

Search for the decay $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ at LHCb

Ryan Newcombe

Imperial College London
Department of Physics

A thesis submitted in part fulfilment of the requirements
for the degree of Doctor of Philosophy at
Imperial College London, April 2022

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

This work presents a search for the decay $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ using data gathered with the LHCb experiment from 2016 to 2018. No significant signal is observed and a limit is placed on the ratio of branching fractions

$$R(p\bar{p}) = \frac{B^+ \rightarrow p\bar{p}\tau^+\nu_\tau}{B^+ \rightarrow p\bar{p}\mu^+\nu_\mu} < 1.46$$

at a 95% confidence level, which represents the first limit on a semileptonic ratio of this kind involving decays that are suppressed by the most poorly understood element of the CKM matrix; V_{ub} . An investigation into the $B \rightarrow p\bar{p}$ transition form factors in $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays, which was instrumental in facilitating the search for $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$, is also presented using the 2011–2012 LHCb data sample.

For Tansy

Declaration

The work detailed in this thesis was carried out between October 2017 and March 2022. It is the result of original studies carried out by me, with the support of Ulrik Egede, Mitesh Patel, Mark Smith and the LHCb group at Imperial College London. The early version of the fitting framework was set up by Mark Smith and developed further by myself. All results and figures that are not the product of my own work are appropriately referenced.

This thesis has not been submitted for any other qualification.

Ryan Newcombe, April 2022

Acknowledgements

I thank my supervisors Mitesh Patel and Mark Smith for their support and patience throughout my PhD. Their guidance and insight was completely fundamental to the development of the work. I am also grateful to Ulrik Egede for his supervising role toward the start of my PhD. Thanks also go to members of the wider LHCb group at Imperial for their company and input along the way.

Special thanks go to my friends from my HEP PhD cohort at Imperial. Your friendship was crucial in maintaining my motivation. I am ever-grateful to the many members of Imperial College Chamber Choir that I have the tremendous fortune of calling dear friends. Being a member of the choir and singing alongside such exceptionally-talented musicians was, in no uncertain terms, one of the great privileges of my life.

A special thank you goes to Shameena Bonomally for her invaluable friendship over the last few years. You have been a constant source of counsel and encouragement and I couldn't have asked for a better Friday-night takeaway partner.

Finally, I thank my family; Owen, Jack, Mum and Dad. I love you more than you'll ever know.

*Nid wy'n gofyn bywyd moethus,
Aur y byd na'i berlau mân,
Gofyn wyf am galon hapus,
Calon onest, calon lân.*

- Gwyrosydd

Contents

Abstract	iii
Declaration	vii
Acknowledgements	ix
1 Introduction	2
2 Theory	7
2.1 Introduction to the Standard Model	7
2.2 Symmetries in the Standard Model	8
2.3 Spontaneous symmetry breaking	10
2.4 The Higgs mechanism	11
2.5 Electroweak sector	13
2.5.1 Chirality and fermion representations	13
2.5.2 Electroweak gauge fields	15
2.5.3 Fermion masses	16
2.6 Flavour-changing currents	18
2.7 Modelling $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays	20
3 The LHCb experiment	24
3.1 The Large Hadron Collider	24
3.2 The LHCb detector	25
3.3 The VELO	27
3.4 Particle identification and the RICH detectors	28
3.5 The tracking stations and magnet	32
3.6 Calorimetry	33
3.7 Muon stations	33

3.8	Trigger system	34
4	Analysis strategy	36
4.1	Rest-frame approximation	36
4.2	Choice of fit variables	37
4.3	Form-factor fit	38
4.4	Removal of partially-reconstructed decays	38
4.5	Corrections to simulation	40
4.6	Calculating $R(p\bar{p})$	41
4.7	Result blinding	41
5	Data samples and simulation corrections	42
5.1	Data samples	42
5.2	Simulation samples	43
5.3	Corrections to B kinematics	43
5.4	Trigger corrections	44
5.5	Tracking corrections	47
5.6	Particle identification corrections	47
5.7	Corrections to background kinematics	50
6	Backgrounds	52
6.1	Partially reconstructed backgrounds	52
6.2	Misidentified proton background	53
6.3	Misidentified muon background	55
6.4	Combinatorial background	57
7	Selection	61
7.1	Trigger and stripping criteria	61
7.2	Charged isolation	62
7.3	Neutral isolation	65
7.4	Particle identification criteria	68
7.5	Opening diproton angle	68
7.6	Optimisation	68
8	Form factor fit	74
8.1	Physics model	74
8.2	Implementation	76

8.3	Fit methodology	76
8.4	Reweighting strategy	78
8.5	Data, simulation and selection	82
8.6	Results	83
9	Efficiency ratio	92
9.1	Acceptance cut on decay products	93
9.2	Reconstruction and stripping	93
9.3	HLT2	94
9.4	L0 and HLT1	94
9.5	Offline selection	94
9.6	Total efficiency ratio	95
10	Fit to data	96
10.1	Fit parameters	96
10.2	Fit to data	103
10.3	Validation of data fit	109
11	Systematic uncertainties	111
11.1	Systematic uncertainties summary	112
11.2	Tracking efficiency	113
11.3	Kinematic reweighting	113
11.4	PID efficiency	115
11.5	Form factor	117
11.6	Simulation statistics	120
11.7	Misidentified proton shape	120
11.8	Misidentified muon shape	120
11.9	Part-reco background shape	121
11.10	Combinatorial background shape	124
12	Results	126
12.1	Fit yields	126
12.2	Limit setting	128
13	Conclusion	131
	Bibliography	133

A	Stripping selection	142
B	Neutral isolation BDT input features	144
C	Reweighted templates for form factor fit	145
D	Stripping selection for form factor fit	148
E	Templates for the fit to data	150
F	Part-reco background shape variations	159

List of Figures

1.1	HFLAV average of $R(D)$ and $R(D^*)$	4
1.2	Lowest order Feynman diagram for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$	5
1.3	Differential branching fraction in $m(p\bar{p})$ for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays . . .	6
2.1	A summary of the particles of the SM	8
2.2	Differential branching fractions from pQCD model	22
3.1	A schematic illustration of the LHCb detector	25
3.2	The production cross-section of $b\bar{b}$ quark pairs at LHCb	26
3.3	Cross-sections of the VELO	27
3.4	Impact parameter definition	28
3.5	RICH 1 and RICH 2 cross-sections	29
3.6	Cherenkov angles for different particle types	30
3.7	DLL efficiency as a function of track momentum	31
3.8	Comparison of ProbNNp and DLL performance	31
3.9	Reconstructed VELO hits	32
3.10	Trigger system flow chart	35
4.1	Data vs. simulation comparison of $p_T(B)$ and $p(B)$ for $B^+ \rightarrow J/\psi K^+$.	40
5.1	Fits to the $m(\mu^+\mu^-K^+)$ distribution from the 2018 $B^+ \rightarrow J/\psi K^+$ data	44
5.2	Comparison of $p_T(B)$ and $p(B)$ for $B^+ \rightarrow J/\psi K^+$ before and after reweighting for	45
5.3	L0Muon trigger efficiency as a function of $p_T(\mu)$	46
5.4	Tracking efficiency ratio between data and simulation	47
5.5	PID efficiency projections from PIDCalib	49
5.6	Differential branching fractions for $B^0 \rightarrow D^-\bar{p}\bar{p}\pi^+$ and $B^0 \rightarrow D^{*-}\bar{p}\bar{p}\pi^+$	50
5.7	Part-reco reweighting to pQCD model	51

6.1	Weights for the p mis-ID sample in the $p - \eta$ plane	56
6.2	Weights for the μ mis-ID sample in the $p - \eta$ plane	57
6.3	Event mixing validation	59
6.4	Combinatorial reweighting	60
7.1	Training performance for the charged isolation BDT	63
7.2	Feature importance for the charged isolation BDT	64
7.3	Training performance for the neutral isolation BDT	66
7.4	Feature importance for the neutral isolation BDT	67
7.5	M_{corr} vs. $\cos \theta_{p\bar{p}}$	69
7.6	Uncertainty $\sigma(\mathcal{R}(p\bar{p}))$ as a function of the isolation BDTs	72
7.7	Uncertainty $\sigma(\mathcal{R}(p\bar{p}))$ as a function ProbNN $_p^{\text{comb.}}$ and ProbNN $_\mu^{\text{comb.}}$	73
8.1	Comparison between PHSP and pQCD in q^2 and $m(p\bar{p})$	79
8.2	Shape variations in $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ due to the $D_{\parallel}^2$ FF parameter	80
8.3	Comparison between PHSP before and after reweighting to pQCD model	81
8.4	Legend for the form factor fit	84
8.5	First and second form factor fit iteration projections	85
8.6	Third and fourth form factor fit iteration projections	86
8.7	Comparison between PHSP and pQCD using D_{fit}^2 in $m(p\bar{p})$ and q^2	88
8.8	Comparison between PHSP and pQCD using D_{fit}^2 in E_μ^* and $\cos \theta_\ell$	89
8.9	Comparison in $m(p\bar{p})$ and q^2 between D_{start}^2 and D_{fit}^2	90
8.10	Comparison in E_μ^* and $\cos \theta_\ell$ between D_{start}^2 and D_{fit}^2	91
10.1	Projections of $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ simulation histogram	98
10.2	Projections of $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ simulation histogram	99
10.3	Shape variations for signal component	101
10.4	Shape variations for normalisation component	102
10.5	Legend and projection in m_{miss}^2 for the main fit	104
10.6	Projection in E_μ^* for the q^2 for the main fit	105
10.7	Profile log-likelihood in $R(p\bar{p})$	106
10.8	Correlation matrix for the main fit	108
10.9	Main fit validation	110
11.1	Alternative reweighting of $B^+ \rightarrow J/\psi K^+$ in $p_{\text{T}}(B)$ and $p(B)$	114
11.2	PID efficiency projections using alternative binning	116

11.3	Shape comparison between nominal and alternative form factor parameter for $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$	118
11.4	Shape comparison between nominal and alternative form factor parameter for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$	119
11.5	The $m(p\bar{p})$ distribution for $B^0 \rightarrow p\bar{p}D^0$ from BaBar	121
11.6	Comparison between nominal and alternative reweighting for part-reco backgrounds	122
11.7	Shape comparison between nominal and alternative reweighting for $B^{+/0} \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X$	123
11.8	Comparison between nominal and alternative combinatorial background shape	125
12.1	CLs p -value distributions	129
12.2	CLs test statistic distribution	130
C.1	PHSP and pQCD comparison for $B^+ \rightarrow p\bar{\Delta}\mu^+\nu_\mu$	146
C.2	PHSP and pQCD comparison for $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$	147
E.1	Projections of data histogram	151
E.2	Projections of $B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu$ histogram	152
E.3	Projections of $B^+ \rightarrow p(\Delta \rightarrow \bar{p}X)X\mu^+\nu_\mu$ histogram	153
E.4	Projections of $B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$ histogram	154
E.5	Projections of $B^0 \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X$ histogram	155
E.6	Projections of p mis-ID 1 histogram	156
E.7	Projections of μ mis-ID histogram	157
E.8	Projections of combinatorial histogram	158
F.1	Shape variations for $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$	160
F.2	Shape variations for $B^+ \rightarrow p\bar{\Delta}\mu^+\nu_\mu$	161

List of Tables

6.1	Part-reco background summary for the analysis	53
6.2	Track categories for the fake- p sample	54
6.3	Fake- μ track categories	56
6.4	Combinatorial candidate selection	58
7.1	Trigger criteria summary	62
7.2	Yield estimate summary for optimisation	71
7.3	Optimal cut summary	71
8.1	Part-reco background summary for form factor fit	82
8.2	Selection summary for the form factor fit	83
8.3	Form factor fit result summary	84
9.1	Summary of the DPC efficiencies	93
9.2	Summary of the reconstruction and stripping efficiencies	93
9.3	Summary of the HLT2 efficiencies	94
9.4	Summary of the L0&HLT1 efficiencies	94
9.5	Summary of the offline efficiencies	95
10.1	Fit parameter treatment summary	97
11.1	Efficiency ratio systematic uncertainties summary	112
11.2	Shape systematic uncertainties summary	112
12.1	Yield summary from the main fit	127
12.2	Summary of limits	128
A.1	Stripping selection summary	143
B.1	Neutral isolation BDT input summary	144

D.1 Summary of stripping selection for form factor fit	149
--	-----

Chapter 1

Introduction

The 20th century played host to a transformation in our understanding of the Universe. A prosperous era of discovery, successive hypotheses were put to the test through generations of experimentation, with the result being an unparalleled theory of Nature; the Standard Model (SM) of particle physics [1–4]. Despite its place as the cornerstone of theoretical particle physics, it is widely-accepted that the SM is far from the whole picture. A number of key fundamental questions remain unanswered, such as the nature of dark matter and dark energy [5, 6].

A rich global programme of experimentation in high energy physics (HEP) now exists as a result of these problems. By recreating the conditions of the early Universe, collider experiments look to examine and test the SM at the TeV scale. Over the last decade, the programme at the Large Hadron Collider (LHC) at the European Centre for Nuclear Research (CERN) has dominated the collider physics landscape. By analysing data taken during its first run from 2010-2012, both the ATLAS [7] and CMS [8] experiments reported the observation of a neutral scalar, the production and decay of which is consistent with that of the Higgs boson of the SM. Although it served as yet another verification of the SM, this observation signalled the completion of the theory as we know it. The focus is firmly on addressing the inescapable shortcomings of the SM with the search for new physics (NP).

Despite the LHC having completed its second run of data taking over the years 2015 – 2018, there have been no conclusive direct observations of NP thus far. The LHCb experiment [9] at the LHC looks to the heavy-flavour sector for NP signals, namely in the phenomena that arise from b and c hadrons. The physics programme at LHCb has resulted in a string of world-leading measurements, including the CP-violating phase in the decay of B_s^0 mesons [10] and the CP-asymmetry in $B_s^0 - \bar{B}_s^0$ mixing [11]. In

recent years, a particular set of measurements involving processes which proceed via both flavour-changing neutral currents (FCNC) and flavour-changing charged currents (FCCC) has drawn considerable attention. This group of decays involve a b -flavoured meson in the initial state and a final state with a lighter meson of different flavour and two leptons - a *semileptonic* decay. There is a growing number of results that point to the possibility of lepton flavour universality violation (LFUV) in some of these decays, with anomalies having been observed both at LHCb [12–17], and at the so-called b -factory experiments BaBar [18, 19] and Belle [20, 21].

A major motivation for carrying out these LFUV measurements stems from the fact that the observables in question, ratios of branching fractions (BFs) for modes involving leptons from different generations in the final state, are attractive from both a theoretical and experimental perspective. It is often the case that a cancellation of uncertainties arises when calculating these quantities within the framework of the SM. This yields precise predictions of the quantities in question, making them manifestly sensitive to any deviations that may arise due to NP. Experimentally, measurements can benefit from a cancellation of systematic effects induced by the detector or the choice of analysis procedure. Thus, in principle, these observables are sensitive probes for NP signals that affect flavour-changing phenomena.

The branching fractions ratios (BFRs) $R(X)$ are defined as

$$R(X) = \frac{\mathcal{B}(B \rightarrow X \ell' \nu_{\ell'})}{\mathcal{B}(B \rightarrow X \ell \nu_{\ell})}, \quad (1.1)$$

where the numerator and denominator will be referred to as the *signal* and *normalisation* modes, respectively. The final states of the decay modes have a common hadronic component X , but differ in the flavour of the $\ell^{(\prime)} - \nu_{\ell^{(\prime)}}$ pair. The most notable examples are the ratios

$$R(D^{(*)}) = \frac{\mathcal{B}(\bar{B}^0 \rightarrow D^{(*)+} \tau^- \bar{\nu}_{\tau})}{\mathcal{B}(\bar{B}^0 \rightarrow D^{(*)+} \mu^- \bar{\nu}_{\mu})}, \quad (1.2)$$

which have been measured by the LHCb [12–14], BaBar [18] and Belle collaborations [21–23]. By convention, the definition also applies to the charge-conjugate modes unless stated otherwise. These decays proceed via $W^{\pm}$ emission with a $b \rightarrow c$ transition. A major motivation for these measurements is the fact that many NP models [24–29] include enhanced couplings to the third generation. Moreover, the predictions for both $R(D)$ [30–34] and $R(D^*)$ [33–36] typically have a precision of a few percent, making them ideal probes for NP. The Heavy Flavour Averaging Group

(HFLAV) combined the aforementioned measurements to form an average which accounts for correlations between the uncertainties [37]. This can be seen in Fig. 1.1. There is an indication of LFUV with an excess of events in the tauonic mode compared

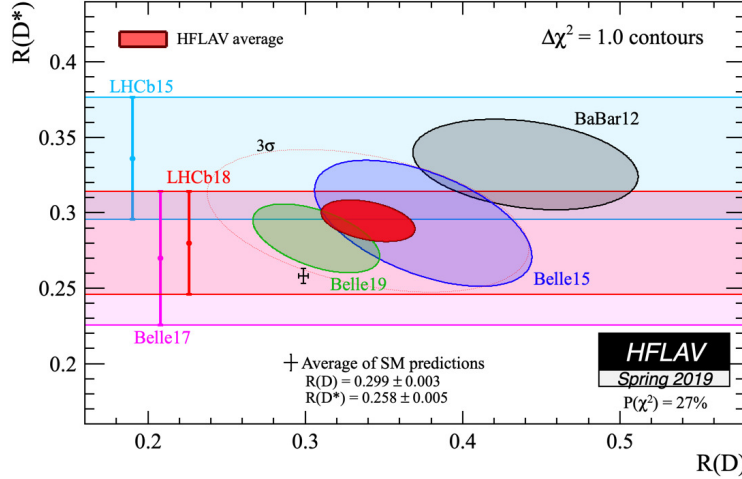


Figure 1.1: Contours of constant χ^2 for the HFLAV combination of the $R(D)$ and $R(D^*)$ measurements at LHCb and the b -factories. There is clear tension between the contour of the average in red, and the average of the SM predictions. Figure taken from Ref. [37].

to the SM calculations, with a 3.1σ tension between the averages of the predictions and the measurements. An analogous quantity $R(J/\psi)$ involving $B_c^\pm$ has also been measured at the LHCb experiment [38], which also showed a 2σ tension with the SM.

If these anomalies are due to NP, then it is desirable to probe its flavour structure; could the NP also affect a $b \rightarrow u$ process? If so, how does the effect compare to that of the $b \rightarrow c$ processes? This information is critical in guiding the formulation of the theoretical models of any NP. This thesis details the study of a new BFR $R(p\bar{p})$ involving a $b \rightarrow u$ transition, which is defined as

$$R(p\bar{p}) = \frac{\mathcal{B}(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)}. \quad (1.3)$$

The main advantage of these modes over other possible semileptonic $b \rightarrow u$ modes, such as $B \rightarrow \pi\ell\nu_\ell$, is that the probability of having random combinations of $p\bar{p} + \mu$ vs. $\pi + \mu$ tracks, or *combinatorial background*, is much lower. In the proton-proton (pp) collisions at the LHC, the production of lighter hadrons, such as pions, is far greater than that of heavier hadrons, such as protons. By extension, there is less chance of

combining a $p\bar{p}$ system with a muon which mimics a $B^\pm$ decay.

The $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ mode was first observed by the LHCb collaboration [39], and was found to have a BF

$$\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu) = (5.27^{+0.23}_{-0.24} \pm 0.21 \pm 0.15) \times 10^{-6}, \quad (1.4)$$

where the first uncertainty is statistical, the second is systematic and the third is the uncertainty on the BF of the normalisation mode chosen for the analysis. The leading-order Feynman diagram for this decay is shown in Fig. 1.2. The BF was also measured

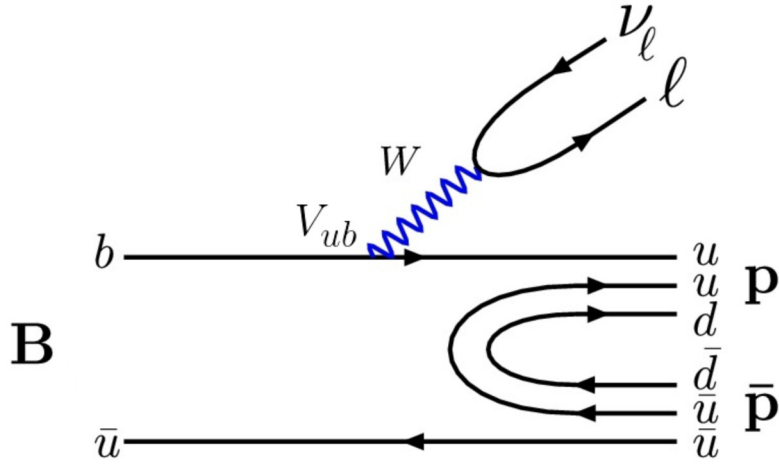


Figure 1.2: Lowest order Feynman diagram for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$. The decay is Cabibbo-suppressed and proceeds via a $b \rightarrow u\ell\nu_\ell$ transition, which introduces the CKM factor V_{ub} (Cabibbo-suppression and the CKM matrix is discussed in more detail in Sec. 2). Figure taken from Ref. [40].

differentially in the diproton invariant mass $d\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)/dm(p\bar{p})$ in order to investigate the so-called *threshold enhancement* effect seen in other similar modes with a $p\bar{p}$ system in the final state [41–44]. This is shown in Fig. 1.3. It was found that $d\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)/dm(p\bar{p})$ does indeed peak near the kinematic threshold. This observation informed further work detailed in this thesis regarding the implementation of a theoretical model outlined in Ref. [40] for decays of the form $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$. This implementation was motivated by the requirement of having simulation samples which capture the kinematic characteristics of such decays for the measurement of $R(p\bar{p})$.

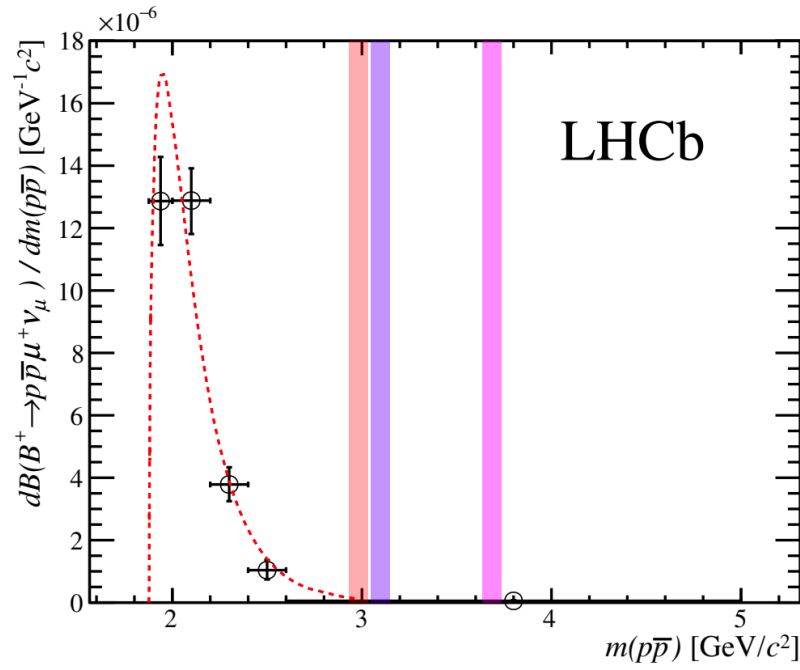


Figure 1.3: Differential branching fraction in $m(p\bar{p})$ for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays. The threshold enhancement effect is clear. Also shown are the $\eta_c \rightarrow p\bar{p}$ (red), $J/\psi \rightarrow p\bar{p}$ (blue) and $\psi(2S) \rightarrow p\bar{p}$ (pink) vetoes. The data are overlayed with a prediction (red-dashed line) using a theoretical model from Ref. [40], which is discussed in more detail in Sec. 2.7. Figure taken from Ref. [39].

Chapter 2

Theory

This chapter describes the theoretical concepts relevant to the measurement of $R(p\bar{p})$. The chapter begins by introducing the Standard Model, the electroweak and Higgs sectors, before focussing on $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays and NP models. Much of this chapter is based on Refs. [45–47].

2.1 Introduction to the Standard Model

The Standard Model [1–4] is a theoretical description of the fundamental particles and their interactions. The constituents of the SM and their main properties are summarised in Fig. 2.1, and are categorised according to their quantum numbers. They are divided into fermions and bosons based on whether they have half odd integer or integer spin, respectively. For each fermion there exists an oppositely-charged counterpart with the same mass; its anti-particle. The fermions each have spin 1/2 in the SM, and are further sub-divided into leptons and quarks based on which charges they carry, whether that charge is electromagnetic, or both electromagnetic and colour charges. In addition, there is a mass hierarchy within the fermions represented by the *generation* to which they belong.

There are two types of leptons. The charged leptons carry an electromagnetic charge $q = -1$ and are the electron e , the muon μ and the tau τ . They are said to come in three different *flavours*. The neutral leptons, or neutrinos, are electrically neutral but also come in the three different flavours of their charged partners; the electron neutrino ν_e , the muon neutrino ν_μ and the tau neutrino ν_τ . The quarks are classified as being either up-type or down-type depending on whether they have charge $q = +2/3$ or $q = -1/3$, respectively. The three up-type quarks are the up u , charm

Fermions				Bosons	
generation	1	2	3	vector	scalar
charged leptons ($q = -1$)	e 511 keV	μ 106 MeV	τ 1.78 GeV	γ 0	h 125 GeV
neutrinos ($q = 0$)	ν_e < 2 eV	ν_μ < 2 eV	ν_τ < 2 eV	$W^\pm$ 80.4 GeV	
up-type quarks ($q = +2/3$)	u 2.2 MeV	c 1.30 GeV	t 170 GeV	Z^0 91.2 GeV	
down-type quarks ($q = -1/3$)	d 4.7 MeV	s 96 MeV	b 4.6 GeV	g 0	

Figure 2.1: A summary of the particles of the SM. Their masses are quoted in natural unites where $c = 1$, and the charges q are quoted in units of the electromagnetic charge e . Figure taken from Ref. [48].

c and top t quarks, whilst the three down-type quarks are the down d , strange s and bottom b .

The vector bosons in the SM all have spin 1. The photon γ mediates the electromagnetic (EM) force. The $W^\pm$ and the Z^0 bosons mediate the weak interaction and the gluons g are the carriers of the strong interaction. The Higgs h is a spin 0, or *scalar*, boson that is responsible for the mass-generation mechanism for the fermions (except for neutrinos) in the SM, and the $W^\pm$ and Z^0 bosons.

2.2 Symmetries in the Standard Model

The SM is a quantum field theory (QFT) with a Lagrangian formalism. The aforementioned particles are thought of as quantised excitations of fields, which are treated as the components of the Lagrangian $\mathcal{L}_{\text{SM}}$. The central ideas that guided the formulation of the SM were the properties of $\mathcal{L}_{\text{SM}}$ under certain transformations of its constituent fields. There is said to be a *symmetry* of $\mathcal{L}_{\text{SM}}$ if upon performing some transformation the Lagrangian remains unchanged. Likewise, there is a breaking of the symmetry if this is not the case. Whilst this may seem fairly simple and a mere technical detail of the mathematical formulation, this crucial idea has profound implications; this principle gives rise to the rich range of phenomena associated with the SM. It would therefore be fair to say that the (in)variance of $\mathcal{L}_{\text{SM}}$ under certain transformations is *the* core idea of the SM.

The particular type of transformation relevant in the SM are *gauge* transformations.

The term gauge here is inherited from the earlier theory of classical electromagnetism. It was found that it was possible to include redundant degrees of freedom, thereby choosing a gauge, into the theory without altering the physical observables. An analogue of this requirement is made in the SM. If some change is made to the fields, such as a shift in its complex phase, then this should not change the observable properties of the theory.

This is demonstrated with the example of a local $U(1)$ transformation on the electron field ψ_e :

$$\psi_e \rightarrow \psi'_e = e^{ig_e\theta(x)}\psi_e, \quad (2.1)$$

where the complex phase $\theta(x)$ is a function of the spacetime coordinate x and g_e is the EM charge. The kinetic part of the EM Lagrangian

$$\mathcal{L}_{\text{EM}}^{\text{kin.}} = -i\bar{\psi}_e\gamma_\mu\partial^\mu\psi_e - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - m_e\bar{\psi}_e\psi_e \quad (2.2)$$

is proportional to the covariant derivative of ψ_e . The invariance $\mathcal{L}_{\text{EM}}^{\text{kin.}}$ under the transformation in Eq. 2.1 is preserved by the inclusion of the EM field tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (2.3)$$

which is proportional to the covariant derivative of the photon field A_μ . Under a local $U(1)$ transformation, A_μ transforms according to

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\theta. \quad (2.4)$$

Thus, the requirement that $\mathcal{L}_{\text{EM}}^{\text{kin.}}$ remains invariant under local $U(1)$ transformations has necessitated the introduction of the *gauge field* A_μ ; the photon. The Lagrangian is invariant under these transformations, and therefore has a $U(1)$ gauge symmetry. The properties of A_μ essentially counteract any changes induced to the Lagrangian by the transformation of ψ_e .

In the SM, this idea is extended to more complex groups of transformations. Each of the EM, weak and strong interactions correspond to the groups $U(1)$, $SU(2)$ and $SU(3)$, respectively. The overall gauge group of the SM can be written as a direct product of these three groups as $SU(3) \otimes SU(2) \otimes U(1)$.

2.3 Spontaneous symmetry breaking

A number of key experimental observations provided stringent constraints for theorists in the development of the SM. The EM interaction was observed to be long range, which meant that the photon must be massless. Conversely, the weak interaction was found to be short range, which meant that the $W^\pm$ and Z^0 bosons must have mass. Moreover, it was inferred that this mass must be approximately 100 times that of the proton mass. This presented a challenge for theorists at the time, particularly in light of their hopes to devise a unified theoretical framework for the EM and weak interactions; the electroweak sector.

This is where the idea of *spontaneous symmetry breaking* (SSB) comes in. This is a way of preserving the gauge symmetry of the Lagrangian, but where the vacuum state does not share this symmetry. This is best illustrated with the example of a massive real scalar field $\Phi(x)$ of mass m moving in a potential

$$V(\Phi) = \frac{1}{2}m^2\Phi^2 + \frac{\lambda}{4}\Phi^4, \quad (2.5)$$

where $\lambda > 0$ and whose Lagrangian $\mathcal{L}$ is

$$\mathcal{L}(\Phi) = \frac{1}{2}(\partial^\mu\Phi)(\partial_\mu\Phi) - V(\Phi). \quad (2.6)$$

A term of the form $1/2m^2\Phi^2$ is commonly referred to as a mass term for the field Φ . It is clear that $\mathcal{L}(\Phi)$ is symmetric under $\Phi \rightarrow -\Phi$. Given that SSB in this context pertains to the breaking of symmetry in ground states, it is instructive to examine the behaviour of $V(\Phi)$ near its minimum. Two cases will be considered: $m^2 > 0$ and $m^2 < 0$. The physical interpretation of $m^2 < 0$ is unimportant here, where purely illustration is the aim. For $m^2 > 0$, $V(\Phi)$ has a minimum at $\Phi = 0$ and the vacuum expectation value (VEV) of Φ is 0. However, for $m^2 < 0$, the potential can be written

$$V(\Phi) = \frac{\lambda}{4}(\Phi^2 - v^2)^2, \quad (2.7)$$

where $v = \sqrt{-m^2/\lambda}$. The minima of the potential are at $V(\Phi = +v)$ and $V(\Phi = -v)$ and the field has a nonzero VEV. The properties of the Lagrangian are examined perturbatively near the minimum. Re-writing $\Phi(x) = v + f(x)$, where $f(x)$ is some

small perturbation about $\Phi = +v$, the Lagrangian can be written

$$\mathcal{L}(f) = \frac{1}{2} (\partial^\mu f) (\partial_\mu f) - \lambda \left(v^2 f^2 + v f^3 + \frac{1}{4} f^4 \right). \quad (2.8)$$

The presence of the cubic term in f indicates that the Lagrangian is not invariant under $f \rightarrow -f$. The original symmetry of the Lagrangian has been broken around the ground state. Moreover, it can be inferred that $f(x)$ represents a massive vacuum fluctuation with mass $m_f^2 = 2\lambda v^2$. This is where the beginnings of the interplay between SSB and mass generation can be seen. However, it is worth stressing that the original symmetry $\mathcal{L}(\Phi) = \mathcal{L}(-\Phi)$ persists.

However, the Goldstone theorem states that the spontaneous breaking of a symmetry leads to a number of massless spin 0, or Nambu-Goldstone [49,50], bosons. A series of such bosons hadn't been observed at the time, nor indeed was there any indication or evidence for their existence. There was no known mechanism for generating gauge boson masses in a way that preserved the gauge symmetry of the Lagrangian and that didn't predict the existence of a number of Nambu-Goldstone bosons. This further muddled the waters, and it was widely held that devising a theory which avoided this inevitability was hopeless. Indeed, Weinberg is alleged to have remarked 'Nothing will come of nothing; speak again' during his sabbatical to Imperial College London in 1961/62 [51].

2.4 The Higgs mechanism

Inspired by analagous problems in superconductivity [52], a number of groups simultaneously proposed a mechanism for generating gauge boson masses through SSB in a way that preserved gauge invariance and didn't lead to the existence of a number of redundant Nambu-Goldstone bosons [53–55]. This mechanism is most commonly known as the Higgs mechanism today.

Central to the Higgs mechanism is the introduction of a *complex* scalar field ϕ ; the Higgs field. This is similar to the example of the *real* scalar field described in Sec. 2.3. The distinction between the real and scalar fields in these two examples is important when it comes to making the expansion about the vacuum. Clearly, in the real case, it is sufficient to include only a real part $f(x)$ in the expansion. However, in the complex case, it is necessary to introduce both a real *and* complex component. The key idea is that the complex part of the expansion, which would otherwise be interpreted as

a Nambu-Goldstone boson, can be absorbed into the Lagrangian through a choice of gauge. It is essentially eliminated through a change of variables, which is motivated by the desire to be able to write down a form of the Lagrangian which relates directly to *physically observable* fields. That is, the resulting Lagrangian includes terms that can be interpreted intuitively as either (say) mass or kinetic terms for the fields.

In a simple $U(1)$ theory with gauge field A_μ , as in Eq. 2.3, with a Higgs field ϕ , the Lagrangian can be expressed as

$$\mathcal{L} = D_\mu \phi^\dagger D^\mu \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - V(\phi), \quad (2.9)$$

where the Higgs potential $V(\phi)$ is given by

$$V(\phi) = \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2, \quad (2.10)$$

and where μ and λ are positive constants. The gauge derivative D_μ is given by

$$D_\mu = \partial_\mu \phi - ig_e A_\mu. \quad (2.11)$$

The field ϕ has a nonzero VEV $v = \mu/\sqrt{\lambda}$. Expanding about this point requires a real φ_1 and complex φ_2 shift, yielding

$$\phi(\varphi_1, \varphi_2) = v + \varphi_1 + i\varphi_2. \quad (2.12)$$

In this context it is useful to understand φ_1 and complex φ_2 themselves as fields. In order to see how φ_2 gets absorbed into the Lagrangian, it is instructive to make a change of variables in defining a new field

$$B_\mu = A_\mu - \frac{1}{g_e v} \partial_\mu \varphi_2. \quad (2.13)$$

This is tantamount to making a choice of gauge. Upon substituting this into the Lagrangian, Eq. 2.9 becomes

$$\mathcal{L} = \frac{1}{2} \partial^\mu \varphi_1 \partial_\mu \varphi_1 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} v^2 \varphi_1^2 + \frac{1}{2} g_e^2 v^2 B_\mu B^\mu + \text{cubic and quartic terms}. \quad (2.14)$$

Thus, both the φ_1 field and the B_μ have acquired mass, whilst the φ_2 is nowhere to be seen. This idea is extended in the context of the SM to a more complex transformation group. However, the principle remains the same. In this example, the field φ_1

corresponds to the Higgs boson, and the B_μ to a massive gauge boson.

2.5 Electroweak sector

It wasn't until three years after the publication of the Higgs mechanism papers that a unified electroweak theory came about. Building on earlier work by Glashow, in a series of papers in 1967 Weinberg and Salam (GWS) devised such a theory that both provided masses to the weak gauge bosons and preserved gauge symmetry. As mentioned before, the Higgs mechanism of the SM is akin to the example involving a $U(1)$ theory, but extended to the $SU(2) \times U(1)$ symmetry group.

2.5.1 Chirality and fermion representations

A key experimental observation that guided the theory was that the weak interaction violated parity conservation. It was seen by Wu and collaborators that the electrons emitted during the β decay of polarised cobalt nuclei were not isotropic [56]; they were preferentially emitted in a direction opposite to that of the nuclear spin. This can be explained by a weak interaction which is chirality-dependent. That is, it couples only to left-handed (LH) fermions and right-handed (RH) anti-fermions. The LH and RH components of the fermion field ψ are defined as

$$\psi_L = P_L \psi = \frac{(1 + \gamma^5)}{2} \psi \quad (2.15)$$

$$\psi_R = P_R \psi = \frac{(1 - \gamma^5)}{2} \psi, \quad (2.16)$$

where P_L and P_R are the LH and RH projection operators. The idea of this chirality-dependent coupling is encoded into EW theory in terms of the transformation properties of the fermions. The LH fermions *do* transform under $SU(2)$ whilst the RH fermions do not. A general $SU(2)$ transformation in the isospin representation can be written in terms of a coupling constant g , a rotation vector $\boldsymbol{\theta}$ and the Pauli matrices $\boldsymbol{\sigma}$ as

$$\psi_L \rightarrow \psi'_L = e^{-\frac{1}{2}g\boldsymbol{\theta}\cdot\boldsymbol{\sigma}} \psi_L. \quad (2.17)$$

As will be seen later, the constant g can be interpreted as the weak coupling. Thus, if $g \neq 0$, then an $SU(2)$ transformation proceeds as Eq. 2.17 which subsequently results in the presence of weak interaction terms in the Lagrangian. However, if $g = 0$, then

no transformation occurs.

This translates into particular representations for the fermions in the SM. Given the presence of the Pauli matrices above, it is clear that $SU(2)$ acts only on two-component vectors or *spinors* in the isospin representation. The most convenient representation of the fermions is in terms of eigenvectors of σ_3 , with eigenvalues $I_3 = \pm 1/2$. The three left-handed lepton generations $i = 1, 2, 3$ are expressed as doublets

$$L^i = \begin{pmatrix} e \\ \nu_e \end{pmatrix}_L, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L, \quad (2.18)$$

and the doublets representing the three quark flavours are

$$Q^i = \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L. \quad (2.19)$$

Conversely, the right-handed components of the fermions are grouped into nine singlets with $I_3 = 0$

$$\begin{aligned} u_R^i &= u_R, c_R, t_R, \\ d_R^i &= d_R, s_R, b_R, \\ l_R^i &= e_R, \mu_R, \tau_R. \end{aligned} \quad (2.20)$$

Notably, there are no singlets for RH neutrinos as they are yet to be observed experimentally.

The representations above play a fundamental role in the host of phenomena exhibited by the SM. This, in conjunction with the fact that the appropriate gauge group for EW theory is $SU(2) \times U(1)$, gives rise to LFU. Thus, LFU is not a true ‘symmetry’ of the SM. However, it can be understood as a manifestation of the underlying structure of the SM. Despite this, the term symmetry is often used when it comes to LFU. This is misleading and it should not be conflated with the true gauge symmetries of the SM.

2.5.2 Electroweak gauge fields

To preserve the gauge symmetry of the Lagrangian under $SU(2) \times U(1)$ transformations, the fields W_μ^a , where $a = 1, 2, 3$, and B_μ such that the gauge derivative of the EW sector takes the form

$$D_\mu = \left(\partial_\mu + \frac{i}{2} g W_\mu^a \sigma^a + \frac{i}{2} g' Y B_\mu \right), \quad (2.21)$$

where the hypercharge Y is related to the electric charge g_e according to

$$Y = g_e - I_3. \quad (2.22)$$

This is related to the particular representation of $U(1)$ and is chosen such that the correct charges g_e are reproduced in the photon interactions terms. The SM Higgs ϕ is introduced as a complex $SU(2)$ doublet

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}. \quad (2.23)$$

The Higgs Lagrangian $\mathcal{L}_{\text{Higgs}}$ can then be written

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \quad (2.24)$$

where $V(\phi)$ has the same definition as in Eq. 2.10. In order to see how the physical gauge bosons arise from the W_μ^a and B_μ fields, it is necessary to examine the Lagrangian with the Higgs field at its VEV

$$\phi_{\text{vac.}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.25)$$

The first term in $\mathcal{L}_{\text{Higgs}}$ becomes

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{2} \left(m_W^2 (W_\mu^+)^{\dagger} W^{\mu,+} + m_W^2 (W_\mu^-)^{\dagger} W^{\mu,-} + m_Z^2 (Z_\mu^0)^{\dagger} Z^{0,\mu} \right), \quad (2.26)$$

where the physical $W_\mu^\pm$ and Z_μ^0 fields are linear superpositions of the W_μ^a and B_μ fields

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2), \\ Z_\mu^0 &= \frac{1}{\sqrt{g^2 + g'^2}} (gW_\mu^3 - g'B_\mu), \end{aligned} \quad (2.27)$$

and their masses are given by

$$\begin{aligned} m_W &= \frac{vg}{2}, \\ m_Z &= \frac{v\sqrt{g^2 + g'^2}}{2}. \end{aligned} \quad (2.28)$$

The Higgs field has successfully generated mass terms for the $W_\mu^\pm$ and Z_μ^0 fields without the appearance of a mass term for the photon field A_μ as hoped. Comparing the form of the photon field

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (g'W_\mu^3 + gB_\mu) \quad (2.29)$$

and that of the Z_μ^0 is suggestive of a relation between the two. This can be expressed in terms of the weak mixing angle θ_W as

$$\begin{pmatrix} Z_\mu^0 \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \quad (2.30)$$

and where

$$\begin{aligned} \cos \theta_W &= \frac{g}{\sqrt{g^2 + g'^2}}, \\ \sin \theta_W &= \frac{g'}{\sqrt{g^2 + g'^2}}. \end{aligned} \quad (2.31)$$

The masses of the $W_\mu^\pm$ and Z_μ^0 are also related by $M_W = M_Z \cos \theta_W$.

2.5.3 Fermion masses

As already alluded to, in the SM fermions acquire mass through an interaction term with the Higgs boson in the Yukawa Lagrangian $\mathcal{L}_Y$. For the first lepton generation,

this is given by

$$\mathcal{L}_Y^e = -\sqrt{2}\lambda_e \left[\begin{pmatrix} \bar{\nu}_e & \bar{e} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R \begin{pmatrix} \phi^{+,*} & \phi^{0,*} \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right], \quad (2.32)$$

where λ_e is the Yukawa coupling. Making an expansion of the Higgs field about the VEV

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (2.33)$$

and then evaluating Eq. 2.32 gives

$$\mathcal{L}_Y^e = -\lambda_e h (\bar{e}_L e_R + \bar{e}_R e_L) - \lambda_e v (\bar{e}_L e_R + \bar{e}_R e_L). \quad (2.34)$$

The first term constitutes an interaction between the electron and the Higgs boson h , and the second term represents a mass term for the electron with $m_e = \lambda_e v$ whilst the neutrino has remained massless.

The generation of quark masses is more involved. In contrast to the leptons, the goal here is to provide a mass term to both components of the quark doublets. Considering the first generation, a mass term for both the u and d quarks is required. However, working with the expansion as in Eq. 2.33, it is clear that only the d -type quarks will acquire mass due to the presence of only one non-zero component in the Higgs doublet.

In generating u -type quark masses, it is therefore necessary to perform a transformation on $\phi(x)$ such that its components are inverted. This is done by acting on $\phi(x)$ with $i\sigma_2$. The Yukawa Lagrangian for the first generation of quarks then reads

$$\begin{aligned} \mathcal{L}_Y^q = & -\sqrt{2}\lambda_d \left[\begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R + \bar{d}_R \begin{pmatrix} \phi^{+,*} & \phi^{0,*} \end{pmatrix} \begin{pmatrix} u \\ d \end{pmatrix}_L \right] \\ & -\sqrt{2}\lambda_u \left[\begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L (i\sigma_2) \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R + \bar{d}_R \begin{pmatrix} \phi^{+,*} & \phi^{0,*} \end{pmatrix} (i\sigma_2)^\dagger \begin{pmatrix} u \\ d \end{pmatrix}_L \right]. \end{aligned} \quad (2.35)$$

The u and d quark masses are then $m_u = \lambda_u v$ and $m_d = \lambda_d v$ after substituting Eq. 2.33 into $\mathcal{L}_Y^q$. The full Yukawa Lagrangian for the three fermion generations necessitates

the introduction of three additional non-diagonal 3×3 matrices λ_{ij}^ℓ , λ_{ij}^u and λ_{ij}^d ; the Yukawa coupling matrices. The full Yukawa Lagrangian is then given by

$$\mathcal{L}_Y = v \left[-\lambda_{ij}^u \bar{u}_L^i u_R^j - \lambda_{ij}^d \bar{d}_L^i d_R^j - \lambda_{ij}^\ell \bar{\ell}_R^i \ell_R^j + \text{h.c.} \right], \quad (2.36)$$

where the corresponding anti-particles are included through the Hermitian conjugate (h.c.) terms.

2.6 Flavour-changing currents

The interactions between the weak gauge bosons and the fermions manifest themselves when the gauge derivative D_μ in Eq. 2.21 acts on the LH doublets L^i and Q^i , as well as the RH singlets. The interaction term for leptons takes the form:

$$\begin{aligned} \mathcal{L}_\ell = & \frac{g_e}{\sqrt{2} \sin \theta_W} \left[W_\mu^+ \bar{\nu}_L^i \gamma^\mu \ell_L^i + W_\mu^- \bar{\ell}_L^i \gamma^\mu \nu_L^i \right] \\ & + g_e A_\mu \bar{\ell}^i \gamma^\mu \ell^i \\ & + \frac{g_e}{\tan \theta_W} Z_\mu^0 \bar{\ell}^i \gamma^\mu \ell^i \\ & + \frac{g_e}{\sin \theta_W \cos \theta_W} Z_\mu^0 \left[\frac{1}{2} \bar{\nu}_L^i \gamma^\mu \nu_L^i - \frac{1}{2} \bar{\ell}_L^i \gamma^\mu \ell_L^i \right], \end{aligned} \quad (2.37)$$

where the first line corresponds to the *charged current* and the second to the coupling to the photon field in the EM current. The terms in the third and fourth lines involving the Z_μ^0 boson are the *neutral currents*. It is clear that the coupling of both the $W_\mu^\pm$ and Z_μ^0 bosons to each generation i of the leptons is the same. This is lepton flavour universality (LFU).

However, there are some important distinctions to be made when it comes to the equivalent interaction terms for the quarks:

$$\begin{aligned} \mathcal{L}_q = & \frac{g_e}{\sqrt{2} \sin \theta_W} \left[W_\mu^+ \bar{u}_L^i \gamma^\mu d_L^i + W_\mu^- \bar{d}_L^i \gamma^\mu u_L^i \right] \\ & + g_e A_\mu \left[\frac{2}{3} \bar{u}_i \gamma^\mu u_i - \frac{1}{3} \bar{d}_i \gamma^\mu d_i \right] \\ & - g_e \tan \theta_W Z_\mu^0 \left[\frac{2}{3} \bar{u}_i \gamma^\mu u_i - \frac{1}{3} \bar{d}_i \gamma^\mu d_i \right] \\ & + \frac{g_e}{\sin \theta_W \cos \theta_W} Z_\mu^0 \left[\frac{1}{2} \bar{u}_L^i \gamma^\mu u_L^i - \frac{1}{2} \bar{d}_L^i \gamma^\mu d_L^i \right]. \end{aligned} \quad (2.38)$$

The charged current terms involve couplings which are bilinear in the up-type quark fields and the down-type quark fields. For any process that involves a coupling of the $W_\mu^\pm$ fields to the quarks necessitates vertices involving quarks of different flavours; a *flavour-changing* current.

Moreover, it is possible for these currents to involve quarks of different generations in a phenomenon called *quark mixing*. Looking at Eq. 2.36, it can be seen that the Yukawa matrices $\lambda_{ij}^{u,d}$ contain non-zero off-diagonal elements. The Yukawa couplings are proportional to some admixture of quarks of different generations. This is the representation of the quarks in the *flavour basis*.

It is possible to make a change of basis to the *mass basis*. Here, the off-diagonal elements of the Yukawa coupling matrices are zero, making it more convenient to read off the mass terms for each of the quarks. This is done with the use of four unitary matrices U_u , U_d , S_u and S_d :

$$\lambda^u = U_u \Lambda^u S_u^\dagger, \quad \lambda^d = U_d \Lambda^d S_d^\dagger. \quad (2.39)$$

The matrices $\Lambda^{u,d}$ are diagonal and the quarks fields transform as:

$$\begin{aligned} u_L &\rightarrow u'_L = U_u u_L \\ d_L &\rightarrow d'_L = U_d d_L \\ u_R &\rightarrow u'_R = S_u u_R \\ d_R &\rightarrow d'_R = S_d d_R \end{aligned} \quad (2.40)$$

Naturally, this has consequences for the form of some of the terms in the Lagrangian. Notably, for L_q in Eq. 2.38, the charged-current terms pick up a factor $V_{\text{CKM}}^{ij} = \{U_u^\dagger U_d\}^{ij}$ and become:

$$\mathcal{L}_q^{\text{CC}} = \frac{g_e}{\sqrt{2} \sin \theta_W} \left[W_\mu^+ \bar{u}_L^i \gamma^\mu V^{ij} d_L^j + W_\mu^- \bar{d}_L^i \gamma^\mu (V^\dagger)^{ij} u_L^j \right] \quad (2.41)$$

The Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM} [57, 58] captures the amount of mixing between the different quark generations through the moduli of its elements

V_{CKM}^{ij} . This is most clearly seen in the Wolfenstein parametrisation [59] of the matrix:

$$V_{\text{CKM}} \approx \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}, \quad (2.42)$$

where $\lambda = \sin \theta_{12} \approx 0.22$, $A\lambda^2 = \sin \theta_{23}$ and $A\lambda^3(\rho - i\eta) = \sin \theta_{13} \exp(-i\delta)$ and θ_{12} , θ_{23} and θ_{13} are the Euler angles, and δ is a complex phase which is responsible for CP violation in the quark sector. It should be noted that this is only an approximate representation of the true matrix, but is easier to interpret. Clearly, the off-diagonal elements of V_{CKM} are smaller than those on the diagonal, resulting in smaller couplings between the first and third generations of quarks. This is termed Cabibbo suppression. The largest couplings are between quarks of the same generation. Understanding the broad range of phenomena which arises from the CKM matrix is one of the main goals of flavour physics.

2.7 Modelling $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays

The $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays proceed via a flavour-changing charged current. A $b \rightarrow u\ell^+\nu_\ell$ transition takes place at the quark level and the decay is Cabibbo-suppressed with a factor V_{ub} entering at the $W^\pm$ – quark vertex. Whilst it is rather straightforward to describe the decay at the quark level, a formulation of the transition to a $p\bar{p}$ final state is more complex.

Quantum chromodynamics (QCD) is the gauge theory of the SU(3) symmetry group and describes the strong interaction. The massless gauge boson of the strong interaction is the gluon and the strength of the coupling to quarks is given by α_s . Crucially, α_s depends on the energy scale of the interaction. The coupling becomes stronger at lower energy scales and smaller at higher energy scales. As is the case for EW calculations, it is commonplace to define processes involving QCD in terms of a perturbative expansion of the coupling. The fact that α_s becomes large at lower energy scales means that perturbative techniques are no longer valid for low energy (or *slow*) systems.

However, in the absence of other methods, perturbative techniques can be used approximately. This is the approach taken to model the hadronic transition in $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$

decays by using the *perturbative QCD (pQCD) counting rules* [40, 60, 61]. It is worth noting that the literature on $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays is scarce, and that the theoretical description is very much in its infancy. An effective Hamiltonian $\mathcal{H}_{\text{eff}}$ for the $b \rightarrow u\ell\nu_\ell$ transition is written as a product of two $V - A$ currents, one for each of the $W^\pm$ -quark and $W^\pm$ -lepton vertices, as

$$\mathcal{H}_{\text{eff}} = \frac{G_F V_{ub}}{\sqrt{2}} \bar{u}\gamma_\mu(1 - \gamma_5)b \bar{\ell}\gamma^\mu(1 - \gamma_5)\nu_\ell, \quad (2.43)$$

where Fermi's constant G_F and the CKM factor V_{ub} determines the strength of the interaction. This is then promoted to a hadron-level amplitude $\mathcal{A}(B^+ \rightarrow p\bar{p}\ell^+\nu_\ell)$ which takes the factorized form

$$\mathcal{A}(B^+ \rightarrow p\bar{p}\ell^+\nu_\ell) = \frac{G_F V_{ub}}{\sqrt{2}} \langle p\bar{p} | \bar{u}\gamma_\mu(1 - \gamma_5)b | B \rangle \ell\gamma^\mu(1 - \gamma_5)\nu_\ell, \quad (2.44)$$

where the vector and axial-vector components of the $B \rightarrow p\bar{p}$ transition matrix element are respectively

$$\begin{aligned} \langle Y\bar{Y} | \bar{q}\gamma_\mu b | B \rangle = \\ i\bar{u}(p_Y) [g_1\gamma_\mu + g_2i\sigma_{\mu\nu}q^\nu + g_3q_\mu + g_4(p_{\bar{Y}} + p_Y)_\mu + g_5(p_{\bar{Y}} - p_Y)_\mu] \gamma_5 v(p_{\bar{Y}}), \end{aligned} \quad (2.45)$$

$$\begin{aligned} \langle Y\bar{Y} | \bar{q}\gamma_\mu\gamma_5 b | B \rangle = \\ i\bar{u}(p_Y) [f_1\gamma_\mu + f_2i\sigma_{\mu\nu}q^\nu + f_3q_\mu + f_4(p_{\bar{Y}} + p_Y)_\mu + f_5(p_{\bar{Y}} - p_Y)_\mu] v(p_{\bar{Y}}), \end{aligned} \quad (2.46)$$

for a general baryon pair $Y\bar{Y}$ and with $q = p_B - p_Y - p_{\bar{Y}}$. Both components take the form of a sum of all possible bilinear products of the various combinations of four-momenta and the spinors of the diproton pair $\bar{u}(q_Y)$ and $v(q_{\bar{Y}})$. According to the pQCD counting rules [60, 61], the vector and axial-vector form factors (FFs) f_i and g_i , respectively, are expressed as

$$f_i = \frac{D_{f_i}}{t^n}, \quad g_i = \frac{D_{g_i}}{t^n}, \quad (2.47)$$

where $t = (q_Y + q_{\bar{Y}})^2$ is the squared dibaryon mass and n is an index. The constants D_{f_i} and D_{g_i} are taken from measurements of other fully hadronic modes with a dibaryon system in the final state [62–66]. In fact, the D_{f_i} and D_{g_i} are themselves

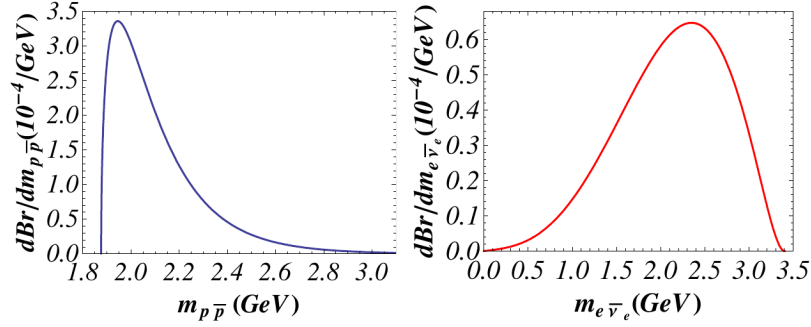


Figure 2.2: Differential branching fractions in $m(p\bar{p})$ and $m(e\nu_e)$ as predicted by the pQCD model. There is a clear enhancement in $m(p\bar{p})$ near the kinematic threshold. Figure taken from Ref. [40]

expressed as linear superpositions of six other constants, which are:

$$D_{g_1} = \frac{5}{3}D_{\parallel} - \frac{1}{3}D_{\parallel\bar{\parallel}}, \quad (2.48a)$$

$$D_{f_1} = \frac{5}{3}D_{\parallel} + \frac{1}{3}D_{\parallel\bar{\parallel}}, \quad (2.48b)$$

$$D_{g_j} = \frac{5}{3}D_{\parallel}^j = -D_{f_j}, \quad (2.48c)$$

where $j = 2, 3, 4, 5$, and they take the values [67, 68]

$$(D_{\parallel}, D_{\parallel\bar{\parallel}}) = (67.7 \pm 16.3, -280.0 \pm 35.9) \text{ GeV}^5, \quad (2.49a)$$

$$(D_{\parallel}^2, D_{\parallel}^3) = (-187.3 \pm 26.6, -840.1 \pm 132.1) \text{ GeV}^4, \quad (2.49b)$$

$$(D_{\parallel}^4, D_{\parallel}^5) = (-10.1 \pm 10.8, -157.0 \pm 27.1) \text{ GeV}^4. \quad (2.49c)$$

Examples of the resulting differential branching fractions (DBFs) for $Y\bar{Y} = p\bar{p}$ in diproton mass $m(p\bar{p})$ and dilepton mass $m(\ell\nu_\ell)$ for $\ell = e$ are shown in Fig. 2.2. Whilst the model succeeds in reproducing the $m(p\bar{p})$ threshold enhancement effect shown for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ in Fig. 1.3, there are some shortcomings. Notably, the model results in a prediction for the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ BF which is approximately an order of magnitude higher than was measured. However, in the absence of any other theoretical description of the decay, the model is used for this work.

As will be discussed in more detail in Sec. 8, a significant part of the work in this thesis involved the implementation of this model for the measurement of $R(p\bar{p})$. In addition, a measurement of the FF parameters in Eq. 2.48 for $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays is made in order to mitigate the approximation of using the numerical values obtained

from fully hadronic decay modes in Eq. 2.49.

Chapter 3

The LHCb experiment

This chapter introduces the Large Hadron Collider (LHC) and the LHCb experiment. The constituent subdetectors of LHCb are described, with particular attention paid to those most relevant to the work in this thesis. Finally, a summary of the data collection process for physics analyses at LHCb is provided.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular particle accelerator with a 27km circumference at CERN, situated 50-157m beneath the Swiss-French border. Whilst there is some utility for the LHC as a heavy-ion collider, the main use is to generate pp collisions at the highest centre-of-mass energies $\sqrt{s}$ ever produced. The LHC operated at $\sqrt{s} = 7$ TeV in 2011, $\sqrt{s} = 8$ TeV in 2012 and at $\sqrt{s} = 13$ TeV from 2015 to 2018. The operational periods from 2011 to 2012 and from 2015 to 2018 are known as Run 1 and Run 2, respectively.

A network of smaller accelerators serve to accelerate the proton beams before being injected into the LHC where they reach the nominal centre-of-mass energy. In the LHC, the two beams circulate in opposite directions and are collided at four main points along the accelerator, each of which is occupied by a detector. The ATLAS [69] and CMS [70] experiments are so-called general purpose detectors (GPDs) with a cylindrical design, concentric around the beam pipe. The ALICE [71] experiment is designed to study heavy-ion collisions, whilst the LHCb [72] experiment is dedicated to studying flavour physics and gathered the data analysed in this thesis.

Two more quantities are introduced here: the luminosity, or instantaneous luminosity, $\mathcal{L}$ and the integrated luminosity $\int \mathcal{L} dt$. The former is a measure of the number

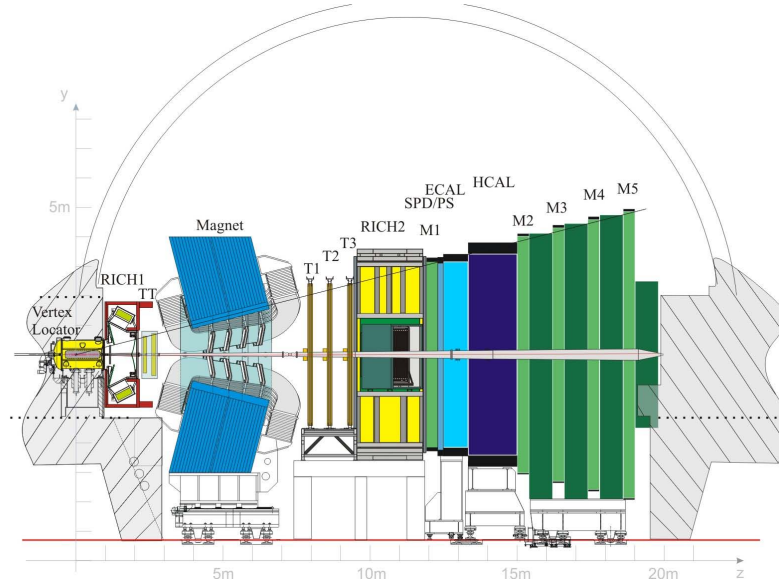


Figure 3.1: A schematic illustration of the LHCb detector cross-section where all of the subdetectors have been labelled. The beam pipe points along the z -axis, the y -axis is vertical and the x -axis is perpendicular to the page. Figure taken from Ref. [72].

of pp interactions produced per unit area, per unit time. As the name suggests, the latter corresponds to $\mathcal{L}$ integrated over time and is a measure of the number of pp interactions generated in a given time period.

3.2 The LHCb detector

The design of the LHCb detector, as shown in Fig. 3.1, is optimised for the study of particles containing b or c quarks, or heavy flavour particles. In contrast to the GPDs with their cylindrical design, LHCb covers the azimuthal beam angle range $10 < \theta < 250$ mrad, or the pseudorapidity range $2 < \eta < 5$, where η is related to θ according to

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (3.1)$$

The motivation for this design is most clearly demonstrated in Fig. 3.2 which shows the production cross section $\sigma(b\bar{b})$ of $b\bar{b}$ quark pairs. It is clear that the cross-section is at its largest near the $\theta \approx 0$ and $\theta \approx \pi$, or respectively the forward and backward, regions. It is for this reason that LHCb is referred to as a forward-arm spectrometer.

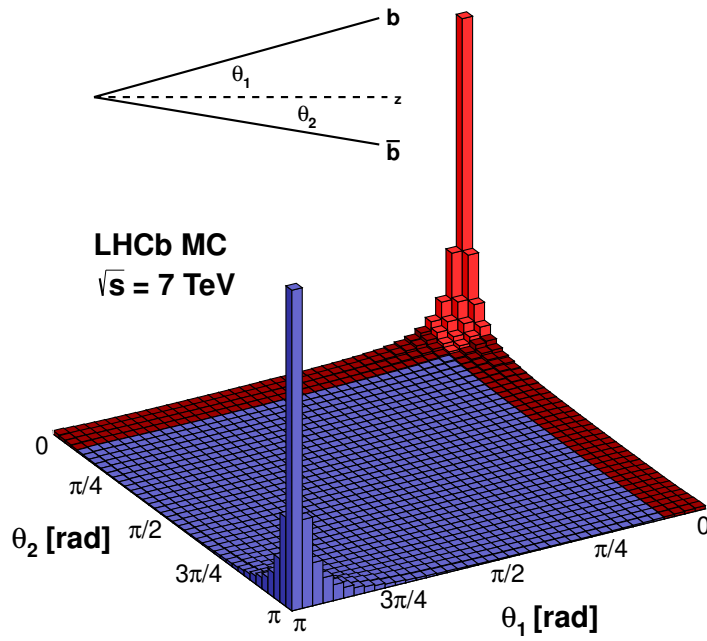


Figure 3.2: The production cross-section of $b\bar{b}$ quark pairs, taken from simulation, as a function of the opening azimuthal angles θ_1 and θ_2 of the b and $\bar{b}$ quarks, respectively, with respect to the z -axis. The red-shaded region represents the LHCb acceptance. Figure taken from Ref. [73].

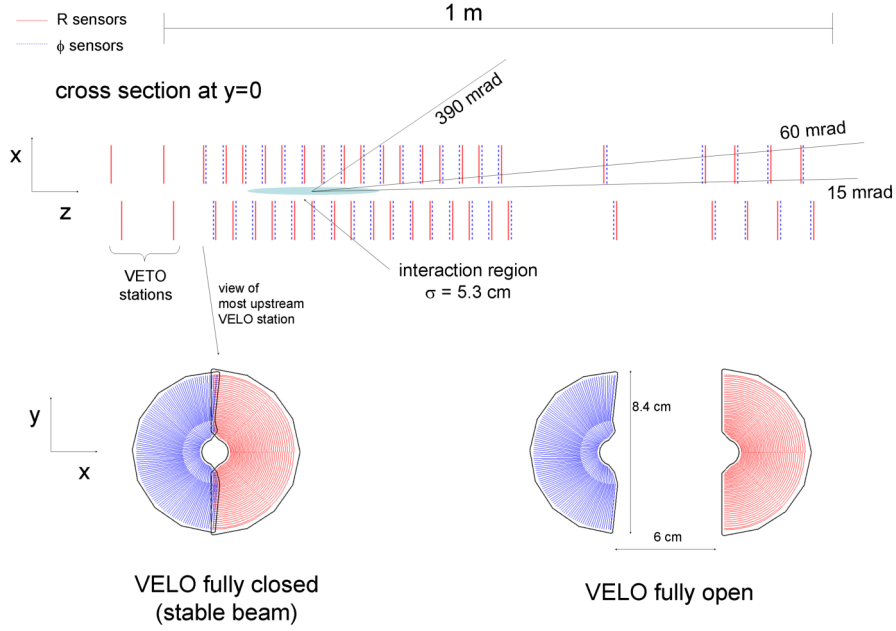


Figure 3.3: Schematic cross-sections of the VELO in the (top) $x - z$ plane and the (bottom) $y - x$ plane in the (left) closed and (right) open positions. The azimuthal sensors are shown in blue, whilst the radial sensors are in red. Figure taken from Ref. [72].

3.3 The VELO

The VERtEX LOcator (VELO) surrounding the pp interaction point is a tracking detector [74]. Hadrons containing b or c quarks are produced at a primary vertex (PV) as a result of a pp interaction. The quarks propagate for a short distance, typically $\approx 1\text{cm}$ for b -hadrons, before decaying to a collection of daughter particles at the secondary vertex (SV). Thus, one of the key experimental characteristics of a heavy flavour decay is a displaced SV from the PV. The primary purpose of the VELO is to find the trajectory of charged particles such that the SV can be distinguished from the PV in a given pp interaction.

As illustrated in Fig. 3.3, the VELO consists of 21 modules each split into semi-circular halves. The two halves meet approximately along the $y - z$ plane where $x = 0$. During the injection phase of the two proton beams, the trajectories of which are initially unstable, the VELO is in the open position where the halves of each module are $\approx 6\text{ cm}$ apart. Once the beams are stabilised for data taking, the modules are positioned $\approx 8\text{ mm}$ from the beam line.

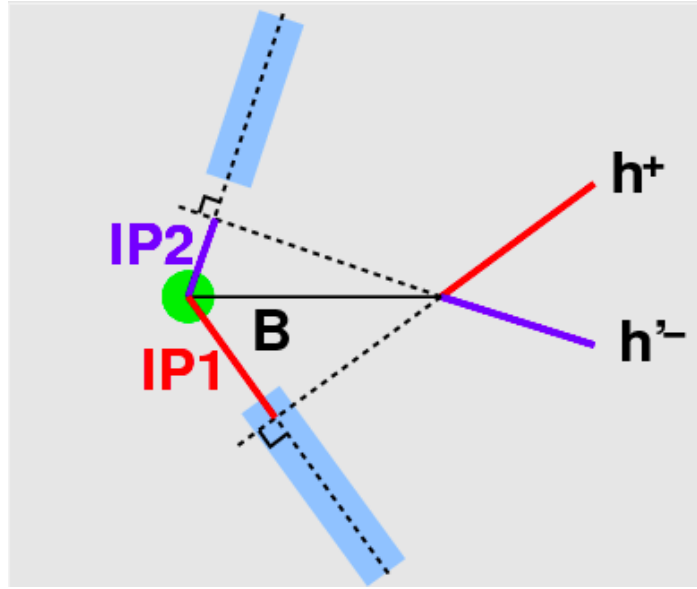


Figure 3.4: Illustration of the meaning of the impact parameter (IP) at LHCb. A B meson decays to two hadrons h^+ and h^- , with impact parameters IP1 and IP2, respectively. The IP of a particle is the closest perpendicular distance between its line of flight and the PV. Figure taken from Ref. [75].

The modules consist of both R and ϕ sensors, that provide radial and azimuthal information, respectively. This configuration of the sensors results in excellent resolutions on the position of the PV, and also on the impact parameter (IP). As can be seen from Fig. 3.4, the IP is defined as the minimum perpendicular distance between a daughter particle's line of flight and the PV. Particles that originate from a heavy flavour decay typically have higher IP than those that originate from the PV. It is therefore an important quantity in LHCb analyses. The PV and IP resolution is typically $\mathcal{O}(10)\mu\text{m}$, but this can vary principally as a function of the particle momentum and the number of tracks originating from the PV.

3.4 Particle identification and the RICH detectors

At LHCb, particle identification (PID) is aided by the Ring Imaging Cherenkov (RICH) detectors [76]. Efficient PID is crucial for the measurement of $R(p\bar{p})$ in order to distinguish the $p\bar{p}$ daughters from other particles in the busy hadronic environment surrounding the PV. There are two such detectors, labelled RICH 1 and RICH 2 in Fig. 3.1, where the former is optimised for the measurement of lower momentum particles whilst the second is designed for those of higher momentum.



Figure 3.5: Schematic cross-sections of the (left) RICH 1 and (right) RICH 2 detectors in the $y - z$ plane. Figure taken from Ref. [72].

The core idea behind the RICH detectors relates to the propagation of particles in a fluid. Cherenkov radiation is emitted when charged particles travel faster than the phase speed of light in that medium. A cone of light concentric with the line of flight of the particle is formed, where the Cherenkov angle θ_c denotes the apex angle of the cone. It is related to the mass m and momentum p of the particle and the refractive index n of the medium through

$$\cos(\theta_c) = \frac{\sqrt{m^2 c^2 + p^2}}{np} \quad (3.2)$$

As can be seen from Fig. 3.5, RICH 1 and RICH 2 are filled with different gases, making them suitable for different momentum ranges. The particles traverse the gas, emitting Cherenkov light cones that are focused by a series of mirrors. A system of hybrid photon detectors (HPDs) serves to measure the rings of light which are formed by the incident cones. Fig. 3.6 shows θ_c as a function of p for different particle types. There is a clear difference in the Cherenkov angles produced by protons, above the momentum threshold, and those of other track types. This is vital for the measurement of $R(p\bar{p})$ in order to correctly identify the products of the $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$. Furthermore, it is possible to construct a likelihood function $\mathcal{L}_X$ under the hypothesis that a given

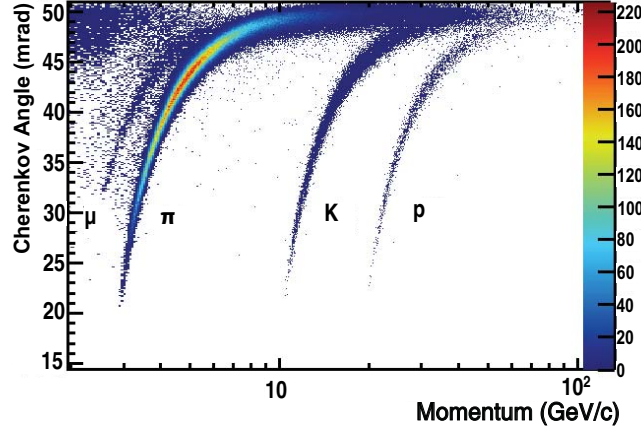


Figure 3.6: Reconstructed Cherenkov angles θ_c as a function of momentum for different particle species in RICH 1. Figure taken from Ref. [77].

track is of particle type X . This is done by combining information from the RICH detectors and the tracking. The difference in the logarithms of these likelihoods (DLLs) are defined as

$$\text{DLL}_X = \log(\mathcal{L}_X) - \log(\mathcal{L}_\pi). \quad (3.3)$$

The π mass hypothesis can therefore be understood as a benchmark hypothesis. DLLs are widely used as discriminants in LHCb physics analyses in order to reduce the amount of contamination from track mis-identification (misID). The efficiencies of two DLL criteria for protons to be correctly identified as protons ($p \rightarrow p$) and for kaons to be mis-identified as protons ($K \rightarrow p$) are shown in Fig. 3.7. There is a clear increase in the $p \rightarrow p$ efficiency above the Cherenkov threshold before it decreases at higher momenta. On the contrary, the $K \rightarrow p$ misID rate increases at high momenta. This is due to the fact that the Cherenkov angles for each of the different track types converge at very high momentum, as is clear from Fig. 3.6, resulting in less discrimination power. An additional PID discriminant is constructed using a machine learning (ML) technique, which uses a multi-variate algorithm (MVA) termed a neural network (NN). The MVA takes various properties of the particle track as inputs, such as the track momentum and transverse momentum, in order to form a quantity ProbNNX which is akin to a probability that a track is a given type X [79]. The improvement that the ProbNNs provide over using DLL alone is illustrated in Fig. 3.8. Understanding the PID response of the detector for different particle species is a crucial part of LHCb analyses. The measurement of $R(p\bar{p})$ is no exception.

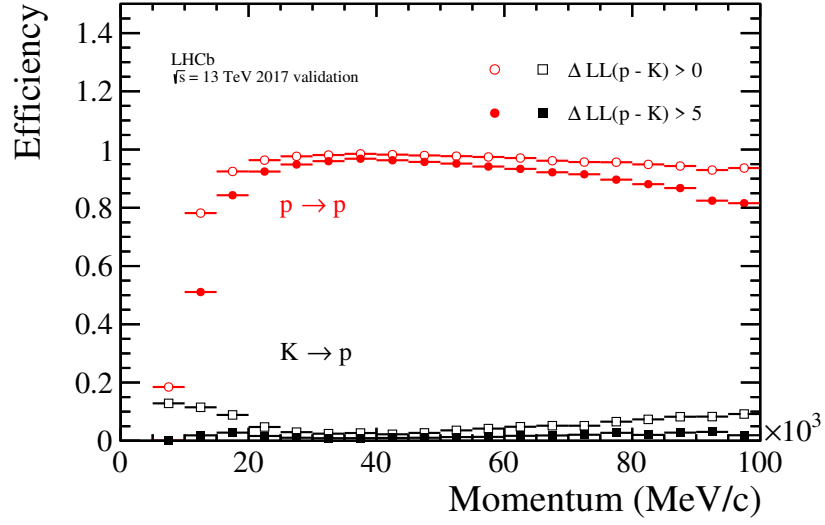


Figure 3.7: Efficiency of (open data points) $DLL_p - DLL_K > 0$ and (filled data points) $DLL_p - DLL_K > 5$ as a function of track momentum for (red) $p \rightarrow p$ and (black) $K \rightarrow p$ identification during the 2017 data-taking period. Figure taken from Ref. [78].

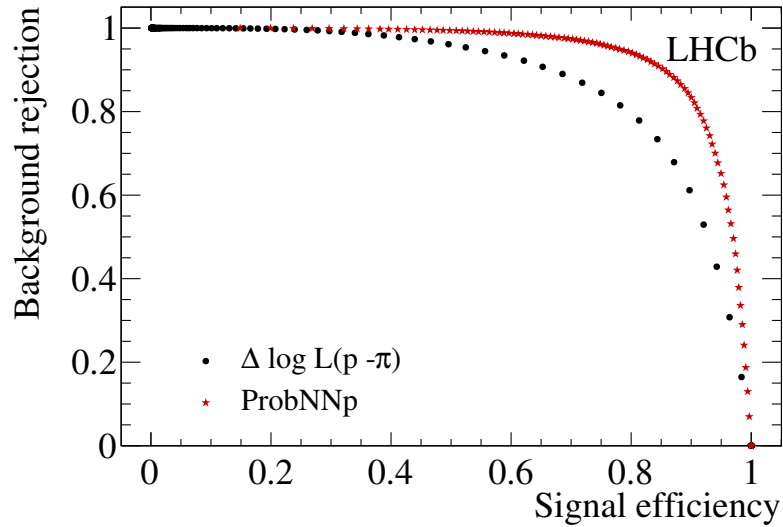


Figure 3.8: Background rejection as a function of signal efficiency as determined from a study which used random combinations of tracks as a background and a simulation sample containing protons as signal. There is a clear improvement upon including the ProbNNp information. Figure taken from Ref. [77].

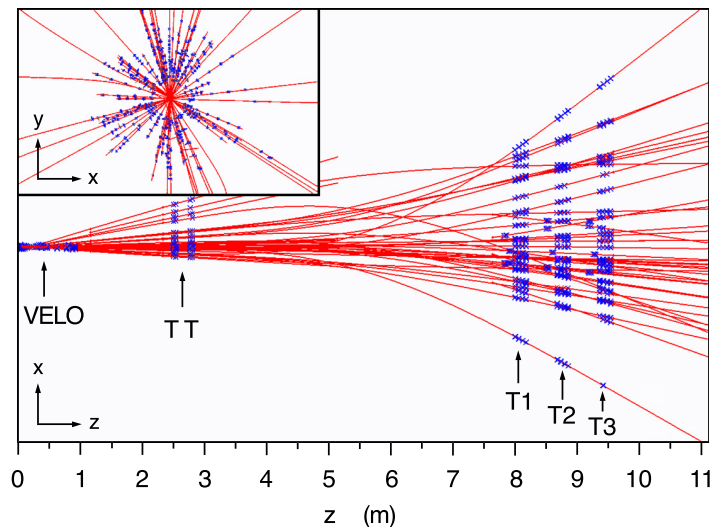


Figure 3.9: A display of the (blue crosses) reconstructed hits of an event in the $x - z$ plane. The insert is a display of the VELO hits in the transverse plane. Figure taken from Ref. [77].

3.5 The tracking stations and magnet

Several other tracking detectors are placed downstream of the VELO. Their main purpose is to measure the trajectories of charged particles. The Tracker Turicensis (TT) lies immediately downstream of the VELO, and upstream of a non-superconducting dipole magnet whose field strength is ≈ 1.1 T at its centre. Further downstream are the tracking stations T1-3.

The field provided by the magnet aligns with the y -axis. When charged particles traverse the field, their trajectories are bent under the action of the Lorentz force. The degree of this bending is a function of the particle's momentum. Thus, accurate measurements of the particle tracks before and after the magnet facilitate momentum measurements. This motivates the orientation of the trackers as shown in Fig. 3.9. Moreover, the polarity of the magnet is inverted periodically such that approximately equal amounts of the data are taken in both field configurations. This is to monitor the performance of both halves of the LHCb detector and to mitigate the effects of any detector asymmetries. The TT and the regions of T1-3 closest to the beam constitute the Inner Tracker (IT) [80], whilst those further from the beam make up the Outer Tracker (OT) [81]. This distinction is motivated by the different technologies used. Precise silicon microstrip sensors are used in the IT due to the greater density of tracks near the beam. Straw tubes filled with gas are used in the OT, which are less precise

than their silicon neighbours, but are also cheaper and achieve a resolution which is sufficient in the lower multiplicity regions.

An algorithm based on a Kalman filter [82] is used to combine the hits in the separate tracking stations into tracks. The efficiency of this track reconstruction is typically greater than 95% [83], which results in a momentum resolution in the range 0.4% to 0.6% [84].

3.6 Calorimetry

Whilst the nominal purpose of calorimeters is to measure the energy of particles, they also provide position and PID information at LHCb. The hadronic calorimeter (HCAL) consists of layers of scintillating plastic inter-spaced between absorber layers of iron [85]. When hadrons traverse the HCAL, there is a higher probability that they interact with the iron absorber layers than the plastic layers. The interaction with the iron nuclei can create hadrons that undergo further interactions in the absorber layers leading to the production of a *hadronic shower*. Scintillation photons are produced in the layers of plastic by the ionizing components of the hadronic shower. These photons are transported along wavelength-shifting fibres to a system of photomultiplier tubes (PMTs).

A similar principle applies to the electromagnetic calorimeter (ECAL), with the main difference being the choice of lead, rather than iron, as the absorber material. This has the advantage of a lower radiation length which results in the ECAL being able to contain the electromagnetic showers entirely. The HCAL is not thick enough to fully contain hadronic showers. Despite this, it retains a sufficient energy deposit in order to be important for hadronic trigger decisions. The ECAL plays a role in distinguishing between γ and π^0 candidates. This will play a role in part of the data selection process, as will be discussed in more detail in Sec. 7.3.

3.7 Muon stations

Owing to their large mass and the fact that energy losses due to material interactions are inversely proportional to the particle mass, muons traverse most of the LHCb detector unimpeded. The muon stations [86] are therefore positioned furthest downstream from the interaction region, with the exception of one of the stations. The M1 station sits between RICH 2 and the ECAL, whilst M2-5 are positioned downstream

of the HCAL.

Broadly speaking, the muon system is designed to simultaneously stop particles (other than muons) from traversing it and to be able to efficiently detect muon tracks. 80cm-thick layers of iron are inter-spaced between systems of multi-wire proportional chambers in M2-5, whilst triple-gas electron multipliers are used in M1. The latter is a technology which is more resistant to radiation damage, given its closer proximity to the interaction region.

The output of the muon stations is used for picking up events which contain muons with high transverse momenta. This is an important criterion for the trigger. PID information can also be inferred from the presence, or lack thereof, of hits in the muon stations for a given track. This is captured by the binary `isMuon` variable. A `TRUE` value of `isMuon` can arise if there are hits left in some of the muon stations for lower momentum tracks, or all of the stations M2-5 for those of higher momentum.

3.8 Trigger system

The 40 MHz crossing rate of the protons at the LHC makes it unfeasible to record every pp interaction. At LHCb, a *triggering* system is used to effectively filter the interactions and retain only those that are of interest for physics analyses. The trigger system is tiered, where increasingly sophisticated techniques are used to infer whether a given event is interesting or not. The low-level (L0) trigger is based on responses from the subdetectors such as the calorimeters or the muon stations that have fast electronics and read-out times.

The high-level trigger is subdivided into two tiers, HLT1 and HLT2, and uses software to perform a reconstruction of the event in real-time. HLT1 uses information from the VELO, whilst HLT2 uses information from all of the subdetectors. An overview of the Run 2 trigger is given in Fig. 3.10. The HLT trigger is gathered into different *lines* which have different criteria in terms of what constitutes an interesting event. This is to facilitate data taking for the broad range of analyses which are undertaken at LHCb. As will be discussed in more detail in Sec. 7, this analysis makes use of a number of trigger lines. After HLT2, the data rate is sufficiently low so as to be suitable for storage.

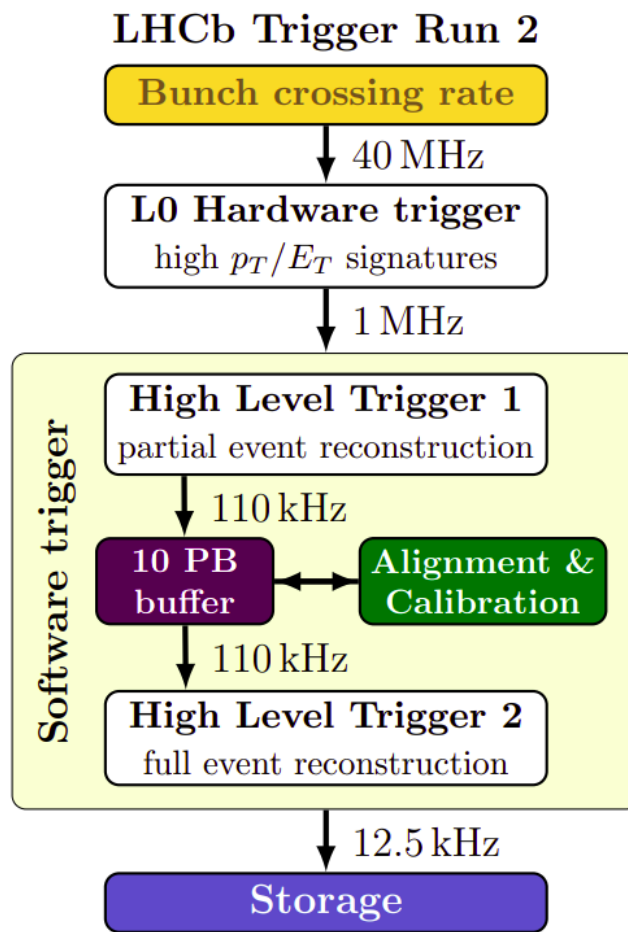


Figure 3.10: Flow chart of the Run 2 trigger system where the reduction in the data rate at each stage is clear. Figure taken from Ref. [87].

Chapter 4

Analysis strategy

This chapter details the methodology for the measurement of $R(p\bar{p})$. Firstly, the rest-frame approximation (RFA) for tauonic decays, which mitigates the experimental impact of the missing final-state neutrinos, is introduced. This is followed by a motivation for the choice of discriminating variables which are used in an extended maximum-likelihood fit to measure $R(p\bar{p})$. A general discussion of the MVA techniques used to help reduce the level of partially-reconstructed (part-reco) backgrounds is then given, before an overview of the FF fit. The procedure for correcting the simulation samples follows. In order to avoid biasing the procedure, efforts are made to not observe the result as the work unfolds; the analysis is performed *blind*. An overview of the protocol used to blind the measurement concludes this section.

4.1 Rest-frame approximation

The τ^+ lepton is reconstructed using the muonic decay $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau \nu_\tau$. The $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau \nu_\tau$ decay is chosen due to its high branching fraction of 17.4% [88] and the fact that the experimentally visible final state of $p\bar{p}\mu$ is the same as that of $B^+ \rightarrow p\bar{p}\mu^+ \nu_\mu$. Therefore $B^+ \rightarrow p\bar{p}\mu^+ \nu_\mu$ is ideally-suited as a normalisation channel for the analysis. However, the presence of three neutrinos in the final state of the $B^+ \rightarrow p\bar{p}(\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau) \nu_\tau$ makes it impossible to determine the four momentum of the B accurately since the neutrinos escape the LHCb detector without interaction. Therefore, it is necessary to estimate the momentum, which can be done by constructing an approximate rest frame for the B .

According to special relativity, the proper velocity of particle X of mass m_X and

momentum p_X is given by

$$\gamma^X \beta^X = \frac{p_X}{m_X}. \quad (4.1)$$

The B rest frame is approximated by assuming that the component of the proper velocity of the visible (vis.) component of the decay along the z -axis $\gamma^{\text{vis}} \beta_z^{\text{vis}}$ is equal to that of the B :

$$\gamma^{\text{vis}} \beta_z^{\text{vis}} \approx \gamma^B \beta_z^B. \quad (4.2)$$

This leads to the following relation between m_B and the invariant-mass m_{vis} of the reconstructed $p\bar{p}\mu$ system:

$$p_z^B \approx \frac{m_B}{m_{\text{vis}}} p_z^{\text{vis}}. \quad (4.3)$$

In conjunction with knowledge of the unit-vector which points along the line of flight of the B , this allows the other components of p^B to be estimated. Denoting the angle subtended by this vector with the z -axis as α , the magnitude of the momentum is given by

$$|p^B| = \frac{m_B}{m_{\text{vis}}} (p^{\text{vis}})_z \sqrt{1 + \tan^2 \alpha}. \quad (4.4)$$

Finally, the so-called corrected mass M_{corr} is introduced:

$$M_{\text{corr}} = \sqrt{m_{\text{vis}}^2 + (p_T^{\text{vis.}})^2} + |p_T^{\text{vis.}}|, \quad (4.5)$$

where $m_{\text{vis.}}^2$ is the squared invariant mass of the visible decay products and $p_T^{\text{vis.}}$ is the transverse momentum of the visible decay products with respect to the line of flight of the B . The corrected mass was first defined in Ref. [89] and is a re-definition of the visible mass that manifestly includes information about the missing momentum carried away by any ‘invisible’ decay products. In the case of a single final-state neutrino from a B decay, the corrected mass exhibits a peak at $M_{\text{corr}} \approx M_B = 5279 \text{ MeV}$. For semileptonic decays, M_{corr} is favoured over $m_{\text{vis.}}$ as it provides greater separation of signal from part-reco background.

4.2 Choice of fit variables

Ultimately, the analysis presented in this thesis requires performing a 3D extended maximum-likelihood fit to the LHCb data distributions in a range of variables, with $R(p\bar{p})$ being the parameter of interest. The nature of the $B^+ \rightarrow p\bar{p}(\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau) \nu_\tau$ final state being an open system due to the missing neutrinos results in a resolution

that leads to the presence of a number of backgrounds and making a conventional fit to the invariant mass of the decay products unfeasible. It is therefore of paramount importance to choose fit variables which lend themselves to distinguishing between signal and background.

The choice of fit variables is motivated by exploiting the discriminating power of the kinematics associated with the $B^+ \rightarrow p\bar{p}(\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau)\nu_\tau$ decay chain. The squared missing-mass $m_{\text{miss}}^2 = (p_B - p_p - p_{\bar{p}} - p_\mu)^2$ aims to capture the missing four-momentum resulting from the presence of the three final-state neutrinos. The squared dilepton invariant-mass $q^2 = (p_B - p_p - p_{\bar{p}})^2$ exploits the characteristics of the decay of the virtual $W^\pm$. Finally, the muon energy in the B rest-frame E_μ^* is chosen as the muons originating from the τ decay tend to be softer than those originating directly from a B decay.

4.3 Form-factor fit

As discussed in Sec. 2.7, the pQCD model as presented in Ref. [40] uses FFs taken from other fully hadronic decay modes with a dibaryonic final state. This is clearly a crude approximation. This is improved upon by using a fit to the Run 1 data to determine estimates for the FF parameters for $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays. The motivation for doing this is to produce accurate simulation samples for the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ for the fit to the Run 2 data sample.

In analogy with the $R(p\bar{p})$ fit, a 3D extended maximum-likelihood fit to q^2 , m_{miss}^2 and E_μ^* is used to determine the FF parameters. The q^2 spectrum is particularly influenced by the FFs. This can be understood by thinking of the $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays as two-body decay in the B rest frame to a virtual W of invariant mass q and a diproton system. There will clearly be a high correlation between the W and the $p\bar{p}$ in this realisation of the system. Therefore, the q^2 spectrum will manifestly be sensitive to the FFs, which encode the dependence of the matrix element on the kinematics of the $p\bar{p}$ system.

4.4 Removal of partially-reconstructed decays

A major challenge in semileptonic analyses concerns the removal of part-reco backgrounds. These decays lead to a (visible) final state of the form $p\bar{p}\mu X$, where X is a placeholder for any number of particles which escape the reconstruction. An example

of such a decay is $B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu$, where the N^* resonances are simply orbitally excited versions of the nucleons, and can be charged or neutral. Likewise, the missing component X can be charged or neutral.

The aim of *charged isolation* is to identify those decays to a $p\bar{p}\mu$ state which are well isolated from other charged particles. That is, there are no VELO tracks which are likely to have originated from the B decay vertex. Equivalently, this can be expressed as the attempt to remove ‘suspicious’ tracks in the vicinity of the SV. This is done with the use of a MVA called a Boosted Decision Tree (BDT) to classify the tracks as such. The BDT takes various properties of the track as inputs, such as the p_T of the track and the opening angle of the track and the $p\bar{p}\mu$ system. As an output, the BDT provides a probability that a given track originated from the SV. In fact, the BDT is implemented such that the probabilities of the four most suspicious tracks are saved. In principle, the component X can include more than simply one additional track.

This analysis makes use of the BDT that was implemented for the LHCb measurement of the CKM element $|V_{ub}|$ [90]. The BDT was trained using simulation samples of $\Lambda_b^0 \rightarrow (\Lambda_c^+ \rightarrow pX)\mu^-\bar{\nu}_\mu$ and $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ as signal and background samples, respectively. In that analysis, the charged isolation BDT (charged BDT) was found to remove 90% of background tracks, whilst retaining 80% of the signal. More details of the implementation for the $R(p\bar{p})$ analysis follow in Sec. 7.2.

In addition to this, a novel isolation technique was developed for the measurement of $R(p\bar{p})$ concerning the removal of part-reco decays with additional neutral particles; *neutral isolation*. An example of such a decay is $B^+ \rightarrow p\bar{\Lambda}_c^-(\rightarrow \bar{p}K^0\pi^0)\mu^+\nu_\mu$. These backgrounds are particularly problematic due to both their relatively high BFs and the fact that the efficiency of reconstructing π^0 or γ at LHCb is worse than that of reconstructing tracks.

Neutral objects leave signatures in the ECAL, either through the showering of a γ or through the strong decay of $\pi^0 \rightarrow \gamma\gamma$, and a subsequent shower. The neutral isolation BDT (neutral BDT) works in a similar manner to the charged isolation BDT. Naturally, in contrast to considering the tracks in the VELO, the π^0 and γ candidates in the ECAL are used for the neutral isolation. A BDT is trained to determine whether π^0 or γ candidates are likely to have originated from the SV or not. By taking various kinematic properties of the neutral candidates as inputs, such as their p_T with respect to the z -axis and also with respect to the B momentum, the BDT assigns a probability to the γ or π^0 which reflects how likely it is that they originated from the SV. More details on the implementation of the neutral BDT are presented in Sec. 7.3.

4.5 Corrections to simulation

It is of paramount importance to minimise the differences observed between the real data produced at LHCb, and the simulation of the collisions and the subsequent detector response. This is due to the fact that the simulation samples are used to construct the probability density functions (PDFs) for the likelihood fits, and to calculate accurate efficiency ratios for the determination of $R(p\bar{p})$. Any poor description of a particular aspect of the data could result in a poor fit quality, inaccurate efficiencies, or large systematic uncertainties. The corrections take the form of a re-weighting of

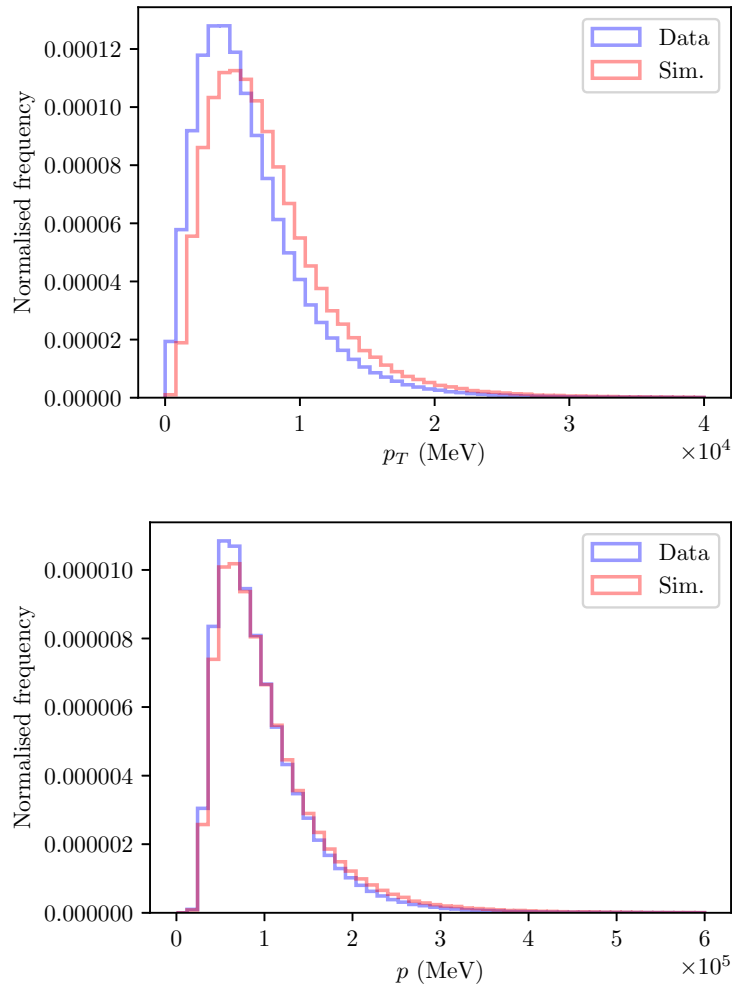


Figure 4.1: Comparison of the (top) $p_T(B)$ and (bottom) $p(B)$ distributions for $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ (blue) data and (red) simulation from 2018. The B mesons are harder in simulation than in data.

distributions. For example, it is known that the spectra of the B p_T produced by the

PYTHIA event generator [91], which is used to simulate the pp collision process, are too high. This is shown in Fig. 4.1. However, it is possible to select a pure sample of $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ because of the high selection efficiency for this decay chain. The B kinematics from the simulation samples can then be compared with those of the pure data sample. A correction weight can be applied which is the ratio of the two spectra.

A number of such corrections are applied. In addition to the B kinematics, corrections are applied to account for the inaccurate PID efficiencies, tracking efficiencies, trigger efficiencies, as well as corrections to the kinematics of some of the background modes [83, 92].

4.6 Calculating $R(p\bar{p})$

The ratio $R(p\bar{p})$ can be expressed as a function of the efficiencies $\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)$ and $\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)$ and the yields $\mathcal{N}(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)$ and $\mathcal{N}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)$ as

$$R(p\bar{p}) = \frac{\mathcal{N}(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)} \times \frac{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)}{\mathcal{N}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)}. \quad (4.6)$$

$R(p\bar{p})$ is extracted directly from the fit to data as the parameter of interest (POI), whilst the ratio of efficiencies $\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)/\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)$ is calculated from the simulation samples with the corrective weights applied. This will be discussed in more detail in Sec. 9.

4.7 Result blinding

Whilst performing the analysis, the numerical value of $R(p\bar{p})$ is blinded, as is convention across LHCb analyses. This is to avoid biasing the selection of the data and the development of the fit. More specifically, the value of $R(p\bar{p})$ from the fit is multiplied by a random number between 0.5 and 1.5 for the duration of the analysis.

Chapter 5

Data samples and simulation corrections

This chapter details the data and simulation samples used for the analysis. An overview of the simulation process is followed by an account of the corrections made to the simulation, including the reweighting of the B kinematics and corrections to the PID, trigger and tracking efficiencies. There are also corrections made to the kinematics of the background processes that are expected to be dominant.

5.1 Data samples

The analysis presented in this thesis makes use of the data collected with the LHCb detector in 2016, 2017 and 2018 at $\sqrt{s} = 13 \text{ TeV}$ when a dedicated trigger line to select the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ final state was in place. This was not the case in 2015, hence these data are not used. The Run 1 data are not used in order to preserve statistical independence from the measurement of the form factor parameters described in Sec. 8. The integrated luminosity of the data is approximately 5.1 fb^{-1} and the cross-section for $B^\pm$ production was $\sigma(pp \rightarrow B^\pm X) \approx 86.6 \mu\text{b}$ [93]. This corresponds to approximately 4.8×10^{11} $B^\pm$ pairs in the LHCb detector acceptance. After applying the full chain of selection criteria, the sample consists of approximately 3×10^5 candidate decays.

5.2 Simulation samples

Several pieces of software are used to simulate events at LHCb. The PYTHIA package [91] is used for the simulation of the pp collisions and for the production and subsequent hadronisation of the b quarks. The decays of the $B^\pm$ hadrons themselves are simulated by the EVTGEN [94] package. Within EVTGEN, there are several theoretical models that can be used to simulate $B^\pm$ decays to a broad range of daughter particles. The pQCD model described in Sec. 2.7 was incorporated into EVTGEN in order to simulate $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays. In the absence of an appropriate model for a given decay, the $B^\pm$ is said to be simulated ‘according to phase space (PHSP)’. In reality, this simply means that the matrix element is assumed to be a constant and the decay is allowed to proceed provided the decay product four-momenta fall within the available phase space for the decay. Phase space modelling was used for some of the part-reco backgrounds, and the drawbacks of this approach are discussed at the end of this section.

The simulation of the LHCb detector is performed using the GEANT4 [95] package. Prior to the simulation of the detector response, the particle four-momenta are referred to as being at truth-level or generator-level (gen-level). All samples were generated using 2018 conditions for this analysis. A systematic uncertainty, described in Sec. 11, is estimated in order to capture the shortcomings of this approach.

5.3 Corrections to B kinematics

The kinematics of B -mesons produced at LHCb are imperfectly simulated. Notably, the momentum $p(B)$ and transverse-momentum $p_T(B)$ are known to be relatively poorly reproduced by the simulation. These distributions are corrected in the simulation by applying a set of kinematic weights w_{kin} derived from $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ data samples. A fit to the $m(K^+\mu^+\mu^-)$ distribution, obtained from the $B^+ \rightarrow J/\psi K^+$ data samples, is performed using the *sPlot* technique [96]. This allows the $p(B)$ and $p_T(B)$ distributions to be obtained.

The fit to the $m(K^+\mu^+\mu^-)$ distribution is shown in Fig. 5.1. The $B^+ \rightarrow J/\psi K^+$ simulation sample is reweighted to the data sample in $p(B)$ and $p_T(B)$ at gen-level. The corrections are applied at gen-level to all simulation samples because the B^+ meson is not fully-reconstructed in semileptonic decays. The results of this reweighting are shown in Figure 5.2. The reweighting was carried out using GBReweighter [97]. This is a BDT-based approach, which enables the reweighting of simulation templates in

multiple variables simultaneously.

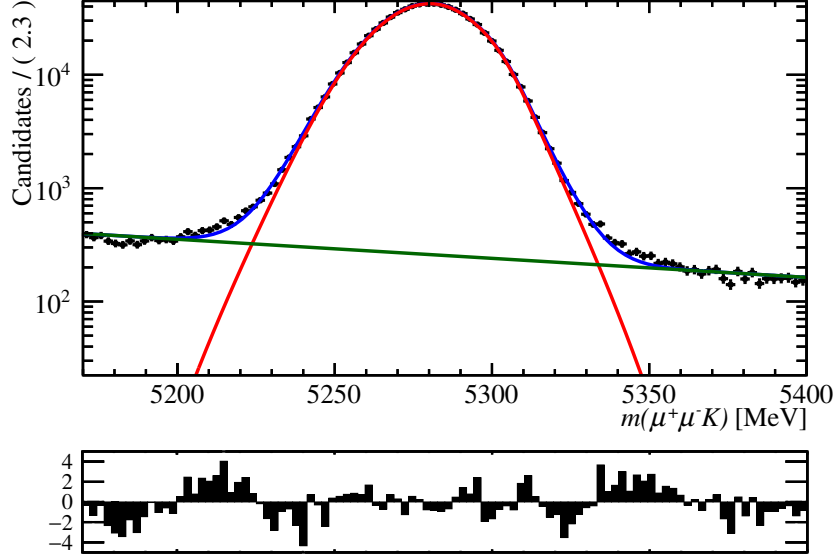


Figure 5.1: Fits to the $m(\mu^+\mu^-K^+)$ distribution from the 2018 $B^+ \rightarrow J/\psi K^+$ data used for the kinematic corrections of the simulation.

5.4 Trigger corrections

There are also inaccuracies in the simulation of the trigger system at LHCb, and a correction is applied. The correction of the L0Muon trigger efficiency also takes the form of a weight w_{trig} . These weights are taken to be the ratio of efficiencies measured in data and simulation. In analogy with the kinematic corrections, an sWeighted $B^+ \rightarrow J/\psi K^+$ data sample is used to calculate the L0Muon efficiency using the so-called TISTOS technique. This allows the determination of the efficiency ϵ_{TOS} of a trigger-on-signal (TOS) trigger decision according to

$$\epsilon_{\text{TOS}} = \frac{\epsilon_{\text{TIS\&TOS}}}{\epsilon_{\text{TIS}}}, \quad (5.1)$$

where $\epsilon_{\text{TIS\&TOS}}$ is the efficiency for the event to be both triggered-independent-of-signal (TIS) and TOS, and ϵ_{TIS} is the TIS efficiency.

In a given event, the trigger can be fired by any of the tracks. TOS refers to the case when the trigger is fired by part of the $p\bar{p}\mu^+$ system which forms the signal candidate. TIS refers to the case when the trigger is fired by a track other than one

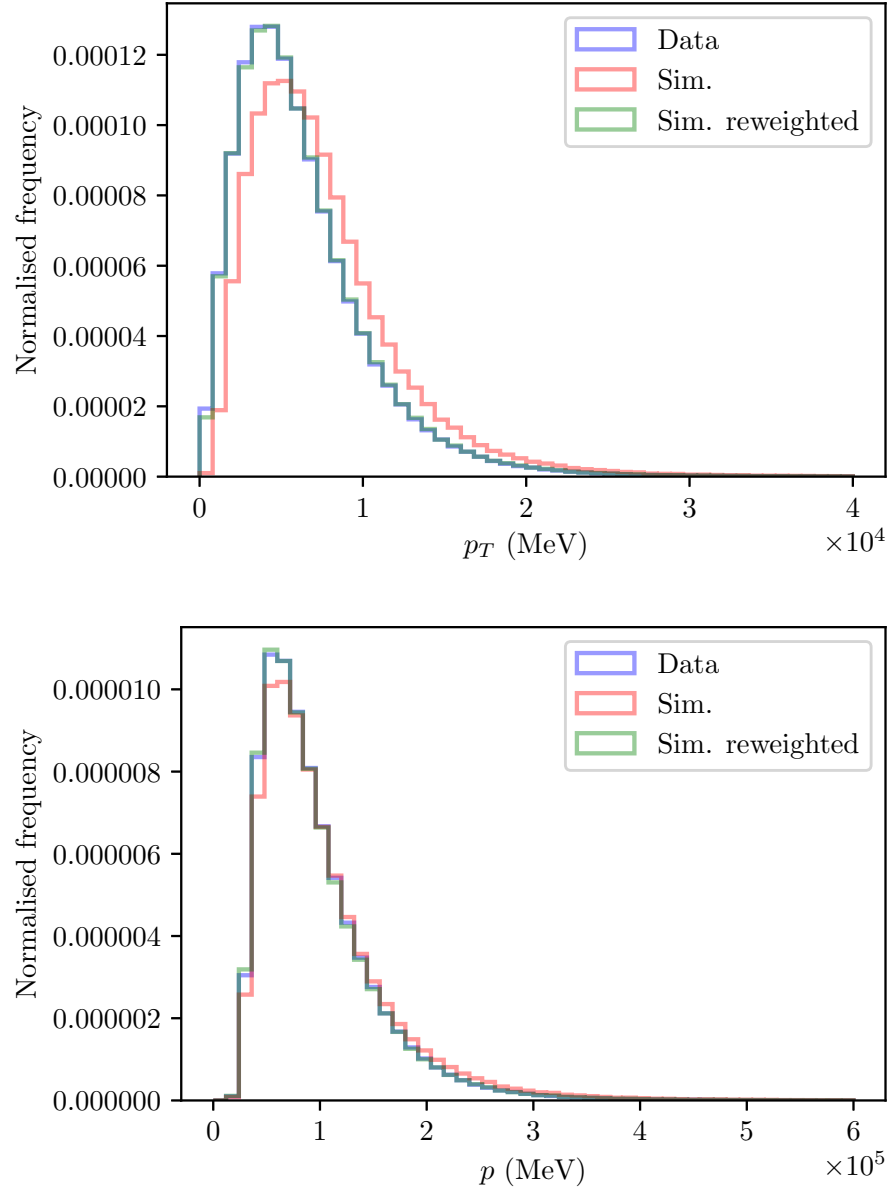


Figure 5.2: A comparison of the $B^+ \rightarrow J/\psi K^+$ simulation before and after reweighting with the sWeighted $B^+ \rightarrow J/\psi K^+$ data in $p_T(B)$ (left) and $p(B)$ (right) for the 2018 samples.

of these. TIS&TOS denotes the case where these two scenarios are both true.

This method is used to calculate the efficiency $\epsilon_{\text{TOS}}^{\text{data}}$ in data and $\epsilon_{\text{TOS}}^{\text{MC}}$ in simulation as a function of the transverse muon momentum $p_{\text{T}}(\mu)$ for the muon with the same sign as the kaon. The L0Muon trigger efficiency is particularly sensitive to $p_{\text{T}}(\mu)$ as the muon chambers rely on the muon traversing the muon stations. This is shown in Fig. 5.3, and there is generally good agreement save for some larger discrepancies towards lower values of $p_{\text{T}}(\mu)$. Whilst this may contribute to some reduction of the signal statistics for the softer muons originating from the $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ decay, the high branching fraction for this process of around 17% is expected to compensate for the potential loss of statistics. The set of weights w_{trig} , defined as

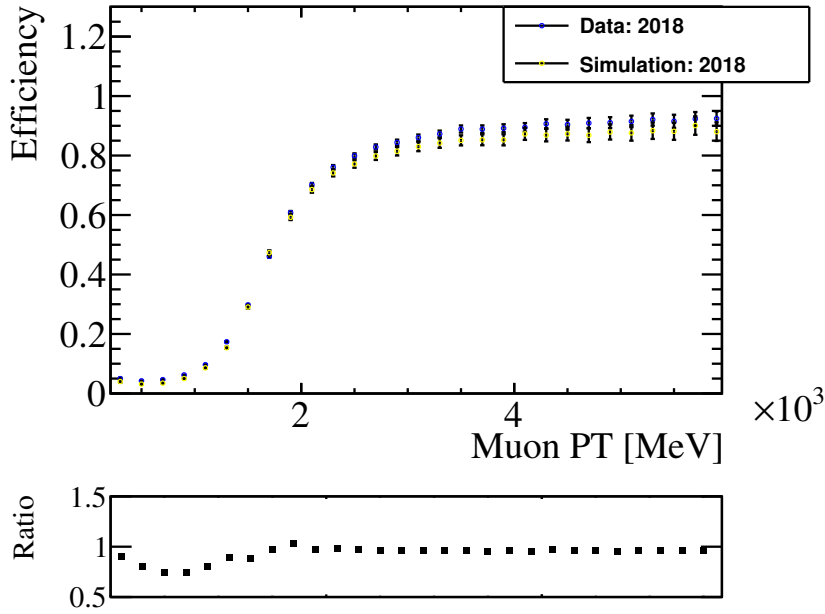


Figure 5.3: The L0Muon trigger efficiency as determined from $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+$ (blue) data and (yellow) simulation samples as a function of the transverse momentum of the muon with the same sign as the kaon. The weight applied to the simulation is the ratio of efficiencies.

$$w_{\text{trig}} = \frac{\epsilon_{\text{TOS}}^{\text{mc}}}{\epsilon_{\text{TOS}}^{\text{data}}}, \quad (5.2)$$

are used to correct the simulation samples in bins of $p_{\text{T}}(\mu)$.

5.5 Tracking corrections

The tracking efficiency is also known to be inaccurate in the simulation. A correction is applied in the form of a product of weights for each of the three tracks in the final-state, yielding

$$w_{\text{track}} = w_{\text{track}}^p \times w_{\text{track}}^{\bar{p}} \times w_{\text{track}}^{\mu^+}. \quad (5.3)$$

These weights represent the ratio of the tracking efficiency in simulation and data in bins of p and η . The production of these efficiency ratios is done centrally at LHCb [83], and their distribution is shown in Figure 5.4. The agreement is generally good across

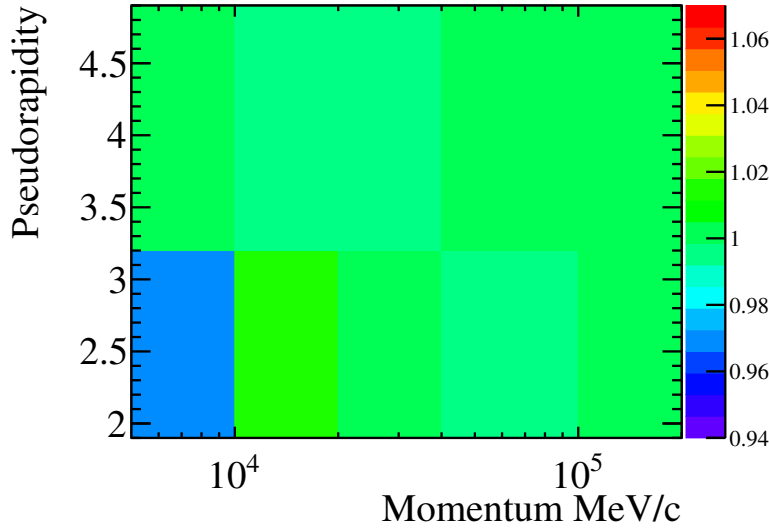


Figure 5.4: The ratio of tracking efficiencies in data and simulation in the $p - \eta$ plane taken from the calibration histograms [83]

the $p - \eta$ plane. However, for momenta $p \leq 10$ GeV, the tracking efficiency is slightly underestimated.

5.6 Particle identification corrections

During the selection (described in more detail in Sec. 7.4), a cut is placed on a so-called ProbNN combination for each of the final-state particles. These variables are derived from a Neural Network (NN) which provides a probability that a given particle is a certain type: p , K , π , e or μ . A probability ProbNNghost is also assigned to reflect how likely the track is to be a fake or *ghost* track. This arises when hits from separate

particles or noise in the subdetector systems mimic a real track.

However, the PID performance of the LHCb detector is poorly simulated. Specifically, the interaction of particles other than those of the signal, or the *underlying event*, with the RICH detectors is inaccurately reproduced by the simulation. Thus, rather than placing cut-based criteria on the PID variables from simulation, a weight is applied to each candidate from simulation which represents the efficiency of such a cut.

The composite variables $\text{ProbNN}_p^{\text{combined}}$ and $\text{ProbNN}_\mu^{\text{combined}}$ are defined as

$$\begin{aligned} \text{ProbNN}_p^{\text{combined}} = & \text{ProbNN}p(1 - \text{ProbNN}\pi)(1 - \text{ProbNN}K)(1 - \text{ProbNN}e) \\ & \times (1 - \text{ProbNNghost}), \end{aligned} \quad (5.4)$$

$$\begin{aligned} \text{ProbNN}_\mu^{\text{combined}} = & \text{ProbNN}\mu(1 - \text{ProbNN}\pi)(1 - \text{ProbNN}K)(1 - \text{ProbNN}e) \\ & \times (1 - \text{ProbNN}p)(1 - \text{ProbNNghost}). \end{aligned} \quad (5.5)$$

The PIDCalib [92] package is used to determine the PID efficiency. At LHCb, this is a standard tool used to determine the PID efficiency of a given criterion using calibration samples from data. For example, a sample of $\Lambda_c^+ \rightarrow pK^-\mu^+$ is used to determine PID efficiencies on protons. This is a decay which is readily accessible experimentally where the decay products can be identified based on kinematics alone. The efficiency of a given cut can be determined with a fit to the invariant mass $m(pK^-\mu^+)$ of the final-state particles. A sample of $J/\psi \rightarrow \mu^+\mu^-$ decays is used for muons for analogous reasons.

The PID performance depends primarily on where the track is in the detector, i.e. its position in η , and also its momentum p . Consequently, the efficiency is determined as a histogram in bins of p and η , which is shown for both protons and muons in Fig. 5.5.

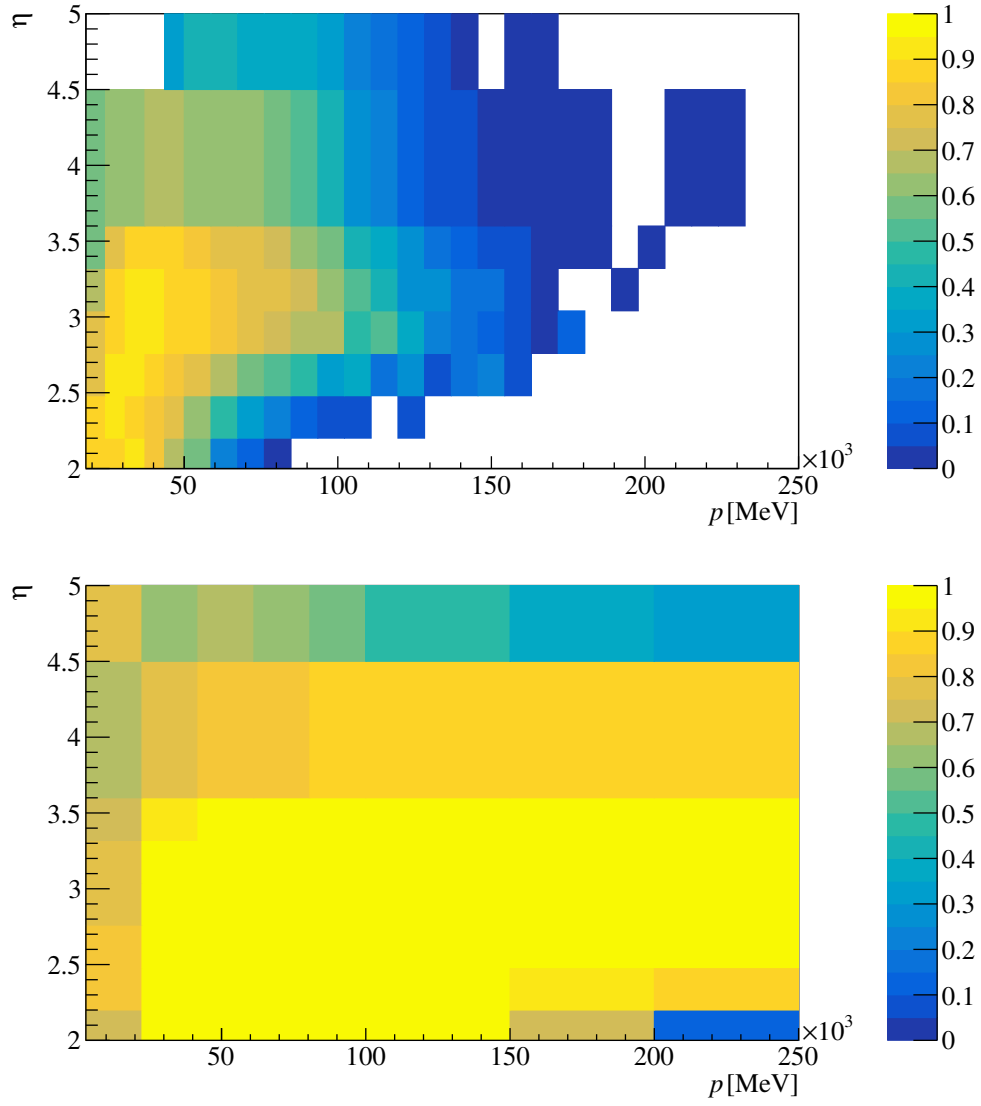


Figure 5.5: Projections of the PID efficiency as determined using PIDCalib in the $p - \eta$ plane for (top) protons with $\text{ProbNN}_p^{\text{combined}} > 0.5$ and (bottom) muons with $\text{ProbNN}_\mu^{\text{combined}} > 0.4$, using calibration samples from 2018.

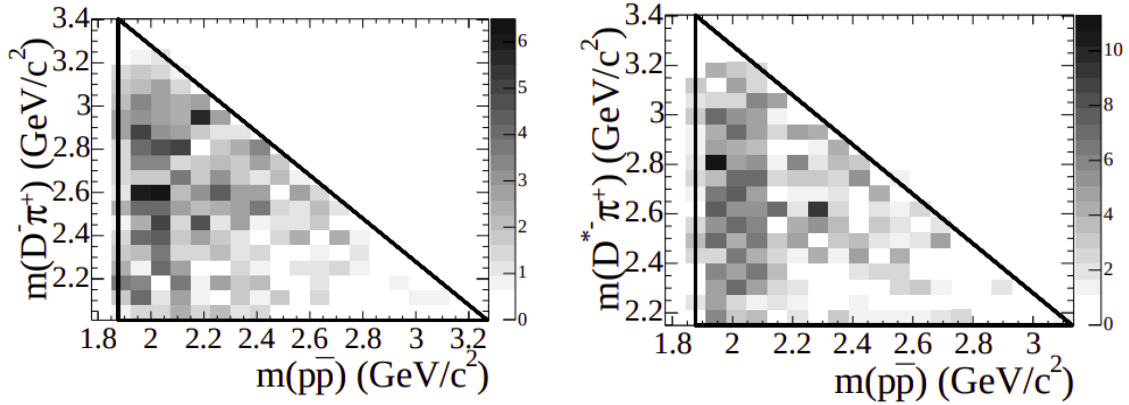


Figure 5.6: Differential branching fractions in (left) $m(D^-\pi^+)$ vs. $m(p\bar{p})$ for $B^0 \rightarrow D^- p\bar{p}\pi^+$ and (right) $m(D^{*-}\pi^+)$ vs. $m(p\bar{p})$ for $B^0 \rightarrow D^{*-} p\bar{p}\pi^+$ as observed by the BaBar collaboration. There is a clear enhancement near the kinematic threshold in $m(p\bar{p})$ in both cases. Figure taken from Ref. [98].

5.7 Corrections to background kinematics

As mentioned in Sec. 5.2, in the absence of an appropriate physics model for a given decay mode, the PHSP approximation is used. However, much of the structure in the kinematic distributions of the decay products is lost when final-state particles are missed in the reconstruction of part-reco backgrounds. Consequently, there is less importance placed on the necessity of accurate physics models for part-reco backgrounds.

That being said, corrections are made where possible, particularly for the part-reco decays that are expected to be dominant. Decays of the form $B \rightarrow p\bar{p}DX$ can be approximated as a two-body decay to a $p\bar{p}$ and DX system. This is analogous to thinking of the $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays as a two-body decay to a $p\bar{p}$ pair and a virtual $W^\pm$. In both cases, a threshold enhancement effect is observed in the $m(p\bar{p})$ spectrum, as seen in both Fig. 1.3 for $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ and in Fig. 5.6 for $B \rightarrow p\bar{p}DX$ decays.

A similar argument applies to decays of the form $B^+ \rightarrow p\bar{\Lambda}_c^-\mu^+\nu_\mu$, where an enhancement is expected towards the $m(p\bar{\Lambda}_c^-)$ kinematic threshold. As a result, the $m(p\bar{p})$ and $m(p\bar{\Lambda}_c^-)$ distributions of the $B \rightarrow p\bar{p}DX$ and $B^+ \rightarrow p\bar{\Lambda}_c^-\mu^+\nu_\mu$ simulation samples, respectively, are reweighted to the $m(p\bar{p})$ distribution from the pQCD model. The results of this reweighting are shown in Fig. 5.7. A systematic uncertainty associated with this reweighting is described in Sec. 11.9.

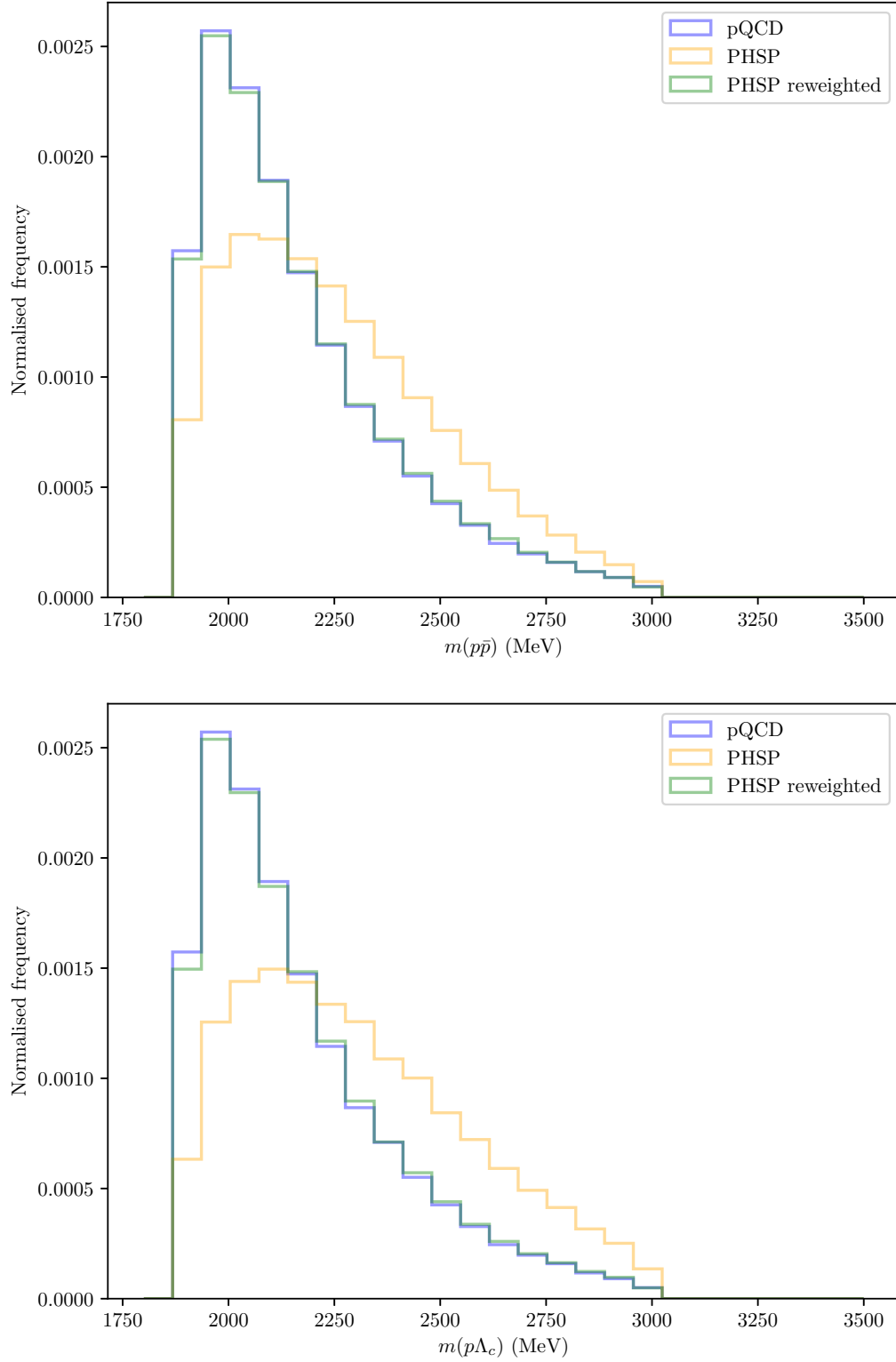


Figure 5.7: (Top) $m(p\bar{p})$ and (bottom) $m(p\Lambda_c^+)$ distributions before and after reweighting to the pQCD model for the $B \rightarrow p\bar{p}DX$ and $B^+ \rightarrow p\bar{\Lambda}_c^-\mu^+\nu_\mu$ simulation samples.

Chapter 6

Backgrounds

As was seen for the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ BF analysis, the dominant source of background for $R(p\bar{p})$ is partially-reconstructed (part-reco) decays. The composition of the simulation samples used to model these decays will be described in this chapter. A subdominant source of background originates from other particles, such as a π or K being misidentified as either a proton or a muon. A data-driven method is used to derive a shape for these components in the fit. It is also possible to constrain their yields through an unfolding procedure which will be described in greater detail later in this section. Finally, combinatorial background arises when combinations of random tracks are mistaken to have originated from a B^+ decay. The treatment of this source of background with a data-driven method is also described.

6.1 Partially reconstructed backgrounds

Part-reco background arises when additional particles X to a $p\bar{p}\mu X$ final-state are not reconstructed. There are a number of such decays which are possible, and they are summarised in Table. 6.1. Many of the backgrounds are poorly understood, and in some cases the BF is not known. The dominant channels are those that involve a $b \rightarrow c$ transition. The BFs for $B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$, $B^+ \rightarrow p(\bar{\Lambda}_c^{*-} \rightarrow \bar{p}X)\mu^+\nu_\mu$, $B^+ \rightarrow p\bar{p}DX$, $B^+ \rightarrow p\bar{p}D\mu^+\nu_\mu$ and $B^0 \rightarrow (\bar{\Lambda}_c^- \rightarrow \mu^-pX)\bar{p}X$ are comparable at the level $\approx 10^{-4}$, with relatively high selection efficiencies at the stripping level. However, the charged isolation is highly effective at removing these components. The neutral isolation also serves to remove the components with additional neutrals in the final state. This is outlined in more detail in Secs. 7.2 and 7.3. The Cabibbo-suppressed decays involving a $b \rightarrow u$ transition $B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu$ and $B^+ \rightarrow p(\bar{\Delta} \rightarrow \bar{p}X)\mu^+\nu_\mu$ are

Decay	Decay model	BF order
$B^{+/-} \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X$	PHSP	$\approx 10^{-3}$
$B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$	PHSP	$\approx 10^{-4}$
$B^0 \rightarrow (\bar{\Lambda}_c^- \rightarrow \mu^-pX)\bar{p}X$	PHSP	$\approx 10^{-4}$
$B^+ \rightarrow p(\Delta \rightarrow \bar{p}X)X\mu^+\nu_\mu$	pQCD	$\approx 10^{-6}$
$B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu$	pQCD	$\approx 10^{-6}$
$B^+ \rightarrow p(\bar{\Lambda}_c^{*-} \rightarrow \bar{p}X)\mu^+\nu_\mu$	PHSP	$\approx 10^{-4}$
$B^+ \rightarrow p\bar{p}D^0\mu^+\nu_\mu X$	PHSP	$\approx 10^{-4}$
$B_s^0 \rightarrow (D_s^+ \rightarrow \mu^+X)p\bar{p}X$	PHSP	$\approx 10^{-3}$
$B_s^0 \rightarrow (D^0 \rightarrow K\mu^+X)p\bar{p}X$	PHSP	$\approx 10^{-4}$

Table 6.1: A summary of the part-reco backgrounds for $R(p\bar{p})$, where missing final-state particles are represented by X . The decays are listed in an approximate order of their expected dominance.

subdominant given their lower BF's which, although unmeasured, are estimated to be $\approx 10^{-6}$, i.e. a similar order of magnitude to the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ BF. Their removal is also facilitated by the isolation tools.

The B_s^0 mode BF's are not known but analogous decays with a $\pi^+\pi^-$ system in the final state rather than a $p\bar{p}$ system have BF's in the region of $10^{-3} - 10^{-4}$ [88]. There is potential for this to lead to a significant contribution from these modes in the sample. However, both the presence of additional charged tracks and the fact that there are approximately 75% fewer B_s^0 produced at LHCb compared to $B^\pm$ [99] will lead to a significant reduction in their contribution.

6.2 Misidentified proton background

The mis-identification of other particles, such as a π or K , as either a p or μ is a subdominant background for this analysis. The procedure for treating these backgrounds with a data-driven method will be described in this section.

The candidates satisfying the requirements of a stripping line, which is akin to that of the nominal data sample but where PID requirements are placed only on one of the protons, is referred to as the fake- p sample. It is expected that the fake- p tracks to fall into three main categories; either π -like, K -like or p -like. μ -like and e -like tracks

Category	Criteria
p -like	$\text{ProbNN}_p^{\text{combined}} > 0.5$
K -like	$\text{ProbNN}_K^{\text{combined}} > 0.3 \ \& \ !p\text{-like}$
π -like	$\text{ProbNN}_\pi^{\text{combined}} > 0.3 \ \& \ !p\text{-like} \ \& \ !K\text{-like}$
U	$!\pi\text{-like} \ \& \ !p\text{-like} \ \& \ !K\text{-like}$

Table 6.2: The scheme used to form the track categories for the fake- p sample.

in this sample are not considered. These are anticipated to be a negligible component given that the stripping line ensures that the proton tracks do not satisfy the isMuon criterion. A remaining component consists of so-called ‘uncategorised’ (U) tracks, which is anticipated to be largely dominated by ghost tracks. The criteria of this categorisation are given in Table. 6.2.

It is possible to estimate the number of true p , K , π and U tracks in the fake- p sample as follows. Taking the p -like tracks as an example, the probability of a particle X to be mis-identified as a p is reflected by the efficiency $\epsilon(X \rightarrow p)$. As was discussed in Sec. 5.6, this efficiency depends on the momentum and pseudorapidity of the track. In a given bin of p and η , the total probability $p(X \rightarrow p)$ of a track to be mis-identified is given by

$$p(X \rightarrow p) = \frac{N_K \epsilon(K \rightarrow p) + N_\pi \epsilon(\pi \rightarrow p) + N_U \epsilon(U \rightarrow p)}{N_K + N_\pi + N_U}, \quad (6.1)$$

where N_X is the true number of particles of type X . This probability is applied as a weight to each of the candidates in the fake- p sample in order to derive a template for this component in the fit.

Additionally, the sum of these weights, multiplied by an appropriate factor which accounts for the *prescaling* of the stripping line can act as a constraint on the p mis-ID yield in the fit. A stripping line is prescaled for looser selections in order to keep the retention rate low such that it is feasible to keep the candidate data on tape. Prescales lie in the range $0 - 1$, and e.g. for a prescale of 0.2, 80% of the candidates are thrown away at random.

The problem of determining the true numbers N_X of each track type on a bin-by-bin basis can be cast in terms of the observed numbers N'_X in the following matrix

form:

$$\begin{pmatrix} N'_\pi \\ N'_K \\ N'_p \\ N'_U \end{pmatrix} \approx \begin{pmatrix} \epsilon(\pi \rightarrow \pi) & \epsilon(K \rightarrow \pi) & \epsilon(p \rightarrow \pi) & \epsilon(U \rightarrow \pi) \\ \epsilon(\pi \rightarrow K) & \epsilon(K \rightarrow K) & \epsilon(p \rightarrow K) & \epsilon(U \rightarrow K) \\ \epsilon(\pi \rightarrow p) & \epsilon(K \rightarrow p) & \epsilon(p \rightarrow p) & \epsilon(U \rightarrow p) \\ \epsilon(\pi \rightarrow U) & \epsilon(K \rightarrow U) & \epsilon(p \rightarrow U) & \epsilon(U \rightarrow U) \end{pmatrix} \begin{pmatrix} N_\pi \\ N_K \\ N_p \\ N_U \end{pmatrix}. \quad (6.2)$$

Denoting the matrix of true track numbers as $\mathbf{N}$, the efficiency matrix as $\mathbf{E}$ and observed numbers as $\mathbf{N}'$, Eq. (6.2) can be written more concisely as

$$\mathbf{N}' \approx \mathbf{E}\mathbf{N}. \quad (6.3)$$

It follows that determining the matrix $\mathbf{N}$ reduces to a problem of matrix inversion:

$$\mathbf{N} \approx \mathbf{E}^{-1}\mathbf{N}'. \quad (6.4)$$

The matrix $\mathbf{E}$ is derived from calibration samples using the PIDCalib package for the π -like, K -like and p -like categories. No such calibration samples exist for the U -like tracks. The efficiencies $\epsilon(X \rightarrow U)$ are determined by requiring that the sum of each columns in $\mathbf{E}$ sum to 1, i.e. that the track must be identified as one of the four categories. A systematic uncertainty is assigned in order to evaluate the accuracy of this assumption, and is described in more detail in Sec. 11. The method of Bayesian unfolding [100] is used in order to perform the matrix inversion. By treating the matrix $\mathbf{E}$ as a response matrix, an iterative method from the RooBayesUnfold package is used [101]. The distributions of the weights defined in Eq. (6.1) are shown in Fig. 6.1.

6.3 Misidentified muon background

The treatment of the μ mis-ID background is along similar lines to that of the p mis-ID. However, the yield resulting from the unfolding method is not reliable. This is due to the fact that the muonic trigger decisions are not applied to the fake- μ sample, in line with the absence of PID criteria applied to this sample. Hence, the template shape is taken from the unfolding procedure, while the yield is left to float in the fit.

The categorisation of the fake- μ differs slightly from the hadronic case. It is as-

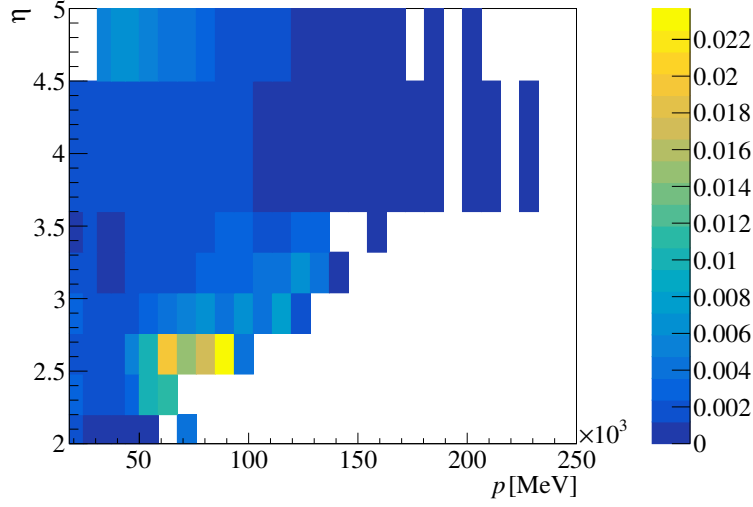


Figure 6.1: The distribution of the weights in the $p - \eta$ plane which are applied to the fake- p data sample in order to determine the p mis-ID template.

Category	Criteria
p -like	$\text{ProbNN}_{\mu}^{\text{combined}} > 0.4$
K -like	$\text{ProbNN}_K^{\text{combined}} > 0.3$ & ! μ -like
π -like	$\text{ProbNN}_{\pi}^{\text{combined}} > 0.3$ & ! μ -like & ! K -like
U	! π -like & ! μ -like & ! K -like

Table 6.3: The scheme used to form the track categories for the fake- μ sample.

sumed that the number of protons misidentified as muons is negligible given that the isMuon requirement is applied in the nominal data sample. Hence, the fake- μ tracks are divided into μ -like, K -like, π -like and U -like categories, according to the criteria in Table. 6.3. The weights are given by

$$p(X \rightarrow \mu) = \frac{N_K \epsilon(K \rightarrow \mu) + N_{\pi} \epsilon(\pi \rightarrow \mu) + N_U \epsilon(U \rightarrow \mu)}{N_K + N_{\pi} + N_U}, \quad (6.5)$$

where the N_X are determined by solving a matrix equation as in Eq. (6.4) using the same unfolding algorithms. The distribution of the weights is shown in Fig. 6.2.

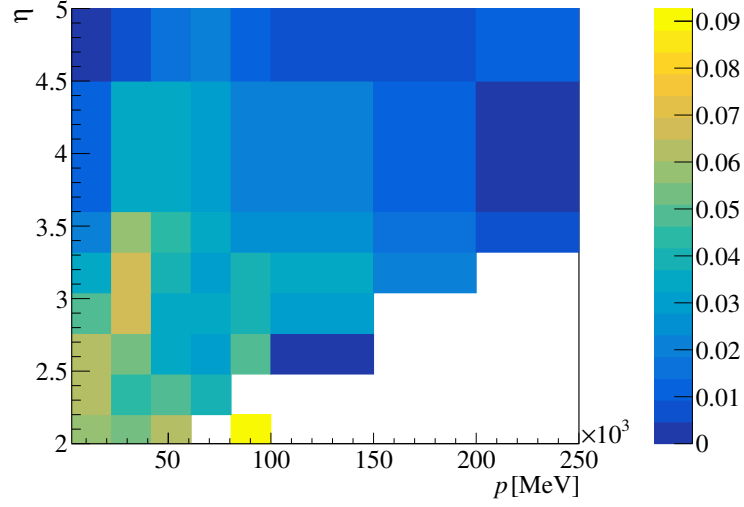


Figure 6.2: The distribution of the weights in the $p - \eta$ plane that are applied to the fake- μ data sample in order to determine the μ mis-ID template. The weights have been inflated by an order of magnitude for illustration purposes.

6.4 Combinatorial background

A data-driven technique is used to derive a proxy for the combinatorial background in the fit. Given the nature of combinatorial background, which results from random combinations of tracks, the *event mixing* method is used. This involves mixing tracks from different events and constructing a candidate which satisfies a looser version of the full selection. Most notably, the criteria placed on quantities that are derived online, such as those relating to the SV, are lost in the derivation of this template by construction. To this end, the template is validated in a control region in order to verify its accuracy despite the modified selection. It is anticipated that the dominant source of combinatorial background in this analysis arises when a $\bar{p}\mu$ system is combined with a random p . The probability of having a combination of three random $p + \bar{p} + \mu$ tracks is expected to be much lower.

In order to implement the event mixing, the B flight vector is taken from candidate n . A $\bar{p}\mu$ combination is taken from candidate $n + 1$ and a proton from candidate $n + 2$. The four-momenta of these tracks are combined in order to form a $p\bar{p}\mu$ candidate. Furthermore, the relative displacements of the $p\bar{p}\mu$ from the B of their original event is retained. This is tantamount to rotating the tracks such that they are approximately collinear with the B from event n . A summary of the selection placed on the kinematics of the $p\bar{p}\mu$ combination is given in Table. 6.4.

Selection
$m(p\bar{p}\mu) < 5279 \text{ MeV}$
$3000 < M_{\text{corr}} < 6000 \text{ MeV}$
$m(p\bar{p}) < 3000 \text{ MeV}$
$\text{ProbNN}_p^{\text{combined}} > 0.5$
$\text{ProbNN}_\mu^{\text{combined}} > 0.4$

Table 6.4: The selection criteria on the combinatorial candidates derived from event mixing.

The validation of the event mixing is shown in Fig. 6.3. This is performed in the so-called upper mass-sideband region $m(p\bar{p}\mu) > 5279 \text{ MeV}$ that will be largely dominated by combinatorial background. Both part-reco and mis-ID backgrounds occupy the lower $m(p\bar{p}\mu)$ region. In the construction of the nominal template, candidates in the range $m(p\bar{p}\mu) < 5279 \text{ MeV}$, or the signal region, are used.

The event-mixing sample is further reweighted to the data in the control region in order to resolve the outstanding discrepancies. The reweighting is performed in the following three variables; the transverse momentum p_T^{flight} of the $p\bar{p}\mu$ system with respect to the line of flight of the parent B , $m(p\bar{p})$ and the magnitude of the $p\bar{p}$ momentum. These are chosen for their high degree of correlation with the fit variables and are found to accurately reproduce the shapes of the control sample.

The degree of cancellation of the individual transverse momenta of the final-state particles is captured by the resulting value of $p_T^{\text{flight}}(p\bar{p}\mu)$. This cancellation has a direct impact on the total four-momentum of the $p\bar{p}\mu$ system, which is closely related to m_{miss}^2 . Furthermore, the kinematics of the $p\bar{p}$ system has a direct influence on the values of E_μ^* and q^2 . This is most clearly seen from a purely numerical perspective by inspecting their definitions in Sec. 4.2. The resulting template in the signal region is shown in Fig. 6.4. The combinatorial background occupies the high q^2 and low E_μ^* regions, whilst the shape in m_{miss}^2 is broad.

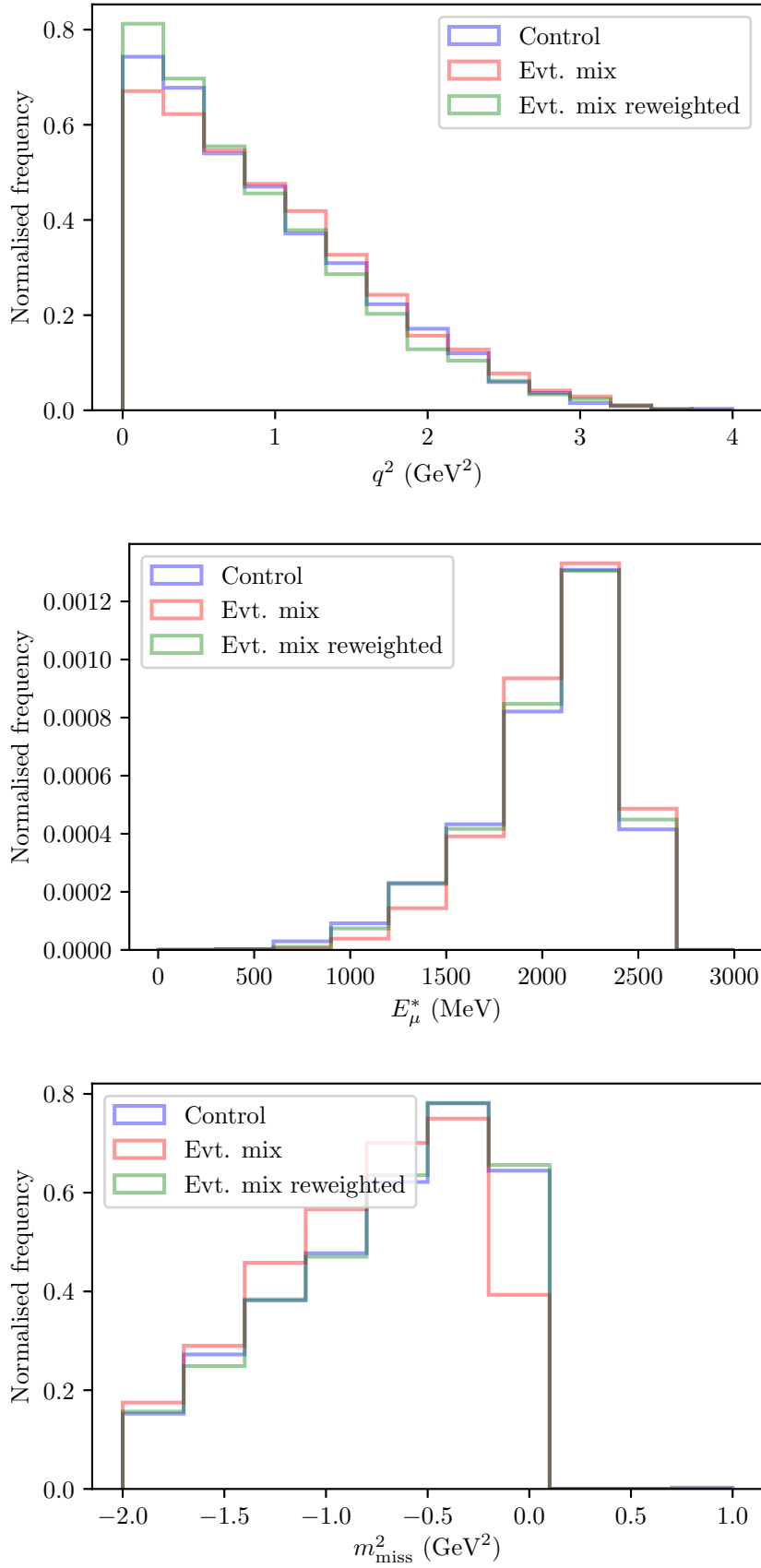


Figure 6.3: Validation of the event mixing template shape in the control region.

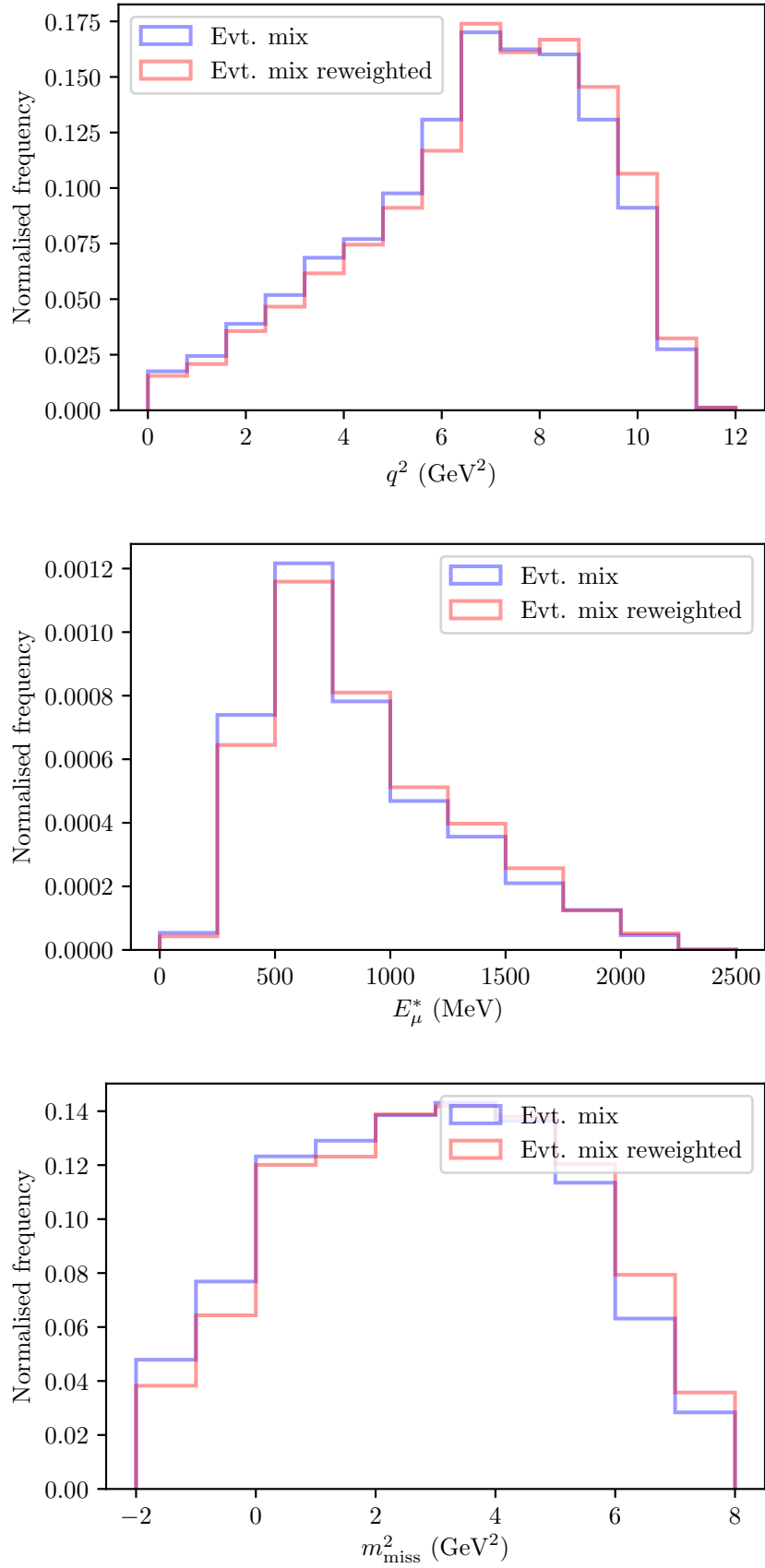


Figure 6.4: Combinatorial background shape in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 , before and after reweighting.

Chapter 7

Selection

The stages of the selection are described in this chapter, concluding with an account of how the criteria are optimised. The stripping and trigger criteria are presented first, followed by a description of the implementation of the charged and neutral isolation tools. The PID requirements are then outlined, before concluding with an explanation for the requirement placed on the opening angle of the $p\bar{p}$ system in the B rest-frame, and the cut optimisation.

7.1 Trigger and stripping criteria

A number of trigger criteria are placed on the candidates, which are outlined in Table 7.1. The candidates are required to satisfy both the L0 and HLT2 requirements and either of the HLT1 requirements. A further set of criteria are applied in the form of a *stripping line*. This is a bespoke suite of cuts placed on the individual daughter tracks, as well as their combination in order to further reduce background contamination. The stripping selection is broadly commensurate with that of the HLT2 trigger line, and an exhaustive list of the criteria is given in Appendix A.

Broadly speaking, the stripping line selects a $p\bar{p}$ pair with (anti)proton momentum greater than 15 GeV that are well-displaced from the PV and that form a good quality vertex. The momentum selection is such that the protons are above the thresholds for Cherenkov radiation in RICH 1 and 2, with the aim of reducing misID background. A loose momentum selection is placed on the muon of $p(\mu) > 3 \text{ GeV}$ to try and maximise the efficiency of the softer muons that are to be expected from the $B^+ \rightarrow p\bar{p}(\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau) \nu_\tau$ decay chain. In addition, the muons are required to have good quality tracks and to be well-displaced from the PV. Relatively loose require-

Trigger	Name	Description
L0	L0Muon	Select track in muon stations with $p_T \gtrsim 1.5$ GeV
HLT1	Hlt1TrackMVA	Select displaced track with $p_T \gtrsim 1.6$ GeV
	Hlt1TrackMuon	Select displaced track with $p_T \gtrsim 1.0$ GeV which satisfies isMuon requirement
HLT2	Hlt2XcMuXForTauB2XcMu	Select $p\bar{p}\mu$ system with high momentum $p\bar{p}$ pair and loose μ selection

Table 7.1: Summary of the trigger criteria used for the data used in the measurement of $R(p\bar{p})$.

ments are made on the $p\bar{p}\mu$ vertex quality, as the τ is expected to fly a short distance before decaying to a muon. The SV is required to be well-displaced from the PV.

7.2 Charged isolation

As discussed in Sec. 4.4, a BDT is used to identify the four tracks in the VELO, besides those of the signal, that are suspected to have originated from the SV through assigning them a corresponding probability. These BDT probabilities are subsequently used as input features to an additional BDT which aims to classify candidates based on how likely they are to be a part-reco background with additional charged tracks.

A simulation sample of $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ is used to represent the signal component of the training sample. A cocktail of $B \rightarrow p\bar{p}DX$, $B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{p}D\mu^+\nu_\mu$ simulation samples is used as a background proxy, where each component is weighted in order to reflect the expected yields. The corrections outlined in Sec. 5 are also applied prior to training. The BDT is trained with k -folding cross-validation [102], with a number $k = 10$ folds. This is a way of dividing up the data randomly into k random samples in order to avoid biasing the algorithm. Candidates where no extra tracks are found by the charged isolation are removed from the training. The training performance of the BDT is shown in Fig. 7.1. The relative discrimination power provided by each of the charged BDT input features is shown in Fig. 7.2. Here, the *score* is a measure of how the feature improves the classification on average across the ensemble of trees in the algorithm. The BDT output for the most likely extra track

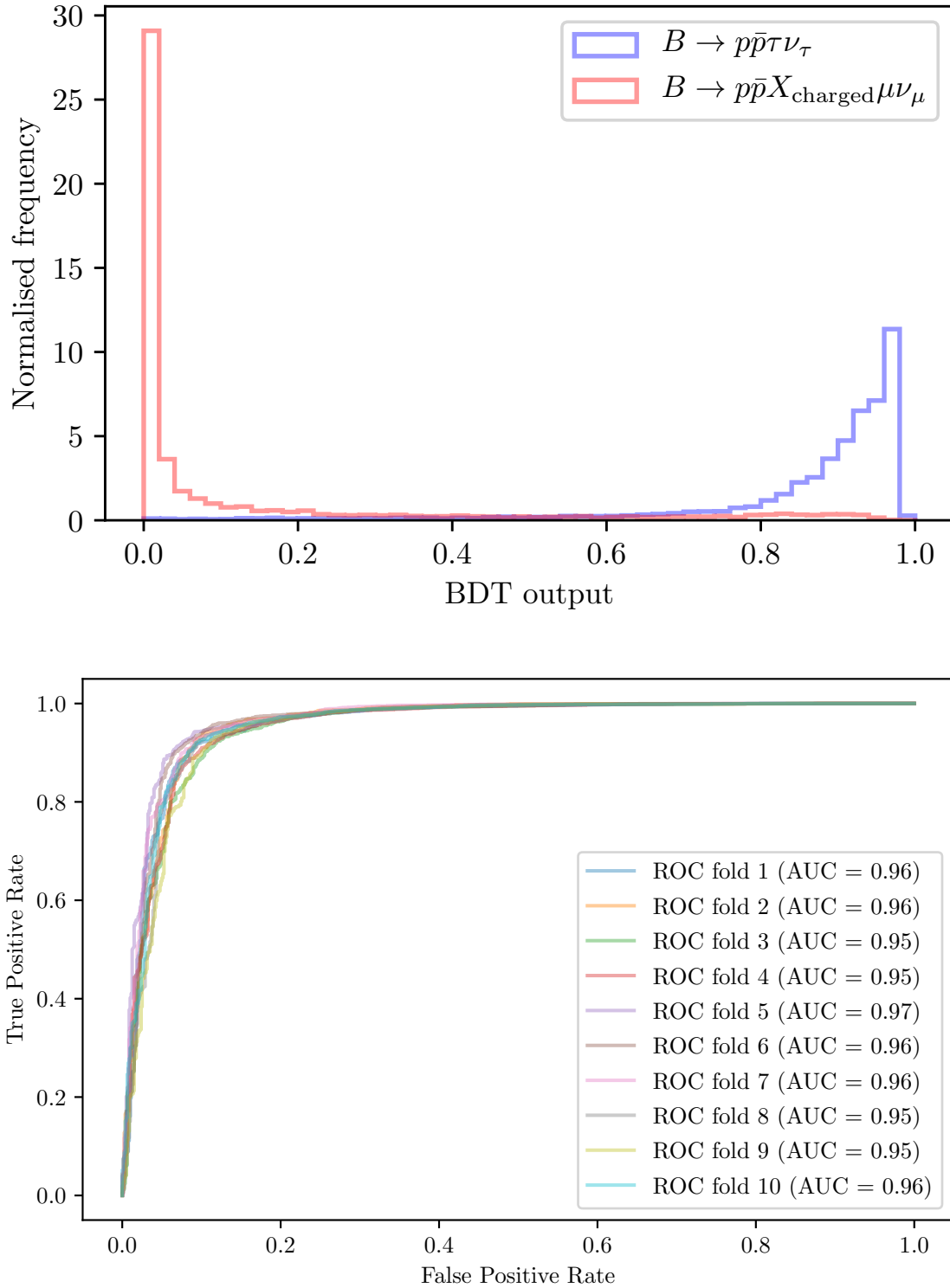


Figure 7.1: (Top) The BDT output distributions for the signal and background sample and (bottom) the true positive classification rate as a function of the false positive rate (the receiver operating characteristic (ROC) curve) for each of the k training folds. The area under the curve (AUC) is a measure of the BDT performance, where an ideal classifier has $\text{AUC} = 1.00$. A small proportion of the background candidates appear signal-like as the track reconstruction efficiency is not 100%, resulting in the small red peak around 0.9.

proves to be the most important feature when it comes to classifying the candidates. The other BDT scores contribute less to the classification, as expected.

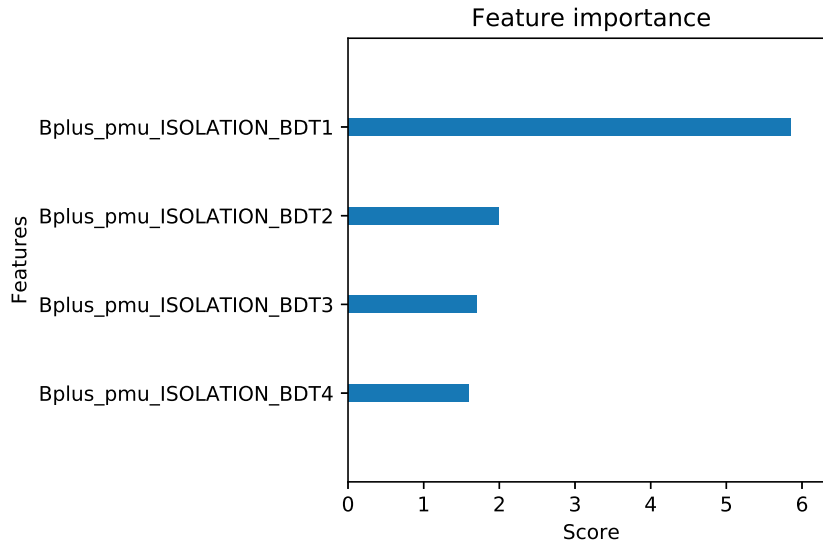


Figure 7.2: Input feature importance for the charged isolation BDT. The features correspond to the additional track probabilities labelled $1 \rightarrow 4$. The highest scoring track BDT output from the vertex isolation tool bears the most discriminating power.

7.3 Neutral isolation

Many part-reco backgrounds involve final states with extra neutral particles, such as $B^+ \rightarrow p\bar{A}_c^-(\rightarrow \bar{p}K^0\pi^0)\mu^+\nu_\mu$. A novel neutral isolation BDT is used for this analysis. The main concept is similar to that of the charged isolation BDT. However, rather than using information from the VELO for the additional charged tracks, the neutral isolation tool takes information from the ECAL for γ or π^0 candidates, such as their momenta and relative displacements from the flight vector, to assign them a probability which reflects how likely it is that they originated from the SV. An exhaustive list of the information used to assign these probabilities is provided in Appendix. B. These probabilities are used as input to another BDT trained on simulation.

A simulation sample of $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ is used as a signal proxy, whilst a cocktail of $B^+ \rightarrow p\bar{A}_c^-(\rightarrow \bar{p}K^0\pi^0)\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{\Delta}(\rightarrow p\pi^0)\mu^+\nu_\mu$ simulation samples is used as a background. The BDT itself is taken from the XGBoost [103] library, and the training is validated using k-folding cross-validation, with a number $k = 10$ folds. The training performance is shown in Fig. 7.3. As expected, the separation power is not as good as that of the charged BDT. However, it is clear that the BDT does provide some signal-background discrimination. The background simulation is concentrated at lower values of the BDT output, whilst the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ is shifted towards higher values. The importance of the four most important features in terms of discrimination power for the neutral BDT is shown in Fig. 7.4. The BDT output for the most likely photon candidate proves to be most discriminating, with an appreciable contribution from the BDT scores of the less likely photon candidates. The number of π^0 candidates surrounding the candidate also contributes to the separation.

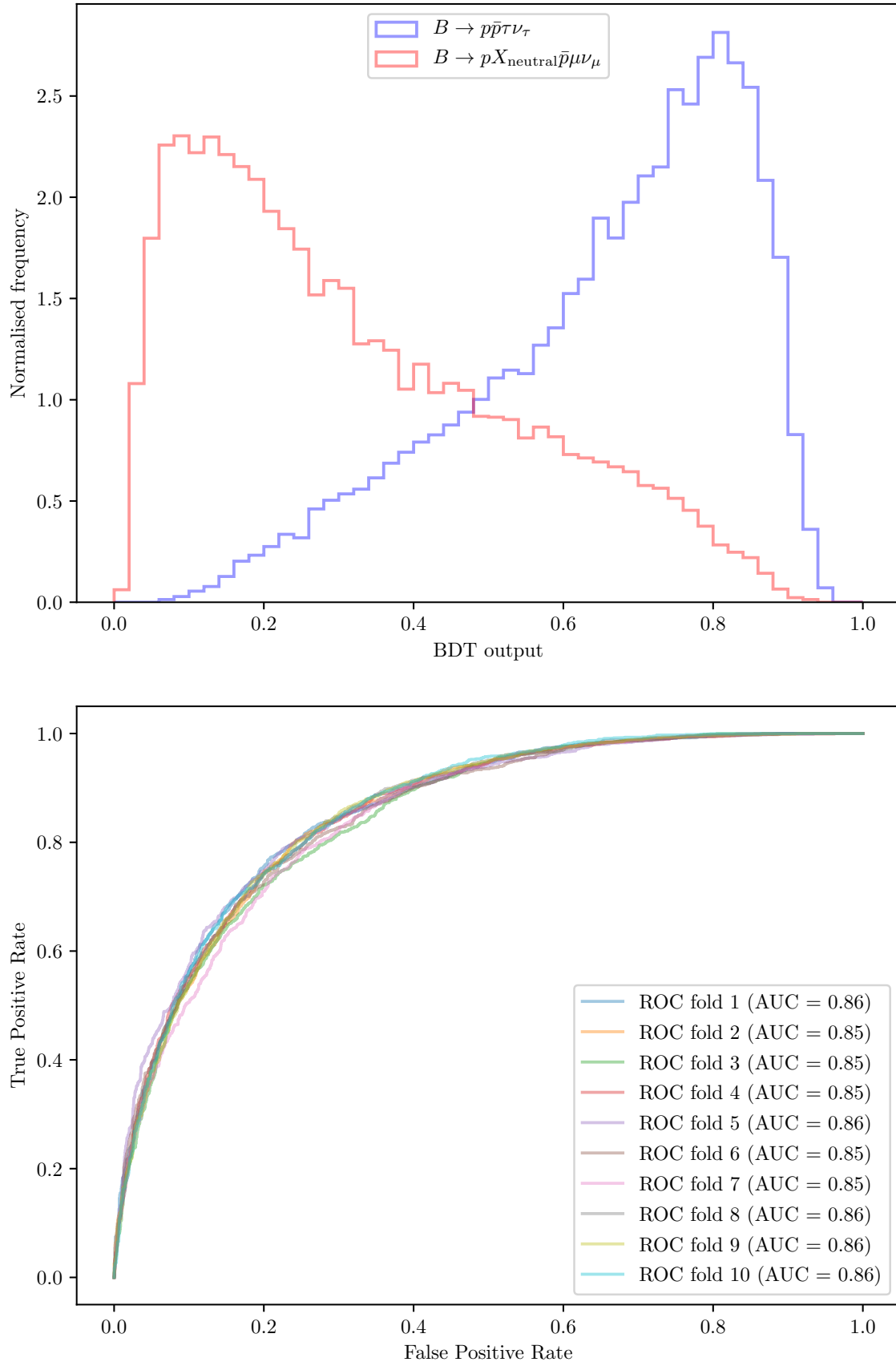


Figure 7.3: (Top) The BDT output distributions for the signal and background sample and (bottom) the ROC curves for each of the folds used in the training of the neutral isolation BDT.

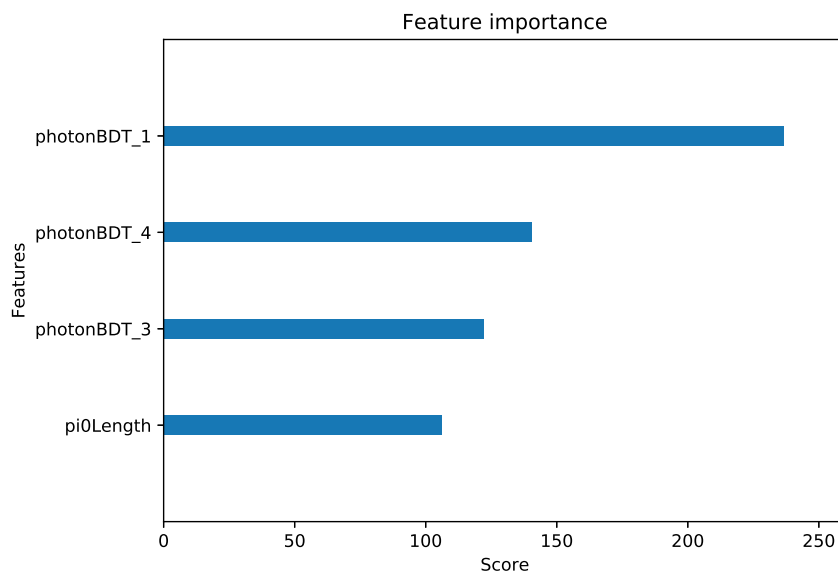


Figure 7.4: Importances for the four most discriminating features for the neutral isolation BDT. The highest scoring photon candidate BDT output is the most important feature.

7.4 Particle identification criteria

As alluded to in Sec. 5.6, a cut is placed on the $\text{ProbNN}_p^{\text{combined}}$ and $\text{ProbNN}_\mu^{\text{combined}}$ variables in order to reduce the contamination from p mis-ID and μ mis-ID candidates, respectively. The definitions are stated below once again for convenience:

$$\begin{aligned} \text{ProbNN}_p^{\text{combined}} = & \text{ProbNN}p(1 - \text{ProbNN}\pi)(1 - \text{ProbNN}K)(1 - \text{ProbNN}e) \\ & \times (1 - \text{ProbNNghost}), \end{aligned} \quad (7.1)$$

$$\begin{aligned} \text{ProbNN}_\mu^{\text{combined}} = & \text{ProbNN}\mu(1 - \text{ProbNN}\pi)(1 - \text{ProbNN}K)(1 - \text{ProbNN}e) \\ & \times (1 - \text{ProbNN}p)(1 - \text{ProbNNghost}). \end{aligned} \quad (7.2)$$

7.5 Opening diproton angle

The opening angle of the $p\bar{p}$ system in the B rest-frame provides further discriminating power. For $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays in the rest-frame of the B , the emission of the W and the diproton is such that momentum is conserved. That is, the momentum vectors of the W and the combined $p\bar{p}$ system cancel. This means that, on average, the p and $\bar{p}$ are highly collinear. By extension, the opening angle $\theta_{p\bar{p}}$ subtended by their respective momentum vectors is likely to be small. However, when either the p or $\bar{p}$ originate from a $\bar{\Lambda}_c^-$ (for example), as in $B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$ decays, then this high degree of collinearity is broken.

It can be seen in Fig. 7.5 that this is borne out in the data. There is a clear concentration of candidates with lower values of $M_{\text{corr}} < 4500$ and with $-1 < \cos \theta_{p\bar{p}} < 0$. As a result, only candidates with $\cos \theta_{p\bar{p}} > 0$ are selected.

7.6 Optimisation

After the stripping selection, cuts are placed on the charged isolation BDT, the neutral isolation BDT and the $\text{ProbNN}_p^{\text{combined}}$ and $\text{ProbNN}_\mu^{\text{combined}}$ quantities. A grid-search optimisation is performed over these selection variables in order to determine the optimal cut-points. The cuts in the stripping selection are fixed and are therefore excluded from the optimisation. The opening diproton angle is also excluded as the cut point is derived from a physical argument regarding the kinematics of the decay

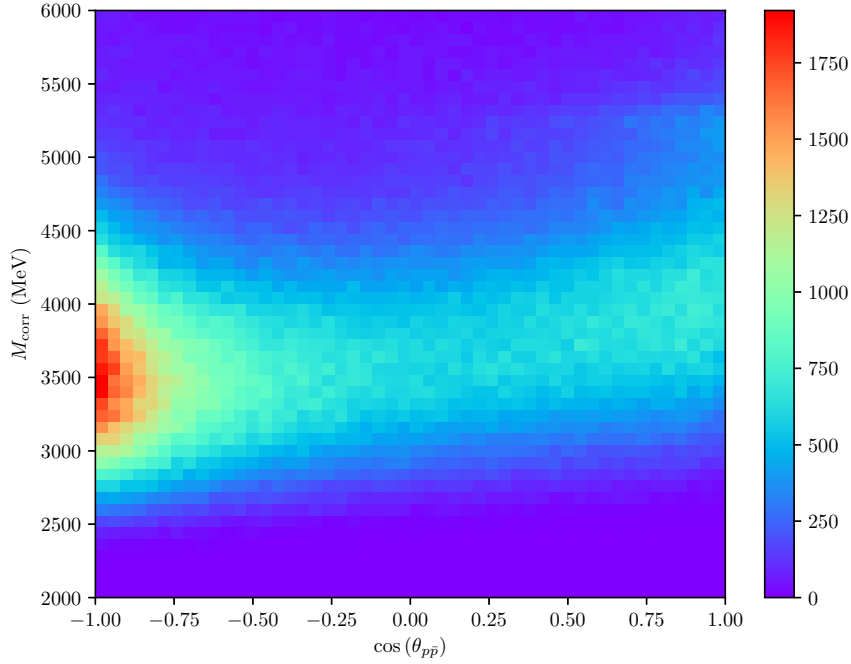


Figure 7.5: The distribution of M_{corr} vs $\cos\theta_{p\bar{p}}$ for the full data set with only the $\text{ProbNN}_p^{\text{combined}} > 0.5$ and $\text{ProbNN}_\mu^{\text{combined}} > 0.4$ requirements applied. There is a clear accumulation of candidates with $M_{\text{corr}} < 4500$ for $\cos\theta_{p\bar{p}} < 0$.

products.

The optimisation involves defining a range over which to optimise the variables, e.g. the charged BDT is optimised in the 0 – 1 range. A number of sampling points from this range are chosen for each variable, resulting in a grid where each point represents a unique configuration of the variables. A number of pseudo- (or toy) experiments are performed at each point on the grid. Expected yields $N(Y)$ for a given decay mode Y are estimated using the integrated luminosity $\mathcal{L}$, the $B^\pm$ production cross-section $\sigma(B^\pm) = 86.6 \pm 8.3 \mu\text{b}$ [93], the branching fraction $\mathcal{B}(Y)$ and selection efficiency $\epsilon(Y)$ according to

$$N(Y) = \mathcal{L} \times \sigma(B^\pm) \times \mathcal{B}(Y) \times \epsilon(Y). \quad (7.3)$$

These estimated yields are used as normalisation factors for each of the templates in the construction of a PDF. A toy data sample is generated by randomly sampling this PDF. A fit is then performed to this data sample using the same PDF. A total of 500 fits are performed for each cut configuration. The cuts are optimised using the mean of the absolute uncertainty $\sigma(\mathcal{R}(p\bar{p}))$ on the raw value of $R(p\bar{p})$ from the fits as a figure of merit (FOM). The optimal cut-point is then taken to be that with the lowest mean

value of $\sigma(\mathcal{R}(p\bar{p}))$.

The yields after stripping and trigger requirements estimated for each of the fit components are shown in Table. 7.2, alongside the assumed BF. The BF is quoted as the product of the BF for the decay of the B and any subsequent decay resulting in missing particles X . Unless stated otherwise, the BFs are taken from Ref. [88]. It is worth noting that the yield estimates used to estimate the best working point have no further bearing on the result. In the case that the chosen working point was in fact sub-optimal, the result would remain unbiased. The final fit is validated with pseudo-experiments to assess any inherent bias as described in Sec. 10.3.

The $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ and $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ BFs are taken from the LHCb measurement [39], and accounting for the $\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau$ decay in the signal mode. For the $B^+ \rightarrow p(\bar{\Delta} \rightarrow \bar{p}X)\mu^+\nu_\mu$ and $B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu$ decays, the $\bar{\Delta} \rightarrow \bar{p}X$ and $\bar{N}^* \rightarrow \bar{p}X$ BFs are assumed to be 100%, and that of the B decay is taken to be the same as the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ BF, as they also involve a $b \rightarrow u$ transition.

The $B^{+/-0} \rightarrow p\bar{p}DX$ BFs are unmeasured but can be estimated by summing the BF for several exclusive modes such as $B^0 \rightarrow D^-p\bar{p}\pi^-$, $B^0 \rightarrow D^{*-}p\bar{p}\pi^-$, $B^+ \rightarrow D^0p\bar{p}\pi^+$ etc. The $D \rightarrow \mu^+X$ BF is taken to be 6.5%, whilst that of $B^+ \rightarrow p\bar{p}D\mu^+\nu_\mu$ is taken to be the same as that for $B^{+/-0} \rightarrow p\bar{p}DX$. The $B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$, $B^+ \rightarrow p\bar{\Lambda}_c^-\tau^+\nu_\tau$ and $B^+ \rightarrow p(\bar{\Lambda}_c^{*-} \rightarrow \bar{p}X)\mu^+\nu_\mu$ yields are estimated using a BF from an unpublished search for the $B^+ \rightarrow p\bar{\Lambda}_c^-\mu^+\nu_\mu$ decay (and by assuming LFU for the tauonic mode), which sets a limit of $\mathcal{B}(B^+ \rightarrow p\bar{\Lambda}_c^-\mu^+\nu_\mu) < 2.5 \times 10^{-4}$ at a 90% confidence level [104]. The $\bar{\Lambda}_c^- \rightarrow \bar{p}X$ and $\bar{\Lambda}_c^{*-} \rightarrow \bar{p}X$ BFs are taken to be 12%.

The $\bar{B}_s^0 \rightarrow D_s^+p\bar{p}\pi^-$ and $\bar{B}_s^0 \rightarrow D^0K^0p\bar{p}$ BFs are also unknown. These are estimated by using the equivalent B^0 mode, where the spectator s -quark is replaced by a d -quark. The BF $\mathcal{B}(\bar{B}^0 \rightarrow D^+p\bar{p}\pi^-) \approx 3.3 \times 10^{-4}$ is used as an estimate for $\mathcal{B}(\bar{B}_s^0 \rightarrow D_s^+p\bar{p}\pi^-)$. The $D_s^+ \rightarrow \tau^+\nu_\tau$ and $D_s^+ \rightarrow \mu^+\nu_\mu$ BFs are taken to be 5.48% and 5.49×10^{-3} , respectively. After accounting for the $\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau$ decay, the tauonic and muonic contributions from the D_s^+ are approximately equal. However, $\mathcal{B}(\bar{B}^0 \rightarrow D^0\pi^0p\bar{p})$ is unmeasured and cannot be used to estimate $\mathcal{B}(\bar{B}_s^0 \rightarrow D^0K^0p\bar{p})$. In this case, $\mathcal{B}(\bar{B}^0 \rightarrow D^0\pi^0p\bar{p}) \approx \mathcal{B}(\bar{B}^0 \rightarrow D^0\pi^+\pi^-p\bar{p})$ is taken as an approximation with $\mathcal{B}(\bar{B}^0 \rightarrow D^0\pi^+\pi^-p\bar{p}) \approx 3.0 \times 10^{-4}$ and $D^0 \rightarrow \mu X \approx 7\%$. The estimates for the p mis-ID and μ mis-ID yields are calculated from the unfolding procedure described in Sec. 7.4. The results of the optimisation are shown in Figs. 7.6 and 7.7, and the optimal cut points are given in Table. 7.3.

Mode	Branching fraction	Approximate Yield
$B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$	9.19×10^{-7}	5,000
$B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$	5.28×10^{-6}	43,000
$B^+ \rightarrow p\bar{\Delta}\mu^+\nu_\mu$	5.28×10^{-6}	16,000
$B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$	5.28×10^{-6}	24,000
$B^+ \rightarrow p\bar{p}DX$	1.59×10^{-4}	180,000
$B^0 \rightarrow p\bar{p}DX$	2.74×10^{-4}	200,000
$B^+ \rightarrow p\bar{p}D\mu^+\nu_\mu$	1.59×10^{-4}	220,000
$B^+ \rightarrow p\bar{\Lambda}_c^-\mu^+\nu_\mu$	6.00×10^{-5}	22,000
$B^+ \rightarrow p\bar{\Lambda}_c^{*-}\mu^+\nu_\mu$	6.00×10^{-5}	13,000
$\bar{B}_s^0 \rightarrow D_s^+p\bar{p}\pi^-$	4.96×10^{-6}	14,000
$\bar{B}_s^0 \rightarrow D^0K^0p\bar{p}$	2.59×10^{-5}	18,000
p mis-ID	N/A	4400
μ mis-ID	N/A	1400

Table 7.2: A summary of the yields and assumed BF's for the cut optimisation. The yields serve as inputs to the toy fits.

Selection cut
Charged BDT > 0.9
Neutral BDT > 0.4
ProbNN _p ^{combined} > 0.5
ProbNN _{μ} ^{combined} > 0.4

Table 7.3: A summary of the optimal cut points for each selection variable considered in the optimisation.

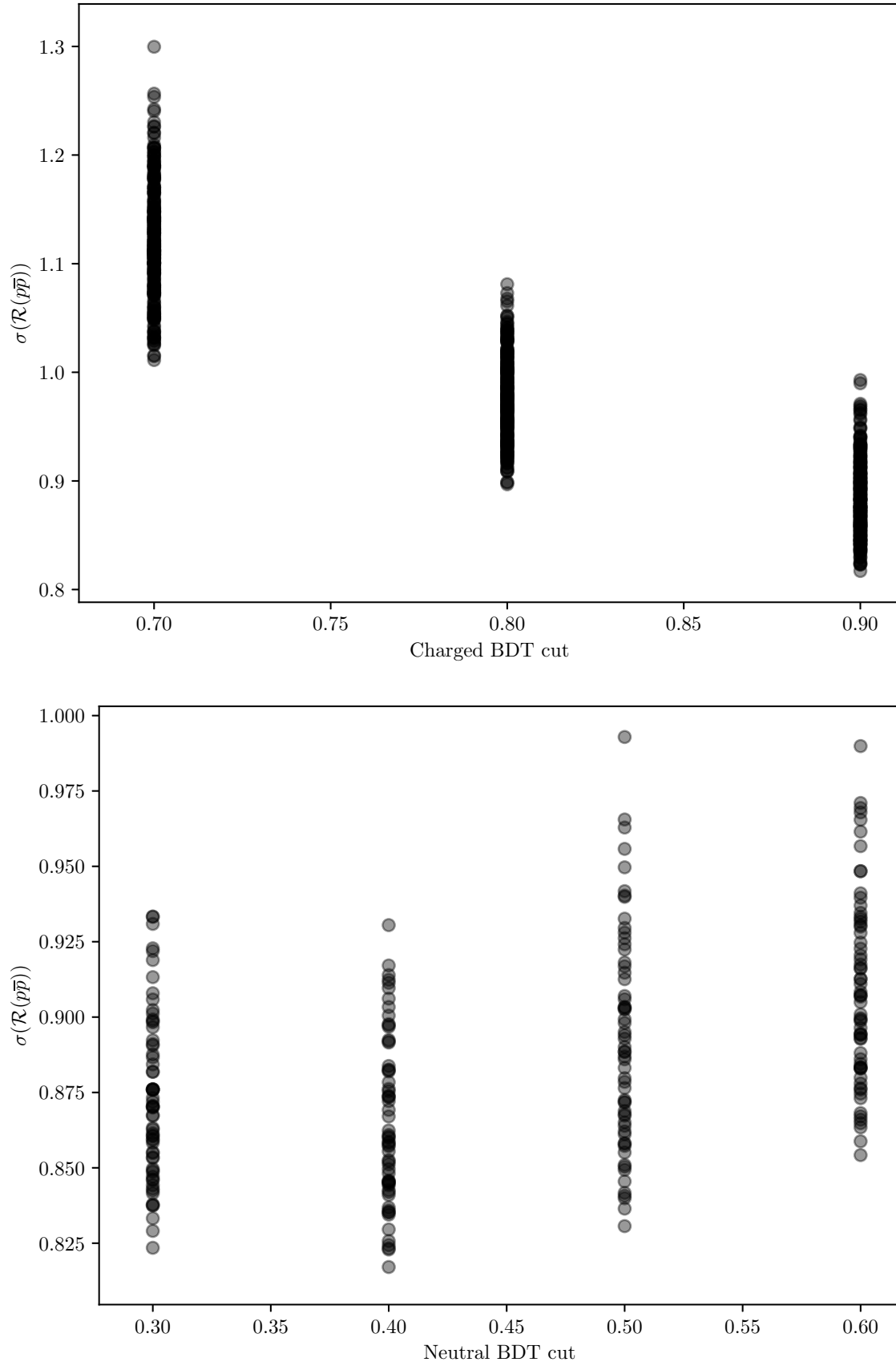


Figure 7.6: Average uncertainty $\sigma(\mathcal{R}(p\bar{p}))$ as a function of the (top) charged and (bottom) neutral isolation BDT cut values. Each point represents the average uncertainty from 500 toy fits at a given configuration of the cut values.

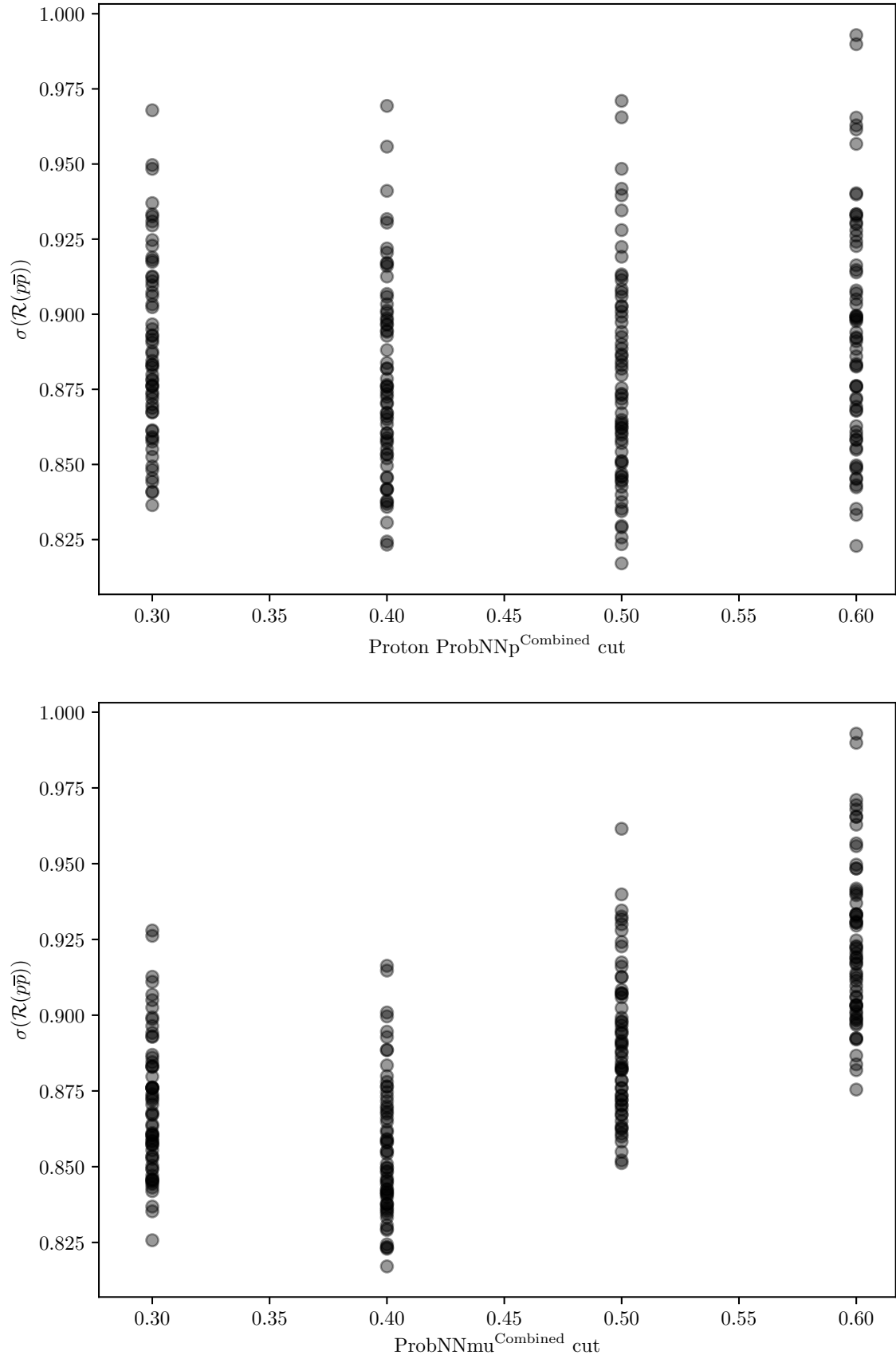


Figure 7.7: Average uncertainty $\sigma(\mathcal{R}(p\bar{p}))$ from 500 fits to toy data samples as a function of the (top) ProbNN_p^{combined} and (bottom) ProbNN_μ^{combined} cut values.

Chapter 8

Form factor fit

This chapter describes the fit used to extract estimates of the FF parameters as defined in Sec. 2.7 for the $B \rightarrow p\bar{p}$ transition in $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays. This is crucial in obtaining accurate shapes for the signal, normalisation and certain part-reco backgrounds shapes in the fit to determine $R(p\bar{p})$. Firstly, a reminder of the main aspects of the pQCD model are presented, followed by a brief account of the implementation of the model and the methodology used to determine the FF parameters. A reweighting strategy used to obtain distributions representing certain configurations of the FF parameters is then described, followed by a description of the data and simulation samples used for the fit. The chapter concludes with the results of the FF fit.

8.1 Physics model

The main aspects of the pQCD model described in more detail in Sec. 2.7 are presented here once more for convenience. The vector and axial vector components of the hadronic current in the transition of a B to a baryon pair $Y\bar{Y}'$ are parameterised in terms of the FFs

$$f_i = \frac{D_{f_i}}{t^n}, \quad g_i = \frac{D_{g_i}}{t^n}, \quad (8.1)$$

where $i = 1, \dots, 5$ and the constants D_{f_i} and D_{g_i} can be expressed as linear superpositions of the parameters

$$(D_{\parallel}, D_{\perp}) = (67.7 \pm 16.3, -280.0 \pm 35.9) \text{ GeV}^5, \quad (8.2a)$$

$$(D_{\parallel}^2, D_{\parallel}^3) = (-187.3 \pm 26.6, -840.1 \pm 132.1) \text{ GeV}^4, \quad (8.2b)$$

$$(D_{\parallel}^4, D_{\parallel}^5) = (-10.1 \pm 10.8, -157.0 \pm 27.1) \text{ GeV}^4, \quad (8.2c)$$

as

$$D_{g_1} = \frac{5}{3}D_{||} - \frac{1}{3}D_{\perp}, \quad (8.3)$$

$$D_{f_1} = \frac{5}{3}D_{||} + \frac{1}{3}D_{\perp}, \quad (8.4)$$

$$D_{g_j} = \frac{5}{3}D_{||}^j = -D_{f_j}. \quad (8.5)$$

The vector and axial-vector components of the hadronic current are (respectively):

$$\begin{aligned} \langle Y\bar{Y}|\bar{q}\gamma_\mu b|B\rangle = \\ i\bar{u}(p_Y)[g_1\gamma_\mu + g_2i\sigma_{\mu\nu}q^\nu + g_3q_\mu + g_4(p_{\bar{Y}} + p_Y)_\mu + g_5(p_{\bar{Y}} - p_Y)_\mu]\gamma_5 v(p_{\bar{Y}}), \end{aligned} \quad (8.6)$$

$$\begin{aligned} \langle Y\bar{Y}|\bar{q}\gamma_\mu\gamma_5 b|B\rangle = \\ i\bar{u}(p_Y)[f_1\gamma_\mu + f_2i\sigma_{\mu\nu}q^\nu + f_3q_\mu + f_4(p_{\bar{Y}} + p_Y)_\mu + f_5(p_{\bar{Y}} - p_Y)_\mu]v(p_{\bar{Y}}). \end{aligned} \quad (8.7)$$

Using the values in Eq. 8.2 would be approximate for the $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ case as they are determined from fits to $B \rightarrow p\bar{p}X$ data, where X is a single hadron [62–67]. In order to mitigate this approximation, the parameters are extracted by performing a fit to the Run 1 data sample.

The problem can be simplified further by reducing the set of parameters to fit for. It can be seen from the form of the hadronic currents given in Eqs. (8.6) and (8.7) that $D_{||}$ and $D_{\perp}$ have a sub-leading effect on the distribution shapes with respect to the other parameters. This is due to the fact that there is a lack of dependence on any momenta in the numerator of the terms which are proportional to g_1 and f_1 .

It is also argued by the authors of Ref. [40] that the terms proportional to $D_{||}^3$ do not contribute to the transition. Given that the muon originates from a highly boosted B meson, it is taken to have an exclusively left-handed polarization $\epsilon_-^\mu(q)$. This, when taking the sum over all possible polarizations of the final-state products when calculating the branching fraction, leads to a vanishing contribution. The sum takes the form of a dot-product $\epsilon \cdot q$. Since the left-handed polarization vector $\epsilon_-^\mu(q)$ is transverse to the dilepton momentum q , the result is zero. Hence, the set of parameters can be reduced to three; $D_{||}^2$, $D_{||}^4$ and $D_{||}^5$.

8.2 Implementation

In order to carry out the form factor fit, the model outlined above was implemented in EVTGEN with the capability to generate $B^+ \rightarrow Y\bar{Y}'\ell^+\nu_\ell$ decays for any pair of baryons, with spins $J = 1/2, 3/2$. A comparison between some of the kinematic distributions produced by the pQCD model and those of the PHSP model are shown in Fig. 8.1. There is a clear concentration of events near the $m(p\bar{p})$ kinematic threshold, which is in accordance with what has already been observed in the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decay in Fig. 1.3. The parameters of the model are provided as external arguments, the values of which are free to be varied by the user.

Clearly, the values of the parameters influence the shapes of the distributions which are produced. Thus, by varying the model parameters, it is possible to produce any number of distribution shapes. As will be discussed below, it is this freedom which facilitates the determination of the form factor parameters for the $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays.

8.3 Fit methodology

The FF parameters are determined through an extended maximum-likelihood fit to the Run 1 data samples in m_{miss}^2 , E_μ^* and q^2 . The selection differs from that of the Run 2 sample used for the $R(p\bar{p})$ fit. Broadly speaking, the selection is tighter in order to better isolate the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ mode. The aim is to achieve a sample that is rich in $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays so that it can be used to reliably estimate the FF parameters. The q^2 spectrum is particularly sensitive to the FF parameters, and the m_{miss}^2 and E_μ^* provide discriminating power over background contributions given the characteristic peak of $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ near $m_{\text{miss}}^2 \approx 0$ and the spectrum of highly energetic muons originating from the B decay.

Given the inaccuracies of the pQCD model, the work presented here is approximate. An uncertainty is assigned to account for variations in the shape of the relevant templates due to the FF parameters in the fit for $R(p\bar{p})$. A more detailed discussion is provided in Sec. 11. The HistFactory template fitting package is used to perform the fit itself [105]. This is a software package which allows fits to be performed taking a series of histograms as an input from which to construct a PDF. None of the Run 2 data samples used to fit $R(p\bar{p})$ are included in the form factor fit in order to preserve the statistical independence of the results.

The fit was performed iteratively as a result of the mechanism by which HistFactory

determines the best-fit values for the FF parameters. Within HistFactory, it is possible to include some freedom in the shape of a given template through the AddHistoSys function. This amounts to providing an additional pair of templates which represent shape variations of some ‘nominal’ template. HistFactory fits for an additional nuisance parameter by performing an interpolation between these templates.

In this case, the best-fit value of this nuisance parameter can be understood as a deviation of a given FF parameter from its nominal value. This is because the varied templates are chosen to be those which result from fixing the value of a given model parameter to a number $\sigma_{\text{low}} = +A\sigma$ and $\sigma_{\text{high}} = +B\sigma$ from the nominal value, where σ is the uncertainty quoted on the parameter as in Eq. (8.2), and A and B are numbers where $B > A$. Examples of shape variations for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ which result from setting $\sigma_{\text{low}} = D_{||}^2 - 5\sigma$ and $\sigma_{\text{high}} = D_{||}^2 + 5\sigma$ illustrated in Fig. 8.2.

In this work, the template shapes which are allowed to vary are those which represent components that proceed via a $b \rightarrow u$ transition at quark-level, and are thus possible to describe by the pQCD model. Specifically, this pertains to the following fit components: $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$, $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$, $B^+ \rightarrow p\bar{\Delta}\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$. Moreover, the nuisance parameters representing the FF deviations are shared between each of these components. This constitutes the approximation that the hadronic transition which takes place in these decays is similar enough to be mediated by the same set of FFs.

These deviations are then used as guidance for the choice of shape variations to include in subsequent iterations of the fit. The ‘starting’ configuration of the templates is such that the nominal templates are taken to be those which correspond to setting the FF parameters to their nominal values as per Eq. (8.2), with the shapes allowed to vary between those with $\sigma_{\text{low}} = -5\sigma$ and $\sigma_{\text{high}} = +5\sigma$. The inclusion of this large degree of freedom in the initial template shapes is an attempt to mitigate the approximation of taking the nominal FF parameters from the fits to the fully hadronic data.

Subsequent iterations of the fit proceed in a similar manner. The deviations in the FF parameters resulting from the previous iteration are used as guidance to produce the ‘nominal’ template in the iteration that follows. To illustrate this, an example is considered where the first iteration results in a deviation of $+4\sigma$ in $D_{||}^2$, with deviations in $D_{||}^4$ and $D_{||}^5$ being consistent with zero. The next iteration would use the template reweighted with $D_{||}^2$ set to $+4\sigma$ as the ‘nominal’ template, and will be free to vary between the shapes corresponding to $\sigma_{\text{low}} = +3\sigma$ and $\sigma_{\text{high}} = +5\sigma$. That is, the shapes

are allowed to vary between $\pm 1\sigma$ of the value resulting from the previous iteration. The results of this iteration are then used as input to the one that follows, with a view to iteratively converge onto values for the FF parameters. Essentially, the reason for the iterative nature of the fit is that, a priori, it isn't possible to determine how much freedom to include in the shape variations given that the form factor parameters are not known.

8.4 Reweighting strategy

As already discussed, each iteration of the FF fit relies on templates which have been reweighted to variations of the nominal pQCD model. This is achieved by a gen-level reweighting of the PHSP simulation templates to instances of the pQCD model using `EVTGEN`. The PHSP distributions were reweighted using `GBReweighter`. The ability to reweight in multiple dimensions is of particular importance here given that the feasibility of the FF fit hinges on being able to fully reproduce the pQCD templates in a number of dimensions.

The following set of gen-level variables are used for the reweighting: $m(p\bar{p})$, q^2 , $\cos\theta_Y$ and $\cos\theta_\ell$, and an example of the reweighting is shown for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays in Figure 8.3, and for the other modes in Appendix. C. This is in line with the set of variables which are conventionally used to parameterise four-body phase space. The latter two variables denote the angle subtended by the momentum of the baryon in the dibaryon rest-frame with respect to the line of flight of the dibaryon system in the B rest-frame, and the angle subtended by the momentum of the charged lepton in the dilepton rest-frame with respect to the line of flight of the dilepton system in the B rest-frame, respectively.

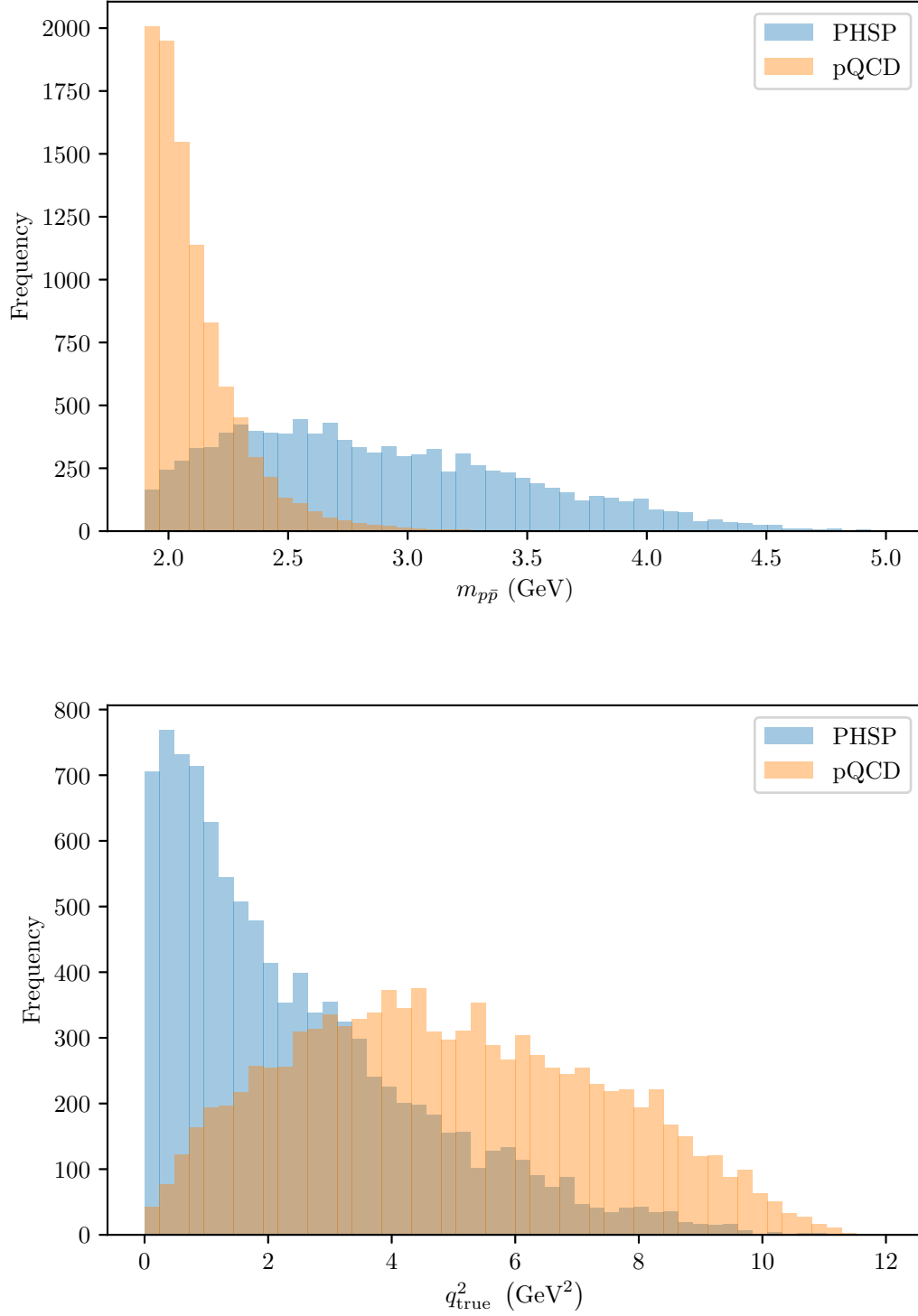


Figure 8.1: Comparison between the distributions obtained using phase-space (PHSP) simulation and those obtained from the physics model (pQCD) in (top) the diproton mass $m(p\bar{p})$ and (bottom) q_{true}^2 for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays.

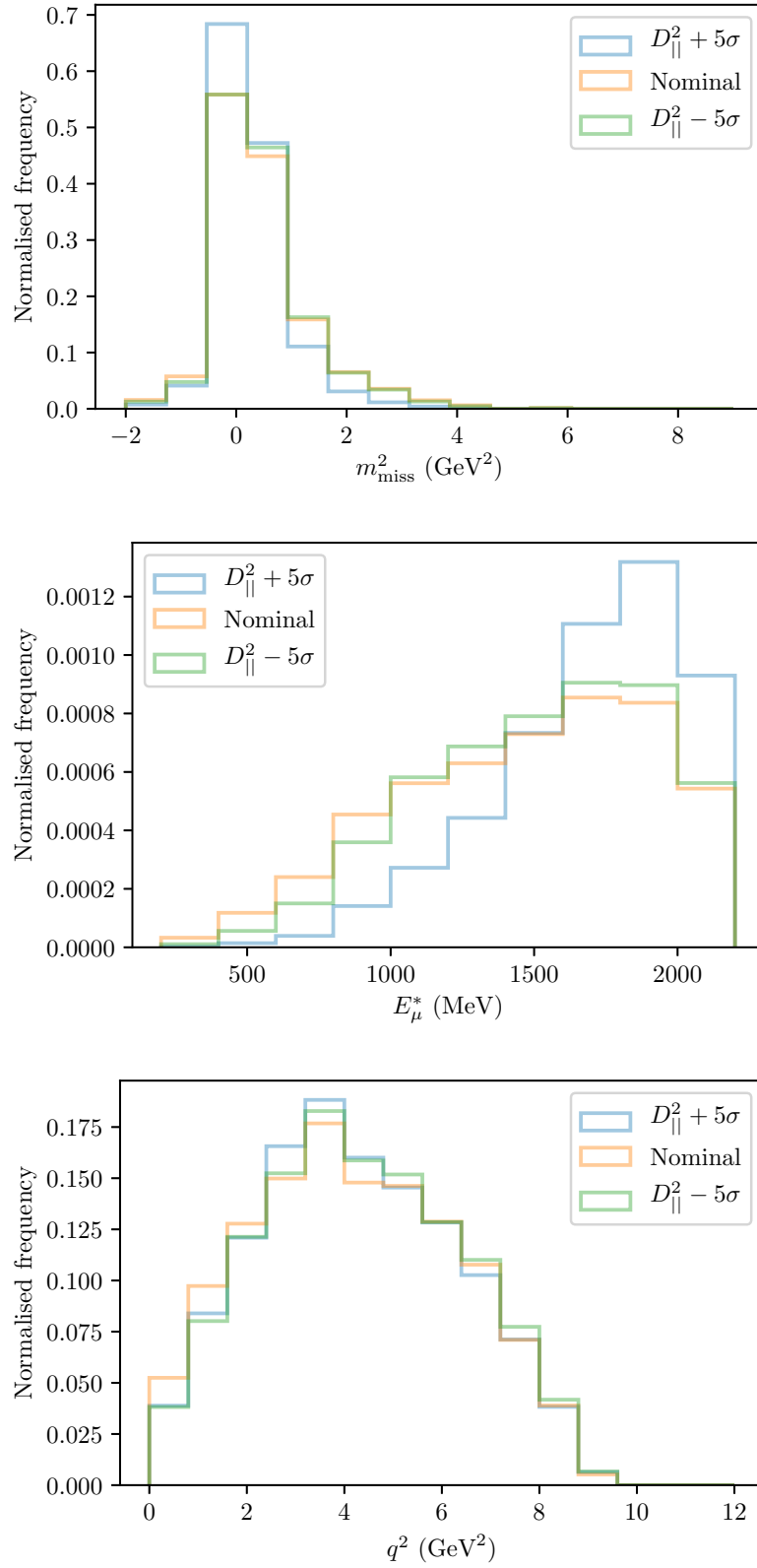


Figure 8.2: Template shape variations for the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ component when $D_{\parallel}^2$ is set to $\pm 5\sigma$ from the nominal value quoted in Eq. (8.2). HistFactory interpolates between these two shapes in order to determine the pull in $D_{\parallel}^2$.

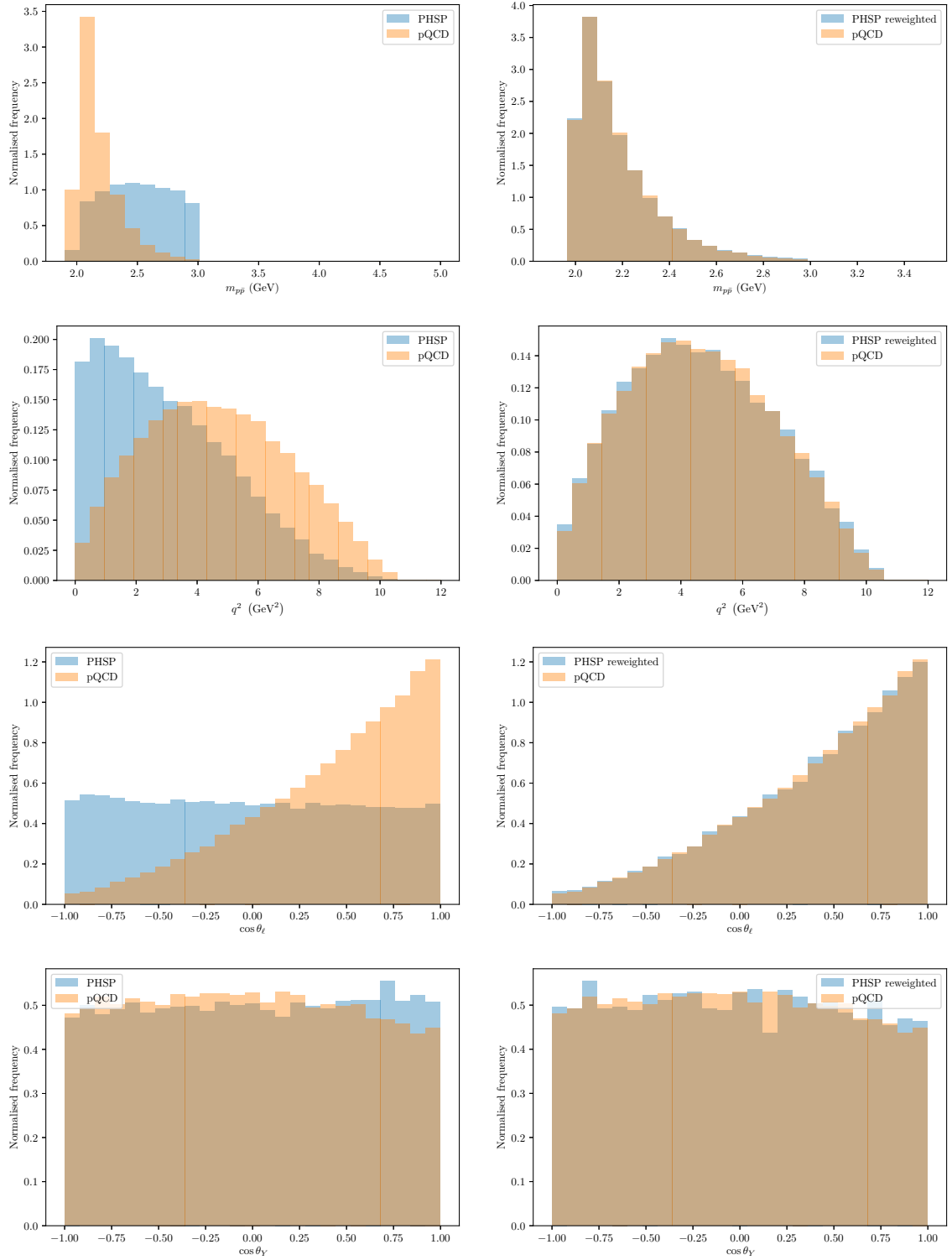


Figure 8.3: Comparisons between PHSP before (left) and after (right) reweighting to the pQCD model in (descending order) $m(p\bar{p})$, q^2 , $\cos \theta_\ell$ and $\cos \theta_Y$ for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays. It is clear that the reweighting algorithm successfully reproduces the pQCD distributions in multiple dimensions.

8.5 Data, simulation and selection

The Run 1 data sample used for the FF fit corresponds to approximately 3.0 fb^{-1} and 1.3×10^{11} $B^\pm$ mesons. The simulation samples used to generate proxies for the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ part-reco background components in the fit were generated using 2012 conditions. A summary of the background processes included in the fit is given in Table. 8.1. The number of part-reco backgrounds included for the FF fit is less than that of the fit for $R(p\bar{p})$ due to the tighter corrected mass window of $4600 < M_{\text{corr}} < 5500 \text{ MeV}$. A significant proportion of part-reco background candidates occupy the lower M_{corr} range, with the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ signal peaking around $M_B \approx 5279 \text{ MeV}$. The simulation corrections described in Sec. 5 are also applied, but using samples

Decay	Decay model	BF order
$B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$	PHSP	$\approx 10^{-4}$
$B^{+/\text{0}} \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X$	PHSP	$\approx 10^{-3}$
$B^0 \rightarrow (\bar{\Lambda}_c^- \rightarrow \mu^-pX)\bar{p}X$	PHSP	$\approx 10^{-4}$
$B^+ \rightarrow p(\Delta \rightarrow \bar{p}X)X\mu^+\nu_\mu$	pQCD	$\approx 10^{-6}$
$B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu$	pQCD	$\approx 10^{-6}$
$B^+ \rightarrow p(\bar{\Lambda}_c^{*-} \rightarrow \bar{p}X)\mu^+\nu_\mu$	PHSP	$\approx 10^{-4}$

Table 8.1: A summary of the part-reco backgrounds included in the fit for the FF parameters. The decays are listed in an approximate order of their expected dominance.

from Run 1 to generate the corrections where necessary. The p mis-ID and μ mis-ID templates are derived using the unfolding procedure described in Secs. 6.2 and 6.3, using fake- p , fake- μ samples and efficiencies from 2012. The combinatorial background template is taken from a so-called inverted $\chi_{\text{vtx.}}^2$ sample. The $\chi_{\text{vtx.}}^2$ value denotes the χ^2 value of the fit used to obtain the vertex and reflects the quality of the vertex. The inverted $\chi_{\text{vtx.}}^2 > 9$ cut is applied to the data to derive the template for the combinatorial background, and the candidates in the nominal data sample satisfy $\chi_{\text{vtx.}}^2 < 9$. to preserve statistical independence. The former sample is expected to be dominated by combinatorial background.

The selection follows that of the LHCb measurement of the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ BF described in Ref. [39]. There are many aspects that are similar to the selection for the $R(p\bar{p})$ measurement, but with some adaptations to maximise the efficiency on $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ rather than $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$. A different stripping line is used and an

account of the details is given in Appendix D. Generally speaking, the kinematic selection on the μ is tighter compared to the line used for $R(p\bar{p})$, and the selection on the $p\bar{p}\mu$ combination is tighter in terms of the requirements placed on the SV.

In addition, analogous L0 and HLT1 trigger requirements are placed on the candidates. However, the general purpose *topological* HLT2 trigger line is used [106]. It is one of the most widely-used trigger lines in LHCb analyses, and is a general-purpose trigger which uses a BDT to identify combinations of tracks that are consistent with originating from a SV that is well-displaced from the PV. This is the topology expected for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays.

A BDT is used for charged isolation during the offline selection, as well as cuts placed on $\text{ProbNN}_p^{\text{combined}}$ and $\text{ProbNN}_\mu^{\text{combined}}$. However, the cut values are adjusted so that they are in accordance with the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ BF measurement. No neutral isolation BDT is used here since it is not necessary in order to achieve a sample sufficiently pure in $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays. A summary of the offline cuts is given in Table. 8.2.

Offline selection cut
Charged BDT > 0.6
$\text{ProbNN}_p^{\text{combined}} > 0.45$
$\text{ProbNN}_\mu^{\text{combined}} > 0.15$
$4600 < M_{\text{corr}} < 5500 \text{ MeV}$

Table 8.2: A summary of the selection for each variable in the FF fit.

8.6 Results

The results of the iterative fitting procedure described in Sec. 8.3 are presented below. A total of four iterations were required before converging on the FF parameter values. Although the fit initially included shape variations in the three parameters $D_{||}^2$, $D_{||}^4$ and $D_{||}^5$, the single parameter $D_{||}^2$ induced enough variation in the shapes of the templates in order to fit the data. This was reflected in the results of the first fit iteration. Whilst there was a significant pull in $D_{||}^2$, the pulls in the $D_{||}^4$ and $D_{||}^5$ were consistent with zero. It is for this reason that variations in $D_{||}^4$ and $D_{||}^5$ were neglected in subsequent iterations of the fit. This corroborates arguments made by the authors of Ref. [40],

where arguments are made on both physical and numerical grounds as to why $D_{||}^2$ dominates over the other parameters.

The projections in m_{miss}^2 , E_μ^* and q^2 of each of the first and second iterations are shown in Fig. 8.5 whilst those of the third and fourth are shown in Fig. 8.6. The legend for the fits is shown in Fig. 8.4. The numerical results of each of the fits are summarised in Table. 8.3. As can be seen, the final value D_{fit}^2 is for $\sigma_{D_{||}^2} = +5.6\sigma$ from the starting value. This is a significant tension with respect to the value D_{start}^2 quoted in Eq. (8.2). However, this disagreement is not surprising as the quoted values were derived from fully hadronic decay modes.

The converged value D_{fit}^2 was used in the simulation production for all of the V_{ub} modes. Due to their negligible influence on the distribution shapes, the remaining parameter values are left unchanged from those quoted in Eq. (8.2).

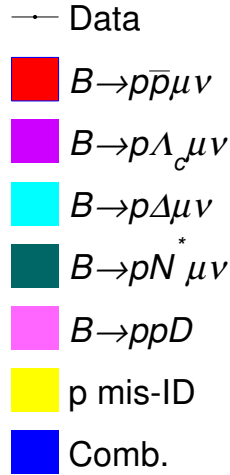


Figure 8.4: Legend for the form factor fit.

Iteration	$(\sigma_{\text{low}}, \sigma_{\text{high}})$	$\sigma_{D_{ }^2}$
1	$(-5, +5)$	4.0
2	$(+3, +5)$	4.5
3	$(+3.5, +5.5)$	5.5
4	$(+4.5, +6.5)$	5.6

Table 8.3: A summary of the results of each of the FF fit iterations. The range of variation is also given, in units of the uncertainty on $D_{||}^2$ quoted in Eq. (8.2)

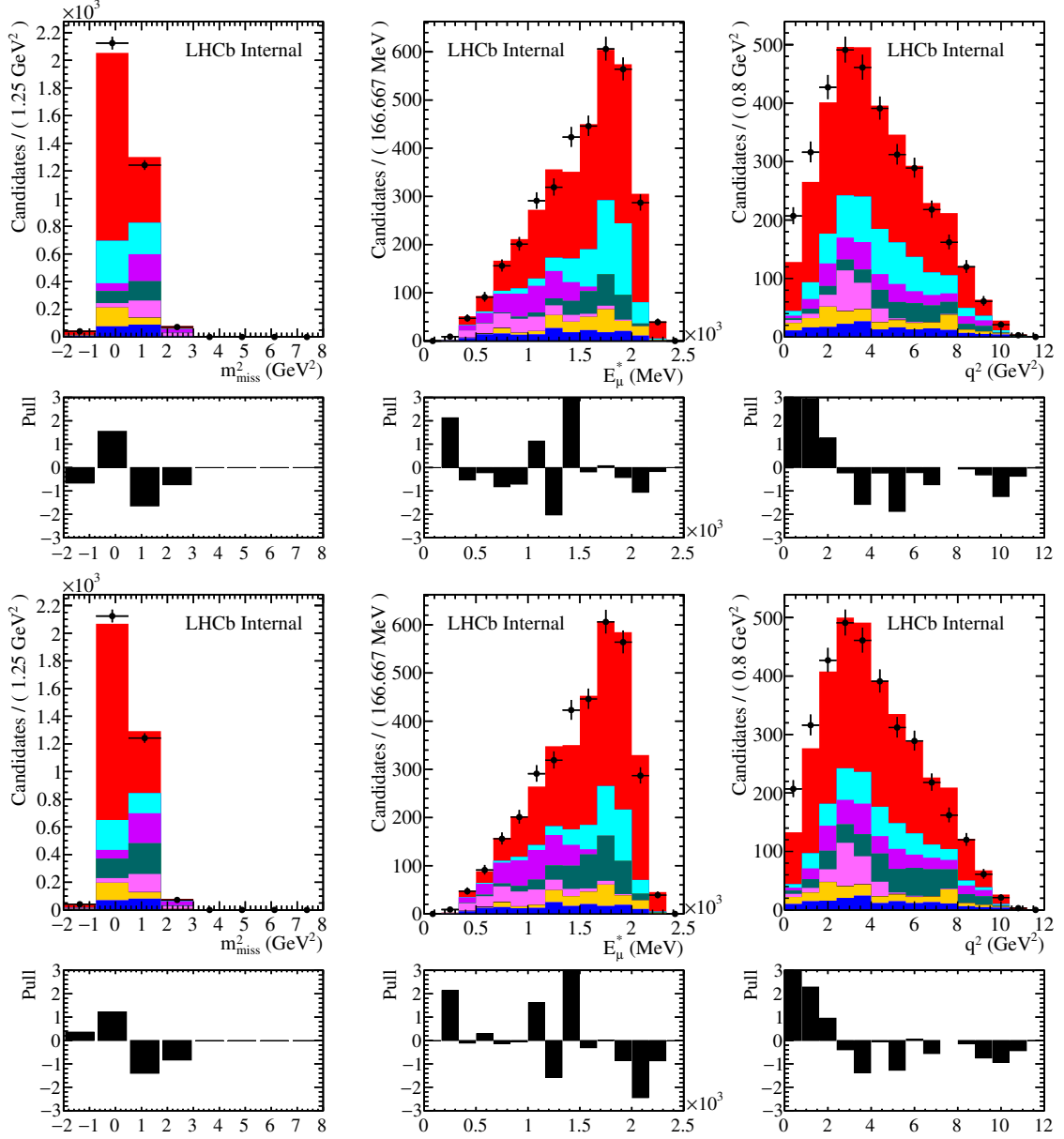


Figure 8.5: Projections of the (top) first and (bottom) second fit iterations in m_{miss}^2 , E_{μ}^* and q^2 to the Run 1 data sample for the $B \rightarrow p\bar{p}$ transition form factor parameters.

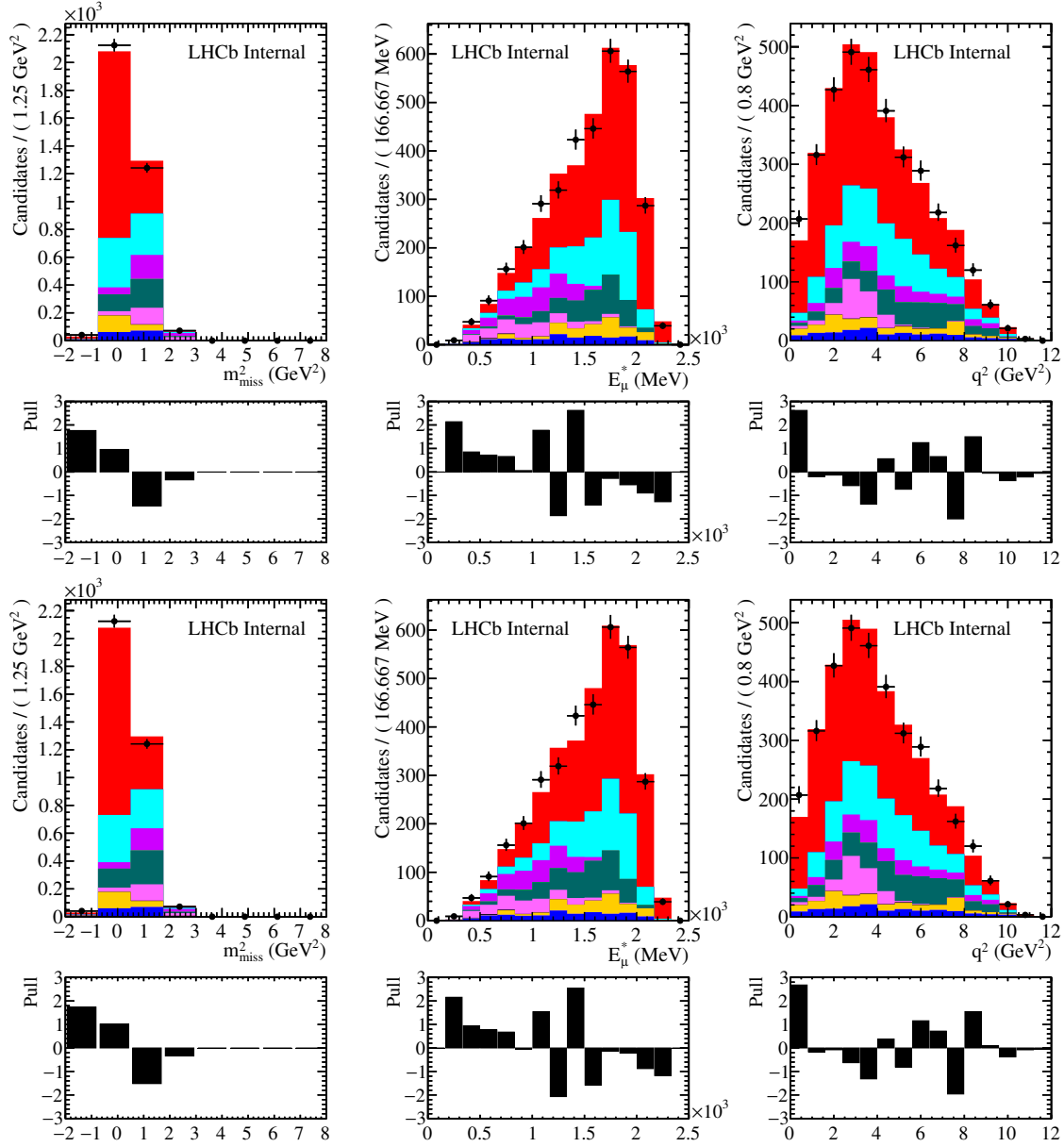


Figure 8.6: Projections of the (top) third and (bottom) fourth fit iterations in m_{miss}^2 , E_{μ}^* and q^2 to the Run 1 data sample for the $B \rightarrow p\bar{p}$ transition form factor parameters.

Comparisons between truth-level distributions between PHSP simulation and those using D_{fit}^2 are provided in Figs. 8.7 and 8.8 in order to illustrate the degree of change in the distributions as a result of the FF. Additional comparisons are given in Figs. 8.9 and 8.10 between simulation using D_{start}^2 and D_{fit}^2 in order to illustrate the effect of the iterative fit.

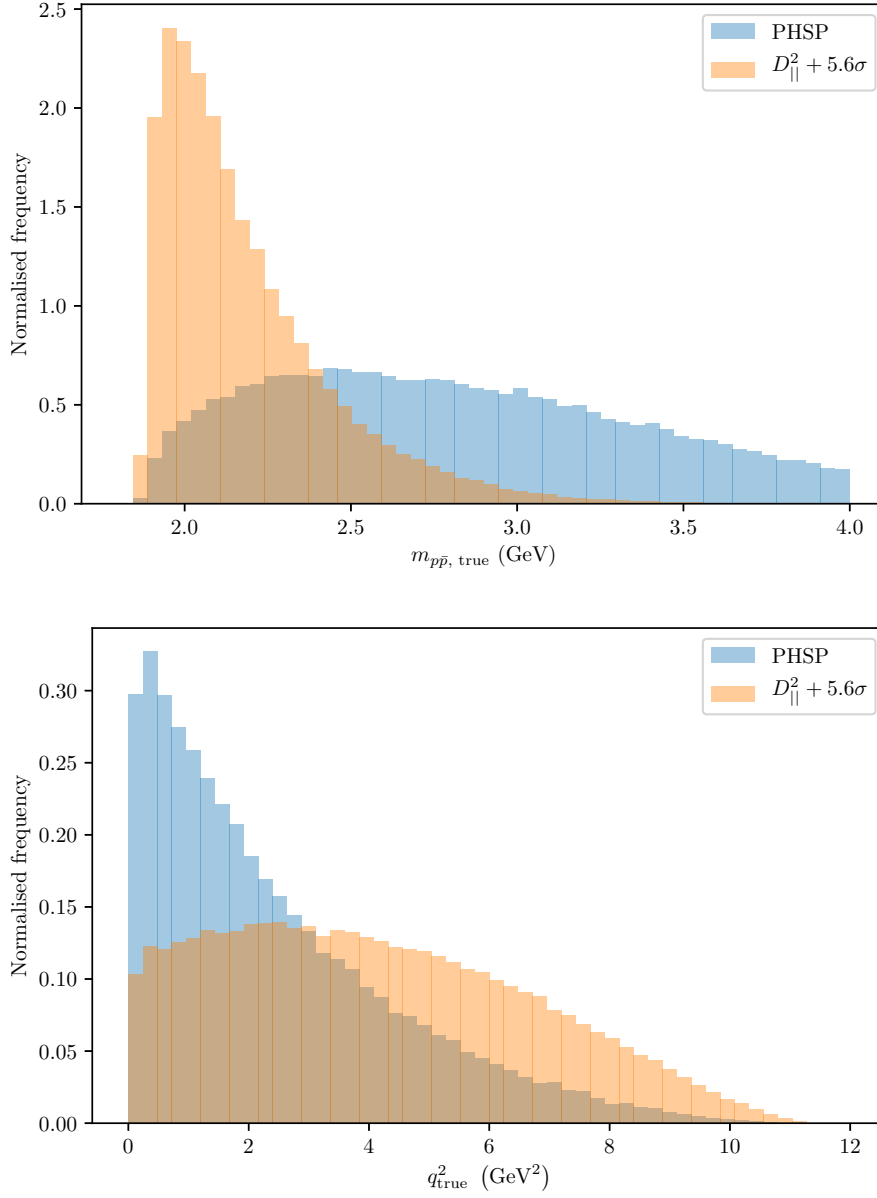


Figure 8.7: Comparisons between PHSP truth-level simulation distributions and those using D_{fit}^2 in $m(p\bar{p})$ and q^2 . The contrast between the broader distribution of the PHSP simulation in $m(p\bar{p})$ and that of the pQCD model is clear.

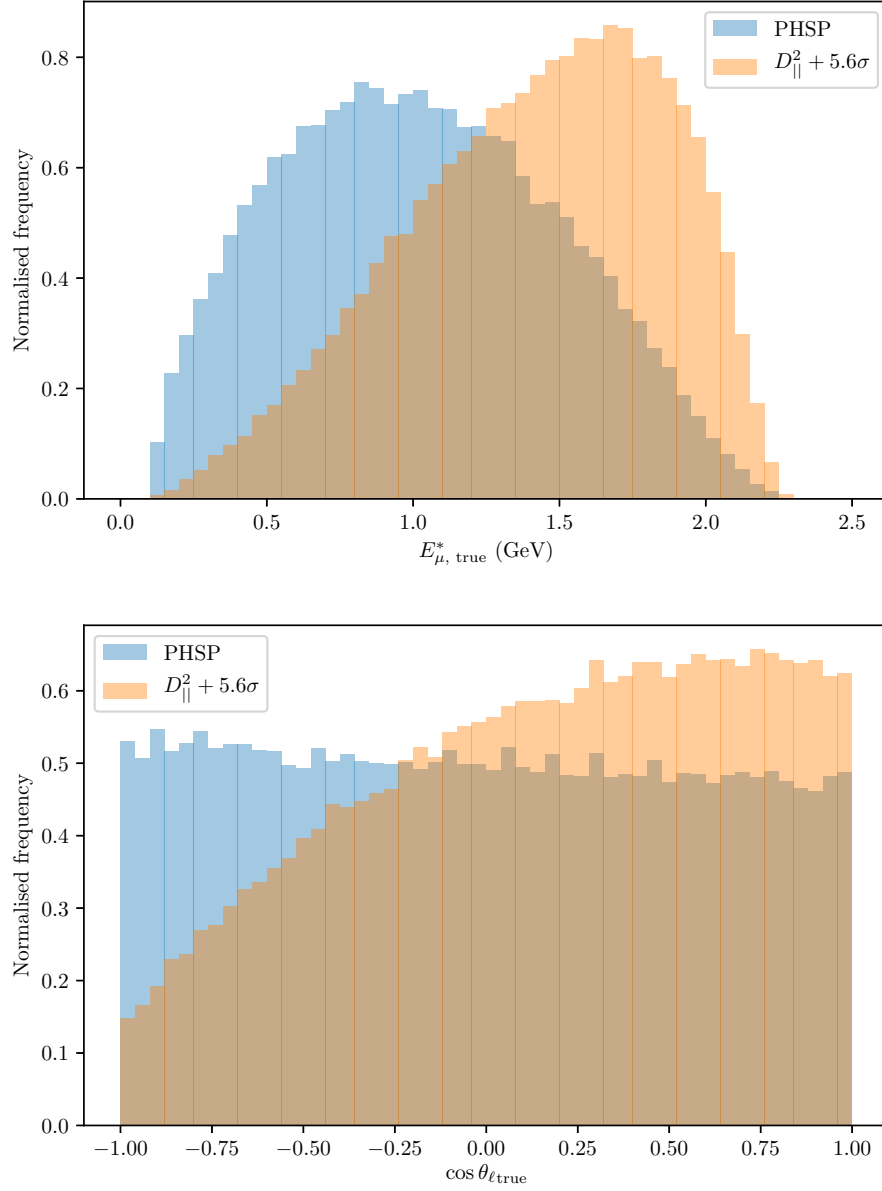


Figure 8.8: Comparisons between PHSP truth-level simulation distributions and those using D_{fit}^2 in E_{μ}^* and $\cos \theta_{\ell}$.

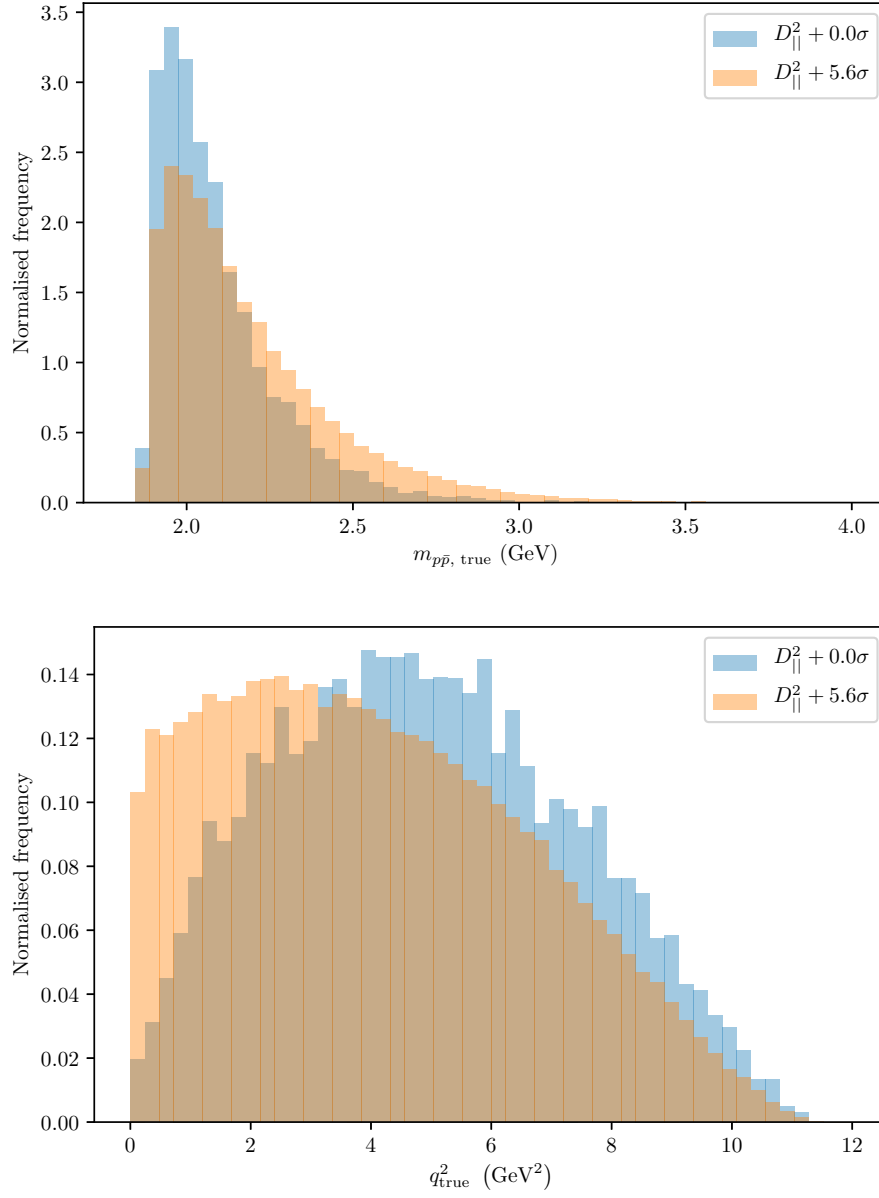


Figure 8.9: Comparisons between truth-level simulation distributions using D_{start}^2 and those using D_{fit}^2 in $m(p\bar{p})$ and q^2 . The $D_{||}^2$ parameter has a clear effect on the distributions shapes.

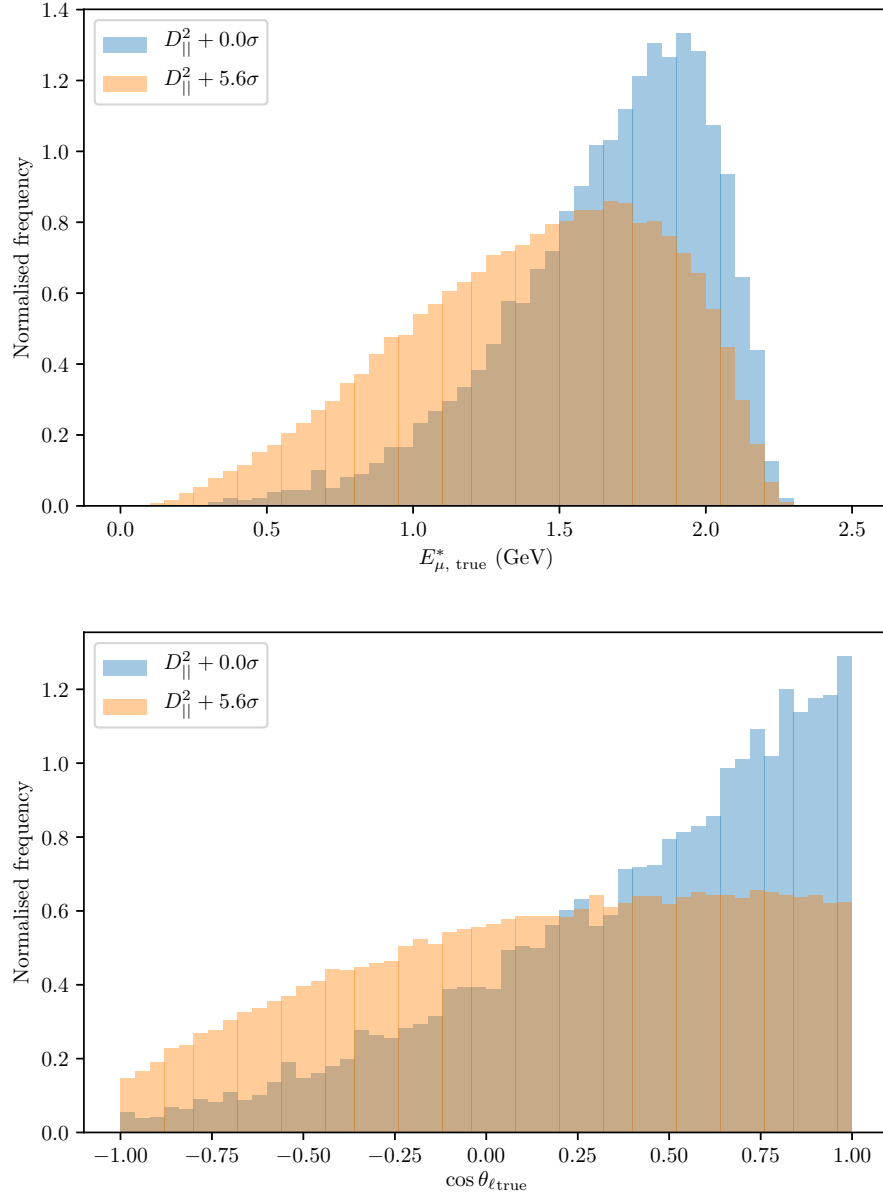


Figure 8.10: Comparisons between truth-level simulation distributions using D_{start}^2 and those using D_{fit}^2 in E_{μ}^* and $\cos \theta_{\ell}$.

Chapter 9

Efficiency ratio

This chapter details the calculation of the ratio of signal and normalisation efficiencies that is required to determine $R(p\bar{p})$ from the maximum likelihood fit to data. The efficiency ratio takes the factorised form:

$$\begin{aligned} \frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)} &= \frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{DPC}}} \\ &\times \frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{RECO\&STRIP|DPC}}} \\ &\times \frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{HLT2|RECO\&STRIP,DPC}}} \\ &\times \frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{L0\&HLT1|HLT2,RECO\&STRIP,DPC}}} \\ &\times \frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{SEL|L0\&HLT1,HLT2,RECO\&STRIP,DPC}}}, \end{aligned}$$

where DPC represents the decay product acceptance (DecProdCut) cut on the simulation samples, RECO&STRIP represents the reconstruction and stripping, HLT2 and L0&HLT1 represent the trigger and SEL represents the offline selection, including the PID cuts. In all cases, the efficiency ratios are evaluated from simulation, with the relevant weights included at each stage. More details on each calculation are given in what follows, where both the absolute and relative efficiencies are given. Various sources of systematic uncertainty are inherent in the calculation of the efficiency ratio. These are discussed in more detail in Sec. 11. For convenience, the conditional part of the subscripts above are dropped in what follows.

9.1 Acceptance cut on decay products

At generator-level, a cut is applied to the decay products of the B^+ such that they fall within the acceptance of the LHCb detector. A cut on the pseudorapidity is made of $2.0 < \eta < 5.0$. The efficiency of this cut for the signal and normalisation modes is taken from the simulation samples with weights that correct the B^+ kinematics (see Sec. 5.3) applied. A statistical uncertainty is associated with the efficiency due to the total number of simulated events.

Efficiency	Value
$\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)_{\text{DPC}}$	0.2140 ± 0.0003
$\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{DPC}}$	0.2037 ± 0.0003
$\frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{DPC}}}$	1.050 ± 0.002

Table 9.1: A summary of the absolute and relative DPC efficiencies.

9.2 Reconstruction and stripping

The reconstruction and stripping efficiency is determined by using fully reconstructed simulation with the stripping cuts and the kinematic weights applied. The efficiencies are shown in Table. 9.2. As expected, the efficiency is lower for $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ than for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ as the softer muons originating from the $\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau$ decay are more difficult to reconstruct.

Efficiency	Value
$\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)_{\text{RECO\&STRIP}}$	0.01942 ± 0.00007
$\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{RECO\&STRIP}}$	0.03706 ± 0.00009
$\frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{RECO\&STRIP}}}$	0.524 ± 0.002

Table 9.2: A summary of the absolute and relative reconstruction and stripping efficiencies.

9.3 HLT2

The HLT2 selection places a number of cuts on the reconstructed candidates which are broadly in line with those of the stripping line. As can be seen from Table 9.3, the HLT2 efficiency for both modes is above 99%.

Efficiency	Value
$\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)_{\text{HLT2}}$	0.9962 ± 0.0002
$\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{HLT2}}$	0.9967 ± 0.0001
$\frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{HLT2}}}$	0.9994 ± 0.0003

Table 9.3: A summary of the absolute and relative HLT2 efficiencies.

9.4 L0 and HLT1

The efficiency of the L0 and HLT 1 triggering requirements are determined using the set of weights w_{trig} , as outlined in Sec. 5.4. The efficiencies are shown in Table 9.4. The efficiency for $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ is lower than that of $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ as the L0Muon trigger efficiency is lower for the softer muons originating from the $\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau$ decay.

Efficiency	Value
$\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)_{\text{L0\&HLT1}}$	0.5275 ± 0.0018
$\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{L0\&HLT1}}$	0.6796 ± 0.0012
$\frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{L0\&HLT1}}}$	0.7762 ± 0.0030

Table 9.4: A summary of the absolute and relative L0&HLT1 efficiencies.

9.5 Offline selection

The offline selection includes cuts placed on both the charged and neutral isolation BDTs, the opening angle of the diproton system $\cos\theta_{p\bar{p}}$ and the PID cuts on

$\text{ProbNN}_p^{\text{combined}}$ and $\text{ProbNN}_\mu^{\text{combined}}$. The efficiency of the PID cuts is determined using the weights outlined in Sec. 5.6. This efficiency also includes cuts placed on the fit variables m_{miss}^2 , q^2 and E_μ^* that are used to define the fit range. The efficiencies are shown in Table 9.5.

Efficiency	Value
$\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)_{\text{SEL}}$	0.1244 ± 0.0016
$\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{SEL}}$	0.1665 ± 0.0012
$\frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_{\text{SEL}}}$	0.7476 ± 0.0110

Table 9.5: A summary of the absolute and relative offline efficiencies.

9.6 Total efficiency ratio

The total efficiency ratio is determined by taking the product of the efficiencies outlined in this section, and is found to be

$$\frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)} = 0.3078 \pm 0.0049.$$

Chapter 10

Fit to data

This chapter outlines the details of the 3D extended maximum-likelihood fit used to determine $R(p\bar{p})$ from the data. Firstly, the treatment of the fit parameters is discussed, and an overview of the templates that are used to construct the likelihood is then given. This is followed by the results of the fit to data, before concluding with the validation of the fit model.

10.1 Fit parameters

The fit parameters are summarised in Table 10.1. $R(p\bar{p})$ is left free to float in the fit, as well as the yield $\mathcal{N}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)$ of the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ component and all of the part-reco background components. The p mis-ID template is divided into two parts; one where the shape is derived from the fake- p sample in which the proton shares its charge with the muon, and the other where the proton and muon have opposite sign. These are denoted p mis-ID 1 and p mis-ID 2 respectively. A dedicated HLT2 trigger line is used to select p mis-ID 2, but no such line is available to determine p mis-ID 1. Therefore p mis-ID 2 is constrained using the unfolding procedure in Sec. 6.2, whilst the yield of p mis-ID 1 is left free to float. The constraint on the p mis-ID 2 yield is Gaussian, centred on the yield obtained from the unfolding procedure, and with a width corresponding to the uncertainty on this yield. The yields of the μ mis-ID and combinatorial background are left free to float in the fit. Projections of the templates for the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ and $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ components are shown in Figs. 10.1 and 10.2 respectively, and those of the dominant background components are given in Appendix E. The distributions for the signal component are broad, whilst those of the normalisation mode have a more distinctive shape, particularly in m_{miss}^2 .

Parameter	Constr. (C)/Floating (F)	Sample
$R(p\bar{p})$	F	Simulation
$\mathcal{N}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)$	F	Simulation
$\mathcal{N}(B^{+/0} \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X)$	F	Simulation
$\mathcal{N}(B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu)$	F	Simulation
$\mathcal{N}(B^0 \rightarrow (\bar{\Lambda}_c^- \rightarrow \mu^- pX)\bar{p}X)$	F	Simulation
$\mathcal{N}(B^+ \rightarrow p(\Delta \rightarrow \bar{p}X)X\mu^+\nu_\mu)$	F	Simulation
$\mathcal{N}(B^+ \rightarrow p(\bar{N}^{*-} \rightarrow \bar{p}X)\mu^+\nu_\mu)$	F	Simulation
$\mathcal{N}(B^+ \rightarrow p(\bar{\Lambda}_c^{*-} \rightarrow \bar{p}X)\mu^+\nu_\mu)$	F	Simulation
$\mathcal{N}(B^+ \rightarrow p\bar{p}D^0\mu^+\nu_\mu X)$	F	Simulation
$\mathcal{N}(B_s^0 \rightarrow (D_s^+ \rightarrow \mu^+ X)p\bar{p}X)$	F	Simulation
$\mathcal{N}(B_s^0 \rightarrow (D^0 \rightarrow K\mu^+ X)p\bar{p}X)$	F	Simulation
$\mathcal{N}(p \text{ mis-ID1})$	F	Fake- p
$\mathcal{N}(p \text{ mis-ID2})$	C	Fake- $\bar{p}$
$\mathcal{N}(\mu \text{ mis-ID})$	F	Fake- μ
$\mathcal{N}(\text{Comb.})$	F	Event mixing

Table 10.1: A summary of the parameters, their treatment and the samples used for each component in the fit used to determine $R(p\bar{p})$.

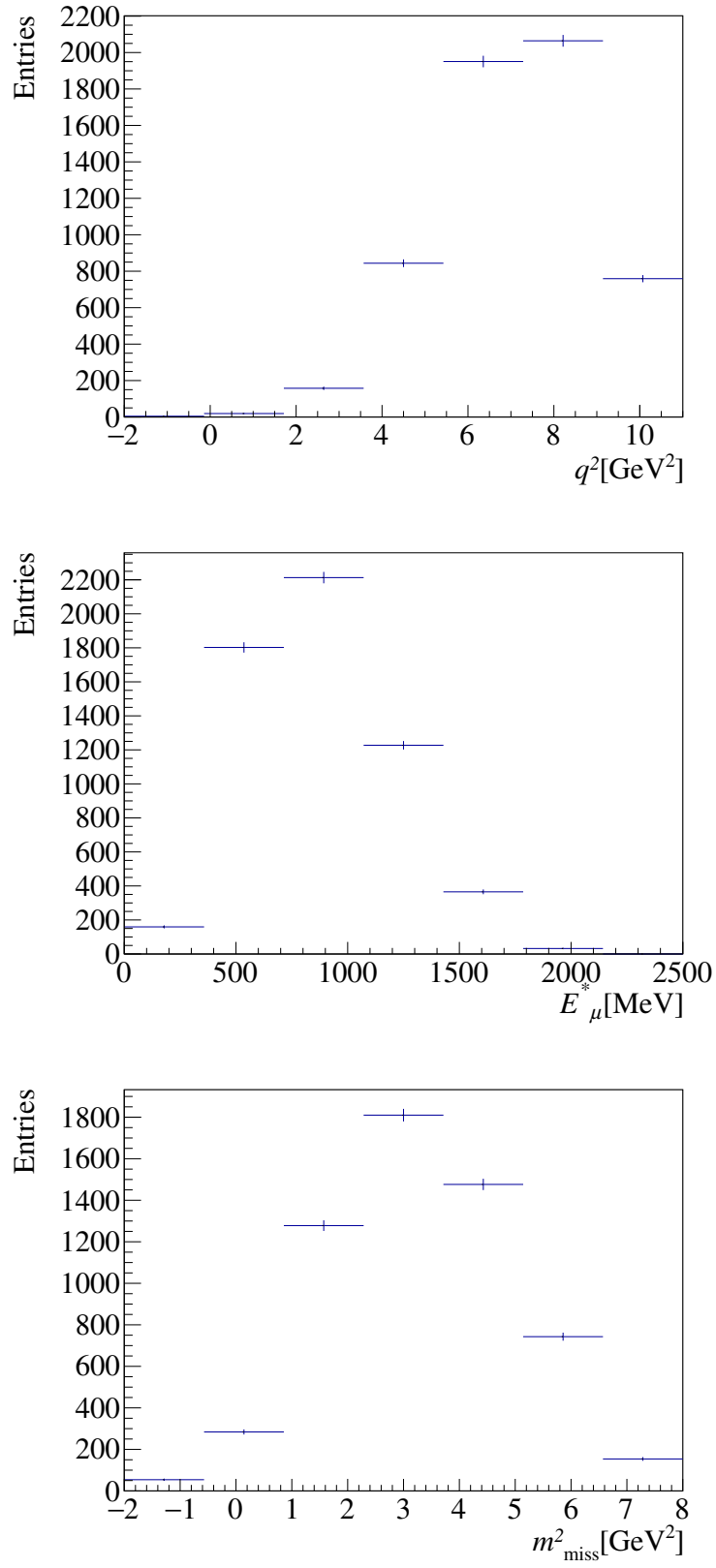


Figure 10.1: Projections of the 3D histogram containing the selected $B^+ \rightarrow p \bar{p} \tau^+ \nu_\tau$ simulation sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

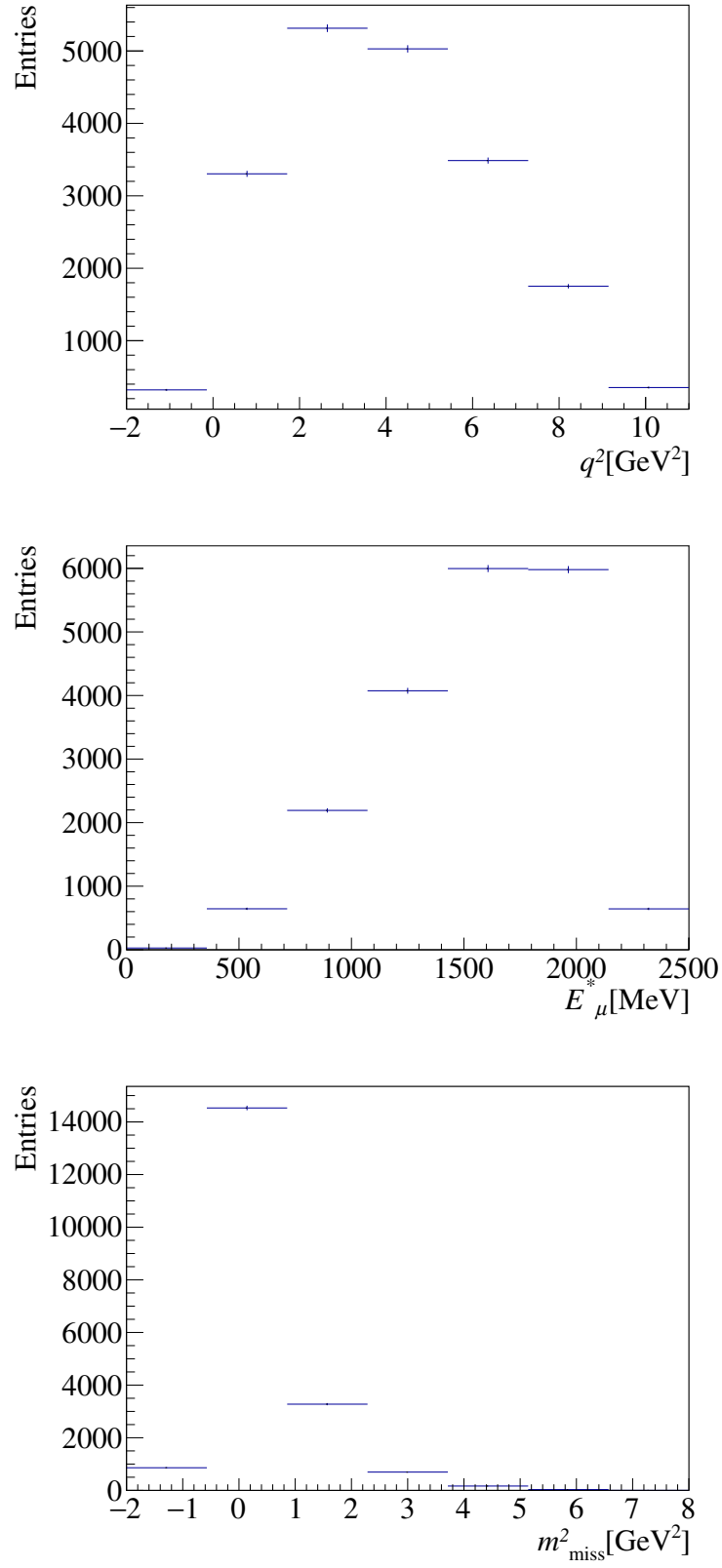


Figure 10.2: Projections of the 3D histogram containing the selected $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ simulation sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

A final nuisance parameter $\alpha_{D_{||}^2}$ is introduced that results from template shape variations included for the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$, $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$, $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{\Delta}\mu^+\nu_\mu$ components. The shapes of these components are free to vary using the AddHistoSys function in an analogous way to the FF fit outlined in Sec. 8. The shape variations represent $\pm 1\sigma$ variations in the $D_{||}^2$ parameter about the central value obtained from the FF fit, where σ is the uncertainty quoted for $D_{||}^2$ in Eq. (8.2). Hence, the value of the nuisance is the deviation from the central value, in units of the original theory uncertainty. The shape variations for the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ and $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ components are shown in Figs. 10.3 and 10.4 and those for the $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{\Delta}\mu^+\nu_\mu$ are given in Appendix. F. The FF parameter predominantly affects the q^2 and E_μ^* shapes, whilst the m_{miss}^2 shape is constrained by the available phase space of the decay products.

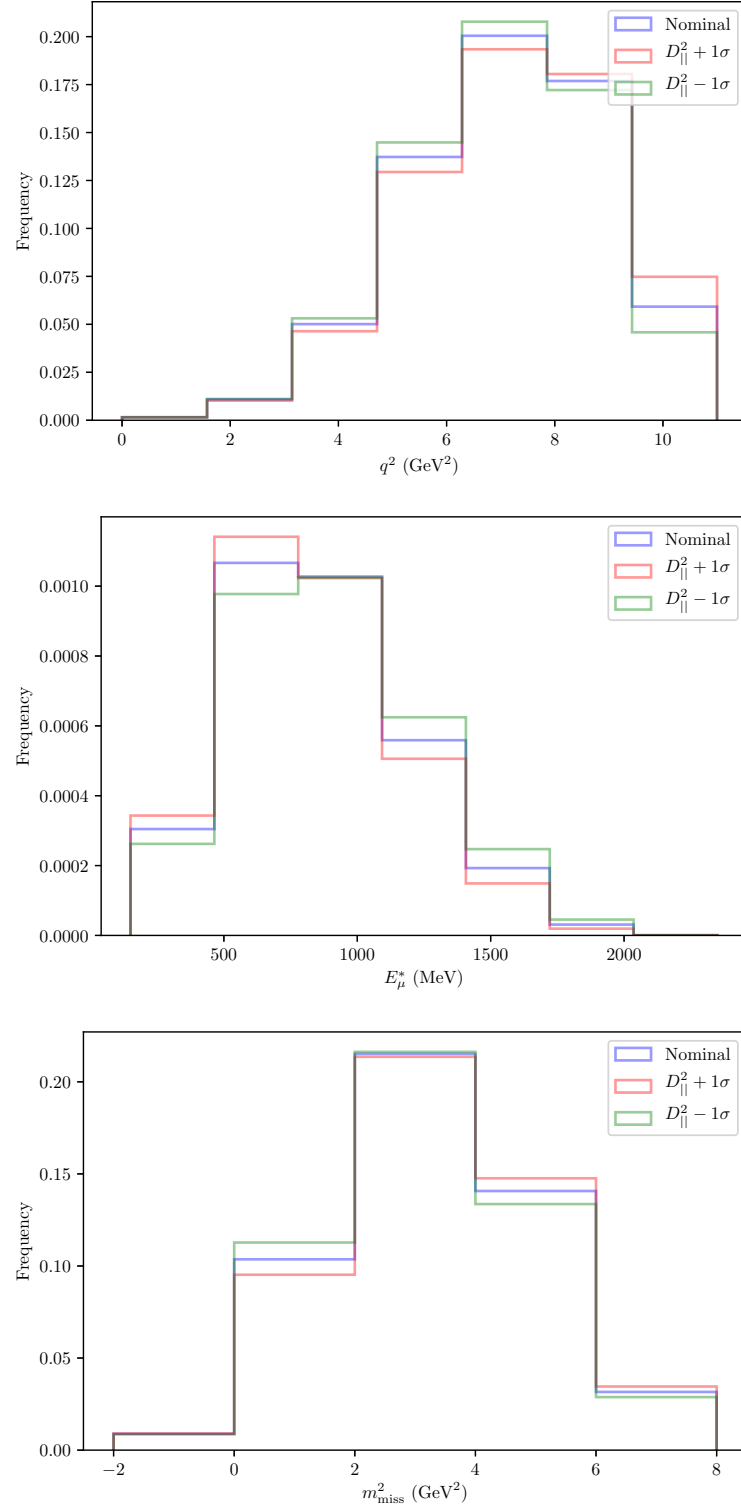


Figure 10.3: Shape variations included for the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ template resulting from $\pm 1\sigma$ fluctuations in the FF parameter $D_{||}^2$ for (top) q^2 , (middle) E_{μ}^* and (bottom) m_{miss}^2 .

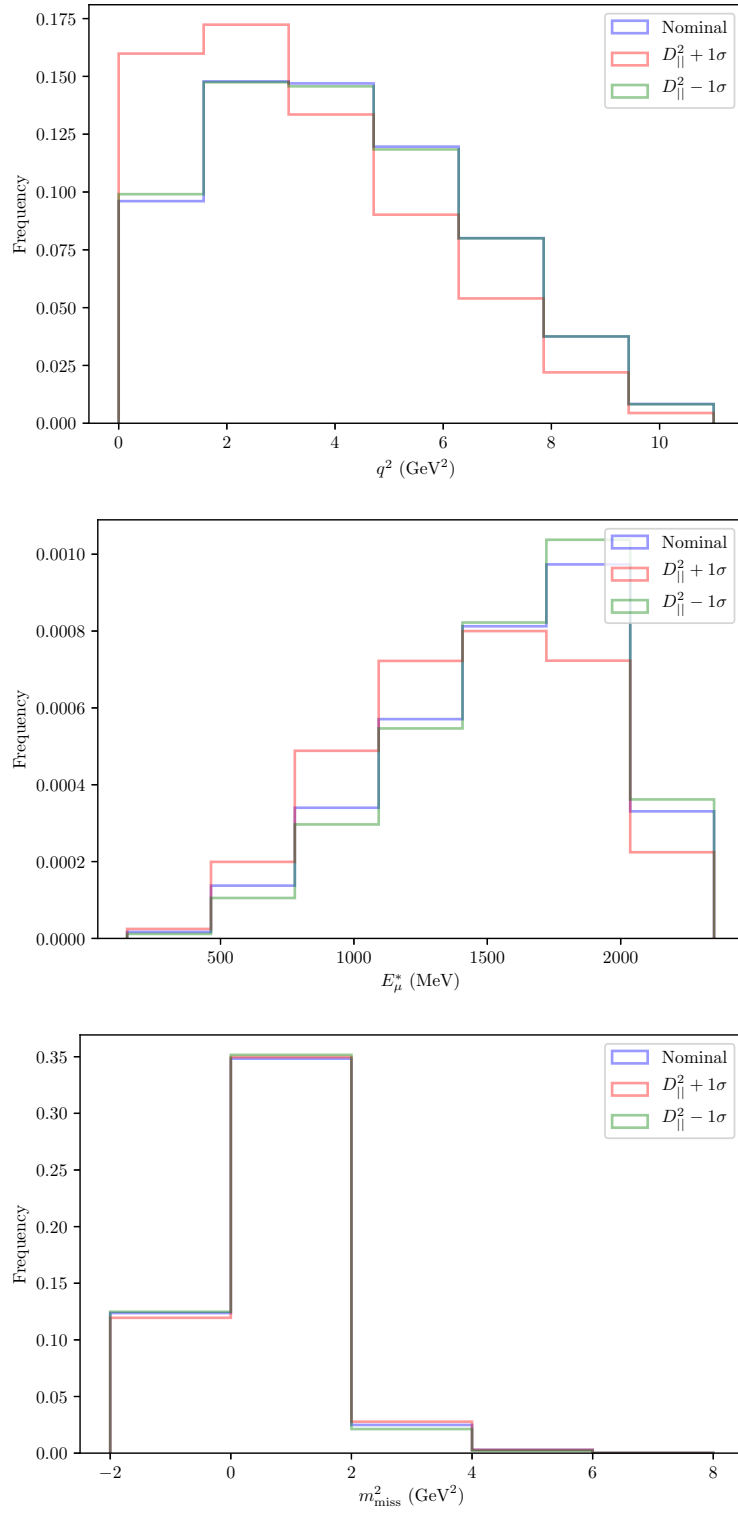


Figure 10.4: Shape variations included for the $B^+ \rightarrow p\bar{p}\mu^+\nu_{\mu}$ template resulting from $\pm 1\sigma$ fluctuations in the FF parameter $D_{||}^2$ for (top) q^2 , (middle) E_{μ}^* and (bottom) m_{miss}^2 .

10.2 Fit to data

The likelihood function $\mathcal{L}(R(p\bar{p}))$ is given by

$$\begin{aligned} \mathcal{L}(R(p\bar{p})) = & \text{Pois} \left(n \middle| R(p\bar{p})c_{\tau\mu} + \sum_i \mathcal{N}(X_i) \right) \\ & \times \left[\prod_{e=1}^n \frac{R(p\bar{p})c_{\tau\mu}f_{\tau}(x_e, y_e, z_e) + \sum_i \mathcal{N}(X_i)f_i(x_e, y_e, z_e)}{R(p\bar{p})c_{\tau\mu} + \sum_i \mathcal{N}(X_i)} \right] \\ & \times \mathcal{G}(N(p \text{ mis-ID2}), \sigma(p \text{ mis-ID2}) | \mathcal{N}(p \text{ mis-ID2})), \end{aligned} \quad (10.1)$$

where $\text{Pois}(n|A)$ is the Poisson probability of obtaining n events given A expected events, the index i runs over the fit components and $e = 1, \dots, n$ runs over the events in the data sample. The variables $\mathcal{N}(X_i)$ represent the nuisance parameters in the fit and correspond to the yields of the background components outlined in Table. 10.1. The functions $f_{\tau}(x_e, y_e, z_e)$ and $f_i(x_e, y_e, z_e)$ are the PDFs of the signal and background components as obtained from simulation, respectively. The Gaussian constraint on the yield of the p mis-ID 2 component is represented by function $\mathcal{G}(\mu, \sigma|x)$, where $\mu = N(p \text{ mis-ID2})$ is the expected yield as obtained from the unfolding procedure, $\sigma = \sigma(p \text{ mis-ID2})$ is the associated uncertainty and $x = \mathcal{N}(p \text{ mis-ID2})$ is the variable that is subject to the constraint. The coefficient $c_{\tau\mu}$ is given by

$$c_{\tau\mu} = \mathcal{N}(B^+ \rightarrow p\bar{p}\mu^+\nu_{\mu}) \times \mathcal{B}(\tau^+ \rightarrow \mu^+\nu_{\mu}\bar{\nu}_{\tau}) \times \frac{\epsilon(B^+ \rightarrow p\bar{p}\tau^+\nu_{\tau})}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_{\mu})} \quad (10.2)$$

and facilitates the definition of $R(p\bar{p})$ as the parameter of interest, rather than the raw yield of the $B^+ \rightarrow p\bar{p}\tau^+\nu_{\tau}$ component as is the case for the nuisance parameters. The negative logarithm of this likelihood is minimised by the Minuit minimisation package [107]. This minimisation is equivalent to the maximisation of the likelihood, as is the objective in a maximum-likelihood fit.

The projections of the three dimensional fit to data are shown in Figs. 10.5 and 10.6. Upon convergence, the best-fit value for $R(p\bar{p})$ is in fact negative and consistent with zero within 1.7σ , where σ is the statistical uncertainty. Fig. 10.7 shows the likelihood minimized at a range of different $R(p\bar{p})$ values, or the *profile* likelihood in $R(p\bar{p})$, where the minimum is clearly visible.

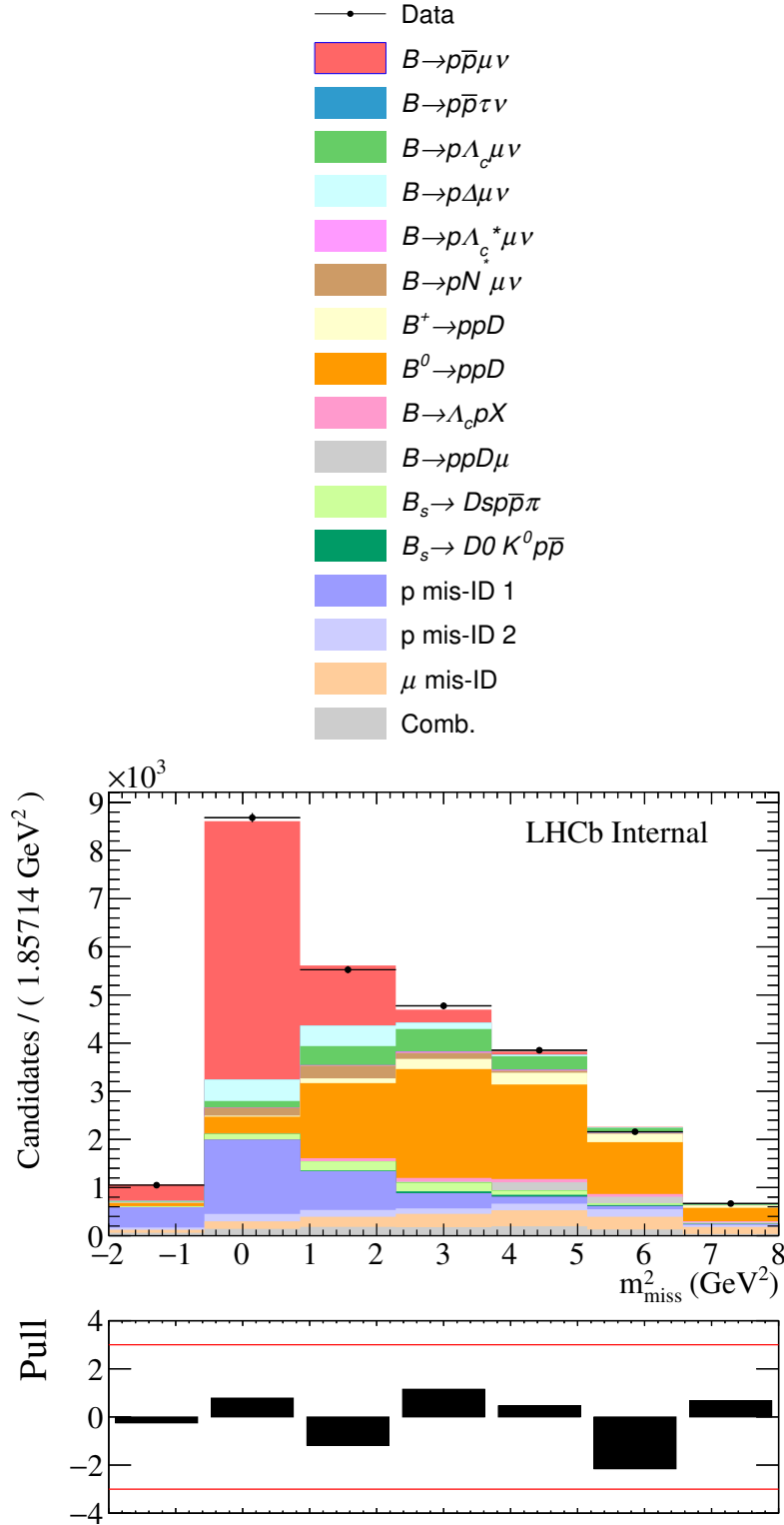
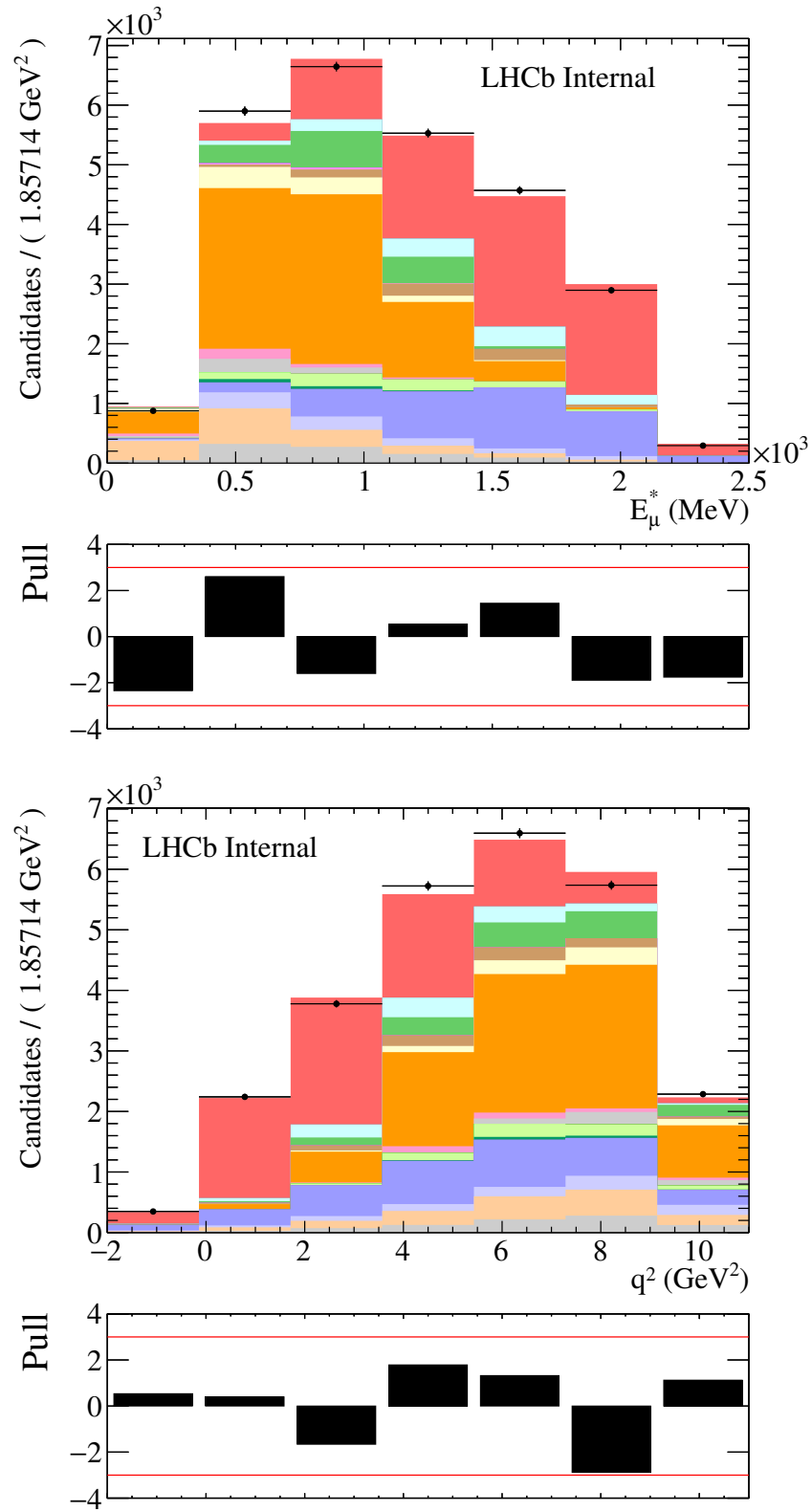


Figure 10.5: (Top) Legend for the fit projections. (Bottom) Projection of the fit for $R(p\bar{p})$ in m^2_{miss} .

Figure 10.6: Projection of the fit in (top) E_μ^* and (bottom) q^2 .

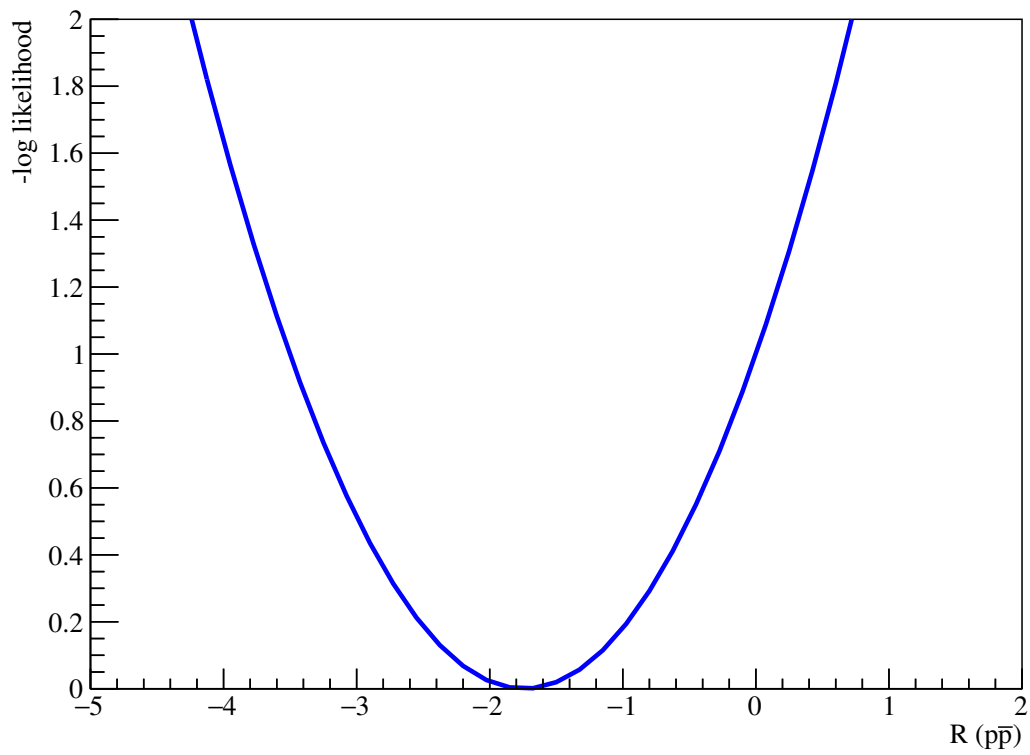


Figure 10.7: The negative log-likelihood minimised at different values of $R(p\bar{p})$.

The correlations between the fit parameters are shown in Fig. 10.8. Correlations arise due to similarities in the shapes of the fit components. The highest correlations are between modes which are most similar in terms of the kinematics of the decay products; namely the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$, $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{\Delta}\mu^+\nu_\mu$ components. The fit converges despite a few large correlations.

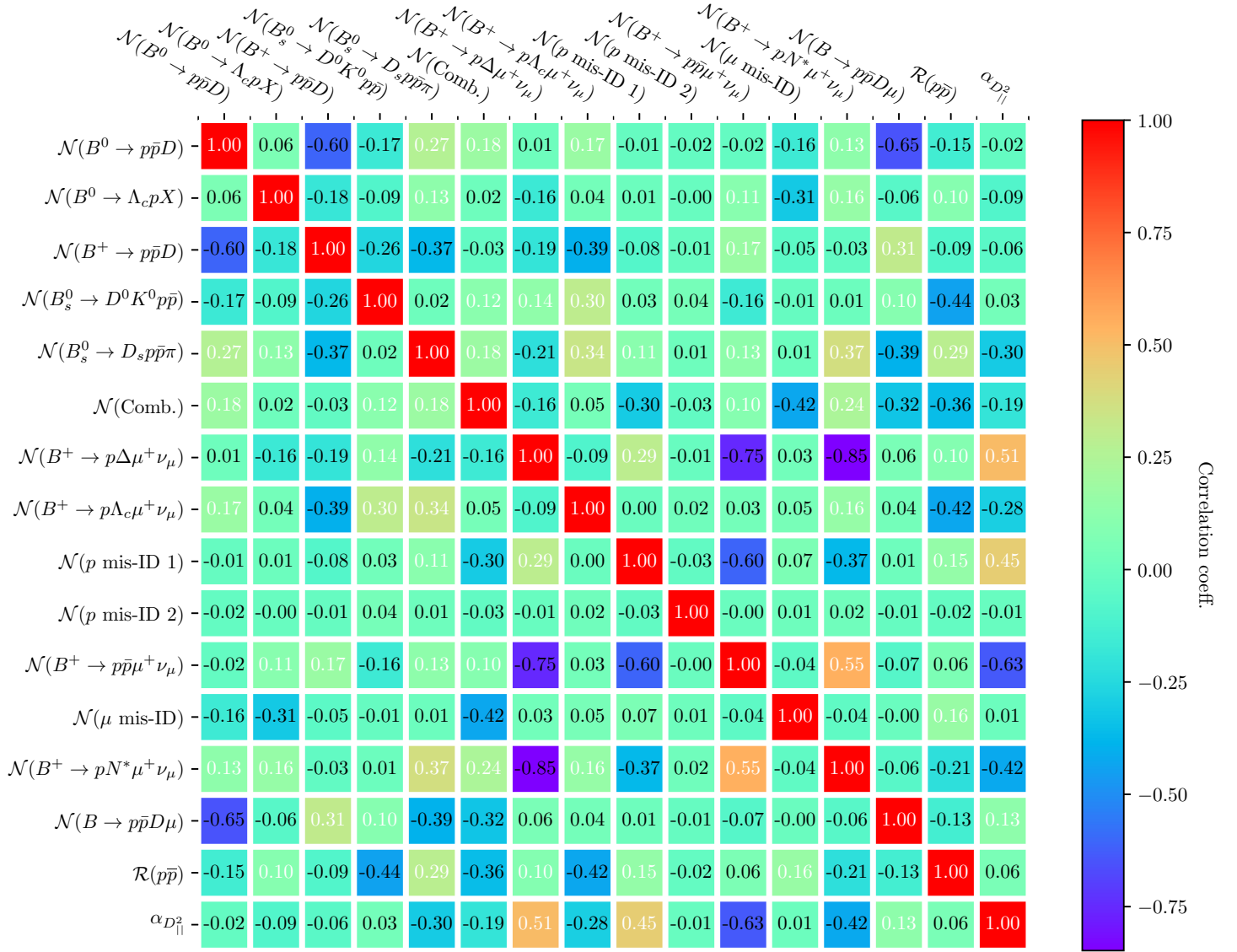


Figure 10.8: Correlation matrix of the fit to data for $R(p\bar{p})$. The largest correlations are for the components which are similar in terms of the kinematics of the final-state decay products.

10.3 Validation of data fit

The fit is validated by performing a number of pseudo-experiments. The purpose of validating the fit is to assess any inherent bias in the fit model and the coverage of the statistical uncertainty. A toy data set is generated by randomly sampling the model from the converged fit to data. A fit is performed to this toy data sample using the nominal fit model. The pull is recorded from a total of 2500 pseudo-experiments, and the mean of the pull distribution represents the bias of the fit model. The degree to which the standard deviation differs from 1 indicates the amount of under/over-coverage of the statistical uncertainty. The results of the fit validation are shown in Fig. 10.9 and indicate a bias of around 2% and a slight over-coverage of the statistical uncertainty. A Gaussian function is used to fit the pull distribution for illustration purposes only. The mean of the pull distribution is assigned as a systematic uncertainty to reflect the fit bias. The over-coverage is neglected given the large statistical uncertainty.

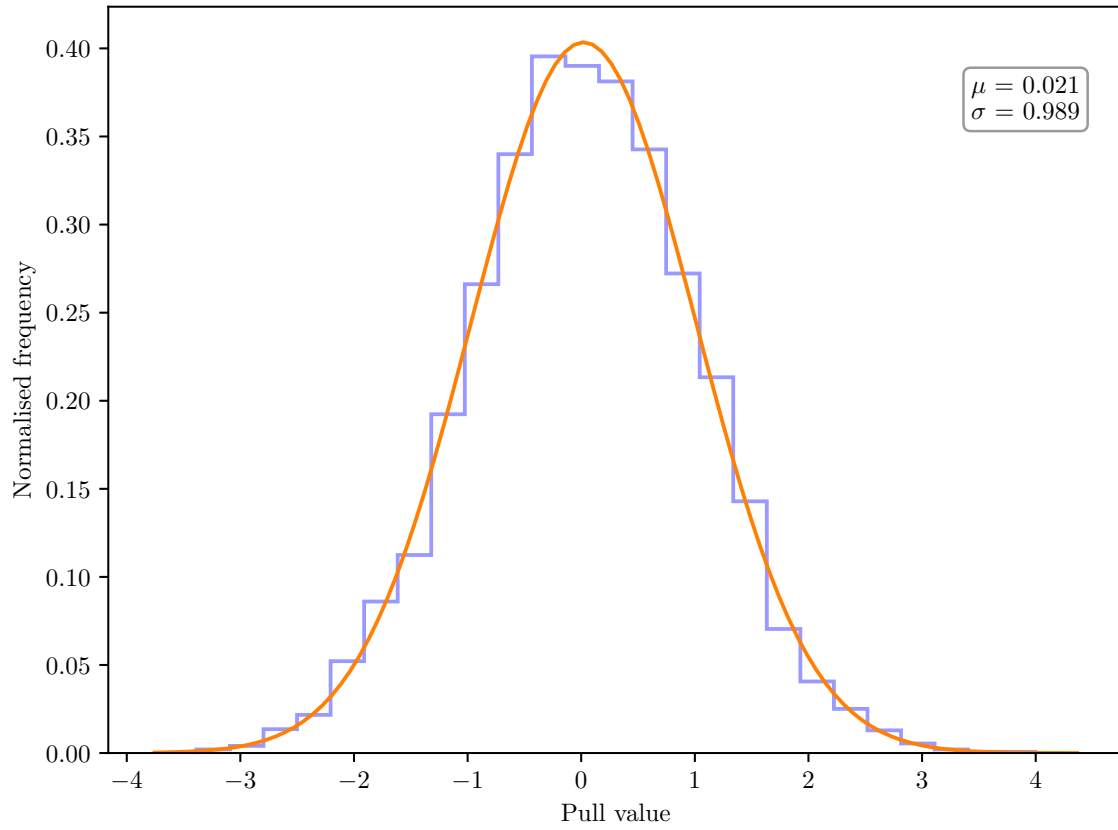


Figure 10.9: (Blue) Distribution of pulls on $R(p\bar{p})$ from fits to toy data samples generated and fitted using the nominal fit model with a (orange) Gaussian fit overlaid.

Chapter 11

Systematic uncertainties

This chapter outlines the sources of systematic uncertainty in the analysis. There are two types of systematic uncertainty; those associated with the efficiency ratio and those arising due to uncertainties in the template shapes used in the fit to data. Firstly, a summary of all of the systematic uncertainties is given. This is followed by an account of the uncertainties on the efficiency ratio in sections 11.2 to 11.5. These arise due to uncertainties associated with calibration samples or the simulation corrections. The efficiency ratio is re-computed a number of times using weights which result from some re-sampling of the nominal weights based on their uncertainties. The spread in the efficiency ratio resulting from this re-sampling is used to derive the systematic uncertainty.

A description of the uncertainties from the template shapes in the fit follows in sections 11.6 to 11.10. These are assessed by performing a number of pseudo-experiments. A toy data sample is generated from the nominal fit model. An alternative model, which constitutes a variation on the nominal model based on the uncertainty in question, is used to fit the toy data sample. The nominal model is also used to fit the same data sample. The difference in the best-fit value of $R(p\bar{p})$ between the nominal and alternative model is recorded for each pseudo-experiment. The mean of the distribution of these differences is taken as the systematic uncertainty. The uncertainties on the efficiency ratio are smaller, due to the cancellation resulting from the visible final state that $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ and $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ both have in common. Here, the largest uncertainty relates to the choice of parameterisation for the $B \rightarrow p\bar{p}$ form factors, as described in Sec. 8.

11.1 Systematic uncertainties summary

Table. 11.1 summarises the systematic uncertainties on the efficiency ratio. Those arising due to shape uncertainties in the fit are given in Table. 11.2. The dominant source of systematic uncertainty is due to the limited statistics of the simulation samples on the template shapes, and contributes an uncertainty of 40%. There is also a significant uncertainty of around 13% due to variations in the shape of the part-reco backgrounds due to the choice of physics model for these decays. However, this is not surprising as the $B^0 \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X$ component is the dominant part-reco background, and variations in its shape are likely to introduce significant changes to the behaviour of the fit.

Source	Uncertainty (%)
Tracking	1.00
Kinematic Reweighting	0.21
PID	0.29
Form factor	6.46
Total	6.74

Table 11.1: A summary of the systematic uncertainties on the efficiency ratio, quoted as a percentage of the total efficiency ratio, as given in Sec. 9.6.

Source	Uncertainty (%)
Sim. Stat.	40.0
Fit Bias	2.10
p mis-ID shape	1.37
μ mis-ID shape	0.70
Part-reco bkg. shape	13.93
Comb. bkg. shape	4.87
Total	43.16

Table 11.2: A summary of the systematic uncertainties due to shape variations, quoted as a percentage of the value of $R(p\bar{p})$ as obtained from the fit to data.

11.2 Tracking efficiency

The tracking correction histogram described in Sec. 5.5 has a systematic uncertainty in each bin of 0.8% [83]. The content of each bin is varied by taking a random number from a Gaussian distribution centred on the nominal value and with a width of 0.8% to produce a fluctuated histogram, and a total of 100 such histograms are produced. The efficiency ratio is re-computed 100 times using these histograms, and the spread in the distribution of the efficiency ratios is taken to be the systematic uncertainty.

11.3 Kinematic reweighting

There is an uncertainty associated with the choice of reweighting method for the kinematics of the B mesons. As described in Sec. 5.3, the BDT-based GBReweigher is used. An alternative reweighting is used based on the BinsReweigher algorithm [97]. This reweights distributions based on the ratio of bin contents between the original and target distributions. The result of the reweighting using this alternative method is shown in Fig. 11.1. The difference in the efficiency ratio between the two reweighting algorithms is taken as the systematic uncertainty.

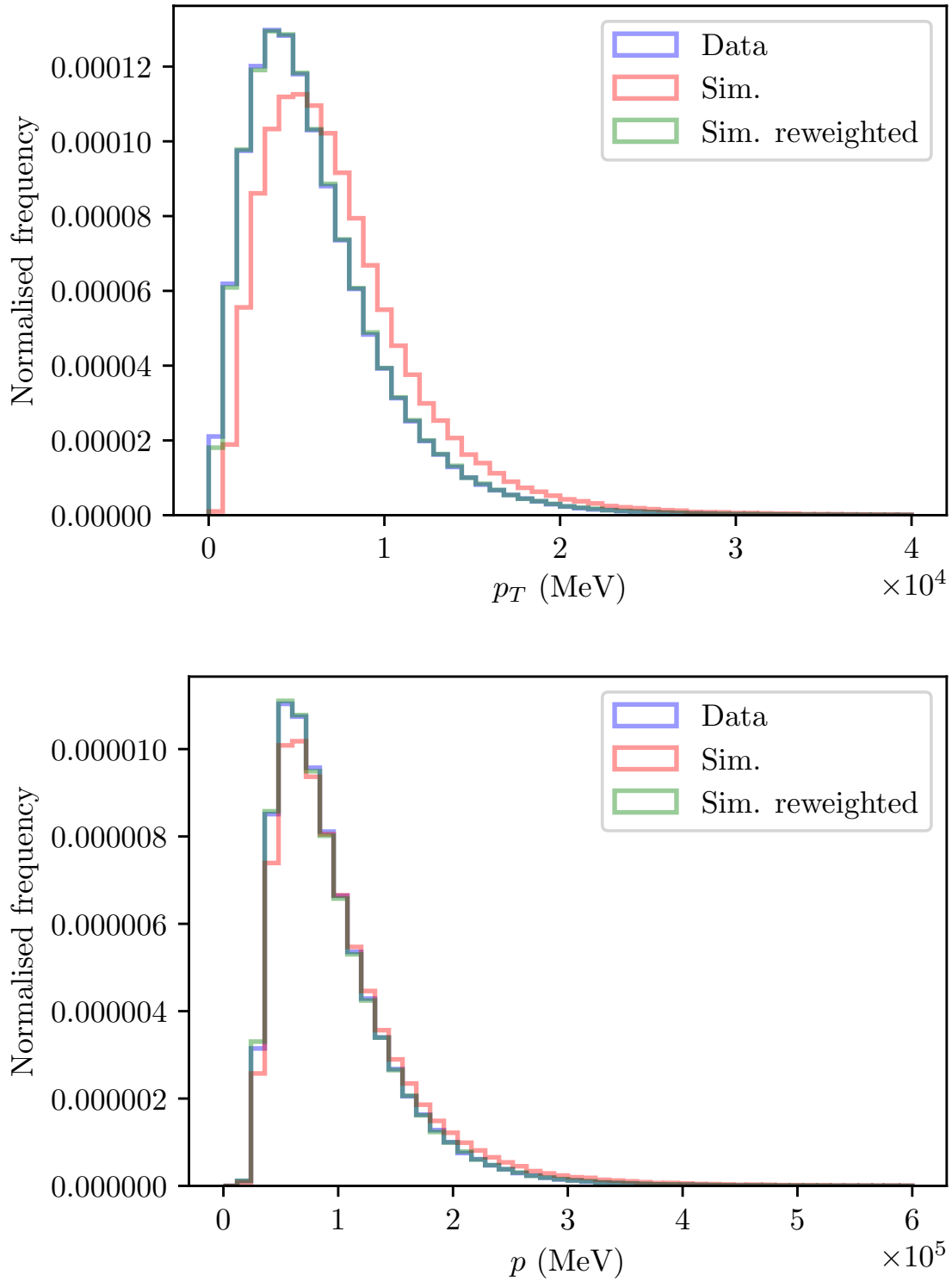


Figure 11.1: A comparison of the $B^+ \rightarrow J/\psi K^+$ before and after reweighting with the sWeighted $B^+ \rightarrow J/\psi K^+$ data in (top) $p_T(B)$ and (bottom) $p(B)$ using the Bin-sReweigher algorithm.

11.4 PID efficiency

The uncertainty on the PID efficiencies can be broken down into two contributions. The first comes from the statistics of the calibration histograms obtained from PID-Calib, as outlined in Sec. 5.6. However, this was found to contribute an uncertainty of $< 0.5\%$. The second comes from the choice of binning scheme used for the efficiency histograms. The systematic uncertainty is assessed by taking the difference in the efficiency ratio when using the nominal binning scheme and the alternative binning scheme shown in Fig. 11.2.

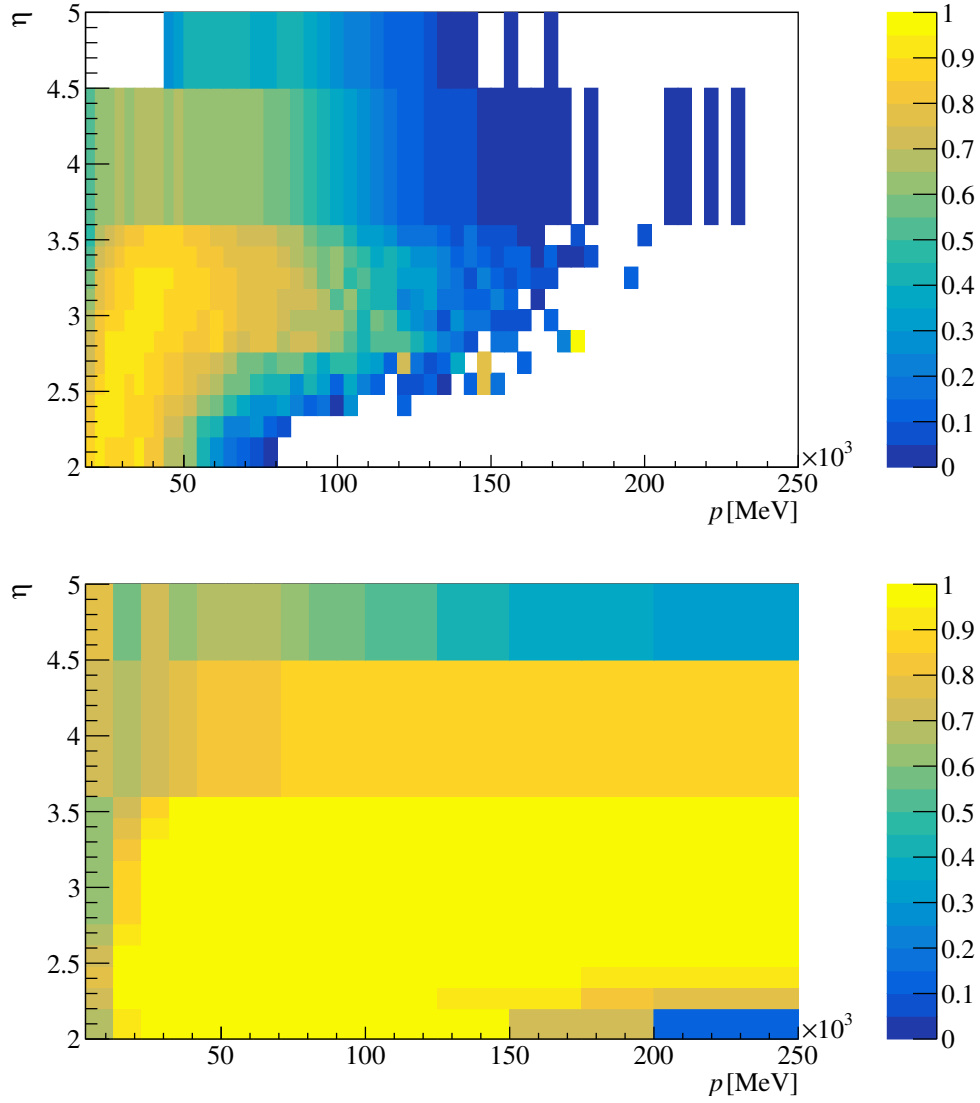


Figure 11.2: Projections of the PID efficiency as determined using PIDCalib in the $p - \eta$ plane for (top) protons with $\text{ProbNN}_p^{\text{combined}} > 0.5$ and (bottom) muons with $\text{ProbNN}_\mu^{\text{combined}} > 0.4$, using calibration samples from 2018 and the alternative binning scheme.

11.5 Form factor

As described in Sec. 8.1, an approximation is made in neglecting some of the form factor parameters, namely $D_{\parallel}^4$ and $D_{\parallel}^5$, that are expected to have sub-dominant contributions to the distribution shapes of $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays when compared with the $D_{\parallel}^2$ parameter. A systematic uncertainty is determined in order to assess the accuracy of this approximation.

The $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ and $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ templates are reweighted to variations of the nominal pQCD model, where the parameters $D_{\parallel}^4$ and $D_{\parallel}^5$ are set to $D_{\parallel}^4 = \pm 5\sigma$ and $D_{\parallel}^5 = \pm 5\sigma$, where σ is the uncertainty quoted on the parameters in Eq. 8.2. The efficiency ratio is re-calculated using these alternative weights, and the spread in the efficiency ratio is taken as the systematic uncertainty. Examples of the resulting shape variations are shown in Fig. 11.3 and 11.4.

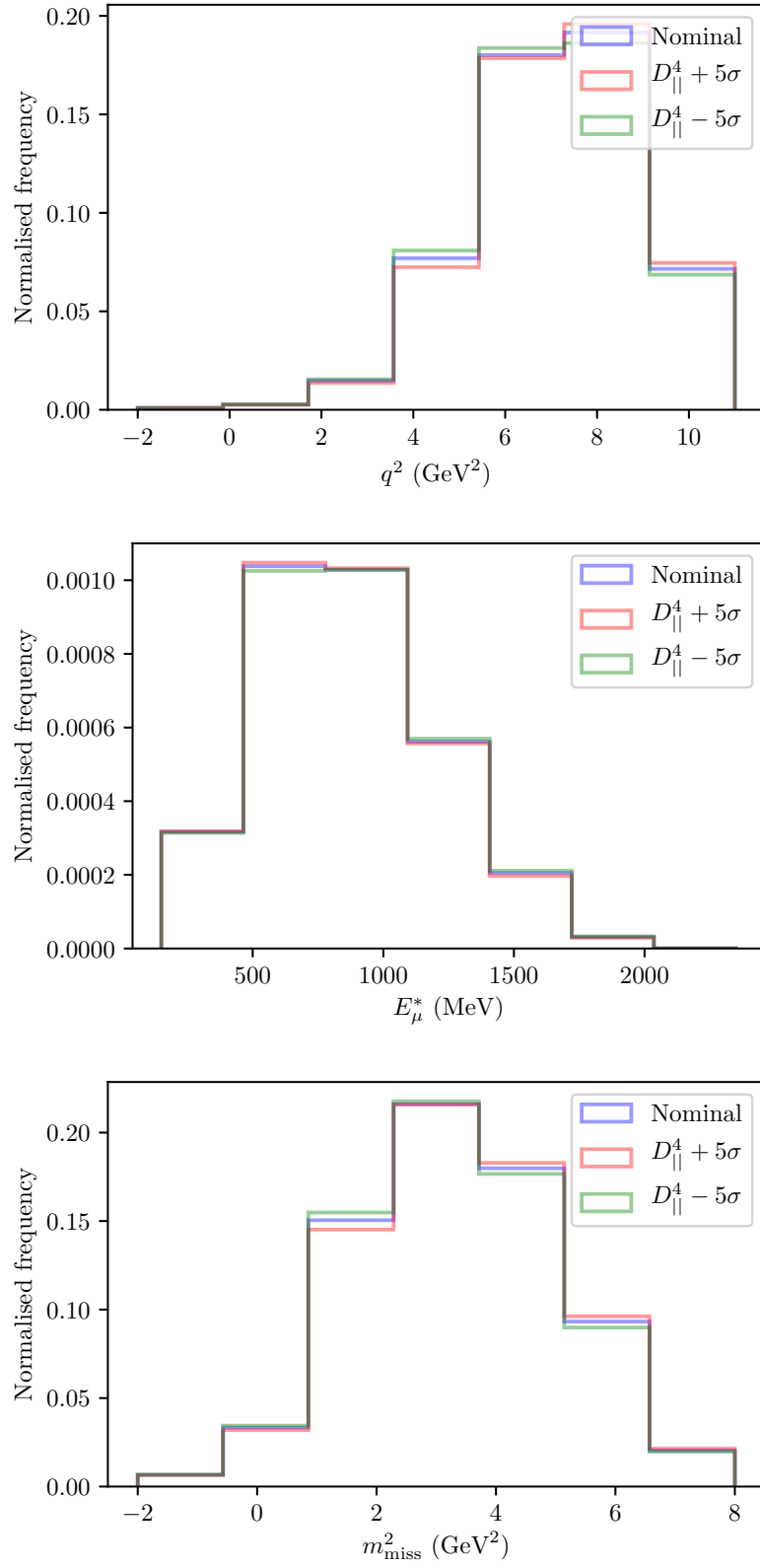


Figure 11.3: Comparison between the nominal template projections for $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ and those with alternative weights based on variations of the pQCD model with $D_{\parallel}^4 \pm 5\sigma$.

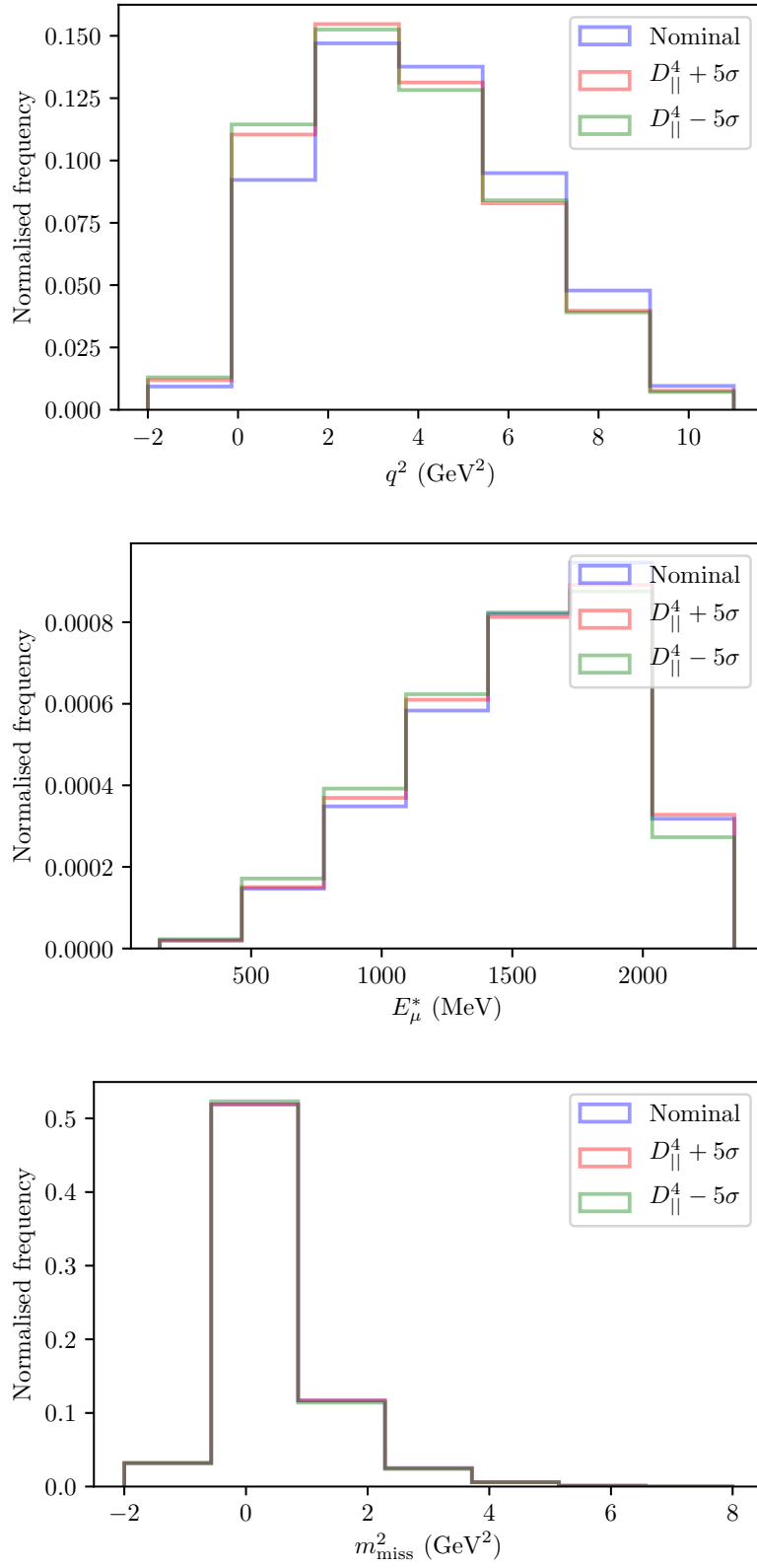


Figure 11.4: Comparison between the nominal template projections for $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ and those with alternative weights based on variations of the pQCD model with $D_{||}^4 \pm 5\sigma$.

11.6 Simulation statistics

The uncertainty due to the simulation statistics is assessed using the so-called Beeston-Barlow Lite (BBL) method [108]. This is used to propagate the statistical error due to the limited statistics of the fit templates to the overall uncertainty on $R(p\bar{p})$. The quadrature difference

$$\sigma_{\text{Sim. Stat}} = \sqrt{\sigma_{\text{BB On}}^2 - \sigma_{\text{BB Off}}^2}$$

between the uncertainty with the BBL method on ($\sigma_{\text{BB On}}$) and the uncertainty with the BBL method off ($\sigma_{\text{BB Off}}$) gives the contribution to the overall uncertainty. This proves to be the dominant source of systematic uncertainty for the analysis. This is due to insufficient simulation statistics for some of the dominant part-reco background modes.

11.7 Misidentified proton shape

There is an uncertainty related to the shape of the p mis-ID components in the fit that has two main origins. Firstly, as described in Sec. 6.2, the efficiency $\epsilon(U \rightarrow X)$ for a ghost track to be misidentified as a hadron X is nominally assumed to be negligible, i.e. $\epsilon(U \rightarrow X) = 0$. A systematic uncertainty is assessed where the efficiency $\epsilon(U \rightarrow X)$ is varied in the range $(0, \epsilon(\pi \rightarrow \pi))$, for tracks in the π -like category. The efficiency $\epsilon(\pi \rightarrow \pi)$ is replaced by the equivalent efficiency in the K -like and p -like categories.

Secondly, there is an uncertainty related to the limited statistics of the PID calibration samples. The p mis-ID shapes are consequently varied by re-sampling the calibration histograms based on this statistical uncertainty.

11.8 Misidentified muon shape

The origin of the uncertainty of the shape of the μ mis-ID component is directly analogous to the uncertainty on the p mis-ID shape; it derives from both the ghost track efficiencies $\epsilon(U \rightarrow X)$ and the finite statistics of the control samples. An analogous procedure is used to derive shape variations for the μ mis-ID component in evaluating the systematic uncertainty.

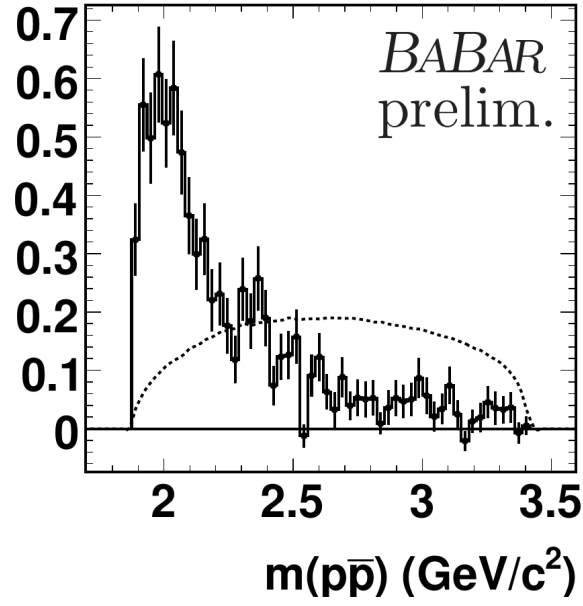


Figure 11.5: The $m(p\bar{p})$ spectrum for $B^0 \rightarrow p\bar{p}D^0$ decays from the observation made by BaBar. Figure taken from Ref. [109].

11.9 Part-reco background shape

A systematic uncertainty is associated with the reweighting in $m(p\bar{p})$ of some of the part-reco backgrounds outlined in Sec. 5.7. Nominally, the $m(p\bar{p})$ spectrum of the $B \rightarrow p\bar{p}DX$ samples is reweighted to that of the pQCD model. An alternative reweighting strategy is used in order to assess the systematic uncertainty. The $m(p\bar{p})$ spectrum shown in Fig. 11.5 is from the BaBar observation of $B^0 \rightarrow p\bar{p}D^0$ [109]. An alternative $m(p\bar{p})$ reweighting is derived from this by generating an approximation of the distribution based on the data. A comparison with the nominal $m(p\bar{p})$ distribution is shown in Fig. 11.6, and a comparison that demonstrates this variation compared to the nominal template shapes is shown in Fig. 11.7. The mean difference in the $R(p\bar{p})$ values from a set of pseudo-experiments based on the nominal and alternative models is taken as the systematic uncertainty.

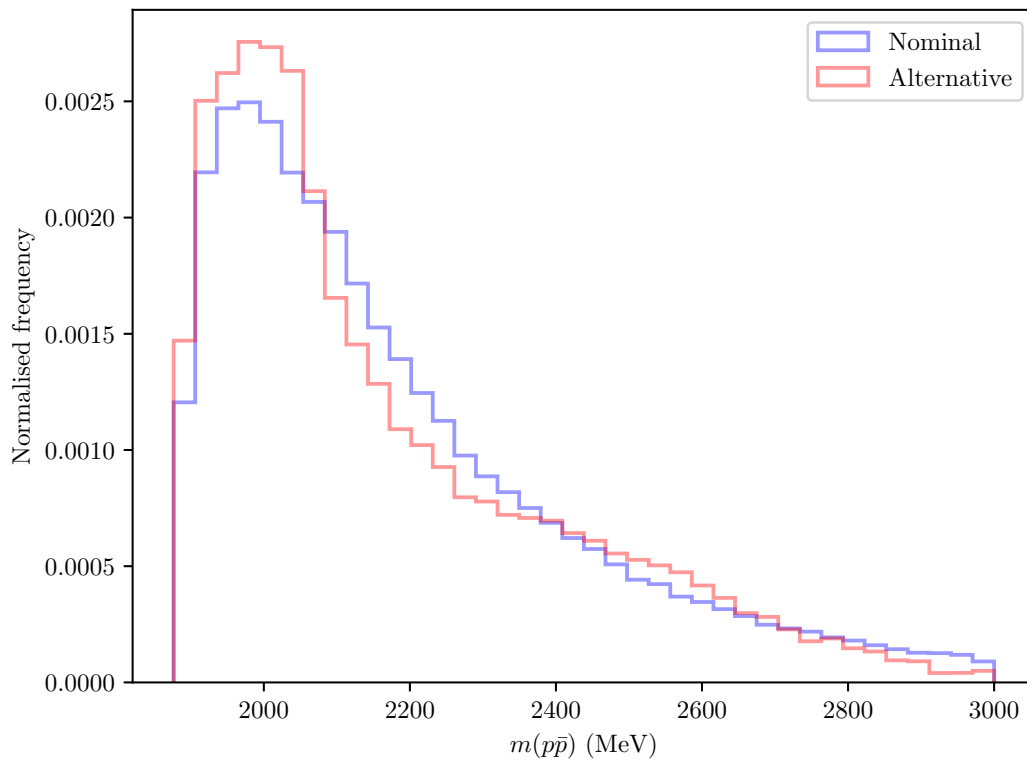


Figure 11.6: Comparison between the (blue) nominal and (red) alternative $m(p\bar{p})$ distributions used for the reweighting of the part-reco backgrounds, derived from the pQCD model and the BaBar data, respectively.

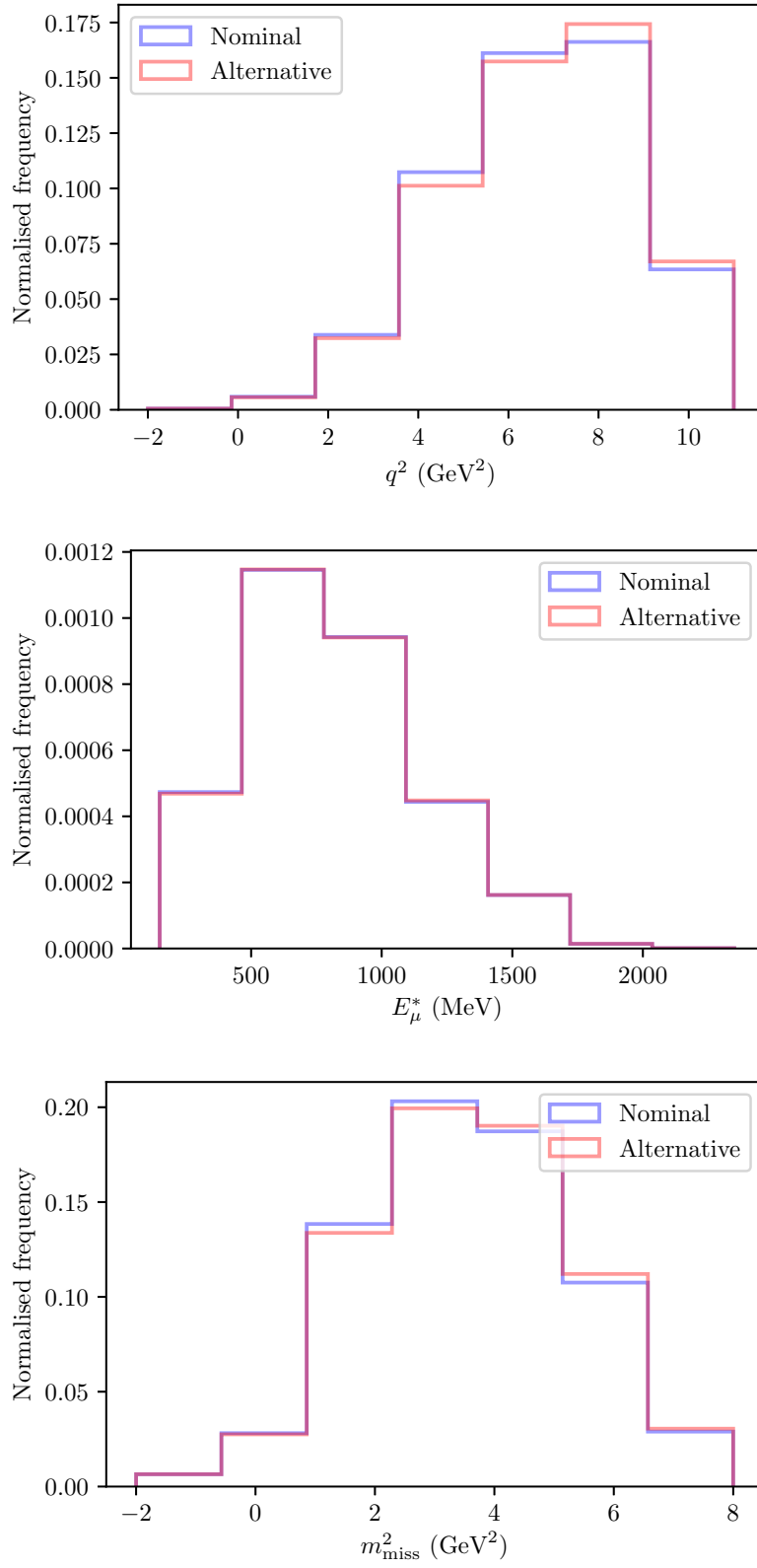


Figure 11.7: Comparison between the (blue) nominal and (red) alternative template projections for the $B^{+/0} \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X$ component.

11.10 Combinatorial background shape

A variation of the nominal event mixing method described in Sec. 6.4 is used to derive a systematic uncertainty due to the shape of the combinatorial background template. Rather than combining $\bar{p}\mu + p$ from separate events as in the nominal case, an alternative template is derived by forming combinations of $\bar{p} + \mu + p$, where each track is from a separate event. A comparison between these two templates is shown in Fig. 11.8.

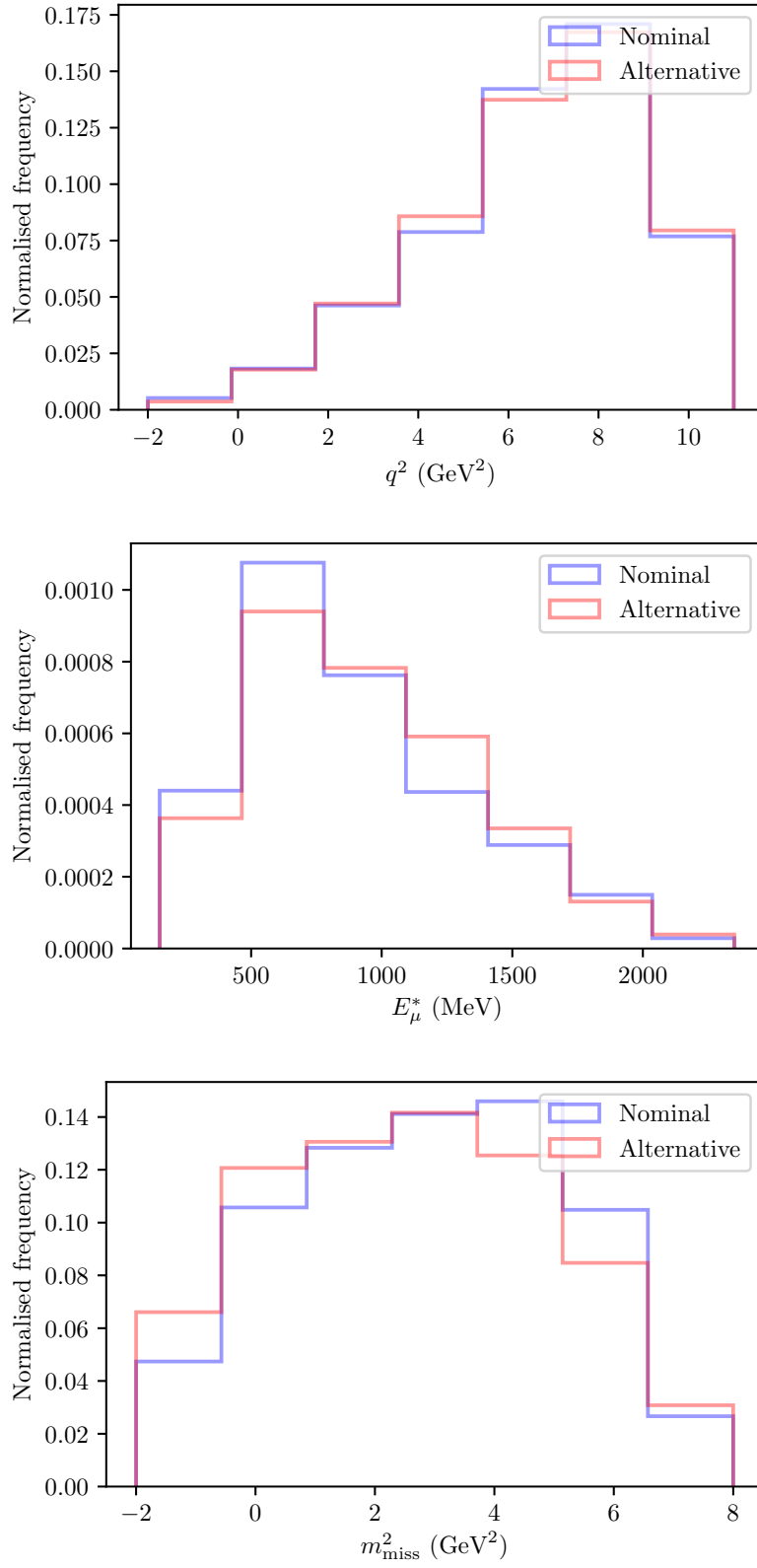


Figure 11.8: Comparison between the (blue) nominal and (red) alternative combinatorial template projections in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 , derived from the event mixing method.

Chapter 12

Results

This chapter details the main results of the analysis. A summary of the yields from the fit to data is presented, before concluding with a discussion of the limit placed on $R(p\bar{p})$.

12.1 Fit yields

The yields for each of the components obtained from the fit to data are given in Table 12.1. The statistical uncertainties are significant for the majority of the parameters in the fit. This is due to the fact that all-but-one of the parameters are left free to float. The uncertainty on the constrained parameter $\mathcal{N}(p \text{ mis-ID2})$ is notably lower.

It is clear that the majority of the candidates in the data sample are due to background processes, with the part-reco backgrounds being the dominant contribution, as expected. There is also a significant amount of p mis-ID background, particularly for p mis-ID 1. There are decays with high branching fractions that can contribute to the p mis-ID 1 component when partially-reconstructed, such as the semileptonic decay $\bar{\Lambda}_b^0 \rightarrow (\bar{\Lambda}_c^- \rightarrow \bar{p}K^+\pi^-)\mu^+\nu_\mu X$, where $\mathcal{B}(\bar{\Lambda}_b^0 \rightarrow \bar{\Lambda}_c^- \mu^+ \nu_\mu X) \approx 10\%$ and $\mathcal{B}(\bar{\Lambda}_c^- \rightarrow \bar{p}K^+\pi^-) \approx 6\%$ [88]. The contribution of this component alone could be significant due to both the abundant production of Λ_b^0 baryons at LHCb [110] and the presence of only one additional pion track.

Parameter	Yield
$R(p\bar{p})$	-1.73 ± 1.23
$\mathcal{N}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)$	7260 ± 300
$\mathcal{N}(B^+ \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X)$	843 ± 599
$\mathcal{N}(B^0 \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X)$	7790 ± 610
$\mathcal{N}(B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu)$	1740 ± 390
$\mathcal{N}(B^0 \rightarrow (\bar{\Lambda}_c^- \rightarrow \mu^- pX)\bar{p}X)$	301 ± 110
$\mathcal{N}(B^+ \rightarrow p(\Delta \rightarrow \bar{p}X)X\mu^+\nu_\mu)$	969 ± 628
$\mathcal{N}(B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu)$	960 ± 582
$\mathcal{N}(B^+ \rightarrow p\bar{p}D^0\mu^+\nu_\mu X)$	443 ± 214
$\mathcal{N}(B_s^0 \rightarrow (D_s^+ \rightarrow \mu^+ X)p\bar{p}X)$	779 ± 416
$\mathcal{N}(B_s^0 \rightarrow (D^0 \rightarrow K\mu^+ X)p\bar{p}X)$	328 ± 355
$\mathcal{N}(p \text{ mis-ID1})$	3140 ± 270
$\mathcal{N}(p \text{ mis-ID2})$	787 ± 31
$\mathcal{N}(\mu \text{ mis-ID})$	1360 ± 180
$\mathcal{N}(\text{Comb.})$	1040 ± 340

Table 12.1: A summary of the yields and their associated statistical uncertainties obtained from the fit to data that is described in Sec. 10.

12.2 Limit setting

As no significant signal is observed, a limit is set on $R(p\bar{p})$ using the CLs method [111]. This facilitates the computation of an upper limit by deriving a p -value based on the distribution of a test-statistic derived from pseudo-experiments. The pseudo-experiments use the model from the converged fit to data. The CLs p -values are shown in Fig. 12.1. Systematic uncertainties on the efficiency ratio are accounted for by placing a Gaussian constraint on the efficiency ratio, with a width given by the total systematic uncertainty. The uncertainties associated with the template shapes are incorporated by performing the pseudo-experiments using fluctuated versions of the nominal distribution based on the shape uncertainties. Table. 12.2 summarises the observed and expected 90% and 95% CLs exclusion limits.

The observed limit is lower than the median of the expected sensitivity. This is expected in the presence of a downward fluctuation of the signal [112]. This is best understood when considering the distributions of the test statistic. Pseudo-experiments are performed using a fixed value of $R(p\bar{p})$ (S+B), alongside background-only (B-only) hypotheses using the model from the fit to data. An example of the distributions is shown in Fig. 12.2 for $R(p\bar{p}) = 1$. The test statistic $q_{\text{obs.}}$ is also calculated for the observed data. The observed CLs p -value is calculated as the ratio of the integrals of the S+B and B-only distributions from $q_{\text{obs.}}$ to the end of the upper tail. Given that the data is consistent with a downward fluctuation of the signal or, equivalently, an upward fluctuation of the background, then $q_{\text{obs.}}$ is bound to sit somewhere in the upper tail of the B-only distribution. This results in smaller CLs p -values and therefore a tighter limit than the expected median.

Confidence level	Expected limit	Observed limit
90%	2.06	1.17
95%	2.49	1.46

Table 12.2: A summary of the limits on $R(p\bar{p})$ obtained from the CLs method.

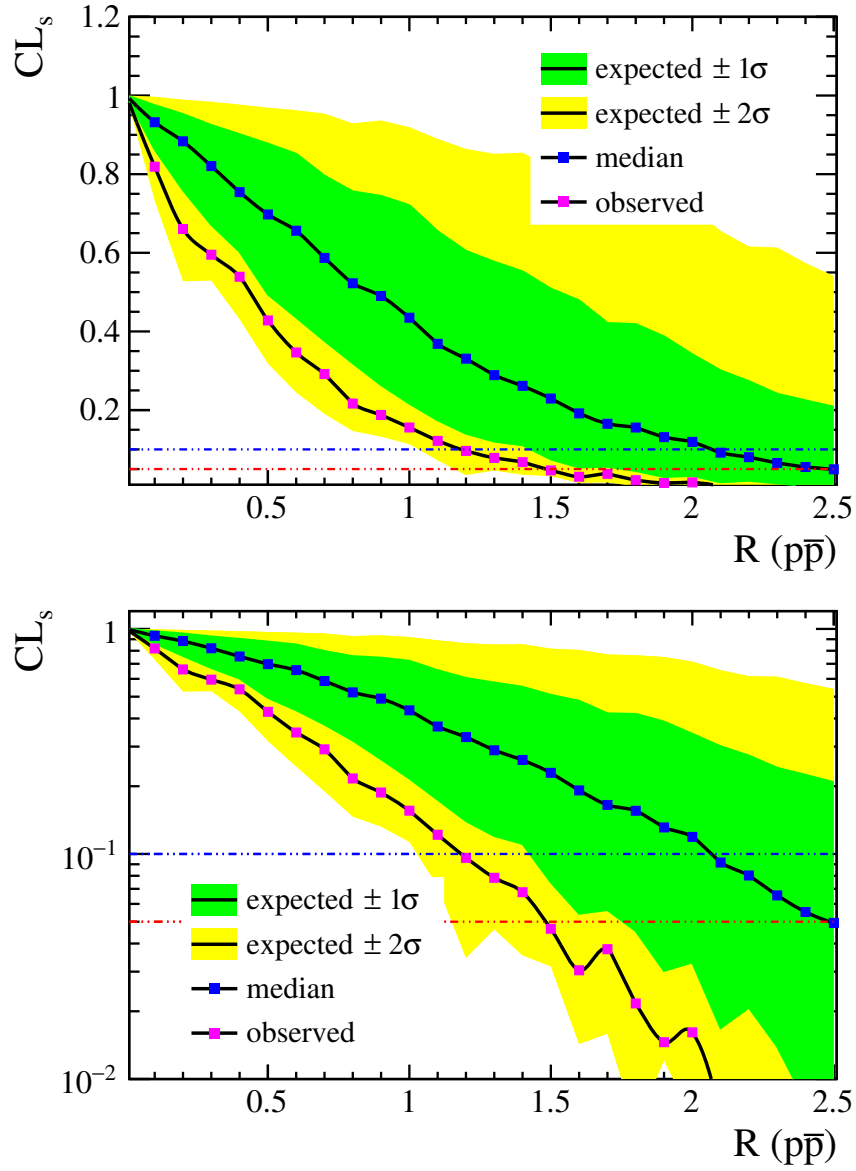


Figure 12.1: Expected and observed (blue horizontal line) 90% and (red horizontal line) 95% CL_s exclusion limits accounting for all systematic uncertainties, with (top) non-logarithmic and (bottom) logarithmic vertical axes.

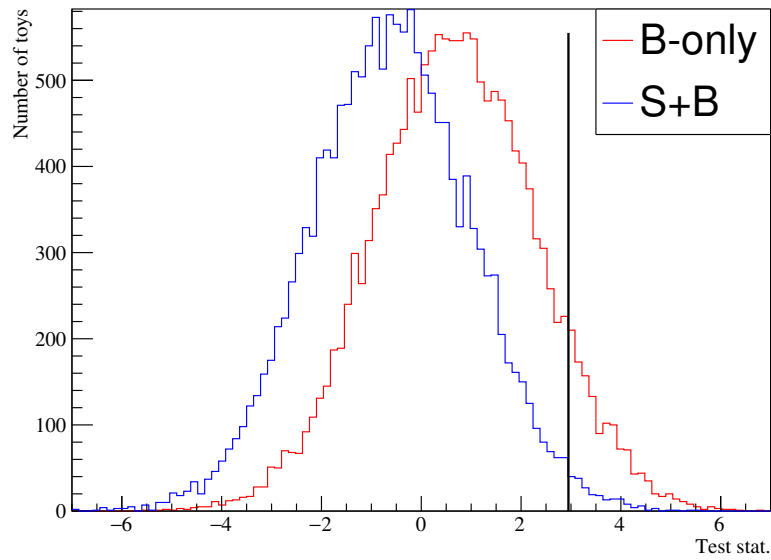


Figure 12.2: (Blue) Signal+background (S+B) and (red) background-only (B-only) test statistic distributions as obtained from the pseudo-experiments for the CLs method. The CLs p -value is given by the ratio of the integral of the two distributions from the observed test statistic value (black line) up to the end of the upper tails.

Chapter 13

Conclusion

The work in this thesis constitutes the first search for the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ decay and limit placed on the branching fraction ratio $R(p\bar{p})$. A measurement of the $B \rightarrow p\bar{p}$ hadronic form factors in $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays is also presented. The analysis was performed using 5.1 fb^{-1} of data collected by the LHCb collaboration from 2016 to 2018. No significant signal is observed and the data are consistent with a negative downward-fluctuation of the signal or an upward fluctuation of the background. A limit of

$$R(p\bar{p}) < 1.46$$

is placed at the 95% confidence level, corresponding to an upper limit of

$$\mathcal{B}(B^+ \rightarrow p\bar{p}\tau^+\nu_\tau) < 7.69 \times 10^{-6}$$

on the $B^+ \rightarrow p\bar{p}\tau^+\nu_\tau$ branching fraction using the LHCb measurement of $\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)$ in Ref. [39]. This is consistent with measurements of other analogous ratios, such as $R(D^{(*)})$ [12–14], where the branching fraction ratios are less than 1. This would also be the naïve expectation based on phase space considerations alone, given the significantly higher mass of the τ lepton over the μ .

In terms of the scope for future improvements, the measurement would further benefit from a thorough investigation into the part-reco background processes. Constraining the yields of these components would benefit the statistical precision, and this could be achieved with a control sample enriched in these decays. This sample can be obtained by using the charged isolation tool to obtain a sample rich in candidates that contain additional charged tracks (in contrast to the nominal use-case, where the aim is to reduce the contamination). Deducing the composition of such a

sample in terms of the contributing part-reco processes can provide information about the levels of these backgrounds in the nominal sample, particularly in the cases where the branching fraction is not known.

It would also be important to reduce the systematic uncertainties where possible. The dominant source could be reduced with the production of larger simulation samples, at the cost of an increased allocation of computing resources by up to as much as an order of magnitude. Further work on the theoretical side would also improve the understanding of $B^+ \rightarrow p\bar{p}\ell^+\nu_\ell$ decays. An updated measurement of $\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)$ and the $B \rightarrow p\bar{p}$ form factors using the full LHCb data set would serve as crucial input to theorists working in this areas.

As the LHCb experiment progresses to another period of data taking over the years 2022–2024 [113] and the Belle experiment continues to gather their 2021–2025 data set, the need to explore the structure of any new phenomena underpinning the b -anomalies is ever-growing. Measurements like those of ratios such as $R(p\bar{p})$ will play a crucial role in this endeavour. Regardless of the resolution of the anomalous results, the years to come are certain to be pivotal in the history of flavour physics.

Bibliography

- [1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579.
- [2] S. Weinberg, *A model of leptons*, Physical Review Letters **19** (1967) 1264.
- [3] A. Salam, *Weak and electromagnetic interactions*, Conf. Proc. **C680519** (1968) 367.
- [4] M. Veltman *et al.*, *Regularization and renormalization of gauge fields*, Nuclear Physics B **44** (1972) 189.
- [5] D. Clowe *et al.*, *A direct empirical proof of the existence of dark matter*, The Astrophysical Journal Letters **648** (2006) L109.
- [6] P. J. E. Peebles and B. Ratra, *The cosmological constant and dark energy*, Reviews of modern physics **75** (2003) 559.
- [7] ATLAS Collaboration, *Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc*, Physics Letters B **716** (2012) 1.
- [8] CMS Collaboration, *Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc*, Physics Letters B **716** (2012) 30.
- [9] LHCb Collaboration, *Letter of intent: A dedicated lhc collider beauty experiment for precision measurements of cp-violation*, CERN/LHCC (1995) 95.
- [10] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays*, Phys. Rev. Lett. **114** (2015) 041801, [arXiv:1411.3104](#).
- [11] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in $B_s^0 - \bar{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, [arXiv:1605.09768](#).

- [12] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$* , Phys. Rev. Lett. **115** (2015) 111803, Publisher's Note *ibid.* **115** (2015) 159901, [arXiv:1506.08614](#).
- [13] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the $\mathcal{B}(B^0 \rightarrow D^{*-}\tau^+\nu_\tau)$ and $\mathcal{B}(B^0 \rightarrow D^{*-}\mu^+\nu_\mu)$ branching fractions using three-prong τ -lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, [arXiv:1708.08856](#).
- [14] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavor universality by the measurement of the $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays*, Phys. Rev. **D97** (2018) 072013, [arXiv:1711.02505](#).
- [15] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ decays*, JHEP **05** (2020) 040, [arXiv:1912.08139](#).
- [16] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \rightarrow K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **122** (2019) 191801, [arXiv:1903.09252](#).
- [17] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \rightarrow K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, [arXiv:1705.05802](#).
- [18] BABAR Collaboration, B. Aubert *et al.*, *Measurement of branching fractions and rate asymmetries in the rare decays $B \rightarrow K^{(*)}\ell^+\ell^-$* , Physical Review D **86** (2012) 032012.
- [19] BABAR Collaboration, B. Aubert *et al.*, *Measurements of the semileptonic decays $B \rightarrow D\ell\bar{\nu}$ and $B \rightarrow D^*\ell\bar{\nu}$ using a global fit to $D X\ell\bar{\nu}_\ell$ final states*, Physical Review D **79** (2009) 012002.
- [20] Belle Collaboration, A. Abdesselam *et al.*, *Measurement of the differential branching fraction and forward-backward asymmetry for $B \rightarrow K^{(*)}\ell^+\ell^-$* , Physical Review Letters **103** (2009) 171801.
- [21] Belle Collaboration, A. Abdesselam *et al.*, *Measurement of the branching ratio of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Physical Review D **92** (2015) 072014.
- [22] Belle Collaboration, A. Abdesselam *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$* , Phys. Rev. Lett. **118** (2017) 211801, [arXiv:1612.00529](#).

- [23] Belle Collaboration, A. Abdesselam *et al.*, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, [arXiv:1904.08794](#).
- [24] R. Alonso, B. Grinstein, and J. Martin Camalich, *Lepton universality violation and lepton flavor conservation in B -meson decays*, JHEP **10** (2015) 184, [arXiv:1505.05164](#).
- [25] A. Greljo, G. Isidori, and D. Marzocca, *On the breaking of Lepton Flavor Universality in B decays*, JHEP **07** (2015) 142, [arXiv:1506.01705](#).
- [26] R. Barbieri, G. Isidori, A. Pattori, and F. Senia, *Anomalies in B -decays and $U(2)$ flavour symmetry*, Eur. Phys. J. C **76** (2016) 67, [arXiv:1512.01560](#).
- [27] L. Calibbi, A. Crivellin, and T. Ota, *Effective Field Theory Approach to $b \rightarrow s\ell\ell'$, $B \rightarrow K^{(*)}\nu\bar{\nu}$ and $B \rightarrow D^{(*)}\tau\nu$ with Third Generation Couplings*, Phys. Rev. Lett. **115** (2015) 181801, [arXiv:1506.02661](#).
- [28] M. Bauer and M. Neubert, *Minimal Leptoquark Explanation for the $R_{D^{(*)}}$, R_K , and $(g-2)_g$ Anomalies*, Phys. Rev. Lett. **116** (2016) 141802, [arXiv:1511.01900](#).
- [29] S. Fajfer and N. Košnik, *Vector leptoquark resolution of R_K and $R_{D^{(*)}}$ puzzles*, Phys. Lett. B **755** (2016) 270, [arXiv:1511.06024](#).
- [30] MILC, J. A. Bailey *et al.*, *$B \rightarrow D\ell\nu$ form factors at nonzero recoil and $|V_{cb}|$ from 2+1-flavor lattice QCD*, Phys. Rev. D **92** (2015) 034506, [arXiv:1503.07237](#).
- [31] HPQCD, H. Na *et al.*, *$B \rightarrow D\ell\nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$* , Phys. Rev. D **92** (2015) 054510, [arXiv:1505.03925](#), [Erratum: Phys.Rev.D 93, 119906 (2016)].
- [32] D. Bigi and P. Gambino, *Revisiting $B \rightarrow D\ell\nu$* , Phys. Rev. D **94** (2016) 094008, [arXiv:1606.08030](#).
- [33] 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 **95** (2017) 115008, [arXiv:1703.05330](#), [Erratum: Phys.Rev.D 97, 059902 (2018)].

- [34] S. Jaiswal, S. Nandi, and S. K. Patra, *Extraction of $|V_{cb}|$ from $B \rightarrow D^{(*)}\ell\nu_\ell$ and the Standard Model predictions of $R(D^{(*)})$* , JHEP **12** (2017) 060, [arXiv:1707.09977](#).
- [35] D. Bigi, P. Gambino, and S. Schacht, *$R(D^*)$, $|V_{cb}|$, and the Heavy Quark Symmetry relations between form factors*, JHEP **11** (2017) 061, [arXiv:1707.09509](#).
- [36] Belle Collaboration, A. Abdesselam *et al.*, *Precise determination of the CKM matrix element $|V_{cb}|$ with $\bar{B}^0 \rightarrow D^{*+}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, [arXiv:1702.01521](#).
- [37] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b -hadron, c -hadron, and τ -lepton properties as of 2018*, [arXiv:1909.12524](#), updated results and plots available at <https://hflav.web.cern.ch>.
- [38] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \rightarrow J/\psi \tau^+ \nu_\tau)/\mathcal{B}(B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu)$* , Phys. Rev. Lett. **120** (2018) 121801, [arXiv:1711.05623](#).
- [39] LHCb collaboration, R. Aaij *et al.*, *Observation of the semileptonic decay $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$* , JHEP **03** (2020) 146, [arXiv:1911.08187](#).
- [40] C. Q. Geng and Y. K. Hsiao, *Semileptonic $B^- \rightarrow p\bar{p}\ell^-\bar{\nu}_\ell$ decays*, Phys. Lett. **B704** (2011) 495–498, [arXiv:1107.0801](#).
- [41] BES Collaboration, J. Z. Bai *et al.*, *Observation of a near-threshold enhancement in the $p\bar{p}$ mass spectrum from radiative $J/\psi \rightarrow \gamma p\bar{p}$ decays*, Phys. Rev. Lett. **91** (2003) 022001, [arXiv:hep-ex/0303006](#).
- [42] J. Haidenbauer, U. G. Meissner, and A. Sibirtsev, *Near threshold $p\bar{p}$ enhancement in the $J/\psi \rightarrow \omega p\bar{p}$ decay*, Phys. Lett. **B666** (2008) 352–354, [arXiv:0804.1469](#).
- [43] LHCb collaboration, R. Aaij *et al.*, *Evidence for CP violation in $B^+ \rightarrow p\bar{p}K^+$ decays*, Phys. Rev. Lett. **113** (2014) 141801, [arXiv:1407.5907](#).
- [44] Belle Collaboration, M. Z. Wang *et al.*, *Observation of $B^+ \rightarrow p\bar{p}\pi^+$, $B^0 \rightarrow p\bar{p}K^0$, and $B^+ \rightarrow p\bar{p}K^{*+}$* , Phys. Rev. Lett. **92** (2004) 131801, [arXiv:hep-ex/0310018](#).
- [45] M. D. Schwartz, *Quantum field theory and the standard model*, Cambridge University Press, 2014.

- [46] M. Thomson, *Modern particle physics*, Cambridge University Press, 2013.
- [47] M. E. Peskin, *An introduction to quantum field theory*, CRC press, 2018.
- [48] T. Humair, *Testing lepton universality in penguin decays of beauty mesons using the LHCb detector*, 2019. Presented 02 May 2019.
- [49] J. Goldstone, *Field theories with superconductor solutions*, Il Nuovo Cimento (1955-1965) **19** (1961) 154.
- [50] Y. Nambu, *Quasi-particles and gauge invariance in the theory of superconductivity*, Physical Review **117** (1960) 648.
- [51] T. W. B. Kibble, *Spontaneous symmetry breaking in gauge theories*, Phil. Trans. Roy. Soc. Lond. A **373** (2014) 20140033.
- [52] P. W. Anderson, *Plasmons, gauge invariance, and mass*, Phys. Rev. **130** (1963) 439.
- [53] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [54] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.
- [55] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [56] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.
- [57] M. Kobayashi and T. Maskawa, *ν -violation in the renormalizable theory of weak interaction*, Progress of Theoretical Physics **49** (1973) 652.
- [58] N. Cabibbo, *Unitary symmetry and leptonic decays*, Physical Review Letters **10** (1963) 531.
- [59] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.
- [60] G. P. Lepage and S. J. Brodsky, *Exclusive processes in quantum chromodynamics: The form factors of baryons at large momentum transfer*, Physical Review Letters **43** (1979) 545.

- [61] G. P. Lepage and S. J. Brodsky, *Exclusive processes in perturbative quantum chromodynamics*, Physical Review D **22** (1980) 2157.
- [62] Belle Collaboration, A. Abdesselam *et al.*, *Observation of $\bar{B}^0 \rightarrow D^{(*)} p \bar{p}$* , Physical Review Letters **89** (2002) 151802.
- [63] CLEO Collaboration, Q. He *et al.*, *First observation of the decays $B^0 \rightarrow D^{*-} p \bar{p} \pi^+$ and $B^0 \rightarrow D^{*-} p n$* , Physical Review Letters **86** (2001) 2732.
- [64] Belle Collaboration, A. Abdesselam *et al.*, *Observation of the decay $\bar{B}^0 \rightarrow D_s^{*+} \Lambda \bar{p}$* , Physical Review D **76** (2007) 051102.
- [65] Belle Collaboration, A. Abdesselam *et al.*, *Observation of $B^- \rightarrow J/\psi \Lambda \bar{p}$ and searches for $B^- \rightarrow J/\psi \sigma^0 \bar{p}$ and $B^0 \rightarrow J/\psi p \bar{p}$ decays*, Physical Review D **72** (2005) 051105.
- [66] BaBar Collaboration, B. Aubert *et al.*, *Evidence for $B^+ \rightarrow J/\psi p \Lambda^-$ and search for $B^0 \rightarrow J/\psi p \bar{p}$* , Physical Review Letters **90** (2003) 231801.
- [67] C.-H. Chen, H.-Y. Cheng, C. Geng, and Y. Hsiao, *Charmful three-body baryonic b decays*, Physical Review D **78** (2008) 054016.
- [68] C. Geng and Y. Hsiao, *Angular distributions in three-body baryonic b decays*, Physical Review D **74** (2006) 094023.
- [69] ATLAS Collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. 437 p, Also published by CERN Geneva in 2010.
- [70] CMS Collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment*, JINST **3** (2008) S08004. 361 p, Also published by CERN Geneva in 2010.
- [71] ALICE Collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC. A Large Ion Collider Experiment*, JINST **3** (2008) S08002. 259 p, Also published by CERN Geneva in 2010.
- [72] LHCb Collaboration, A. Alves *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005, Also published by CERN Geneva in 2010.
- [73] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots*, .

- [74] LHCb collaboration, *LHCb VELO (VERtex LOcator): Technical Design Report*, CERN-LHCC-2001-011, 2001.
- [75] LHCb collaboration, R. Aaij *et al.*, *Measurement of the effective $B_s^0 \rightarrow K^+ K^-$ lifetime*, Phys. Lett. **B707** (2012) 349, [arXiv:1111.0521](#).
- [76] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037, 2000.
- [77] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](#).
- [78] LHCb Collaboration, C. Marin Benito, *PID strategy and performance at LHCb in Run 2*, PoS **ICHEP2018** (2019) 687. 4 p.
- [79] LHCb, D. Derkach *et al.*, *Machine-Learning-based global particle-identification algorithms at the LHCb experiment*, J. Phys. : Conf. Ser. **1085** (2018) 042038. 5 p.
- [80] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, 2002.
- [81] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, 2001.
- [82] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Journal of Basic Engineering **82** (1960) 35.
- [83] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, [arXiv:1408.1251](#).
- [84] M. Needham, *Performance of the LHCb Track Reconstruction Software*, , CERN, Geneva, 2008.
- [85] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, 2000.
- [86] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010, 2001.

- [87] LHCb Collaboration, R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC. Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2018) P04013. 43 p, [arXiv:1812.10790](#), 46 pages, 35 figures, 1 table. All figures and tables are available at <https://cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-DP-2019-001.html>.
- [88] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. D **98** (2018) 030001, and 2019 update.
- [89] K. Abe *et al.*, *Measurement of R_b using a vertex mass tag*, Physical review letters **80** (1998) 660.
- [90] LHCb collaboration, R. Aaij *et al.*, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, Nature Physics **11** (2015) 743, [arXiv:1504.01568](#).
- [91] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, [arXiv:0710.3820](#).
- [92] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.
- [93] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s} = 7$ and 13 TeV*, JHEP **12** (2017) 026, [arXiv:1710.04921](#).
- [94] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [95] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [96] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](#).
- [97] A. Rogozhnikov, *Reweighting with boosted decision trees*, in *Journal of Physics: Conference Series*, **762**, 012036, IOP Publishing, 2016.
- [98] BaBar Collaboration, B. Aubert *et al.*, *Measurements of the decays $B^0 \rightarrow \bar{D}^0 p \bar{p}$, $B^0 \rightarrow \bar{D}^{*0} p \bar{p}$, $B^0 \rightarrow D^- p \bar{p} \pi^+$, and $B^0 \rightarrow D^{*-} p \bar{p} \pi^+$* , Physical Review D **74** (2006) .

- [99] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, [arXiv:1301.5286](#).
- [100] G. D’Agostini, *Improved iterative bayesian unfolding*, arXiv preprint [arXiv:1010.0632](#) (2010).
- [101] T. Adye, *Unfolding algorithms and tests using RooUnfold*, [arXiv:1105.1160](#).
- [102] P. A. Lachenbruch and M. R. Mickey, *Estimation of error rates in discriminant analysis*, Technometrics **10** (1968) 1.
- [103] T. Chen *et al.*, *Xgboost: extreme gradient boosting*, R package version 0. 4-2 **1** (2015) 1.
- [104] T. Leddig, *Search for the decay $B^- \rightarrow \Lambda_c^+ \bar{p} \ell^- \bar{\nu}_\ell$ with the BaBar detector*, 2015.
- [105] K. Cranmer *et al.*, *Histfactory: A tool for creating statistical models for use with roofit and roostats*, , 2012.
- [106] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, [arXiv:1210.6861](#).
- [107] F. James, M. Winkler *et al.*, *Minuit user’s guide*, CERN, Geneva **23** (2004).
- [108] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219.
- [109] BaBar Collaboration, B. Aubert *et al.*, *Observation and study of baryonic decays: $B \rightarrow D^{(*)} p \bar{p}$, $D^{(*)} p \bar{p} \pi$, and $D^{(*)} p \bar{p} \pi \pi$* , arXiv preprint [arXiv:0908.2202](#) (2009).
- [110] LHCb collaboration, R. Aaij *et al.*, *Study of the productions of Λ_b^0 and $\bar{B}^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \rightarrow J/\psi p K^-$ branching fraction*, Chin. Phys. C **40** (2016) 011001, [arXiv:1509.00292](#).
- [111] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. **G28** (2002) 2693.
- [112] LHCb collaboration, R. Aaij *et al.*, *Search for the rare decay $B^+ \rightarrow \mu^+ \mu^- \mu^+ \nu_\mu$* , Eur. Phys. J. **C79** (2018) 675, [arXiv:1812.06004](#).
- [113] I. Belyaev *et al.*, *The history of LHCb*, The European Physical Journal H **46** (2021) 1.

Appendix A

Stripping selection

Particle	Selection
p	$p > 15 \text{ GeV}$
	$p_{\text{T}} > 800 \text{ MeV}$
	$\chi^2_{IP} \text{ wrt PV} > 9$
	Track $\chi^2/DOF < 6$
	$DLL(p - \pi) > 2$
	$DLL(p - K) > 2$
	$P(\text{ghost}) < 0.35$
μ^+	$P(\text{ghost}) < 0.35$
	Track $\chi^2/DOF < 4$
	$p > 3000 \text{ MeV}$
	$\chi^2_{IP} \text{ wrt PV} > 16$
$p\bar{p}$	Vertex $\chi^2/DOF < 10$
	$DIRA > 0.999$
	$\chi^2 \text{ PV distance} > 25$
B^+	Vertex $\chi^2/DOF < 15$
	$DIRA > 0.999$
	$\chi^2 \text{ PV distance} > 50$
	$p_{\text{T}} > 1500 \text{ MeV}$
	$m(p\bar{p}\mu^+) > 1000 \text{ MeV}$

Table A.1: Summary of the criteria made in the stripping selection for the measurement of $R(p\bar{p})$.

Appendix B

Neutral isolation BDT input features

Neutral object	Feature	Definition
γ	photon_E	Energy
	photon_pt	p_T
	photon_bframe_pt	p_T relative to visible B^+ momentum
	ap_photon_bframe_pt	p_T relative to $\bar{p}$ momentum
	ap_photon_dist_sph	Distance in spherical coords. from $\bar{p}$, in a cone centred on the $\bar{p}$ momentum
	photon_dist_cart	Distance in Cartesian coords. from B^+
	photon_dist_cone	Distance in (η, ϕ) from B^+
π^0	pi0_M	Mass
	pi0_pt	p_T
	pi0_bframe_pt	p_T relative to visible B^+ momentum
	pi0_dist_cart	Distance in Cartesian coords. from B^+
	ap_pi0_dist_sph	Distance in spherical coords. from $\bar{p}$, in a cone centred on the $\bar{p}$ momentum
	ap_pi0_bframe_pt	p_T relative to $\bar{p}$ momentum

Table B.1: The features used as inputs to the neutral isolation BDTs. The term ‘relative’ here refers to directionality, and not relative magnitudes.

Appendix C

Reweighted templates for form factor fit

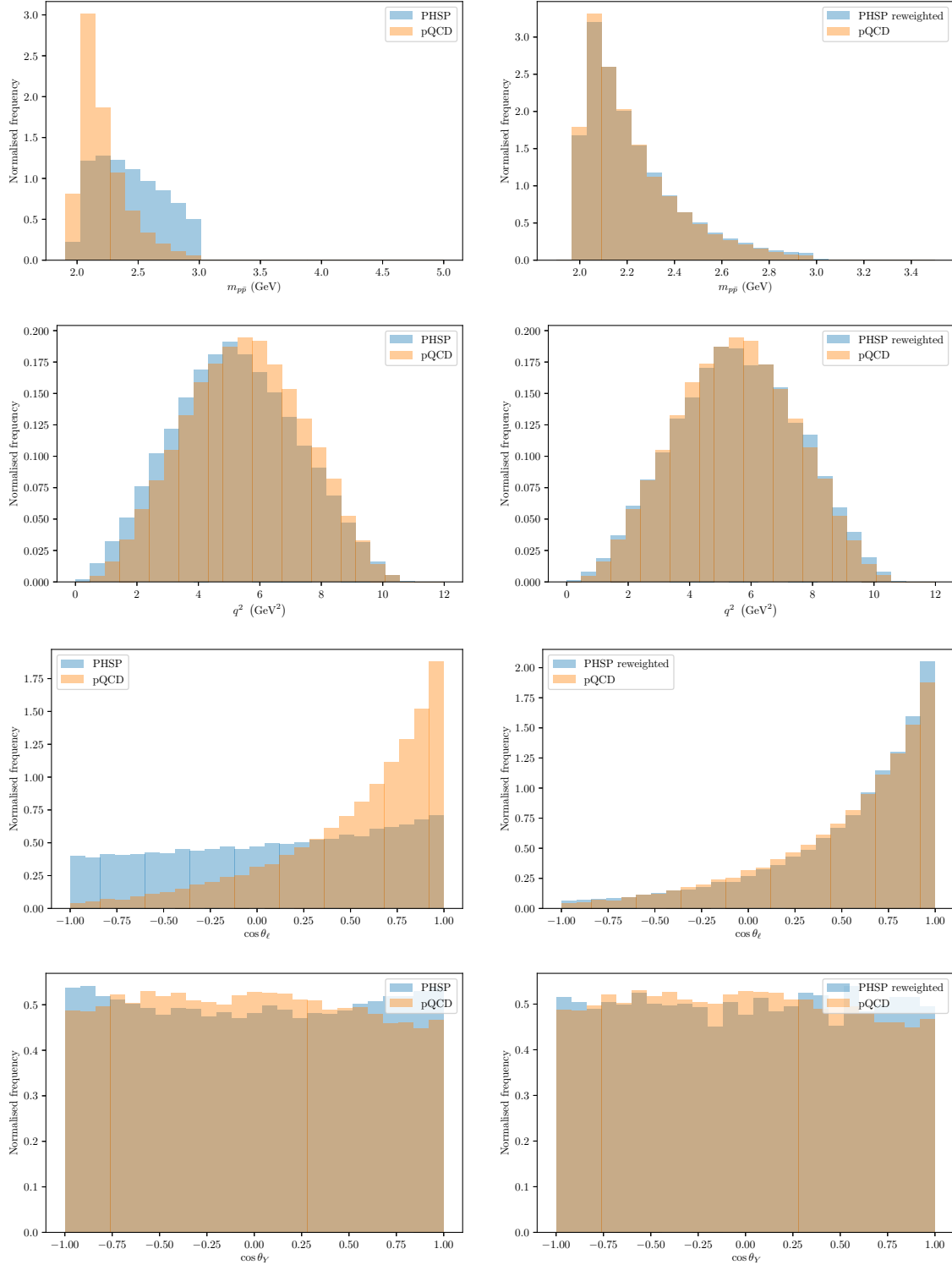


Figure C.1: Comparisons between PHSP before (left) and after (right) reweighting to the pQCD model in (descending order) $m(p\bar{p})$, q^2 , $\cos \theta_\ell$ and $\cos \theta_Y$ for $B^+ \rightarrow p \bar{\Delta} \mu^+ \nu_\mu$ decays.

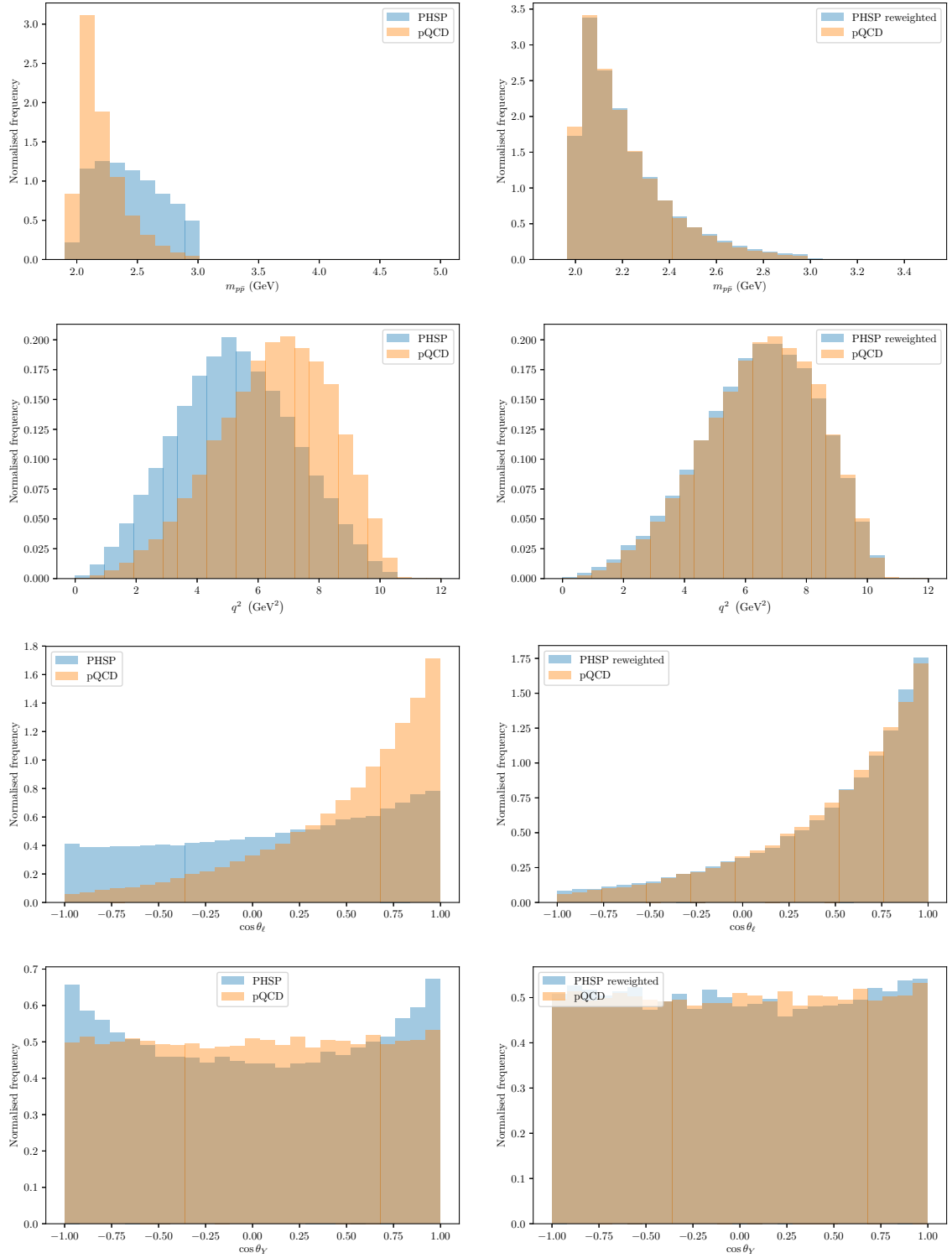


Figure C.2: Comparisons between PHSP before (left) and after (right) reweighting to the pQCD model in (descending order) $m(p\bar{p})$, q^2 , $\cos\theta_\ell$ and $\cos\theta_Y$ for $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$ decays.

Appendix D

Stripping selection for form factor fit

Particle	Selection
p	$p > 18 \text{ GeV}$
	$p_T > 1500 \text{ MeV}$
	$\chi^2_{IP} \text{ wrt PV} < 169$
	Track $\chi^2/DOF < 6$
	$DLL(p - \pi) > 10$
	$DLL(p - K) > 5$
	$P(\text{ghost}) < 0.35$
μ^+	$P(\text{ghost}) < 0.35$
	Track $\chi^2/DOF < 4$
	$p > 3000 \text{ MeV}$
	$p_T > 1500 \text{ MeV}$
	$\chi^2_{IP} \text{ wrt PV} > 16$
$p\bar{p}$	Vertex $\chi^2/DOF < 4$
	$DIRA > 0.999$
	$p_T > 1500 \text{ MeV}$
	$\chi^2 \text{ PV distance} > 150$
B^+	Vertex $\chi^2/DOF < 15$
	$DIRA > 0.9994$
	$\chi^2 \text{ PV distance} > 150$
	$p_T > 1500 \text{ MeV}$
	$m(p\bar{p}\mu^+) < 7000 \text{ MeV}$

Table D.1: Summary of the criteria made in the stripping selection for the measurement of the $B \rightarrow p\bar{p}$ transition form factors.

Appendix E

Templates for the fit to data

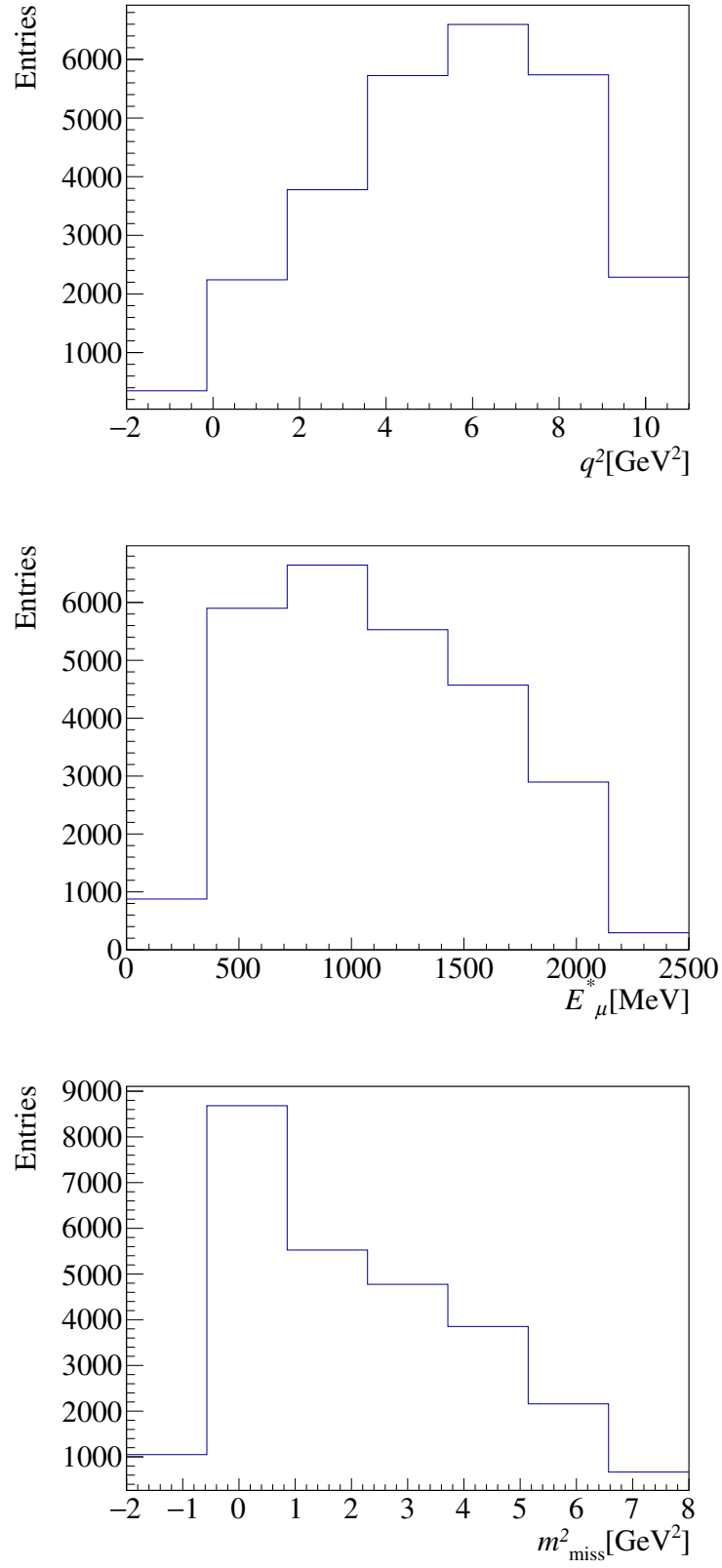


Figure E.1: Projections of the 3D histogram containing the selected data candidates in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

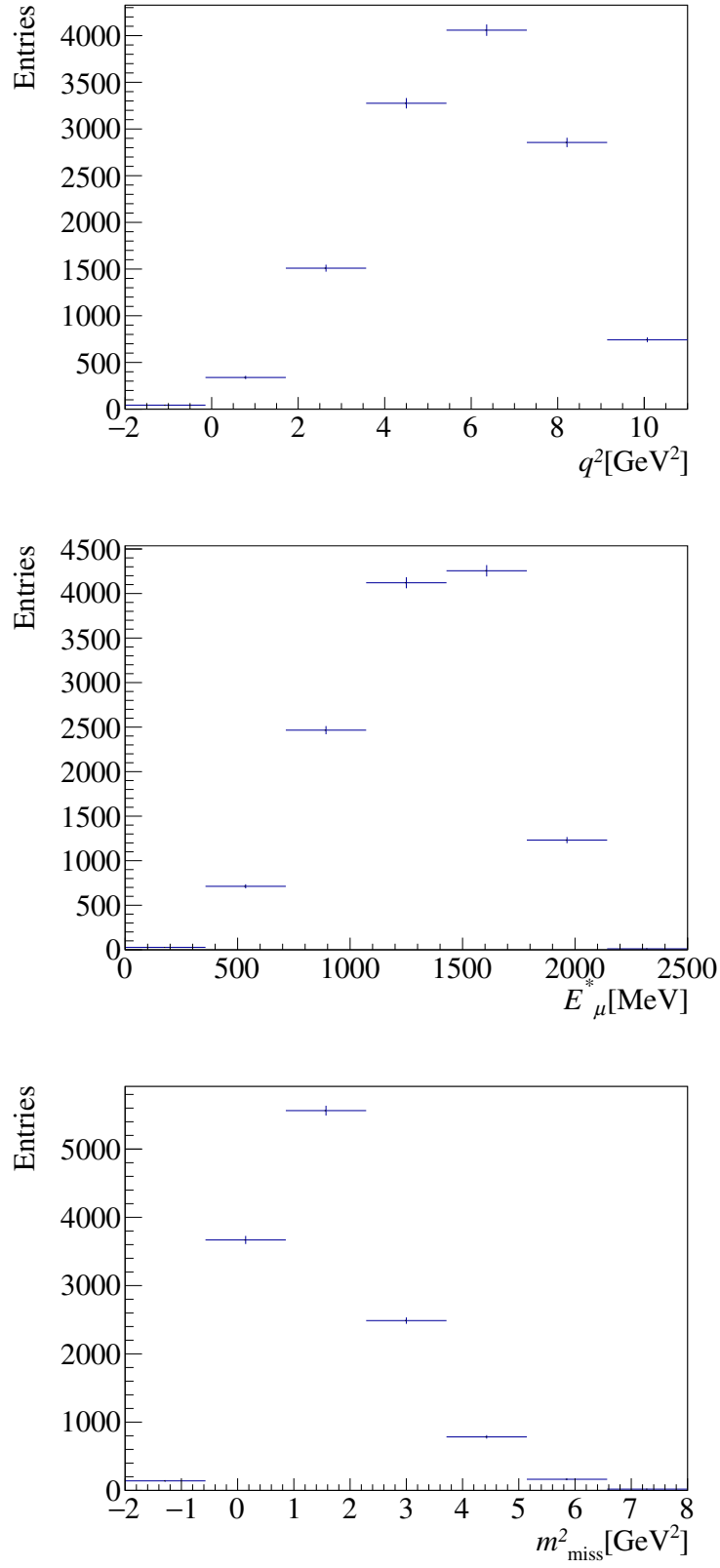


Figure E.2: Projections of the 3D histogram containing the selected $B^+ \rightarrow p(\bar{N}^* \rightarrow \bar{p}X)\mu^+\nu_\mu$ simulation sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

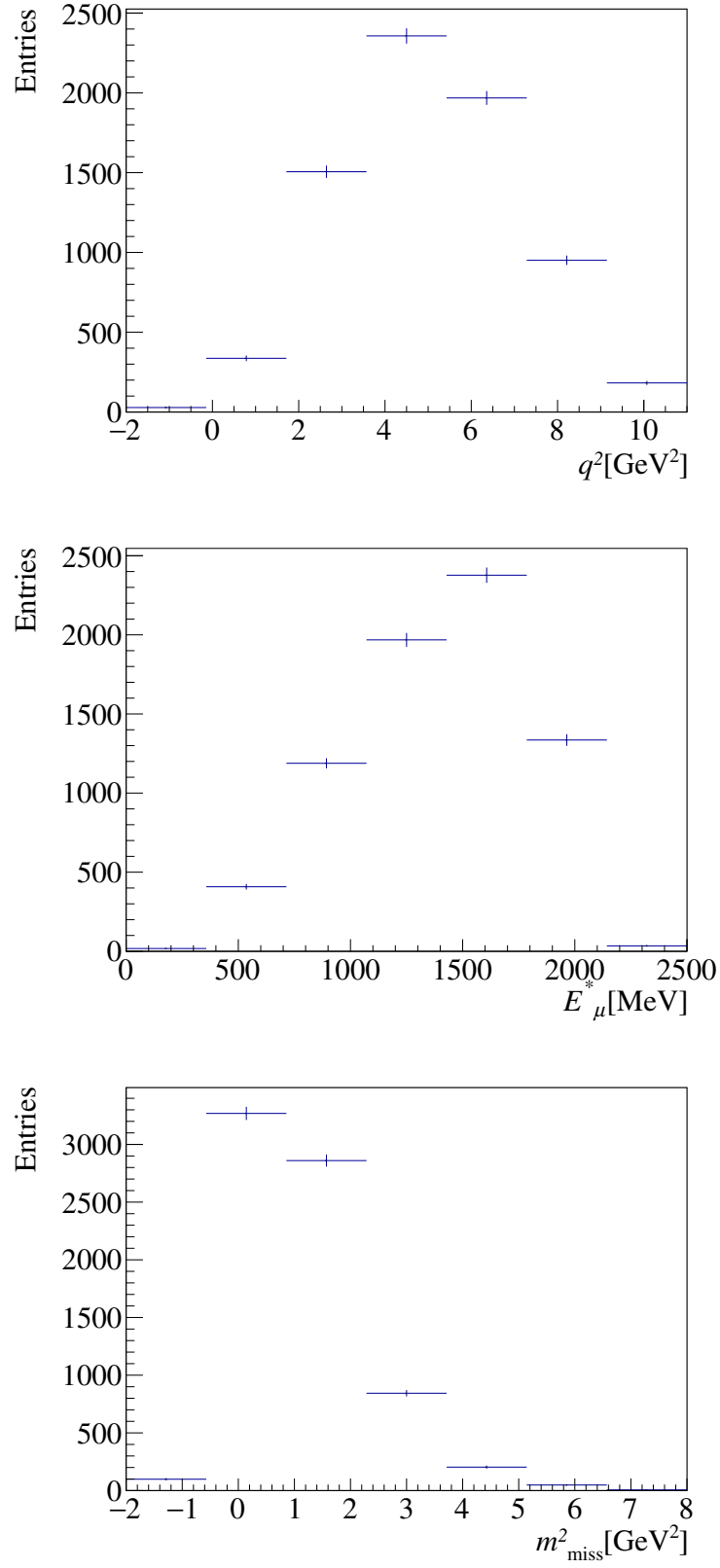


Figure E.3: Projections of the 3D histogram containing the selected $B^+ \rightarrow p(\Delta \rightarrow \bar{p}X)X\mu^+\nu_\mu$ simulation sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

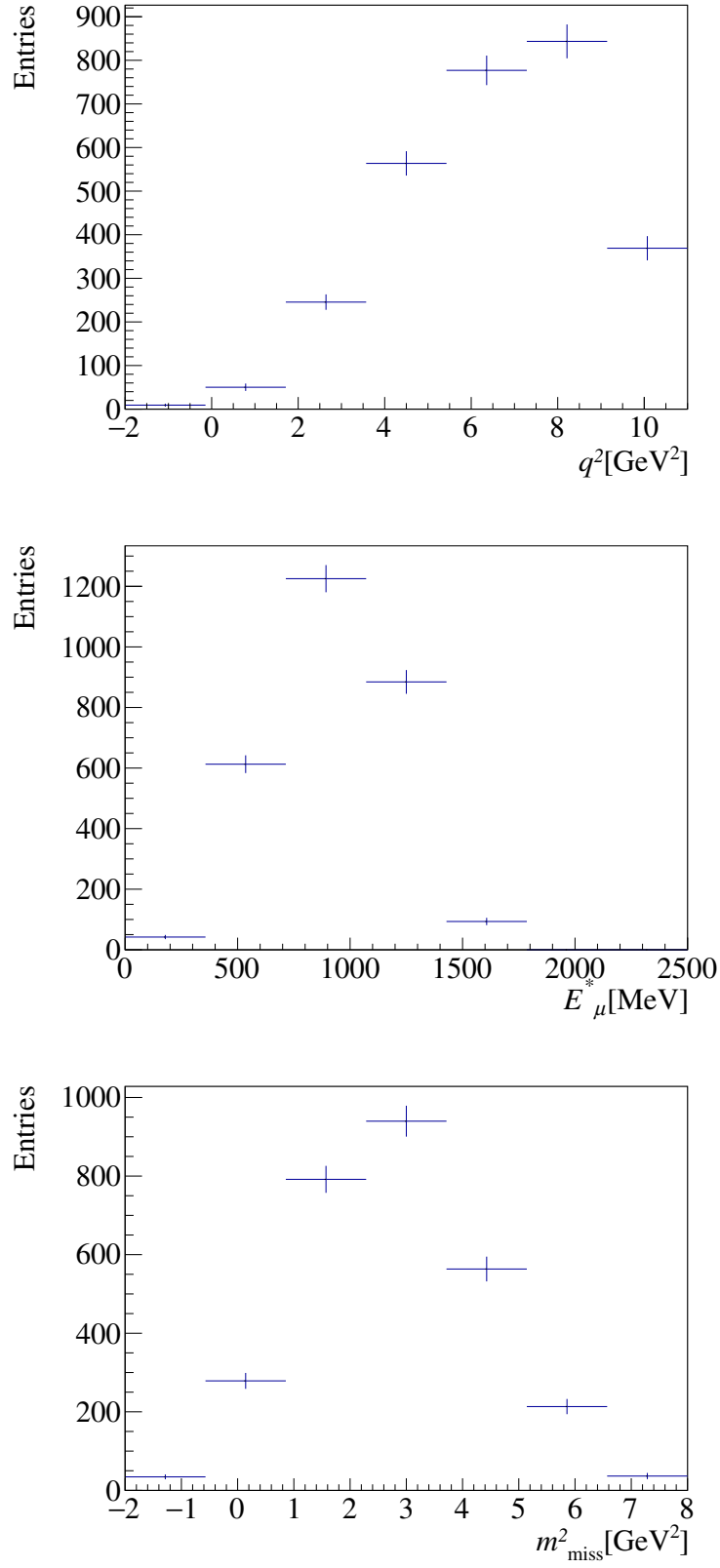


Figure E.4: Projections of the 3D histogram containing the selected $B^+ \rightarrow p(\bar{\Lambda}_c^- \rightarrow \bar{p}X)\mu^+\nu_\mu$ simulation sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

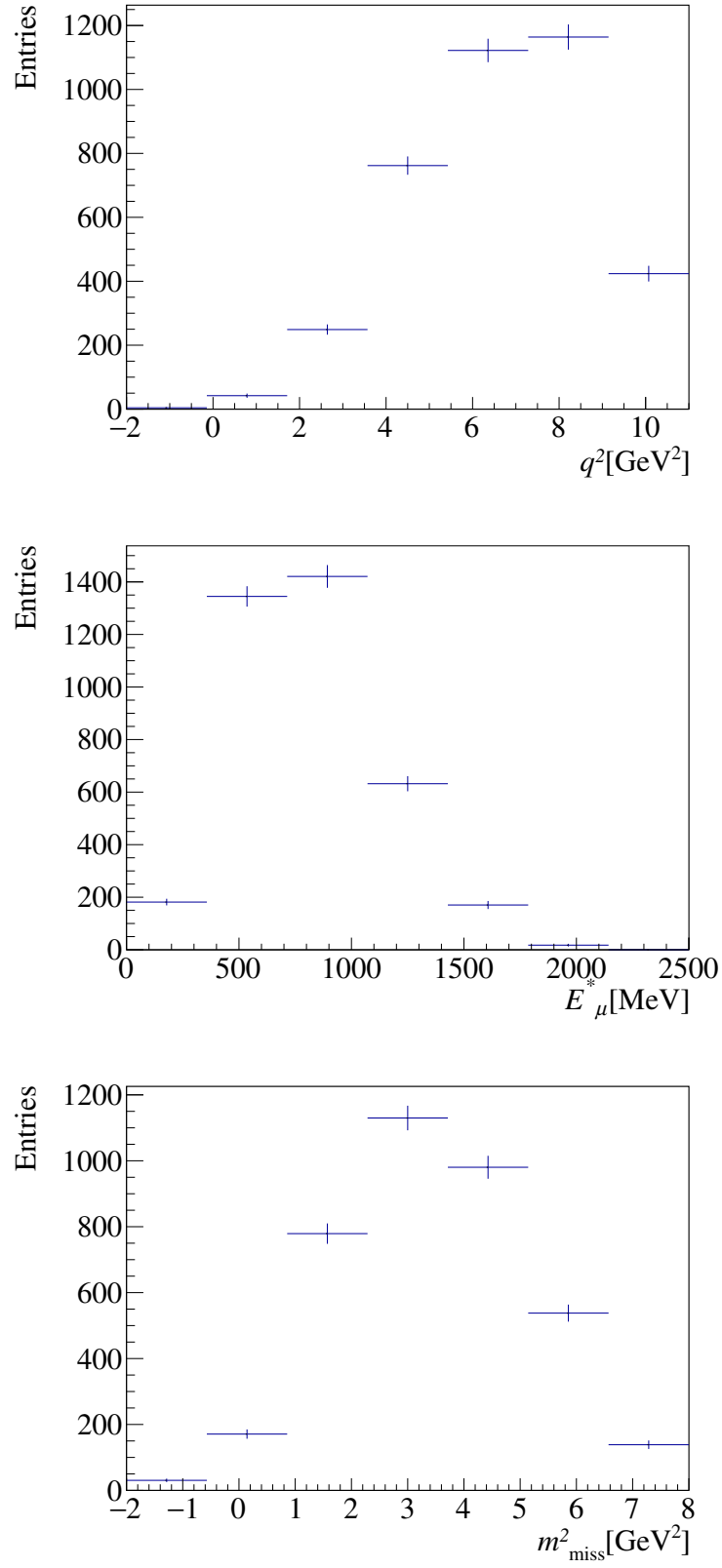


Figure E.5: Projections of the 3D histogram containing the selected $B^0 \rightarrow p\bar{p}(D \rightarrow \mu^+\nu_\mu X)X$ simulation sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

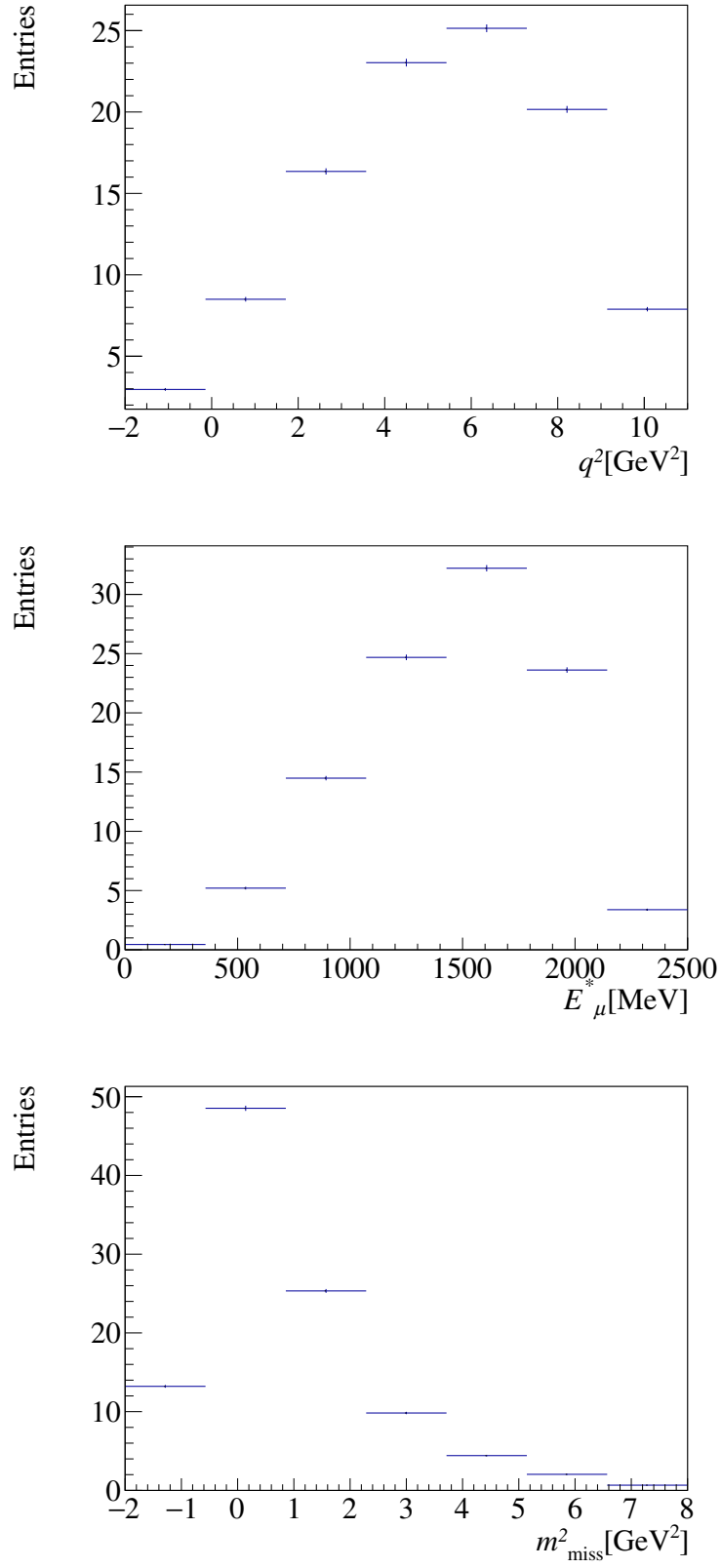


Figure E.6: Projections of the 3D histogram containing the selected p mis-ID 1 sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

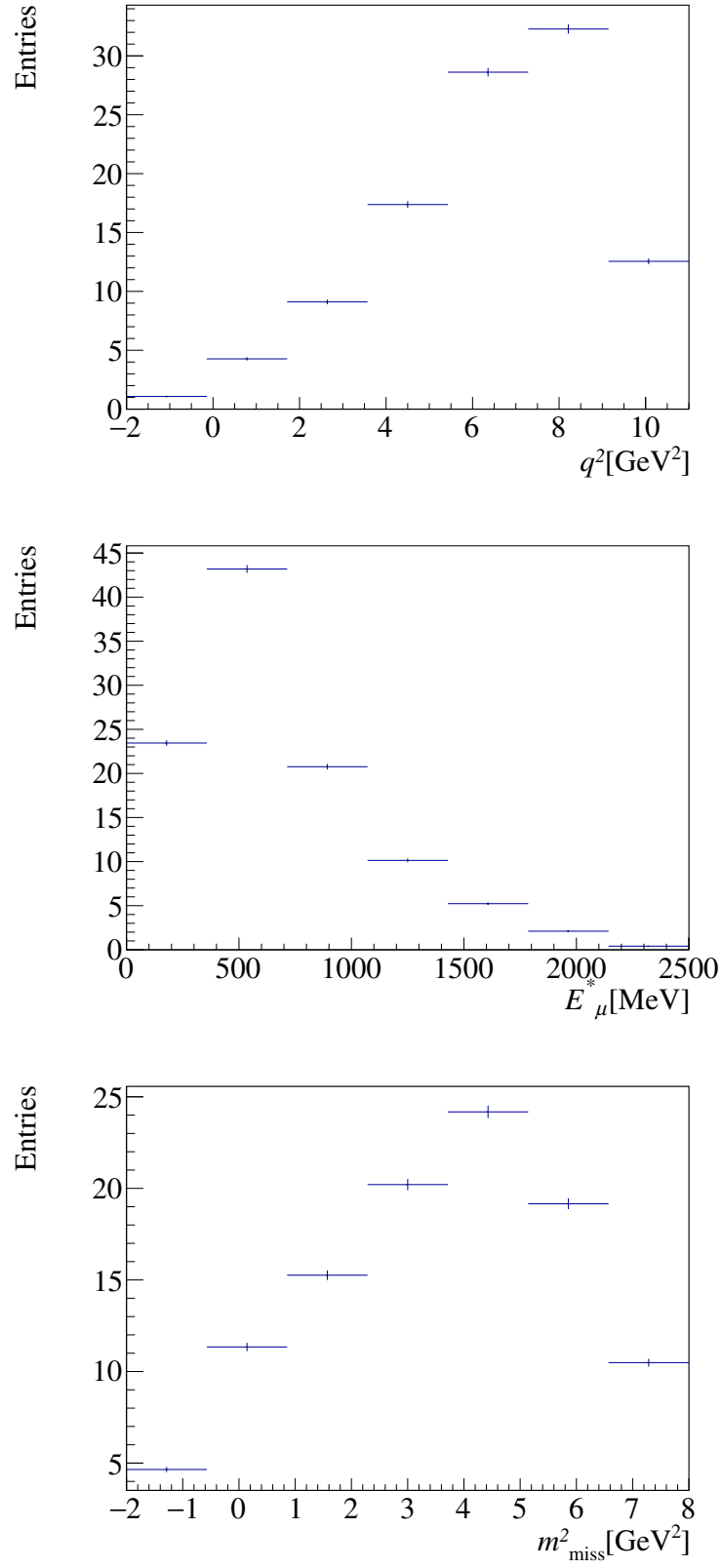


Figure E.7: Projections of the 3D histogram containing the selected μ mis-ID sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

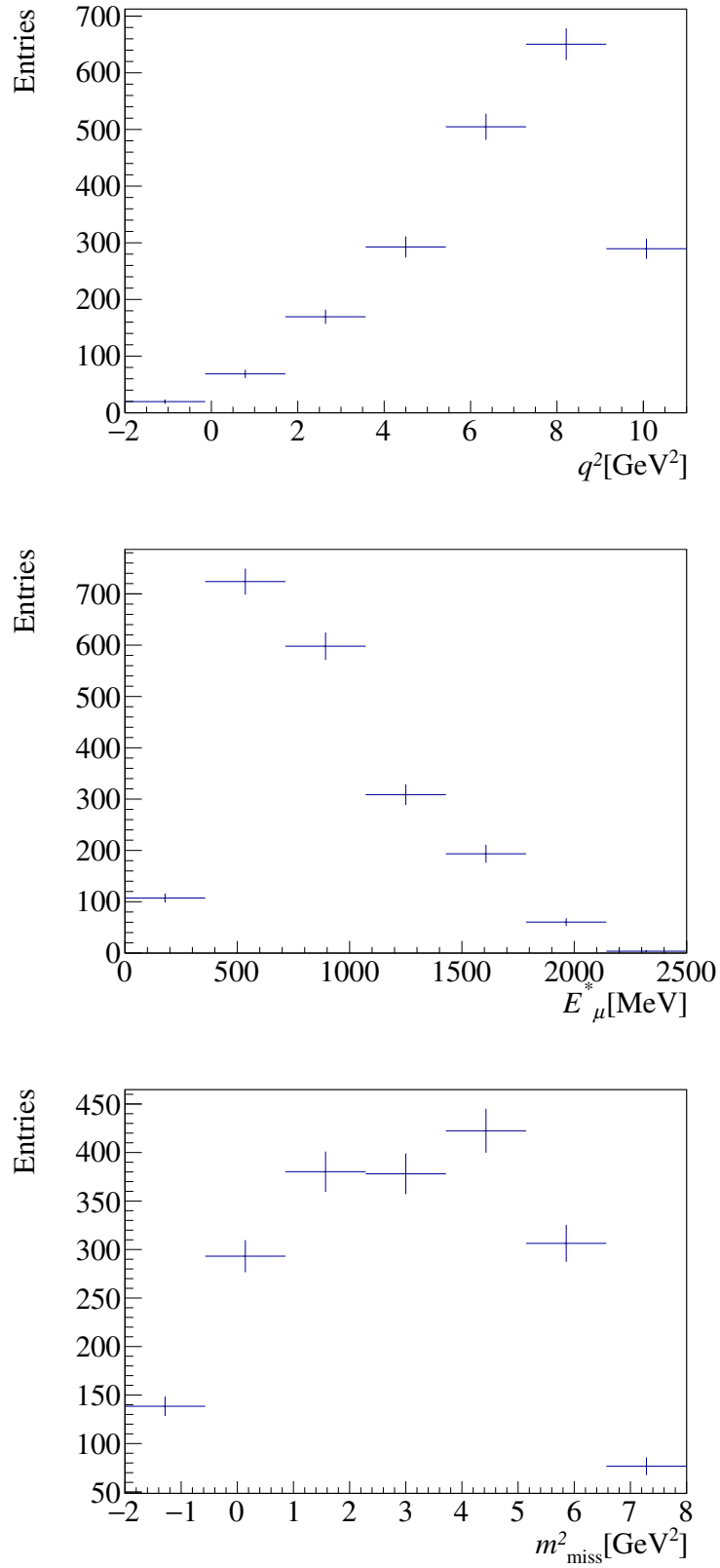


Figure E.8: Projections of the 3D histogram containing the selected combinatorial sample in (top) q^2 , (middle) E_μ^* and (bottom) m_{miss}^2 .

Appendix F

Part-reco background shape variations

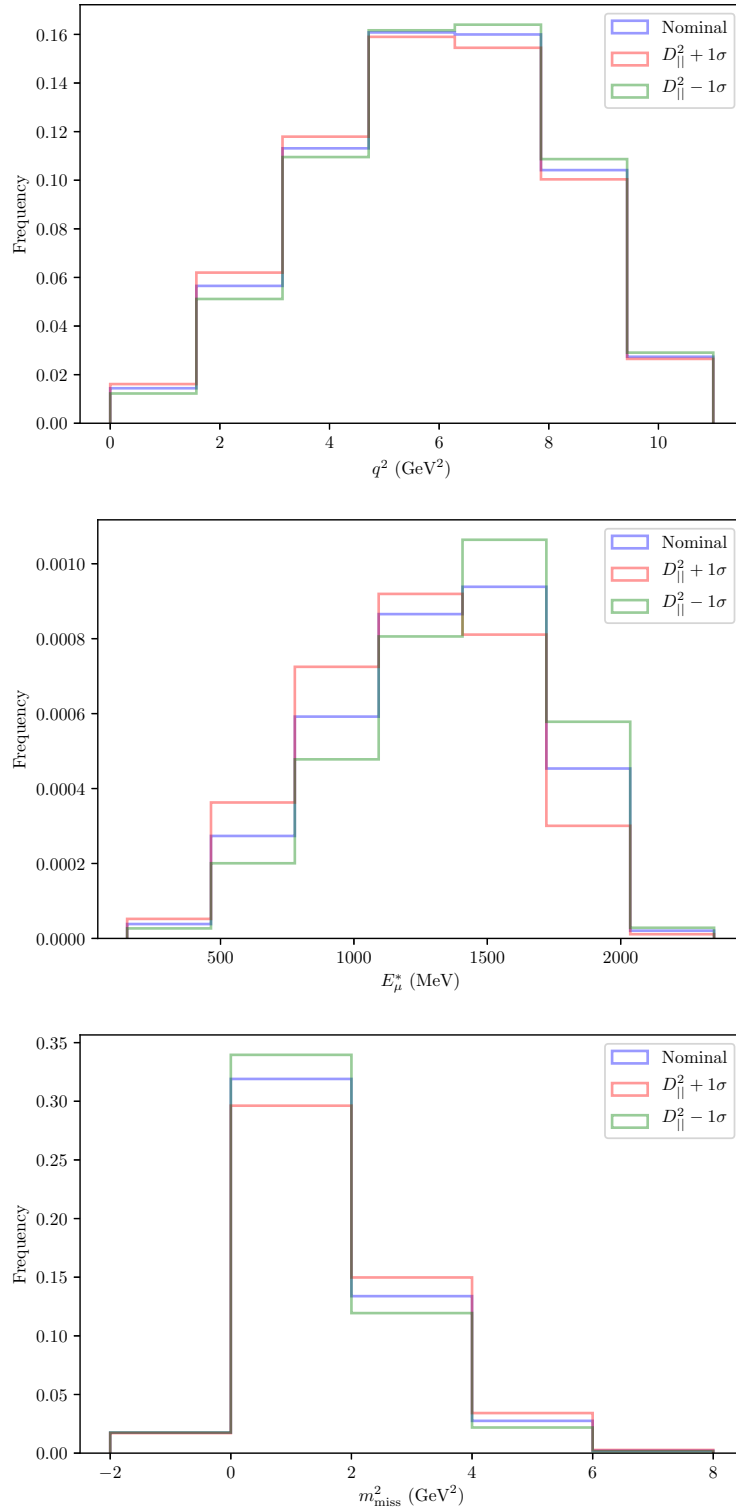


Figure F.1: Shape variations included for the $B^+ \rightarrow p\bar{N}^*\mu^+\nu_\mu$ template resulting from $\pm 1\sigma$ fluctuations in the FF parameter $D_{||}^2$ for (top) q^2 , (middle) E_{μ}^* and (bottom) m_{miss}^2 .

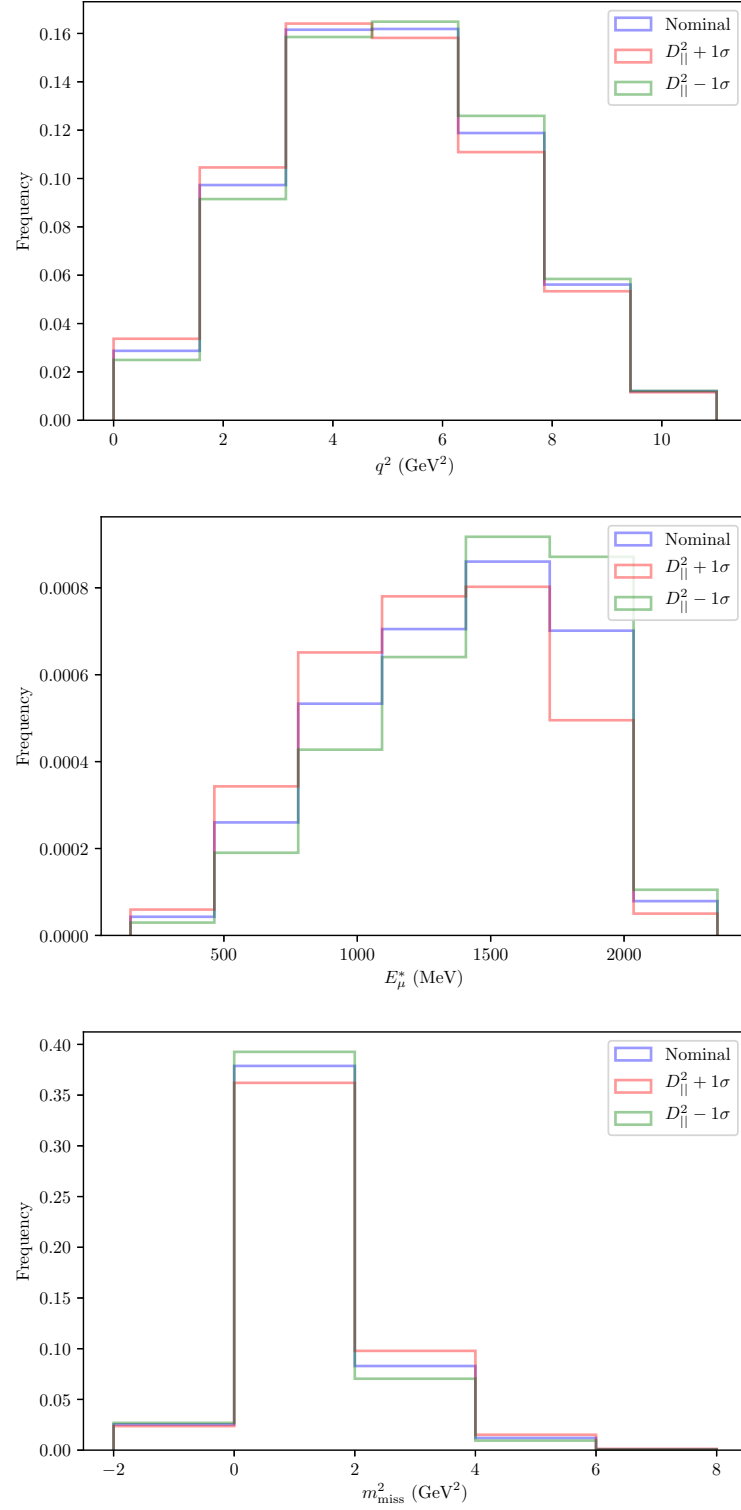


Figure F.2: Shape variations included for the $B^+ \rightarrow p\bar{\Delta}\mu^+\nu_{\mu}$ template resulting from $\pm 1\sigma$ fluctuations in the FF parameter $D_{||}^2$ for (top) q^2 , (middle) E_{μ}^* and (bottom) m_{miss}^2 .