Measurements of q^2 moments of inclusive $B \to X_c \ell \nu$ decays toward an alternative determination of $|V_{cb}|$

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Raynette VAN TONDER
aus
Pretoria, Südafrika

Bonn 2021

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn.

1. Gutachter:	Prof. Dr. Florian BERNLOCHNER
2. Gutachter:	Prof. Dr. Jochen DINGFELDER
Tag der Promotion:	22.11.2021
Erscheinungsjahr:	2021

"I go down to the shore in the morning
and depending on the hour the waves
are rolling in or moving out,
and I say, oh, I am miserable,
what shall –
what should I do? And the sea says
in its lovely voice:
Excuse me, I have work to do."

Mary Oliver, A Thousand Mornings

RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

Abstract

Mathematisch-Naturwissenschaftlichen Fakultät
Physikalisches Institut

Doctor rerum naturalium

Measurements of q^2 moments of inclusive $B \to X_c \ell \nu$ decays toward an alternative determination of $|V_{cb}|$

by Raynette VAN TONDER

This thesis presents the first measurement of the first to the fourth moments of the four-momentum transfer squared q^2 spectrum for inclusive $B \to X_c \ell \nu$ decays, where $\ell = e, \mu$. The determination of these moments and their systematic uncertainties are crucial experimental inputs for a novel, alternative method to determine the magnitude of the CKM matrix element V_{cb} using a reduced set of hadronic matrix elements in the heavy quark expansion. The analysis investigates the complete $\Upsilon(4S)$ data set collected by the Belle detector, corresponding to 711 fb^{-1} of integrated luminosity and a total of $(772 \pm 10) \times 10^6$ B meson pairs. To identify and reconstruct the kinematic variables of inclusive semileptonic B meson decays, machine learning techniques are employed to fully reconstruct one of the B mesons in purely hadronic decay modes. Consequently, this approach allows for the kinematics of the remaining B meson to be inferred using conservation laws leading to the explicit reconstruction of the hadronic X_c system and, in turn, the q^2 spectrum. The moments are measured with progressively increasing threshold selections on q^2: from $q^2 > 3.0\,\text{GeV}^2$ to $q^2 > 10.0\,\text{GeV}^2$ while incrementing the threshold in $0.5\,\text{GeV}^2$ steps. To ensure an unbiased result, the measured moments are unfolded and corrected for reconstruction and selection effects. The measured moments are reported separately for electron and muon final states, allowing for a test of lepton flavour universality. No deviation from the expectation of unity is observed. In addition, the presented measurement is combined with theoretical predictions of the $\langle q^n \rangle$ moments and the semileptonic decay rate up to order $1/m_b^4$ to perform a simultaneous fit and extract $|V_{cb}|$ in a data-driven manner. A value of $|V_{cb}| \times 10^3 = 41.7 \pm 1.2$ is obtained, which is consistent with previous determinations within experimental uncertainty.

Acknowledgements

The completion of this research would not have been possible without the insights and encouragement of various groups and individuals. I would like to take a moment to thank them.

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Florian Bernlochner. Thank you for giving me the opportunity to work on many interesting research topics during this endeavour. Additionally, thank you for unwavering support and guidance. I also wish to thank Prof. Dr. Jochen Dingfelder for agreeing to be the second reviewer of this work. Thank you for perceptive observations during group meetings and always having a few minutes for impromptu chats.

I would also like to thank Dr. Lu Cao and Dr. William Sutcliffe for invaluable contributions to the project and for patiently answering endless naïve questions. Additionally, I extend special thanks to Dr. Matteo Fael, Dr. Pablo Goldenzweig, Dr. Peter Lewis, and Dr. Keri Vos for insightful suggestions and useful discussions regarding the analysis. Furthermore, I acknowledge my Belle Review Committee for their constructive and critical review of the analysis methodology: Prof. Kevin Varvell, Dr. Oskar Hartbrich and Dr. Eiasha Waheed. I am also grateful to Dr. Christoph Schwanda for valuable advice and for moral boosting zoom calls in the midst of the pandemic.

To my colleagues in the KIT and Bonn Belle (II) groups, thank you for non-work related banter and jokes while waiting for the coffee machine (or kettle) as well as numerous quests to find delicious food and beer. Thanks should also go to the system admins of both institutes for efficient management of the cluster. Furthermore, I gratefully acknowledge the assistance of Ms. Bärbel Bräunling and Ms. Andrea Fürstenberg. Thank you both for dealing with the administrative paperwork that had to be done throughout my stay in Germany.

Many thanks to my colleagues for reading sections of this thesis and providing detailed feedback: Martin Angelsmark, Lu Cao, Svenja Granderath, Florian Herren, Alina Manthei, Markus Prim, William Sutcliffe and Max Welsch.

I would like to extend my sincere thanks to my parents, Raymond and Lorette van Tonder, for urging me to pursue my dreams and achieve my goals. Thank you for care parcels from home containing all my favourite South African goodies. To all my friends and family scattered across the world, thank you for the overwhelming amount of encouragement that I received in the form of motivational messages, voice notes, phone calls and even postcards. I cannot name everybody here, but know that I am deeply touched.

Lastly, I would like to thank Flo for all the support and encouragement you have shown me throughout this research and the completion of this thesis –the culmination of years of work! Thank you for the countless adventures we shared, and for many more to come.

Shantih.

Contents

List of Figures

List of Tables

List of Abbreviations

ACC	**A**erogel **C**erenkov **C**ounters
BASF	**B**elle **A**nalysi**S** **F**ramework
BEH	**B**rout-**E**nglert-**H**iggs mechanism
CC	**C**harged **C**urrent
CDC	**C**entral **D**rift **C**hamber
CKM	**C**abibbo-**K**obayashi-**M**askawa
DAQ	**D**ata **Acq**uisition
DSSD	**D**ouble-sided **S**ilicon **S**trip **D**etector
DST	**D**ata **S**ummary **T**apes
ECL	**E**lectromagnetic **Ca**lorimeter
EFC	**E**xtreme **F**orward **C**alorimeter
FOM	**F**igure **o**f **M**erit
FR	**F**ull **R**econstruction
GDL	**G**lobal **D**ecision **L**ogic
HER	**H**igh **E**nergy **R**ing
HQE	**H**eavy **Q**uark **E**xpansion
HQET	**H**eavy **Q**uark **E**ffective **T**heory
IP	**I**nteration **P**oint
LER	**L**ow **E**nergy **R**ing
MC	**M**onte **C**arlo
NC	**N**eutral **C**urrent
OPE	**O**perator **P**roduct **E**xpansion
PDF	**P**robability **D**ensity **F**unction
PID	**P**artical **id**entification
QCD	**Q**uantum **C**hromo**d**ynamics
QED	**Q**uantum **E**lectro**d**ynamics
QFT	**Q**uantum **F**ield **T**heory
RMS	**R**oot-**M**ean-**S**quare
RPC	**R**esistive **P**late **C**hamber
RPI	**Rep**arametrization **I**nvariance
SEQ	**Se**quence **C**ontrol
SM	**S**tandard **M**odel
SVD	**S**ilicon **V**ertex **D**etector
TOF	**T**ime **o**f **F**light Counters
TSC	**T**rigger **S**cintillation **C**ounter

In memory of Elsie Tonkin.
Never a dull moment in the company of an adventurous Ouma like you.

Chapter 1

Introduction

The disagreement between inclusive and exclusive measurements of the Cabibbo-Kobayashi-Maskawa [1, 2] matrix element $|V_{cb}|$ has posed a longstanding puzzle that still remains to be solved. One of the best means to determine $|V_{cb}|$ is by studying semileptonic B meson decays to hadronic final states containing a charm quark and a lepton-neutrino pair: $B \to X_c \ell \nu$. These relatively abundant decays are dominated by tree-level processes, which are expected to remain unaffected by new physics contributions in most models and thus offer theoretically clean avenues to measure $|V_{cb}|$. Additionally, semileptonic B decays are theoretically better understood than decays containing purely hadronic final states, due to the factorisation of the leptonic and hadronic final states. Current determinations of $|V_{cb}|$ focus on two complementary approaches that make use of different theoretical assumptions and experimental reconstruction techniques. While the exclusive approach focuses on the reconstruction of a specific decay mode, the inclusive approach aims to measure the sum of all possible final states entailing the same quark-level transition.

The most precise exclusive determinations of $|V_{cb}|$ to date are performed by measuring partial or total differential decay rates of $B \to D \ell \nu$ and $B \to D^* \ell \nu$ decays [3]. In order to extract a value of $|V_{cb}|$ from the measured differential decay rates, theoretical predictions of hadronic transition form factors are required. The estimation of these functions involve non-perturbative methods, resulting in complicated calculations and larger theoretical uncertainties on the extracted value of $|V_{cb}|$ depending on the chosen parametrization. In contrast, inclusive determinations of $|V_{cb}|$ exploit the fact that the total decay rate can be expanded into a manageable number of non-perturbative matrix elements using the heavy quark expansion (HQE) [4, 5]. With this technique the total semileptonic decay rate is expressed as a series of terms proportional to increasing powers of the inverse bottom quark mass $1/m_b^n$, while corrections proportional to the strong coupling constant, $\alpha_s(m_b)$, are systematically included. A simultaneous global fit is then performed in order to extract a value of $|V_{cb}|$, together with the non-perturbative parameters of the HQE. However, with each increasing order of the expansion in $1/m_b^n$, the number of non-perturbative parameters increases and finally proliferates, which complicates their extraction at order $1/m_b^4$ and higher. Therefore, accessing higher orders in the expansion requires the modelling of HQE parameters to gain a quantitative idea of the possible size of these contributions.

A novel method for the extraction of inclusive $|V_{cb}|$ is proposed in [6]: by exploiting a symmetry within the HQE, known as reparametrization invariance, the coefficients of different HQE operators can be linked by specific linear combinations of the most general HQE parameters. Making use of these relations, the full set of thirteen non-perturbative parameters can be drastically reduced to a set involving only

eight parameters at the order of $1/m_b^4$. This reduction of the total number of non-perturbative parameters allows for the extraction of $|V_{cb}|$ from inclusive decays in a model-independent manner using a complementary data-driven approach. New measurements are required to determine this reduced set of parameters, since the four-momentum transfer squared q^2 spectrum is found to be the unique observable that satisfies the key prerequisite giving rise to the reparametrization invariance.

This thesis presents the first measurement of the first to the fourth moments of the four-momentum transfer squared q^2 spectrum for inclusive $B \to X_c \ell \nu$ decays, where $\ell = e, \mu$. The analysis investigates the complete $\Upsilon(4S)$ data set collected by the Belle detector, corresponding to 711 fb^{-1} of integrated luminosity and a total of $(772 \pm 10) \times 10^6$ B meson pairs. To identify and reconstruct the kinematic variables of inclusive semileptonic B meson decays, one of the B mesons present in the event is fully reconstructed in purely hadronic decay modes with the aid of neural networks. Consequently, this approach allows for the kinematics of the remaining B meson to be inferred using conservation laws, leading to the explicit reconstruction of the hadronic X_c system and, in turn, the explicit reconstruction of the q^2 spectrum. The first measurement of the first moment of the q^2 spectrum was reported in [7] with a lepton energy requirement of 1 GeV[1]. However, this selection effectively reintroduces the full set of hadronic matrix elements, since the lepton energy is not a reparametrization invariant quantity. Furthermore, the moments of higher order are found to be more sensitive to the higher order terms of the HQE. This thesis reports a first systematic study with progressively increasing threshold selections on the q^2 distribution, while the third and fourth order q^2 moments are presented for the first time. The final q^2 moments are measured separately for electron and muon $B \to X_c \ell \nu$ final states, allowing for a test of lepton flavour universality. Furthermore, a non-linear transformation is applied to the nominal measured moments in order to directly determine the central moments, which are less correlated compared to the systematic correlations of the nominal moments. In addition, this thesis presents a preliminary test of the proposed alternative method to determine inclusive $|V_{cb}|$ in a fully data-driven manner: the presented measurement is combined with theoretical predictions of the $\langle q^n \rangle$ moments and the semileptonic decay rate up to order $1/m_b^4$ to perform a simultaneous fit and extract $|V_{cb}|$. Subsequently, a determination of inclusive $|V_{cb}|$ together with three heavy quark parameters at order $1/m_b^4$ using conservative theory uncertainties is reported.

Chapters of this thesis are structured in the following manner. First, the foundational theoretical background concerning semileptonic B meson decays is introduced in Chapter 2, followed by an overview of the KEKB accelerator complex and the Belle detector in Chapter 3. Additionally, the recorded data set and simulated samples used in this analysis are described. Next, the reconstruction algorithms employed by the Belle Collaboration to identify candidate final state particles and subsequently reconstruct B meson decays are summarised in Chapter 4. Contributions from various background processes are subtracted in an unbinned approach by making use of event-by-event weights. These background subtraction weights are determined by performing an extended maximum likelihood fit. Chapter 5 describes the implementation of the complete background subtraction procedure, while Chapter 6 outlines the analysis strategy to extract the $\langle q^n \rangle$ moments. The calibration procedure

[1] Natural units are used throughout this thesis: $\hbar = c = 1$.

that aims to correct the measured moments from detector resolution and acceptance effects is introduced, and is verified on independent simulated samples. Furthermore, the statistical and systematic uncertainties affecting the measured moments are obtained and described. Finally, the measured $\langle q^n \rangle$ moments are presented and compared to the expectation of the exclusive $B \to X_c \ell \nu$ composition by making use of simulated samples. A preliminary simultaneous fit is performed to fit the measured $\langle q^n \rangle$ moments to the theoretical predictions at order $1/m_b^4$ in order to extract a value of $|V_{cb}|$. The final chapter provides a summary of the investigations and results obtained in this thesis.

Additional material relating to corrections applied to outdated simulated samples is considered in Appendix A. Furthermore, distributions of various kinematic variables that are used to discriminate signal events from background processes are shown in Appendix B. The event-wise signal probability weights for each considered threshold selection on the q^2 distribution for both electron and muon candidates are given in Appendix C. In addition, all material related to the calibration procedure and verification thereof for muon candidates is given in Appendix D. Since the assumed $B \to X_c \ell \nu$ composition is a dominant systematic uncertainty on the measured moments, the $B \to X_c \ell \nu$ modelling systematics are summarised in Appendix E. Statistical and experimental correlations for both the measured nominal and central moments of different threshold selections and orders are presented in Appendix F. Moreover, comparisons of the measured nominal moments for both electron and muon candidates to simulated samples are shown in Appendix G. The measured $\langle q^n \rangle$ moments for muon candidates are fitted to the theoretical predictions at order $1/m_b^4$ in Appendix H. Lastly, studies concerning measured $\langle q^n \rangle$ moments for threshold selections between $0.0\,\text{GeV}^2$ and $2.5\,\text{GeV}^2$ are condensed in Appendix I.

Chapter 2

Theoretical background

Since its completion in the 1970s, the Standard Model of Particle Physics (SM) has become one of the most successful and widely accepted theories in physics. The aim of the SM is to provide a description of the universe at the very smallest scale through the fundamental particles that comprise all known matter. These elementary particles and the forces that govern their interactions form the basis of the SM, which combines quantum electrodynamics (QED) [8], the Glashow-Weinberg-Salam theory of electroweak processes [9–11] and quantum chromodynamics (QCD) [12, 13] to formulate a renormalisable Quantum Field Theory (QFT) [14]. Furthermore, to allow for massive vector bosons and fermions the electroweak symmetry is broken in the SM by the Brout-Englert-Higgs (BEH) mechanism [15–17]. Even though the SM is extremely successful in describing a plethora of experimental observations, as well as predicting a wide range of different phenomena, the theory fails to explain several open questions in a satisfactory manner. In particular, the SM fails to explain the dominance of matter over anti-matter in the universe, nor accounts for the observations of dark matter and dark energy. In addition, a renormalisable QFT does not currently exist for gravitation, which therefore remains to be included in the current framework of the SM. The failure to explain these observations, together with additional conceptual problems, suggests that the SM is an incomplete theory, forming the low energy limit of a more fundamental, complete theory.

2.1 The Standard Model of Particle Physics

In the SM elementary particles are characterised by their intrinsic properties: spin, mass, and charge. Fermions carry half-integer spin and are further divided into leptons and quarks. These two types of particles are classified into three generations and account for a total of twelve different elementary particles. On the other hand, particles known as vector bosons carry integer spin. Each vector boson is responsible for mediating at least one of the fundamental forces described by the SM: electromagnetic, strong, and weak. Particles carrying electric charge will partake in electromagnetic interactions, described by the famously successful theory of QED, and are mediated by the massless photon, γ. Strong interactions are described by QCD and are mediated by the exchange of massless gluons –eight in total. Analogous to the role of the electric charge in the electromagentic interaction, particles participating in strong interactions carry a colour charge, which comes in three different flavours: red, blue and green. Therefore, only quarks and gluons are affected by the strong interaction. Lastly, the weak force is conveyed by the massive, charged W bosons and neutral Z bosons, and it affects both leptons and quarks.

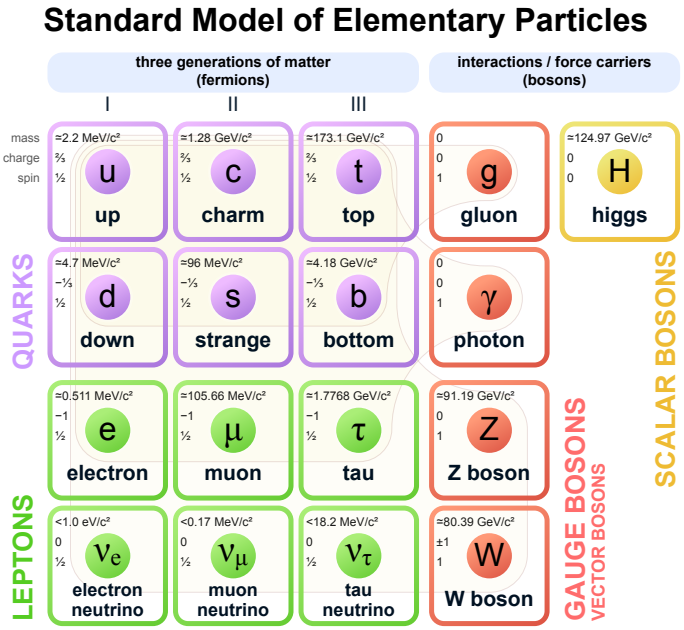

FIGURE 2.1: Classification of elementary particles in the Standard Model [23]. Fermions are divided into quarks (violet) and leptons (green), that can be further divided into three generations (columns). Vector bosons (red) are grouped together with the scalar Higgs (yellow). Each particle's mass, charge and spin is also given.

The SM is a gauge theory [18, 19] that requires the SM Lagrangian to be invariant under complex phase transformations of a set of global symmetries. Each of these global symmetries is connected to an underlying conservation law of nature, as proven by Noether's theorem [20]. For example, the conservation of linear and angular momentum are related to translational and rotational invariance. The local symmetry group that describes the SM is given by:

$$SU(3)_C \times SU(2)_L \times U(1)_Y \,. \tag{2.1}$$

Here, the strong interaction is represented by the $SU(3)_C$ symmetry group with the subscript C denoting colour charge, while the unified electroweak interaction is described by the $SU(2)_L \times U(1)_Y$ symmetry group. The subscripts L and Y refer to the left-handed fields and weak hypercharge, respectively. By requiring gauge invariance under the electroweak $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group, the prediction of massless spin-1 vector bosons arises. While both mediators for the electromagnetic and strong interactions are found to be massless, the mediators for the weak interaction are observed to be very massive. In order to accommodate these massive bosons the electroweak symmetry is spontaneously broken by the presence of a spin-0 field, leading to the prediction of the Higgs boson. The observation of this elementary particle in 2012 by the ATLAS and CMS Collaborations at the Large Hadron Collider [21, 22] provided the final missing piece of the SM. Unlike the vector bosons, the Higgs boson is not associated to any force. A classification of all known elementary particles predicted by the SM as well as their intrinsic properties is shown in Figure 2.1.

In the following sections the theoretical formulations of the strong and electroweak interactions are briefly described, which form the basis of all further calculations concerning semileptonic B decays. Furthermore, the Cabibbo-Kobayashi-Maskawa (CKM) mechanism is introduced together with the current strategy to overconstrain the unitarity condition of the CKM matrix in an effort to find possible hints of physics beyond the SM. The overview of the strong interaction given in Section 2.1.1, as well as the subsequent description of heavy quark methods, is a summary of the discussions found in [24–28]. Furthermore, the overview of the electroweak theory in Section 2.1.3, the summary of charge conjugation and parity (CP) invariance in the SM and the description of the CKM unitarity triangle in Section 2.1.4 are discussed in much greater detail in [27–30].

2.1.1 Strong interactions

Strong interactions in the SM are built on colour symmetry, which is described by a non-Abelian Lie group $SU(3)_C$ that has eight generators corresponding to the gluons. By assuming local gauge symmetry, the QCD Lagrangian can be written as:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G^A_{\mu\nu} G^{\mu\nu}_A + \sum_{q=\{u,d,s\}} \bar{q}(i\not{D} - m_q)q + \sum_{Q=\{c,b,t\}} \bar{Q}(i\not{D} - m_q)Q\,, \tag{2.2}$$

where q and Q are triplets of the light and heavy quarks, m_q is the quark mass matrix, $G^A_{\mu\nu}$ is the gluon field strength tensor and the covariant derivative for the QCD gauge invariance D is given by,

$$iD_\mu = i\partial_\mu - gA^A_\mu T^A = i\partial_\mu - gA_\mu\,. \tag{2.3}$$

Here, g is the strong coupling constant, A^A_μ denotes the eight gluon fields, while the eight generators of the gauge group are given by T^A. The distinction between light (q) and heavy quarks (Q) in Eqn. 2.2 is introduced in anticipation for the discussion in the following section. One of the successes of QCD was the realisation that the coupling constant, defined as

$$g_s = \sqrt{4\pi\alpha_s}\,, \tag{2.4}$$

is in fact not a constant, but depends on the separation distance between interacting particles. An interplay of two different processes contributes towards this dependence and thus the overall effective coupling strength, α_s. On the one hand, virtual quark-antiquark pairs create a shielding effect that reduces the value of α_s for increasing distances. On the other hand, since gluons themselves carry colour charge, an additional contribution stemming from gluon self-interactions creates an opposing anti-screening effect. The effect of the gluon anti-screening exceeds that of the quark-antiquark shielding, leading to an increase in the magnitude of α_s with increasing distance. This behaviour results in the confinement of quarks, which offers a possible explanation to the fact that no single gluon has been observed yet by experiments. In contrast, α_s decreases at short distances, leading to the phenomenon of asymptotic freedom in which quarks are allowed to behave as free particles. At leading order, the effective running coupling constant can be written explicitly as a function of the energy scale:

$$\alpha_s(q^2) = \frac{12\pi}{(33 - 2N_q)\ln(q^2/\Lambda_{\text{QCD}})} \tag{2.5}$$

where N_q is the effective number of quark flavours, q^2 is the energy scale of the process or the momentum transfer squared, and the scale at which QCD becomes strongly coupled is defined by the constant Λ_{QCD}. At this point non-perturbative effects become important and perturbation theory breaks down. Experimentally, $\Lambda_{\text{QCD}} \approx 200\,\text{MeV}$ and sets the scale for non-perturbative strong interaction effects [24].

2.1.2 Heavy quark methods

Methods based on the fact that some of the quark masses are large compared to Λ_{QCD} are standard strategies in B physics. In the case of semileptonic B decays, the B meson can be treated as a bound system containing a heavy quark Q and light degrees of freedom, i.e., the light quark and soft-gluon interactions. Since the mass of the b quark is approximately 4 GeV [30] and thus much larger than Λ_{QCD}, the heavy quark moves non-relativistically within this system. The momentum of the heavy quark $\vec{p_Q}$ must balance the momentum of the light constituents $\vec{p}_{\text{light}}$ such that:

$$|\vec{p_Q}| = |\vec{p}_{\text{light}}| \sim \Lambda_{\text{QCD}}\,. \tag{2.6}$$

For the case of heavy-light systems the typical momentum transfer to the light constituents is of order Λ_{QCD}, while the size of the hadron is approximately $1/\Lambda_{\text{QCD}}$. Thus,

$$|\vec{v}_Q| = \frac{|\vec{p}_Q|}{m_Q} \sim \frac{\Lambda_{\text{QCD}}}{m_Q}\,. \tag{2.7}$$

In the limit $m_Q \to \infty$, the configuration of the light degrees of freedom within a hadron containing a heavy quark with velocity v is almost completely undisturbed if this quark is replaced by another heavy quark with the same velocity. Relativistic effects depending on the mass of the heavy quark become negligible and the heavy quark behaves as a stationary source of the colour field. Therefore, in the heavy quark limit both the flavour and spin of the heavy quark decouple from the dynamics of the system, leading to an $SU(2N_h)$ spin-flavour symmetry group. These heavy quark symmetries, which are obfuscated in the QCD Lagrangian, are exploited by Heavy Quark Effective Theory (HQET): an effective field theory specifically constructed to describe the dynamics of heavy quarks inside heavy mesons.

In contrast to the QCD Lagrangian, HQET is constructed in a manner such that only inverse powers of m_Q appear in the effective Lagrangian. The first step is the observation that a heavy quark bound within a hadron moves with the velocity of the hadron and is almost on-shell, $p_Q^2 = m_Q^2$. It is convenient to write the momentum p_Q^μ as a sum of two parts

$$p^\mu = m_Q v^\mu + k^\mu\,, \tag{2.8}$$

where $m_Q v^\mu$ is the large on-shell contribution, with v^μ the four-velocity of the heavy quark, and $k^\mu \sim \Lambda_{\text{QCD}}$ originates from the non-perturbative interactions binding the hadron together. The soft interactions of the heavy quark with the light degrees of freedom change the residual momentum by a small amount of order $\Delta k \sim \Lambda_{\text{QCD}}$, but the corresponding changes in the heavy quark velocity vanish as $\Lambda_{\text{QCD}}/m_Q \to 0$. To separate the heavy and light degrees of freedom, the four-velocity is used to split the heavy quark field Q into large- and small-components, h_v and H_v, given by

$$h_v(x) = e^{im_Q v\cdot x} P_+ Q(x) \qquad \text{and} \qquad H_v(x) = e^{im_Q v\cdot x} P_- Q(x)\,, \tag{2.9}$$

where $P_\pm$ are projection operators defined as: $P_\pm = (1 \pm \not{v})/2$. The heavy quark field can thus be expressed as:

$$Q_x = e^{im_Q v\cdot x}[h_v(x) + H_v(x)]\,, \tag{2.10}$$

with the new fields satisfying

$$\not{v} h_v = h_v \qquad \text{and} \qquad \not{v} H_v = -H_v\,, \tag{2.11}$$

due to the projection operators. These equations can be inserted into the heavy quark contribution of the QCD Lagrangian (Eqn. 2.2) to obtain:

$$\mathcal{L}_Q = \underbrace{\bar{h}_v(iv\cdot D)h_v}_{\text{massless}} - \underbrace{\bar{H}_v(iv\cdot D + 2m_Q)H_v}_{\text{heavy d.o.f with } 2m_Q} + \underbrace{\bar{h}_v(i\not{D}_\perp)H_v + \bar{H}_v(i\not{D}_\perp)v_v}_{\text{coupling}}\,, \tag{2.12}$$

where $D^\mu_\perp = D^\mu - (v\cdot)v^\mu$ is orthogonal to the heavy quark velocity such that: $v \cdot D_\perp = 0$. In this formulation, it becomes apparent that h_v describes the massless degrees of freedom due to the residual momentum k of the heavy quark. In contrast, all heavy degrees of freedom are described by H_v with twice the mass of the heavy quark. The third and fourth terms in Eqn. 2.12 describe pair annihilation or creation of heavy quarks and antiquarks. The heavy degrees of freedom H_v can be eliminated by using the equation of motion. Thus, taking the variation of the Lagrangian with respect to the field H_v,

$$(iv\cdot D + 2m_Q)H_v = i\not{D}_\perp h_v \qquad \text{and thus} \qquad H_v = \frac{1}{iv\cdot D + 2m_Q} i\not{D}_\perp h_v\,. \tag{2.13}$$

Here it becomes clear that the small-component field H_v is of order $1/m_Q$. By inserting this solution into Eqn. 2.12, the "non-local effective Lagrangian" is obtained,

$$\begin{aligned}\mathcal{L}_Q &= \bar{h}_v(iv\cdot D)h_v + i\not{D}_\perp h_v \frac{1}{iv\cdot D + 2m_Q} i\not{D}_\perp h_v \\ &= \bar{h}_v(iv\cdot D)h_v + \frac{1}{2m_Q} i\not{D}_\perp h_v \sum_{n=0}^{\infty}\left(-\frac{iv\cdot D}{2m_Q}\right)^n i\not{D}_\perp h_v\,.\end{aligned} \tag{2.14}$$

In the second step a geometric series of the denominator is used to expand the second term into operators of increasing orders, which are suppressed by increasing powers of $1/2m_Q$. This step yields the heavy quark expansion, which becomes applicable in the heavy quark limit $m_Q \to \infty$. The HQET Lagrangian can schematically be written in the following manner:

$$\mathcal{L}_{\text{HQET}} = \mathcal{L}_0 + \mathcal{L}_1 + \mathcal{O}(1/m_Q^2)\,, \tag{2.15}$$

where the $1/m_Q$ contribution can be shown to be [25],

$$\mathcal{L}_1 = \frac{1}{2m_Q}\bar{h}_v\left((iD_\perp)^2 - \frac{g_s}{2}\sigma_{\mu\nu}\cdot G^{\mu\nu}\right)h_v\,. \tag{2.16}$$

Here, the first term can be identified as the non-relativistic motion of the heavy quark, while the second term describes the colour-magnetic coupling of the heavy quark spin rotation $\sigma_{\mu\nu}$ to the gluon field.

2.1.3 Electroweak interactions

The electroweak theory unifies the electromagnetic and weak interactions into one single gauge theory described by the symmetry group $SU(2)_L \times U(1)$. Three vector bosons, $W^\mu = (W^{\mu 1}, W^{\mu 2}, W^{\mu 3})$, are associated with the $SU(2)$ symmetry group, while a vector boson, known as B, is associated with the $U(1)$ symmetry. This theory includes both charged currents (CC) and neutral currents (NC) of the electromagnetic and weak interactions. The weak CC interactions are mediated by the $W^\pm$ bosons, which are linear combinations of the $W^{\mu 1}$ and $W^{\mu 2}$ vector bosons in the following manner:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2)\,. \tag{2.17}$$

Similarly, it can be shown that the two neutral states, W^3 and B, are mutually orthogonal linear combinations of the fields that mediate the electromagnetic and neutral currents. The neutral states mix to produce two linear combinations: a massless combination corresponding to the photon and a massive linear combination corresponding to the Z^0. The transformation can be expressed in terms of the two coupling constants, g and g', where the fields that mediate the electromagnetic and neutral currents are denoted as A and Z, respectively:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & -g' \\ g' & g \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}\,. \tag{2.18}$$

This transformation may also be expressed as a rotation through the weak mixing angle θ_w, which has been experimentally determined to be approximately 29°:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_\omega & -\sin\theta_\omega \\ \sin\theta_\omega & \cos\theta_\omega \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}\,. \tag{2.19}$$

Due to the observation of massive $W^\pm$ and Z^0 bosons, the electroweak symmetry $SU(2)_L \times U(1)$ is broken in the SM by the BEH mechanism, which introduces a new, complex scalar field Φ with a potential $V(\Phi)$ [15–17]. This scalar field transforms as a self-interacting doublet under $SU(2)_L$ with four real degrees of freedom. The scalar field together with its given potential are shown in Eqn. 2.20 and 2.21.

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.20}$$

$$V(\Phi) = -\frac{1}{2}\mu^2(\phi^*\phi) + \frac{\lambda^2}{4}(\phi^*\phi)^2 \tag{2.21}$$

The potential function V has rotational symmetry with minima that occur on a circle in ϕ space at:

$$|\Phi| = \frac{v}{\sqrt{2}}\,, \qquad \text{with radius } v = \frac{\mu}{\lambda}\,. \tag{2.22}$$

In principle, any state satisfying $|\Phi|^2 = v^2/2$ could be the ground state in the above-mentioned potential, however a specific gauge that yields a massless photon, massive $W^\pm$ and Z^0 bosons and a Higgs particle is a convenient choice. Therefore, the following ground state is chosen such that,

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \qquad \text{with } v = \sqrt{\mu^2/\lambda}\,. \tag{2.23}$$

By expanding about the ground state, the perturbations can be viewed as excitations of the particle field

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{2.24}$$

where $H = H(x)$ is a real-valued scalar field representing the physical Higgs boson and $H(x) = 0$ is the ground state of the potential. By assigning the particular ground state given in Eqn. 2.29, the SM symmetry of $SU(3)_C \times SU(2)_L \times U(1)_Y$ is broken into $U(1)_{em}$. This effect is known as spontaneous symmetry breaking, since the symmetry of the underlying Lagrangian is not respected by a particular ground state of the system. The effect occurs when a system tends to the lowest energy state without any external influence, hence the term "spontaneous". In total, three of the four degrees of freedom associated with the $SU(2)_L \times U(1)_Y$ gauge group are spontaneously broken. This implies that these three bosons have non-trivial transformations of the ground state and indicate the existence of three massless scalar particles, known as Goldstone bosons. The existence of one or more Goldstone bosons is a general phenomenon that always accompanies the spontaneous symmetry breaking of a continuous global symmetry [31] [32].

The Higgs field couples to the W_μ and B_μ gauge bosons associated with the electroweak symmetry group through the covariant derivative that appears in the kinetic term of the Higgs Lagrangian,

$$\mathcal{L} = \frac{1}{2}(D_\mu \phi)^* (D^\mu \phi) + \frac{1}{2}\mu^2(\phi^* \phi) - \frac{\lambda^2}{4}(\phi^* \phi)^2 . \tag{2.25}$$

As a result, the two charged and neutral massless Goldstone degrees of freedom mix with the corresponding broken generators of $SU(2)_L \times U(1)_Y$ to become the longitudinal polarisation modes of the physical vector bosons, $W^\pm$ and Z^0, as described in [30]. The fourth generator remains unbroken, since it is associated with the unbroken gauge symmetry of $U(1)_{em}$, which corresponds to the massless photon. Therefore, from the initial four degrees of freedom of the BEH field, two are absorbed by the $W^\pm$ vector bosons and another one is absorbed by the Z^0 vector boson such that these bosons become massive, as seen in Eqn. 2.26 and 2.27, respectively.

$$M_W^2 = \frac{g^2 v^2}{4} \tag{2.26}$$

$$M_Z^2 = \frac{(g'^2 + g^2)v^2}{4} \tag{2.27}$$

The final remaining degree of freedom is the physical Higgs boson, a new scalar particle with mass $m_H = \sqrt{2\lambda}v$ where λ is the Higgs self-coupling parameter in $V(\Phi)$.

2.1.4 CP violation and the unitarity triangle

Unlike the other SM interactions, the weak interaction violates both parity transformations and charge conjugation. Prior to 1956, parity invariance was assumed to be one of the fundamental conservation laws of physics. However, Lee and Yang revealed that although ample evidence for parity invariance in strong and electromagnetic processes existed, there was no confirmation for weak interactions [33]. Soon thereafter, the first evidence of parity violation in the weak interaction was

demonstrated by the famous experiment carried out by Wu and collaborators using β-decays in polarised Cobalt-60 [34]. The violation was neither a small effect, nor was it limited to the beta decay in cobalt. It became clear that parity violation was in fact a signature of the weak interaction. As a consequence, charged currents of the weak interaction only act on left-handed fermions and right-handed anti-fermions, while the neutral currents act on both left-handed and right-handed fermions. For each fermion, there is a weak isospin current, $j_\mu = (j^1_\mu, j^2_\mu, j^3_\mu)$, that couples to W^μ with a dimensionless coupling constant of strength g_ω. Parity violation is expressed by a projector P_L in the weak charged current as:

$$j_\mu = \bar{\ell}\gamma_\mu P_L \nu \tag{2.28}$$

which projects the fermion fields ℓ and ν onto their left-handed states $(\ell, \nu)_L = P_L(\ell, \nu)$, while P_L is defined by the Dirac gamma matrices as $P_L = \frac{1}{2}(1 + \gamma^5)$. While the discovery of parity violation was devastating, the combined operation of both charge and parity (CP) transformations was assumed to be infallible. This assumption was also shown to be false: the violation of CP symmetry was first observed in the decays of the neutral kaon system by Cronin and Fitch in 1964 [35].

Meanwhile, additional puzzles became evident concerning weak interactions. In the case where the weak interaction is mediated by a $W^\pm$ boson, the strength of the coupling to the neutrino-lepton pair across all generations of leptons is found to be the same, a property known as lepton flavour universality. However, the analogous situation is not observed in the quark sector. Here, weak transitions were observed to occur across quark generations. This observation led Cabibbo to formulate the concept of quark mixing, where the weak eigenstates of quarks that couple to the $W^\pm$ bosons are superpositions of the quark mass eigenstates [1]. The mixing of the first and second generation of quarks is allowed through a 2×2 unitary matrix expressed in terms of a single parameter: the Cabibbo angle θ_C [36]. Maskawa and Kobayashi extended this formulation to three generations arguing that CP violation arises naturally in this theory, which was impossible to include using only two generations [2]. The resulting 3×3 matrix transformation that relates the weak eigenstates (d', s', b') to the mass eigenstates (d, s, b) is known as the CKM matrix, defined as:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{2.29}$$

where the probability of a flavour transition from the i-th generation up-type quark to a j-th generation down-type quark is proportional to the CKM matrix element $|V_{ij}|^2$. The values of the matrix elements exhibit a clear hierarchy leading to CKM suppression of transitions between different generations. This can be visualised by making use of the Wolfenstein parametrization [37], which expresses the CKM matrix as a Taylor expansion around $\lambda = |V_{us}| \approx 0.225$:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}, \tag{2.30}$$

where A, ρ and η are additional parameters of order unity. Here, it can bee seen that off-diagonal matrix elements are relatively small, while the diagonal elements are close to unity. The suppression is maximal for the couplings between the first

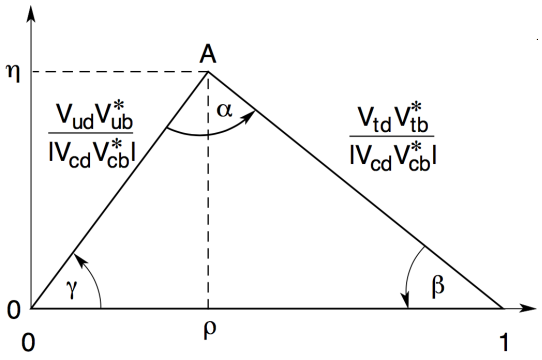

FIGURE 2.2: The triangle expressing the unitarity condition for the first and third columns of the CKM matrix [38].

and third generation quarks, ub and td. In addition, the complex phase η is found to reside only in the two outermost off-diagonal elements, implying that η must be non-zero for CP violation to occur in the quark sector.

Due to the conservation of probability, the CKM matrix satisfies the unitary condition $VV^\dagger = I$, resulting in the following relations between the matrix elements:

$$\sum_i = V_{ij}V_{ik}^* = \delta_{jk}\,, \qquad \text{and} \qquad \sum_i = V_{ij}V_{kj}^* = \delta_{ik}\,. \tag{2.31}$$

The six vanishing unitarity conditions are graphically represented as unitarity triangles, while the conditions obtained from neighbouring rows and columns are nearly degenerate. Since the lengths and sides of the triangles are invariant under phase transformations, these values can be measured using suitable experiments. Furthermore, the areas of all the triangles are equal, which determine the size of CP violation within the SM. The most commonly used triangle is found by imposing the unitarity relation on the first and third columns:

$$V_{ud}V_{ub}^* + V_{td}V_{tb}^* + V_{cd}V_{cb}^* = 0\,. \tag{2.32}$$

Each term in the sum is of the order λ^3, which enables the measurement of the angles of the triangle if the complex phase is non-zero. Dividing this expression by the most precisely measured term, $V_{cd}V_{cb}^*$, results in a triangle with unit base and two fixed vertices at $(0,0)$ and $(0,1)$. The apex of the triangle is denoted as (ρ,η). By using the parametrization given in Eqn. 2.30 and introducing two additional parameters, $\bar{\rho} = \rho(1-\lambda^2/2)$ and $\bar{\eta} = \eta(1-\lambda^2/2)$, the sides of the triangle are redefined as

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} = \bar{\rho} + i\bar{\eta}\,, \qquad \text{and} \qquad \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 1 - \bar{\rho} - i\bar{\eta}\,. \tag{2.33}$$

A schematic of the unitarity triangle is shown in Figure 2.2.

Various processes are studied to determine the sides and angles of the unitarity triangle, which are used in combination to ultimately overconstrain the unitarity condition. For instance, a combined global fit in the $(\bar{\rho},\bar{\eta})$ plane can be performed using a compendium of measurements and predictions as input. A non-exhaustive list of input results is given below [39]:

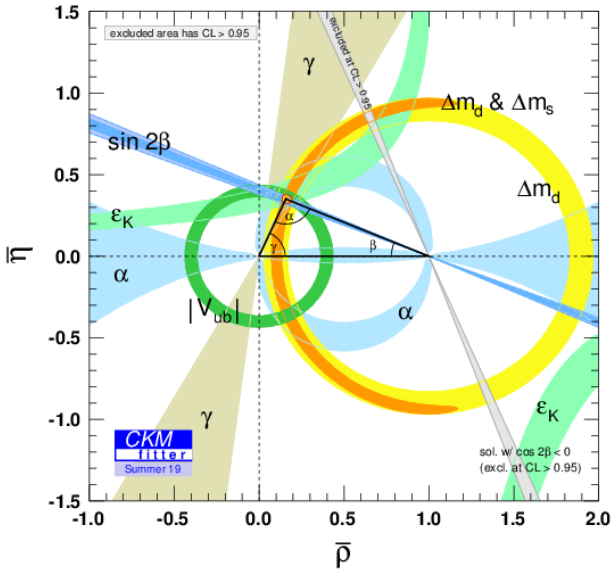

FIGURE 2.3: Experimental constraints on the unitarity triangle as provided by the CKMfitter group [39].

- ϵ_K and $\sin 2\beta$: measured by studying CP violation in the neutral kaon system or the B meson system, respectively.
- Δm_d and Δm_s: determined by studying the mass splitting from $B_{(s)}$ meson mixing.
- $\sin 2\alpha$: constrained significantly by measuring CP violation in $B^0 \to \rho^+\rho^-$ decays.
- $|V_{ub}|$ and $|V_{cb}|$: measured by studying semileptonic B decays.

There are several approaches to combine the available experimental data: CKMfitter [39] uses frequentist statistics, while UTfit [40] employs a Bayesian approach. Both methods yield similar outcomes [30]. A result of the global fit to determine the position of the unitarity triangle's apex, performed by the CKMfitter group, is shown in Figure 2.3. A deviation of the apex position from the theoretical expectation could hint at the existence of possible new physics processes that have not been predicted by the SM.

Measurements of the $|V_{ub}|$ and $|V_{cb}|$ matrix elements are of particular interest, since these values are determined using tree-level dominated processes that are expected to remain unaffected by new physics contributions. Therefore, a comparison of precise measurements of $|V_{ub}|$ and $|V_{cb}|$ is crucial to uncover a possible violation of the unitarity condition and provide an unbiased measure of the amount of CP violation in the quark sector. One of the best modes to determine both $|V_{ub}|$ and $|V_{cb}|$ is semileptonic $B \to X\ell\nu$ decays. In the following sections, the different experimental and theoretical strategies to precisely determine $|V_{cb}|$ by studying $B \to X_c\ell\nu$ are discussed.

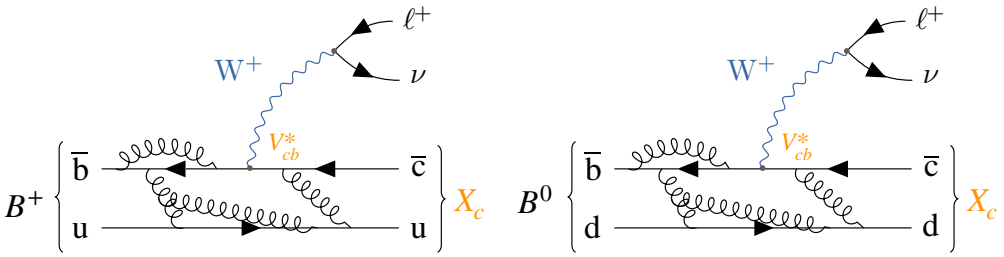

FIGURE 2.4: Feynman diagram of semileptonic $B \to X_c \ell \nu$ decays for charged (left) and neutral (right) B mesons.

2.2 Semileptonic B decays

Semileptonic $B \to X_c \ell \nu$ processes, depicted[1] in Figure 2.4, involve the decay of B mesons into a lepton-neutrino pair together with a set of charmed hadrons, where X_c denotes the hadronic system. These relatively abundant decays are dominated by tree-level processes and therefore offer a theoretically clean avenue to determine the magnitude of the CKM matrix element $|V_{cb}|$. Since the resulting leptons are unaffected by strong interactions, the leptonic and hadronic final states factorise up to small electroweak corrections. Consequently, these processes are theoretically better understood than purely hadronic decays. This is due to the fact that perturbation theory is found to be invalid for the long-distance strong interactions that occur in decays involving hadronic final states, whereas the theory is perfectly valid for calculations of higher-order corrections at tree-level for electroweak interactions.

A key characteristic of semileptonic B decays is the relative simplicity due the factorisation of the leptonic and hadronic final states. In the SM, semileptonic flavour-changing weak decays are mediated by the exchange of charged $W^\pm$ bosons between a quark and lepton charged current according to the electroweak interaction Lagrangian:

$$\mathcal{L}_{\text{int}} = \frac{g}{\sqrt{2}}(j_\mu^- W_+^\mu + j_\mu^+ W_-^\mu). \tag{2.34}$$

The hadronic current is written as,

$$j_\mu = \bar{c}\gamma_\mu P_L V_{\text{cb}} b\,, \tag{2.35}$$

where V_{cb} represents the CKM matrix element for $c \to b$ quark transitions, while the leptonic current is expressed in the following manner:

$$j_\mu = \bar{\ell}_L \gamma_\mu P_L \bar{\nu}_{\ell,L}\,. \tag{2.36}$$

With the exception of the top quark, all quarks are considered light compared to the mass of the $W^\pm$ boson. Therefore, the $W_\mu^\pm$ can be considered as infinitely heavy in decays of weak hadrons [42], leading to the combination of Eqn. 2.35 and 2.36 into a local effective Hamiltonian of the form,

$$H_{\text{eff}} = \frac{4G_F}{\sqrt{2}} V_{\text{cb}} (\bar{c}\gamma_\mu P_L b)(\bar{\ell}_L \gamma_\mu P_L \bar{\nu}_{\ell,L}) + \text{h.c.} \tag{2.37}$$

[1] Feynman diagrams in this thesis are drawn with the package described in [41].

where G_F is the Fermi coupling constant defined as,

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}\,. \tag{2.38}$$

The simplicity of the leptonic current is overshadowed by the complicated hadronic term, which cannot be calculated in a simple manner. Different approaches to simplify and determine the hadronic matrix element exist that depend on the reconstruction of the $B \to X_c \ell \nu$ final state.

Values of $|V_{cb}|$ are extracted from the study of semileptonic B meson decays by making use of two complementary approaches. While the exclusive approach focuses on the reconstruction of a specific decay mode, the inclusive approach aims to measure the sum of all possible final states with the same quark-level transition. Inclusive determinations of $|V_{cb}|$ exploit the fact that the total decay rate can be expanded into a manageable number of non-perturbative matrix elements using the heavy quark expansion (HQE). A simultaneous combined fit is then performed in order to extract a value of $|V_{cb}|$ as well as the non-perturbative parameters of the HQE. On the other hand, the most precise exclusive determinations of $|V_{cb}|$ to date are performed by measuring partial or total differential decay rates of $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays [3]. To translate the differential decay rates into a measurement of $|V_{cb}|$, theoretical predictions of hadronic form factors are required. Consequently, exclusive determinations provide complementary information and serve as independent cross-checks of the inclusive determinations.

In the next section, the decay kinematics of semileptonic B mesons are briefly discussed. The kinematic decay variables are important to describe both exclusive and inclusive final states that are considered in subsequent sections. Overviews of the experimental and theoretical descriptions of semileptonic B decays are presented in great detail in [24, 28, 42], which the following sections summarise.

2.2.1 Decay kinematics

Semileptonic decays for a fixed mass m_X of the final-state hadronic system can be described by a choice of several kinematic variables. Two key kinematic quantities, the four-momentum transfer squared of the lepton-neutrino system q^2 and the energy of the charged lepton E_ℓ, are defined as:

$$q^2 = (p_\ell + p_\nu)^2 = (p_B - p_X)^2\,, \qquad E_\ell = \frac{p_B p_\ell}{m_B}\,. \tag{2.39}$$

Here, m_B and p_B correspond to the nominal mass and momentum of the B meson, while p_ℓ, p_ν and p_X denote the momentum of the lepton, neutrino and the hadronic final state, respectively. The two variables are not independent: Figure 2.5 (left) shows the boundaries of the allowed region in the $q^2 - E_\ell$ plane for the specific case of $B \to D^*\ell\nu$ decays. At high q^2 most of the available energy is shared between the masses of the virtual W^* boson and the daughter hadron. As a result the W^* is produced nearly at rest, while the lepton and neutrino of the decay are produced back-to-back. Therefore, the daughter hadron receives a negligible momentum exchange. This scenario is known as the "zero recoil" configuration, corresponding to $q^2_{\text{max}} = (m_B - m_x)^2$, and is typically the most favourable for the formation of low-mass mesons. In contrast, the lepton and neutrino are produced in parallel at

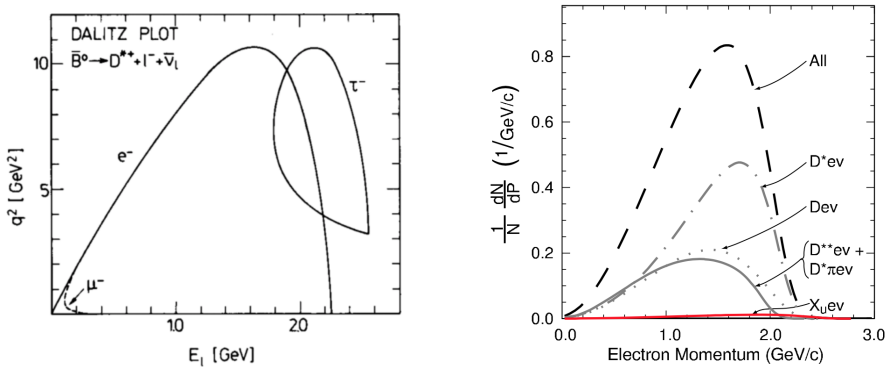

FIGURE 2.5: The allowed kinematic region in the $q^2 - E_\ell$ plane for $B \to D^* \ell\nu$ decays (left) [43] and lepton momentum spectra for $B \to X_c \ell\nu$ decays (right) [44].

a minimum value $q^2_{\text{min}} = m^2_\ell \approx 0$. For the formation of a bound state in this configuration, gluons must be exchanged to transfer momentum to the light degrees of freedom of the meson. Consequently, q^2_{min} is typically the least favoured configuration for the formation of a meson. Since each of the exclusive semileptonic decay modes have spectra with different end-points, different $B \to X_c \ell\nu$ decays will thus dominate in different regions of q^2 phase space. Figure 2.5 (right) shows the lepton momentum spectra for different $B \to X_c \ell\nu$ decays.

As mentioned in Section 2.1.2, four-velocities rather than momenta are used to describe the dynamics of systems containing heavy quarks. For B and $D^{(*)}$ mesons, the four-velocities are defined as:

$$v_B = \frac{p_B}{m_B}, \qquad v_{D^{(*)}} = \frac{p_{D^{(*)}}}{m_{D^{(*)}}}, \tag{2.40}$$

where $p_{D^{(*)}}$ and $m_{D^{(*)}}$ are the momentum and and nominal mass of the $D^{(*)}$ meson, respectively. In addition, the recoil variable w is defined as the scalar product of the 4-velocities of the B and $D^{(*)}$ mesons and is related to q^2 in the following manner:

$$w = v_B \cdot v_{D^{(*)}} = \frac{m^2_B + m^2_{D^{(*)}} - q^2}{2 m_B m_{D^{(*)}}}. \tag{2.41}$$

The point $w = 1$ corresponds to the maximum momentum transfer to the leptons $q^2_{\text{max}} = (m_B - m_{D^{(*)}})^2$, while the maximum value of w is reached at $q^2 = 0$.

Semileptonic $B \to D^* \ell\nu$ decays, where a vector D^* meson decays to two pseudoscalars, are described completely by four independent kinematic variables. The four variables most commonly used are q^2 or w along with the three angular variables, illustrated in Figure 2.6 and defined as follows:

- θ_ℓ: the angle between the direction of the lepton and the direction opposite the B meson in the virtual W rest frame.
- θ_v: the angle between the direction of the D^0 meson and the direction opposite the B meson in the D^* rest frame.

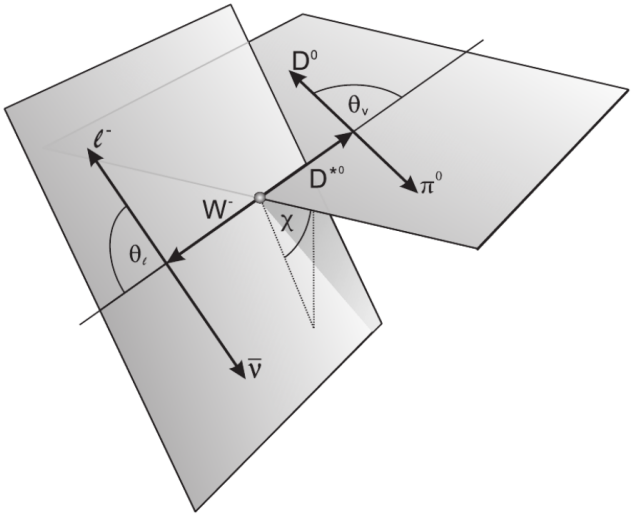

FIGURE 2.6: Definition of the angles θ_ℓ, θ_v and χ that characterise the $B \to D^* \ell\nu$ decay. These angles are used to describe any semileptonic decay in which the vector meson decays to two pseudoscalars [45].

- χ: the angle between the two planes formed by the decays of the W and the D^* meson, defined in the rest frame of the B meson.

2.2.2 Exclusive decays

The decays of $B \to D\ell\nu$ and $B \to D^*\ell\nu$ are the most frequent semileptonic B decays into a charm meson. Neglecting higher order electroweak corrections, the decay amplitude for these processes is determined by the tree-level matrix element of the local effective Hamiltonian in Eqn. 2.37 and can be expressed as:

$$\mathcal{M}(B \to D^{(*)}\ell\nu) = -i\frac{G_F}{\sqrt{2}} V_{cb} (\bar{\ell}\gamma^\mu P_L \nu_\ell) \langle D^{(*)}(p_{D^{(*)}}) | \bar{c}\gamma_\mu P_L b | B(p_B) \rangle \,. \tag{2.42}$$

The $V - A$ structure of the weak current dictates that only the vector current contributes for the decay of a pseudoscalar B meson to a pseudoscalar D meson final state, whereas both the vector and axial-vector current contribute for a vector D^* final state. Since the strong interaction between the quarks has to be taken into account for hadronic currents, the decay rate calculations are noticeably more complicated. The strong coupling becomes large for low momentum transfer, and thus higher order corrections have to be taken into consideration. As a result, analytical solutions are unavailable and the hadronic current is approximated by employing Lorentz invariant functions of q^2 known as hadron transition form factors. These functions describe how strong interactions modify the underlying weak decay, which require non-perturbative methods to calculate.

The treatment of the form factors is an ongoing theoretical research topic and different form factor parametrizations are available in the literature. Recent studies suggest that the underlying simplification assumptions introduced in these parametrizations should be re-evaluated, since the use of more generalised forms has hinted at higher observed values of $|V_{cb}|$ [46–49]. In addition, a fresh study aiming to address the tensions between inclusive and exclusive measurements of $|V_{cb}|$ computed the

form factors of $B \to D^* \ell\nu$ decays for nonzero values of the recoil parameter by utilizing lattice QCD [50].

Two of the most commonly used form factor parametrizations for $B \to D^{(*)}\ell\nu$ decays are discussed below, followed by a brief description of the orbitally excited D^{**} states. The form factor parametrizations for the $B \to D\ell\nu$ decay are described in [28, 51], while the parametrizations for the $B \to D^*\ell\nu$ decay are outlined in [28, 45, 52]. In addition, a comprehensive overview of the orbitally excited D^{**} states is presented in [24], while the form factor parametrization used in this analysis is proposed in [53].

$B \to D^{(*)}\ell\nu$ decays

For the $B \to D\ell\nu$ decay, the hadronic current can be decomposed in terms of the vector $f_+(q^2)$ and scalar $f_0(q^2)$ form factors in the following manner:

$$\langle D(p_D)|\bar{c}\gamma_\mu b|B(p_B)\rangle = f_+(q^2)\left((p_B+p_D)^\mu - \frac{m_B^2-m_D^2}{q^2}q^\mu\right) + f_0(q^2)\frac{m_B^2-m_D^2}{q^2}q^\mu\,. \tag{2.43}$$

In the limit of negligible lepton masses, the differential decay rate for $B \to D\ell\nu$ decays does not depend on $f_0(q^2)$ and can therefore to a very good approximation be described by only one form factor:

$$\frac{\mathrm{d}\Gamma(B\to D\ell\nu)}{\mathrm{d}w} = \frac{G_F^2|V_{cb}|^2 m_D^3}{48\pi^3}(m_B+m_D)^2(w^2-1)^{3/2}|\eta_{\mathrm{EW}}\mathcal{G}(w)|^2\,. \tag{2.44}$$

The form factor $\mathcal{G}(w)$ is given by

$$\mathcal{G}(w)^2 = \frac{4r}{(1+r^2)}f_+(w)^2\,, \tag{2.45}$$

where $r = m_D/m_B$ and η_{EW} is the electroweak correction calculated to be 1.0066 [54].

The parametrization of the $B \to D^*\ell\nu$ decay is more complicated, since the D^* vector meson decays to two pseudoscalars. Therefore, the hadronic current is described by four transition form factors: $A_i(q^2)$ with $i = 0, 1, 2$ and V, which can be expressed as follows:

$$\begin{aligned}
\langle D^*(p_{D^*},\epsilon_M)|\bar{c}\gamma_\mu b|B(p_B)\rangle &= \frac{2iV(q^2)}{m_B^2+m_{D^*}^2}\epsilon_{\mu\nu\alpha\beta}\epsilon^{*\nu}p_B^\alpha p_{D^*}^\beta\,,\\
\langle D^*(p_{D^*},\epsilon)|\bar{c}\gamma_\mu\gamma^5 b|B(p_B)\rangle &= 2m_{D^*}A_0(q^2)\frac{\epsilon^*\cdot q}{q^2}q_\mu + (m_B+m_{D^*})A_1(q^2)\left(\epsilon_\mu^* - \frac{\epsilon^*\cdot q}{q^2}q_\mu\right)\\
&\quad - A_2(q^2)\frac{\epsilon^*\cdot q}{m_B+m_{D^*}}\left((p_B+p_{D^*})_\mu - \frac{m_B^2+m_{D^*}^2}{q^2}q_\mu\right)
\end{aligned} \tag{2.46}$$

where $\epsilon_{\mu\nu\alpha\beta}$ is the Levi-Civita symbol and ϵ^* is the complex conjugated polarization vector of the D^* meson. Using the above equations, the form factors can subsequently be expressed in terms of helicity amplitudes in the following manner [55]:

$$H_\pm(q^2) = (m_B+m_{D^*})A_1(q^2) \mp \frac{2m_B}{m_B+m_{D^*}}|p_{D^*}|V(q^2)\,, \tag{2.47}$$

$$H_0(q^2) = \frac{1}{2m_{D^*}\sqrt{q^2}}\left((m_B^2 - m_{D^2}^2 - q^2)(m_B + m_{D^*})A_1(q^2) - \frac{4m_B^2|p_{D^*}|^2}{m_B + m_{D^*}}A_2(q^2)\right),$$
$$H_s(q^2) = \frac{2m_B|p_{D^*}|}{\sqrt{q^2}}A_0(q^2). \tag{2.48}$$

As in the previous case for $B \to D\ell\nu$ decays, the corresponding lepton current vanishes in the limit of negligible lepton masses, leaving only three helicity amplitudes: $H_\pm$ and H_0. Therefore, the differential decay rate can be written in terms of the four kinematic variables described in Section 2.2.1 as:

$$\begin{aligned}\frac{d\Gamma(B \to D^*(\to D\pi)\ell\nu)}{dw\,d\theta_\ell\,d\theta_\nu\,d\chi} = \frac{6m_B m_{m_{D^*}}^2}{8(4\pi)^2}G_F^2|V_{cb}|^2\sqrt{w-1}(1 - 2wr + r^2) \times \mathcal{B}(D^* \to D\pi)\\ \{(1-\cos\theta_\ell)^2\sin\theta_\nu{}^2 H_+^2(w) + ((1+\cos\theta_\ell)^2\sin\theta_\nu{}^2 H_-^2(w)\\ + 4\sin\theta_\ell{}^2\cos\theta_\nu{}^2 H_0^2(w) - 2\sin\theta_\ell{}^2\sin\theta_\nu{}^2\cos 2\chi H_+(w)H_-(w)\\ - 4\sin\theta_\ell(1-\cos\theta_\ell)\sin\theta_\nu\cos\theta_\nu\cos\chi H_+(w)H_0(w)\\ - 4\sin\theta_\ell(1+\cos\theta_\ell)\sin\theta_\nu\cos\theta_\nu\cos\chi H_-(w)H_0(w)\}\end{aligned} \tag{2.49}$$

The CLN parametrization

Up until recently, the form factor parametrization by Caprini, Lellouch and Neubert (CLN) [56] was the most commonly used approach. The CLN approach relies on the spin-flavour symmetry of HQET to derive relations between the form factors and reduce the number of free parameters. A general approach is to express the form factors as an expansion in the complex variable z defined as:

$$z(w) = \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}}, \tag{2.50}$$

while the coefficients of the expansion are subject to constraints from unitarity and analyticity [42]. Using this parametrization, the $B \to D\ell\nu$ form factor in Eqn. 2.45 can be expressed as:

$$\mathcal{G}(z) = \mathcal{G}(1)\left(1 - 8\rho_D^2 z + (51\rho_D^2 - 10)z^2 - (252\rho_D^2 - 84)z^3\right), \tag{2.51}$$

where form factor at zero recoil $\mathcal{G}(1)$ and the linear slope ρ_D^2 are the free parameters to be determined.

The remaining form factors describing $B \to D^*\ell\nu$ decays can be written in terms of a universal form factor $h_{A_1}(w)$ and two ratios $R_{1,2}(w)$ using the following relations derived from heavy quark symmetry,

$$A_1 = \frac{w+1}{2}r' h_{A_1}(w), \qquad A_2 = \frac{R_2(w)}{r'}h_{A_1}(w), \qquad V = \frac{R_1(w)}{r'}h_{A_1}(w), \tag{2.52}$$

where $r' = 2\sqrt{m_B m_{D^*}}/(m_B + m_{D^*})$. Subsequently, the helicity amplitudes can be expressed in terms of the universal form factor $h_{A_1}(w)$, the form factor ratios $R_{1,2}(w)$

and the slope parameter $\rho^2_{D^*}$:

$$\begin{aligned} h_{A_1}(w) &= h_{A_1}(1)\left(1 - 8\rho^2_{D^*}z + (53\rho^2_{D^*} - 15)z^2 - (231\rho^2_{D^*} - 91)z^3\right), \\ R_1(w) &= R_1(1) - 0.12(w-1) + 0.05(w-1)^2, \\ R_2(w) &= R_2(1) - 0.11(w-1) + 0.06(w-1)^2. \end{aligned} \tag{2.53}$$

Additional electroweak corrections are taken into consideration by introducing the transformation: $h_{A_1}(1) \to h_{A_1}(1)\eta_{\rm EW}$. Therefore, the CLN description parametrizes the differential decay rates of $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays by $\{G(1), \rho^2_D\}$, and $\{h_{A_1}(1)\eta_{\rm EW}, \rho^2_{D^*}, R_1(1), R_2(1)\}$. The values of these parameters are not calculated theoretically and must be extracted though an experimental analysis of the differential $B \to D^{(*)}\ell\nu$ spectrum.

The BGL parametrization

A more general parametrization proposed by Boyd, Grinstein and Lebed (BGL) [57] relies only on QCD dispersion relations. In this approach the form factors can be written as a series in powers of z. Thus, the form factors describing $B \to D\ell\nu$ using this parametrization take the form:

$$f_i(z) = \frac{1}{P_i(z)\phi_i(z)} \sum_{n=0}^{N} a_{i,n} z^n, \qquad i = +, 0. \tag{2.54}$$

Here, $P_i(z)$ are given by the "Blaschke factors" that contain the B_c or B_c^* poles, while $\phi_i(z)$ are the "outer functions". The free parameters of the expansion are $a_{i,n}$ and N denotes the order at which the series is truncated. By explicitly choosing the values to be $P_i(z) = 1$ with the outer functions,

$$\begin{aligned} \phi_+(z) &= 1.1213(1+z)^2(1-z)^{1/2}[(1+r)(1-z) + 2\sqrt{r}(1+z)]^{-5}, \\ \phi_0(z) &= 0.5299(1+z)^2(1-z)^{3/2}[(1+r)(1-z) + 2\sqrt{r}(1+z)]^{-4}. \end{aligned} \tag{2.55}$$

the unitarity bound on the coefficients $a_{i,n}$ takes the simple form

$$\sum_{n=0}^{N} |a_{i,n}|^2 \leq 1, \qquad \forall\, N. \tag{2.56}$$

The helicity amplitudes describing $B \to D^*\ell\nu$ decays are written slightly differently in the BGL approach and are given in terms of three form factors, i.e. $\mathcal{F}_1(w)$, $f(w)$ and $g(w)$:

$$\begin{aligned} H_0(w) &= \mathcal{F}_1(w)/\sqrt{q^2}, \\ H_\pm(w) &= f(w) \mp m_B m_{D^*}\sqrt{w^2-1}\, g(w). \end{aligned} \tag{2.57}$$

These three BGL form factors can, in turn, be written in a similar manner as the above-mentioned expansion,

$$\begin{aligned} f(z) &= \frac{1}{P_{1+}(z)\phi_f(z)} \sum_{n=0}^{N} a_n^f z^n, \\ \mathcal{F}_1(z) &= \frac{1}{P_{1+}(z)\phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{N} a_n^{\mathcal{F}_1} z^n, \\ g(z) &= \frac{1}{P_{1-}(z)\phi_g(z)} \sum_{n=0}^{N} a_n^g z^n. \end{aligned} \tag{2.58}$$

The outer functions ϕ_i for $i = g, f, \mathcal{F}$ are as follows:

$$\begin{aligned}
\phi_g(z) &= \sqrt{\frac{n_I}{3\pi\chi_1^T(0)}}\frac{2^4 r^2(1+z)^2(1-z)^{-1/2}}{[(1+r)(1-z)+2\sqrt{2}(1+z)]^4}\,,\\
\phi_f(z) &= \frac{4r}{m_B^2}\sqrt{\frac{n_I}{3\pi\chi_+^T(0)}}\frac{(1+z)(1-z)^{3/2}}{[(1+r)(1-z)+2\sqrt{2}(1+z)]^4}\,,\\
\phi_{\mathcal{F}_1}(z) &= \frac{4r}{m_B^3}\sqrt{\frac{n_I}{6\pi\chi_+^T(0)}}\frac{(1+z)(1-z)^{5/2}}{[(1+r)(1-z)+2\sqrt{2}(1+z)]^5}\,,
\end{aligned} \tag{2.59}$$

where $\chi_+^T(0)$ and $\chi_-^T(0)$ are additional constraints given in [52] and n_I represents the number of spectator quarks. A relation between the form factors $\mathcal{F}_1(0) = (m_B - m_{D^*})f(0)$ can be obtained by exploiting the dynamics at zero recoil ($w = 1$ or $z = 0$). As with the previous case, the coefficients in each of the expansions are subject to constraints from unitarity. Similar to the parameters of the CLN description, the values of the coefficients that parametrize the differential decay rates for $B \to D^{(*)}\ell\nu$ decays in the BGL description must also be determined by studying experimental data.

$B \to D^{**}\ell\nu$ decays

Understanding the modelling of heavy-light quark meson $Q\bar{q}$ systems as well as their mass spectra is important while studying semileptonic B decays, since these decays produce charmed mesons ($Q = c$) almost exclusively. A direct consequence of the spin-flavour symmetry that arises in the heavy quark limit $m_Q \to \infty$ is that these heavy-light bound systems resemble a similar configuration to that of hydrogen. Not only is the total angular momentum of the hadron **J** conserved, but also the spin of the heavy quark in the $m_Q \to \infty$ limit. Therefore, the spin of the light degrees of freedom, which is defined by

$$\mathbf{S}_\ell \equiv \mathbf{J} - \mathbf{S}_Q\,, \tag{2.60}$$

is also a conserved quantity in the heavy quark limit. As a result, the excited states of charmed mesons can be classified by spin-symmetry doublets containing states with total spin:

$$j_\pm = s_\ell \pm 1/2\,, \tag{2.61}$$

which is obtained by combining the spin of the light degrees of freedom with the spin of the heavy quark $s_Q = 1/2$. For a spin-parity $s_\ell = 1/2$, the two lightest states are the narrow $1S$ mesons: D and D^*, with total spin 0 and 1, respectively. The next highest orbitally excited mesons correspond to $L = 1$ excitations in the non-relativistic quark model. Here, L denotes the quantum number of the orbital angular momentum. The two states D_1 and D_2^* correspond to $s_\ell = 3/2$ with total spin 1 and 2. These resonances are narrow with widths of a few tens of MeV. Furthermore, two states correspond to $s_\ell = 1/2$: D_0^* and D_1^* with total spin 0 and 1. In contrast to the pair, these resonances are broad and have widths of a few hundred MeV. The mass spectrum of these orbitally excited 1P states, collectively referred to as D^{**}, is shown in Figure 2.7. Experimental knowledge of $B \to D^{**}\ell\nu$ decays is still rather poor, especially for the broad states. Since the branching fractions of D^{**} mesons themselves are not well known, the product of the $B \to D^{**}\ell\nu$ and $D^{**} \to D^{(*)}\pi$ branching fractions are generally quoted. Both Belle [58] and BaBar [59] performed

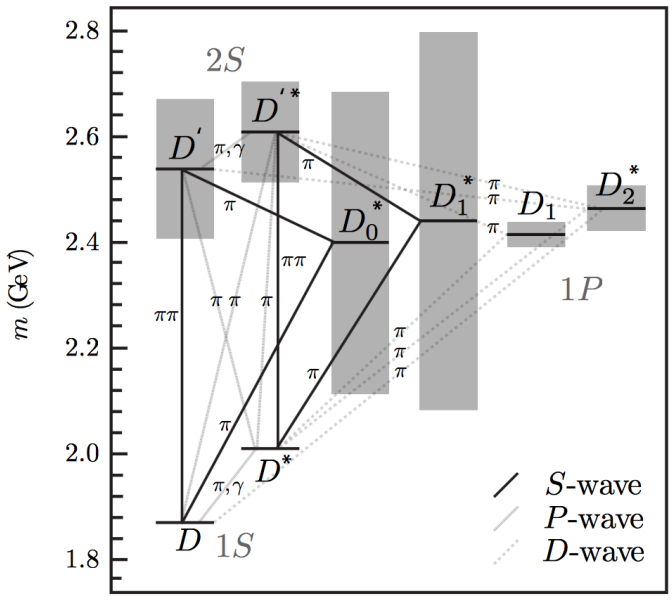

FIGURE 2.7: Masses and widths of the orbitally excited D^{**} mesons. The D_1 and D_2^* states have a narrow width, while the widths of the D_0^* and D_1^* states are much broader [60].

analyses of $B \to D^{(*)}\pi\ell\nu$ decays to select $B \to D^{**}\ell\nu$ candidates, where the D^{**} mesons decay specifically to $D^*\pi$. A concise summary of the experimental efforts to measure these decays is given in [42].

The decay kinematics of semileptonic $B \to D^{**}\ell\nu$ processes are investigated by Leibovich, Ligeti, Stewart and Wise (LLSW) [61]. In this model, the differential decay widths are given as a function of the recoil parameter and the angle between the charmed meson and the lepton, similar to the depiction given in Figure 2.6. To describe these decays, a new set of axial and vector form factors is introduced: f_i, $k_i(w)$, $g_\pm(w)$ and $g_i(w)$ where $i = V, A$. In addition, different sets of functions parametrizes the form factors describing the narrow and broad resonances. The hadronic matrix elements between the B meson and D_1 or D_2^* mesons can once again be separated into the vector and axial-vector currents:

$$\begin{aligned}
\frac{\langle D_1(v',\epsilon)|\bar{c}\gamma^\mu b|B(v)\rangle}{\sqrt{m_{D_1}}m_B} &= f_{V_1}\epsilon^{*\mu} + (f_{V_2}v^\mu + f_{V_3}v'^\mu)(\epsilon^*\cdot v)\,,\\
\frac{\langle D_1(v',\epsilon)|\bar{c}\gamma^\mu\gamma_5 b|B(v)\rangle}{\sqrt{m_{D_1}}m_B} &= if_A\epsilon^{\mu\alpha\beta\gamma}\epsilon^*_\alpha v_\beta v'_\gamma\,,\\
\frac{\langle D_2^*(v',\epsilon)|\bar{c}\gamma^\mu b|B(v)\rangle}{\sqrt{m_{D_2^*}}m_B} &= ik_V\epsilon^{\mu\alpha\beta\gamma}\epsilon^*_{\alpha\sigma}v^\sigma v_\beta v'_\gamma\,,\\
\frac{\langle D_2^*(v',\epsilon)|\bar{c}\gamma^\mu\gamma_5 b|B(v)\rangle}{\sqrt{m_{D_2^*}}m_B} &= ik_{A_1}\epsilon^{*\mu\alpha}v_\alpha + (k_{A_2}v^\mu + k_{A_3}v'^\mu)(e^*_{\alpha\beta}v^\alpha v^\beta)\,.
\end{aligned} \tag{2.62}$$

Similarly, the hadronic matrix elements for the broad states are written as:

$$\begin{aligned}
\frac{\langle D_0^*(v',\epsilon)|\bar{c}\gamma^\mu b|B(v)\rangle}{\sqrt{m_{D_0^*}}m_B} &= 0\,,\\
\frac{\langle D_0^*(v',\epsilon)|\bar{c}\gamma^\mu\gamma_5 b|B(v)\rangle}{\sqrt{m_{D_0^*}}m_B} &= g_+(v^\mu+v'_\mu)+g_-(v^\mu-v'_\mu)\,,\\
\frac{\langle D_2^*(v',\epsilon)|\bar{c}\gamma^\mu b|B(v)\rangle}{\sqrt{m_{D_2^*}}m_B} &= g_{V_1}\epsilon^{*\mu}+(g_{V_2}v^\mu+g_{V_3v'^\mu})(\epsilon^*\cdot v)\,,\\
\frac{\langle D_2^*(v',\epsilon)|\bar{c}\gamma^\mu\gamma_5 b|B(v)\rangle}{\sqrt{m_{D_2^*}}m_B} &= ig_A\epsilon^{\mu\alpha\beta\gamma}\epsilon^*_\alpha v_\beta v'_\gamma\,,
\end{aligned} \tag{2.63}$$

where v and v' are four-velocities of the B and D^{**} mesons, and the polarisation of the charmed meson is given by $\epsilon^{*\mu}$ or $\epsilon^{*\mu\alpha}$. Using the free-quark decay rate $\Gamma_0 = G_F^2|V_{cb}|^2m_B^5/(192\pi^3)$ and $r = m_{D^{**}}/m_B$, the double differential decay rates are given by:

$$\begin{aligned}
\frac{\mathrm{d}\Gamma_{D_1}}{\mathrm{d}w\mathrm{d}\theta} &= 3\Gamma_0 r^3\sqrt{w^2-1}\Big(\sin\theta^2[(w-r)f_{V_1}+(w^2-1)(f_{V_3}+rf_{V_2})]^2\\
&\quad+(1-2wr+r^2)[(1+\cos\theta^2)[f_{V_1}^2+(w^2-1)f_A^2]-4\cos\theta\sqrt{w^2-1}f_{V_1}f_A]\Big)\,,\\
\frac{\mathrm{d}\Gamma_{D_2^*}}{\mathrm{d}w\mathrm{d}\theta} &= \frac{3}{2}\Gamma_0 r^3(w^2-1)^{3/2}\Big(\frac{4}{3}\sin\theta^2[(w-r)k_{A_1}+(w^2-1)(k_{A_3}+rk_{A_2})]^2\\
&\quad+(1-2wr+r^2)[(1+\cos\theta^2)[k_{A_1}^2+(w^2-1)k_V^2]-4\cos\theta\sqrt{w^2-1}k_{A_1}k_V]\Big)\,,\\
\frac{\mathrm{d}\Gamma_{D_1^*}}{\mathrm{d}w\mathrm{d}\theta} &= 3\Gamma_0 r^3\sqrt{w^2-1}\Big(\sin\theta^2[(w-r)g_{V_1}+(w^2-1)(g_{V_3}+rg_{V_2})]^2\\
&\quad+(1-2wr+r^2)[(1+\cos\theta^2)[g_{V_1}^2+(w^2-1)g_A^2]-4\cos\theta\sqrt{w^2-1}g_{V_1}g_A]\Big)\,,\\
\frac{\mathrm{d}\Gamma_{D_0^*}}{\mathrm{d}w\mathrm{d}\theta} &= 3\Gamma_0 r^3(w^2-1)^{3/2}\sin\theta^2[(1+r)g_+-(1-r)g_-]^2\,.
\end{aligned} \tag{2.64}$$

Even though a large number of form factors was introduced within the above equations, this can be reduced to only a single form factor for the two spin-parity states, $\tau_{1/2}$ and $\tau_{3/2}$, in the heavy quark limit of an expansion in powers of $1/m_Q$ ($Q = c, b$). The complete expansions of the form factors for all four D^{**} states are omitted here for readability, however they are documented in their entirety in [53]. These form factors are still poorly understood; while an estimate of the contributions for the various D^{**} states may be obtained by constraining $\tau_{1/2}$ and $\tau_{3/2}$ with sum rules, significantly smaller decay rates are predicted than observed in current data. Further discussions and possible solutions for this puzzle are found in [60].

2.2.3 Inclusive decays

Rather than explicitly reconstructing a specific final state, an inclusive approach ignores all the details of the final hadronic state X_c and sums over all final states containing a c quark instead. Here X_c can be a single-particle state, like a D meson, or a multi-particle state, like $D\pi$ for example. The simplest approximation of a semileptonic $B \to X_c\ell\nu$ decay is to only consider the decay of the b quark, i.e. $b \to c\ell\nu$.

Using the theoretical framework of HQET to perform an operator product expansion, this approximation can be shown to be the leading result in the limit $m_b \to \infty$. Derivations of this result can be found in [24, 25, 38, 62], while key steps are summarised in the following section.

The mass of the b quark is of particular importance, since theoretical predictions of several quantities depend on this value. However, unlike leptons, quarks are confined inside physical hadrons and thus are not observed as physical particles. As a result, any information regarding their mass must be inferred through their influence on hadron properties. Therefore, several mass definitions employing different renormalization schemes can be used to describe the mass of the b quark. The difference between mass schemes can be determined as a series in powers of α_s. Transformations between some of the different schemes are possible by following simple guidelines. The two main schemes that are used in semileptonic $b \to c\ell\nu$ decays are the 1S and kinetic schemes, since these schemes are defined at an energy scale μ, below which HQET is valid. The kinetic scheme defines a μ-dependent mass from a sum rule for the non-relativistic kinetic energy of the heavy quark m_b^{kin}, where values of μ is below m_b. In the 1S scheme, the b quark is defined as half the perturbative contribution to the mass of a $J^{PC} = 1^{--}$ and $^{2s+1}L_J = {}^3S_1$ bound $b\bar{b}$ state. This bottomonium state is known as the $\Upsilon(1S)$ meson, for which precise measurements exist.

The Operator Product Expansion

Starting from the low energy effective Hamiltonian given by Eqn. 2.37, the triple differential decay rate can be written in terms of the momentum transfer squared of the virtual W boson q^2 and the lepton and neutrino energies E_ℓ and E_{ν_ℓ} as follows:

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2\mathrm{d}E_\ell\mathrm{d}E_{\nu_\ell}} = \frac{1}{4}\sum_{X_c}\sum_{\text{lepton spins}}\frac{|\langle X_c\ell\nu|H_{\text{eff}}|\bar{B}\rangle|^2}{2m_B}(2\pi)^3\delta^4(p_B - \underbrace{(p_\ell - p_{\nu_\ell})}_{=q} - p_{X_c})\,. \tag{2.65}$$

Since leptons do not interact strongly, the matrix element can once again be factorised into two parts: a complicated hadronic contribution $B \to W^* X_c$ and an easily calculable leptonic contribution $W^* \to \ell\nu$. Subsequently, Eqn. 2.65 can be further simplified,

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2\mathrm{d}E_\ell\mathrm{d}E_{\nu_\ell}} = 2G_F^2|V_{cb}|^2 W_{\alpha\beta}L^{\alpha\beta}\,, \tag{2.66}$$

with $W_{\alpha\beta}$ and $L^{\alpha\beta}$ denoting the hadronic and the leptonic tensor, respectively. The hadronic tensor depends on the B four-momentum $p_B = m_b v$ as well as the momentum transfer q, and is given by

$$W_{\alpha\beta} = \sum_{X_c}\frac{(2\pi)^3}{m_B}\delta^4(p_B - q - p_{X_c})\langle\bar{B}(p_B)|J_L^{\dagger\alpha}|X_c(p_{X_x})\rangle\langle X_c(p_{X_x})|J_L^{\beta}|\bar{B}(p_B)\rangle\,. \tag{2.67}$$

All strong interaction physics relevant for inclusive semileptonic B decays are parameterized by the hadronic tensor, which cannot easily be simplified and must be approximated using an Operator Product Expansion (OPE). The first step is to use the Optical Theorem to relate $W_{\alpha\beta}$ to the time-ordered product of currents $T_{\alpha\beta}$ such

that,

$$W_{\alpha\beta} = -\frac{1}{\pi}\mathrm{Im}T_{\alpha\beta} = -\frac{1}{\pi}\mathrm{Im}\int \mathrm{d}^4x e^{-iq\cdot x}\frac{\langle \bar{B}|T[J^\dagger_{L\alpha}(x)J_{L\beta}(0)]|\bar{B}\rangle}{2m_B}\,. \tag{2.68}$$

Next, the time-ordered product of currents can be expressed in terms of an expansion of matrix elements of local operators $\mathcal{O}_i$,

$$\int \mathrm{d}^4x e^{-iq\cdot x}T[J^\dagger_{L\alpha}(x)J_{L\beta}(0)] = \sum_{n,i}\frac{1}{m_B^n}\mathcal{C}_{n,i}\mathcal{O}_{n,i}\,. \tag{2.69}$$

Here, $\mathcal{C}_n$ represents a set of coefficients that can be confidently calculated in terms of α_s using perturbative methods in the limit of large momenta ($q \gg \Lambda_{\mathrm{QCD}}$). These coefficients are often referred to as Wilson coefficients. On the other hand, the matrix elements of the operators $\mathcal{O}_n$ encode all non-perturbative inputs into the decay rate. The normalized forward matrix elements of the local operators are defined as:

$$\langle O_n\rangle_B = \frac{\langle B|O_n|B\rangle}{2m_B}\,. \tag{2.70}$$

The goal is thus to systematically expand $\langle O_n\rangle_B$ for increasing powers of $1/m_B$ and increasing complexity, while applying the formalism of HQET. The transition is implemented by replacing the physical b quark fields by the corresponding heavy quark fields defined by the Lagrangian in Eqn. 2.14, resulting in the light degrees of freedom only becoming relevant for corrections at expansion orders of $1/m_b^2$ and higher. Details of the expansions will not be shown here, however, the main results are briefly summarised.

At leading order ($n = 0$) it can be shown that all operators of higher order corrections up to $\mathcal{O}(1/m_b)$ are related to the matrix element containing the operator $\langle\bar{b}b\rangle_B$, or even vanish completely due to parity invariance of the strong interaction. Therefore, there are no unknown matrix elements at leading order and C_0 yields the decay rate of the free b quark. All operators at order $n = 1$ can be written in terms of higher order operators using the equation of motion of the b quark. Consequently, there are no additional contributions of order $1/m_b$ to the decay rate of the free quark. The first non-trivial non-perturbative contributions appear at order $n = 2$ with an operator given by $\langle\bar{b}g_s\sigma_{\mu\nu}G^{\mu\nu}b\rangle_B$, which contains the gluon field. Using these operators, the semileptonic decay rate can be written as:

$$\Gamma(B\to X_c\ell\nu) = \frac{G_F^2}{192\pi^3}m_b^5\left(c_3\langle\bar{b}b\rangle_B + c_5\frac{\langle\bar{b}g_s\sigma_{\mu\nu}G^{\mu\nu}b\rangle_B}{m_b^2} + \mathcal{O}(1/m_b^3)\right)\,, \tag{2.71}$$

where the results of expanding the matrix elements in powers of $1/m_b$ are given by

$$\langle\bar{b}b\rangle_B = 1 + \frac{\lambda_1 + 3\lambda_2}{2m_b^2} + \mathcal{O}(1/m_b^3)\,, \tag{2.72}$$

$$\langle\bar{b}g_s\sigma_{\mu\nu}G^{\mu\nu}b\rangle_B = \frac{6\lambda_2}{m_b^2} + \mathcal{O}(1/m_b^3)\,. \tag{2.73}$$

These two matrix elements can be directly compared to Eqn. 2.16 and are found to be, in fact, the matrix elements of $\mathcal{L}_1$. Therefore, the two HQE parameters have a simple physical interpretation: λ_1 corresponds to the residual kinetic energy due to

the motion of the heavy quark, while λ_2 corresponds to the chromomagnetic moment of the heavy quark.

The total inclusive semileptonic decay rate is subsequently calculated by combining each of the individual contributions introduced in this section and can be written in compact form as follows:

$$\Gamma = \frac{G_F^2 m_b^5}{192\pi^3}|V_{cb}|^2 \left[1 + \frac{\lambda_1}{2m_b^2} + \frac{3\lambda_2}{2m_b^2}(2\rho\frac{\mathrm{d}}{\mathrm{d}\rho} - 3)\right] f(\rho)\,, \tag{2.74}$$

where the short distance coefficients take the form,

$$f(\rho) = 1 - 8\rho + 8\rho^3 - \rho^4 - 12\rho^2 \ln\rho\,, \qquad \text{and} \qquad \rho = \frac{m_c^2}{m_b^2}\,. \tag{2.75}$$

Here, the first term is the leading term in the $m_b \to \infty$ limit and yields the free quark decay rate. The next two terms are $1/m_b^2$ corrections, since the corrections at $1/m_b$ vanish. The HQE can be further utilised to calculate the next set of corrections to the total decay rate for order $1/m_b^3$, however, each higher order introduces additional operators and thus, matrix elements. Therefore, for calculations of $\mathcal{O}(1/m_b^3)$, two additional matrix elements of another two local operators arise. These matrix elements are discussed in the following section in the context of the kinetic scheme.

The kinetic scheme

The total inclusive decay rate shown in Eqn. 2.74 can also be expressed in terms of the kinetic operators up to order $1/m_b^3$ as [63]:

$$\Gamma = \frac{G_F^2 m_b^5}{192\pi^3}|V_{cb}|^2(1 + A_{\text{ew}})A_{\text{pert}}(r,\mu)\left[z_0(r)\left(1 - \frac{\mu_\pi^2(\mu) - \mu_G^2(\mu) + \frac{\rho_D^3(\mu)+\rho_{LS}^3(\mu)}{m_b(\mu)}}{2m_b^2(\mu)}\right)\right.$$
$$\left. - 2(1-r)^4\frac{\mu_G^2(\mu) - \frac{\rho_D^3(\mu)+\rho_{LS}^3(\mu)}{m_b(\mu)}}{m_b^2(\mu)} + d(r)\frac{\rho_D^3}{m_b^3} + \mathcal{O}(1/m_b^4)\right]\,, \tag{2.76}$$

where $z_0(r)$ is the same tree-level expression in Eqn. 2.75 and the expression $d(r)$ is given by

$$d(r) = 8\ln r + \frac{34}{3} + \frac{32}{3}r - 8r^2 + \frac{32}{3}r^3 - \frac{10}{3}^4\,. \tag{2.77}$$

The quantities A_{ew} and A_{pert} accounts for electroweak and perturbative corrections, respectively, and are calculated in [63]. The parameter μ is chosen to be 1 GeV, which sets the scale for short-distance dynamics in the OPE. In the kinetic scheme, the matrix operators corresponding to the kinetic and chromomagnetic operators are called μ_π^2 and μ_G^2. They are defined as:

$$2m_B\mu_\pi^2 \equiv -\langle B|\bar{b}_v(iD)^2 b_v|B\rangle\,,$$
$$2m_B\mu_G^2 \equiv \langle B|\bar{b}_v(iD^\mu)(iD^\nu)(i\sigma_{\mu\nu})b_v|B\rangle\,. \tag{2.78}$$

Corrections of order $1/m_b^3$ arise from the appearance of two additional operators, which are known as the Darwin and spin-orbital terms:

$$2m_B\rho_D^3 \equiv -\langle B|\bar{b}_v(iD_\mu)(ivD)(iD^\mu)b_v|B\rangle\,,$$
$$2m_B\rho_{LS}^3 \equiv \langle B|\bar{b}_v(iD^\mu)(ivD)(iD^\nu)(i\sigma_{\mu\nu})b_v|B\rangle\,. \tag{2.79}$$

For calculations of even higher order corrections ($n \geq 4$) the number of non-perturbative parameters proliferates, which complicates their extraction at order $1/m_b^4$ and higher. Therefore, accessing higher orders in the expansion requires the modelling of HQE parameters, which has been investigated in [64].

In order to determine $|V_{cb}|$ from inclusive decays, a simultaneous fit is performed to not only extract $|V_{cb}|$, but also the parameters of the HQE and the quark masses. The HQE parameters are sensitive to distributions of various kinematic variables such as the lepton energy, the hadronic mass or the hadronic energy spectrum of $B \to X_c \ell \nu$ decays. Therefore, the spectral moments of these inclusive distributions are measured to characterize the shapes of these variables, which can also be expressed in a similar manner as the total rate using the OPE. The expressions describing these observables are known [65–72]:

- to leading order in the HQE with next-to-next-to-leading order precision in α_s,
- up to the HQE of $\mathcal{O}(1/m_b^2)$ with next-to-leading order in α_s,
- up to the HQE of $\mathcal{O}(1/m_b^5)$ at tree-level with respect to the strong interaction.

By combining all available measurements of the total $B \to X_c \ell \nu$ rate, the lepton energy and the hadronic mass moments [7, 73–78], the non-perturbative HQE parameters and $|V_{cb}|$ can be determined by carrying out a global fit [79, 80]. A detailed overview of this method and various measurements of spectral moments from different experiments is given in [42].

2.3 Current status of $|V_{cb}|$ (and beyond!)

The current world averages of $|V_{cb}|$ from exclusive and inclusive determinations are [3]:

$$|V_{cb}^{\text{excl.}}| = (39.25 \pm 0.56) \times 10^{-3}, \tag{2.80}$$

$$|V_{cb}^{\text{incl.}}| = (42.19 \pm 0.78) \times 10^{-3}, \tag{2.81}$$

where the total uncertainty is the sum of both experimental and theoretical contributions. Both world averages exhibit a disagreement of approximately three standard deviations with one another. This disagreement between both measurements is limiting the reach of current searches for loop-level new physics in the CKM sector of the SM[2]. Recent measurements of exclusive $|V_{cb}|$ tend to agree better with the inclusive value, however the uncertainties of these values also tend to be larger. The current inclusive $|V_{cb}|$ average is measured with a relative precision of approximately 2%, which is dominated by the theoretical uncertainty including an additional 1.4% to account for missing higher order corrections [63, 82].

An alternative method for the extraction of inclusive $|V_{cb}|$ is proposed in [6]: by exploiting a symmetry within the HQE, known as reparametrization invariance (RPI)[3], the coefficients of different HQE operators can be linked forming linear combinations of the most general HQE parameters. Thus, the number of independent parameters for specific observables, especially total decay rates, are drastically reduced.

[2] See for example [81].

[3] A textbook example is given in [24].

The calculated reduced set involves only three parameters up to order $1/m_b^3$, while only five additional inputs are necessary once the contributions for order $1/m_b^4$ are included. New measurements are needed to determine this reduced set of eight parameters, as the key prerequisite that gives rise to the RPI in the total rate is violated in measurements of moments of lepton energy, hadronic mass, and hadronic energy spectra. For semileptonic $B \to X_c \ell \nu$ decays, the q^2 spectrum is found to be the unique observable that satisfies the RPI requirement. Since the lepton energy is not a RPI quantity, the reduction of parameters using the proposed strategy only works by measuring the moments of the q^2 distribution without a selection on the lepton energy, or by imposing progressively increasing requirements directly on the q^2 distribution. Furthermore, the lowest requirement should preferably be set as low as possible to not only ensure the selection of an inclusive $B \to X_c \ell \nu$ sample, but to also reduce the theoretical uncertainties. Combining measurements of the moments of the q^2 distribution with the alternative setup of the HQE would allow and extraction of $|V_{cb}|$ up to an order of $1/m_b^4$ without the modelling of higher orders, resulting in a complementary and fully data-driven determination.

Chapter 3

The Belle Experiment at the precision frontier

In order to investigate the SM in the search of a possible, more fundamental theory, a complementary strategy involving several approaches is carried out by physicists. First, the "energy frontier" experiments make use of powerful colliders to accelerate particles to unprecedented energies. The main goal of these experiments is to search for new, heavy particles or interactions that have thus far eluded detection. Next, at the "cosmic frontier" the universe itself serves as a laboratory by utilising satellites or ground based telescopes. Finally, the "intensity frontier" experiments create low energy, albeit intense, beams to search for deviations from the SM with high precision measurements, or search for new physics by studying rare processes.

An international collaboration known as the Belle Experiment played a key role in the intensity frontier by successfully establishing CP violation in the $B\bar{B}$ system. The main goal of this experiment was to study CP violation in the flavour sector as well as study rare or forbidden B meson decays. The Belle Experiment was conducted at the High Energy Accelerator Research Organisation (KEK), which is located in Tsukuba, Japan. An electron-positron asymmetric-energy accelerator known as KEKB allowed the collection of the world's largest sample of $\Upsilon(4S)$ energy data by the Belle detector during the course of its lifetime from 1999 to 2010. Subsequently, a total of 711 fb^{-1} integrated luminosity, corresponding to 772 million B mesons, was produced at KEKB leading to the aptly named B-factory. Belle's legacy as a B-factory was only recently superseded by its successor, Belle II, which broke the world record of the total collected instantaneous luminosity in June 2020 [83].

This chapter presents a brief overview of the KEKB accelerator facility, followed by a description of the Belle detector and its main detector sub-systems.

3.1 The KEKB Accelerator

The KEKB accelerator complex comprises an injector line and two storage rings with a circumference of 3 km. Furthermore, the injector line consists of two linear accelerators (LINAC), which are connected by a 180° bending magnet. A schematic overview of the accelerator infrastructure at KEKB is shown in Figure 3.1. The following section briefly describes the accelerator complex and is a summary of [84]. Additionally, an overview of the KEKB accelerator is also given in [85]. Furthermore, the subsequent section on luminosity is a summary of the concepts described in [86] and [87].

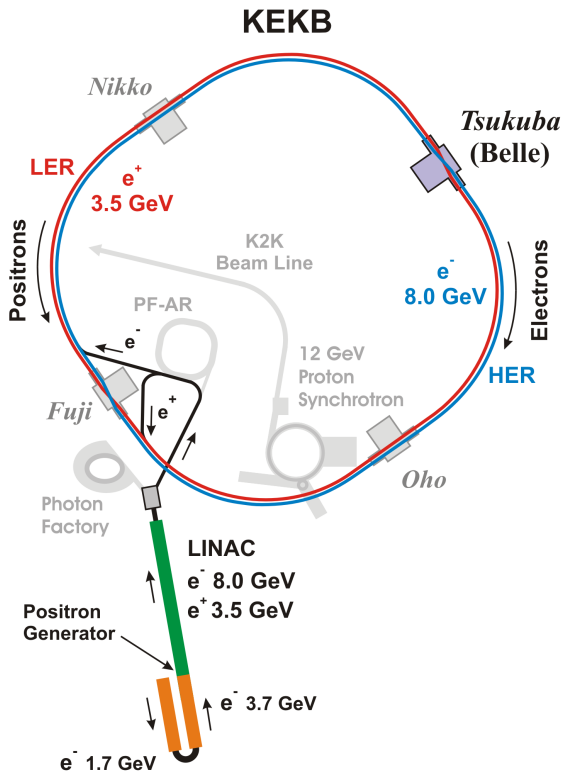

FIGURE 3.1: A schematic of the KEKB accelerator complex [88].

The process begins by producing electrons with an electron gun. In order to create positrons the electron beam with an energy of 3.7 GeV is targeted at a tungsten plate. Next, the electrons are accelerated to 8 GeV, while the captured positrons are accelerated to 3.5 GeV. Subsequently, the electron and positron beams are separated and injected into the High Energy Ring (HER) and Low Energy Ring (LER), respectively. The maximal current reached in each storage ring is 1.6 A for HER and 1.1 A for LER, while each beam consists of approximately 1584 individual bunches. These two particle beams follow opposing trajectories and may collide at a single point, known as the interaction point (IP), at a beam crossing angle of 11 mrad to reduce beam background.

Neglecting the electron mass, the components of the beam four-momentum can be written as follows:

$$p_H = \begin{pmatrix} E_H \\ 0 \\ E_H \sin\theta_{beam} \\ E_H \cos\theta_{beam} \end{pmatrix} \quad \text{and} \quad p_L = \begin{pmatrix} E_L \\ 0 \\ 0 \\ -E_L \end{pmatrix} \tag{3.1}$$

Therefore, using an invariant mass of $P = p_H + p_L$ the resulting centre of mass energy can be estimated as

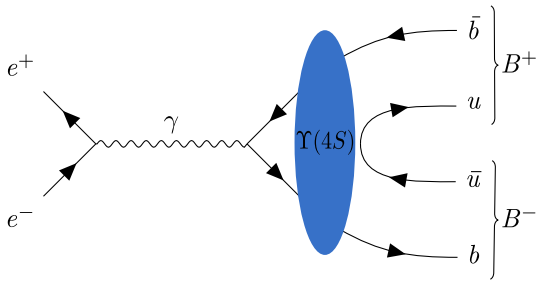

FIGURE 3.2: The Feynman diagram of the production and decay of the $\Upsilon(4S)$.

$$\begin{aligned}\sqrt{s} &= \sqrt{P_\mu p^\mu} \\ &= \sqrt{2E_H E_L(1+\cos\theta_{beam})} \\ &\approx 2\sqrt{E_H E_L} \\ &\approx 10.58\,\text{GeV}.\end{aligned} \tag{3.2}$$

Here, the resulting centre of mass energy corresponds to the $\Upsilon(4S)$ resonance: a $b\bar{b}$ state which decays via the strong force almost exclusively to a $B\bar{B}$ meson pair [89]. Figure 3.2 shows the Feynman diagram of the production and decay of the $\Upsilon(4S)$ resonance. Due to the asymmetry of the beam energies, the created final states experience a Lorentz boost of $\beta\gamma = 0.425$ in the forward direction with respect to the laboratory frame. Consequently, the average flight length of the two B mesons is increased by approximately 200 μm, which simplifies the separation of the two decay vertices.

The KEKB accelerator facility was not only designed to operate at the centre of mass energy of the $\Upsilon(4S)$ resonance, but also additional $\Upsilon(nS)$ resonances. Cross sections of different $\Upsilon(nS)$ resonances, originally measured by the CLEO collaboration [90], are given in Figure 3.3. In the context of this thesis, however, only the $\Upsilon(4S)$ resonance is of importance. Electron-positron collisions not only produce $\Upsilon(nS)$ resonances, but also various non-resonant processes. The largest of these background processes is light-quark production $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$ and is often referred to as "continuum". Further additional processes that might occur are: Bhabha scattering, tau and muon pair production and two-photon events.

Since the mass of the $\Upsilon(4S)$ is only marginally larger than the mass of the two B mesons, no additional particles are produced during the hadronization of the b quarks in the decay of the $\Upsilon(4S)$. Furthermore, in the centre of mass frame of the $\Upsilon(4S)$ system the two B mesons are produced back-to-back with a momentum of $p_B \approx 320\,\text{MeV}$ [42], i.e. nearly at rest production. This kinematic property is exclusive to the $\Upsilon(4S)$ resonance. Due to the precisely known initial state, coupled with the knowledge that each event will produce exactly two B mesons, this setup provides an experimentally clean environment to study B mesons decays with lower backgrounds than generally encountered at hadron colliders. B-factories are unique laboratories where rare B meson decays containing neutrinos in the final states may be precisely measured, due to the exact knowledge of initial and final states.

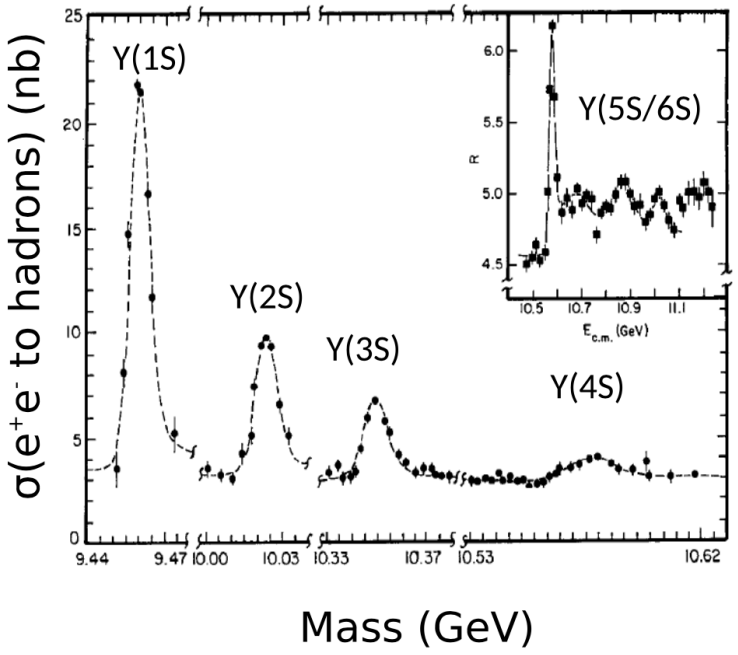

FIGURE 3.3: The scan of the first $\Upsilon(nS)$ resonances measured by the Cleo Collaboration [90].

Given a specific cross section for a process of interest, the rate of interactions produced per second can be described by:

$$\frac{dN}{dt} = L\sigma \, . \tag{3.3}$$

In this particular case dN/dt corresponds to the rate of e^+e^- collisions, while σ is the interaction cross section for $\Upsilon(4S)$ production at the $\Upsilon(4S)$ resonance energy. The instantaneous luminosity, denoted as L, refers to the number of collisions occurring at an interaction point of a particle collider per unit time, per cross sectional area, and is therefore useful to characterize the performance of an accelerator. For beam-beam collisions, the instantaneous luminosity can be expressed as follows:

$$L = fn\frac{N_1 N_2}{A} \tag{3.4}$$

where f is the rate per second that n number of bunches, consisting of N_1 and N_2 particles in opposing beams with overlapping area A collide. Assuming equal beam parameters and flat beams, this expression is re-parameterized at KEKB as [91]

$$L = \frac{\gamma_\pm}{2er_e}\frac{I_\pm \zeta_{y\pm}}{\beta^*_y \pm}\frac{R_L}{R_{\zeta_Y}} \, , \tag{3.5}$$

where γ, e and r_e are the Lorentz factor, the elementary electric charge and classical electron radius. The beam current, beam-beam parameter and the vertical beta function at the point of interaction are denoted as I, ζ_y and β^*_y, while the suffix $\pm$ indicates the electron $(-)$ or positron $(+)$ beam. Lastly, the parameters R_L and R_{ζ_Y} represent geometrical reduction factors for the luminosity and the beam-beam parameter.

The aim of the operation of a particle collider is to optimise the integrated luminosity, which is defined as the maximum achieved luminosity. The total integrated luminosity can be obtained directly from the instantaneous luminosity using the following equation [87]:

$$\mathcal{L}_{int} = \int_0^T L(t')dt' \tag{3.6}$$

where the integral is taken over the total amount of sensitive time, excluding possible periods of dead time. Rather than interrupting the facility's data taking operations to "fill" the two storage rings, the injector line continuously injects the rings with their respective electron and positron beams. Keeping the beam currents in the storage rings nearly constant during the experiment allows for nearly uninterrupted data acquisition with the Belle detector.

An accelerator's energy and luminosity are the most important figures in experimental accelerator physics [86]. The peak design luminosity of KEKB is of the order $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and this goal was not only reached in 2004, but even exceeded. A maximum peak luminosity of $L = 2.11 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ was achieved by the accelerator during its run period. With the cross section of $\Upsilon(4S)$ production given by $\sigma_{\Upsilon(4S)} = 1.2\,\mathrm{nb}$, a total of 25 $\Upsilon(4S)$ events were produced every second.

3.2 The Belle detector

The Belle detector is a multi-layered, general purpose detector designed to detect final state particles as well as their kinematics with high precision, allowing for precise investigations of B meson decays. It was designed to cover a large solid angle and is nearly a full 4π detector. Additionally, the detector consists of seven sub-detector systems that are specialised in the detection of a plethora of different types of particles. The main components of the Belle detector are the Silicon Vertex Detectors (SVD), the Central Drift Chamber (CDC), the Aerogel Cherenkov Counters (ACC), the Time of Flight Counters (TOF), the calorimeter systems, the muon and K_L detection system (KLM), and lastly the trigger system. A superconducting solenoid encompasses all sub-detector systems up to the KLM, providing a 1.5 T magnetic field. The nearly homogenous field covers a a cylindrical volume of 4.4 m in length and 3.4 m in diameter. Additionally, the iron structure of the Belle detector serves as a return path for the magnetic field, as well as absorber material for the KLM. A schematic of the Belle detector can be seen in Figure 3.4 which has been sliced in order for each of the main components to be visible.

This section will provide a brief summary of the main sub-detectors. The following sections are based on the comprehensive descriptions of the Belle detector and its sub-components found in [92].

3.2.1 Coordinate system

Within the Belle detector a right-handed coordinate system is used with the nominal IP taken as the origin: the x-axis points horizontally away from the IP, the y-axis points directly upwards and the z-axis points along the beam tunnel in the direction of the electron beam. The polar angle θ is subtended from the positive z-axis, while the azimuthal angle ϕ is subtended from the positive x-axis and lies in the $x-y$

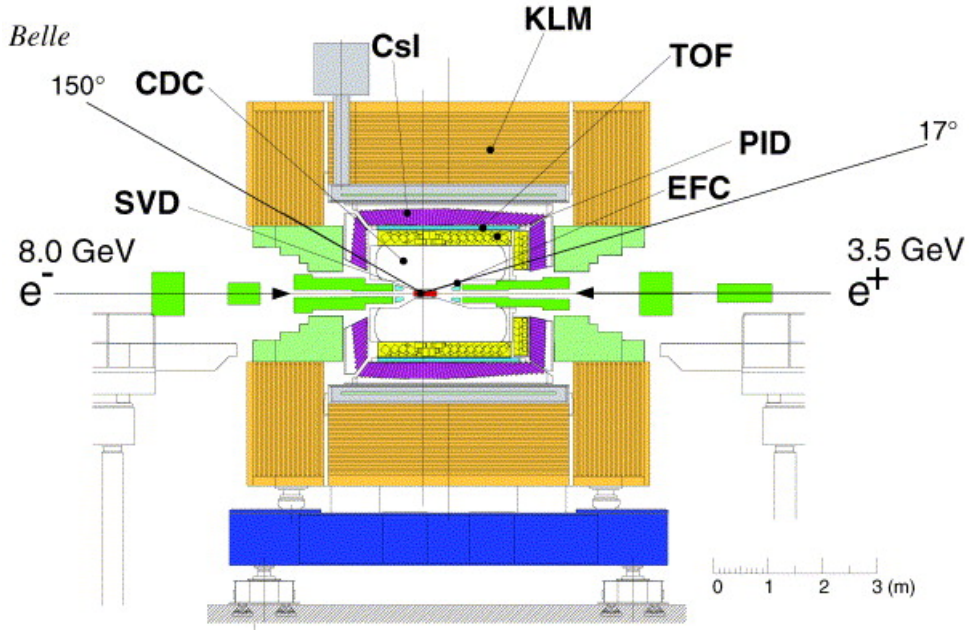

FIGURE 3.4: A sideview of the Belle detector, highlighting the main sub-components [92].

plane. Cylindrical coordinates are used in this plane, where the radius is measured from the origin as $r = \sqrt{x^2 + y^2}$.

3.2.2 Interaction Point

The beam pipe encloses the IP and maintains the accelerator vacuum. Since the main limiting factor of the resolution on the z-vertex position is multiple Coulomb scattering of particles originating in the beam pipe, the material budget at the interaction region is minimised. To this end, the material of the beam pipe is changed from aluminium to a double-wall beryllium cylinder. Each wall has a thickness of only 0.5 mm, while a 2.5 mm gap between the walls is constantly flushed with helium gas for cooling purposes. The cooling is crucial to the operation of the SVD, since the vertex detectors must be placed as close to the IP as possible to improve the vertex resolution. However, the close proximity to the beam pipe, which reaches beam-induced heating of several hundred watts, might damage the detectors without constant cooling.

3.2.3 Silicon Vertex Detector

Not only is the SVD designed to precisely measure the vertices of B meson decays, but it also aids the CDC in charged particle tracking. Two different designs of the SVD were deployed during the lifetime of the Belle detector. The first version, SVD1, covered a region of $\theta \in [23^\circ, 140^\circ]$ and was used until July 2003, after which it was replaced with the second version, SVD2, which covered a larger region of $\theta \in [17^\circ, 150^\circ]$.

Both versions of the SVD measure incident charged particles by employing double-sided silicon strip detectors (DSSD): a p-n junction that is operated under a reverse

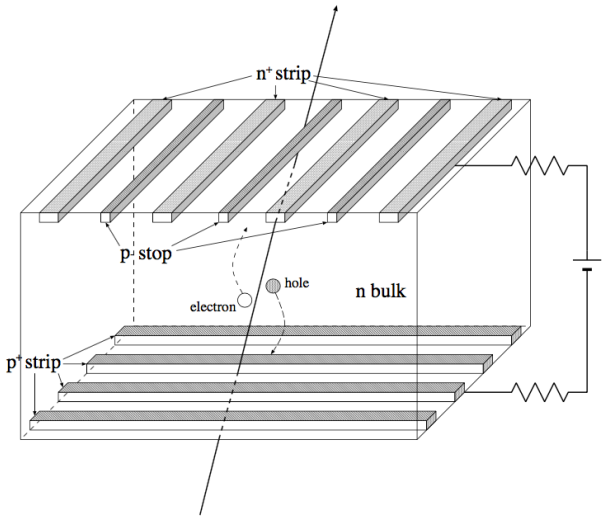

FIGURE 3.5: Schematic view of a DSSD [93].

bias of 75 V to a create full depletion zone. A charged particle traversing through the DSSD liberates electrons from the valence band into the conduction band, creating electron-hole pairs along its trajectory. These electron-hole pairs induce currents in the p^+ and n^+ strips that are situated along the surface of the bulk, on directly opposing sides of the DSSD. A DSSD has the dimensions of $57.5 \times 33.5 \times 0.3\,\mathrm{mm}^3$ and consists of 1280 sense strips with 640 readout pads on each side. An illustration of the operation of a DSSD is depicted in Figure 3.5.

The first version of the detector, SVD1, consists of three concentric layers arranged in a cylindrical barrel structure placed at radii of 30, 45.4 and 60.5 mm from the IP. Each layer in turn comprises 8, 10 and 14 ladders in the inner, middle and outer layers, respectively. Furthermore, each ladder contains two to four DSSDs. Therefore, the first version of the detector contains a total of 102 DSSDs together with a total of 81920 readout channels. On the other hand, the second version, SVD2, includes a fourth layer in order to allow for the reconstruction of charged tracks by making use of only SVD hits. Furthermore, the radius of the beam pipe is also reduced from 20 to 15 mm in order to place the SVD even closer to the IP. The layers of SVD2 are then located at radii of 20, 43.5, 70 and 88 mm. With the additional layer, SVD2 consists of 138 DSSDs and a total number of readout channels of 110592. A detailed description of the SVD upgrade can be found in [94].

3.2.4 Central Drift Chamber

The CDC is the principle component of the Belle tracking system: its main design purpose is to measure the trajectories of charged particles, as well as the specific ionisation energy loss, dE/dx. Due to the presence of the 1.5 T magnetic field, produced by the superconducting solenoidal coil, which completely immerses the CDC, charged particles follow a helicoidal trajectory when traversing through the detector. Additionally, the curvature of the particle's track is inversely proportional to the momentum of the particle.

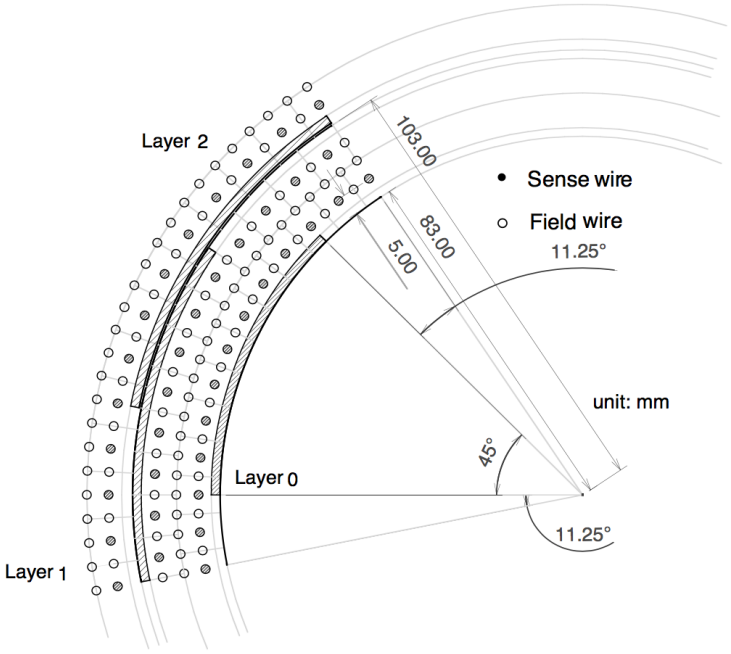

FIGURE 3.6: Cell structure in the cathode sector of the Central Drift Chamber [92].

Due to the asymmetric beam energies, the centre of mass system carries a non-zero momentum in the z direction. Consequently, the CDC is constructed asymmetrically as well and is divided into three geometric sections: the cathode, the conical-shaped inner and the toroidal-shaped outer, which ensures an angular coverage of $\theta \in [17^{\circ}, 150^{\circ}]$. The CDC consists of 32 axial layers and 18 small angle stereo layers each providing different spatial information of a charged particle's trajectory. The main purpose of the axial layers is to measure the $r - \phi$ position. On the other hand, stereo layers are inclined at a small angle with respect to the beam pipe and measure the z position in conjunction with the axial layers. The spatial resolution in $r - \phi$ is 130 μm, and is better than 2 mm in the z direction. Together with the three cathode strip layers the CDC contains a total of 8400 drift cells. Each drift cell, in turn, comprises a positively biased sense wire which is surrounded by eight negatively biased field wires strung along the beam direction. Figure 3.6 provides an illustration of the cell structure in the cathode section of the detector. In oder to minimise multiple Coulomb scattering, the cells are immersed in a 50% Helium and 50% Ethane gas mixture, which has a relatively long radiation length of 640 m. Additionally, the Ethane component of the mixture increases the specific ionisation energy loss resolution, which in turn improves the particle identification capabilities of charged particles.

A particle traversing through the CDC ionises the gas mixture along its path. The ionised electrons and positive ions then proceed to drift toward the anode and cathode sense wires, respectively, instigating further ionisations resulting in electron and positive ion avalanches. The electrons are collected by the sense wire and a hit is recorded by the CDC readout electronics. Due the presence of an homogenous electric field within the detector, the ionised particles accelerate until a saturation velocity is reached. At this point the energy gained from the electric field is equalled

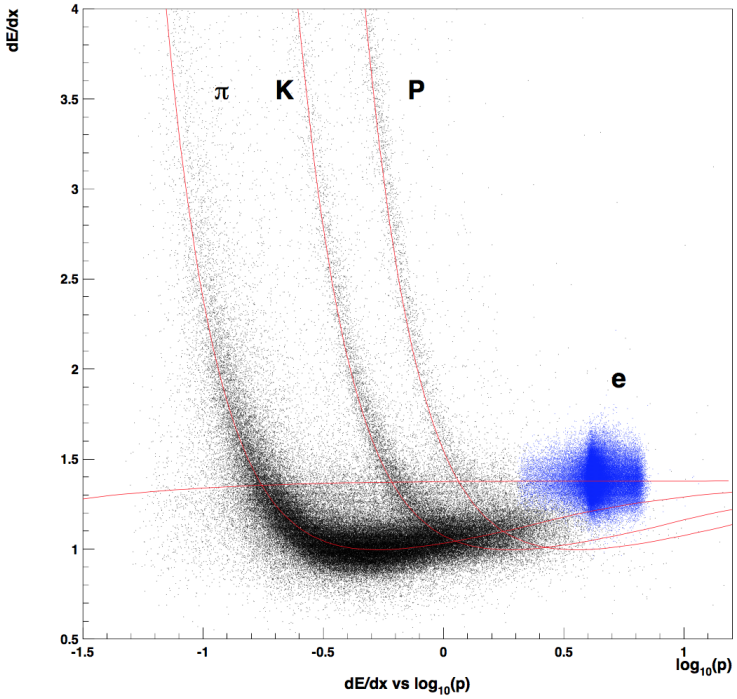

FIGURE 3.7: Truncated mean of dE/dx vs. momentum. The points are measurements taken during accelerator operations, while the lines are the expected distributions for each different particle type.

by the energy lost from interaction with the surrounded gas. Once this saturation velocity is reached, a linear correlation exists between the drift time of the ionised particle and its velocity. Therefore, the exact determination of the incident time directly leads to an accurate determination of a particle's position. The axial wires provide transverse momentum (p_T) information, while the small-angle-stereo layers provide z-axis information.

The hit amplitude recorded by the sense wire is used to determine the total energy of ionisation and, in turn, a charged particle's energy loss due to ionisation in the drift cell. Since the ionisation energy loss is dependent on the velocity of a particle at a given momentum, dE/dx varies with respect to the mass of the particle. For each CDC hit, the ionisation loss is measured along the trajectory and combined in order to determine the truncated average dE/dx of the track. The truncated mean dE/dx vs different particle's momenta observed in collision data is shown in Figure 3.7. The CDC distinguishes pions from kaons for momenta up to 0.8 GeV with a separation of 3σ.

3.2.5 Time of Flight Counters

The TOF measures the velocity of particles in a intermediate momentum range of 0.8 GeV to 1.2 GeV. Approximately 90% of the tracks produced in $\Upsilon(4S)$ decays fall into this momentum range. The incident particle's velocity is measured by determining its time-of-flight, as well as the flight length which is provided by the CDC's measurement of track parameters. The particle's velocity, in combination

with the momentum provided by the CDC, allows for the determination of the particle's mass, and therefore, the type.

The TOF is based on the concept of scintillation: the property of certain chemical compounds to emit short light pulses after excitation by the passage of charged particles or by photons of high energy. Consisting of 64 concentrically arranged modules as a radius of 1.2 m from the IP, the TOF covers a polar angle range of $\theta \in [34^\circ, 120^\circ]$. Each module comprises two trapezoid-shaped time of flight counters as well as one Trigger Scintillation Counter (TSC) separated by a 1.5 cm gap. Photomultiplier tubes placed directly on the TOF and TSC counters collect the scintillation light, which enables the sub-detector to measure the time of flight of a particle originating at the IP up until it's passage through the scintillator. Time intervals are measured within a precision of 100 ps. Primarily dedicated to the discrimination between kaons and pions, the subsystem's separation power is better than 3σ for momenta below 1.2 GeV.

3.2.6 Aerogel Cherenkov Counter

The final component of the particle identification system, the silica ACC extends the momentum coverage beyond the reach of the dE/dx measurements in the CDC and the time of flight measurements from the TOF. As suggested by the subsystem's name, the ACC's functionality relies on the Cherenkov principle: when a charged particle's velocity exceeds the speed of light in the medium it is traversing through, it emits Cherenkov radiation. As a result the emitted light appears as a coherent waveform at a fixed angle with respect to the particle's trajectory. The threshold energy for emitting Cherenkov photons can be related to a particle's velocity by,

$$|v| \geq \frac{c}{n} \tag{3.7}$$

where n is the refractive index of the medium. Therefore, selecting media with appropriate refractive indices allows for discrimination between pions and kaons. For the ACC, the refractive index was chosen such that pions with momenta between 1.2 and 3.5 GeV produce Cherenkov radiation, while kaons within the same momentum range do not.

The ACC is divided in two distinct regions, namely the barrel and the endcap, and is designed to cover a polar angle range of $\theta \in [17^\circ, 127^\circ]$. While the barrel contains 960 counter modules segmented into 60 cells in the ϕ direction, the forward endcap contains 228 counter modules arranged into five concentric layers. An ACC module consists of five Aerogel tiles that are stacked inside a thin aluminium box. In order to allow for efficient detection of Cherenkov light, photomultiplier tubes are attached directly to the aerogel boxes.

Since pions are some of the most ubiquitous particles in hadronic events, the ACC's performance is determined by its ability to identify kaons amongst pions. The ACC provides good K/π separation, with a kaon efficiency of 73% and a pion-to-kaon fake rate of 7% [95].

3.2.7 Calorimeters

A calorimeter system provides energy and position measurements of electrons and photons originating from B meson decays. The Belle calorimeter system consists

of two main calorimeters: the electromagnetic calorimeter (ECL) and the extreme forward calorimeter (EFC). Both calorimeters make use of a sampling methodology based design, in which the detectors consist of absorber and active materials. Absorber materials are usually of high density and aid in particle shower development, while the active materials collect and measure the deposited energy.

The main purposes of the ECL is the detection of photons, as well as the identification of electron tracks. It consists of a highly segmented array of 8736 Caesium Iodide crystals doped with Thallium (CsI(TI)). Since the Thallium shifts the excitation light to the visible spectrum, a pair of silicon photodiodes are placed at the rear of each crystal to collect the light. A typical crystal is shaped like a tower with a length of 30 cm, which is equivalent to 1.62 radiation lengths (X_0) for electrons and photons. This length was specifically chosen in order to minimise energy resolution deterioration at high energies due to shower leakage at the back of the crystal. Additionally, the crystals are designed such that a photon entering the centre of a crystal would deposit 80% of its energy in that crystal. The crystals are arranged in three sections: the forward end cap ($\theta \in [12.4^\circ, 31.4^\circ]$), the barrel ($\theta \in [32.2^\circ, 128.7^\circ]$), and the backward end cap ($\theta \in [130.7^\circ, 155.1^\circ]$). A crystal in the barrel region has a forward and backward face measuring an area of $55 \times 55\,\mathrm{mm}^2$ and $65 \times 65\,\mathrm{mm}^2$, while crystals in the forward and backward end caps vary from $44.5 - 70.8$ mm and from $54 - 82$ mm, respectively. In order to prevent particles from escaping through gaps between the crystals, the barrel region's crystals are slightly tilted at an angle of approximately 1.3° in the θ and ϕ directions. The ECL has the capability to measure photon energies up to a lower limit of 50 MeV. Additionally, since pions deposit less energy in the crystals than electrons, the ECL plays a crucial role in distinguishing electrons tracks from hadrons. The rate of mis-identifying a pion as an electron is less than 1% for momenta above 2 GeV.

The EFC extends the range of photon and electron calorimetry energy measurements to the extreme forward and backward regions. These regions correspond to a polar angle of $\theta \in [6.4^\circ, 11.5^\circ]$ in the forward direction and $\theta \in [163.3^\circ, 171.2^\circ]$ in the backward region. Additionally, the EFC is also instrumented as a beam monitor for KEKB, as well as a luminosity monitor for Belle. It is located on the front faces of the KEKB accelerator compensation solenoid magnet cryostats and completely surrounds the beam. Due to its proximity to the IP, the EFC is exposed to a high dose of radiation. Consequently, the EFC is constructed from radiation hard Bismuth Germanate in order to ensure a long lifetime of the sub-detector system. For both forward and backward cones, the detector is separated into 32 azimuthal and 5 polar sections. Each crystal is tower shaped and is tilted towards the IP.

3.2.8 KLM Detector

Located in the outermost layer of the Belle detector, the KLM is designed to identify K_L mesons and muons with high efficiency for momenta greater than 600 MeV. Long-lived neutral particles like K_L mesons originating from the IP will typically traverse undetected through the inner sub-components of the detector before reaching the KLM. When a K_L meson interacts with matter, a shower of ionising particles is created. The KLM induces these showers by providing a minimum of 3.9 interaction lengths, while the ECL provides an additional 0.8 interaction lengths. While the shower position enables the determination of the flight direction of the K_L, fluctuations in the total shower size prevents an accurate energy measurement.

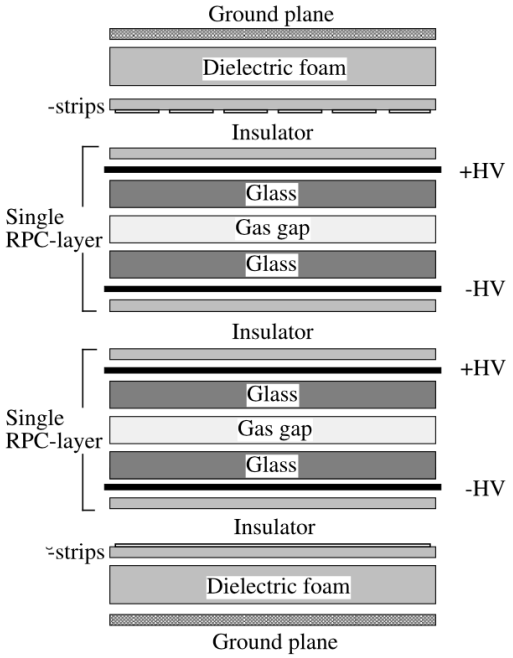

FIGURE 3.8: Cross section of a KLM super layer.

On the other hand, since muons neither interact strongly nor suffer from significant Bremsstrahlung radiation loss, given sufficient energy these particles will pass through the KLM. Therefore, any track matched with a particle traversing several layers of the KLM may be identified as a muon. Furthermore, muons may be separated from charged hadrons, particularly $K^{\pm}$ and $\pi^{\pm}$, by comparing the energy deposited by the charged tracks along their paths. In contrast to strongly interacting particles, muons deposit a smaller fraction of their energy prior to reaching the KLM. Muons with momenta below 500 MeV do not reach the KLM and must be discarded.

In order to detect the passage of charged particles, the KLM makes use of Resistive Plate Chambers (RPC). As their name suggests, RPCs consist of two resistive plates separated by a gas filled gap. This gap between the plates is held at an uniform electric field. A typical detector layer consists of a super layer of two glass-electrode RPC modules, interleaved between high voltage biased plates, insulators and external pickup strips. The layout of a super layer is shown in Figure 3.8. The entire KLM comprises alternating layers of detector super layers and thick iron plates with a thickness of 4.7 cm. The barrel region is octagonally shaped and consists of 15 detector layers and 4 iron layers, while the forward and backward regions consist of 14 detector layers each. In total the barrel and end cap regions contain 240 and 122 RPC modules, respectively.

An ionising particle traversing the gas filled gap in the single layer RPC induces a streamer in the gas which results in a local discharge of the plates. The discharge, in turn, induces a signal on the external orthogonal pickup strips located on either side of the pair of RPCs. These pickup strips, which are typically 5 cm wide, provide $\phi - z$ and $\theta - \phi$ information in the barrel end cap regions, respectively. The KLM provides a total coverage of $\theta \in [20^{\circ}, 155^{\circ}]$ and an angular resolution from the IP that is better

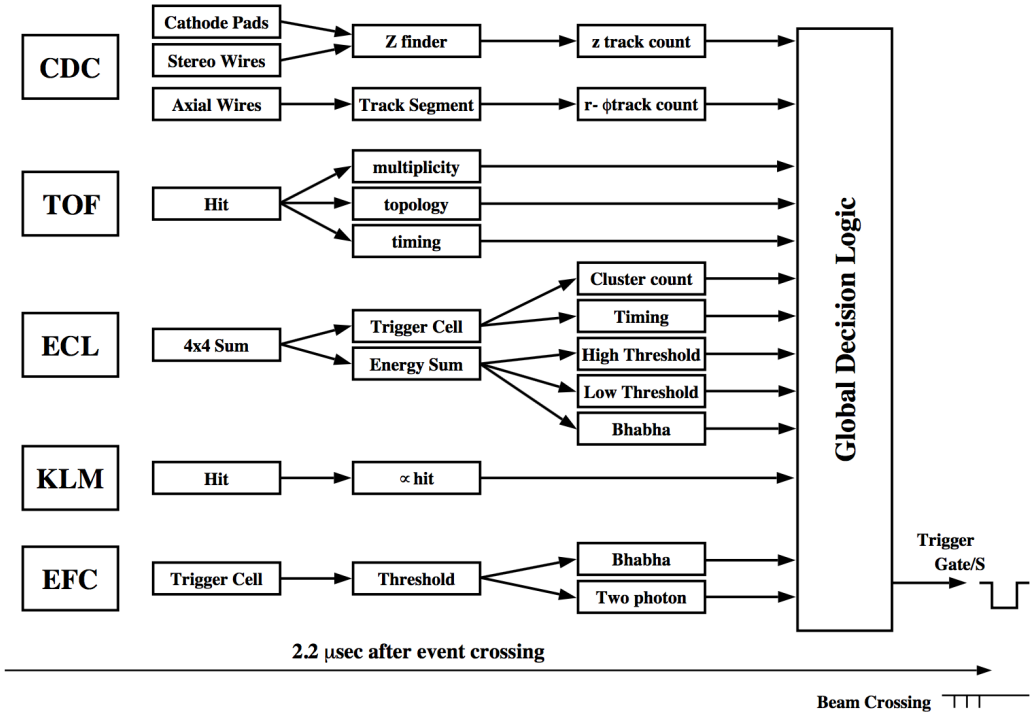

FIGURE 3.9: The Level-1 trigger system for the Belle detector.

than 10 mrad. Additionally, the muon identification efficiency is greater than 90% for momenta greater than 1.5 GeV, with a mis-identification rate smaller than 5%.

3.2.9 Trigger and Data Acquisiton System

In the Belle experiment, the main interesting physics events are considered to be $e^+e^- \to \Upsilon(4S)$, $e^+e^- \to c\bar{c}$, $e^+e^- \to \tau\bar{\tau}$ processes. Additionally, events of interest also include two photons ($e^+e^- \to \gamma\gamma$), Bhabha ($e^+e^- \to e^+e^-$) and μ-pair processes, which are used for luminosity measurements as well as detector calibration. The main sources of background events stem from interactions between the beams and residual gas molecules in the beam pipe, synchrotron radiation, and cosmic ray events. At an instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ the trigger rate for physics events of interest is around 100 Hz, while the trigger system maintains a total trigger rate of 200 Hz. Since the rate of beam background events is sensitive to the actual accelerator conditions, it is difficult to calculate a reliable estimate. The Belle trigger system consists of the Level-1 (L1) hardware trigger and the Level-3 (L3) and Level-4 (L4) software triggers.

An overview of the L1 trigger system is shown in Figure 3.9. The L1 trigger consists of sub-trigger systems that are governed by a single central trigger system called the Global Decision Logic (GDL). The GDL receives information from the individual trigger systems of each of the detector sub-systems within 1.85 μs after a collision event in order to issue a global trigger decision 2.2 μs after a collision. Sub-detector triggers are based on track or cluster information. The CDC provides $r-\phi$ and $r-z$ track signal trigger signals, while the EFC triggers on Bhabha and two-photon events. The ECL provides two complementary triggers based on the total energy

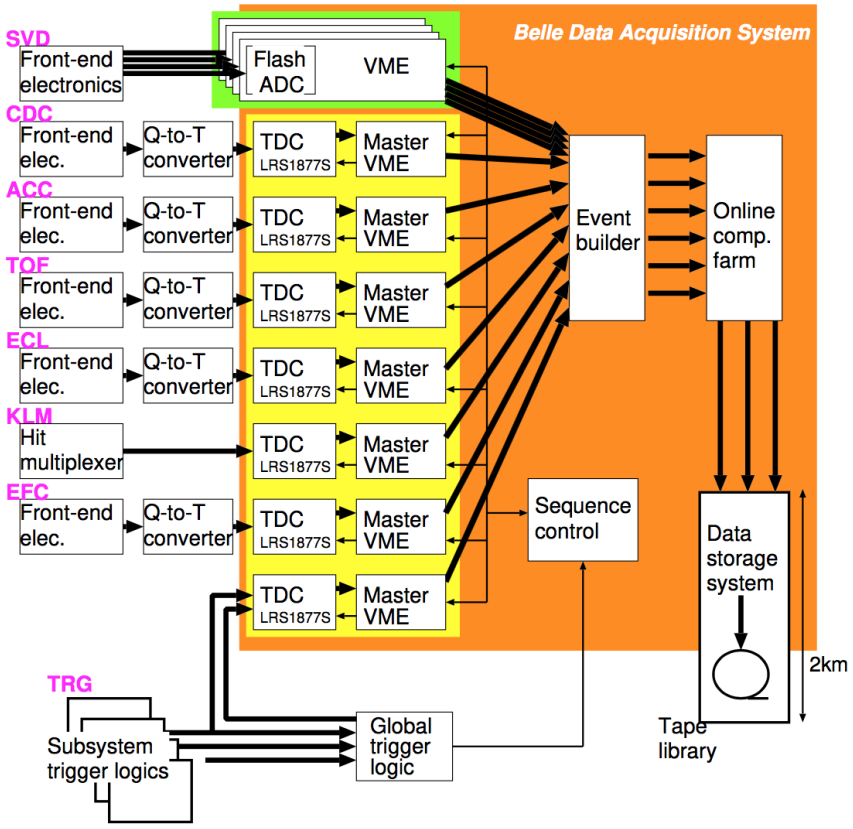

FIGURE 3.10: Overview of the Belle DAQ system.

deposits and cluster counting of crystal hits. Furthermore, the KLM provides a high efficiency trigger for muon tracks. Timing information of the event is provided by the TOF, however when unavailable the ECL is used for the trigger timing. The combined efficiency of the L1 trigger is 99.5%. A more detailed description can be found in [96].

Once the sequence control (SEQ) receives a trigger decision from the GDL, the sub-detector information from the L1 triggered events is collected by the Data Acquisition (DAQ) system. An overview of the DAQ system is shown in Figure 3.10. The system is segmented into 7 sub-systems dedicated to each of the different sub-detectors and running in parallel. An event-builder combines data from each sub-system into a single event record. The resulting output is transferred to an online PC farm, where the raw data is converted into offline data format and processed by the L3 trigger. The purpose of the L3 trigger is to further reduce background events. A fast track fitting algorithm is used to find tracks originating from the IP. It requires at least one track with an impact parameter of less than 5 cm and the total energy deposit in the ECL to be greater than 3 GeV. The L3 trigger retains physics events of interest with 99% efficiency, while reducing the total event rate by approximately $40\% - 50\%$ [97].

At the next step of the process, the L4 trigger is applied to the raw data. To pass this filter, events are required to have at least one track with $p_T > 300$ MeV and an

TABLE 3.1: Summary of the data sets recorded with the Belle detector at KEKB [100].

Resonance	On-peak luminosity (fb^{-1})	Off-peak luminosity (fb^{-1})	Number of resonances
$\Upsilon(1S)$	5.7	1.8	102×10^6
$\Upsilon(2S)$	24.9	1.7	158×10^6
$\Upsilon(3S)$	2.9	0.25	11×10^6
$\Upsilon(4S)$	711.0	89.4	772×10^6 $B\bar{B}$
$\Upsilon(5S)$	121.4	1.7	7.1×10^6 $B_s\bar{B}_s$

impact parameter in r and $|z|$ to be less than 1 cm and 4 cm, respectively. Furthermore, a total energy deposit in the ECL greater than 4 GeV is also a requirement. Events passing these selection criteria are converted from raw data to fully reconstructed data and are stored in files known as Data Summary Tapes (DST) in the PANTHER [98] file format.

The large quantities of data are further reduced to isolate specific physics processes of interest and stored in Mini-DST (MDST) files. The analysis presented in this work makes use of a skim, which imposes loose requirements on B meson events. The production of the MDST files as well as the analysis of Belle data are carried out in the Belle Analysis Framework (BASF) [99].

3.3 Recorded data set

The Belle detector recorded a data sample of $\mathcal{L} \approx 1\,\text{ab}^{-1}$ during its successful 10-year run period. While the largest fraction of the data sample was recorded at the $\Upsilon(4S)$ resonance, data at the other $\Upsilon(nS)$ resonances were also recorded. In order to study non-resonant contributions, smaller data samples were recorded at energies below the energy of each of the resonances and are referred to as "off-resonance" data samples. A summary of the collected data sets is given in Table 3.1.

To carry out precise measurements of rare semileptonic B meson decays, or to properly investigate possible physics processes beyond the SM, it is crucial to compare data measured by the Belle detector subsystems to theoretical predictions. This can be achieved by simulating data samples based on the Monte Carlo (MC) procedure –hence the name of MC samples for simulated data [101]. These MC samples can also be used to reflect the response and performance of the current detector, by comparing well-understood predictions of SM processes with the estimations produced from full detector simulations.

In order to generate MC samples two key components are required: an event generator and a detector response simulator. Belle makes use of the EvtGen [102] event generator to simulate events of $\Upsilon(4S)$ decays, while hadronic continuum events are produced with PYTHIA [103]. Event generators create a list of particles from e^+e^- interactions, as well as subsequent decays, according to a decay table. This decay table contains information regarding the decay models, modes, branching fractions,

lifetimes, etc., of all possible particles involved in the decay of the $\Upsilon(4S)$ and subsequent decays. The generated list, containing information such as particle's four vector momentum, is passed on to the detector response simulator. In order to simulate particle interactions with the detector, GEANT3 [104] is employed to model the geometry of the detector and particle interactions with matter. Furthermore, electromagnetic final-state radiation, where photons are radiated into the final state $B \to X\ell\nu\gamma$, is simulated using the PHOTOS package [105, 106].

Since the existing Belle MC samples were generated circa 2010, the branching ratios and theoretical models of key processes are outdated. Furthermore, various particle properties need to be updated to the latest values for more accurate modelling of both signal and background processes. The details of the updated or, in some cases, newly produced MC samples are described in the following sections.

3.3.1 Monte Carlo samples

The official MC samples provided by the Belle Collaboration are produced in several sets. These sets are referred to as "streams" and correspond to a total integrated luminosity equivalent to that of the data sample recorded by the Belle detector. The advantages of incorporating several streams for an analysis can range from overcoming statistical limitations to the validation of analysis and fitting procedures. Additionally, the produced MC samples are generated to be dependent on specific experiment numbers to take into account variable detector performance and beam conditions for each experimental run. This section describes the simulated signal and background processes that are used in this analysis. The modelling of $B \to D^{**}\ell\nu$ decays, non-resonant $B \to X_c\ell\nu$ processes, and $B \to X_u\ell\nu$ decays is treated identically to the detailed description found in [107].

Generic: This sample contains all possible processes that may occur in an $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ event, where the B meson decays via the dominating quark-level transition $b \to cW$. Therefore, both $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays, which form the largest signal contribution in the analysis, are present in this sample. The generic MC sample consists of a total of ten streams that is separated into decays of neutral B^0 mesons, and charged $B^\pm$ mesons.

$\boldsymbol{B \to D^{**}\ell\nu}$: Since the particle masses and widths of the four orbitally excited charmed mesons (Section 2.2.2) are outdated in the existing generic MC, new samples are generated using the values from [30] to replace the $B \to D^{**}\ell\nu$ events in the generic sample. The branching fraction values are adopted from [3] and are corrected to account for missing isospin-conjugated decay modes by following the prescription proposed in [53]. The branching fractions of the subsequent D^{**} decays to final states containing charged and neutral pions are also considered in the event simulation. As previously discussed, current available measurements only target $D^{**0} \to D^{(*)+}\pi^-$ decay modes. Therefore, the missing isospin modes are taken into account with a factor of

$$f_\pi = \frac{\mathcal{B}(\overline{D}^{**0} \to D^{(*)-}\pi^+)}{\mathcal{B}(\overline{D}^{**0} \to \overline{D}^{(*)}\pi)} = \frac{2}{3}. \tag{3.8}$$

Furthermore, the measurements of the $B \to D_2^*\ell\nu$ process given in [3] are converted to be only with respect to the $\overline{D}_2^{*0} \to D^{*-}\pi^+$ final states, while

$\overline{D}_2^{*0} \to D^- \pi^+$ contributions are taken into account by applying a factor of of [30]

$$f_{D_2^*} = \frac{\mathcal{B}(\overline{D}_2^{*0} \to D^- \pi^+)}{\mathcal{B}(\overline{D}_2^{*0} \to D^{*-} \pi^+)} = 1.54 \pm 0.15 \,. \tag{3.9}$$

The world average of the $B \to D_1^* \, \ell\nu$ decay is given by a combination of several measurements in [3], however poor agreement is observed. Notably, the measurement of [58] is in conflict with the measured branching fractions of [59] and [74]. The conflicting measurement is thus excluded and a new average is calculated for this analysis:

$$\mathcal{B}(B^+ \to \overline{D}_1^{*0} (\to D^{*-} \pi^+) \, \ell^+ \nu_\ell) = (0.28 \pm 0.06) \times 10^{-2} \,. \tag{3.10}$$

This value deviates slightly from the average given in [30], which not only excludes the measurement of [58], but also [74]. The world average of $B \to D_1 \, \ell\nu$ does not consider contributions from three-body decays of the form $D_1 \to D\pi\pi$. These final states are taken into account by making use of a factor [108]:

$$f_{D_1} = \frac{\mathcal{B}(\overline{D}_1 \to D^{*-} \pi^+)}{\mathcal{B}(\overline{D}_1 \to \overline{D}^0 \pi^+ \pi^-)} = 2.32 \pm 0.54 \,. \tag{3.11}$$

In addition, the contribution of $D_1 \to D\pi\pi$ is subtracted from the measured non-resonant plus resonant $B \to D\pi\pi\ell\bar{\nu}$ branching fraction of [109]. In order to account for the missing isospin-conjugated modes of the three-hadron final states, the prescription from [109] is adopted. An average isospin correction factor is quoted to be:

$$f_{\pi\pi} = \frac{\mathcal{B}(\overline{D}^{**0} \to \overline{D}^{(*)0} \pi^+ \pi^-)}{\mathcal{B}(\overline{D}^{**0} \to \overline{D}^{(*)} \pi\pi)} = \frac{1}{2} \pm \frac{1}{6} \,. \tag{3.12}$$

Furthermore, the following constraints are implicitly assumed:

$$\begin{aligned} \mathcal{B}(\overline{D}_2^* \to \overline{D}\pi) + \mathcal{B}(\overline{D}_2^* \to \overline{D}^*\pi) &= 1 \,, \\ \mathcal{B}(\overline{D}_1 \to \overline{D}^*\pi) + \mathcal{B}(\overline{D}_1 \to \overline{D}\pi\pi) &= 1 \,, \\ \mathcal{B}(\overline{D}_1^* \to \overline{D}^*\pi) = 1 \,, \quad \text{and} \quad \mathcal{B}(\overline{D}_0^* \to \overline{D}\pi) &= 1 \,. \end{aligned} \tag{3.13}$$

For the remaining $B \to D^{(*)} \, \pi \, \pi \, \ell^+ \, \nu$ contributions the measured value of [109] is implemented. A summary of the masses and widths used in the event generation of the new $B \to D^{**} \ell\nu$ sample is given in Table 3.2, while the updated branching fractions are shown in Table 3.3.

Non-resonant $B \to X_c \ell\nu$: The remaining "gap" between the sum of all considered exclusive modes and the inclusive $B \to X_c \ell\nu$ branching fraction is filled in equal parts with $B \to D \, \eta \, \ell^+ \, \nu$ and non-resonant $B \to D^* \, \eta \, \ell^+ \, \nu$ decays that are not included in the current, official generic Belle MC. The final states are simulated assuming that they are produced by the decay of two intermediate broad resonant states, D^{**}_{gap}, with masses and widths identical to D_0^* and D_1^*. Currently, no experimental evidence suggests the existence of an additional broad state in semileptonic transitions, however this estimate provides a better kinematic description of the initial three-body decay, $B \to D^{**}_{\text{gap}} \, \ell\nu$,

TABLE 3.2: Masses and widths of the four different D^{**} mesons used in the event generation of the new $B \to D^{**}\ell\nu$ sample. The values are taken from [30], while the values given in parentheses denote the original values that were used in the simulation of the official Belle MC.

	Charged		Neutral	
D^{**} Type	Mass (GeV)	Width (GeV)	Mass (GeV)	Width (GeV)
D_1	2.4230 (2.4270)	0.0200 (0.0280)	2.4208 (2.4222)	0.0317 (0.0189)
D_2^*	2.4654 (2.4590)	0.0467 (0.0250)	2.4607 (2.4589)	0.0475 (0.0230)
D_1^*	2.4450 (2.4223)	0.2503 (0.4120)	2.4450 (2.4223)	0.2503 (0.4120)
D_0^*	2.3000 (2.3080)	0.2700 (0.2760)	2.3000 (2.3080)	0.2700 (0.2760)

than alternative models. An example of an alternative model is based on the equidistribution of all final-state particles in phase-space, which predicts lower lepton momenta than expected from semileptonic B decays. Comparisons of kinematic distributions for the different $B \to D^{**}_{\text{gap}}\,\ell\nu$ models are shown in Figure 3.11. A summary of the assumed branching fractions for the above-mentioned hypothetical decays is given in Table 3.3. Throughout this analysis, an uncertainty of 100% is assumed for these samples. An overview of the current status concerning the gap between the sum of all exclusive states and the inclusive branching fraction, together with possible solutions, is detailed in [60].

Continuum: The generic continuum sample contains simulated events for $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$ processes. A total of six streams are available, which consists of two individual samples containing simulated events for $e^+e^- \to c\bar{c}$ processes and events with light quarks u, d, and s, respectively. Since the modelling of quark fragmentation is difficult, analyses with a large number of continuum background events often rely on the off-resonance data sample for accurate modelling of these processes.

$\boldsymbol{B \to X_u \ell \nu}$: Semileptonic $B \to X_u\ell\nu$ decays are modelled as a mixture of specific exclusive modes and non-resonant contributions. A method originally proposed by [110], known as the "hybrid" method, is adopted to combine the inclusive and exclusive $b \to u\ell\nu$ processes. The implementation of this approach is described in [111].

3.3.2 Monte Carlo corrections

Since the generation of the official Belle MC, several decay models describing key signal processes have become outdated, while branching fractions of various signal and background decays have been measured with much greater precision. Furthermore, different reconstruction and particle identification efficiencies between data and MC are observed, which led to the overall improvement of the initial detector response models over the course of the Belle detector's lifetime. To account for these shortcomings, the MC samples are corrected by applying event-by-event weights to

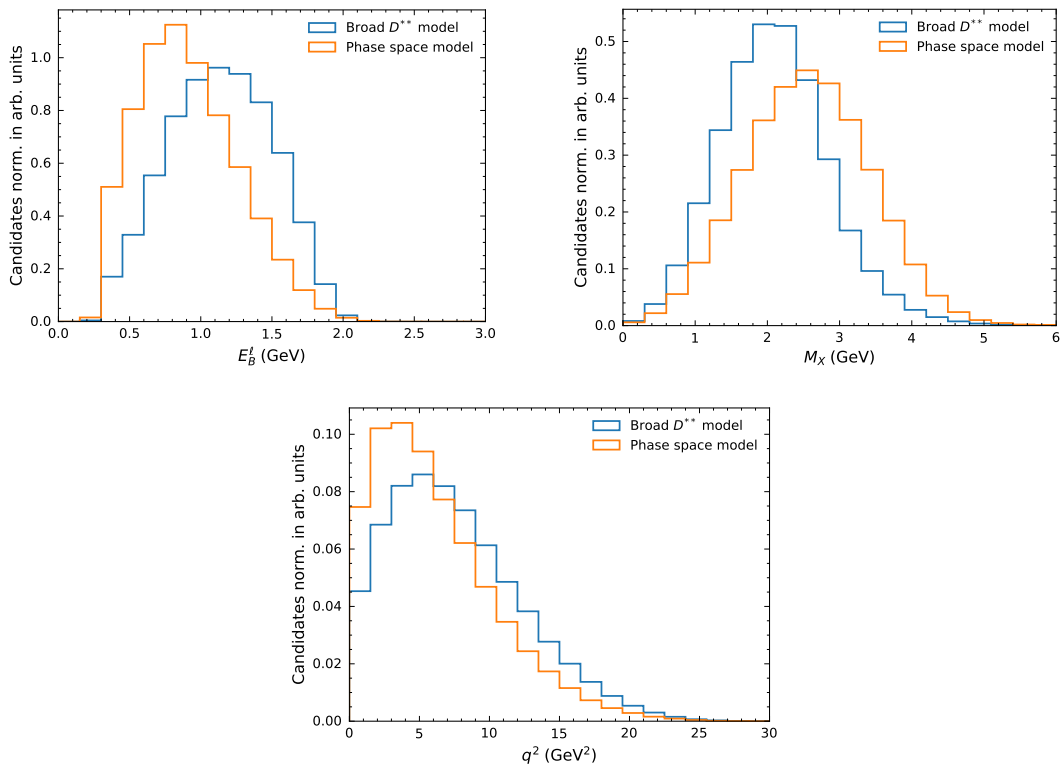

FIGURE 3.11: Comparisons between the different $B \to D^{**}_{\text{gap}}\,\ell\bar{\nu}$ decay models for the reconstructed lepton energy in the signal B rest frame E_ℓ^B (upper left), the M_X (upper right), and the q^2 (bottom) distributions.

the affected samples or decays. Depending on the type of correction under consideration, these weights are either calculated analytically or obtained through data-driven methods. The uncertainties on the calculated correction weights are treated as different sources of systematic errors. An outline of the corrections considered in this analysis is given in the next sections, while the implementation of the uncertainties on the correction weights as systematic errors is discussed in Section 5 and 6.

Branching fractions

Due to the presence of numerous different decay channels in the simulated MC samples, it is simply not feasible to update the branching fractions of all involved processes. Rather, the branching fractions of crucial signal processes, as well as dominant background components, are updated to the current world averages quoted in [3] and [30]. Table 3.3 shows a summary of the updated branching fractions of semileptonic $B \to X\ell\nu$ decays. Furthermore, secondary leptons originating from semileptonic D meson decays form a large contribution to the secondary lepton background component. Consequently, the branching fractions of these decays are scaled to the latest values in [30] (shown in Table A.1), while the uncertainties on these values are assigned as an additional systematic error on this background component.

TABLE 3.3: The updated branching fractions for the main signal and background processes. Values of the $B \to X_c \ell \nu$ branching fractions are taken from the world averages given in [3], while the total inclusive $B \to X_u \ell \nu$ branching fractions is from [30].

$\mathcal{B}$	Value B^+	Value B^0
$B \to X_c \ell^+ \nu_\ell$		
$B \to D \ell^+ \nu_\ell$	$(2.5 \pm 0.1) \times 10^{-2}$	$(2.3 \pm 0.1) \times 10^{-2}$
$B \to D^* \ell^+ \nu_\ell$	$(5.4 \pm 0.1) \times 10^{-2}$	$(5.1 \pm 0.1) \times 10^{-2}$
$B \to D_0^* \ell^+ \nu_\ell$	$(0.420 \pm 0.075) \times 10^{-2}$	$(0.390 \pm 0.069) \times 10^{-2}$
$(\hookrightarrow D\pi)$		
$B \to D_1^* \ell^+ \nu_\ell$	$(0.423 \pm 0.083) \times 10^{-2}$	$(0.394 \pm 0.077) \times 10^{-2}$
$(\hookrightarrow D^*\pi)$		
$B \to D_1 \ell^+ \nu_\ell$	$(0.422 \pm 0.027) \times 10^{-2}$	$(0.392 \pm 0.025) \times 10^{-2}$
$(\hookrightarrow D^*\pi)$		
$B \to D_2^* \ell^+ \nu_\ell$	$(0.116 \pm 0.011) \times 10^{-2}$	$(0.107 \pm 0.010) \times 10^{-2}$
$(\hookrightarrow D^*\pi)$		
$B \to D_2^* \ell^+ \nu_\ell$	$(0.178 \pm 0.024) \times 10^{-2}$	$(0.165 \pm 0.022) \times 10^{-2}$
$(\hookrightarrow D\pi)$		
$\rho(D_2^* \to D^*\pi, D_2^* \to D\pi) = 0.693$		
$B \to D_1 \ell^+ \nu_\ell$	$(0.242 \pm 0.100) \times 10^{-2}$	$(0.225 \pm 0.093) \times 10^{-2}$
$(\hookrightarrow D\pi\pi)$		
$B \to D\pi\pi \ell^+ \nu_\ell$	$(0.06 \pm 0.06) \times 10^{-2}$	$(0.06 \pm 0.06) \times 10^{-2}$
$B \to D^*\pi\pi \ell^+ \nu_\ell$	$(0.216 \pm 0.102) \times 10^{-2}$	$(0.201 \pm 0.095) \times 10^{-2}$
$B \to D\eta \ell^+ \nu_\ell$	$(0.396 \pm 0.396) \times 10^{-2}$	$(0.399 \pm 0.399) \times 10^{-2}$
$B \to D^*\eta \ell^+ \nu_\ell$	$(0.396 \pm 0.396) \times 10^{-2}$	$(0.399 \pm 0.399) \times 10^{-2}$
$B \to X_c \ell^+ \nu_\ell$	$(10.8 \pm 0.4) \times 10^{-2}$	$(10.1 \pm 0.4) \times 10^{-2}$
$B \to X_u \ell^+ \nu_\ell$	$(2.2 \pm 0.3) \times 10^{-3}$	$(2.0 \pm 0.3) \times 10^{-3}$

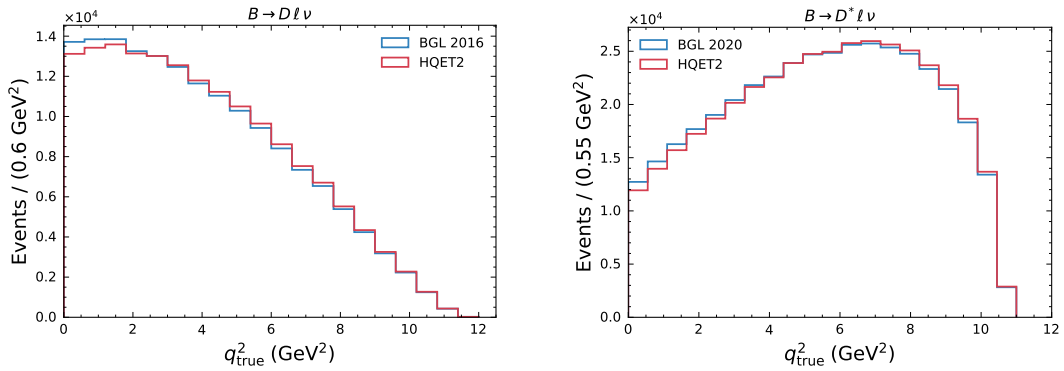

FIGURE 3.12: Comparison of the original and reweighted q^2 distributions for $B \to D\ell\nu$ (left) and $B \to D^*\ell\nu$ (right) decays. The original MC samples were simulated using the CLN form factor parametrization (red) that are reweighted to the BGL (blue) model.

Form factor corrections

Since the shapes of kinematic distributions are dictated by the choice of form factor parametrization for a given kinematic decay model, the correct modelling of $B \to X_c \ell \nu$ decays plays a crucial role in signal and background separation. A mismodelling of a decay process may not only result in poor modelling of the distributions that are chosen to be signal extraction variables, but also affect the overall selection efficiency of the process. Therefore, the main exclusive signal decays are updated to the latest theoretical models. To this end, event-wise correction weights are calculated from the normalised differential decay rates of the new and old models according to the following equation:

$$w(\vec{v}) = \frac{\Gamma_{\text{old}}}{\Gamma_{\text{new}}} \frac{d\Gamma(\vec{v})_{\text{new}}/dv}{d\Gamma(\vec{v})_{\text{old}}/dv} \,. \tag{3.14}$$

Here, the weight w is a function of a set of kinematic variables $\vec{v}$, while the differential decays rates are either determined from analytical expressions or from MC distributions. An overview of the different reweighting strategies implemented for different signal decays is given below.

- $\boldsymbol{B \to D^{(*)} \ell \nu}$: These decays completely dominate inclusive semileptonic $B \to X_c \ell \nu$ decays and are simulated using the CLN parametrization (Section 2.2.2) with the `HQET2` EvtGen package. However, it has recently become clear that the CLN parametrization no longer provides an adequate description of the $B \to D\ell\nu$ data from current experiments and is superseded by the model-independent BGL parametrization. Event-by-event correction weights are calculated using the library by [112] to evaluate Eqn. 3.14 with the analytical expressions of the differential decay rates. As previously discussed, the differential decay rate for $B \to D\ell\nu$ decays is described by the recoil variable w, defined in Eqn. 2.41, and only one form factor to a very good approximation. The most recent form factor central values and uncertainties, obtained from the fit result in [51], are implemented to determine the weights. For $B \to D^*\ell\nu$ decays, the BGL implementation proposed by [113] is utilised, while using the form factor central values and uncertainties obtained from the fit results of [52].

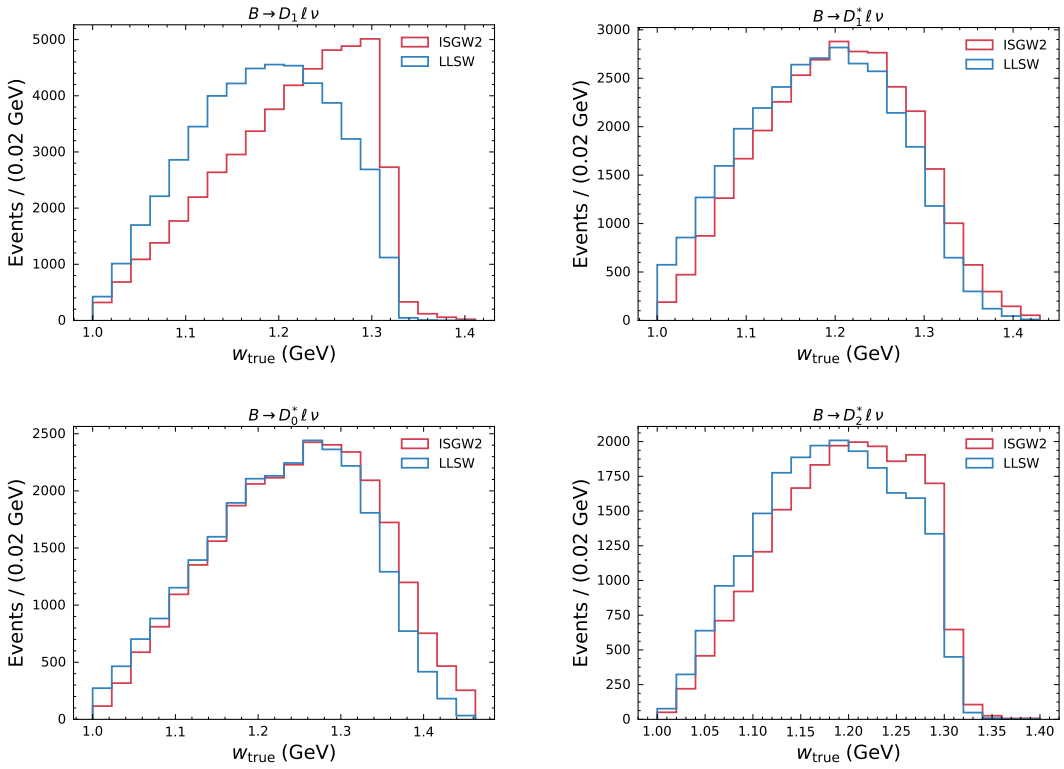

FIGURE 3.13: Comparison of the original and reweighted w distributions for each of the four $B \to D^{**}\ell\nu$ decays. The original MC samples were simulated using the ISGW form factor parametrization (red) that are reweighted to the LLSW (blue) model.

Here, several parametrizations with different truncations were provided and the most precise set with the best χ^2 value is adopted. Since the differential rate for $B \to D^*\ell\nu$ decays is described completely by four independent kinematic decay variables, the reweighting is performed in four dimensions. The four variables most commonly used are q^2 or w together with the three angular variables, defined in Section 2.2.1 and illustrated in Figure 2.6. Comparisons of the original and newly reweighted q^2 distributions for both $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays are shown in Figure 3.12, while the additional kinematic variables are shown in Figure A.1.

$B \to D^{**}\ell\nu$ and "gap": The MC samples for the $B \to D^{**}\ell\nu$ decays are generated with the `ISGW2` EvtGen package, based on the quark model described in [114]. This model describes the data poorly and has been superseded by the LLSW model (Section 2.2.2). Since the differential decay rates for $B \to D^{**}\ell\nu$ decays are not implemented analytically, the ratio given in Eqn. 3.14 is rather estimated from an interpolated ratio. This is achieved by employing a similar method outlined in [115]. The interpolation is approximated by creating generator-level MC histograms of the relevant kinematic decay variables. Since the shape of the reconstructed q^2 distribution is the most important distribution for this analysis, the generator-level histograms are created using only one of the kinematic decay variables describing $B \to D^{**}\ell\nu$ decays, namely the recoil variable w. Furthermore, the range and number of bins used

to construct the histograms of w vary for each D^{**} type. In order to determine these histograms, dedicated $B \to D^{**}\ell\nu$ MC samples, corresponding to one million events, are generated using the ISGW2 and LLSW decay models for each of the four D^{**} types. The central values of the LLSW form factor parameters proposed in [53] are adopted to simulate the new samples. Comparisons of the updated q^2 distributions for each of the D^{**} types are shown in Figure 3.13, while the histogram parameters for each D^{**} type are summarised in Table A.2.

As described in Section 3.3.1, the non-resonant $B \to X_c\ell\nu$ contribution is simulated assuming that the final states are the decay products of two intermediate broad resonant states with masses and widths identical to D_0^* and D_1^*. The advantage of this model is that these $B \to D_{\text{gap}}^{**}\ell\nu$ decays are correctly described by the LLSW form factor parametrization. Therefore, the dedicated gap MC samples are also updated from the default ISGW2 to the LLSW model by making use of the implementation described above.

In order to estimate the systematic uncertainty due to the choice of form factor parametrization, the experimental correlations between the fitted form factor parameters must be taken into consideration. To this end, a new basis of orthogonal parameters is defined that contains uncorrelated uncertainties. Each of the new parameters corresponds to the eigenvalues λ_i and eigenvectors $\vec{e}_i$ of the covariance matrix V. This matrix describes the correlations between the central values of the fitted parameters $\vec{x}_{\text{nom}}$ and is usually provided in addition to the fit results. Subsequently, a set of one sigma variations is calculated in the newly defined basis with the equation:

$$\vec{x}^{\pm}_{\text{var},i} = \vec{x}_{\text{nom}} \pm \sqrt{\lambda_i}\vec{e}_i\,, \tag{3.15}$$

These varied parameters are used as inputs for Eqn. 3.14 to calculate sets of varied correction weights which are, in turn, utilised in the analysis to compute independent contributions of systematic uncertainty from each of the form factor parameters.

Particle identification efficiency corrections

Since the identification of charged tracks and neutral clusters is dependent on the geometry of the Belle detector, differences in the identification efficiency between the simulated MC samples and data are observed due to inaccurate modelling of the detector response. In order to take the observed differences into consideration, calibrations are performed on precisely measured decays, known as control samples, to provide sets of efficiency corrections that are used to adjust the existing MC samples. The most important Particle identification (PID) corrections applied to the simulated MC samples for this analysis are described below, while the algorithms used to reconstruct and identify various final state particles with information of collision events recorded by the Belle detector are described in greater detail in Section 4.1.

Lepton PID efficiency: The lepton identification efficiency is studied in [116] by using a control sample of the two-photon process $e^+e^- \to e^+e^-\ell^+\ell^-$. Since this sample is a clean channel with a low multiplicity of tracks, the efficiency corrections are validated on a sample with hadrons in the final states, namely inclusive $B \to XJ/\psi(\to \ell^+\ell^-)$. In order to account for different acceptance regions due to the detector geometry as well as differences due to the track

momentum, the efficiency corrections are available in bins of lepton momentum p_ℓ in the lab frame and polar angle θ. The bin-wise corrections are given by the efficiency ratio,

$$R_i = \frac{\epsilon_i^{\text{data}}}{\epsilon_i^{\text{MC}}}\,, \tag{3.16}$$

where ϵ_i^{MC} and ϵ_i^{data} are the efficiencies in MC and data, respectively. The uncertainty associated with the efficiency corrections consists of contributions stemming from two different sources: a statistical error σ_i^{stat} due to the size of MC and data samples, while the two systematic uncertainties, $\sigma_i^{\text{sys,1}}$ and $\sigma_i^{\text{sys,2}}$, stem from the reconstruction of the two decay channels. In order to estimate the total uncertainty due to the lepton identification correction factors, the bin-wise efficiency ratios together with their respective uncertainties are used to create an ensemble of variation ratios. To this end, statistical fluctuations are simulated by generating varied errors that is randomly drawn from a Gaussian distribution such that:

$$\Delta\sigma_i^{\text{stat}} = \mathcal{G}(0, \sigma_i^{\text{stat}})\,. \tag{3.17}$$

One the other hand, since the systematic uncertainties are assumed to be fully correlated across all bins, the total systematic uncertainty is multiplied with a global scaling factor that is randomly drawn from a unit normal distribution:

$$\Delta\sigma_i^{\text{sys}} = \mathcal{G}(0,1) \times \sigma_i^{\text{sys}}\,, \tag{3.18}$$

after estimating the total systematic error by adding the two error sources in quadrature. Subsequently, sets of varied efficiency ratios are then calculated in the following manner:

$$R_i^{\text{var}} = R_i^{\text{nom}} + \Delta\sigma_i^{\text{stat}} + \Delta\sigma_i^{\text{sys}}\,. \tag{3.19}$$

The corrections are applied to all correctly reconstructed leptons, while the variation weights are utilised to evaluate the total systematic uncertainty due to the PID corrections.

Hadron PID and fake rate: The efficiency and fake rate corrections of kaon and pion identification are investigated [117] by making use of the $D^{*+} \to D^0(\to K^-\pi^+)\pi^+_{\text{slow}}$ process. The procedure is similar to the lepton identification efficiency discussed in the previous section. Furthermore, the mis-identification rate of hadrons reconstructed as leptons is studied by [118] using the same control sample. Since this decay has a low available phase space, the pions that originate from the D^* decay have a low momentum in the range of 50 MeV $< p_{\pi_{\text{slow}}} <$ 200 MeV in the lab frame and are thus known as slow pions. A clear distinction between correctly reconstructed and background events is visible in the distribution of the mass difference $\Delta m = m(D^*) - m(D)$, which is fitted to obtain the number of events per bin. Similar to the previous PID corrections, the resulting efficiency ratios are provided in bins of lepton momentum, evaluated in the lab frame and the polar angle. All kaon and pion tracks are corrected, while varied efficiency ratios are generated using the same prescription outlined in the previous section.

Track finding efficiency: In order to account for the track finding efficiency for high momentum tracks, the number of partially and fully reconstructed D^* decays

in MC and data is investigated and compared in [119]. Here, a systematic uncertainty associated to the track finding efficiency is estimated to be 0.35% per track. Therefore, the multiplicity of reconstructed tracks with $p_T > 200\,\text{MeV}$ is determined and the recommended weight of 0.35% per track is applied as an additional systematic. One weight per event is calculated, which is then used to obtain a pair of varied weights.

Chapter 4

Event reconstruction

Many different kinds of particles can be produced in e^+e^- collisions, however many of these particles cannot be observed directly. Often the type of particle and its kinematic properties can only be inferred by examining its interactions with various components of the Belle detector. Some particles, such as B mesons, have short lifetimes and decay before they can reach any sensitive component of the detector and therefore cannot be measured directly. On the other hand, the decay products of the B meson might have lifetimes that are long enough to interact with the Belle detector. However, decays involving neutrinos are hard to reconstruct, since neutrinos traverse the entire detector without interacting with any sub-detector systems. This means that there are only a small number of final state particles that interact with various sub-detector components. Final state particles in B meson decays that are stable enough to be observed[1] by the Belle detector are: $e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$, protons, neutrons and photons.

The algorithms employed by the Belle Collaboration to reconstruct and identity final state particles are briefly outlined in Section 4.1. This is followed by a discussion of the various strategies utilized by B-factory experiments to reconstruct B meson decays in Section 4.2. Since this analysis make use of a technique known as hadronic tagging, Section 4.2.1 describes the full reconstruction of the candidate B meson in the event in purely hadronic decay modes with the aid of neural networks. This approach enables the explicit reconstruction of the inclusive hadronic X_c system and, in turn, the explicit reconstruction of the q^2 spectrum, discussed in 4.2.2. The reconstruction strategy employed by this analysis is similar to the method outlined in [107]. Lastly, Section 4.3 introduces the main processes that can mimic the experimental signature of signal $B \to X_c \ell \nu$ decays and summarises the event signal selection criteria. These requirements are optimised to not only reject background processes, but also ensure the selection of well-reconstructed candidates.

4.1 Particle identification

In order to identify final state particles, a probability density function (PDF) is calculated for each component of the Belle detector. The PDFs are combined to create a likelihood account for the particle candidate, which peaks close to unity. A description of the discriminating variables that are employed for particle identification is given in the following sections and are summaries of the in-depth discussions found in [92].

[1] In order to avoid the difficulties in understanding the reconstruction efficiencies of K_L^0 mesons, these particles are not explicitly reconstructed or used in this measurement.

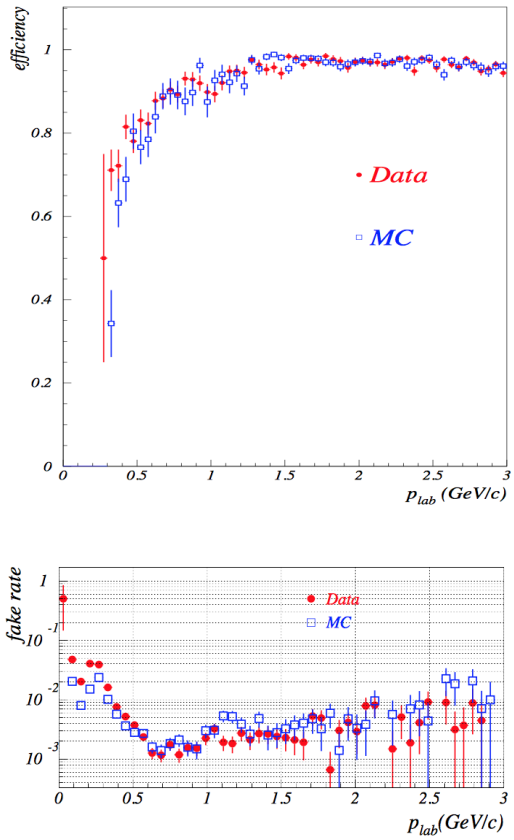

FIGURE 4.1: Efficiency of electron identification (top) and the pion fake rate (bottom) in radiative Bhabha events as a function of lab momentum for $\mathcal{L}_{eid} > 0.5$. Data is represented by filled red circles and MC by blue squares [120].

4.1.1 Electron identification

Several discriminating variables are used to associate a track and calorimeter cluster with an electron candidate. The following list of the most important identifying features is a summary of the discussion found in [120]:

- **Track to cluster matching**: electron tracks extrapolated to the ECL are required to match with the position of the ECL cluster. The matching is assessed with a χ^2-like variable based on the separation between the extrapolated track and the centre of the ECL cluster.
- **E/p**: the ratio between the energy deposited in the ECL and the momentum measured by the CDC. Since the mass of the electron is negligible compared to other final state particles, this ratio is close to unity for electrons.
- **E9/E25**: since electrons produce a narrow shower in the ECL, the transverse shower shape is compared with the ratio of the deposited energy in 3×3 (E9) and 5×5 (E25) crystals. This variable peaks at 0.95 for electrons, while the value tends to peak at lower values for hadrons.
- **dE/dx**: the specific energy loss due to ionisation of a charged track's trajectory. The energy loss is dependent on a particle's velocity β and provides excellent separation between electrons and pions for momenta greater than 0.5 GeV.

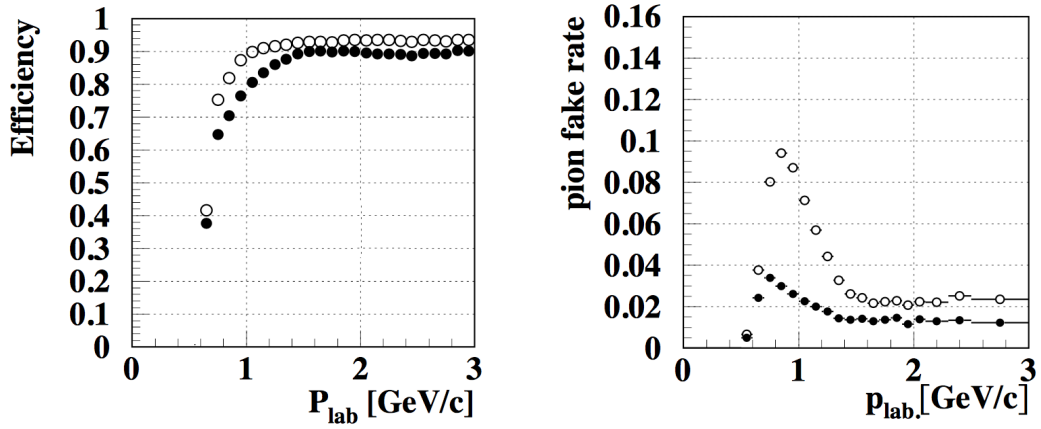

FIGURE 4.2: Efficiency of muon identification (left) and the pion fake rate (right) as a function of lab momentum over the entire polar angle region. Open circles represent $\mathcal{L}_{\mu id} > 0.1$, while filled circles denote $\mathcal{L}_{\mu id} > 0.9$ [121].

- **ACC light yield**: the presence or absence of photoelectrons due to the Cherenkov effect can aid in the particle identification of a passing particle, since the threshold for emitting photons differs for different particle candidates. This threshold is of the order of MeV for electrons, while it lies in the momentum range of $0.5 - 1$ GeV for pions. The separation of electrons from pions is therefore only possible for momenta below 1 GeV.

A likelihood between electron and pion candidates is constructed by combining the PDFs from each of the discriminating variables. Subsequently, an overall likelihood for electron identification is defined as the sum of products from a single variable:

$$\mathcal{L}_{eid} = \frac{\prod_i \mathcal{L}_i^e}{\prod_i \mathcal{L}_i^e + \prod_i \mathcal{L}_i^\pi} \tag{4.1}$$

where i runs over all input parameters. The efficiency for electron selection and the pion fake rate as a function of the momentum in the lab frame, measured in radiative Bhabha events, are shown in Figure 4.1.

4.1.2 Muon identification

To identify a track produced by a muon candidate, the reconstructed track is extrapolated to the KLM and associated hits within 25 cm of the track are located. Two variables are utilised to calculate PDFs for muon identification: the difference between the measured and expected range of the track ΔR and the normalized transverse deviations of all hits associated with the track χ^2_r. Assuming a weak correlation between the two quantities, the final PDF is a product of these separate PDFs. The likelihood for muon identification is calculated as follows:

$$\mathcal{L}_{\mu id} = \frac{\mathcal{L}_\mu}{\mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K} \tag{4.2}$$

where $\mathcal{L}_\mu$, $\mathcal{L}_\pi$ and $\mathcal{L}_K$ are the probabilities from muons, pions and kaons, respectively. The efficiency of muon selection and the pion fake rate for two different $\mathcal{L}_{\mu id}$

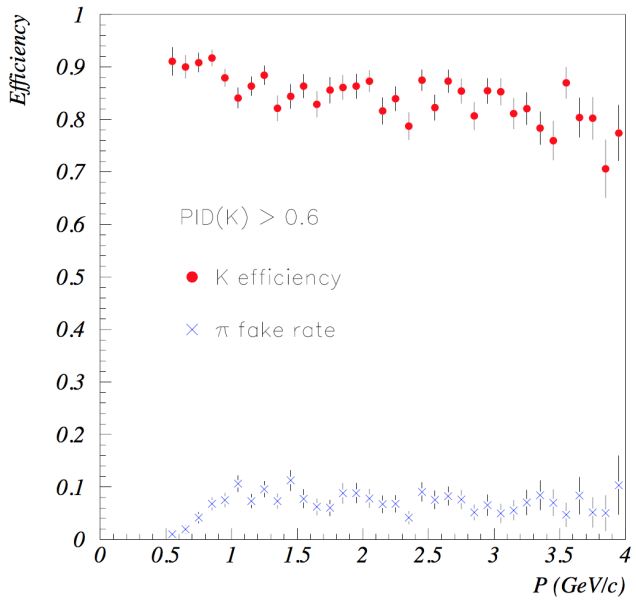

FIGURE 4.3: Kaon efficiency and pion fake rate for the barrel region of the ACC, measured using $D^{*+} \rightarrow D^0(\rightarrow K\pi) + \pi^+$ decays. A likelihood ratio selection of PID(K) ≥ 0.6 is applied [92].

selections as a function of lab momentum, measured in the $e^+e^- \rightarrow e^+e^-, \mu^+\mu^-$ processes, are shown in Figure 4.2.

4.1.3 Hadron identification

In order to identify hadron candidates and to separate pions from kaons, the combined information is required:

- **ACC light yield**: The refractive indices of the aerogel Cerenkov radiators are optimised for the separation of hadrons with a momentum in the range of $1.2 < p < 3.5$ GeV. The likelihoods for different particle candidates are determined from the total light yield that was obtained.
- **Expected time of flight**: The TOF is used to measure particle velocities by making use of the time required for the particle to traverse a certain distance. It is useful for the separation between pions and kaons that carry low momenta below 1.2 GeV. The likelihood is calculated from the difference between the expected time of flight for a particle candidate and the measured time.
- **dE/dx**: the likelihood obtained from the specific ionisation loss is calculated from the difference between a particle candidate's expected specific ionisation loss and the quantity that was measured in the CDC.

The global probability for a track to correspond to either a kaon or a pion is determined as the product of the three likelihood functions. Therefore, a particle is identified as a kaon or pion by selecting a probability from the likelihood ratio:

$$PID(K) = \frac{\mathcal{P}_k}{\mathcal{P}_k + \mathcal{P}_\pi} \tag{4.3}$$

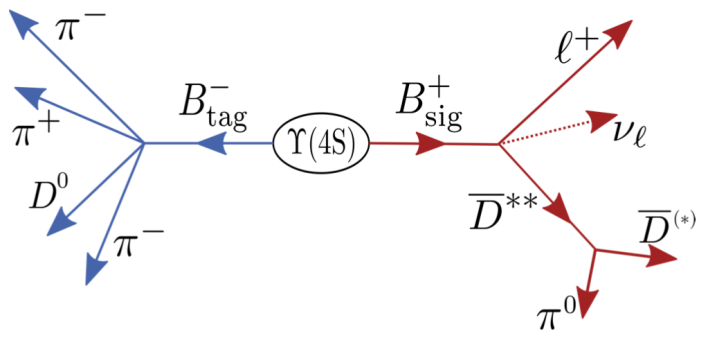

FIGURE 4.4: Signal- and tag-side reconstruction of an $\Upsilon(4S)$ event. One B meson decays in a signal channel $B^+ \to \bar{D}^{**}\ell^+\nu_\ell$ (red), while the companion B meson decays in a hadronic channel $B \to D^0\pi^+\pi^-\pi^-$ (blue), which is reconstructed as the tag-side for this event [115].

where tracks with values of $\mathcal{P}(K)$ close to unity are regarded as kaons, while tracks with values close to zero are identified as pions. The kaon selection efficiency together with the pion fake rate for PID(K) ≥ 0.6 is shown in Figure 4.3.

4.1.4 Reconstruction of photon clusters

Photons are reconstructed from calorimeter clusters that have not been matched to any nearby charged tracks. The ECL is specifically constructed such that a large fraction of the energy of an electromagnetic shower produced by a photon is deposited in this sub-detector system. Consequently, the energy of the photon candidate is obtained by locating the crystal with the largest energy deposit and summing the energies of 3×3 and 5×5 surrounding counters to determine the cluster energy. The position of the cluster is obtained by calculating the "centre of gravity" of the energy, while the momentum vector of the photon candidate is determined from the position and energy of the cluster.

4.2 Reconstruction of B meson decays

Once the final state particles have been reconstructed and identified, their momenta can be summed to create intermediate state particles, which in turn can be assigned to a specific B meson decay. Since semileptonic B decays contain an undetected neutrino with unmeasured momentum in the final state, these decays cannot easily be reconstructed. However, the perfect hermicity of the Belle detector can be exploited to impose kinematic constraints on the decay in order to infer the momentum of the neutrino. In addition, these kinematic constraints benefit from the unique event topology of an $\Upsilon(4S)$ decay, since $B\bar{B}$ pairs are produced without the presence of any additional particles and the initial state energy is precisely known.

B-factories employ three strategies to reconstruct B meson decays: inclusive, semileptonic, and hadronic tagging methods. The inclusive tagging method involves reconstructing only the B meson decay of interest, denoted as the signal B or B_{sig}. Since no additional requirements are placed on the companion B meson, this method yields

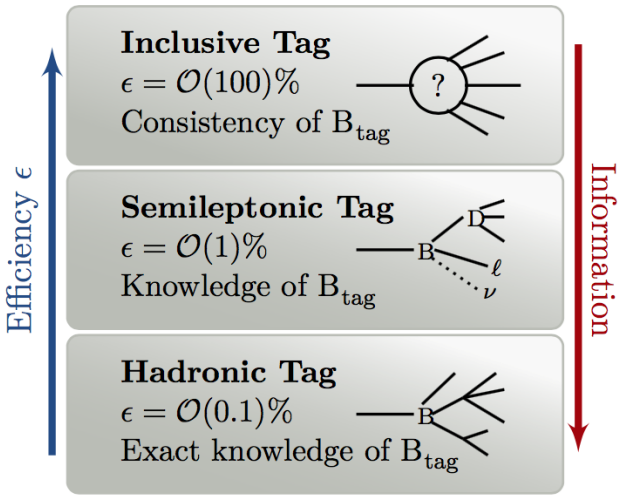

FIGURE 4.5: Different reconstruction techniques of the signal B meson, together with the efficiency and purity for each method. The inclusive method only reconstructs the signal B_{sig}, which results in a large sample of signal and background events. On the other hand, semileptonic and hadronic tagging requires the reconstruction of the companion B_{tag}, which improves the background rejection but yields a smaller signal sample [122].

a large sample of signal candidates, while at the same time incuring a large sample of background events. Therefore, the inclusive tagging method offers a high signal efficiency together with a low purity. The remaining two methods involve the reconstruction of the companion B meson in exclusive channels, denoted as the tag B or B_{tag}. Semileptonic tagging involves reconstructing the B_{tag} in semileptonic decays to charm mesons such as $D^{\pm}$, D^0, $D^{*\pm}$, and D^{*0} among others. Due to the presence of an additional neutrino, the kinematic constraints that can be inferred are restricted and only partial information of B_{sig} can be determined from the B_{tag}. Finally, hadronic tagging involves fully reconstructing the B_{tag} by making use of hadronic decays containing charm mesons of the form $B \to \bar{D}^* n\pi$, $B \to \bar{D}^* D_S^*$ or $B \to J/\psi\, Km\pi$. Here, n and m indicate a number ($n, m < 10$) of charged and neutral pions, respectively. An illustration of the signal- and tag-side reconstruction of an $\Upsilon(4S)$ event using the hadronic tagging method is shown in Figure 4.4. The signal efficiency for hadronic tagging is the smallest of all the reconstruction techniques, of the order 0.1%, and therefore a large data set is required in order to produce precise results. On the other hand, this method provides a considerably high signal purity, together with complete knowledge of the event. Hence, hadronic tagging allows for the determination of the direction, momentum, flavour and charge of B_{sig}. Furthermore, this method ensures that various kinematic variables of the signal side can be accessed, for example the hadronic mass spectrum of the charm system M_X as well as the squared four-momentum transfer q^2. A summary of the three tagging methods together with the interplay of efficiency and purity, is depicted in Figure 4.5. The analysis presented in this thesis utilises hadronic tagging, which is described in further detail in the following section.

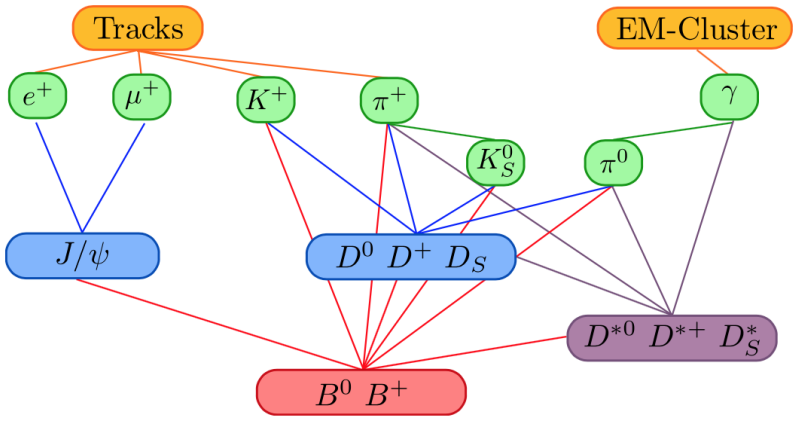

FIGURE 4.6: The four stages of the Full Reconstruction algorithm represented by different colours [123].

4.2.1 NeuroBayes Full Reconstruction

At Belle, B_{tag} candidates are reconstructed using the Full Reconstruction (FR) algorithm, which is based on the NeuroBayes package [123]. The algorithm reconstructs charged and neutral B candidates by making use of a hierarchical approach with four different reconstruction stages, as illustrated in Figure 4.6. At each stage a NeuroBayes artificial neural network combines all available information of a specific particle candidate to generate a single scalar variable referred to as the classifier output. This variable can be interpreted as the probability for a given particle candidate to be reconstructed correctly. The classifier output for each reconstructed particle becomes an input for the neural network in the following stage.

The first stage starts with the reconstruction of final state particles as well as neutral pions and kaons. Each particle candidate that passes the identification criteria is classified by a dedicated neural network that uses various discriminating variables as input including momentum, time-of-flight measurements as well as flight direction, specific energy loss in the CDC, Cherenkov light yield in the ACC for charged particles, and shower shape variables for photons. Multiple candidates are possible, however each track signature may only contribute once in the formation of a B_{tag} candidate. In the next stage, the reconstructed particles are combined to form intermediate state particles or resonances, namely J/ψ, D^0 and $D_s^{\pm}$ candidates. Vertex fits are performed for these candidates in order to improve momentum and spatial resolution. The fit results, together with information of daughter particles and the invariant mass of the candidate, are used as input for the next classifying stage. Excited charmed mesons are reconstructed in the third stage using similar input variables as the previous stage, with the exception of the invariant mass of the candidate. For this stage, the mass difference $\Delta m = m_{D_{(S)^*}} - m_{D_{(S)}}$ is employed instead. Finally, B meson candidates are reconstructed at the fourth and final stage, while making use of variables such as network outputs of the daughter particles, the mass of the D mesons or Δm, and the angle between the B meson momentum and the beam. The combination of all decay modes that are included in the reconstruction algorithm leads to a total of 1104 exclusive B meson decay channels, resulting in an efficiency of 0.28% and 0.18% for charged and neutral B meson pairs [124].

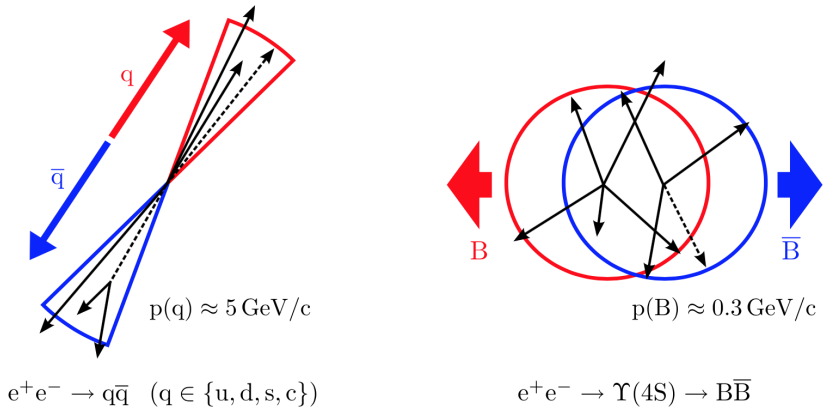

FIGURE 4.7: Comparison between the event shape of a continuum event (left) in comparison to the event shape of a $B\bar{B}$ event (right) in the centre of mass reference frame [125].

After the full hadronic reconstruction the NeuroBayes package also provides a further suppression of continuum events, since a large fraction of events recorded by the Belle detector originates from these processes. At the production centre of mass energy of the $\Upsilon(4S)$ resonance, the cross-section for continuum events $e^+e^- \to q\bar{q}$ has a value of $\sigma(q\bar{q}) = 2.8\,\text{nb}$, while the cross-section for $e^+e^- \to \Upsilon(4S)$ is given by $\sigma(\Upsilon(4S)) = 1.2\,\text{nb}$ [124]. Kinematic constraints from the accelerator as well as different variables describing the event shape can be utilised to discriminate B meson decays from continuum processes. Since B mesons decay almost at rest in the centre of mass frame, their decay products are isotropically distributed in the detector. On the other hand, light quarks in continuum events are produced with large initial momenta in back-to-back topology that results in events exhibiting two jets of light hadrons. A comparison of the event shapes for the two different processes is given in Figure 4.7. There are numerous variables to further discriminate continuum events from $B\bar{B}$ events, which are described in [124] and are briefly summarised in the following list. Each event consists of a set of N particles with momenta $\mathbf{p}_i$ with $i \in [1, 2, \ldots, N]$.

- **Thrust**: The thrust axis $\mathbf{T}$ is defined as the unit vector along which the projection of all momenta is maximised. The scalar thrust variable T is defined as:

$$T = \frac{\sum_{i=1}^{N} |\mathbf{T} \cdot \mathbf{p}_i|}{\sum_{i=1}^{N} |\mathbf{p}_i|} \,. \tag{4.4}$$

  A thrust value of $T = 1/2$ indicates a spherical distribution of the particle momenta, while a value of $T = 1$ corresponds to a jet-like structure.

- **cos θ_B**: The angle between the momentum of the reconstructed B meson and the beam axis is denoted by $\cos\theta_B$, which is distributed as $1 - \cos\theta_B{}^2$. This distribution is flat for continuum events.

- **CLEO cones**: A set of momentum flow variables introduced by the CLEO collaboration in a study of charmless B meson decays [126]. These cones are defined along 10° intervals around the B meson's thrust axis. In these cones, the

scalar momenta of all particles pointing in this interval are summed. For continuum events the momentum flow is clustered inside the CLEO cones with small opening angles.

- **Fox Wolfram Moments**: The Fox Wolfram moments describe the phase-space distribution of energy and momenta in the event. The k-th Fox Wolfram moment is defined as

$$H_k = \sum_{i,j}^{N} |\mathbf{p}_i||\mathbf{p}_j| P_k(\cos\theta_{ij}), \tag{4.5}$$

  where θ_{ij} is the angle between $\mathbf{p}_i$ and $\mathbf{p}_j$ and P_k is the k-th order Legendre polynomial. For collimated jets, the reduced Fox-Wolfram moments, $R_k = H_k/H_0$, take values close to zero and one for odd and even values of k, respectively. R_2 is one of the most powerful variables that can be used to separate $B\bar{B}$ from continuum events.

- **Super Fox Wolfram Moments**: The Super Fox Wolfram moments are calculated by calculating the Fox Wolfram Moments with respect to the particles belonging to the reconstructed B meson (labelled s) and the rest of the event (labelled o). Possible combinations are then R_k^{oo}, R_k^{so} and R_k^{ss}.

The NeuroBayes Full Reconstruction provides an additional algorithm which includes the above-mentioned variables as input for the training of an additional neural network allowing for the separation of continuum and $B\bar{B}$ events. The final classifier score is the network output of the B mesons, denoted as $\mathcal{O}_{\text{FR}}$, and the output discriminating between continuum events, $\mathcal{O}_{\text{Cont}}$. Both outputs take values between 0 and 1 and correspond to the probability of selecting a well-reconstructed B_{tag} candidate. In each event a single B_{tag} candidate is then selected according to the highest $\mathcal{O}_{\text{FR}}$ score value.

The B_{tag} candidates are selected based on requirements of the beam-constrained mass M_{bc}, which further reduces continuum events as well as combinatorial backgrounds, hence ensuring the selection of well-reconstructed candidates. By exploiting the precisely known beam-energy and using the information obtained from fully reconstructing the charge and momentum of the event's decay products, the four-momentum of the B_{tag} candidate is easily determined. The beam-constrained mass of the event is then calculated using the following equation:

$$M_{bc} = \sqrt{E_{\text{beam}}^{*2} - p_B^{*2}}\,. \tag{4.6}$$

Here, $E_{\text{beam}}^{*} = \sqrt{s}/2$ is the beam energy and p_B^{*} is the reconstructed three-momentum of the B_{tag} candidate. Both quantities are evaluated in the centre of mass frame, denoted here by the star. An additional variable of interest is the deviation from the beam energy ΔE which is defined as the difference of the reconstructed B_{tag} candidate's energy and the beam energy:

$$\Delta E = E_B^{*} - E_{\text{beam}}^{*}\,. \tag{4.7}$$

Since this variable is already used as an input for the final stage of the Full Reconstruction algorithm, no further requirements are imposed in this analysis.

After the reconstruction of the B_{tag} candidate all remaining tracks and clusters in the event are used to reconstruct the signal side. By using the momentum of the B_{tag}

candidate together with the precisely known initial beam-momentum, the signal B meson rest frame is defined as:

$$p_{\text{sig}} = p_{e^+e^-} - \left(\sqrt{m_B^2 + |\mathbf{p}_{\text{tag}}|^2}, \mathbf{p}_{\text{tag}}\right). \tag{4.8}$$

Here, $p_{e^+e^-}$ denotes the four-momentum of the colliding electron-positron pair. The reconstruction of the signal side will be described in the following section.

4.2.2 Signal side reconstruction

In order to suppress beam background processes and ensure the selection of quality tracks, final state particle candidates are required to pass several loose selection criteria. Due to the magnetic field surrounding the Belle detector, low momentum tracks spiral inside the CDC and may lead to the reconstruction of duplicate tracks. These tracks can either be two tracks of the same charge within close proximity of each other, or two oppositely charged tracks with a back-to-back trajectory. To identify and remove duplicated tracks similar criteria to those outlined in [127] are implemented. A pair of tracks is regarded as duplicated tracks if their momenta transverse to the beam direction are below 275 MeV and the momentum difference of the two tracks is less than 100 MeV. Additionally, the polar angle between the tracks is $\theta < 15^\circ$ for equally charged tracks, while $\theta > 165^\circ$ for oppositely charged tracks. Once a pair of duplicated tracks is identified, the track which origin lies closer to the interaction point is kept. All remaining charged tracks must meet selection requirements dependent on their transverse momenta in order to suppress beam background processes. These requirements are summarised in Table 4.1. Lepton and hadron identification is then performed on the accepted tracks with stricter requirements.

Well-reconstructed leptons are selected based on the probability likelihood described in Sections 4.1.1 and 4.1.2. Leptons are required to originate with close proximity to the interaction point and lie within the acceptance of the tracking detectors. In order to be identified, electrons must reach the ECL which requires a transverse momentum of $p_T > 0.3\,\text{GeV}$, while muons must traverse the detector and reach the KLM requiring $p_T > 0.6\,\text{GeV}$. Electrons are more susceptible to the Bremsstrahlung process than other charged final state particles, which reduces the reconstructed electron momentum. In order to account for the lost momentum due to the Bremsstrahlung process, a cone of 5° is applied around the flight direction of each electron candidate to identify an energetic Bremsstrahlung photon with $E_\gamma < 1\,\text{GeV}$. Subsequently, the four-momentum of the associated photon is added to the momentum of the corresponding electron candidate. In the case where multiple Bremsstrahlung photons are present, the most energetic photon is selected.

Charged tracks that fail the selection and identification requirements of lepton candidates are reconstructed as kaons or pions based on the probability likelihood described in Section 4.1.3. In addition, kaon and pion candidates are also required to originate from the vicinity of interaction point. Tracks that fail the selection and identification requirements for charged final state particles, summarised in Table 4.2, are discarded.

Photons are identified as neutral clusters in the ECL without an associated charged track. Since the ECL can be separated into different regions according to the polar

TABLE 4.1: Summary of p_T-dependent selection criteria for charged tracks to suppress beam background.

p_T range (GeV)	$\lvert dr \rvert$ (cm)	$\lvert dz \rvert$ (cm)
$p_T < 0.25$	< 20	< 100
$0.25 < p_T < 0.5$	< 5	< 50
$p_T > 0.5$	< 10	< 20

TABLE 4.2: Summary of the selection and identification criteria for charged and neutral particles on the signal side. Final state particles that fail these requirements are discarded.

Particle	ID prob.	dr (cm)	$\lvert dz \rvert$ (cm)	p_T (GeV)	θ
e	eID > 0.9	< 0.5	< 2	> 0.3	(17°, 150°)
μ	muID > 0.9	< 0.5	< 2	> 0.6	(25°, 145°)
K	Atckpi > 0.6	< 1	< 2	-	-
π	Atckpi < 0.4	< 1	< 2	-	-

angle of the photon in the lab frame θ_γ, the energy thresholds to select photon candidates are dependent on different angular regions. The minimum threshold requirements are 50 MeV in the barrel region ($32^\circ < \theta_\gamma < 130^\circ$), 150 MeV in the forward endcap ($130^\circ < \theta_\gamma < 150^\circ$), and 100 MeV in the backward endcap ($17^\circ < \theta_\gamma < 32^\circ$). Neutral pions are reconstructed from a pair of photon candidates with a requirement on the invariant mass to lie between $m_{\gamma\gamma} \in [0.12, 0.15]$ GeV.

The four-momentum of the hadronic system p_X is reconstructed from the sum of the remaining unassigned charged particles and neutral energy depositions after identifying a signal lepton, such that

$$p_X = \sum_i \left(\sqrt{m_\pi^2 + |\mathbf{p_i}|^2}, \mathbf{p_i} \right) + \sum_j \left(E_j, \mathbf{k_j} \right) . \tag{4.9}$$

Where $E_j = |\mathbf{k_j}|$ is the energy of the neutral clusters and all charged particles with momentum $\mathbf{p_i}$ are assumed to be pions. Furthermore, the hadronic mass of the X system, which is later used to discriminate $B \to X_c \ell \nu$ decays from background processes, is reconstructed using

$$M_X = \sqrt{(p_X)^\mu \, (p_X)_\mu} . \tag{4.10}$$

Explicitly reconstructing the X system allows for the direct reconstruction of several kinematic variables of interest using conservation laws. The total observed charge of the event is simply reconstructed as:

$$|Q_{\text{tot}}| = |Q_{B_{\text{tag}}} + Q_X + Q_\ell| . \tag{4.11}$$

In addition, the four-momentum of the neutrino in the event is estimated by exploiting the fact that the initial state is known and the hadronic tag side is fully reconstructed. Thus, the missing four-momentum can be calculated as follows:

$$P_{\text{miss}} = (E_{\text{miss}}, \mathbf{p}_{\text{miss}}) = p_{\text{sig}} - p_X - p_\ell \,, \tag{4.12}$$

where the four-momentum of the measured lepton and the combined X_c is also used. This four-vector, in the case of a semileptonic $B \to X_c \ell \nu$ decay with a correctly reconstructed lepton and X_c system, will fulfil,

$$P^2_{\text{miss}} = M^2_{\text{miss}} = E^2_{\text{miss}} - (|\mathbf{p}_{\text{miss}}|)^2 = 0 \tag{4.13}$$

which is equivalent to:

$$E_{\text{miss}} - |\mathbf{p}_{\text{miss}}| = 0 \tag{4.14}$$

with $E_{\text{miss}} > 0$ and $|\mathbf{p}_{\text{miss}}| > 0$ within the detector resolution.

Finally, the four-momentum transfer squared is reconstructed by once again making use of the fully reconstructed hadronic tag side and is calculated as:

$$\begin{aligned} q^2 &= (p_B - p_X)^2 \\ &= (p_l + p_\nu)^2 \end{aligned} \tag{4.15}$$

where p_X and p_B are the four-momentum vectors of the reconstructed hadronic mass system and the B_{tag} candidate, respectively.

4.3 Event selection

A variety of processes, with a similar experimental signature to that of signal $B \to X_c \ell \nu$ decays, or detector effects can contribute to the inclusive semileptonic event samples. In addition to continuum $e^+e^- \to q\bar{q}$ events, the following background classes are also considered:

- **Secondary background:** Leptons that do not originate from a semileptonic B decay, but from a secondary decay, for example from semileptonic decays of D mesons or J/ψ.
- **Fake leptons:** Kaons and pions that are mis-identified as leptons. The largest contributor to this background source are mis-identified muons.
- $\boldsymbol{B \to X_u \ell \nu}$**:** This background process exhibits a similar experimental signature to $B \to X_c \ell \nu$ decays with the presence of a high momentum lepton and undetected neutrino in the final state. Even though $B \to X_c \ell \nu$ decays are of $\mathcal{O}(100)$ times more likely than inclusive charmless semileptonic B decays, the latest experimental uncertainty for the $B \to X_u \ell \nu$ branching fraction, estimated to be approximately 14%, results in the process forming a large contribution to the overall systematics of the final measurement.

Secondary decays, fake leptons and continuum events are the dominating background processes at low q^2, while $B \to X_u \ell \nu$ becomes more prevalent at high q^2 regions of phase space.

Various analysis selection criteria are imposed to not only minimise the total contributions from background processes, but also to increase the signal to background

ratio. Additionally, the requirements ensure that only events originating from well-reconstructed $B \to X_c \ell \nu$ decays are selected for the analysis. In order to suppress continuum events, which as previously discussed is one of the dominant background processes for semileptonic B decays, the NeuroBayes output discriminating between continuum events $\mathcal{O}_{\text{Cont}}$ is optimised. To determine the optimum selection requirement the Figure of Merit (FOM), which is defined as:

$$\text{FOM} = \frac{S}{\sqrt{S+B}} \tag{4.16}$$

is maximised. Here, S and B denote the number of signal and background events, respectively. Since the Full Reconstruction efficiency differs for continuum and $B\bar{B}$ events, the optimisation is also evaluated for the classifier output, $\mathcal{O}_{\text{FR}}$, which signifies the probability of selecting a well-reconstructed B meson. Both optimisations yield similar results and a requirement of $\mathcal{O}_{\text{FR}} > 10^{-4}$ and $\mathcal{O}_{\text{Cont}} > 10^{-4}$ is chosen. For correctly reconstructed B meson candidates the beam-constrained mass peaks at the nominal B mass, while conversely this distinctive peaking structure is absent for continuum processes. Therefore, to further reject continuum processes only events with a beam-constrained mass greater than 5.27 GeV on the tag B side are selected. Additionally, this requirement aids in removing poorly reconstructed B_{tag} candidates. Since the charge of the signal lepton is equal to the sign of the charge of the b quark for semileptonic B decays, leptons originating from secondary decays are rejected by requiring that the charge correlation between the reconstructed lepton on the signal side and the B_{tag} candidate is correct. This can be achieved by requiring that $Q_{\text{tag}} \times Q_{\text{lep}} < 0$ where Q_{tag} is the charge of the B_{tag} candidate, and Q_{lep} the charge of the signal lepton. Furthermore, the selection requirement is only applied on reconstructed B^+ decays, since this condition does not strictly hold for B^0 decays due to flavour mixing. With all selection criteria on the hadronic tag side complete, the remainder of the requirements focuses on the signal side.

Since $E_{\text{miss}} \approx |\mathbf{p}_{\text{miss}}|$ is true for correctly reconstructed semileptonic $B \to X_c \ell \nu$ decays, the selection of correctly reconstructed signal events is ensured by requiring that $|E_{\text{miss}} - p_{\text{miss}}| < 0.5\,\text{GeV}$. Furthermore, this selection requirement rejects regions of phase space that are poorly understood and inadequately described by the generated MC samples, improving the resolution of the reconstructed q^2 distribution significantly. In addition, the total observed charge of the event is required to be $|Q_{\text{tot}}| \leq 1$, allowing for a charge imbalance in events with low momentum tracks. This requirement is not only aimed towards the selection of well-reconstructed semileptonic decays, but also the rejection of leptons originating from secondary leptons. Since events containing multiple signal leptons are more likely to originate from secondary, charm or J/ψ decays, exactly one signal lepton per event is required.

As can be seen in Figures 4.1 and 4.2, the measured efficiencies of electron and muon identification drastically decreases below a momentum of 1 GeV in the laboratory frame. Conversely, the pion fake rate increases monotonically below this point for muons, before dropping off at approximately 0.6 GeV, while the fake rate increases steadily for electrons below a momentum of 0.5 GeV. Due to the relationship between q^2 and lepton momentum, only events passing the selection requirement of $q^2 > 3\,\text{GeV}^2$ are selected, which excludes all leptons with a momentum of approximately 0.3 GeV in the B meson rest frame. This region corresponds to a region of

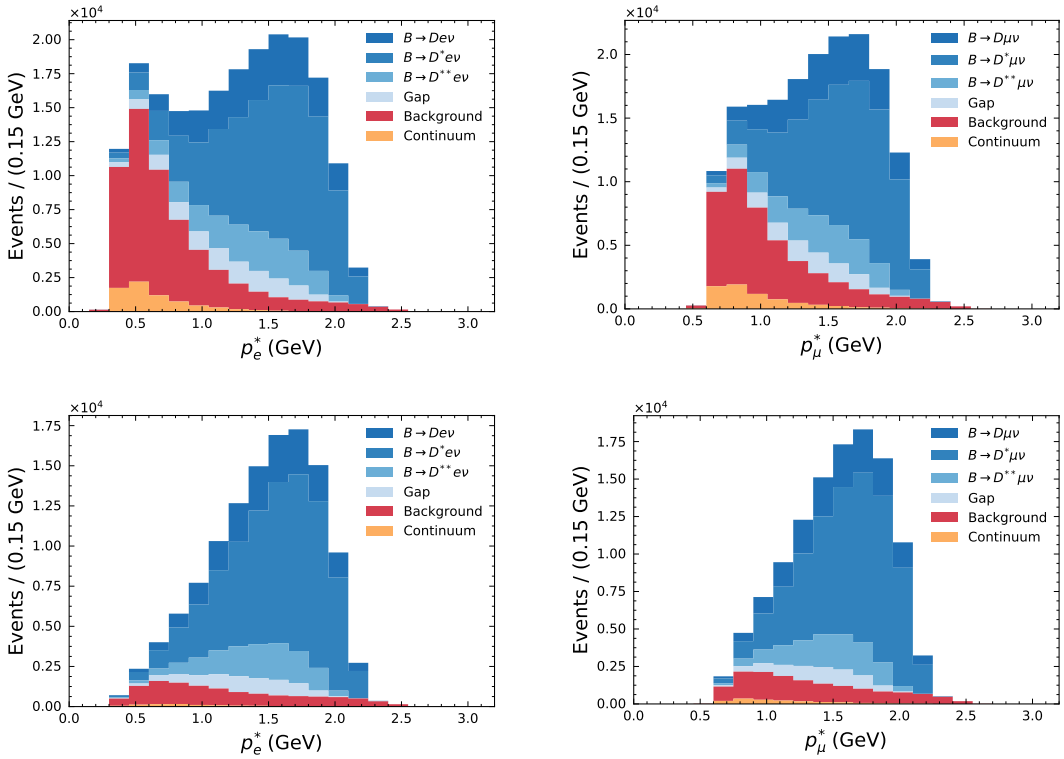

FIGURE 4.8: The reconstructed lepton momentum MC distributions in the B meson rest frame for electrons (left) and muons (right) before (top) and after (bottom) applying the $q^2 > 3.0\,\mathrm{GeV}^2$ selection requirement. Each of the exclusive signal modes, continuum, and "background" which consists of fakes, secondary leptons and $B \to X_u \ell\nu$ decays is shown.

phase space in the laboratory frame at which the Belle detector operates efficiently in the reconstruction and identification of leptons that originate from signal decays. The region of phase space below $3.0\,\mathrm{GeV}^2$ is dominated by leptons originating from secondary decays and continuum processes, as well as fake lepton candidates, which are reduced by this final selection requirement. Figure 4.8 shows the reconstructed lepton momentum distributions for electrons and muons in the B meson rest frame, with and without a requirement of $q^2 > 3.0\,\mathrm{GeV}^2$. Here one can see that this requirement excludes a large amount of background, which mainly stems from leptons originating from secondary decays and fake lepton candidates.

A summary of the impact of each of the signal selection criteria on reconstructed MC events for each of the exclusive signal modes as well as the main background processes is presented in Table 4.3. The values are given in terms of cumulative efficiencies and are normalized to the initial yield obtained after the pre-selection requirements. Furthermore, the initial yield for each signal or background component is provided in the top row of the table. Figure B.1 shows the distributions of the kinematic variables that appear in Table 4.3. The overall efficiency for the selection of signal $B \to X_c \ell\nu$ decays is 27.2%, while the background reduction is 87.6% for $B \to X_u \ell\nu$ decays and 98.7% for the remaining background events.

TABLE 4.3: Comparison of the cumulative selection efficiencies (%) for signal and background processes. The values are normalized to the initial pre-selection yields, which are shown for each of the samples.

	$B \to Dl\nu$	$B \to D^* l\nu$	$B \to D^{**} l\nu$	**Gap modes**
Number of pre-selected events	443758	1353696	561481	333395
Selection				
$\mathcal{O}_{\mathrm{FR}}$ & $\mathcal{O}_{\mathrm{Cont}} > 10^{-4}$	40.34	37.59	30.40	28.93
$Q_{\mathrm{tag}} \times Q_{\mathrm{lep}} < 0$	39.54	36.56	29.32	27.87
$N_{\mathrm{lep}} = 1$	36.53	33.75	26.94	25.53
$M_{bc} > 5.27\,\mathrm{GeV}$	28.77	26.26	18.67	19.49
$\lvert E_{miss} - p_{miss}\rvert < 0.5\,\mathrm{GeV}$	12.48	12.79	8.07	7.61
$\lvert Q_{\mathrm{tot}}\rvert \leq 1$	12.21	12.26	7.67	7.16
$q^2 > 3.0\,\mathrm{GeV}^2$	8.76	9.67	4.84	3.92
	Secondaries	**Fakes**	**Continuum**	$B \to X_u l\nu$
Number of pre-selected events	2284619	878375	5133948	59753
Selection				
$\mathcal{O}_{\mathrm{FR}}$ & $\mathcal{O}_{\mathrm{Cont}} > 10^{-4}$	23.71	23.11	5.39	35.87
$Q_{\mathrm{tag}} \times Q_{\mathrm{lep}} < 0$	16.23	18.17	4.08	35.10
$N_{\mathrm{lep}} = 1$	14.56	16.21	3.90	32.31
$M_{bc} > 5.27\,\mathrm{GeV}$	8.49	7.84	0.77	23.89
$\lvert E_{miss} - p_{miss}\rvert < 0.5\,\mathrm{GeV}$	2.98	3.84	0.36	15.75
$\lvert Q_{\mathrm{tot}}\rvert \leq 1$	2.66	3.43	0.32	15.41
$q^2 > 3.0\,\mathrm{GeV}^2$	0.52	0.70	0.05	12.42

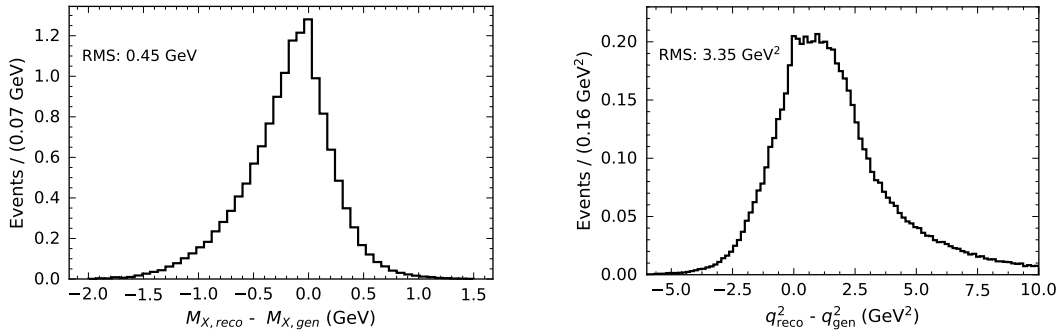

FIGURE 4.9: The resolution of the reconstructed M_X (left) and q^2 (right) distributions for $B \to X_c \ell \nu$ signal decays is shown as a residual with respect to the generator-level distributions. The RMS for both variables is given.

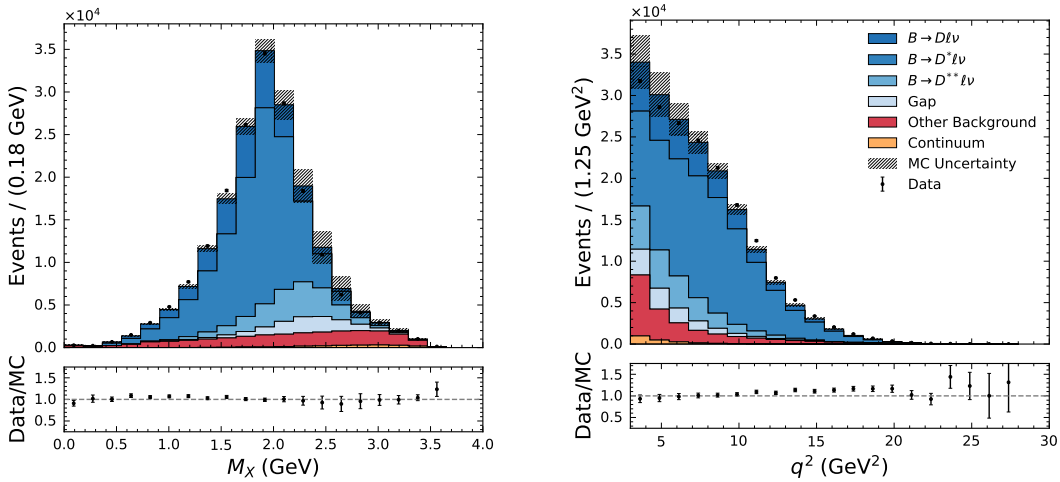

FIGURE 4.10: Comparison of the reconstructed M_X (left) and q^2 (right) distributions with simulated events. The error band of the simulated samples incorporate the full set of systematic uncertainties discussed in Section 3.3.2.

The resolution of both the reconstructed M_X and q^2 distributions for $B \to X_c \ell \nu$ decays is shown in Figure 4.9 as residuals with respect to the generator-level distributions. The resolution for M_X has a root-mean-square (RMS) deviation of 0.45 GeV and exhibits a tail towards negative values of the residuals. Constituents of the hadronic X system that have not been reconstructed correctly or missing charged and neutral particles contribute to the distortion of the resolution of the reconstructed M_X distribution. The four-momentum transfer squared q^2 exhibits a poor resolution, which is caused by a combination of errors in the tag-side B and the X reconstruction. The RMS deviation for q^2 is 3.35 GeV^2. The core resolution is dominated by the tagging resolution, whereas the large positive tail is dominated by the resolution of the reconstruction of the X system. Figure 4.9 compares the selected events with the expected simulated events, obtained from the simulated signal and background MC samples described in 3.3.1. Here, the small contribution from the continuum

processes is normalized using the off-resonance data sample, while the remaining simulated events are normalized to the number of reconstructed events originating from $\Upsilon(4S) \rightarrow B\bar{B}$ decays. The error band of the simulated samples incorporates the full set of systematic uncertainties stemming from various MC corrections, discussed in Section 3.3.2. Despite poor knowledge of the non-resonant component of $B \rightarrow X_c \ell \nu$ decays, the overall observed agreement between the reconstructed events selected in data and the simulated events is fair within experimental uncertainty.

Chapter 5

Background subtraction

Even though the signal selection criteria suppresses a large amount of background events, a small residual amount of fakes and secondaries remain in the low q^2 region. In addition, $B \to X_u \ell \nu$ decays have a very similar experimental signature to the signal $B \to X_c \ell \nu$ decays and become a more significant background component for higher selection values on q^2. In order to determine the remaining background contributions after the final signal selection criteria has been applied, a 1-dimensional binned χ^2 fit in the M_X distribution is performed. Both signal and background yields are determined from the fit parameters, however only the total fitted background yield is utilised to calculate event-by-event signal probability weights, w. The event-wise signal probability function completes the final step to subtract residual background events and is required to calculate the q^2 moments.

The fitting procedure as well as the treatment of statistical and systematic uncertainties in the fit model are described in Section 5.1, while the signal and background fit templates are presented in Section 5.1.4. The description of the method likelihood in the following section is a short summary of the detailed discussion found in [128]. To validate the fit procedure and test for possible bias in the yield extraction, various pseudo-experiments are carried out. The results, which confirm that the fit procedure is unbiased, are summarised in Section 5.2.1. An example fit to the M_X distribution in data with the nominal signal selection criteria is subsequently performed in Section 5.2.2 to extract the remaining contributions of background processes present in the recorded collision events. Finally, the event-wise signal probability weights as a function of lower q^2 selections are calculated in Section 5.3. Unless otherwise stated, all studies are conducted with the nominal selection criteria for the first and loosest q^2 threshold selection of $q^2 > 3.0\,\mathrm{GeV}^2$ in the electron channel.

5.1 Fit model

The aim of the fitting procedure is to find the set of model parameters $\vec{\theta}$ that provides the best agreement between the observed data and the underlying model. Due to the lack of analytical probability density distributions describing background components and non-resonant contributions for signal decays, a binned likelihood fit is performed by using templates representing each background and signal shape. This method exploits the shape differences in a discriminating variable that are observed between individual components to extract the estimated yields for each process.

The number of events in each individual bin of the template histogram is described by a Poisson distribution, $\mathcal{P}(n_i|v_i(\vec{\theta}))$. Here, n_i and v_i denote the number of observed and expected events in bin i, respectively. When considering large data sets,

the underlying Poisson distributions are well-approximated by a Gaussian distribution. The optimal set of parameters that maximises the probability of the observed data with an underlying model hypothesis is given as a maximum likelihood function $\mathcal{L}$, which is constructed as the product of each of the bin-wise Gaussian probabilities:

$$\mathcal{L}(n_i; v_i(\vec{\theta})) = \prod_i^{\text{bins}} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(n_i - v_i(\vec{\theta}))^2}{2\sigma_i^2} \right), \tag{5.1}$$

with a mean of $\mu = n_i$ and the standard deviation given as $\sigma_i = \sqrt{n_i}$.

Taking the logarithm of the PDF in order to reduce computationally expensive products and exponentials gives the log-likelihood function,

$$-2\ln\mathcal{L}(\vec{\theta}) = \sum_i \left(\frac{n_i - v_i(\vec{\theta})}{\sigma_i} \right)^2. \tag{5.2}$$

This equation is maximised by finding the values of the parameters $\vec{\theta}$ that minimise the quantity,

$$\chi^2(\vec{\theta}) = \sum_{i=1}^{N} \left(\frac{n_i - v_i(\vec{\theta})}{\sigma_i} \right)^2, \tag{5.3}$$

which is easily recognised as the well-known method of least squares. In order to generalise this method, Eqn. 5.1 can be replaced with a multivariate Gaussian distribution leading to the generalised form of the χ^2 method,

$$\chi^2(\vec{\theta}) = (\vec{n} - v(\vec{\theta}))^T C^{-1} (\vec{n} - v(\vec{\theta})) \,. \tag{5.4}$$

The covariance matrix C describes the correlations across various bins and is here simply represented by a diagonal matrix, since the bins are known to be uncorrelated for the observed data.

The number of expected events v_i in a specific bin i is defined as:

$$v_i = \sum_k^{\text{templates}} f_{ik} n_k \,, \tag{5.5}$$

where n_k is the total number of expected events from a process described by template k and f_{ik} is the fraction of events found in bin i for template k. This fraction is determined from the underlying template histogram such that:

$$f_{ik} = \frac{n_{ik}}{\sum_j^{\text{bins}} n_{jk}} \,, \tag{5.6}$$

where n_{ik} is the number of expected events from template k in bin i. A total of three signal and background templates are estimated using simulated MC samples and incorporated into the fit procedure. In addition, the signal template is further divided into sub-templates for each of the exclusive $B \to X_c \ell \nu$ decays in order to form a signal multi-template. The total the number of events in a given bin in terms of the considered templates is thus given by,

$$v = n_{\text{sig}} f_{i,\text{sig}} + n_{\text{bkg}} f_{i,\text{bkg}} + n_{\text{cont}} f_{i,\text{cont}} \,. \tag{5.7}$$

Here, n_{sig} is the total number of $B \to X_c \ell \nu$ signal events, which is divided into different sub-fractions for each of the exclusive decay modes. Furthermore, n_{bkg} denotes

the background events stemming from secondary semileptonic decays, $B \to X_u \ell \nu$ decays, and from hadrons mis-identified as leptons. The number of continuum events is denoted as n_{cont}. Further discussions regarding the distributions of the different signal and background fit templates are given in Section 5.1.4.

Systematic uncertainties arising from MC efficiency corrections such as lepton and hadron identification efficiencies, tracking efficiencies and form factor corrections for signal decays allow the overall shape of a template to change. These shape uncertainties are included in the likelihood function by introducing nuisance parameters θ_{ik} that are incorporated by altering Eqn. 5.6 such that,

$$f_{ik} \to \frac{n_{ik}(1+\epsilon_{ik}\theta_{ik})}{\sum_j^{\text{bins}} n_{jk}(1+\epsilon_{jk}\theta_{jk})} \,. \tag{5.8}$$

A nuisance parameter θ_{ik} is introduced for each template k and bin i, while the relative uncertainty of template k in bin i is denoted by ϵ_{ik}. Although a template's shape might change slightly by the fit pulling on a nuisance parameter of a given systematic uncertainty, the denominator of the fraction in the expression above ensures that the overall normalization of the template remains constant. Furthermore, the nuisance parameters are constrained by a multivariate Gaussian distribution $\mathcal{N}(\theta_{jk}|0,\Sigma_{jk})$ with mean $\mu = 0$ and a correlation matrix, Σ_{jk}. This correlation matrix is obtained from the total covariance matrix, which is the combined sum of individual covariance matrices corresponding to each of the systematic and statistical sources. The associated bin-to-bin correlations arising from systematic uncertainties are thus accounted for in the fit with this method.

A Poisson constraint is imposed on the expected yield parameter for the continuum process and is implemented by adding an additional term $\mathcal{P}_{\text{cont}}(n_{\text{cont}}|\tilde{n}_{\text{cont}})$ to the likelihood function. Here $\tilde{n}_{\text{cont}}$ denotes the expected continuum yield as determined from off-resonance collision events, which is adjusted to reflect the number of continuum events for each q^2 threshold selection under consideration. In addition, a multivariate Gaussian distribution $\mathcal{N}(f_l|\tilde{f}_l,\Sigma_{f_l})$ is included to constrain the expected fraction of events f for each of the exclusive $B \to X_c \ell \nu$ processes to the expectations estimated from MC events. In this term $\tilde{f}_l$ is the experimental branching fraction for the respective process, while Σ_{f_l} is the corresponding error covariance matrix.

Including all systematic uncertainties together with the additional constraints, the final likelihood function takes the form:

$$\mathcal{L} = \prod_i^{\text{bins}} \mathcal{P}(n_i|\nu_i(\theta)) \times \prod_k^{\text{templates}} \mathcal{N}(\theta_k|0,\Sigma_k) \times \mathcal{P}_{\text{cont}}(n_{\text{cont}}|\tilde{n}_{\text{cont}}) \times \prod_l^{f_{\text{constr.}}} \mathcal{N}(f_l|\tilde{f}_l,\Sigma_{f_l}) \,. \tag{5.9}$$

The treatment of systematic uncertainties as nuisance parameters as well as the implementation of both the additional constraints are outlined in the following sections.

5.1.1 Systematic uncertainties

Shape uncertainties related to the choice of form factor parametrization for $B \to Dl\nu$, $B \to D^* l\nu$, and $B \to D^{**} l\nu$ decays are incorporated into the fit by determining covariance matrices for each of the form factor parameters associated to the specific

process. This is achieved by making use of the sets of varied event-wise weights, described in Section 3.3.2, and creating varied histograms for the specific signal process. The independent contribution of systematic uncertainty for a given form factor parameter is subsequently estimated as follows:

$$\sigma_i = \frac{h_{\text{up},i} - h_{\text{down},i}}{2}, \tag{5.10}$$

where $h_{\text{up},i}$ and $h_{\text{down},i}$ are the number of events found in bin i of the varied histograms created from a pair of up and down variations. This formulation allows for a change of sign in the case where $h_{\text{down},i} < h_{\text{up},i}$, which introduces a negative correlation between bins with opposite signs. The error source stemming from each model parameter is fully correlated or anti-correlated across all bins by taking the outer product of the bin-wise uncertainty vector to obtain the associated covariance matrix. Each of the individual covariance matrices are summed in order to determine the total covariance matrix relating to $B \to Dlv$, $B \to D^* lv$, and $B \to D^{**} lv$ form factor uncertainties. A similar strategy is employed to include the uncertainties of the branching fractions of secondary leptons originating from semileptonic D decays, summarised in Table A.1, in the fit.

Lepton and hadron identification efficiencies as well as tracking efficiencies are considered in the fit by applying the respective variation weights discussed in Section 3.3.2, and creating new varied histogram templates. By creating histograms for all considered templates with different sets of variation weights, a dedicated global covariance matrix governing shape uncertainties for each of the respective particle identification efficiency corrections can be estimated for the fit. This global matrix is calculated by constructing a matrix M that contains varied histograms for each template with rows spanning the total number of variations for a given efficiency correction. Together with the matrix P, which consists of the equivalent nominal histograms in every row, the global covariance matrix for a given efficiency correction is simply estimated in the following manner:

$$G = \frac{(M-P)^T (M-P)}{N_{\text{var}}}. \tag{5.11}$$

Here, N_{var} denotes the total number of variations for the efficiency correction under study. Since the dimensions of M and P are given by $N_{\text{vars}} \times N'$ where $N' = N_{\text{temps}} \times N_{\text{bins}}$, the global covariance matrix for a given PID correction has dimension $N' \times N'$, as expected.

5.1.2 Continuum constraint

A Poisson constraint is imposed on the expected continuum yield parameter that enters the fit with value n_{cont}. The normalization of the continuum MC is determined using off-resonance data that consists of 79 fb^{-1}. Since the data was collected at a centre of mass energy approximately 60 MeV lower than the on-resonance data set, the data points are shifted to the expected kinematic end-point of the on-resonance M_{bc} distribution. The shift value ΔM_{bc} is calculated by determining the mean difference between on- and off-resonance data sets, which is found to be $\Delta M_{bc} = 0.0289 \pm 0.0004$ GeV.

The constraint value for n_{cont} in the fit is subsequently estimated from the number of expected events in off-resonance data, which is calculated as follows:

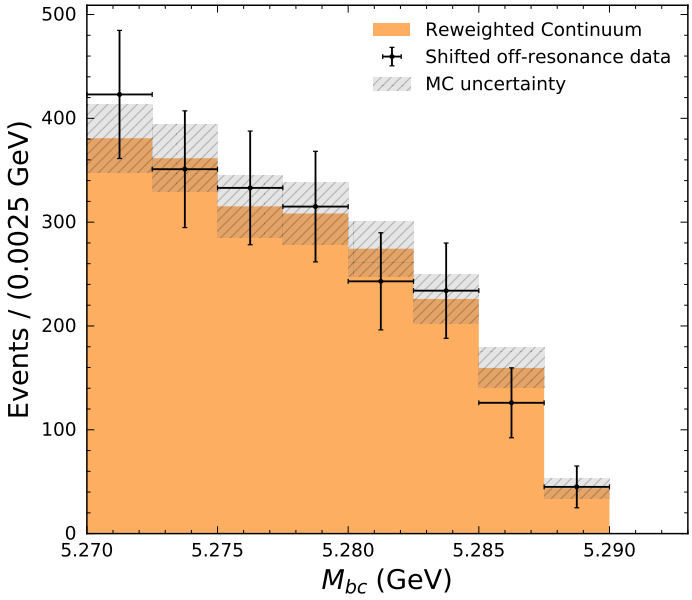

FIGURE 5.1: Comparison between the off-resonance data and reweighted continuum MC M_{bc} distributions for $q^2 > 3.0\,\mathrm{GeV}^2$. Good agreement within statistical and systematic uncertainties is observed.

$$N_{\text{expected}} = N_{\text{off}} \times \frac{711\text{fb}^{-1}}{79\text{fb}^{-1}} \,. \tag{5.12}$$

Here, N_{off} is the number of observed off-resonance events passing the selection criteria, which is scaled up by the ratio of on-resonance to off-resonance luminosity factors. The uncertainty in the shift value ΔM_{bc} is taken into account by varying the shift value around its error and shifting the off-resonance continuum to obtain varied N_{expected} estimates. Thus, the error due to the M_{bc} shift is determined by:

$$\sigma_{\text{shift}} = \frac{|N_{\text{expected, up}} - N_{\text{expected, down}}|}{2} \,. \tag{5.13}$$

In order to determine the final uncertainty of the continuum yield extracted from the fit procedure, the uncertainty on the yield parameter and the error due to the M_{bc} shift are added in quadrature.

The ratio between N_{expected} and the initial number of events found in the continuum MC sample $N_{\text{cont, MC}}$ is calculated as a correction factor,

$$\mathrm{r}_{\text{continuum}} = \frac{N_{\text{expected}}}{N_{\text{cont, MC}}} \,, \tag{5.14}$$

which is used to study the agreement between the shifted off-resonance data and simulated MC distributions. Figure 5.1 compares the M_{bc} distributions for the reweighted continuum MC and the measured off-resonance data. An example correction factor of $r = 0.822 \pm 0.057$ is calculated for the nominal selection requirements, while considering the combined electron and muon final state channels.

5.1.3 Branching fraction constraints

The branching fractions of key $B \to X_c \ell \nu$ processes are propagated into the fit, while the associated errors on these values are implemented as Gaussian constraints for each of the exclusive signal templates. Both the updated branching fraction values and their respective estimated errors are summarised in Table 3.3. In contrast to exclusive signal decays, the non-resonant gap states are corrected by a factor of 1 ± 0.1. This correction is equivalent to assuming that the branching fraction of this component is unchanged, but allowing it to fluctuate by 100% in the fit.

The uncertainties due to branching fractions are propagated into the sub-fraction parameter covariance matrix for $B \to X_c \ell \nu$ decays. The sub-fractions are determined from MC with the following equation,

$$f_i = \frac{N_i}{\sum_k N_k}\,, \tag{5.15}$$

where f_i and N_i is the fraction and number of events observed in MC for a specific exclusive decay. Additionally, the yields for each decay process are updated to the latest state-of-the-art branching fractions by calculating:

$$N_{i^{\text{new}}} = N_i^{\text{old}} \frac{\mathcal{B}_i^{\text{new}}}{\mathcal{B}_i^{\text{old}}}, \tag{5.16}$$

where $\mathcal{B}_i^{\text{new}}$ denotes the newest branching fractions given by [3], while $\mathcal{B}_i^{\text{old}}$ represents the outdated branching fractions. Furthermore, the uncertainty stemming from MC statistics as well as from the updated branching fractions is taken into consideration. Since the expected yields and their respective errors are uncorrelated, the covariance matrix simply contains diagonal elements such that:

$$\Sigma^N = \text{diag}(\sigma^2_{\mathcal{B}_D}, \sigma^2_{\mathcal{B}_{D^*}}, \ldots, \sigma^2_{\mathcal{B}_{\text{gap}}}). \tag{5.17}$$

In order to subsequently determine the covariance matrix for the sub-fractions Σ^f one makes use of the Jacobian formulation of error propagation together with the matrix equation:

$$\Sigma^f = \mathcal{J}^T \Sigma^N \mathcal{J} \tag{5.18}$$

where $\mathcal{J}$ is a Jacobian matrix computed using $\mathcal{J} = \partial f_i / \partial N_i$. This method of propagating the uncertainties of the branching fractions for exclusive decay channels was also implemented in the fitting procedure outlined in [129], which gave consistent results. A summary of the uncertainties due to the branching fractions of $B \to X_c \ell \nu$ processes for both the electron and muon final states is given in Table 5.1. The slight difference between the computed uncertainties of the electron and muon channels is due to the differences in MC statistics and lepton and hadron identification efficiencies.

5.1.4 Fit templates

The remaining number of events originating from background processes in the recorded collision events are determined by performing the above-mentioned likelihood fit in 20 equidistant bins of the M_X distribution as a progression of increasing threshold selections on q^2. In order to avoid empty bins the range of each fit is adjusted for each incremental step in q^2. Furthermore, the fits are performed separately for the

TABLE 5.1: Summary of the uncertainties due to the branching fractions of exclusive $B \to X_c \ell \nu$ processes, separated into the electron and muon final states.

Process	Uncertainty (%)	Process	Uncertainty (%)
$B \to D\, e\, \nu$	5.429	$B \to D\, \mu\, \nu$	5.380
$B \to D^*\, e\, \nu$	4.665	$B \to D^*\, \mu\, \nu$	4.609
$B \to D_1\, e\, \nu$	9.348	$B \to D_1\, \mu\, \nu$	9.374
$B \to D_1^*\, e\, \nu$	7.586	$B \to D_1^*\, \mu\, \nu$	7.520
$B \to D_0^*\, e\, \nu$	7.247	$B \to D_0^*\, \mu\, \nu$	7.212
$B \to D_2^*\, e\, \nu$	5.491	$B \to D_2^*\, \mu\, \nu$	5.480

electron and muon final states due to different mis-identification efficiencies in these channels. As already mentioned, the final likelihood function, shown in Eqn. 5.9, depends on three expectation values v_i of signal and background templates. In addition, the signal template is divided into sub-templates for each of the exclusive $B \to X_c \ell \nu$ decays. The likelihood function is therefore also dependent on six additional signal sub-fractions f while the gap template is constrained by these fractions such that:

$$f_{\text{gap}} = 1 - \sum_m f_m \,. \tag{5.19}$$

Here, the sum runs is performed all exclusive signal states. Template histograms for each of the processes considered in the likelihood fit are created from MC samples. Furthermore, the statistical and systematic uncertainties for each of the fit templates are parametrized by a total of 180 nuisance parameters θ_i with Gaussian constraints. Figure 5.2 shows the template histograms for background processes, while Figures 5.3 and 5.4 show the histograms for signal decays. Additionally, the systematic correlation matrix for each template is also given. The composition of individual templates as well as the considered systematic uncertainties are described below.

- **Signal:** The signal template forms a multi-template consisting of seven templates for the exclusive signal decays: $B \to D\ell\nu$, $B \to D^*\ell\nu$, $B \to D^{**}\ell\nu$ as well as the non-resonant gap contribution. Furthermore, separate templates for each of the individual $B \to D^{**}\ell\nu$ decays, $D^{**} = \{D_1, D_1^*, D_0^*, D_2^*\}$, are also considered. The $B \to D\ell\nu$ and $B \to D^*\ell\nu$ templates exhibit a clear, distinctive peak at $M_X \approx 2.0\,\text{GeV}$ with longer tails extending towards low values of M_X. On the other hand, the distributions for the $B \to D^{**}\ell\nu$ decays as well as the gap template exhibit much broader distributions with peaks between $2.0\,\text{GeV} < M_X < 2.5\,\text{GeV}$ depending on the specific process. Each of the signal templates are constrained to their expected branching fractions, while the gap template is allowed to fluctuate by 100%. Systematic uncertainties considered for all signal templates include MC statistics, lepton and hadron identification efficiency corrections as well as tracking efficiencies. In addition, form factor corrections corresponding to the specific signal process and model are also taken into account.

- **Background:** This template comprises the combination of all remaining background components; namely fake lepton candidates, leptons originating from secondary decays as well as from $B \to X_u \ell \nu$ decays. Unlike the signal templates, this template exhibits a broad structure that peaks at higher values of M_X. The uncertainties for this template include MC statistics, lepton and

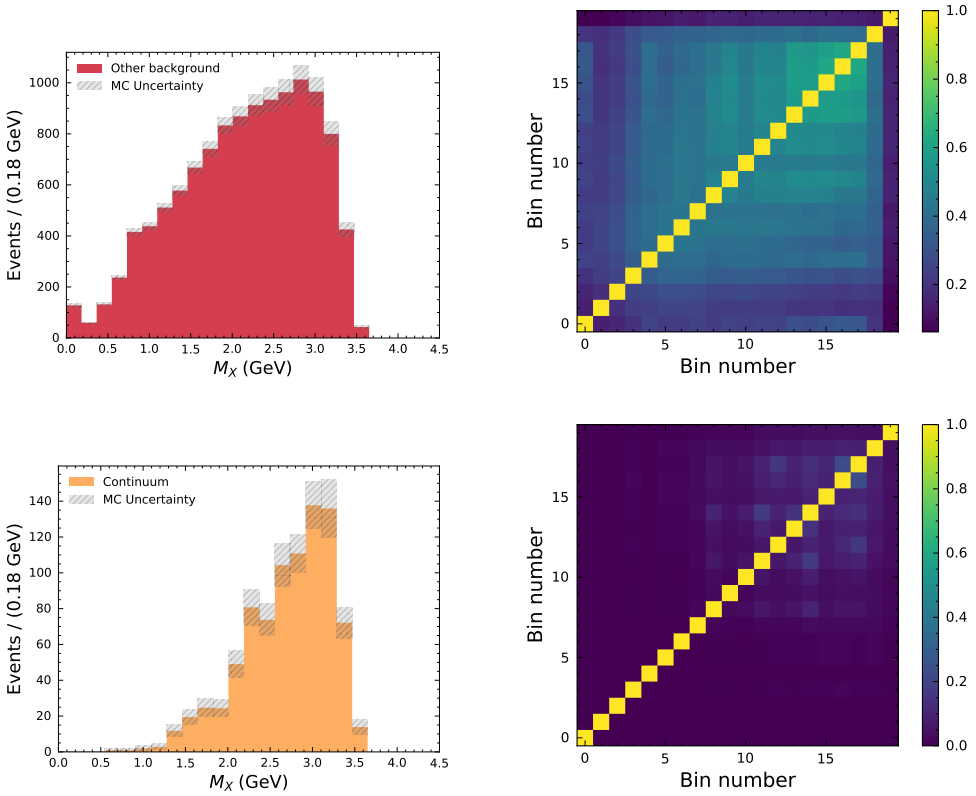

FIGURE 5.2: Template histograms for continuum and background processes together with the corresponding error correlation matrix, calculated on MC samples. The MC uncertainties include uncertainties from MC statistics, lepton and hadron identification efficiency corrections, and the tracking efficiency correction. The uncertainty due to the branching fractions for secondary leptons originating from semileptonic D decays is included for the secondary component of the background template.

hadron identification efficiency corrections, and tracking efficiency. Furthermore, the branching fraction uncertainties for secondary leptons originating from semileptonic D decays are included as an additional systematic for the secondary component.

- **Continuum:** The continuum template consists of two individual components. While one sample contains simulated events for $e^+e^- \to c\bar{c}$ processes, the other contains events with light quarks u, d, and s. This template exhibits a broad peak at $M_X \approx 3.0\,\text{GeV}$, which drops off at lower values for M_X. Although the normalization of this template is constrained to the expectation calculated from off-resonance continuum data, the template may still vary within its statistical and systematic uncertainties. Only lepton and hadron identification efficiency corrections and the tracking efficiency correction are considered for this template.

With the fit procedure setup complete, the minimization of the final likelihood function is performed using the Sequential Least Squares Programming algorithm described in [130]. The implementation of this algorithm is provided by the iMINUIT software package [131].

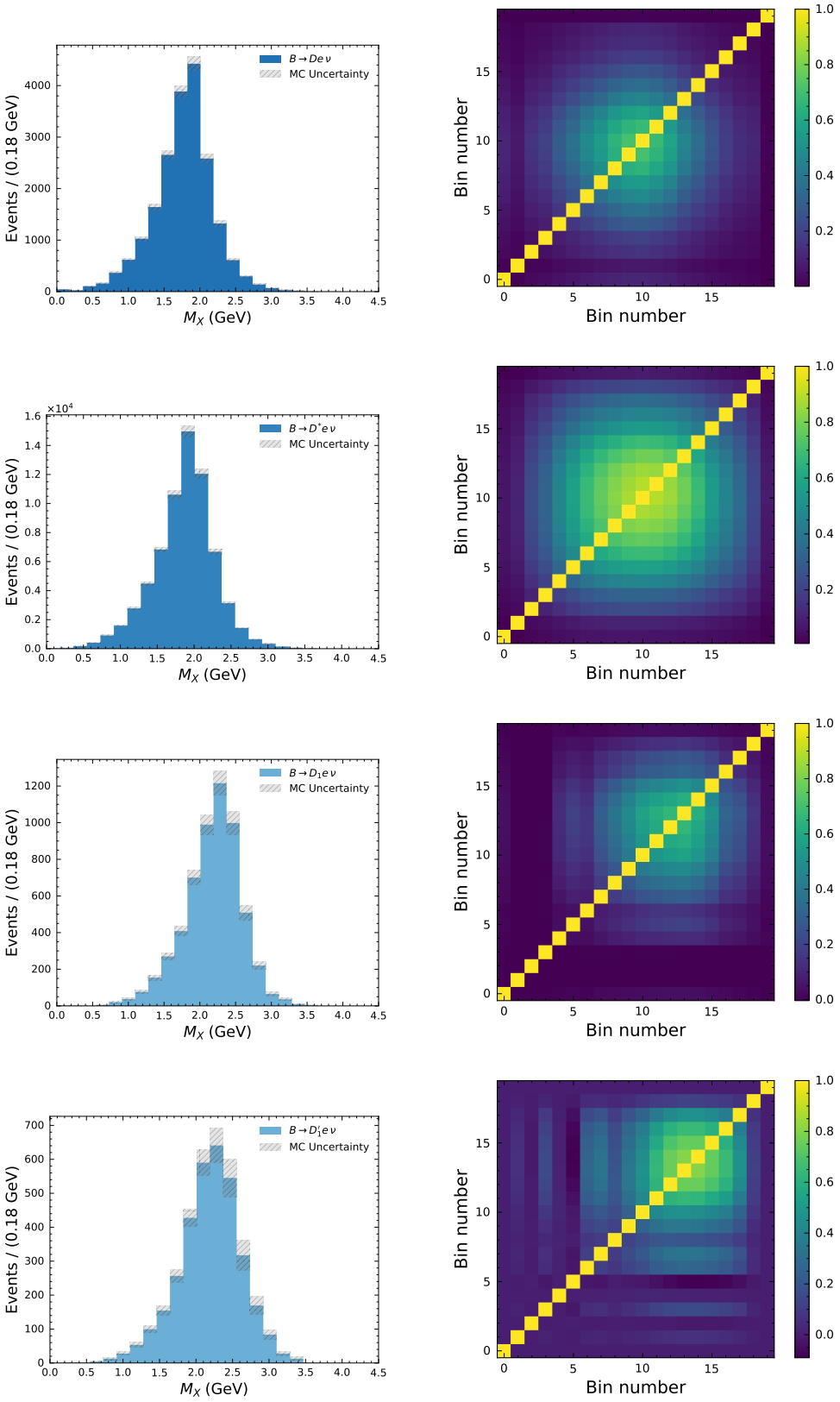

FIGURE 5.3: Template histograms for signal processes together with the corresponding error correlation matrix, calculated on MC samples. The MC uncertainties include uncertainties from MC statistics, form factor corrections, lepton and hadron identification efficiency corrections, and the tracking efficiency correction.

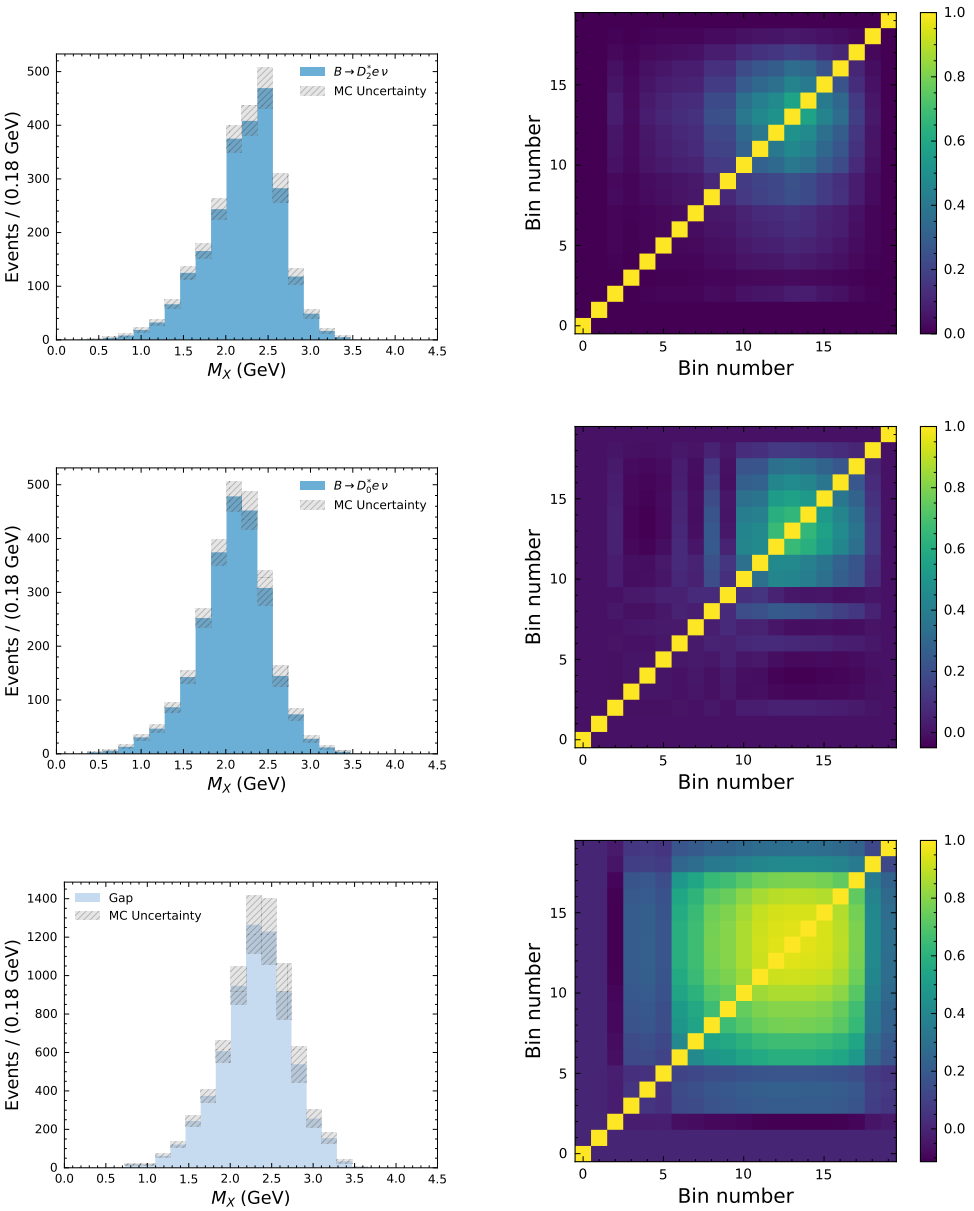

FIGURE 5.4: Template histograms for signal processes together with the corresponding error correlation matrix, calculated on MC samples. The MC uncertainties include uncertainties from MC statistics, form factor corrections, lepton and hadron identification efficiency corrections, and the tracking efficiency correction.

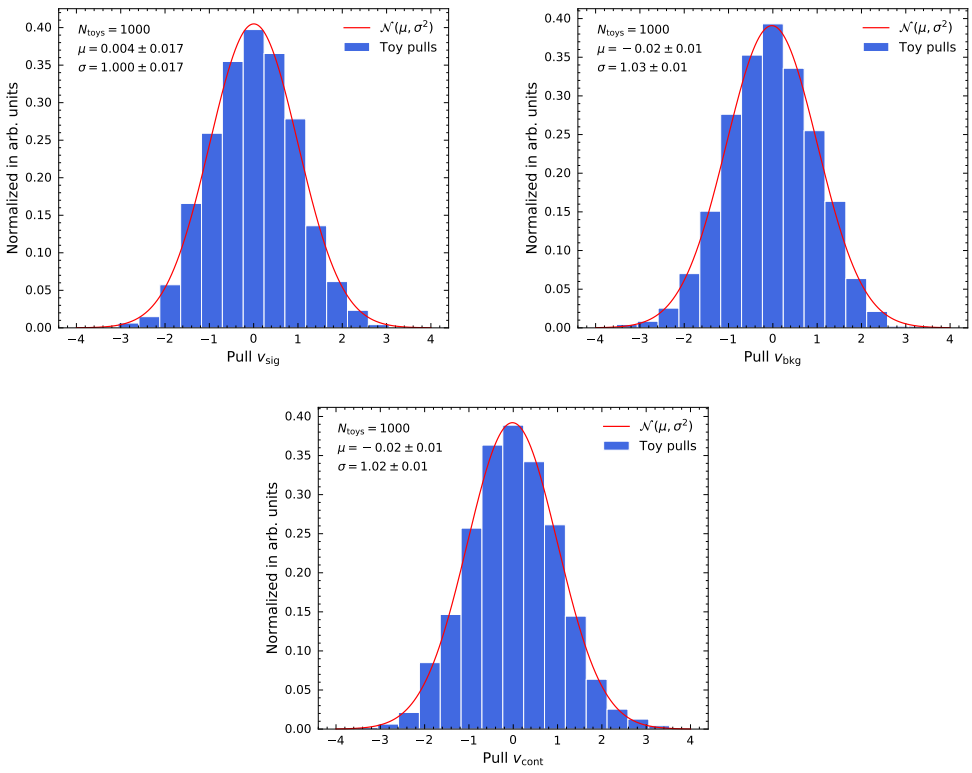

FIGURE 5.5: Normalized pull distributions for each of the fit yield parameters determine with $N = 1000$ pseudo-experiments. The mean and variance are obtained by fitting a Gaussian curve to the distributions.

5.2 Fit validation & results

5.2.1 Toy experiments

Various pseudo-experiments are performed to not only test the stability of the fit model, but also uncover potential bias in the extraction of the yields. To this end, pseudo-data sets are generated by sampling from a Poisson distribution for each bin, with the expectation value set to the MC expectation, in order to simulate statistical fluctuations. The fit procedure is then performed on each pseudo-data set, while extracting the pull for each fitted parameter. For a given parameter v in the fit, the pull is given by:

$$\text{Pull} = \frac{v^{\text{fit}} - v^{\text{expected}}}{\sigma_v^{\text{fit}}} . \tag{5.20}$$

For a completely unbiased fit, the distribution of pulls should be described by a unit normal distribution, i.e. a Gaussian distribution with $\mu = 0$ and $\sigma = 1$. A pull distribution that deviates from the mean and standard deviation of a unit normal distribution suggests a bias in the extraction of the yield parameters, or an incorrect estimation of the extracted yield uncertainty. Figure 5.5 shows the normalized pull distributions for each of the yield parameters in the fit, while the mean and variance are determined by fitting a Gaussian curve to the distributions. Neither the extracted yield parameters nor the estimated parameter errors exhibit any observed bias.

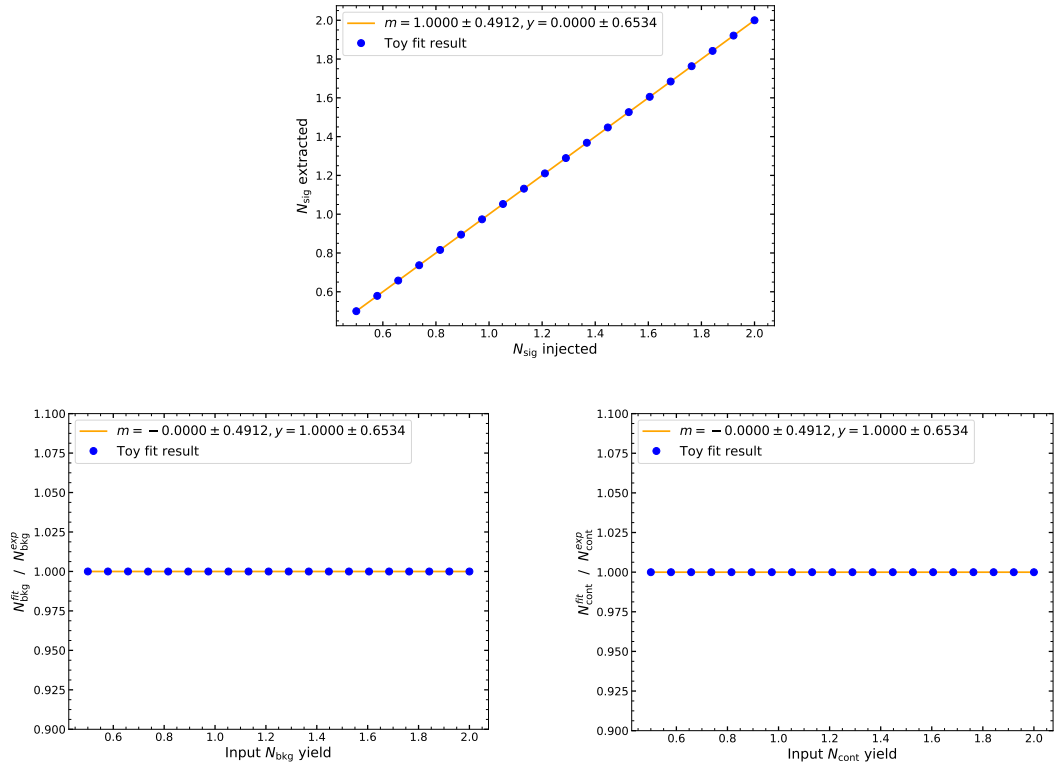

FIGURE 5.6: Linearity test results for the signal extraction (top), and the results to test for possible correlations between the signal extraction and the background normalization (bottom). A linear polynomial is fitted to the extracted yield parameters and results for the slope m and intercept y are shown.

A similar strategy is employed to perform a linearity test to determine whether any correlation between the expected and extracted values exist. A stable fit is verified by demonstrating that the fitted yield parameter depends linearly with gradient, $m = 1$, on the true parameter yield that was initially used as input. Subsequently, the linearity test is performed on pseudo-experiments while scaling the expected signal yield parameter from half to double the nominal yield value. In addition, possible bias between the extracted signal yield and the normalization of the background components is investigated by scaling the expected background yield parameters by the same incremental steps. Figure 5.6 shows the scaled expectation values for the signal and background yield parameters, together with the extracted signal values for each pseudo-experiment. Linear polynomials are fitted to the extracted yield parameters. The fitted slopes and intercepts for each test indicate an unbiased and stable fit.

The fit procedure is validated by carrying out the fit on a so-called Asimov data sample. Here, the measured data corresponds exactly to the MC expectation, leading to a perfect fit result in the case of a correct fit setup. The successful fit result in Figure 5.7 verifies that the Asimov data points in each bin matches exactly with the expected number of MC events. In addition, the bin-wise pull distribution, which is calculated in each bin as the difference between the expectation and the fitted prediction divided by the total error, is shown in the bottom panel and is consistently zero as expected. The uncertainty of a nuisance parameter gives an indication of the fit's dependence on a particular nuisance parameter as well as the parameter's

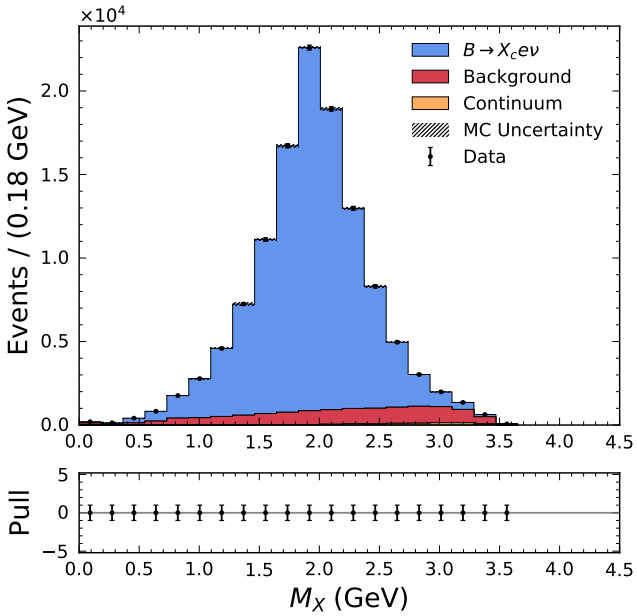

FIGURE 5.7: Post-fit M_X distribution after fitting the templates to an Asimov data sample.

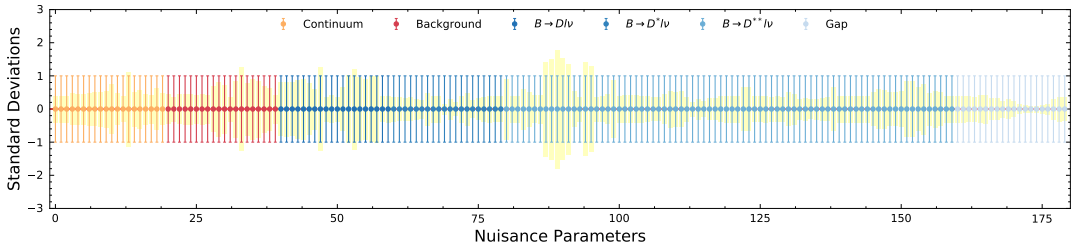

FIGURE 5.8: Nuisance parameters for each of the signal and background templates after performing the fit on an Asimov data sample. The error bars correspond to the initial 1σ error assigned to the nuisance parameters, while the yellow shading corresponds to the post-fit errors.

contribution to the overall uncertainty. Furthermore, the uncertainty signifies how the observed data constrains a specific nuisance parameter. In the absence of any constraints, nuisance parameters have uncertainties of exactly one standard deviation, whereas a tight or loose constraint from data is reflected by a small or large uncertainty. Figure 5.8 shows the parameter pulls for the nuisance parameters of each signal and background template. An upward or downward 1σ pull in the bin for a given template is represented by a bin parameter value of 1 or -1. For an Asimov fit, no significant pull on any of the nuisance parameters should be observed, as verified in Figure 5.8.

5.2.2 Fit results on data

The reconstructed MC and data M_X distributions are shown in Figure 5.9, together with the post-fit distribution of the previously discussed three component template

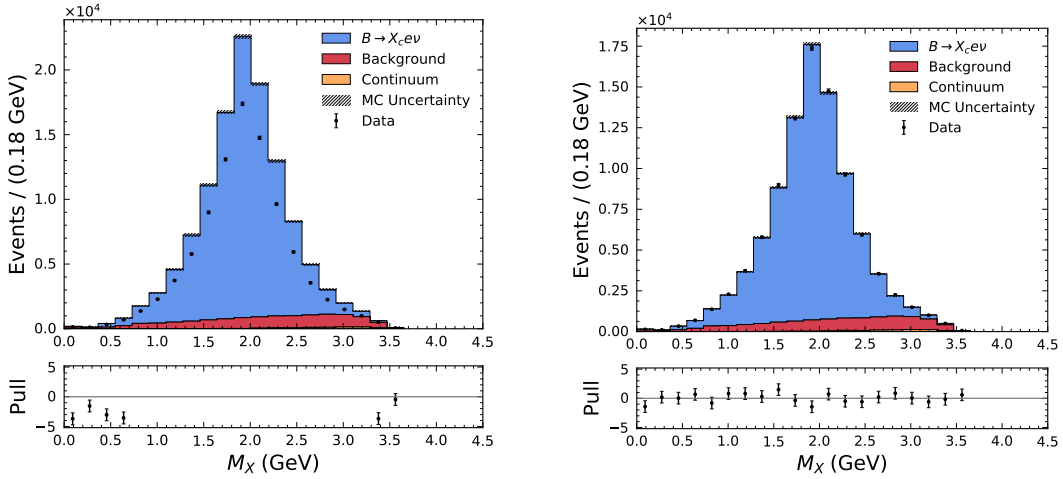

FIGURE 5.9: The reconstructed, pre-fit MC and data M_X distributions (left) and the post-fit distribution of the three component template fit (right).

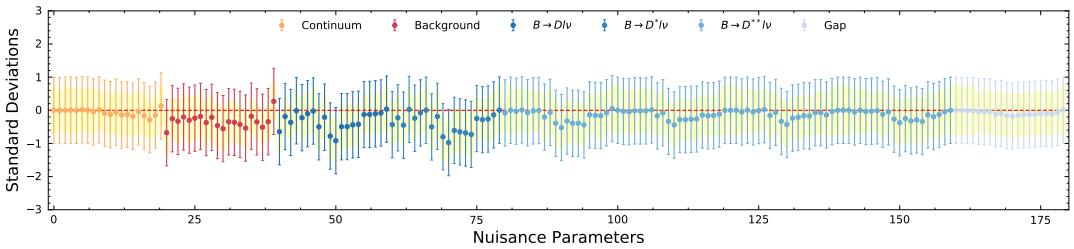

FIGURE 5.10: Nuisance parameters for each of the signal and background templates after performing the fit on data with the nominal selection criteria. The error bars correspond to the initial 1σ error assigned to the nuisance parameters, while the yellow shading corresponds to the post-fit errors.

fit. An initial mismatch between data and MC in the reconstructed, pre-fit distribution is observed, which is due to the different reconstruction efficiencies of the FR in data and MC. The MC uncertainty shown in the pre-fit distribution includes all of the previously discussed sources of error for each of the signal and background templates, while the error bars for the post-fit distribution are given by the global post-fit covariance matrix. The bin-wise pulls for the post-fit distribution, shown in the bottom panel, do not exceed 2σ, which indicates a good consistency between data and the fitted result. Figure 5.10 shows the parameter pulls for the nuisance parameters of each signal and background template. The nuisance parameters are generally consistent with 0σ, with slight downward pulls for the $B \to Dl\nu$ and $B \to D^* l\nu$ templates. These pulls, however, do not exceed 1σ and the overall well-behaved pull on the nuisance parameters shows that the fit templates describe the data adequately.

By using normalizations of the continuum and remaining background processes, obtained from the fit results, bin-wise signal probabilities $w(q^2)$ can be calculated in the next section.

5.3 Signal probability weights

The moments of the measured q^2 distribution are calculated as a weighted mean where the event-wise weight function $w_i(q^2)$ represents the probability for a given event with a certain value of q^2 to be a signal event. Therefore, the event-wise signal probability function ensures the effective subtraction of residual background events in the recorded data. To this end, the estimated number of residual background and continuum events, obtained from the normalizations of the fitted M_X distribution, are used to construct binned signal probability weights as a function of q^2. These weights are determined for each q^2 threshold selection and are defined in the following manner:

$$w_i(q^2) = \frac{N^i_{\text{Total}} - N^i_{\text{Bkg}}}{N^i_{\text{Total}}}, \tag{5.21}$$

where N^i_{Total} is the total amount of events in a certain bin i, and N^i_{Bkg} is the sum of continuum processes and residual background components in the same bin. To ensure independence from binning effects, the binned signal probabilities for each q^2 threshold selection are fitted with a polynomial function of a given order n to obtain continuous signal probability functions in q^2. The polynomial function is determined by performing a least squared χ^2 minimization, while the order of the polynomial is determined using a nested hypothesis test. A polynomial of order n over $n-1$ is accepted in the case where the improvement in the final calculated χ^2 is larger than one. Shape uncertainties of the two background distributions are propagated directly into the χ^2 minimization by calculating the total systematic covariance matrix, as previously described Section 5.1.1. In addition, the uncertainty due to the background normalization is taken into account by estimating the total relative error of the bin-wise signal probability weights and taking the outer product of the bin-wise uncertainty vector to obtain the associated covariance matrix. The calculation of this covariance matrix is briefly outlined below.

Since the fitted yields of the continuum and residual background are correlated, the post-fit covariance matrix is used to determine the uncertainty on the total background yield with the given formula:

$$\sigma_{\text{Total Bkg}} = \sqrt{\sigma^2_{\text{Cont}} + \sigma^2_{\text{Bkg}} + 2 \cdot \rho \cdot \sigma_{\text{Cont}} \sigma_{\text{Bkg}}}\,, \tag{5.22}$$

where ρ is the non-zero correlation coefficient given by the post-fit covariance matrix. Next, the relative background error, $r = \sigma_{\text{Total Bkg}} / N_{\text{Total Bkg}}$, is calculated which in turn is used to determine modified background subtraction weights:

$$\tilde{w}_i(q^2) = \frac{N^i_{\text{Total}} - N^i_{\text{Bkg}}(1+r)}{N^i_{\text{Total}}}\,. \tag{5.23}$$

Here, the bin-wise errors are taken as the difference between the modified and nominal background subtraction weights. Subsequently, the associated covariance matrix is calculated by taking the outer product of the bin-wise error vector.

Figure 5.11 shows the extracted q^2 distribution in data, together with the estimated residual background from the fitted M_X distribution for $B \to X_c\, e\, \nu$ and $B \to X_c\, \mu\, \nu$ decays. Furthermore, the corresponding bin-wise signal probability weights with

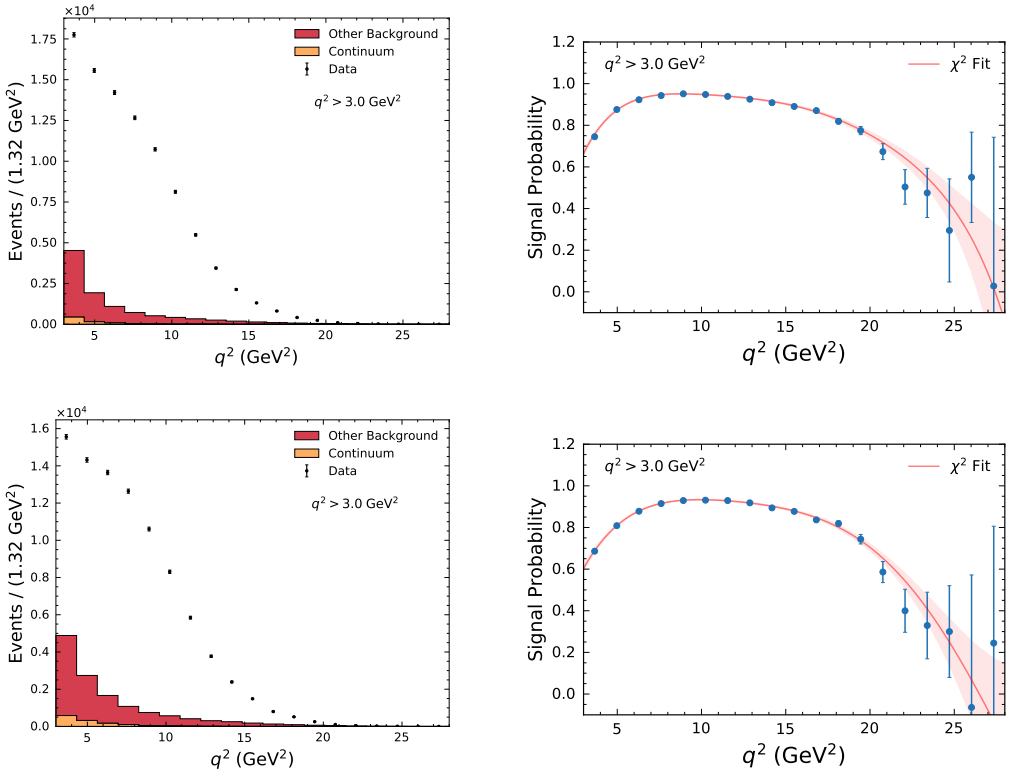

FIGURE 5.11: The extracted post-fit background and data M_X distributions (left) for $B \to X_c\, e\, \nu$ (top) and $B \to X_c\, \mu\, \nu$ (bottom) decays. The corresponding bin-wise signal probability weights, $w_i(q^2)$ together with a fitted polynomial of a given order n (right) are displayed for $q^2 > 3.0\,\text{GeV}^2$.

the fitted polynomial curves are also shown. The resulting polynomials $w(q^2)$ are required to have positive or zero event weights. Therefore, negative values at the end-points of the q^2 spectrum, where mainly background events contribute, are set to zero to avoid unrealistic probability weights. The signal probability functions for all q^2 threshold selections between $q^2 > 0.0\,\text{GeV}^2$ and $q^2 > 10.0\,\text{GeV}^2$ for progressively increasing incremental steps of $0.5\,\text{GeV}^2$ are found in Appendix C. These probability curves are utilised in the following section to complete the calculation of the q^2 moments.

Chapter 6

Measurement of q^2 moments

This section describes the developed methodology to calculate the moments of the q^2 distribution $\langle q^n \rangle$ with $n = 2, 4, 6, 8$. The moments are measured as a set of threshold selections on the reconstructed q^2 distribution that start at $q^2 > 3.0\,\text{GeV}^2$ and progressively increase with incremental steps of $0.5\,\text{GeV}^2$ up to $q^2 > 10.0\,\text{GeV}^2$. In addition, the lower q^2 threshold selections between $0.0\,\text{GeV}^2$ and $2.5\,\text{GeV}^2$ are also studied to verify the analysis strategies employed to calibrate and extract the measured moments. The studies concerning these low q^2 threshold selections are found in Appendix I.

The background subtracted moments of order n are calculated as a weighted mean summing over all selected events:

$$\langle q^n \rangle = \frac{\sum_i w_i(q^2)(q^n_{\text{calib},i})}{\sum_i w_i(q^2)} \times \mathcal{C}_{\text{cal}} \times \mathcal{C}_{\text{acc}}\,, \tag{6.1}$$

where $w_i(q^2)$ are the signal probability weights calculated as a continuous function of q^2, $q^n_{\text{calib},i}$ is the event-wise calibrated q^2 value to the power of n, and $\mathcal{C}_{\text{calib}}$ and $\mathcal{C}_{\text{true}}$ denote additional correction factors. The reconstructed q^2 distribution is distorted by missing charged and neutral particles in the reconstruction of the hadronic tag-side and by finite detector acceptance and resolution effects. To correct for reconstruction and selection effects and ensure an unbiased measurement, the $\langle q^n \rangle$ moments are unfolded with correction methods based on generator-level MC samples. The selected calibration method has yielded reliable results in past measurements of the moments of the hadronic mass, and combined mass-and-energy spectra of $B \to X_c \ell \nu$ decays, found in [132, 133]. Consequently, this calibration method is adjusted and tested to meet the requirements of this analysis. Both the calibration method and the calculation of the additional correction factors, which further correct the moments from selection and acceptance effects, are described in Section 6.1. Various sources of statistical and systematic uncertainties are introduced by the background subtraction and calibration procedures and affect the final measured $\langle q^n \rangle$ moments. The evaluation of the statistical and systematic uncertainties considered in the measurement is discussed in Section 6.2. The measurement of the $\langle q^n \rangle$ moments of $B \to X_c \ell \nu$ decays are presented separately for electron and muon candidates in Section 6.3, allowing for a test of lepton flavour universality. Lastly, Section 6.4 describes the interpretation of the measured $\langle q^n \rangle$ moments in the framework of the HQE. A preliminary, simultaneous fit is performed to the theoretical predictions based on the reduced set of HQE parameters at $\mathcal{O}(1/m_b^4)$ and a value of $|V_{cb}|$ is subsequently extracted. Unless otherwise stated, all calibration studies shown in the following sections are calculated on simulated signal decays for electron candidates, while the calibrations derived for muon candidates are given in Appendix D.

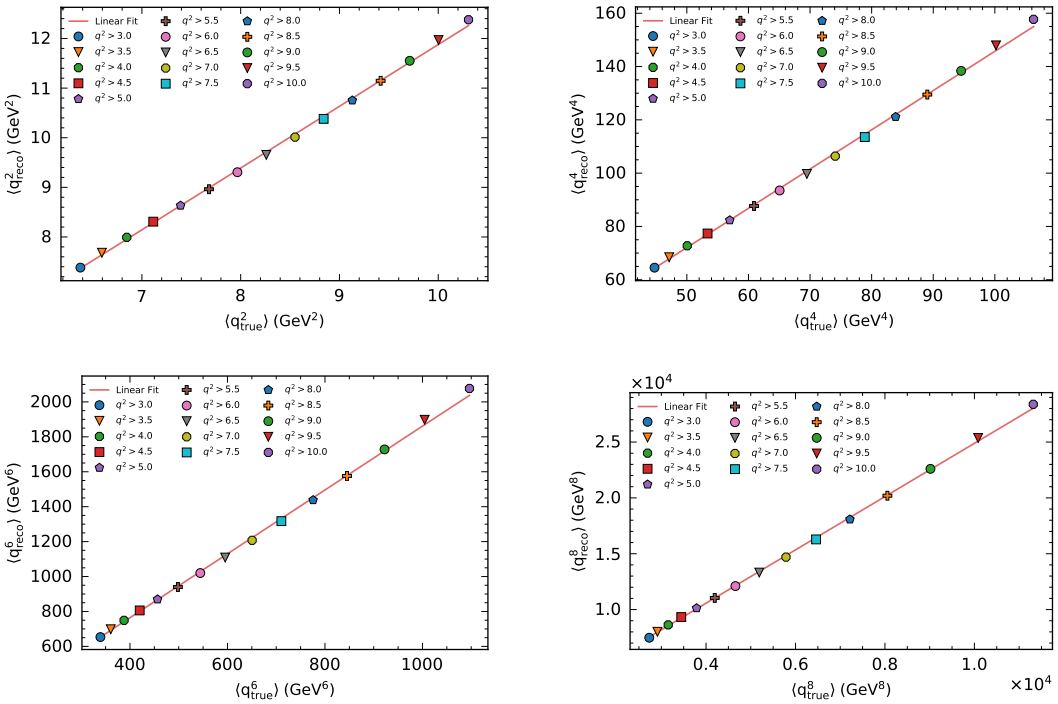

FIGURE 6.1: Calibration curves for $\langle q^{2,4,6,8} \rangle$ for electron candidates. The extracted $\langle q^n_{,\text{reco}} \rangle$ versus $\langle q^n_{,\text{true}} \rangle$ as a function of q^2 threshold selections together with the fitted linear calibration curves are shown.

6.1 Calibration

The main goal of the calibration method is to correct the measured moments of the selected data sample to the underlying true moments. This is achieved by exploiting the linear relationship between the moments of the reconstructed q^2 distribution $\langle q^n_{,\text{reco}} \rangle$ and the true distribution $\langle q^n_{,\text{true}} \rangle$. Therefore, each calibration curve is based on the following linear approximation:

$$q^n_{\text{reco}} = c_n + m_n \cdot q^n_{\text{true}} \qquad \text{for } n = 2, 4, 6, 8, \tag{6.2}$$

where m_n and c_n are parameters of the linear curve that are determined for each order of moment. The final calibration is subsequently applied as an event-by-event linear correction function, which is given as a function of the reconstructed q^n. Furthermore, two additional global correction factors are introduced to correct for any residual bias as well as selection and acceptance effects. The parameters in the function shown in Eqn. 6.2 are determined by utilising a separate, independent $B \to X_c \ell \nu$ MC sample, which corresponds to one stream of the data set recorded by the Belle detector.

For each q^2 threshold selection, the true underlying $\langle q^n_{,\text{true}} \rangle$ and the reconstructed moments $\langle q^n_{,\text{reco}} \rangle$ are calculated on an independent simulated signal MC sample. A strong linear relationship is found for all orders of moments under study and linear calibration functions are derived by fitting linear curves to the calculated moments. Therefore, this method leads to the extraction of exactly one calibration curve for

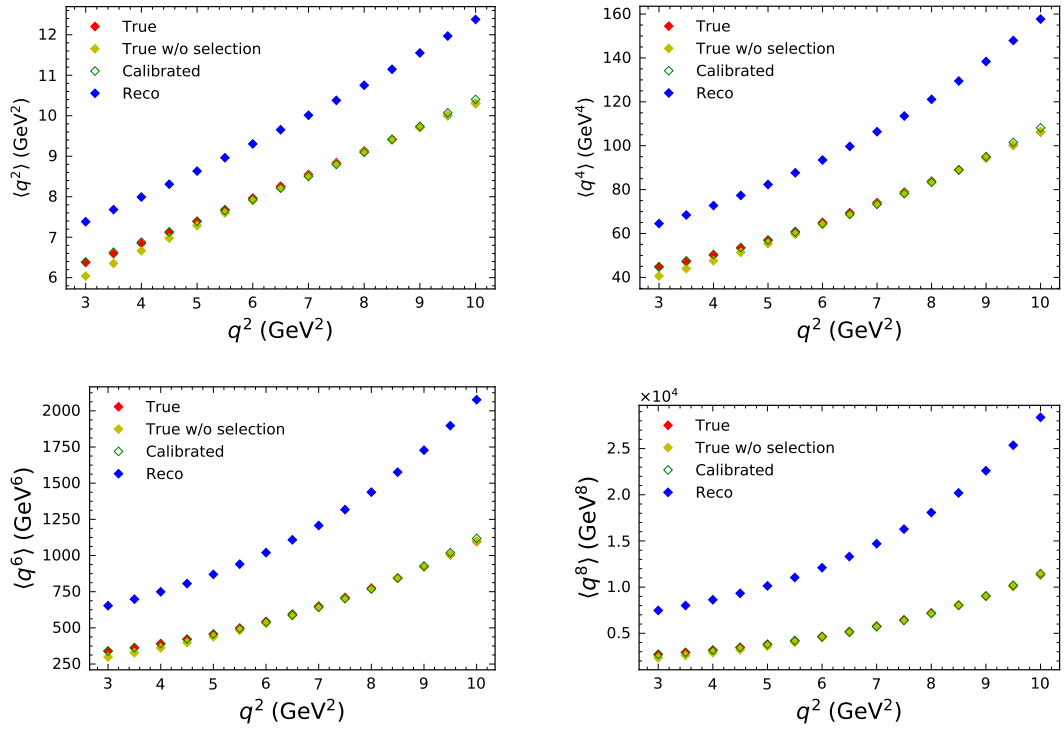

FIGURE 6.2: Moments of q^n calculated on simulated $B \to X_c e\nu$ decays for different q^2 threshold selections. The reconstructed, calibrated and true moments after all selection criteria have been applied are compared. Additionally, the true moments without any selection criteria are shown as the yellow diamonds.

each order of $\langle q^n \rangle$, reducing the possible statistical error due to the calibration functions. The extracted moments for $\langle q^{2,4,6,8} \rangle$ are shown in Figure 6.1, together with the fitted calibration curves. Due to a greater presence of final state radiative photons in the electron channel, a small difference between the fitted slope and intercept of the obtained calibration curves for electron and muon candidates is observed.

After obtaining a set of calibration curves for each order of moment $\langle q^n \rangle$, the reconstructed q^2 distribution is subsequently corrected by making use of the linear calibration function in Eqn. 6.2 to calculate the calibrated values q^2_{calib}. For each event with a given reconstructed q^2, the corresponding calibrated value is determined by simply inverting the linear calibration function:

$$q^n_{\text{calib}} = \frac{(q^n_{\text{reco}}) - c_n}{m_n}, \tag{6.3}$$

where $n = 2, 4, 6, 8$, and m and c denote the fitted slope and intercept of the linear correction function, respectively. The performance of this calibration method is shown in Figure 6.2, where the reconstructed, calibrated and true underlying moments are compared as a function of q^2 threshold selections. These moments are calculated on the same independent simulated signal sample used to derive the linear calibration functions.

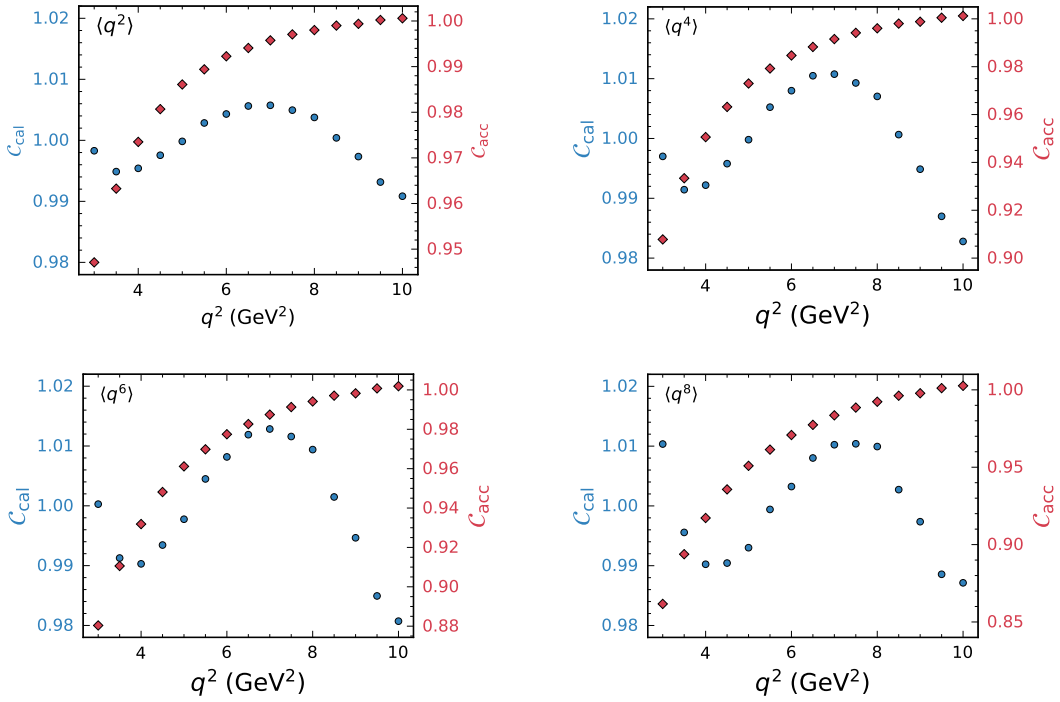

FIGURE 6.3: Remaining bias after application of the calibration curves on simulated $B \to X_c e\nu$ decays as a function of q^2 threshold selection. Shown here are the bias correction factors $\mathcal{C}_{\text{cal}}$ (blue) and $\mathcal{C}_{\text{acc}}$ (red).

6.1.1 Residual bias corrections

Even though the above-mentioned calibration method corrects most of the reconstructed moments to the underlying true moments, small deviations still remain after the correction procedure. The observed bias is due to imperfections in the calibration method, i.e. small non-linearities that the extracted curves do not account for. Furthermore, the remaining bias increases for higher q^2 threshold selections due to limited statistics in these regions of phase space. In order to eliminate any possible residual bias, sets of correction factors for each q^2 threshold selection are determined, while using the same independent simulated signal sample discussed in the previous section. Two distinct factors are considered: $\mathcal{C}_{\text{cal}}$ corrects the remaining bias due to imperfections in the calibration method, while $\mathcal{C}_{\text{acc}}$ corrects for a possible intrinsic bias introduced by finite detector resolution as well as the chosen signal selection criteria.

The correction factor $\mathcal{C}_{\text{cal}}$ accounts for the observed residual bias after applying the derived calibration functions to the reconstructed q^2 distribution to obtain calibrated values. A global correction factor is calculated for each individual q^2 threshold selection and each order of moment as follows:

$$\mathcal{C}_{\text{cal}} = \frac{\langle q^n_{\text{true}} \rangle}{\langle q^n_{\text{calib}} \rangle} \,. \tag{6.4}$$

The determined correction factors vary greatly depending on the specific q^2 threshold selection, with the largest observed bias found for higher q^2 regions. In these

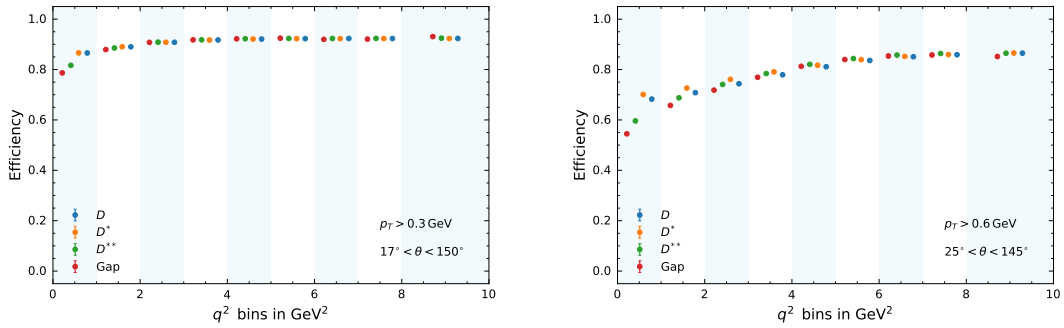

FIGURE 6.4: Selection efficiencies of different $B \to X_c e\nu$ (left) and $B \to X_c \mu\nu$ (right) components with basic transverse momentum and angular requirements in bins of q^2. The error bars include only the uncertainty due to limited MC statistics. Different q^2 bins are indicated by the alternating shaded regions.

regions of phase space, the limited statistics as well as the poor q^2 resolution contributes to a large residual bias after performing the calibration procedure. Figure 6.3 shows the calculated correction factors, where the overall residual bias is observed to be less than a percent for the first moment $\langle q^2 \rangle$. On the other hand, the largest residual bias is found for the higher moments, where the calibrated and true underlying moments differ by approximately 2%.

Since the calibration curves are derived by calculating the reconstructed $\langle q^n_{\text{reco}} \rangle$ and true $\langle q^n_{\text{true}} \rangle$ moments after applying all signal selection criteria, the true moments are affected by an intrinsic selection bias. This effect is observed in Figure 6.4, which shows the selection efficiencies in bins of q^2 for different exclusive $B \to X_c \ell\nu$ decays with the basic angular and transverse momentum requirements outlined in Table 4.2. Different selection and acceptance efficiencies are observed for different $B \to X_c \ell\nu$ final states, especially for the poorly constrained D^{**} and non-resonant decays in the low q^2 region. Therefore, the calculated underlying true moments are biased with respect to the true moments calculated on events selected only with a true q^2 threshold selection. In addition, theoretical interpretations are based on calculations that do not account for photons emitted by final state radiation. However, these photons are not only simulated in MC, but also present in data. Consequently, this additional, albeit tiny, effect as well as detector resolution and selection effects are considered by applying a global correction factor $\mathcal{C}_{\text{acc}}$ for each q^2 threshold selection and order of moment. The bias correction factor $\mathcal{C}_{\text{acc}}$ is calculated on simulated signal samples and is defined as:

$$\mathcal{C}_{\text{acc}} = \frac{\langle q^n_{\text{true without cut}} \rangle}{\langle q^n_{\text{true}} \rangle} \,. \tag{6.5}$$

Here, $\langle q^n_{\text{true}} \rangle$ is the underlying true moment calculated after all selection criteria have been applied, while $\langle q^n_{\text{true without cut}} \rangle$ is the true moment calculated without any selection requirements except a true q^2 threshold selection. The true moments $\langle q^n_{\text{true without cut}} \rangle$ are calculated by making use of a dedicated generator-level signal sample that does not include effects due to final state radiation. Figure 6.3 shows the calculated $\mathcal{C}_{\text{acc}}$ correction factors for each considered q^2 threshold selection and

moment order. The observed factors are once again clearly dependent on the specific q^2 threshold selection: the poorly constrained non-resonant decays are more abundant in low regions of the q^2 spectrum, leading to an estimated bias correction factor of approximately 5% for the first moments $\langle q^2 \rangle$ that increases for higher orders. One the other hand, the well-understood $D\ell\nu$ and $D^*\ell\nu$ decays dominate in the high regions of q^2 phase space resulting in smaller corrections tending toward unity for higher q^2 threshold selections. A similar trend in the calculated $\mathcal{C}_{\text{acc}}$ correction factors is reported in [134] where a similar methodology is utilized to measure the hadronic mass moments of inclusive $B \to X_c\ell\nu$ decays.

6.1.2 Verification of the calibration procedure

In order to verify the full background subtraction and calibration methodology, sets of calibrated and background subtracted q^2 moments are calculated on ensembles of independent simulated samples to determine any bias in the extraction method. The verification procedure is briefly outlined below:

1. Linear calibration functions are derived by making use of an independent sample of signal MC corresponding to one stream of the data set recorded by the Belle detector.

2. An additional, independent sample of generic MC containing both signal and background processes is divided into two equal samples. Both the continuous signal probability weights and the bias correction factors are calculated by using one of the halved samples.

3. The remaining halved sample is treated as "mock data" and is used to test the moment extraction method. Calibrated and background subtracted q^2 "mock moments" are calculated using Eqn. 6.1. The above-mentioned linear calibration functions and the additional bias correction factors, $\mathcal{C}_{\text{cal}}$ and $\mathcal{C}_{\text{acc}}$, are applied to the q^2 distribution in combination with the Asimov signal probability weights.

This study tests the complete extraction formalism; if the true generator-level moments are obtained after the background subtraction and calibration steps, then the procedure can be trusted to produce reliable results. In order to determine any additional remaining bias due to the extraction method, the bootstrapping algorithm [135, 136] is implemented. A bootstrapped sample is determined by repeatedly sampling the nominal data set with replacement, repeating the verification procedure described above and calculating new bootstrapped moments $\langle q^n_{\text{mock}} \rangle$. A sample size of 500 pseudo-data sets is used to evaluate the relative error of the remaining bias, which is calculated as:

$$\sigma_{\text{non-closure}} = \frac{\langle q^n_{\text{mock}} \rangle - \langle q^n_{\text{gen}} \rangle}{\langle q^n_{\text{gen}} \rangle}, \tag{6.6}$$

where $\langle q^n_{\text{gen}} \rangle$ denotes the true generator-level moments. Figure 6.5 shows box plots of the resulting relative errors for electron candidates. The box plots show the standard 0.25%, 0.5%, and 0.75%-percentiles, while the whiskers correspond to the standard 1.5 times the interquartile range. While a small remaining bias is still observed, it is not considered to be statistically significant compared to the leading sources of statistical and systematic uncertainties, further discussed in Section 6.2. The observed residual bias is possibly introduced by imperfections in the interpolation of

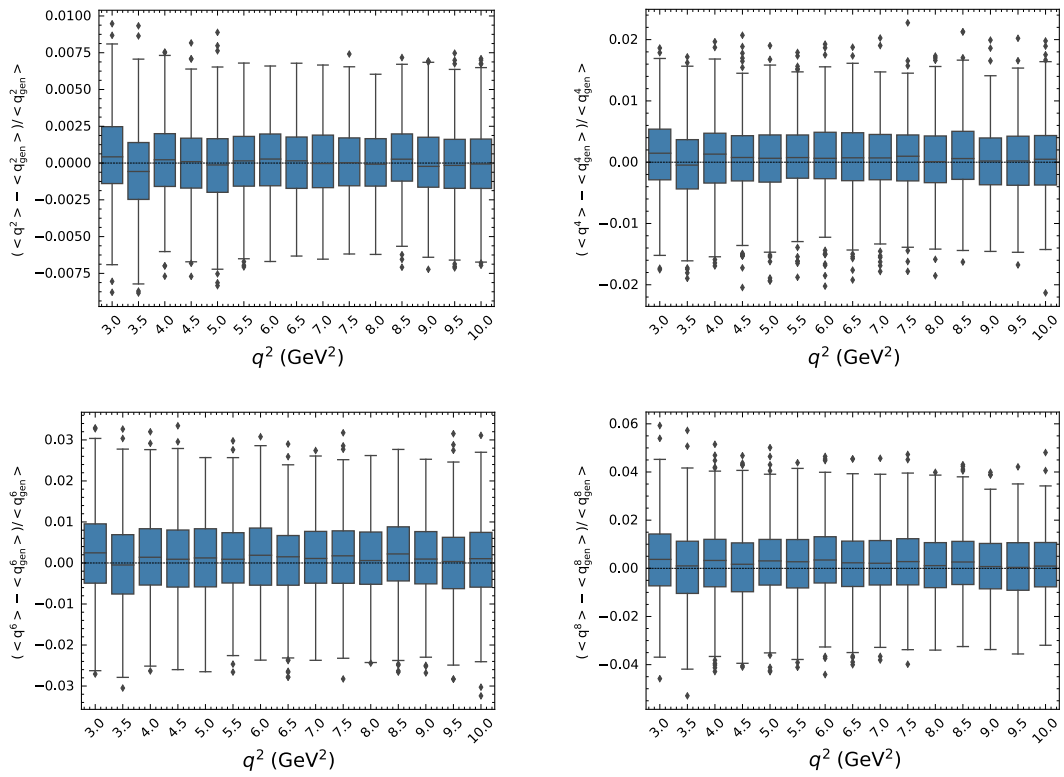

FIGURE 6.5: Box plot of the sample test relative error distributions for q^n moments as a function of q^2 threshold selections for electron candidates.

the bin-wise signal probabilities. To obtain a quantitative estimate of the bias, the mean value of the relative error distribution is estimated and added as an additional non-closure uncertainty on the final measured moments. Overall this test gives us confidence that the analysis strategy yields unbiased unfolded q^n moments.

6.2 Statistical and systematic uncertainties

Various sources of statistical and systematic uncertainty arising from the background subtraction and calibration procedures are estimated in this section. The total estimated error on the final measured moments $\langle q^n \rangle$ is obtained by adding the individual uncertainty sources in quadrature. Tables 6.1 and 6.2 summarise the relative statistical and systematic uncertainties on the measured moments $\langle q^n \rangle$ for both electron and muon candidates, given in permille.

6.2.1 Statistical error on extracted moments

The statistical error on the extracted moments, which is directly related to the amount of observed events in simulated samples and data, is determined as the variance of a weighted mean in the following manner:

$$\begin{aligned} V(\langle q^n \rangle) &= V\left(\frac{\sum w_i(q^2)(q^n)}{\sum_i w_i(q^2)}\right) \\ &= \frac{1}{\left(\sum_i w_i(q^2)\right)^2} V\left(\sum w_i(q^2)(q^n)\right) \\ &= \frac{\sum_i w_i(q^2)^2}{\left(\sum_i w_i(q^2)\right)^2} V(q^n) \\ &= \frac{\sum_i w_i(q^2)^2}{\left(\sum_i w_i(q^2)\right)^2} \langle q^{2n} \rangle - \langle q^n \rangle^2 \,. \end{aligned} \tag{6.7}$$

Here, the overall statistical variance is given by the calibrated, but unweighted, q^n distribution. Since the errors due to the background subtraction, calibration and bias correction factors are estimated separately as systematic effects, the treatment of the probability density weights, $w_i(q^2)$, as constants in the aforementioned equation is sound.

6.2.2 Statistical error on the signal probability weights

Since the continuous signal probability functions are determined by fitting the binned signal probabilities for each q^2 threshold selection with a polynomial function of a given order n, each of the fitted coefficients have correlated statistical errors. In order to propagate the correlated uncertainties correctly, the eigenvectors and eigenvalues of the post-fit covariance matrix are calculated to extract orthogonal variations corresponding to each of the fitted function's coefficients. The varied parameters, $c_i^{\pm}$, are calculated by making use of Eqn. 3.15, where i corresponds to the number of eigen variations for a given polynomial function.

To estimate the impact of the uncertainties on the fitted parameters, the calculation of the measured moments $\langle q^n \rangle$ is repeated for each of the varied polynomial coefficients. An uncertainty for each variation is then estimated as follows:

$$\sigma_{\langle q^n_{c_i} \rangle} = \frac{|\langle q^n_{c_i^+} \rangle - \langle q^n_{c_i^-} \rangle|}{2} \,. \tag{6.8}$$

The total uncertainty due to the background subtraction weights is estimated by adding all individual uncertainties in quadrature.

6.2.3 Statistical error on the linear calibration functions

Similar to the previous error calculation, the fitted coefficients of the linear calibration curves also have correlated statistical errors. Therefore, varied parameters are again calculated by making use of Eqn. 3.15. In total, two independent variations of the fitted parameters are determined for this error source. The impact of the uncertainties on the fitted calibration curve is estimated by repeating the calculation of the moments $\langle q^n \rangle$ and estimating an uncertainty for each variation with Eqn. 6.8. The total uncertainty due to the determination of the calibration function is calculated in the same manner as described in the previous section.

6.2.4 Statistical error on the additional bias correction factors

The statistical uncertainty due to the additional bias correction factors is estimated by varying the bias correction factors within one standard deviation and repeating the moments $\langle q^n \rangle$ calculation. The resulting statistical uncertainty due to the global correction factors $\mathcal{C}_{\text{cal}}$ and $\mathcal{C}_{\text{acc}}$ is estimated as half the difference between the two varied moments.

6.2.5 Uncertainty due to the $B \to X_u \ell\nu$ branching fraction

As summarised in Table 3.3, the latest experimental uncertainty for the $B \to X_u \ell\nu$ branching fraction is estimated to be approximately 14% [30]. To propagate the uncertainty of the branching fraction into the measurement of the moments $\langle q^n \rangle$, the corresponding $B \to X_u \ell\nu$ simulated background component is rescaled with scaling factors given by:

$$s = s_{\text{nominal}} \pm \sigma_{\mathcal{B}} \,. \tag{6.9}$$

New varied signal probability curves are subsequently derived using the altered MC sample and the calculation of the moments is repeated. Finally, the resulting uncertainty is estimated in the same manner as described in the previous section.

6.2.6 Uncertainty due to the $B \to X_c \ell\nu$ composition

Since the developed methodology to calculate the moments of the q^2 distribution $\langle q^n \rangle$ relies very strongly on MC simulations, the final measured moments depend strongly on the choice of $B \to X_c \ell\nu$ modelling that was employed to simulate the signal decays. With the exception of the $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays, higher mass contributions of the full semileptonic decay width are scarcely known. Furthermore, the composition of the simulated signal MC affects the background normalization as well as the calibration of the measured moments. The systematic uncertainty related to the chosen signal model is evaluated by changing the assumed composition, i.e. by dropping each of the exclusive $B \to D^{**}\ell\nu$ states. In addition, two different methods are implemented to estimate the impact of the non-resonant contribution of the X_c system. The first approach involves completely dropping the $B \to D^{(*)}\pi\pi\ell\nu$ and $B \to D^{(*)}\eta\ell\nu$ states, while the second method replaces the current gap model with the scaled-up contributions of the two broad D^{**} states (D_1' and D_0^*). In addition, the systematic uncertainties due to the $B \to D\ell\nu$ and $B \to D^*\ell\nu$ branching fractions are estimated by varying the corresponding MC sample within the experimental uncertainty of the current branching fractions, shown in Table 3.3, while using the same scaling factors introduced in Eqn. 6.9. The simulation model-dependence of both the derived calibration curves and the residual bias corrections is studied as well by carrying out each of the subsequent analysis steps outlined in previous sections.

Since the calibration curves are also sensitive to the chosen signal model, the systematic uncertainty arising from this dependency is taken into account by calculating new calibration curves for each of the altered compositions of the signal model. Furthermore, new residual bias correction factors are also calculated with the modified signal decay models. Therefore, six different sets of calibration curves and bias corrections for the $B \to D^{**}\ell\nu$ and non-resonant states are derived, while four sets of varied calibration curves are calculated for $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays.

Subsequently, the estimated uncertainties due to the $B \to D\ell\nu$ and $B \to D^*\ell\nu$ branching fractions are calculated as described in previous sections, while uncertainties due to the modelling of the $B \to D^{**}\ell\nu$ and non-resonant states are estimated by calculating:

$$\sigma_{X_c\text{component}} = |\langle q^n_{,\text{nominal}}\rangle - \langle q^n_{,X_c\text{ component}}\rangle| \,. \tag{6.10}$$

Summaries of the uncertainties due to the modelling of each exclusive state for the moments $\langle q^{2,4,6,8}\rangle$ are found in Appendix E. The total systematic uncertainty due to the $B \to X_c\ell\nu$ composition is estimated by adding all individual uncertainties in quadrature.

6.2.7 Uncertainty due to the $B \to X_c\ell\nu$ modelling

The final measured moments not only depend on the $B \to X_c\ell\nu$ composition, but also on the choice of specific exclusive kinetic decay models. In order to take into account the uncertainty due to the choice of form factor parametrization for exclusive $B \to X_c\ell\nu$ decays, the event-wise form factor weights are varied with the sets of variation weights, described in Section 3.3.2, for each of the specific form factor parameters. New sets of calibration curves as well as bias correction factors are calculated for the each of the varied parameters associated with the $B \to Dl\nu$, $B \to D^*l\nu$, and $B \to D^{**}l\nu$ decays. The moments $\langle q^n\rangle$ calculation is subsequently repeated with the sets of varied calibration and bias correction factors. The error due to a given form factor parameter is estimated by making use of Eqn. 6.8. Summaries of the uncertainties due to the choice of form factor parametrization describing exclusive $B \to X_c\ell\nu$ decays for the moments $\langle q^{2,4,6,8}\rangle$ are also provided in Appendix E. The final uncertainty due to the choice of $B \to X_c\ell\nu$ form factor parameters is estimated by adding individual form factor parameter errors from each of the respective processes in quadrature.

6.2.8 Efficiency correction uncertainties

As mentioned in Section 3.3.2, the track finding efficiency has an associated systematic uncertainty of 0.35% per track due to the difference between the value in MC and data. Therefore, the error due to the tracking efficiency is estimated by varying the nominal tracking weight by 0.35% per track. Varied calibration curves are derived for the varied track efficiency correction weights, which in turn are used to calculate new varied sets of bias correction factors. The moments $\langle q^n\rangle$ are extracted with the new varied calibrations, and the total uncertainty due to the difference in track finding efficiency between MC and data is estimated by once again making use of Eqn. 6.8.

The systematic uncertainty due to lepton and hadron identification efficiency corrections is estimated by applying the sets of associated variation weights, described in Section 3.3.2. Varied calibration curves and new sets of bias correction factors are calculated for each of the variations leading to a total of 20 different calibrations per PID type. The analysis procedure is repeated in order to extract an ensemble of varied moments q^n, while the size of the uncertainty is evaluated by taking the RMS of the observed moment variations.

The reconstruction efficiencies of the Full Reconstruction algorithm, described in Section 4.2.1, differ for neutral and charged B mesons in data and simulated MC samples. This difference arises due to inaccurate modelling of the detector response

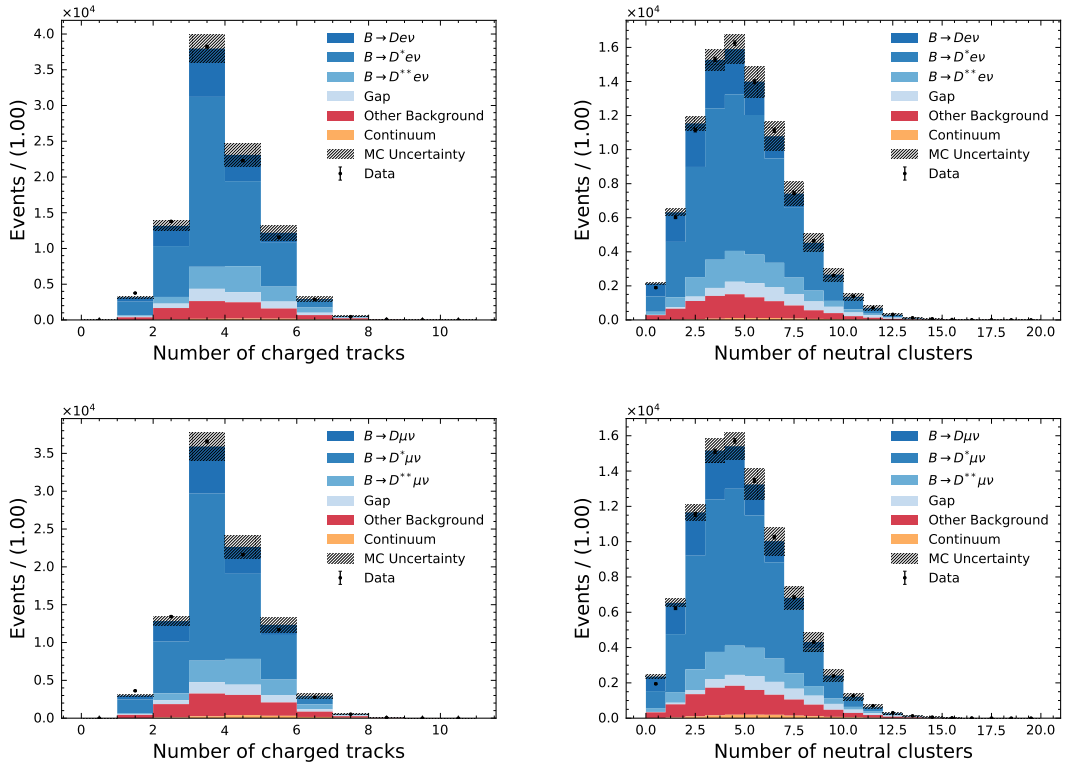

FIGURE 6.6: Comparisons between the data and MC distributions of the number of charged tracks and neutral clusters in the electron (top) and muon (bottom) channels.

or incorrect branching fractions of the exclusive B meson decay channels included in the reconstruction algorithm. The systematic uncertainty due to the $B^0/B^\pm$ reconstruction efficiency is taken into account by correcting for the difference in MC and repeating the calibration procedure to determine a varied set of q^2 moments. Subsequently, the full difference between the varied and nominal moments is estimated as the systematic uncertainty.

6.2.9 MC modelling of tracks and clusters in the X system

As can be seen in Figure 6.6, a slight disagreement between the expected and observed number of charged tracks and neutral clusters in the X system is observed. In order to estimate the impact of a potential mis-modelling of the X system resolution in MC, bin-wise weight functions are determined in a region where signal dominates: $M_X < 3\,\text{GeV}$ and $p_\ell > 1\,\text{GeV}$. The weights are calculated separately for electron and muon candidates as the bin-wise ratio of the normalized MC and data distribution. Since the number of charged tracks and neutral clusters are independent, separate sets of weight functions are calculated for each of these distributions and applied independently to the signal MC. Subsequently, new calibrations, bias correction factors and varied sets of moments are determined and the absolute value of the difference between the varied and nominal moments is taken as the estimated error.

Since the $E_{\text{miss}} - |\mathbf{p}_{\text{miss}}|$ distribution is not only used to select correctly reconstructed semileptonic $B \to X_c \ell \nu$ decays, but also to improve the resolution of the reconstructed q^2, a similar strategy as described above is employed to take into account any mis-modelling of the $E_{\text{miss}} - |\mathbf{p}_{\text{miss}}|$ distribution. The impact of this systematic uncertainty is estimated by calculating correction functions based on the interpolated bin-wise ratios between the normalized MC and data distributions, while applying the bin-wise weights derived for the number of charged tracks and neutral clusters to avoid double counting.

6.2.10 MC modelling of the q^2 distribution

To estimate an uncertainty on the modeling of the q^2 distribution, the reconstructed distribution is scaled by 1%. This shift in the q^2 distribution corresponds to the observed differences in the mean values of the energy and momentum of the reconstructed X system. The calibration procedure is repeated in order to redetermine calibration functions as well as bias correction factors, while varied sets of background subtraction weights are determined for the shifted distribution. Subsequently, varied sets of shifted moments are extracted and the uncertainty due to the modelling of the q^2 distribution is evaluated by taking the full difference to the nominal moments.

6.2.11 Statistical and systematic correlations

Since the determined moments $\langle q^n \rangle$ are measured with progressively increasing threshold selections on the q^2 distribution, moments of different cuts share a subset of events and are therefore statistically correlated. The correlations between the moments $\langle q^n \rangle$ for different q^2 threshold selections are important if these measurements are to be used to extract HQE parameters in a simultaneous fit to the theoretical moment predictions. This fit is beyond the scope of this work, however, measured correlations are included for completeness. The bootstrapping algorithm is implemented in order to determine the measured correlations between the moments for different q^2 threshold selections. Similar to the strategy employed in Section 6.1.2, the bootstrapped sample is determined by sampling the nominal data set with replacement and repeatedly calculating new values of $\langle q^n \rangle$. The correlation values are then calculated by making use of the Pearson correlation coefficient to estimate the correlation matrix of the measured moments $\langle q^n \rangle$. Figure 6.7 shows an example of the statistical correlations for electron candidates. Moments of the same order with similar selections on q^2 are highly correlated, while moments of higher order with identical q^2 selections contain more independent information as the difference in the order increases.

Additionally, a covariance for the lepton and hadron identification efficiencies is estimated in a similar manner as discussed above, however here the ensemble of varied moments described in Section 6.2.8 is used for the calculation. The MC non-closure uncertainty discussed in Section 6.1.2 is assumed to be uncorrelated across different moment orders as well as different q^2 threshold selections. Therefore, only a diagonal covariance matrix is constructed to describe this error source. For the remaining systematic uncertainties considered in previous sections, identical sources are fully correlated or anti-correlated across bins. The complete systematic covariance matrix is determined by combining the statistical and all systematic covariance matrices

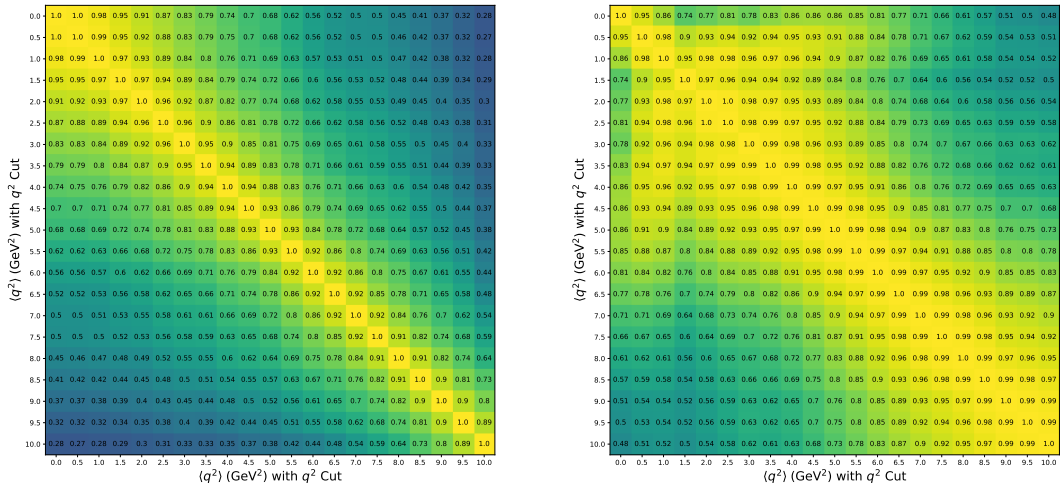

FIGURE 6.7: Statistical correlations (left) and experimental correlations (right) between measured $\langle q^2 \rangle$ moments for electron candidates.

stemming from the individual systematic error sources. An example of the experimental correlations for electron candidates is shown in Figure 6.7. The statistical and experimental correlation matrices for the measured $\langle q^n \rangle$ moments for both electron and muon candidates are given in Appendix F.

TABLE 6.1: Summary of statistical and systematic uncertainties for the moments $\langle q^{2,4,6,8} \rangle$ for electron candidates. The values are given as the relative error in permille.

q^2 selection in GeV^2	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$ in GeV^2	6.21	6.51	6.81	7.10	7.41	7.71	8.01	8.30	8.60	8.88	9.19	9.48	9.78	10.07	10.38
Stat. error (data)	1.42	1.36	1.31	1.27	1.24	1.21	1.18	1.17	1.16	1.16	1.17	1.19	1.22	1.25	1.29
Bkg. subtraction	1.06	0.76	0.57	0.40	0.37	0.41	0.48	0.55	0.60	0.63	0.67	0.70	0.77	0.83	0.91
$B \to X_u \ell\nu$ BF	1.80	1.66	1.52	1.10	0.78	0.57	0.42	0.29	0.23	0.19	0.16	0.13	0.12	0.15	0.10
$B \to X_c \ell\nu$ BF	4.94	4.72	4.90	4.55	4.47	4.30	3.77	3.47	3.28	2.99	2.49	1.96	1.41	1.27	1.16
Non-resonant model	13.20	11.54	10.14	8.33	6.93	5.78	4.50	3.54	2.76	2.21	1.63	1.24	0.81	0.77	0.66
$B \to X_c \ell\nu$ FF	1.63	1.47	1.27	1.10	0.94	0.84	0.79	0.75	0.69	0.63	0.56	0.49	0.42	0.40	0.37
N_{tracks} res.	4.67	4.40	4.17	3.91	3.72	3.50	3.31	3.13	2.95	2.83	2.66	2.50	2.40	2.24	2.13
N_γ res.	0.50	0.49	0.44	0.43	0.40	0.37	0.36	0.36	0.36	0.35	0.33	0.31	0.31	0.32	0.32
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	0.71	0.69	0.73	0.68	0.77	0.86	0.89	0.94	0.97	1.06	1.14	1.20	1.30	1.29	1.31
q^2 scale	8.98	6.77	6.12	5.77	5.70	5.50	5.12	4.86	4.72	4.62	4.25	4.33	3.91	3.90	3.94
MC non-closure	0.06	0.09	0.03	0.01	0.01	0.03	0.02	0.01	0.01	0.01	0.00	0.03	0.00	0.01	0.00
Cal. function	0.13	0.08	0.03	0.02	0.08	0.12	0.17	0.22	0.26	0.30	0.34	0.38	0.42	0.45	0.49
Stat. bias corr.	1.22	1.18	1.14	1.11	1.08	1.06	1.04	1.03	1.02	1.02	1.02	1.03	1.05	1.07	1.10
PID eff.	0.17	0.16	0.14	0.13	0.13	0.11	0.10	0.09	0.09	0.08	0.08	0.08	0.07	0.06	0.05
Track eff.	0.42	0.39	0.36	0.33	0.31	0.29	0.27	0.25	0.23	0.22	0.20	0.19	0.18	0.16	0.15
$B^0/B^\pm$ tag eff.	0.21	0.25	0.35	0.41	0.52	0.44	0.45	0.37	0.55	0.74	0.72	0.73	0.60	0.47	0.45
Sys. error (total)	17.62	15.11	13.72	11.98	10.88	9.89	8.66	7.83	7.27	6.85	6.16	5.87	5.29	5.17	5.13
Total rel. error in ‰	17.68	15.17	13.78	12.04	10.95	9.96	8.74	7.92	7.36	6.95	6.27	5.99	5.43	5.32	5.29
$\langle q^4 \rangle$ in GeV^4	42.99	46.27	49.74	53.41	57.42	61.64	65.96	70.46	75.15	79.92	85.20	90.50	96.07	101.73	108.10
Stat. error (data)	3.23	3.13	3.04	2.96	2.88	2.82	2.77	2.74	2.72	2.72	2.74	2.78	2.84	2.92	3.01
Bkg. subtraction	1.78	1.35	1.14	0.99	1.06	1.17	1.33	1.49	1.58	1.65	1.71	1.78	1.93	2.05	2.23
$B \to X_u \ell\nu$ BF	3.84	3.52	3.16	2.29	1.64	1.21	0.90	0.63	0.52	0.43	0.37	0.31	0.29	0.34	0.24
$B \to X_c \ell\nu$ BF	9.45	9.45	9.93	9.48	9.41	9.10	8.13	7.54	7.11	6.45	5.40	4.31	3.18	2.88	2.62
Non-resonant model	25.19	22.36	19.84	16.57	13.95	11.72	9.25	7.36	5.81	4.68	3.49	2.69	1.82	1.73	1.49
$B \to X_c \ell\nu$ FF	3.13	2.84	2.51	2.22	1.95	1.81	1.72	1.64	1.52	1.41	1.26	1.12	0.98	0.92	0.85
N_{tracks} res.	10.44	9.95	9.49	8.98	8.56	8.09	7.67	7.26	6.84	6.52	6.13	5.75	5.48	5.11	4.84
N_γ res.	1.18	1.16	1.07	1.05	1.00	0.93	0.91	0.90	0.90	0.86	0.82	0.79	0.78	0.81	0.80
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	2.17	2.11	2.15	2.05	2.20	2.32	2.36	2.43	2.47	2.62	2.75	2.83	3.00	2.95	2.93
q^2 scale	18.61	15.14	14.05	13.28	13.10	12.57	11.79	11.15	10.79	10.59	9.81	9.95	9.04	9.03	9.15
MC non-closure	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00	0.00
Cal. function	0.27	0.15	0.03	0.08	0.20	0.31	0.42	0.52	0.61	0.70	0.79	0.88	0.95	1.03	1.10
Stat. bias corr.	2.68	2.63	2.57	2.52	2.47	2.43	2.40	2.37	2.35	2.35	2.35	2.37	2.40	2.45	2.50
PID eff.	0.36	0.33	0.31	0.29	0.28	0.24	0.22	0.20	0.20	0.19	0.18	0.17	0.14	0.12	0.10
Track eff.	0.90	0.85	0.80	0.75	0.71	0.66	0.62	0.58	0.53	0.50	0.46	0.43	0.40	0.36	0.34
$B^0/B^\pm$ tag eff.	0.81	0.89	1.03	1.12	1.30	1.14	1.15	1.00	1.32	1.65	1.60	1.58	1.30	1.03	0.97
Sys. error (total)	34.95	30.89	28.49	25.42	23.45	21.54	19.21	17.53	16.35	15.48	14.04	13.42	12.18	11.90	11.82
Total rel. error in ‰	35.10	31.04	28.65	25.59	23.63	21.72	19.41	17.74	16.58	15.72	14.30	13.71	12.50	12.25	12.20
$\langle q^6 \rangle$ in GeV^6	326.23	355.51	387.59	423.92	465.17	510.57	558.65	610.98	667.93	728.29	797.82	870.48	949.14	1031.95	1128.69
Stat. error (data)	5.84	5.70	5.56	5.40	5.26	5.13	5.03	4.96	4.91	4.89	4.89	4.94	5.03	5.14	5.27
Bkg. subtraction	2.55	2.18	2.20	2.14	2.38	2.54	2.81	3.05	3.14	3.22	3.30	3.37	3.62	3.81	4.11
$B \to X_u \ell\nu$ BF	6.43	5.82	5.07	3.66	2.62	1.97	1.45	1.04	0.88	0.74	0.65	0.55	0.51	0.59	0.42
$B \to X_c \ell\nu$ BF	13.83	14.41	15.25	14.81	14.75	14.32	12.97	12.09	11.37	10.31	8.70	7.01	5.31	4.81	4.38
Non-resonant model	35.93	32.23	28.84	24.43	20.79	17.60	14.08	11.34	9.04	7.33	5.55	4.33	3.03	2.85	2.48
$B \to X_c \ell\nu$ FF	4.42	4.10	3.73	3.41	3.12	2.96	2.84	2.72	2.54	2.36	2.12	1.90	1.68	1.58	1.44
N_{tracks} res.	17.45	16.76	16.09	15.31	14.63	13.87	13.17	12.46	11.75	11.16	10.46	9.79	9.28	8.63	8.12
N_γ res.	2.15	2.11	1.99	1.95	1.87	1.76	1.72	1.69	1.68	1.61	1.54	1.48	1.47	1.49	1.47
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	4.56	4.44	4.43	4.27	4.41	4.53	4.53	4.58	4.59	4.75	4.88	4.93	5.10	4.96	4.86
q^2 scale	29.62	25.28	23.94	22.61	22.35	21.35	20.15	19.01	18.35	18.07	16.85	17.06	15.55	15.56	15.83
MC non-closure	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cal. function	0.42	0.22	0.02	0.18	0.38	0.57	0.75	0.92	1.09	1.24	1.39	1.53	1.65	1.77	1.87
Stat. bias corr.	4.68	4.60	4.52	4.44	4.36	4.29	4.23	4.18	4.14	4.11	4.10	4.12	4.15	4.20	4.27
PID eff.	0.57	0.53	0.50	0.47	0.45	0.40	0.37	0.33	0.33	0.30	0.29	0.27	0.23	0.19	0.16
Track eff.	1.47	1.40	1.33	1.25	1.18	1.10	1.04	0.97	0.90	0.84	0.78	0.72	0.67	0.61	0.57
$B^0/B^\pm$ tag eff.	1.79	1.87	2.03	2.13	2.34	2.11	2.09	1.86	2.29	2.71	2.58	2.50	2.06	1.64	1.52
Sys. error (total)	52.77	47.68	44.62	40.49	37.88	35.07	31.80	29.25	27.40	26.06	23.83	22.86	20.84	20.39	20.28
Total rel. error in ‰	53.09	48.01	44.96	40.85	38.24	35.44	32.20	29.67	27.84	26.52	24.33	23.39	21.44	21.03	20.95
$\langle q^8 \rangle$ in GeV^8	2663.52	2914.20	3192.80	3531.01	3925.20	4375.31	4862.26	5414.29	6036.81	6723.88	7545.39	8439.78	9431.26	10512.44	11818.99
Stat. error (data)	9.77	9.58	9.38	9.11	8.84	8.58	8.40	8.23	8.09	8.00	7.93	7.93	8.01	8.13	8.25
Bkg. subtraction	4.02	3.88	4.31	4.24	4.67	4.85	5.20	5.52	5.51	5.58	5.65	5.71	6.05	6.31	6.77
$B \to X_u \ell\nu$ BF	9.95	8.85	7.47	5.32	3.79	2.87	2.11	1.52	1.32	1.13	1.01	0.87	0.80	0.91	0.64
$B \to X_c \ell\nu$ BF	19.15	20.11	21.08	20.54	20.40	19.78	18.08	16.93	15.92	14.46	12.29	10.03	7.78	7.06	6.42
Non-resonant model	45.70	41.23	37.11	31.73	27.22	23.21	18.80	15.31	12.33	10.08	7.73	6.11	4.40	4.13	3.61
$B \to X_c \ell\nu$ FF	5.57	5.32	5.03	4.75	4.49	4.33	4.19	4.01	3.76	3.50	3.16	2.85	2.54	2.37	2.16
N_{tracks} res.	25.81	24.89	24.00	22.90	21.91	20.80	19.77	18.71	17.64	16.71	15.64	14.60	13.78	12.78	11.97
N_γ res.	3.51	3.45	3.30	3.22	3.10	2.94	2.87	2.81	2.77	2.66	2.55	2.44	2.41	2.42	2.37
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	7.95	7.73	7.63	7.37	7.46	7.52	7.44	7.42	7.34	7.44	7.51	7.47	7.59	7.30	7.07
q^2 scale	42.51	37.31	35.84	33.74	33.46	31.82	30.22	28.43	27.42	27.12	25.46	25.75	23.55	23.65	24.18
MC non-closure	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cal. function	0.62	0.31	0.01	0.31	0.62	0.93	1.20	1.47	1.72	1.95	2.18	2.39	2.57	2.74	2.89
Stat. bias corr.	7.59	7.45	7.31	7.16	7.02	6.89	6.76	6.65	6.56	6.48	6.43	6.40	6.40	6.44	6.51
PID eff.	0.79	0.74	0.70	0.66	0.63	0.56	0.53	0.48	0.47	0.43	0.41	0.37	0.32	0.27	0.22
Track eff.	2.11	2.02	1.93	1.83	1.73	1.62	1.53	1.43	1.33	1.24	1.15	1.06	0.98	0.89	0.83
$B^0/B^\pm$ tag eff.	3.03	3.10	3.26	3.34	3.56	3.24	3.18	2.86	3.35	3.81	3.59	3.43	2.81	2.24	2.05
Sys. error (total)	72.27	66.18	62.60	57.37	54.26	50.52	46.44	42.99	40.44	38.66	35.64	34.31	31.44	30.82	30.73
Total rel. error in ‰	72.93	66.86	63.30	58.09	54.97	51.25	47.19	43.77	41.24	39.48	36.51	35.21	32.44	31.87	31.81

TABLE 6.2: Summary of statistical and systematic uncertainties for the moments $\langle q^{2,4,6,8} \rangle$ for muon candidates. The values are given as the relative error in permille.

q^2 selection in GeV^2	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$ in GeV^2	6.25	6.54	6.83	7.13	7.42	7.72	8.02	8.32	8.61	8.91	9.21	9.51	9.80	10.09	10.40
Stat. error (data)	1.51	1.45	1.39	1.34	1.30	1.27	1.24	1.22	1.21	1.20	1.20	1.22	1.23	1.26	1.29
Bkg. subtraction	1.34	1.12	0.90	0.71	0.59	0.53	0.49	0.49	0.57	0.63	0.65	0.70	0.76	0.77	0.82
$B \to X_u \ell \nu$ BF	2.18	2.04	1.75	1.48	1.35	1.04	0.76	0.54	0.38	0.29	0.19	0.16	0.12	0.05	0.05
$B \to X_c \ell \nu$ BF	4.82	5.02	5.14	5.14	5.05	5.00	4.67	4.05	3.51	3.11	2.66	2.21	1.75	1.36	1.16
Non-resonant model	14.25	12.72	11.04	9.28	7.83	6.62	5.42	4.00	3.02	2.28	1.65	1.43	1.04	0.86	0.78
$B \to X_c \ell \nu$ FF	1.43	1.30	1.16	1.03	0.91	0.85	0.82	0.74	0.69	0.62	0.54	0.48	0.42	0.39	0.35
N_{tracks} res.	5.66	5.31	4.96	4.65	4.36	4.06	3.78	3.52	3.29	3.06	2.85	2.66	2.51	2.38	2.20
N_γ res.	0.39	0.38	0.34	0.31	0.30	0.28	0.30	0.31	0.32	0.28	0.27	0.26	0.25	0.27	0.29
$E_{\text{miss}} - \lvert \mathbf{p}_{\text{miss}} \rvert$ shape	1.29	1.26	1.21	1.17	1.15	1.11	1.04	1.05	1.06	1.09	1.16	1.20	1.30	1.33	1.29
q^2 scale	9.48	7.15	6.65	6.65	6.12	5.91	5.83	5.48	5.26	4.69	4.27	4.42	3.91	3.94	4.38
MC non-closure	0.19	0.11	0.12	0.11	0.11	0.05	0.05	0.06	0.08	0.07	0.11	0.04	0.04	0.06	0.02
Cal. function	0.13	0.08	0.03	0.02	0.07	0.12	0.17	0.22	0.26	0.31	0.35	0.39	0.43	0.47	0.51
Stat. bias corr.	1.32	1.27	1.23	1.19	1.16	1.13	1.10	1.08	1.07	1.06	1.06	1.06	1.07	1.09	1.11
PID eff.	0.16	0.14	0.14	0.13	0.13	0.12	0.11	0.10	0.10	0.10	0.09	0.08	0.08	0.07	0.06
Track eff.	0.44	0.42	0.39	0.36	0.34	0.31	0.29	0.27	0.25	0.23	0.21	0.20	0.18	0.17	0.15
$B^0/B^\pm$ tag eff.	0.46	0.58	0.50	0.44	0.51	0.40	0.28	0.34	0.36	0.38	0.29	0.23	0.20	0.12	0.47
Sys. error (total)	18.99	16.65	15.03	13.62	12.22	11.19	10.17	8.86	7.97	7.06	6.30	6.09	5.44	5.27	5.50
Total rel. error in ‰	19.05	16.71	15.09	13.68	12.29	11.26	10.25	8.94	8.06	7.16	6.41	6.21	5.58	5.42	5.65
$\langle q^4 \rangle$ in GeV^4	43.52	46.73	50.16	53.81	57.65	61.75	66.09	70.71	75.36	80.34	85.63	90.96	96.39	102.18	108.50
Stat. error (data)	3.45	3.33	3.22	3.12	3.03	2.95	2.88	2.83	2.80	2.79	2.79	2.81	2.86	2.91	2.97
Bkg. subtraction	2.41	2.10	1.76	1.48	1.36	1.34	1.30	1.34	1.53	1.65	1.67	1.80	1.90	1.90	2.00
$B \to X_u \ell \nu$ BF	4.66	4.39	3.80	3.20	2.90	2.23	1.61	1.14	0.81	0.61	0.39	0.34	0.26	0.11	0.10
$B \to X_c \ell \nu$ BF	9.55	10.30	10.68	10.78	10.67	10.50	9.84	8.63	7.52	6.67	5.69	4.73	3.78	2.97	2.52
Non-resonant model	27.34	24.67	21.67	18.48	15.72	13.36	11.00	8.24	6.30	4.82	3.53	3.05	2.27	1.88	1.69
$B \to X_c \ell \nu$ FF	2.73	2.51	2.29	2.08	1.90	1.80	1.75	1.61	1.49	1.36	1.20	1.07	0.95	0.87	0.79
N_{tracks} res.	12.45	11.81	11.13	10.51	9.89	9.25	8.63	8.04	7.51	6.98	6.49	6.04	5.66	5.34	4.90
N_γ res.	0.95	0.92	0.85	0.79	0.78	0.74	0.77	0.78	0.78	0.71	0.69	0.66	0.64	0.67	0.71
$E_{\text{miss}} - \lvert \mathbf{p}_{\text{miss}} \rvert$ shape	3.35	3.26	3.12	3.01	2.92	2.81	2.66	2.65	2.64	2.67	2.76	2.81	2.95	2.98	2.85
q^2 scale	19.97	16.30	15.38	15.26	14.19	13.68	13.33	12.59	11.97	10.81	9.86	10.03	8.97	9.17	10.05
MC non-closure	0.06	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.03	0.01	0.01	0.01	0.01
Cal. function	0.27	0.16	0.04	0.07	0.19	0.30	0.41	0.52	0.62	0.71	0.81	0.90	0.98	1.06	1.14
Stat. bias corr.	2.92	2.85	2.77	2.70	2.64	2.58	2.53	2.48	2.45	2.43	2.42	2.42	2.43	2.46	2.49
PID eff.	0.35	0.32	0.31	0.29	0.29	0.26	0.25	0.23	0.22	0.22	0.20	0.18	0.18	0.16	0.13
Track eff.	0.97	0.91	0.86	0.81	0.76	0.70	0.65	0.61	0.56	0.52	0.48	0.44	0.41	0.38	0.34
$B^0/B^\pm$ tag eff.	0.92	1.09	0.95	0.84	0.94	0.74	0.51	0.59	0.62	0.63	0.43	0.29	0.22	0.42	1.12
Sys. error (total)	38.08	34.23	31.41	28.91	26.30	24.24	22.18	19.65	17.77	15.90	14.27	13.73	12.36	12.08	12.51
Total rel. error in ‰	38.23	34.39	31.58	29.08	26.48	24.42	22.37	19.85	17.99	16.14	14.54	14.02	12.69	12.43	12.86
$\langle q^6 \rangle$ in GeV^6	331.70	360.44	392.58	428.49	467.70	511.77	560.28	614.62	670.90	733.57	803.64	876.13	952.91	1038.72	1135.12
Stat. error (data)	6.22	6.03	5.84	5.66	5.49	5.34	5.20	5.08	5.01	4.96	4.94	4.96	5.02	5.07	5.16
Bkg. subtraction	3.62	3.28	2.94	2.67	2.67	2.77	2.71	2.77	3.08	3.26	3.24	3.48	3.57	3.52	3.68
$B \to X_u \ell \nu$ BF	7.78	7.30	6.34	5.26	4.74	3.62	2.57	1.81	1.27	0.95	0.59	0.52	0.41	0.15	0.14
$B \to X_c \ell \nu$ BF	14.82	16.14	16.68	16.81	16.67	16.30	15.29	13.54	11.86	10.53	9.00	7.50	6.04	4.78	4.06
Non-resonant model	39.15	35.58	31.54	27.20	23.35	19.95	16.53	12.55	9.72	7.52	5.59	4.82	3.64	3.02	2.71
$B \to X_c \ell \nu$ FF	3.86	3.65	3.42	3.19	3.00	2.88	2.82	2.61	2.44	2.23	1.98	1.78	1.58	1.44	1.30
N_{tracks} res.	20.45	19.54	18.55	17.59	16.63	15.59	14.58	13.60	12.70	11.79	10.93	10.14	9.46	8.85	8.10
N_γ res.	1.77	1.73	1.63	1.54	1.52	1.45	1.46	1.47	1.46	1.34	1.30	1.25	1.21	1.24	1.29
$E_{\text{miss}} - \lvert \mathbf{p}_{\text{miss}} \rvert$ shape	6.30	6.10	5.85	5.63	5.43	5.22	4.94	4.87	4.79	4.76	4.83	4.82	4.95	4.92	4.65
q^2 shift	32.17	27.63	26.32	26.02	24.37	23.45	22.66	21.49	20.24	18.50	16.91	16.97	15.32	15.86	17.17
MC non-closure	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
Cal. function	0.42	0.24	0.04	0.16	0.35	0.54	0.73	0.91	1.08	1.24	1.40	1.55	1.68	1.80	1.92
Stat. bias corr.	5.10	4.98	4.86	4.75	4.64	4.53	4.43	4.34	4.27	4.21	4.17	4.15	4.14	4.16	4.19
PID eff.	0.57	0.53	0.51	0.48	0.48	0.44	0.41	0.39	0.36	0.36	0.33	0.29	0.28	0.25	0.21
Track eff.	1.57	1.50	1.42	1.34	1.26	1.18	1.09	1.02	0.94	0.87	0.80	0.73	0.67	0.62	0.56
$B^0/B^\pm$ tag eff.	1.16	1.36	1.18	1.03	1.15	0.85	0.53	0.63	0.65	0.63	0.32	0.10	0.02	0.97	1.98
Sys. error (total)	58.03	53.09	49.34	45.97	42.29	39.22	36.10	32.45	29.49	26.65	24.04	23.03	20.88	20.59	21.17
Total rel. error in ‰	58.36	53.43	49.68	46.32	42.65	39.58	36.47	32.84	29.91	27.11	24.54	23.56	21.47	21.21	21.79
$\langle q^8 \rangle$ in GeV^8	2717.22	2963.88	3248.31	3578.45	3947.44	4384.73	4878.23	5458.95	6072.92	6780.95	7616.67	8497.60	9466.03	10603.31	11917.23
Stat. error (data)	10.35	10.07	9.78	9.47	9.19	8.89	8.63	8.36	8.19	8.05	7.94	7.91	7.94	7.95	7.99
Bkg. subtraction	5.57	5.25	4.98	4.80	4.99	5.23	5.02	5.06	5.51	5.71	5.58	6.00	5.96	5.81	6.02
$B \to X_u \ell \nu$ BF	11.94	11.10	9.61	7.82	7.00	5.31	3.66	2.53	1.76	1.30	0.79	0.69	0.59	0.20	0.16
$B \to X_c \ell \nu$ BF	21.51	22.91	23.24	23.14	22.84	22.14	20.76	18.50	16.31	14.53	12.43	10.40	8.44	6.74	5.74
Non-resonant model	49.93	45.52	40.56	35.22	30.45	26.13	21.80	16.75	13.12	10.26	7.73	6.66	5.10	4.25	3.79
$B \to X_c \ell \nu$ FF	4.91	4.76	4.60	4.40	4.23	4.12	4.03	3.75	3.52	3.23	2.88	2.59	2.31	2.09	1.89
N_{tracks} res.	29.72	28.51	27.15	25.82	24.47	22.99	21.54	20.09	18.76	17.40	16.09	14.89	13.83	12.86	11.73
N_γ res.	2.95	2.89	2.75	2.62	2.58	2.46	2.46	2.44	2.39	2.22	2.16	2.07	2.00	2.01	2.06
$E_{\text{miss}} - \lvert \mathbf{p}_{\text{miss}} \rvert$ shape	10.18	9.83	9.42	9.05	8.69	8.33	7.89	7.70	7.50	7.35	7.33	7.21	7.26	7.11	6.66
q^2 scale	46.61	41.26	39.53	39.00	36.70	35.23	33.82	32.22	30.11	27.83	25.47	25.28	23.04	24.16	25.90
MC non-closure	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cal. function	0.63	0.34	0.04	0.26	0.55	0.86	1.14	1.42	1.68	1.92	2.17	2.37	2.56	2.74	2.90
Stat. bias corr.	8.21	8.00	7.79	7.59	7.38	7.18	6.99	6.83	6.67	6.54	6.43	6.35	6.29	6.27	6.27
PID eff.	0.81	0.77	0.74	0.70	0.69	0.64	0.60	0.56	0.52	0.51	0.47	0.42	0.40	0.35	0.30
Track eff.	2.25	2.16	2.05	1.95	1.84	1.72	1.60	1.49	1.38	1.27	1.16	1.06	0.98	0.89	0.81
$B^0/B^\pm$ tag eff.	0.97	1.19	0.98	0.80	0.94	0.58	0.18	0.29	0.30	0.26	0.14	0.42	0.59	1.79	3.06
Sys. error (total)	79.98	73.90	69.18	64.95	60.23	56.09	51.86	47.19	43.06	39.32	35.61	34.06	31.06	30.93	31.65
Total rel. error in ‰	80.64	74.58	69.87	65.64	60.93	56.79	52.58	47.93	43.83	40.13	36.48	34.97	32.06	31.93	32.64

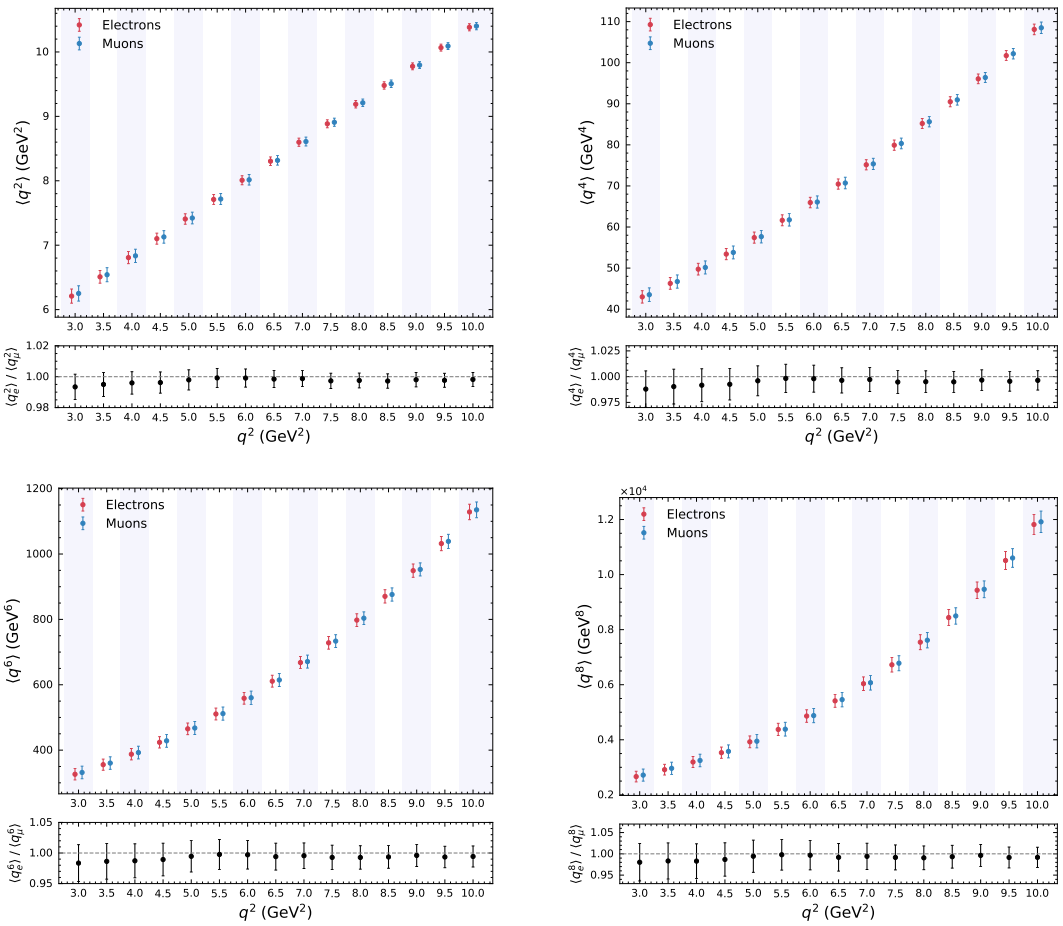

FIGURE 6.8: Comparison of the measured $\langle q^{2,4,6,8} \rangle$ moments for electron and muon candidates with full statistical and systematic errors. The expectation of lepton flavour universality is tested: in the shown ratio of electron to muon moments many of the associated systematic uncertainties cancel and all reported moments are compatible with the expectation of unity. Note that the individual electron and muon moments are strongly correlated.

6.3 Results

The extracted $\langle q^{2,4,6,8} \rangle$ moments measured as a progression of increasing threshold selections on the q^2 distribution for both electron and muon candidates are summarised in Tables 6.1 and 6.2, respectively. Furthermore, the expectation of lepton flavour universality is tested for moments of the same order: Figure 6.8 compares the first to fourth order measured moments for electron and muon candidates. In the shown ratio many of the associated systematic uncertainties cancel and all reported moments are compatible with the expectation of unity.

Dominant systematic uncertainties stem from the uncertainty associated with the modelling of the $B \to X_c \ell \nu$ composition, especially the non-resonant components. The estimated uncertainties associated with these sources for the first moment, with the lowest selection of $q^2 > 3.0\,\text{GeV}^2$ for electron candidates, are found to be 0.49% and 1.32%, respectively. This is followed in size by the uncertainty associated with

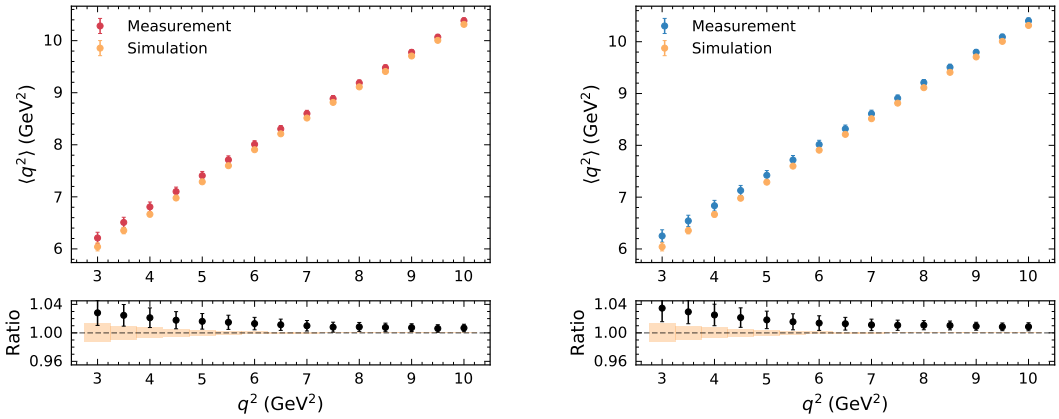

FIGURE 6.9: Comparison of the simulated and measured moments for electron (left) and muon (right) candidates.

the modelling of the number of charged particles in the X system as well as the modelling of the q^2 distribution. These sources of uncertainty remain leading systematics across all q^2 selections for both muon and electron candidates. On the other hand, the uncertainty due to the modelling of the $B \to X_u \ell\nu$ background component is found to be a leading systematic mainly for low q^2 selections. While the statistical uncertainties due to the additional correction factors, the determination of the linear calibration functions and the reconstruction efficiency of neutral and charged B mesons are small contributions to the overall systematic error for low q^2 selections, these sources increase in significance as the selection criteria progress to higher values of q^2 and the number of signal events gradually decrease. Conversely, the uncertainties due to the lepton and hadron identification efficiency, the overall track finding efficiency and the modelling of the number of neutral clusters in the X system are already small for the lower q^2 selections and gradually decrease further as stricter q^2 selections are imposed. The smallest source of systematic error across all q^2 selections is the estimated residual bias due to the extraction method.

The measured moments are compared to simulated moments calculated on generator-level signal MC. Figure 6.9 compares the first measured moments to the simulated moments for electron and muon candidates, while the comparisons for the measured moments of higher orders are given in Appendix G. For the generator-level moments, uncertainties on the $B \to X_c \ell\nu$ composition as well as form factor variations are considered. The first moment with the lowest threshold selection of $q^2 > 3.0\,\text{GeV}^2$ is observed to be approximately one standard deviation higher in data than in the simulated signal samples. The agreement steadily improves with each successive q^2 threshold selection, until the best possible agreement is observed for high values of q^2, where $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays dominate. This trend is found to be consistent between electron and muon candidates and also for moments of higher order. This suggests that the chosen modelling of the $B \to X_c \ell\nu$ composition tends to overestimate the number of signal events in low regions of the q^2 spectrum. A comparable trend has been reported by [137] in which the inclusive lepton spectrum is analysed using a $B \to X_c \ell\nu$ model similar to the one used in this work. In addition, similar observations have also been reported by [138] where the moments

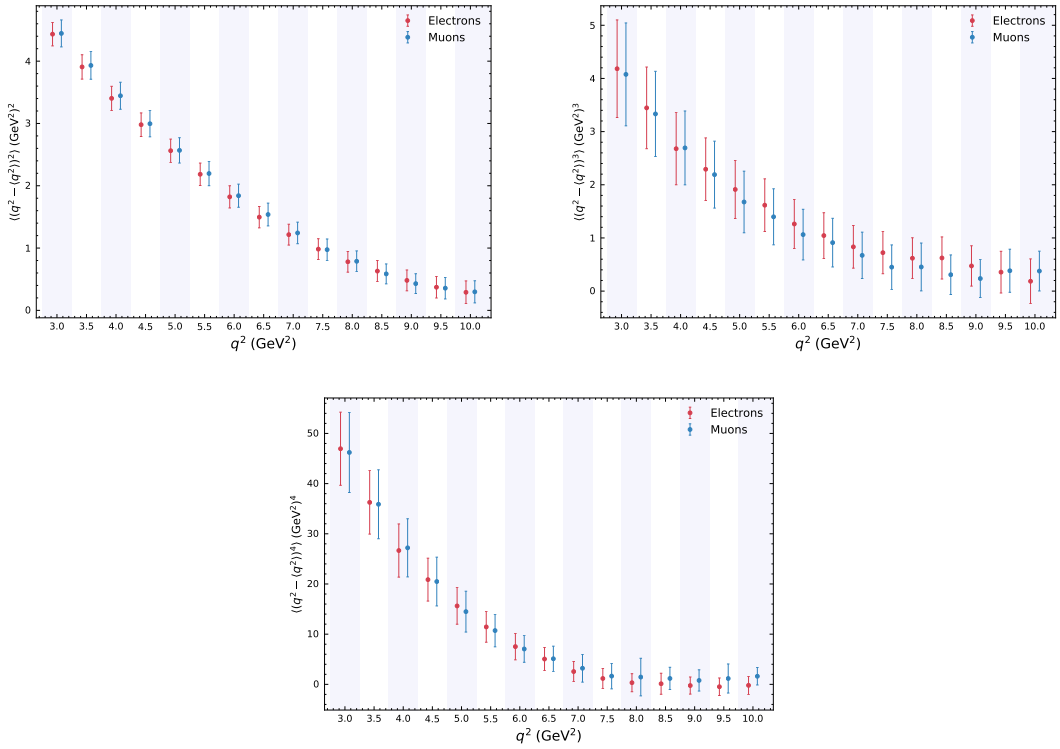

FIGURE 6.10: The extracted central moments for both electron and muon candidates calculated from the measured nominal moments with Eqn. 6.11.

of the lepton energy, hadronic mass, and hadronic energy spectra were investigated to determine the exclusive composition of $B \to X_c \ell \nu$. Here, it was observed that the data prefer to increase the yield of the $B \to X_c \ell \nu$ components that produce a more energetic lepton momentum spectrum, resulting in higher values of the moments at low regions of q^2. The direct measurement of the q^2 moments presented in this work supports this explanation.

In addition to measuring the nominal q^2 moments, the central or normalized moments are also extracted. The central moments have the advantage of becoming slightly less correlated with respect to the correlations of the nominal moments, especially for the higher moments. To calculate the central moments directly from the measured nominal moments, the following non-linear transformations are implemented:

$$\begin{pmatrix} \langle q^2 \rangle \\ \langle q^4 \rangle \\ \langle q^6 \rangle \\ \langle q^8 \rangle \end{pmatrix} \quad \to \quad \begin{pmatrix} \langle q^2 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^2 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^3 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^4 \rangle \end{pmatrix} . \tag{6.11}$$

The new systematic covariance matrix $\mathcal{C}'$ for the vector of central moments is calculated by making use of the Jacobian matrix $\mathcal{J}$ for the transformation, together with the initial experimental covariance matrix $\mathcal{C}$ describing the nominal moments, such that

$$\mathcal{C}' = \mathcal{J}^T \mathcal{C} \mathcal{J} . \tag{6.12}$$

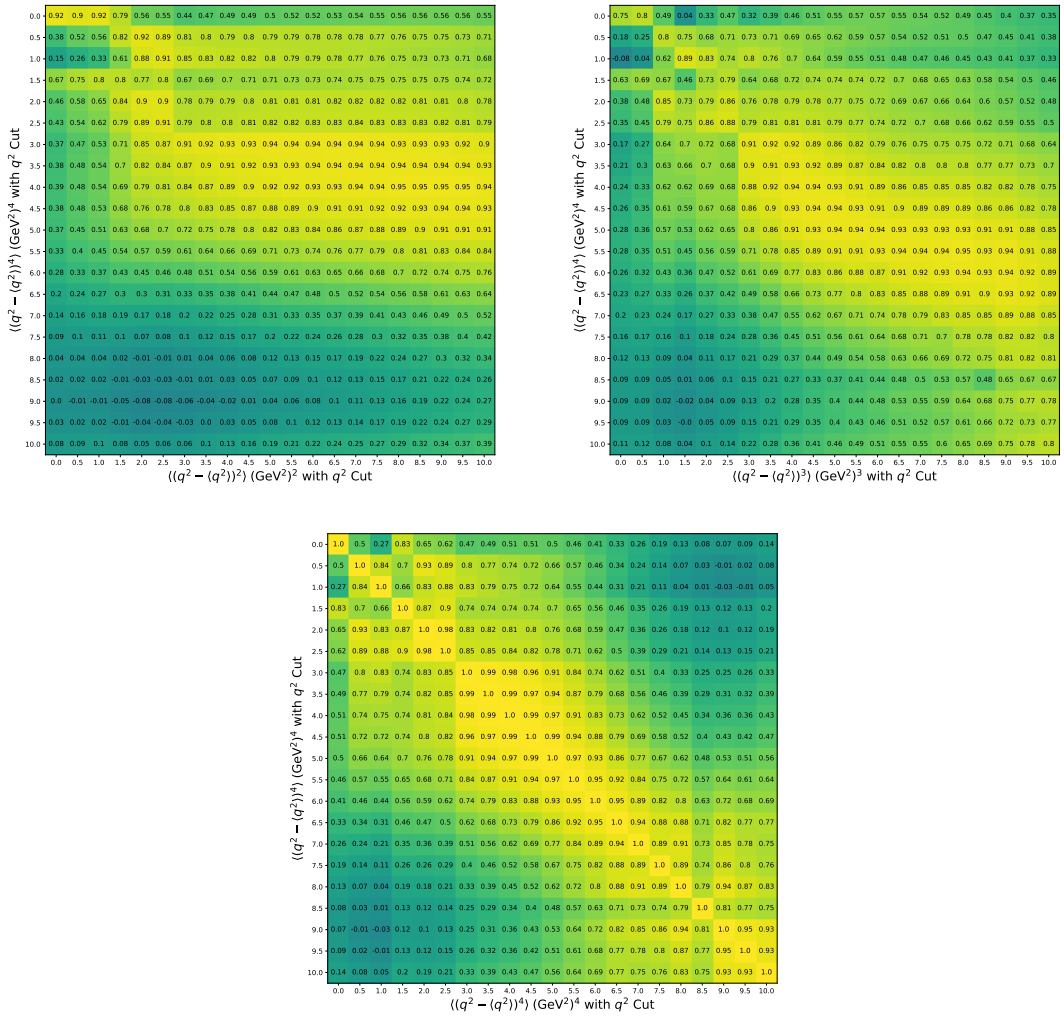

FIGURE 6.11: Correlations between the measured central moments for electron candidates.

This approximation of the uncertainties of the central moments yields the same results as given by performing Gaussian error propagation. Figure 6.10 shows the extracted central moments for both electron and muon candidates after applying the non-linear transformations with Eqn. 6.11. The central moments become slightly less correlated with respect to the correlations of the nominal moments, especially for the higher moments, while negative correlations between different orders of moments are also observed. Figure 6.11 shows an example of the correlations between different orders of the extracted central moments, while the complete experimental correlation matrices for the determined central moments are given in Appendix F.

6.4 Interpretation of the measurement

In this section the CKM matrix element $|V_{cb}|$ is extracted based on the reduced set of HQE parameters discussed in Section 2.3, while using the total $B \to X_c \ell \nu$ rate together with preliminary measured q^2 moments presented in [139]. These measurements are calculated by making use of a similar method as described in the

TABLE 6.3: Summary of the HQE parameters restricted in all fits by Gaussian error constraints.

Parameter	Value	σ
$m_c^{\overline{\text{MS}}}$	1.093 GeV	0.010 GeV
μ_π^2	1.423 GeV^2	0.24 GeV^2
μ_G^2	0.362 GeV^2	0.21 GeV^2

previous sections, however the minimum threshold selection on q^2 was chosen to be $q^2 > 3.6\,\text{GeV}^2$. Furthermore, the results presented in this section were obtained through close collaboration with M. Fael, K. Olschewsky and K. Keri Vos, who provided theoretical calculations in the form of a preliminary software package described in [140].

An identical strategy to the approach in [79] is proposed to extract the HQE parameters as well as $|V_{cb}|$ from a global fit to the lepton energy, hadronic mass or hadronic energy spectra of $B \to X_c \ell \nu$ decays. The moments of the q^2 distribution are defined as

$$\langle (q^2)^n \rangle_{q^2_{\text{cut}}} \equiv \frac{\int_{q^2_{\text{cut}}}^{q^2_{\text{max}}} (q^2)^n (d\Gamma/dq^2) dq^2}{\int_{q^2_{\text{cut}}}^{q^2_{\text{max}}} (d\Gamma/dq^2) dq^2}, \tag{6.13}$$

where q^2_{cut} denotes the lower limit of integration. By making use of the calculable dependence of the q^2 moments on the threshold selections as well as the theoretical prediction of the total semileptonic rate, $|V_{cb}|$ can also be extracted as an additional free parameter in the global fit. To this end, the ratio R^* is defined as follows

$$R^* \equiv \frac{\Gamma_{q^2_{\text{cut}}}}{\Gamma_{\text{tot}}}. \tag{6.14}$$

This quantity relates the measurement of the rate with a threshold selection on q^2 to the total semileptonic rate without any threshold selection, which allows for the determination of $|V_{cb}|$.

The fit method designed to extract HQE parameters from the moments measurement has been documented in [141], which is based on a χ^2 function constructed as follows:

$$\chi^2 = (\vec{x}_{\text{pred}} - \vec{x}_{\text{meas}})^T C^{-1} (\vec{x}_{\text{pred}} - \vec{x}_{\text{meas}}) + \sum_i \frac{(p_i - p_i^c)^2}{(\sigma_i^c)^2}. \tag{6.15}$$

Here, the vectors $\vec{x}_{\text{meas}}$ and $\vec{x}_{\text{pred}}$ contain the measured moments included in the fit and the corresponding moments determined from theory predictions. Furthermore, the total covariance matrix C is defined as the sum of the experimental and theoretical covariance matrices: $C = C_{\text{stat}} + C_{\text{sys}} + C_{\text{theory}}$. The final term in Eqn. 6.15 includes p_i^c and denotes Gaussian constraints on nuisance parameters that are used to introduce external information to the fit. Table 6.3 summarises the constraints imposed on various fit parameters. Furthermore, for the fits to the measured moments of higher order, additional constraints are applied on m_b^{kin} and the following HQE parameters:

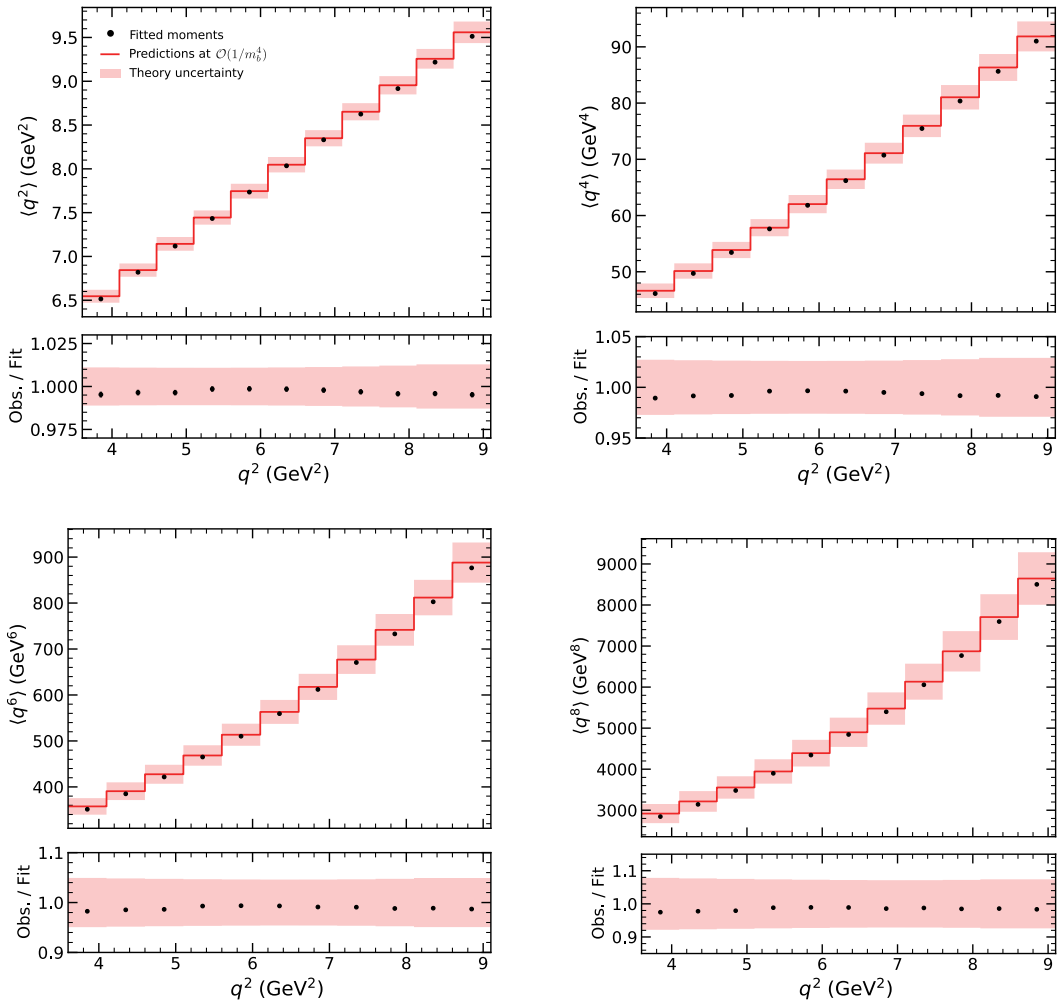

FIGURE 6.12: The fitted $\langle q^{2,4,6,8} \rangle$ moments for electron candidates and the theoretical predictions at $\mathcal{O}(1/m_b^4)$. While both experimental and theoretical uncertainties are shown in the ratio, the results are dominated by theoretical systematic uncertainties. The fit converges with a final χ^2 of 80.5 for 83 degrees of freedom.

$$s_B^4 = (-0.13 \pm 5)\,\text{GeV}^4\,, \qquad s_E^4 = (-0.072 \pm 5)\,\text{GeV}^4\,, \qquad s_{qB}^4 = (-0.80 \pm 5)\,\text{GeV}^4\,. \tag{6.16}$$

The values for all restricted parameters are determined in [6] by converting the values quoted in [79] to the RPI basis.

The theoretical covariance matrix is constructed by assuming a decorrelation of $0.96^{|i-j|}$ across increasing threshold selections i and j for theoretical uncertainties. Furthermore, a correlation of $0.2^{|k-l|}$ is assumed for moments of different orders k and l. The fit result is also compared to a theory covariance that assumes stronger correlations of theoretical uncertainties between moments of different orders: $0.99^{|k-l|}$ and a decorrelation between selections of $0.97^{|i-j|}$. Subsequently, an additional uncertainty is estimated for the assumed theoretical covariance as the difference between the two extracted central values of $|V_{cb}|$ obtained from the different scenarios. In addition, a shift of 10 MeV on m_b^{kin} is used to determine an additional uncertainty on

TABLE 6.4: The fitted HQE parameters as well as $|V_{cb}|$, where the quotes errors include both experimental and theoretical uncertainties. Correlation coefficients for all parameters are summarised below the results.

	$\lvert V_{cb}\rvert$	m_b^{kin} GeV	$m_c^{\overline{\text{MS}}}$ GeV	μ_G^2 GeV^2	s_B^4 GeV^4	r_E^4 GeV^4	s_{qB}^4 GeV^4	s_E^4 GeV^4	r_G^4 GeV^4	$\tilde{\rho}_D^3$ GeV^3	μ_π^2 GeV^2
Value	0.0417	4.5277	1.0936	0.4272	3.5764	0.0314	-1.2158	-1.9477	3.2288	-0.6649	0.4315
σ	0.0009	0.0453	0.0099	0.1268	0.7459	0.0168	1.5897	0.4255	0.1616	0.0424	0.2387
	1.00	-0.77	-0.04	-0.73	-0.53	-0.32	0.09	0.16	-0.48	-0.40	0.02
		1.00	0.18	0.93	0.67	0.45	-0.14	-0.20	0.62	0.56	0.14
			1.00	-0.01	-0.01	0.01	-0.01	-0.05	-0.01	-0.05	0.01
				1.00	0.64	0.38	-0.12	-0.22	0.58	0.46	0.13
					1.00	0.24	0.33	-0.04	0.55	0.26	0.10
						1.00	0.05	0.04	0.40	0.44	0.08
							1.00	-0.31	-0.13	0.07	0.01
								1.00	-0.38	-0.40	-0.05
									1.00	0.14	0.11
										1.00	0.08
											1.00

the total rate to account for missing higher order contributions.

The χ^2 function is numerically minimised and all quoted errors on the fitted parameters are obtained from $\Delta\chi^2 = 1$ contours, while the correlations are estimated directly from the Hessian matrix at the best fit point. Figure 6.12 shows the fit results for the first to the fourth q^2 moments over a threshold selection range of $q^2 \in [3.6, 8.6]$ GeV^2 in incremental steps of 0.5 GeV^2 for the electron candidates. Overall good agreement between the fitted moments and the theoretical predictions is observed with a final χ^2 of 80.5 for 83 degrees of freedom. Table 6.4 summarises the results of the fitted HQE parameters together with $|V_{cb}|$, which is found to be:

$$|V_{cb}| \times 10^3 = 41.7 \pm 1.2\,. \tag{6.17}$$

Here, the final quoted uncertainty is the linear sum of the experimental and theoretical errors. This determination of $|V_{cb}|$ agrees within experimental uncertainty with the current world average of inclusive $|V_{cb}|$ given by Eqn. 2.81, as well as previous determinations, shown in Figure 6.13. Due to conservative theoretical systematic uncertainties, the estimated total error on the final value is larger than previous measurements. Further work regarding the treatment of theoretical systematic uncertainties and correlations is required to refine this novel method, which would no doubt produce an updated future result that would provide complementary information to previous determinations of $|V_{cb}|$.

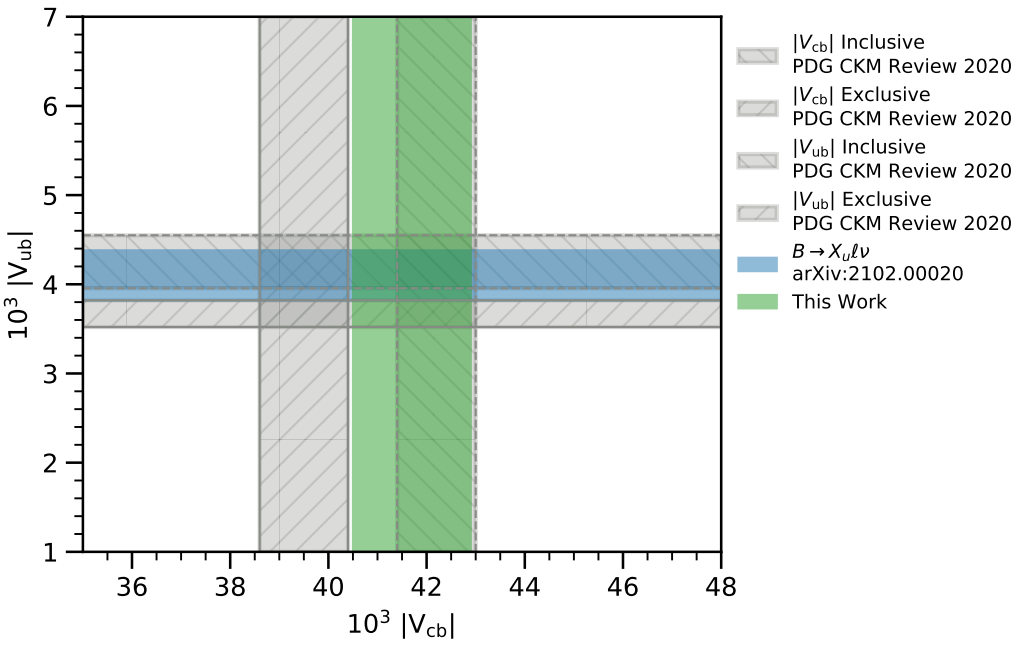

FIGURE 6.13: Comparison of this result with previous determinations of inclusive and exclusive $|V_{cb}|$ and $|V_{ub}|$.

Chapter 7

Conclusion and outlook

This thesis presents the first systematic study of the first to the fourth moments of the four-momentum transfer squared q^2 spectrum in semileptonic B meson decays to hadronic final states containing a charm quark: $B \to X_c \ell \nu$, where $\ell = e, \mu$. The study investigates the full Belle data set corresponding to 711 fb^{-1} of integrated luminosity at the $\Upsilon(4S)$ resonance and a total of $(772 \pm 10) \times 10^6$ B meson pairs. To identify and reconstruct the hadronic X_c system, the Full Reconstruction algorithm is employed to fully reconstruct one of the B mesons present in the event in purely hadronic decay modes with the aid of neural networks. Consequently, this approach allows for the kinematics of the remaining B meson to be inferred using conservation laws, leading to the explicit reconstruction of the hadronic X_c system and, in turn, the explicit reconstruction of the q^2 spectrum. Background contributions stemming from continuum, semileptonic cascade and $B \to X_u \ell \nu$ processes, as well as mis-identified leptons are subtracted by making use of the hadronic mass system M_X in an unbinned approach with event-wise signal probability weights. Since the reconstructed M_X and q^2 distributions are distorted by missing charged and neutral particles in the reconstruction of the hadronic tag-side and by finite detector acceptance and resolution effects, the measured moments are unfolded with correction methods based on generator-level MC samples to ensure an unbiased measurement. The full background subtraction and calibration procedure is verified on ensembles of statistically independent MC samples and no statistical significant biases in the unfolded moments are observed. Various sources of statistical and systematic uncertainties on the measurement that are introduced by the background subtraction and calibration procedures are investigated. Dominant contributions stem from the assumed $B \to X_c \ell \nu$ composition and from the modelling of the resolution of the X system. These contributions are followed in significance by the subtraction of the $B \to X_u \ell \nu$ process, which exhibits a very similar experimental signature to the signal decay of interest. The final q^2 moments are measured separately for electron and muon $B \to X_c \ell \nu$ final states, allowing for a test of lepton flavour universality and no deviation in the moments from the expectation of unity is observed. Furthermore, a non-linear transformation is applied to the nominal measured moments in order to directly determine the central moments, which are less correlated compared to the systematic correlations of the nominal moments.

The $\langle q^n \rangle$ moments are crucial experimental inputs for a novel approach to determine the quark mixing parameter $|V_{cb}|$ from inclusive decays, outlined in [6]. By exploiting reparametrization invariance the set of non-perturbative hadronic matrix elements present in the heavy quark expansion is greatly reduced from thirteen to eight at $\mathcal{O}(1/m_b^4)$, allowing for a model-independent determination of $|V_{cb}|$. Using the software package provided by [140], which includes theoretical predictions of

the $\langle q^n \rangle$ moments and the semileptonic decay rate up to $\mathcal{O}(1/m_b^4)$, a simultaneous fit to preliminary measured $\langle q^n \rangle$ moments presented in [139] is performed. Conservative theory uncertainties are employed to determine $|V_{cb}|$ together with three heavy quark parameters at $\mathcal{O}(1/m_b^4)$. A value of $|V_{cb}|$ is found to be:

$$|V_{cb}| \times 10^3 = 41.7 \pm 1.2\,. \tag{7.1}$$

Here, the final quoted uncertainty is the linear sum of the experimental and theoretical errors. This result is consistent with the current world average of inclusive $|V_{cb}|$ as well as previous determinations. The successful fit serves as an important test for the novel, complementary method to determine inclusive $|V_{cb}|$ in a fully data-driven manner.

Even though the $\langle q^n \rangle$ moments are already measured with high experimental precision, several improvements in the analysis strategy could lead to an even more precise measurement as well as an improvement in the precision of the extracted value of $|V_{cb}|$. Since the dominant contribution on the measured q^2 composition stems from heavier charmed final states and high multiplicity decays that are poorly constrained by current measurements, an improvement in the current understanding of these decays by exploiting the larger data set expected to be collected by the Belle Collaboration's successor, Belle II, could improve the overall precision of the measurement. As mentioned, one of the leading sources of systematic uncertainty on the measured q^2 moments is observed to be the modelling of the $B \to X_u \ell \nu$ decays. Events originating from this background component are first rejected by imposing selection criteria on decay kinematics, after which the remaining component is further suppressed by making use of the background subtraction procedure. A possible future improvement of the analysis strategy is outlined in [142]. Rather than subtracting the $B \to X_u \ell \nu$ component, the authors suggest measuring the full $B \to X \ell \nu$ spectrum and obtaining the $B \to X_u \ell \nu$ contribution precisely from within the HQE. This strategy has the potential to not only increase the precision of the measured $\langle q^n \rangle$ moments, but also reduce the uncertainty on the measured value of $|V_{cb}|$ determined from a global fit to the moments. Due to the high precision of the measured $\langle q^n \rangle$ moments, the treatment of theoretical uncertainties on the extraction of $|V_{cb}|$ is imperative. An updated determination of $|V_{cb}|$ using the measured nominal and central moments presented in this thesis is expected with anticipation and excitement.

Appendix A

Monte Carlo corrections

TABLE A.1: The updated branching fractions of semileptonic D meson decays. All values are from [30], while the values of the unmeasured non-resonant decays (labeled with "†") are assumed based on the measured isospin-symmetry counterpart.

$\mathcal{B}$	Value D^+	Value D^0
$D \to K^\star e^+ \nu$	$(5.40 \pm 0.10) \times 10^{-2}$	$(2.15 \pm 0.16) \times 10^{-2}$
$D \to K^\star \mu^+ \nu$	$(5.27 \pm 0.15) \times 10^{-2}$	$(1.89 \pm 0.24) \times 10^{-2}$
$D \to K e^+ \nu$	$(8.73 \pm 0.10) \times 10^{-2}$	$(3.54 \pm 0.035) \times 10^{-2}$
$D \to K \mu^+ \nu$	$(8.76 \pm 0.19) \times 10^{-2}$	$(3.41 \pm 0.04) \times 10^{-2}$
$D \to \pi e^+ \nu$	$(3.72 \pm 0.17) \times 10^{-3}$	$(2.91 \pm 0.04) \times 10^{-3}$
$D \to \pi \mu^+ \nu$	$(3.50 \pm 0.15) \times 10^{-3}$	$(2.67 \pm 0.12) \times 10^{-3}$
$D \to \eta \ell^+ \nu$	$(1.11 \pm 0.07) \times 10^{-3}$	-
$D \to \rho e^+ \nu$	$(2.18^{+0.17}_{-0.25}) \times 10^{-3}$	$(1.5 \pm 0.12) \times 10^{-2}$
$D \to \rho \mu^+ \nu$	$(2.4 \pm 0.4) \times 10^{-3}$	$(1.5 \pm 0.12) \times 10^{-2}$
$D \to \omega \ell^+ \nu$	$(1.69 \pm 0.11) \times 10^{-3}$	-
$D \to K^- \pi^+ e^+ \nu$ non-resonant	$(0.7 \pm 0.7) \times 10^{-2}$	-
$D \to K^- \pi^+ \mu^+ \nu$ non-resonant	$(1.9 \pm 0.5) \times 10^{-3}$	-
$D \to K \pi^0 e^+ \nu$ non-resonant	$(3.5 \pm 3.5) \times 10^{-3}$ †	$(7.9 \pm 1.67) \times 10^{-4}$
$D \to K \pi^0 \mu^+ \nu$ non-resonant	$(0.95 \pm 0.95) \times 10^{-3}$ †	$(7.9 \pm 1.67) \times 10^{-4}$
$D \to K^0 \pi^- \ell^+ \nu$ non-resonant	-	$(1.58 \pm 0.33) \times 10^{-3}$ †
$D \to K_2^\star \ell^+ \nu$	0	0
$D \to K_1 \ell^+ \nu$	0	-

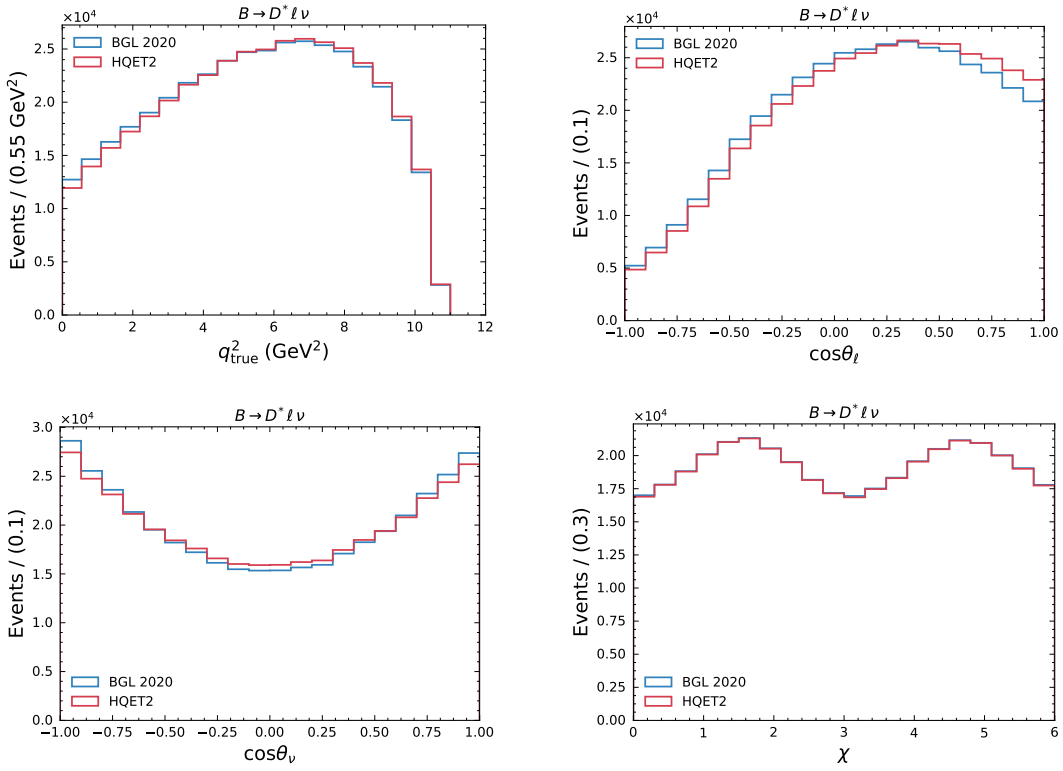

FIGURE A.1: Comparison of kinematic variables describing $B \to D^* \ell \nu$ decays. The variables shown are q^2, $\cos\theta_\ell$, $\cos\theta_\nu$ and χ. The original MC samples were simulated using the CLN form factor parametrization (red) that are reweighted to the BGL (blue) model.

TABLE A.2: Variables, number of bins and range of the generator-level histograms used to reweight each of the D^{**} types from the ISGW2 to the LLSW form factor model.

D^{**} Type	Variable	Number of bins	Range
D_1	w	50	[1.0,1.42]
D_2^*	w	50	[1.0,1.42]
D_1^*	w	50	[1.0,1.425]
D_0^*	w	50	[1.0,1.425]

Appendix B

Distributions of kinematic variables

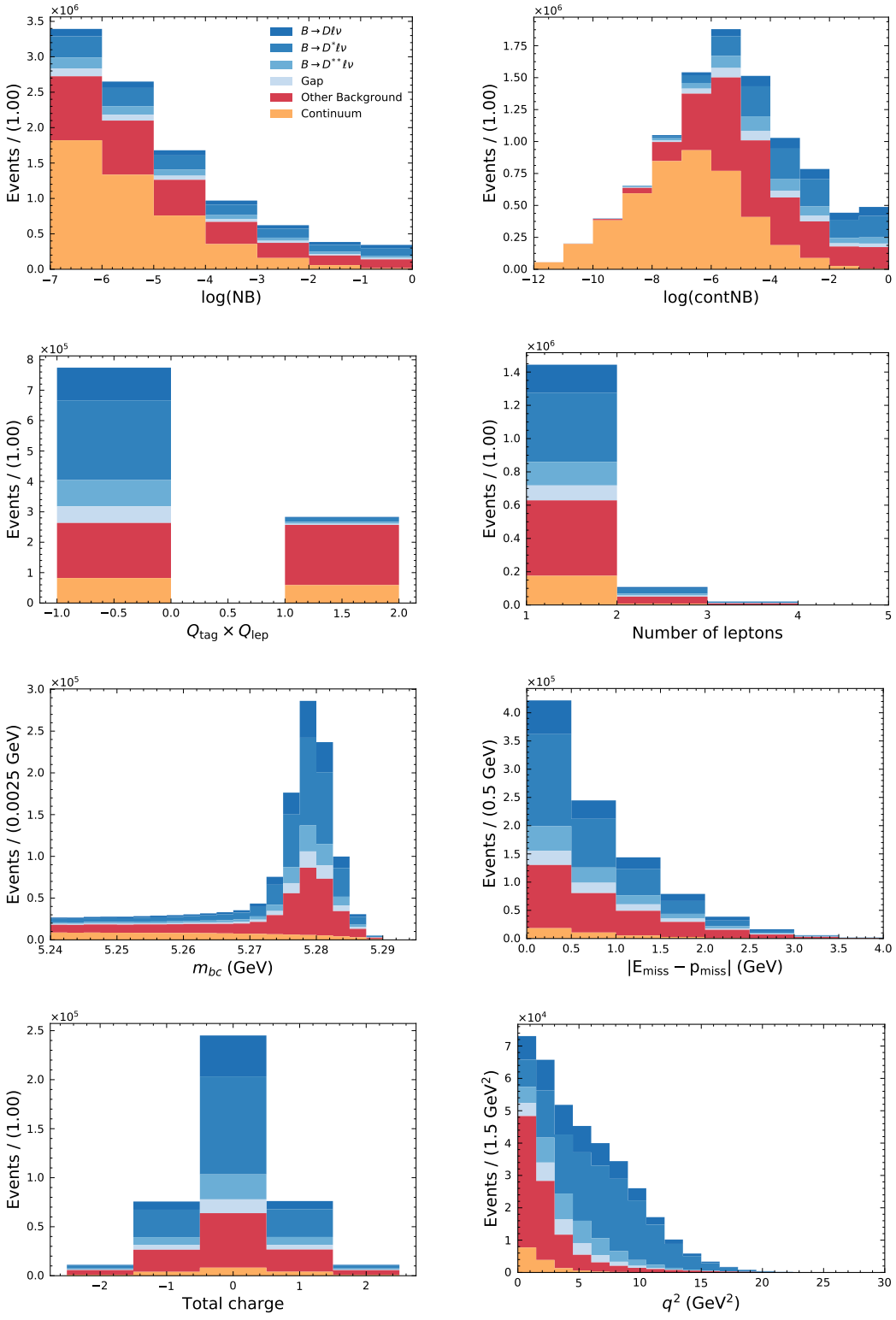

FIGURE B.1: Reconstructed MC distributions of the kinematic variables for which signal selection requirements are applied, as listed in Table 4.3. Each of the exclusive signal modes, continuum, and "other background" which consists of fakes, secondary leptons and $B \to X_u \ell\nu$ decays is shown. The selection criteria have been applied cumulatively, except those affecting the variable under investigation.

Appendix C

Signal probability weights

C.1 Signal probability functions for electron candidates

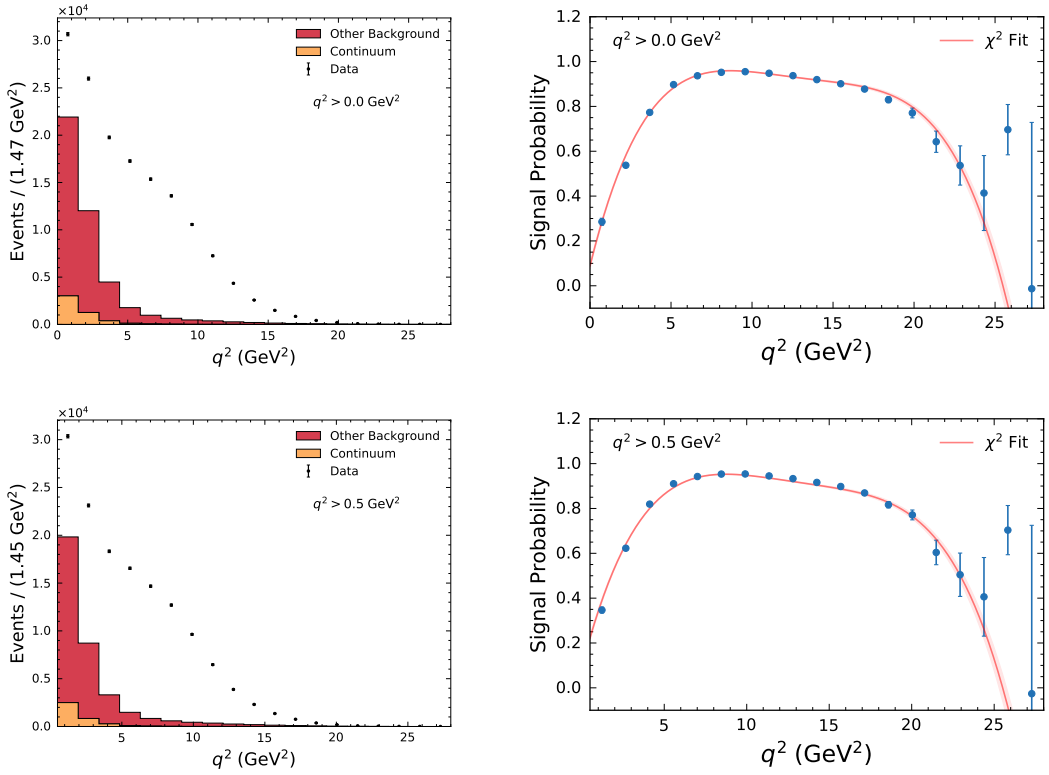

FIGURE C.1: The left column shows the extracted post-fit and data q^2 distributions for electron candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

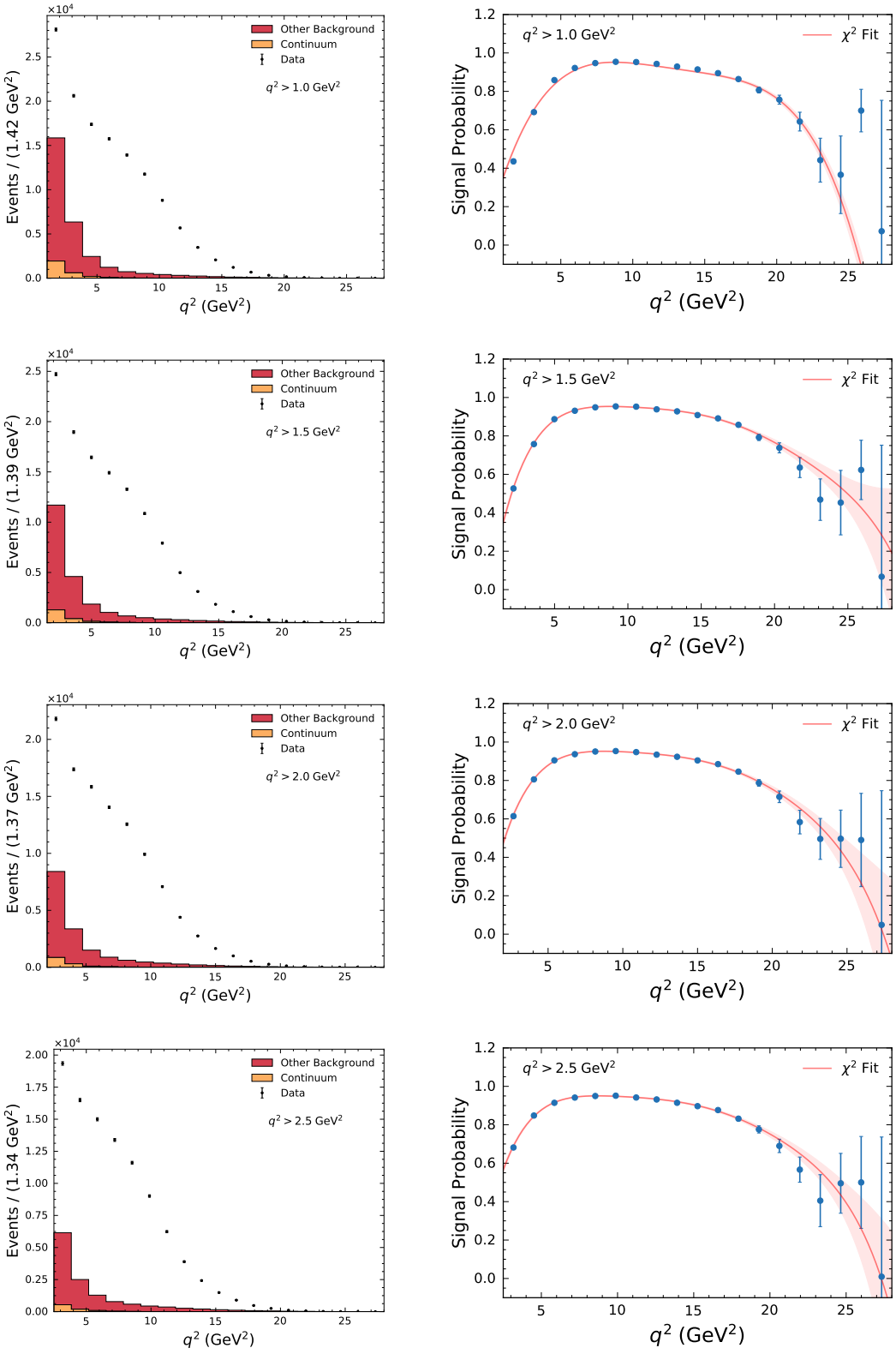

FIGURE C.2: The left column shows the extracted post-fit and data q^2 distributions for electron candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

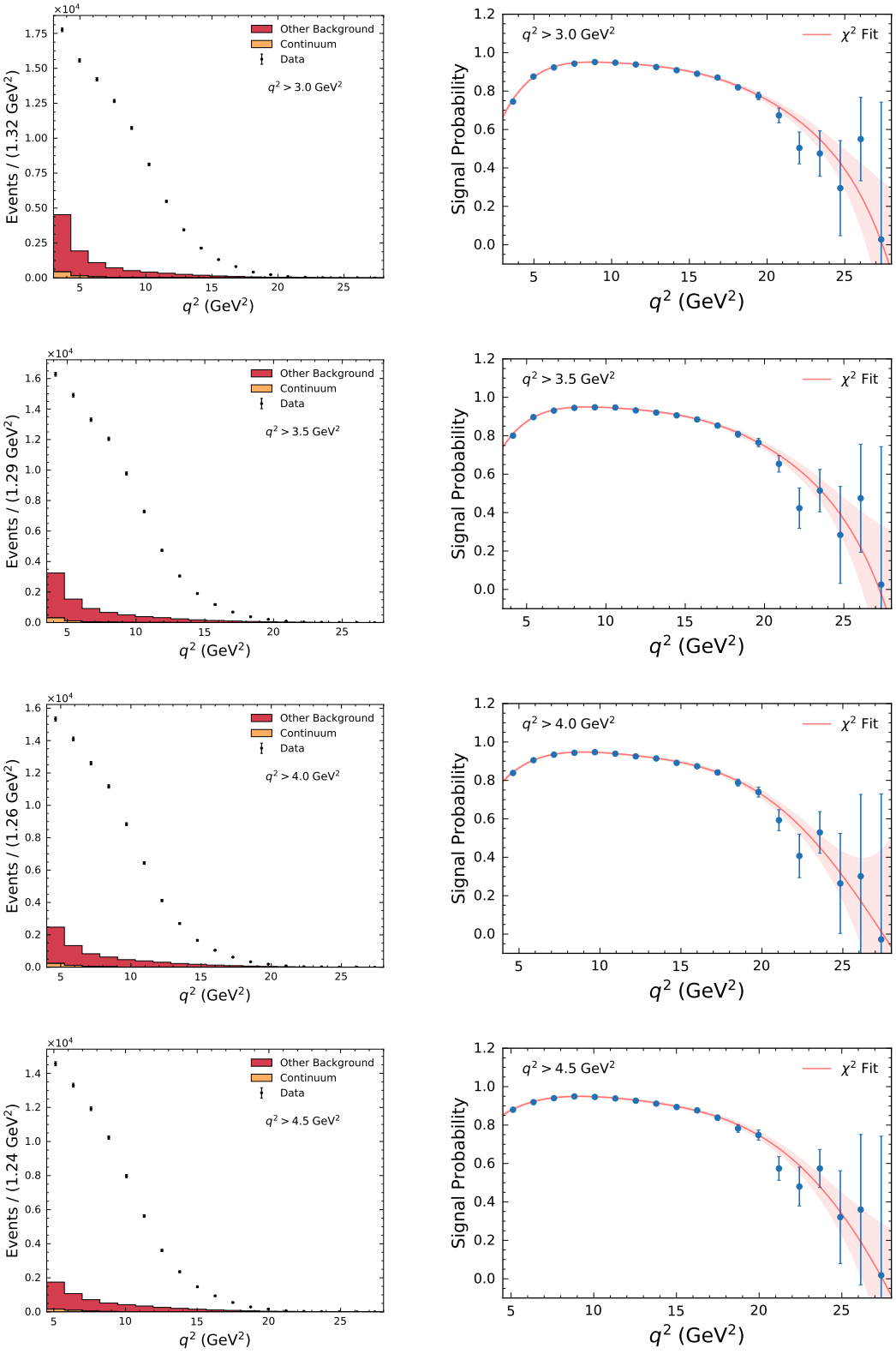

FIGURE C.3: The left column shows the extracted post-fit and data q^2 distributions for electron candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

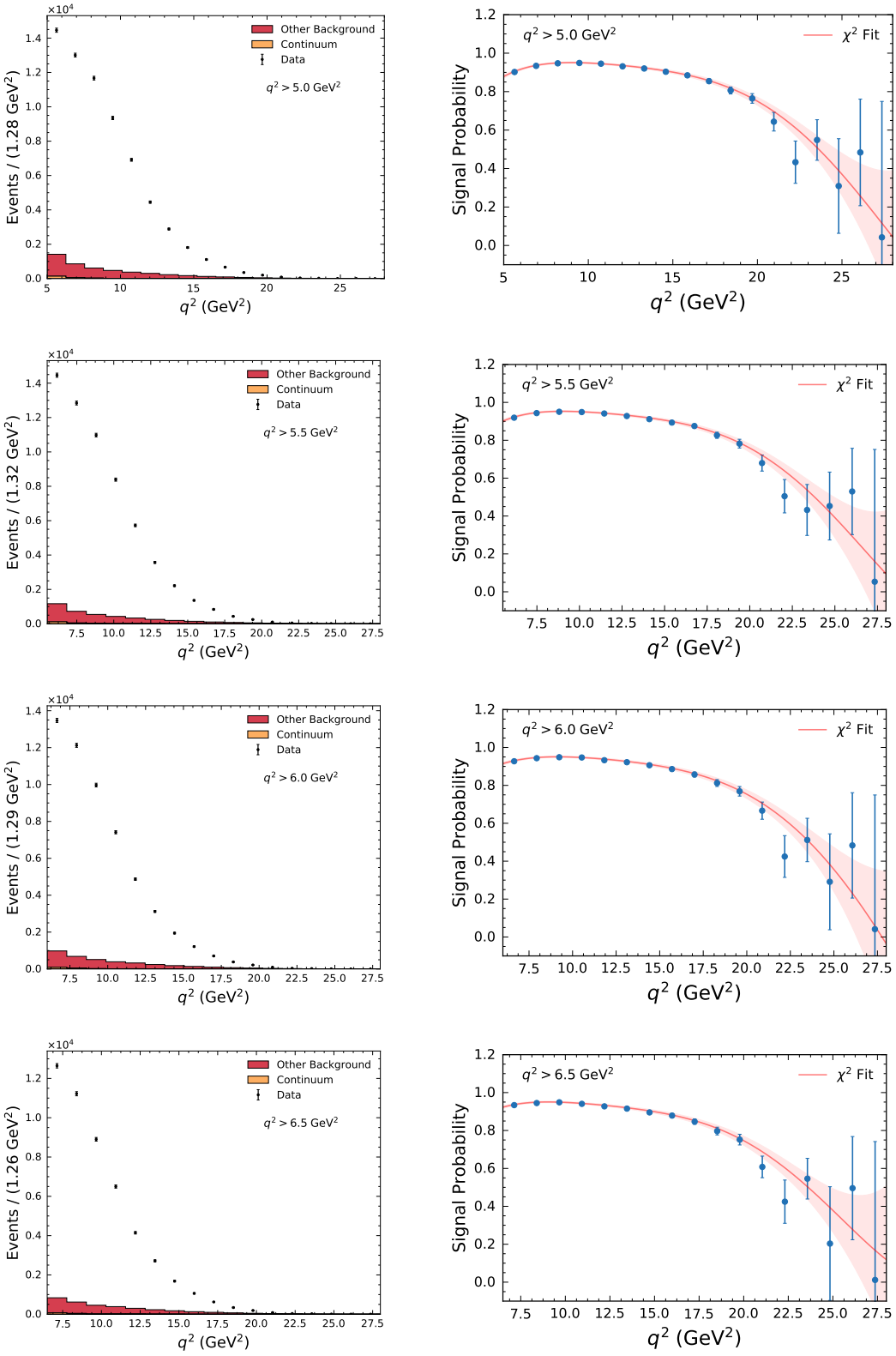

FIGURE C.4: The left column shows the extracted post-fit and data q^2 distributions for electron candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

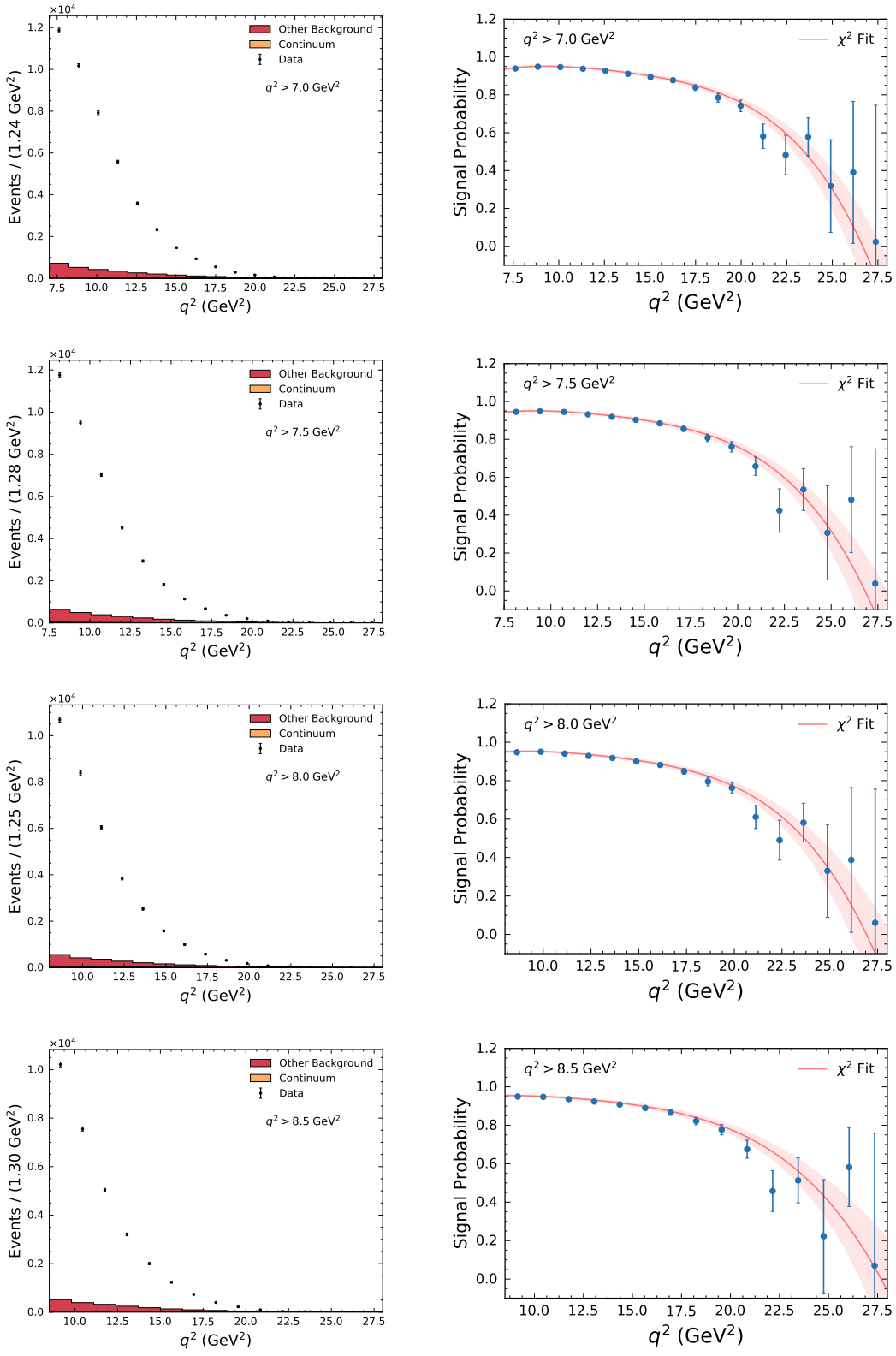

FIGURE C.5: The left column shows the extracted post-fit and data q^2 distributions for electron candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

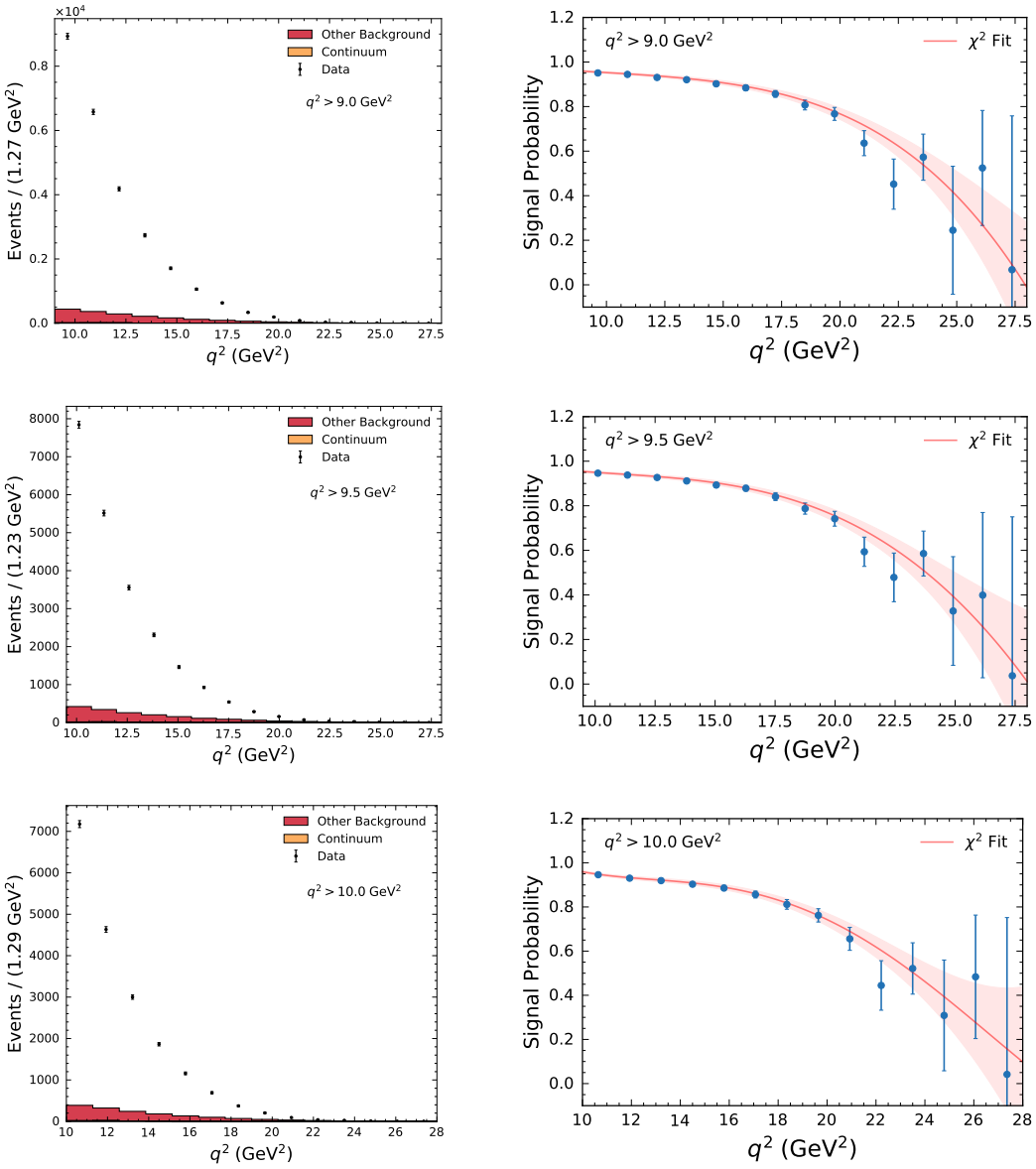

FIGURE C.6: The left column shows the extracted post-fit and data q^2 distributions for electron candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

C.2 Signal probability functions for muon candidates

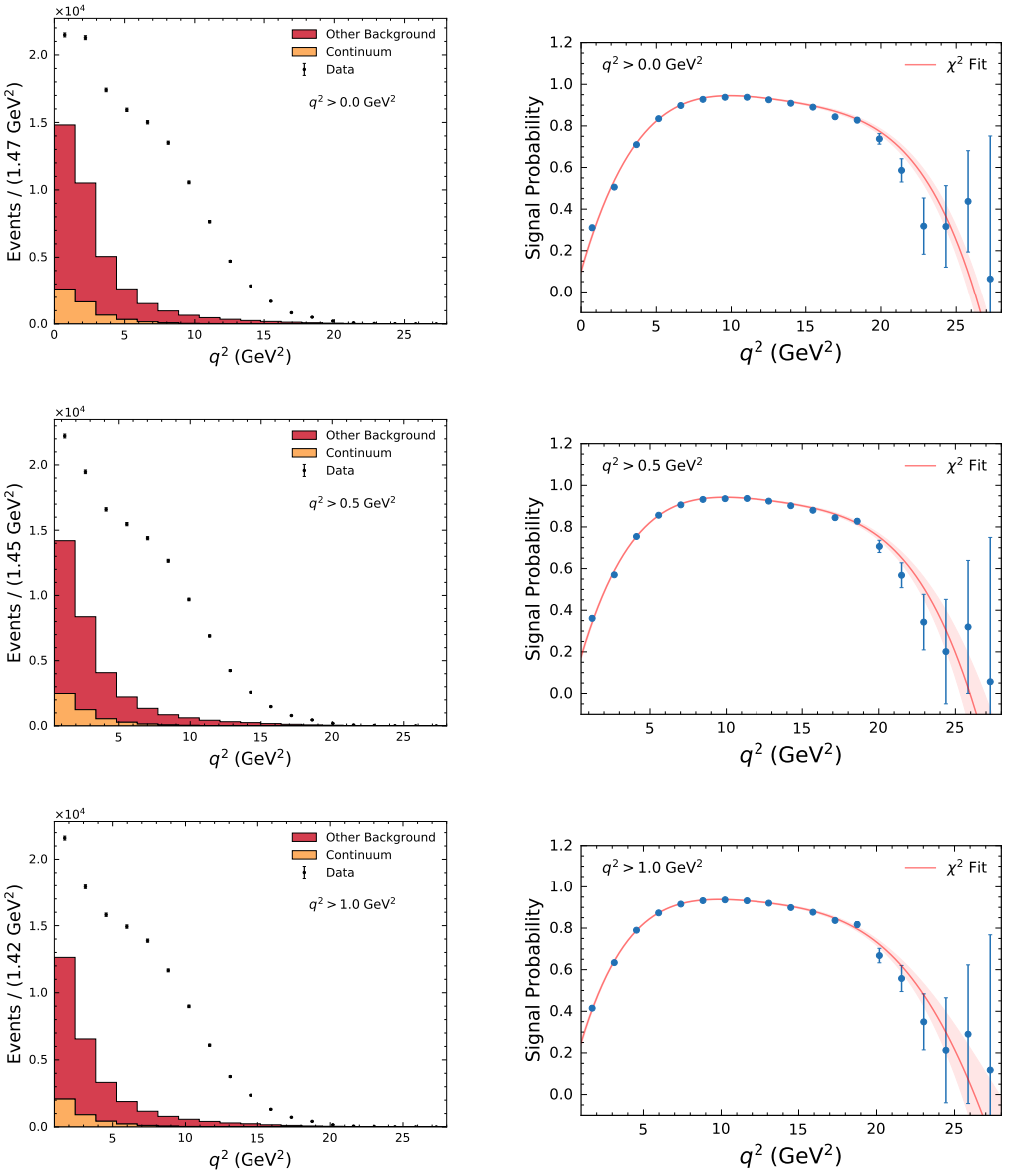

FIGURE C.7: The left column shows the extracted post-fit and data q^2 distributions for muon candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

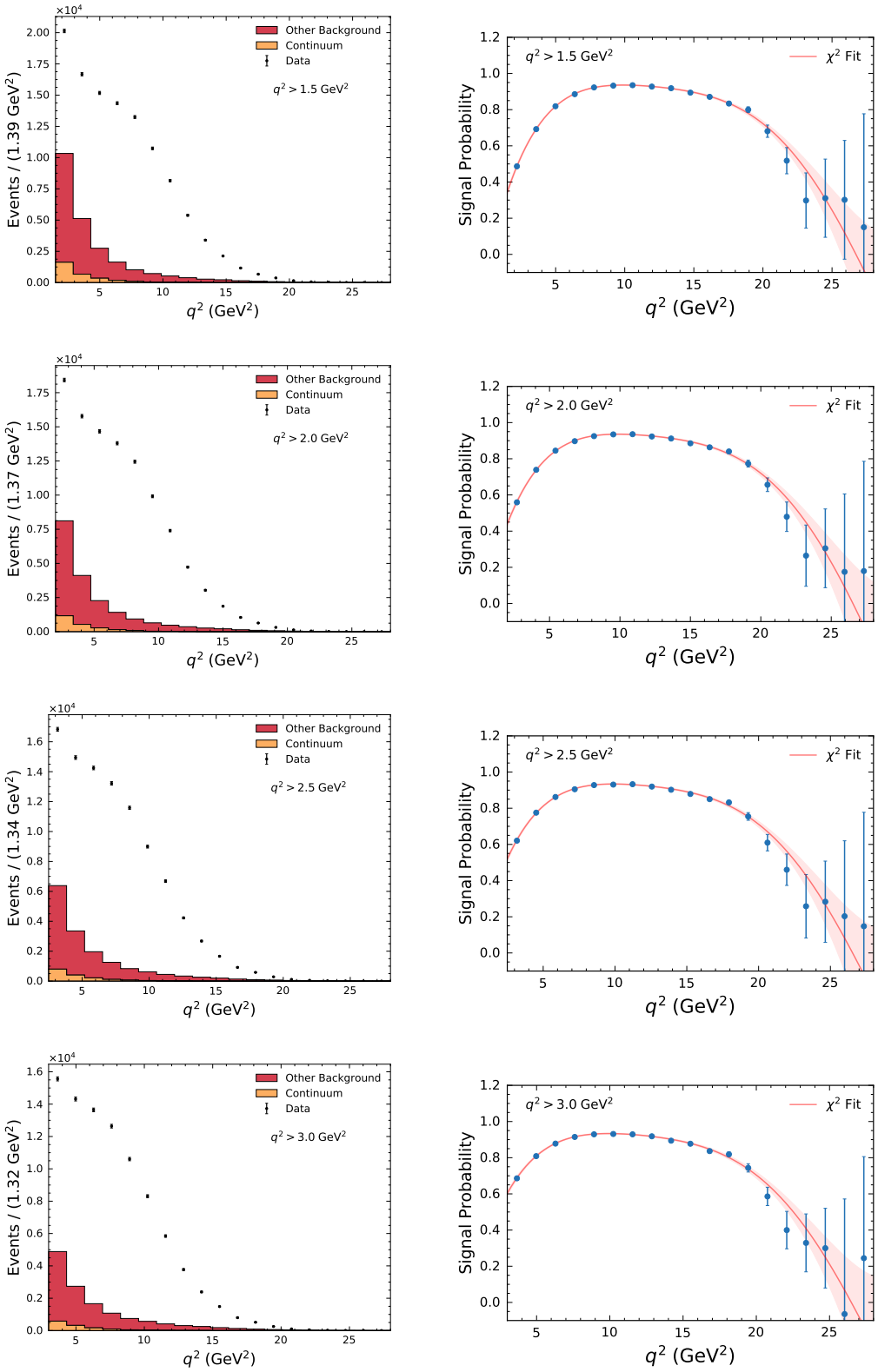

FIGURE C.8: The left column shows the extracted post-fit and data q^2 distributions for muon candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

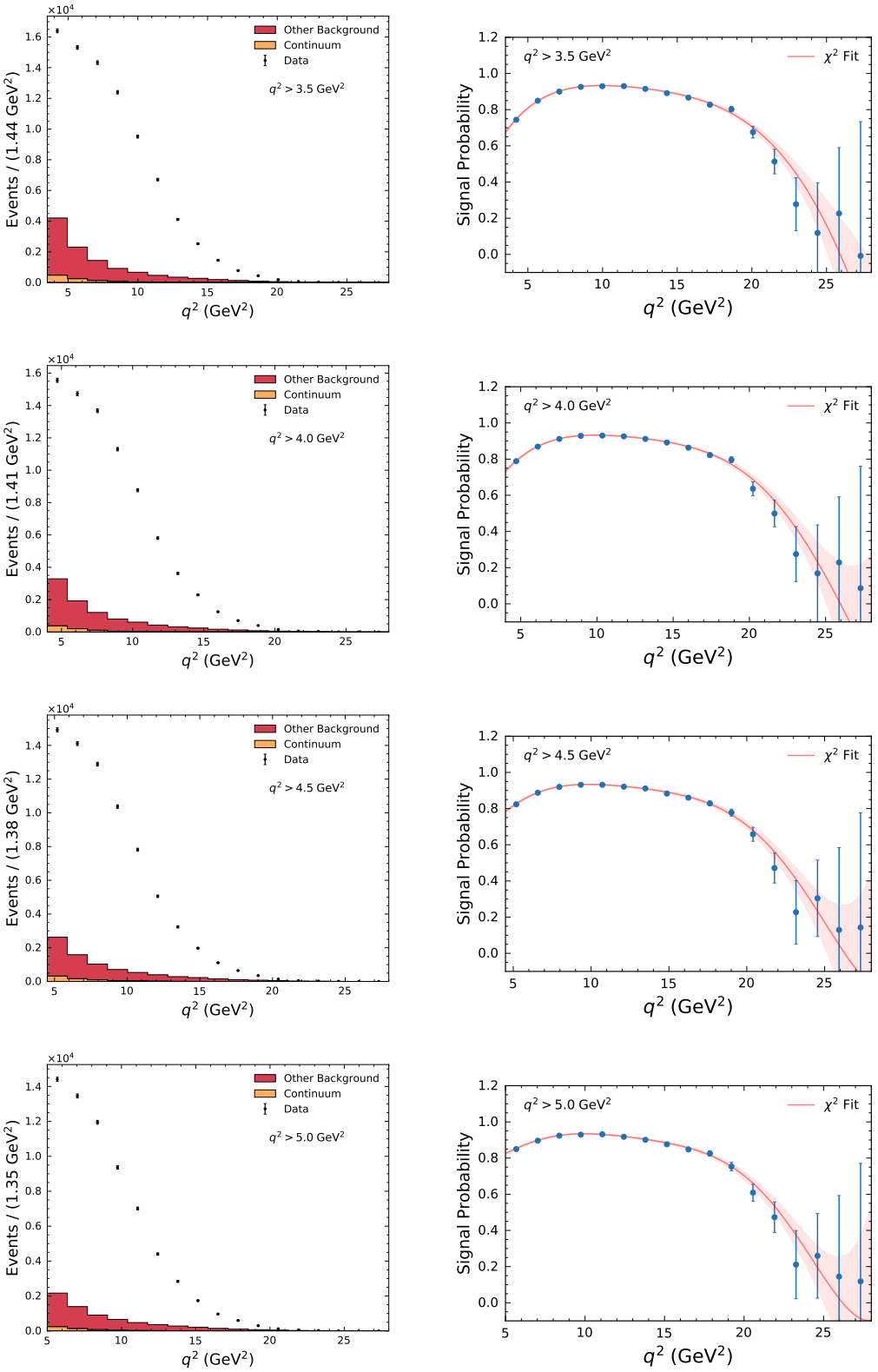

FIGURE C.9: The left column shows the extracted post-fit and data q^2 distributions for muon candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

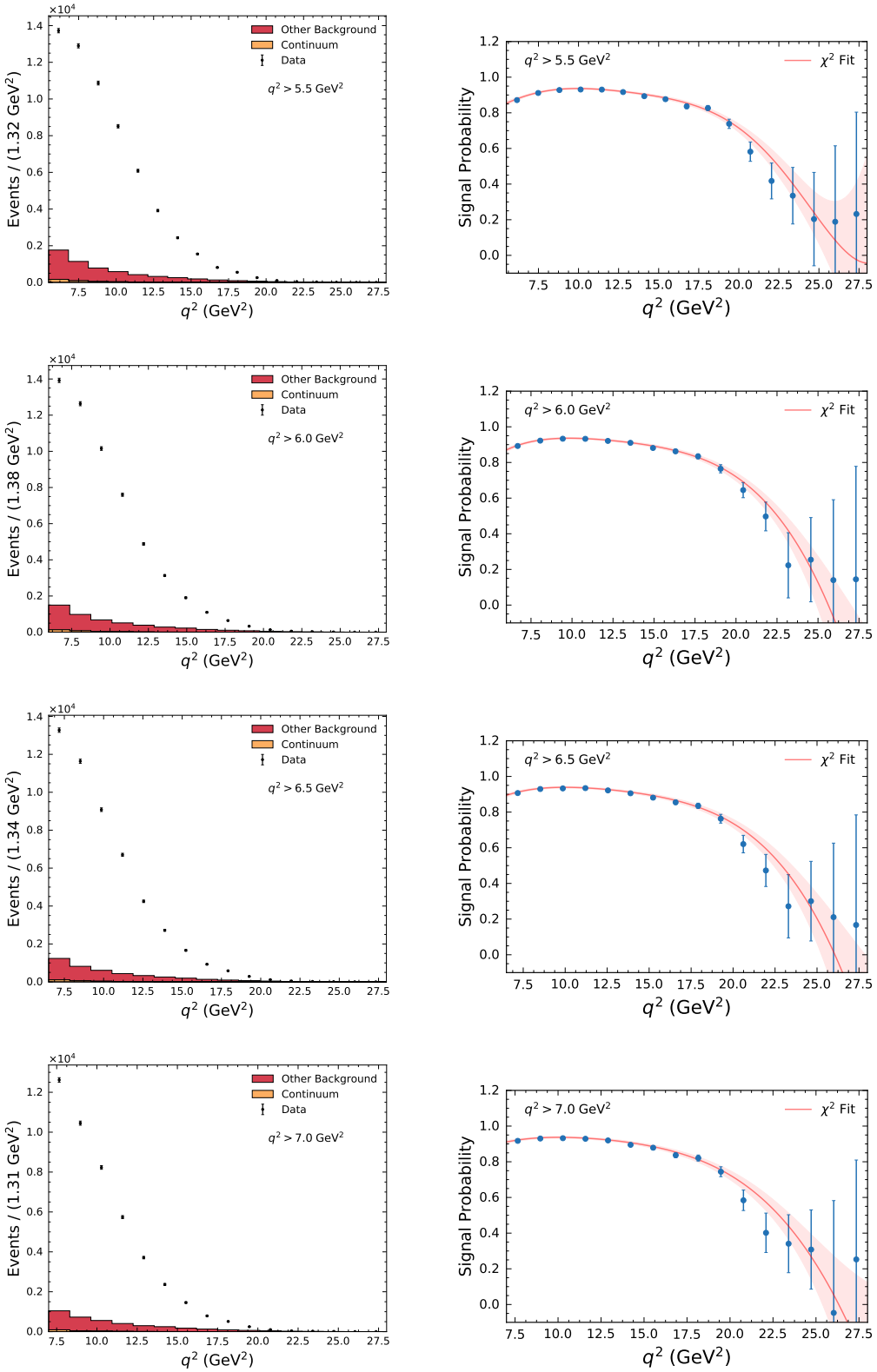

FIGURE C.10: The left column shows the extracted post-fit and data q^2 distributions for muon candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

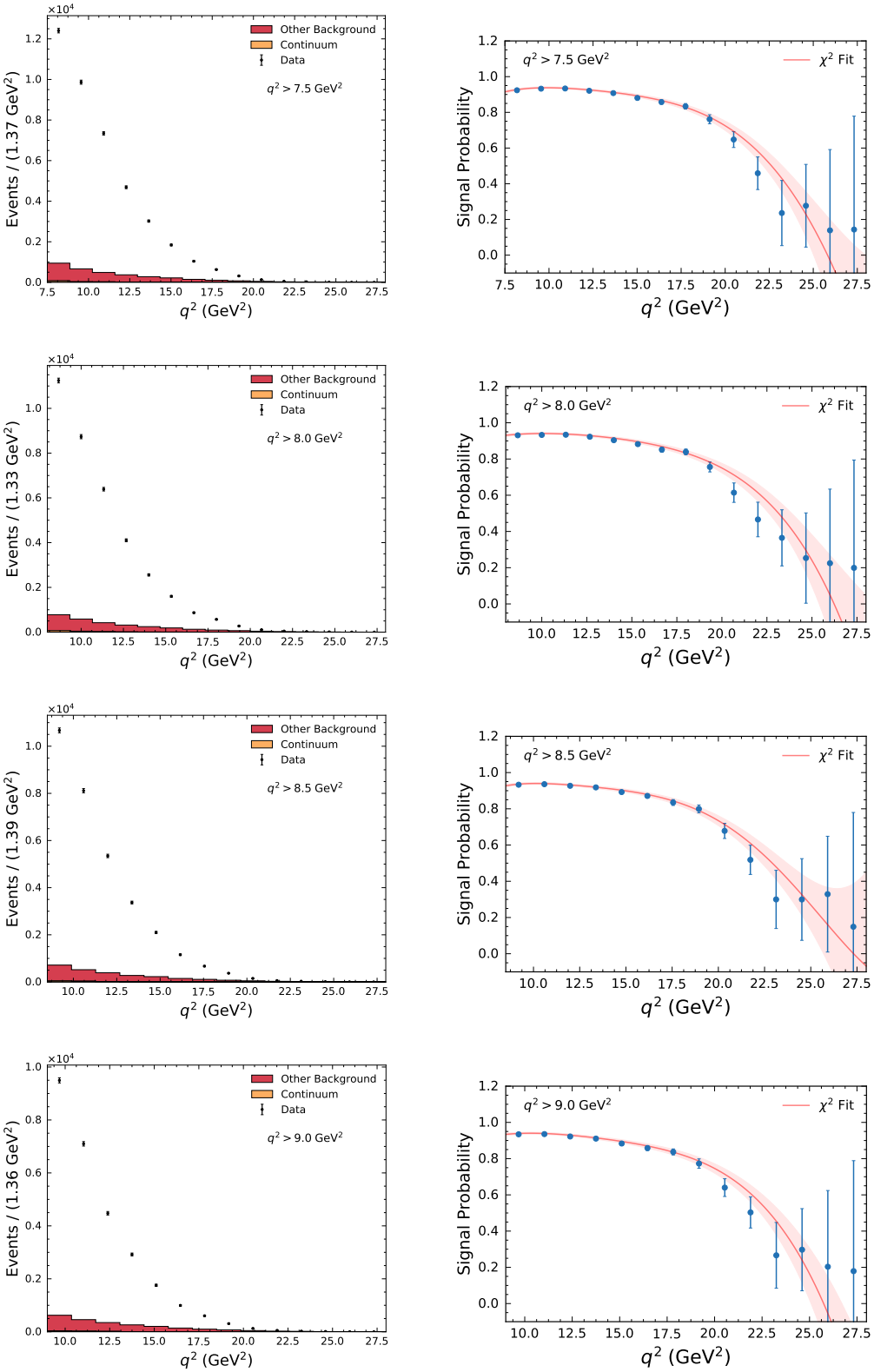

FIGURE C.11: The left column shows the extracted post-fit and data q^2 distributions for muon candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

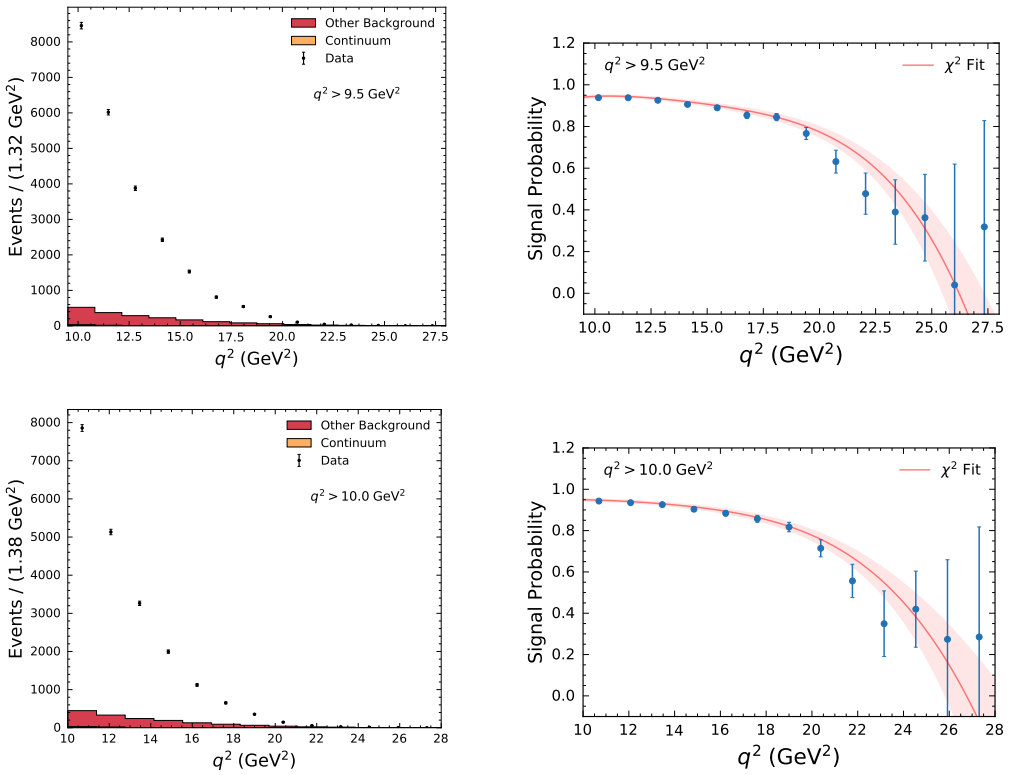

FIGURE C.12: The left column shows the extracted post-fit and data q^2 distributions for muon candidates for different lower q^2 selections. The corresponding bin-wise signal probability weights, $w_i(q^2)$, together with a fitted polynomial of a given order n are shown in the right column.

Appendix D

Calibration studies for muon candidates

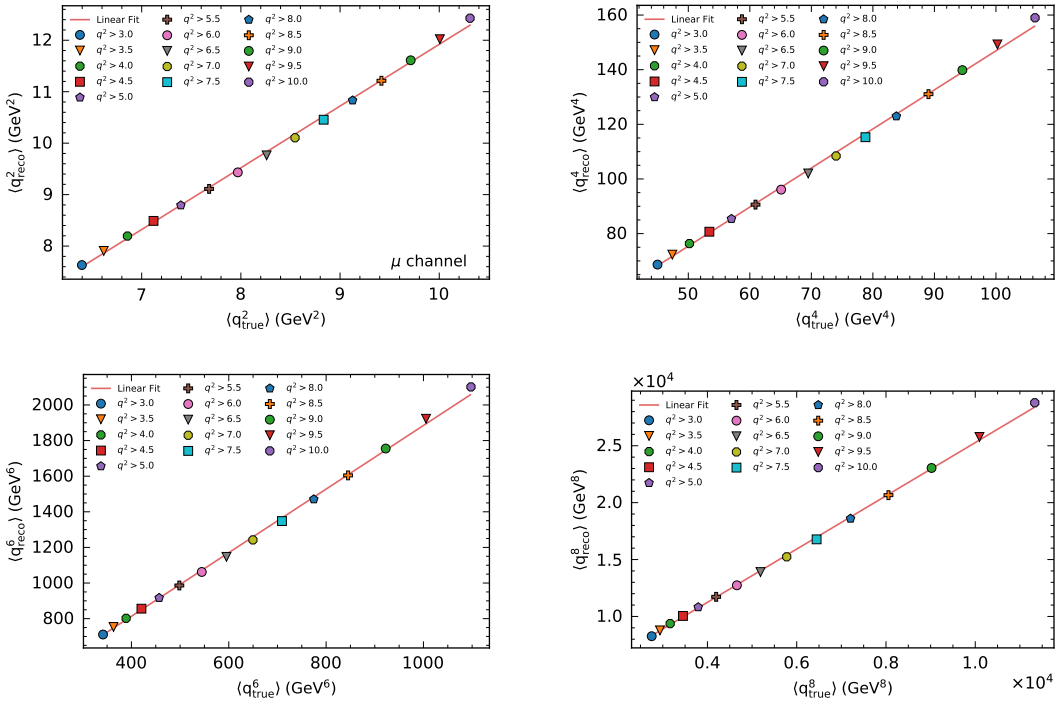

FIGURE D.1: Calibration curves for $\langle q^{2,4,6,8} \rangle$ for muon candidates. Shown here are the the extracted $\langle q^n_{,\text{reco}} \rangle$ versus $\langle q^n_{,\text{true}} \rangle$ as a function of lower q^2 cut, together with the fitted linear calibration curves.

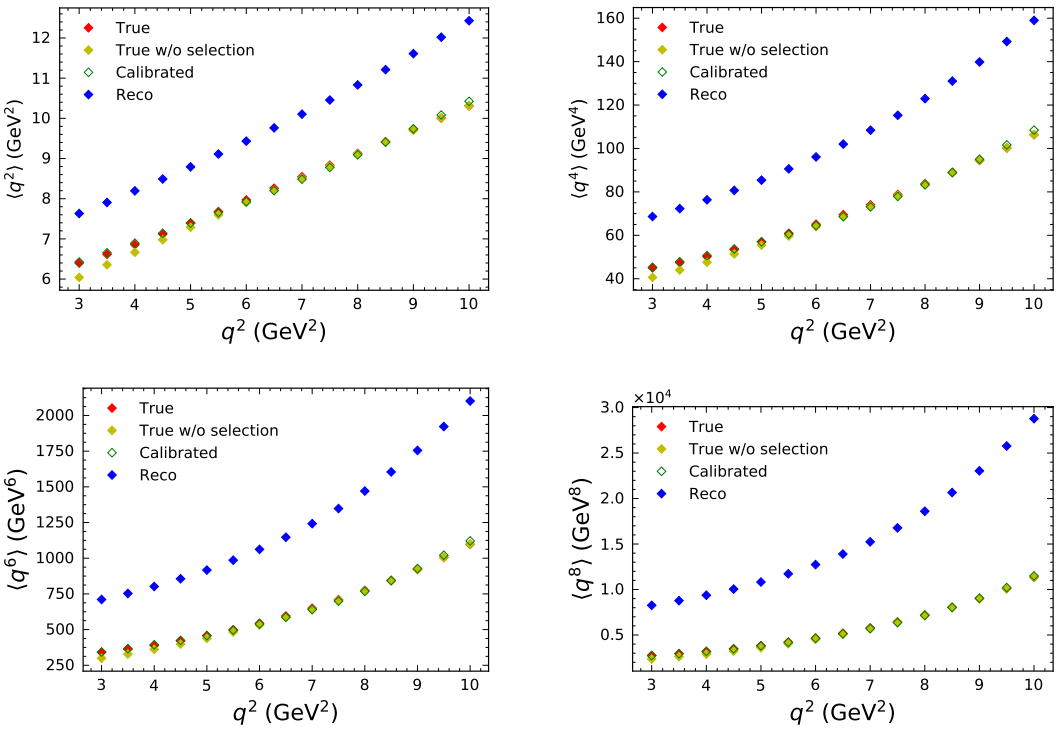

FIGURE D.2: Moments of q^n calculated on simulated $B \to X_c \mu \nu$ decays for different lower cuts on q^2. Shown here are the reconstructed, calibrated and true moments after all selection criteria have been applied. Additionally, the true moments without any selection criteria are shown as the yellow diamonds.

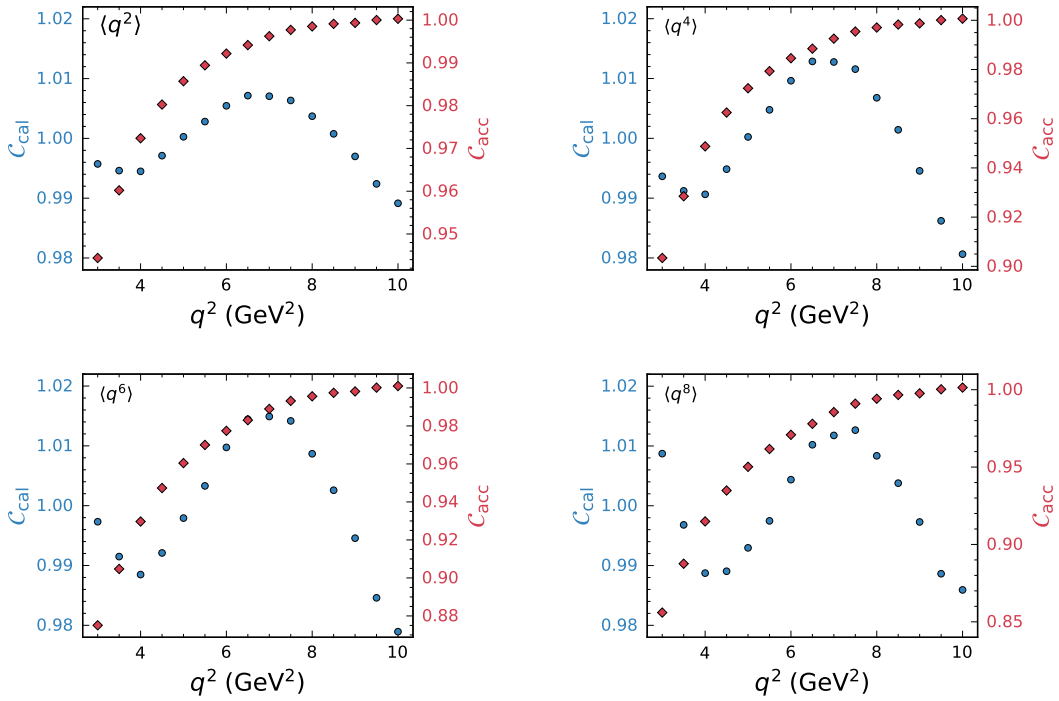

FIGURE D.3: Remaining bias after application of the calibration curves on simulated $B \to X_c \mu\nu$ decays as a function of q^2. Shown here are the bias correction factors $\mathcal{C}_{\text{cal}}$ (blue) and $\mathcal{C}_{\text{acc}}$ (red).

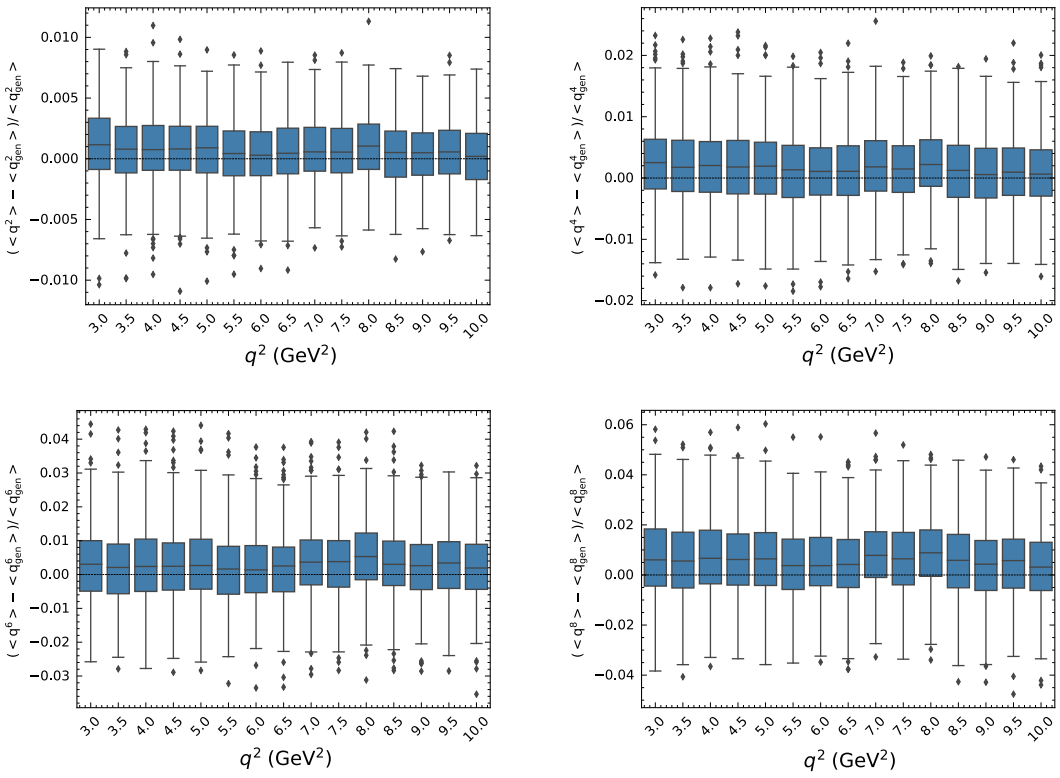

FIGURE D.4: Box plot of the sample test relative error distributions for q^n moments as a function of lower level q^2 selections in the muon channel.

Appendix E

$B \to X_c \ell \nu$ modelling systematics

TABLE E.1: Summary of systematic uncertainties due to the $B \to X_c \ell \nu$ composition for the moments $\langle q^{2,4,6,8} \rangle$ for electron candidates. The values are given as the total relative error in permille.

q^2 selection in GeV^2	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$																					
$B \to Dl\nu$	0.412	0.358	0.326	0.276	0.215	0.174	0.081	0.063	0.051	0.042	0.005	0.010	0.019	0.026	0.034	0.039	0.053	0.058	0.066	0.064	0.051
$B \to D^* l\nu$	1.075	0.814	0.660	0.554	0.496	0.442	0.538	0.472	0.407	0.340	0.283	0.218	0.159	0.090	0.020	0.033	0.060	0.075	0.080	0.106	0.137
$B \to D_1 l\nu$	4.269	5.247	5.300	4.844	4.554	3.854	4.248	3.262	2.639	1.793	1.033	0.333	0.213	0.870	1.299	1.436	1.319	1.011	0.724	0.582	0.512
$B \to D_1' l\nu$	9.339	5.682	3.369	1.518	0.291	0.639	0.286	0.797	1.116	1.420	1.724	1.732	1.569	1.390	1.203	0.991	0.891	0.753	0.612	0.528	0.531
$B \to D_0^* l\nu$	9.351	5.828	3.714	2.196	1.123	0.426	0.549	0.044	0.187	0.446	0.624	0.602	0.544	0.560	0.597	0.535	0.486	0.345	0.164	0.167	0.194
$B \to D_2^* l\nu$	1.891	1.816	1.702	1.387	1.059	0.713	1.014	0.814	0.357	0.116	0.342	0.486	0.666	0.678	0.858	0.754	0.535	0.391	0.412	0.404	0.349
Gap	23.445	11.239	4.948	0.989	1.496	3.021	2.160	3.178	3.931	3.886	3.918	3.833	3.311	2.924	2.556	2.241	1.772	1.408	0.943	0.892	0.794
Total rel. error in ‰	27.339	14.973	9.009	5.820	5.074	5.030	4.942	4.719	4.898	4.546	4.469	4.295	3.773	3.467	3.281	2.988	2.491	1.963	1.413	1.273	1.164
$\langle q^4 \rangle$																					
$B \to Dl\nu$	0.807	0.720	0.634	0.558	0.477	0.392	0.077	0.052	0.034	0.019	0.046	0.071	0.088	0.101	0.118	0.127	0.153	0.163	0.178	0.173	0.144
$B \to D^* l\nu$	1.613	1.388	1.212	1.047	0.915	0.809	1.100	0.960	0.824	0.682	0.560	0.426	0.304	0.167	0.031	0.071	0.123	0.151	0.161	0.211	0.271
$B \to D_1 l\nu$	5.701	6.406	6.423	6.122	5.212	4.226	7.361	5.592	4.385	2.807	1.373	0.065	0.953	2.149	2.914	3.125	2.848	2.209	1.615	1.320	1.164
$B \to D_1' l\nu$	10.838	6.903	4.315	2.103	0.363	0.935	1.375	2.232	2.774	3.284	3.771	3.745	3.405	3.026	2.632	2.196	1.971	1.678	1.381	1.203	1.189
$B \to D_0^* l\nu$	10.677	7.243	4.663	2.272	0.917	0.242	0.673	0.192	0.612	1.074	1.386	1.341	1.222	1.238	1.288	1.149	1.036	0.747	0.391	0.393	0.441
$B \to D_2^* l\nu$	2.549	2.172	1.961	1.649	1.183	0.645	1.474	1.074	0.272	0.563	0.973	1.235	1.548	1.553	1.859	1.637	1.200	0.910	0.940	0.915	0.792
Gap	25.386	13.270	4.538	2.080	6.635	9.440	5.415	7.139	8.404	8.322	8.321	8.082	7.048	6.250	5.472	4.794	3.828	3.069	2.127	2.003	1.787
Total rel. error in ‰	28.914	17.016	11.383	9.058	9.461	10.154	9.447	9.452	9.934	9.479	9.409	9.102	8.134	7.537	7.106	6.449	5.404	4.307	3.183	2.875	2.622
$\langle q^6 \rangle$																					
$B \to Dl\nu$	0.926	0.818	0.706	0.607	0.504	0.397	0.020	0.053	0.078	0.098	0.183	0.216	0.237	0.253	0.275	0.285	0.319	0.329	0.347	0.334	0.283
$B \to D^* l\nu$	1.895	1.717	1.581	1.440	1.317	1.206	1.563	1.384	1.200	1.002	0.823	0.627	0.446	0.244	0.047	0.103	0.180	0.221	0.236	0.312	0.400
$B \to D_1 l\nu$	5.017	5.863	6.047	5.896	5.045	4.054	8.608	6.459	4.938	2.883	0.974	0.785	2.162	3.774	4.788	5.028	4.568	3.589	2.684	2.228	1.971
$B \to D_1' l\nu$	9.181	5.359	2.881	0.700	1.061	2.407	3.375	4.341	4.949	5.538	6.093	5.987	5.460	4.874	4.266	3.602	3.232	2.770	2.305	2.023	1.972
$B \to D_0^* l\nu$	9.510	6.199	3.716	1.315	0.038	1.252	0.188	0.824	1.317	1.887	2.270	2.188	2.005	2.008	2.050	1.830	1.644	1.209	0.688	0.686	0.747
$B \to D_2^* l\nu$	1.615	1.310	1.212	0.983	0.554	0.012	1.135	0.665	0.320	1.375	1.895	2.228	2.623	2.605	2.991	2.644	1.996	1.562	1.584	1.531	1.327
Gap	18.074	6.416	1.988	8.458	12.949	15.710	10.104	12.005	13.426	13.182	13.044	12.583	11.066	9.858	8.660	7.593	6.127	4.956	3.546	3.326	2.973
Total rel. error in ‰	22.665	13.143	11.187	12.061	14.123	15.535	9.447	9.452	9.934	9.479	9.409	9.102	8.134	7.537	7.106	6.449	5.404	4.307	3.183	2.875	2.622
$\langle q^8 \rangle$																					
$B \to Dl\nu$	0.444	0.347	0.268	0.151	0.021	0.073	0.221	0.269	0.304	0.332	0.432	0.469	0.489	0.504	0.524	0.529	0.563	0.566	0.578	0.551	0.472
$B \to D^* l\nu$	2.185	2.037	1.945	1.883	1.829	1.747	1.871	1.698	1.503	1.280	1.067	0.826	0.594	0.334	0.079	0.117	0.222	0.279	0.303	0.405	0.522
$B \to D_1 l\nu$	6.378	7.751	8.119	7.756	7.629	6.647	7.779	5.673	4.195	1.987	0.134	2.144	3.747	5.656	6.848	7.092	6.448	5.143	3.934	3.316	2.947
$B \to D_1' l\nu$	3.431	0.077	2.552	4.675	6.197	7.447	6.240	7.084	7.583	8.110	8.622	8.387	7.664	6.872	6.054	5.171	4.644	4.007	3.370	2.977	2.867
$B \to D_0^* l\nu$	8.493	5.030	2.773	1.066	0.352	1.331	0.942	1.897	2.321	2.882	3.258	3.110	2.856	2.835	2.858	2.561	2.300	1.730	1.062	1.051	1.115
$B \to D_2^* l\nu$	0.287	0.131	0.003	0.188	0.487	0.911	0.044	0.420	1.421	2.548	3.083	3.424	3.852	3.795	4.226	3.757	2.912	2.339	2.337	2.243	1.950
Gap	7.230	4.690	11.048	14.508	17.223	18.842	16.212	17.759	18.961	18.323	17.900	17.141	15.179	13.583	11.986	10.540	8.594	7.018	5.168	4.836	4.330
Total rel. error in ‰	13.487	10.567	14.354	17.241	19.923	21.455	13.833	14.413	15.245	14.805	14.753	14.317	12.966	12.087	11.373	10.312	8.699	7.013	5.315	4.811	4.378

TABLE E.2: Summary of systematic uncertainties due to the $B \to X_c \ell \nu$ composition for the moments $\langle q^{2,4,6,8} \rangle$ for muon candidates. The values are given as the total relative error in permille.

q^2 selection in GeV^2	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$																					
$B \to Dl\nu$	0.561	0.493	0.427	0.372	0.316	0.260	0.143	0.102	0.065	0.048	0.026	0.005	0.006	0.018	0.022	0.022	0.017	0.019	0.004	0.010	0.008
$B \to D^* l\nu$	1.033	0.778	0.618	0.496	0.415	0.362	0.458	0.399	0.357	0.295	0.231	0.175	0.113	0.055	0.003	0.048	0.088	0.106	0.129	0.139	0.156
$B \to D_1 l\nu$	3.467	4.116	4.161	3.989	3.400	2.797	3.454	2.402	1.652	1.047	0.076	0.290	0.735	1.432	1.442	1.608	1.655	1.339	1.105	0.768	0.681
$B \to D_1' l\nu$	9.695	5.947	3.751	2.053	0.816	0.055	0.081	0.517	0.738	1.142	1.266	1.335	1.263	1.127	1.067	0.866	0.673	0.474	0.361	0.180	0.056
$B \to D_0^* l\nu$	9.213	5.929	3.743	1.913	0.966	0.200	0.191	0.163	0.397	0.724	0.839	0.967	0.961	0.875	0.770	0.727	0.639	0.466	0.457	0.431	0.265
$B \to D_2^* l\nu$	2.076	1.708	1.528	1.306	0.998	0.654	0.733	0.310	0.085	0.315	0.531	0.577	0.734	0.837	0.756	0.669	0.507	0.411	0.311	0.331	0.305
Gap	25.434	13.802	6.334	1.186	2.110	4.031	3.234	4.340	4.782	4.829	4.778	4.672	4.268	3.406	2.819	2.316	1.786	1.566	1.186	0.960	0.827
Total rel. error in ‰	29.043	16.785	9.403	5.223	4.345	4.974	4.817	5.017	5.142	5.141	5.048	4.999	4.672	4.049	3.511	3.111	2.656	2.206	1.755	1.363	1.157
$\langle q^4 \rangle$																					
$B \to Dl\nu$	0.807	0.720	0.634	0.558	0.477	0.392	0.205	0.142	0.082	0.053	0.014	0.022	0.043	0.064	0.071	0.072	0.061	0.065	0.037	0.048	0.040
$B \to D^* l\nu$	1.613	1.388	1.212	1.047	0.915	0.809	0.946	0.819	0.718	0.584	0.452	0.334	0.208	0.091	0.011	0.112	0.188	0.219	0.264	0.282	0.315
$B \to D_1 l\nu$	5.701	6.406	6.423	6.122	5.212	4.226	5.547	3.711	2.333	1.196	0.548	1.236	2.044	3.289	3.268	3.524	3.545	2.876	2.374	1.686	1.491
$B \to D_1' l\nu$	10.838	6.903	4.315	2.103	0.363	0.935	0.758	1.489	1.875	2.550	2.744	2.832	2.659	2.370	2.214	1.796	1.394	0.992	0.754	0.391	0.145
$B \to D_0^* l\nu$	10.677	7.243	4.663	2.272	0.917	0.242	0.117	0.738	1.158	1.719	1.914	2.121	2.083	1.894	1.671	1.561	1.361	1.006	0.967	0.899	0.560
$B \to D_2^* l\nu$	2.549	2.172	1.961	1.649	1.183	0.645	0.936	0.202	0.218	0.918	1.302	1.384	1.653	1.815	1.636	1.447	1.117	0.914	0.711	0.744	0.683
Gap	25.386	13.270	4.538	2.080	6.635	9.440	7.616	9.427	10.157	10.200	10.027	9.709	8.857	7.154	5.962	4.927	3.837	3.350	2.570	2.091	1.799
Total rel. error in ‰	30.301	18.011	10.389	7.451	8.638	10.447	9.548	10.302	10.678	10.776	10.674	10.504	9.839	8.632	7.523	6.669	5.690	4.730	3.784	2.968	2.522
$\langle q^6 \rangle$																					
$B \to Dl\nu$	0.926	0.818	0.706	0.607	0.504	0.397	0.197	0.113	0.034	0.005	0.057	0.105	0.131	0.158	0.166	0.165	0.147	0.151	0.107	0.120	0.102
$B \to D^* l\nu$	1.895	1.717	1.581	1.440	1.317	1.206	1.334	1.172	1.032	0.845	0.654	0.481	0.294	0.121	0.030	0.177	0.287	0.333	0.397	0.422	0.471
$B \to D_1 l\nu$	5.017	5.863	6.047	5.896	5.045	4.054	5.677	3.496	1.811	0.378	1.871	2.764	3.823	5.466	5.384	5.668	5.609	4.567	3.776	2.737	2.414
$B \to D_1' l\nu$	9.181	5.359	2.881	0.700	1.061	2.407	2.132	2.933	3.346	4.145	4.340	4.395	4.102	3.655	3.386	2.751	2.137	1.537	1.172	0.634	0.275
$B \to D_0^* l\nu$	9.510	6.199	3.716	1.315	0.038	1.252	1.098	1.792	2.271	2.946	3.160	3.394	3.300	2.997	2.651	2.461	2.132	1.596	1.512	1.386	0.876
$B \to D_2^* l\nu$	1.615	1.310	1.212	0.983	0.554	0.012	0.403	0.455	0.955	1.823	2.298	2.385	2.718	2.902	2.619	2.322	1.825	1.509	1.203	1.236	1.132
Gap	18.074	6.416	1.988	8.458	12.949	15.710	13.400	15.328	16.016	15.887	15.492	14.874	13.560	11.072	9.294	7.731	6.077	5.292	4.110	3.358	2.885
Total rel. error in ‰	23.100	12.166	8.191	10.580	14.020	16.499	14.817	16.142	16.677	16.806	16.674	16.299	15.287	13.537	11.864	10.535	8.996	7.496	6.036	4.776	4.063
$\langle q^8 \rangle$																					
$B \to Dl\nu$	0.954	0.821	0.680	0.552	0.421	0.290	0.123	0.011	0.092	0.143	0.207	0.264	0.293	0.321	0.325	0.317	0.288	0.286	0.225	0.235	0.201
$B \to D^* l\nu$	1.977	1.838	1.747	1.650	1.567	1.487	1.563	1.412	1.270	1.061	0.835	0.621	0.385	0.162	0.038	0.232	0.376	0.437	0.522	0.556	0.621
$B \to D_1 l\nu$	2.021	3.057	3.503	3.662	3.107	2.400	3.746	1.662	0.057	1.354	3.807	4.736	5.915	7.816	7.658	7.932	7.759	6.355	5.271	3.894	3.430
$B \to D_1' l\nu$	6.373	2.711	0.454	1.521	3.088	4.262	4.154	4.804	5.077	5.855	5.970	5.942	5.514	4.911	4.525	3.692	2.877	2.098	1.613	0.916	0.455
$B \to D_0^* l\nu$	7.209	4.065	1.784	0.458	1.611	2.681	2.737	3.301	3.683	4.343	4.501	4.703	4.533	4.110	3.648	3.374	2.914	2.209	2.071	1.879	1.207
$B \to D_2^* l\nu$	0.256	0.466	0.401	0.465	0.741	1.155	0.896	1.695	2.124	3.018	3.492	3.531	3.883	4.057	3.674	3.272	2.615	2.187	1.781	1.800	1.646
Gap	8.215	2.856	10.626	16.450	20.296	22.461	20.508	21.985	22.238	21.687	20.937	19.930	18.156	14.950	12.645	10.593	8.397	7.311	5.747	4.716	4.051
Total rel. error in ‰	13.002	6.757	11.500	17.022	20.902	23.222	21.509	22.912	23.239	23.142	22.841	22.135	20.758	18.498	16.308	14.526	12.432	10.401	8.438	6.736	5.742

TABLE E.3: Summary of systematic uncertainties due to the $B \to X_c \ell \nu$ modelling for the moments $\langle q^{2,4,6,8} \rangle$ for electron candidates. The values are given as the total relative error in permille.

q^2 selection in GeV^2	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$																					
$B \to Dl\nu$	0.380	0.229	0.140	0.092	0.075	0.064	0.094	0.080	0.068	0.057	0.045	0.037	0.032	0.028	0.025	0.029	0.030	0.033	0.038	0.046	0.061
$B \to D^* l\nu$	0.827	0.784	0.778	0.768	0.773	0.774	0.894	0.871	0.841	0.812	0.769	0.722	0.670	0.614	0.549	0.493	0.438	0.400	0.369	0.343	0.310
$B \to D_1 l\nu$	1.330	0.846	0.516	0.427	0.468	0.503	0.520	0.489	0.463	0.416	0.333	0.274	0.223	0.184	0.128	0.093	0.091	0.080	0.066	0.057	0.047
$B \to D_1' l\nu$	1.207	0.805	0.383	0.040	0.192	0.354	0.418	0.466	0.451	0.420	0.327	0.201	0.112	0.023	0.052	0.100	0.116	0.114	0.096	0.098	0.097
$B \to D_0^* l\nu$	1.101	0.782	0.503	0.252	0.069	0.064	0.117	0.177	0.203	0.210	0.199	0.163	0.124	0.081	0.016	0.024	0.057	0.082	0.072	0.076	0.081
$B \to D_2^* l\nu$	0.612	0.352	0.158	0.030	0.080	0.147	0.183	0.208	0.201	0.180	0.143	0.104	0.058	0.009	0.050	0.063	0.055	0.041	0.037	0.039	0.035
Gap	0.135	0.580	1.042	1.282	1.367	1.222	1.167	0.924	0.636	0.352	0.068	0.168	0.300	0.370	0.381	0.363	0.299	0.233	0.149	0.146	0.126
Total rel. error in ‰	2.379	1.761	1.550	1.578	1.655	1.582	1.632	1.466	1.271	1.101	0.936	0.839	0.788	0.746	0.685	0.632	0.558	0.493	0.424	0.402	0.367
$\langle q^4 \rangle$																					
$B \to Dl\nu$	0.508	0.359	0.260	0.199	0.173	0.152	0.201	0.172	0.147	0.124	0.100	0.082	0.074	0.065	0.059	0.066	0.069	0.074	0.084	0.099	0.131
$B \to D^* l\nu$	1.871	1.884	1.896	1.868	1.866	1.853	2.040	1.974	1.894	1.814	1.711	1.604	1.491	1.371	1.237	1.122	1.009	0.927	0.860	0.801	0.727
$B \to D_1 l\nu$	1.371	0.932	0.695	0.722	0.834	0.903	0.949	0.896	0.846	0.757	0.609	0.502	0.413	0.344	0.247	0.200	0.201	0.176	0.146	0.126	0.103
$B \to D_1' l\nu$	1.248	0.822	0.318	0.120	0.444	0.680	0.787	0.854	0.815	0.744	0.565	0.328	0.157	0.016	0.150	0.238	0.266	0.258	0.218	0.219	0.214
$B \to D_0^* l\nu$	1.306	0.971	0.639	0.316	0.062	0.145	0.235	0.334	0.374	0.380	0.352	0.280	0.204	0.119	0.006	0.082	0.145	0.191	0.168	0.175	0.182
$B \to D_2^* l\nu$	0.630	0.355	0.126	0.046	0.190	0.288	0.349	0.384	0.365	0.322	0.249	0.174	0.086	0.015	0.115	0.139	0.121	0.092	0.084	0.086	0.075
Gap	0.823	1.344	1.879	2.160	2.245	1.997	1.977	1.532	1.008	0.488	0.033	0.458	0.694	0.815	0.822	0.774	0.638	0.499	0.329	0.318	0.274
Total rel. error in ‰	3.158	2.846	2.864	2.972	3.080	2.971	3.132	2.840	2.506	2.218	1.953	1.805	1.719	1.637	1.518	1.409	1.255	1.121	0.979	0.925	0.846
$\langle q^6 \rangle$																					
$B \to Dl\nu$	0.544	0.415	0.332	0.284	0.266	0.247	0.300	0.265	0.233	0.201	0.168	0.142	0.129	0.116	0.106	0.116	0.120	0.127	0.141	0.163	0.211
$B \to D^* l\nu$	3.076	3.141	3.193	3.169	3.190	3.185	3.343	3.258	3.148	3.022	2.862	2.693	2.514	2.323	2.112	1.928	1.746	1.609	1.492	1.387	1.260
$B \to D_1 l\nu$	1.175	0.821	0.758	0.905	1.073	1.169	1.242	1.180	1.117	1.002	0.811	0.674	0.565	0.479	0.359	0.319	0.327	0.287	0.239	0.206	0.170
$B \to D_1' l\nu$	1.068	0.648	0.141	0.300	0.644	0.903	1.034	1.105	1.045	0.939	0.694	0.369	0.132	0.103	0.294	0.415	0.448	0.429	0.365	0.362	0.348
$B \to D_0^* l\nu$	1.257	0.934	0.610	0.291	0.048	0.206	0.320	0.434	0.480	0.481	0.438	0.336	0.228	0.107	0.068	0.178	0.266	0.327	0.292	0.298	0.303
$B \to D_2^* l\nu$	0.544	0.272	0.047	0.124	0.281	0.390	0.464	0.502	0.474	0.413	0.312	0.207	0.087	0.051	0.195	0.225	0.196	0.152	0.138	0.139	0.121
Gap	1.070	1.607	2.167	2.476	2.572	2.290	2.314	1.756	1.084	0.405	0.282	0.847	1.159	1.311	1.303	1.216	1.006	0.791	0.536	0.512	0.439
Total rel. error in ‰	3.912	3.829	3.996	4.155	4.302	4.222	4.422	4.101	3.733	3.410	3.119	2.956	2.841	2.717	2.536	2.359	2.119	1.905	1.682	1.579	1.443
$\langle q^8 \rangle$																					
$B \to Dl\nu$	0.539	0.433	0.375	0.349	0.348	0.341	0.388	0.358	0.326	0.292	0.252	0.220	0.203	0.185	0.171	0.183	0.187	0.195	0.212	0.241	0.304
$B \to D^* l\nu$	4.392	4.512	4.614	4.599	4.675	4.706	4.740	4.683	4.587	4.440	4.239	4.015	3.773	3.507	3.210	2.941	2.672	2.461	2.276	2.107	1.911
$B \to D_1 l\nu$	0.993	0.704	0.776	0.990	1.191	1.305	1.389	1.328	1.269	1.145	0.936	0.789	0.680	0.591	0.467	0.449	0.466	0.410	0.343	0.297	0.245
$B \to D_1' l\nu$	0.948	0.536	0.036	0.387	0.730	0.992	1.120	1.192	1.123	0.997	0.709	0.324	0.044	0.246	0.478	0.623	0.657	0.624	0.535	0.524	0.497
$B \to D_0^* l\nu$	1.185	0.875	0.570	0.276	0.053	0.215	0.333	0.451	0.500	0.497	0.445	0.323	0.192	0.044	0.172	0.309	0.417	0.490	0.440	0.442	0.443
$B \to D_2^* l\nu$	0.482	0.213	0.021	0.167	0.326	0.438	0.512	0.552	0.521	0.449	0.330	0.206	0.067	0.108	0.284	0.318	0.278	0.220	0.198	0.197	0.170
Gap	0.828	1.380	1.966	2.322	2.451	2.185	2.199	1.613	0.884	0.132	0.651	1.303	1.660	1.829	1.800	1.671	1.387	1.097	0.761	0.721	0.619
Total rel. error in ‰	4.878	4.904	5.121	5.282	5.482	5.474	5.568	5.325	5.032	4.751	4.488	4.329	4.188	4.012	3.759	3.502	3.162	2.854	2.538	2.369	2.158

TABLE E.4: Summary of systematic uncertainties due to the $B \to X_c \ell \nu$ modelling for the moments $\langle q^{2,4,6,8} \rangle$ for muon candidates. The values are given as the total relative error in permille.

q^2 selection in GeV^2	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$																					
$B \to Dl\nu$	0.324	0.206	0.130	0.082	0.061	0.050	0.079	0.067	0.059	0.051	0.044	0.040	0.036	0.036	0.040	0.045	0.056	0.065	0.080	0.089	0.104
$B \to D^* l\nu$	0.817	0.792	0.781	0.763	0.757	0.749	0.847	0.822	0.795	0.764	0.730	0.687	0.632	0.573	0.516	0.463	0.416	0.382	0.346	0.322	0.293
$B \to D_1 l\nu$	1.044	0.646	0.401	0.374	0.430	0.468	0.474	0.463	0.428	0.391	0.314	0.266	0.227	0.194	0.136	0.091	0.102	0.093	0.087	0.070	0.063
$B \to D_1' l\nu$	0.836	0.481	0.141	0.161	0.380	0.519	0.499	0.523	0.525	0.460	0.377	0.260	0.139	0.040	0.055	0.101	0.122	0.105	0.101	0.076	0.076
$B \to D_0^* l\nu$	1.055	0.756	0.485	0.252	0.105	0.031	0.109	0.167	0.195	0.192	0.175	0.150	0.109	0.058	0.007	0.035	0.065	0.074	0.087	0.088	0.079
$B \to D_2^* l\nu$	0.516	0.298	0.132	0.020	0.097	0.152	0.187	0.217	0.219	0.198	0.162	0.119	0.058	0.009	0.037	0.053	0.042	0.037	0.029	0.029	0.026
Gap	0.199	0.408	0.792	1.037	1.072	0.953	0.900	0.668	0.400	0.177	0.047	0.261	0.430	0.424	0.426	0.384	0.290	0.233	0.169	0.143	0.112
Total rel. error in ‰	1.995	1.465	1.299	1.376	1.441	1.409	1.433	1.300	1.158	1.029	0.914	0.846	0.820	0.743	0.687	0.622	0.540	0.481	0.425	0.388	0.355
$\langle q^4 \rangle$																					
$B \to Dl\nu$	0.449	0.337	0.257	0.196	0.163	0.143	0.183	0.160	0.141	0.125	0.111	0.103	0.096	0.095	0.103	0.112	0.133	0.151	0.179	0.195	0.223
$B \to D^* l\nu$	1.838	1.844	1.841	1.814	1.793	1.764	1.923	1.855	1.782	1.701	1.615	1.514	1.394	1.269	1.151	1.040	0.942	0.868	0.788	0.732	0.667
$B \to D_1 l\nu$	0.981	0.652	0.530	0.629	0.747	0.817	0.841	0.823	0.762	0.695	0.564	0.480	0.415	0.360	0.257	0.197	0.228	0.208	0.193	0.159	0.144
$B \to D_1' l\nu$	0.909	0.551	0.164	0.217	0.514	0.717	0.921	0.948	0.937	0.813	0.656	0.438	0.213	0.028	0.147	0.232	0.266	0.229	0.217	0.164	0.161
$B \to D_0^* l\nu$	1.367	1.067	0.760	0.464	0.257	0.087	0.207	0.300	0.344	0.335	0.302	0.252	0.173	0.078	0.022	0.096	0.151	0.164	0.185	0.184	0.163
$B \to D_2^* l\nu$	0.481	0.260	0.075	0.105	0.221	0.298	0.355	0.398	0.393	0.351	0.281	0.200	0.087	0.033	0.089	0.117	0.096	0.084	0.068	0.067	0.059
Gap	0.309	0.898	1.327	1.615	1.642	1.452	1.410	1.014	0.549	0.150	0.244	0.630	0.927	0.906	0.896	0.803	0.610	0.491	0.359	0.301	0.237
Total rel. error in ‰	2.750	2.502	2.471	2.570	2.623	2.553	2.729	2.514	2.287	2.077	1.897	1.796	1.751	1.606	1.495	1.362	1.197	1.071	0.952	0.869	0.792
$\langle q^6 \rangle$																					
$B \to Dl\nu$	0.522	0.428	0.362	0.310	0.280	0.260	0.300	0.272	0.248	0.226	0.206	0.194	0.184	0.183	0.192	0.205	0.233	0.258	0.296	0.318	0.358
$B \to D^* l\nu$	2.970	3.003	3.025	3.015	3.004	2.974	3.123	3.034	2.930	2.805	2.670	2.508	2.319	2.120	1.933	1.755	1.593	1.466	1.331	1.228	1.113
$B \to D_1 l\nu$	0.766	0.535	0.582	0.780	0.936	1.027	1.064	1.048	0.974	0.893	0.733	0.631	0.557	0.499	0.371	0.325	0.379	0.347	0.321	0.268	0.243
$B \to D_1' l\nu$	1.068	0.724	0.341	0.059	0.370	0.601	1.203	1.226	1.202	1.035	0.823	0.528	0.222	0.037	0.273	0.387	0.429	0.370	0.346	0.264	0.254
$B \to D_0^* l\nu$	1.579	1.297	0.999	0.701	0.482	0.285	0.277	0.383	0.435	0.421	0.375	0.305	0.195	0.064	0.073	0.176	0.251	0.265	0.291	0.284	0.248
$B \to D_2^* l\nu$	0.368	0.157	0.033	0.199	0.320	0.403	0.470	0.515	0.504	0.446	0.351	0.239	0.086	0.077	0.156	0.193	0.161	0.141	0.116	0.114	0.101
Gap	0.384	0.969	1.400	1.701	1.736	1.532	1.486	1.013	0.442	0.075	0.574	1.076	1.462	1.420	1.389	1.240	0.951	0.767	0.568	0.475	0.374
Total rel. error in ‰	3.687	3.557	3.563	3.636	3.669	3.594	3.864	3.649	3.418	3.189	2.996	2.883	2.820	2.608	2.438	2.232	1.978	1.775	1.582	1.439	1.305
$\langle q^8 \rangle$																					
$B \to Dl\nu$	0.586	0.508	0.460	0.421	0.404	0.392	0.424	0.401	0.378	0.355	0.331	0.317	0.303	0.300	0.310	0.324	0.357	0.386	0.432	0.459	0.509
$B \to D^* l\nu$	4.144	4.205	4.260	4.281	4.299	4.294	4.369	4.297	4.195	4.052	3.885	3.671	3.416	3.137	2.874	2.616	2.372	2.177	1.970	1.804	1.625
$B \to D_1 l\nu$	0.600	0.440	0.593	0.833	1.006	1.109	1.139	1.130	1.058	0.979	0.815	0.717	0.658	0.615	0.488	0.476	0.556	0.510	0.471	0.399	0.362
$B \to D_1' l\nu$	1.427	1.092	0.714	0.331	0.043	0.255	1.324	1.343	1.315	1.126	0.882	0.535	0.172	0.140	0.426	0.563	0.609	0.527	0.488	0.377	0.357
$B \to D_0^* l\nu$	1.822	1.554	1.271	0.982	0.765	0.561	0.306	0.413	0.466	0.448	0.397	0.314	0.183	0.027	0.142	0.268	0.358	0.372	0.398	0.383	0.332
$B \to D_2^* l\nu$	0.289	0.088	0.087	0.252	0.370	0.453	0.513	0.560	0.548	0.481	0.371	0.240	0.063	0.139	0.235	0.278	0.236	0.208	0.174	0.168	0.149
Gap	0.128	0.680	1.112	1.428	1.484	1.303	1.187	0.711	0.120	0.436	0.994	1.561	2.001	1.936	1.882	1.680	1.301	1.054	0.793	0.663	0.526
Total rel. error in ‰	4.831	4.713	4.699	4.730	4.752	4.701	4.908	4.764	4.596	4.403	4.235	4.120	4.033	3.755	3.520	3.234	2.883	2.592	2.312	2.093	1.889

Appendix F

Correlations between measured q^n moments

F.1 Statistical correlations for electron candidates

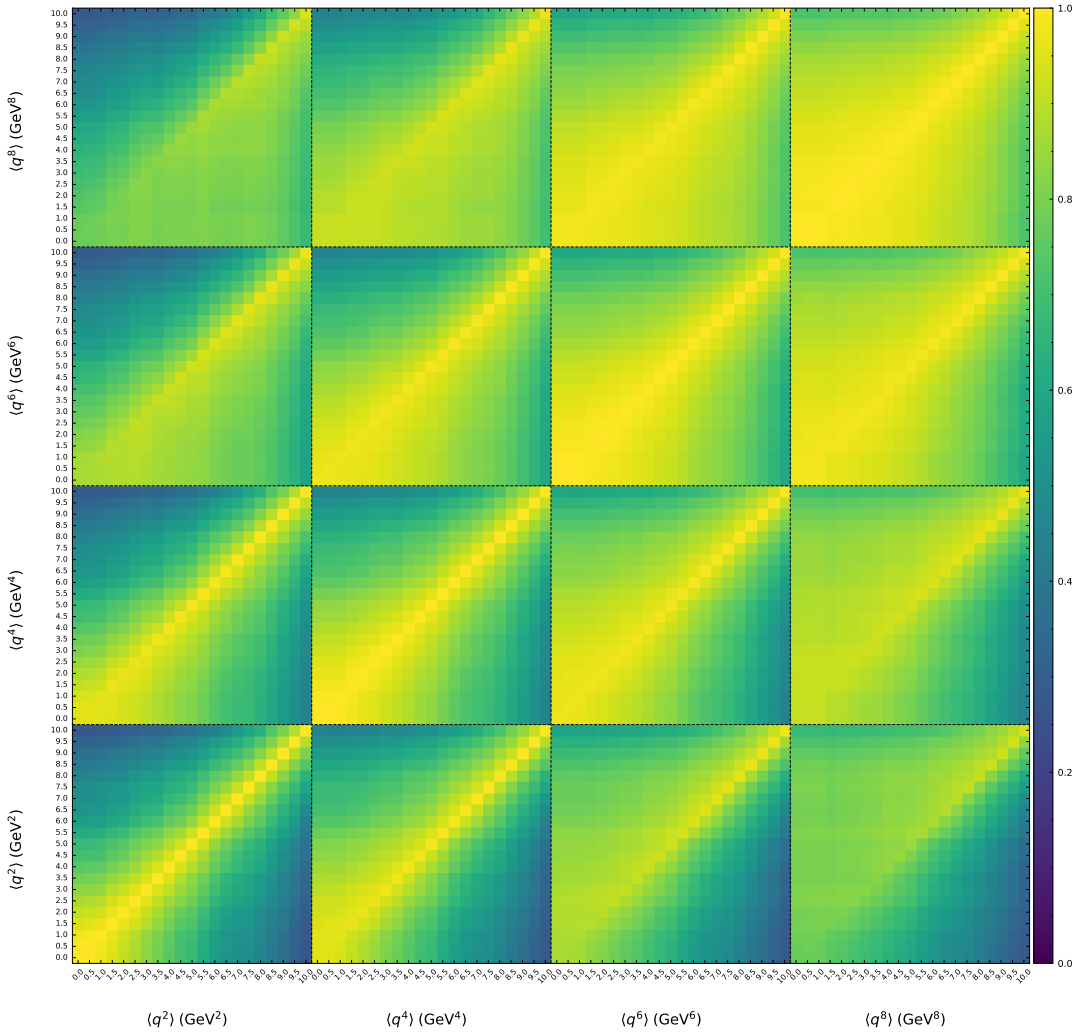

FIGURE F.1: Statistical correlations between measured $\langle q^n \rangle$ moments for the electron channel.

F.2 Experimental correlations for electron candidates

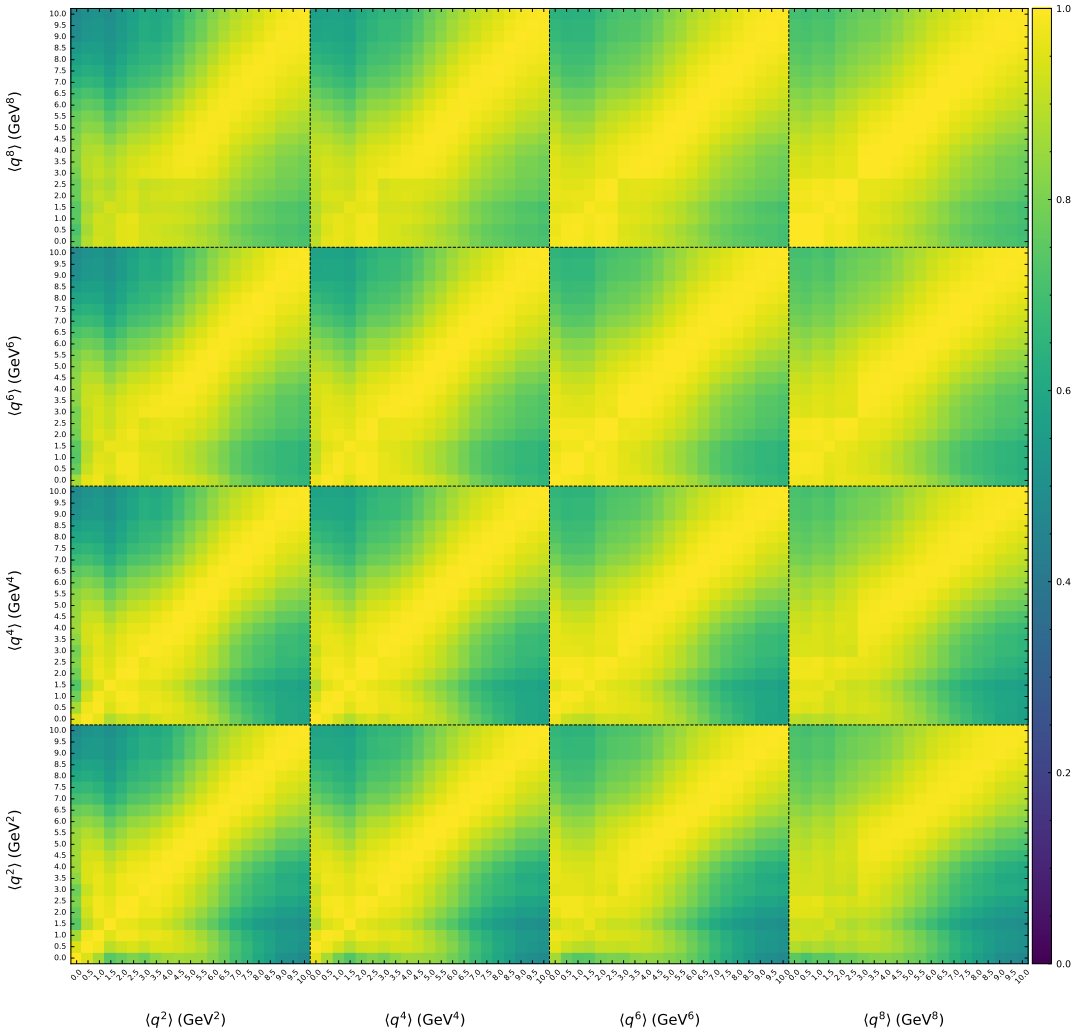

FIGURE F.2: Full experimental correlations between measured $\langle q^n \rangle$ moments for the electron channel.

F.3 Experimental correlations for normalized electron candidates

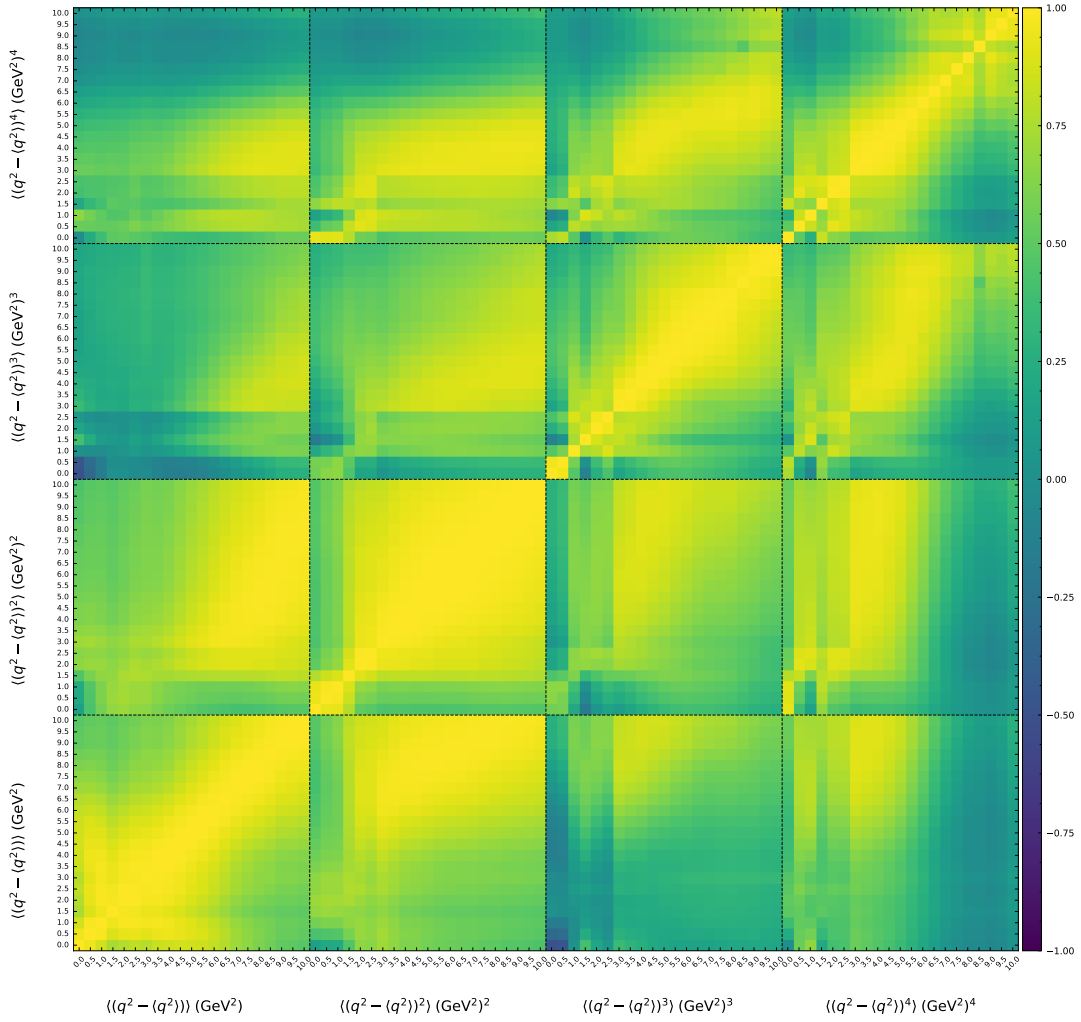

FIGURE F.3: Full experimental correlations between measured $\langle(q^2 - \langle q^2\rangle)^n\rangle$ central moments for the electron channel.

F.4 Statistical correlations for muon candidates

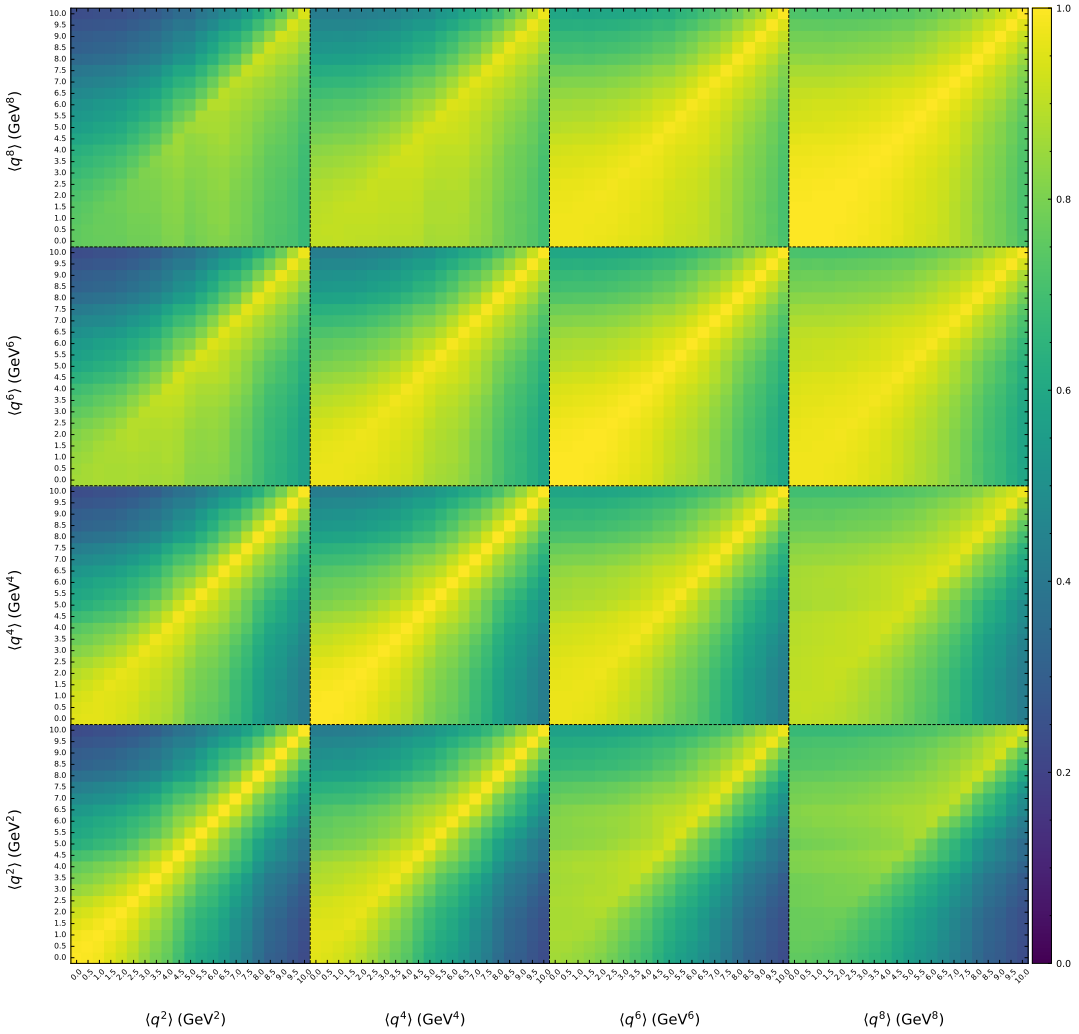

FIGURE F.4: Statistical correlations between measured $\langle q^n \rangle$ moments for the muon channel.

F.5 Experimental correlations for muon candidates

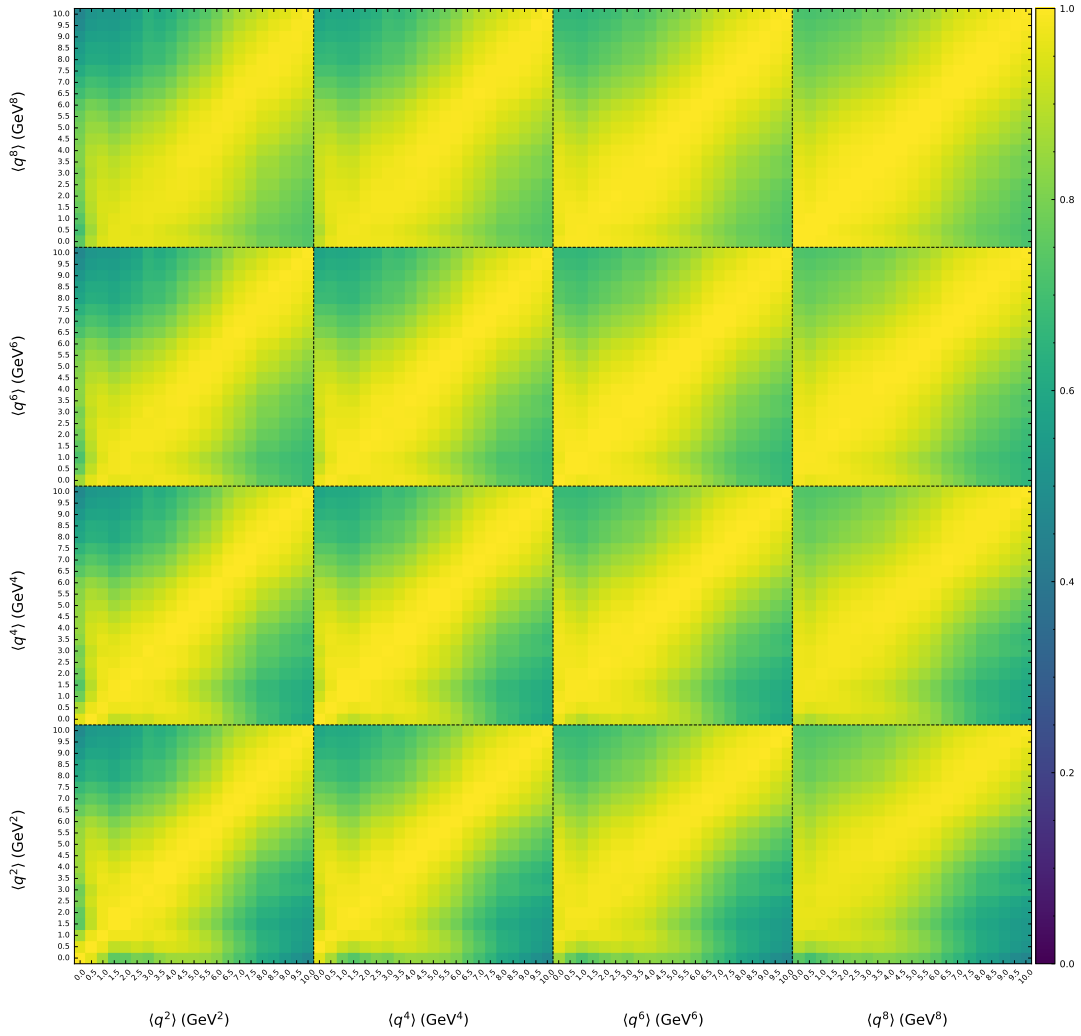

FIGURE F.5: Full experimental correlations between measured $\langle q^n \rangle$ moments for the muon channel.

F.6 Experimental correlations for normalized muon candidates

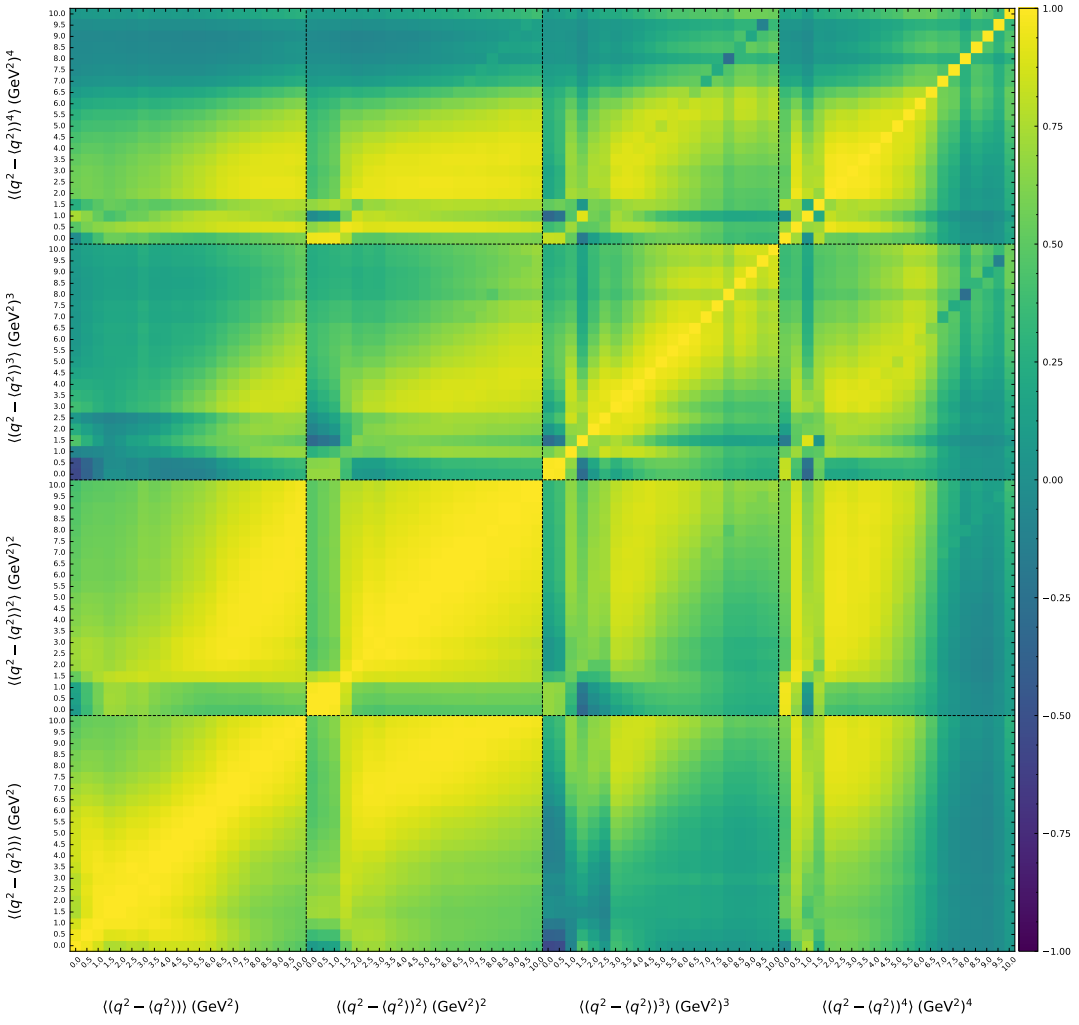

FIGURE F.6: Full experimental correlations between measured $\langle(q^2 - \langle q^2\rangle)^n\rangle$ central moments for the muon channel.

Appendix G

Comparison with simulated and measured moments

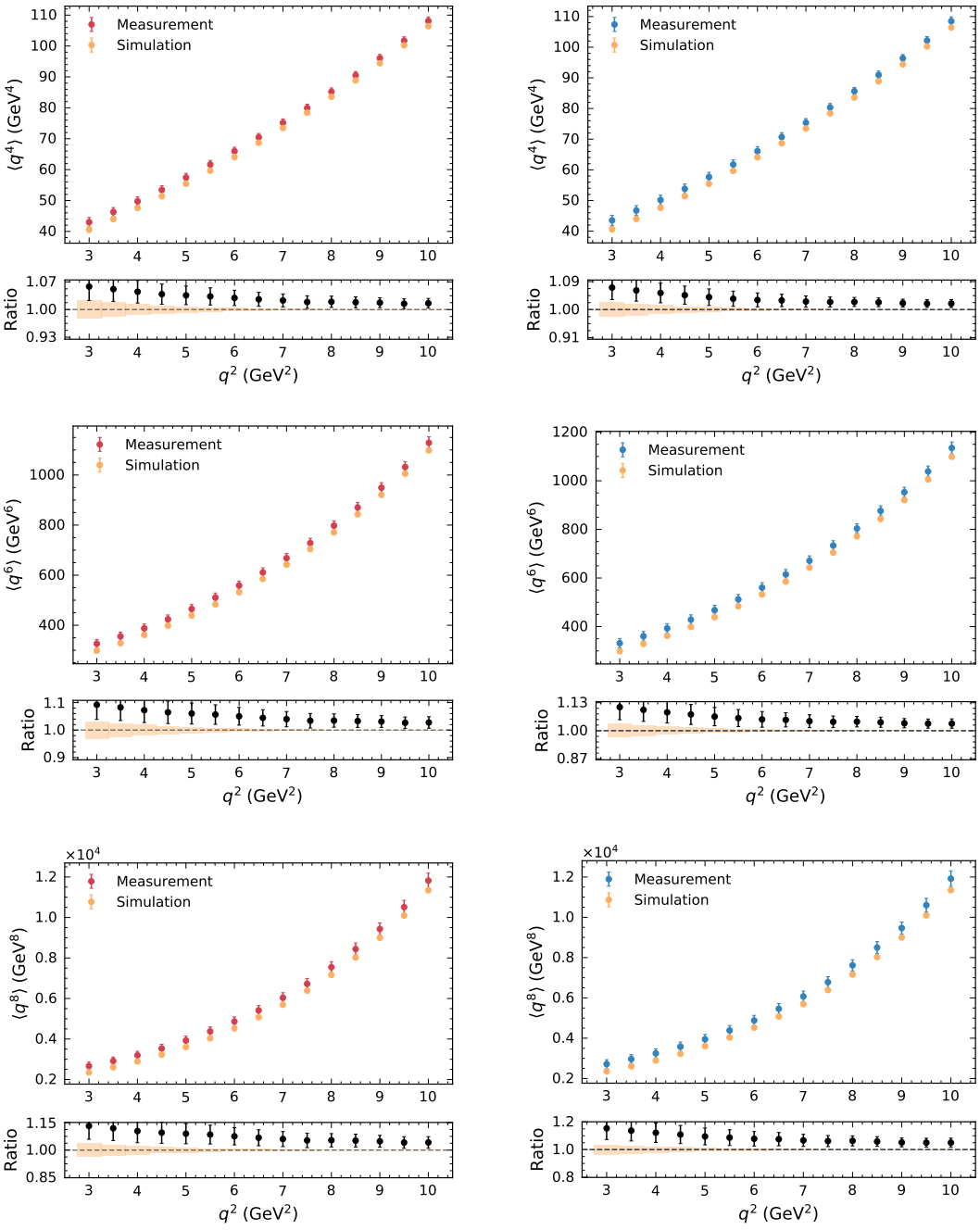

FIGURE G.1: Comparison of the simulated and measured higher order moments for electron (left) and muon (right) candidates.

Appendix H

Fitted q^2 moments for muon candidates

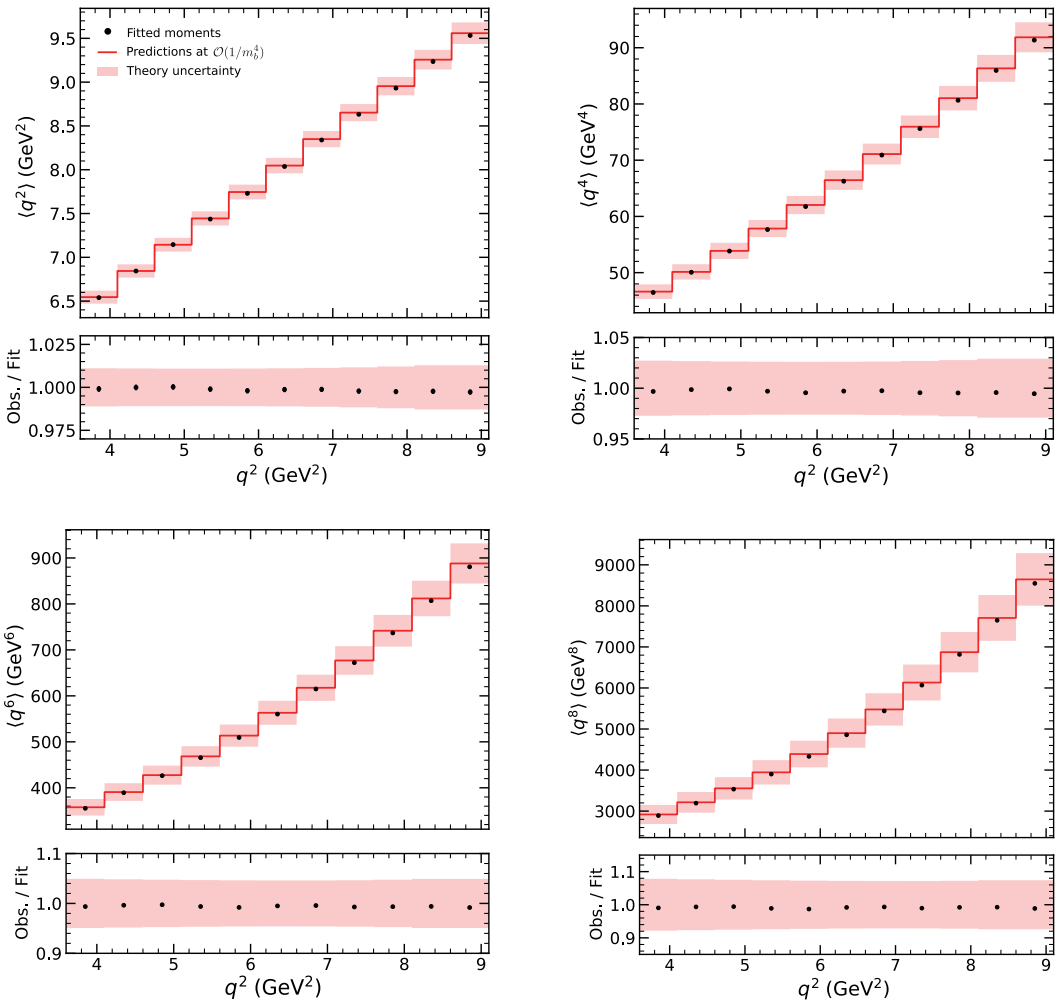

FIGURE H.1: The fitted $\langle q^{2,4,6,8} \rangle$ moments for muon candidates and the theoretical predictions at $\mathcal{O}(1/m_b^4)$. While both experimental and theoretical uncertainties are shown in the ratio, the results are dominated by theoretical systematic uncertainties. The fit converges with a final χ^2 of 80.5 for 83 degrees of freedom.

Appendix I

Studies for low q^2 moments

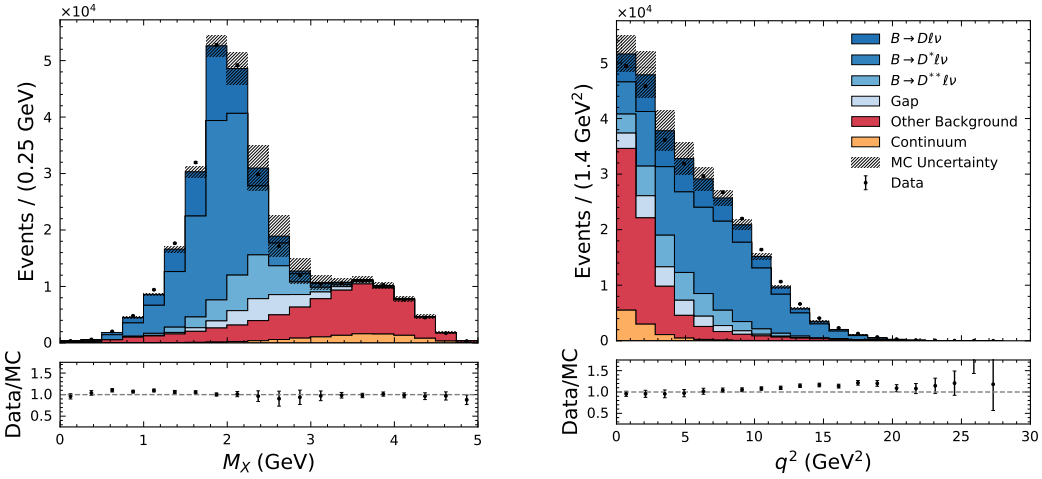

FIGURE I.1: Comparison of the reconstructed M_X (left) and q^2 (right) distributions with simulated events. The error band of the simulated samples incorporate the full set of systematic uncertainties discussed in Section 3.3.2.

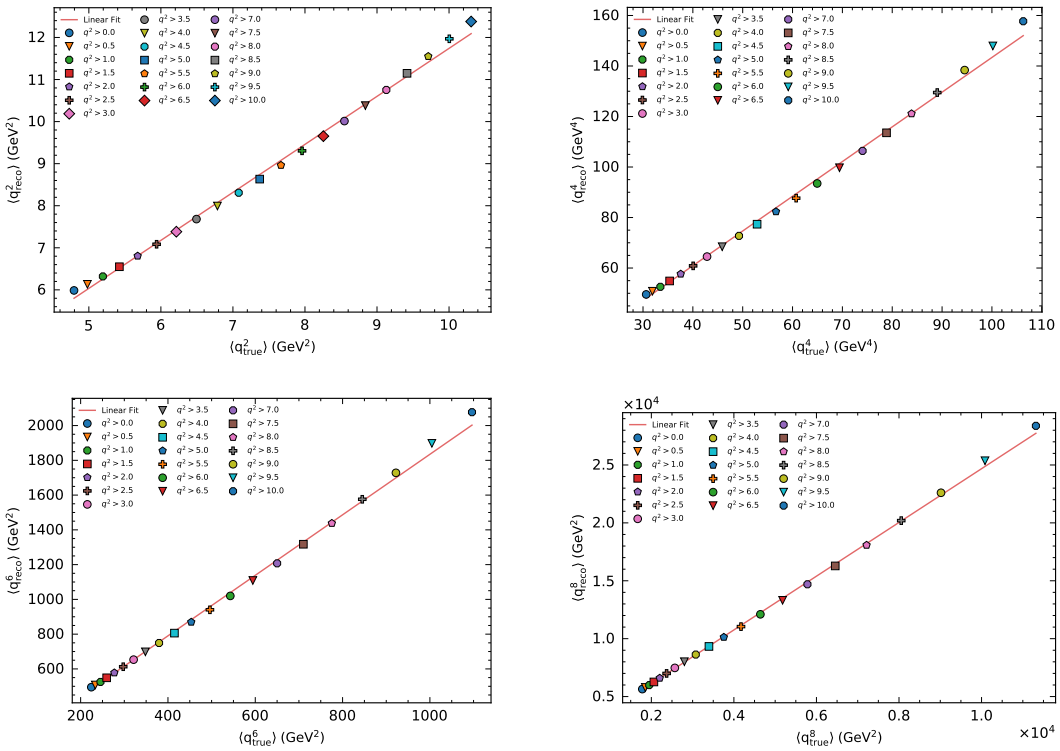

FIGURE I.2: Calibration curves for $\langle q^{2,4,6,8} \rangle$ for electron candidates. Shown here are the the extracted $\langle q^n_{,\text{reco}} \rangle$ versus $\langle q^n_{,\text{true}} \rangle$ as a function of lower q^2 cut, together with the fitted linear calibration curves.

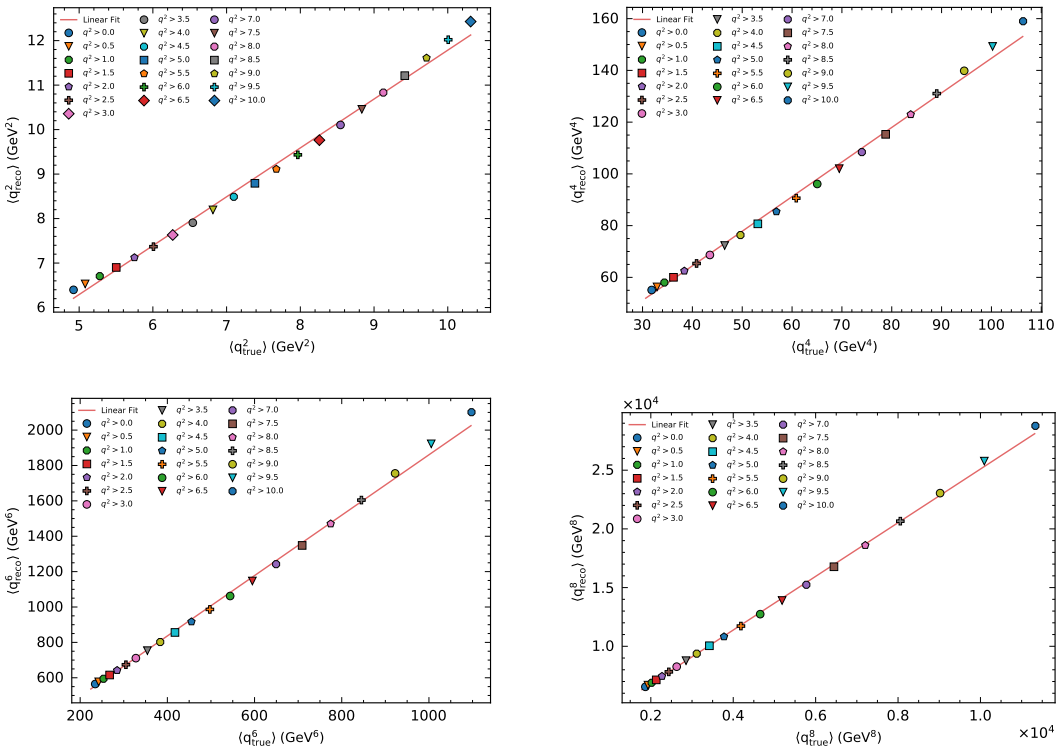

FIGURE I.3: Calibration curves for $\langle q^{2,4,6,8} \rangle$ for muon candidates. Shown here are the the extracted $\langle q^n_{,\text{reco}} \rangle$ versus $\langle q^n_{,\text{true}} \rangle$ as a function of lower q^2 cut, together with the fitted linear calibration curves.

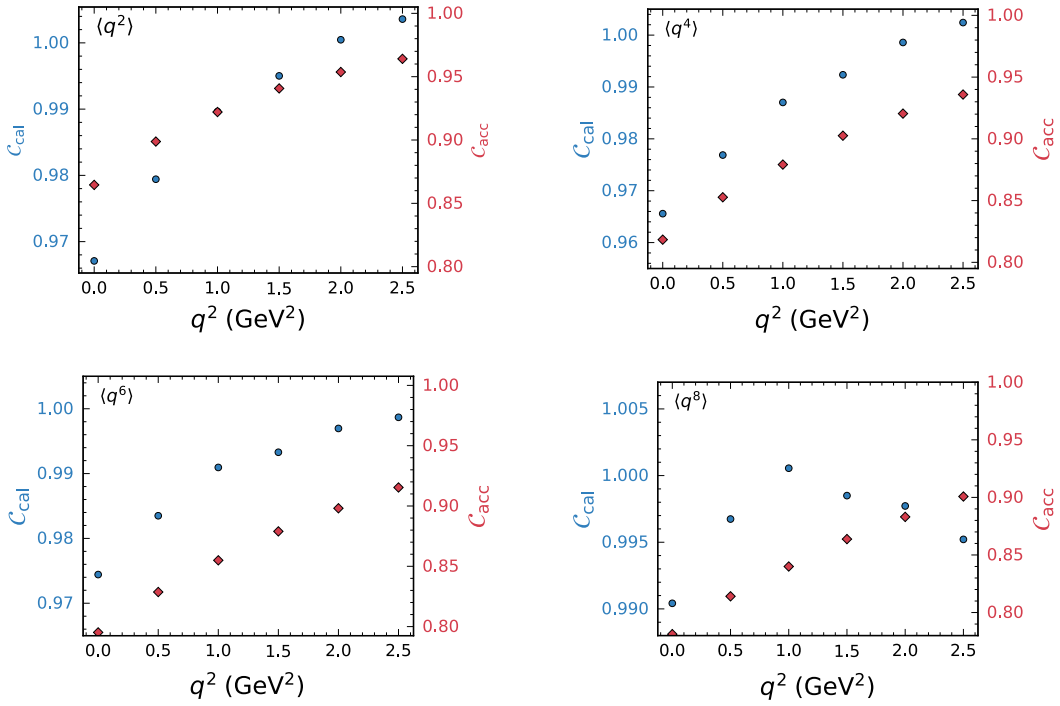

FIGURE I.4: Remaining bias after application of the calibration curves on simulated $B \to X_c e\nu$ decays as a function of q^2. Shown here are the bias correction factors $\mathcal{C}_{\text{cal}}$ (blue) and $\mathcal{C}_{\text{acc}}$ (red).

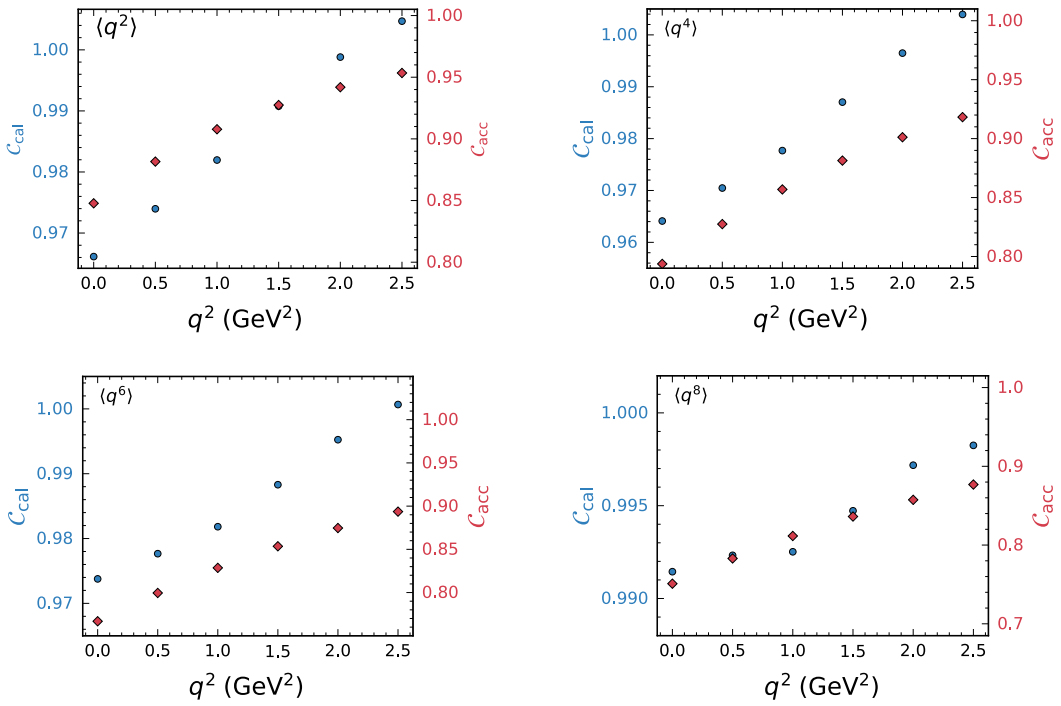

FIGURE I.5: Remaining bias after application of the calibration curves on simulated $B \to X_c \mu \nu$ decays as a function of q^2. Shown here are the bias correction factors $\mathcal{C}_{\text{cal}}$ (blue) and $\mathcal{C}_{\text{acc}}$ (red).

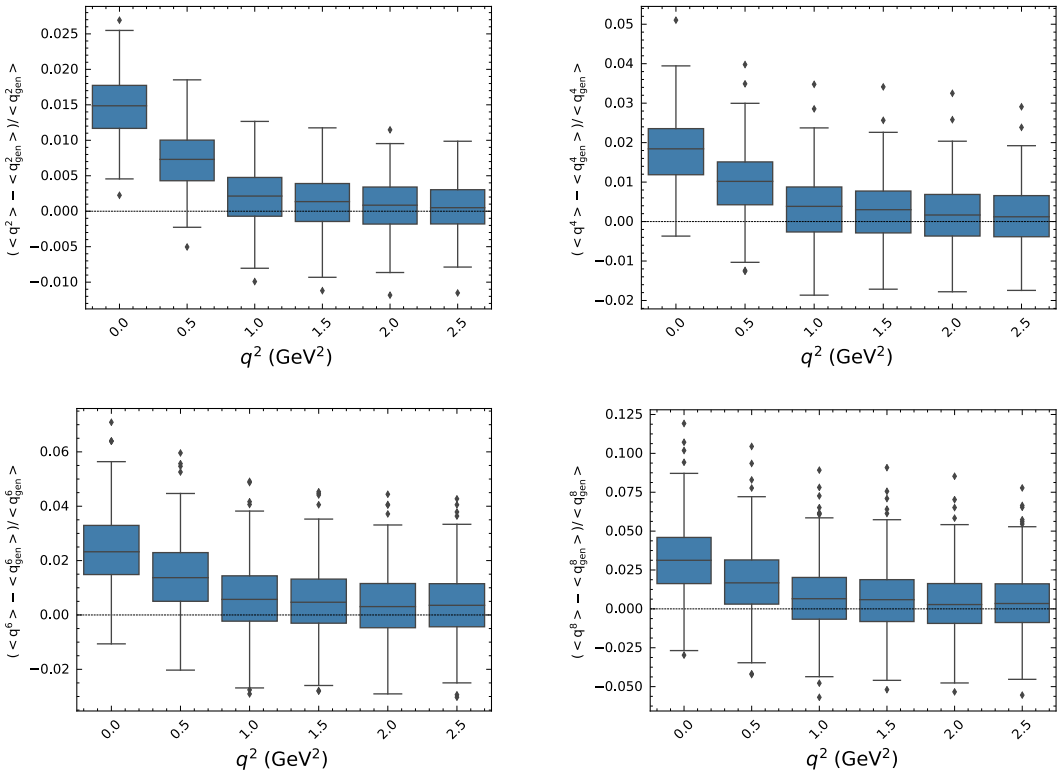

FIGURE I.6: Box plot of the sample test relative error distributions for q^n moments as a function of lower level q^2 selections in the electron channel.

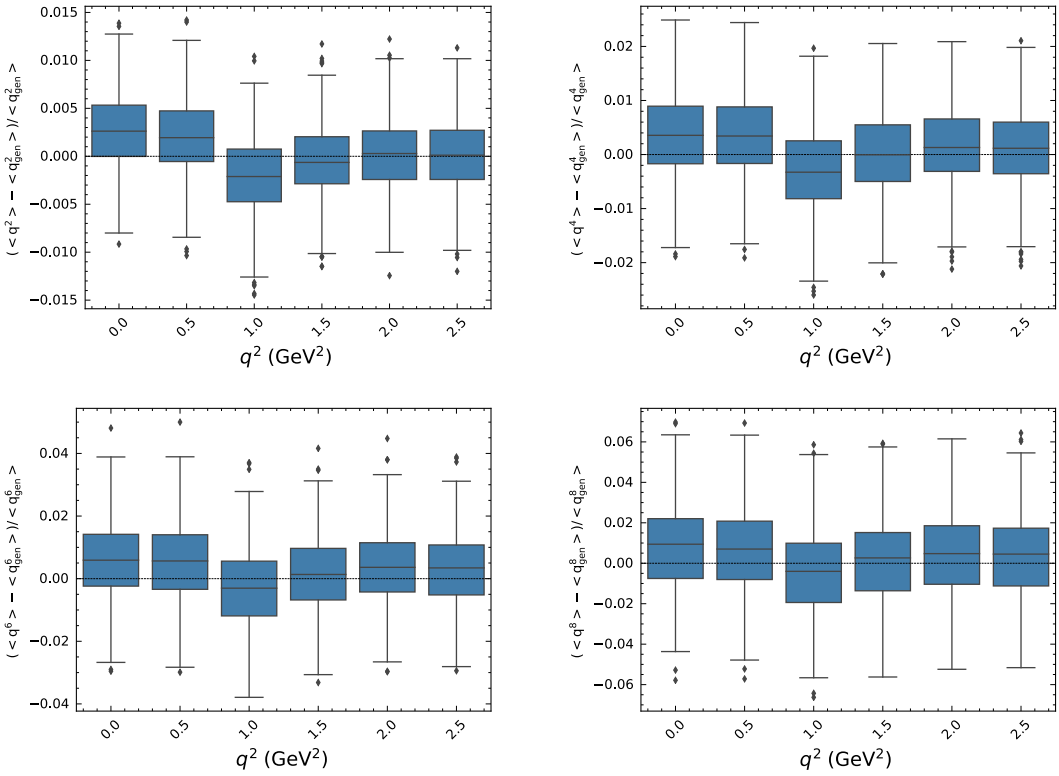

FIGURE I.7: Box plot of the sample test relative error distributions for q^n moments as a function of lower level q^2 selections in the muon channel.

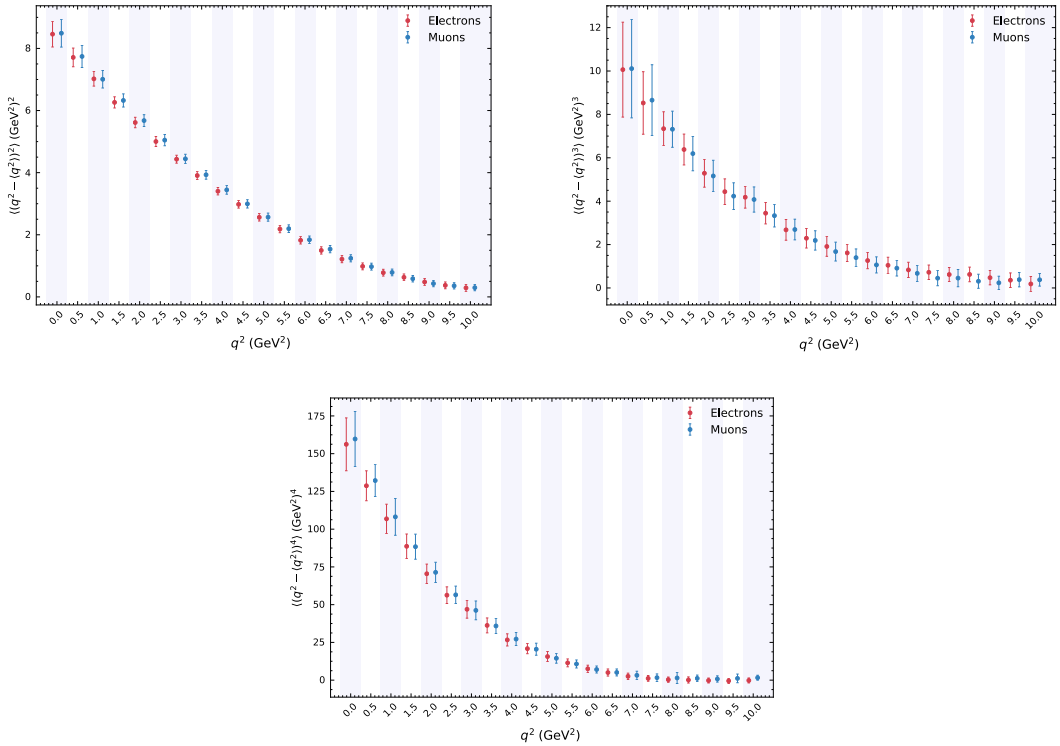

FIGURE I.8: The extracted central moments for both electron and muon candidates calculated from the measured nominal moments using Eqn. 6.11.

TABLE I.1: Summary of statistical and systematic uncertainties for the moments $\langle q^{2,4,6,8} \rangle$ for electron candidates. The values are given as the relative error in permille.

q^2 selection in GeV^2	0.0	0.5	1.0	1.5	2.0	2.5
$\langle q^2 \rangle$ in GeV^2 4.37	4.67	4.97	5.35	5.64	5.95	
Stat. error (data)	2.18	2.02	1.89	1.78	1.68	1.60
Bkg. subtraction	3.23	2.86	2.48	2.50	1.97	1.48
$B \to X_u \ell\nu$ BF	1.11	1.71	1.79	9.03	2.40	2.15
$B \to X_c \ell\nu$ BF	27.34	14.97	9.01	5.82	5.07	5.03
Non-resonant model	24.99	25.38	23.92	21.75	19.40	16.97
$B \to X_c \ell\nu$ FF	2.38	1.76	1.55	1.58	1.66	1.58
N_{tracks} res.	8.06	7.66	7.21	6.68	6.23	5.85
N_γ res.	1.12	1.27	1.33	1.12	1.36	1.35
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	1.89	2.47	2.82	2.93	3.19	3.19
q^2 scale	9.94	9.73	9.49	8.90	8.38	7.90
MC non-closure	3.39	1.54	0.40	0.26	0.16	0.08
Cal. function	0.18	0.16	0.13	0.07	0.03	0.01
Stat. bias corr.	1.75	1.70	1.64	1.57	1.50	1.44
PID eff.	0.34	0.32	0.29	0.26	0.24	0.21
Track eff.	0.73	0.69	0.65	0.59	0.55	0.51
$B^0/B^\pm$ tag eff.	0.04	0.03	0.08	0.24	0.18	0.21
Sys. error (total)	39.66	32.40	28.64	27.09	23.20	20.82
Total error	39.72	32.46	28.71	27.15	23.26	20.88
$\langle q^4 \rangle$ in GeV^4	27.53	29.55	31.68	34.83	37.45	40.37
Stat. error (data)	4.20	4.02	3.90	3.74	3.64	3.53
Bkg. subtraction	4.20	3.93	3.59	3.75	3.09	2.41
$B \to X_u \ell\nu$ BF	3.20	4.16	4.24	14.84	4.97	4.59
$B \to X_c \ell\nu$ BF	28.91	17.02	11.38	9.06	9.46	10.15
Non-resonant model	44.05	44.26	42.32	39.18	35.67	31.84
$B \to X_c \ell\nu$ FF	3.16	2.85	2.86	2.97	3.08	2.97
N_{tracks} res.	16.16	15.71	15.14	14.31	13.60	12.93
N_γ res.	3.71	3.88	3.95	3.65	3.95	3.90
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	7.37	8.01	8.40	8.46	8.75	8.66
q^2 scale	20.93	20.72	20.42	19.44	18.63	17.78
MC non-closure	0.65	0.33	0.11	0.08	0.04	0.03
Cal. function	0.28	0.25	0.20	0.09	0.01	0.07
Stat. bias corr.	3.48	3.44	3.37	3.29	3.21	3.12
PID eff.	0.62	0.59	0.56	0.52	0.48	0.44
Track eff.	1.41	1.37	1.31	1.22	1.15	1.09
$B^0/B^\pm$ tag eff.	0.55	0.57	0.71	0.91	0.83	0.86
Sys. error (total)	59.97	55.30	52.01	50.41	45.19	41.69
Total error	60.12	55.45	52.16	50.55	45.34	41.84
$\langle q^6 \rangle$ in GeV^6	204.19	218.73	234.35	259.51	279.94	304.07
Stat. error (data)	6.92	6.72	6.61	6.40	6.31	6.19
Bkg. subtraction	4.91	4.64	4.33	4.69	3.98	3.24
$B \to X_u \ell\nu$ BF	6.20	7.56	7.61	20.82	8.17	7.66
$B \to X_c \ell\nu$ BF	22.67	13.14	11.19	12.06	14.12	15.54
Non-resonant model	59.30	59.25	56.89	52.98	48.77	44.08
$B \to X_c \ell\nu$ FF	3.91	3.83	4.00	4.15	4.30	4.22
N_{tracks} res.	25.09	24.61	23.95	22.83	21.95	21.05
N_γ res.	8.21	8.39	8.45	8.04	8.35	8.24
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	16.12	16.79	17.16	17.08	17.29	17.06
q^2 scale	33.31	33.12	32.77	31.38	30.39	29.27
MC non-closure	0.12	0.06	0.03	0.02	0.01	0.01
Cal. function	0.37	0.33	0.25	0.08	0.04	0.17
Stat. bias corr.	5.79	5.73	5.65	5.55	5.44	5.32
PID eff.	0.88	0.86	0.82	0.76	0.72	0.67
Track eff.	2.12	2.07	2.00	1.89	1.81	1.72
$B^0/B^\pm$ tag eff.	1.60	1.62	1.78	1.97	1.87	1.90
Sys. error (total)	78.84	76.63	74.21	72.98	67.01	62.97
Total error	79.14	76.92	74.50	73.26	67.30	63.28
$\langle q^8 \rangle$ in GeV^8	1663.85	1775.83	1899.15	2114.84	2275.67	2474.78
Stat. error (data)	10.92	10.69	10.59	10.26	10.22	10.09
Bkg. subtraction	5.85	5.59	5.28	6.02	5.34	4.60
$B \to X_u \ell\nu$ BF	10.44	12.32	12.30	28.94	12.40	11.69
$B \to X_c \ell\nu$ BF	13.49	10.57	14.35	17.24	19.92	21.45
Non-resonant model	72.40	72.07	69.20	64.37	59.53	54.12
$B \to X_c \ell\nu$ FF	4.88	4.90	5.12	5.28	5.48	5.47
N_{tracks} res.	35.11	34.60	33.81	32.26	31.23	30.09
N_γ res.	15.27	15.46	15.48	14.85	15.12	14.90
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	28.73	29.40	29.70	29.30	29.37	28.91
q^2 scale	47.29	47.15	46.71	44.69	43.54	42.12
MC non-closure	0.02	0.01	0.00	0.00	0.00	0.00
Cal. function	0.49	0.43	0.32	0.04	0.12	0.32
Stat. bias corr.	9.12	9.02	8.88	8.72	8.54	8.35
PID eff.	1.14	1.12	1.08	1.00	0.95	0.90
Track eff.	2.87	2.82	2.74	2.60	2.50	2.39
$B^0/B^\pm$ tag eff.	2.92	2.93	3.10	3.28	3.16	3.16
Sys. error (total)	101.09	100.67	98.71	97.75	90.58	86.03
Total error	101.68	101.24	99.28	98.28	91.16	86.62

TABLE I.2: Summary of statistical and systematic uncertainties for the moments $\langle q^{2,4,6,8} \rangle$ for muon candidates. The values are given as the relative error in permille.

q^2 selection in GeV^2	0.0	0.5	1.0	1.5	2.0	2.5
$\langle q^2 \rangle$ in GeV^2	4.45	4.76	5.07	5.37	5.68	5.98
Stat. error (data)	2.38	2.21	2.06	1.93	1.82	1.72
Bkg. subtraction	3.81	3.59	3.14	2.67	2.24	1.84
$B \to X_u \ell\nu$ BF	2.59	2.93	3.04	2.98	2.83	2.67
$B \to X_c \ell\nu$ BF	29.04	16.78	9.40	5.22	4.34	4.97
Non-resonant model	22.96	23.34	22.87	21.76	20.14	18.07
$B \to X_c \ell\nu$ FF	2.00	1.46	1.30	1.38	1.44	1.41
N_{tracks} res.	9.36	8.90	8.35	7.85	7.34	6.86
N_γ res.	0.07	0.04	0.06	0.06	0.08	0.09
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	1.42	1.93	2.12	2.33	2.28	2.22
q^2 scale	10.53	10.34	9.98	9.72	9.35	8.58
MC non-closure	0.61	0.40	0.41	0.09	0.05	0.03
Cal. function	0.16	0.14	0.11	0.07	0.03	0.01
Stat. bias corr.	1.90	1.84	1.77	1.70	1.62	1.56
PID eff.	0.28	0.27	0.25	0.23	0.22	0.20
Track eff.	0.75	0.71	0.67	0.62	0.58	0.54
$B^0/B^\pm$ tag eff.	0.81	0.92	0.99	1.18	0.88	0.68
Sys. error (total)	40.02	32.32	28.48	26.17	24.29	22.20
Total error	40.09	32.40	28.55	26.24	24.36	22.26
$\langle q^4 \rangle$ in GeV^4	28.33	30.44	32.74	35.20	37.90	40.85
Stat. error (data)	4.69	4.47	4.28	4.11	3.95	3.80
Bkg. subtraction	5.22	5.16	4.73	4.22	3.70	3.18
$B \to X_u \ell\nu$ BF	5.03	5.69	5.97	5.97	5.75	5.54
$B \to X_c \ell\nu$ BF	30.30	18.01	10.39	7.45	8.64	10.45
Non-resonant model	42.24	42.37	41.55	39.85	37.32	33.99
$B \to X_c \ell\nu$ FF	2.75	2.50	2.47	2.57	2.62	2.55
N_{tracks} res.	18.60	18.03	17.29	16.56	15.74	14.92
N_γ res.	0.25	0.22	0.24	0.24	0.27	0.29
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	4.72	5.27	5.46	5.68	5.55	5.39
q^2 scale	22.47	22.19	21.70	21.27	20.65	19.39
MC non-closure	0.12	0.11	0.09	0.00	0.04	0.02
Cal. function	0.26	0.21	0.16	0.09	0.02	0.06
Stat. bias corr.	3.80	3.75	3.67	3.58	3.49	3.38
PID eff.	0.54	0.52	0.50	0.47	0.45	0.42
Track eff.	1.47	1.42	1.35	1.29	1.22	1.15
$B^0/B^\pm$ tag eff.	1.58	1.71	1.78	2.01	1.58	1.28
Sys. error (total)	60.46	55.22	52.13	49.81	47.35	44.20
Total error	60.64	55.40	52.30	49.98	47.52	44.37
$\langle q^6 \rangle$ in GeV^6	211.98	227.49	244.46	263.32	284.76	309.04
Stat. error (data)	7.74	7.45	7.22	7.02	6.82	6.63
Bkg. subtraction	6.42	6.52	6.10	5.57	5.03	4.47
$B \to X_u \ell\nu$ BF	8.17	9.22	9.61	9.59	9.26	8.98
$B \to X_c \ell\nu$ BF	23.10	12.17	8.19	10.58	14.02	16.50
Non-resonant model	58.13	57.92	56.70	54.50	51.31	47.15
$B \to X_c \ell\nu$ FF	3.69	3.56	3.56	3.64	3.67	3.59
N_{tracks} res.	28.50	27.82	26.95	26.04	25.00	23.90
N_γ res.	0.59	0.55	0.58	0.57	0.60	0.62
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	9.15	9.71	9.87	10.05	9.83	9.56
q^2 scale	35.98	35.63	35.03	34.46	33.61	31.98
MC non-closure	0.03	0.02	0.01	0.00	0.01	0.01
Cal. function	0.33	0.27	0.18	0.09	0.03	0.15
Stat. bias corr.	6.34	6.26	6.16	6.04	5.90	5.75
PID eff.	0.81	0.79	0.76	0.73	0.70	0.66
Track eff.	2.20	2.15	2.07	2.00	1.91	1.82
$B^0/B^\pm$ tag eff.	1.96	2.09	2.18	2.43	1.93	1.57
Sys. error (total)	79.22	76.36	74.31	72.33	69.63	65.83
Total error	79.59	76.72	74.66	72.67	69.96	66.16
$\langle q^8 \rangle$ in GeV^8	1744.59	1867.12	2000.34	2151.09	2324.52	2523.62
Stat. error (data)	12.09	11.72	11.45	11.21	10.98	10.74
Bkg. subtraction	8.03	8.47	8.04	7.49	6.93	6.38
$B \to X_u \ell\nu$ BF	12.36	14.05	14.46	14.35	13.78	13.34
$B \to X_c \ell\nu$ BF	13.00	6.76	11.50	17.02	20.90	23.22
Non-resonant model	71.72	71.13	69.45	66.71	62.88	57.99
$B \to X_c \ell\nu$ FF	4.83	4.71	4.70	4.73	4.75	4.70
N_{tracks} res.	39.28	38.48	37.44	36.33	35.03	33.66
N_γ res.	1.13	1.09	1.11	1.10	1.12	1.14
$E_{\text{miss}} - \lvert\mathbf{p}_{\text{miss}}\rvert$ shape	14.66	15.22	15.32	15.43	15.08	14.67
q^2 scale	51.22	50.79	50.06	49.30	48.18	46.17
MC non-closure	0.00	0.00	0.00	0.00	0.00	0.00
Cal. function	0.41	0.32	0.21	0.07	0.10	0.28
Stat. bias corr.	9.92	9.79	9.61	9.41	9.19	8.95
PID eff.	1.09	1.07	1.04	1.00	0.96	0.91
Track eff.	2.98	2.91	2.83	2.74	2.63	2.52
$B^0/B^\pm$ tag eff.	1.77	1.93	2.02	2.30	1.76	1.36
Sys. error (total)	100.24	98.99	97.47	95.48	92.33	87.82
Total error	100.96	99.68	98.14	96.13	92.98	88.47

References

[1] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531.

[2] M. Kobayashi and T. Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[3] HFLAV Collaboration. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018". In: *Eur. Phys. J. C* 81.3 (2021), p. 226. DOI: 10.1140/epjc/s10052-020-8156-7. arXiv: 1909.12524 [hep-ex].

[4] J. Chay, H. Georgi, and B. Grinstein. "Lepton energy distributions in heavy meson decays from QCD". In: *Phys. Lett. B* 247 (1990), pp. 399–405. DOI: 10.1016/0370-2693(90)90916-T.

[5] Ikaros I. Y. Bigi et al. "QCD predictions for lepton spectra in inclusive heavy flavor decays". In: *Phys. Rev. Lett.* 71 (1993), pp. 496–499. DOI: 10.1103/PhysRevLett.71.496. arXiv: hep-ph/9304225.

[6] M. Fael, T. Mannel, and K. Keri Vos. "V_{cb} determination from inclusive $b \to c$ decays: an alternative method". In: *JHEP* 02 (2019), p. 177. DOI: 10.1007/JHEP02(2019)177. arXiv: 1812.07472 [hep-ph].

[7] CLEO Collaboration. "Moments of the B meson inclusive semileptonic decay rate using neutrino reconstruction". In: *Phys. Rev. D* 70 (2004), p. 032002. DOI: 10.1103/PhysRevD.70.032002. arXiv: hep-ex/0403052.

[8] R. P. Feynman. "Mathematical formulation of the quantum theory of electromagnetic interaction". In: *Phys. Rev.* 80 (1950). Ed. by L. M. Brown, pp. 440–457. DOI: 10.1103/PhysRev.80.440.

[9] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[10] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[11] A. Salam and J. C. Ward. "Electromagnetic and weak interactions". In: *Phys. Lett.* 13 (1964), pp. 168–171. DOI: 10.1016/0031-9163(64)90711-5.

[12] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys. Rev. Lett.* 30 (1973). Ed. by J. C. Taylor, pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[13] D. J. Gross and F. Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Phys. Rev. Lett.* 30 (1973). Ed. by J. C. Taylor, pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[14] G. 't Hooft and M. J. G. Veltman. "Regularization and Renormalization of Gauge Fields". In: *Nucl. Phys. B* 44 (1972), pp. 189–213. DOI: 10.1016/0550-3213(72)90279-9.

[15] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[16] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[17] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[18] C. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". In: *Phys. Rev.* 96 (1954). Ed. by J.-P. Hsu and D. Fine, p. 191. DOI: 10.1103/PhysRev.96.191.

[19] A. Salam and J. C. Ward. "Gauge theory of elementary interactions". In: *Phys. Rev.* 136 (1964), B763–B768. DOI: 10.1103/PhysRev.136.B763.

[20] E. Noether. "Invariant variation problems". In: *Trans. Theory. Stat. Phys.* 1.3 (1971), 186–207. ISSN: 1532-2424. DOI: 10.1080/00411457108231446.

[21] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[22] CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[23] *Wikipedia*. [Online; accessed 2021.07.29]. 2019. URL: https://en.wikipedia.org/wiki/Standard_Model.

[24] A. V. Manohar and M. B. Wise. *Heavy Quark Physics*. Cambridge University Press, 2007.

[25] M. Neubert. "B decays and the heavy quark expansion". In: *Adv. Ser. Direct. High Energy Phys.* 15 (1998). Ed. by A. J. Buras and M. Lindner, pp. 239–293. DOI: 10.1142/9789812812667_0003. arXiv: hep-ph/9702375.

[26] F. J. Roschk. "Effects of Higher Orders in $1/m_b$ in Semileptonic B-Meson Decays". MA thesis. Dresden University of Technology, 2004.

[27] D. Griffiths. *Introduction to Elementary Particles*. Portland, USA: WILEY-VCH, 2008.

[28] J. D. Richman and P. R. Burchat. "Leptonic and semileptonic decays of charm and bottom hadrons". In: *Rev. Mod. Phys.* 67 (1995), pp. 893–976. DOI: 10.1103/RevModPhys.67.893. arXiv: hep-ph/9508250.

[29] A. Bettini. *Introduction to Elementary Particle Physics*. Cambridge University Press, 2014.

[30] Particle Data Group. "Review of Particle Physics". In: *Prog. Theor. Exp. Phys.* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[31] Y. Nambu. "Quasi-particles and Gauge Invariance in the Theory of Superconductivity". In: *Phys. Rev.* 117 (1960). Ed. by J. C. Taylor, pp. 648–663. DOI: 10.1103/PhysRev.117.648.

[32] J. Goldstone. "Field Theories with 'Superconductor' Solutions". In: *Nuovo Cim.* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722.

[33] T. D. Lee and C. N. Yang. "Question of Parity Conservation in Weak Interactions". In: *Phys. Rev.* 104 (1956), pp. 254–258. DOI: 10.1103/PhysRev.104.254.

[34] C. S. Wu et al. "Experimental Test of Parity Conservation in β Decay". In: *Phys. Rev.* 105 (1957), pp. 1413–1414. DOI: 10.1103/PhysRev.105.1413.

[35] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138.

[36] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285.

[37] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett.* 51 (1983), p. 1945. DOI: 10.1103/PhysRevLett.51.1945.

[38] H. R. Quinn and P. F. Harrison. *The BABAR physics book: Physics at an asymmetric B factory*. Oct. 1998. DOI: 10.2172/979931.

[39] CKMfitter Group. "CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories". In: *Eur. Phys. J. C* 41.1 (2005), pp. 1–131. DOI: 10.1140/epjc/s2005-02169-1. arXiv: hep-ph/0406184. URL: http://ckmfitter.in2p3.fr.

[40] UTfit Collaboration. "The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of $\Delta m_{(s)}$ and BR($B \to \tau\nu_\tau$)". In: *JHEP* 10 (2006), p. 081. DOI: 10.1088/1126-6708/2006/10/081. arXiv: hep-ph/0606167. URL: http://www.utfit.org/UTfit/WebHome.

[41] J. Ellis. "TikZ-Feynman: Feynman diagrams with Ti*k*Z". In: *Comput. Phys. Commun.* 210 (2017), pp. 103–123. DOI: 10.1016/j.cpc.2016.08.019. arXiv: 1601.05437 [hep-ph].

[42] J. Dingfelder and T. Mannel. "Leptonic and semileptonic decays of B mesons". In: *Rev. Mod. Phys.* 88.3 (2016), p. 035008. DOI: 10.1103/RevModPhys.88.035008.

[43] J. G. Korner and G. A. Schuler. "Exclusive Semileptonic Heavy Meson Decays Including Lepton Mass Effects". In: *Z. Phys. C* 46 (1990), p. 93. DOI: 10.1007/BF02440838.

[44] BaBar Collaboration. "Measurement of the inclusive electron spectrum in charmless semileptonic B decays near the kinematic endpoint and determination of $|V_{ub}|$". In: *Phys. Rev. D* 73 (2006), p. 012006. DOI: 10.1103/PhysRevD.73.012006. arXiv: hep-ex/0509040.

[45] Belle Collaboration. "Precise determination of the CKM matrix element $|V_{cb}|$ with $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle". In: *arXiv preprint* (Feb. 2017). arXiv: 1702.01521 [hep-ex].

[46] D. Bigi and P. Gambino. "Revisiting $B \to D\ell\nu$". In: *Phys. Rev. D* 94.9 (2016), p. 094008. DOI: 10.1103/PhysRevD.94.094008. arXiv: 1606.08030 [hep-ph].

[47] D. Bigi, P. Gambino, and S. Schacht. "A fresh look at the determination of $|V_{cb}|$ from $B \to D^*\ell\nu$". In: *Phys. Lett. B* 769 (2017), pp. 441–445. DOI: 10.1016/j.physletb.2017.04.022. arXiv: 1703.06124 [hep-ph].

[48] F. U. Bernlochner et al. "Tensions and correlations in $|V_{cb}|$ determinations". In: *Phys. Rev. D* 96.9 (2017), p. 091503. DOI: 10.1103/PhysRevD.96.091503. arXiv: 1708.07134 [hep-ph].

[49] D. Ferlewicz, P. Urquijo, and E. Waheed. "Revisiting fits to $B^0 \to D^{*-}\ell^+\nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at non-zero recoil". In: *arXiv preprint* (Aug. 2020). arXiv: 2008.09341 [hep-ph].

[50] Fermilab Lattice and MILC Collaborations. "Semileptonic form factors for $B \to D^*\ell\nu$ at nonzero recoil from 2 + 1-flavor lattice QCD". In: *arXiv preprint* (May 2021). arXiv: 2105.14019 [hep-lat].

[51] Belle Collaboration. "Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$". In: *Phys. Rev. D* 93.3 (2016), p. 032006. DOI: 10.1103/PhysRevD.93.032006. arXiv: 1510.03657 [hep-ex].

[52] Belle Collaboration. "Measurement of the CKM matrix element $|V_{cb}|$ from $B^0 \to D^{*-}\ell^+\nu_\ell$ at Belle". In: *Phys. Rev. D* 100.5 (2019), p. 052007. DOI: 10.1103/PhysRevD.100.052007. arXiv: 1809.03290 [hep-ex].

[53] F. U. Bernlochner and Z. Ligeti. "Semileptonic $B_{(s)}$ decays to excited charmed mesons with e, μ, τ and searching for new physics with $R(D^{**})$". In: *Phys. Rev. D* 95.1 (2017), p. 014022. DOI: 10.1103/PhysRevD.95.014022. arXiv: 1606.09300 [hep-ph].

[54] A. Sirlin. "Large m_W, m_Z Behavior of the O(α) Corrections to Semileptonic Processes Mediated by W". In: *Nucl. Phys. B* 196 (1982), pp. 83–92. DOI: 10.1016/0550-3213(82)90303-0.

[55] G. Caria. "Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a Semileptonic Tag at the Belle Experiment". PhD thesis. University of Melbourne, 2019.

[56] I. Caprini, L. Lellouch, and M. Neubert. "Dispersive bounds on the shape of $\bar{D} \to D^{(*)}\ell\bar{\nu}$ form factors". In: *Nucl. Phys. B* 530 (1998), pp. 153–181. DOI: 10.1016/S0550-3213(98)00350-2. arXiv: hep-ph/9712417.

[57] C. G. Boyd, B. Grinstein, and R. F. Lebed. "Constraints on form-factors for exclusive semileptonic heavy to light meson decays". In: *Phys. Rev. Lett.* 74 (1995), pp. 4603–4606. DOI: 10.1103/PhysRevLett.74.4603. arXiv: hep-ph/9412324 [hep-ph].

[58] Belle Collaboration. "Study of $B \to D^{**}\ell\nu$ with full reconstruction tagging". In: *Phys. Rev. D* 77 (2008), p. 091503. DOI: 10.1103/PhysRevD.77.091503. arXiv: 0711.3252 [hep-ex].

[59] BaBar Collaboration. "Measurement of the Branching Fractions of $\bar{B} \to D^{**}l^-\bar{\nu}_\ell$ Decays in Events Tagged by a Fully Reconstructed B Meson". In: *Phys. Rev. Lett.* 101 (2008), p. 261802. DOI: 10.1103/PhysRevLett.101.261802. arXiv: 0808.0528 [hep-ex].

[60] F. U. Bernlochner, Z. Ligeti, and S. Turczyk. "A Proposal to solve some puzzles in semileptonic B decays". In: *Phys. Rev.* D85 (2012), p. 094033. DOI: 10.1103/PhysRevD.85.094033. arXiv: 1202.1834 [hep-ph].

[61] A. K. Leibovich et al. "Semileptonic B decays to excited charmed mesons". In: *Phys. Rev. D* 57 (1998), pp. 308–330. DOI: 10.1103/PhysRevD.57.308. arXiv: hep-ph/9705467.

[62] T. Mannel. "Effective Field Theories for Heavy Quarks: Heavy Quark Effective Theory and Heavy Quark Expansion". In: *Proceedings, Les Houches summer school: EFT in Particle Physics and Cosmology*. Ed. by S. Davidson et al. Vol. 108. Lecture Notes of the Les Houches Summer School. Oxford University Press, Apr. 2020. DOI: 10.1093/oso/9780198855743.003.0009.

[63] D. Benson et al. "Imprecated, yet impeccable: On the theoretical evaluation of $\Gamma(B \to X_c \ell \nu)$". In: *Nucl. Phys. B* 665 (2003), pp. 367–401. DOI: 10.1016/S0550-3213(03)00452-8. arXiv: hep-ph/0302262.

[64] P. Gambino, K. J. Healey, and S. Turczyk. "Taming the higher power corrections in semileptonic B decays". In: *Phys. Lett. B* 763 (2016), pp. 60–65. DOI: 10.1016/j.physletb.2016.10.023. arXiv: 1606.06174 [hep-ph].

[65] M. Jezabek and Johann H. Kuhn. "QCD Corrections to Semileptonic Decays of Heavy Quarks". In: *Nucl. Phys. B* 314 (1989), pp. 1–6. DOI: 10.1016/0550-3213(89)90108-9.

[66] V. Aquila et al. "Perturbative corrections to semileptonic b decay distributions". In: *Nucl. Phys. B* 719 (2005), pp. 77–102. DOI: 10.1016/j.nuclphysb.2005.04.031. arXiv: hep-ph/0503083.

[67] A. Pak and A. Czarnecki. "Heavy-to-heavy quark decays at NNLO". In: *Phys. Rev. D* 78 (2008), p. 114015. DOI: 10.1103/PhysRevD.78.114015. arXiv: 0808.3509 [hep-ph].

[68] K. Melnikov. "$\mathcal{O}(\alpha_s^2)$ corrections to semileptonic decay $b \to c l \bar{\nu}_l$". In: *Phys. Lett. B* 666 (2008), pp. 336–339. DOI: 10.1016/j.physletb.2008.07.089. arXiv: 0803.0951 [hep-ph].

[69] T. Becher, H. Boos, and E. Lunghi. "Kinetic corrections to $B \to X_c \ell \bar{\nu}$ at one loop". In: *JHEP* 12 (2007), p. 062. DOI: 10.1088/1126-6708/2007/12/062. arXiv: 0708.0855 [hep-ph].

[70] T. Mannel, A. A. Pivovarov, and D. Rosenthal. "Inclusive semileptonic B decays from QCD with NLO accuracy for power suppressed terms". In: *Phys. Lett. B* 741 (2015), pp. 290–294. DOI: 10.1016/j.physletb.2014.12.058. arXiv: 1405.5072 [hep-ph].

[71] A. Alberti, P. Gambino, and S. Nandi. "Perturbative corrections to power suppressed effects in semileptonic B decays". In: *JHEP* 01 (2014), p. 147. DOI: 10.1007/JHEP01(2014)147. arXiv: 1311.7381 [hep-ph].

[72] T. Mannel, S. Turczyk, and N. Uraltsev. "Higher Order Power Corrections in Inclusive B Decays". In: *JHEP* 11 (2010), p. 109. DOI: 10.1007/JHEP11(2010)109. arXiv: 1009.4622 [hep-ph].

[73] CDF Collaboration. "Measurement of the moments of the hadronic invariant mass distribution in semileptonic B decays". In: *Phys. Rev. D* 71 (2005), p. 051103. DOI: 10.1103/PhysRevD.71.051103. arXiv: hep-ex/0502003.

[74] DELPHI Collaboration. "Determination of heavy quark non-perturbative parameters from spectral moments in semileptonic B decays". In: *Eur. Phys. J. C* 45 (2006), pp. 35–59. DOI: 10.1140/epjc/s2005-02406-7. arXiv: hep-ex/0510024.

[75] BaBar Collaboration. "Measurement of the electron energy spectrum and its moments in inclusive $B \to Xe\nu$ decays". In: *Phys. Rev. D* 69 (2004), p. 111104. DOI: 10.1103/PhysRevD.69.111104. arXiv: hep-ex/0403030.

[76] BaBar Collaboration. "Measurement and interpretation of moments in inclusive semileptonic decays $\bar{B} \to X_c \ell^- \bar{\nu}$". In: *Phys. Rev. D* 81 (2010), p. 032003. DOI: 10.1103/PhysRevD.81.032003. arXiv: 0908.0415 [hep-ex].

[77] Belle Collaboration. "Moments of the electron energy spectrum and partial branching fraction of $B \to X_c e \nu$ decays at Belle". In: *Phys. Rev. D* 75 (2007), p. 032001. DOI: 10.1103/PhysRevD.75.032001. arXiv: hep-ex/0610012.

[78] Belle Collaboration. "Moments of the Hadronic Invariant Mass Spectrum in $B \to X_c \ell \nu$ Decays at Belle". In: *Phys. Rev. D* 75 (2007), p. 032005. DOI: 10.1103/PhysRevD.75.032005. arXiv: hep-ex/0611044.

[79] P. Gambino and C. Schwanda. "Inclusive semileptonic fits, heavy quark masses, and V_{cb}". In: *Phys. Rev. D* 89.1 (2014), p. 014022. DOI: 10.1103/PhysRevD.89.014022. arXiv: 1307.4551 [hep-ph].

[80] A. Alberti et al. "Precision Determination of the Cabibbo-Kobayashi-Maskawa Element V_{cb}". In: *Phys. Rev. Lett.* 114.6 (2015), p. 061802. DOI: 10.1103/PhysRevLett.114.061802. arXiv: 1411.6560 [hep-ph].

[81] UTFit Collaboration. "Updates on the Standard Model and beyond SM Unitarity Triangle fits and neutral charm mixing results by UTfit". 40th International Conference on High Energy Physics (ICHEP 2020). 2020. URL: https://indico.cern.ch/event/868940/contributions/3815693/attachments/2082161/3497433/bona-utfit.pdf.

[82] P. Gambino. "B semileptonic moments at NNLO". In: *JHEP* 09 (2011), p. 055. DOI: 10.1007/JHEP09(2011)055. arXiv: 1107.3100 [hep-ph].

[83] *SuperKEKB collider achieves the world's highest luminosity*. [Online; accessed 2021.08.25]. 2020. URL: https://www.kek.jp/en/newsroom/2020/06/26/1400/.

[84] *KEK B-Factory Design Report*. KEK Report 95-7. 1995.

[85] S. Kurokawa and E. Kikutani. "Overview of the KEKB accelerators". In: *Nucl. Instrum. Meth. A* 499.1 (2003), pp. 1–7. DOI: 10.1016/S0168-9002(02)01771-0.

[86] G. Barr et al. *Particle Physics in the LHC Era*. Vol. 24. Oxford University Press, 2016, pp. 62–64.

[87] W. Herr and B. Muratori. "Concept of luminosity". In: *CERN Accelerator School: Intermediate Course on Accelerator Physics*. CERN. 2006, pp. 361–378.

[88] *Wikipedia*. [Online; accessed 2021.07.29]. 2013. URL: https://en.wikipedia.org/wiki/KEKB_(accelerator).

[89] CLEO Collaboration. "Measurement of the B Semileptonic Branching Fraction with Lepton Tags". In: *Phys. Rev. Lett.* 76.10 (1996), p. 1570. DOI: 10.1103/PhysRevLett.76.1570.

[90] D. Besson and T. Skwarnicki. "Upsilon Spectroscopy: Transitions in the Bottomonium System". In: *Annu. Rev. Nucl. Part. Sci.* 43.1 (1993), pp. 333–378. DOI: 10.1146/annurev.ns.43.120193.002001.

[91] Y. Ohnishi et al. "Beam Background and MDI Design for SuperKEKB/Belle-II". In: *Conf. Proc. C* 110904 (2011), pp. 3703–3705.

[92] Belle Collaboration. "The Belle Detector". In: *Nucl. Instrum. Meth. A* 479.1 (2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[93] P. Goldenzweig. "A Study of Neutral B Meson Decays to ωK^{*0} at Belle". PhD thesis. University of Cincinnati, 2008.

[94] Z. Natkaniec et al. "Status of the Belle silicon vertex detector". In: *Nucl. Instrum. Meth. A* 560.1 (2006), pp. 1–4. DOI: 10.1016/j.nima.2005.11.228.

[95] T. Iijima et al. "Aerogel Cherenkov counter for the BELLE detector". In: *Nucl. Instrum. Meth. A* 453.1-2 (2000), pp. 321–325. DOI: 10.1016/S0168-9002(00)00652-5.

[96] Y. Ushiroda et al. "Development of the central trigger system for the BELLE detector at the KEK B-factory". In: *Nucl. Instrum. Meth. A* 438.2-3 (1999), pp. 460–471. DOI: 10.1016/S0168-9002(99)00823-2.

[97] T. Hojo, M. Hazumi, and R. Itoh. "Level 3 trigger at Belle". Belle note 421. 2001.

[98] N. Katayama et al. "Belle computing model". In: *Comput. Phys. Commun.* 110.1-3 (1998), pp. 22–25. DOI: 10.1016/S0010-4655(97)00148-3.

[99] R. Itoh. "BASF-Belle AnalysiS Framework". In: *9th International Conference on Computing in High-Energy Physics (CHEP 1997)*. 1997.

[100] J. Brodzicka et al. "Physics Achievements from the Belle Experiment". In: *Prog. Theor. Exp. Phys.* 2012 (2012), p. 04D001. DOI: 10.1093/ptep/pts072. arXiv: 1212.5342 [hep-ex].

[101] N. Metropolis and S. Ulam. "The monte carlo method". In: *J. Am. Stat. Assoc.* 44.247 (1949), pp. 335–341. DOI: 10.1080/01621459.1949.10483310.

[102] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462.1-2 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[103] T. Sjöstrand et al. "High-energy physics event generation with PYTHIA 6.1". In: *Comput. Phys. Commun.* 135 (2001), pp. 238–259. DOI: 10.1016/S0010-4655(00)00236-8. arXiv: hep-ph/0010017.

[104] R. Brun et al. *GEANT 3: user's guide Geant 3.10, Geant 3.11*. Tech. rep. CERN, 1987.

[105] E. Barberio, B. van Eijk, and Z. Wąs. "PHOTOS –a universal Monte Carlo for QED radiative corrections in decays". In: *Comput. Phys. Commun.* 66.1 (1991), pp. 115–128. DOI: 10.1016/0010-4655(91)90012-A.

[106] E. Barberio and Z. Wąs. "PHOTOS –a universal Monte Carlo for QED radiative corrections: version 2.0". In: *Comput. Phys. Commun.* 79.2 (1994), pp. 291–308. DOI: 10.1016/0010-4655(94)90074-4.

[107] Belle Collaboration. "Measurements of Partial Branching Fractions of Inclusive $B \to X_u \ell^+ \nu_\ell$ Decays with Hadronic Tagging". In: *Phys. Rev. D* 104.1 (2021), p. 012008. DOI: 10.1103/PhysRevD.104.012008. arXiv: 2102.00020 [hep-ex].

[108] LHCb Collaboration. "Measurements of the Branching fractions for $B_{(s)} \to D_{(s)}\pi\pi\pi$ and $\Lambda_b^0 \to \Lambda_c^+\pi\pi\pi$". In: *Phys. Rev. D* 84 (2011). [Erratum: Phys.Rev.D 85, 039904 (2012)], p. 092001. DOI: 10.1103/PhysRevD.84.092001. arXiv: 1109.6831 [hep-ex].

[109] BaBar Collaboration. "Observation of $\overline{B} \to D^{(*)}\pi^+\pi^-\ell^-\overline{\nu}$ decays in e^+e^- collisions at the $\Upsilon(4S)$ resonance". In: *Phys. Rev. Lett.* 116.4 (2016), p. 041801. DOI: 10.1103/PhysRevLett.116.041801. arXiv: 1507.08303 [hep-ex].

[110] C. Ramirez, J. F. Donoghue, and G. Burdman. "Semileptonic $b \to u$ decay". In: *Phys. Rev.* D 41 (5 1990), pp. 1496–1503. DOI: 10.1103/PhysRevD.41.1496.

[111] Belle Collaboration. "Search for $B^+ \to \mu^+\nu_\mu$ and $B^+ \to \mu^+ N$ with inclusive tagging". In: *Phys. Rev. D* 101.3 (2020), p. 032007. DOI: 10.1103/PhysRevD.101.032007. arXiv: 1911.03186 [hep-ex].

[112] M. Prim. *b2-hive/eFFORT v0.1.0*. Version v0.1.0. July 2020. DOI: 10.5281/zenodo.3965699. URL: https://doi.org/10.5281/zenodo.3965699.

[113] B. Grinstein and A. Kobach. "Model-Independent Extraction of $|V_{cb}|$ from $\bar{B} \to D^* \ell \overline{\nu}$". In: *Phys. Lett.* B 771 (2017), pp. 359–364. DOI: 10.1016/j.physletb.2017.05.078. arXiv: 1703.08170 [hep-ph].

[114] D. Scora and N. Isgur. "Semileptonic meson decays in the quark model: An update". In: *Phys. Rev. D* 52 (1995), pp. 2783–2812. DOI: 10.1103/PhysRevD.52.2783. arXiv: hep-ph/9503486.

[115] M. Welsch. "Measurement of the branching fraction of $B \to D^{**}(\to D^{(*)}\pi^0 \ell \nu_\ell)$ in the context of a $R(D^{(*)})$ analysis at the Belle experiment". MA thesis. Karlsruhe Institute of Technology, 2019.

[116] L. Hinz et al. "Lepton ID efficiency correction and systematic error". Belle Note 954. 2006.

[117] S. Nishida et al. "Study of Kaon and Pion Identification Using Inclusive D^* Sample". Belle Note 779. 2005.

[118] P. Urquijo. "Semi-inclusive semileptonic B decays". Belle Note 1251. 2012.

[119] B. Bhuyan. "High P_T tracking efficiency using partially reconstructed D^* decays". Belle Note 1165. 2010.

[120] K. Hanagaki et al. "Electron identification in Belle". In: *Nucl. Instrum. Meth. A* 485 (2002), pp. 490–503. DOI: 10.1016/S0168-9002(01)02113-1. arXiv: hep-ex/0108044.

[121] A. Abashian et al. "Muon identification in the Belle experiment at KEKB". In: *Nucl. Instrum. Meth. A* 491.1-2 (2002), pp. 69–82. DOI: 10.1016/S0168-9002(02)01164-6.

[122] T. Keck. "Machine learning algorithms for the Belle II experiment and their validation on Belle data". PhD thesis. Karlsruhe Institute of Technology, 2017.

[123] M. Feindt et al. "A Hierarchical NeuroBayes-based Algorithm for Full Reconstruction of B Mesons at B Factories". In: *Nucl. Instrum. Meth. A* 654 (2011), pp. 432–440. DOI: 10.1016/j.nima.2011.06.008. arXiv: 1102.3876 [hep-ex].

[124] Ed. A.J. Bevan, B. Golob, Th. Mannel, S. Prell, and B.D. Yabsley. "The Physics of the B Factories". In: *Eur. Phys. J. C* 74 (2014), p. 3026. DOI: 10.1140/epjc/s10052-014-3026-9. arXiv: 1406.6311 [hep-ex].

[125] M. Röhrken. "Time-dependent CP violation measurements in neutral B meson to double-charm decays at the Japanese Belle Experiment". PhD thesis. Karlsruhe Institute of Technology, 2012.

[126] CLEO Collaboration. "Search for exclusive charmless hadronic B decays". In: *Phys. Rev. D* 53 (1996), pp. 1039–1050. DOI: 10.1103/PhysRevD.53.1039. arXiv: hep-ex/9508004.

[127] H. Kakuno. "Measurement of $|V_{ub}|$ through Charmless Semileptonic B-meson Decay". PhD thesis. Tokyo Institute of Technology, 2003.

[128] G. Cowan. *Statistical Data Analysis*. Oxford university press, 1998.

[129] Belle II Collaboration. "A calibration of the Belle II hadronic tag-side reconstruction algorithm with $B \to X \ell \nu$ decays". In: *arXiv preprint* (Aug. 2020). arXiv: 2008.06096 [hep-ex].

[130] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.

[131] H. Dembinski, P. Ongmongkolkul, et al. *scikit-hep/iminuit*. Dec. 2020. DOI: 10.5281/zenodo.3949207. URL: https://doi.org/10.5281/zenodo.3949207.

[132] J. E. Sundermann. "Improved Measurements of Hadronic Mass Moments in Decays $B \to X_c l\nu$ and Determination of $|V_{cb}|$". PhD thesis. Dresden University of Technology, 2006.

[133] V. Klose. "Measurement and interpretation of moments of the combined hadronic mass and energy spectrum in inclusive semileptonic B-meson decays". PhD thesis. Dresden University of Technology, 2007.

[134] Belle II Collaboration. "Measurement of Hadronic Mass Moments $\langle M_X^n \rangle$ in $B \to X_c \ell\nu$ Decays at Belle II". In: *arXiv preprint* (Sept. 2020). arXiv: 2009.04493 [hep-ex].

[135] B. Efron. "Bootstrap Methods: Another Look at the Jackknife". In: *Ann. Stat.* 7.1 (1979), pp. 1 –26. DOI: 10.1214/aos/1176344552.

[136] K. G. Hayes, M. L. Perl, and B. Efron. "Application of the bootstrap statistical method to the tau-decay-mode problem". In: *Phys. Rev. D* 39 (1 1989), pp. 274–279. DOI: 10.1103/PhysRevD.39.274.

[137] BaBar Collaboration. "Measurement of the inclusive electron spectrum from B meson decays and determination of $|V_{ub}|$". In: *Phys. Rev. D* 95.7 (2017), p. 072001. DOI: 10.1103/PhysRevD.95.072001. arXiv: 1611.05624 [hep-ex].

[138] F. U. Bernlochner et al. "Constraints on exclusive branching fractions $\mathcal{B}_i(B^+ \to X_c^i l^+ \nu)$ from moment measurements in inclusive $B \to X_c l\nu$ decays". In: *Eur. Phys. J. C* 74.6 (2014), p. 2914. DOI: 10.1140/epjc/s10052-014-2914-3. arXiv: 1402.2849 [hep-ph].

[139] R. van Tonder. "Inclusive determination of $|V_{cb}|$ at Belle". In: *55th Rencontres de Moriond on Electroweak Interactions and Unified Theories*. May 2021. arXiv: 2105.08001 [hep-ex].

[140] K. Olschewsky, M. Fael, and K. Keri Vos. "Inclusive determination of V_{cb} using a new method". 55th Rencontres de Moriond on Electroweak Interactions and Unified Theories (Moriond EW 2021). 2021. URL: https://moriond.in2p3.fr/2021/EW/slides/3_flavour_ysf_09_olschewsky.pdf.

[141] O. Buchmüller and H. Flächer. "Fit to moment from $B \to X_c \ell\bar{\nu}$ and $B \to X_s \gamma$ decays using heavy quark expansions in the kinetic scheme". In: *Phys. Rev. D* 73 (2006), p. 073008. DOI: 10.1103/PhysRevD.73.073008. arXiv: hep-ph/0507253.

[142] T. Mannel, M. Rahimi, and K. Keri Vos. "Impact of background effects on the inclusive V_{cb} determination". In: *JHEP* 09 (2021), p. 051. DOI: 10.1007/JHEP09(2021)051. arXiv: 2105.02163 [hep-ph].