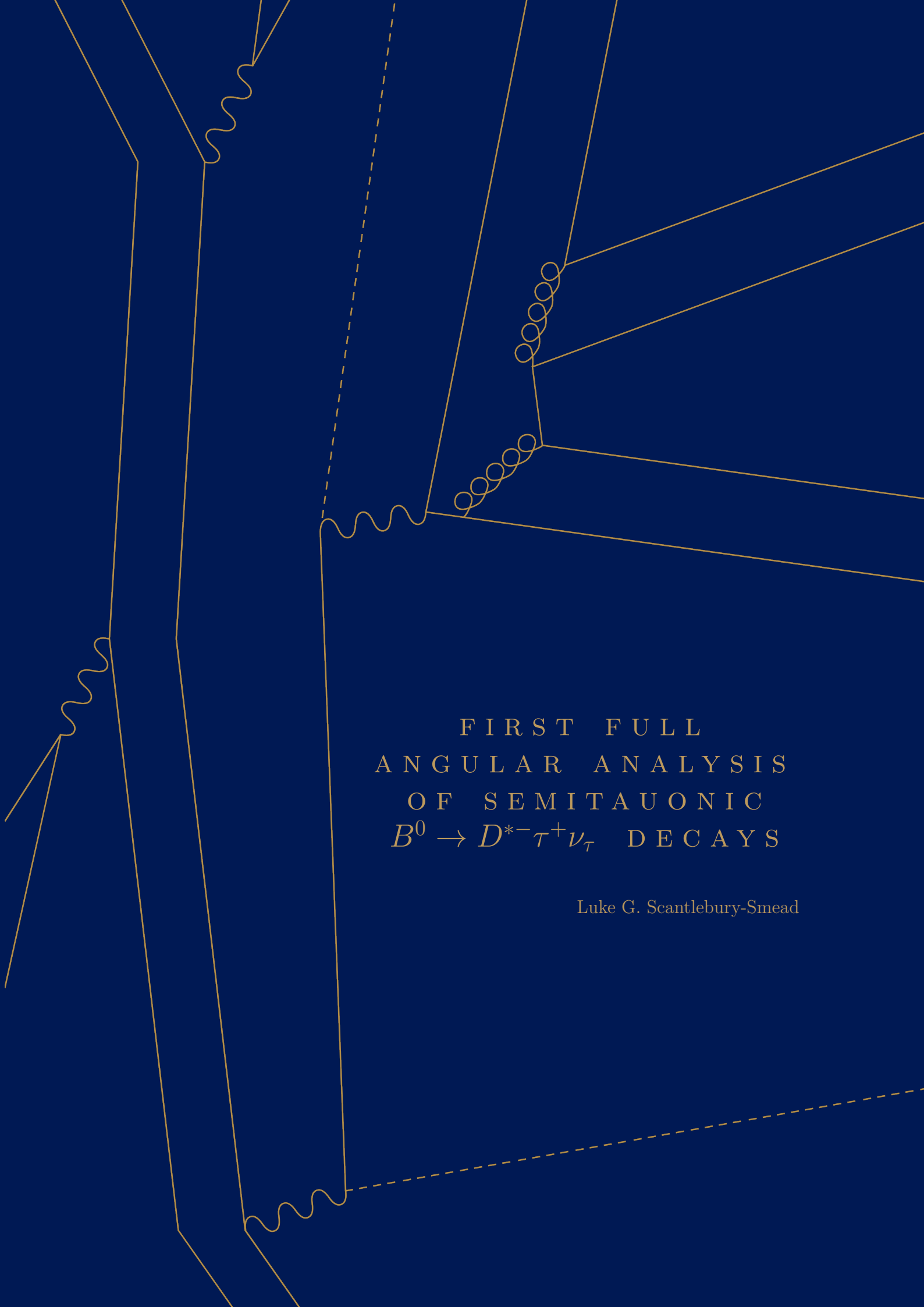


FIRST FULL ANGULAR ANALYSIS OF SEMITAUONIC $B^0 \to D^{*-}\tau^+\nu_\tau$ DECAYS

Luke G. Scantlebury-Smead

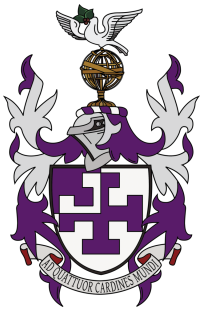
FIRST FULL ANGULAR ANALYSIS OF SEMITAUONIC $B^0 \to D^{*-}\tau^+\nu_\tau$ DECAYS



Luke G. Scantlebury-Smead

St. Cross College

University of Oxford



A thesis submitted in fulfilment of the requirements for the degree of *Doctor of Philosophy* at the University of Oxford.

External Examiner	Supervisor	Internal Examiner
Dr. Tom Blake	Dr. Malcolm John	Prof. Daniela Bortoletto
University of Warwick	University of Oxford	University of Oxford

ABSTRACT

A novel template-based technique is used to measure the q^2-integrated, rate normalised angular coefficients of the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay with the full Run 1 + 2 LHCb dataset. These coefficients capture the full angular information of the decay and have never before been measured in a comprehensive way. The technique involves creating multidimensional templates from simulated LHCb data, suitably weighted and corrected for the simulation's imperfection. The selection of the signal decay is optimised to achieve a $(28.3 \pm 1.3 \pm 0.9)\%$ signal fraction in a partially-reconstructed decay, where the first error is statistical and the second is systematic. Several data-driven control studies are performed to improve the simulation, especially for the large, double-charm backgrounds. To measure the angular coefficients, a simultaneous multi-dimensional fit is performed, utilising the 3D decay-angle distribution plus a simultaneous fit to two background-suppressing variables. The measured values of the q^2-integrated, rate normalised angular coefficients are,

$$
\begin{aligned}
\langle I_{1c} \rangle &= 0.612 \pm 0.108\ (\text{stat.}) \pm 0.060\ (\text{syst.}) \\
\langle I_{1s} \rangle &= 0.611 \pm 0.066\ (\text{stat.}) \pm 0.032\ (\text{syst.}) \\
\langle I_{2c} \rangle &= 0.123 \pm 0.251\ (\text{stat.}) \pm 0.153\ (\text{syst.}) \\
\langle I_{2s} \rangle &= 0.689 \pm 0.159\ (\text{stat.}) \pm 0.087\ (\text{syst.}) \\
\langle I_{3} \rangle &= -0.149 \pm 0.141\ (\text{stat.}) \pm 0.077\ (\text{syst.}) \\
\langle I_{4} \rangle &= 0.040 \pm 0.141\ (\text{stat.}) \pm 0.079\ (\text{syst.}) \\
\langle I_{5} \rangle &= -0.086 \pm 0.072\ (\text{stat.}) \pm 0.039\ (\text{syst.}) \\
\langle I_{6c} \rangle &= -0.158 \pm 0.191\ (\text{stat.}) \pm 0.117\ (\text{syst.}) \\
\langle I_{6s} \rangle &= -0.200 \pm 0.090\ (\text{stat.}) \pm 0.050\ (\text{syst.}) \\
\langle I_{7} \rangle &= 0.045 \pm 0.077\ (\text{stat.}) \pm 0.039\ (\text{syst.}) \\
\langle I_{8} \rangle &= 0.130 \pm 0.163\ (\text{stat.}) \pm 0.094\ (\text{syst.}) \\
\langle I_{9} \rangle &= 0.112 \pm 0.137\ (\text{stat.}) \pm 0.075\ (\text{syst.})
\end{aligned}
$$

From these measurements, the longitudinal polarisation fraction of the vector charm meson, $f_L^{D^*}$, and the forward-backward asymmetry, A_{FB}, are calculated to be,

$$f_L^{D^*} = \quad 0.428 \ \pm \ 0.036 \ \text{(stat.)} \ \pm \ 0.016 \ \text{(syst.)}$$
$$A_{\rm FB} = -0.209 \ \pm \ 0.081 \ \text{(stat.)} \ \pm \ 0.050 \ \text{(syst.)}$$

where the 9.2% relative precision on $f_L^{D^*}$ is an improvement compared with the 15% relative uncertainty achieved by previous experiments while the A_{FB} result is a world-first measurement. These measurements, in their final published form, will test the Standard Model by providing comprehensive sensitivity to new-physics effects in $b \to c\tau\overline{\nu}$ transitions.

PREFACE

This thesis describes an innovative approach that utilises templates to measure the q^2-integrated angular coefficients of the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay, utilising the complete 9 fb^{-1} Run 1 + 2 LHCb dataset. I have performed additional work during the course of the doctoral programme that is not described in this thesis. All of my work was completed during the years 2018-2023 and is as follows.

Feasibility study of $\mathcal{R}(D^{**})$ (2018 - 2019): I began by exploring the $B \to D^{(*)(*)}\tau\nu$ datasets, aiming to investigate additional decay modes beyond the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay mode. The work was focused on the $B^+ \to D^{**0}\tau^+\nu_\tau$ decay mode, and I aimed to estimate the number of $B^+ \to D^{**0}\tau^+\nu_\tau$ candidates in the full Run 1 + 2 LHCb dataset after a selection was applied. This estimation was performed by measuring the number of $B^+ \to D^{**0}D_s^+$ and $B^0 \to D^{*-}D_s^+$ events in the dataset and using the ratio of the yields and relative integrated luminosity to scale the Run 1 measured yield of $B^0 \to D^{*-}\tau^+\nu_\tau$ decays. This estimated an expected yield of around 250 events. I was awarded the Denys Wilkinson Prize for Outstanding First Year Performance for this work.

VELO microchannel component production (2019 - 2021): A new set of data has begun to be collected and will constitute the Run 3 dataset, where LHCb aims to record a factor of 2-3 times more data. This additional data will reduce the statistical uncertainty on the angular observables measured in this thesis and will provide the opportunity to understand background decays at a new level. A critical requirement for the analysis of the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay, the subject of this thesis, is superb vertexing of the B^0 and τ decays. At LHCb, the VErtex LOcator (VELO) detector provides a measurement of these vertices and has been upgraded for the Run 3 data-taking period, where the silicon pixel detectors are cooled using a novel liquid CO_2 silicon microchannel method.

I played a significant role, before and during the COVID-19 pandemic, in constructing the silicon microchannel cooling system used by the VELO to maintain the silicon detectors' temperature. Soldering the connectors that transport the liquid CO_2 to the silicon microchannel cooling plates was a significant technical challenge

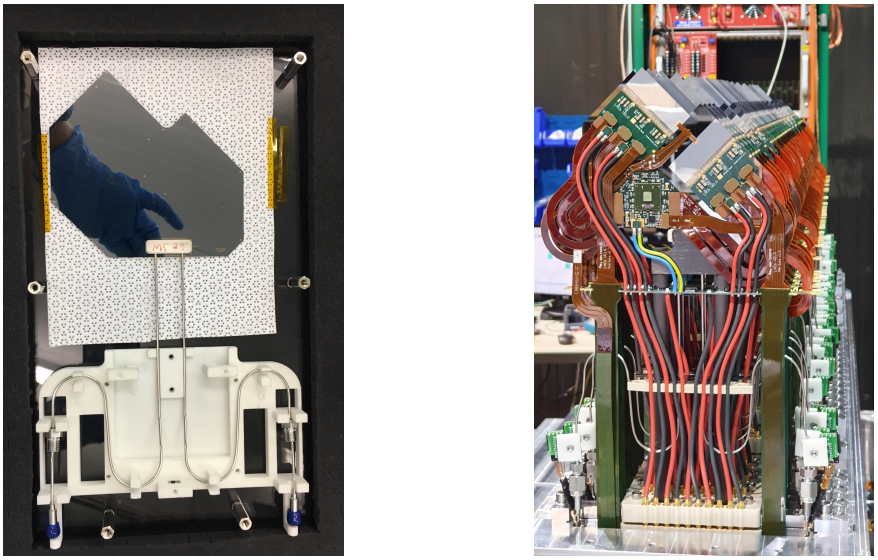


Figure 1: A CO_2 connector soldered to a silicon microchannel cooling plate, ready for additional construction steps to produce a full VELO module (left). One-half of the new VELO detector, which is to be used in Run 3 data taking (right).

with many intermediate steps. I contributed several months during my time at Oxford and CERN, often moving between the two locations as the work demanded, performing several necessary functions for the VELO microchannel cooling system production. I performed quality assurance for all parts, prepared parts for mechanical or chemical processing steps, performed some of those steps personally, and then ensured the resulting cooling system met the stringent criteria to be used in the LHCb detector. Each constructed microchannel cooling plate constituted tens of hours of labour. A constructed cooling system, composed of a CO_2 connector soldered to a silicon microchannel cooling plate can be seen in Figure 1 as well as one of the fully constructed VELO halves, ready to be installed in the LHCb cavern.

VELO slow controls and monitoring (2021 - 2022): I developed a temperature monitoring system for the VELO during Run 3 operations. This monitoring involves over 1200 temperature readings from different parts of the VELO and different parts of the individual silicon detector sensors. The temperature monitoring was used daily by experts during the VELO commissioning and installation phases and it continues to be used by VELO shift experts during normal operation. This system is composed of a main monitoring panel which can be seen in Figure 2, however, every button on the main panel creates a new window with more specialised,

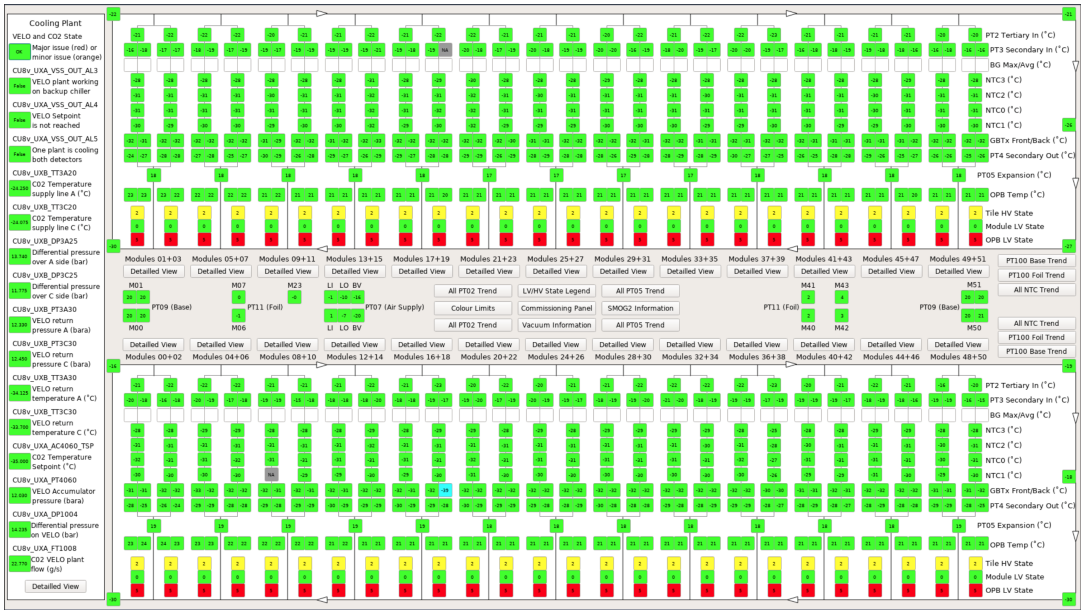


Figure 2: The main temperature monitoring panel during normal operation, when the VELO is cooled to -30°C.

detailed, information. Additionally, each square with colour can be pressed to create a plot of the temperature of that sensor over time.

Data driven background studies for the angular analysis (2019 - 2023): This work involves the study of several background modes that are present in the final angular signal fit and are described in Chapter 7. The studies of $B \to D^{*-}D^{0}X$, $B \to D^{*-}D^{+}X$, and $B \to D^{*-}D_{s}^{+}X$ were initially performed by Anthony Correia during the period of 2020-2021, but have since been developed further and the datasets have changed as the event selection evolved. Additional control studies have also been performed, for the $D_{s}^{+} \to \pi^{+}\pi^{-}\pi^{+}X$, $B \to D^{*-}\pi^{+}\pi^{-}\pi^{+}$, and $B \to D^{*-}\pi^{+}\pi^{-}\pi^{+}X$ datasets. This work is crucial in order to measure and understand the backgrounds in the signal fit, derive corrections for the imperfect simulation samples, and provide confidence in the background shapes used in the signal fit.

First implementation of the angular analysis (2019-2023): The implementation of the angular analysis was based on a paper authored by Dr Donal Hill, a post-doctoral researcher who developed the template-based angular analysis technique and worked closely with me throughout the analysis. The first implementation of the signal fit in this analysis was developed by Dr Hill and followed the paper

methodology (referenced in this thesis), where a four-dimensional fit is performed in the three decay angles and a fourth discriminating variable. A significant amount of time was taken to develop this fit beyond this point, going through many iterations with continual feedback from the background studies. The signal fit methodology has ultimately been altered to split the fit into a simultaneous multidimensional fit, where a 3D angular fit is performed simultaneously with a 2D fit using discriminating variables. A comprehensive assessment of the systematic uncertainties associated with the fit was performed.

Current status (2023): The analysis presented in this thesis is currently under review by the LHCb collaboration. As part of the internal procedure, the analysis is blind and the results presented here are therefore preliminary and unknown to those that continue this work towards publication. It is expected that this work will be expanded to include a measurement of $\mathcal{R}(D^*)$ which is not presented in this thesis. This will require external branching fraction measurements and the use of simulation to measure the efficiencies of decay modes. The use of the simulation in this way means the measurement of $\mathcal{R}(D^*)$ will not be strictly model-independent. The implementation of the method presented in this thesis and the techniques used to understand the background will remain the core of this examination of semitauonic decays.

This thesis describes a new analysis of data collected by the LHCb collaboration during 2011-2018. The analysis was effected in close partnership with the post-doctoral researcher Dr Dónal Hill, but the thesis itself is produced solely by the author. The thesis has not been submitted for another qualification to this or any other university.

ACKNOWLEDGEMENTS

I would like to extend my deepest appreciation to my supervisor, Malcolm John, for his invaluable guidance, support, and encouragement throughout my research journey. From the early stages of my first-year project to the final stages of my thesis, his expertise, constructive feedback, and commitment to excellence have been instrumental in shaping my academic and professional growth. Malcolm has challenged me to think more deeply, critically, and creatively about my research and has helped me to develop the skills and confidence necessary to succeed. I am grateful for his mentorship and for the countless hours he has spent supporting me and guiding me towards academic excellence.

I would also like to express my sincere appreciation for the contributions made by Dónal Hill, who has been a friend and mentor to me throughout my research journey. His unwavering support, valuable insights, and thoughtful guidance have been critical in helping me to refine my research and to develop a deeper understanding of particle physics. I am grateful for the many enriching conversations and ideas we have shared. Beyond his academic expertise, Dónal has also provided me with invaluable personal and professional guidance. His kindness, empathy, and genuine interest in my well-being have made him not only a colleague but also a trusted friend. I am grateful for his support, encouragement, and friendship, and I feel privileged to have had him as a mentor throughout my research journey.

I am deeply indebted to the academic community at the University of Oxford, which includes not only my fellow doctoral students but also those who came before me and those who continue their academic pursuits. From my earliest days at the university, the senior students generously imparted their knowledge and expertise, quickly orienting me to the office culture of camaraderie and mutual support. Meanwhile, my peers in the doctoral program have been invaluable partners in learning; we have shared experiences that will remain unforgettable and cherished long after our studies have ended. Their company, laced with moments of laughter and joy, has enriched my academic and personal growth in countless ways. I am deeply grateful for the opportunity to have been a part of this vibrant community and

for the friendships that I have formed along the way. In particular, I would like to thank Jenny and Rizwaan for all the wonderful memories and their incredible support during my doctorate. Cycling hundreds of kilometres across Switzerland, countless conversations in the office, and many, many more memories, I want to thank you both for your friendship.

I am grateful to have had the opportunity to work alongside the distinguished VELO group and to have contributed to their remarkable upgrade project. The challenges posed by the technical aspects of this endeavour were intellectually stimulating, and the sense of camaraderie among the team members was invigorating. I extend my heartfelt thanks to Wiktor Byczyński, whose technical proficiency and patience were invaluable in my development. His expertise was imparted with meticulous attention to detail, and I am immensely grateful for his guidance. I also wish to express my appreciation to Viktor Coco and Paula Collins, whose mentorship and support helped me strive for excellence, and always fostered an environment of friendship within the VELO group.

I must acknowledge that there were many other individuals who contributed to my journey in ways that I cannot fully summarize here. I would like to extend my heartfelt thanks to my LTA cohort for the incredible experiences we shared together during our time in Switzerland, exploring so much of the country and experiencing the diverse culture. From engaging discussions to breathtaking hikes in the Alps, every moment spent with this group was characterized by a spirit of friendship and inclusiveness. The diverse perspectives, insights, and experiences that each person brought enriched my learning and inspired me to broaden my horizons. I am deeply grateful for the support and encouragement of this exceptional group of individuals, and I will always look back on this journey with fondness and appreciation.

I would like to express my heartfelt appreciation to my parents, family, and friends for the unwavering support, love, and inspiration that they have provided me throughout my doctoral journey. My parents instilled in me a love of science from an early age, and their encouragement and guidance have been instrumental in shaping my academic pursuits. Their unflagging support has enabled me to pursue my passion for research and reach this important milestone in my academic journey.

I hope I have made them proud. My partner Alice, too, deserves my sincerest thanks for embracing the adventure that this doctorate has been. Her love and encouragement have been a constant source of inspiration, and I could not have reached this point without her support. Finally, I would like to thank all of my friends who have been an essential part of my life throughout my doctorate. Your enthusiasm for my research, encouragement, and friendship have been vital, and I am grateful for the countless memories and experiences that we have shared.

CONTENTS

I PROLOGUE 1

1 INTRODUCTION 2

1.1 The Standard Model . . . 2

1.2 Lepton Flavour Universality . . . 3

2 THE LHCB EXPERIMENT 7

2.1 The Tracking System . . . 9

2.2 The RICH Detectors . . . 11

2.3 The Calorimeters . . . 12

2.4 The Muon Detectors . . . 13

2.5 The LHCb Trigger and Data Acquisition System . . . 13

3 SEMILEPTONIC DECAYS 15

3.1 Angular Formalism . . . 17

4 METHODOLOGY OF THE ANGULAR MEASUREMENT 29

4.1 The Signal . . . 29

4.2 Kinematic Reconstruction . . . 30

4.3 Decay Angle Reconstruction . . . 33

4.4 Bias in $\cos\theta_L$ . . . 36

4.5 Angular Coefficient Parameterisation . . . 38

4.6 Backgrounds . . . 39

II LHCB DATA ANALYSIS 42

5 EVENT SELECTION 43

5.1 Geometrical Acceptance . . . 45

5.2 Trigger . . . 45

5.3 GRID-Processing Preselection . . . 47

5.4 Local-Processing Preselection . . . 49

5.5 Multivariate Analysis . . . 51

5.6 Particle Identification Selection . . . 58

5.7 Isolating Genuine D^* Candidates . . . 59
5.8 Normalisation Selection . . . 70
5.9 Efficiency . . . 70

6 NORMALISATION MEASUREMENT 71
6.1 $m(\pi^+\pi^-\pi^+)$ D_s Peak Fit . . . 71
6.2 $m(D^{*-}\pi^+\pi^-\pi^+)$ Fit with ${}_s$Weights . . . 73
6.3 Full $m(D^{*-}\pi^+\pi^-\pi^+)$ Fits to Simulation . . . 73
6.4 Full $m(D^{*-}\pi^+\pi^-\pi^+)$ Fit . . . 75

7 DATA-DRIVEN BACKGROUND CONTROL STUDIES 77
7.1 D_s^+ Decay Study . . . 78
7.2 $B \to D^{*-}D_s^+(X)$ control mode . . . 85
7.3 $B \to D^{*-}D^0(X)$ and $B \to D^{*-}D^+(X)$ Control Modes . . . 100
7.4 $B \to D^{*-}3\pi$ Control Mode for L0 Reweighting . . . 112
7.5 Comparison of $b\bar{b} \to D^*3\pi X$ inclusive simulation and data . . . 113

III ANGULAR COEFFICIENT MEASUREMENT 118

8 ANGULAR ANALYSIS 119
8.1 Signal Model Used in Simulation . . . 120
8.2 Signal Template Creation . . . 123
8.3 Angular Fit Components . . . 138
8.4 Fit to Data . . . 146
8.5 Parameter Uncertainty . . . 151
8.6 Pseudo-Experiment Fits . . . 156
8.7 Numerical Results . . . 163

9 SYSTEMATICS 166
9.1 Fit Based Systematics . . . 166
9.2 Summary of Systematic Uncertainties . . . 173

10 CONCLUSION 177
10.1 Final Results . . . 177
10.2 Interpretation . . . 178
10.3 Final word . . . 181

BIBLIOGRAPHY 182

IV APPENDIX 190

A HELICITY AMPLITUDES 191

A.1 Leptonic Helicity Amplitudes 191

A.2 Hadronic Helicity Amplitudes 193

A.3 Angular Coefficient Definitions 195

B SUBDECAY CATEGORIES 196

Part I

PROLOGUE

1 INTRODUCTION

1.1 THE STANDARD MODEL

The Standard Model (SM) is the most robust and thoroughly tested theory of particle physics, unifying elementary particles and forces, and is the result of decades of development and discovery [1–8]. While many particle physics experiments have confirmed its predictions, the SM is incomplete as it does not explain the asymmetry of matter and antimatter, gravity, or particle candidates for dark matter and dark energy.

The SM unites fundamental particles and three forces: weak, strong, and electromagnetic. Elementary particles fall into two categories: fermions and bosons, with spin being half-integer or integer, respectively. Bosons act as mediators and include gluons for the strong force, $W^{\pm}$, Z^0 for weak interaction, and photons for electromagnetic interaction. The Higgs boson provides mass to particles. Fermions divide into quarks, which interact via electroweak and strong forces, and leptons, which only interact through the electroweak force. There are three generations of fermions, with increasing masses. Quarks include up, down, charm, strange, top, and bottom. Similarly, the classification of leptons follows the three families of a lepton and its corresponding neutrino. The e, μ and τ have charge e while the neutrinos have no charge.

Electroweak interactions can be described with a gauge theory with the symmetry group SU(2)×U(1), where the SU(2) group has generators that form three gauge bosons W_{μ}^{k}, their coupling g, and the quantum number T, weak isospin. The U(1) group results in the gauge boson B_{μ}, its coupling g', and quantum number Y, hypercharge. These four gauge bosons are massless, however, the SU(2)×U(1) symmetry is spontaneously broken by the Higgs mechanism and gives masses to the weak mediators [4–6]. The physical $W^{\pm}$, Z^0, and γ bosons are linear combinations of the gauge bosons W_{μ}^{k} and B_{μ}.

1.2 LEPTON FLAVOUR UNIVERSALITY

In the Standard Model the couplings of the electroweak bosons, g and g', to the leptons are independent of the flavour of the lepton. Leptons do not couple to the mediators of the strong force, gluons. The remaining gauge boson that couples to the leptons is the Higgs, which does have flavour-dependent couplings in the form of the Yukawa couplings. These statements are combined to form the concept of Lepton Flavour Universality. Couplings to leptons that are dependent on the lepton flavour, beyond the Yukawa couplings, would be extensions to the Standard Model and evidence of this in an experiment would be evidence of Beyond the Standard Model (BSM) physics.

Since the electroweak coupling to each of the three charged leptons (e, μ, and τ) is identical, the transition rates involving these leptons are expected to be the same up to phase-space effects related to the lepton mass. SM predictions for the branching fraction ratios of semitauonic B-meson decays [1] relative to decays involving a lighter lepton,

$$\mathcal{R}(D^{(*)}) \equiv \frac{\mathcal{B}(B^{0,+} \to D^{(*)-,0}\,\tau^+\,\nu_\tau)}{\mathcal{B}(B^{0,+} \to D^{(*)-,0}\,\mu^+\,\nu_\mu)}\,, \tag{1.1}$$

are known to percent-level precision [9], since uncertainties due to hadronic effects cancel to a large extent. Measurements of such ratios have been performed and provide tests of the SM and offer sensitivity to Lepton Flavour Universality Violation (LFUV), which would be an unambiguous sign of BSM. A key objective in flavour physics is to make measurements with uncertainties that are comparable to that of theoretical predictions in order to look for discrepancies. In $\mathcal{R}(D^{(*)})$ measurements, BSM theories with flavour-dependent couplings, such as models with an extended Higgs sector [10], leptoquarks [11], or an extended gauge sector [12], have the potential to explain discrepancies that may be observed between theory and experiment, as shown in Figure 1.1.

The first measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay was made by the Belle collaboration in 2007 [13]. Both the Belle and BaBar collaborations have published measurements of the branching fractions $\mathcal{B}(B^+ \to \overline{D}^0\tau^+\nu_\tau)$ and $\mathcal{B}(B^+ \to \overline{D}^{*0}\tau^+\nu_\tau)$

[1] The inclusion of charge-conjugate processes is implied throughout this thesis.

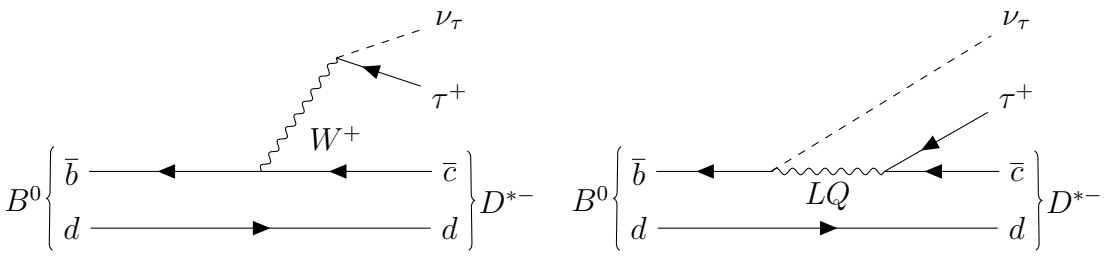


Figure 1.1: Feynman diagrams showing the Standard Model $b \to c\ell\nu$ process (left) and a possible BSM leptoquark process (right).

[14, 15]. Potential discrepancies in $\mathcal{R}(D^{(*)})$-type measurements began when the BaBar collaboration published results of $\mathcal{R}(D^*)$ that were in tension with the SM expectations up to the level of 3.4σ [16]. The Belle collaboration has also published results for $\mathcal{R}(D^*)$, which are compatible with the SM to within 1σ [17]. The BaBar and Belle measurements were performed using the e^+e^- colliders SLAC and KEKB respectively, which operated at the energy of the $\Upsilon(4S)$ resonance (10.58 GeV) which decays into pairs of $B^0\ \overline{B}^0$ or $B^+\ B^-$ mesons. The events were tagged by fully reconstructing the decay of one of the b-hadrons to a hadronic final state, while the other b decay was used to search for the signal.

The LHCb collaboration first measured $\mathcal{R}(D^*)$ in 2015 with muonic tau decays [18] and found results compatible with those from BaBar. A simultaneous measurement of $\mathcal{R}(D^*)$ and of the τ polarisation, using hadronic tagging and reconstruction of the $\tau^- \to \pi^-\overline{\nu}_\tau$ and $\tau^- \to \rho^-\overline{\nu}_\tau$ decays, was published by the Belle Collaboration [19]. The leptonic decay of the τ lepton to a μ or an e was used in these measurements, while a measurement of $\mathcal{R}(D^{*-})$ from the LHCb collaboration [20] was performed using the three-prong τ decays, $\tau^+ \to \pi^+\pi^-\pi^+\,\overline{\nu}_\tau$ and $\tau^+ \to \pi^+\pi^-\pi^+\pi^0\,\overline{\nu}_\tau$. This measurement was reported to be 1.1σ higher than the SM prediction. More recently, the Belle collaboration published measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with events tagged using semileptonic decays which were compatible with the SM at 0.2σ and 1.1σ respectively [21]. In 2022, LHCb presented the simultaneous measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, which was the first time LHCb has performed a measurement of this type and used the muonic τ decay. The value of $\mathcal{R}(D^*)$ was found to be $0.281 \pm 0.018 \pm 0.024$, and the value of $\mathcal{R}(D)$ was found to be

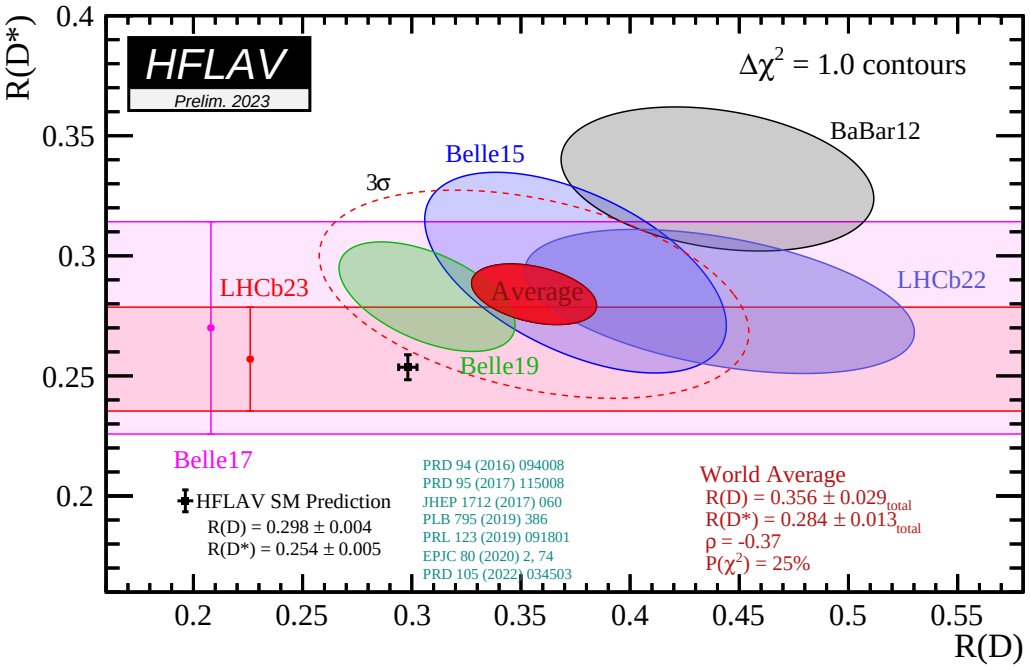


Figure 1.2: The most recent global fit of measurements of $\mathcal{R}(D)$ vs $\mathcal{R}(D^*)$ [24] including the soon to be published latest results from LHCb [22,23] The combination of these results is in tension with calculations of the SM prediction at a significance of 3.1σ [24].

$0.441 \pm 0.060 \pm 0.066$. These values are compatible with the current world average and the SM predictions to within 2σ [22]. An updated measurement of $\mathcal{R}(D^*)$ using the hadronic τ decay and the data collected in 2015 and 2016 has been presented by LHCb. The value of $\mathcal{R}(D^*)$ found was $0.247 \pm 0.015 \pm 0.015 \pm 0.012$ where the first error is statistical, the second systematic, and the third systematic due to external branching fraction inputs [23]. When this result is combined with the Run 1 result the value of $\mathcal{R}(D^*)$ was $0.257 \pm 0.012 \pm 0.014 \pm 0.012$ [23]. The combination of these results, as shown in Figure 1.2, is in tension with calculations of the SM prediction at a significance of 3.1σ [24].

Although decay rate measurements are useful tools for detecting LFUV, it is necessary to study the kinematic and angular properties of these semileptonic decays in order to fully characterise the $b \to c\ell\nu$ transition. Many polarisation and asymmetry observables have been shown to discriminate between the SM and BSM scenarios [25–27]. A measurement of the D^* longitudinal polarisation fraction performed by Belle [28] found $F_L^{D^*} = 0.60 \pm 0.09$ which is in agreement within 2σ with the the SM predicted value, $F_L^{D^*} = 0.462 \pm 0.005$, derived from the predicted values

of other angular observables [29]. In order to obtain complete kinematic information on the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay, the full angular decay rate must be studied [29].

This thesis presents the first measurement of the q^2-integrated angular coefficients for the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay. The measurement utilises the application of a technique developed for the angular analysis of partially reconstructed decays, where histogram templates in the angular variable space are derived from simulation and are used in the fit to encapsulate the resolution and reconstruction effects due to missing momentum [30]. The signal decay is reconstructed using a six-track final state; the τ lepton is reconstructed in the three-prong $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\,\overline{\nu}_\tau$ modes (the π^0 is not reconstructed), while the D^{*-} meson is reconstructed via $D^{*-} \to D^0\,(\to K^+\pi^-)\,\pi^-$. The LHC and the LHCb detector are described in Chapter 2. The theory of the Standard Model as it pertains to semileptonic decays and the angular description of the signal decay is presented in Chapter 3. The analysis methodology for the measurement of the angular coefficients is described in Chapter 4, including the consideration of the backgrounds in the analysis. The event selection used for the signal and normalisation fits is discussed in Chapter 5. In Chapter 6 the normalisation measurement is described. Studies of the largest backgrounds in the analysis are reported in Chapter 7. The procedure for the creation of angular templates and details of the signal measurement is given in Chapter 8. Systematic uncertainties for the signal measurement are reported in Chapter 9. Finally, the results and their interpretation are given in Chapter 10.

2 THE LHCB EXPERIMENT

The Large Hadron Collider (LHC) has a circumference of 27 km and was designed to use a centre-of-mass energy of $\sqrt{s} = 14$ TeV and luminosity of 10^{34} cm^{-2}s^{-1}. The ring is composed of 1232 dipole magnets, which bend the beams, and 392 quadrupole magnets which focus the beams. There are also 16 radio-frequency cavities, which accelerate the protons to close to the speed of light.

The LHC is composed of eight sections, each section has a straight region, a bending region, and an interaction region (IR). IR1 houses one of the two high luminosity experiments at CERN, the ATLAS (A Toroidal LHC ApparatuS) experiment [31]. ATLAS is a general-purpose detector designed to observe processes that involve massive particles which were not observable using lower-energy accelerators. The ALICE (A Large Ion Collider Experiment) experiment is situated at IR2 and is designed for the study of quark-gluon plasma produced in heavy ion collisions [32]. The second high luminosity experiment, CMS (Compact Muon Solenoid), is located at IR5. The CMS experiment is a general-purpose detector like ATLAS, with a wide range of physics goals, including measurements of Higgs boson properties [33]. The final interaction region, IR8, houses the LHCb experiment.

The LHCb experiment is a single-arm forward spectrometer with a design chosen to exploit the high cross-section of b quark pairs in the forward and backward directions [34]. LHCb is a precision experiment, its goals require the accurate reconstruction of the full decay chain of any b or c hadron. The operating luminosity of the LHC results in substantially more than one collision per bunch crossing and a track multiplicity that is too high to reliably reconstruct b and c hadron decays. To prevent this problem, LHCb uses a reduced beam luminosity which is achieved by defocusing the beams at the interaction point (IP) [35]. This ensures that the collected data are predominately events with one interaction per bunch crossing and it is expected that 10^{12} $b\bar{b}$ pairs would be produced in 10^7s, corresponding to one year of data taking.

In order to perform a measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay, several criteria must be met. The τ lepton must be accurately reconstructed (allowing for the

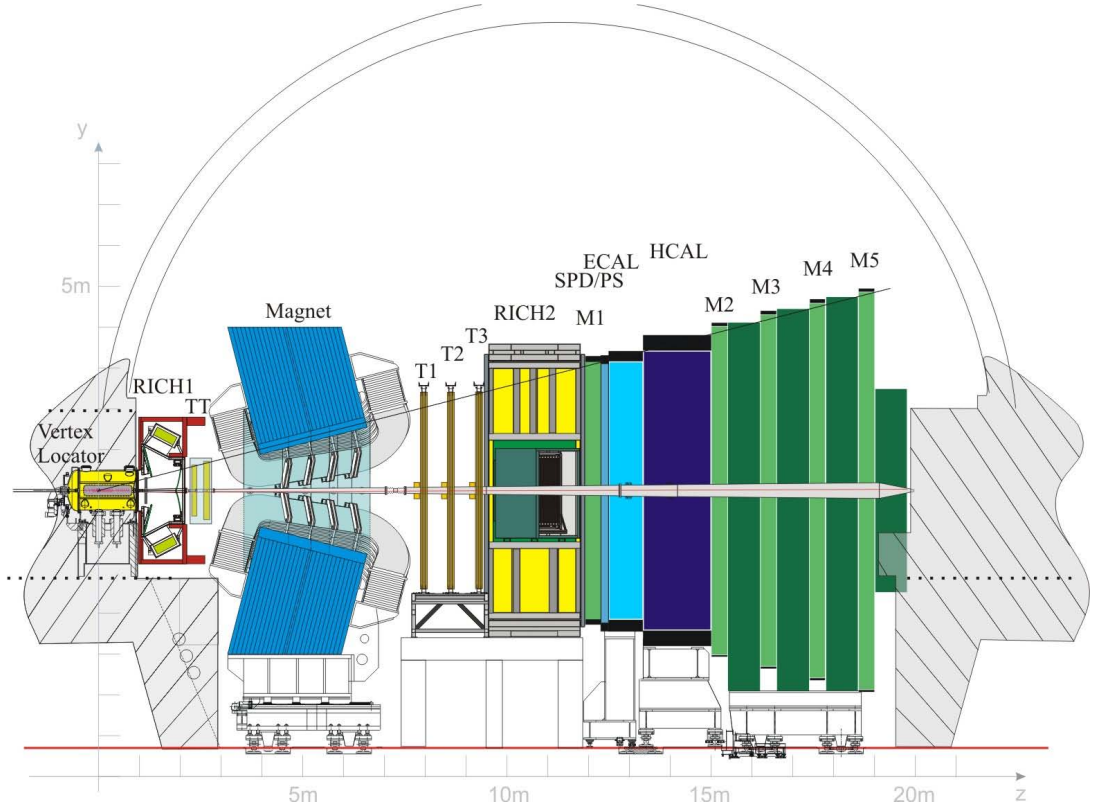


Figure 2.1: Schematic of the LHCb detector viewed in the non-bending plane [36]. The tracking system is composed of the VELO, TT, and T1-3 subdetectors. The particle identification detectors, RICH1 and RICH2, are situated amongst the tracking detectors. Energy measurements are performed by the calorimeters, ECAL and HCAL. Finally, the muon stations, M1-5, detect muons that pass through the previous subdetectors.

limitations due to the missing neutrino in the τ decay) and the vertex determined with high accuracy. It is also essential to identify the three pions in the τ decay, as well as the pions and kaon in the $D^{*-} \to D^0(\to K^+\pi^-)\pi^-$ decay. The decay angles, which form the cornerstone of the angular decay rate, are determined from the four-momenta of the decay particles; the tracking of particles and the momentum measurements are crucial in this calculation, as well as the energy measurements. Backgrounds with additional particles are rejected using the information from the tracking and calorimeter detectors. The LHCb detector satisfies these requirements, where each sub-detector contributes an essential aspect of these criteria.

The layout of the LHCb detector is shown in Figure 2.1. There are several sub-detectors of the experiment; the VErtex LOcator (VELO) is situated at the IP and is the first detector particles encounter. The first Ring Imaging CHerenkov counter (RICH1) is located downstream of the VELO, followed by the Tracker Turicensis

(TT). A dipole magnet then deflects charged particle trajectories before they enter three tracking detectors. Particles then encounter the second RICH detector (RICH2) before the first of the muon detection stations (M1). Following M1 is the calorimeter system composed of a scintillator pad detector (SPD) and Preshower (PS), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). Finally, the remaining muon detection systems are situated behind the HCAL.

2.1 THE TRACKING SYSTEM

The LHCb tracking system comprises the VELO and four tracking stations, with the VELO and TT upstream of the dipole magnet and the T1-3 stations downstream. The LHCb experiment requires vertex reconstruction to identify secondary vertices in b and c hadron decays. The VELO was designed to detect charged particles and reconstruct their paths close to the interaction region [37]. The VELO has two sets of 21 detector modules, each with an R-sensor and a ϕ-sensor. The VELO halves are retractable to protect against radiation damage; when the beam spread is large and the orbit unstable, the two VELO halves are separated by 6cm. In normal operation, the detector modules must be close to the beam, leaving a 7mm hole for the beam to pass through with a precision better than 4μm, achieved through measuring the beam position with the VELO while the two halves are brought together [38].

The VELO achieved hit resolutions of 4μm to 20μm, and a vertex resolution of 13μm in the transverse plane and 71μm along the beam axis for vertices with 25 tracks [38]. Impact parameter resolution was measured to be better than 35μm for particles with transverse momentum of $> 1\,\text{GeV/c}$. At lower momentum, the resolution suffers due to significant multiple scattering in detector material at lower momentum tracks.

In the signal decay, the vertices are constructed from significantly fewer than 25 tracks; the τ vertex has three tracks while the B^0 is ultimately reconstructed from the reconstructed τ and D^* totalling six tracks. In reconstructed simulation, the B^0 flight distance from the primary vertex in the z direction has an average of 23mm which is measured with an average uncertainty of 1.0mm. The B^0 can be separated

from the PV in the z direction with an average significance of 32σ. In the case of the τ vertex, the average flight in the z direction from the B^0 vertex is 3.5mm measured with an average uncertainty of 1.3mm. This results in an average separation with the B^0 vertex of 3.3σ significance. The excellent vertexing of the B^0 and the τ is crucial in the analysis since the flight distance significance of the 3π system (the τ in the signal case) is a powerful discriminating variable, used to remove extremely large backgrounds (as discussed in Section 5.4) and is also used in the signal fit to data.

The LHCb detector uses a water-cooled dipole magnet to provide momentum information. The magnet consists of two saddle-shaped coils weighing 27 tonnes each, contained within a 1500 tonne yoke, and is sloped to match the detector's angular acceptance. The magnet's magnetic field is in the y-direction, which concentrates the Lorentz force in the horizontal (x, z) plane and bends tracks in this plane. The dipole magnet provides momentum information by inducing curvature in charged particle trajectories through its magnetic field. The ability to invert the magnetic field in the LHCb detector is crucial for achieving one of its main goals: measuring CP violation in b or c hadrons [39, 40]. Systematic effects from detector asymmetry can have a significant impact on the results, but by regularly inverting the magnetic field, these effects can be reversed, measured, and corrected.

The T1-3 stations are composed of two parts: a silicon micro-strip section close to the beamline known as the Inner Tracker (IT) and a straw-tube detector section in the outer region called the Outer Tracker (OT). The IT and TT sub-detectors form the Silicon Tracker system. The straw-tube detectors in the outer region are used since silicon sensors are not needed in these lower occupancy regions. The purpose of the tracking stations is to provide high-precision measurements of charged particle trajectories, before and after the dipole magnet. The momentum of the particles can be calculated from the bending caused by the magnetic field using $p = qBr$, where q, B, and r are the electric charge, magnetic field, and radius of curvature, respectively. Track reconstruction algorithms search for straight line segments in the VELO, TT, and T1-3, with the highest quality candidates chosen based on the χ^2 of the track matching. The spatial resolution of the Silicon Tracker system was

found to be about 50 μm. The OT uses drift tubes that have an efficiency of 99.2% to detect hits [41]. The tubes provide momentum measurements with a resolution of $\sim 0.4\%$ and mass reconstruction of b-hadrons within 10 MeV/c^2. In Run 2, the performance of the OT was improved by 20% through the use of a novel alignment procedure [42].

As previously described, the decay angles in this analysis are determined from the four-momenta of the particles in the decay; high resolution three-momenta measurements are therefore essential. Due to the missing neutrinos at the τ and B^0 decay vertices, the reconstruction of the momenta is not perfect and reconstruction effects are introduced, this is discussed in Section 4.2. In contrast, the particles produced in the fully reconstructed D^* and D^0 decay vertices are measured with high accuracy, and the momenta are measured with high precision. This allows for high quality reconstruction of the D^* and D^0 particles, which can be seen in Section 5.7. High efficiency for hit detection is also essential: there are six final state particles in the decay meaning that any inefficiency in track detection is compounded to the sixth power.

2.2 THE RICH DETECTORS

Charged particles emit Cherenkov radiation when they travel through a medium faster than the speed of light in that medium. The angle of this radiation relative to the particle's direction of motion is related to its velocity, with $\cos(\theta) = \frac{1}{\beta n}$, where $\beta = v/c$, c is the speed of light in a vacuum, and n is the refractive index of the medium. RICH detectors use spherical mirrors to focus this radiation into rings on the detection plane. By measuring the ring radius which is proportional to the Cherenkov angle, the velocity of the particle can be determined and combined with momentum information from tracking detectors to identify the particle. LHCb uses two RICH detectors: RICH1 covers 1-60 GeV/c momentum range using aerogel and C4F10 radiators, and RICH2 covers 15-100 GeV/c momentum range using a CF4 radiator [39]. Averaging over the momentum range 2 - 100 GeV/c, using the criteria that the delta log-likelihood (DDL)$(K - \pi) > 0$, the kaon efficiency and pion misidentification fraction are found to be $\sim$ 95 % and $\sim$ 10 %, respectively.

Altering the criteria to DDL$(K-\pi) > 5$ resulted in a reduced kaon efficiency of 85 % and a reduced pion misidentification fraction of 3% during Run 1 operation [43]. During Run 2 operation, it was found that the probability to correctly identify each charged hadron is consistently above 90% for a 5% misidentification probability [44].

The ability to separate kaons and pions in b-hadron decays is an important requirement for the LHCb experiment [45]. This separation is essential in this analysis, with a particular focus on the three tracks from the τ vertex. The largest background in the signal fit comes from decays of the type $B \to D^* D_s X$, where the $D_s X$ system is responsible for three or more tracks. D_s decays involving kaons are more likely than the corresponding decays involving only pions. For example, the $D_s \to K^+K^-\pi^+$ decay has a branching fraction of $(5.38 \pm 0.10)\%$ while the $D_s \to \pi^+\pi^-\pi^+$ decay has a branching fraction of $(1.08 \pm 0.04)\%$ [46]. By placing particle identification (PID) requirements on the three tracks in the 3π system the decays involving kaons can be reduced significantly, despite being more common than the decay involving pions. Once the 3π system has been subjected to these PID requirements, additional selection steps can be taken to reduce backgrounds involving pions.

2.3 THE CALORIMETERS

The LHCb calorimeter system provides position information of particles and measures the transverse energy, which is used to identify electrons, photons, and hadrons in addition to the RICH detectors. The calorimeters also aid in the reconstruction of π^0 mesons. The LHCb calorimeter system includes an electromagnetic calorimeter (ECAL), hadron calorimeter (HCAL), scintillating pad detector (SPD) and pre-shower detector (PS). The ECAL is followed by the HCAL, while the SPD and PS are situated upstream of the ECAL [47]. The energy measurements provided by the calorimeter system are used in the analysis to construct the four-momenta of the particles, as previously discussed. Additionally, a key element of the signal signature is that it is well isolated, with no additional tracks close to the B^0, D^*, D^0, or τ vertices. As discussed in Section 5.5.3 this isolation is exploited to reduce backgrounds that contain additional tracks or neutral particles. The additional charged

tracks can be detected by the tracking detectors as well as the calorimeters, but the neutral particles (almost always π^0 particles) are only detected in the calorimeters.

The SPD/PS detector is made up of a lead converter and two planes of scintillator pads. The scintillator pads are divided into inner, middle, and outer sections. The first scintillator pad distinguishes between charged and neutral particles, while the second detects the deposited energy from electrons and photons interacting with the lead plate. The SPD/PS detector is primarily used to differentiate electrons from neutral pions. Hadrons deposit only a small portion of their energy in the ECAL and instead have their main shower in the HCAL. The HCAL's primary purpose is to provide hadronic high-energy clusters for the L0 trigger due to its moderate resolution. Hadrons are Minimum Ionising Particles in the SPD, PS, and ECAL, with almost all of their energy released in the HCAL [48].

2.4 THE MUON DETECTORS

The muon system, consisting of five stations named M1-M5, distinguishes muons from hadrons and provides a 20% resolution measurement of the muon transverse momentum used in the L0 hardware trigger [49]. In Run 2, the tracking efficiency of muons was found to be 95% [50]. The last four stations are interleaved with an iron absorber to stop hadrons not fully absorbed by the HCAL, and only muons with momentum greater than 6 GeV/c reach the last station. Muons are not present in the signal decay for this analysis, nor are they found in the backgrounds, however, different analysis strategies for the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay utilise the $\tau^+ \to \mu^+\nu_\mu\overline{\nu}_\tau$ decay. In these cases, the muon detectors are essential in identifying the signal signature.

2.5 THE LHCB TRIGGER AND DATA ACQUISITION SYSTEM

To prevent exceeding available disk space while retaining the most relevant events for analysis, a prompt and dependable trigger system is required to efficiently select and store interesting events at the LHC collider's 40 MHz bunch crossing frequency [51]. The LHCb trigger system aims to achieve an event selection rate of 12.5 kHz.

The LHCb trigger system includes the Level-0 Trigger (L0), a hardware-based trigger that reduces the rate from 40 MHz to 1 MHz, and uses p_T and E_T information from the muon stations and calorimeters. The High-Level Trigger (HLT1), a software-based trigger that filters reconstructed events based on inclusive selections to reduce the rate to 150 kHz. The HLT1 begins reconstructing VELO track segments and selects tracks using stricter p_T conditions. The output of HLT1 is processed by the HLT2, which performs more precise event reconstruction and selects events based on inclusive or exclusive requirements, producing an output rate of 12.5 kHz that is saved to permanent storage [52, 53]. The information here involves full track reconstruction and the constructions of 2, 3, and 4 track vertices, with requirements enforced on the distance between the tracks and the proposed vertex.

3 SEMILEPTONIC DECAYS

This chapter delivers a theoretical outline of the Standard Model (SM) with regard to semileptonic decays. An angular description of the rate of $B \to D^* \tau \nu$ decays is also given. This theoretical overview is based on Thompson [54], Neubert [55], and angular observable papers [29, 56].

The weak interaction of quarks can be described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [57]. Using this matrix, the weak eigenstates (x') of the quarks are related to the mass eigenstates (x),

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{3.1}$$

and the weak charged-current vertices involving quarks are given by

$$-i\frac{g}{\sqrt{2}}(\overline{u},\ \overline{c},\ \overline{t})\gamma^{\mu}\frac{1}{2}(1-\gamma^5) \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{3.2}$$

where γ^{μ} are the Dirac matrices and $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. The decay this thesis investigates is the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay, where a $b \to c\overline{\tau}\nu$ transition takes place. The Feynman diagram of this transition without the spectator quark is shown in Figure 3.1. In reality, quarks do not propagate as free particles but hadronise due to the strong force. This means the initial and final states involving quarks must be described in terms of mesons or baryons. In this thesis, the decay considered has initial and final quark states that are mesons.

In bound quark states, strong interactions are constantly taking place between the quarks that make up the hadron (valence quarks). The mediators of the strong force, gluons, self-interact so additional interactions also occur between the gluons in the hadron. Feynman diagrams involving loops can be drawn, where the gluon produces a quark-antiquark pair that interacts with the gluons in the hadron before annihilating into a gluon (virtual quarks). These interactions are shown in Figure 3.2, but additional higher-order interactions can also take place.

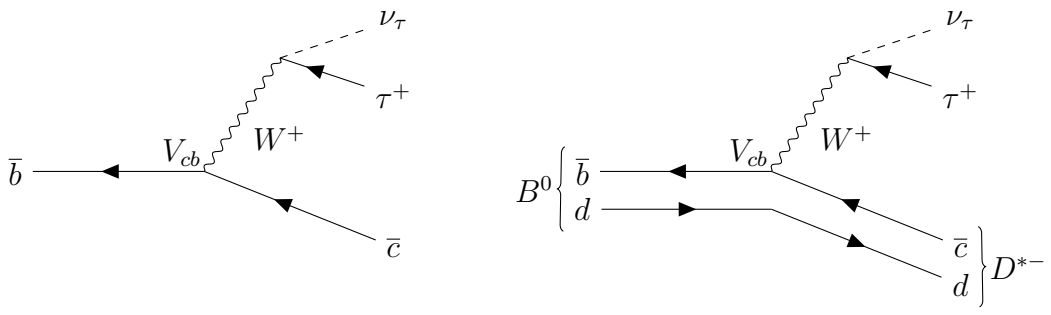


Figure 3.1: Feynman diagrams showing the SM $b \to c\overline{\tau}\nu_\tau$ transition (left) and the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay (right).

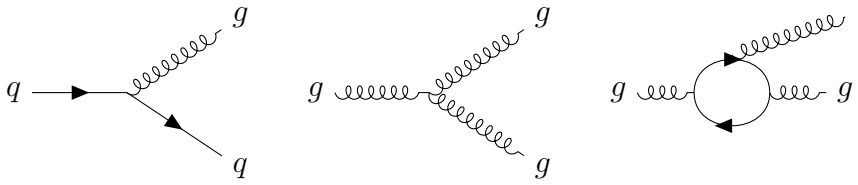


Figure 3.2: Feynman diagrams showing some simple strong interactions inside hadrons.

The description of the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay involves the complex interaction between the strong and weak interactions. Form factors are parameters that describe the effects of this complex interaction, beyond the simple weak interaction described in the $b \to c\overline{\tau}\nu_\tau$ Feynman diagram. These form factors can be measured experimentally, are expressed in terms of invariant quantities such as the momentum transfer q^2, and can have multiple parameterizations. The electroweak energy scale is $\Lambda_{\text{EW}} \approx m_W$ which is above the energy scale of the $b \to c\overline{\tau}\nu_\tau$ transition. As such, the transition can be treated as point-like and effective field theories approximating the weak decay as a four-fermion vertex can be used. The scale of the strong interaction is around $\Lambda_{\text{QCD}} \approx 0.2\,\text{GeV}$, but is only relevant when considering the internal dynamics of the B^0 and D^{*-} mesons, not the $b \to c\overline{\tau}\nu_\tau$ transition since the $\overline{\tau}$ and ν_τ leptons do not interact via the strong force.

3.1 ANGULAR FORMALISM

The SM Hamiltonian for the $b \to c\overline{\tau}\nu$ decay can be written as a product of the two vertices, bWc and $\tau W\nu$, where the neutrino is left-handed

$$\mathcal{H} = \sqrt{2}G_F V_{cb} \left[\bar{c}\gamma_\mu(1-\gamma_5)b\right] \left[\overline{\tau}_L\gamma_\mu\nu_L\right] \tag{3.3}$$

where the first term in the product is similar to Equation 3.2 and describes the $b \to cW$ transition, and the second term is the equivalent description of the $W \to \tau\nu$ transition. Here the Fermi constant is used, which can be written in terms of the weak coupling constant used in Equation 3.2, $G_F/\sqrt{2} = g^2/8m_W$. The Hamiltonian can be expanded to the effective Hamiltonian that includes additional couplings,

$$\begin{aligned}\mathcal{H}_{eff} = \sqrt{2}G_F V_{cb} \left[(1+g_V)(\bar{c}\gamma_\mu b)(\overline{\tau}_L\gamma_\mu\nu_L) + (-1+g_A)(\bar{c}\gamma_\mu\gamma_5 b)(\overline{\tau}_L\gamma_\mu\nu_L)\right. \\ + g_S(\bar{c}b)(\overline{\tau}_L\nu_L) + g_P(\bar{c}\gamma_5 b)(\overline{\tau}_L\nu_L) \\ \left. + g_T(\bar{c}\sigma_{\mu\nu}b)(\overline{\tau}_L\sigma^{\mu\nu}\nu_L) + g_{T5}(\bar{c}\sigma_{\mu\nu}\gamma_5 b)(\overline{\tau}_L\sigma^{\mu\nu}\nu_L)\right]\end{aligned} \tag{3.4}$$

where g_i with $i \in [S,\ P,\ V,\ A,\ T,\ T5]$ are the couplings to fields beyond the SM electroweak field and represent scalar, pseudo-scalar, vector, axial vector, tensor, and pseudo-tensor couplings respectively. It can be seen that by setting all $g_i = 0$, the SM effective Hamiltonian is recovered; the only terms remaining are the terms containing the $\bar{c}\gamma_\mu b$ and $-\bar{c}\gamma_\mu\gamma_5 b$ terms, which can be written as $\bar{c}\gamma_\mu(1-\gamma_5)b$ and can be seen in Equation 3.3. To derive the decay rate for the decay $B \to D^*V^*(\to \tau\nu)$, the matrix element must be first calculated

$$\mathcal{M} = \langle D^*V^*|\mathcal{H}_{eff}|B^0\rangle \tag{3.5}$$

which can be split into the different coupling amplitudes

$$\mathcal{M}_X = \langle D^*|J^X_{\text{had}}|B^0\rangle\langle\tau\nu|J_{\text{lep},\ X}|0\rangle \tag{3.6}$$

where X depends on the coupling $[S,\ P,\ V,\ A,\ T,\ T5]$ and $|0\rangle$ is the vacuum state. In the signal decay, the D^* resonance decays to $D\pi$ so the total amplitude is

$$\mathcal{M}_X = \langle D\pi|D^*\rangle\langle D^*|J^\mu_{\text{had}}|B^0\rangle\langle\tau\nu|J_{\text{lep},\ \mu}|0\rangle BW_{D^*} \tag{3.7}$$

where the propagation of the intermediate D^* resonance is given by the Breit-Wigner function

$$BW_{D^*} = \frac{1}{m_{D\pi}^2 - m_{D^*}^2 + i m_{D^*}\Gamma_{D^*}}. \tag{3.8}$$

The coupling amplitudes can be written in terms of hadronic and leptonic helicity amplitudes. These amplitudes can be separated into the different coupling types, and for the leptonic helicity amplitudes the couplings are pairs with their pseudo counterparts. For example, the $V-A$ leptonic helicity amplitude is defined as

$$L^{\lambda_\tau}_{V-A,\ \lambda}(q^2, \chi, \theta_\ell) = \eta_\mu(\lambda)\langle\tau(\lambda_\tau)\overline{\nu}|\overline{\tau}\gamma^\mu(1-\gamma_5)\nu|0\rangle \tag{3.9}$$

where $\eta_\mu(\lambda)$ is the polarisation vector of the W, λ is the helicity of the W in the $\overline{B}$ reference frame, and λ_τ is the helicity of the τ in the W reference frame. This $V-A$ leptonic helicity amplitude can be written as an explicit angular formula using the polarisation vectors, Dirac spinors, and Weyl representation of the gamma matrices. The W boson has four different polarisation states, the left and right polarisations, $\pm$, which are also helicity states, and the longitudinal and auxiliary polarisation states, $0,\ t$. The τ has two polarisation states, the left and right polarisations, $\pm$. This gives a total of 4×2 separate $V-A$ leptonic helicity amplitudes. Similarly, there are two $S-P$ amplitudes and sixteen $T-T_5$ amplitudes. For example, when the W boson and τ both have positive helicity, the leptonic helicity amplitude $L^{+}_{V-A,\ +}(q^2,\ \chi,\ \theta_\ell)$ can be calculated using

$$\eta^\mu(+) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\i\\0\end{pmatrix}, \quad \eta^\mu(0) = \frac{1}{\sqrt{q^2}}\begin{pmatrix}|\vec{p}|\\0\\0\\-E\end{pmatrix}, \quad \eta^\mu(t) = \frac{1}{\sqrt{q^2}}\begin{pmatrix}E\\0\\0\\-|\vec{p}|\end{pmatrix}, \tag{3.10}$$

for the W polarisation vectors [58], where $\vec{p}$ and E are the W boson momentum and energy respectively, the Dirac spinors are given by

$$u(\lambda = \pm 1) = \begin{pmatrix}\sqrt{E \mp |\mathbf{p}|}\xi_\pm\\ \sqrt{E \pm |\mathbf{p}|}\xi_\pm\end{pmatrix} \tag{3.11}$$

$$v(\lambda = \pm 1) = \begin{pmatrix}-\sqrt{E \pm |\mathbf{p}|}\xi_\mp\\ \sqrt{E \mp |\mathbf{p}|}\xi_\mp\end{pmatrix} \tag{3.12}$$

where E is the energy and $\mathbf{p}$ is the three-momentum of the τ or ν, with the helicity eigenspinors, $\xi_{\mp}$ defined as

$$\xi_{+} = \begin{pmatrix} \cos\theta_\ell/2 \\ \sin\theta_\ell/2 e^{i\chi} \end{pmatrix}, \; \xi_{-} = \begin{pmatrix} -\sin\theta_\ell/2 e^{-i\chi} \\ \cos\theta_\ell/2 \end{pmatrix}. \tag{3.13}$$

and the gamma matrices in the Weyl representation are

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{pmatrix}, \; \gamma^5 = \begin{pmatrix} -I_2 & 0 \\ 0 & I_2 \end{pmatrix} \tag{3.14}$$

where $\sigma^\mu = (I_2, \; \sigma^i)$, $\overline{\sigma}^\mu = (I_2, \; -\sigma^i)$, σ^i are the Pauli matrices, and I_2 is the 2×2 identity matrix. After substituting in these definitions and multiplying out the tensors, this gives

$$L^{+}_{V-A,\;+}(q^2, \; \chi, \; \theta_\ell) = \pm\sqrt{2} m_\tau \beta_\tau \sin\theta_\ell e^{-2i\chi} \tag{3.15}$$

where θ_ℓ and χ are the angles describing the decay and are defined in Figure 3.3 and $\beta_\tau = \sqrt{1 - m_\tau^2/q^2}$. Similar expressions can be found for all the leptonic helicity amplitudes L_X and are given in Appendix A.1.

The hadronic helicity amplitudes can be defined in a similar way to the leptonic amplitudes, for example, the vector helicity amplitude is defined by itself and not in a $V - A$ form as

$$H^{\lambda_{D^*}}_{V,\;\lambda}(q^2) = (g_V + 1)\eta^*_\mu(\lambda)\langle D^*(\lambda_{D^*})|\bar{c}\gamma^\mu b|\overline{B}\rangle. \tag{3.16}$$

where g_V is the additional coupling beyond the SM, λ is the helicity of the W boson, $\eta^*_\mu(\lambda)$ is the complex conjugate of the W polarisation vector, and λ_{D^*} is the

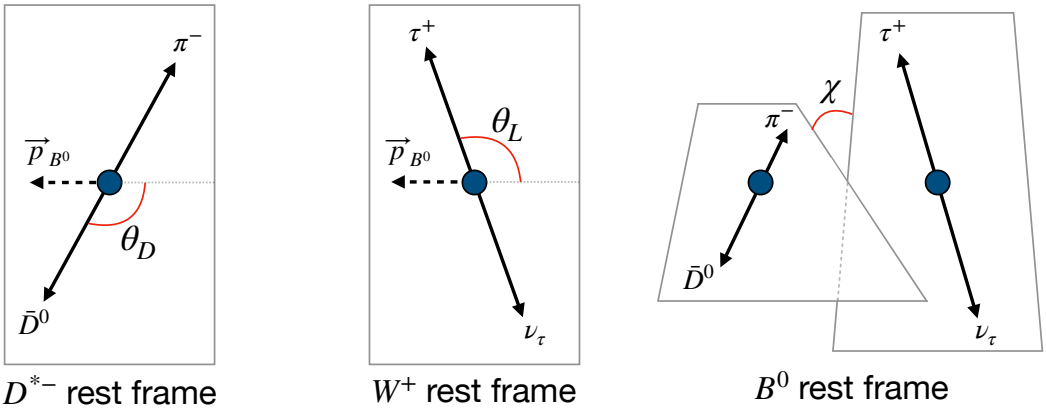


Figure 3.3: Diagram of the three decay angles for $B^0 \to D^{*-}(\to \overline{D}^0\pi^-)\tau^+\nu_\tau$.

helicity of the D^*. Rather than defining these helicity amplitudes as explicit angular functions, as was done with the leptonic amplitudes, these hadronic amplitudes must be expressed in terms of hadronic form factors. Since B^0 is a pseudo-scalar with a helicity of 0 and helicity is conserved, the helicities of the W and D^* are related. This reduces the number of possible helicity combinations, for example, the vector amplitude only has four combinations of helicity and the amplitudes involving the two combinations of the left and right polarisations are given by

$$H^{\pm}_{V,\,\pm}(q^2) = \mp(g_V + 1)\frac{\lambda_{BD^*}(q^2)}{m_B + m_{D^*}}V(q^2) \tag{3.17}$$

where $V(q^2)$ is the form factor that parameterises this vector interaction, and the lambda function $\lambda(x^2, y^2, z^2) = [x^2 - (y - z)^2][x^2 - (y + z)^2]$ is used to give $\lambda(m_B^2,\ M_{D^*}^2,\ q^2) = m_B^4 + m_{D^*}^4 + q^4 - 2(m_B^2 m_{D^*}^2 + m_B^2 q^2 + m_{D^*}^2 q^2)$. This form factor $V(q^2)$ is a function of q^2 and in general, form factors can be parameterised in several ways [59–63]. The other coupling types have different numbers of helicity combinations, the axial coupling also has four combinations, the pseudo-scalar has one, and the tensor/pseudo-tensor couplings have eight combinations each. Several form factors appear in the full list of hadronic helicity amplitudes and are given in Appendix A.2.

The total matrix element can be used to calculate the full angular distribution. This includes amplitudes from the $D^* \to D\pi$ decay, which introduce additional form factors, and terms dependent on θ_D and χ, defined in Figure 3.3, in a similar way to the θ_ℓ and χ terms introduced in the leptonic amplitudes described above. For the $\overline{B} \to D^*(\to D\pi)\ell\overline{\nu}$ decay, the full angular distribution is given by

$$\frac{d^5\Gamma}{dq^2 dm_{D\pi} d\Omega} = \frac{\sqrt{\lambda(m_B^2,\ M_{D^*}^2,\ q^2)}}{256(2\pi)^6 m_B^3}\left(1 - \frac{m_\ell^2}{q^2}\right)\frac{|\hat{p}_D|}{m_{D\pi}}\left|\mathcal{M}(\overline{B} \to D^*(\to D\pi)\ell\overline{\nu})\right|^2 \tag{3.18}$$

where the angular phase space involving the three decay angles is defined as $d\Omega = d\cos\theta_D d\cos\theta_\ell d\chi$, and the magnitude of the D three-momentum can be written as $\sqrt{\lambda(m_{D\pi}^2,\ m_D^2,\ m_\pi^2)}/2m_{D\pi}$. After summing over all the helicity amplitudes, for all the coupling types, and then integrating over $m_{D\pi}^2$ around the D^* resonance

gives

$$\begin{aligned}\frac{d^4\Gamma}{dq^2 d\left(\cos\theta_D\right) d\left(\cos\theta_L\right) d\chi} = \frac{9}{32\pi}\Big(&I_{1c}(q^2)\cos^2\theta_D + I_{1s}(q^2)\sin^2\theta_D \\ &+ \Big[I_{2c}(q^2)\cos^2\theta_D + I_{2s}(q^2)\sin^2\theta_D\Big]\cos 2\theta_L \\ &+ \Big[I_{6c}(q^2)\cos^2\theta_D + I_{6s}(q^2)\sin^2\theta_D\Big]\cos\theta_L \\ &+ \Big[I_3(q^2)\cos 2\chi + I_9(q^2)\sin 2\chi\Big]\sin^2\theta_L\sin^2\theta_D \\ &+ \Big[I_4(q^2)\cos\chi + I_8(q^2)\sin\chi\Big]\sin 2\theta_L\sin 2\theta_D \\ &+ \Big[I_5(q^2)\cos\chi + I_7(q^2)\sin\chi\Big]\sin\theta_L\sin 2\theta_D\Big)\end{aligned} \tag{3.19}$$

where the hadronic and leptonic amplitudes have been substituted in. Here, the angular dependence has been separated from the q^2 and form factor dependence using the newly defined $I_x(q^2)$. This gives the angular distribution in a form where the 3D angular structure of the decay becomes apparent, and the angular coefficients, $I_x(q^2)$ are functions of q^2. These functions are defined in terms of the q^2, helicity amplitudes, the branching ratio $\mathcal{B}(D^* \to D\pi)$, V_{cb}, the Fermi constant G_F, and the masses of the B, D^*, and lepton. The full definitions of the $I_x(q^2)$ are given in Appendix A.3. Since these $I_x(q^2)$ depend on the hadronic helicity amplitudes which in turn depend on the form factors of the different coupling types, the $I_x(q^2)$ functions contain all form factor and q^2 dependence. For example, a BSM interaction involving a pseudo-scalar coupling would introduce terms to the decay rate that depend on the pseudo-scalar form factor $A_0(q^2)$. The form factor affects the pseudo-scalar helicity amplitudes $H_P\left(q^2\right)$. This helicity amplitude appears in the definition of $\tilde{H}_t(q^2)$ which in turn appears in the definition of the angular coefficients $I_{1c}(q^2)$, $I_5(q^2)$, $I_{6c}(q^2)$, and $I_7(q^2)$. This means the presence of a pseudo-scalar coupling would affect the q^2 dependence on these four $I_x(q^2)$.

In this analysis, the q^2-integrated, rate normalised values of the $I_x(q^2)$ are measured. This thesis uses a shorthand notation for these values

$$\langle I_x\rangle = \frac{1}{\Gamma(B^0 \to D^{*-}\tau^+\nu_\tau)}\int_{q^2} I_x(q^2)dq^2 \quad \to \quad I_x. \tag{3.20}$$

Except for where it is explicitly stated, I_x is used to refer to these q^2-integrated, rate normalised values in this thesis [1]. Similarly to the example above, the presence

[1] In the abstract and conclusion, the $\langle I_x\rangle$ form is used for explicit clarity.

of a pseudo-scalar coupling would alter the q^2 integrated values of I_{1c}, I_5, I_{6c}, and I_7. The SM values for the q^2 integrated I_x used in the model for generating the signal simulation in this analysis are given in Table 8.1 and the method of their measurement is described in Section 8.1. Each of the angular coefficients I_x multiplies an angular function, as seen in Equation 3.19. For example, the I_{1s} coefficient multiplies the angular function $\sin^2\theta_D$. In general, these functions are 3D in the decay angles, however, some particular angular functions may be flat in one or more decay angles. The 1D angular functions for each of the three decay angles can be seen for all I_x in Figures 3.4, 3.5, and 3.6. The construction of the 1D projections of the total model is shown in Figure 3.7. Partial models are first shown where only a subset of angular functions is summed using the SM predictions for the I_x.

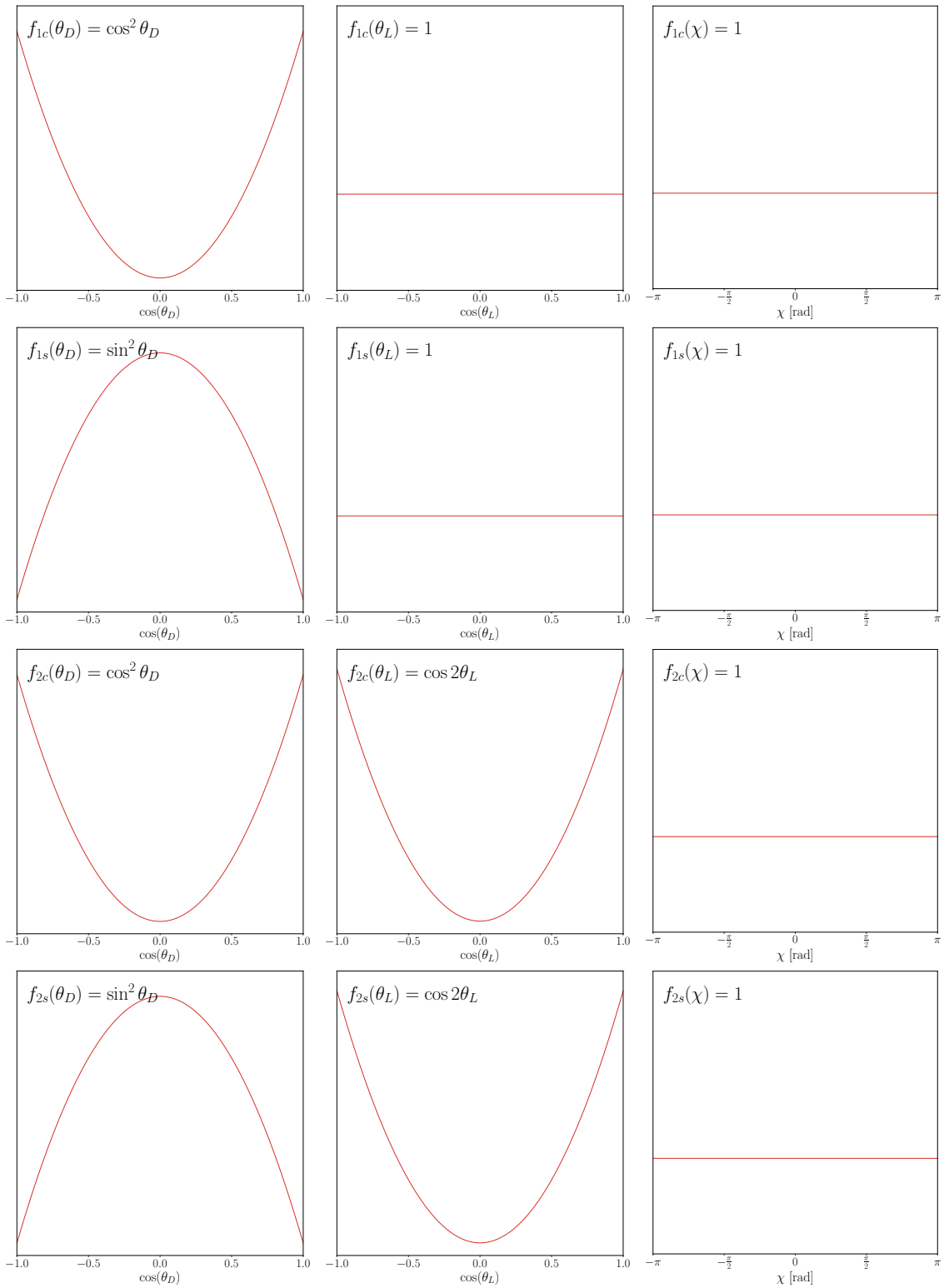


Figure 3.4: Shape in $\cos\theta_D$ (left), $\cos\theta_L$ (centre), and χ (right) functions of the I_x coefficients, where $x \in [1c,\ 1s,\ 2c,\ 2s]$. These functions are denoted $f_x(\phi)$, where x is the angular coefficient index and ϕ is the angle that the function depends on.

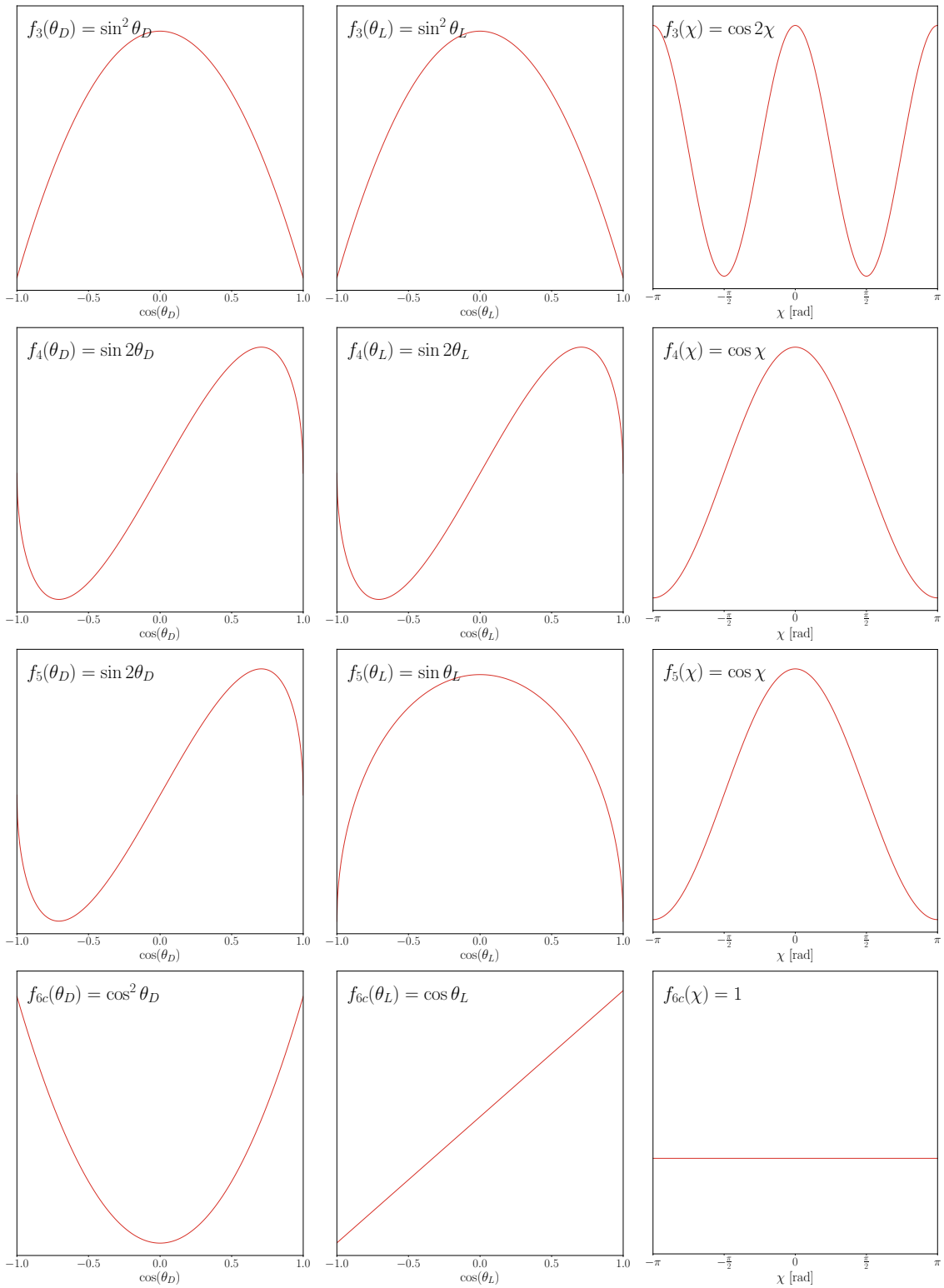


Figure 3.5: Shape in $\cos\theta_D$ (left), $\cos\theta_L$ (centre), and χ (right) functions of the I_x coefficients, where $x \in [3,\ 4,\ 5,\ 6c]$. These functions are denoted $f_x(\phi)$, where x is the angular coefficient index and ϕ is the angle that the function depends on.

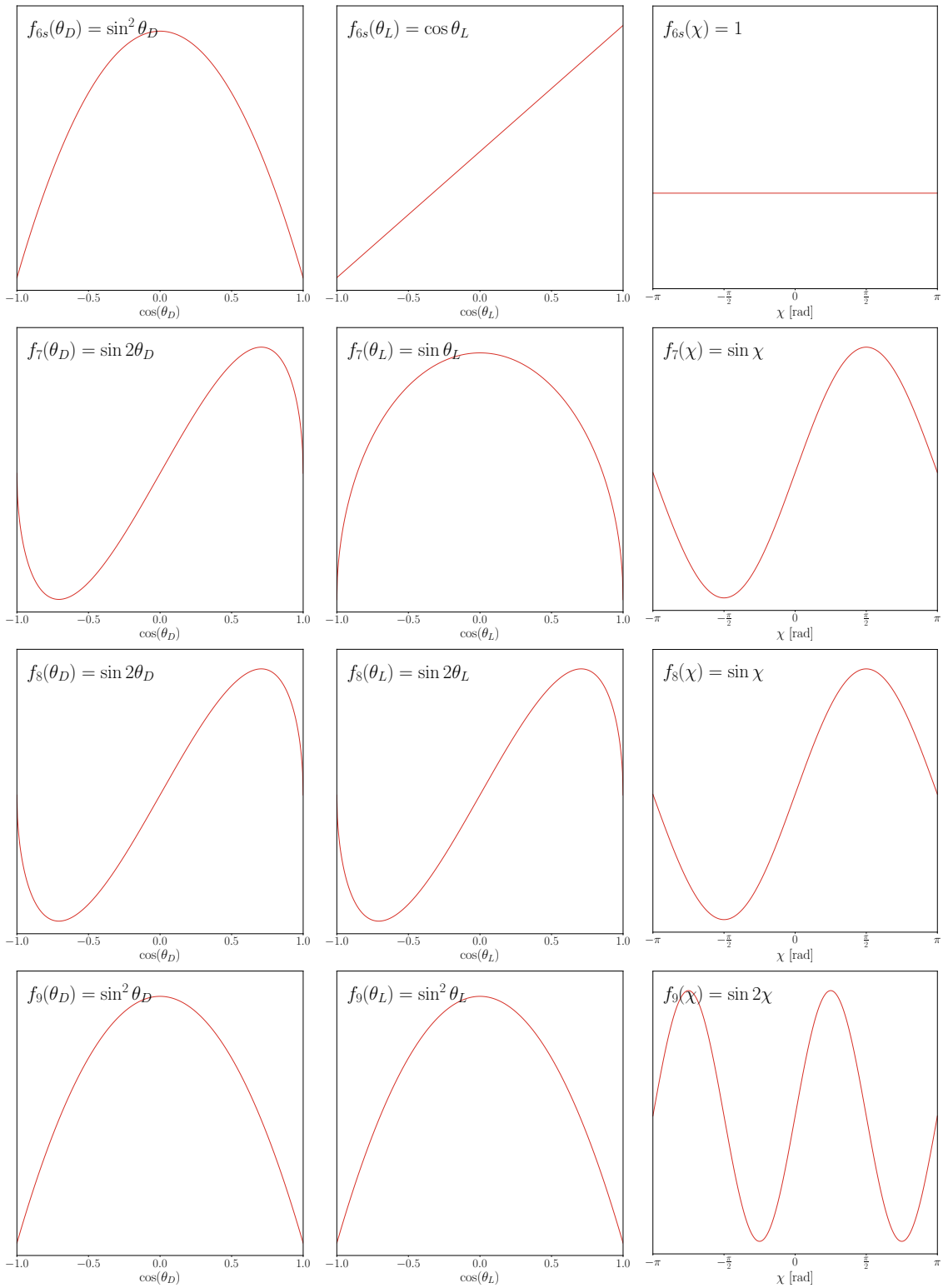


Figure 3.6: Shape in $\cos\theta_D$ (left), $\cos\theta_L$ (centre), and χ (right) functions of the I_x coefficients, where $x \in [6s,\ 7,\ 8,\ 9]$. These functions are denoted $f_x(\phi)$, where x is the angular coefficient index and ϕ is the angle that the function depends on.

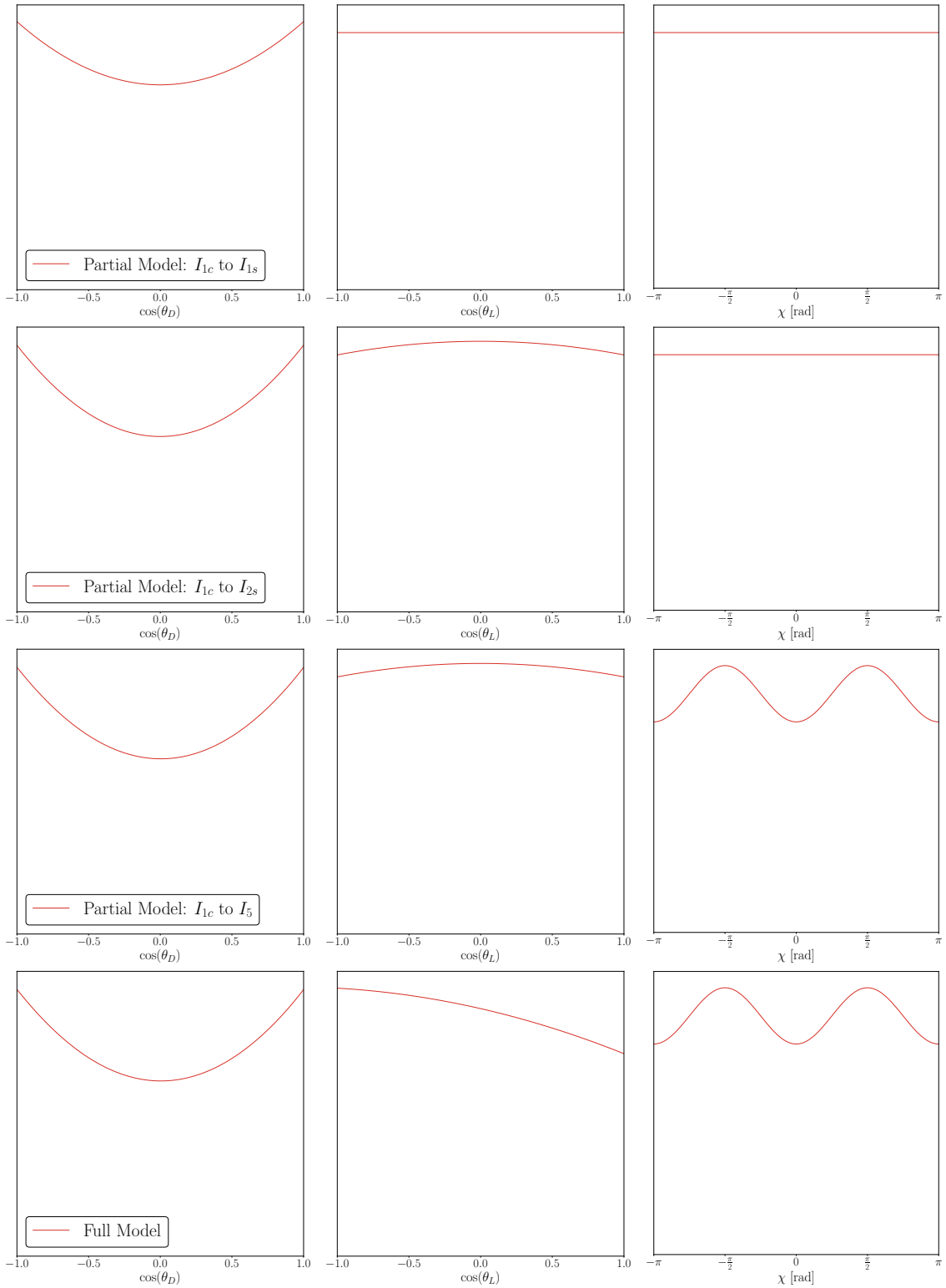


Figure 3.7: Projections of the full angular model (bottom), where the angular functions are summed using the SM values of the angular coefficients I_x. The first three rows show partial models, where only some of the angular functions are summed: I_{1c} to I_{1s} (first), I_{1c} to I_{2s} (second), I_{1c} to I_5 (third). These partial models help show how the total model changes as you add the contributions of more angular functions.

The angular coefficients can be used to calculate additional observables, such as the longitudinal polarisation fraction of the vector charm meson, $f_L^{D^*}$, the forward-backward asymmetry, A_{FB}, or several other asymmetries that isolate particular I_x in their definitions. Their definitions in terms of the angular coefficients are as follows,

$$
\begin{aligned}
f_L^{D^*} &= \frac{3I_{1c} - I_{2c}}{3I_{1c} - I_{2c} + 6I_{1s} - 2I_{2s}} \\
A_{FB} &= \frac{3}{4}\frac{I_{6c} + 2I_{6s}}{3I_{1c} - I_{2c} + 6I_{1s} - 2I_{2s}} \\
A_C^1 &= \frac{4}{3}\frac{I_3}{A_L + A_T} \\
A_C^2 &= \frac{I_5}{A_L + A_T} \\
A_C^3 &= \frac{I_4}{A_L + A_T} \\
A_T^1 &= \frac{4}{3}\frac{I_9}{A_L + A_T} \\
A_T^2 &= \frac{I_7}{A_L + A_T} \\
A_T^3 &= \frac{I_8}{A_L + A_T}
\end{aligned}
$$

where the numbers A_L and A_T are multiples of the longitudinal and transverse D^* polarisation decay rates, they also defined in terms of the angular coefficients,

$$
\begin{aligned}
A_L &= I_{1c} - \frac{1}{3}I_{2c} \\
A_T &= 2\left(I_{1s} - \frac{1}{3}I_{2s}\right).
\end{aligned}
$$

The full decay rate Γ can also be written in terms of the I_x and the convention that $\Gamma = 1$ is used, making the I_x rate-normalised,

$$
\Gamma = \frac{1}{4}\left(3I_{1c} + 6I_{1s} - I_{2c} - 2I_{2s}\right) = 1. \tag{3.21}
$$

It is important to note that the angular distribution can reveal BSM effects even if $R(D^*)$ becomes fully compatible with the SM [29]. It can be seen in Equation 3.21 that the integrated full decay rate only depends on some of the angular coefficients, meaning that the other coefficients may deviate from SM predictions and go undetected when only measuring the total decay rate. Additionally, multiple coefficients

that the decay rate does depend on can disagree with SM, but give a value of the decay rate that is compatible.

Angular analysis of B-meson decays, such as $B^0 \to K^{*0}\mu^+\mu^-$, where parametric probability density functions (PDF) are used to fit the angular distribution, is well established [64]. This technique has the advantage of including all form factor information in the angular coefficients, eliminating systematic uncertainty due to a choice of form factor scheme. Studies of $B \to K^{(*)}\mu^+\mu$ decays have been performed by several collaborations. The BaBar collaboration measured rate asymmetries and angular distributions in 2006, finding agreement with the SM [65]. Subsequently, the CDF collaboration measured the angular distributions in 2012, also finding no deviation from the SM [66]. In contrast, LHCb performed a measurement in 2014 that showed the branching fractions favoured lower values than the SM while the isospin symmetries were compatible with SM predictions [67]. The $B^0 \to K^{*0}\mu^+\mu^-$ decay has also been studied by the ATLAS and CMS experiments [68,69]. The Belle collaboration found a 2.6σ tension with the SM in the $B^+ \to K^{*+}\mu^+\mu$ decay [70]. Similar measurements have been performed for the $\Lambda_b^0 \to \Lambda\mu^+\mu$ decay at LHCb, with the differential branching fractions showing tension with the SM in the low q^2 region [71]. LHCb further investigated this phenomenon using the full Run 1 + 2 data sets, corresponding to 9 fb^{-1} of data, and found the SM to be consistent with measurements in the CP-averaged angular observables for the $B^+ \to K^{*+}\mu^+\mu$ decay [72].

These techniques leveraging angular information are very powerful, as previously discussed. The parametric PDFs can be derived theoretically, in a general form that includes all sources of possible BSM physics. However, for semitauonic decays, the presence of multiple neutrinos causes information loss and severe distortions to angular variable distributions, making a parametric fit impossible for data analysis. The following chapter described the methodology used to make the measurement of the angular observables possible for partially reconstructed decays.

4 METHODOLOGY OF THE ANGULAR MEASUREMENT

In this chapter, the methodology of the I_x measurements are described. Section 4.1 describes the signal decay with three-prong τ decays and a normalisation channel. The kinematic reconstruction under the signal hypothesis is defined in Section 4.2. The reconstruction of the decay angles is addressed in Section 4.3. Section 4.5 details the angular coefficient parameterisation of the signal decay. Finally, the backgrounds involved in the measurement are discussed in Section 4.6.

4.1 THE SIGNAL

For this analysis, the signal decay mode is $B^0 \to D^{*-}\tau^+\nu_\tau$, where the D^{*-} is reconstructed via the three track final state $D^{*-} \to (\overline{D}^0 \to K^+\pi^-)\pi^-$. The decay $D^{*-} \to \overline{D}^0\pi^-$ has a branching ratio of $(67.7 \pm 0.5)\%$ while the $\overline{D}^0 \to K^+\pi^-$ decay has a branching ratio of $(3.946 \pm 0.030)\%$ [46]. The hadronic $\tau \to \pi^+\pi^-\pi^+(\pi^0)\overline{\nu}_\tau$ decays are employed in this analysis, which have branching ratios of $(9.31 \pm 0.05)\%$ and $(4.62 \pm 0.05)\%$, respectively [46]. The two neutrinos produced in the full decay chain are not detected, resulting in only partially reconstructed τ and B^0 mass peaks for the signal mode and sculpting of the decay angle distributions. Therefore, the final visible state for the signal mode is $5\pi^\pm\, K^+$. This can be seen in Figure 4.1 (left) where the signal mode topology is shown. The three-prong τ decays enable the reconstruction of the τ vertex, which in turn allows for the decay angles to be reconstructed as described in Section 4.3.

The normalisation mode $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ is used since it has the same visible final state as the signal decay as seen in Figure 4.1 (right). This mode is used in the selection optimisation to provide an estimate of the signal yield, this process is described in Section 5.5.4. The decay is also used in a comparison of data and simulation to correct the L0 trigger composition in the simulation, as described in Section 7.4.

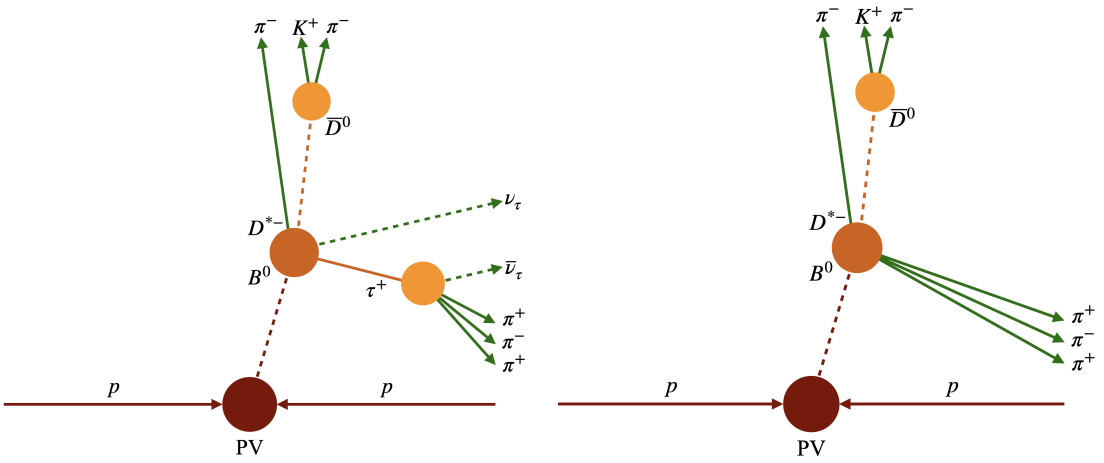


Figure 4.1: Diagram of the topology of the $B^0 \to D^{*-}\tau^+\nu_\tau$ signal decay (left) and the $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ decay mode (right). The green lines indicate the final state particles, and the orange and red are intermediate states. Solid lines represent charged particles while dashed lines are neutral.

4.2 KINEMATIC RECONSTRUCTION

Due to the neutrinos produced at the B^0 and τ decay vertices and a lack of constraints on the initial conditions of the collision, neither the B^0 nor τ momenta can be fully reconstructed. In this analysis, they are approximated using the decay vertices of the $\pi^+\pi^-\pi^+$ and $D^{*-}\pi^+\pi^-\pi^+$ systems. The $\pi^+\pi^-\pi^+$ and $D^{*-}\pi^+\pi^-\pi^+$ vertices allow for the reconstruction of the τ and B^0 flight directions, where the B^0 flight direction is defined by the line from the primary vertex to the $D^{*-}\pi^+\pi^-\pi^+$ vertex and the τ flight direction is similarly defined as the line from the $D^{*-}\pi^+\pi^-\pi^+$ vertex to the $\pi^+\pi^-\pi^+$ vertex. Using the known masses of these particles, their momenta can be estimated up to a two-fold ambiguity in the lab frame. The τ momentum magnitude in the lab frame can be written as

$$|\vec{p}_\tau| = \frac{(m_{3\pi}^2)|\vec{p}_{3\pi}|\cos\theta_{\tau,3\pi} \pm E_{3\pi}\sqrt{(m_\tau^2 - m_{3\pi}^2)^2 - 4m_\tau^2|\vec{p}_{3\pi}|^2\sin^2\theta_{\tau,3\pi}}}{2(E_{3\pi}^2 - |\vec{p}_{3\pi}|^2\cos^2\theta_{\tau,3\pi})} \tag{4.1}$$

where $m_{3\pi}$, $\vec{p}_{3\pi}$, and $E_{3\pi}$ are the reconstructed invariant mass, momentum, and energy of the $\pi^+\pi^-\pi^+$ system, m_τ is the known τ mass, and $\theta_{\tau,3\pi}$ is the angle between the $\pi^+\pi^-\pi^+$ system momentum vector and the flight vector of the τ. In the absence of the neutrino, $\vec{p}_{3\pi}$ and $\vec{p}_\tau$ would be aligned and $\theta_{\tau,3\pi} = 0$ reducing

Equation 4.1 to a single trivial solution. Equation 4.1 can also be reduced to a single nontrivial solution when $\theta_{\tau,3\pi}$ takes the maximum allowed value

$$\theta^{\max}_{\tau,3\pi} = \arcsin\left(\frac{m_\tau^2 - m_{3\pi}^2}{2m_\tau|\vec{p}_{3\pi}|}\right) \tag{4.2}$$

such that the square root term is zero, this solution is chosen as it is the most conservative estimate of the reconstructed τ momentum, where the neutrino in the decay is not reconstructed. Additionally, this solution provides one non-ambiguous estimate of the τ momentum, using a non-trivial solution for $\theta_{\tau,3\pi}$. When combined with the τ line of flight, this provides an estimate of the τ momentum components with minimal bias. Figure 4.2 shows the resulting biases which are unavoidable when the momenta are only partially reconstructed. The B^0 momentum is estimated using a similar method, using the reconstructed $D^{*-}\tau^+$ system

$$|\vec{p}_{B^0}| = \frac{(m_Y^2)|\vec{p}_Y|\cos\theta_{B^0,Y} \pm E_Y\sqrt{(m_{B^0}^2 - m_Y^2)^2 - 4m_{B^0}^2|\vec{p}_Y|^2\sin^2\theta_{B^0,Y}}}{2(E_Y^2 - |\vec{p}_Y|^2\cos^2\theta_{B^0,Y})} \tag{4.3}$$

where Y represents the $D^{*-}\tau^+$ system, and the τ^+ is reconstructed as described above using Equation 4.1. The maximum opening angle between the B^0 flight vector and the reconstructed $D^{*-}\tau^+$ momentum vector, $\theta^{\max}_{B^0,Y}$, is similarly chosen such that the square-root term is zero

$$\theta^{\max}_{B^0,Y} = \arcsin\left(\frac{m_{B^0}^2 - m_Y^2}{2m_{B^0}|\vec{p}_Y|}\right). \tag{4.4}$$

When used in Equation 4.3 to produce a single solution and the B^0 line of flight, the B^0 momentum components can be estimated. These estimates can be used to obtain Lorentz invariant variables associated with the decay, such as $q^2 \equiv (p_{B^0} - p_{D^{*-}})^2 = (p_\tau + p_{\nu_\tau})^2$ and the τ lifetime, τ_τ. Figure 4.2 shows signal simulation true and reconstructed distributions for $|p_\tau|$, $|p_{B^0}|$, q^2, and τ_τ, the resolution for these variables, and the two-dimensional distributions of the true and reconstructed quantities. The bias and resolution for these variables are reported in Table 4.1.

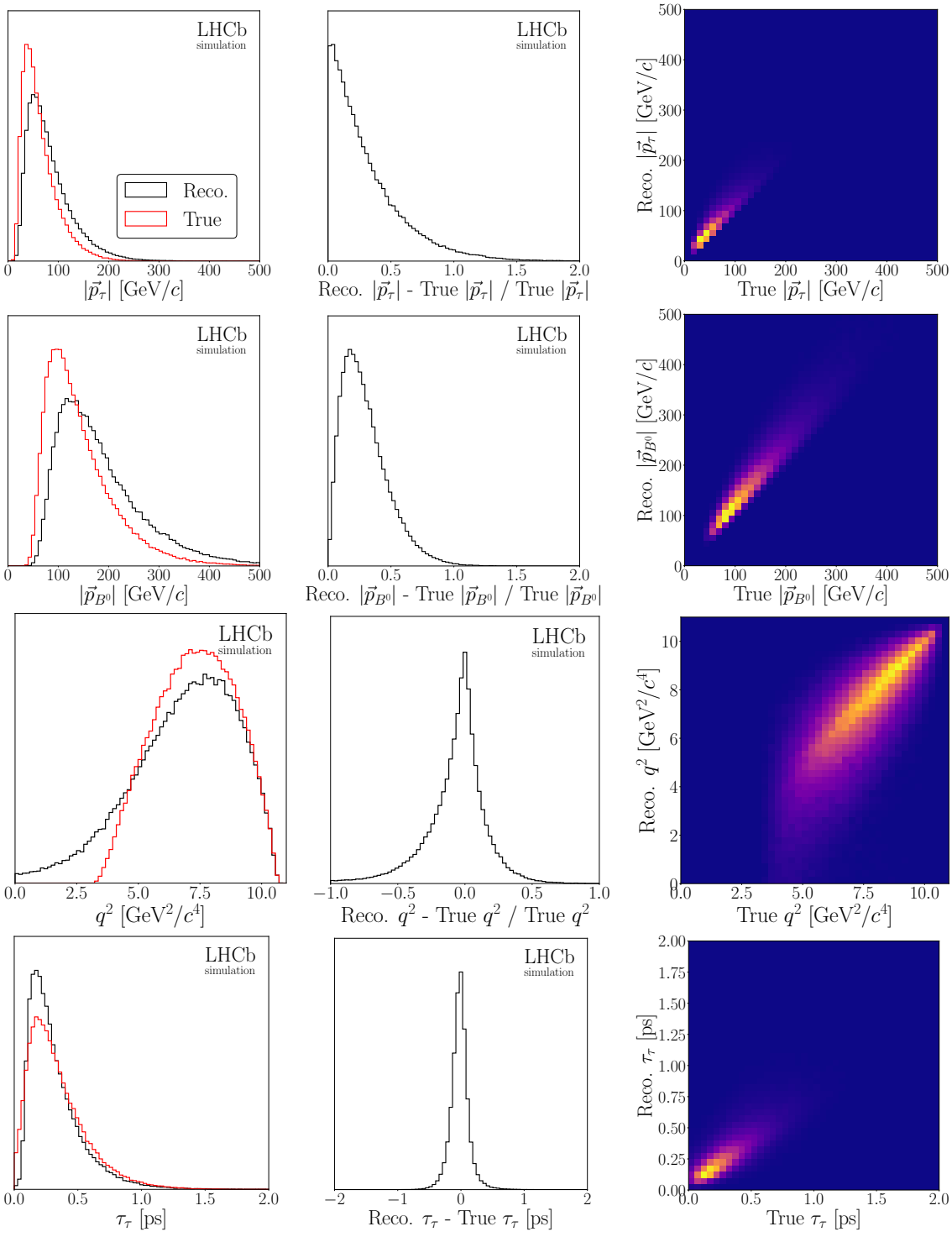


Figure 4.2: Signal simulation true and reconstructed distributions (left), resolutions where the $|p_\tau|$, $|p_{B^0}|$, and q^2 distributions are normalised (middle), and the two-dimensional distributions of the true and reconstructed quantities, where lighter colours represent a higher density of events (right) for $|p_\tau|$ (first row), $|p_{B^0}|$ (second row), q^2 (third row), and τ_τ (fourth row).

Table 4.1: Summary of the bias and resolution for $|p_\tau|$, $|p_{B^0}|$, q^2, and τ_τ. These are defined as the mean and standard deviation of $(x_{\text{Reco.}} - x_{\text{True}})/x_{\text{True}}$ for $x \in \{|p_\tau|, |p_{B^0}|, q^2\}$ and $(x_{\text{Reco.}} - x_{\text{True}})$ for $x = \tau_\tau$.

Variable	True mean	Resolution μ	Resolution σ
$\lvert p_\tau \rvert$	66.8 GeV/c	0.33	0.30
$\lvert p_{B^0} \rvert$	151.2 GeV/c	0.30	0.19
q^2	7.1 GeV$^2/c^4$	-0.07	0.24
τ_τ	0.33 ps	-0.02 ps	0.15 ps

4.3 DECAY ANGLE RECONSTRUCTION

The signal decay can be parameterised through three decay angles: θ_D is the angle between the $\overline{D}^0$ flight direction and the opposite of the B^0 flight direction, in the D^{*-} rest frame; θ_L is the angle between the τ^+ flight direction and the opposite of the B^0 flight direction in the virtual W^+ boson rest-frame; and χ is the angle between the plane containing the $\overline{D}^0$ and π^- from the D^{*-} decay and the plane containing the τ^+ and ν_τ from the W^+ decay, in the B^0 rest frame. These angles are defined as follows and follow the definition in Reference [73]

$$\cos\theta_D = \left(\hat{p}_{\overline{D}^0}^{(D^{*-})}\right) \cdot \left(\hat{p}_{D^{*-}}^{(B^0)}\right) = \left(\hat{p}_{\overline{D}^0}^{(D^{*-})}\right) \cdot \left(-\hat{p}_{B^0}^{(D^{*-})}\right) \tag{4.5}$$

$$\cos\theta_L = \left(\hat{p}_{\tau^+}^{(W^+)}\right) \cdot \left(\hat{p}_{W^+}^{(B^0)}\right) = \left(\hat{p}_{\tau^+}^{(W^+)}\right) \cdot \left(-\hat{p}_{B^0}^{(W^+)}\right) \tag{4.6}$$

$$\cos\chi = \left(\hat{p}_{\tau^+}^{(B^0)} \times \hat{p}_{\nu_\tau}^{(B^0)}\right) \cdot \left(\hat{p}_{\overline{D}^0}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right) \tag{4.7}$$

$$\sin\chi^{B^0} = -\left[\left(\hat{p}_{\tau^+}^{(B^0)} \times \hat{p}_{\nu_\tau}^{(B^0)}\right) \times \left(\hat{p}_{\bar{D}^0}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right)\right] \cdot \hat{p}_{D^{*-}}^{(B^0)} \tag{4.8}$$

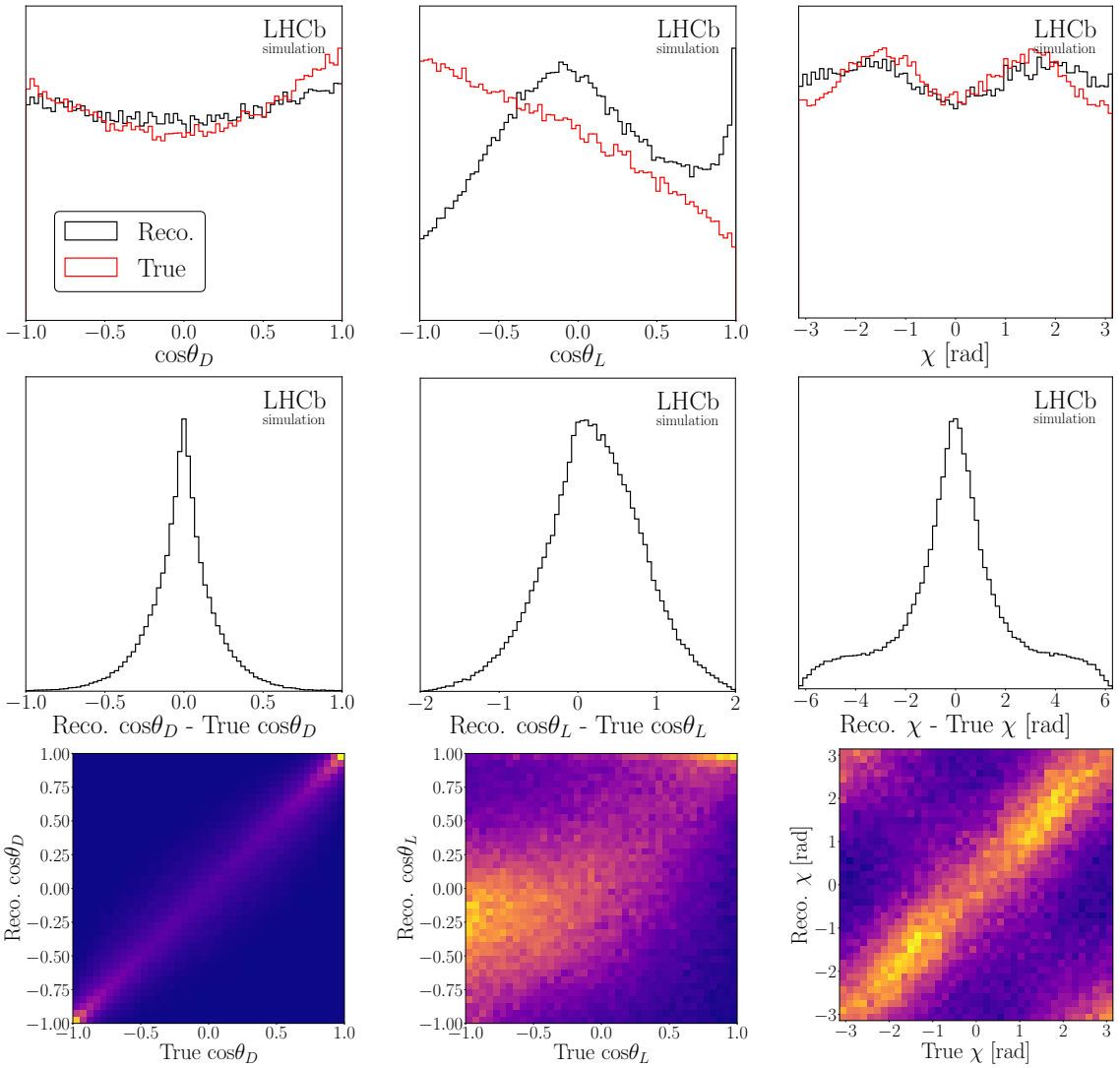
$$\sin\chi^{\bar{B}^0} = +\left[\left(\hat{p}_{\tau^-}^{(\bar{B}^0)} \times \hat{p}_{\bar{\nu}_\tau}^{(\bar{B}^0)}\right) \times \left(\hat{p}_{D^0}^{(\bar{B}^0)} \times \hat{p}_{\pi^-}^{(\bar{B}^0)}\right)\right] \cdot \hat{p}_{D^{*+}}^{(\bar{B}^0)} \tag{4.9}$$

where $\hat{p}_X^{(Y)}$ denotes the unit vector of the direction of the particle X in the rest frame of system Y. In every case, the particle momenta are first boosted to the B^0 frame. In this basis, the definition of the angles in the $\overline{B}^0$ decay is a CP transformation of the angles defined for the B^0 decay. This results in the relative sign in $\sin\chi$ being negative for B^0 candidates and positive for $\bar{B}^0$ candidates, where the B-meson flavour is tagged by the D^*-meson charge. The three decay angles were shown previously in Figure 3.3.

Table 4.2: Summary of the bias and resolution for $\cos\theta_D$, $\cos\theta_L$, and χ. These are defined as the mean and standard deviation of $(x_{\text{Reco.}} - x_{\text{True}})$ for $x \in \{\cos\theta_D, \cos\theta_L, \text{ and } \chi\}$.

Variable	Resolution μ	Resolution σ
$\cos\theta_D$	-0.01	0.22
$\cos\theta_L$	0.20	0.63
χ	0.01 rad	2.29 rad

The decay angles suffer from significant resolution and reconstruction effects since the τ^+ and B^0 momentum vectors are inputs to the decay angle calculations. These effects can be seen in Figure 4.3 where the true and reconstructed distributions for the angles are compared and their resolution distributions are shown for signal simulation which have preliminary selection criteria applied. The bias and resolution for the decay angles are reported in Table 4.2. It can be seen that $\cos\theta_D$ is the best reconstructed of the decay angles while there are large resolution effects present in the $\cos\theta_L$ and χ reconstruction. These resolution effects in the decay angles result in significant migration of events within the angular phase space. This is shown in Figure 4.4 where the two-dimensional projections of the true and reconstructed angular phase space are presented. The density difference between the true and reconstructed projections is shown, where the regions of increased (red) and decreased (blue) density are caused by the reconstruction. It can be seen that the event migration generally decreases density variation over the phase space, with an additional bias towards positive $\cos\theta_L$ values present.



(a) $\cos\theta_L$ reconstruction. (b) $\cos\theta_L$ reconstruction. (c) χ reconstruction.

Figure 4.3: The true and reconstructed distributions (top), resolutions (middle), and the two-dimensional distributions of the true and reconstructed quantities, where lighter colours represent a higher density of events (bottom) for $\cos\theta_D$ (a), $\cos\theta_L$ (b), and χ (c). For both the true and the reconstructed distributions, only events that have passed preliminary selection criteria are included.

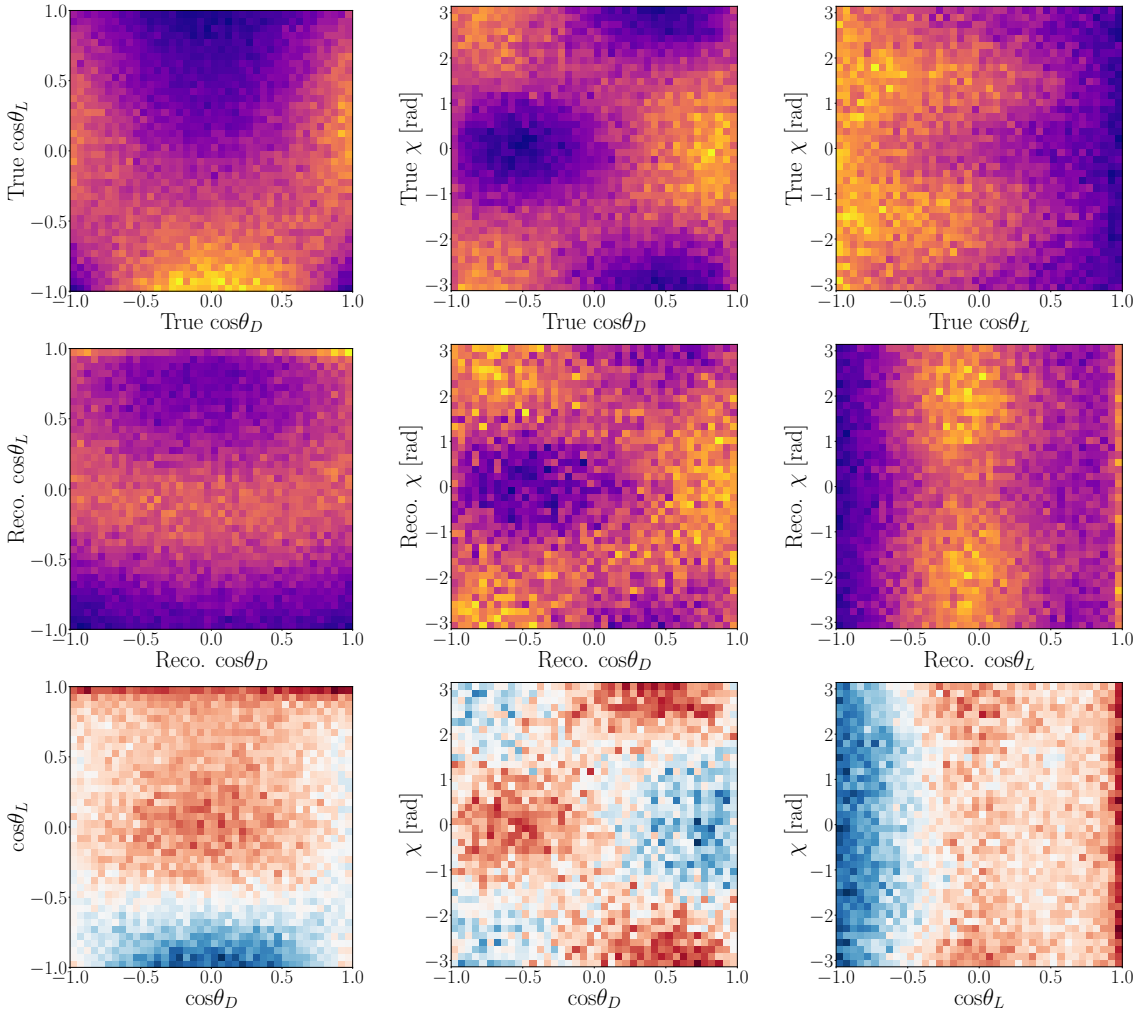


Figure 4.4: The two-dimensional projections of the true (top) and reconstructed (middle) angular distributions where the lighter colours represent areas of higher density. The density difference in each bin (bottom) shows where the angular phase space density increases (red) or decreases (blue) due to the reconstruction.

4.4 BIAS IN $\cos\theta_L$

The kinematically-constrained reconstruction of the τ lepton and especially the B^0 meson leads to significant bias in the $\cos\theta_L$ distribution. The higher values of $\cos\theta_L$ imply that the τ lepton departs the B^0 decay in a direction that is more colinear with the negative direction of the B^0 meson, as calculated in the rest-frame of the virtual W boson. The difference between $\theta^{max}_{\tau,3\pi}$ and $\theta^{max}_{B^0,Y}$ and their true values, as measured in signal simulation, are shown in Figure 4.5 and show an upward bias in $\theta_{\tau,3\pi}$ and a downward bias in $\theta_{B^0,Y}$.

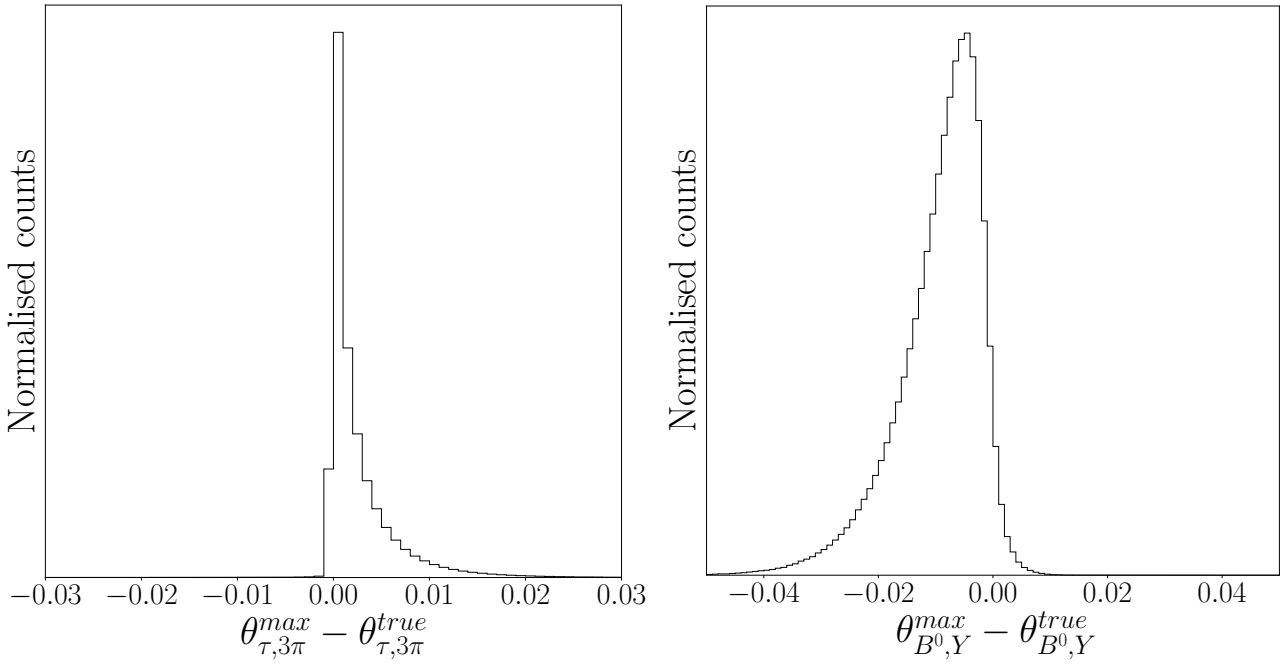


Figure 4.5: Left: The difference between $\theta^{max}_{\tau,3\pi}$ and the true value as measured in signal simulation. Right: The difference between $\theta^{max}_{B^0,Y}$ and the true value.

A negative correlation between $\cos\theta_L$ and the lab-frame reconstruction angle $\theta^{\max}_{B0,Y}$ is shown in Figure 4.6 (left), which is produced using signal simulation, and shows that the $\theta^{max}_{B^0,Y}$ value used in the reconstruction (as defined in Equation 4.4) nearly always gives a smaller defection angle that it should. This is an unavoidable necessity of the partial reconstruction. The smaller apparent deflection is kinematically equivalent to a larger apparent momentum of the $D^*\tau$ system. The over estimation of the $D^*\tau$ system's momentum comes from the necessary choice of $\theta^{max}_{\tau,3\pi}$ in Equation 4.2, which systematically chooses a higher value for $|p_\tau|$.

To study the choice of using $\theta^{max}_{\tau,3\pi}$ and $\theta^{max}_{B^0,Y}$, the truth-level information in the signal simulation is utilised, and the two-valued solutions to Equations 4.1 and 4.3 can be used, instead of always relying on the maximum angle choices as defined in Equations 4.2 and 4.4. If the $\sqrt{X}$ terms in Equations 4.1 and 4.3 are negative, then the nominal (N) solution can be taken (ie. using $\theta^{max}_{\tau,3\pi}$ and $\theta^{max}_{B^0,Y}$), otherwise, the choice of the $-\sqrt{X}$ or $+\sqrt{X}$ solution can be determined through comparison to the true momentum and selecting the closest choice. This gives three possible choices for the p_τ reconstruction and three for the p_{B^0} reconstruction, totaling 9 combinations which are denoted as (X, Y) where X is the method of p_{B^0} reconstruction and

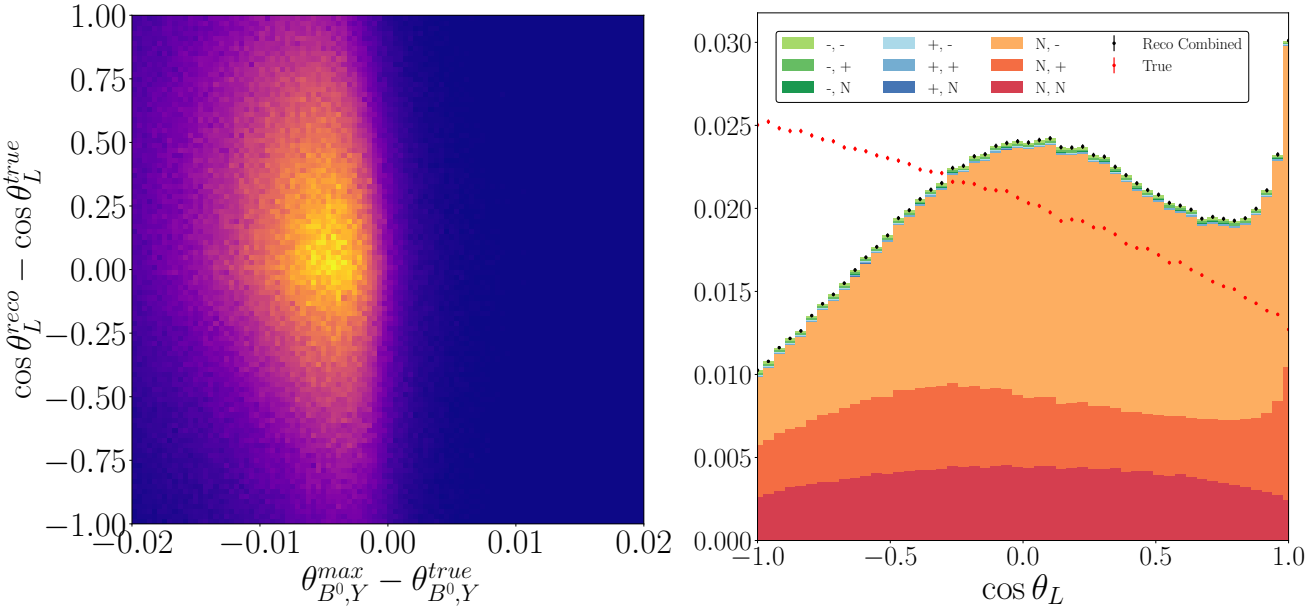


Figure 4.6: Left: The difference between the reconstructed $\cos\theta_L$ and the true value vs the difference between $\theta^{max}_{B^0,Y}$ as defined in Equation 4.4 and the true angle of $\theta_{B^0,Y}$. This shows a negative correlation between $\cos\theta_L$ and the lab-frame reconstruction angle $\theta_{B0,Y}$. Right: The $\cos\theta_L$ distribution where the total distribution has been split into 9 categories based on different reconstruction choices for the B^0 and τ momentum. These combinations denoted as (X, Y) where X is the method of p_{B^0} reconstruction and Y is the method of p_τ reconstruction. The signal simulation samples where used to produce both of these figures.

Y is the method of p_τ reconstruction. These combinations can be seen in Figure 4.6 (right) where the forward-biasing is dominated by the $(N, -)$ component. In this case, the $-\sqrt{X}$ solution of Equation 4.1 should have, ideally, been chosen. This means the $\theta^{max}_{\tau,3\pi}$ choice biases to higher momentum. This overestimate nearly always results in the non-physical $\sqrt{-X}$ solution in Equation 4.3.

4.5 ANGULAR COEFFICIENT PARAMETERISATION

The full angular decay rate is given by Equation 3.19 as derived in [29], with θ_D, θ_L, and χ defined in Figure 3.3. In this equation, there is a sum of twelve independent angular functions multiplied by a corresponding angular coefficient $I_x(q^2)$ (where $x \in \{1c, 1s, 2c, 2s, 3, 4, 5, 6c, 6s, 7, 8, 9\}$) which contain the dependence on q^2, form factors, and fundamental couplings. As previously mentioned, this analysis focuses on the q^2 integrated angular coefficients I_x. Due to the reconstruction effects that produce event migration in the angular phase space, as discussed in Sections

4.2, 4.3, a parametric fit to the reconstructed decay angles cannot be used to measure the q^2 integrated I_x coefficients, as is the case for fully reconstructed analysis such as $B^0 \to K^{*0}\mu^+\mu^-$ [73]. Introducing a correction to the reconstruction biases using simulation by necessity leads to a dependence on the model used in the simulation to derive this correction. Instead, a model-independent approach is utilised where the signal simulation is used to produce angular histogram templates to measure the I_x and reconstruction effects are incorporated directly into these templates. This is described in detail in Section 8.2.

4.6 BACKGROUNDS

The backgrounds in the signal fit can be separated into three general categories. The largest background category is known as the Double Charm background, due to its components having the form $B \to D^{*-}h_c(X)$, where $B \in \{B^0, B^\pm, B_s^0\}$, $h_c \in \{D^0, D^+, D_s^+\}$ is a charm hadron, and X are other potentially non-reconstructed particles. This background category is further divided based on the species of h_c, where the D_s^+ component is the largest. All three of these mesons produce a vertex that is detached from the B vertex, which mimics the flight of a τ lepton in the signal mode before it undergoes a three-prong decay. Additionally, these decays can produce six or more tracks, with extra tracks not being reconstructed at the B and/or h_c vertices. Extra neutral particles can also be produced in these decays which are then not reconstructed. In either case, these extra particles result in missing energy, mimicking the two neutrinos in the signal decay.

Background from $B \to D^{*-}\pi^+\pi^-\pi^+(X)$ decays, referred to as Prompt, enters the signal fit for the same reasons as the Double Charm background. Such decays involve the production of three pions at the B decay vertex and thus can be suppressed by requiring that the τ vertex is well displaced from the B^0 vertex. However, the high flight tail of this background which is caused by detector resolution effects still enters the signal fit. This can be seen in Section 5.4 where the flight is used to reduce this background and distributions of different simulation samples are compared and discussed.

Another contributing background has the form $B \to D^{**-}\tau^+\nu_\tau$, and is referred to as Feed-down. Such decays are similar to signal, but where the charm meson is produced in a higher angular momentum state. The excited charm mesons decay very quickly, producing additional particles at the B vertex that are not reconstructed. These are partially suppressed through the application of an multivariate analysis (MVA) described in Section 5.5.3 and the remaining contribution is estimated relative to the signal fraction, as described in Section 8.3.2.

The final background in the signal fit is combinatorial in nature, where particles from different decays are combined into a signal candidate in error. There are many sources of combinatorial background: an event with a genuine D^* that is combined with a τ candidate where one or more of the pions come from another decay; an event with a genuine D^0 that is then combined with a slow pion from another event, forming a fake D^* candidate; or an event where a kaon and a pion from different events are combined into a fake D^0 candidate. These backgrounds are suppressed through the use of the MVA detailed in Section 5.5.1 and the application of ${}_s$Weights, a statistical tool to remove measured backgrounds as shown in Section 5.7.

The backgrounds in the normalisation fit are small compared to the normalisation peak, but several backgrounds are accounted for in the normalisation fit as described in Chapter 6. The normalisation peak is found at the known B^0 mass since the decay is fully reconstructed. Also peaking at the B^0 mass, the fully reconstructed $B^0 \to D^{*-}D_s^+$ decay is measured prior to the normalisation fit and is subsequently fixed in the normalisation fit. The partially reconstructed $B^0 \to D^{*-}D_s^{*+}$ decay is included in the $B^0 \to D^{*-}D_s^+$ decay measurement and the normalisation measurement. There is also a small misidentified background from $B^0 \to D^{*-}K^-\pi^+\pi^+$ decays, where the K^- has been misidentified as a π^-. Lastly, there is also a combinatorial background present in the normalisation fit. Backgrounds due to the decay of baryons, such as Λ_b are not considered since they are removed with particle identification requirements on the three pions from the nominal τ vertex.

As shown in Reference [30], an unbiased and model-independent measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ angular coefficients is possible with the current Run 1 + 2 LHCb dataset, providing a necessary number of signal events are retained after selection,

with a certain level of purity. The next chapter describes the selection steps utilised to achieve the required signal yield with higher purity than the required level in the methodology paper.

Part II

LHCb data analysis

5 EVENT SELECTION

The event selection process is designed to carefully choose the most relevant events that meet specific criteria. The initial stages of this process focus on the detector trigger, which aims to reduce the readout rate of the detector by only saving interesting events [53]. In order to achieve this, the trigger applies a set of criteria that are designed to identify events that are more likely to be of interest, based on various factors such as the energy or momentum of the particles involved. Once the trigger has been applied, the GRID-computing based selection is applied [74]. This step applies a more specific, but still very general set of loose selections criteria, which are designed to further reduce the number of irrelevant events. Specifically, these criteria help to eliminate combinatorial background, which refers to events where random combinations of particles satisfy previous selections. The GRID-computing based selection also helps to eliminate prompt background events, which refer to events of the type $B^0 \to D^* 3\pi X$. After the GRID-computing based selection has been applied, the next step is the 'offline selection'. This step is tailored specifically for the particular analysis and signal decay being studied, and involves the application of specific conditions and data treatments. The aim of this step is to further reduce the combinatorial background, as well as all the other backgrounds in the analysis, such as the large $B \to D^* D_s X$ background.

The data and simulation samples for the signal, normalisation, and control modes are handled identically for most of the selection workflow until the final stages of the offline selection. The normalisation mode is the fully reconstructed $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ decay. In this decay, the $\pi^+\pi^-\pi^+$ are produced at the B vertex, which differs from the signal decay. In the signal decay the $\pi^+\pi^-\pi^+$ are produced at the τ vertex, separated from the B meson vertex. There are several control mode studies with different control selections.

The fully reconstructed $B^0 \to D^{*-}(D_s^+ \to \pi^+\pi^-\pi^+)$ decay is used as a control mode to investigate data and simulation agreement in several variables. It exhibits a peak in the reconstructed mass of the $\pi^+\pi^-\pi^+$ system, $m(\pi^+\pi^-\pi^+)$ at the D_s mass and at the B^0 mass in $m(D^{*-}\pi^+\pi^-\pi^+)$. The signal selection is described in

the following sections, and differences between it and the normalisation and control mode selections are highlighted. The only exception to this treatment of the data and simulation is the data used to study the combinatorial background, which is composed of "Wrong Sign" (WS) candidates, where the reconstructed D^* and τ candidates have the same electric charge. These candidates are reconstructed with a dedicated algorithm in the LHCb GRID-computing pre-selection phase; the signal candidates are reconstructed using opposite-charge D^* and τ candidates. The treatment of the WS data then follows the same workflow as the rest of the data and simulation samples.

The selection workflow can be separated into two distinct sections. The 'online' selection is largely fixed and cannot be easily altered and not on a short timescale. It includes the geometrical acceptance of the LHCb detector, the real-time trigger selection, and conditions applied in the stripping and filtering phases. These selection stages occur 'online' during the data-taking process and subsequent GRID-computation. The 'offline' selection includes a set of additional criteria and selection techniques to further reduce background and increase signal purity.

This analysis requires large simulated datasets in order to model the shapes of the background in the signal fit. Some of the backgrounds have high branching fractions, but low efficiencies when the signal selection is applied. This requires large simulated samples to ensure suitable statistics in the signal fit. The samples generated for the Run I + II dataset required 5.6 billion generated events, where 199 million of these are saved to disk after a combination of unfiltered, stripping filtered, and custom filtering processes. In order to generate such a large quantity of events, the rapid simulation technique ReDecay is used to decrease simulation production time [75].

From the proton-proton collision rate of 40MHz, trillions of $b\bar{b}$ pairs are produced at LHCb each year. Due to the acceptance of the detector, not all of these collisions are recorded. The LHCb trigger selects interesting events and reduces the rate of recorded events by three orders of magnitude [53]. After additional selection criteria are applied 'online' utilising the GRID computing power, the number of candidates saved to disk for analysis is 25 million. This represents 419GB of data with each

candidate taking up 16kB. Further requirements enforced on the data reduce the number of candidates to $33,010 \pm 220$.

5.1 GEOMETRICAL ACCEPTANCE

Due to the size, shape, and location of the LHCb detector in relation to the pp collision point, the geometrical coverage of the detector is much less than 100%. In particular, particles must satisfy the angular requirement $10 < \theta < 300(250)$mrad in the bending (non-bending) plane. This geometrical acceptance affects all signal and background decays in the analysis and the value of the efficiency is measured at the simulation production stage. For the signal and normalisation modes, the geometric efficiency is around 16%, with full details given in Table 5.5.

5.2 TRIGGER

For the real data collected by LHCb, events must pass a subset of trigger conditions before they are saved. In order to measure the efficiencies of the trigger requirements for the signal and normalisation simulation datasets, these trigger requirements are postponed until later. This means there is a disparity in how the data and simulation are treated, however, the actual trigger conditions are not changed and the total selection efficiency is not impacted since it is the product of the efficiencies of each selection stage. This does mean that the measured efficiencies for the application of the trigger conditions in the simulation are not representative of the real data, since the trigger conditions are applied at a later stage of the selection for the simulation compared to the data. From the simulation, the trigger efficiencies are measured to be $60 - 90\%$ depending on the decay.

As previously stated, the trigger selection stage reduces the readout rate of the experiment and entails the application of several generic selection criteria. Many analyses of different decays may use the same trigger selection criteria and then apply more specific conditions later on in the analysis chain. The trigger consists of three stages: the Level 0 trigger decision (L0) is implemented in hardware, and requires high transverse energy (> 3.5 GeV) in the hadronic calorimeter of the

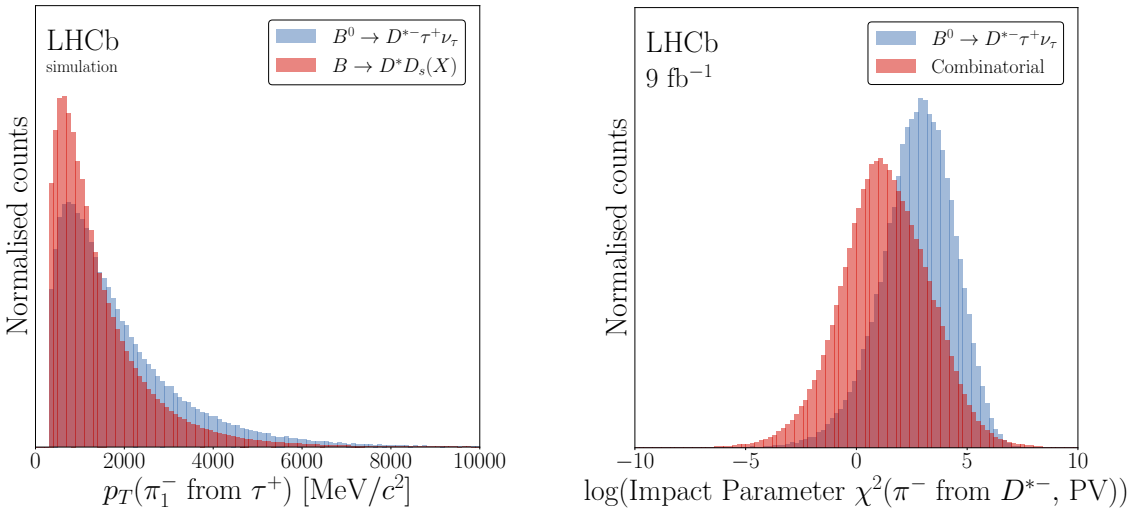


Figure 5.1: The distribution of the transverse momentum of the π^- from the τ^+ decay and the π^- from the D_s^+ decay in the signal and $B \to D^* D_s X$ simulation, respectively (left). The impact parameter normalised by its uncertainty for the π^- from the D^{*-} decay in the signal simulation and combinatorial background data (right).

LHCb detector, either from a signal decay product or another particle in the event. The next stage is the High-Level Trigger 1 (HLT1), which is the first software trigger. It first reconstructs the tracks produced by charged particles, then it selects tracks, and pairs of tracks, that have a high impact parameter with respect to the primary vertex, as well as high transverse momentum. The final trigger stage is the High-Level Trigger 2 (HLT2), where two, three, and four-track combinations are reconstructed and required to pass track and vertex quality conditions, and kinematic requirements.

These trigger conditions describe the signal decay well, where there are three-track vertices from the $\tau^+ \to \pi^+\pi^-\pi^+$ decay, and two, or three-track vertices from the $D^0 \to K^+\pi^-$ or $D^{*-} \to D^0(\to K^+\pi^-)\pi^-$ decays. The transverse momentum of the π^- from the τ^+ decay in the signal simulation can be seen in Figure 5.1, compared to the same distribution in $B \to D^* D_s X$ simulation. It can be seen that the peak of these distributions is around 1 GeV/c which satisfies the HLT1 condition of $p_T > 0.5\,\text{GeV}/c$. The impact parameter normalised by its uncertainty for the π^- from the D^{*-} decay is also shown in Figure 5.1 for the signal simulation and combinatorial background. The separation in these distributions shows how the impact parameter is a powerful variable to remove unwanted backgrounds.

The efficiencies for each stage of the trigger selection are shown in Table 5.5. One of the L0 conditions requires another particle in the decay, not coming from the signal B^0, to trigger the event. For example, a high momentum hadron produced in the decay of the other B hadron in the event. This category is thus least biased towards the signal since none of the signal tracks triggered the L0 decision.

5.2.1 L0 TRIGGER EFFICIENCY CORRECTIONS

In order to correct differences in the L0 trigger efficiencies for the data and simulation, the normalisation data set and simulation are compared. There are multiple L0 conditions used and the proportion of events in each L0 category are compared, and weights are calculated as

$$w_C = \frac{f_C(\text{Data})}{f_C(\text{Simulation})} \tag{5.1}$$

where C is either the TOS, TIS, or both L0 categories and f_C is the proportion of events in that category. The label TOS refers to events where an object in the signal decay triggered the L0 decision (Trigger On Signal) while TIS refers to events where objects not included in the signal decay, but still part of the event, triggered the L0 decision (Trigger Independent of Signal). The calculated weights for the three L0 categories are

$$w_{\text{TOS}} = 1.013 \qquad w_{\text{TIS}} = 1.090 \qquad w_{\text{both}} = 0.895$$

and are applied to the simulation samples when making templates for the signal fit. This process makes small adjustments to the L0 composition. The derivation of these weights is discussed in detail in Section 7.4

5.3 GRID-PROCESSING PRESELECTION

The next stage of the online selection is the stripping preselection which reduces the size of the signal and normalisation simulation to around $1.0 - 1.5\%$ of the size after the trigger selection. The stripping is a set of generic selection criteria that are usually used by more than one analysis, and a particular set of criteria within the stripping is known as a 'stripping line'. One stripping line is used for all

data and simulation samples except the WS data. This first stripping line creates a B^0 candidate by combining a $D^{*-} \to \overline{D}^0(\to K^+\pi^-)\pi^-$ candidate with a $\tau^+ \to \pi^+\pi^-\pi^+\overline{\nu}_\tau$ candidate. In the candidate reconstruction, loose requirements are placed on all tracks and reconstructed particles; these requirements are designed to be highly efficient on signal while rejecting combinatorial background from random track combinations. These requirements range from minimum p_T of tracks and reconstructed particles, to minimum vertex quality criteria.

The second stripping line is used to select the WS data and is identical to the first stripping line apart from the fact that it combines same-sign D^* and τ candidates to emulate combinatorial background. Additional criteria are also implemented at this stage in order to reduce offline processing time and disk space when the data is saved locally. These include tighter requirements on the transverse momentum of particles, vertex quality constraints, and a logical 'or' that selects events that have high flight distance significance (thus entering the signal fit) or a B^0 mass close to the fully reconstructed value (thus entering the normalisation fit).

This collection of selection criteria aims to reduce many backgrounds in the analysis. In particular, many of the vertex quality and transverse momentum conditions are tailored to reduce combinatorial background. Other vertex or impact parameter conditions aim to reduce the prompt background, $B \to D^*3\pi X$, where the 3π is produced at the B decay vertex, rather than a secondary vertex as in the signal decay. The flight significance requirement on the 3π vertex also heavily reduces the prompt background.

To begin, a $K^-\pi^+$ system is formed through requirements on the track momenta, vertex quality, ghost track probability, distance from the Primary Vertex (PV), and particle identity. A $\overline{D}^0$ candidate is created from this system though additional requirements on the difference between the system invariant mass and the known mass of the $\overline{D}^0$, the Distance Of Closest Approach (DOCA) of the K^- and π^+, and production vertex quality. Similar track quality requirements are applied to the π^- from the D^{*-} and a D^{*-} candidate is formed from the $\overline{D}^0\pi^-$ system by requirements on the vertex quality, the momentum of the system, the $m(D^* - D^0) \equiv m(D^{*-}) - m(\overline{D}^0)$, and the mass of the D^{*-}.

The $\tau^+ \to \pi^+\pi^-\pi^+\overline{\nu}_\tau$ side of the decay is reconstructed from a $\pi^+\pi^-\pi^+$ system, where each pion has requirements on its transverse momentum, impact parameter χ^2 with respect to PVs in the event, track quality, and particle identity. The $\pi^+\pi^-\pi^+$ system itself has requirements on the reconstructed mass $m(\pi^+\pi^-\pi^+)$, track quality, the angle between the particle momentum vector and the line of flight from the best PV to the particle decay vertex, the DOCA of each $\pi^\pm\pi^\pm$ pair, the min$[m(\pi^+\pi^-)]$, impact parameter χ^2 with respect to PVs in the event, and the transverse momentum. Finally, the B^0 candidate is formed by combining the D^{*-} and τ^+ candidates with requirements on the difference between the reconstructed invariant mass and the known mass of the B^0, the angle between the particle momentum vector and the line of flight from the best PV to the particle decay vertex, and the DOCA between the D^{*-} and the τ candidates. The combined efficiency of the stripping criteria and the additional conditions is shown in Table 5.5.

5.4 LOCAL-PROCESSING PRESELECTION

The total offline selection reduces the dataset from 25,571,955 events to $33,010 \pm 220$ reported in Table 5.4. The first stage of the offline selection is a set of conditions in a number of variables, summarised in Table 5.1. Due to the partially reconstructed nature of the signal decay, tight windows on the τ and B^0 masses cannot be set; however, limits of $m(\pi^+\pi^-\pi^+) \in [500, 1700]$ MeV/c^2 and $m(B^0) < 4700$ MeV/c^2 are used. The D^{*-} candidate is loosely constrained with requirements on $m(D^* - D^0) \equiv m(D^{*-}) - m(D^0) \in [140,\ 160]$ MeV/c^2 and $|m(D^0) - m(D^0(\text{PDG}))| < 100$ MeV/c^2.

Table 5.1: The list of the requirements applied in the first offline selection phase to $B^0 \to D^{*-}\tau^+\nu_\tau$ candidates. The efficiencies for these requirements are reported in Table 5.5.

Variable	Condition
$m(\pi^+\pi^-\pi^+)$	$\in [500,\ 1700]$ MeV/c^2
$m(B^0)$	< 4700 MeV/c^2
$m(D^* - D^0) \equiv m(D^{*-}) - m(D^0)$	$\in (142,\ 148)$ MeV/c^2
$\lvert m(D^0) - m(D^0(\text{PDG}))\rvert$	< 40 MeV/c^2
PV(τ^+)	$=$ PV($\overline{D}^0$)
τ^+ flight significance from IP	> 10
τ^+ flight significance from B^0 vertex	$\in [3, 20]$

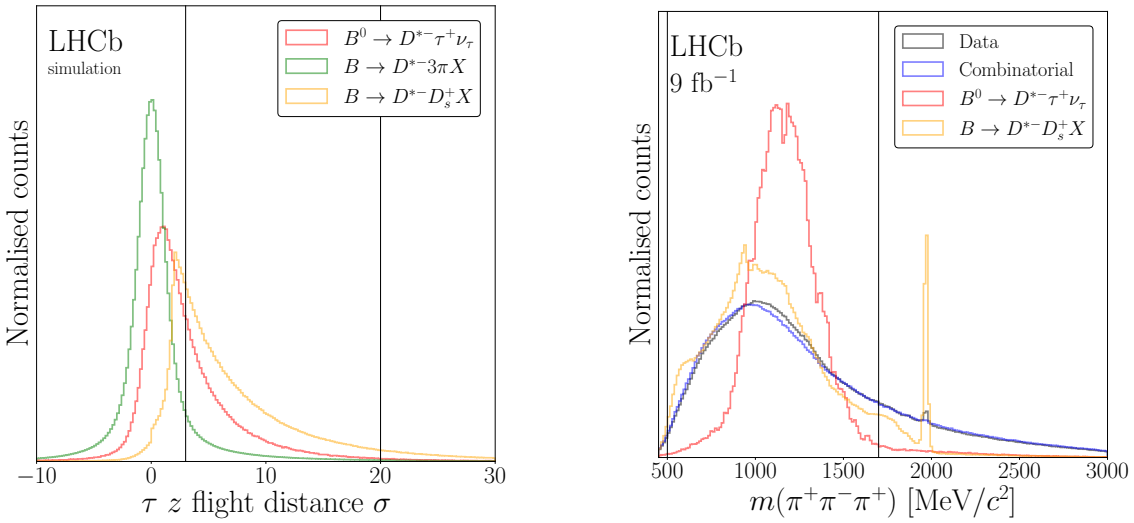


Figure 5.2: The tau flight significance (left) and $m(\pi^+\pi^-\pi^+)$ (right) distributions for several data and simulation samples. The tau flight significance distribution compares the signal, prompt and $B \to D^* D_s X$ simulation samples. The $m(\pi^+\pi^-\pi^+)$ distribution compares the data, combinatorial data, signal simulation, and $B \to D^* D_s X$ simulation samples. The black vertical lines in each figure show the upper and lower bounds of the signal region.

Further reduction in the combinatorial background from fake D^* candidates that fall within these mass windows is discussed in Section 5.7. The PV associated with the τ candidate is required to be the same as the PV associated with the D^0 candidate. The τ vertex is required to lie more than 10σ downstream of the PV, where σ is the combined uncertainty on the τ vertex and the PV. Similarly, the τ vertex is required to lie 3σ downstream of the B^0 vertex in the z direction, where σ is the combined uncertainty on the reconstructed τ and B^0 candidate vertices. These two criteria relating to the τ vertex are not applied for the normalisation fit; additionally, the τ and B^0 mass windows are altered as described in Section 5.8. The efficiencies for these criteria are fully reported in Table 5.5. In general, the efficiency of these requirements on the signal and normalisation simulation is very high.

Distributions for some of the selection variables can be seen in Figure 5.2. These show different combinations of data and simulation samples to highlight the differences between the signal simulation and certain backgrounds that we aim to remove. In the tau flight significance distribution, the signal simulation is compared to the $B \to D^* D_s X$ simulation and the inclusive $B \to D^* 3\pi X$ simulation. The later sample contains many prompt events where the 3π is produced at the B decay vertex.

In the tau flight significance distribution, it can be seen that the prompt sample has a larger tail at low flight and a significant proportion of these events are removed with the selection condition. $B \to D^* D_s X$ events are also removed with this requirement, but due to the longer lifetime of the D_s, the efficiency for this sample with this requirement applied is actually higher than the signal efficiency. It can be seen that the signal, prompt, and $B \to D^* D_s X$ distributions have different shapes within the signal region. This is important since this distribution is used in the signal fit and the differences in these distributions are exploited to measure the relative fractions of the signal and background. The $m(\pi^+\pi^-\pi^+)$ distributions for the data, combinatorial data, signal simulation, and $B \to D^* D_s X$ simulation are also shown. It can be seen that the significant contribution of the fully reconstructed $D_s^+ \to \pi^+\pi^-\pi^+$ background is removed with the selection requirements. The large tails in this distribution for the data and combinatorial data are removed with the selection requirements, at the expense of a significantly smaller tail in the signal distribution. The signal distribution peaks at a value much lower than the measured τ mass. This is due to the partial reconstruction of the τ, where the neutrino in the decay is not reconstructed.

5.5 MULTIVARIATE ANALYSIS

The next stage of the selection uses three MVA algorithms to enhance signal purity by suppressing three types of backgrounds. Specifically, Boosted Decision Trees (BDT) use properties of the signal candidate and underlying event to reject background. The BDT training is performed using the scikit-learn and XGBoost packages [76]. Signal and background samples are required as well as a list of discriminating variables to distinguish between them. For all three BDT types, the $B^0 \to D^{*-}\tau^+(\to \pi^+\pi^-\pi^+)\nu_\tau$ simulation is used as the signal sample. The sample is truth-matched by requiring the reconstructed τ and D^* candidates to be associated with true generated τ and D^* particles, and that both of these particles originate from the same B^0 meson. The background samples differ between BDTs and are described in the relevant sections below. The selection up to this point is applied

to both the signal and background samples prior to the training, in order to focus the BDTs on background events that most resemble the signal.

The input variables are selected and curated in multiple stages. Initially, a set of variables is selected by measuring the separation of signal and background in the full set of variables and ranking them from most to least separating. The degree of separation is measured using the figure of merit

$$d = \frac{|\overline{x}_1 - \overline{x}_2|}{s}, \quad \text{with} \quad s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}, \tag{5.2}$$

where $\overline{x}_1$ and $\overline{x}_2$ are the means and s_1 and s_2 are standard deviations of distributions 1 and 2, s is the pooled standard deviation, and n_1 and n_2 are the sample sizes [77]. The list of ranked variables is then inspected, and variables are rejected if they are correlated with any previously accepted variables in the signal or background distributions such that

$$|P_{x,y}| \equiv \left| \frac{cov(x, y)}{\sigma_x \sigma_y} \right| > 0.99. \tag{5.3}$$

This step eliminates variables that bring no additional orthogonal information and also removes any duplicate information. The three BDTs are then trained as described below and their outputs are calculated for all data and simulation samples. For each BDT type, two separate MVA algorithms are trained. The data and simulation samples used to train the BDTs are split into two halves, where the first half is used to train the first BDT, and the second half is used to train the second BDT. When the BDTs are applied to the data and simulation, events that were used to train the first BDT have the second BDT applied to them, and vice versa for the events used to train the second BDT. This technique is known as cross training and it allows the full use of data and simulation statistics without applying the BDT to the same events used for training the BDT. For simulation and data samples that were not used in the BDT training, a random choice of the first or second BDT is made for each event. Having good agreement between the two BDT output distributions is crucial to prevent systematic shape differences between data and simulation outputs. To ensure that there are no significant differences in the input variable distributions used to train the two BDTs, a check is performed. This in-

Table 5.2: The hyper-parameters used to train the selection BDTs. The average of the χ^2 for the signal and background distributions is also shown, where the χ^2 values are calculated from the ratio of the test output distributions for the first and second BDTs.

BDT	Max Depth	Learning Rate	Max Iterations	Average χ^2
Combinatorial	4	0.80	2140	0.947
Tau	3	0.95	1460	0.986
Isolation	2	0.90	4940	1.059

volves calculating the ratios of the input variable distributions for the two samples, with a ratio of 1.0 indicating agreement between the distributions.

A check is performed to ensure agreement between the output distributions of the first and second BDTs. The ratio of the testing BDT distributions is calculated and a ratio of 1.0 indicates agreement between the BDTs. To optimize performance, a scan was conducted to determine the best hyper-parameters for the training process, including the learning rate and maximum tree depth. The learning rate was set to 0.05 and increased incrementally up to 1.0, with maximum tree depths ranging from 2 to 4. A total of 75 hyper-parameter combinations were tested per BDT type, with 10,000 iterations used in the training process. The set is reduced by excluding trainings where the test loss has not reached a minimum during the training based on the test and train loss metric calculated.

The remaining trainings are studied by analyzing the loss curves, and an optimal number of iterations is chosen for each set of hyper-parameters based on the minimal test loss. A second round of training is then performed for the remaining combinations of learning rate and max depth to maximize the BDTs' performance. The test output distributions are compared, and the hyper-parameters that minimize the χ^2 of the ratio compared to a flat line at 1.0 are chosen as the selection BDTs for the analysis. Table 5.2 summarises the chosen hyper-parameters for each BDT type and quotes the average χ^2 for the signal and background BDT ratios.

5.5.1 COMBINATORIAL BACKGROUND SUPPRESSION

The goal of the first BDT is to remove combinatorial background events, where the reconstructed τ and D^* candidates originate from different parent vertices. The background sample used to train this BDT is the WS data, where a reconstructed τ

and D^* of the same charge are combined to form a B^0 candidate. This sample, and the signal sample, are selected via all previously described criteria. The variables used to train this BDT contain track and vertex quality information, as well as momentum information for the particles. The testing distributions are created by applying the first BDT to the sample that was used to train the second BDT and vice versa. Due to the cross-training, only the testing distributions for each BDT are used later in the analysis. These testing distributions are shown in Figure 5.3 as well as the ratio between them. The flat distribution in the ratio indicates that the BDT output distributions agree between the two BDTs.

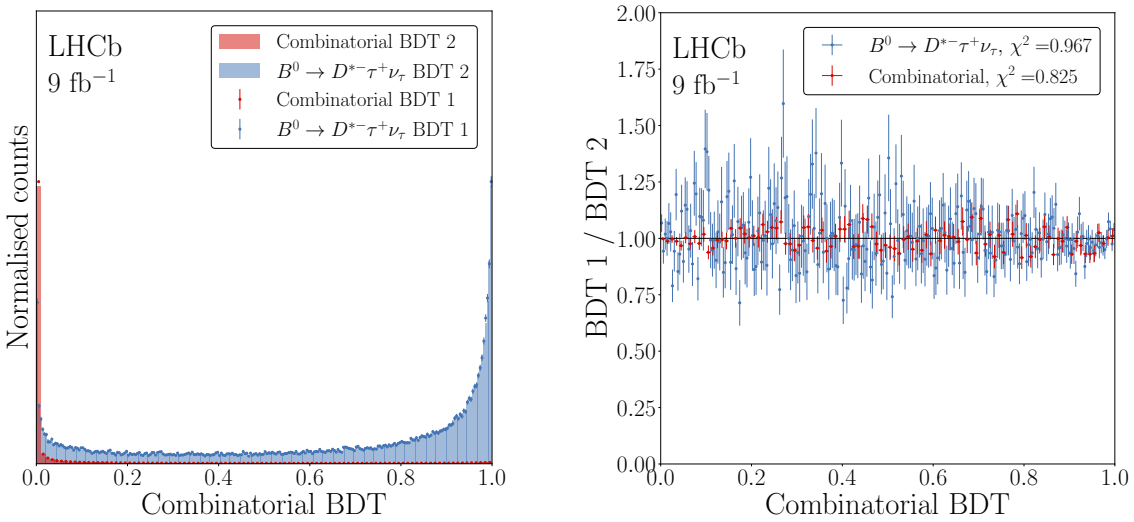


Figure 5.3: The testing distributions of the two BDTs (left) show good agreement. This can be seen in the ratio of the two distributions (right).

5.5.2 τ vs. D_s enhancement

The second BDT aims to reject $B \to D^{*-}D_s^+(X)$ events. This background is reduced considerably using the following MVA selection, though remains the largest component of the signal fit after the full selection. The background sample used to train this BDT is composed of a mixture of truth-matched $B \to D^{*(*)}D_s(X)$ simulation samples, where $B \in \{B^0,\ B^\pm,\ B_s\}$. This BDT aims to exploit the differences between the signal and background samples in the three-prong $\pi^+\pi^-\pi^+$ system produced in τ^+ and D_s^+ decays. The training variables consist of momentum and track quality information for the reconstructed τ candidate and the three pions.

More separating variables, such as the $m(\pi^{\pm}\pi^{\pm})$ invariant-mass combinations, are not used in the training to avoid strong correlations between these variables and the τ vs. D_s BDT response. These variables are used in the control study detailed in Section 7.1 so avoiding strong correlations between the BDT output and the variables significantly reduces biases in their distributions when applying a restriction on the BDT output. The testing distributions are shown in Figure 5.4 as well as the ratio between them. The flat distribution in the ratio indicates that the BDT output distributions agree between the two BDTs.

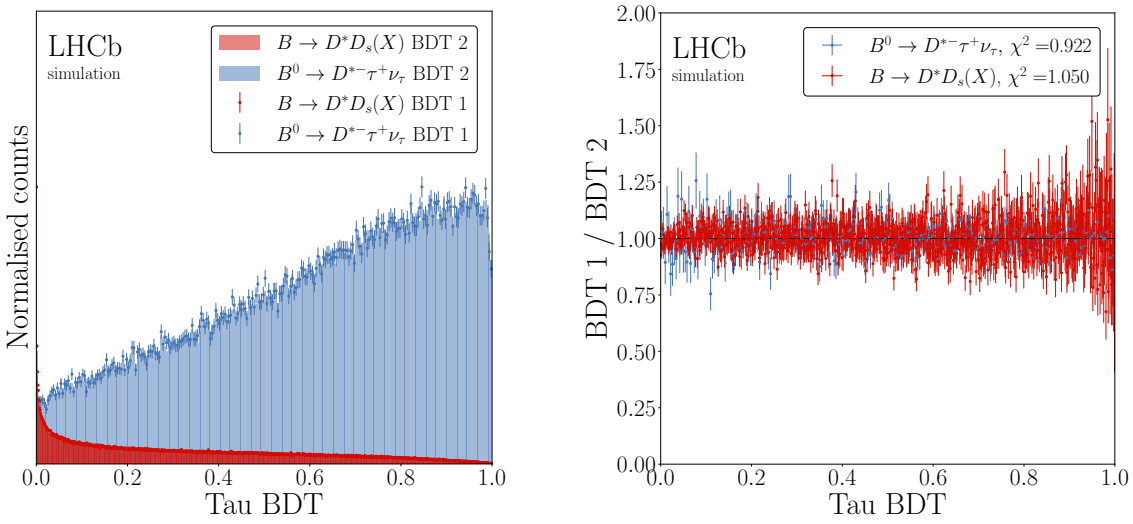


Figure 5.4: The testing distributions of the two BDTs (left) show good agreement. This can be seen in the ratio of the two distributions (right).

5.5.3 ISOLATION

The third BDT is designed to further suppress $B \to D^{*-}D_s^+(X)$ events, through the use of isolation variables. The background training sample is the same as the τ vs D_s BDT: $B \to D^{*(*)}D_s(X)$ simulated samples, where $B \in \{B^0,\ B^{\pm},\ B_s\}$. The variables used to train this BDT are known as isolation variables. They are constructed by drawing a cone in three-dimensional space around reconstructed particle candidates and measuring various quantities related to the additional tracks or neutral particles found within the cone. These variables include the number of tracks or neutral particles in the cone and p_T asymmetries. The testing distributions

are shown in Figure 5.5 as well as the ratio between them. The flat distribution in the ratio indicates that the BDT output distributions agree between the two BDTs.

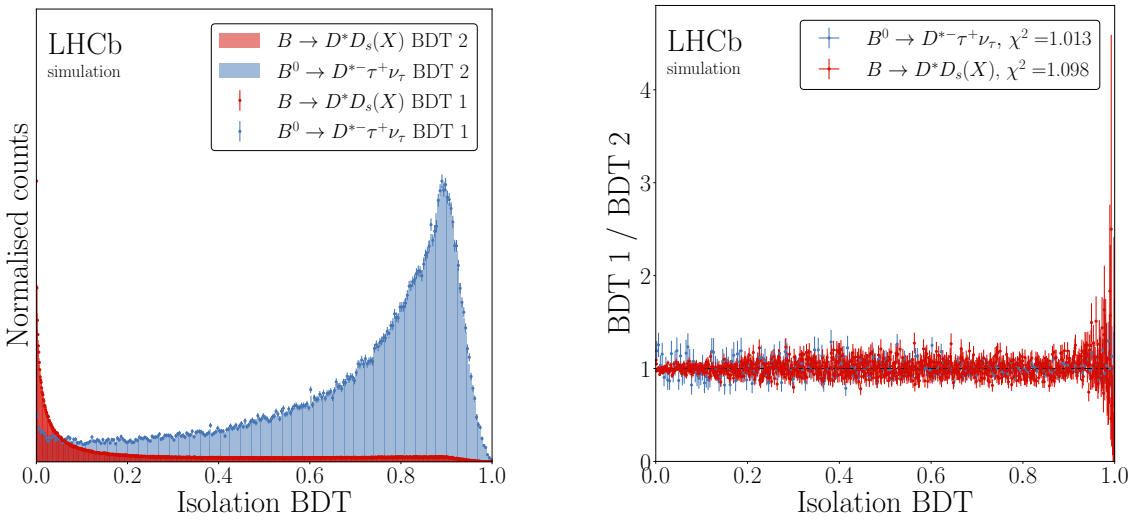


Figure 5.5: The testing distributions of the two BDTs (left) show good agreement. This can be seen in the ratio of the two distributions (right).

5.5.4 NUMERICAL METHOD FOR MVA OPTIMISATION

The requirements on the three BDT outputs are optimised simultaneously using a numerical method. The aim of the optimisation is to maximise the metric defined as the product of the signal fraction and the signal significance

$$m = \frac{N(B^0 \to D^{*-}\tau^+\nu_\tau)}{N(\text{Data})} \cdot \frac{N(B^0 \to D^{*-}\tau^+\nu_\tau)}{\sqrt{N(\text{Data})}} \tag{5.4}$$

where $N(\text{data})$ is the number of events in the data sample used in the optimisation after all requirements have been applied, and $N(B^0 \to D^{*-}\tau^+\nu_\tau)$ is the number of expected signal events in the data sample. This is estimated using the result of the normalisation fit (see Chapter 6) and scaling the normalisation yield for the difference in branching ratio and efficiency for the selection:

$$N(B^0 \to D^{*-}\tau^+\nu_\tau) = N(B^0 \to D^{*-}\pi^+\pi^-\pi^+) * \frac{\epsilon(B^0 \to D^{*-}\tau^+\nu_\tau)}{\epsilon(B^0 \to D^{*-}\pi^+\pi^-\pi^+)} * r_{\mathcal{B}} \tag{5.5}$$

where the ϵ terms represent the full signal and normalisation mode efficiencies derived from simulation, and

$$r_{\mathcal{B}} = \frac{\mathcal{B}(B^0 \to D^{*-}\,\tau^+\,\nu_\tau) * [\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+\overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau)]}{\mathcal{B}(B^0 \to D^{*-}\,\pi^+\pi^-\pi^+)}, \tag{5.6}$$

where the value used for $\mathcal{B}(B^0 \to D^{*-}\,\tau^+\,\nu_\tau)$ is the world average value, multiplied by a random number in the range [0.75, 1.25] in order to blind the optimisation to the true value. During the BDT training, the $\tau^+ \to \pi^+\pi^-\pi^+\overline{\nu}_\tau$ signal mode is used as the signal sample in order to maximise the selection performance for this mode. In reality, the signal sample is composed of a mixture of the $\tau^+ \to \pi^+\pi^-\pi^+\overline{\nu}_\tau$ and $\tau^+ \to \pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau$ decays; as such, the π^0 mode is also considered in the signal yield estimation. The signal efficiency used is a weighted average of the two contributing τ decay modes, taking into account their relative branching fractions and efficiencies.

The normalisation fit as described in Chapter 6 is performed on the dataset with the selection applied, but with no BDT requirements. This fit provides a value for the normalisation yield with no BDT selection applied. This value is then scaled when BDT conditions are applied, using the efficiency for the BDT requirements when applied to the normalisation simulation. Although there is reasonable data/simulation agreement in the BDT outputs, this method will not be completely accurate. Therefore, after the optimisation has finished, the normalisation fit is rerun with the optimised BDT requirements applied, in order to determine the signal purity precisely. This value will still not be exact, due to the blinding of the $\mathcal{B}(B^0 \to D^{*-}\,\tau^+\,\nu_\tau)$ value.

Each time the metric optimisation function is called for a different set of BDT requirements, the signal efficiency is calculated, the data has the requirements temporarily applied, and the normalisation yield and efficiency are calculated. This allows the optimisation metric to be calculated and subsequently maximised by the numerical method. The optimisation is penalised if the BDT requirement values result in the estimated signal yield being lower than a target signal yield, $n_{\text{signal}}^{\text{target}}$, by multiplying the optimised metric by a factor of

$$1 + 0.75 * \log(n_{\text{signal}}/n_{\text{signal}}^{\text{target}}) \tag{5.7}$$

where the value of $n_{\text{signal}}^{\text{target}}$ was chosen to be 7500. This value is motivated by Reference [30] where a signal yield of 8000 was used, but with lower purity than that achieved

by this selection. The optimisation finds BDT requirement values of

$$\text{Combinatorial BDT} > 0.16$$

$$\tau \text{ vs } D_s \text{ BDT} > 0.32$$

$$\text{Isolation BDT} > 0.71$$

The signal selection as described above results in an estimated 7600 ± 2000 signal events in the full 9 fb^{-1} Run 1 + 2 LHCb dataset, with a signal purity of $(18.7 \pm 4.9)\%$, or a signal fraction of $(15.8 \pm 4.1)\%$. The uncertainties on these values come from the range of possible values after dividing by the range of possible blinding factors for $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$. The purity is further improved with requirements on the particle identity of the $\pi^\pm$ from the $\pi^+\pi^-\pi^+$ system, discussed in Section 5.6, and the removal of fake D^* candidates as described in Section 5.7.

5.6 PARTICLE IDENTIFICATION SELECTION

The next stage of the selection involves a series of requirements on the particle identity of the $\pi^\pm$ in the $\pi^+\pi^-\pi^+$ system, as well as requirements on their momentum and transverse momentum. The p and p_T requirements are required to align the data and simulation with the calibration samples used by the PIDCalib2 package [78], which was used to calculate corrective weights for the simulation and is an update to the widely used PIDCalib package [79]. These weights are used to correct the efficiency of the particle identity requirements on the simulation. All three pions have requirements of $p \in [3, 100]$ GeV/c and $p_T \in [0.3, 10]$ GeV/c applied. The two pions with the same charge have a loose PIDK < 4.0 requirement applied to them while the single pion with an opposite charge has a much stricter PIDK < 1.0 requirement applied, where the PIDK variable is calculated from the difference in the log-likelihood of a particle being a kaon and the log-likelihood of a particle being a pion.

5.7 ISOLATING GENUINE D^* CANDIDATES

After the above selection has been applied to the data and simulation, there is still some considerable combinatorial background remaining. This is largely due to the loose mass windows on the $m(D^* - D^0)$ and $m(D^0)$ variables. The loose values of these requirements enable a two-dimensional fit to be performed which measures the fake D^0 and fake D^* combinatorial background rates. By measuring these background yields in an invariant-mass fit, ${}_s$Weights can be calculated and used to remove these components from the final signal fit [80].

First, an unbinned two-dimensional fit in $m(D^* - D^0)$ and $m(D^0)$ is performed on the signal simulation, with the full selection applied. The $m(D^0)$ distribution is modelled using a sum of two Crystal Ball functions, while the $m(D^* - D^0)$ distribution is described using the sum of two Crystal Ball functions and a Gaussian function. The 2D distribution is modelled using a product of the two one-dimensional functions. The one-dimensional projections of this fit can be seen in Figure 5.6. The fit is then performed in data, where the genuine D^* shape is constrained from the previous simulation fit: the tail parameters and internal PDF fractions are Gaussian constrained to the values found in the simulation fit, while the mean and width parameters vary freely. Two background shapes are used to account for different combinatorial sources, fake D^0 and fake D^*. The fake D^0 component arises from events where unrelated K^- and π^+ are combined and does not peak in either fit variable distribution; this contribution is modelled using exponential functions both in $m(D^0)$ and $m(D^* - D^0)$. The fake D^* component arises from events where genuine D^0 candidates are combined with an unrelated π^-, resulting in a peak in the $m(D^0)$ distribution but no peak in the $m(D^* - D^0)$ distribution. This contribution is modelled using the same PDF as the signal component in $m(D^0)$, and a `RooDstD0BG` PDF in $m(D^* - D^0)$. This information is summarised in Table 5.3, and the fit results to data are shown in Figure 5.7.

The measured rate of combinatorial background due to fake D^0 and fake D^* events was measured to be 9%, with full details given in Table 5.4. Using the results of the 2D fit to data, ${}_s$Weights were calculated to isolate the genuine D^* component in the data [80]. The distributions of several high-level variables in data are shown in

Table 5.3: The PDF information for the components in the genuine D^* measurement fit as well as whether each component peaks in the relevant variable. The total PDF in the 2D fit for each component is the product of the two one-dimensional PDFs.

Component	$m(D^* - D^0)$		$m(D^0)$	
	Peak	PDF	Peak	PDF
Genuine D^*	✓	Sum of two Crystal Ball PDFs and a Gaussian	✓	Sum of two Crystal Ball PDFs
Fake D^*	X	`RooDstD0BG` PDF	✓	Sum of two Crystal Ball PDFs
Fake D^0	X	Exponential	X	Exponential

Table 5.4: The results of the two-dimensional fit to data which measures the genuine D^* component of the data sample. The number of genuine D^* candidates is reported as well as the combined number of fake D^0 and fake D^* candidates. The fraction of fake candidates that are fake D^* candidates is also reported.

Fit Variable	Fit Result
N(Genuine)	$33,010 \pm 220$
N(Fake)	6910 ± 150
f(Fake D^*)	$(54.7 \pm 1.4)\%$

Figures 5.8 and 5.9, and the same distributions after the application of the ${}_s$Weights in Figures 5.10 and 5.11, to illustrate the background reduction achieved by applying the ${}_s$Weights.

There are other sources of combinatorial background that are not accounted for in this stage, namely genuine D^* candidates that are combined with a $\pi^+\pi^-\pi^+$ system where one or more of the $\pi^\pm$ come from an unrelated event, or events where a genuine D^* is combined with a genuine $\pi^+\pi^-\pi^+$ system, but the D^* and $\pi^+\pi^-\pi^+$ system come from different B decays. These backgrounds are accounted for and discussed in Section 8.3.

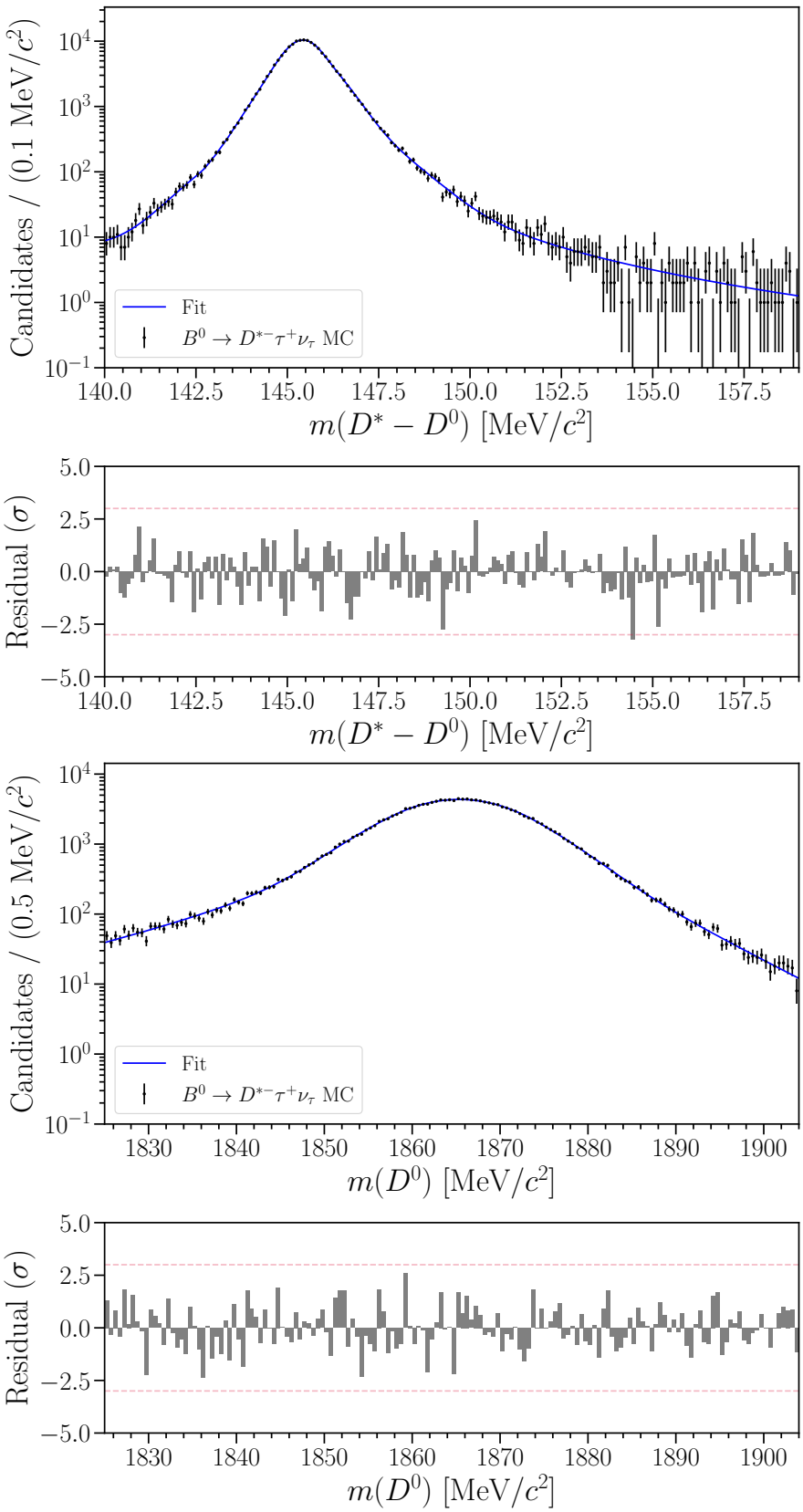


Figure 5.6: The one-dimensional projections of the $m(D^* - D^0)$ (top) and $m(D^0)$ (bottom) variables used in the two-dimensional fit to isolate genuine D^* candidates. This is the fit to the $B^0 \to D^{*-}\tau^+\nu_\tau$ simulation where the samples for all years have been combined.

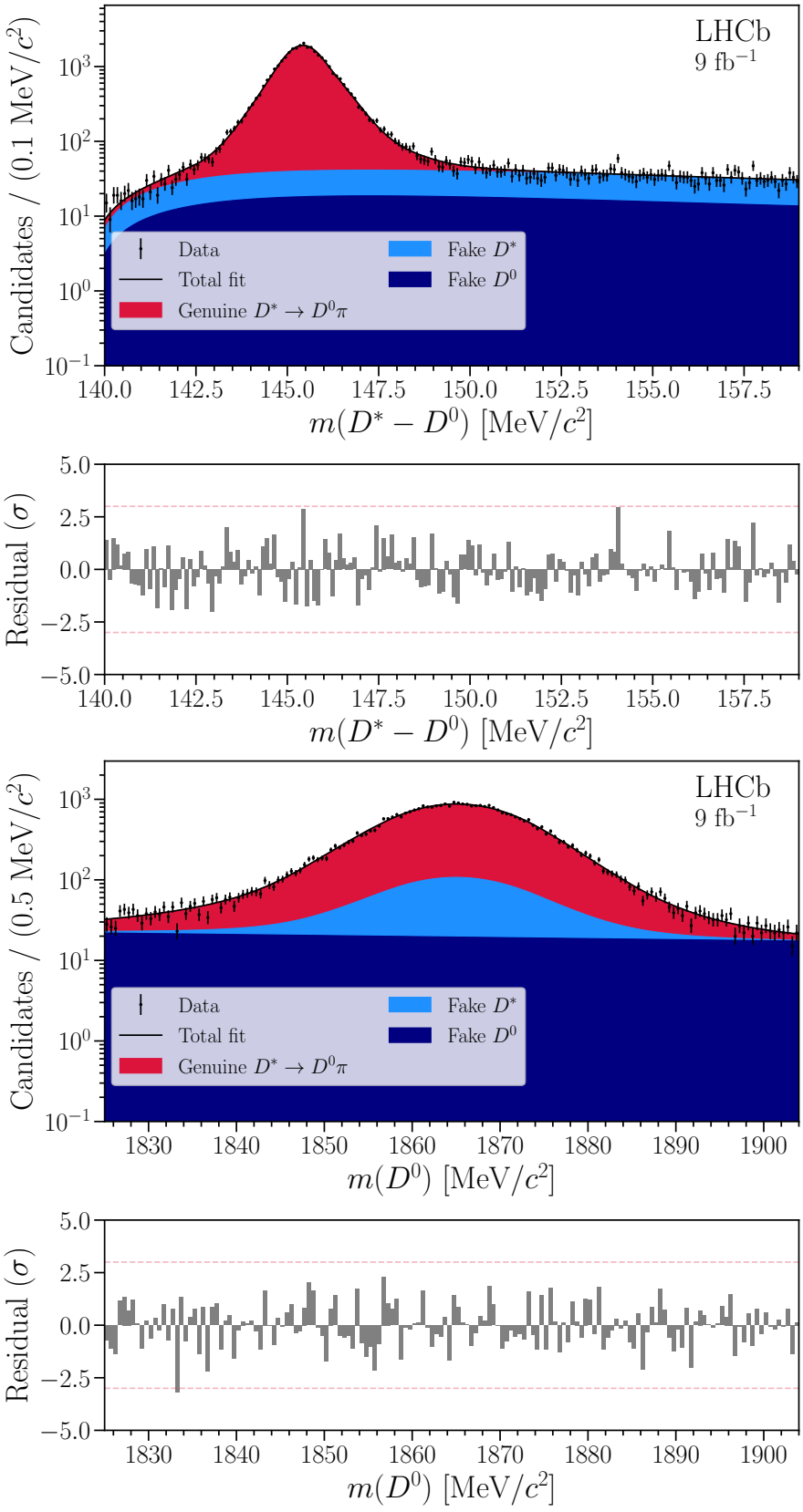


Figure 5.7: The one-dimensional projections of the $m(D^*-D^0)$ (top) and $m(D^0)$ (bottom) variables used in the two-dimensional fit to isolate genuine D^* candidates. This is the fit to the data sample with the full selection applied and the samples for all years have been combined.

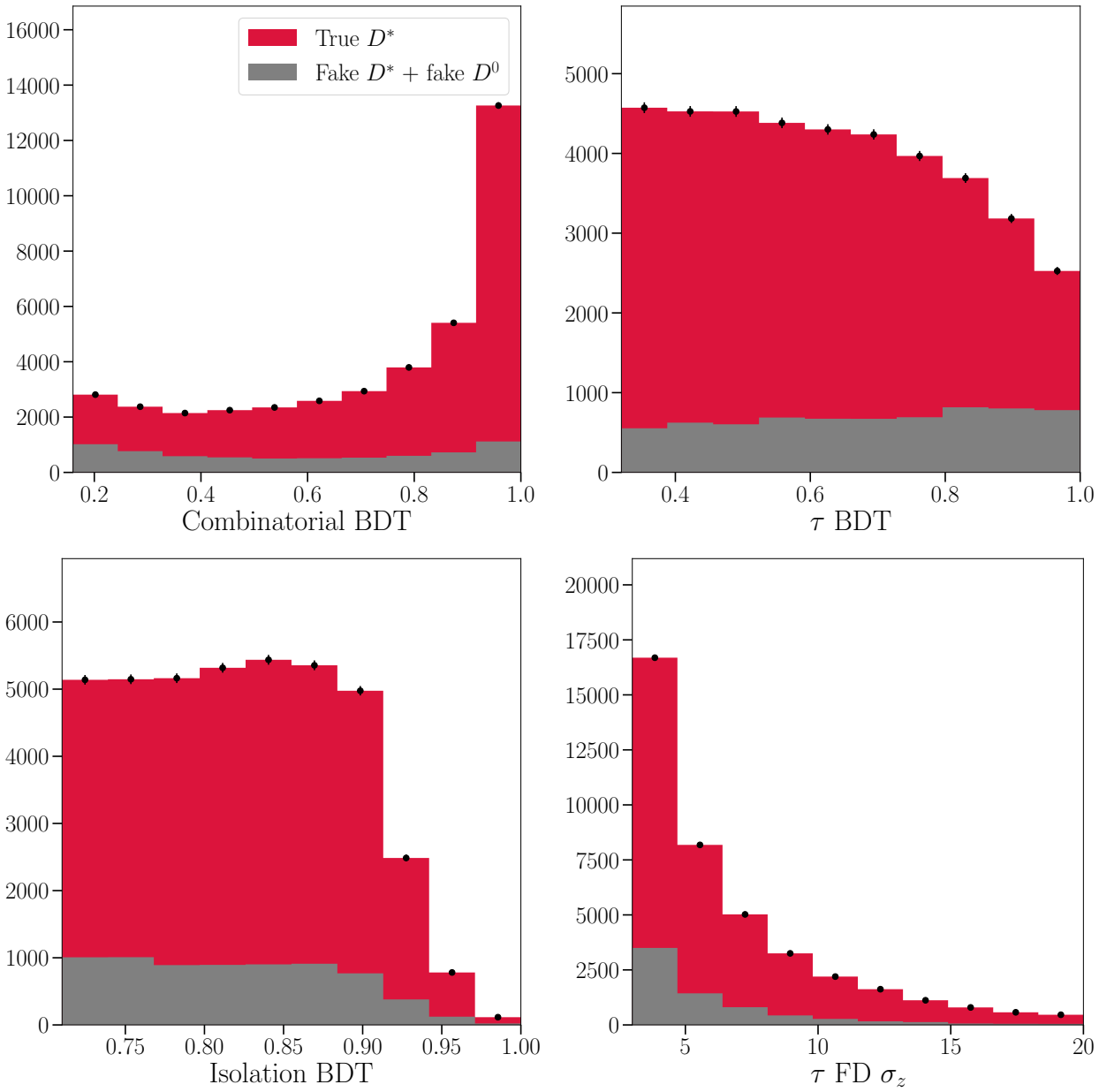


Figure 5.8: High-level variable distributions in data (points) for candidates passing all selection requirements and considered in the 2D $m(D^* - D^0)$ vs. $m(D^0)$ fit. The $_s$Weighted true D^* contribution is shown in red, while the combined fake D^0 and fake D^* contributions are shown in grey.

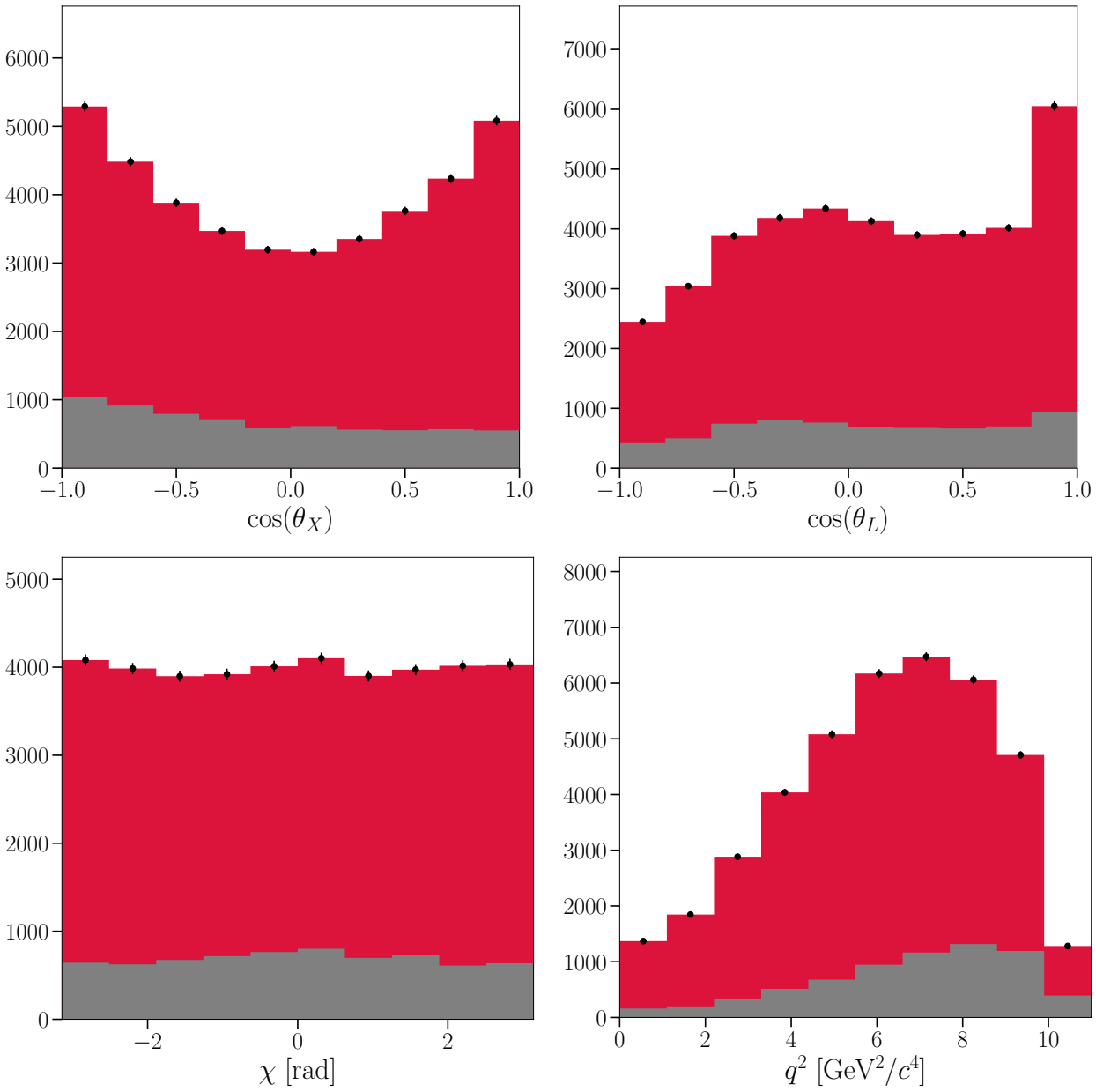


Figure 5.9: High-level variable distributions in data (points) for candidates passing all selection requirements and considered in the 2D $m(D^* - D^0)$ vs. $m(D^0)$ fit. The ${}_s$Weighted true D^* contribution is shown in red, while the combined fake D^0 and fake D^* contributions are shown in grey (continued).

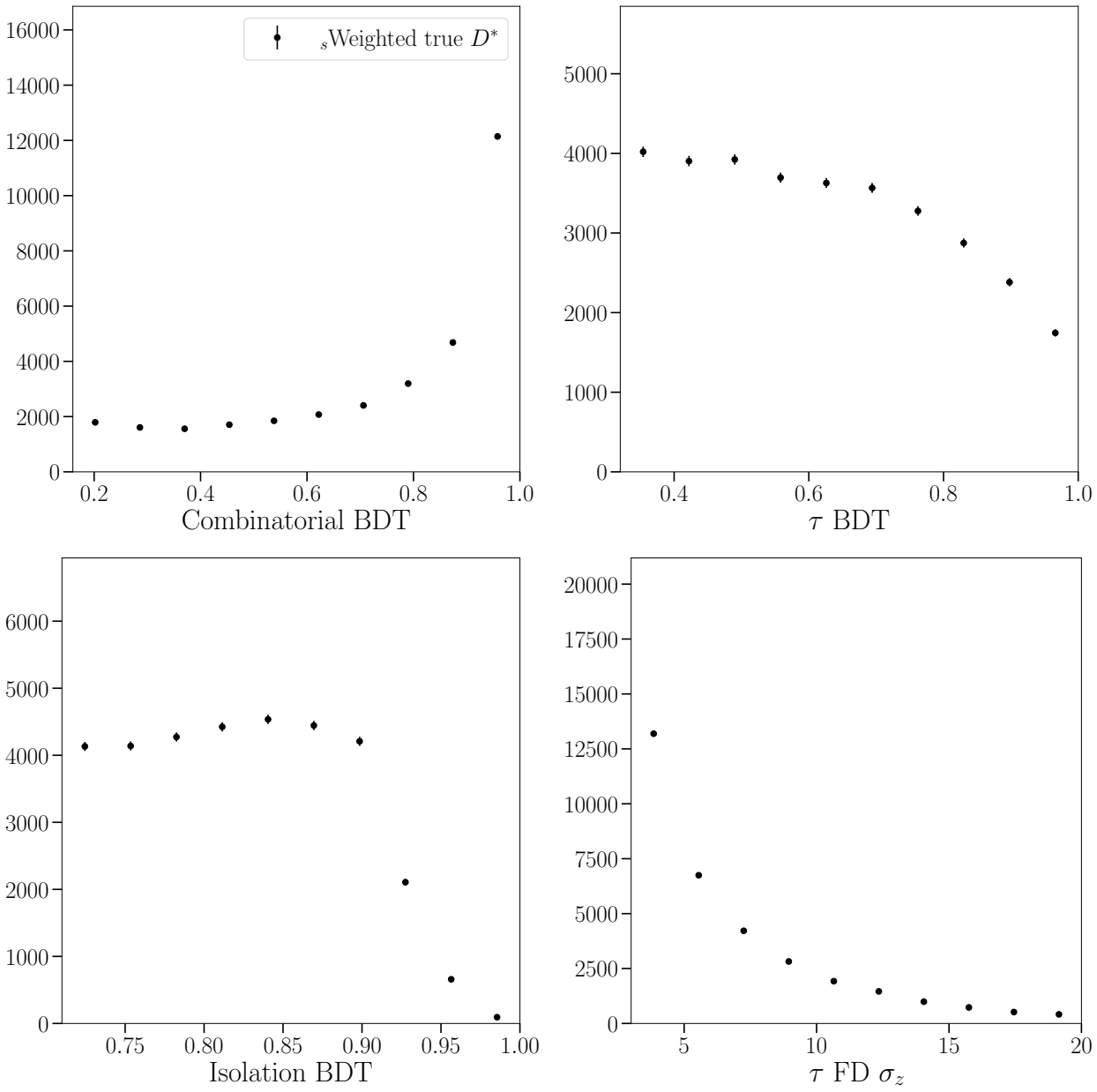


Figure 5.10: High-level variable distributions in data for $_s$Weighted true D^* candidates.

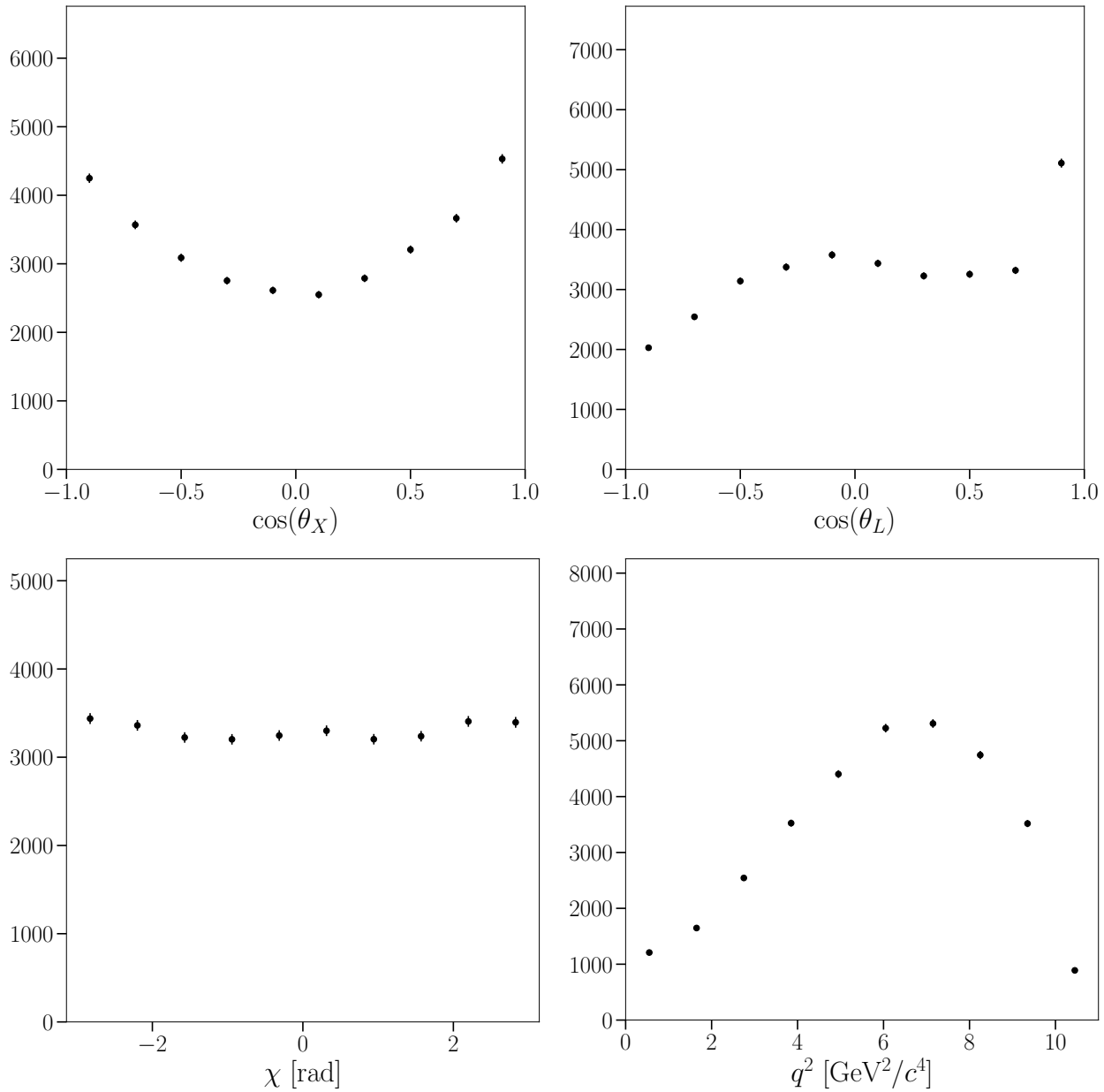


Figure 5.11: High-level variable distributions in data for ${}_s$Weighted true D^* candidates (continued).

5.7.1 DATA STABILITY UNDER $_s$WEIGHT VARIATIONS

The $_s$Weights calculated on data above, which are used to statistically subtract the fake D^* and fake D^0 combinatorial backgrounds, have an associated uncertainty arising from finite knowledge of the signal and background 2D fit PDFs. To determine this uncertainty, all PDF parameters from the 2D fit are varied within their uncertainties according to the fit covariance matrix, and a new set of $_s$Weights is derived using the varied PDFs. This process is repeated 100 times, to form a distribution of possible $_s$Weight values for each data event.

To assess the impact of the $_s$Weight uncertainty on the angular fit distributions, the data are binned with the default $_s$Weights applied as shown in Figures 5.12 and 5.13. The data are also binned with the 100 varied sets of $_s$Weights applied, and the variation in bin content values is displayed as filled blue rectangles. In general, the $_s$Weights variation causes changes in bin content well below the statistical uncertainty in each bin. As such, the $_s$Weighted data distributions used in the angular fit are not considered to be subject to a systematic uncertainty arising from the use of $_s$Weights.

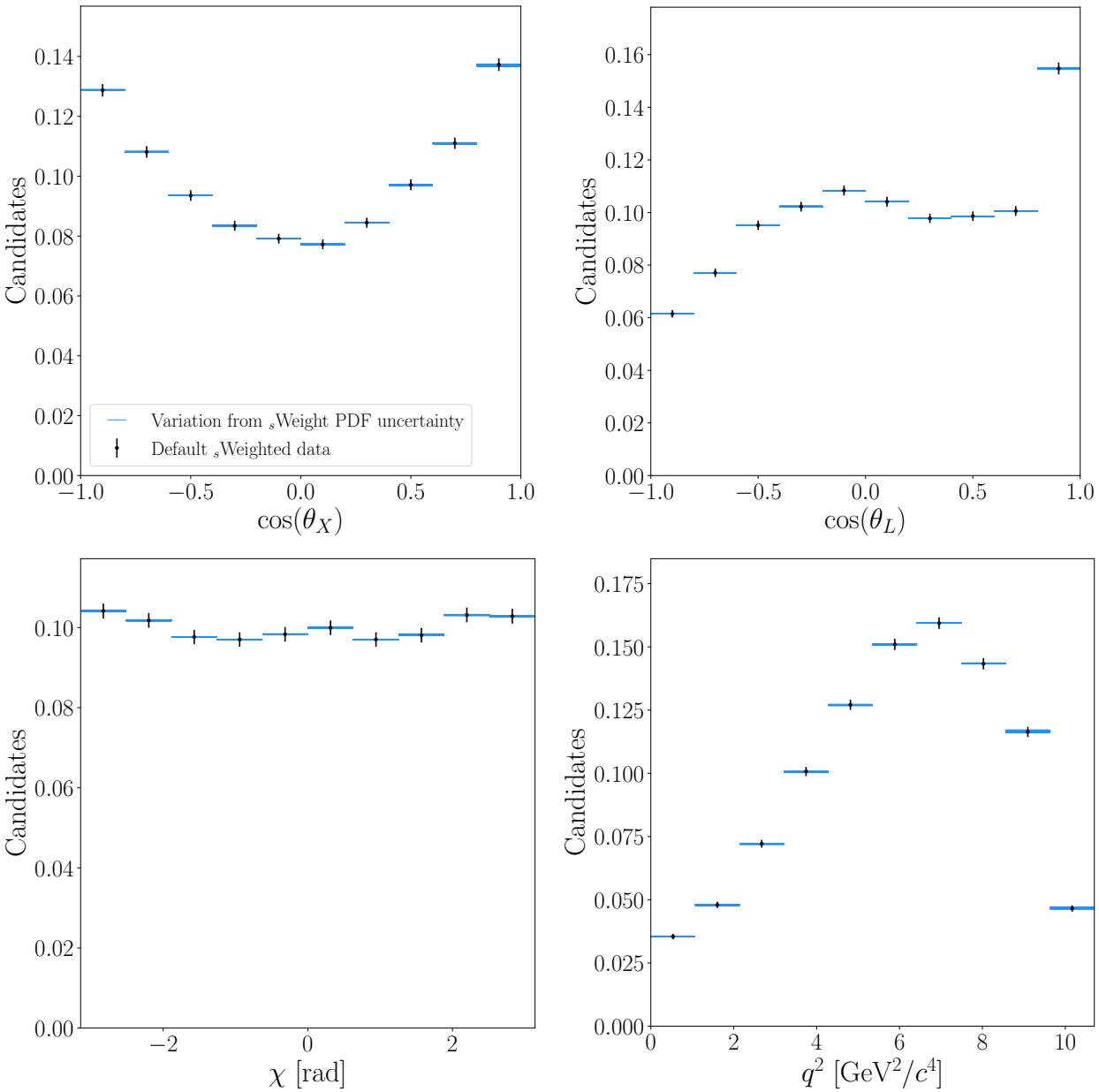


Figure 5.12: Data distributions used in the angular fit with 2D fit $_s$Weights applied (points). The bin content variations caused by variation in the $_s$Weights are shown by the blue shaded rectangles.

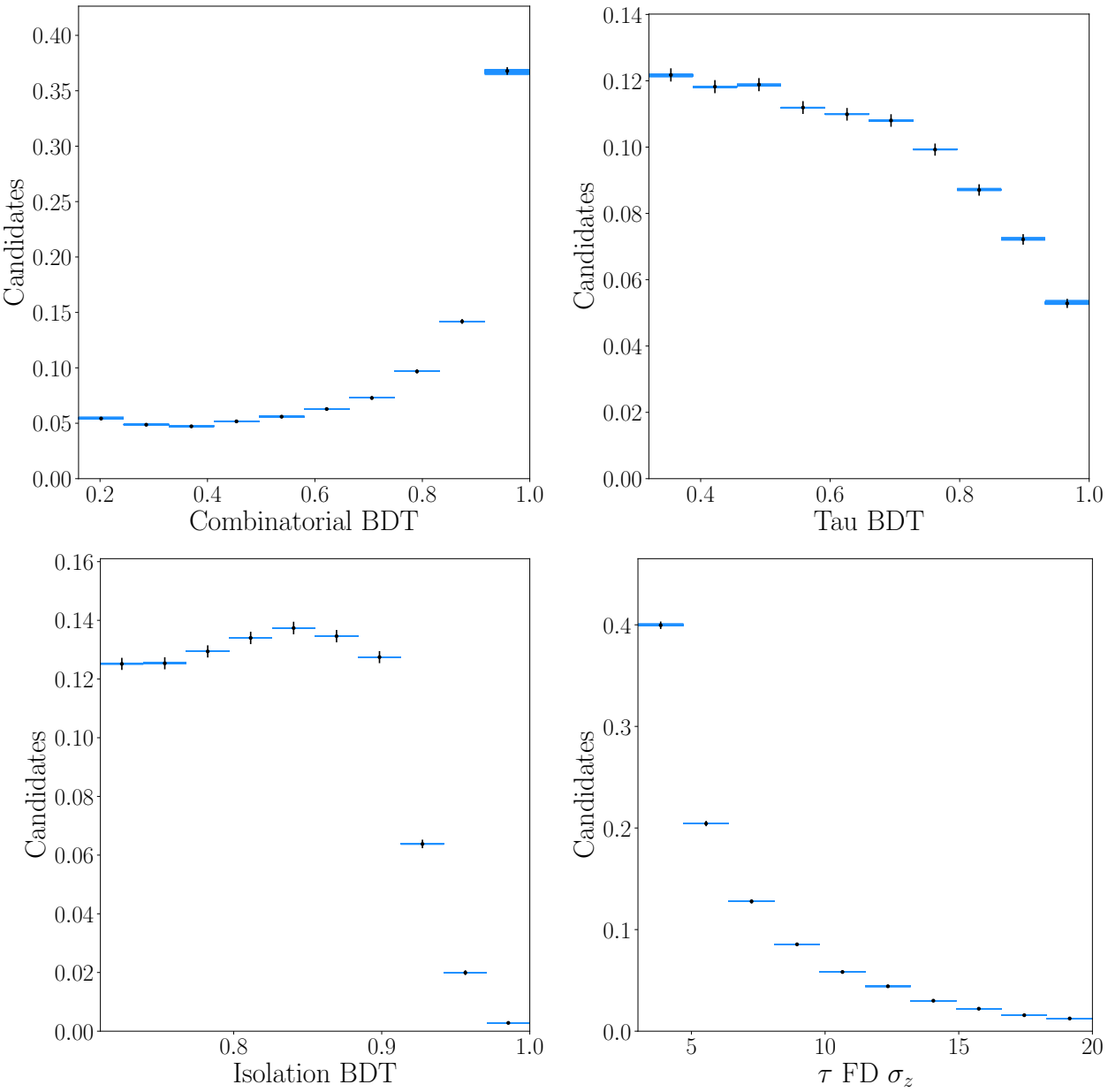


Figure 5.13: Data distributions used in the angular fit with 2D fit $_s$Weights applied (points). The bin content variations caused by variation in the $_s$Weights are shown by the blue shaded rectangles (continued).

5.8 NORMALISATION SELECTION

The selection of the normalisation mode is designed to be as similar to the signal selection as possible; the signal $m(B^0)$ requirement is replaced with the normalisation $m(B^0)$ requirement: $m(B^0) \in [5150, 5400]$ MeV/c^2, while the flight distance and $m(\pi^+\pi^-\pi^+)$ requirements are removed. The requirements on $m(D^0)$ and $m(D^* - D^0)$ are altered to $m(D^0) = m(D^0)_{\text{PDG}} \pm 25\,\text{GeV}/c^2$ and $m(D^* - D^0) \in [142.0, 148.0]\,\text{GeV}/c^2$ respectively. All three optimised BDT requirements are still applied to the normalisation sample.

5.9 EFFICIENCY

Table 5.5 shows the efficiency of the two signal modes when the optimised signal selection is applied to them.

Table 5.5: The efficiencies for the signal modes with the signal selection applied. The relative efficiencies show the efficiency of each requirement individually.

Variable	Condition	Absolute Efficiency		Relative Efficiency	
		$\tau^+ \to 3\pi^\pm$	$\tau^+ \to 3\pi^\pm\pi^0$	$\tau^+ \to 3\pi^\pm$	$\tau^+ \to 3\pi^\pm\pi^0$
Online Selection					
Geometrical Acceptance	Pass	0.163	0.156	-	-
Preselection	Pass	9.036×10^{-4}	6.850×10^{-4}	0.006	0.004
L0 Trigger	Pass	6.644×10^{-4}	6.377×10^{-4}	0.928	0.931
`HLT1` Trigger	Pass	8.064×10^{-4}	6.079×10^{-4}	0.961	0.953
`HLT2` Trigger	Pass	5.904×10^{-4}	4.156×10^{-4}	0.732	0.684
Offline Selection					
Truth Matching	Pass	4.32×10^{-4}	2.94×10^{-4}	0.732	0.707
$m(D^* - D^0)$	$\in [142.0, 148.0]\,\text{GeV}/c^2$	4.32×10^{-4}	2.94×10^{-4}	0.999	0.999
$m(D^0)$	$m(D^0)_{\text{PDG}} \pm 25\,\text{GeV}/c^2$	4.32×10^{-4}	2.94×10^{-4}	1.000	1.000
τ PV	$= D^0$ PV	4.11×10^{-4}	2.80×10^{-4}	0.953	0.954
$m(B^0)$	$< 4700\,\text{GeV}/c^2$	4.06×10^{-4}	2.80×10^{-4}	0.988	0.999
$m(\tau^+)$	$< 1600\,\text{GeV}/c^2$	4.04×10^{-4}	2.80×10^{-4}	0.993	1.000
τ FD z σ	$\in [3.0, 20.0]$	2.12×10^{-4}	1.53×10^{-4}	0.525	0.548
τ-PV separation significance	> 4.0	2.12×10^{-4}	1.53×10^{-4}	1.000	1.000
$\pi^\pm$ from τ decay	$p \in [3, 100]\,\text{GeV}/c$	1.92×10^{-4}	1.40×10^{-4}	0.905	0.916
$\pi^\pm$ from τ decay	$p_T \in [0.3, 10]\,\text{GeV}/c$	1.85×10^{-4}	1.35×10^{-4}	0.963	0.960
π_0^+ from τ decay	PIDK < 4.0	1.82×10^{-4}	1.32×10^{-4}	0.983	0.979
π_1^- from τ decay	PIDK < 1.0	1.74×10^{-4}	1.25×10^{-4}	0.957	0.948
π_2^+ from τ decay	PIDK < 4.0	1.71×10^{-4}	1.22×10^{-4}	0.983	0.979
Combinatorial BDT	> 0.16	1.44×10^{-4}	9.01×10^{-5}	0.840	0.736
Tau BDT	> 0.33	1.19×10^{-4}	6.05×10^{-5}	0.827	0.671
Isolation BDT	> 0.71	7.68×10^{-5}	2.53×10^{-5}	0.647	0.419

6 NORMALISATION MEASUREMENT

The normalisation fit is performed in three stages. The aim of the first two stages, described below, is to measure the fully reconstructed $B^0 \to D^{*-}(D_s^+ \to \pi^+\pi^-\pi^+)$ yield so that it can be fixed in the final stage which measures the normalisation yield. This fully reconstructed decay peaks at the B^0 meson mass, and thus contributes as a background under the $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ normalisation mode peak. The contribution from this mode is small, however, due to the $\sim 1\%$ $D_s^+ \to \pi^+\pi^-\pi^+$ branching fraction.

6.1 $m(\pi^+\pi^-\pi^+)$ D_s PEAK FIT

The first stage of the normalisation measurement is an unbinned fit to the $m(\pi^+\pi^-\pi^+)$ distribution around the D_s^+ mass, where a temporary requirement of $m(\pi^+\pi^-\pi^+) \in m(D_s^+)_{PDG} \pm 150$ MeV/c^2 is applied. The $m(\pi^+\pi^-\pi^+) = m(D_s^+)$ peak is fitted using a Crystal Ball function and the combinatorial background is fitted with an exponential function. The results of this fit can be seen in Figure 6.1 (top).

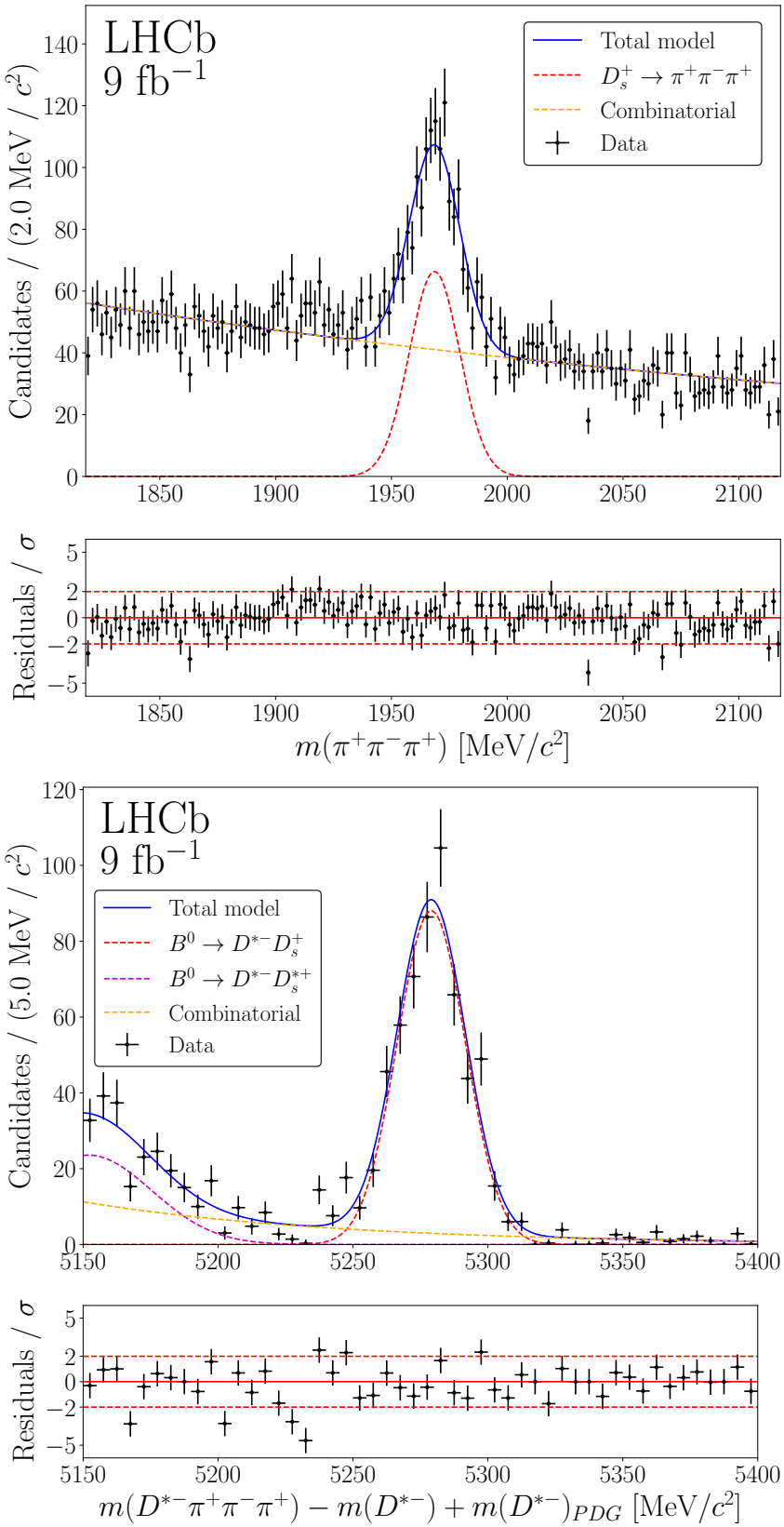


Figure 6.1: Top: Fit to $m(\pi^+\pi^-\pi^+)$, to determine the D_s^+ contribution. Bottom: Fit to $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ with $_s$Weights applied, to determine the $B^0 \to D^{*-}D_s^+$ contribution.

6.2 $m(D^{*-}\pi^+\pi^-\pi^+)$ FIT WITH $_s$WEIGHTS

In the second stage of the normalisation fit, the $m(D^{*-}\pi^+\pi^-\pi^+)$ distribution is altered to $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ to improve the invariant-mass resolution and an unbinned fit is performed. A temporary requirement of $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG} \in (5150,\ 5400)$ MeV/c^2 is applied, and $_s$Weights are assigned to the data based on the $m(\pi^+\pi^-\pi^+)$ fit in order to isolate the contribution from genuine D_s^+ decays [80]. The $_s$Weighted distribution is then fitted using Gaussian distributions for the $B^0 \to D^{*-}D_s^+$ peak and the partially reconstructed $B^0 \to D^{*-}D_s^{*+}$ background, and an exponential function for the combinatorial background. The result of this fit can be seen in Figure 6.1 (bottom plot).

6.3 FULL $m(D^{*-}\pi^+\pi^-\pi^+)$ FITS TO SIMULATION

Before the normalisation peak is measured in data, an unbinned fit to the normalisation simulation is performed to measure the tail parameters of the normalisation peak. A fit is performed to the $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ distribution where a Double Crystal Ball PDF is used to the normalisation peak and an exponential PDF is used for the Combinatorial background. The result of this fit can be seen in Figure 6.2 (top). The values of the tail shape parameters are used to fix the corresponding parameters in the data fit.

An unbinned fit to the $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ distribution for a $B^0 \to D^{*-}K^+\pi^-\pi^+$ simulation sample is performed. This decay can enter the signal region through the misidentification of the K^+ as a π^+. The result of this fit can be seen in Figure 6.2 (bottom). The values of all shape parameters are used to fix the corresponding parameters in the data fit.

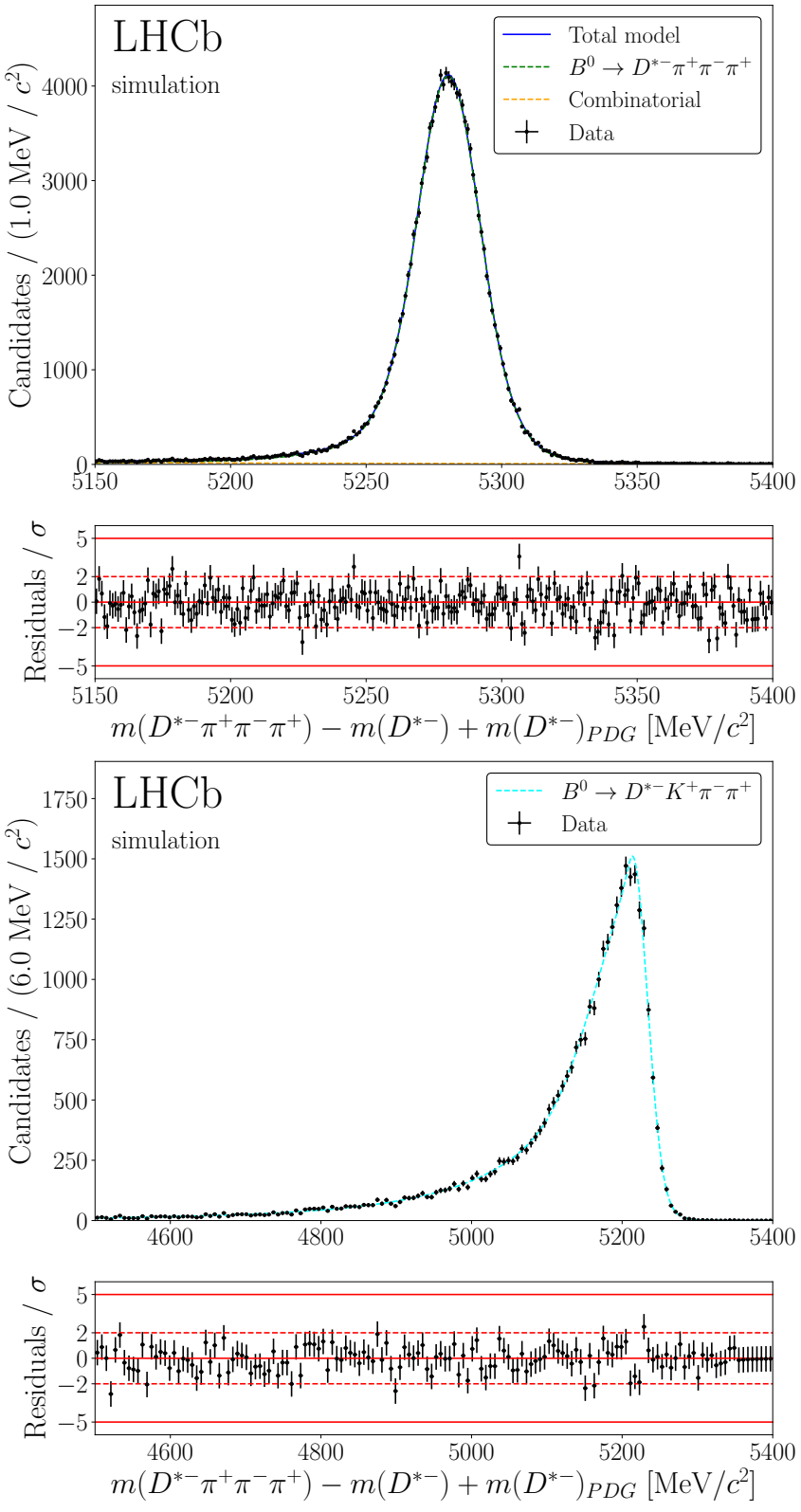


Figure 6.2: Top: Fit to $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ for the normalisation simulation. Bottom: Fit to $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ for the $B^0 \to D^{*-}K^+\pi^-\pi^+$ simulation.

6.4 FULL $m(D^{*-}\pi^+\pi^-\pi^+)$ FIT

The final stage of the normalisation fit measures the normalisation yield, which is used to calculate the signal yield in the MVA optimisation as described in Section 5.5.4. An unbinned fit is performed to the $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ distribution without ${}_s$Weights applied. A temporary requirement of $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG} \in (5150,\ 5400)$ MeV/c^2 is applied. The $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ normalisation peak is fitted with a double Crystal Ball function, where the tail parameters are fixed to the values found in the fit to simulation above, and the yield and shape of the $B^0 \to D^{*-}D_s^+$ peak is fixed using the results of the previous fit described above in Section 6.2. A misidentified $B^0 \to D^{*-}K^-\pi^+\pi^+$ contribution is modelled using a double Crystal Ball where all shape parameters are fixed from the fit to the simulation sample described above, but the yield is Gaussian constrained to be a fraction of the normalisation yield. This fraction is estimated using the following three factors:

- The ratio of branching fractions for the normalisation and misidentified decays: $\mathcal{B}(B^0 \to D^{*-}K^+\pi^-\pi^+)/\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+) = 0.065 \pm 0.006$.
- The efficiency of the K^+ passing the particle identification requirements for the normalisation selection: PIDK < 4 efficiency $= 0.2 \pm 0.1$. This efficiency is not measured using the simulation so a large uncertainty is assigned to its value.
- The efficiency of the simulation when applying the window on $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$; the requirement on the $m(B^0)$ distribution has an efficiency $= 0.573 \pm 0.003$.

This results in a Gaussian constrained fraction of 0.007 ± 0.004. Finally, an exponential PDF is used to model the combinatorial background. The results of the fit can be seen in Figure 6.3; the fit with the optimised BDT requirements applied is also shown. The measured yield of the normalisation peak is $41,830 \pm 270$ for the optimised BDT requirements.

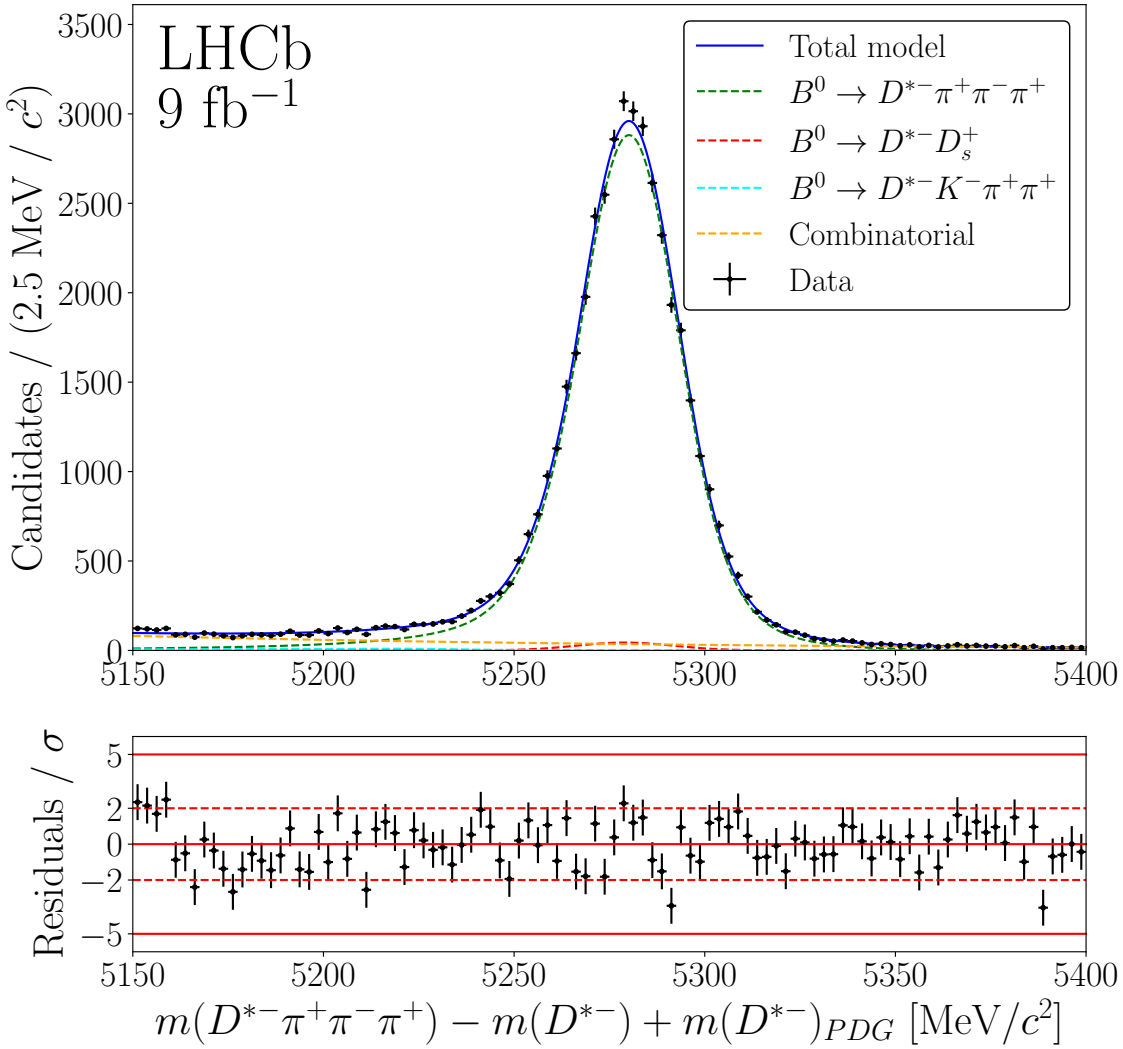


Figure 6.3: Final stage of the normalisation fit: fitting the normalisation peak in $m(D^{*-}\pi^+\pi^-\pi^+) - m(D^{*-}) + m(D^{*-})_{PDG}$ with the optimised BDT requirements applied.

7 DATA-DRIVEN BACKGROUND CONTROL STUDIES

Before the signal fit is performed, the simulation for the largest backgrounds is corrected using data-driven methods. The data samples used for this purpose are called control samples and several control samples are created.

A generic control selection is created by using the signal selection and removing the requirements on $m(B^0)$, $m(\pi^+\pi^-\pi^+)$, the τ vs D_s BDT, and the Isolation BDT. The optimised Combinatorial BDT requirement is still applied in this generic control selection. Each control study described below adds additional requirements on top of this base selection and will be described in the relevant sections.

First, the resonant structure of $D_s^+ \to \pi^+\pi^-\pi^+(X)$ decays is corrected, where (X) are additional particles that may be produced, but are not reconstructed. This study provides corrective weights for the $B \to D^* D_s(X)$ simulation and is described in Section 7.1. Next the relative amounts of $B \to D^* D_s(X)$ decays are measured and provide constraints for use in the signal fit. This is detailed in Section 7.2. Section 7.3 describes the isolation of $B \to D^* D^+(X)$ and $B \to D^* D^0(X)$ control samples, which are used to produce corrective weights that increase data/simulation agreement for the relevant simulation samples. These two control studies are described as one since they use the same methodology. Section 7.4 details the use of $B^0 \to D^* 3\pi$ data and simulation to correct the L0 trigger composition of all simulation samples. Lastly, a comparison of $b\bar{b} \to D^* 3\pi X$ inclusive simulation and data is performed in order to derive data-driven corrections for the 3π kinematics and z flight distance significance in Section 7.5.

7.1 D_s^+ DECAY STUDY

For this control study, the generic control selection is used, with additional requirements specified in Table 7.1 applied. These requirements apply an inverted τ vs D_s BDT condition, the optimised Combinatorial and Isolation BDT requirements, an altered tau flight significance requirement, and windows on the four fit variables: $m(\pi^+\pi^-\pi^+)$, $m(\pi^+\pi^+)$, $\min[m(\pi^+\pi^-)]$, and $\max[m(\pi^+\pi^-)]$. The altered tau flight significance requirement is used due to the higher D_s lifetime compared to the τ, so a higher upper limit captures a larger proportion of D_s decays.

This control study measures the resonant structure of $D_s^+ \to \pi^+\pi^-\pi^+(X)$ decays. The measured fractions of these modes are compared to the fractions that exist in the simulated samples, and a set of corrective weights is derived which ensures that the simulation aligns well with the data. The $B \to D^* D_s(X)$ simulated samples eventually used in the signal fit in Chapter 8 have these corrective weights applied.

The decay fractions are measured using a simultaneous binned maximum likelihood template fit to the fit variables mentioned above. The variables $m(\pi^+\pi^+)$, $\min[m(\pi^+\pi^-)]$, and $\max[m(\pi^+\pi^-)]$ are the reconstructed masses of each of the $\pi^\pm\pi^\pm$ pairs in the $\pi^+\pi^-\pi^+$ system. These variables capture the resonant structure in the $D_s \to \pi^+\pi^-\pi^+(X)$ decays. The $m(\pi^+\pi^-\pi^+)$ variable enables the fit to distinguish the fully reconstructed $D_s^+ \to \pi^+\pi^-\pi^+$ and $D^+ \to \pi^+\pi^-\pi^+$ decays from partially reconstructed $B \to D^*\pi^+\pi^-\pi^+(X)$ decays, which can involve a D_s, D^+, D^0, τ, or other particles as intermediate states. The templates for the $D_s^+ \to \pi^+\pi^-\pi^+(X)$ decays come from the $B \to D^* D_s(X)$ simulation, and a "non-D_s" template is taken

Table 7.1: The additional requirements applied to the $D_s^+ \to \pi^+\pi^-\pi^+(X)$ control data sample and simulation. These requirements are applied on top of the generic control selection as described above.

Variable	Requirement
tau flight significance	$\in [3.0,\ 25.0]$
$m(D^{*-}\pi^+\pi^-\pi^+)$	$< 5400\,\text{MeV}/c^2$
$m(\pi^+\pi^-\pi^+)$	$\in [500,\ 2100]\,\text{MeV}/c^2$
$m(\pi^+\pi^+)$	$\in [280,\ 2000]\,\text{MeV}/c^2$
$\min[m(\pi^+\pi^-)]$	$\in [280,\ 1500]\,\text{MeV}/c^2$
$\max[m(\pi^+\pi^-)]$	$\in [280,\ 2000]\,\text{MeV}/c^2$

from the inclusive $b\bar{b} \to D^{*-}\pi^+\pi^-\pi^+(X)$ simulation which is used to model all sources of $B \to D^*\pi^+\pi^-\pi^+(X)$ that do not involve the production of a D_s meson.

The fit model is split into two halves, the D_s PDF which encompasses all of the $D_s^+ \to \pi^+\pi^-\pi^+(X)$ decays and the non-D_s PDF which encompasses all of the $\pi^+\pi^-\pi^+(X)$ decays that do not contain a D_s in their decay chains. The relative amount of these PDFs is controlled by the D_s fraction fit parameter: f_{D_s}.

The non-D_s PDF is not subdivided further and the internal fractions of this PDF are fixed to the native values in the simulation. The D_s PDF is subdivided into many subcategories which are normalised by their subcategory fractions. A subcategory X has an internal D_s PDF fraction $f_{D_s^+ \to X}$ and therefore a total PDF normalisation of $f_{D_s} \cdot f_{D_s^+ \to X}$. These subcategories are defined in Table B.1. The subcategory, $D_s^+ \to X\pi^+\pi^-\pi^+$, does not have a subcategory fraction, but is instead normalised by the sum of the other subcategory fractions subtracted from unity to ensure the D_s PDF fractions sum to unity:

$$f_{D_s^+ \to X\pi^+\pi^-\pi^+} = 1 - \sum_X f_{D_s^+ \to X}. \tag{7.1}$$

The fit model is first validated through the creation of many pseudo-experiments where simulation samples are used as the data sample to be fit. A pseudo-data sample is created from the simulation used to make the templates. The expected statistics of each sample are calculated from the statistics of the control data sample, and the default template fraction in the simulation:

$$N(X) = N(\text{data}) \cdot f_X. \tag{7.2}$$

where f_X is the fraction of subcategory X in the simulation. Then a random number is chosen using the Poisson distribution for $N(X)$, $P(N)$, and that number of events is randomly selected from the relevant template. These events are added to the pseudo-data sample and removed from the fit template. This results in a pseudo-data sample that is statistically independent of the fit templates, has roughly the same statistics as the real data sample, and has decay fractions close to the expected values.

Then the fit is performed on the pseudo-data sample, and the fit results can be compared to the true values. The pull values are defined as

$$\frac{p_{\text{true}} - p_{\text{fit}}}{\sigma_{p_{\text{fit}}}},$$

for a fit variable p, and are calculated for all fit parameters. This process was repeated 1000 times, creating a new pseudo-data sample each time, enabling pull distributions for the parameters to be produced. Since the fit is performed to four variables simultaneously, the statistical error for the fit parameters will be underestimated. The pull distributions for the pseudo-experiments can be used to correct biases in the central value of the fit parameters, but also the underestimation of the statistical uncertainty. The following transformations are applied:

$$p^c = p - \mu^p_{\text{pull}} \times \sigma$$
$$\sigma^c = \sigma^p_{\text{pull}} \times \sigma$$

where $p \pm \sigma$ is the default fit result for a parameter p, μ^p_{pull} is the mean of its pull distribution, and σ^p_{pull} the width of its pull distribution.

The fit projections for the fit to the control data sample are shown in Figures 7.1 and 7.2. The decay fractions present in the simulation compared to the fit fractions are shown in Table B.1 and the fit parameters are shown in Figure 7.3. The derived correction factors for the D_s simulation are shown in Figure 7.4.

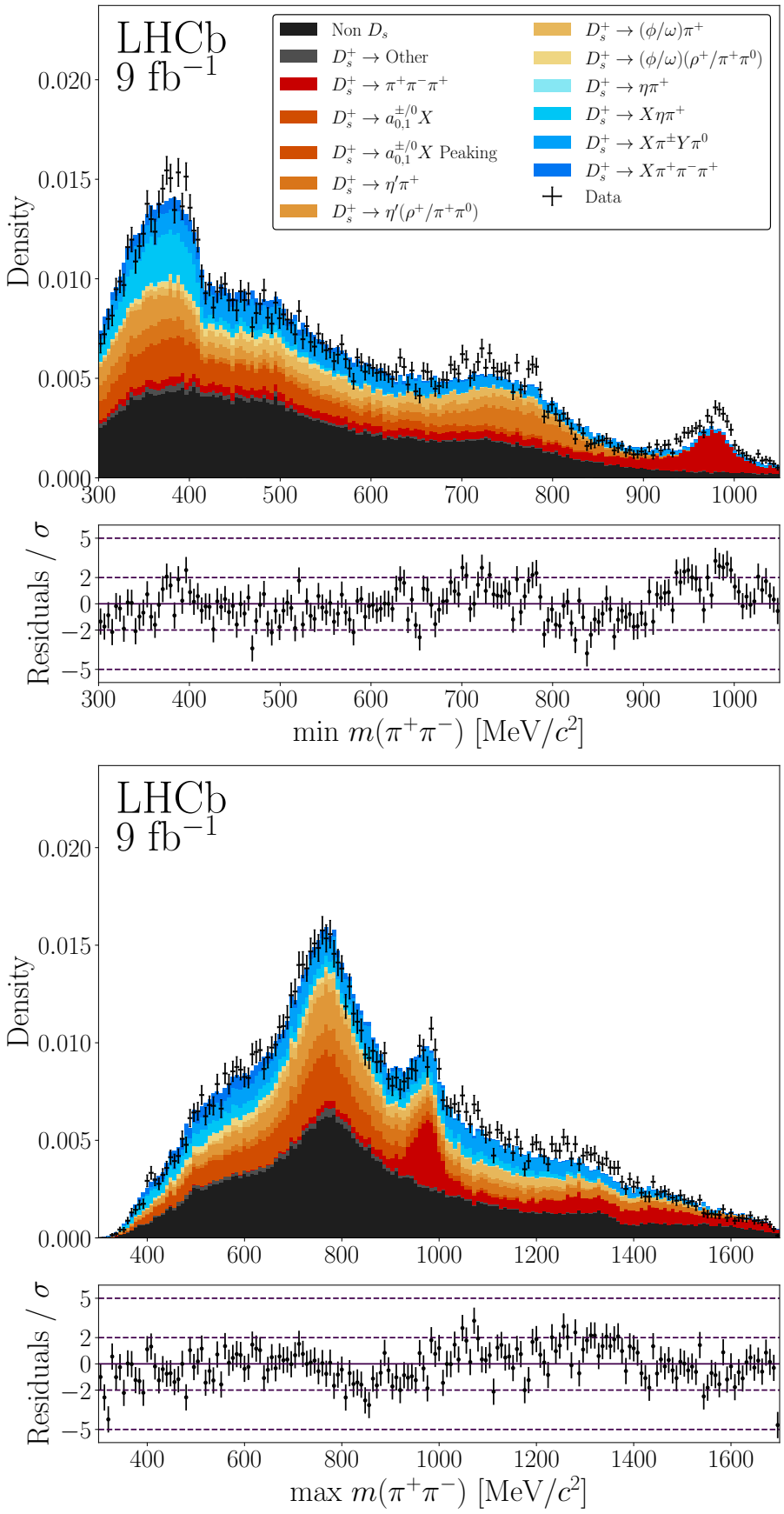


Figure 7.1: The fit projections for two of the four fit variables: min$[m(\pi^+\pi^-)]$ and max$[m(\pi^+\pi^-)]$ for the control data fit.

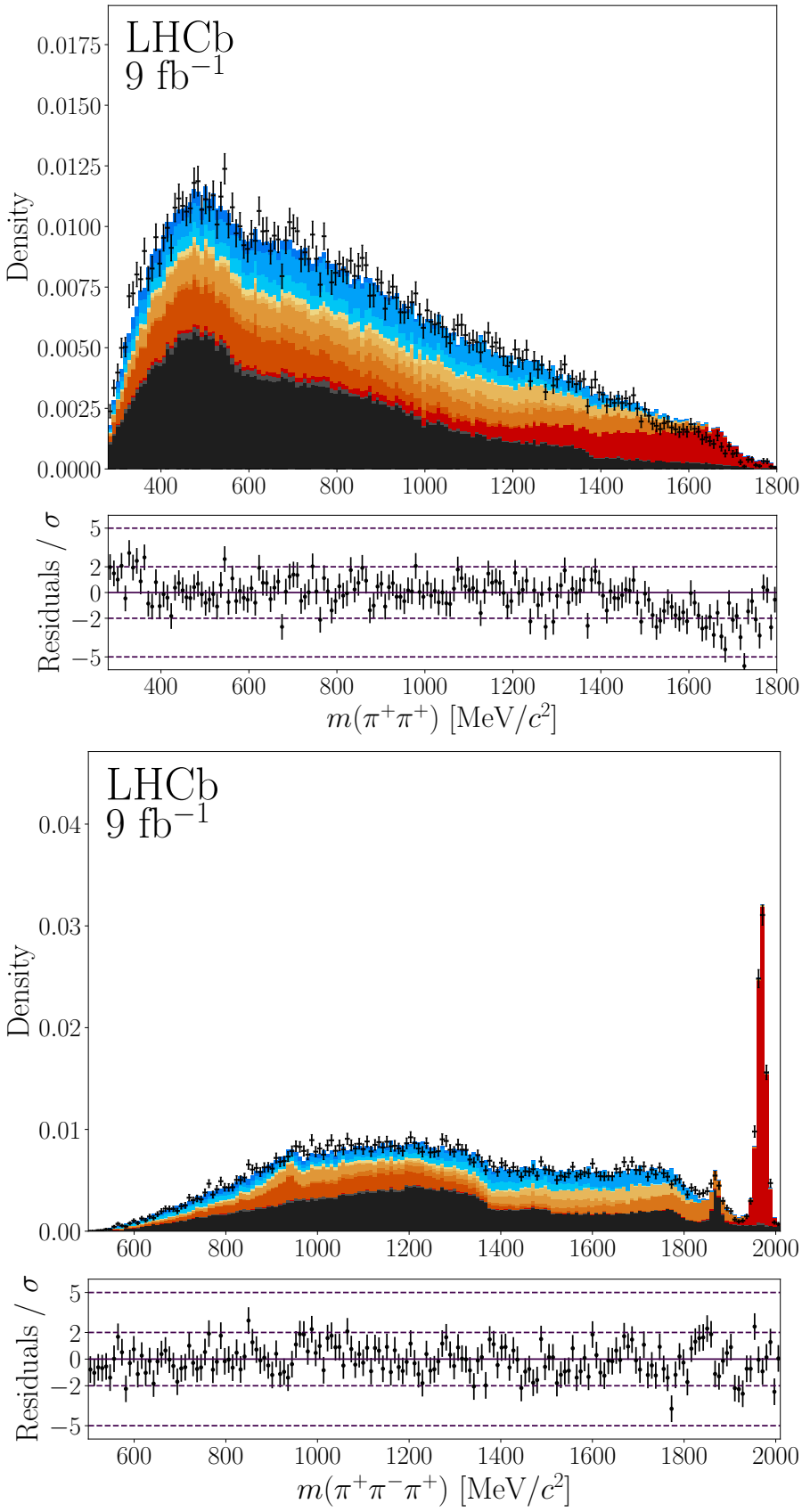


Figure 7.2: The fit projections for two of the four fit variables: $m(\pi^+\pi^+)$ and $m(\pi^+\pi^-\pi^+)$ for the control data fit.

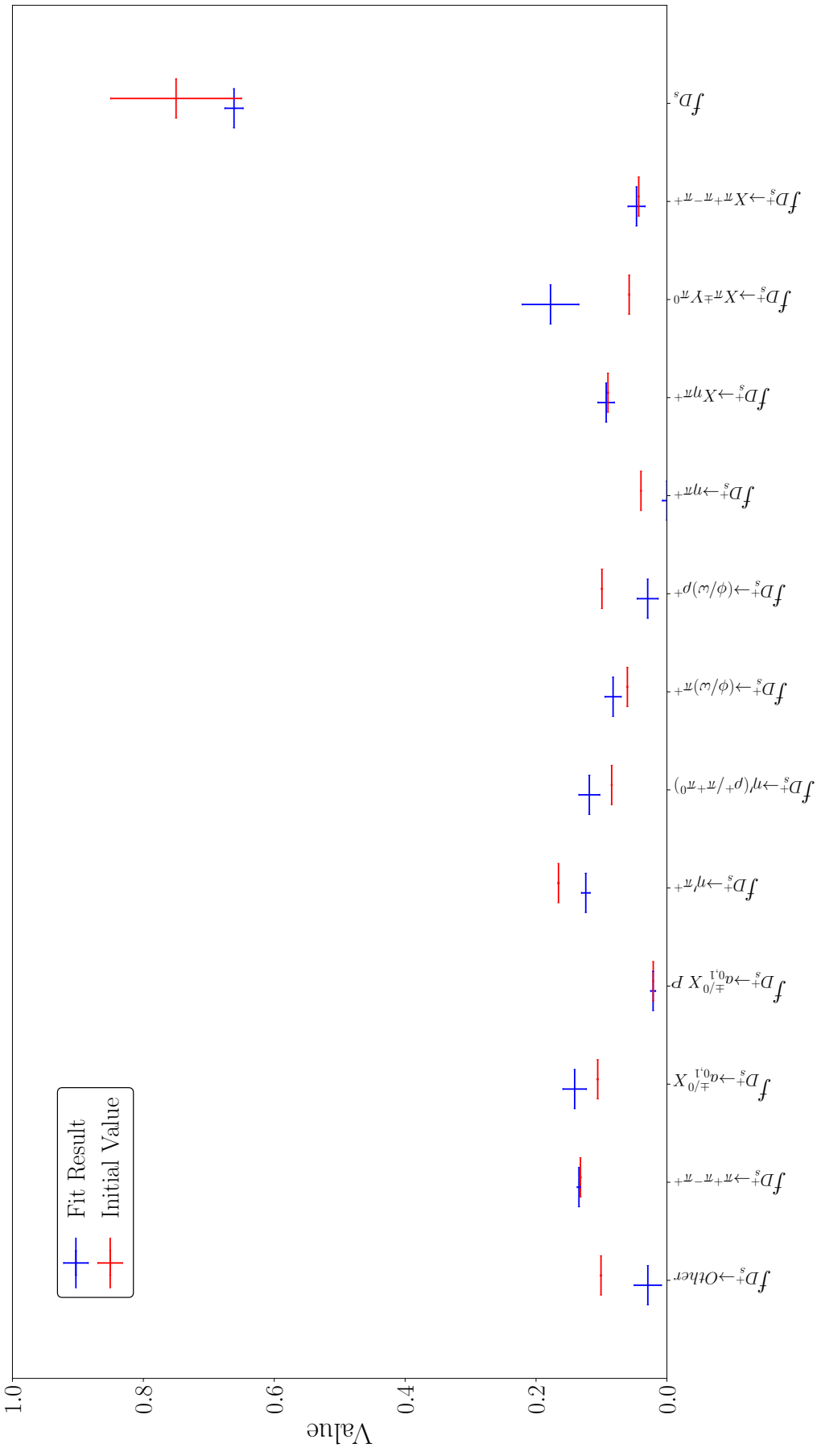


Figure 7.3: The fit parameters measured in the control data fit (blue). These are compared to their native values in the simulation to derive a correction factor (red).

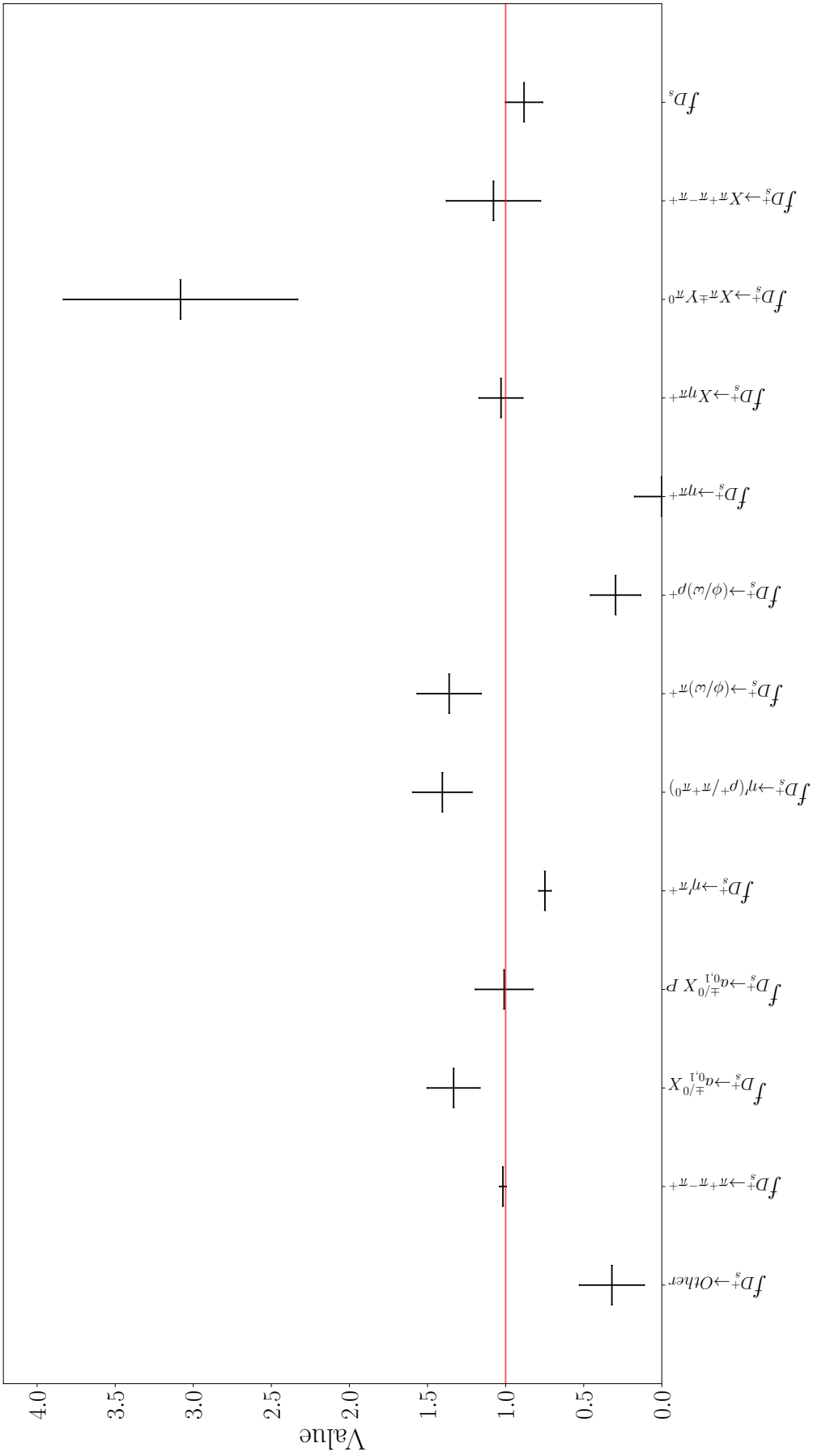


Figure 7.4: The correction factors measured in the control data fit. These factors are used as weights to correct the D_s decay fractions in the D_s simulation.

7.2 $B \to D^{*-} D_s^+(X)$ CONTROL MODE

Given that the $B \to D^* D_s X$ double-charm background is the largest single contribution in data, a detailed study of its composition is strongly motivated. In addition to correcting the simulated $D_s \to 3\pi X$ fractions as detailed in Section 7.1, the fractions of each B decay contributing to the total $B \to D^* D_s X$ component must also be aligned between data and simulation. To this end, control samples of $B \to D^*(D_s^+ \to 3\pi)X$ decays are selected in data and simulation, enabling a measurement of the B decay fractions in data via a fit to the $m(D^* D_s)$ distribution.

After these fractions are determined, a comparison between data and simulation is performed across several high-level variables including the decay angles, 3π flight, q^2, and the selection BDTs. This comparison enables a set of correction weights to be determined for the full $B \to D^* D_s X$ simulation sample, in order to align the data and simulation as much as possible. This reweighting procedure is designed to correct for residual differences in decay dynamics (e.g. imperfect knowledge of helicity amplitudes) between data and simulation, as well for differences in detector resolution and acceptance.[1]

7.2.1 SELECTION OF $B \to D^*(D_s^+ \to 3\pi)X$

To study the fractional contribution from different B-hadron decays in the $B \to D^* D_s X$ background, it is necessary to isolate a fully reconstructed D_s meson. The $D_s^+ \to 3\pi$ decay is chosen given it shares the same final state as the $D^* \tau \nu$ signal mode, and thus can be selected under the same conditions as signal albeit with looser BDT requirements to allow for more complete data/simulation alignment.

To select pure samples of $D_s^+ \to 3\pi$ decays in data and simulation, fits to the $m(3\pi)$ distribution of selected candidates are performed in order to derive ${}_s$Weights. Prior to fitting, $B \to D^* D_s X$ simulation is split into component B decays: $B^0 \to D^{*-} D_s^+$, $B^0 \to D^{*-} D_s^{*+}$, $B^0 \to D^{*-} D_{s0}(2317)^{*+}$, $B^0 \to D^{*-} D_{s1}(2460)^+$, $B \to D^{**} D_s^+ X$, and $B_s^0 \to D^{*-} D_s^+ X$. This enables the simulation samples for each mode to be used in a template fit to data, in order to measure the B decay fractions. The

[1] The initial version of this control study was performed by Anthony Correia.

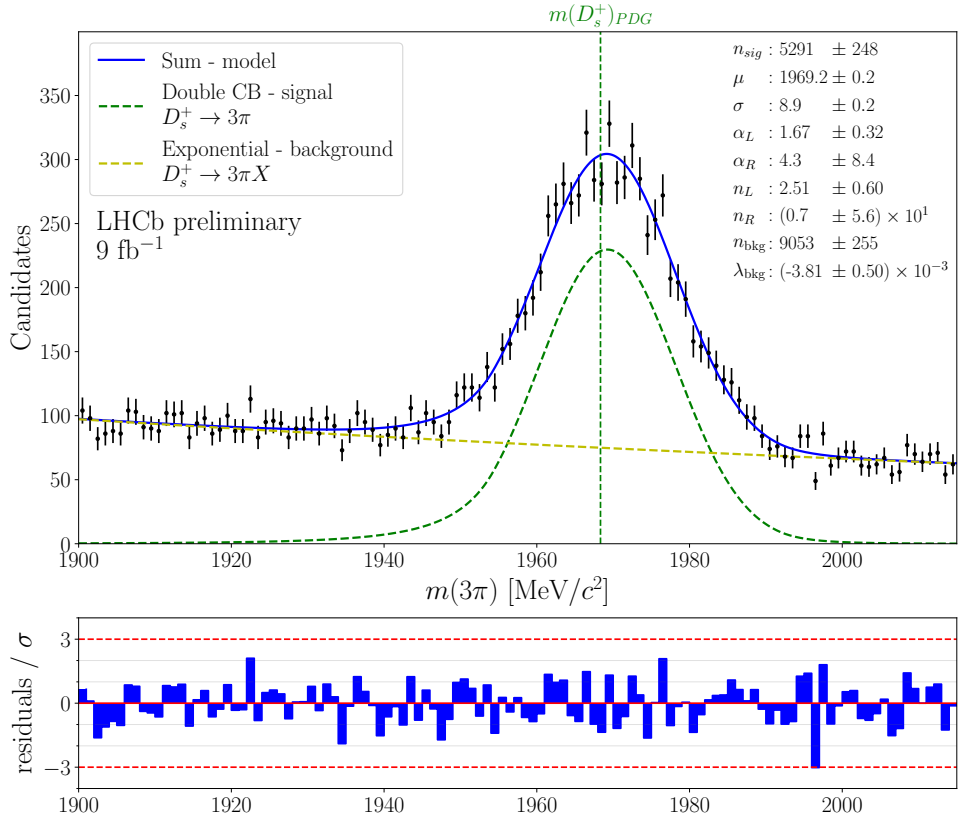


Figure 7.5: Invariant-mass fit to the $m(3\pi)$ distribution in data, the signal peak is fitted with a Double Crystal-Ball function.

$m(3\pi)$ fit to data and simulation is shown in Figure 7.5. A total $D_s^+ \to 3\pi$ yield of $5,291 \pm 248$ is measured in data, providing sufficient statistics for a measurement of the B decay fractions and a subsequent data-driven correction of the $B \to D^* D_s X$ simulation.

7.2.2 FIT TO $m(D^* D_s)$

The $B \to D^* D_s X$ simulation is generated with fixed internal fractions for the B-decay modes listed above. To align the simulation with data, these fractions can be measured in data using a fit to the $m(D^* D_s)$ distribution, and the results are used to reweight the simulation fractions. This approach is possible due to the different invariant-mass distributions of modes such as fully reconstructed $B^0 \to D^{*-} D_s^+$ and partially reconstructed $B^0 \to D^{*-} D_s^{*+}$.

A binned maximum-likelihood fit is performed to the $m(D^* D_s)$ distribution in data, where a sample of $B \to D^*(D_s \to 3\pi)X$ decays is isolated using the $m(3\pi)$

${}_s$Weights described above. For each decay mode, a binned template is derived from simulation, created with ${}_s$Weights from the $m(3\pi)$ fit applied. Exceptions to this are the $B^0 \to D^{*-} D_s^+$ and $B^0 \to D^{*-} D_s^{*+}$ modes, where the following PDF parameterisations are used:

- $B^0 \to D^{*-} D_s^+$: A double Crystal Ball PDF, with tail parameters determined from a fit to a dedicated sample of $B \to D^*(D_s \to KK\pi)$ simulated decays as used in the LHCb measurement of the longitudinal fraction f_L in $B^0 \to D^{*-} D_s^{*+}$ decays [81].

- $B^0 \to D^{*-} D_s^{*+}$: RooHORNSdini and RooHILLSdini PDFs [82] are taken from the dedicated LHCb analysis of $B^0 \to D^{*-} D_s^{*+}$ decays [81], where RapidSim samples are generated separately for $D_s^* \to D_s\gamma$ and $D_s^* \to D_s\pi^0$ in both the transverse and longitudinal polarisation states. The $D_s^* \to D_s\gamma$ and $D_s^* \to D_s\pi^0$ sub-components are combined using PDG branching fractions and simulation efficiencies. Separate PDFs for longitudinal and transverse polarised decays are combined using the LHCb measurement of f_L in this decay, namely $f_L = 0.578 \pm 0.015$.

In the $m(D^*D_s)$ fit to data, the yield of the $B^0 \to D^* D_s^*$ component varies freely. This component is chosen since it is the largest contribution in the fit, but the choice is arbitrary. All other decay modes are measured relative to this contribution with freely varying ratios. The relative peak height parameter ξ of the $B^0 \to D^* D_s^*$ PDFs also varies freely, to enable the modelling of small mass-dependent acceptance effects not present in the RapidSim simulation. The fit to data is shown in Figure 7.6, and the fit results are summarised in Table 7.2.

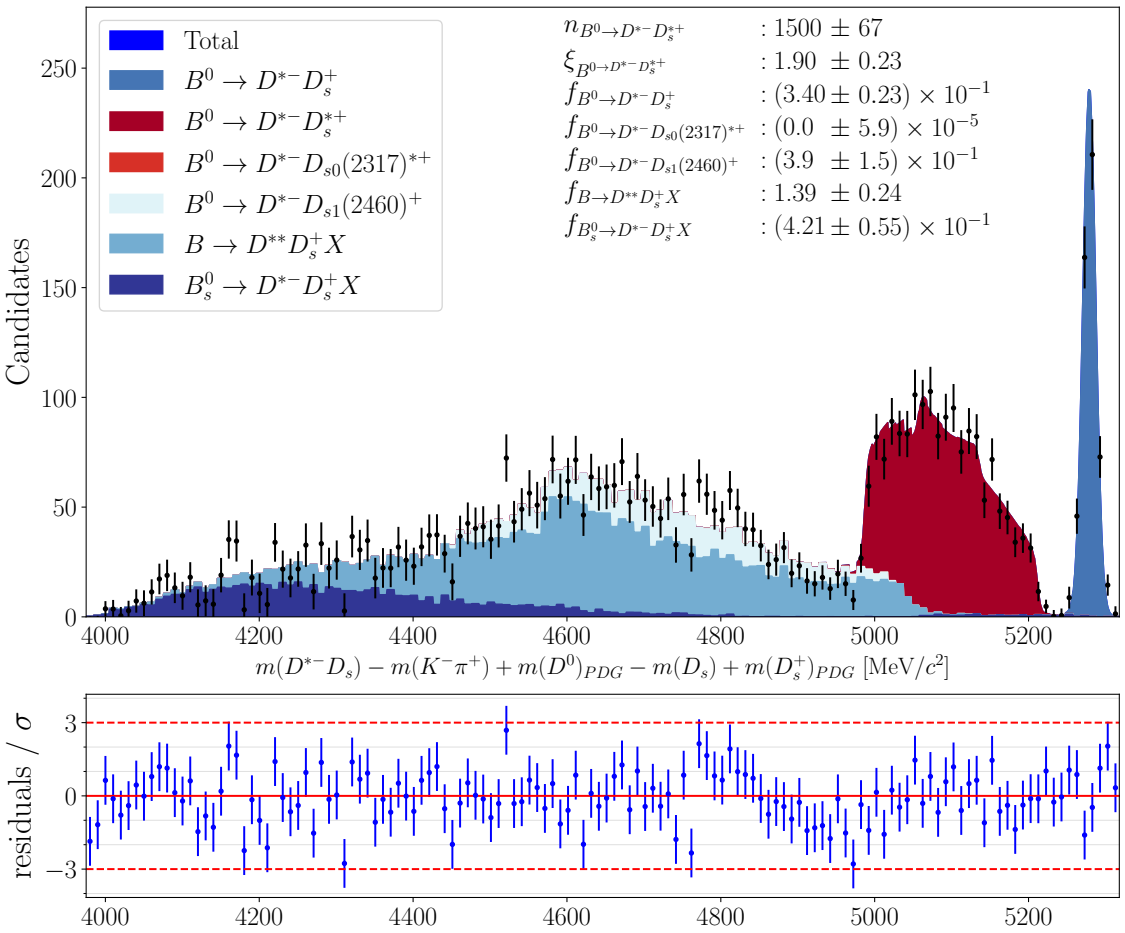


Figure 7.6: Invariant-mass fit to the $m(D^* D_s)$ distribution of the $B \to D^*(D_s \to 3\pi)$ control sample in data. From this fit, the fractions of each contributing B decay mode are measured relative to the $B^0 \to D^* D_s^*$ contribution, allowing the generated simulation fractions to be reweighted to align with data.

Table 7.2: Results of the $m(D^* D_s)$ fit to data. The fit variables contain one absolute yield, five relative fractions, and a shape parameter.

Variable	Fitted Value
$n_{B^0 \to D^{*-} D_s^{*+}}$	1500 ± 67
$\xi_{B^0 \to D^{*-} D_s^{*+}}$	1.90 ± 0.23
$f_{B^0 \to D^{*-} D_s^+}$	$(3.40 \pm 0.23) \times 10^{-1}$
$f_{B^0 \to D^{*-} D_{s0}(2317)^{*+}}$	$(0.0 \pm 5.9) \times 10^{-5}$
$f_{B^0 \to D^{*-} D_{s1}(2460)^+}$	$(3.9 \pm 1.5) \times 10^{-1}$
$f_{B \to D^{**} D_s^+ X}$	1.39 ± 0.24
$f_{B_s^0 \to D^{*-} D_s^+ X}$	$(4.21 \pm 0.55) \times 10^{-1}$

To assess the stability of the $m(D^*D_s)$ fit, and derive corrections for any bias observed, a series of pseudo-experiment fits are performed. Pseudo-experiments are generated from the final data fit PDF and pull distributions in each of the freely varying fit parameters are created. The pull means and widths are used to correct the default fit results as follows:

$$p^c = p - \mu^p_{\text{pull}} \times \sigma \tag{7.3}$$

$$\sigma^c = \sigma^p_{\text{pull}} \times \sigma \tag{7.4}$$

where $p \pm \sigma$ is the default fit result for a parameter p, μ^p_{pull} is the mean of its pull distribution, and σ^p_{pull} the width of its pull distribution. These corrections are generally small, with the largest μ^p_{pull} being 0.67 and the σ^p_{pull} furthest from 1.0 being 0.86. The fit parameter central values and uncertainties after pull corrections are summarised in Table 7.3.

To calculate the full uncertainties on the $B \to D^*D_sX$ component fractions, which are used to correct the simulation fractions, systematic uncertainties from the $m(D^*D_s)$ fit are also evaluated. Sources of systematic uncertainty include:

- The use of a fixed shape parameters for the $B^0 \to D^{*-}D_s$ fully reconstructed peak PDF;
- The use of a fixed shape parameters for the $B^0 \to D^{*-}D_s^*$ PDFs, and the use of a fixed value for f_L;
- The statistical uncertainty on the templates used to describe the remaining components in the fit.

Table 7.3: Results of the $m(D^*D_s)$ fit to data after pull bias corrections to the central values and uncertainties.

Variable	Fitted Value
$n_{B^0 \to D^{*-}D_s^{*+}}$	1514 ± 57
$f_{B^0 \to D^{*-}D_s^+}$	$(3.39 \pm 0.22) \times 10^{-1}$
$f_{B^0 \to D^{*-}D_{s0}(2317)^{*+}}$	$(0.0 \pm 3.6) \times 10^{-5}$
$f_{B^0 \to D^{*-}D_{s1}(2460)^+}$	$(4.5 \pm 1.5) \times 10^{-1}$
$f_{B \to D^{**}D_s^+X}$	1.32 ± 0.24
$f_{B_s^0 \to D^{*-}D_s^+X}$	$(4.26 \pm 0.55) \times 10^{-1}$

Table 7.4: Total uncertainties for each component yield and fraction in the $m(D^*D_s)$ fit. The total uncertainties are considered when reweighting the $B \to D^*D_sX$ simulation fractions.

Variable	Final Result
$n_{B^0 \to D^{*-} D_s^{*+}}$	1514 ± 58
$f_{B^0 \to D^{*-} D_s^+}$	$(3.39 \pm 0.23) \times 10^{-1}$
$f_{B^0 \to D^{*-} D_{s0}(2317)^{*+}}$	$(0.0 \pm 3.8) \times 10^{-3}$
$f_{B^0 \to D^{*-} D_{s1}(2460)^+}$	$(4.5 \pm 1.6) \times 10^{-1}$
$f_{B \to D^{**} D_s^+ X}$	1.32 ± 0.25
$f_{B_s^0 \to D^{*-} D_s^+ X}$	$(4.26 \pm 0.58) \times 10^{-1}$

The $B^0 \to D^*D_s$ and $B^0 \to D^*D_s^*$ systematic uncertainties are separately evaluated by running many fits to data, each time randomly varying the associated fixed parameters within their uncertainties. In the case of the shape parameters, the uncertainties are taken from the respective fits to simulation, while the uncertainty on f_L is taken from the LHCb measurement. For the template statistical uncertainties, many alternative templates are created for each mode by Poisson varying the bin contents of the default histograms: many fits to data are then performed using the varied histograms. In all cases, distributions of the resulting fit parameters are created, and the widths of the distributions are taken to represent the systematic uncertainties. The total uncertainties for each parameter are then determined by a sum in quadrature over the individual systematic uncertainties and the pull-corrected fit statistical uncertainties. The fit parameters and their total uncertainties are summarised in Table 7.4.

7.2.3 CORRECTION OF $B \to D^*D_sX$ SIMULATION

Once the $B \to D^*D_sX$ simulation is corrected to align with the data using the B decay fractions measured above, it is possible to perform a comparison of data and simulation using the $B \to D^*(D_s \to 3\pi)X$ control samples. This comparison can be used to identify remaining discrepancies present due to imperfect knowledge of the decay dynamics (e.g. no polarisation measurements exist for $B \to D^{**}D_sX$ decays), the internal composition of the $B_s \to D^*D_sX$ sample, and differences in detector resolution and acceptance. A comparison is performed across a range of high-level

variables, namely the decay angles, q^2, and the selection BDTs. The results are shown in Figures 7.7, 7.8, 7.9, and 7.10, where each sub-decay is shown in the same colour as the $m(D^*D_s)$ fit in Figure 7.6 and with a yield corresponding to the yield measured in data.

It is possible to correct for the discrepancies observed, in order to derive a set of weights which can act to align the full $B \to D^*D_sX$ simulated sample with data. The following procedure is used to compare the $B \to D^*(D_s \to 3\pi)X$ data and simulation shown in Figures 7.7 and 7.8, in order to derive correction weights:

- A list of variables to be corrected is defined, containing the decay angles, q^2, and the three selection BDT distributions.
- A binned ratio of data/simulation histograms is created in the first variable x. The simulation histogram is created by combining the $B \to D^*(D_s \to 3\pi)X$ mode subsamples for each B decay, weighted using the $m(D^*D_s)$ fit results to align their content with data.
- The data/simulation ratio histogram is fitted with a cubic spline function $f(x)$, with smoothing applied according to the number of bins used and the uncertainties on each bin considered in the spline evaluation.
- The spline is used to apply per-event corrections to the simulation control sample, where $w_x^i = f(x^i)$ for each event i.
- A binned ratio of data/simulation is created in the second variable y, where the correction weights w_x for the previous variable x are applied to simulation. The above process is repeated sequentially until splines for all variables have been calculated.

The net result of this procedure is a set of correction splines, which when applied to each $B \to D^*D_sX$ simulation event provide a single weight that acts to align all of the considered variables simultaneously. This sequential reweighting approach offers good control over the reweighting performance, and more robust results than those achieved using a binned or BDT-based reweighting procedure.

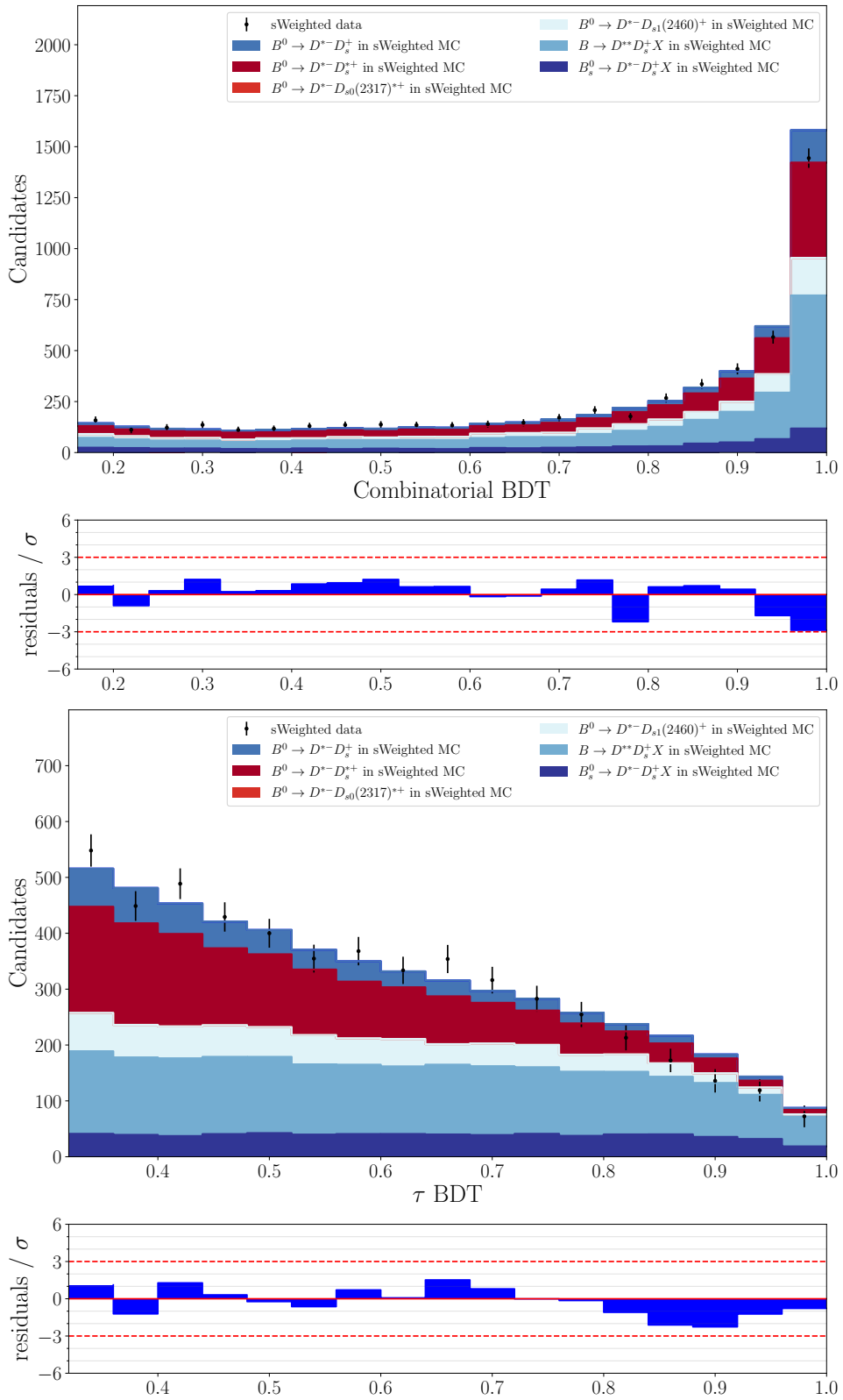


Figure 7.7: Comparison of $B \to D^*(D_s \to 3\pi)X$ data (points) and simulation (shaded histograms). The remaining discrepancies observed are corrected with a sequential reweighting procedure, the results of which are applied to the full $B \to D^* D_s X$ simulation sample.

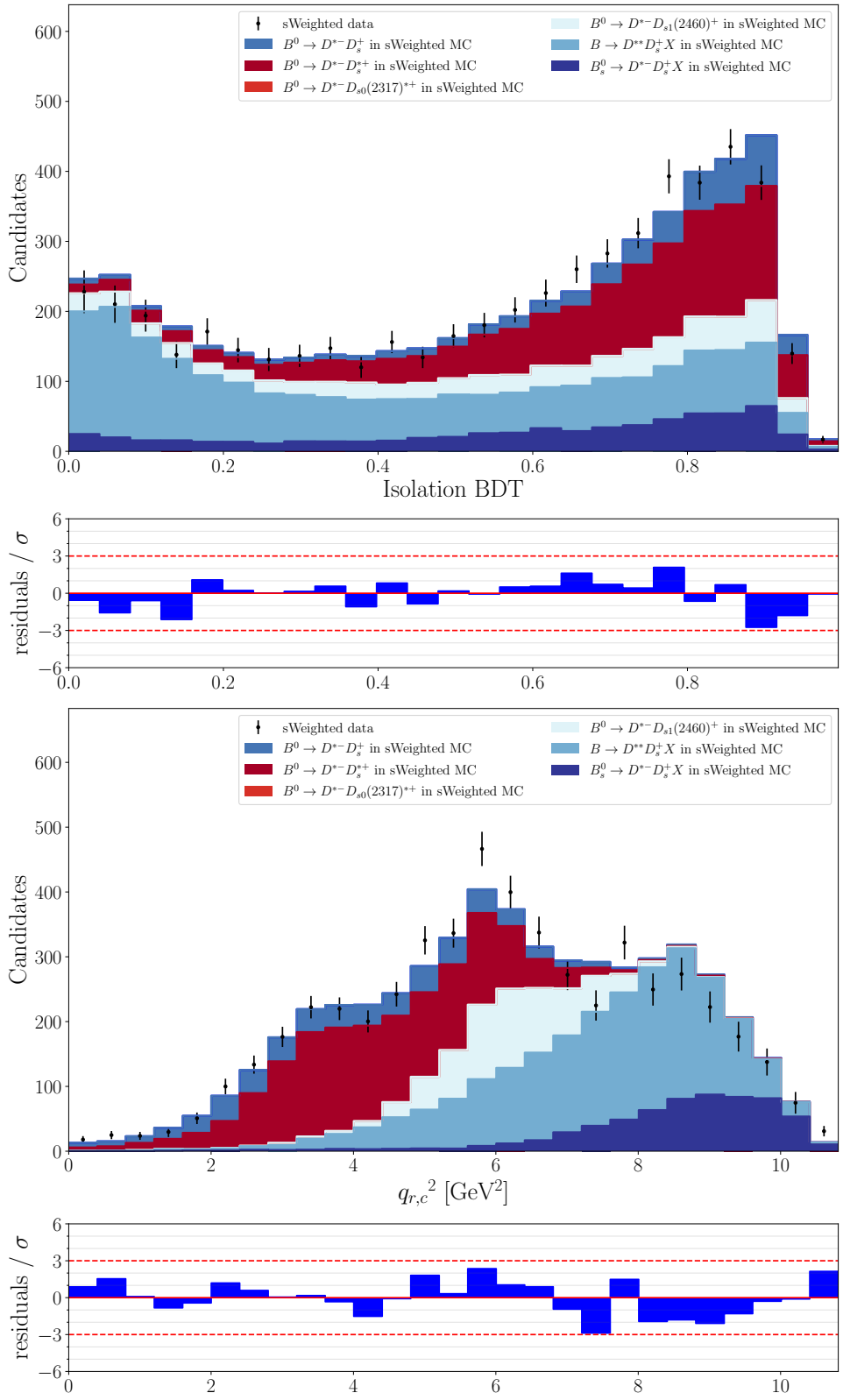


Figure 7.8: Comparison of $B \to D^*(D_s \to 3\pi)X$ data (points) and simulation (shaded histograms). The remaining discrepancies observed are corrected with a sequential reweighting procedure, the results of which are applied to the full $B \to D^* D_s X$ simulation sample (continued).

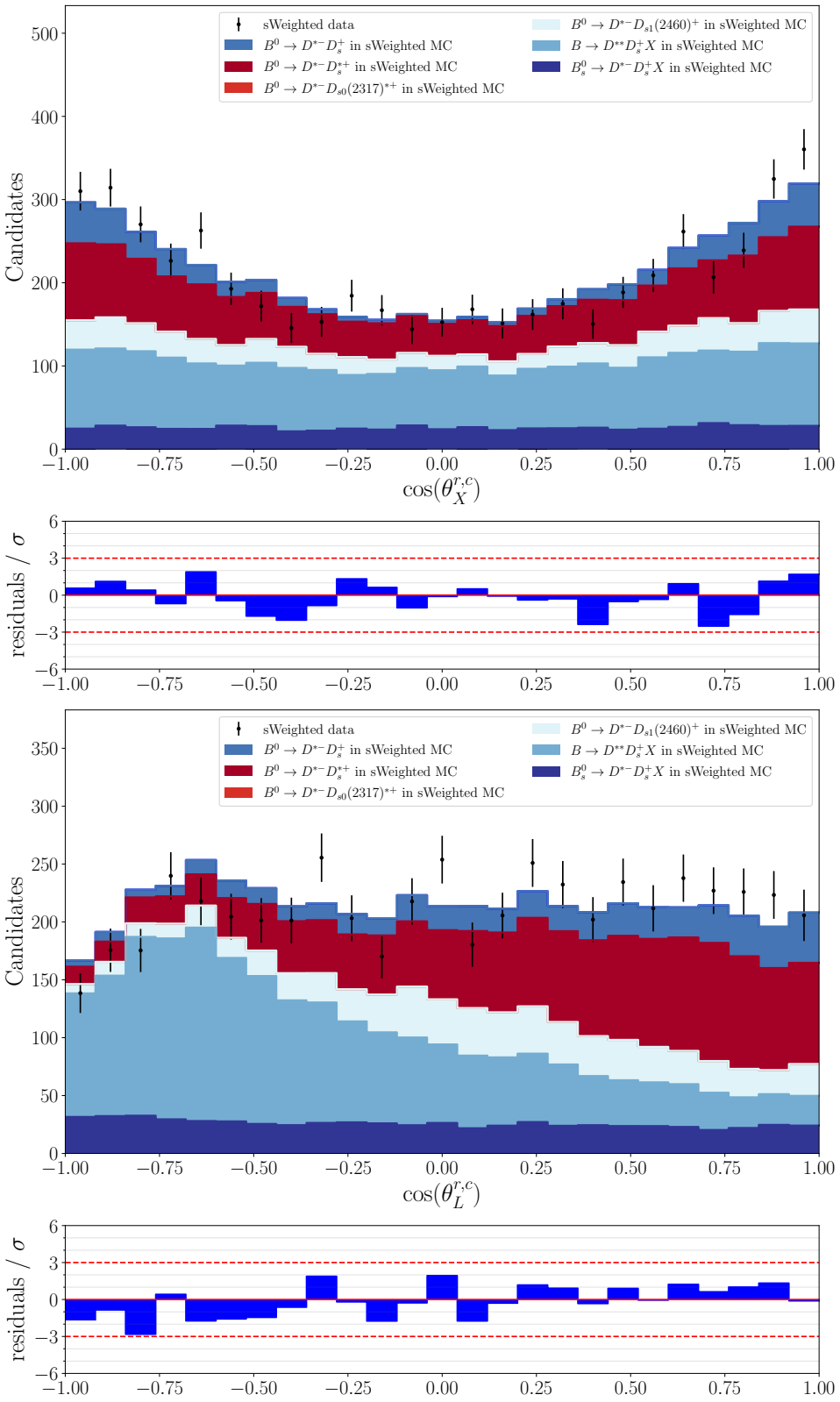


Figure 7.9: Comparison of $B \to D^*(D_s \to 3\pi)X$ data (points) and simulation (shaded histograms). The remaining discrepancies observed are corrected with a sequential reweighting procedure, the results of which are applied to the full $B \to D^* D_s X$ simulation sample (continued).

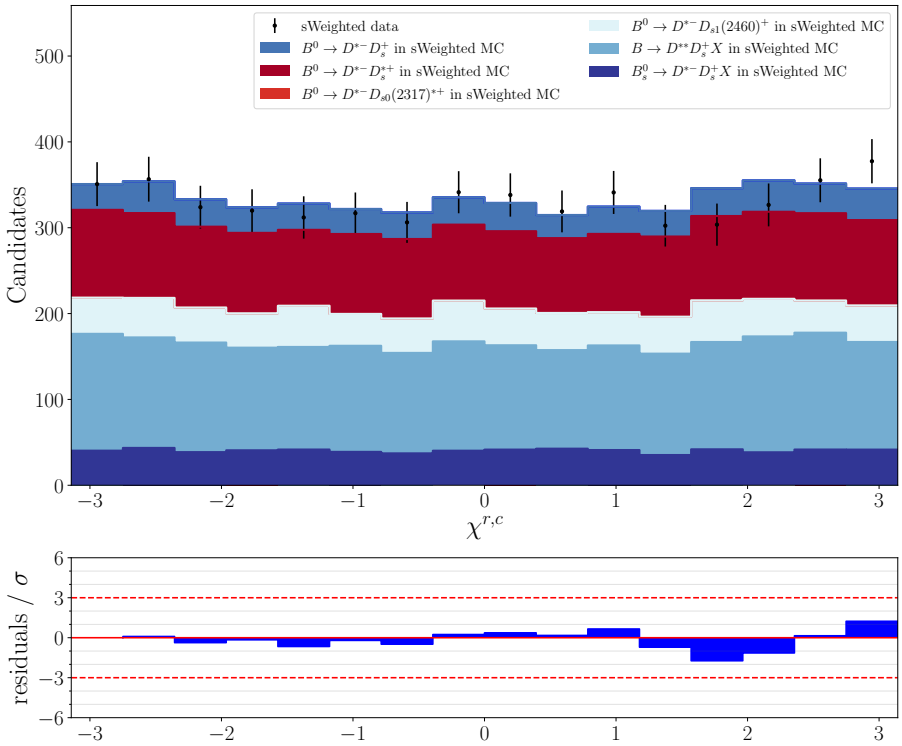


Figure 7.10: Comparison of $B \to D^*(D_s \to 3\pi)X$ data (points) and simulation (shaded histograms). The remaining discrepancies observed are corrected with a sequential reweighting procedure, the results of which are applied to the full $B \to D^* D_s X$ simulation sample (continued).

No significant changes to the reweighting output are found with altered variable ordering. The spline fits are shown in Figures 7.11 and 7.12, where the fits to the default data/simulation ratios are shown by the solid curve. The data and simulation samples are also bootstrapped many times (randomly sampled with replacement) in order to form a range of systematically varied ratios, each of which is fitted with a spline function. The results of many systematic fits, which are used to assess the systematic uncertainty from the spline reweighting, are shown by the shaded regions. To check that the sequential reweighting procedure is robust, a comparison of simulation with data before and after reweighting is performed using the $D_s \to 3\pi$ control samples. In all variables considered, the alignment between data and simulation is improved as a result of the reweighting, as shown in Figures 7.13 and 7.14. The splines derived from these control studies are used to correct the fully inclusive $B \to D^*(D_s \to 3\pi X)X$ samples, which are used in the angular fit to create background templates.

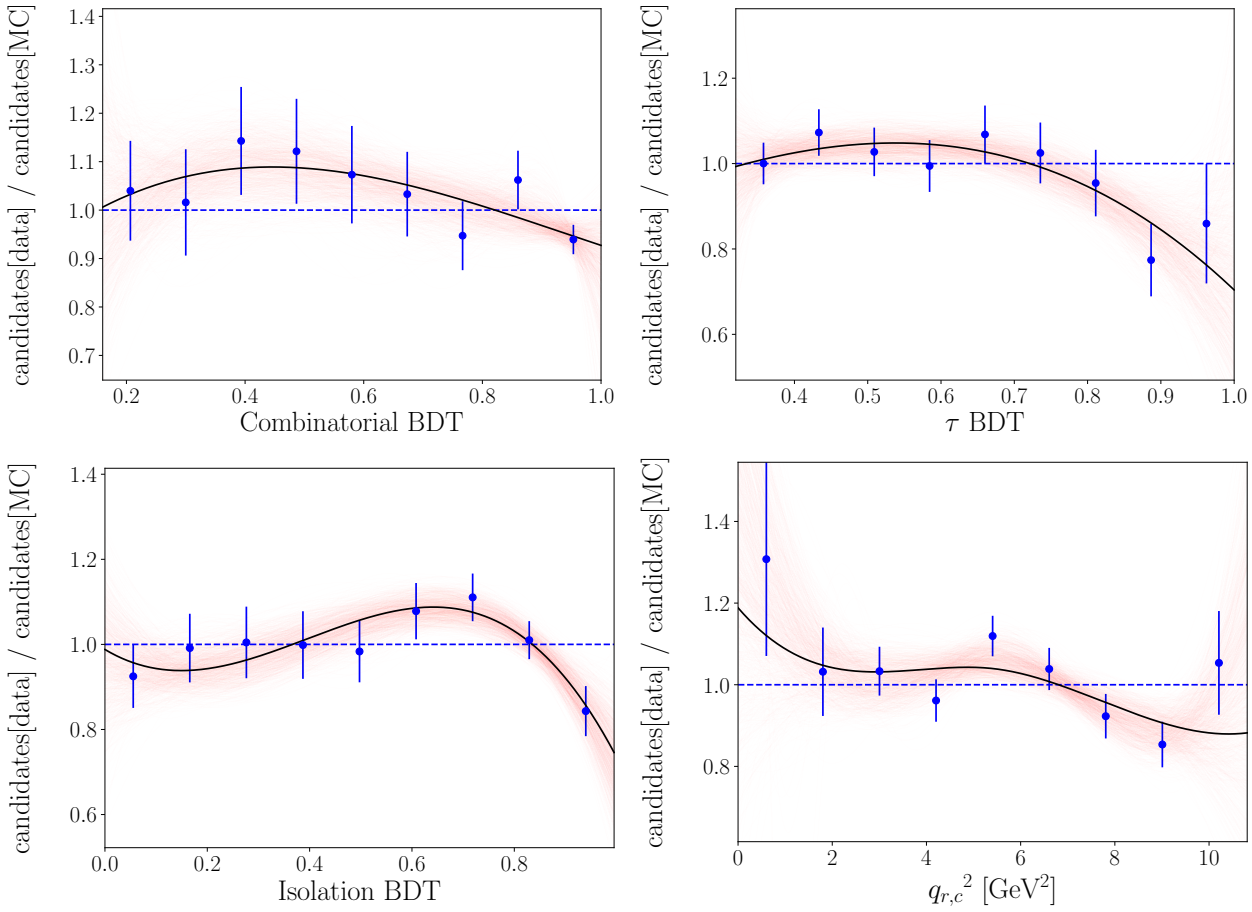


Figure 7.11: Binned data/simulation ratios for $B \to D^*(D_s \to 3\pi)X$, with cubic spline fits overlaid. The ratios shown are determined in an iterative manner, with spline-based corrections for the previous variables applied to simulation prior to their calculation. The spline fit to the default data/simulation ratio is shown by the solid curve, and the shaded region shows the fits to many bootstrapped data/simulation ratios for systematic uncertainty evaluation.

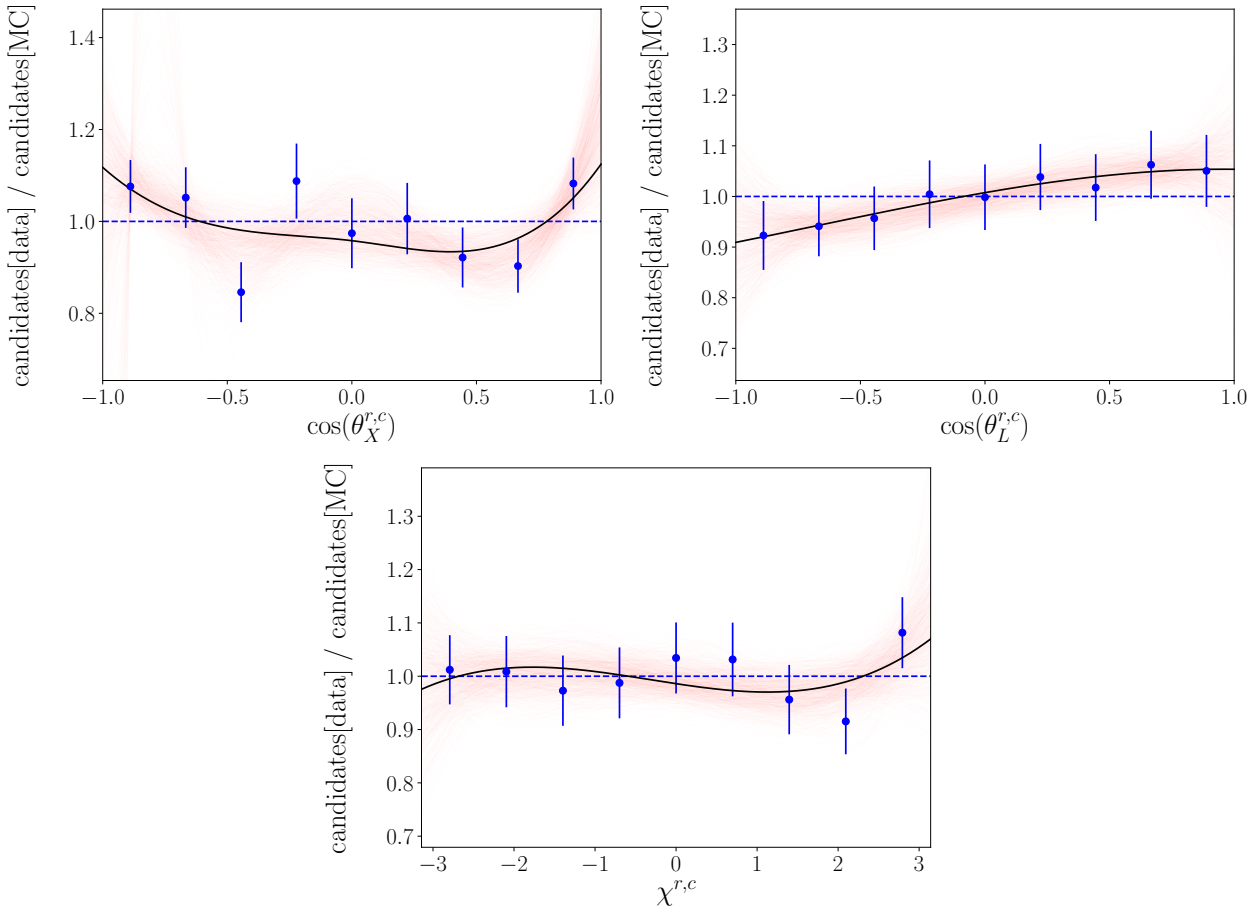


Figure 7.12: Binned data/simulation ratios for $B \to D^*(D_s \to 3\pi)X$, with cubic spline fits overlaid. The ratios shown are determined in an iterative manner, with spline-based corrections for the previous variables applied to simulation prior to their calculation. The spline fit to the default data/simulation ratio is shown by the solid curve, and the shaded region shows the fits to many bootstrapped data/simulation ratios for systematic uncertainty evaluation (continued).

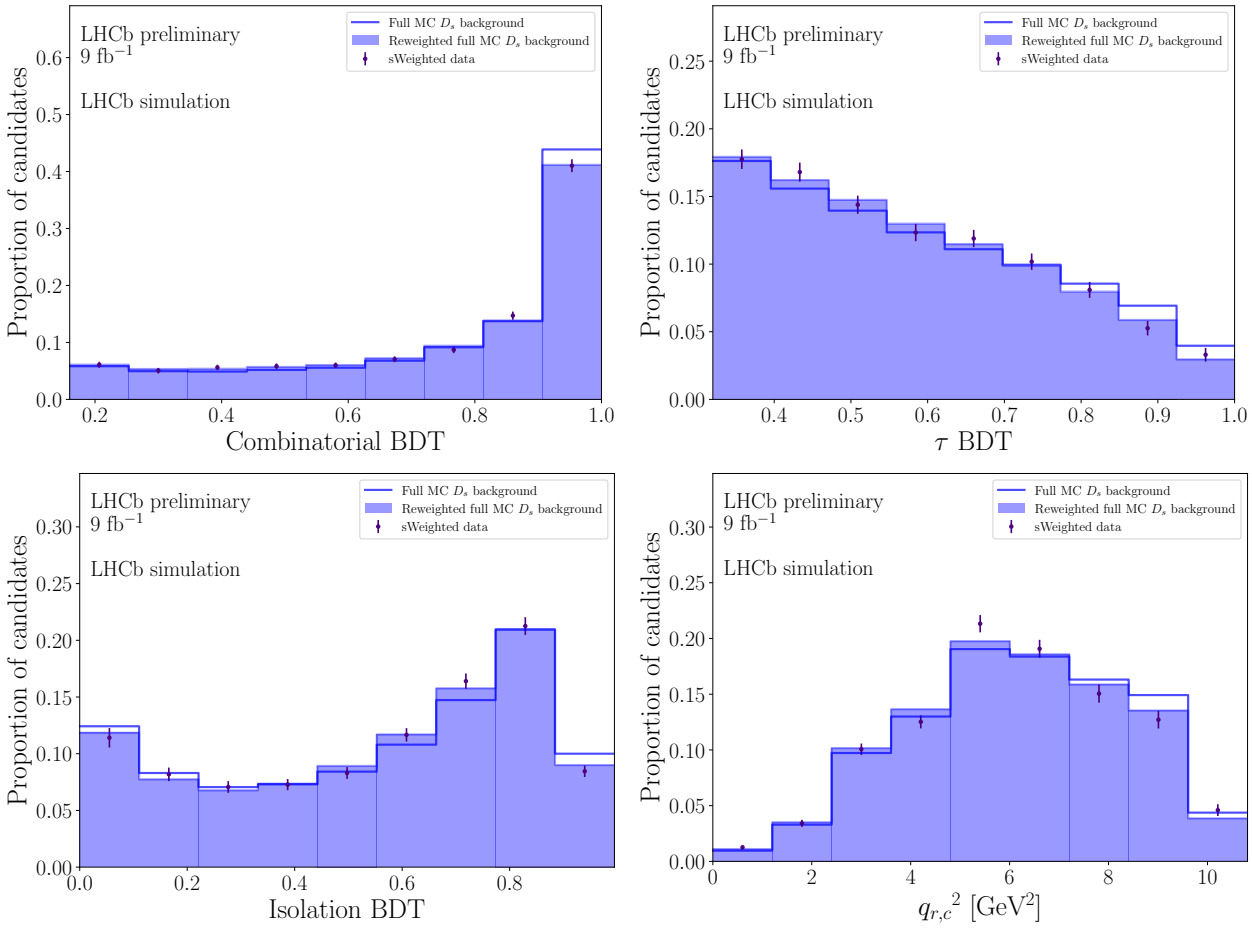


Figure 7.13: Comparison of $B \to D^*(D_s \to 3\pi)X$ data (points) and simulation before (blue) and after (blue shaded) the sequential reweighting splines are applied. The agreement in all variables is improved as a result of the reweighting. The splines are subsequently used to correct the full $B \to D^*(D_s \to 3\pi X)X$ sample.

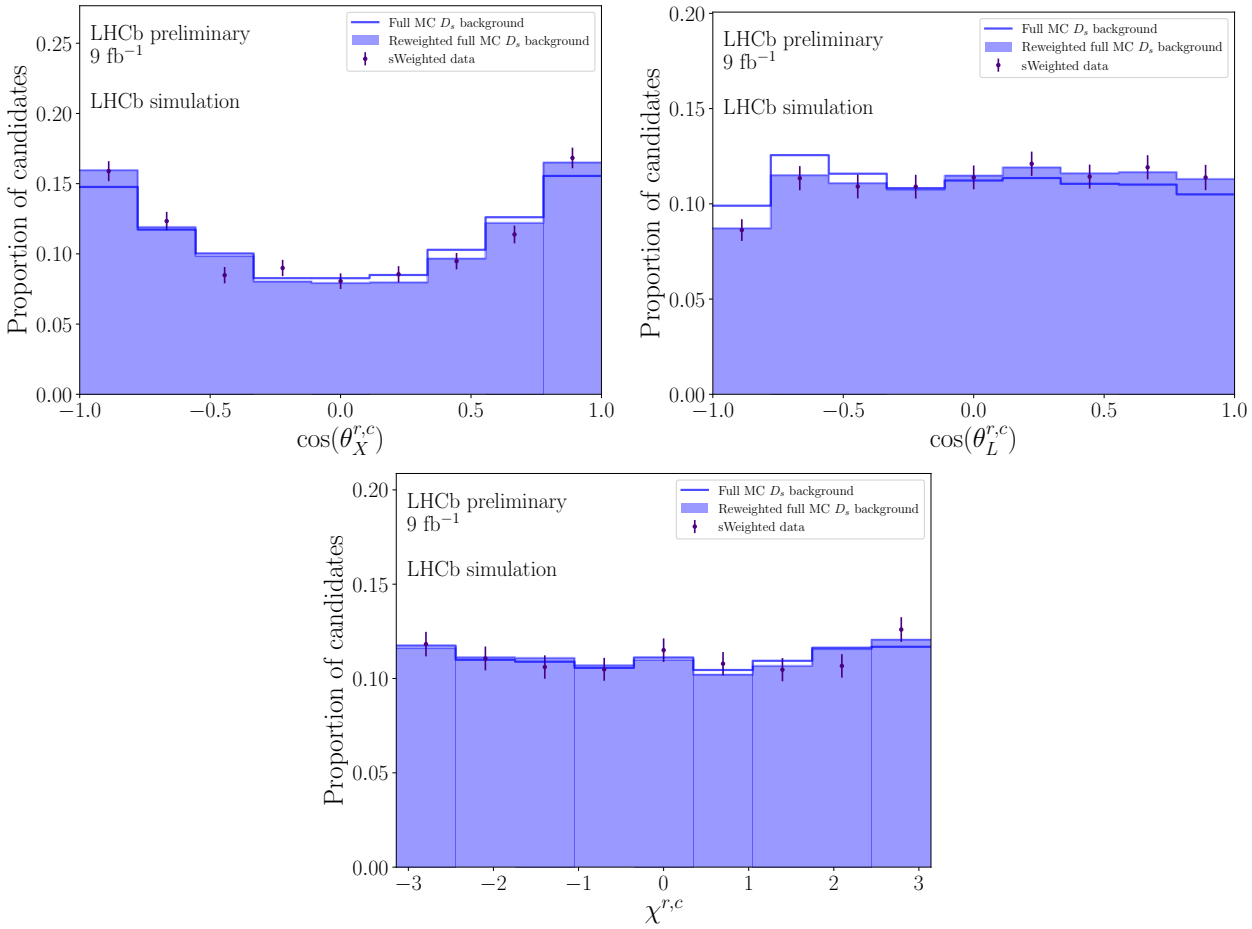


Figure 7.14: Comparison of $B \to D^*(D_s \to 3\pi)X$ data (points) and simulation before (blue) and after (blue shaded) the sequential reweighting splines are applied. The agreement in all variables is improved as a result of the reweighting. The splines are subsequently used to correct the full $B \to D^*(D_s \to 3\pi X)X$ sample (continued).

7.3 $B \to D^{*-}D^0(X)$ AND $B \to D^{*-}D^+(X)$ CONTROL MODES

The D^0 and D^+ double-charm backgrounds contribute sub-dominantly relative to the D_s^+ background. These backgrounds are generated inclusively in simulation, to final states including at least a D^* meson and an associated D^0 or D^+ meson, where the $D^{0/+}$ then decays to at least three charged pions. Decays with missing momentum are thus naturally included in the generation, arising either from the b-hadron decay or from the $D^{0/+}$ decay. To align the simulation with data for both of these backgrounds, control samples with fully reconstructed charm hadrons are selected as follows:

- $B \to D^*D^0X$: charged kaons are combined with selected 3π candidates via the isolation tool, and a sample of $B \to D^*(D^0 \to K3\pi)X$ decays selected using the $m(K3\pi)$ distribution.
- $B \to D^*D^+X$: for 3π candidates with positive (negative) charge, the pion of negative (positive) charge is assigned the kaon mass hypothesis. Samples of $B \to D^*(D^+ \to K^-\pi^+\pi^+)X$ decays are then selected using the $m(K\pi\pi)$ distribution.

After the control samples are selected in data and simulation, they are statistically compared in order to derive a set of correction weights. These weights, when applied to the full $B \to D^*D^{0/+}X$ samples, act to align the simulation with data across a range of high-level variables including the decay angles, 3π flight, q^2, and selection BDTs. Further details of the full procedure are provided below.[2]

7.3.1 SELECTION OF FULLY RECONSTRUCTED SAMPLES

To isolate $D^0 \to K3\pi$ decays both in data and the $B \to D^*D^0X$ simulated samples, unbinned invariant-mass fits to their $m(K3\pi)$ distributions are performed as shown in Figure 7.15. Prior to the fits, candidates with $m(D^*3\pi)$ values above 4700 MeV are excluded since such candidates do not arise from single B decays involving a missing kaon (they are combinatorial in nature) and their inclusion undermines

[2]The initial version of this control study was performed by Anthony Correia.

subsequent data/simulation comparisons. The fit to simulation is performed first, enabling tail parameters from the double Crystal Ball signal PDF to be Gaussian constrained in the subsequent data fit. ${}_s$Weights derived from these mass fits are used to isolate pure $B \to D^*(D^0 \to K3\pi)X$ samples in data and simulation, which can then be statistically compared in order to derive simulation correction weights.

To isolate a pure sample of $B \to D^*(D^+ \to K^-\pi\pi)X$ decays in simulation, true simulation information is used to select decays where a genuine $D^+ \to K^-\pi^+\pi^+$ decay was reconstructed as a 3π. To select $B \to D^*(D^+ \to K^-\pi\pi)X$ decays in data, an ${}_s$Weight fit to the $m(K\pi\pi)$ distribution of selected candidates is performed. As in the D^0 case above, an initial fit to the $D^+ \to K^-\pi^+\pi^+$ peak in simulation is first performed in order to derive Gaussian constraints for the tail parameters. The fits to simulation and data are shown in Figure 7.16.

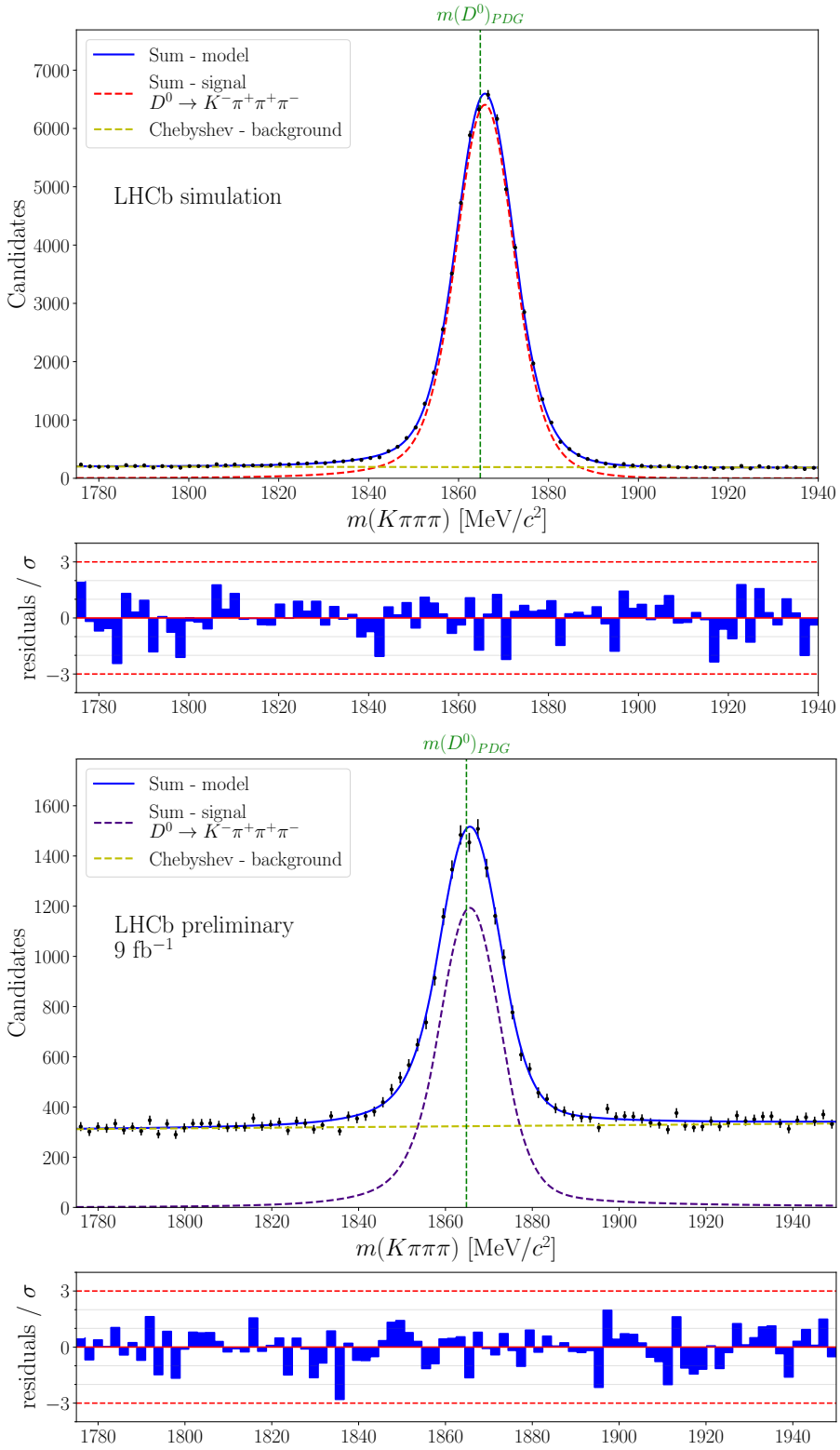


Figure 7.15: Invariant-mass fits in simulation (top) and data (bottom) to the $m(K3\pi)$ distribution of selected candidates. $_s$Weights derived from these fits are used to select control samples of $B \to D^*(D^0 \to K3\pi)X$ decays.

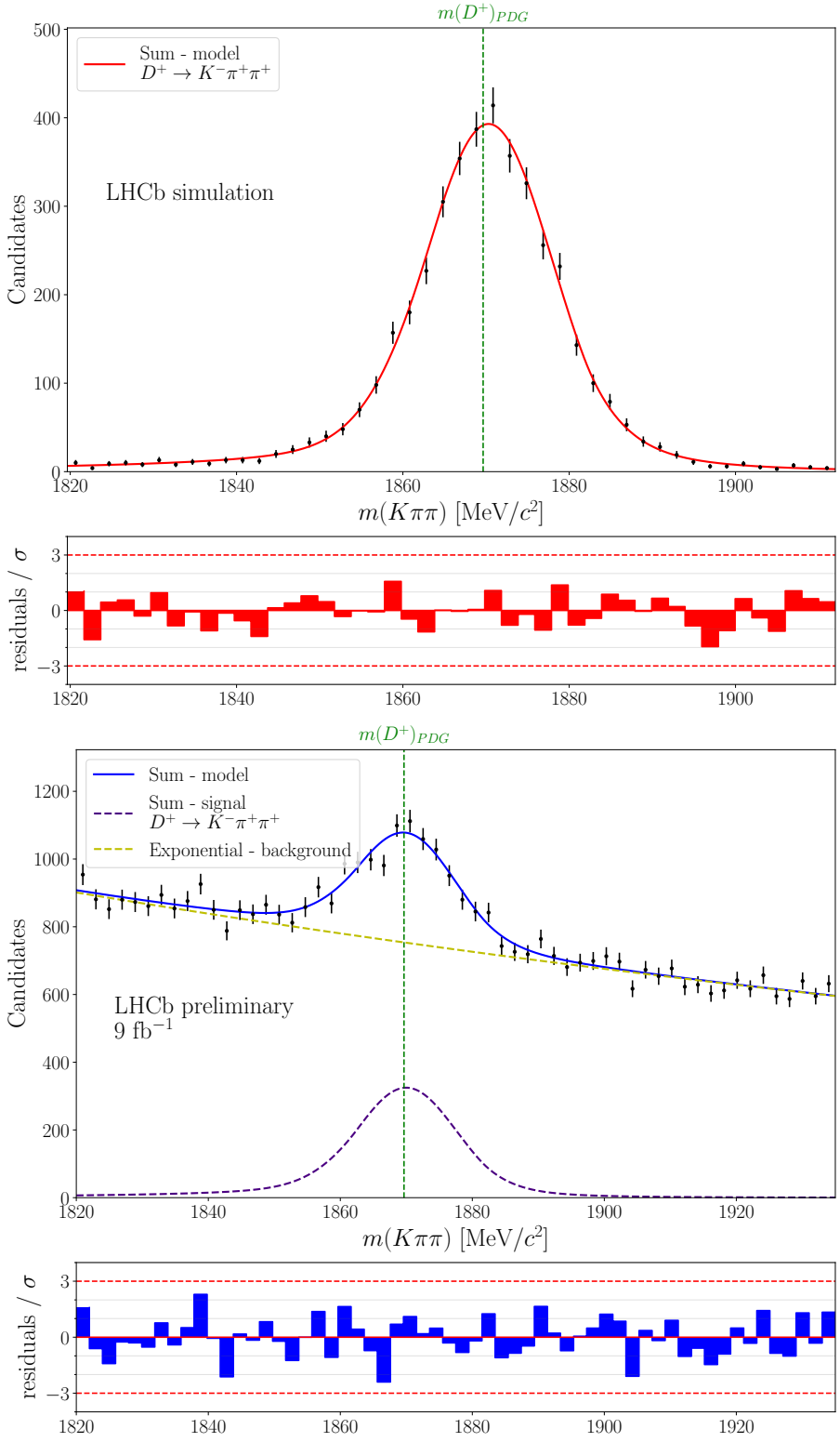


Figure 7.16: Invariant-mass fits in simulation (top) and data (bottom) to the $m(K\pi\pi)$ distribution of selected candidates. ${}_s$Weights derived from the data fit are used to select a control sample of $B \to D^*(D^+ \to K\pi\pi)X$ decays in data, which can be compared to the equivalent truth-matched sample in simulation.

7.3.2 CORRECTION OF SIMULATION

After the selection of the fully reconstructed charm sub-samples as described above, data and simulation are compared across a range of key analysis variables. The purpose of the comparison is to derive a set of weights which can be applied to the full $B \to D^*D^{0/+}X$ simulation samples, in order to better align them with data. This procedure aims to correct differences between data and simulation arising from an incomplete set of B decays in the inclusive generation, imperfect modelling of the decay dynamics in simulation (e.g. D^* polarisation, helicity amplitudes), and differences in the detector resolution and acceptance.

The same reweighting procedure as the $B \to D^*D_sX$ control studies is used to derive correction weights. The net result of this procedure is a set of correction splines, which when applied to each $B \to D^*D^{0/+}X$ simulation event provide a single weight that acts to align all of the considered variables simultaneously. This sequential reweighting approach offers good control over the reweighting performance, and more robust results than those achieved using a binned or BDT-based reweighting procedure. No significant changes to the reweighting output are found with altered variable ordering. In the case of $B \to D^*(D^0 \to K3\pi)X$, the Isolation BDT is not included in the reweighting: the addition of a charged kaon retains only candidates with low Isolation BDT scores, which do not cover the full variable range. However, in the limited Isolation BDT range covered by the $D^0 \to K3\pi$ control samples, excellent agreement between data and simulation is observed as seen in Figures 7.19 and 7.20.

The spline fits for $B \to D^*D^0X$ are shown in Figure 7.17, and the fits for $B \to D^*D^+X$ are shown in Figure 7.18, where the spline fits to the default data/simulation ratios are shown by the solid curve. The data and simulation samples are also bootstrapped many times in order to form a range of systematically varied ratios, each of which is fitted with a spline function. The results of many systematic fits, which are used to assess the systematic uncertainty from the spline reweighting, are shown by the shaded regions in Figures 7.17 and 7.18.

To check that the sequential reweighting procedure is robust, a comparison of simulation with data before and after reweighting is performed using the $D^0 \to K3\pi$

and $D^+ \to K\pi\pi$ control samples. In all variables considered, the alignment between data and simulation is improved as a result of the reweighting, as shown in Figures 7.19, 7.20, 7.21, and 7.22. The splines derived from these control studies are used to correct the fully inclusive $B \to D^*(D^0 \to 3\pi X)X$ and $B \to D^*(D^+ \to 3\pi X)X$ samples, which are used in the angular fit to create background templates.

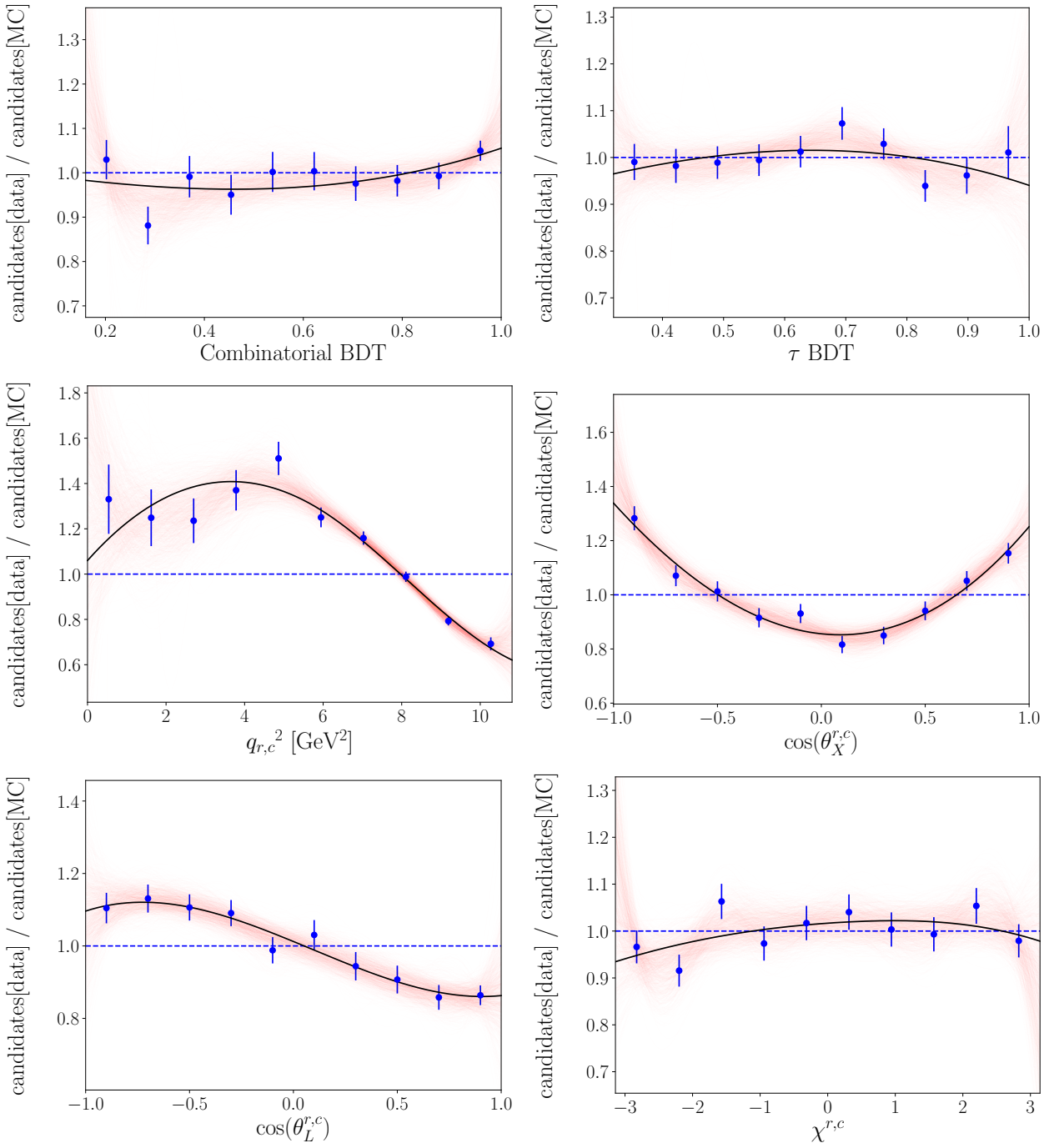


Figure 7.17: Binned data/simulation ratios for $B \to D^*(D^0 \to K3\pi)X$, with cubic spline fits overlaid. The ratios shown are determined in an iterative manner, with spline-based corrections for the previous variables applied to simulation prior to their calculation. The spline fit to the default data/simulation ratio is shown by the black curve, and the orange shaded regions show the systematically varied spline fits to many bootstrapped ratios.

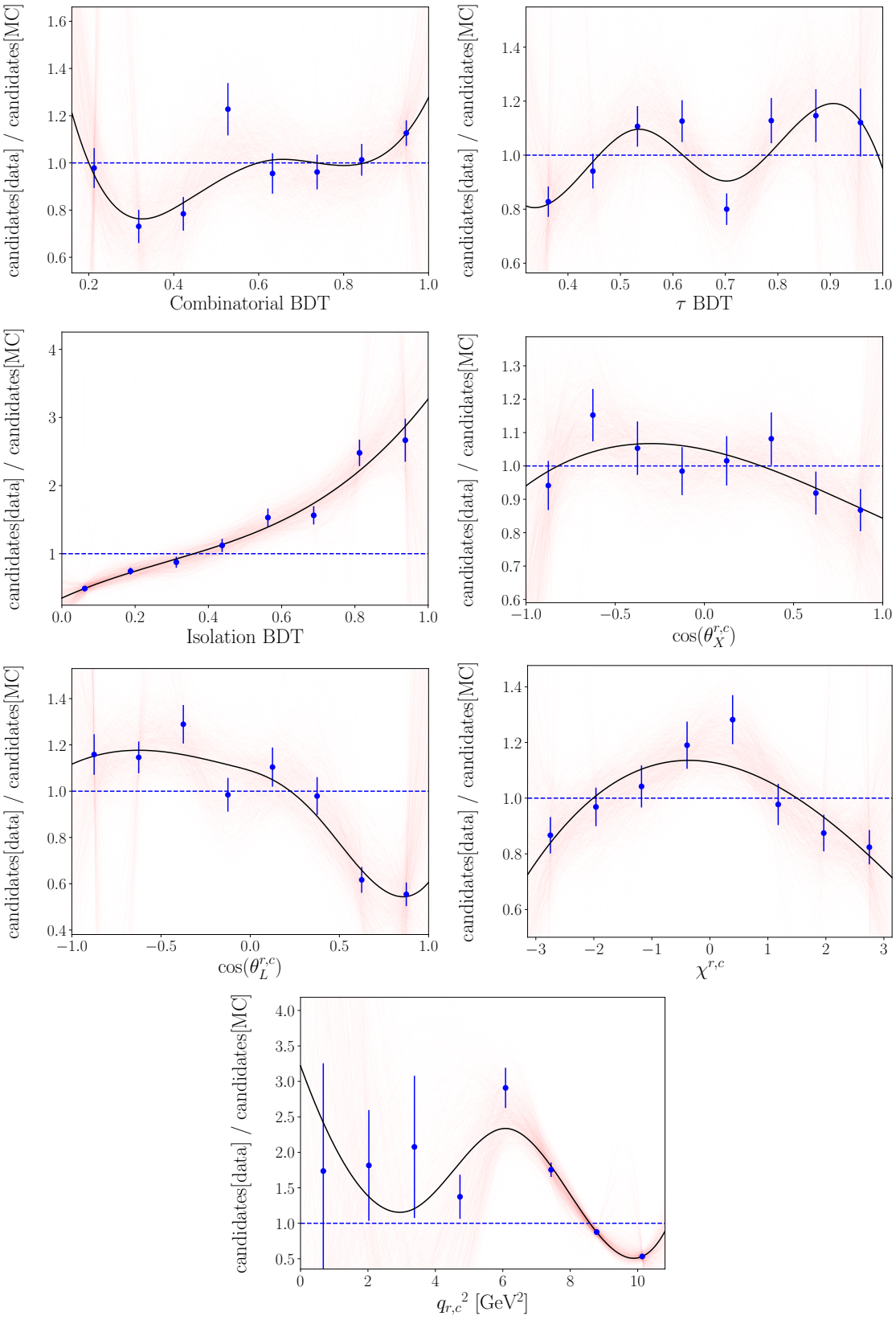


Figure 7.18: Binned data/simulation ratios for $B \to D^*(D^+ \to K\pi\pi)X$, with cubic spline fits overlaid. The ratios shown are determined in an iterative manner, with spline-based corrections for the previous variables applied to simulation prior to their calculation. The spline fit to the default data/simulation ratio is shown by the black curve, and the orange-shaded regions show the systematically varied spline fits to many bootstrapped ratios.

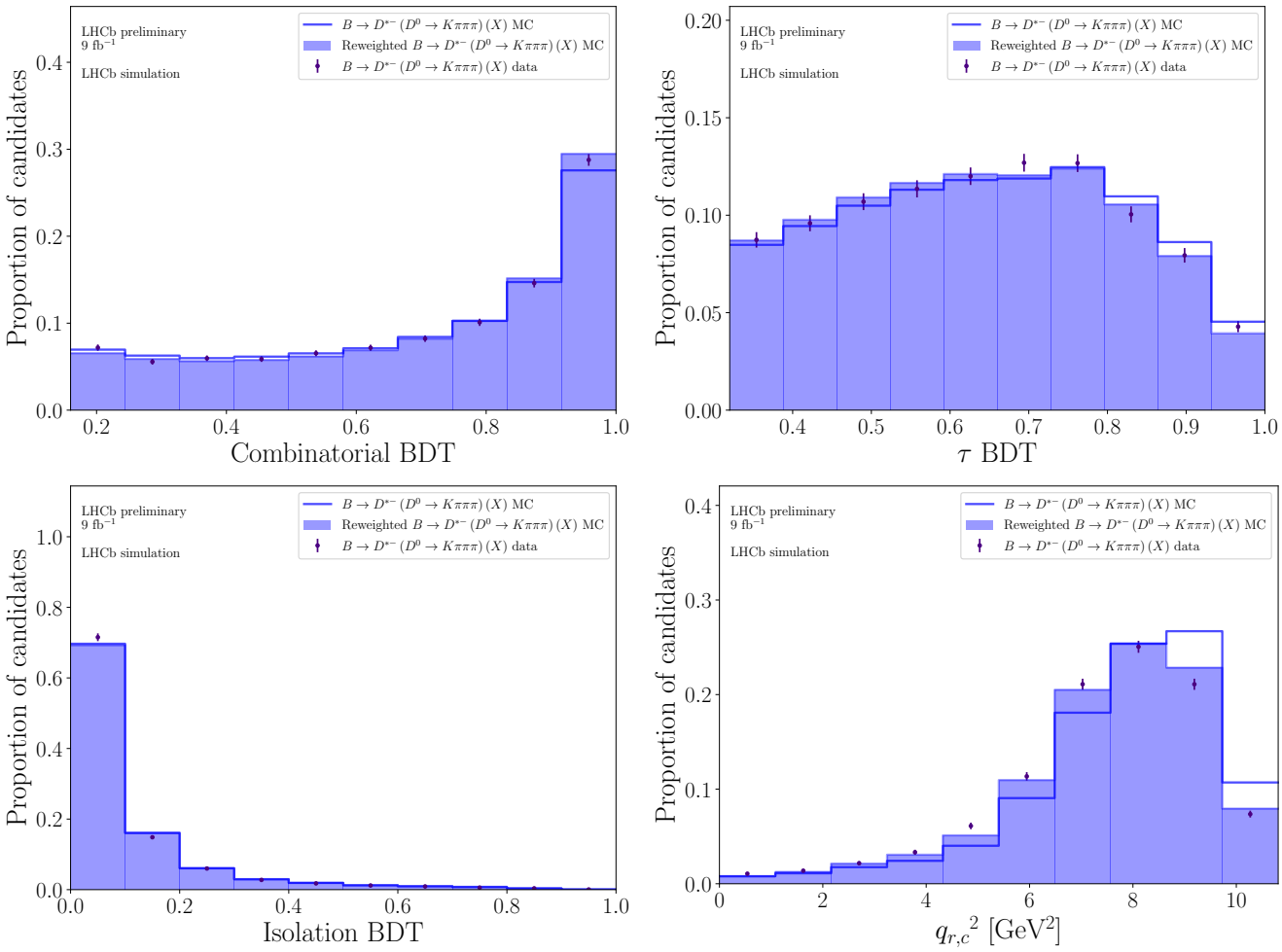


Figure 7.19: Comparison of data (points) with simulation both before (blue) and after (blue shaded) the spline-based reweighting is applied to the $B \to D^*(D^0 \to K3\pi)X$ simulation. The improvement in the agreement between data and simulation indicates the iterative reweighting procedure functions as expected. The Isolation BDT is not included in the $B \to D^* D^0 X$ reweighting, its distribution is included to show the good agreement irrespective of the reweighting.

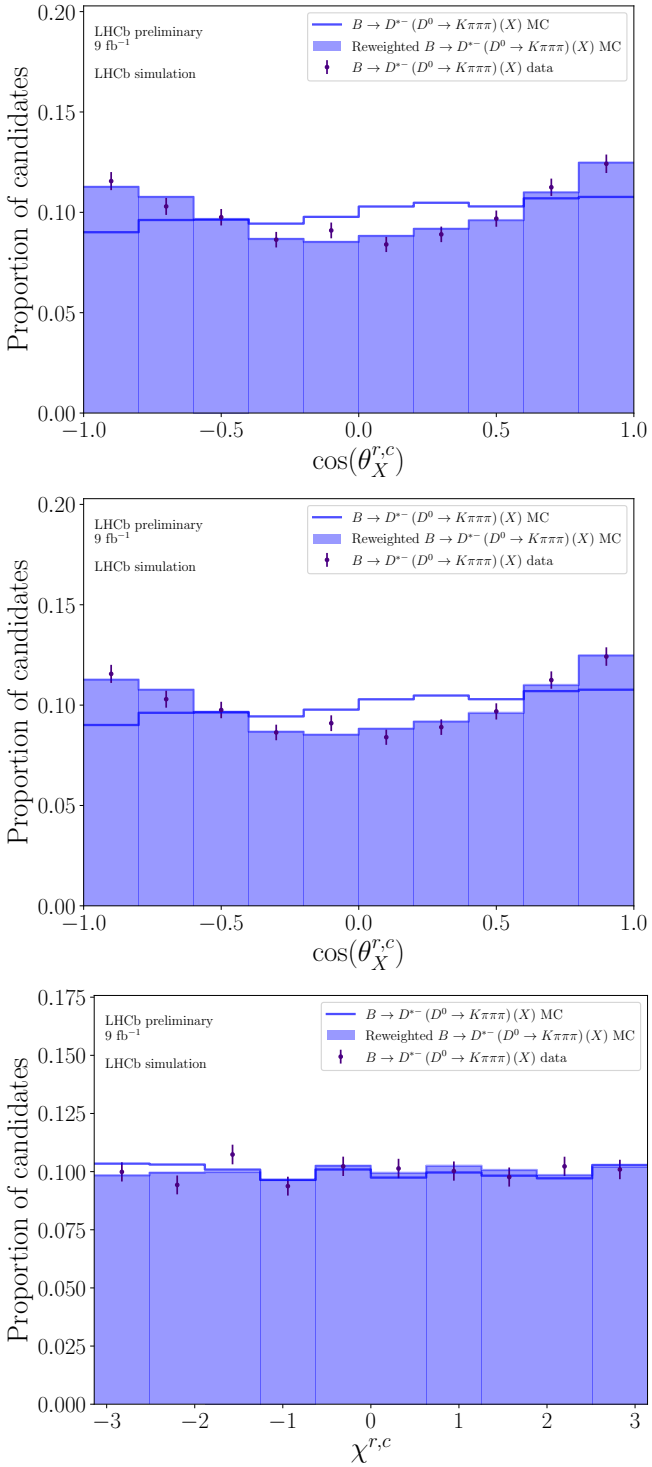


Figure 7.20: Comparison of data (points) with simulation both before (blue) and after (blue shaded) the spline-based reweighting is applied to the $B \to D^*(D^0 \to K3\pi)X$ simulation. The improvement in the agreement between data and simulation indicates the iterative reweighting procedure functions as expected (continued).

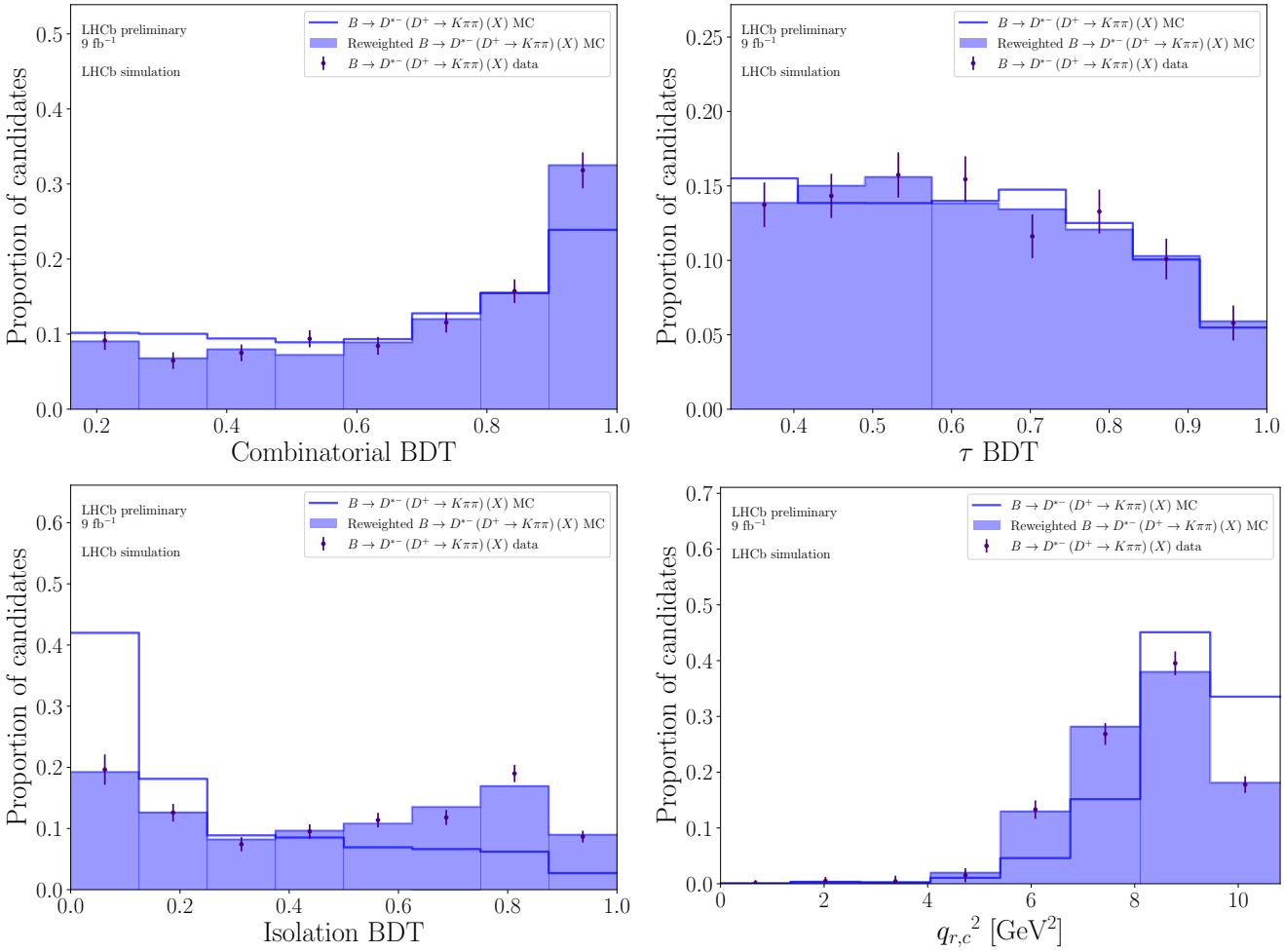


Figure 7.21: Comparison of data (points) with simulation both before (blue) and after (blue shaded) the spline-based reweighting is applied to the $B \to D^*(D^+ \to K\pi\pi)X$ simulation. The improvement in the agreement between data and simulation is evident across all variables, indicating that the iterative reweighting procedure functions as expected.

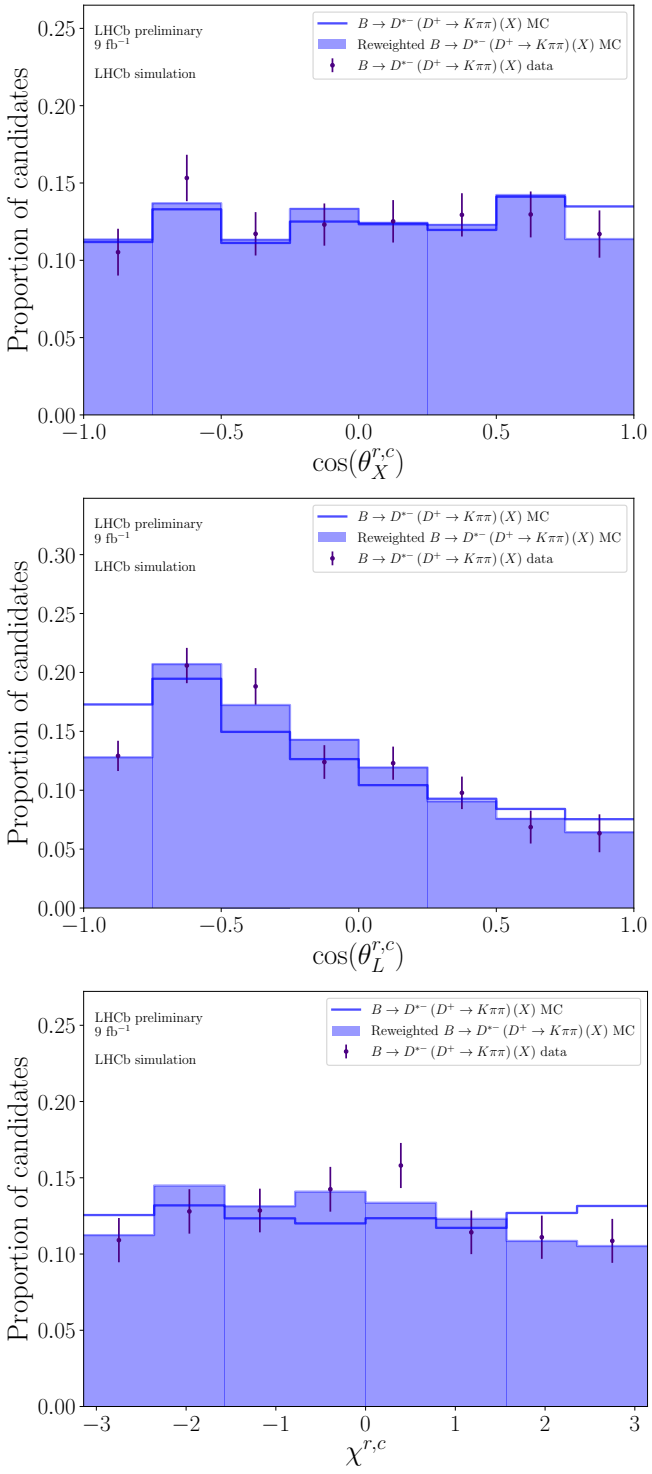


Figure 7.22: Comparison of data (points) with simulation both before (blue) and after (blue shaded) the spline-based reweighting is applied to the $B \to D^*(D^+ \to K\pi\pi)X$ simulation. The improvement in the agreement between data and simulation is evident across all variables, indicating that the iterative reweighting procedure functions as expected (continued).

7.4 $B \to D^{*-}3\pi$ control mode for L0 reweighting

A comparison of $B \to D^{*-}3\pi$ data and simulation is performed in order to compare their relative L0 trigger category compositions. The following mutually exclusive L0 categories are defined: L0 Hadron TOS only, L0 Global TIS only, and finally L0 Hadron TOS and L0 Global TIS together. The proportion of data and simulation in each L0 category are calculated with binomial uncertainties assigned, and the data/simulation ratios are calculated. These ratios can be used to reweight the L0 proportions in the simulated samples used to create templates for the angular fit, in order to align with data.

For the comparison, $B \to D^{*-}3\pi$ events in data are isolated using ${}_s$Weights from the normalisation fit described in Chapter 6. This sample has the full event selection applied, bar the 3π z flight distance significance requirement. The simulated sample has the same selection requirements applied, with truth-matching requirements applied to the D^* and the B. Events in the simulation are weighted according to year, in order to align the total simulated sample composition with the luminosity and cross-sections in data. To calculate the trigger category compositions in data, the ${}_s$Weights are summed in a given trigger category and divided by the total sum of ${}_s$Weights. In simulation, the sum of weights for events in a given trigger category is divided by the total sum of weights across all L0 categories.

The L0 category fractions are shown in Table 7.5, along with the ratios used to correct all simulated samples used to create templates in the angular fit. The TOS and TIS exclusive categories are slightly enhanced in data relative to simulation, with a corresponding under-representation of events triggering both TOS and TIS. The L0 weights assigned to the simulation are varied within their corresponding uncertainties as a source of systematic uncertainty in the angular fit.

Table 7.5: Proportion of $B \to D^*3\pi$ data and simulation events in each L0 category, and the data/simulation ratios calculated from these values. The data/simulation ratios are used to correct the L0 composition of all simulated samples used in the angular fit.

L0 category	Data proportion (%)	Simulation proportion (%)	Data/Simulation
TOS only	0.481 ± 0.008	0.474 ± 0.005	1.013 ± 0.020
TIS only	0.273 ± 0.007	0.251 ± 0.004	1.090 ± 0.035
TOS & TIS	0.246 ± 0.007	0.275 ± 0.004	0.895 ± 0.029

7.5 Comparison of $b\bar{b} \to D^*3\pi X$ inclusive simulation and data

The $B \to D^*DX$ control studies detailed above rely upon the use of fully reconstructed $3h$ systems, namely from:

- $D^0 \to K3\pi$ in $B \to D^*D^0X$;
- $D^+ \to K\pi\pi$ in $B \to D^*D^+X$;
- $D_s \to 3\pi$ in $B \to D^*D_sX$;
- Prompt 3π in $B \to D^*3\pi$.

Although these modes provide excellent control of the B decay distributions (angles, q^2, BDTs), they do not offer sufficient control over the $3\pi X$ dynamics present in the final angular fit dataset. To align the 3π kinematics and flight distributions in simulation and data prior to the angular fit, a comparison of data and inclusive $b\bar{b} \to D^*3\pi X$ simulation is performed. The comparison is based on the same sequential spline reweighting techniques as presented earlier and uses the following variables:

- 3π momentum (visible momentum of the reconstructed 3π candidate, not the τ momentum estimate);
- 3π transverse momentum (visible transverse momentum of the reconstructed 3π candidate, not the τ momentum estimate);
- 3π z flight distance significance.

The benefit of using inclusive simulation for this comparison is that a wide range of $3\pi X$ modes are included, which is more representative of the $3\pi X$ content present in data. The use of inclusive simulation is justified by comparing the 3π system mass distributions in data and inclusive simulation, as shown in Figure 7.23. For this comparison, data with ${}_s$Weights applied from the 2D ΔM vs. $m(D^0)$ fit are displayed, corresponding to the final dataset used in the angular fit. The inclusive simulation with a true D^* requirement and all selection requirements applied is also shown. A good agreement between data and inclusive simulation is seen, indicating that the $3\pi X$ content in the inclusive simulation is representative of data. As such, a

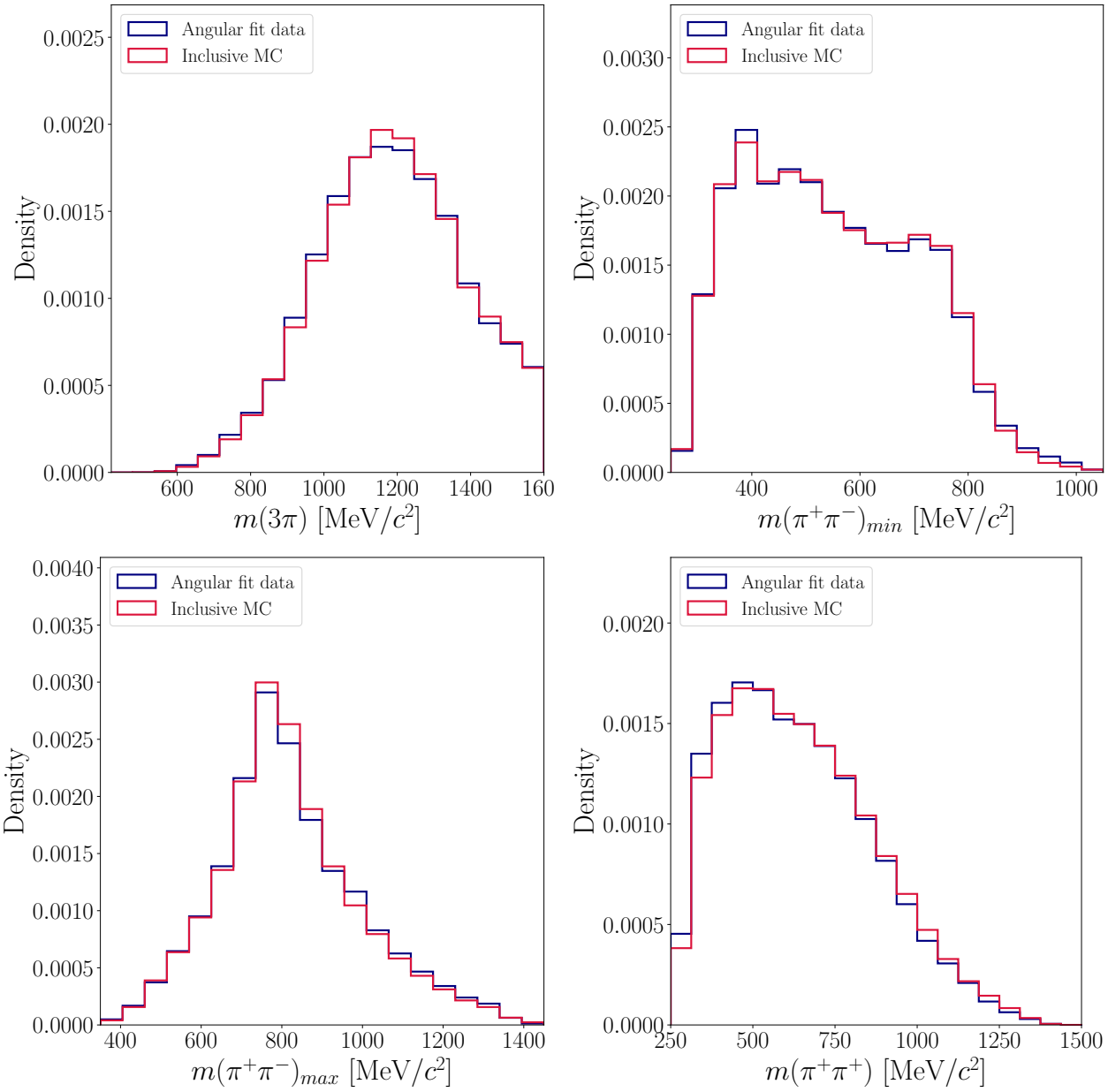


Figure 7.23: Comparison of 3π system mass distributions in the final angular fit dataset (blue histograms) and in inclusive $b\bar{b} \to D^*3\pi X$ simulation (red histograms). The good agreement shown indicates that the $3\pi X$ content in the inclusive simulation is representative of data, and thus that a data/simulation comparison is justified in order to derive simulation correction weights.

comparison of the 3π kinematics and flight in order to derive data-driven correction weights is deemed appropriate.

The spline fits performed to the data/inclusive simulation ratios are shown in Figure 7.24, and a comparison of data and inclusive simulation before and after the reweighting procedure is shown in Figure 7.25. The splines derived from this

comparison are subsequently applied to all simulated samples used in the angular fit to derive templates, in order to align the 3π kinematics and flight distributions with data. Two sources of systematic uncertainty on the spline corrections are considered:

- The finite statistics of the data and inclusive simulation samples used to form the binned ratios;
- The specific signal fraction present in the inclusive simulation sample could introduce a degree of signal model dependence in the spline corrections.

To account for the first of these effects, both the data and inclusive simulation samples are bootstrapped many times in order to create a statistically varied set of ratios. To account for the second effect, the signal component of each bootstrapped simulation sample is assigned a random uniform weight in the range $(0, 2)$. This ensures that the signal fraction varies between zero and twice its default value in the bootstrapped inclusive simulation sample, which propagates a large signal uncertainty into the bootstrapped ratios. Finally, spline fits to each ratio are performed, building an envelope of spline variations around the default curves as shown in Figure 7.24.

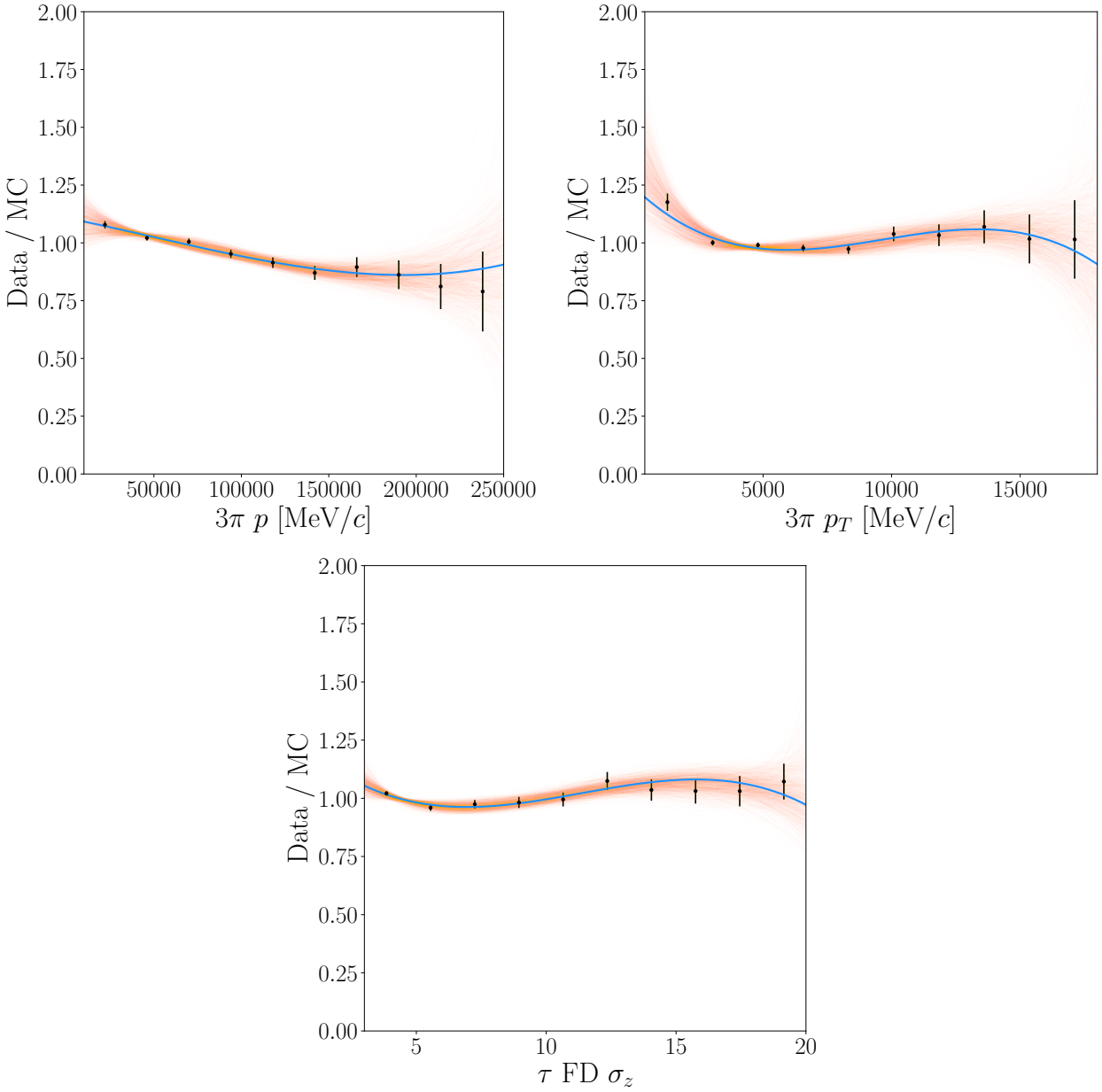


Figure 7.24: Spline fits to data/inclusive simulation ratios in the 3π kinematics and flight. The spline fit to the default data/simulation ratio is shown in blue. The variation in splines found is shown by the orange-shaded region where spline fits were run with bootstrapping applied to the data and simulation samples.

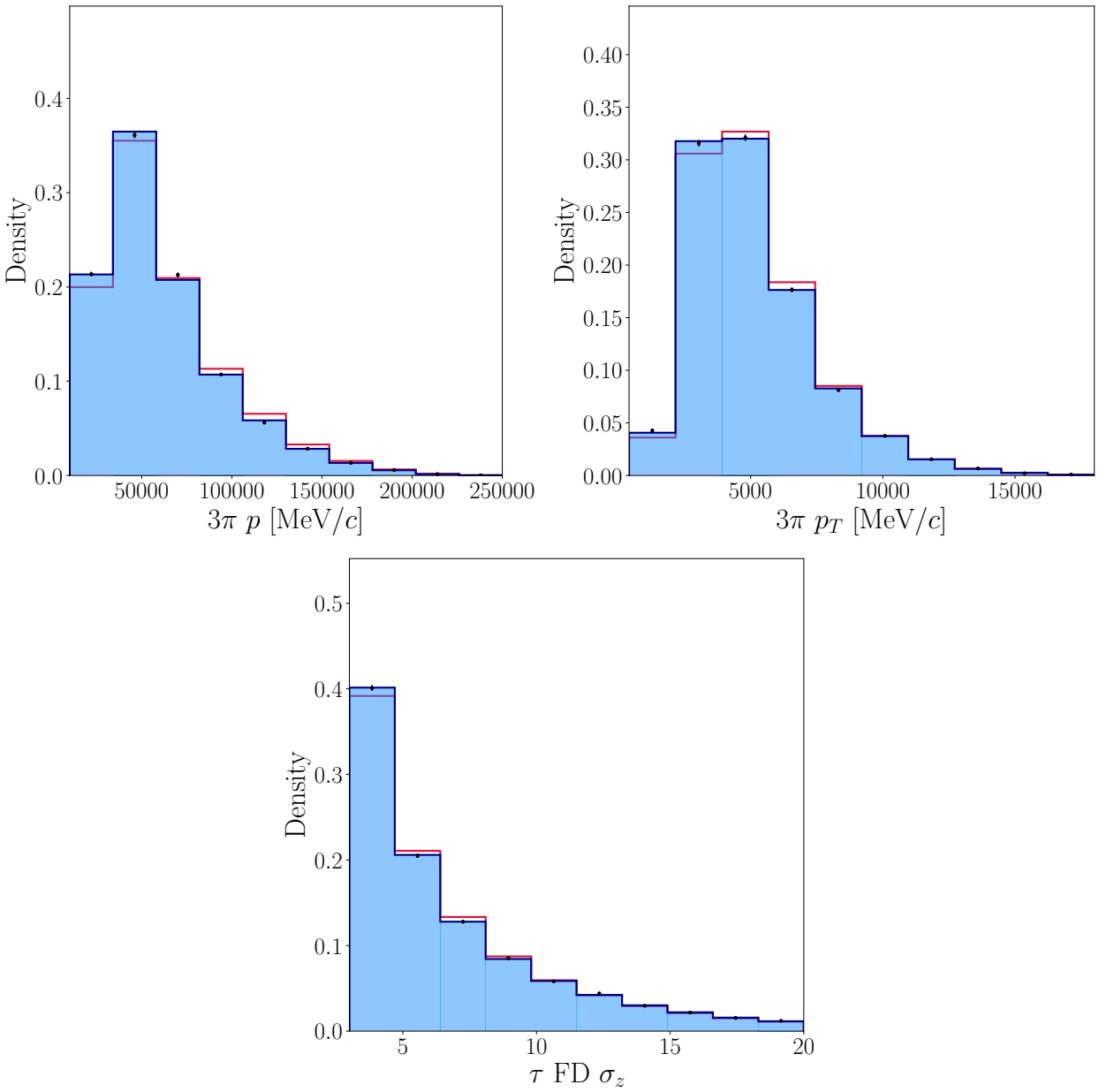


Figure 7.25: Comparison of data (points) with inclusive simulation before (red histograms) and after (shaded blue histograms) the sequential spline reweighting is applied. The same splines are used to correct all simulated samples from which templates are derived in the angular fit.

Part III

ANGULAR COEFFICIENT MEASUREMENT

8 ANGULAR ANALYSIS

To measure the angular coefficients, a binned maximum-likelihood fit to the data passing all selection requirements is performed using zfit [83]. The data histograms used in the fit are formed by applying per-event $_s$Weights derived from the 2D fit to $m(D^* - D^0)$ and $m(D^0)$, in order to statistically subtract the fake D^* and fake D^0 contributions (see Section 5.7). The fit is performed simultaneously across the following multidimensional distributions:

- Reconstructed decay angles $\cos(\theta_X)$, $\cos(\theta_L)$, and χ;
- Reconstructed τ candidate z flight distance significance and the Isolation BDT.

In this setup, the signal fraction and angular coefficients are shared across both fit categories: the decay angle 3D distribution provides sensitivity to the I_x coefficients, while the other 2D distribution helps to constrain the total signal yield and background fractions. The parameterisation of the total signal template as a sum of twelve component histograms, each scaled by a corresponding I_x observable, is detailed in Section 8.2 and follows the formalism previously published by the authors in Reference [30]. The component histograms for each multidimensional fit category are projected out of the full signal simulation sample in an identical manner using per-event weights. In this way, the change in total signal template shape due to variations in the I_x are fully captured not only in the decay angle distributions but also in the tau flight significance and BDT distributions. As such, no assumption of the underlying signal model is made in the fit.

Since the fit uses the same data across both multidimensional categories, the fit statistical uncertainties reported by Hesse are underestimated [84]. To correct for this double-counting effect, many bootstrapped fits to the default data are performed by sampling the data with replacement. The widths of the resulting bootstrap distributions for each freely varying fit parameter reflect the genuine statistical uncertainty arising from the finite data sample since the random variation introduced is common across both 3D fit categories. The bootstrapping is further discussed in Section 8.5.

Fit stability is assessed via a series of pseudo-experiments, where pseudo-experiment datasets are generated via per-bin Poisson variation of the final data fit total PDF. These studies check for potential biases in the central values and uncertainties reported by the fit and are distinct from the double-counting effects evaluated in the bootstrapping studies. The pseudo-experiment fit results are detailed in Section 8.6.

The data fit uses a set of fixed template shapes to model the signal and each background component. These templates are subject to per-bin statistical uncertainties, and uncertainties arising from the application of per-event weights to align data and simulation. These effects are treated as sources of systematic uncertainty on the observable results. All sources of systematic uncertainty are discussed and evaluated in Chapter 9.

8.1 SIGNAL MODEL USED IN SIMULATION

The set of signal templates required for the angular fit relies on a per-event weighting procedure, as described in Section 8.2 below. An important input to this procedure is the unbinned angular decay rate distribution of the signal at the generator-level i.e. the signal model generated in EvtGen [85]. This model, referred to as M in the following, is used in the assignment of per-event weights to the signal simulation, in order to project out the angular distributions for individual I_x components of the full decay rate. To derive the form of M, an unbinned fit to the true decay angle distribution is performed on a sample of 1 million generator-level $B \to D^*\tau\nu$ signal events. The sample is produced using RapidSim and is generated with no detector acceptance or resolution effects included. The ISGW2 model is used in the B-decay generation, which is identical to the model used in the generation of all fully simulated signal simulation samples used in the analysis [86]. It is essential that no detector acceptance or resolution effects are included, otherwise, a parametric fit would not be possible.

An unbinned 3D maximum-likelihood fit to the true decay angles is performed using zfit, where a decay rate PDF is defined according to Equation 3.19. In the fit, the normalisation condition $\Gamma = \frac{1}{4}\left(3I_{1c} + 6I_{1s} - I_{2c} - 2I_{2s}\right) = 1$ is implemented

Variable	Fitted Value
I_{1c}	$(5.326 \pm 0.011) \times 10^{-1}$
I_{1s}	$(3.9517 \pm 0.0059) \times 10^{-1}$
I_{2c}	$(-1.627 \pm 0.017) \times 10^{-1}$
I_{2s}	$(6.576 \pm 0.099) \times 10^{-2}$
I_3	$(-1.097 \pm 0.010) \times 10^{-1}$
I_4	$(-1.411 \pm 0.010) \times 10^{-1}$
I_5	$(2.8000 \pm 0.0093) \times 10^{-1}$
I_{6c}	$(3.239 \pm 0.018) \times 10^{-1}$
I_{6s}	$(-2.528 \pm 0.010) \times 10^{-1}$
I_7	$(0.1 \pm 1.1) \times 10^{-3}$
I_8	$(1.6 \pm 1.1) \times 10^{-3}$
I_9	$(-1.5 \pm 1.0) \times 10^{-3}$

Table 8.1: Values of the angular coefficients measured in an unbinned maximum-likelihood fit to the generator-level $B \to D^*\tau\nu$ angular decay rate. These values are used to define the generator-level model M, which was employed via the ISGW2 option in EvtGen.

in the fit via the relation $I_{1c} = \frac{1}{3}(4 - 6I_{1s} + I_{2c} + 2I_{2s})$, which fixes the value of I_{1c} relative to the other angular coefficients. All other angular coefficients vary freely in the fit. The total signal PDF is normalised to 1.0 and then scaled by the data yield. The one-dimensional fit projections are shown in Figure 8.1, and the measured values of the I_x are given in Table 8.1. All I_x are measured with high precision, such that the generator-level signal model M is known with negligible uncertainty. The values measured align well with published I_x SM predictions [29], with differences attributable to differences in the form-factor schemes. This analysis is agnostic to the specific choice of signal model M used in the simulation generation, as detailed in Section 8.2.

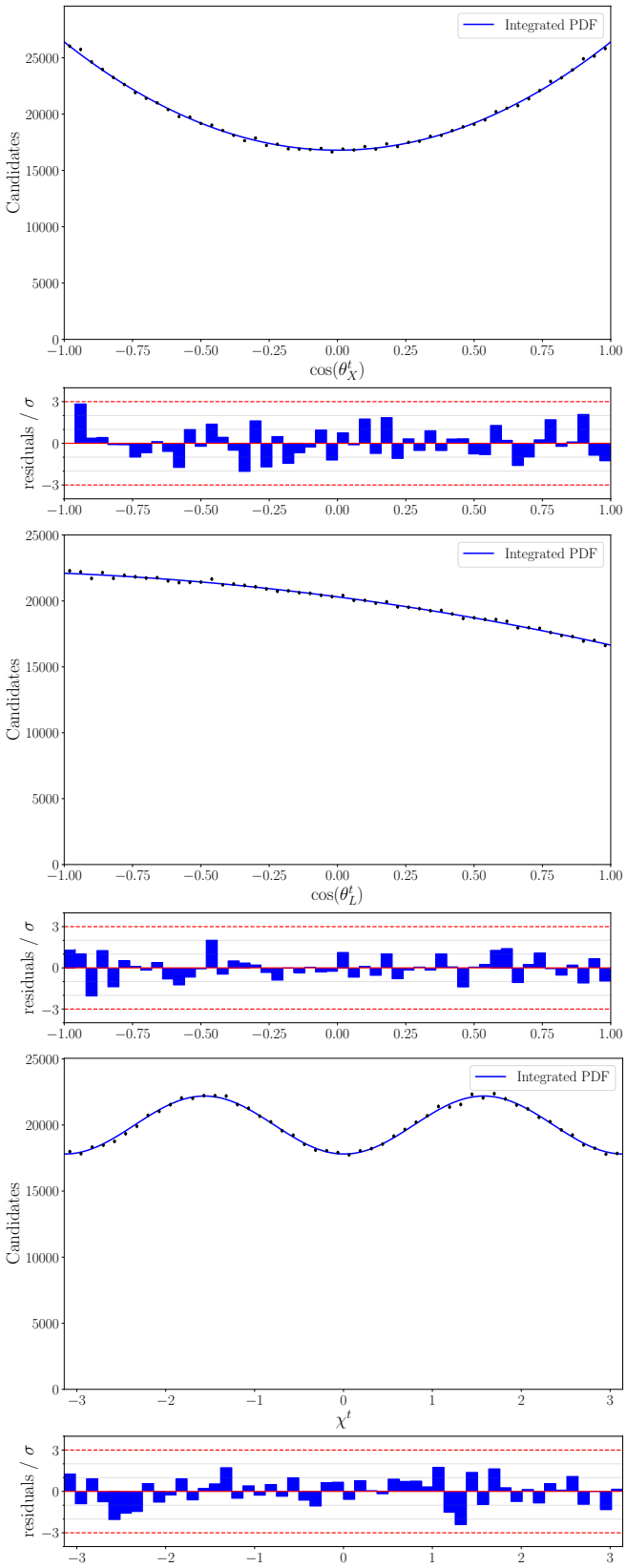


Figure 8.1: One-dimensional projections of the unbinned 3D maximum-likelihood to the generator-level signal angular distribution. This fit is used to measure the I_x coefficients used in the ISGW2 signal generation in EvtGen, which are used to define the generator-level model M.

8.2 SIGNAL TEMPLATE CREATION

The full angular decay rate as described by Equation 3.19 is a sum of twelve independent functions in $(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{true}}$ parameter space. Since there is missing energy at two vertices from the neutrinos, the signal decay is partially reconstructed. As a result, the reconstruction of the three angular variables is distorted by resolution and reconstruction effects as shown in Figures 4.3 and 4.4. Therefore, a template fit must be used in this analysis, contrasting with the parametric fits performed in $B^0 \to K^{*0}\mu^+\mu^-$ angular analyses for instance. This template fit is performed in $(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{reco}}$ space, and the twelve angular functions are transformed into twelve multidimensional templates each scaled by a corresponding I_x coefficient. The templates contain all of the reconstruction effects directly in the PDF, and allow the signal distribution in data to be modelled in a model-independent manner via a sum of the twelve templates each multiplied by a freely varying I_x parameter.

To help constrain the total signal yield, the three-dimensional angular fit is performed simultaneously with a 2D fit to the tau flight significance and the Isolation BDT distributions. The total signal template in these variables is also formed by a sum over twelve templates, each scaled by their corresponding I_x coefficients. The (tau flight significance, Isolation BDT) templates are projected out of the signal simulation in an identical manner to the corresponding angular templates, as described below. The fact that the total signal template is always formed by a sum of twelve freely varying sub-components ensures that no signal model assumption is made in any part of the fit.

The twelve component template histograms are created by firstly creating twelve continuous angular functions, each one corresponding to a single term in the full angular decay rate (e.g. $f_{I_{1s}} = \sin^2\theta_D$). These twelve functions are then divided by the total model, M, given in Equation 3.19, where M is calculated using the I_x values used in the signal simulation in EvtGen (see Section 8.1 for details on how the generator-level I_x are determined). The ratios $R_{I_x} = f_{I_x}/M$ are model-independent since the model used to produce the simulation is common to both the numerator and denominator.

The R_{I_x} functions are then used to assign twelve weights to each event in the fully simulated signal sample, with weights assigned according to the position of an event in $(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{true}}$ space. An event with position $(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{true}}$ will be assigned twelve weights $w_{I_x} = R_{I_x}(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{true}}$, where each weight corresponds to one of the terms in the full angular decay rate.

Each set of weights is then applied to the signal simulation when creating a histogram in reconstructed angular space, $(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{reco}}$, resulting in twelve independent angular templates h_{I_x}. The same weights are used to create a set of complementary templates in (tau flight significance, Isolation BDT). In addition to the w_{I_x} weights, per-event weights are also assigned to align data with simulation: details of these weights are given for each fit component in Section 8.3. The resulting set of signal templates includes all acceptance, selection, reconstruction, resolution, and event migration effects modelled in the simulation, as well as the additional adjustments applied to align data with the simulation. These effects are visible as the differences between $h_{I_{1s}}$ and $f_{I_{1s}}$ in Figure 8.2, where the template creation procedure for the I_{1s} term is illustrated.

8.2.1 CONSISTENCY OF SIGNAL TEMPLATES

A central feature of the template-based angular fit approach is the model independence of the signal templates. As described above, templates for each angular term in the fit are projected out of the total signal simulation sample using per-event weights. In the calculation of the per-event weights, the full model used in the generation of the simulation sample acts as a denominator. The net result should be a series of I_x templates which have no residual dependence on the model used to generate the signal simulation sample. This independence should include both q^2 and the decay angles, even though the reweighting is performed using only the 3D angular decay rate.

To demonstrate the model independence of the angular I_x templates, the template creation procedure is replicated in two statistically independent RapidSim samples generated with the ISGW2 and HQET2 models, [86, 87]. These models differ in their underlying calculations of the differential decay rate via differences in

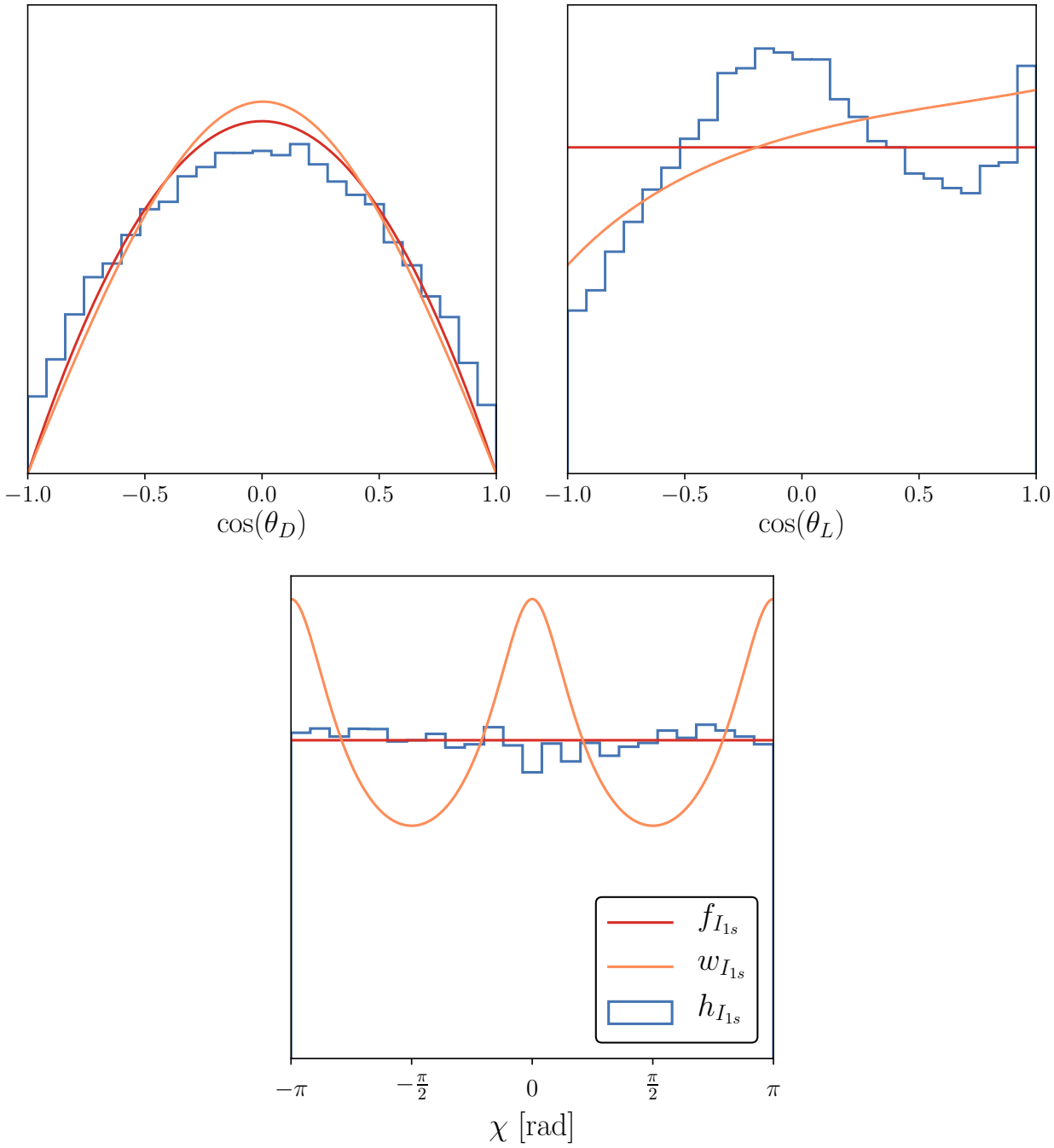


Figure 8.2: One-dimensional projections of the template creation procedure. The angular function $f_{I_{1s}}$ (red), is divided by the total model M, to obtain the weight function $w_{I_{1s}}$ (orange) which is calculated in the true angular space, $(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{true}}$. The final fit templates, $h_{I_{1s}}$ (blue), are obtained by applying the weights to the signal simulation in the reconstructed angular space $(\cos(\theta_D),\ \cos(\theta_L),\ \chi)_{\text{reco}}$. The resulting angular template $h_{I_{1s}}$ includes the acceptance, selection, reconstruction, resolution, and event migration effects, which appear as differences relative to the true $f_{I_{1s}}$ shape. As such, all reconstruction effects are directly embedded into the signal I_x component histograms, which can thus be applied directly to data.

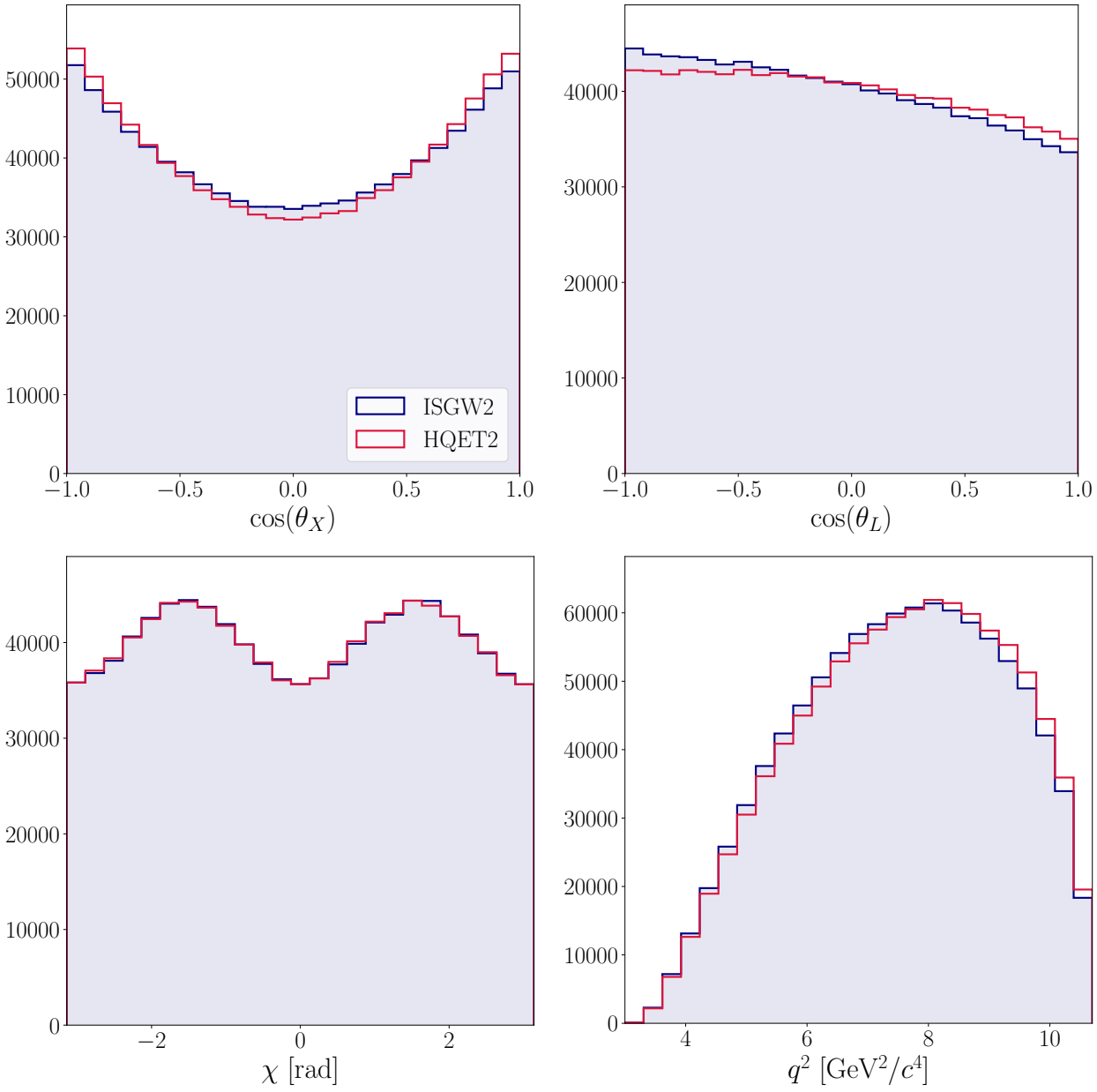


Figure 8.3: Comparison of the truth-level decay angle and q^2 signal distributions in RapidSim samples generated using ISGW2 and HQET2 form-factor models.

form-factor prescription, which leads to different angular and q^2 distributions. The truth-level angular and q^2 distributions of both samples are compared in Figure 8.3.

To determine the model parameters (q^2-integrated I_x coefficients) used in the ISGW2 and HQET2 models, unbinned fits to the 3D truth-level decay angle distributions are performed in an identical fashion to Section 8.1. The 1D projections of the fit to the ISGW2 sample are shown in Figure 8.4, and the fit to the HQET2 sample is shown in Figure 8.5. The I_x values measured in the ISGW2 and HQET2

Variable	Fitted Value ISGW2	Fitted Value HQET2
I_{1c}	$(5.326 \pm 0.011) \times 10^{-1}$	$(5.602 \pm 0.011) \times 10^{-1}$
I_{1s}	$(3.9517 \pm 0.0059) \times 10^{-1}$	$(3.7982 \pm 0.0058) \times 10^{-1}$
I_{2c}	$(-1.627 \pm 0.017) \times 10^{-1}$	$(-1.675 \pm 0.017) \times 10^{-1}$
I_{2s}	$(6.576 \pm 0.099) \times 10^{-2}$	$(6.346 \pm 0.098) \times 10^{-2}$
I_3	$(-1.097 \pm 0.010) \times 10^{-1}$	$(-1.080 \pm 0.010) \times 10^{-1}$
I_4	$(-1.411 \pm 0.010) \times 10^{-1}$	$(-1.416 \pm 0.010) \times 10^{-1}$
I_5	$(2.8000 \pm 0.0093) \times 10^{-1}$	$(2.8118 \pm 0.0093) \times 10^{-1}$
I_{6c}	$(3.239 \pm 0.018) \times 10^{-1}$	$(3.504 \pm 0.018) \times 10^{-1}$
I_{6s}	$(-2.528 \pm 0.010) \times 10^{-1}$	$(-2.351 \pm 0.010) \times 10^{-1}$
I_7	$(0.1 \pm 1.1) \times 10^{-3}$	$(-1.1 \pm 1.1) \times 10^{-3}$
I_8	$(1.6 \pm 1.1) \times 10^{-3}$	$(1.0 \pm 1.1) \times 10^{-3}$
I_9	$(-1.5 \pm 1.0) \times 10^{-3}$	$(0.3 \pm 1.0) \times 10^{-3}$

Table 8.2: True I_x values used in the ISGW2 and HQET2 signal sample generations.

fits are listed in Table 8.2. These I_x values are used to define the true models $M(\text{ISGW2})$ and $M(\text{HQET2})$ used in the generation of both samples.

Using the true models, per-event weights are assigned to both RapidSim samples following the procedure described earlier. A set of 12 I_x template histograms is then created from each sample, using the per-event weights to project out shapes for each angular term in reconstructed angular space. If signal model cancellation is complete during I_x template creation, the templates derived from the ISGW2 and HQET2 samples should be consistent with each other within statistical variations.

To check the template consistency across samples, the I_x templates derived from the ISGW2 and HQET2 samples are compared bin-by-bin. For a given I_x, pull values for each bin i are calculated using

$$p_i = [v(\text{ISGW2})_i - v(\text{HQET2})_i] / \sqrt{\epsilon(\text{ISGW2})_i^2 + \epsilon(\text{HQET2})_i^2}$$

where the terms v represent the bin contents of bin i, and the terms ϵ represent the bin uncertainties given by the square root of the sum of weights in the bin. The values of p_i are stored in a histogram, in order to check the distribution of pull values across all template bins. If the ISGW2 and HQET2 templates for a given I_x are fully consistent within statistical fluctuations, the pull distribution should be close to a unit Gaussian.

The pull distributions for each I_x are shown in Figures 8.6, 8.7, and 8.8, where a unit Gaussian is also shown in pink for comparison. The pull distributions observed

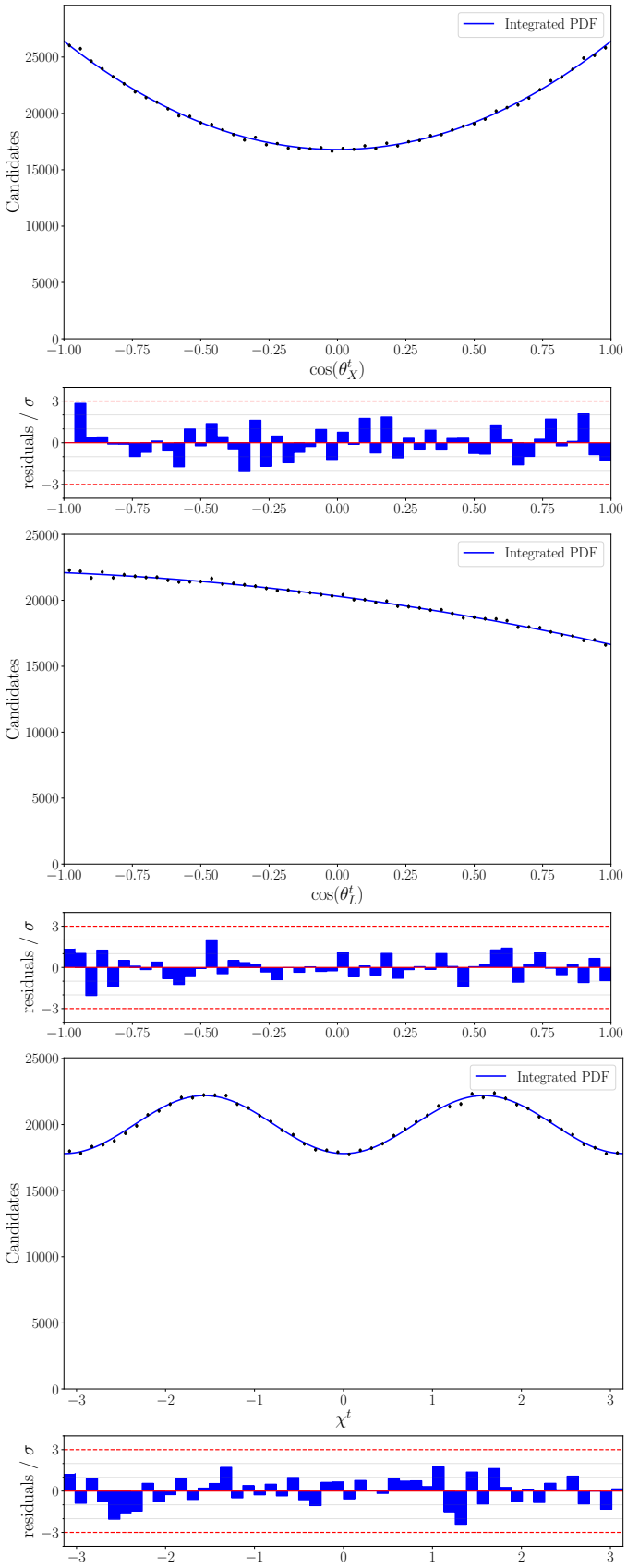


Figure 8.4: One-dimensional projections of the unbinned 3D maximum-likelihood to the ISGW2 truth-level signal angular distribution.

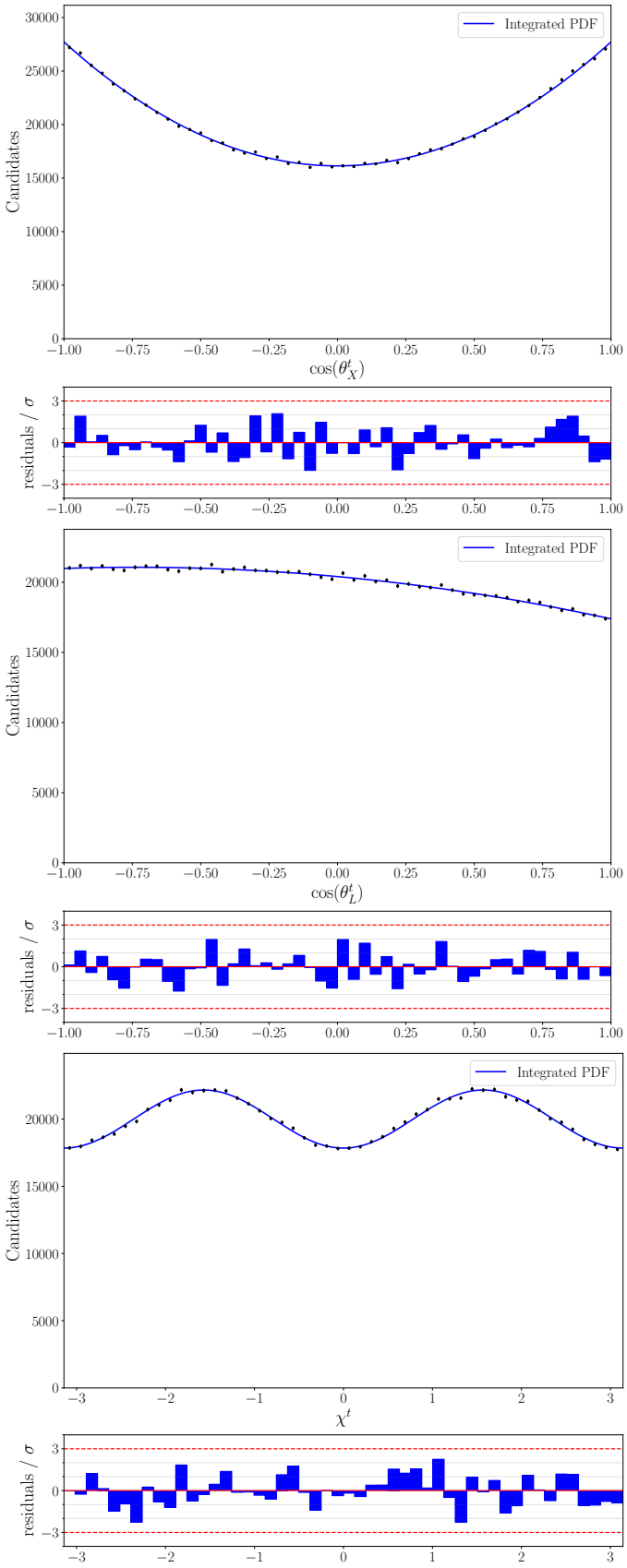


Figure 8.5: One-dimensional projections of the unbinned 3D maximum-likelihood to the HQET2 truth-level signal angular distribution.

are highly consistent with the unit Gaussian, with a mean close to zero and a width close to unity. As such, the templates derived from the ISGW2 and HQET2 samples are fully consistent within their statistical fluctuations. No residual model information is thus present in the templates, providing confidence that the I_x template histogram method indeed operates with a set of model-independent templates.

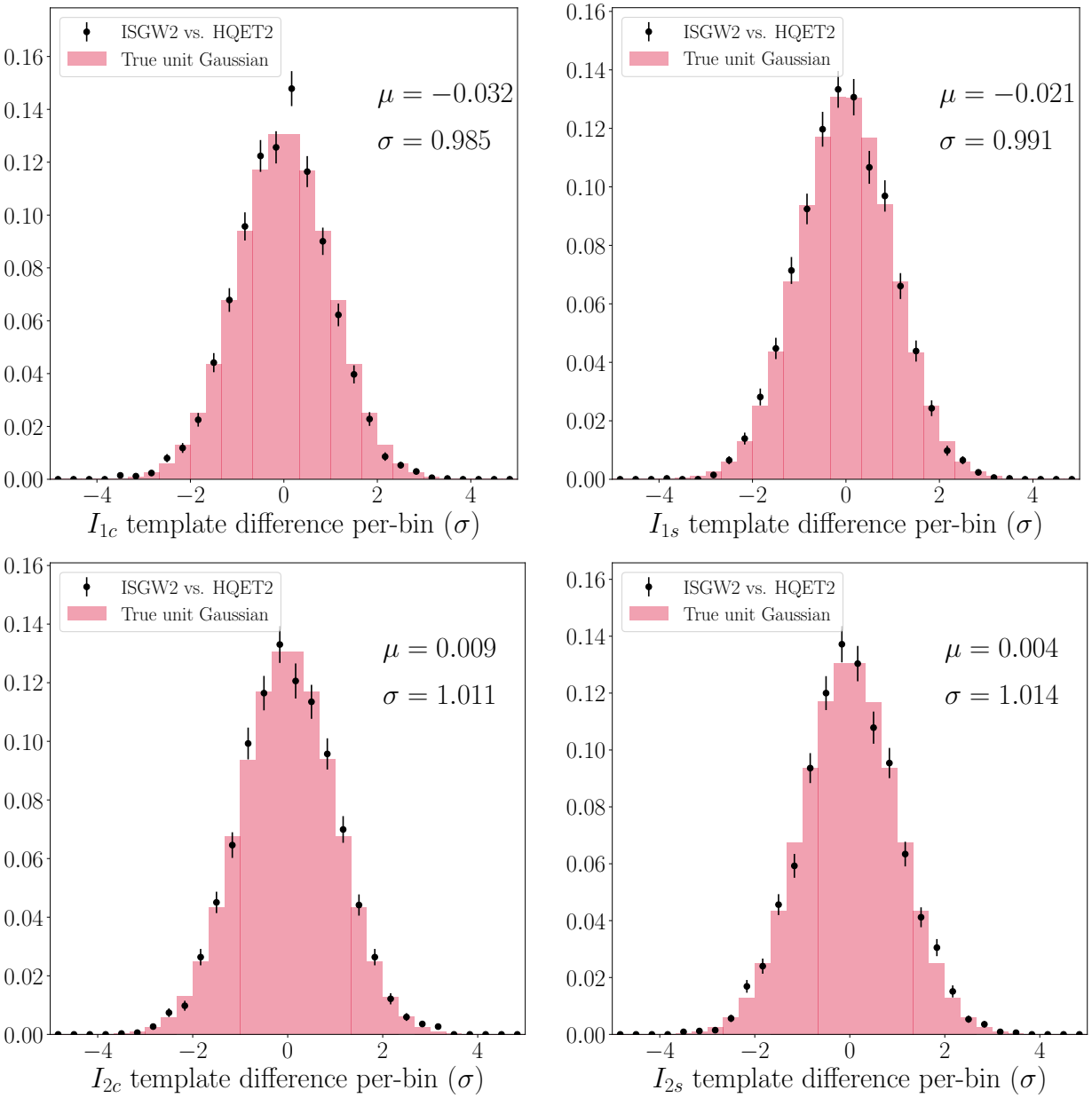


Figure 8.6: Normalised residual distributions for the comparison of I_x histograms generated from independent ISGW2 and HQET2 signal simulation samples. Both sets of templates are found to be fully consistent within their statistical uncertainties, demonstrating that the I_x template histogram method operates with a set of model-independent signal templates. The pull distribution for each I_x are found to be identical, which is expected given that all I_x templates are generated from the same ISGW2 and HQET2 samples.

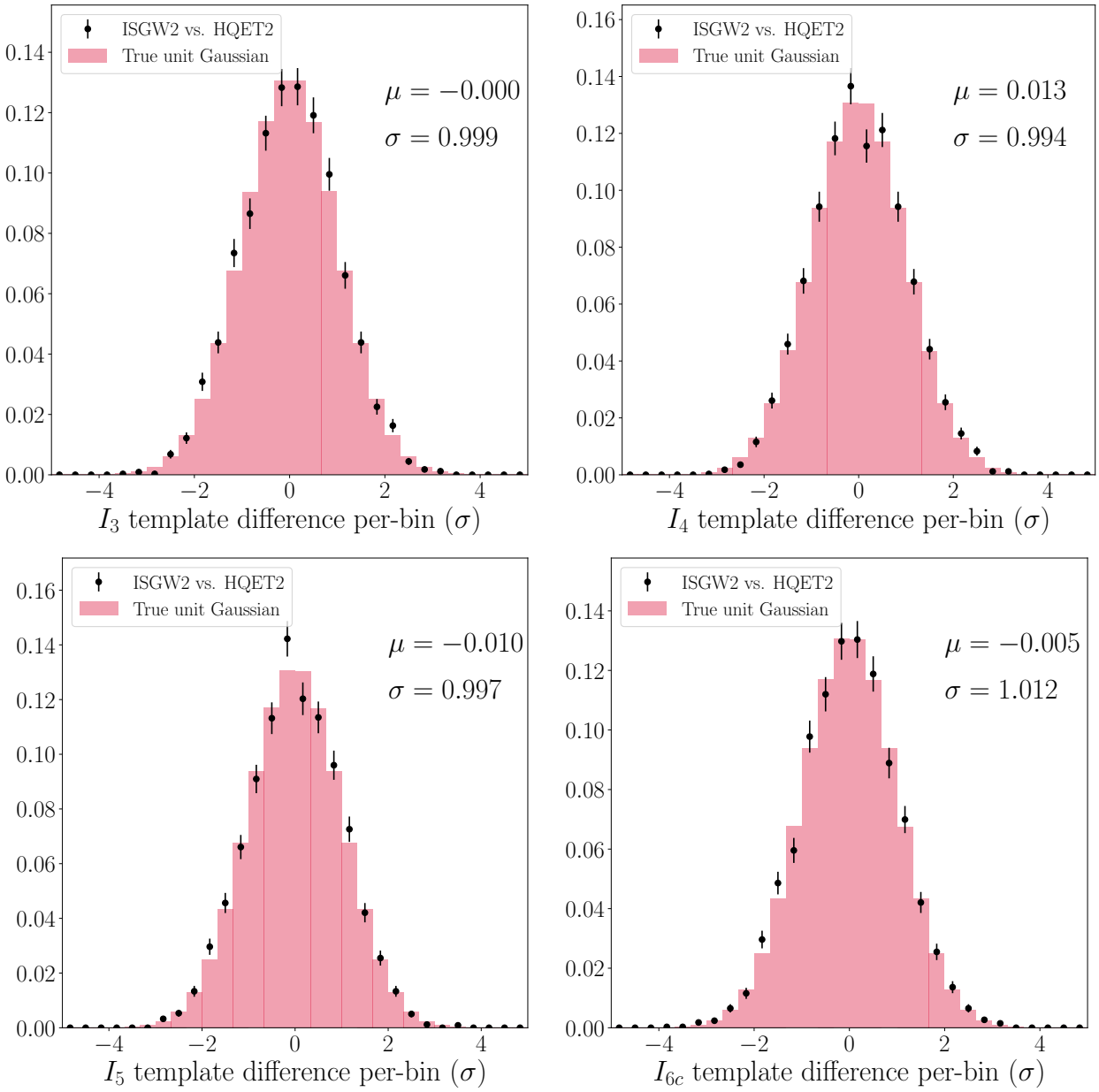


Figure 8.7: Normalised residual distributions for the comparison of I_x histograms generated from independent ISGW2 and HQET2 signal simulation samples. Both sets of templates are found to be fully consistent within their statistical uncertainties, demonstrating that the I_x template histogram method operates with a set of model-independent signal templates. The pull distribution for each I_x are found to be identical, which is expected given that all I_x templates are generated from the same ISGW2 and HQET2 samples (continued).

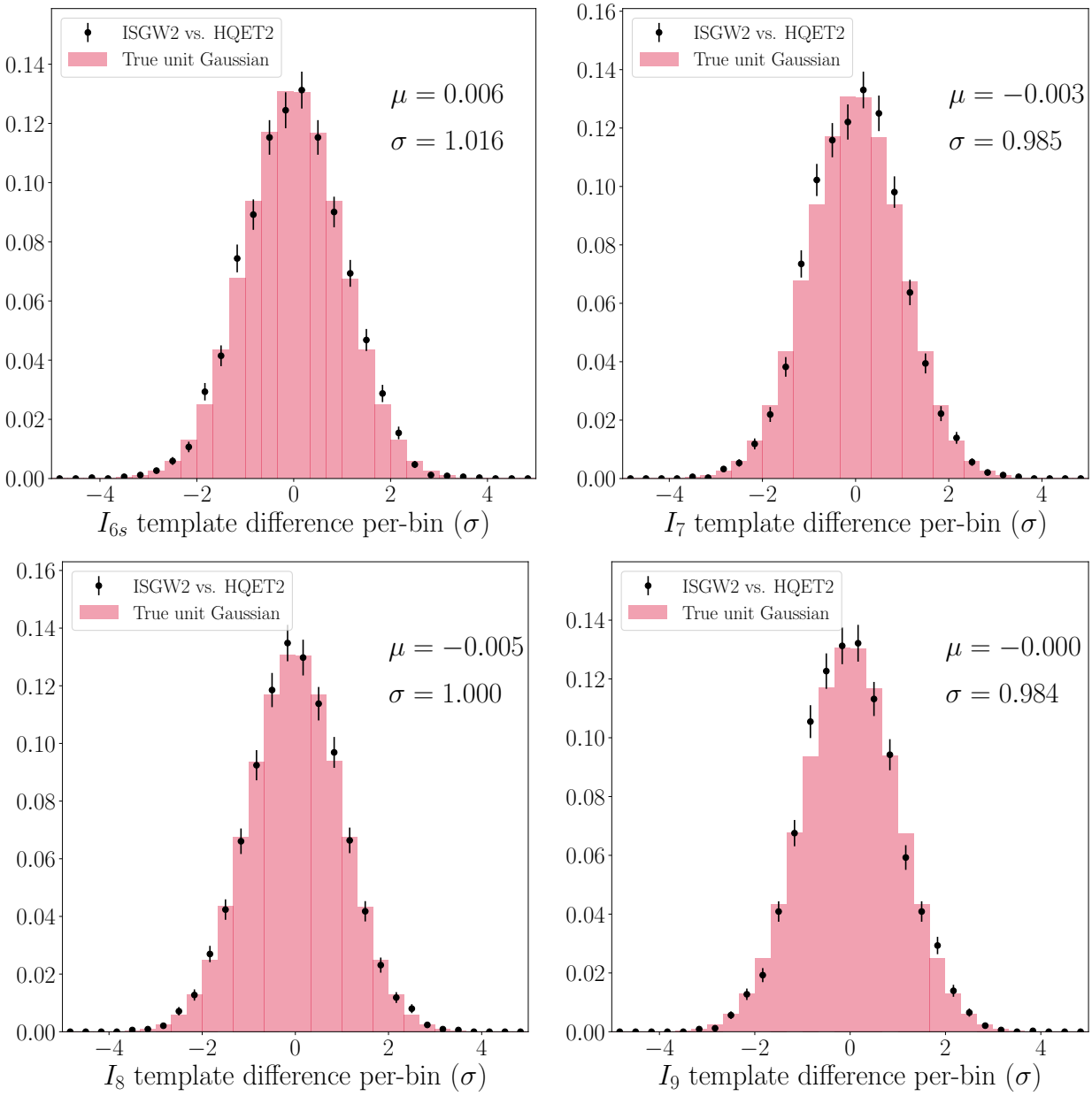


Figure 8.8: Normalised residual distributions for the comparison of I_x histograms generated from independent ISGW2 and HQET2 signal simulation samples. Both sets of templates are found to be fully consistent within their statistical uncertainties, demonstrating that the I_x template histogram method operates with a set of model-independent signal templates. The pull distribution for each I_x are found to be identical, which is expected given that all I_x templates are generated from the same ISGW2 and HQET2 samples (continued).

8.2.2 IMPACT OF SELECTION ON TEMPLATE MODEL-INDEPENDENCE

Another potential source of model dependence in the signal templates is the application of selection requirements to simulation (e.g. trigger requirements, stripping, BDTs), which are applied in order to align with the requirements applied to data. Selection requirements can in general warp the distributions used in the final angular fit, and could potentially cause different warping in signal samples generated from different models. To test for a potential bias from this source, a reconstructed $q^2 > 6$ requirement is applied to both the ISGW2 and HQET2 RapidSim samples prior to template creation. Due to underlying differences in q^2 distribution as shown in Figure 8.9, this requirement does not have the same absolute efficiency when applied to both samples and acts to remove a considerable proportion of low q^2 events.

However, the I_x templates projected out of both the ISGW2 and HQET2 samples remain statistically consistent after application of the q^2 requirement, as shown in Figure 8.11 for I_{1s} (all other I_x show identical behaviour). Several factors should be considered in order to understand this result:

- The underlying true q^2 distributions are encoded in the true models M, which are used to assign per-event weights during template creation. The true models act as denominators in the w_{I_x} weights and are unchanged by the reconstructed q^2 requirement. The models M are always defined purely at the generator level, with no requirements applied.

- The I_x templates are created from simulation samples surviving the reconstructed $q^2 > 6$ requirement. This requirement is not used in the signal selection and was chosen to introduce a significant q^2 acceptance. The proportions of events surviving this requirement are different in the ISGW2 and HQET2 samples, due to their different true q^2 distributions. However, the differences in the surviving true q^2 distribution are cancelled when the per-event weights w_{I_x} are applied since the weights encode the true q^2 distributions within the true models M.

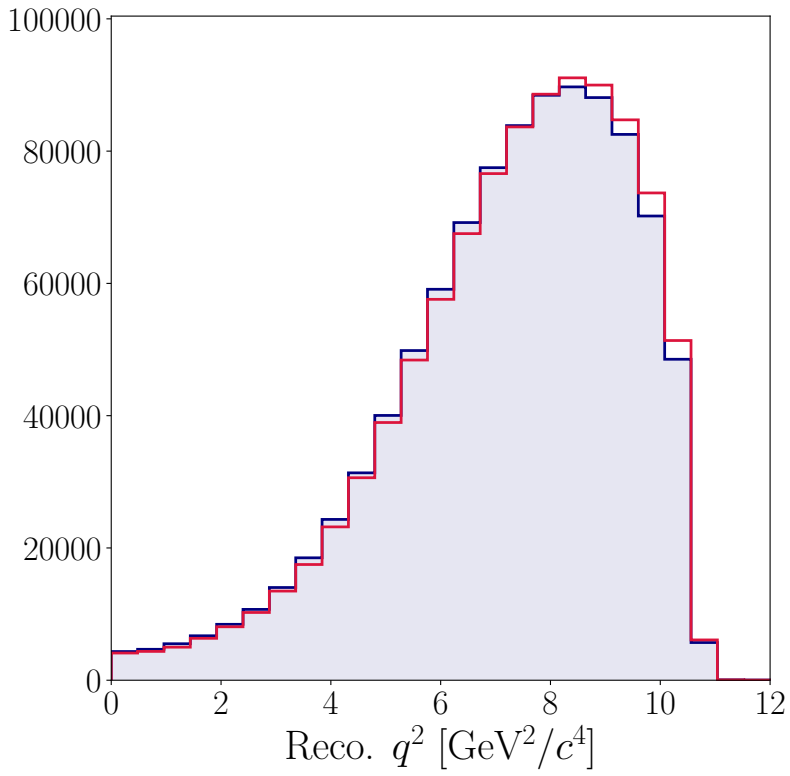


Figure 8.9: Comparison of the reconstructed q^2 distribution in ISGW2 and HQET2 RapidSim signal samples. A requirement of reconstructed $q^2 > 6$ is applied in order to test the model independence of templates in the presence of selection requirements.

- The net result is that both the ISGW2 and HQET2 template shapes are changed by the reconstructed q^2 requirement, but they are changed in the same way, leaving them consistent with each other. Such a change in template shape due to selection requirements is the desired behaviour since simulation templates used in the angular fit should reflect the total warping introduced by reconstruction and selection effect present in data. However, the shape changes are found not to depend on the underlying model used to generate the simulation.

In Figure 8.10, an illustration of the template consistency across models is shown, as well as a demonstration of the consistency across models when selection requirements are applied. One-dimensional projections of the following template histograms are displayed for the I_{1c} term:

- ISGW2 template with no requirement on reconstructed q^2 (red)

- HQET2 template with no requirement on reconstructed q^2 (blue)
- ISGW2 template with a reconstructed $q^2 > 6$ requirement applied (green)
- HQET2 template with a reconstructed $q^2 > 6$ requirement applied (orange)

The template shapes for ISGW2 (red) and HQET2 (blue) with no q^2 requirement applied are found to be consistent within their statistical fluctuations, in agreement with the results shown in Figures 8.6, 8.7, and 8.8. In addition, the templates for ISGW2 (green) and HQET2 (orange) with a reconstructed $q^2 > 6$ requirement applied are found to be consistent with each other. The template shapes are clearly altered by the application of the reconstructed q^2 requirement: importantly, the q^2 requirement alters the template shapes in the same way for the ISGW2 and HQET2 signal samples, such that model-independence of the I_x templates is maintained.

The consistency of template shapes across signal models, even in the presence of selection requirements, demonstrates the robustness of the I_x template method as a means of describing the signal distribution in the angular fit. An important consequence of the above test is that the I_x observables measured in the angular fit do not need to be unfolded for q^2 acceptance effects post-fit. The q^2 acceptance, along with the acceptance due to all other selection requirements, is encoded in the final shape of the I_x signal template shapes. When these templates are used to fit data where the same acceptance effects are present, they provide an unbiased measurement of the q^2-integrated I_x coefficient values. Given this, it is not envisioned to publish a q^2 acceptance for the analysis, since theorists can use the measured I_x values directly in any comparison with theory.

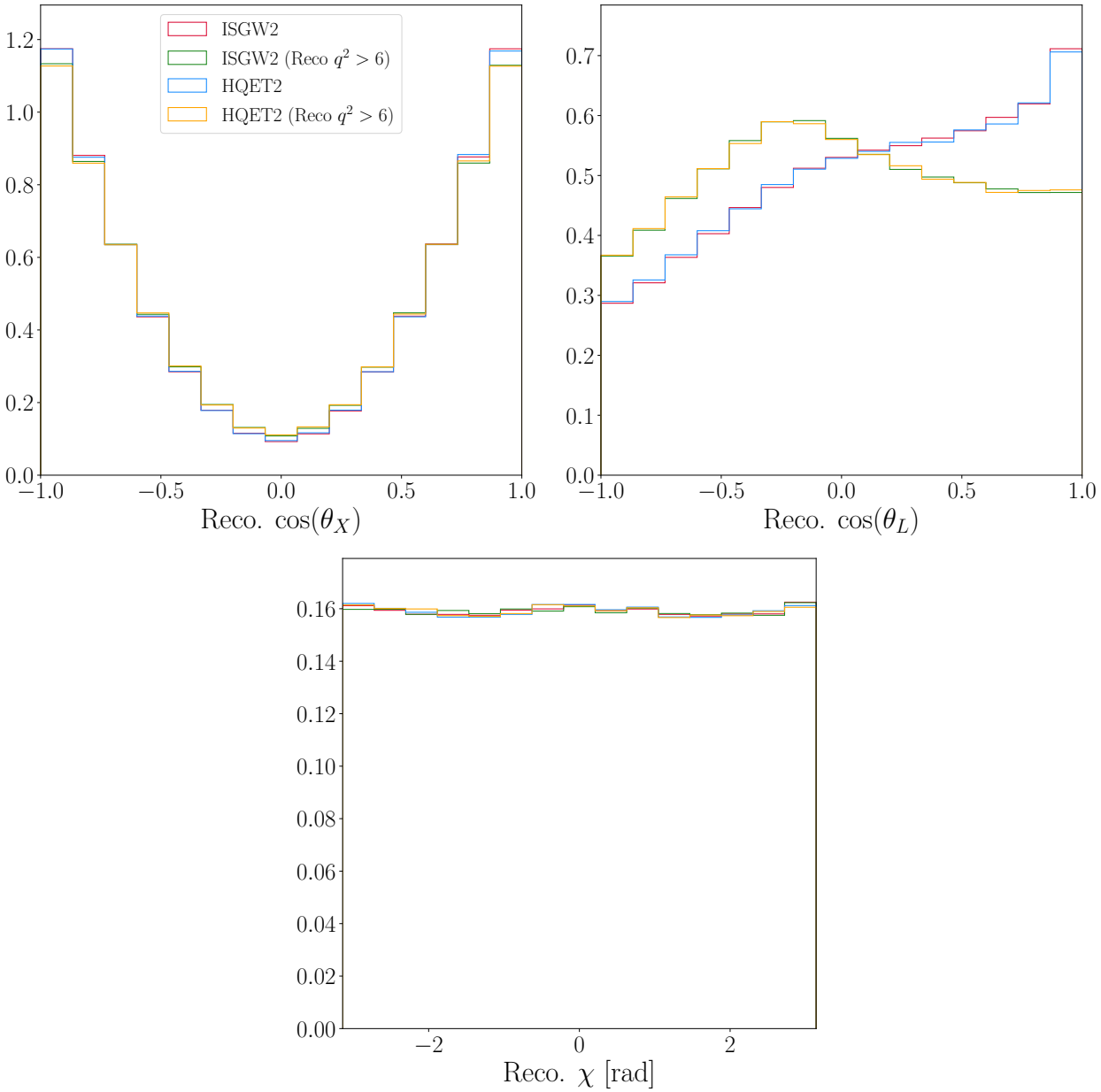


Figure 8.10: Comparison of I_{1c} template shapes derived from the ISGW2 and HQET2 signal models, both with and without a requirement on reconstructed q^2 applied. The template shapes derived from both signal samples are consistent within their statistical uncertainties, and selection requirements on both samples are found to influence the template shapes in the same way. Together these observations provide confidence that the I_x template method delivers a set of model-independent signal templates which can be used in the angular fit to data to measure unbiased I_x coefficient values.

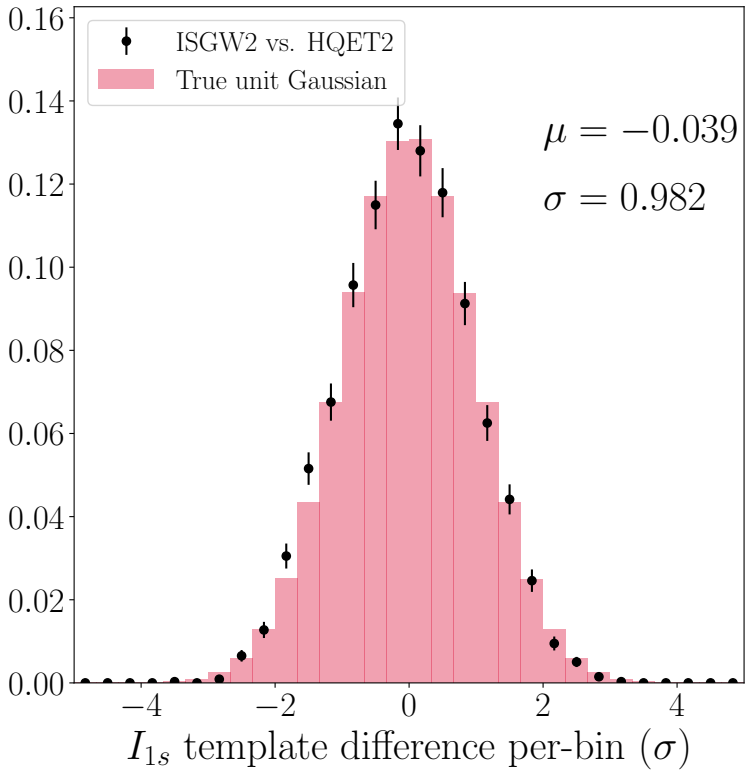


Figure 8.11: Pull distribution for the comparison of I_{1s} histograms generated from independent ISGW2 and HQET2 signal simulation samples, where a reconstructed $q^2 > 6$ requirement is applied to both samples before template creation. Both sets of templates are found to be fully consistent within their statistical uncertainties, demonstrating that the templates derived from simulation remain model-independent even in the presence of selection requirements. The distributions for all other I_x templates are also consistent with unit Gaussians.

8.3 ANGULAR FIT COMPONENTS

Below, a description of each angular fit component is provided, followed by a definition of the total fit PDF. For all simulated samples, a simulation-truth requirement is applied to the D^* meson to ensure that only true D^* candidates are modelled in the simulation. This is required in order to align with the data, which has ${}_s$Weights applied to remove the fake D^* and fake D^0 contributions.

8.3.1 $B \to D^*\tau\nu$ SIGNAL

The angular fit to data is performed using a sum of twelve signal templates, each scaled by an I_x coefficient, where the sum of all twelve templates is itself scaled by a freely varying signal fraction f_{sig}. The normalisation condi-

tion $\Gamma = \frac{1}{4}(3I_{1c} + 6I_{1s} - I_{2c} - 2I_{2s}) = 1$ is incorporated by requiring $I_{1c} = \frac{1}{3}(4 - 6I_{1s} + I_{2c} + 2I_{2s})$, which fixes the value of I_{1c} relative to the other coefficients. The total signal PDF is then normalised to an integral of 1, then scaled by the signal fraction f_{sig}.

A combination of $\tau \to 3\pi\nu$ and $\tau \to 3\pi\pi^0\nu$ decays within the total signal template is modelled via a fixed parameter $f_{3\pi} = (87.79 \pm 0.14)\%$, which is determined using the full selection efficiencies for each mode and their known branching ratios. The value of $f_{3\pi}$ contributes negligibly as a systematic uncertainty on the total template shapes given its small uncertainty.

In addition to the w_{I_x} weights assigned to the signal simulation to form I_x sub-templates, per-event PID efficiency weights, L0 weights, $3\pi X$ weights, and year weights are also assigned. The per-event PID efficiencies are assigned according to PIDCalib, using reference histograms defined in track momentum and pseudorapidity. The total efficiency for each event is given by the product of all three τ pion PID efficiencies. Every event is assigned an absolute PID-related systematic uncertainty of 0.6%, corresponding to a fully correlated uncertainty of 0.2% per track. This uncertainty covers the use of finite PID calibration samples, the use of ${}_s$Weights in the calibration samples, and the use of a binned procedure for efficiency assignment. The systematic uncertainty is propagated through the fit by regenerating the templates many times under variation of the per-event PID efficiencies, and running the fit to data many times (see Chapter 9).

The L0 weights are assigned to align the L0 composition of the simulation with data, as detailed in Section 7.4. The $3\pi X$ weights are assigned to align the 3π kinematics and flight significance in simulation and data, as described in Section 7.5. The year weights assigned act to balance the per-year composition of the signal sample, ensuring that the relative contribution from each year is aligned with the data. This is achieved using the measured cross-sections and luminosities for each data-taking year, and the relative generated signal sample sizes for each year. The weights are calculated such that the simulation proportions are reweighted to the expected proportions in data. If there has been too much simulation generated for a particular year in proportion to the other data-taking years, the year weights for

that sample will be less than 1 in order to reduce its relative proportion in the total simulation dataset.

In summary, the signal PDF is parameterised by the following:

- Freely varying signal fraction f_{sig}.
- Eleven freely varying angular coefficients I_x, with the twelfth coefficient I_{1c} fixed according to the normalisation condition.
- Fixed relative amount of $\tau \to 3\pi\nu$ and $\tau \to 3\pi\pi^0\nu$ decays via the parameter $f_{3\pi}$.
- Per-event PID, L0, $3\pi X$, and year-balancing weights applied during template creation.

8.3.2 $B \to D^{**}\tau\nu$ FEED-DOWN BACKGROUND

An important class of background to consider in semileptonic $R(X)$ measurements are the so-called "feed-down" contributions from higher charm resonances. In this analysis, it is possible for decays modes of the general type $B \to (D^{**} \to D^*X)\tau\nu$ to contribute, where one or more low-momentum particles X from various D^{**} decays are not reconstructed. It is possible to estimate the total contribution from such backgrounds relative to signal using a combination of:

- SM predictions for $R(D^{**}) = \mathcal{B}(B \to D^{**}\tau\nu)/\mathcal{B}(B \to D^{**}\ell\nu)$
- SM prediction for $R(D^*)$
- Measured value of $\mathcal{B}(B \to D^*\ell\nu)$
- Measurements of various $B \to D^{**}\ell\nu$ branching fractions
- Total selection efficiencies for feed-down and signal modes, taken from simulation

The goal of this study is to arrive at a value for $f_{D^{**}\tau\nu}$, which will be used in the angular fit to constrain the fraction of the feed-down background relative to the

signal fraction. For instance, $f_{D^{**}\tau\nu} = 0.1$ and $f_{\text{sig}} = 0.2$ would correspond to a total feed-down fraction of $0.1 \times 0.2 = 0.02$ in the angular fit.

The values of $R(D^{**})$ used are given in Table 8.3, and are taken from Reference [88]. Estimates of $\mathcal{B}(B \to D^{**}\tau\nu)$ branching fractions for each excited charm state can be found by multiplying these predictions by corresponding $\mathcal{B}(B \to D^{**}\ell\nu)$ values. The values used are summarised in Table 8.4, where all values are quoted inclusive of the $D^{**} \to D^{*}X$ process. Values for $\mathcal{B}(B^{+/0} \to D_0^*\ell\nu)$ are not available and are assumed to take the value $(0.50 \pm 0.25) \times 10^{-3}$ as done in Reference [88].

All B^0 modes are additionally scaled by a factor of 0.5 to account for isospin; this is necessary since a charged D^* is required in the final state. The only branching fraction value available for B_s^0 decays in the PDG is $\mathcal{B}(B_s^0 \to D_1\ell\nu) = (2.7 \pm 0.7) \times 10^{-3}$; values for the other excited charm states are inferred using this value and ratios of the equivalent B^0 decays. All B_s^0 values are additionally scaled by the LHCb measurement of f_s/f_d to account for relative B_s^0 production [89]. An average value across each data-taking year is calculated by weighting the published f_s/f_d values for each collision energy by the corresponding luminosity and cross-section values: an average value of $f_s/f_d = (25.06 \pm 0.35)\%$ is found.

A total expected SM feed-down rate $\mathcal{B}(B \to D^{**}\tau\nu)$ is calculated by summing over all of the modes considered above, where a value of $(0.91 \pm 0.16) \times 10^{-3}$ is determined. By comparison, the total rate of $B^0 \to D^*\tau\nu$ signal expected in the SM is determined to be $(13.1 \pm 0.4) \times 10^{-3}$, using $R(D^*) = 0.258 \pm 0.005$ from HFLAV and $\mathcal{B}(B^0 \to D^*\ell\nu) = (5.06 \pm 0.12)\%$ from the PDG. The relative total rate of feed-down decays compared to the signal is thus $(7.0 \pm 1.2)\%$. However, this value must be scaled by the relative efficiency of feed-down decays passing the full analysis selection relative to signal decays. The relative efficiency, taken from simulation, is found to be 0.27 ± 0.01, where a sum over separate feed-down simulation samples is performed weighting each mode by its expected fractional contribution (using the above branching fraction estimates).

Finally, an estimate of $f_{D^{**}\tau\nu} = 0.07 \times 0.27 = (1.89 \pm 0.33)\%$ is determined, where the uncertainty quoted is inclusive of the uncertainties on all input quantities used in the above calculations. The expected feed-down rate is thus small relative to the

	Value
$R(D_1^{'})$	0.05 ± 0.02
$R(D_1)$	0.10 ± 0.02
$R(D_2^*)$	0.07 ± 0.01
$R(D_0^*)$	0.05 ± 0.02

Table 8.3: SM $R(D^{**})$ predictions for several excited charm states (from [88]).

Mode	$\mathcal{B}(\times 10^{-3})$	Source
$B^+ \to D_1^{'}\ell\nu$	2.7 ± 0.6	PDG
$B^+ \to D_1\ell\nu$	3.03 ± 0.20	PDG
$B^+ \to D_2^*\ell\nu$	1.01 ± 0.24	PDG
$B^+ \to D_0^*\ell\nu$	0.50 ± 0.25	Assumed
$B^0 \to D_1^{'}\ell\nu$	(2.80 ± 0.28) $\times 0.5$	PDG
$B^0 \to D_1\ell\nu$	(3.1 ± 0.9) $\times 0.5$	PDG
$B^0 \to D_2^*\ell\nu$	(0.68 ± 0.12) $\times 0.5$	PDG
$B^0 \to D_0^*\ell\nu$	(0.50 ± 0.25) $\times 0.5$	Assumed
$B_s^0 \to D_1\ell\nu$	2.7 ± 0.7	PDG
$B_s^0 \to D_1^{'}\ell\nu$	2.4 ± 0.1	Assumed from D_1 mode and B^0 ratio
$B_s^0 \to D_2^*\ell\nu$	0.59 ± 0.25	Assumed from D_1 mode and B^0 ratio
$B_s^0 \to D_0^*\ell\nu$	0.44 ± 0.28	Assumed from D_1 mode and B^0 ratio

Table 8.4: $B \to D^{**}\ell\nu$ branching fractions used in the determination of $B \to D^{**}\tau\nu$ feed-down background rates. All values quoted are inclusive of the $D^{**} \to D^*X$ process.

signal, and notably smaller than estimates made in previous $R(D^*)$ measurements. The value of $f_{D^{**}\tau\nu}$ is used in a fit as a Gaussian constraint on the total contribution from feed-down decays.

8.3.3 $B \to D^*D_sX$ BACKGROUND

The largest single component of the angular fit arises from $B \to D^*(D_s \to 3\pi X)X$ decays. A dedicated set of $D_s \to 3\pi X$ correction weights are derived in Section 7.1 to align the simulation $D_s \to 3\pi X$ model with data. In addition, the individual B-decay contributions in the full $B \to D^*D_sX$ sample are aligned with data according to the control studies detailed in Section 7.2. In the angular fit, templates are created for each contributing B-decay mode within the total $B \to D^*D_sX$ background, and their relative fractions are Gaussian constrained according to the values measured in data (see Table 7.4). In creating the templates, weights are assigned per $D_s \to 3\pi X$ decay mode to align the $D_s \to 3\pi X$ model with data

as per Section 7.1. Per-event weights are also assigned to each event using the correction splines described in Section 7.2.3, in order to align simulation and data. Finally, PID efficiency, L0, $3\pi X$, and year-balancing weights are assigned in the same manner as described above for the signal template.

The total $B \to D^* D_s X$ template created by the above procedure is scaled by a freely varying fraction $f_{D^* D_s X}$ in the angular fit. The $D_s \to 3\pi X$ weights assigned are varied as a source of systematic uncertainty, as are the spline correction, PID efficiency, L0, and $3\pi X$ weights (see Chapter 9).

8.3.4 $B \to D^* D^+ X$ BACKGROUND

The $B \to D^* D^+ X$ template is taken from simulation, with a freely varying fraction $f_{D^* D^+ X}$ used to scale the contribution from this background in the fit. During template creation, per-event weights are assigned according to the splines detailed in Section 7.3, in order to align the simulation with data. Per-event PID efficiency, L0, $3\pi X$, and year-balancing weights are also assigned following the same procedure as described above. A correction weight of 1/15 is also assigned to the $D^+ \to K_s^0 3\pi$ sub-component, since the generated fraction of this mode in the simulation is 15 times larger than expected. The spline, PID, L0, and $3\pi X$ weights are assigned as a source of systematic uncertainty in Chapter 9.

8.3.5 $B \to D^* D^0 X$ BACKGROUND

The $B \to D^* D^0 X$ template is taken from simulation, with a freely varying fraction $f_{D^* D^0 X}$ used to scale the contribution from this background in the fit. During template creation, per-event weights are assigned according to the splines detailed in Section 7.3, in order to align the simulation with data. Per-event PID efficiency, L0, $3\pi X$, and year-balancing weights are also assigned following the same procedure as described above. The spline, PID, L0, and $3\pi X$ weights are assigned as a source of systematic uncertainty in Chapter 9.

8.3.6 $B \to D^* 3\pi X$ PROMPT BACKGROUND

Even after the tau flight significance > 3 requirement is applied, a residual contribution from the prompt background ($3\pi X$ produced at the B-decay vertex) is present in the data. This contribution is modelled using the inclusive $b\bar{b} \to D^* 3\pi X$ simulated sample, where simulation-truth requirements are applied to the 3π system to require that it does not originate from a flying parent such as a charm meson. Per event PID efficiency, L0, $3\pi X$, and year-balancing weights are assigned during template creation, and the uncertainties on these weights are considered as sources of systematic uncertainty in Chapter 9.

8.3.7 COMBINATORIAL BACKGROUND

Fake D^* and fake D^0 combinatorial background is statistically subtracted from the angular fit dataset via ${}_s$Weights. However, other sources of combinatorial background remain in data, namely from random 3π combinations and combinations of D^* and 3π candidates originating from different B-hadron decays. These contributions are modelled in the data fit using a sample of wrong-sign $D^{*+}3\pi^+$ decays taken from data. A 2D fit to $m(D^* - D^0)$ and $m(D^0)$ is performed on this data in an identical manner to the angular fit dataset, in order to isolate a sample of true D^* decays. Templates for the angular fit are then formed using this sample, with per-event ${}_s$Weights applied. The total contribution from such decays is small in the angular fit, and is allowed to vary freely under the condition that all component fractions must sum to unity. The angular projections of each background in the signal fit can be seen in Figure 8.12.

8.3.8 TOTAL FIT PDF

The total angular fit PDF is given by a linear sum of the signal and background components, with fractions assigned to each component template as described above. All fit parameters are shared across the two simultaneous fit categories, and the templates formed for each fit category are created in an identical fashion modulo

the variables used. The total fit PDF template is defined as follows:

$$\begin{aligned}\mathcal{P}(V) = & f_{\text{sig}} \times \Sigma_x I_x \times (f_{3\pi} \times h_{I_x}^{3\pi}(V) + (1 - f_{3\pi}) \times h_{I_x}^{3\pi\pi^0}(V)) \\ & + f_{\text{sig}} \times f_{D^{**}} \times h_{D^{**}\tau\nu}(V) \\ & + f_{D^*D_sX} \times \Sigma_B f_B \times h_B(V) \\ & + f_{D^*D^+X} \times h_{D^*D^+X}(V) \\ & + f_{D^*D^0X} \times h_{D^*D^0X}(V) \\ & + f_{D^*3\pi X} \times h_{D^*3\pi X}(V) \\ & + (1 - f_{\text{sig}} - f_{\text{sig}} \times f_{D^{**}} - f_{D^*D_sX} - f_{D^*D^+X} - f_{D^*D^0X} - f_{D^*3\pi X}) \times h_{\text{comb.}}(V)\end{aligned}$$

where V refers to the variables chosen for the binning, namely the reconstructed decay angles for the first fit category and the reconstructed tau flight significance and Isolation BDT for the second category. In all cases, $h(V)$ refers to a normalised template histogram defined in the variables V and are fixed template shapes. The sum Σ_x refers to the sum over the twelve signal I_x coefficients, while the sum Σ_B refers to the sum over the component B decays of the $B \to D^* D_s X$ background. The total PDF $\mathcal{P}(V)$ is multiplied by N_{data}, the sum of ${}_s$Weights in the angular fit dataset, in order for each bin of the fit PDF to be represented as a yield. This allows a per-bin comparison of the fitted yield with the observed yield in data, which forms the basis of the binned maximum-likelihood fit performed. In order to perform the simultaneous fit across both sets of variables, the log-likelihoods for each category are summed and passed to a Minuit minimisation routine in zfit. The minimiser adjusts the free parameters of the fit until a minimum in the summed log-likelihood is found, which corresponds to a maximum in the likelihood. The negative log-likelihoods used in the fit are defined as:

$$\text{NLL}(V) = \Sigma_i (N_i^{\mathcal{P}} - N_i^{\text{data}} + N_i^{\text{data}} \times \log(N_i^{\text{data}} / N_i^{\mathcal{P}}))$$

where V is a given set of variables, and the index i runs over the individual bins of the fit. The PDF yield in bin i is $N_i^{\mathcal{P}}$ and the data yield in bin i, calculated as the sum of ${}_s$Weights in that bin, is N_i^{data}. A summary of the PDF parameters is provided in Table 8.5, including information on which parameters are freely varied, constrained, or fixed.

Table 8.5: Summary of the angular fit PDF parameters.

Parameter	Component	Information
f_{sig}	Signal	Freely varies
I_x	Signal	11 freely vary with I_{1c} fixed by normalisation
$f_{3\pi}$	Signal	Fixed to $(87.79 \pm 0.14)\%$
$f_{D^{**}}$	Feed-down	Gaussian constrained to $(1.89 \pm 0.33)\%$
$f_{D^*D_sX}$	$B \to D^* D_s X$	Freely varies
f_B	$B \to D^* D_s X$	All Gaussian constrained to Table 7.4 values
$f_{D^*D^+X}$	$B \to D^* D^+ X$	Freely varies
$f_{D^*D^0X}$	$B \to D^* D^0 X$	Freely varies
$f_{D^*3\pi X}$	$B \to D^* 3\pi X$	Freely varies
$f_{\text{comb.}}$	Combinatorial	Constrained by sum of all components to unity

8.4 FIT TO DATA

The binned maximum likelihood fit to data is performed simultaneously across the 3D distribution of reconstructed decay angles and the 2D distribution of reconstructed tau flight significance and Isolation BDT. The total PDF, $\mathcal{P}(V)$, used to describe the binned data is detailed in Section 8.3.8. To define the bin boundaries for each fit variable, the signal simulation is binned separately in each variable such that each bin is uniformly populated. This non-uniform binning is chosen to ensure a roughly equal population of signal events across the full fit phase space, which also ensures that each data bin in the fit is adequately populated to avoid empty bins.

Ten bins are chosen per fit variable in the 3D angular distribution and thirty bins are chosen per fit variable in the 2D complementary distribution, which corresponds to a total of 1000 bins in the decay angles and 900 bins in the complementary variable space. The total ${}_s$Weighted data yield from the 2D fit to $m(D^* - D^0)$ and $m(D^0)$ is $33,010 \pm 220$: the total yield of the data histogram in the fit is thus 33,010 when summed over all bins in a given category.[1] Using 1000 bins, an average data yield of around 33 events per bin is expected.

[1]The dataset in each fit category is identical, and thus both data histograms sum to the same number of events across all bins.

8.4.1 FIT PROJECTIONS

The one-dimensional fit projections are shown in Figures 8.12 and 8.13, for the decay angles category and complementary variables category of the fit respectively. The total $_s$Weighted data yield in each bin is shown by points, with errors calculated according to the square root of the sum of squared weights in each bin. The total fit PDF in each bin is shown by a black solid line, and each fit component is shown by a filled histogram of a different colour. Below each fit projection, the residuals between the data and PDF is given for each bin: The projections indicate good agreement between data and the total PDF, and the total fit χ^2/dof is found to be 1.058.

A more complete visualisation of the fit quality can be achieved by plotting each bin of the fit unrolled along a single axis. This allows the data and PDF yields in all bins of the fit to be displayed, as well as their residuals. The unrolled plots are shown in Figures 8.14 and 8.15, and demonstrate good behaviour of the fit across the full phase space considered.

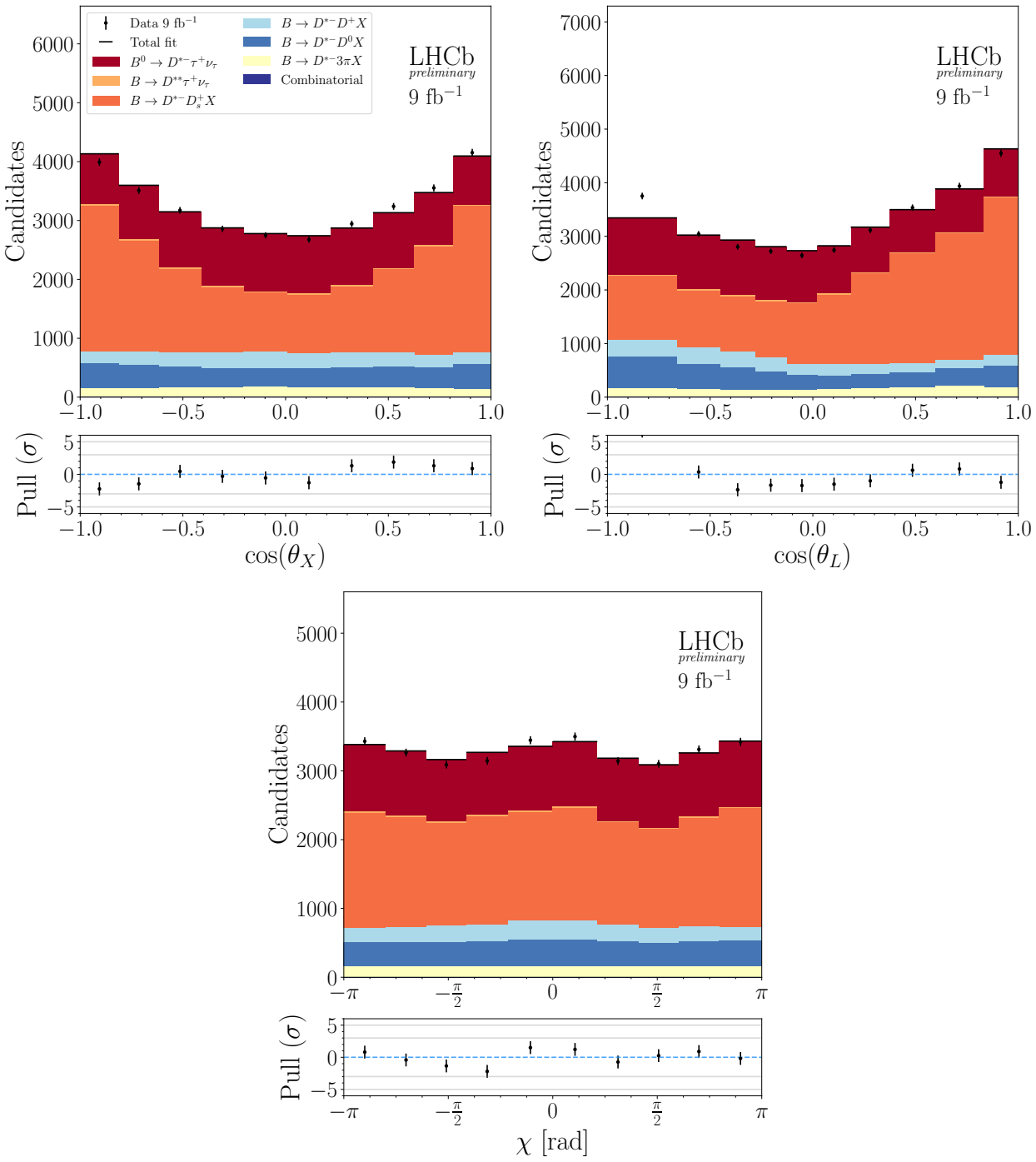


Figure 8.12: One-dimensional projections of the angular fit to data (decay angles category).

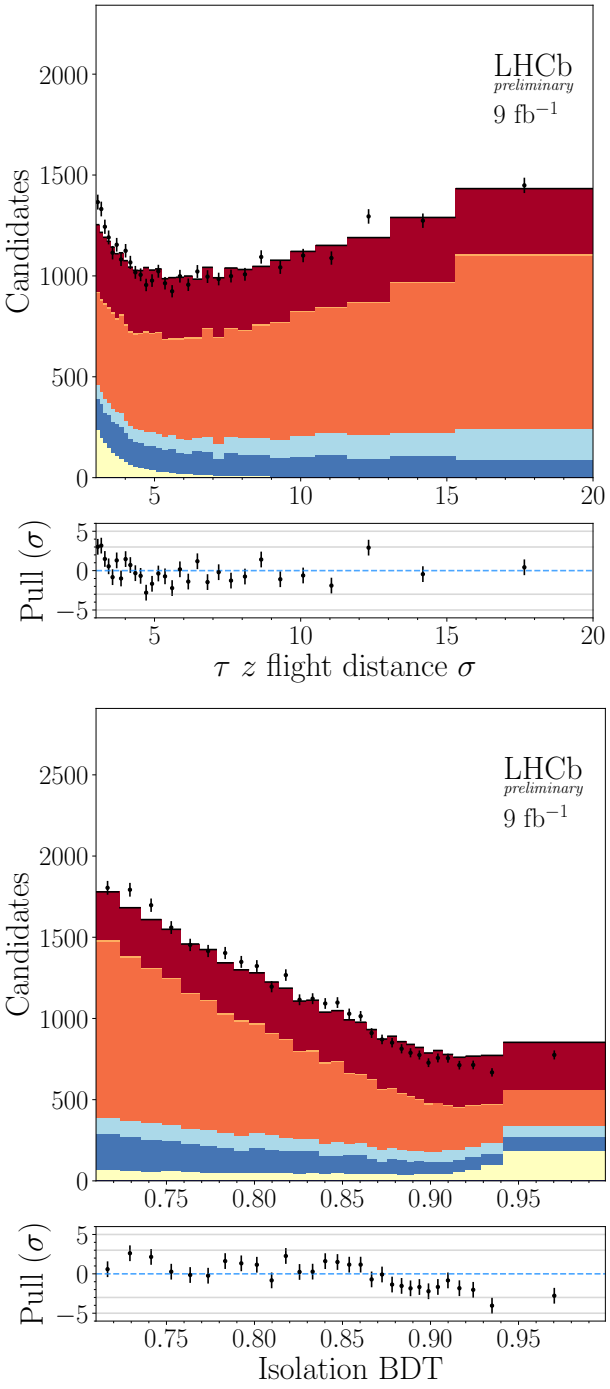


Figure 8.13: One-dimensional projections of the angular fit to data (complementary variables category).

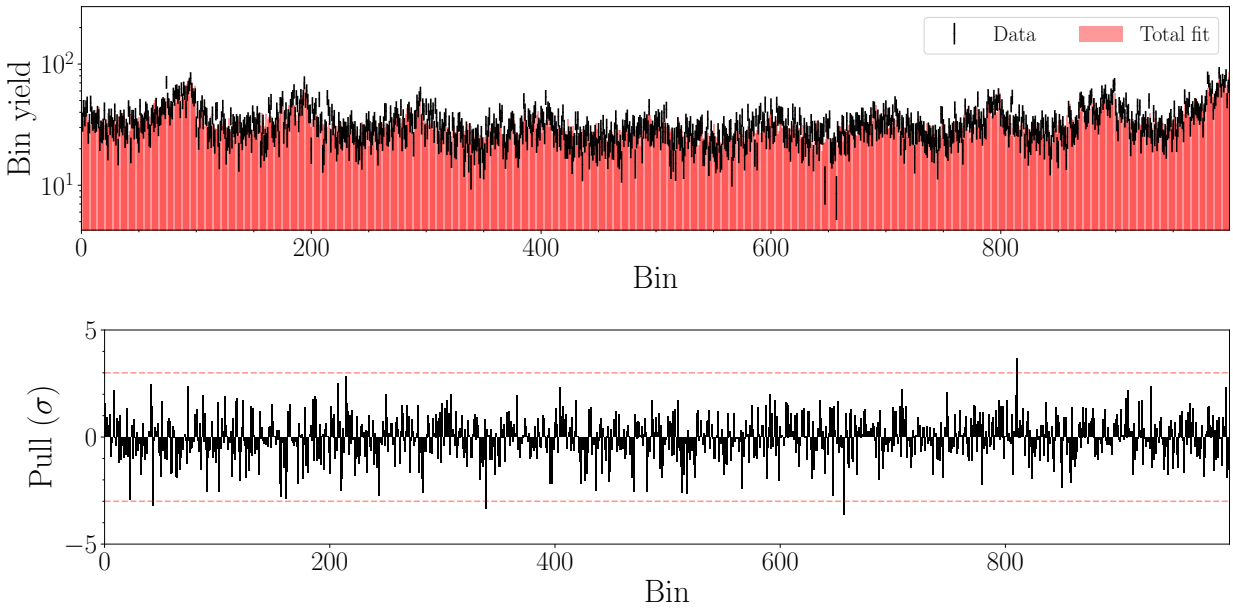


Figure 8.14: Each bin of the angular fit unrolled along a single axis (decay angles category).

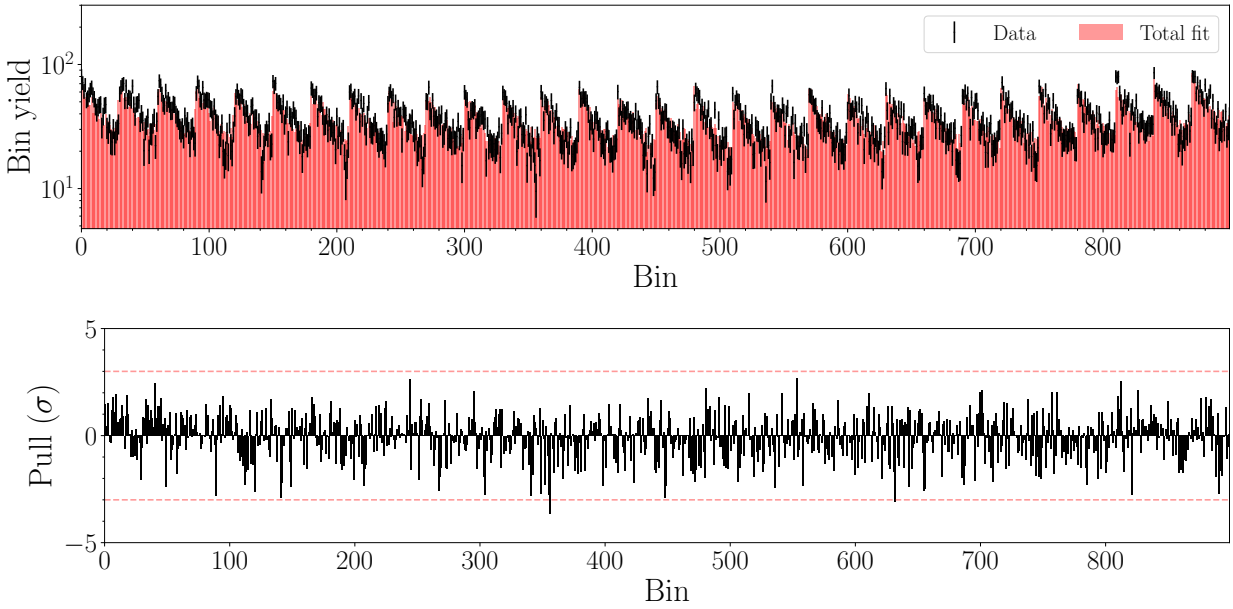


Figure 8.15: Each bin of the angular fit unrolled along a single axis (complementary variables category).

8.5 PARAMETER UNCERTAINTY

In the simultaneous fit to data, the 3D reconstructed angular distribution is fitted simultaneously with the 2D distribution in reconstructed tau flight significance and Isolation BDT. The same data are used in both categories of the fit, and the uncertainties computed by Hesse do not account for this double-counting effect. As a result, the parameter uncertainties reported by Hesse are underestimates.

To assess the genuine statistical uncertainties on each parameter due to the finite fit dataset, 2000 bootstrapped data samples are created by randomly sampling the data with replacement [90]. Fits to each of the bootstrapped datasets are performed, where the statistical fluctuations introduced by the bootstrapping are common to both 3D fit categories. The spread in parameter results across many bootstrap fits thus provides a natural measure of the fit statistical uncertainties, including the effect of the data double-counting.

To measure the bootstrap parameter uncertainties, Gaussian fits to the bootstrap distributions for each parameter are performed after statistical subtraction of the mean values. Mean subtraction is performed in order to centre each bootstrap distribution at zero, in order to avoid any interpretation of central values from the bootstrap fits. The bootstrap distributions for the angular coefficients, the signal fraction, and each freely varying background fraction are shown in Figures 8.16, 8.17, 8.18, and 8.19. Derived parameters such as I_{1c} are also naturally derived from the bootstrap distributions, via calculations performed per bootstrap fit using the other measured I_x values. All bootstrap distributions show the expected Gaussian behaviour, and the widths are measured with negligible uncertainties. These widths are henceforth taken to be the fit statistical uncertainties for each parameter.

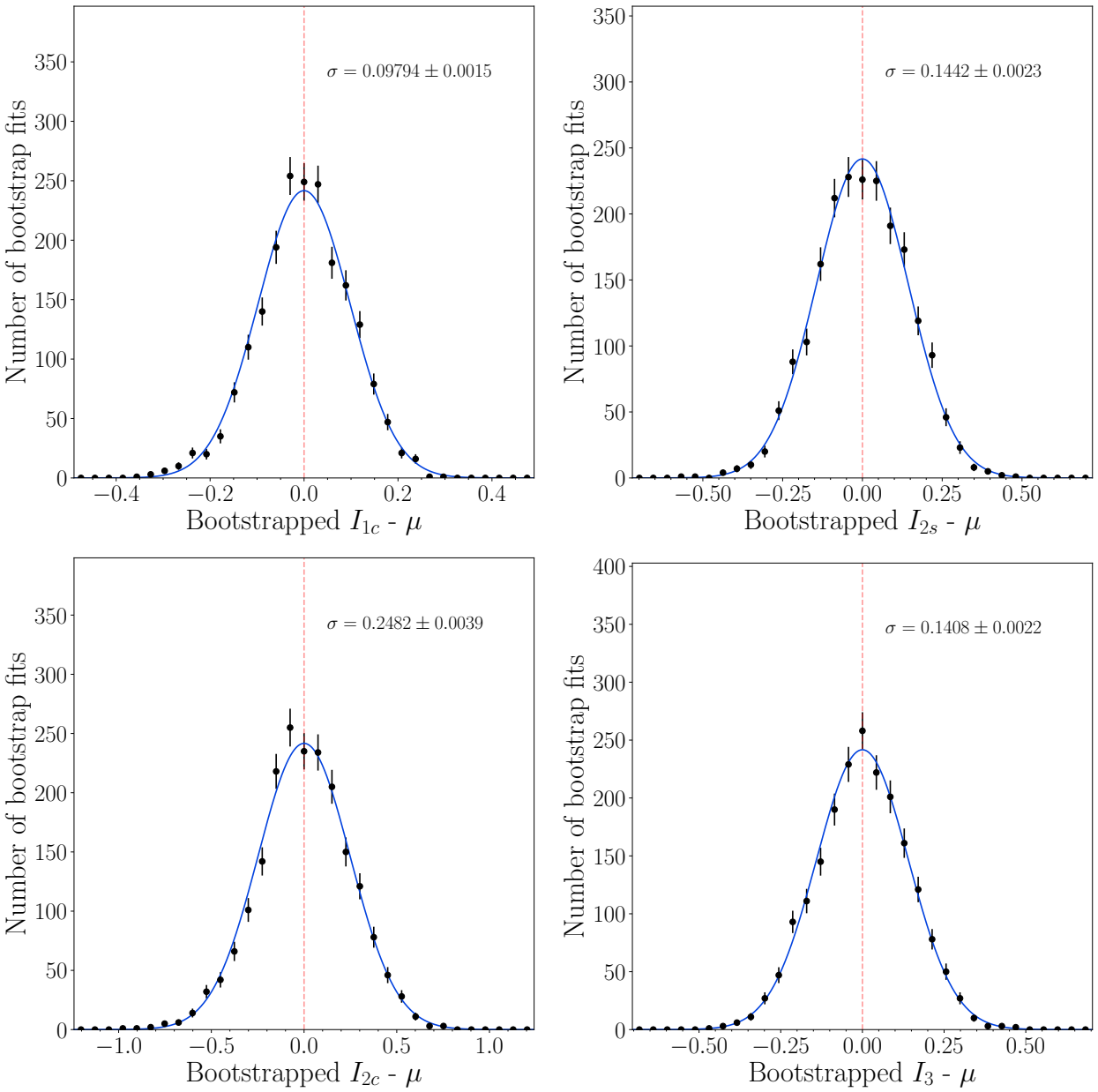


Figure 8.16: Bootstrap distributions for each fit parameter after mean subtraction. The widths measured by Gaussian fits are taken as the fit statistical uncertainties henceforth.

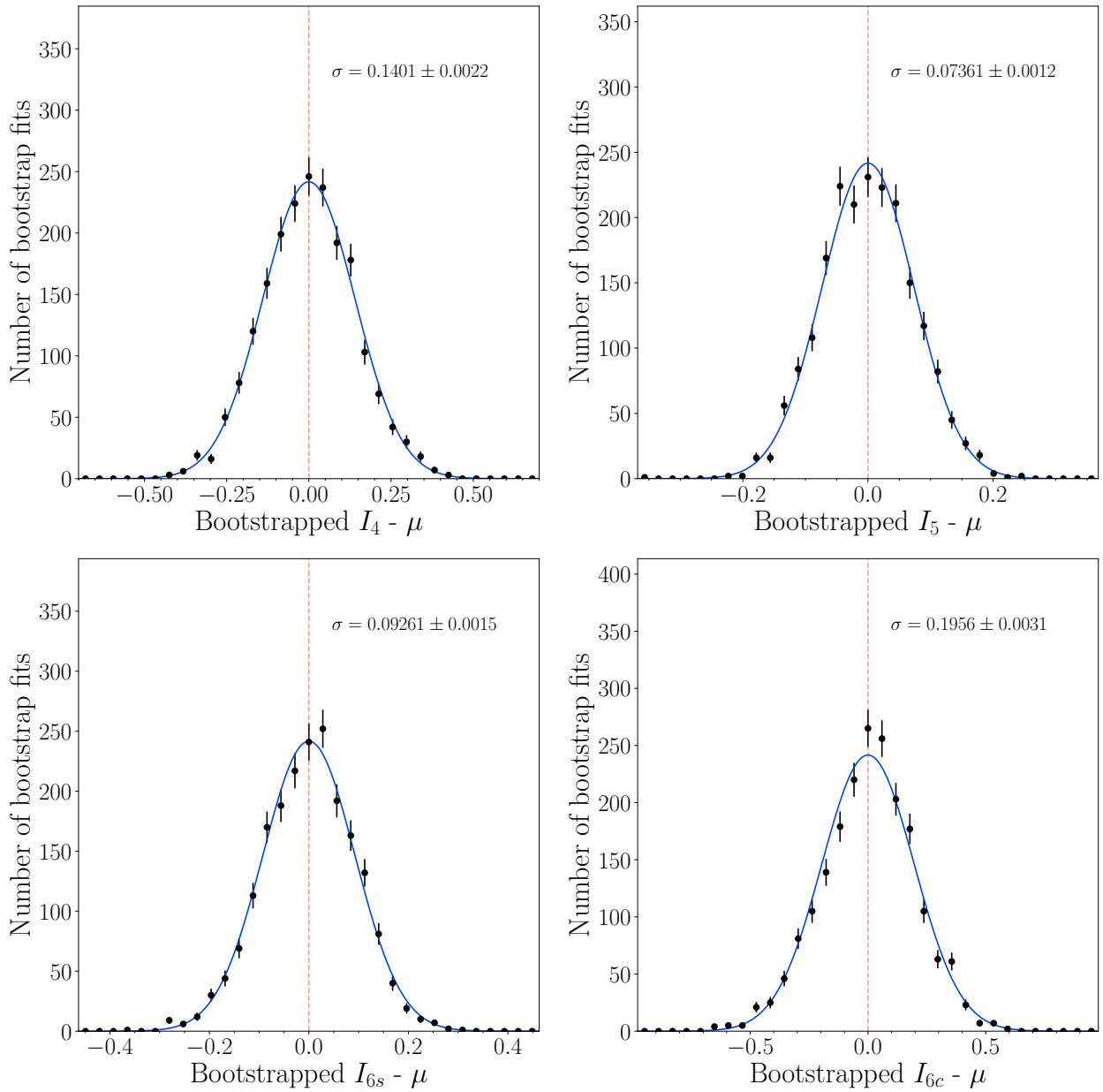


Figure 8.17: Bootstrap distributions for each fit parameter after mean subtraction. The widths measured by Gaussian fits are taken as the fit statistical uncertainties henceforth (continued).

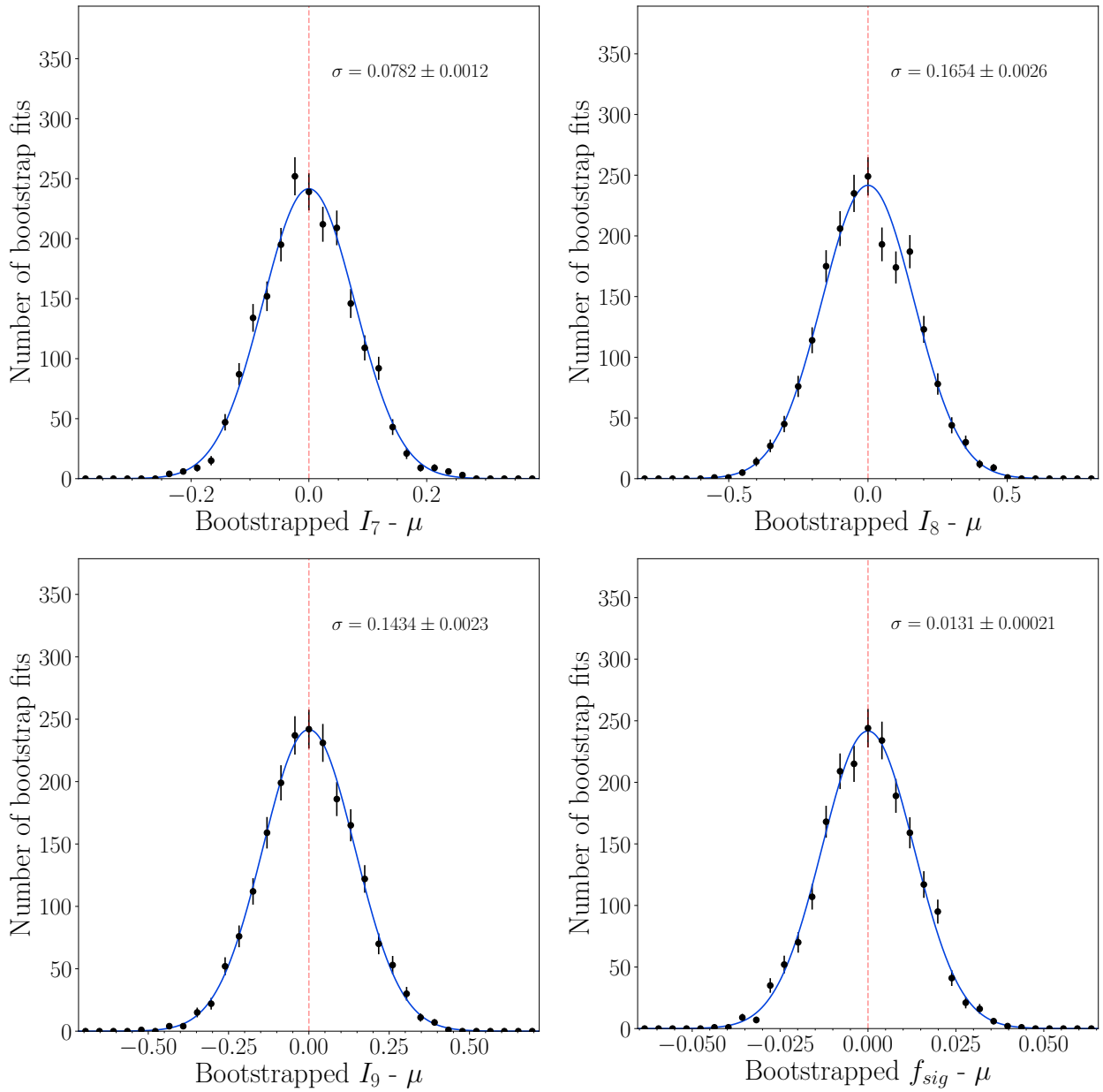


Figure 8.18: Bootstrap distributions for each fit parameter after mean subtraction. The widths measured by Gaussian fits are taken as the fit statistical uncertainties henceforth (continued).

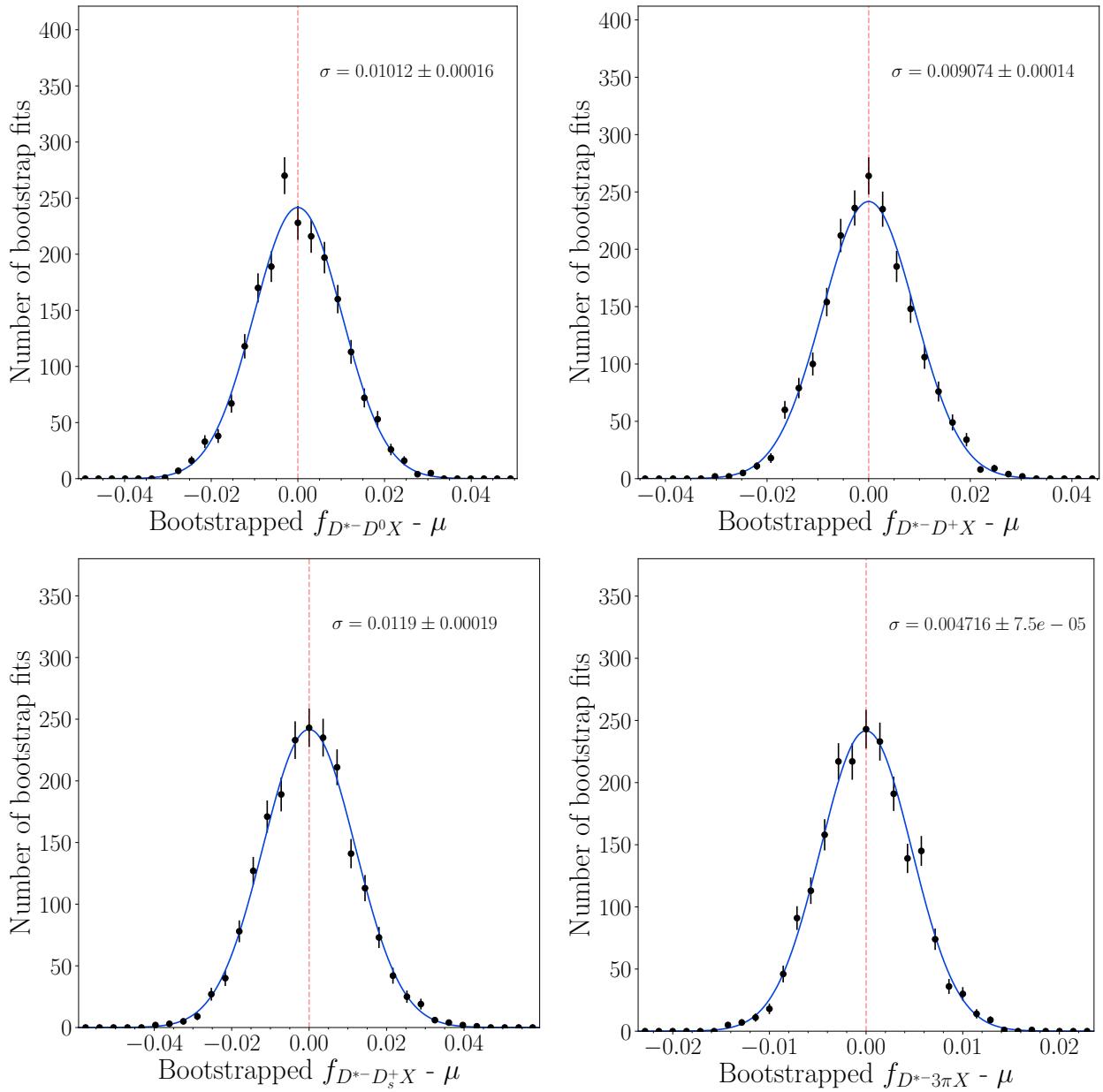


Figure 8.19: Bootstrap distributions for each fit parameter after mean subtraction. The widths measured by Gaussian fits are taken as the fit statistical uncertainties henceforth (continued).

8.6 PSEUDO-EXPERIMENT FITS

To check for biases in the fit procedure, a series of pseudo-experiments are performed. To create pseudo-experiment datasets, the final normalised data fit PDF is first multiplied by the total data yield to determine the PDF yield per bin. A pseudo-experiment sample is then created by Poisson varying each bin of the total PDF according to its yield. Fits are then performed to each pseudo-experiment sample, in order to build pull distributions for the free parameters of the fit. Parameters measured without bias should exhibit pull distributions with a mean of zero and a width of unity, where the pull value for a given parameter x as measured in one pseudo-experiment fit is

$$P_x = \frac{x_{\text{fitted}} - x_{\text{generated}}}{\sigma_x}$$

where x_{fitted} is the fit result for parameter x, $x_{\text{generated}}$ is the value used in the PDF for pseudo-experiment generation, and σ_x is the uncertainty as measured by Hesse in the fit. One-dimensional projections of an example pseudo-experiment fit are shown in Figures 8.20 and 8.21.

The pull distributions for the I_x coefficients, the signal fraction, and the freely varying background fractions are shown in Figures 8.22, 8.23, 8.24, and 8.25: all parameters exhibit pull means and widths consistent with the unit Gaussian, and as such the fit is deemed to be free of statistical biases. The pseudo-experiment samples generated here introduce statistically independent variations in the two 3D fit categories, and thus do no include the data double-counting effect present in real data. The double-counting effect is fully addressed by the bootstrapping procedure in Section 8.5, however.

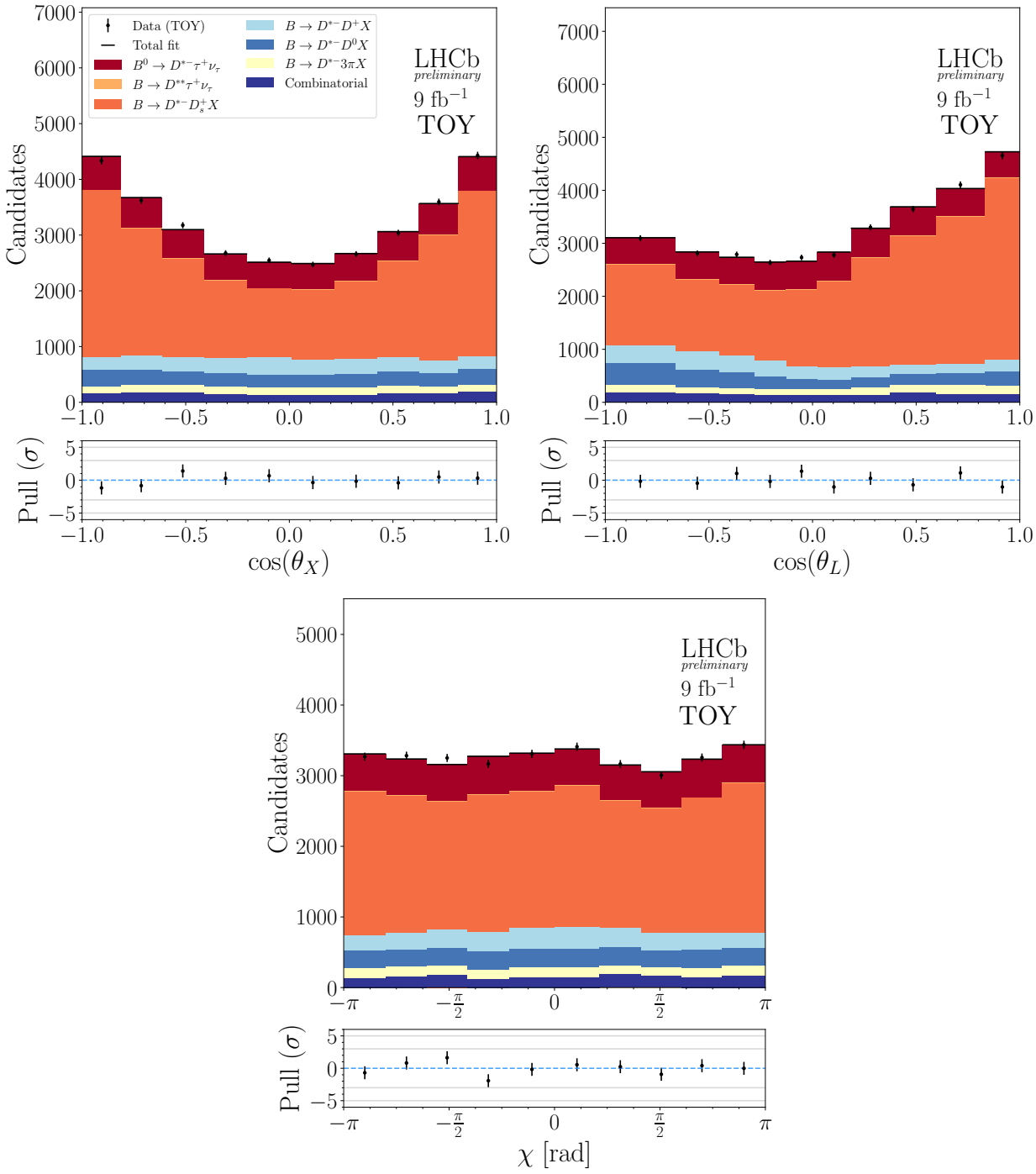


Figure 8.20: One-dimensional projections of the angular fit to a single pseudo-experiment dataset (decay angles category).

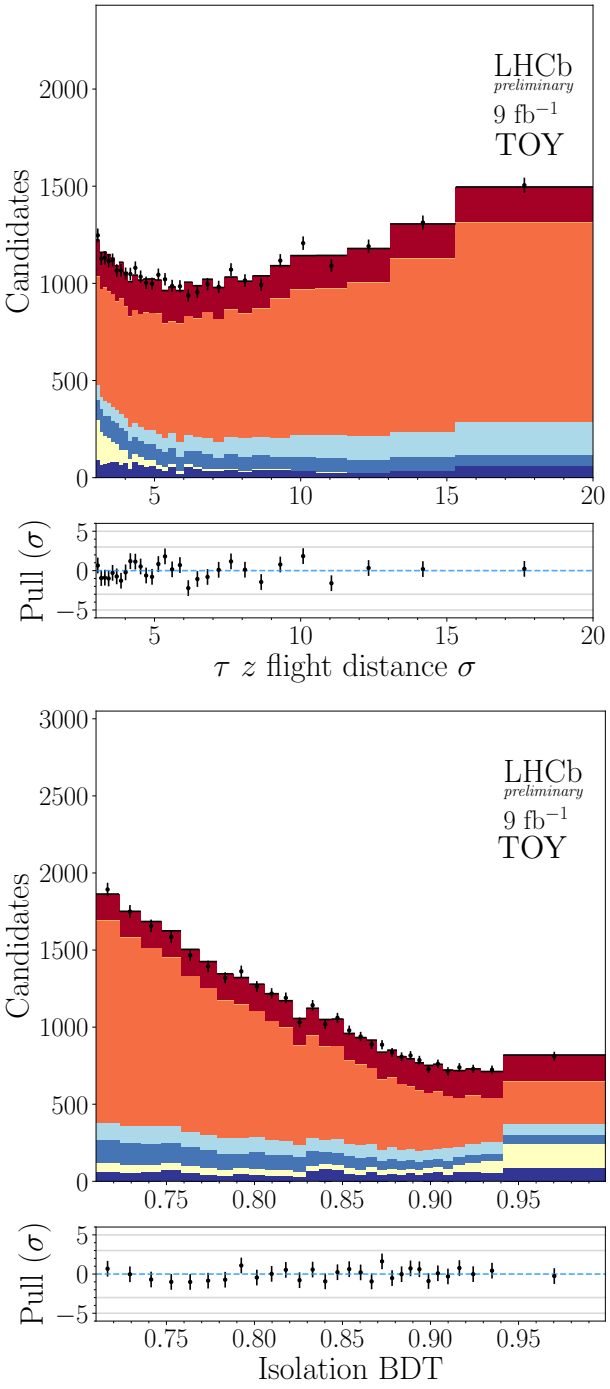


Figure 8.21: One-dimensional projections of the angular fit to a single pseudo-experiment dataset (complementary variables category).

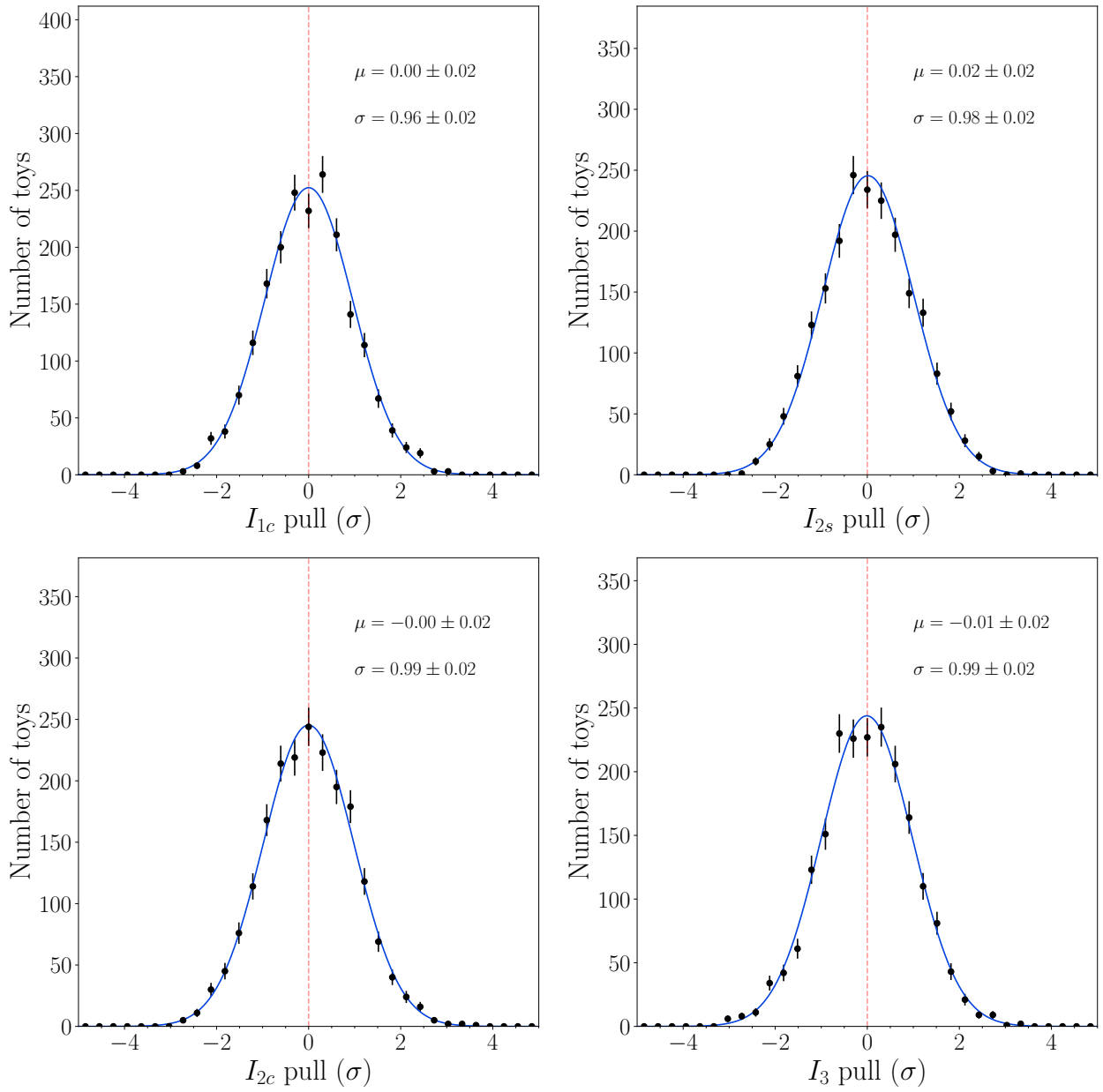


Figure 8.22: Pull distributions for each fit parameter as measured in pseudo-experiment fits. The means and widths measured by Gaussian fits are consistent with unit Gaussians for all fit parameters.

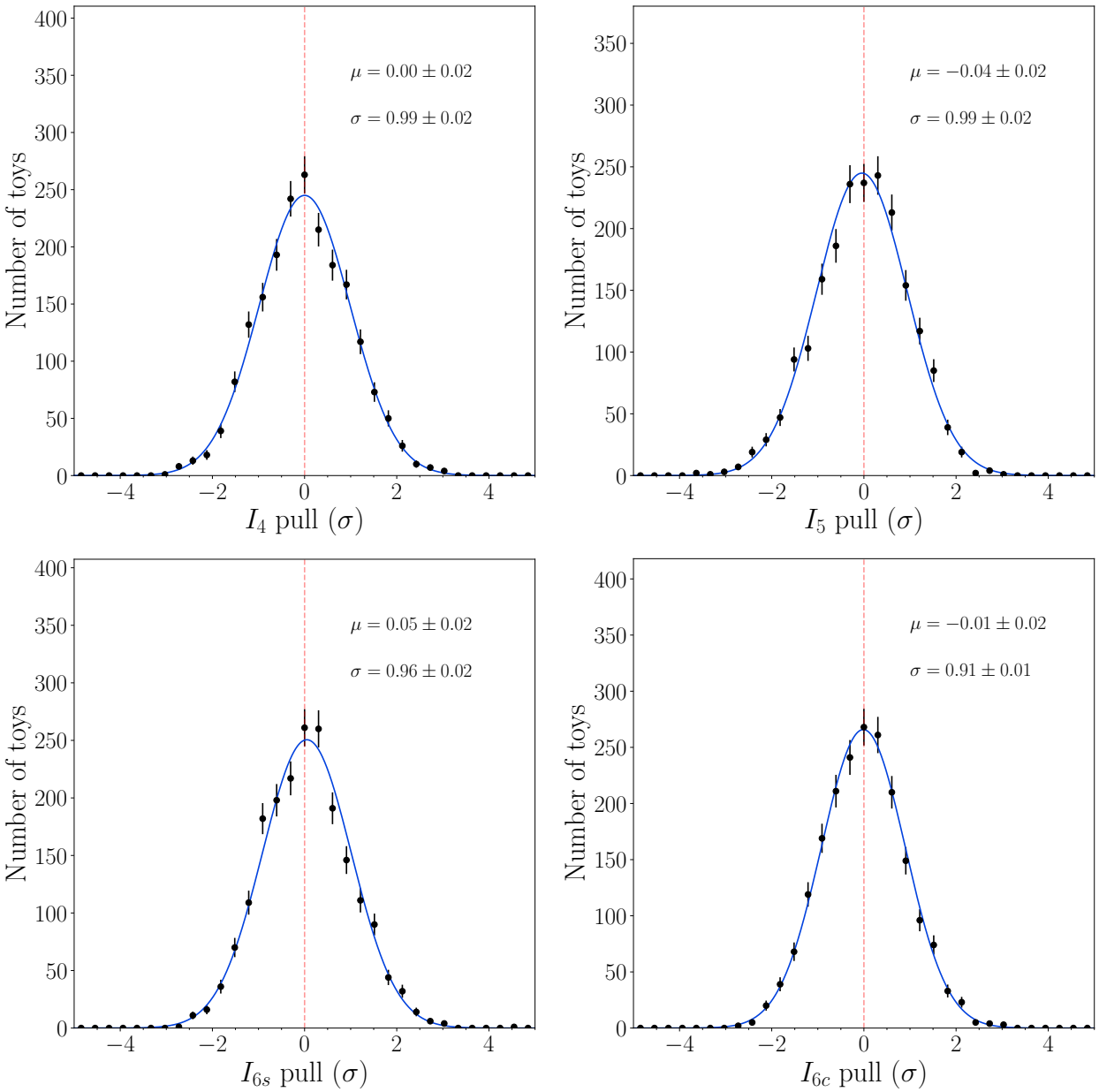


Figure 8.23: Pull distributions for each fit parameter as measured in pseudo-experiment fits. The means and widths measured by Gaussian fits are consistent with unit Gaussians for all fit parameters (continued).

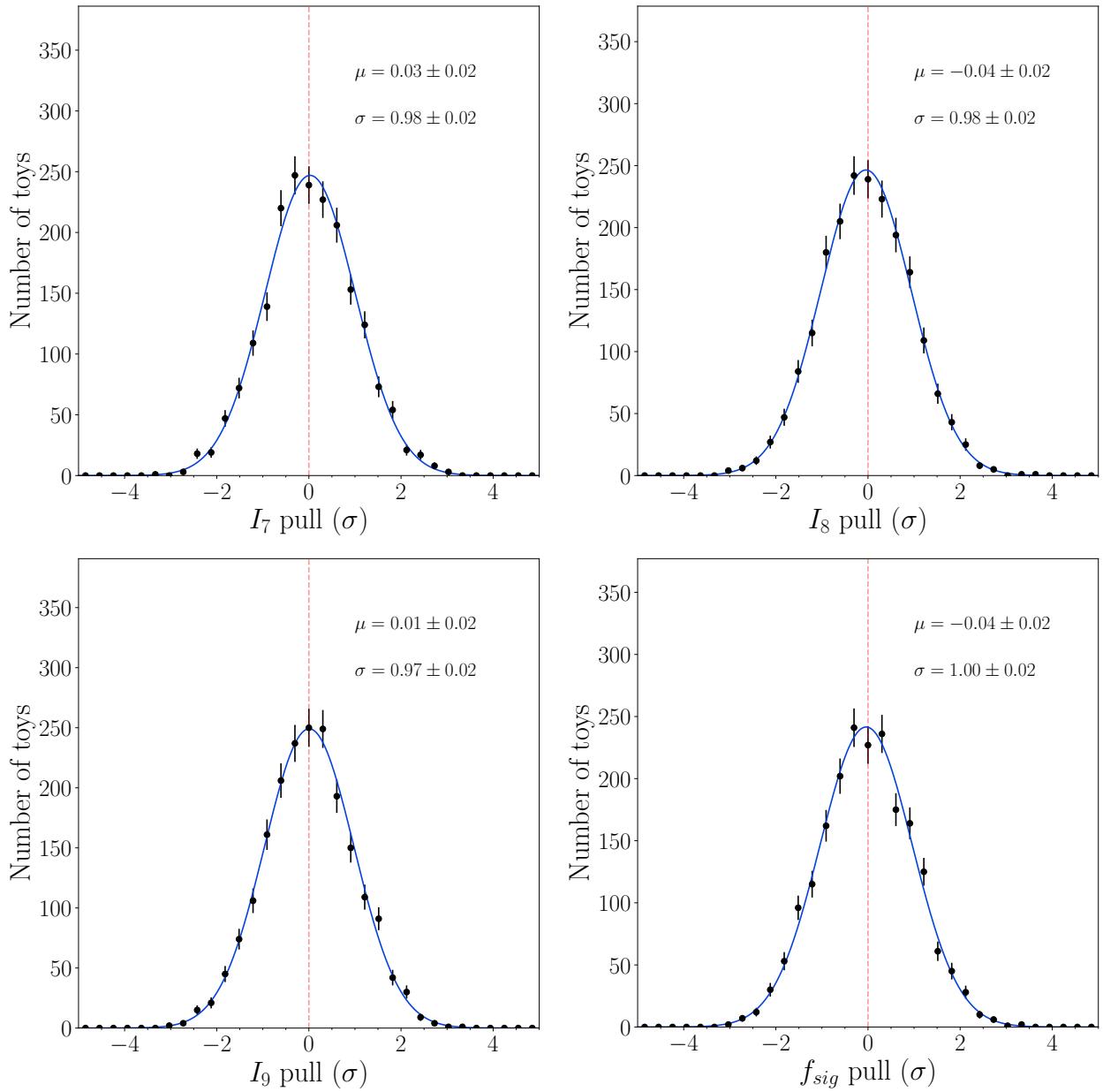


Figure 8.24: Pull distributions for each fit parameter as measured in pseudo-experiment fits. The means and widths measured by Gaussian fits are consistent with unit Gaussians for all fit parameters (continued).

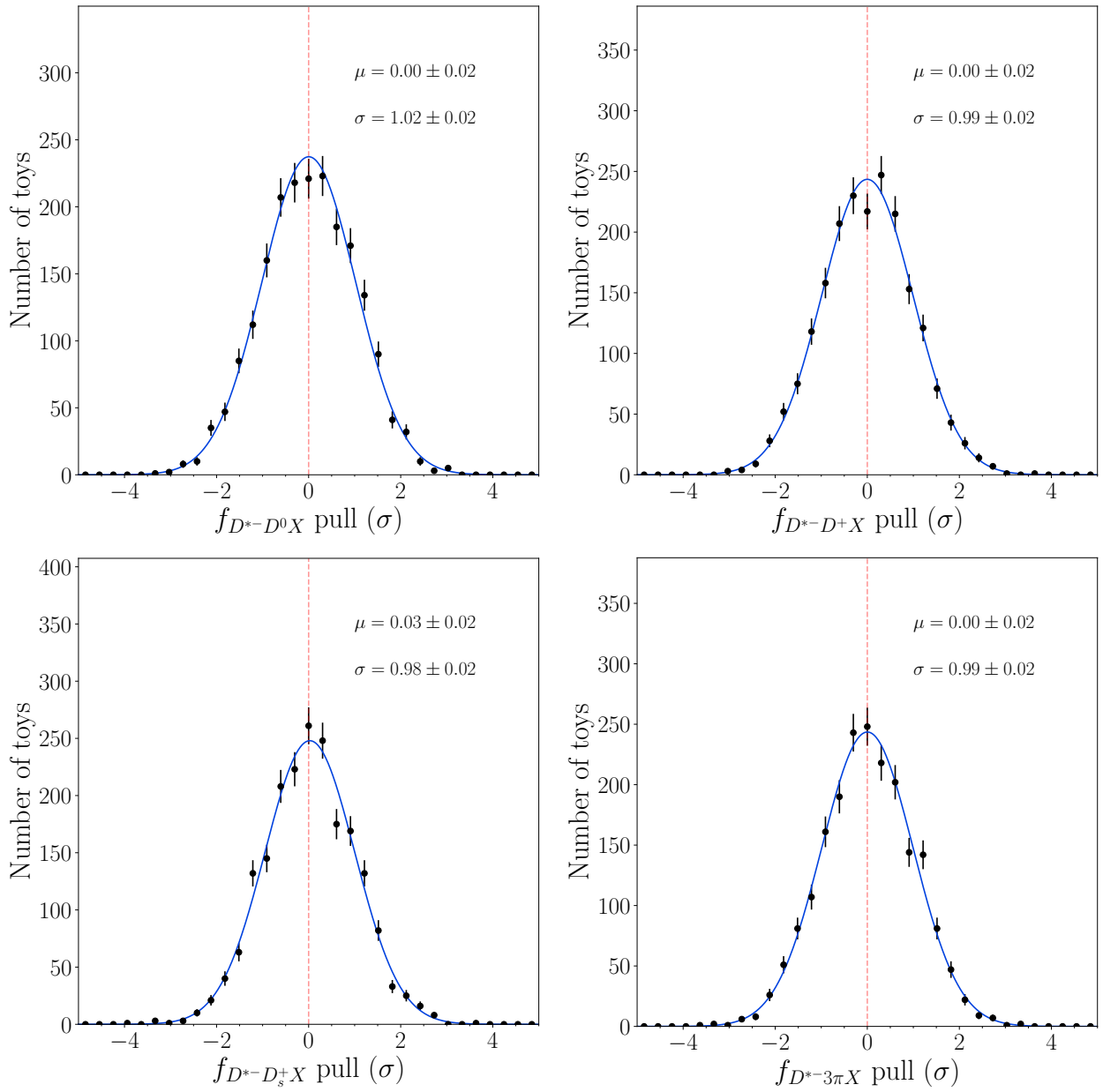


Figure 8.25: Pull distributions for each fit parameter as measured in pseudo-experiment fits. The means and widths measured by Gaussian fits are consistent with unit Gaussians for all fit parameters (continued).

8.7 NUMERICAL RESULTS

The numerical results of the angular fit to data are summarised in Table 8.6, where the final values of all free and Gaussian-constrained parameters in the fit are given. As previously stated, the q^2 integrated values of the $I_x(q^2)$ are measured and are referred to as I_x. The uncertainties quoted are taken from the data fit bootstrapping procedure detailed in Section 8.5, in order to correctly account for the use of the same data events in both 3D categories of the simultaneous fit.

The relative statistical precision on the signal fraction is found to be 4.5%, while the angular coefficients are measured with large relative uncertainties. This level of precision is expected given the limited sample size and purity of the data, and is commensurate with the performance detailed in the $B^0 \to D^*\tau\nu$ angular fit methods paper [30]. In particular, the signal fraction statistical precision is larger than a naive extrapolation from the 6.6% statistical precision published in the 3 fb^{-1} hadronic $R(D^*)$ analysis [20]. In the 9 fb^{-1} Run 1 + 2 sample, the total dataset size is around five times larger than in the 3 fb^{-1} analysis: a reduction in signal fraction uncertainty from 6.6% to 3% might thus be expected. The 4.5% precision achieved in this analysis is a factor of 1.5 larger, which is attributed to the following features of the measurement.

The signal model has full freedom in the angular fit, where variation in both the underlying decay amplitudes and hadronic form-factors is permitted via the I_x coefficients. This introduces additional covariance in the data fit, and thus a larger degree of uncertainty on the total signal rate. In the previous LHCb measurement of $R(D^*)$ using hadronic τ decays, the SM signal distribution was assumed in the signal fit with no associated systematic uncertainty, although a systematic uncertainty of 0.7% relative was assigned for the choice of form-factor scheme [20].

In this analysis, the feed-down background contribution is constrained in the fit as well as the internal B-decay fractions of the $B \to D^* D_s X$ background; all other background rates vary freely. In the 3 fb^{-1} analysis, the feed-down and combinatorial background rates were fixed, and constraints were additionally applied to the rate of $B \to D^* 3\pi X$ prompt background and a subset of the $B \to D^* D^0 X$ background

Table 8.6: Parameters of the angular fit to data and their final numerical values. The free parameter uncertainties quoted are taken from many fits to bootstrapped data samples, to account for the use of the same data events in both 3D categories of the simultaneous fit.

Parameter	Value	Free / Constrained
I_{1s}	0.611 ± 0.066	Freely varies
I_{2c}	0.123 ± 0.251	Freely varies
I_{2s}	0.689 ± 0.159	Freely varies
I_3	-0.149 ± 0.141	Freely varies
I_4	0.040 ± 0.141	Freely varies
I_5	-0.086 ± 0.072	Freely varies
I_{6c}	-0.158 ± 0.191	Freely varies
I_{6s}	-0.200 ± 0.090	Freely varies
I_7	0.045 ± 0.077	Freely varies
I_8	0.130 ± 0.163	Freely varies
I_9	0.112 ± 0.137	Freely varies
f_{sig}	0.283 ± 0.013	Freely varies
$f_{D^*D^0X}$	0.111 ± 0.010	Freely varies
$f_{D^*D^+X}$	0.071 ± 0.009	Freely varies
$f_{D^*3\pi X}$	0.046 ± 0.005	Freely varies
$f_{D^*D_sX}$	0.481 ± 0.012	Freely varies
$f_{D^{**}\tau\nu}$	0.023 ± 0.004	Gaussian constrained
$f_{B^0\to D^*D_s}$	0.147 ± 0.012	Gaussian constrained
$f_{B^0\to D^*D_{s0}(2317)^*}$	0.000 ± 0.001	Gaussian constrained
$f_{B\to D^{**}D_sX}$	0.220 ± 0.015	Gaussian constrained
$f_{B^0\to D^*D_{s1}(2460)}$	0.040 ± 0.012	Gaussian constrained
$f_{B_s\to D^{**}D_sX}$	0.158 ± 0.015	Gaussian constrained

(the so-called "same-vertex" component). Such fixed and constrained terms remove uncertainty from the fit and reduce the statistical uncertainty on the signal yield.

The selection BDTs used in the analysis are trained using a cross-validation procedure with hyper-parameter tuning, to ensure that the signal and background simulation samples used both in BDT training and fit template creation do not exhibit any over-training artefacts. In addition, the choice of BDT training variables avoids the use of any variables which appear in the double-charm control study fits which minimises the dependence of the control studies on the selection, particularly since the control regions are selected with different requirements than the final signal fit data. An additional benefit is that the BDT distribution used in the final angular fit has a weaker dependence on the underlying $D_s \to 3\pi X$ modelling assumptions. Overall, the choices made in BDT implementation result in a set of BDTs with

lower separation power than in previous $R(D^*)$ measurements. However, after spline reweighting, all BDT distributions are well aligned between data and simulation, and scale well into the signal fit region as evidenced by the good description of the data BDT distribution in the angular fit.

9 SYSTEMATICS

The observables measured in the angular fit, namely the angular coefficients I_x and the signal fraction f_{sig}, are subject to fit-related systematic uncertainties. These uncertainties are discussed and quantified below. These uncertainties affect all fitted parameters and derived values, in particular, the I_x coefficients.

9.1 FIT BASED SYSTEMATICS

Since the angular coefficients are measured directly by the fit there are no external inputs used after the fit is performed to convert fit parameters to observables. This is in contrast to an $\mathcal{R}(D^*)$ measurement using hadronic tau decays, where branching fraction measurements and simulation efficiencies are used to convert the measured signal fractions into a measurement of $\mathcal{R}(D^*)$. Due to the direct measurements of the I_x in the signal fit, the only systematic uncertainties are those arising from the fit itself.

9.1.1 FINITE TEMPLATE SAMPLE SIZE

All contributions in the angular fit are modelled by template histogram PDFs, where the PDF bin contents are defined from finite simulation or data control samples. As such, the signal and background PDFs are subject to systematic uncertainties due to finite statistics. To assess this systematic uncertainty, the samples used to generate fit templates are bootstrapped (random sampling with replacement), and many fits to data are run with templates derived from the bootstrapped samples. Bootstrapping ensures that the templates in both 3D categories of the simultaneous fit are subject to the same statistical variations. Randomly varying each template PDF bin does not couple the statistical fluctuations across each of the 3D fit categories, and as such bootstrapping is preferred.

Fits to data are run 2000 times with all templates systematically varied at the same time, providing a measure of the total systematic effect due to finite template statistics. Distributions for each fit observable are built from the set of 2000 system-

atic fits, and the widths of the distributions are taken to represent the systematic uncertainty.

9.1.2 $D_s \to 3\pi X$ MODELLING

The $B \to D^* D_s X$ template is derived from a dedicated generation containing a dedicated set of $D_s \to 3\pi X$ decays. A control fit to data is performed in order to measure the fractional contributions from different D_s decays to the $3\pi X$ final state, as described in Section 7.1. The results of this fit are used to assign weights per decay mode to the $B \to D^* D_s X$ simulation, in order to align the contributions from each decay with those observed in the data. These weights are subject to systematic uncertainties arising from the finite precision with which the decay mode fractions are measured in data.

Fits to data are run 2000 times, where the D_s decay mode weights for each decay mode in the $B \to D^* D_s X$ simulation samples are varied randomly within their uncertainties prior to building the total template. This introduces variation in the total $B \to D^* D_s X$ fit template, which can thus impact both the signal fraction and I_x measurements. The widths of the systematic distributions for each observable under variation of the D_s weights are taken to represent the systematic uncertainties from this source.

9.1.3 L0 TRIGGER PROPORTIONS IN SIMULATION

The proportion of simulation events falling in each L0 trigger category (L0 Hadron TOS, L0 Global TIS, and both) are compared to data using the $B^0 \to D^* 3\pi$ mode (see Section 7.4). The proportions of events in each L0 category are calculated both in data and simulation, and data/simulation ratios are used to derive a set of correction weights for all simulation samples. Each ratio is subject to a systematic uncertainty, arising from the finite $B^0 \to D^* 3\pi$ statistics in data and simulation.

During fit template creation, L0 weights are assigned to all simulated samples based on the above $B^0 \to D^* 3\pi$ comparison. To assess the systematic uncertainty, the correction weights for each L0 category are varied at random within their un-

certainties, and templates are created using the varied weights. Many fits to data are performed under such variations, and the widths of the observable distributions formed are taken to represent the systematic uncertainties.

9.1.4 PID CALIBRATION

In the analysis, particle identification requirements are applied to each of the pions in the 3π system in order to minimise the contributions from the misidentified background. As the PIDK distributions in the simulation are not well reproduced, direct requirements on the pion PIDK distributions are not applied. Instead, PID efficiencies are assigned per-event to simulation based on efficiency histograms created in PIDCalib2. Efficiencies are assigned to each 3π pion based on its momentum and pseudorapidity, and the total PID efficiency for an event is then given by the product of the individual pion PID efficiencies. These efficiencies are treated as per-event weights during template creation, as they act to adjust the template distributions in the same way as the direct set of PIDK requirements do when applied to data.

All per-event PID efficiencies in the simulation are assigned a total systematic uncertainty of 0.6%, which arises from a combination of fully correlated 0.2% uncertainties on each pion in the 3π system. The total uncertainty assigned covers the following contributing sources of systematic uncertainty:

- The use of finite data PID calibration samples;
- The application of ${}_s$Weights in PIDCalib when creating efficiency histograms;
- The use of a finite set of variables within which to parameterise the pion PID efficiency variation;
- The use of a particular binning scheme in momentum and pseudorapidity for PID efficiency assignment.

During template creation, the PID efficiency weights for each event are independently varied according to an absolute uncertainty of 0.6%. This variation is introduced across all template samples at the same time, providing an overall measure of

the PID systematic uncertainty. Many fits are run under variations of the per-event PID weights, and the widths of the observable distributions formed are taken to represent the systematic uncertainties.

9.1.5 $3\pi X$ KINEMATICS AND FLIGHT MODELLING

In Section 7.5, a comparison of inclusive $b\bar{b} \to D^{*}3\pi X$ simulation and data is performed after the full selection is applied. The purpose of this comparison is to derive a set of correction weights for the 3π system kinematics and flight, which are not controlled via the other double-charm control studies in the analysis. Ratios of data and inclusive simulation are formed in the 3π momentum, transverse momentum, and tau flight significance, and spline fits are performed in order to derive a set of correction functions to be applied to all simulation samples used for template creation in the angular fit.

The splines are subject to systematic uncertainties arising from the finite data and inclusive simulation samples used, as well as from the specific signal fraction present in the inclusive simulation sample. To determine the systematic uncertainty on the splines, many spline fits are performed with bootstrapping applied to both the data and inclusive simulation samples. The simulation sample used in this study contains signal decays following SM predictions. During the bootstrapping, the signal component in each bootstrapped simulation sample is assigned a random uniform weight in the range $(0, 2)$, which will change the signal fraction in the simulation, in order to propagate a large signal fraction uncertainty to the binned data/simulation ratios.

The net result is a distribution of $3\pi X$ correction weights for every simulation event used in template creation. To determine the systematic uncertainty, the default $3\pi X$ weights for each event are varied randomly within the standard deviation of their individual $3\pi X$ weight distributions. This random weight variation is applied to all simulation samples during template creation, and many fits to data are run using the varied templates. The widths of the observable distributions formed from these systematic fits are taken to represent the systematic uncertainties arising from the $3\pi X$ correction splines.

9.1.6 $B \to D^* D^+ X$ MODELLING

In Section 7.3, $B \to D^*(D^+ \to K\pi\pi)X$ data and simulation are compared in order to derive a set of correction weights for the full $B \to D^* D^+ X$ simulation sample. This comparison relies on the same spline-based method described above for the $3\pi X$ corrections, and is subject to uncertainty arising from the finite $B \to D^*(D^+ \to K\pi\pi)X$ data and simulation samples. As for the $3\pi X$ case, many spline fits are performed under bootstrapping of the data and simulation control samples, to form a distribution of correction weights for each $B \to D^* D^+ X$ simulation event: the widths of these distributions are taken to represent the uncertainty on the correction weights for each event. During angular fit template creation, each event weight is varied randomly within its uncertainty, and fits to data are performed under many varied templates. The widths of the observable distributions formed from these systematic fits are taken to represent the systematic uncertainties arising from the $B \to D^*(D^+ \to K\pi\pi)X$ correction splines.

9.1.7 $B \to D^* D^0 X$ MODELLING

To align the $B \to D^* D^0 X$ simulation sample with data, a set of correction weights is assigned based on a comparison of $B \to D^*(D^0 \to K3\pi)X$ data and simulation. The spline-based procedure is subject to the same type of systematic uncertainties as the $B \to D^* D^+ X$ corrections described above, and systematic uncertainties are evaluated in an identical manner via random variations introduced to the correction weights.

9.1.8 $B \to D^* D_s X$ MODELLING

A dedicated fit to the $m(D^*(D_s \to 3\pi))$ distribution in data is performed in Section 7.2.2 to determine the B-decay fractions of the total $B \to D^* D_s X$ sample. These fractions are subsequently used in the angular fit as Gaussian constraints, with widths representative of the total uncertainty on each fraction including systematic uncertainties.

After the B-decay fractions are measured, a comparison of $B \to D^*(D_s \to 3\pi)X$ data and simulation is performed in order to derive correction weights for the full $B \to D^* D_s X$ simulation sample. This procedure is based on the same sequential spline fitting method detailed above for $B \to D^* D^+ X$, and is subject to the same systematic uncertainties. Many fits to data are performed under variations of the per-event correction weights, and the widths of the observable distributions found are taken to represent the systematic uncertainties.

9.1.9 MODEL DEPENDENCE

A limitation of the I_x template method is that the summed templates in supporting variables don't capture the full q^2 dependence in the signal. This only poses a problem when supporting variables used in the fit themselves depend upon q^2.

In this measurement, the Isolation BDT is used which does not depend on q^2 (it uses event-level information not specific to the signal) and thetau flight significance which does depend slightly on q^2 due to the momentum dependence of the flight reconstruction.

The change in the supporting variable distributions under variations in the signal mode q^2 distribution was studied. The q^2 distribution was modified through a reweighting procedure where these weights were calculated with

$$w_{q^2} = 1 + \alpha \frac{q^2}{r_{q^2}}$$

where α is a random number in the range [-0.75, 0.75] and r_{q^2} is the range the q^2 values can take; $r_{q^2} = 10.7\text{GeV}^2/c^4$. These weights were then applied to the signal simulation to study the change in the supporting variables. The reweighting procedure was repeated 1000 times, where a new value of α is used each time, to build up a spread in the q^2 and supporting distributions. This ensemble of distributions can be seen in Figure 9.1.

It can be seen that the Isolation BDT is resilient to variations in q^2 while the tau flight significance does exhibit some variation. The weights that correspond to the maximum variation in q^2, ie. $\alpha = \pm 0.75$ are used to calculate ratios of the default

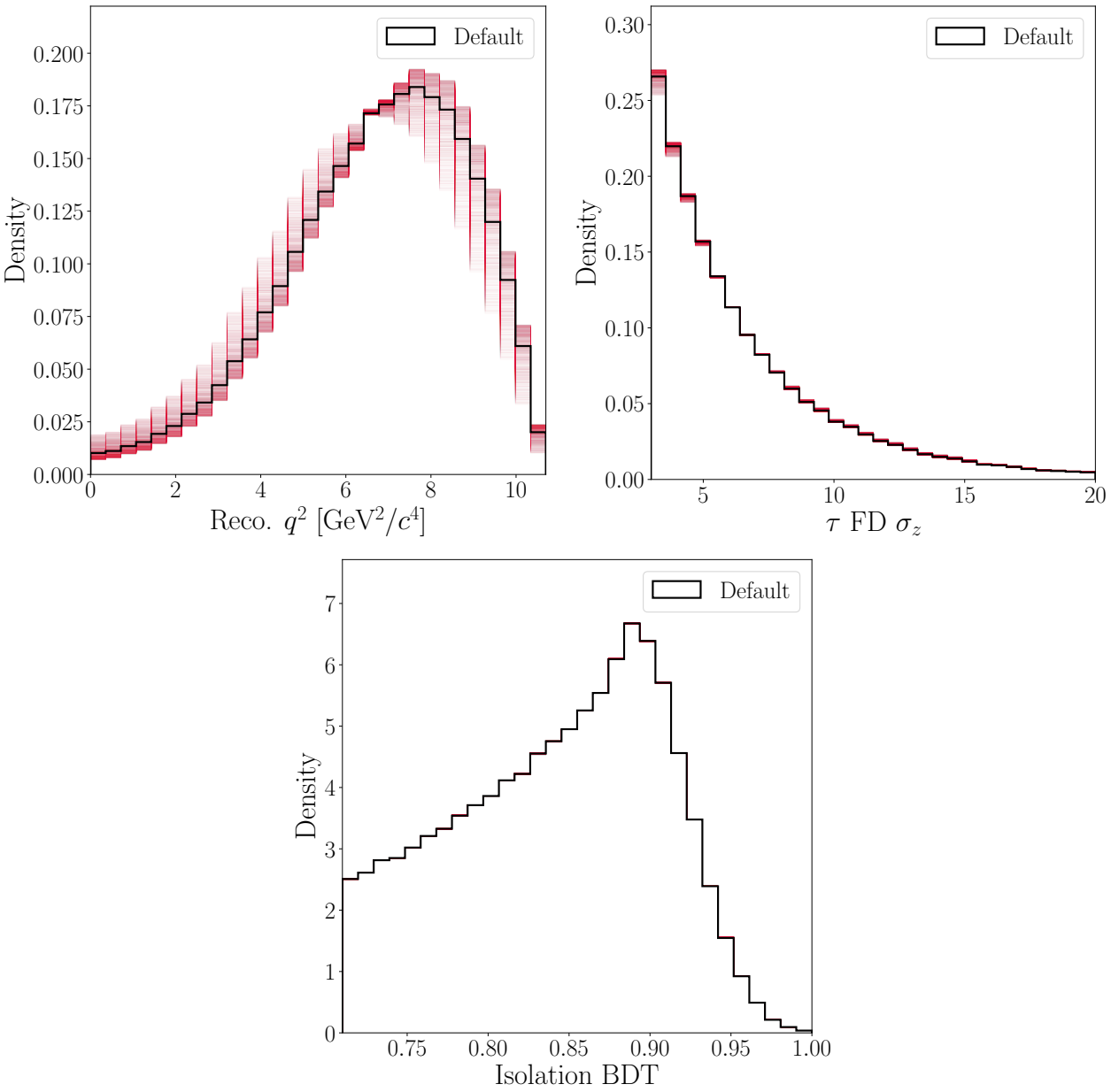


Figure 9.1: Comparison of the default signal simulation distributions for q^2 (left), tau flight significance (right) and Isolation BDT (bottom). It can be seen that the Isolation BDT distribution is resilient to variations in q^2.

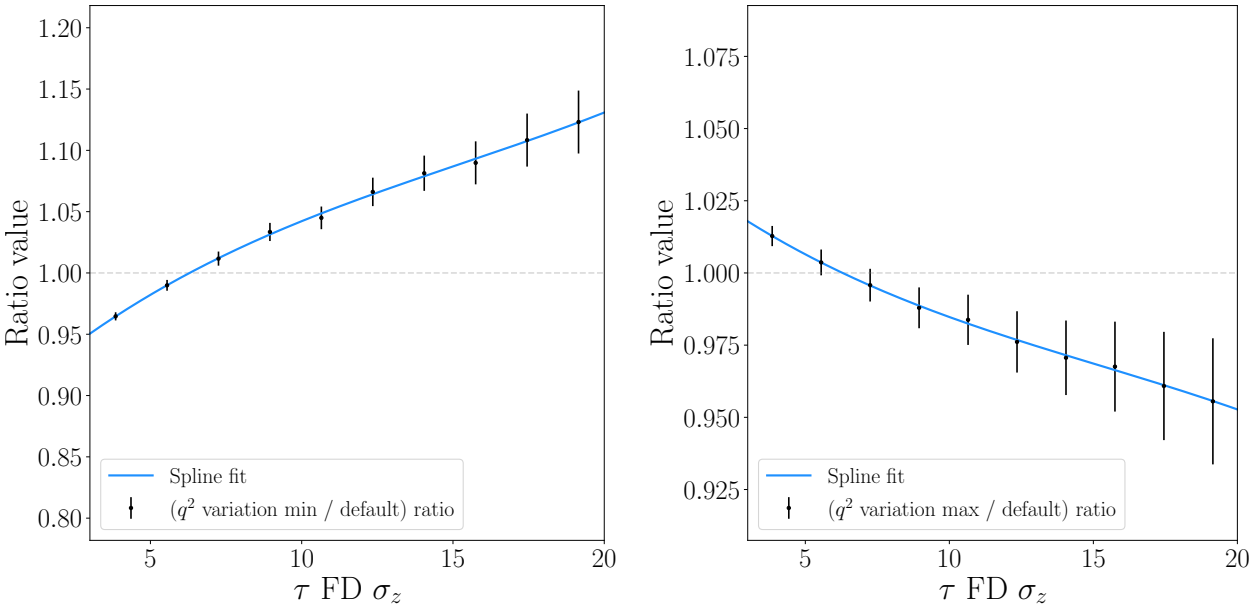


Figure 9.2: The ratio of the default signal simulation to the reweighted distributions. To obtain the reweighted distributions, values of $\alpha = -0.75$ (left) and $\alpha = 0.75$ (right) are used.

signal simulation to the reweighted cases. The ratios are found in Figure 9.2. These ratios are fitted with splines which are then used in a similar spline-based procedure described in Section 9.1.5 to evaluate the model-dependent systematic uncertainty for the measurement.

9.2 SUMMARY OF SYSTEMATIC UNCERTAINTIES

The systematic uncertainties from each individual source are summarised in Table 9.1 for each observable. The total systematic uncertainties are also given, which are calculated using a sum in quadrature over the contributing uncertainties. Finally, the statistical uncertainties for each observable are included as a reference. The correlation matrices for the statistical and systematic uncertainties on the fit parameters are given in Tables 9.2 and 9.3.

Table 9.1: Fit-related systematic uncertainties for the angular coefficients and signal fraction. The corresponding statistical uncertainties for each observable are also given for comparison.

Systematic	I_{1s}	I_{2c}	I_{2s}	I_3	I_4	I_5	I_{6c}	I_{6s}	I_7	I_8	I_9	f_{sig}
Simulation statistics	0.0294	0.1501	0.0809	0.0762	0.0782	0.0380	0.1102	0.0476	0.0381	0.0919	0.0741	0.0083
$D_s \to 3\pi X$ modelling	0.0122	0.0134	0.0276	0.0070	0.0080	0.0040	0.0291	0.0106	0.0049	0.0169	0.0070	0.0026
L0 trigger	0.0023	0.0087	0.0068	0.0027	0.0025	0.0009	0.0124	0.0013	0.0011	0.0033	0.0023	0.0011
PID	0.0004	0.0018	0.0010	0.0009	0.0008	0.0004	0.0011	0.0005	0.0003	0.0011	0.0008	0.0001
$3\pi X$ modelling	0.0019	0.0101	0.0053	0.0049	0.0047	0.0021	0.0063	0.0027	0.0019	0.0061	0.0049	0.0004
$B \to D^* D^+ X$ modelling	0.0034	0.0076	0.0089	0.0049	0.0044	0.0020	0.0057	0.0049	0.0021	0.0047	0.0038	0.0017
$B \to D^* D^0 X$ modelling	0.0013	0.0060	0.0034	0.0029	0.0029	0.0016	0.0056	0.0021	0.0016	0.0034	0.0029	0.0005
$B \to D^* D_s X$ modelling	0.0040	0.0214	0.0090	0.0075	0.0089	0.0049	0.0207	0.0073	0.0045	0.0103	0.0070	0.0010
Model dependence	0.0004	0.0024	0.0012	0.0011	0.0011	0.0004	0.0012	0.0005	0.0004	0.0014	0.0010	0.0001
Total ($\sigma_{\text{syst.}}$)	0.032	0.153	0.087	0.077	0.079	0.039	0.117	0.050	0.039	0.094	0.075	0.009
$\sigma_{\text{stat.}}$	0.066	0.251	0.159	0.141	0.141	0.072	0.191	0.090	0.077	0.163	0.137	0.013

Table 9.2: Correlation matrix of the statistical uncertainties for each observable measured in the angular fit to data.

	I_{1s}	I_{2c}	I_{2s}	I_3	I_4	I_5	I_{6c}	I_{6s}	I_7	I_8	I_9	f_{sig}
I_{1s}	1											
I_{2c}	−0.49	1										
I_{2s}	0.93	−0.46	1									
I_3	0.03	−0.04	0.14	1								
I_4	0.07	−0.09	0.07	−0.13	1							
I_5	−0.15	0.05	−0.07	0.45	0.05	1						
I_{6c}	−0.45	0.34	−0.25	0.02	0.16	−0.25	1					
I_{6s}	0.21	−0.19	0.13	−0.03	−0.21	0.11	−0.32	1				
I_7	−0.02	0.00	−0.01	0.00	−0.01	0.02	0.03	0.02	1			
I_8	0.02	−0.02	0.01	−0.04	0.04	0.01	−0.04	−0.03	−0.03	1		
I_9	0.04	−0.02	0.04	0.00	−0.02	0.00	−0.02	−0.01	0.40	−0.05	1	
f_{sig}	−0.28	0.09	−0.19	0.06	−0.10	0.04	0.21	0.09	−0.02	−0.04	−0.05	1

Table 9.3: Correlation matrix of the systematic uncertainties for each observable measured in the angular fit to data.

	I_{1s}	I_{2c}	I_{2s}	I_3	I_4	I_5	I_{6c}	I_{6s}	I_7	I_8	I_9	f_{sig}
I_{1s}	1											
I_{2c}	0.47	1										
I_{2s}	0.95	0.51	1									
I_3	−0.34	−0.15	−0.20	1								
I_4	0.76	0.46	0.77	−0.38	1							
I_5	−0.72	−0.50	−0.78	0.43	−0.68	1						
I_{6c}	−0.23	0.19	0.07	0.44	0.01	−0.21	1					
I_{6s}	−0.10	0.09	0.17	0.33	0.01	−0.23	0.82	1				
I_7	0.22	0.28	0.33	0.21	0.19	−0.24	0.40	0.33	1			
I_8	0.40	0.15	0.28	−0.41	0.32	−0.23	−0.43	−0.37	−0.16	1		
I_9	0.48	0.42	0.58	0.11	0.39	−0.44	0.32	0.28	0.65	−0.02	1	
f_{sig}	−0.63	−0.53	−0.80	0.12	−0.69	0.79	−0.47	−0.56	−0.36	−0.09	−0.54	1

10 CONCLUSION

10.1 FINAL RESULTS

The results of this work, where the q^2-integrated, rate normalised, angular coefficients, including an assessment of systematic uncertainties are:

$$\begin{aligned}
\langle I_{1s} \rangle &= 0.611 \pm 0.066\ (\text{stat.}) \pm 0.032\ (\text{syst.}) \\
\langle I_{2c} \rangle &= 0.123 \pm 0.251\ (\text{stat.}) \pm 0.153\ (\text{syst.}) \\
\langle I_{2s} \rangle &= 0.689 \pm 0.159\ (\text{stat.}) \pm 0.087\ (\text{syst.}) \\
\langle I_{3} \rangle &= -0.149 \pm 0.141\ (\text{stat.}) \pm 0.077\ (\text{syst.}) \\
\langle I_{4} \rangle &= 0.040 \pm 0.141\ (\text{stat.}) \pm 0.079\ (\text{syst.}) \\
\langle I_{5} \rangle &= -0.086 \pm 0.072\ (\text{stat.}) \pm 0.039\ (\text{syst.}) \\
\langle I_{6c} \rangle &= -0.158 \pm 0.191\ (\text{stat.}) \pm 0.117\ (\text{syst.}) \\
\langle I_{6s} \rangle &= -0.200 \pm 0.090\ (\text{stat.}) \pm 0.050\ (\text{syst.}) \\
\langle I_{7} \rangle &= 0.045 \pm 0.077\ (\text{stat.}) \pm 0.039\ (\text{syst.}) \\
\langle I_{8} \rangle &= 0.130 \pm 0.163\ (\text{stat.}) \pm 0.094\ (\text{syst.}) \\
\langle I_{9} \rangle &= 0.112 \pm 0.137\ (\text{stat.}) \pm 0.075\ (\text{syst.})
\end{aligned}$$

All angular coefficients are statistically limited at present, owing to the relatively small size and limited purity of the final angular fit dataset. These uncertainties will continue to improve in future, with considerably more data expected during Run 3 and 4 of LHCb. The dominant systematic uncertainty on all angular coefficients is due to finite simulation template statistics, which will also improve in future as larger simulated samples are required to keep up with the increase in dataset size.

A number of further angular observables may also be derived from the I_x results:

$$\begin{aligned}
\langle I_{1c} \rangle &= 0.612 \pm 0.108\ (\text{stat.}) \pm 0.060\ (\text{syst.}) \\
f_L^{D^*} &= 0.428 \pm 0.036\ (\text{stat.}) \pm 0.016\ (\text{syst.}) \\
A_{FB} &= -0.209 \pm 0.081\ (\text{stat.}) \pm 0.050\ (\text{syst.}) \\
A_C^1 &= -0.149 \pm 0.132\ (\text{stat.}) \pm 0.077\ (\text{syst.}) \\
A_C^2 &= -0.064 \pm 0.052\ (\text{stat.}) \pm 0.029\ (\text{syst.}) \\
A_C^3 &= 0.030 \pm 0.098\ (\text{stat.}) \pm 0.060\ (\text{syst.}) \\
A_T^1 &= 0.112 \pm 0.132\ (\text{stat.}) \pm 0.075\ (\text{syst.}) \\
A_T^2 &= 0.034 \pm 0.055\ (\text{stat.}) \pm 0.029\ (\text{syst.}) \\
A_T^3 &= 0.098 \pm 0.118\ (\text{stat.}) \pm 0.071\ (\text{syst.})
\end{aligned}$$

Notably, the relative precision on $f_L^{D^*}$ is found to be 9.2%, which is an improvement compared to the 15% relative uncertainty achieved by Belle [28]. A world-first measurement of the forward-background asymmetry A_{FB} is also presented, albeit with a large absolute uncertainty of 0.095, calculated from adding the statistical and systematic uncertainties in quadrature. Considerable improvements in these observables can be expected in future iterations of this analysis.

10.2 INTERPRETATION

When comparing the measured q^2 integrated values for the angular coefficients I_x, eight are compatible with the SM at, or below the 1.5σ level. I_{6c} is compatible to 2.3σ. I_{1s}, I_{2s}, and I_5 show larger discrepancies at the 3.2σ, 3.4σ, and 4.7σ level, respectively. Combined, the I_x are compatible with the SM at a level of 2.2σ. It can be seen in the correlation matrices in Tables 9.2 and 9.3 that I_{1s} and I_{2s} are highly correlated in both cases. Similarly, I_5 is anti-correlated with both I_{1s} and I_{2s}, albeit at a lower level than the correlation between I_{1s} and I_{2s}.

The derived value of $f_L^{D^*}$ is found to be compatible with the SM at a level below 1σ and compatible with the previously measured value of 0.60 ± 0.09 by Belle to within 1.7σ [28]. As previously mentioned, the world's first measurement of A_{FB} has a large relative uncertainty and is compatible with the SM at a level of 1.6σ.

The other derived terms, A_C^a and A_T^a, where $a \in [1,\ 2,\ 3]$, are asymmetry quantities which isolate individual I_x in their definitions [29]. The A_C^a asymmetries isolated the I_3, I_5, and I_4, coefficients respectively, while the A_T^a asymmetries isolate I_9, I_7, and I_8. Consequently, the A_T^a asymmetries are all predicted to be zero in the SM and the measured values are consistent with this. The A_C^a asymmetries have a slight combined tension with the SM at the level of 2.9σ, however, this tension is driven by the value of A_C^2 which isolates the I_5 coefficient which has been seen to disagree with the SM. The other two A_C^a are compatible with the SM: A_C^1 agrees at the 0.3σ level while A_C^3 is compatible with the SM at 1.2σ. The comparisons between the measured I_x, and the derived quantities with their SM predictions are shown in Figure 10.1.

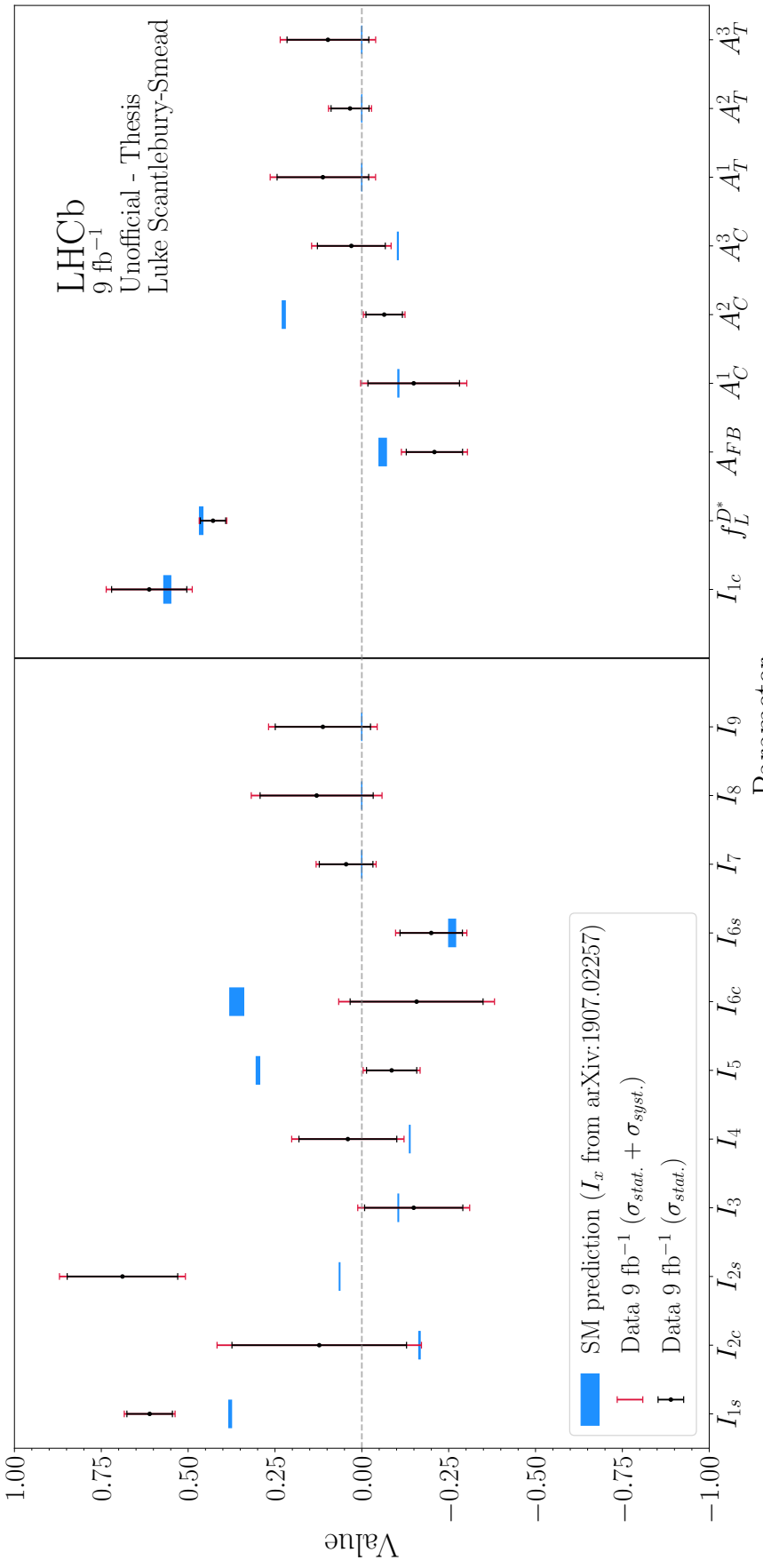


Figure 10.1: The comparison between the measured I_x, and the derived quantities with their SM predictions.

10.3 FINAL WORD

This thesis presents the world's first measurement of the q^2-integrated angular coefficients of the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay. A novel technique is employed, which uses model-independent templates that allow for the partial reconstruction of this semitauonic decay. The analysis of the angular structure is motivated by tensions with the SM of the overall decay rate.

At the time of writing, the analysis is under review by the LHCb collaboration. As part of the internal procedure the analysis is blind and the results presented here are therefore preliminary and unknown to those that continue this work towards publication. Of the twelve I_x measurements, three show marked discrepancies from the SM expectation. Whether or not these excursions persist through the final stages of the review, the implementation of the method and the techniques used to understand the background will remain the core of this novel examination of semitauonic decays.

As shown in Reference [30], an increase in statistics would improve the I_x measurements more than a naive $\sqrt{n_{\text{sig}}}$ scaling, due to the fact that the increased statistics would allow for a larger number of bins in each variable in the multidimensional fit. The recently-upgraded LHCb experiment will collect data at a much faster rate. The semitauonic dataset should more than double during Run 3 and, by the end of Run 4, LHCb will have a dataset approaching 50 fb^{-1}, which is six times the dataset that is used in the present work. With this increase in statistics, the expected statistical precision on $f_L^{D^*}$, for example, will be around 2%. Additionally, with more collected data, systematic uncertainties related to the background decays will also decrease as more precise measurements of auxiliary decays are made.

BIBLIOGRAPHY

[1] C. N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys. Rev.*, vol. 96, pp. 191–195, Oct 1954. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.96.191

[2] R. P. Feynman and M. Gell-Mann, "Theory of the Fermi Interaction," *Phys. Rev.*, vol. 109, pp. 193–198, Jan 1958. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.109.193

[3] S. L. Glashow, "Partial-symmetries of weak interactions," *Nuclear Physics*, vol. 22, no. 4, pp. 579–588, 1961. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0029558261904692

[4] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.508

[5] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.321

[6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.585

[7] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264

[8] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Non-Abelian Gauge Theories," *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, Jun 1973. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343

[9] S. Jaiswal, S. Nandi, and S. K. Patra, "Extraction of $|V_{cb}|$ from $B \to D^{(*)}\ell\nu_\ell$ and the standard model predictions of $\mathcal{R}(D^{(*)})$," *Journal of High Energy Physics*, vol. 2017, no. 12, p. 60, 12 2017. [Online]. Available: https://doi.org/10.1007/JHEP12(2017)060

[10] M. Tanaka, "Charged Higgs effects on exclusive semi-tauonic B decays," *Zeitschrift für Physik C Particles and Fields*, vol. 67, no. 2, pp. 321–326, 6 1995. [Online]. Available: https://doi.org/10.1007/BF01571294

[11] W. Buchmüller, R. Rückl, and D. Wyler, "Leptoquarks in lepton-quark collisions," *Physics Letters B*, vol. 191, no. 4, pp. 442 – 448, 1987. [Online]. Available: http://www.sciencedirect.com/science/article/pii/037026938790637X

[12] P. Asadi and D. Shih, "Maximizing the impact of new physics in $b \to c\tau\nu$ anomalies," *Phys. Rev. D*, vol. 100, p. 115013, Dec 2019. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.100.115013

[13] A. Matyja *et al.*, "Observation of $B^0 \to D^{*-}\tau^+\nu_\tau$ decay at Belle," *Phys. Rev. Lett.*, vol. 99, p. 191807, 11 2007. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.99.191807

[14] B. Aubert *et al.*, "Observation of the semileptonic decays $B \to D^*\tau^-\overline{\nu}_\tau$ and evidence for $B \to D\tau^-\overline{\nu}_\tau$," *Phys. Rev. Lett.*, vol. 100, p. 021801, 1 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.100.021801

[15] A. Bozek *et al.*, "Observation of $B^+ \to \overline{D}^{*0}\tau^+\nu_\tau$ and evidence for $B^+ \to \overline{D}^0\tau^+\nu_\tau$ at Belle," *Phys. Rev. D*, vol. 82, p. 072005, 10 2010. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.82.072005

[16] J. P. Lees *et al.*, "Measurement of an excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ decays and implications for charged higgs bosons," *Phys. Rev. D*, vol. 88, p. 072012, 10 2013. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.88.072012

[17] M. Huschle *et al.*, "Measurement of the branching ratio of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ relative to $\overline{B} \to D^{(*)}\ell^-\overline{\nu}_\ell$ decays with hadronic tagging at Belle," *Phys. Rev. D*, vol. 92, p. 072014, 10 2015. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.92.072014

[18] R. Aaij *et al.*, "Measurement of the ratio of branching fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$," *Phys. Rev. Lett.*, vol. 115, p. 111803, 9 2015. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.115.111803

[19] S. Hirose *et al.*, "Measurement of the τ lepton polarization and $\mathcal{R}(D^*)$ in the decay $\overline{B} \to D^{*+}\tau^-\overline{\nu}_\tau$ with one-prong hadronic τ decays at Belle," *Phys. Rev. D*, vol. 97, p. 012004, 1 2018. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.97.012004

[20] R. Aaij *et al.*, "Test of lepton flavor universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays," *Phys. Rev. D*, vol. 97, p. 072013, 4 2018. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.97.072013

[21] G. Caria *et al.*, "Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a Semileptonic Tagging Method," *Phys. Rev. Lett.*, vol. 124, p. 161803, Apr 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.124.161803

[22] R. Aaij *et al.*, "New Test of Lepton Universality using the First Simultaneous Measurement of $R(D)$ and $R(D^*)$ Observables at LHCb," https://lhcb-outreach.web.cern.ch/2022/10/18/new-test-of-lepton-universality-using-the-first-simultaneous-measurement-of-rd-and-rd-observables-at-lhcb/, October 2022.

[23] R. Aaij *et al.*, "Test of Lepton Flavour Universality using a Measurement of $R(D)$ with Hadronic τ Decays," https://lhcb-outreach.web.cern.ch/2023/03/21/test-of-lepton-flavour-universality-using-a-measurement-of-rd-with-hadronic-%cf%84-decays/, March 2023.

[24] Y. Amhis *et al.*, "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021," 6 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2206.07501

[25] D. Bardhan, P. Byakti, and D. Ghosh, "A closer look at the R_D and R_{D^*} anomalies," *Journal of High Energy Physics*, vol. 2017, no. 1, p. 125, Jan 2017. [Online]. Available: https://doi.org/10.1007/JHEP01(2017)125

[26] R. Alonso, J. M. Camalich, and S. Westhoff, "Tau properties in $B \to D\tau\nu$ from visible final-state kinematics," *Physical Review D*, 2017. [Online]. Available: https://doi.org/10.1103/PhysRevD.95.093006

[27] P. Asadi, M. R. Buckley, and D. Shih, "Asymmetry observables and the origin of $R_{D^{(*)}}$ anomalies," *Physical Review D*, 2018. [Online]. Available: https://doi.org/10.1103/PhysRevD.99.035015

[28] A. Abdesselam *et al.*, "Measurement of the D^{*-} polarization in the decay $B^0 \to D^{*-}\tau^+\nu_\tau$," 3 2019. [Online]. Available: https://doi.org/10.48550/arXiv.1903.03102

[29] D. Becirevic, M. Fedele, I. Nisandzic, and A. Tayduganov, "Lepton flavor universality tests through angular observables of $\overline{B} \to D^{(*)}\ell\overline{\nu}$ decay modes," 2019. [Online]. Available: https://doi.org/10.48550/arXiv.1907.02257

[30] D. Hill, M. John, W. Ke, and A. Poluektov, "Model-independent method for measuring the angular coefficients of $B^0 \to D^{*-}\tau^+\nu_\tau$ decays," *JHEP*, vol. 11, p. 133, 2019. [Online]. Available: https://doi.org/10.1007/JHEP11(2019)133

[31] G. Aad *et al.*, "The ATLAS experiment at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08003

[32] K. Aamodt *et al.*, "The ALICE experiment at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08002, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08002

[33] S. Chatrchyan *et al.*, "The CMS experiment at theCERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08004, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08004

[34] A. A. A. Jr *et al.*, "The LHCb Detector at the LHC," *Journal of Instrumentation*, vol. 3, no. 08, p. S08005, aug 2008. [Online]. Available: https://dx.doi.org/10.1088/1748-0221/3/08/S08005

[35] F. Follin and D. Jacquet, "Implementation and experience with luminosity levelling with offset beam," p. 5, October 2014. [Online]. Available: http://cds.cern.ch/record/1955354

[36] R. Lindner, "LHCb layout_2. LHCb schema_2," 2 2008, LHCb Collaboration. [Online]. Available: http://cds.cern.ch/record/1087860

[37] P. R. Barbosa-Marinho *et al.*, "LHCb VELO (VErtex LOcator): Technical Design Report," Geneva, 2001. [Online]. Available: http://cds.cern.ch/record/504321

[38] R. Aaij *et al.*, "Performance of the LHCb Vertex Locator," *Journal of Instrumentation*, vol. 9, no. 09, p. P09007, 2014. [Online]. Available: http://stacks.iop.org/1748-0221/9/i=09/a=P09007

[39] A. A. A. Jr *et al.*, "The LHCb detector at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08005, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08005

[40] S. Amato *et al.*, "LHCb magnet: Technical Design Report," Geneva, 2000. [Online]. Available: http://cds.cern.ch/record/424338

[41] R. Arink *et al.*, "Performance of the LHCb Outer Tracker," *Journal of Instrumentation*, vol. 9, no. 01, p. P01002, 2014. [Online]. Available: http://stacks.iop.org/1748-0221/9/i=01/a=P01002

[42] P. d'Argent *et al.*, "Improved performance of the LHCb outer tracker in LHC run 2," *Journal of Instrumentation*, vol. 12, no. 11, pp. P11 016–P11 016, nov 2017. [Online]. Available: https://doi.org/10.1088%2F1748-0221%2F12%2F11%2Fp11016

[43] M. Adinolfi *et al.*, "Performance of the LHCb RICH detector at the LHC," *The European Physical Journal C*, vol. 73, no. 5, p. 2431, 5 2013. [Online]. Available: https://doi.org/10.1140/epjc/s10052-013-2431-9

[44] R. Calabrese *et al.*, "Performance of the LHCb RICH detectors during LHC run 2," *Journal of Instrumentation*, vol. 17, no. 07, p. P07013, jul 2022. [Online]. Available: https://doi.org/10.1088%2F1748-0221%2F17%2F07%2Fp07013

[45] S. Amato *et al.*, "LHCb RICH: Technical Design Report," Geneva, 2000. [Online]. Available: http://cds.cern.ch/record/494263

[46] R. L. Workman *et al.*, "Review of Particle Physics," *Progress of Theoretical and Experimental Physics*, vol. 2022, no. 8, 08 2022, 083C01. [Online]. Available: https://doi.org/10.1093/ptep/ptac097

[47] S. Amato *et al.*, "LHCb calorimeters: Technical Design Report," Geneva, 2000. [Online]. Available: https://cds.cern.ch/record/494264

[48] C. A. Beteta *et al.*, "Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC," 2020. [Online]. Available: https://doi.org/10.48550/arXiv.2008.11556

[49] A. A. A. Jr *et al.*, "Performance of the LHCb muon system," *Journal of Instrumentation*, vol. 8, no. 02, p. P02022, 2013. [Online]. Available: http://stacks.iop.org/1748-0221/8/i=02/a=P02022

[50] M. Pili, "Performance of the LHCb detector in the Run 2," *PoS*, vol. ICHEP2020, p. 767, 2021. [Online]. Available: https://cds.cern.ch/record/2783948

[51] R. Antunes-Nobrega *et al.*, "LHCb trigger system: Technical Design Report," Geneva, 2003, revised version number 1 submitted on 2003-09-24 12:12:22. [Online]. Available: http://cds.cern.ch/record/630828

[52] R. Aaij *et al.*, "The LHCb trigger and its performance in 2011," *Journal of Instrumentation*, vol. 8, no. 04, p. P04022, 2013. [Online]. Available: http://stacks.iop.org/1748-0221/8/i=04/a=P04022

[53] R. Aaij *et al.*, "Design and performance of the LHCb trigger and full real-time reconstruction in run 2 of the LHC," *Journal of Instrumentation*, vol. 14, no. 04, pp. P04 013–P04 013, apr 2019. [Online]. Available: https://doi.org/10.1088%2F1748-0221%2F14%2F04%2Fp04013

[54] M. Thompson, *Modern Particle Physics*, 2nd ed. Cambridge University Press, 2013.

[55] M. Neubert, *Heavy-Quark Effective Theory*. Springer, 1993. [Online]. Available: https://doi.org/10.1007/978-3-662-02993-3

[56] S. Fajfer, J. F. Kamenik, I. Niandi, and A. Prasov, "Angular distributions of $B^0 \to D^*\tau\nu$ decays and search of New Physics," *Phys. Rev.*, vol. D85, p. 094025, 2012. [Online]. Available: https://doi.org/10.1103/PhysRevD.85.094025

[57] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.10.531

[58] K. Hagiwara, A. Martin, and M. Wade, "Exclusive semileptonic B-meson decays," *Nuclear Physics B*, vol. 327, no. 3, pp. 569–594, 1989. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0550321389903064

[59] C. G. Boyd, B. Grinstein, and R. F. Lebed, "Constraints on form factors for exclusive semileptonic heavy to light meson decays," *Phys. Rev. Lett.*, vol. 74, pp. 4603–4606, Jun 1995. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.74.4603

[60] I. Caprini, L. Lellouch, and M. Neubert, "Dispersive bounds on the shape of $B \to D^{(*)}\ell\nu$ form factors," *Nuclear Physics B*, vol. 530, no. 1, pp. 153–181, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0550321398003502

[61] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, "Combined analysis of semileptonic B decays to D and D^*: $R(D^{(*)})$, $|V_{cb}|$, and new physics," *Phys. Rev. D*, vol. 95, p. 115008, Jun 2017. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.95.115008

[62] M. Neubert, "Heavy-quark symmetry," *Physics Reports*, vol. 245, no. 5, pp. 259–395, 1994. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370157394900914

[63] A. V. Manohar and M. B. Wise, *Heavy Quark Physics*, ser. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2000.

[64] A. K. Alok, A. Datta, A. Dighe, M. Duraisamy, D. Ghosh, and D. London, "New physics in $b \to s\mu^+\mu^-$: CP-violating observables," *Journal of High Energy Physics*, vol. 2011, no. 11, p. 122, Nov 2011. [Online]. Available: https://doi.org/10.1007/JHEP11(2011)122

[65] B. Aubert *et al.*, "Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$," *Phys. Rev. D*, vol. 73, p. 092001, May 2006. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.73.092001

[66] T. Aaltonen *et al.*, "Measurements of the Angular Distributions in the Decays $B \to K^{(*)}\mu^+\mu^-$ at CDF," *Phys. Rev. Lett.*, vol. 108, p. 081807, Feb 2012. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.108.081807

[67] R. Aaij *et al.*, "Differential branching fractions and isospin asymmetries of B $\to K^{(*)}\mu^+\mu^-$ decays," *Journal of High Energy Physics*, vol. 2014, no. 6, p. 133, Jun 2014. [Online]. Available: https://doi.org/10.1007/JHEP06(2014)133

[68] M. Aaboud *et al.*, "Angular analysis of $B^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Journal of High Energy Physics*, vol. 2018, no. 10, p. 47, Oct 2018. [Online]. Available: https://doi.org/10.1007/JHEP10(2018)047

[69] A. Sirunyan *et al.*, "Measurement of angular parameters from the decay $B^0 \to K^0\mu^+\mu$ in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Physics Letters B*, vol. 781, pp. 517–541, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370269318303149

[70] S. Wehle *et al.*, "Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$," *Phys. Rev. Lett.*, vol. 118, p. 111801, 3 2017. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.118.111801

[71] R. Aaij *et al.*, "Erratum to: Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu$ decays," *Journal of High Energy Physics*, vol. 2018, no. 9, p. 145, Sep 2018. [Online]. Available: https://doi.org/10.1007/JHEP09(2018)145

[72] R. Aaij *et al.*, "Angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay," *Phys. Rev. Lett.*, vol. 126, p. 161802, Apr 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.126.161802

[73] R. Aaij *et al.*, "Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$," *Journal of High Energy Physics*, 2013. [Online]. Available: https://doi.org/10.1007/JHEP08(2013)131

[74] A. Maier *et al.*, "User analysis of LHCb data with Ganga," *Journal of Physics: Conference Series*, vol. 219, no. 7, p. 072008, apr 2010. [Online]. Available: https://dx.doi.org/10.1088/1742-6596/219/7/072008

[75] D. Müller, M. Clemencic, G. Corti, and M. Gersabeck, "Redecay: a novel approach to speed up the simulation at lhcb," *The European Physical Journal C*, vol. 78, no. 12, p. 1009, Dec 2018. [Online]. Available: https://doi.org/10.1140/epjc/s10052-018-6469-6

[76] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011. [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html

[77] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, 1988. [Online]. Available: https://doi.org/10.4324/9780203771587

[78] D. Cervenkov, "First performance of the real-time reconstruction at LHCb," *PoS*, vol. ICHEP2022, p. 686, 2022.

[79] L. Anderlini, A. Contu, C. R. Jones, S. S. Malde, D. Muller, S. Ogilvy, J. M. Otalora Goicochea, A. Pearce, I. Polyakov, W. Qian, B. Sciascia, R. Vazquez Gomez, and Y. Zhang, "The PIDCalib package," CERN, Geneva, Tech. Rep., 2016. [Online]. Available: http://cds.cern.ch/record/2202412

[80] M. Pivk and F. R. Le Diberder, "SPlot: A Statistical tool to unfold data distributions," *Nucl. Instrum. Meth. A*, vol. 555, pp. 356–369, 2005. [Online]. Available: https://doi.org/10.1016/j.nima.2005.08.106

[81] R. Aaij *et al.*, "Angular analysis of $B^0 \to D^{*-}D_s^{*+}$ decays with $D_s^{*+} \to D_s^+\gamma$ decays," *Journal of High Energy Physics*, vol. 2021, no. 6, p. 177, Jun 2021. [Online]. Available: https://doi.org/10.1007/JHEP06(2021)177

[82] D. Hill and P. Gandini, "Probability density functions for describing partially reconstructed B-hadron decays," CERN, Geneva, Tech. Rep., 2020. [Online]. Available: https://cds.cern.ch/record/2740994

[83] J. Eschle, A. Puig Navarro, R. Silva Coutinho, and N. Serra, "zfit: Scalable pythonic fitting," *SoftwareX*, vol. 11, Jan 2020. [Online]. Available: https://doi.org/10.1016/j.softx.2020.100508

[84] F. James and M. Roos, "MINUIT: a system for function minimization and analysis of the parameter errors and corrections," *Comput. Phys. Commun.*, vol. 10, no. 6, pp. 343–67, 1975. [Online]. Available: https://cds.cern.ch/record/310399

[85] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Methods Phys. Res.*, vol. 462, no. 1, pp. 152–155, 2001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900201000894

[86] D. Scora and N. Isgur, "Semileptonic meson decays in the quark model: An update," *Phys. Rev. D*, vol. 52, pp. 2783–2812, Sep 1995. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.52.2783

[87] K. Abe *et al.*, "Measurement of $\mathcal{B}(\overline{B}^0 \to D^+\ell^-\overline{\nu})$ and determination of $|V_{cb}|$," *Physics Letters B*, vol. 526, no. 3, pp. 258–268, 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370269301014836

[88] F. U. Bernlochner, M. F. Sevilla, D. J. Robinson, and G. Wormser, "Semitauonic b-hadron decays: A lepton flavor universality laboratory," *Rev. Mod. Phys.*, vol. 94, no. 1, p. 015003, 2022. [Online]. Available: https://doi.org/10.1103/RevModPhys.94.015003

[89] R. Aaij *et al.*, "Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B^0_s decay branching fractions," *Phys. Rev. D*, vol. 104, p. 032005, Aug 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.104.032005

[90] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, ser. Monographs on Statistics and Applied Probability. Boca Raton, Florida, USA: Chapman & Hall/CRC, 1993, no. 57.

Part IV

APPENDIX

A HELICITY AMPLITUDES

A.1 LEPTONIC HELICITY AMPLITUDES

The leptonic helicity amplitudes are defined for each of the coupling types [29],

$$\begin{aligned} L^{\lambda_\ell}_{V-A,\lambda}\left(q^2,\chi,\theta_\ell\right) &= \eta_\mu(\lambda)\left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\gamma^\mu\left(1-\gamma_5\right)\nu\right|0\right\rangle \\ L^{\lambda_\ell}_{S-P}\left(q^2,\chi,\theta_\ell\right) &= \left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\left(1-\gamma_5\right)\nu\right|0\right\rangle \\ L^{\lambda_\ell}_{T-T5,\lambda\lambda'}\left(q^2,\chi,\theta_\ell\right) &= -L^{\lambda_\ell}_{T-T5,\lambda'\lambda} = -i\eta_\mu(\lambda)\eta_\nu\left(\lambda'\right)\left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}^{\mu\nu}\left(1-\gamma_5\right)\nu\right|0\right\rangle \end{aligned} \tag{A.1}$$

Using polarisation vectors, Dirac spinors, and the Weyl representation of the gamma-matrices, the explicit formulas for the leptonic helicity amplitudes can be found. The vector-type amplitudes are

$$\begin{aligned} L^{+}_{V-A,+}\left(q^2,\chi,\theta_\ell\right) &= \pm\sqrt{2}m_\ell\beta_\ell\sin\theta_\ell e^{-2i\chi} \\ L^{+}_{V-A,-}\left(q^2,\chi,\theta_\ell\right) &= \pm\sqrt{2}m_\ell\beta_\ell\sin\theta_\ell \\ L^{+}_{V-A,0}\left(q^2,\chi,\theta_\ell\right) &= 2m_\ell\beta_\ell\cos\theta_\ell e^{-i\chi} \\ L^{+}_{V-A,t}\left(q^2,\chi,\theta_\ell\right) &= -2m_\ell\beta_\ell e^{-i\chi} \\ L^{-}_{V-A,\pm}\left(q^2,\chi,\theta_\ell\right) &= \sqrt{2q^2}\beta_\ell\left(1\pm\cos\theta_\ell\right)e^{\mp i\chi} \\ L^{-}_{V-A,0}\left(q^2,\chi,\theta_\ell\right) &= -2\sqrt{q^2}\beta_\ell\sin\theta_\ell \\ L^{-}_{V-A,t}\left(q^2,\chi,\theta_\ell\right) &= 0, \end{aligned} \tag{A.2}$$

the scalar-type amplitudes are,

$$\begin{aligned} L^{+}_{S-P}\left(q^2,\chi,\theta_\ell\right) &= -2\sqrt{q^2}\beta_\ell e^{-i\chi} \\ L^{-}_{S-P}\left(q^2,\chi,\theta_\ell\right) &= 0, \end{aligned} \tag{A.3}$$

and the tensor-type amplitudes are,

$$
\begin{aligned}
L^{+}_{T-T5,+0}\left(q^2,\chi,\theta_\ell\right) &= \sqrt{2q^2}\beta_\ell \sin\theta_\ell e^{-2i\chi} \\
L^{+}_{T-T5,-0}\left(q^2,\chi,\theta_\ell\right) &= \sqrt{2q^2}\beta_\ell \sin\theta_\ell \\
L^{+}_{T-T5,+-}\left(q^2,\chi,\theta_\ell\right) &= -L^{+}_{T-T5,0t} = 2\sqrt{q^2}\beta_\ell \cos\theta_\ell e^{-i\chi} \\
L^{+}_{T-T5,+t}\left(q^2,\chi,\theta_\ell\right) &= \mp\sqrt{2q^2}\beta_\ell \sin\theta_\ell e^{-2i\chi} \\
L^{+}_{T-T5,-t}\left(q^2,\chi,\theta_\ell\right) &= \mp\sqrt{2q^2}\beta_\ell \sin\theta_\ell \\
L^{-}_{T-T5,\pm 0}\left(q^2,\chi,\theta_\ell\right) &= \pm\sqrt{2}m_\ell\beta_\ell \left(1 \pm \cos\theta_\ell\right) e^{\mp i\chi} \\
L^{-}_{T-T5,+-}\left(q^2,\chi,\theta_\ell\right) &= -L^{-}_{T-T5,0t} = -2m_\ell\beta_\ell \sin\theta_\ell \\
L^{-}_{T-T5,\pm t}\left(q^2,\chi,\theta_\ell\right) &= -\sqrt{2}m_\ell\beta_\ell \left(1 \pm \cos\theta_\ell\right) e^{\mp i\chi}.
\end{aligned}
\tag{A.4}
$$

A.2 HADRONIC HELICITY AMPLITUDES

The hadronic helicity amplitudes are defined for each coupling type [29],

$$
\begin{aligned}
H^{\lambda_M}_{V(A),\lambda}\left(q^2\right) &= \left(g_{V(A)} \pm 1\right) \eta^*_\mu(\lambda) \left\langle M\left(\lambda_M\right) \left|\bar{c}\gamma^\mu\left(\gamma_5\right) b\right| \bar{B}\right\rangle, \\
H_{S(P),\lambda}\left(q^2\right) &= g_{S(P)} \left\langle M\left(\lambda_M = 0\right) \left|\bar{c}\left(\gamma_5\right) b\right| \bar{B}\right\rangle \\
H^{\lambda_M}_{T(T5),\lambda\lambda'}\left(q^2\right) &= -H^{\lambda_M}_{T,\lambda'\lambda}\left(q^2\right) = i g_{T(T5)} \eta^*_\mu(\lambda) \eta^*_\nu\left(\lambda'\right) \left\langle M\left(\lambda_M\right) \left|\bar{c}\sigma^{\mu\nu}\left(\gamma_5\right) b\right| \bar{B}\right\rangle.
\end{aligned}
\tag{A.5}
$$

which can be written out explicitly in terms of the coupling strength g_X, the masses of the particles, q^2, and the form factors of each coupling type,

$$
\begin{aligned}
H^{\pm}_{V,\pm}\left(q^2\right) &= \mp\left(g_V + 1\right) \frac{\sqrt{\lambda(m_B^2,\ m_{D^*}^2,\ q^2)}}{m_B + m_{D^*}} V\left(q^2\right), \\
H^{\pm}_{A,\pm}\left(q^2\right) &= -\left(g_A - 1\right)\left(m_B + m_{D^*}\right) A_1\left(q^2\right), \\
H_0\left(q^2\right) \equiv H^0_{A,0}\left(q^2\right) &= \left(g_A - 1\right) \frac{8 m_B m_{D^*}}{\sqrt{q^2}} A_{12}\left(q^2\right), \\
H_t\left(q^2\right) \equiv H^0_{A,t}\left(q^2\right) &= \left(g_A - 1\right) \sqrt{\frac{\lambda(m_B^2,\ m_{D^*}^2,\ q^2)}{q^2}} A_0\left(q^2\right), \\
H_P\left(q^2\right) &= -g_P \frac{\sqrt{\lambda(m_B^2,\ m_{D^*}^2,\ q^2)}}{m_b + m_c} A_0\left(q^2\right) \\
H^{\pm}_{T,\pm 0}\left(q^2\right) &= \mp\frac{g_T}{g_{T5}} H^{\pm}_{T5,\pm t}\left(q^2\right) = \pm g_T \frac{m_B^2 - m_{D^*}^2}{\sqrt{q^2}} T_2\left(q^2\right), \\
H^{\pm}_{T,\pm t}\left(q^2\right) &= \mp\frac{g_T}{g_{T5}} H^{\pm}_{T5,\pm 0}\left(q^2\right) = \pm g_T \sqrt{\frac{\lambda(m_B^2,\ m_{D^*}^2,\ q^2)}{q^2}} T_1\left(q^2\right), \\
H^0_{T,+-}\left(q^2\right) &= -\frac{g_T}{g_{T5}} H^0_{T5,0t}\left(q^2\right) = -g_T \frac{4 m_B m_{D^*}}{m_B + m_{D^*}} T_{23}\left(q^2\right), \\
H^0_{V,0}\left(q^2\right) &= H^0_{V,t}\left(q^2\right) = H_S\left(q^2\right) = H^0_{T,0t}\left(q^2\right) = H^0_{T5,+-}\left(q^2\right) = 0.
\end{aligned}
\tag{A.6}
$$

These hadronic amplitudes can be redefined to combine the scalar and tensor type amplitudes with the vector type amplitudes,

$$\begin{aligned}
\widetilde{H}_{V/A,t}^{\lambda_M(=0)} &\equiv H_{V/A,t}^{0} + \frac{\sqrt{q^2}}{m_\ell} H_{S/P} \\
\widetilde{H}_{\pm}^{\lambda_M,+} &\equiv H_{V,\pm}^{\lambda_M} + H_{A,\pm}^{\lambda_M} - 2\frac{\sqrt{q^2}}{m_\ell}\left(H_{T,\pm t}^{\lambda_M} \pm H_{T,\pm 0}^{\lambda_M} + H_{T5,\pm t}^{\lambda_M} \pm H_{T5,\pm 0}^{\lambda_M}\right), \\
\widetilde{H}_{0}^{\lambda_M,+} &\equiv H_{V,0}^{\lambda_M} + H_{A,0}^{\lambda_M} - 2\frac{\sqrt{q^2}}{m_\ell}\left(H_{T,+-}^{\lambda_M} + H_{T,0t}^{\lambda_M} + H_{T5,+-}^{\lambda_M} + H_{T5,0t}^{\lambda_M}\right), \\
\widetilde{H}_{\pm}^{\lambda_M,-} &\equiv H_{V,\pm}^{\lambda_M} + H_{A,\pm}^{\lambda_M} - 2\frac{m_\ell}{\sqrt{q^2}}\left(H_{T,\pm t}^{\lambda_M} \pm H_{T,\pm 0}^{\lambda_M} + H_{T5,\pm t}^{\lambda_M} \pm H_{T5,\pm 0}^{\lambda_M}\right), \\
\widetilde{H}_{0}^{\lambda_M,-} &\equiv H_{V,0}^{\lambda_M} + H_{A,0}^{\lambda_M} - 2\frac{m_\ell}{\sqrt{q^2}}\left(H_{T,+-}^{\lambda_M} + H_{T,0t}^{\lambda_M} + H_{T5,+-}^{\lambda_M} + H_{T5,0t}^{\lambda_M}\right), \\
\widetilde{H}_{t}^{\lambda_M} &\equiv \widetilde{H}_{V,t}^{\lambda_M} + \widetilde{H}_{A,t}^{\lambda_M}.
\end{aligned} \tag{A.7}$$

Finally, additional linear combinations can be created in order to write the decay rate in a more compact form,

$$\begin{aligned}
\widetilde{H}_{\pm}^{+}\left(q^2\right) &\equiv H_{\pm}\left(q^2\right) - 2\frac{\sqrt{q^2}}{m_\ell} H_{T,\pm}\left(q^2\right) \\
\widetilde{H}_{0}^{+}\left(q^2\right) &\equiv H_{0}\left(q^2\right) - 2\frac{\sqrt{q^2}}{m_\ell} H_{T,0}\left(q^2\right), \\
\widetilde{H}_{\pm}^{-}\left(q^2\right) &\equiv H_{\pm}\left(q^2\right) - 2\frac{m_\ell}{\sqrt{q^2}} H_{T,\pm}\left(q^2\right), \\
\widetilde{H}_{0}^{-}\left(q^2\right) &\equiv H_{0}\left(q^2\right) - 2\frac{m_\ell}{\sqrt{q^2}} H_{T,0}\left(q^2\right), \\
\widetilde{H}_{t}\left(q^2\right) &\equiv H_{t}\left(q^2\right) + \frac{\sqrt{q^2}}{m_\ell^2} H_{P}\left(q^2\right),
\end{aligned} \tag{A.8}$$

with

$$\begin{aligned}
H_{\pm}\left(q^2\right) &\equiv H_{V,\pm}^{\pm}\left(q^2\right) + H_{A,\pm}^{\pm}\left(q^2\right), \\
H_{T,\pm}\left(q^2\right) &\equiv H_{T,\pm t}^{\pm}\left(q^2\right) \pm H_{T,\pm 0}^{\pm}\left(q^2\right) + H_{T5,\pm t}^{\pm}\left(q^2\right) \pm H_{T5,\pm 0}^{\pm}\left(q^2\right), \\
H_{T,0}\left(q^2\right) &\equiv H_{T,+-}^{0}\left(q^2\right) + H_{T5,0t}^{0}\left(q^2\right).
\end{aligned} \tag{A.9}$$

This allows the angular coefficient functions I_x defined below in Appendix A.3 to be expressed in a significantly compact form.

A.3 ANGULAR COEFFICIENT DEFINITIONS

The definitions of the q^2 dependent angular coefficients $I_x(q^2)$ are as follows [29]:

$$I_{1c}(q^2) = 2N\left[|\tilde{H}_0^-|^2 + \frac{m_\ell^2}{q^2}|\tilde{H}_0^+|^2 + 2\frac{m_\ell^2}{q^2}|\tilde{H}_t|^2\right] \tag{A.10}$$

$$I_{1s}(q^2) = \frac{N}{2}\left[3\left(|\tilde{H}_+^-|^2 + |\tilde{H}_-^-|^2\right) + \frac{m_\ell^2}{q^2}\left(|\tilde{H}_+^+|^2 + |\tilde{H}_-^+|^2\right)\right] \tag{A.11}$$

$$I_{2c}(q^2) = 2N\left[-|\tilde{H}_0^-|^2 + \frac{m_\ell^2}{q^2}|\tilde{H}_0^+|^2\right] \tag{A.12}$$

$$I_{2s}(q^2) = \frac{N}{2}\left[|\tilde{H}_+^-|^2 + |\tilde{H}_-^-|^2 - \frac{m_\ell^2}{q^2}\left(|\tilde{H}_+^+|^2 + |\tilde{H}_-^+|^2\right)\right] \tag{A.13}$$

$$I_3(q^2) = -2N\mathcal{R}e\left[\tilde{H}_+^-\tilde{H}_-^{-*} - \frac{m_\ell^2}{q^2}\tilde{H}_+^+\tilde{H}_-^{+*}\right] \tag{A.14}$$

$$I_4(q^2) = N\mathcal{R}e\left[\left(\tilde{H}_+^- + \tilde{H}_-^-\right)\tilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\tilde{H}_+^+ + \tilde{H}_-^+\right)\tilde{H}_0^{+*}\right] \tag{A.15}$$

$$I_5(q^2) = 2N\mathcal{R}e\left[\left(\tilde{H}_+^- - \tilde{H}_-^-\right)\tilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\tilde{H}_+^+ + \tilde{H}_-^+\right)\tilde{H}_t^*\right] \tag{A.16}$$

$$I_{6c}(q^2) = 8N\frac{m_\ell^2}{q^2}\mathcal{R}e\left[\tilde{H}_0^+\tilde{H}_t^*\right] \tag{A.17}$$

$$I_{6s}(q^2) = 2N\left[|\tilde{H}_+^-|^2 - |\tilde{H}_-^-|^2\right] \tag{A.18}$$

$$I_7(q^2) = 2N\mathcal{I}m\left[\left(\tilde{H}_+^- + \tilde{H}_-^-\right)\tilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\tilde{H}_+^+ - \tilde{H}_-^+\right)\tilde{H}_t^*\right] \tag{A.19}$$

$$I_8(q^2) = N\mathcal{I}m\left[\left(\tilde{H}_+^- - \tilde{H}_-^-\right)\tilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\tilde{H}_+^+ - \tilde{H}_-^+\right)\tilde{H}_0^{+*}\right] \tag{A.20}$$

$$I_9(q^2) = 2N\mathcal{I}m\left[\tilde{H}_+^-\tilde{H}_-^{-*} - \frac{m_\ell^2}{q^2}\tilde{H}_+^+\tilde{H}_-^{+*}\right] \tag{A.21}$$

$$\tag{A.22}$$

where

$$N \equiv N(q^2) = \mathcal{B}(D^* \to D\pi)\frac{G_F^2|V_{cb}|^2}{48(2\pi)^3 m_B^3}q^2\sqrt{\lambda(m_B^2,\ m_{D^*}^2,\ q^2)}\left(1 - \frac{m_\ell^2}{q^2}\right) \tag{A.23}$$

and $H_\lambda^{\lambda_\ell}$ are linear combinations of the hadronic helicity amplitudes, defined in Appendix A.2 and discussed in detail in Reference [29].

B SUBDECAY CATEGORIES

The definitions of the subcategories used in the D_s PDF are given in Table B.1. The first column shows the label of the subcategory while the second column shows the list of D_s decays that make up the subcategory. The third column shows the native fraction of the subcategory in the simulation. The final column shows the relative amounts of each D_s decay inside its subcategory. The relative amounts in each subcategory are determined by the simulation generation.

Table B.1: The definition of the subcategories used in the D_s PDF.

Group	Decay	Simulation Fraction (%)	Split (%)
$D_s^+ \to$ Other	$D_s^+ \to$ Other	10.188	59.667
	$D_s^+ \to f_0(1500)\pi^+$		8.830
	$D_s^+ \to f_0(1370)\pi^+$		1.152
	$D_s^+ \to f_1(1420)\pi^+$		0.202
	$D_s^+ \to f_0(980)\pi^+$		0.010
	$D_s^+ \to \tau^+\nu_\tau$		30.139
$D_s^+ \to \pi^+\pi^-\pi^+$	$D_s^+ \to \pi^+\pi^-\pi^+$	12.803	100.000
$D_s^+ \to a_{0,1}^{\pm/0}X$	$D_s^+ \to a_1^+\eta$	12.714	34.283
	$D_s^+ \to a_1^+\phi$		15.362
	$D_s^+ \to a_1^+\omega$		9.959
	$D_s^+ \to a_1^+K^0$		9.746
	$D_s^+ \to a_1^+f_0(980)$		6.350
	$D_s^+ \to a_1^+f_2(1270)$		4.047
	$D_s^+ \to a_1^+\eta'$		2.493
	$D_s^+ \to a_1^+a_0^0$		1.325
	$D_s^+ \to a_0^0\pi^+$		8.247
	$D_s^+ \to a_0^+\pi^0$		4.753
	$D_s^+ \to a_0^0\pi^+\pi^-\pi^+$		0.923
	$D_s^+ \to a_0^+\rho^0$		0.773
	$D_s^+ \to a_0^-\pi^+\pi^+$		0.694
	$D_s^+ \to a_0^0\rho^+$		0.529

Table B.1: The definition of the subcategories used in the D_s PDF.

Group	Decay	Simulation Fraction (%)	Split (%)
	$D_s^+ \to a_0^+\pi^+\pi^-$		0.517
$D_s^+ \to \eta'\pi^+$	$D_s^+ \to \eta'\pi^+$	16.493	99.720
	$D_s^+ \to \eta(1475)\pi^+$		0.280
$D_s^+ \to \eta'(\rho^+/\pi^+\pi^0)$	$D_s^+ \to \eta'\rho^+$	8.671	97.184
	$D_s^+ \to \eta'\pi^+\pi^0$		2.816
$D_s^+ \to (\phi/\omega)\pi^+$	$D_s^+ \to \phi\pi^+$	6.278	79.169
	$D_s^+ \to \omega\pi^+$		20.831
$D_s^+ \to (\phi/\omega)(\rho^+/\pi^+\pi^0)$	$D_s^+ \to \omega\pi^+\pi^0$	10.062	35.617
	$D_s^+ \to \omega\rho^+$		40.148
	$D_s^+ \to \phi\rho^+$		24.235
$D_s^+ \to \eta\pi^+$	$D_s^+ \to \eta\pi^+$	3.810	100.000
$D_s^+ \to X\eta\pi^+$	$D_s^+ \to \eta\rho^+$	8.709	66.770
	$D_s^+ \to \eta\pi^+\pi^0$		4.451
	$D_s^+ \to \eta\pi^+\sigma$		11.862
	$D_s^+ \to \eta\pi^+\omega$		11.764
	$D_s^+ \to \eta\pi^+ f_0(980)$		5.153
$D_s^+ \to X\pi^\pm Y\pi^0$	$D_s^+ \to \pi^+\pi^-\pi^+\pi^0$	6.123	86.044
	$D_s^+ \to \pi^+\pi^-\pi^+\pi^-\pi^+$		12.580
	$D_s^+ \to \pi^+\pi^-\pi^+\pi^-\pi^+\pi^0$		1.229
	$D_s^+ \to \pi^+\pi^-\pi^+\pi^-\pi^+\pi^0\pi^0$		0.147
$D_s^+ \to X\pi^+\pi^-\pi^+$	$D_s^+ \to K^0\pi^+\pi^-\pi^+$	4.148	37.053
	$D_s^+ \to \omega\pi^+\pi^-\pi^+$		30.948
	$D_s^+ \to K^0K^0\pi^+\pi^-\pi^+$		13.068
	$D_s^+ \to \phi\pi^+\pi^-\pi^+$		9.961
	$D_s^+ \to \eta'\pi^+\pi^-\pi^+$		8.632
	$D_s^+ \to K^+K^-\pi^+\pi^-\pi^+$		0.338