Beauty meson to double open charm decays with the LHCb detector

Fionn Caitlin Ròs Bishop

Darwin College

University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

April 2023

Declaration

This thesis is the result of my own work except where explicitly specified in the preface. It is not the same as any work that has previously been submitted for any degree or other qualification. It does not exceed the prescribed word limit for the Faculty of Physical Sciences and Chemistry Degree Committee.

Abstract

Measurements of beauty meson to double open charm decays are presented. These measurements use data collected in proton-proton collisions by the LHCb detector at the Large Hadron Collider (LHC) between 2011 and 2018.

Firstly a search for sixteen rare decays of the B_c^+ meson to two charm mesons is performed. Studies of the B_c^+ meson can improve understanding of heavy quark dynamics, and its decay to two charm mesons could be useful in measuring the Cabibbo-Kobayashi-Maskawa quark-mixing phase γ in future LHC data taking periods. The first evidence of the $B_c^+ \to D_s^+ \overline{D}^0$ decay is found with a significance of 3.4 standard deviations and its branching fraction is measured. Upper confidence limits on the branching fractions of the other fifteen decays are improved.

Secondly, the charge-parity asymmetries ($\mathcal{A}^{CP}$) and branching fractions of decays of the B^- meson to two charm mesons are measured. No evidence of charge-parity violation is seen, but the first measurements of $\mathcal{A}^{CP}(B^- \to D_s^{*-} D^0)$ and $\mathcal{A}^{CP}(B^- \to D_s^- D^{*0})$ are made and the precision is improved on the $\mathcal{A}^{CP}$ of a further five decay modes. Moreover, ratios between the branching fractions of doubly charmed decays of the B^- meson are measured for the first time. These results will help to constrain physics beyond the Standard Model.

Finally, a generic beauty hadron neural network and exclusive selections for doubly charmed beauty meson decays are developed, which are implemented in the trigger for future data taking. The improvement in selection efficiency using these selections in the upgraded trigger system is evaluated. Additionally, prospects for measurements of doubly charmed beauty meson decays with future LHCb datasets are discussed. Sensitivity to rare doubly charmed B_c^+ meson decays and high-precision measurements of charge-parity violation will be enabled by the upgraded LHCb detector and trigger system.

Beauty meson to double open charm decays with the LHCb detector
Fionn Caitlin Ròs Bishop

Acknowledgements

Firstly, I would like to express my gratitude to my supervisor, Val Gibson, for the opportunity to undertake this PhD, and for the support throughout the process. Furthermore, I am deeply grateful to Rolf Oldeman, for his extraordinary generosity with his time, and for teaching me so much about being a particle physicist.

I am immensely thankful to the Huo Family Foundation for the award of a P C Ho scholarship, without which this research would not have been possible. I also acknowledge Darwin College for providing additional travel grants that have helped me further my education.

It has been an honour and a pleasure to be part of the Cambridge High Energy Physics group. I would like to thank Paula Álvarez Cartelle for the sage advice and Blaise, John, and Lakshan for the less sage but still excellent advice. I am indebted to Steve Wotton for maintaining such a wonderful computer system (among his many accomplishments). I extend my heartfelt gratitude to my excellent companions in terminal-PhD suffering: David, Dom, and Ben. Furthermore, I am grateful to Josh, Thomas, and Raoul for being such fun office-mates in the final months of my PhD and not talking too much while I was writing this thesis. Finally, I express my sincere appreciation to everyone who respected the importance of going to the pub and the sanctity of the lunch break.

A twelve month attachment at CERN was without doubt the highlight of my PhD and a fantastic opportunity to meet collaborators and participate in data taking in the control room. I am grateful to the UK Liaison Office for organising the associated administration and accommodation. I also extend my appreciation to the whole LTA cohort for the many enjoyable lunchtimes, day trips, potlucks, and enlightening conversations about physics. A special thanks to the majestic country of Switzerland for providing so many wonderful places to explore.

My success has relied on the support of my family and friends: most importantly, my parents, to whom I owe so much. My sister is inspiring in her tenacity and the cats have been a constant source of happiness and comfort. I am eternally grateful to my Virtual Euroclub family for helping me to find joy at the hardest times and for making a pandemic bearable. Finally, I extend my appreciation to all of my friends in Cambridge, with whom I have not been able to spend enough time in these last few years. I will miss you and I hope that we will find each other again on our new adventures.

Preface

This thesis documents the research that I have performed since October 2019 as part of the LHCb collaboration on *beauty meson to double open charm decays*. Firstly, a search for double charm decays of the B_c^+ meson is presented using the data collected by the LHCb experiment in proton-proton collisions between 2011 and 2018 [1]. This is followed by measurements of charge-parity violation and branching fractions for double charm decays of the B^- meson using the same dataset [2]. Finally, selections for doubly charmed beauty meson decays to be evaluated in real-time during future LHCb data taking have been developed.

The majority of the research described in this thesis is my own work. However, these measurements were performed in collaboration with Dr Rolf Oldeman, with contributions to the early stages of the analysis from Dr Alison Tully and Dr Chris Jones, and under the supervision of Professor Val Gibson. Therefore it is necessary to describe some work that has been performed by my collaborators for the sake of clarity. My contributions to Chapter 5 are:

- Studies of the inclusion of the charm meson decay modes that include neutral particles.
- Correcting simulated samples for the B^+ meson kinematics and event multiplicity, B_c^+ meson kinematics, and four-body D^0 meson decay distribution.
- Homogenisation of simulation versions.
- Weighting simulated samples to emulate the kinematics of partially reconstructed decays.
- Adaptation of the multivariate analysis for measuring the branching fraction of B^- meson decays.
- Optimising the selection of data according to the multivariate analysis output.
- Evaluating the charge dependence of the multivariate analysis.
- Analysis and treatment of multiple candidates.

The contributions of my collaborators to Chapter 5 are:

- Development of stripping lines and preselection.
- Preparation of data and simulation samples.
- Correcting momentum scale, momentum resolution, and PID variables for simulated samples.
- Development and training of the multivariate analysis.

Chapters 6-7 are my own work except for the evaluation of single charm background yields and of signal and normalisation efficiencies. The correction of production and detection asymmetries and selection efficiencies described in Sections 8.2-8.3 relies on measurements performed by LHCb collaborators, which are either published as measurements or available in internal tools, namely:

- Measurement of the production asymmetry in kinematic bins.
- Measurement of the hardware trigger global TIS asymmetry in kinematic bins.
- Measurement of the hardware trigger hadron TOS efficiency in bins of track transverse energy and HCAL region, and evaluation of HCAL cell miscalibration and occupancy.
- Internal tool to evaluate kaon-pion detection asymmetry for given signal decay kinematics.
- Measurement of the pion detection asymmetry in kinematic bins.
- Internal tool to evaluate particle ID efficiencies in kinematic bins.
- Measurement of the tracking efficiency in kinematic bins.

In these sections, my contributions are:

- Weighting of the binned measurements by signal kinematics.
- Optimisation of binning for the evaluation of the kaon-pion detection asymmetry and particle ID efficiencies.
- Validation of the weighting procedure in the evaluation of the kaon-pion detection asymmetry.
- Determination of the effective PID requirement.
- Evaluation of the hardware trigger TOS and TISnotTOS efficiencies in bins of kinematics of the B^- meson.

The remainder of Chapter 8 is my own work.

The development of a generic b-hadron neural network for real-time selections in future data taking (Section 9.2) was performed in collaboration with the Beauty to Open Charm working group. My contribution was the neural network training. The remainder of Chapter 9 is my own work.

The field of particle physics and its open questions are introduced in Chapter 1. Chapter 2 provides the theoretical background that motivates the study of beauty meson to double charm decays, as well as frameworks used to extract physical information from the data in later chapters. Chapter 3 introduces the Large Hadron Collider and then the LHCb detector, which are the machine used to accelerate and collide protons at high energies and the detector that records the collision events used in the analyses described in this thesis, respectively.

The measurements of beauty meson to double open charm decays are introduced in Chapter 4, and then presented in full detail in the following chapters. The data and simulation samples, and the selection applied to reduce backgrounds, are described in Chapter 5. In Chapter 6, the models used to describe the invariant-mass distributions of signal and background decays in the data are explained. Chapter 7 completes the description of the search for doubly charmed B_c^+ meson decays, and Chapter 8 details the measurement of charge-parity violation and branching fractions in doubly charmed B^- meson decays.

Chapter 9 explains the development of real-time selections for beauty meson to double charm decays for future data taking and outlines experimental prospects for the remainder of LHC data taking. The results are summarised in Chapter 10, and prospects for improving knowledge on beauty meson to double charm decays in future data taking at the Large Hadron Collider are also discussed.

Contents

Declaration **iii**

Abstract **v**

Acknowledgements **vii**

Preface **xi**

1 Introduction **1**

2 Theoretical background **5**

2.1 The Standard Model 5

2.2 Weak interactions of quarks in the Standard Model 5

2.3 Charge-parity violation 7

2.3.1 Direct charge-parity violation 7

2.3.2 Charge-parity violation in neutral meson mixing 8

2.3.3 Charge-parity violation in the interference of neutral meson mixing and decay 9

2.4 Unitarity of the CKM matrix 9

2.4.1 Extraction of the CKM phase γ 11

2.5 Doubly charmed B_c^+ meson decays 13

2.5.1 Extraction of γ with B_c^+ meson decays 15

2.6 Charge-parity violation in doubly charmed B^- decays 15

3 The LHCb detector at the LHC **19**

3.1 The Large Hadron Collider 19

3.2 The LHCb detector 23

3.2.1 Tracking 26

3.2.2 Particle identification 32

3.2.3 Trigger system 37

3.2.4 Offline reconstruction and selection 42

3.3 Upgrade I 46

3.3.1 Hardware upgrades 46

3.3.2 Data processing upgrades 46

3.4 Simulation 47

4 Analysis overview **51**
4.1 Search for doubly charmed B_c^+ meson decays 51
4.2 Analysis of doubly charmed B^- meson decays 52

5 Data samples and selection **55**
5.1 Data samples 55
5.2 Simulation samples 57
5.3 Corrections to simulation 59
5.3.1 Momentum scale and resolution 59
5.3.2 Particle identification variables 60
5.3.3 Gradient-boosted reweighting 60
5.3.4 Background subtraction in data 61
5.3.5 B^+ kinematics and event multiplicity 62
5.3.6 B_c^+ kinematics 63
5.3.7 Angular distribution of four-body D^0 decay 70
5.3.8 Homogenisation of simulation versions 72
5.3.9 Weighting to emulate the kinematics of partially reconstructed decays 72
5.4 Comparison of simulated and true signal distributions 73
5.5 Trigger selection 76
5.6 Stripping selection 77
5.7 Preselection 77
5.8 Multivariate selection 80
5.8.1 Multivariate classifier 80
5.8.2 Optimisation of the selection 84
5.8.3 Charge dependence 85
5.9 Multiple candidates 85

6 Modelling of signal and background decays **91**
6.1 Fully reconstructed $B_{(c)}^+ \to DD$ decays 92
6.2 Partially reconstructed $B \to D^* D^{(*)}$ decays 94
6.2.1 $B \to DD$ decays with one missing particle 95
6.2.2 $B \to DD$ decays with two missing particles 98
6.2.3 $B_c^+ \to DD$ decays 100
6.2.4 Polarisation fractions 100
6.2.5 $B_s^0 \to DD$ yields 100
6.3 Partially reconstructed $B \to D^{**} D$ decays 102
6.4 Single charm and charmless backgrounds 103
6.5 Cross-feed 106
6.5.1 $\overline{D}^0 \to D^0$ cross-feed 106

6.5.2 $D_s^+ \leftrightarrow D^+$ cross-feed . . . 107
6.6 Combinatorial background . . . 108

7 Search for doubly charmed B_c^+ meson decays 113
7.1 Signal and normalisation efficiencies . . . 114
7.2 Fit to data . . . 115
7.2.1 Fit method . . . 115
7.2.2 Model . . . 115
7.2.3 Blinding . . . 116
7.2.4 Fit results . . . 116
7.3 Systematic uncertainties . . . 118
7.3.1 Systematic uncertainties on the normalisation . . . 118
7.3.2 Systematic uncertainties on the model . . . 122
7.4 Results . . . 124
7.4.1 Signal significances and branching fractions . . . 124
7.4.2 Confidence limits on branching fractions . . . 126

8 Analysis of doubly charmed B^- meson decays 131
8.1 Raw asymmetries and yields . . . 131
8.1.1 Fit model . . . 131
8.1.2 Fit results . . . 133
8.1.3 Validation of modelling of partially reconstructed decays . . . 135
8.1.4 Bias and coverage . . . 136
8.2 Charge asymmetry corrections . . . 138
8.2.1 Production asymmetry . . . 139
8.2.2 Hardware trigger global TIS asymmetry . . . 141
8.2.3 Hardware trigger hadron TOS asymmetry . . . 141
8.2.4 Kaon-pion detection asymmetry . . . 144
8.2.5 Pion detection asymmetry . . . 148
8.2.6 PID asymmetry . . . 149
8.3 Efficiency corrections . . . 152
8.3.1 Hardware trigger efficiency . . . 153
8.3.2 Tracking efficiency . . . 154
8.3.3 PID efficiency . . . 155
8.4 Systematic uncertainties . . . 156
8.5 Charge-parity asymmetry results . . . 162
8.5.1 Validation . . . 164
8.6 Branching fraction results . . . 167
8.6.1 Validation . . . 167

8.6.2 Stability tests . . . 167

9 Beauty to double charm decays in Run 3 and beyond 171

9.1 Trigger line specification . . . 171

9.2 Generic beauty hadron neural network . . . 172

9.3 Beauty to double charm trigger lines . . . 173

9.4 Experimental prospects in Run 3 and beyond . . . 182

10 Summary and outlook 185

10.1 Summary . . . 185

10.2 Outlook . . . 187

A Comparison of simulation and data 189

B Input variables to the BDT 197

C Likelihood scans 199

D CLs curves for branching fractions 203

E Bias and coverage on raw asymmetry 207

F The LHCb detector in cross stitch 211

References 213

Chapter 1

Introduction

At the scale of fundamental particles, our universe is described by the *Standard Model* (SM) of particle physics [3–5]. The SM firstly contains twelve spin-half fermions, consisting of six *quarks* (u, d, s, c, t, b), three charged *leptons* (e, μ, τ), and three neutral leptons, or *neutrinos* $(\nu_e, \nu_\mu, \nu_\tau)$. For each of these fermions an antiparticle exists, whose quantum numbers are reversed in sign relative to the particle. These fermions are divided into three *generations*, each containing one up-type and one down-type quark, one charged lepton, and one neutrino. The SM also describes several interactions through which fermions interact, namely the *strong*, *weak*, and *electromagnetic* (EM) interactions. These interactions are mediated by four spin-one *gauge bosons*. The weak and electromagnetic interactions unify at sufficiently high energy as the *electroweak* interaction. The strong interaction is mediated by the gluon, and the electroweak interaction by the $W^\pm$ and Z bosons and the photon (γ). Finally, the Higgs field gives particles mass, and is associated with a spin-zero Higgs boson particle [6–8].

Since the conception of the SM, particle physics experiments have sought to test this theory, firstly by discovering new particles predicted by the SM, and then by comparing their properties with predictions: although the SM has nineteen free parameters, the number of observables is much higher, allowing the SM to be overconstrained and thereby tested. To date, the SM has enjoyed great success at predicting the properties of particles and their interactions, including before their observation. The production of W and Z bosons was observed in proton-antiproton collisions in 1982 and 1983, respectively, with masses and cross-sections consistent with the SM predictions [9–12]. The verification of the self-consistency of the SM through precision measurements of electroweak parameters continued in high energy electron-positron collisions. On the side of the fermions, the observation of the top quark (t) in proton-antiproton collisions in 1995 [13, 14] and the tau neutrino in 2000 [15] completed the third generation of matter. The observation of the SM particles was complete with the discovery of the Higgs boson in proton-proton (pp) collisions in 2012 [16, 17].

Despite these successes, the SM leaves a number of questions unanswered. For example, in the fermion sector, there is no explanation as to why there are three generations of matter or why the generations span such a large mass range. Moreover, the distribution of astronomical matter inferred from gravitational behaviour is inconsistent with the visible distribution of matter. This observation cannot be explained using SM particles, motivating the hypothesis of weakly interacting *dark matter* particles beyond the SM [18–20].

Furthermore, the SM does not explain the baryon asymmetry of the universe (BAU) given that matter and antimatter were produced in equal quantities in the early universe. The baryon density asymmetry relative to the photon density of the universe is measured using observations of the cosmic microwave background [20–22] and the relative abundances of light elements in astrophysical systems [21–23]. Baryon asymmetry is permitted if the three Sakharov conditions are satisfied [24]. The first condition is the existence of processes that violate baryon number conservation, which is accommodated in the SM [25]. The second condition, a departure from thermal equilibrium in the early universe, can be explained in the SM but is incompatible with the measured Higgs boson mass [26, 27].

The final condition is the violation of charge (C) and charge-parity (CP) symmetries. Charge transformation is the reversal of the sign of all quantum charges, in effect transforming a particle into its antiparticle and vice versa. Parity transformation is the reversal of all spatial coordinates, which transforms left-handed into right-handed fields and vice versa. However, even assuming a departure from thermal equilibrium, the degree of CP symmetry violation (CPV) observed in the SM is many orders of magnitude too small to produce the observed BAU [28–30].

The issues with the SM, which include the inconsistency of the observed BAU with the measured SM parameters, imply a need for new physics (NP) beyond the Standard Model (BSM), including a larger source of CPV. The ATLAS and CMS experiments at the Large Hadron Collider (LHC) seek to produce heavy BSM particles in pp collisions, but none have been observed to date. Other experiments, including the LHCb detector at the LHC, precisely measure the behaviour of SM particles. BSM particles may perturb this behaviour through quantum loops, allowing their presence to be identified even at energies too low to produce them directly.

Moreover, even processes so far described accurately by the SM cannot necessarily be described precisely because the strong interaction cannot be evaluated as a perturbation expansion in low energy processes, such as in the bound states of quarks and their decays. Improved understanding of the bound states and strong interactions of quarks can help improve the theoretical tools used to describe these non-perturbative processes. This will enable a more precise description of the behaviour of hadronic particles, which can in turn assist with the identification of BSM particles in quantum loops.

This thesis presents two measurements of the decays of beauty (B) mesons that consist of a b quark plus a lighter antiquark,[1] to two charm (D) mesons that consist of a c quark plus a lighter antiquark, with the LHCb detector. Firstly, a search for decays of the B_c^+ meson ($\overline{b}c$) to two charm mesons ($B_c^+ \to DD$) is performed [1] using data collected in pp collisions by the LHCb detector [31] at the LHC. Consisting of two heavy quarks, the B_c^+ meson is poorly

[1] The inclusion of charge-conjugate processes is implied throughout except in the discussion of asymmetries.

understood both theoretically and experimentally. Measurements of its hadronic decays may improve understanding of its QCD dynamics. Furthermore, hadronic beauty meson decays could in general be affected by BSM physics [32–34]. Finally, when larger samples of these decays are obtained, they may be useful to test the self-consistency of CPV in the SM. Since existing e^+e^- colliders do not operate at sufficiently high centre-of-mass energies to produce the B_c^+ meson, and identification of these rare decays above a large hadronic background requires excellent momentum resolution and particle species identification, the LHCb detector provides a unique opportunity to search for these decays.

Secondly, a measurement of CPV and branching fractions for the decays of the B^- meson ($b\overline{u}$) to two charm mesons ($B^- \to DD$) decays is presented [2]. These measurements could help to identify BSM physics exhibiting a high degree of CPV, which may help to explain the baryon asymmetry problem.

The LHCb detector has recently completed upgrades to permit the detection of five times more pp collision events at once. This requires redesign of the trigger and readout system through which events are saved for analysis. The development of selections implemented in the trigger system to select beauty meson to double charm decays will also be described in this thesis.

CHAPTER 2

Theoretical background

2.1 The Standard Model

Interactions in the SM are defined by a local $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ gauge symmetry. Each of the constituent gauge symmetries results in spin-one gauge bosons that mediate the interactions between the SM particles. The $\mathrm{SU}(3)_\mathrm{C}$ gauge symmetry of the SM gives rise to the strong interaction, described by the theory of *quantum chromodynamics* (QCD). QCD is mediated by eight massless gluons. Gluons couple to particles carrying *colour* charge, namely the quarks and themselves.

The $\mathrm{SU}(2)_\mathrm{L}$ and $\mathrm{U}(1)_\mathrm{Y}$ gauge symmetries result in W^i bosons ($i = 1, 2, 3$) and one B boson, respectively. Additionally, the SM contains the complex scalar Higgs field that spontaneously breaks the $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ symmetry to a $\mathrm{U}(1)_\mathrm{QED}$ symmetry through the *Higgs mechanism* [6–8]. This symmetry breaking produces the observable electroweak mediators. Massive $W^\pm$ bosons are linear combinations of the W^1 and W^2 bosons, and mediate charged-current weak interactions. The massive Z boson and massless photon are linear combinations of the W^3 and the B bosons that mediate neutral-current weak interactions and EM interactions, respectively. The Higgs mechanism additionally results in a physical spin-zero Higgs boson in the SM.

The constituent particles of the SM are the twelve fermions identified in Chapter 1 plus the four gauge bosons and one Higgs boson described above. These particles and their properties are specified in Figure 2.1.

2.2 Weak interactions of quarks in the Standard Model

The quark states which couple to the strong and EM interactions are equal to the quark mass eigenstates. This is also true for the up-type quark states in their couplings to the weak interaction. In constrast, the weak down-type eigenstates (d', s', b') are related to the mass eigenstates (d, s, b) by the unitary *Cabibbo-Kobayashi-Maskawa matrix* [36, 37] (CKM matrix, or V_{CKM}) as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.1}$$

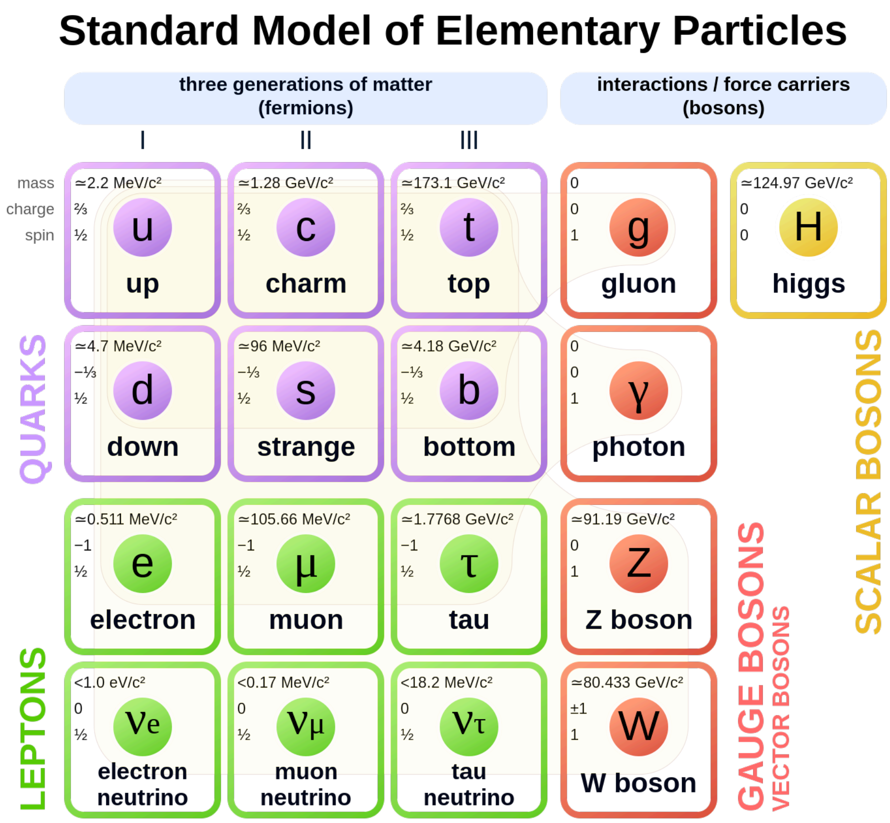

Figure 2.1: The elementary particles of the SM and their masses, EM charges, and spin [35].

where

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{2.2}$$

and the matrix elements $V_{qq'}$ are complex scalars. Off-diagonal elements in V_{CKM} enable flavour changes in charged-current weak interactions. The matrix V_{CKM} can be parameterised in terms of three mixing angles (θ_{12}, θ_{13} and θ_{23}), and six complex phases. The relative phase between quark fields is unphysical, so these phases can be absorbed into the redefinition of five quark fields, leaving one physical complex phase (δ), allowing V_{CKM} to be expressed as [38]

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.3}$$

where $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. The mixing angles and complex phase are free parameters of the SM that must be measured experimentally. The current world average values of the magnitudes of the elements of V_{CKM} are [39]

$$|V_{CKM}| = \begin{pmatrix} 0.97435^{+0.00005}_{-0.00006} & 0.22500^{+0.00024}_{-0.00021} & 0.00367^{+0.00009}_{-0.00007} \\ 0.22487^{+0.00024}_{-0.00021} & 0.97352 \pm 0.00006 & 0.0415^{+0.0004}_{-0.0006} \\ 0.00852^{+0.00008}_{-0.00015} & 0.0407^{+0.0004}_{-0.0006} & 0.999142^{+0.000018}_{-0.000023} \end{pmatrix}. \tag{2.4}$$

The CKM matrix is strongly hierarchical with $s_{13} \ll s_{23} \ll s_{12} \ll 1$, so weak coupling is strongest between quarks in the same generation. Consequently, decays involving coupling between quarks in the same generation are referred to as *Cabibbo-favoured* (CF). Decays involving coupling between adjacent generations are *Cabibbo-supressed* (CS) and *doubly Cabibbo-supressed* (DCS) decays involve a coupling between the first and third generations.

2.3 Charge-parity violation

In the SM, CPV manifests itself in the weak interactions of quarks due to the imaginary part of V_{CKM}. CPV was first observed in the decays of kaons [40–42]. The phenomenon was subsequently observed in beauty mesons in 2001 [43, 44] and in charm mesons in 2019 [45]. The remainder of this section further describes the manifestations of CPV in the weak interactions of quarks.

2.3.1 Direct charge-parity violation

The rate Γ of the decay of an initial state i to a final state f is related to the amplitude $A(i \to f)$ as

$$\Gamma(i \to f) \propto |A(i \to f)|^2. \tag{2.5}$$

The total amplitude is the sum over the amplitudes of all possible diagrams,

$$A(i \to f) = \sum_i A_i e^{i(\delta_i + \phi_i)}, \tag{2.6}$$

where A_i are the magnitudes, δ_i are the CP-even strong phases from strong interactions, and ϕ_i are the CP-odd weak phases arising from V_{CKM}, of each diagram. Under a CP transformation, the sign of the weak phase is reversed, and the strong phase is unchanged:

$$A(\bar{i} \to \bar{f}) = \sum_i A_i e^{i(\delta_i - \phi_i)}. \tag{2.7}$$

The decay $i \to f$ exhibits CPV if its rate is different to that of the CP-conjugate process, that is if

$$|A(i \to f)|^2 - |A(\bar{i} \to \bar{f})|^2 \neq 0. \tag{2.8}$$

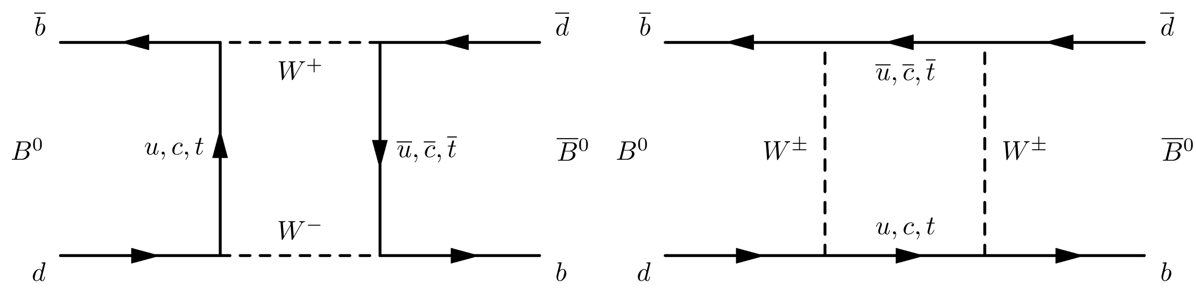

Figure 2.2: Box diagrams contributing to B^0-$\overline{B}^0$ oscillations [46].

If only one diagram contributes to Equation 2.6, then the left hand side of Equation 2.8 is always equal to zero and *CP* symmetry is conserved, regardless of the existence of a weak phase. The observation of direct CPV in a decay requires the existence of at least two amplitudes with different strong phases $\delta_1 \neq \delta_2$ and weak phases $\phi_1 \neq \phi_2$.

The degree of CPV in a decay is quantified by the *CP* asymmetry

$$\mathcal{A}^{CP} = \frac{\Gamma(i \to f) - \Gamma(\bar{i} \to \bar{f})}{\Gamma(i \to f) + \Gamma(\bar{i} \to \bar{f})}. \tag{2.9}$$

2.3.2 Charge-parity violation in neutral meson mixing

Neutral K^0, D^0, B^0, and B^0_s mesons oscillate as a function of time into their antiparticles through flavour-changing-neutral-current box diagrams (Figure 2.2). Consequently the flavour eigenstates mix so that the mass eigenstates

$$M_H = qM^0 + p\overline{M}^0 \tag{2.10}$$

and

$$M_L = pM^0 - q\overline{M}^0 \tag{2.11}$$

are linear combinations of the flavour eigenstates. The mass eigenstates differ both in mass,

$$\Delta m = m_H - m_L, \tag{2.12}$$

and in width,

$$\Delta\Gamma = \Gamma_H - \Gamma_L. \tag{2.13}$$

These oscillations violate *CP* symmetry if the rate of $M^0 \to \overline{M}^0$ oscillations differs from the rate of $\overline{M}^0 \to M^0$ oscillations. CPV in mixing can be isolated from direct CPV by considering the asymmetry in the decay rate to a flavour-specific final state f, for example a semileptonic final state, which identifies the flavour at the time of decay. The rate of the decay to the final state $\bar{f}$ that is wrong-flavour relative to the produced M^0 meson is proportional to the rate of the process $M^0 \to \overline{M}^0$. Therefore CPV in mixing is quantified by measuring

$$a_{\mathrm{sl}} = \frac{\Gamma(\overline{M}^0 \to f) - \Gamma(M^0 \to \overline{f})}{\Gamma(\overline{M}^0 \to f) + \Gamma(M^0 \to \overline{f})}, \tag{2.14}$$

where $\overset{(-)}{M}{}^0$ denotes the initial flavour.

2.3.3 Charge-parity violation in the interference of neutral meson mixing and decay

The interference of neutral meson mixing and flavour-non-specific decays induces a third manifestation of CPV. The rates of the four possible decays oscillate as a function of the decay time (t). In the $\overset{(-)}{B}{}^0 \to \overset{(-)}{f}$ decay system, where $\Delta\Gamma/\Gamma \ll 1$ [47], the decay rates evolve as:

$$\begin{aligned}
\frac{d\Gamma(B^0 \to f)}{dt} &= \frac{e^{-t/\tau}}{8\tau}(1+\mathcal{A}^{CP})(1-(S+\Delta S)\sin(\Delta m t)+(C+\Delta C)\cos(\Delta m t)),\\
\frac{d\Gamma(\overline{B}^0 \to f)}{dt} &= \frac{e^{-t/\tau}}{8\tau}(1+\mathcal{A}^{CP})(1+(S+\Delta S)\sin(\Delta m t)-(C+\Delta C)\cos(\Delta m t)),\\
\frac{d\Gamma(B^0 \to \overline{f})}{dt} &= \frac{e^{-t/\tau}}{8\tau}(1-\mathcal{A}^{CP})(1-(S-\Delta S)\sin(\Delta m t)+(C-\Delta C)\cos(\Delta m t)),\\
\frac{d\Gamma(\overline{B}^0 \to \overline{f})}{dt} &= \frac{e^{-t/\tau}}{8\tau}(1-\mathcal{A}^{CP})(1+(S-\Delta S)\sin(\Delta m t)-(C-\Delta C)\cos(\Delta m t)),
\end{aligned} \tag{2.15}$$

where τ is the B^0 meson lifetime and C, S, ΔC, and ΔS are decay-specific parameters. $\mathcal{A}^{CP}$ and C quantify CPV in the decay, S quantifies the CPV in mixing, ΔC quantifies the flavour specificity of the decay, and ΔS is related to the strong phase difference between the interfering amplitudes [48, 49].

2.4 Unitarity of the CKM matrix

Unitarity of the CKM matrix requires that $\sum_i V_{ij}V^*_{ik} = \delta_{jk}$ and $\sum_j V_{ij}V^*_{kj} = \delta_{ik}$. The six constraints with $j \neq k$ and $i \neq k$ can be represented in the complex plane as six *unitarity triangles* (UTs). The most commonly used UT is from the constraint

$$1 + \frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}} + \frac{V_{td}V^*_{tb}}{V_{cd}V^*_{cb}} = 0, \tag{2.16}$$

which can be represented as a triangle in the complex plane (Figure 2.3). Its three angles are labelled as

$$\alpha = \arg(-V_{td}V^*_{tb}/V_{ud}V^*_{ub}), \tag{2.17}$$

$$\beta = \arg(-V_{cd}V^*_{cb}/V_{td}V^*_{tb}), \tag{2.18}$$

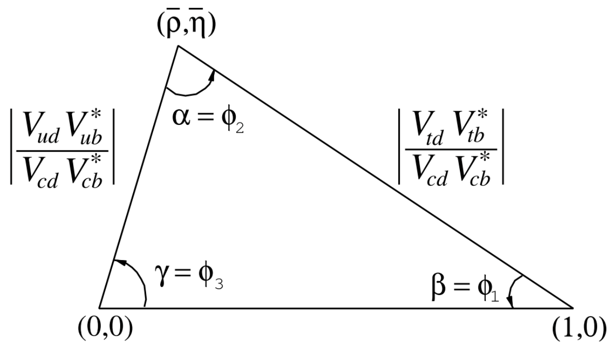

Figure 2.3: A diagram of the most commonly used unitarity triangle in the complex plane, corresponding to Equation 2.16 [22].

$$\gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*). \tag{2.19}$$

The vertices are $(0,0)$, $(1,0)$, and $(\overline{\rho},\overline{\eta})$, where

$$\overline{\rho} + i\overline{\eta} = -(V_{ud}V_{ub}^*/V_{cd}V_{cb}^*). \tag{2.20}$$

The location of the third vertex is not predicted by the SM and must be measured experimentally. Experimental constraints on this vertex are shown in Figure 2.4. Overconstraining the UT by measuring more than two of the parameters shown in Figure 2.4 provides a powerful test of the SM, since differences between two determinations of the UT would indicate that the measurement of the input parameters is biased, potentially by BSM physics.

In particular, the UT determined from the quantities extracted from tree-level decays (γ and $|V_{ub}|/|V_{cb}|$) can be compared to that determined from quantities extracted from loop-level decays (β and $\Delta m_d/\Delta m_s$), as demonstrated in Figure 2.5. This is a powerful test of the SM because NP is expected to affect tree-level and loop-level decays differently. Loop-level decays are generally believed to be the most sensitive to NP since their amplitudes are suppressed in the SM. However, large NP contributions to tree-level heavy quark decays, which could allow a $10°$ shift in γ, are also permitted by existing experimental constraints [50, 51]. Moreover, discrepancies between the measured branching fractions of singly charmed beauty meson decays and SM predictions have been identified [32–34].

The two UT determinations are presently consistent with each other. To improve the sensitivity of this test, it is essential to improve precision on the input parameters, particularly γ, which is one of the least well-known CKM matrix parameters and has a negligible irreducible theoretical uncertainty on its extraction of $\delta\gamma/\gamma < \mathcal{O}(10^{-7})$ [52].

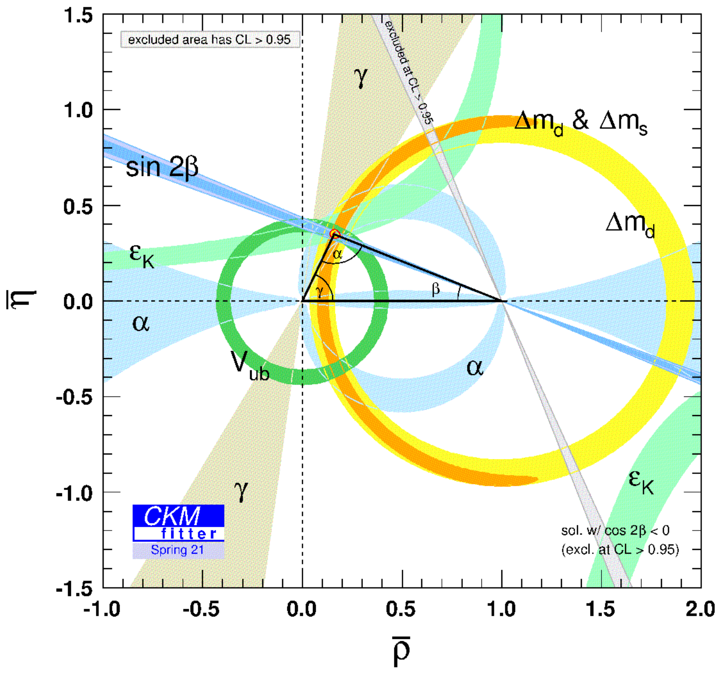

Figure 2.4: Experimental constraints at 95% confidence level on $(\overline{\rho}, \overline{\eta})$ [39].

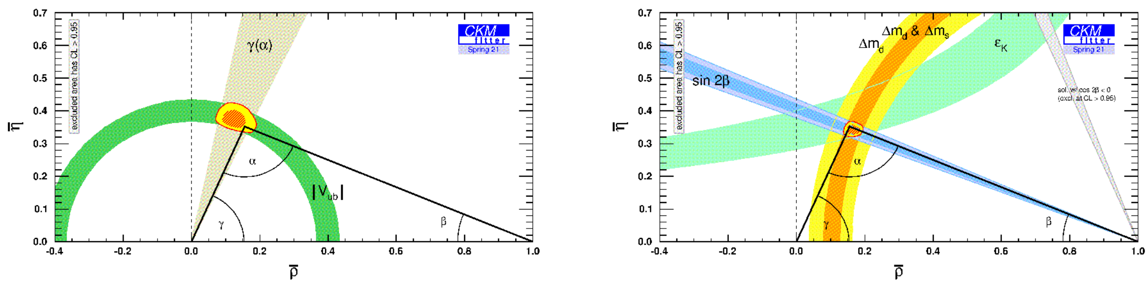

Figure 2.5: Experimental constraints at 95% confidence level on $(\overline{\rho}, \overline{\eta})$ using inputs measured in (left) tree-level and (right) loop-level decays [39].

2.4.1 Extraction of the CKM phase γ

The phase γ can be extracted from the interference between $b \to cW^-$ and $b \to uW^-$ transitions, for example in $B \to Dh$ decays where h is a kaon or pion and D is a linear combination of the D^0 and $\overline{D}^0$ mesons that decays to a final state f (Figure 2.6). For a two-body final state f, the amplitudes are related as

$$\frac{A(B^+ \to [\overline{D}^0]_f h^+)}{A(B^+ \to [D^0]_f h^+)} = \frac{r_D e^{-i\delta_D}}{r_B e^{i(\delta_B + \gamma)}}, \tag{2.21}$$

where the hadronic paramters r_B and δ_B are the ratio and strong phase difference between the $B^+ \to D^0 h^+$ and $B^+ \to \overline{D}^0 h^+$ amplitudes, and r_D and δ_D are the ratio and strong phase difference between the $D^0 \to f$ and $\overline{D}^0 \to f$ amplitudes. For the charge-conjugate process,

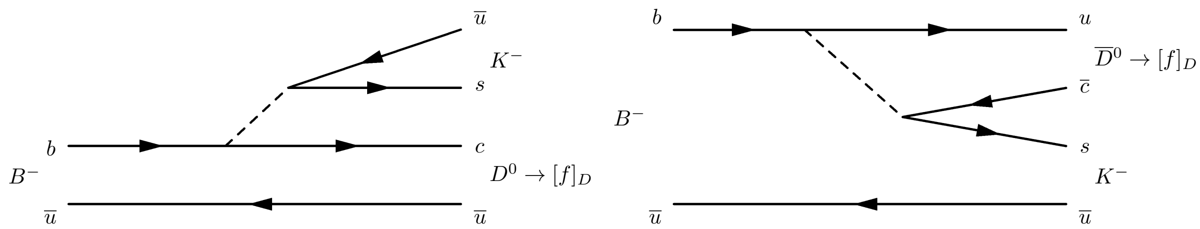

Figure 2.6: Feynman diagrams for the $B^- \to D^0 K^-$ and $B^- \to \overline{D}^0 K^-$ decays [46].

the sign of the weak phase γ is reversed, and the sign of the strong phase is unchanged. Therefore the total rate depends on the charge of the $B^\pm$ as

$$\Gamma(B^\pm \to [\overset{(-)}{f}]_D h^\pm) \propto r_B^2 + r_D^2 + 2 r_D r_B \cos(\delta_B + \delta_D \pm \gamma), \tag{2.22}$$

demonstrating that direct CPV is present in these decays and is a function of γ. To constrain all hadronic parameters and therefore extract γ from the measured CPV, it is necessary to analyse multiple beauty and charm meson decay modes, and perform a combined fit to extract γ from the information provided by each decay mode [53]. Different charm meson decay modes require different analysis methods, which will be outlined in the remainder of this section.

The *Gronau-London-Wyler method* [54, 55] uses decays where D is a CP eigenstate,

$$D^0_\pm = \frac{D^0 \pm \overline{D}^0}{\sqrt{2}}. \tag{2.23}$$

In this case, $r_D = 1$ and $\delta_D = 0$, eliminating uncertainties from these parameters. The CP-even D^0_+ can be identified in its decay to $\pi^+\pi^-$ or K^+K^-, and the CP-odd D^0_- can be identified in its decay to $K^0_S\pi^0$, $K^0_S\rho^0$, $K^0_S\omega$, or $K^0_S\phi$. This method is illustrated in Figure 2.7 using $B^\pm \to DK^\pm$ as an example, which shows the relationship between the amplitudes in the complex plane. If $\Gamma(B^+ \to D^0_+ K^+)$, $\Gamma(B^- \to D^0_+ K^-)$, $\Gamma(B^+ \to D^0 K^+) = \Gamma(B^- \to \overline{D}^0 K^-)$, and $\Gamma(B^+ \to \overline{D}^0 K^+) = \Gamma(B^- \to D^0 K^-)$ are known, then the phase γ can be derived. This method can be extended to multibody decays which are not pure CP eigenstates, such as $D \to \pi^+\pi^-\pi^0$. In this case the sensitivity to γ is diluted according to the fraction of CP-even content [56].

Despite lower yields, decays with $h = K$ produce better precision than decays with $h = \pi$ as r_B is closer to unity, so the interference is larger. Even so, $r_B \approx 0.1$, which limits sensitivity to γ. This is avoided in the *Atwood-Dunietz-Soni method* [57, 58] by choosing f such that $r_B \sim r_D$, for example $f = K^-\pi^+$. The cost is reduced precision from the uncertainty on r_D and δ_D. This method is also extended to multibody decays, with the dilution due to strong phases between intermediate resonances parameterised by the coherence factor [58].

The *Bondar-Poluektov-Giri-Grossman-Soffer-Zupan method* [59, 60] uses multibody self-conjugate decays (such as $D \to K^0_S\pi^+\pi^-$), and analyses the distributions of decay

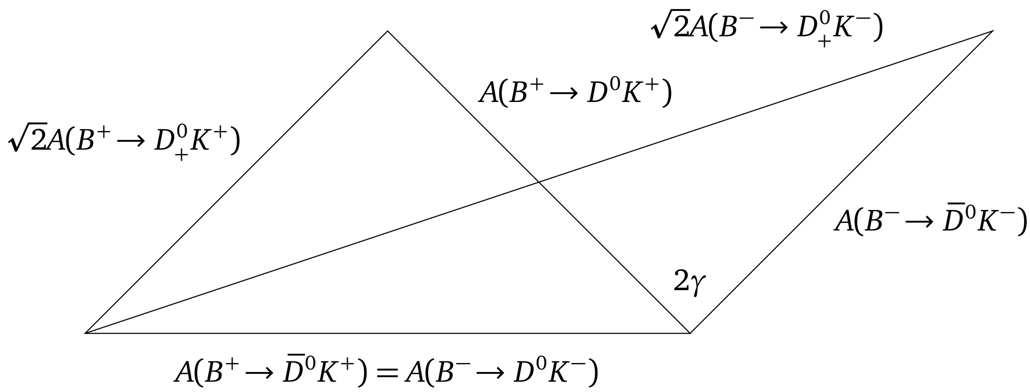

Figure 2.7: Amplitudes of $B^\pm \to DK^\pm$ decays represented in the complex plane. Knowledge of the magnitude of each amplitude is sufficient to extract the CKM phase γ, which is the angle between $A(B^+ \to D^0 K^+)$ and $A(B^- \to \overline{D}{}^0 K^-)$. Relative lengths are not correct.

candidates in the Dalitz plane. The Dalitz plane is the plane spanned by the two-body invariant masses in a three-body decay, for example $m(K^0_S \pi^+)$ and $m(K^0_S \pi^-)$. Information on the Dalitz distributions allows the dependence of r_D and δ_D over the phase space to be captured, avoiding the dilution that occurs when using multibody charm decays with the GLW or ADS methods.

Finally, $\gamma - 2\beta_s$, where $\beta_s \equiv \arg(-V_{ts}V^*_{tb}/V_{cs}V^*_{cb})$, can be extracted from the time-dependent rates of $B^0_s \to D^\mp_s K^\pm$ decays [61–63]. Combining results from all measured beauty and charm decay modes, and using dedicated measurements of the charm hadronic parameters at LHCb, CLEO, and BESIII [64–68], the current world average measurement is $\gamma = \left(63.8^{+3.5}_{-3.7}\right)^\circ$ [69].

2.5 Doubly charmed B_c^+ meson decays

The B_c^+ meson has a mass of $6274.47 \pm 0.32\,\text{MeV}/c^2$ [70] and a lifetime of $0.510 \pm 0.009\,\text{ps}$ [22], which is several times shorter than that of the B^+, B^0 or B^0_s. The requirement for the production and subsequent hadronisation of two heavy quarks makes the production of a B_c^+ meson rare, and it was only observed for the first time in 1998 at the CDF experiment [71].

Since $m_{B_c^+}$ is below the threshold to produce a beauty plus a charm meson, the B_c^+ meson only decays through the weak interaction. Decays of the $\overline{b}$ or c quarks, as well as weak annihilation, are possible. The absence of strong or EM decay channels, and the presence of two heavy quarks, makes the B_c^+ particularly interesting in the study of QCD dynamics.

One interesting category is the tree-level decays of the B_c^+ meson to two charm mesons. Decays of the form $B_c^+ \to D^+_{(s)}\overline{D}{}^0$ are CF compared to decays of the form $B_c^+ \to D^+_{(s)}D^0$, and so these decays are referred to as *right-sign* (RS) and *wrong-sign* (WS), respectively. The RS decays receive contributions from colour-suppressed and annihilation

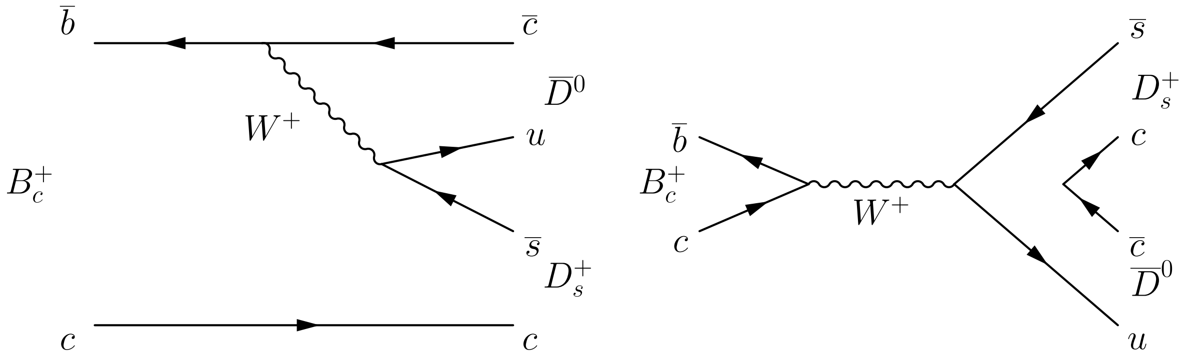

Figure 2.8: The (left) colour-suppressed and (right) annihilation diagrams contributing to the CF $B_c^+ \to D_s^+ \overline{D}^0$ decay.

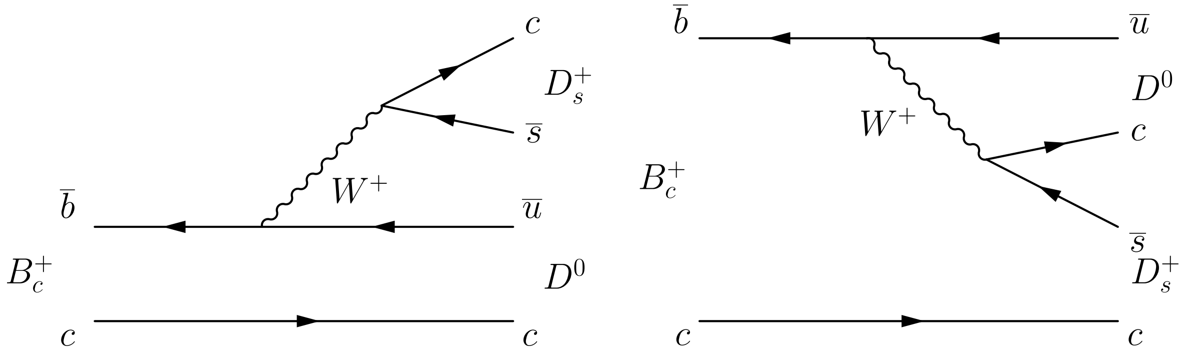

Figure 2.9: The (left) colour-allowed and (right) colour-suppressed diagrams contributing to the CS $B_c^+ \to D_s^+ D^0$ decay.

diagrams (Figure 2.8). The WS decays are dominated by the colour-allowed diagram, with a subleading contribution from the colour-suppressed diagram (Figure 2.9).

Decays of the B_c^+ meson to two charm mesons have not yet been observed, despite being tree-level processes. This is mostly due to the small fragmentation fraction of the B_c^+ (Table 3.2), but also because the CF $B_c^+ \to D^+ \overline{D}^0$ decay is colour suppressed relative to the CF $B^+ \to D_s^+ \overline{D}^0$ decay. Finally, it is difficult to separate the short-lived B_c^+ meson from random combinations of tracks in pp collisions.

The branching fractions of $B_c^+ \to DD$ decays have been predicted by multiple groups (Table 2.1). Predicted values depend on the theory tools and the included diagrams, and span over an order of magnitude. This large variation in branching fraction predictions in part motivates the experimental study of $B_c^+ \to DD$ decays to help probe heavy quark dynamics.

A further motivation for the study of these decays is their potential utility in measuring the CKM phase γ, which will be explained in the remainder of this section.

Table 2.1: Predicted branching fractions of selected $B_c^+ \to DD$ decays, in units of 10^{-6}.

Channel	Ref. [72]	Ref. [73]	Ref. [74]	Ref. [75]	Ref. [76]
$B_c^+ \to D_s^+ \overline{D}^0$	2.3 ± 0.5	4.8	1.7	2.1	0.27
$B_c^+ \to D_s^+ D^0$	3.0 ± 0.5	6.6	2.5	7.4	0.22
$B_c^+ \to D^+ \overline{D}^0$	32 ± 7	53	32	33	1.3
$B_c^+ \to D^+ D^0$	0.10 ± 0.02	0.32	0.11	0.31	0.008
$B_c^+ \to D^{*+} \overline{D}^0$	12 ± 3	49	17	9	1.6
$B_c^+ \to D^{*+} D^0$	0.09 ± 0.02	0.40	0.38	0.44	0.018

2.5.1 Extraction of γ with B_c^+ meson decays

Since the CF $B_c^+ \to D_s^+ \overline{D}^0$ decay is colour suppressed and the CS $B_c^+ \to D_s^+ D^0$ decay is colour allowed, the ratio between the interfering amplitudes in $B_c^+ \to D_s^+ D_\pm^0$ decays is expected to be $r_B \sim \mathcal{O}(1)$. For this reason, $B_c^+ \to D_s^+ D$ decays have been proposed as an alternative mode for the GLW method outlined in Section 2.4.1, with much better sensitivity to γ at a given yield [73, 74, 77, 78]. Angular observables in $B_c^+ \to D_s^{*+} D^*$ decays could provide further information on γ [79].

Another benefit is that the strong phase difference δ_B is expected to vanish [79], although it may be non-zero in reality due to resonance effects [80]. In the case that $\delta_B = 0$, the CP asymmetry would vanish, but by measuring the charge-averaged ratios

$$R_\pm = 2\frac{\Gamma(B_c^+ \to D_s^+ D_\pm^0) + \Gamma(B_c^- \to D_s^- D_\pm^0)}{\Gamma(B_c^+ \to D_s^+ \overline{D}^0) + \Gamma(B_c^- \to D_s^- D^0)}, \tag{2.24}$$

γ could be extracted with no dependence on the hadronic parameters through the expression [79]

$$\cos\gamma = \frac{1}{4}\frac{R_+ - R_-}{\sqrt{\frac{1}{2}(R_+ + R_-) - 1}}. \tag{2.25}$$

2.6 Charge-parity violation in doubly charmed B^- decays

Direct CPV manifests in the decays of the B^- meson to two charm mesons due to the interference between the colour-allowed tree diagrams and the suppressed QCD penguin and annihilation diagrams (Figure 2.10). The SM predicts small CP asymmetries (Table 2.2) due to large differences in the sizes of the interfering amplitudes: $\mathcal{A}^{CP}(B^- \to D_s^{(*)-} D^{(*)0})$ are expected to be $\mathcal{O}(0.1\%)$, and $\mathcal{A}^{CP}(B^- \to D^{(*)-} D^{(*)0})$ are expected to be $\mathcal{O}(1\%)$. These asymmetries could be enhanced by an order of magnitude in BSM models, for example in R-parity violating minimal supersymmetry [81], in minimal supergravity [82], or with four generations of quarks [83].

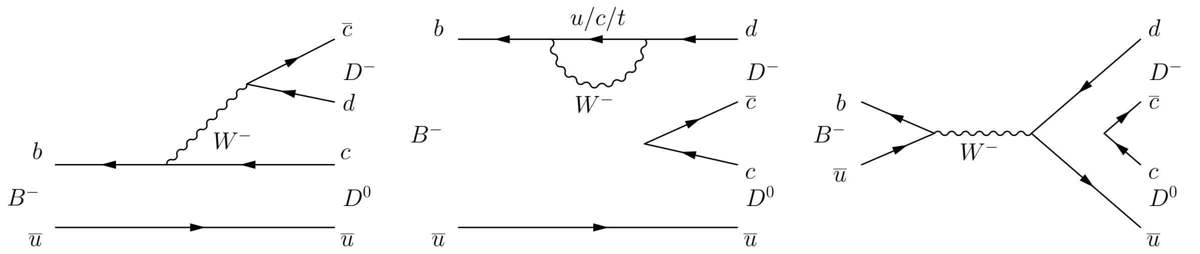

Figure 2.10: (Left) colour-allowed tree, (centre) QCD penguin, and (right) annihilation diagrams contributing to $B^- \to D^- D^0$. Similar diagrams apply for excited charm mesons and, with the d quark replaced by an s quark, for decays with the D^- replaaced by a D_s^-.

Table 2.2: Predicted *CP* asymmetries for B^- meson decays to two charm mesons in percent.

Decay	Ref. [81]	Ref. [84]	Ref. [82]	Ref. [85]	Ref. [83]
$B^- \to D_s^- D^0$	-0.28 ± 0.06	-	$-0.26^{+0.05}_{-0.04}$	0.30	-0.14 ± 0.25
$B^- \to D_s^{*-} D^0$	-0.065 ± 0.005	-	$-0.07^{+0.03}_{-0.02}$	0.08	-0.01 ± 0.10
$B^- \to D_s^- D^{*0}$	0.045 ± 0.015	-	$0.03^{+0.02}_{-0.02}$	-0.03	0.08 ± 0.03
$B^- \to D^- D^0$	4.95 ± 1.08	$0.6^{+0.6}_{-0.1}$	$4.4^{+1.1}_{-0.4}$	-4.8	-
$B^- \to D^{*-} D^0$	1.19 ± 0.16	$0.1^{+0.6}_{-0.1}$	$1.2^{+0.4}_{-0.3}$	-1.4	-
$B^- \to D^- D^{*0}$	-0.80 ± 0.35	$-0.5^{+0.1}_{-0.4}$	$-0.6^{+0.4}_{-0.2}$	0.4	-
$B^- \to D^{*-} D^{*0}$	1.19 ± 0.16	$0.2^{+0.0}_{-0.1}$	$1.2^{+0.4}_{-0.3}$	-1.4	-

The amplitude of the QCD penguin diagrams is difficult to predict. Therefore, although observing a *CP* asymmetry of several percent in a CF decay would provide convincing evidence of NP, it would be difficult to distinguish between enhanced QCD penguin amplitudes and NP for observation of a moderate but larger than predicted *CP* asymmetry [86, 87]. This discrimination can be improved through a simultaneous analysis of *CP* observables and branching fractions for all double charm decays of the B^- and $B^0_{(s)}$ mesons. Sum rules relating amplitudes, and therefore expressions relating observables, can be established using SU(3) flavour symmetry. Small corrections are applied for SU(3) symmetry breaking and highly suppressed diagrams [86]. For example, $B^- \to D_s^- D^0$ and $\bar{B}^0 \to D_s^- D^+$ decays have the same amplitude except for small corrections from annihilation and EW penguin diagrams,

$$\frac{\mathcal{A}(B^- \to D_s^- D^0)}{\mathcal{A}(\bar{B}^0 \to D_s^- D^+)} = 1 + \mathcal{O}(10^{-3}). \tag{2.26}$$

QCD penguin amplitudes and SU(3) symmetry breaking corrections cancel in the ratio between these amplitudes. Similarly,

$$\frac{\mathcal{A}(B^- \to D^- D^0)}{\mathcal{A}(\bar{B}^0 \to D^- D^+) + \mathcal{A}(\bar{B}^0 \to \bar{D}^0 D^0)} = 1 + \mathcal{O}(10^{-2}). \tag{2.27}$$

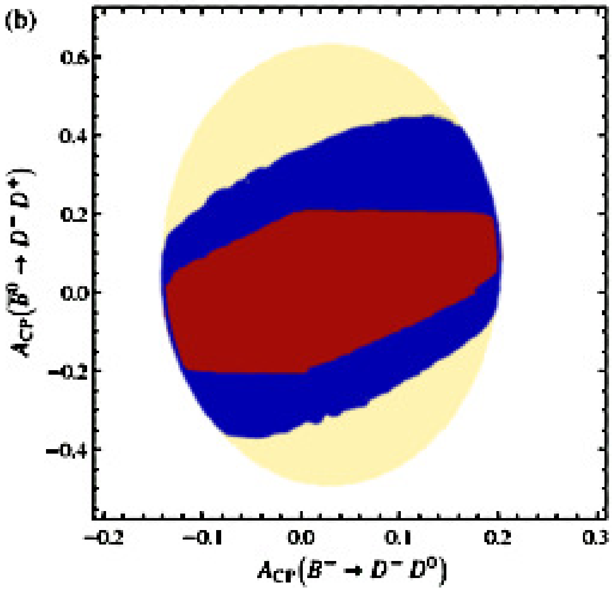

Figure 2.11: (Yellow) Experimental constraints on $\mathcal{A}^{CP}(B^- \to D^- D^0)$ and $C(\bar{B}^0 \to D^- D^+)$ (here denoted as $\mathcal{A}^{CP}(\bar{B}^0 \to D^- D^+)$). (Red) Constraints at 95% confidence level from a global fit to observables in $B \to DD$ decays using amplitude sum rules and (dark blue) similar constraints allowing for enhanced amplitudes from QCD penguin diagrams [86].

These relations can be translated into correlations between *CP* observables and branching fractions of different decay modes. For example, in Figure 2.11, the experimental constraints on *CP* asymmetries are compared to constraints using all measured beauty meson to double charm decays with amplitude sum rules. Although the constraints are currently consistent with each other, there are regions of phase space that are permitted by experimental constraints but not when enforcing amplitude sum rules, illustrating how BSM effects could be identified using more precise measurements. These examples demonstrate how a comparison of isospin-related decays may allow the detection of isospin symmetry-violating NP even in the presence of enhanced QCD penguin amplitudes. To separate between isospin symmetry-violating SM contributions and NP, decays to states involving excited charm mesons can also be measured. Measurements of annihilation-dominated decays would furthermore be particularly helpful, since the effect of NP may be more visible against a smaller SM amplitude.

In summary, a comprehensive set of measurements of the *CP* observables and branching fractions for double charm decays of the B^- and $B^0_{(s)}$ mesons would overstrain sum rules relating the amplitudes of these decays. This would improve control over hadronic uncertainties and thereby allow the SM to be tested to much higher precision than could be achieved by considering individual observables.

Chapter 3

The LHCb detector at the LHC

The results described in this thesis use data collected by the LHCb, or *Large Hadron Collider beauty* detector. The LHCb detector is one of the four large detectors analysing high-energy collisions at the LHC, and it is optimised for the study of heavy hadrons containing a beauty or charm quark. The LHC will be introduced in Section 3.1, and the LHCb detector in Section 3.2. The LHCb detector has been upgraded between 2018 and 2022 to enable the collection of larger datasets, and these upgrades are outlined in Section 3.3. Simulation of events and their interaction with the LHCb detector is essential to interpret the collected data, and Section 3.4 describes the generation of these simulated datasets.

3.1 The Large Hadron Collider

The LHC, in the CERN accelerator complex (Figure 3.1), is the largest particle collider in the world with a circumference of 26.7 km. The most common configuration is collisions between proton bunches in two counter-rotating beams. However, collisions with one or both beams replaced by lead or xenon ion beams have also been performed. It is also the highest energy particle collider. The design *centre-of-mass energy* ($\sqrt{s}$) for proton-proton (pp) collisions is 14 TeV; collisions at $\sqrt{s} = 13.6$ TeV have already been achieved.

The LHC ring consists of eight arcs and eight straight sections. The arcs are each lined with 154 superconducting twin-bore magnets, which produce the magnetic field of up to 8.33 T required to constrain the high-momentum beams to their central orbits. Each twin-bore magnet contains two beam pipes and two sets of coils (Figure 3.2). To achieve the electrical current required to produce this magnetic field the superconducting magnets are cooled to below 2 K using superfluid helium [89]. The straight sections are approximately 528 m long and lined by radio-frequency cavities oscillating at 400 MHz, which accelerate the beams to compensate for energy loss from synchrotron radiation in the arcs. *Crossing points* are located in the centre of each straight section. The two beams share a beam pipe for approximately 130 m at the crossing points. Four crossing points are instrumented with particle physics experiments and are also referred to as *interaction points*, and collisions are suppressed at the other four. Each interaction point has one large experiment, and optionally additional smaller experiments. The main experiments are:

- **LHCb** [31]: A forward-arm spectrometer designed for the study of weak decays of heavy hadrons containing a beauty or charm quark. This experiment is described in more detail in Section 3.2.

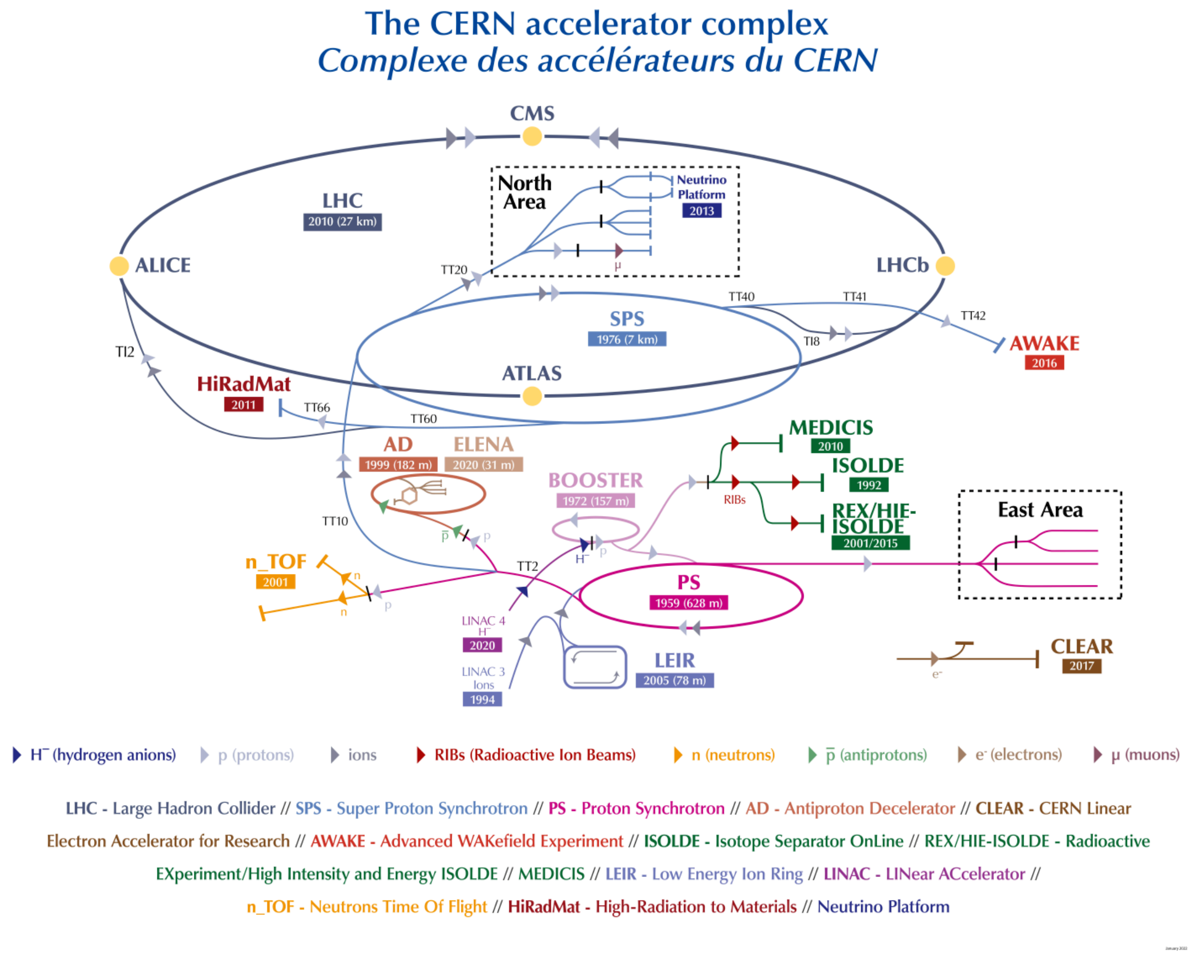

Figure 3.1: Layout of the CERN accelerator complex in 2022 [88].

- **ATLAS** [91]: A general-purpose particle physics experiment that jointly discovered the Higgs boson. Its main purpose is precision measurements of electroweak physics and the Higgs boson, and direct searches for BSM particles at the TeV scale such as supersymmetric particles. Its physics programme also encompasses studies of flavour physics, proton-ion, and ion-ion collisions.

- **CMS** [92]: The other general-purpose particle physics experiment at the LHC that discovered the Higgs boson jointly with the ATLAS experiment. Its physics goals are similar to the ATLAS experiment. This overlap was important to provide confidence in the discovery of the Higgs boson and will be equally important if either experiment discovers a new fundamental particle in the future.

- **ALICE** [93]: An experiment dedicated to studying the high-energy collisions of nuclei with the aim of improving understanding of QCD at at high energy densities. This encompasses studies of the quark gluon plasma, the state of matter believed to have formed just after the Big Bang.

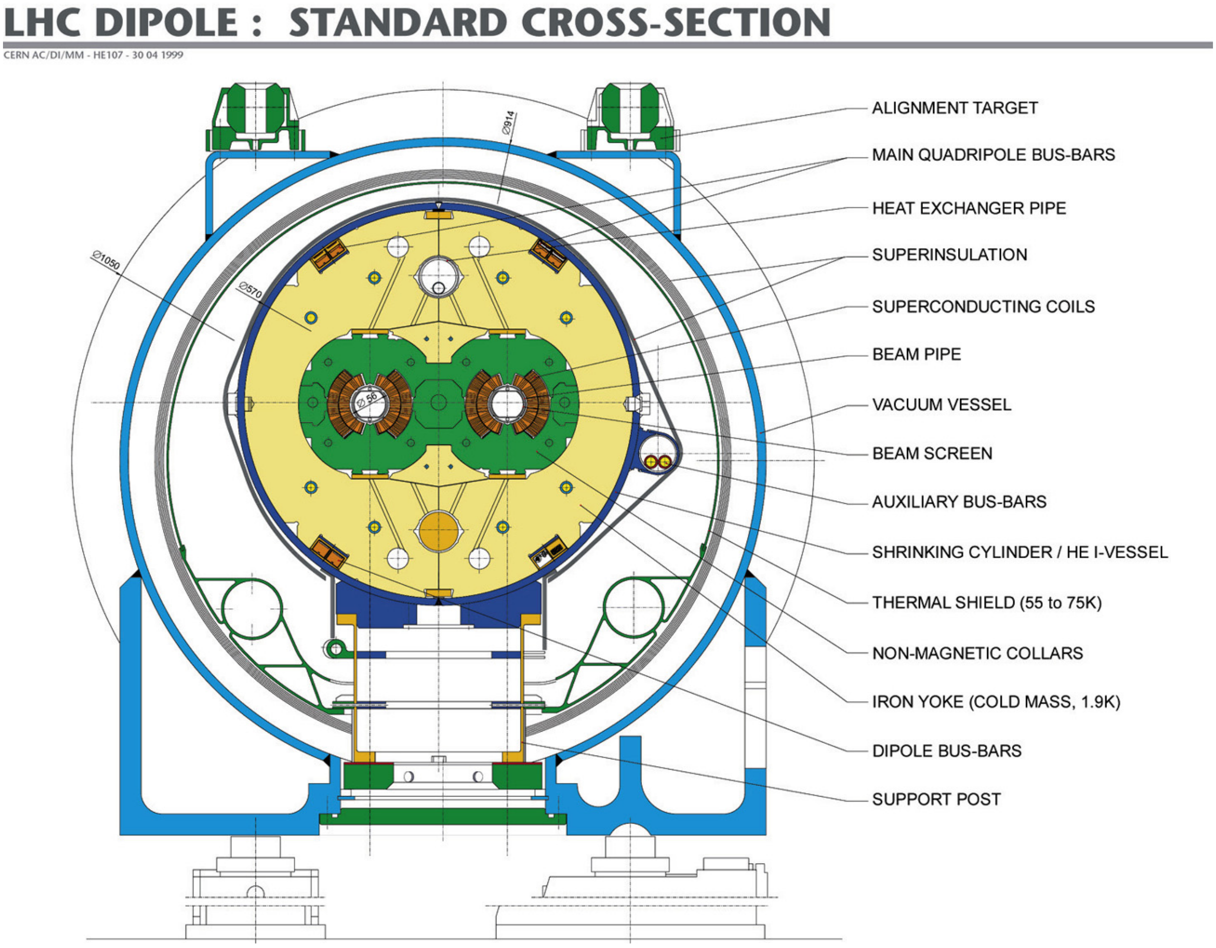

Figure 3.2: Cross-section of an LHC twin-bore dipole magnet. Two beam pipes at vacuum are each surrounded by superconducting coils that produce a strong magnetic field. The beam pipes and coils are surrounded by an iron yoke cooled to below 2 K [90].

The protons are produced by ionising hydrogen gas in an electric field. Before injection into the LHC they are accelerated in several stages: First to 50 MeV in the LINAC2, then 1.4 GeV in the Booster, then 25 GeV in the PS and finally 450 GeV in the SPS [94]. From 2020, LINAC2 has been replaced by LINAC4, which accelerates negative hydrogen ions to 160 MeV. The negative ions are subsequently transformed into protons [95]. An LHC fill is the injection of proton bunches that circulate for up to twelve hours. During the fill the number of protons decreases, mostly as a result of collisions [94]. Once the number has fallen sufficiently, the beam is dumped and a new fill is started.

The luminosity ($\mathcal{L}$) is the ratio between the number of interactions per unit time and the interaction cross-section (σ), and is therefore proportional to the number of collision events in a bunch crossing. The design luminosity of the LHC for pp collisions was $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. However, collisions with luminosity up to double this value were achieved between 2015 and 2018. The luminosity will be increased to $5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ by approximately 2030.

LHCb benefits from a lower luminosity than the ATLAS and CMS experiments since a small number of interactions per bunch crossing accommodates precise vertex and track

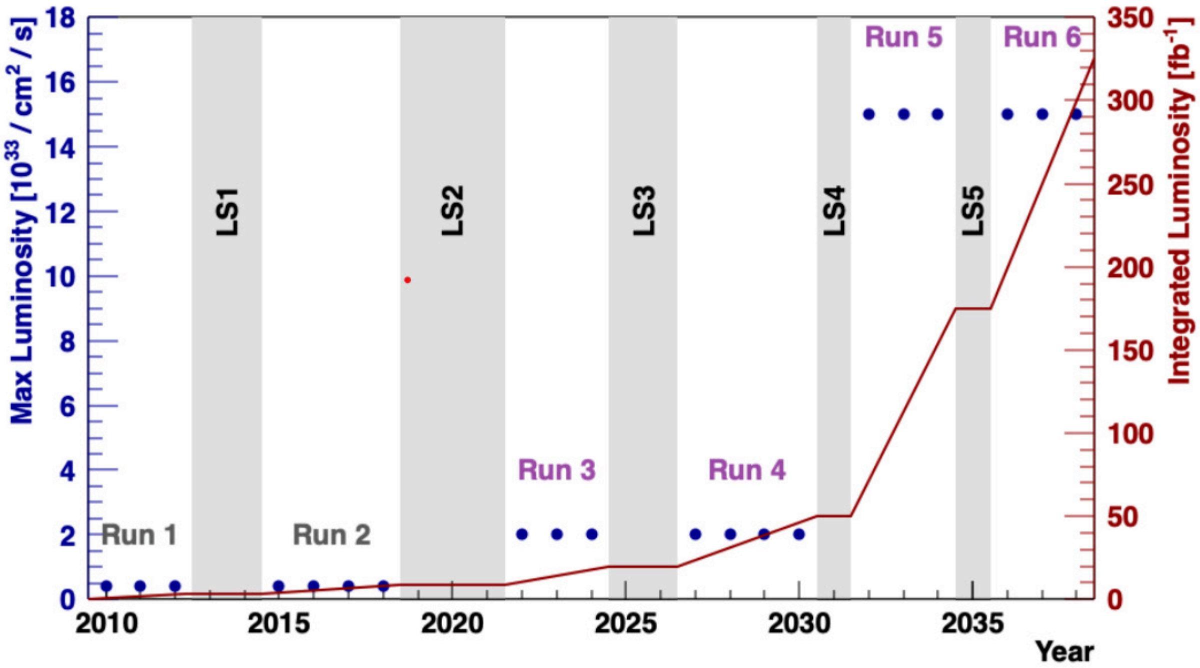

Figure 3.3: Evolution of the LHCb peak and time-integrated luminosity [100]. Exact years are only indicative and subject to changes in the LHC schedule.

reconstruction [96]. The luminosity at the LHCb interaction point is reduced with a *levelling* process where the overlap between the colliding bunches is reduced locally [97]. Levelling is also used more generally in the LHC to maintain a constant luminosity as the maximum possible luminosity naturally declines, to reduce variations in detector occupancy, and to allow the trigger configuration to remain constant during a fill [98]. The achieved and expected peak luminosities and time-integrated luminosities ($\int \mathcal{L}dt$) at the LHCb interaction point are shown in Figure 3.3, where peak luminosity denotes the luminosity at the start of each fill. Peak luminosities of up to $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ were used in LHCb data taking until 2018. In 2022, the peak luminosity has been increased to $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The peak luminosity will be increased again to $1.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ after 2030. The *pile-up* is the average number of *pp* interactions per visible event, which depends mostly on the instantaneous luminosity but also on $\sqrt{s}$. The average pile-up was approximately 2 for data taking at $\mathcal{L} = 4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and is expected to increase to 6 for post-2022 data taking at $\mathcal{L} = 2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [99] and 40 for post-2030 data taking at $\mathcal{L} = 1.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [100].

The LHC schedule is divided into a series of operational *runs* separated by *long shutdowns* (LS). Major upgrades to the experiments and the LHC are implemented in each LS. To date Run 1 and Run 2 have been completed and are mostly differentiated from each other by a significant increase in $\sqrt{s}$. Run 3 began in 2022 and is marked by a five-fold increase in luminosity at the LHCb interaction point and a small increase in $\sqrt{s}$. The past and expected schedule of LHC operational runs is given in Table 3.1.

Table 3.1: Past and expected schedule of LHC operational runs, and anticipated peak and integrated luminosities [101, 102]. The expected time period for Run 6 has not been denoted in Ref. [101].

Run	Time period	$\sqrt{s}$/ TeV	LHC peak $\mathcal{L}$ [cm^{-2} s^{-1}]	LHCb peak $\mathcal{L}$ [cm^{-2} s^{-1}]	LHCb $\int \mathcal{L} dt$ [fb^{-1}]
Run 1	2010-12	$7-8$	8×10^{33}	4×10^{32}	3
Run 2	2015-18	13	2×10^{34}	4×10^{32}	6
Run 3	2022-25	13.6	2×10^{34}	2×10^{33}	14
Run 4	2029-32	14	5×10^{34}	2×10^{33}	27
Run 5	2035-38	14	5×10^{34}	1.5×10^{34}	125
Run 6	-	14	5×10^{34}	1.5×10^{34}	150

3.2 The LHCb detector

The LHCb detector is a single-arm forward spectrometer at the LHC optimised for the study of heavy flavour physics. In particular, the detector was designed to precisely measure the CKM matrix and *CP* violation, to investigate rare beauty hadron decays involving leptons, and to perform heavy flavour spectroscopy [103]. However, LHCb has also shown excellent performance in areas including electroweak, heavy ion, and fixed target physics.

The layout of the LHCb detector is shown in Figure 3.4. The beam pipe is along the z-axis. The interaction point is at $(x, y, z) = (0, 0, 0)$ and particles traverse the detector in the positive z direction. A dipole magnet bends charged particle trajectories in the (x, z) plane to allow for precise momentum measurements. Subdetectors at lower (higher) z than the magnet are referred to as *up(down)stream* of the magnet.

The detector acceptance is an angle of 15—300(250) mrad in the bending (non-bending) plane of the magnet. This corresponds to a pseudorapidity (η) of $2.1 < \eta < 4.9$, where

$$\eta \equiv -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{3.1}$$

for a polar angle θ. This narrow acceptance covers 27% of produced b quarks (Figure 3.5), which tend to be boosted along the beam pipe by the dominant gluon-gluon fusion production mechanism [104]. Quarks produced in pp collisions rapidly hadronise with each other and form colour-singlet hadrons. The production cross-sections of mesons formed by the hadronisation of a b quark with a lighter antiquark with $2.0 < \eta < 4.5$ and $p_{\mathrm{T}} < 40\,\mathrm{GeV}/c$, where p_{T} is the momentum transverse to the beam axis, are given in Table 3.2. The differential production cross-sections decrease with increasing p_{T} and comparatively few mesons are produced with $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$. All cross-sections are approximately doubled at $\sqrt{s} = 13\,\mathrm{TeV}$ compared to $\sqrt{s} = 7\,\mathrm{TeV}$. The *fragmentation fraction* (f) expresses the fraction of b quarks that hadronise into each species. The ratio of the B_c^- to B^- (f_c/f_u) and the B_s^0 to B^0 (f_s/f_d) fragmentation fractions is approximately

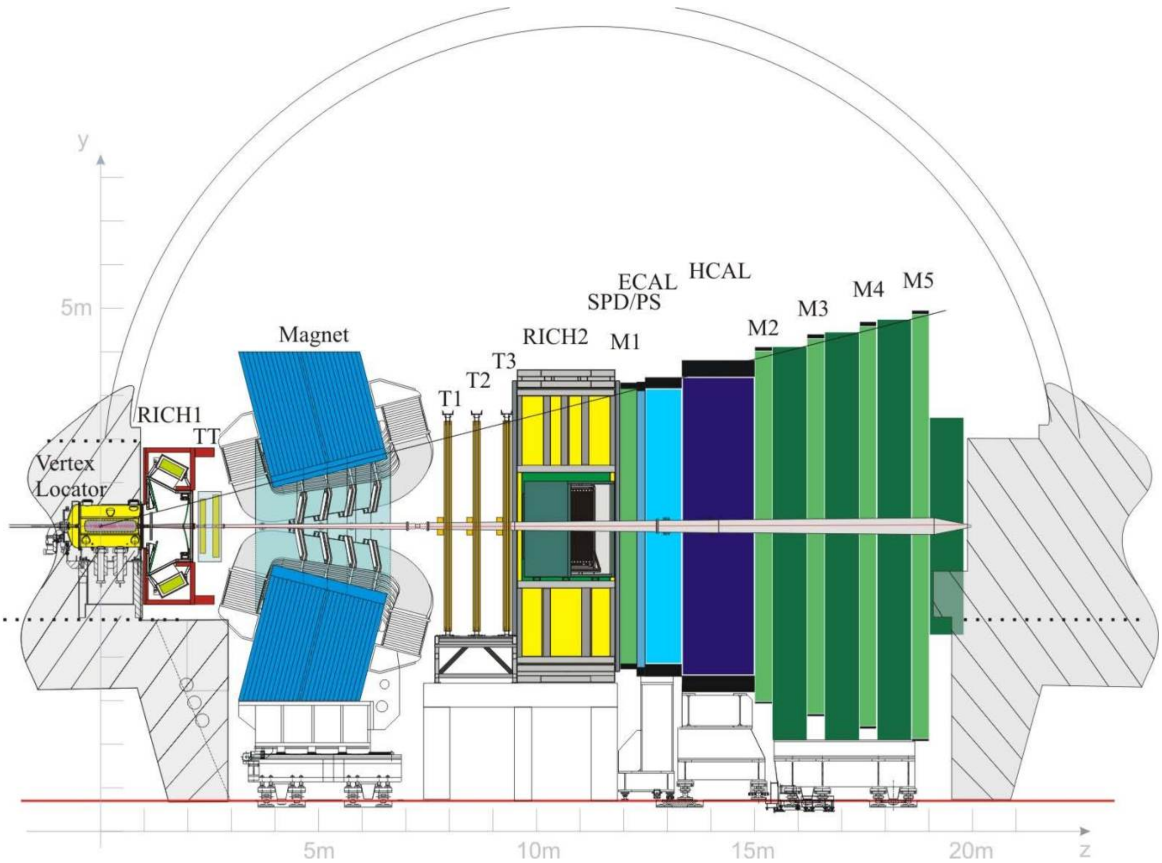

Figure 3.4: A schematic diagram of the LHCb detector at the LHC [31]. The beam pipe is along the z-axis. The interaction point is $(x, y, z) = (0, 0, 0)$.

unchanged at different centre-of-mass energies from 7 to 13 TeV [105, 106].

Physics measurements with the LHCb detector generally require full reconstruction of the decay vertices, identification of particle species, and knowledge of momentum to achieve good mass resolution. This information also allows interesting signal candidates to be separated from uninteresting events, first by the trigger system, which reduces the event rate to a sufficient level that events can be stored permanently, then in offline selection. This information is acquired with a series of subdetectors, which each have a specific purpose. Progressing downstream, these are as follows:

- **Vertex Locator (VELO):** A silicon microstrip detector surrounding the interaction region. Its primary aim is to identify secondary vertices displaced from the location of the collision (*primary vertex*, PV) consistent with long-lived hadrons.

- **Ring Imaging Cherenkov detector 1 (RICH1):** This subdetector identifies the species of charged particles according to their Cherenkov light cones.

- **Tracker Turicensis (TT):** A tracking detector upstream of the magnet.

- **Magnet:** A dipole magnet bends charged particle trajectories to enable momentum measurements.

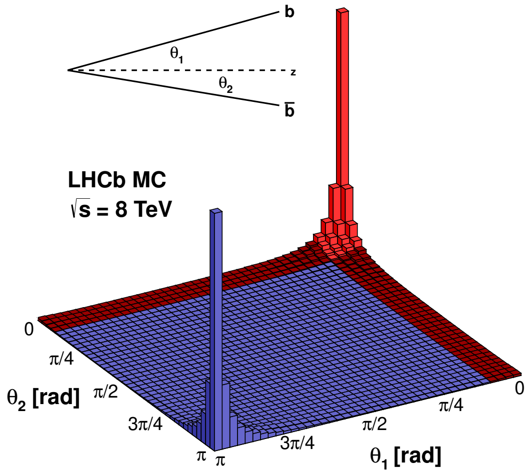

Figure 3.5: Simulated production distribution of b quarks as a function of polar angle from the beam pipe [107].

Table 3.2: Production cross-sections of B mesons in the LHCb acceptance at $\sqrt{s} = 7$ and 13 TeV. Charge conjugates are included. No measurement for $\sigma(B^0)$ at 13 TeV is available, however it is expected to scale with $\sqrt{s}$ similarly to $\sigma(B^-)$. Additionally no measurement for $\sigma(B_s^0)$ at 13 TeV is available, however $\frac{f_s}{f_d}$ is known [105] and is similar to its value at 7 TeV.

Species	σ / μb 7 TeV	13 TeV	Source
B^-	43.0 ± 3.0	86.6 ± 6.4	[108]
B^0	38.1 ± 6.0	–	[109]
B_s^0	10.3 ± 0.8	–	[105]
B_c^-	0.31 ± 0.07	0.65 ± 0.17	[106]

- **Tracking stations (T1-3):** Three tracking stations downstream of the magnet able to measure momentum precisely.
- **RICH2:** A second RICH detector, optimised for higher momentum particles than RICH1.
- **Calorimeter (CALO):** A series of calorimeters that aid in particle identification (PID) and provide energy measurements.
- **Muon system (MUON):** Five muon stations (M1-5) that identify and measure the momentum of muons.

The tracking detectors are described in more detail in Section 3.2.1 and the particle identification detectors in Section 3.2.2. Section 3.2.3 explains how the information from the subdetectors is used by the trigger system to decide which events to save, and Section 3.2.4 describes how the data is processed offline.

3.2.1 Tracking

Vertex Locator

The VELO is a silicon microstrip tracking detector surrounding the interaction region. This subsystem measures the relative position of production and decay vertices, which allows the identification of b and c hadrons that tend to have long lifetimes, and therefore the rejection of backgrounds from random combinations of tracks. The decay time of beauty and charm hadrons can also be measured using information from the VELO, which for example is essential in time-dependent measurements such as of the B_s^0 mass difference from flavour oscillations [110].

The VELO is 1 m long in the direction of the beam pipe and is divided in two, with one half on each side of the beam pipe. Each half contains 21 semicircular modules in series along the beam direction. There is a small overlap between the halves to ensure full angular coverage and to assist with alignment [111]. Each module consists of one R and one Φ n^+-on-n sensor, which measure radial distance from the beam axis and azimuthal angle around the beam axis, respectively. Each sensor has a diameter of 84 mm and contains 2048 strips. The diameter is such that all tracks inside the LHCb acceptance will cross at least three VELO stations [111].

An inner hole in the sensors of radius 8 mm accommodates the beam. The sensors must not be so close to the unstable beams during injection, so the VELO is retracted in these periods [31, 112]. The VELO is enclosed and separated from the beam vacuum within a vacuum vessel. The beam-facing surface is a 0.3 mm aluminium sheet chosen for its low material budget and the protection that it provides the sensors from beam-induced electromagnetic fields [113]. The aluminium sheet is corrugated to allow overlap between VELO halves and to further reduce the material budget [112, 114].

Dipole magnet

A particle of charge q moving with velocity $\vec{v}$ in a magnetic field $\vec{B}$ is deflected in the direction perpendicular to its velocity and the magnetic field by the Lorentz force,

$$\vec{F} = q\left(\vec{v} \times \vec{B}\right). \tag{3.2}$$

A particle moving through a magnetic field of length L is consequently deflected by an angle

$$\alpha = \frac{q}{p_\perp} BL, \tag{3.3}$$

where $p_\perp$ is the particle momentum perpendicular to the magnetic field [115]. Therefore the momentum of a charged particle can be derived if its trajectory before and after passing through a magnetic field is known. In the LHCb detector, a dipole magnet with an integrated

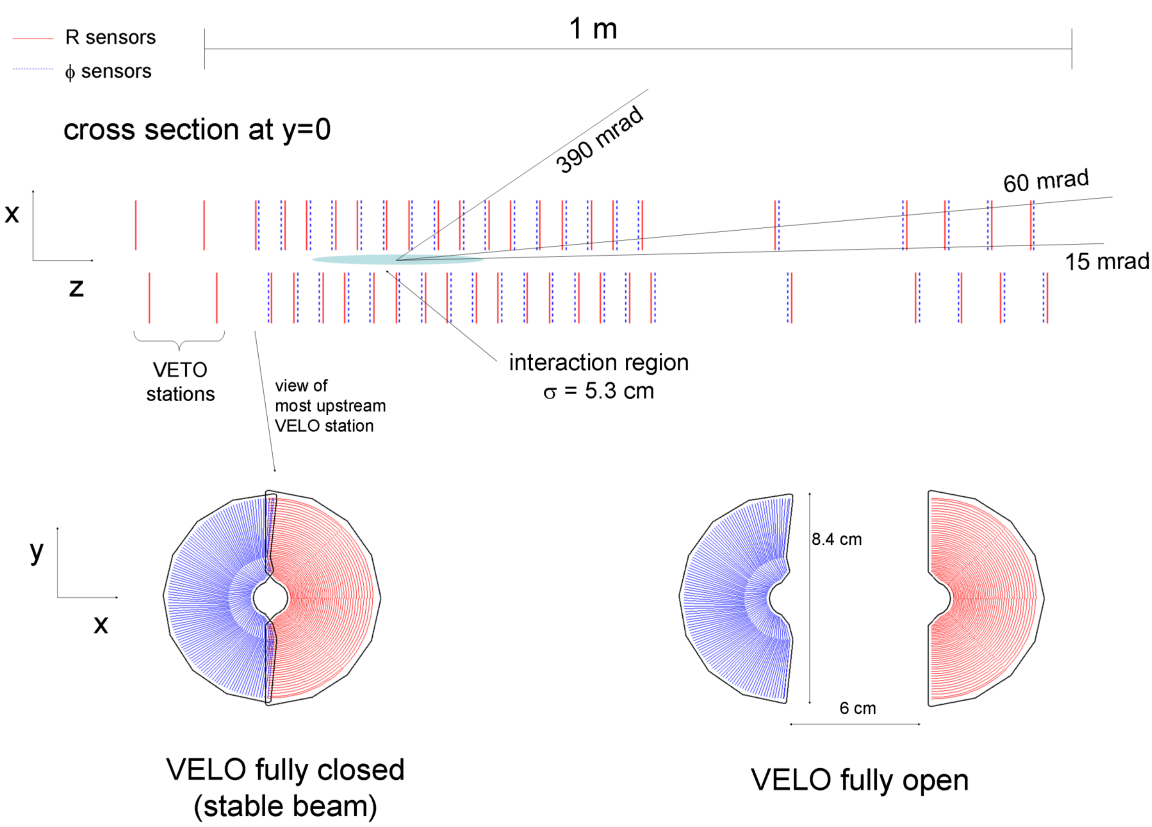

Figure 3.6: (Top) Layout of the VELO sensors when fully closed in the (x, z) plane relative to the interaction region. (Bottom) Front face of a (red) R and (blue) Φ sensor in the (left) fully closed and (right) fully open positions.

field of 4 Tm bends charged particle trajectories to enable momentum measurements using tracks reconstructed both up and downstream of the magnet.

Two saddle-shaped aluminium coils are mounted inside a window-frame yoke of laminated steel. Each coil is formed of fifteen pancakes that are electrically connected in series. Each pancake is wound from fifteen turns of aluminium. The aluminium has a circular channel in the centre of its cross-section for water cooling. This layout is illustrated in Figure 3.7. The coils carry 5.8 kA resulting in power consumption of 4.2 MW. [31, 116].

Oppositely charged particles are deflected in opposite directions, so left-right detector differences lead to charge asymmetries in detection. These asymmetries are cancelled to a large extent by reversing the magnet polarity with a frequency of approximately two weeks, corresponding to the frequency of detector and trigger calibrations. The two polarities are referred to as *MagUp* and *MagDown*. Approximately the same quantity of data was collected with each magnet polarity in each year from 2012 to 2018. In 2011, 30% more data was collected with the down polarity [117].

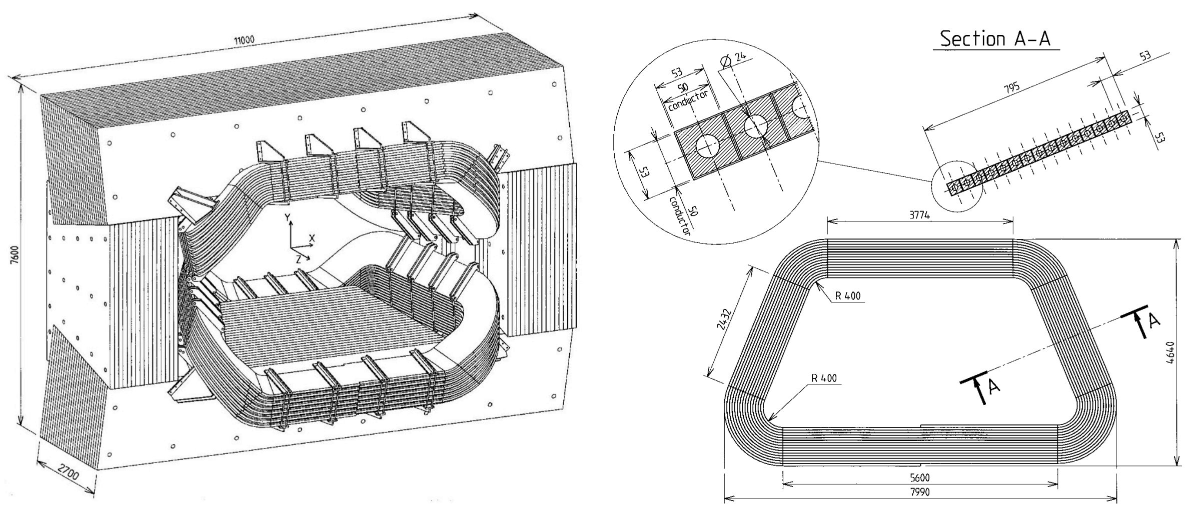

Figure 3.7: (Left) Layout of the dipole magnet. (Right) cross-section of a pancake showing the fifteen turns, and the square cross-section of the aluminium conductor with a circular channel in the centre for water cooling [116].

Tracker Turicensis

The TT is a silicon tracker upstream of the dipole magnet covering the full LHCb acceptance. It assists in matching track segments between the upstream and downstream detectors and provides p_{T} measurements, which are used by the trigger system to make decisions. It is also essential for reconstructing the trajectories of low-momentum particles that are bent out of acceptance before the downstream tracking stations by the dipole magnet and of long-lived neutral particles that decay outside of the VELO [118, 119].

The tracker consists of four detection layers of 500 µm thick p^+-on-n silicon microstrip sensors with strip pitch 200 µm. The layers have an (x-u-v-x) arrangement where the strips are vertical in the first and fourth layers and rotated by ±5° around the beam axis in the second and third layers. The layers are grouped into two pairs with 27 cm separation between the pairs (Figure 3.8). The primarily vertical orientation of the strips allows for precise measurements in the bending plane, which are required for good momentum resolution. The small rotations and the separation of the detection layers enable three-dimensional trajectory reconstruction [31, 118, 120]. The single hit resolution is 59 µm and the hit efficiency is 99.3% for charged particles with $p > 10$ GeV/c [121].

Inner tracker

The IT forms the central region around the beam pipe of the three T stations. It is used in place of the less expensive gaseous straw tube technology used for the outer region of the T stations to achieve higher granularity. This is necessary to maintain sufficiently low occupancy in this region of high particle flux. It is formed of four detector boxes surrounding the beam pipe, occupying a region 120 cm wide and 40 cm high (Figure 3.9). Silicon microstrip sensors identical to those in the TT, except for their thickness, are

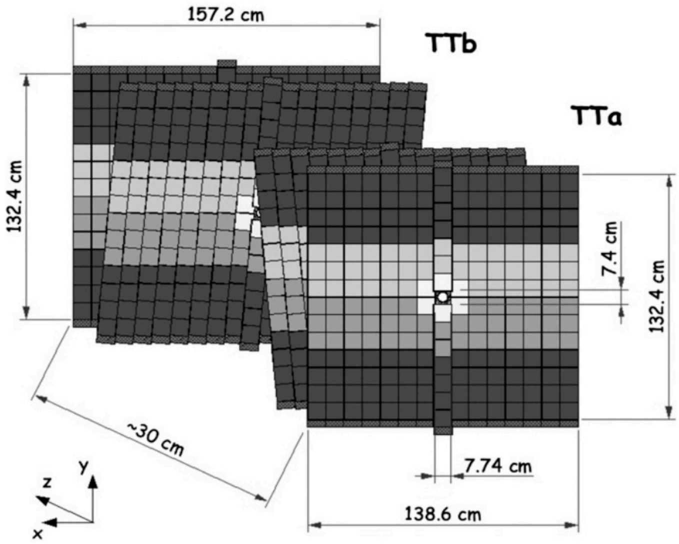

Figure 3.8: Layout of the Tracker Turicensis. The square elements are silicon sensors that each contain 512 strips [122].

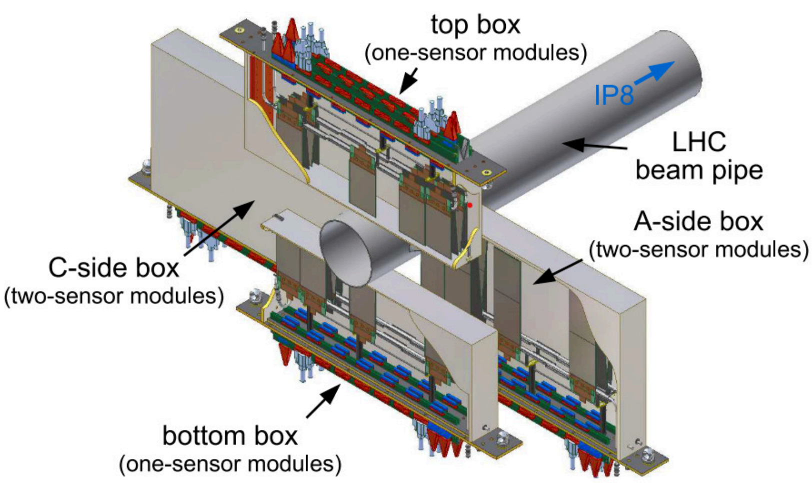

Figure 3.9: Arrangement of the four IT detector boxes relative to the LHC beam pipe [31].

employed. Sensor thicknesses of 320-410 μm are chosen as a compromise between maximising the signal-to-noise ratio and minimising material budget. The four layers are orientated in an (x-u-v-x) configuration similarly to the TT [31, 120]. The single hit resolution is 50 μm and the hit efficiency is 99.7% for charged particles with $p > 10\,\mathrm{GeV}/c$ [121].

Outer tracker

The OT is a gaseous straw tube tracking detector that covers the full detector acceptance except for the small region covered by the IT and forms the remainder of the T stations. An $Ar/CO_2/O_2$ mixture provides drift time of below 50 ns in all tubes, which minimises overlap between bunch crossings. The straw tubes have an inner diameter of 4.9 mm and are 2.4 m long. Each station consists of four module layers in an (x-u-v-x) configuration, which

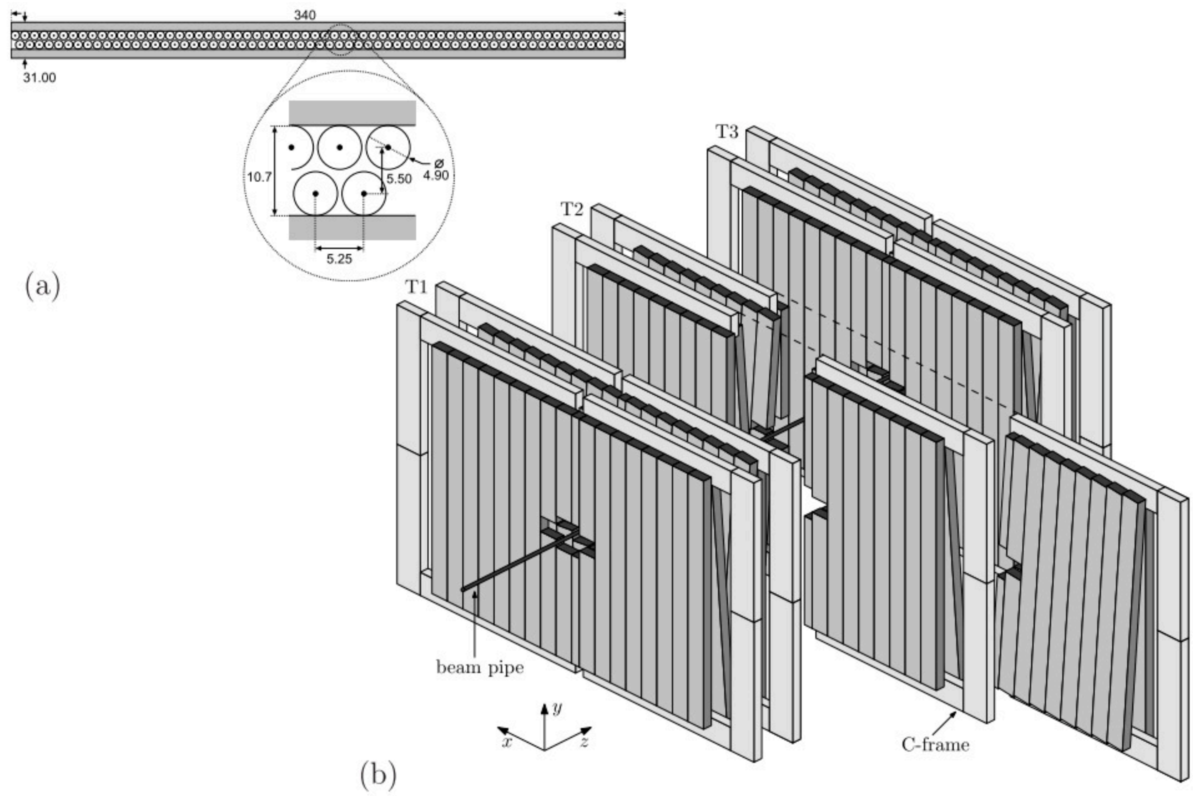

Figure 3.10: (a) Cross-section of an OT module layer. (b) Arrangement of OT modules in stations and shape of OT stations with a cross-shaped gap around the beam pipe for the IT [123].

each contain two staggered layers of straw tubes (Figure 3.10). The single-hit resolution is 200 μm and the single-hit efficiency is 99.2% in the central half of the straw [123].

Track reconstruction

Tracks are reconstructed by combining approximate straight line segments of hits in the upstream and downstream tracking detectors. The types of reconstructed track are illustrated in Figure 3.11 and are defined as [98]:

- **Long tracks:** Hits in the VELO and T stations, and optionally also TT. These tracks have the best momentum resolution and are preferentially used for physics analysis.

- **Upstream tracks:** Hits in the VELO and TT only. These tracks are associated with low momentum particles. Their reconstruction allows the corresponding background to be subtracted when deriving information from the RICH1 detector.

- **Downstream tracks:** Hits in the TT and T stations. These tracks are used to study long-lived particles.

- **VELO tracks:** Hits in the VELO only, which applies to large-angle or backward tracks. These tracks are useful in reconstructing the primary vertex.

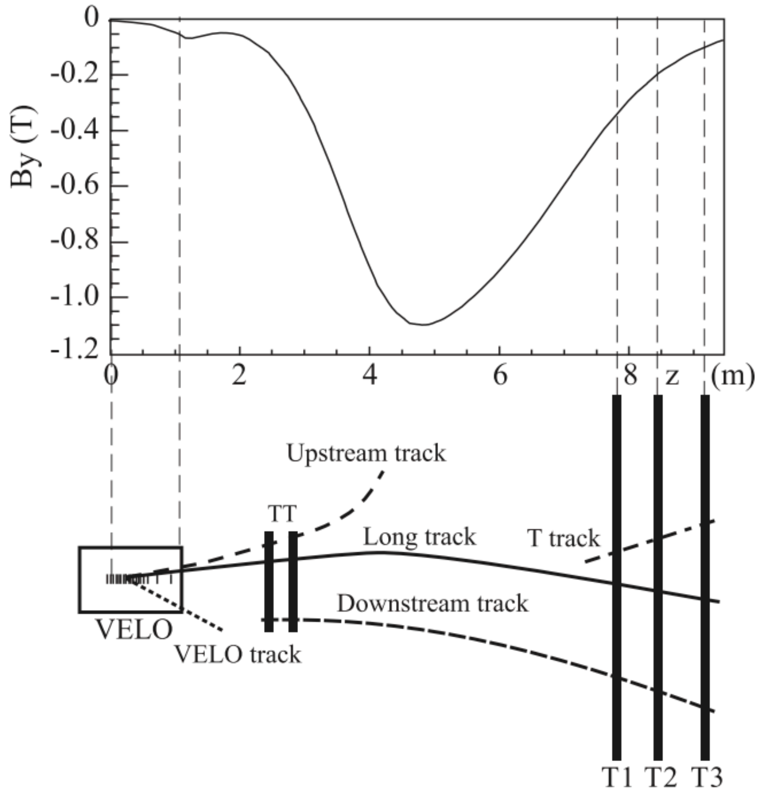

Figure 3.11: Types of reconstructed tracks defined by the tracking systems in which hits are present, and comparison of the location of tracking systems to the magnetic field distribution [98].

- **T tracks:** Hits in the T stations only. These tracks are generally the result of secondary interactions. Similarly to upstream tracks, T tracks are used to subtract the corresponding backgrounds when analysing the output from the RICH2 detector.

- **Ghost tracks:** Do not correspond to real charged particles. Most ghost tracks result from the incorrect matching between track segments in the VELO and T stations.

Two algorithms are employed to reconstruct long tracks, which both begin by identifying VELO tracks. The forward tracking algorithm combines VELO tracks with a T station hit, which is sufficient to fully define a possible trajectory. The algorithm then selects tracks by identifying the proposed trajectories that are associated with a cluster of T station hits [124]. The track matching algorithm combines track segments from the VELO and T stations [125]. Both algorithms subsequently add TT track segments that are consistent with the interpolated trajectories. Tracks are then fitted with a Kalman filter [126] taking multiple scattering and energy loss due to ionisation into account, to improve estimation of the track parameters $(x, y, dx/dz, dy/dz, q/p)$ [98]. Finally, ghost tracks are rejected using a neural network that evaluates a *ghost probability* using the fit quality; track kinematics; subdetector occupancies; and the number of hits on a track and the difference relative to the expected number [127].

The momentum resolution is 0.5% for particles with $p < 20\,\mathrm{GeV}/c$, and rises to 1% at

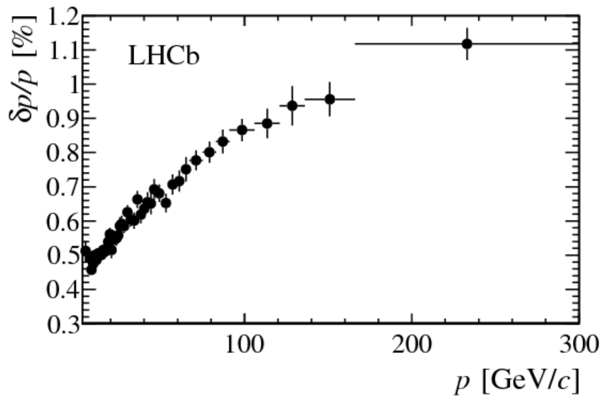

Figure 3.12: Relative momentum resolution as a function of momentum for long tracks [98].

150 GeV/c (Figure 3.12). The *tracking efficiency*, defined as the probability that the trajectory of a charged particle that has passed through the full tracking system is reconstructed, is 96% for tracks with $p > 5$ GeV/c and events with less than 200 tracks (Figure 3.13) [128].

Momentum scale calibration

Imperfect knowledge of the magnetic field map and tracking system alignment may bias the reconstructed momentum and therefore also invariant masses. Momentum scale is calibrated by comparing measured to known reconstructed mass values in $J/\psi \to \mu^+\mu^-$ and $B^+ \to J/\psi K^+$ decays [129, 130].

3.2.2 Particle identification

Ring imaging Cherenkov detectors

The RICH detectors provide discrimination between charged particle species based on their produced Cherenkov light. This discrimination is essential to distinguish between signal and background events with the same topology. For example, $B_s^0 \to D_s^- K^+$ decays used to measure the CKM phase γ [63] would be almost invisible alongside the Cabbibo-favoured $B_s^0 \to D_s^- \pi^+$ decays without the discrimination between charged pions and kaons provided by the RICH detectors. The PID information from the RICH also allows backgrounds from random combinations of tracks with incorrect flavour content to be reduced. These tracks are mostly pions that are produced abundantly in pp collisions.

As charged particles move through a medium with a refractive index (n) higher than a vacuum, they emit a cone of radiation, known as a *Cherenkov ring*, given their speed (v) is greater than c/n. The angle between the cone from the particle trajectory, the *Cherenkov angle* (θ_c), depends on n and v as

$$\cos\theta_c = \frac{c}{nv}. \tag{3.4}$$

The velocity is extracted by measuring the cone angle. The particle species is

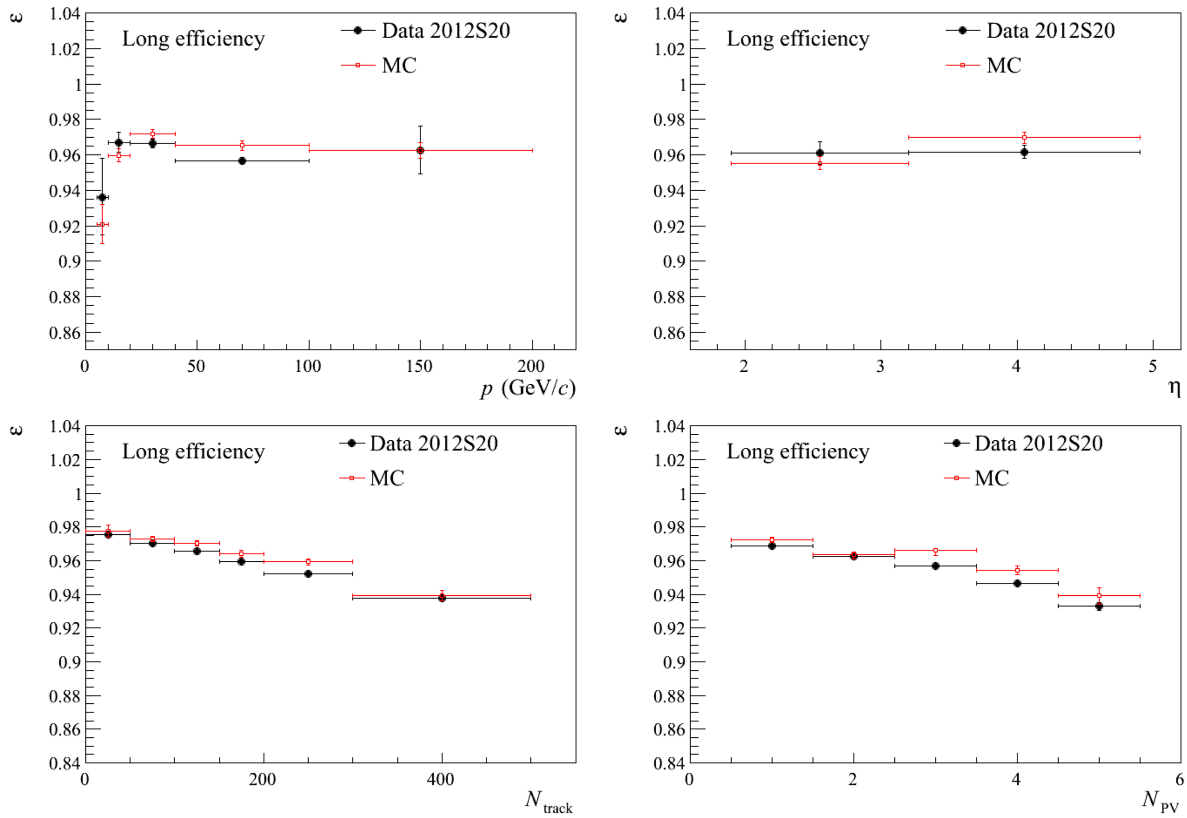

Figure 3.13: Long track efficiency as a function of (top left) track momentum, (top right) track pseudorapidity, (bottom left) number of tracks in the event, and (bottom right) number of primary vertices in the event in 2012 data and simulation. Uncertainties are statistical only [128].

subsequently determined by combining the velocity with the momentum measured by the tracking detectors. Figure 3.14 shows θ_c as a function of momentum for several particle species in three radiators that have been used in the RICH detectors. Below a momentum threshold, which depends on the particle species and radiator, no Cherenkov ring is produced. θ_c also asymptotically approaches a maximum at high momenta. Therefore at sufficiently low or high momentum it is impossible to discriminate between particle species hypotheses. Consequently, multiple radiators are required to cover the entire desired momentum range.

The RICH system is split into two subdetectors: RICH1 upstream of the magnet and RICH2 downstream. RICH1 has dimensions of $2.4 \times 2.4 \times 1.0\,\mathrm{m}^3$ resulting in an acceptance of 25 − 300 mrad. It is filled with C_4F_{10} ($n = 1.0014$), which provides positive identification for kaons in the range $9.3 < p < 40\,\mathrm{GeV}/c$. In the range $2.6 < p < 9.3\,\mathrm{GeV}/c$, pions can be positively identified and kaons or heavier species are identified as tracks without associated Cherenkov rings [131]. In Run 1, a 50 mm aerogel wall ($n = 1.03$) was also located in RICH1 to positively identify kaons with momenta as low as $2\,\mathrm{GeV}/c$ [132]. The aerogel was removed for Run 2, since it produced a low photon yield per detected track and large rings, which would become increasingly difficult to analyse in higher pile-up. Its removal also reduced the material budget by $0.05X_0$ where X_0 is the radiation length, and increased the

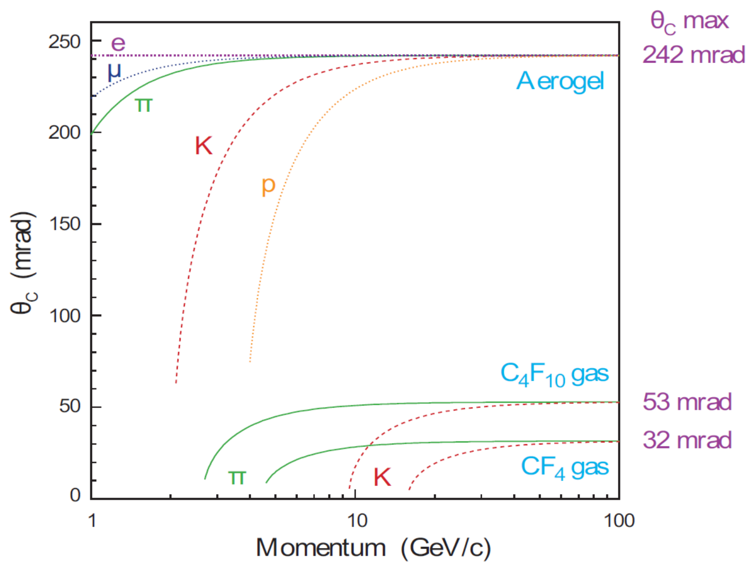

Figure 3.14: Cherenkov angle θ_c as a function of momentum for several particle species in the radiators used in the RICH detectors [31].

reconstruction speed by halving the number of Cherenkov ring candidates [131, 133, 134].

RICH2 is filled with CF_4 ($n = 1.0005$), with 10% CO_2 to reduce scintillation [134]. It provides kaon-pion discrimination for $15 < p < 100\,\text{GeV}/c$ in the acceptance $15 < \theta < 120\,\text{mrad}$ [132]. This limited acceptance is acceptable since high-momentum particles produced in the decays of forward-boosted heavy hadrons tend to have a low polar angle and are less deflected by the magnet.

Figure 3.15 illustrates the design of both RICH detectors in Run 1. Cherenkov light is detected by an array of hybrid photodiode (HPD) detectors [135] located outside of the acceptance by way of a spherical focusing mirror and a flat mirror [136].

Owing to the high background from other overlapping tracks in the event, it is not possible in general to precisely determine the Cherenkov angle associated with a specific track. Instead a global algorithm using all tracks and RICH hits produces a likelihood-based PID discriminant for each track [132, 137]. The algorithm first assumes that all tracks are pions since this is the most probable species. An event likelihood is evaluated given the observed and expected photon hit distributions, assuming Poisson statistics. The expected distribution is evaluated based on the reconstructed tracks and their assumed species, and the expected background from sources such as electronic noise. The background contribution is determined by comparing the expected and observed signal in each HPD. Then for each track separately the mass hypothesis is changed to e, μ, K and p, leaving all other hypotheses unchanged. The nominal hypothesis is updated to that which produces the largest increase in the likeihood. The background estimation and likelihood maximisation are repeated iteratively until no further improvement in likelihood is found.

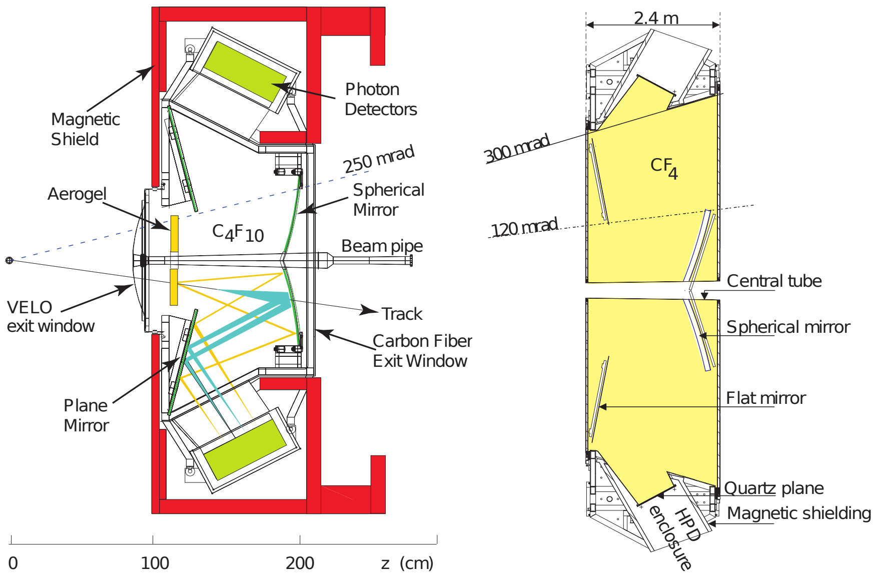

Figure 3.15: Schematic diagram of the (left) RICH1 and (right) RICH2 detectors in Run 1 [31].

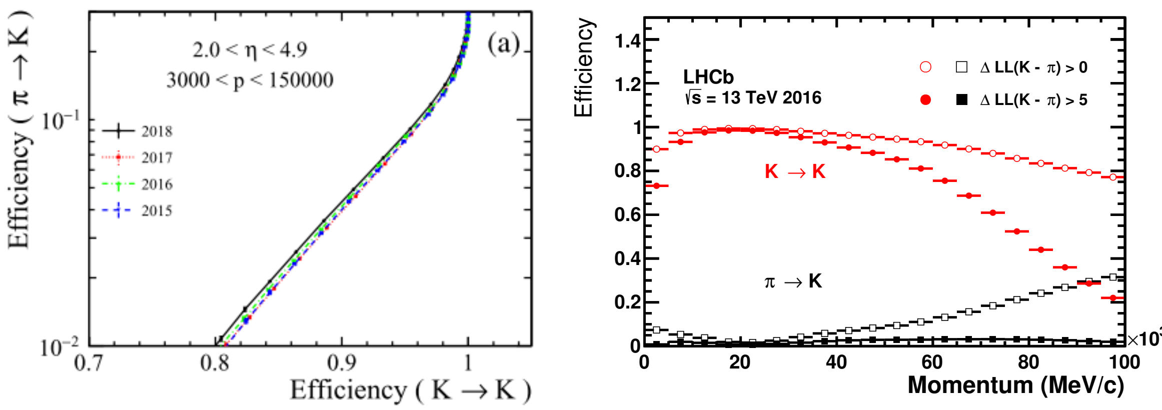

Figure 3.16: (Left) Fractional retention of pions as a function of kaon efficiency when using $\mathrm{DLL}_{K\pi}$ requirements as a discriminant [138]. (Right) Kaon and pion efficiency as a function of momentum for different $\mathrm{DLL}_{K\pi}$ requirements in 2016 magnet down data [139].

The algorithm output is a set of $\mathrm{DLL}_{x\pi}$ values,

$$\mathrm{DLL}_{x\pi} = \log(\mathcal{L}(x)) - \log(\mathcal{L}(\pi)), \tag{3.5}$$

which are the differences in overall event log-likelihood when the track hypothesis is changed from a pion to another species (x). PID criteria are imposed with requirements on DLL values. The use of this algorithm with information from the RICH detectors results in the excellent kaon-pion discrimination shown in Figure 3.16. Discriminating power is momentum-dependent and degrades at both high and low momentum. It is also more weakly dependent on the event multiplicity and track pseudorapidity. Overall, in the range $3 < p < 150\,\mathrm{GeV}/c$, $2.0 < \eta < 4.9$, 95% pion rejection is achieved at 90% kaon efficiency.

Calorimeters

The CALO system provides positive identification and energy measurements for electrons, photons, and neutral pions that convert into a pair of photons. The positive identification of and discrimination between neutral pions and photons is necessary for many measurements sensitive to BSM physics, and increases the number of final states available for measuring the CKM phase γ. The reconstruction of electron energy and bremstrahlung is essential for tests of lepton universality [140]. The calorimeters also allow for rapid triggering on high transverse energy (E_T) hadron, electron and photon candidates that often originate in heavy flavour decays [31].

The CALO system is divided into three subsystems: The scintillator pad detector (SPD)/preshower (PS), electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). Each system uses scintillator material to measure energy deposits produced in electromagnetic or hadronic showers. Showers excite the scintillator material, which then emits scintillation light with energy proportional to that of the shower. The scintillation light is transmitted by wavelength-shifting fibres to photomultipliers [31].

The SPD/PS and ECAL dimensions and segmentation into cells are related projectively. Since the hit density varies by over two orders of magnitude transverse to the beam pipe, the scintillator pads are divided into inner, middle and outer regions, with the finest segmentation in the inner region and the coarsest segmentation in the outer region. The HCAL dimensions are related projectively but the segmentation is coarser due to the larger size of the hadronic showers, and it is segmented into only an inner and outer section. The outer dimensions are equivalent to the detector acceptance [31]. The composition and segmentation of each subsystem is illustrated in Figure 3.17.

The SPD and PS are near-identical high-granularity planes of scintillator pads separated by a 15 mm($2.5X_0$) lead converter. Only charged particles produce ionisation in the SPD. The lead plate initiates electron and photon showers. By comparing energy deposits in the SPD and PS, charged and neutral particles can be distinguished. The SPD hit multiplicity is a proxy for the event multiplicity, which is also used to veto the most complex events that would be difficult to analyse even offline [141]. The PS increases the longitudinal segmentation of the electromagnetic shower, which improves discrimination between electrons and the large charged pion background [31].

The ECAL is a shashlik-type calorimeter consisting of alternating scintillator tiles and lead plates. It measures the energy and position of electrons and photons and allows triggering on high-E_T electrons and photons. Electrons and photons shower at the lead plates, similarly to in the SPD/PS. The thickness is $25X_0$, which is required to fully contain the electromagnetic shower and therefore optimise the energy resolution, which is measured to be [141]

$$\frac{\sigma_E}{E} = \frac{(9.0 \pm 0.5)\%}{\sqrt{E/\,\mathrm{GeV}}} \oplus (0.8 \pm 0.2)\% \oplus \frac{0.003}{E \sin\theta/\,\mathrm{GeV}}. \tag{3.6}$$

The HCAL has a similar design, but uses iron instead of lead to initiate hadronic showers. The total thickness is 5.6 interaction lengths due to space limitations. The energy resolution

$$\frac{\sigma_E}{E} = \frac{(67 \pm 5)\%}{\sqrt{E/\,\mathrm{GeV}}} \oplus (9 \pm 2)\% \tag{3.7}$$

is comparatively poor [141].

As well as providing input to the hardware trigger for high-E_{T} charged hadrons, a comparison of energy deposits in the PS and HCAL allows for discrimination between electrons and charged pions [142].

Muon system

The triggering on and identification of muons is important for reconstructing decays with a J/ψ in the final state, for example $B^0 \to J/\psi K^0_{\mathrm{S}}$, which is used to measure the CKM angle β [143]. Muon reconstruction is also necessary for studying leptonic or semileptonic decays that may be sensitive to BSM physics, such as $B^0_{(s)} \to \mu^+\mu^-$ [144].

The MUON system [145] consists of five rectangular muon tracking stations, denoted as M1-5, in series along the beam axis with projective geometry. M1 is upstream of the calorimeter and M2-5 are downstream. M1-3 have high spatial resolution in the bending plane. This allows p_{T} measurements with 20% resolution [31]. M2-5 are interleaved with 80 cm iron absorbers to attenuate hadrons that were not completely stopped by the calorimeters, and therefore select penetrating muons. The total absorber thickness is 20 nuclear interaction lengths, of which 6.6 interaction lengths is in the calorimeters. The layout of the MUON stations relative to the calorimeters is shown in Figure 3.18.

Over 99% of the muon system is instrumented with multi-wire proportional chambers (MWPCs) [146]. At very high flux, MWPCs cannot be segmented sufficiently finely to maintain adequately low occupancy in a channel [115], so the inner region of M1 is instead instrumented with triple-gas electron multiplier (GEM) detectors. Both types of detector are filled with an $Ar/CO_2/CF_4$ mixture [146]. Each station has a detection efficiency of over 99% [146].

3.2.3 Trigger system

A trigger system is required to reduce the event rate sufficiently to allow the output to be saved. LHCb employs a multi-level trigger system (Figure 3.19) where each level performs an improved event reconstruction at lower rate using information from more subdetectors.

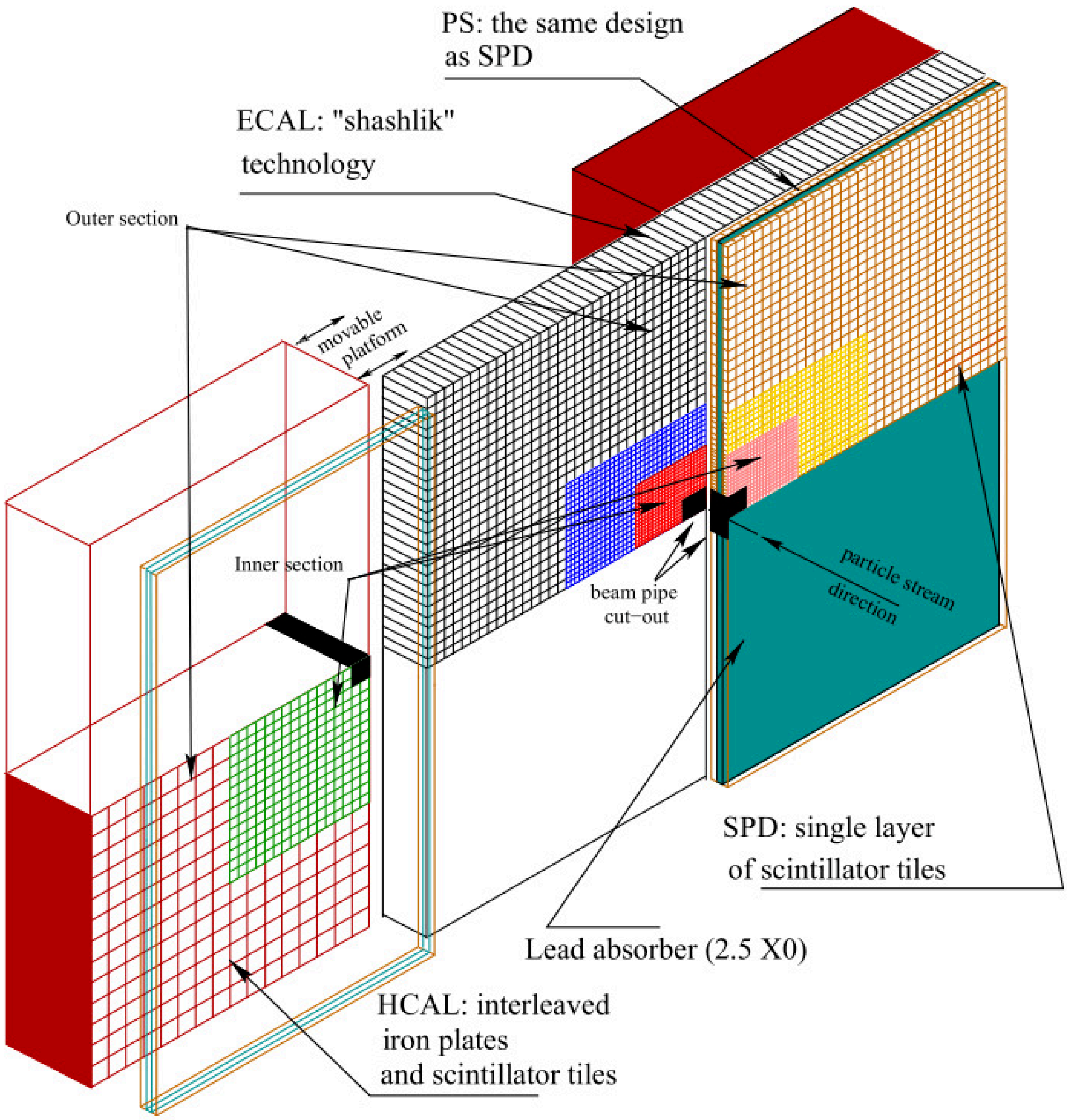

Figure 3.17: Diagram of the calorimeter system with particle stream direction indicated [141].

Real-time alignment and calibration was also introduced at the beginning of Run 2, enabling offline-quality reconstruction in the final level of the software trigger, which improved discrimination between signal candidates and background, and allowed partial events to be saved, thereby reducing bandwidth relative to signal rate.

The trigger system implements a set of algorithms at each filtering stage that decide whether to retain events and which part of the event to retain. Each algorithm is referred to as a *trigger line*. A candidate for a signal decay is *Triggered On Signal* (TOS) relative to a trigger line if the candidate itself fired this trigger line, or *Triggered Independent of Signal* (TIS) if the event excluding the candidate fired the trigger line.

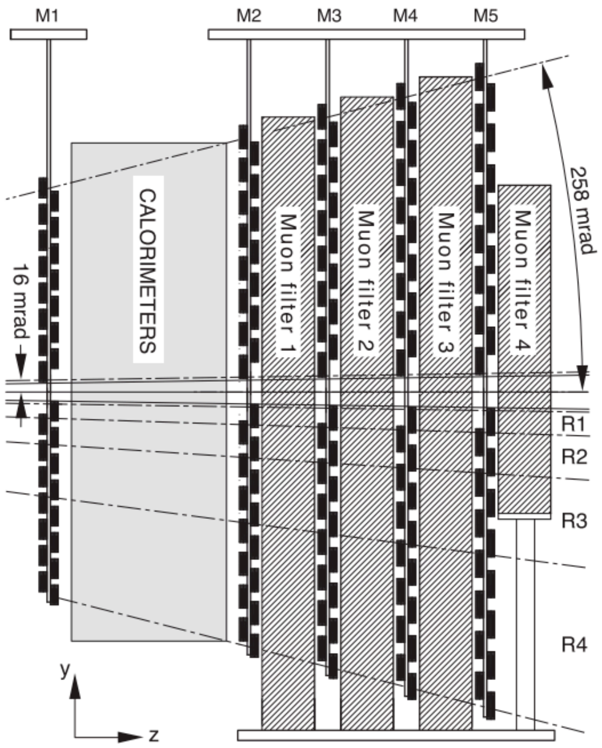

Figure 3.18: Layout of the five MUON stations relative to the calorimeters. The z-axis is parallel to the beam pipe and the y-axis is in the non-bending plane [31].

L0 trigger

The L0 hardware trigger filters events to a rate of 1 MHz using information from the calorimeters and MUON system. It is implemented on field-programmable gate arrays with a 4 µs latency. The hadron, electron and photon triggers fire if a transverse energy deposit in a 2×2 cluster of cells exceeds a threshold. Transverse energy is defined as $E_{\mathrm{T}} = E \sin\theta$, where θ is the angle between the z-axis and a line from the centre of a cell to the interaction point. The thresholds are adjusted to maintain the output rate under different LHC conditions, and the threshold for the hadron trigger is approximately 3.5 GeV/c. The muon and dimuon triggers fire upon the identification of a muon with $p_{\mathrm{T}} > 1.76\,\mathrm{GeV}/c$, or two muons with $p_{T,1} \times p_{T,2} > 1.6\,\mathrm{GeV}^2/c^2$, in the MUON system. A requirement on the number of hits in the SPD rejects high-multiplicity events that are slow to reconstruct and where discrimination between signal and background is particularly difficult. A maximum of 600(450) hits is imposed for the majority of trigger lines in Run 1(2) [148, 149]. The efficiency as a function of the true E_{T} of the particle rises sharply above the threshold and is higher for deposits in the outer region of the HCAL (Figure 3.20).

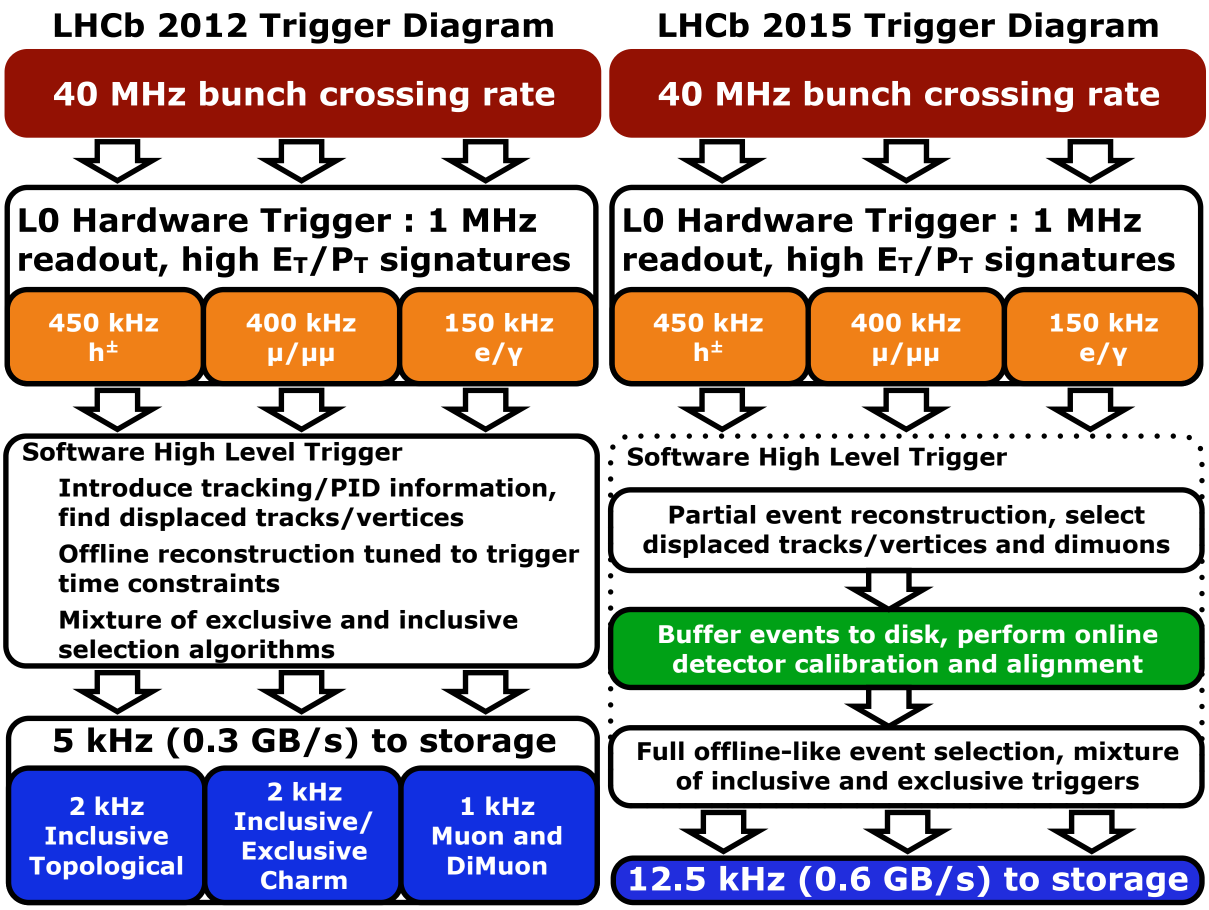

Figure 3.19: Diagram of the stages in the (left) Run 1 and (right) Run 2 trigger system, and the output rate of each stage [147].

HLT1

The events passing the L0 trigger are transferred to an event filter farm of 27000 physical cores on 1700 nodes [148] where the software trigger runs on central processing units (CPUs). The HLT1 stage of the software trigger partially reconstructs the event using information from the VELO, tracking system and MUON system. The forward tracking algorithm reconstructs long tracks, a Kalman filter is applied, and ghost tracks are rejected (Section 3.2.1). Muons are identified by adding hits in the MUON stations that are near the projected trajectory of long tracks.

Events are then filtered according to a small set of trigger lines. Most important for the selection of hadronic B decays are the two inclusive trigger lines that select events with one or two good quality high-p_{T} tracks that are significantly detached from the PV [148, 150]. Additional trigger lines select decays involving muons; low-multiplicity events for the study of central exclusive production; and decays useful for alignment and calibration. The TOS efficiency of the inclusive trigger lines for several beauty meson decays is shown in Figure 3.21.

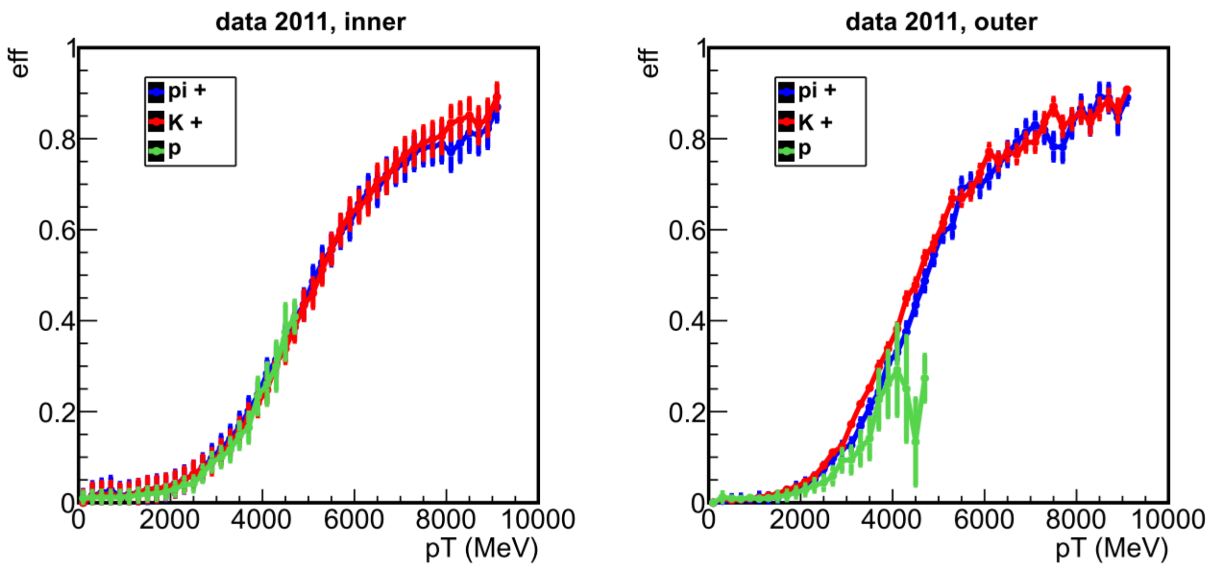

Figure 3.20: Hadron hardware trigger efficiency as a function of track p_T in the (left) inner and (right) outer regions of the HCAL for (blue) π^+, (red) K^+ and (green) p [149].

Real-time alignment and calibration

During Run 2, the output of HLT1 was written to a 10 PB disk buffer, sufficient to store two weeks of HLT1 output. Real-time detector alignment and calibration was run on the disk buffer for each fill using the output of dedicated HLT1 lines, to compute alignment and calibration constants, which were written to a conditions database (Figure 3.22). This permitted offline-quality event reconstruction in the final trigger stage [151]. The introduction of a buffer moreover allowed events to be processed between fills, which accommodated higher HLT1 output rates [148].

HLT2

The final trigger stage uses information from all subdetectors to perform a more complete event reconstruction than in HLT1 including particle identification; calorimeter cluster reconstruction including shower energy measurements and neutral cluster reconstruction; and improved track reconstruction with better momentum resolution. This allows the calculation of quantities such as track momenta, shower energies and DLL variables. In Run 2, the real-time alignment and calibration allowed offline-quality event reconstruction at this point, which permitted higher signal efficiencies.

Beauty hadron decays are mostly selected using inclusive *topological* trigger lines that occupy 40% of the HLT2 output rate. A rectangular selection followed by a Boosted Decision Tree (BDT) selects two to four-track vertices with high p_T, large displacement from the PV, and a topology consistent with a beauty hadron decay [148, 152–154]. The TOS efficiency of these topological trigger lines for certain beauty hadron decays is shown in Figure 3.23. The remaining 60% of the output rate is occupied by muon and dimuon trigger lines; trigger

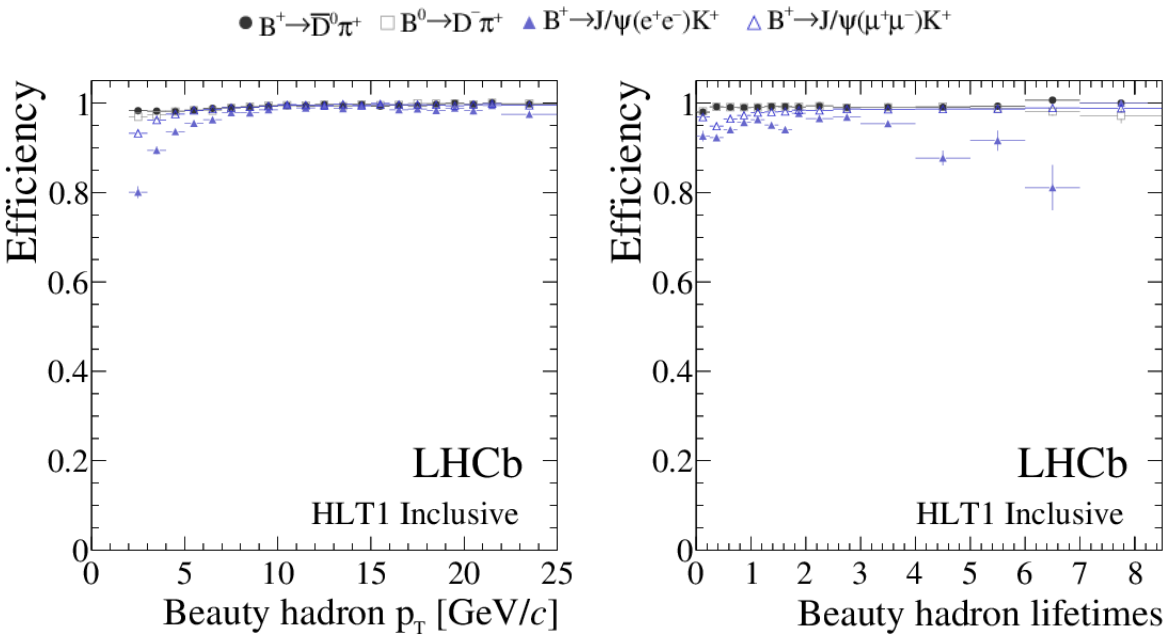

Figure 3.21: TOS efficiency of the HLT1 inclusive trigger lines in Run 2 as a function of (left) p_T and (right) rest frame lifetime for selected beauty meson decays [148].

lines for low-multiplicity events to study central exclusive production; and exclusive lines selecting specific decays for studies including charm physics, electroweak physics, exotic searches, and calibration of track reconstruction and PID [148].

Selected events are written to magnetic tape. Since event reconstruction in HLT2 was of offline-quality during Run 2, it was possible to only save part of the event. As little as only the signal candidate could be saved, referred to as the *Turbo* model, which would reduce the event size by an order of magnitude. This model was mostly used for charm decays. Additional reconstructed objects (*TurboSP*) or the full reconstruction (*Turbo++*) could also be saved if important for the subsequent physics analysis [151].

3.2.4 Offline reconstruction and selection

Subsequent stages of the dataflow are implemented offline, and are shown relative to the online stages in Figure 3.24. Reconstruction is performed on the saved raw events, followed by further selection and the calculation of higher-level quantities. The final output is ROOT [155] *ntuple* files suitable for use in physics analysis. This section describes the offline reconstruction stages. Different software applications are used at each stage, and are built on GAUDI [156], an object oriented framework that provides a common infrastructure and environment for the software applications used by LHCb.

The BRUNEL application reconstructs the raw events saved by the HLT2 trigger [157] using alignment and calibration constants for the corresponding fill. In Run 2, this is of similar quality to the reconstruction in HLT2 but the algorithms in this offline reconstruction may be updated. The reconstructed events are saved in the Data Summary Tape (DST)

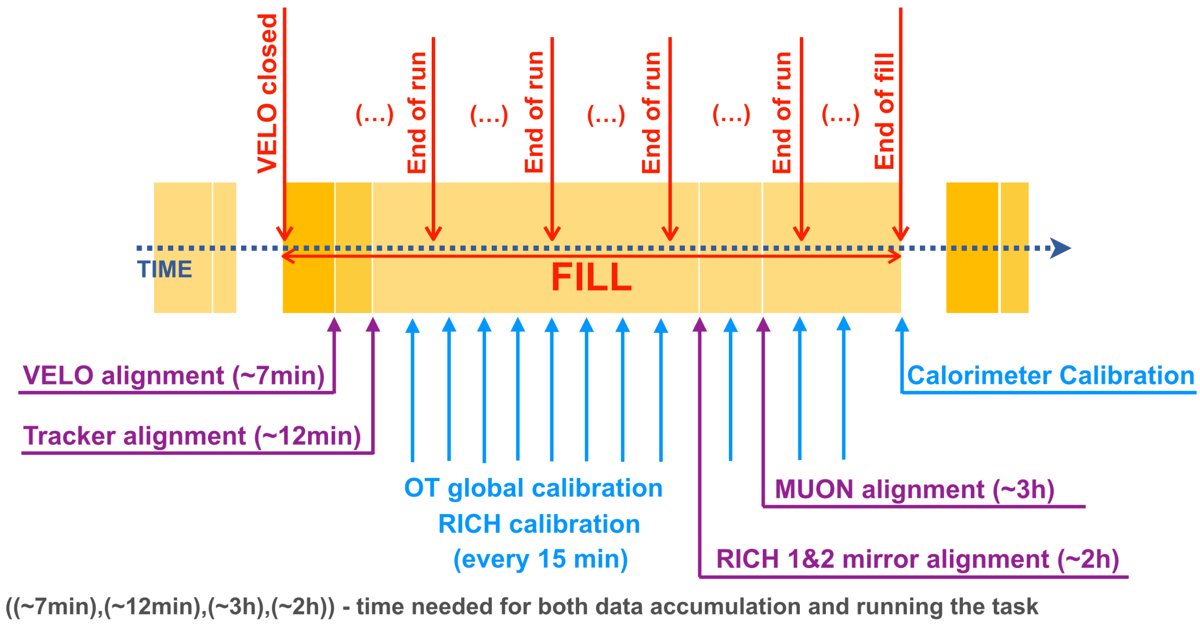

Figure 3.22: Real-time alignment and calibration procedure used in a fill for 2018 data taking.

format.

The resulting dataset is still too large to be usable by individual analysts. Instead the data is filtered further in the DaVinci application. Firstly DaVinci analyses the *ProtoParticles* from the reconstruction to locate primary and secondary vertices and identify composite particle candidates. Then the data is filtered by loose *stripping line* exclusive selections. Stripping lines are grouped into streams, and the events selected by all stripping lines in a stream are written to the same file in DST format. To minimise duplication of events between streams and therefore reduce file sizes, stripping lines with common physics, for example containing a beauty hadron decay, are in the same stream.

Finally, analysts can use DaVinci to process the data in a stream and produce ntuples, which are useful for physics analysis. Only the events that were selected by a specified set of stripping lines are retained. Algorithms can be applied to evaluate and save higher-level quantities for signal decay candidates, which are useful in physics analysis or further selection. One such algorithm is a *Decay Tree Fit* (DTF) [158], which improves the momentum and therefore mass resolution by incorporating knowledge of the decay chain. Constraints that can be used include fixing daughter particle masses to their known values and requiring that the momenta of all tracks point towards the corresponding production vertices. A Kalman filter applies these constraints and extracts more precise estimates of the vertex positions, particle momenta, and decay times.

Events saved in the Turbo stream by the HLT2 trigger (Section 3.2.3) do not go through the reconstruction and stripping stages, and can be accessed directly when making ntuples.

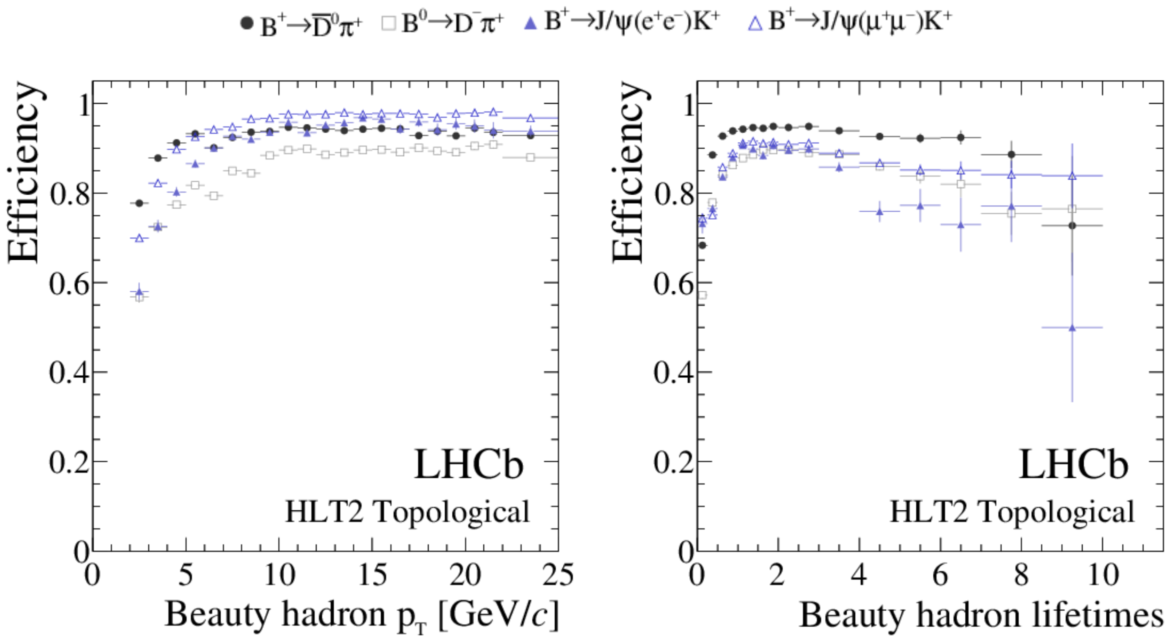

Figure 3.23: TOS efficiency of the HLT2 topological trigger lines in Run 2 as a function of (left) p_{T} and (right) rest frame lifetime for selected beauty meson decays [148].

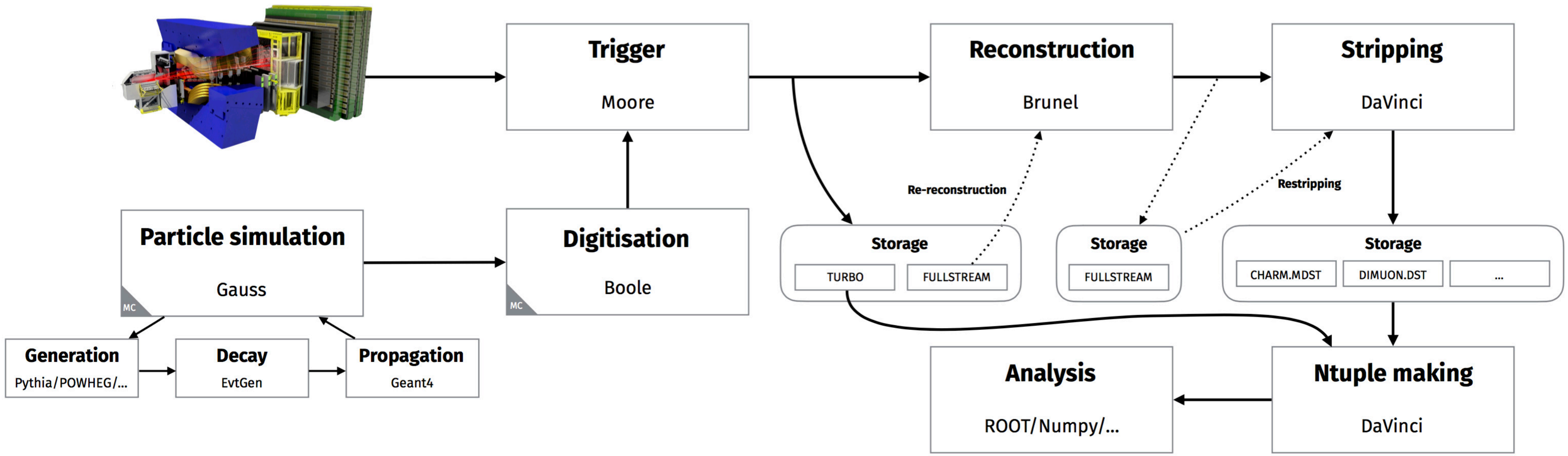

Figure 3.24: Stages of processing and filtering for data from the detector and simulation in Run 2. The software application used at each stage is also specified [159].

3.3 Upgrade I

To accelerate the growth of the LHCb dataset, the instantaneous luminosity has been increased by a factor of five for Run 3 data taking, as discussed in Section 3.1. The LHCb detector itself has been substantially upgraded to cope with the resulting higher pile-up. To maintain sufficiently low occupancies under increased event multiplicity, the granularity of many subdetectors has been increased. Furthermore, the data processing strategy has been largely redesigned. To maintain an acceptable event rate, it would have been necessary to increase the L0 trigger thresholds, which would saturate the triggered yield of interesting decays. Therefore the hardware trigger is replaced by a full-software trigger, which relies on the selection of good-quality displaced tracks at the first level of filtering (HLT1). This increased demand on the HLT1 trigger required further upgrades to the tracking detectors to enable improved track reconstruction and impact parameter measurement. The read-out electronics were also replaced to accommodate readout at the bunch crossing frequency of 40 MHz. These improvements, implemented between the end of Run 2 and the start of Run 3, are collectively referred to as *Upgrade I* [99].

3.3.1 Hardware upgrades

Owing to their importance in the new full-software trigger, all tracking detectors have been entirely replaced to achieve improved track reconstruction speed and precision. The VELO now uses hybrid pixel sensors instead of microstrips; the inner radius has been reduced to 5.1 mm; and the inner foil radius has been reduced to 3.5 mm [114]. These changes improve resolution, and the uniformity enables faster track reconstruction. A new higher granularity, more radiation hard, silicon microstrip Upstream Tracker (UT) replaced the TT. A Scintillating Fibre tracker (SciFi) replaced the T stations. 2.5 m long fibres read out by silicon photo-multipliers offer high efficiency, resolution and radiation hardness with a low material budget [160].

The SPD, PS and M1 have been removed owing to their reduced importance with the full-software trigger. Otherwise, the conceptual design of the PID detectors is unchanged. All read-out electronics have been upgraded to accommodate 40 MHz triggerless readout and the RICH HPDs have been replaced with multianode photomultiplier tubes. The RICH1 optical layout has been redesigned to reduce the peak occupancy [99].

3.3.2 Data processing upgrades

With no hardware trigger, the first stage of the software trigger is required to be an inclusive trigger able to reduce the data rate from 30 MHz to 1 MHz and process each event in 13 ms [131, 161]. Conceptually the upgraded HLT1 trigger is similar to the original HLT1 trigger, and it performs a similar partial reconstruction using only the tracking detectors. However, improvements to the reconstruction algorithms and tracking detectors, and implementation of the trigger on 200 graphics processing unit (GPU) cards [99] allow the substantial increase in rate despite the higher multiplicity environment.

The 7% of events with the highest UT+SciFi occupancy are rejected, playing a similar role to the requirement on the number of hits in the SPD implemented in the previous hardware trigger [162]. Tracks are then reconstructed using information from the VELO, UT, and SciFi. Finally, events are filtered according to a set of trigger lines, which are similar to those used in Runs 1 and 2. Trigger lines selecting a single displaced track or displaced two-track vertex with high p_{T} inclusively select most beauty and charm hadron decays. There are also lines to select decays involving muons; decays useful for calibration; and lines which occupy a smaller rate designed specifically for hard-to-select signals.

The output data is written to a disk buffer where real-time alignment and calibration is performed to accommodate offline-quality reconstruction, similarly to in Run 2. The data is then filtered by the HLT2 trigger implemented on CPUs [99]. The HLT2 trigger firstly performs a full event reconstruction including reconstructing charged particle tracks and calorimeter clusters, and deriving PID variables. A set of trigger lines determine which events to save based on rectangular and multivariate cuts, and which subset of the event to save. The majority of lines are inclusive topogical lines or select exclusive decays.

The major difference with respect to Run 2 is that, owing to the success of the Turbo model in Run 2 and the need to reduce the output bandwidth of HLT2 in view of higher luminosity and trigger efficiencies, the TurboSP model has become the default for Run 3. 68% of events are expected to be written through the TurboSP model, occupying 25% of bandwidth. In particular, most beauty hadron decays will now also use the TurboSP model. This will additionally avoid the inefficiency of the topological triggers. The full event is still written for the output of inclusive topological triggers and for physics cases that are particularly sensitive to misalignment and therefore may benefit from offline recalibration, such as electroweak physics analyses. The output of HLT2 is written to tape at 10 GB/s.

The stripping of Run 1-2 has been replaced by a *sprucing* stage [99] that re-reconstructs the event and applies exclusive selections for data saved in the Full stream such that the bandwidth is reduced sufficiently to write the data to disk. Since this stage is performed offline, the timing constraints on reconstruction are less strict than in HLT2. Sprucing will run concurrently with data taking and in an annual re-sprucing.

The full data flow in Run 3 is shown in Figure 3.25.

3.4 Simulation

Simulating the differential production distribution and the detector response for signal and background decays using Monte Carlo (MC) methods is critical for analysis of LHCb data.

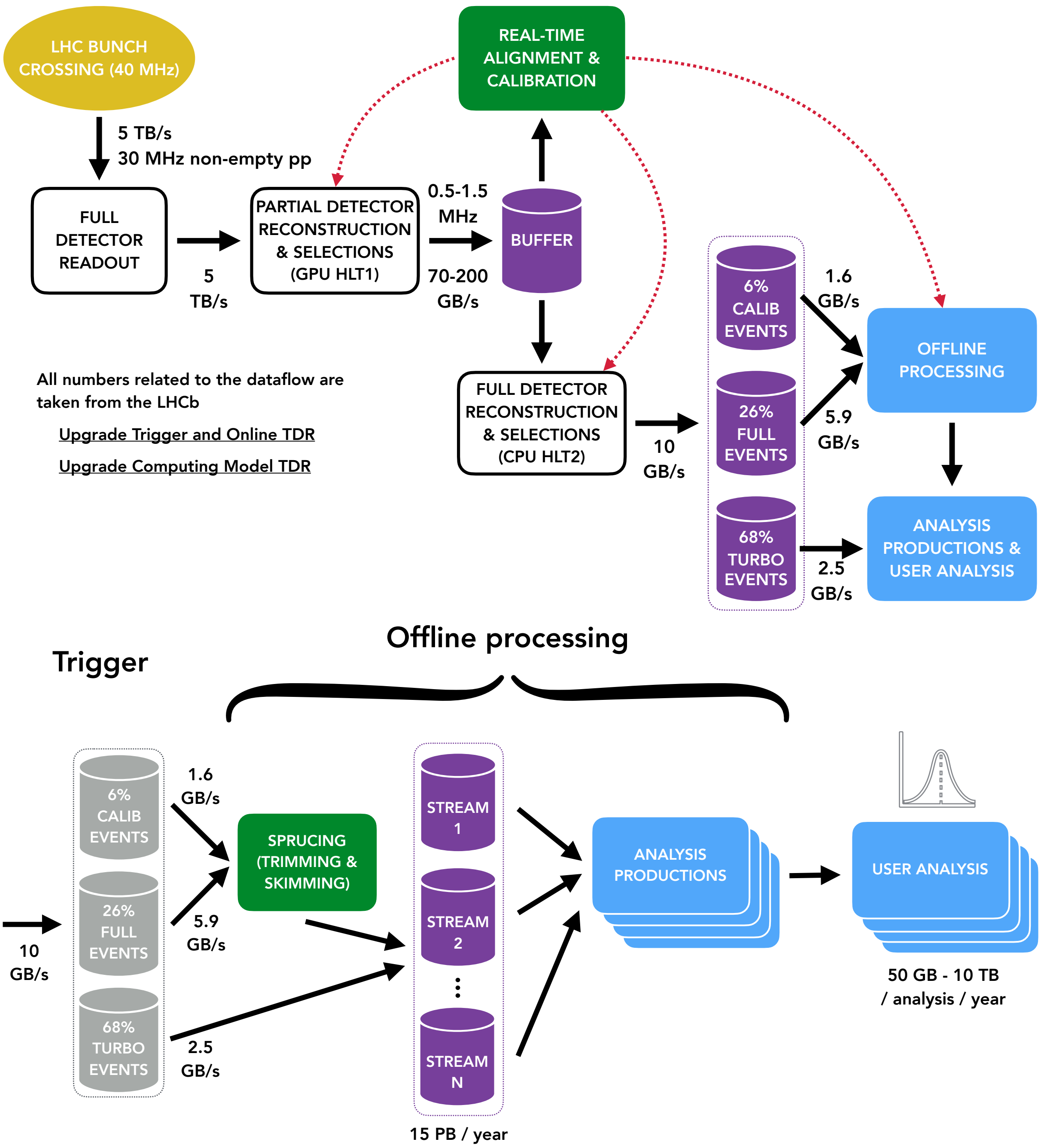

Figure 3.25: Dataflow following Upgrade 1, with detail given for (top) the real-time online processing and (bottom) the offline processing [163].

The GAUSS application [164] controls this simulation by calling specialised applications.

A generator application simulates the pp interaction and subsequent hadronisation. The general-purpose generator PYTHIA [165] is used for most LHCb simulation. The specialised BCVEGPY [166] generator is typically used to simulate B_c^+ meson production as these are rarely produced by PYTHIA. EVTGEN [167] is then called to simulate the signal particle decay. PHOTOS [168] generates final state radiation. The distribution of detector hits as a result of the propagation through and interaction with the detector of the final state particles is simulated by GEANT4 [169, 170].

To ensure that the trigger and offline data processing will be modelled correctly, the simulated hits are converted into the same format as the read-out of real events. This digitisation is performed by BOOLE, which simulates the detector response, read-out electronics, and hardware trigger [157].

Simulated data also contains *truth information*, which records the physics history of the generated event and relates the hits to particles [157]. This information can be useful when using the simulation in physics analysis as the reconstructed candidate may not be the true signal decay in the event.

Simulation is labelled according to the version of Gauss used in the generation. The label takes the form SimXXy, where XX is a two-digit number and y is a letter. The major version denoted by XX reflects updates updates in the decay generation software and in GEANT4, and major detector description updates. The minor version denoted by y is incremented upon smaller changes.

Analysis overview

4.1 Search for doubly charmed B_c^+ meson decays

This thesis firstly presents a search for sixteen decays of the B_c^+ meson to two open charm mesons with the data collected in pp collisions by the LHCb detector between 2011 and 2018 [1]. A previous search using only data collected in 2011 and 2012 found no evidence of these decays [171].

These sixteen decays are all possible combinations of a $D^{(*)+}$ or $D_s^{(*)+}$ meson with a $D^{(*)0}$ or $\overline{D}^{(*)0}$ meson. D_s^{*+} and D^{*0} mesons decay to a D_s^+ and D^0 meson, respectively, primarily by emitting a soft $X^0 = \{\pi^0, \gamma\}$, which are very difficult to reconstruct in the LHCb detector due to the finite energy and spatial resolution of the ECAL and the large background from the underlying event [172]. Therefore decays involving these mesons are *partially reconstructed* in the analysis, meaning that only the charged final state tracks are reconstructed. The missing momentum results in a reconstructed mass distribution at lower mass than m_B and with a standard deviation of typically $100\,\text{MeV}/c^2$. The wider mass distribution decreases the statistical sensitivity to these decays in the presence of backgrounds. The D^{*+} meson may also decay emitting a neutral particle, but it primarily decays to $D^0\pi^+$ with a branching fraction of $(66.7\pm0.5)\%$ [22]. Therefore searches for decays involving a D^{*+} are performed by fully reconstructing the decay to $D^0\pi^+$.

To cancel systematic uncertainties associated with luminosity and absolute signal efficiencies, the branching fractions are measured using high yield $B^+ \to D_{(s)}^{(*)+}\overline{D}^0$ decays as normalisation modes. Note that $B^+ \to D_{(s)}^{(*)+}D^0$ decays are forbidden at tree level. The experimentally accessible observable proportional to the branching fraction of fully reconstructed $B_c^+ \to D_{(s)}^{(*)+}\overset{(-)}{D}{}^0$ decays is

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_{(s)}^{(*)+}\overset{(-)}{D}{}^0)}{\mathcal{B}(B^+ \to D_{(s)}^{(*)+}\overline{D}{}^0)} = \frac{N(B_c^+)}{N(B^+)}\frac{\varepsilon(B^+)}{\varepsilon(B_c^+)}, \tag{4.1}$$

and similar expressions are applicable for partially reconstructed decays. N are the yields in data, extracted from fits of a model to the invariant mass distribution of $B_{(c)}^+$ candidates. ε are the efficiencies accounting for acceptance, reconstruction, and selection, evaluated with samples of simulated signal decays. The value of f_c/f_u has been measured in the LHCb acceptance for $p_{\mathrm{T}}(B) > 4\,\text{GeV}/c$ [106]. The number of generated events in the kinematic range of interest $p_{\mathrm{T}}(B) > 4\,\mathrm{GeV}/c, 2.0 < y < 4.5$ ($N_{\mathrm{gen,kin}}$) is the denominator for ε,

$$\varepsilon = \frac{N_{\mathrm{sel}}}{N_{\mathrm{gen,kin}}}, \tag{4.2}$$

to ensure that the efficiency of this requirement cancels in the branching fraction measurement.

4.2 Analysis of doubly charmed B^- meson decays

A measurement of *CP* asymmetries , defined as

$$\mathcal{A}^{CP}(B^- \to D_{(s)}^{(*)-} D^{(*)0}) \equiv \frac{\mathcal{B}(B^- \to D_{(s)}^{(*)-} D^{(*)0}) - \mathcal{B}(B^+ \to D_{(s)}^{(*)+} \overline{D}^{(*)0})}{\mathcal{B}(B^- \to D_{(s)}^{(*)-} D^{(*)0}) + \mathcal{B}(B^+ \to D_{(s)}^{(*)+} \overline{D}^{(*)0})}, \tag{4.3}$$

and branching fractions in $B^- \to DD$ decays is performed, again using the data collected in pp collisions by the LHCb detector between 2011 and 2018. Similarly to the search for $B_c^+ \to DD$ decays, both fully and partially reconstructed decays are analysed.

The $\mathcal{A}^{CP}$ are measured for fully reconstructed decays and decays with one missing particle. The $\mathcal{A}^{CP}$ to be measured, the channel in which they will be measured, and their current experimental status are shown in Table 4.1. $\mathcal{A}^{CP}(B^- \to D_s^- D^0)$ and $\mathcal{A}^{CP}(B^- \to D^- D^0)$ were measured with the Run 1 LHCb dataset [173] and were found to be consistent with both *CP* symmetry and SM expectations. The $B^- \to D^{*-} D^0$ decay is accessible in both the $D^+ \overline{D}^0$ and $D^{*+} \overline{D}^0$ channels, but better precision is possible in the $D^{*+} \overline{D}^0$ channel where it can be fully reconstructed with a higher branching fraction. The stripping line requirement $m(B) > 4800\,\mathrm{MeV}/c^2$ prevents the analysis of decays with two missing particles.

When all asymmetries are smaller than 5%, the *CP* asymmetry is approximated to very

Table 4.1: The *CP* asymmetries measured in this analysis, the channels in which they are measured, and the current status of their measurement [22].

Channel	Decay	$\mathcal{A}^{CP}$ world average / %	Source
$D_s^+ \overline{D}^0$	$B^- \to D_s^- D^0$	-0.4 ± 0.7	[173]
	$B^- \to D_s^{*-} D^0$	-	-
	$B^- \to D_s^- D^{*0}$	-	-
$D^+ \overline{D}^0$	$B^- \to D^- D^0$	1.6 ± 2.5	[173–175]
	$B^- \to D^- D^{*0}$	13 ± 18	[175]
$D^{*+} \overline{D}^0$	$B^- \to D^{*-} D^0$	$-\ 6 \pm 13$	[175]
	$B^- \to D^{*-} D^{*0}$	-15 ± 11	[175]

Table 4.2: Current world averages on the branching fractions of $B^- \to D^- D^0$ decays [22].

Quantity	World average
$\mathcal{B}(B^- \to D_s^- D^0)$	$(9.0 \pm 0.9) \times 10^{-3}$
$\mathcal{B}(B^- \to D^- D^0)$	$(3.8 \pm 0.4) \times 10^{-4}$
$\mathcal{B}(B^- \to D^{*-} D^0)$	$(3.9 \pm 0.5) \times 10^{-4}$

good precision by subtracting the *production asymmetry* of the B^- meson,

$$\mathcal{A}^{\mathrm{P}} \equiv \frac{\sigma(B^-) - \sigma(B^+)}{\sigma(B^-) + \sigma(B^+)}, \tag{4.4}$$

and the *detection asymmetry* of the final state particles,

$$\mathcal{A}^{\mathrm{D}} \equiv \frac{\varepsilon(B^-) - \varepsilon(B^+)}{\varepsilon(B^-) + \varepsilon(B^+)}, \tag{4.5}$$

from the *raw asymmetry* measured in the data,

$$\mathcal{A}^{CP} = \mathcal{A}^{\mathrm{raw}} - \mathcal{A}^{\mathrm{P}} - \mathcal{A}^{\mathrm{D}}, \tag{4.6}$$

$$\mathcal{A}^{\mathrm{raw}} \equiv \frac{N(B^-) - N(B^+)}{N(B^-) + N(B^+)}. \tag{4.7}$$

Production and detection asymmetries are evaluated using measurements from calibration data samples that are corrected for the kinematics of the signal decays. Additionally, *CP*-averaged branching fractions of the fully reconstructed decays are also measured. Only ratios of branching fractions can be measured with high precision due to large uncertainties on luminosities, the B^+ production cross-section, and absolute efficiencies. This thesis describes a measurement of the ratio between the $B^- \to D^- D^0$ and $B^- \to D_s^- D^0$ branching fractions, and the $B^- \to D^{*-} D^0$ and $B^- \to D^- D^0$ branching fractions.

$$R(D^- D^0 / D_s^- D^0) = \frac{\mathcal{B}(B^- \to D^- D^0)}{\mathcal{B}(B^- \to D_s^- D^0)} \frac{\mathcal{B}(D^- \to \pi^- \pi^- K^+)}{\mathcal{B}(D_s^- \to K^+ K^- \pi^-)}, \tag{4.8}$$

and

$$R(D^{*-} D^0 / D^- D^0) = \frac{\mathcal{B}(B^- \to D^{*-} D^0)}{\mathcal{B}(B^- \to D^- D^0)} \frac{\mathcal{B}(D^{*-} \to \overline{D}^0 \pi^-) \mathcal{B}(\overline{D}^0 \to K^+ \pi^-)}{\mathcal{B}(D^- \to \pi^- \pi^- K^+)}. \tag{4.9}$$

These branching fractions have been measured as absolute branching fractions, most recently by Belle [174, 176] and BaBar [175, 177], but with a fractional uncertainty of approximately 10%. LHCb has also measured the branching fraction of the $B^- \to D_s^- D^0$ decay relative to $B^0 \to D_s^- D^+$ [178]. The current world averages on the absolute branching fractions are shown in Table 4.2. A higher precision measurement of the ratios of $B^- \to DD$ branching fractions will be useful in the combined analysis of $B \to DD$ decays.

Branching fraction ratios are measured as the efficiency-corrected yield ratio,

$$\frac{\mathcal{B}(B^- \to D^- D^0)}{\mathcal{B}(B^- \to D_s^- D^0)} = \frac{N(B^- \to D^- D^0)}{N(B^- \to D_s^- D^0)} \frac{\varepsilon(B^- \to D_s^- D^0)}{\varepsilon(B^- \to D^- D^0)}, \tag{4.10}$$

where N is the yield in selected data and ε is the acceptance, reconstruction and selection efficiency, which is evaluated using simulated signal samples and corrected for mismodelling in simulation using samples of calibration data.

Data samples and selection

This chapter describes the data and simulation samples for the analyses discussed in this thesis. The properties of the dataset are outlined in Section 5.1 and the choice of the decay modes in which the charm mesons are reconstructed is explained. Sections 5.2-5.4 describe the simulation samples and the correction of mismodelling in the simulation. Sections 5.5-5.8 explain how backgrounds are reduced in the data. Finally, Section 5.9 discusses the treatment of events producing multiple candidates in the selected data.

5.1 Data samples

The analyses described in this thesis use the data collected by the LHCb detector in pp collisions during Run 1 and Run 2 of the LHC. This corresponds to an integrated luminosity of $1.0\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$ in 2011; $2.0\,\text{fb}^{-1}$ at $\sqrt{s} = 8\,\text{TeV}$ in 2012; and $6.0\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$ in 2015-2018.

D^0, D^+ and D_s^+ mesons are reconstructed in specific decay modes, and the decay products are referred to as *final states*. The decays with the highest branching fraction occur at tree level and are either fully hadronic (for example $D^0 \to K^-\pi^+$) or semileptonic (for example $D^0 \to K^-\ell^+\nu$). The LHCb detector is optimised for the reconstruction of charged kaons and pions. Therefore, in this thesis, charm mesons are reconstructed in their CF charged hadronic decay modes: $D^0 \to K^-\pi^+(\pi^-\pi^+)$, $D^+ \to K^-\pi^+\pi^+$, and $D_s^+ \to K^+K^-\pi^+$. The four body D^0 decay has a branching fraction double that of the two body decay, primarily due to the larger available phase space. However, the available momentum is split between more particles, so the reconstruction and selection efficiency of each final state particle is lower. Overall, $1.5-2$ times fewer $D^0 \to K^-\pi^+\pi^-\pi^+$ decays than $D^0 \to K^-\pi^+$ decays are observed. Therefore, $B^+_{(c)} \to D^{*+}\overset{(-)}{D}{}^{(*)0}$ decays where both D^0 mesons decay to $K^-\pi^+\pi^-\pi^+$ are not included in the following analyses. The D^{*+} meson is reconstructed through its decay into a D^0 and a soft π^+.

Neutrinos cannot be observed by the LHCb detector, and therefore the resolution on any mass observable for a semileptonic final state is very poor compared to a charged hadronic final state. With poor mass resolution, separation between the low-yield $B_c^+ \to DD$ decays and background, or percent-precision measurements of $B^- \to DD$ decays, would not be possible. Therefore semileptonic D decay modes have not been included in these analyses, despite the potential yield gains.

The possible gain in sample size through reconstructing final states with a K_S^0 or π^0

Table 5.1: Branching fraction of selected D^0, D^+ and D_s^+ meson decays [22].

Decay mode		$\mathcal{B}/\%$
$D^0 \to$	$K^-\pi^+$	3.947 ± 0.030
	$K^-\pi^+\pi^0$	14.4 ± 0.5
	$K^-\pi^+\pi^0\pi^0$	8.86 ± 0.23
	$K^-\pi^+\pi^-\pi^-$	8.22 ± 0.14
	$K^-\pi^+\pi^-\pi^-\pi^0$	4.3 ± 0.4
	$K_S^0\pi^+\pi^-$	2.80 ± 0.18
	$K_S^0\pi^+\pi^-\pi^0$	5.2 ± 0.6
$D^+ \to$	$K^-\pi^+\pi^+$	9.38 ± 0.16
	$K^-\pi^+\pi^+\pi^0$	6.25 ± 0.18
	$K^+K^-\pi^+$	0.968 ± 0.018
	$K^+K^-\pi^+\pi^0$	0.662 ± 0.032
	$K_S^0\pi^+\pi^0$	7.36 ± 0.21
	$\pi^+\pi^+\pi^-$	0.327 ± 0.018
	$\pi^+\pi^+\pi^-\pi^0$	1.16 ± 0.08
$D_s^+ \to$	$K^+K^-\pi^+$	5.38 ± 0.10
	$K^+K^-\pi^+\pi^0$	5.50 ± 0.24
	$\pi^+\pi^+\pi^-$	1.08 ± 0.04
	$K^+K^+K^-$	0.0215 ± 0.0020

can also be considered. Table 5.1 shows that branching fractions to these final states can be substantial: for example, $\mathcal{B}(D^0 \to K^-\pi^+\pi^0) = (14.4 \pm 0.5)\%$ compared to $\mathcal{B}(D^0 \to K^-\pi^+) = (3.946 \pm 0.030)\%$. However, it is difficult to reconstruct a π^0 or a K_S^0 in the LHCb detector. A π^0 decays into a pair of photons, which are reconstructed in the ECAL. At p_T below 500 MeV/c, the π^0 identification efficiency is low, and the background from neutral pions created in fragmentation and from random combinations of photons is high. A K_S^0 is reconstructed in its decay to $\pi^+\pi^-$, which has a branching fraction of $(69.20 \pm 0.05)\%$ [22]. It has a long lifetime of (89.564 ± 0.033) ps [22] so two thirds of K_S^0 mesons are reconstructed in LHCb as a pair of downstream tracks. Only one third are reconstructed as a pair of long tracks, which have better momentum, vertex resolution, and trigger efficiency [179].

The decay channels $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K_S^0\pi^+\pi^-$ are used to investigate whether improvements to signal sensitivity could be achieved by using decay modes with neutral pions or kaons in spite of the experimental difficulties. $B^+ \to D^+\overline{D}^0$ candidates with $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$, and $D^0 \to K_S^0\pi^+\pi^-$ in the 2018 dataset are obtained from the output of the `StrippingBc2DD0D2HHHD02XBeauty2CharmLine` stripping lines where `X={HH, KHHH, HHPI0, KSHH}`. The `D02HH` and `D02KHHH` lines are described in Section 5.6, and the other lines are identical except for changes to the D^0 selection. The requirements applied in these stripping selections to the D^0 candidates are standard requirements used in many LHCb stripping lines. The selection described in

Table 5.2: $B^+ \to D^+\overline{D}^0$ candidate yields in 2018 data for several D^0 final states obtained using the method described in the text.

D^0 final state	Yield
$K^-\pi^+$	2099 ± 102
$K^-\pi^+\pi^-\pi^+$	1428 ± 108
$K^-\pi^+\pi^0$	192 ± 47
$K^0_S\pi^+\pi^-$	126 ± 72

Section 5.7 is also applied, as well as $\pm 20\,\mathrm{MeV}/c^2$ K^0_S and π^0 mass windows, which have a high signal efficiency.

The $m(D^+\overline{D}^0)$ distributions for the selected data are shown in Figure 5.1. The signal yields in Table 5.2 are extracted by subtracting the sideband yields $30 < |m - m_{B^+}| < 60\,\mathrm{MeV}/c^2$ from the signal window yields $|m - m_{B^+}| < 30\,\mathrm{MeV}/c^2$.

Both the $K^-\pi^+\pi^0$ and $K^0_S\pi^+\pi^-$ final states have less than 10% of the yield of the $K^-\pi^+$ final state. These yields and the corresponding branching fractions imply efficiency ratios

$$\varepsilon(D^0 \to K^-\pi^+\pi^0)/\varepsilon(D^0 \to K^-\pi^+) = 0.025 \pm 0.006, \tag{5.1}$$

$$\varepsilon(D^0 \to K^0_S\pi^+\pi^-)/\varepsilon(D^0 \to K^-\pi^+\pi^-\pi^+) = 0.37 \pm 0.22. \tag{5.2}$$

The low efficiency for the $K^-\pi^+\pi^0$ final state results from the $p_{\mathrm{T}}(\pi^0) > 500\,\mathrm{MeV}/c$ requirement applied in the stripping selection to remove the large low-energy π^0 and γ background in the calorimeters, where the π^0 reconstruction efficiency is anyway low. Although the reconstruction efficiency is not as low for the $K^0_S\pi^+\pi^-$, it is still lower than for a four-track final state not involving an intermediate long-lived K^0_S, and there are no final states involving the K^0_S with a branching fraction high enough to compensate for the reduced efficiency. Overall, no useful gain in signal yield can be achieved by including final states with neutral particles.

5.2 Simulation samples

Samples of simulated signal and background decays are essential for the training of a multivariate classifier to improve background rejection, the evaluation of efficiencies, and knowledge of invariant mass and kinematic distributions. Simulation samples were produced as described in Section 3.4 using PYTHIA and BCVEGPY to generate B^+ and B_c^+ mesons, respectively. EVTGEN simulated the decay chains. Two body decays were generated according to a phase space model including particle spin. Three body decays followed the resonance structure measured in the Dalitz plane. Four body decays were generated according to either a phase space model (Sim08) or an incoherent sum of intermediate states (Sim09). Final state radiation was implemented by PHOTOS, and

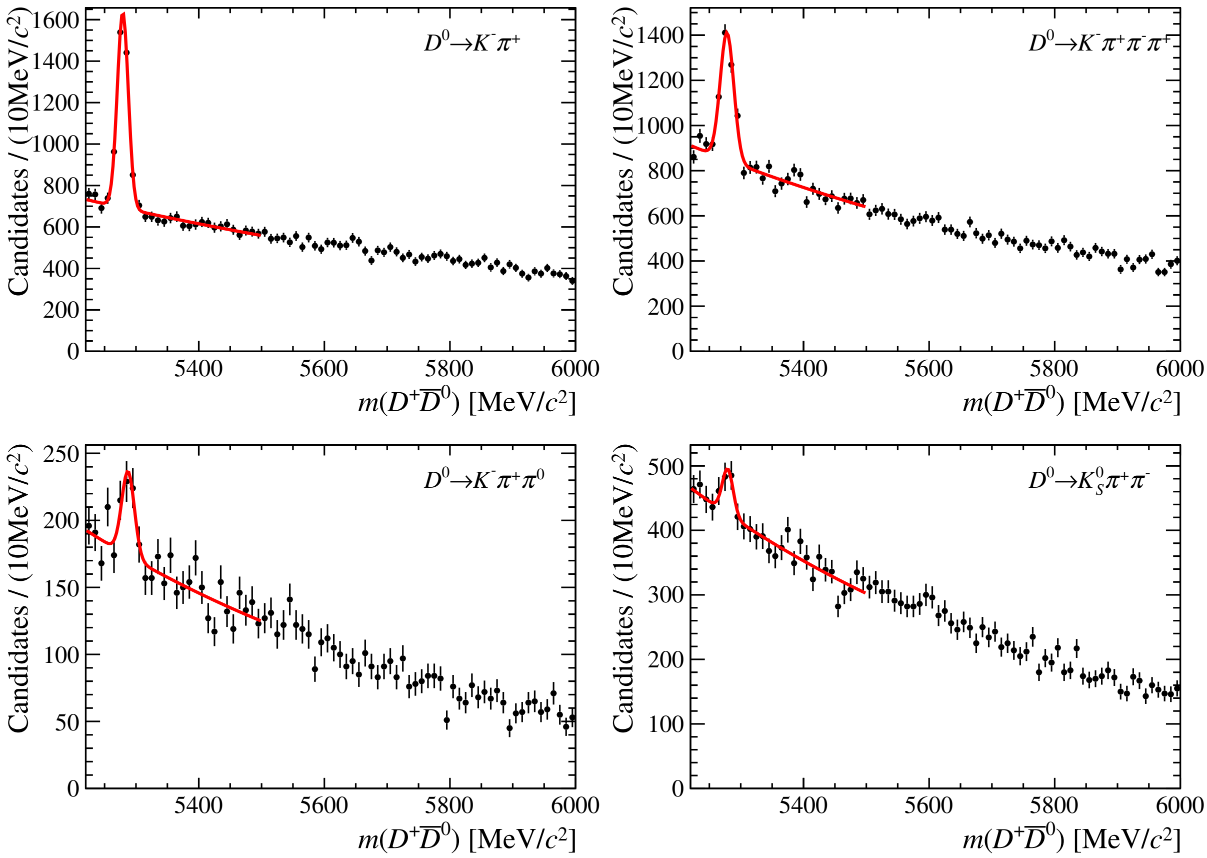

Figure 5.1: $B^+ \to D^+\overline{D}^0$ candidates selected by relevant stripping lines in 2018 data. The D^0 is reconstructed as (top left) $K^-\pi^+$, (top right) $K^-\pi^+\pi^-\pi^+$, (bottom left) $K^-\pi^+\pi^0$, and (bottom right) $K^0_S\pi^+\pi^-$. The preselection described in Section 5.7 is applied, in addition to $\pm 20\,\mathrm{MeV}/c^2$ K^0_S and π^0 mass windows. The fit of a Gaussian plus an exponential is drawn (red) for illustration.

GEANT4 simulated the detector response.

This analysis uses simulated samples of each RS fully reconstructed $B^+_{(c)} \to D^{(*)+}_{(s)}\overline{D}^0$ signal decay and partially reconstructed $B^+_c \to D^{(*)+}_{(s)}\overline{D}^{(*)0}$ decay. WS decays have identical kinematics and therefore need not be simulated. For partially reconstructed decays, a simulation sample was produced for both $D^* \to D\pi^0$ and $D^* \to D\gamma$ decays, and for both polarisations of $B \to D^*D^*$ decays. Simulation was generated under 2012, 2015, and 2016 conditions, with a 1:4 ratio between the generated yield in 2015 and 2016. Additionally, a sample of $B^+ \to \overline{D}^0K^+K^-\pi^+$ decays with $\overline{D}^0 \to K^+\pi^-$ was generated under 2012 conditions to help model this peaking background.

$B^+_{(c)} \to D^+_{(s)}\overline{D}^0$ and $B^+ \to \overline{D}^0K^+K^-\pi^+$ decays under 2012 conditions were generated with Sim08; all other simulation was generated with Sim09. LHCb simulation produced centrally is often *stripping-filtered* to reduce file sizes, meaning that only events selected by a specified stripping line are saved. Stripping decisions are evaluated but the corresponding selection is not applied for *stripping-flagged* simulation samples. The Sim08 samples described above are stripping-flagged and the Sim09 samples are stripping-filtered. Stripping-flagged samples are required to evaluate the efficiency of the

PID requirements applied in the stripping selection, which is important for high precision branching fraction measurements. For this purpose, stripping-flagged samples of $B^+ \to D^+_{(s)}\overline{D}^0$ in 2017 conditions and $B^+ \to D^{*+}\overline{D}^0$ in 2015 conditions with $\overline{D}^0 \to K^+\pi^-$ are also used in the measurement of $B^- \to DD$ branching fractions.

Simulation of the detector response is time consuming, and in certain cases the inclusion of detector reponse does not significantly modify the important distributions. *Generator-level simulation* samples are produced for these cases, which include only the production and decay processes, allowing the rapid production of simulation for a large range of decays. Generator-level simulation is used to model low-mass backgrounds in the analysis of $B^- \to DD$ decays from $B \to D^*D^*$ decays with two missing particles or $B \to D^{**}D$ decays with at least one missing particle. The invariant mass distributions of these decays are already much broader ($\sigma > 100\,\text{MeV}/c^2$) than the B mass resolution ($\sigma \sim 22\,\text{MeV}/c^2$ [180]) due to the unreconstructed particles and the non-negligible width of D^{**} states, so an accurate description of the detector response is not necessary. To improve correspondence between distributions in the generator-level simulation and data, the B meson and all final state particles are required to be within the LHCb acceptance ($2.0 < y(B) < 4.5$, $10 < \theta < 400\,\text{mrad}$ for all final state particles), and the $p_{(\mathrm{T})}$ requirements in the stripping (Section 5.6) and preselection (Section 5.7) are applied. 350 thousand events were generated under 2016 conditions, of which approximately 10% pass the applied cuts, which do not significantly modify the invariant mass distribution in practice. To emulate detector resolution, the invariant mass of the B meson is smeared by a double Gaussian whose shape parameters are obtained from a fit to full simulation of fully reconstructed decays to the same final state.

5.3 Corrections to simulation

Distributions are often mismodelled in simulation. This can bias the signal efficiencies and charge asymmetries, which depend strongly on the kinematics and PID variables, for example. It is necessary to correct simulation for any mismodelling that may bias these quantities. This section will explain these corrections.

5.3.1 Momentum scale and resolution

Mismodelling of momentum resolution in simulation is corrected by smearing momenta using a standard tool at the time of ntuple creation. Additionally, to improve agreement in D mass distributions between simulation and data, the momenta of all particles are scaled downwards by 0.038% in Run 1 and 0.052% in Run 2 simulation [181], and invariant masses are updated accordingly.

5.3.2 Particle identification variables

Due to the complexity of the RICH detectors, $\mathrm{DLL}_{K\pi}$ variables also tend to be mismodelled in simulation. These variables are corrected with a function that transforms $\mathrm{DLL}_{K\pi}$ values as a function of track p_T and η. The transformation is obtained using calibration samples and implemented in the PIDCorr package [182]. The corrected $\mathrm{DLL}_{K\pi}$ is the value in the calibration sample that results in the same efficiency as the original value in simulation, at a given value of track p_T and η. This requires the distribution of the calibration and simulation samples to be parameterised as a function of $(\mathrm{DLL}_{K\pi}, p_\mathrm{T}, \eta)$ using a kernel density estimate. This method preserves but does not correct dependence of PID variables on other parameters. In particular, it preserves correlations in PID variables between tracks in the same candidate.

5.3.3 Gradient-boosted reweighting

Simulation is corrected such that several distributions match those in data (Sections 5.3.5-5.3.8) with a *Gradient Boosted Reweighter* (GBR) [183] method. A GBR is a series of *boosted decision trees* (BDTs) [184] that evaluates a weight for each candidate according to its location in the phase space such that the weighted distribution (the corrected simulation) in the phase space matches a target distribution (the background-subtracted data). In general, a decision tree (DT) splits the dataset in one variable in order to optimise some criterion, creating two *nodes*. Each node is then split in the same manner. This process is repeated until another criterion is met, and the final nodes are referred to as *leaves*. The number of times the splitting is repeated is the *depth* of the DT. In the context of a GBR, the split is chosen to maximise

$$\chi^2 = \frac{(W_{\mathrm{MC},1} - W_{\mathrm{data},1})^2}{W_{\mathrm{MC},1} + W_{\mathrm{data},1}} + \frac{(W_{\mathrm{MC},2} - W_{\mathrm{data},2})^2}{W_{\mathrm{MC},2} + W_{\mathrm{data},2}}, \tag{5.3}$$

where $W_{\mathrm{MC,i}}$ and $W_{\mathrm{data,i}}$ are the total weights of candidates on node i in simulation and data, respectively, normalised so that $W_{\mathrm{MC},1} + W_{\mathrm{MC},2} = W_{\mathrm{data},1} + W_{\mathrm{data},2}$. In this way, the DT finds the division which maximise the differences between simulation and data on each leaf.

The weight of each simulated candidate is then updated so that the total weight of the simulated candidates on each leaf is equal to the total weight in data:

$$w_\mathrm{MC} \rightarrow w_\mathrm{MC} \times e^y, \tag{5.4}$$

$$y = \ln \frac{W_{\mathrm{data,i}}}{W_{\mathrm{MC,i}}}. \tag{5.5}$$

Using this method, a sufficiently deep DT could correct the simulated distribution. However, the resulting weight function would be unstable with respect to statistical fluctuations. A boosted series of DTs produces a much more stable weight function. In the context of a GBR, *boosting* refers to the use of the weights updated by one DT in the creation of the next DT.

The variable y may be scaled by a *learning rate* between 0 and 1. A lower learning rate mitigates the risk of *overtraining*, where the reweighter learns to follow statistical fluctuations in either dataset, but more trees are needed for the reweighting rule to converge. An overly high number of trees or tree depth also results in overtraining. Since a small number of candidates on a leaf produces large and unphysical weights, the minimum number of candidates on each leaf can be specified. Finally, a fraction of the data can be selected at random to train each DT, which further reduces overtraining.

The hyperparameters for training this reweighter are 200 trees; a learning rate of 0.1; a maximum depth of 2; a minimum of 1000 candidates on each leaf; and the selection of 70% of the dataset at random to train each tree. Since asymmetries and efficiencies vary slowly as a function of the distributions to be corrected, they will not be significantly biased as a result of overtraining, and therefore the choice of hyperparameters is not critical. *k-folding* [185] is employed to reduce bias, where the weight function applied to each fifth of the dataset is determined by training with the other four-fifths.

5.3.4 Background subtraction in data

Signal distributions in data are obtained using *sideband subtraction*. This is a robust background subtraction method where the background distributions from the $m(B^+)$ sidebands are subtracted from the signal window. The signal window is $|m(B^+)-m_{B^+}| < 30\,\mathrm{MeV}/c^2$ and the sidebands are defined as $30 < |m(B^+)-m_{B^+}| < 60\,\mathrm{MeV}/c$, limited by the presence of partially reconstructed $B^+ \to DD$ decays below $5220\,\mathrm{MeV}/c^2$. The signal model described in Section 6.1 plus an exponential to describe background from random combinations of tracks is fitted to the data. The relative background yields in the sidebands and the signal window are extracted from this fit and converted into the weights for sideband candidates needed to cancel the background in the signal window. In practice, if the signal window candidates are assigned a weight of 1, the sideband weight is within 0.5% of -1.

A more sophisticated method that results in smaller statistical uncertainties is the *sPlot* method [186]. This method assigns signal weights as a continuous function of $m(B)$ according to the fit of a signal plus background model to the $m(B)$ distribution. The result of applying these weights to the data is the background-subtracted distribution in any variable that is uncorrelated with $m(B)$. However, correlation can introduce a bias into the resulting distribution. The absence of correlation is not a trivial requirement: For example, mass resolution is correlated with p_{T}. Consequently, a bias of $140\,\mathrm{MeV}/c$ on $\left\langle p_{\mathrm{T}}(B^+ \to D_s^+ \overline{D}^0)\right\rangle = 11930 \pm 20\,\mathrm{MeV}/c$ is observed in Run 2 data as a result of using the *sPlot* method instead of sideband subtraction. Therefore the *sPlot* method is not used in these analyses.

Sideband subtraction could also introduce bias if the signal window is not much larger than the signal width. No bias is expected since the $\pm 30\,\text{MeV}/c^2$ window is three times the standard deviation in $m(B^+)$. This is confirmed by noting that using a $\pm 40\,\text{MeV}/c^2$ signal window with $20\,\text{MeV}/c^2$ sidebands results in only a $10\,\text{MeV}/c$ shift in $p_{\mathrm{T}}(B^+)$. Consequently, sideband subtraction is used for all background subtraction in these analyses.

5.3.5 B^+ kinematics and event multiplicity

B meson production as a function of $p_{\mathrm{T}}(B)$, $y(B)$, and event multiplicity is often mismodelled in simulation. Production and detection asymmetries tend to be smaller at high p_{T}, and signal efficiencies are in general larger. Detection asymmetries are larger at low y as a larger volume of material is traversed. A high-multiplicity event is more difficult to reconstruct, so tracking and PID efficiencies are smaller. The number of tracks in the event ($nTracks$) is used as a proxy for the event multiplicity. Therefore mismodelling of these quantities biases signal efficiencies and charge asymmetries, and must be corrected. For a decay with a clear signal in data, such as for $B^+ \to DD$ decays, this can be achieved by applying a weight to each simulated candidate so that the $(p_{\mathrm{T}}(B), y(B), nTracks)$ distribution matches that in background-subtracted data.

The B production spectrum and event multiplicity are assumed to be mismodelled in the same way in all final states, so the weight function is determined using the high yield $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K\pi$ final state, separately for Run 1 and Run 2. The selection described in Sections 5.5-5.7 is applied to both data and simulation samples. The unweighted and weighted distributions are compared to background-subtracted data in Figure 5.2, which demonstrates that this procedure corrects substantial mismodelling in $p_{\mathrm{T}}(B^+)$ and $nTracks$, and small discrepancies in $y(B^+)$. 2D projections of the mean weights are shown in Figure 5.3. The relative changes in efficiency between different final states, which does not depend on the normalisation of the weights, is shown in Table 5.3.

When measuring branching fraction ratios, it is essential that the effect of this weighting on the denominator of the efficiencies in Equations 4.2 and 8.18 cancels or otherwise vanishes. In the measurement of $B^- \to DD$ branching fractions relative to other $B^- \to DD$ decays, the effect cancels naturally, as the production spectrum does not depend on the decay products. Conversely when measuring the branching fraction of $B_c^+ \to DD$ decays relative to $B^+ \to DD$ decays, the weights must be normalised such that the denominator is unchanged. The $nTracks$ distribution prior to reconstruction and selection is not available, and therefore it is not possible to weight the simulation such that the denominator is unchanged. Therefore, in the measurement of the $B_c^+ \to DD$ branching fractions, simulation is not weighted in $nTracks$ for either B^+ or B_c^+ simulation.

Table 5.3: Fractional changes in efficiency relative to the $KK\pi, K\pi$ final state when applying weights as a function of $(p_{\mathrm{T}}(B), y(B), nTracks)$.

Final state	$\Delta\varepsilon/\epsilon$	
	Run 1	Run 2
$KK\pi, K3\pi$	$2.82 \pm 0.14\%$	$1.22 \pm 0.06\%$
$K\pi\pi, K\pi$	$0.13 \pm 0.00\%$	$-1.28 \pm 0.03\%$
$K\pi\pi, K3\pi$	$3.27 \pm 0.18\%$	$-0.36 \pm 0.20\%$
$K\pi\pi_s, K\pi$	$2.39 \pm 0.07\%$	$-0.04 \pm 0.03\%$
$K\pi\pi_s, K3\pi$	$4.79 \pm 0.20\%$	$-0.51 \pm 0.13\%$
$K3\pi\pi_s, K\pi$	$4.65 \pm 0.21\%$	$0.09 \pm 0.09\%$

This is acceptable, firstly because the statistical uncertainties on these measurements are expected to be much larger than the bias from event multiplicity mismodelling, and additionally because the effect of mismodelling on the efficiencies is expected to cancel between two decays with almost identical final states to a much higher degree than the systematic uncertainties related to limited knowledge on the true B_c^+ production distribution.

5.3.6 B_c^+ kinematics

Discrepancies between the true and simulated B_c^+ production spectra are also expected and must be corrected to obtain unbiased efficiencies. A forward-backward asymmetry in BCVEGPY under 2012 conditions[1] was corrected using weights obtained in $(p_{\mathrm{T}}(B_c^+), \eta(B_c^+))$ bins using simulated samples both with and without the asymmetry. Weights as a function of the decay time t,

$$w = \frac{\tau_{\mathrm{gen}}}{\tau_{\mathrm{PDG}}} \exp(t/\tau_{\mathrm{gen}} - t/\tau_{\mathrm{PDG}}), \tag{5.6}$$

where τ_{gen} is the simulated lifetime and τ_{PDG} is the world average lifetime [22], correct the decay time distribution.

A linear dependence of $\frac{f_c}{f_u+f_d}$ on p_{T} has been measured in the LHCb acceptance for $4 < p_{\mathrm{T}} < 25\,\mathrm{GeV}/c$ [106], parameterised as

$$\frac{f_c}{f_u + f_d} = p_1 + p_2(p_{\mathrm{T}} - 7200\,\mathrm{MeV}/c), \tag{5.7}$$

with the fitted coefficients (p_1, p_2) at $\sqrt{s} = 7, 13\,\mathrm{TeV}$ listed in Table 5.4. This function is extrapolated to $p_{\mathrm{T}} > 25\,\mathrm{GeV}/c$ by assuming it is valid until it becomes negative and is subsequently zero. No dependence on $\eta(B_c^+)$ was observed in Ref. [106].

The production distribution in $(p_{\mathrm{T}}(B_c^+), \eta(B_c^+))$ can be corrected so that the p_{T} and y-dependences of the ratio of B_c^+ to B^+ yields in generator-level simulation matches the

[1]This asymmetry was due to a bug in BCVEGPY at the time that the simulation samples were generated.

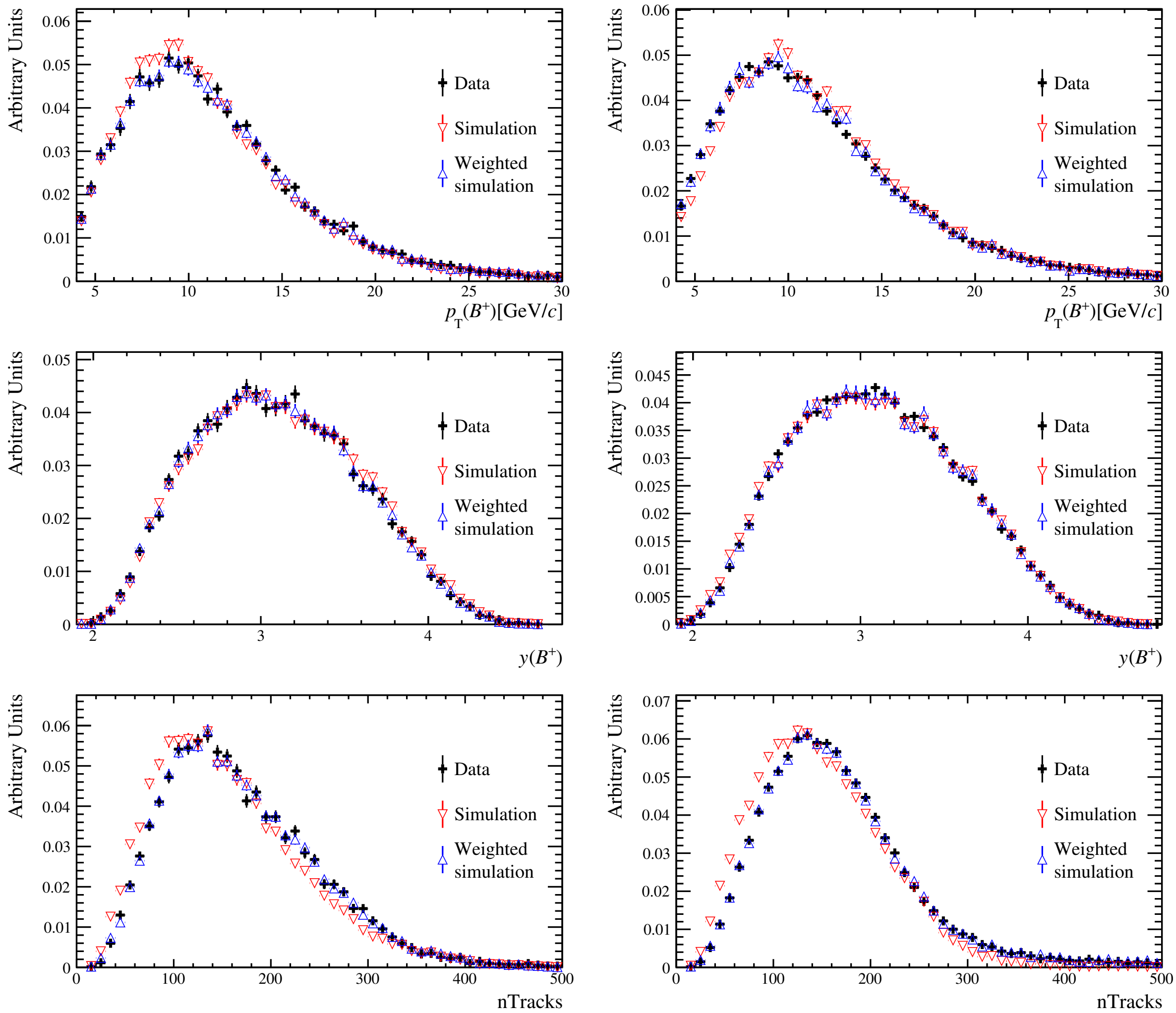

Figure 5.2: Distributions of (top) $p_{\mathrm{T}}(B^+)$, (middle) $y(B^+)$, and (bottom) $nTracks$ in (black) background-subtracted data, (red) simulation and (blue) corrected simulation. For $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K\pi$ candidates in (left) Run1 and (right) Run2. Each histogram is normalised to unit integral.

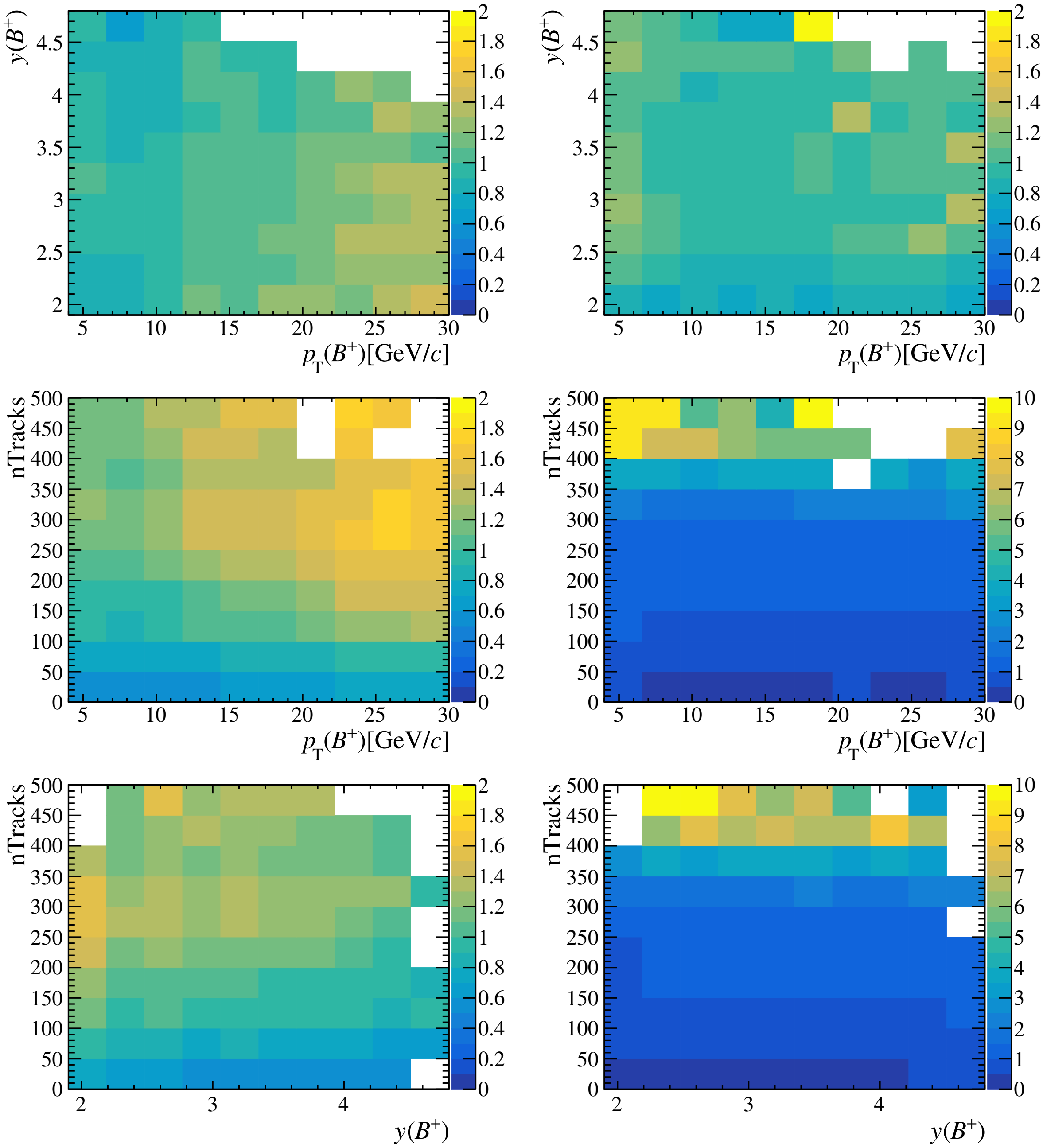

Figure 5.3: Weights applied to B^+ simulation to improve agreement with background-subtracted data in (left) Run 1 and (right) Run 2. Averaged in (top row) in (p_T, y), (middle row) $(p_T, nTracks)$ and (bottom row) $(y, nTracks)$ bins. Large weights of up to ten are present for Run 2 due to a tail of high-multiplicity ($nTracks > 400$) events in data that is effectively absent in simulation. Since the fraction of such events is small in both data and simulation, these large weights do not significantly increase uncertainties associated with the simulation sample size.

Table 5.4: Parameters describing the p_{T}-dependence of $\frac{f_c}{f_u+f_d}$ as defined in Eq. 5.7 [106].

$\sqrt{s}$	p_1	$p_2/10^{-5}(\mathrm{MeV}/c)^{-1}$
7 TeV	$3.82 \pm 0.09 \pm 0.05$	$-6.2 \pm 1.7 \pm 1.1$
13 TeV	$4.13 \pm 0.05 \pm 0.04$	$-9.7 \pm 0.8 \pm 1.0$

Table 5.5: Parameters of a Tsallis function fitted to the p_{T}-distribution of generated B_c^+ and B^+ mesons at $\sqrt{s} = 8, 13$ TeV. The fits are illustrated in Figure 5.4.

Parameter	2012	2016
$T(B_c^+)/\mathrm{MeV}/c^2$	$(1.411 \pm 0.026) \times 10^3$	$(1.504 \pm 0.026) \times 10^3$
$N(B_c^+)$	11.0 ± 0.6	10.2 ± 0.5
$T(B^+)/\mathrm{MeV}/c^2$	$(1.589 \pm 0.029) \times 10^3$	$(1.324 \pm 0.024) \times 10^3$
$N(B^+)$	9.3 ± 0.4	6.98 ± 0.19

measured dependence of f_c/f_u. The corrections on $(p_{\mathrm{T}}(B^+), y(B^+))$ and the B_c^+ forward-backward asymmetry are applied prior to performing this correction.

The B production distribution with respect to $p_{\mathrm{T}}(B)$ can be modelled by a Tsallis function [187, 188]:

$$T(p_{\mathrm{T}}) = \frac{C p_T}{\left(1 + \frac{\sqrt{p_T^2+M^2}-M}{T \cdot N}\right)^N}. \tag{5.8}$$

The Tsallis function $T(p_{\mathrm{T}})$ is fitted to the generator-level distributions with M fixed to $M(B_{(c)}^+)$ and treating T and N as free parameters, which vary as a function of $\sqrt{s}$ and between B_c^+ and B^+. The fitted $T(p_{\mathrm{T}})$ are shown in Figure 5.4. Additionally, the ratios of the B_c^+ to B^+ production spectra are plotted and the Tsallis function ratios are compared to the measured p_{T}-dependence of f_c/f_u. The fitted parameters are given in Table 5.5.

Although Ref. [106] studied f_c/f_u as a function of η, the $2.0 < y(B) < 4.5$ requirement used in this analysis will introduce non-linearities in the η-distribution of the ratio, and it is reasonable to assert that the absence of evidence for an η-dependence implies the absence of evidence for a y-dependence. Therefore the ratio of B_c^+ to B^+ yields in generator-level simulation is fitted to a linear function of $y(B)$,

$$\frac{N_{B_c^+}(y)}{N_{B^+}(y)} = q_1 + q_2 y. \tag{5.9}$$

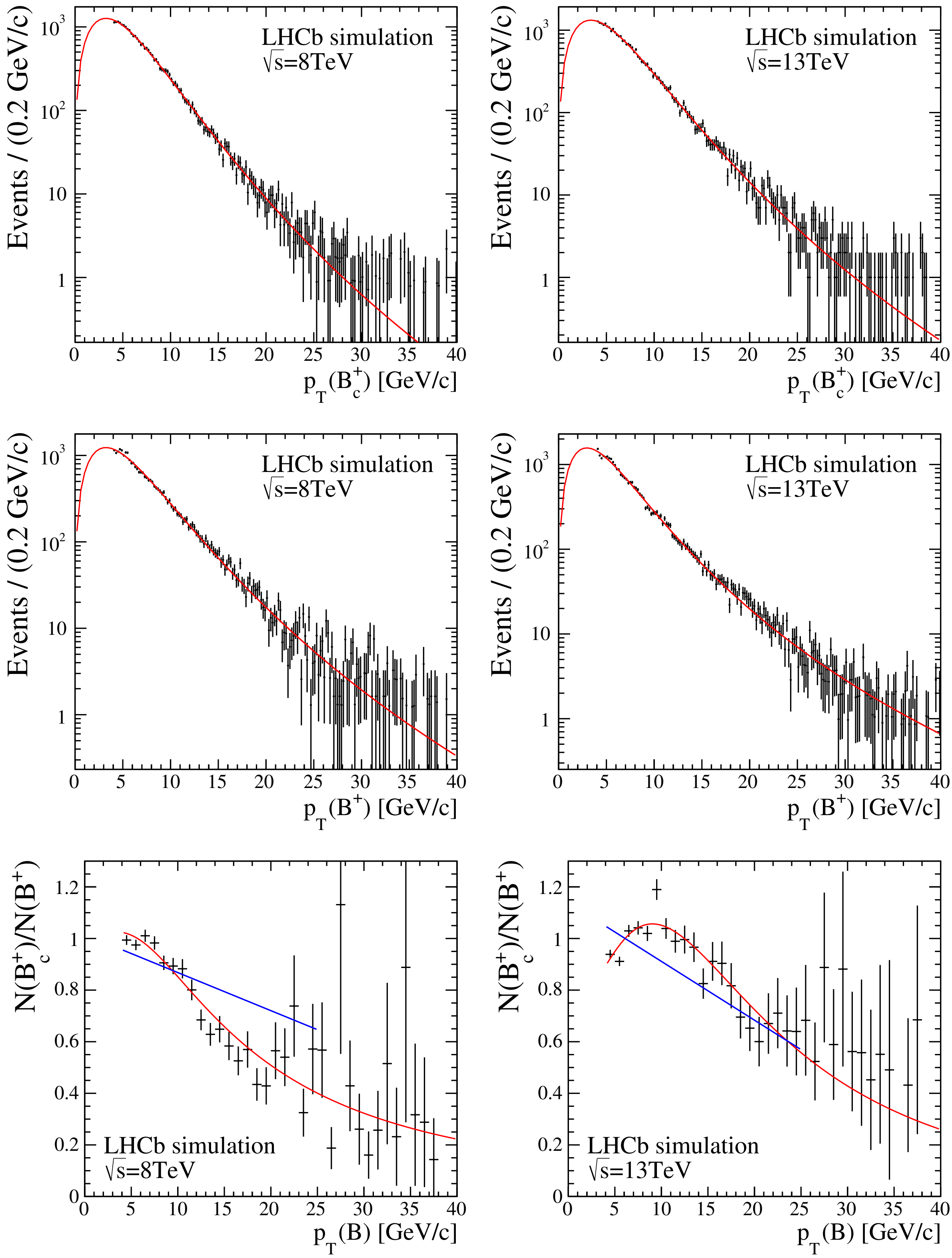

Figure 5.4: (top) Generated B_c^+ yield, (middle) generated B^+ yield and (bottom) their ratio as a function of p_{T} in the kinematic region of interest: $2.0 < \eta < 4.5$, $p_{\mathrm{T}} > 4\,\mathrm{GeV}/c$. For (left) 2012 simulation at 8 TeV and (right) 2016 simulation at 13 TeV. Weights to correct for asymmetric B_c^+ production and mismodelling of B^+ production have been applied. The fit of a Tsallis function to each generator-level distribution is overlaid (red), corresponding to the fitted parameters in Table 5.5. The measured $\frac{f_c}{f_u+f_d}$ as a linear function of p_{T} is also overlaid (blue), corresponding to the parameters in Table 5.4. The simulated ratio is consistent with the measured ratio, within uncertainties on the measurement and from the size of the simulation sample.

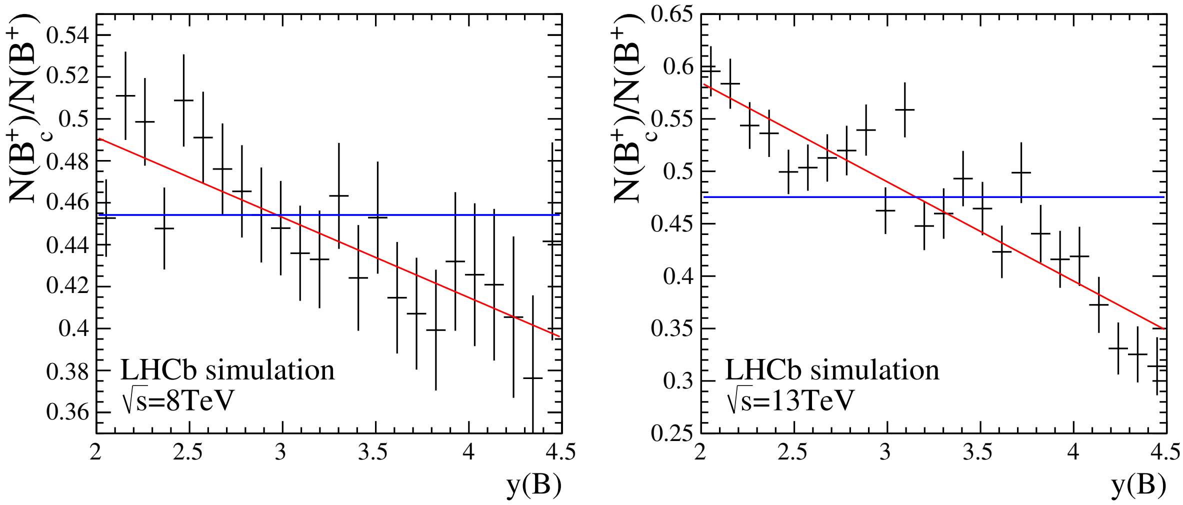

Figure 5.5: Ratio of generated B_c^+ yield to generated B^+ yield as a function of y in the kinematic region of interested: $2.0 < \eta < 4.5$, $p_{\mathrm{T}} > 4\,\mathrm{GeV}/c$. For (left) 2012 simulation at 8 TeV and (right) 2016 simulation at 13 TeV. Weights to correct for asymmetric B_c^+ production and B^+ kinematic weights have been applied. A linear fit is overlaid (red), corresponding to the fitted parameters in Table 5.6. The measured uniform $\frac{f_c}{f_u+f_d}$ as a function of y is also overlaid.

Table 5.6: Parameters of a linear fit to the y-distribution of generated B_c^+ and B^+ mesons at $\sqrt{s} = 8, 13$ TeV. The fits are illustrated in Figure 5.5.

Parameter	2012	2016
q_1	0.568 ± 0.024	0.774 ± 0.023
q_2	-0.038 ± 0.008	-0.095 ± 0.007

The fits are shown in Figure 5.5 and parameters are tabulated in Table 5.6. B_c^+ simulation is corrected by applying the weights

$$w = \begin{cases} (p_1 + p_2(p_{\mathrm{T}} - 7200\,\mathrm{MeV}/c)) \cdot \frac{T_{B^+}(p_{\mathrm{T}})}{T_{B_c^+}(p_{\mathrm{T}})} \cdot \frac{1}{q_1+q_2|y|} & p_{\mathrm{T}} \leq -\frac{p_1}{p_2} + 7200\,\mathrm{MeV}/c \\ 0 & p_{\mathrm{T}} > -\frac{p_1}{p_2} + 7200\,\mathrm{MeV}/c. \end{cases} \tag{5.10}$$

These weights are normalised so that the B_c^+ yield in the kinematic region of interest in generator-level simulation is unchanged. Figure 5.6 compares p_{T} and y distributions between weighted and unweighted data, and shows the normalised 2D weights. Table 5.7 shows the fractional changes in the B_c^+ and B^+ efficiencies, and their ratio, as a result of reweighting in p_{T} and y.

The p_{T} and y-dependence of f_c/f_u has also been measured using fully hadronic decays [189]. In contrast to Ref. [106], a dependence on y was observed. The potential effect of a change in y-dependence on the efficiencies is assessed by repeating the

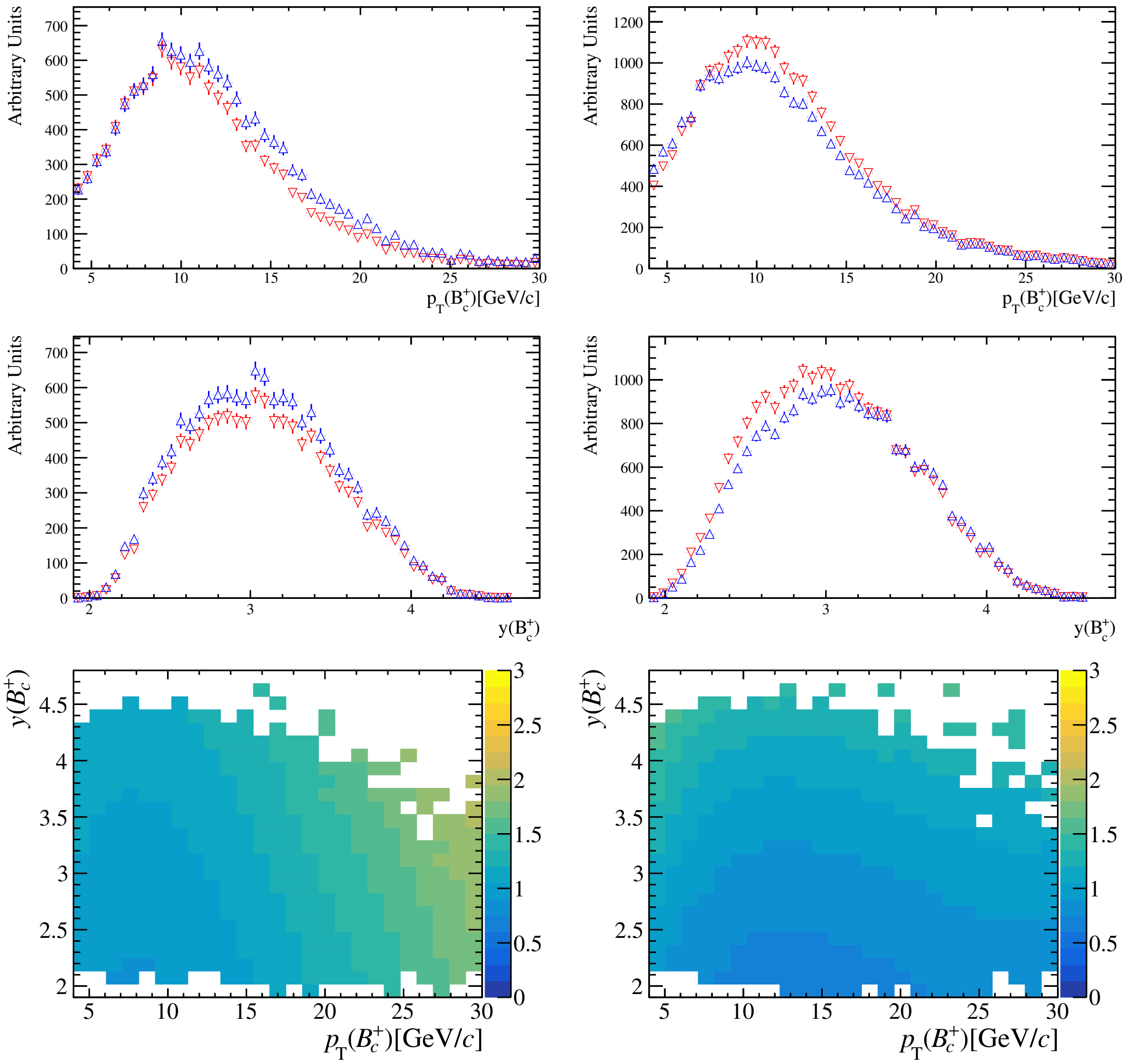

Figure 5.6: Comparison of (red) B_c^+ simulation and (blue) corrected B_c^+ simulation in (top) $p_T(B_c^+)$ and (middle) $y(B_c^+)$, and (bottom) corresponding 2D distribution of weights. For $B^+ \to D_s^+ \bar{D}^0$, $\bar{D}^0 \to K^+ \pi^-$ candidates in (left) Run1 and (right) Run2.

Table 5.7: Fractional change in efficiencies and efficiency ratio as a result of weighting as a function of $(p_{\mathrm{T}}(B), y(B))$. Both $\epsilon(B_c^+)$ and $\epsilon(B^+)$ receive sizeable corrections that cancel to a large extent in the ratio. Differences between Run 1 and Run 2 are predominantly due to changes in the B meson production distributions in simulation and statistical uncertainty on the measured f_c/f_u.

Period	Final state	$\Delta\epsilon(B_c^+)$	$\Delta\epsilon(B^+)$	$\Delta(\epsilon(B_c^+)/\epsilon(B^+))$
	$KK\pi, K\pi$	5.1%	1.9%	3.1%
	$KK\pi, K3\pi$	9.0%	4.5%	4.4%
	$K\pi\pi, K\pi$	5.1%	2.7%	2.3%
Run 1	$K\pi\pi, K3\pi$	8.4%	4.6%	3.6%
	$K\pi K\pi$	5.9%	3.4%	2.5%
	$K\pi K3\pi$	9.6%	5.1%	4.3%
	$K3\pi K\pi$	8.4%	4.4%	3.8%
	$KK\pi, K\pi$	−6.7%	−6.3%	−0.4%
	$KK\pi, K3\pi$	−6.5%	−6.8%	0.4%
	$K\pi\pi, K\pi$	−6.7%	−6.1%	−0.7%
Run 2	$K\pi\pi, K3\pi$	−6.9%	−6.8%	0.0%
	$K\pi, K\pi$	−6.2%	−5.9%	−0.3%
	$K\pi, K3\pi$	−6.4%	−6.8%	0.4%
	$K3\pi, K\pi$	−6.4%	−6.4%	0.0%

correction to B_c^+ kinematics described in this section without considering y. The change in efficiency is $1-2\%$, which is smaller than the systematic uncertainty assigned to the B_c^+ kinematic correction in Section 7.3.1.

5.3.7 Angular distribution of four-body D^0 decay

The angular distribution of the $D^0 \to K^-\pi^+\pi^-\pi^+$ decay is modelled in Sim08 simulation according to a uniform distribution in phase space, when reality it depends strongly upon the presence of intermediate resonances in the decay. The comparison of simulation and background-subtracted data in Figure 5.7 shows this difference as the absence or presence, respectively, of resonant peaks in the two- or three-body invariant mass combinations of the decay products. This mismodelling of the angular distribution may bias the momentum distributions of the final state particles, and therefore the efficiencies. Therefore Sim08 simulation is weighted to match background-subtracted data in the ten two- or three-body invariant mass combinations. The CF $D_s^+\overline{D}^0$ decay is used for both the $D_s^+\overline{D}^0$ and $D^+\overline{D}^0$ channels. Otherwise, the methodology is the same as in Section 5.3.5. The weighted and unweighted simulation is compared to background-subtracted data in Figure 5.7, which demonstrates that significant differences are corrected by this procedure. The weights are normalised so that the yield in generator-level simulation is unchanged. Applying these normalised weights to simulation increases the efficiencies by up to 2%.

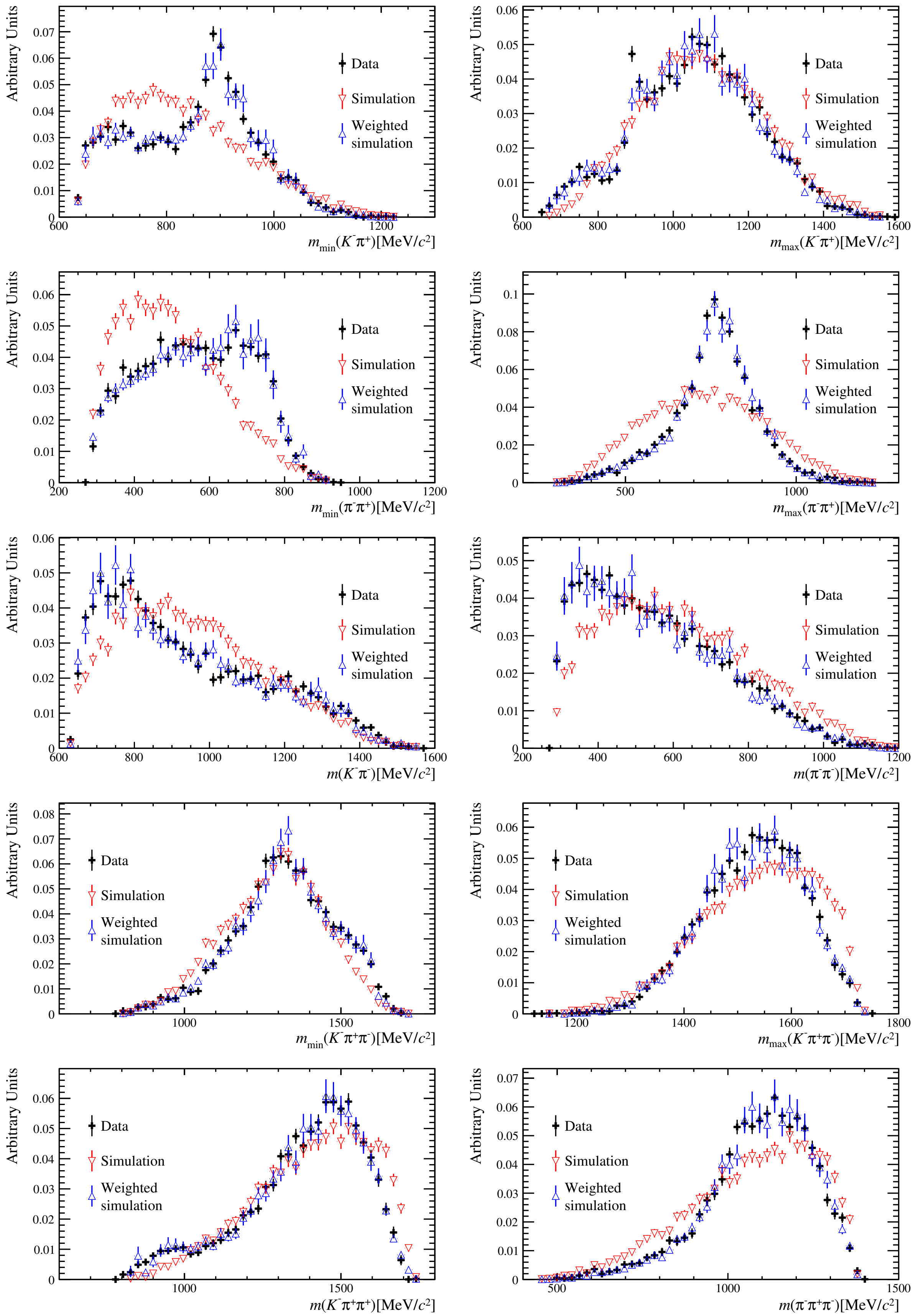

Figure 5.7: Angular distributions in (black) background-subtracted data, (red) simulation and (blue) corrected simulation. Each histogram is normalised to unit integral.

5.3.8 Homogenisation of simulation versions

As discussed in Section 5.2, 2017 stripping-flagged simulation samples are used to obtain PID efficiencies. The simulation versions are Sim09f and Sim09h for $B^+ \to D_s^+ \overline{D}^0$ and $D^+ \overline{D}^0$, respectively, whereas the remainder of the Sim09 simulation of B^+ decays is Sim09c. The distributions of B meson kinematics and $nTracks$ are known to differ even between these minor versions, in particular due to changes in the treatment of resonant states in hadronisation and decay processes [190]. The Sim09c samples have been corrected as described in Section 5.3.5, so the Sim09f and Sim09h samples must be corrected to match the Sim09c samples. Correction weights are obtained as a function of $(p_{\mathrm{T}}(B), y(B), nTracks)$ using a GBR similarly to in previous sections. The normalisation of these weights is not important because they are only used to obtain per-candidate PID efficiencies. The effect on this correction on the relevant distributions is shown in Figure 5.8.

5.3.9 Weighting to emulate the kinematics of partially reconstructed decays

The charge asymmetries in production and detection depend on the kinematics of the candidate. Full simulation is used to model the kinematics of fully reconstructed decays. Full simulation was not produced for the partially reconstructed $B^+ \to DD$ decays. Regardless, good precision can be obtained on the production and detection asymmetries for partially reconstructed decays by weighting the full simulation of the fully reconstructed decays to achieve a softer spectrum.

To compare the distribution of $p(B)$ between partially and fully reconstructed signal, background subtraction is performed using the charm sidebands ($25 < |m(D) - m_D| < 75\,\mathrm{MeV}/c^2$) in the mass range occupied by partially reconstructed signal ($5040 < m(B) < 5140\,\mathrm{MeV}/c^2$) and fully reconstructed signal ($5250 < m(B) < 5310\,\mathrm{MeV}/c^2$) (Figure 5.9). An appropriate weight function for modelling the partially reconstructed signal kinematics is determined from the ratio of the $p(B)$ distributions. A power law of the form p^{-a} describes the ratio well and is therefore chosen as the weight function. The ratio is fitted to a power law for each channel, final state and for Run 1 and 2 (Table 5.8). The overall correction is a power law p^{-a} whose power is the weighted average over all final states and data taking periods.

However, this weighting does not in general correct the momentum distributions of the final state particles, which are required for the detection asymmetries. This is seen in Figure 5.10, which also shows that the size of the shift due to weighting is similar to the size of the discrepancy. Therefore, for corrections that depend on the kinematics of the final state particles, the difference between applying this power law weighting rule and applying

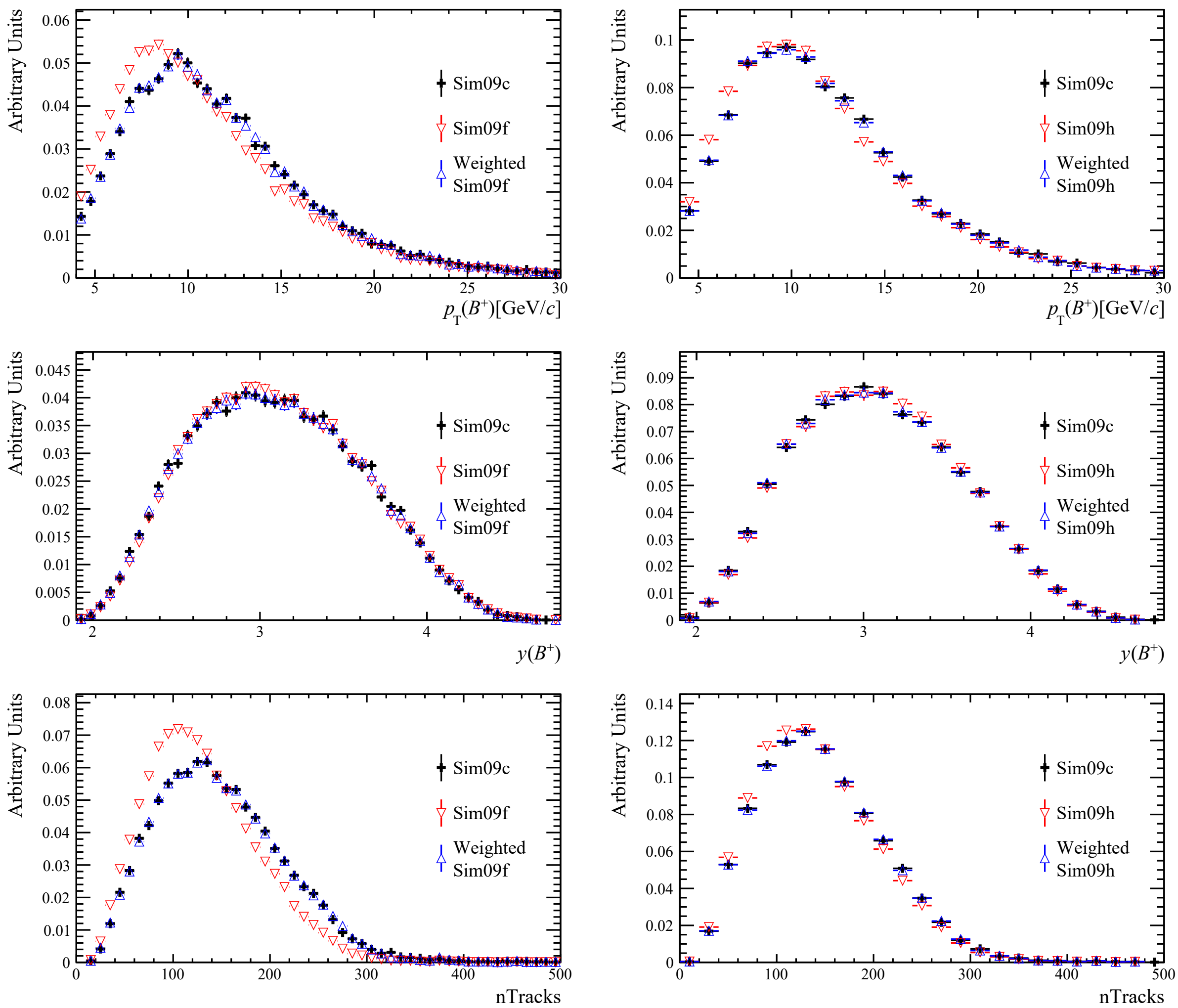

Figure 5.8: Comparison of (blue) corrected and (red) uncorrected (left) Sim09f and (right) Sim09h simulation to (black) Sim09c simulation. Large differences in p_T and $nTracks$ are corrected. Each histogram is normalised to unit integral.

no weights will be assigned as a systematic uncertainty for partially reconstructed decays.

5.4 Comparison of simulated and true signal distributions

Corrected simulation is compared to background-subtracted data in Appendix A for the highest signal yield dataset: $B^+ \to D_s^+ \overline{D}^0$ decays in Run 2 data. Several variables compared in this Appendix have not yet been defined. These are:

- **Decay time (t):** The time between the production and decay of the particle in the particle's rest frame. This is the distance between the production and decay vertices measured in the VELO. The decay time can be negative due to finite resolution in the VELO.

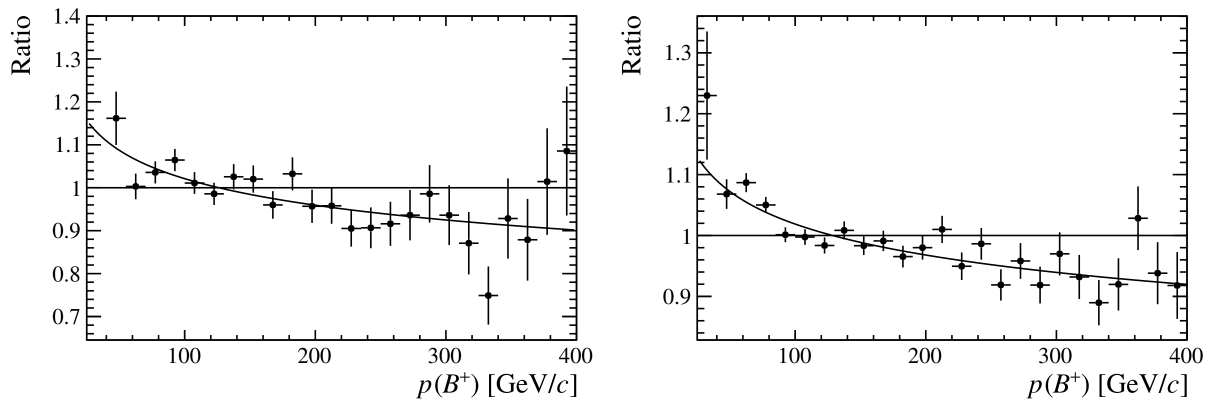

Figure 5.9: Ratio of $p(B^-)$ spectrum between partially reconstructed region $5040 < m(B^-) < 5140\,\mathrm{MeV}/c^2$ and fully reconstructed region $5250 < m(B^-) < 5310\,\mathrm{MeV}/c^2$ in background-subtracted data. For $B^- \to D_s^- D^0$, $D^0 \to K^-\pi^+$ candidates in (left) Run1 and (right) Run2 data. The spectrum is softer in partially reconstructed than in fully reconstructed decays. The ratio can be described by a power law (fitted curve).

- **Decay time significance ($t/\Delta t$):** The measured decay time divided by its uncertainty.
- **Vertex χ^2 ($\chi^2_{\rm vtx}$):** The χ^2 of a vertex fit.
- **Impact parameter (IP):** The shortest distance between the extrapolated trajectory of the particle and a vertex, in general the primary vertex.
- **Impact parameter χ^2 ($\chi^2_{\rm IP}$):** The change in the χ^2 of the vertex fit when the particle is removed from the vertex. This quantity behaves similarly to $(\mathrm{IP}/\sigma_{\rm IP})^2$.
- **Track χ^2 per degree of freedom($\chi^2_{\rm trk}$/ndf):** The χ^2 per degree of freedom quality of the track fit.
- **Dalitz variables (m_{ij}):** The invariant mass of the combination of two decay products in a three-body decay. Three-body decays often proceed through a decay to one of the decay products and an intermediate resonance that decays to the remaining decay products, for example $D_s^+ \to (\phi(1020) \to K^+K^-)\pi^+$. Therefore genuine three-body decays often exhibit strong peaks in the Dalitz variables. This can be used to distinguish true signal candidates involving three-body decays from backround candidates with random combinations from tracks, which have a flatter distribution in the Dalitz variables.

The D invariant masses and kinematics, the track kinematics, and PID variables are all well modelled following the corrections applied in Section 5.3. A small degree of mismodelling is observed for $t/\Delta t$, $\chi^2_{\rm IP}$, $\chi^2_{\rm vtx}$, and Dalitz variables. The $\chi^2_{\rm trk}$/ndf is substantially mismodelled. Examples of good modelling of the track kinematics after corrections and poor modelling of $\chi^2_{\rm trk}$/ndf are shown in Figure 5.11.

Table 5.8: Result of the fit of a power law to the ratio in $p(B)$ of partially to fully reconstructed background-subtracted data.

Dataset	Final State	a
Run 1	$KK\pi, K\pi$	0.089 ± 0.017
	$KK\pi, K3\pi$	0.034 ± 0.026
	$K\pi\pi, K\pi$	0.11 ± 0.07
	$K\pi\pi, K3\pi$	0.20 ± 0.16
	$K\pi\pi_s, K\pi$	0.40 ± 0.16
	$K\pi\pi_s, K3\pi$	0.29 ± 0.16
	$K3\pi\pi_s, K\pi$	0.11 ± 0.18
Run 2	$KK\pi, K\pi$	0.072 ± 0.008
	$KK\pi, K3\pi$	0.047 ± 0.012
	$K\pi\pi, K\pi$	0.15 ± 0.04
	$K\pi\pi, K3\pi$	0.13 ± 0.05
	$K\pi\pi_s, K\pi$	-0.05 ± 0.07
	$K\pi\pi_s, K3\pi$	0.10 ± 0.09
	$K3\pi\pi_s, K\pi$	0.23 ± 0.09

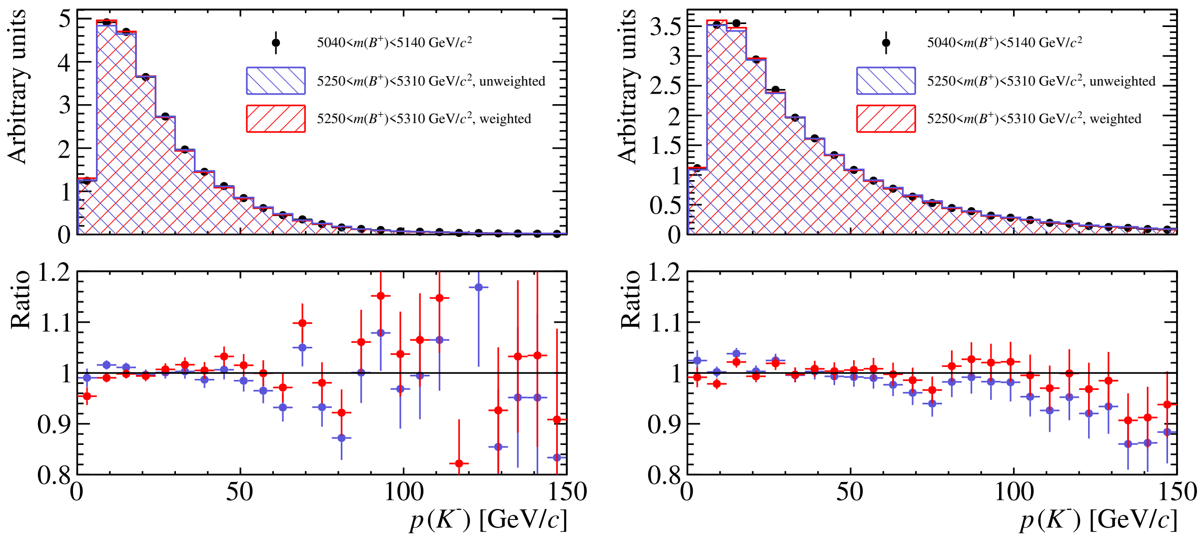

Figure 5.10: Momentum distribution of (left) the K^- from the D_s^- and (right) the K^- from the D^0. In (points) the mass region occupied by partially reconstructed signal and (filled) the mass region occupied by fully reconstructed signal, both with (red) and without (blue) $p(B)$ weights applied. The bottom pannel shows the ratio between data in the fully reconstructed and partially reconstructed regions. The agreement is generally poor both with and without weights.

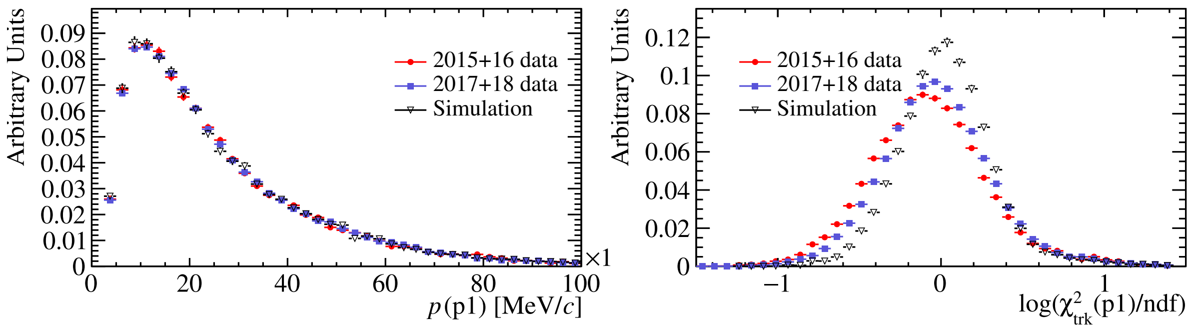

Figure 5.11: Comparison of (left) $p(K^+)$ after momentum scale and resolution corrections, and (right) $\log(\chi^2_{\rm trk}/{\rm ndf}(K^+))$ for the K^+ from the D_s^+ between background-subtracted data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \overline{D}{}^0$, $D^0 \to K^-\pi^+$ decays.

The effect of mismodelling $\chi^2_{\rm trk}/{\rm ndf}$ is mitigated by only applying a loose selection on this variable in the trigger and stripping selection (Section 5.6). However, even this loose requirement could introduce a bias to a high precision branching fraction measurement as discussed in Section 8.4. For the remaining mismodelled variables, the bias induced by this mismodelling is also expected to be small in general, since the relevant selection requirements have a high signal efficiency and only branching fraction ratios or charge asymmetries are considered. However, care must be taken when including these variables in the selection. The choice of variables in a multivariate discriminant will be discussed in Section 5.8.1, and the biases introduced through the use of these variables will be quantified in Sections 7.3.1 and 8.4.

5.5 Trigger selection

Candidates for these analyses are selected by specific trigger lines at each level of the trigger system. At L0, either the candidate fires the hadron trigger (L0 Hadron TOS) or the rest of the event fires any L0 trigger (L0 Global TIS). The fraction of candidates in each category depends on the momentum of the final state particles but is approximately 60% in each category after full selection is applied, as some events are triggered in both ways.

At HLT1, an inclusive trigger described in Section 3.2.3 selects a single good quality high-$p_{\rm T}$ track displaced from the PV. At HLT2, the inclusive topological triggers select the majority of candidates. A trigger line selecting the $\phi(1020) \to K^+K^-$ resonance additionally selects candidates with a D_s^+, since 42% of $D_s^+ \to K^+K^-\pi^+$ decays proceed through this narrow resonance [22]. This trigger line selects two kaon candidates with good $\chi^2_{\rm trk}/{\rm ndf}$, high $p_{\rm T}$, high $\chi^2_{\rm IP}$, and a small distance of closest approach (DOCA) to each other. The kaon pair is combined to form a $\phi(1020)$ candidate, which is selected if it is within a $\pm 20\,{\rm MeV}/c^2$ mass window, has high $p_{\rm T}$, and has a high $\chi^2_{\rm vtx}$. The candidate is required to be TOS to both the HLT1 and HLT2 decisions.

5.6 Stripping selection

The output of the following exclusive stripping lines select $B_c^+ \to DD$ candidates from data saved by the triggers described in the previous section.

- `StrippingBc2DD0D2HHHD02HHBeauty2CharmLine`
- `StrippingBc2DD0D2HHHD02KHHHBeauty2CharmLine`
- `StrippingBc2DstD0Dst2D0PID02HHD02HHBeauty2CharmLine`
- `StrippingBc2DstD0Dst2D0PID02HHD02KHHHBeauty2CharmLine`
- `StrippingBc2DstD0Dst2D0PID02KHHHD02HHBeauty2CharmLine`

These stripping lines apply the selection specified in Table 5.9. The *vertex separation* χ^2 (χ^2_{VS}) is the χ^2 significance of the separation between the decay vertex of the particle and the PV. The *direction angle* θ is the angle between the particle momentum and the line from the PV to the particle decay vertex; $\cos(\theta)$ is also referred to as DIRA. The first block in this table correspond to requirements either on the B or on all tracks in the candidate, or the number of long tracks in the event. The second block specifies requirements on each D daughter or on all tracks forming the D candidate. The combination of the D^0 with a soft π^+ produces a poor quality vertex, so requirements on the D vertex quality or geometry are not applied for the D^{*+}. Instead a cut on the mass difference $\Delta m(D^{*+}) = m(D^{*+}) - m(D^0)$ is applied to reject background from the combination of a D^0 and π^+ where one or both are not from the signal decay. The same requirements are applied to the D^0 from the D^{*+} decay as is applied to the other D^0. The final block lists the requirements on each final state track in the candidate.

5.7 Preselection

Additional selection is applied to reduce the background from specific decays or from random combinations of tracks. To address the second category, the B candidate is required to have $p_{\text{T}} > 4\,\text{GeV}/c$ and $\chi^2_{\text{IP}} < 15$; pions must have $\text{DLL}_{K\pi} < 10$; and kaons must have $\text{DLL}_{K\pi} > -5$. The invariant mass of all charm meson candidates must be within $25\,\text{MeV}/c^2$ of their known mass [22]. An opening angle of greater than 0.5 mrad between all tracks in each candidate is required to avoid the multiple use of a single VELO track.

There is an $\mathcal{O}(1\%)$ charge asymmetry in the interactions of kaons with detector material that must be corrected when measuring the $\mathcal{A}^{CP}$. The corresponding asymmetry for pions is also corrected, but is much smaller and consequently a precise correction is less important. The interaction cross-section is smoothly varying for $p > 3\,\text{GeV}/c$ and

Table 5.9: Cuts applied in the stripping to select $B^+_{(c)} \to D^{(*)+}_{(s)} \overline{D}^0$ candidates.

Variable	Requirement
B and event	
Number of long tracks	< 500
$m(B)$	$> 4800\,\text{MeV}/c^2$; $< 6800\,\text{MeV}/c^2$
$t(B)$	$> 0.05\,\text{ps}$
$\chi^2_{\text{vtx}}/\text{ndf}\ (B)$	< 10
$\chi^2_{\text{IP}}(B)$	< 20 (25 in Run 1)
$\cos(\theta_B)$	> 0.999
$\sum_{trk} p_{\text{T}}$	$> 5000\,\text{MeV}/c$
$\min(\chi^2_{\text{trk}}/\text{ndf})$	< 2.5 (4 in Run1)
max(track p_{T})	$> 1700\,\text{MeV}/c$
max(track p)	$> 10000\,\text{MeV}/c$
max(track IP)	$> 0.1\,\text{mm}$
max(track χ^2_{IP})	> 16
Each *D* daughter	
$m(D^+_{(s)})$	$> m_{D^+} - 100\,\text{MeV}/c^2$; $< m_{D^+_s} + 100\,\text{MeV}/c^2$
$\lvert m(D^0) - m_{D^0}\rvert$	$< 100\,\text{MeV}/c^2$
$\Delta m(D^{*+})$	$< 200\,\text{MeV}/c^2$
$\chi^2_{\text{vtx}}/\text{ndf}\ (D^+_{(s)},D^0,D^{*+})$	< 10
$\chi^2_{\text{VS}}\ (D^+_{(s)},D^0,D^{*+})$	> 36
$\cos(\theta_{D^+_{(s)},D^0,D^{*+}})$	> 0
$\sum_{trk} p_{\text{T}}$	$> 1800\,\text{MeV}/c$
max(track p_{T})	$> 500\,\text{MeV}/c$
max(track p)	$> 5000\,\text{MeV}/c$
DOCA for all track combinations	$< 0.5\,\text{mm}$
Each final state track	
p_{T}	$> 100\,\text{MeV}/c$
p	$> 1000\,\text{MeV}/c$
$\chi^2_{\text{trk}}/\text{ndf}$	< 4 (3 in Run 1)
χ^2_{IP}	> 4
$\text{DLL}_{K\pi}(\pi)$	< 10 (20 in Run 1 or D^{*+} decay product)
$\text{DLL}_{K\pi}(K)$	> -5 (10 in Run 1 or D^{*+} decay product)
Ghost probability	< 0.4

varies on small momentum scales below this [22]. To ensure the accuracy of this important correction, kaons with momenta below 3 GeV/c are not used in the *CP* asymmetry measurement.

Cross-feed between decay channels involving a D_s^+ and D^+ is possible through $K-\pi$ misidentification and forms a peaking background. To reduce $D_s^+ \to D^+$ cross-feed, a $D^+ \to K^-\pi^+\pi^+$ candidate is rejected if, when one of the π^+ is assigned a K^+ hypothesis, the resulting $K^-K^+\pi^+$ combination has a mass within 25 MeV/c^2 of $m_{D_s^+}$, and either $m(K^-K^+)$ is within 10 MeV/c^2 of $m_{\phi(1020)}$ or the original π^+ has $\mathrm{DLL}_{K\pi} > -10$. $D^+ \to D_s^+$ cross-feed is reduced by rejecting a $D_s^+ \to K^-K^+\pi^+$ candidate if $m(K^-K^+)$ is not within 10 MeV/c^2 of $m_{\phi(1020)}$, $\mathrm{DLL}_{K\pi}(K^+) < 10$, and the $K^-\pi^+\pi^+$ combination has a mass within 25 MeV/c^2 of m_{D^+} when the K^+ is assigned a π^+ hypothesis.

$D_s^+ \to K^-K^+\pi^+$ and $D^0 \to K^-\pi^+(\pi^-\pi^+)$ candidates can be duplicated if a kaon and pion are both consistent with the PID requirements for both kaon and pion hypotheses, namely $-5 < \mathrm{DLL}_{K\pi} < 10$, and both particle species assignments produce charm mesons within ± 25 MeV/c^2 of the known mass. For the D_s^+, the candidate with $|m(K^+K^-) - m_{\phi(1020)}| <$ 10 MeV/c^2 is retained. If neither or both candidates satisfy this condition, the candidate with $\mathrm{DLL}_{K\pi}(K) > \mathrm{DLL}_{K\pi}(\pi)$ is retained. For the D^0, the candidate with $\mathrm{DLL}_{K\pi}(K) > \mathrm{DLL}_{K\pi}(\pi)$ is retained.

Although the single charm decay $B^+ \to \bar{D}^0 K^-\pi^+\pi^+$ is forbidden at tree level, $B^+ \to \bar{D}^0\pi^-\pi^+\pi^+$ has a branching fraction over one hundred times larger than $B^+ \to \bar{D}^0(D^+ \to \pi^+\pi^+K^-)$ [22] and produces a large peaking background in this channel above m_{B^+} through $\pi \to K$ misidentification. In addition to modifying the background shape, the presence of this background when training a multivariate discriminant (Section 5.8) would reduce the efficacy of this discriminant in rejecting the main background from random combinations of tracks. This background can be rejected by requiring $t/\Delta t(D^+) > 3$. The $B^+ \to \bar{D}^0 K^-K^+\pi^+$ decay produces a peaking background in the $\bar{D}^0 D_s^+$ channel. However, the D_s^+ meson has half the lifetime of the D^+ meson [22], so a looser cut $t/\Delta t(D_s^+) > -3$ is required to maintain a high signal efficiency. The resulting background rejection is not sufficiently low, so this background must be included when modelling the invariant mass distribution (Section 6.4). $B^+ \to D_{(s)}^+ K^+\pi^-$ decays proceed through a V_{ub} annihilation with low branching fraction and consequently only a loose requirement on the D^0 lifetime significance $t/\Delta t(D^0) > -3$ is necessary.

$B_{(s)}^0 \to D_{(s)}^-\pi^+$ decays can additionally form a peaking background when combined with one or three charged tracks from the underlying event. Similarly to the single charm backgrounds, this modifies the background shape and reduces the ability of a multivariate discriminant to reject background from random combinations of tracks. This background is vetoed by requiring that $|m(D_s^+\pi^-(\pi^+\pi^-)) - m_{B_s^0}| > 30$ MeV/c^2 and $|m(D^{(*)+}\pi^-(\pi^+\pi^-)) - m_{B^0}| > 30$ MeV/c^2. It is not necessary to consider $B^0 \to \bar{D}^0\pi^+\pi^-$ decays that require π^- misidentification in addition to an extra track from the event to be reconstructed in one of the analysed final states, or other $B^0 \to \bar{D}^0 h^+ h^-$ decays with a lower branching fraction.

5.8 Multivariate selection

Even after applying preselection, the level of background from random combinations of tracks is still too high to permit a search for the rare $B_c^+ \to DD$ decays that lack a distinctive signature. The precision on measurements of $B^- \to DD$ decays can also be improved with an additional rejection of background. Further background rejection is achieved with a *multivariate analysis* (MVA) that quantifies the signal-likeness of a candidate according to its location in a multidimensional parameter space. The mapping between the set of input parameters and output response is obtained through a machine learning algorithm that uses the parameter values in a signal and a background training sample as input. The configuration of the MVA is discussed in Section 5.8.1. The selection of data according to MVA output is explained in Section 5.8.2. Its dependence on the B charge is investigated in Section 5.8.3.

5.8.1 Multivariate classifier

An adaptive BDT [191] classifier has been found to be most discriminating between signal and background data for these decay modes [192]. BDTs have already been introduced in Section 5.3.5; in the context of classification, each DT splits the training data into nodes to achieve the best signal-background separation and each leaf is classified depending on whether the training data on the leaf is predominantly signal or background. Adaptive boosting gives events misclassified during the training of one DT a higher weight when training the next DT. When the BDT is applied to another dataset, the output is a *response* value for each candidate, which is obtained by averaging over candidate classification for each DT. We transform the BDT response such that it is a uniform distribution between 0 and 1 for the signal training sample, so the requirement on the BDT response is equal to the signal rejection. Training and classification is implemented in the TMVA package [193]. $k = 5$-folding reduces bias [185].

Different BDTs are trained for each final state and separately for Run 1 and Run 2, but RS and WS candidates use the same BDT. Large samples of signal-like and background-like data are required to train the BDT. The signal training sample is signal simulation passing all selection described in the previous sections, within $\pm 40\,\text{MeV}/c^2$ of the known B mass and $\pm 25\,\text{MeV}/c^2$ of the known D masses, to avoid training with low-quality signal candidates. The corrections described in Section 5.3 are applied. Separate BDTs are trained for the analyses of $B_c^+ \to DD$ and $B^- \to DD$ decays, and the signal training sample corresponds to the appropriate B meson. The background training sample is both RS and WS candidates

Table 5.10: The number of events in the signal and background training samples for the BDT. Note that when applying $k = 5$-folding, 80% of the events are used in training.

Period	Final state	Signal	Background
Run 1	$KK\pi, K\pi$	2.4×10^4	7.2×10^5
	$KK\pi, K3\pi$	7.6×10^3	9.6×10^5
	$K\pi\pi, K\pi$	1.6×10^4	2.6×10^5
	$K\pi\pi, K3\pi$	5.8×10^3	4.9×10^5
	$K\pi, K\pi$	6.4×10^4	1.0×10^4
	$K\pi\pi, K3\pi$	3.2×10^3	1.8×10^4
	$K\pi, K\pi$	2.4×10^4	1.6×10^4
Run 2	$KK\pi, K\pi$	3.0×10^4	1.5×10^7
	$KK\pi, K3\pi$	1.2×10^4	2.9×10^6
	$K\pi\pi, K\pi$	2.4×10^4	5.4×10^5
	$K\pi\pi, K3\pi$	9.3×10^3	9.4×10^5
	$K\pi, K\pi$	6.0×10^4	3.0×10^4
	$K\pi\pi, K3\pi$	5.4×10^3	5.5×10^4
	$K\pi, K\pi$	2.0×10^4	5.1×10^4

in the $5350 < m(B) < 6200\,\text{MeV}/c^2$ sideband between m_{B^+} and $m_{B_c^+}$. The charm mass windows are extended to $\pm75\,\text{MeV}/c^2$ to increase the size of the training sample. This does not strongly modify the composition of the background sample as most background candidates are either single-charm or charmless, as opposed to combinations of two random charm mesons [192]. The training sample sizes are given in Table 5.10.

The input variables are:

- $\chi^2_{\text{vtx}}(B)$
- $\chi^2_{\text{vtx}}(D)$ for both D mesons
- $\chi^2_{\text{IP}}(B)$
- $t/\Delta t(B)$
- $t/\Delta t(D)$ for both D mesons
- $m(D)$ for both D mesons
- $m(K^+K^-)$ and $m(\pi^+K^-)$: the Dalitz variables for the D_s^+ decays
- $\max(m(\pi^+K^-))$ and $\min(m(\pi^+K^-))$: the Dalitz variables for the D^+ decays
- $\Delta m(D^{*+})$
- min(track p_{T}) for each D meson
- min(track χ^2_{IP}) for each D meson

- min(track $\mathrm{DLL}_{K\pi}$ (K)) for each D meson
- max(track $\mathrm{DLL}_{K\pi}$ (π)) for each D meson

Logarithms are applied to $t/\Delta t$, p_{T}, track χ^2_{IP}, and $\mathrm{DLL}_{K\pi}$ variables to compress long tails, after a subtraction to ensure positive definite values. Charm masses are folded in by adding (subtracting) $50\,\mathrm{MeV}/c^2$ if they are above (below) the $\pm25\,\mathrm{MeV}/c^2$ window around the known mass. This avoids the training of the classifier to reject charm sideband events entirely using charm masses, and maintains a large sample of selected charm sideband candidates for background studies. The distributions in signal and background of the transformed input variables is shown in Appendix B. No further improvement in performance was observed by uniformising or decorrelating the variables [192]. The distribution of BDT response for signal and background samples is shown in Figure 5.12.

As discussed in Section 5.4, several of these variables are mismodelled in simulation. This dataset is too small to assess the degree of mismodelling for $\Delta m(D^{*+})$, but in general it is known to be poorly modelled. This is unimportant in the $\mathcal{A}^{CP}$ measurement where biases cancel in the charge difference. Much more care is required when measuring branching fraction ratios. When measuring the branching fraction of $B_c^+ \to DD$ decays, it is not necessary to know this bias to high precision due to the large statistical uncertainty in this measurement. In the measurement of branching fractions in $B^- \to DD$ decays, the bias on the branching fraction ratios as a consequence of mismodelling can be quantified as the change when correcting the mismodelled variables to match data (Section 8.4). To measure this bias with acceptable precision, the weighting rule is determined using the CF $B^- \to D_s^- D^0$ decay. This is not possible when correcting the properties of the charged charm meson. Therefore the $t/\Delta t$ and χ^2_{vtx} of the charged charm meson, Dalitz variables, $\Delta m(D^{*+})$, and $\mathrm{DLL}_{K\pi}$ are removed from the BDT when measuring the $B^- \to DD$ branching fractions. This does not substantially reduce the performance.

The performance of a BDT can be quantified as the integral under the receiver operator characteristic (ROC) curve of background rejection as a function of signal efficiency. The maximum depth of the DTs was limited to 2 to avoid overtraining, which with 800 trees produced a ROC integral of 0.995 when training to select $B_c^+ \to D_s^+ \overline{D}{}^0$ decays in Run 2 data compared to 0.992 with a maximum depth of 1. No improvement in performance was observed with a higher number of trees was observed. For each DT, only four variables are considered at random, which further mitigates overtraining. This does not reduce the performance compared to considering all variables for every tree.

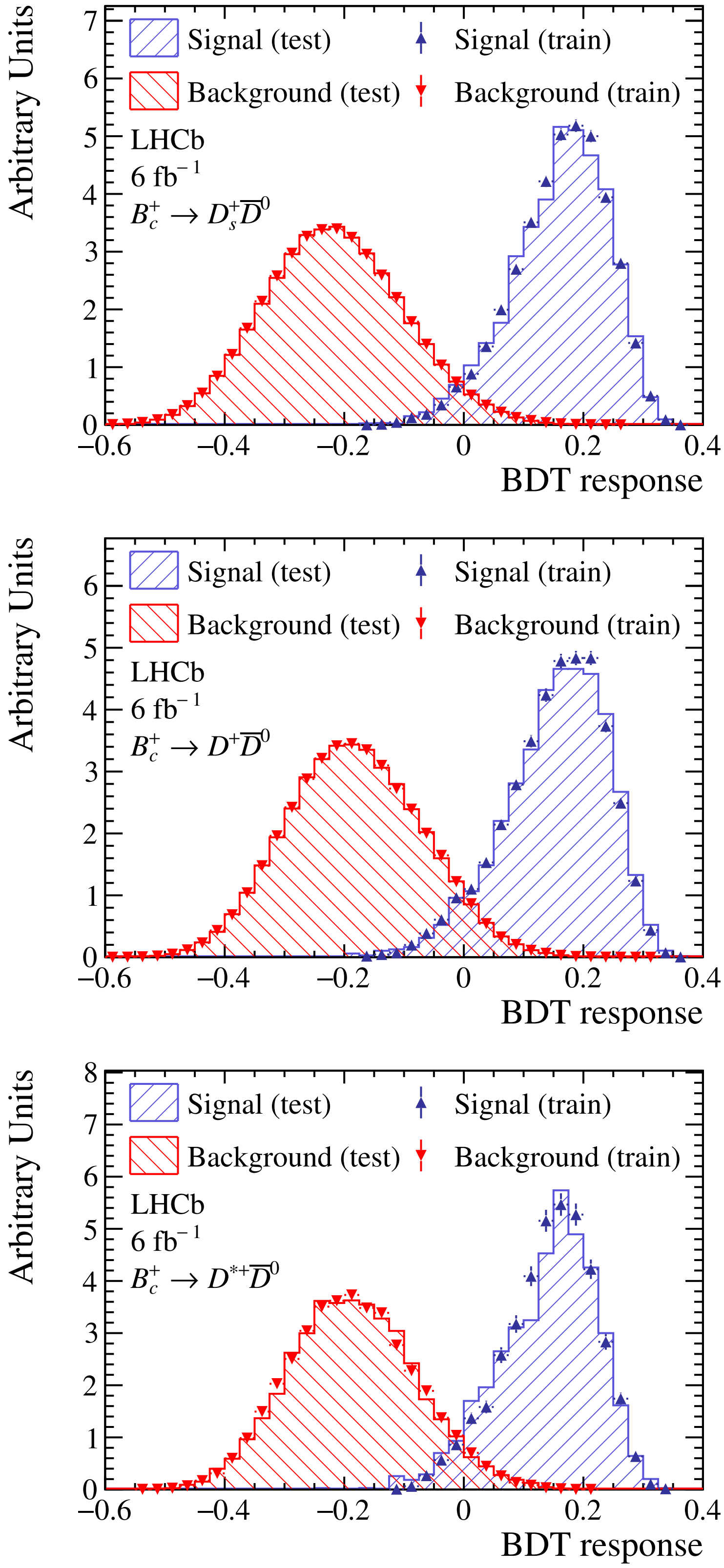

Figure 5.12: BDT response in Run 2 signal and background samples in the (first row) $B_c^+ \to D_s^+\overline{D}^0$ (second row) $B_c^+ \to D^+\overline{D}^0$ and (third row) $B_c^+ \to D^{*+}\overline{D}^0$ channels with $D^0 \to K^-\pi^+$ [1].

5.8.2 Optimisation of the selection

The two analyses, a search for the very rare $B_c^+ \to DD$ decays and a measurement of the high-yield $B^+ \to DD$ decays, require very different selection strategies. The following section describes how the response to the BDT is used to further reduce the level of background in the data in both cases.

Optimisation of the selection of $B_c^+ \to DD$ decays

The statistical power of the search can be improved by splitting the data into subsamples according to BDT response and performing a fit in each subsample while sharing certain parameters including the signal strength between subsamples. Although the sensitivity is correlated with the number of subsamples, the complexity of the fit also increases as the data is more finely divided. Both the statistical sensitivity and fit complexity must be considered when defining the subsamples. To assess the optimum divisions, the expected 90 and 95% confidence limits on several branching fractions are obtained using three different schemes.

In scheme A, the lowest purity data according to BDT response containing 30(10)% of signal is discarded for the $D_{(s)}^{+}\overset{(-)}{D}{}^{0}$ ($D^{*+}\overset{(-)}{D}{}^{0}$) channels. The remaining data is divided into three subsamples with low, medium, and high purity data, containing 20%, 20%, and 30(50)% of signal, respectively for the $D_{(s)}^{+}\overset{(-)}{D}{}^{0}$ ($D^{*+}\overset{(-)}{D}{}^{0}$) channels. Less data is discarded for the $D^{*+}\overset{(-)}{D}{}^{0}$ channels where the background is already lower. In scheme B, the lowest purity data containing 10% of signal is discarded, and the remaining data is divided into nine subsamples, each containing 10% of the signal. This scheme effectively represents the highest sensitivity that can be achieved using this method with a very complex fit. Scheme C is a simple cut rejecting the lowest purity data containing 40% of the signal, and treating the remainder of the data together. This choice of threshold is taken from the search for $B_c^+ \to DD$ decays in Run 1 LHCb data [171]. The expected limits at 90(95)% confidence level (C.L.) using each scheme are given in Table 5.11. This illustrates that a large improvement in sensitivity is achieved using three bins compared to a simple cut, and only a small improvement in sensitivity is achieved by extending this to nine bins. Therefore the three bin scheme A is chosen to achieve high sensitivity while avoiding an unecessarily complex fit.

Optimisation of the selection of $B^+ \to DD$ decays

Since $B^+ \to DD$ decays already have a high signal-background ratio, little sensitivity is gained by dividing the data into subsamples when analysing these decays. Instead, data with BDT response below a threshold is discarded, where the threshold is chosen to maximize the *significance* (S) of the fully reconstructed signal,

$$S = \frac{N_S}{\sqrt{N_S + N_B}}, \tag{5.11}$$

where N_S is the signal yield within $\pm 40\,\mathrm{MeV}/c^2$ of $m(B^-)$. To avoid sensitivity to fluctuations in the BDT response distribution of signal candidates in data, $N_S = \epsilon_S N_{S,0}$, where ϵ_S is the signal efficiency from simulation, and $N_{S,0}$ is the signal yield with no requirement on the BDT response. N_B is the background yield within $\pm 20\,\mathrm{MeV}/c^2$ of $m(B^-)$. It is obtained from extrapolating the fit of an exponential function to the high-mass sideband $[5350, 6200]\,\mathrm{MeV}/c^2$.

The dependence of ε and S on the BDT response cut for the $\mathcal{A}^{CP}$ BDT is shown in Figure 5.13, and the values at the optimal cut for both BDTs are given in Tables 5.12-5.13. The branching fraction BDT perfoms only slightly worse than the $\mathcal{A}^{CP}$ BDT.

5.8.3 Charge dependence

To verify that the BDT does not introduce a charge asymmetry, the asymmetry in the efficiency of a BDT requirement is assessed. In Figure 5.14, the charge asymmetry in the BDT response distribution after normalising the data so that there is no overall charge asymmetry, is plotted. The charge asymmetries induced by the BDT requirements are shown in Table 5.14 and there is no evidence of a non-zero asymmetry in any dataset.

5.9 Multiple candidates

Some events in a selected dataset will contain *multiple candidates*. For rare processes (on average much less than one candidate per pp collision) like those studied with the LHCb detector, the probability of more than one true signal process per event is negligible. However, $B^0 \to D^{*+} D^{*-}$ decays reconstructed as $D^{*+}\bar{D}^0$ can produce two genuine signal candidates differing only in the D^0 and $\bar{D}^0$ assignment and reconstructed charge. Another

Table 5.11: Expected limits at 90(95)% C.L. assuming statistical uncertainties only, in Run 2 data. Limits are evaluated in each of three division schemes, which are defined in the text.

Parameter of interest	Scheme A	Scheme B	Scheme C
$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^+ \bar{D}^0)}{\mathcal{B}(B^+ \to D^+ \bar{D}^0)}$	$3.1\,(3.8) \times 10^{-3}$	$2.8\,(3.5) \times 10^{-3}$	$4.0\,(5.0) \times 10^{-3}$
$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^+ \bar{D}^0)}{\mathcal{B}(B^+ \to D_s^+ \bar{D}^0)}$	$1.4\,(1.8) \times 10^{-4}$	$1.4\,(1.8) \times 10^{-4}$	$2.5\,(3.0) \times 10^{-4}$
$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^{*+} \bar{D}^0)}{\mathcal{B}(B^+ \to D^{*+} \bar{D}^0)}$	$4.7\,(6.0) \times 10^{-3}$	$4.5\,(5.7) \times 10^{-3}$	$6.2\,(7.8) \times 10^{-3}$

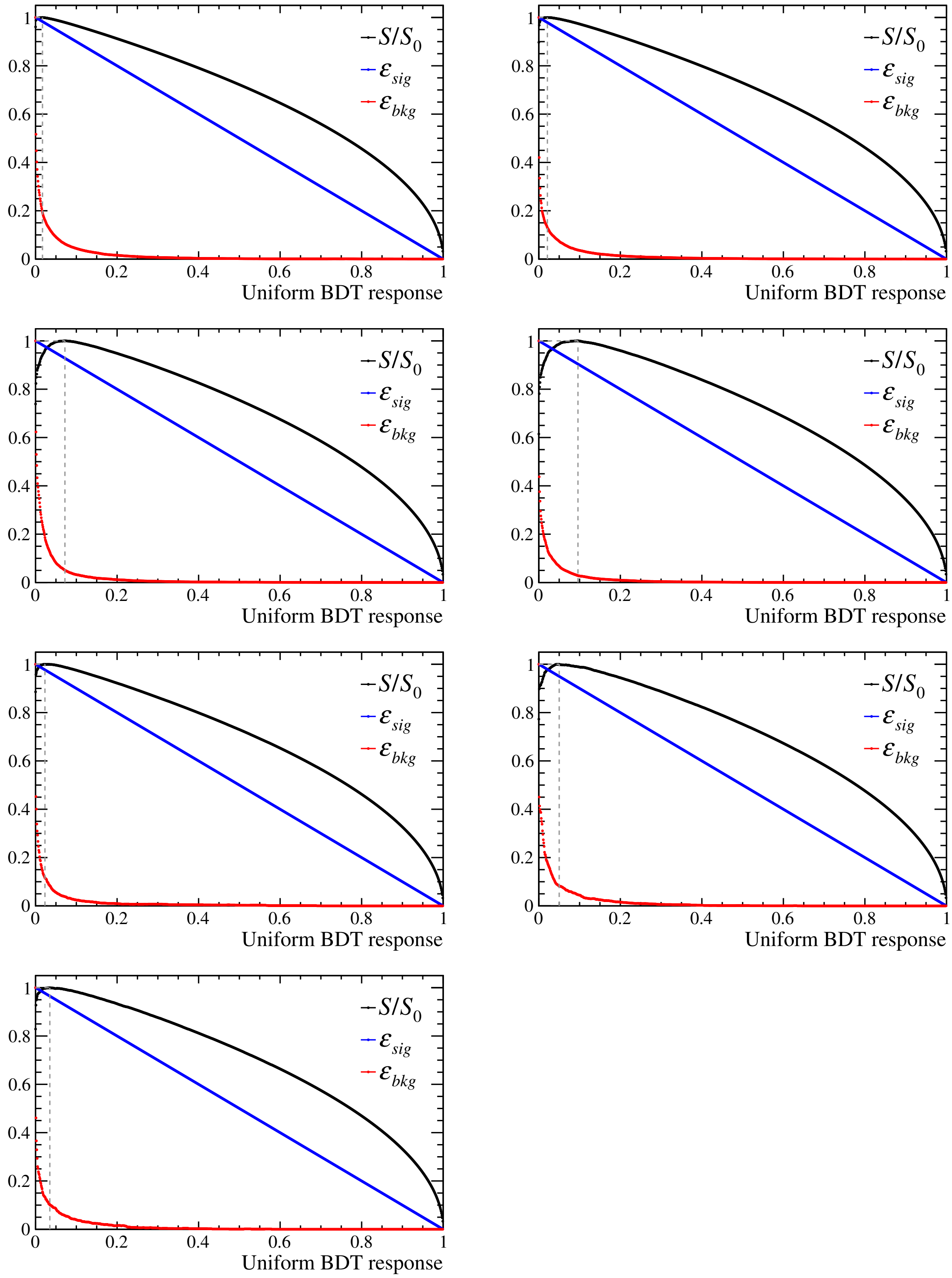

Figure 5.13: Dependence of the signal efficiency ε_{sig}, background retention ε_{bkg}, and the increase in significance S/S_0 on the cut value of the $\mathcal{A}^{CP}$ BDT for Run2 data. The dashed line shows the cut value that optimises S. First row are $D_s^- D^0$ final states, second row $D^- D^0$, third row $D^{*-} D^0$ with $D^0 \to K\pi$, fourth row $D^{*-} D^0$ with $D^0 \to K3\pi$. Left are $D^0 \to K^-\pi^+$, right are $D^0 \to K^-\pi^+\pi^-\pi^+$.

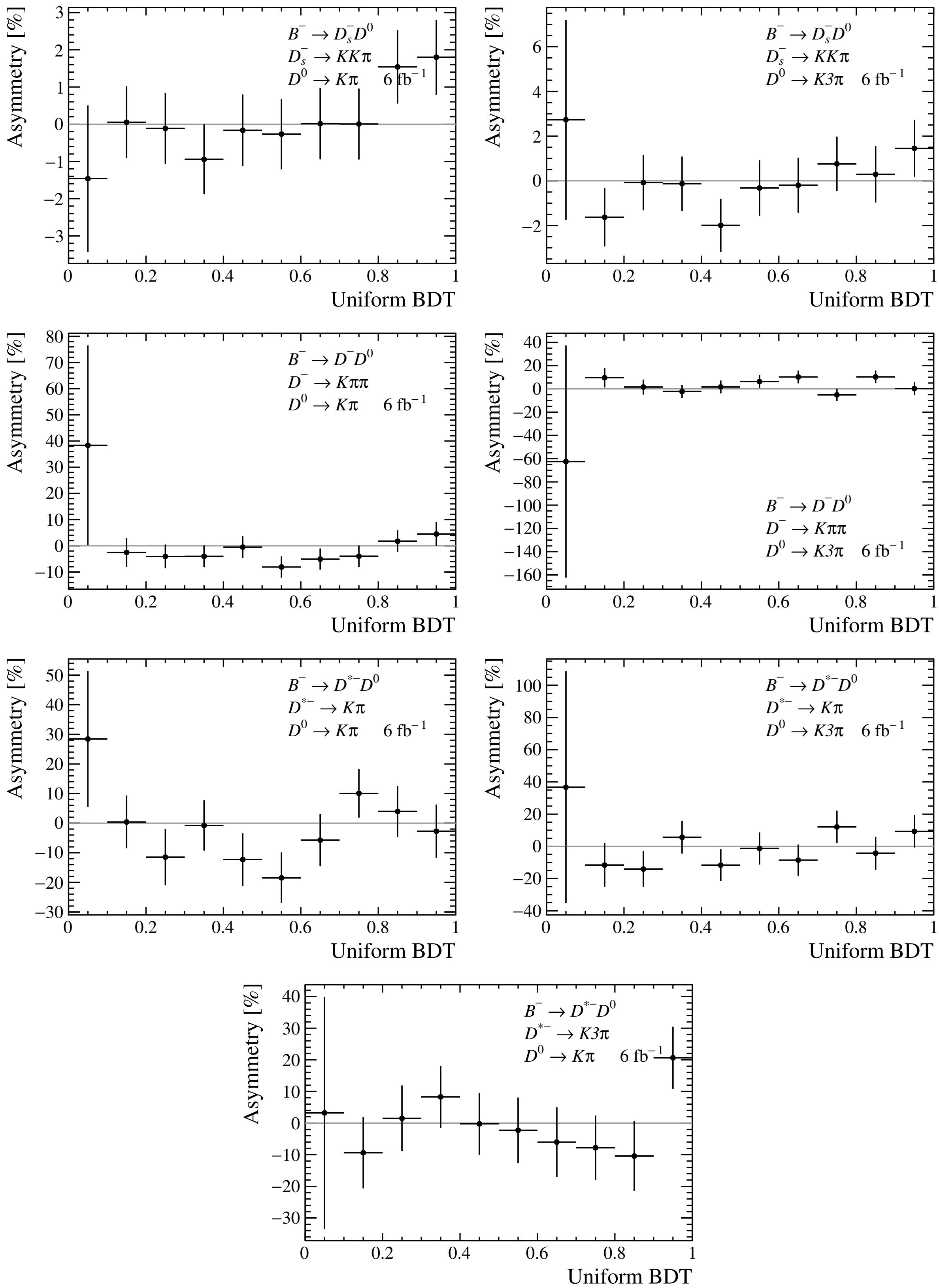

Figure 5.14: Charge asymmetry of the BDT response for background-subtracted signal in Run 2 data, normalised so that the overall asymmetry is zero.

Table 5.12: Signal efficiency (ϵ_S), background retention (ϵ_B) and relative significance compared to the significance under no BDT cut (S/S_0) at the optimal cut of the $\mathcal{A}^{CP}$ BDT.

Period	Final State	ϵ_S	ϵ_B	S/S_0
	$KK\pi, K\pi$	0.98	0.15	1.09
	$KK\pi, K3\pi$	0.97	0.11	1.15
	$K\pi\pi, K\pi$	0.90	0.03	1.62
Run 1	$K\pi\pi, K3\pi$	0.87	0.02	2.08
	$K\pi\pi_s, K\pi$	0.95	0.09	1.19
	$K\pi\pi_s, K3\pi$	0.93	0.06	1.40
	$K3\pi\pi_s, K\pi$	0.95	0.08	1.31
	$KK\pi, K\pi$	0.98	0.19	1.04
	$KK\pi, K3\pi$	0.98	0.13	1.11
	$K\pi\pi, K\pi$	0.93	0.05	1.35
Run 2	$K\pi\pi, K3\pi$	0.90	0.03	1.63
	$K\pi\pi_s, K\pi$	0.98	0.12	1.13
	$K\pi\pi_s, K3\pi$	0.95	0.08	1.29
	$K3\pi\pi_s, K\pi$	0.96	0.10	1.21

source of multiple candidates is through different PID hypothesis assignments, and the least probable candidates can be vetoed, such as in Section 5.7. After applying these vetoes, in general most multiple candidates are formed by combining a subset of the tracks in one candidate with other tracks in the event to form another candidate.

Table 5.15 shows the fraction of events passing all selection for the $\mathcal{A}^{CP}$ measurement and in the signal region that contain multiple candidates, and the composition of these multiple candidates. As previously discussed, $B^0 \to D^{*+}D^{*-}$ decays result in a large fraction of candidates that differ only in the D^0 and $\overline{D}^0$ assignment in the $D^{*+}\overline{D}^0$ channels. Otherwise, the majority of multiple candidates differ from the other candidates in the event by only one track, and no more than 2.5% of events contain multiple candidates overall. In the data selected for the $B_c^+ \to DD$ search, the majority of multiple candidates again differ from the other candidates in the event by only one track, and no more than 4.7% of events contain multiple candidates.

When measuring a branching fraction, rejecting all but one of the multiple candidates helps to avoid undercoverage in confidence intervals, and also reduces the backgroud level. Therefore, in the branching fraction measurements discussed in this thesis, only one candidate per event is retained, and this candidate is chosen at random after all selection has been applied. For the $B_c^+ \to DD$ search, the BDT samples above the minimum threshold for analysis are combined when analysing and rejecting multiple candidates.

In contrast, the removal of multiple candidates will bias an $\mathcal{A}^{CP}$ measurement [194] since multiple candidates are less likely for the charge with the lower yield. Therefore

Table 5.13: Signal efficiency (ϵ_S), background retention (ϵ_B) and relative significance compared to the significance under no BDT cut (S/S_0) at the optimal cut of the branching fraction BDT.

Period	Final State	ϵ_S	ϵ_B	S/S_0
Run 1	$KK\pi, K\pi$	0.98	0.37	1.06
	$KK\pi, K3\pi$	0.96	0.22	1.12
	$K\pi\pi, K\pi$	0.91	0.06	1.58
	$K\pi\pi, K3\pi$	0.89	0.04	1.98
	$K\pi\pi_s, K\pi$	0.97	0.16	1.17
	$K\pi\pi_s, K3\pi$	0.88	0.08	1.34
	$K3\pi\pi_s, K\pi$	0.93	0.16	1.24
Run 2	$KK\pi, K\pi$	0.99	0.43	1.03
	$KK\pi, K3\pi$	0.97	0.23	1.09
	$K\pi\pi, K\pi$	0.93	0.08	1.33
	$K\pi\pi, K3\pi$	0.91	0.06	1.57
	$K\pi\pi_s, K\pi$	0.96	0.13	1.12
	$K\pi\pi_s, K3\pi$	0.93	0.11	1.26
	$K3\pi\pi_s, K\pi$	0.94	0.15	1.18

multiple candidates are not rejected in the $\mathcal{A}^{CP}$ measurement. This causes small undercoverage and slightly increases the background level, but both of these effects are negligible given that the fraction of multiple candidates is small, which is seen to be the case in Table 5.15. Note that the presence of multiple candidates in the $B^0 \to D^{*+}D^{*-}$ background instead causes overcoverage in measurements of the $B^- \to D^{*-}D^{*0}$ signal.

Table 5.14: Charge asymmetry (A_{BDT}) of the BDT requirement for $B^{\mp}$ candidates in data.

Final state	$A_{\mathrm{BDT}}/\%$ Run 1	$A_{\mathrm{BDT}}/\%$ Run 2
$KK\pi$, $K\pi$	-0.31 ± 0.29	-0.15 ± 0.11
$KK\pi$, $K3\pi$	0.9 ± 0.5	0.33 ± 0.22
$K\pi\pi$, $K\pi$	-4 ± 4	1.3 ± 1.4
$K\pi\pi$, $K3\pi$	-1 ± 8	-2.3 ± 2.8
$K\pi$, $K\pi$	3 ± 4	1.7 ± 1.5
$K\pi$, $K3\pi$	2 ± 7	-0.5 ± 3.1
$K3\pi$, $K\pi$	-2 ± 7	-0.3 ± 2.3

Table 5.15: Fraction f of events containing multiple candidates in data selected for the $\mathcal{A}^{CP}$ measurements split by channel, final state and between Run 1 and 2. All candidates counted here have passed preselection cuts; have a BDT response greater than the minimum threshold for use in the analysis; and are within the $4870 < m(B^+) < 5400\,\text{MeV}/c^2$ and $|m - m_D| < 25\,\text{MeV}/c^2$ mass windows that constitute the signal region. The fraction of events with multiple candidates in the same final state f_{FS}; the fraction with multiple candidates that differ by no more than one track $f_{\Delta=1trk}$; and the fraction that differ only in the D^0 and $\bar{D}^0$ assignment $f_{D^0\leftrightarrow\bar{D}^0}$ are also shown.

Period	Channel	Final state	f	f_{FS}	$f_{\Delta=1trk}$	$f_{D^0\leftrightarrow\bar{D}^0}$
Run 1	$B^+ \to D_s^+\bar{D}^0$	$K\pi$	0.5%	0.5%	0.4%	0.0%
	$B^+ \to D_s^+\bar{D}^0$	$K3\pi$	2.0%	2.0%	1.3%	0.0%
	$B^+ \to D^+\bar{D}^0$	$K\pi$	0.8%	0.3%	0.4%	0.0%
	$B^+ \to D^+\bar{D}^0$	$K3\pi$	1.7%	1.3%	1.0%	0.0%
	$B^+ \to D^{*+}\bar{D}^0$	$K\pi, K\pi$	7.5%	7.5%	0.6%	7.0%
	$B^+ \to D^{*+}\bar{D}^0$	$K3\pi, K\pi$	13.7%	1.6%	1.1%	11.7%
	$B^+ \to D^{*+}\bar{D}^0$	$K\pi, K3\pi$	15.3%	2.8%	1.7%	12.1%
Run 2	$B^+ \to D_s^+\bar{D}^0$	$K\pi$	0.6%	0.6%	0.5%	0.0%
	$B^+ \to D_s^+\bar{D}^0$	$K3\pi$	2.5%	2.5%	1.7%	0.0%
	$B^+ \to D^+\bar{D}^0$	$K\pi$	0.7%	0.3%	0.6%	0.0%
	$B^+ \to D^+\bar{D}^0$	$K3\pi$	2.1%	1.7%	1.4%	0.0%
	$B^+ \to D^{*+}\bar{D}^0$	$K\pi, K\pi$	8.8%	8.8%	0.7%	8.2%
	$B^+ \to D^{*+}\bar{D}^0$	$K3\pi, K\pi$	16.8%	2.6%	1.9%	14.1%
	$B^+ \to D^{*+}\bar{D}^0$	$K\pi, K3\pi$	17.7%	3.1%	2.1%	14.9%

CHAPTER 6

Modelling of signal and background decays

The raw yields required to evaluate branching fractions and *CP* asymmetries are determined with best precision by fitting a probability density function (PDF) to the invariant mass distribution of the beauty meson candidates in data. This chapter will describe the contribution to the PDF from each signal, normalisation, and background component. The PDFs used to model the invariant mass distribution of each component are summarised in Table 6.1, and the remainder of this chapter will describe the contribution to the overall PDF from each component in more detail. Many of these PDFs are informed by the invariant mass distributions in simulated decays.

Table 6.1: Summary of PDFs used to describe the invariant mass distribution of each component.

Component	Model	Section
Fully reconstructed $B^+_{(c)} \to DD$	Gaussian function plus double-sided Crystal Ball function	6.1
$B^+, B^0_{(s)} \to D^* D^{(*)}$ with one missing particle	Parabola function corrected for resolution and selection effects	6.2.1
$B^+, B^0_{(s)} \to D^* D^*$ with two missing particles	Kernel Density Estimate	6.2.2
Partially reconstructed $B^+_c \to D^* D^{(*)}$	Kernel Density Estimate	6.2.3
Partially reconstructed $B \to D^{**} D$	Kernel Density Estimate	6.3
Single charm background	Double-sided Crystal Ball function	6.4
$\overline{D}^0 \to D^0$ cross-feed	Gaussian function plus double-sided Crystal Ball function	6.5.1
$D^+_s \leftrightarrow D^+$ cross-feed	Sum of two Gaussian functions	6.5.2
Combinatorial background	Exponential or exponential plus constant	6.6

6.1 Fully reconstructed $B^+_{(c)} \to DD$ decays

The invariant mass distribution of fully reconstructed $B^+_{(c)} \to D^+_s \overline{D}^0$, $D^+\overline{D}^0$ and $D^{*+}\overline{D}^0$ decays are described by the sum of a narrow Gaussian (Gaus) and a broader double-sided Crystal Ball (DSCB) function. A DSCB function [195] has a Gaussian core that becomes power law tails of power n_i at distance $\alpha_i\sigma$ from the peak. The tails account for Bremstrahlung radiation and inaccurate reconstruction of momenta. Its functional form is

$$\mathrm{DSCB}(m; m_0, \sigma, \alpha_1, n_1, \alpha_2, n_2) \propto \begin{cases} A_1\left(B_1 - \frac{m-m_0}{\sigma}\right)^{-n_1} & \text{for } m - m_0 < -\alpha_1\sigma \\ \exp\left(-\frac{1}{2}\left(\frac{m-m_0}{\sigma}\right)^2\right) & \text{for } -\alpha_1\sigma \leq m - m_0 \leq \alpha_2\sigma \\ A_2\left(B_2 + \frac{m-m_0}{\sigma}\right)^{-n_2} & \text{otherwise} \end{cases} \tag{6.1}$$

where

$$A_i = \left(\frac{n_i}{\alpha_i}\right)^{n_i} \exp\left(-\frac{\alpha_i^2}{2}\right), \tag{6.2}$$

and

$$B_i = \frac{n_i}{\alpha_i} - \alpha_i. \tag{6.3}$$

The sum $f(m)$ is written as

$$f(m) = f_{\mathrm{Gaus}}\mathrm{Gaus}(m; m_0, \kappa\sigma) + (1 - f_{\mathrm{Gaus}})\mathrm{DSCB}(m; m_0, \sigma, \alpha_1, n_1, \alpha_2, n_2), \tag{6.4}$$

$$\mathrm{Gaus}(m; m_0, \sigma) \propto \exp\left(-\frac{1}{2}\left(\frac{m-m_0}{\sigma}\right)^2\right), \tag{6.5}$$

where f_{Gaus} is the fraction of the integral from the Gaussian. The peak position m_0 is shared between the Gaussian and the DSCB. The width of the Gaussian is related to the width of the DSCB core by a scale factor $\kappa < 1$.

The mass resolution is better for higher values of the BDT reponse, so in the search for $B^+_c \to DD$ decays, the width σ is a linear function of the mean BDT response of signal in the data sample ($\overline{\mathrm{BDT}}$),

$$\sigma = \sigma_0 - \sigma_{\mathrm{BDT}}\overline{\mathrm{BDT}}. \tag{6.6}$$

The fit to the invariant mass distribution of $B^+_c \to D^+_s \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ candidates in each BDT sample is shown in Figure 6.1, which demonstrates the correlation between BDT response and mass resolution. Fits of this model to the invariant mass distribution of simulated $B^+ \to DD$ candidates passing selection for the $\mathcal{A}^{CP}$ measurement in different final states are shown in Figure 6.2.

The statistical power of the dataset is only sufficient to constrain $m_0(B^+)$ and $\sigma_0(B^+)$, which are therefore free parameters when the model is fitted to data. The parameters κ, f_{Gaus}, α_i, n_i, and σ_{BDT} are fixed to values obtained from fitting the model to simulation. The

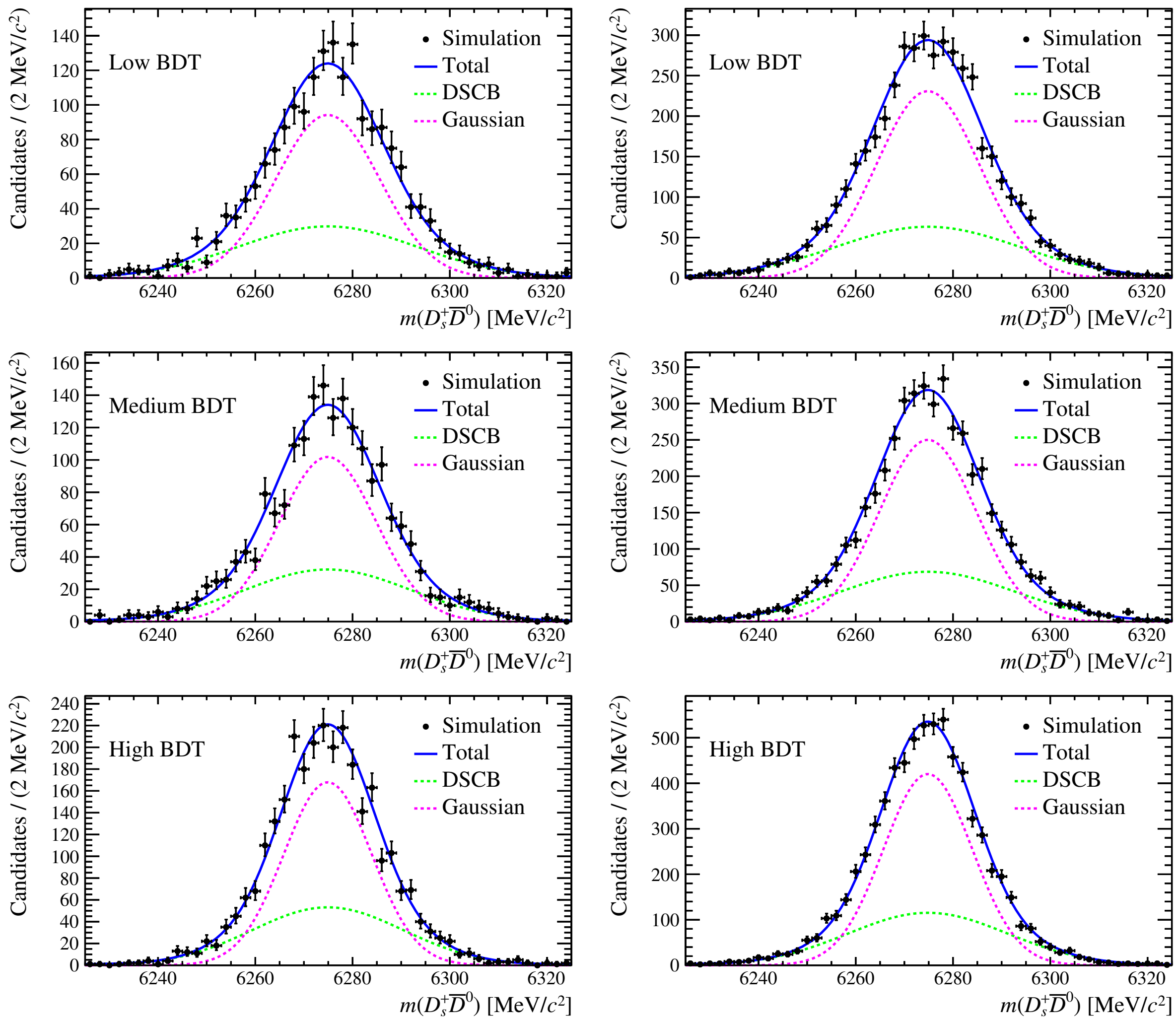

Figure 6.1: Invariant mass distribution of $B_c^+ \to D_s^+ \overline{D}^0$, $D^0 \to K^- \pi^+$ candidates in (left) Run 1 and (right) Run 2 simulation, in the (top) low, (middle) medium, and (bottom) high BDT samples. The fitted BDT-dependent Gaussian plus DCSB model described in the text is also drawn.

ratio of σ_0 between data and simulation is required to be the same for B_c^+ and $B^+ \to DD$ decays,

$$\frac{\sigma_0(B_c^+, \text{data})}{\sigma_0(B_c^+, \text{MC})} = \frac{\sigma_0(B^+, \text{data})}{\sigma_0(B^+, \text{MC})}, \tag{6.7}$$

and the difference between the peak positions for B^+ and B_c^+ signal is fixed to the known mass difference [22, 70],

$$m_0(B_c^+) - m_0(B^+) = m_{B_c^+} - m_{B^+}. \tag{6.8}$$

In general, $m_0(B^+)$ is mismodelled in simulation by up to $1\,\text{MeV}/c^2$ and $\sigma_0(B^+)$ by up to 3%. Exemplar fits to data in the mass range near m_{B^+}, dominated by this model but also including single charm (Section 6.4) and combinatorial (Section 6.6) backgrounds, are shown in Figure 6.3.

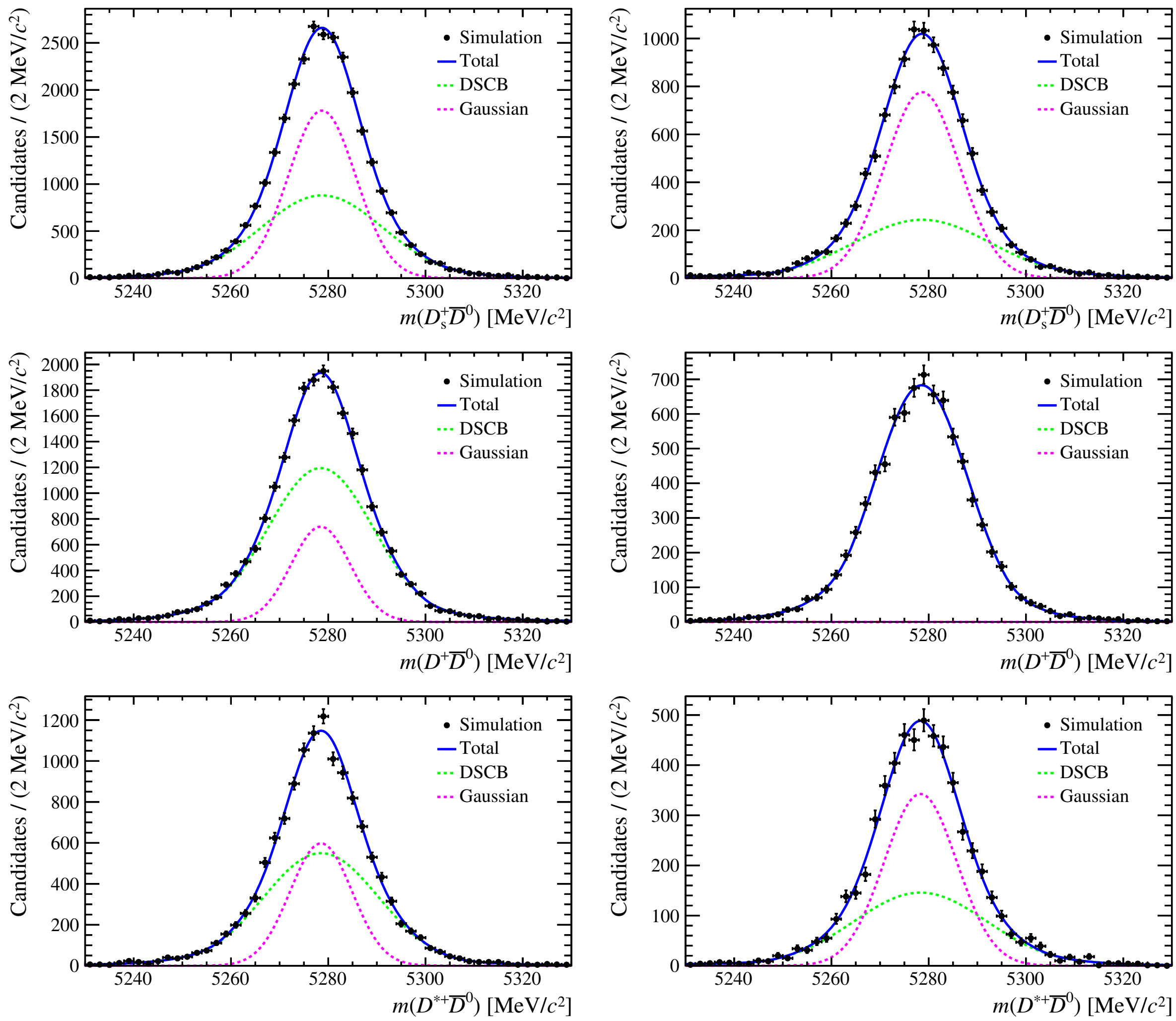

Figure 6.2: Invariant mass distribution of (top) $B^+ \to D_s^+ \overline{D}^0$, (middle) $D^+\overline{D}^0$, and (bottom) $D^{*+}\overline{D}^0$ candidates with (left) $\overline{D}^0 \to K^+\pi^-$ and (right) $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ in Run 2 simulation passing selection for the $\mathcal{A}^{CP}$ measurement. The fitted Gaussian plus DCSB model described in the text is also drawn.

6.2 Partially reconstructed $B \to D^* D^{(*)}$ decays

Signal decays with one or two missing photons or pions have broader and more complicated reconstructed mass distributions at lower mass. Different methods are used to model B_c^+, B^+, B^0 and B_s^0 decays, and decays with one or two missing particles. The appropriate method depends on the importance of an accurate model, the possibility of modelling the shape with an analytic function, and the availability of simulation.

The invariant mass distribution depends on the number and species of the missing particles, and the polarisation of the charm mesons in $B \to D^* D^*$ decays. These contributions are modelled separately, and weighted in the model by the $D^* \to DX^0$ branching fractions [22] and the fraction of longitudinally polarised $B \to D^* D^*$ decays,

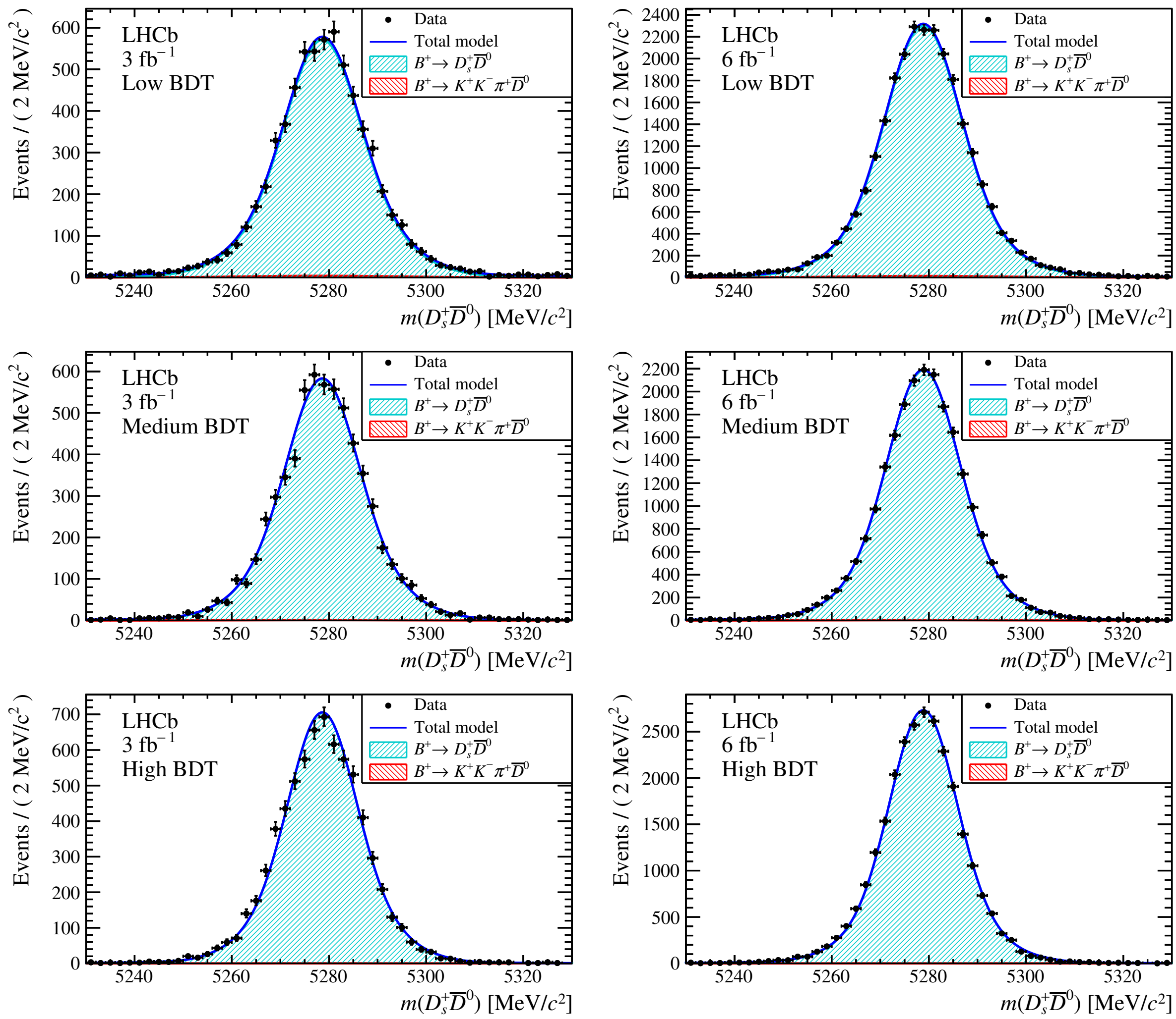

Figure 6.3: Selected $B^+ \to D_s^+ \overline{D}^0$, $D^0 \to K^- \pi^+$ candidates. For the (top) low, (middle) medium, and (bottom) high BDT sample, in (left) Run 1 data and (right) Run 2 data [1].

known as the polarisation fraction (f_L). Sections 6.2.1-6.2.3 explain how each contribution is modelled under different circumstances. Section 6.2.4 discusses the treatment of the polarisation fractions. The precision on $\mathcal{A}^{CP}$ measurements can be improved by constraining the yields of B_s^0 relative to B^0 decays, which is described in Section 6.2.5.

6.2.1 $B \to DD$ decays with one missing particle

$B^- \to D^* D^{(*)}$ decays with one missing particle form a high-yield signal component in the CP asymmetry measurement. To ensure that the fitted model accurately describes these high-yield signal decays, the reconstructed mass is described using an analytic function with free parameters. $B^0_{(s)} \to D^{(*)-}_{(s)} D^{*+}$ decays with $D^{*+} \to D^0 \pi^+$ where the π^+ is not reconstructed form a peaking background and can be modelled in the same way. The true visible mass is a parabola. This parabola is multiplied by a mass-dependent reconstruction and selection efficiency function, and the product is convolved with a resolution function [196]. Each of these elements will be described in the remainder of this section.

True visible mass distribution

In a $B \to D^*D$ decay or a longitudinally polarised $B \to D^*D^*$ decay, the D^* has zero helicity. Conservation of angular momentum and parity requires that if the D^* decays into a D and a pion, the differential decay rate is

$$\frac{d\Gamma}{d\cos\theta} \propto \cos^2\theta, \tag{6.9}$$

where θ is the angle between the pion momentum in the D^* rest frame and the D^* momentum in the B rest frame. This means that the pion is most likely to be produced parallel or antiparallel to the D^* momentum. If the D^* instead decays into a D and a photon,

$$\frac{d\Gamma}{d\cos\theta} \propto \sin^2\theta, \tag{6.10}$$

so the photon is most likely to be produced perpendicular to the D^* momentum. In a transversely polarised $B \to D^*D^*$ decay, the D^* has helicity $\lambda = \pm 1$ and the angular distribution is identical to Equation 6.10 for a $D^* \to D\pi$ decay. For a $D^* \to D\gamma$ decay,

$$\frac{d\Gamma}{d\cos\theta} \propto 1 + \cos^2\theta, \tag{6.11}$$

so that the photon is most likely to be produced parallel or antiparallel to the D^* momentum, but the miminum is shallower than in the other angular distributions [197, 198].

The visible mass $m(D_1D_2)$ for the decay chain $B \to D_1(D^* \to D_2X^0)$ depends only on $\cos\theta$ as

$$m(D_1D_2)^2 = m_{D_1}^2 + m_{D_2}^2 + 2E_1^{'}E_2^{'} - 2p_1^{'}p_2^{'}\cos\theta, \tag{6.12}$$

where $E_i^{'}$ and $p_i^{'}$ are the energy and momentum of D_i in the rest frame of the D^*. Therefore the decay rate as a function of the visible mass is

$$\frac{d\Gamma}{dm} \propto \begin{cases} \left(m - \frac{a+b}{2}\right)^2 & \text{for } a < m < b \\ 0 & \text{otherwise} \end{cases} \tag{6.13}$$

neglecting small terms of $\mathcal{O}((m - \frac{a+b}{2})^3) \ll 1$, where

$$a^2 = m_{D_1}^2 + m_{D_2}^2 + 2E_1^{'}E_2^{'} - 2p_1^{'}p_2^{'}, \tag{6.14}$$

$$b^2 = m_{D_1}^2 + m_{D_2}^2 + 2E_1^{'}E_2^{'} + 2p_1^{'}p_2^{'}. \tag{6.15}$$

In practice a and b are shifted by the same value, ΔM, that the fully reconstructed peak is shifted from m_{B^+}, primarily due to momentum miscalibration. The angular distribution in Equation 6.10 leads to the mass distribution

$$\frac{d\Gamma}{dm} \propto \begin{cases} -(m-a)(m-b) & \text{for } a < m < b \\ 0 & \text{otherwise} \end{cases}, \tag{6.16}$$

and the angular distribution in Equation 6.11 leads to the mass distribution

$$\frac{d\Gamma}{dm} \propto \begin{cases} \left(m - \frac{a+b}{2}\right)^2 + \left(\frac{a-b}{2}\right)^2 & \text{for } a < m < b \\ 0 & \text{otherwise} \end{cases}. \tag{6.17}$$

Mass-dependence of efficiency

Reconstructed final state particles carry less momentum when the reconstructed invariant mass is smaller, so the reconstruction and selection efficiency is correlated with the reconstructed mass. This is modelled as a linear function of m as

$$\varepsilon(m) \propto \frac{1-\xi}{b-a} m + \frac{b\xi - a}{b-a}, \tag{6.18}$$

where $0 < \xi < 1$ is the relative efficiency between $m = a$ and $m = b$. The relative efficiency gradient, $\frac{d\varepsilon}{dm}/\varepsilon(m = m_{B^+})$, is shared between all decays with one missing particle in a given final state. This is enforced by writing ξ as

$$\xi = \frac{\varepsilon(m = m_{B^+}) + (a - m_{B^+}) d\varepsilon/dm}{\varepsilon(m = m_{B^+}) + (b - m_{B^+}) d\varepsilon/dm}. \tag{6.19}$$

Resolution

The resolution function is a quadruple Gaussian that consists of a double Gaussian core with tails on either side,

$$\begin{aligned} R(m) = {} & f\,\text{Gaus}(m; m_0, \kappa\sigma) + (1 - f - f_1 - f_2)\text{Gaus}(m; m_0, \sigma) \\ & + f_1 \text{Gaus}(m; m_0 - s_\mu \sigma, s_\sigma \sigma) + f_2 \text{Gaus}(m; m_0 + s_\mu \sigma, s_\sigma \sigma) \end{aligned}. \tag{6.20}$$

The core parameters m_0, κ, σ, and $f = f_{\text{Gaus}}$ are fixed to their values from the Gaussian plus DSCB fit to fully reconstructed simulation in the previous section. f_1, f_2, s_μ, and s_σ are obtained from fits of the quadruple Gaussian to fully reconstructed simulation, as shown in Figure 6.4.

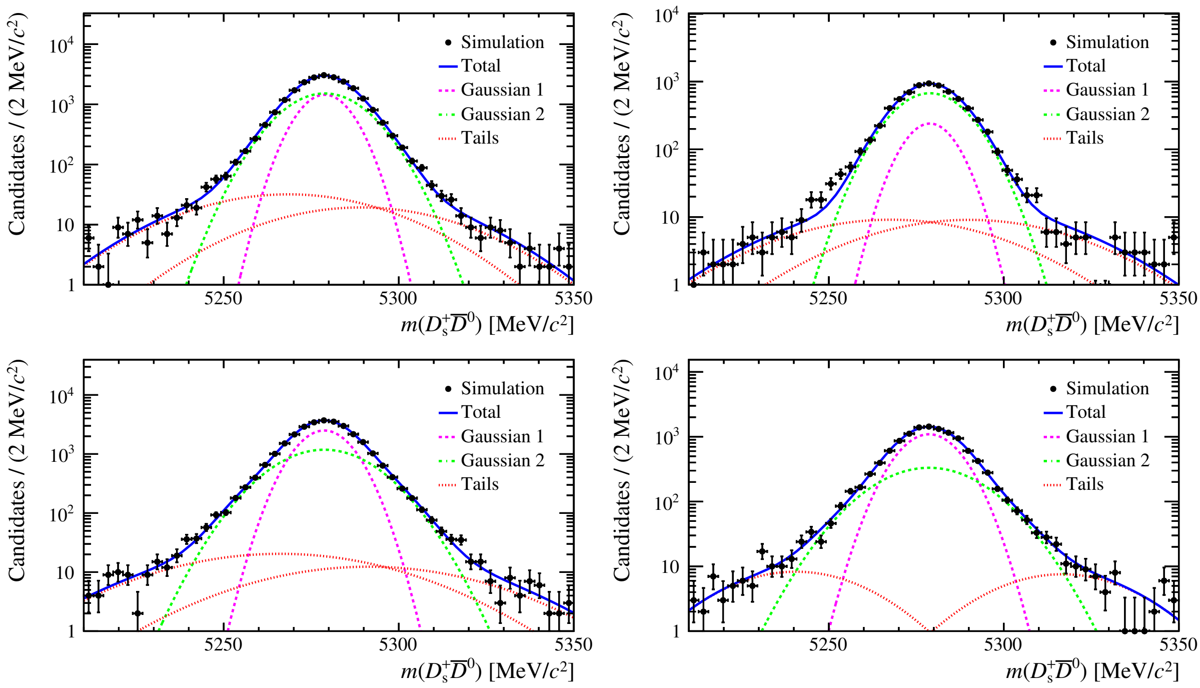

Figure 6.4: Fits of the quadruple Gaussian resolution function $R(m)$ (Equation 6.20) to the invariant mass distributions of fully reconstructed $B^+ \to D_s^+\bar{D}^0$ decays. For (left) $\bar{D}^0 \to K^+\pi^-$ and (right) $\bar{D}^0 \to K^+\pi^-\pi^+\pi^-$ in (top) Run 1 and (bottom) Run 2 simulation.

6.2.2 $B \to DD$ decays with two missing particles

The reconstructed mass distribution of $B^-, B^0, B_s^0 \to D^*D^*$ decays with two missing particles cannot be straightforwardly described with an analytic function. Since these decays do not form the signal in this analysis and only marginally overlap with signal decays, their accurate modelling is also not critical: a large change in the shape causes only a small change in $\mathcal{A}^{CP}$ (Section 8.4). Therefore these decays are modelled as a *kernel density estimate* (KDE) [199] of the reconstructed mass in generator-level simulation produced as described in Section 5.2. Examplar fits are shown in Figure 6.5.

The efficiency again depends on the reconstructed mass, and this effect is not fully modelled in the generator-level simulation samples. Therefore each template is again multiplied by a linear efficiency function, with a free relative efficiency gradient $\frac{d\varepsilon}{dm}/\varepsilon(m = m_{B^+})$ shared between all decays in a final state with two missing particles.

To ensure the stability of the fit, the relative yields of $B^- \to DD$ and $B^0 \to DD$ decays are fixed using the known branching fractions [22] and assuming equal B^- and B^0 production.

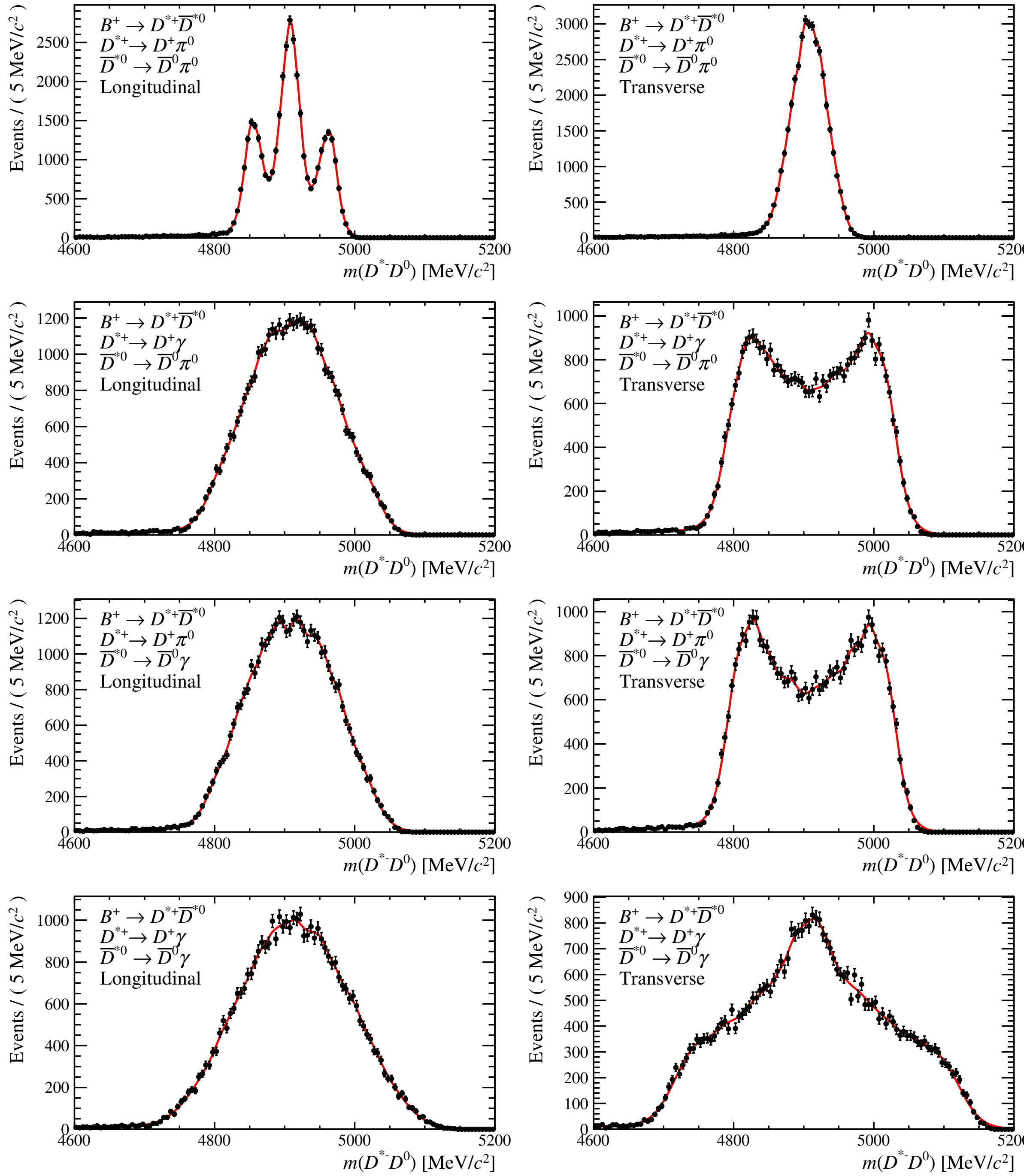

Figure 6.5: Distributions of the mass reconstructed when missing two soft neutral particles, in generator-level simulation of $B^- \to D^{*-}D^{*0}$ decays. The fitted KDE is also drawn (red). Decays involving a B^0 or a D_s^+ are similar except for small mass shifts.

Table 6.2: Values of f_L used to model B^+, B^0 and $B^0_s \to D^* D^*$ decays, and the source of these values.

Decay	Value	Source
$B^- \to D_s^{*-} \overline{D}^{*0}$	0.52 ± 0.06	[81–85]
$B^- \to D^{*-} \overline{D}^{*0}$	0.54 ± 0.09	[81, 82, 84, 85]
$\overline{B}^0 \to D_s^{*-} D^{*+}$	0.578 ± 0.015	[197]
$B^0 \to D^{*-} D^{*+}$	0.624 ± 0.031	[200]
$B_s^0 \to D_s^{*-} D^{*+}$	0.50 ± 0.09	[81, 82, 84]
$B_s^0 \to D^{*-} D^{*+}$	0.5 ± 0.5	-

6.2.3 $B_c^+ \to DD$ decays

$B_c^+ \to DD$ decays are expected to have yields of no more than a few candidates, so a model with free parameters is not required to describe the data well. Instead the model for each $B_c^+ \to DD$ is a KDE of the simulated invariant mass distribution. Examples are shown in Figure 6.6.

6.2.4 Polarisation fractions

The relative yield of longitudinally and transversely polarised $B \to D^* D^*$ decays is determined according to the polarisation fraction, f_L. Values of $f_L(\overline{B}^0 \to D_{(s)}^{*-} D^{*+})$ have been measured [197, 200], and these measurements as absolute fractions differ by up to 0.1 from theoretical predictions [81–85]. Values of $f_L(B^- \to D_{(s)}^{*-} D^{*0})$ and $f_L(B_s^0 \to D_s^{*-} D^{*+})$ have been predicted but not measured [81–85]. In this case f_L is taken as the mean of the predictions, and the uncertainty is the quadrature sum of the standard deviation and the difference between experiment and theory in B^0 decays. No measurement or prediction is available for $f_L(B_s^0 \to D^{*-} D^{*+})$, so it is assigned a value of 0.5 ± 0.5.

To avoid a strong dependence of the results on theory predictions, $f_L(B^- \to D^{*-} D^{*0})$ is treated as a free parameter in the fit to data in the $D^{*-} D^0$ channel. Otherwise, f_L is set to the values in Table 6.2.

Values of $f_L(B_c^+ \to D_{(s)}^{*+} \overline{D}^{*0})$ have been predicted [72]. Nevertheless, to avoid strong dependence of results on one theory prediction, f_L is treated as a free parameter. A full description of the treatment of f_L is given in Section 7.3.

6.2.5 $B_s^0 \to DD$ yields

The yields of B_s^0 decays are related to the corresponding B^0 yields as

$$\frac{N(B_s^0)}{N(B^0)} = \frac{f_s}{f_d} \frac{\mathcal{B}(B_s^0)}{\mathcal{B}(B^0)} \frac{\varepsilon(B_s^0)}{\varepsilon(B^0)}. \tag{6.21}$$

Values of the ratio $\varepsilon(B_s^0)/\varepsilon(B^0)$ in the range [1.0, 1.1] have been measured in multiple

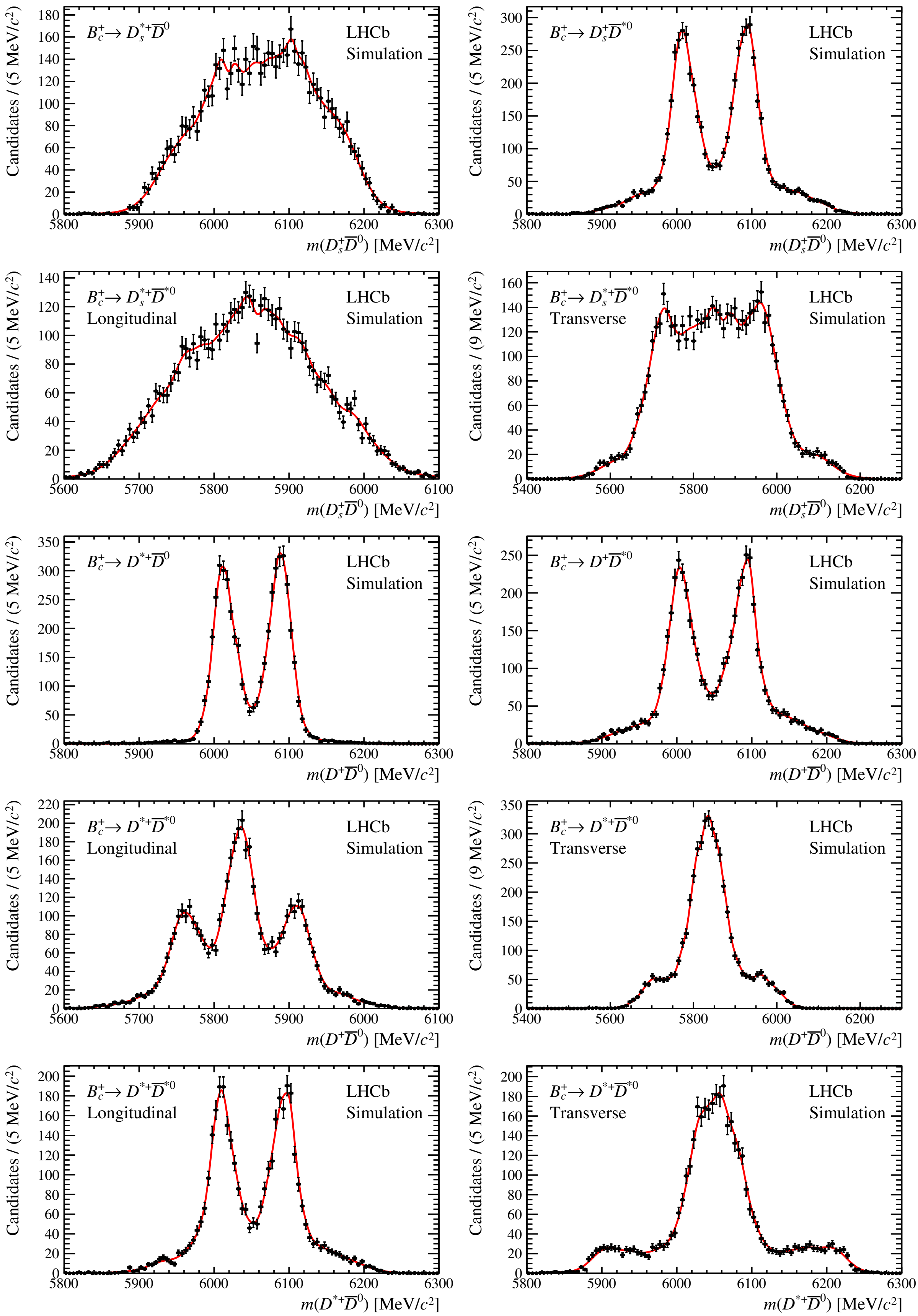

Figure 6.6: Invariant mass distributions of Run 2 simulated (first row) $B_c^+ \to D_s^{(*)+}\overline{D}^{(*)0}$ decays with one missing particle, (second row) $B_c^+ \to D_s^{*+}\overline{D}^{*0}$ decays with two missing particles, (third row) $B_c^+ \to D^{(*)+}\overline{D}^{(*)0}$ decays with one missing soft particle, (fourth row) $B_c^+ \to D^{*+}\overline{D}^{*0}$ decays with two missing particles and (fifth row) $B_c^+ \to D^{*+}\overline{D}^{*0}$ decays with one missing particle, summing over contributions from each missing particle species and polarisation. The fitted KDE is also drawn (red) [1].

Table 6.3: Values of branching fractions and $\frac{f_s}{f_d}$ used to relate yields of B_s^0 to yields of B^0 decays.

Quantity	Value	Source
$\mathcal{B}(B_s^0 \to D_s^- D^{*+})$	$(2.9 \pm 0.7) \times 10^{-4}$	[81,82,84]
$\mathcal{B}(B_s^0 \to D_s^{*-} D^{*+})$	$(6.7 \pm 2.2) \times 10^{-4}$	[81,82,84]
$\mathcal{B}(\bar{B}^0 \to D_s^- D^{*+})$	$(8.0 \pm 1.1) \times 10^{-3}$	[22]
$\mathcal{B}(\bar{B}^0 \to D_s^{*-} D^{*+})$	$(1.77 \pm 0.14) \times 10^{-2}$	[22]
$\frac{\mathcal{B}(B_s^0 \to D^{\mp} D^{*\pm})}{\mathcal{B}(B^0 \to D^{\mp} D^{*\pm})}$	0.137 ± 0.018	[202]
$\frac{\mathcal{B}(B_s^0 \to D^{*-} D^{*+})}{\mathcal{B}(B^0 \to D^{*-} D^{*+})}$	0.269 ± 0.034	[201]
$\frac{f_s}{f_d}$ (7 TeV)	0.2390 ± 0.0076	[105]
$\frac{f_s}{f_d}$ (8 TeV)	0.2385 ± 0.0075	[105]
$\frac{f_s}{f_d}$ (13 TeV)	0.2539 ± 0.0079	[105]

data taking periods and final states when fully reconstructing all charm mesons [201,202]. Therefore, 1.05 ± 0.05 is used as the efficiency ratio in Equation 6.21. The input branching fractions and values of $\frac{f_s}{f_d}$ are given in Table 6.3, which are taken from measurements if possible, or otherwise from theoretical predictions.

6.3 Partially reconstructed $B \to D^{**}D$ decays

$B \to D^{**}D$ decays, where D^{**} is the umbrella term for excited charm mesons with orbital angular momentum $L \geq 1$ [203], can enter the low-mass end of the signal region when they are partially reconstructed after emitting one or more pions and photons. This peaking background should be modelled. The branching fractions of $B \to D^{**}D$ decays are not known in general. As a representive model, the branching fractions of the decays involving the four P-wave states are assumed to be equal, and states with higher L are not considered. The assumption of equality is motivated by the similar branching fractions of $B^+ \to D_{sJ}^+(2457)\bar{D}^0$, $B^+ \to D_{s0}^{*+}(2317)\bar{D}^0$, and $B^+ \to D_{s1}^+(2536)\bar{D}^0$ [22]. The P-wave states are the $D_0^*(2300)$, $D_1(2420)$, $D_1^{'} \equiv D_1(2430)$, $D_2(2460)$, $D_{s0}^*(2317)$, $D_{s1}(2460)$, $D_{s1}(2536)$, and $D_{s1}^*(2573)$. The $D_0^*(2300)$ and $D_1(2430)$ have large widths, and the remainder are narrow. It is not necessary to include the $D_{s1}(2536)$ and $D_{s1}^*(2573)$ in the model as they emit a kaon in almost 100% of the decays, resulting in a different flavour content at much lower reconstructed mass.

A template is obtained for each possible decay chain, for both B^- and B^0 mesons, by fitting a KDE to generator-level simulation similarly as in Section 6.2.2. Since these decays occupy a similar mass range to the $B \to DD$ decays with two missing particles, the template is multiplied by the same efficiency slope.

The branching fractions of all considered D^{**} states are known with good precision (Table 6.4). The templates for each decay chain contributing to a final state are summed

Table 6.4: Branching fractions that determine the weight of each decay chain in the $B \to D^{**}D$ template.

Quantity	Value	Source
$\mathcal{B}(D^{*0} \to D^0\pi^0)$	0.647 ± 0.009	[22]
$\mathcal{B}(D^{*0} \to D^0\gamma)$	0.353 ± 0.009	[22]
$\mathcal{B}(D^{*+} \to D^+\pi^0)$	0.307 ± 0.005	[22]
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	0.677 ± 0.005	[22]
$\mathcal{B}(D^{*+} \to D^+\gamma)$	0.016 ± 0.004	[22]
$\mathcal{B}(D_s^{*+} \to D_s^+\pi^0)$	0.0058 ± 0.007	[22]
$\mathcal{B}(D_s^{*+} \to D_s^+\gamma)$	0.935 ± 0.007	[22]
$\mathcal{B}(D_{s0}^+ \to D_s^+\pi^0)$	1	[22]
$\mathcal{B}(D_{s1}^+ \to D_s^+\gamma)$	0.18 ± 0.04	[177]
$\mathcal{B}(D_{s1}^+ \to D_s^{*+}\pi^0)$	0.48 ± 0.11	[177]
$\mathcal{B}(D^{**} \to D^{(*)}\pi^+)/\mathcal{B}(D^{**} \to D^{(*)}\pi^0)$	2	isospin
$\mathcal{B}(D_2^* \to D\pi)/\mathcal{B}(D_2^* \to D^*\pi)$	1.52 ± 0.14	[22]

with weights determined by the branching fractions and assuming equal B^- and B^0 production. The resulting template in each channel, after summing over all D^{**} species, is shown in Figure 6.7.

6.4 Single charm and charmless backgrounds

Single charm and charmless decays, which have the same final state as the double charm decays but without one or both intermediate charm mesons, may form peaking backgrounds under the double charm decays. As discussed in Section 5.7, these backgrounds can be reduced to a large extent using charm lifetime requirements and mass windows. Figure 6.8 shows that there is a peak only in the D_s^+ sideband, from single charm $B^+ \to \overline{D}^0K^+K^-\pi^+$ and charmless $B^+ \to K^+K^+K^-\pi^+\pi^-$ decays, which is expected since these decays have the highest branching fractions. After halving the charmless yield to account for the size of the mass range, it is only 6% of the single charm yield, so the charmless background can be neglected.

The single charm background could only be rejected further by introducing a large signal inefficiency, so it is instead modelled in the fit to data. A DSCB describes the $m(\overline{D}^0K^+K^-\pi^+)$ distribution well, and because the yield is much smaller than the signal yield, a more complicated description is not necessary. Whereas the DTF improves mass resolution if the decay chain is correct, it randomly changes the momenta of tracks not corresponding to the correct decay chain, thereby degrading the mass resolution for single charm decays by the same amount. The tail parameters, the mass resolution without the DTF, and the increase in resolution due to the application of the DTF (σ_{DTF}), are obtained through a simultaneous fit to the simulated raw mass (m_{raw}) and DTF mass (m_{DTF})

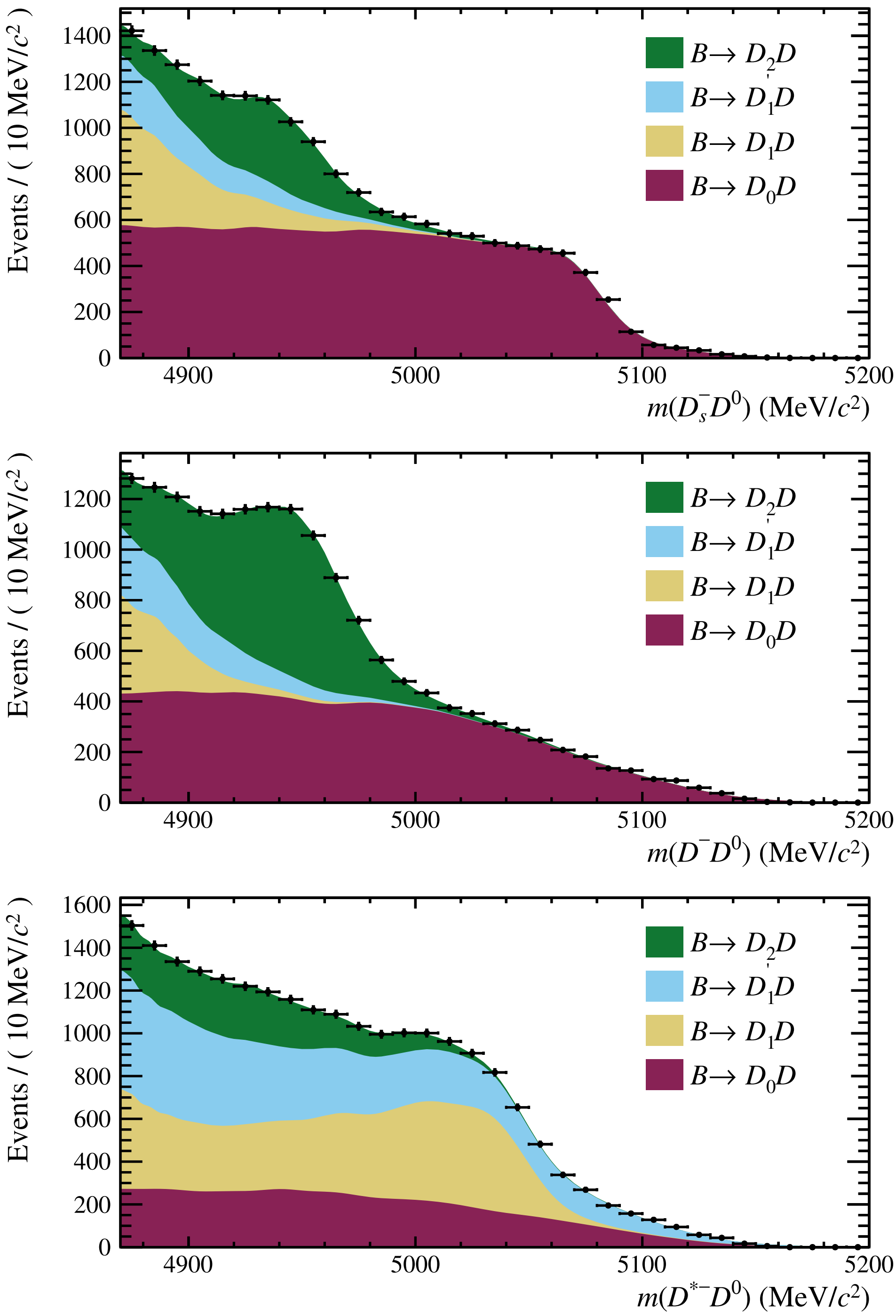

Figure 6.7: Generator-level simulation of $B \to D^{**}D$ decays that are reconstructed as (top) $D_s^- D^0$, (centre) $D^- D^0$ and (bottom) $D^{*-} D^0$. All four D^{**} species and their decay chains are included, weighted by branching fractions, assigning an equal branching fraction to each of the D^{**} species and assuming equal B^- and B^0 production. The contributions to the fitted KDE from each D^{**} species are also shaded.

distributions. For the fit to m_{raw}, the $m(D_s^+)$ window is extended to $\pm 75\,\text{MeV}/c^2$ to improve the statistical precision on the fitted parameters. The window cannot be extended for the fit to m_{DTF}, since it depends on $m(D_s^+)$. Figure 6.9 shows the model fitted to both m_{raw} and m_{DTF}.

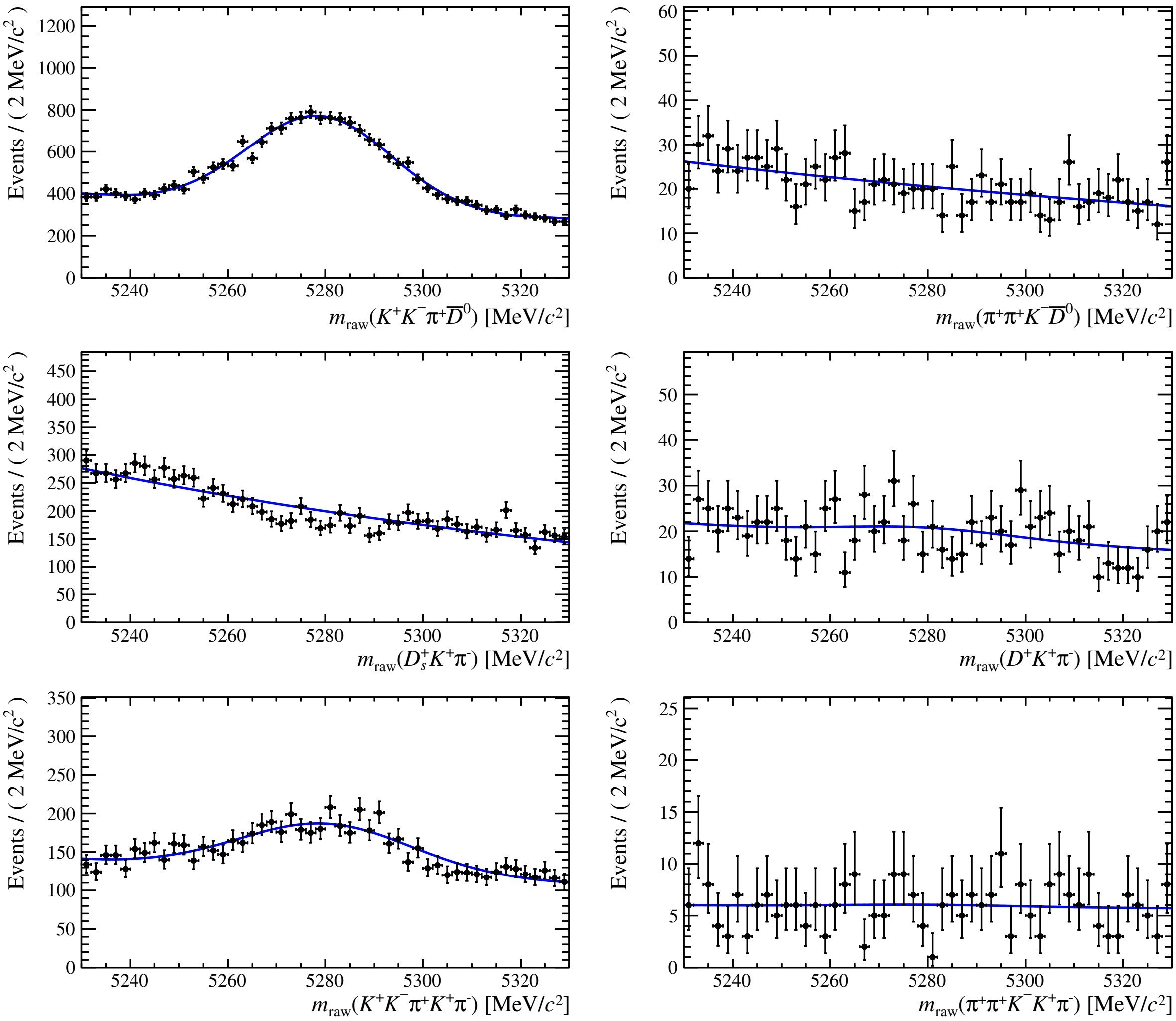

Figure 6.8: Mass distribution without the DTF in the (top) $D^+_{(s)}$ sidebands, (middle) $\overline{D}^0$ sidebands, and (bottom) $D^+_{(s)}$ and $\overline{D}^0$ sidebands. For the (left) $D_s^+\overline{D}^0$ and (right) $D^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ final states in Run 2 data passing selection for the $\mathcal{A}^{CP}$ measurement.

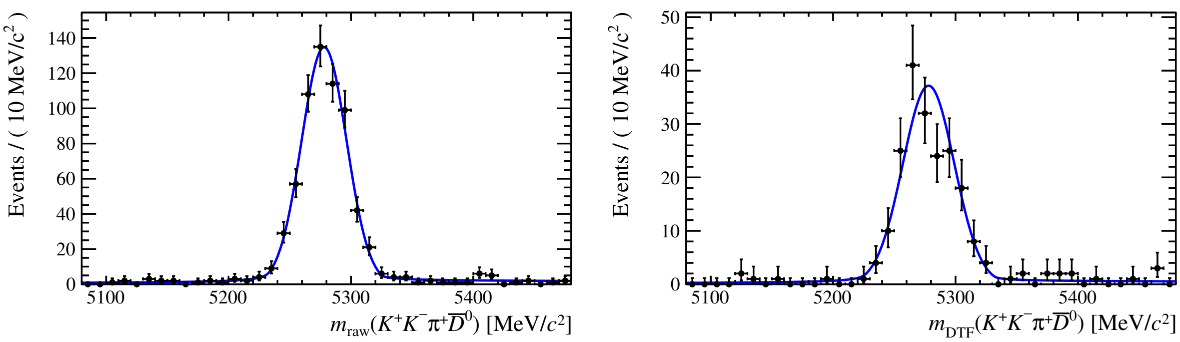

Figure 6.9: The (left) raw and (right) constrained mass distributions of simulated $B^+ \to \overline{D}^0 K^+K^-\pi^+$ decays. The fitted DSCB model described in the text is also drawn.

The DSCB peak position is shared with the fully reconstructed $B^+ \to DD$ decays in the fit to data. As discussed in Section 6.1, mass resolution tends to be mismodelled in simulation. The change in mass resolution introduced by the DTF is however expected to be correctly modelled. Therefore the resolution of the single charm decay in m_{DTF} is the simulated resolution in m_{raw} scaled by the data/simulation ratio of the double charm resolution in m_{raw}, smeared by the simulated change in resolution due to the DTF application:

$$\sigma^2_{DKK\pi}(\mathrm{DTF,data}) = \frac{\sigma^2_{DD}(\mathrm{DTF,data}) - \sigma^2_{\mathrm{DTF}}}{\sigma^2_{DD}(\mathrm{DTF,MC}) - \sigma^2_{\mathrm{DTF}}} \sigma^2_{DKK\pi}(\mathrm{raw,MC}) + \sigma^2_{\mathrm{DTF}}. \tag{6.22}$$

The yield of this background is determined using a fit of this model plus an exponential to describe background from random combinations of tracks to the raw mass distribution in the $25 < |m(D_s^+) - m_{D_s^+}| < 75\,\mathrm{MeV}/c^2$ sidebands (Figure 6.8). The yield is divided by two to account for the width of the sidebands. In the $\mathcal{A}^{CP}$ measurement, the yield is approximately 6% of that of the fully reconstructed signal. The BDT response for single charm decays is skewed towards lower values than for double charm signal, so this background is less important in the search for $B_c^+ \to DD$ decays where tighter requirements on the BDT response are imposed, but it is still modelled.

6.5 Cross-feed

6.5.1 $\overline{D}^0 \to D^0$ cross-feed

$B^+ \to D_{(s)}^{(*)+} \overline{D}^0$ decays create a cross-feed into the WS $B^+ \to D_{(s)}^{(*)+} D^0$ channels through either DCS $\overline{D}^0 \to K^-\pi^+(\pi^-\pi^+)$ decays or CF $\overline{D}^0$ decays with the misidentification of both a kaon and an opposite sign pion. This cross-feed is suppressed using a veto (Section 5.7) but some background may remain. The DCS cross-feed has the same invariant-mass distribution as the RS decays, and its yield in BDT sample i (N^i_{DCS}) is constrained by scaling the RS yield (N^i_{RS}) by the ratio of the DCS to CF D^0 branching fractions as

$$N^i_{\mathrm{DCS}} = \frac{\mathcal{B}(\overline{D}^0 \to K^-\pi^+(\pi^-\pi^+))}{\mathcal{B}(\overline{D}^0 \to K^+\pi^-(\pi^+\pi^-))} N^i_{\mathrm{RS}}. \tag{6.23}$$

Double species misidentification degrades the mass resolution but does not affect the peak position. Therefore the double misidentification cross-feed is described with the same invariant-mass distribution model as the DCS cross-feed, but the width σ_0 is scaled by a factor greater than one. Cross-feed from the CS $B^+ \to D^{(*)+}\overline{D}^0$ decays are small, so the scale factor is determined when fitting the model to data in the $B^+ \to D_s^+ D^0$ channel and the same scale factor is used for each channel. A different scale factor is obtained for Run 1 and Run 2 data and for each D^0 final state. Fitted scale factors are given in Table 6.5.

Fits of the total model in the mass range around m_{B^+} to the $D_s^+ D^0$ channel, which includes both sources of cross-feed as well as combinatorial background (Section 6.6), are shown in Figure 6.10. These fits show the increased width of doubly misidentified decays and that their yield is lower at higher BDT response.

6.5.2 $D_s^+ \longleftrightarrow D^+$ cross-feed

A reconstructed D^+ and D_s^+ differ in the identification of one of the tracks as a pion or kaon. In this way, CF $B^+ \to D_s^+ \overline{D}^0$ decays can produce a peaking background below the CS $B^+ \to D^+ \overline{D}^0$ peak, and there is cross-feed between the $B_c^+ \to D^+ \overline{D}^0$ and $B_c^+ \to D_s^+ \overline{D}^0$ decays. Vetoes remove most of this cross-feed (Section 5.7), and the charm mass windows and BDT requirements provide further rejection, but this peaking background may still be present in the data.

The number of $D^+ \overline{D}^0$ candidates in the $D_s^+ \overline{D}^0$ channel is found in simulation to be no more than 0.4% of the number in the $D^+ \overline{D}^0$ channel after all selection, including the minimum BDT requirement. Assuming the predicted branching fraction ratio (Section 2.5), the yield of CF $B_c^+ \to D^+ \overline{D}^0$ cross-feed in the CS $D_s^+ \overline{D}^0$ channel is expected to be no more than 5% of the signal yield. The yield of CS cross-feed to the CF channel is expected to be yet smaller. Therefore this background can be neglected in the search for $B_c^+ \to DD$ decays as long as the yields are not large.

The yield of $B^+ \to D_s^+ \overline{D}^0$ to $B^+ \to D^+ \overline{D}^0$ cross-feed is similarly calculated to be 4% of the signal yield. In the precise measurement of *CP* asymmetries and branching fractions of the B^- decays, this is not negligible, so it is important to include this background when modelling the data. The expected cross-feed yield is estimated by scaling the ratio between the yields of simulated $B^+ \to D_s^+ \overline{D}^0$ decays selected in the $D^+ \overline{D}^0$ and $D_s^+ \overline{D}^0$ channels by the selected yield in the $D_s^+ \overline{D}^0$ channel in data (Table 6.6).

This background is modelled as the sum of two Gaussians, and the shape parameters are extracted from a fit to the simulated mass distribution (Figure 6.11). Since this background has a low selected yield in simulation, the shape parameters are shared between both final states and data taking periods.

Table 6.5: Scale factors relating the widths of the distribution of $B^+ \to D_s^+ D^0$ candidates from double misidentification to candidates from true $B^+ \to D_s^+ \overline{D}^0$ decays. Values are obtained from a fit in the $B^+ \to D_s^+ D^0$ channel and are split between Run 1 and 2 data and by D^0 final state.

Final state	Run 1	Run 2
$D^0 \to K^+ \pi^-$	1.2 ± 0.2	1.9 ± 0.2
$D^0 \to K^+ \pi^- \pi^+ \pi^-$	1.9 ± 0.4	1.2 ± 0.2

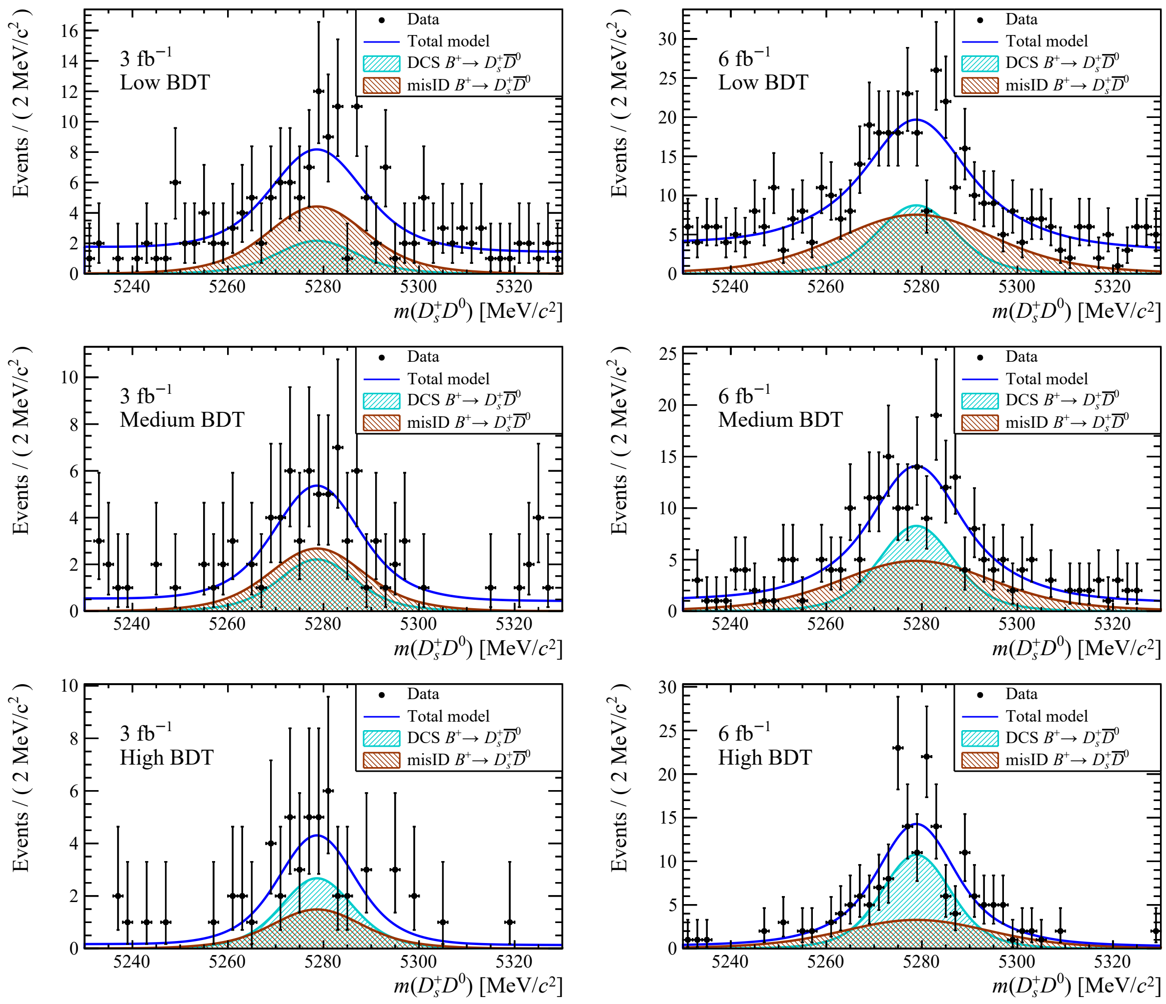

Figure 6.10: Invariant mass distributions of selected $B^+ \to D_s^+ D^0$ candidates in the region near m_{B^+} in (left) Run 1 and (right) Run 2 data. For the (top row) low, (middle row) medium, and (top row) high BDT samples. The model containing contributions from $B^+ \to D_s^+ \overline{D}^0$ cross-feed and combinatorial background is also drawn.

Table 6.6: Expected $B^- \to D_s^- D^0$ to $D^- D^0$ cross-feed yields.

Period	Final state	Expected yield
Run 1	$D^0 \to K^- \pi^+$	43 ± 7
	$D^0 \to K^- \pi^+ \pi^- \pi^+$	12 ± 5
Run 2	$D^0 \to K^- \pi^+$	233 ± 30
	$D^0 \to K^- \pi^+ \pi^- \pi^+$	125 ± 27

6.6 Combinatorial background

Random combinations of tracks form a *combinatorial background* that is smoothly falling as function of invariant mass. A very high fraction is rejected by the selection, but is still clearly present in the selected data. In the analysis of $B^- \to DD$ decays, this background is

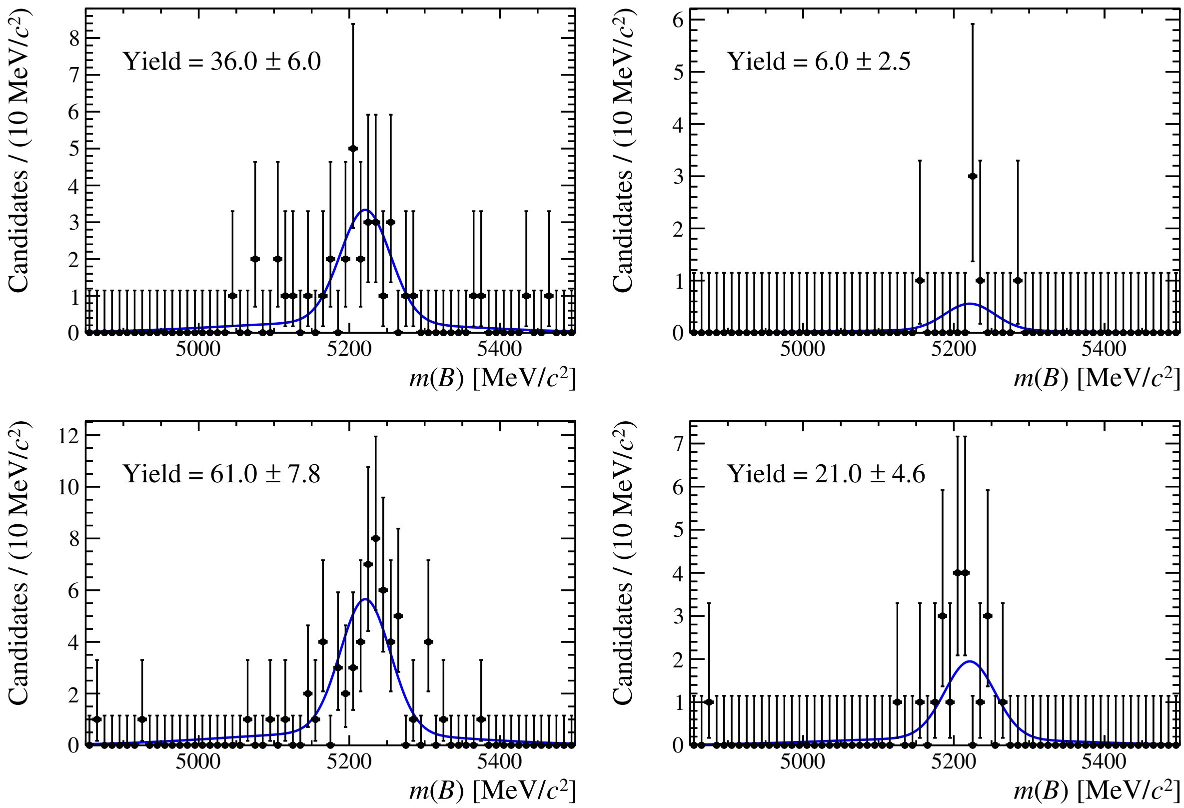

Figure 6.11: Invariant mass distribution for $D_s^+\overline{D}^0$ decays selected in the $D^+\overline{D}^0$ channel and the fit of a double gaussian. For (left) $\overline{D}^0 \to K^+\pi^-$ and (right) $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ final states in (first row) Run 1 and (second row) Run 2 data.

described well by an exponential $f(m) = e^{-\gamma^{\pm}m}$. $B^{\pm}$ candidates are modelled with different decay constants as the underlying event is charge-asymmetric. In the search for $B_c^+ \to DD$ decays, an additional degree of freedom is required to ensure that the background is accurately modelled over the larger mass range ($5230 < m(B) < 6700\,\text{MeV}/c^2$). An exponential function plus a constant is chosen,

$$f(m) = c + e^{-\gamma m}. \tag{6.24}$$

The sensitivity of the search is improved by sharing the combinatorial background shape parameters between all BDT samples when fitting the model to data. In this way, the background shape in the high signal purity sample is informed by the high yield background in the low signal purity sample. The validity of the shape model and the assumption of BDT-independence of the shape is tested by introducing a linear dependence on the mean signal BDT response in the subsample ($0 < \overline{\text{BDT}} < 1$) into the exponent,

$$\gamma = \gamma_0(1 - \gamma_{\overline{\text{BDT}}}\overline{\text{BDT}}), \tag{6.25}$$

and fitting the modified model to data in the charm sidebands (for example.

Table 6.7: Values of $\gamma_{\overline{\mathrm{BDT}}}$ (Equation 6.25) obtained from fitting the BDT-dependent exponential plus constant model to the charm sidebands. The χ^2 for consistency with $\gamma_{\overline{\mathrm{BDT}}} = 0$ is also calculated.

Final state	$\gamma_{\overline{\mathrm{BDT}}}$ Run 1	$\gamma_{\overline{\mathrm{BDT}}}$ Run 2
$D_s^+\overline{D}^0$, $\overline{D}^0 \to K\pi$	0.60 ± 0.25	-0.4 ± 0.5
$D_s^+ D^0$, $D^0 \to K\pi$	-0.2 ± 0.9	-0.3 ± 0.8
$D_s^+\overline{D}^0$, $\overline{D}^0 \to K3\pi$	-0.3 ± 0.6	0.4 ± 0.2
$D_s^+ D^0$, $D^0 \to K3\pi$	-1.8 ± 2.6	-0.2 ± 0.4
$D^+\overline{D}^0$, $\overline{D}^0 \to K\pi$	-0.7 ± 1.3	-0.0 ± 0.8
$D^+ D^0$, $D^0 \to K\pi$	0.9 ± 0.4	0.4 ± 1.1
$D^+\overline{D}^0$, $\overline{D}^0 \to K3\pi$	-11 ± 16	-1.3 ± 2.2
$D^+ D^0$, $D^0 \to K3\pi$	$-\ 4 \pm 10$	0.4 ± 0.9
$D^{*+}\overline{D}^0$, $D^0 \to K\pi$, $\overline{D}^0 \to K\pi$	0.5 ± 0.5	1.0 ± 0.4
$D^{*+} D^0$, $D^0 \to K\pi$, $D^0 \to K\pi$	0.8 ± 0.6	0.0 ± 1.2
$D^{*+}\overline{D}^0$, $D^0 \to K3\pi$, $\overline{D}^0 \to K\pi$	0.0 ± 0.8	0.6 ± 0.5
$D^{*+} D^0$, $D^0 \to K3\pi$, $D^0 \to K\pi$	-0.6 ± 1.4	0.0 ± 0.8
$D^{*+}\overline{D}^0$, $D^0 \to K\pi$, $\overline{D}^0 \to K3\pi$	-1.3 ± 1.3	0.2 ± 0.5
$D^{*+} D^0$, $D^0 \to K\pi$, $D^0 \to K3\pi$	-1.3 ± 5.4	0.7 ± 0.5
χ^2/ndf	15.7/14	17.0/14

$|m(D^+) - m_{D^+}| < 75\,\mathrm{MeV}/c^2$ & $|m(D^0) - m_{D^0}| < 75\,\mathrm{MeV}/c^2$ &! ($|m(D^+) - m_{D^+}| < 25\,\mathrm{MeV}/c^2$ & $|m(D^0) - m_{D^0}| < 25\,\mathrm{MeV}/c^2$) for D^+D^0). Exemplar fits in Figure 6.12 demonstrate that the exponential plus constant function describes this background well. Table 6.7 shows the fitted $\gamma_{\overline{\mathrm{BDT}}}$ in each dataset. By calculating the χ^2 for consistency with $\gamma_{\overline{\mathrm{BDT}}} = 0$, no evidence of a BDT-dependence of the background shape is observed.

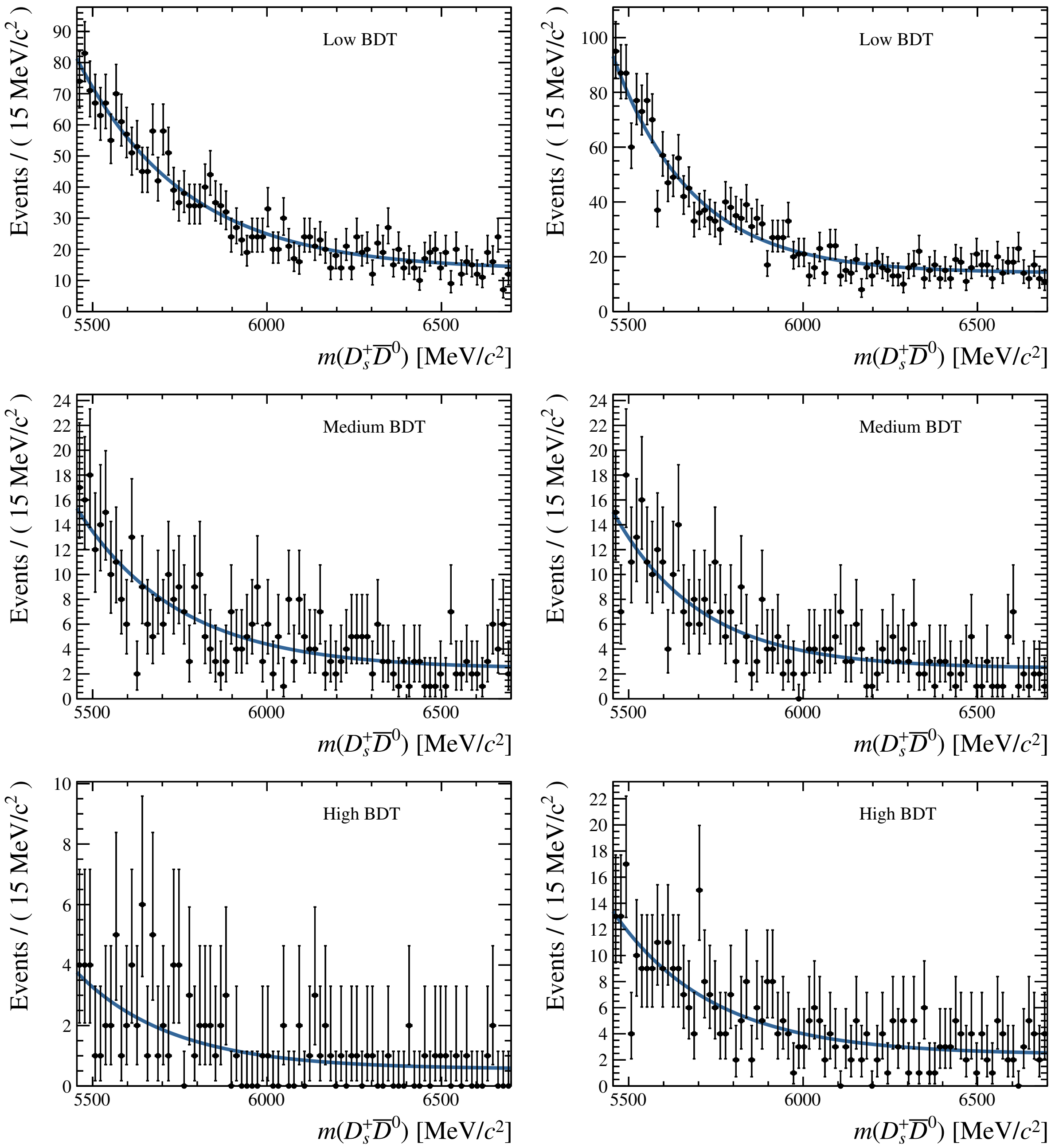

Figure 6.12: $B^+ \to D_s^+ \overline{D}^0$ candidates with $\overline{D}^0 \to K3\pi$ in the charm mass sidebands in (left) Run 1 and (right) Run 2 data in the (top row) low BDT, (middle row) medium BDT, and (top row) high BDT samples. The fit to the BDT-dependent exponential plus constant model described in Equations 6.24-6.25 is also drawn.

Chapter 7

Search for doubly charmed B_c^+ meson decays

This chapter completes the discussion of the search for sixteen $B_c^+ \to DD$ decays that was introduced in Section 4.1. Branching fraction ratios are measured for decays with a significance of greater than three standard deviations, and an upper confidence limit is set on the branching fraction of decays with a signifance of less than five standard deviations. For the fully reconstructed decays, the measured ratios are

$$R(D_{(s)}^{(*)+} \overset{(-)}{D}{}^{0}) \equiv \frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to D_{(s)}^{(*)+} \overset{(-)}{D}{}^{0})}{\mathcal{B}(B^+ \to D_{(s)}^{(*)+} \bar{D}^0)} = \frac{N(B_c^+)}{N(B^+)} \frac{\varepsilon(B^+)}{\varepsilon(B_c^+)}. \tag{7.1}$$

As established in Section 6.2, the invariant-mass distributions of partially reconstructed $B_c^+ \to D_{(s)}^{*+} \overset{(-)}{D}{}^{0}$ and $B_c^+ \to D_{(s)}^{+} \overset{(-)}{D}{}^{*0}$ decays overlap. Their branching fractions can be measured separately by assuming that the signal in the corresponding invariant mass range arises entirely from each decay, as

$$R_+^{'}(D_{(s)}^{+} \overset{(-)}{D}{}^{0}) \equiv \frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to D_{(s)}^{*+} \overset{(-)}{D}{}^{0})}{\mathcal{B}(B^+ \to D_{(s)}^{+} \bar{D}^0)} = \frac{1}{\mathcal{B}(D_{(s)}^{*+} \to D_{(s)}^{+} X^0)} \frac{N(B_c^+)}{N(B^+)} \frac{\varepsilon(B^+)}{\varepsilon(B_c^+)}, \tag{7.2}$$

$$R_0^{'}(D_{(s)}^{+} \overset{(-)}{D}{}^{0}) = \frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to D_{(s)}^{+} \overset{(-)}{D}{}^{*0})}{\mathcal{B}(B^+ \to D_{(s)}^{+} \bar{D}^0)} = \frac{N(B_c^+)}{N(B^+)} \frac{\varepsilon(B^+)}{\varepsilon(B_c^+)}. \tag{7.3}$$

$B_c^+ \to D^* D^*$ decays are observed in both the $D^{*+} \overset{(-)}{D}{}^{0}$ channels when a $\overset{(-)}{D}{}^{*0}$ emits a soft neutral particle, and in the $D_{(s)}^{+} \overset{(-)}{D}{}^{0}$ channels when both charm mesons emit a soft neutral particle. The measured ratios are

$$R^{'}(D^{*+} \overset{(-)}{D}{}^{0}) \equiv \frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to D^{*+} \overset{(-)}{D}{}^{*0})}{\mathcal{B}(B^+ \to D^{*+} \bar{D}^0)} = \frac{N(B_c^+)}{N(B^+)} \frac{\varepsilon(B^+)}{\varepsilon(B_c^+)}, \tag{7.4}$$

$$R^{''}(D_{(s)}^{+} \overset{(-)}{D}{}^{0}) \equiv \frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to D_{(s)}^{*+} \overset{(-)}{D}{}^{*0})}{\mathcal{B}(B^+ \to D_{(s)}^{+} \bar{D}^0)} = \frac{1}{\mathcal{B}(D_{(s)}^{*+} \to D_{(s)}^{+} X^0)} \frac{N(B_c^+)}{N(B^+)} \frac{\varepsilon(B^+)}{\varepsilon(B_c^+)}. \tag{7.5}$$

Efficiencies of signal and normalisation modes are evaluated in Section 7.1. Section 7.2 explains how the data is modelled and how signal and normalisation yields are extracted by fitting the model to data. Sources of systematic uncertainty on the branching fractions

Table 7.1: Acceptance, reconstruction and selection efficiencies of $B_{(c)}^+ \to D_{(s)}^{(*)+}\overline{D}^0$ decays in the kinematic range of interest, as defined in Equation 4.2, and the ratio between the B_c^+ and B^+ efficiencies. The uncertainties are only from the finite size of the simulation samples.

Period	Channel	Final state	$\varepsilon(B_c^+)/10^{-3}$	$\varepsilon(B^+)/10^{-3}$	$\varepsilon(B_c^+)/\varepsilon(B^+)$
Run 1	$D_s^+\overline{D}^0$	$KK\pi, K\pi$	9.28 ± 0.12	19.94 ± 0.19	0.465 ± 0.008
	$D_s^+\overline{D}^0$	$KK\pi, K3\pi$	2.64 ± 0.05	6.45 ± 0.09	0.409 ± 0.010
	$D^+\overline{D}^0$	$K\pi\pi, K\pi$	7.48 ± 0.11	16.36 ± 0.17	0.457 ± 0.008
	$D^+\overline{D}^0$	$K\pi\pi, K3\pi$	2.11 ± 0.05	5.16 ± 0.08	0.410 ± 0.011
	$D^{*+}\overline{D}^0$	$K\pi, K\pi$	5.21 ± 0.08	14.33 ± 0.20	0.364 ± 0.008
	$D^{*+}\overline{D}^0$	$K\pi, K3\pi$	1.78 ± 0.04	5.11 ± 0.11	0.348 ± 0.011
	$D^{*+}\overline{D}^0$	$K3\pi, K\pi$	1.79 ± 0.04	5.28 ± 0.10	0.339 ± 0.010
Run 2	$D_s^+\overline{D}^0$	$KK\pi, K\pi$	9.73 ± 0.12	20.55 ± 0.19	0.473 ± 0.007
	$D_s^+\overline{D}^0$	$KK\pi, K3\pi$	2.71 ± 0.05	7.01 ± 0.09	0.386 ± 0.008
	$D^+\overline{D}^0$	$K\pi\pi, K\pi$	8.26 ± 0.10	17.19 ± 0.17	0.480 ± 0.008
	$D^+\overline{D}^0$	$K\pi\pi, K3\pi$	2.15 ± 0.04	5.41 ± 0.07	0.396 ± 0.009
	$D^{*+}\overline{D}^0$	$K\pi, K\pi$	5.74 ± 0.09	18.51 ± 0.20	0.310 ± 0.006
	$D^{*+}\overline{D}^0$	$K\pi, K3\pi$	1.97 ± 0.06	7.26 ± 0.12	0.271 ± 0.009
	$D^{*+}\overline{D}^0$	$K3\pi, K\pi$	1.92 ± 0.05	6.64 ± 0.10	0.289 ± 0.009

are discussed in Section 7.3. Section 7.4 then presents the measurements of and confidence limits on the branching fractions.

7.1 Signal and normalisation efficiencies

Branching fraction measurements rely on knowledge of the acceptance, reconstruction, and selection efficiency for both signal and normalisation modes (Equation 7.1). These efficiencies are calculated by dividing the yield of selected simulation events in each BDT sample after applying the corrections in Section 5.3.1-5.3.6 by the number of events generated with a signal or normalisation candidate in the kinematic range of interest (Equation 4.2).

B_c^+ decays have 30-50% of the efficiency of the corresponding B^+ decays due to the shorter lifetime of the B_c^+. The efficiencies for B_c^+ and B^+ decays, summed over all BDT samples, are displayed in Table 7.1. Partially reconstructed decays have a lower efficiency than fully reconstructed decays by 10-20% due to the lower momentum of the reconstructed particles.

The efficiency of L0 trigger requirements is poorly modelled in simulation, and this is discussed in detail in Section 8.3.1. However, it is reasonable to assume that this mismodelling will cancel in the ratio between B_c^+ and B^+ decays to the same final state to a degree much better than the precision of this analysis. This is supported by observing

Table 7.2: Fraction of $D_s^+\bar{D}^0$ candidates in selected simulation passing the L0 Global TIS (f_{TIS}) and Hadron TOS (f_{TOS}) requirements. Calculated using simulation that has passed all preselection cuts and with a BDT response greater than the minimum threshold (Section 5.8.2).

Period	Final state	$f_{\text{TIS}}(B_c^+)$	$f_{\text{TIS}}(B^+)$	$f_{\text{TOS}}(B_c^+)$	$f_{\text{TOS}}(B^+)$
Run 1	$\bar{D}^0 \to K\pi$	0.587 ± 0.009	0.596 ± 0.006	0.651 ± 0.010	0.661 ± 0.006
	$\bar{D}^0 \to K3\pi$	0.630 ± 0.019	0.626 ± 0.010	0.614 ± 0.018	0.621 ± 0.010
Run 2	$\bar{D}^0 \to K\pi$	0.593 ± 0.006	0.590 ± 0.005	0.666 ± 0.006	0.656 ± 0.006
	$\bar{D}^0 \to K3\pi$	0.653 ± 0.011	0.641 ± 0.008	0.610 ± 0.011	0.605 ± 0.008

that the fraction of simulated candidates passing the L0 Global TIS and L0 Hadron TOS requirements is equal within statistical uncertainties for B_c^+ and B^+ decays (Table 7.2). Therefore mismodelling of L0 trigger efficiencies is not considered further in this analysis.

7.2 Fit to data

7.2.1 Fit method

The branching fraction ratios are extracted from data using an extended unbinned maximum likelihood fit of the PDF $F(m;\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the set of free parameters in the PDF. The parameters that best describe the data are found by maximising

$$L(\boldsymbol{\theta}) = \frac{\nu(\boldsymbol{\theta})^n e^{-\nu(\boldsymbol{\theta})}}{n!} \prod_{i=1}^{n} F(m_i;\boldsymbol{\theta}), \tag{7.6}$$

where $\nu(\boldsymbol{\theta})$ is the expected total yield of the PDF as a function of $\boldsymbol{\theta}$, and the product is over all candidates in the dataset. This fit is performed simultaneously over all BDT samples and final states, sharing certain shape and signal strength parameters. Maximum likelihood estimation is only fully unbiased in the asymptotic limit as the number of candidates in the dataset tends to infinity. In the extreme case of no candidates in the signal window, the estimated signal yield diverges to negative infinity. This is of concern in this analysis, since there are no candidates within $\pm 20\,\text{MeV}/c^2$ of $m_{B_c^+}$ in some datasets. Therefore the yields of all components are required to be positive. This introduces an explicit bias to the signal yields. To avoid bias on the measurements, branching fractions are only measured for decay modes with a signal significance (S) greater than three standard deviations.

7.2.2 Model

The components modelled in this search are the fully reconstructed signal and normalisation decays (Section 6.1), partially reconstructed signal decays (Section 6.2), single charm background (Section 6.4), $\bar{D}^0$ cross-feed to D^0 (Section 6.5.1), and combinatorial background (Section 6.6). The shape parameters to be shared have been specified in these sections.

The signal ($N_{B_c^+}^{i,\text{FS}}$) and normalisation ($N_{B^+}^{i,\text{FS}}$) yields in each BDT sample (i) and final state (FS) are related through the branching fraction ratio R (Equation 7.1) and the efficiencies ($\varepsilon^{i,\text{FS}}$) as

$$N_{B_c^+}^{i,\text{FS}} = R\frac{\varepsilon_{B_c^+}^{i,\text{FS}}}{\varepsilon_{B^+}^{i,\text{FS}}}N_{B^+}^{i,\text{FS}}. \tag{7.7}$$

In this way, R is shared between all BDT samples and final states, which allows R to be extracted straightforwardly and with maximum precision. The quantity that relates $N_{B_c^+}^{i,\text{FS}}$ and R is henceforth referred to as the *normalisation factor*,

$$\mathcal{N}^{i,\text{FS}} = \frac{\varepsilon_{B_c^+}^{i,\text{FS}}}{\varepsilon_{B^+}^{i,\text{FS}}}N_{B^+}^{i,\text{FS}}. \tag{7.8}$$

The yields of doubly misidentified $\overline{D}^0$ decays and combinatorial background are free parameters in each BDT sample. The single charm yield in each BDT sample is fixed as described in Section 6.4. The yield of DCS $\overline{D}^0$ decay cross-feed is constrained to its expected value calculated in Section 6.5.1, which means that the PDF is multiplied by a normal distribution $\exp\left(-(x-\mu)^2/2\sigma^2\right)$ where μ is the expected value of the parameter x to be constrained and σ is the uncertainty on the expected value.

7.2.3 Blinding

The data in the signal window ($[5500, 6325]\,\text{MeV}/c^2$ for $D_{(s)}^+ \overset{(-)}{D}{}^0$; $[5845, 6325]\,\text{MeV}/c^2$ for $D^{*+} \overset{(-)}{D}{}^0$) was blinded during the development of the method to avoid bias. All of the data presented in this chapter is now unblinded.

A normalisation plus background model was fitted to the data excluding the mass range to be blinded. This model was interpolated into the signal window. The signal window was divided into twenty equal windows. To blind the data, the data in each window was replaced by fake data randomly generated in each window according to the expected background shape and yield from the interpolated fit and Poisson statistics. This blinding method allows the analysis workflow to be fully tested, including with the true data outside of the signal window.

7.2.4 Fit results

The fitted yields of B_c^+ candidates in the selected data are given in Tables 7.3-7.5, including statistical uncertainties and the systematic uncertainties described in Section 7.3. The invariant mass distributions in each channel and the highest purity BDT samples summing over Run 1 and Run 2 and all D^0 final states are shown in Figures 7.1.

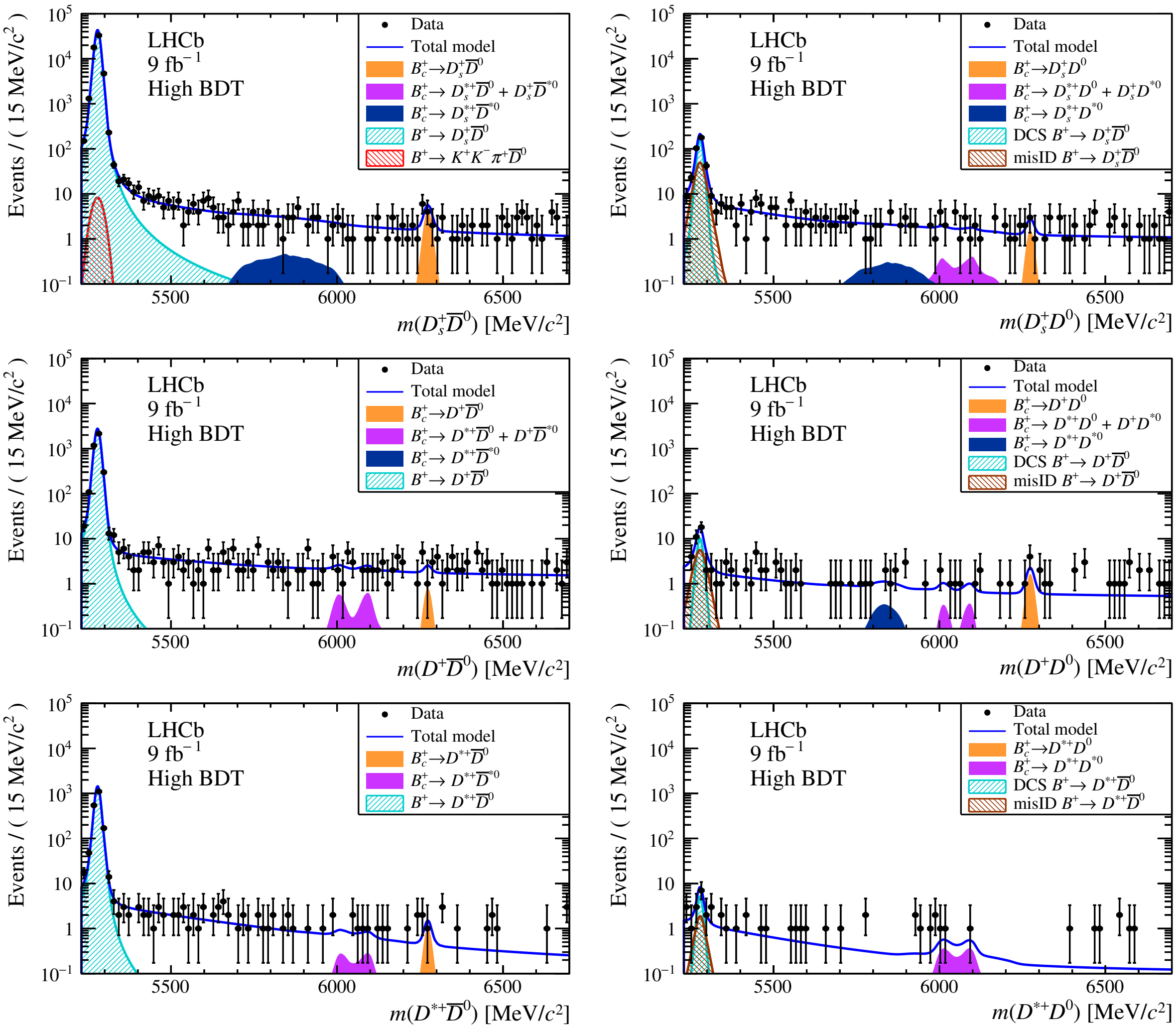

Figure 7.1: Invariant mass distribution of (top left) $D_s^+\overline{D}^0$, (top right) $D_s^+D^0$, (middle left) $D^+\overline{D}^0$, (middle right) D^+D^0, (bottom left) $D^{*+}\overline{D}^0$, and (bottom right) $D^{*+}D^0$ candidates summing over all final states in Run 1+2 data in the high purity BDT samples. The fitted model described in the text is also drawn.

Figure 7.2 shows the invariant mass distributions of $B_{(c)}^+ \to D_s^+\overline{D}^0$ candidates in Run 2 data for each BDT sample and D^0 final state, corresponding to the datasets involved in a single fit. This demonstrates how the highest purity BDT sample provides the best sensitivity to B_c^+ signal, whereas the lowest purity BDT sample helps to constrain the shape of the combinatorial background.

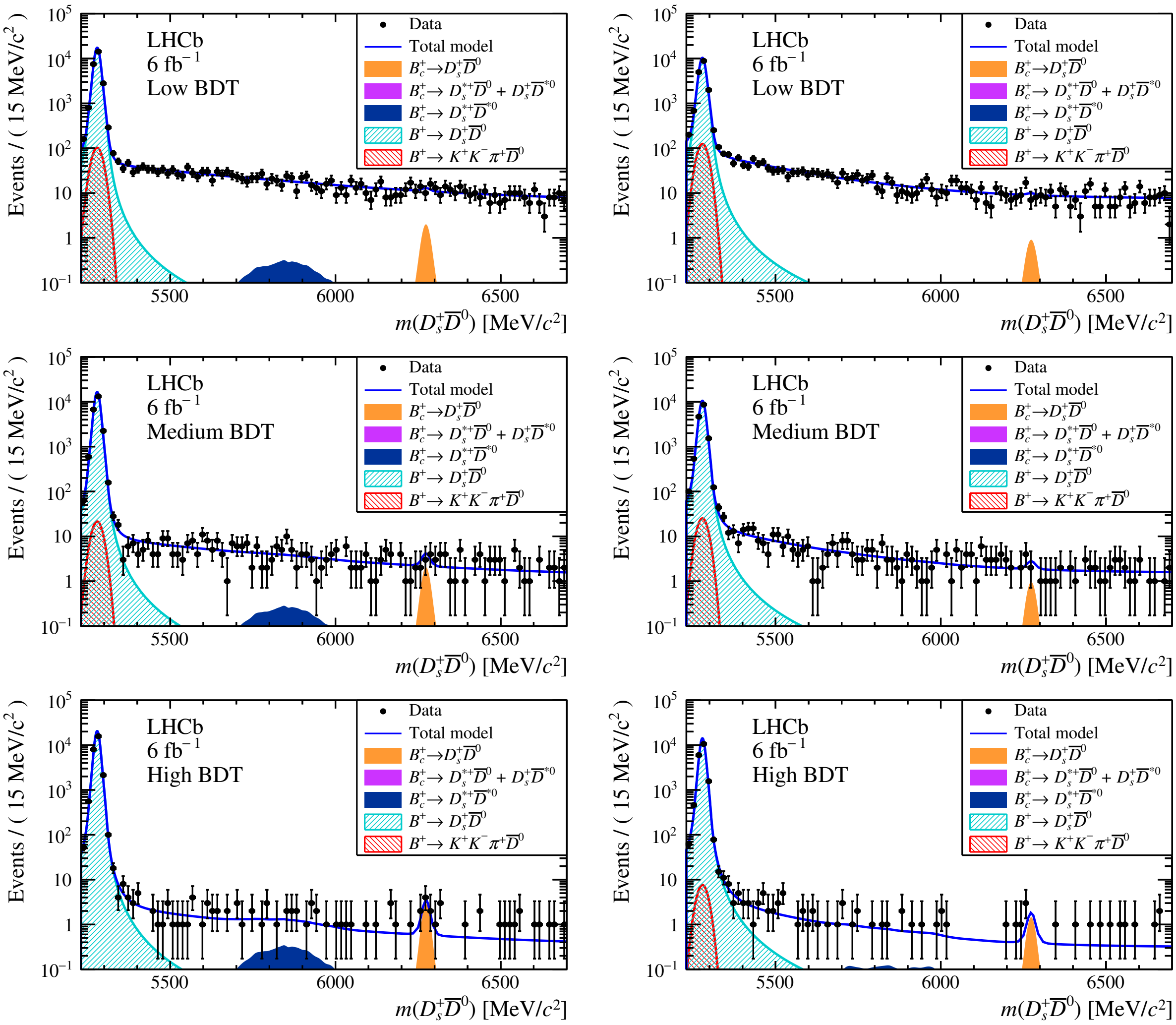

Figure 7.2: Invariant mass distribution of $B_{(c)}^+ \to D_s^+\overline{D}^0$ candidates in Run 2 data with (left) $D^0 \to K^-\pi^+$ and (right) $D^0 \to K^-\pi^+\pi^-\pi^+$ for data corresponding to the (top) lowest, (middle) intermediate, and (bottom) highest BDT sample. The fitted model described in the text is also drawn and corresponds to a single combined fit to all BDT samples and D^0 final states.

7.3 Systematic uncertainties

7.3.1 Systematic uncertainties on the normalisation

Systematic uncertainties account for limited knowledge on effects including the efficiencies, modelling of invariant mass distributions, and D^* branching fractions. Many of these systematic uncertainties can be expressed as a relative uncertainty on R. The effect of these uncertainties is incorporated into the fit described in the previous section and consquently into the measurement of R by allowing $\mathcal{N}$ to vary by writing

$$\mathcal{N} = \mathcal{N}_0 \times \left(1 + \frac{\Delta\mathcal{N}}{\mathcal{N}_0}\right)^{\alpha}, \tag{7.9}$$

Table 7.3: Yields of $B_c^+ \to D_s^{(*)+}\overline{D}^{(*)0}$ and $B_c^+ \to D_s^{(*)+}D^{(*)0}$ decays in the selected data, summing over all BDT samples. Both statistical and systematic uncertainties are included.

Period	Final state	$N(D_s^+\overline{D}^0)$	$N(D_s^{*+}\overline{D}^0 + D_s^+\overline{D}^{*0})$	$N(D_s^{*+}\overline{D}^{*0})$
Run 1	$K\pi, KK\pi$	$0.0^{+1.6}_{-0.0}$	$0.0^{+1.7}_{-0.0}$	$0.0^{+3.0}_{-0.0}$
	$K3\pi, KK\pi$	$0.0^{+0.8}_{-0.0}$	$0.0^{+0.9}_{-0.0}$	$0.0^{+1.5}_{-0.0}$
Run 2	$K\pi, KK\pi$	$12.8^{+5.1}_{-4.4}$	$0.0^{+4.9}_{-0.0}$	$12.5^{+17.4}_{-12.5}$
	$K3\pi, KK\pi$	$7.1^{+2.8}_{-2.4}$	$0.0^{+2.7}_{-0.0}$	$6.8^{+9.5}_{-6.8}$
		$N(D_s^+D^0)$	$N(D_s^{*+}D^0 + D_s^+D^{*0})$	$N(D_s^{*+}D^{*0})$
Run 1	$K\pi, KK\pi$	$0.6^{+2.3}_{-0.6}$	$0.0^{+2.3}_{-0.0}$	$0.0^{+5.2}_{-0.0}$
	$K3\pi, KK\pi$	$0.3^{+1.2}_{-0.3}$	$0.0^{+1.2}_{-0.0}$	$0.0^{+2.9}_{-0.0}$
Run 2	$K\pi, KK\pi$	$3.9^{+3.3}_{-2.6}$	$6.7^{+8.3}_{-6.7}$	$6.2^{+15.7}_{-6.2}$
	$K3\pi, KK\pi$	$2.1^{+1.8}_{-1.4}$	$3.7^{+4.5}_{-3.7}$	$3.4^{+8.5}_{-3.4}$

Table 7.4: Yields of $B_c^+ \to D^{(*)+}\overline{D}^{(*)0}$ and $B_c^+ \to D^{(*)+}D^{(*)0}$ decays in the selected data, summing over all BDT samples. Both statistical and systematic uncertainties are included.

Period	Final state	$N(D^+\overline{D}^0)$	$N(D^{*+}\overline{D}^0 + D^+\overline{D}^{*0})$	$N(D^{*+}\overline{D}^{*0})$
Run 1	$K\pi, K\pi\pi$	$0.7^{+2.1}_{-0.7}$	$0.0^{+2.5}_{-0.0}$	$0.0^{+16.8}_{-0.0}$
	$K3\pi, K\pi\pi$	$0.3^{+1.1}_{-0.3}$	$0.0^{+1.3}_{-0.0}$	$0.0^{+9.1}_{-0.0}$
Run 2	$K\pi, K\pi\pi$	$1.9^{+3.9}_{-1.9}$	$9.2^{+9.8}_{-8.5}$	$0.0^{+15.5}_{-0.0}$
	$K3\pi, K\pi\pi$	$1.0^{+2.0}_{-1.0}$	$4.9^{+5.2}_{-4.5}$	$0.0^{+8.2}_{-0.0}$
		$N(D^+D^0)$	$N(D^{*+}D^0 + D^+D^{*0})$	$N(D^{*+}D^{*0})$
Run 1	$K\pi, K\pi\pi$	$2.8^{+0.2}_{-1.6}$	$0.0^{+0.0}_{-0.0}$	$19.4^{+16.3}_{-13.2}$
	$K3\pi, K\pi\pi$	$1.6^{+0.0}_{-0.9}$	$0.0^{+0.0}_{-0.0}$	$10.6^{+8.9}_{-7.2}$
Run 2	$K\pi, K\pi\pi$	$2.2^{+2.6}_{-1.9}$	$3.9^{+6.6}_{-3.9}$	$0.0^{+15.3}_{-0.0}$
	$K3\pi, K\pi\pi$	$1.2^{+1.5}_{-1.0}$	$2.0^{+3.5}_{-2.0}$	$0.0^{+8.0}_{-0.0}$

where $\mathcal{N}_0$ is the nominal value of $\mathcal{N}$, $\Delta\mathcal{N}$ is its uncertainty, and α is a free parameter in the fit constrained by a normal distribution with zero mean and unit standard deviation. This results in a *log-normal* distribution of $\mathcal{N}$, which means that $\log\mathcal{N}$ is normally distributed. When $\Delta\mathcal{N}/\mathcal{N}_0 \ll 1$, a normal distribution with mean $\mathcal{N}_0$ and standard deviation $\Delta\mathcal{N}$ is recovered.

All systematic uncertainties except that due to the finite size of the simulation samples are correlated between BDT samples and D^0 final states. These correlations are incorporated into the measurement by decomposing the log-normal variation into a correlated and uncorrelated component, such that

$$\mathcal{N}^{i,\mathrm{FS}} = \mathcal{N}_0^{i,\mathrm{FS}} \times \left(1 + \frac{\Delta\mathcal{N}_{\mathrm{corr}}^{i,\mathrm{FS}}}{\mathcal{N}_0^{i,\mathrm{FS}}}\right)^{\alpha_{\mathrm{corr}}} \times \left(1 + \frac{\Delta\mathcal{N}_{\mathrm{uncorr}}^{i,\mathrm{FS}}}{\mathcal{N}_0^{i,\mathrm{FS}}}\right)^{\alpha_{\mathrm{uncorr}}^{i,\mathrm{FS}}}, \tag{7.10}$$

where $\Delta\mathcal{N}_{(\mathrm{un})\mathrm{corr}}$ are the uncertainties from (un)correlated systematics. The variation

Table 7.5: Yields of $B_c^+ \to D^{(*)+}\overline{D}^{(*)0}$ and $B_c^+ \to D^{(*)+}D^{(*)0}$ decays in the selected data, summing over all BDT samples. Both statistical and systematic uncertainties are included.

Period	Final state	$N(D^{*+}\overline{D}^0)$	$N(D^{*+}\overline{D}^{*0})$
Run 1	$K\pi, K\pi\pi_s$	$0.8^{+1.3}_{-0.8}$	$0.0^{+1.9}_{-0.0}$
	$K\pi, K3\pi\pi_s$	$0.6^{+0.9}_{-0.5}$	$0.0^{+1.3}_{-0.0}$
	$K3\pi, K\pi\pi_s$	$0.5^{+0.9}_{-0.5}$	$0.0^{+1.3}_{-0.0}$
Run 2	$K\pi, K\pi\pi_s$	$0.6^{+1.3}_{-0.6}$	$1.9^{+3.1}_{-1.9}$
	$K\pi, K3\pi\pi_s$	$0.4^{+0.9}_{-0.4}$	$1.3^{+2.2}_{-1.3}$
	$K3\pi, K\pi\pi_s$	$0.4^{+0.8}_{-0.4}$	$1.2^{+2.0}_{-1.2}$
		$N(D^{*+}D^0)$	$N(D^{*+}D^{*0})$
Run 1	$K\pi, K\pi\pi_s$	$0.0^{+0.4}_{-0.0}$	$1.3^{+1.7}_{-1.3}$
	$K\pi, K3\pi\pi_s$	$0.0^{+0.2}_{-0.0}$	$1.0^{+1.2}_{-1.0}$
	$K3\pi, K\pi\pi_s$	$0.0^{+0.2}_{-0.0}$	$0.9^{+1.1}_{-0.9}$
Run 2	$K\pi, K\pi\pi_s$	$0.0^{+0.4}_{-0.0}$	$0.8^{+1.7}_{-0.8}$
	$K\pi, K3\pi\pi_s$	$0.0^{+0.3}_{-0.0}$	$0.5^{+1.2}_{-0.5}$
	$K3\pi, K\pi\pi_s$	$0.0^{+0.3}_{-0.0}$	$0.5^{+1.2}_{-0.5}$

corresponding to the correlated component (α_{corr}) is shared between data samples, whereas a variation corresponding to the uncorrelated component ($\alpha^{i,\text{FS}}_{\text{uncorr}}$) is assigned to each dataset.

The remainder of this section describes the sources of systematic uncertainty that are treated as an uncertainty on the normalisation, and explains how their size is evaluated. The uncertainties are tabulated in Table 7.6. Unless specified uncertainties are given for the fully reconstructed signal, but are of similar size for partially reconstructed signal.

Finite size of simulation sample: The uncertainty associated with the simulation sample size is propagated.

PID variable correction: Kinematic differences between B_c^+ and B^+ decays may result in small differences in the PID efficiency ratio. A systematic uncertainty is assigned by comparing the efficiency ratio with and without the application of PID corrections. Although conservative, this evaluation of the systematic uncertainty is acceptable since it is still small compared to other sources of systematic uncertainty.

D^{*+} branching fractions: The uncertainty on $\mathcal{B}(D^{*+} \to D^+X^0)$ is propagated.

B_c^+ lifetime: The uncertainty on $\tau(B_c^+)$ is propagated to an uncertainty on the weights.

Single charm background: To account for statistical uncertainty and possibly invalid assumptions when extracting the yield of the single-charm background to $B^+ \to D_s^+\overline{D}^0$ from the sidebands, this yield is assigned a 100% uncertainty. Although this is a conservative estimate, it is still one of the smaller systematic uncertainties.

Multiple candidate rejection: The rate and distribution of multiple candidates are poorly described by simulation, and this mismodelling does not necessarily cancel between the signal and normalisation modes [194]. The fraction of candidates in the signal region removed by the multiple candidate removal and no other preselection is assigned as a systematic uncertainty.

Hit resolution parameterisation: Hit resolution in the VELO is mismodelled in 2017-18 simulation [190]. This biases the χ^2_{IP} distribution by approximately −20% and therefore biases simulated efficiencies. The change in efficiency ratio as a result of increasing the χ^2_{IP} requirement on all final state particles by 20% is assigned as a systematic uncertainty.

Fully reconstructed signal model: Toy studies demonstrate that uncertainty on the fully reconstructed signal shape leads to a multiplicative uncertainty on the signal yield, which can consequently be absorbed into the normalisation factor. Toy data is generated under a normalisation-plus-background hypothesis according to Poisson statistics. A complete fit (signal, normalisation and background) is then performed for each toy under a double sided Crystal Ball plus Gaussian (DCSBG) and double Gaussian (DG) signal shape hypothesis. The DSCB tail parameters are fixed to $\alpha_1 = +1.5$ and $\alpha_2 = -1.5$, and all other shape parameters are fixed to their values from simulation. Since only positive signal yields are considered in this analysis, toys where the fitted branching fraction ratio is zero are discarded. Owing to the low expected background yield, this can occur for a large fraction of toys: For certain Run 1 datasets containing $B^+ \to D^{*+}\overline{D}^0$ candidates where the expected background yield within 15 MeV of $m_{B_c^+}$ is less than 1 in the low purity and 0.01 in the high purity BDT samples, up to 90% of toys must be discarded. The number of generated toys is adjusted acccording to the fraction of discarded toys to ensure an ensemble of adequate size to precisely evaluate this systematic uncertainty. The distribution of the ratio of the fitted branching fraction ratio under the DSCBG hypothesis relative to the DG hypothesis peaks a few percent above unity, as expected from the slowly decaying tails of the DSCBG. Only a weak dependence of the peak position on the background yield is observed. This indicates that uncertainty on the fully reconstructed signal model can be treated as a fractional uncertainty on the signal yield and therefore absorbed into the normalisation factor. The systematic uncertainty is obtained by subtracting unity from the median ratio.

B^+ production spectrum: There are systematic uncertainties on $\varepsilon(B^+)$ associated with the background subtraction and the statistical uncertainty of the training datasets. The uncertainty associated with background subtraction is calculated as the change in efficiency when a linear function instead of an exponential is used in background subtraction. The statistical uncertainty is calculated as the change in efficiency when the number of trees in the reweighter is halved from 200 to 100. The overall systematic is dominated by the statistical uncertainty.

B_c^+ production spectrum: The systematic uncertainty on $\varepsilon(B_c^+)$ from the kinematic correction is dominated by uncertainties on the measured slope of $\frac{f_c}{f_u+f_d}$ [106]. This uncertainty is evaluated as the change in efficiency when the parameters p_1 and p_2 are varied by their uncertainty.

Table 7.6: Effective fractional contributions of the systematic uncertainties on the normalisation in percent. A weighted average is performed over all BDT samples and D^0 decay modes. Weights are the normalisation yield in each dataset.

Final state	$D_s^+ \overset{(-)}{D}{}^0$		$D^+ \overset{(-)}{D}{}^0$		$D^{*+} \overset{(-)}{D}{}^0$	
	Run 1	Run 2	Run 1	Run 2	Run 1	Run 2
B_c^+ signal shape	9.4	3.8	4.8	5.3	2.8	3.9
B_c^+ production spectrum	3.7	2.4	3.9	2.4	4.2	2.9
B^+ production spectrum	0.5	0.9	0.6	1.0	0.6	1.1
Hit resolution parameterisation	–	1.5	–	1.2	–	2.2
R simulation sample size	1.2	1.0	1.4	1.1	1.5	1.5
$R'_{(+,0)}$ simulation sample size	1.4	0.9	2.1	1.2	1.1	1.1
R'' simulation sample size	1.5	0.8	1.7	0.9	–	–
B_c^+ lifetime	1.3	1.4	1.3	1.3	2.1	2.6
PID efficiencies	1.6	1.2	2.8	0.8	2.2	1.4
Multiple $B_{(c)}^+$ candidates	0.4	0.4	0.6	0.5	1.4	1.2
Data-simulation differences	0.1	0.1	0.1	0.1	0.1	0.2
$B^+ \to \overline{D}{}^0 K^+ K^- \pi^+$	0.7	0.5	–	–	–	–
$\mathcal{B}(D^{*+} \to D^+ X^0)$	–	–	1.5	1.5	–	–
R total	10.4	5.3	7.2	6.6	6.3	6.5
$R'_{(+,0)}$ total	4.6	3.7	5.7	3.8	5.5	5.0
R'' total	4.6	3.7	5.5	3.7	–	–

Data-simulation agreement in BDT variables: There is a systematic uncertainty on the relative efficiency between BDT samples due to data-Simulation differences in the variables used in the BDT. This is evaluated by fitting a linear function to the ratio between data and simulation of the BDT response. The systematic uncertainty is the fractional change in the efficiency ratio when both B_c^+ and B^+ simulation are weighted by this function, which is normalised so that the number of candidates with no BDT requirement is unchanged.

7.3.2 Systematic uncertainties on the model

Systematic uncertainties due to model choice that cannot be absorbed into the systematic on the normalisation are included as nuisance parameters in the model. This section will outline the sources of these systematic uncertainties and how they are incorporated into the model. Including these model uncertainties results in an increase of the upper limits calculated as described in Section 7.4.2 on the branching fraction ratios of 7% on average.

Combinatorial background model: The shape of the combinatorial background is not known beyond considering the goodness of fit to the data for different models. This systematic uncertainty is incorporated with the discrete profiling method, where the model choice is indexed with a discrete nuisance parameter that is chosen to maximise the likelihood [204]. The model choices are an exponential and an exponential plus a constant.

Normalisation mode model: There is an uncertainty of the shape of the normalisation mode that affects the normalisation yield and may also affect the combinatorial background shape. The overall width and peak position are already free parameters in the fit to data. The tail parameters α_i have an uncertainty from the finite simulation sample size and a systematic uncertainty from possible mismodelling in simulation. The α_i do not differ between data and simulation by more than approximately 0.2 (Table 7.7). Following the earlier conservative approach to small systematic uncertainties, each α_i is assigned an uncertainty that is the quadrature sum of their statistical uncertainty and 0.2. In the fit to data, the α_i are allowed to vary within Gaussian constraints, where the width of each constraint is the uncertainty.

Fully reconstructed signal width: The ratio $\sigma(B_c^+)/\sigma(B^+)$ is determined from simulation with a statistical uncertainty of $15-20\%$, which is assigned as a systematic uncertainty on the fit to data.

Fully reconstructed signal peak position: There is an uncertainty on the difference between $\overline{m}(B_c^+)$ and $\overline{m}(B^+)$ due to uncertainty on the measured masses ($m_{B_c^+} = 6274.5 \pm 0.3\,\mathrm{MeV}/c^2$ [70], $m_{B^+} = 5279.34 \pm 0.12$ [22]) and 0.03% momentum scale uncertainty [130]. Therefore the difference is assigned a systematic uncertainty of $0.5\,\mathrm{MeV}/c^2$.

Partially reconstructed signal shape: Partially reconstructed signal decay shapes have a systematic uncertainty from the statistical uncertainty on the simulated shape and from the fraction of the yield that is attributed to each of the two overlapping contributions. Although the fractions have been predicted [72, 73, 75], these predictions are not used to avoid strong dependence of the results on theory predictions. In the fit to data, the uncertainty is incorporated by treating the fraction as a free parameter.

When setting confidence limits on the branching fractions with pseudoexperiments, greater model stability is required. For $B_c^+ \to D^*D^*$ decays, a limit is found under the $f_L = 0$ and $f_L = 1$ hypotheses and the least stringent limit is quoted. For $B_c^+ \to D^*D$ decays, separate limits are found for $B_c^+ \to D_{(s)}^{*+}\overset{(-)}{D}{}^0$ and $B_c^+ \to D_{(s)}^{+}\overset{(-)}{D}{}^{*0}$ decays by assuming that the contribution arises entirely from each decay. To account for the statisical uncertainty on the simulated shapes, a small amount of the overlapping decay is allowed to contribute by constraining instead of fixing the fraction. The width of the constraint is equal to the fraction of the overlapping decay that must be added to increase the χ^2 between the original KDE and the simulation admixture by 1σ.

Cross-feed yield: The DCS $B^+ \to DD$ cross-feed yield has an uncertainty from the statistical uncertainty on the RS yield and uncertainty on the D^0 branching fractions.

Table 7.7: Comparison of fitted DSCB tail parameters α_i between data and simulation for $B^+ \to D_s^+ \overline{D}^0$ final states in Run 2.

Final State	α_1 Data	α_1 Simulation	α_2 Data	α_2 Simulation
$KK\pi\ K\pi$	2.29 ± 0.04	2.28 ± 0.08	2.51 ± 0.05	2.53 ± 0.12
$KK\pi\ K3\pi$	2.14 ± 0.04	2.07 ± 0.12	2.38 ± 0.05	2.55 ± 0.17

7.4 Results

7.4.1 Signal significances and branching fractions

The significance (S) of signal is calculated using the results of the extended maximum likelihood fits in Section 7.2.4 and Wilks' theorem [205] as

$$S = \sqrt{2\ln\left(\frac{L_{s+b}}{L_b}\right)}, \tag{7.11}$$

where L_{s+b} and L_s are the maximised extended likelihoods defined in Equation 7.6 under the signal plus background and background only hypotheses, respectively. The significances are given in Table 7.8, including for the Run 1+2 combination, which is determined in a simultaneous fit to the Run 1 and Run 2 dataset. Evidence of $B_c^+ \to D_s^+ \overline{D}^0$ signal is seen in Run 2 data with a significance of 3.7 standard deviations, which is reduced to 3.4 standard deviations in the Run 1+2 combination. This corresponds to a measurement in Run 2 data of

$$R(D_s^+ \overline{D}^0) = (3.6^{+1.5+0.3}_{-1.2-0.2}) \times 10^{-4},$$

where the uncertainties are statistical and systematic, respectively. No other evidence of signal is seen.

The absolute branching fraction can also be calculated using external inputs for the normalisation mode branching fractions [22] and f_c/f_u [106] as

$$\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0) = (4.3^{+1.8+0.4}_{-1.4-0.2} \pm 1.1) \times 10^{-4},$$

where the first and second uncertainties are statistical and internal systematic uncertainties. The third uncertainty is from the external inputs, which is dominated by the uncertainty on the theoretical prediction of $\mathcal{B}(B_c^+ \to J/\psi \mu^- \overline{\nu}_\mu)$ on which f_c/f_u depends. Since f_c/f_u is a function of $\sqrt{s}$, only the absolute branching fraction is measured for the combined Run 1+2 dataset,

$$\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0) = (3.5^{+1.5+0.3}_{-1.3-0.2} \pm 1.0) \times 10^{-4}.$$

Table 7.8: Significances of signal decays calculated using Wilks' theorem.

Decay	Run 1	Run 2	Run 1+2
$B_c^+ \to D_s^+ \overline{D}^0$	0.0	3.7	3.4
$B_c^+ \to D_s^{*+} \overline{D}^0 + D_s^+ \overline{D}^{*0}$	0.0	0.0	0.0
$B_c^+ \to D_s^{*+} \overline{D}^{*0}$	0.0	0.6	0.0
$B_c^+ \to D_s^+ D^0$	0.3	1.5	1.5
$B_c^+ \to D_s^{*+} D^0 + D_s^+ D^{*0}$	0.0	0.8	0.5
$B_c^+ \to D_s^{*+} D^{*0}$	0.0	0.4	0.2
$B_c^+ \to D^+ \overline{D}^0$	0.4	0.5	0.6
$B_c^+ \to D^{*+} \overline{D}^0 + D^+ \overline{D}^{*0}$	0.0	1.1	0.9
$B_c^+ \to D^{*+} \overline{D}^{*0}$	0.0	0.0	0.0
$B_c^+ \to D^+ D^0$	2.3	1.2	2.2
$B_c^+ \to D^{*+} D^0 + D^+ D^{*0}$	0.0	0.7	0.4
$B_c^+ \to D^{*+} D^{*0}$	1.3	0.0	0.0
$B_c^+ \to D^{*+} \overline{D}^0$	1.1	0.7	1.0
$B_c^+ \to D^{*+} \overline{D}^{*0}$	0.2	0.8	0.8
$B_c^+ \to D^{*+} D^0$	0.0	0.0	0.0
$B_c^+ \to D^{*+} D^{*0}$	1.9	1.6	1.8

Comparison with Table 2.1 demonstrates that this measurement is two orders of magnitude larger than SM predictions. Profile likelihood scans of $R(D_s^+ \overline{D}^0)$ in Run 2 data and $\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0)$ in Run 1+2 data are shown in Figure 7.3. Profile likelihood scans for the branching fraction of all decays in Run 1+2 data are given in Appendix C.

Verification of significances

Wilks' theorem holds exactly only in the asymptotic limit. More generally, S is related to the one-sided p-value through

$$S = \sqrt{2}\mathrm{erfc}^{-1}(2p), \tag{7.12}$$

where erfc is the complementary error function

$$\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt, \tag{7.13}$$

and p is the probability that a background-only pseudodataset varies more from the background-only hypothesis than the observed data. The variation from the background-only hypothesis is quantified with the *delta log likelihood,*

$$\Delta LL \equiv \ln L_{s+b} - \ln L_b. \tag{7.14}$$

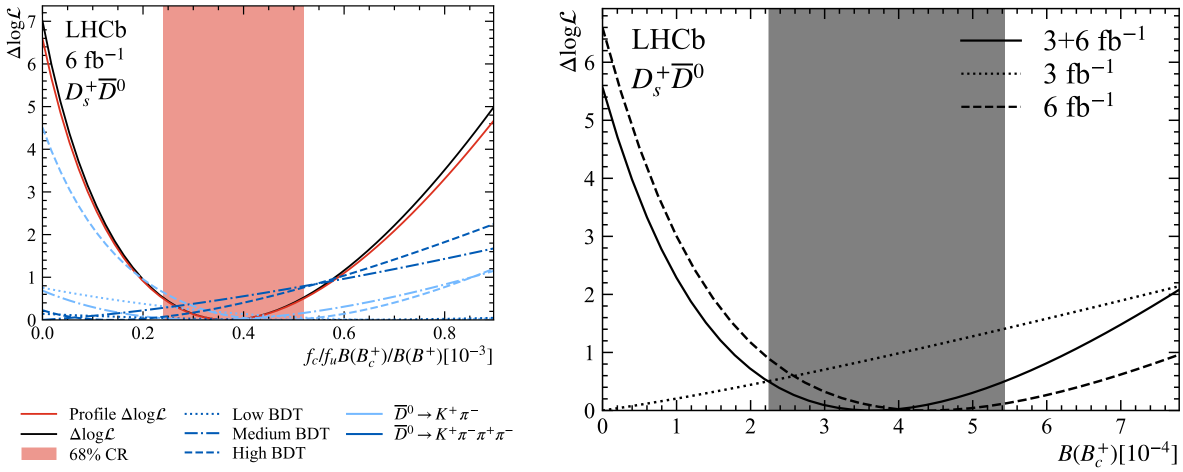

Figure 7.3: (Left) Delta log (black) likelihood and (red) profile likelihood for $R(D_s^+\overline{D}^0)$ in Run 2 data, and (blue) delta log profile likelihoods from individual final states and BDT samples. (Right) Delta log profile likelihood for $\mathcal{B}(B_c^+ \to D_s^+\overline{D}^0)$ in (dotted) Run 1, (dashed) Run 2, and (solid) Run 1+2 data. The 68% confidence interval is shaded in both figures.

To verify the significance of the evidence for $B_c^+ \to D_s^+\overline{D}^0$ signal, background-only pseudodatasets are generated according to the fitted model. A signal plus background model is fitted to each pseudodataset and the distribution of ΔLL in the pseudoexperiments is plotted and compared to the observed $\Delta LL = 6.6$ (Figure 7.4). It is not feasible to generate enough toys to calculate the p-value by counting (none of 5000 generated pseudoexperiments has a ΔLL larger than the observed value), so the p-value is estimated by extrapolating the fit of an exponential to the tail of the distribution ($\Delta LL > 2$). Using this method $S = 3.7\sigma$ is found, in agreement with Wilks' theorem.

7.4.2 Confidence limits on branching fractions

The CLs method

Upper confidence limits are evaluated on the branching fraction of all decays using the CLs method [206] implemented in GammaCombo [207]. The CLs method is a conventional statistical method of excluding a range of signal strengths (μ) at confidence level (C.L.) $1-\alpha$. It is a modified frequentist method that exhibits overcoverage, so that the true value of μ is excluded in fewer than a fraction α of experiments.

The confidence limit on μ is the value of μ for which

$$CL_s < 1-\alpha, \tag{7.15}$$

where

$$CL_s \equiv \frac{CL_{s+b}}{CL_b}, \tag{7.16}$$

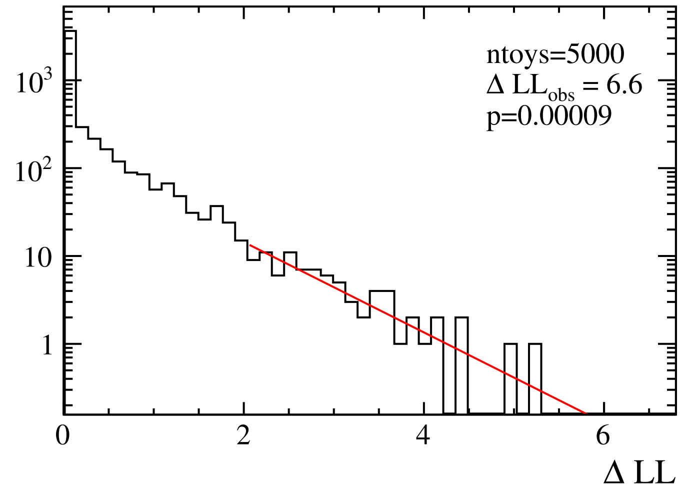

Figure 7.4: ΔLL distributions of toys for $B_c^+ \to D_s^+ \overline{D}^0$ in Run 2 data. The p-value is estimated by fitting an exponential (red line) to the tail of the distribution.

$$CL_{s+b} = \int_{-\infty}^{t_{\mu,\text{obs}}} \frac{dP(t_\mu|\mu)}{dt_\mu} dt_\mu, \tag{7.17}$$

$$CL_b = \int_{-\infty}^{t_{\mu,\text{obs}}} \frac{dP(t_\mu|0)}{dt_\mu} dt_\mu. \tag{7.18}$$

The *test statistic* t_μ is defined to increase monotonically for increasingly signal-like datasets, and $t_{\mu,\text{obs}}$ is the value of the test statistic for the observed data as a function of the signal strength μ. $dP(t_\mu|\mu)/dt_\mu$ is the distribution of t_μ in pseudoexperiments generated with signal strength μ. These pseudoexperiments are generated from the model fitted to the observed data, but with the value of the global observables generated according to their constraint PDF.

A modified likelihood ratio

$$t_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \text{otherwise} \end{cases}, \tag{7.19}$$

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{7.20}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the maximum likelihood estimates of μ and $\boldsymbol{\theta}$, and $\hat{\hat{\boldsymbol{\theta}}}$ is the maximum likelihood estimate of $\boldsymbol{\theta}$ at a given μ, is used as the test statistic. This test statistic is appropriate for the determination of upper limits (as opposed to two-sided confidence intervals) since datasets with $\hat{\mu} > \mu$ are not used to reject μ. Note also that μ and $\hat{\mu}$ are required to be positive, as discussed in Section 7.2.1.

Observed CLs limits are sensitive to fluctuations in the data. Expected CLs limits, assuming that the true signal strength is zero, can be extracted from the expected CL_s

Table 7.9: Upper limits on the branching fraction ratios at 90(95)% C.L. for Run 1 and Run 2 data, in units of 10^{-3}.

R	Run 1	Run 2
$R(D_s^+\overline{D}^0)$	0.45 (0.58)	0.57 (0.62)
$R'_+(D_s^+\overline{D}^0)$	0.45 (0.75)	0.36 (0.42)
$R'_0(D_s^+\overline{D}^0)$	0.64 (0.71)	0.27 (0.36)
$R''(D_s^+\overline{D}^0)$	1.1 (1.5)	0.9 (1.1)
$R(D_s^+D^0)$	0.51 (0.62)	0.22 (0.25)
$R'_+(D_s^+D^0)$	0.76 (0.89)	0.59 (0.72)
$R'_0(D_s^+D^0)$	0.74 (0.88)	0.49 (0.60)
$R''(D_s^+D^0)$	1.6 (2.3)	0.9 (1.0)
$R(D^+\overline{D}^0)$	8 (11)	3.5 (4.4)
$R'_+(D^+\overline{D}^0)$	45 (52)	26 (33)
$R'_0(D^+\overline{D}^0)$	16 (20)	11 (12)
$R''(D^+\overline{D}^0)$	90 (110)	21 (28)
$R(D^+D^0)$	12 (14)	2.9 (3.6)
$R'_+(D^+D^0)$	17 (33)	18 (21)
$R'_0(D^+D^0)$	8.2 (9.8)	6.6 (7.3)
$R''(D^+D^0)$	82 (94)	17 (20)
$R(D^{*+}\overline{D}^0)$	26 (28)	6.9 (8.4)
$R'(D^{*+}\overline{D}^0)$	31 (44)	16 (19)
$R(D^{*+}D^0)$	8 (13)	3.6 (4.4)
$R'(D^{*+}D^0)$	36 (45)	12 (15)

curve, which is the median CL_s value as a function of μ when substituting the observed data with the ensemble of background-only pseudodatasets. The expected CLs limit is useful to denote the expected sensitivity of a search.

Confidence limits on the branching fractions

Upper limits on R, for which the lowest uncertainty is achieved, are evaluated at 90(95)% C.L. (Table 7.9). The limits using Run 1 data, in particular for partially reconstructed decays, are more stringent than in the previous analysis using Run 1 data [171], mostly due to improved constraints on the combinatorial background shape as a result of dividing the data into BDT samples.

Upper limits at 90(95)% C.L. on the absolute branching fractions, which are the most interesting quantity from a physics perspective, are given below. These limits use the Run 2 dataset only, since inclusion of the Run 1 data does not significantly improve sensitivity, and the evaluation of limits corresponding to data from a single operational run is already very computationally demanding. The reported upper limits on decays with a D^{*+} meson use the channel with fully reconstructed $D^{*+} \to D^0\pi^+$ decays because the sensitivity is better than in the channel with a partially reconstructed $D^{*+} \to D^+X^0$ decay. The corresponding observed and expected CL_s curves are shown in Appendix D.

$$\mathcal{B}(B_c^+ \to D_s^+\overline{D}^0) < 7.2\,(8.4) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D_s^+D^0) < 3.0\,(3.7) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D^+\overline{D}^0) < 1.9\,(2.5) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D^+D^0) < 1.4\,(1.8) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D_s^{*+}\overline{D}^0) < 5.3\,(5.7) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D_s^+\overline{D}^{*0}) < 4.6\,(5.6) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D_s^{*+}D^0) < 0.9\,(1.0) \times 10^{-3};$$
$$\mathcal{B}(B_c^+ \to D_s^+D^{*0}) < 6.6\,(8.4) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D^{*+}\overline{D}^0) < 3.8\,(4.8) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D^{*+}D^0) < 2.0\,(2.4) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D^+\overline{D}^{*0}) < 6.5\,(8.2) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D^+D^{*0}) < 3.7\,(4.6) \times 10^{-4};$$
$$\mathcal{B}(B_c^+ \to D_s^{*+}\overline{D}^{*0}) < 1.3\,(1.5) \times 10^{-3};$$
$$\mathcal{B}(B_c^+ \to D_s^{*+}D^{*0}) < 1.3\,(1.6) \times 10^{-3};$$
$$\mathcal{B}(B_c^+ \to D^{*+}\overline{D}^{*0}) < 1.0\,(1.3) \times 10^{-3};$$
$$\mathcal{B}(B_c^+ \to D^{*+}D^{*0}) < 7.7\,(8.9) \times 10^{-4}.$$

Analysis of doubly charmed B^- meson decays

This chapter details the analysis of $B^- \to DD$ decays that was introduced in Section 4.2. Raw asymmetries are obtained and validated in Section 8.1. Production and detection asymmetries are calculated in Section 8.2, and efficiencies are corrected in Section 8.3. Systematic uncertainties are evaluated in Section 8.4. Sections 8.5-8.6 present and validate results on the *CP* asymmetries and branching fraction ratios, respectively.

8.1 Raw asymmetries and yields

Yields and raw charge asymmetries are extracted in the extended maximum likelihood fit of a model to the invariant mass distribution of data candidates in the range $4870 < m(D_{(s)}^{(*)-}D^0) < 5400\,\text{MeV}/c^2$, in bins of width $1\,\text{MeV}/c^2$. The fit is performed simultaneously between the datasets containing B^+ and B^- candidates, sharing shape parameters except where specified in Chapter 6. A separate fit is performed for Run 1 and 2 and for each final state.

During the development of the analysis, the data was charge-blinded by reversing the charge for even run numbers. This ensured that the analysis would be blind to the charge of a specific candidate but that correlations in charge between candidates in the same event would be preserved. The data presented in this chapter is now unblinded.

Section 8.1.1 describes the model. Fit results are presented in Section 8.1.2 and validated in Sections 8.1.3-8.1.4.

8.1.1 Fit model

The invariant mass distribution contains contributions from fully reconstructed $B^- \to DD$ decays (Section 6.1), partially reconstructed $B \to DD$ decays (Section 6.2), partially reconstructed $B \to D^{**}D$ decays (Section 6.3), single-charm background (Section 6.4), $D_s^- \to D^-$ cross-feed (Section 6.5.2), and combinatorial background (Section 6.6).

The charge-summed yield of each fully reconstructed decay and each B^- or B^0 decay with one missing particle is a free parameter. The sum of the yields of all B^- and $B^0 \to D^*D^*$ decays with two missing particles, and the sum of the yields of all $B \to D^{**}D$ decays, are free parameters, with the relative yield from each decay chain fixed as specified in Sections 6.2.2

and 6.3 to ensure fit stability. The yield attributed to the single-charm, cross-feed and B_s^0 decay backgrounds have been described in Sections 6.4, 6.5.2, and 6.2.5, respectively. The combinatorial background yield is a free parameter.

All signal modes have a free $\mathcal{A}^{\rm raw}$. Precision on $\mathcal{A}^{CP}$ is improved by constraining the $\mathcal{A}^{\rm raw}$ of background decays where possible. $\mathcal{A}^{\rm raw}(B^- \to D^{*-}D^{(*)0})$ are constrained using measurements in the $D^{*-}D^0$ channel corrected for $\mathcal{A}^{\rm P} + \mathcal{A}^{\rm D}$ (Section 8.2). $\mathcal{A}^{\rm raw}(B_{(s)}^0 \to D^{*-}D^{*+})$ must be equal to $\mathcal{A}^{\rm D}$.

The asymmetry in $\Gamma(\overset{(-)}{B^0} \to D^-D^{*+})$ oscillates as described in Section 2.3.3. The time-integrated $\mathcal{A}^{\rm raw}$ is dominated by $\mathcal{A}^{CP} = (1.3 \pm 1.4)\%$ [22] plus $\mathcal{A}^{\rm D}$. The time-dependent $\mathcal{A}^{\rm raw}(t)$ also contains

$$\mathcal{A}^{\rm P}(\Delta C \cos \Delta m_d t - \Delta S \sin \Delta m_d t) + \mathcal{O}(0.01\%), \tag{8.1}$$

where t is the B^0 decay time, $\varepsilon(t)$ is the decay-time dependent efficiency, and $\Delta m_d = 0.5065 \pm 0.0019\,\text{ps}^{-1}$ [22] is the B^0 mass eigenstate difference. Therefore the time-integrated asymmetry contains

$$\frac{\int_0^\infty e^{-t/\tau(B^0)}\varepsilon(t)\mathcal{A}^{\rm P}(\Delta C \cos \Delta m_d t - \Delta S \sin \Delta m_d t)dt}{\int_0^\infty e^{-t/\tau(B^0)}\varepsilon(t)dt}, \tag{8.2}$$

where $\tau(B^0) = 1.519 \pm 0.004\,\text{ps}$ [22] is the B^0 lifetime. Since $\tau(B^0)$ is similar to the oscillation period $T = 2\pi/\Delta m_d$, the oscillating asymmetry partially averages to zero when integrating over decay time. Given the values of $\mathcal{A}^{\rm P} \sim \mathcal{O}(1\%)$ [48] and $\Delta C \sim \Delta S \sim \mathcal{O}(0.1)$ [22], the contribution from Equation 8.2 can be neglected relative to $\mathcal{A}^{CP}$.

$\mathcal{A}^{\rm D} + \mathcal{A}^{CP}(B^0 \to D_s^- D^{*+})$ has been measured using the fully reconstructed decay chain at LHCb [48]. This is corrected for the soft pion detection asymmetry. The same analysis measured $\mathcal{A}^{\rm P}(B^0) = (-1.1 \pm 0.9)\%$ in Run 1 and $(0.4 \pm 0.5)\%$ in Run 2, which is diluted through B^0 oscillations by a multiplicative factor $0 < d < 1$. Choosing $d = 0.5 \pm 0.5$ does not lead to a significant systematic uncertainty (Section 8.4), so no attempt is made to determine a more precise value.

$\mathcal{A}^{CP}$ for the remaining double charm background decays have not been measured. Predictions of $\mathcal{A}^{CP}(B^- \to D_s^{*-}D^{*0})$ and $\mathcal{A}^{CP}(B_{(s)}^0 \to D_s^{(*)-}D^{*+})$ [81–85] are not used to avoid a strong dependence of the result on theoretical predictions. Instead the $\mathcal{A}^{\rm raw}$ are treated as free parameters. To ensure fit stability, the $\mathcal{A}^{\rm raw}(B_s^0 \to D_s^{(*)-}D^{*+})$ are fixed to 0 and assigned a systematic uncertainty of $\pm 100\%$.

The $\mathcal{A}^{\rm raw}$ of cross-feed is constrained by measurements in the $D_s^- D^0$ channel. $\mathcal{A}^{CP}(B^- \to [K^-\pi^+]_D K^-\pi^-K^+)$ has not been measured. It is assigned a value of $(0.0 \pm 2.2)\%$, which is double the uncertainty on the measured

Table 8.1: $\mathcal{A}^{\text{raw}}$ in percent and statistical uncertainty.

$\mathcal{A}^{\text{raw}}$	Run	$D^0 \to K\pi$	$D^0 \to K3\pi$	
$\mathcal{A}^{\text{raw}}(B^- \to D_s^- D^0)$	1	-2.0 ± 0.6	-1.6 ± 0.9	
	2	-1.03 ± 0.31	-1.1 ± 0.4	
$\mathcal{A}^{\text{raw}}(B^- \to D_s^{*-} D^0)$	1	1.0 ± 3.1	7 ± 6	
	2	-2.7 ± 1.4	-3.7 ± 2.2	
$\mathcal{A}^{\text{raw}}(B^- \to D_s^- D^{*0})$	1	-5.3 ± 2.3	-2.1 ± 2.5	
	2	0.6 ± 1.2	-0.5 ± 1.4	
$\mathcal{A}^{\text{raw}}(B^0 \to D_s^- D^{*+})$	1	-2.5 ± 1.0	-2.4 ± 1.1	
	2	-1.4 ± 0.7	-1.2 ± 0.7	
		$D^0 \to K\pi$	$D^0 \to K3\pi$	
$\mathcal{A}^{\text{raw}}(B^- \to D^- D^0)$	1	0.5 ± 2.9	4 ± 4	
	2	0.6 ± 1.4	2.9 ± 1.9	
$\mathcal{A}^{\text{raw}}(B^- \to D^- D^{*0})$	1	1 ± 5	$-\ 2 \pm 10$	
	2	-0.9 ± 2.8	-3.2 ± 3.3	
$\mathcal{A}^{\text{raw}}(B^- \to D^{*-} D^0)$	1	2.7 ± 1.9	2.4 ± 1.9	
	2	2.2 ± 2.0	2.1 ± 2.0	
$\mathcal{A}^{\text{raw}}(B^0 \to D^- D^{*+})$	1	1.9 ± 1.4	1.6 ± 1.4	
	2	1.3 ± 1.4	1.3 ± 1.4	
		$D^0 \to K\pi$ $D^0 \to K\pi$	$D^0 \to K\pi$ $D^0 \to K3\pi$	$D^0 \to K3\pi$ $D^0 \to K\pi$
$\mathcal{A}^{\text{raw}}(B^- \to D^{*+} D^0)$	1	9 ± 6	-4 ± 7	3 ± 7
	2	2.0 ± 2.8	2.2 ± 3.3	1.9 ± 3.3
$\mathcal{A}^{\text{raw}}(B^- \to D^{*+} D^{*0})$	1	-8 ± 7	-1 ± 9	5 ± 9
	2	4.7 ± 3.5	-0 ± 4	-2 ± 4
$\mathcal{A}^{\text{raw}}(B^0 \to D^{*+} D^{*-})$	1	0.5 ± 0.6	0.2 ± 0.6	0.6 ± 0.7
	2	0.14 ± 0.23	-0.02 ± 0.28	0.23 ± 0.22

$\mathcal{A}^{CP}(B^- \to [K^-\pi^+]_D \pi^-\pi^+\pi^-) = (-0.2 \pm 1.1)\%$ [208], as these decays have the same flavour structure. The raw asymmetry of the sum of all $B \to D^{**}D$ decay yields and the combinatorial background are free parameters.

8.1.2 Fit results

The invariant mass distributions of $B^\pm$ candidates, and the fitted model, are plotted in Figure 8.1. The $\mathcal{A}^{\text{raw}}$ and signal yields obtained from these fits are presented in Table 8.1 and Table 8.2, respectively. $\mathcal{A}^{\text{raw}}(B^- \to D_s^- D^{*0})$ and $\mathcal{A}^{\text{raw}}(B^- \to D_s^{*-} D^0)$ have a -65% correlation as the invariant mass distributions overlap, and otherwise correlations between the $\mathcal{A}^{\text{raw}}$ of decays in the same channel are no more than 10%.

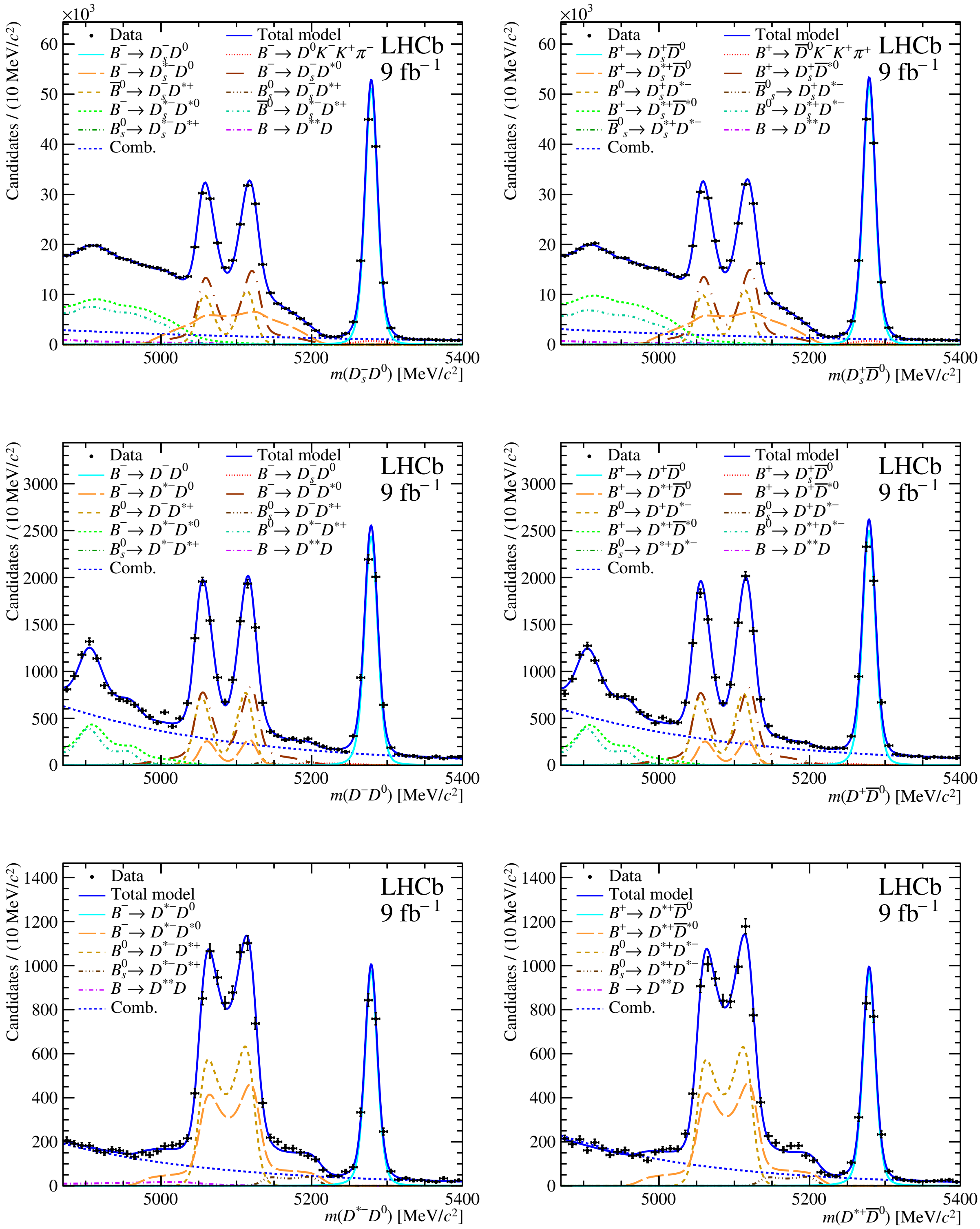

Figure 8.1: Invariant mass distribution of (top row) $B^-\to D_s^-D^0$ candidates, (middle row) $B^-\to D^-D^0$ candidates, and (bottom row) $B^-\to D^{*-}D^0$ candidates, separated into (left) B^- and (right) B^+ candidates. The fitted model described in the text is also drawn. The data and model is a sum over all datasets [2].

Table 8.2: Charge-summed fitted yields of signal candidates selected for the measurement of $\mathcal{A}^{CP}$, summed over Run 1 and 2 and over D^0 meson final states.

Decay	Yield
$B^- \to D_s^- D^0$	$(2.305 \pm 0.005) \times 10^5$
$B^- \to D_s^{*-} D^0$	$(1.905 \pm 0.030) \times 10^5$
$B^- \to D_s^- D^{*0}$	$(2.14 \pm 0.04) \times 10^5$
$B^- \to D^- D^0$	$(1.149 \pm 0.012) \times 10^4$
$B^- \to D^- D^{*0}$	$(1.22 \pm 0.08) \times 10^4$
$B^- \to D^{*-} D^0$	$(2.9 \pm 0.6) \times 10^3$
$B^- \to D^{*-} D^0$	$(4.41 \pm 0.07) \times 10^3$
$B^- \to D^{*-} D^{*0}$	$(8.5 \pm 0.6) \times 10^3$

8.1.3 Validation of modelling of partially reconstructed decays

To validate the modelling of partially reconstructed decays with one missing soft particle $X_s = \{\gamma, \pi^0, \pi^+\}$ from a D^*, their branching fractions relative to the fully reconstructed decays,

$$\begin{aligned} R(D^*D) &= \frac{\mathcal{B}(B \to D^*D)\mathcal{B}(D^* \to DX_s)}{\mathcal{B}(B^- \to D^0 D)} \\ &= \frac{N(B \to (D^* \to DX_s)D)}{N(B^- \to D^0 D)} \frac{\varepsilon(B^- \to D^0 D)}{\varepsilon(B \to (D^* \to DX_s)D)}, \end{aligned} \tag{8.3}$$

are measured. The efficiency ratio is estimated from the gradient $d\varepsilon/dm$ of the linear efficiency slope obtained in the fit to data,

$$\frac{\varepsilon(B \to (D^* \to DX_s)D)}{\varepsilon(B^- \to DD)} = 1 - \frac{d\varepsilon}{dm}\left(m_{B^-} - \frac{a+b}{2}\right), \tag{8.4}$$

where a and b are defined in Section 6.2.1. Since $B^0 \to D^{*-}D^{*+}$ candidates in the $D^{*-}D^0$ channel can be reconstructed twice, the efficiency is doubled.

All systematic uncertainties considered for $\mathcal{A}^{\text{raw}}$ (Section 8.4) are considered in this validation. An additional systematic uncertainty to account for non-linearity of the efficiency slope is evaluated as the change in the branching fraction when the efficiency gradient for decays with two instead of one missing particle is used. Finally, the statistical uncertainty on the efficiency gradient is propagated to an uncertainty on the branching fraction ratio. The non-linearity of the efficiency slope forms the dominant systematic uncertainty.

The measured values of $R(D^*D)$ are presented in Table 8.3, including the combined measurement using all datasets. All measurements are consistent with expectations (Table 8.4) and between datasets. This combination requires the correlation in uncertainties between datasets to be specified. These correlations are not well known for many sources of uncertainty, including the dominant systematic uncertainties from the description of the efficiency slope and combinatorial background, so the correlations that represent the least information in the covariance matrix must be chosen. Assuming 100% correlations on systematic uncertainties results in the least information only if the correlated uncertainties are smaller than uncorrelated uncertainties [209]. To obtain the covariance matrix with the least information, each covariance matrix element σ_{ij} is required to be no more than than both σ_{ii} and σ_{jj} [209].

8.1.4 Bias and coverage

This section describes the implementation of pseudoexperiments to assess the bias and coverage properties of the $\mathcal{A}^{\text{raw}}$ measurements, and discusses the results of these pseudoexperiments. Generating pseudodata requires an understanding of the correlations between candidates, including knowledge of the fraction of $B^0 \to D^{*+}D^{*-}$ decays producing double candidates, which will be evaluated first.

Counting of duplicated $B^0 \to D^{*+}D^{*-}$ decays

The fraction of selected $B^0 \to D^{*-}D^{*+}$ decays that produce two candidates is evaluated by fitting to two subsets of the data. The first subset contains all pairs of candidates that differ only by whether the soft pion is assigned to the D^0 or $\overline{D}^0$. In the second subset, these pairs are removed.

If both soft pions are reconstructed, the average momentum of the decay excluding one soft pion is lower than if only one soft pion is reconstructed, which creates a negative correlation between the reconstructed mass and efficiency. Therefore, in the first subset, the efficiency gradient is allowed to take the opposite sign, and in the second subset, the B^+ and $B^0_{(s)}$ decays are allowed to have different efficiency slopes.

Exemplar fits to each subset are shown in Figure 8.2. The number of $B^0 \to D^{*-}D^{*+}$ decays producing a single candidate, N_{single} and the number producing a double candidate, N_{double} are extracted from these fits and used to calculate the fraction of selected $B^0 \to D^{*-}D^{*+}$ decays that produce two candidates, f_{double} (Table 8.5).

Pseudoexperiments

Pseudodatasets are generated with one signal $\mathcal{A}^{\text{raw}}$ set to 0%, 2% or 10%, and all other unconstrained asymmetries set to 0%. Global observables, which are the parameters that are constrained using external inputs, are generated randomly according to their Gaussian PDFs. All other parameters are set to the parameters obtained from the fit to data.

Toy datasets are initially generated according to Poisson statistics with $N(B^0_{(s)} \to D^{*-}D^{*+})$ equal to the number of decays with only one candidate. Then the number of decays with two

Table 8.3: Measured values of R with statistical and systematic uncertainties.

R	Final state	Run 1	Run 2	Run 1+2
$D_s^- D^{*0}$	$KK\pi, K\pi$	$1.20 \pm 0.06 \pm 0.41$	$1.15 \pm 0.02 \pm 0.26$	$1.15 \pm 0.02 \pm 0.26$
	$KK\pi, K3\pi$	$1.45 \pm 0.10 \pm 0.44$	$1.23 \pm 0.04 \pm 0.25$	$1.23 \pm 0.04 \pm 0.25$
	combination	$1.23 \pm 0.06 \pm 0.41$	$1.22 \pm 0.04 \pm 0.25$	$1.22 \pm 0.04 \pm 0.25$
$D_s^{*-} D^0$	$KK\pi, K\pi$	$1.08 \pm 0.04 \pm 0.36$	$1.08 \pm 0.03 \pm 0.25$	$1.08 \pm 0.03 \pm 0.25$
	$KK\pi, K3\pi$	$0.85 \pm 0.09 \pm 0.26$	$1.03 \pm 0.03 \pm 0.20$	$1.02 \pm 0.03 \pm 0.20$
	combination	$0.87 \pm 0.08 \pm 0.26$	$1.03 \pm 0.03 \pm 0.21$	$1.03 \pm 0.03 \pm 0.21$
$D_s^- D^{*+}$	$KK\pi, K\pi$	$0.77 \pm 0.04 \pm 0.27$	$0.71 \pm 0.02 \pm 0.17$	$0.71 \pm 0.02 \pm 0.17$
	$KK\pi, K3\pi$	$0.54 \pm 0.06 \pm 0.17$	$0.61 \pm 0.03 \pm 0.12$	$0.60 \pm 0.03 \pm 0.12$
	combination	$0.56 \pm 0.06 \pm 0.17$	$0.61 \pm 0.03 \pm 0.13$	$0.61 \pm 0.03 \pm 0.13$
$D^- D^{*0}$	$K\pi\pi, K\pi$	$1.30 \pm 0.26 \pm 0.28$	$1.29 \pm 0.12 \pm 0.16$	$1.29 \pm 0.11 \pm 0.16$
	$K\pi\pi, K3\pi$	$1.09 \pm 0.42 \pm 0.48$	$1.27 \pm 0.16 \pm 0.28$	$1.26 \pm 0.15 \pm 0.28$
	combination	$1.26 \pm 0.23 \pm 0.28$	$1.29 \pm 0.10 \pm 0.17$	$1.29 \pm 0.10 \pm 0.16$
$D^{*-} D^0$	$K\pi\pi, K\pi$	$0.29 \pm 0.19 \pm 0.07$	$0.30 \pm 0.08 \pm 0.05$	$0.30 \pm 0.08 \pm 0.05$
	$K\pi\pi, K3\pi$	$0.38 \pm 0.29 \pm 0.16$	$0.42 \pm 0.12 \pm 0.09$	$0.42 \pm 0.11 \pm 0.09$
	combination	$0.32 \pm 0.16 \pm 0.08$	$0.33 \pm 0.07 \pm 0.05$	$0.33 \pm 0.06 \pm 0.05$
$D^- D^{*+}$	$K\pi\pi, K\pi$	$1.06 \pm 0.20 \pm 0.19$	$0.93 \pm 0.09 \pm 0.07$	$0.94 \pm 0.08 \pm 0.08$
	$K\pi\pi, K3\pi$	$0.79 \pm 0.27 \pm 0.28$	$0.76 \pm 0.12 \pm 0.14$	$0.76 \pm 0.11 \pm 0.15$
	combination	$0.99 \pm 0.16 \pm 0.20$	$0.89 \pm 0.07 \pm 0.08$	$0.91 \pm 0.07 \pm 0.08$
$D^{*-} D^{*0}$	$K\pi\pi_s, K\pi$	$4.07 \pm 0.64 \pm 0.82$	$2.49 \pm 0.67 \pm 1.02$	$3.47 \pm 0.47 \pm 0.79$
	$K\pi\pi_s, K3\pi$	$2.53 \pm 0.72 \pm 1.53$	$3.54 \pm 0.48 \pm 0.65$	$3.46 \pm 0.44 \pm 0.66$
	$K3\pi\pi_s, K\pi$	$2.82 \pm 0.92 \pm 0.83$	$2.44 \pm 0.31 \pm 1.08$	$2.60 \pm 0.42 \pm 0.90$
	combination	$3.51 \pm 0.49 \pm 0.69$	$3.25 \pm 0.36 \pm 0.68$	$3.44 \pm 0.32 \pm 0.62$
$D^{*-} D^{*+}$	$K\pi\pi_s, K\pi$	$1.83 \pm 0.28 \pm 0.38$	$1.36 \pm 0.21 \pm 0.53$	$1.66 \pm 0.20 \pm 0.38$
	$K\pi\pi_s, K3\pi$	$1.21 \pm 0.30 \pm 0.68$	$1.49 \pm 0.20 \pm 0.29$	$1.47 \pm 0.18 \pm 0.30$
	$K3\pi\pi_s, K\pi$	$1.42 \pm 0.39 \pm 0.38$	$0.92 \pm 0.13 \pm 0.41$	$1.08 \pm 0.15 \pm 0.36$
	combination	$1.63 \pm 0.22 \pm 0.32$	$1.35 \pm 0.14 \pm 0.30$	$1.46 \pm 0.13 \pm 0.28$

Table 8.4: Expected R values [22].

R	Expectation
$D_s^- D^{*0}$	0.91 ± 0.21
$D_s^{*-} D^0$	0.84 ± 0.20
$D_s^- D^{*+}$	0.60 ± 0.10
$D^- D^{*0}$	1.7 ± 0.5
$D^{*-} D^0$	0.33 ± 0.06
$D^- D^{*+}$	1.09 ± 0.16
$D^{*-} D^{*0}$	2.1 ± 0.5
$D^{*-} D^{*+}$	1.39 ± 0.21

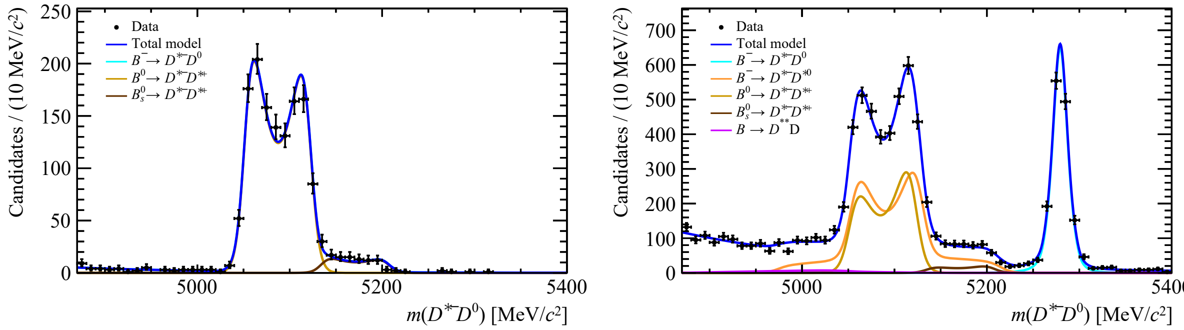

Figure 8.2: Fit to $B \to D^{*-}D^0$, $D^0 \to K^-\pi^+$, $\bar{D}^0 \to K^+\pi^-$ candidates in Run 2 data. For (left) the subset with only $B \to D^{*+}D^{*-}$ candidates and (right) with duplicated $B \to D^{*+}D^{*-}$ candidates removed. The fitted model is also drawn and is described in the text.

candidates is generated according to Poisson statistics, and toy datasets for $B^0_{(s)} \to D^{*-}D^{*+}$ decays with this generated yield are generated with $\mathcal{A}^{\text{raw}} = 0$.

The model is then fitted to the dataset. The bias $x - x_{gen}$ and pull $(x - x_{gen})/\sigma$, where x is the fitted value, x_{gen} is the generated value and σ is the uncertainty on x, are obtained for each parameter of interest. 1000 pseudoexperiments are performed for each set of generation parameters. The central value bias, and the mean (μ) and standard deviation (σ) of the pull distributions, are given in Tables E.1-E.6. Any bias is at least five times smaller than statistical uncertainties and can therefore be neglected. Some overcoverage is observed, in part due to the presence of double $B^0 \to D^{*-}D^{*+}$ candidates. This overcoverage is no larger than 20% and no correction is applied.

8.2 Charge asymmetry corrections

Detection asymmetries (Equation 4.5) result from the asymmetry of nuclear interactions of final state particles with the matter in the detector. The L0 trigger requirements, track reconstruction, and PID requirements also induce detection asymmetries. Furthermore there is an asymmetry in the production of B^- mesons (Equation 4.4). These asymmetries

Table 8.5: The number of $B^0 \to D^{*-}D^{*+}$ decays producing a single candidate, N_{single}; the number producing a double candidate, N_{double}; and the fraction of selected $B^0 \to D^{*-}D^{*+}$ decays that produce two candidates, f_{double}.

Period	Final State	N_{single}	N_{double}	f_{double}
Run 1	$K\pi, K\pi$	511 ± 69	133 ± 11	0.207 ± 0.026
	$K\pi, K3\pi$	401 ± 53	75 ± 8	0.158 ± 0.023
	$K3\pi, K\pi$	289 ± 88	79 ± 9	0.22 ± 0.05
Run 2	$K\pi, K\pi$	1766 ± 247	646 ± 25	0.268 ± 0.028
	$K\pi, K3\pi$	1431 ± 137	463 ± 21	0.244 ± 0.020
	$K3\pi, K\pi$	968 ± 220	454 ± 21	0.32 ± 0.05

must be evaluated to correct $\mathcal{A}^{\text{raw}}$ to obtain $\mathcal{A}^{CP}$ (Equation 4.6). This section will explain how each contribution is evaluated.

8.2.1 Production asymmetry

The asymmetry of the initial pp state induces a production asymmetry that is expected to vary as a function of $\sqrt{s}$, p_{T}, and y [210]. LHCb has measured $\mathcal{A}^{\text{P}}(B^-)$ in pp collisions at $\sqrt{s} = 7, 8, 13\,\text{TeV}$ as a function of $(p_{\text{T}}(B^-), y(B^-))$ by correcting the raw asymmetry in $B^+ \to J/\psi K^+$ decays for the detection and known $\mathcal{A}^{CP}$ (Figures 8.3-8.4) [211, 212]. The dominant uncertainties are from the finite size of the calibration samples and a $\pm 0.3\%$ uncertainty on $\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$ [22].

Corrections to $\mathcal{A}^{CP}$ (Table 8.6) are obtained by weighting the binned measurements by the distributions of simulated candidates passing all selection and within $\pm 40\,\text{MeV}/c^2$ of m_{B^-}. The correction for partially reconstructed candidates differs by approximately 0.03%.

Table 8.6: $\mathcal{A}^{\text{P}}(B^-)$ in percent and statistical plus systematic uncertainty, excluding the uncertainty on $\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$.

Channel	Final State	7 TeV	8 TeV	13 TeV
$B^- \to D_s^- D^0$	$KK\pi, K\pi$	-0.7 ± 0.6	-1.3 ± 0.4	-1.1 ± 0.5
$B^- \to D_s^- D^0$	$KK\pi, K3\pi$	-0.7 ± 0.7	-1.5 ± 0.5	-1.2 ± 0.5
$B^- \to D^- D^0$	$K\pi\pi, K\pi$	-0.7 ± 0.6	-1.3 ± 0.4	-1.1 ± 0.5
$B^- \to D^- D^0$	$K\pi\pi, K3\pi$	-0.8 ± 0.7	-1.4 ± 0.5	-1.2 ± 0.5
$B^- \to D^{*-} D^0$	$K\pi\pi_s, K\pi$	-0.6 ± 0.6	-1.4 ± 0.5	-1.1 ± 0.5
$B^- \to D^{*-} D^0$	$K\pi\pi_s, K3\pi$	-0.9 ± 0.8	-1.5 ± 0.5	-1.2 ± 0.5
$B^- \to D^{*-} D^0$	$K3\pi\pi_s, K\pi$	-0.8 ± 0.7	-1.5 ± 0.5	-1.1 ± 0.5

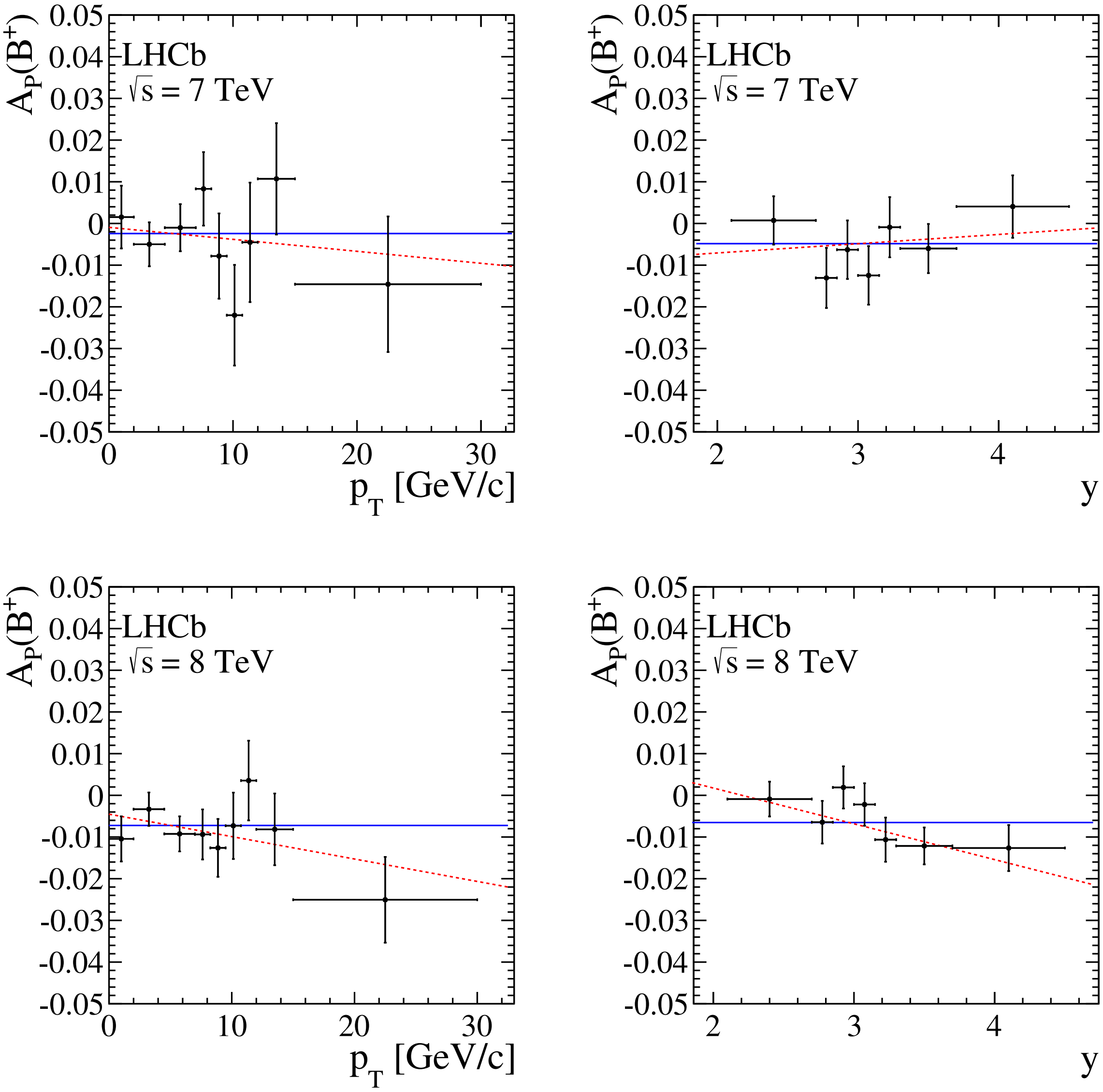

Figure 8.3: $\mathcal{A}^{\mathrm{P}}(B^-)$ as a function of $p_{\mathrm{T}}(B^-)$ (left) and $y(B^-)$ (right) at (top) $\sqrt{s} = 7$ TeV and (bottom) 8 TeV [211]. Uncertainties are both statistical and systematic, and a 0.3% uncertainty from $\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$ is correlated between all bins.

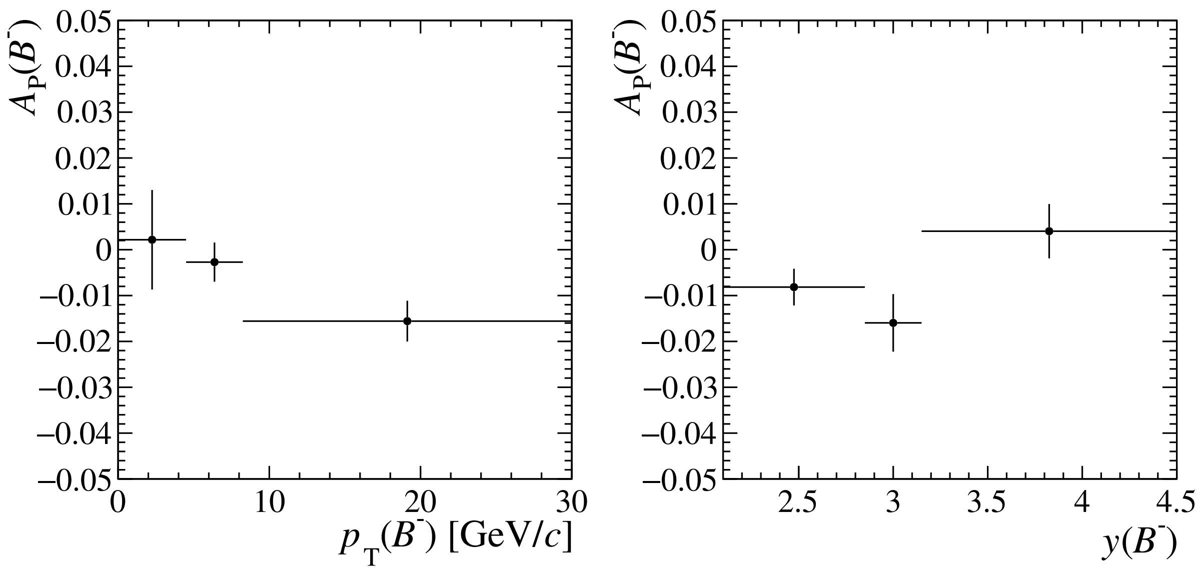

Figure 8.4: $\mathcal{A}^{\mathrm{P}}(B^-)$ as a function of $p_{\mathrm{T}}(B^-)$ (left) and $y(B^-)$ (right) at $\sqrt{s} = 13\,\mathrm{TeV}$ [212]. Uncertainties are both statistical and systematic, and a 0.3% uncertainty from $\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$ is correlated between all bins.

8.2.2 Hardware trigger global TIS asymmetry

The requirement that the event excluding the signal candidate passes an L0 trigger introduces a *CP* asymmetry because the event always contains a b hadron of opposite flavour to the signal candidate. The asymmetry

$$\mathcal{A}^{\mathrm{TIS}} = \frac{\varepsilon_{\mathrm{TIS}}(B^-) - \varepsilon_{\mathrm{TIS}}(B^+)}{\varepsilon_{\mathrm{TIS}}(B^-) + \varepsilon_{\mathrm{TIS}}(B^+)}, \tag{8.5}$$

where $\varepsilon_{\mathrm{TIS}}(B^-)$ is the efficiency for a candidate to be TIS to any L0 trigger, has been measured as a function of $p_{\mathrm{T}}(B^-)$ by LHCb at $\sqrt{s} = 7, 8\,\mathrm{TeV}$ using $B^- \to D^0 \mu^- \overline{\nu}_\mu$ decays that passed the L0 muon trigger [213]. The p_{T}-integrated asymmetry was consistent with zero.

$\mathcal{A}^{\mathrm{TIS}}$ in signal is obtained by weighting these measurements by the p_{T} distribution of selected candidates passing the L0 Global TIS requirement within $\pm 40\,\mathrm{MeV}/c^2$ of m_{B^-}. The 8 TeV measurement is used to obtain a correction for Run 2 data. The resulting $\mathcal{A}^{\mathrm{TIS}}$ values are are given in Table 8.7. $\mathcal{A}^{CP}$ is corrected by $\mathcal{A}^{\mathrm{TIS}}$ multiplied by the fraction of selected candidates in background subtracted data passing the L0 Global TIS requirement (approximately 60%).

8.2.3 Hardware trigger hadron TOS asymmetry

The efficiency of requiring that a hadron fires the L0 Hadron trigger depends to first order on the true E_{T} at the HCAL surface and whether the energy deposit is in the inner or outer

Table 8.7: $\mathcal{A}^{\mathrm{TIS}}$ in percent for signal candidates passing the L0 Global TIS requirement.

Final State	Polarity	2011	2012	2015-18
$B^- \to D_s^- D^0$	Up	0.10 ± 0.34	0.06 ± 0.20	0.05 ± 0.21
$KK\pi$	Down	0.31 ± 0.28	-0.14 ± 0.20	-0.17 ± 0.21
$K\pi$	Combined	0.22 ± 0.18	-0.04 ± 0.14	-0.06 ± 0.15
$B^- \to D_s^- D^0$	Up	0.12 ± 0.42	0.09 ± 0.25	0.08 ± 0.26
$KK\pi$	Down	0.58 ± 0.36	-0.17 ± 0.25	-0.22 ± 0.26
$K3\pi$	Combined	0.38 ± 0.34	-0.04 ± 0.18	-0.07 ± 0.19
$B^- \to D^- D^0$	Up	0.11 ± 0.35	0.05 ± 0.20	0.05 ± 0.21
$K\pi\pi$	Down	0.34 ± 0.29	-0.15 ± 0.20	-0.18 ± 0.21
$K\pi$	Combined	0.24 ± 0.20	-0.05 ± 0.14	-0.06 ± 0.15
$B^- \to D^- D^0$	Up	0.10 ± 0.45	0.09 ± 0.26	0.08 ± 0.27
$K\pi\pi$	Down	0.60 ± 0.38	-0.20 ± 0.27	-0.23 ± 0.27
$K3\pi$	Combined	0.38 ± 0.34	-0.06 ± 0.19	-0.07 ± 0.19
$B^- \to D^{*-} D^0$	Up	0.05 ± 0.38	0.07 ± 0.22	0.06 ± 0.24
$K\pi\pi_s$	Down	0.40 ± 0.33	-0.18 ± 0.23	-0.22 ± 0.24
$K\pi$	Combined	0.25 ± 0.23	-0.06 ± 0.16	-0.08 ± 0.17
$B^- \to D^{*-} D^0$	Up	0.07 ± 0.49	0.08 ± 0.29	0.08 ± 0.28
$K\pi\pi_s$	Down	0.65 ± 0.42	-0.25 ± 0.29	-0.25 ± 0.28
$K3\pi$	Combined	0.40 ± 0.37	-0.08 ± 0.20	-0.09 ± 0.20
$B^- \to D^{*-} D^0$	Up	0.08 ± 0.42	0.08 ± 0.25	0.07 ± 0.26
$K3\pi\pi_s$	Down	0.52 ± 0.36	-0.18 ± 0.25	-0.24 ± 0.27
$K\pi$	Combined	0.33 ± 0.30	-0.05 ± 0.18	-0.08 ± 0.19

HCAL. This dependence, which is plotted in Section 3.2.3, is measured separately for kaons and pions of each charge as the fraction of $D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$ decays passing the L0 Global TIS requirement where the kaon or pion also fires the L0 hadron trigger.

At higher order, overlapping energy deposits from other hadrons in the candidate increase the efficiency. The additional E_{T} from overlap is estimated by modelling the shower shape as a double Gaussian and the fraction of tracks not reaching the HCAL as an exponential function of E_{T}. The shower shape parameters and the decay constant are derived from simulation.

The efficiency measurements for Run 1 [149] do not account for this effect, so the measured efficiency curve incorporates the average E_{T} overlap from the remainder of the $D^{*+} \to D^0\pi^+$ candidate. Consequently the measured efficiency curves in Run 1 are higher than the true efficiency curves for a single charged hadron. Therefore this analysis does not correct for overlap when calculating efficiencies for Run 1, and instead assigns a systematic uncertainty to cover the effect of this correction.

The underlying event produces a small amount of additional E_{T} that has been measured

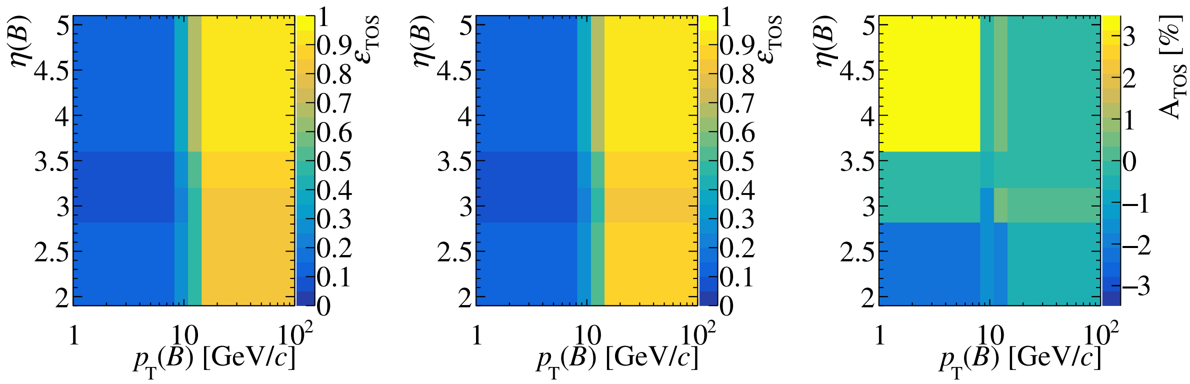

Figure 8.5: (Left) $\varepsilon_{\mathrm{TOS}}(B^-)$, (centre) $\varepsilon_{\mathrm{TOS}}(B^+)$, and (right) $\mathcal{A}^{\mathrm{TOS}}$, as a function of $(p_{\mathrm{T}}(B), \eta(B))$. For $B^- \to D_s^- D^0$, $D^0 \to K^- \pi^+$ decays in 2016 MagUp data.

in Run 2 only as a function of position in the HCAL, using events passing the L0 Muon trigger. Finally, the relative miscalibration of each HCAL cell is derived for Run 2 by comparing the location of the turn-on in the efficiency curve in each cell to the known E_{T} threshold, and efficiencies are adjusted accordingly.

The candidate efficiency is the logical OR of the efficiencies ε_i of the tracks in the candidate,

$$\varepsilon_{\mathrm{TOS}}(B) = 1 - \prod_i (1 - \varepsilon_i). \tag{8.6}$$

The asymmetry

$$\mathcal{A}^{\mathrm{TOS}} = \frac{\varepsilon_{\mathrm{TOS}}(B^-) - \varepsilon_{\mathrm{TOS}}(B^+)}{\varepsilon_{\mathrm{TOS}}(B^-) + \varepsilon_{\mathrm{TOS}}(B^+)} \tag{8.7}$$

in signal is calculated in bins of $(p_{\mathrm{T}}(B), \eta(B))$ (Figure 8.5) by applying this method to simulated candidates passing the L0 Global TIS requirement, in addition to all selection. Each dimension is divided into two (four) bins containing equal yield for Run 1(2) simulation. $\mathcal{A}^{\mathrm{TOS}}$ for signal candidates passing the L0 Hadron TOS requirement is calculated by weighting this measurement by the distribution of selected simulation candidates passing the L0 Hadron TOS decision. $\mathcal{A}^{\mathrm{TOS}}$ differs by less than 0.01% for partially reconstructed decays. $\mathcal{A}^{CP}$ is corrected by $\mathcal{A}^{\mathrm{TOS}}$ multiplied by the fraction of selected candidates in background subtracted data passing the L0 Hadron TOS requirement (approximately 60%). No calibration data is available for 2015, so 2016 data is used instead.

Table 8.8 shows $\mathcal{A}^{\mathrm{TOS}}$ for each final state and magnet polarity in 2012 and 2016. The quoted uncertainties are from the finite sizes of the calibration and signal simulation samples.

The uncertainty from the simulation sample size is large, up to 0.5%, when considering only a single magnet polarity. The magnetic field shifts the x-component of the momentum by approximately 1.2 GeV/c. This introduces a large shift in the efficiency, which is in the opposite direction depending on whether a B^- or B^+ hypothesis is

Table 8.8: $\mathcal{A}^{\text{TOS}}$ in percent and uncertainty is from the finite size of the simulation and calibration samples.

Year	Final state	Up	Down	Combined
2012	$KK\pi, K\pi$	-0.13 ± 0.13	-0.08 ± 0.13	-0.10 ± 0.02
	$KK\pi, K3\pi$	-0.34 ± 0.20	0.15 ± 0.20	-0.10 ± 0.01
	$K\pi\pi, K\pi$	0.04 ± 0.16	-0.03 ± 0.16	0.01 ± 0.02
	$K\pi\pi, K3\pi$	-0.14 ± 0.27	0.17 ± 0.27	0.01 ± 0.01
	$K\pi\pi_s, K\pi$	0.11 ± 0.25	-0.07 ± 0.25	0.02 ± 0.01
	$K\pi\pi_s, K3\pi$	0.13 ± 0.31	-0.10 ± 0.31	0.01 ± 0.02
	$K3\pi\pi_s, K\pi$	0.40 ± 0.37	-0.39 ± 0.37	0.01 ± 0.04
2016	$KK\pi, K\pi$	-0.22 ± 0.20	-0.10 ± 0.19	-0.16 ± 0.03
	$KK\pi, K3\pi$	-0.46 ± 0.34	0.15 ± 0.33	-0.15 ± 0.03
	$K\pi\pi, K\pi$	-0.21 ± 0.20	0.02 ± 0.20	-0.10 ± 0.06
	$K\pi\pi, K3\pi$	-0.24 ± 0.39	0.12 ± 0.39	-0.06 ± 0.05
	$K\pi\pi_s, K\pi$	-0.14 ± 0.23	0.19 ± 0.22	0.02 ± 0.03
	$K\pi\pi_s, K3\pi$	-0.35 ± 0.46	0.38 ± 0.44	0.01 ± 0.04
	$K3\pi\pi_s, K\pi$	-0.59 ± 0.45	0.45 ± 0.44	-0.07 ± 0.04

considered, and therefore a large charge asymmetry. When reversing the magnetic field, trajectories are bent in the opposite direction. Therefore the uncertainty associated with the finite size of the simulation sample is strongly anticorrelated between magnet polarities. Therefore, when averaging over magnet polarities, this uncertainty mostly cancels, and the uncertainty associated with the calibration sample size becomes dominant. An example of this effect is shown in Figure 8.6.

8.2.4 Kaon-pion detection asymmetry

The interaction cross-section of hadrons with matter has a momentum-dependent charge asymmetry. Due to variations in material distribution, the asymmetry in detector interactions also depends on pseudorapidity. This asymmetry is particularly large for kaons, and smaller but non-zero for pions [22, 214].

The best precision on the kaon detection asymmetry is achieved by measuring the detection asymmetry of a $K^-\pi^+$ pair,

$$\mathcal{A}^{\text{K}\pi} = \frac{\varepsilon(K^-\pi^+) - \varepsilon(K^+\pi^-)}{\varepsilon(K^-\pi^+) + \varepsilon(K^+\pi^-)}, \tag{8.8}$$

in the difference of the raw asymmetries of the CF $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to \overline{K}^0\pi^+$ decays [214–216]. These raw asymmetries are decomposed as

$$\mathcal{A}^{\text{raw}}(D^+ \to K^-\pi^+\pi^+_{\text{trig}}) = \mathcal{A}^{\text{P}}(D^+) + \mathcal{A}^{\text{D}}(K^-\pi^+) + \mathcal{A}^{\text{D}}(\pi^+_{\text{trig}}), \tag{8.9}$$

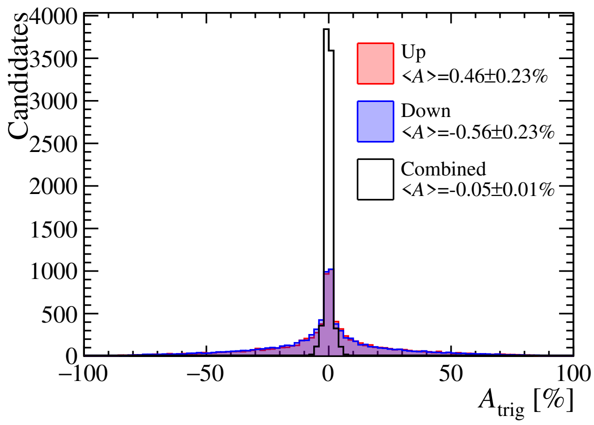

Figure 8.6: Distribution of $\mathcal{A}^{\mathrm{TOS}}$ weighted by $\varepsilon_{\mathrm{TOS}}$ in simulated $B^- \to D^- D^0$, $D^0 \to K3\pi$ candidates in 2016. For each magnet polarity, the average asymmetry is $\mathcal{O}(0.1\%)$ but consistent with zero. The polarity-averaged asymmetry is $\mathcal{O}(0.01\%)$. The quoted uncertainty is from only the finite size of the simulation sample.

and

$$\mathcal{A}^{\mathrm{raw}}(D^+ \to \overline{K}^0 \pi^+_{\mathrm{trig}}) = \mathcal{A}^{\mathrm{P}}(D^+) + \mathcal{A}^{\mathrm{raw}}(\overline{K}^0) + \mathcal{A}^{\mathrm{D}}(\pi^+_{\mathrm{trig}}), \tag{8.10}$$

where π^+_{trig} triggered the HLT1 line that selects events with a good quality track. If the D^+ and π^+_{trig} kinematics match between the two calibration samples, then the D^+ production, π^+ detection, and trigger asymmetries cancel in the difference:

$$\mathcal{A}^{\mathrm{raw}}(D^+ \to K^- \pi^+ \pi^+_{\mathrm{trig}}) - \mathcal{A}^{\mathrm{raw}}(D^+ \to \overline{K}^0 \pi^+_{\mathrm{trig}}) = \mathcal{A}^{\mathrm{D}}(K^- \pi^+) - \mathcal{A}^{\mathrm{raw}}(\overline{K}^0). \tag{8.11}$$

To determine $\mathcal{A}^{\mathrm{K}\pi}$ for the signal decay, the kinematics of the $K^-\pi^+$ in the $D^+ \to K^-\pi^+\pi^+$ calibration sample must additionally match the kinematics in the signal sample. Therefore binned weights are firstly applied to the $D^+ \to K^-\pi^+\pi^+$ sample so that the $(p(K^-), p_{\mathrm{T}}(\pi^+), \eta(K^-), \eta(\pi^+))$ distribution matches that in the signal sample. Then weights are applied to the $D^+ \to \overline{K}^0\pi^+$ sample so that the $(p_{\mathrm{T}}(D^+)$, $p_{\mathrm{T}}(\pi^+_{\mathrm{trig}})$, $\eta(D^+)$, $\phi(\pi^+{}_{\mathrm{trig}}))$ distribution matches that in the $D^+ \to K^-\pi^+\pi^+$ sample.

The $\mathcal{A}^{\mathrm{raw}}$ for both calibration samples are extracted by fitting a signal plus background model to the weighted $m(D^+)$ distributions. The bins (Table 8.9) were chosen to be as narrow as possible while avoiding large weights that would reduce the stability of the $\mathcal{A}^{\mathrm{raw}}$ determination. The binning for the first weighting is much coarser due to the lower statistics of the signal sample compared to the calibration samples. The weighting procedure was considered to be acceptably stable since the standard deviation of the weights was never more than ten times the mean weight. An example of the weighting procedure is given in Figure 8.7, which demonstrates that large differences, particularly in p and p_{T}, are corrected by the weighting.

If the calibration samples are not weighted accurately, the asymmetries will be biased. The finite size of the bins necessitates the validation of this weighting procedure. The

Table 8.9: Bin edges used in the weighting of calibration samples. The top three rows correspond to binning for the first weighting and the bottom four rows correspond to binning for the second weighting.

Variable	Bin low edges
$p(K)$[GeV/c]	[0, 7, 10, 12, 15, 18, 21, 30, 40, 50, 70]
$p_{\mathrm{T}}(\pi)$[GeV/c]	[0, 0.5, 1, 1.5, 2, 3]
$\eta(K,\pi)$	[1.8, 2.9, 3.4]
$\eta(D)$	[2.55, 2.68, 2.78, 2.87, 2.95, 3.03, 3.11, 3.18, 3.26, 3.33, 3.40, 3.47, 3.54, 3.61, 3.69, 3.77, 3.87, 3.98, 4.14]
$p_{\mathrm{T}}(D)$[GeV/c]	[1.9, 2.6, 2.9, 3.0, 3.2, 3.4, 3.5, 3.7, 3.9, 4.0, 4.2, 4.4, 4.6, 4.9, 5.2, 5.5, 5.9, 6.4, 7.1, 8.3]
$p_{\mathrm{T}}(\pi_{\mathrm{trig}})$[GeV/$c$]	[1.47, 1.67, 1.73, 1.79, 1.85, 1.92, 1.99, 2.07, 2.16, 2.25, 2.35, 2.47, 2.60, 2.74, 2.91, 3.13, 3.39, 3.74, 4.24, 5.15]
$\phi(\pi_{\mathrm{trig}})$	[-3.15, -3.00, -2.85,..., 3.15]

procedure is validated by inserting artificial detection and production asymmetries with kinematic variations into the calibration samples and verifying that the obtained signal $\mathcal{A}^{K\pi}$ agrees with expectations.

The exact form of the artificial $\mathcal{A}^{K\pi}$ is not critical, but it is important that it is qualitatively similar to and varies at approximately the same rate as the true $\mathcal{A}^{K\pi}$. The dependence on the kaon and pion momenta is determined in a fit to measurements of $\mathcal{A}^{K\pi}$ obtained as described above. A linear η dependence is chosen to introduce a small but non-negligible variation qualitatively in line with physical expectations. The artificial detection asymmetries are

$$\mathcal{A}^{\mathrm{D}}(K^-) = -(1+0.5(\eta-3))\left(0.4+\frac{6}{p/\,\mathrm{GeV}/c}\right), \tag{8.12}$$

$$\mathcal{A}^{\mathrm{D}}(\pi^+) = (1+0.5(\eta-3))\left(-0.6+\frac{5}{p/\,\mathrm{GeV}/c}\right), \tag{8.13}$$

and $\mathcal{A}^{\mathrm{P}}(D^+) = -0.96\%$ [217]. For each D^+ charge q, asymmetry weights are calculated for each candidate regardless of true charge in the $D^+ \to K^-\pi^+\pi^+$ sample,

$$w^{\mathcal{A}}(q) = \frac{1+q\left(\mathcal{A}^{\mathrm{P}}(D^+)+\mathcal{A}^{\mathrm{D}}(K^-)+\mathcal{A}^{\mathrm{D}}(\pi^+)+\mathcal{A}^{\mathrm{D}}(\pi^+)\right)}{2}, \tag{8.14}$$

and the $D^+ \to \overline{K}^0\pi^+$ sample,

$$w^{\mathcal{A}}(q) = \frac{1+q\left(\mathcal{A}^{\mathrm{P}}(D^+)+\mathcal{A}^{\mathrm{D}}(\pi^+)\right)}{2}. \tag{8.15}$$

The yields for $q = \pm 1$ are determined in a fit to the calibration dataset, with each candidate weighted by the product of the kinematic weights and $w^{\mathcal{A}}(q)$. These yields are

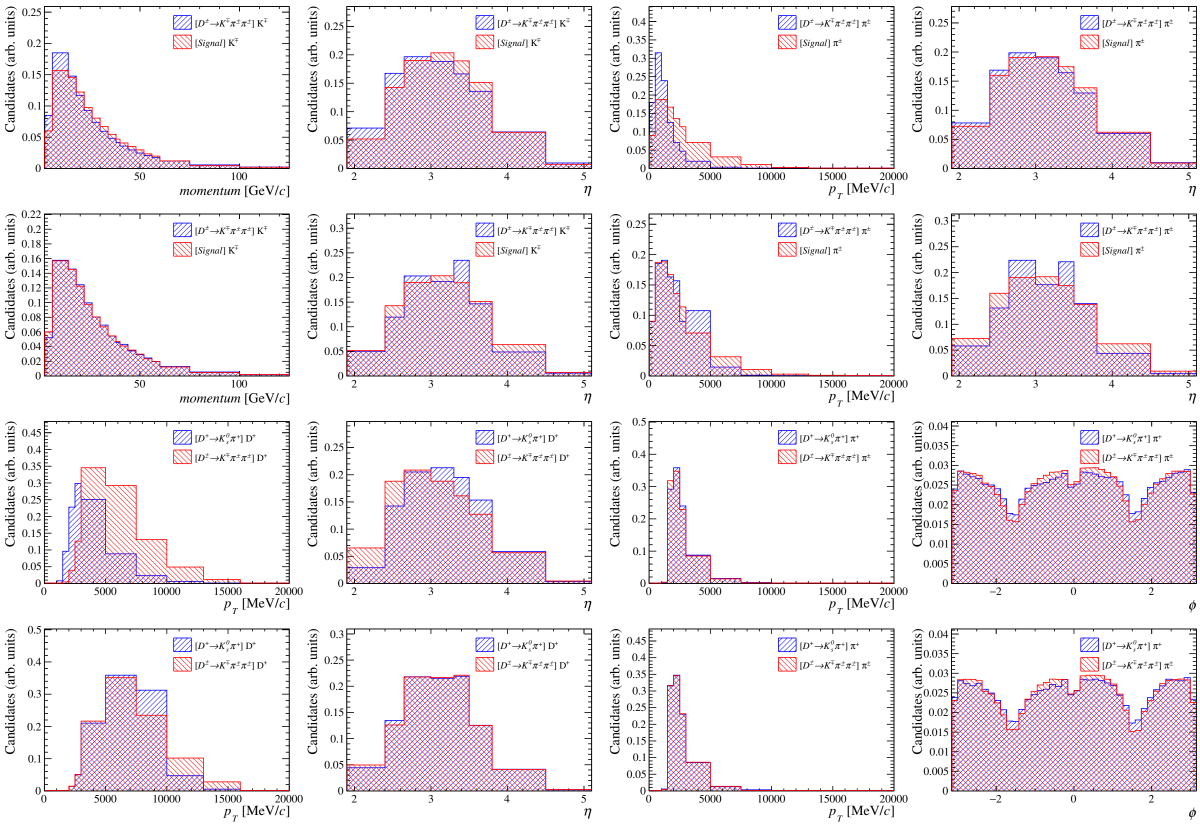

Figure 8.7: (First two rows) distributions of (from left) $p(K)$, $\eta(K)$, $p_{\mathrm{T}}(\pi)$, and $\eta(\pi)$ in (red) $B^- \to D_s^- D^0$ signal simulation and (blue) $D^+ \to K^-\pi^+\pi^+$ calibration data. (Second two rows) distributions of (from left) $p_{\mathrm{T}}(D^+)$, $\eta(D^+)$, $p_{\mathrm{T}}(\pi)$ and $\phi(\pi)$ in (red) $D^+ \to K^-\pi^+\pi^+$ and (blue) $D^+ \to K^0_{\mathrm{S}}\pi^+$ calibration data. The signal and calibration distributions are compared (first and third rows) before and (second and fourth rows) after weighting. For 2016 MagUp data.

used to obtain $\mathcal{A}^{K\pi}$. This procedure ensures that the statistical uncertainty in this validation is negligible.

The measured and expected asymmetries are compared for selected $K^-\pi^+$ pairs, which are chosen to ensure that the full range of kinematics is investigated. The differences, averaged over years and magnet polarities, are shown in Table 8.11. Overall there is no strong indication of a bias, and differences are smaller than the systematic uncertainties assigned to this asymmetry.

The residual $\mathcal{A}^{\mathrm{raw}}(\overline{K}^0)$ was estimated by simulating the propagation of the neutral kaons through the detector and their flavour oscillations when interacting with the material [214]. $\mathcal{A}^{\mathrm{raw}} = (-0.080 \pm 0.005)\%$ in 2011 and $(-0.082 \pm 0.005)\%$ in 2012 were calculated [218], with the uncertainty dominated by the material budget material. For Run 2 the asymmetry was calculated as a function of the momentum of the pion from the $\overline{K}^0$ [216]. This measurement was folded over the momentum distribution in the weighted calibration sample to obtain a correction of approximately $(-0.065 \pm 0.005)\%$ when the calibration samples were weighted to match a high-momentum two-body $D^0 \to K^-\pi^+$

Table 8.10: $\mathcal{A}^{\mathrm{K}\pi} - \mathcal{A}^{\mathrm{raw}}(\overline{K}^0)$ in percent for $K^-\pi^+$ pairs in signal decays. Uncertainties are from the finite size of the calibration samples. The asymmetry in each row is for the underlined kaon and averaging over all opposite-sign pions. In the final row of each channel, the total contribution to the detection asymmetries from $K\pi$ pairs is calculated, treating the uncertainties as fully correlated.

Final state	Run 1	Run 2
$D_s^-(\underline{K^-}\pi^-K^+)D^0(K^-\pi^+)$	-0.85 ± 0.26	-0.79 ± 0.08
$D_s^-(K^-\pi^-\underline{K^+})D^0(K^-\pi^+)$	-0.98 ± 0.13	-0.85 ± 0.06
$D_s^-(K^-\pi^-K^+)D^0(\underline{K^-}\pi^+)$	-0.80 ± 0.25	-0.71 ± 0.10
$D_s^-(K^-\pi^-K^+)D^0(K^-\pi^+)$	-0.67 ± 0.39	-0.65 ± 0.13
$D_s^-(\underline{K^-}\pi^-K^+)D^0(K^-\pi^+\pi^-\pi^+)$	-0.95 ± 0.15	-0.84 ± 0.04
$D_s^-(K^-\pi^-\underline{K^+})D^0(K^-\pi^+\pi^-\pi^+)$	-0.96 ± 0.16	-0.82 ± 0.06
$D_s^-(K^-\pi^-K^+)D^0(\underline{K^-}\pi^+\pi^-\pi^+)$	-0.96 ± 0.12	-0.82 ± 0.05
$D_s^-(K^-\pi^-K^+)D^0(K^-\pi^+\pi^-\pi^+)$	-0.94 ± 0.11	-0.84 ± 0.04
$D^-(\pi^-\pi^-\underline{K^+})D^0(K^-\pi^+)$	-0.98 ± 0.18	-0.83 ± 0.06
$D^-(\pi^-\pi^-K^+)D^0(\underline{K^-}\pi^+)$	-0.76 ± 0.23	-0.70 ± 0.11
$D^-(\pi^-\pi^-K^+)D^0(K^-\pi^+)$	0.23 ± 0.08	0.13 ± 0.07
$D^-(\pi^-\pi^-\underline{K^+})D^0(K^-\pi^+\pi^-\pi^+)$	-0.97 ± 0.14	-0.78 ± 0.06
$D^-(\pi^-\pi^-K^+)D^0(\underline{K^-}\pi^+\pi^-\pi^+)$	-0.95 ± 0.12	-0.81 ± 0.04
$D^-(\pi^-\pi^-K^+)D^0(K^-\pi^+\pi^-\pi^+)$	0.02 ± 0.03	-0.03 ± 0.02
$D^{*-}(\underline{K^+}\pi^-\pi_s^+)D^0(K^-\pi^+)$	-0.80 ± 0.19	-0.64 ± 0.10
$D^{*-}(K^+\pi^-\pi_s^+)D^0(\underline{K^-}\pi^+)$	-0.74 ± 0.27	-0.70 ± 0.10
$D^{*-}(K^+\pi^-\pi_s^-)D^0(K^-\pi^+)$	0.06 ± 0.15	-0.06 ± 0.01
$D^{*-}(\underline{K^+}\pi^-\pi_s^+)D^0(K^-\pi^+\pi^-\pi^+)$	-0.84 ± 0.17	-0.68 ± 0.09
$D^{*-}(K^+\pi^-\pi_s^+)D^0(\underline{K^-}\pi^+\pi^-\pi^+)$	-0.98 ± 0.14	-0.82 ± 0.05
$D^{*-}(K^+\pi^-\pi_s^-)D^0(K^-\pi^+\pi^-\pi^+)$	-0.13 ± 0.11	-0.13 ± 0.04
$D^{*-}(\underline{K^+}\pi^-\pi^+\pi^-\pi_s^+)D^0(K^-\pi^+)$	-1.00 ± 0.10	-0.83 ± 0.04
$D^{*-}(K^+\pi^-\pi^+\pi^-\pi_s^+)D^0(\underline{K^-}\pi^+)$	-0.84 ± 0.18	-0.70 ± 0.09
$D^{*-}(K^+\pi^-\pi^+\pi^-\pi_s^-)D^0(K^-\pi^+)$	0.17 ± 0.11	0.13 ± 0.06

decay or $(-0.071 \pm 0.005)\%$ when weighted to match lower momentum particles.

8.2.5 Pion detection asymmetry

The pion detection asymmetry is defined as

$$\mathcal{A}^{\pi} = \frac{\varepsilon(\pi^+) - \varepsilon(\pi^-)}{\varepsilon(\pi^+) + \varepsilon(\pi^-)}, \tag{8.16}$$

where $\varepsilon(\pi^\pm)$ is the detection efficiency for a $\pi^\pm$. $\varepsilon(\pi^+)$ is measured as the fraction of $D^{*+} \to (D^0 \to K^-\pi^+\pi^-\pi^+)\pi_s^+$ decays that can be partially reconstructed as $K^-\pi^+\pi^-\pi_s^+$ requiring only one π^+, where the full final state can also be reconstructed [219].

Table 8.11: Absolute difference between the measured and expected $\mathcal{A}^{K\pi}$ in percent for selected $K^-\pi^+$ pairs when the asymmetries in Equations. 8.12-8.13 are applied.

$K\pi$	Run 1	Run 2
$\underline{K^-}\pi^-K^+, K^-\underline{\pi^+}$	0.04	0.07
$K^-\underline{\pi^-K^+}, K^-\pi^+$	–0.04	–0.01
$K^-\pi^-K^+, \underline{K^-\pi^+}$	0.02	0.04
$K^-\pi^-K^+, \underline{K^-\pi^+}\pi^-\pi^+$	–0.06	–0.01
$\underline{K^+\pi^-}\pi_s^-, K^-\pi^+$	0.03	0.03
$\underline{K^+\pi^-}\pi^+\pi^-\pi_s^-, K^-\pi^+$	–0.08	0.00

$\mathcal{A}^{\pi}$ for each final state pion (Table 8.12) is evaluated by weighting the measurement with the momentum distribution in selected simulated candidates. For 2016-2018 data, only momentum-integrated measurements are available. To obtain $\mathcal{A}^{\pi}$, these integrated measurements are shifted by the change in asymmetry induced by folding Run 2 simulation over the 2012 measurement. The correction in partially reconstructed signal differs by less than 0.01%.

8.2.6 PID asymmetry

Owing to the dependence of the PID response on the reconstruction of the charged particle tracks, there may be a charge asymmetry in the PID response,

$$\mathcal{A}^{\text{PID}} = \frac{\varepsilon_{\text{PID}}(B^- \to f) - \varepsilon_{\text{PID}}(B^+ \to \overline{f})}{\varepsilon_{\text{PID}}(B^- \to f) + \varepsilon_{\text{PID}}(B^+ \to \overline{f})}. \tag{8.17}$$

PID requirements are placed on all final state particles in the preselection, and cross-feed vetoes subject some particles to even tighter PID requirements (Section 5.7). The inclusion of PID variables in the BDT (Section 5.8) further tightens the requirements, but it is no longer possible to define exact PID cuts. Therefore the effective PID cut is defined as the cut on $\text{DLL}_{K\pi}$ that produces the same mean $\text{DLL}_{K\pi}$ as the optimum BDT cut (Figure 8.8). This effective cut accounts for PID requirements from all of the above sources. Effective cuts are given in Table 8.13.

Furthermore a PID asymmetry is absorbed into $\mathcal{A}^{K\pi}$ and $\mathcal{A}^{\pi}$ since the calibration samples used to derive these asymmetries are subjected to PID requirements that are often tighter than the requirements in this analysis. The PID requirements applied to the calibration samples for $\mathcal{A}^{K\pi}$ are shown in Table 8.14. The relevant pions in the $\mathcal{A}^{\pi}$ calibration sample have $\text{DLL}_{K\pi} < 6(10)$ in Run 1(2).

The software package `PIDCalib` [220] uses calibration data samples [182] to determine the charge-dependent efficiency of the effective and detection asymmmetry PID requirements for each track in bins of track p, track η, and $nTracks$. Since PID efficiencies

Table 8.12: $\mathcal{A}^{\pi}$ for all pions in signal decays in percent. Uncertainties are from the finite size of the calibration samples.

Final state	Run 1	Run 2
$D_s^-(K^-\underline{\pi^-}K^+)D^0(K^-\pi^+)$	-0.19 ± 0.07	-0.04 ± 0.04
$D_s^-(K^-\pi^-K^+)D^0(K^-\underline{\pi^+})$	-0.15 ± 0.09	-0.02 ± 0.04
$D_s^-(K^-\underline{\pi^-}K^+)D^0(K^-\pi^+\pi^-\pi^+)$	-0.19 ± 0.07	-0.04 ± 0.04
$D_s^-(K^-\pi^-K^+)D^0(K^-\underline{\pi^+}\pi^-\pi^+)$	-0.19 ± 0.06	-0.05 ± 0.04
$D_s^-(K^-\pi^-K^+)D^0(K^-\pi^+\underline{\pi^-}\pi^+)$	-0.19 ± 0.06	-0.04 ± 0.04
$D_s^-(K^-\pi^-K^+)D^0(K^-\pi^+\pi^-\underline{\pi^+})$	-0.20 ± 0.06	-0.05 ± 0.04
$D^-(\underline{\pi^-}\pi^-K^+)D^0(K^-\pi^+)$	-0.17 ± 0.06	-0.03 ± 0.04
$D^-(\pi^-\underline{\pi^-}K^+)D^0(K^-\pi^+)$	-0.17 ± 0.06	-0.03 ± 0.04
$D^-(\pi^-\pi^-K^+)D^0(K^-\underline{\pi^+})$	-0.15 ± 0.09	-0.02 ± 0.04
$D^-(\underline{\pi^-}\pi^-K^+)D^0(K^-\pi^+\pi^-\pi^+)$	-0.17 ± 0.06	-0.03 ± 0.04
$D^-(\pi^-\underline{\pi^-}K^+)D^0(K^-\pi^+\pi^-\pi^+)$	-0.17 ± 0.06	-0.03 ± 0.04
$D^-(\pi^-\pi^-K^+)D^0(K^-\underline{\pi^+}\pi^-\pi^+)$	-0.20 ± 0.06	-0.05 ± 0.04
$D^-(\pi^-\pi^-K^+)D^0(K^-\pi^+\underline{\pi^-}\pi^+)$	-0.19 ± 0.06	-0.04 ± 0.04
$D^-(\pi^-\pi^-K^+)D^0(K^-\pi^+\pi^-\underline{\pi^+})$	-0.19 ± 0.06	-0.05 ± 0.04
$D^{*-}(K^+\underline{\pi^-}\pi_s^-)D^0(K^-\pi^+)$	-0.14 ± 0.09	-0.01 ± 0.04
$D^{*-}(K^+\pi^-\underline{\pi_s^-})D^0(K^-\pi^+)$	-0.35 ± 0.09	-0.17 ± 0.04
$D^{*-}(K^+\pi^-\pi_s^-)D^0(K^-\underline{\pi^+})$	-0.14 ± 0.10	-0.02 ± 0.04
$D^{*-}(K^+\underline{\pi^-}\pi_s^-)D^0(K^-\pi^+\pi^-\pi^+)$	-0.14 ± 0.10	-0.01 ± 0.04
$D^{*-}(K^+\pi^-\underline{\pi_s^-})D^0(K^-\pi^+\pi^-\pi^+)$	-0.34 ± 0.09	-0.16 ± 0.04
$D^{*-}(K^+\pi^-\pi_s^-)D^0(K^-\underline{\pi^+}\pi^-\pi^+)$	-0.19 ± 0.06	-0.04 ± 0.04
$D^{*-}(K^+\pi^-\pi_s^-)D^0(K^-\pi^+\underline{\pi^-}\pi^+)$	-0.19 ± 0.06	-0.04 ± 0.04
$D^{*-}(K^+\pi^-\pi_s^-)D^0(K^-\pi^+\pi^-\underline{\pi^+})$	-0.19 ± 0.06	-0.04 ± 0.04
$D^{*-}(K^+\underline{\pi^-}\pi^+\pi^-\pi_s^-)D^0(K^-\pi^+)$	-0.20 ± 0.06	-0.05 ± 0.04
$D^{*-}(K^+\pi^-\underline{\pi^+}\pi^-\pi_s^-)D^0(K^-\pi^+)$	-0.19 ± 0.06	-0.04 ± 0.04
$D^{*-}(K^+\pi^-\pi^+\underline{\pi^-}\pi_s^-)D^0(K^-\pi^+)$	-0.20 ± 0.06	-0.05 ± 0.04
$D^{*-}(K^+\pi^-\pi^+\pi^-\underline{\pi_s^-})D^0(K^-\pi^+)$	-0.35 ± 0.09	-0.16 ± 0.04
$D^{*-}(K^+\pi^-\pi^+\pi^-\pi_s^-)D^0(K^-\underline{\pi^+})$	-0.15 ± 0.11	-0.02 ± 0.04

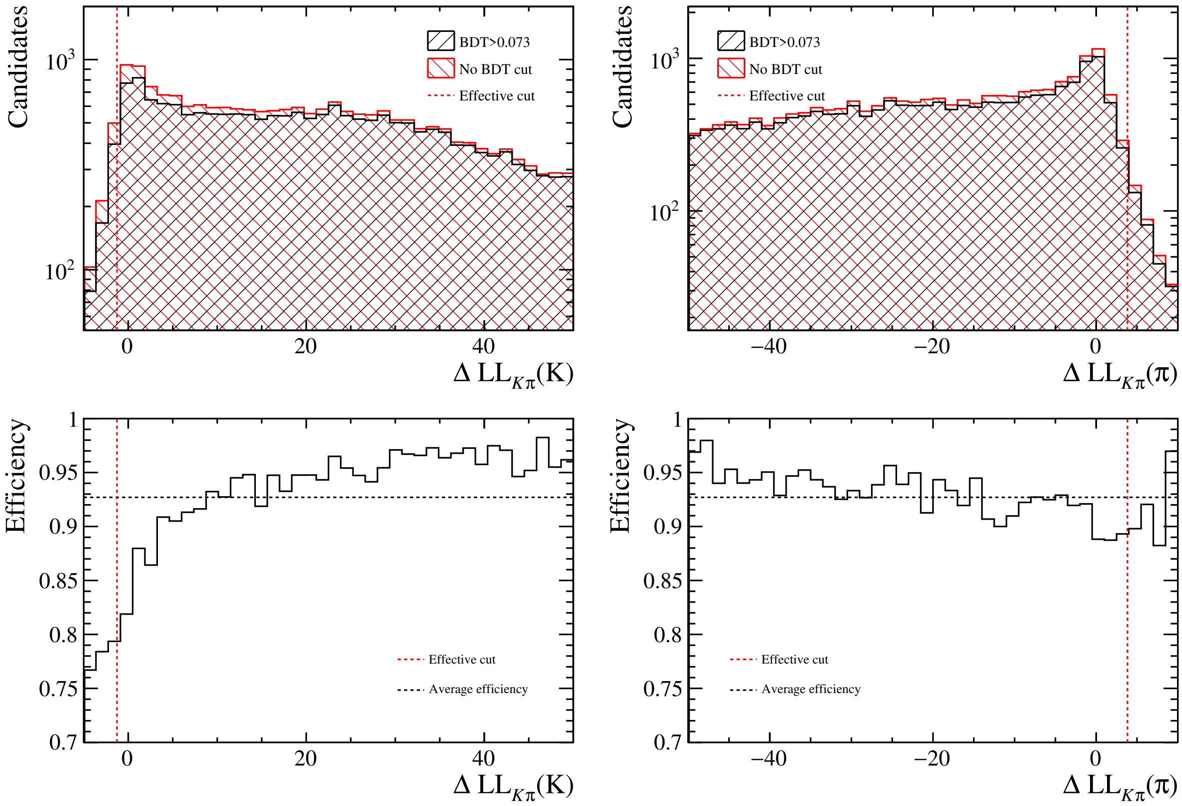

Figure 8.8: (Top) Distributions of $\mathrm{DLL}_{K\pi}$ (red) before and (black) after the BDT cut and (bottom) efficiency of the BDT cut as a function of $\mathrm{DLL}_{K\pi}$ in corrected simulation. For (left) the kaon and (right) the pion from the D^0 in Run 2-simulated $B^- \to D^- D^0$ decays. The location of the effective cut and the average efficiency are also indicated.

vary most strongly as a function of p, the calibration data is most finely divided in this variable. In particular, the data is very finely divided below 10 GeV/c where the efficiency varies most rapidly, and there are bin edges at the pion and kaon Cherenkov ring thresholds of 2.6 and 9.3 GeV/c. The division is much coarser in η and $nTracks$ to avoid a large number of bins with low calibration yield. The bin low edges are:

- p: [1.0, 2.6, 4.5, 7.0, 9.3, 15.6, 19.0, 24.4, 29.8, 35.2, 40.6, 51.4, 56.8, 62.2, 67.6, 73.0, 78.4, 83.8, 89.2, 94.6, 100.0] GeV/c
- η: [1.500, 2.375, 3.250, 4.125]
- $nTracks$: [0, 150]

The PID asymmetry is evaluated on simulated candidates within ±40 MeV/c^2 of m_{B^-} that pass all selection except for the BDT requirement, since the BDT introduces PID requirements. Per-candidate efficiencies are determined for each track under both the positive and negative charge hypotheses. The product of the track efficiencies are the candidate efficiencies, which are used to determine the per-candidate PID asymmetries.

Table 8.13: Effective $\mathrm{DLL}_{K\pi}$ requirements on each final state particle.

Period	Final state	$\mathrm{DLL}_{K\pi}(K) >$	$\mathrm{DLL}_{K\pi}(\pi) <$
Run 1	$KK\pi, K\pi$	$(-2.0, -2.4, -2.8)$	$(8.7, 8.6)$
	$KK\pi, K3\pi$	$(-2.0, -1.8, -1.8)$	$(8.8, 9.5, 9.5, 9.2)$
	$K\pi\pi, K\pi$	$(-1.1, -0.6)$	$(7.6, 8, 6.7)$
	$K\pi\pi, K3\pi$	$(-0.5, 0.6)$	$(7.6, 8, 7.0, 7.2, 8)$
	$K\pi\pi_s, K\pi$	$(-3.4, -1.4)$	$(8.0, 9.9, 8)$
	$K\pi\pi_s, K3\pi$	$(-2.4, -0.1)$	$(6.6, 9.9, 7.7, 7.4, 7.9)$
	$K3\pi\pi_s, K\pi$	$(-2.1, -1.8)$	$(9.3, 8.9, 9.1, 10.0, 8.6)$
Run 2	$KK\pi, K\pi$	$(-2.9, -3.0, -3.1)$	$(6.2, 6.7)$
	$KK\pi, K3\pi$	$(-2.8, -2.7, -2.0)$	$(6.0, 5.9, 5.9, 6.1)$
	$K\pi\pi, K\pi$	$(-1.1, -1.2)$	$(2.7, 3.3, 3.8)$
	$K\pi\pi, K3\pi$	$(-0.9, 0.6)$	$(3.0, 2.8, 3.1, 3.6, 3.1)$
	$K\pi\pi_s, K\pi$	$(-3.2, -1.9)$	$(7.2, 9.5, 5.7)$
	$K\pi\pi_s, K3\pi$	$(-2.0, -0.4)$	$(5.8, 10.0, 4.1, 4.0, 3.9)$
	$K3\pi\pi_s, K\pi$	$(-1.8, -1.7)$	$(7.0, 6.0, 7.2, 9.6, 5.7)$

Table 8.14: Requirements on $\mathrm{DLL}_{K\pi}$ in the calibration samples used to calculate $A_D(K\pi)$. Requirements are tighter for earlier years.

Years	PID requirements
2011-12	$\mathrm{DLL}_{K\pi}(K) > 7$, $\mathrm{DLL}_{K\pi}(\pi) < 0$
2015	$\mathrm{DLL}_{K\pi}(K) > 5$, $\mathrm{DLL}_{K\pi}(\pi) < 5$
2016-18	$\mathrm{DLL}_{K\pi}(K) > -10$, $\mathrm{DLL}_{K\pi}(\pi) < 7$

The total $\mathcal{A}^{\mathrm{PID}}$ is the mean of the per-candidate asymmetries weighted by the per-candidate charge-averaged efficiencies. The total correction to $\mathcal{A}^{CP}$ is the difference $\Delta\mathcal{A}^{\mathrm{PID}}$ between the $\mathcal{A}^{\mathrm{PID}}$ from the effective PID requirement and the total $\mathcal{A}^{\mathrm{PID}}$ absorbed into $\mathcal{A}^{\mathrm{K}\pi}$ and $\mathcal{A}^{\pi}$ (Tables 8.15-8.16). The correction for partially reconstructed decays differs by less than 0.01%.

8.3 Efficiency corrections

Efficiencies of acceptance, reconstruction, and selection are essential to measure branching fraction ratios, and are determined using samples of simulated decays as

$$\varepsilon = \frac{\sum_i w_i}{N_{gen}} \tag{8.18}$$

where the sum is over all selected candidates in simulation within $\pm 40\,\mathrm{MeV}$ of m_{B^-}. N_{gen} is the number of generated events. The weights w_i correct for both the mismodelling of

Table 8.15: $\mathcal{A}^{\rm PID}$ of the effective PID requirement and absorbed into detection asymmetries, and the associated $\Delta\mathcal{A}^{\rm PID}$, in Run 1 data.

Final State	Polarity	$\mathcal{A}^{\rm PID}$ / % Effective Cut	$\mathcal{A}^{\rm D}$	$\Delta\mathcal{A}^{\rm PID}$/%
$B^- \to D_s^- D^0$	Up	0.00 ± 0.06	0.17 ± 0.10	-0.17 ± 0.06
$KK\pi$	Down	-0.03 ± 0.06	-0.33 ± 0.10	0.30 ± 0.05
$K\pi$	Combined	-0.02 ± 0.05	-0.09 ± 0.08	0.07 ± 0.04
$B^- \to D_s^- D^0$	Up	0.00 ± 0.12	0.12 ± 0.15	-0.11 ± 0.06
$KK\pi$	Down	-0.03 ± 0.12	-0.30 ± 0.15	0.27 ± 0.06
$K3\pi$	Combined	-0.02 ± 0.08	-0.10 ± 0.11	0.08 ± 0.04
$B^- \to D^- D^0$	Up	0.06 ± 0.14	0.12 ± 0.15	-0.07 ± 0.03
$K\pi\pi$	Down	-0.04 ± 0.14	-0.13 ± 0.14	0.09 ± 0.03
$K\pi$	Combined	0.00 ± 0.10	-0.01 ± 0.11	0.01 ± 0.02
$B^- \to D^- D^0$	Up	0.02 ± 0.09	0.06 ± 0.12	-0.04 ± 0.04
$K\pi\pi$	Down	-0.05 ± 0.08	-0.05 ± 0.11	0.00 ± 0.04
$K3\pi$	Combined	-0.01 ± 0.06	0.00 ± 0.08	-0.02 ± 0.03
$B^- \to D^{*-} D^0$	Up	-0.02 ± 0.41	-0.16 ± 0.48	0.15 ± 0.13
$K\pi\pi_s$	Down	-0.05 ± 0.40	-0.02 ± 0.44	-0.03 ± 0.13
$K\pi$	Combined	-0.03 ± 0.29	-0.09 ± 0.34	0.06 ± 0.09
$B^- \to D^{*-} D^0$	Up	-0.04 ± 0.24	-0.18 ± 0.32	0.14 ± 0.11
$K\pi\pi_s$	Down	-0.06 ± 0.21	0.08 ± 0.28	-0.14 ± 0.10
$K3\pi$	Combined	-0.05 ± 0.16	-0.04 ± 0.21	-0.01 ± 0.07
$B^- \to D^{*-} D^0$	Up	-0.00 ± 0.24	0.08 ± 0.25	-0.09 ± 0.14
$K3\pi\pi_s$	Down	-0.02 ± 0.24	-0.07 ± 0.23	0.05 ± 0.15
$K\pi$	Combined	-0.01 ± 0.18	0.00 ± 0.17	-0.02 ± 0.10

kinematic and angular distributions (Section 5.3) and of the efficiency of L0 requirements, track reconstruction, and PID requirements. This section will describe the evaluation of the weights to correct each these efficiencies.

8.3.1 Hardware trigger efficiency

The efficiencies of L0 trigger requirements are not modelled accurately in simulation, so per-candidate efficiency correction weights are determined using measurements derived from calibration samples. The efficiency of the L0 Hadron TOS requirement ($\varepsilon_{\rm TOS}$) is determined in bins of $(p_{\rm T}(B), \eta(B))$ using the method described in Section 8.2.3, and binned data-simulation corrections for $\varepsilon_{\rm TOS}$ are derived. Data-simulation corrections for the requirement that a candidate passes L0 Global TIS but not L0 Hadron TOS ($\varepsilon_{\rm TISnotTOS}$) are obtained using $\varepsilon_{\rm TISnotTOS} = 1 - \varepsilon_{\rm TOS}$ for candidates that have passed the L0 Global TIS decision. Per-candidate weights for all selected simulated candidates are determined from

Table 8.16: $\mathcal{A}^{\mathrm{PID}}$ of the effective PID requirement and absorbed into detection asymmetries, and the associated $\Delta\mathcal{A}^{\mathrm{PID}}$, in Run 2 data.

Final State	Polarity	$\mathcal{A}^{\mathrm{PID}}$ / % Effective Cut	$\mathcal{A}^{\mathrm{D}}$	$\Delta\mathcal{A}^{\mathrm{PID}}$/%
$B^- \to D_s^- D^0$	Up	0.03 ± 0.03	0.03 ± 0.03	0.01 ± 0.02
$KK\pi$	Down	0.12 ± 0.03	0.13 ± 0.03	-0.01 ± 0.02
$K\pi$	Combined	0.08 ± 0.02	0.08 ± 0.02	-0.00 ± 0.01
$B^- \to D_s^- D^0$	Up	-0.00 ± 0.03	0.01 ± 0.05	-0.02 ± 0.05
$KK\pi$	Down	0.09 ± 0.04	0.07 ± 0.04	0.01 ± 0.02
$K3\pi$	Combined	0.04 ± 0.02	0.04 ± 0.03	-0.00 ± 0.03
$B^- \to D^- D^0$	Up	0.06 ± 0.08	0.05 ± 0.06	0.01 ± 0.04
$K\pi\pi$	Down	-0.08 ± 0.08	-0.00 ± 0.06	-0.08 ± 0.05
$K\pi$	Combined	-0.01 ± 0.06	0.03 ± 0.04	-0.03 ± 0.03
$B^- \to D^- D^0$	Up	0.09 ± 0.04	0.06 ± 0.03	0.03 ± 0.03
$K\pi\pi$	Down	-0.07 ± 0.06	-0.02 ± 0.04	-0.04 ± 0.05
$K3\pi$	Combined	0.01 ± 0.04	0.02 ± 0.03	-0.01 ± 0.03
$B^- \to D^{*-} D^0$	Up	0.14 ± 0.12	0.15 ± 0.12	-0.01 ± 0.01
$K\pi\pi_s$	Down	0.12 ± 0.13	0.05 ± 0.11	0.07 ± 0.04
$K\pi$	Combined	0.13 ± 0.10	0.10 ± 0.09	0.03 ± 0.02
$B^- \to D^{*-} D^0$	Up	0.08 ± 0.07	0.08 ± 0.07	-0.00 ± 0.02
$K\pi\pi_s$	Down	0.03 ± 0.07	0.06 ± 0.06	-0.03 ± 0.02
$K3\pi$	Combined	0.05 ± 0.05	0.07 ± 0.05	-0.02 ± 0.01
$B^- \to D^{*-} D^0$	Up	0.10 ± 0.16	0.13 ± 0.15	-0.03 ± 0.04
$K3\pi\pi_s$	Down	0.14 ± 0.14	0.07 ± 0.13	0.07 ± 0.04
$K\pi$	Combined	0.12 ± 0.11	0.10 ± 0.10	0.02 ± 0.03

these binned corrections. Figure 8.9 shows a typical example of the binned efficiencies and weights, and Table 8.17 shows the average weight for each final state.

8.3.2 Tracking efficiency

The tracking efficiency, introduced in Section 3.2.1, has been measured in bins of track (p, η) with $J/\psi \to \mu^+\mu^-$ decays using a *tag-and-probe* technique [128]. One of the muons, the *tag* leg, is fully reconstructed. To reconstruct $m(\mu^+\mu^-)$ with sufficient resolution to separate the decays from random combinations of tracks, the other muon, the *probe* leg, must be reconstructed in two of the VELO, TT, T stations, and MUON stations. The tracking efficiency in the other tracking detectors is then the fraction of probe legs that match a long track. For example, the long track efficiency is determined using tracks present in both the TT and MUON as probes. Data-simulation correction weights are the ratio of the efficiencies after the same method is applied to simulation.

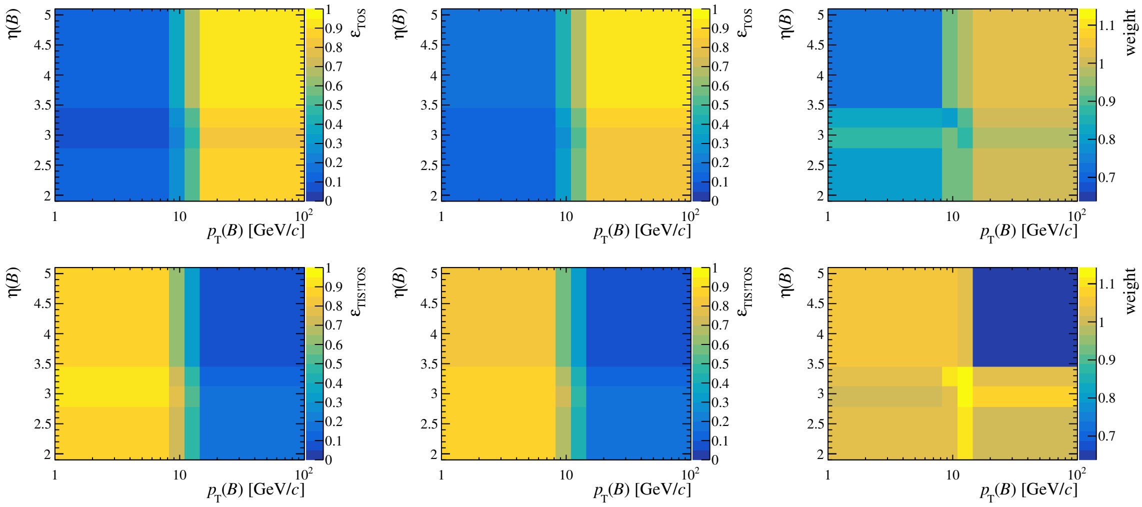

Figure 8.9: (Top) $\varepsilon_{\mathrm{L0TOS}}$ and (bottom) $\varepsilon_{\mathrm{L0TISnotTOS}}$ in (left) data and (centre) simulation, and (right) corresponding data-simulation weights as a function of $(p_{\mathrm{T}}(B), \eta(B))$. For the $D_s^- D^0$, $D^0 \to K^- \pi^+$ final state in the 2016 MagUp dataset.

Table 8.17: Corrections to $\varepsilon_{\mathrm{L0}}$ including uncertainties from the finite size of calibration and simulation samples.

Final state	Run 1	Run 2
$KK\pi, K\pi$	0.976 ± 0.003	0.974 ± 0.003
$KK\pi, K3\pi$	0.985 ± 0.005	0.983 ± 0.004
$K\pi\pi, K\pi$	0.978 ± 0.004	0.971 ± 0.003
$K\pi\pi, K3\pi$	0.990 ± 0.007	0.988 ± 0.005
$K\pi\pi_s, K\pi$	0.997 ± 0.009	0.970 ± 0.004
$K\pi\pi_s, K3\pi$	1.028 ± 0.014	0.990 ± 0.008
$K3\pi\pi_s, K\pi$	1.003 ± 0.016	0.983 ± 0.007

Per-candidate weights are derived by applying corrections to each final state particle. The calibration sample only contains tracks with $p > 5\,\mathrm{GeV}$ so a correction of 1.00 ± 0.05 is assigned for tracks with $p < 5\,\mathrm{GeV}$. The resulting efficiency corrections are shown in Table 8.18.

8.3.3 PID efficiency

The efficiency of PID requirements on a candidate ($\varepsilon_{\mathrm{PID}}$) can be determined accurately and precisely as described in Section 8.2.6. The PID requirements include:

- The PID requirements applied in stripping using uncorrected $\mathrm{DLL}_{K\pi}$.
- The PID requirements applied in preselection using corrected $\mathrm{DLL}_{K\pi}$.
- The efficiency of the PID-dependent components of the cross-feed vetoes.

Table 8.18: Corrections to the tracking efficiency in each final state and operational run, including uncertainties from the finite size of calibration and simulation samples.

Final state	Run 1	Run 2
$KK\pi, K\pi$	1.03 ± 0.01	0.98 ± 0.01
$KK\pi, K3\pi$	1.05 ± 0.02	0.97 ± 0.02
$K\pi\pi, K\pi$	1.03 ± 0.01	0.98 ± 0.01
$K\pi\pi, K3\pi$	1.05 ± 0.02	0.97 ± 0.02
$K\pi\pi_s, K\pi$	1.03 ± 0.02	0.98 ± 0.02
$K\pi\pi_s, K3\pi$	1.06 ± 0.03	0.97 ± 0.03
$K3\pi\pi_s, K\pi$	1.06 ± 0.04	0.97 ± 0.03

The majority of the simulated samples for this analysis were generated in 'Filtered' mode, so the PID requirements in stripping have been irreversibly applied. Stripping 'Flagged' samples are available for $B^- \to D^-_{(s)}D^0$, $D^0 \to K\pi(\pi\pi)$ in 2012 conditions, $B^- \to D^{*-}D^0$, $\overline{D}^0 \to K^+\pi^-$, $D^0 \to K^-\pi^+$ in 2015 conditions and $B^- \to D^-_{(s)}D^0$, $D^0 \to K\pi$ in 2017 conditions.

To correct the PID efficiency in filtered samples, data-simulation corrections for the PID efficiency of each final state particle in $(p, \eta, nTracks)$ bins are derived using stripping-flagged samples. Per-candidate weights are evaluated using these binned corrections. Corrections for all pions in a $D^0 \to K^-\pi^+\pi^-\pi^+$ decay use corrections from the pion in the $D^0 \to K^-\pi^+$ decay. Corrections are obtained for $B^- \to D^{*-}D^0$ decays in Run 1 using $B^- \to D^-D^0$ simulation, using the first pion in the $D^- \to \pi^-\pi^-K^+$ decay to correct the pions from the $\overline{D}^0$ and the second pion to correct the soft pion from the D^{*-}. Average weights are shown in Table 8.19.

The exception is for $B^- \to D^-_{(s)}D^0$ decays in Run 1, where the stripping-flagged samples are used elsewhere in the analysis. In this case, no PID requirements are applied to the simulation sample and the weights are efficiencies instead of data-simulation corrections. Average efficiencies are given in Table 8.20.

To validate the corrections obtained using different final states, ε_{PID} for $D^-_{(s)}D^0$, $D^0 \to K^-\pi^+\pi^-\pi^+$ in Run 1 when using the $D^0 \to K^-\pi^+$ final state is compared to the true ε_{PID} obtained using PIDCalib and stripping-flagged simulation samples. A similar test is performed using the D^-D^0 channel to correct the $D^{*-}D^0$ channel in Run 2. In both cases, the efficiencies are consistent within 2σ (Table 8.21).

8.4 Systematic uncertainties

This section describes the evaluation of systematic uncertainties on $\mathcal{A}^{CP}$ and R. The numerical values following a weighted average over D^0 decay modes are presented in Tables 8.22-8.23.

Table 8.19: Data-simulation correction weights, including uncertainties from the finite size of the calibration and $B \to DD$ calibration simulation samples.

Channel	Final state	Run 1	Run 2
$D_s^- D^0$	$D^0 \to K\pi$	-	0.996 ± 0.003
	$D^0 \to K3\pi$	-	0.999 ± 0.005
$D^- D^0$	$D^0 \to K\pi$	-	0.998 ± 0.005
	$D^0 \to K3\pi$	-	1.003 ± 0.008
$D^{*-} D^0$	$D^{*-} \to K\pi\pi_s$, $D^0 \to K\pi$	0.978 ± 0.005	0.998 ± 0.003
	$D^{*-} \to K\pi\pi_s$, $D^0 \to K3\pi$	0.981 ± 0.012	1.002 ± 0.005
	$D^{*-} \to K3\pi\pi_s$, $D^0 \to K\pi$	0.973 ± 0.006	0.993 ± 0.004

Table 8.20: $\varepsilon_{\mathrm{PID}}$ for Run 1 data, including uncertainties from the finite size of the calibration samples.

Channel	Final state	$\varepsilon_{\mathrm{PID}}/\%$
$D_s^- D^0$	$D^0 \to K\pi$	88.19 ± 0.12
	$D^0 \to K3\pi$	83.21 ± 0.41
$D^- D^0$	$D^0 \to K\pi$	78.88 ± 0.14
	$D^0 \to K3\pi$	73.30 ± 0.34

Systematic uncertainties on the raw asymmetries and yields

Systematic uncertainties on the raw asymmetries and yields due to the choice of invariant-mass model are assessed conservatively as the change in central value when an alternative model is used. These alternative models are specified below. Additionally, the systematic uncertainties associated with fixed model parameters are evaluated as the change in central value when these parameters are varied by their uncertainties.

Fully and partially (one missing particle) $B \to DD$ model: The resolution function is replaced by a double Gaussian.

$B \to D^* D^*$ (two missing particles) model: The resolution smearing applied to the generator-level simulation (Section 5.2) is removed.

Partially reconstructed $B \to D^{**} D$ model: The resolution smearing is removed, and

Table 8.21: Bias on $\varepsilon_{\mathrm{PID}}$ as a result of using the weighted correction method. The quoted uncertainty includes the finite size of the calibration and simulation samples, and the finite size of the calibration and signal bins. Correlations in uncertainties between the methods are included.

Final state	Period	Bias
$KK\pi, K3\pi$	Run 1	0.993 ± 0.011
$K\pi\pi, K3\pi$	Run 1	0.983 ± 0.011
$K\pi\pi_s, K\pi$	Run 2	1.004 ± 0.003

the fraction of the model from each of the four components ($D^{**} = D_0, D_1, D_1^{'}, D_2$) is varied between 0 and 1.

Combinatorial background model: A linear model is used.

Cross-feed yield, asymmetry, and model: The yield and asymmetry are varied by their uncertainties. A single Gaussian model is used.

Single charm model: The yield and Gaussian width are varied by their statistical uncertainties.

External inputs: Branching fractions, polarisation fractions, asymmetries, and B_s^0 to B^0 efficiency ratios that are fixed from external inputs are varied by adding and subtracting their uncertainties. The average change in central value is assigned as a systematic uncertainty.

$\mathcal{A}^P(B^0)$ dilution: The dilution is varied between 0 and 1.

Systematic uncertainties on the asymmetry and efficiency corrections

All corrections have the following sources of systematic uncertainty in common:

Calibration sample size and systematic uncertainties: The statistical and systematic uncertainties on the binned measurements are propagated.

Simulation sample size: The systematic uncertainty is the uncertainty on the (weighted) mean of the per-simulated-candidate asymmetries or weights.

Finite bin size: The bias due to the finite size of the bins is estimated by doubling the bin size in all dimensions.

Further uncertainties are specific to an asymmetry or efficiency correction:

$\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$: The 0.3% uncertainty [22] is propagated.

$\mathcal{A}^{\text{raw}}(\overline{K}^0)$: The uncertainty, dominated by material map uncertainty, is propagated. Additionally, this asymmetry is observed to vary as a function of the daughter pion momentum, and this dependence is not included in the Run 1 evaluation. Therefore the 0.009% difference between the asymmetries from weighted and unweighted calibration samples [218] is included as a systematic uncertainty for the Run 1 measurement.

2015 calibration data: No $\mathcal{A}^\pi$ or ε_{TOS} measurements are available for 2015 data. Systematic uncertainties are evaluated as the standard deviation of the 2016-2018 corrections.

Effective PID requirement: The uncertainty on the effective PID requirement is estimated as the change in the PID requirement that further doubles the PID inefficiency. This uncertainty is propagated to an uncertainty on $\mathcal{A}^{\text{PID}}$.

PID efficiency independence: The PID efficiencies of each of the tracks in a candidate are assumed to be independent, but this is not valid as the Cherenkov rings may overlap. The associated bias on $\mathcal{A}^{\text{PID}}$ is assessed by comparing the simulated asymmetry induced by applying all PID requirements at once to the sum of the simulated track PID asymmetries. The statistical uncertainty on this bias is evaluated using bootstrap resampling [221]. The bias is consistent with zero and the uncertainty is assigned as a systematic uncertainty. The standard deviation of the biases on PID efficiencies, which are evaluated as the fractional difference between the PID efficiency and the product of track PID efficiencies, is assigned as a systematic uncertainty on all branching fraction ratios.

Kinematic distributions for $\mathcal{A}^{\rm PID}$: As described in Section 8.2.6, the BDT requirement is not applied when evaluating $\mathcal{A}^{\rm PID}$ because the BDT uses PID variables. The absence of the BDT requirement modifies kinematic distributions, which may bias $\mathcal{A}^{\rm PID}$. This bias is assessed as the change in $\mathcal{A}^{\rm PID}$ when the BDT requirement is applied.

$\mathcal{A}^{\rm TOS}$ and $\varepsilon_{\rm L0}$ binning variable: $\mathcal{A}^{\rm TOS}$ and $\varepsilon_{\rm L0}$ are determined in bins of $(p_{\rm T}(B), y(B))$, but in reality depend on the $E_{\rm T}$ and HCAL position of all final state tracks. The bias associated with this simplification is estimated as the change in asymmetry correction when binning the measurement in the maximum $E_{\rm T}$ of the final state particles that are within HCAL acceptance and assigned as a systematic uncertainty.

$\varepsilon_{\rm TOS}$ corrections: The calibration data available for evaluation of $\varepsilon_{\rm TOS}$ in Run 1 does not permit the evaluation of higher order corrections from the overlap of other tracks in the decay, HCAL occupancy from the underlying event, and relative miscalibration of HCAL cells. The resulting bias is assessed by considering the effect of these corrections in Run 2.

Hadronic interactions: Approximately 14% of kaon and 11% of pion tracks cannot be reconstructed because of hadronic interactions. These interactions are modelled in simulation, however, there is a 10% uncertainty on the material budget. Therefore there is a fractional systematic uncertainty on the tracking efficiency for each final state particle of 10% multiplied by the fraction of tracks that cannot be reconstructed [128].

Tracking efficiency calibration simulation: The simulation samples used to obtain the data-simulation weights are corrected in $nTracks$. A systematic uncertainty of 0.4(0.8)% per track in Run 1(2) is found by correcting the calibration simulation in different variables.

Additional uncertainties on the efficiency ratio are:

Finite size of simulation sample: The uncertainty associated with the simulation sample size is propagated.

Multiple candidate removal efficiency: The occurence of multiple candidates is poorly described in simulation [194]. The change in efficiency ratio when all multiple candidates are retained is assigned as a systematic uncertainty.

Mismodelling of BDT input variables: $\chi^2_{\rm vtx}(D)$, $t/\Delta t(D)$ and $\chi^2_{\rm IP}(B)$ are mismodelled in simulation (Appendix A), which biases the efficiency of the requirement on BDT response. As discussed in Section 5.8.1, $\chi^2_{\rm vtx}(D^+)$ and $t/\Delta t(D^+)$ are removed from the BDT for the branching fraction measurement for this reason. $\chi^2_{\rm vtx}(D^0)$, $t/\Delta t(D^0)$ and $\chi^2_{\rm IP}(B)$ can be corrected using the CF decay, treating $D^0 \to K\pi$ and $D^0 \to K3\pi$ separately. A systematic uncertainty is assigned as the difference with and without this correction.

Mismodelling of χ^2_{trk}/ndf: χ^2_{trk}/ndf is mismodelled in simulation (Figure A.7). Since the χ^2_{trk} cuts applied in stripping lines are not fully efficient, this may bias the efficiencies. This bias is assessed by fitting a bifurcated Crystal Ball function to the χ^2_{trk}/ndf distributions in data and simulation, sharing the tail parameters between all tracks in all final states to ensure fit stability. The inefficiency of the stripping requirements in data and simulation are estimated by extrapolating the fitted PDF and used to calculate the bias. Inefficiencies of up to 8% and biases of up to 4% are observed, which cancel to a large extent in the efficiency ratio.

Systematic uncertainties from weighting

Uncertainties on the weighting of the simulation to agree with background-subtracted data and on the weighting of simulation to model partially reconstructed decays (Section 5.3) result in uncertainties on all corrections:

Reweighting training sample size: The uncertainty from the finite size of the training samples is evaluated as the standard deviation in the average weight when the training dataset is bootstrap resampled [221] one hundred times.

Background subtraction: Background subtraction is performed using a linear background moddel.

Partially reconstructed weight parameterisation: The power law weight function in $p(B)$ is replaced by a Gaussian plus constant weight function in $p_{\mathrm{T}}(B)$.

Partially reconstructed weight statistical uncertainty: The power law is varied by its statistical uncertainty.

Partially reconstructed weights for final state particles: As discussed in Section 5.3.9, the weight function for partially reconstructed decays does not in general correct the kinematics of final state particles, but the size of the shift due to weighting is similar to the size of the discrepancy. Therefore, for corrections that depend on the kinematics of the final state particles, the difference with and without the partially reconstructed weights is assigned as a systematic uncertainty.

Table 8.22: Systematic uncertainties on $\mathcal{A}^{CP}$ in percent, averaged over all D^0 decay modes.

Source	$D_s^- D^0$		$D^- D^0$		$D^{*-} D^0$	
	Run 1	Run 2	Run 1	Run 2	Run 1	Run 2
$\mathcal{A}^{\text{raw}}(B^- \to D_{(s)}^{(*)-} D^0)$						
Double charm model	0.01	0.02	0.11	0.08	0.21	0.18
Single charm model	0.05	0.05	-	-	-	-
Cross-feed model	-	-	0.00	0.00	-	-
Combinatorial model	0.01	0.00	0.14	0.13	0.52	0.15
External inputs	0.05	0.06	0.01	0.00	0.04	0.01
Total	0.07	0.08	0.18	0.15	0.56	0.23
$\mathcal{A}^{\text{raw}}(B^- \to D_{(s)}^{(*)-} D^{*0})$						
Double charm model	0.14	0.23	1.10	1.04	1.51	1.19
Cross-feed model	-	-	0.00	0.00	-	-
Combinatorial model	0.08	0.04	1.14	0.92	4.49	1.02
External inputs	0.55	0.38	0.14	0.12	0.11	0.11
Total	0.57	0.45	1.59	1.40	4.74	1.57
$\mathcal{A}^{\text{raw}}(B^- \to D_s^{*-} D^0)$						
Double charm model	0.78	0.51				
Combinatorial model	0.55	0.22				
External inputs	0.89	0.73				
Total	1.31	0.91				
$\mathcal{A}^{\text{P}}$	0.42	0.43	0.41	0.43	0.48	0.48
$\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$	0.30	0.30	0.30	0.30	0.30	0.30
$\mathcal{A}^{\text{K}\pi}$	0.28	0.11	0.04	0.04	0.10	0.00
$\mathcal{A}^{\pi}$	0.09	0.09	0.06	0.06	0.18	0.17
$\mathcal{A}^{\text{PID}}$	0.29	0.03	0.25	0.11	0.55	0.10
$\mathcal{A}^{\text{TIS}}$	0.08	0.10	0.08	0.10	0.09	0.11
$\mathcal{A}^{\text{TOS}}$	0.01	0.03	0.01	0.02	0.01	0.01
Weighting	0.01	0.00	0.04	0.00	0.01	0.00
Part. rec. weighting	0.03	0.02	0.02	0.01	0.03	0.01
Total	0.67	0.55	0.58	0.55	0.82	0.61

Table 8.23: Fractional systematic uncertainties on branching fraction ratios in percent, combined over all D^0 decay modes.

Source	$R(B^- \to D^- D^0)$		$R(B^- \to D^{*-} D^0)$	
	Run 1	Run 2	Run 1	Run 2
Double charm model	0.8	0.1	1.2	0.4
Single charm model	0.3	0.3	-	-
Cross-feed model	0.0	0.0	0.0	0.0
Combinatorial model	0.1	0.0	0.5	0.2
Hardware trigger efficiency	0.6	0.1	1.4	0.2
Tracking efficiency	0.4	0.3	1.4	1.1
PID efficiency	0.8	0.6	1.6	0.7
Simulation sample size	0.8	0.9	1.1	1.1
Multiple candidate removal	0.1	0.1	0.2	0.4
MVA variable mismodelling	0.8	0.1	1.0	0.4
Track χ^2 mismodelling	0.5	0.1	0.5	0.1
Weighting	0.1	0.0	0.1	0.0
Total	1.8	1.2	3.3	1.9

8.5 Charge-parity asymmetry results

In this section, *CP* asymmetry measurements are presented and validated, and the measurements using data from each final state and operational run are averaged. The *CP* asymmetries are obtained as specified in Equation 4.6 by subtracting the corrections derived in Section 8.2 from the raw asymmetries measured in Section 8.1.

When averaging measurements from multiple datasets, weights

$$\boldsymbol{w} = \frac{\boldsymbol{V}^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\boldsymbol{V}^{-1}\mathbf{1}}, \tag{8.19}$$

where $\boldsymbol{V}$ is the covariance matrix, and $\mathbf{1}$ is a vector of ones, are assigned to each dataset. These weights minimise the uncertainty on the average [222]. However, correlations between datasets for many sources of systematic uncertainty are not precisely known, including for sources of large uncertainty, such as the combinatorial background shape. To obtain $\boldsymbol{V}$, 100% correlation in uncertainty between datasets is assumed for all sources of systematic uncertainty where a correlation might be expected. If the correlated uncertainties are not large compared to uncorrelated uncertainties, this assumption corresponds to the least additional information in the covariance matrix, which results in the largest and therefore most conservative uncertainty on the combined measurement, reflecting the limited knowledge of the correlations. Conversely, if correlated uncertainties are large compared to uncorrelated uncertainties, higher correlations increase the

Table 8.24: Weights for each data taking period and final state in their combination.

$\mathcal{A}^{CP}$	Final State	Run 1	Run 2
$D_s^- D^0$	$KK\pi, K\pi$	0.20	0.52
	$KK\pi, K3\pi$	0.12	0.16
$D_s^{*-} D^0$	$KK\pi, K\pi$	0.11	0.59
	$KK\pi, K3\pi$	0.02	0.28
$D_s^- D^{*0}$	$KK\pi, K\pi$	0.12	0.45
	$KK\pi, K3\pi$	0.12	0.31
$D^- D^0$	$K\pi\pi, K\pi$	0.13	0.53
	$K\pi\pi, K3\pi$	0.07	0.27
$D^- D^{*0}$	$K\pi\pi, K\pi$	0.17	0.52
	$K\pi\pi, K3\pi$	0.01	0.31
$D^{*-} D^0$	$K\pi\pi_s, K\pi$	0.08	0.34
	$K\pi\pi_s, K3\pi$	0.05	0.24
	$K3\pi\pi_s, K\pi$	0.05	0.24
$D^{*-} D^{*0}$	$K\pi\pi_s, K\pi$	0.00	0.43
	$K\pi\pi_s, K3\pi$	0.03	0.29
	$K3\pi\pi_s, K\pi$	0.00	0.24

information in the covariance matrix; this regime is indicated by the presence of negative weights in Equation 8.19 [209].

To address the presence of negative weights due to limited knowledge of correlations, the measurement corresponding to the most negative weight is removed from the average and the weights are re-computed. This process is repeated until no negative weights remain [209]. The removed measurements correspond to measurements with large correlated systematic uncertainties, and therefore contribute little to the overall measurement and erroneously reduce the overall uncertainty if they are included. In practice, negative weights are only obtained for $\mathcal{A}^{CP}(B^- \to D^{*-}D^{*0})$. The weights derived using this procedure are shown in Table 8.24.

Measured *CP* asymmetries, including weighted averages obtained as described above, are presented in Table 8.25. No evidence of *CP* violation is seen. This table also shows the relative size of the statistical and systematic uncertainties for each dataset. Correlations between the $\mathcal{A}^{CP}$ measurements are given in Table 8.26. Tables 8.27 and 8.28 also specify the correlations in the statistical and systematic uncertainties, respectively.

Table 8.25: $\mathcal{A}^{CP}$ measurements in percent in each final state and operational run, and the combined measurements. The first uncertainty is statistical, the second is internal systematic, and each measurement has an additional uncertainty of 0.3% from $\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$ that is not tabulated.

Decay	Final state	Run 1	Run 2	Run 1+2
$D_s^- D^0$	$KK\pi, K\pi$	$-0.4 \pm 0.6 \pm 0.6$	$0.8 \pm 0.3 \pm 0.4$	$0.5 \pm 0.3 \pm 0.4$
	$KK\pi, K3\pi$	$0.4 \pm 0.9 \pm 0.6$	$1.0 \pm 0.4 \pm 0.5$	$0.8 \pm 0.4 \pm 0.5$
	combination	$-0.1 \pm 0.5 \pm 0.6$	$0.8 \pm 0.3 \pm 0.4$	$0.5 \pm 0.2 \pm 0.5$
$D_s^{*-} D^0$	$KK\pi, K\pi$	$2.5 \pm 3.1 \pm 1.2$	$-0.9 \pm 1.4 \pm 1.0$	$-0.3 \pm 1.3 \pm 1.0$
	$KK\pi, K3\pi$	$8.5 \pm 5.9 \pm 2.2$	$-1.6 \pm 2.2 \pm 0.9$	$-0.6 \pm 2.0 \pm 1.0$
	combination	$3.7 \pm 2.7 \pm 1.3$	$-1.1 \pm 1.2 \pm 0.9$	$-0.5 \pm 1.1 \pm 1.0$
$D_s^- D^{*0}$	$KK\pi, K\pi$	$-3.7 \pm 2.3 \pm 0.9$	$2.4 \pm 1.2 \pm 0.6$	$1.1 \pm 1.1 \pm 0.6$
	$KK\pi, K3\pi$	$-0.1 \pm 2.5 \pm 0.8$	$1.6 \pm 1.4 \pm 0.6$	$1.2 \pm 1.2 \pm 0.6$
	combination	$-2.0 \pm 1.7 \pm 0.8$	$2.1 \pm 0.9 \pm 0.6$	$1.1 \pm 0.8 \pm 0.6$
$D^- D^0$	$K\pi\pi, K\pi$	$1.2 \pm 2.9 \pm 0.5$	$1.7 \pm 1.4 \pm 0.4$	$1.6 \pm 1.3 \pm 0.4$
	$K\pi\pi, K3\pi$	$5.0 \pm 4.1 \pm 0.6$	$4.2 \pm 1.9 \pm 0.6$	$4.3 \pm 1.7 \pm 0.6$
	combination	$2.4 \pm 2.4 \pm 0.5$	$2.6 \pm 1.1 \pm 0.4$	$2.5 \pm 1.0 \pm 0.4$
$D^- D^{*0}$	$K\pi\pi, K\pi$	$1.9 \pm 5.3 \pm 1.1$	$0.2 \pm 2.8 \pm 1.6$	$0.6 \pm 2.5 \pm 1.4$
	$K\pi\pi, K3\pi$	$-0.9 \pm 9.7 \pm 4.5$	$-2.0 \pm 3.3 \pm 2.4$	$-1.9 \pm 3.1 \pm 2.5$
	combination	$1.4 \pm 4.7 \pm 1.7$	$-0.6 \pm 2.2 \pm 1.4$	$-0.2 \pm 2.0 \pm 1.4$
$D^{*-} D^0$	$K\pi\pi_s, K\pi$	$9.5 \pm 5.6 \pm 0.9$	$3.0 \pm 2.8 \pm 0.5$	$4.2 \pm 2.5 \pm 0.6$
	$K\pi\pi_s, K3\pi$	$-2.9 \pm 7.2 \pm 1.0$	$3.5 \pm 3.3 \pm 0.6$	$2.3 \pm 3.0 \pm 0.6$
	$K3\pi\pi_s, K\pi$	$3.9 \pm 7.1 \pm 1.0$	$2.8 \pm 3.3 \pm 0.5$	$3.0 \pm 3.0 \pm 0.5$
	combination	$4.6 \pm 3.7 \pm 0.9$	$3.1 \pm 1.8 \pm 0.5$	$3.3 \pm 1.6 \pm 0.6$
$D^{*-} D^{*0}$	$K\pi\pi_s, K\pi$	$-7.4 \pm 7.1 \pm 13.0$	$5.7 \pm 3.5 \pm 1.5$	$5.0 \pm 3.3 \pm 1.6$
	$K\pi\pi_s, K3\pi$	$0.0 \pm 8.7 \pm 3.1$	$0.9 \pm 3.7 \pm 1.9$	$0.8 \pm 3.4 \pm 2.0$
	$K3\pi\pi_s, K\pi$	$5.6 \pm 8.9 \pm 7.5$	$-1.5 \pm 3.8 \pm 2.6$	$-1.4 \pm 3.7 \pm 2.7$
	combination	$1.7 \pm 6.3 \pm 4.8$	$2.4 \pm 2.1 \pm 1.6$	$2.3 \pm 2.1 \pm 1.7$

8.5.1 Validation

As part of the unblinding procedure, the differences in the *CP* asymmetries between magnet polarities (Table 8.29), final states (Tables 8.30-8.32), and operational runs (Tables 8.33) were measured. Additionally, the measurements of $\mathcal{A}^{CP}(B^- \to D_{(s)}^- D^0)$ with Run 1 data were compared to the earlier LHCb measurement that used only Run 1 data [173]. In all cases, a difference of more than two standard deviations would trigger further investigation, and a difference of more than five standard deviations would be considered an error that would need to be resolved before the analysis could progress.

A difference of 2.0σ is observed for $\mathcal{A}^{CP}(B^- \to D^- D^0)$ when comparing magnet polarities and for $\mathcal{A}^{CP}(B^- \to D_s^- D^{*0})$ when comparing Run 1 and Run 2. All other comparisons pass the unblinding check of consistency within 2σ. This is deemed acceptable since 23 independent comparisons are performed.

Table 8.26: Total correlations between $\mathcal{A}^{CP}$ measurements after combining Run 1 and 2 and final states.

	$D_s^- D^0$	$D_s^{*-} D^0$	$D_s^- D^{*0}$	$D^- D^0$	$D^- D^{*0}$	$D^{*-} D^0$	$D^{*-} D^{*0}$
$D_s^- D^0$	1.000						
$D_s^{*-} D^0$	0.335	1.000					
$D_s^- D^{*0}$	0.431	−0.282	1.000				
$D^- D^0$	0.386	0.186	0.234	1.000			
$D^- D^{*0}$	0.183	0.227	0.180	0.195	1.000		
$D^{*-} D^0$	0.286	0.161	0.177	0.166	0.130	1.000	
$D^{*-} D^{*0}$	0.190	0.227	0.211	0.168	0.311	0.189	1.000

Table 8.27: Correlations in statistical uncertainty between $\mathcal{A}^{CP}$ measurements after combining Run 1 and 2 and final states.

	$D_s^- D^0$	$D_s^{*-} D^0$	$D_s^- D^{*0}$	$D^- D^0$	$D^- D^{*0}$	$D^{*-} D^0$	$D^{*-} D^{*0}$
$D_s^- D^0$	1.000						
$D_s^{*-} D^0$	0.092	1.000					
$D_s^- D^{*0}$	−0.025	−0.628	1.000				
$D^- D^0$	−0.000	0.000	0.000	1.000			
$D^- D^{*0}$	0.000	0.000	0.000	0.060	1.000		
$D^{*-} D^0$	0.000	−0.000	−0.000	−0.000	−0.000	1.000	
$D^{*-} D^{*0}$	0.000	0.000	0.000	0.000	0.000	0.054	1.000

Table 8.28: Correlations in systematic uncertainties excluding $\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$ between $\mathcal{A}^{CP}$ measurements after combining Run 1 and 2 and final states.

	$D_s^- D^0$	$D_s^{*-} D^0$	$D_s^- D^{*0}$	$D^- D^0$	$D^- D^{*0}$	$D^{*-} D^0$	$D^{*-} D^{*0}$
$D_s^- D^0$	1.000						
$D_s^{*-} D^0$	0.434	1.000					
$D_s^- D^{*0}$	0.714	−0.007	1.000				
$D^- D^0$	0.895	0.512	0.736	1.000			
$D^- D^{*0}$	0.282	0.536	0.438	0.530	1.000		
$D^{*-} D^0$	0.846	0.516	0.708	0.962	0.574	1.000	
$D^{*-} D^{*0}$	0.273	0.476	0.488	0.550	0.834	0.616	1.000

Table 8.29: Difference in $\mathcal{A}^{CP}$ between magnet polarities.

$\mathcal{A}^{CP}$	Run 1	Run 2	Run 1+2
$D_s^- D^0$	$-1.3 \pm 1.0 \pm 0.4$	$0.6 \pm 0.5 \pm 0.4$	$0.0 \pm 0.5 \pm 0.3$
$D_s^{*-} D^0$	$-5.2 \pm 5.4 \pm 0.8$	$2.6 \pm 2.4 \pm 0.7$	$1.4 \pm 2.2 \pm 0.6$
$D_s^- D^{*0}$	$2.5 \pm 3.3 \pm 0.6$	$0.1 \pm 1.7 \pm 0.5$	$0.9 \pm 1.5 \pm 0.5$
$D^- D^0$	$5.7 \pm 4.8 \pm 0.6$	$3.9 \pm 2.2 \pm 0.3$	$4.2 \pm 2.0 \pm 0.3$
$D^- D^{*0}$	$-16.9 \pm 13.2 \pm 2.3$	$-0.3 \pm 4.3 \pm 0.6$	$-2.8 \pm 3.9 \pm 0.5$
$D^{*-} D^0$	$5.0 \pm 7.5 \pm 0.7$	$-2.0 \pm 3.6 \pm 0.4$	$-0.8 \pm 3.2 \pm 0.4$
$D^{*-} D^{*0}$	$14.5 \pm 11.0 \pm 3.8$	$4.0 \pm 5.0 \pm 1.1$	$3.2 \pm 4.8 \pm 1.2$

Table 8.30: Difference $\Delta(f_i, f_j)$ in $\mathcal{A}^{CP}$ between final states f_i and f_j in Run 1 data.

$\mathcal{A}^{CP}$	$\Delta(K\pi, K3\pi)$		
$D_s^- D^0$	$-0.8 \pm 1.1 \pm 0.4$		
$D_s^{*-} D^0$	$-5.9 \pm 6.6 \pm 1.8$		
$D_s^- D^{*0}$	$-3.6 \pm 3.4 \pm 0.6$		
$\mathcal{A}^{CP}$	$\Delta(K\pi, K3\pi)$		
$D^- D^0$	$-3.8 \pm 5.0 \pm 0.2$		
$D^- D^{*0}$	$2.8 \pm 11.0 \pm 3.5$		
$\mathcal{A}^{CP}$	$\Delta(K\pi K\pi, K\pi K3\pi)$	$\Delta(K\pi K\pi, K3\pi K\pi)$	$\Delta(K\pi K3\pi, K3\pi K\pi)$
$D^{*-} D^0$	$12.4 \pm 9.1 \pm 0.3$	$5.6 \pm 9.0 \pm 0.4$	$-6.8 \pm 10.1 \pm 0.3$
$D^{*-} D^{*0}$	$-7.5 \pm 11.2 \pm 10.3$	$-13.1 \pm 11.4 \pm 5.7$	$-5.6 \pm 12.4 \pm 4.7$

Table 8.31: Difference $\Delta(f_i, f_j)$ in $\mathcal{A}^{CP}$ between final states f_i and f_j in Run 2 data.

$\mathcal{A}^{CP}$	$\Delta(K\pi, K3\pi)$		
$D_s^- D^0$	$-0.2 \pm 0.5 \pm 0.2$		
$D_s^{*-} D^0$	$0.7 \pm 2.6 \pm 0.5$		
$D_s^- D^{*0}$	$0.8 \pm 1.8 \pm 0.2$		
$\mathcal{A}^{CP}$	$\Delta(K\pi, K3\pi)$		
$D^- D^0$	$-2.5 \pm 2.3 \pm 0.3$		
$D^- D^{*0}$	$2.1 \pm 4.4 \pm 2.6$		
$\mathcal{A}^{CP}$	$\Delta(K\pi K\pi, K\pi K3\pi)$	$\Delta(K\pi K\pi, K3\pi K\pi)$	$\Delta(K\pi K3\pi, K3\pi K\pi)$
$D^{*-} D^0$	$-0.5 \pm 4.3 \pm 0.2$	$0.1 \pm 4.3 \pm 0.3$	$0.6 \pm 4.7 \pm 0.2$
$D^{*-} D^{*0}$	$4.8 \pm 5.1 \pm 1.3$	$7.2 \pm 5.2 \pm 2.5$	$2.4 \pm 5.3 \pm 1.3$

Table 8.32: Difference $\Delta(f_i, f_j)$ in $\mathcal{A}^{CP}$ between final states f_i and f_j in Run 1+2 data.

$\mathcal{A}^{CP}$	$\Delta(K\pi, K3\pi)$		
$D_s^- D^0$	$-0.3 \pm 0.5 \pm 0.2$		
$D_s^{*-} D^0$	$0.3 \pm 2.4 \pm 0.6$		
$D_s^- D^{*0}$	$-0.1 \pm 1.6 \pm 0.2$		
$\mathcal{A}^{CP}$	$\Delta(K\pi, K3\pi)$		
$D^- D^0$	$-2.7 \pm 2.1 \pm 0.2$		
$D^- D^{*0}$	$2.5 \pm 4.0 \pm 2.4$		
$\mathcal{A}^{CP}$	$\Delta(K\pi K\pi, K\pi K3\pi)$	$\Delta(K\pi K\pi, K3\pi K\pi)$	$\Delta(K\pi K3\pi, K3\pi K\pi)$
$D^{*-} D^0$	$1.9 \pm 3.9 \pm 0.2$	$1.2 \pm 3.9 \pm 0.2$	$-0.7 \pm 4.3 \pm 0.2$
$D^{*-} D^{*0}$	$4.3 \pm 4.8 \pm 0.9$	$6.4 \pm 5.0 \pm 2.0$	$2.1 \pm 5.1 \pm 1.2$

8.6 Branching fraction results

Combining measurements from all datasets with weights evaluated using the method described in the previous section results in measurements of

$$R(B^- \to D^- D^0) = (7.25 \pm 0.09 \pm 0.09) \times 10^{-2},$$

and

$$R(B^- \to D^- D^0) = 0.271 \pm 0.007 \pm 0.005,$$

where the first uncertainty is statistical and the second is systematic. The correlation between these measurements is −47%. These values are in good agreement with the world averages of $(7.4 \pm 1.1) \times 10^{-2}$ and 0.29 ± 0.05, respectively [22].

8.6.1 Validation

To validate the measurement, the ratio

$$R(D^0) = \frac{\mathcal{B}(D^0 \to K^- \pi^+)}{\mathcal{B}(D^0 \to K^- \pi^+ \pi^- \pi^+)} \tag{8.20}$$

is measured to be $0.524 \pm 0.005 \pm 0.019$ and compared to the world average of 0.480 ± 0.007 [22], which is dominated by a measurement from CLEO [223]. The uncertainty is dominated by uncertainty on the material budget.

Measurements of the branching fraction ratios and $R(D^0)$ are compared between Run 1 and 2 and between the decay channels in Tables 8.34-8.36. To validate the L0 trigger efficiency corrections, the branching fractions are measured separately for TOS and TISnotTOS candidates in Tables 8.37-8.39. When dividing the dataset in this way, the uncertainty is dominated by the uncertainty associated with corrections to $\varepsilon_{\mathrm{L0}}$. In all cases, the measurements agree to better than 3 standard deviations.

8.6.2 Stability tests

The stability of the central value against different selection is evaluated by using no MVA cut, a tight MVA cut (half the efficiency of the nominal cut) and a tight PID cut ($\mathrm{DLL}_{K\pi}(K) > -1$, $\mathrm{DLL}_{K\pi}(\pi) < 1$). Tables 8.40-8.42 show the relative bias on each value of $\mathcal{R}$. This bias is defined as the value under the modified selection divided by the nominal value. Correlations between the datasets are accounted for. The significance of the bias is also quoted. When applying a tight MVA cut, the systematic uncertainty from reweighting of MVA variables becomes dominant. When applying tight PID cuts, systematic uncertainties from the binning of PID efficiency calibration samples and PID efficiency track independence become dominant.

Table 8.33: Difference in $\mathcal{A}^{CP}$ between Run 1 and Run 2.

$\mathcal{A}^{CP}$	Δ
$D_s^-D^0$	$-0.9 \pm 0.6 \pm 0.4$
$D_s^{*-}D^0$	$4.8 \pm 3.0 \pm 0.7$
$D_s^-D^{*0}$	$-4.1 \pm 1.9 \pm 0.7$
D^-D^0	$-0.1 \pm 2.6 \pm 0.3$
D^-D^{*0}	$2.0 \pm 5.2 \pm 0.6$
$D^{*-}D^0$	$1.5 \pm 4.2 \pm 0.7$
$D^{*-}D^{*0}$	$-0.7 \pm 6.6 \pm 3.5$

Table 8.34: $R(D^-D^0/D_s^-D^0)/10^{-2}$ in each final state and for each data taking period. The first uncertainty is statistical and the second is systematic.

Final state	Run 1	Run 2
$K\pi$	$6.88 \pm 0.24 \pm 0.12$	$7.35 \pm 0.12 \pm 0.11$
$K3\pi$	$6.93 \pm 0.38 \pm 0.23$	$7.40 \pm 0.18 \pm 0.15$

Table 8.35: $R(D^{*-}D^0/D^-D^0)$ in each final state and for each data taking period. The first uncertainty is statistical and the second is systematic.

Final state	Run 1	Run 2
$K\pi$	$0.328 \pm 0.023 \pm 0.011$	$0.256 \pm 0.009 \pm 0.005$
$K3\pi$	$0.316 \pm 0.033 \pm 0.015$	$0.272 \pm 0.012 \pm 0.008$

Table 8.36: $R(D^0)$ as measured in each decay channel and for each data taking period. The first uncertainty is statistical and the second is systematic.

Channel	Run 1	Run 2
$B^- \to D_s^-D^0$	$0.562 \pm 0.007 \pm 0.030$	$0.518 \pm 0.003 \pm 0.020$
$B^- \to D^-D^0$	$0.558 \pm 0.036 \pm 0.030$	$0.514 \pm 0.015 \pm 0.021$
$B^- \to D^{*-}{}_{[K\pi\pi_s]}D^0$	$0.579 \pm 0.062 \pm 0.035$	$0.485 \pm 0.024 \pm 0.020$
$B^- \to D^{*-}D^0{}_{[K\pi]}$	$0.558 \pm 0.059 \pm 0.032$	$0.513 \pm 0.028 \pm 0.026$

Table 8.37: $R(D^-D^0/D_s^-D^0)/10^{-2}$ in each final state and for each data taking period, separated by L0 trigger decision. The first uncertainty is statistical and the second is systematic.

Final state	Run 1		Run 2	
	TOS	TISnotTOS	TOS	TISnotTOS
$K\pi$	$6.90 \pm 0.30 \pm 0.19$	$7.00 \pm 0.41 \pm 0.19$	$7.18 \pm 0.15 \pm 0.15$	$7.66 \pm 0.20 \pm 0.16$
$K3\pi$	$6.88 \pm 0.50 \pm 0.38$	$6.93 \pm 0.58 \pm 0.43$	$7.42 \pm 0.24 \pm 0.23$	$7.42 \pm 0.27 \pm 0.22$

Table 8.38: $R(D^{*-}D^0/D^-D^0)$ in each final state and for each data taking period, separated by L0 trigger decision. The first uncertainty is statistical and the second is systematic.

Final state	Run 1		Run 2	
	TOS	TISnotTOS	TOS	TISnotTOS
$K\pi$	$0.316 \pm 0.026 \pm 0.013$	$0.346 \pm 0.044 \pm 0.019$	$0.254 \pm 0.011 \pm 0.006$	$0.273 \pm 0.016 \pm 0.010$
$K3\pi$	$0.281 \pm 0.040 \pm 0.018$	$0.420 \pm 0.063 \pm 0.029$	$0.254 \pm 0.015 \pm 0.010$	$0.318 \pm 0.024 \pm 0.012$

Table 8.39: $R(D^0)$ as measured in each decay channel and for each data taking period, separated by L0 trigger decision. The first uncertainty is statistical and the second is systematic.

Channel	Run 1		Run 2	
	TOS	TISnotTOS	TOS	TISnotTOS
$B^- \to D_s^- D^0$	$0.576 \pm 0.010 \pm 0.070$	$0.551 \pm 0.011 \pm 0.045$	$0.508 \pm 0.004 \pm 0.022$	$0.538 \pm 0.005 \pm 0.024$
$B^- \to D^- D^0$	$0.577 \pm 0.048 \pm 0.066$	$0.556 \pm 0.056 \pm 0.045$	$0.492 \pm 0.019 \pm 0.023$	$0.556 \pm 0.024 \pm 0.028$
$B^- \to D^{*-}{}_{[K\pi\pi_s]} D^0$	$0.649 \pm 0.093 \pm 0.071$	$0.459 \pm 0.078 \pm 0.046$	$0.492 \pm 0.031 \pm 0.023$	$0.478 \pm 0.040 \pm 0.027$
$B^- \to D^{*-} D^0{}_{[K\pi]}$	$0.590 \pm 0.078 \pm 0.057$	$0.504 \pm 0.091 \pm 0.044$	$0.483 \pm 0.031 \pm 0.031$	$0.616 \pm 0.052 \pm 0.056$

Table 8.40: Bias on $R(D^-D^0/D_s^-D^0)$ for different selections and corresponding significance

Run	Final state	No MVA	Tight MVA	Tight PID
Run 1	$K\pi$	$1.062 \pm 0.044 \pm 0.006(+1.4\sigma)$	$0.925 \pm 0.032 \pm 0.004(-2.4\sigma)$	$1.032 \pm 0.023 \pm 0.024(+1.0\sigma)$
	$K3\pi$	$1.096 \pm 0.095 \pm 0.027(+1.0\sigma)$	$1.008 \pm 0.052 \pm 0.031(+0.1\sigma)$	$1.048 \pm 0.047 \pm 0.034(+0.8\sigma)$
Run 2	$K\pi$	$0.993 \pm 0.015 \pm 0.004(-0.5\sigma)$	$0.977 \pm 0.015 \pm 0.010(-1.2\sigma)$	$0.989 \pm 0.009 \pm 0.024(-0.4\sigma)$
	$K3\pi$	$0.989 \pm 0.028 \pm 0.017(-0.3\sigma)$	$0.945 \pm 0.021 \pm 0.020(-1.9\sigma)$	$1.031 \pm 0.014 \pm 0.025(+1.1\sigma)$

Table 8.41: Bias on $R(D^{*-}D^0/D^-D^0)$ for different selections and corresponding significance

Run	Final state	No MVA	Tight MVA	Tight PID
Run 1	$K\pi$	$0.968 \pm 0.053 \pm 0.011(-0.6\sigma)$	$0.969 \pm 0.070 \pm 0.010(-0.4\sigma)$	$0.891 \pm 0.054 \pm 0.055(-1.4\sigma)$
	$K3\pi$	$0.850 \pm 0.099 \pm 0.017(-1.5\sigma)$	$1.163 \pm 0.104 \pm 0.029(+1.5\sigma)$	$0.776 \pm 0.085 \pm 0.039(-2.4\sigma)$
Run 2	$K\pi$	$1.048 \pm 0.023 \pm 0.003(+2.0\sigma)$	$1.006 \pm 0.034 \pm 0.016(+0.1\sigma)$	$0.998 \pm 0.019 \pm 0.022(-0.1\sigma)$
	$K3\pi$	$1.063 \pm 0.045 \pm 0.011(+1.4\sigma)$	$0.986 \pm 0.044 \pm 0.019(-0.3\sigma)$	$0.931 \pm 0.027 \pm 0.021(-2.0\sigma)$

Table 8.42: Bias on $R(D^0)$ for different selections and corresponding significance.

Run	Channel	No MVA	Tight MVA	Tight PID
Run 1	$B^- \to D_s^- D^0$	$0.969 \pm 0.007 \pm 0.035(-0.9\sigma)$	$1.073 \pm 0.013 \pm 0.039(+1.8\sigma)$	$0.928 \pm 0.009 \pm 0.023(-2.9\sigma)$
	$B^- \to D^- D^0$	$0.939 \pm 0.090 \pm 0.053(-0.6\sigma)$	$0.985 \pm 0.060 \pm 0.051(-0.2\sigma)$	$0.915 \pm 0.045 \pm 0.028(-1.6\sigma)$
	$B^- \to D^{*-}{}_{[K\pi\pi_s]} D^0$	$1.069 \pm 0.092 \pm 0.041(+0.7\sigma)$	$0.820 \pm 0.080 \pm 0.048(-1.9\sigma)$	$1.051 \pm 0.121 \pm 0.034(+0.4\sigma)$
	$B^- \to D^{*-} D^0{}_{[K\pi]}$	$1.002 \pm 0.073 \pm 0.042(+0.0\sigma)$	$0.923 \pm 0.096 \pm 0.024(-0.8\sigma)$	$0.863 \pm 0.089 \pm 0.054(-1.3\sigma)$
Run 2	$B^- \to D_s^- D^0$	$1.001 \pm 0.003 \pm 0.017(+0.1\sigma)$	$1.007 \pm 0.006 \pm 0.012(+0.5\sigma)$	$0.983 \pm 0.003 \pm 0.029(-0.6\sigma)$
	$B^- \to D^- D^0$	$1.005 \pm 0.032 \pm 0.025(+0.1\sigma)$	$1.041 \pm 0.028 \pm 0.017(+1.3\sigma)$	$0.942 \pm 0.015 \pm 0.027(-1.8\sigma)$
	$B^- \to D^{*-}{}_{[K\pi\pi_s]} D^0$	$0.991 \pm 0.035 \pm 0.023(-0.2\sigma)$	$1.062 \pm 0.052 \pm 0.026(+1.1\sigma)$	$1.010 \pm 0.031 \pm 0.035(+0.2\sigma)$
	$B^- \to D^{*-} D^0{}_{[K\pi]}$	$0.990 \pm 0.030 \pm 0.021(-0.3\sigma)$	$1.100 \pm 0.055 \pm 0.018(+1.7\sigma)$	$0.976 \pm 0.029 \pm 0.039(-0.5\sigma)$

Beauty to double charm decays in Run 3 and beyond

To prepare for a five-fold increase in instantaneous luminosity, the LHCb detector has been upgraded in advance of Run 3, as described in Section 3.3. These upgrades included the removal of the hardware trigger and the use of real-time alignment and calibration to enable exclusive real-time selection of decay chains in the final stage of the software trigger. This redevelopment of the trigger system is expected to produce large increases in signal efficiencies, particularly for fully hadronic final states such as those described in these thesis, which suffered from low hardware trigger efficiencies.

This chapter firstly describes the development of HLT2 trigger selections to select $B_{(c)}^+$ meson decays to two charm mesons in Run 3 within the new trigger paradigm. The trigger line requirements in terms of final states and target event rates are specified in Section 9.1. In Section 9.2, a generic b-hadron neural network is developed to reduce the rate of background candidates that are not kinematically and geometrically consistent with a b-hadron candidate. The utility of this neural network extends beyond double charm decays, and its output is currently used as a discriminating variable in over two hundred HLT2 lines selecting beauty hadron to open charm decays in Run 3. In Section 9.3, trigger line thresholds on discriminating variables are optimised for signal efficiency and event rate.

Up to four further data taking periods, at progressively higher luminosities and centre-of-mass energies, are foreseen for the LHC and the LHCb experiment, as specified in Section 3.1. The substantial increases in luminosity, in particular, will vastly increase the size of the LHCb dataset and accelerate the pace at which the LHCb experiment is able to advance understanding of fundamental physics. Experimental propsects for beauty meson to double charm decays in future data taking periods, based on the scheduled integrated luminosities and the expected gain in signal efficiencies evaluated in this chapter, are discussed in Section 9.4.

9.1 Trigger line specification

Trigger lines are developed for the exclusive selection of $B_{(c)}^+$ decays to $D_{(s)}^{(*)+}\overset{(-)}{D}{}^0$ final states. D^+ mesons are selected in the $K^-\pi^+\pi^+$ or $K^-K^+\pi^+$ final states; D_s^+ mesons are selected in the $K^-K^+\pi^+$ final state; D^{*+} mesons are selected in the $D^0\pi^+$ final state; and D^0 mesons are selected in the $K^{\mp}\pi^{\pm}(\pi^{+}\pi^{-})$, $K^{-}K^{+}(\pi^{+}\pi^{-})$, and $\pi^{+}\pi^{-}\pi^{+}\pi^{-}$ final states. For B_c^+ decays, the D^0 mesons are additionally selected in the CP-odd $K_S^0\pi^+\pi^-$, $K_S^0K^+K^-$, and $K_S^0K^{\mp}\pi^{\pm}$ final states that will be useful for a potential measurement of γ. $K_S^0 \to \pi^+\pi^-$ candidates are reconstructed as either a pair of long tracks ($K^0_{\mathrm{S}\,[\mathrm{LL}]}$) or a pair of downstream tracks ($K^0_{\mathrm{S}\,[\mathrm{DD}]}$) in separate trigger lines.

The HLT2 output bandwidth limit is 10 GB/s, which restricts the rate at which trigger lines can select events. In 2022 and early 2023 data taking, the instantaneous luminosity is similar to that in Run 2, which is five times lower than the planned eventual instantaneous luminosity for Run 3 data taking. Furthemore, detector commissioning, calibration, and alignment will not be complete until later in 2023. Moreover, the quality of the agreement between data and simulation is not known, and reconstruction and identification tools must be tuned on data before they are reliable. Consequently, tight selections, even with a high signal efficiency in simulation, may introduce a large signal rejection in data.

Owing to the low instantaneous luminosity and the limited reliability of tight selections, the initial development of trigger selections described in this chapter targets a background efficiency of below 0.01%, corresponding to a 100 Hz event selection rate for a 1 MHz HLT1 output rate. These selections are optimised to achieve an event rate below the target while maximising signal efficiency. Once the improvements described above are complete, the selections must be revised to further reduce the background efficiency.

9.2 Generic beauty hadron neural network

An MVA is trained to discriminate between b-hadron decays and candidates from combinatorial background. This MVA is evaluated at the final filtering stage for HLT2 lines, and takes as an input the full b-hadron candidate. The training variables are the η, $\log(p_{\mathrm{T}})$, $\log(1-\mathrm{DIRA})$, $\log(\chi^2_{\mathrm{vtx}})$, $\log(\chi^2_{\mathrm{IP}})$, and $\log(\mathrm{IP})$ of the b hadron. The MVA is trained using simulated signal and minimum bias candidates selected by the HLT2 line corresponding to the signal decay for $B_s^0 \to (D_s^- \to K^-K^+\pi^-)K^+$, $B_s^0 \to (D_s^- \to K^-K^+\pi^-)\pi^+$, $B^- \to (D^0 \to K^-K^+)K^-$, $B^- \to (D^0 \to K_S^0\pi^+\pi^-)K^-$, and $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\pi^-$ decays.

The MVA architecture is a *monotonic Lipschitz neural network* [224]. This architecture is based on a *feed-forward neural network*, which consists of layers of *nodes*. Each node outputs a response as a function of the output of all nodes in the previous layer. The input layer consists of one node for each training variable that outputs the value of the corresponding variable; this is followed by *hidden* layers of nodes, and an output layer that evaluates and outputs a scalar response.

A monotonic Lipschitz neural network is robust to fluctuations in the training data and small changes in the experimental conditions, such as the calibration. This property is enforced by constraining the functions that transform node inputs into outputs, and therefore constraining the magnitude of the gradient of the response function with respect to the input paramaters. This robustness reduces bias in the simulated selection efficiency from disagreement between data and simulation, which decreases systematic uncertainties in future physics analyses. Monotonicity of the response function in any subset of of the input variables is enforced by constraining the sign of the gradient. This property allows a high efficiency to be maintained for signal candidates with extreme values of the input variables that are a priori known to be interesting for physics analysis, even if these candidates are not well represented in the signal training data.

The MVA is trained with two thousand background candidates and five thousand signal candidates. It will be retrained with HLT2 output data once this becomes available, to both increase the size of the training dataset and to better describe the background distributions. The response is required to increase monotonically in p_{T} and decrease monotonically in all other variables. The distributions of the training variables in signal and background are given in Figure 9.1. The signal response as a function of each variable (evaluated at the median of the signal distribution in all other variables) is also shown in this figure, illustrating the monotonic and smooth variation of the response in each dimension. Figure 9.2 shows the response in two-dimensional planes of the discriminating variables (again evaluated at the median of the signal distribution in all other variables), and the contour corresponding to the requirement of a response greater than 0.2 applied in Section 9.3. The output distribution and ROC curve are shown in Figure 9.3.

Overtraining of the MVA can be assessed using a two-sample Kolmogorov-Smirnov test [225], which tests if two distributions are compatible with originating from the same parent distribution. If two distributions are compatible, p-values for the test are expected to be randomly distributed according to a uniform distribution between zero and one. For this neural network, the p-values are 0.11 and 0.65 for the signal and background datasets, respectively, indicating no evidence of overtraining.

9.3 Beauty to double charm trigger lines

The trigger line selections must reduce the event rate to below 100 Hz, where the event rate

$$R = 1\ \text{MHz} \times \varepsilon_{\text{minbias}} \tag{9.1}$$

is evaluated as the HLT1 output rate, 1 MHz, multiplied by the selection efficiency on HLT1-filtered minimum bias simulation, $\varepsilon_{\text{minbias}}$. Meanwhile, the signal efficiency ε_{sig} should be maximised. ε_{sig} is the yield of selected simulated candidates matching truth criteria divided by the number of simulated signal decays where all final state particles are within LHCb acceptance.

Separate trigger lines are developed for B^+ and B_c^+ meson decays. This allows the lines selecting B_c^+ meson candidates to take advantage of the lower background at higher

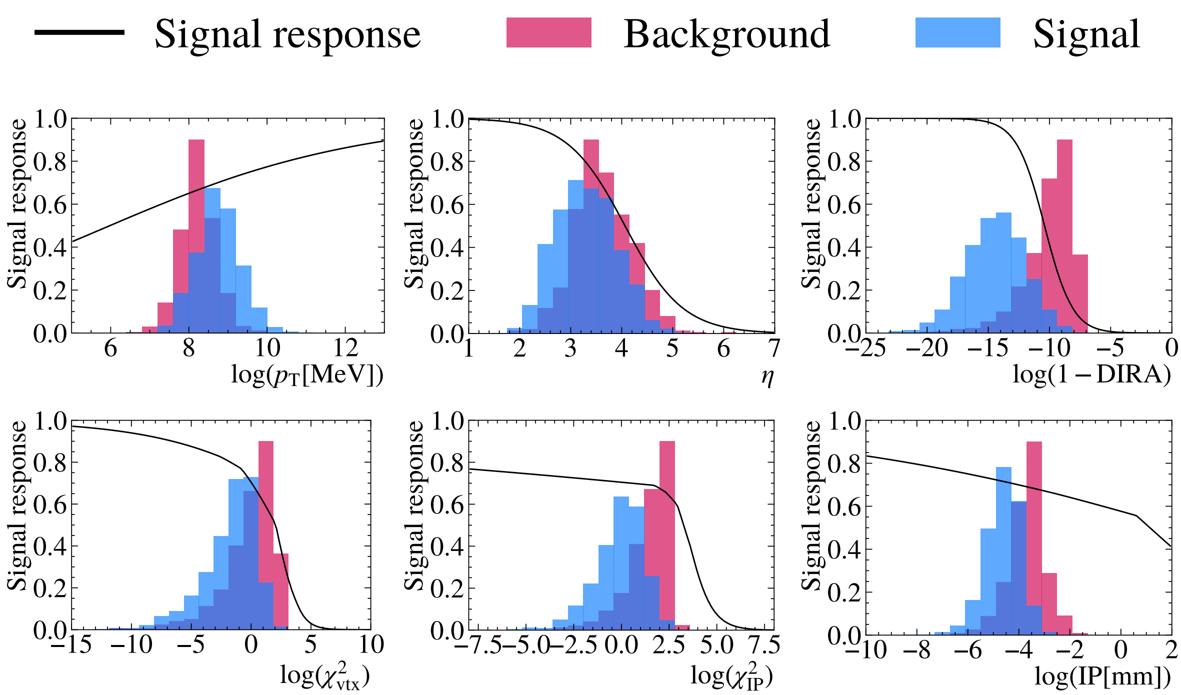

Figure 9.1: Distributions of input variables to the b-hadron MVA in signal and background training samples, and signal response as a function of each parameter. The signal response is higher at high p_{T} and η; at DIRA close to one where the directions of the candidate momentum and displacement are similar; for good-quality vertices; and for small transverse displacements from the PV.

invariant mass to employ looser selections on the displacement of the B_c^+ meson.

Mass windows for the $B_{(c)}^+$ and D mesons are chosen to include the full range that may be useful for physics measurements. The window $m(B^+) \in [4800, 5650]\,\mathrm{MeV}/c^2$ includes $B^+ \to DD$ decays that are fully reconstructed or partially reconstructed with one missing particle; the upper mass sideband for training an MVA in offline selection; and a lower mass range to constrain peaking backgrounds that partially overlap with signal decays. The window $m(B_c^+) \in [5050, 6700]\,\mathrm{MeV}/c^2$ covers fully and partially reconstructed $B_c^+ \to DD$ decays and fully reconstructed $B^+ \to DD$ decays that can be used as normalisation modes. $D_{(s)}^+$ and D^0 candidates are required to have a reconstructed mass within $100\,\mathrm{MeV}/c^2$ of their known mass.

Selection variables are based on those used in the stripping selection for Run 1+2 data. The displacement along the z-axis of the D meson from the $B_{(c)}^+$ meson, $\Delta z(D)$, is also considered. The $\chi^2_{\mathrm{trk}}/\mathrm{ndf}$ and ghost probability are not considered at this stage as these must be trained on data before they are used in selections.

An iterative procedure is established to optimise the thresholds. Firstly, samples of simulated $B_{(c)}^+ \to D_s^+ \overline{D}^0$ and $B_{(c)}^+ \to D^{*+} \overline{D}^0$ decays with $\overline{D}^0 \to K^+\pi^-$, and minimum bias events, are selected using the loose thresholds in Table 9.1. Additionally, standard thresholds that are not studied are applied, requiring a DOCA smaller than 0.5 mm

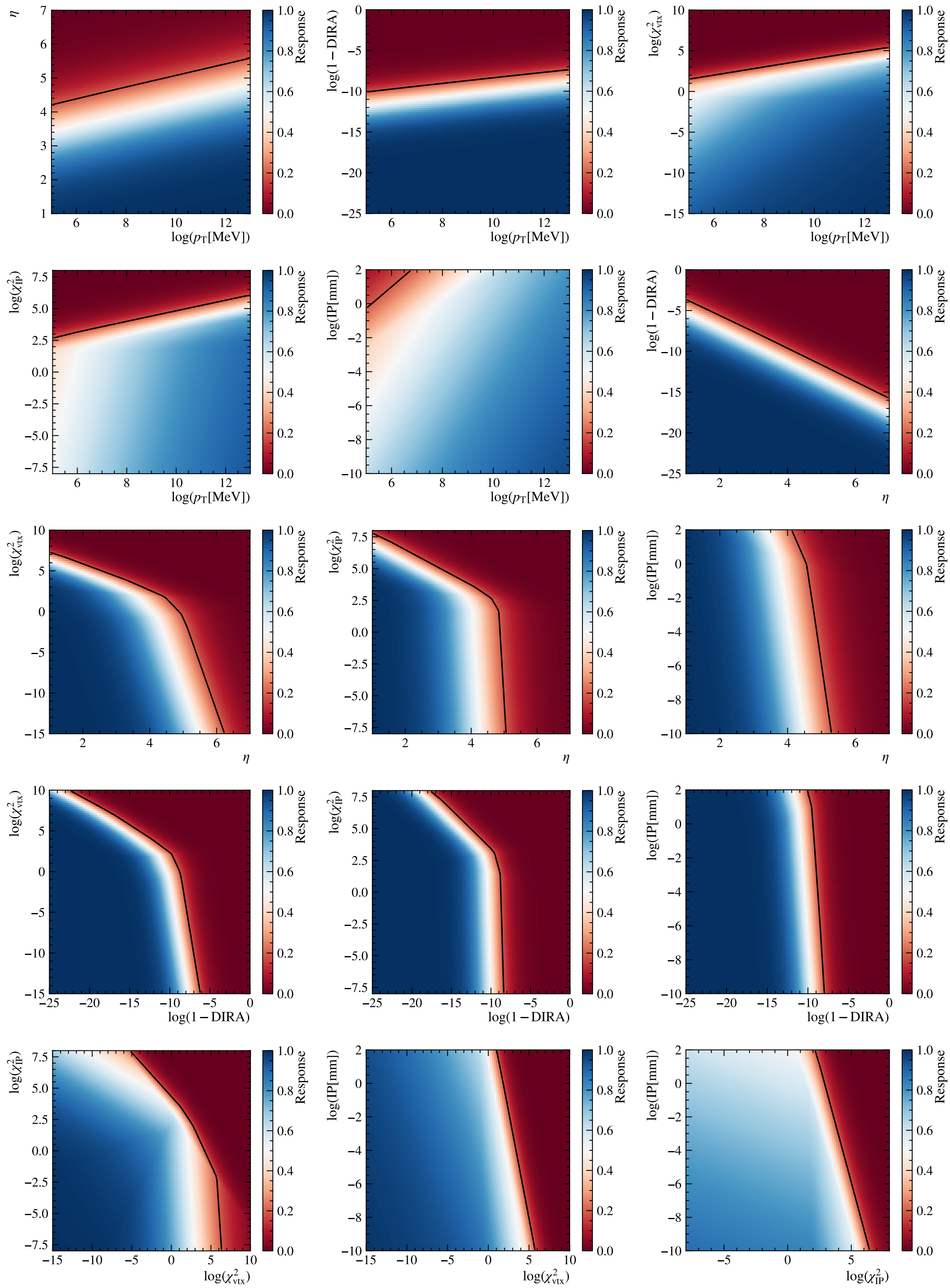

Figure 9.2: Response to the b-hadron MVA in two-dimensional planes of the discriminating variables. The contour corresponding to a response of 0.2 is also drawn (black).

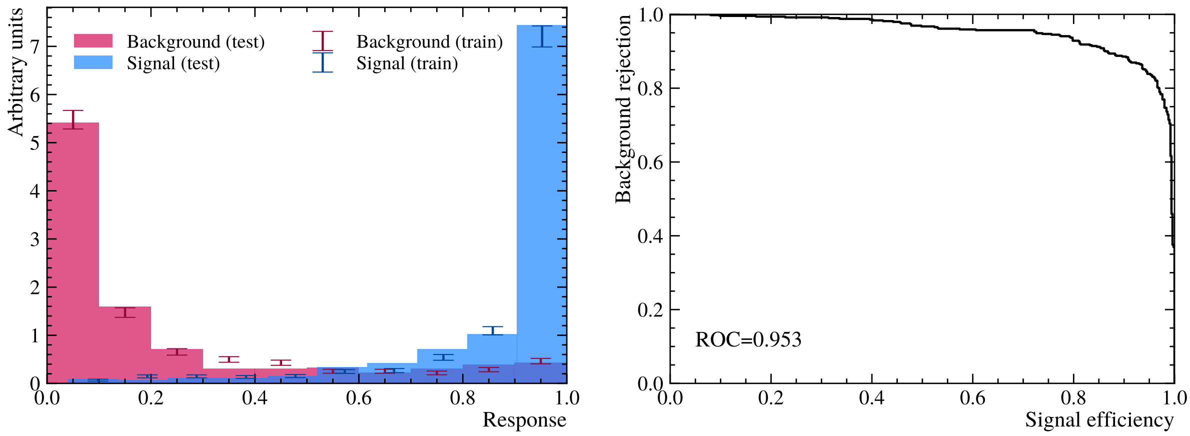

Figure 9.3: (Left) Response of signal and background samples to the b-hadron MVA. (Right) Corresponding background rejection as a function of signal efficiency.

between all tracks in a vertex and between the D^0 and π^+ forming a D^{*+}; $\cos(\theta_B) > 0.999$; $\cos(\theta_{D^+_{(s)},D^0}) > 0$; and $\chi^2_{\text{IP}} > 4$ for all final state tracks.

ε_{sig} and R are plotted as a function of the threshold values in Figures 9.4-9.5. Since this study is performed using only candidates selected with a specific final state ($D_s^+ \to K^+K^-\pi^+$ and $D^0 \to K^-\pi^+$), R is smaller than the true event rate, but still indicates the variation as a function of threshold values. These plots are used to decide tighter selections. This process is repeated until an acceptable rate is achieved.

The final thresholds deployed in HLT2 for data taking in 2022 are given in Table 9.2. This includes the standard selections on the K^0_S meson that are not studied in this thesis. The resulting event rates, including all charm decay channels, are given in Table 9.3. Since different lines may select the same event, the total event rate for all lines, 420 Hz, is lower than the sum of the event rates for each line.

Table 9.4 compares the signal trigger efficiencies to the efficiencies of the comparable selections in Run 2, namely the efficiency of trigger and stripping selections plus the $\text{DLL}_{K\pi}$ and $p_\text{T}(B)$ requirements applied in the preselection (Sections 5.5-5.7). The efficiency is between a factor of 2.3 to 5.9 higher in Run 3, benefitting from the removal of the hardware and topological trigger selections, and the removal of the requirements on $\chi^2_\text{VS}(D^{*+})$ and $\cos(\theta_{D^{*+}})$.

Table 9.1: Loose cuts applied to select simulated candidates used in the optimisation of selections in $B^+_{(c)} \to DD$ decays.

Variable	Requirement
B and event	
$\chi^2_{\mathrm{IP}}(B)$	< 50
$\chi^2_{\mathrm{VS}}(B)$	-
$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}\ (B)$	< 50
$t(B)$	$> 0.05\,\mathrm{ps}$
$\sum_{trk} p_{\mathrm{T}}$	$> 1000\,\mathrm{MeV}/c$
MVA response	-
Each D daughter	
$\Delta m(D^{*+})$	$[90, 400]\,\mathrm{MeV}/c^2$
$\chi^2_{\mathrm{IP}}\ (D^+_{(s)}, D^0)$	-
$\chi^2_{\mathrm{IP}}\ (D^{*+})$	-
$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}\ (D^+_{(s)}, D^0, D^{*+})$	< 30
$\chi^2_{\mathrm{VS}}\ (D^+_{(s)}, D^0)$	-
$\chi^2_{\mathrm{VS}}\ (D^{*+})$	-
$\sum_{trk} p_{\mathrm{T}}$	$> 1800\,\mathrm{MeV}/c$
$\Delta z(D^+_{(s)}, D^0)$	-
Each final state track	
p_{T}	$> 100\,\mathrm{MeV}/c$
p	$> 1\,\mathrm{GeV}/c$
χ^2_{IP}	> 4
$\mathrm{DLL}_{K\pi}(\pi)$	< 20
$\mathrm{DLL}_{K\pi}(K)$	> -20
Soft pions	
$\mathrm{DLL}_{K\pi}(\pi)$	-
p_{T}	-
p	-
χ^2_{IP}	-

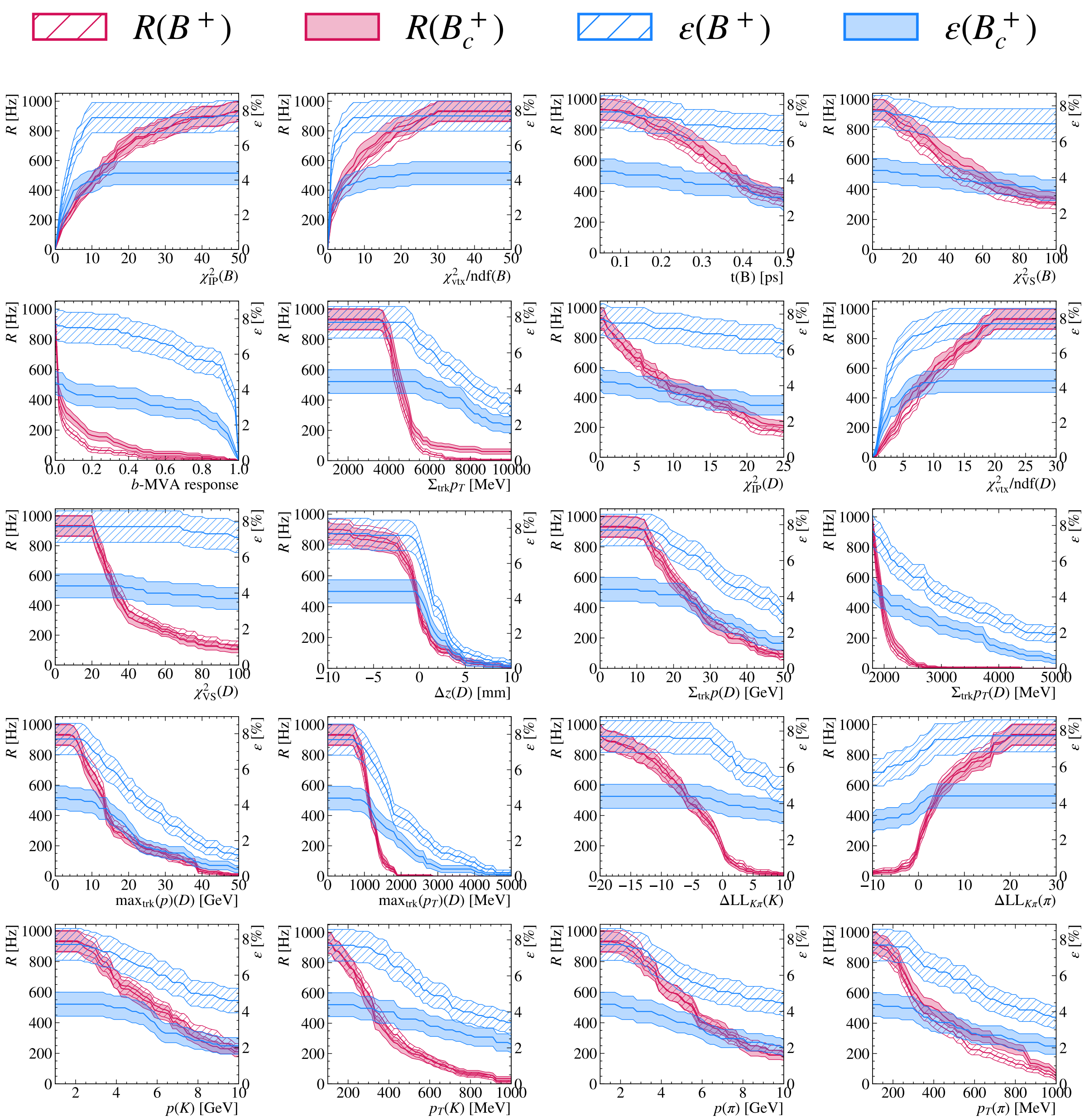

Figure 9.4: Event rate and signal efficiency as a function of selection thresholds in the selection of $B^+_{(c)} \to D^+_s \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ candidates.

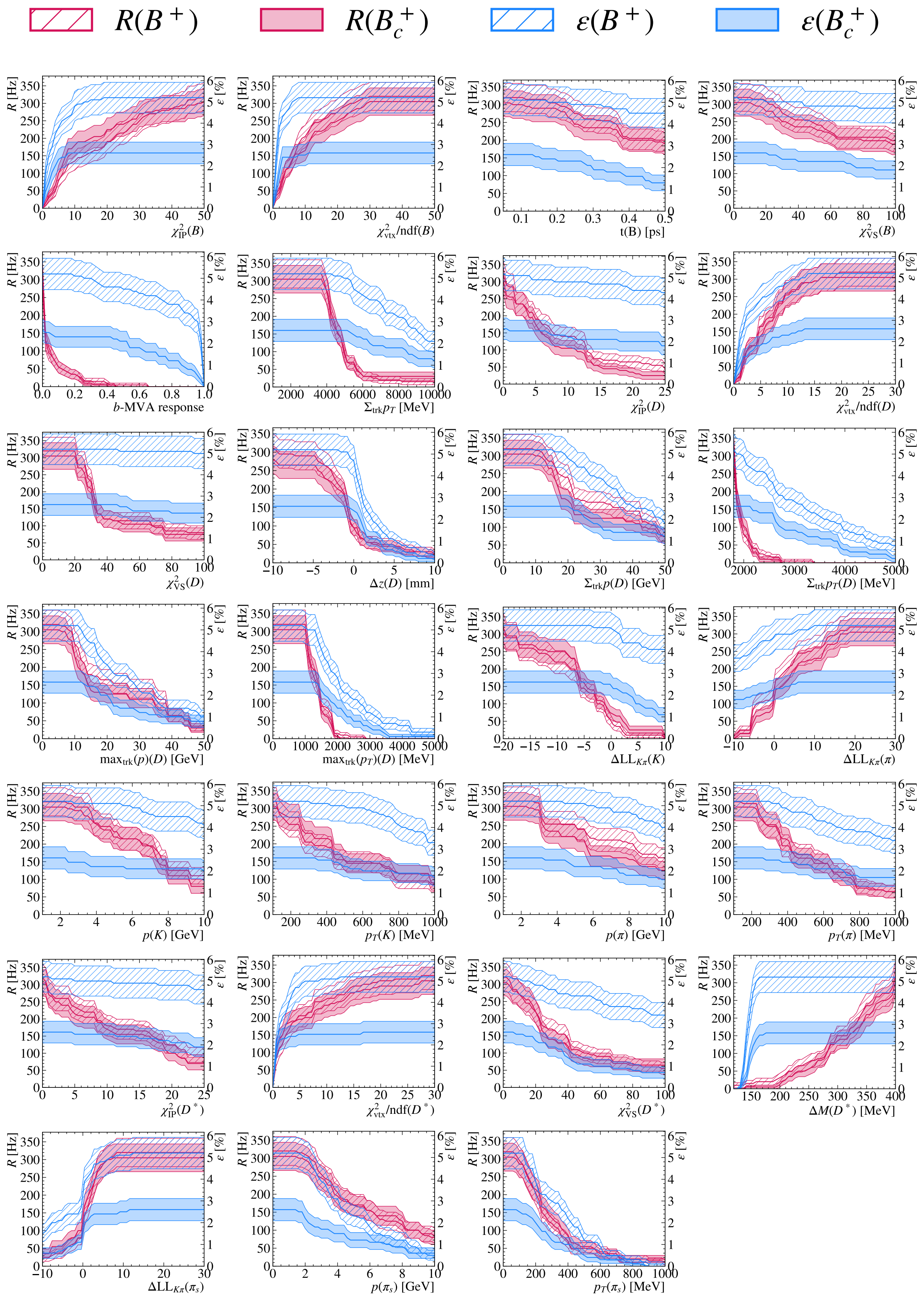

Figure 9.5: Event rate and signal efficiency as a function of selection thresholds in the selection of $B^+_{(c)} \to D^{*+}\overline{D}^0$, $D^0 \to K^-\pi^+$, $\overline{D}^0 \to K^+\pi^-$ candidates.

Table 9.2: Thresholds applied in HLT2 lines for selecting $B^+_{(c)} \to DD$ decays.

Variable	Requirement
B and event	
$\chi^2_{\text{IP}}(B)$	< 20
$\chi^2_{\text{VS}}(B)$	-
$\chi^2_{\text{vtx}}/\text{ndf}\ (B)$	< 20
$t(B)$	$> 0.05\,\text{ps}$
$\cos(\theta_B)$	> 0.999
$\sum_{trk} p_{\text{T}}$	$> 5000(5500)\,\text{MeV}/c$
MVA response	> 0.2
Each *D* daughter	
$\Delta m(D^{*+})$	$[90, 180]\,\text{MeV}/c^2$
$\chi^2_{\text{IP}}\ (D^+_{(s)}, D^0, D^{*+})$	-
$\chi^2_{\text{vtx}}/\text{ndf}\ (D^+_{(s)}, D^0, D^{*+})$	< 10
$\chi^2_{\text{VS}}\ (D^+_{(s)}, D^0)$	$> 64(49)$
$\chi^2_{\text{VS}}\ (D^{*+})$	-
$\sum_{trk} p_{\text{T}}$	$> 1800\,\text{MeV}/c$
$\Delta z(D^+_{(s)}, D^0)$	$> -1\,\text{mm}$
Each final state track	
p_{T}	$> 250\,\text{MeV}/c$
p	$> 2\,\text{GeV}/c$
$\text{DLL}_{K\pi}(\pi)$	< 5
$\text{DLL}_{K\pi}(K)$	> -5
Soft pions	
$\text{DLL}_{K\pi}(\pi)$	< 20
p_{T}	-
p	-
χ^2_{IP}	-
$K^0_{\text{S}\,[\text{DD}]}$	
χ^2_{DOCA}	< 25
$p(\pi)$	$> 2\,\text{GeV}/c$
$\chi^2_{\text{IP}}(\pi)$	> 4
$\chi^2_{\text{vtx}}/\text{ndf}$	< 25
$K^0_{\text{S}\,[\text{LL}]}$	
χ^2_{DOCA}	< 30
$p(\pi)$	$> 2\,\text{GeV}/c$
$\chi^2_{\text{IP}}(\pi)$	> 9
$\chi^2_{\text{vtx}}/\text{ndf}$	< 30
χ^2_{VS}	> 4

Table 9.3: Event selection rates of HLT2 lines. h refers to a pion or kaon.

Channel	Final state	$R(B^+)$/Hz	$R(B_c^+)$/Hz
$D_s^+\overline{D}^0$	$\overline{D}^0 \to h^+h^-$	40 ± 20	30 ± 20
	$\overline{D}^0 \to 2h^+2h^-$	70 ± 30	60 ± 20
	$\overline{D}^0 \to K^0_{S[DD]}h^+h^-$	–	0 ± 10
	$\overline{D}^0 \to K^0_{S[LL]}h^+h^-$	–	20 ± 10
$D^+\overline{D}^0$	$\overline{D}^0 \to h^+h^-$	70 ± 30	60 ± 20
	$\overline{D}^0 \to 2h^+2h^-$	70 ± 30	100 ± 30
	$\overline{D}^0 \to K^0_{S[DD]}h^+h^-$	–	0 ± 10
	$\overline{D}^0 \to K^0_{S[LL]}h^+h^-$	–	10 ± 10
$D^{*+}\overline{D}^0$ $D^{*+} \to D^0\pi^+$	$D^0 \to h^+h^-$, $\overline{D}^0 \to h^+h^-$	10 ± 10	10 ± 10
	$D^0 \to h^+h^-$, $\overline{D}^0 \to 2h^+2h^-$	40 ± 20	80 ± 30
	$D^0 \to h^+h^-$, $\overline{D}^0 \to K^0_{S[DD]}h^+h^-$	–	0 ± 0
	$D^0 \to h^+h^-$, $\overline{D}^0 \to K^0_{S[LL]}h^+h^-$	–	10 ± 10
	$D^0 \to 2h^+2h^-$, $\overline{D}^0 \to h^+h^-$	40 ± 20	40 ± 20
	$D^0 \to 2h^+2h^-$, $\overline{D}^0 \to 2h^+2h^-$	10 ± 10	50 ± 20
	$D^0 \to 2h^+2h^-$, $\overline{D}^0 \to K^0_{S[DD]}h^+h^-$	–	0 ± 10
	$D^0 \to 2h^+2h^-$, $\overline{D}^0 \to K^0_{S[LL]}h^+h^-$	–	0 ± 10

Table 9.4: Simulated signal efficiencies in percent of (Run 3) trigger selections and (Run 2) trigger and stripping selections, $\mathrm{DLL}_{K\pi}(K) > -5$, $\mathrm{DLL}_{K\pi}(\pi) < 10$, and $p_\mathrm{T}(B) > 4\,\mathrm{GeV}/c$. Uncertainties are statistical only and simulation has not been corrected for mismodelling.

Decay	Run 3	Run 2
$B^+ \to D_s^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	5.3 ± 0.2	2.27 ± 0.02
$B_c^+ \to D_s^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	3.3 ± 0.1	1.15 ± 0.01
$B^+ \to D^{*+}\overline{D}^0$, $D^0 \to K^-\pi^+$, $\overline{D}^0 \to K^+\pi^-$	5.2 ± 0.2	1.52 ± 0.01
$B_c^+ \to D^{*+}\overline{D}^0$, $D^0 \to K^-\pi^+$, $\overline{D}^0 \to K^+\pi^-$	3.0 ± 0.1	0.51 ± 0.01

9.4 Experimental prospects in Run 3 and beyond

In this chapter, a b-hadron neural network and exclusive HLT2 trigger lines have been developed to select beauty hadron to open charm decays in Run 3. The gain in signal efficiency as a result of the redesign of the trigger system has been demonstrated. In this section, the prospects for doubly charmed beauty meson decays in Run 3 and beyond, bolstered by the improved efficiency from the upgraded trigger system, are assessed.

The expected selected yield of B_c^+ meson decays is estimated by scaling the selected yield of normalisation candidates in Run 2 (R2) data, $N(B^+, \mathrm{R2})$, as

$$N(B_c^+) = \frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to DD)}{\mathcal{B}(B^+ \to DD)}\frac{\varepsilon(B_c^+)}{\varepsilon(B^+, \mathrm{R2})}\frac{\int \mathcal{L} dt}{\int_{\mathrm{R2}} \mathcal{L}_{\mathrm{R2}} dt} N(B^+, \mathrm{R2}). \tag{9.2}$$

The range of signal yields that could be selected by the end of each data taking period, assuming the predictions on the branching fractions in Table 2.1 and the integrated luminosities in Table 3.1, are displayed in Table 9.5. The quoted yields are cumulative, so include the signal yield from the given and all preceeding datasets. The efficiency ratio is taken from Table 9.4 and assumed to remain constant from Run 3 onwards. Variations in the production cross-section between 13 and 14 TeV are neglected. The quoted range includes uncertainty from the range of the branching fraction predictions, f_c/f_u, the measured $B^+ \to DD$ branching fractions, and the efficiency increase. The uncertainty is dominated by the range of predictions of $B_c^+ \to DD$ branching fractions.

Table 9.5 also shows projections assuming the much larger value of $\mathcal{B}(B_c^+ \to D_s^+ \overline{D}{}^0)$ measured in Chapter 7. These yields are evaluated by scaling the fitted yield in Run 1+2 data for the integrated luminosity, production cross-section, and efficiency. The statistical uncertainty of the Run 1+2 yield and the efficiencies are included.

If the branching fraction measured in Chapter 7 is of the correct order of magnitude,

Table 9.5: Range of expected selected yields for $B_c^+ \to DD$ decay modes by the end of each operational run. For the measured $\mathcal{B}(B_c^+ \to D_s^+ \overline{D}{}^0) = (3.5 \pm 1.7) \times 10^{-4}$ and for the branching fractions predicted in Table 2.1.

Decay	Run 3	Run 4	Run 5	Run 6
$\mathcal{B}(B_c^+ \to D_s^+ \overline{D}{}^0) = (3.5 \pm 1.7) \times 10^{-4}$				
$B_c^+ \to D_s^+ \overline{D}{}^0$	[79, 189]	[207, 497]	[797, 1926]	[1506, 3640]
Predicted $\mathcal{B}$				
$B_c^+ \to D_s^+ \overline{D}{}^0$	[0.1, 2.2]	[0.2, 5.9]	[1, 23]	[1, 43]
$B_c^+ \to D_s^+ D^0$	[0.1, 3.5]	[0.2, 9.1]	[1, 35]	[1, 66]
$B_c^+ \to D^+ \overline{D}{}^0$	[0, 19]	[1, 49]	[3, 189]	[5, 356]
$B_c^+ \to D^+ D^0$	[0.00, 0.11]	[0.00, 0.29]	[0.0, 1.1]	[0.0, 2.2]

Table 9.6: Expected total yields of selected $B^- \to DD$ decays by the end of each operational run.

Yield	Run 3	Run 4	Run 5	Run 6
$N(B^- \to D_s^- D^0)/10^6$	1.29 ± 0.04	3.25 ± 0.12	12.3 ± 0.5	23.2 ± 0.9
$N(B^- \to D^- D^0)/10^4$	6.49 ± 0.21	16.3 ± 0.6	61.7 ± 2.4	116 ± 5

Table 9.7: Expected statistical uncertainty on $\mathcal{A}^{\text{raw}}$ in percent by the end of each operational run.

Decay	Run 3	Run 4	Run 5	Run 6
$B^- \to D_s^- D^0$	0.11	0.07	0.04	0.03
$B^- \to D^- D^0$	0.4	0.3	0.1	0.1

an observation of the $B_c^+ \to D_s^+ \overline{D}^0$ decay in the Run 3 dataset should be achievable. Conversely, if the branching fractions are in accordance with the most pessimistic predictions in Ref. [76], it is unlikely that any $B_c^+ \to DD$ decays can be observed at the LHC. However, predictions from the majority of sources [72–75] suggest good prospects for observing several $B_c^+ \to DD$ decays within the lifetime of the LHC: The $B_c^+ \to D^+ \overline{D}^0$ decay may be observed as early as Run 3 if the true branching fraction is in accordance with optimistic predictions [73]. Observation of the $B_c^+ \to D_s^+ \overset{(-)}{D}{}^0$ decays is also likely, although not until Run 5 according to the branching fraction predictions [72–75].

Selected yields for $B^- \to DD$ decays are predicted by scaling the fitted signal yields from Chapter 8 for increased integrated luminosities and increased efficiencies. The corresponding statistical uncertainties on the $\mathcal{A}^{\text{raw}}$ are also estimated by assuming that they scale as $1/\sqrt{N}$. These are displayed, again cumulatively, in Tables 9.6-9.7.

An integrated luminosity of $14\,\text{fb}^{-1}$ collected in Run 3 would provide more than a five-fold increase in signal yield. The substantial increase in instantaneous luminosity again at the start of Run 5 will allow the signal yield to grow to one hundred times its current value assuming $250\,\text{fb}^{-1}$ collected during Run 5+6. Assuming that uncertainty on the production and detection asymmetries will also scale with the size of the calibration datasets, these yield increases would allow measurements of $\mathcal{A}^{CP}$ with more than double the precision at the end of Run 3 and ten times improved precision at the end of Run 6.

Given the values of the SM predictions on $\mathcal{A}^{CP}$, measuring evidence for or observing CP violation in one or more decay modes is probable by the end of Run 3. By the end of LHC running, discrepancies from the SM of $\mathcal{O}(10^{-3(4)})$ in the CS (CF) decay modes will be measurable, which would permit the identification of deviations at this level due to CP-violating NP.

Summary and outlook

10.1 Summary

The SM has enjoyed a high degree of success in predicting the properties and behaviour of particles to date. However, some phenomena remain unexplained: for example, CPV measured at the quark level is insufficient to explain the observed baryon asymmetry of the universe. Moreover, satisfactory methods to make predictions in the non-perturbative regimes of the SM, for example in hadronic bound states, have not yet been established.

This thesis advances understanding of these problems through the study of double charm decays of charged beauty mesons using data collected by the LHCb detector in pp collisions at the LHC between 2011 and 2018. Firstly, a search was performed for sixteen double charm decays of the B_c^+ meson. This search provided evidence for the $B_c^+ \to D_s^+ \overline{D}^0$ decay with a significance of 3.4 standard deviations, corresponding to a branching fraction of

$$\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0) = (3.5^{+1.5+0.3}_{-1.3-0.2} \pm 1.0) \times 10^{-4}. \tag{10.1}$$

This branching fraction is two orders of magnitude higher than predicted with the SM. If this value is confirmed using data collected by the LHCb detector in future data taking periods, it would challenge understanding of the behaviour of the complex B_c^+ meson within the SM, or potentially the validity of the SM itself. Upper limits at 90(95)% C.L. on the branching fraction of all sixteen decays were established as

$$\begin{aligned}
\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0) &< 7.2\,(8.4) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D_s^+ D^0) &< 3.0\,(3.7) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D^+ \overline{D}^0) &< 1.9\,(2.5) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D^+ D^0) &< 1.4\,(1.8) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D_s^{*+} \overline{D}^0) &< 5.3\,(5.7) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^{*0}) &< 4.6\,(5.6) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D_s^{*+} D^0) &< 0.9\,(1.0) \times 10^{-3}, \\
\mathcal{B}(B_c^+ \to D_s^+ D^{*0}) &< 6.6\,(8.4) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D^{*+} \overline{D}^0) &< 3.8\,(4.8) \times 10^{-4}, \\
\mathcal{B}(B_c^+ \to D^{*+} D^0) &< 2.0\,(2.4) \times 10^{-4},
\end{aligned}$$

$$\mathcal{B}(B_c^+ \to D^+ \overline{D}^{*0}) < 6.5\,(8.2) \times 10^{-4},$$
$$\mathcal{B}(B_c^+ \to D^+ D^{*0}) < 3.7\,(4.6) \times 10^{-4},$$
$$\mathcal{B}(B_c^+ \to D_s^{*+} \overline{D}^{*0}) < 1.3\,(1.5) \times 10^{-3},$$
$$\mathcal{B}(B_c^+ \to D_s^{*+} D^{*0}) < 1.3\,(1.6) \times 10^{-3},$$
$$\mathcal{B}(B_c^+ \to D^{*+} \overline{D}^{*0}) < 1.0\,(1.3) \times 10^{-3},$$
$$\mathcal{B}(B_c^+ \to D^{*+} D^{*0}) < 7.7\,(8.9) \times 10^{-4}.$$

Secondly, CP asymmetries of decays of the B^- meson to two charm mesons were measured to be

$$\mathcal{A}^{CP}(B^- \to D_s^- D^0) = (+0.5 \pm 0.2 \pm 0.5 \pm 0.3)\%,$$
$$\mathcal{A}^{CP}(B^- \to D_s^{*-} D^0) = (-0.5 \pm 1.1 \pm 1.0 \pm 0.3)\%,$$
$$\mathcal{A}^{CP}(B^- \to D_s^- D^{*0}) = (+1.1 \pm 0.8 \pm 0.6 \pm 0.3)\%,$$
$$\mathcal{A}^{CP}(B^- \to D^- D^0) = (+2.5 \pm 1.0 \pm 0.4 \pm 0.3)\%,$$
$$\mathcal{A}^{CP}(B^- \to D^{*-} D^0) = (+3.3 \pm 1.6 \pm 0.6 \pm 0.3)\%,$$
$$\mathcal{A}^{CP}(B^- \to D^- D^{*0}) = (-0.2 \pm 2.0 \pm 1.4 \pm 0.3)\%,$$
$$\mathcal{A}^{CP}(B^- \to D^{*-} D^{*0}) = (+2.3 \pm 2.1 \pm 1.7 \pm 0.3)\%,$$

where the first uncertainty is statistical, the second is systematic, and the third is from $\mathcal{A}^{CP}(B^+ \to J/\psi K^+)$. The branching fraction ratios

$$\frac{\mathcal{B}(B^- \to D^- D^0)}{\mathcal{B}(B^- \to D_s^- D^0)} \frac{\mathcal{B}(D^- \to K^+ \pi^- \pi^-)}{\mathcal{B}(D_s^- \to K^+ K^- \pi^-)} = (7.25 \pm 0.09 \pm 0.09) \times 10^{-2}$$

and

$$\frac{\mathcal{B}(B^- \to D^{*-} D^0)}{\mathcal{B}(B^- \to D^- D^0)} \frac{\mathcal{B}(D^{*-} \to \overline{D}^0 \pi^-)\mathcal{B}(\overline{D}^0 \to K^+ \pi^-)}{\mathcal{B}(D^- \to K^+ \pi^- \pi^-)} = 0.271 \pm 0.07 \pm 0.05$$

were also measured, where the first uncertainty is statistical, and the second is systematic. These are the most precise measurements of CP asymmetries and branching fraction ratios for double charm decays of the B^- meson, and include the first measurements of $\mathcal{A}^{CP}(B^- \to D_s^{*-} D^0)$ and $\mathcal{A}^{CP}(B^- \to D_s^- D^{*0})$. Although no evidence of CP violation is seen, these results will help to constrain CP-violating NP.

10.2 Outlook

Following the completion of Run 1 and Run 2 data taking at the LHC, four further operational data taking runs are forseen over the next twenty years with higher luminosities, higher centre-of-mass energies, and upgrades to the LHCb detector. Run 3 data taking commenced in 2022 with a target instantaneous luminosity at the LHCb interaction point five times higher than in Run 2, and the LHCb detector has been upgraded to accommodate the associated increase in pile-up. Furthermore, a redesign of the trigger system, including the removal of the hardware trigger and the wide-scale implementation of exclusive selections in the final level of the software trigger, has permitted a three-fold increase in efficiency for beauty meson to double charm decays. The development of exclusive selections for these decays has been presented in Chapter 9.

A further ten-fold increase in instantaneous luminosity at the LHCb interaction point is planned for Run 5 data taking, commensurate with further upgrades to the LHCb detector (Upgrade II [100]) to accomodate yet higher pile-up. The yields of beauty meson to double charm decays available for physics analysis by the end of each operational run have been predicted in Section 9.4, indicating an increase of two orders of magnitude by the end of LHC running.

If the value of $\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0)$ measured in Chapter 7 is the correct order of magnitude, an observation of this decay is likely with the data that will be collected in Run 3. This value is two orders of magnitude larger than predictions based on the SM, so an observation would challenge understanding of the behaviour of the B_c^+ meson in the SM, or understanding of the SM itself. To improve understanding of the reason for this discrepancy, it would be essential to measure the branching fraction more precisely in future datasets. The observation of $B_c^+ \to D_s^{(*)+} \overline{D}^{(*)0}$ decays would also be likely, and measurements of the corresponding branching fractions as well as $f_L(B_c^+ \to D_s^{*+} \overline{D}^{*0})$ would provide further phenomenological insight.

The evidence for the $B_c^+ \to D_s^+ \overline{D}^0$ decay suggests that the branching fractions of other doubly charmed B_c^+ meson decays may also differ from expectations, and therefore may be observed sooner than anticipated. Even if the SM predictions are in fact accurate, an observation of the $B_c^+ \to D^+ \overline{D}^0$ decay is possible in Run 3 or 4, and the $B_c^+ \to D_s^+ \overset{(-)}{D}{}^0$ decays in Run 5 or 6.

The feasability of measuring the CKM phase γ in doubly charmed B_c^+ meson decays depends entirely on their true phenomenology: The dominant diagrams for the $B_c^+ \to D_s^+ \overline{D}^0$ and $B_c^+ \to D_s^+ D^0$ decays must have similar amplitudes, and one must contain a $b \to u$ transition and the other a $b \to c$ transition. It seems unlikely that these conditions would still hold if the true value of $\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0)$ is two orders of magnitude larger than expectations, but a full phenomenological understanding is required to draw definitive conclusions.

Conversely, if the SM predictions are accurate, only small yields would be required to measure γ due to the large degree of interference. An uncertainty of $\sigma_\gamma \sim 18^\circ$ would be expected using data corresponding to 300 fb^{-1} of integrated luminosity [192]. Although this is much lower precision than the expectation of $\sigma_\gamma \sim 0.35^\circ$ using all other decay modes with the same dataset [102], a comparison of γ measured using different beauty mesons would provide complementary sensitivity to NP since different BSM diagrams and couplings may contribute to the decays of different beauty mesons [69].

On the side of B^- meson decays to double charm final states, prospects for measuring CPV as early as Run 3 are promising. Furthermore, by the end of LHC running, precision of 0.03% and 0.1% is anticipated on $\mathcal{A}^{CP}$ $(B^- \to D_s^- D^0)$ and $\mathcal{A}^{CP}$ $(B^- \to D^- D^0)$, respectively. To maximise sensitivity to NP, it is essential that measurements of the isospin partner decays are also updated. These measurements will benefit similarly from the improvements in trigger efficiency. In future data taking periods, measurements of *CP* observables for the suppressed annihilation decays of $B^0_{(s)}$ mesons will become possible, further enhancing sensitivity to NP. Overall, with the collection of fourteen times more signal decays possible with the LHCb Upgrade I detector, and a hundred times more with the LHCb Upgrade II detector, the prospects for probing *CP*-violating NP in doubly charmed beauty meson decays with future LHCb datasets are very encouraging.

Comparison of simulation and data

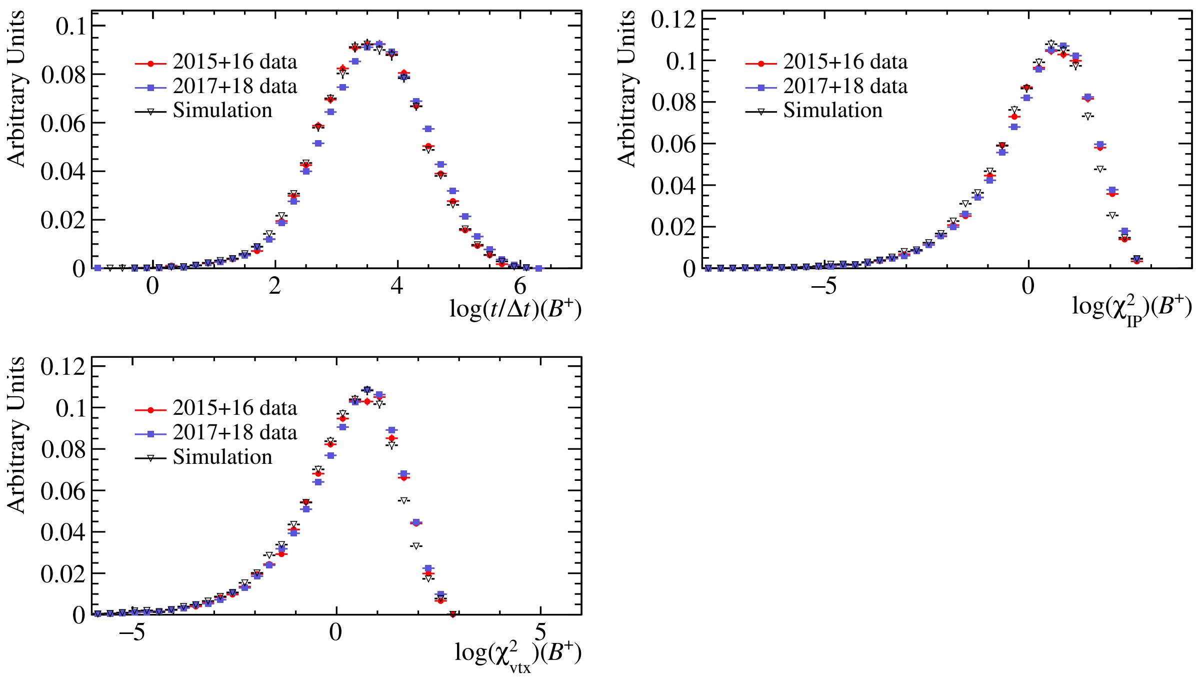

Figure A.1: Comparison of variables related to the reconstructed B^+ between data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \overline{D}^0$, $D^0 \to K^- \pi^+$ decays.

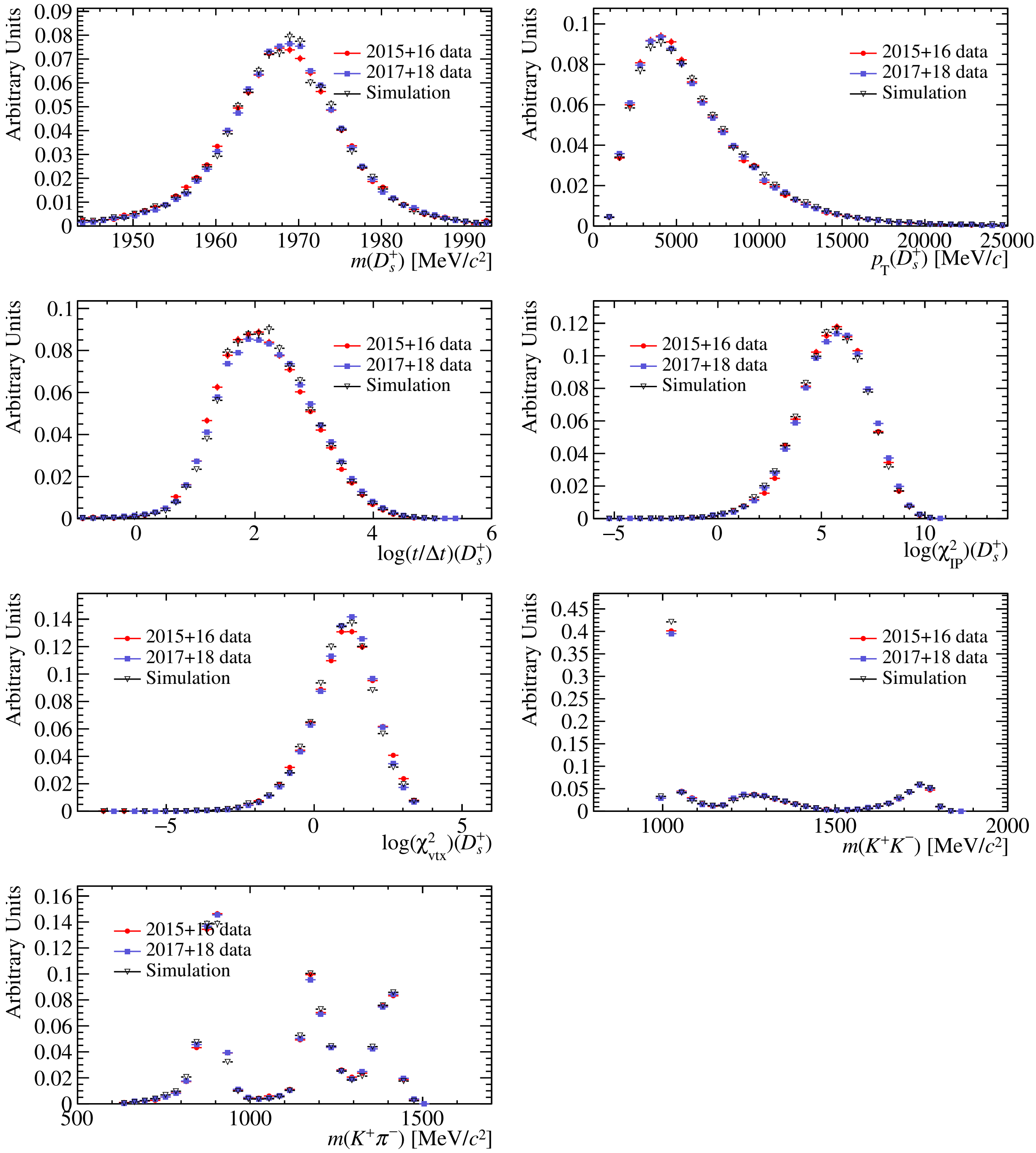

Figure A.2: Comparison of variables related to the reconstructed D_s^+ between data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \bar{D}^0$, $D^0 \to K^- \pi^+$ decays.

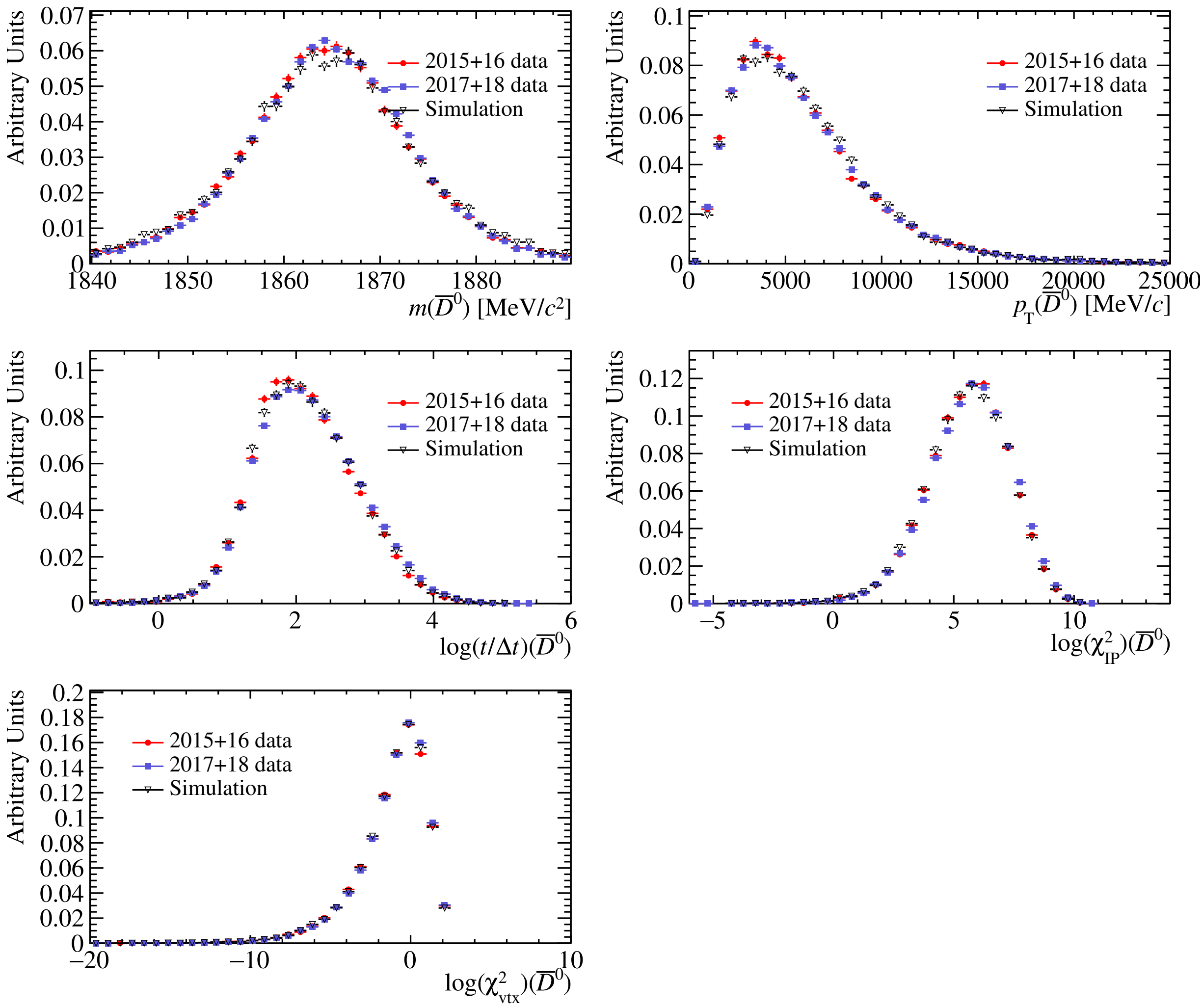

Figure A.3: Comparison of variables related to the reconstructed D^0 between data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \overline{D}^0$, $D^0 \to K^- \pi^+$ decays.

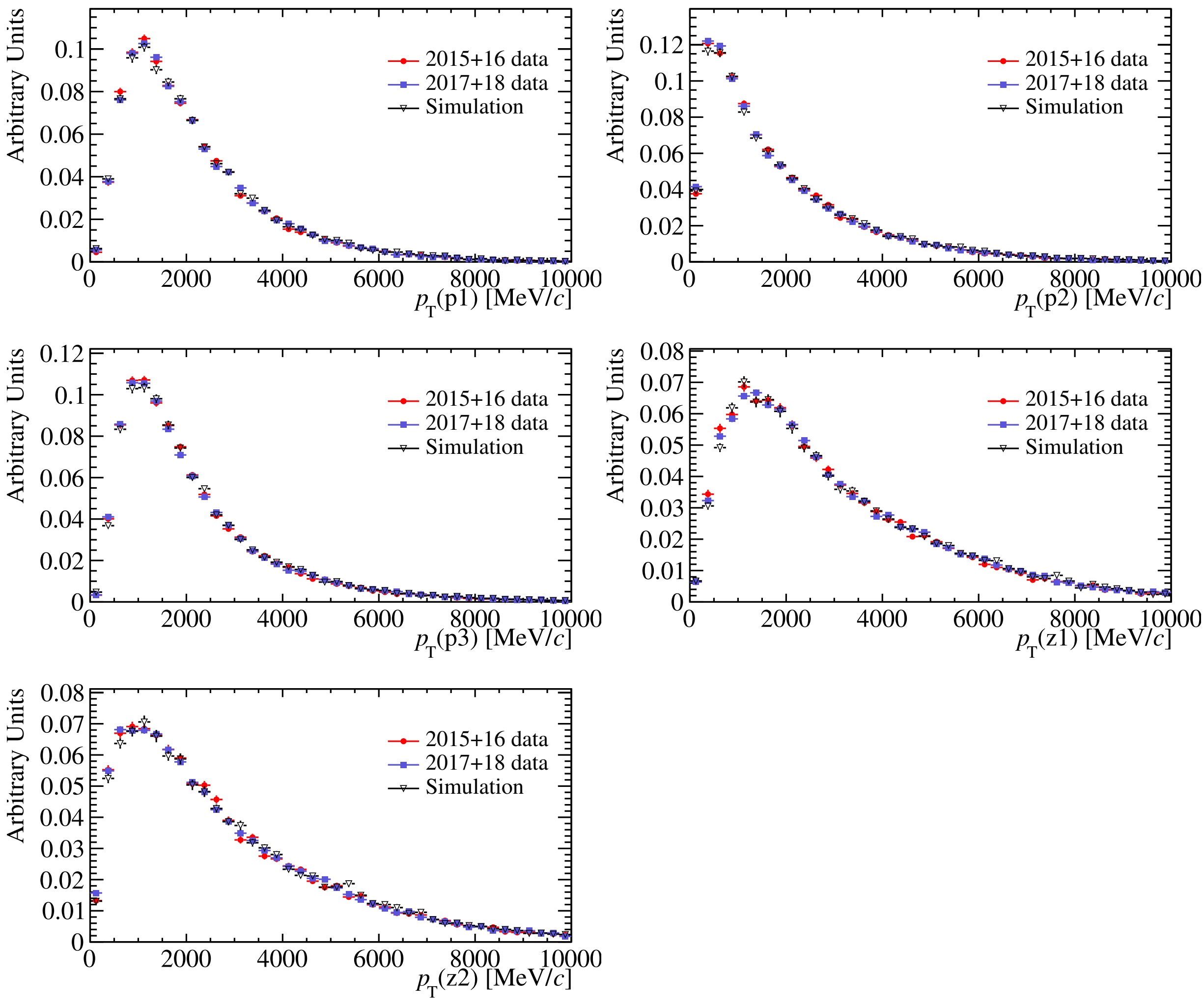

Figure A.4: Comparison of final state particle p_T distributions between data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \bar{D}^0$, $D^0 \to K^- \pi^+$ decays. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D^+, respectively. The labels z1 and z2 stand for the K^- and π^+ daughters of D^0, respectively.

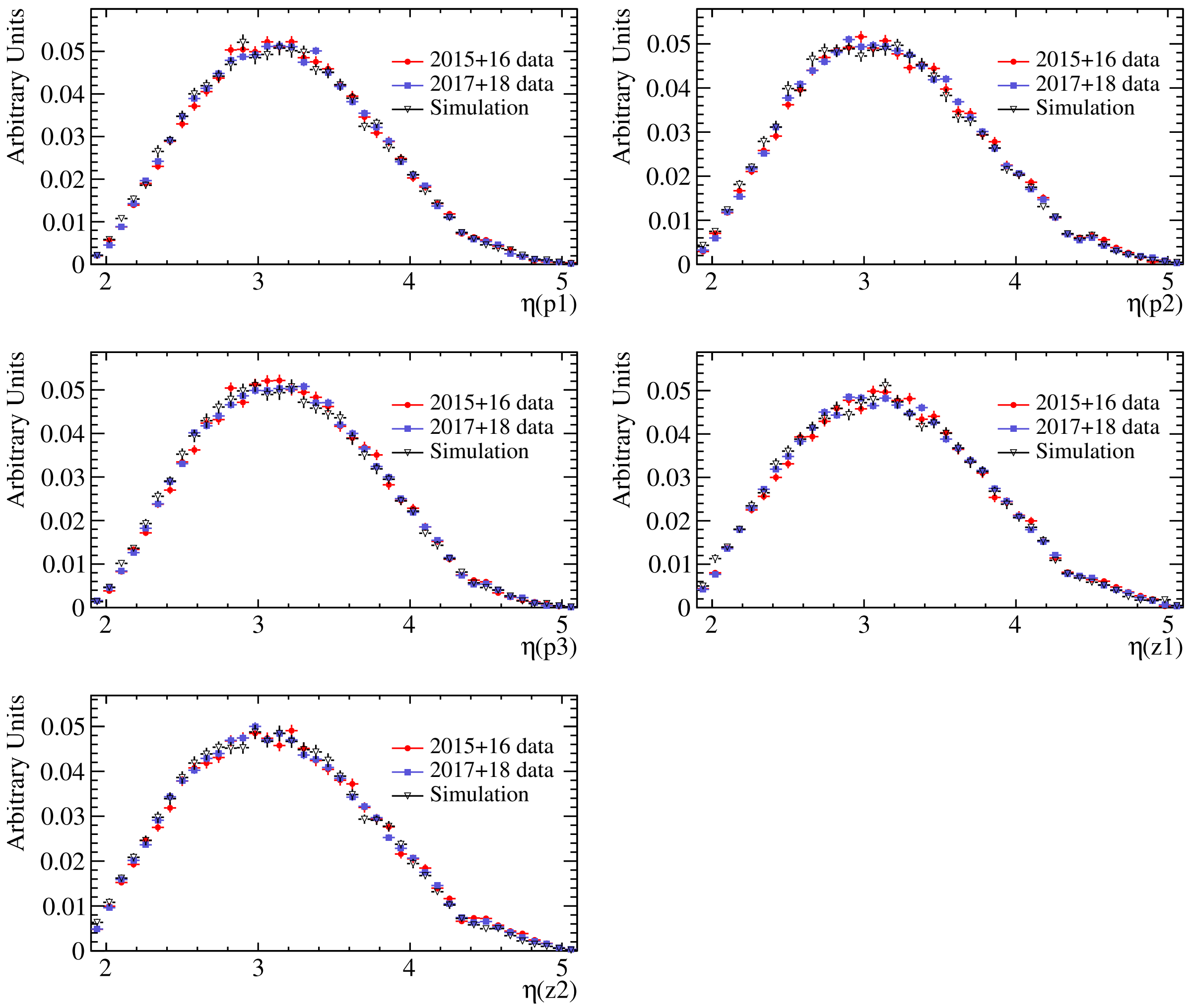

Figure A.5: Comparison of final state particle η distributions between data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \bar{D}^0$, $D^0 \to K^- \pi^+$ decays. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D^+, respectively. The labels z1 and z2 stand for the K^- and π^+ daughters of D^0, respectively.

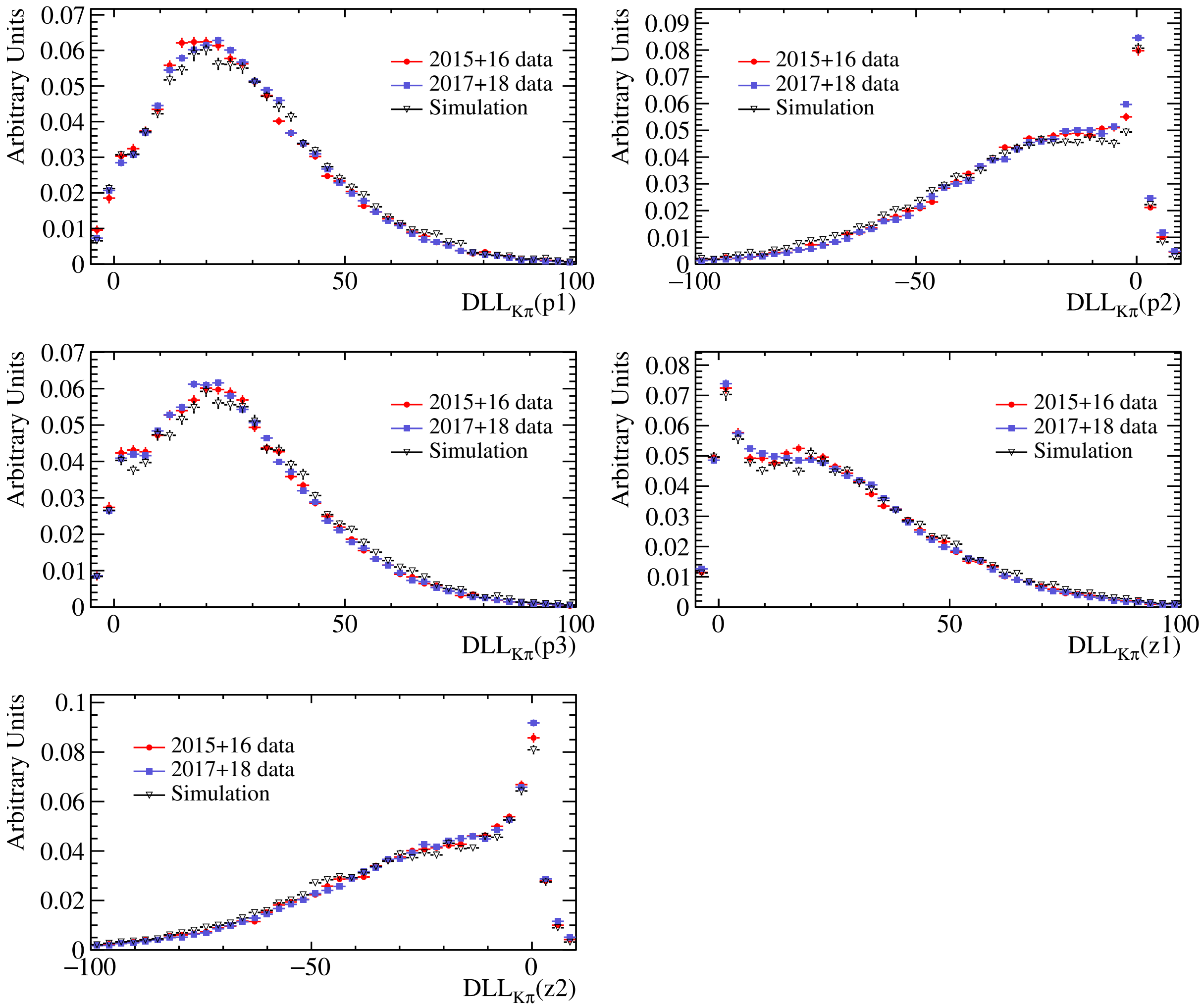

Figure A.6: Comparison of corrected PID distributions between data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \bar{D}^0$, $D^0 \to K^- \pi^+$ decays. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D^+, respectively. The labels z1 and z2 stand for the K^- and π^+ daughters of D^0, respectively.

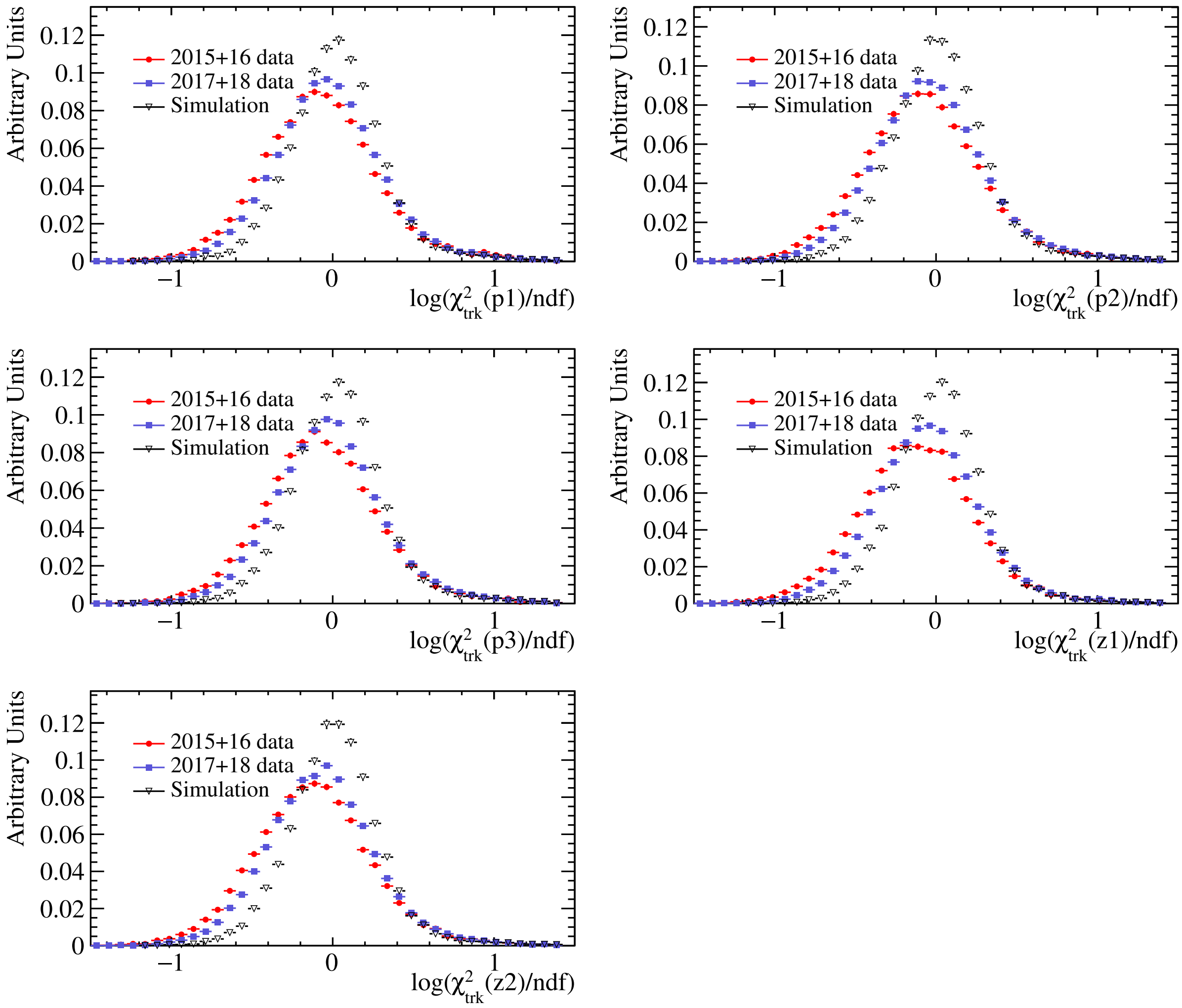

Figure A.7: Comparison of $\log(\chi^2_{\mathrm{trk}}/\mathrm{ndf})$ distributions between data in 2015+16 (red), 2017+18 (blue), and simulation (black) for $B^+ \to D_s^+ \overline{D}^0$, $D^0 \to K^- \pi^+$ decays. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D^+, respectively. The labels z1 and z2 stand for the K^- and π^+ daughters of D^0, respectively.

Input variables to the BDT

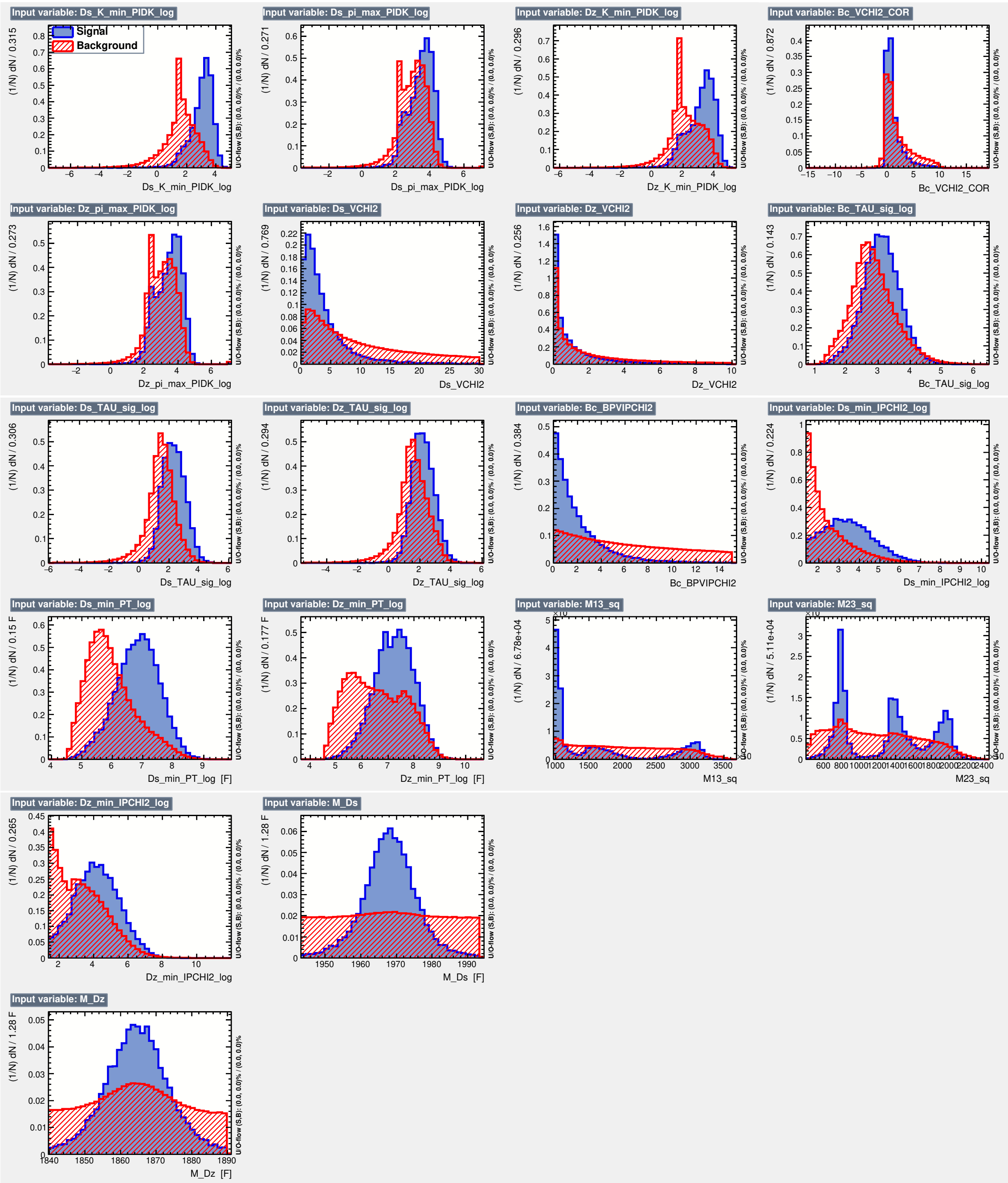

Figure B.1: Distributions in (blue) signal and (red) background of the input variables to the MVA for $B_c^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ candidates in Run 2 data.

Likelihood scans

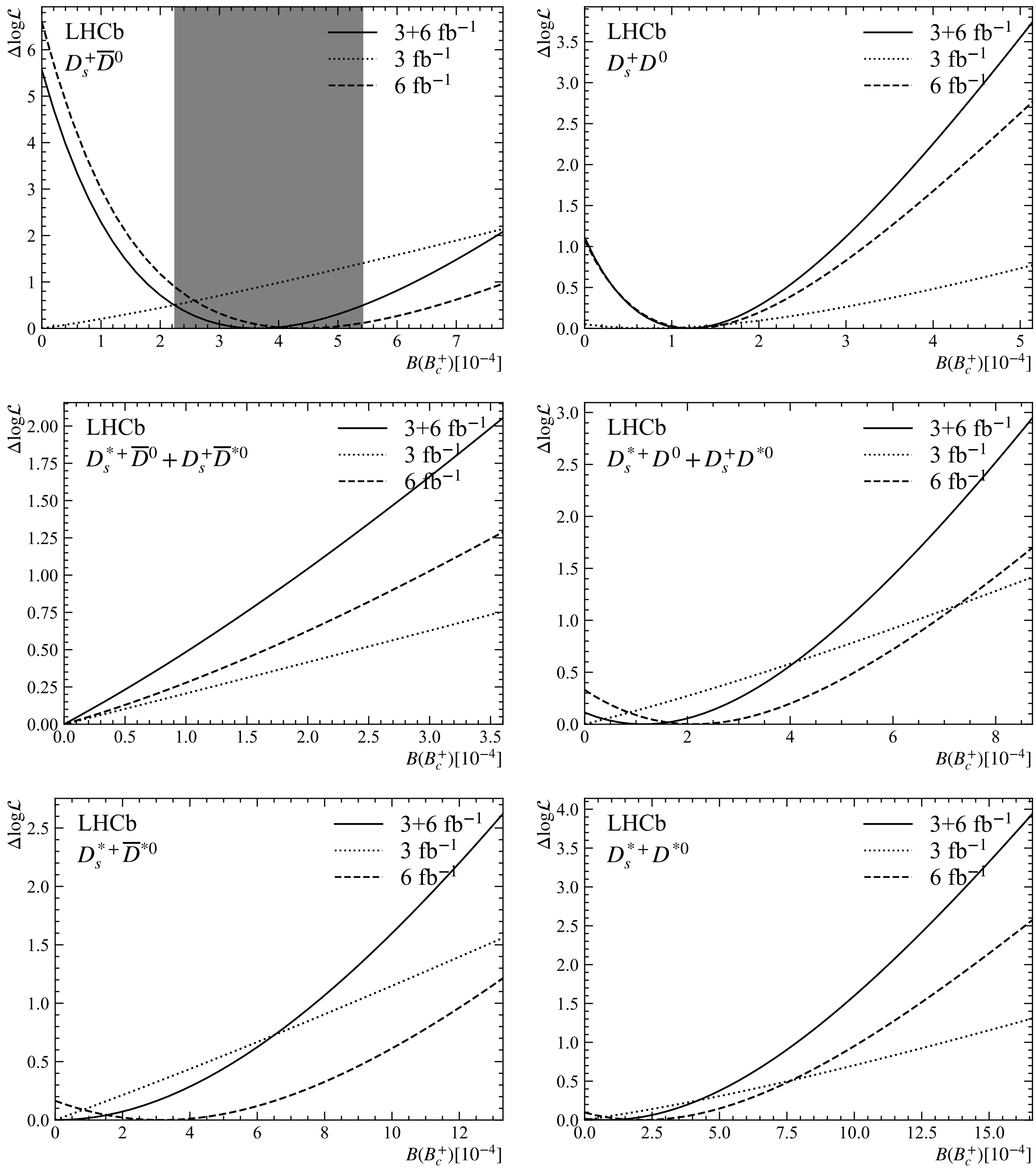

Figure C.1: Negative log profile likelihood in (dotted line) Run 1, (dashed line) Run 2 and (solid line) Run 1+2 combination for $B_c^+ \to D_s^+ D$ decays. Where the combined significance is greater than 2σ, the 68% CR is shaded. The top row is for $B_c^+ \to DD$ decays; the middle row is for $B_c^+ \to D^* D$ decays; the bottom row is for $B_c^+ \to D^* D^*$ decays. The left column is for RS decays; the right column is for WS decays.

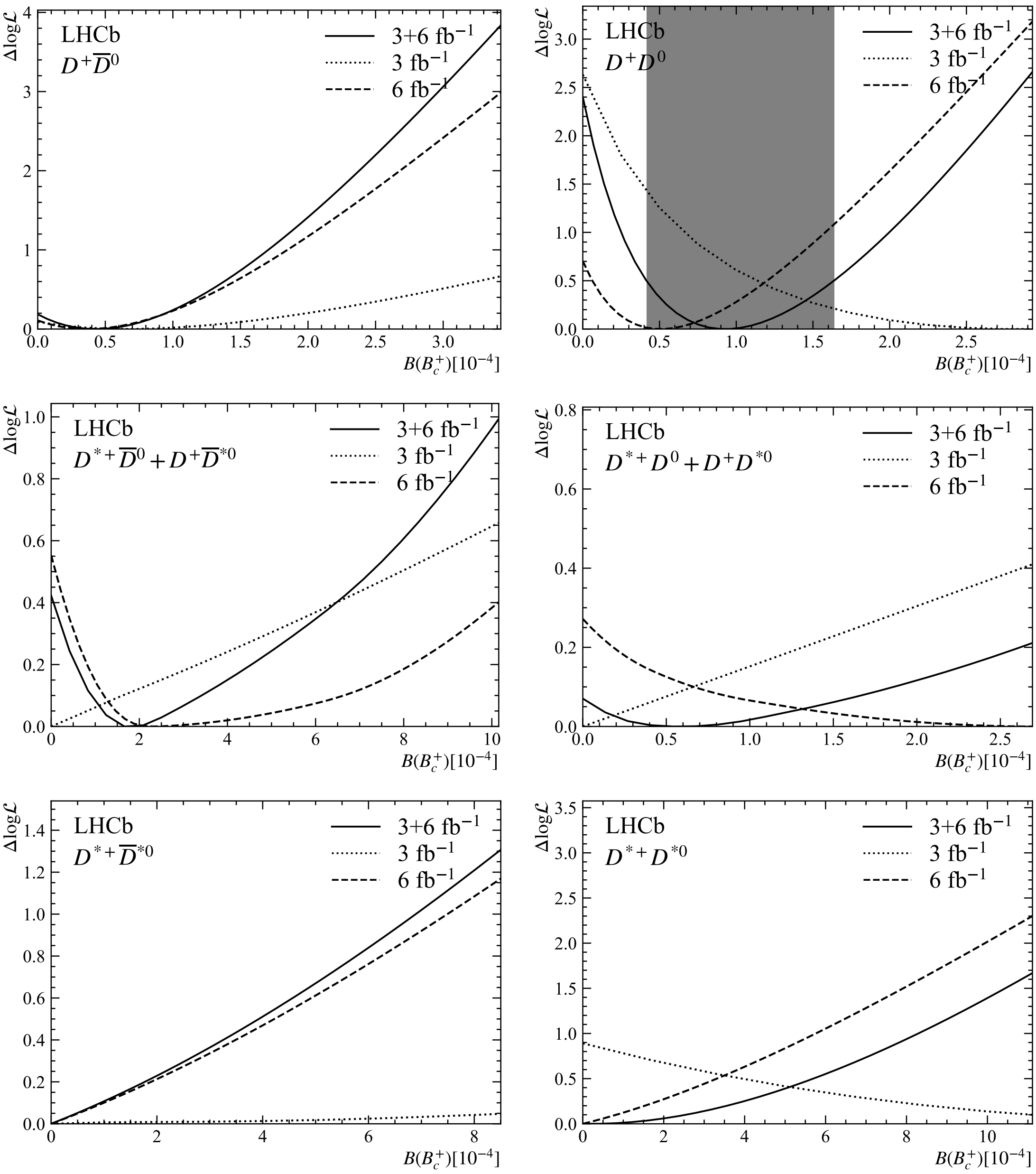

Figure C.2: Negative log profile likelihood in (dotted line) Run 1, (dashed line) Run 2 and (solid line) Run 1+2 combination for $B_c^+ \to D^+D$ decays. Where the combined significance is greater than 2σ, the 68% CR is shaded. The top row is for $B_c^+ \to DD$ decays; the middle row is for $B_c^+ \to D^*D$ decays; the bottom row is for $B_c^+ \to D^*D^*$ decays. The left column is for RS decays; the right column is for WS decays.

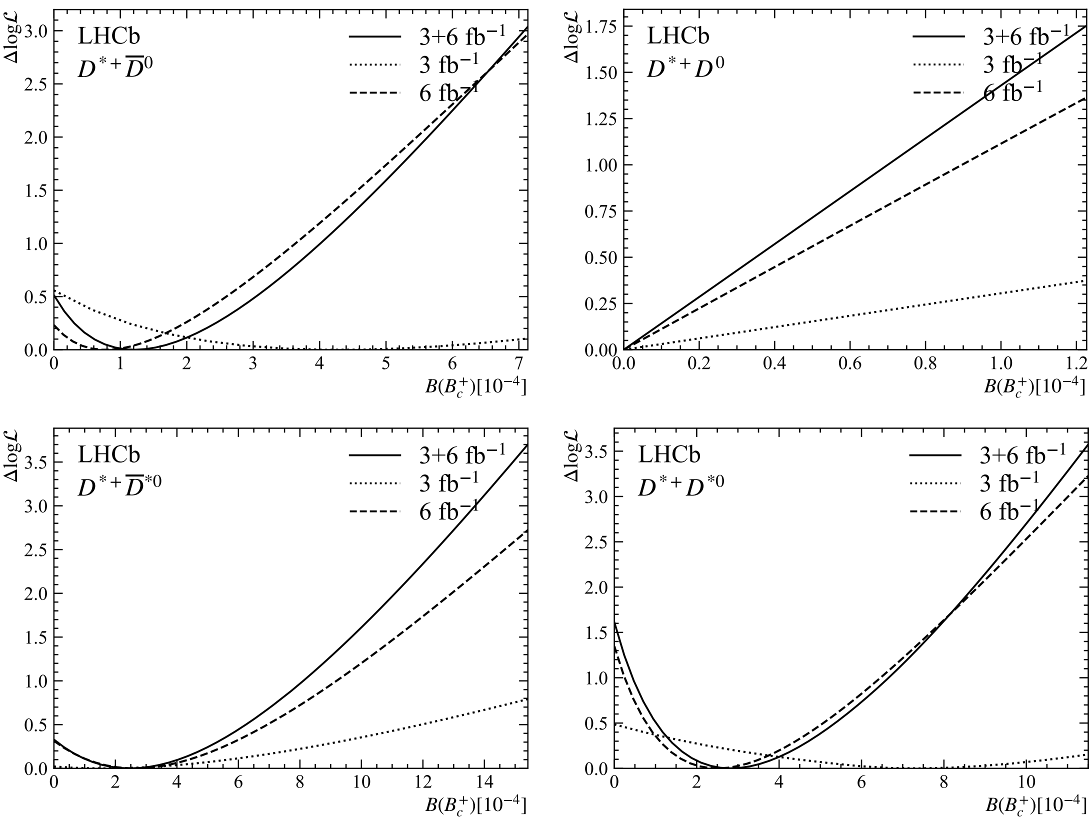

Figure C.3: Negative log profile likelihood in (dotted line) Run 1, (dashed line) Run 2 and (solid line) Run 1+2 combination for $B_c^+ \to D^{*+}D$ decays. Where the combined significance is greater than 2σ, the 68% CR is shaded. The top row is for $B_c^+ \to D^*D$ decays; the bottom row is for $B_c^+ \to D^*D^*$ decays. The left column is for RS decays; the right column is for WS decays.

CLs curves for branching fractions

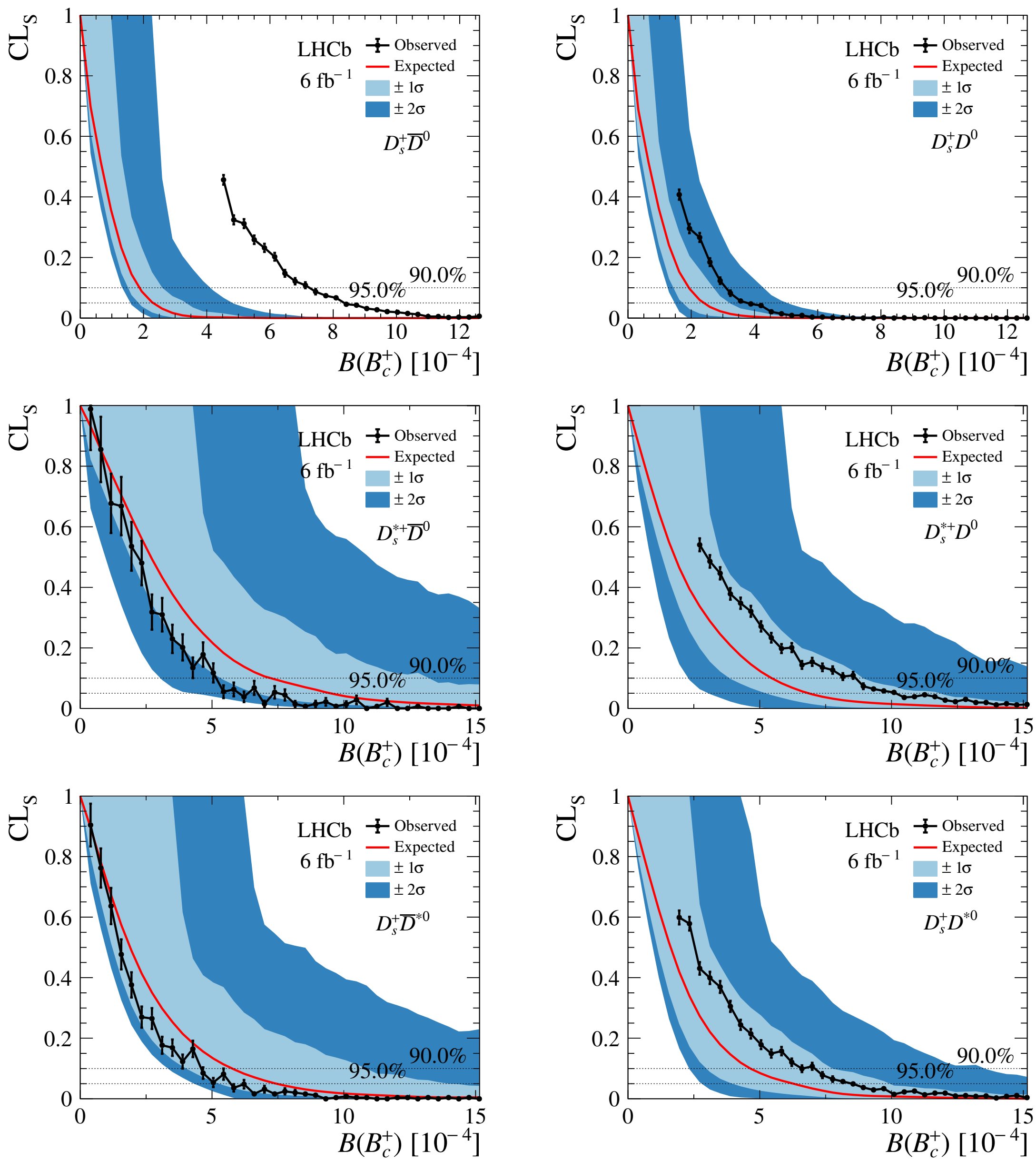

Figure D.1: CL$_s$ distributions for the branching fractions of (top left) $B_c^+ \to D_s^+\overline{D}^0$, (top right) $B_c^+ \to D_s^+D^0$, (middle left) $B_c^+ \to D_s^{*+}\overline{D}^0$, (middle right) $B_c^+ \to D_s^{*+}D^0$, (bottom left) $B_c^+ \to D_s^+\overline{D}^{*0}$, and (bottom right) $B_c^+ \to D_s^+D^{*0}$ decays in Run 2 data.

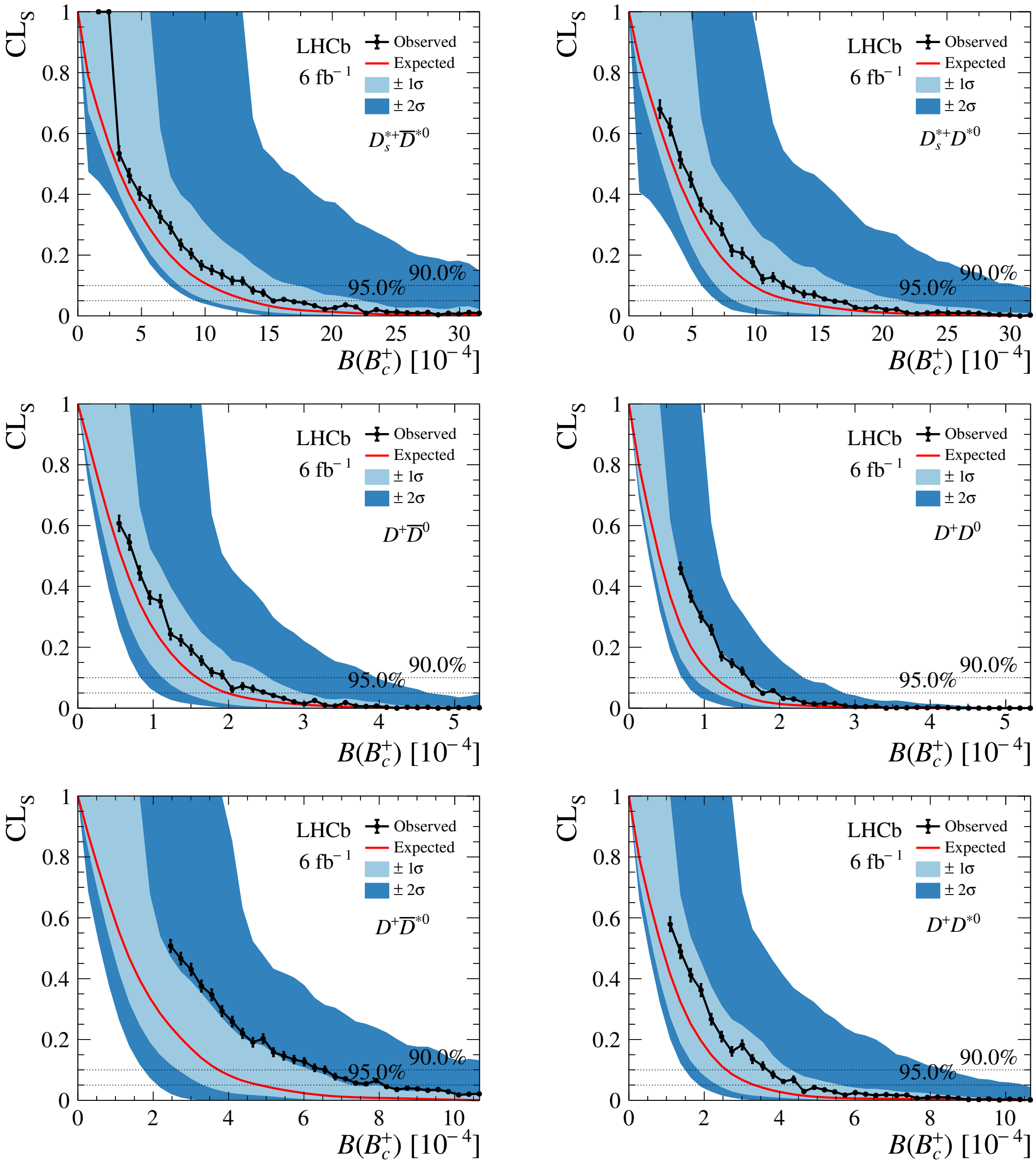

Figure D.2: CL$_s$ distributions for the branching fractions of (top left) $B_c^+ \to D_s^{*+}\overline{D}^{*0}$, (top right) $B_c^+ \to D_s^{*+}D^{*0}$, (middle left) $B_c^+ \to D^+\overline{D}^0$, (middle right) $B_c^+ \to D^+D^0$, (bottom left) $B_c^+ \to D^{*+}\overline{D}^0, D^+\overline{D}^{*0}$, and (bottom right) $B_c^+ \to D^{*+}D^0, D^+D^{*0}$ decays in Run 2 data.

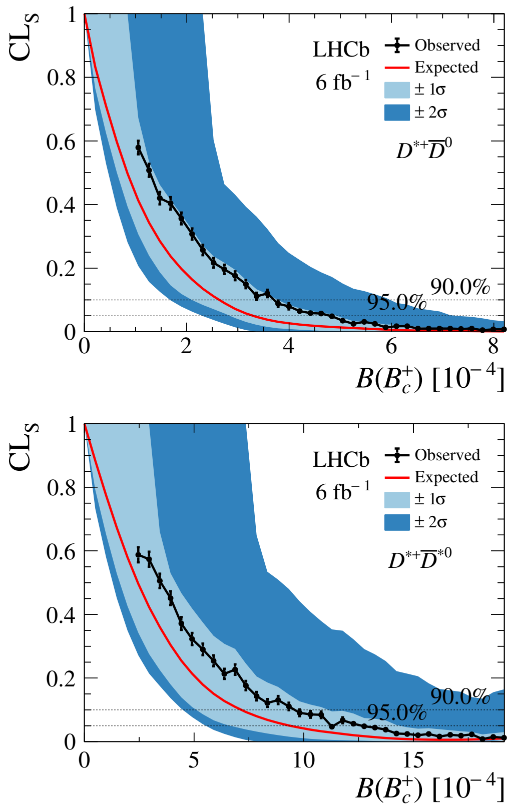

Figure D.3: CL_s distributions for the branching fractions of (top left) $B_c^+ \to D^{*+}\overline{D}^0$, (top right) $B_c^+ \to D^{*+}D^0$, (bottom left) $B_c^+ \to D^{*+}\overline{D}^{*0}$, and (bottom right) $B_c^+ \to D^{*+}D^{*0}$ decays in Run 2 data.

Bias and coverage on raw asymmetry

Table E.1: Bias on central value, and mean and standard deviation of pull distributions, from fits to pseudoexperiments generated under multiple $\mathcal{A}^{\text{raw}}$ hypotheses. For $B^- \to D_s^- D^0$ candidates in Run 1 data.

Generated $\mathcal{A}^{\text{raw}}$	Fitted $\mathcal{A}^{\text{raw}}$	$KK\pi, K\pi$ $\mu - \mathcal{A}^{\text{raw}}_{\text{gen}}/\%$	$KK\pi, K\pi$ μ_{pull}	$KK\pi, K\pi$ σ_{pull}	$KK\pi, K3\pi$ $\mu - \mathcal{A}^{\text{raw}}_{\text{gen}}/\%$	$KK\pi, K3\pi$ μ_{pull}	$KK\pi, K3\pi$ σ_{pull}
0%	$D_s^- D^0$	-0.03 ± 0.02	-0.05 ± 0.03	0.99 ± 0.02	-0.02 ± 0.03	-0.01 ± 0.03	1.02 ± 0.02
	$D_s^{*-} D^0$	0.03 ± 0.10	0.01 ± 0.03	0.99 ± 0.02	0.12 ± 0.19	0.02 ± 0.03	0.95 ± 0.02
	$D_s^- D^{*0}$	-0.07 ± 0.07	-0.03 ± 0.03	1.00 ± 0.02	0.03 ± 0.08	-0.00 ± 0.03	0.97 ± 0.02
$D_s^- D^0 = 2\%$	$D_s^- D^0$	-0.01 ± 0.02	-0.01 ± 0.03	1.02 ± 0.02	-0.01 ± 0.03	-0.02 ± 0.03	0.96 ± 0.02
	$D_s^{*-} D^0$	0.04 ± 0.10	0.02 ± 0.03	0.97 ± 0.02	-0.30 ± 0.19	-0.04 ± 0.03	0.96 ± 0.02
	$D_s^- D^{*0}$	-0.02 ± 0.07	-0.02 ± 0.03	0.96 ± 0.02	0.15 ± 0.08	0.04 ± 0.03	0.98 ± 0.02
$D_s^- D^0 = 10\%$	$D_s^- D^0$	0.02 ± 0.02	0.03 ± 0.03	0.97 ± 0.02	0.04 ± 0.03	0.03 ± 0.03	1.03 ± 0.02
	$D_s^{*-} D^0$	0.13 ± 0.10	0.05 ± 0.03	0.99 ± 0.02	0.01 ± 0.19	0.00 ± 0.03	0.97 ± 0.02
	$D_s^- D^{*0}$	-0.02 ± 0.08	-0.02 ± 0.03	1.00 ± 0.02	-0.02 ± 0.08	-0.02 ± 0.03	1.01 ± 0.02
$D_s^{*-} D^0 = 2\%$	$D_s^- D^0$	0.01 ± 0.02	0.01 ± 0.03	0.99 ± 0.02	-0.00 ± 0.03	-0.00 ± 0.03	0.99 ± 0.02
	$D_s^{*-} D^0$	0.01 ± 0.10	-0.00 ± 0.03	0.99 ± 0.02	0.03 ± 0.19	-0.01 ± 0.03	0.97 ± 0.02
	$D_s^- D^{*0}$	0.08 ± 0.07	0.03 ± 0.03	0.98 ± 0.02	0.05 ± 0.08	0.00 ± 0.03	0.96 ± 0.02
$D_s^{*-} D^0 = 10\%$	$D_s^- D^0$	-0.02 ± 0.02	-0.03 ± 0.03	1.01 ± 0.02	0.02 ± 0.03	0.02 ± 0.03	0.98 ± 0.02
	$D_s^{*-} D^0$	0.01 ± 0.10	-0.02 ± 0.03	0.99 ± 0.02	-0.36 ± 0.19	-0.06 ± 0.03	0.93 ± 0.02
	$D_s^- D^{*0}$	-0.01 ± 0.08	-0.01 ± 0.03	0.99 ± 0.02	0.16 ± 0.08	0.04 ± 0.03	0.98 ± 0.02
$D_s^- D^{*0} = 2\%$	$D_s^- D^0$	-0.04 ± 0.02	-0.06 ± 0.03	0.96 ± 0.02	-0.01 ± 0.03	-0.01 ± 0.03	0.97 ± 0.02
	$D_s^{*-} D^0$	-0.08 ± 0.10	-0.03 ± 0.03	0.99 ± 0.02	-0.22 ± 0.19	-0.04 ± 0.03	0.98 ± 0.02
	$D_s^- D^{*0}$	0.02 ± 0.07	-0.00 ± 0.03	0.96 ± 0.02	0.09 ± 0.09	0.01 ± 0.03	1.02 ± 0.02
$D_s^- D^{*0} = 10\%$	$D_s^- D^0$	-0.01 ± 0.02	-0.01 ± 0.03	1.00 ± 0.02	-0.01 ± 0.03	-0.00 ± 0.03	0.99 ± 0.02
	$D_s^{*-} D^0$	0.17 ± 0.10	0.06 ± 0.03	0.95 ± 0.02	-0.11 ± 0.19	0.01 ± 0.03	0.98 ± 0.02
	$D_s^- D^{*0}$	-0.04 ± 0.07	-0.05 ± 0.03	0.94 ± 0.02	-0.02 ± 0.09	-0.05 ± 0.03	0.99 ± 0.02

Table E.2: Bias on central value, and mean and standard deviation of pull distributions, from fits to pseudoexperiments generated under multiple $\mathcal{A}^{\mathrm{raw}}$ hypotheses. For $B^- \to D_s^- D^0$ candidates in Run 2 data.

Generated $\mathcal{A}^{\mathrm{raw}}$	Fitted $\mathcal{A}^{\mathrm{raw}}$	$KK\pi, K\pi$ $\mu - \mathcal{A}^{\mathrm{raw}}_{\mathrm{gen}}/\%$	$KK\pi, K\pi$ μ_{pull}	$KK\pi, K\pi$ σ_{pull}	$KK\pi, K3\pi$ $\mu - \mathcal{A}^{\mathrm{raw}}_{\mathrm{gen}}/\%$	$KK\pi, K3\pi$ μ_{pull}	$KK\pi, K3\pi$ σ_{pull}
0%	$D_s^- D^0$	-0.01 ± 0.01	-0.02 ± 0.04	1.03 ± 0.03	-0.01 ± 0.01	-0.03 ± 0.03	0.98 ± 0.02
	$D_s^{*-} D^0$	-0.03 ± 0.05	-0.03 ± 0.03	0.99 ± 0.02	0.05 ± 0.07	0.02 ± 0.03	0.99 ± 0.02
	$D_s^- D^{*0}$	0.01 ± 0.04	0.02 ± 0.03	0.99 ± 0.02	0.01 ± 0.04	0.01 ± 0.03	1.03 ± 0.02
$D_s^- D^0$ $= 2\%$	$D_s^- D^0$	0.01 ± 0.01	0.03 ± 0.04	1.00 ± 0.02	-0.02 ± 0.01	-0.05 ± 0.03	0.97 ± 0.02
	$D_s^{*-} D^0$	-0.13 ± 0.05	-0.08 ± 0.04	1.01 ± 0.02	-0.06 ± 0.07	-0.03 ± 0.03	1.01 ± 0.02
	$D_s^- D^{*0}$	0.03 ± 0.04	0.02 ± 0.04	1.07 ± 0.03	0.00 ± 0.05	0.01 ± 0.03	1.05 ± 0.02
$D_s^- D^0$ $= 10\%$	$D_s^- D^0$	0.00 ± 0.01	-0.01 ± 0.03	0.99 ± 0.02	-0.02 ± 0.01	-0.04 ± 0.03	1.01 ± 0.02
	$D_s^{*-} D^0$	0.06 ± 0.05	0.06 ± 0.04	1.03 ± 0.03	0.00 ± 0.07	0.00 ± 0.03	0.97 ± 0.02
	$D_s^- D^{*0}$	-0.01 ± 0.04	-0.01 ± 0.04	1.05 ± 0.03	-0.00 ± 0.04	-0.00 ± 0.03	0.96 ± 0.02
$D_s^{*-} D^0$ $= 2\%$	$D_s^- D^0$	-0.01 ± 0.01	-0.02 ± 0.03	1.01 ± 0.02	0.03 ± 0.01	0.08 ± 0.03	0.94 ± 0.02
	$D_s^{*-} D^0$	0.01 ± 0.05	0.00 ± 0.03	0.99 ± 0.02	0.02 ± 0.07	0.01 ± 0.03	0.97 ± 0.02
	$D_s^- D^{*0}$	-0.01 ± 0.04	-0.01 ± 0.03	0.99 ± 0.02	-0.01 ± 0.04	-0.01 ± 0.03	0.99 ± 0.02
$D_s^{*-} D^0$ $= 10\%$	$D_s^- D^0$	0.00 ± 0.01	0.01 ± 0.03	0.99 ± 0.02	0.00 ± 0.01	0.02 ± 0.03	1.00 ± 0.02
	$D_s^{*-} D^0$	0.02 ± 0.05	0.02 ± 0.03	1.00 ± 0.02	-0.06 ± 0.07	-0.03 ± 0.03	0.98 ± 0.02
	$D_s^- D^{*0}$	0.03 ± 0.04	0.02 ± 0.03	0.95 ± 0.02	0.00 ± 0.04	-0.00 ± 0.03	0.96 ± 0.02
$D_s^- D^{*0}$ $= 2\%$	$D_s^- D^0$	0.01 ± 0.01	0.03 ± 0.04	1.01 ± 0.02	-0.00 ± 0.01	-0.01 ± 0.03	0.99 ± 0.02
	$D_s^{*-} D^0$	-0.04 ± 0.05	-0.03 ± 0.03	0.95 ± 0.02	0.03 ± 0.07	0.01 ± 0.03	1.00 ± 0.02
	$D_s^- D^{*0}$	0.01 ± 0.04	-0.00 ± 0.03	0.93 ± 0.02	0.04 ± 0.05	0.03 ± 0.03	1.02 ± 0.02
$D_s^- D^{*0}$ $= 10\%$	$D_s^- D^0$	0.00 ± 0.01	0.01 ± 0.04	1.01 ± 0.03	-0.02 ± 0.01	-0.04 ± 0.03	1.01 ± 0.02
	$D_s^{*-} D^0$	-0.01 ± 0.05	-0.01 ± 0.03	1.00 ± 0.02	-0.03 ± 0.07	-0.01 ± 0.03	0.96 ± 0.02
	$D_s^- D^{*0}$	-0.01 ± 0.04	-0.02 ± 0.04	1.01 ± 0.03	0.02 ± 0.04	-0.00 ± 0.03	1.00 ± 0.02

Table E.3: Bias on central value, and mean and standard deviation of pull distributions, from fits to pseudoexperiments generated under multiple $\mathcal{A}^{\text{raw}}$ hypotheses. For $B^- \to D^- D^0$ candidates in Run 1 data.

Generated $\mathcal{A}^{\text{raw}}$	Fitted $\mathcal{A}^{\text{raw}}$	$K\pi\pi, K\pi$ $\mu - \mathcal{A}^{\text{raw}}_{\text{gen}}/\%$	μ_{pull}	σ_{pull}	$K\pi\pi, K3\pi$ $\mu - \mathcal{A}^{\text{raw}}_{\text{gen}}/\%$	μ_{pull}	σ_{pull}
0%	D^-D^0	-0.15 ± 0.09	-0.05 ± 0.03	0.98 ± 0.02	0.07 ± 0.12	0.02 ± 0.03	0.97 ± 0.02
	D^-D^{*0}	-0.08 ± 0.18	0.01 ± 0.03	0.92 ± 0.02	-0.12 ± 0.33	0.04 ± 0.03	0.84 ± 0.02
D^-D^0	D^-D^0	0.11 ± 0.09	0.04 ± 0.03	0.99 ± 0.02	-0.09 ± 0.13	-0.03 ± 0.03	1.01 ± 0.02
$= 2\%$	D^-D^{*0}	0.17 ± 0.17	0.06 ± 0.03	0.91 ± 0.02	0.09 ± 0.34	0.04 ± 0.03	0.84 ± 0.02
D^-D^0	D^-D^0	-0.03 ± 0.09	-0.02 ± 0.03	0.99 ± 0.02	0.07 ± 0.12	0.02 ± 0.03	0.98 ± 0.02
$= 10\%$	D^-D^{*0}	-0.20 ± 0.19	0.01 ± 0.03	0.97 ± 0.02	0.25 ± 0.35	0.08 ± 0.03	0.87 ± 0.02
D^-D^{*0}	D^-D^0	-0.08 ± 0.09	-0.03 ± 0.03	1.03 ± 0.02	0.23 ± 0.13	0.05 ± 0.03	1.04 ± 0.02
$= 2\%$	D^-D^{*0}	-0.01 ± 0.18	0.00 ± 0.03	0.94 ± 0.02	-0.16 ± 0.35	-0.00 ± 0.03	0.82 ± 0.02
D^-D^{*0}	D^-D^0	-0.10 ± 0.09	-0.04 ± 0.03	0.99 ± 0.02	0.04 ± 0.12	0.01 ± 0.03	0.99 ± 0.02
$= 10\%$	D^-D^{*0}	0.22 ± 0.19	-0.05 ± 0.03	0.96 ± 0.02	0.86 ± 0.37	-0.13 ± 0.03	0.89 ± 0.02

Table E.4: Bias on central value, and mean and standard deviation of pull distributions, from fits to pseudoexperiments generated under multiple $\mathcal{A}^{\text{raw}}$ hypotheses. For $B^- \to D^- D^0$ candidates in Run 2 data.

Generated $\mathcal{A}^{\text{raw}}$	Fitted $\mathcal{A}^{\text{raw}}$	$K\pi\pi, K\pi$ $\mu - \mathcal{A}^{\text{raw}}_{\text{gen}}/\%$	μ_{pull}	σ_{pull}	$K\pi\pi, K3\pi$ $\mu - \mathcal{A}^{\text{raw}}_{\text{gen}}/\%$	μ_{pull}	σ_{pull}
0%	D^-D^0	0.04 ± 0.05	0.03 ± 0.03	1.04 ± 0.02	0.05 ± 0.06	0.03 ± 0.03	1.01 ± 0.02
	D^-D^{*0}	-0.01 ± 0.09	-0.02 ± 0.03	1.01 ± 0.02	-0.09 ± 0.11	-0.02 ± 0.03	0.95 ± 0.02
D^-D^0	D^-D^0	0.09 ± 0.04	0.07 ± 0.03	1.02 ± 0.02	-0.00 ± 0.06	-0.00 ± 0.03	0.96 ± 0.02
$= 2\%$	D^-D^{*0}	0.01 ± 0.09	-0.00 ± 0.03	0.99 ± 0.02	0.14 ± 0.12	0.03 ± 0.03	0.98 ± 0.02
D^-D^0	D^-D^0	-0.06 ± 0.04	-0.05 ± 0.03	1.02 ± 0.02	0.02 ± 0.06	0.02 ± 0.03	1.01 ± 0.02
$= 10\%$	D^-D^{*0}	0.09 ± 0.09	0.03 ± 0.03	1.00 ± 0.02	0.07 ± 0.11	0.02 ± 0.03	0.97 ± 0.02
D^-D^{*0}	D^-D^0	0.06 ± 0.04	0.04 ± 0.03	0.98 ± 0.02	0.01 ± 0.06	0.01 ± 0.03	1.05 ± 0.02
$= 2\%$	D^-D^{*0}	0.08 ± 0.09	0.02 ± 0.03	0.96 ± 0.02	0.01 ± 0.12	-0.02 ± 0.03	1.00 ± 0.02
D^-D^{*0}	D^-D^0	0.06 ± 0.04	0.05 ± 0.03	1.03 ± 0.02	-0.02 ± 0.06	-0.02 ± 0.03	1.01 ± 0.02
$= 10\%$	D^-D^{*0}	-0.07 ± 0.10	-0.07 ± 0.03	0.96 ± 0.02	-0.02 ± 0.12	-0.06 ± 0.03	0.97 ± 0.02

Table E.5: Bias on central value, and mean and standard deviation of pull distributions, from fits to pseudoexperiments generated under multiple $\mathcal{A}^{\rm raw}$ hypotheses. For $B^- \to D^{*-}D^0$ candidates in Run 1 data.

Generated $\mathcal{A}^{\rm raw}$	Fitted $\mathcal{A}^{\rm raw}$	$K\pi,K\pi$ $\mu-\mathcal{A}^{\rm raw}_{\rm gen}/\%$	$\mu_{\rm pull}$	$\sigma_{\rm pull}$	$K\pi,K3\pi$ $\mu-\mathcal{A}^{\rm raw}_{\rm gen}/\%$	$\mu_{\rm pull}$	$\sigma_{\rm pull}$	$K3\pi,K\pi$ $\mu-\mathcal{A}^{\rm raw}_{\rm gen}/\%$	$\mu_{\rm pull}$	$\sigma_{\rm pull}$
0%	$D^{*-}D^0$	0.18 ± 0.17	0.03 ± 0.03	0.98 ± 0.02	0.05 ± 0.23	0.00 ± 0.03	1.01 ± 0.02	-0.05 ± 0.22	-0.01 ± 0.03	0.99 ± 0.02
	$D^{*-}D^{*0}$	-0.10 ± 0.18	-0.02 ± 0.03	0.93 ± 0.02	0.56 ± 0.30	0.04 ± 0.03	0.87 ± 0.02	-0.08 ± 0.30	-0.01 ± 0.03	0.84 ± 0.02
$D^{*-}D^0$	$D^{*-}D^0$	0.15 ± 0.17	0.03 ± 0.03	0.98 ± 0.02	0.20 ± 0.23	0.03 ± 0.03	1.04 ± 0.02	-0.17 ± 0.21	-0.02 ± 0.03	1.00 ± 0.02
$= 2\%$	$D^{*-}D^{*0}$	0.10 ± 0.18	0.02 ± 0.03	0.93 ± 0.02	0.10 ± 0.30	0.02 ± 0.03	0.87 ± 0.02	0.15 ± 0.30	0.01 ± 0.03	0.84 ± 0.02
$D^{*-}D^0$	$D^{*-}D^0$	-0.21 ± 0.17	-0.03 ± 0.03	0.98 ± 0.02	-0.11 ± 0.23	-0.01 ± 0.03	1.03 ± 0.02	-0.04 ± 0.21	0.00 ± 0.03	0.98 ± 0.02
$= 10\%$	$D^{*-}D^{*0}$	-0.48 ± 0.18	-0.08 ± 0.03	0.92 ± 0.02	-0.16 ± 0.29	-0.03 ± 0.03	0.87 ± 0.02	-0.48 ± 0.32	-0.02 ± 0.03	0.87 ± 0.02
$D^{*-}D^{*0}$	$D^{*-}D^0$	0.05 ± 0.17	0.01 ± 0.03	0.98 ± 0.02	0.36 ± 0.22	0.05 ± 0.03	0.99 ± 0.02	0.05 ± 0.22	0.01 ± 0.03	1.01 ± 0.02
$= 2\%$	$D^{*-}D^{*0}$	-0.02 ± 0.17	-0.01 ± 0.03	0.90 ± 0.02	-0.16 ± 0.28	-0.04 ± 0.03	0.84 ± 0.02	-0.69 ± 0.31	-0.07 ± 0.03	0.86 ± 0.02
$D^{*-}D^{*0}$	$D^{*-}D^0$	-0.11 ± 0.18	-0.02 ± 0.03	1.01 ± 0.02	-0.21 ± 0.23	-0.03 ± 0.03	1.01 ± 0.02	-0.35 ± 0.23	-0.05 ± 0.03	1.02 ± 0.02
$= 10\%$	$D^{*-}D^{*0}$	-0.20 ± 0.18	-0.09 ± 0.03	0.91 ± 0.02	0.73 ± 0.31	-0.09 ± 0.03	0.90 ± 0.02	-0.24 ± 0.32	-0.19 ± 0.03	0.91 ± 0.02

Table E.6: Bias on central value, and mean and standard deviation of pull distributions, from fits to pseudoexperiments generated under multiple $\mathcal{A}^{\rm raw}$ hypotheses. For $B^- \to D^{*-}D^0$ candidates in Run 2 data.

Generated $\mathcal{A}^{\rm raw}$	Fitted $\mathcal{A}^{\rm raw}$	$K\pi,K\pi$ $\mu-\mathcal{A}^{\rm raw}_{\rm gen}/\%$	$\mu_{\rm pull}$	$\sigma_{\rm pull}$	$K\pi,K3\pi$ $\mu-\mathcal{A}^{\rm raw}_{\rm gen}/\%$	$\mu_{\rm pull}$	$\sigma_{\rm pull}$	$K3\pi,K\pi$ $\mu-\mathcal{A}^{\rm raw}_{\rm gen}/\%$	$\mu_{\rm pull}$	$\sigma_{\rm pull}$
0%	$D^{*-}D^0$	-0.06 ± 0.08	-0.02 ± 0.03	0.97 ± 0.02	0.00 ± 0.10	0.00 ± 0.03	0.97 ± 0.02	-0.01 ± 0.11	-0.00 ± 0.03	1.02 ± 0.02
	$D^{*-}D^{*0}$	-0.12 ± 0.10	-0.03 ± 0.03	0.89 ± 0.02	-0.03 ± 0.11	-0.02 ± 0.03	0.85 ± 0.02	0.05 ± 0.10	0.01 ± 0.03	0.87 ± 0.02
$D^{*-}D^0$	$D^{*-}D^0$	0.03 ± 0.09	0.01 ± 0.03	1.00 ± 0.02	-0.19 ± 0.11	-0.05 ± 0.03	1.02 ± 0.02	0.09 ± 0.10	0.01 ± 0.03	1.01 ± 0.02
$= 2\%$	$D^{*-}D^{*0}$	-0.04 ± 0.10	-0.01 ± 0.03	0.87 ± 0.02	-0.07 ± 0.12	-0.03 ± 0.03	0.92 ± 0.02	0.22 ± 0.10	0.05 ± 0.03	0.87 ± 0.02
$D^{*-}D^0$	$D^{*-}D^0$	-0.01 ± 0.08	0.00 ± 0.03	0.97 ± 0.02	0.03 ± 0.11	0.01 ± 0.03	1.01 ± 0.02	0.06 ± 0.10	0.01 ± 0.03	0.97 ± 0.02
$= 10\%$	$D^{*-}D^{*0}$	0.05 ± 0.10	0.01 ± 0.03	0.91 ± 0.02	0.10 ± 0.11	0.02 ± 0.03	0.92 ± 0.02	0.02 ± 0.10	0.01 ± 0.03	0.87 ± 0.02
$D^{*-}D^{*0}$	$D^{*-}D^0$	-0.08 ± 0.09	-0.02 ± 0.03	1.01 ± 0.02	-0.20 ± 0.10	-0.06 ± 0.03	0.98 ± 0.02	-0.10 ± 0.11	-0.03 ± 0.03	1.02 ± 0.02
$= 2\%$	$D^{*-}D^{*0}$	0.06 ± 0.09	0.01 ± 0.03	0.86 ± 0.02	0.08 ± 0.11	-0.00 ± 0.03	0.89 ± 0.02	0.33 ± 0.10	0.06 ± 0.03	0.88 ± 0.02
$D^{*-}D^{*0}$	$D^{*-}D^0$	-0.06 ± 0.09	-0.03 ± 0.03	1.02 ± 0.02	0.03 ± 0.10	0.01 ± 0.03	0.98 ± 0.02	0.01 ± 0.11	0.02 ± 0.03	1.01 ± 0.02
$= 10\%$	$D^{*-}D^{*0}$	0.08 ± 0.10	-0.04 ± 0.03	0.89 ± 0.02	0.57 ± 0.12	0.06 ± 0.03	0.91 ± 0.02	0.13 ± 0.11	-0.04 ± 0.03	0.90 ± 0.02

The LHCb detector in cross stitch

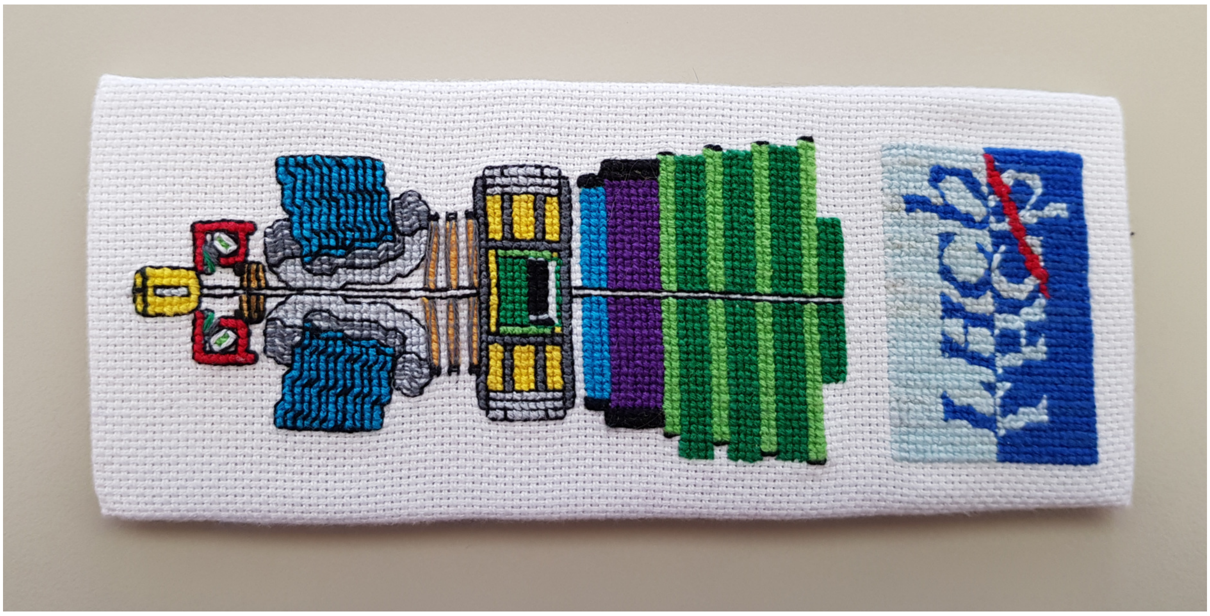

Figure F.1: Depiction of the Upgrade I LHCb detector in cross stitch. Author's own work, based on schematic diagram in Ref. [99].

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Updated search for B_c^+ decays to two charm mesons*, JHEP **12** (2021) 117, `arXiv:2109.00488`.

[2] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetries and branching fractions of B^- decays to two charm mesons*, `arXiv:2306.09945`, submitted to JHEP.

[3] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579.

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[5] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

[6] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[9] UA1, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. B **122** (1983) 103.

[10] UA2, M. Banner *et al.*, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. B **122** (1983) 476.

[11] UA1, G. Arnison *et al.*, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, Phys. Lett. B **126** (1983) 398.

[12] UA2, P. Bagnaia *et al.*, *Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider*, Phys. Lett. B **129** (1983) 130.

[13] CDF, F. Abe *et al.*, *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626, `arXiv:hep-ex/9503002`.

[14] D0, S. Abachi *et al.*, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632, `arXiv:hep-ex/9503003`.

[15] DONUT, K. Kodama *et al.*, *Observation of tau neutrino interactions*, Phys. Lett. B **504** (2001) 218, `arXiv:hep-ex/0012035`.

[16] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, `arXiv:1207.7214`.

[17] CMS, S. Chatrchyan *et al.*, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, `arXiv:1207.7235`.

[18] D. Clowe *et al.*, *A direct empirical proof of the existence of dark matter*, Astrophys. J. Lett. **648** (2006) L109, `arXiv:astro-ph/0608407`.

[19] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr. *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/*, Astrophys. J. **238** (1980) 471.

[20] Planck, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020) A6, `arXiv:1807.06209`, [Erratum: Astron.Astrophys. 652, C4 (2021)].

[21] J. M. Cline, *Baryogenesis*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime*, 9, 2006. `arXiv:hep-ph/0609145`.

[22] Particle Data Group, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[23] R. H. Cyburt, B. D. Fields, K. A. Olive, and T.-H. Yeh, *Big Bang Nucleosynthesis: 2015*, Rev. Mod. Phys. **88** (2016) 015004, `arXiv:1505.01076`.

[24] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[25] G. 't Hooft, *Symmetry Breaking Through Bell-Jackiw Anomalies*, Phys. Rev. Lett. **37** (1976) 8.

[26] A. I. Bochkarev and M. E. Shaposhnikov, *Electroweak Production of Baryon Asymmetry and Upper Bounds on the Higgs and Top Masses*, Mod. Phys. Lett. A **2** (1987) 417.

[27] M. Dine *et al.*, *Towards the theory of the electroweak phase transition*, Phys. Rev. D **46** (1992) 550, `arXiv:hep-ph/9203203`.

[28] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pène, *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994) 795, `arXiv:hep-ph/9312215`.

[29] M. B. Gavela *et al.*, *Standard model CP violation and baryon asymmetry. Part 2: Finite temperature*, Nucl. Phys. **B430** (1994) 382, `arXiv:hep-ph/9406289`.

[30] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. **D51** (1995) 379, `arXiv:hep-ph/9404302`.

[31] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008), no. LHCb-DP-2008-001 S08005.

[32] M. Bordone *et al.*, *A puzzle in $\bar{B}^0_{(s)} \to D^{(*)+}_{(s)}\{\pi^-, K^-\}$ decays and extraction of the f_s/f_d fragmentation fraction*, Eur. Phys. J. **C80** (2020), no. 10 951, `arXiv:2007.10338`.

[33] S. Iguro and T. Kitahara, *Implications for new physics from a novel puzzle in $\bar{B}^0_{(s)} \to D^{(*)+}_{(s)}\{\pi^-, K^-\}$ decays*, Phys. Rev. **D102** (2020), no. 7 071701, `arXiv:2008.01086`.

[34] F.-M. Cai, W.-J. Deng, X.-Q. Li, and Y.-D. Yang, *Probing new physics in class-I B-meson decays into heavy-light final states*, `arXiv:2103.04138`.

[35] *Standard Model of Elementary Particles*, `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`. Accessed: 04/01/2023.

[36] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[37] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[38] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802.

[39] CKMfitter group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, `arXiv:hep-ph/0406184`, updated results and plots available at `http://ckmfitter.in2p3.fr/`.

[40] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K^0_2 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[41] KTeV, A. Alavi-Harati *et al.*, *Observation of direct CP violation in $K_{S,L} \to \pi\pi$ decays*, Phys. Rev. Lett. **83** (1999) 22, `arXiv:hep-ex/9905060`.

[42] NA48, V. Fanti *et al.*, *A New measurement of direct CP violation in two pion decays of the neutral kaon*, Phys. Lett. B **465** (1999) 335, `arXiv:hep-ex/9909022`.

[43] BaBar, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, `arXiv:hep-ex/0107013`.

[44] Belle, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.

[45] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in charm decays*, Phys. Rev. Lett. **122** (2019) 211803, arXiv:1903.08726.

[46] *Feynman Diagram Library*, https://www.physik.uzh.ch/~che/FeynDiag/index.php. Accessed: 27/01/2023.

[47] BaBar, D. Boutigny *et al.*, *The BABAR physics book: Physics at an asymmetric B factory*, 10, 1998. doi: 10.2172/979931.

[48] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B^0 \to D^{*\pm}D^{\mp}$ decays*, JHEP **03** (2020) 147, arXiv:1912.03723.

[49] Belle, M. Rohrken *et al.*, *Measurements of Branching Fractions and Time-dependent CP Violating Asymmetries in $B^0 \to D^{(*)\pm}D^{\mp}$ Decays*, Phys. Rev. D **85** (2012) 091106, arXiv:1203.6647.

[50] J. Brod, A. Lenz, G. Tetlalmatzi-Xolocotzi, and M. Wiebusch, *New physics effects in tree-level decays and the precision in the determination of the quark mixing angle γ*, Phys. Rev. D **92** (2015), no. 3 033002, arXiv:1412.1446.

[51] A. Lenz and G. Tetlalmatzi-Xolocotzi, *Model-independent bounds on new physics effects in non-leptonic tree-level decays of B-mesons*, JHEP **07** (2020) 177, arXiv:1912.07621.

[52] J. Brod and J. Zupan, *The ultimate theoretical error on γ from $B \to DK$ decays*, JHEP **01** (2014) 051, arXiv:1308.5663.

[53] LHCb collaboration, R. Aaij *et al.*, *Simultaneous determination of CKM angle γ and charm mixing parameters*, JHEP **12** (2021) 141, arXiv:2110.02350.

[54] M. Gronau and D. London, *How to determine all the angles of the unitarity triangle from $B_d^0 \to DK_S$ and $B_s^0 \to D\phi$*, Phys. Lett. B **253** (1991) 483.

[55] M. Gronau and D. Wyler, *On determining a weak phase from CP asymmetries in charged B decays*, Phys. Lett. B **265** (1991) 172.

[56] M. Nayak *et al.*, *First determination of the CP content of $D \to \pi^+\pi^-\pi^0$ and $D \to K^+K^-\pi^0$*, Phys. Lett. B **740** (2015) 1, arXiv:1410.3964.

[57] D. Atwood, I. Dunietz, and A. Soni, *Enhanced CP violation with $B \to KD^0(\overline{D}^0)$ modes and extraction of the CKM angle γ*, Phys. Rev. Lett. **78** (1997), no. FERMILAB-PUB-96-448-T, JLAB-THY-96-19 3257, arXiv:hep-ph/9612433.

[58] D. Atwood, I. Dunietz, and A. Soni, *Improved methods for observing CP violation in $B^{\pm} \to KD$ and measuring the CKM phase γ*, Phys. Rev. D **63** (2001), no. AMES-HET-00-09, BNL-HET-00-31 036005, `arXiv:hep-ph/0008090`.

[59] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, *Determining γ using $B^{\pm} \to DK^{\pm}$ with multibody D decays*, Phys. Rev. D **68** (2003) 054018, `arXiv:hep-ph/0303187`.

[60] Belle, A. Poluektov *et al.*, *Measurement of phi_3 with Dalitz plot analysis of $B^{\pm} \to D^{(*)}K^{\pm}$ decay*, Phys. Rev. D **70** (2004) 072003, `arXiv:hep-ex/0406067`.

[61] R. Aleksan, I. Dunietz, and B. Kasyer, *Determining the CP-violating phase γ*, Z. Phys. C **54** (1992) 653.

[62] R. Fleischer, *New strategies to obtain insights into CP violation through $B_s \to D_s^{\pm}K^{\mp}$, $D_s^{*\pm}K^{\mp}$, ... and $B_d \to D^{\pm}\pi^{\mp}$, $D^{*\pm}\pi^{\mp}$, ... Decays*, Nucl. Phys. B **671** (2003), no. CERN-TH-2003-084 459, `arXiv:hep-ph/0304027`.

[63] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \to D_s^{\mp}K^{\pm}$ decays*, JHEP **03** (2018) 059, `arXiv:1712.07428`.

[64] CLEO, J. Insler *et al.*, *Studies of the decays $D^0 \to K_S^0 K^- \pi^+$ and $D^0 \to K_S^0 K^+ \pi^-$*, Phys. Rev. D **85** (2012) 092016, `arXiv:1203.3804`, [Erratum: Phys.Rev.D 94, 099905 (2016)].

[65] J. Libby *et al.*, *New determination of the $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ coherence factors and average strong-phase differences*, Phys. Lett. B **731** (2014) 197, `arXiv:1401.1904`.

[66] T. Evans *et al.*, *Improved determination of the $D \to K^-\pi^+\pi^+\pi^-$ coherence factor and associated hadronic parameters from a combination of $e^+e^- \to \psi(3770) \to c\bar{c}$ and $pp \to c\bar{c}X$ data*, Phys. Lett. B **757** (2016) 520, `arXiv:1602.07430`, [Erratum: Phys.Lett.B 765, 402–403 (2017)].

[67] LHCb collaboration, R. Aaij *et al.*, *Studies of the resonance structure in $D^0 \to K_S^0 K^{\pm}\pi^{\mp}$ decays*, Phys. Rev. **D93** (2016) 052018, `arXiv:1509.06628`.

[68] BESIII, M. Ablikim *et al.*, *Measurement of the $D \to K^-\pi^+\pi^+\pi^-$ and $D \to K^-\pi^+\pi^0$ coherence factors and average strong-phase differences in quantum-correlated $D\bar{D}$ decays*, JHEP **05** (2021) 164, `arXiv:2103.05988`.

[69] LHCb collaboration, *Simultaneous determination of the CKM angle γ and parameters related to mixing and CP violation in the charm sector*, LHCb-CONF-2022-003, CERN-LHCb-CONF-2022-003.

[70] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the B_c^+ meson mass*, JHEP **07** (2020) 123, `arXiv:2004.08163`.

[71] CDF, F. Abe *et al.*, *Observation of the B_c meson in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **81** (1998) 2432, `arXiv:hep-ex/9805034`.

[72] Z. Rui, Z. Zhitian, and C.-D. Lu, *The Double charm decays of B_c Meson in the Perturbative QCD Approach*, Phys. Rev. **D86** (2012) 074019, `arXiv:1203.2303`.

[73] V. V. Kiselev, *Gold plated mode of CP violation in decays of B_c meson from QCD sum rules*, J. Phys. G **30** (2004) 1445, `arXiv:hep-ph/0302241`.

[74] M. A. Ivanov, J. G. Korner, and O. N. Pakhomova, *The Nonleptonic decays $B_c^+ \to D_s^+\bar{D}^0$ and $B_c^+ \to D_s^+D^0$ in a relativistic quark model*, Phys. Lett. **B555** (2003) 189, `arXiv:hep-ph/0212291`.

[75] M. A. Ivanov, J. G. Korner, and P. Santorelli, *Exclusive semileptonic and nonleptonic decays of the B_c meson*, Phys. Rev. **D73** (2006) 054024, `arXiv:hep-ph/0602050`.

[76] S. Naimuddin *et al.*, *Nonleptonic two-body Bc-meson decays*, Phys. Rev. D **86** (2012) 094028.

[77] M. Masetti, *CP violation in B_c decays*, Phys. Lett. **B286** (1992), no. ROME-871-1992 160.

[78] R. Fleischer and D. Wyler, *Exploring CP violation with B_c decays*, Phys. Rev. **D62** (2000), no. DESY-00-052, ZH-TH-7-00 057503, `arXiv:hep-ph/0004010`.

[79] A. K. Giri, R. Mohanta, and M. P. Khanna, *Determination of the angle gamma from nonleptonic $B_c \to D_sD^0$ decays*, Phys. Rev. **D65** (2002) 034016, `arXiv:hep-ph/0104009`.

[80] A. F. Falk, Y. Nir, and A. A. Petrov, *Strong phases and $D^0 - \overline{D}^0$ mixing parameters*, JHEP **12** (1999), no. IASSNS-HEP-99-104, JHU-TIPAC-99010 019, `arXiv:hep-ph/9911369`.

[81] C. S. Kim, R.-M. Wang, and Y.-D. Yang, *Studying Double Charm Decays of $B_{(u,d)}$ and B_s Mesons in the MSSM with R-parity Violation*, Phys. Rev. D **79** (2009) 055004, `arXiv:0812.4136`.

[82] L.-X. Lu, Z.-J. Xiao, S.-W. Wang, and W.-J. Li, *The Double charm decays of B Mesons in the mSUGRA model*, Commun. Theor. Phys. **56** (2011) 125, `arXiv:1008.4987`.

[83] Y.-G. Xu and R.-M. Wang, *Studying the fourth generation quark contributions to the double charm decays $B_{(s)} \to D_{(s)}^{(*)}D_s^{(*)}$*, Int. J. Theor. Phys. **55** (2016), no. 12 5290.

[84] R.-H. Li, X.-X. Wang, A. I. Sanda, and C.-D. Lu, *Decays of B meson to two charmed mesons*, Phys. Rev. D **81** (2010) 034006, arXiv:0910.1424.

[85] H.-F. Fu, G.-L. Wang, Z.-H. Wang, and X.-J. Chen, *Semi-leptonic and non-leptonic B meson decays to charmed mesons*, Chin. Phys. Lett. **28** (2011) 121301, arXiv:1202.1221.

[86] M. Jung and S. Schacht, *Standard model predictions and new physics sensitivity in $B \to DD$ decays*, Phys. Rev. D **91** (2015), no. 3 034027, arXiv:1410.8396.

[87] M. Gronau, J. L. Rosner, and D. Pirjol, *Small amplitude effects in $B^0 \to D^+D^-$ and related decays*, Phys. Rev. D **78** (2008), no. EFI-08-13 033011, arXiv:0805.4601.

[88] E. Lopienska, *The CERN accelerator complex, layout in 2022*, https://cds.cern.ch/record/2800984.

[89] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004. doi: 10.5170/CERN-2004-003-V-1.

[90] AC Team, *Diagram of an LHC dipole magnet*, https://cds.cern.ch/record/40524, Jun, 1999.

[91] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003.

[92] The CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004.

[93] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08002.

[94] L. Evans and P. Bryant, *LHC machine*, Journal of Instrumentation **3** (2008) S08001.

[95] M. Vretenar *et al.*, *Linac4 design report*, vol. 6 of *CERN Yellow Reports: Monographs*, CERN, Geneva, 2020. doi: 10.23731/CYRM-2020-006.

[96] M. Pepe Altarelli and F. Teubert, *B Physics at LHCb*, Int. J. Mod. Phys. A **23** (2008) 5117, arXiv:0802.1901.

[97] B. Muratori and T. Pieloni, *Luminosity levelling techniques for the LHC*, in *ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders*, pp. 177–181, 2014. arXiv:1410.5646. doi: 10.5170/CERN-2014-004.177.

[98] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[99] LHCb collaboration, R. Aaij *et al.*, *The LHCb Upgrade I*, `arXiv:2305.10515`, submitted to JINST.

[100] LHCb collaboration, *LHCb Framework TDR for the LHCb - Upgrade II Opportunities in flavour physics, and beyond, in the HL-LHC era*, CERN-LHCC-2021-012.

[101] *Schedule and luminosity forecasts*, `https://lhc-commissioning.web.cern.ch/schedule/HL-LHC-plots.htm`. Accessed: 04/12/2022.

[102] LHCb, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, tech. rep., CERN, Geneva, 2016. ISBN 978-92-9083-494-6, doi: 10.17181/CERN.QZRZ.R4S6.

[103] H. Dijkstra, H. J. Hilke, T. Nakada, and T. Ypsilantis, *LHCb Letter of Intent, LHCb Collaboration*, tech. rep., CERN, Geneva, 1995.

[104] M. Pepe Altarelli, *CP violation in charm and beauty decays at LHCb*, Nuclear Physics B - Proceedings Supplements **241-242** (2013) 43, Proceedings of the Fourth Workshop on Theory, Phenomenology and Experiments in Heavy Flavour Physics.

[105] LHCb collaboration, R. Aaij *et al.*, *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B_s^0 decay branching fractions*, Phys. Rev. **D104** (2021) 032005, `arXiv:2103.06810`.

[106] LHCb collaboration, R. Aaij *et al.*, *Measurement of the B_c^- production fraction and asymmetry in 7 and 13 TeV pp collisions*, Phys. Rev. **D100** (2019) 112006, `arXiv:1910.13404`.

[107] LHCb collaboration, C. Elsässer, *$\overline{b}b$ production angle plots*, `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html`. Accessed: 02/11/2022.

[108] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^{\pm}$ production cross-section in pp collisions at $\sqrt{s}$ =7 and 13 TeV*, JHEP **12** (2017) 026, `arXiv:1710.04921`.

[109] LHCb collaboration, R. Aaij *et al.*, *Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s}$ =7 TeV*, JHEP **08** (2013) 117, `arXiv:1306.3663`.

[110] LHCb collaboration, R. Aaij *et al.*, *Precise determination of the $B_s^0-\bar{B}_s^0$ oscillation frequency*, Nature Physics **18** (2022) 1, `arXiv:2104.04421`.

[111] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[112] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011.

[113] N. van Bakel and M. Ferro-Luzzi, *Shielding of the VELO detectors from the LHC beam high-frequency fields : preliminary considerations*, LHCb-NOTE-2001-081.

[114] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021.

[115] H. Kolanoski and N. Wermes, *Particle Detectors: Fundamentals and Applications*, Oxford University Press, 06, 2020. doi: 10.1093/oso/9780198858362.001.0001.

[116] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007.

[117] M. Vesterinen, *Considerations on the LHCb dipole magnet polarity reversal*, CERN-LHCb-PUB-2014-006. On behalf of the LHCb collaboration.

[118] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030.

[119] O. Steinkamp, *Layout and R&D for an All-Silicon TT Station*, LHCb-2002-056.

[120] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029.

[121] M. Tobin, *The LHCb Silicon Tracker*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **732** (2013) 168, Vienna Conference on Instrumentation 2013.

[122] J. J. van Hunen, *The LHCb tracking system*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **572** (2007), no. 1 149, Frontier Detectors for Frontier Physics.

[123] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, `arXiv:1311.3893`.

[124] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, tech. rep., CERN, Geneva, 2007.

[125] M. Needham and J. Van Tilburg, *Performance of the track matching*, tech. rep., CERN, Geneva, 2007.

[126] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987), no. 2 444.

[127] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, LHCb-PUB-2017-011.

[128] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[129] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron masses*, Phys. Lett. **B708** (2012) 241, `arXiv:1112.4896`.

[130] LHCb collaboration, R. Aaij *et al.*, *Measurements of the Λ_b^0, Ξ_b^-, and Ω_b^- baryon masses*, Phys. Rev. Lett. **110** (2013) 182001, `arXiv:1302.1072`.

[131] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022.

[132] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[133] LHCb collaboration, *Letter of Intent for the LHCb Upgrade*, CERN-LHCC-2011-001.

[134] A. Papanestis and C. D'Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **876** (2017) 221, The 9th international workshop on Ring Imaging Cherenkov Detectors (RICH2016).

[135] M. Alemi *et al.*, *First operation of a hybrid photon detector prototype with electrostatic cross-focussing and integrated silicon pixel readout*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **449** (2000), no. 1 48.

[136] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037.

[137] R. Forty, *RICH pattern recognition for LHCb*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **433** (1999), no. 1 257.

[138] R. Calabrese *et al.*, *Performance of the LHCb RICH detectors during LHC Run 2*, Journal of Instrumentation **17** (2022) P07013.

[139] LHCb collaboration, *PID Conference Plots*, `https://twiki.cern.ch/twiki/bin/view/LHCb/PIDConferencePlots`.

[140] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in $b \to s\ell^+\ell^-$ decays*, `arXiv:2212.09152`, to appear in Phys. Rev. Lett.

[141] C. Abellan Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, `arXiv:2008.11556`, submitted to JINST.

[142] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036.

[143] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B^0 \to J/\psi K^0_S$ decays*, Phys. Rev. Lett. **115** (2015) 031601, `arXiv:1503.07089`.

[144] LHCb collaboration, R. Aaij *et al.*, *Analysis of neutral B-meson decays into two muons*, Phys. Rev. Lett. **128** (2022) 041801, `arXiv:2108.09284`.

[145] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010.

[146] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[147] LHCb collaboration, *Trigger schemes*, `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html`.

[148] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019), no. LHCb-DP-2019-001 P04013, `arXiv:1812.10790`.

[149] A. Martin Sanchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 Calorimeter Trigger*, CERN-LHCb-PUB-2011-026.

[150] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

[151] R. Aaij *et al.*, *A comprehensive real-time analysis model at the LHCb experiment*, JINST **14** (2019) P04006, `arXiv:1903.01360`.

[152] M. Williams *et al.*, *The HLT2 Topological Lines*, CERN-LHCb-PUB-2011-002.

[153] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, CERN-LHCb-PUB-2011-016.

[154] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, `arXiv:1210.6861`.

[155] *ROOT Data Analysis Framework*, `https://root.cern.ch`. Accessed: 03/02/2023.

[156] G. Barrand *et al.*, *GAUDI A software architecture and framework for building HEP data processing applications*, Computer Physics Communications **140** (2001), no. 1 45, CHEP2000.

[157] LHCb collaboration, *LHCb computing: Technical Design Report*, CERN-LHCC-2005-019.

[158] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, `arXiv:physics/0503191`.

[159] *The LHCb data flow in Run 2*, `https://lhcb.github.io/starterkit-lessons/first-analysis-steps/run-2-data-flow.html`. Accessed: 27/11/2022.

[160] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001.

[161] R. Aaij *et al.*, *Allen: A high level trigger on GPUs for LHCb*, Comput. Softw. Big Sci. **4** (2020), no. 1 7, `arXiv:1912.09161`.

[162] LHCb collaboration, *LHCb Upgrade GPU High Level Trigger Technical Design Report*, CERN-LHCC-2020-006.

[163] LHCb collaboration, *RTA and DPA dataflow diagrams for Run 1, Run 2, and the upgraded LHCb detector*, LHCb-FIGURE-2020-016.

[164] LHCb, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[165] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

[166] C.-H. Chang, C. Driouichi, P. Eerola, and X. G. Wu, *BCVEGPY: An Event generator for hadronic production of the B_c meson*, Comput. Phys. Commun. **159** (2004) 192, `arXiv:hep-ph/0309120`.

[167] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[168] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur. Phys. J. C **45** (2006), no. IFJPAN-V-05-01, CERN-PH-TH-2005-091 97, `arXiv:hep-ph/0506026`.

[169] GEANT4, S. Agostinelli *et al.*, *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003), no. SLAC-PUB-9350, FERMILAB-PUB-03-339, CERN-IT-2002-003 250.

[170] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006), no. SLAC-PUB-11870 270.

[171] LHCb collaboration, R. Aaij *et al.*, *Search for B_c^+ decays to two charm mesons*, Nucl. Phys. **B930** (2018) 563, `arXiv:1712.04702`.

[172] O. Deschamps *et al.*, *Photon and neutral pion reconstruction*, LHCb-2003-091.

[173] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in $B^- \to D_s^- D^0$ and $B^- \to D^- D^0$ decays*, JHEP **05** (2018) 160, `arXiv:1803.10990`.

[174] Belle collaboration, I. Adachi *et al.*, *Measurement of the branching fraction and charge asymmetry of the decay $B^+ \to D^+ \overline{D}^0$ and search for $B^0 \to D^0 \overline{D}^0$*, Phys. Rev. **D77** (2008) 091101, `arXiv:0802.2988`.

[175] BaBar collaboration, B. Aubert *et al.*, *Measurement of branching fractions and CP-violating charge asymmetries for B meson decays to D(*) anti-D(*), and implications for the CKM angle gamma*, Phys. Rev. **D73** (2006) 112004, `arXiv:hep-ex/0604037`.

[176] Belle, G. Majumder *et al.*, *Evidence for $B^0 \to D^+ D^-$ and observation of $B^- \to D^0 D^-$ and $B^- \to D^0 D^{*-}$ decays*, Phys. Rev. Lett. **95** (2005) 041803, `arXiv:hep-ex/0502038`.

[177] BaBar, B. Aubert *et al.*, *Study of $B \to D^{(*)} D_{s(J)}^{(*)}$ Decays and Measurement of D_s^- and $D_{sJ}(2460)$ Branching Fractions*, Phys. Rev. D **74** (2006) 031103, `arXiv:hep-ex/0605036`.

[178] LHCb collaboration, R. Aaij *et al.*, *First observations of $\overline{B}_s^0 \to D^+ D^-, D_s^+ D^-$ and $D^0 \overline{D}^0$ decays*, Phys. Rev. **D87** (2013) 092007, `arXiv:1302.5854`.

[179] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K_S^0 K_S^0$ decays*, Phys. Rev. **D104** (2021) L031102, `arXiv:2105.01565`.

[180] LHCb Collaboration, *LHCb Performance Numbers*, `https://lhcb.web.cern.ch/speakersbureau/html/PerformanceNumbers.html`. Accessed: 19/06/2020.

[181] A. Tully. Private communication, 2019.

[182] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2019) 1, `arXiv:1803.00824`.

[183] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016), no. 1 012036, `arXiv:1608.05806`, `https://github.com/arogozhnikov/hep_ml`.

[184] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[185] A. Blum, A. Kalai, and J. Langford, *Beating the Hold-out: Bounds for K-fold and Progressive Cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, (New York, NY, USA), pp. 203–208, ACM, 1999. doi: 10.1145/307400.307439.

[186] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[187] C. Tsallis, *Possible Generalization of Boltzmann-Gibbs Statistics*, J. Statist. Phys. **52** (1988) 479.

[188] C. Tsallis, *Nonextensive statistics: Theoretical, experimental and computational evidences and connections*, Braz. J. Phys. **29** (1999) 1.

[189] LHCb collaboration, R. Aaij *et al.*, *Measurement of B_c^+ production in proton-proton collisions at $\sqrt{s}$ =8 TeV*, Phys. Rev. Lett. **114** (2015) 132001, `arXiv:1411.2943`.

[190] LHCb collaboration, *Differences between Sim09 versions*, `https://twiki.cern.ch/twiki/bin/view/LHCb/Sim09Differences`. Accessed: 08/06/2023.

[191] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.

[192] A. M. Tully, *Doubly charmed B decays with the LHCb experiment*, PhD thesis, 2019.

[193] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis with ROOT*, PoS **ACAT** (2007) 040.

[194] P. Koppenburg, *Statistical biases in measurements with multiple candidates*, `arXiv:1703.01128`.

[195] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[196] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP observables in $B^\pm \to D^{(*)}K^\pm$ and $B^\pm \to D^{(*)}\pi^\pm$ decays*, Phys. Lett. **B777** (2018) 16, `arXiv:1708.06370`.

[197] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of $B^0 \to D^{*-}D_s^{*+}$ with $D_s^{*+} \to D_s^+\gamma$ decays*, JHEP **06** (2021) 177, `arXiv:2105.02596`.

[198] G. Kramer and W. F. Palmer, *Branching ratios and CP asymmetries in the decay $B \to VV$*, Phys. Rev. D **45** (1992) 193.

[199] K. Cranmer, *Kernel estimation in high-energy physics*, Computer Physics Communications **136** (2001) 198, `arXiv:hep-ex/0011057`.

[200] Belle, B. Kronenbitter *et al.*, *First observation of CP violation and improved measurement of the branching fraction and polarization of $B^0 \to D^{*+}D^{*-}$ decays*, Phys. Rev. D **86** (2012) 071103, `arXiv:1207.5611`.

[201] LHCb collaboration, R. Aaij *et al.*, *Observation of the $B_s^0 \to D^{*+}D^{*-}$ decay*, arXiv:2210.14945, to appear in JHEP.

[202] LHCb collaboration, R. Aaij *et al.*, *Observation of the $B_s^0 \to D^{*\pm}D^{\mp}$ decay*, JHEP **03** (2021) 099, arXiv:2012.11341.

[203] H.-X. Chen *et al.*, *A review of the open charm and open bottom systems*, Rept. Prog. Phys. **80** (2017), no. 7 076201, arXiv:1609.08928.

[204] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, *Handling uncertainties in background shapes: the discrete profiling method*, JINST **10** (2015), no. 04 P04015, arXiv:1408.6865.

[205] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

[206] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. **G28** (2002) 2693.

[207] M. Kenzie *et al.*, *GammaCombo: A statistical analysis framework for combining measurements, fitting datasets and producing confidence intervals*, doi: 10.5281/zenodo.3371421.

[208] LHCb collaboration, R. Aaij *et al.*, *Study of $B^- \to DK^-\pi^+\pi^-$ and $B^- \to D\pi^-\pi^+\pi^-$ decays and determination of the CKM angle γ*, Phys. Rev. **D92** (2015) 112005, arXiv:1505.07044.

[209] A. Valassi and R. Chierici, *Information and treatment of unknown correlations in the combination of measurements using the BLUE method*, Eur. Phys. J. C **74** (2014) 2717, arXiv:1307.4003.

[210] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. C **17** (2000) 137, arXiv:hep-ph/0005110.

[211] LHCb collaboration, R. Aaij *et al.*, *Measurement of B^0, B_s^0, B^+ and Λ_b^0 production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.

[212] LHCb collaboration, R. Aaij *et al.*, *Direct CP violation in charmless three-body decays of $B^\pm$ mesons*, arXiv:2206.07622, to appear in Phys. Rev. D.

[213] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^\pm$ production asymmetry and the CP asymmetry in $B^\pm \to J/\psi K^\pm$ decays*, Phys. Rev. **D95** (2017) 052005, arXiv:1701.05501.

[214] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays*, JHEP **07** (2014) 041, arXiv:1405.2797.

[215] A. Davis *et al.*, *Measurement of the instrumental asymmetry for $K^- \pi^+$-pairs at LHCb in Run 2*, Tech. Rep. LHCb-PUB-2018-004, CERN, Geneva, 2018.

[216] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-integrated CP asymmetry in $D^0 \to K^- K^+$ decays*, arXiv:2209.03179, to appear in Phys. Rev. Lett.

[217] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $D^\pm$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B718** (2013) 902, arXiv:1210.4112.

[218] S. Stahl, *Measurement of CP asymmetry in muon-tagged $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays at LHCb*, PhD thesis, Heidelberg U., 2014, doi: 10.11588/heidok.00017241.

[219] LHCb collaboration, R. Aaij *et al.*, *Measurement of the D_s^+– D_s^- production asymmetry in 7 TeVpp collisions*, Phys. Lett. **B713** (2012) 186, arXiv:1205.0897.

[220] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.

[221] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.

[222] L. Lyons, D. Gibaut, and P. Clifford, *How to Combine Correlated Estimates of a Single Physical Quantity*, Nucl. Instrum. Meth. A **270** (1988) 110.

[223] CLEO, G. Bonvicini *et al.*, *Updated measurements of absolute D^+ and D^0 hadronic branching fractions and $\sigma(e^+e^- \to D\overline{D})$ at $E_{\rm cm} = 3774$ MeV*, Phys. Rev. D **89** (2014), no. 7 072002, arXiv:1312.6775, [Erratum: Phys.Rev.D 91, 019903 (2015)].

[224] O. Kitouni, N. Nolte, and M. Williams, *Robust and Provably Monotonic Networks*, in *35th Conference on Neural Information Processing Systems*, 11, 2021. arXiv:2112.00038.

[225] *scipy.stats.kstest*, https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kstest.html. Accessed: 10/06/2023.