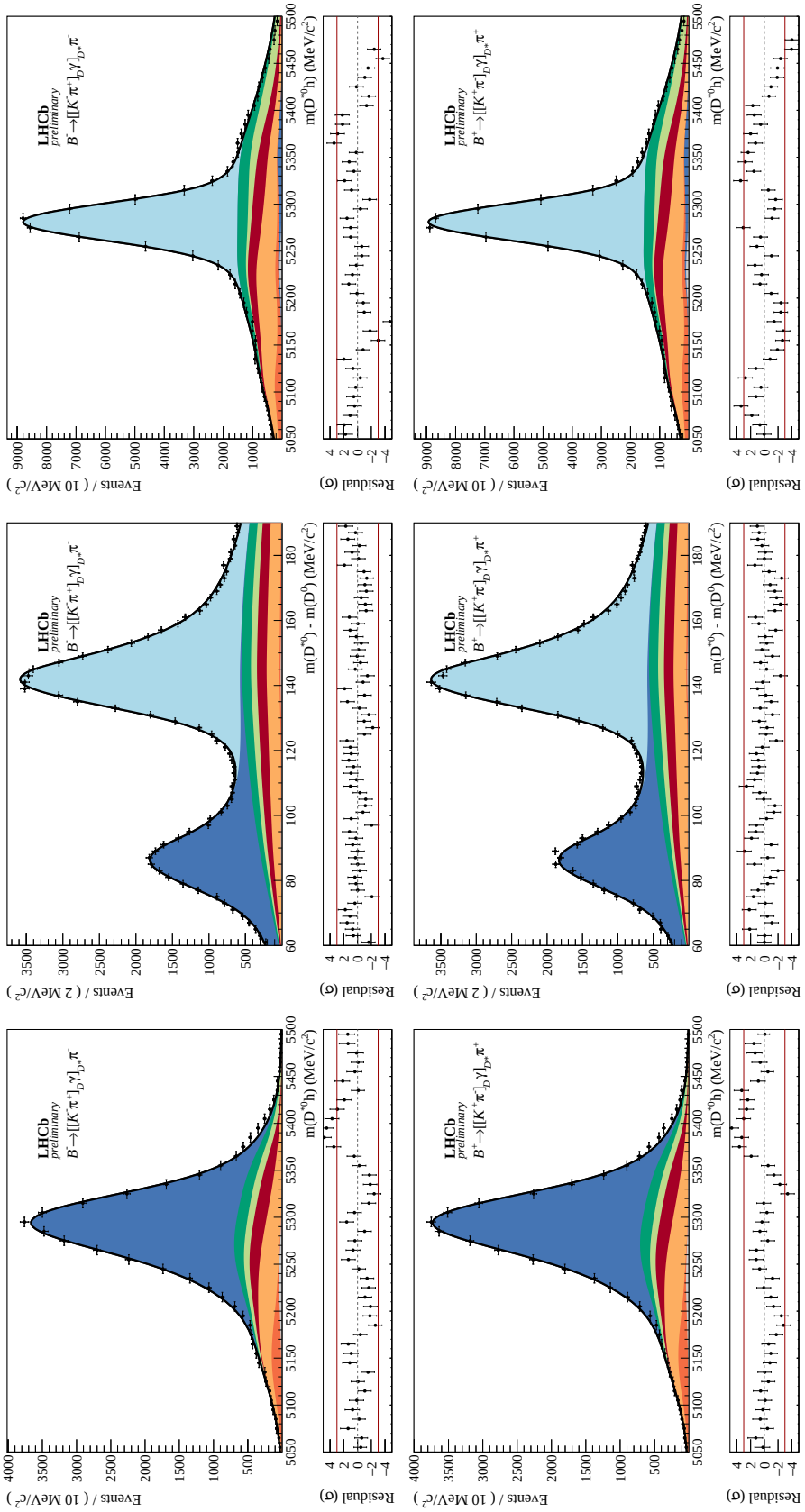


Figure 4.11: Invariant mass fit to $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_{D\gamma})\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

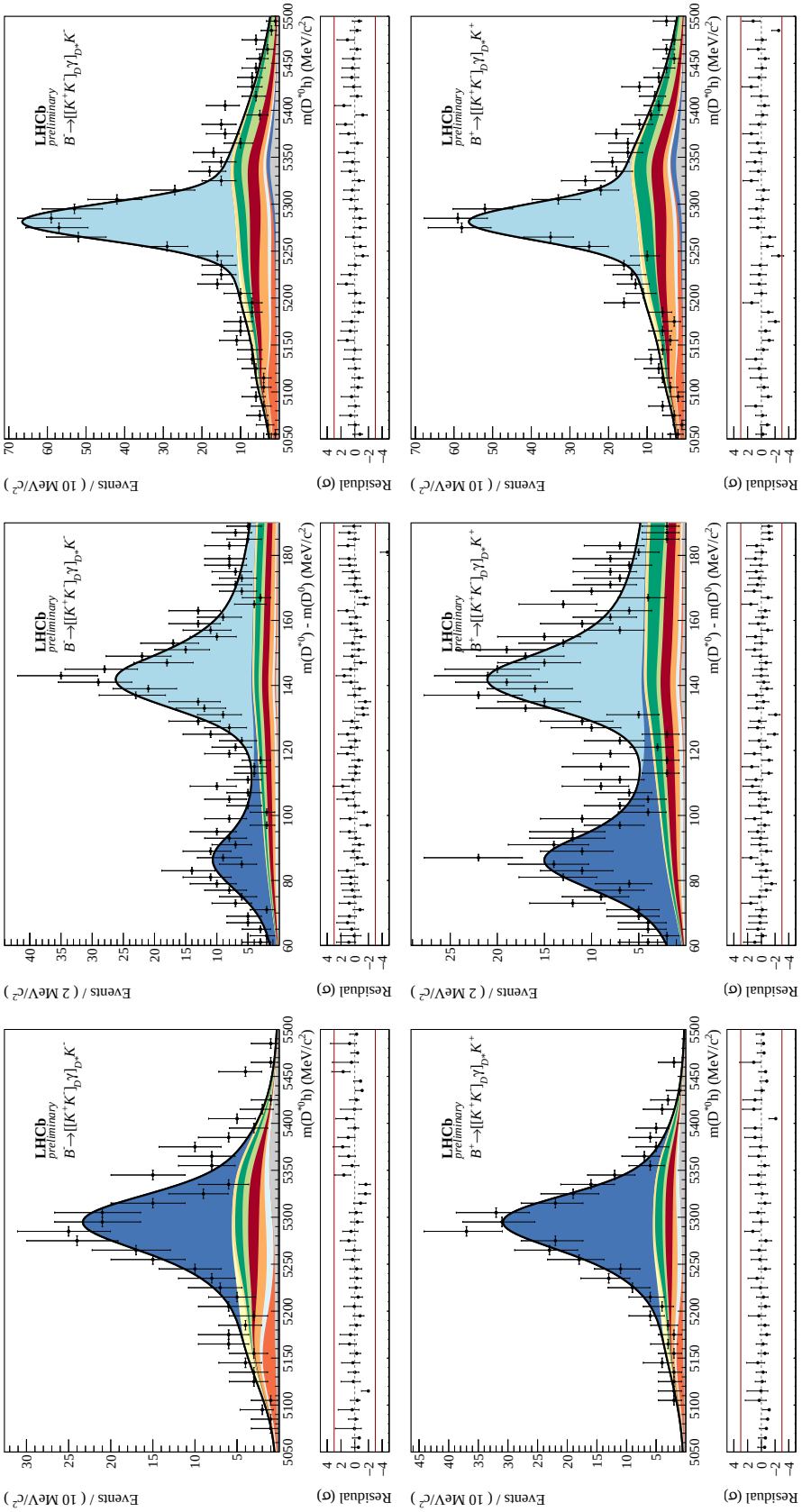


Figure 4.12: Invariant mass fit to $B^\pm \to (D^* \to [K^\pm K^\mp]_D\gamma)K^\pm$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

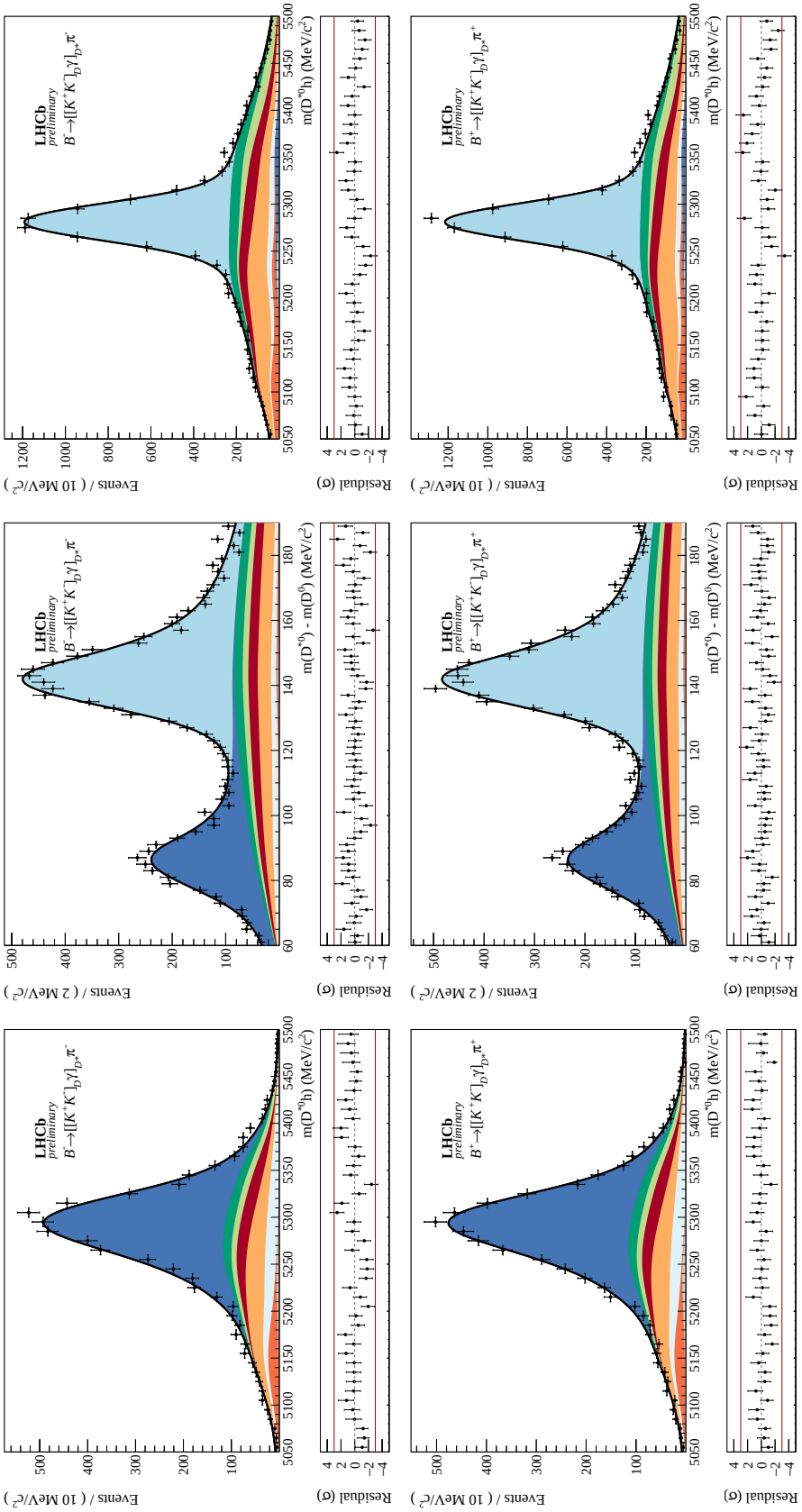


Figure 4.13: Invariant mass fit to $B^{\pm} \to (D^* \to [K^{\pm}K^{\mp}]_{D\gamma})\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

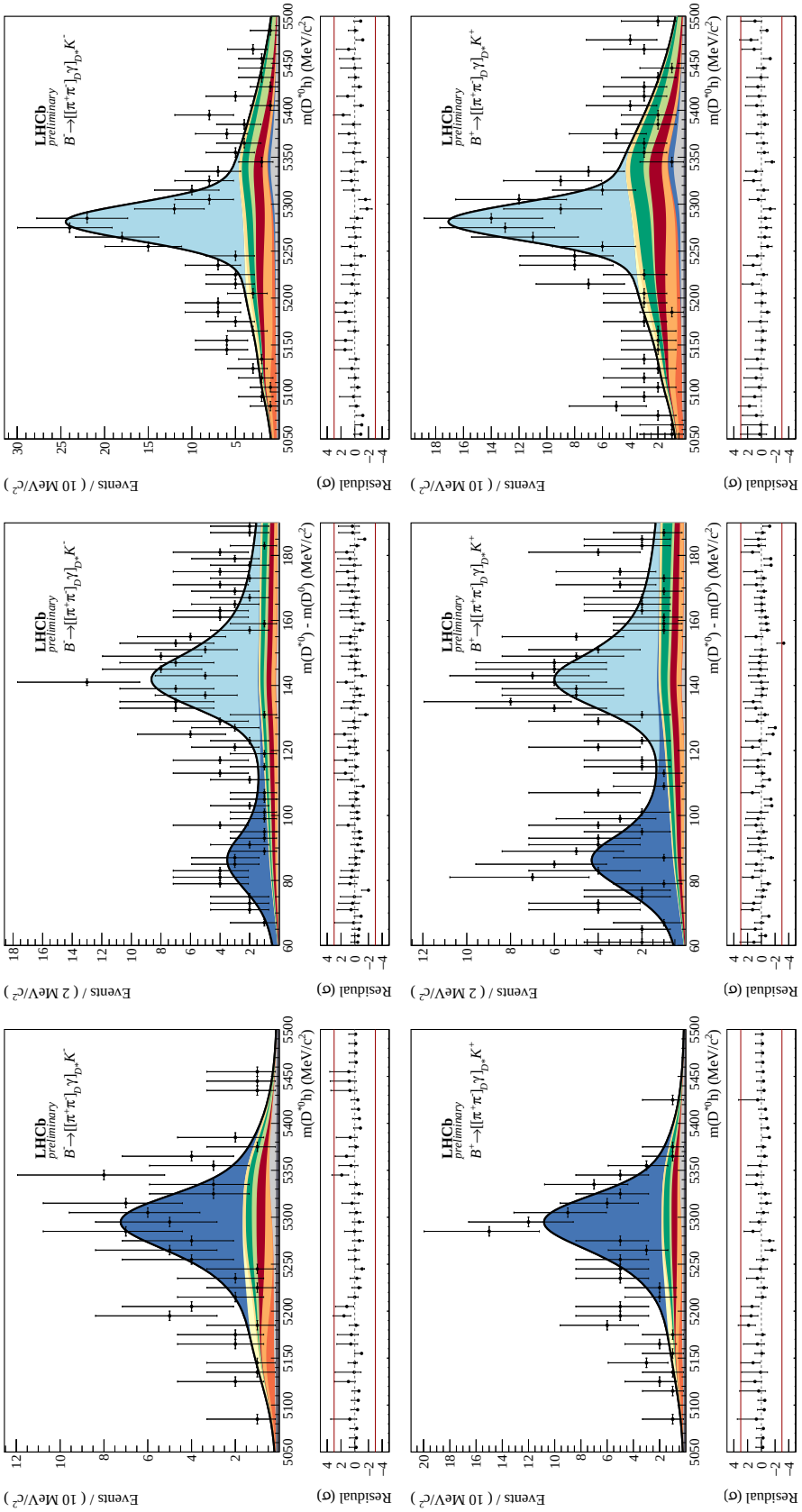


Figure 4.14: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}\pi^{\mp}]_{D\gamma})K^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

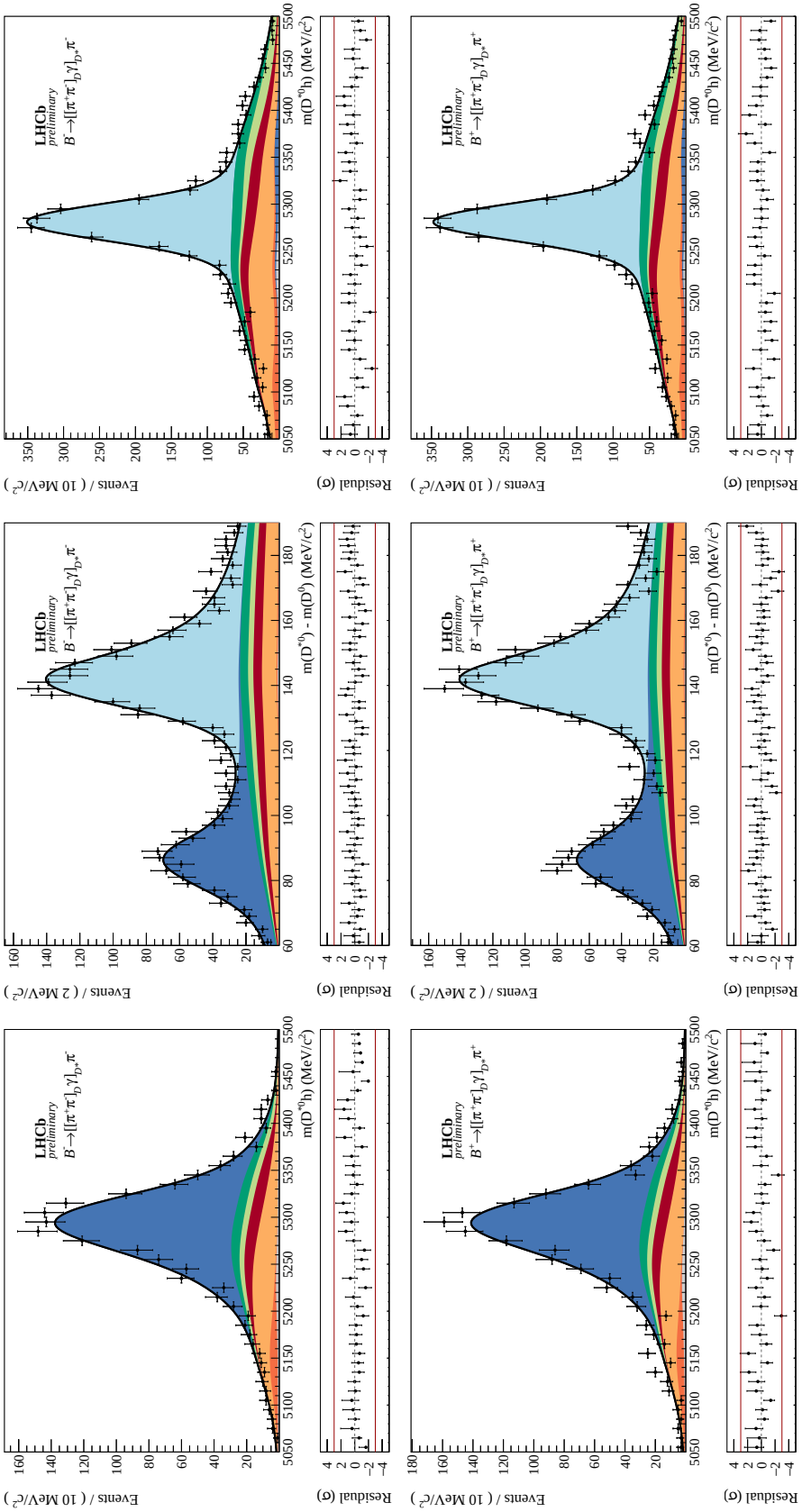


Figure 4.15: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

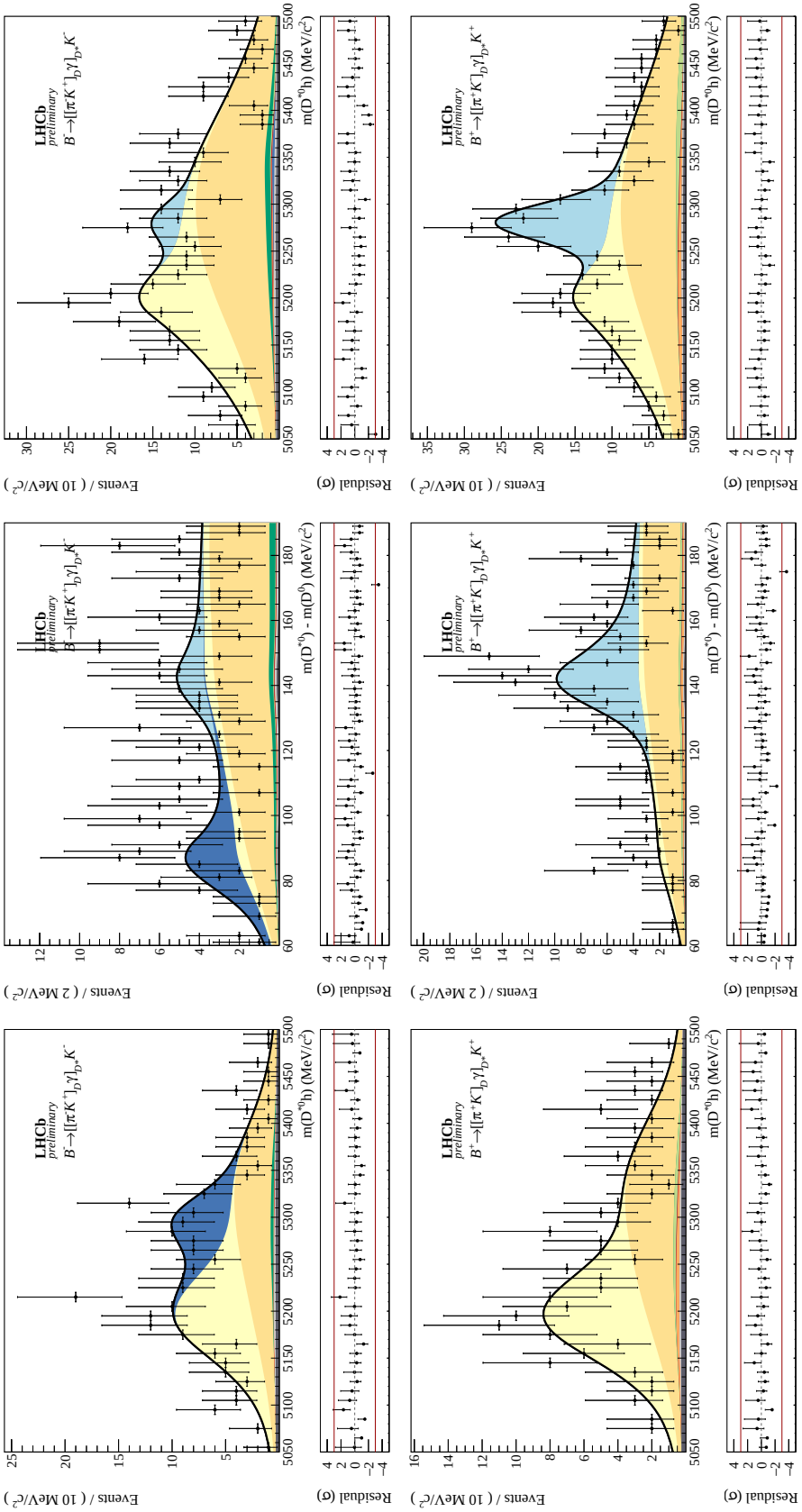


Figure 4.16: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_{D\gamma})K^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

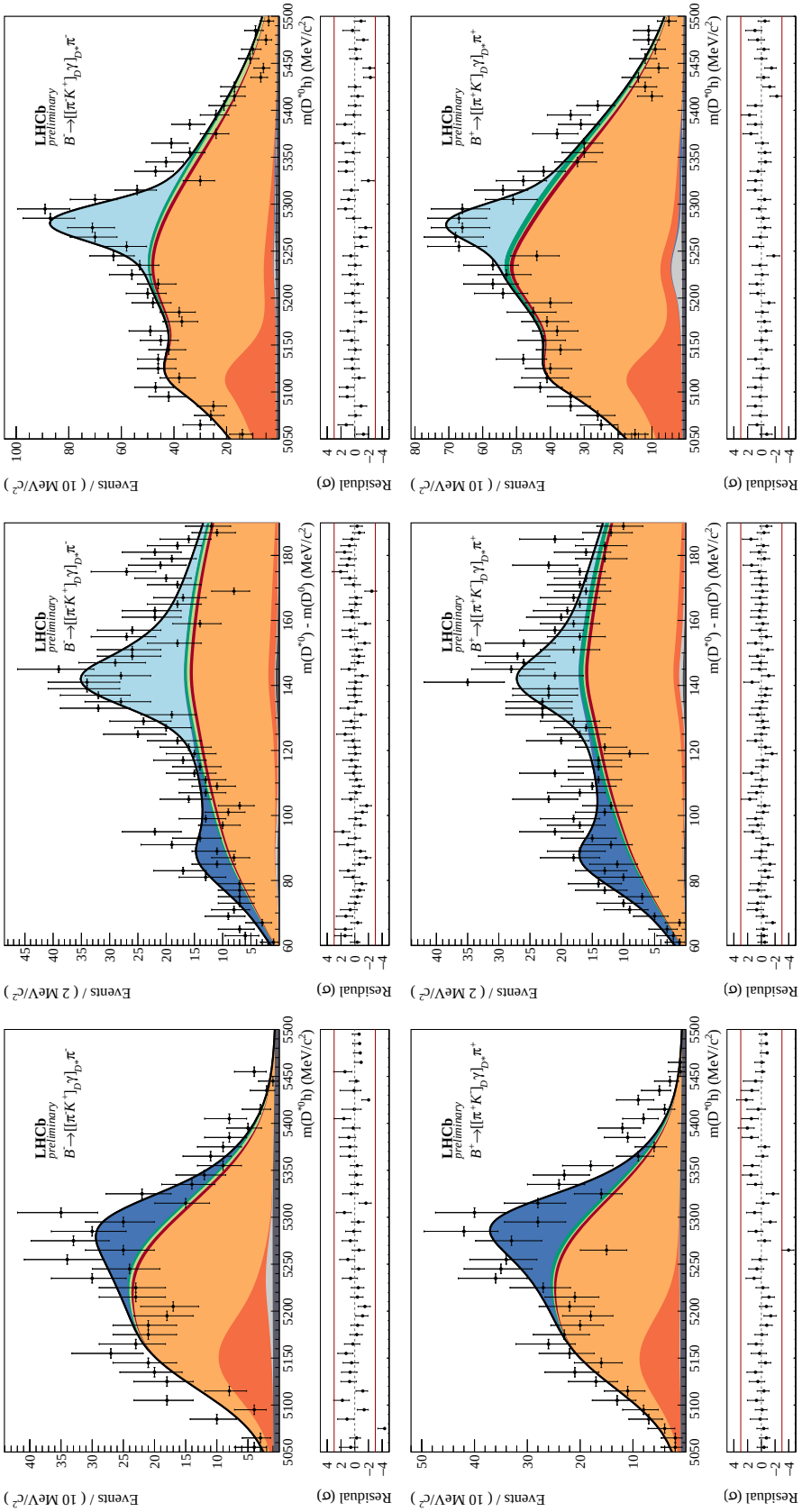


Figure 4.17: Invariant mass fit to $B^{\pm} \to (D^* \to [\pi^{\pm}K^{\mp}]_{D\gamma})\pi^{\pm}$ candidates in 2011-2018 data, split by B charge. The fit is simultaneous across the partially reconstructed $D\pi^0$ slice of $m(B)$ (left), Δm (middle) and the fully reconstructed $D\gamma$ slice of $m(B)$ (right). B^- (B^+) meson candidates are displayed on the top (bottom). A legend is given in Fig. 4.1.

4.2.1 Signal Yields

The summed over charge yields for each signal mode, measured by the $D\pi^0$ and $D\gamma$ data fits, are given in Tabs. 4.2 and 4.3, respectively. The quoted uncertainties are statistical and have been corrected for double counting using the method described in Sec. 4.4.

Decay	D mode	Yield
$B^\pm \to (D\pi^0)_{D^*}\pi^\pm$	$K\pi$	25825 ± 218
	KK	3273 ± 62
	$\pi\pi$	956 ± 32
	πK	108 ± 14
$B^\pm \to (D\pi^0)_{D^*}K^\pm$	$K\pi$	1889 ± 60
	KK	256 ± 22
	$\pi\pi$	75 ± 7
	πK	28 ± 9

Table 4.2: Signal yields, summed over B charge, measured by the $D\pi^0$ data fit.

Decay	D mode	Yield
$B^\pm \to (D\gamma)_{D^*}\pi^\pm$	$K\pi$	82984 ± 406
	KK	10874 ± 121
	$\pi\pi$	3220 ± 57
	πK	368 ± 38
$B^\pm \to (D\gamma)_{D^*}K^\pm$	$K\pi$	5875 ± 104
	KK	720 ± 37
	$\pi\pi$	213 ± 11
	πK	147 ± 23
$B^\pm \to (D\pi^0)_{D^*}\pi^\pm$	$K\pi$	58598 ± 373
	KK	7279 ± 135
	$\pi\pi$	2182 ± 54
	πK	245 ± 33
$B^\pm \to (D\pi^0)_{D^*}K^\pm$	$K\pi$	4236 ± 100
	KK	557 ± 38
	$\pi\pi$	167 ± 12
	πK	70 ± 19

Table 4.3: Signal yields, summed over B charge, measured by the $D\gamma$ data fit.

4.3 Fit validation with toys

The performance and stability of both the $D\pi^0$ and $D\gamma$ invariant-mass fits are evaluated using toy experiments. For each D^* final state, the total simultaneous PDF is first fit to data, then many toy datasets are generated from this PDF; these datasets are therefore created with all floating parameters initialised to the *final* values of the data fit. Fits are performed to these datasets, where all PDF parameters are initialised to the same *initial* values used in the data fit. Collecting the results of these toy experiments, pull distributions can be plotted for each *CP* observable:

$$\mathcal{P} = \begin{cases} \dfrac{\mathcal{O}_{gen} - \mathcal{O}_{toy}}{\Delta^{+}_{toy}}, & \text{if } \mathcal{O}_{toy} \leq \mathcal{O}_{gen} \\ \dfrac{\mathcal{O}_{toy} - \mathcal{O}_{gen}}{\Delta^{-}_{toy}}, & \text{if } \mathcal{O}_{toy} > \mathcal{O}_{gen}. \end{cases} \tag{4.29}$$

Here, $\mathcal{O}_{toy}$ is the final value of the observable found by the toy fit; $\mathcal{O}_{gen}$ is the value the toy dataset was generated with (*i.e.* the final value of the fit to data); $\Delta^{+(-)}_{toy}$ is the positive (negative) uncertainty on $\mathcal{O}$ measured by the toy fit.

For a given CP observable, a $\mathcal{P}$ distribution with mean consistent with 0 demonstrates that the fit returns an unbiased estimate of this observable; a $\mathcal{P}$ distribution with width consistent with 1 demonstrates that the fit returns the correct associated statistical uncertainty on this parameter. The formula for calculating pulls with asymmetric errors was derived in Ref. [92]; these are used as slight non-parabolic behaviour of the likelihood profile at low event yields in the D^*K ADS mode shows up as asymmetry in $\mathcal{P}$ if not accounted for. The difference between raw positive and negative errors on $R^{\pi K}_{K^\pm}$ for both the $D\pi^0$ and $D\gamma$ signal extraction fits is $< 5\%$, therefore single double-sided errors are quoted in the final results.

2500 toy datasets have been generated for the $D\pi^0$ and $D\gamma$ D1D PDFs separately, and the pull distributions for each *CP* observable fit to using a Gaussian function with freely varying mean and width; the results are provided in Tabs. 4.4 and 4.5, respectively. These demonstrate that both invariant-mass fits are able to measure all parameters without significant biases to either central values or statistical uncertainties.

Observable	Pull μ	Pull σ
$A_K^{K\pi,\pi^0}$	0.03 ± 0.02	0.98 ± 0.01
A_K^{CP,π^0}	-0.01 ± 0.02	0.98 ± 0.01
A_π^{CP,π^0}	0.03 ± 0.02	1.00 ± 0.02
R^{CP,π^0}	-0.04 ± 0.02	1.00 ± 0.02
$R_{K/\pi}^{K\pi,\pi^0}$	-0.08 ± 0.02	0.96 ± 0.01
$R_{K^-}^{\pi K,\pi^0}$	-0.06 ± 0.02	1.01 ± 0.02
$R_{K^+}^{\pi K,\pi^0}$	-0.09 ± 0.02	1.01 ± 0.02
$R_{\pi^-}^{\pi K,\pi^0}$	-0.04 ± 0.02	1.00 ± 0.02
$R_{\pi^+}^{\pi K,\pi^0}$	-0.01 ± 0.02	0.98 ± 0.01

Table 4.4: Measured pull means and widths for each *CP* observable in the $D\pi^0$ invariant-mass fit.

Observable	Pull μ	Pull σ
$A_K^{K\pi,\gamma}$	-0.01 ± 0.02	0.99 ± 0.01
$A_K^{K\pi,\pi^0}$	-0.02 ± 0.02	1.02 ± 0.01
$A_K^{CP,\gamma}$	0.00 ± 0.02	1.01 ± 0.01
$A_\pi^{CP,\gamma}$	0.02 ± 0.02	1.00 ± 0.01
A_K^{CP,π^0}	0.01 ± 0.02	1.01 ± 0.01
A_π^{CP,π^0}	0.01 ± 0.02	1.01 ± 0.01
$R^{CP,\gamma}$	-0.04 ± 0.02	0.98 ± 0.01
R^{CP,π^0}	0.01 ± 0.02	0.99 ± 0.01
$R_{K/\pi}^{K\pi,\gamma}$	0.02 ± 0.02	0.97 ± 0.01
$R_{K/\pi}^{K\pi,\pi^0}$	-0.04 ± 0.02	1.01 ± 0.01
$R_{K^-}^{\pi K,\gamma}$	-0.01 ± 0.02	1.01 ± 0.01
$R_{K^+}^{\pi K,\gamma}$	0.03 ± 0.02	1.00 ± 0.01
$R_{\pi^-}^{\pi K,\gamma}$	-0.03 ± 0.02	1.00 ± 0.01
$R_{\pi^+}^{\pi K,\gamma}$	-0.00 ± 0.02	0.99 ± 0.01
$R_{K^-}^{\pi K,\pi^0}$	-0.03 ± 0.02	1.03 ± 0.01
$R_{K^+}^{\pi K,\pi^0}$	-0.06 ± 0.02	0.99 ± 0.01
$R_{\pi^-}^{\pi K,\pi^0}$	0.01 ± 0.02	1.00 ± 0.01
$R_{\pi^+}^{\pi K,\pi^0}$	0.02 ± 0.02	0.99 ± 0.01

Table 4.5: Measured pull means and widths for each *CP* observable in the $D\gamma$ invariant-mass fit.

4.4 Correction of statistical uncertainties

The invariant-mass fit procedure described in Sec. 3.2 is performed simultaneously across 2 or 3 samples of $m(B)$ and Δm, which overlap around the signal peaks. The raw fit errors on the observables given in Sec. 4.2 do not account for the double counting of candidates that have been included in more than one sample,[1] therefore a bootstrapping procedure is employed to calculate the true statistical error on each observable. This method is outlined in the section below.

4.4.1 Bootstrapping method

Toy datasets generated to validate the fit models in Sec. 4.3 are Poisson fluctuated independently within each mass sample. In order to assess the effect of double counting, the fluctuation of events in the $m(B)$ and Δm samples must be correlated within the box region. For this purpose, toy datasets are generated across the full 2D $m(B) - \Delta m$ plane using the *bootstrapping* method [93], whereby events from the original data sample of size N (*i.e.* the LHCb data) are randomly sampled N times with replacement. These 2D bootstrap datasets are then sliced up using the box cuts described in Sec. 3.2.2, and are fit to with the total simultaneous D1D PDF (where all parameters are initialised to their initial values in the data fit).

Fits are performed to 2500 datasets generated in this way, for each D^* final state, and the distribution of measurements for each observable is fit to using a Gaussian function with freely varying mean and width. The distributions are shown for $B^\pm \to (D^* \to D\gamma)h^\pm$ mode observables reconstructed in the $D\gamma$ final state in Fig. 4.18; the other bootstrap distributions can be found in Appendix A. It can be seen that the set of measurements for each observable is centered on the central value of the nominal fit result; the spread of measurements for each observable, however, is larger than the associated statistical uncertainty listed in Sec. 4.2. This is because the bootstrap distributions encompass the effect of double counting, where as the

[1] Note that although the $D\gamma$ invariant-mass fit is performed simultaneously across Δm and two phase space regions of $m(B)$, the latter do not overlap, so there is no 'triple counting' of events.

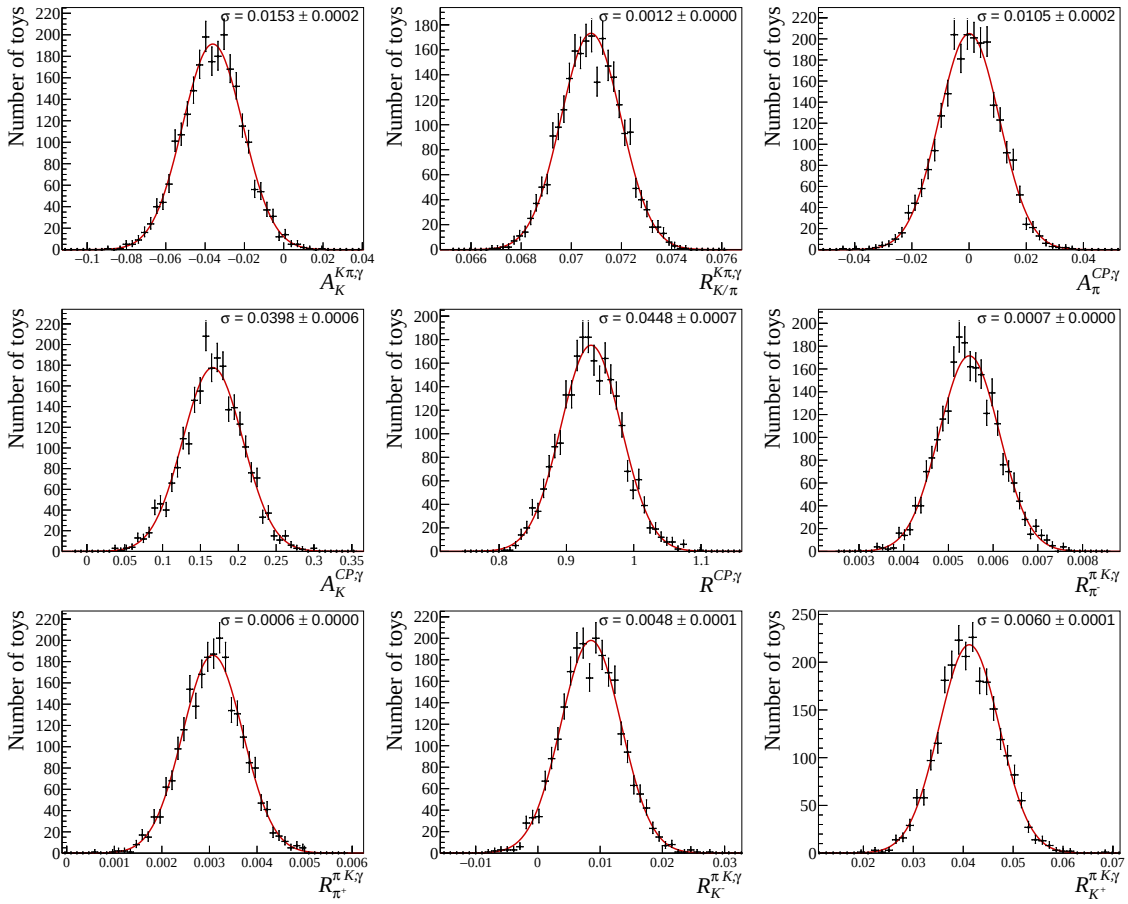


Figure 4.18: Bootstrap distributions for each $D^* \to D\gamma$ *CP* observable in the $D\gamma$ data fit.

raw `RooFit` results do not. The widths of the bootstrap distributions are therefore taken to represent the true statistical uncertainties on the observables; these can be viewed in list format with the final results in Sec. 4.6.

Validation of this method is given by looking at the pull distributions of the bootstrap fits. These are defined by the equation:

$$\mathcal{P}' = \begin{cases} \dfrac{\mathcal{O}_{nominal} - \mathcal{O}_{bootstrap}}{\Delta^{+}_{bootstrap}}, & \text{if } \mathcal{O}_{bootstrap} \leq \mathcal{O}_{nominal} \\ \dfrac{\mathcal{O}_{bootstrap} - \mathcal{O}_{nominal}}{\Delta^{-}_{bootstrap}}, & \text{if } \mathcal{O}_{bootstrap} > \mathcal{O}_{nominal}, \end{cases} \tag{4.30}$$

where $\mathcal{O}_{bootstrap}$ is the final value of the observable found by the fit to the bootstrap sample; $\mathcal{O}_{nominal}$ is the nominal value of the fit to data; and $\Delta^{+(-)}_{bootstrap}$ is the raw positive (negative) uncertainty on the observable $\mathcal{O}$, measured by the bootstrap fit. The width of each pull, $\sigma_{\mathcal{P}'}$, relates to the raw fit error on a given observable, Δ_{raw},

and the true statistical uncertainty on that observable, Δ_{stat}, in the following way:

$$\Delta_{stat} = \Delta_{raw} \times \sigma_{\mathcal{P}'}. \tag{4.31}$$

Employing Poisson errors, in the limit of that all events are double counted and background levels are negligible, the width of each pull is expected be $\sigma_{\mathcal{P}'} = \sqrt{2}$. The proportion of $D\pi^0$ signal candidates that are present in both mass samples in the $D\pi^0$ fit is 88%; the proportion of $D\gamma$ ($D\pi^0$) signal candidates that are present in both mass samples in the $D\gamma$ fit is 86% (63%). The widths are therefore expected to fall within the range $\sigma_{\mathcal{P}'} \in [1, \sqrt{2}]$. In addition, in order to demonstrate that the observables are measured in an unbiased way, the pulls should be centered on zero.

The pull distributions for each *CP* observable are fit to using a Gaussian function with freely varying mean and width; the plots are provided for the $B^{\pm} \to (D^* \to D\gamma)h^{\pm}$ mode observables reconstructed in the $D\gamma$ final state in Fig. 4.19. Similar

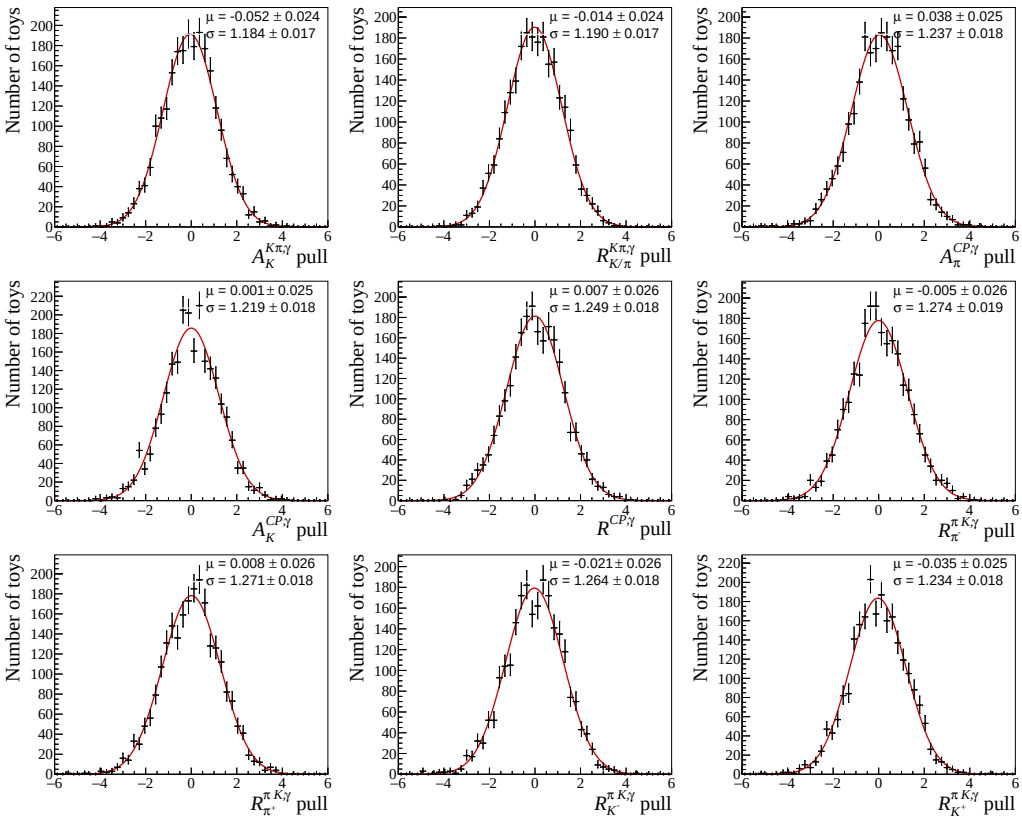


Figure 4.19: Pull distribution from 2D toys for each $D^* \to D\gamma$ *CP* observable in the $D\gamma$ data fit.

Observable	Pull μ	Pull σ
$A_K^{K\pi,\pi^0}$	-0.01 ± 0.03	1.26 ± 0.02
A_K^{CP,π^0}	0.03 ± 0.02	1.15 ± 0.02
A_π^{CP,π^0}	-0.06 ± 0.03	1.25 ± 0.02
R^{CP,π^0}	0.01 ± 0.03	1.28 ± 0.02
$R_{K/\pi}^{K\pi,\pi^0}$	-0.05 ± 0.03	1.25 ± 0.02
$R_{K^-}^{\pi K,\pi^0}$	-0.03 ± 0.03	1.25 ± 0.02
$R_{K^+}^{\pi K,\pi^0}$	-0.05 ± 0.03	1.29 ± 0.02
$R_{\pi^-}^{\pi K,\pi^0}$	-0.06 ± 0.03	1.26 ± 0.02
$R_{\pi^+}^{\pi K,\pi^0}$	-0.03 ± 0.03	1.27 ± 0.02

Table 4.6: Measured pull widths from 2D toys for each *CP* observable in the $D\pi^0$ invariant-mass fit. The pulls for each observable are computed on the same set of toys and thus the uncertainties in the table are not independent.

Observable	Pull μ	Pull σ
$A_K^{K\pi,\gamma}$	-0.06 ± 0.02	1.19 ± 0.02
$A_K^{K\pi,\pi^0}$	0.01 ± 0.02	1.18 ± 0.02
$A_K^{CP,\gamma}$	0.00 ± 0.03	1.22 ± 0.02
$A_\pi^{CP,\gamma}$	0.03 ± 0.03	1.23 ± 0.02
A_K^{CP,π^0}	-0.00 ± 0.02	1.21 ± 0.02
A_π^{CP,π^0}	0.02 ± 0.03	1.23 ± 0.02
$R^{CP,\gamma}$	0.00 ± 0.03	1.25 ± 0.02
R^{CP,π^0}	0.03 ± 0.02	1.17 ± 0.02
$R_{K/\pi}^{K\pi,\gamma}$	-0.01 ± 0.02	1.19 ± 0.02
$R_{K/\pi}^{K\pi,\pi^0}$	-0.01 ± 0.03	1.22 ± 0.02
$R_{K^-}^{\pi K,\gamma}$	-0.02 ± 0.03	1.26 ± 0.02
$R_{K^+}^{\pi K,\gamma}$	-0.03 ± 0.03	1.23 ± 0.02
$R_{\pi^-}^{\pi K,\gamma}$	-0.00 ± 0.03	1.27 ± 0.02
$R_{\pi^+}^{\pi K,\gamma}$	0.01 ± 0.03	1.27 ± 0.02
$R_{K^-}^{\pi K,\pi^0}$	-0.04 ± 0.03	1.26 ± 0.02
$R_{K^+}^{\pi K,\pi^0}$	-0.06 ± 0.03	1.31 ± 0.02
$R_{\pi^-}^{\pi K,\pi^0}$	-0.03 ± 0.03	1.27 ± 0.02
$R_{\pi^+}^{\pi K,\pi^0}$	-0.01 ± 0.03	1.27 ± 0.02

Table 4.7: Measured pull widths from 2D toys for each *CP* observable in the $D\gamma$ invariant-mass fit. The pulls for each observable are computed on the same set of toys and thus the uncertainties in the table are not independent.

results are obtained for the fully and partially reconstructed $D\pi^0$ observables, as can be seen in Tabs. 4.6 and 4.7, where the pull means and widths are collated for *CP* observables measured by the $D\pi^0$ and $D\gamma$ invariant-mass fits, respectively. All means are centered on zero and the width of each Gaussian lies within the expected range. A systematic uncertainty to account for the variation in pull width across *CP* observables in assigned in Sec. 4.5.

4.5 Systematic uncertainties

All observables measured in this analysis are ratios of topologically identical final states, therefore most sources of systematic uncertainty cancel. The majority of those that remain enter as fixed parameters in the invariant-mass fits, where each parameter has an associated uncertainty; what must be determined is how these uncertainties translate onto the *CP* observables of interest.

To evaluate this effect, the signal extraction fits are run 1000 times, where, for each run, a given set of parameters are varied randomly by sampling each from a Gaussian function with mean representing the nominal parameter value and width corresponding to the assigned uncertainty. A distribution of each observable is built up by recording the final value after each run, where outliers are removed by requiring a Z-score of less than 3:

$$Z = \frac{\mathcal{O} - \mu}{\sigma}. \tag{4.32}$$

Here, for a given observable, $\mathcal{O}$ represents the value after a systematics run, μ the measured value in the fit to data with all fixed parameters at their nominal values, and σ the corrected statistical uncertainty. The root mean square (RMS) of these distributions are then taken to be the systematic uncertainties on the observables associated with the set of parameters that were varied.

There are two categories of systematic uncertainty that are evaluated using different methods, which will be explained in the subsection below: the uncertainty assigned to the correction of statistical errors, covered in Sec. 4.4, and the uncertainty

assigned to modelling $B_s^0 \to D^* K^- \pi^+$ decays with a hybrid simulation sample, as described in Sec. 3.3.4.

4.5.1 Sources of systematic uncertainty

Fixed parameters in the invariant-mass fits and their assigned uncertainties are detailed in this subsection; the assigned parameter uncertainties apply to both the $D\pi^0$ and $D\gamma$ invariant-mass fits, unless specified. The parameters within each category listed below are varied simultaneously to account for correlated effects on the final observables; explicit correlations between parameters are not modelled.

For PDF parameters fixed from fits to simulation, the central values and uncertainties are taken directly from these results.

$B^\pm \to (D^* \to D\pi^0)h^\pm$ PDFs

The tail parameters and fraction of each Crystal Ball shape in the composite PDF, for both $m(B)$ and Δm samples, are fixed from fits to simulation. The ratio of peak width in D^*K with respect to $D^*\pi$ is fixed in both data fits to the value $\pm$ uncertainty found when it is allowed to float.

$B^\pm \to (D^* \to D\gamma)h^\pm$ PDFs

The tail parameters and fraction of each Crystal Ball shape in the composite PDF, for both $m(B)$ and Δm samples, are fixed from fits to simulation. The ratio of peak width in D^*K with respect to $D^*\pi$ is fixed in both data fits to the value $\pm$ uncertainty found when it is allowed to float.

Mis- and partially reconstructed background PDFs

All PDF parameters describing the mis- and partially reconstructed backgrounds listed in Sec. 3.3 are fixed from fits to simulation, apart from the mis-reconstructed $B \to D\pi^\pm\pi$ component and the partially reconstructed $B \to D^*\pi^\pm\pi$ component in the ADS mode. These have different phase space distributions compared to the same decay modes in the other fit categories as mis-identified favoured B^0 decays dominate;

the PDF parameters in the $D^*\pi$ ADS mode are therefore given some freedom. The means and widths of the $m(B)$ PDFs are allowed to float in the fits to data then are fixed to the measured values ± uncertainties. The same procedure is performed to the A parameters of the $m(D^*) - m(D^0)$ mass difference PDFs in Δm.

Fixed parameters associated with combinatorial backgrounds are also included in this parameter set as they sit underneath the mis- and partially reconstructed modes. These components are modelled using a wide Gaussian in $m(B)$, with fixed mean $(5200 \pm 100)\,\text{MeV}/c^2$ and width $(300 \pm 100)\,\text{MeV}/c^2$. The $\epsilon_{\Delta m}$ box efficiency, calculated from the B mass distribution, is determined by integrating the Gaussian function over the box cut range and is assigned a 10% error. In Δm, the peaking (true D^{*0}) and flat (fake D^{*0}) components are modelled using fully and mis-reconstructed signal PDFs, respectively; the $\epsilon_{m(B)}$ box efficiencies, calculated from Δm distributions, are a normalised sum of the corresponding fully and mis-reconstructed signal efficiencies. The uncertainties on the Δm PDF parameters and $\epsilon_{m(B)}$ box efficiencies are therefore shared with the signal components.

The weighted average of the measured combinatorial asymmetries in the partially reconstructed $B^\pm \to D^{(*)}h^\pm$ analysis [31] is (1.3±8.6)%. The combinatorial component in this work is therefore assumed to exhibit zero *CP* asymmetry, with a 10% uncertainty. Combinatorial yields float in the favoured mode, then are fixed relative to this in all other modes using the measured yields in the PR analysis; these fixed ratios are assigned uncertainties calculated from the statistical errors on the PR yields.

$B_s^0 \to D^* K^- \pi^+$ PDFs

$B_s^0 \to D^* K^- \pi^+$ decays are considered in the D^*K ADS and GLW modes. All PDF parameters are fixed from fits to simulation samples.

For the cases where the true neutral D^* decay is reconstructed, hybrid $B_s^0 \to D^* K^- \pi^+$ samples are used to describe the $K^-\pi^+$ phase space: $(80 \pm 10)\%$ $B_s^0 \to D^*(K^{*0}(892) \to K^-\pi^+)$ decays and $(20 \pm 10)\%$ $B_s^0 \to D^* K^- \pi^+$ decays generated over a flat, square Dalitz space (see Sec. 3.3.4 for details). To evaluate the systematic

uncertainty associated with this phase space modelling, the invariant-mass fits to data are run with the PDF parameters and box efficiencies for the $B_s^0 \to D^* K^- \pi^+$ components calculated from hybrid simulation samples with a 70 : 30 split and a 90 : 10 split. The difference between the final values of the observables measured by the data fits when run with these alternative setups, with respect to the observables measured using the default setup (with an 80 : 20 split), are taken to be the systematic uncertainties. The positive and negative uncertainties on the observables calculated using this method are almost equal, so the largest uncertainty is chosen and attributed symmetrically to each observable.

Mis-identified and crossfeed PDFs

Components for favoured mode decays mis-reconstructed in the ADS mode are included for each signal channel; all PDF parameters are fully fixed from fits to simulated favoured mode samples with both D decay product mass hypotheses swapped. Mis-identified PDF parameters are fully fixed from fits to simulated samples reconstructed with the incorrect companion mass hypothesis.

Box efficiencies

The box efficiencies given in Tab. 3.5, determined from simulation samples, are used to split total component yields into 2(3) individual mass samples in the $D\pi^0$ ($D\gamma$) data fit; the errors given in this table are calculated from the finite size of simulation samples.

Signal PDFs in the invariant-mass fits to data are well constrained in both B mass and Δm samples, therefore their means and widths are able to float freely and their tail parameters are floated then fixed to the value preferred by the data. They therefore provide good candidates to investigate the difference between mass shapes (and therefore box efficiencies) in data and simulation.

The integral of each signal PDF in the data fits, over the mass range of the relevant box cut, is evaluated and compared to the same integral from the fit to simulation. The average difference of these integrals across all signal PDFs is 0.8%. This is taken to represent the error associated with transferring the box efficiencies,

measured using simulation, onto data.

Companion hadron PID cuts have the potential to influence invariant-mass distributions, and therefore box efficiencies. To account for the fact that efficiencies are calculated before the application of PID, the box efficiencies for each simulation sample were evaluated before and after cutting on the stored, uncalibrated $\Delta LL_{K/\pi}$ variable. The largest difference compared to the default efficiencies, calculated before PID cuts, was found to be 0.8%.

The total uncertainty assigned to each box efficiency when evaluating systematic uncertainties is the sum in quadrature of the 0.8% uncertainty assigned for MC/data differences, the 0.8% uncertainty assigned for PID effects, and each binomial error given in Tab. 3.5. Each box efficiency is varied as a systematic at the same time as the corresponding PDF component as both are determined by the invariant-mass distribution of the associated decay mode.

Selection efficiencies

The signal selection efficiencies listed in Tab. 3.6, determined from simulation samples, are used to correct raw yields in the signal extraction fits. The background selection efficiencies listed in Tab. 3.7 are used to constrain background components. In both cases, efficiencies are varied as a source of systematic uncertainty within the given errors, which are determined from finite simulation sample sizes.

PID efficiencies

Kaon and pion PID efficiencies are determined by correcting $\Delta LL_{K/\pi}$ distributions in simulation samples with data-driven calibration samples, as described in Sec. 3.1.6. The errors are given in Tabs. 3.3 and 3.4, which account for the use of binned efficiencies and finite sample sizes.

Fixed branching fractions and rates

The residual rate of doubly mis-identified favoured mode candidates infiltrating the ADS sub-samples, taken from the partially reconstructed $B^{\pm} \to D^{(*)}h^{\pm}$ analysis [31],

is fixed in both invariant-mass fits to the value ± uncertainty given in Sec. 3.3.7.

The fit components describing mis-reconstructed $\Lambda_b^0 \to \Sigma_c^+ h^-$ and $B \to Dh^\pm\pi^-$ decays are constrained in the D^*K sub-samples by the corresponding yields in the $D^*\pi$ sub-samples using the fixed branching fractions listed in Tab. 4.1.

The partially reconstructed $B_s^0 \to D^*K^-\pi^+$ and $B^\pm \to D^*h^\pm\pi$ background components are described by composite PDFs , as detailed in Sec. 3.3.4. The branching fractions used to fix the relative amount of the contributing decay modes are varied within their known errors when evaluating systematic uncertainties.

Fixed asymmetries

The following asymmetries are fixed in the invariant-mass fits and varied as a source of systematic uncertainty within the quoted errors: the detection asymmetries given in Sec. 4.1.1, the *CP* asymmetry between $D \to K^+K^-$ and $D \to \pi^+\pi^-$ decays described in Sec. 4.1.2, and the fixed physics background asymmetries summarised in Tab. 4.1.

Fixed *CP* ratios

Fixed values of R_{CP} and $R_h^{\pi K}$ for the background components, summarised in Tab. 4.1, are varied within the given uncertainties when evaluating systematics.

Statistical error correction

The increase in statistical uncertainty due to the double counting of events, represented by the pull width of bootstrap distributions (see Sec. 4.4), should be approximately uniform across all *CP* observables. Differences in background levels across the different decay channels can lead to differences in the correction factor: the higher the background level, the lower the correction is expected to be (as a smaller fraction of the statistical error is Poissonian).

The root mean square (RMS) of the range of pull widths given in Tabs. 4.6 and 4.7, multiplied by the raw statistical uncertainty for each observable, is therefore conservatively assigned as the systematic uncertainty associated with the bootstrapping method. The RMS is calculated separately for fully reconstructed $D\pi^0$ measurements,

fully reconstructed $D\gamma$ measurements and partially reconstructed $D\pi^0$ measurements, in line with the different box efficiencies of these modes.

4.5.2 Systematic uncertainties for each observable

A breakdown of the systematic uncertainties from the sources listed above can be found in Appendix B. Summaries of the systematics for each observable are given in Tabs. 4.8 and 4.9 for the $D\pi^0$ and $D\gamma$ data fits, respectively, where the uncertainties have been grouped into the following categories:

- $PDFs$: fixed PDF parameters and box efficiencies.
- ϵ_{sel}: fixed selection efficiencies determined from simulation.
- ϵ_{PID}: fixed kaon and pion PID efficiencies.
- *Rates*: relative decay rates fixed from branching fraction measurements and the fixed background level of favoured to ADS mode crossfeed peaking backgrounds.
- *Asyms*: fixed asymmetries.
- *CP Ratios*: fixed *CP* ratios.
- *Corr*: systematic assigned to the correction of statistical uncertainties.

All systematics are given as a percentage of the statistical uncertainty of the associated observable, in order to highlight their relative importance. The total systematic

Observable	$PDFs$	ϵ_{sel}	ϵ_{PID}	*Rates*	*Asyms*	*CP Ratios*	*Corr*	Total
$R^{\pi K,\pi^0}_{K^-}$	17.91	2.12	1.80	2.72	5.85	5.86	3.76	20.46
$R^{\pi K,\pi^0}_{K^+}$	14.54	1.13	2.19	1.10	5.85	7.66	3.76	18.05
$R^{\pi K,\pi^0}_{\pi^-}$	44.98	0.48	2.28	0.59	4.17	3.76	3.76	45.55
$R^{\pi K,\pi^0}_{\pi^+}$	47.11	0.59	1.62	1.64	4.65	3.74	3.76	47.69
R^{CP,π^0}	17.89	18.97	2.37	3.59	2.02	1.67	3.76	26.82
$R^{K\pi,\pi^0}_{K/\pi}$	57.91	45.96	39.68	5.77	2.07	2.05	3.76	84.24
$A^{K\pi,\pi^0}_{K}$	14.35	0.60	0.66	0.55	6.07	0.73	3.76	16.08
A^{CP,π^0}_{K}	12.75	3.18	1.89	2.18	3.31	1.88	3.76	14.47
A^{CP,π^0}_{π}	5.57	0.43	0.42	0.42	9.24	0.63	3.76	11.47

Table 4.8: Systematic uncertainties for each observable measured in the $D\pi^0$ data fit. Uncertainties are given as a percentage of the statistical uncertainty, and the total is evaluated by summing all columns in quadrature.

uncertainty for each observable, printed in the right hand column of each table, represents the sum in quadrature of the individual sources.

All observables are statistically limited. It can be seen that the systematic uncertainties for the ratio observables are larger than those of the asymmetries. This is due to the presence of different D decay modes in the numerator and denominator of the former: the systematic effect of different background contributions across the D decay mode categories do not cancel. The systematic uncertainties only approach the statistical in the case of $R_{K/\pi}^{K\pi,\pi^0}$ and $R_{K/\pi}^{K\pi,\gamma}$, the ratios of favoured mode yields in D^*K and $D^*\pi$, due to the presence of a different companion particle on the top and bottom of the ratio. This means that selection and PID efficiency corrections do not cancel, inflating the systematic uncertainties compared to the other observables.

Observable	*PDFs*	ϵ_{sel}	ϵ_{PID}	*Rates*	*Asyms*	*CP Ratios*	*Corr*	Total
$R_{K^-}^{\pi K,\gamma}$	27.48	0.62	0.48	1.86	8.20	4.06	2.95	29.19
$R_{K^+}^{\pi K,\gamma}$	18.10	0.95	3.49	5.24	8.41	10.19	2.95	23.48
$R_{\pi^-}^{\pi K,\gamma}$	21.74	0.10	0.49	0.54	1.94	0.13	2.95	22.04
$R_{\pi^+}^{\pi K,\gamma}$	20.86	0.12	0.83	0.64	2.02	0.31	2.95	21.20
$R_{K^-}^{\pi K,\pi^0}$	28.59	0.89	2.50	1.22	7.11	3.87	4.34	30.17
$R_{K^+}^{\pi K,\pi^0}$	39.32	1.12	1.23	1.62	8.46	10.70	4.34	41.91
$R_{\pi^-}^{\pi K,\pi^0}$	38.29	0.23	1.25	1.37	1.72	0.32	4.34	38.62
$R_{\pi^+}^{\pi K,\pi^0}$	39.00	0.24	1.07	1.27	1.64	0.37	4.34	39.31
$R^{CP,\gamma}$	19.19	17.75	5.66	3.45	4.91	4.70	2.95	27.97
R^{CP,π^0}	37.24	14.24	12.21	4.26	8.77	4.20	4.34	43.25
$R_{K/\pi}^{K\pi,\gamma}$	52.57	49.19	65.75	18.64	2.99	3.10	2.95	99.40
$R_{K/\pi}^{K\pi,\pi^0}$	61.54	37.66	54.56	13.49	3.41	3.50	4.34	91.69
$A_K^{K\pi,\gamma}$	6.98	0.37	1.03	0.47	12.06	0.18	2.95	14.30
$A_K^{K\pi,\pi^0}$	9.86	0.48	0.39	2.06	10.32	0.39	4.34	15.08
$A_K^{CP,\gamma}$	14.60	0.65	0.28	1.42	8.64	0.12	2.95	17.29
$A_\pi^{CP,\gamma}$	5.82	0.10	0.20	0.66	15.59	0.32	2.95	16.92
A_K^{CP,π^0}	22.33	1.18	0.66	2.39	19.86	1.51	4.34	30.36
A_π^{CP,π^0}	10.15	0.11	0.59	0.37	11.04	0.22	4.34	15.63

Table 4.9: Systematic uncertainties for each observable measured in the $D\gamma$ data fit. Uncertainties are given as a percentage of the statistical uncertainty, and the total is evaluated by summing all columns in quadrature.

4.6 Results

Invariant-mass fits to the $m(D^*h)$ and Δm distributions of data reconstructed in the $D\pi^0$ and $D\gamma$ final states are shown in Figs. 4.2 to 4.9 and Figs. 4.10 to 4.17, respectively; the measured signal yields are given in Tabs. 4.2 and 4.3. The final results for the $B \to (D^* \to D\pi^0)h^\pm$ *CP* observables measured using data reconstructed in the $D\pi^0$ final state are:

$$
\begin{aligned}
R_{K^-}^{\pi K,\pi^0} &= 0.0215 \pm 0.0069 \pm 0.0014 \\
R_{K^+}^{\pi K,\pi^0} &= 0.0077 \pm 0.0059 \pm 0.0011 \\
R_{\pi^-}^{\pi K,\pi^0} &= 0.0039 \pm 0.0007 \pm 0.0003 \\
R_{\pi^+}^{\pi K,\pi^0} &= 0.0044 \pm 0.0008 \pm 0.0004 \\
R^{CP,\pi^0} &= 1.0692 \pm 0.0853 \pm 0.0228 \\
R_{K/\pi}^{K\pi,\pi^0} &= 0.0732 \pm 0.0023 \pm 0.0019 \\
A_K^{K\pi,\pi^0} &= -0.0031 \pm 0.0303 \pm 0.0049 \\
A_K^{CP,\pi^0} &= -0.1804 \pm 0.0564 \pm 0.0083 \\
A_\pi^{CP,\pi^0} &= 0.0195 \pm 0.0179 \pm 0.0020.
\end{aligned}
$$

The final results for the $B \to (D^* \to D\gamma)h^\pm$ and $B \to (D^* \to D\pi^0)h^\pm$ *CP* observables measured using data reconstructed in the $D\gamma$ final state are:

$$
\begin{aligned}
R_{K^-}^{\pi K,\gamma} &= 0.0086 \pm 0.0048 \pm 0.0014 \\
R_{K^+}^{\pi K,\gamma} &= 0.0414 \pm 0.0060 \pm 0.0014 \\
R_{\pi^-}^{\pi K,\gamma} &= 0.0056 \pm 0.0007 \pm 0.0001 \\
R_{\pi^+}^{\pi K,\gamma} &= 0.0032 \pm 0.0006 \pm 0.0001 \\
R_{K^-}^{\pi K,\pi^0} &= 0.0333 \pm 0.0075 \pm 0.0023 \\
R_{K^+}^{\pi K,\pi^0} &= -0.0001 \pm 0.0052 \pm 0.0022 \\
R_{\pi^-}^{\pi K,\pi^0} &= 0.0033 \pm 0.0007 \pm 0.0003 \\
R_{\pi^+}^{\pi K,\pi^0} &= 0.0051 \pm 0.0008 \pm 0.0003 \\
R^{CP,\gamma} &= 0.9357 \pm 0.0448 \pm 0.0125
\end{aligned}
$$

$$
\begin{aligned}
R^{CP,\pi^0} &= \quad 1.0586 \pm 0.0644 \pm 0.0277 \\
R^{K\pi,\gamma}_{K/\pi} &= \quad 0.0708 \pm 0.0012 \pm 0.0012 \\
R^{K\pi,\pi^0}_{K/\pi} &= \quad 0.0723 \pm 0.0016 \pm 0.0015 \\
A^{K\pi,\gamma}_{K} &= -0.0354 \pm 0.0153 \pm 0.0022 \\
A^{K\pi,\pi^0}_{K} &= \quad 0.0563 \pm 0.0191 \pm 0.0029 \\
A^{CP,\gamma}_{K} &= \quad 0.1525 \pm 0.0398 \pm 0.0069 \\
A^{CP,\gamma}_{\pi} &= -0.0126 \pm 0.0105 \pm 0.0018 \\
A^{CP,\pi^0}_{K} &= -0.1922 \pm 0.0468 \pm 0.0142 \\
A^{CP,\pi^0}_{\pi} &= \quad 0.0083 \pm 0.0148 \pm 0.0023,
\end{aligned}
$$

where the first quoted uncertainties are statistical and the second are systematic. The statistical uncertainties have been corrected for double-counting using the procedure described in Sec. 4.4.

The measured values of $R^{\pi K,\pi^0/\gamma}_{h^+}$ and $R^{\pi K,\pi^0/\gamma}_{h^-}$ can be transformed into the *CP* parameters $R^{\pi K,\pi^0/\gamma}_{h}$ and $A^{\pi K,\pi^0/\gamma}_{h}$ using Eqs. (1.38) and (1.39), respectively:

$$
\begin{aligned}
R^{\pi K,\pi^0}_{K} &= \quad 0.0155 \quad \pm 0.0037 \\
R^{\pi K,\gamma}_{K} &= \quad 0.0250 \quad \pm 0.0045 \\
R^{\pi K,\pi^0}_{\pi} &= \quad 0.00418 \pm 0.00047 \\
R^{\pi K,\gamma}_{\pi} &= \quad 0.00443 \pm 0.00048 \\
A^{\pi K,\pi^0}_{K} &= \quad 0.697 \quad \pm 0.228 \\
A^{\pi K,\gamma}_{K} &= -0.656 \quad \pm 0.163 \\
A^{\pi K,\pi^0}_{\pi} &= -0.127 \quad \pm 0.083 \\
A^{\pi K,\gamma}_{\pi} &= \quad 0.270 \quad \pm 0.106,
\end{aligned}
$$

where statistical and systematic uncertainties have been combined according to correlations between $R^{\pi K,\pi^0/\gamma}_{h^+}$ and $R^{\pi K,\pi^0/\gamma}_{h^-}$. Single values for $R^{\pi K,\pi^0}_{h}$ and $A^{\pi K,\pi^0}_{h}$ are

obtained by taking the error-weighted average:

$$\mu = \frac{\sum_i x_i/\sigma_i^2}{\sum_i 1/\sigma_i^2} \tag{4.33}$$

$$\sigma = \frac{1}{\sqrt{\sum_i 1/\sigma_i^2}} \tag{4.34}$$

of measurements made by both the $D\gamma$ and $D\pi^0$ signal-extraction fits. Here, for a given observable, x_i represents the measurement from a single fit, and σ_i is the combined statistical and systematic uncertainty on the measurement; μ and σ represent the average value and error of the combination.

The statistical significance of the D^*K ADS modes can be calculated using Wilk's theorem [94]:

$$S = \sqrt{-2\ln\left(\frac{\mathcal{L}_0}{\mathcal{L}_1}\right)}, \tag{4.35}$$

where $\mathcal{L}_0$ is the extended maximum likelihood value for the nominal D1D fit model, and $\mathcal{L}_1$ is the extended maximum likelihood value for an alternative D1D model with $R_{h^+}^{\pi K,\pi^0/\gamma}$ and $R_{h^-}^{\pi K,\pi^0/\gamma}$ set to zero. The likelihood value is larger if the model better describes the data, therefore the more unlikely the alternative model is correct, the higher the statistical significance. The significance of fully reconstructed $B^\pm \to (D^* \to [\pi^\pm K^\mp]_D\gamma)K^\pm$ signal is measured to be 8.8σ, constituting the *first observation* of this decay channel. The significance of partially reconstructed $B^\pm \to (D^* \to [\pi^\pm K^\mp]_D\pi^0)K^\pm$ signal is measured to be 5.9σ; a significant result, but this mode was first observed by the partially reconstructed $B^\pm \to D^{(*)}h^\pm$ analysis [31]. The significance of fully reconstructed $B^\pm \to (D^* \to [\pi^\pm K^\mp]_D\pi^0)K^\pm$ signal is measured to be 3.7σ.

4.6.1 Correlation matrices

The statistical and systematic correlation matrices for *CP* observables are given in Tabs. 4.10 to 4.13.

	$A_K^{K\pi,\pi^0}$	A_K^{CP,π^0}	A_π^{CP,π^0}	R^{CP,π^0}	$R_{K^-}^{\pi K,\pi^0}$	$R_{K^+}^{\pi K,\pi^0}$	$R_{\pi^-}^{\pi K,\pi^0}$	$R_{\pi^+}^{\pi K,\pi^0}$
$A_K^{K\pi,\pi^0}$	1.00	0.03	0.08	−0.00	0.01	−0.01	0.02	−0.02
A_K^{CP,π^0}		1.00	0.03	0.08	−0.00	−0.01	0.01	−0.01
A_π^{CP,π^0}			1.00	0.00	0.02	−0.02	0.03	−0.03
R^{CP,π^0}				1.00	0.04	0.00	−0.00	−0.00
$R_{K^-}^{\pi K,\pi^0}$					1.00	0.23	−0.03	0.01
$R_{K^+}^{\pi K,\pi^0}$						1.00	0.01	−0.02
$R_{\pi^-}^{\pi K,\pi^0}$							1.00	0.28
$R_{\pi^+}^{\pi K,\pi^0}$								1.00

Table 4.10: Statistical correlation matrix for observables measured in the $D^* \to D\pi^0$ invariant-mass fit.

	$R_{K^-}^{\pi K,\pi^0}$	$R_{K^+}^{\pi K,\pi^0}$	$R_{\pi^-}^{\pi K,\pi^0}$	$R_{\pi^+}^{\pi K,\pi^0}$	R^{CP,π^0}	$A_K^{K\pi,\pi^0}$	A_K^{CP,π^0}	A_π^{CP,π^0}
$R_{K^-}^{\pi K,\pi^0}$	1.00	0.88	0.67	0.67	0.28	0.68	0.62	0.41
$R_{K^+}^{\pi K,\pi^0}$		1.00	0.43	0.43	0.04	0.43	0.29	0.34
$R_{\pi^-}^{\pi K,\pi^0}$			1.00	1.00	0.60	0.93	0.92	0.46
$R_{\pi^+}^{\pi K,\pi^0}$				1.00	0.60	0.93	0.92	0.46
R^{CP,π^0}					1.00	0.53	0.71	0.18
$A_K^{K\pi,\pi^0}$						1.00	0.94	0.74
A_K^{CP,π^0}							1.00	0.55
A_π^{CP,π^0}								1.00

Table 4.11: Systematic correlation matrix for observables measured in the $D^* \to D\pi^0$ invariant-mass fit.

	$A_K^{K\pi,\gamma}$	$A_K^{K\pi,\pi^0}$	$A_K^{CP,\gamma}$	$A_\pi^{CP,\gamma}$	A_K^{CP,π^0}	A_π^{CP,π^0}	$R^{CP,\gamma}$	R^{CP,π^0}	$R_{K^-}^{\pi K,\gamma}$	$R_{K^+}^{\pi K,\gamma}$	$R_{\pi^-}^{\pi K,\gamma}$	$R_{\pi^+}^{\pi K,\gamma}$	$R_{K^-}^{\pi K,\pi^0}$	$R_{K^+}^{\pi K,\pi^0}$	$R_{\pi^-}^{\pi K,\pi^0}$	$R_{\pi^+}^{\pi K,\pi^0}$
$A_K^{K\pi,\gamma}$	1.00	−0.14	0.02	0.06	0.02	0.04	−0.00	0.00	0.01	−0.01	0.00	−0.00	0.00	−0.00	0.00	−0.00
$A_K^{K\pi,\pi^0}$		1.00	0.02	0.05	0.01	0.04	−0.00	0.01	0.01	−0.00	0.00	−0.00	0.01	−0.00	0.00	−0.00
$A_K^{CP,\gamma}$			1.00	0.02	−0.18	0.03	−0.06	−0.00	0.00	−0.00	0.00	−0.00	0.00	−0.00	0.00	−0.00
$A_\pi^{CP,\gamma}$				1.00	0.02	0.20	−0.00	−0.00	0.01	−0.01	0.01	−0.01	0.01	−0.01	0.00	−0.00
A_K^{CP,π^0}					1.00	−0.01	0.01	0.12	0.00	−0.00	0.00	−0.00	0.00	0.00	0.00	0.00
A_π^{CP,π^0}						1.00	0.00	0.00	0.01	−0.00	0.00	−0.00	0.00	−0.00	0.00	−0.00
$R^{CP,\gamma}$							1.00	0.03	0.03	0.05	−0.00	−0.00	0.01	0.01	0.00	0.00
R^{CP,π^0}								1.00	0.00	0.00	−0.00	−0.00	0.04	0.00	0.00	0.00
$R_{K^-}^{\pi K,\gamma}$									1.00	0.20	−0.00	0.00	0.03	0.17	0.02	−0.00
$R_{K^+}^{\pi K,\gamma}$										1.00	0.00	−0.01	0.15	0.05	−0.01	0.02
$R_{\pi^-}^{\pi K,\gamma}$											1.00	0.00	0.01	0.00	0.23	0.00
$R_{\pi^+}^{\pi K,\gamma}$												1.00	0.00	0.01	−0.00	0.23
$R_{K^-}^{\pi K,\pi^0}$													1.00	0.14	−0.05	−0.00
$R_{K^+}^{\pi K,\pi^0}$														1.00	−0.00	−0.03
$R_{\pi^-}^{\pi K,\pi^0}$															1.00	0.03
$R_{\pi^+}^{\pi K,\pi^0}$																1.00

Table 4.12: Statistical correlation matrix for observables measured in the $D^* \to D\gamma$ invariant-mass fit.

	$R^{\pi K,\gamma}_{K^-}$	$R^{\pi K,\gamma}_{K^+}$	$R^{\pi K,\gamma}_{\pi^-}$	$R^{\pi K,\gamma}_{\pi^+}$	$R^{CP,\gamma}$	$A^{K\pi,\gamma}_{K}$	$A^{CP,\gamma}_{K}$	$A^{CP,\gamma}_{\pi}$	$R^{\pi K,\pi^0}_{K^-}$	$R^{\pi K,\pi^0}_{K^+}$	$R^{\pi K,\pi^0}_{\pi^-}$	$R^{\pi K,\pi^0}_{\pi^+}$	R^{CP,π^0}	$A^{K\pi,\pi^0}_{K}$	A^{CP,π^0}_{K}	A^{CP,π^0}_{π}
$R^{\pi K,\gamma}_{K^-}$	1.00	0.86	0.85	0.85	0.47	0.51	0.80	0.39	0.79	0.86	0.85	0.85	0.70	0.61	0.72	0.65
$R^{\pi K,\gamma}_{K^+}$		1.00	0.60	0.60	0.29	0.45	0.60	0.38	0.74	0.85	0.60	0.59	0.49	0.47	0.58	0.51
$R^{\pi K,\gamma}_{\pi^-}$			1.00	1.00	0.65	0.48	0.88	0.31	0.41	0.51	1.00	1.00	0.89	0.66	0.75	0.67
$R^{\pi K,\gamma}_{\pi^+}$				1.00	0.64	0.48	0.88	0.31	0.41	0.51	1.00	1.00	0.89	0.65	0.75	0.67
$R^{CP,\gamma}$					1.00	0.28	0.56	0.15	0.10	0.19	0.65	0.65	0.87	0.36	0.47	0.38
$A^{K\pi,\gamma}_{K}$						1.00	0.82	0.98	0.21	0.22	0.43	0.43	0.46	0.95	0.92	0.96
$A^{CP,\gamma}_{K}$							1.00	0.71	0.35	0.42	0.86	0.86	0.81	0.92	0.97	0.93
$A^{CP,\gamma}_{\pi}$								1.00	0.16	0.15	0.26	0.25	0.28	0.89	0.85	0.90
$R^{\pi K,\pi^0}_{K^-}$									1.00	0.97	0.42	0.41	0.26	0.21	0.30	0.28
$R^{\pi K,\pi^0}_{K^+}$										1.00	0.52	0.51	0.36	0.24	0.36	0.31
$R^{\pi K,\pi^0}_{\pi^-}$											1.00	1.00	0.89	0.61	0.72	0.63
$R^{\pi K,\pi^0}_{\pi^+}$												1.00	0.89	0.61	0.71	0.63
R^{CP,π^0}													1.00	0.57	0.70	0.60
$A^{K\pi,\pi^0}_{K}$														1.00	0.97	0.99
A^{CP,π^0}_{K}															1.00	0.97
A^{CP,π^0}_{π}																1.00

Table 4.13: Systematic correlation matrix for observables measured in the $D^* \to D\gamma$ invariant-mass fit.

4.6.2 Comparison with previous measurements

The corresponding measurements of $R_h^{\pi K,\pi^0/\gamma}$ and $A_h^{\pi K,\pi^0/\gamma}$ made by the partially reconstructed $B^\pm \to D^{(*)}h^\pm$ analysis [31] are:

$$\begin{aligned}
R_K^{\pi K,\pi^0} &= \quad 0.0118 \quad \pm 0.0034 \\
R_K^{\pi K,\gamma} &= \quad 0.0163 \quad \pm 0.0373 \\
R_\pi^{\pi K,\pi^0} &= \quad 0.00471 \pm 0.00077 \\
R_\pi^{\pi K,\gamma} &= \quad 0.00429 \pm 0.00138 \\
A_K^{\pi K,\pi^0} &= \quad 0.717 \quad \pm 0.286 \\
A_K^{\pi K,\gamma} &= -0.558 \quad \pm 1.349 \\
A_\pi^{\pi K,\pi^0} &= -0.149 \quad \pm 0.059 \\
A_\pi^{\pi K,\gamma} &= \quad 0.079 \quad \pm 0.128.
\end{aligned}$$

To provide a point of reference when comparing the results of the two analyses, the signal statistics are also presented. The number of ADS mode candidates measured in the partially reconstructed analysis are as follows:

$$\begin{aligned}
N_{tot}(B^\pm \to (D^* \to [\pi^\pm K^\mp]_D \pi^0) K^\pm) &= 1124 \pm 231 \\
N_{tot}(B^\pm \to (D^* \to [\pi^\pm K^\mp]_D \gamma) K^\pm) &= 674 \pm 931,
\end{aligned}$$

where the quoted uncertainties are statistical. The corresponding yields measured in this analysis (combining partially and fully reconstructed $D\pi^0$ candidates) are:

$$\begin{aligned}
N_{tot}(B^\pm \to (D^* \to [\pi^\pm K^\mp]_D \pi^0) K^\pm) &= 98 \pm 28 \\
N_{tot}(B^\pm \to (D^* \to [\pi^\pm K^\mp]_D \gamma) K^\pm) &= 147 \pm 23.
\end{aligned}$$

These results demonstrate that, despite significantly smaller signal yields due to the low neutral reconstruction efficiency at LHCb, the fully reconstructed technique developed in this thesis measures *CP* observables for the $B^\pm \to D^* h^\pm$ modes with significantly higher purity in the $D\gamma$ final state, and with competitive accuracy in the $D\pi^0$ final state, compared to the partially reconstructed method.

To quantify the compatibility of the results of the two analyses, a comparison across all *CP* observables is performed. For each observable, the level of agreement is determined by calculating the absolute difference in central value between the fully and partially reconstructed measurements, divided by the sum in quadrature of the statistical uncertainties. To make these comparisons, $B^{\pm} \to (D^{*} \to D\pi^{0})h^{\pm}$ observables measured by both the $D\pi^{0}$ and $D\gamma$ invariant mass fits are combined using a weighted average with mean and statistical uncertainty defined by Eqs. (4.33) and (4.34), respectively. The values are summarised in Tab. 4.14, where the χ^2/ndf is found be to 0.69, indicating consistency between the analyses.

Observable	Agreement (σ)	χ^2
$R_{K^-}^{\pi K,\gamma}$	0.14	0.02
$R_{K^+}^{\pi K,\gamma}$	0.55	0.30
$R_{\pi^-}^{\pi K,\gamma}$	0.80	0.63
$R_{\pi^+}^{\pi K,\gamma}$	0.72	0.52
$R_{K^-}^{\pi K,\pi^0}$	1.09	1.18
$R_{K^+}^{\pi K,\pi^0}$	0.01	0.00
$R_{\pi^-}^{\pi K,\pi^0}$	0.59	0.35
$R_{\pi^+}^{\pi K,\pi^0}$	0.75	0.56
$R^{CP,\gamma}$	0.21	0.05
R^{CP,π^0}	0.21	0.04
$A_K^{K\pi,\gamma}$	1.51	2.27
$A_K^{K\pi,\pi^0}$	1.11	1.23
$A_K^{CP,\gamma}$	0.44	0.19
$A_\pi^{CP,\gamma}$	0.72	0.52
A_K^{CP,π^0}	1.77	3.14
A_π^{CP,π^0}	0.01	0.00
χ^2/ndf		$11.01/16 = 0.69$

Table 4.14: Table to quantify the compatibility of *CP* observables measured by this analysis and the partially reconstructed analysis [31]. The agreement levels of each observable are indicated in standard deviations, and the total χ^2/ndf is given in the final row.

Visual comparisons of the results of this analysis and previous measurements made by LHCb and the *B*-factories are shown in Figs. 4.20 to 4.27. The error bars on the LHCb measurements include both statistical and systematic errors, where the former are represented by black bars and the latter red bars. The 68% and 95%

confidence regions are constrained by B-factory measurements, and are represented by the dark and light blue shaded regions, respectively. These are calculated using the equations for the observables in terms of the fundamental parameters derived in Sec. 1.4. Values for $r_B^{D^*K}$ and $\delta_B^{D^*K}$ are taken from Ref. [44], which combines measurements made in previous ADS/GLW and GGSZ analyses preformed by BaBar and Belle. The CKM angle γ is taken from the latest LHCb combination [23], $\gamma = (65.4^{+3.8}_{-4.2})^\circ$. No assumptions are made about $r_B^{D^*\pi}$ or $\delta_B^{D^*\pi}$: both are represented by uniform distributions over the ranges $0 - 0.02$ and $0 - 180^\circ$, respectively.

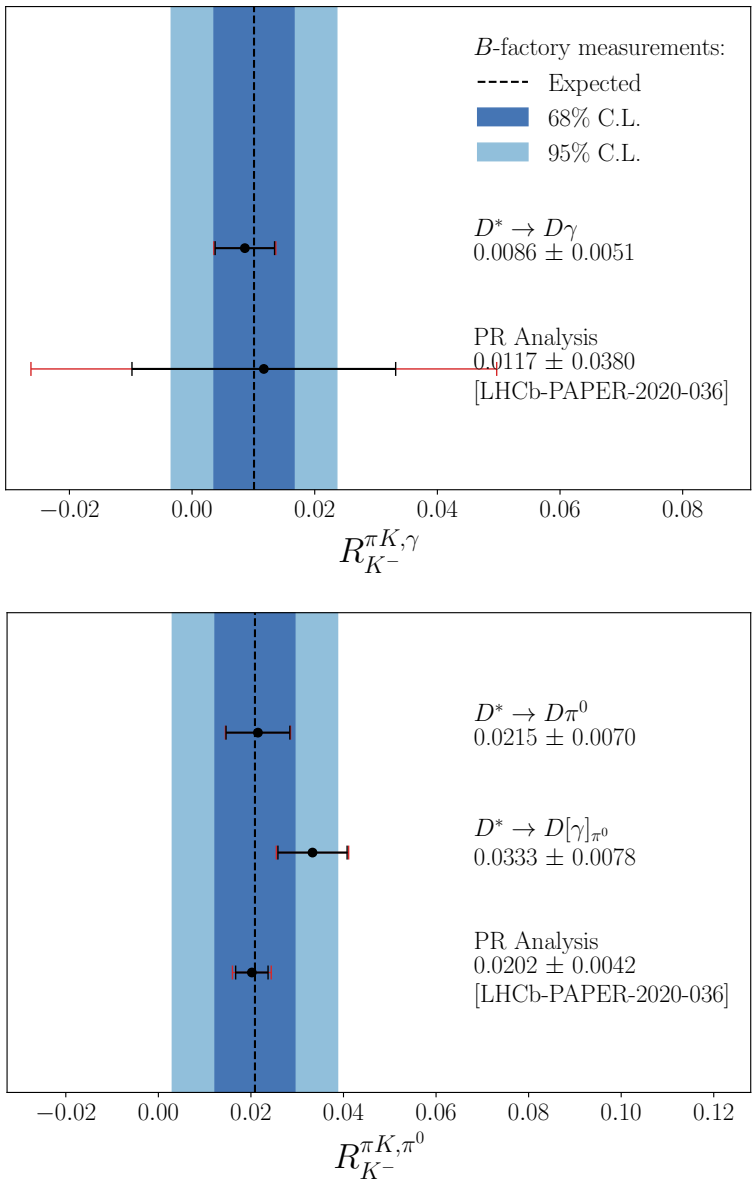


Figure 4.20: Comparison of $R_{K^-}^{\pi K,\gamma}$ (top) and $R_{K^-}^{\pi K,\pi^0}$ (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from measurements of $r_B^{D^*K}$ and $\delta_B^{D^*K}$ using the ADS/GLW and GGSZ methods by BaBar and Belle [44]. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

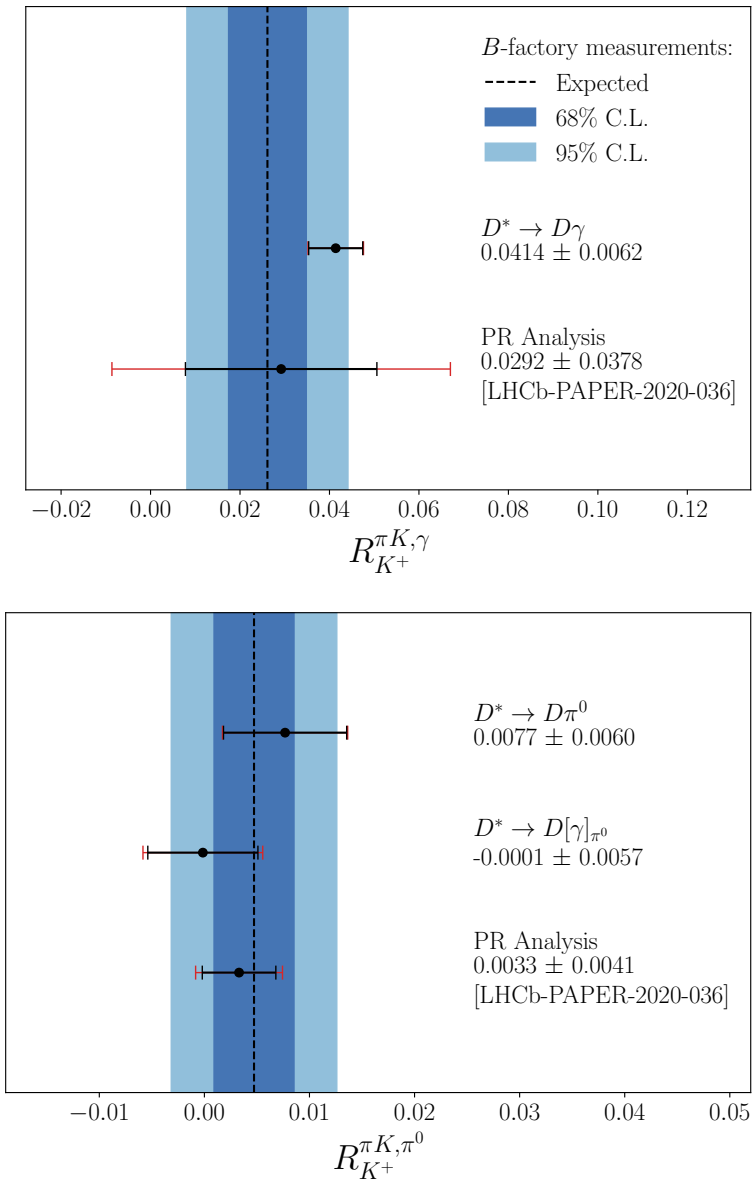


Figure 4.21: Comparison of $R_{K+}^{\pi K,\gamma}$ (top) and $R_{K+}^{\pi K,\pi^0}$ (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from measurements of $r_B^{D^*K}$ and $\delta_B^{D^*K}$ using the ADS/GLW and GGSZ methods by BaBar and Belle [44]. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

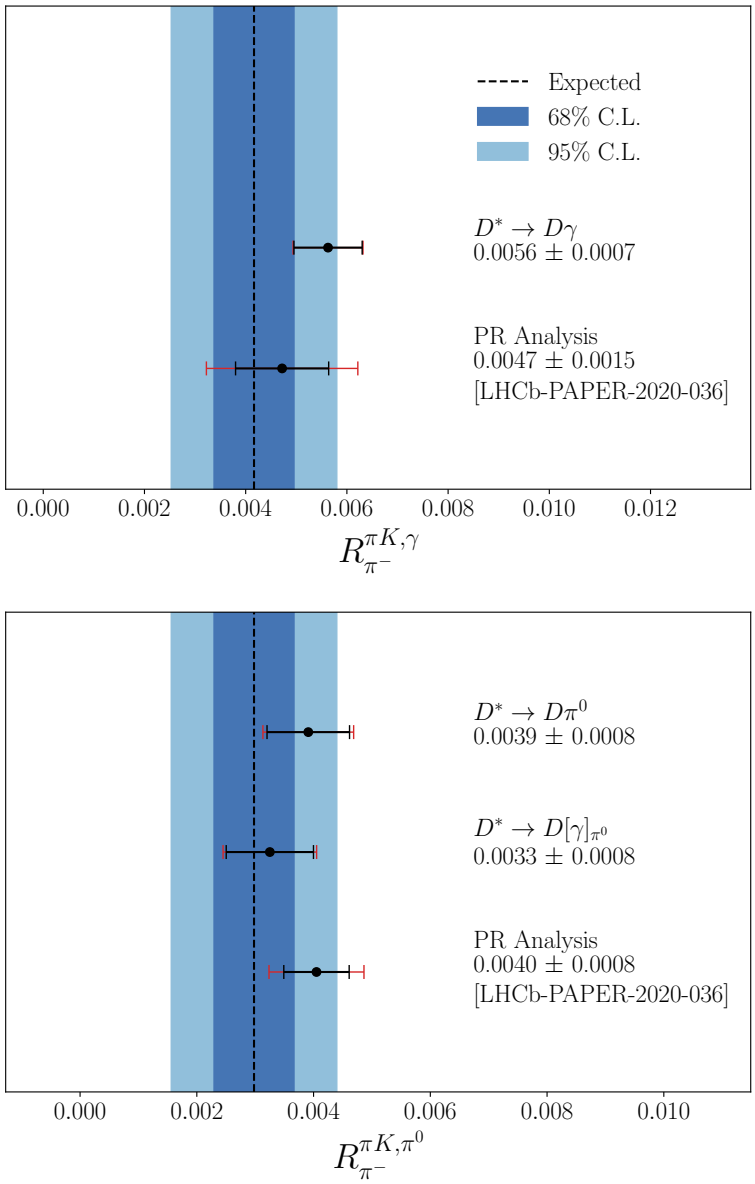


Figure 4.22: Comparison of $R^{\pi K,\gamma}_{\pi^-}$ (top) and $R^{\pi K,\pi^0}_{\pi^-}$ (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from knowledge of the hadronic parameters $r_B^{D^*\pi}$ and $\delta_B^{D^*\pi}$, which are modelled using uniform distributions over the ranges $0 - 0.02$ and $0 - 180°$, respectively. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

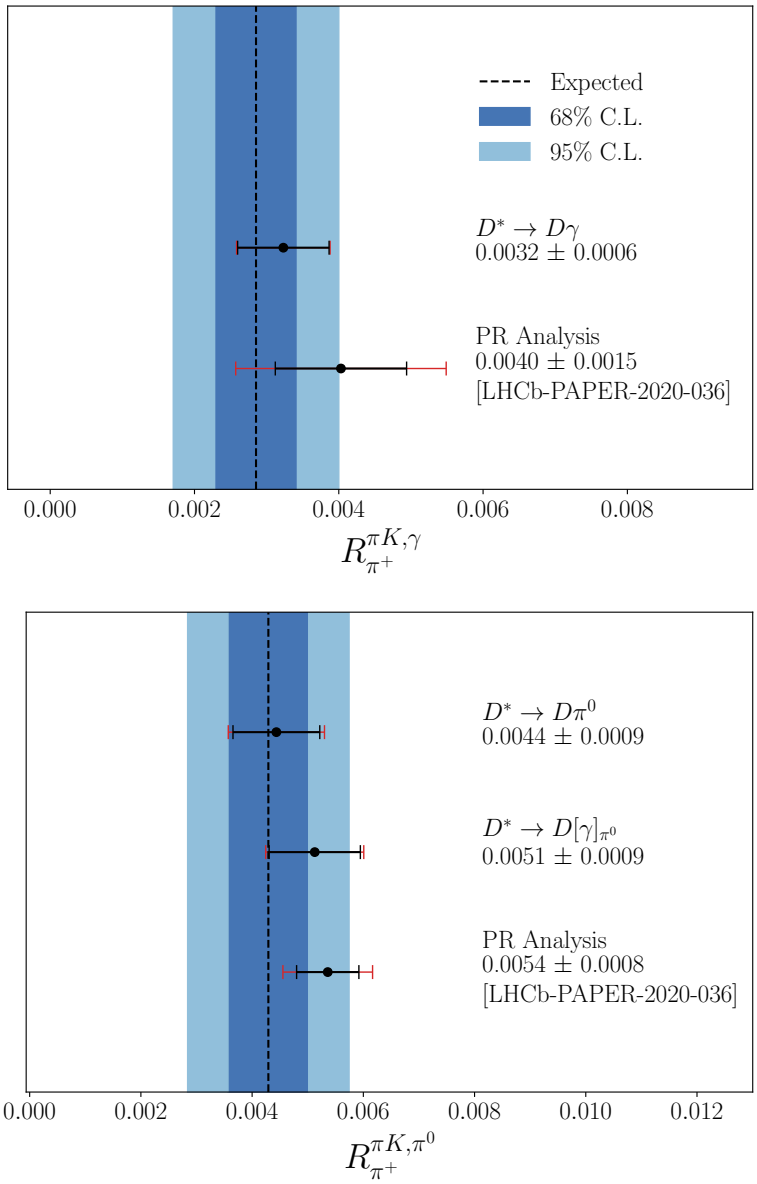


Figure 4.23: Comparison of $R^{\pi K,\gamma}_{\pi^+}$ (top) and $R^{\pi K,\pi^0}_{\pi^+}$ (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from knowledge of the hadronic parameters $r_B^{D^*\pi}$ and $\delta_B^{D^*\pi}$, which are modelled using uniform distributions over the ranges $0-0.02$ and $0-180°$, respectively. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

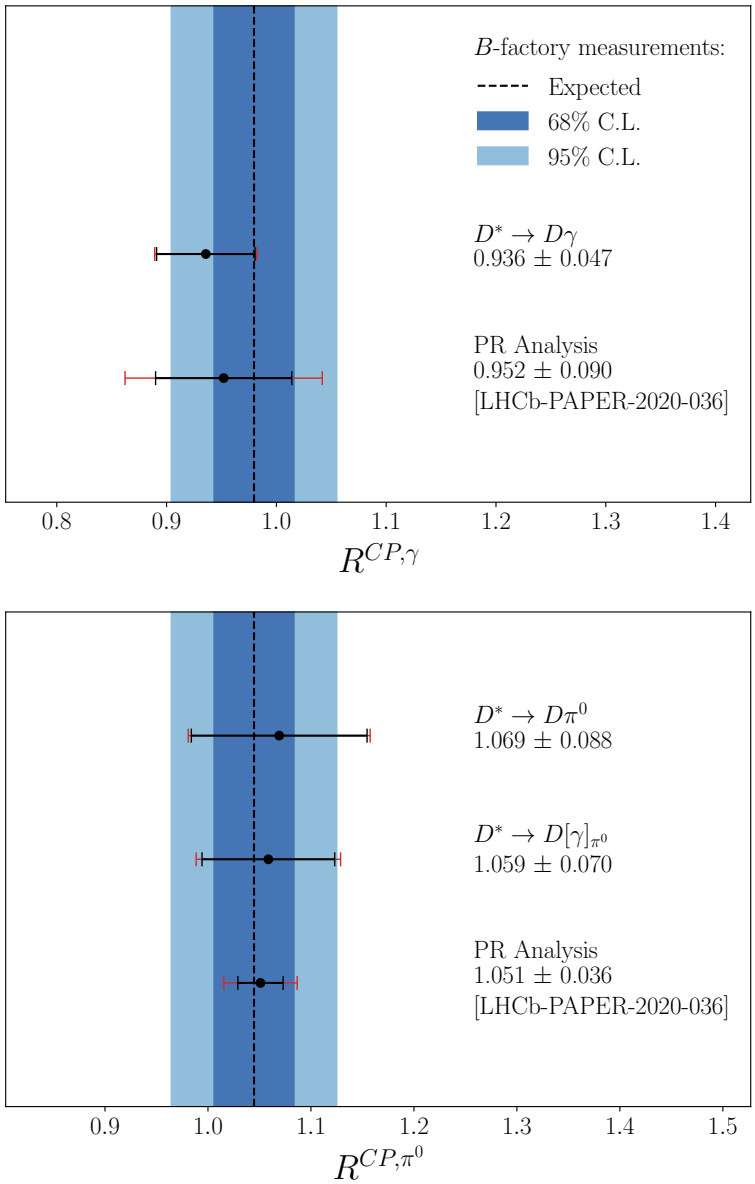


Figure 4.24: Comparison of $R^{CP,\gamma}$ (top) and R^{CP,π^0} (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from measurements of $r_B^{D^*K}$ and $\delta_B^{D^*K}$ using the ADS/GLW and GGSZ methods by BaBar and Belle [44]. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to the measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

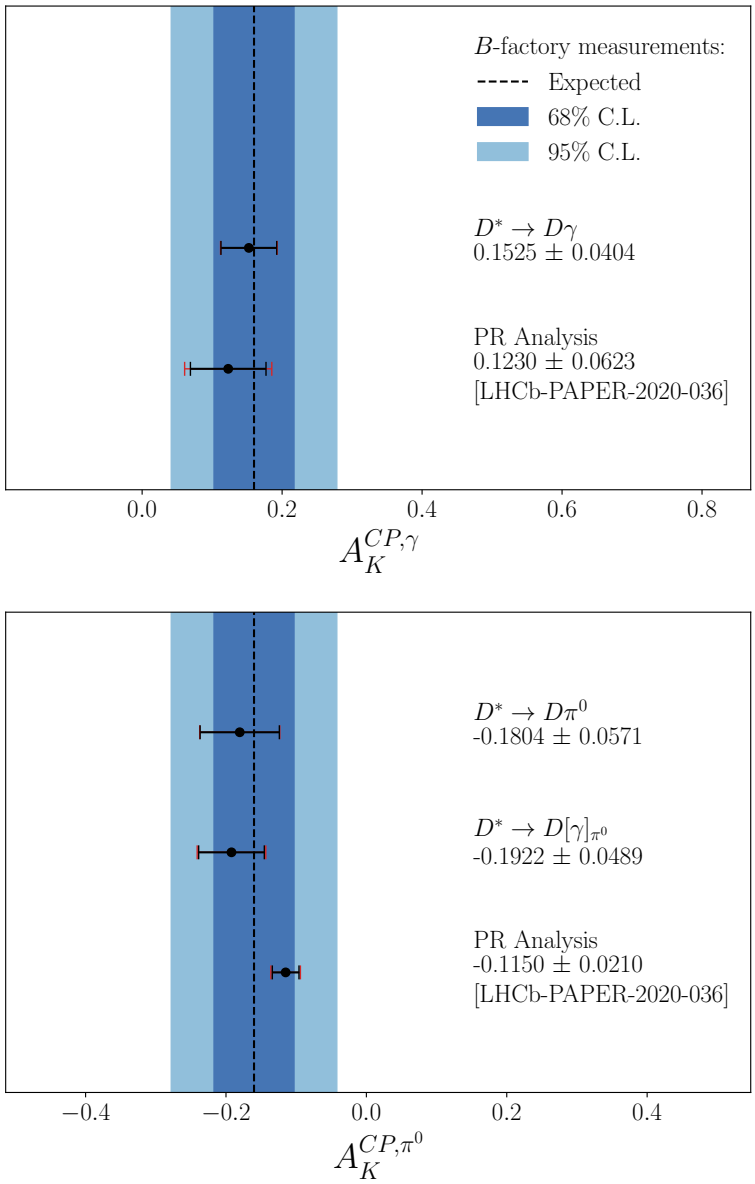


Figure 4.25: Comparison of $A_K^{CP,\gamma}$ (top) and A_K^{CP,π^0} (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from measurements of $r_B^{D^*K}$ and $\delta_B^{D^*K}$ using the ADS/GLW and GGSZ methods by BaBar and Belle [44]. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to the measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

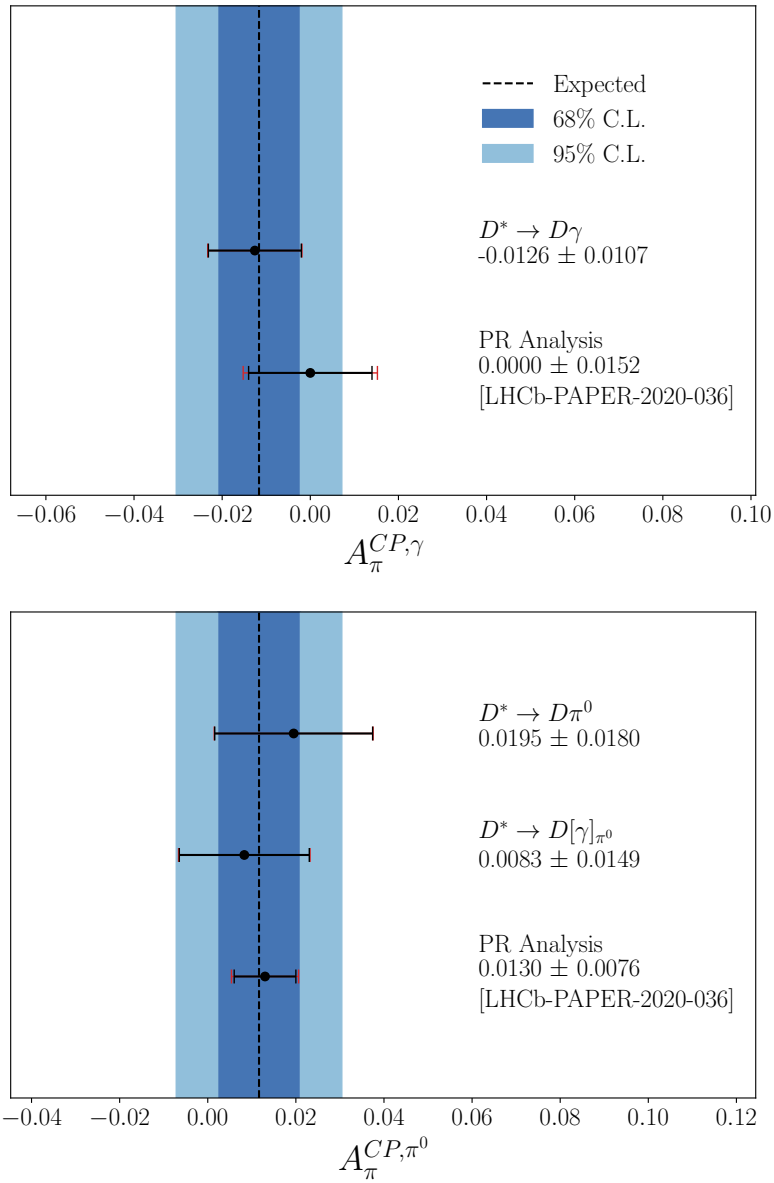


Figure 4.26: Comparison of $A_\pi^{CP,\gamma}$ (top) and A_π^{CP,π^0} (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from knowledge of the hadronic parameters $r_B^{D^*\pi}$ and $\delta_B^{D^*\pi}$, which are modelled using uniform distributions over the ranges $0 - 0.02$ and $0 - 180°$, respectively. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to the measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

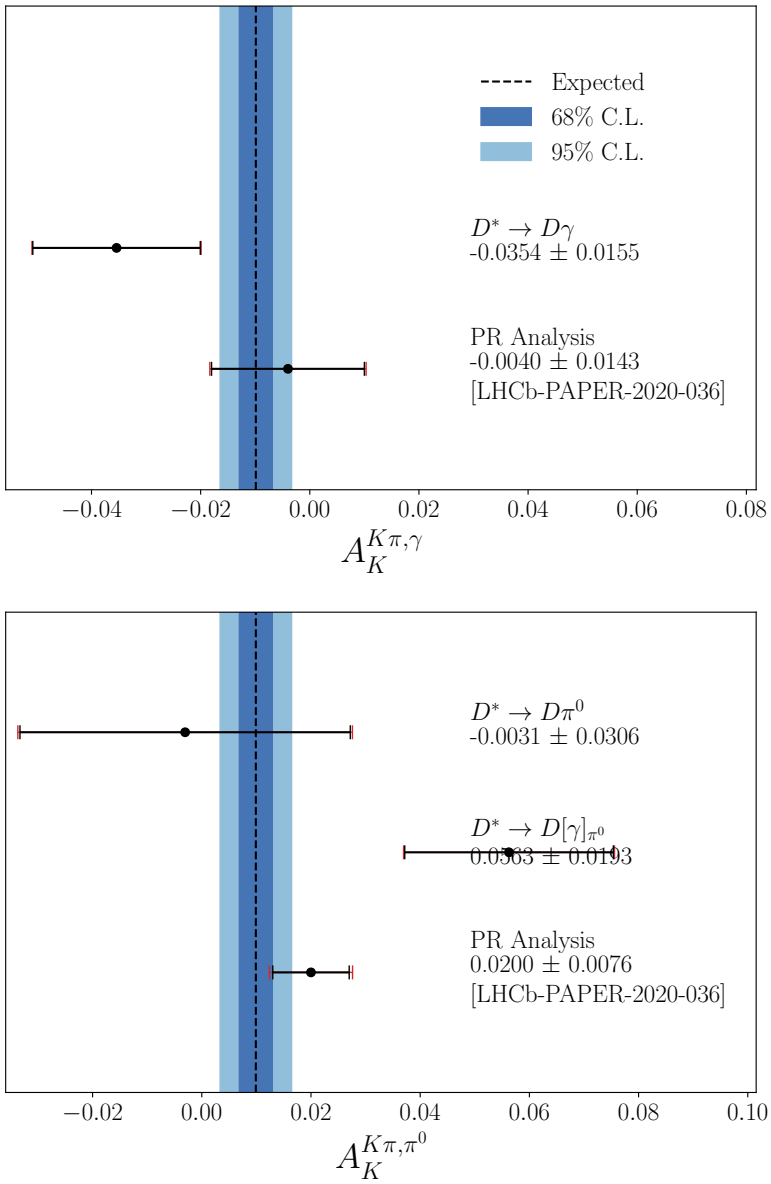


Figure 4.27: Comparison of $A_K^{K\pi,\gamma}$ (top) and $A_K^{K\pi,\pi^0}$ (bottom) with results from the partially reconstructed (PR) analysis [31] and expectations from knowledge of the hadronic parameters $r_B^{D^*\pi}$ and $\delta_B^{D^*\pi}$, which are modelled using uniform distributions over the ranges $0 - 0.02$ and $0 - 180°$, respectively. The result labelled with $D^* \to D\gamma$ corresponds to the measurement from fully reconstructed $D\gamma$ signal; the result labelled with $D^* \to D\pi^0$ corresponds to the measurement from fully reconstructed $D\pi^0$ signal; the result labelled with $D^* \to D[\gamma]_{\pi^0}$ corresponds to the measurement from partially reconstructed $D\pi^0$ signal. The error bars include both statistical (black bar) and systematic (red bar) uncertainties.

5

Interpretation in terms of γ and hadronic parameters

The sensitivity of the measured observables, A_{obs}, to the fundamental parameters $(\gamma, r_B^{D^*K}, \delta_B^{D^*K}, r_B^{D^*\pi}, \delta_B^{D^*\pi})$ are evaluated using the profile likelihood method. The set of parameters, θ, are defined:

$$\theta = (\gamma, r_B^{D^*K}, \delta_B^{D^*K}, r_B^{D^*\pi}, \delta_B^{D^*\pi}, r_D, \delta_D, x, y, \alpha), \tag{5.1}$$

where the charm inputs, r_D, δ_D, x and y are taken from HFLAV 2020 [40], and Gaussian constrained to their values:

$$r_D = (0.3445 \pm 0.0023)\% \tag{5.2}$$

$$\delta_D = (188.9 \pm 8.4)^\circ, \tag{5.3}$$

x and y are given in Eqs. (1.29) and (1.30), respectively; $\alpha = 0.88 \pm 0.03$ is the analysis specific D-mixing correction co-efficient, whose origin is explained in Appendix C. It is assumed that the likelihood function for θ, given A_{obs}, can be represented by a multivariate Gaussian:

$$\mathcal{L}(\theta|A_{obs}) \propto \exp\left(-\frac{1}{2}(A(\theta) - A_{obs})^T V_{cov}^{-1}(A(\theta) - A_{obs})\right), \tag{5.4}$$

where $A(\theta)$ represents the relations between the observables and the fundamental parameters (derived in Sec. 1.4), and V_{cov} is the sum of the statistical and systematic

covariance matrices between the observables, which can be calculated from the correlation matrices given in Sec. 4.6.1. For Gaussian likelihoods, the following relation holds:

$$\chi^2(\theta|A_{obs}) = -2\ln\mathcal{L}(\theta|A_{obs}) \tag{5.5}$$

$$= (A(\theta) - A_{obs})^T V_{cov}^{-1}(A(\theta) - A_{obs}) + c, \tag{5.6}$$

where c is a constant that is independent of θ. It is this χ^2 function that is minimised to determine the best estimates for the central values of the fundamental parameters of interest; the minimum value of the χ^2 function is denoted $\chi^2_{min} \equiv \chi^2(\hat{\theta}|A_{obs})$.

The `PROB` method is used to construct frequentist confidence intervals for some subset, ϕ, of the full parameter set θ; the remaining *nuisance* parameters are denoted $\eta = \theta - \phi$. The confidence level (CL) for a specific set of values, ϕ_0, is determined by again minimising the χ^2 function whilst constraining $\phi = \phi_0$. This results in a new minimum $\hat{\theta}' = (\phi_0, \hat{\eta}')$, which satisfies $\chi^2(\hat{\theta}'|A_{obs}) \geq \chi^2_{min}$. In the approximation that the fundamental parameter estimates $\hat{\theta}$ have been sampled from Gaussians centred on the true values[1], the variable:

$$\Delta\chi^2(\phi_0|A_{obs}) = \chi^2(\hat{\theta}'|A_{obs}) - \chi^2_{min} \tag{5.7}$$

follows a χ^2 distribution with n degrees of freedom, where n represents the number of parameters in the subset ϕ. This process is repeated for different values of ϕ_0, mapping out $\Delta\chi^2$ across the parameter space. The values of $\Delta\chi^2$ can then be converted to confidence levels using the cumulative distribution function of a χ^2 with n degrees of freedom, F_n:

$$\mathrm{CL}(\phi_0|A_{obs}) = F_n(\Delta\chi^2(\phi_0|A_{obs})). \tag{5.8}$$

This method is named `PROB` after the `ROOT` function that performs this conversion; the logic follows that given in the documentation of Ref. [95].

Using only the *CP* observable measurements presented in this thesis, two-dimensional confidence regions for the parameter pairs $\phi = (\gamma, \delta_B^{D^*\pi})$, $(r_B^{D^*\pi}, \delta_B^{D^*\pi})$,

[1]This is the case for maximum likelihood estimates in asymptotically large samples; the datasets considered in this thesis are deemed large enough to give meaningful results.

$(\gamma, \delta_B^{D^*K})$, and $(r_B^{D^*K}, \delta_B^{D^*K})$ are drawn in Fig. 5.1, where the colour intensity of the filled regions represents the CL value at each point in 2D space. The $\Delta\chi^2 = 2.30$, 6.18 and 11.8 contours, corresponding to the 68.3%, 95.5% and 99.7% confidence levels, respectively, are indicated by the red lines. For comparison, the CL contours and best-fit points for the latest LHCb γ combination [23] are also shown in black, where constraints on the hadronic parameters come from the partially reconstructed $B^\pm \to D^{(*)}h^\pm$ analysis [31]. In general, the confidence intervals show good agreement between the γ combination and the results of this analysis.

Confidence regions have also been mapped out in $(\gamma, \delta_B^{D^*K})$ space in Fig. 5.2, comparing results from the fully reconstructed analysis using only measurements of *CP* observables in $B^\pm \to (D^* \to D\pi^0)h^\pm$ decays in (a); using only measurements of *CP* observables in $B^\pm \to (D^* \to D\gamma)h^\pm$ decays in (b); using the results of the partially reconstructed analysis (dominated by decays to the $D\pi^0$ final state) in (c); and using a combination of all 3 in (d). It was initially suggested that the ADS method applied to $B^\pm \to D^*h^\pm$ decays could provide a single, unambiguous measurement of γ [36]. This is due to the strong phase difference of π between the two D^* final states, leading to 4 ADS mode observables, defined by Eqs. (1.27) and (1.28), with only 3 unknowns: γ, $r_B^{D^*h}$ and $\delta_B^{D^*h}$ (the traditional ADS method using $B^\pm \to Dh^\pm$ decays provides just 2 independent equations with 3 unknowns). By comparing plots (a) and (b) in Fig. 5.2, it is evident that the constraints provided by both the $D\pi^0$ and $D\gamma$ final states, respectively, do not resolve the trigonometric ambiguities. Multiple minima for the χ^2 function still exist, and it is not possible to extract a single solution for γ using the ADS/GLW analysis of $B^\pm \to D^*h^\pm$ decays in isolation.

The contour plots of Fig. 5.2 also give evidence that the partially and fully reconstructed analyses posses complementary strengths, and offer powerful constraints on γ and the hadronic parameters when combined. Another χ^2 minimisation is therefore performed, this time using all inputs to the latest LHCb combination [23], alongside the fully reconstructed observables. Best estimates for γ and the hadronic parameters $r_B^{D^*\pi}$, $\delta_B^{D^*\pi}$, $r_B^{D^*K}$ and $\delta_B^{D^*K}$ are found, and the `PROB` method is employed

for each of the parameters individually to extract the 68.3% confidence levels:

$$\gamma = (64.5^{+3.5}_{-3.8})^{\circ}$$
$$r_B^{D^*K} = (11.50^{+0.81}_{-0.86})\%$$
$$\delta_B^{D^*K} = (309.4^{+7.0}_{-8.4})^{\circ}$$
$$r_B^{D^*\pi} = (0.69^{+0.33}_{-0.18})\%$$
$$\delta_B^{D^*\pi} = (102^{+34}_{-47})^{\circ}.$$

Comparing these results to those measured by the latest γ combination, which were given in Chapter 1 and are repeated here as a reference:

$$\gamma = (65.4^{+3.8}_{-4.2})^{\circ}$$
$$r_B^{D^*K} = (9.9^{+1.6}_{-1.9})\%$$
$$\delta_B^{D^*K} = (310^{+12}_{-23})^{\circ}$$
$$r_B^{D^*\pi} = (0.95^{+0.85}_{-0.61})\%$$
$$\delta_B^{D^*\pi} = (139^{+22}_{-86})^{\circ},$$

it can be seen that with the inclusion of the *CP* observables measured in this work, world leading constraints are achieved. It should be noted that constraints on γ come from the many contributing analyses included in the combination; constraints on the D^*h hadronic parameters are dominated by the results of the PR analysis and those presented in this thesis. The best parameter estimates show that with the inclusion of the FR analysis in the γ combination, the uncertainties on γ are reduced by 9%; the uncertainties on $r_B^{D^*K}$ are reduced by 52%; taking a simple average of the positive and negative errors on $\delta_B^{D^*K}$ gives a 66% reduction in the average uncertainty; for $r_B^{D^*\pi}$, the average uncertainty is reduced by 65%; for $\delta_B^{D^*\pi}$, the average uncertainty is reduced by 25%.

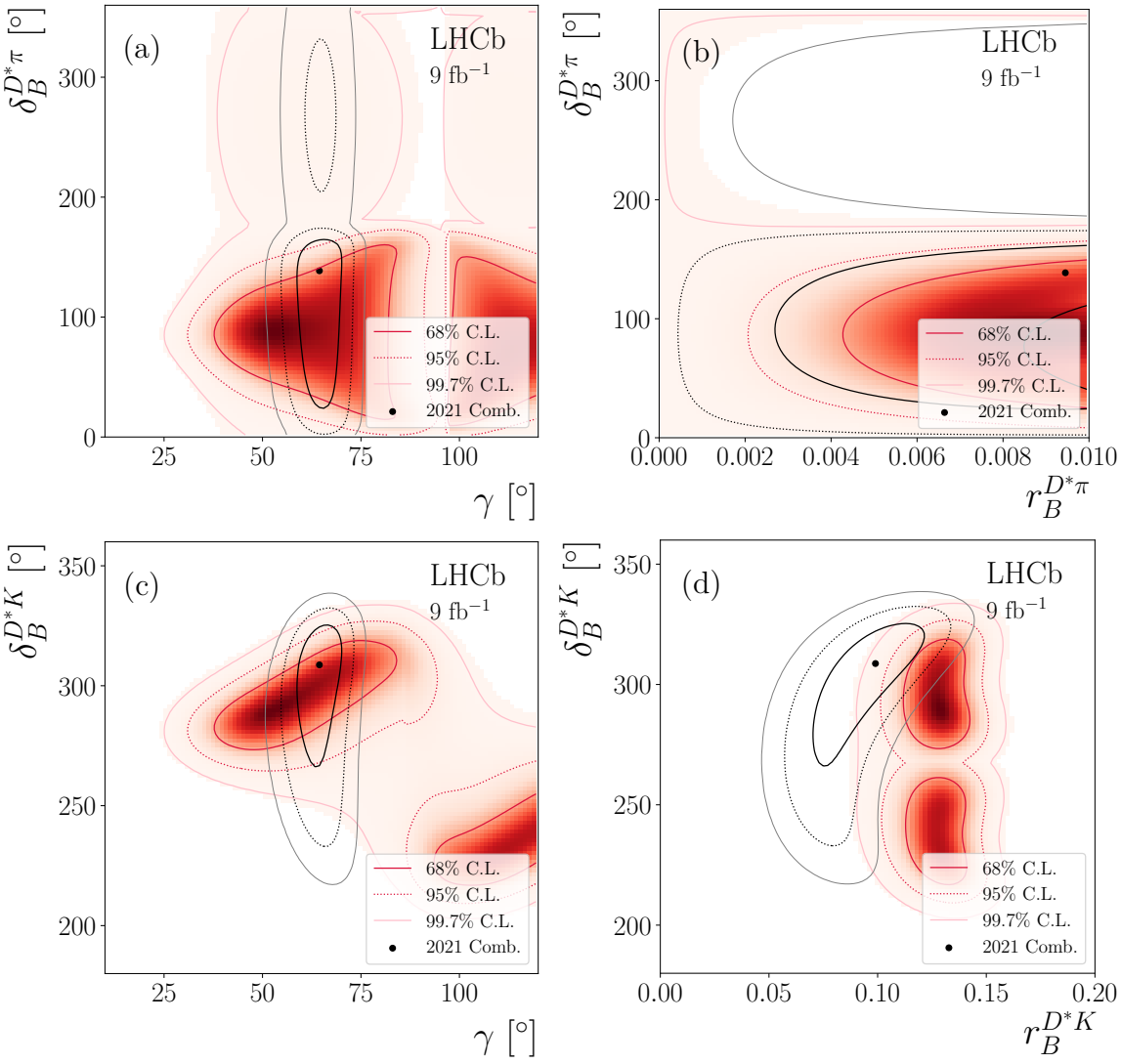


Figure 5.1: 2D confidence regions for the fundamental parameters $(\gamma, r_B^{D^*\pi})$ in (a), $(\delta_B^{D^*\pi}, r_B^{D^*\pi})$ in (b), $(\gamma, r_B^{D^*K})$ in (c) and $(\delta_B^{D^*K}, r_B^{D^*K})$ in (d). The constraints provided by the results of this thesis are indicated in red, and the black/grey lines depict the confidence regions from the latest LHCb γ combination [23].

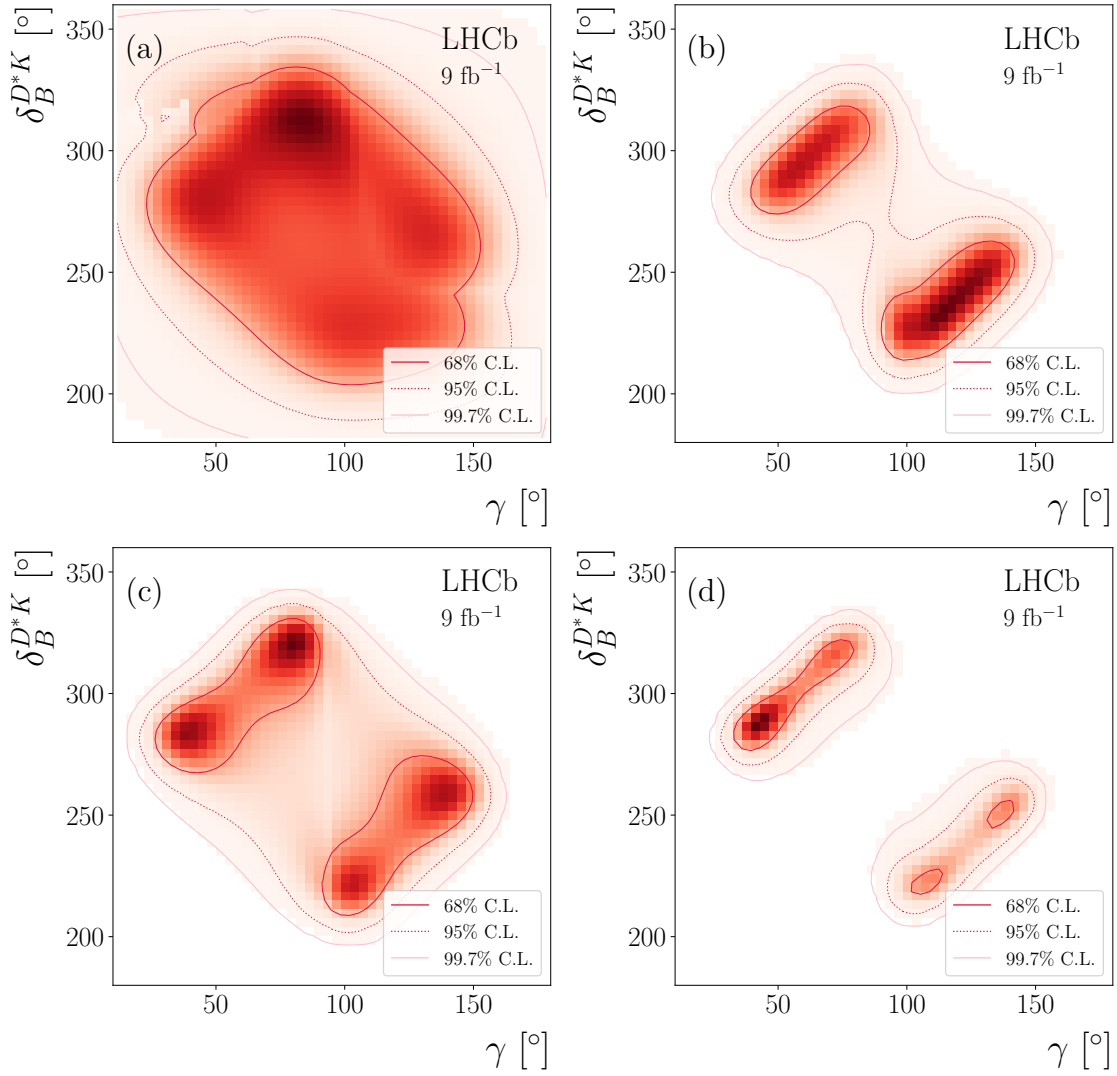


Figure 5.2: 2D confidence regions for the fundamental parameters $(\gamma, \delta_B^{D^*K})$ constrained by the $B^\pm \to (D^* \to D\pi^0)h^\pm$ observables measured using the fully reconstructed analysis in (a), by the $B^\pm \to (D^* \to D\gamma)h^\pm$ observables measured using the fully reconstructed analysis in (b), by the observables measured using the partially reconstructed analysis in (c) (these are dominated by $B^\pm \to (D^* \to D\pi^0)h^\pm$ results) [31], and using the results of both analyses in (d).

5.1 Conclusion

CP violation studies of fully reconstructed $B^{\pm} \to D^{*}h^{\pm}$ decays, $(h \in \pi, K)$, have been conducted for the first time at a hadron collider, where the 2-body D^0 meson final states $K^{-}\pi^{-}$, $K^{-}K^{+}$, $\pi^{-}\pi^{+}$, and $K^{-}\pi^{+}$ are considered. The D^{*} meson is reconstructed in the $D\pi^0$ and $D\gamma$ final states, which possess a strong phase difference of π between them, doubling the number of *CP* violation observables compared to traditional $B^{\pm} \to Dh^{(*)\pm}$ analyses. The data used in this work corresponds to a total integrated luminosity of $8.7\,\text{fb}^{-1}$ of pp collision data collected by the LHCb experiment: $3\,\text{fb}^{-1}$ taken at centre-of-mass energies of 7 TeV and 8 TeV (Run 1); and $5.7\,\text{fb}^{-1}$ at 13 TeV (Run 2).

The first observation of the suppressed $B^{\pm} \to (D^{*} \to [\pi^{\mp}K^{\pm}]_{D}\gamma)K^{\pm}$ decay channel is achieved with a statistical significance of 8.8σ. 16 *CP* violation observables are measured, which, when combined with existing measurements used in the latest LHCb combination [23], offer world leading constraints on the hadronic parameters $(r_B^{D^*K}, \delta_B^{D^*K}, r_B^{D^*\pi}, \delta_B^{D^*\pi})$, and a 9% reduction in the uncertainty on the tree-level determination of the CKM angle γ. These measurements will therefore be of high importance in future combinations, driving down the uncertainty on this Standard Model standard candle.

The techniques developed in this thesis will be reused in future studies of D^{*0} mesons: the D^{*0} reconstruction process; and the error-corrected double-1D fitting method, which is particularly applicable to analyses that require simultaneous fits across multiple correlated variables in order to avoid large systematic uncertainties.

Appendices

A

Bootstrap distributions

The distribution of measurements for each $D^* \to D\pi^0$ observable from fits to 2500 datasets generated using the bootstrapping method [93] are provided in Figs. A.1 and A.2. The standard deviation printed on each plot represents the true statistical uncertainty on the observable, encompassing the effect of double counting.

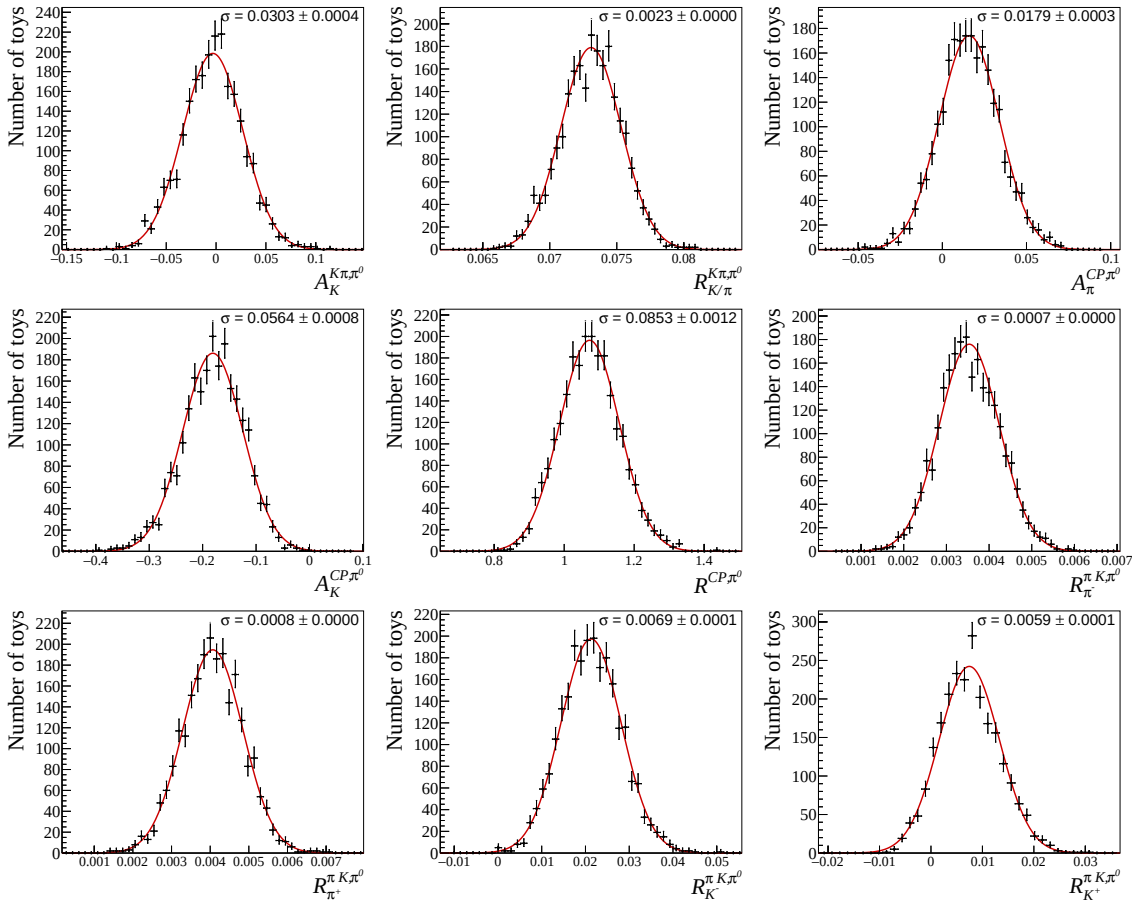


Figure A.1: Bootstrap distributions for each *CP* observable in the $D\pi^0$ data fit.

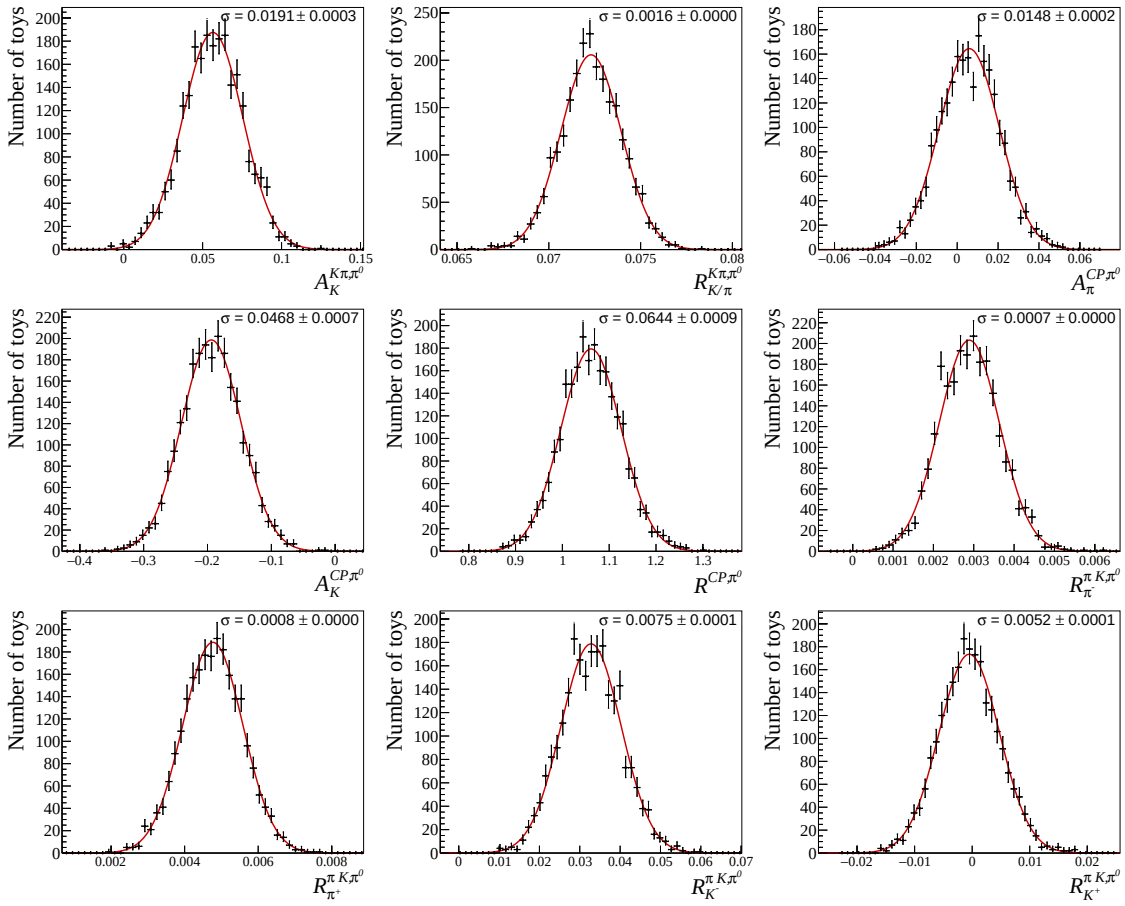


Figure A.2: Bootstrap distributions for each $D^* \to D\pi^0$ *CP* observable in the $D\gamma$ data fit.

B

Breakdown of systematic uncertainties

The sources of systematic uncertainty listed in Sec. 4.5 are broken down individually in Tabs. B.1 to B.2.

	$A_K^{K\pi,\pi^0}$	A_K^{CP,π^0}	A_π^{CP,π^0}
$B^\pm \to (D^* \to D\pi^0)h^\pm$ PDFs	0.0010	0.0026	0.0003
$B^\pm \to (D^* \to D\gamma)h^\pm$ PDFs	0.0000	0.0000	0.0000
Mis. and part. reco. PDFs	0.0042	0.0067	0.0010
$B_s^0 \to D^{(*)}K^\mp\pi^\pm$ PDFs	0.0002	0.0004	0.0001
Mis-ID and crossfeed PDFs	0.0002	0.0012	0.0001
ϵ_{sel}	0.0002	0.0018	0.0001
ϵ_{PID}	0.0002	0.0011	0.0001
Fixed branching fractions and rates	0.0002	0.0012	0.0001
Fixed asymmetries	0.0018	0.0019	0.0017
Fixed CP ratios	0.0002	0.0011	0.0001
Statistical Error Correction	0.0011	0.0022	0.0007
$\sigma(Syst.)$	0.0049	0.0083	0.0020
$\frac{\sigma(Syst.)}{\sigma(Stat.)}$	0.1608	0.1447	0.1147

Table B.1: Systematic uncertainties for each asymmetry measured in the $D\pi^0$ data fit. Uncertainties are given in absolute terms, where the dominant contribution is highlighted in red. The total systematic uncertainty is evaluated by summing all rows in quadrature. The ratio of the total systematic to statistical uncertainty is given in the final row of the table.

	$R_{K^-}^{\pi K,\pi^0}$	$R_{K^+}^{\pi K,\pi^0}$	$R_{\pi^-}^{\pi K,\pi^0}$	$R_{\pi^+}^{\pi K,\pi^0}$	R^{CP,π^0}	$R_{K/\pi}^{K\pi,\pi^0}$
$B^\pm \to (D^* \to D\pi^0)h^\pm$ PDFs	0.0005	0.0001	0.0000	0.0000	0.0035	0.0006
$B^\pm \to (D^* \to D\gamma)h^\pm$ PDFs	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mis. and part. reco. PDFs	0.0008	0.0005	0.0003	0.0004	0.0146	0.0011
$B_s^0 \to D^{(*)}K^\mp\pi^\pm$ PDFs	0.0008	0.0007	0.0000	0.0000	0.0011	0.0000
Mis-ID and crossfeed PDFs	0.0001	0.0001	0.0000	0.0000	0.0017	0.0002
ϵ_{sel}	0.0001	0.0001	0.0000	0.0000	0.0161	0.0010
ϵ_{PID}	0.0001	0.0001	0.0000	0.0000	0.0020	0.0009
Fixed branching fractions and rates	0.0002	0.0001	0.0000	0.0000	0.0031	0.0001
Fixed asymmetries	0.0004	0.0003	0.0000	0.0000	0.0017	0.0000
Fixed CP ratios	0.0004	0.0005	0.0000	0.0000	0.0014	0.0000
Statistical Error Correction	0.0003	0.0002	0.0000	0.0000	0.0032	0.0001
$\sigma(Syst.)$	0.0014	0.0011	0.0003	0.0004	0.0228	0.0019
$\frac{\sigma(Syst.)}{\sigma(Stat.)}$	0.2046	0.1805	0.4555	0.4769	0.2682	0.8424

Table B.2: Systematic uncertainties for each ratio measured in the $D\pi^0$ data fit. Uncertainties are given in absolute terms, where the dominant contribution is highlighted in red. The total systematic uncertainty is evaluated by summing all rows in quadrature. The ratio of the total systematic to statistical uncertainty is given in the final row of the table.

	$A_K^{K\pi,\gamma}$	$A_K^{K\pi,\pi^0}$	$A_K^{CP,\gamma}$	$A_\pi^{CP,\gamma}$	A_K^{CP,π^0}	A_π^{CP,π^0}
$B^\pm \to (D^* \to D\pi^0)h^\pm$ PDFs	0.0001	0.0003	0.0008	0.0001	0.0022	0.0002
$B^\pm \to (D^* \to D\gamma)h^\pm$ PDFs	0.0001	0.0006	0.0005	0.0001	0.0015	0.0002
Mis. and part. reco. PDFs	0.0011	0.0018	0.0058	0.0006	0.0101	0.0015
$B_s^0 \to D^{(*)}K^\mp\pi^\pm$ PDFs	0.0000	0.0000	0.0001	0.0000	0.0003	0.0000
Mis-ID and crossfeed PDFs	0.0001	0.0002	0.0002	0.0001	0.0006	0.0001
ϵ_{sel}	0.0001	0.0001	0.0003	0.0000	0.0006	0.0000
ϵ_{PID}	0.0002	0.0001	0.0001	0.0000	0.0003	0.0001
Fixed branching fractions and rates	0.0001	0.0004	0.0006	0.0001	0.0011	0.0001
Fixed asymmetries	0.0019	0.0020	0.0035	0.0016	0.0093	0.0016
Fixed CP ratios	0.0000	0.0001	0.0000	0.0000	0.0007	0.0000
Statistical Error Correction	0.0005	0.0008	0.0012	0.0003	0.0020	0.0006
$\sigma(Syst.)$	0.0022	0.0029	0.0069	0.0018	0.0142	0.0023
$\frac{\sigma(Syst.)}{\sigma(Stat.)}$	0.1430	0.1508	0.1729	0.1692	0.3036	0.1563

Table B.3: Systematic uncertainties for each asymmetry measured in the $D\gamma$ data fit. Uncertainties are given in absolute terms, where the dominant contribution is highlighted in red. The total systematic uncertainty is evaluated by summing all rows in quadrature. The ratio of the total systematic to statistical uncertainty is given in the final row of the table.

	$R^{\pi K,\gamma}_{K^-}$	$R^{\pi K,\gamma}_{K^+}$	$R^{\pi K,\gamma}_{\pi^-}$	$R^{\pi K,\gamma}_{\pi^+}$	$R^{\pi K,\pi^0}_{K^-}$	$R^{\pi K,\pi^0}_{K^+}$	$R^{\pi K,\pi^0}_{\pi^-}$	$R^{\pi K,\pi^0}_{\pi^+}$	$R^{CP,\gamma}$	R^{CP,π^0}	$R^{K\pi,\gamma}_{K/\pi}$	$R^{K\pi,\pi^0}_{K/\pi}$
$B^\pm \to (D^* \to D\pi^0)h^\pm$ PDFs	0.0001	0.0001	0.0000	0.0000	0.0001	0.0002	0.0000	0.0000	0.0014	0.0052	0.0001	0.0004
$B^\pm \to (D^* \to D\gamma)h^\pm$ PDFs	0.0001	0.0001	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	0.0014	0.0015	0.0004	0.0001
Mis. and part. reco. PDFs	0.0012	0.0008	0.0001	0.0001	0.0011	0.0013	0.0003	0.0003	0.0083	0.0232	0.0004	0.0009
$B^0_s \to D^{(*)}K^\mp\pi^\pm$ PDFs	0.0006	0.0007	0.0000	0.0000	0.0018	0.0017	0.0000	0.0000	0.0009	0.0017	0.0000	0.0000
Mis-ID and crossfeed PDFs	0.0001	0.0001	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	0.0007	0.0013	0.0002	0.0001
ϵ_{sel}	0.0000	0.0001	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	0.0079	0.0091	0.0006	0.0006
ϵ_{PID}	0.0000	0.0002	0.0000	0.0000	0.0002	0.0001	0.0000	0.0000	0.0025	0.0078	0.0008	0.0009
Fixed branching fractions and rates	0.0001	0.0003	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	0.0015	0.0027	0.0002	0.0002
Fixed asymmetries	0.0004	0.0005	0.0000	0.0000	0.0005	0.0005	0.0000	0.0000	0.0022	0.0056	0.0000	0.0001
Fixed CP ratios	0.0002	0.0006	0.0000	0.0000	0.0003	0.0006	0.0000	0.0000	0.0021	0.0027	0.0000	0.0001
Statistical Error Correction	0.0001	0.0002	0.0000	0.0000	0.0003	0.0002	0.0000	0.0000	0.0013	0.0028	0.0000	0.0001
$\sigma(Syst.)$	0.0014	0.0014	0.0001	0.0001	0.0023	0.0022	0.0003	0.0003	0.0125	0.0277	0.0012	0.0015
$\frac{\sigma(Syst.)}{\sigma(Stat.)}$	0.2919	0.2348	0.2204	0.2120	0.3017	0.4191	0.3862	0.3931	0.2797	0.4325	0.9940	0.9169

Table B.4: Systematic uncertainties for each ratio measured in the $D\gamma$ data fit. Uncertainties are given in absolute terms, where the dominant contribution is highlighted in red. The total systematic uncertainty is evaluated by summing all rows in quadrature. The ratio of the total systematic to statistical uncertainty is given in the final row of the table.

C

D lifetime acceptance functions

The 'ideal-case' corrections for D mixing are calculated assuming a uniform selection efficiency over D lifetime [39]. In practice, it is necessary to measure the analysis specific D lifetime acceptance functions in order to correct the 'ideal-case' equations. This correction takes the form of the multiplicative factor α, as shown in Eqs. (1.27) and (1.28).

To determine the functional form of the D decay time acceptance function for each of the signal modes, fits to the D proper time distributions of signal simulation samples are performed. The fit PDFs are composite: an exponential function is convoluted with a Gaussian resolution model, where the τ parameter of the exponential is fixed to the known D lifetime ($\tau_D = 0.41\,\text{ps}$) and the mean and width of the resolution function vary freely. This function is then multiplied by an acceptance function of the form

$$f(t) = c \times \frac{(a \times t)^n}{1 + (a \times t)^n} \times e^{-t/b}, \tag{C.1}$$

where t is the D proper time, c is a normalisation term, the parameters a and n describe the loss of events at low lifetime values, and b accounts for a loss of events at higher lifetimes. In the fits, all acceptance function parameters vary freely.

Fits to simulation samples of fully reconstructed $B^\pm \to (D^* \to [K^\pm\pi^\mp]_D\gamma)\pi^\pm$ decays, partially reconstructed $B^\pm \to (D^* \to [K^\pm\pi^\mp]_D\pi^0)\pi^\pm$ decays and fully

reconstructed $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ decays are shown in Fig. C.1, with the measured acceptance functions displayed underneath. The acceptance function parameters a, b and n are given in Tab. C.1, as are the means and widths of the Gaussian resolution models. It can be seen that the D meson selection efficiency falls at low decay times; this is due to the D flight distance significance cut detailed in Sec. 3.1.3, used to remove backgrounds from wrong combinations of tracks.

Using these functional forms, the D system can be integrated over time for each signal mode in order to extract the 'ideal-case' correction factor for D-mixing terms. The errors on each correction are evaluated by varying the acceptance function parameters within the given uncertainties. The individual values obtained for each signal channel are consistent within 1σ, therefore the weighted average, calculated using Eqs. (4.33) and (4.34), is taken to be the correction co-efficient for the D-mixing terms of ADS observables:

$$\alpha = 0.88 \pm 0.03. \tag{C.2}$$

Parameter	Value (MC)	Parameter	Value (MC)	Parameter	Value (MC)
b [ps]	1.52 ± 0.17	b [ps]	1.24 ± 0.14	b [ps]	1.35 ± 0.16
a [ps^{-1}]	25.61 ± 1.17	a [ps^{-1}]	26.32 ± 0.90	a [ps^{-1}]	31.58 ± 2.07
n	1.85 ± 0.08	n	2.01 ± 0.08	n	2.09 ± 0.16
μ [ps]	$-\ 1.03 \pm 0.37$	μ [ps]	$-\ 1.22 \pm 0.29$	μ [ps]	$-\ 0.55 \pm 0.30$
σ [ps]	0.67 ± 0.13	σ [ps]	0.75 ± 0.10	σ [ps]	0.50 ± 0.12

Table C.1: Fitted parameter values from D lifetime fits to simulation samples of fully reconstructed $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\gamma)\pi^{\pm}$ decays on the left, partially reconstructed $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ decays in the centre and fully reconstructed $B^{\pm} \to (D^{*} \to [K^{\pm}\pi^{\mp}]_{D}\pi^{0})\pi^{\pm}$ decays on the right.

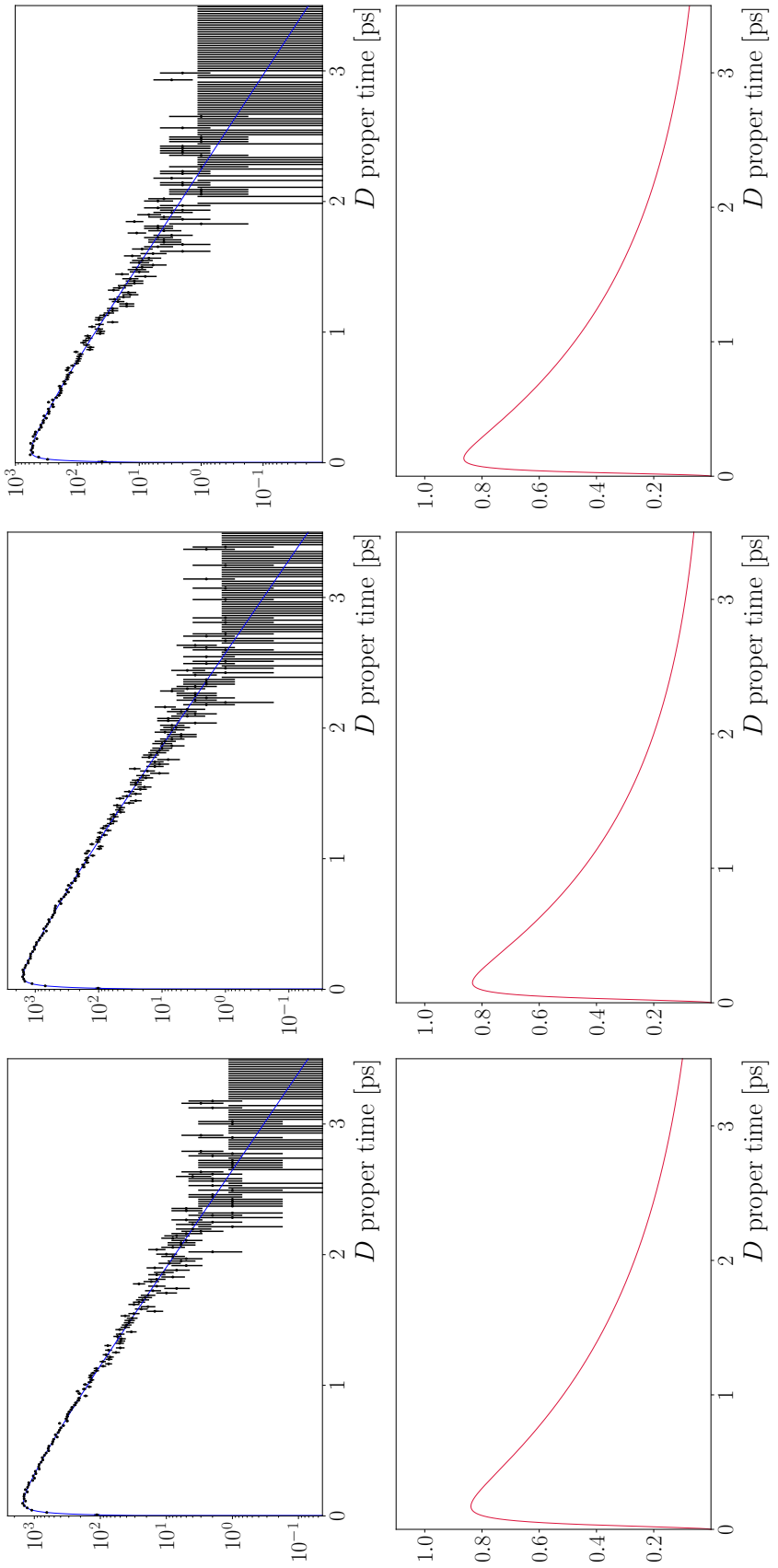


Figure C.1: A fit to the D proper lifetime distribution in simulated fully reconstructed $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_{D\gamma})\pi^{\pm}$ events (left), partially reconstructed $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_{D\pi^0})\pi^{\pm}$ events (centre) and fully reconstructed $B^{\pm} \to (D^* \to [K^{\pm}\pi^{\mp}]_{D\pi^0})\pi^{\pm}$ events (right). The corresponding acceptance functions are shown below, with parameter values fixed to those found in the lifetime fits.

References

[1] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. "Matter and antimatter in the universe". In: *New Journal of Physics* 14.9 (Sept. 2012), p. 095012. URL: http://dx.doi.org/10.1088/1367-2630/14/9/095012.

[2] Andrei D Sakharov. "Violation CP variance, C asymmetry, and baryon asymmetry of the universe". In: 34.5 (May 1991), pp. 392–393. URL: https://doi.org/10.1070/pu1991v034n05abeh002497.

[3] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[4] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. URL: https://doi.org/10.1143/PTP.49.652.

[5] Emmy Noether. "Invariant Variation Problems". In: *Gott. Nachr.* 1918 (1918), pp. 235–257. arXiv: physics/0503066.

[6] G Lüders. "On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories". In: *Dan. Mat. Fys. Medd.* 28 (1954), pp. 1–17. URL: https://cds.cern.ch/record/1071765.

[7] C. S. Wu et al. "Experimental Test of Parity Conservation in β Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[8] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[9] Bernard Aubert et al. "Observation of CP violation in the B^0 meson system". In: *Phys. Rev. Lett.* 87 (2001), p. 091801. arXiv: hep-ex/0107013.

[10] Kazuo Abe et al. "Observation of large CP violation in the neutral B meson system". In: *Phys. Rev. Lett.* 87 (2001), p. 091802. arXiv: hep-ex/0107061.

[11] K. Trabelsi. "Study of direct CP in charmed B decays and measurement of the CKM angle gamma at Belle". In: (2013). arXiv: 1301.2033 [hep-ex].

[12] Denis Derkach. "Combination of gamma measurements from BaBar". In: (2013). arXiv: 1301.3283 [hep-ex].

[13] R. Aaij et al. "Observation of CP violation in $B^\pm \to DK^\pm$ decays". In: *Physics Letters B* 712.3 (2012), pp. 203–212. URL: https://www.sciencedirect.com/science/article/pii/S0370269312004893.

[14] R Aaij et al. "First observation of CP violation in the decays of B_s^0 mesons". In: *Phys. Rev. Lett.* 110.22 (2013), p. 221601. arXiv: 1304.6173 [hep-ex].

[15] R Aaij et al. "Observation of CP Violation in Charm Decays". In: *Phys. Rev. Lett.* 122 (21 May 2019), p. 211803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.211803.

[16] K. Abe et al. "Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations". In: *Nature* 580.7803 (Apr. 2020), pp. 339–344. URL: http://dx.doi.org/10.1038/s41586-020-2177-0.

[17] C. L. Bennett et al. "Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Result". In: *The Astrophysical Journal Supplement Series* 208.2 (Sept. 2013), p. 20. URL: http://dx.doi.org/10.1088/0067-0049/208/2/20.

[18] Ling-Lie Chau and Wai-Yee Keung. "Comments on the Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett.* 53 (19 Nov. 1984), pp. 1802–1805. URL: https://link.aps.org/doi/10.1103/PhysRevLett.53.1802.

[19] P.A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. URL: https://pdg.lbl.gov/.

[20] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[21] Lincoln Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett.* 51 (21 Nov. 1983), pp. 1945–1947. URL: https://link.aps.org/doi/10.1103/PhysRevLett.51.1945.

[22] Joachim Brod and Jure Zupan. "The ultimate theoretical error on γ from $B \to DK$ decays". In: *JHEP* 01 (2014), p. 051. arXiv: 1308.5663 [hep-ph].

[23] Roel Aaij et al. "Simultaneous determination of CKM angle γ and charm mixing parameters". In: (Oct. 2021). arXiv: 2110.02350 [hep-ex].

[24] J. Charles et al. "Current status of the standard model CKM fit and constraints on $\Delta F = 2$ new physics". In: *Physical Review D* 91.7 (Apr. 2015). URL: http://dx.doi.org/10.1103/PhysRevD.91.073007.

[25] R Aaij et al. "Test of lepton universality in beauty-quark decays". In: (2021). arXiv: 2103.11769 [hep-ex].

[26] R Aaij et al. "Tests of lepton universality using $B^0 \to K^0_S \ell^+ \ell^-$ and $B^+ \to K^{*+} \ell^+ \ell^-$ decays". In: (Oct. 2021). arXiv: 2110.09501.

[27] R Aaij et al. "Test of lepton universality with $\Lambda_b^0 \to pK^- \ell^+ \ell^-$ decays". In: *JHEP* 2020 (Dec. 2019), 040. 26 p. arXiv: 1912.08139.

[28] A. Abdesselam and orhters. "Test of lepton flavor universality and search for lepton flavor violation in $B \to K \ell\ell$ decays". In: *Journal of high energy physics* 3 (Mar. 2021), 105 (1–19). arXiv: 1908.01848. URL: https://doi.org/10.1007/JHEP03(2021)105.

[29] S. Wehle et al. "Test of Lepton-Flavor Universality in $B \to K^* \ell^+ \ell^-$ Decays at Belle". In: *Phys. Rev. Lett.* 126 (16 Apr. 2021), p. 161801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.126.161801.

[30] Claudia Cornella et al. "Reading the footprints of the B-meson flavor anomalies". In: *Journal of High Energy Physics* 2021.8 (Aug. 2021). URL: http://dx.doi.org/10.1007/JHEP08(2021)050.

[31] R. Aaij et al. "Measurement of CP observables in $B^{\pm} \to D^{(*)}K^{\pm}$ and $B^{\pm} \to D^{(*)}\pi^{\pm}$ decays using two-body D final states". In: (2020). arXiv: 2012.09903 [hep-ex].

[32] R Aaij et al. "Measurement of the CKM angle γ in $B^{\pm} \to DK^{\pm}$ and $B^{\pm} \to D\pi^{\pm}$ decays with $D \to K^0_S h^+ h^-$". In: *JHEP* 2102 (Oct. 2020), 169. 36 p. arXiv: 2010.08483. URL: https://cds.cern.ch/record/2742273.

[33] Anjan Giri et al. "Determining gamma using $B^{\pm} \to DK^{\pm}$ with multibody D decays". In: *Phys. Rev. D* 68 (2003), p. 054018. arXiv: hep-ph/0303187.

[34] A Bondar. *Proceedings of BINP special analysis meeting on Dalitz analysis, 24-26 Sep. 2002*. unpublished. 2002.

[35] A. Poluektov et al. "Measurement of ϕ_3 with Dalitz plot analysis of $B^{\pm} \to D^{(*)}K^{\pm}$ decay". In: *Phys. Rev. D* 70 (2004), p. 072003. arXiv: hep-ex/0406067.

[36] Alex Bondar and Tim Gershon. "On ϕ_3 measurements using $B^- \to D^*K^-$ decays". In: *Phys. Rev. D* 70 (9 Nov. 2004), p. 091503. URL: https://link.aps.org/doi/10.1103/PhysRevD.70.091503.

[37] Michael Gronau and Daniel Wyler. "On determining a weak phase from charged B decay asymmetries". In: *Physics Letters B* 265.1 (1991), pp. 172–176. URL: https://www.sciencedirect.com/science/article/pii/037026939190034N.

[38] David Atwood, Isard Dunietz, and Amarjit Soni. "Enhanced CP Violation with $B \to KD^0(\overline{D}^0)$ Modes and Extraction of the Cabibbo-Kobayashi-Maskawa Angle γ". In: *Physical Review Letters* 78.17 (Apr. 1997), pp. 3257–3260. URL: http://dx.doi.org/10.1103/PhysRevLett.78.3257.

[39] Matteo Rama. "Effect of $D - \overline{D}$ mixing in the extraction of γ with $B^- \to D^0K^-$ and $B^- \to D^0\pi^-$ decays". In: *Phys. Rev. D* 89.1 (2014), p. 014021. arXiv: 1307.4384 [hep-ex].

[40] Yasmine Sara Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018". In: (2019). updated results and plots available at https://hflav.web.cern.ch/. arXiv: 1909.12524 [hep-ex].

[41] K. Abe et al. "Study of $B^{\pm} \to D(CP)K^{\pm}$ and $D^*(CP)K^{\pm}$ decays". In: *Physical Review D* 73.5 (Mar. 2006). URL: http://dx.doi.org/10.1103/PhysRevD.73.051106.

[42] Bernard Aubert et al. "Measurement of Ratios of Branching Fractions and CP-Violating Asymmetries of $B^{\pm} \to D^*K^{\pm}$ Decays". In: *Phys. Rev. D* 78 (2008), p. 092002. arXiv: 0807.2408 [hep-ex].

[43] P. del Amo Sanchez et al. "Evidence for Direct CP Violation in the Measurement of the Cabbibo-Kobayashi-Maskawa Angle γ with $B^{\pm} \to D^{(*)}K^{\pm}$ Decays". In: *Physical Review Letters* 105.12 (Sept. 2010). URL: http://dx.doi.org/10.1103/PhysRevLett.105.121801.

[44] 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". In: *Physical Review D* 87.5 (Mar. 2013). URL: http://dx.doi.org/10.1103/PhysRevD.87.052015.

[45] The LHCb Collaboration. "The LHCb Detector at the LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005–S08005. URL: https://doi.org/10.1088/1748-0221/3/08/s08005.

[46] Christian Elsässer. *$\bar{b}b$ production angle plots*. URL: %7B%5Chref%7Bhttps://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb%5C_ProductionAngles.html%7D%7Bhttps://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb%5C_ProductionAngles.html%7D%7D.

[47] Cinzia De Melis. "The CERN accelerator complex. Complexe des accélérateurs du CERN". In: (Jan. 2016). General Photo. URL: https://cds.cern.ch/record/2119882.

[48] R Aaij et al. "Performance of the LHCb Vertex Locator". In: *Journal of Instrumentation* 9.09 (Sept. 2014), P09007–P09007. URL: http://dx.doi.org/10.1088/1748-0221/9/09/P09007.

[49] P R Barbosa-Marinho et al. "LHCb VELO (VErtex LOcator): Technical Design Report". In: Technical design report. LHCb (2001). URL: https://cds.cern.ch/record/504321.

[50] S Amato et al. "LHCb magnet: Technical Design Report". In: Technical design report. LHCb (2000). URL: https://cds.cern.ch/record/424338.

[51] Roel Aaij et al. "LHCb Detector Performance". In: *Int. J. Mod. Phys. A* 30.07 (2015), p. 1530022. arXiv: 1412.6352 [hep-ex].

[52] P R Barbosa-Marinho et al. "LHCb inner tracker: Technical Design Report". In: Technical design report. LHCb (2002). URL: https://cds.cern.ch/record/582793.

[53] Ph. d'Argent et al. "Improved performance of the LHCb Outer Tracker in LHC Run 2". In: *JINST* 12.11 (2017), P11016. arXiv: 1708.00819 [physics.ins-det].

[54] P R Barbosa-Marinho et al. "LHCb outer tracker: Technical Design Report". In: Technical design report. LHCb (2001). URL: https://cds.cern.ch/record/519146.

[55] Antonis Papanestis and Carmelo D'Ambrosio. "Performance of the LHCb RICH detectors during the LHC Run II". In: *Nucl. Instrum. Meth. A* 876 (2017). Ed. by P. Krizan et al., pp. 221–224. arXiv: 1703.08152 [physics.ins-det].

[56] S Amato et al. "LHCb RICH: Technical Design Report". In: Technical design report. LHCb (2000). URL: https://cds.cern.ch/record/494263.

[57] M. Adinolfi et al. "Performance of the LHCb RICH detector at the LHC". In: *Eur. Phys. J. C* 73 (2013), p. 2431. arXiv: 1211.6759 [physics.ins-det].

[58] Carlos Abellán Beteta et al. "Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC". In: (Aug. 2020). arXiv: 2008.11556 [physics.ins-det].

[59] A Arefev et al. "Beam Test Results of the LHCb Electromagnetic Calorimeter." In: (May 2008). URL: https://cds.cern.ch/record/1103500.

[60] A. A. Alves Jr. et al. "Performance of the LHCb muon system". In: *JINST* 8 (2013), P02022. arXiv: 1211.1346 [physics.ins-det].

[61] P R Barbosa-Marinho et al. "LHCb muon system: Technical Design Report". In: Technical design report. LHCb (2001). URL: https://cds.cern.ch/record/504326.

[62] "LHCb muon system: second addendum to the Technical Design Report". In: Technical design report. LHCb (2005). Submitted on 9 Apr 2005. URL: https://cds.cern.ch/record/831955.

[63] Roel Aaij et al. "Measurement of the $B^{\pm}$ production cross-section in pp collisions at $\sqrt{s} = 7$ and 13 TeV". In: *JHEP* 12 (2017), p. 026. arXiv: 1710.04921 [hep-ex].

[64] David Hutchcroft. "VELO pattern recognition". In: (Mar. 2007). URL: https://cds.cern.ch/record/1023540/files/lhcb-2007-013.pdf.

[65] M Needham and J Van Tilburg. "Performance of the track matching". In: (Mar. 2007). URL: https://cds.cern.ch/record/1020304.

[66] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth. A* 262.2 (1987), pp. 444–450. URL: https://www.sciencedirect.com/science/article/pii/0168900287908874.

[67] J Van Tilburg. *Track simulation and reconstruction in LHCb*. 2005. URL: https://cds.cern.ch/record/885750.

[68] Andreas Jaeger et al. "Measurement of the track finding efficiency". In: (Apr. 2012).

[69] R. Forty. "RICH pattern recognition for LHCb". In: *Nucl. Instrum. Meth. A* 433 (1999). Ed. by A. Breskin, R. Chechik, and T. Ypsilantis, pp. 257–261.

[70] A Powell. "Particle Identification at LHCb. Particle ID in LHCb". In: (Jan. 2011), 5 p. URL: https://cds.cern.ch/record/1322666.

[71] LHCb collaboration. *PID performance plots*. https://twiki.cern.ch/twiki/bin/view/LHCb/PIDConferencePlot.

[72] O Deschamps et al. "Photon and neutral pion reconstruction". In: (Sept. 2003). URL: https://cds.cern.ch/record/691634.

[73] "Selected HLT2 reconstruction performance for the LHCb upgrade". In: (June 2021). URL: https://cds.cern.ch/record/2773174.

[74] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. arXiv: hep-ph/0603175.

[75] I Belyaev et al. *Handling of the generation of primary events in Gauss, the LHCb simulation framework*. Tech. rep. Geneva: CERN, Nov. 2010. URL: https://cds.cern.ch/record/1307917.

[76] David J. Lange. "The EvtGen particle decay simulation package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. URL: https://www.sciencedirect.com/science/article/pii/S0168900201000894.

[77] Piotr Golonka and Zbigniew Was. "PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays". In: *Eur. Phys. J. C* 45 (2006), pp. 97–107. arXiv: hep-ph/0506026.

[78] J. Allison et al. "Geant4 Developments and Applications". In: *IEEE Transactions on Nuclear Science* 53 (Feb. 2006), pp. 270–278.

[79] L. Breiman et al. *Classification and Regression Trees*. Belmont, California, USA: Wadsworth international group, 1984.

[80] Li-Gang Xia. "Understanding the boosted decision tree methods with the weak-learner approximation". In: (2018). arXiv: 1811.04822 [hep-ex].

[81] Jan Therhaag. "TMVA: Toolkit for multivariate data analysis". In: *AIP Conf. Proc.* 1504.1 (2012). Ed. by Theodore E. Simons and George Maroulis, pp. 1013–1016.

[82] Andreas Hoecker et al. "TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide." In: (2009). arXiv: physics/0703039 [physics].

[83] R. Aaij et al. "Measurement of *CP* observables in $B^{\pm} \to DK^{\pm}$ and $B^{\pm} \to D\pi^{\pm}$ with two- and four-body D decays". In: *Phys. Lett.* B760 (2016), p. 117. arXiv: 1603.08993 [hep-ex].

[84] Wouter Verkerke and David P. Kirkby. "The RooFit toolkit for data modeling". In: *eConf* C0303241 (2003). Ed. by L. Lyons and Muge Karagoz. arXiv: physics/0306116.

[85] Roger Barlow. "Extended maximum likelihood". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 297.3 (1990), pp. 496–506. URL: https://www.sciencedirect.com/science/article/pii/0168900290913348.

[86] R Aaij et al. "Test of lepton flavor universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays". In: *Phys. Rev. D* 97 (7 Apr. 2018), p. 072013. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.072013.

[87] Roel Aaij et al. "Dalitz plot analysis of $B^0 \to \overline{D}^0\pi^+\pi^-$ decays". In: *Phys. Rev. D* 92.3 (2015), p. 032002. arXiv: 1505.01710 [hep-ex].

[88] Roel Aaij et al. "Dalitz plot analysis of $B_s^0 \to \bar{D}^0 K^-\pi^+$ decays". In: *Phys. Rev. D* 90.7 (2014), p. 072003. arXiv: 1407.7712 [hep-ex].

[89] R. Aaij, B. Adeva, and et. al. "Measurement of the $B^{\pm}$ production asymmetry and the CP asymmetry in $B^{\pm} \to J/\psi K^{\pm}$ decays". In: *Phys. Rev. D* 95 (5 Mar. 2017), p. 052005. URL: https://link.aps.org/doi/10.1103/PhysRevD.95.052005.

[90] R. Aaij et al. "Measurement of *CP* observables in $B^{\pm} \to DK^{*\pm}$ decays using two- and four-body D final states". In: *Journal of High Energy Physics* 2017.11 (Nov. 2017). URL: http://dx.doi.org/10.1007/JHEP11(2017)156.

[91] R. Aaij et al. "Measurement of *CP* observables in the process $B^0 \to DK^{*0}$ with two- and four-body D decays". In: *Journal of High Energy Physics* 2019.8 (Aug. 2019). URL: http://dx.doi.org/10.1007/JHEP08(2019)041.

[92] Luc Demortier and Louis Lyons. *Everything you always wanted to know about pulls.* Tech. rep. Link to CDF note. The Rockefeller University and University of Oxford, 2002.

[93] B. Efron. "Bootstrap Methods: Another Look at the Jackknife". In: *The Annals of Statistics* 7.1 (1979), pp. 1–26. URL: https://doi.org/10.1214/aos/1176344552.

[94] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. URL: http://www.jstor.org/stable/2957648.

[95] M. Kenzie et al. *Gamma combo user manual.* Can be obtained, along with the code, at: https://gammacombo.github.io/.