Measurement of Lepton Mass Squared Moments of $B \to X_c \ell \nu_\ell$ Decays with the Belle II Experiment and Determination of the CKM Matrix Element $|V_{cb}|$

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

von
Maximilian Welsch
aus
Ludwigsburg

Bonn, 28.07.2022

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:	13.12.2022
Erscheinungsjahr:	2023

Abstract

This thesis presents measurements of the first to fourth moments of the lepton mass squared $q^2 = (p_\ell + p_{\nu_\ell})^2$ distribution of $B \to X_c \ell \nu_\ell$ decays for $\ell = e, \mu$ and with X_c denoting a hadronic system containing a charm quark. These results use a sample of e^+e^--collisions at the $\Upsilon(4S)$ resonance corresponding to $62.8\,\mathrm{fb}^{-1}$ of integrated luminosity and collected by the Belle II experiment in 2019 and 2020. To identify the X_c system and reconstruct q^2, one of the B mesons from an $\Upsilon(4S) \to B\overline{B}$ decay is fully reconstructed in a hadronic decay mode using a multivariate B tagging algorithm. The yields of the remaining $e^+e^- \to q\overline{q}$ continuum and $B\overline{B}$ background events are determined with a binned likelihood fit to the hadronic mass m_X distribution. To correct the bias in the mean value of the reconstructed q^2 distribution, an event-wised calibration method is applied to the reconstructed q^2 values. The q^2 moments are calculated as a weighted mean using the calibrated q^2 values summing all selected events. In this approach, the weights denote a signal probability and effectively subtract the contribution of the remaining background events. This work reports raw and central q^2 moments measured as functions of lower q^2 thresholds from $q^2 > 1.5\,\mathrm{GeV}^2$ up to $q^2 > 8.5\,\mathrm{GeV}^2$, probing up to 77 % of the accessible $B \to X_c \ell \nu_\ell$ phase space. This is the first measurement of moments in the experimentally challenging range of $q^2 \in [1.5, 2.5]$ GeV.

In addition, this thesis presents a first determination of $|V_{cb}|$ using q^2 moments of inclusive $B \to X_c \ell \nu_\ell$ decays. Like the prediction of total rate, the expressions for spectral moments of inclusive semileptonic B decays can be calculated using the heavy quark expansion and depend on a set of non-perturbative parameters which can only be determined from measurements. At higher orders in the expansion, one observes a proliferation of these parameters which further complicates their extraction. The q^2 moments and the total rate are reparameterization invariant quantities and depend on a reduced set of non-perturbative parameters. This reduced set of parameters opens a new path to extract these parameters up to $\mathcal{O}(1/m_b^4)$ in the HQE purely from data and thereby reducing the uncertainty on $|V_{cb}|$. A combined fit to the q^2 moments measured in this work and by the Belle collaboration is used to extract $|V_{cb}|$. The fit results in $|V_{cb}| = (41.69 \pm 0.63) \times 10^{-3}$. This result presents an important independent cross-check of, and is consistent with, the previous state-of-the-art inclusive $|V_{cb}|$ determinations based on lepton energy and hadronic mass moments.

Contents

CHAPTER 1

Introduction

The Standard Model of particle physics (SM) describes the fundamental building blocks of matter, the leptons and quarks, as well as their interactions via the electromagnetic, weak and strong force. In this model, the forces are mediated with the exchange of gauge bosons: photons, gluons, $W^{\pm}$, and Z. With the latest milestone observation of the Higgs boson in 2012 by the LHC experiments ATLAS and CMS [1, 2], the SM is now completed as a self-consistent theory. Over the last decades, the quantitative SM predictions have been in excellent agreement with various experimental measurements. However, there are several phenomena that are not explained by the SM and it is clear that this theory cannot be the final answer to our description of nature on a sub-atomic scale. One of them is the matter-antimatter asymmetry in our matter-dominated universe. A necessary condition for explaining this asymmetry is the existence of charge (C) and charge-parity (CP) violating effects [3]. Without the latter, there would be a CP conjugate process for every process creating matter over antimatter and the necessary asymmetry could not be achieved.

In the SM, the coupling of different quark flavors in weak charged currents is parameterized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [4, 5]. The CKM matrix is a 3×3 unitary matrix relating the weak interaction eigenstates to the mass eigenstates of the down-type d, s, b quarks. Here, CP violating effects are introduced by the irreducible complex phase of the CKM matrix. But these effects are not sufficient to explain the observed matter-antimatter asymmetry. Additional CP violating processes could be introduced with physics beyond the SM predicting new heavy particles. If such new heavy particles couple to quarks in loop-processes, measurements constraining the unitary condition of the CKM matrix might differ from the SM prediction. To test the unitary condition, precise determinations of the CKM matrix elements $|V_{ub}|$, $|V_{cb}|$, and the CKM unitary triangle angle γ are necessary. All three parameters can be measured in tree-level decays. They determine the position of the apex in the unitary triangle and thus, the amount of CP violation present in the weak interaction of quarks described by the SM.

The magnitude of the CKM matrix element V_{cb} can be studied via semileptonic $b \to c\ell\nu_{\ell}$ ($\ell = e, \mu$) tree-level transitions. The decay rate and decay dynamics of semileptonic B meson decays offer a theoretically clean approach to measure $|V_{cb}|$. In such decays, the leptonic and hadronic currents of the matrix element factorize up to a small electroweak correction factor. Consequently, these final states are theoretically easier to understand than fully hadronic decays involving $|V_{cb}|$.

Regarding semileptonic B meson decays, there are two approaches to measure $|V_{cb}|$. The first one focuses on decays to exclusive final-states. Here, results based $B \to D\ell\nu_{\ell}$ and $B \to D^{*}\ell\nu_{\ell}$ decays provide the most precise values. The second approach is based on the study of inclusive $B \to X_c \ell \nu_\ell$ decays where the charmed hadronic final-state is not further specified. The inclusive determination of $|V_{cb}|$ relies on expressing the total decay rate as an expansion in $1/m_b$, the inverse bottom quark mass, using the heavy quark expansion (HQE). At each order in $1/m_b$, the prediction for the total rate depends on a small number of non-perturbative matrix elements and perturbatively calculable Wilson coefficients. Like the rate, expressions for the spectral moments of inclusive observables can be expanded using the HQE and described with the same set of non-perturbative parameters. To date, the expressions for the rate and the moments are available with next-to-next-to-leading and next-to-leading order precision in α_s at leading order in the HQE and up to $\mathcal{O}(1/m_b^3)$ in the HQE, respectively. Currently, no perturbative corrections for the Wilson coefficients are known for the HQE up to $\mathcal{O}(1/m_b^5)$ [6–13]. The non-perturbative HQE parameters and $|V_{cb}|$ are determined simultaneously in a global fit [14, 15] to measurements of the $B \to X_c \ell \nu_\ell$ branching fraction and moments of the hadronic mass m_X and lepton energy E_ℓ distributions [16–22].

The current world averages [23] for $|V_{cb}|$ determined using the exclusive and inclusive approaches are

$$|V_{cb}|_{\text{excl}} = (39.5 \pm 0.9) \times 10^{-3} \quad \text{and} \tag{1.1}$$

$$|V_{cb}|_{\text{incl}} = (42.2 \pm 0.8) \times 10^{-3}, \tag{1.2}$$

respectively. The world averages differ by about three standard deviations. The origin of this discrepancy is currently unknown and commonly referred to as the V_{cb} puzzle. Recent exclusive $|V_{cb}|$ determinations [24–27] are in better agreement with the world average of inclusively determined values. However, these analyses are based on a more general form factor parameterization and have larger uncertainties.

The approximate 2% uncertainty on the inclusive $|V_{cb}|$ world average is mainly driven by theory uncertainties related to the truncation of the HQE and perturbative corrections [28, 29]. Higher order HQE parameters have to be included in the global fit to further reduce this uncertainty. This is complicated by the proliferation of non-perturbative matrix elements at $\mathcal{O}(1/m_b^{4,5})$ in the expansion. At $\mathcal{O}(1/m_b^4)$, there are already 13 parameters present in the expressions for the rate and spectral m_X and E_ℓ moments. The authors of [30] propose a novel and alternative approach to determine $|V_{cb}|$ inclusively avoiding the proliferation of HQE parameters at higher orders in the expansion. By exploiting reparameterization invariance (RPI), the authors show that the number of non-perturbative matrix elements can be reduced to 8 for the HQE expression of the total rate at $\mathcal{O}(1/m_b^4)$. Unfortunately, existing measurements of m_X and E_ℓ moments cannot be used, since their predictions are not reparameterization invariant and the reduction of HQE parameters is not valid. However, RPI is retained in the predictions for spectral moments of the lepton mass squared $q^2 = (p_\ell + p_{\nu_\ell})^2$.

This thesis presents the measurement of the raw and central spectral moments of the lepton mass squared $\langle q^{2n} \rangle$ with $n = 1$–4, with a progression of lower thresholds on q^2 ranging from $q^2 > 1.5\,\text{GeV}^2$ up to $q^2 > 8.5\,\text{GeV}^2$. The analysis is based on the Belle II data set recorded in 2019 and 2020 at a center-of-mass energy of $\sqrt{s} = 10.58\,\text{GeV}$ corresponding to an integrated luminosity of $62.8\,\text{fb}^{-1}$. The measurement is restricted to the region $q^2 > 1.5\,\text{GeV}^2$, since events passing this threshold have signal leptons that can be reliably identified by the detector. A similar measurement, using the full Belle data set has been recently reported by the Belle collaboration [31] with a lowest q^2 threshold of $q^2 > 3.0\,\text{GeV}^2$, covering 58% of the accessible $B \to X_c \ell \nu_\ell$ phase space. The results reported here have a comparable precision and the measurement is extended to include the experimentally challenging region of $q^2 \in [1.5, 2.5]$ GeV, covering up to 77% of the available phase space. Moments with a lower q^2 threshold retain more information about the inclusive $B \to X_c \ell \nu_\ell$ process.

Furthermore, this thesis contains a first extraction of inclusive $|V_{cb}|$ from q^2 moments based on a simultaneous fit to the measurements from this work and from the Belle collaboration. Besides $|V_{cb}|$, the fit determines two $\mathcal{O}(1/m_b^4)$ HQE parameters for the first time in a purely data-driven way. This alternative determination of $|V_{cb}|$ presents an independent cross-check of previous inclusive $|V_{cb}|$ extractions based on new data and a new method using only q^2 moments.

The remainder of this thesis is organized as follows: Chapter 2 gives an overview of the foundations of the SM, the CKM matrix and the theoretical description of inclusive $B \to X_c \ell \nu_\ell$. In addition, recent experimental findings comprising the determination of $|V_{cb}|$ determinations are discussed. The experimental setup, the SuperKEKB e^+e^- accelerator and the Belle II detector are described in Chapter 3. The reconstruction and selection of semileptonic $B \to X_c \ell \nu_\ell$ decays as well as the application of a kinematic fit to the whole $\Upsilon(4S) \to B\overline{B}$ event are explained in Chapter 4. Next, Chapter 5 describes the necessary steps to measure the q^2 moments in more detail. Chapter 6 introduces the method and discusses the results of a first determination of $|V_{cb}|$ based on q^2 moments measured in this work and by the Belle collaboration. Finally, Chapter 7 summarizes the results and presents the conclusions of this thesis.

CHAPTER 2

Theoretical Foundations and Experimental Status of $|V_{cb}|$

This chapter introduces the theoretical background and concepts used in the description of inclusive semileptonic B meson decays and the calculation of observables reconstructed in this measurement. First a short overview of the standard model of particle physics is given in Section 2.1 following the summary given in [32]. Sections 2.2 and 2.3 serve as a reminder of the mathematical formulation of the SM and the effective field theory methods used in heavy quark physics. The presentation of the quantum field theoretical concepts follows [33–35]. Lastly, 2.4 summarizes the current experimental status of measurements of $|V_{cb}|$.

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics (SM) describes the interactions between the fundamental building blocks of matter, the elementary particles. Mathematically, it is formulated as relativistic quantum field theory. The SM Lagrangian is invariant under local gauge transformations defined by the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times U(1)_Y$ gauge group. The invariance under $\mathrm{SU}(3)_C$ results in the strong interactions between quarks and also leads to the formation of protons and neutrons in nuclei. This force is described by the theory of Quantum Chromodynamics (QCD). The gauge group $\mathrm{SU}(2)_L \times U(1)_Y$ unifies the theory of weak and electromagnetic forces. The latter is also known as Quantum Electrodynamics (QED) which is the result of describing electromagnetic interactions in a relativistic field theory.

The SM includes the unified theory of electroweak interactions mediated by the photon, $W^{\pm}$, and Z bosons and the strong interaction mediated by eight gluons. The photon and the gluons are massless, while $W^{\pm}$ and Z are massive particles with masses of $m_W = (80.377 \pm 0.012)$ GeV and $m_Z = (91.188 \pm 0.020)$ GeV [23], respectively. All gauge boson are spin-1 particles. Besides the gauge bosons, the other fundamental particles are classified as quarks and leptons. Both particle types are fermions carrying spin-1/2. They occur in three generations or families of particles that can be arranged in a hierarchial structure based on the particle masses. For the fermions, the masses range from $\mathcal{O}(1)$ MeV to $\mathcal{O}(100)$ GeV [23]. The neutrinos are assumed to be massless in the SM. For each fermion, there is a corresponding anti-fermion with the same mass but opposite quantum numbers. An overview of all SM elementary particles and their corresponding mass, electric charge, and spin is given in Figure 2.1.

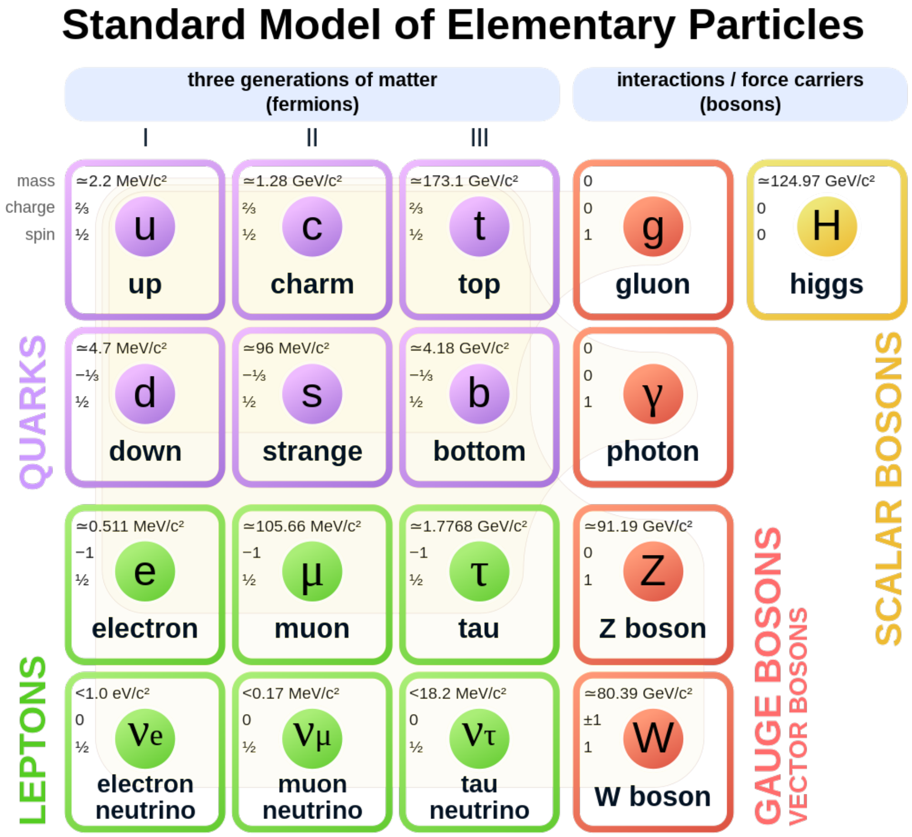


Figure 2.1: Overview of SM fermions and bosons and their corresponding masses, electric charge and spin. Taken from [36].

The coupling of the gauge bosons to fermions is proportional to different kinds of charges. Photons couple to the electric charge q, while gluons couple to charges referred to as color charges. The coupling of W and Z bosons is different for chiral left- and right-handed particles. While the W boson couples only to left-handed fermions, the Z boson interacts with left- and right-handed particles but with different strengths. All gauge bosons except the photon carry charges as well, resulting in different kinds of self-couplings between the gauge bosons.

The consistency of the SM requires the existence of an additional spin-0 particle. The coupling between the scalar particle and the fermions and gauge bosons is proportional to the mass of the corresponding particle. Experimentally, this particle was observed in 2012 independently by the ATLAS and CMS experiments at the LHC with a mass $m_H = (125.25 \pm 0.17)$ GeV [23] commonly referred to as the Higgs particle [1, 2].

Quarks can change their flavor in weak decays. These flavor changing processes prefer transitions within one quark family, e.g. $c \rightarrow s$. Transitions between quark families are suppressed. Transitions between the first and second generation are suppressed with one order of magnitude and transitions between the first and third generation even by two orders of magnitude compared to flavor changing transitions within the same quark family [23]. This phenomenon is described by the quark mixing matrix, also known as CKM matrix, relating the weak interaction and mass eigenstates of the down-type quarks. Flavor changing transitions between leptons are almost universal between the lepton families.

The measured width of the Z boson does not allow for a fourth light neutrino $m_\nu < m_Z$. Searches for additional fermion generations did not show any success. Observations of neutrino-oscillations show that neutrinos must have a non-vanishing mass. The consequent lepton mixing matrix can be incorporated into the SM analogous to the CKM matrix.

Both the photon and the gluon are massless. While the electromagnetic force has effectively an infinite range, the strong force only has an effective range of $O(10^3)$ fm. The origin of this short range are self-interactions between gluons. Compared to the QED potential $V_{\text{QED}} \propto 1/r$, the QCD color potential at large distances can be described by $V_{\text{QCD}} \propto r$. Consequently, it would require an infinite amount of energy to separate two colored particles at macroscopic distances. Nevertheless, it is possible to separate a quark pair spatially. However, no free quarks result from this process. Instead, new quark pairs are created, which together with the original quarks form hadrons. This process is also known as hadronization. Experimentally, no free quarks are observed and colored object always arrange themselves in bound hadronic states that are effectively colorless. This is explained by the confinement hypothesis which states that free particles only occur in a “color neutral” state.

The SM describes the electromagnetic and weak forces in a unified theory. As a result, the weak and electric coupling strength are related to each other with the electroweak mixing angle θ_W. While the parity P and charge-parity C are conserved in electromagnetic and strong interactions, the weak force violates both P and the combined CP parity.

Overall, the SM has 18 free parameters: three coupling constants describing the strengths of the gauge interactions, six quark and three charged lepton masses, four CKM parameters and two parameters describing the Higgs boson properties. In principle, one can also include a phase leading to CP violation in strong interactions; but it is extremely small and often taken to be zero. The number of free parameters would even rise to 26 when one includes the neutrino masses and the corresponding parameters of the lepton mixing matrix.

Although quantitative predictions by the SM are in very good agreement with the findings of various experiments, there are still some open questions that cannot be explained or answered directly by the SM:

- Why are there exactly three generations of fermions?
- What is the origin of CP violation?
- What is the mechanism responsible for the creation of neutrino masses?
- What is dark matter?
- ...

2.2 Electroweak Symmetry Breaking and the CKM Matrix

2.2.1 Local Gauge Invariance and Quantum Electrodynamics

The Lagrangian for the spinor field $\psi(x)$ describing a free Dirac particle with spin 1/2 is given by

$$\mathcal{L}_D = i\overline{\psi}(\gamma^\mu \partial_\mu - m)\psi. \tag{2.1}$$

The Lagrangian $\mathcal{L}_D$ is not invariant under the U(1) local phase transformation

$$\psi(x) \to \psi'(x) = \mathrm{e}^{iq_e\chi(x)}\psi(x). \tag{2.2}$$

However, gauge invariance can be restored by replacing the derivative ∂_μ with the covariant derivative D_μ,

$$\partial_\mu \to D_\mu = \partial_\mu + iq_e A_\mu, \tag{2.3}$$

with A_μ being a new massless field that transforms as

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu \chi. \tag{2.4}$$

Thus, the gauge invariant Lagrangian can be written as

$$\mathcal{L} = i\overline{\psi}(\gamma^\mu\partial_\mu - m)\psi - q_e\overline{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.5}$$

which contains an interaction term between the spin-1/2 particle and the field A_μ and a kinetic term $F_{\mu\nu}F^{\mu\nu}$ describing the dynamics of the massless spin-1 field A_μ. The Lagrangian in Equation (2.5) describes the quantum electrodynamics (QED), the quantum field theory of electromagnetic interactions of fermions via the coupling to the photon. The coupling strength is given by the electric charge q_e.

2.2.2 Electroweak Unification

Analogous to the electromagnetic interaction, weak and strong interactions can be obtained by requiring $\mathcal{L}_D$ to be invariant under local non-abelian SU(2) and SU(3) gauge transformations, respectively. The SU(2) local phase transformation is defined as

$$\varphi(x) \to \varphi'(x) = \mathrm{e}^{ig_W\alpha_k(x)T^k}\varphi(x) \tag{2.6}$$

with the $T^k = \sigma^k/2$ denoting the three generators of the SU(2). The generators are proportional to 2×2 Pauli matrices σ_i. The $\mathrm{SU}(2)_L$ symmetry is introduced via the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + ig_W T^k W^k_\mu(x). \tag{2.7}$$

Here, the $W^k_\mu(x)$ are three new gauge fields. To retain gauge invariance, the fields W^k have to transform under a SU(2) as

$$W^k_\mu \to W'^k_\mu = W^k_\mu\partial_\mu\alpha^k - g_W f^{ijk}\alpha^i W^j_\mu, \tag{2.8}$$

with $f^{ijk} = \epsilon^{ijk}$ being the Levi-Civita tensor, the structure constants of the SU(2) group defined by the commutation relations $[T_i, T_j] = if_{ijk}T_k$. The gauge invariant form of the field strength tensor for the W fields can be written as

$$W^{\mu\nu}_i = \partial^\mu W^\nu_i - \partial^\nu W^\mu_i - g_W\epsilon_{ijk}W^\mu_j W^\nu_k. \tag{2.9}$$

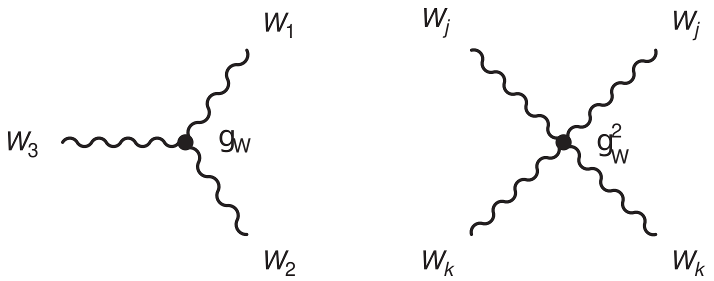


Figure 2.2: Triple and quartic $\mathrm{SU}(2)_L$ gauge boson self-interaction vertices. Taken from [33].

The corresponding kinetic term in the Lagrangian

$$\mathcal{L}_W = -\frac{1}{4} W_i^{\mu\nu} W^i_{\mu\nu}, \tag{2.10}$$

leads to self-interaction terms with triple and quartic gauge boson couplings depicted in Figure 2.2.

Since the generators of the SU(2) are the 2×2 Pauli matrices, the wave function $\varphi(x)$ has two components and is referred to as the weak isospin doublet. The $W^\pm$ bosons couple different left-handed fermion flavors with a charge difference of one elementary charge e.g. the first generation leptons

$$\varphi(x) = \begin{pmatrix} \nu_e(x) \\ e^-(x) \end{pmatrix}. \tag{2.11}$$

Here ν_e and e^- have a total weak isospin $I_W = 1/2$. The third component of the weak isospin is $I_W^3(\nu_e) = 1/2$ and $I_W^3(e^-) = -1/2$. Right handed fermions are unaffected by the $\mathrm{SU}(2)_L$ local gauge transformation. They are represented as weak isospin e_R singlets with $I_W = 0$. In the SM, there are no right-handed neutrinos. The isospin doublets are composed of the weak interaction eigenstates. The interaction term arising from the local gauge invariance is given by

$$i\frac{g_W}{2}\sigma^k \gamma_\mu W_k^\mu \varphi_L, \tag{2.12}$$

with φ_L denoting a weak isospin doublet of left-handed particles. The left-handed and right-handed fermion wave functions can be obtained with projection operators $P_L = (1-\gamma_5)/2$ and $P_R = (1+\gamma_5)/2$ as $\psi_L = P_L\psi$ and $\psi_R = P_R\psi$, respectively.

The weak four-vector current is defined as

$$j_\mu^k = \frac{g_W}{2}\overline{\varphi}_L \gamma_\mu \sigma^k \varphi_L. \tag{2.13}$$

With the weak isospin raising and lowering operators $\sigma_\pm = (\sigma_1 \pm i\sigma_2)/2$, the physical four-vector currents corresponding to the exchange of a $W^\pm$ boson are given by

$$j_\pm^\mu = \frac{g_W}{\sqrt{2}}\overline{\varphi}_L \gamma_\mu \sigma_\pm \varphi_L. \tag{2.14}$$

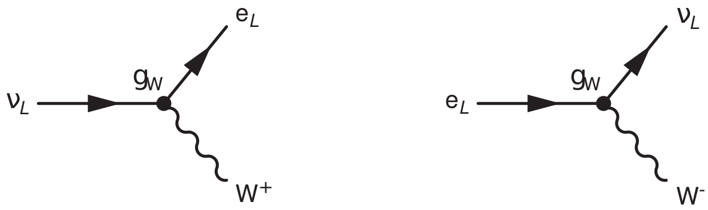


Figure 2.3: Weak charged current interaction vertices between e^- and ν_e weak eigenstates. Taken from [33].

With $\varphi_L = (\nu_L, f_l)$, the charged currents take the familiar form

$$j_+^\mu = \frac{g_W}{\sqrt{2}} \bar{\nu} \gamma_\mu \frac{1}{2}(1-\gamma^5) f \quad \text{and} \quad j_-^\mu = \frac{g_W}{\sqrt{2}} \bar{f} \gamma_\mu \frac{1}{2}(1-\gamma^5) \nu. \tag{2.15}$$

The physical $W^\pm$ bosons are identified with the linear combinations

$$W^{\pm}{}_{\mu} = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2). \tag{2.16}$$

Figure 2.3 depicts the weak charged current interaction between the e^- and ν_e weak eigenstates.

Consequently, the weak current can be written in the form

$$j_\mu^a W_a^\mu \equiv j_+^\mu W_\mu^+ + j_-^\mu W_\mu^- + j_3^\mu W_\mu^3, \tag{2.17}$$

with j_+^μ and j_-^μ being the two charged currents. There exists a weak neutral current with

$$\begin{aligned} j_\mu^3 &= \frac{g_W}{2} \bar{\varphi}_L \gamma_\mu \sigma_3 \varphi_L. \\ &= I_W^3(\nu) g_W \bar{\nu} \gamma_\mu \frac{1}{2}(1-\gamma^5)\nu + I_W^3(f) g_W \bar{f} \gamma_\mu \frac{1}{2}(1-\gamma^5) f. \end{aligned} \tag{2.18}$$

The experimental data shows that the Z boson couples to both left-handed and right-handed particles. Thus, the exchange of a Z boson cannot be directly associated to the neutral current defined in Equation (2.18) and the physical Z boson cannot be identified with the W^3 boson.

In the unification of electromagnetic and weak forces in the electroweak (EW) model of the local U(1) symmetry associated to electromagnetic interactions is replaced with the $\mathrm{U}(1)_Y$ local gauge symmetry. The local phase transformation is given by

$$\psi(x) \to \psi'(x) = e^{ig' \frac{Y}{2} \xi(x)} \psi(x), \tag{2.19}$$

and a new gauge field B_μ is introduced. The field B_μ couples to the weak hyper-charge Y with the coupling to fermions defined as

$$g' \frac{Y}{2} \gamma_\mu B^\mu \psi. \tag{2.20}$$

Here, the electric charge q_e is replaced with $g'Y/2$. In the unified EW theory, the photon and Z boson

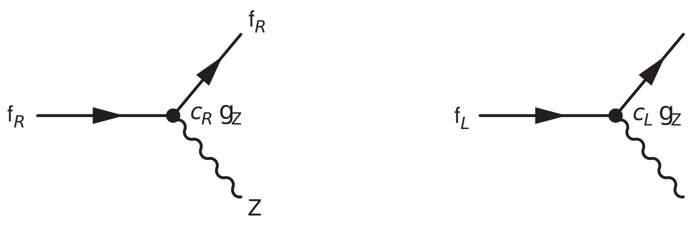


Figure 2.4: Weak neutral current interaction vertices for the coupling of the physical Z boson to chiral fermion states. Taken from [33].

are given by linear combinations of the fields B_μ and W^3_μ, defined as

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix}. \tag{2.21}$$

Here θ_W is the weak mixing angle. The angle can be measured in a number of ways, e.g. with studying the process $e^+e^- \to Z \to f\bar{f}$ and the current world average is $\sin^2\theta_W = 0.23122 \pm 0.00004$ [23]. The $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry implies that particles in a weak isospin doublet must have the same hyper-charge Y. The electromagnetic and weak coupling strengths are related via

$$e = g_W \sin\theta_W, \tag{2.22}$$

while the relation between g' and e is defined as

$$e = g' \cos\theta_W. \tag{2.23}$$

The value for the hyper-charge is given by

$$Y = 2(q - I^3_W), \tag{2.24}$$

with q denoting the electromagnetic charge in units of the elementary charge e. Since the couplings in the EW model are chosen to reproduce the electromagnetic couplings, the couplings of the Z boson to fermions are entirely specified. With the coupling to the physical Z boson defined as

$$g_Z = \frac{g_W}{\cos\theta_W} = \frac{e}{\sin\theta_W \cos\theta_w}, \tag{2.25}$$

the neutral Z current can be written as

$$j^\mu_Z = g_Z(c_L \bar{f}_L \gamma^\mu f_L + c_R \bar{f}_R \gamma^\mu f_R), \quad \text{with} \tag{2.26}$$

$$c_L = I^3_W(f) - q_f \sin^2\theta_W \quad \text{and} \tag{2.27}$$

$$c_R = -q_f \sin^2\theta_W. \tag{2.28}$$

The weak neutral current vertices for coupling of the physical Z boson to chiral fermion states is depicted in Figure 2.4.

Table 2.1: Charge, I_W^3, and weak hyper-charge assignments of the SM fermions. The couplings to the Z boson are calculated assuming $\sin^2\theta_W = 0.23146$. All values are taken from [33].

Fermion	Q	I_W^3	Y_L	Y_R	c_L	c_R	c_V	c_A
ν_e, ν_μ, ν_τ	0	$+\frac{1}{2}$	-1	0	$+\frac{1}{2}$	0	$+\frac{1}{2}$	$+\frac{1}{2}$
e^-, μ^-, τ^-	-1	$-\frac{1}{2}$	-1	-2	-0.27	$+0.23$	-0.04	$-\frac{1}{2}$
u, c, t	$+\frac{2}{3}$	$+\frac{1}{2}$	$+\frac{1}{3}$	$+\frac{4}{3}$	$+0.35$	-0.15	$+0.19$	$+\frac{1}{2}$
d, s, b	$-\frac{1}{3}$	$-\frac{1}{2}$	$+\frac{1}{3}$	$-\frac{2}{3}$	-0.42	$+0.08$	-0.35	$-\frac{1}{2}$

Alternatively, the neutral Z current can be expressed with an vector and axial-vector couplings as

$$j_Z^\mu = \frac{g_Z}{2}\bar{f}(c_V\gamma^\mu - c_A\gamma^\mu\gamma^5)f, \quad \text{with} \tag{2.29}$$

$$c_V = (c_L + c_R) = I_W^3(f) - 2q_f \sin^2\theta_W \quad \text{and} \tag{2.30}$$

$$c_A = (c_L - c_R) = I_W^3(f). \tag{2.31}$$

The interaction term between fermions and the Z boson can be written as

$$\frac{g_Z}{2}\gamma_\mu Z^\mu(c_V - c_A\gamma^5)\psi. \tag{2.32}$$

The fermion quantum numbers of the EW model and couplings to the Z bosons are listed in Table 2.1.

2.2.3 Gauge Boson and Fermion Mass Generation Mechanisms

Quadratic mass terms for the gauge bosons of the form

$$\frac{m_Z^2}{2}Z_\mu Z^\mu \tag{2.33}$$

break the local $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry of the EW Lagrangian. This is not a problem for QED and the strong interaction, since the photon and the gluons are massless bosons. However, the gauge bosons of the weak interaction are massive. Also the fermion mass terms in the QED Lagrangian of the form

$$-m_f\bar{f}f = -m_f(\bar{f}_R f_L + \bar{f}_L f_R) \tag{2.34}$$

are not invariant under the local $\mathrm{SU}(2)_L$ transformation with left-handed particles transforming as weak isospin doublets and right-handed particles as singlets.

In the SM, particle masses can be generated with the interaction with the Higgs field. The mass terms can be obtained from spontaneous symmetry breaking of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge symmetry of the EW Lagrangian. The minimal Higgs model embedded in the unified EW model is describe by a weak isospin doublet of two complex scalar fields

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.35}$$

The Higgs Lagrangian is given by

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - \frac{\lambda}{4} \left(\phi^\dagger \phi - \frac{v^2}{2} \right)^2. \tag{2.36}$$

The potential is minimized with $\phi^\dagger \phi = v^2/2$. For the photon to remain massless, the non-zero vacuum expectation value is chosen to be

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.37}$$

In the unitary gauge, the Higgs doublet can be written as

$$\phi(x) = \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.38}$$

Around the vacuum expectation value, the Higgs potential evaluates to

$$\frac{\lambda}{4}(h^2 + \sqrt{2} v h)^2. \tag{2.39}$$

Thus, the mass of the Higgs boson is given by

$$m_h = \sqrt{\frac{\lambda}{2}} v. \tag{2.40}$$

The gauge boson mass terms arise by writing the Higgs Lagrangian in a form that is invariant under local $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ transformations by replacing the derivatives with the appropriate covariant derivatives as

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i g_W T^k W^k_\mu(x) + i g' \frac{Y}{2} B_\mu. \tag{2.41}$$

By expanding the kinetic term, one obtains

$$\begin{aligned}(D_\mu \phi)^\dagger (D^\mu \phi) =& \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{8} g_W^2 (W^1_\mu + i W^2_\mu)(W^{1\mu} - i W^{2\mu})(v+h)^2 \\ &+ \frac{1}{8}(g_W W^3_\mu - g' B_\mu)(g_W W^{3\mu} - g' B^\mu)(v+h)^2.\end{aligned} \tag{2.42}$$

The terms quadratic in the gauge fields give the mass terms. For W^1 and W^2, the mass terms appear as

$$\frac{1}{2} m_W^2 W^1_\mu W^{1\mu} \quad \text{and} \quad \frac{1}{2} m_W^2 W^2_\mu W^{2\mu}, \tag{2.43}$$

and thus the mass of the W boson is defined by the EW coupling constant g_W and the vacuum expectation value of the Higgs field v

$$m_W = \frac{1}{2} g_W v. \tag{2.44}$$

The quadratic terms for the W^3 and B fields can be written as

$$\frac{v^2}{8}(g_W W^3_\mu - g' B_\mu)(g_W W^{3\mu} - g' B^\mu) = \frac{v^2}{8} \begin{pmatrix} W^3_\mu & B_\mu \end{pmatrix} \begin{pmatrix} g_W^2 & -g_W g' \\ -g_W g' & g'^2 \end{pmatrix} \begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix}. \tag{2.45}$$

The physical boson field propagate as independent mass-eigenstates. Thus the physical masses can be obtained by diagonalizing the mixing matrix, yielding

$$\frac{v^2}{8} \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & g'^2 + g_W^2 \end{pmatrix} \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix}. \tag{2.46}$$

The masses for the Z boson can be identified as

$$m_Z = \frac{1}{2} v \sqrt{g'^2 + g_W^2}, \tag{2.47}$$

and the photon remains massless. The physical gauge boson fields correspond to the normalized eigenvectors

$$A_\mu = \frac{g' W^3_\mu + g_W B_\mu}{\sqrt{g_W^2 + g'^2}}, \tag{2.48}$$

$$Z_\mu = \frac{g_W W^3_\mu - g' B_\mu}{\sqrt{g_W^2 + g'^2}}. \tag{2.49}$$

Writing the ratio of the gauge couplings of the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ as

$$\frac{g'}{g_W} = \tan\theta_W, \tag{2.50}$$

one obtains the relation from Equation (2.21). With Equation (2.50) , the masses of the Z and the $W^\pm$ bosons are related via the expression

$$m_Z = \frac{g_W v}{2\cos\theta_W} = \frac{m_W}{\cos\theta_W} \tag{2.51}$$

To generate mass terms for fermions in the standard model, a Yukawa coupling of left-handed $\mathrm{SU}(2)$ doublets L and right-handed $\mathrm{SU}(2)$ R singlets to the complex scalar Higgs doublet is introduced. The terms $\overline{L}\phi R$ and $(\overline{L}\phi R)^\dagger = \overline{R}\phi^\dagger L$ are invariant under $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ transformations. After spontaneous symmetry breaking and the Higgs doublet in the unitary gauge $\phi(x) = (0, v + h(x))/\sqrt{2}$, the coupling terms for electrons in the Lagrangian correspond to

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= -\frac{g_e}{\sqrt{2}}\left(\begin{pmatrix}\overline{\nu_e} & \overline{e}\end{pmatrix}_L \begin{pmatrix}0 \\ v+h\end{pmatrix} e_R + \overline{e}_R \begin{pmatrix}0 & v+h\end{pmatrix}\begin{pmatrix}\nu_e \\ e\end{pmatrix}_L\right) \\ &= -\frac{g_e}{\sqrt{2}} v(\overline{e}_L e_R + \overline{e}_R e_L) - \frac{g_e}{\sqrt{2}} h(\overline{e}_L e_R + \overline{e}_R e_L). \end{aligned} \tag{2.52}$$

Here, the electron mass can be identified with

$$m_e = \frac{g_e v}{\sqrt{2}}, \tag{2.53}$$

which arises through the coupling of massless left- and right-handed chiral fermions and the non-zero vacuum expectation of the Higgs field. With the coupling in Equation (2.55), masses can be generated for charged leptons and down-type quarks, while the neutrinos or up-type quarks remain massless. But the same mechanism can be used by introducing a coupling to the conjugate Higgs doublet

$$\phi_c = -i\sigma_2\phi^* = \begin{pmatrix}-\phi^{0*} \\ \phi^-\end{pmatrix}, \tag{2.54}$$

and the corresponding Lagrangian for the up-quark after symmetry breaking becomes

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= -\frac{g_u}{\sqrt{2}}\left(\begin{pmatrix}\overline{u} & \overline{d}\end{pmatrix}_L \begin{pmatrix}(v+h) \\ 0\end{pmatrix} u_R + \overline{u}_R \begin{pmatrix}(v+h) & 0\end{pmatrix}\begin{pmatrix}u \\ d\end{pmatrix}_L\right) \\ &= -\frac{g_u}{\sqrt{2}} v(\overline{u}_L u_R + \overline{u}_R u_L) - \frac{g_u}{\sqrt{2}} h(\overline{u}_L u_R + \overline{u}_R u_L). \end{aligned} \tag{2.55}$$

Again, the mass for the up-type quark is given by

$$m_u = \frac{g_u v}{\sqrt{2}}. \tag{2.56}$$

Generalizing the Yukawa coupling Lagrangian to more than one quark generation results in

$$\mathcal{L}_{\text{Yukawa}} = -g_d^{ij}\overline{Q}_L'^i \phi d_R'^j + g_u^{ij}\overline{Q}_L'^i \phi_c u_R'^j + \text{h.c.}, \tag{2.57}$$

with the weak SU(2) doublets and singlets for each quark generation $i = 1, 2, 3$

$$Q_L'^i = \begin{pmatrix}u'^i \\ d'^i\end{pmatrix}_L, \; u_R'^i, \text{ and } d_R'^i. \tag{2.58}$$

The prime indicates the fields as weak eigenstates that do not need to correspond to the physical mass eigenstates. Here, the generalized couplings g_d^{ij} and g_u^{ij} are hermitian matrices corresponding to the mass matrices after symmetry breaking $M_d^{ij} = vg_d^{ij}/\sqrt{2}$ and $M_u^{ij} = vg_u^{ij}/\sqrt{2}$. Separate unitary transformations in flavor space for left- and right-handed fields

$$\begin{aligned} u'^i_L &= A^{ij}_u u^j_L, \quad & u'^i_R &= B^{ij}_u u^j_R, \\ d'^i_L &= A^{ij}_d d^j_L, \quad & d'^i_R &= B^{ij}_d d^j_R, \end{aligned} \tag{2.59}$$

allow a diagonalization of the mass matrices

$$B^\dagger_u M_u A_u = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \quad \text{and} \quad B^\dagger_d M_d A_d = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix}. \tag{2.60}$$

The unprimed states correspond to the mass eigenstates of the quarks of the free-particle Lagrangian. Consequently, the weak interaction isospin doublet can be written as

$$\begin{pmatrix} u'_L \\ d'_L \end{pmatrix} = \begin{pmatrix} A_u u_L \\ A_d d_L \end{pmatrix} = A_u \begin{pmatrix} u_L \\ V d_L \end{pmatrix}, \tag{2.61}$$

with the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix V defined by

$$V = A^\dagger_u A_d. \tag{2.62}$$

Expressing the SM Lagrangian in quark mass eigenstates leaves the kinetic terms and couplings to the photon and Z boson fields unchanged. However the coupling to the $W^\pm$ bosons changes to

$$\frac{g_W}{\sqrt{2}} W^\pm{}_\mu \bar{u}'^i_L \gamma^\mu d'^i_L = \frac{g_W}{\sqrt{2}} W^\pm{}_\mu \bar{u}^i_L \gamma^\mu V^{ij} d^j_L. \tag{2.63}$$

This allows flavor-changing charged currents at tree level interactions.

2.2.4 The CKM Matrix

The CKM matrix is a 3 × 3 unitary matrix that relates the quark mass eigenstates to the eigenstate of the weak interaction

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.64}$$

and its elements define the relative strength of the charged current interaction. The transition rates related to flavor changing currents are proportional to $|V_{ij}|^2$. Overall, the CKM matrix can be parameterized with three mixing angles θ_{ij} and one complex phase δ, and a standard choice is [37]

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\mathrm{e}^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\mathrm{e}^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s13\mathrm{e}^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\mathrm{e}^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s13\mathrm{e}^{i\delta} & s_{23}c_{13} \end{pmatrix}, \tag{2.65}$$

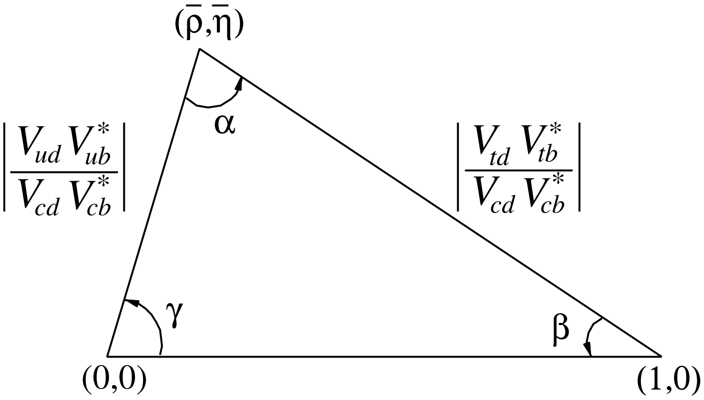


Figure 2.5: Sketch of the unitary triangle. Taken from [23]

with $s_{ij} = \sin\theta_{ij}$ and $c_i = \cos\theta_{ij}$. The complex phase δ is necessary to introduce CP-violating effects in the SM quark sector. Experiments show that $s_{13} \ll s_{23} \ll s_{12} \ll 1$, and thus this hierarchy can be expressed using the Wolfenstein parameterization [38] of the CKM matrix with four real parameters A, λ, η, and ρ as

$$V = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.66}$$

In this representation, the parameters are defined as

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \tag{2.67}$$

$$s_{23} = A\lambda^2 = \lambda\left|\frac{V_{cb}}{V_{us}}\right|, \quad \text{and} \tag{2.68}$$

$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho}+i\bar{\eta})\sqrt{1-A^2\lambda^4}}{\sqrt{1-\lambda^2}(1-A^2\lambda^4(\bar{\rho}+i\bar{\eta}))} = V_{ub}^*, \tag{2.69}$$

with $(\bar{\rho}+i\bar{\eta}) = -(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$ being independent of the phase convention.

The unitary condition of the CKM matrix imposes the relations $V_{ij}V_{ik}^* = \delta_{jk}$ and $V_{ij}V_{kj}^* = \delta_{ik}$. Here, the six equations equal to zero can be expressed geometrically as triangles in the complex plane. All triangles have the same area, which itself is a measure of the amount of CP-violation in the SM. The most commonly used relation is

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{2.70}$$

with each triangle side normalized to $V_{cd}V_{cb}^*$ the vertices are at exactly $(0,0)$, $(1,0)$, and at $(\bar{\rho},\bar{\eta})$. The triangle representation of Equation (2.70) is shown in Figure 2.5. Here, the angles of the unitary triangle are defined as

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \tag{2.71}$$

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \tag{2.72}$$

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{2.73}$$

Experimental measurements of branching fractions only allow for a determination of the magnitude of the CKM matrix elements. To determine the complex phase, measurements have to be carried out that are directly sensitive to the transition amplitudes instead of the amplitudes squared .

2.3 Heavy Quark Physics

2.3.1 Quantum Chromodynamics

Quantum Chromodynamics is a non-abelian gauge theory described by the SU(3) gauge symmetry group. The invariance under local SU(3) transformations

$$\psi \longrightarrow \psi'(x) = \mathrm{e}^{ig_s\alpha_k(x)T^k}\psi(x), \tag{2.74}$$

of the free Dirac Lagrangian is achieved with the covariant derivative defined as

$$\partial_\mu \longrightarrow D_\mu = \partial_\mu + ig_s T^k G_\mu^k(x). \tag{2.75}$$

Here, eight new gauge fields $G_\mu^k(x)$ are introduced. The generators of the SU(3) $T^k = \lambda^k/2$ are related to the 3×3 Gell-Mann matrices λ^k. The fields represent the gluons, the gauge bosons of the strong force. Like the photon field in the QED, the gluon fields are massless. Under a local SU(3) gauge transformation, the gluon fields G_μ^k transform as

$$G_\mu^k \longrightarrow G_\mu'^k = G_\mu^k \partial_\mu \alpha^k - g_s f^{ijk}\alpha^i G_\mu^j. \tag{2.76}$$

Here f^{ijk} denote the structure constants of the SU(3) defined by the commutation relations $[T_i, T_j] = if_{ijk}T_k$.

The quark-gluon interaction term is given by

$$-i\frac{g_s}{2}\lambda^k\gamma^\mu G_\mu^k\psi. \tag{2.77}$$

Compared to the single charge in QED, there are three conserved charges in QCD, commonly referred to as the color charges r, g, and b. Thus, quarks exist in three orthogonal color states. Since QCD is invariant under local SU(3) transformations in this color space, the strength of the strong force is independent of the color charge. The quark wave functions can be extended with a color part $\psi \longrightarrow c_i\psi$,

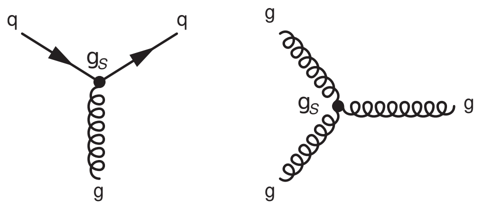


Figure 2.6: QCD interaction vertices arising from the requirement of local SU(3) gauge invariance. Taken from [33].

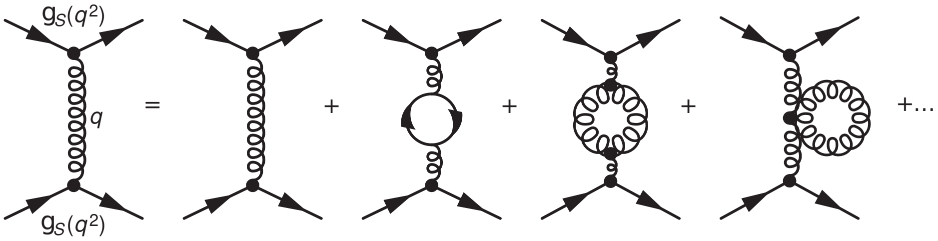


Figure 2.7: Higher-order corrections for the gluon propagator including virtual quark and gluon loops. Taken from [33].

represented as three orthogonal vectors

$$c_1 = r = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad c_2 = g = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad c_3 = b = \begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{2.78}$$

In this representation, the quark current associated with the quark-gluon vertex can be written as

$$j_q^\mu = \frac{g_s}{2}\overline{\psi} c_j^\dagger \lambda^a \gamma^\mu c_i \psi. \tag{2.79}$$

The left-most diagram in Figure 2.6 shows the quark-gluon interaction vertex.

Since the generators do not commute, additional self-interaction terms for the gluons arise like in the EW model. The kinetic term for the gluons in the QCD Lagrangian is given by

$$\begin{aligned} \mathcal{L}_G &= -\frac{1}{4} G_i^{\mu\nu} G_{\mu\nu}^i, \quad \text{with} \\ G_i^{\mu\nu} &= \partial^\mu G_i^\nu - \partial^\nu G_i^\mu - g_s f_{ijk} G_j^\mu G_k^\nu. \end{aligned} \tag{2.80}$$

Here, the last term proportional to $G_j^\mu G_k^\nu$ results in triple and quartic gluon vertices, depicted in the two right-most diagrams in Figure 2.6.

Despite the experimental evidence for the existence of quarks, there are no direct observations of free quarks. This absence of free quarks in nature is explained with the color confinement hypothesis. This hypothesis states that colored particles are always confined to color neutral or color singlet states. The origin of the color confinement is believed to be related to the gluon-gluon self-interactions, since they carry color charges themselves. Though, currently there exists no analytic proof of the concept. Consequently, quarks and gluons are always confined to colorless hadrons.

The QCD coupling strength g_s is not a constant quantity. It depends on the energy scale μ or momentum transfer q^2 of the considered interaction. The energy scale dependence of g_s is closely related to the concept of renormalization, especially the treatment of divergent higher-order corrections to the gluon propagator in the form of virtual quark and gluon loops. These corrections are schematically depicted in Figure 2.7. The energy scale dependence is controlled by the renormalization group equation

$$\beta(g_s) = \mu \frac{\mathrm{d}g_s}{\mathrm{d}\mu} \tag{2.81}$$

. The β function can be expressed as an perturbative series and the first non-vanishing term is given by

$$\beta(g_s) = -\frac{g_s^3}{16\pi^2}(11 - \frac{2}{3}N_q), \tag{2.82}$$

with N_q being the number of quark flavors. The solution of the renormalization group equation can be used to calculate the energy scale dependence of the fine structure constant $\alpha_s(\mu) = g_s^2(\mu)/4\pi$. With the first non-vanishing term of the QCD β function, the running of α_s is described by

$$\alpha_s(\mu) = \frac{12\pi}{(33 - N_q)\ln(\mu^2/\Lambda_{\mathrm{QCD}}^2)}. \tag{2.83}$$

Here, Λ_{QCD} denotes the energy scale where the QCD becomes strongly coupled and perturbation theory is no longer applicable. At low energy scales, $\mu \approx 1\,\mathrm{GeV}$, α_s is of $O(1)$ and perturbation theory is not applicable. This low-energy regime describes e.g. strong interactions in bound hadronic states. At high energy scales $\mu \approx 100\,\mathrm{GeV}$, one finds $\alpha_s \approx 0.1$, which is small enough to again employ perturbation theory. This feature of QCD is known as asymptotic freedom, allowing the treatment of quarks as free particles e.g. in the discussion of deep inelastic scattering or proton-proton collisions. Thus, QCD is divided into a low-energy non-perturbative and a high-energy regime where perturbation theory can be applied.

2.3.2 Heavy Quark Effective Theory

The typical size of a meson containing a heavy quark Q, like the charm or bottom quark with $m_Q \gg \Lambda_{\mathrm{QCD}}$, and a light quark $\overline{q}$ with $m_q \ll \Lambda_{\mathrm{QCD}}$ is of the order $O(1/\Lambda_{\mathrm{QCD}})$. In this system, the typical momentum transfer between the light and heavy quark due to non-perturbative QCD effects is of the order $O(\Lambda_{\mathrm{QCD}})$. Consequently, the velocity v of the heavy quark is almost unaffected by the QCD dynamics since $\Delta v = \Delta p/m_Q$ and in the limit $m_Q \to \infty$, the heavy quark velocity does not change. In this limit, the heavy quark acts a static color source and the meson dynamics reduces to interactions of the light quark with this static potential. Further, the heavy quark mass becomes completely irrelevant for the interaction between the light degrees of freedom and different heavy quark flavors within the $Q\overline{q}$ system. This symmetry, unchanged dynamics under exchange of heavy quark flavors, is known as heavy quark flavor symmetry. In addition, the static heavy quark only interacts with gluons via its chromomelectric charge which is spin independent leading to the heavy quark spin symmetry. Here the dynamics are unchanged under arbitrary transformations of the heavy quark spin.

The QCD Lagrangian for a heavy quark field Q

$$\mathcal{L}_Q = Q(i\gamma_\mu D^\mu - m_Q)Q \tag{2.84}$$

does not manifest the heavy quark spin and flavor symmetry for $m_Q \to \infty$. However, in the effective field theory formulation of QCD, known as heavy quark effective theory (HQET), the heavy quark symmetries become apparent in the limit $m_Q \to \infty$ [33–35]. The HQET Lagrangian is constructed in such a way, that only inverse powers of m_Q appear. In contrast, the full QCD Lagrangian contains only positive powers of m_Q.

The HQET provides a convenient description for the dynamics of hadrons containing a single heavy quark The momentum of a heavy quark inside a $Q\overline{q}$ system is always close to the on-shell four momentum $p^\mu = m_Q v^\mu$ with v^μ denoting the four-velocity of the heavy quark. Thus it can be written as

$$p^\mu = m_Q v^\mu + k^\mu, \tag{2.85}$$

with k^μ being the residual momentum obtained from non-perturbative interactions with the light degrees of freedom inside the $Q\overline{q}$ system. Typically, k^μ is of the order Λ_{QCD}.

In the heavy quark limit, the Dirac propagator of the heavy quark field simplifies to

$$i\frac{p_\mu\gamma^\mu + m_Q}{p^2 - m_Q^2 + i\epsilon} = i\frac{m_Q v_\mu\gamma^\mu + m_Q + k_\mu\gamma^\mu}{2m_Q v_\mu k^\mu + k^2 + i\epsilon} \overset{m_Q\to\infty}{=} i\frac{1 + v_\mu\gamma^\mu}{2v_\mu k^\mu + i\epsilon}. \tag{2.86}$$

Here, the propagator contains the velocity-dependent projection operator $(1 + v_\mu\gamma^\mu)/2$. In the rest-frame of the heavy quark, this operator projects onto the upper two particle components of the Dirac spinor.

To obtain the effective Lagrangian, it is convenient to write the heavy quark field as the sum of two velocity-dependent fields as

$$Q(x) = \mathrm{e}^{-im_Q v_\mu x^\mu}\left(Q_v(x) + \mathfrak{Q}_v(x)\right), \tag{2.87}$$

with

$$Q_v(x) = \mathrm{e}^{im_Q v_\mu x^\mu}\frac{1 + v_\mu\gamma^\mu}{2}Q(x) \quad \text{and} \quad \mathfrak{Q}_v(x) = \mathrm{e}^{im_Q v_\mu x^\mu}\frac{1 - v_\mu\gamma^\mu}{2}Q(x). \tag{2.88}$$

The exponential pre-factor subtracts the on-shell momentum from the heavy quark momentum. In the rest frame with $v^\mu = (1, 0, 0, 0)$, Q_v and $\mathfrak{Q}_v$ correspond to the upper and lower two components of the quark field Q. The former annihilates a heavy quark with velocity v, while the latter creates a heavy anti-quark with velocity v. Contributions from the quark field $\mathfrak{Q}_v$ are suppressed with $O(1/m_Q)$ and only the field Q_v produces effects at leading order

$$\frac{1 + v_\mu\gamma^\mu}{2}Q(x) = Q(x) + O\left(\frac{1}{m_Q}\right) \quad \text{and} \quad \frac{1 - v_\mu\gamma^\mu}{2}Q(x) = 0 + O\left(\frac{1}{m_Q}\right). \tag{2.89}$$

By substituting the expression for Q in Equation (2.87) into the QCD Lagrangian in Equation (2.84), one obtains

$$\mathcal{L} = \overline{Q}_v (i v_\mu D^\mu) Q_v - \overline{\mathfrak{Q}}_v (i v_\mu D^\mu + 2m_Q) \mathfrak{Q}_v + \overline{Q}_v (iD_\perp^\mu \gamma_\mu) \mathfrak{Q}_v + \overline{\mathfrak{Q}}_v (iD_\perp^\mu \gamma_\mu) Q_v, \tag{2.90}$$

with $D_\perp^\mu = D^\mu - (v_\nu D^\nu) v^\mu$. $D_\perp^\mu$ is orthogonal to D^μ with respect to the heavy quark velocity v^μ. $\overline{Q}_v$ describes massless degrees of freedom, while $\mathfrak{Q}_v$ corresponds to excitations with two times the heavy quarks mass. Using the variation of the Lagrangian with respect to $\overline{\mathfrak{Q}}_v$

$$(i v_\mu D^\mu + 2m_Q) \mathfrak{Q}_v = iD_\perp^\mu \gamma_\mu Q_v, \tag{2.91}$$

these heavy degrees of freedom can be eliminated. With this relation, the HQET Lagrangian takes the form

$$\begin{aligned}
\mathcal{L}_{\text{HQET}} &= \overline{Q}_v (i v_\mu D^\mu) Q_v + \overline{Q}_v iD_\perp^\mu \gamma_\mu \frac{1}{2m_Q + iD_\nu v^\nu} iD_\perp^\kappa \gamma_\kappa Q_v \\
&= \overline{Q}_v (i v_\mu D^\mu) Q_v + \frac{1}{2m_Q} \sum_{n=0}^{\infty} \overline{Q}_v iD_\perp^\mu \gamma_\mu \left(-i \frac{v_\nu D^\nu}{2m_Q} \right)^n iD_\perp^\kappa \gamma_\kappa Q_v \\
&= \overline{Q}_v (i v_\mu D^\mu) Q_v - \frac{1}{2m_Q} \overline{Q}_v (D_\perp^2 + \frac{g}{2} \sigma_{\mu\nu} G^{\mu\nu}) Q_v + O\left(\frac{1}{m_Q^2} \right).
\end{aligned} \tag{2.92}$$

In the limit $m_Q \rightarrow \infty$, the effective Lagrangian does not depend on the mass or the spin of the heavy quark and thus has a manifest heavy quark spin-flavor symmetry. Non-perturbative correction terms breaking the spin-flavor symmetry are suppressed with $O(1/m_Q)$. At order $1/m_Q$ in the expansion, the first term can be identified with the gauge-covariant extension of the kinetic energy, while the second term describes the color-magnetic coupling of the heavy quark spin to the gluon field. The former explicitly depends on the heavy quark mass an thus breaks the heavy quark flavor symmetry, while the second term breaks both heavy quark flavor and spin symmetries.

The choice of the residual momentum k contributing to the heavy quark momentum $p_Q = m_Q v + k$ is not unique. A small change in v of order Λ_{QCD}/m_Q can be compensated by a shift in the residual momentum k

$$v \rightarrow v + \frac{\epsilon}{m_Q} \quad \text{and} \quad k \rightarrow k - \epsilon. \tag{2.93}$$

To satisfy $v^2 = 1$, it must hold $v_\nu \epsilon^\nu = 0$. Consequently, the heavy quark spinor Q_v has to transform as

$$Q_v \rightarrow Q_v + \delta Q_v = \left(1 + \frac{\epsilon_\mu \gamma^\mu}{2m_Q} \right) Q_v, \tag{2.94}$$

to retain the constraint $v_\mu \gamma^\mu Q_v = Q_v$. The HQET Lagrangian is invariant under the combined transformation

$$v \rightarrow v + \frac{\epsilon}{m_Q} \quad \text{and} \quad Q_v \rightarrow \mathrm{e}^{i \epsilon_\mu x^\mu} \left(1 + \frac{\epsilon_\mu \gamma^\mu}{2m_Q} \right) Q_v, \tag{2.95}$$

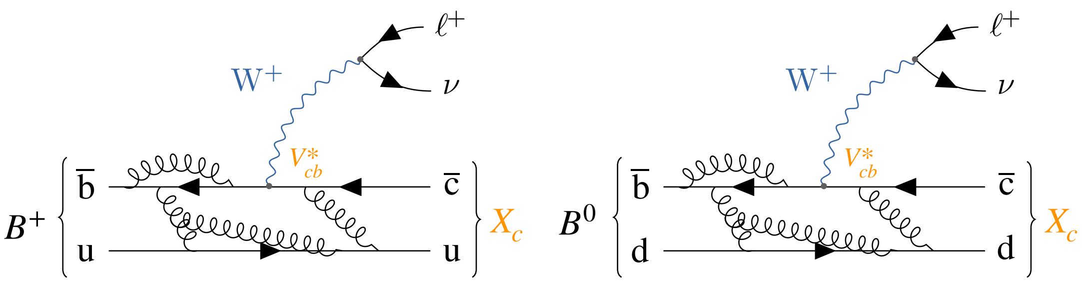


Figure 2.8: Feynman diagram for the semileptonic $B \longrightarrow X_c \ell \nu_\ell$ decay.

with the pre-factor $\mathrm{e}^{i\epsilon_\mu x^\mu}$ causing the shift in the residual momentum $k \longrightarrow k - \epsilon$. In fact, the first two terms in Equation (2.92) change as

$$\begin{aligned} \mathcal{L}_0 &= \overline{Q}_v(iv_\mu D^\mu)Q_v \longrightarrow \mathcal{L}_0 + \frac{1}{m_Q}\overline{Q}_v(i\epsilon_\mu D^\mu)Q_V, \quad \text{and} \\ \mathcal{L}_1 &= -\frac{1}{2m_Q}\overline{Q}_v(D_\perp^2 + \frac{g}{2}\sigma_{\mu\nu}G^{\mu\nu})Q_v \longrightarrow \mathcal{L}_1 - \frac{1}{m_Q}\overline{Q}_v(i\epsilon_\mu D^\mu)Q_V. \end{aligned} \tag{2.96}$$

As a result, the sum of both Lagrangians $\mathcal{L}_0 + \mathcal{L}_1$ is invariant under reparameterization. Since the transformation in Equation (2.95) involves m_Q, reparameterization invariance connects different orders $1/m_Q$ in the expansion of the HQET Lagrangian [39].

2.3.3 Inclusive $B \longrightarrow X_c \ell \nu_\ell$ Decays

The semileptonic decay $B \longrightarrow X_c \ell \nu_\ell$, with X_c denoting any hadronic system containing a charm quark, is described by the effective Hamiltonian

$$H_W = \frac{4G_F}{\sqrt{2}} V_{cb}(\bar{c}\gamma^\mu P_L b)(\bar{\ell}\gamma_\mu P_L \nu) = \frac{4G_F}{\sqrt{2}} V_{cb} J_q^\mu J_{\ell\mu} + \text{h.c.}. \tag{2.97}$$

Here, $G_F/\sqrt{2} = g_W^2/8m_W^2$ denotes the Fermi constant, the effective coupling strength in weak decays. The Hamiltonian directly depends on the CKM matrix element V_{cb}. Since leptons do not couple to gluons, the leptonic and hadronic currents J_ℓ^μ amd J_q^μ factorize. The hadronic current also involves non-perturbative strong interactions from the heavy quarks interacting with the light degrees of freedom inside the hadrons.

The Feynman diagrams for the B^+ and B^0 decay are depicted in Figure 2.8. Here, the final-state consists of a lepton-neutrino pair and the hadronic X_c system. The latter can be a single particle state like a $D^{(*,**)}$ meson or a multi-particle state like $D\pi\pi$.

Usually, one chooses $q^2 = (p_\ell + p_{\nu_\ell})^2$ which corresponds to mass of the virtual W boson.

In the B rest frame, the differential decay rate is given by

$$\begin{aligned}\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2\mathrm{d}E_\ell\mathrm{d}E_{\nu_\ell}} &= \frac{1}{4}\sum_{X_c}\sum_{\substack{\text{lepton}\\\text{spins}}}\frac{\left|\langle X_c\ell\nu_\ell|H_W|B\rangle\right|^2}{2}(2\pi)^3\delta^4(p_B - \underbrace{(p_\ell + p_{\nu_\ell})}_{=q^2} - p_{X_c}) \\ &= 2G_F^2|V_{cb}|^2 W_{\mu\nu}L^{\mu\nu}. \end{aligned} \tag{2.98}$$

The weak matrix element factorizes into a leptonic and a hadronic matrix element. The spin and hadronic state averaged result is usually written in terms of the leptonic and hadronic tensors, defined as

$$L^{\mu\nu} = 2(p_\ell^\mu p_{\nu_\ell}^\nu + p_\ell^\nu p_{\nu_\ell}^\mu - g^{\mu\nu} p_{\ell\alpha} p_{\nu_\ell}^\alpha - i\epsilon^{\eta\nu\lambda\mu} p_{\ell\eta} p_{\nu_\ell\lambda}), \tag{2.99}$$

and

$$W^{\mu\nu} = \sum_{X_c}\frac{(2\pi)^3}{2m_B}\delta^4(p_B - q^2 - p_{X_c})\langle\bar{B}(p_B)|J_q^{\dagger\mu}|X_c(p_{X_c})\rangle\langle X_c(p_{X_c})|J_q^\nu|\bar{B}(p_B)\rangle. \tag{2.100}$$

The hadronic tensor incorporates all strong interactions relevant in the decay $B \longrightarrow X_c\ell\nu_\ell$ and it explicitly depends on the B meson momentum and the momentum transfer to the hadronic system q^2.

With the optical theorem, $W^{\mu\nu}$ can be related to the imaginary part of the forward scattering amplitude

$$W^{\mu\nu} = -\frac{1}{\pi}\mathrm{Im}\,T^{\mu\nu} = -\frac{1}{\pi}\mathrm{Im}\int \mathrm{d}^4x\mathrm{e}^{-iq_\alpha x^\alpha}\frac{\langle\bar{B}|T[J_q^{\dagger\mu}(x)J_q^\nu(0)]|\bar{B}\rangle}{2m_B}. \tag{2.101}$$

With the Operator Product Expansion, the time ordered product can be expanded in terms of local operators and coefficients that dependent on the decay kinematics

$$\int \mathrm{d}^4x\mathrm{e}^{-iq_\alpha x^\alpha}T[J_q^{\dagger\mu}(x)J_q^\nu(0)] = \sum_{n,i}\frac{1}{m_b^n}C_{n,i}O_{n+3,i}. \tag{2.102}$$

Here $O_{l,i}$ denotes a set (labeled by i) of operators with mass dimension l. The coefficient $C_{n,i}$ can be reliably calculated in QCD perturbation theory (with respect to α_s). With the forward matrix element of this expression, it is possible to calculate the decay rate in terms of Wilson coefficients and matrix elements describing non-perturbative QCD effects. This is known as the Heavy Quark Expansion (HQE). Thus the total decay rate can be expressed as an expansion in $1/m_b$ as

$$\Gamma = \Gamma_0 + \frac{1}{m_b}\Gamma_1 + \frac{1}{m_b^2}\Gamma_2 + \frac{1}{m_b^3}\Gamma_3 + \cdots. \tag{2.103}$$

The non-perturbative matrix elements are commonly referred to as HQE parameters. Usually, they take the form of

$$\langle B(p_B)|\bar{b}_v(iD\mu_1)(iD_2^\mu)\dots b_v(iD\mu_n)\Gamma b_v|B(p_B)\rangle, \tag{2.104}$$

with Γ being some Dirac matrix. At $n = 0$, the operators have dimension three and the operators can be related to the matrix element

$$\langle B(p_B)|\bar{b}\gamma_\mu b|B(p_B)\rangle, \tag{2.105}$$

which is normalized in full QCD and the Wilson coefficient C_0 denotes the decay rate of a free b quark. There are no contributions of dimension four operators, since it is possible to express these operators in terms of dimension five operators using the equation of motion for the b quark. At $n = 2$, two non-perturbative contributions are present. These matrix elements can be expressed as

$$\begin{aligned} 2m_B\mu_\pi^2 &= \langle B(p_B)|\bar{b}_v(iD)^2 b_v|B(p_B)\rangle, \\ 2m_B\mu_G^2 &= \langle B(p_B)|\bar{b}_v\sigma_{\mu\nu}(iD^\mu)(iD^\nu)b_v|B(p_B)\rangle. \end{aligned} \tag{2.106}$$

Here, b_v denotes the heavy quark field Q_v introduced in Equation (2.88). Analogous to the first $1/m_B$ term in the HQET Lagrangian, the matrix elements μ_π^2 and μ_G^2 correspond to the kinetic and chromo-magnetic energies of the heavy quark inside the B meson. At dimension six ($n = 3$), again two matrix elements contribute

$$\begin{aligned} 2m_B\rho_D^3 &= \langle B(p_B)|\bar{b}_v(iD_\mu)(iv_\alpha D^\alpha)(iD^\mu)b_v|B(p_B)\rangle, \\ 2m_B\rho_{LS}^3 &= \langle B(p_B)|\bar{b}_v\sigma_{\mu\nu}(iD^\mu)(iv_\alpha D^\alpha)(iD^\nu)b_v|B(p_B)\rangle, \end{aligned} \tag{2.107}$$

referred to as the Darwin and spin-orbit terms. For higher orders ($n \geq 4$), the number of non-perturbative matrix elements proliferates.

Similar to the total rate Γ, spectral moments $\langle M^k\rangle$ of an observable M in inclusive decays can be calculated via phase space integration of the differential rate multiplied by a weight function $w(v, p_\ell, p_{\nu_\ell})$ with $v = p_B/m_B$ as

$$\langle M^k[w]\rangle = \int \mathrm{d}\Phi w^k(v, p_\ell, p_{\nu_\ell})W^{\mu\nu}L_{\mu\nu}. \tag{2.108}$$

To calculate hadronic mass m_X and lepton energy E_ℓ moments, the weight function is set to $w = (m_B v - q)^2$ and $w = v_\mu p_\ell^\mu$, respectively. For the calculation of lepton mass squared q^2 moments, the weight function corresponds to $w = (p_\ell + p_{\nu_\ell})^2 = q^2$. Analogous to the total rate, the expression for the spectral moments can be written as an expansion in $1/m_b$ using the HQE

$$\langle M^k\rangle = \langle M_0^k\rangle + \frac{1}{m_b}\langle M_1^k\rangle + \frac{1}{m_b^2}\langle M_2^k\rangle + \frac{1}{m_b^3}\langle M_3^k\rangle + \cdots, \tag{2.109}$$

with the same set of HQE parameters.

2.4 Experimental Status of $|V_{cb}|$

This subsection summarizes the current status of experimental determinations of the CKM matrix element $|V_{cb}|$. All numerical values are taken from the review in [40] unless stated otherwise.

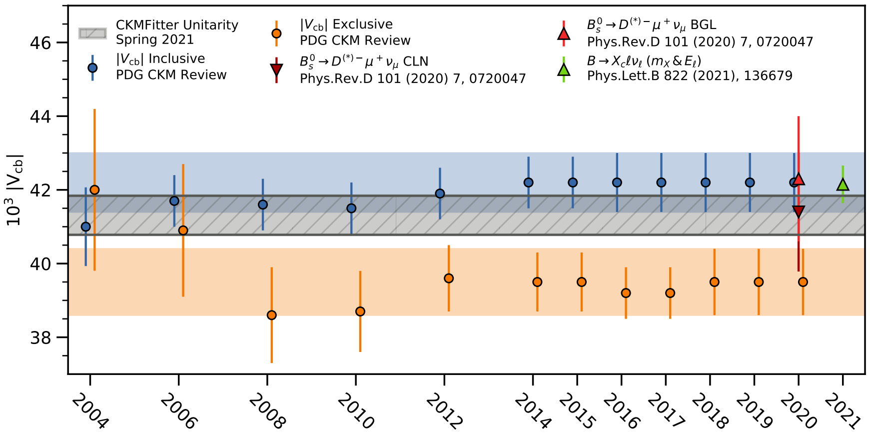


Figure 2.9: Inclusive and exclusive $|V_{cb}|$ measurements over time. The values correspond to the averages reported by the PDG [23]. In addition, the average result between exclusive and inclusive determinations of $|V_{cb}|$ reported by the CKMFitter group [41] is shown for comparison.

Precise determinations of $|V_{cb}|$ are essential to test the CKM sector of the SM. $|V_{cb}|$ directly affects the length of the unitary triangle side between the apex and the origin and thus, the position of the former. $|V_{cb}|$ can be determined from semileptonic B meson decays via CKM favored $b \to c\ell\nu_\ell$ transitions at tree-level. One can distinguish between $|V_{cb}|$ determinations using exclusive decays modes, e.g. the decays $B \to D^{(*)}\ell\nu_\ell$, and determinations based on the inclusive decay $B \to X_c\ell\nu_\ell$. In the latter, there is no separation of specific final states and the sum of all possible hadronic systems are considered at once. Both approaches are based on analyzing the differential or partial decay rates which are sensitive to $|V_{cb}|^2$.

The differential decay rates for exclusive $B \to D^{(*)}\ell\nu_\ell$ decays are described with respect to the recoil variable

$$w = v_{B\mu} v^{\mu}_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{m_B m_{D^{(*)}}}, \tag{2.110}$$

the product of the four-velocities of the B and $D^{(*)}$ mesons. The point $w = 1$ corresponds to the maximum momentum transfer to the lepton-neutrino system. For $q^2 = 0$, the w spectrum has a distinct kinematically allowed endpoint.

Assuming massless leptons, the differential decay rate for $B \to D^*\ell\nu_\ell$ decays is proportional to

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}w} \propto |V_{cb}|^2 \eta_{\mathrm{EW}}^2 \chi(w) \mathcal{F}^2(w), \tag{2.111}$$

with $\chi(w)$ is a known phase space factor, and η_{EW} is a small electroweak correction [42]. The combination of $\chi(w)$ and the form factor $\mathcal{F}^2(w)$ can be described by three independent functions $h_{A_1}(w)$, $R_1(w)$, and $R_2(w)$. In the CLN form factor parameterization [43], these functions are parameterized by the four parameters $h_{A_1}(1)$, ρ^2, $R_1(1)$, and $R_2(1)$. Measurements usually provide a result for the combination $\eta_{\mathrm{EW}}\mathcal{F}(1)|V_{cb}|$. With external input from lattice QCD for $\mathcal{F}(1)$ [44–46], one can translate the result into a value for $|V_{cb}|$. The current HLFAV average states

$$|V_{cb}| = (38.76 \pm 0.42|_{\mathrm{exp}} \pm 0.55|_{\mathrm{th}}) \times 10^{-3}. \tag{2.112}$$

Here, the first and second uncertainties denote experimental and theoretical uncertainties on $|V_{cb}|$.

In addition to the CLN parameterization, a more general prescription for the form factors is provided by the BGL parameterization [25, 47]. There are published analyses by Belle [48] and BaBar [49] that extract $|V_{cb}|$ using BGL. Due to the limited set of measurements, no combination of $|V_{cb}|$ is performed by HFLAV. Though, the results from both analyses are consistent with the $|V_{cb}|$ average based on the CLN parameterization.

Similar to $B \to D^{*}\ell\nu_{\ell}$ decays, the differential decay rate for $B \to D\ell\nu_{\ell}$ is proportional to

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}w} \propto |V_{cb}|^2 \eta_{\mathrm{EW}}^2 (w^2 - 1)^{3/2} \mathcal{G}^2(w). \tag{2.113}$$

Here, the form factor $\mathcal{G}(w)$ is described by a single function $f_{+}(w)$. In the CLN parameterization, this function is parameterized by the two parameters $\mathcal{G}(1)$ and ρ^2 . With lattice input for $\mathcal{G}(1)$ [50], $|V_{cb}|$ is determined from the form factor fit results for $\eta_{\mathrm{EW}}\mathcal{G}(1)|V_{cb}|$. The current HFLAV report states

$$|V_{cb}| = (39.58 \pm 0.94|_{\mathrm{exp}} \pm 0.37|_{\mathrm{th}}) \times 10^{-3} \tag{2.114}$$

as their current average for $|V_{cb}|$ from $B \to D\ell\nu_{\ell}$ decays.

Also, no combination of measurements based on the BGL parameterization is reported for $B \to D\ell\nu_{\ell}$. The differential w spectrum is measured by Belle [51] and BaBar [52], but the latter result does not include the full covariance matrix necessary to combine the results.

In addition to studies of exclusive semileptonic B meson decays, semileptonic decays of B_s^0 are studied at the LHCb experiment. A recent analysis of $B_s^0 \to D_s^{(*)-}\mu^{+}\nu_{\mu}$ performed by LHCb [53] extract values for $|V_{cb}|$ based on CLN and BGL parameterization for the hadronic form factors. The results $|V_{cb}| = (41.4 \pm 1.6)$ (CLN) and $|V_{cb}| = (42.3 \pm 1.7)$ (BGL) ar in agreement with $|V_{cb}|$ values obtained by B^0 and B^{+} decays.

The determination of $|V_{cb}|$ from inclusive $B \to X_c\ell\nu_{\ell}$ decays is based on the HQE calculation for the partial rate and observables like lepton energy and hadronic mass moments. Calculations using different definitions for b quark mass and other theoretical quantities based on the kinetic [28, 29, 54] and 1S [55] renormalization schemes are available. Different sets of HQE parameters are used to describe the inclusive observables in the two renormalization schemes, allowing no easy comparison of numerical results between both sets.

$|V_{cb}|$ and the HQE parameters are determined in a global fit to measurements of the partial rate and spectral moments measured as functions of lower thresholds on the lepton energy. The analysis based on the kinetic scheme (following the procedure in [14]) obtains

$$|V_{cb}| = (42.19 \pm 0.78) \times 10^{-3}, \tag{2.115}$$

while the fit based on the predictions in the 1S scheme (following the procedure in [56]) results in

$$|V_{cb}| = (41.98 \pm 0.45) \times 10^{-3}. \tag{2.116}$$

The current determinations are able extract to $|V_{cb}|$ and HQE parameters including HQE terms up to $\mathcal{O}(1/m_b^3)$ in a model independent way. As stated before, higher orders in the HQE lead to a proliferation of non-perturbative matrix elements further complicating the extraction from data. The size of these higher order terms can be estimated with the lowest-lying state saturation approximation [11]. The fit in [57] includes the estimates for $\mathcal{O}(1/m_b^{(4,5)})$ HQE parameters in the global fit using Gaussian priors. This results in small reduction of $|V_{cb}|$ of 0.25% compared to the fit without the inclusion higher order $1/m_b^4$ and $1/m_b^5$ terms in the HQE [15].

A recent analysis in [58] reports the current most precise value for $|V_{cb}|$

$$|V_{cb}| = (42.16 \pm 0.51) \times 10^{-3}, \tag{2.117}$$

including $\mathcal{O}(\alpha_s^3)$ corrections for calculations of the total width [59–61]. This study updates the results from [15] and the inclusion of the higher order α_s^3 corrections for the total width lead to a reduction of the uncertainty of $|V_{cb}|$.

A summary of different $|V_{cb}|$ measurements over time is shown in Figure 2.9. In this figure, the given values show the world averages reported by the PDG [23] over time. The latest PDG world average for $|V_{cb}|$ determined using exclusive approaches is

$$|V_{cb}|_{\text{excl}} = (39.5 \pm 0.9) \times 10^{-3}. \tag{2.118}$$

For inclusive determinations of $|V_{cb}|$, the current world average is

$$|V_{cb}|_{\text{incl}} = (42.2 \pm 0.8) \times 10^{-3}. \tag{2.119}$$

The observed $\sim 3\sigma$ tension between $|V_{cb}|_{\text{excl}}$ and $|V_{cb}|_{\text{incl}}$ is a longstanding discrepancy also known as the V_{cb} puzzle [62].

CHAPTER 3

Experimental Setup

This work analyzes data collected with the Belle II detector located at the SuperKEKB B-factory at the High Energy Accelerator Research Organization in Tsukuba, Japan. Belle II is the successor of the Belle detector, which observed e.g. time-dependent and direct CP violation in B decays. The main physics goals of Belle II are the search for new physics at the intensity frontier and the measurement of the CKM parameters with improved precision. In addition, there is also a competitive non-flavour physics program ongoing with indirect searches for dark matter and meson spectroscopy [63].

To reach these goals, both the accelerator and the detector received substantial upgrades, which will be discussed in Sections 3.1 and 3.2, respectively. Section 3.3 gives a short overview of the final-state particle reconstruction algorithms at Belle II. Lastly, Section 3.4 describes the analyzed data samples.

3.1 The SuperKEKB e^+e^--Collider

This section gives an overview of the SuperKEKB e^+e^--collider following the description given in [64, 65]. The SuperKEKB is a e^+e^--collider located at KEK, the High Energy Accelerator Research Organization, in Tsukuba, Japan. SuperKEKB is the successor and upgrade of the KEKB B-factory [66].

The accelerator is mainly operating at a center-of-mass (CM) energy of $\sqrt{s} = 10.58\,\text{GeV}$, corresponding to the $\Upsilon(4S)$ mass. The $\Upsilon(4S)$ is an excited $b\bar{b}$ state and decays almost exclusively into a pair of B mesons. In the CM frame, both B mesons are produced almost at rest. The electron and positron beams have asymmetric beam energies of 7 GeV and 4 GeV, respectively. This results in a Lorentz boost $\beta\gamma = 0.28$ of the CM system with respect to the laboratory frame. The boosted CM system is necessary to achieve a spatial separation between the short-lived B meson pair. Otherwise, it would not be possible to measure the distance of the B meson decay vertex with respect to the interaction point (IP). These measurements are essential for precision measurements of CP asymmetries in B decays, one of the major physics goals of the Belle II experiment.

A schematic overview of the SuperKEKB accelerator complex is shown in Figure 3.1. It consists of an injector linear accelerator with a 1 GeV positron damping ring and the 7 GeV electron and 4 GeV positron ring. Low-emittance and high-current electron bunches are produced with a photo-cathode radio-frequency gun in the pre-injector at the beginning of the linear accelerator. The electron beam is accelerated up to 7 GeV and is then injected into the high-energy electron ring. Positron bunches are produced in the collision of electrons with a tungsten target. These electrons are produced by the old

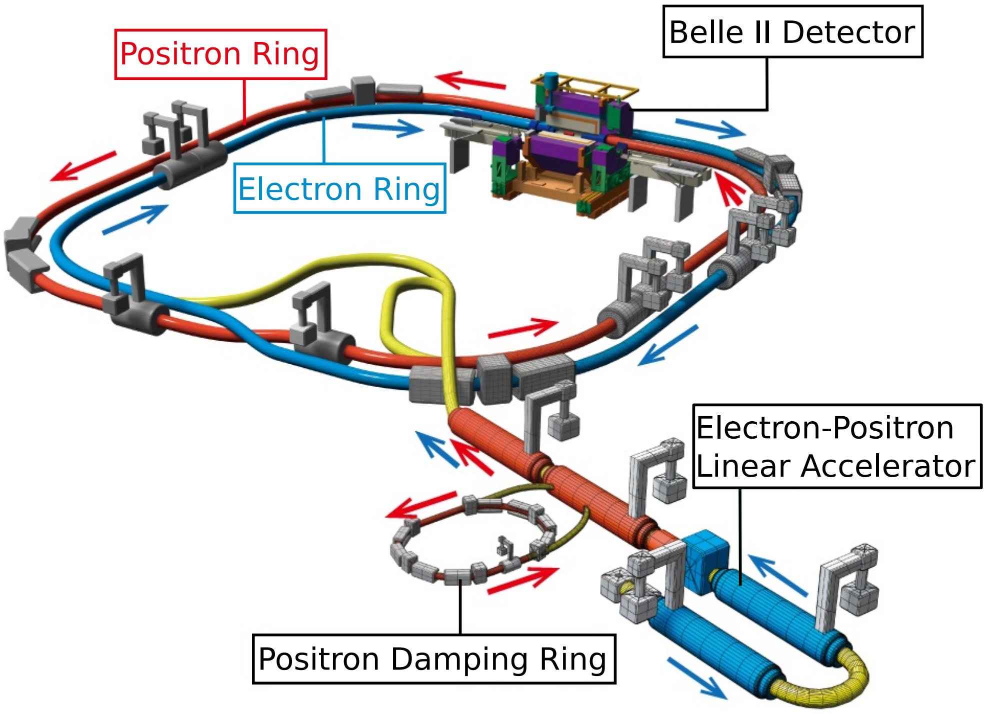


Figure 3.1: Schematic overview of the SuperKEKB accelerator complex. Adapted from [64].

KEKB pre-injector. The damping ring is used to reduce the emittance of the positron beam and its energy spread. Then, the positron bunches are re-injected into the linear accelerator and is accelerated up to 4 GeV in the second half of the latter before they are injected into the low-energy positron ring.

The design luminosity of the SuperKEKB accelerator is $8 \times 10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is 40 times than the luminosity of its predecessor KEKB. The luminosity increase is based on the nano-beam collision scheme [67]. This scheme proposes to squeeze the vertical beta function at the IP β_y^* by minimizing the effective longitudinal beam overlap at the IP. The effective overlap region characterized by a large crossing angle ($\phi = 41.5\,\mathrm{mrad}$) and a small horizontal beam size ($\sigma_x^* = 7.75\,\mu\mathrm{m}$ for the electron beam). This results in a small vertical beam size ($\sigma_y^* = 59\,\mathrm{nm}$). Overall, β_y^* is reduced by a factor of 20 at SuperKEKB compared to KEKB.

SuperKEKB holds the world record for peak luminosity with $3.1 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [68]. The total recorded luminosity by the Belle II experiment is approximately $400\,\mathrm{fb}^{-1}$.

3.2 The Belle II Detector

The Belle II detector designed as a general-purpose 4π detector with multiple detector sub-systems arranged in a cylindrical structure around IP. It is used for measurements of particle energies and momenta in the range of $\mathcal{O}(10\,\mathrm{MeV})$ up to a few GeV as well as identification of charged particles. This summary of the Belle II detector follows roughly the detailed presentation given in [65].

The detector presents a substantial upgrade to the Belle detector [69]. The upgrade allows the Belle II detector to maintain the same performance as the Belle detector with at the same time highly

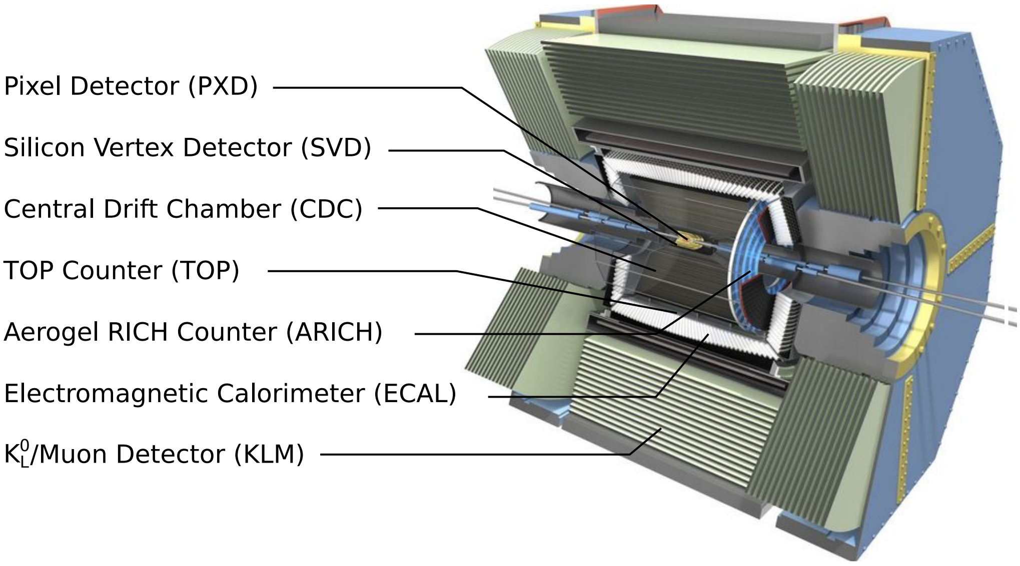


Figure 3.2: Schematic overview of the Belle II detector. Adapted from [70].

increased radiation background levels. Due to the new innermost pixel-detector, Belle II provides a better vertex resolution. Since the strip silicon vertex detector occupies a larger volume, the efficiency for reconstructing K_S^0 is also improved. Further, new particle identification detectors in the barrel and end-cap region extend the good K/π separation to the geometrical acceptance limits of the detector. Lastly, the noise pile up in the electromagnetic calorimeter is reduced by new readout electronics.

The origin of the Belle II coordinate system is defined at the nominal IP. The z-axis of the laboratory frame is defined as the central axis of the solenoid with the positive direction in direction of the electron beam. The y and x axes point vertically upwards in the detector and horizontally away from the center of the SuperKEKB rings, respectively. Due to the cylindrical detector structure, it is convenient to use a cylindrical coordinate system. Here, the polar angle θ and the longitudinal and transverse directions are defined with respect to the z-axis.

A schematic overview of the Belle II detector is shown in Figure 3.2. It consists of a tracking system including the pixel detector (PXD), a double-sided silicon vertex detector (SVD), and the central drift chamber (CDC). Outside of the CDC, a time-of-propagation counter (TOP) and aerogel ring-imaging Cherenkov detector (ARICH) cover the barrel and forward end-cap regions, respectively. Both detectors are important for the identification of charged particles. The electromagnetic calorimeter (ECL) is located outside the TOP and inside the remaining volume of a superconducting solenoid with a field strength of 1.5 T. The K_L^0 and muon detector (KLM) is placed outside of the coil.

The remaining part of this subsection provides a short description of the individual sub-detector components. For a more detailed overview, the reader is referred to [65]. An introduction to interaction of particles with detector material as well as the operating principles of particle detectors can be found in [71].

Interaction Region

Electrons and positrons collide at the IP with a non-zero crossing angle. This large angle ensures a quick separation of the beams. In addition, it allows the magnets of the final beam focusing system to be closer to the IP. The IP is enclosed by a double-walled beryllium pipe with an inner radius of 10 mm to shield against particle shower backgrounds.

Pixel Detector

Due to the increased beam background levels, the innermost layer of the tracking detector can no longer be realized with a silicon strip detector due to the large occupancy. Thus the innermost layer of the detector is based on silicon pixel sensors. The pixels sensors provide much more channels and therefore a smaller occupancy. Meanwhile, the SVD is moved outwards to safer regions with respect to the background levels at radii beyond 38 mm.

Precise determinations of B meson decay vertices are important for measurements of CP violation in neutral B meson oscillations at the Belle II experiment. To reduce the size of the PXD and thus the effect of multiple scattering, the Belle II pixel sensors are based on the depleted field effect transistor (DEPFET) technology allowing thin sensors of 75 μm. The advantage of the DEPFET detector is combining the detection and amplification of the signals within one device [72]. With this technology, the readout electronics of the PXD can be placed outside the acceptance region avoiding the need for active cooling inside the detector.

The detector has been designed with two PXD layers at radii of 14 mm and 22 mm. The inner and outer layer consists of 8 and 12 planar sensors each covering an angular acceptance of $\theta \in [17°, 150°]$, respectively. Before the first long shutdown of the Belle II experiment in summer 2022, only one PXD layer is installed inside the detector. A full replacement of the current PXD and the installation of the second layer is planned for the end of 2022.

Silicon Vertex Detector

The SVD consists of four layers of double-sided silicon strip detectors. The radii of the first and fourth SVD layer are 38 mm and 140 mm, respectively. The SVD covers the full angular acceptance of $\theta \in [17°, 150°]$. While the first SVD layer is arranged entirely parallel to the z-axis, the forward sections of the second to fourth layer are tilted with respect to the beam axis. This reduces the overall required material budget.

Tracks reconstructed in the central drift chamber (CDC) are extrapolated to the PXD using SVD data. Together with the PXD, it is possible to reconstruct low-$|\boldsymbol{p}_\mathrm{T}|$ tracks that do not even reach the CDC. This is especially helpful for a efficient reconstruction of D^* meson daughters.

Central Drift Chamber

The CDC is a multi-wire proportional drift chamber with inner and outer radii of 16 cm and 113 cm, respectively. It encloses the two inner silicon tracking detectors and is filled with a He (50%) and C_2H_6 (50%) gas mixture.

The design of the CDC follows mainly the structure of its predecessor in the Belle detector. In total 14 336 sense and 42 240 field wires are arranged in 56 radial layers. The layers are further grouped into nine super-layers consisting of alternating groups of axial and stereo layers. The wires in the axial layers are tensioned parallel to the z-axis, while the wires in the stereo layers are slightly tilted with respect to the z-axis. The separate axial and stereo layers allow for a three-dimensional track reconstruction.

In addition to measurements used for track reconstruction, the CDC contributes to the charged particle identification through $\mathrm{d}E/\mathrm{d}x$ measurements and delivers information for the trigger system. Overall, the CDC covers the full angular acceptance region $\theta \in [17°, 150°]$.

Time-of-Propagation Counter

The TOP surrounds the CDC in the barrel region and is designed for the charged particle identification. It consists of 16 quartz radiators with a thickness of 2 cm each.

The TOP measures the time of propagations of Cherenkov photons that are internally reflected inside the quartz bar. At the end of each quartz bar, an array of photo-multiplier tubes are used to measure the (x, y) coordinates and arrival time of incoming photons. Both measurements combined allow the reconstruction of the Cherenkov angle θ_c. With the estimated momentum of the incoming particle, the particle's mass can be constrained with the relation $\cos\theta_c = 1/(n\beta)$, with n and β denoting the refraction index and the particle's velocity in units of c, respectively. Even though the TOP is tuned for an optimal separation between kaons and pions, other charged particles can also be identified.

Aerogel Ring-Imaging-Cherenkov Counter

In the forward end-cap region, the ARICH detector is used to separate kaons and pions over almost their entire momentum spectra. In addition, the information provided by the ARICH measurements is also useful to discriminate between low momentuma pions, muons, and electrons. The main component of the ARICH is an aerogel radiator, the active material producing Cherenkov photons from charged particles. An array of position sensitive photon detectors are use to detect single Cherenkov photons efficiently. An expansion volume with a length of 20 cm is placed between the active material and the photon detector, allowing the ring image of the Cherenkov photons to form on the photon detector surface.

The active radiator is realized with two 20 mm thick layers of silica aerogel. The refractive indices of the first and second layer are $n = 1.055$ and $n = 1.065$, respectively. The array of photon detectors consists of square shaped hybrid avalanche photo-detectors.

Electromagnetic Calorimeter

The ECL is responsible for high resolutions measurements of photons energies within the range of 20 MeV up to 4 GeV. Further, the ECL is used for the determination of the photon position, identification of electrons and generation of trigger signals.

The calorimeter is separated in a barrel section and annular end-caps in the forward and backward regions. Overall, it covers an polar angle of $\theta \in [12.4°, 155.1°]$. Between the barrel and the end-cap regions, there is a gap of approximately 1°.

The active material of the ECL consists of 8 736 CsI(TI) crystals, which are reused from the ECL of the Belle detector. The crystals are shaped like truncated pyramids with an average cross section of 36 cm^2 and a length of 30 cm, corresponding to a radiation length of $16.1X_0$. Photo diodes are glued to the end of each crystal for the readout of the scintillation light.

Detector Solenoid and Iron Structure

The magnetic field, with an field strength of 1.5 T, is provided by a superconducting solenoid. The solenoid encloses a cylindrical volume with 3.4 m in diameter and 4.4 m and length. For the magnetic flux return, the iron structure of the Belle detector is reused. In addition, the iron structure also provides support for the sub-detector components and acts as absorber material for the KLM.

K_L^0/Muon Detector

The KLM is designed for the detection of K_L^0 and muons over a broad momentum range. Particles need a minimal momentum of $|\boldsymbol{p}| \approx 600\,\mathrm{MeV}$ to reach the KLM. Due to the absence of a magnetic field, charged particles travel in a straight line trough the KLM until they escape ($|\boldsymbol{p}| > 1.5\,\mathrm{GeV}$) or range out due to energy depositions. Trough interaction with the absorber material, K_L^0 mesons create a hadronic shower.

The KLM consists of alternating layers of iron plates (4.7 cm thickness) and active detector elements. In addition to the 0.8 interaction lengths of the ECL , the iron plates add 3.9 interaction lengths for hadrons traversing orthogonally to the detector planes. Glass electrode resistive plate chambers detect charged particles in the barrel region of the KLM. Due to the high background rates in the end-cap region and long dead time of the resistive plate chambers, scintillator strips and silicon photo diodes are used as active detector material in the end-cap region of the KLM.

Because of large fluctuations in the development of hadronic showers, KLM clusters allow only the reconstruction of the direction of the K_L^0 candidate, not its energy. The KLM in the barrel region covers a polar angle of $\theta \in [45^\circ, 125^\circ]$. The coverage is extended to $\theta \in [20^\circ, 155^\circ]$ by the detectors placed in the end-caps.

3.3 Final-State Particle Reconstruction

3.3.1 Track Reconstruction

The main tasks of the track reconstruction is the extraction of a charged particle's momentum $\boldsymbol{p}$ from the track curvature in the magnetic field and its point of closets approach $\boldsymbol{x}$ to the IP. For the track reconstruction, energy depositions or detector hits of charged particles provide spatial information that can be combined to form tracks allowing the inference of the kinematic properties. Several track finding and fitting algorithms are employed to reconstruct charged particle trajectories originating from primary or secondary decay vertices. The track reconstruction can be separated into two steps: track finding and track fitting. A detailed description of the track finding at Belle II is given in [63, 73].

The track finding consists of several pattern recognition algorithms that combine detector hits in the PXD, SVD, and CDC into a single track candidate. This step is complicated by a large amount of additional hits in the detector originating from machine induced backgrounds or detector noise. There are different pattern recognition algorithms used for track finding in the CDC and the silicon based vertex detectors. In the CDC, a global and local track finding approach based on the Legendre algorithm [74] and a cellular automaton [75] are employed, respectively. The resulting hits are merged and a CDC only track is fitted using a deterministic annealing filter (DAF) [76]. A standalone algorithm for the track finding of low momentum particles is used with SVD hits only [77]. A Combinatorial Kalman Filter (CKF) [78] is used to relate CDC tracks to SVD hits. Again CDC and SVD hits are combined and fitted with the DAF. A second CKF is used to extrapolate and combines CDC and SVD tracks to PXD hits.

The final step in the track reconstruction is the track fit with a DAF implemented in the `GenFit2` [76] package. This track fit is performed with a specific particle hypothesis considering energy loss and material effects. At Belle II, the pion, kaon, and proton mass hypotheses are used.

A manuscript describing the performance of the tracking algorithms with Belle II data is currently in preparation [79].

3.3.2 Photon Reconstruction

In each event, the energy and timing information of the ECL crystals are used to reconstruct photon candidates. A dedicated clustering algorithm is applied to group individual crystal into ECL clusters. This algorithm is used to separate signal photon clusters from surrounding energy depositions originating from beam background. In a first step, ECL crystals are grouped into connected regions, which are sets of crystals with a significant energy deposition isolated from other connected regions. Then the connected regions are further split up into clusters based on local maxima in an iterative process. These local maxima are crystals with an energy greater than the energy of its direct neighbors within the connected region. Before one obtains the energy of the incident photon, the final cluster energy has to be corrected for leakage of the electromagnetic shower out of the back of the calorimeter and energy depositions in inactive detector material. This correction depends on the photon energy and the detector region.

Photon candidates are identified with reconstructed ECL clusters that are unassociated with a reconstructed charged particle trajectory. The shower shape is used to discriminate between electromagnetic showers and showers from neutral and charged hadron interactions. The former show a cylindrically symmetric shower shape, while the latter have asymmetric shower shapes. The showers induced from hadron interactions often produce secondary ECL clusters that are also not matched to a reconstructed track.

A more in-depth description of the photon reconstruction is given in [63, 80].

3.3.3 Charged Particle Identification

Tracks are identified as electrons, muons or charged hadrons by combining information from several sub-detectors into a single likelihood-ratio like identification value. This ratio combines the individual particle identification (PID) likelihoods of electrons, muons, kaons, pions, protons, and deuterons. For a certain particle α, PID quantity is defined as

$$\mathrm{PID}(\alpha) = \frac{\mathcal{L}_\alpha}{\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_d}. \tag{3.1}$$

The identification of charged hadrons mainly relies on measurements from the CDC, TOP, and ARICH detectors. Hadrons with momenta below 0.7 GeV are primarily identified from $\mathrm{d}E/\mathrm{d}x$ measurements from the CDC. The SVD also provides independent $\mathrm{d}E/\mathrm{d}x$ measurements, but these are not included in the Belle II data processing at the time of writing. Above momenta of 0.7 GeV, hadrons are primarily identified using the TOP and ARICH measurements. Muons can be reliably identified with the KLM, based on differences of the longitudinal penetration depth and transverse scattering of extrapolated tracks. The primary feature for identifying electrons is the $E/|\boldsymbol{p}|$ value. Here E and $|\boldsymbol{p}|$ denote the measured energy in the ECL and the estimated track momentum, respectively.

The details of the likelihood construction for individual particle hypotheses and sub-detectors are described in [63]. A manuscript studying the performance of the charged particle identification at Belle II is currently in preparation [81].

3.4 Recorded and Simulated Data Samples

3.4.1 Data Samples

The analysis in this work is based on the data collected in 2019 and 2020 by the Belle II detector. The integrated luminosity of the on-resonance data set, recorded at $\sqrt{s} = 10.58\,\mathrm{GeV}$, corresponds to $68.8\,\mathrm{fb}^{-1}$ [82] This is equivalent to approximately recorded $(68.2 \pm 0.9) \times 10^6$ B meson events. In addition, $9.2\,\mathrm{fb}^{-1}$ of data recorded at a CM energy of 60 MeV below the $\Upsilon(4S)$ mass is used to constrain the contribution from $e^+e^- \to q\overline{q}$ continuum processes. This data set is referred to as off-resonance data in the following.

Monte Carlo (MC) samples are used to determine reconstruction efficiencies and acceptance effects as well as to estimate background contaminations. MC samples of B meson decays are generated with the `EvtGen` [83]. The $e^+e^- \to q\overline{q}$ continuum processes are simulated with `KKMC` [84] and `PYTHIA8` [85]. For charged particles, the effect of electromagnetic final-state radiation (FSR) is simulated using `PHOTOS` [86]. Particle interactions with the detector material are simulated with `GEANT4` [87].

The simulated data samples are produced centrally by the collaboration within the MC campaign `MC13`. The MC generation, detector simulation, and event reconstruction of all data sets is based on `release-04` of the open-source Belle II analysis software framework `basf2` [88]. Run-independent $B\overline{B}$ and continuum MC samples corresponding to an integrated luminosity of $200\,\mathrm{fb}^{-1}$ are used in this analysis. In addition, a signal MC sample containing 80×10^6 $B\overline{B}$ events with a $B \to X_c \ell \nu_\ell$ decay on the signal and a generic B decay on the tag-side are used for certain parts of the analysis chain.

3.4.2 Simulation of $B \to X_c \ell \nu_\ell$ Decays

Several steps in the measurement of the q^2 moments rely on accurate modeling of $B \to X_c \ell \nu_\ell$ decays. The m_X shape of the total inclusive $B \to X_c \ell \nu_\ell$ spectrum serves as a template in the binned likelihood to the m_X distribution determining signal and background yields. The measured q^2 moments are calibrated and corrected for efficiency effects based on the q^2 distribution obtained from the simulation of $B \to X_c \ell \nu_\ell$ decays.

The inclusive $B \to X_c \ell \nu_\ell$ spectrum is simulated with several exclusive decay modes. The X_c composition is dominated by the exclusive $B \to D \ell \nu_\ell$ and $B \to D^* \ell \nu_\ell$ decays. The $B \to D \ell \nu_\ell$ decays are modeled using the BGL form-factor parameterization proposed by [89] with central values and uncertainties for the form-factor parameters taken from the fit result in [51]. To simulate $B \to D^* \ell \nu_\ell$ decays, the BGL implementation from [25, 47] is used. The central values and uncertainties for the form-factor parameters are taken from the fit results in [48]. The average branching fractions reported in [40] assuming isopsin symmetry are used for the simulation of both decays.

Table 3.1: Branching fractions used in the simulation of $B \to X_c \ell \nu_\ell$ decays. The values given correspond to the isospin-averaged branching fractions of the individual measurements listed in [40]. The total $B \to X_c \ell \nu_\ell$ branching fraction is taken from [23].

Decay	$\mathcal{B}(B^+)$	$\mathcal{B}(B^0)$
$B \to D \ell \nu_\ell$	$(2.4 \pm 0.1) \times 10^{-2}$	$(2.2 \pm 0.1) \times 10^{-2}$
$B \to D^* \ell \nu_\ell$	$(5.5 \pm 0.1) \times 10^{-2}$	$(5.1 \pm 0.1) \times 10^{-2}$
$B \to D_1 \ell \nu_\ell$	$(6.6 \pm 1.1) \times 10^{-3}$	$(6.2 \pm 1.0) \times 10^{-3}$
$B \to D_2^* \ell \nu_\ell$	$(2.9 \pm 0.3) \times 10^{-3}$	$(2.7 \pm 0.3) \times 10^{-3}$
$B \to D_0^* \ell \nu_\ell$	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.7) \times 10^{-3}$
$B \to D_1' \ell \nu_\ell$	$(4.2 \pm 0.9) \times 10^{-3}$	$(3.9 \pm 0.8) \times 10^{-3}$
$B \to D \pi\pi \ell \nu_\ell$	$(0.6 \pm 0.9) \times 10^{-3}$	$(0.6 \pm 0.9) \times 10^{-3}$
$B \to D^* \pi\pi \ell \nu_\ell$	$(2.2 \pm 1.0) \times 10^{-3}$	$(2.0 \pm 1.0) \times 10^{-3}$
$B \to D \eta \ell \nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \to D^* \eta \ell \nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \to X_c \ell \nu_\ell$	$(10.8 \pm 0.4) \times 10^{-2}$	$(10.1 \pm 0.4) \times 10^{-2}$

Semileptonic B decays to orbitally excited D meson states, denoted as $D^{**} = D_1, D_2^*, D_0^*, D_1'$, are modeled with the heavy-quark-symmetry-based form-factor prescriptions proposed in [90]. Values for the masses and widths of individual D^{**} mesons are taken from [23]. The branching fractions for $B \to D^{**} \ell \nu_\ell$ decays are adopted from [40]. The prescription outlined in [90] is used to account for missing isospin-conjugated and other established decay modes observed in studies of fully hadronic B decays.

The measured $B \to D^{(*)} \pi\pi \ell \nu_\ell$ branching fraction [91] is almost saturated by the contribution of $B \to D_1 (\to D\pi\pi) \ell \nu_\ell$ decays. The remaining part of this branching fractions is filled with non-resonant $B \to D^{(*)} \pi\pi \ell \nu_\ell$ decays.

No additional non-resonant $B \to D^{(*)} \pi \ell \nu_\ell$ decays are considered, since the contribution of $B \to D_I (\to D^{(*)} \pi) \ell \nu_\ell$ already saturates the overall measured $B \to D^{(*)} \pi \ell \nu_\ell$ branching fraction. The small contributions from $B^- \to D_s^+ K^- \ell^- \nu_\ell$ decays are neglected in this analysis.

Overall, the sum of exclusive branching fractions does not saturate the measured inclusive $B \to X_c \ell \nu_\ell$ branching fraction [23]. The missing contribution is referred to as the ‘gap’. This gap is filled with non-resonant $B \to D^{(*)} \eta \ell \nu_\ell$ decays. An uncertainty of 100% is assigned to the branching fraction of the gap modes. Both non-resonant decays, $B \to D^{(*)} \pi\pi \ell \nu_\ell$ and $B \to D^{(*)} \eta \ell \nu_\ell$, are simulated assuming an uniform distribution of the final-state momenta in the available phase space.

For a step-by-step description of the calculation of isospin-averaged $B \to X_c \ell \nu_\ell$ branching fractions, the reader is referred to the resources provided in [92]. Table 3.1 summarizes the $B \to X_c \ell \nu_\ell$ branching fractions for different X_c final-states used in the simulation. Figure 3.3 shows the resulting q^2 spectrum evaluated without reconstruction effects for the different X_c final-states. The hatched uncertainty band represents the shape uncertainty on the normalized distribution due to the assumed branching fraction uncertainties. Contributions from $B \to D^* \ell \nu_\ell$ dominate at high q^2 , whereas $B \to D^{**} \ell \nu_\ell$ and non-resonant X_c ($B \to D^{(*)} \pi\pi \ell \nu_\ell$ and $B \to D^{(*)} \eta \ell \nu_\ell$) have sizeable contributions at low q^2.

The uncertainties on the BGL form-factor parameters used for the simulation of $B \to D^{(*)} \ell \nu_\ell$

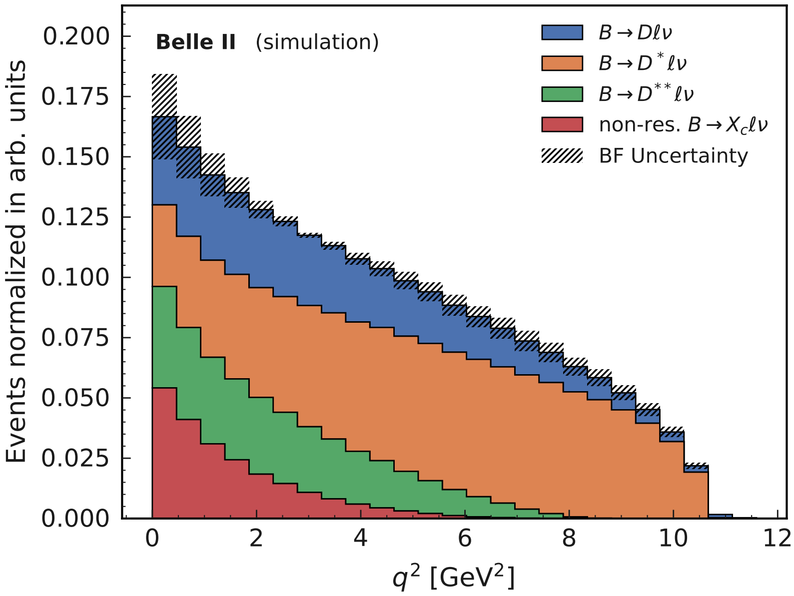


Figure 3.3: Normalized q^2 spectrum for different X_c final-states on generator level without reconstruction effects. The hatched uncertainty band shows the impact of the branching fraction uncertainties on the normalized histograms.

decays are propagated with sets of event weights representing variations within one standard deviation. These weights are determined by a one-dimensional reweighing of the w spectrum with a set of orthogonal parameter variations. The orthogonal parameter variations are calculated as

$$\boldsymbol{p}_i^{\pm} = \boldsymbol{p}^{\text{nom}} \pm \sqrt{\lambda_i}\boldsymbol{e}_i \quad (i = 1, \ldots, N_{\text{params}}), \tag{3.2}$$

with $\boldsymbol{p}^{\text{nom}}$ being the vector of nominal parameter values. Here λ_i and $\boldsymbol{e}_i$ denote the i-th eigenvalue and eigenvector of the covariance matrix of the form-factor parameters. The weights are determined according to the equation

$$w_i^{\pm}(w) = \frac{\Gamma^{\text{nom}}}{\Gamma_i^{\pm}} \frac{\frac{\mathrm{d}\Gamma}{\mathrm{d}w}\left(\boldsymbol{p}_i^{\pm}\right)}{\frac{\mathrm{d}\Gamma}{\mathrm{d}w}\left(\boldsymbol{p}^{\text{nom}}\right)}. \tag{3.3}$$

The formula in Equation (3.3) is evaluated with the generator level value for w. The factor $\Gamma^{\text{nom}}/\Gamma^{\pm}$ ensures that the overall normalization of the total rate remains unchanged. In Equation (3.3), the analytical expression for the differential decay rates $\mathrm{d}\Gamma/\mathrm{d}w$ are used. The expressions for the decay rates are implemented in the `eFFort` package [93].

3.4.3 Lepton Identification Efficiency and Hadron-to-Lepton Fake Rate Corrections

The charged lepton identification efficiency and hadron-to-lepton fake rate are corrected using data-driven weights to account for reconstruction efficiency differences between the simulated and

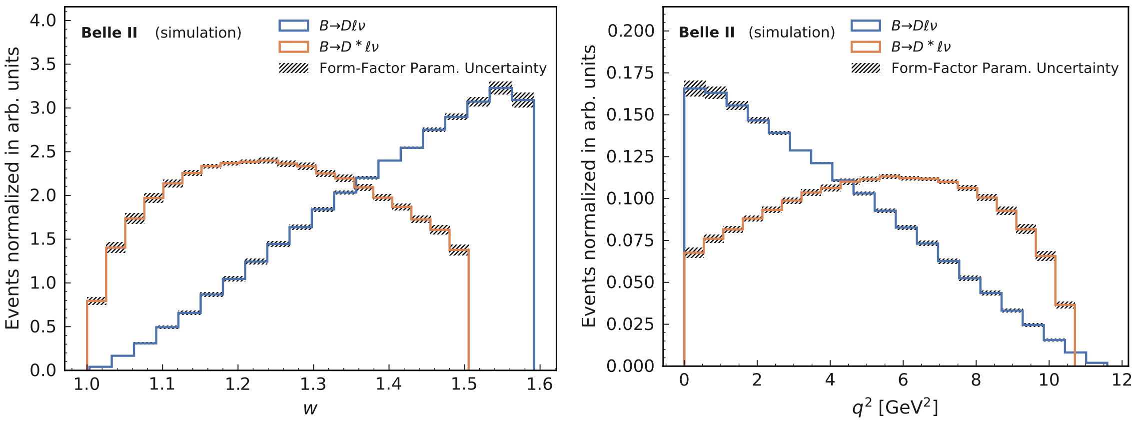


Figure 3.4: Normalized w (left) and q^2 (right) spectra of simulated $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ decays on generator level without reconstruction effects. The hatched uncertainty band shows the impact of the form-factor parameter uncertainties on the normalized histograms.

measured data samples. The corrections are obtained in the studies presented in [94] based on on- and off-resonance data sets corresponding to an integrated luminosity 62.8 fb^{-1} and 8.4 fb^{-1}, respectively.

The electron identification efficiency is studied using samples of $e^+e^- \to e^+e^-(\gamma)$ and $e^+e^- \to e^+e^-e^+e^-$ events as well as $J/\psi \to e^+e^-$ decays. For muons, the particle identification efficiency is corrected using samples of $e^+e^- \to \mu^+\mu^-(\gamma)$ and $e^+e^- \to e^+e^-\mu^+\mu^-$ events and events with $J/\psi \to \mu^+\mu^-$ decays. The hadron-to-lepton fake rates are studied with samples of $K_S^0 \to \pi^+\pi^-$, $D^{*+} \to D^0\pi^+$, and $e^+e^- \to \tau^+\tau^-$ decays. The efficiencies and fake rates ϵ in data and MC are studied as a function of the laboratory-frame momentum $|\boldsymbol{p}|$ and the polar angle θ. To correct the efficiency differences between data and MC, the weight

$$w(|\boldsymbol{p}|, \theta) = \frac{\epsilon_{\text{data}}(|\boldsymbol{p}|, \theta)}{\epsilon_{\text{MC}}(|\boldsymbol{p}|, \theta)} \tag{3.4}$$

is applied for the signal lepton candidate of the $B \to X_c\ell\nu_\ell$ decay in each event. The correction factors applied in this analysis are available at [95] in form of `csv` tables. Here, the charge-averaged corrections are used for a PID selection working point of $\text{PID}(\ell) > 0.9$ as defined in Equation (3.1).

The distribution of applied lepton identification efficiency and hadron-to-lepton fake rate corrections are shown in Figure 3.5. The average efficiency correction for electrons and muons is 0.95 and 0.89, respectively. For the hadron-to-muon fake rate, the average correction is 0.98. The average hadron-to-electron fake rate correction is 1.50, which is driven by the kaon to electron fake rate correction tables.

The uncertainty on the efficiency and fake rate correction is considered with a set of variations around the nominal correction values. These varied weights are generated per $|\boldsymbol{p}|$–θ phase space bin k via

$$w'_k = w_k^{\text{nom}} + \mathcal{G}\left(0, \left(\sigma_k^{\text{stat}}\right)^2\right) + \mathcal{G}\left(\mathbf{0}, C^{\text{sys}}\right), \tag{3.5}$$

with $\mathcal{G}$ denoting the Gaussian distribution. Here, the statistical uncertainty on the correction factor is assumed to be independent for each phase space bin, while the systematic uncertainties are assumed

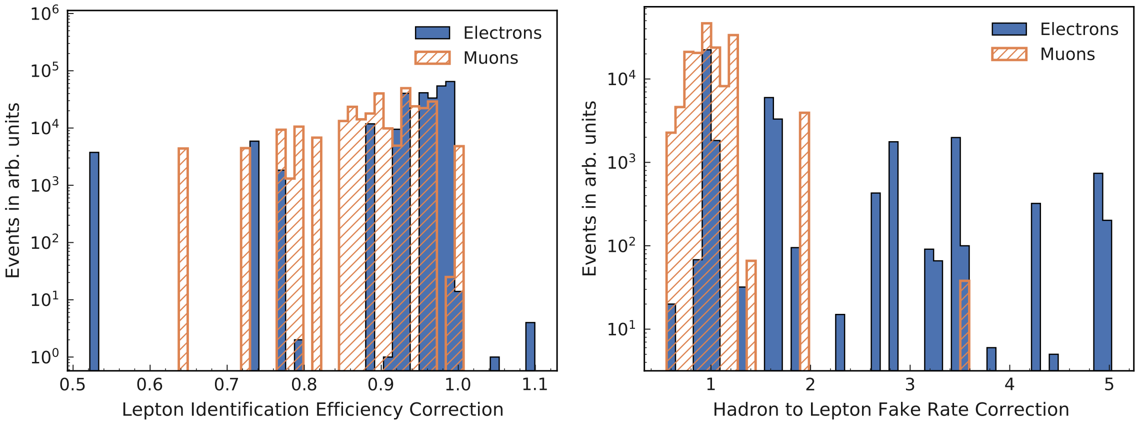


Figure 3.5: Applied lepton identification efficiency (left) and hadron-to-lepton fake rate (right) corrections.

to be fully correlated between all bins. The systematic variations are generated with a multivariate Gaussian distribution and C^{sys} denotes the systematic covariance matrix of the phase space bins k.

Figures 3.6 and 3.7 show normalized distributions in the $|\boldsymbol{p}|-\theta$ plane for signal lepton and hadron-to-lepton fake candidates, respectively. Not the entire $|\boldsymbol{p}|-\theta$ phase space is covered by the provided correction tables. The corrections for the kaon to lepton fake rate suffer from poorly described systematic uncertainties. Thus, $|\boldsymbol{p}|-\theta$ regions with uncertainties greater than 10 are discarded in the application of the correction weight and treated as uncovered regions. Uncovered lepton candidates are assigned a correction weight of $w = 1$. A total averaged uncertainty is assumed for all uncovered lepton candidates calculated as

$$\sigma^2_{\text{avg}} = \overline{\sigma^2}_{\text{stat\&sys}} + \overline{(1 - w^{\text{nom}})}^2, \tag{3.6}$$

with averages over all covered efficiency and fake rate correction bins. The variations are then generated as

$$w'_k = 1 + \mathcal{G}(0, \sigma^2_{\text{avg}}). \tag{3.7}$$

About 2.2×10^{-3} % of the electron and 1.3 % of the muon candidates are not covered by the efficiency correction tables. The average uncertainty assigned to the former is 0.05 and the latter is 0.13. All considered pion to electron fakes are covered by the correction tables, while 0.9 % of the pion to muon fakes are not covered. Here, an average uncertainty of 0.30 is assigned. Regarding the kaon to lepton fakes, 88.3 % of the electron and 2.5 % of the muon fakes are not covered by the available correction tables. Also the average uncertainties for these corrections are large, with 3.31 for electrons and 0.59 for muons.

The set of varied lepton identification efficiency and hadron-to-lepton fake rate correction weights allow an estimation of the covariance matrices for e.g. binned distributions. Unless stated otherwise, a set of 50 variations is uses throughout this analysis.

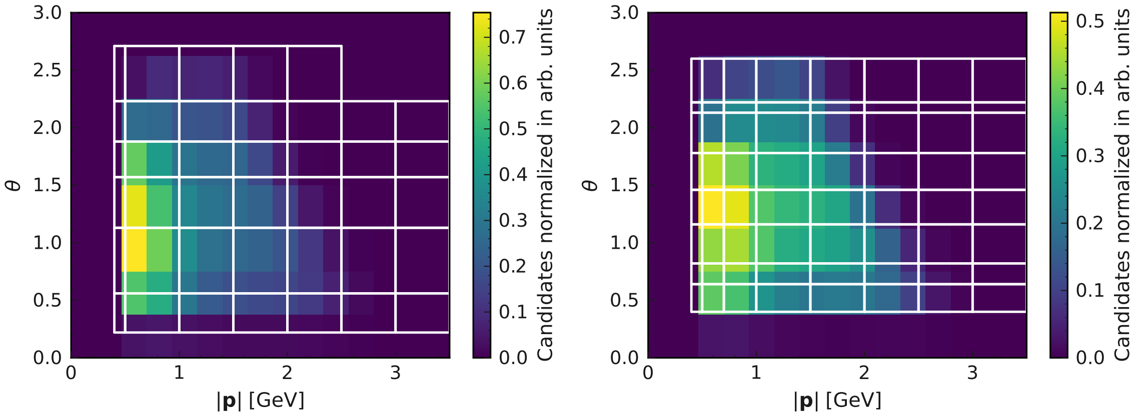
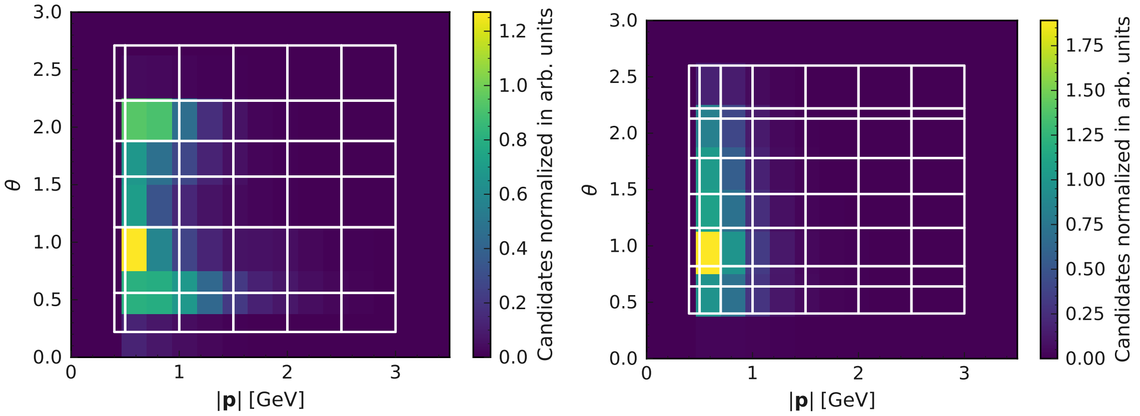
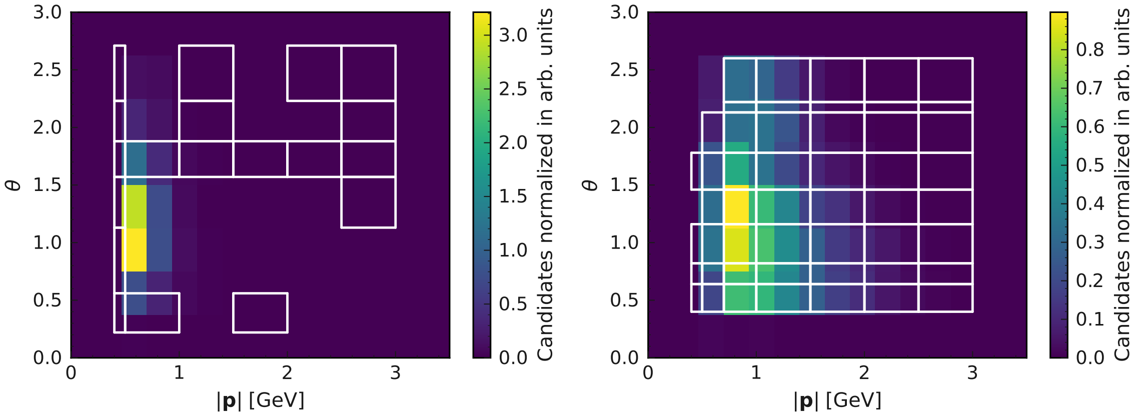


Figure 3.6: Normalized distributions of signal electron and muon candidates in the $|\boldsymbol{p}|$–θ phase space. The white lines indicate the coverage of the lepton efficiency correction tables.



(a) $\pi^- \to \ell^-$



(b) $K^- \to \ell^-$

Figure 3.7: Normalized distributions of hadron-to-electron (left) and muon (right) fake candidates in the $|\boldsymbol{p}|$–θ phase space. The sub-figures (a) and (b) show the distributions for pion to lepton and hadron-to-lepton fakes, respectively. The white lines indicate the coverage of the lepton efficiency correction tables.

CHAPTER 4

Inclusive Reconstruction of $B \to X_c \ell \nu_\ell$ Decays and Event Selection

This chapter gives an overview of the event reconstruction and selection of $B \to X_c \ell \nu_\ell$ decays in $\Upsilon(4S) \to B\overline{B}$ events. The reconstruction of one of the B mesons with an automated tagging algorithm in a fully hadronic decay chain is discussed in Section 4.1. Section 4.2 describes the procedure of reconstructing the semileptonic $B \to X_c \ell \nu_\ell$ signal decay. Here, the hadronic X system is not reconstructed in exclusive charmed meson states. Instead an inclusive approach is used combining all remaining charged particle trajectories and ECL clusters that are not used in the reconstruction of the tag-side B and signal lepton candidate. An overview of the event selection criteria is given in Section 4.3. Lastly, Section 4.4 introduces the kinematic fit used to improve the resolution and reduce the bias,introduced by detector resolution and mis-reconstruction effects, of the reconstructed q^2 spectrum.

4.1 Tag-Side Reconstruction

With a branching fraction of $\mathcal{B}(\Upsilon(4S) \to B\overline{B}) > 96\,\%$ [23], the $\Upsilon(4S)$ resonance decays almost exclusively into a pair of charged or neutral B mesons. The $\Upsilon(4S)$ mass $m_{\Upsilon(4S)} = 10.58\,\mathrm{GeV}$ lies just above the $B\overline{B}$ production threshold. Thus, both B mesons are produced almost at rest and back-to-back in the CM frame and no other additional particles are produced in the decay. Both B mesons further decay via various intermediate states into charged and neutral final-state particles that are considered stable within the Belle II detector. It is a common approach to reconstruct both B meson decays, which can be conceptually separated into a tag and signal-side decay. Due to the well-known initial state of the e^+e^--collision, reconstructing the tag-side B_{tag} yields strong constraints on the flavor and kinematic properties of the second B meson, the signal-side B_{sig}. There are several approaches for reconstructing the tag-side using different B meson decay modes, resulting in varying achievable tagging efficiencies and purities. The signal-side B is reconstructed in the decay chain of interest. This situation is schematically shown in Figure 4.1.

In the case of a semileptonic signal decay, the final state is characterized by a lepton-neutrino pair. Since neutrinos do not interact with the detector material, their momentum cannot be directly inferred. Therefore, fully reconstructing a semileptonic B decay is not possible, which makes studying such decays quite challenging. In this analysis, the B_{tag} meson in reconstructed in a fully hadronic

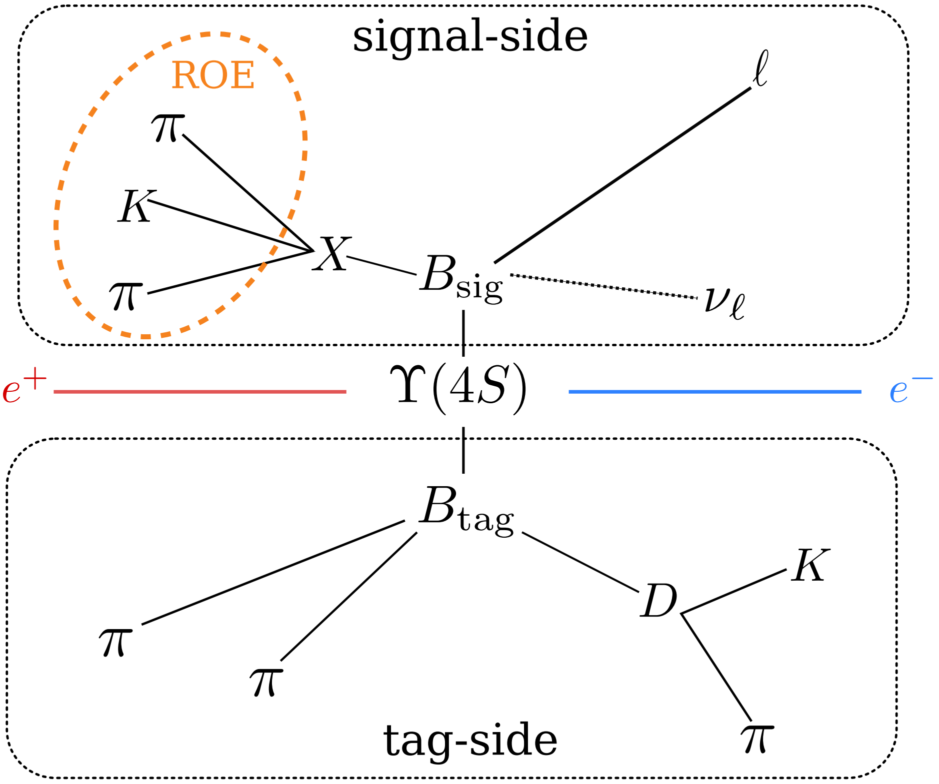


Figure 4.1: Schematic overview of an $\Upsilon(4S) \to B\overline{B}$ event.

decay chain, which yields the largest amount of information about the signal-side decay. This avoids additional neutrinos from e.g. semileptonic D meson decays in the final state. In this case, the B_{sig} four-momentum can be directly inferred from the tag-side four-momentum in CM frame via

$$p^*_{B_{\mathrm{tag}}} + p^*_{B_{\mathrm{sig}}} = 0. \tag{4.1}$$

Hadronic B decays typically have small branching fractions of $\mathcal{O}(10^{-3})$. So only a small number of events can be reconstructed using the hadronic tagging approach, resulting in a low tagging efficiency below 0.1 %. When studying an exclusive decay like $B_{\mathrm{sig}} \to D\ell\nu_\ell$, the full decay chain of the $\Upsilon(4S)$ is reconstructed. Here, no additional tracks and ECL clusters should remain that are not used in the reconstruction of the signal or tag-side B mesons. This completeness requirement imposes a powerful constraint to remove background events with distinct final-state particles from the signal decay. In the case one chooses an inclusive approach to reconstruct the signal decay like $B \to X_c\ell\nu_\ell$, the hadronic system X is not further specified. But it is possible to identify the hadronic X system with all remaining charged particle tracks and ECL clusters not used in the reconstruction of the signal lepton ℓ or tag-side B_{tag} candidates.

The B_{tag} candidate is reconstructed in a fully hadronic decay chain using the Full Event Interpretation (FEI) algorithm [96]. The FEI automatically reconstructs exclusive B meson decay chains in a hierarchical approach using multivariate classifiers to identify correctly reconstructed particle

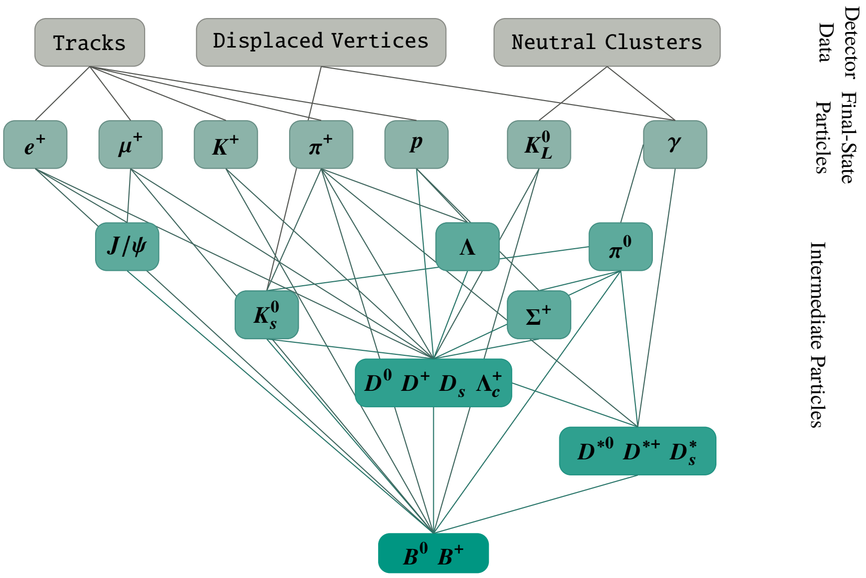


Figure 4.2: Schematic overview of the hierarchical reconstruction approach of the FEI. Adapted from [96].

candidates. In the current application, a speed-optimized implementation of gradient-boosted decision trees [97] is used for the classification.

Figure 4.2 shows a schematic overview of the hierarchical structure used by the FEI. The first stage consists of the reconstruction of final-state particles (e^+, μ^+, K^+, π^-, p, K_L^0, γ) from tracks assuming different mass hypotheses, neutral clusters, and displaced vertices. In the following stages, intermediate particles (J/ψ, π^0, K_S^0, $D^{(*)}$, $D_s^{(*)}$, Λ, Λ_c^+, Σ^+) are recombined in multiple distinct decay channels. In the last stage, B meson candidates are formed combining intermediate and final-state particles. Overall, more than 100 exclusive B decay channels are reconstructed resulting in $\mathcal{O}(10\,000)$ distinct decay chains. For each final-state particle, intermediate particle decay, and B meson decay individual gradient-boosted decision trees are used to estimate a signal probability to separate correctly and incorrectly reconstructed particle candidates. The input features of each classifier consist of kinematic and vertex fit information of the corresponding particle candidate and its daughter particles and the set of discriminating variables is unique for each classifier. The signal probabilities of daughter particles serve as additional input features for the classifiers.

To reduce to combinatorial complexity of the B_{tag} reconstruction, the FEI applies two sets of pre and post-selection criteria. Before estimating the signal probabilities, loose selections on e.g. the photon energy, muon identification probability, or the invariant mass of D mesons are applied. In addition, a best-candidate selection is used to further reduce the number of candidates. After the classification, background-like candidates are removed by a selection on the signal probability. A second best-candidate selection is applied to keep the number of candidates produced in the next stage manageable. The final signal probability of the B_{tag} candidate classifiers $\mathcal{P}_{\text{FEI}}$ combines the available information and can be used to for the selection of the tag-side B meson candidates.

At the time of writing, no dedicated study of the FEI performance on Belle II has been performed. Compared to the Full Reconstruction [98], the hadronic tagging algorithm used by the Belle collaboration, the FEI improves the tagging efficiency by $30-50\%$ depending on the purity [96]. These results are based on studies using the full data set recorded by the Belle experiment. This improvement is mainly attributed to the increased number of additional decay channels and improved candidate selection criteria. A first calibration of the tagging efficiency of the hadronic FEI using Belle II data is reported in [99].

Tag-Side Selection

The $\Upsilon(4S) \to B\overline{B}$ decay is a two body decay into two particles of the same mass m_B. In the CM frame, the energy of a correctly reconstructed B candidate should correspond to half of the CM energy or beam energy in the CM frame E^*_{beam}

$$E^*_B = E^*_{\text{beam}} = \frac{\sqrt{s}}{2}. \tag{4.2}$$

The beam-constrained mass m_{bc} is defined as

$$m_{\text{bc}} = \sqrt{{E^*_{\text{beam}}}^2 - |\boldsymbol{p}^*_{B_{\text{tag}}}|^2}. \tag{4.3}$$

Here, the B energy is substituted with the beam energy. m_{bc} only depends on the momenta of the B candidate's daughter particles and is independent of the daughters' mass hypotheses. For correctly reconstructed B candidates, m_{bc} should have values close to the nominal B meson mass m_B. Another important variable to judge the reconstruction quality of reconstructed B candidates is the energy difference

$$\Delta E = E^*_{B_{\text{tag}}} - E^*_{\text{beam}} \tag{4.4}$$

in the CM frame. The energy difference ΔE should be zero within detector resolution effects for correctly reconstructed B mesons. Contrary to m_{bc}, ΔE is sensitive to the assumed particle hypotheses of the daughter particles because it explicitly depends on the daughters' energies and thus, the assumed mass hypotheses for charged particles.

This analysis uses centrally produced sets of simulated and recorded data with pre-reconstructed B_{tag} candidates, called FEI skims. In the skimmed version of the data sets, the hadronic FEI algorithm is applied and loose selection criteria are used to further reduce the contribution from background processes. Table 4.1 lists the individual event-level pre-selection and B_{tag} candidate selection criteria. Events are required to have at least three well-reconstructed tracks and ECL clusters to suppress low-multiplicity backgrounds. Here, well-reconstructed tracks are defined as tracks with a minimum transverse momentum of $|\boldsymbol{p}_{\text{T}}| > 0.1\,\text{GeV}$ and IP constraints $d_0 < 0.5\,\text{cm}$ and $z_0 < 2\,\text{cm}$, with d_0 and z_0 being the distance and the z-coordinate of the point-of-closest-approach to the origin of the coordinate system. Well-reconstructed ECL clusters are defined as clusters within the CDC acceptance $\theta \in [17^\circ, 150^\circ]$ and with a lower energy threshold $E > 0.1\,\text{GeV}$. Further, the visible energy in the CM frame E^*_{vis} must be greater than 4 GeV and the total energy in the ECL is required to be between 2 and 7 GeV. Additionally, a loose requirement on R_2, the ratio of the second and zeroth Fox-Wolfram moments [100], is used to suppress contributions from continuum processes. A selection of $R_2 < 0.4$ is applied. Tag-side B meson candidates are required to have m_{bc} and $|\Delta E_{B_{\text{tag}}}|$ values greater than 5.24 GeV and smaller than 0.2 GeV, respectively. B_{tag} candidates are required to pass a minimum signal probability threshold of $\mathcal{P}_{\text{FEI}} > 1 \times 10^{-3}$. To further improve the reconstruction quality of the

Table 4.1: Event-level and tag-side candidate selection criteria applied in th FEI skim.

	Variable	Selection Criteria
Event-Level	N_{tracks}	≥ 3
	$N_{\text{ECL}_{\text{clusters}}}$	≥ 3
	E^{*}_{vis}	$> 4\,\text{GeV}$
	$E_{\text{ECL}_{\text{total}}}$	$\in [2\,\text{GeV}, 7\,\text{GeV}]$
	R_2	< 0.4
Tag-Side Candidate	$m_{\text{bc},B_{\text{tag}}}$	$> 5.24\,\text{GeV}$
	$\lvert\Delta E_{B_{\text{tag}}}\rvert$	$< 0.2\,\text{GeV}$
	$\mathcal{P}_{\text{FEI}}$	$> 1 \times 10^{-2}$

tag-side in this analysis, the requirement on the FEI signal probability is tightened to $\mathcal{P}_{\text{FEI}} > 1 \times 10^{-2}$.

4.2 Signal-Side Reconstruction

The semileptonic decay of the signal-side B meson is identified by selecting an electron or muon candidate with laboratory frame momentum greater than 0.5 GeV, at least one hit in the CDC, and a polar angle within the CDC acceptance. Their respective trajectories are constrained to originate near the IP by requiring $d_r < 1\,\text{cm}$ and $|d_z| < 2\,\text{cm}$. Here, d_r and d_z denote the distances to the IP transverse to and along the z-axis, respectively. In addition, electron and muon candidates are required to have a minimal particle identification probability of 0.9.

Electrons are susceptible to bremsstrahlung radiation through interactions with dense detector material. Two effects can be observed if such bremsstrahlung photons are not recovered: a degradation of the electron four-momentum resolution and the assignment of radiated photons to the inclusive X system. To account for the energy of electrons lost to bremsstrahlung photons, the four-momenta of such photons are added to the four-momenta of electrons. Bremsstrahlung photons are identified using the electron track, extrapolating its PXD and SVD hits and the estimated track intersections with the beam pipe and inner wall of the CDC to the ECL to search for clusters. ECL clusters with energies between 2% and 100% of the electron energy and without any other track association are identified as potential bremsstrahlung photons. All clusters that lie within three times the expected resolutions in polar and azimuthal angles are used to correct the electron candidate. These clusters are then removed from consideration for the remainder of the analysis. A more detailed description of the bremsstrahlung recovery algorithm is available in [101].

Tracks of low-momentum particles have sufficiently small radii such that the particles reenter or never leave the sensitive volume of the tracking detectors. These particles loop inside the detector and lead to multiple reconstructed tracks with similar helix parameters. They appear as tracks with the same charge assignment within close proximity or with opposite charge assignments with a back-to-back topology. It is important to identify and remove such duplicated track candidates in order to not assign additional particles to the inclusive X system. For identifying duplicated track candidates, the proximity and momenta of all low-momentum tracks with $|\boldsymbol{p}_{\text{T}}| < 275\,\text{MeV}$ are compared. Pairs of tracks are selected with a maximal momentum difference of $|\boldsymbol{p}_1 - \boldsymbol{p}_2| < 100\,\text{MeV}$ and an angle between $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ smaller than 15° (greater than 165°) for same (opposite) charge tracks. Among the possible duplicated tracks, the track with the smallest value of the heuristic relation $(5 \times d_r)^2 + |d_z|^2$ is selected [102] and the other tracks are discarded in the remaining analysis.

Signal leptons are combined with B_{tag} candidates to $B_{\text{tag}}\ell$ pairs. For charged tag-side B mesons, the signal-side lepton is required to have a charge opposite to that of the B_{tag} candidate. In the case of a neutral B_{tag} candidate, no charge requirement is imposed to allow for the possibility of neutral B meson mixing. Here, the charge correlation $Q_b \times Q_\ell < 1$ between the b quark constituent of the B_{tag} candidate and signal lepton is no longer given. Overall, the reconstruction channels $B^+\ell^-$, $B^0\ell^-$, and $B^0\ell^+$ with $\ell = e, \mu$ are considered. Since no exclusive charmed states are explicitly reconstructed, the hadronic X system in the semileptonic $B \to X_c \ell \nu_\ell$ decay is identified with the rest-of-event (ROE) of the $B_{\text{tag}}\ell$ combination. The ROE contains all particles in the event that are not used in the reconstruction of $B_{\text{tag}}\ell$ pair. It is constructed with the remaining ℓ^+, K^+, π^+, and γ candidates. The charged kaon candidates are selected using the same track requirements as for the signal leptons and additionally require a kaon identification probability greater than 0.6. For the charged pion candidates, a looser IP constraint of $d_r < 2\,\text{cm}$ and $|d_z| < 4\,\text{cm}$ is applied to allow for the inclusion of daughters of K_S^0 decays. To suppress beam-background induced ECL depositions, photons are required to pass an ECL region-dependent selection on the transverse energy E_T and the output of a multivariate classifier $\mathcal{P}_{\text{Zernike}}$ combining eleven Zernike moments [103] of the cluster. The lower threshold for the selections are $E_T > 20\,\text{MeV}$ and $\mathcal{P}_{\text{Zernike}} > 0.35$ in the forward end-cap ($\theta \in [12.4°, 31.4°]$), $E_T > 30\,\text{MeV}$ and $\mathcal{P}_{\text{Zernike}} > 0.15$ in the barrel region ($\theta \in [32.2°, 128.7°]$), and $E_T > 20\,\text{MeV}$ and $\mathcal{P}_{\text{Zernike}} > 0.40$ in the backward end-cap ($\theta \in [130.7°, 155.7°]$). Table 4.2 summarizes all selection criteria applied to the considered final-state particles.

The four-momentum of the hadronic X system is calculated from the summed four-momenta of charged particles and photons that are part of the ROE

$$p_X = \sum_{\substack{\text{charged} \\ \in \text{ROE}}} p_{\text{charged}} + \sum_{\substack{\gamma \\ \in \text{ROE}}} p_\gamma. \tag{4.5}$$

With this expression, the hadronic mass is directly calculated as

$$m_X = \sqrt{p_X^2}. \tag{4.6}$$

A fit to the hadronic mass m_X distribution is used in a later stage of the analysis in order to determine the signal $B \to X_c \ell \nu_\ell$ and background normalizations. Further, the total charge of the event is determined as

$$Q_{\text{evt}} = Q_{B_{\text{tag}}} + Q_\ell + Q_X. \tag{4.7}$$

The neutrino four momentum can be inferred with the missing four-momentum and is reconstructed as

$$p_{\text{miss}} = (E_{\text{miss}}, \boldsymbol{p}_{\text{miss}}) = p_{e^+e^-} - p_{B_{\text{tag}}} - p_\ell - p_X. \tag{4.8}$$

Table 4.2: List of all final-state particles and their corresponding selection criteria.

Final-State Particle	Selection Criteria	
e^+, μ^+, K^+, π^+	$\theta \in [17°, 150°]$ $N_{\mathrm{CDCHits}} > 0$	
e^+, μ^+, K^+	$d_r < 1\,\mathrm{cm}$ $\lvert d_z \rvert < 2\,\mathrm{cm}$	
π^+	$d_r < 2\,\mathrm{cm}$ $\lvert d_z \rvert < 4\,\mathrm{cm}$	
e^+	$\mathcal{P}_e > 0.9$ $\lvert \boldsymbol{p} \rvert > 0.5\,\mathrm{GeV}$	
μ^+	$\mathcal{P}_\mu > 0.9$ $\lvert \boldsymbol{p} \rvert > 0.5\,\mathrm{GeV}$	
K^+	$\mathcal{P}_K > 0.6$	
γ	$\theta \in [12.4°, 31.4°]$	$E_T > 20\,\mathrm{MeV}$ $\mathcal{P}_{\mathrm{Zernike}} > 0.35$
	$\theta \in [32.2°, 128.7°]$	$E_T > 30\,\mathrm{MeV}$ $\mathcal{P}_{\mathrm{Zernike}} > 0.15$
	$\theta \in [130.7°, 155.7°]$	$E_T > 20\,\mathrm{MeV}$ $\mathcal{P}_{\mathrm{Zernike}} > 0.40$

The lepton mass squared, the key quantity explored in this analysis, is calculated as

$$q^2 = (p^*_{B_{\mathrm{sig}}} - p^*_X)^2. \tag{4.9}$$

Here the CM signal-side four-momentum is calculated using the reconstructed B_{tag} momentum and the beam energy in the CM frame assign

$$p^*_{B_{\mathrm{sig}}} = (\sqrt{s}/2, -\boldsymbol{p}^*_{B_{\mathrm{tag}}}). \tag{4.10}$$

4.3 Event Selection

Simulated events are categorized by the MC decay matched to the signal lepton candidate. The applied matching algorithm is based on `PDG` codes from the MC particle numbering scheme [23]. It uses the `basf2` meta variables `matchedMC`, `mcMother`, and `mcDaughter` to retrieve the `PDG` codes of the corresponding MC particles. A signal lepton originating from a semileptonic B decay is identified with a `matchedMC(PDG)` and `mcMother(PDG)` code compatible with that of a light lepton and a B meson, respectively. The semileptonic B decays are further distinguished by analyzing `mcDaughter(PDG)` codes of the respective MC B meson. In addition, if the `matchedMC(PDG)` code of the lepton candidate corresponds to other particles than light leptons or the matched MC decay is not a semileptonic B

decay, several background categories can be identified.

In this analysis, the following categories are used to distinguish events:

$\mathbf{B \to X_c \ell \nu_\ell}$ The signal lepton is matched to a semileptonic B decay with a charmed hadronic system X_c. This category describes the signal component in this analysis. It can be further separated into different sub-categories based on the charmed meson matched to X_c with $X_c = \{D, D^*, D^{**}, \text{non-resonant } X_c\}$.

$\mathbf{B\overline{B}}$ Background All other events originating from $B\overline{B}$ events are considered as background. They can be further broken down into:

- **$\mathbf{B \to X_u \ell \nu_\ell}$** The signal lepton originates from a semileptonic $b \to u\ell\nu_\ell$ transition. Here, the hadronic system X_u also contains resonant and non-resonant contributions. This category constitutes an important background at low and high values of m_X and q^2, respectively.
- **$\mathbf{B \to X\tau\nu_\tau}$** This category contains events where the signal lepton does not directly originate from a semileptonic B decay but is matched to a τ decay from a semitauonic $b \to q\tau\nu_\tau$ transition. This component is characterized by additional neutrinos from the τ decay in the final state.
- **Secondary Leptons** Here, the signal lepton does not originate from a semileptonic B decay but from another secondary decay like semileptonic D or J/ψ decays. This category is a major background component especially at low q^2 values.
- **hadron-to-Lepton Fakes** Charged kaons and pions from the tag-side or e.g. $B \to DK/\pi$ decays can be mis-identified as lepton candidates. Also this component constitutes one of the major backgrounds in the low q^2 region.
- **Remaining $\mathbf{B\overline{B}}$ Background** All remaining $B\overline{B}$ events that do not fit in any of categories described above are combined in this category.

$\mathbf{e^+e^- \to q\overline{q}}$ Continuum Processes $e^+e^- \to q\overline{q}$ events with $q = \{u, d, s, c\}$ have a significant contribution to the e^+e^- production cross section at $\sqrt{s} = m_{\Upsilon(4S)}$.

A combination of rectangular selection requirements is chosen to reject background events and improve reconstruction qualities of the B_{tag} and X candidates.

To further improve the reconstruction quality of the tag-side, only B_{tag} candidates with $m_{\text{bc}} > 5.27\,\text{GeV}$ and $-0.15\,\text{GeV} < \Delta E < 0.10\,\text{GeV}$ are selected. These selection criteria define the tag-side signal region and suppress mis-reconstructed B_{tag} candidates. A better quality of the B_{tag} candidate reconstruction reduces the chance of mis-assigning particles to the inclusive X system.

To suppress contribution from continuum processes, differences in the angular distributions of particles produced in $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ and $e^+e^- \to q\overline{q}$ events are exploited. In the CM frame, B meson decay products are distributed isotropically since the B mesons are produced almost at rest. In contrast, in $q\overline{q}$ events, particles are produced in two back-to-back jets of light hadrons since both quarks are produced with high momenta. The thrust axis $\boldsymbol{T}$ for a set of N particles, e.g. particles forming a B meson candidate, is defined as the unit vector along which the projection of their momenta is maximal

$$\max_{|\boldsymbol{T}|=1} \left(\sum_{i=1}^{N} \boldsymbol{p}_i \cdot \boldsymbol{T} \right). \tag{4.11}$$

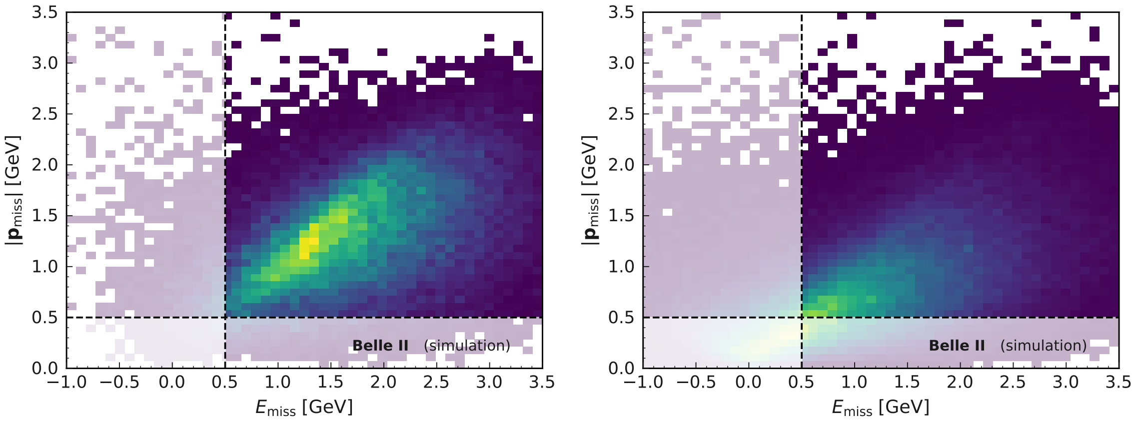


Figure 4.3: Two-dimensional histograms of E_{miss} and $\boldsymbol{p}_{\mathrm{miss}}$ for signal $B \longrightarrow X_c \ell \nu_\ell$ (left) and background (right) events. The shaded area indicate the selection requirements of $E_{\mathrm{miss}} > 0.5\,\mathrm{GeV}$ and $|\boldsymbol{p}_{\mathrm{miss}}| > 0.5\,\mathrm{GeV}$.

A single selection based on $|\cos\theta_{\mathrm{T}}|$ is used to suppress continuum background. Here, θ_T is the angle between the thrust axis of the tag-side B candidate decay products and all other particles in the events not associated with the tag-side B meson in the CM frame [104]. For $B\overline{B}$ events, $|\cos\theta_{\mathrm{T}}|$ is expected to follow a uniform distribution between 0 and 1, while continuum events peak at values close to 1 in the $|\cos\theta_{\mathrm{T}}|$ distribution. An upper threshold of $|\cos\theta_{\mathrm{T}}| < 0.7$ is used to suppress contributions from continuum processes.

The total event charge is required to be $|Q_{\mathrm{evt}}| \leq 1$. The allowed charge imbalance accounts for possible missing tracks from low-momentum particles which have only a minor impact on the reconstruction quality of the inclusive X system. In addition, at least one charged particle has to be part of the inclusive X system.

Further, events are required to pass the criteria $E_{\mathrm{miss}} > 0.5\,\mathrm{GeV}$ and $|\boldsymbol{p}_{\mathrm{miss}}| > 0.5\,\mathrm{GeV}$. Figure 4.3 shows two-dimensional histograms of E_{miss} and $\boldsymbol{p}_{\mathrm{miss}}$ for $B \longrightarrow X_c \ell \nu_\ell$ and background events. Both, the missing energy and the missing momentum are estimates for the energy and momentum of the prompt neutrino in the $B \longrightarrow X_c \ell \nu_\ell$ decay. The neutrinos from semileptonic decays tend to have higher energies and momenta.

To further suppress the small contributions from exclusive and inclusive $B \longrightarrow X_u \ell \nu_\ell$ decays, the low m_X region is rejected by requiring $m_X > 0.5\,\mathrm{GeV}$. This selection requirement is applied to the hadronic mass distribution calculated after the kinematic fit discussed in Section 4.4. Figure 4.4 shows the distributions for the individual observables used in the event selection. The rectangular selection requirements are indicated with the shaded areas. Each distribution is shown without the application of any of the event selection criteria describe above.

The fraction of remaining events for the signal $B \longrightarrow X_c \ell \nu_\ell$, $B\overline{B}$ background and $e^+e^- \longrightarrow q\overline{q}$ continuum events after consecutively applying the individual selection criteria are listed in Table 4.3. After the application of all event selection criteria, the overall retention rate for $B \longrightarrow X_c \ell \nu_\ell$ signal events is 41.15‰. The $B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$ background can be reduced to a contribution of 22.50‰ and 3.94‰, respectively.

After applying the event selection criteria, the average $B_{\mathrm{tag}}\ell$ pair multiplicity per event is 1.5. In case of multiple $B_{\mathrm{tag}}\ell$ combinations per event, a best-candidate selection is performed. The signal lepton candidates are ranked according to their momenta. Only the $B_{\mathrm{tag}}\ell$ pair with the highest lepton

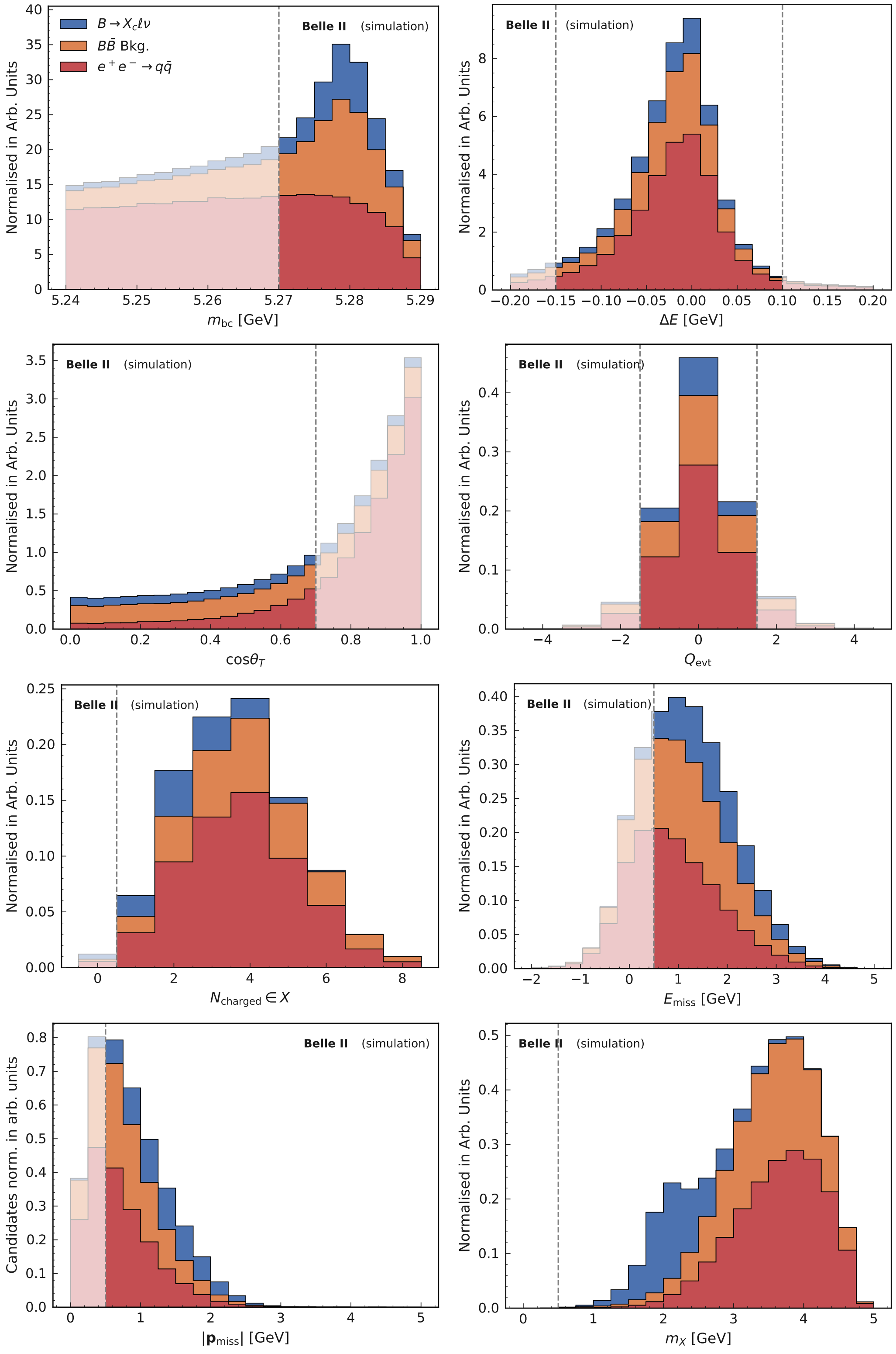


Figure 4.4: Normalized distributions of observables used in the event selection. The individual selection requirements on the respective variables are indicated with the shaded areas. All distributions are shown without any selection criteria applied.

Table 4.3: Fraction of events remaining after each selection step for the signal $B \to X_c \ell \nu_\ell$, $B\overline{B}$ background and $e^+e^- \to q\overline{q}$ continuum components. All values are given in %.

	$B \to X_c \ell \nu_\ell$	$B\overline{B}$ Bkg.	$e^+e^- \to q\overline{q}$
$m_\mathrm{bc} > 5.27\,\mathrm{GeV}$	75.56	64.70	41.44
$-0.15\,\mathrm{GeV} < \Delta E < 0.10\,\mathrm{GeV}$	72.23	62.10	39.67
$\lvert\cos\theta_\mathrm{T}\rvert < 0.7$	50.37	40.88	8.22
$\lvert Q_\mathrm{evt}\rvert \leq 1$	47.38	35.29	7.14
$N_\mathrm{charged} \in X > 0$	45.45	34.95	7.05
$E_\mathrm{miss} > 0.5\,\mathrm{GeV}$	42.90	27.75	5.04
$\lvert \boldsymbol{p}_\mathrm{miss}\rvert > 0.5\,\mathrm{GeV}$	41.19	22.53	3.94
$m_X > 0.5\,\mathrm{GeV}$	41.15	22.50	3.94

momentum is considered in the remaining analysis. When the same lepton is combined with multiple tag-side candidates, one B_tag candidate is chosen randomly.

4.4 Kinematic Fit

With the simple approach, to reconstruct the inclusive X system with all remaining charged particles and photons in the ROE of the $B_\mathrm{tag}\ell$ pair, only a poor resolution is achieved in observables related to the inclusive system. Possible causes for this poor resolution are mis-assignments of particles from the B_tag candidate to the inclusive X system (and vice-versa), particles of the X system that could not be reconstructed, or the addition of particles not belonging to the event like e.g. beam-background induced photon candidates. Figure 4.5 shows the distribution of the residuals of q^2 and m_X for $B \to X_c \ell \nu_\ell$ signal MC. The residuals are defined as the difference between the reconstructed and generator level values $\Delta(x) = x_\mathrm{reco} - x_\mathrm{gen}$. The $\Delta(q^2)$ and $\Delta(m_X)$ distributions are asymmetric and exhibit long tails towards positive and negative values. The root-mean-square of the residual distribution

$$\Delta(x)_\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\Delta(x_i)^2} \tag{4.12}$$

is used to measure the resolution, while the mean $\overline{\Delta(x)}$ is used to quantify the bias of the reconstructed observable. The bias and resolution of the reconstructed q^2 are $3.43\,\mathrm{GeV}^2$ and $5.76\,\mathrm{GeV}^2$, respectively. For m_X, the bias and resolution are $-0.36\,\mathrm{GeV}$ and $0.68\,\mathrm{GeV}$, respectively.

At an e^+e^- collider, the well-defined initial-state yields kinematic constraints on the semileptonic decay of the signal B meson. These constraints can be exploited, to improve the resolution and reduce the bias of q^2 and m_X by performing a kinematic fit of the entire event. The kinematic fit itself is a χ^2 fit of the particle four-momenta with a set of external physical constraints. In the case of a fully reconstructed tag-side and a $B \to X_c \ell \nu_\ell$ decay on the signal-side, there are eleven measured parameters $\boldsymbol{y}$ and three unmeasured $\boldsymbol{a}$. The measured parameters $\boldsymbol{y}$ consist of the B_tag and X four-momenta p_{B_tag}

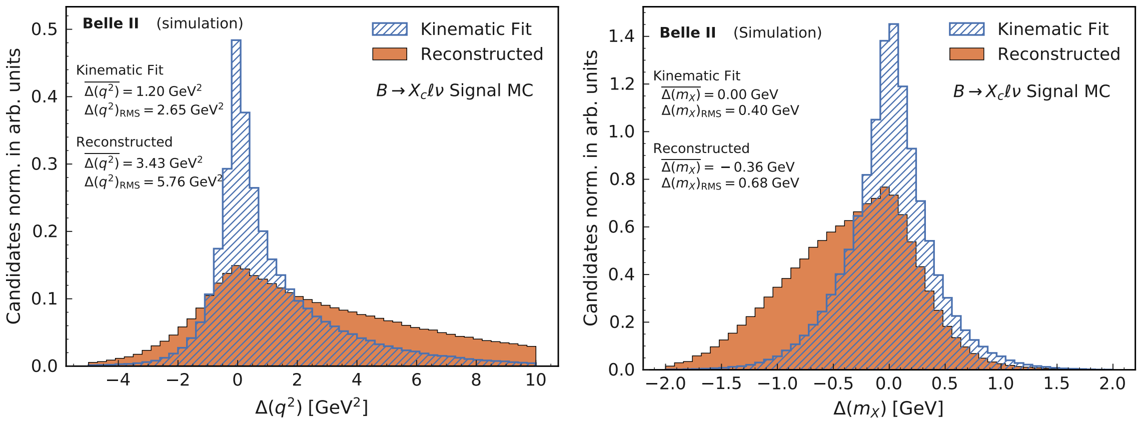


Figure 4.5: Comparison of q^2 and m_X residuals for $B \to X_c \ell \nu_\ell$ signal events before and after the kinematic fit. The resolution and the bias of the reconstructed observable is measured with the mean and root-mean-square of the residuals distribution, respectively.

and p_X, as well as the momentum of the signal lepton $\boldsymbol{p}_\ell$

$$\boldsymbol{y} = (p_{B_{\mathrm{tag}}}, p_X, \boldsymbol{p}_\ell), \tag{4.13}$$

while the unmeasured parameters $\boldsymbol{a}$ are given by the neutrino momentum

$$\boldsymbol{a} = \boldsymbol{p}_\nu. \tag{4.14}$$

The energies of the lepton and neutrino are calculated from the momenta using the energy-momentum relation and assuming $p_\ell^2 = m_\ell^2$ and $p_\nu^2 = 0$.

The physical constraints can be formulated as a set of equations that are fulfilled by the true parameter values $\boldsymbol{y}_{\mathrm{true}}$ and $\boldsymbol{a}_{\mathrm{true}}$, such that it holds

$$g_i(\boldsymbol{y}_{\mathrm{true}}, \boldsymbol{a}_{\mathrm{true}}) = 0 \quad \text{or} \tag{4.15}$$

$$g_i(\boldsymbol{y}_{\mathrm{true}}, \boldsymbol{a}_{\mathrm{true}}) \geq 0. \tag{4.16}$$

Due to detector resolution and mis-reconstruction effects, the measured parameters $\boldsymbol{y}$ will not solve these equations. Corrected parameter values $\boldsymbol{y}'$ and values for the unmeasured parameters, $\boldsymbol{a}$ can be obtained by minimizing the χ^2 function

$$\chi^2(\boldsymbol{y}', \boldsymbol{a}) = (\boldsymbol{y}' - \boldsymbol{y})^\intercal C^{-1} (\boldsymbol{y}' - \boldsymbol{y}) - \sum_i \lambda_i g_i(\boldsymbol{y}', \boldsymbol{a}). \tag{4.17}$$

Here, formally the constraints are considered directly in the cost function using a set of Lagrange multipliers λ_i. The constraint minimization of the χ^2 function is performed using the `SLSQP` algorithm [105] implemented in `SciPy` [106]. The fit is performed with the following constraints:

$$g_1 = E_{e^+e^-} - E_{B_\text{tag}} - E_X - E_\ell - E_\nu = 0, \tag{4.18}$$

$$g_{2,3,4} = \boldsymbol{p}_{e^+e^-} - \boldsymbol{p}_{B_\text{tag}} - \boldsymbol{p}_X - \boldsymbol{p}_\ell - \boldsymbol{p}_\nu = 0, \tag{4.19}$$

$$g_5 = p^2_{B_\text{tag}} - m_{B^2} = 0, \tag{4.20}$$

$$g_6 = (p_X - p_\ell - p_\nu)^2 - m_{B^2} = 0, \quad \text{and} \tag{4.21}$$

$$g_7 = p^2_X \geq 0. \tag{4.22}$$

Constraints g_1 and $g_{2,3,4}$ impose energy and momentum conservation to the event. In the constraints g_5 and g_6, the mass of the B_sig and B_tag candidates are constrained to the same value of $m_B = 5.279\,\text{GeV}$. The last constraint g_7 is used to maintain a physical meaning of the hadronic mass m_X. Without g_7, it can occur that p^2_X becomes negative due to numerical reasons during the minimization process.

The covariance matrix C in Equation (4.17) is constructed as a block-diagonal matrix of the form

$$C = \begin{pmatrix} C_{B_\text{tag}} & 0 & 0 \\ 0 & C_X & 0 \\ 0 & 0 & C_\ell \end{pmatrix}, \tag{4.23}$$

with C_{B_tag}, C_X, and C_ℓ being the covariance matrices of the B_tag, X, and ℓ four and three-momenta, respectively. The covariance matrix for the lepton momentum C_ℓ is given by the fit result of the underlying track fit. For the tag-side B meson, a covariance matrix would be available as a result of a successful vertex fit. However, differences in the rate of a successful vertex fit in data and MC are observed rendering this variable unusable in the selection process. The covariance matrix for the inclusive X system could be calculated as the sum of the covariance matrices of its final-state particles. But this might underestimate the uncertainty of the four-momentum since one of the main mis-reconstruction effects is missing particles. Thus, the resolution of the B_tag and X four-momenta is studied using MC samples, and appropriate covariance matrices are constructed. For the X system, the residuals distributions $\Delta(p_X)$ of the four-momentum are studied in bins of

- the particle multiplicity $X_\text{multi} \in (0,5], (5,7], (7,10], (10,13], (13,30]$ and
- the polar angle $\theta \in (0,0.52], (0.52,1.05], (1.05,1.57], (1.57,2.09], (2.09,\pi]$.

The residuals of the tag-side four momenta are in bins of $m_\text{bc} \in (5.270, 5.278]$, $(5.278, 5.280]$, $(5.280, 5.282]$ and $(5.282, 5.30]$ GeV. Figure 4.6 shows examples of the normalized energy residuals distributions for the tag-side candidate and X system in one respective phase space bin. The energy residuals distributions are asymmetric and do not follow a Gaussian distribution. Hence, the root-mean-square is chosen over the empirical variance to estimate the resolution for the four-momenta.

The B_tag and X four-momentum covariance matrices are constructed as

$$C = DRD \tag{4.24}$$

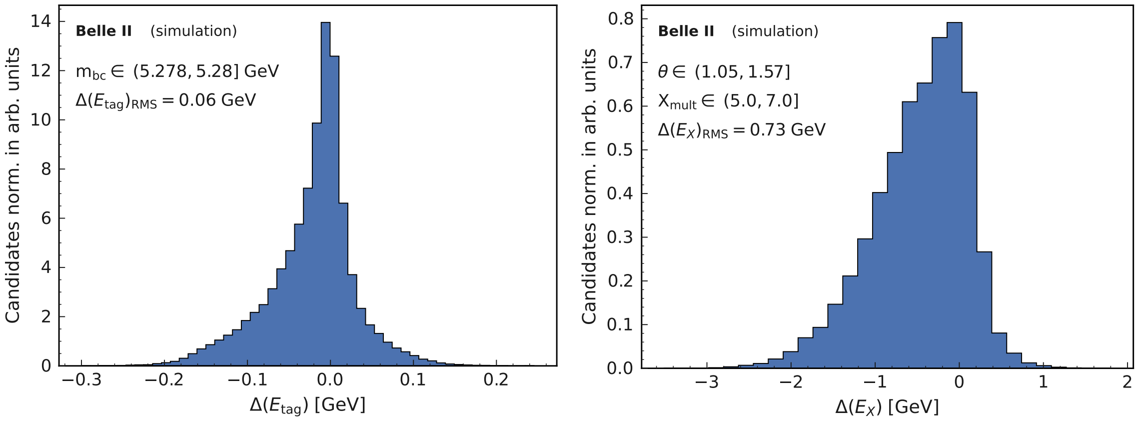


Figure 4.6: Examples of the B_{tag} (left) and X (right) energy residuals distributions in on phase space bin. The resolution energy is measured with the root-mean-square of the residuals distribution.

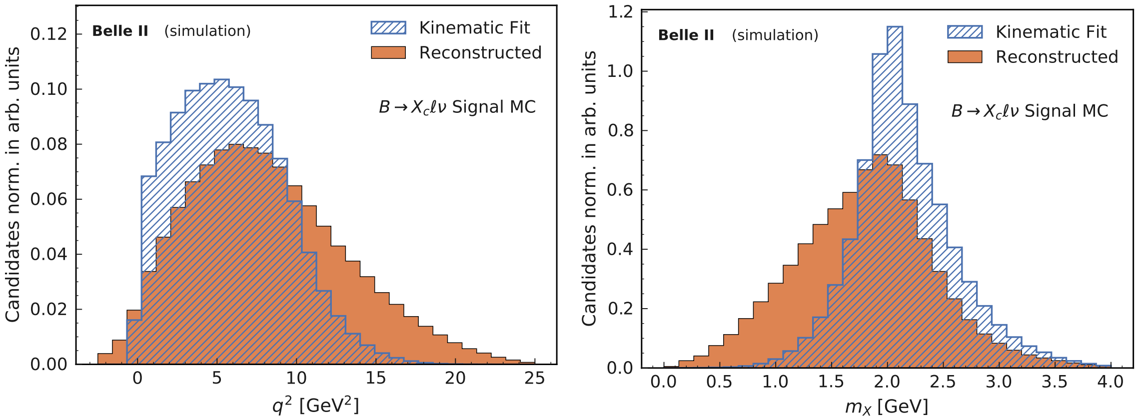


Figure 4.7: Comparisons of the q^2 and m_X spectra for $B \to X_c \ell \nu_\ell$ signal events before and after the kinematic fit.

with the diagonal matrix

$$D = \text{diag}\left(\Delta(E)_{\text{RMS}}, \Delta(p_x)_{\text{RMS}}, \Delta(p_y)_{\text{RMS}}, \Delta(p_z)_{\text{RMS}}\right) \tag{4.25}$$

and R denoting the correlation matrix of the four-momentum components estimated with the Pearson correlation coefficients. Since the kinematic fit is applied before the final event-selection, all event-selection criteria except the lower threshold on m_X and the best-candidate selection are applied to the MC sample used to determine the covariance matrices.

The effect of the kinematic fit on the q^2 and m_X residuals is shown in Figure 4.5. After the kinematic fit, the distributions of the residuals are more symmetric and the tails are reduced. The root-mean-square values of the q^2 and m_X distributions improve to 2.65 GeV^2 and 0.40 GeV, respectively. For q^2, the bias is reduced to 1.20 GeV^2. The residual distribution of m_X is unbiased after applying the kinematic fit. Figure 4.7 shows comparisons of the q^2 and m_X spectra for $B \to X_c \ell \nu_\ell$ signal events before and after the kinematic fit. Due to the energy and momentum conservation constraint, the q^2 spectrum after kinematic fit begins at $q^2 = 0\,\mathrm{GeV}^2$ and negative q^2 values can be avoided. Additionally, the drop in the spectrum at $q^2 \approx 11.6\,\mathrm{GeV}^2$, the kinematically allowed endpoint of the q^2 spectrum, is more prominent. In the m_X spectrum, a more distinct peak around $m_X \approx 2\,\mathrm{GeV}$, representing the mass of the D and D^* mesons, is visible. Figure 4.8 shows the resolution $\Delta(q^2)_{\mathrm{RMS}}$ and bias $\overline{\Delta(q^2)}$ of the reconstructed q^2 spectrum for $B \to X_c \ell \nu_\ell$ signal decays before and after applying the kinematic fit in bins of $E_{\mathrm{miss}} - |\boldsymbol{p}_{\mathrm{miss}}|$, $N_{\mathrm{charged}} \in X$ and $N_\gamma \in X$. The kinematic fit yields significant improvements of the q^2 resolution and bias in nearly all bins. However, a slight dependence of the q^2 resolution and bias on the region of phase space is still observed.

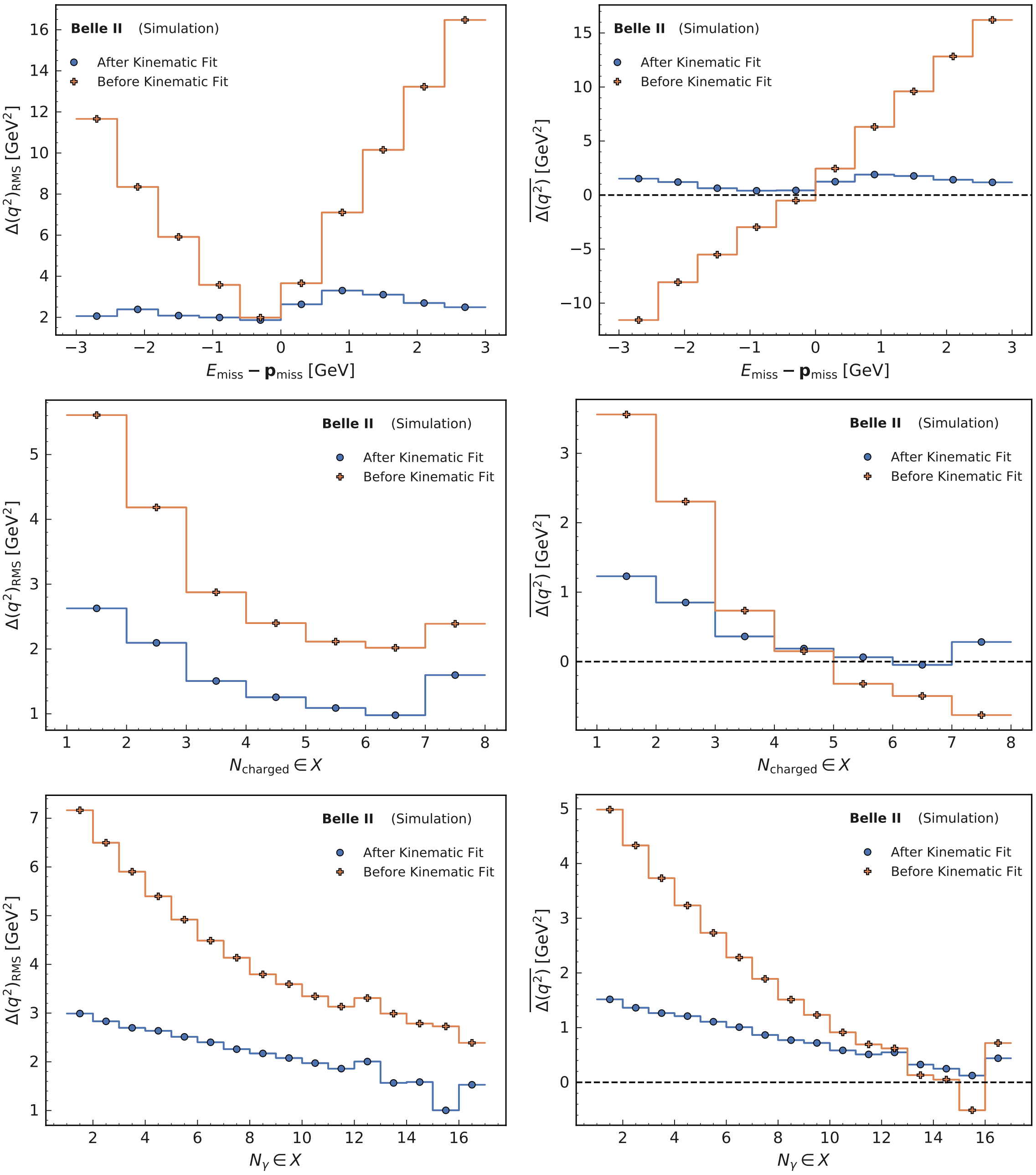


Figure 4.8: Resolution $\Delta(q^2)_{\text{RMS}}$ (left) and bias $\overline{\Delta(q^2)}$ (right) of the reconstructed q^2 spectrum for $B \to X_c \ell \nu_\ell$ signal decays before and after applying the kinematic fit.

CHAPTER 5

Measurement of q^2 Moments

This chapter describes the measurement of raw and central q^2 moments of inclusive semileptonic $B \to X_c \ell \nu_\ell$ decays. The raw moments $\langle q^{2n} \rangle$ of order $n = 1$–4 are measured as functions of lower thresholds of q^2 itself. In addition, the central moments $\langle (q^2 - \langle q^2 \rangle)^n \rangle$ with $n = 2$–4 are calculated directly from measured raw q^2 moments.

The measurement is separated into two steps:

1. Determination of the background normalization with a binned likelihood fit to the m_X distribution and construction of a signal probability weight function $w(q^2)$ using the measured q^2 spectrum and the q^2 distribution of the expected remaining background.
2. Application of an event-wise calibration $q^{2n}_{\mathrm{reco}} \to q^{2n}_{\mathrm{calib}}$ correcting detector resolution and selection efficiency effects distorting the moments of the reconstructed q^2 spectrum.

The background subtracted q^2 moment of order n is calculated as a weighted mean of the calibrated q^{2n}_{calib} values as

$$\langle q^{2n} \rangle = \frac{\sum_i^{N_{\mathrm{data}}} w(q_i^2) \times q^{2n}_{\mathrm{calib},i}}{\sum_j^{N_{\mathrm{data}}} w(q_j^2)} \times C_{\mathrm{calib}} \times C_{\mathrm{gen}}, \tag{5.1}$$

with sums over all events. The weight $w(q_i^2)$ assigns a signal probability to each event calculated from a continuous function of the reconstructed q^2 values. Further, the two additional factors are applied to correct for remaining biases. The factor C_{calib} corrects for imperfections in the calibration procedure, while C_{gen} corrects an intrinsic bias of the q^2 moments due to differences in detector acceptance and selection efficiencies of exclusive $B \to X_c \ell \nu_\ell$ final-states. The analysis strategy is adapted from the measurements of the hadronic mass m_X and combined hadronic mass-and-energy n_X^2 moments performed by the BaBar collaboration [107–109].

Both steps, the background subtraction and the calibration of the q^2 moments, heavily depend on MC simulations of signal and background processes. A reliable modeling of the remaining background in the q^2 distribution is crucial in order to not under- or overestimate the signal contribution. Further, the simulation of the higher mass X_c states has a large impact on the shape of the $B \to X_c \ell \nu_\ell$ q^2 spectrum, especially at low q^2 values. To ensure reliable measurements, the smallest lower q^2 threshold considered in this analysis is $q^2 > 1.5\,\mathrm{GeV}^2$. The decision for this value is based two key arguments:

1. Comparisons between measured and simulated q^2 spectra for different lower q^2 thresholds.

2. A study of the impact of basic selection requirements on the efficiencies of different X_c final-states for different lower q^2 thresholds.

The remainder of this chapter is structured as follows: Section 5.1 discusses the details and the results of the binned likelihood fit to the m_X distribution. In addition, it explains the determination of the signal probability function $w(q^2)$. The concept of the event-wise q^2 calibration is introduced in Section 5.2. Ensemble tests of the analysis procedure based on simulated data samples are performed in Section 5.3. The estimation of statistical and systematic uncertainties as well as their correlations are summarized in Sections 5.4 and 5.5. Section 5.6 presents the results of this measurement: the extracted raw and central q^2 moments. Lastly, final stability checks of the measurement are discussed in Section 5.7.

5.1 Background Subtraction

5.1.1 Continuum Yield Estimation using Off-Resonance Data

The contribution from $e^+e^- \longrightarrow q\bar{q}$ continuum processes in the selected sample are constrained using a data set recorded at a CM energy of $\sqrt{s} = 10.52\,\mathrm{GeV}$ corresponding to an integrated luminosity of $9.2\,\mathrm{fb}^{-1}$. This data set mainly consists of $e^+e^- \longrightarrow q\bar{q}$ events and is referred to as off-resonance data in the following.

Due to the difference of 60 MeV in the CM energies in the on and off-resonance data samples, the m_{bc} and ΔE event-selection criteria have to be adjusted accordingly. For the beam-constrained mass m_{bc}, the endpoint of the spectrum for off-resonance data is shifted by $\Delta m_{\mathrm{bc}} \approx 30\,\mathrm{MeV}$. Therefore, the off-resonance m_{bc} values are shifted by Δm_{bc} prior to applying the selection $m_{\mathrm{bc}} > 5.27\,\mathrm{GeV}$. To translate the ΔE selection applied to the on-resonance sample into requirements for the off-resonance data set, the ΔE values are transformed in a slightly different way by calculating

$$\Delta E' = \Delta E - \mathrm{mode}(\Delta E), \tag{5.2}$$

with $\mathrm{mode}(\Delta E)$ denoting the mode of the ΔE distribution. For the on and off-resonance data sample $\mathrm{mode}(\Delta E)$ is found to be $-4\,\mathrm{MeV}$ and $22\,\mathrm{MeV}$, respectively. The translated ΔE selection corresponds to $-0.146\,\mathrm{GeV} < \Delta E' < 0.104\,\mathrm{GeV}$. Variations of the m_{bc} shift and transformation and selection for $\Delta E'$ are studied but the changes on the number of selected off-resonance events is negligible compared to the statistical uncertainty.

The number of continuum events in the on-resonance sample $\hat{N}^{q\bar{q}}_{\text{on-res.}}$ is estimated by scaling the observed number of events in the off-resonance sample according to

$$\hat{N}^{q\bar{q}}_{\text{on-res.}} = \mu_{\text{off-res.}} N_{\text{off-res.}}. \tag{5.3}$$

The scaling factor $\mu_{\text{off-res.}}$ is given by

$$\mu_{\text{off-res.}} = \frac{\mathcal{L}_{\text{on-res.}}}{\mathcal{L}_{\text{off-res.}}} \times \frac{s_{\text{off-res.}}}{s_{\text{on-res.}}}, \tag{5.4}$$

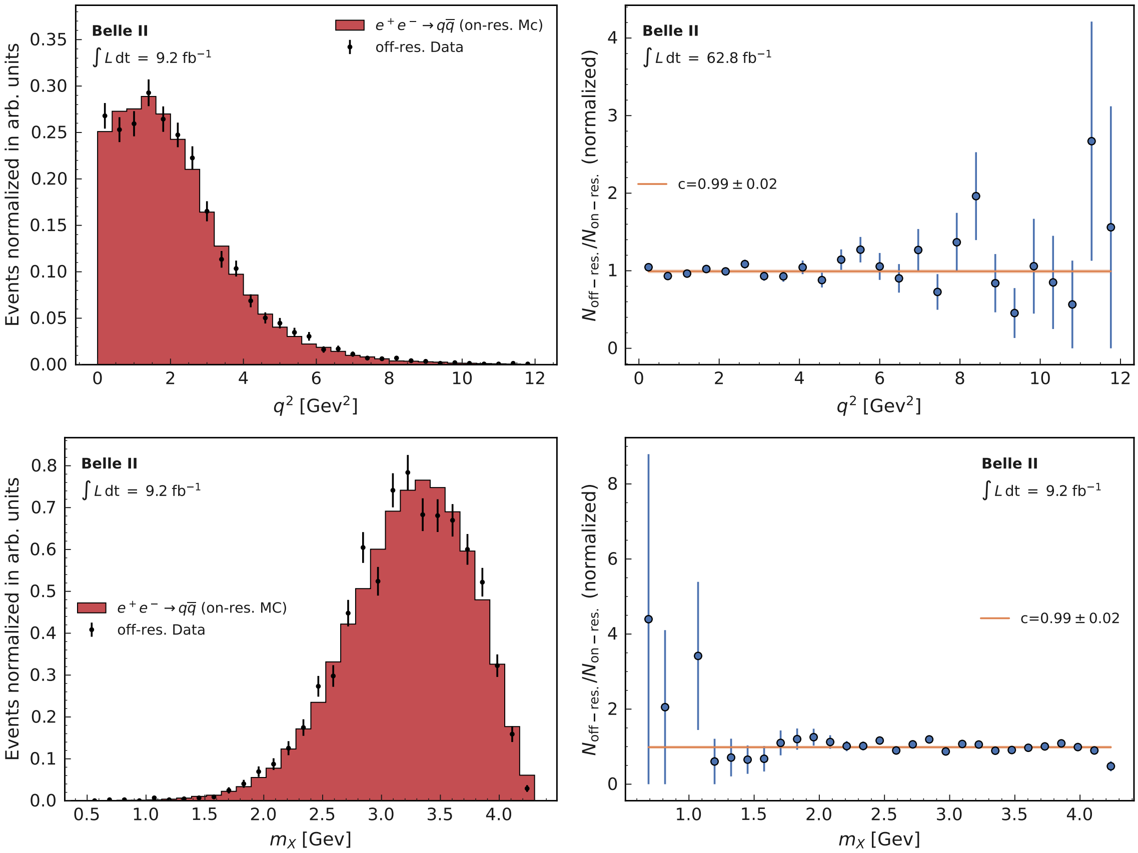


Figure 5.1: Normalized q^2 (top) and m_X (bottom) distributions for on-resonance $e^+e^- \to q\bar{q}$ MC and off-resonance data and their bin-wise ratio.

and takes into account the difference in the integrated luminosity of both data samples and the energy dependence of the $e^+e^- \to q\bar{q}$ production cross section. The statistical uncertainty is calculated via error-propagation

$$\sigma_{\hat{N}^{q\bar{q}}_{\text{on-res.}}} = \mu_{\text{off-res.}} \sqrt{N_{\text{off-res.}}}. \tag{5.5}$$

Figure 5.1 shows a comparison of the normalized q^2 and m_X distributions for on-resonance $e^+e^- \to q\bar{q}$ MC and off-resonance data. The q^2 and m_X spectra are calculated from the resulting four-vectors of the kinematic fit. To check whether the small difference in the CM energy affects the q^2 and m_X values, a fit of $y = c$ to the bin-wise ratio $n_{\text{off-res.}}/n_{\text{on-res.}}$ is performed. The ratios are compatible with unity, showing a good agreement between the shapes of on-resonance MC and off-resonance data in both distributions.

Table 5.1 summarizes the estimated numbers of continuum events in the on-resonance data sample $\hat{N}^{q\bar{q}}_{\text{on-res.}}$ and their statistical uncertainty $\sigma_{\hat{N}^{q\bar{q}}_{\text{on-res.}}}$ in the three $B\ell$ reconstruction channels for different lower q^2 thresholds. They are used as external input in the likelihood fit to the m_X distribution to constrain the $e^+e^- \to q\bar{q}$ contributions.

Table 5.1: Estimated number of continuum events $\hat{N}^{q\bar{q}}_{\text{on-res.}}$ and their statistical uncertainty $\sigma_{\hat{N}^{q\bar{q}}_{\text{on-res.}}}$ in the three $B\ell$ reconstruction channels for different lower q^2 thresholds.

	$B^+\ell^-$		$B^0\ell^-$		$B^0\ell^+$	
q^2_{th} [GeV^2]	$\hat{N}^{q\bar{q}}_{\text{on-res.}}$	$\sigma_{\hat{N}^{q\bar{q}}_{\text{on-res.}}}$	$\hat{N}^{q\bar{q}}_{\text{on-res.}}$	$\sigma_{\hat{N}^{q\bar{q}}_{\text{on-res.}}}$	$\hat{N}^{q\bar{q}}_{\text{on-res.}}$	$\sigma_{\hat{N}^{q\bar{q}}_{\text{on-res.}}}$
0.0	12964.61	295.80	5520.59	193.02	5338.37	189.81
0.5	11257.14	275.63	4865.95	181.22	4501.51	174.30
1.0	9657.65	255.30	4137.07	167.09	3846.86	161.13
1.5	7902.94	230.95	3279.96	148.78	3111.24	144.90
2.0	6161.73	203.92	2598.32	132.42	2328.36	125.35
2.5	4609.49	176.38	1808.70	110.48	1774.96	109.45
3.0	3131.48	145.38	1316.03	94.24	1228.30	91.05
3.5	2247.38	123.16	1039.33	83.75	911.10	78.41
4.0	1565.74	102.80	755.88	71.42	654.64	66.47
4.5	1181.05	89.28	587.15	62.95	445.43	54.83
5.0	863.86	76.35	404.93	52.28	330.70	47.24
5.5	647.89	66.13	276.70	43.21	242.96	40.49
6.0	452.18	55.24	161.97	33.06	182.22	35.07
6.5	357.69	49.13	107.98	27.00	128.23	29.42
7.0	242.96	40.49	74.24	22.38	94.48	25.25
7.5	175.47	34.41	53.99	19.09	74.24	22.38
8.0	148.48	31.66	40.49	16.53	53.99	19.09
8.5	101.23	26.14	20.25	11.69	40.49	16.53

5.1.2 Fit of the m_X Distribution

The overall background normalization is determined with a template likelihood fit. The template fit uses shape differences of simulated signal and background processes in the binned distribution of a discriminating variable to extract the yields $\boldsymbol{n}$ of signal and background.

The statistical model assumes that the number of observed events d_i in a bin i of the discriminating variable follows Poisson distribution

$$\mathcal{P}(d_i|\nu_i) = \frac{\nu_i^{d_i}}{d_i!}\mathrm{e}^{-\nu_i}, \tag{5.6}$$

with ν_i denoting the expected number of events in bin i. The likelihood function is defined as the product of independent Poisson probabilities for all bins i

$$L(\boldsymbol{n}) = \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}(d_i|\nu_i(\boldsymbol{n})), \tag{5.7}$$

and only depends on the different process yields $\boldsymbol{n}$. For N_p contributing processes, the number of expected events in bin i is given by

$$\nu_i(\boldsymbol{n}) = \sum_{k=1}^{N_p} n_k f_{ik}, \tag{5.8}$$

with f_{ik} being the fraction of events of process k reconstructed in bin i. These fractions are estimated as the ratio

$$f_{ik} = \frac{h_{ik}}{\sum_j^{N_\text{bins}} h_{jk}}, \tag{5.9}$$

where h_{ik} denotes the number of reconstructed events of process k in bin i obtained from MC simulation.

Template shape uncertainties are directly incorporated into the likelihood function. This is achieved by introducing nuisance parameters θ_{ik} for each template k and bin i. The fraction of events f_{ik} in Equation (5.8) is replaced by

$$f_{ik} = \frac{h_{ik} + \sigma_{ik}\theta_{ik}}{\sum_j h_{jk} + \sigma_{jk}\theta_{jk}}. \tag{5.10}$$

Here, σ_{ik} denotes the uncertainty of the template histogram h_{ik}. The nuisance parameters are constrained by a multivariate Gaussian distribution $\mathcal{N}(\boldsymbol{\theta}_k|\boldsymbol{0}, R_k)$ with mean value $\boldsymbol{0}$ and R_k being the correlation matrix of h_{ik}. The correlation matrix is obtained from the total covariance matrix C_k of template k. Here, the total covariance matrix C_k is given by the sum of all considered covariance matrices describing systematic shape uncertainties of template k.

Additionally, a given yield parameter n_k can be constrained to an external measurement $\tilde{n}_k$ with uncertainty $\sigma_{\tilde{n}_k}$ with an univariate Gaussian distribution $\mathcal{N}(n_k|\tilde{n}_k, \sigma_{\tilde{n}_k})$. Including shape uncertainties and constraints on yield parameters, the likelihood function takes the form

$$L(\boldsymbol{n}, \boldsymbol{\theta}) = \prod_{i=1}^{N_\text{bins}} \mathcal{P}(d_i|\nu_i(\boldsymbol{n}, \boldsymbol{\theta})) \times \prod_{k=1}^{N_p} \mathcal{N}(\boldsymbol{\theta}_k|\boldsymbol{0}, R_k) \times \prod_{l=1}^{N_\text{constr.}} \mathcal{N}(n_l|\tilde{n}_l, \sigma_{\tilde{n}_l}). \tag{5.11}$$

The binned likelihood fit of the m_X distribution is carried out separately for the $B^+\ell^-$, $B^0\ell^-$, and $B^0\ell^+$ reconstruction channels to account for efficiency differences in the FEI algorithm. Electron and muon channels are not separated in the fit. The fits are performed for lower q^2 thresholds from $q^2 > 0.0\,\text{GeV}^2$ up to $q^2 > 8.5\,\text{GeV}^2$ in $0.5\,\text{GeV}^2$ increments. To reduce the dependence on the modeling of signal and background templates, the fit is carried out in five equidistant m_X bins only. For each channel and q^2 threshold an appropriate binning is chosen to cover the range of all observed q^2 values in data and MC. The likelihood function is numerically optimized with respect to the yield $\boldsymbol{n}$ and nuisance parameters $\boldsymbol{\theta}$ using the `Minuit` algorithm [110] implemented in `scikit-hep/iminuit` [111].

The fit distinguishes between the following three processes (compare to Section 4.2):

1. $B \to X_c \ell \nu_\ell$ signal (with yield n_{X_c}),
2. $B\overline{B}$ background dominated by secondary leptons and hadronic B decays with K or π candidates misidentified as signal leptons ($n_{B\overline{B}}$), and

Table 5.2: Systematic uncertainties included in the m_X fit. The markers ✓ and ✗ indicate whether a template is affected by a uncertainty source or not, respectively.

Source	$B \to X_c \ell \nu_\ell$ Signal	$B\overline{B}$ Background	$e^+e^- \to q\overline{q}$ Continuum
MC sample size	✓	✓	✓
$\mathcal{B}(B \to X_c \ell \nu_\ell)$	✓	✗	✗
$B \to D^{(*)} \ell \nu_\ell$ form factor	✓	✗	✗
$\mathcal{B}(B \to X_u \ell \nu_\ell)$	✗	✓	✗
Tracking efficiency	✓	✓	✓
PID efficiency & fake rate	✓	✓	✓

3. $e^+e^- \to q\overline{q}$ continuum processes (n_{qq}). The yields of this process are constrained to the values given in Table 5.1.

For each template, covariance matrices describing systematic uncertainties are constructed in three ways starting from a vector $\boldsymbol{\sigma}$ quantifying the uncertainty in each template bin:

1. For a source of uncertainty that is assumed to be uncorrelated, the covariance matrix is given by the diagonal matrix $C = \mathrm{diag}(\boldsymbol{\sigma}^2)$.
2. The covariance matrix for fully (anti)-correlated uncertainties is constructed using the outer product $C = \boldsymbol{\sigma} \otimes \boldsymbol{\sigma}$. The vector $\boldsymbol{\sigma}$ contains the bin-wise uncertainties and is calculated from varied histograms $h_i^{\mathrm{up/down}}$ representing variations within one standard deviation as

$$\sigma_i = \frac{h_i^{\mathrm{up}} - h_i^{\mathrm{down}}}{2}. \tag{5.12}$$

3. The covariance matrix is estimated using the sample covariance of an ensemble of varied template histograms h'_k

$$C_{ij} = \frac{1}{N-1} \sum_{k=1}^{N} (h'_{ki} - \bar{h}'_i)(h'_{kj} - \bar{h}'_j) \tag{5.13}$$

   with k and N denoting the variation and ensemble size, respectively.

The varied histograms are determined by modifying the overall event weight. Here, the nominal weight of a certain systematic correction is replaced with its varied value as

$$w_{\mathrm{evt}} \to \begin{cases} w_{\mathrm{evt}} \times \frac{w_{\mathrm{sys}}^{\mathrm{up/down}}}{w_{\mathrm{sys}}^{\mathrm{nom}}} & \text{or,} \\ w_{\mathrm{evt}} \times \frac{w'_{\mathrm{sys}}}{w_{\mathrm{sys}}^{\mathrm{nom}}} \end{cases} \tag{5.14}$$

A list of all systematic uncertainties included in the m_X fit is given in Table 5.2. The only source of uncorrelated uncertainties considered is statistical uncertainty related the size of the simulated samples used to construct the templates. Here, the uncertainty for bin i is given by

$$\sigma_i = \sqrt{\sum_{l \in i} w_l^2}, \tag{5.15}$$

with w_l denoting the weight of event l falling into bin i. All templates are affected by the uncertainty on the MC sample size. The composition uncertainties of $B \to X_c \ell \nu_\ell$ decays are determined with the branching fraction uncertainties listed in Table 3.1. The BGL form factor parameter uncertainties for $B \to D \ell \nu_\ell$ and $B \to D^* \ell \nu_\ell$ decays are evaluated using a orthogonal set of parameter variations for each decay (see Section 3.4.2). For the $B \to X_u \ell \nu_\ell$ component included in the $B\overline{B}$ background template, a uncertainty of 14% is assigned to the branching fraction [23]. The track selection efficiency is studied in [112] using τ-pair events. The study suggests to assign an uncertainty of 0.69% per track on the signal-side. The lepton identification efficiency and hadron-to-lepton fake rate uncertainties are estimated using ensembles of corrections varied within their respective statistical and systematic uncertainties (see Section 3.3.3).

The pre- and post-fit m_X distributions in the three reconstruction channels of the fits for $q^2 > 0.0\,\mathrm{GeV}^2$ are presented in Figure 5.2. For the post-fit projections, the signal and background templates are scaled to the fitted event yields and the shapes are transformed according to the post-fit values of the nuisance parameters. For higher values of the q^2 thresholds, the pre-fit and post-fit m_X distributions are shown Figures A.1 to A.17 in Appendix A. The uncertainty bands in the pre-fit and post-fit distributions are obtained in a slightly different way. For the pre-fit distribution, the uncertainties of the different templates are considered uncorrelated. Hence the uncertainties of the three templates for a bin i are added in quadrature

$$\sigma_i^{\text{pre-fit}^2} = \sum_k^{N_p} \sigma_{ik}^{\text{pre-fit}^2}. \tag{5.16}$$

As a result of the likelihood fit, the nuisance parameters $\boldsymbol{\theta}$ are constrained which can be seen by the reduced uncertainties $\sigma_{\theta_{ik}}^{\text{post-fit}}$ compared to the uncertainty of the pre-fit expectation of $\sigma_{\theta_{ik}}^{\text{pre-fit}} = 1$. This can be translated to a post-fit uncertainty in bin i for template k by scaling the pre-fit uncertainty according to

$$\sigma_{ik}^{\text{post-fit}} = \sigma_{\theta_{ik}}^{\text{post-fit}} \times \sigma_{ik}^{\text{pre-fit}}. \tag{5.17}$$

Taking the correlations between the nuisance parameters of different templates into account, the post-fit uncertainty for bin i is given by

$$\sigma_i^{\text{post-fit}^2} = \sum_k^{N_p} \sigma_{ik}^{\text{post-fit}^2} + \sum_k^{N_p} \sum_{l(l \neq k)}^{N_p} \rho_{\theta_{ik},\theta il} \sigma_{ik}^{\text{post-fit}} \sigma_{il}^{\text{post-fit}}. \tag{5.18}$$

Figure 5.3 shows the post-fit values and uncertainties of the bin-wise nuisance parameters $\boldsymbol{\theta}$ as well as their pre-fit expectation for $q^2 > 0.0\,\mathrm{GeV}^2$. Nearly all nuisance parameters are constrained and pulled away from the pre-fit expectation except in the first m_X bin. The nuisance parameters of the $B\overline{B}$ background template exhibit the largest pulls by ≈ 1-σ with respect to their post-fit uncertainty. In the region $m_X > 2.5\,\mathrm{GeV}$ the fit is also able to constrain the $B \to X_c \ell \nu_\ell$ nuisance parameters. This is not unexpected since the $B \to X_c \ell \nu_\ell$ composition uncertainties for the higher mass states is rather conservative by assigning a 100% uncertainty on the non-resonant part of the X_c spectrum. The continuum template post-fit nuisance parameters also show a positive pull, but they are only slightly constrained by the fit. The resulting post-fit nuisance parameter values and uncertainties for higher q^2 threshold values are shown in Figures A.18 to A.34. To quantify the agreement between measured and post-fit m_X distribution, the test statistic

$$\chi_P^2 = 2 \sum_i^{N_\text{bins}} d_i \log \frac{d_i}{\nu_i(\hat{\boldsymbol{n}}, \hat{\boldsymbol{\theta}})} + \nu_i(\hat{\boldsymbol{n}}, \hat{\boldsymbol{\theta}}) - d_i \tag{5.19}$$

suggested in [113] is used. For Poisson distributed data, χ_P^2 follows a χ^2 distribution with $N_\text{bins} - m$ degrees of freedom. Here, m denotes the number of free fit parameters, which are the $B \to X_c \ell \nu_\ell$ signal and $B\overline{B}$ background yields. In this case, the degrees of freedom is three. Overall, the post-fit m_X distributions describe the measured spectrum adequately with observed p-values ranging from 0.09 to 0.98. the fitted signal and background yields as well as the p-values of the post-fit m_X distributions for all reconstruction channels and lower q^2 thresholds are summarized in Tables 5.3 to 5.5.

The obtained $e^+e^- \to q\overline{q}$ and $B\overline{B}$ background yields are used in the calculation of the binned signal probability $w_i(q^2)$ as explained in Section 5.1.4. In addition, the resulting $B \to X_c \ell \nu_\ell$ signal yields are used to correct the contribution of the three reconstruction channels to the overall q^2 signal spectrum for the q^2 calibration introduced in Section 5.2.

5.1.3 Comparison of Measured and Simulated q^2 Spectrum

For the background subtraction and calibration, a reliable description of the q^2 spectrum in MC essential. The calibration of the q^2 spectrum depends on the composition of the $B \to X_c \ell \nu_\ell$ spectrum as well as the kinematic properties of the simulated X_c components. The background subtraction is based on the simulated shapes and the fitted yields of the remaining background components. To study the agreement between the measured and simulated q^2 distribution, a two-sample χ^2 and Kolmogorov-Smirnov (KS) test are used. Here, the simulated signal and background components are normalized to the yields obtained from the m_X fit. The χ^2 test statistic is defined as

$$\chi^2 = \boldsymbol{r}^\top (C_\text{data} + C_\text{MC})^{-1} \boldsymbol{r}, \tag{5.20}$$

with the residuals $\boldsymbol{r} = \boldsymbol{d} - \boldsymbol{\nu}_\text{MC}$ denoting the bin-wise difference between the measurement and simulation. The covariance describing the measured spectrum is given by the diagonal matrix

$$C_\text{data} = \text{diag}(\boldsymbol{d}). \tag{5.21}$$

assuming each bin follows an independent Poisson distribution. For the simulated distribution, the covariance matrix is calculated as

$$C_\text{MC} = C_{X_c} + C_{B\overline{B}} + C_{q\overline{q}}. \tag{5.22}$$

The individual covariance matrices for the MC components include the statistical MC uncertainty and the normalization uncertainty. Again, the covariance matrix describing the statistical MC uncertainty is

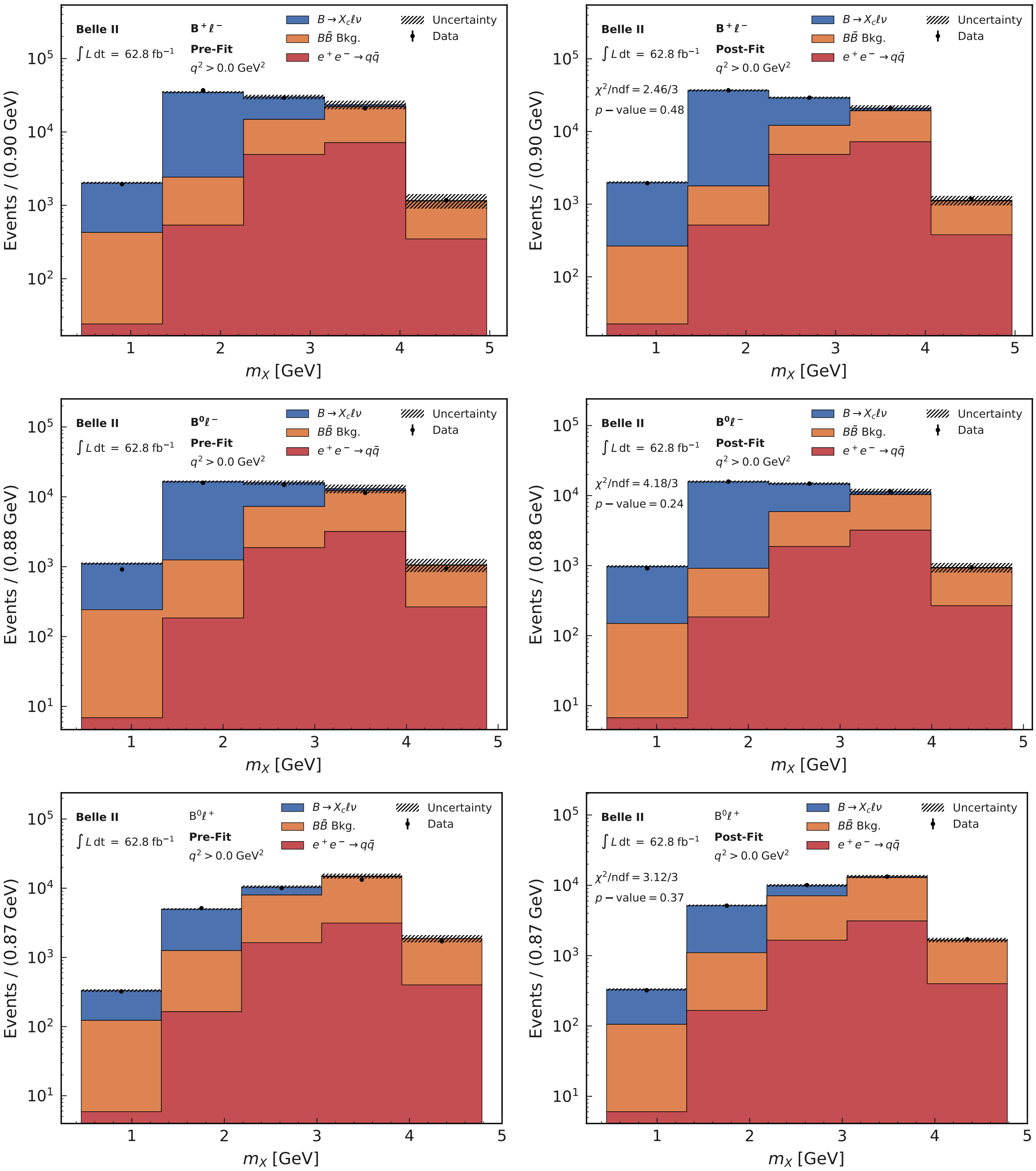


Figure 5.2: m_X distribution for a $q^2 > 0.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

Table 5.3: Resulting number of signal ($\hat{n}_{X_c}$), $B\overline{B}$ background ($\hat{n}_{B\overline{B}}$), and continuum ($\hat{n}_{q\bar{q}}$) events and their respective uncertainties $\sigma_{\hat{n}}$ after the m_X fit for different lower q^2 thresholds in the $B^+\ell^-$ channel.

q^2_{th} [GeV^2]	$\hat{n}_{X_c}$	$\sigma_{\hat{n}_{X_c}}$	$\hat{n}_{B\overline{B}}$	$\sigma_{\hat{n}_{B\overline{B}}}$	$\hat{n}_{q\bar{q}}$	$\sigma_{\hat{n}_{q\bar{q}}}$	χ^2/ndf	p-value
0.0	55678.61	716.99	21623.51	755.93	12967.45	297.89	0.82	0.48
0.5	53465.11	792.29	19426.65	815.88	11254.58	275.43	0.62	0.61
1.0	51043.97	702.25	16621.96	718.91	9654.57	254.78	0.89	0.44
1.5	49011.08	652.85	13377.59	662.75	7914.62	230.20	0.70	0.55
2.0	46776.49	555.98	10179.78	562.85	6182.64	202.83	1.53	0.21
2.5	44902.41	517.30	7093.41	523.57	4630.98	175.21	1.62	0.18
3.0	42624.17	613.72	5104.01	602.08	3153.51	143.59	1.37	0.25
3.5	40408.10	607.21	3398.64	587.25	2264.64	122.13	1.35	0.26
4.0	37694.21	527.49	2399.07	493.48	1573.56	101.94	0.84	0.47
4.5	34583.56	437.39	1878.51	401.87	1189.67	88.80	1.09	0.35
5.0	31617.73	369.48	1466.40	327.83	871.18	75.55	1.02	0.38
5.5	28640.61	320.20	1256.19	270.61	658.59	65.50	1.48	0.22
6.0	25640.76	268.95	1108.53	218.18	457.79	54.65	1.36	0.25
6.5	22864.43	231.43	811.63	173.53	363.99	48.50	2.14	0.09
7.0	20083.54	200.95	738.01	147.67	248.95	39.89	2.29	0.08
7.5	17475.12	181.45	605.46	122.78	182.36	34.22	2.10	0.10
8.0	14951.29	161.11	490.39	102.85	153.54	31.54	2.03	0.11
8.5	12702.71	142.42	351.50	85.35	104.32	26.12	2.07	0.10

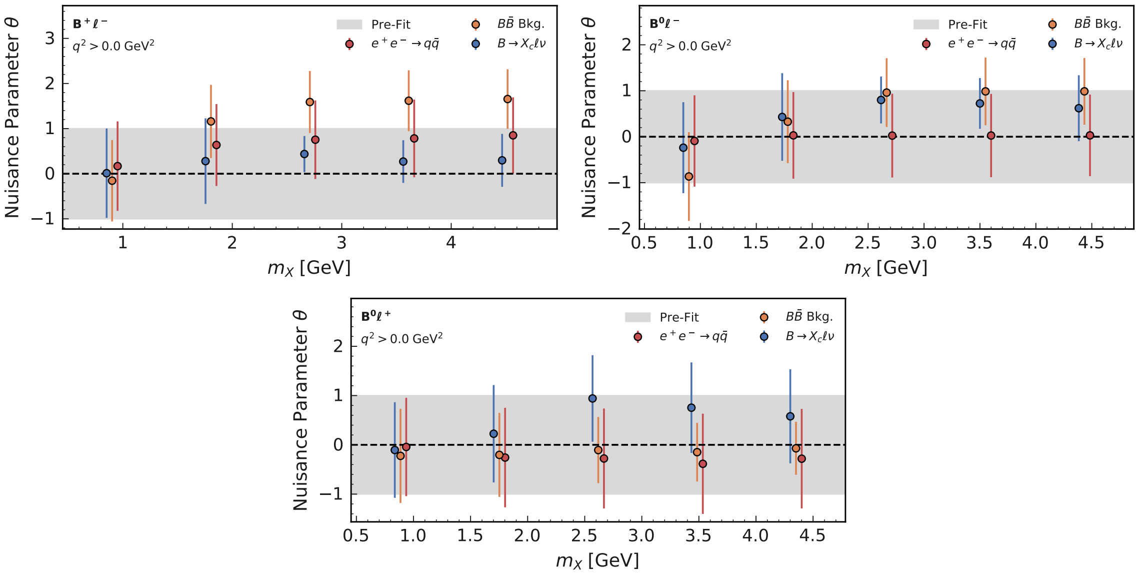


Figure 5.3: Bin-wise nuisance parameters after the m_X fit for $q^2 > 0.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

Table 5.4: Resulting number of signal ($\hat{n}_{X_c}$), $B\overline{B}$ background ($\hat{n}_{B\overline{B}}$), and continuum ($\hat{n}_{q\bar{q}}$) events and their respective uncertainties $\sigma_{\hat{n}}$ after the m_X fit for different lower q^2 thresholds in the $B^0\ell^-$ channel.

q^2_{th} [GeV^2]	$\hat{n}_{X_c}$	$\sigma_{\hat{n}_{X_c}}$	$\hat{n}_{B\bar{B}}$	$\sigma_{\hat{n}_{B\bar{B}}}$	$\hat{n}_{q\bar{q}}$	$\sigma_{\hat{n}_{q\bar{q}}}$	χ^2/ndf	p-value
0.0	25701.18	478.22	12713.74	510.67	5541.91	193.20	1.39	0.24
0.5	25109.56	476.06	11007.47	499.06	4876.41	181.41	1.63	0.18
1.0	23780.24	449.38	9698.34	465.46	4151.31	167.50	1.76	0.15
1.5	22742.38	381.54	7987.38	393.37	3301.96	148.71	2.25	0.08
2.0	21575.96	345.27	6209.61	352.70	2622.77	132.36	1.85	0.14
2.5	20662.26	323.24	4600.20	325.32	1833.93	110.25	1.77	0.15
3.0	19866.01	368.69	2935.69	363.32	1340.53	80.36	1.16	0.32
3.5	18941.05	386.26	1756.97	372.19	1056.29	82.88	0.35	0.79
4.0	17504.15	367.54	1343.58	350.37	771.83	70.42	0.41	0.75
4.5	16072.79	302.21	1121.39	280.71	599.08	61.76	0.15	0.93
5.0	14754.94	257.02	846.60	228.81	409.75	51.46	0.09	0.96
5.5	13489.93	222.21	651.52	191.02	284.39	42.48	0.25	0.86
6.0	12246.78	192.99	479.80	158.73	166.77	32.43	0.42	0.74
6.5	10951.19	173.87	414.03	138.73	113.86	26.65	0.59	0.62
7.0	9712.20	147.99	281.24	108.35	74.50	22.25	1.08	0.35
7.5	8520.35	131.11	220.35	91.95	54.24	18.86	1.55	0.20
8.0	7325.38	119.91	199.46	79.07	41.95	16.55	1.89	0.13
8.5	6244.66	102.56	147.81	65.29	20.51	11.51	2.32	0.07

constructed as diagonal matrix with variances calculated according to Equation (5.15). The covariance matrix for the normalization uncertainty treats the bins as fully correlated and the uncertainty is estimated according to Equation (5.12). Further, the impact of the different shape uncertainties listen in Table 5.2 on the simulated q^2 spectrum are estimated with normalized histograms. Here, the uncertainties are estimated analogous to Equations (5.12) and (5.13) but the varied histograms are normalized to the number of events without the variation. To achieve this, the replacement in Equation (5.14) is modified according to

$$w_{\text{evt}} \rightarrow \begin{cases} w_{\text{evt}} \times \dfrac{w_{\text{sys}}^{\text{up/down}}}{w_{\text{sys}}^{\text{nom}}} \times A^{\text{up/down}} & \text{or,} \\ w_{\text{evt}} \times \dfrac{w'_{\text{sys}}}{w_{\text{sys}}^{\text{nom}}} \times A', \end{cases} \tag{5.23}$$

with a constant factor $A^{\text{up/down}}$ or A' such that

$$\sum_{i}^{N_{\text{bins}}} h_i = \begin{cases} \sum_i^{N_{\text{bins}}} h_i^{\text{up/down}} & \text{or,} \\ \sum_i^{N_{\text{bins}}} h'_i. \end{cases} \tag{5.24}$$

The two-sample KS test statistic is defined as

$$D_{nm} = \max |F_{1,n}(x) - F_{2,m}(x)|, \tag{5.25}$$

Table 5.5: Resulting number of signal ($\hat{n}_{X_c}$), $B\overline{B}$ background ($\hat{n}_{B\overline{B}}$), and continuum ($\hat{n}_{q\overline{q}}$) events and their respective uncertainties $\sigma_{\hat{n}}$ after the m_X fit for different lower q^2 thresholds in the $B^0\ell^+$ channel.

q^2_{th} [GeV^2]	$\hat{n}_{X_c}$	$\sigma_{\hat{n}_{X_c}}$	$\hat{n}_{B\overline{B}}$	$\sigma_{\hat{n}_{B\overline{B}}}$	$\hat{n}_{q\overline{q}}$	$\sigma_{\hat{n}_{q\overline{q}}}$	χ^2/ndf	p-value
0.0	7649.03	338.41	17542.35	411.35	5342.08	189.77	1.04	0.37
0.5	7352.43	310.64	15496.36	379.25	4498.66	173.97	0.63	0.60
1.0	6956.97	285.56	12954.17	347.78	3840.76	161.02	1.21	0.31
1.5	6660.47	192.18	10329.54	261.38	3109.60	144.84	1.35	0.26
2.0	6369.33	242.97	7943.13	286.13	2327.65	125.18	1.40	0.24
2.5	5999.92	195.75	5869.96	235.32	1772.43	109.25	0.57	0.63
3.0	5669.57	185.07	4308.89	213.91	1222.47	90.77	1.24	0.29
3.5	5330.52	193.87	3088.02	213.30	894.83	77.96	1.29	0.27
4.0	5003.91	183.36	2275.35	196.80	648.60	66.14	0.33	0.81
4.5	4475.04	158.20	1905.09	167.38	446.68	54.58	0.50	0.68
5.0	4078.68	161.38	1486.69	167.22	333.25	46.91	0.36	0.78
5.5	3773.04	163.97	1085.43	167.32	239.97	40.09	0.67	0.57
6.0	3321.09	156.76	935.14	157.60	180.08	34.65	0.53	0.66
6.5	2946.33	165.72	771.40	163.69	126.52	28.97	0.56	0.64
7.0	2448.85	168.96	806.46	164.54	96.37	24.87	0.07	0.98
7.5	2229.63	160.77	560.60	155.77	73.83	21.92	0.06	0.98
8.0	2002.60	135.59	404.35	127.74	51.41	18.93	0.59	0.62
8.5	1650.22	116.48	384.19	109.06	40.58	16.48	0.41	0.75

with $F_{1,n}$ and $F_{2,m}$ denoting the empirical distribution functions for each sample. Contrary to the χ^2 test, the KS test uses an un-binned approach but does not consider systematic uncertainties. The p-value is calculated using the asymptotic KS distribution [114] implemented in `SciPy` [106]. Note that the KS test statistic distribution is evaluated with an effective event number $N_{\mathrm{eff}} = \sqrt{nm/(n+m)}$, with n and m being the number of events observed in data and MC, respectively. Here, the number of events for the MC sample is given by the sum of event weights.

Both tests are repeated for different lower q^2 thresholds. The χ^2 test is performed using twelve bins to retain the Gaussian limit describing the number of events in bins with a low population. Figure 5.4 presents the obtained p-values of the both tests as well as the measured and simulated q^2 spectrum for the lowest threshold of $q^2 > 0.0\,\mathrm{GeV}^2$. For higher values of the q^2 threshold, the measured and simulated q^2 distributions are shown in Figures B.2 to B.4. The p-values obtained by the χ^2 test are all consistently above 5%. The KS test shows p-values below 10^{-3} for the two lowest q^2 thresholds of $q^2 > 0.0\,\mathrm{GeV}^2$ and $q^2 > 0.5\,\mathrm{GeV}^2$. For lower thresholds above $q^2 > 1\,\mathrm{GeV}^2$ the resulting p-values are also above 5%, except for the test result for $q^2 > 2.5\,\mathrm{GeV}^2$. The region $q^2 < 3\,\mathrm{GeV}$ has significant contributions from $B \to X_c \ell \nu_\ell$ with low momenta leptons. For these leptons, the particle identification efficiency and the hadron-to-lepton fake rate have large uncertainties. The χ^2 test considers these uncertainties and results in acceptable p-values. Despite the fact that there are some disagreements between measured and simulated q^2 distribution in the low q^2 region, both distributions are still compatible within the considered uncertainties.

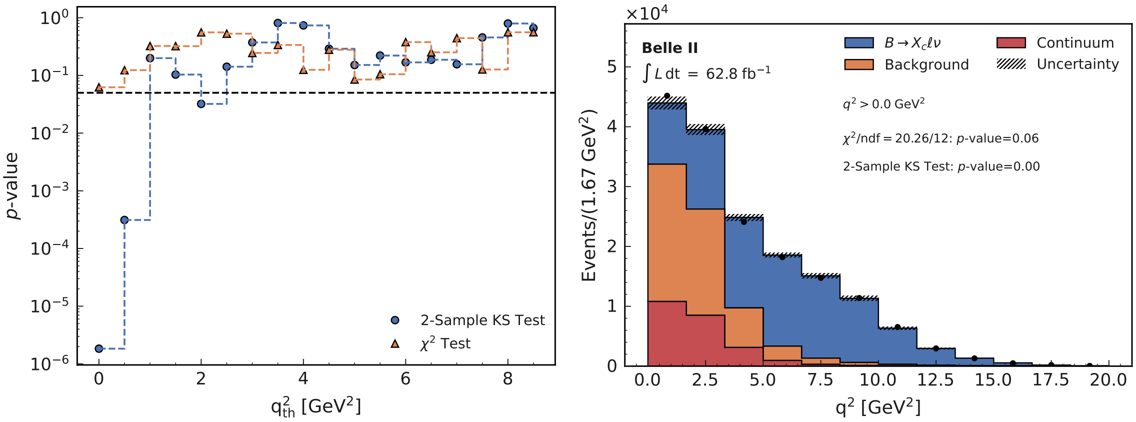


Figure 5.4: Obtained p-values of the χ^2 and KS tests comparing measured and simulated q^2 spectra for different lower q^2 thresholds (left) and the resulting q^2 distribution for the lowest threshold of $q^2 > 0.0\,\mathrm{GeV}^2$ (right).

5.1.4 Background Subtraction with Signal Probability Weights $w(q^2)$

In this analysis, the moments of the q^2 distribution are calculated as a weighted mean according to Equation (5.1). Here, the weight $w_i(q^2)$ represents a signal probability for a single event to contain a $B \to X_c \ell \nu_\ell$ decay. Overall, the weighted mean takes sums over all selected events above a certain q^2 threshold, but the signal probability weights effectively subtract the remaining background directly in the moment calculation.

The weights are calculated with a continuous signal probability function $w(q^2)$, which is determined in two steps:

1. Determine the signal probability in bins of q^2. The bin-wise signal probability is calculated as

$$w_i(q^2) = \frac{d_i - \nu_i(\hat{n}_{B\overline{B}}, \hat{n}_{q\overline{q}})}{d_i}. \tag{5.26}$$

   Here d_i denotes the number of events observed in data in bin i, while $\nu_i(\hat{n}_{B\overline{B}})$ and $\nu_i(\hat{n}_{q\overline{q}})$ are the number of expected $B\overline{B}$ and continuum background events, respectively. The remaining background components are normalized to resulting yields of the m_X fit.

2. A cubic spline is fitted to the bin-wise signal probabilities to determine a continuous description of the latter. The spline function is constructed using the smoothing cubic spline algorithm proposed in [115] and implemented in the `csaps` package [116]. The smoothness of the spline function is controlled by a parameter p. This parameter can be manually chosen within the range $[0, 1]$. $p = 0$ results in a least-squares fit of a straight line to the data, while $p = 1$ gives a cubic spline with exact interpolation.

A single signal probability function is determined for each lower q^2 threshold. Each spline is fitted to the binned signal probability determined in 20 bins of the q^2 distribution. For the nominal estimation of $w(q^2)$, a smoothing parameter $p = 0.1$ is chosen. To retain the probability interpretation, lower and upper boundaries for the spline function are set to 0 and 1, respectively. The impact of different choices for p is discussed in Section 5.5, together with the impact of normalization and shape uncertainties on the remaining background modeling. The estimation of the statistical uncertainty on the signal probability function is explained in Section 5.4.

The validate the background subtraction, the procedure is tested using an Asimov data set and presented in Figure 5.5. The Asimov data set is a toy data set corresponding to the MC expectation. Figure 5.5(a) shows the q^2 distribution of the remaining background ($B\overline{B}$ and continuum) and Asimov data for $q^2 > 0.0\,\mathrm{GeV}^2$ and $q^2 > 3.0\,\mathrm{GeV}^2$. The corresponding binned signal probabilities and the nominal spline fit are presented in Figure 5.5(b). Further the interpolated signal probability is evaluated for each event in the Asimov data set and applied as an additional multiplicative weight. The q^2 distributions of the background subtracted Asimov data are shown in Figure 5.5(c). In addition, the simulated $B \to X_c \ell \nu_\ell$ spectrum is shown in comparison. The background subtraction with the signal probability weights is able to reproduce the expected signal distribution. At low q^2 values, some fluctuations around one in the ratio of background subtracted data and MC expectation are visible. This effect is an artifact of the cubic spline interpolation at the lower boundary of the binned q^2 spectrum. Its impact on the extracted q^2 moments is further discussed in Section 5.3.

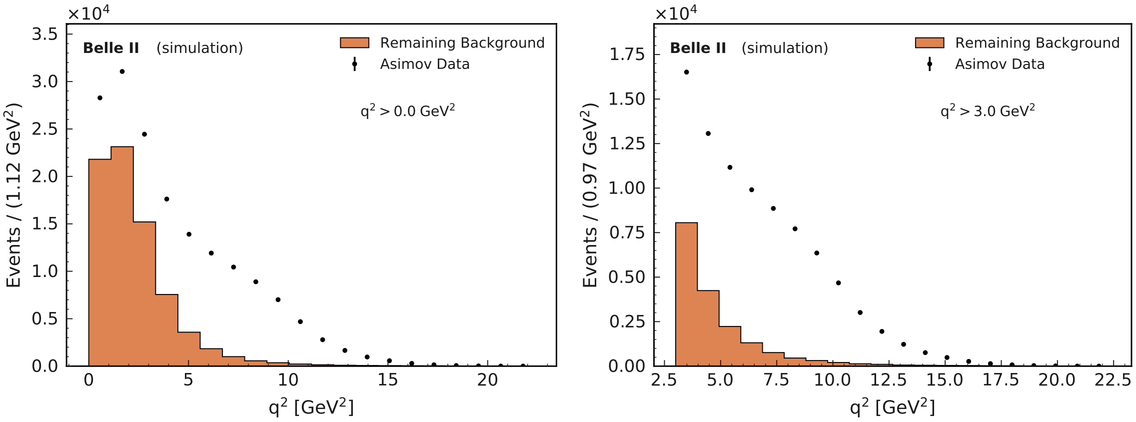
Figure 5.6 shows examples of the background subtraction based on the measured data set for lower q^2 thresholds of $q^2 > 0.0\,\mathrm{GeV}^2$ and $q^2 > 3.0\,\mathrm{GeV}^2$. The q^2 distributions for the remaining background and data are shown in Figure 5.6(a). Here, the MC uncertainty band includes normalization and shape uncertainties. The corresponding binned signal probabilities and the interpolating cubic spline fits are shown in Figure 5.6(b). In addition to the nominal spline fit, the statistical and total uncertainty for the spline fits are shown as the inner and outer uncertainty bands. The uncertainties are estimate using the sample covariance of the respective statistic and systematic variations of the measured sample and background template, respectively. Sections 5.4 and 5.5 discuss the details of both procedures of generation these variations. At low q^2 values, the systematic uncertainty dominates, while the impact of the statistical uncertainty increases with for higher q^2 values. Figure 5.6 shows the background subtracted q^2 spectrum in data and the $B \to X_c \ell \nu_\ell$ MC expectation.

5.2 Calibration of the q^2 Distribution

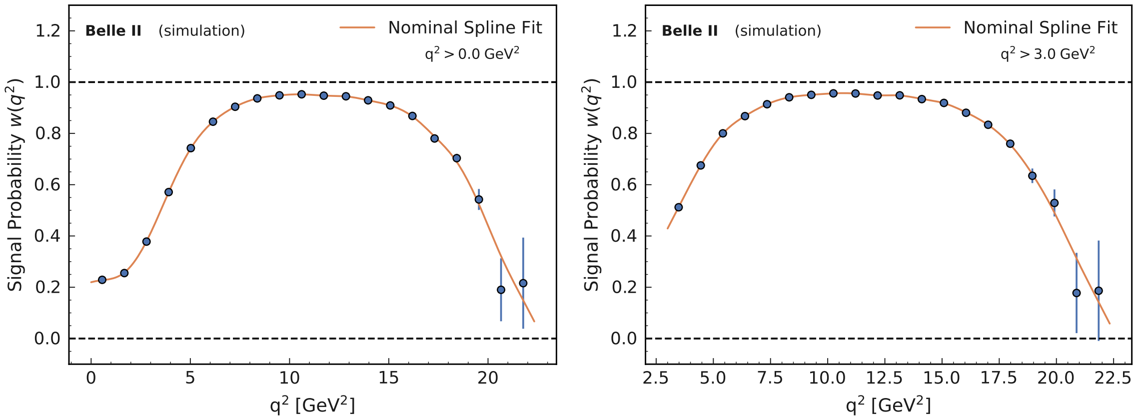
Despite the application of the kinematic fit to the $\Upsilon(4S) \to B\overline{B}$ event and the resulting improvement in both resolution and bias of the reconstructed q^2 distribution, further corrections are necessary in order to measure unbiased moments. A comparison between the q^2 spectra on generator level and reconstructed $B \to X_c \ell \nu_\ell$ events with all selection criteria applied is given in Figure 5.7. The generator level distributions shows a distinct kinematically allowed endpoint around $q^2 \approx 11.6\,\mathrm{GeV}^2$. Due to mis-reconstruction and detector resolution effects, the reconstructed q^2 distribution exhibits higher q^2 values on average as well as an extension beyond the kinematically allowed endpoint of the generator level spectrum.

Figure 5.8 shows the first four q^2 moments as functions of the lower q^2 threshold extracted from the reconstructed and generator level q^2 distribution for $B \to X_c \ell \nu_\ell$ signal MC with all selection criteria applied. In the following, the minimal value of the lower q^2 threshold considered is $q^2 > 1.5\,\mathrm{GeV}^2$. The moments of the reconstructed q^2 spectrum show a continuous positive bias compared to the generator level moments. For the first moment, the observed bias ranges in between 20% and 12%, while the fourth moment exhibit a bias between 143% and 93%.

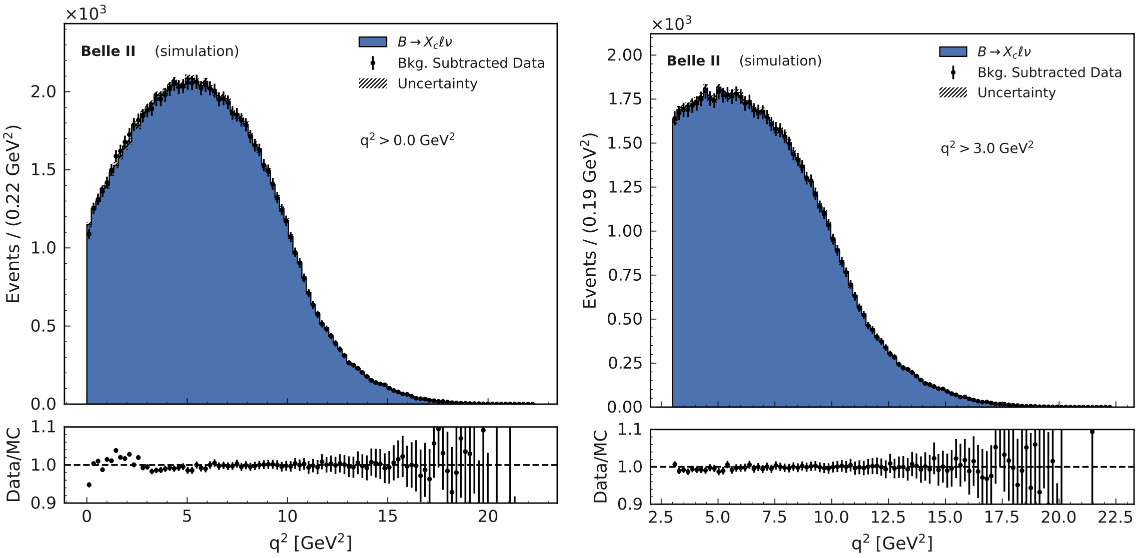
To ensure an unbiased measurement of the q^2 moments, a calibration method is developed and applied to the measured q^2 spectrum correcting the bias introduced by the reconstruction, detector



(a) q^2 distributions for the remaining MC background ($B\overline{B}$ and continuum) and Asimov data.

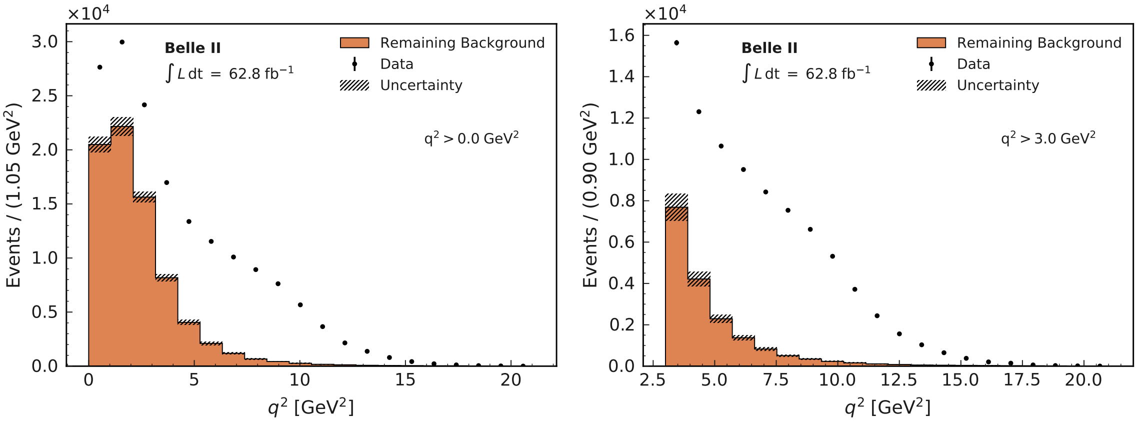


(b) Bin-wise signal probabilities and the nominal spline fit determined using Asimov data.

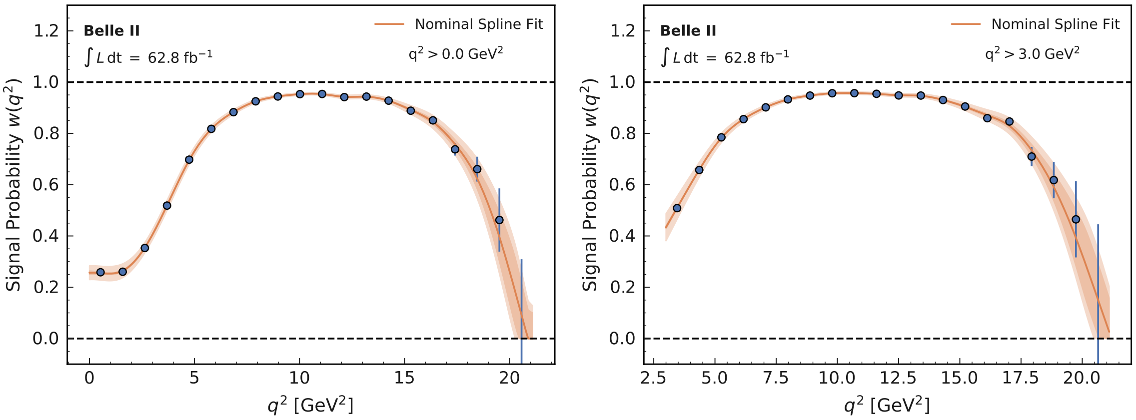


(c) q^2 distributions of the background subtracted Asimov data spectrum and the $B \to X_c \ell \nu_\ell$ MC expectation

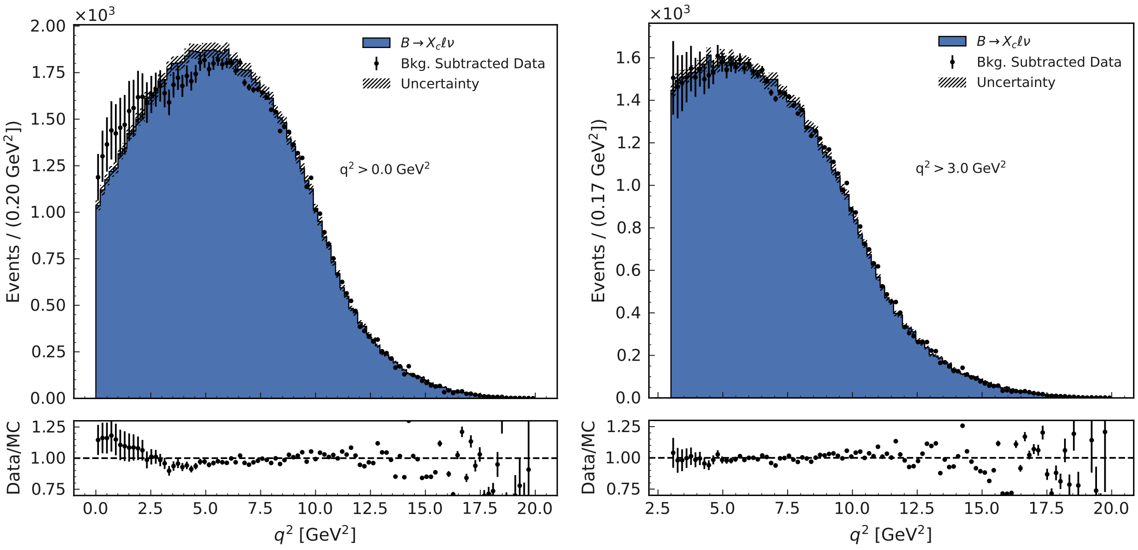
Figure 5.5: Examples of the background subtraction using signal probability weights based on the Asimov data set. The left and right columns show the background subtraction for $q^2 > 0.0\,\mathrm{GeV}^2$ and $q^2 > 3.0\,\mathrm{GeV}^2$, respectively.



(a) q^2 distributions for the remaining MC background ($B\overline{B}$ and continuum) and data.



(b) Bin-wise signal probabilities and the nominal spline fit determined using the measured data set.



(c) q^2 distributions of the background subtracted data spectrum and the $B \rightarrow X_c \ell \nu_\ell$ MC expectation

Figure 5.6: Examples of the background subtraction using signal probability weights based on the measured data set. The left and right columns show the background subtraction for $q^2 > 0.0\,\mathrm{GeV}^2$ and $q^2 > 3.0\,\mathrm{GeV}^2$, respectively.

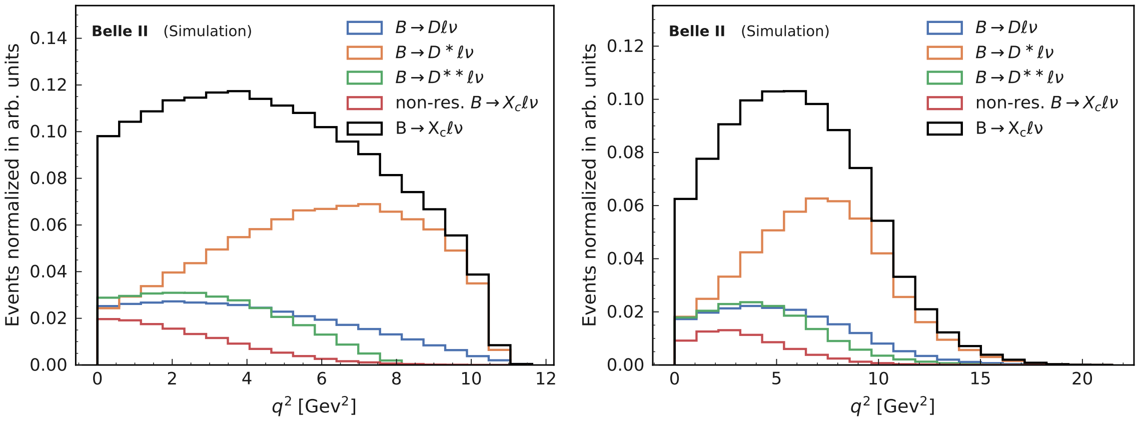


Figure 5.7: Normalized q^2 distributions for generator level (left) and reconstructed (right) $B \to X_c\ell\nu_\ell$ signal MC. The distributions of the individual X_c components are normalized to their respective contribution to the overall $B \to X_c\ell\nu_\ell$ spectrum. All selection criteria are applied.

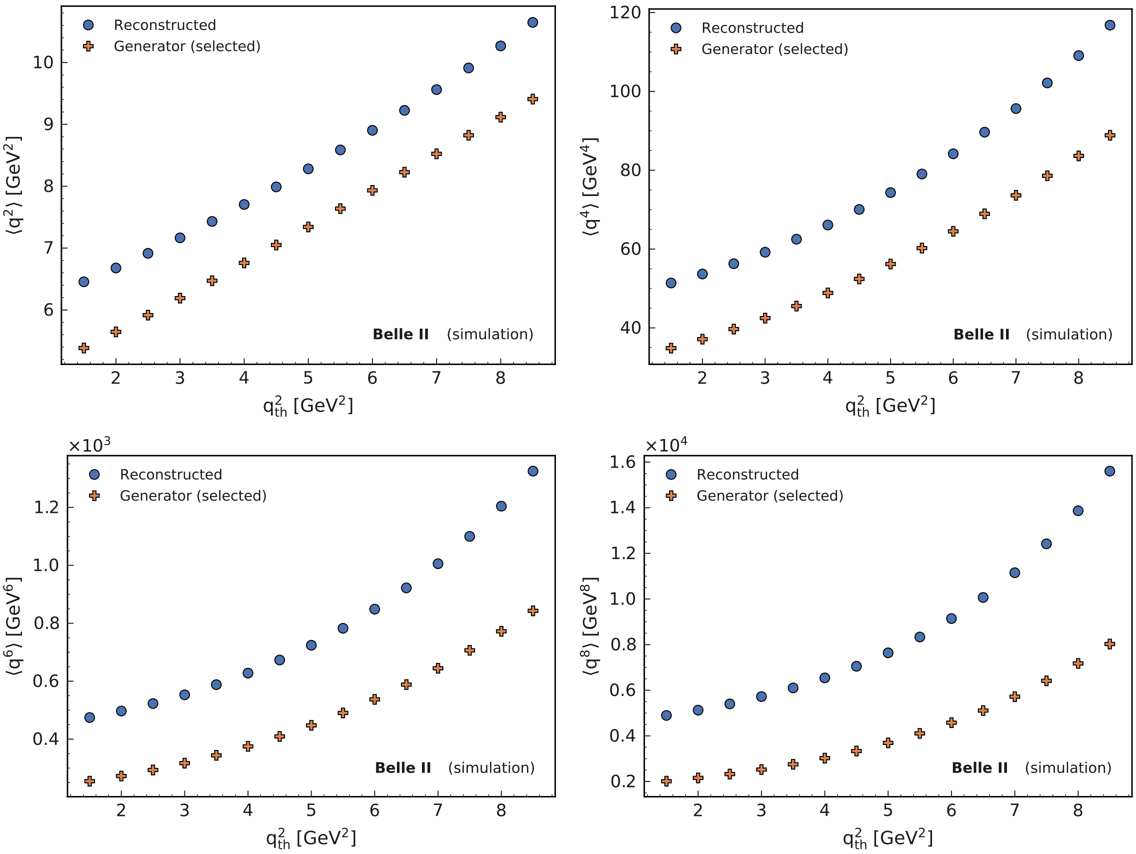


Figure 5.8: First four moments of the reconstructed (blue) and generator level (orange) q^2 distribution for $B \to X_c\ell\nu_\ell$ signal MC as functions of the lower q^2 threshold. All selection criteria are applied.

resolution, and different selection efficiencies for different $B \to X_c \ell \nu_\ell$ final-states. The calibration is performed on an event basis relating measured and true q^2 moments instead of unfolding the binned q^2 distribution. The general principle of the calibration is based on the almost linear relationship between moments of the reconstructed q^2 spectrum $\langle q^{2n}_{\mathrm{reco}} \rangle$ and the generator level distribution with all selection criteria applied. Here, the reconstructed and generator level q^2 moments are calculated with lower q^2 thresholds using the reconstructed and generator level q^2 values, respectively. $\langle q^{2n}_{\mathrm{reco}} \rangle$ and $\langle q^{2n}_{\mathrm{gen,sel}} \rangle$ are calculated using an independent $B \to X_c \ell \nu_\ell$ signal MC sample and without separating any reconstruction channels. Their relationship is parameterized using the linear function

$$\langle q^{2n}_{\mathrm{reco}} \rangle = m_n \times \langle q^{2n}_{\mathrm{gen,sel}} \rangle + c_n \quad (n = 1, 2, 3, 4), \tag{5.27}$$

with m_n and c_n denoting the slope and offset parameters, respectively. The fit is performed using the method of orthogonal distance regression [117] implemented in `SciPy` [106]. This method takes into account the statistical uncertainties on both moments, $\langle q^{2n}_{\mathrm{reco}} \rangle$ and $\langle q^{2n}_{\mathrm{gen,sel}} \rangle$. This fit is repeated for each order of the moments individually. Figure 5.9 shows the relationships between $\langle q^{2n}_{\mathrm{gen,sel}} \rangle$ and $\langle q^{2n}_{\mathrm{gen,sel}} \rangle$ as well as the linear fits. The resulting slopes m_n of these linear fits range between 1.04 for the first and 1.72 for the fourth moment. The corresponding offset parameters c_n are all positive with values between 0.75 GeV^2 and 1 362.90 GeV^8 for the first and fourth moment, respectively.

The actual calibration function is given by the inverse of the relation in Equation (5.27) and replacing $\langle q^{2n}_{\mathrm{reco}} \rangle \to q^{2n}_{\mathrm{reco}}$ and $\langle q^{2n}_{\mathrm{calib}} \rangle \to q^{2n}_{\mathrm{calib}}$. The calibrated value q^{2n}_{calib} is calculated from the reconstructed value q^{2n}_{reco} for each event as

$$q^{2n}_{\mathrm{calib}} = \frac{q^{2n}_{\mathrm{reco}} - c_n}{m_n}. \tag{5.28}$$

Due to the linear nature of the calibration, the q^{2n} distribution are shifted towards smaller q^2 values and also rescaled. This can also result in negative calibrated q^2 values. Overall, the calibration procedure does not correct the shape of the distribution but effectively only changes the mean value of the q^2 spectrum

$$\langle q^{2n}_{\mathrm{calib}} \rangle = \left\langle \frac{q^{2n}_{\mathrm{reco}} - c_n}{m_n} \right\rangle = \frac{\langle q^{2n}_{\mathrm{reco}} \rangle - c_n}{m_n}. \tag{5.29}$$

The reconstructed, calibrated and generator level q^2 distributions for $q^2 > 1.5\,\mathrm{GeV}^2$ are compared in Figure 5.10.

The performance of the calibration is illustrated in Figure 5.11 comparing the q^2 moments after applying the calibration procedure and generator level moment with and without selection criteria applied. Compared to the situation before the calibration shown in Figure 5.8, the remaining bias between $\langle q^{2n}_{\mathrm{calib}} \rangle$ and $\langle q^{2n}_{\mathrm{gen,sel}} \rangle$ is in the order a few percent depending on the order of the moment and the q^2 threshold.

The origin of this bias are small non-linearities in the relation between $\langle q^{2n}_{\mathrm{reco}} \rangle$ and $\langle q^{2n}_{\mathrm{gen,sel}} \rangle$ since the calibration is based on the assumption of a linear behavior. Studies using a calibration based on the use of a quadratic ansatz describing the relation between $\langle q^{2n}_{\mathrm{reco}} \rangle$ and $\langle q^{2n}_{\mathrm{gen,sel}} \rangle$ could not reduce this bias.

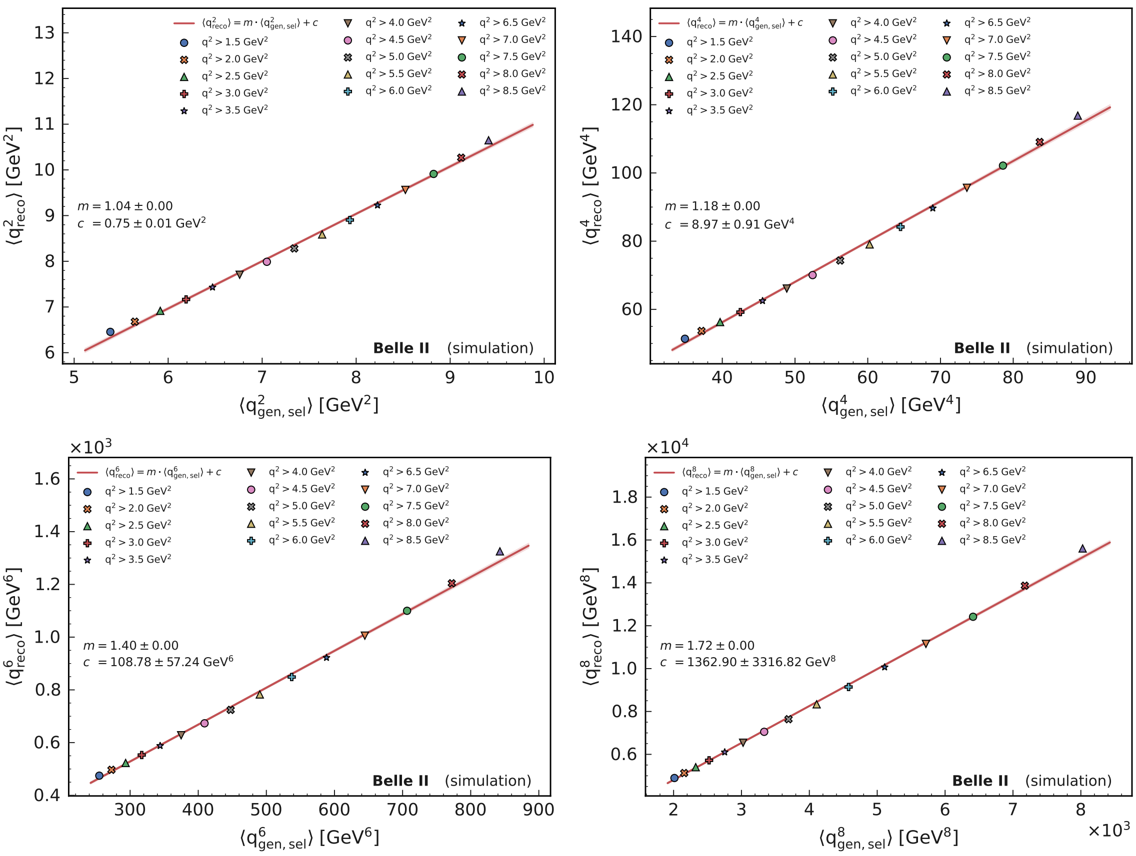


Figure 5.9: The linear calibration curves for q^2 moment. The q^2 moments are shown as functions of the minimum q^2 requirement on the reconstructed and true underlying q^2 distributions.

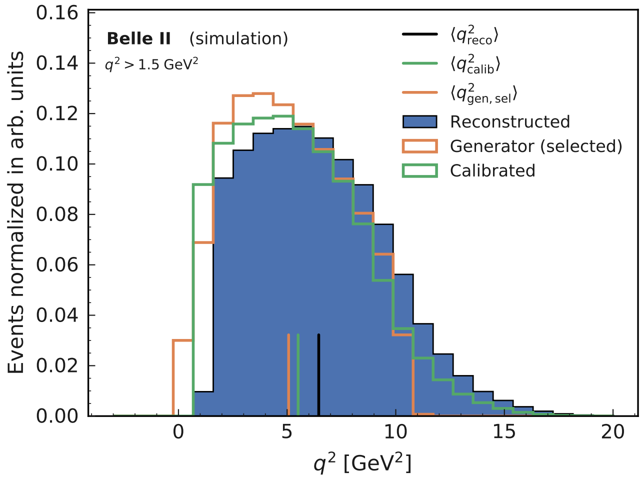


Figure 5.10: Normalized q^2 distributions after the reconstruction (blue), calibration procedure (green) and on generator level (orange). The vertical lines represent the mean values of the different distributions. All selection criteria are applied.

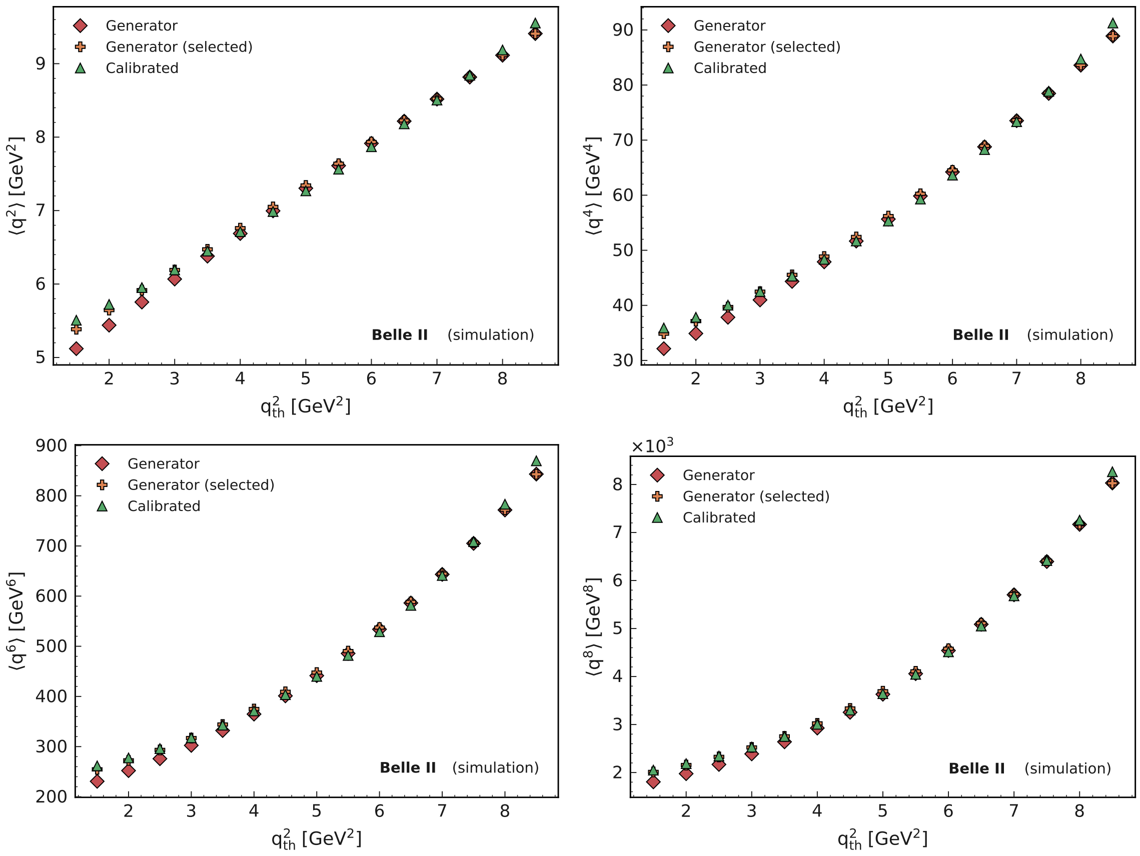


Figure 5.11: Comparisons of the first four q^2 moments as functions of the lower q^2 threshold after applying the calibration procedure (green) and on generator level with the application of all event selection criteria (orange) and without selection (red).

To correct for this remaining bias after the calibration, a bias correction factor C_{calib} is calculated for each moment and lower q^2 threshold. C_{calib} is defined as the ratio of generator level with selection and calibrated q^2 moments

$$C_{\text{calib}} = \frac{\langle q^{2n}_{\text{gen,sel}} \rangle}{\langle q^{2n}_{\text{calib}} \rangle}. \tag{5.30}$$

Figure 5.12 shows the determined calibration bias correction factors of the first to fourth q^2 moments. They vary between 0.96 and 1.03 depending the order of the moment and the lower q^2 threshold.

As pointed out before, the calibration is based on the relationship between the reconstructed moments and the moments on generator level after the application of all selection criteria. As a result the $\langle q^{2n}_{\text{reco}} \rangle$ are corrected to $\langle q^{2n}_{\text{gen,sel}} \rangle$. An additional intrinsic bias is observed between the q^2 moments on generator level with simulation, reconstruction, and event selection applied $\langle q^{2n}_{\text{gen,sel}} \rangle$ and the moments calculated without it $\langle q^{2n}_{\text{gen}} \rangle$. Figure 5.11 shows that this bias is present especially for q^2 moments determined for small values of q^2_{th}. This bias is introduced by different selection and

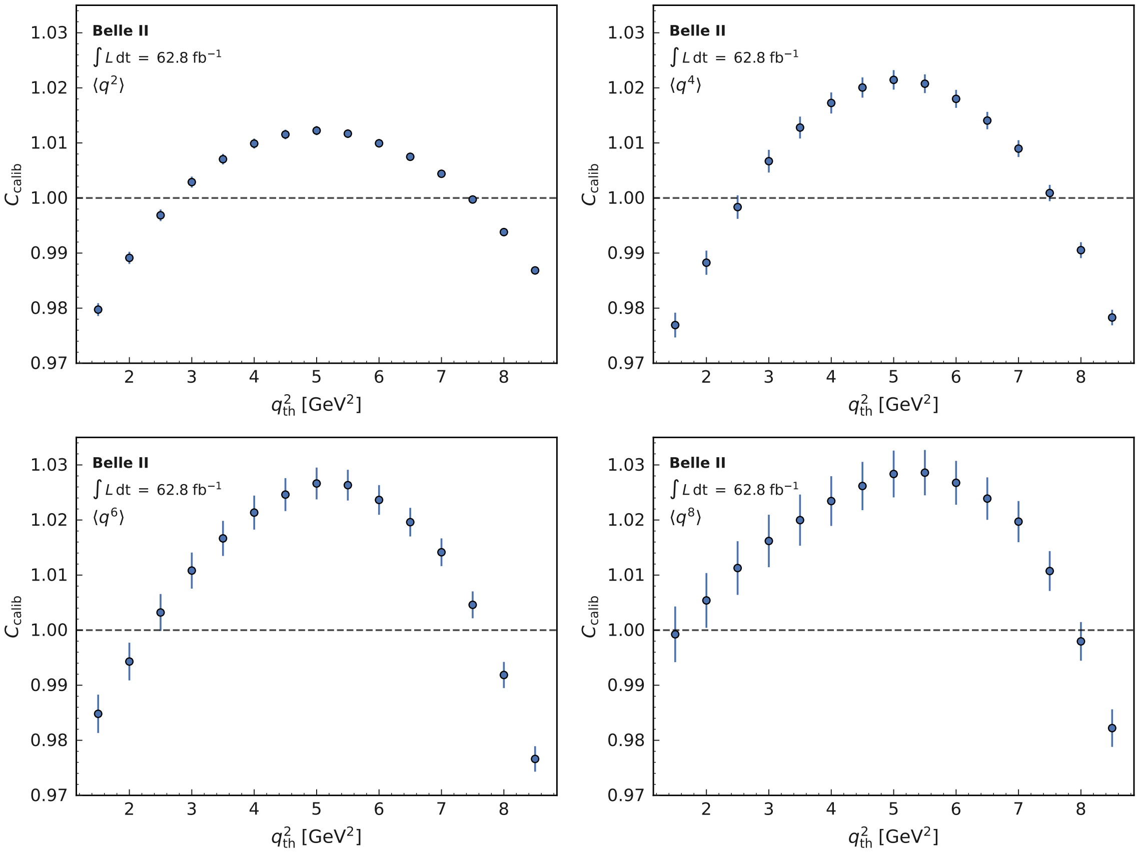


Figure 5.12: Calibration factors C_calib applied in the calculation of the first to fourth q^2 moments.

acceptance efficiencies for the various $B \rightarrow X_c \ell \nu_\ell$ final-states. Figure 5.13 illustrates this effect by showing the different efficiencies for selecting generator events based on a basic lepton track selection requirement of $|\boldsymbol{p}| > 0.5\,\mathrm{GeV}$ and $17° < \theta < 150°$ as functions of the lower q^2 threshold. One observes significantly lower selection efficiencies for the non-resonant X_c final-states at low q^2_th values compared to the resonant $B \rightarrow X_c \ell \nu_\ell$ decays using this selection criteria only. For higher values of q^2_th, the efficiencies for the different final states approach the same value. Another difference between $\langle q^{2n}_\mathrm{gen,sel} \rangle$ and $\langle q^{2n}_\mathrm{gen} \rangle$ is that the generator moments with selection are simulated with possible FSR by `Photos`. However, the theory predictions do not take this FSR into account and this effect has to be corrected as well.

A second correction factor C_gen is determined to correct this intrinsic bias. Again, it is also calculated individually for different lower q^2 thresholds and moment orders and is defined as

$$C_\mathrm{gen} = \frac{\langle q^{2n}_\mathrm{gen} \rangle}{\langle q^{2n}_\mathrm{gen,sel} \rangle}. \tag{5.31}$$

The obtained values for C_gen are shown in Figure 5.14. Again a clear q^2_th dependence is observed, with larger corrections for lower q^2_th values. For the lowest value $q^2 > 1.5\,\mathrm{GeV}^2$ corrections between 12%

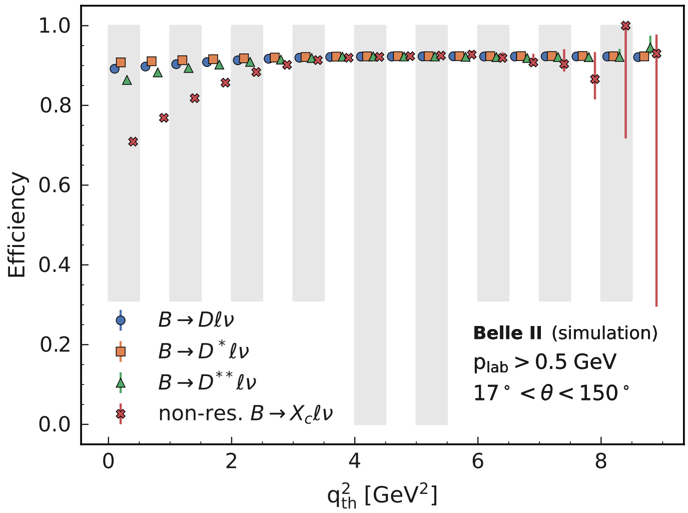


Figure 5.13: Selection efficiencies as functions of q^2 threshold q^2_{th}. The points for different X_c final states and the same lower q^2 threshold are shifted horizontally and the grey and white-bands visually group the same q^2 threshold.

and almost 18% are necessary for $\langle q^2 \rangle$ and $\langle q^8 \rangle$, respectively. With larger q^2_{th} values, C_{gen} approaches unity for all moments. Independent $B \to X_c \ell \nu_\ell$ signal MC samples are used for the determination calibration parameters and the calculation of C_{calib} and C_{gen}.

The numerical values of the bias correction factors C_{calib} and C_{gen} are summarized in Tables C.1 to C.4 in Appendix C.

5.3 Testing the Moment Extraction with Ensemble Tests

The moment extraction procedure is tested with two ensemble tests based on simulated data. The first test checks the calibration of the q^2 moments, while the second test is used to investigate the full moment extraction including q^2 calibration and background subtraction with signal probability weights.

For the ensemble tests, the q^2 moments are repeatedly determined on different subsets of MC. One half of the MC is treated as mock data, while the second half is used for the steps dependent on simulated data. One entity of the ensembles is generated as follows:

1. The calibration coefficients c_n and m_n are determined using the same independent sample of signal $B \to X_c \ell \nu_\ell$ MC as in the actual measurement. With these coefficients, the calibration is applied to a sample of generic $B\overline{B}$ and continuum MC sample. The generic $B\overline{B}$ MC contains both signal and background processes.

2. The combined MC data set with generic $B\overline{B}$ and continuum events is split randomly into two equal size samples. One sample is treated as MC, while the second sample is used as the mock data set.

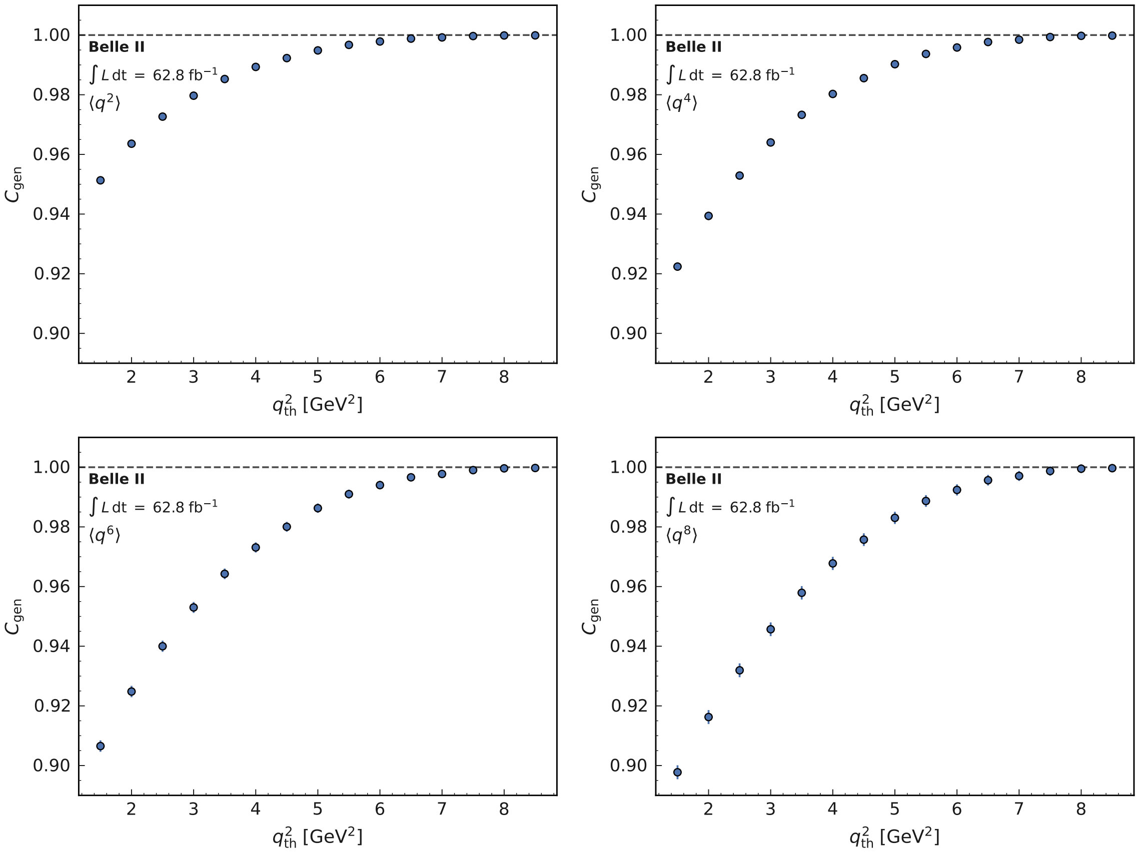


Figure 5.14: Calibration factors C_{gen} applied in the calculation of the first to fourth q^2 moments.

3. The $B \longrightarrow X_c \ell \nu_\ell$ part of the MC sample is then used to determine the bias correction factors C_{calib} and C_{gen}. For the check of only the q^2 calibration, the correct signal probabilities of $w_i = 1$ and $w_i = 0$ are assigned to the signal and background events in the mock data set, respectively. When testing the full moment extraction procedure, the signal probabilities are determined according to the method introduced in Section 5.1.4 using the background events of the MC sample and the mock data.

4. Lastly, the q^2 moments are calculated according to Equation (5.1) with the events in the mock data set.

For each test, an ensembles size of 1 000 is used to study the extraction formalism. The relative difference between measured and generator moments, defined as

$$\Delta_{\text{rel}}(\langle q^{2n} \rangle) = \frac{\langle q^{2n} \rangle - \langle q^{2n}_{\text{gen}} \rangle}{\langle q^{2n}_{\text{gen}} \rangle}, \tag{5.32}$$

is calculated for each ensemble entity to investigate a possible bias.

The results for the check of the calibration procedure only are summarized in Figure 5.15. It shows

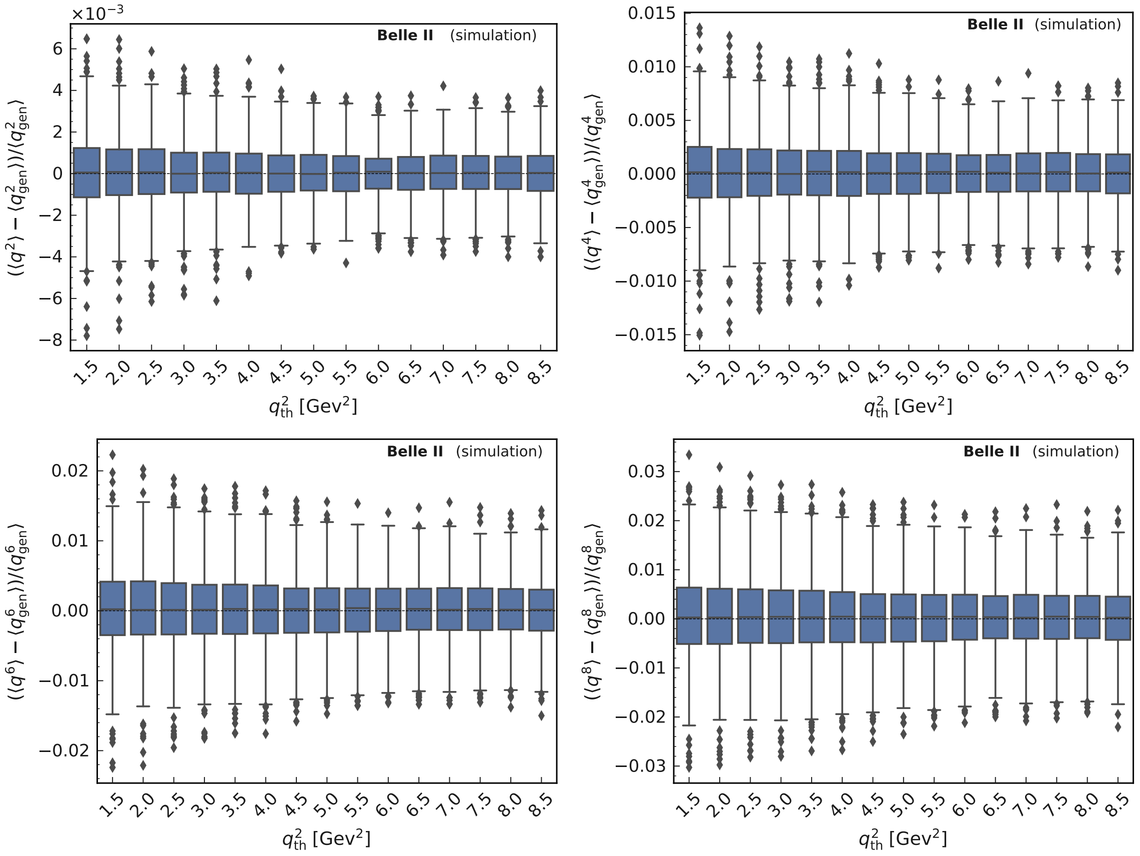


Figure 5.15: Difference of the calculated and generator level q^2 moments normalized to the generator level moments as functions of the lower q^2 thresholds. The box plots show the median as well as the first and third quartiles while the length of the whiskers corresponds to 1.5 times the interquartile range. These values are obtained by a sample test checking the calibration procedure only.

box plots of Δ_{rel} for each moment and lower q^2 threshold. The box plots follow the standard definition showing the median as well as the first and third quartile. The length of the whiskers corresponds to 1.5 times the interquartile range. No bias stemming from the q^2 calibration is observed for all order of moments and lower q^2 threshold further showing a reliable performance of the latter.

Figure 5.16 presents the results of the ensemble test investigating the full extraction procedure (calibration and background subtraction). A small bias is observed for moments of all order especially for lower q^2 thresholds between 1.5 GeV^2 and 3 GeV^2. This bias is introduced by the interpolation of the binned signal probabilities with the cubic spline fit at the lower boundary of the q^2 spectrum. For each moment and lower q^2 threshold, the bias is quantified with the absolute value of the sample mean of the Δ_{rel} distribution. However, the observed non-closure bias is still small compared to the size of other leading systematic uncertainties discussed in Section 5.5. Thus, an additional systematic uncertainty is added instead of correcting for the bias.

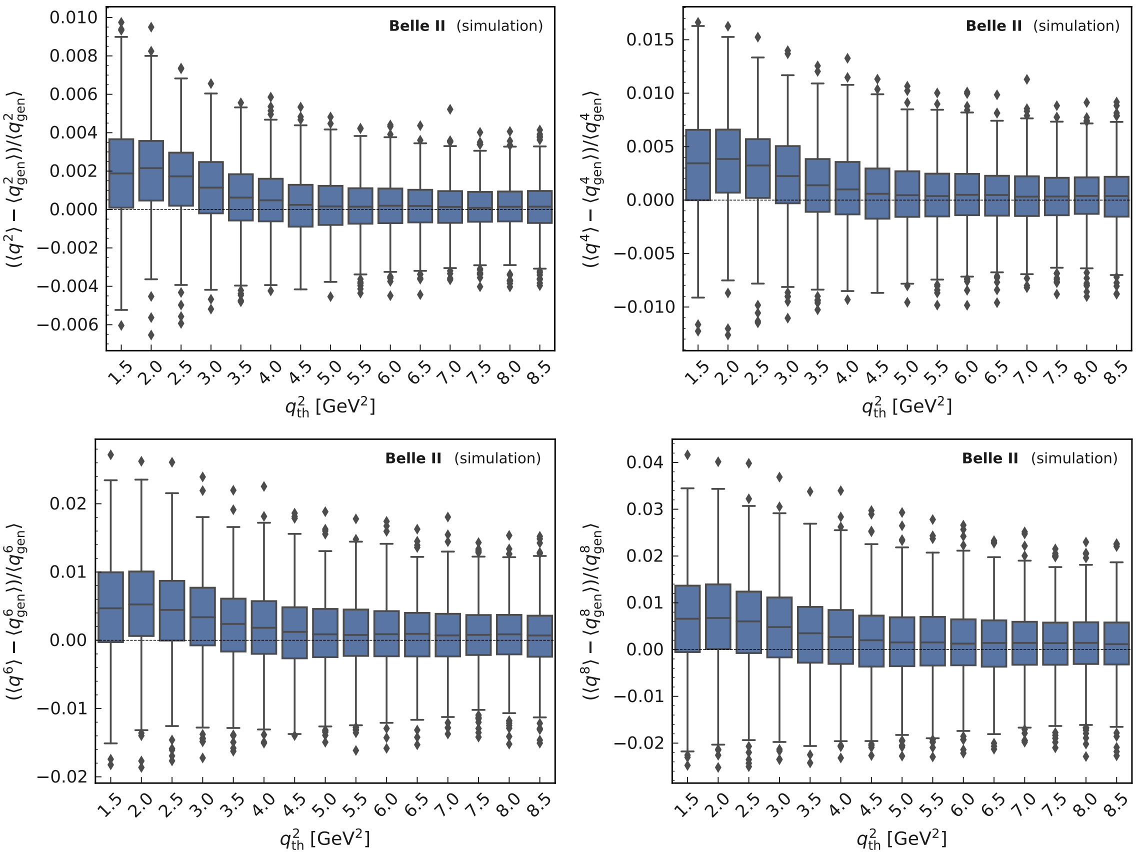


Figure 5.16: Difference of the calculated and generator level q^2 moments normalized to the generator level moments as functions of the lower q^2 thresholds. The box plots show the median as well as the first and third quartiles while the length of the whiskers corresponds to 1.5 times the interquartile range. These values are obtained by a sample test checking the full analysis procedure (calibration and background subtraction).

5.4 Statistical Uncertainties and Correlations

The q^2 moments are measured as functions of progressive lower thresholds on q^2 itself. Subsequently, moments with a higher q^2_{th} value are calculated from a subset of events used in the calculation of moments with lower q^2 thresholds values. Thus, a significant correlation between the different moments is expected. The determination of these correlations is important since they serve as input for the calculation of the total experimental covariance matrix of the moments. This covariance is then used in the extraction of $|V_{cb}|$ and the HQE parameters by a χ^2 fit of the theory predictions to the measured q^2 moments.

In this work, a bootstrapping procedure [118] is used to estimate the statistical uncertainties and correlation between q^2 moments of different order and lower q^2 thresholds. For this approach, a set of B artificial data samples $D_b = d'_1, \ldots, d'_{n_b}$ with $b = 1$–B are generated by using random sampling with replacement from the observed data sample. In addition, the total number of events in each generated data sample n_b is varied assuming a Poisson distribution with a mean equal to the observed number of events. In total $B = 10\,000$ bootstrapped data samples are generated for the estimation of

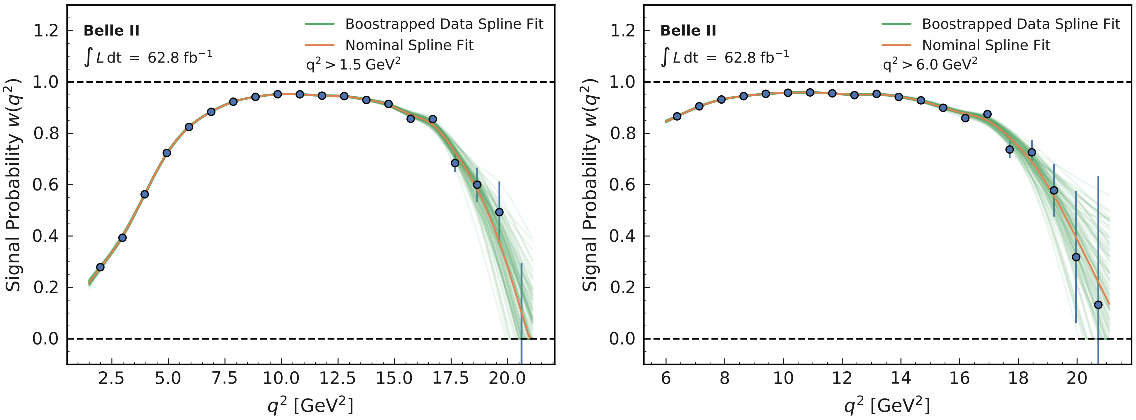


Figure 5.17: Binned signal probability w_i together with the nominal cubic spline fit (orange) and variations of the spline fit (green) determined with bootstrapped data samples for lower q^2 thresholds of 1.5 GeV^2 (left) and 6 GeV^2 (right).

the statistical uncertainty and correlation.

With each bootstrapped data sample D_b, the determination of the signal probability weights is repeated and a new set of q^2 moments is calculated. Figure 5.17 illustrates the variations of the signal probability $w(q^2)$ spline fit for lower q^2 thresholds of 1.5 GeV^2 and 6 GeV^2. Significant variations of the spline fit are observed in the sparsely populated region at high q^2 above 15 GeV^2. For lower q^2 values, the statistical variations have only a negligible impact on $w(q^2)$.

The statistical uncertainty of a moment for a certain q^2 threshold and order is estimated using the empirical sample variance of the set bootstrapped moments. The Pearson correlation coefficients are used to determine the correlation between the moments.

Figure 5.18 gives examples of one and two-dimensional distributions of q^2 moments obtained by the bootstrapping procedure. Here, the relationship of the first moment calculated at different lower q^2 thresholds of 1.5 GeV^2 and 2.0 GeV^2 shows a positive correlation of $\rho = 86.9\%$. Comparing the same moment but with a larger difference between the q^2_{th} values shows that the correlation between these moments declines. A correlation value of $\rho = 30.0\%$ is observed between the first moments calculated for $q^2 > 1.5$ GeV^2 and $q^2 > 7.5$ GeV^2. The correlation between all moments and all lower q^2 thresholds is presented in Figure 5.19. In addition, the numerical values for the correlation coefficients between individual moments are summarized in Figures C.1 to C.4. There is a clear tendency of high correlations between moments with different q^2_{th} values that lie close together. Moments measured with different lower q^2 thresholds that have a larger separation in q^2 tend to have lower correlation values. Also moments of higher order tend to have a higher correlation between moments determined for different q^2_{th} values.

5.5 Systematic Uncertainties

In addition to the statistical uncertainty on the weighted mean, several sources of systematic uncertainties are considered. They mainly affect the background subtraction or the q^2 calibration procedure, which both rely on simulated data.

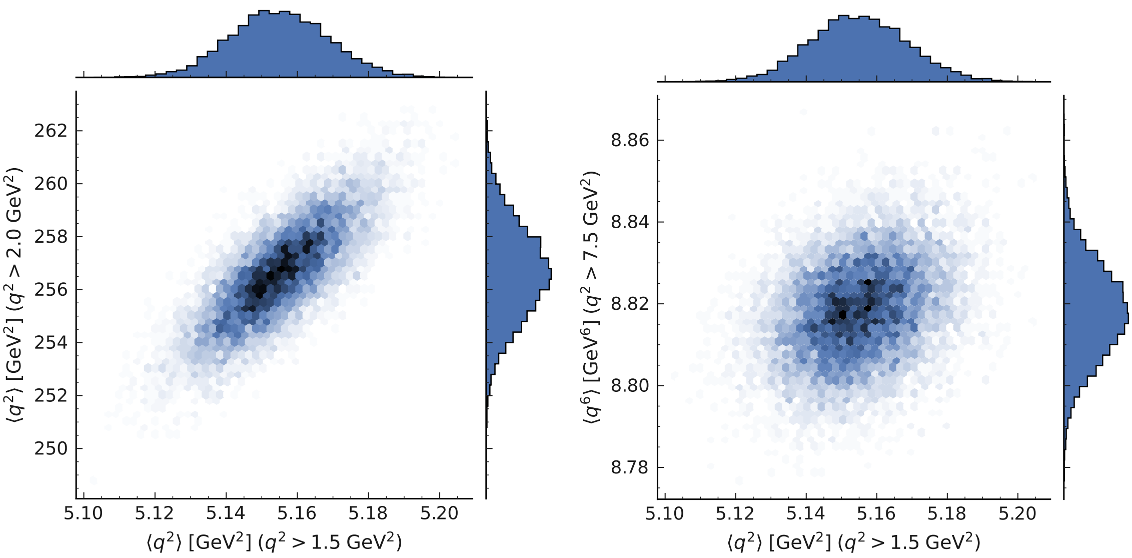


Figure 5.18: One and two-dimensional distributions of q^2 moments obtained by the bootstrapping procedure. The left-hand figure shows values for the first moments with lower q^2 thresholds of $1.5\,\mathrm{GeV}^2$ and $2.0\,\mathrm{GeV}^2$. The figure on the right-hand side illustrates the relationship between $\langle q^2 \rangle$ calculated at lower q^2 threshold of $1.5\,\mathrm{GeV}^2$ and $7.5\,\mathrm{GeV}^2$

5.5.1 Background Subtraction

Spline Smoothing Parameter

As discussed in Section 5.1.4, the smoothness of the cubic spline is controlled by a regularization parameter p. This parameter can be set to values in the range $[0, 1]$, with $p = 0$ and $p = 1$ resulting in a least squares fit of a straight line and exact cubic spline interpolation, respectively. For the nominal results, the default value is chosen to be $p = 0.1$. To asses the impact of different regularization parameters, alternative spline fits are repeated with variations of p from $p = 0.1$ to $p = 1$ in 0.01 increments. With each varied spline flit, new sets of signal probability weights are calculated and the measurement of the q^2 moments is repeated. Examples for the signal probability functions determined from spline fits with varied smoothing parameters p are shown in Figure 5.20 for lower q^2 thresholds of $1.5\,\mathrm{GeV}^2$ and $6\,\mathrm{GeV}^2$. The interpolation between the binned signal probabilities differs only significantly at high q^2 values. At low q^2 values, the spline fit is insensitive to the choice of the smoothing parameter. The covariance matrix of the moments with respect to the smoothing parameter variation is directly estimated using the empirical sample covariance off all these different variations.

Background Normalization and Shape

The cubic spline is fitted to the bin-wise signal probabilities calculated from the measured q^2 spectrum and remaining background expectation according to Equation (5.26). The uncertainty of the background template is given by the normalization and shape uncertainties after the m_X fit. Here, the total uncertainty is described by the background template covariance matrix constructed according to the approach used in Section 5.1.3.

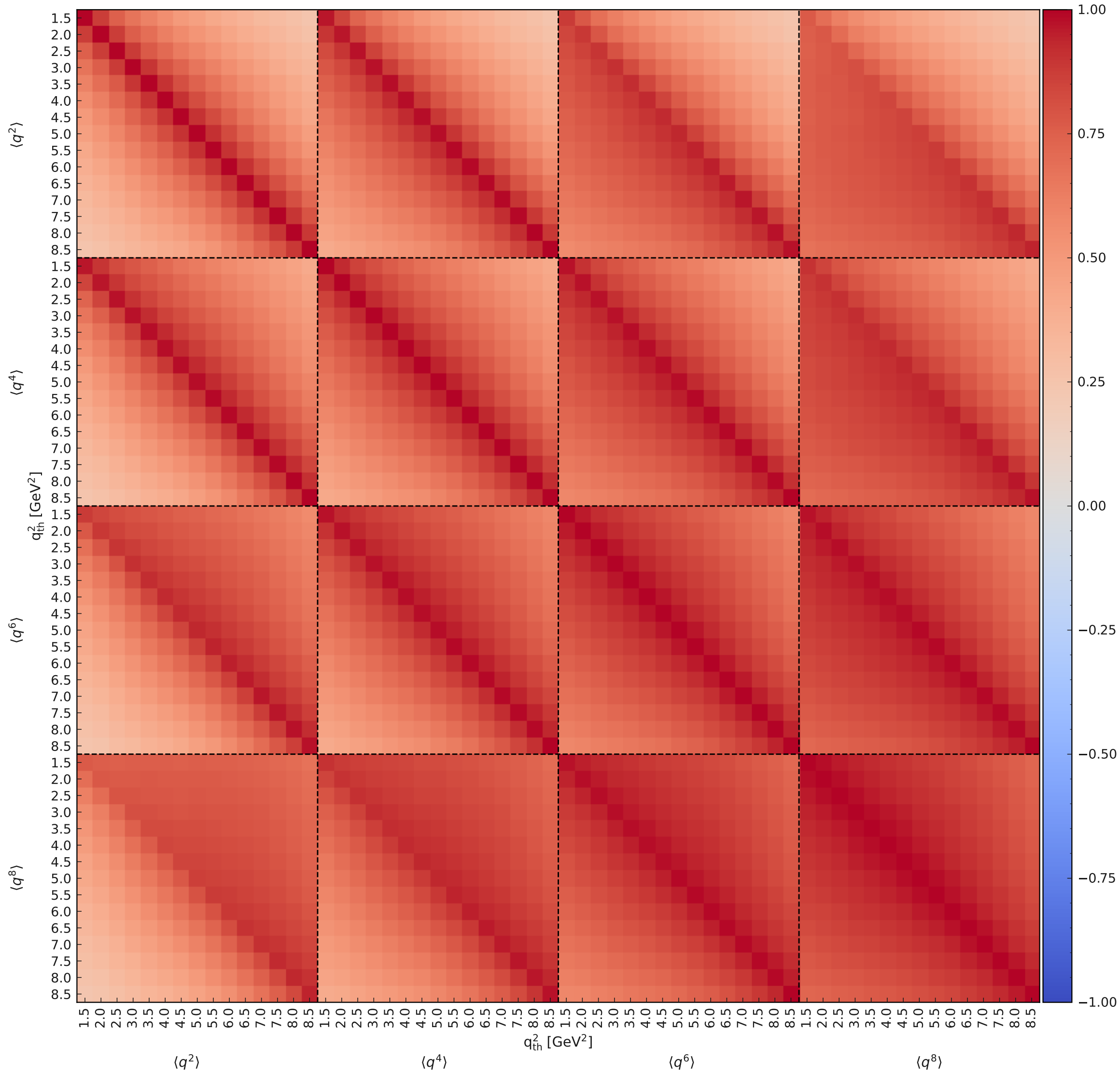


Figure 5.19: Statistical correlations between $\langle q^2 \rangle$ and $\langle q^{2n} \rangle$ for n = 1–4.

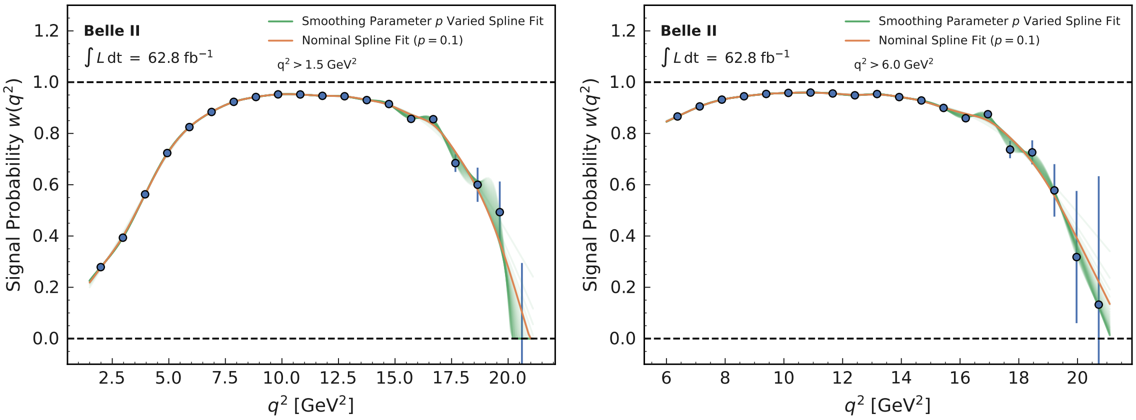


Figure 5.20: Binned signal probability w_i together with the nominal cubic spline fit (orange) and variations of the spline fit (green) determined with different smoothing parameter values p for lower q^2 thresholds of $1.5\,\mathrm{GeV}^2$ (left) and $6\,\mathrm{GeV}^2$ (right).

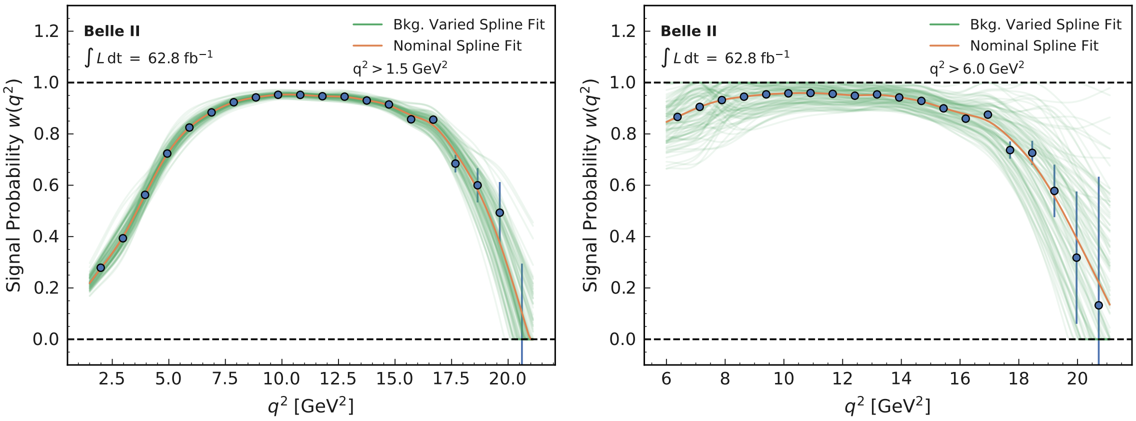


Figure 5.21: Binned signal probability w_i together with the nominal cubic spline fit (orange) and variations of the spline fit (green) determined from background variations for lower q^2 thresholds of $1.5\,\mathrm{GeV}^2$ (left) and $6\,\mathrm{GeV}^2$ (right).

A toy MC approach is used to asses the uncertainty on the signal probability function due to the background normalization and shape uncertainty. Given the nominal background expectation $\nu_i(\hat{n}_{B\bar{B}}, \hat{n}_{q\bar{q}})$ in each q^2 bin and the background covariance C_{bkg}, an ensemble of varied background distributions is generated assuming a multivariate Gaussian distribution $\mathcal{N}(\boldsymbol{\nu}'|\boldsymbol{\nu}, C_{\mathrm{bkg}})$. In total, 1 000 variations of the background expectation are considered. For each variation ν_i', varied signal probability weights are determined by repeating the cubic spline fit.

Exemplary fits of the cubic spline to bin-wise signal probabilities determined from varied background distributions are shown in Figure 5.21 for $q^2 > 1.5\,\mathrm{GeV}^2$ and $q^2 > 6\,\mathrm{GeV}^2$. The calculation of the q^2 moments is repeated for each variation and the covariance matrix is directly estimated using the empirical sample covariance off all these different variations.

5.5.2 q^2 Calibration

Uncertainty of Calibration Coefficients c_n and m_n

The calibration coefficients c_n and m_n are determined using a fit of the relationship between $\langle q^{2n}_{\mathrm{reco}}\rangle$ and $\langle q^{2n}_{\mathrm{gen,sel}}\rangle$. To propagate their uncertainties to the moments, two orthogonal sets of parameter variations of are calculated as

$$\boldsymbol{p}_i^{\pm} = \boldsymbol{p}^{\mathrm{nom}} \pm \sqrt{\lambda_i}\boldsymbol{e}_i \quad (i = 1, 2), \quad \text{with} \tag{5.33}$$

$$\boldsymbol{p}^{\mathrm{nom}} = (c_n, m_n) \tag{5.34}$$

Here λ_i and $\boldsymbol{e}_i$ denote the eigenvalues and eigenvectors of the covariance matrix for the coefficients, respectively. In a next step, the calibration is repeated with these varied coefficients and the uncertainty on the q^2 moments is estimated as

$$\sigma_{\langle q^{2n}_{\boldsymbol{p}_i}\rangle} = \frac{\left|\langle q^{2n}_{\boldsymbol{p}_i^+}\rangle - \langle q^{2n}_{\boldsymbol{p}_i^-}\rangle\right|}{2}. \tag{5.35}$$

This uncertainty is considered as fully correlated between moments of different order and for different lower q^2 thresholds. The sign information of the difference in the numerator is considered as well in the correlation structure when constructing the covariance matrix.

Uncertainty of Bias Correction Factors C_{calib} and C_{gen}

The bias correction factors are determined with the moments $\langle q^{2n}_{\mathrm{reco}}\rangle$, $\langle q^{2n}_{\mathrm{gen,sel}}\rangle$ and $\langle q^{2n}_{\mathrm{gen}}\rangle$ obtained from simulated data samples. Thus, both C_{calib} and C_{gen} are also only known with a finite precision related to the MC sample size. The uncertainty on $C_{\mathrm{calib}} \times C_{\mathrm{gen}}$ is propagated to the q^2 moments by repeating the calculation with the product varied up and down within one standard deviation and taking half the difference between the resulting moments as systematic uncertainty

$$\sigma_{\langle q^{2n}_{C}\rangle} = \frac{\left|\langle q^{2n}_{C^+}\rangle - \langle q^{2n}_{C^-}\rangle\right|}{2}. \tag{5.36}$$

Note that the uncertainty on $\langle q^{2n}_{\mathrm{gen,sel}}\rangle$ cancels out in the product $C_{\mathrm{calib}} \times C_{\mathrm{gen}}$. This uncertainty is considered as fully correlated between moments of different order and for different lower q^2 thresholds.

Composition of the $\mathrm{B} \to \mathrm{X_c}\ell\nu_\ell$ Spectrum

The $B \to X_c\ell\nu_\ell$ spectrum is simulated with exclusive decays according to the branching fractions listed in Table 3.1. The actual composition of these different exclusive X_c states has a direct impact on the q^2 moments in the signal MC. Thus, the calibration of the q^2 distribution is also affected by this composition uncertainty. To evaluate it's impact, new set of calibration coefficients c_n and m_n and bias correction factors C_{calib} and C_{gen} are determined by independently varying the branching fraction of $B \to D\ell\nu_\ell$, $B \to D^*\ell\nu_\ell$ and the four $B \to D^{**}\ell\nu_\ell$ decays within one standard deviation. To assess the effect of the poorly known non-resonant states and gap modes, the calibration procedure is repeated with two different approaches. In the first approach, the contributions from $B \to D^{(*)}\pi\pi\ell\nu_\ell$

and $B \to D^{(*)}\eta\ell\nu_\ell$ are removed. For the second approach, the non-resonant and gap contributions are not only removed, but the branching fraction of the decays $B \to D_0^*\ell\nu_\ell$ and $B \to D_1'\eta\ell\nu_\ell$ are scaled up accordingly to replace them. Even though there is no experimental evidence for additional charm $1P$ states into other final states or the existence of an additional broad state in semileptonic $b \to c\ell\nu_\ell$ transitions, this provides an alternative kinematic description with a three-body decay to the simulation assuming momenta uniformly distributed in phase space . For each variation, the modifications are simultaneously applied to both MC samples used to calculate $\langle q^{2n}_{\mathrm{gen,sel}}\rangle$ and $\langle q^{2n}_{\mathrm{gen}}\rangle$.

The uncertainty of the branching fraction modifications is estimated by repeating the moment calculation and taking half the difference between the varied moments as systematic uncertainty

$$\sigma_{\langle q^{2n}_{\mathcal{B}}\rangle} = \frac{\left|\langle q^{2n}_{\mathcal{B}^+}\rangle - \langle q^{2n}_{\mathcal{B}^-}\rangle\right|}{2}. \tag{5.37}$$

For the two approaches of dropping and replacing the non-resonant and gap modes, the absolute difference to the nominal result is taken as estimate for the uncertainty

$$\sigma_{\langle q^{2n}_{\text{non-res./gap}}\rangle} = \left|\langle q^{2n}_{\text{non-res./gap}}\rangle - \langle q^{2n}\rangle\right|. \tag{5.38}$$

For constructing the respective covariance matrices, the uncertainties are considered to be fully correlated or anti-correlated depending on the sign of the difference in the uncertainty estimation.

BGL Form Factor Parameter Uncertainties of $B \to D^{(*)}\ell\nu_\ell$ Decays

The uncertainty of the BGL form factor parameters for $B \to D^{(*)}\ell\nu_\ell$ decays are evaluated with orthogonal parameter variations $\boldsymbol{p}_i^{\pm}$ (see Section 3.4.3). The form factor parameterization has a small impact on the shape of the q^2 distribution. Since $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ are the dominating contributions in the $B \to X_c\ell\nu_\ell$ spectrum, the influence of the form factor parameter uncertainty is evaluated as an additional systematic uncertainty. For each set of varied parameters, new calibration coefficients and bias correction factors are determined with reweighed signal and generator MC samples. Subsequently, the measurement of the q^2 moments is repeated and the uncertainty for each set of varied parameters is calculated as

$$\sigma_{\langle q^{2n}_{\boldsymbol{p}_{\mathrm{FF},i}}\rangle} = \frac{\left|\langle q^{2n}_{\boldsymbol{p}^+_{\mathrm{FF},i}}\rangle - \langle q^{2n}_{\boldsymbol{p}^-_{\mathrm{FF},i}}\rangle\right|}{2}. \tag{5.39}$$

Again the respective covariance matrices for each set of varied parameters are constructed assuming fully correlated or anti-correlated moments depending on the sign of the difference.

Lepton Identification Efficiency Uncertainty

The uncertainties on the lepton identification efficiency corrections are propagated and estimated by a set of varied correction weights. Section 3.4.3 describes the generation of these variations of the nominal correction within statistical and systematic uncertainties. For each variation, the calibration procedure is repeated extracting new calibration coefficients and bias correction factors. Subsequently, varied q^2 moments are calculated after applying the varied calibration, The covariance matrix is directly estimated with the sample covariance of the varied moments.

Tracking Efficiency Uncertainty

As a result of the track selection efficiency study performed in [112], an uncertainty of 0.69 % is assigned per track on the signal side. The calibration procedure is reevaluated by deriving new calibration coefficients and bias corrections factors with signal MC decays varied within one standard deviation of the tracking efficiency uncertainty. Equivalent of previous MC variations of this type, the uncertainty for the moments is calculated as

$$\sigma_{\langle q^{2n}_{\epsilon_{\text{tracking}}} \rangle} = \frac{\left| \langle q^{2n}_{\epsilon^{+}_{\text{tracking}}} \rangle - \langle q^{2n}_{\epsilon^{-}_{\text{tracking}}} \rangle \right|}{2} . \tag{5.40}$$

The covariance matrix is constructed assuming either fully (anti-)correlated q^2 moments depending on the sign in the difference $\langle q^{2n}_{\epsilon^{+}_{\text{tracking}}} \rangle - \langle q^{2n}_{\epsilon^{-}_{\text{tracking}}} \rangle$. A similar uncertainty estimation for the photon selection efficiency could not be performed, since no studies for the low energy photon selection efficiency (especially photons entering the ROE) were available at the time of finalizing this analysis.

MC Modeling of $N_{\text{charged}} \in$ X, $N_{\gamma} \in$ X and $E_{\text{miss}} - |\boldsymbol{p}_{\text{miss}}|$

The studies in Section 4.4 show that the q^2 resolution and bias are directly affected by the photon and charged particle multiplicities in the inclusive system and the $E_{\text{miss}} - |\boldsymbol{p}_{\text{miss}}|$ values. Differences between data and MC contribute a systematic uncertainty in how final-state particles are assigned to the signal and tag-side. To estimate the impact of a possible mis-modeling in these observables, correction weights are derived in a signal enriched region. This region is selected by the requirements $m_X < 3\,\text{GeV}$ and $p^*_\ell > 1\,\text{GeV}$.

Figure 5.22 shows a slight disagreement in all three observables. For the charged particle multiplicity, the simulated distribution is shifted towards higher values compared to the measured distribution In contrast, more photons are observed in the measured photon multiplicity than in the simulation. The correction weights are calculated as a bin-wise ratio of the respective normalized distributions in data and MC for $N_{\text{charged}} \in X$ and $N_{\gamma} \in X$ independently. Since $E_{\text{miss}} - |\boldsymbol{p}_{\text{miss}}|$ is directly related to the particle multiplicities, both the photon and charged particle multiplicity reweighing are applied before comparing $E_{\text{miss}} - |\boldsymbol{p}_{\text{miss}}|$. Here, the data distribution is slightly shifted towards positive values of $E_{\text{miss}} - |\boldsymbol{p}_{\text{miss}}|$ compared to the simulated spectrum. The differences in data and MC are parameterized by interpolating the bin-wise ratio of the normalized distributions with the same smoothed cubic spline approach used to determine the signal probability weight function described in Section 5.1.4.

Comparisons between the reconstructed q^2 distribution after applying the correction weights derived from the different observables and the nominal $B \to X_c \ell \nu_\ell$ expectation and are shown in Figure 5.22. Applying the $N_{\text{charged}} \in X$ and $E_{\text{miss}} - |\boldsymbol{p}_{\text{miss}}|$ reweighing results in small shift of the q^2 mean value with respect to the nominal expectation. The $N_{\gamma} \in X$ reweighing shows an opposite effect. Here, the reweighed spectrum is shifted towards lower q^2 values compared to the nominal spectrum.

To estimate the uncertainty on the q^2 moments, the calibration procedure is repeated with the different correction weights applied to the signal MC sample, extracting new calibration coefficients and bias corrections factors. As a measure for the uncertainties, the deviations from the varied results

Table 5.6: Size of the non-closure bias Δ_{rel} for q^2 moments of different order and for different lower q^2 thresholds observed in the sample test described in Section 5.3. All values are given in %.

q^2_{th} [GeV^2]	$\Delta_{\text{rel}}(\langle q^2 \rangle)$	$\Delta_{\text{rel}}(\langle q^4 \rangle)$	$\Delta_{\text{rel}}(\langle q^6 \rangle)$	$\Delta_{\text{rel}}(\langle q^8 \rangle)$
1.5	0.186	0.325	0.466	0.631
2.0	0.205	0.363	0.504	0.657
2.5	0.160	0.296	0.421	0.556
3.0	0.112	0.223	0.333	0.455
3.5	0.062	0.134	0.214	0.310
4.0	0.048	0.104	0.170	0.252
4.5	0.020	0.054	0.106	0.180
5.0	0.016	0.045	0.090	0.157
5.5	0.013	0.039	0.082	0.148
6.0	0.018	0.044	0.082	0.141
6.5	0.016	0.039	0.074	0.128
7.0	0.013	0.035	0.070	0.127
7.5	0.010	0.029	0.064	0.120
8.0	0.013	0.034	0.067	0.121
8.5	0.012	0.033	0.065	0.116

to the nominal measurement are used

$$\sigma_{\langle q^{2n}_{\text{reweighed}} \rangle} = \left| \langle q^{2n}_{\text{reweighed}} \rangle - \langle q^{2n} \rangle \right| . \tag{5.41}$$

The uncertainties are assumed to be fully correlated, taking into account the sign of the difference in the moments.

5.5.3 Other Uncertainties

Non-Closure Bias

As discussed in Section 5.3, a small non-closure bias is observed in the sample test checking the background subtraction and q^2 calibration procedure. This bias is introduced by the interpolation of the bin-wise signal probabilities with the cubic spline fit. Due to the small size of the bias, the final measured moments are not corrected but instead the size of the bias is added as an additional uncertainty

$$\sigma_{\text{non-closure}} = \langle q^{2n} \rangle \times \Delta_{\text{rel}}(\langle q^{2n} \rangle). \tag{5.42}$$

Table 5.6 lists the non-closure bias values in percent for moments of different order and for different lower q^2 thresholds. The largest biases are observed for moments with a lower q^2 threshold value between 1 GeV^2 and 3 GeV^2 and moments with higher order tend to have larger bias values. The covariance matrix for this uncertainty is constructed assuming a correlation of 100% between all measured moments.

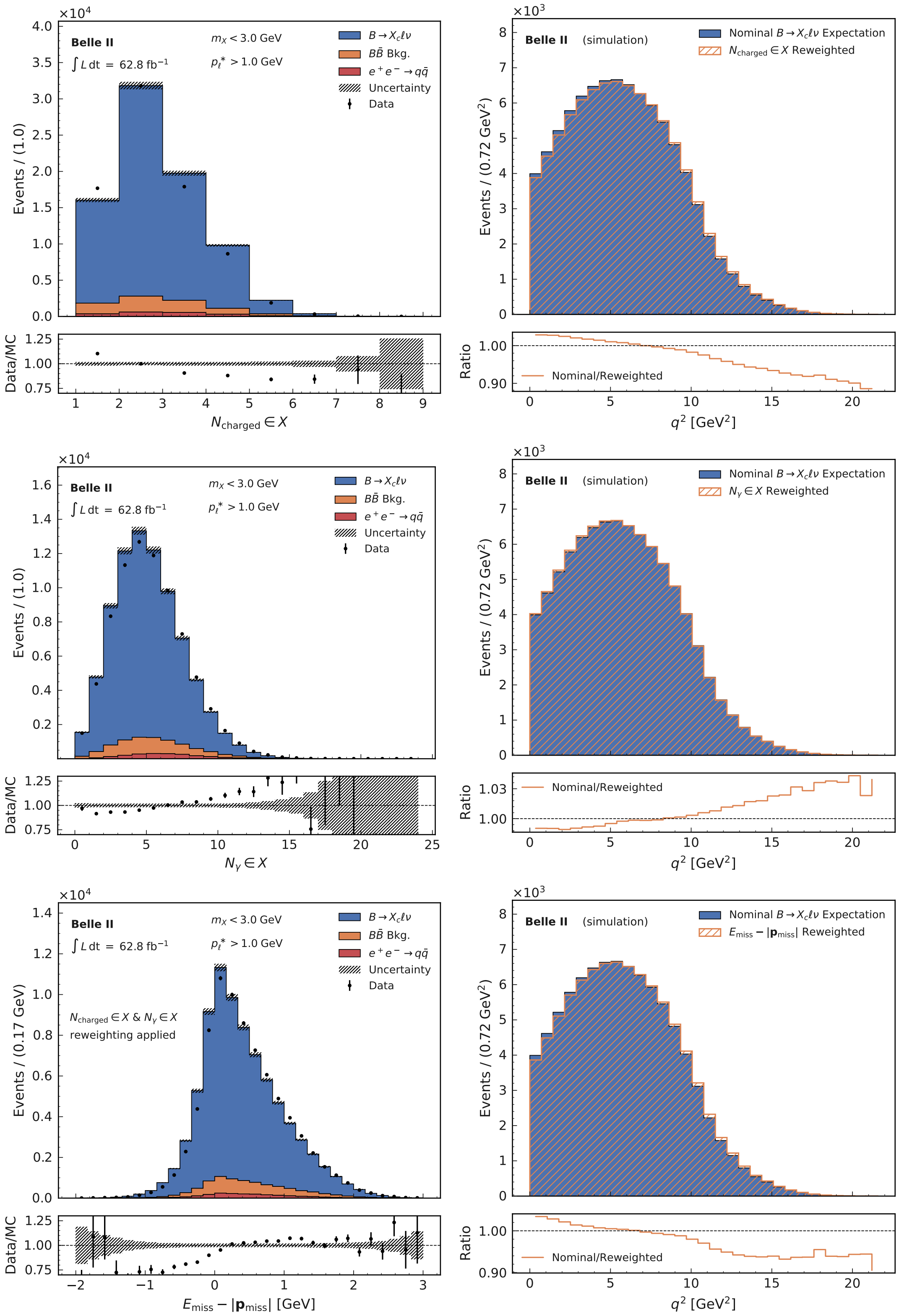


Figure 5.22: Comparison of data and MC distributions of $N_{\text{charged}} \in X$, $N_\gamma \in X$ and $E_{\text{miss}} - |\boldsymbol{p}_{\text{miss}}|$ in the signal enriched region of $m_X < 3\,\text{GeV}$ and $p_\ell^* > 1\,\text{GeV}$ (left). Reconstructed q^2 spectrum with the nominal $B \to X_c \ell \nu_\ell$ expectation and after applying the reweighing (right). The hatched uncertainty band represents the systematic uncertainty on the bin-wise ratio of data and MC, while the error bars show the statistical uncertainty.

5.6 Results

The measured $\langle q^{2n} \rangle$ moments for $n = 1$–4 are presented in Figure 5.23 as functions of lower q^2 thresholds ranging from $q^2 > 1.5\,\mathrm{GeV}^2$ to $q^2 > 8.5\,\mathrm{GeV}^2$ in $0.5\,\mathrm{GeV}^2$ increments. The numerical central values as well as the statistical and systematic uncertainties are listed in Table 5.7.

For all measured moments, the precision of the measurement is dominated by systematic uncertainties. Figure 5.24 shows the total relative systematic uncertainty for the raw q^2 moments. Individual contributions from related systematic uncertainties are grouped together.

The total relative uncertainty for the first moment $\langle q^2 \rangle$ decreases from 2.17% for a q^2 threshold of $q^2 > 1.5\,\mathrm{GeV}^2$ to 0.58% for the highest q^2 threshold $q^2 > 8.5\,\mathrm{GeV}^2$. The same trend is observed for higher order q^2 moments. For each moment, the total systematic uncertainty decreases with increasing q^2 thresholds, whereas the contribution of the statistical uncertainty increases. At low q^2 thresholds and for the first and second moments, the uncertainty from background subtraction, the $B \to X_c \ell \nu_\ell$ modeling and the impact of the mis-modeling of the number of charged particles in the X system are of similar size. The uncertainty due to the BGL form factor parameter for $B \to D^{(*)} \ell \nu_\ell$ decays is small compared to the impact of the branching fraction and X_c composition on the q^2 calibration. At high q^2 thresholds, the uncertainty due to the MC simulation statistics affecting the calibration coefficients and bias correction factors becomes sizable for the first and second q^2 moments. For the third and fourth moments, the dominant systematic uncertainty contribution comes from the mis-modeling of particle multiplicities in the inclusive system and of $E_{\mathrm{miss}} - |\boldsymbol{p}_{\mathrm{miss}}|$. The uncertainty associated with the smoothing parameter of the cubic spline describing the signal probabilities turns out to be negligible for all measurements. The uncertainty associated with the non-closure bias is sub-leading for all moments.

A more detailed breakdown of the relative statistical and systematic uncertainties for the measured raw moments $\langle q^{2n} \rangle$ with $n = 1$–4 and for different lower q^2 thresholds is given in Tables 5.9 to 5.12.

The correlation structure obtained from the full experimental covariance matrix of the measured raw q^2 moments is visualized in Figure 5.26. The observed correlation between individual moments is higher than the pure statistical expectation since several systematic uncertainty sources are assumed to be fully correlated. Numerical values for the correlations between the individual moments are summarized in Figures C.5 to C.8.

Figure 5.23 also shows the moments calculated from the generator $B \to X_c \ell \nu_\ell$ sample in comparison. The simulated moments include uncertainties on the $B \to X_c \ell \nu_\ell$ composition and $B \to D^{(*)} \ell \nu_\ell$ BGL form factor parameters. Although the measured moments tend to slightly higher values compared to the moments obtained from the assumed X_c model at lower values of q^2_{th}, no significant deviations are observed. Measured and simulated q^2 moments of the same order are compared quantitatively with χ^2 tests. Within these tests, the $B \to X_c \ell \nu_\ell$ composition and form factor parameter uncertainties are assumed to be 100% correlated between the measured and simulated q^2 moments. In each test, only measurements with a correlation below 95% are included to obtain numerically stable results since the actual χ^2 test statistic is sensitive to highly correlated measurements. Since the uncertainty is dominated by the systematic contribution, also the total covariance is dominated by the systematic covariance matrices. The latter are constructed under certain assumptions and thus, it is likely possible that these high correlations might not be estimated accurately enough. The resulting p-values range from 27% to 94%.

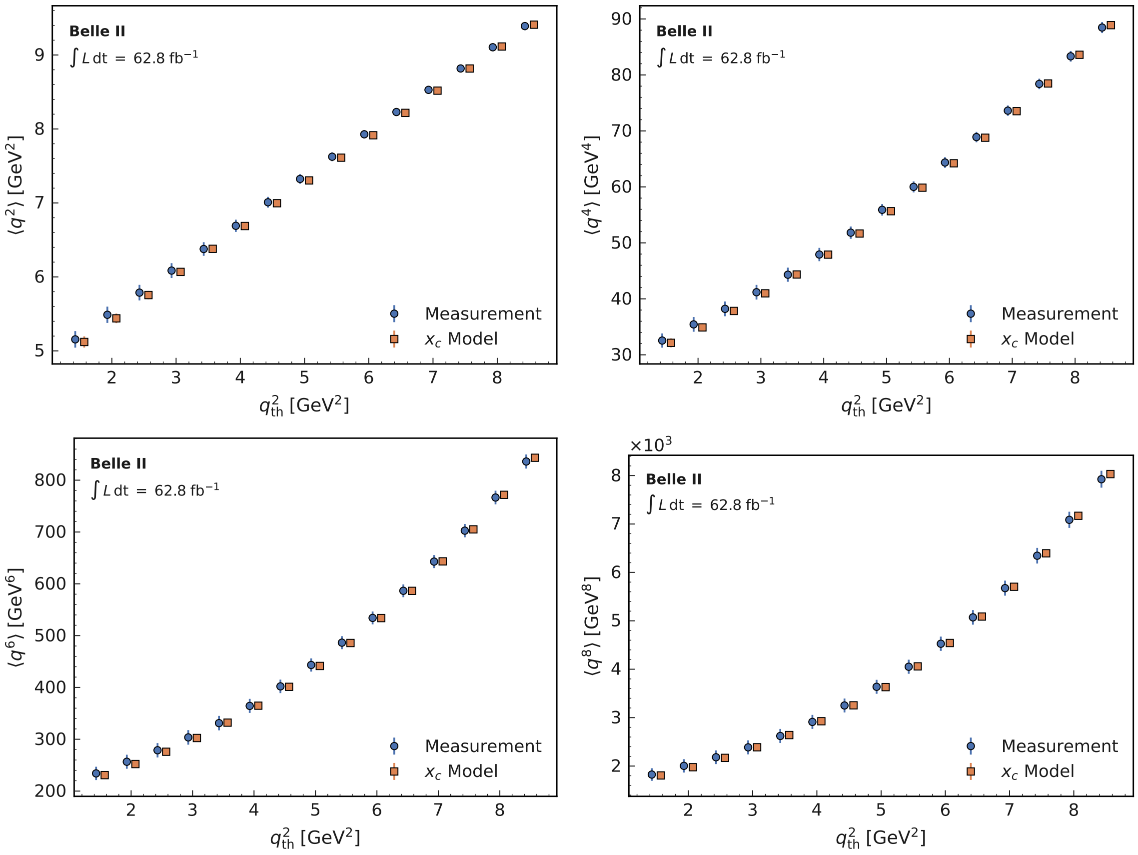


Figure 5.23: q^2 moments (blue) as functions of q^2 threshold with full uncertainties. The simulated moments (orange) are shown for comparison.

The central q^2 moments are calculated by expanding the binomial relation

$$\langle (q^2 - \langle q^2 \rangle)^n \rangle = \sum_{j=0}^{n} \binom{n}{j} (-1)^{n-j} \langle q^{2j} \rangle \langle q^2 \rangle^{n-j}, \tag{5.43}$$

and applying the following non-linear transformation

$$\begin{pmatrix} \langle q^2 \rangle \\ \langle q^4 \rangle \\ \langle q^6 \rangle \\ \langle q^8 \rangle \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} \langle q^2 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^2 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^3 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^4 \rangle \end{pmatrix}. \tag{5.44}$$

The covariance matrix of the central moments C' is calculated using Gaussian uncertainty propagation $C' = J\,C\,J^\intercal$. Here, J is the Jacobian matrix for the transformation in Equation (5.44). The Jacobian matrix is calculated numerically with the `numdifftools` package [119]. Figure 5.25 shows the central q^2 moments of order two to four as functions of lower q^2 thresholds, while their numerical values are summarized in Table 5.8. Comparing measured and simulated central moments

Table 5.7: Numerical values of measured q^2 moments of order one to four as functions of lower q^2 thresholds. The stated uncertainties are separated in a statistical and systematic contribution.

q^2_{th} [Gev^2]	$\langle q^2 \rangle$ [Gev^2]	$\langle q^4 \rangle$ [Gev^4]	$\langle q^6 \rangle$ [Gev^6]	$\langle q^8 \rangle$ [Gev^8]
1.5	5.16 ± 0.01 ± 0.11	32.55 ± 0.16 ± 1.26	234.11 ± 1.75 ± 12.74	1824.48 ± 19.82 ± 128.57
2.0	5.49 ± 0.01 ± 0.11	35.44 ± 0.16 ± 1.31	256.58 ± 1.82 ± 13.44	2003.76 ± 20.91 ± 135.90
2.5	5.79 ± 0.01 ± 0.10	38.21 ± 0.16 ± 1.31	278.78 ± 1.89 ± 13.73	2182.03 ± 21.97 ± 139.69
3.0	6.09 ± 0.01 ± 0.10	41.18 ± 0.16 ± 1.30	303.60 ± 1.96 ± 13.93	2386.22 ± 23.09 ± 142.87
3.5	6.38 ± 0.01 ± 0.09	44.31 ± 0.17 ± 1.25	331.14 ± 2.04 ± 13.76	2621.05 ± 24.37 ± 143.28
4.0	6.69 ± 0.01 ± 0.08	47.92 ± 0.17 ± 1.18	364.36 ± 2.15 ± 13.40	2911.47 ± 26.10 ± 142.76
4.5	7.01 ± 0.01 ± 0.07	51.82 ± 0.18 ± 1.08	402.07 ± 2.29 ± 12.85	3251.23 ± 28.20 ± 141.01
5.0	7.32 ± 0.01 ± 0.06	55.90 ± 0.18 ± 1.02	443.33 ± 2.44 ± 12.51	3636.73 ± 30.52 ± 141.07
5.5	7.62 ± 0.01 ± 0.06	60.00 ± 0.19 ± 0.96	486.42 ± 2.61 ± 12.25	4051.07 ± 33.18 ± 141.75
6.0	7.93 ± 0.01 ± 0.05	64.35 ± 0.20 ± 0.93	534.18 ± 2.79 ± 12.10	4526.33 ± 36.09 ± 143.31
6.5	8.23 ± 0.01 ± 0.05	68.90 ± 0.21 ± 0.90	586.53 ± 2.99 ± 12.00	5071.04 ± 39.44 ± 145.23
7.0	8.53 ± 0.01 ± 0.05	73.62 ± 0.22 ± 0.89	642.87 ± 3.24 ± 12.12	5675.17 ± 43.57 ± 149.30
7.5	8.82 ± 0.01 ± 0.05	78.40 ± 0.23 ± 0.90	702.59 ± 3.49 ± 12.34	6344.75 ± 47.78 ± 153.80
8.0	9.10 ± 0.01 ± 0.05	83.33 ± 0.24 ± 0.92	766.54 ± 3.77 ± 12.62	7085.85 ± 52.71 ± 158.80
8.5	9.39 ± 0.01 ± 0.05	88.47 ± 0.26 ± 0.95	836.02 ± 4.11 ± 12.99	7924.67 ± 58.56 ± 164.49

results in p-values greater than 98%.

The overall correlation structure obtained from the full experimental covariance matrix for the moments $\langle q^2 \rangle$ and $\langle (q^2 - \langle q^2 \rangle)^n \rangle$ with $n = 2$–4 is illustrated in Figure 5.27. Compared to the raw moments, the central moments offer the advantage of less correlation between moments of different orders. Negative correlation values are also observed. The numerical values ot the correlation between individual moments are summarized in Figures C.9 to C.12.

A similar measurement of the q^2 moments to this one was reported by the Belle Collaboration [31]. This analysis presents new additional measurements of raw and central q^2 moments with comparable precision starting at lower q^2 thresholds of 1.5, 2.0 and 2.5 GeV^2. This probes up to 77% of the accessible $B \to X_c \ell \nu_\ell$ phase space and retains more information about the inclusive X_c spectrum compared to the 58% phase space coverage of the Belle measurement starting at a lower q^2 threshold of $q^2 > 3.0\,\text{GeV}^2$. A χ^2 test is used to compare the overlapping measurements of raw moments from both analyses for lower q^2 thresholds between 3.0 GeV^2 and 8.5 GeV^2. Again, only measurements with an observed correlation below 95% are compared. For the test, the systematic uncertainties for the modeling of the $B \to X_c \ell \nu_\ell$ spectrum are assumed to be fully correlated between the Belle and Belle II measurements. The resulting p-values range between 5% and 72%.

All numerical results for the measured raw and central moments as well as their covariance matrices are also provided on HEPData.

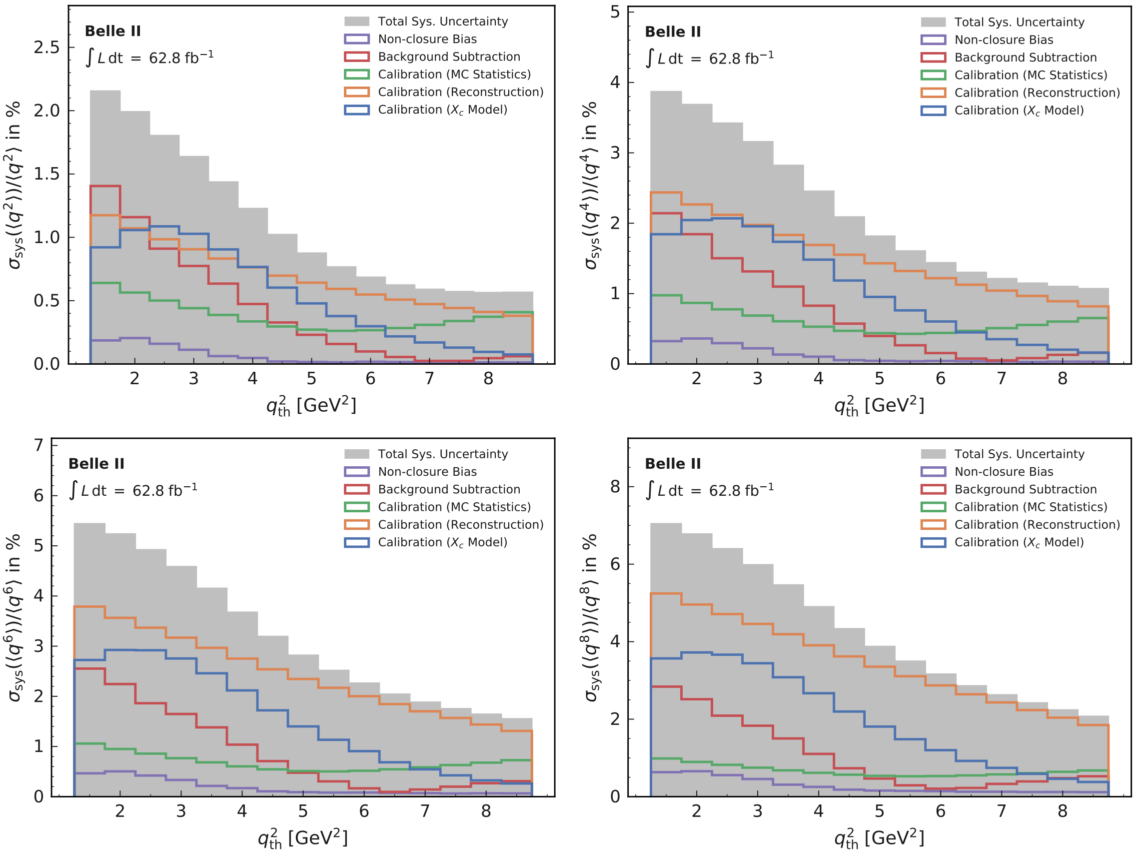


Figure 5.24: Total (gray) and grouped (colored histograms) relative systematic uncertainties of the raw q^2 moments as functions of q^2 threshold are shown.

Table 5.8: Numerical values of the measured centra q^2 moments of order one to four as functions of lower q^2 thresholds. The stated uncertainties are separated in a statistical and systematic contribution.

q^2_{th} [GeV^2]	$\langle(\mathrm{q}^2 - \langle\mathrm{q}^2\rangle)^2\rangle$ [GeV^4]	$\langle(\mathrm{q}^2 - \langle\mathrm{q}^2\rangle)^3\rangle$ [GeV^6]	$\langle(\mathrm{q}^2 - \langle\mathrm{q}^2\rangle)^4\rangle$ [GeV^8]
1.5	5.97 ± 0.04 ± 0.24	4.74 ± 0.23 ± 0.81	68.22 ± 2.12 ± 6.56
2.0	5.32 ± 0.04 ± 0.20	3.74 ± 0.22 ± 0.66	53.57 ± 1.89 ± 5.12
2.5	4.72 ± 0.04 ± 0.18	3.05 ± 0.21 ± 0.56	41.78 ± 1.69 ± 4.14
3.0	4.15 ± 0.04 ± 0.17	2.49 ± 0.20 ± 0.50	32.26 ± 1.49 ± 3.38
3.5	3.65 ± 0.04 ± 0.16	2.07 ± 0.19 ± 0.45	24.98 ± 1.32 ± 2.73
4.0	3.15 ± 0.04 ± 0.15	1.59 ± 0.18 ± 0.39	18.58 ± 1.16 ± 2.14
4.5	2.69 ± 0.04 ± 0.14	1.08 ± 0.18 ± 0.34	13.28 ± 1.02 ± 1.66
5.0	2.26 ± 0.04 ± 0.13	0.73 ± 0.17 ± 0.32	9.40 ± 0.89 ± 1.30
5.5	1.87 ± 0.04 ± 0.12	0.44 ± 0.16 ± 0.31	6.33 ± 0.79 ± 1.17
6.0	1.50 ± 0.04 ± 0.12	0.18 ± 0.16 ± 0.31	3.98 ± 0.69 ± 1.27
6.5	1.18 ± 0.04 ± 0.12	0.06 ± 0.15 ± 0.31	2.63 ± 0.62 ± 1.52
7.0	0.88 ± 0.04 ± 0.12	−0.08 ± 0.15 ± 0.30	1.53 ± 0.57 ± 1.88
7.5	0.65 ± 0.04 ± 0.13	−0.08 ± 0.14 ± 0.27	1.07 ± 0.53 ± 2.23
8.0	0.44 ± 0.04 ± 0.14	−0.08 ± 0.14 ± 0.24	0.69 ± 0.49 ± 2.59
8.5	0.28 ± 0.04 ± 0.16	−0.03 ± 0.14 ± 0.21	0.49 ± 0.45 ± 2.92

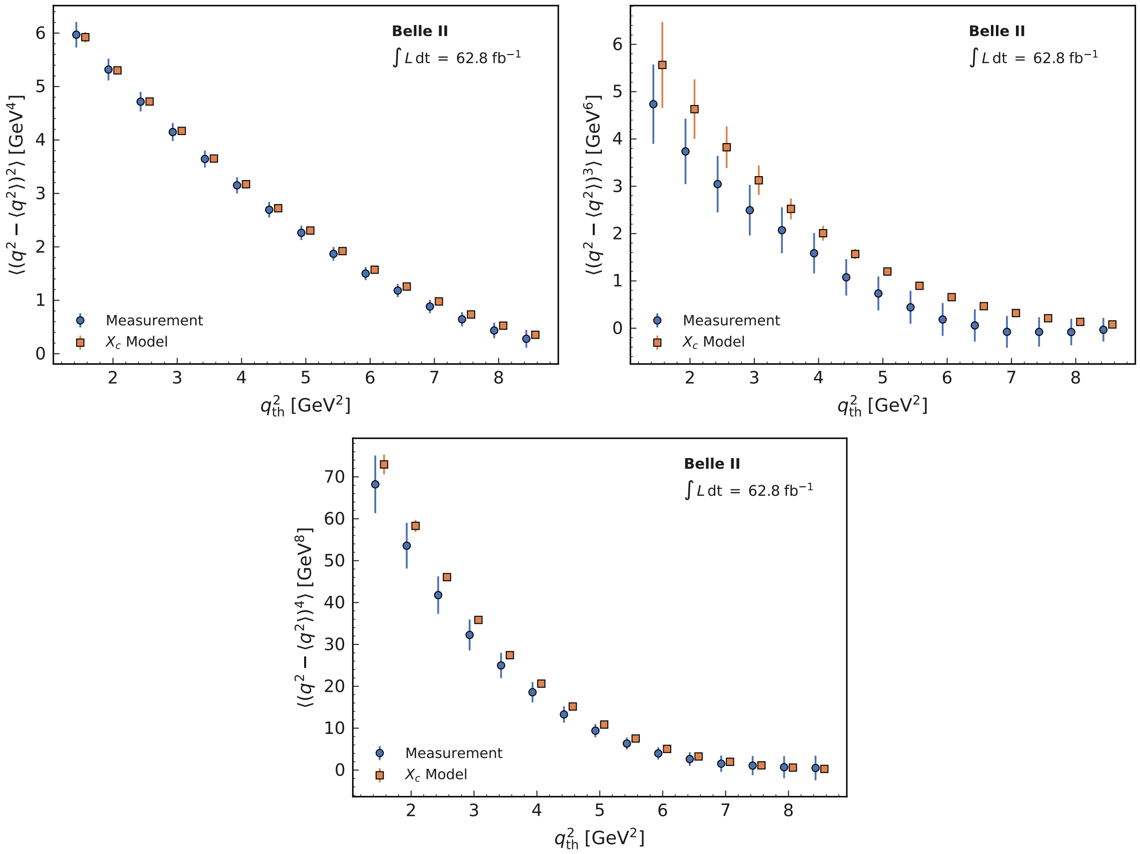


Figure 5.25: Central q^2 moments (blue) as functions of q^2 threshold with full uncertainties. The simulated moments (orange) are shown for comparison.

Table 5.9: Breakdown of systematic uncertainties for the measurement of $\langle q^2 \rangle$. All uncertainties are given as relative uncertainties in %.

	q^2_{th} [Gev^2]	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
Calibration (MC Statistics)	Calib. Curve (Stat. Unc.)	0.63	0.56	0.49	0.43	0.38	0.33	0.29	0.26	0.25	0.26	0.28	0.30	0.33	0.37	0.40
	Bias Corr. (Stat. Unc.)	0.10	0.09	0.09	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.06	0.06
Calibration (X_c Model)	$\mathcal{B}(B \to D\ell\nu)$	0.10	0.09	0.08	0.07	0.06	0.05	0.04	0.04	0.03	0.02	0.02	0.01	0.01	0.00	0.00
	$\mathcal{B}(B \to D^*\ell\nu)$	0.33	0.29	0.24	0.21	0.17	0.14	0.11	0.09	0.07	0.05	0.04	0.03	0.02	0.01	0.00
	$\mathcal{B}(B \to D^{**}\ell\nu)$	0.71	0.63	0.55	0.48	0.40	0.34	0.28	0.23	0.18	0.13	0.10	0.07	0.05	0.03	0.02
	Non-Res. X_c Dropped	0.31	0.63	0.75	0.76	0.69	0.60	0.48	0.39	0.32	0.25	0.18	0.14	0.11	0.08	0.06
	Non-Res. X_c Repl. w/ D_1', D_0^*	0.34	0.49	0.51	0.45	0.37	0.29	0.18	0.10	0.04	0.02	0.00	0.03	0.03	0.03	0.01
	$B \to D\ell\nu$ Form Factor	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$B \to D^*\ell\nu$ Form Factor	0.08	0.07	0.07	0.07	0.06	0.06	0.06	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.03
Calibration (Reconstruction)	PID Uncertainty	0.14	0.12	0.11	0.09	0.08	0.07	0.05	0.04	0.03	0.02	0.02	0.01	0.01	0.01	0.01
	$N_\gamma \in X$ Reweighted	0.30	0.27	0.24	0.22	0.20	0.18	0.16	0.14	0.14	0.13	0.13	0.12	0.11	0.10	0.10
	$N_{\text{charged}} \in X$ Reweighted	1.09	1.00	0.92	0.85	0.78	0.72	0.65	0.60	0.55	0.51	0.47	0.44	0.41	0.38	0.35
	$E_{\text{miss}} - \lvert \boldsymbol{p}_{\text{miss}} \rvert$ Reweighted	0.26	0.22	0.21	0.19	0.18	0.17	0.15	0.15	0.14	0.14	0.13	0.12	0.12	0.11	0.09
	Tracking Efficiency	0.13	0.12	0.11	0.10	0.09	0.09	0.08	0.07	0.06	0.06	0.05	0.05	0.05	0.04	0.04
Background Subtraction	Spline Smooth. Factor	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
	Bkg. Yield & Shape	1.41	1.16	0.91	0.77	0.63	0.47	0.33	0.23	0.16	0.10	0.06	0.02	0.02	0.05	0.06
Other	Non-Closure Bias	0.19	0.20	0.16	0.11	0.06	0.05	0.02	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.01
	Stat. Uncertainty	0.27	0.24	0.21	0.19	0.18	0.17	0.16	0.15	0.14	0.14	0.14	0.13	0.13	0.13	0.13
	Syst. Uncertainty	2.16	1.99	1.80	1.64	1.44	1.23	1.02	0.88	0.77	0.69	0.62	0.59	0.57	0.57	0.57
	Total Uncertainty	2.17	2.01	1.82	1.65	1.45	1.24	1.03	0.89	0.78	0.70	0.64	0.61	0.59	0.58	0.58

Table 5.10: Breakdown of systematic uncertainties for the measurement of $\langle q^4 \rangle$. All uncertainties are given as relative uncertainties in %.

	q^2_{th} [Gev^2]	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
Calibration (MC Statistics)	Calib. Curve (Stat. Unc.)	0.96	0.85	0.75	0.67	0.58	0.50	0.44	0.41	0.40	0.42	0.45	0.49	0.54	0.59	0.64
	Bias Corr. (Stat. Unc.)	0.20	0.19	0.18	0.18	0.17	0.16	0.16	0.15	0.15	0.14	0.14	0.14	0.14	0.14	0.14
Calibration (X_c Model)	$\mathcal{B}(B \to D \ell \nu)$	0.18	0.16	0.15	0.13	0.12	0.10	0.08	0.07	0.06	0.04	0.03	0.02	0.01	0.01	0.01
	$\mathcal{B}(B \to D^* \ell \nu)$	0.60	0.52	0.45	0.39	0.33	0.27	0.22	0.18	0.14	0.10	0.07	0.05	0.03	0.02	0.01
	$\mathcal{B}(B \to D^{**} \ell \nu)$	1.30	1.17	1.04	0.91	0.79	0.67	0.56	0.45	0.36	0.27	0.20	0.14	0.09	0.06	0.05
	Non-Res. X_c Dropped	0.91	1.31	1.47	1.47	1.35	1.18	0.96	0.79	0.64	0.52	0.38	0.30	0.23	0.16	0.13
	Non-Res. X_c Repl. w/ D_1', D_0^*	0.69	0.87	0.89	0.79	0.66	0.51	0.31	0.17	0.07	0.03	0.02	0.06	0.07	0.06	0.03
	$B \to D \ell \nu$ Form Factor	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	$B \to D^* \ell \nu$ Form Factor	0.17	0.16	0.15	0.15	0.14	0.13	0.12	0.12	0.11	0.10	0.10	0.09	0.09	0.08	0.08
Calibration (Reconstruction)	PID Uncertainty	0.25	0.23	0.20	0.17	0.15	0.13	0.10	0.08	0.06	0.05	0.04	0.03	0.02	0.02	0.02
	$N_\gamma \in X$ Reweighted	0.61	0.57	0.52	0.49	0.45	0.40	0.36	0.33	0.32	0.30	0.28	0.26	0.25	0.23	0.22
	$N_{\mathrm{charged}} \in X$ Reweighted	2.27	2.11	1.98	1.85	1.72	1.58	1.46	1.34	1.24	1.14	1.05	0.97	0.90	0.83	0.76
	$E_{\mathrm{miss}} - \lvert \boldsymbol{p}_{\mathrm{miss}} \rvert$ Reweighted	0.53	0.48	0.45	0.42	0.39	0.37	0.34	0.32	0.31	0.30	0.28	0.26	0.24	0.21	0.18
	Tracking Efficiency	0.28	0.26	0.24	0.22	0.20	0.19	0.17	0.16	0.14	0.13	0.12	0.11	0.10	0.09	0.09
Background Subtraction	Spline Smooth. Factor	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
	Bkg. Yield & Shape	2.14	1.84	1.50	1.32	1.10	0.83	0.57	0.40	0.27	0.16	0.08	0.05	0.08	0.13	0.16
Other	Non-Closure Bias	0.32	0.36	0.30	0.22	0.13	0.10	0.05	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.03
	Stat. Uncertainty	0.49	0.46	0.43	0.40	0.38	0.36	0.34	0.33	0.32	0.31	0.30	0.30	0.29	0.29	0.29
	Syst. Uncertainty	3.87	3.69	3.42	3.16	2.82	2.46	2.09	1.82	1.61	1.44	1.30	1.21	1.15	1.10	1.07
	Total Uncertainty	3.90	3.72	3.45	3.18	2.85	2.48	2.12	1.85	1.64	1.47	1.34	1.25	1.19	1.14	1.11

Table 5.11: Breakdown of systematic uncertainties for the measurement of $\langle q^6 \rangle$. All uncertainties are given as relative uncertainties in %.

	q^2_{th} [Gev^2]	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
Calibration (MC Statistics)	Calib. Curve (Stat. Unc.)	1.01	0.90	0.80	0.71	0.62	0.54	0.48	0.44	0.44	0.46	0.49	0.54	0.59	0.64	0.69
	Bias Corr. (Stat. Unc.)	0.31	0.31	0.30	0.29	0.28	0.27	0.26	0.25	0.25	0.24	0.23	0.23	0.23	0.22	0.22
Calibration (X_c Model)	$\mathcal{B}(B \to D\ell\nu)$	0.24	0.22	0.20	0.18	0.16	0.14	0.12	0.10	0.08	0.06	0.04	0.02	0.01	0.01	0.02
	$\mathcal{B}(B \to D^*\ell\nu)$	0.79	0.70	0.62	0.54	0.46	0.39	0.32	0.26	0.20	0.15	0.11	0.07	0.05	0.02	0.02
	$\mathcal{B}(B \to D^{**}\ell\nu)$	1.75	1.60	1.44	1.29	1.12	0.96	0.81	0.67	0.53	0.40	0.30	0.22	0.15	0.10	0.08
	Non-Res. X_c Dropped	1.66	2.03	2.16	2.12	1.94	1.71	1.41	1.17	0.96	0.78	0.58	0.46	0.35	0.25	0.20
	Non-Res. X_c Repl. w/ D_1', D_0^*	0.93	1.12	1.14	1.02	0.85	0.66	0.40	0.21	0.08	0.03	0.04	0.09	0.11	0.10	0.05
	$B \to D\ell\nu$ Form Factor	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01
	$B \to D^*\ell\nu$ Form Factor	0.27	0.26	0.25	0.24	0.23	0.21	0.20	0.19	0.18	0.17	0.17	0.16	0.15	0.14	0.14
Calibration (Reconstruction)	PID Uncertainty	0.34	0.31	0.28	0.24	0.21	0.18	0.15	0.12	0.10	0.08	0.06	0.05	0.04	0.04	0.04
	$N_\gamma \in X$ Reweighted	0.96	0.90	0.84	0.79	0.73	0.67	0.61	0.56	0.53	0.50	0.48	0.44	0.42	0.38	0.36
	$N_{\text{charged}} \in X$ Reweighted	3.53	3.33	3.15	2.97	2.78	2.58	2.38	2.20	2.03	1.87	1.72	1.58	1.46	1.34	1.23
	$E_{\text{miss}} - \lvert\boldsymbol{p}_{\text{miss}}\rvert$ Reweighted	0.81	0.74	0.70	0.67	0.62	0.58	0.54	0.50	0.48	0.46	0.42	0.39	0.35	0.30	0.25
	Tracking Efficiency	0.43	0.40	0.38	0.35	0.33	0.30	0.28	0.26	0.23	0.22	0.20	0.18	0.17	0.15	0.14
Background Subtraction	Spline Smooth. Factor	0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01
	Bkg. Yield & Shape	2.55	2.24	1.86	1.65	1.38	1.04	0.71	0.48	0.30	0.17	0.10	0.14	0.20	0.27	0.31
Other	Non-Closure Bias	0.47	0.50	0.42	0.33	0.21	0.17	0.11	0.09	0.08	0.08	0.07	0.07	0.06	0.07	0.07
	Stat. Uncertainty	0.75	0.71	0.68	0.64	0.62	0.59	0.57	0.55	0.54	0.52	0.51	0.50	0.50	0.49	0.49
	Syst. Uncertainty	5.44	5.24	4.92	4.59	4.16	3.68	3.20	2.82	2.52	2.27	2.05	1.89	1.76	1.65	1.55
	Total Uncertainty	5.49	5.29	4.97	4.63	4.20	3.73	3.25	2.87	2.58	2.32	2.11	1.95	1.82	1.72	1.63

Table 5.12: Breakdown of systematic uncertainties for the measurement of $\langle q^8 \rangle$. All uncertainties are given as relative uncertainties in %.

	q^2_{th} [Gev^2]	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
Calibration (MC Statistics)	Calib. Curve (Stat. Unc.)	0.87	0.77	0.69	0.61	0.54	0.46	0.41	0.38	0.38	0.39	0.42	0.46	0.50	0.55	0.59
	Bias Corr. (Stat. Unc.)	0.46	0.45	0.44	0.43	0.42	0.41	0.39	0.38	0.37	0.36	0.35	0.34	0.34	0.33	0.33
Calibration (X_c Model)	$\mathcal{B}(B \to D\ell\nu)$	0.28	0.26	0.24	0.22	0.19	0.17	0.14	0.12	0.09	0.07	0.04	0.02	0.02	0.01	0.03
	$\mathcal{B}(B \to D^*\ell\nu)$	0.93	0.84	0.74	0.65	0.57	0.48	0.40	0.32	0.25	0.19	0.14	0.09	0.06	0.03	0.02
	$\mathcal{B}(B \to D^{**}\ell\nu)$	2.10	1.94	1.77	1.59	1.40	1.22	1.03	0.86	0.69	0.53	0.40	0.29	0.20	0.14	0.11
	Non-Res. X_c Dropped	2.47	2.76	2.82	2.72	2.48	2.17	1.81	1.51	1.25	1.02	0.78	0.62	0.47	0.35	0.28
	Non-Res. X_c Repl. w/ D_1', D_0^*	1.06	1.26	1.28	1.15	0.96	0.75	0.46	0.24	0.08	0.02	0.07	0.14	0.16	0.14	0.08
	$B \to D\ell\nu$ Form Factor	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.02
	$B \to D^*\ell\nu$ Form Factor	0.38	0.37	0.35	0.34	0.33	0.31	0.30	0.28	0.27	0.26	0.25	0.24	0.23	0.22	0.21
Calibration (Reconstruction)	PID Uncertainty	0.41	0.38	0.34	0.30	0.26	0.23	0.19	0.16	0.13	0.11	0.09	0.07	0.06	0.06	0.05
	$N_\gamma \in X$ Reweighted	1.35	1.27	1.20	1.13	1.06	0.98	0.90	0.83	0.79	0.74	0.70	0.65	0.61	0.56	0.52
	$N_{\text{charged}} \in X$ Reweighted	4.90	4.64	4.41	4.18	3.93	3.67	3.40	3.15	2.91	2.69	2.47	2.27	2.09	1.91	1.74
	$E_{\text{miss}} - \lvert \boldsymbol{p}_{\text{miss}} \rvert$ Reweighted	1.09	1.00	0.95	0.90	0.85	0.79	0.73	0.68	0.64	0.60	0.55	0.50	0.44	0.38	0.31
	Tracking Efficiency	0.58	0.55	0.52	0.49	0.46	0.43	0.40	0.37	0.34	0.31	0.28	0.26	0.24	0.21	0.19
Background Subtraction	Spline Smooth. Factor	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.03	0.02
	Bkg. Yield & Shape	2.84	2.51	2.09	1.83	1.50	1.10	0.73	0.47	0.29	0.21	0.22	0.33	0.39	0.48	0.53
Other	Non-Closure Bias	0.63	0.66	0.56	0.45	0.31	0.25	0.18	0.16	0.15	0.14	0.13	0.13	0.12	0.12	0.12
	Stat. Uncertainty	1.09	1.04	1.01	0.97	0.93	0.90	0.87	0.84	0.82	0.80	0.78	0.77	0.75	0.74	0.74
	Syst. Uncertainty	7.05	6.78	6.40	5.99	5.47	4.90	4.34	3.88	3.50	3.17	2.86	2.63	2.42	2.24	2.08
	Total Uncertainty	7.13	6.86	6.48	6.06	5.55	4.98	4.42	3.97	3.59	3.26	2.97	2.74	2.54	2.36	2.20

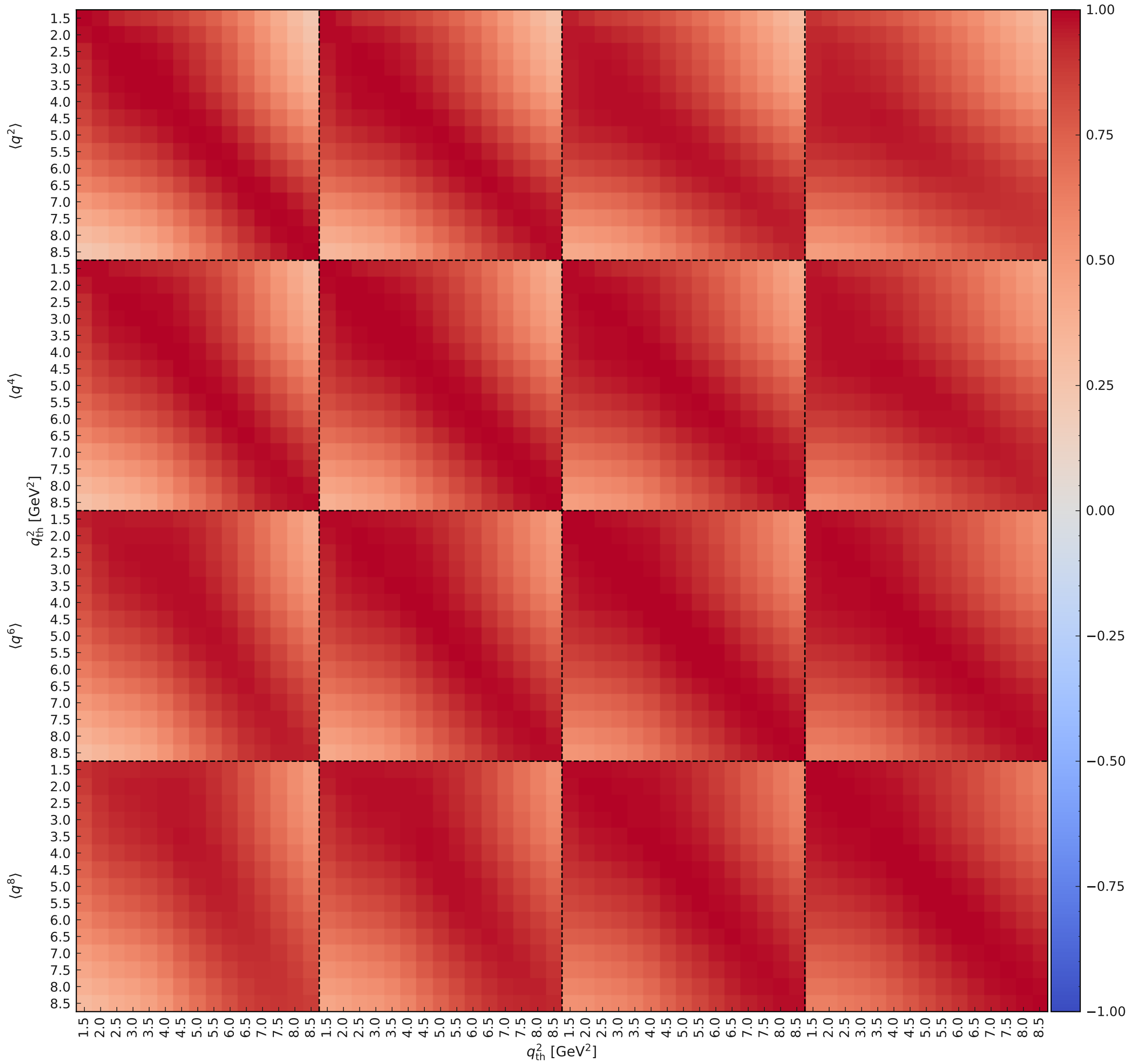


Figure 5.26: Experimental correlations between $\langle q^2 \rangle$ and $\langle q^{2n} \rangle$ for n = 1–4.

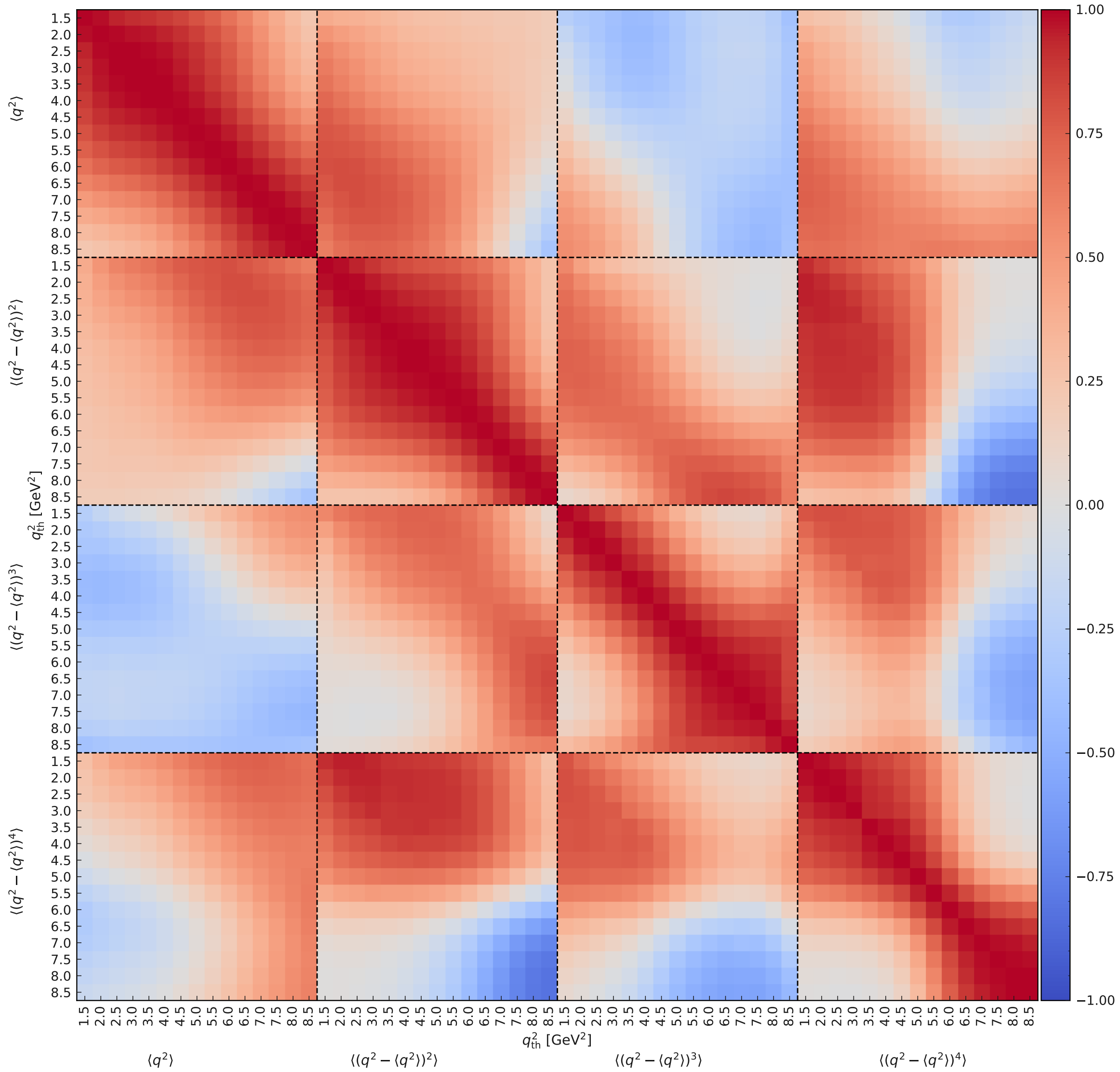


Figure 5.27: Correlations between $\langle q^2 \rangle$ and $\langle (q^2 - \langle q^2 \rangle)^n \rangle$ for n = 2–4 and for central moments of different order.

5.7 Stability Checks

Two additional test are performed to check the stability of measurement. First the measurement is repeated on statistically independent data samples. The second test investigates the impact of artificially altered shapes of the assumed q^2 distribution in the calibration procedure.

5.7.1 Measurement on Independent Data Samples

For this test, two different approaches in splitting the data set are considered: A division by lepton flavor in an electron and muon sub-sample and a separation by the charge of the reconstructed tag-side B meson in a charged and neutral sub-sample.

For both approaches, the moment extraction procedure is repeated with independent determinations of the signal probability function w, q^2 calibration coefficients c_n and m_n and independent bias correction factors C_{calib} and C_{gen}. In the test considering the sample split by lepton flavor, the m_X fit has to be repeated also, since the nominal approach determines the background normalizations only for reconstruction channels with combined electron and muon candidates.

Figure 5.28 shows the first raw q^2 moment as a function of the lower q^2 threshold measured on sub-samples split by lepton flavor and B charge, respectively. The comparisons for higher order moments are shown in Figures D.1 and D.2 Here, the considered uncertainties only include the statistical uncertainty on the moments and the uncertainties on the calibration coefficients and bias corrections factors. The latter are also considered to be independent since different MC sub-samples are used in their determination. Almost all individually measured moments agree within one standard deviation except the q^2 moments measured on the electron and muon subsample for $q^2 > 3.5\,\text{GeV}^2$. Here a deviation up to $1.5\,\sigma$ is observed.

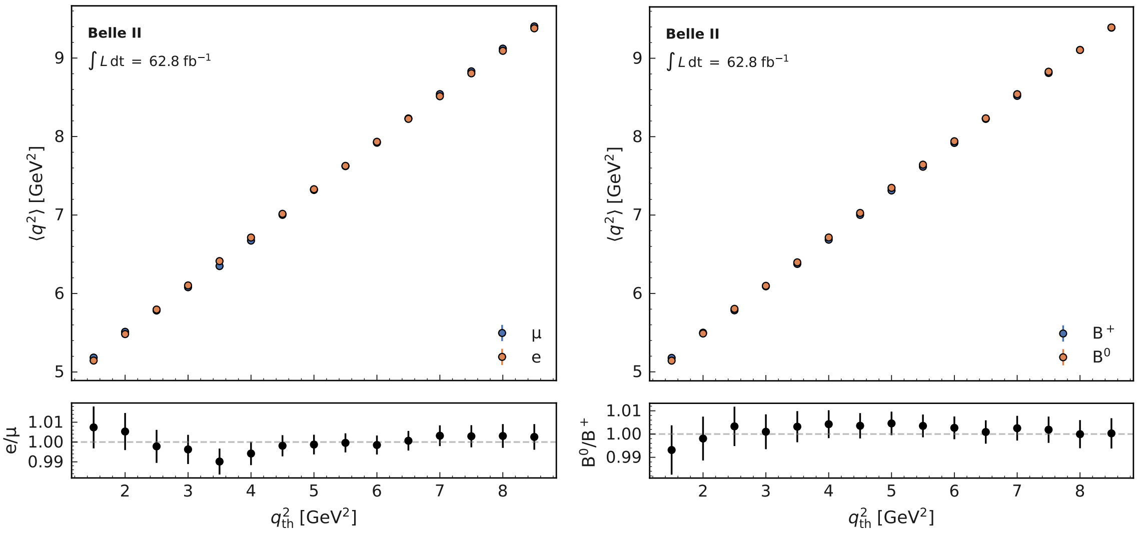


Figure 5.28: Comparison of the first raw q^2 moment as a function of the lower q^2 threshold measured on statistically independent muon and electron (left) and B^0 and B^+ (right) sub-samples.

5.7.2 Robustness Against Altered q^2 Shapes

One important aspect in the analysis is that a certain q^2 signal shape is assumed based on the $B \to X_c \ell \nu_\ell$ simulation. This affects both key steps in the analysis, the determination of the background normalizations with the binned likelihood of the m_X distribution and calibration of the reconstructed q^2 values. In the m_X fit itself, the dependence on the actual signal and background shapes is already reduced by using a very coarse binning of only five bins. Also, in the signal template the $B \to X_c \ell \nu_\ell$ composition and form factor uncertainties are considered as a shape uncertainty.

To further test the dependence of the extracted moments on the assumed $B \to X_c \ell \nu_\ell$ modeling and thus, the assumed shape of the q^2 distribution, several bias test scenarios are studied to check the robustness of the analysis procedure. These tests are setup as follows:

1. Create an artificial mock data set which is based on an altered q^2 distribution. This is achieved by modifying the underlying $B \to X_c \ell \nu_\ell$ modeling.
2. Analyze the mock data set by repeating the m_X fit, determining new signal probability weights and derive new calibration coefficients using the nominal analysis procedure to extract the first to fourth q^2 moments.
3. Compare the relative bias of the measured mock moments to their true underlying generator level moments with the uncertainty budget of the nominal analysis assigned ot the $B \to X_c \ell \nu_\ell$ modeling and background subtraction.

The assigned X_c modeling and background subtraction uncertainties in the nominal analysis should cover the observed residual bias. The residual bias is calculated according to the definition in Equation (5.32).

Four different scenarios modifying the $B \to X_c \ell \nu_\ell$ modeling are tested in this study: In scenario one and two, the contribution from $B \to D^{(*)} \eta \ell \nu_\ell$ decays and the whole non-resonant part of the spectrum are dropped, respectively. The other two scenarios either the $B \to D \ell \nu_\ell$ or the $B \to D^* \ell \nu_\ell$ branching fractions are doubled. Figure 5.29 shows the effect on the different modifications on $q^2_{\text{gen,sel}}$, the generator level q^2 distribution of the $B \to X_c \ell \nu_\ell$ signal MC after the selection. Doubling the $B \to D \ell \nu_\ell$ contribution results in a $q^2_{\text{gen,sel}}$ distribution with its mean value shifted to smaller q^2 values. The remaining modifications have the opposite effect of shifting q^2 to higher values.

Figure 5.30 shows the relative bias of the first q^2 moment extracted from the modified mock data sets compared using the nominal analysis procedure. The nominal X_c modeling and background subtraction uncertainty budget of the nominal measurement is also given as reference. The observed remaining bias between the unfolded and generator level moments is covered by the uncertainties assumed in the measurement. The respective comparisons for higher order moments are illustrated in Figures D.3 to D.5. Overall, these results indicate that the analysis is robust against fairly drastic changes in the q^2 shape and can recover moments with remaining biases smaller than the assigned uncertainties.

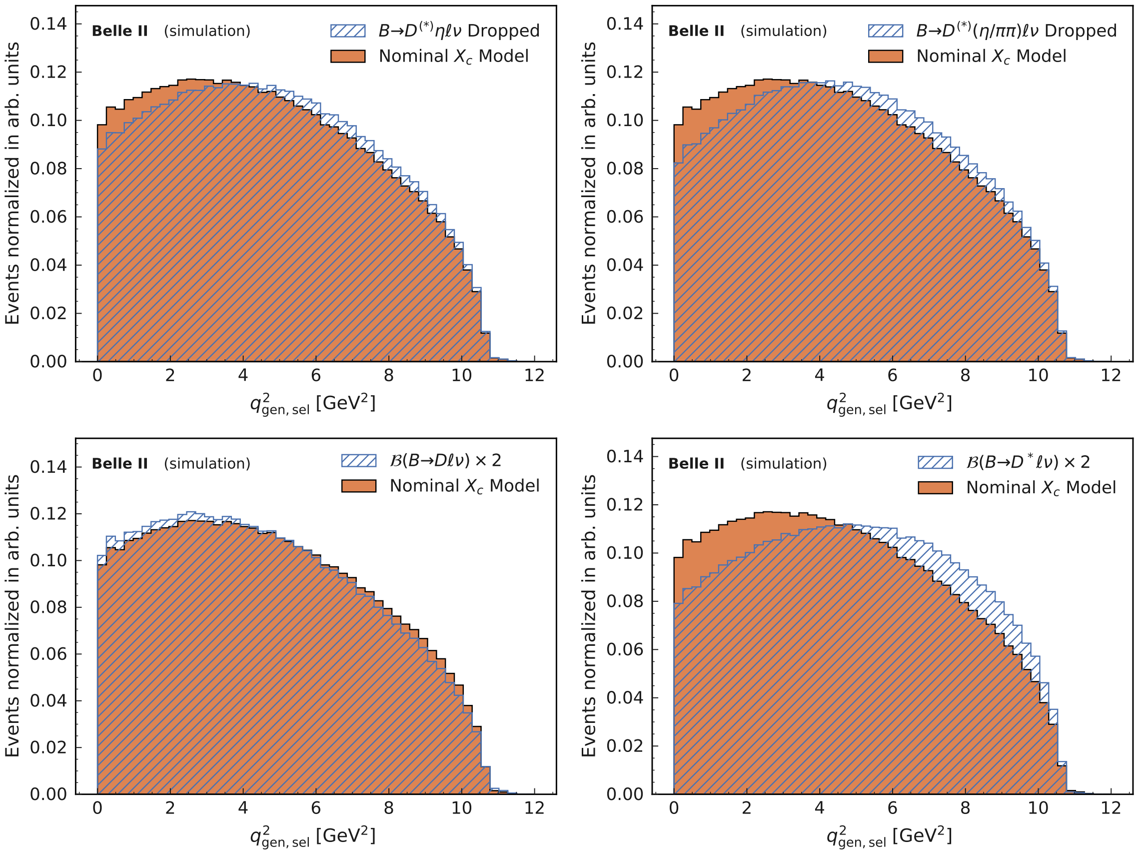


Figure 5.29: Comparison between the generator level q^2 distribution of the nominal $B \longrightarrow X_c \ell \nu_\ell$ model (blue) and artificially altered contributions of individual X_c components (orange). The distribution shows signal MC with all event selection criteria applied.

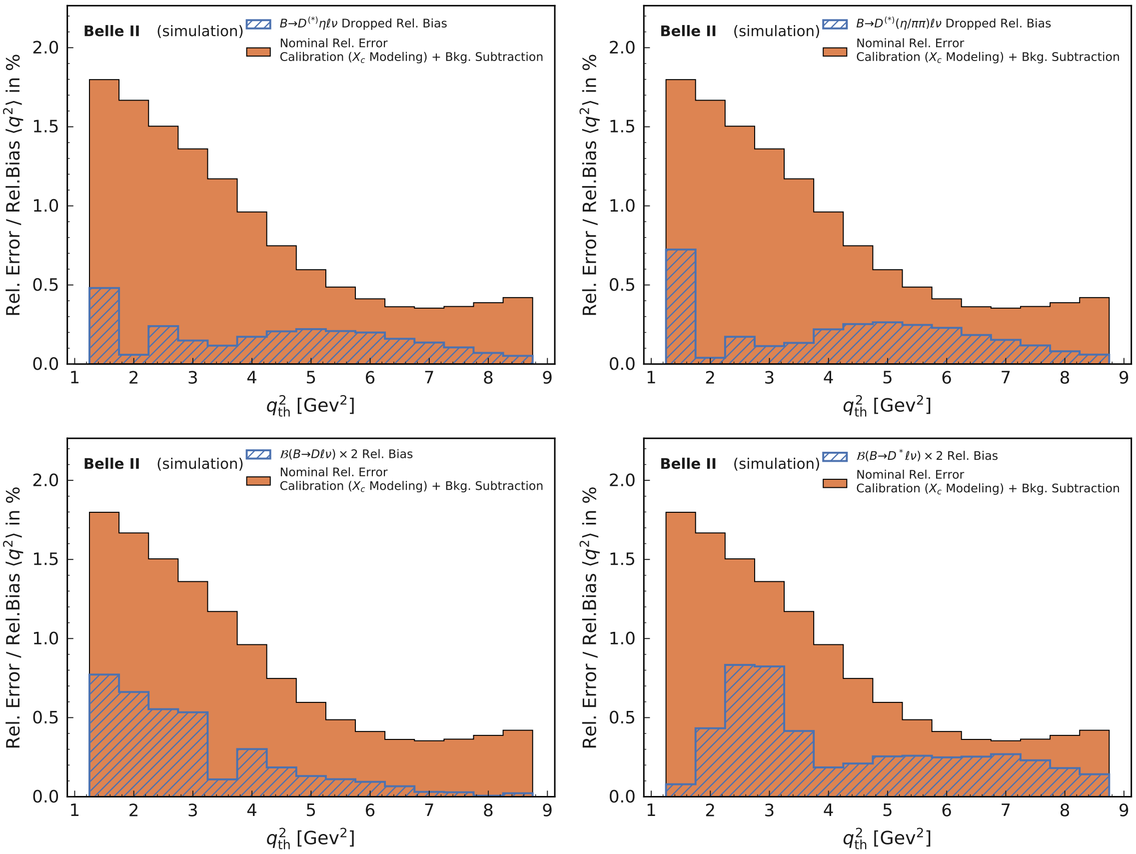


Figure 5.30: Relative bias of the extracted first q^2 moments from the different mock data set with a modified q^2 shape (blue) compared to the X_c modeling and background subtraction uncertainty budget of the nominal result (orange).

CHAPTER 6

First Extraction of Inclusive $|V_{cb}|$ from q^2 Moments

This chapter presents and summarizes the results from the first extraction of $|V_{cb}|$ from q^2 moments in $B \to X_c \ell \nu_\ell$ decays following [120]. These results are based on the work carried out in collaboration with F. Bernlochner, M. Fael, E. Persson, K. Olschewsky, K. K. Vos, and R. van Tonder. All theoretical calculations were performed by M. Fael, K. Olschewsky, and K. K. Vos. The theory predictions for the q^2 moments are provided in form of a software package [121]. The main contribution from this work to [120] is the development and implementation of the fit procedure as well as performing all studies and fits producing the results. The fit combines the measurement of the q^2 moments carried out in this work and recent results reported by the Belle Collaboration [31].

The state-of-the-art method to extract $|V_{cb}|$ from inclusive $B \to X_c \ell \nu_\ell$ decays is a global fit of the partial rate and spectral moments of the hadronic mass m_X and lepton energy E_ℓ. The spectral moments of order n for an observable O are usually measured as functions of a lower threshold on E_ℓ and are defined as

$$\langle O^n \rangle \equiv \int_{E_{\ell,\mathrm{cut}}}^{E_{\ell,\mathrm{max}}} \mathrm{d}O\, O^n \frac{\mathrm{d}\Gamma}{\mathrm{d}O} \bigg/ \int_{E_{\ell,\mathrm{cut}}}^{E_{\ell,\mathrm{max}}} \mathrm{d}O \frac{\mathrm{d}\Gamma}{\mathrm{d}O}. \tag{6.1}$$

Here, the partial rate, and thus also the moments, can be expressed using the Heavy Quark Expansion (HQE) as an expansion in inverse powers of the bottom quark mass m_b. At each order in the expansion, non-perturbative matrix elements, the HQE parameters, and coefficients, calculable in perturbation theory, are introduced.

The current most precise determination of $|V_{cb}|$ following this approach is [58]

$$|V_{cb}| = (42.16 \pm 0.30|_{\mathrm{th}} \pm 0.32|_{\mathrm{exp}} \pm 0.25|_{\Gamma}) \times 10^{-3} = (42.16 \pm 0.51) \times 10^{-3}. \tag{6.2}$$

This determination fits HQE parameters up to order $1/m_b^3$ in a model-independent way. Extracting contributions from higher orders is complicated by the proliferation of non-perturbative matrix elements. Up to order $1/m_b^3$, four non-perturbative matrix elements are presents, while the next order $1/m_b^4$ already introduces nine new parameters.

The main advantage of including only q^2 moments in the fit instead of m_X and E_ℓ moments is that the former are reparameterization invariant observables. As shown in [30, 39], this symmetry relates different HQE parameters at a certain order and thus effectively reduces the number of hadronic matrix elements from nine to five at order $1/m_b^4$.

The remainder of this chapter is structured as follows. Section 6.1 describes the novel method to determine $|V_{cb}|$ from q^2 moments with a reduced set of HQE parameters. Next, Section 6.2 introduces the fit procedure and summarizes the experimental input used in the fit. Lastly, the main findings of this study are presented in Section 6.3.

6.1 $|V_{cb}|$ Determination using q^2 Moments

The total semileptonic rate of $B \to X_c \ell \nu_\ell$ decays and moments of inclusive observables can be calculated in the HQE as a series expanded in inverse powers of the bottom quark mass m_b. In [30, 39], the authors show that the invariance under reparameterization (RPI) of the total rate and q^2 moments is also valid within the HQE. For a detailed discussion of the RPI and its application to inclusive decays, the reader is referred to [122–126]. As a result of the RPI, only eight non-perturbative parameters are present at tree level at order $1/m_b^4$ in the HQE, whose definitions are given in [39].

Until now, spectral moments have been usually measured with a lower threshold on the charged lepton energy E_ℓ. Introducing a lower E_ℓ threshold for q^2 moments would break their RPI. Therefore, the authors of [30] suggest to measure the q^2 moments as functions of lower cuts on q^2 itself, which preserves the RPI of the observable.[1]

The moments of the q^2 spectrum with a lower q^2 threshold are given by

$$Q_n(q^2_{\text{cut}}) \equiv \frac{1}{\Gamma_0} \int \mathrm{d}\Phi \, q^{2n} \, \theta(q^2 - q^2_{\text{cut}}) \, W^{\mu\nu} L_{\mu\nu} = \frac{1}{\Gamma_0} \int_{q^2_{\text{cut}}}^{m_b^2(1-\sqrt{\rho})^2} \mathrm{d}q^2 \, q^{2n} \, \frac{\mathrm{d}\Gamma}{\mathrm{d}q^2}, \tag{6.3}$$

with $\Gamma_0 = G_F^2 m_b^5 |V_{cb}|^2 A_{\text{ew}}/(192\pi^3)$ and $\rho = m_c^2/m_b^2$. The factor $A_{\text{ew}} = (1 + \frac{\alpha}{\pi}\log(m_Z/m_b))^2 \approx 1.01435$ stems from short-distance radiative corrections at the electroweak scale [42].

The normalized moments are defined as

$$\langle q^{2n} \rangle_{q^2 \geq q^2_{\text{cut}}} \equiv \int_{q^2_{\text{cut}}}^{m_b^2(1-\sqrt{\rho})^2} \mathrm{d}q^2 \, q^{2n} \, \frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} \Bigg/ \int_{q^2_{\text{cut}}}^{m_b^2(1-\sqrt{\rho})^2} \mathrm{d}q^2 \, \frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} = \frac{Q_n(q^2_{\text{cut}})}{Q_0(q^2_{\text{cut}})}. \tag{6.4}$$

The first raw q^2 moment is given by

$$q_1(q^2_{\text{cut}}) \equiv \langle q^2 \rangle_{q^2 \geq q^2_{\text{cut}}} \text{ for } n = 1, \tag{6.5}$$

while the central moments

$$q_n(q^2_{\text{cut}}) \equiv \langle (q^2 - \langle q^2 \rangle)^n \rangle_{q^2 \geq q^2_{\text{cut}}} \text{ for } n > 1 \tag{6.6}$$

can be obtained from the raw moments via the binomial expansion in Equation (5.43).

---

[1] In this chapter, the expression cut and threshold are used synonymously.

Table 6.1: Schematic overview of the perturbative corrections available for the partial rate Γ and the q^2 moments. ✓ denotes corrections that are known and built into the code [121]. ✗ indicates corrections that are known, but currently not included in the software package and the corrections indicated with ✗ are not yet available. For references and further information the reader is referred to [120].

Γ	tree	α_s	α_s^2	α_s^3
Partonic	✓	✓	✓	✓
μ_G^2	✓	✓	✗	✗
ρ_D^3	✓	✓	✗	✗
$1/m_b^4$	✓	✗	✗	✗
$m_b^{\text{kin}}/\overline{m}_c$	✗	✓	✓	✓

$\langle q^{2n} \rangle$	tree	α_s	α_s^2	α_s^3
Partonic	✓	✓	✗	✗
μ_G^2	✓	✗	✗	✗
ρ_D^3	✓	✗	✗	✗
$1/m_b^4$	✓	✗	✗	✗

With the HQE, the moments $Q_n(q^2_{\text{cut}})$ can be expressed in a double expansion in α_s and $1/m_b$ as

$$\begin{aligned} Q_n(q^2_{\text{cut}}) = (m_b^2)^n \Bigg\{ & \mu_3 \left[X_0^{(n)} + \left(\frac{\alpha_s}{\pi} \right) X_1^{(n)} + \dots \right] \\ & + \frac{\mu_G^2}{m_b^2} \left[g_0^{(n)} + \left(\frac{\alpha_s}{\pi} \right) g_1^{(n)} + \dots \right] + \frac{\rho_D^3}{m_b^3} \left[d_0^{(n)} + \left(\frac{\alpha_s}{\pi} \right) d_1^{(n)} + \dots \right] \\ & + \frac{r_E^4}{m_b^4} l_{r_E}^{(n)} + \frac{r_G^4}{m_b^4} l_{r_G}^{(n)} + \frac{s_B^4}{m_b^4} l_{s_B}^{(n)} + \frac{s_E^4}{m_b^4} l_{s_E}^{(n)} + \frac{s_{qB}^4}{m_b^4} l_{s_{qB}}^{(n)} \Bigg\}, \end{aligned} \tag{6.7}$$

with $\alpha_s \equiv \alpha_s^{(4)}(\mu_s)$ taken at the renormalization scale μ_s. At $\mathcal{O}(1/m_b^4)$, the moments $Q_n(q^2_{\text{cut}})$ are parameterized by five non-perturbative HQE parameter r_E^4, r_G^4, s_B^4, s_E^4, and s_{qB}^4. The expansion in Equation (6.7) does not explicitly depend on the parameter μ_π^2, as it only appears indirectly via the relation $\mu_3 = 1 + (\mu_G^2 - \mu_\pi^2)/(2m_b^2)$. Due to the normalization in Equation (6.4), μ_π^2 is effectively also suppressed with $1/m_b^4$ and the moments are rather insensitive to μ_π^2. Since μ_π^2 is an important input for the total rate, the parameter is constrained to an external measurement further discussed in Section 6.2.2.

The tree-level expression of the moments are computed and listed in [30]. A schematic overview of available and implemented perturbative corrections for the total rate and the q^2 moments is given in Table 6.1.

Extracting $|V_{cb}|$ with q^2 moments instead of m_X or E_ℓ moments has the advantage of a reduced set of HQE parameters at order $1/m_b^4$. The rate and the moments only depend on the following eight parameters:

- μ_G^2 and μ_π^2 at order $1/m_b^2$,
- ρ_D^3 at order $1/m_b^3$,
- r_E^4, r_G^4, s_E^4, s_B^4, and s_{qB}^4 at order $1/m_b^4$.

Figure 6.1 shows the predictions for the moments q_n as functions of lower q^2 thresholds. Here, the contributions of individual $1/m_b^4$ terms to the prediction are compared to the q_n prediction without

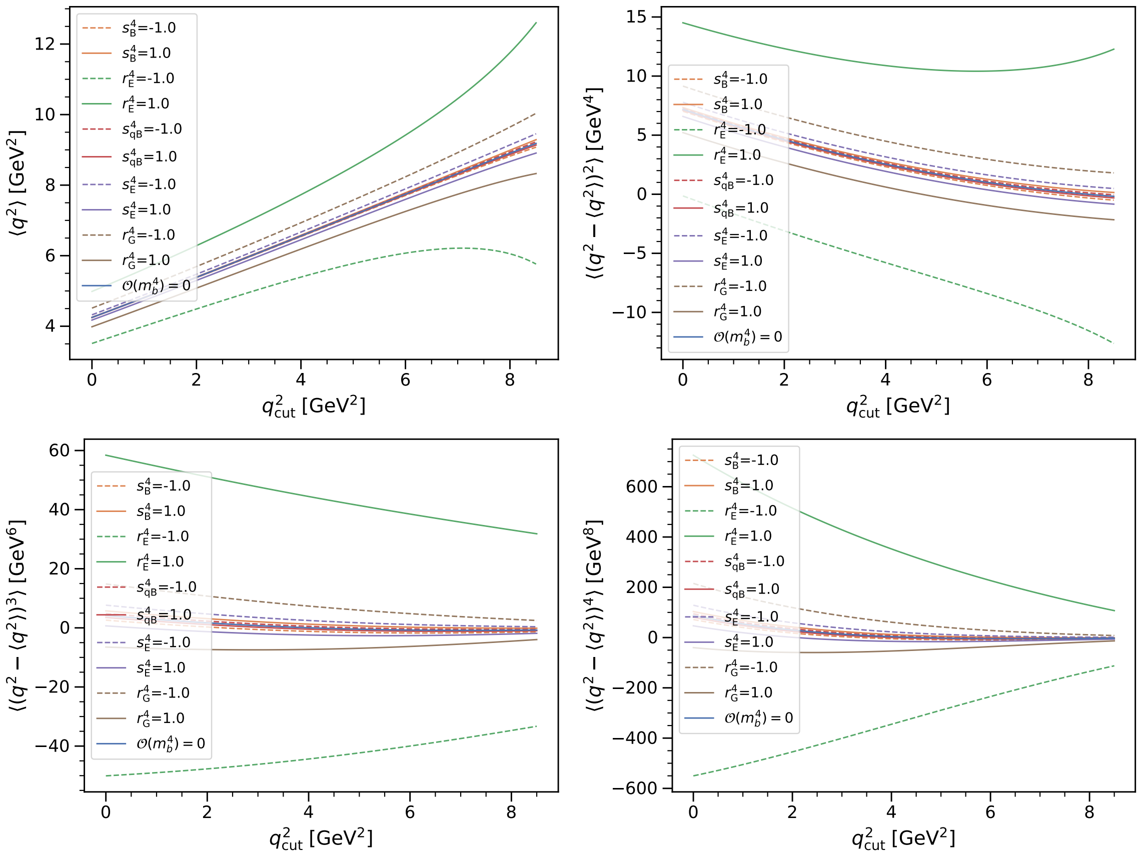


Figure 6.1: Sensitivity of the central q^2 moments for individual contributions of $1/m_b^4$ HQE terms obtained by varying their values between ±1 GeV^4. Variations of the $1/m_b^4$ HQE terms with positive and negative values are shown as solid and dashed lines, respectively.

any contribution of $1/m_b^4$ terms. For each variation, a $1/m_b^4$ parameter is varied between ±1 GeV^4 while simultaneously setting the other $1/m_b^4$ parameters to zero. One observes a large sensitivity to r_E^4 and r_G^4, while the other parameters only lead to small changes in the moments.

Figure 6.2 illustrates the impact of different choices for the $1/m_b^4$ parameters on a hypothetical $|V_{cb}|$ value from a measured $B \to X_c \ell \nu_\ell$ branching fractions for variations between ±1 GeV^4. The branching fraction is fixed to the average value given in Section 6.2.4. Again, the largest shift in $|V_{cb}|$ is observed from r_E^4 and r_G^4, while the other parameters only show a small effect on the total rate. Thus, only the parameters r_E^4 and r_G^4 are included into the default fit setup since a small sensitivity on the other $1/m_b^4$ parameters is expected. They only have a small impact on the prediction of the moments and due to their sub-leading contributions to the total rate also on $|V_{cb}|$. No experimental measurements of the partial rate as a function of lower q^2 cuts are provided by the experiments yet. Hence, the extraction of $|V_{cb}|$ will mainly rely on the input $B \to X_c \ell \nu_\ell$ branching fraction.

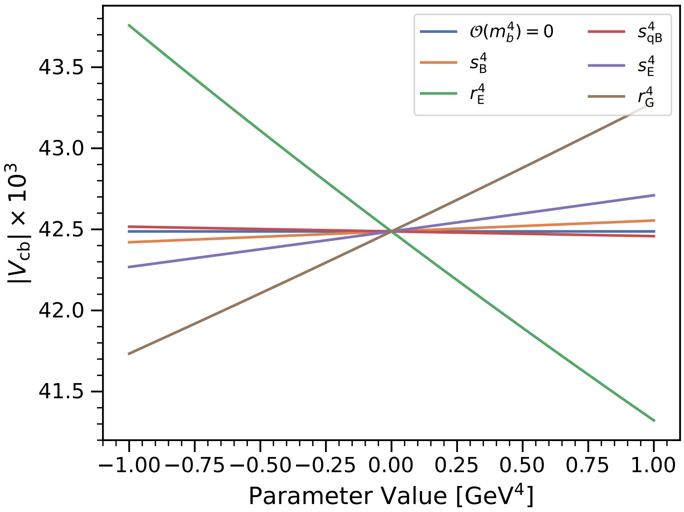


Figure 6.2: Illustration of the effect on $\left|V_{cb}\right|$ by varying the $1/m_b^4$ HQE terms independently between ± 1 GeV^4. The $B \to X_c \ell \nu_\ell$ branching fractions is fixed to the value given in Equation (6.25).

6.2 Fit Procedure

6.2.1 Fit Setup

$|V_{cb}|$ and the HQE parameters are determined in a simultaneous χ^2 fit to the central q^2 moments measured in this work and by the Belle collaboration. Both, theoretical and experimental uncertainties are considered in the fit. The χ^2 function is defined as

$$\begin{aligned} \chi^2(|V_{cb}|, \boldsymbol{\theta}) = & \frac{\left(\mathcal{B} - \Gamma(|V_{cb}|, \boldsymbol{\theta})\tau_B\right)^2}{\sigma_{\mathcal{B}}^2 + \sigma_\Gamma^2} \\ & + \left(\boldsymbol{q}(\boldsymbol{\theta}) - \boldsymbol{q}_{\text{meas}}\right)^\intercal C^{-1} \left(\boldsymbol{q}(\boldsymbol{\theta}) - \boldsymbol{q}_{\text{meas}}\right) \\ & + \sum_{i=1}^{4} \frac{\left(\theta_i - \tilde{\theta}_i\right)^2}{\sigma_{\theta_i}^2}, \end{aligned} \tag{6.8}$$

with $\mathcal{B}$ and Γ being the experimental branching fraction for $B \to X_c \ell \nu_\ell$ decays and the theoretical expression for the decay rate, respectively. The vectors $\boldsymbol{q}(\boldsymbol{\theta})$ and $\boldsymbol{q}_{\text{meas}}$ denote the theory predictions and measurements for the moments q_n for different lower q^2 thresholds, respectively. Here, $\boldsymbol{\theta}$ are the fitted HQE parameters. The average lifetime $\tau_B = (1.579 \pm 0.005)$ ps [23] of charged and neutral B mesons is used to relate the calculated decay rate to the measured branching fraction. The HQE parameters $\theta_i = m_b^{\text{kin}}, \overline{m}_c, \mu_G^2, \mu_\pi^2$ are constrained parameters in the fit, further discussed in Section 6.2.2.

The total covariance matrix is the sum of the theory and experimental covariance matrices $C = C_{\text{theo}} + C_{\text{exp}}$. The experimental covariance matrix C_{exp} describes the statistical and systematic uncertainties of the Belle and Belle II measurements. Some systematic uncertainties are considered to be correlated between both measurements, which is further discussed in Section 6.2.4. The theory covariance matrix C_{theo} is used to incorporate the uncertainty from missing higher-order corrections in α_s and the HQE. Section 6.2.3 discusses the approach used to quantify the theory uncertainty.

A priori, the correlation between the theory predictions of q^2 moments calculated for different q^2 thresholds and orders are unknown. This ignorance about the theory correlation structure imposes an additional uncertainty in the extraction of $|V_{cb}|$. In this work, two additional parameters ρ_{cut} and ρ_{mom} are introduced to parameterize the correlation between moments calculated for different lower q^2 thresholds and orders, respectively. The branching fraction is considered to be uncorrelated to the q^2 moments.

The correlation between two moments q_n of the same order n but calculated for different lower q^2 thresholds q^2_A and q^2_B is parameterized as

$$\rho_n\left(q_n(q^2_A), q_n(q^2_B)\right) = \rho^x_{\text{cut}} \quad \text{with} \quad x = \frac{\left|q^2_A - q^2_B\right|}{0.5\,\text{GeV}^2}. \tag{6.9}$$

This results in a lower correlation between two moments calculated for q^2 thresholds with larger differences $\left|q^2_A - q^2_B\right|$. A comparable functional approach was used in [14] to describe the correlation between moments with different lepton energy thresholds. However, for the fit performed in [109], the theory covariance matrix was constructed assuming fully correlated uncertainties.

In contrast to [14, 109], moments of different orders are not treated as uncorrelated. The correlation of different orders n and m, the correlation coefficient is given by

$$\rho_{nm}\left(q_m(q^2_A), q_n(q^2_B)\right) = \text{sign}(\rho_{\text{mom}}) \cdot \left|\rho_{\text{mom}}\right|^{|m-n|} \cdot \rho_n(q_n(q^2_A), q_n(q^2_B)). \tag{6.10}$$

In the default fit scenario, the correlation parameters ρ_{cut} and ρ_{mom} are not fixed at certain values, but are treated as nuisance parameters in the fit and are allowed to float. With this method, many possible correlations scenarios are considered directly in the fit uncertainties of the HQE parameters and $|V_{cb}|$.

The nuisance parameters ρ_{cut} and ρ_{mom} are constrained to a sensible range using the double Fermi-Dirac function

$$f_{\text{DFD}}(\rho, a, b) = \frac{1}{2(1+e^{w(\rho-b)})(1+e^{-w(\rho-a)})} \quad (b > a), \tag{6.11}$$

with $w = 50$. Here, the arguments a and b denote the minimal and maximal values for allowed range of the correlation parameters ρ.

Only positive correlation values are allowed for ρ_{cut} by using $a = 0$ and $b = 1$. For ρ_{mom}, the minimal value is set to $a = -0.45$ allowing also negative correlations. A lower boundary of $a = -1$ could not be used since the constructed covariance matrix becomes non-singular for $a < -0.45$. Two constraint terms $-2\log f_{\text{DFD}}$ for ρ_{cut} and ρ_{mom} are added to Equation (6.8)

$$\chi^2 \rightarrow \chi^2 + \chi^2_{\text{DFD}} = \chi^2 - 2\log f_{\text{DFD}}(\rho_{\text{cut}}, 0, 1) - 2\log f_{\text{DFD}}(\rho_{\text{mom}}, -0.45, 1). \tag{6.12}$$

Figure 6.3 shows examples of the functional dependence of f_{DFD} and $-2\log f_{\text{DFD}}$ with $a = -1$ and $b = 1$. For values outside the allowed range of $a < \rho < b$, the constraint term adds a large positive contribution to the χ^2 function.

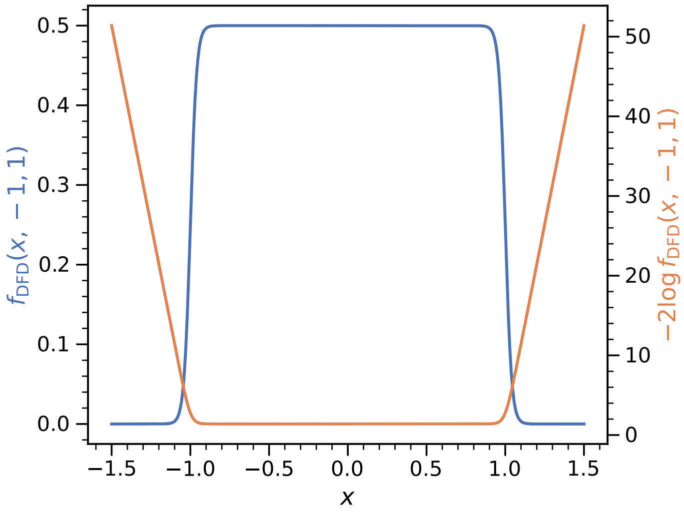


Figure 6.3: Examples of a double Fermi-Dirac function f_{DFD} and $-2\log f_{\mathrm{DFD}}$ with $a = -1$ and $b = 1$.

6.2.2 External Constraints to the Fit

The HQE parameter m_b, m_c, μ_G^2, and μ_π^2 are constrained to external measurements in the fit with Gaussian priors.

The bottom quark mass, defined in the kinetic mass scheme, and is constrained to

$$m_b^{\mathrm{kin}}(1\,\mathrm{GeV}) = (4.565 \pm 0.020)\,\mathrm{GeV}. \tag{6.13}$$

The value for m_b^{kin} is calculated from the bottom and charm quark masses defined in the $\overline{\mathrm{MS}}$ scheme. The calculation follows [60, 61] using $\overline{m}_b(\overline{m}_b) = (4.198 \pm 0.012)\,\mathrm{GeV}$ [23] and $\overline{m}_c(3\,\mathrm{GeV}) = (0.988 \pm 0.007)\,\mathrm{GeV}$ [44].

The charm quark mass has been calculated in lattice QCD [44] and QCD sum rules [127, 128]. In the fit, the charm quark mass is constrained to

$$\overline{m}_c(2\,\mathrm{GeV}) = (1.093 \pm 0.008)\,\mathrm{GeV}. \tag{6.14}$$

This value is calculated using `RunDec` [129, 130] at 4-loop accuracy.

A constraint for μ_G^2 can be obtained from the measured mass difference of the B^* and B mesons. This difference is parameterized by $\mu_G^2(\mu_s)$ as

$$m_{B^*}^2 - m_B^2 = \frac{4}{3} C_{\mathrm{mag}}(\mu_S)\mu_G^2(\mu_s) + O(\alpha_s\mu_G^2, 1/m_b^3), \tag{6.15}$$

with $C_{\mathrm{mag}} = 1$. The definition $\mu_G^2 \equiv \mu_G^2(\mu_s = m_b^{\mathrm{kin}})$ is used throughout this study. Following [131], higher order contributions to the mass difference are ignored and one obtains

$$\mu_G^2 = (0.36 \pm 0.07)\,\mathrm{GeV}^2. \tag{6.16}$$

As stated previously, the theory prediction for the q^2 moments is insensitive to μ_π^2. The parameter is constrained to

$$\mu_\pi^2 = (0.43 \pm 0.24)\,\mathrm{GeV}^2. \tag{6.17}$$

This conservative constraint is obtained from the computation in [57] and inflating the uncertainty by a factor of four.

6.2.3 Theory Uncertainties

The size of the theory uncertainty for the prediction of the q^2 moments is obtained by varying ρ_D^3, μ_G^2 and the scale of α_s. The predictions for the varied moments $q_n^{\mathrm{var}}(q_{\mathrm{th}}^2)$ are compared to a set of nominal predictions $q_n^{\mathrm{nom}}(q_{\mathrm{th}}^2)$. The uncertainty amplitude is estimated as

$$\sigma_{q_n(q_{\mathrm{th}}^2)} = \left| q_n^{\mathrm{var}}(q_{\mathrm{th}}^2) - q_n^{\mathrm{nom}}(q_{\mathrm{th}}^2) \right| \tag{6.18}$$

for each variation. The total amplitude is obtained by summing the individual uncertainties in quadrature.

The nominal predictions use the m_b^{kin} and $\overline{m}_c$ values from Equations (6.13) and (6.14), respectively. In addition, the HQE parameters μ_G^2 and μ_π^2 are set to their respective constraint value (see Equations (6.16) and (6.17)). The strong coupling constant is set to

$$\alpha_s(\mu_s = m_b^{\mathrm{kin}}) = 0.2184. \tag{6.19}$$

This value is determined with `RunDec` with $n_f = 4$ active quark flavors and 5-loop accuracy and the initial value of $\alpha_s(m_Z) = 0.1179 \pm 0.0009$ [23].

Three different variations are considered to quantify the theory uncertainty on the q^2 moments following the methods used in [14, 15]:

1. Missing higher-order corrections in α_s. Here, the scale of $\alpha_s(\mu_s)$ is evaluated at $\mu_s = m_b^{\mathrm{kin}}/2$.
2. Missing higher-order $1/m_b$ corrections. Here, the HQE parameter ρ_D^3 is varied up by 30%.
3. Missing $\alpha_s/m_b^{2,3}$ corrections. Here, the HQE parameter μ_G^2 is varied up by 20%.

For the variation of ρ_D^3, the value $\rho_D^3 = 0.127\,\mathrm{GeV}^3$ from Appendix A in [30] is used. The uncertainty obtained by these variations result in a conservative estimate. In the future, this uncertainty may be reduced by including higher-order corrections in α_s and α_s/m_b to the predictions. Figure 6.4 shows the individual and total relative uncertainty amplitudes σ_{q_n}/q_n^2 for the q^2 moments as functions of the lower q^2 thresholds. The estimated uncertainties from the ρ_D^3 and μ_G^2 variations are larger than the variation of the scale μ_s. At higher q^2 thresholds, above $\approx 6\,\mathrm{GeV}^2$, the uncertainties for the central moments increase rapidly.

In the fit, the theory covariance matrix is constructed as

$$C_{\mathrm{theo}} = \mathrm{diag}(\boldsymbol{\sigma}_{\mathrm{tot}}) R(\rho_{\mathrm{cut}}, \rho_{\mathrm{mom}}) \mathrm{diag}(\boldsymbol{\sigma}_{\mathrm{tot}}), \tag{6.20}$$

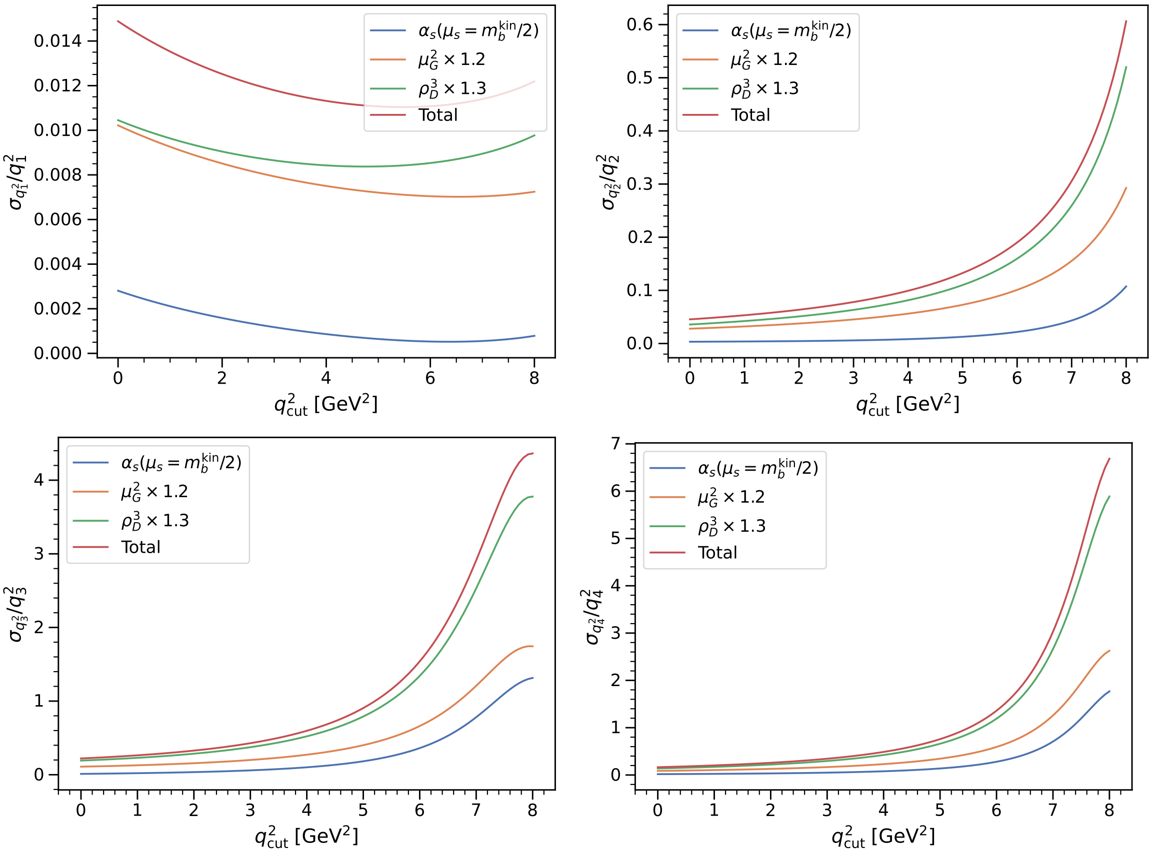


Figure 6.4: Relative uncertainty amplitudes $\sigma_{q_n(q^2_{\rm th})}/q_n^2$ of the q^2 moments as functions of the lower q^2 threshold. The total uncertainty is obtained by summing the individual contributions from the μ_s, μ_G^2, and ρ_D^3 variations in quadrature.

with $\sigma_{\rm tot}$ denoting the total uncertainty amplitude for the branching fraction and the q^2 moments. The correlation matrix R is constructed according to the values of the correlation parameters $\rho_{\rm cut}$ and $\rho_{\rm mom}$ and the prescriptions in Equations (6.9) and (6.10).

6.2.4 Experimental Input

The branching fraction $\mathcal{B}(B \to X_c \ell \nu_\ell)$ is an important input to the fit and a precise knowledge of it is essential in this determination of $|V_{cb}|$. Available measurements either quote partial branching fractions as a function of a lower lepton energy threshold or the full $B \to X \ell \nu_\ell$ branching ratio. Usually, $\mathcal{B}(B \to X \ell \nu_\ell)$ is measured with an implicit cut on the lepton energy and uses an extrapolation back to the full phase space.

Since the lepton energy is not an RPI invariant observable, the measurements of the partial branching fractions cannot be used directly in the fit like in recent $|V_{cb}|$ determinations [14, 58]. No measurements of partial branching fractions as functions of lower q^2 thresholds exist and so the full $B \to X_c \ell \nu_\ell$ branching fraction has to be used in the fit. Only a small number of measurements provide the full branching fraction. Thus, the measured partial branching fraction $\mathcal{B}_{\rm cut}$ results are extrapolated back to

Table 6.2: Available measurements of the inclusive $B \to X\ell\bar{\nu}_\ell$ and $B \to X_c\ell\bar{\nu}_\ell$ branching fractions, extrapolated to the full region using the correction factors in Equation (6.22).

	$\mathcal{B}(B \to X\ell\nu_\ell)$ [%]	$\mathcal{B}(B \to X_c\ell\nu_\ell)$ [%]	In Average
Belle [21] $E_\ell > 0.6\,\mathrm{GeV}$	-	10.54 ± 0.31	✓
Belle [21] $E_\ell > 0.4\,\mathrm{GeV}$	-	10.58 ± 0.32	
CLEO [133] incl.	10.91 ± 0.26	10.72 ± 0.26	
CLEO [133] $E_\ell > 0.6\,\mathrm{GeV}$	10.69 ± 0.25	10.50 ± 0.25	✓
BaBar [134] incl.	10.34 ± 0.26	10.15 ± 0.26	✓
BaBar [20] $E_\ell > 0.6\,\mathrm{GeV}$	-	10.68 ± 0.24	✓
Our Average	-	10.48 ± 0.13	
Average Belle [21] & BaBar [20] ($E_\ell > 0.6\,\mathrm{GeV}$)	-	10.63 ± 0.19	

the full phase space with a correction factor $\Delta(E_{\ell,\mathrm{cut}})$ depending on the lower lepton energy threshold

$$\mathcal{B}(B \to X_c\ell\nu_\ell) = \Delta(E_{\ell,\mathrm{cut}})\mathcal{B}_{\mathrm{cut}}(B \to X_c\ell\nu_\ell). \tag{6.21}$$

For this study, the factor $\Delta(E_{\ell,\mathrm{cut}})$ is determined in the local OPE using the partial rate at leading order in the HQE and including perturbative corrections up to α_s^2. For the lower thresholds $E_{\ell,\mathrm{cut}} = 0.4\,\mathrm{GeV}$ and $E_{\ell,\mathrm{cut}} = 0.6\,\mathrm{GeV}$, the corrections factors are calculated to be

$$\Delta(0.4\,\mathrm{GeV}) = 1.014 \pm 0.001 \quad \text{and} \tag{6.22}$$

$$\Delta(0.6\,\mathrm{GeV}) = 1.047 \pm 0.004, \tag{6.23}$$

respectively.

Measurements of the full $B \to X\ell\nu_\ell$ branching fractions have to be corrected for the $|V_{ub}|^2$ suppressed contribution of $B \to X_u\ell\nu_\ell$ decays as

$$\mathcal{B}(B \to X_c\ell\nu_\ell) = \mathcal{B}(B \to X\ell\nu_\ell) - \mathcal{B}_{\mathrm{cut}}(B \to X_u\ell\nu_\ell)/\epsilon_{\mathcal{B}_{\mathrm{cut}}}. \tag{6.24}$$

Here, the measurement $\mathcal{B}_{\mathrm{cut}}(B \to X_u\ell\nu_\ell) = (0.159 \pm 0.017)\,\%$ for $E_\ell > 1.0\,\mathrm{GeV}$ from [132] is used. The correction factor $\epsilon_{\mathcal{B}_{\mathrm{cut}}} = 0.858 \pm 0.008$ is obtained from private communications with the authors of [132]. Thus, the total $B \to X_u\ell\nu_\ell$ branching fraction is calculated to be $\mathcal{B}(B \to X_u\ell\nu_\ell) = (0.185 \pm 0.020)\,\%$.

Table 6.2 lists the available measurements for $B \to X\ell\nu_\ell$ and $B \to X_c\ell\nu_\ell$ branching fractions. Here, the extrapolations to the full phase space are given. Averaging the listed measurements, one obtains

$$\mathcal{B}(B \to X_c\ell\nu_\ell) = (10.48 \pm 0.13)\,\%, \tag{6.25}$$

which is used as the default branching fraction in the determination of $|V_{cb}|$ presented in this work. The χ^2 of the average with respect to the included measurements is 2.2, corresponding to a p-value of 52%. The measurement of [135] is not included in the average, as the analysis does not quote a partial

Table 6.3: Analyzed measured q^2 moments from Belle and Belle II.

	q^2_{cut} [GeV^2]
Belle [31]	$3, 4, 5, 6, 7, 8$
Belle II (this work)	$1.5, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5$

branching fraction corrected for FSR.

The averaged branching ratio in Equation (6.25) will dominate the central value of $|V_{cb}|$. In [58], the fitted value for the branching fraction is $(10.66 \pm 0.15)\%$. The difference of roughly 1.4σ to the average in Equation (6.25) is mainly caused by including the measurement from [134]. The authors of [58] only include the measurements from Belle [21] and BaBar [20] into their study since they directly use the partial branching fractions in their fit. The average of only the extrapolated branching fractions [21] and [20] results in

$$\mathcal{B}(B \to X_c \ell \nu_\ell) = (10.63 \pm 0.19)\,\%, \tag{6.26}$$

which is compatible with the result from [58].

In this study, a simultaneous fit to the q^2 moment measurements by Belle [31] and Belle II (this work) is used. The fit includes measurements of the first raw and second to fourth central q^2 moments and the average branching fraction given in Equation (6.25). An alternative determination of $|V_{cb}|$ is also performed using the branching fraction used reported in Equation (6.26). Belle presents measurements separated by lepton flavor with a minimal lower q^2 threshold of $q^2 > 3.0\,\mathrm{GeV}^2$ up to $q^2 > 10.5\,\mathrm{GeV}^2$. For this study, the average of the individual Belle electron and muon measurements is calculated. In this average, identical systematic uncertainty sources are assumed to be fully correlated. This work provides q^2 moments measured for lower q^2 thresholds starting at $1.5\,\mathrm{GeV}^2$ up to $8.5\,\mathrm{GeV}^2$. Here, the measurements are not separated between lepton flavor.

Since the measured moments are highly correlated, only a subset of moments of each experiment is used in the fit. The default sets are listed in Table 6.3. Fits with alternative sets of moments do not result in significant variations in the results. The maximal considered value for the lower q^2 threshold is $8\,\mathrm{GeV}^2$, since this avoids the kinematically allowed endpoint in the q^2 spectrum mainly populated by $B \to D^{(*)} \ell \nu_\ell$ decays where the .

Between both experiments, systematic uncertainties related to the branching fraction variations of $B \to D^{(*,**)} \ell \nu_\ell$ decays, the composition of the non-resonant part of the inclusive X_c spectrum, and the BGL form factor parameter uncertainties of $B \to D^{(*)} \ell \nu_\ell$ decays are considered to be fully correlated.

6.3 Results

6.3.1 First $|V_{cb}|$ Determination from q^2 Moments

In the default fit scenario, the theory prediction for the total rate and the q_n for $n = 1$–4 is fitted to the subset of q^2 moment measurements with lower q^2 threshold listed Table 6.3 and the average branching fraction from Equation (6.25). The parameters $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 are free parameters in the fit, while m_b^{kin}, $\overline{m}_c$, μ_G^2, and μ_π^2 are included with Gaussian constraints with the values discussed in

Table 6.4: Results of our default fit using both Belle and Belle II data for $|V_{cb}|$, $m_b^{\text{kin}}(1\,\text{GeV})$, $\overline{m}_c(2\,\text{GeV})$, the HQE parameters, and the correlation parameters ρ_{cut} and ρ_{mom}. All parameters are expressed in GeV at the appropriate power.

	$\lvert V_{cb}\rvert \times 10^3$	m_b^{kin}	$\overline{m}_c$	μ_G^2	μ_π^2	ρ_D^3	r_G^4	$r_E^4 \times 10$	ρ_{cut}	ρ_{mom}
Value	41.69	4.56	1.09	0.37	0.43	0.12	-0.21	0.02	0.05	0.09
Uncertainty	0.59	0.02	0.01	0.07	0.24	0.20	0.69	0.34	$^{+0.03}_{-0.01}$	$^{+0.10}_{-0.10}$

Section 6.2.2. Since prediction of the q^2 moments has limited sensitivity on the remaining $\mathcal{O}(1/m_b^4)$ HQE parameters s_E^4, s_B^4, and s_{qB}^4, they are set to zero and kept fixed during the fit. Following the discussion in Section 6.2, two additional nuisance parameters ρ_{cut} and ρ_{mom}, describing the correlation structure of the theory covariance matrix, are introduced to the fit.

The fit result is given in Table 6.4. Here, the uncertainties for $|V_{cb}|$ and the HQE parameters are estimated from the Hessian matrix. Figure 6.5 shows the post-fit projections of the theory predictions together with the analyzed Belle and Belle II measurements. The minimum of the χ^2 fit is $\chi^2_{\text{min}} = 7.17$ with 49 degrees-of-freedom (dof), indicating an excellent fit. Similar to reports in [14, 15, 58], a fit without the inclusion a theory covariance matrix results in a poor fit with $\chi^2/\text{dof} = 5.02$.

One-dimensional $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ scans are carried out to validate the Hessian uncertainties which are shown in Figure 6.6. For the parameters $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4, the Gaussian approximation of the χ^2 function at the minimum is valid. The uncertainties of the correlation parameter ρ_{cut} and ρ_{mom} showing a non-parabolic behavior, are estimated with pseudo-experiments which are further discussed in Section 6.3.4.

The corresponding correlation matrix is shown in Figure 6.7. $|V_{cb}|$ has a small correlation to $\overline{m}_c$ and the HQE parameters. A larger negative correlation between $|V_{cb}|$ and m_b^{kin} is observed with -0.59, due to the m_b^{kin} dependence of the total rate.

The default fit yields

$$\begin{aligned} |V_{cb}| &= (41.69 \pm 0.27|_{\mathcal{B}} \pm 0.31|_{\Gamma} \pm 0.18|_{\text{Exp.}} \pm 0.17|_{\text{Theo.}} \pm 0.34|_{\text{Constr.}}) \times 10^{-3} \\ &= (41.69 \pm 0.59) \times 10^{-3}, \end{aligned} \tag{6.27}$$

where the uncertainties stem from the experimental branching fraction $\mathcal{B}$, the theoretical uncertainty on the total rate Γ, the experimental and theoretical uncertainties on the q^2 moments and the uncertainty from external constraints.

For a different input $B \to X_c \ell \nu_\ell$ branching fraction, the corresponding $|V_{cb}|$ can be obtained by rescaling

$$|V_{cb}| = \sqrt{\frac{\mathcal{B}(B \to X_c \ell \nu_\ell)}{(10.48 \pm 0.13)\%}} \times (41.69 \pm 0.59) \times 10^{-3}. \tag{6.28}$$

With the value reported in Equation (6.26), $|V_{cb}|$ consequently increases to

$$|V_{cb}| = (41.99 \pm 0.65) \times 10^{-3}. \tag{6.29}$$

This value is also obtained by repeating the default fit using the branching fraction given in

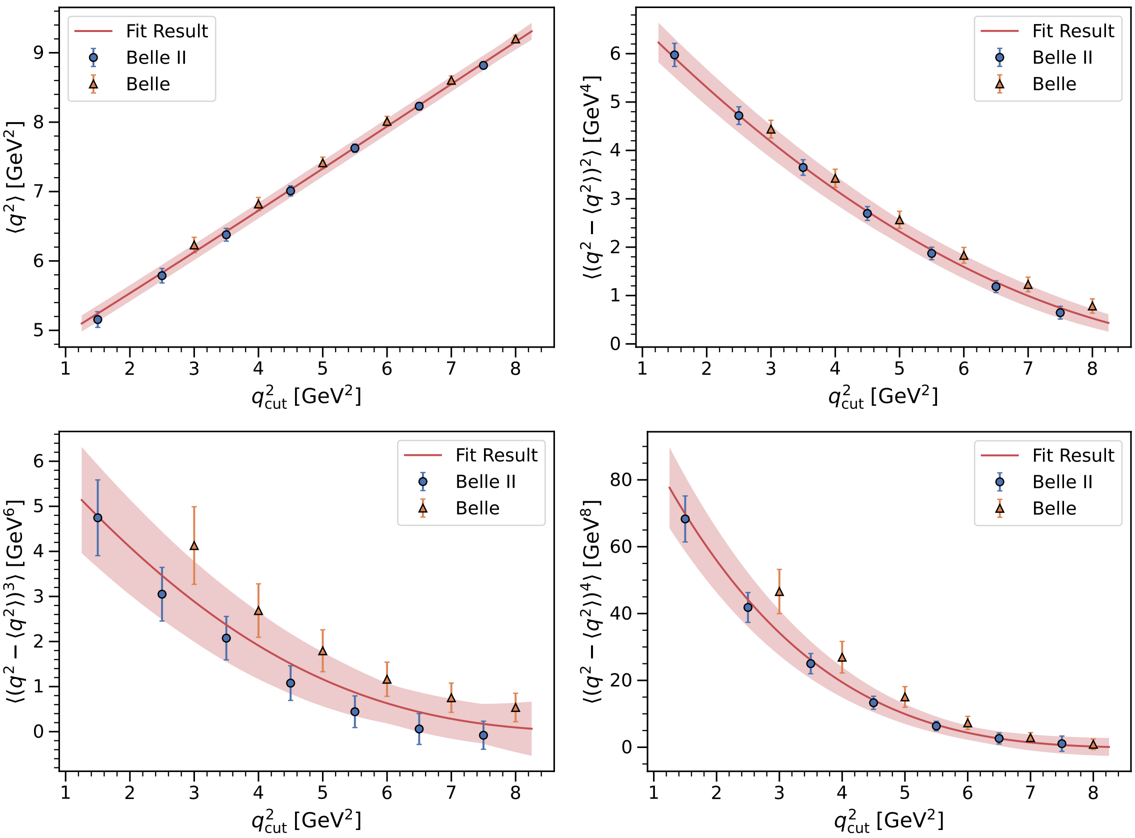


Figure 6.5: Fit projections for the central q^2 moments as a function of the q^2 threshold, combined with the measurement moments from both Belle and Belle II.

Equation (6.26) as an input, while the HQE parameters remain unchanged. The results of this fit are summarized in Table 6.5. The scaled $|V_{cb}|$ value from Equation (6.29) is in excellent agreement with the results reported in [58], especially comparing it to their result including estimates for power corrections:

$$|V_{cb}| = (42.00 \pm 0.53) \times 10^{-3}. \tag{6.30}$$

Fits using measurements from Belle and Belle II individually are summarized in Appendix E, respectively. Figure 6.8 compares the obtained $|V_{cb}|$ and ρ_D^3 values and also shows the corresponding

Table 6.5: Results of our default fit with the $B \to X_c \ell \nu_\ell$ branching fraction value reported in Equation (6.26), using both Belle and Belle II data. All parameters are expressed in GeV at the appropriate power.

	$\lvert V_{cb}\rvert \times 10^3$	m_b^{kin}	$\overline{m}_c$	μ_G^2	μ_π^2	ρ_D^3	r_G^4	$r_E^4 \times 10$	ρ_{cut}	ρ_{mom}
Value	41.99	4.56	1.09	0.37	0.43	0.12	-0.21	0.02	0.05	0.09
Uncertainty	0.65	0.02	0.01	0.07	0.24	0.20	0.69	0.34	0.13	0.99

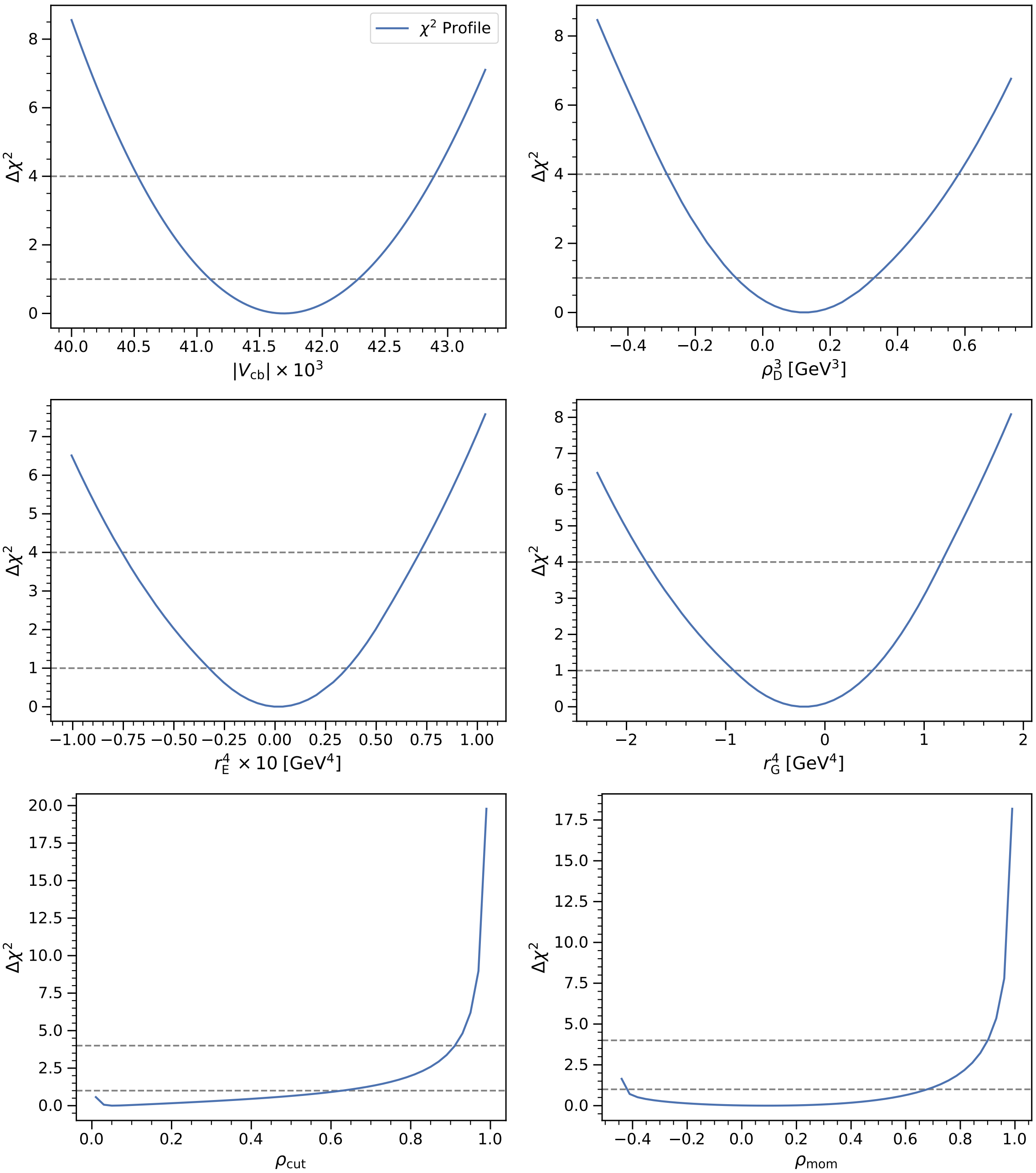
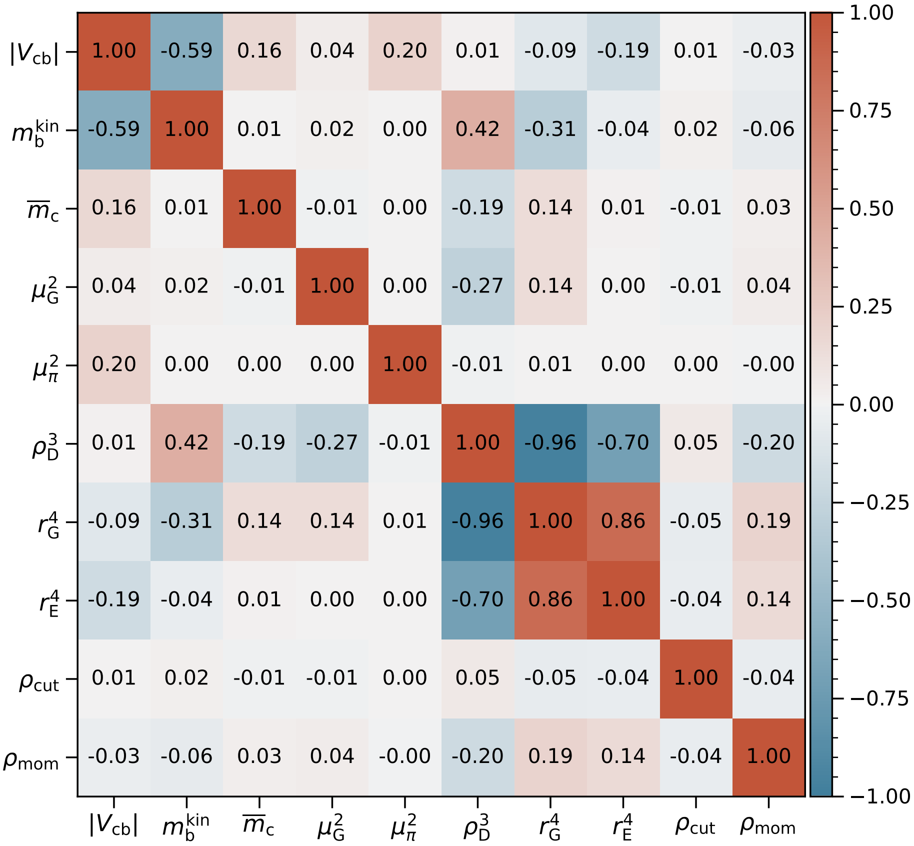


Figure 6.6: One-dimensional χ^2 profile scans for $|V_{cb}| \times 10^3$, ρ_D^3, $r_E^4 \times 10$, and r_G^4. The minimum χ^2_{min} is subtracted from the χ^2 function.



	$\lvert V_{cb}\rvert$	m_b^{kin}	$\overline{m}_c$	μ_G^2	μ_π^2	ρ_D^3	r_G^4	r_E^4	ρ_{cut}	ρ_{mom}
$\lvert V_{cb}\rvert$	1.00	-0.59	0.16	0.04	0.20	0.01	-0.09	-0.19	0.01	-0.03
m_b^{kin}	-0.59	1.00	0.01	0.02	0.00	0.42	-0.31	-0.04	0.02	-0.06
$\overline{m}_c$	0.16	0.01	1.00	-0.01	0.00	-0.19	0.14	0.01	-0.01	0.03
μ_G^2	0.04	0.02	-0.01	1.00	0.00	-0.27	0.14	0.00	-0.01	0.04
μ_π^2	0.20	0.00	0.00	0.00	1.00	-0.01	0.01	0.00	0.00	-0.00
ρ_D^3	0.01	0.42	-0.19	-0.27	-0.01	1.00	-0.96	-0.70	0.05	-0.20
r_G^4	-0.09	-0.31	0.14	0.14	0.01	-0.96	1.00	0.86	-0.05	0.19
r_E^4	-0.19	-0.04	0.01	0.00	0.00	-0.70	0.86	1.00	-0.04	0.14
ρ_{cut}	0.01	0.02	-0.01	-0.01	0.00	0.05	-0.05	-0.04	1.00	-0.04
ρ_{mom}	-0.03	-0.06	0.03	0.04	-0.00	-0.20	0.19	0.14	-0.04	1.00

Figure 6.7: Correlation matrix for $|V_{cb}|$, the HQE parameters, and the correlation parameters ρ_{cut} and ρ_{mom}.

two-dimensional $\Delta\chi^2 = 2.30$ contours indicating good agreement.

6.3.2 Theory Correlations and HQE Parameters versus $|V_{cb}|$

The impact of the correlation parameters ρ_{cut} and ρ_{mom} on $|V_{cb}|$ is of particular interest. To study this, two-dimensional scans of $\Delta\chi^2$ around the minimum are performed. Figure 6.9 shows the $\Delta\chi^2 = 2.30, 4.61$, and 5.99 contours in the $|V_{cb}|$ versus ρ_{cut} and ρ_{mom} plane. In a valid Gaussian approximation, the contour levels would correspond to the two-dimensional 68%, 90% and 95% confidence regions. One observes a small impact of both ρ_{cut} and ρ_{mom} on $|V_{cb}|$ when profiling over a large range for the two parameters. Thus, the uncertainty on $|V_{cb}|$ includes a large range of possible correlation coefficients. In conclusion, the fitted value of $|V_{cb}|$ is stable with respect to both nuisance parameters. As these parameters are a priori unknown, this is an important finding.

Two-dimensional $\Delta\chi^2$ scans are also performed for $|V_{cb}|$ versus ρ_D^3, r_E^4, and r_G^4. Here, no sizable correlations are observed.

Scans over the ρ_{cut} and ρ_{mom} space with fixed correlation parameters are shown in Appendix E. The fit results are given in Figures E.10 to E.13 for the combined fit to Belle and Belle II as well as for individual fits to both measurements.

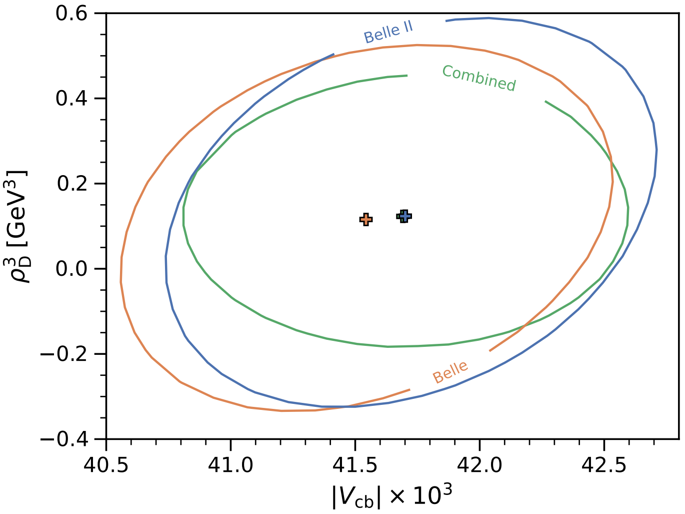


Figure 6.8: Comparison between Belle, Belle II and the combined fit for the correlation between $|V_{cb}|$ and ρ_D^3. The crosses indicate the best-fit points. ρ_D^3 is expressed in GeV at the appropriate power.

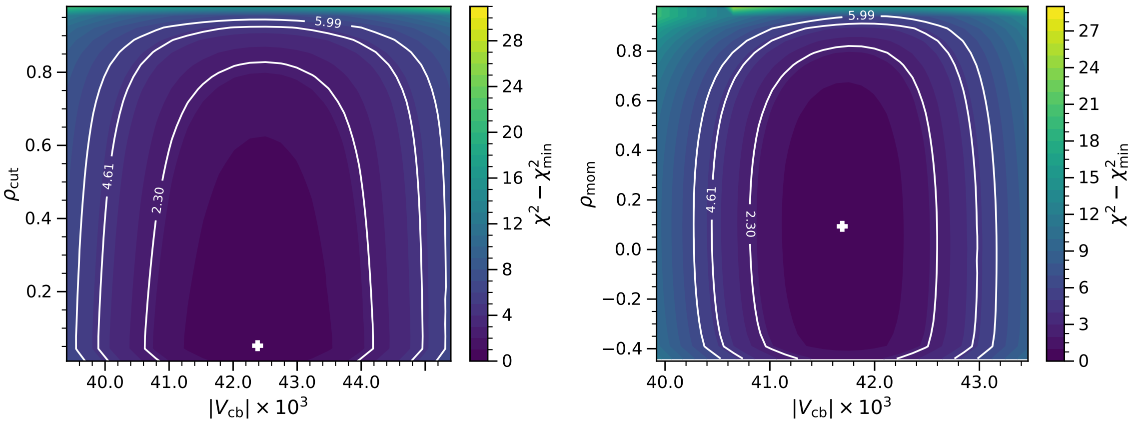


Figure 6.9: Two-dimensional χ^2 profile scans of $|V_{cb}|$ versus ρ_{cut} and ρ_{mom}. The minimum χ^2_{min} is subtracted from the χ^2 function.

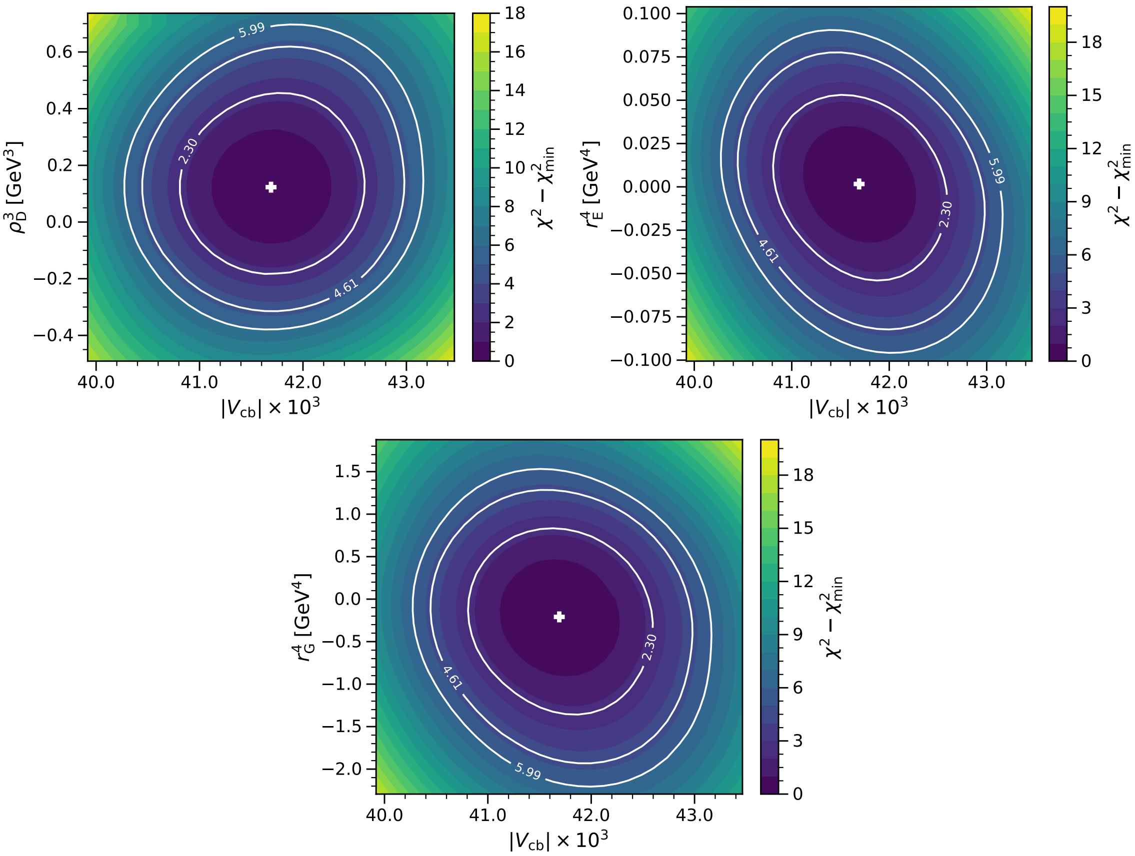


Figure 6.10: Two-dimensional χ^2 profile scans of $|V_{cb}|$ versus ρ_D^3, r_E^4 and r_G^4. The minimum χ^2_{min} is subtracted from the χ^2 function.

6.3.3 Determination of the HQE Parameters

Besides the fit result for $|V_{cb}|$, the determination the HQE parameters is also interesting. In the fits, the parameters μ_G^2 and μ_π^2 are constrained to external inputs. Table 6.4 shows that the fit to the q^2 moments has little further sensitivity to those parameters.

The only free HQE parameters in the default fit are ρ_D^3, r_E^4, and r_G^4. The fit to the q^2 moment gives

$$\rho_D^3 = (0.12 \pm 0.12|_{\text{Exp.}} \pm 0.13|_{\text{Theo.}} \pm 0.11|_{\text{Constr.}})\ \text{GeV}^3 = (0.12 \pm 0.20)\ \text{GeV}^3, \tag{6.31}$$

with the uncertainties stemming from the experimental and theory uncertainty on the moments and on the constraints. The uncertainty on the $B \to X_c \ell \nu_\ell$ branching ratio and the prediction on the total rate has no impact on the uncertainty of the HQE parameters.

This fit reports values for r_E^4 and r_G^4, the only order $1/m_b^4$ parameters included in the fit, for the first time in a purely data-driven way. One finds

Table 6.6: Results of our default fit using both Belle and Belle II data for $|V_{cb}|$, $m_b^{\text{kin}}(1\,\text{GeV})$, $\overline{m}_c(2\,\text{GeV})$, ρ_D^3, and the correlation parameters ρ_{cut} and and ρ_{mom}. The $O(1/m_b^4)$ terms are set to zero. All parameters are expressed in GeV at the appropriate power.

	$\lvert V_{\text{cb}}\rvert \times 10^3$	m_b^{kin}	$\overline{m}_c$	μ_G^2	μ_π^2	ρ_D^3	ρ_{cut}	ρ_{mom}
Value	41.76	4.56	1.10	0.38	0.43	0.03	0.05	0.14
Uncertainty	0.57	0.02	0.01	0.07	0.24	0.02	-	-

$$
\begin{aligned}
r_E^4 &= (0.02 \pm 0.21|_{\text{Exp.}} \pm 0.27|_{\text{Theo.}} \pm 0.00|_{\text{Constr.}}) \times 10^{-1}\ \text{GeV}^4 \\
&= (0.02 \pm 0.34) \times 10^{-1}\ \text{GeV}^4 \quad \text{and} \qquad (6.32)\\
r_G^4 &= (-0.21 \pm 0.42|_{\text{Exp.}} \pm 0.49|_{\text{Theo.}} \pm 0.25|_{\text{Constr.}})\ \text{GeV}^4 \\
&= (-0.21 \pm 0.69)\ \text{GeV}^4 \,. \qquad (6.33)
\end{aligned}
$$

Here, both values are compatible with zero. But r_E^4, the $1/m_b^4$ HQE parameter with the presumably most sensitivity, is constrained to values well below $1\,\text{GeV}^4$ or even Λ_{QCD}^4. Also the result for r_G^4 excludes unexpected large values for this parameter. This implies that the HQE seems well behaved and valid up to order $1/m_b^4$. The result for ρ_D^3 reported in Equation (6.31) is also compatible with zero. This holds as well for the scans with fixed correlation parameters given in Figure E.11. Since here, the definition of ρ_D^3 includes an additional $1/m_b$ contribution [39], the result cannot be directly compared to $\rho_D^3 = (0.185 \pm 0.031)\,\text{GeV}^3$ stated in [58].

To explore the apparent different sensitivities in the fit to the q^2 moments and the fit in [58], the two-dimensional $\Delta\chi^2$ scans of ρ_D^3 versus r_E^4, ρ_D^3 versus r_G^4, and r_E^4 versus r_G^4 are investigated. The corresponding scans are shown in Figure 6.11. All three parameters are highly (anti-)correlated. Also the shape changes of the q_n predictions as functions of the lower q^2 threshold are similar for ρ_D^3, r_E^4, and r_G^4 indicating that they can compensate each other. It seems that at least in a fit to only the q^2 moments, the fit is only sensitive to a linear combination of these HQE parameters. Together with the conservatively estimated theory uncertainty, this may explain the limited sensitivity to ρ_D^3. Interestingly, a fit without the inclusion of $1/m_b^4$ terms yields $\rho_D^3 = (0.03 \pm 0.02)\,\text{GeV}^3$. The fit result is listed in Table 6.6. One observes also slightly higher central value for $|V_{cb}|$ compared to the default fit setup. This determination of ρ_D^3 differs substantially from the one obtained in [58], which may indicate that the q^2 moments actually add additional information on the HQE parameters.

6.3.4 Fit Validation

The fit is validated using simulated pseudo-experiments (toy data). Ensembles of toy data sets are generated using a multivariate Gaussian distribution with mean values set to the theory prediction evaluated at the best-fit points and the full covariance matrix C_{tot}. C_{theo} is constructed using the fitted values for ρ_{cut} and ρ_{mom}. The mean values in the Gaussian distribution used to constrain m_b^{kin}, $\overline{m}_c$, μ_G^2, and μ_π^2 are also varied around their best-fit points with the prior uncertainties. Figure 6.12 in Appendix E shows the pull distributions of $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 which are compatible with standard

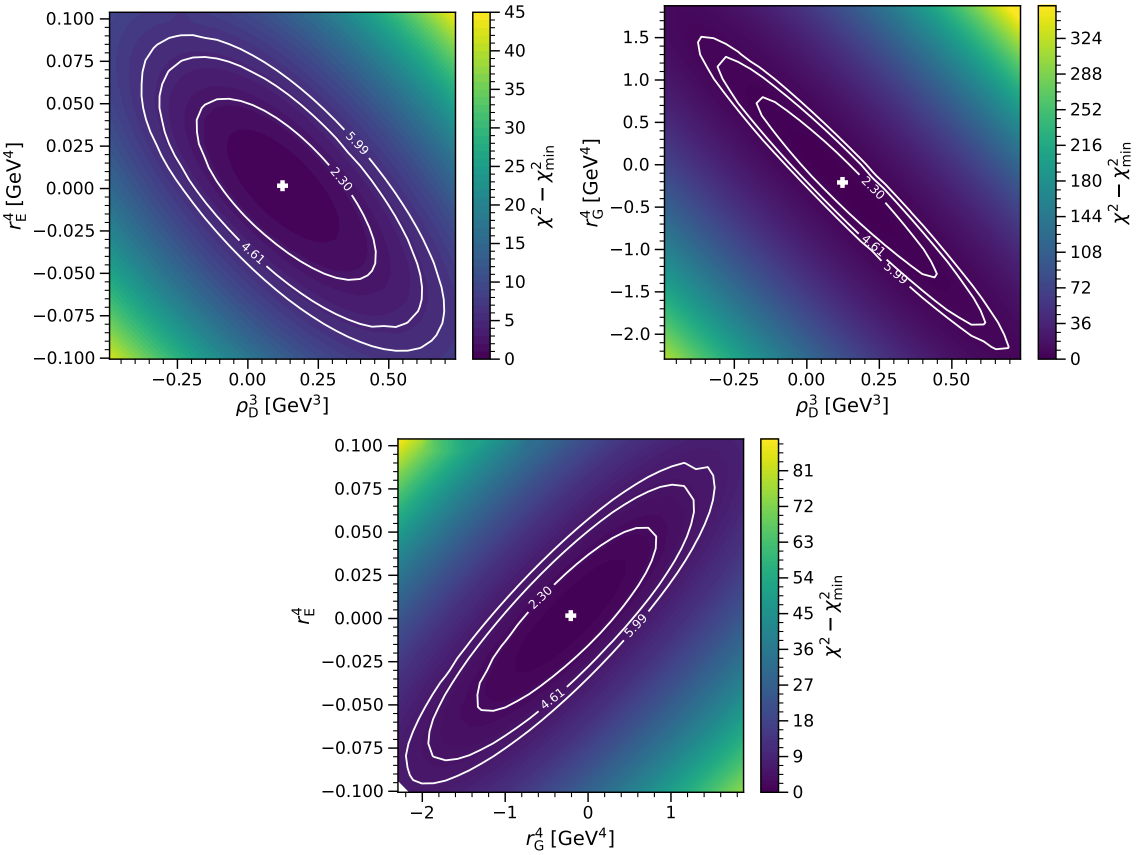


Figure 6.11: Two-dimensional χ^2 profile scans of ρ_D^3 versus r_E^4 (top left) and r_G^4 (top right), and of r_E^4 versus r_G^4 (bottom). The minimum χ^2_{min} is subtracted from the χ^2 function.

normal distributions. Thus, the estimated parameter values and uncertainties are considered to be unbiased. The distributions for the nuisance parameters ρ_{cut} and ρ_{mom} resulting from the toy fits are shown in Figure 6.13. Since they are not Gaussian distributed, an approximated 68% confidence interval around the mode of the distribution is constructed to estimate their uncertainties. These uncertainties are reported in Table 6.4.

The individual contributions to the fit uncertainties are also estimated using the same toy approach used to validate the fit procedure. Here, the ensembles of pseudo-experiments are not generated using the full covariance and variation of the constraint mean values but with separated contributions. The pull distributions for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 resulting from variations using only the experimental and theoretical covariance matrix are shown in Figures E.1 and E.2, respectively. Fits to pseudo-experiments generated with the uncertainty of the experimental branching fraction and the theory prediction for the total rate only result in a variation in $|V_{cb}|$. The corresponding pull distributions are shown in Figures E.3 and E.4. Lastly, the observed parameter pull distributions resulting from variations of the mean values in the constraints are also not Gaussian. These pull distributions are given in Figure E.5.

In the case where all separate pull distributions follow a Gaussian distribution, the relative

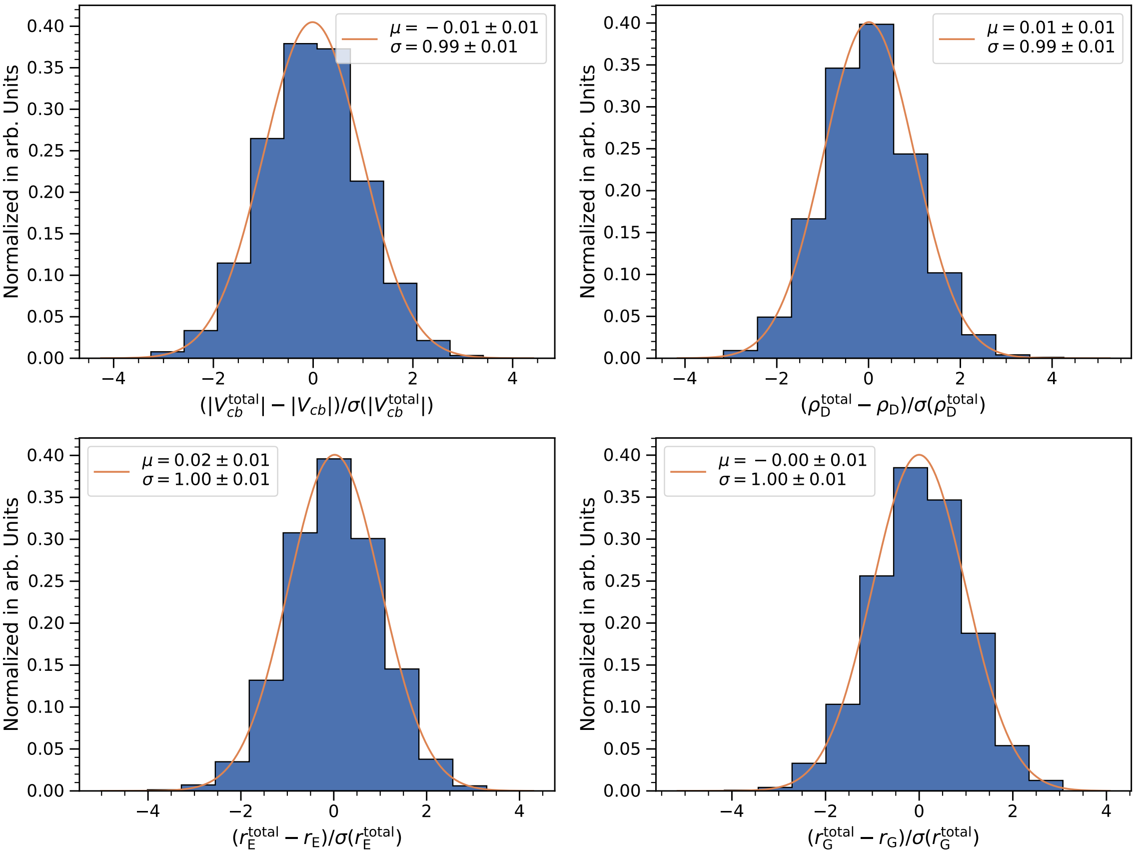


Figure 6.12: Pull distributions for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 calculated from fits to toy data sets sampled using the full covariance matrix and constraint uncertainty.

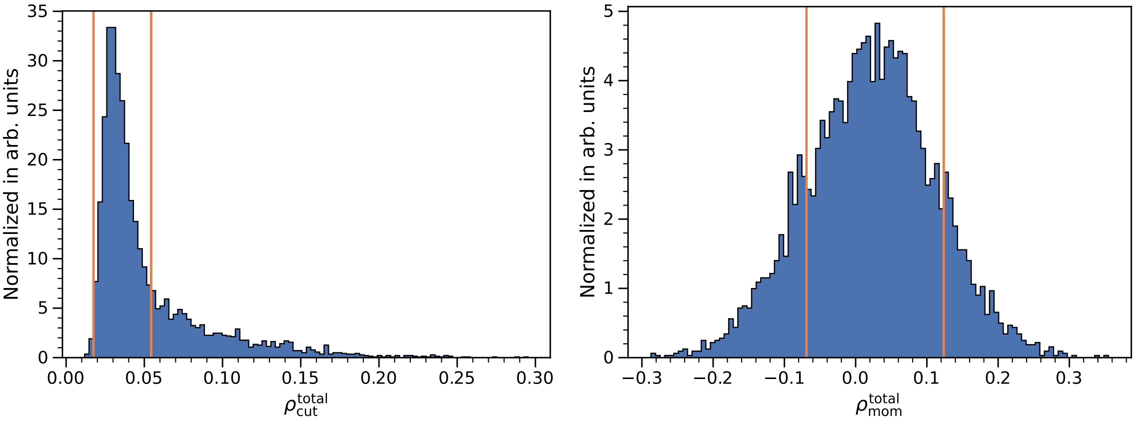


Figure 6.13: Distributions for ρ_{cut} and ρ_{mom} obtained from fits to toy data sets sampled using the full covariance matrix and constraint uncertainty.

Table 6.7: Fit result including all $1/m_b^4$ parameters with a standard normal Gaussian constraint. All parameters are expressed in GeV at the appropriate power.

	$\lvert V_{cb}\rvert \times 10^3$	m_b^{kin}	$\overline{m}_c$	μ_G^2	μ_π^2	ρ_D^3	r_G^4	$r_E^4 \times 10$	s_E^4	s_{qB}^4	s_B^4	ρ_{cut}	ρ_{mom}
Value	41.69	4.56	1.09	0.37	0.43	0.10	-0.12	0.04	-0.04	-0.02	0.04	0.05	0.10
Uncertainty	0.59	0.02	0.01	0.07	0.24	0.18	0.68	0.31	0.95	0.99	0.95	$^{+0.03}_{-0.01}$	$^{+0.10}_{-0.10}$

Table 6.8: Uncertainties on $|V_{cb}|$ estimated from individual variations of the neglected $1/m_b^4$ parameters s_E^4, s_B^4, and s_{qB}^4.

	s_E^4	s_B^4	s_{qB}^4	Total
$\sigma_{\lvert V_{cb}\rvert} \times 10^3$	0.22	0.03	0.07	0.23

contribution of a single uncertainty source σ_{source} would be given by $\sigma_{\text{source}} = \sigma_{\text{source}}^{\text{pull}} \times \sigma_{\text{fit}}$. Here, $\sigma_{\text{source}}^{\text{pull}}$ and σ_{fit} denote the standard deviation of the corresponding pull distribution and the total fit uncertainty, respectively. It should hold that all $\sigma_{\text{source}}^{\text{pull}}$ summed in quadrature are equal to unity. Since the pull distributions related to the variation of the constraint mean values are non-Gaussian, their contribution is estimated as $\sigma_{\text{Constr.}}^{\text{pull}} = \sqrt{1 - \sum(\sigma_{\text{source}}^{\text{pull}})^2}$.

6.3.5 Including all $1/m_b^4$ Terms

To study the impact of the neglected parameters, an additional fit including also s_E^4, s_B^4, and s_{qB}^4 is performed. In this fit, all $1/m_b^4$ parameters are constrained with a standard Gaussian distribution $\mathcal{G}(0, 1)$ to achieve a well-behaved fit. The fit results are given in in Table 6.7. No significant deviations from the default fit scenario are observed. Again, this fit demonstrates the limited sensitivity to the $1/m_b^4$ HQE terms. Only the post-fit uncertainty of r_E^4 and r_G^4 can be reduced from unity while the other parameters show no significant reduction. Most importantly, the same $|V_{cb}|$ value as in the default fit is obtained

$$|V_{cb}| = (41.69 \pm 0.59) \times 10^{-3}. \tag{6.34}$$

To account for the missing $1/m_b^4$ corrections from s_E^4, s_B^4, and s_{qB}^4, an additional uncertainty is added to the $|V_{cb}|$ result of the default fit in Equation (6.27). This uncertainty is studied by the impact on $|V_{cb}|$ of independent variations by ±1 GeV^4 of the neglected parameters as

$$\sigma_{|V_{cb}|}(s_x^4) = \frac{\left||V_{cb}|(s_x^4 = 1\,\text{GeV}^4) - |V_{cb}|(s_x^4 = -1\,\text{GeV}^4)\right|}{2}. \tag{6.35}$$

Figure 6.2 shows the effect of these variations on $|V_{cb}|$. Here, the $B \to X_c \ell \nu_\ell$ branching fraction is fixed at the value given in Equation (6.25). The individual uncertainty estimates are listed in Table 6.8. In total, the additional uncertainty on $|V_{cb}|$ is estimated to be $\sigma_{|V_{cb}|} = 0.23 \times 10^{-3}$. This value is dominated by the variation of s_E^4.

The final result for the determination of $|V_{cb}|$ is therefore

$$|V_{cb}| = (41.69 \pm 0.59|_{\mathrm{fit}} \pm 0.23|_{\mathrm{h.o.}}) \times 10^{-3} = (41.69 \pm 0.63) \times 10^{-3}\,, \tag{6.36}$$

with both uncertainties added in quadrature.

CHAPTER 7

Summary and Conclusion

This thesis presents the measurement of raw and central q^2 moments of inclusive $B \longrightarrow X_c \ell \nu_\ell$ decays with the Belle II experiment. Measurements of the q^2 moments are provided with lower thresholds of q^2 itself ranging from $q^2 > 1.5\,\mathrm{GeV}^2$ up to $q^2 > 8.5\,\mathrm{GeV}^2$. The analysis probes up to 77% of the accessible $B \longrightarrow X_c \ell \nu_\ell$ phase space, improving the findings reported by the Belle Collaboration [31], and includes the experimentally challenging region of $q^2 \in [1.5, 2.5]\,\mathrm{GeV}^2$.

The measurements reported here make use of the Belle II data set, recorded at a CM energy of $\sqrt{s} = 10.58\,\mathrm{GeV}$ in 2019 and 2020, corresponding to a integrated luminosity of $62.8\,\mathrm{fb}^{-1}$. The $\Upsilon(4S) \longrightarrow B\overline{B}$ events are analyzed using a multivariate tagging algorithm to reconstruct one of the B mesons in the $\Upsilon(4S)$ decay in fully hadronic modes. The other B decay is reconstructed inclusively by selecting a well-reconstructed lepton candidate. The hadronic X system is identified with the remaining charged particle and photon candidates not used in the reconstruction of the tag-side B meson and the signal lepton. In combination with the hadronically reconstructed B_{tag} candidate and the inclusive X system, a direct reconstruction of the q^2 spectrum is possible. A kinematic fit using the reconstructed four-momenta of the B_{tag}, X, and lepton candidates is employed to improve the

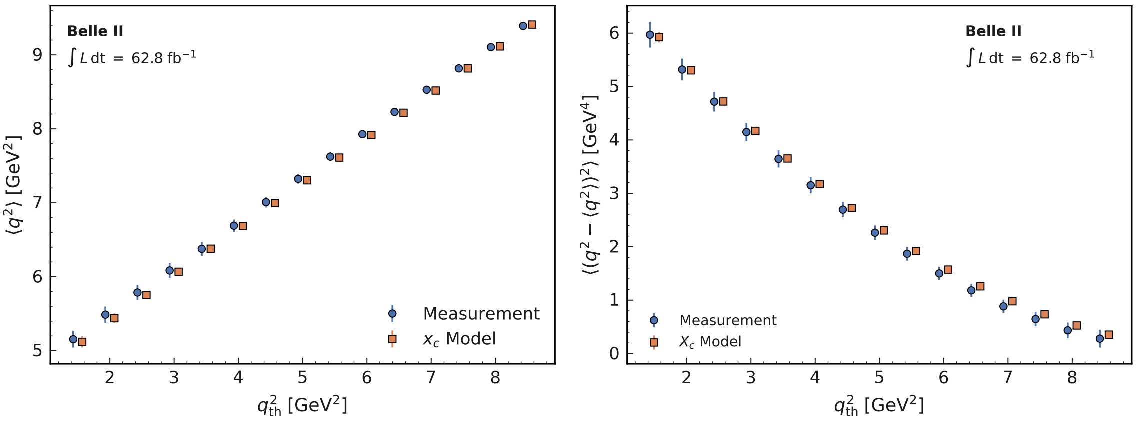


Figure 7.1: First raw and second central q^2 moments (blue) as functions of lower q^2 thresholds. The simulated moments (orange) are shown for comparison.

resolution on the reconstructed q^2 and m_X spectra. Background contributions from other processes are subtracted using an event-wise weight whose value is calculated from the background normalization, as obtained by a fit to the m_X spectrum and simulated background shapes. The reconstructed q^2 distribution is calibrated to remove reconstruction and resolution effects and other biases introduced in the selection. This allows to calculate the raw moments directly from the calibrated q^2 as a weighted mean. Figure 7.1 shows the measured raw and central q^2 moments of order one and two, respectively. No significant differences to the q^2 moments extracted from the simulated $B \to X_c \ell \nu_\ell$ spectrum are observed.

The total uncertainty for the measured q^2 moments is dominated by systematic uncertainties, where the uncertainties associated with the background normalization and shape, composition of the X_c system, and the simulated detector resolution are the dominant sources. A better understanding of the detector simulation will most likely lead to a more precise determination of the q^2 moments in the future. In addition, a better understanding of the remaining $B\overline{B}$ and continuum backgrounds will allow measurements with a q^2 selection below 1.5 GeV^2.

In addition, a first extraction of the magnitude of the fundamental CKM matrix element V_{cb} is performed using q^2 moments, as proposed in [30]. By using measurements of q^2 moments and exploiting reparameterization invariance, the number of non-perturbative matrix elements can be reduced from 13 to only 8 parameters at $\mathcal{O}(1/m_b^4)$ in the HQE. The reduced set of non-perturbative parameters allows a new attempt to probe the $1/m_b^4$ terms in a purely data-driven way. In this investigation, two out of five $1/m_b^4$ HQE terms are included in the fit.

A fit to total inclusive $B \to X_c \ell \nu_\ell$ branching fraction and the q^2 moments results in

$$\begin{aligned} |V_{cb}| &= (41.69 \pm 0.27|_{\mathcal{B}} \pm 0.31|_{\Gamma} \pm 0.18|_{\text{Exp.}} \pm 0.17|_{\text{Theo.}} \pm 0.34|_{\text{Constr.}} \pm 0.23|_{\text{h.o.}}) \times 10^{-3} \\ &= (41.69 \pm 0.63) \times 10^{-3}. \end{aligned} \tag{7.1}$$

Here the uncertainties originate from the experimental uncertainty on the branching fraction $\mathcal{B}$, the theoretical uncertainty on the total rate Γ, the experimental and theoretical uncertainties on the q^2 moments, and the uncertainty from external constraints. The last uncertainty accounts for missing higher order HQE terms not included in the fit. The fit is performed simultaneously to the moments measured in this work and to the moments reported by the Belle collaboration. The result is in good agreement with the current most precise inclusive $|V_{cb}|$ value determined in [58] and provides strong evidence that the inclusive $|V_{cb}|$ determination based on the HQE predictions is reliable and that the uncertainties are well under control. This represents an important finding in the persistent tension between exclusively and inclusively determined $|V_{cb}|$ values.

An extensive study of the a-priori unknown correlation structure of the covariance matrix describing the uncertainties in the theory predictions is performed. The default scenario includes two parameters used to construct the correlation structure directly into the fit as nuisance parameters. Overall, the theory correlations have a negligible impact on $|V_{cb}|$ and the central value remains stable for a large range of possible correlation scenarios.

The default fit scenario includes the HQE parameters ρ_D^3, r_E^4, and r_G^4 as free parameters. The fit to the q^2 moments results in

$$\begin{aligned}\rho_D^3 &= (0.12 \pm 0.12|_{\text{Exp.}} \pm 0.13|_{\text{Theo.}} \pm 0.11|_{\text{Constr.}})\ \text{GeV}^3 \\ &= (0.12 \pm 0.20)\ \text{GeV}^3. \end{aligned} \tag{7.2}$$

The two $1/m_b^4$ parameters r_E^4 and r_G^4 are determined for the first time purely in a data-driven way and one finds

$$\begin{aligned}r_E^4 &= (0.02 \pm 0.21|_{\text{Exp.}} \pm 0.27|_{\text{Theo.}} \pm 0.00|_{\text{Constr.}}) \times 10^{-1}\ \text{GeV}^4 \\ &= (0.02 \pm 0.34) \times 10^{-1}\ \text{GeV}^4 \quad \text{and} \end{aligned} \tag{7.3}$$

$$\begin{aligned}r_G^4 &= (-0.21 \pm 0.42|_{\text{Exp.}} \pm 0.49|_{\text{Theo.}} \pm 0.25|_{\text{Constr.}})\ \text{GeV}^4 \\ &= (-0.21 \pm 0.69)\ \text{GeV}^4. \end{aligned} \tag{7.4}$$

The extracted $1/m_b^4$ parameters are small and compatible with zero within their uncertainties, which implies that the HQE seems well-behaved up to this order.

The difference between the central values of $|V_{cb}|$ determined in this work and in [58] is mainly driven by the different $B \to X_c \ell \nu_\ell$ branching fraction values used as input in the fit. This emphasizes the need for additional new branching fraction measurements by Belle II, especially also as functions of lower q^2 thresholds. Further improvements are expected with the inclusion of currently missing higher order perturbative corrections for the HQE Wilson coefficients as indicated in Table 6.1. It may also be interesting to carry out a $|V_{cb}|$ determination with the combination of the q^2 moments with available measurements of m_X and E_ℓ moments. However, this would require new HQE calculations for the q^2 moments based on the full set of non-perturbative parameters at order $1/m_b^4$. Another possibility is the inclusion of observables like the forward-backward asymmetry [136] or partial moments [137]. All these improvements could be an important possibility to push the precision of inclusively determined $|V_{cb}|$ to below the percent level.

APPENDIX A

Additional Material m_X Fit

This appendix shows the pre- and post-fit m_X distributions in the $B^+\ell^-$, $B^0\ell^-$, and $B^0\ell^+$ reconstruction channels in Figures A.1 to A.17 for lower q^2 thresholds starting at $0.5\,\mathrm{GeV}^2$ up to $8.5\,\mathrm{GeV}^2$ in $0.5\,\mathrm{GeV}^2$ increments. The corresponding post-fit nuisance parameter pull distributions are shown in Figures A.18 to A.34.

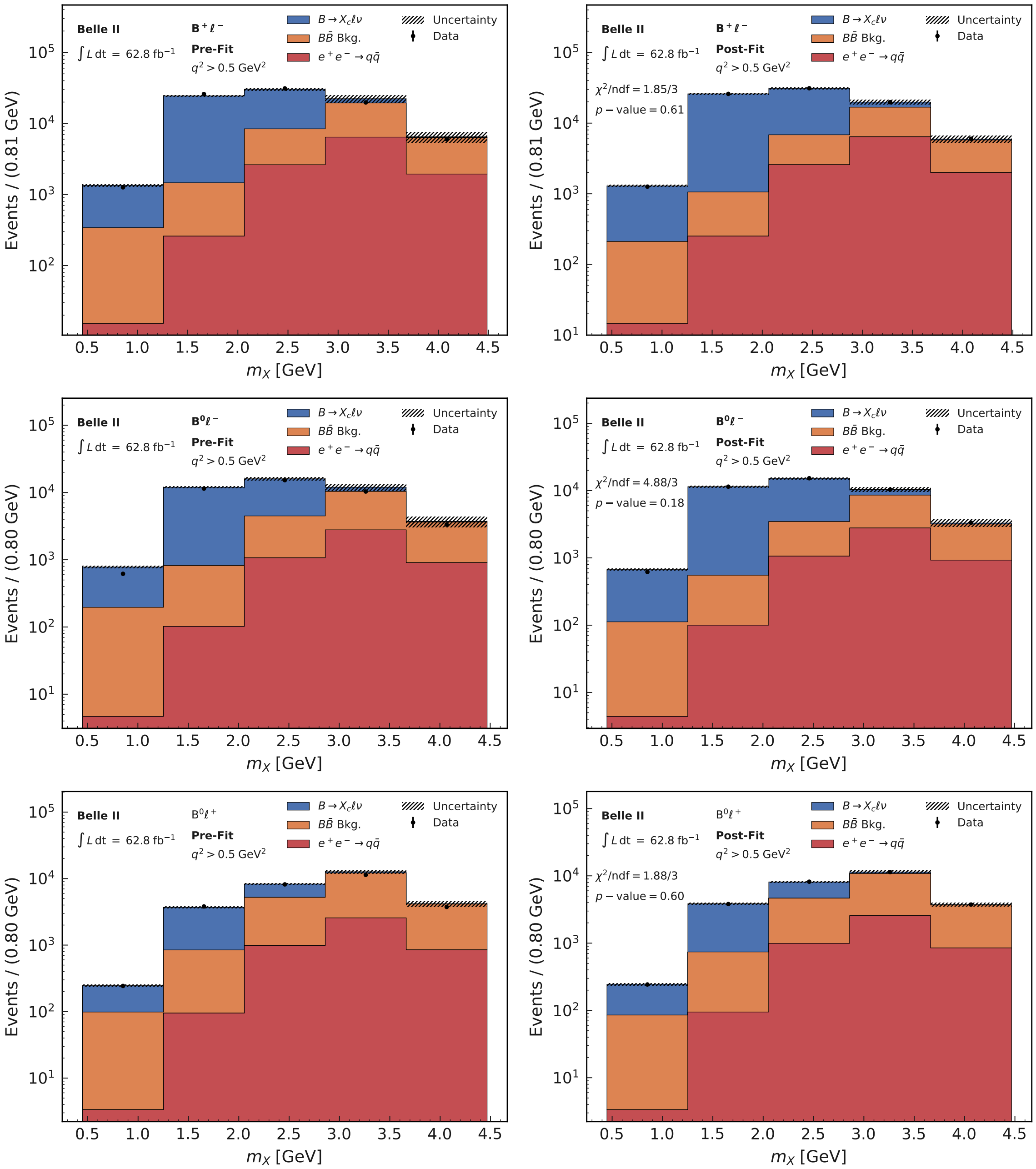


Figure A.1: m_X distribution for a $q^2 > 0.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

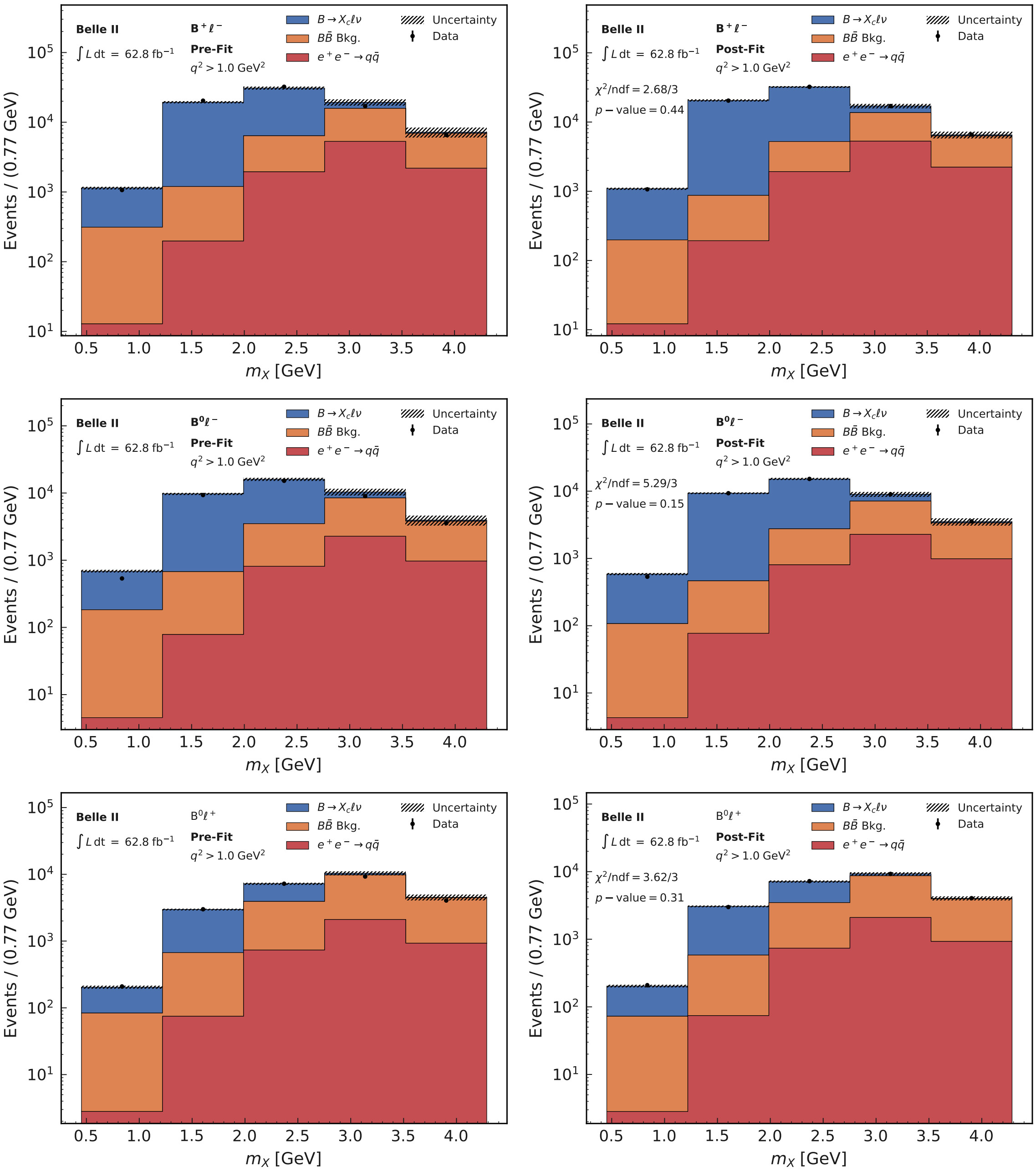


Figure A.2: m_X distribution for a $q^2 > 1.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c \ell \nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

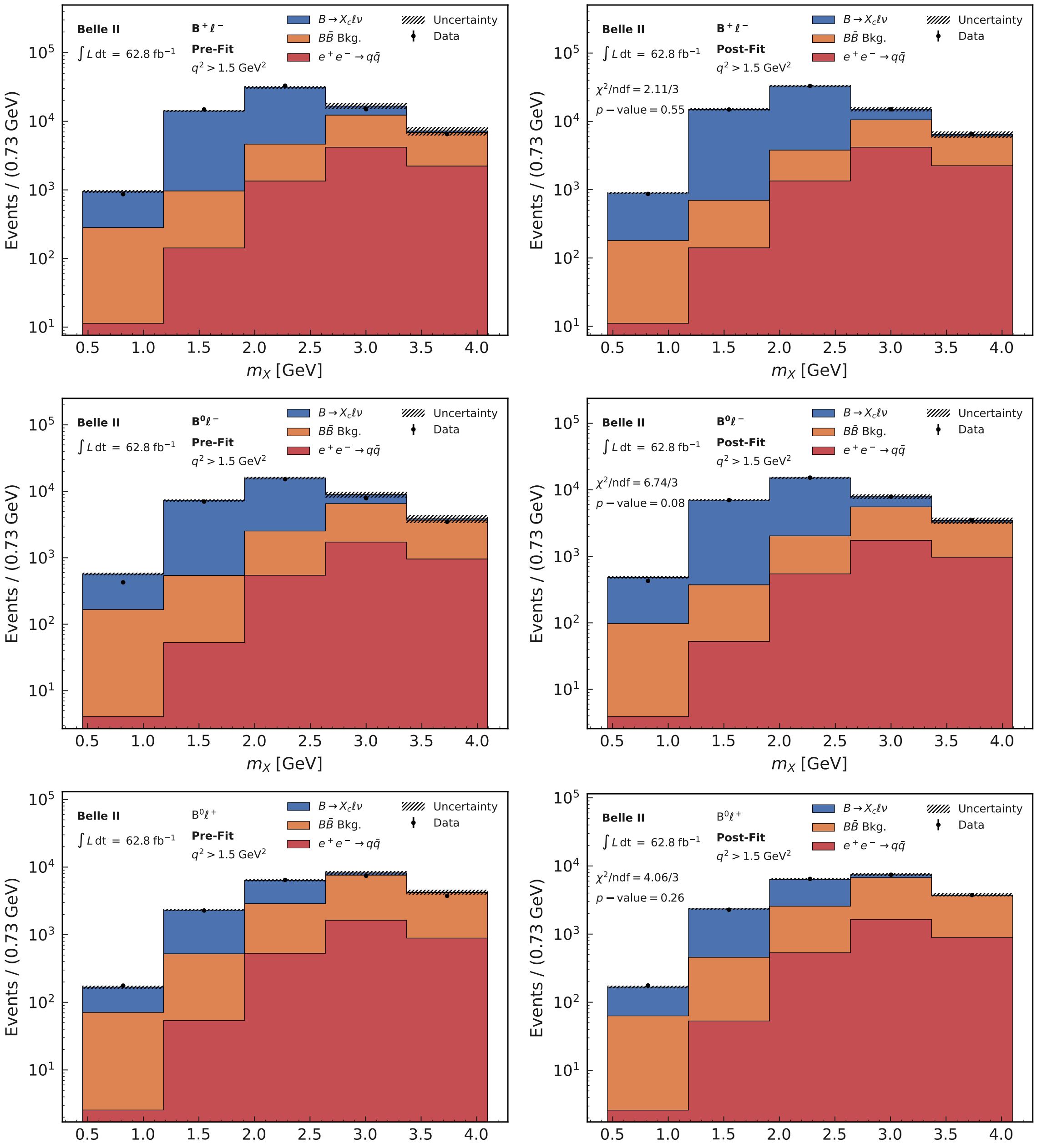


Figure A.3: m_X distribution for a $q^2 > 1.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

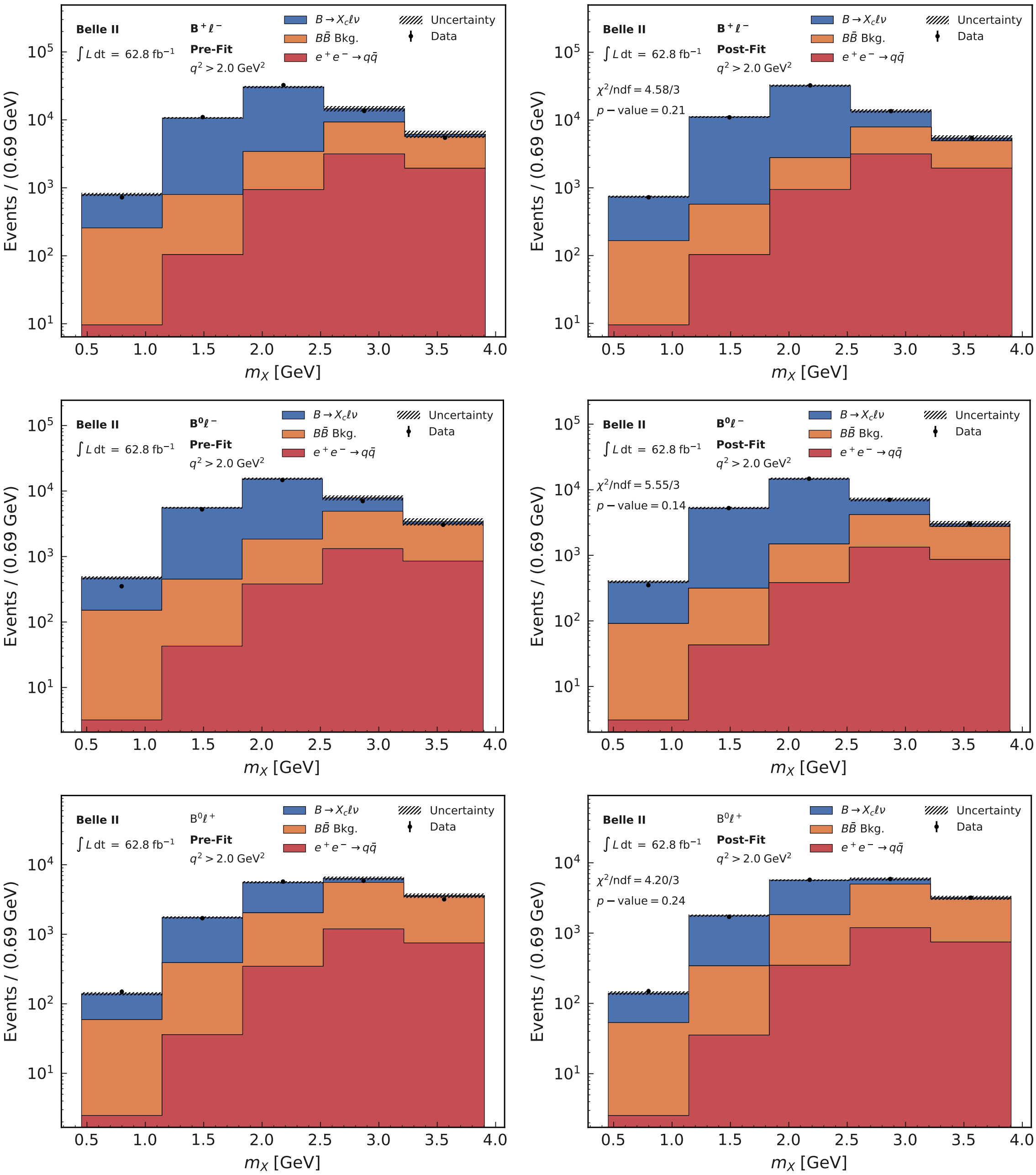


Figure A.4: m_X distribution for a $q^2 > 2.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

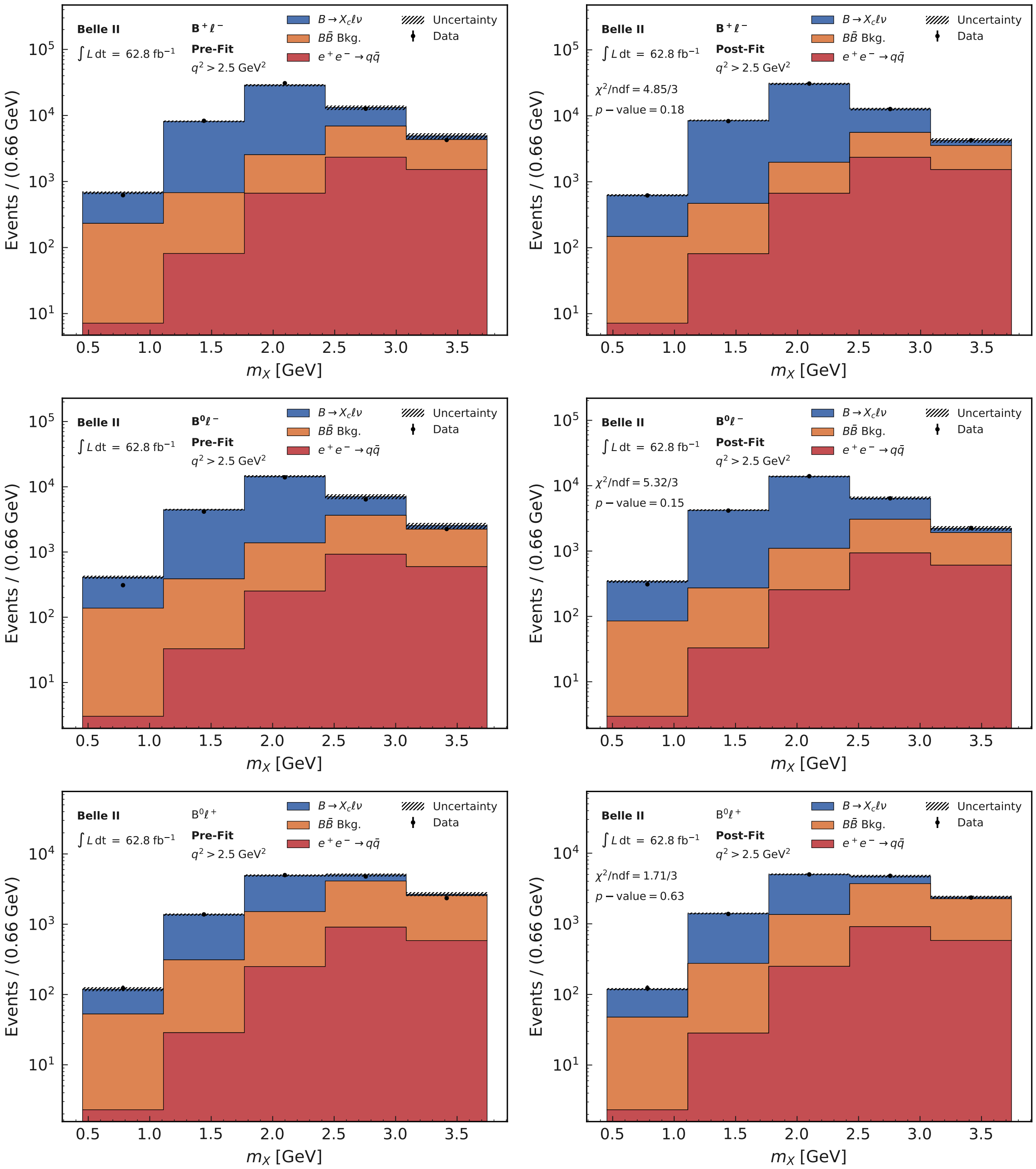


Figure A.5: m_X distribution for a $q^2 > 2.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

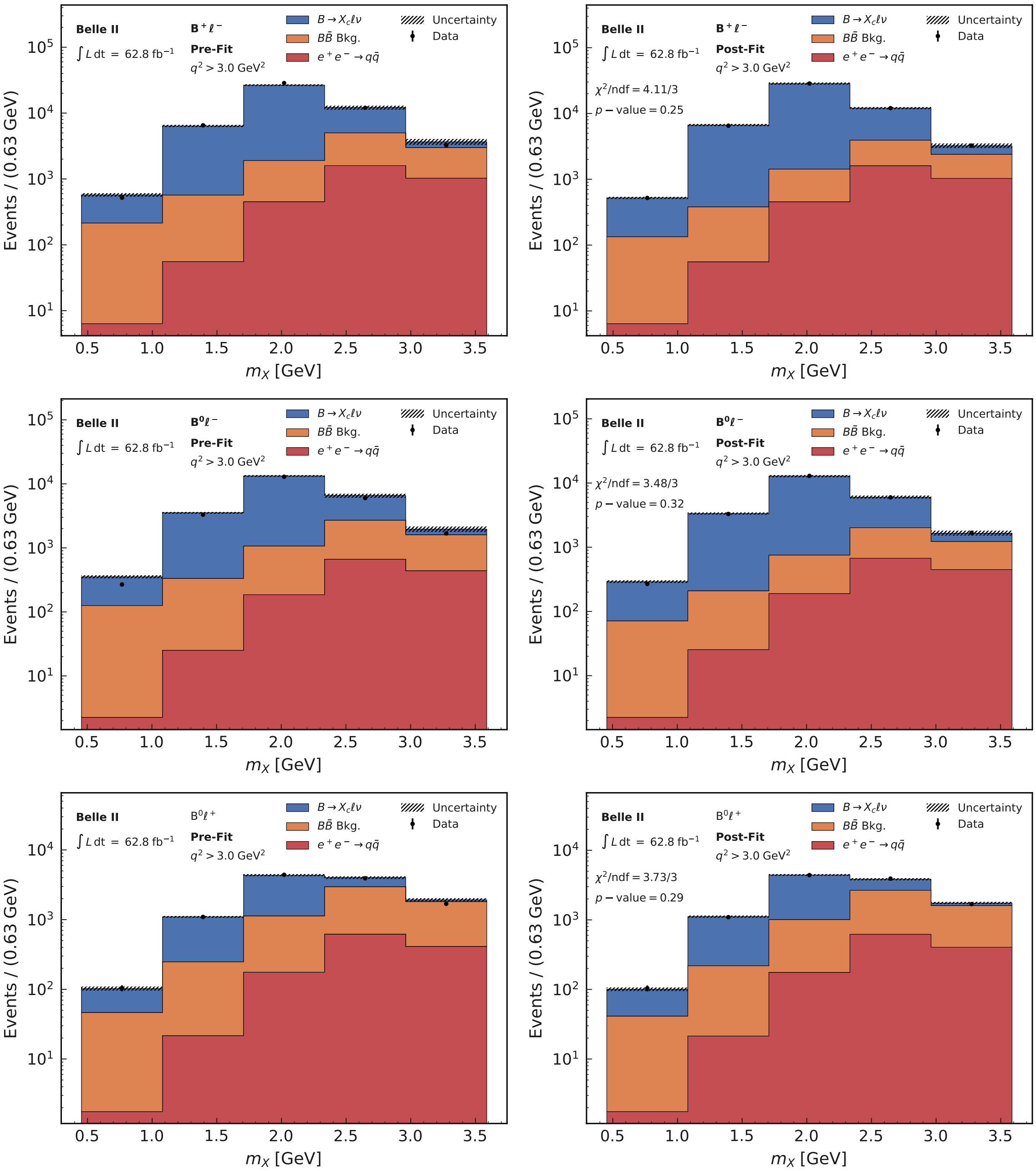


Figure A.6: m_X distribution for a $q^2 > 3.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

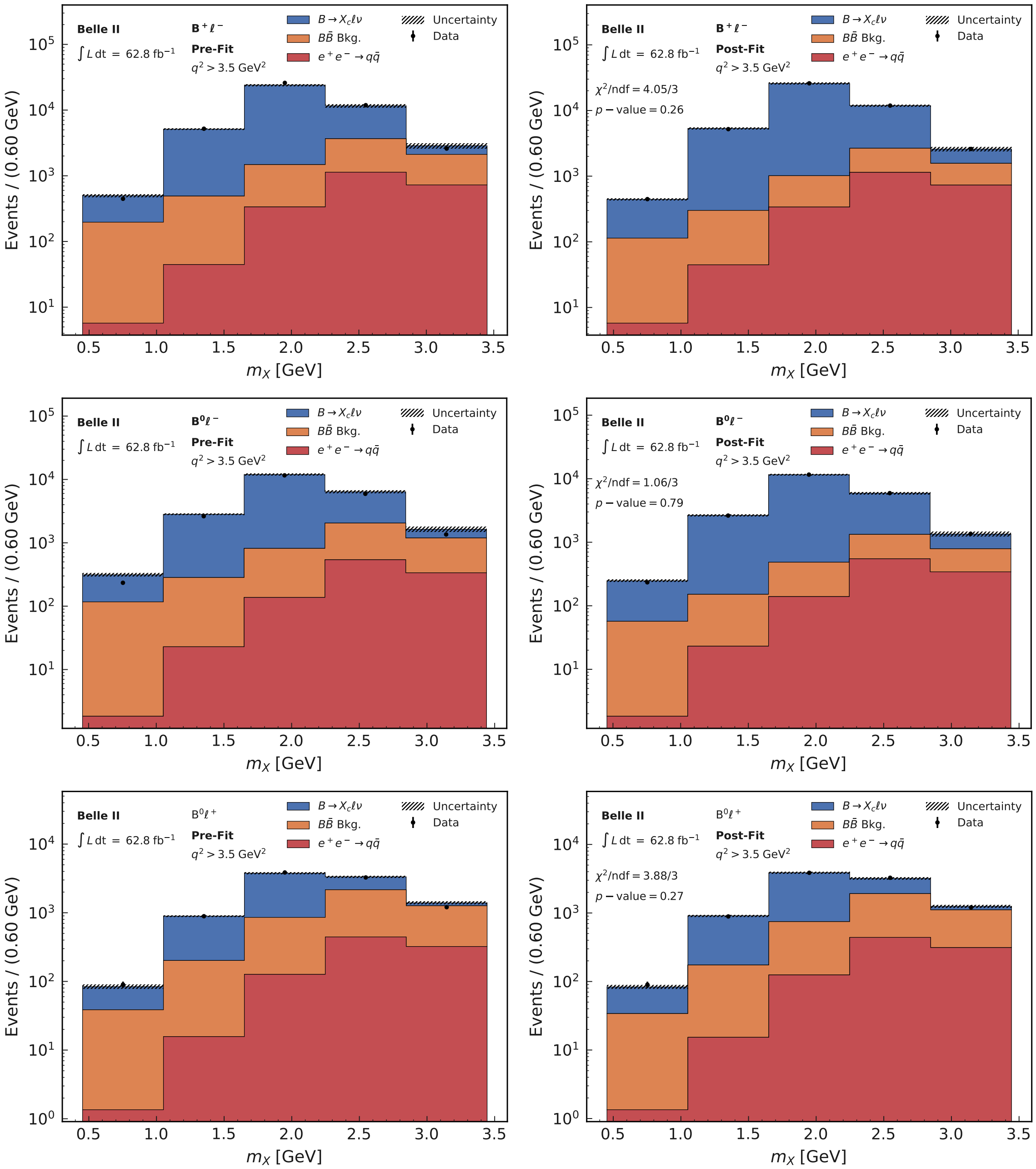


Figure A.7: m_X distribution for a $q^2 > 3.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

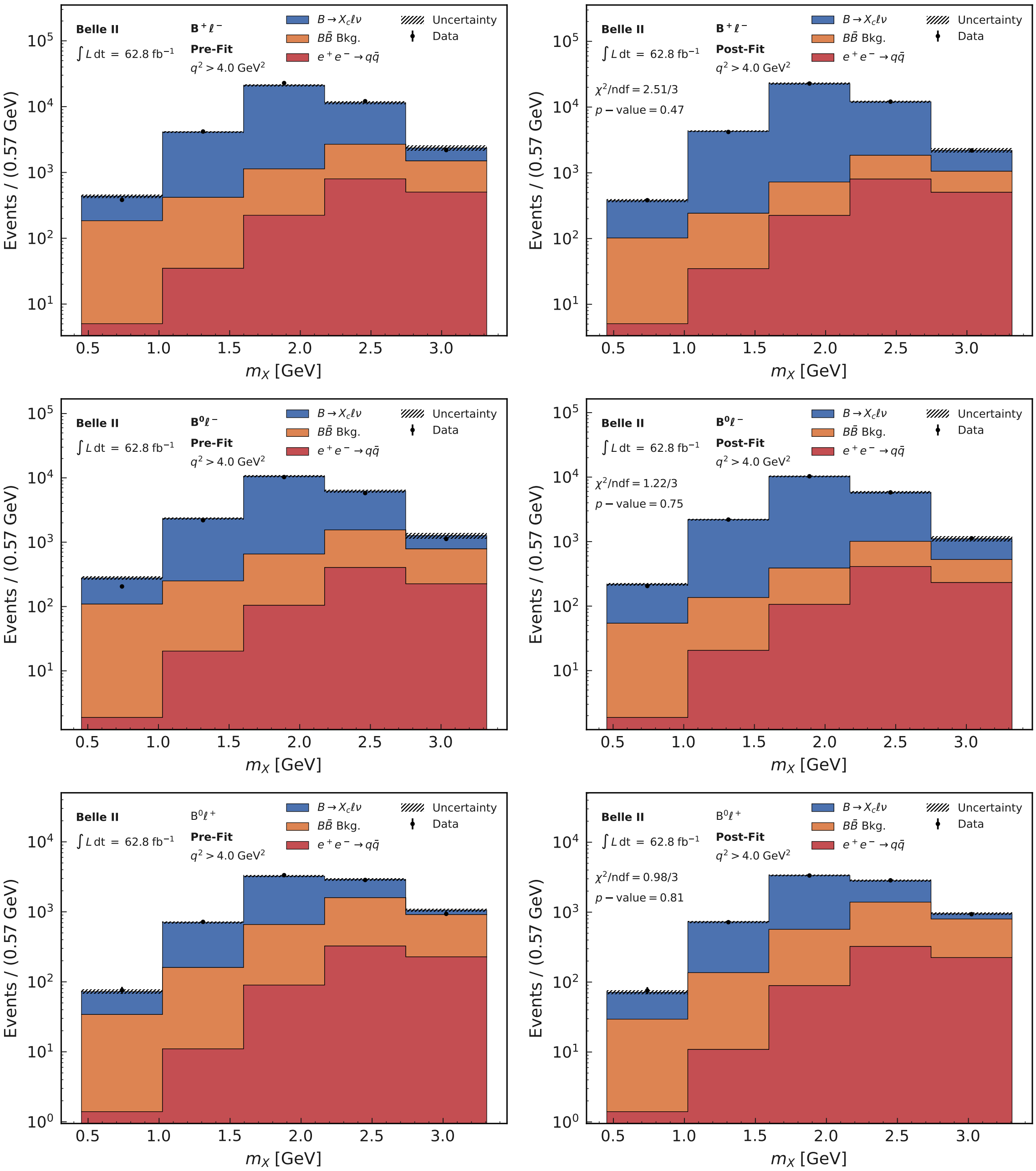


Figure A.8: m_X distribution for a $q^2 > 4.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

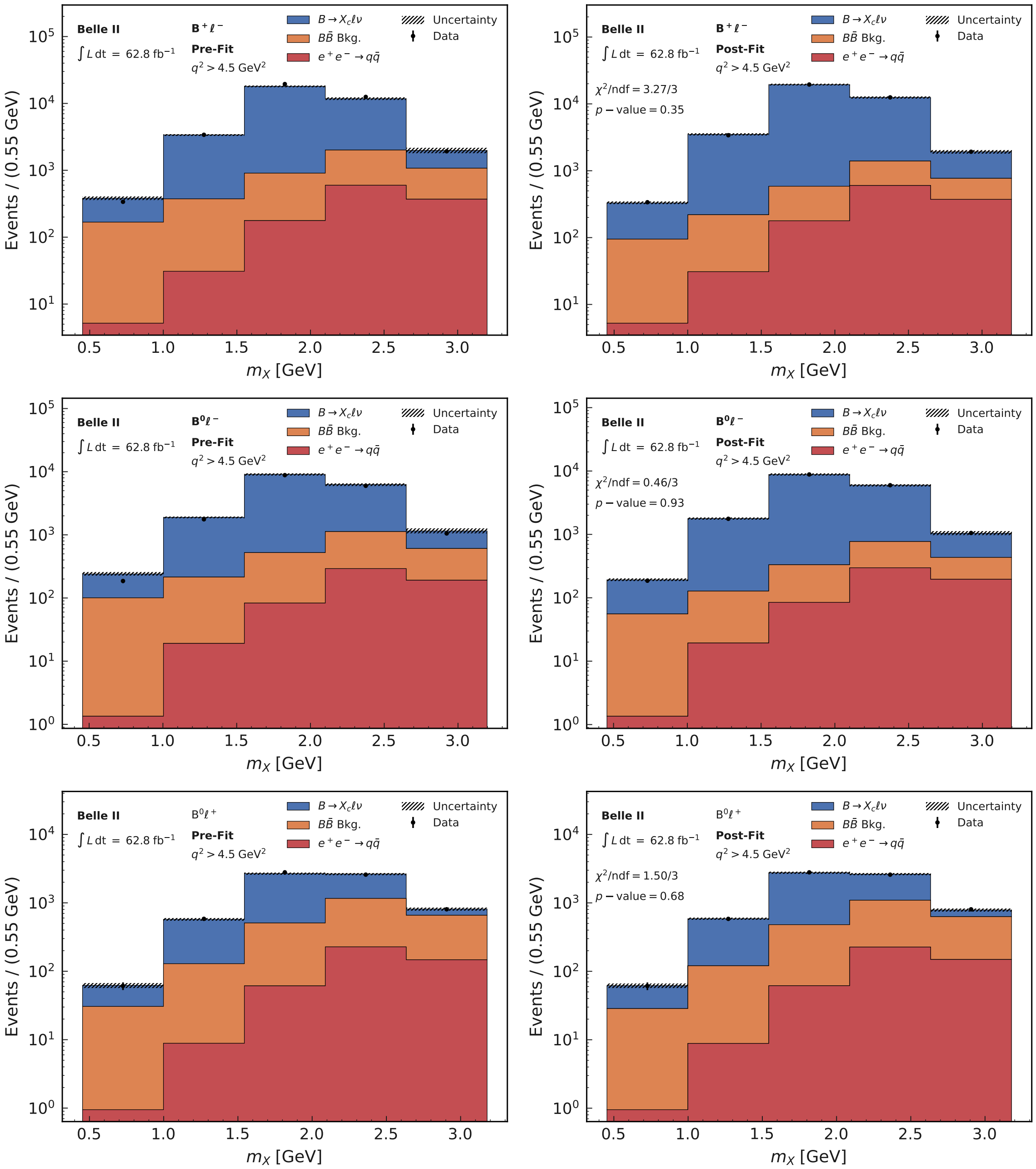


Figure A.9: m_X distribution for a $q^2 > 4.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

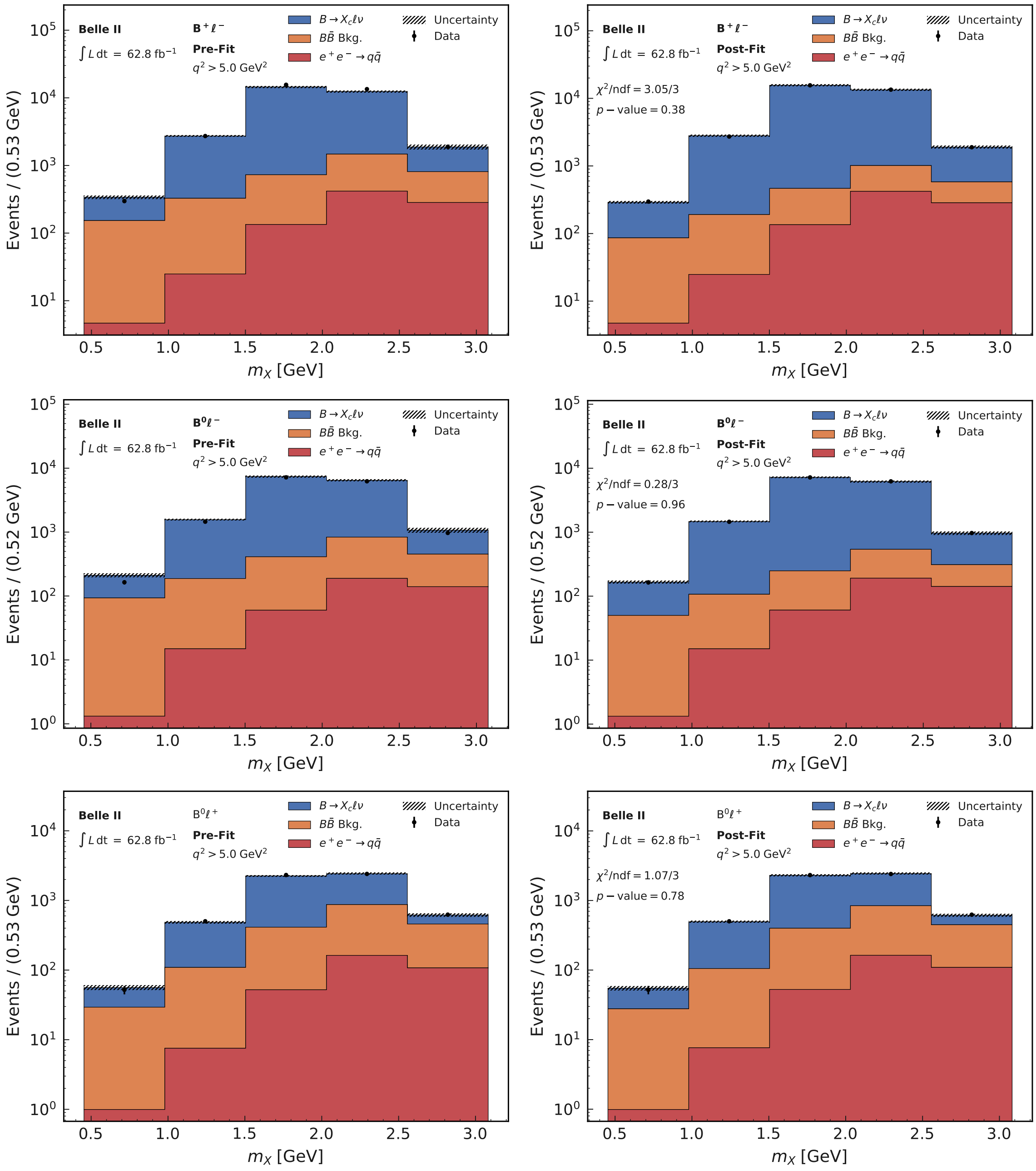


Figure A.10: m_X distribution for a $q^2 > 5.0\,\text{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

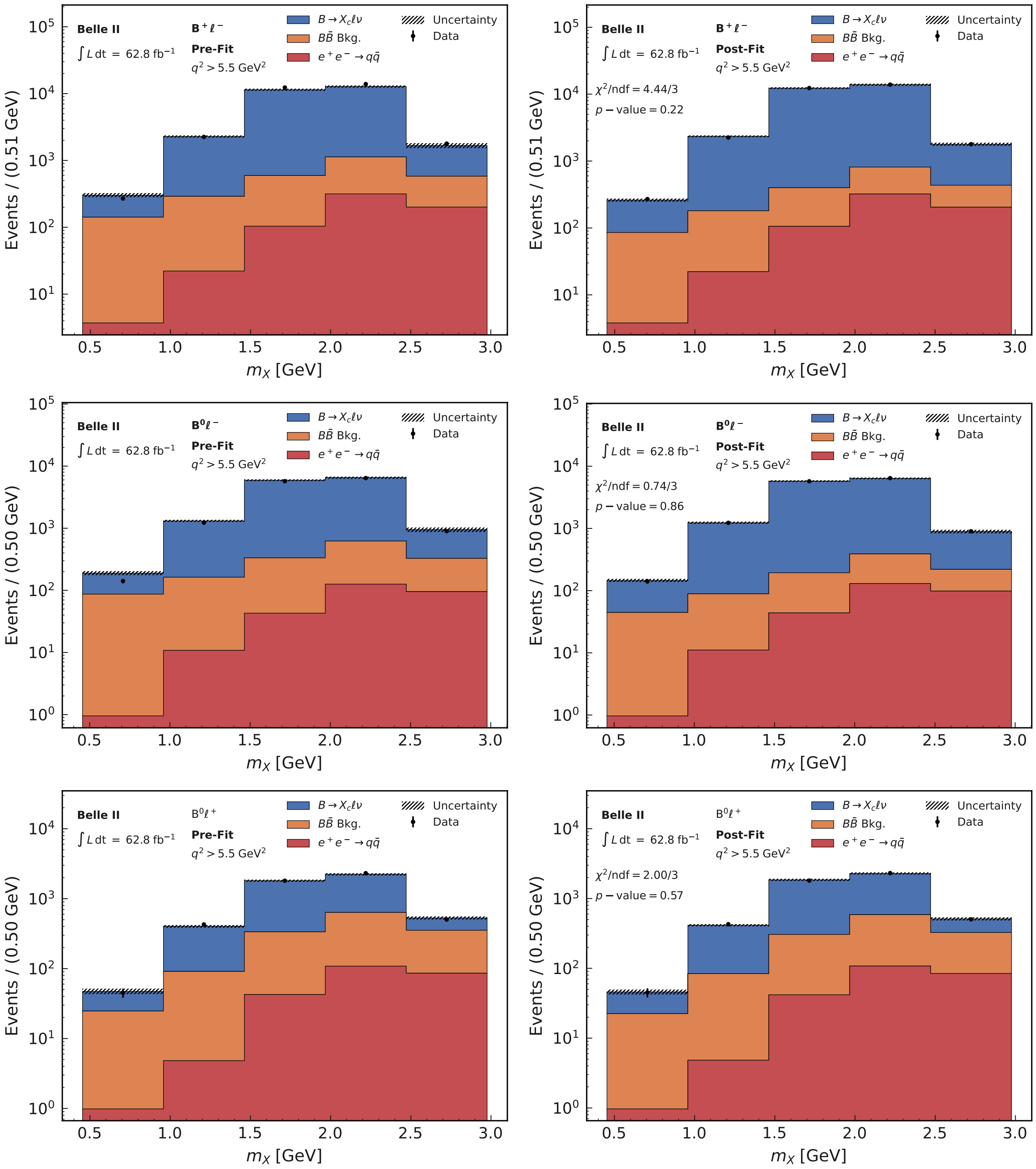


Figure A.11: m_X distribution for a $q^2 > 5.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

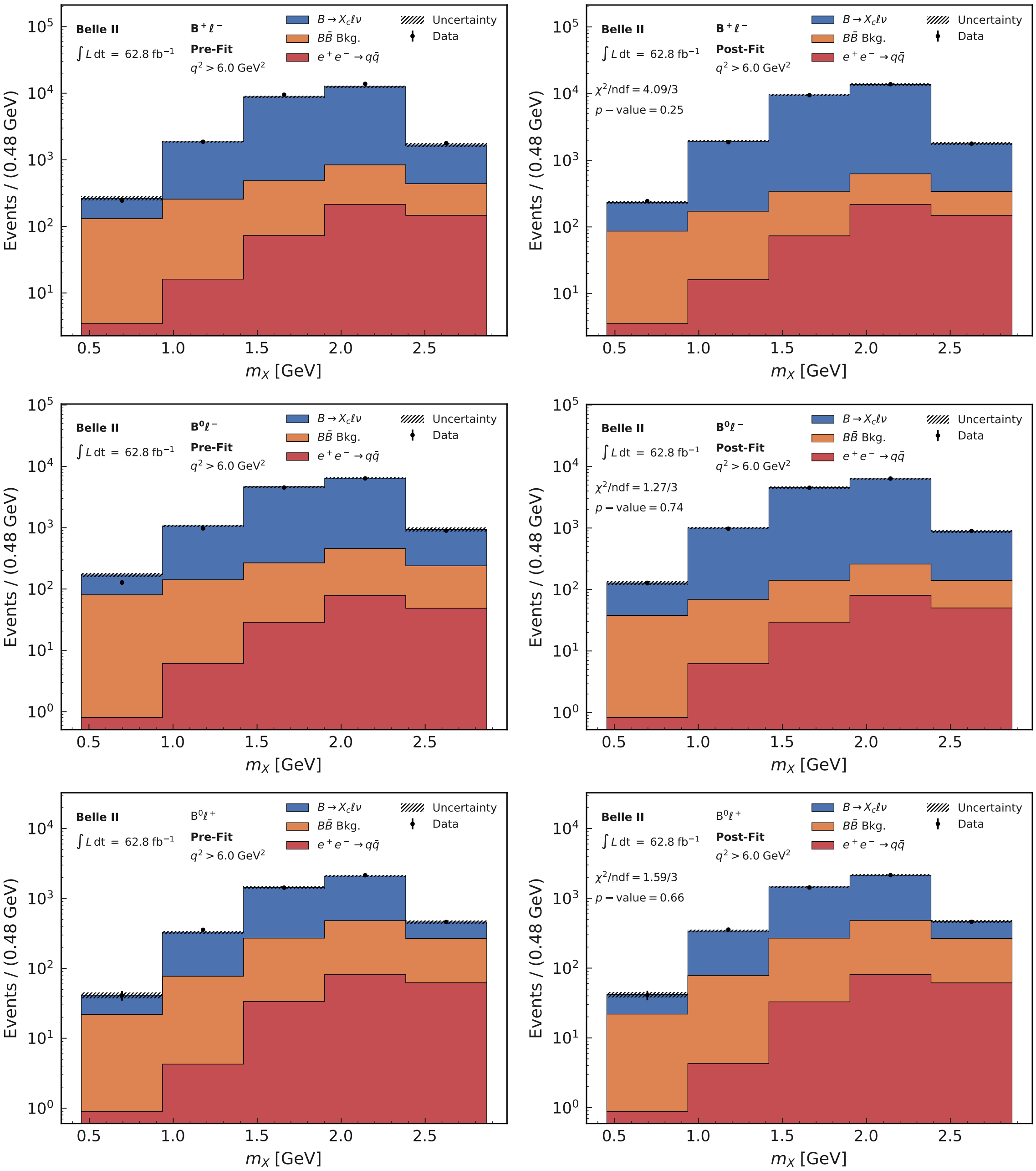


Figure A.12: m_X distribution for a $q^2 > 6.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

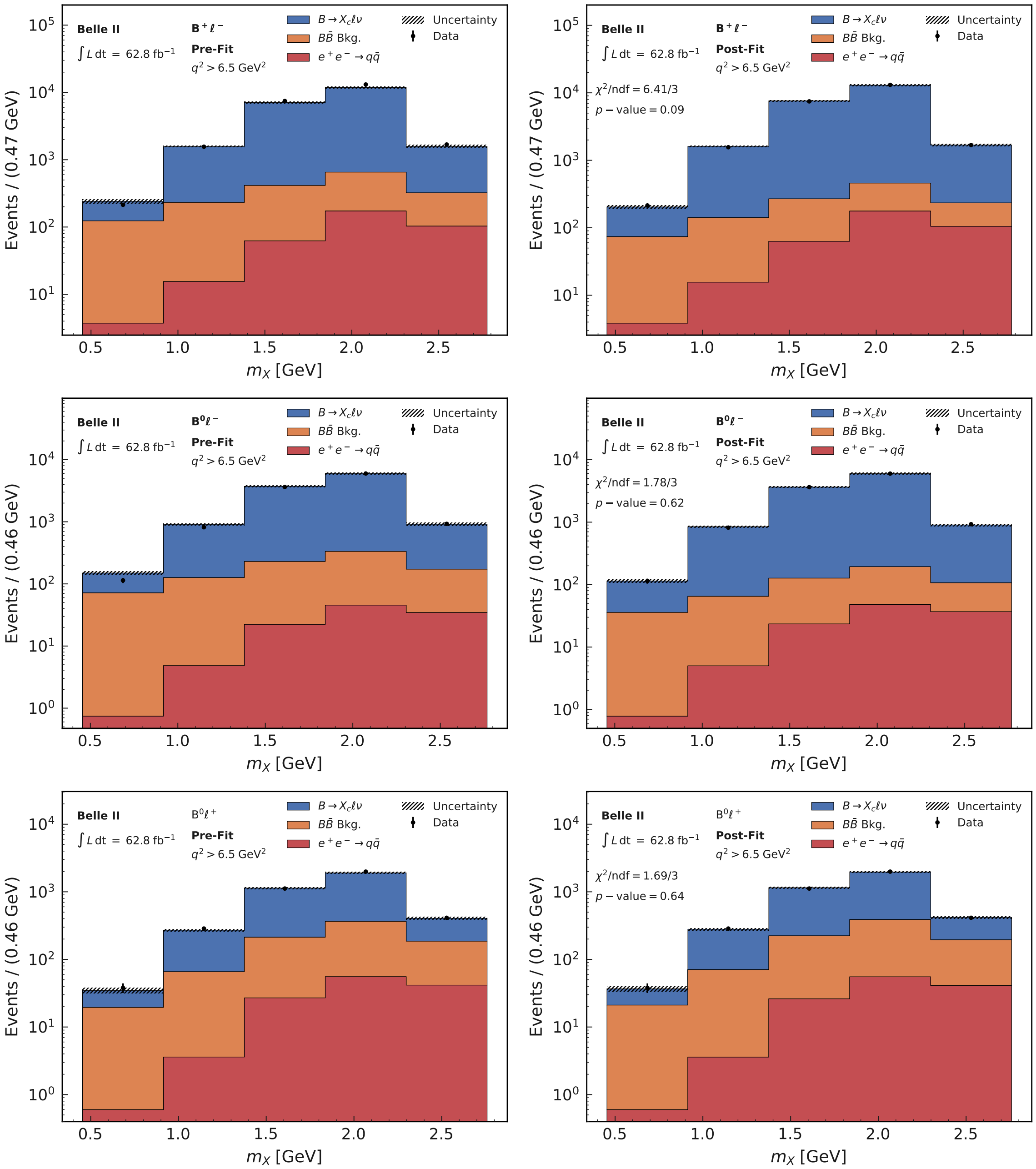


Figure A.13: m_X distribution for a $q^2 > 6.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \to X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \to q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

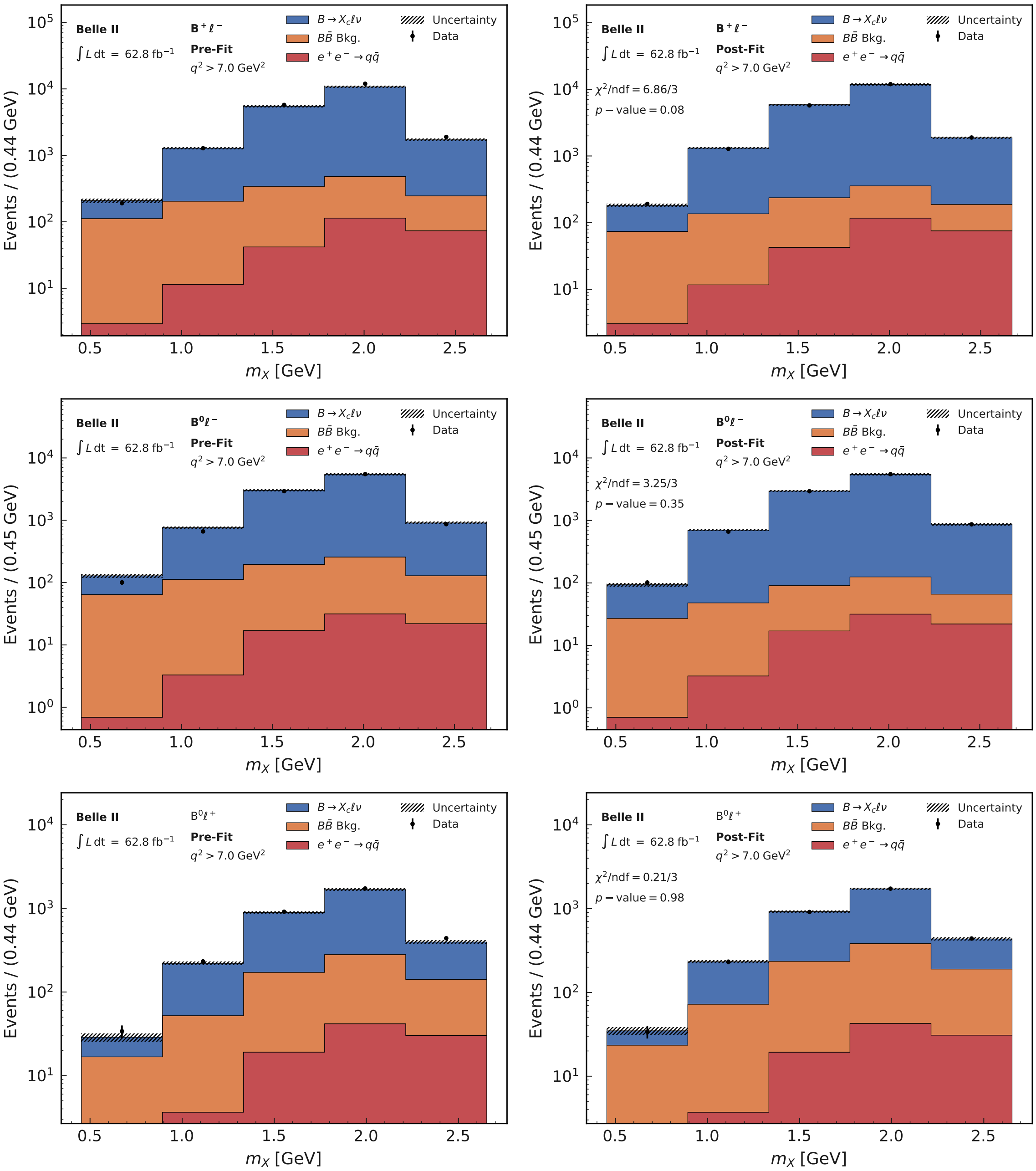


Figure A.14: m_X distribution for a $q^2 > 7.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

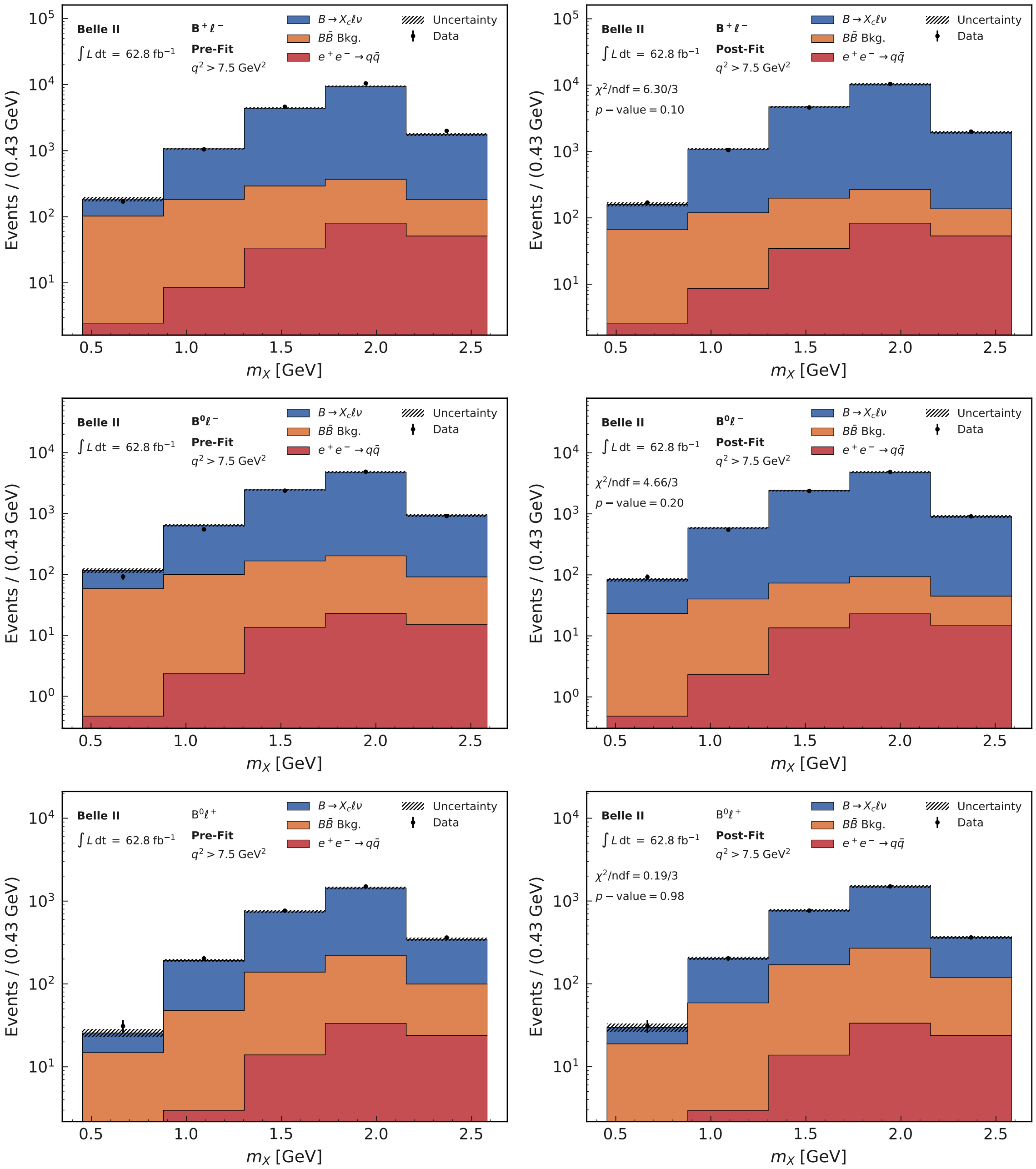


Figure A.15: m_X distribution for a $q^2 > 7.5\,\text{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

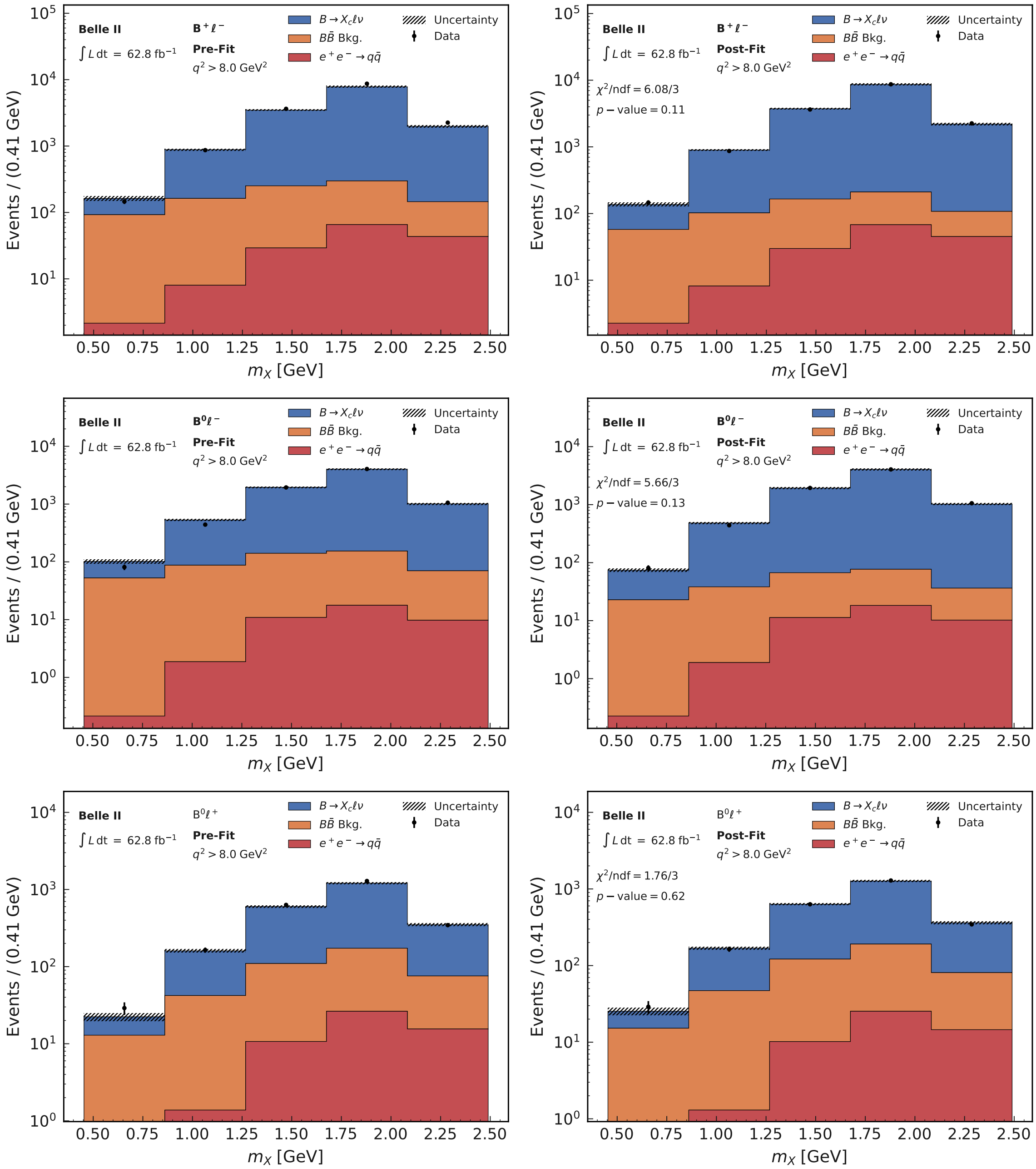


Figure A.16: m_X distribution for a $q^2 > 8.0\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

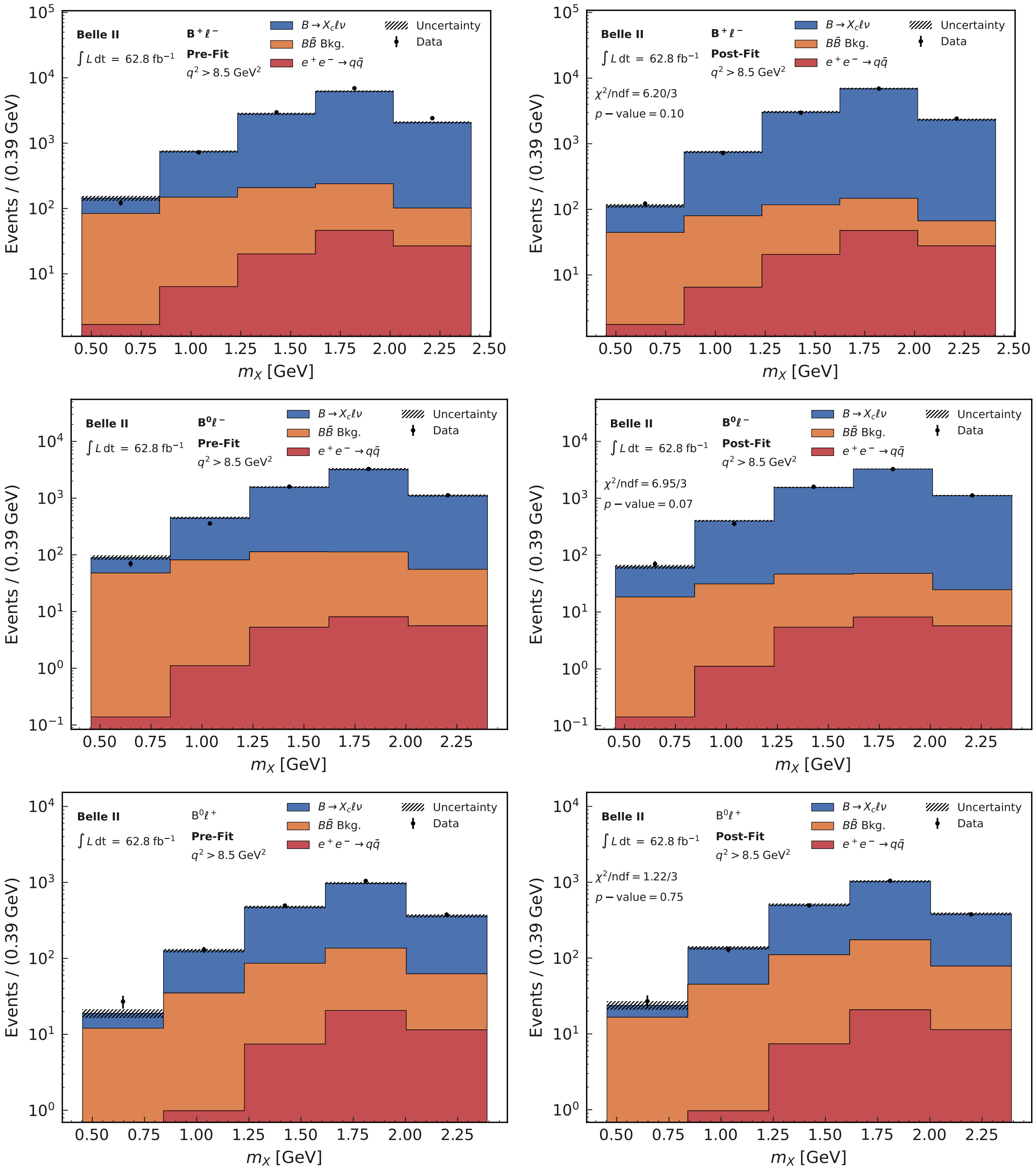


Figure A.17: m_X distribution for a $q^2 > 8.5\,\mathrm{GeV}^2$ before and after the fit in the $B^+\ell^-$ (top row), $B^0\ell^-$ (middle row), and $B^0\ell^+$ (bottom row) reconstruction channels. The measured distribution is shown by the black markers while the expected signal $B \longrightarrow X_c\ell\nu_\ell$ and backgrounds ($B\overline{B}$ and $e^+e^- \longrightarrow q\overline{q}$) are displayed as stacked histograms. The hatched uncertainty bands show the total MC uncertainty before and after the fit.

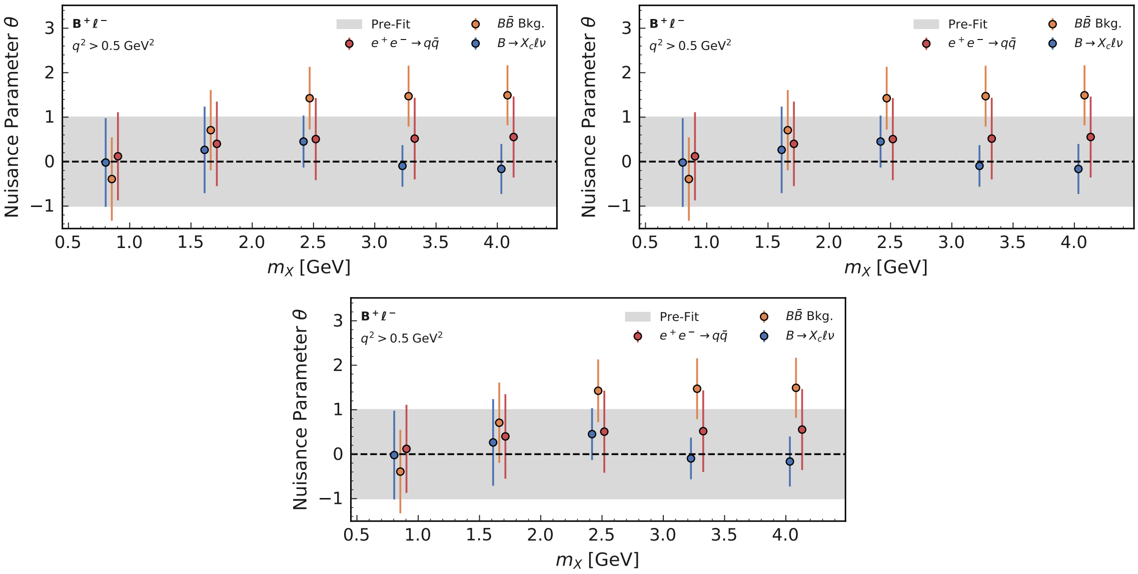


Figure A.18: Bin-wise nuisance parameters after the m_X fit for $q^2 > 0.5\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

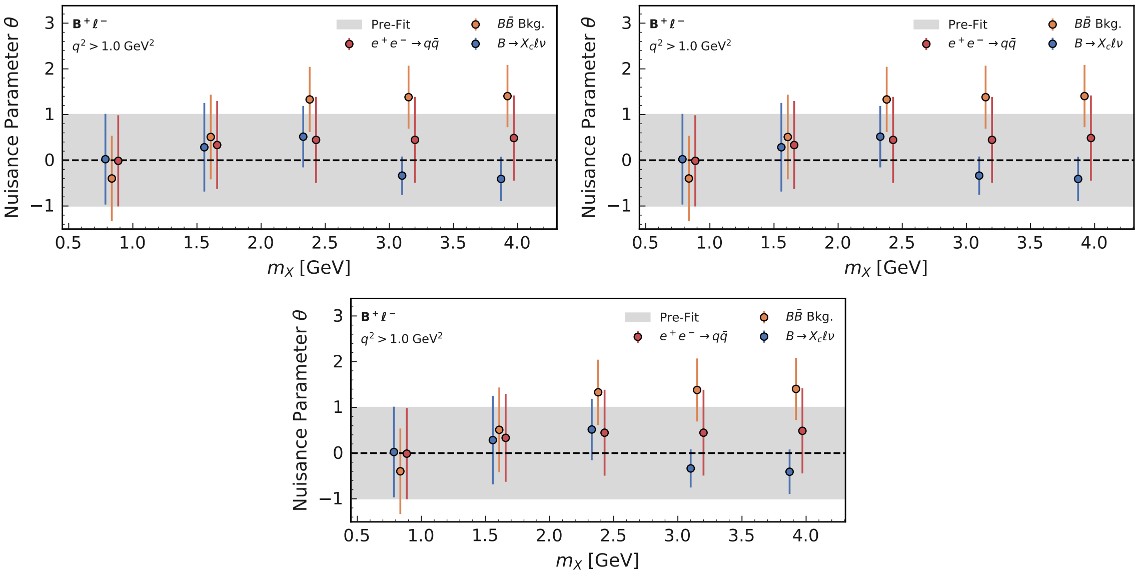


Figure A.19: Bin-wise nuisance parameters after the m_X fit for $q^2 > 1.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

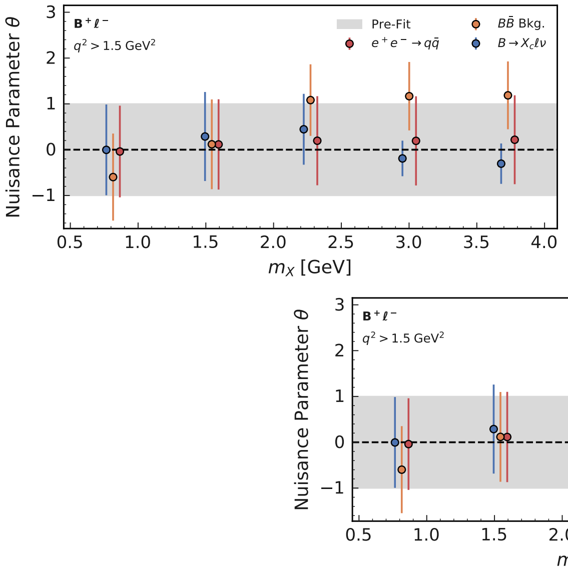


Figure A.20: Bin-wise nuisance parameters after the m_X fit for $q^2 > 1.5\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

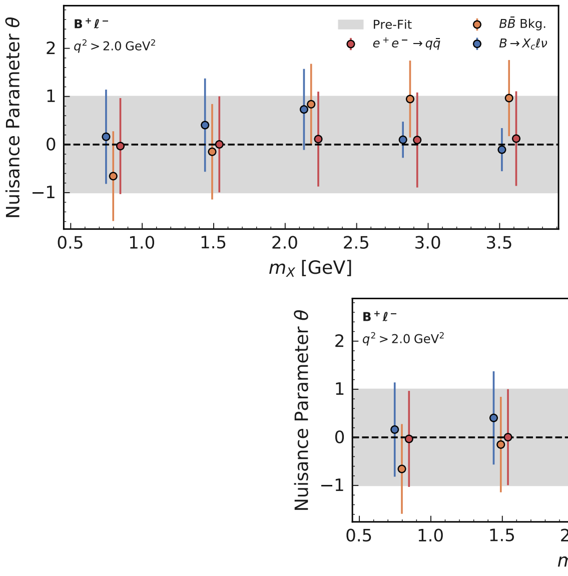


Figure A.21: Bin-wise nuisance parameters after the m_X fit for $q^2 > 2.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

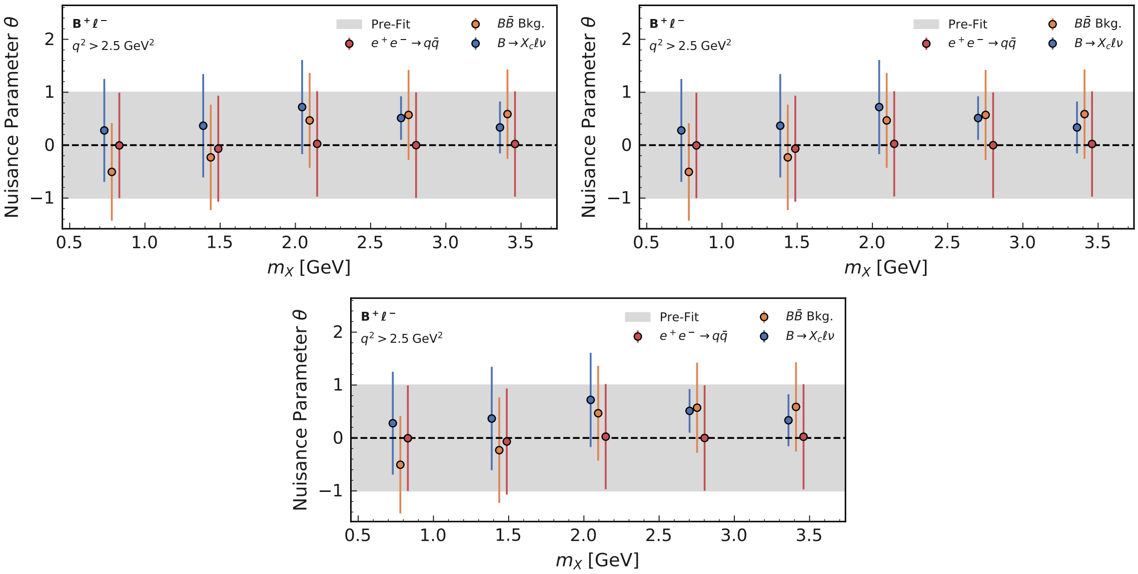


Figure A.22: Bin-wise nuisance parameters after the m_X fit for $q^2 > 2.5\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

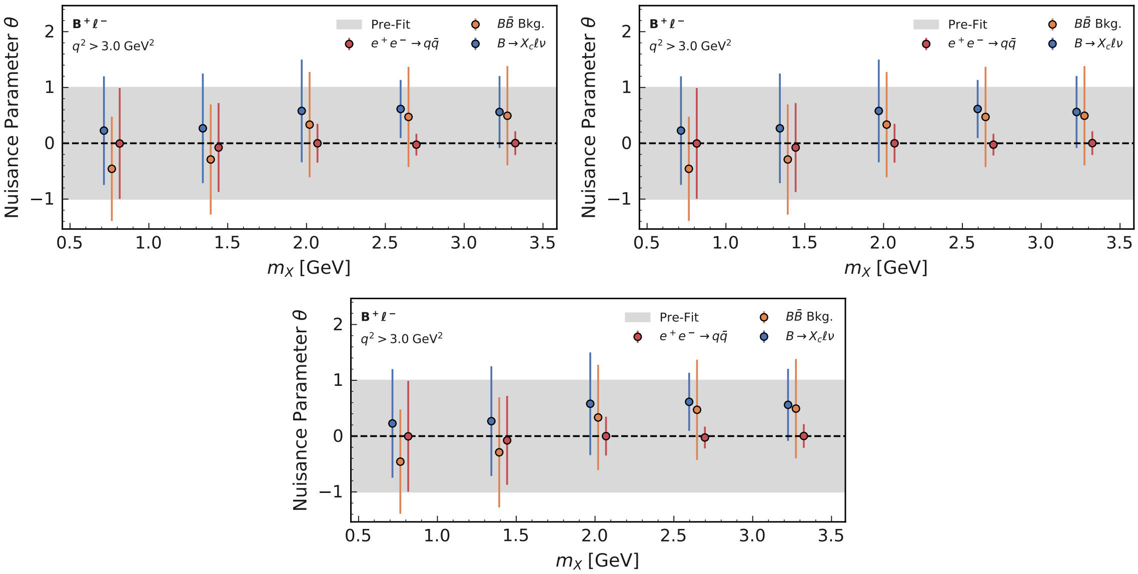


Figure A.23: Bin-wise nuisance parameters after the m_X fit for $q^2 > 3.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

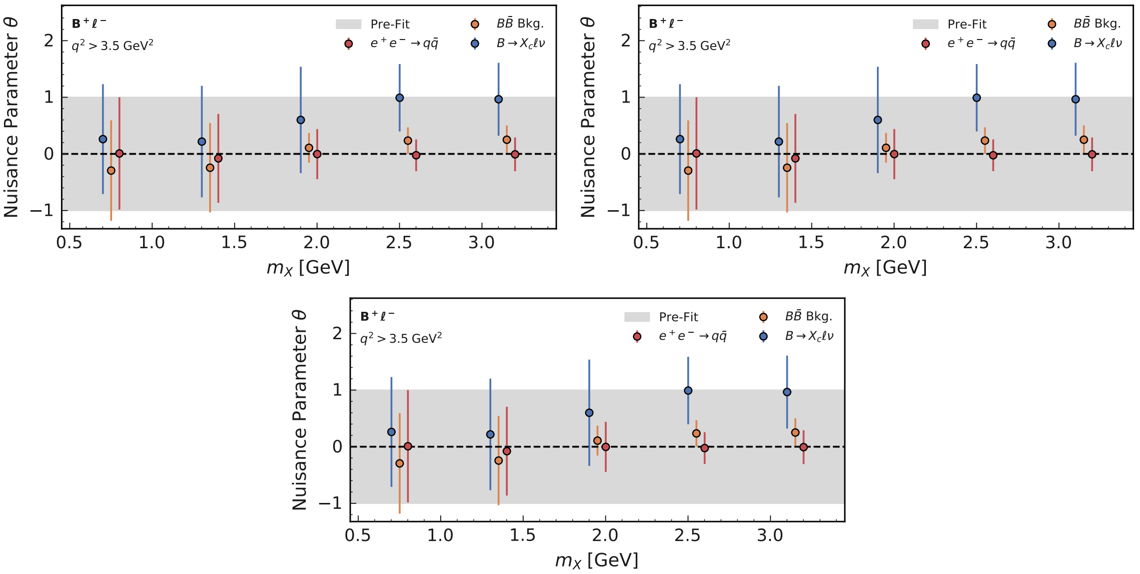


Figure A.24: Bin-wise nuisance parameters after the m_X fit for $q^2 > 3.5\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

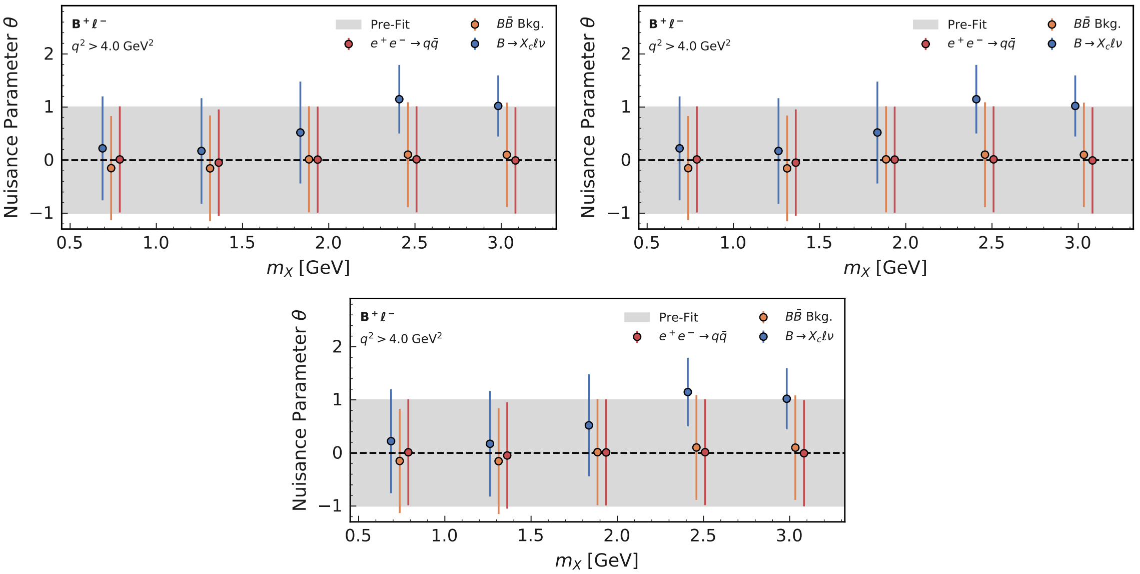


Figure A.25: Bin-wise nuisance parameters after the m_X fit for $q^2 > 4.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

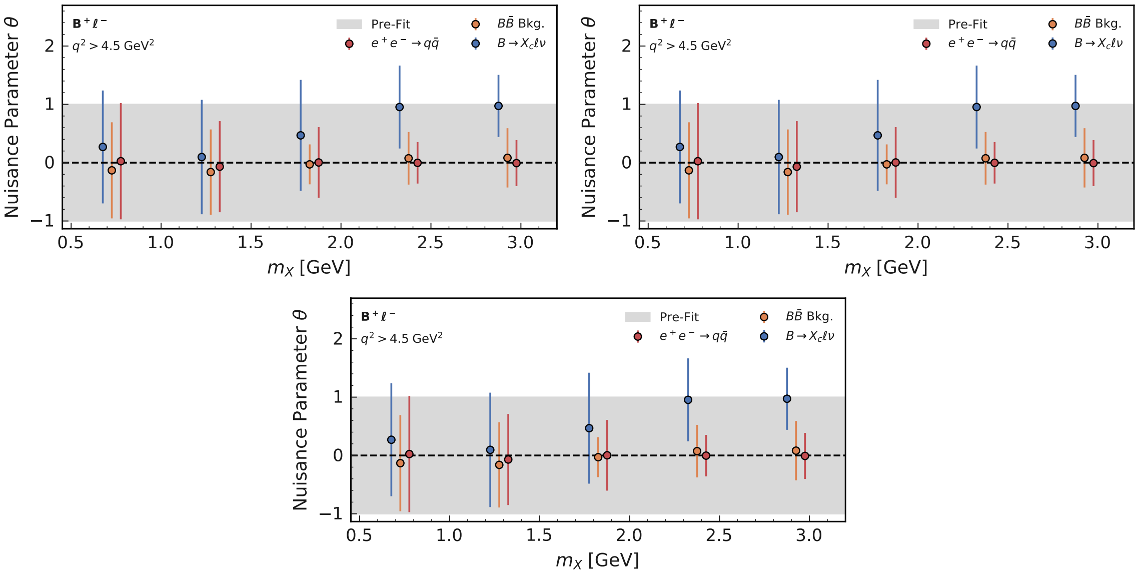


Figure A.26: Bin-wise nuisance parameters after the m_X fit for $q^2 > 4.5\,\text{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

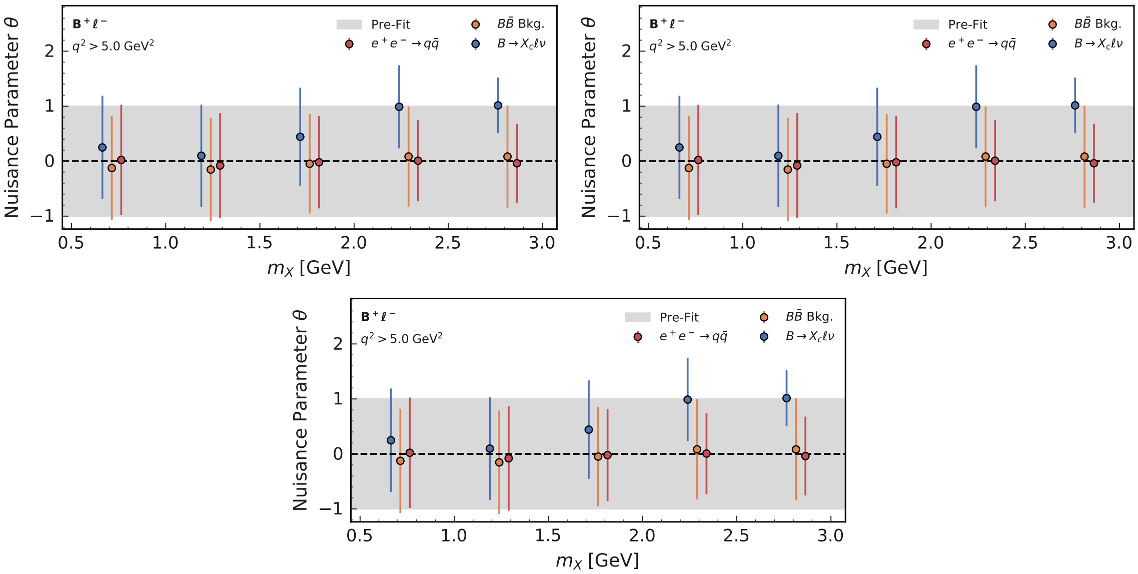


Figure A.27: Bin-wise nuisance parameters after the m_X fit for $q^2 > 5.0\,\text{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

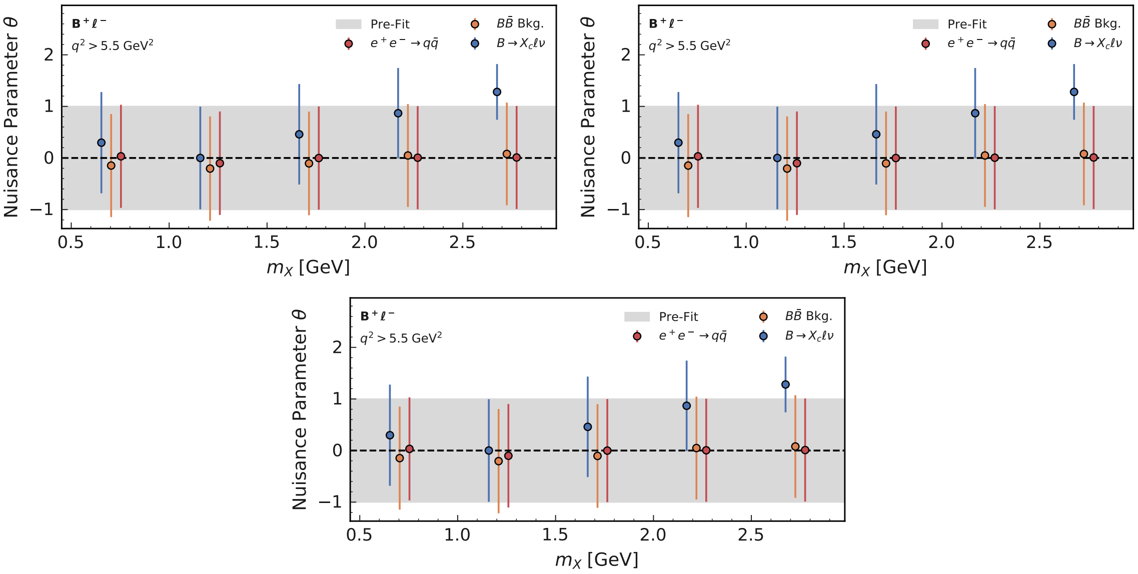


Figure A.28: Bin-wise nuisance parameters after the m_X fit for $q^2 > 5.5\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

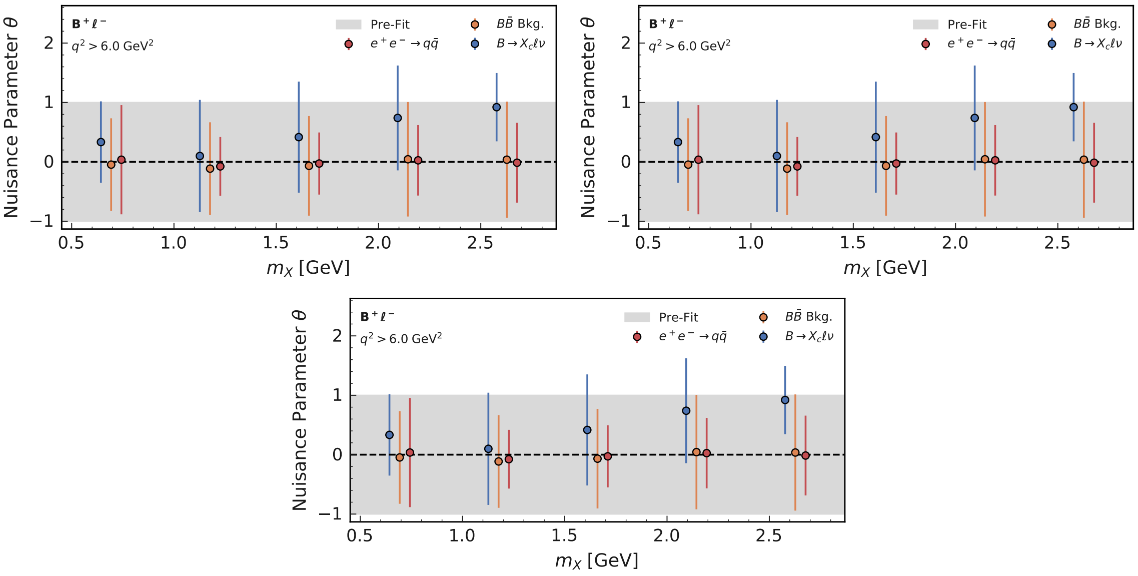


Figure A.29: Bin-wise nuisance parameters after the m_X fit for $q^2 > 6.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

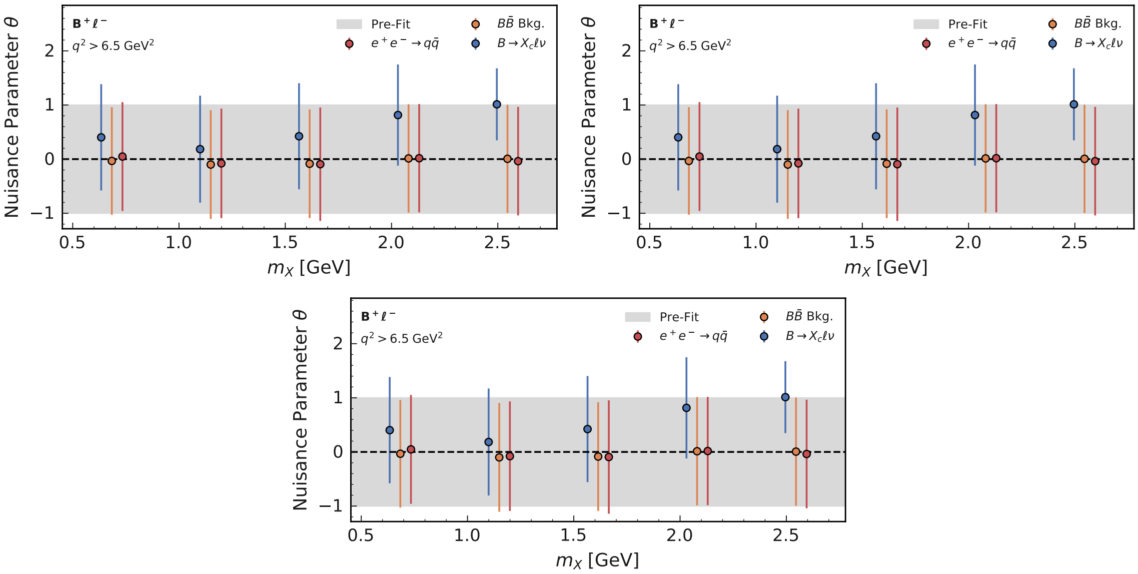


Figure A.30: Bin-wise nuisance parameters after the m_X fit for $q^2 > 6.5\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

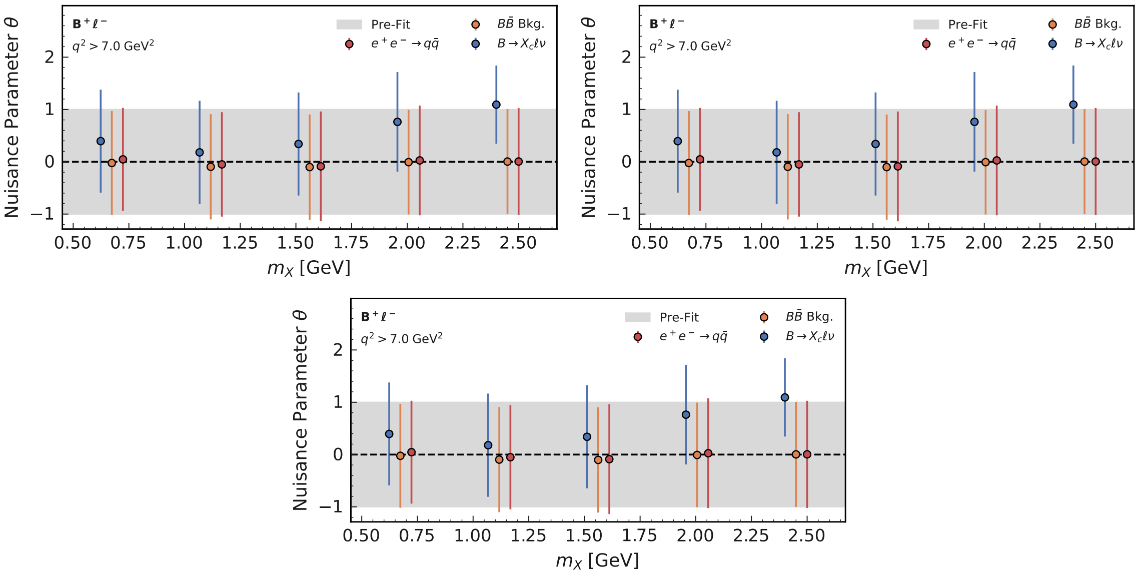


Figure A.31: Bin-wise nuisance parameters after the m_X fit for $q^2 > 7.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

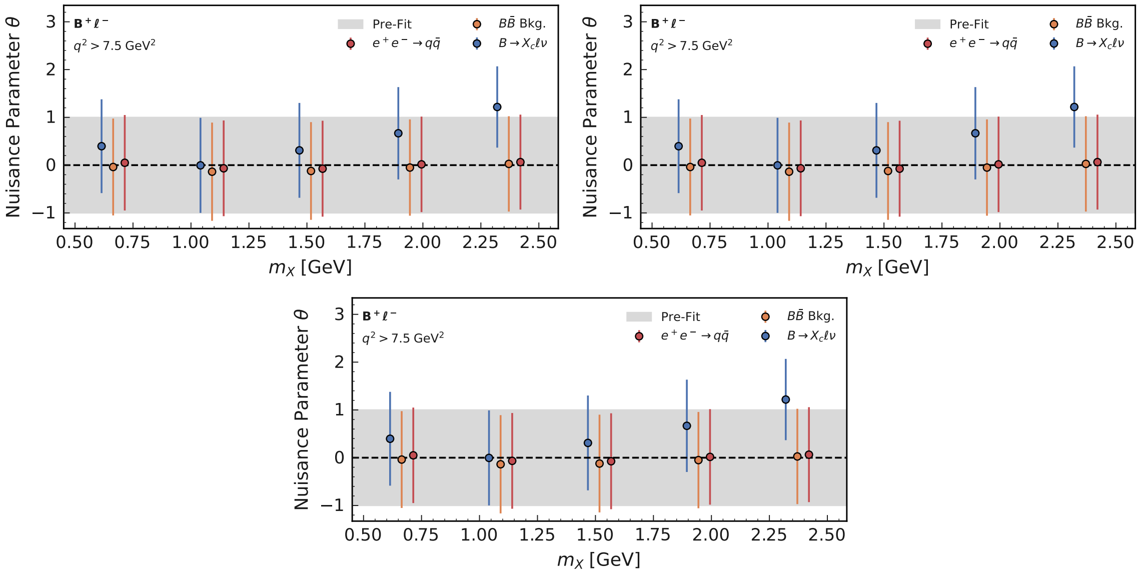


Figure A.32: Bin-wise nuisance parameters after the m_X fit for $q^2 > 7.5\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

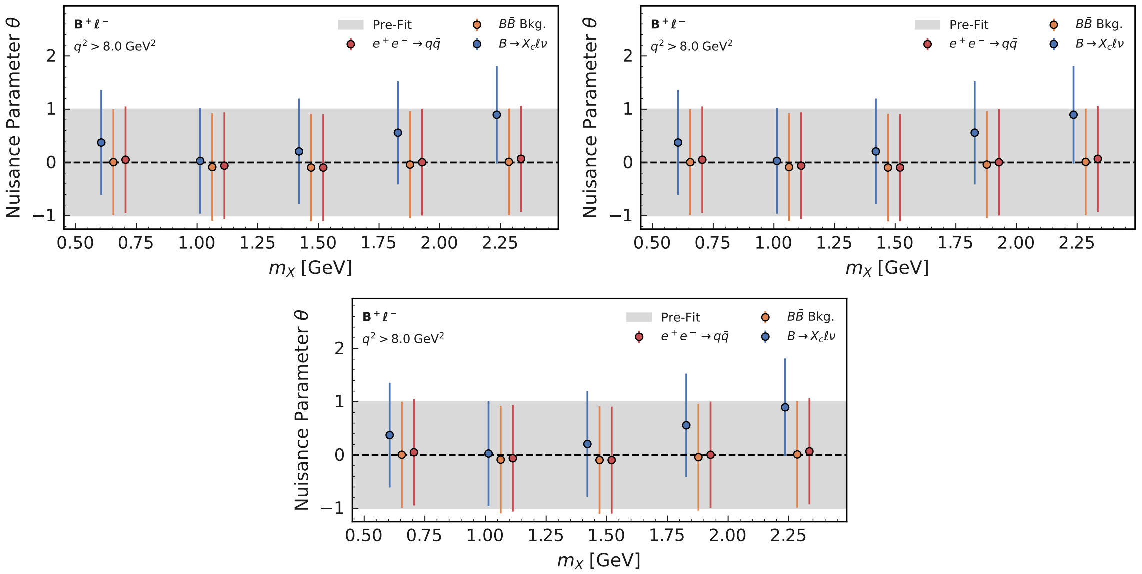


Figure A.33: Bin-wise nuisance parameters after the m_X fit for $q^2 > 8.0\,\mathrm{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

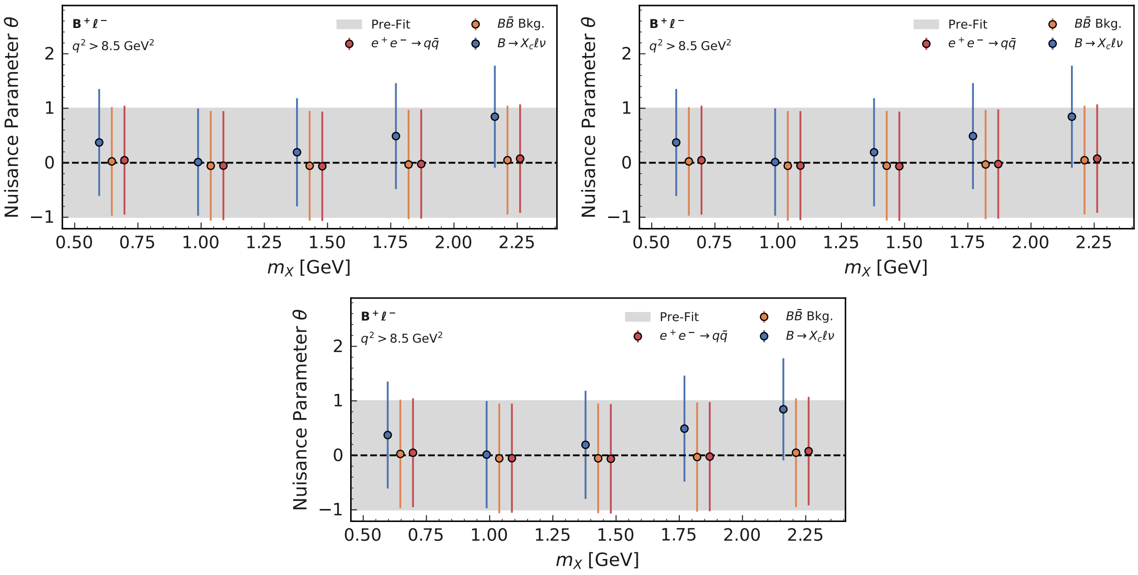


Figure A.34: Bin-wise nuisance parameters after the m_X fit for $q^2 > 8.5\,\text{GeV}^2$. The grey band shows the pre-fit expectation and uncertainty. The markers and error bars show the nuisance parameter value and uncertainty after the fit, respectively.

APPENDIX B

Additional Material q^2 Spectrum Comparison

This appendix contains additional material for the comparison of measured and simulated q^2 spectra for different lower q^2 thresholds starting at 0.5 GeV^2 up to 8.5 GeV^2 in 0.5 GeV^2 increments. Figure B.1 shows the obtained χ^2 values from the binned comparison and the resulting value of the KS test statistic $D_{n,m}$. For a given confidence level α, the null-hypothesis in the KS test is rejected if

$$D_{n,m}(\alpha) > \sqrt{-\log\frac{\alpha}{2} \times \frac{n+m}{2nm}}, \tag{B.1}$$

with n and m denoting the number of events in the two compared samples. Comparisons of the measured and simulated q^2 distributions are shown in Figures B.2 to B.4.

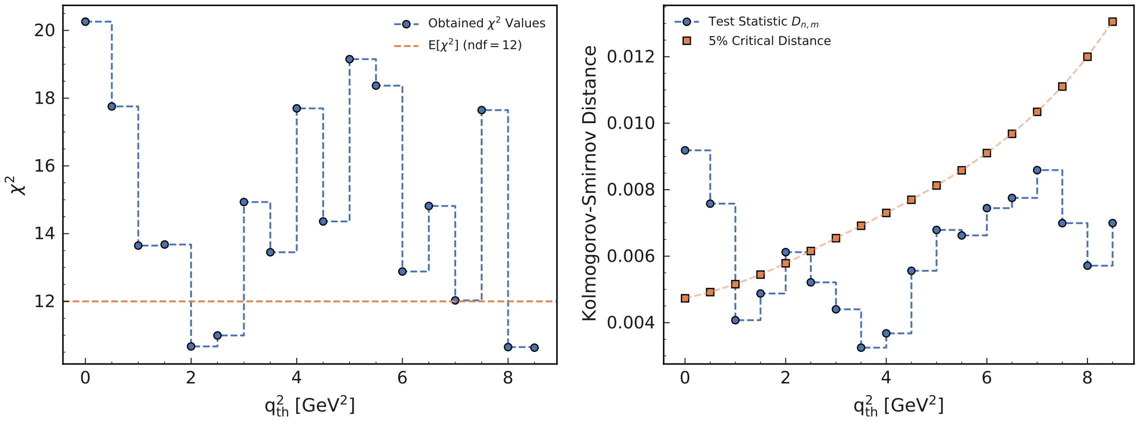


Figure B.1: Observed χ^2 values in the binned comparisons between measured and simulated q^2 distributions (left) and KS distances $D_{n,m}$ resulting from an un-binned comparison (right) as a function of lower q^2 thresholds.

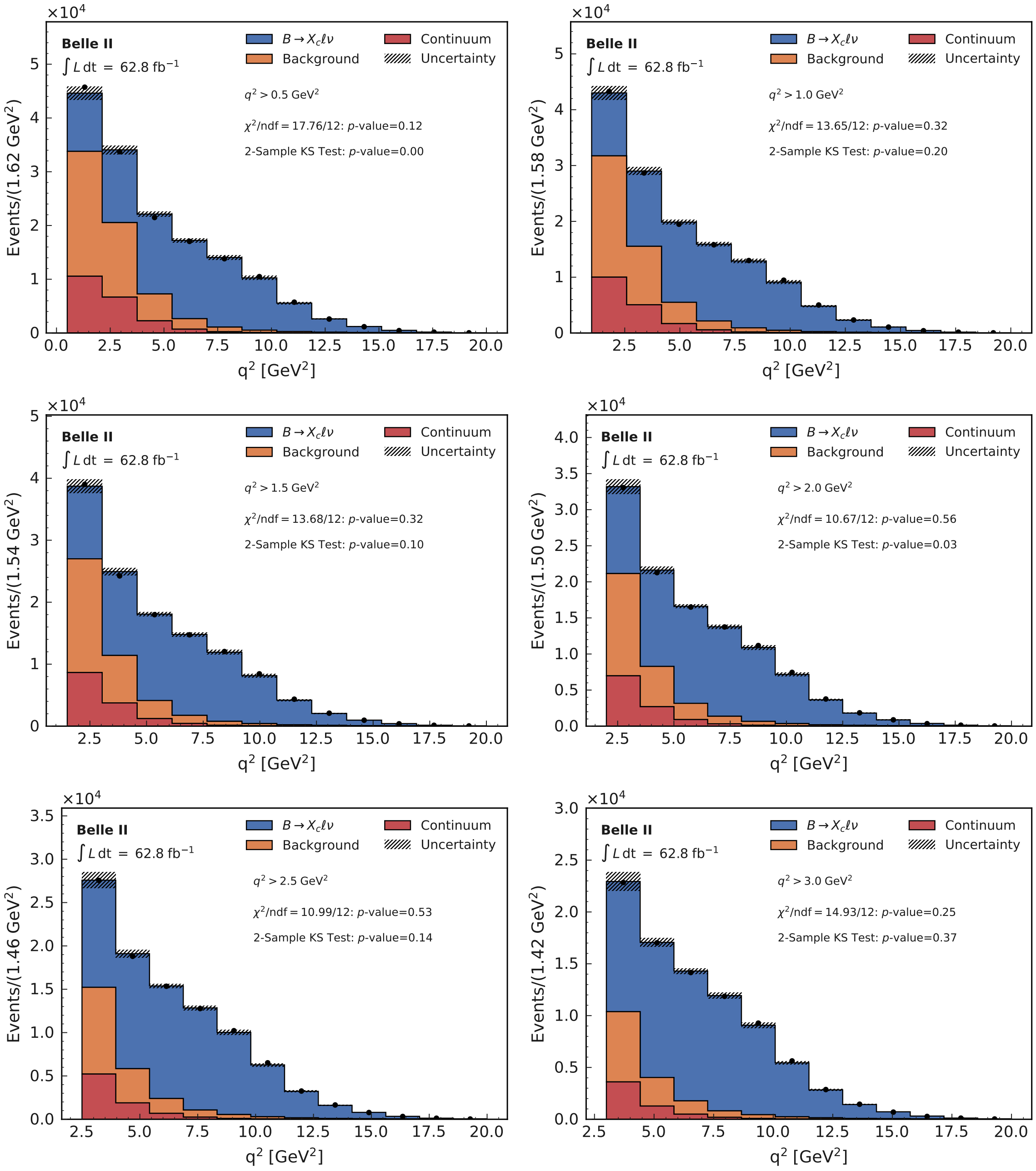


Figure B.2: Comparison of the measured and simulated q^2 distributions for the lower q^2 thresholds between $q^2 > 0.5\,\text{GeV}^2$ and $q^2 > 3.0\,\text{GeV}^2$.

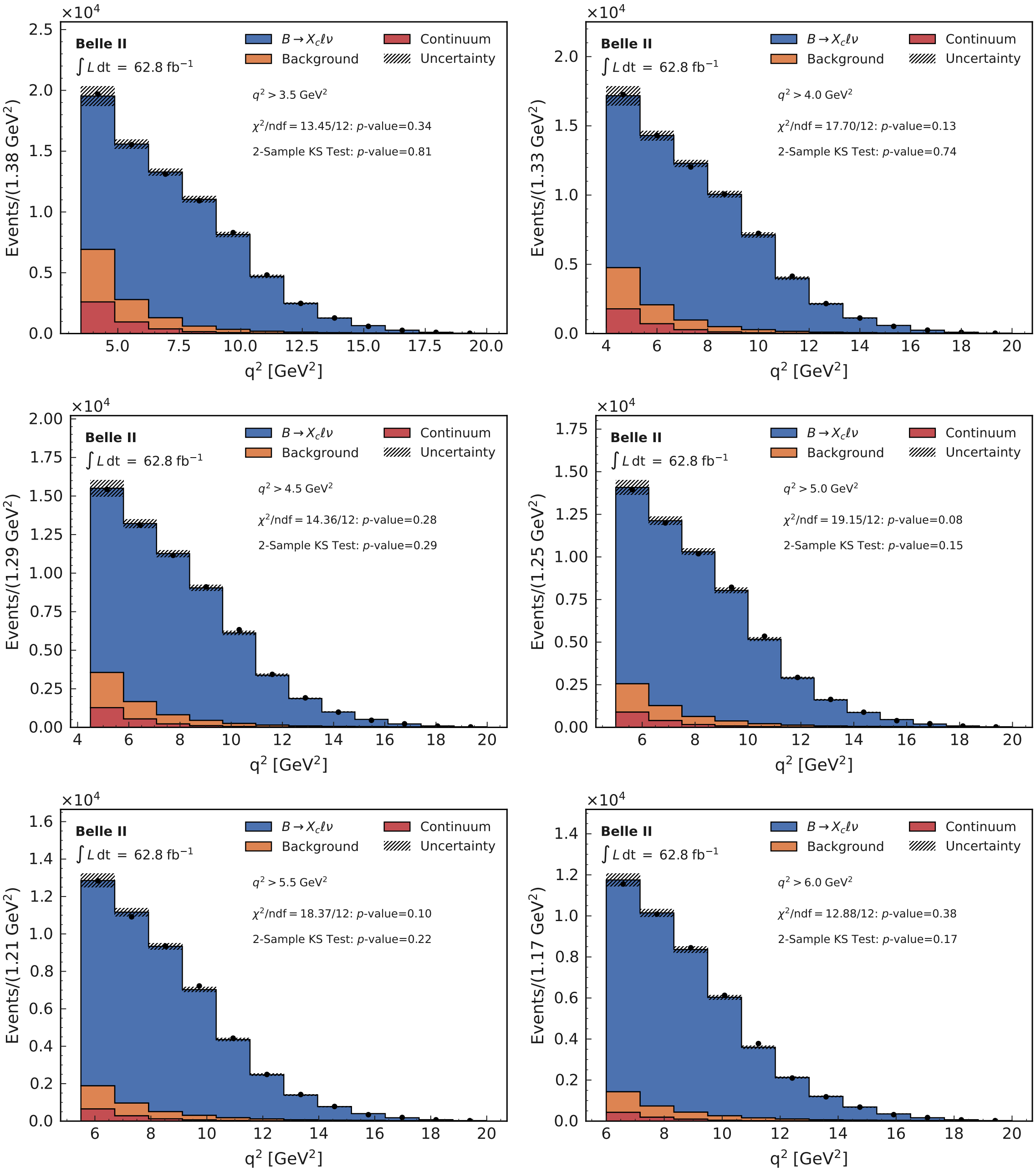


Figure B.3: Comparison of the measured and simulated q^2 distributions for the lower q^2 thresholds between $q^2 > 3.5\,\text{GeV}^2$ and $q^2 > 6.0\,\text{GeV}^2$.

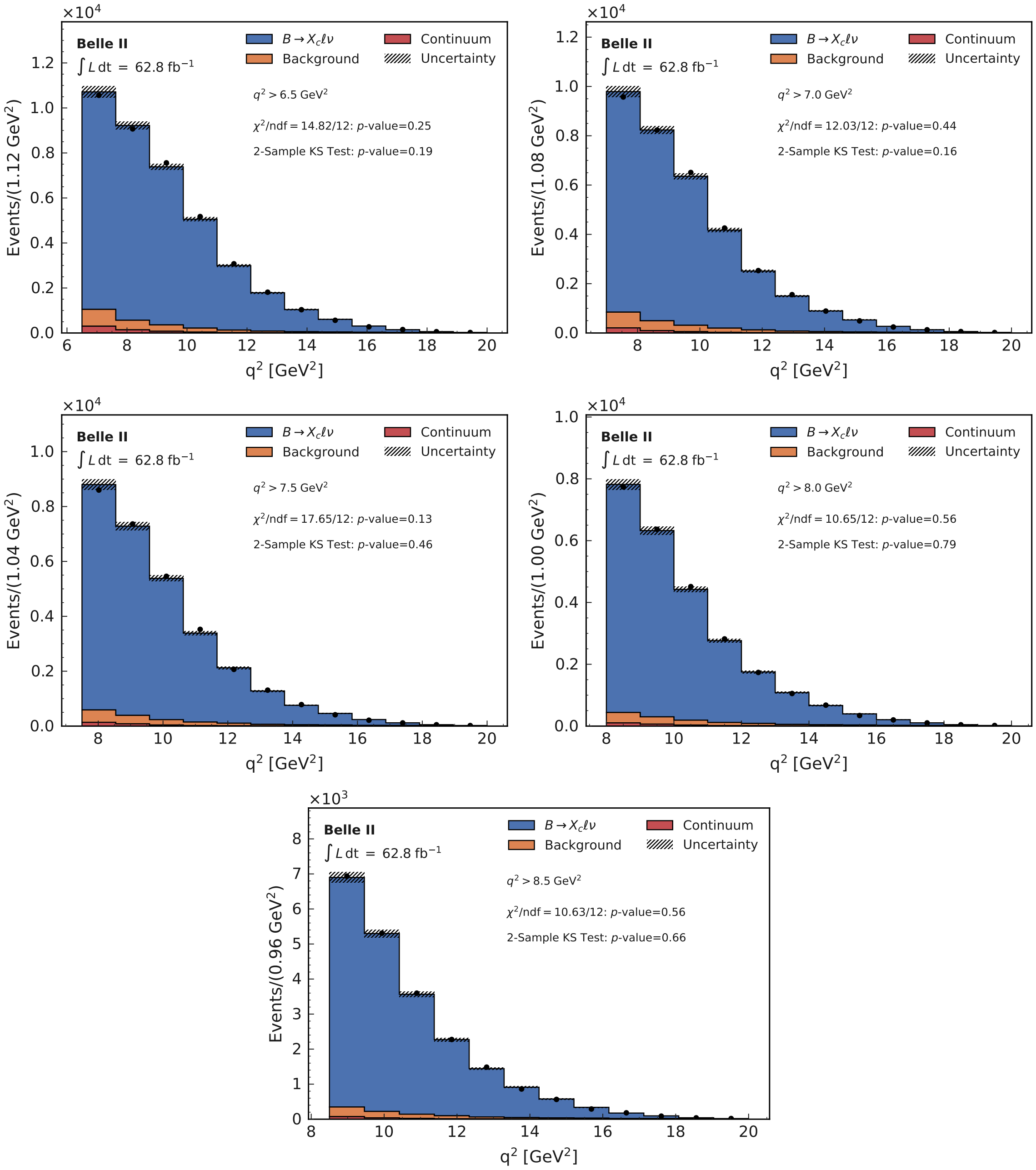


Figure B.4: Comparison of the measured and simulated q^2 distributions for the lower q^2 thresholds between $q^2 > 6.5\,\mathrm{GeV}^2$ and $q^2 > 8.5\,\mathrm{GeV}^2$.

APPENDIX C

Additional Material Results

This appendix presents additional material for the results of the measurement of raw and central q^2 in $B \to X_c \ell \nu_\ell$ decays.

Numerical Values Bias Correction Factors

Tables C.1 to C.4 list the bias correction factors C_{calib}, C_{gen}, and the product $C_{\text{calib}} \times C_{\text{gen}}$ applied in the calculation of the raw q^2 moments. The bias correction factors are determined with an independent signal MC sample that is not used to derive the calibration coefficients c_n and m_n.

Table C.1: Numerical values for the bias correction factors C_{calib}, C_{gen} and $C_{\text{calib}} \times C_{\text{gen}}$ applied in the calculation of the first q^2 moment $\langle q \rangle$ as functions of the lower q^2 threshold.

q^2_{th} [Gev^2]	C_{calib}	C_{gen}	$C_{\text{calib}} \times C_{\text{gen}}$
1.5	0.980 ± 0.001	0.951 ± 0.001	0.932 ± 0.001
2.0	0.989 ± 0.001	0.964 ± 0.001	0.953 ± 0.001
2.5	0.997 ± 0.001	0.973 ± 0.001	0.970 ± 0.001
3.0	1.003 ± 0.001	0.980 ± 0.001	0.983 ± 0.001
3.5	1.007 ± 0.001	0.985 ± 0.001	0.992 ± 0.001
4.0	1.010 ± 0.001	0.989 ± 0.001	0.999 ± 0.001
4.5	1.012 ± 0.001	0.992 ± 0.001	1.004 ± 0.001
5.0	1.012 ± 0.001	0.995 ± 0.001	1.007 ± 0.001
5.5	1.012 ± 0.001	0.997 ± 0.001	1.008 ± 0.001
6.0	1.010 ± 0.001	0.998 ± 0.000	1.008 ± 0.001
6.5	1.007 ± 0.001	0.999 ± 0.000	1.006 ± 0.001
7.0	1.004 ± 0.001	0.999 ± 0.000	1.004 ± 0.001
7.5	1.000 ± 0.001	1.000 ± 0.000	0.999 ± 0.001
8.0	0.994 ± 0.001	1.000 ± 0.000	0.994 ± 0.001
8.5	0.987 ± 0.001	1.000 ± 0.000	0.987 ± 0.001

Table C.2: Numerical values for the bias correction factors C_{calib}, C_{gen} and $C_{\text{calib}} \times C_{\text{gen}}$ applied in the calculation of the first q^2 moment $\langle q^4 \rangle$ as functions of the lower q^2 threshold.

q^2_{th} [Gev^2]	C_{calib}	C_{gen}	$C_{\text{calib}} \times C_{\text{gen}}$
1.5	0.977 ± 0.002	0.922 ± 0.001	0.901 ± 0.002
2.0	0.988 ± 0.002	0.939 ± 0.001	0.928 ± 0.002
2.5	0.998 ± 0.002	0.953 ± 0.001	0.951 ± 0.002
3.0	1.007 ± 0.002	0.964 ± 0.001	0.970 ± 0.002
3.5	1.013 ± 0.002	0.973 ± 0.001	0.986 ± 0.002
4.0	1.017 ± 0.002	0.980 ± 0.001	0.997 ± 0.002
4.5	1.020 ± 0.002	0.986 ± 0.001	1.005 ± 0.002
5.0	1.021 ± 0.002	0.990 ± 0.001	1.011 ± 0.002
5.5	1.021 ± 0.002	0.994 ± 0.001	1.014 ± 0.001
6.0	1.018 ± 0.002	0.996 ± 0.001	1.014 ± 0.001
6.5	1.014 ± 0.002	0.998 ± 0.001	1.012 ± 0.001
7.0	1.009 ± 0.002	0.998 ± 0.001	1.007 ± 0.001
7.5	1.001 ± 0.001	0.999 ± 0.001	1.000 ± 0.001
8.0	0.991 ± 0.001	1.000 ± 0.001	0.990 ± 0.001
8.5	0.978 ± 0.001	1.000 ± 0.001	0.978 ± 0.001

Table C.3: Numerical values for the bias correction factors C_{calib}, C_{gen} and $C_{\text{calib}} \times C_{\text{gen}}$ applied in the calculation of the first q^2 moment $\langle q^6 \rangle$ as functions of the lower q^2 threshold.

q^2_{th} [Gev^2]	C_{calib}	C_{gen}	$C_{\text{calib}} \times C_{\text{gen}}$
1.5	0.985 ± 0.003	0.907 ± 0.002	0.893 ± 0.003
2.0	0.994 ± 0.003	0.925 ± 0.002	0.920 ± 0.003
2.5	1.003 ± 0.003	0.940 ± 0.002	0.943 ± 0.003
3.0	1.011 ± 0.003	0.953 ± 0.002	0.963 ± 0.003
3.5	1.017 ± 0.003	0.964 ± 0.002	0.980 ± 0.003
4.0	1.021 ± 0.003	0.973 ± 0.002	0.994 ± 0.003
4.5	1.025 ± 0.003	0.980 ± 0.002	1.004 ± 0.003
5.0	1.027 ± 0.003	0.986 ± 0.002	1.013 ± 0.003
5.5	1.026 ± 0.003	0.991 ± 0.001	1.017 ± 0.003
6.0	1.024 ± 0.003	0.994 ± 0.001	1.018 ± 0.002
6.5	1.020 ± 0.003	0.997 ± 0.001	1.016 ± 0.002
7.0	1.014 ± 0.003	0.998 ± 0.001	1.012 ± 0.002
7.5	1.005 ± 0.002	0.999 ± 0.001	1.004 ± 0.002
8.0	0.992 ± 0.002	1.000 ± 0.001	0.991 ± 0.002
8.5	0.977 ± 0.002	1.000 ± 0.001	0.976 ± 0.002

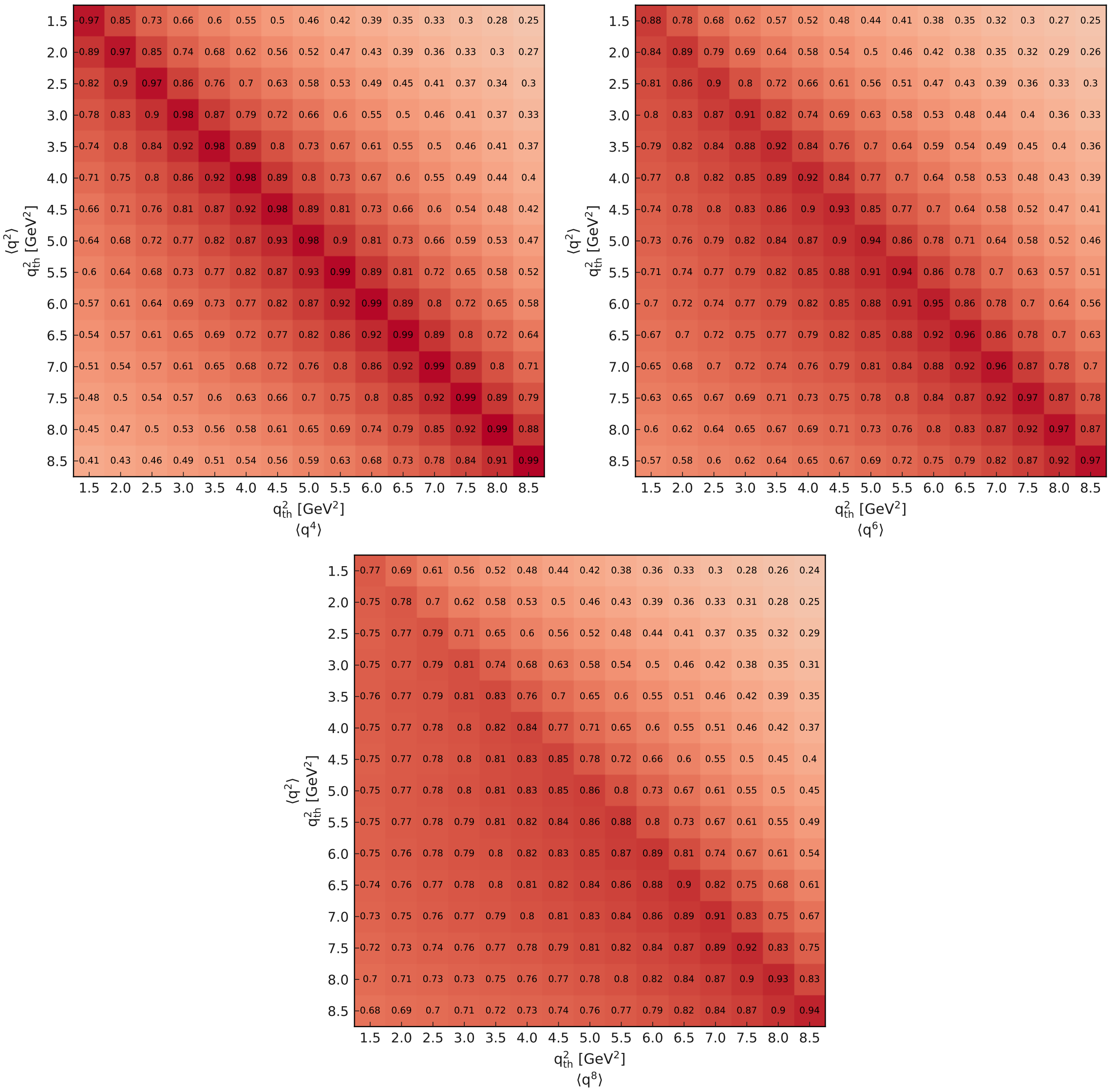
Table C.4: Numerical values for the bias correction factors C_{calib}, C_{gen} and $C_{\text{calib}} \times C_{\text{gen}}$ applied in the calculation of the first q^2 moment $\langle q^8 \rangle$ as functions of the lower q^2 threshold.

q^2_{th} [Gev^2]	C_{calib}	C_{gen}	$C_{\text{calib}} \times C_{\text{gen}}$
1.5	0.999 ± 0.005	0.898 ± 0.002	0.897 ± 0.004
2.0	1.005 ± 0.005	0.916 ± 0.002	0.921 ± 0.004
2.5	1.011 ± 0.005	0.932 ± 0.002	0.942 ± 0.004
3.0	1.016 ± 0.005	0.946 ± 0.002	0.961 ± 0.004
3.5	1.020 ± 0.005	0.958 ± 0.002	0.977 ± 0.004
4.0	1.023 ± 0.005	0.968 ± 0.002	0.990 ± 0.004
4.5	1.026 ± 0.004	0.976 ± 0.002	1.001 ± 0.004
5.0	1.028 ± 0.004	0.983 ± 0.002	1.011 ± 0.004
5.5	1.029 ± 0.004	0.989 ± 0.002	1.017 ± 0.004
6.0	1.027 ± 0.004	0.992 ± 0.002	1.019 ± 0.004
6.5	1.024 ± 0.004	0.996 ± 0.002	1.019 ± 0.004
7.0	1.020 ± 0.004	0.997 ± 0.002	1.017 ± 0.003
7.5	1.011 ± 0.004	0.999 ± 0.002	1.009 ± 0.003
8.0	0.998 ± 0.004	1.000 ± 0.001	0.997 ± 0.003
8.5	0.982 ± 0.003	1.000 ± 0.001	0.982 ± 0.003

Numerical Values of the Correlation Matrices

Raw Moments (Statistical Correlation)

Figures C.1 to C.4 summarize the numerical values for the estimated statistical correlation between the raw q^2 moments of different order and lower q^2 threshold. The full correlation structure is shown in Figure 5.19.

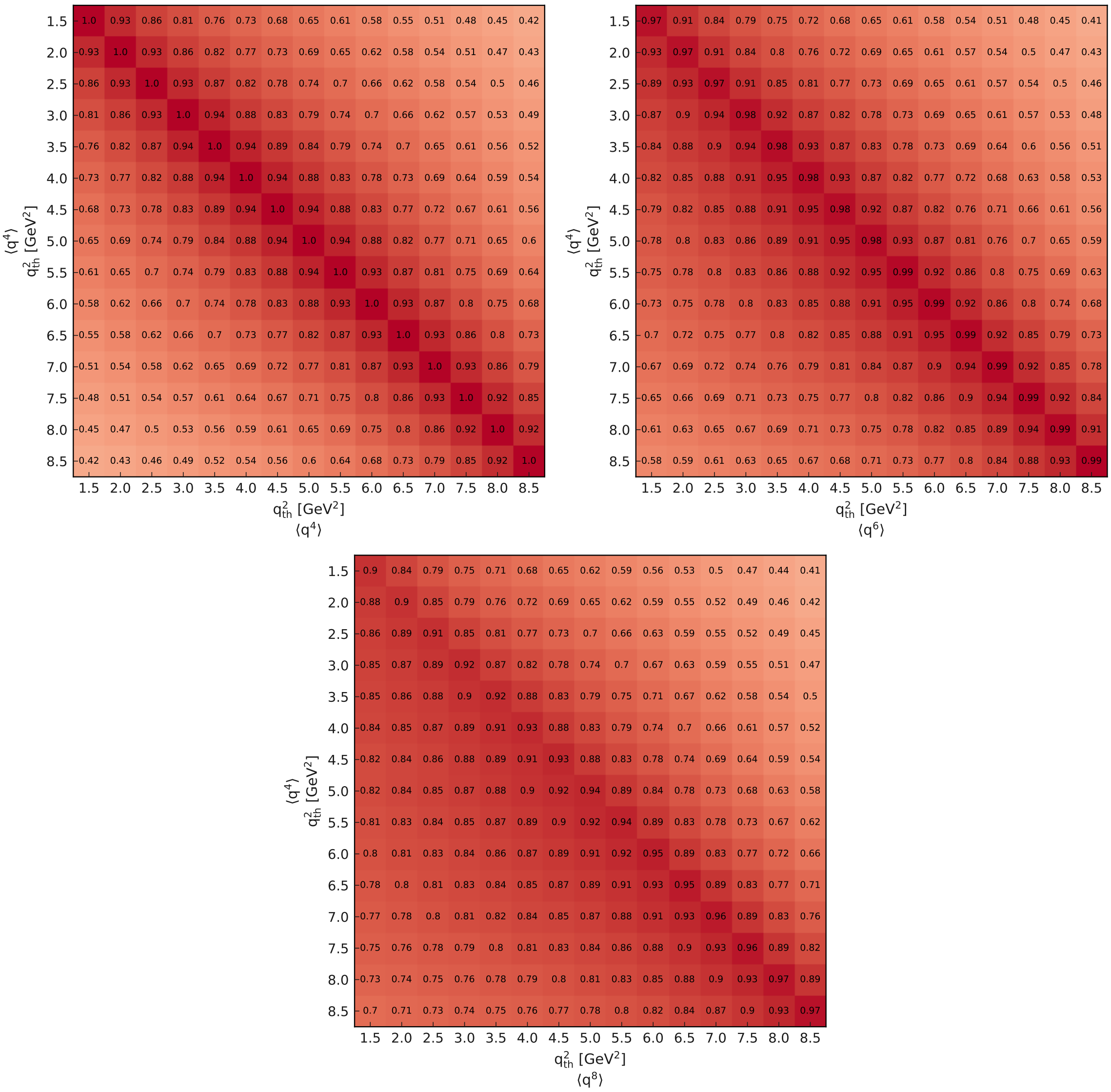


	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
1.5	0.97	0.85	0.73	0.66	0.6	0.55	0.5	0.46	0.42	0.39	0.35	0.33	0.3	0.28	0.25
2.0	0.89	0.97	0.85	0.74	0.68	0.62	0.56	0.52	0.47	0.43	0.39	0.36	0.33	0.3	0.27
2.5	0.82	0.9	0.97	0.86	0.76	0.7	0.63	0.58	0.53	0.49	0.45	0.41	0.37	0.34	0.3
3.0	0.78	0.83	0.9	0.98	0.87	0.79	0.72	0.66	0.6	0.55	0.5	0.46	0.41	0.37	0.33
3.5	0.74	0.8	0.84	0.92	0.98	0.89	0.8	0.73	0.67	0.61	0.55	0.5	0.46	0.41	0.37
4.0	0.71	0.75	0.8	0.86	0.92	0.98	0.89	0.8	0.73	0.67	0.6	0.55	0.49	0.44	0.4
4.5	0.66	0.71	0.76	0.81	0.87	0.92	0.98	0.89	0.81	0.73	0.66	0.6	0.54	0.48	0.42
5.0	0.64	0.68	0.72	0.77	0.82	0.87	0.93	0.98	0.9	0.81	0.73	0.66	0.59	0.53	0.47
5.5	0.6	0.64	0.68	0.73	0.77	0.82	0.87	0.93	0.99	0.89	0.81	0.72	0.65	0.58	0.52
6.0	0.57	0.61	0.64	0.69	0.73	0.77	0.82	0.87	0.92	0.99	0.89	0.8	0.72	0.65	0.58
6.5	0.54	0.57	0.61	0.65	0.69	0.72	0.77	0.82	0.86	0.92	0.99	0.89	0.8	0.72	0.64
7.0	0.51	0.54	0.57	0.61	0.65	0.68	0.72	0.76	0.8	0.86	0.92	0.99	0.89	0.8	0.71
7.5	0.48	0.5	0.54	0.57	0.6	0.63	0.66	0.7	0.75	0.8	0.85	0.92	0.99	0.89	0.79
8.0	0.45	0.47	0.5	0.53	0.56	0.58	0.61	0.65	0.69	0.74	0.79	0.85	0.92	0.99	0.88
8.5	0.41	0.43	0.46	0.49	0.51	0.54	0.56	0.59	0.63	0.68	0.73	0.78	0.84	0.91	0.99

	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
1.5	0.88	0.78	0.68	0.62	0.57	0.52	0.48	0.44	0.41	0.38	0.35	0.32	0.3	0.27	0.25
2.0	0.84	0.89	0.79	0.69	0.64	0.58	0.54	0.5	0.46	0.42	0.38	0.35	0.32	0.29	0.26
2.5	0.81	0.86	0.9	0.8	0.72	0.66	0.61	0.56	0.51	0.47	0.43	0.39	0.36	0.33	0.3
3.0	0.8	0.83	0.87	0.91	0.82	0.74	0.69	0.63	0.58	0.53	0.48	0.44	0.4	0.36	0.33
3.5	0.79	0.82	0.84	0.88	0.92	0.84	0.76	0.7	0.64	0.59	0.54	0.49	0.45	0.4	0.36
4.0	0.77	0.8	0.82	0.85	0.89	0.92	0.84	0.77	0.7	0.64	0.58	0.53	0.48	0.43	0.39
4.5	0.74	0.78	0.8	0.83	0.86	0.9	0.93	0.85	0.77	0.7	0.64	0.58	0.52	0.47	0.41
5.0	0.73	0.76	0.79	0.82	0.84	0.87	0.9	0.94	0.86	0.78	0.71	0.64	0.58	0.52	0.46
5.5	0.71	0.74	0.77	0.79	0.82	0.85	0.88	0.91	0.94	0.86	0.78	0.7	0.63	0.57	0.51
6.0	0.7	0.72	0.74	0.77	0.79	0.82	0.85	0.88	0.91	0.95	0.86	0.78	0.7	0.64	0.56
6.5	0.67	0.7	0.72	0.75	0.77	0.79	0.82	0.85	0.88	0.92	0.96	0.86	0.78	0.7	0.63
7.0	0.65	0.68	0.7	0.72	0.74	0.76	0.79	0.81	0.84	0.88	0.92	0.96	0.87	0.78	0.7
7.5	0.63	0.65	0.67	0.69	0.71	0.73	0.75	0.78	0.8	0.84	0.87	0.92	0.97	0.87	0.78
8.0	0.6	0.62	0.64	0.65	0.67	0.69	0.71	0.73	0.76	0.8	0.83	0.87	0.92	0.97	0.87
8.5	0.57	0.58	0.6	0.62	0.64	0.65	0.67	0.69	0.72	0.75	0.79	0.82	0.87	0.92	0.97

	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
1.5	0.77	0.69	0.61	0.56	0.52	0.48	0.44	0.42	0.38	0.36	0.33	0.3	0.28	0.26	0.24
2.0	0.75	0.78	0.7	0.62	0.58	0.53	0.5	0.46	0.43	0.39	0.36	0.33	0.31	0.28	0.25
2.5	0.75	0.77	0.79	0.71	0.65	0.6	0.56	0.52	0.48	0.44	0.41	0.37	0.35	0.32	0.29
3.0	0.75	0.77	0.79	0.81	0.74	0.68	0.63	0.58	0.54	0.5	0.46	0.42	0.38	0.35	0.31
3.5	0.76	0.77	0.79	0.81	0.83	0.76	0.7	0.65	0.6	0.55	0.51	0.46	0.42	0.39	0.35
4.0	0.75	0.77	0.78	0.8	0.82	0.84	0.77	0.71	0.65	0.6	0.55	0.51	0.46	0.42	0.37
4.5	0.75	0.77	0.78	0.8	0.81	0.83	0.85	0.78	0.72	0.66	0.6	0.55	0.5	0.45	0.4
5.0	0.75	0.77	0.78	0.8	0.81	0.83	0.85	0.86	0.8	0.73	0.67	0.61	0.55	0.5	0.45
5.5	0.75	0.77	0.78	0.79	0.81	0.82	0.84	0.86	0.88	0.8	0.73	0.67	0.61	0.55	0.49
6.0	0.75	0.76	0.78	0.79	0.8	0.82	0.83	0.85	0.87	0.89	0.81	0.74	0.67	0.61	0.54
6.5	0.74	0.76	0.77	0.78	0.8	0.81	0.82	0.84	0.86	0.88	0.9	0.82	0.75	0.68	0.61
7.0	0.73	0.75	0.76	0.77	0.79	0.8	0.81	0.83	0.84	0.86	0.89	0.91	0.83	0.75	0.67
7.5	0.72	0.73	0.74	0.76	0.77	0.78	0.79	0.81	0.82	0.84	0.87	0.89	0.92	0.83	0.75
8.0	0.7	0.71	0.73	0.73	0.75	0.76	0.77	0.78	0.8	0.82	0.84	0.87	0.9	0.93	0.83
8.5	0.68	0.69	0.7	0.71	0.72	0.73	0.74	0.76	0.77	0.79	0.82	0.84	0.87	0.9	0.94

Figure C.1: Statistical correlations between $\langle q^2 \rangle$ and $\langle q^{2n} \rangle$ for $n = 2$–4.



$\langle q^4 \rangle$

	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
1.5	1.0	0.93	0.86	0.81	0.76	0.73	0.68	0.65	0.61	0.58	0.55	0.51	0.48	0.45	0.42
2.0	0.93	1.0	0.93	0.86	0.82	0.77	0.73	0.69	0.65	0.62	0.58	0.54	0.51	0.47	0.43
2.5	0.86	0.93	1.0	0.93	0.87	0.82	0.78	0.74	0.7	0.66	0.62	0.58	0.54	0.5	0.46
3.0	0.81	0.86	0.93	1.0	0.94	0.88	0.83	0.79	0.74	0.7	0.66	0.62	0.57	0.53	0.49
3.5	0.76	0.82	0.87	0.94	1.0	0.94	0.89	0.84	0.79	0.74	0.7	0.65	0.61	0.56	0.52
4.0	0.73	0.77	0.82	0.88	0.94	1.0	0.94	0.88	0.83	0.78	0.73	0.69	0.64	0.59	0.54
4.5	0.68	0.73	0.78	0.83	0.89	0.94	1.0	0.94	0.88	0.83	0.77	0.72	0.67	0.61	0.56
5.0	0.65	0.69	0.74	0.79	0.84	0.88	0.94	1.0	0.94	0.88	0.82	0.77	0.71	0.65	0.6
5.5	0.61	0.65	0.7	0.74	0.79	0.83	0.88	0.94	1.0	0.93	0.87	0.81	0.75	0.69	0.64
6.0	0.58	0.62	0.66	0.7	0.74	0.78	0.83	0.88	0.93	1.0	0.93	0.87	0.8	0.75	0.68
6.5	0.55	0.58	0.62	0.66	0.7	0.73	0.77	0.82	0.87	0.93	1.0	0.93	0.86	0.8	0.73
7.0	0.51	0.54	0.58	0.62	0.65	0.69	0.72	0.77	0.81	0.87	0.93	1.0	0.93	0.86	0.79
7.5	0.48	0.51	0.54	0.57	0.61	0.64	0.67	0.71	0.75	0.8	0.86	0.93	1.0	0.92	0.85
8.0	0.45	0.47	0.5	0.53	0.56	0.59	0.61	0.65	0.69	0.75	0.8	0.86	0.92	1.0	0.92
8.5	0.42	0.43	0.46	0.49	0.52	0.54	0.56	0.6	0.64	0.68	0.73	0.79	0.85	0.92	1.0

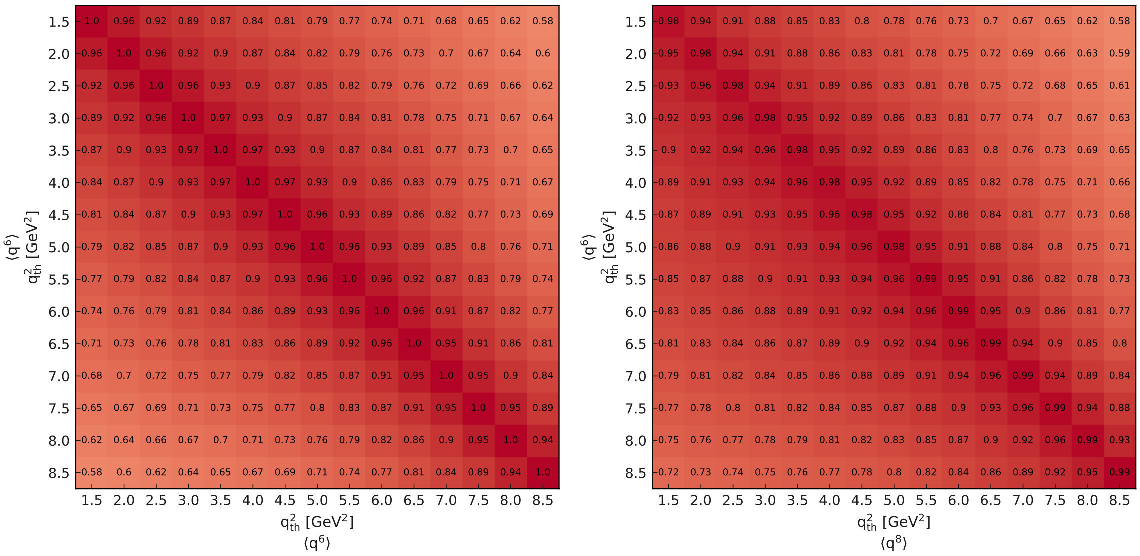
$\langle q^6 \rangle$

	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
1.5	0.97	0.91	0.84	0.79	0.75	0.72	0.68	0.65	0.61	0.58	0.54	0.51	0.48	0.45	0.41
2.0	0.93	0.97	0.91	0.84	0.8	0.76	0.72	0.69	0.65	0.61	0.57	0.54	0.5	0.47	0.43
2.5	0.89	0.93	0.97	0.91	0.85	0.81	0.77	0.73	0.69	0.65	0.61	0.57	0.54	0.5	0.46
3.0	0.87	0.9	0.94	0.98	0.92	0.87	0.82	0.78	0.73	0.69	0.65	0.61	0.57	0.53	0.48
3.5	0.84	0.88	0.9	0.94	0.98	0.93	0.87	0.83	0.78	0.73	0.69	0.64	0.6	0.56	0.51
4.0	0.82	0.85	0.88	0.91	0.95	0.98	0.93	0.87	0.82	0.77	0.72	0.68	0.63	0.58	0.53
4.5	0.79	0.82	0.85	0.88	0.91	0.95	0.98	0.92	0.87	0.82	0.76	0.71	0.66	0.61	0.56
5.0	0.78	0.8	0.83	0.86	0.89	0.91	0.95	0.98	0.93	0.87	0.81	0.76	0.7	0.65	0.59
5.5	0.75	0.78	0.8	0.83	0.86	0.88	0.92	0.95	0.99	0.92	0.86	0.8	0.75	0.69	0.63
6.0	0.73	0.75	0.78	0.8	0.83	0.85	0.88	0.91	0.95	0.99	0.92	0.86	0.8	0.74	0.68
6.5	0.7	0.72	0.75	0.77	0.8	0.82	0.85	0.88	0.91	0.95	0.99	0.92	0.85	0.79	0.73
7.0	0.67	0.69	0.72	0.74	0.76	0.79	0.81	0.84	0.87	0.9	0.94	0.99	0.92	0.85	0.78
7.5	0.65	0.66	0.69	0.71	0.73	0.75	0.77	0.8	0.82	0.86	0.9	0.94	0.99	0.92	0.84
8.0	0.61	0.63	0.65	0.67	0.69	0.71	0.73	0.75	0.78	0.82	0.85	0.89	0.94	0.99	0.91
8.5	0.58	0.59	0.61	0.63	0.65	0.67	0.68	0.71	0.73	0.77	0.8	0.84	0.88	0.93	0.99

$\langle q^8 \rangle$

	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
1.5	0.9	0.84	0.79	0.75	0.71	0.68	0.65	0.62	0.59	0.56	0.53	0.5	0.47	0.44	0.41
2.0	0.88	0.9	0.85	0.79	0.76	0.72	0.69	0.65	0.62	0.59	0.55	0.52	0.49	0.46	0.42
2.5	0.86	0.89	0.91	0.85	0.81	0.77	0.73	0.7	0.66	0.63	0.59	0.55	0.52	0.49	0.45
3.0	0.85	0.87	0.89	0.92	0.87	0.82	0.78	0.74	0.7	0.67	0.63	0.59	0.55	0.51	0.47
3.5	0.85	0.86	0.88	0.9	0.92	0.88	0.83	0.79	0.75	0.71	0.67	0.62	0.58	0.54	0.5
4.0	0.84	0.85	0.87	0.89	0.91	0.93	0.88	0.83	0.79	0.74	0.7	0.66	0.61	0.57	0.52
4.5	0.82	0.84	0.86	0.88	0.89	0.91	0.93	0.88	0.83	0.78	0.74	0.69	0.64	0.59	0.54
5.0	0.82	0.84	0.85	0.87	0.88	0.9	0.92	0.94	0.89	0.84	0.78	0.73	0.68	0.63	0.58
5.5	0.81	0.83	0.84	0.85	0.87	0.89	0.9	0.92	0.94	0.89	0.83	0.78	0.73	0.67	0.62
6.0	0.8	0.81	0.83	0.84	0.86	0.87	0.89	0.91	0.92	0.95	0.89	0.83	0.77	0.72	0.66
6.5	0.78	0.8	0.81	0.83	0.84	0.85	0.87	0.89	0.91	0.93	0.95	0.89	0.83	0.77	0.71
7.0	0.77	0.78	0.8	0.81	0.82	0.84	0.85	0.87	0.88	0.91	0.93	0.96	0.89	0.83	0.76
7.5	0.75	0.76	0.78	0.79	0.8	0.81	0.83	0.84	0.86	0.88	0.9	0.93	0.96	0.89	0.82
8.0	0.73	0.74	0.75	0.76	0.78	0.79	0.8	0.81	0.83	0.85	0.88	0.9	0.93	0.97	0.89
8.5	0.7	0.71	0.73	0.74	0.75	0.76	0.77	0.78	0.8	0.82	0.84	0.87	0.9	0.93	0.97

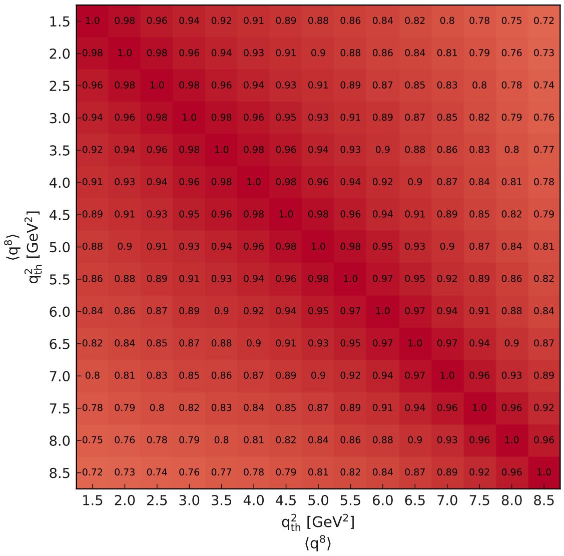
Figure C.2: Statistical correlations between $\langle q^4 \rangle$ and $\langle q^{2n} \rangle$ for $n = 2–4$.



	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
1.5	1.0	0.96	0.92	0.89	0.87	0.84	0.81	0.79	0.77	0.74	0.71	0.68	0.65	0.62	0.58
2.0	0.96	1.0	0.96	0.92	0.9	0.87	0.84	0.82	0.79	0.76	0.73	0.7	0.67	0.64	0.6
2.5	0.92	0.96	1.0	0.96	0.93	0.9	0.87	0.85	0.82	0.79	0.76	0.72	0.69	0.66	0.62
3.0	0.89	0.92	0.96	1.0	0.97	0.93	0.9	0.87	0.84	0.81	0.78	0.75	0.71	0.67	0.64
3.5	0.87	0.9	0.93	0.97	1.0	0.97	0.93	0.9	0.87	0.84	0.81	0.77	0.73	0.7	0.65
4.0	0.84	0.87	0.9	0.93	0.97	1.0	0.97	0.93	0.9	0.86	0.83	0.79	0.75	0.71	0.67
4.5	0.81	0.84	0.87	0.9	0.93	0.97	1.0	0.96	0.93	0.89	0.86	0.82	0.77	0.73	0.69
5.0	0.79	0.82	0.85	0.87	0.9	0.93	0.96	1.0	0.96	0.93	0.89	0.85	0.8	0.76	0.71
5.5	0.77	0.79	0.82	0.84	0.87	0.9	0.93	0.96	1.0	0.96	0.92	0.87	0.83	0.79	0.74
6.0	0.74	0.76	0.79	0.81	0.84	0.86	0.89	0.93	0.96	1.0	0.96	0.91	0.87	0.82	0.77
6.5	0.71	0.73	0.76	0.78	0.81	0.83	0.86	0.89	0.92	0.96	1.0	0.95	0.91	0.86	0.81
7.0	0.68	0.7	0.72	0.75	0.77	0.79	0.82	0.85	0.87	0.91	0.95	1.0	0.95	0.9	0.84
7.5	0.65	0.67	0.69	0.71	0.73	0.75	0.77	0.8	0.83	0.87	0.91	0.95	1.0	0.95	0.89
8.0	0.62	0.64	0.66	0.67	0.7	0.71	0.73	0.76	0.79	0.82	0.86	0.9	0.95	1.0	0.94
8.5	0.58	0.6	0.62	0.64	0.65	0.67	0.69	0.71	0.74	0.77	0.81	0.84	0.89	0.94	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
1.5	0.98	0.94	0.91	0.88	0.85	0.83	0.8	0.78	0.76	0.73	0.7	0.67	0.65	0.62	0.58
2.0	0.95	0.98	0.94	0.91	0.88	0.86	0.83	0.81	0.78	0.75	0.72	0.69	0.66	0.63	0.59
2.5	0.93	0.96	0.98	0.94	0.91	0.89	0.86	0.83	0.81	0.78	0.75	0.72	0.68	0.65	0.61
3.0	0.92	0.93	0.96	0.98	0.95	0.92	0.89	0.86	0.83	0.81	0.77	0.74	0.7	0.67	0.63
3.5	0.9	0.92	0.94	0.96	0.98	0.95	0.92	0.89	0.86	0.83	0.8	0.76	0.73	0.69	0.65
4.0	0.89	0.91	0.93	0.94	0.96	0.98	0.95	0.92	0.89	0.85	0.82	0.78	0.75	0.71	0.66
4.5	0.87	0.89	0.91	0.93	0.95	0.96	0.98	0.95	0.92	0.88	0.84	0.81	0.77	0.73	0.68
5.0	0.86	0.88	0.9	0.91	0.93	0.94	0.96	0.98	0.95	0.91	0.88	0.84	0.8	0.75	0.71
5.5	0.85	0.87	0.88	0.9	0.91	0.93	0.94	0.96	0.99	0.95	0.91	0.86	0.82	0.78	0.73
6.0	0.83	0.85	0.86	0.88	0.89	0.91	0.92	0.94	0.96	0.99	0.95	0.9	0.86	0.81	0.77
6.5	0.81	0.83	0.84	0.86	0.87	0.89	0.9	0.92	0.94	0.96	0.99	0.94	0.9	0.85	0.8
7.0	0.79	0.81	0.82	0.84	0.85	0.86	0.88	0.89	0.91	0.94	0.96	0.99	0.94	0.89	0.84
7.5	0.77	0.78	0.8	0.81	0.82	0.84	0.85	0.87	0.88	0.9	0.93	0.96	0.99	0.94	0.88
8.0	0.75	0.76	0.77	0.78	0.79	0.81	0.82	0.83	0.85	0.87	0.9	0.92	0.96	0.99	0.93
8.5	0.72	0.73	0.74	0.75	0.76	0.77	0.78	0.8	0.82	0.84	0.86	0.89	0.92	0.95	0.99

Figure C.3: Statistical correlations between $\langle q^6 \rangle$ and $\langle q^{2n} \rangle$ for $n = 3, 4$.

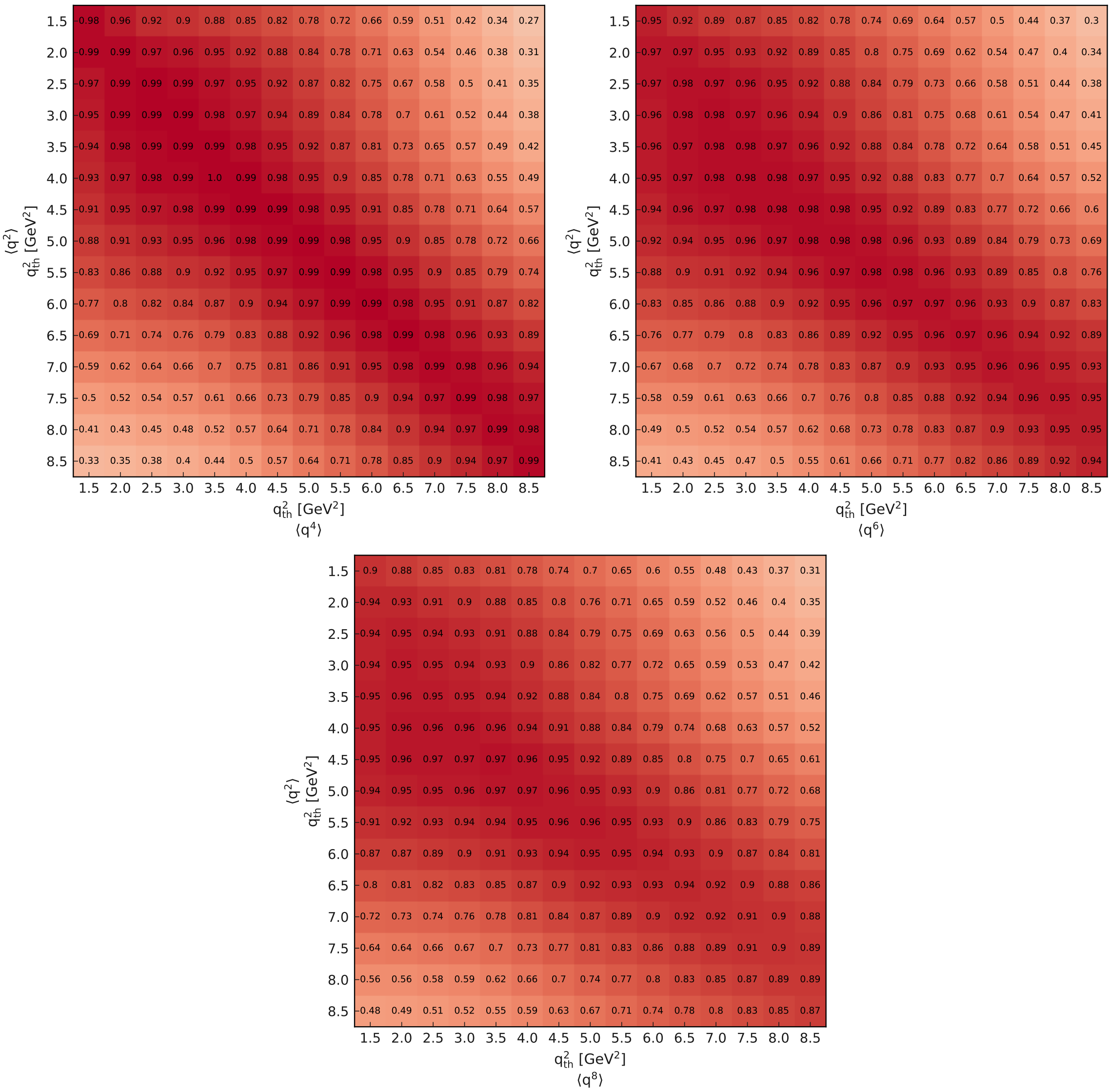


	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
1.5	1.0	0.98	0.96	0.94	0.92	0.91	0.89	0.88	0.86	0.84	0.82	0.8	0.78	0.75	0.72
2.0	0.98	1.0	0.98	0.96	0.94	0.93	0.91	0.9	0.88	0.86	0.84	0.81	0.79	0.76	0.73
2.5	0.96	0.98	1.0	0.98	0.96	0.94	0.93	0.91	0.89	0.87	0.85	0.83	0.8	0.78	0.74
3.0	0.94	0.96	0.98	1.0	0.98	0.96	0.95	0.93	0.91	0.89	0.87	0.85	0.82	0.79	0.76
3.5	0.92	0.94	0.96	0.98	1.0	0.98	0.96	0.94	0.93	0.9	0.88	0.86	0.83	0.8	0.77
4.0	0.91	0.93	0.94	0.96	0.98	1.0	0.98	0.96	0.94	0.92	0.9	0.87	0.84	0.81	0.78
4.5	0.89	0.91	0.93	0.95	0.96	0.98	1.0	0.98	0.96	0.94	0.91	0.89	0.85	0.82	0.79
5.0	0.88	0.9	0.91	0.93	0.94	0.96	0.98	1.0	0.98	0.95	0.93	0.9	0.87	0.84	0.81
5.5	0.86	0.88	0.89	0.91	0.93	0.94	0.96	0.98	1.0	0.97	0.95	0.92	0.89	0.86	0.82
6.0	0.84	0.86	0.87	0.89	0.9	0.92	0.94	0.95	0.97	1.0	0.97	0.94	0.91	0.88	0.84
6.5	0.82	0.84	0.85	0.87	0.88	0.9	0.91	0.93	0.95	0.97	1.0	0.97	0.94	0.9	0.87
7.0	0.8	0.81	0.83	0.85	0.86	0.87	0.89	0.9	0.92	0.94	0.97	1.0	0.96	0.93	0.89
7.5	0.78	0.79	0.8	0.82	0.83	0.84	0.85	0.87	0.89	0.91	0.94	0.96	1.0	0.96	0.92
8.0	0.75	0.76	0.78	0.79	0.8	0.81	0.82	0.84	0.86	0.88	0.9	0.93	0.96	1.0	0.96
8.5	0.72	0.73	0.74	0.76	0.77	0.78	0.79	0.81	0.82	0.84	0.87	0.89	0.92	0.96	1.0

Figure C.4: Statistical correlations between $\langle q^8 \rangle$ and $\langle q^8 \rangle$.

Raw Moments (Full Correlation)

Figures C.5 to C.8 summarize the numerical values for the full experimental (statistical and systematic) correlation between the raw q^2 moments of different order and lower q^2 threshold. The full correlation structure is shown in Figure 5.26.



⟨q⁴⟩

⟨q^2⟩ q^2_{th} [GeV^2] \ q^2_{th} [GeV^2]	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
1.5	0.98	0.96	0.92	0.9	0.88	0.85	0.82	0.78	0.72	0.66	0.59	0.51	0.42	0.34	0.27
2.0	0.99	0.99	0.97	0.96	0.95	0.92	0.88	0.84	0.78	0.71	0.63	0.54	0.46	0.38	0.31
2.5	0.97	0.99	0.99	0.99	0.97	0.95	0.92	0.87	0.82	0.75	0.67	0.58	0.5	0.41	0.35
3.0	0.95	0.99	0.99	0.99	0.98	0.97	0.94	0.89	0.84	0.78	0.7	0.61	0.52	0.44	0.38
3.5	0.94	0.98	0.99	0.99	0.99	0.98	0.95	0.92	0.87	0.81	0.73	0.65	0.57	0.49	0.42
4.0	0.93	0.97	0.98	0.99	1.0	0.99	0.98	0.95	0.9	0.85	0.78	0.71	0.63	0.55	0.49
4.5	0.91	0.95	0.97	0.98	0.99	0.99	0.99	0.98	0.95	0.91	0.85	0.78	0.71	0.64	0.57
5.0	0.88	0.91	0.93	0.95	0.96	0.98	0.99	0.99	0.98	0.95	0.9	0.85	0.78	0.72	0.66
5.5	0.83	0.86	0.88	0.9	0.92	0.95	0.97	0.99	0.99	0.98	0.95	0.9	0.85	0.79	0.74
6.0	0.77	0.8	0.82	0.84	0.87	0.9	0.94	0.97	0.99	0.99	0.98	0.95	0.91	0.87	0.82
6.5	0.69	0.71	0.74	0.76	0.79	0.83	0.88	0.92	0.96	0.98	0.99	0.98	0.96	0.93	0.89
7.0	0.59	0.62	0.64	0.66	0.7	0.75	0.81	0.86	0.91	0.95	0.98	0.99	0.98	0.96	0.94
7.5	0.5	0.52	0.54	0.57	0.61	0.66	0.73	0.79	0.85	0.9	0.94	0.97	0.99	0.98	0.97
8.0	0.41	0.43	0.45	0.48	0.52	0.57	0.64	0.71	0.78	0.84	0.9	0.94	0.97	0.99	0.98
8.5	0.33	0.35	0.38	0.4	0.44	0.5	0.57	0.64	0.71	0.78	0.85	0.9	0.94	0.97	0.99

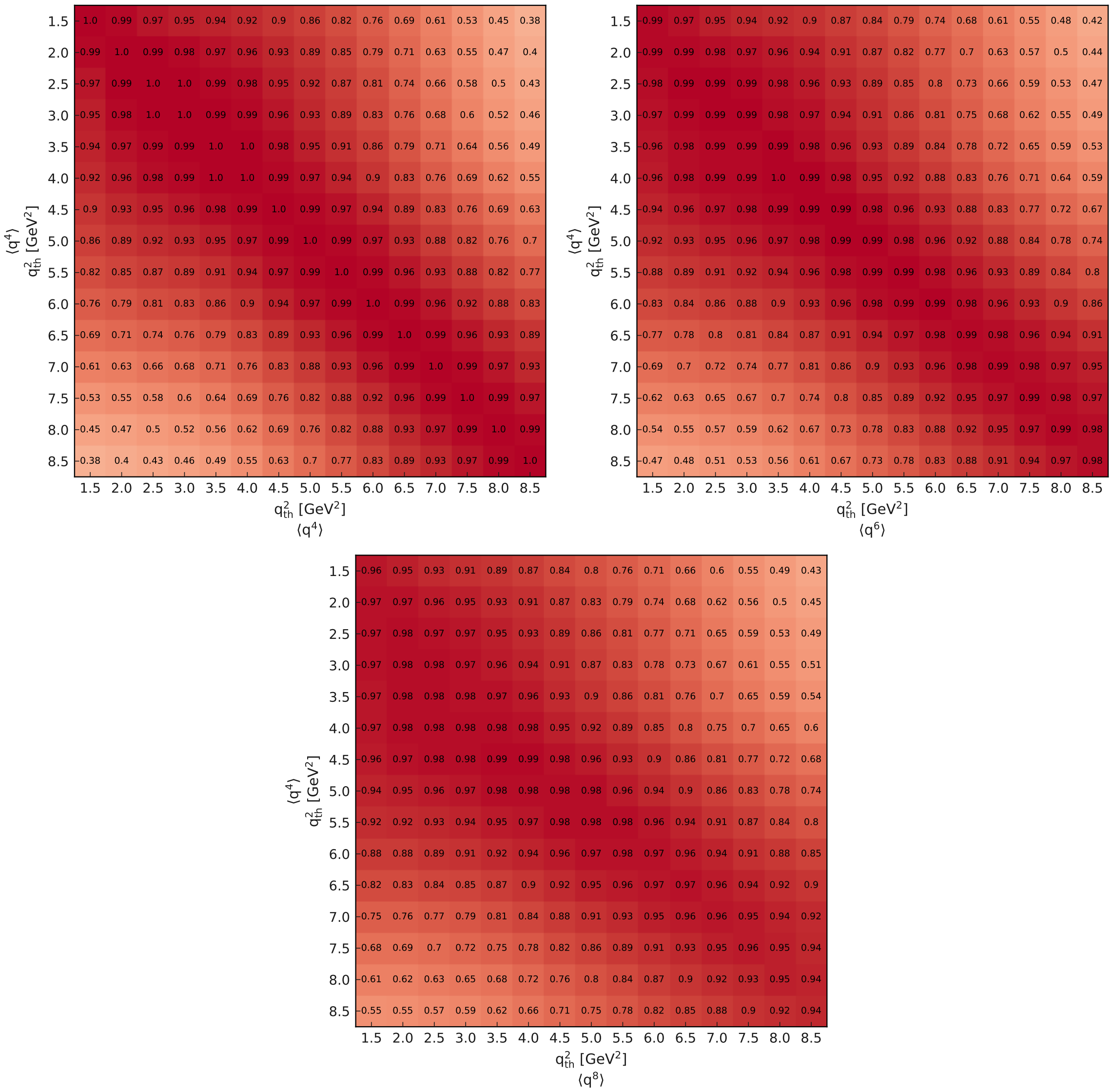
⟨q⁶⟩

⟨q^2⟩ q^2_{th} [GeV^2] \ q^2_{th} [GeV^2]	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
1.5	0.95	0.92	0.89	0.87	0.85	0.82	0.78	0.74	0.69	0.64	0.57	0.5	0.44	0.37	0.3
2.0	0.97	0.97	0.95	0.93	0.92	0.89	0.85	0.8	0.75	0.69	0.62	0.54	0.47	0.4	0.34
2.5	0.97	0.98	0.97	0.96	0.95	0.92	0.88	0.84	0.79	0.73	0.66	0.58	0.51	0.44	0.38
3.0	0.96	0.98	0.98	0.97	0.96	0.94	0.9	0.86	0.81	0.75	0.68	0.61	0.54	0.47	0.41
3.5	0.96	0.97	0.98	0.98	0.97	0.96	0.92	0.88	0.84	0.78	0.72	0.64	0.58	0.51	0.45
4.0	0.95	0.97	0.98	0.98	0.98	0.97	0.95	0.92	0.88	0.83	0.77	0.7	0.64	0.57	0.52
4.5	0.94	0.96	0.97	0.98	0.98	0.98	0.98	0.95	0.92	0.89	0.83	0.77	0.72	0.66	0.6
5.0	0.92	0.94	0.95	0.96	0.97	0.98	0.98	0.98	0.96	0.93	0.89	0.84	0.79	0.73	0.69
5.5	0.88	0.9	0.91	0.92	0.94	0.96	0.97	0.98	0.98	0.96	0.93	0.89	0.85	0.8	0.76
6.0	0.83	0.85	0.86	0.88	0.9	0.92	0.95	0.96	0.97	0.97	0.96	0.93	0.9	0.87	0.83
6.5	0.76	0.77	0.79	0.8	0.83	0.86	0.89	0.92	0.95	0.96	0.97	0.96	0.94	0.92	0.89
7.0	0.67	0.68	0.7	0.72	0.74	0.78	0.83	0.87	0.9	0.93	0.95	0.96	0.96	0.95	0.93
7.5	0.58	0.59	0.61	0.63	0.66	0.7	0.76	0.8	0.85	0.88	0.92	0.94	0.96	0.95	0.95
8.0	0.49	0.5	0.52	0.54	0.57	0.62	0.68	0.73	0.78	0.83	0.87	0.9	0.93	0.95	0.95
8.5	0.41	0.43	0.45	0.47	0.5	0.55	0.61	0.66	0.71	0.77	0.82	0.86	0.89	0.92	0.94

⟨q⁸⟩

⟨q^2⟩ q^2_{th} [GeV^2] \ q^2_{th} [GeV^2]	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
1.5	0.9	0.88	0.85	0.83	0.81	0.78	0.74	0.7	0.65	0.6	0.55	0.48	0.43	0.37	0.31
2.0	0.94	0.93	0.91	0.9	0.88	0.85	0.8	0.76	0.71	0.65	0.59	0.52	0.46	0.4	0.35
2.5	0.94	0.95	0.94	0.93	0.91	0.88	0.84	0.79	0.75	0.69	0.63	0.56	0.5	0.44	0.39
3.0	0.94	0.95	0.95	0.94	0.93	0.9	0.86	0.82	0.77	0.72	0.65	0.59	0.53	0.47	0.42
3.5	0.95	0.96	0.95	0.95	0.94	0.92	0.88	0.84	0.8	0.75	0.69	0.62	0.57	0.51	0.46
4.0	0.95	0.96	0.96	0.96	0.96	0.94	0.91	0.88	0.84	0.79	0.74	0.68	0.63	0.57	0.52
4.5	0.95	0.96	0.97	0.97	0.97	0.96	0.95	0.92	0.89	0.85	0.8	0.75	0.7	0.65	0.61
5.0	0.94	0.95	0.95	0.96	0.97	0.97	0.96	0.95	0.93	0.9	0.86	0.81	0.77	0.72	0.68
5.5	0.91	0.92	0.93	0.94	0.94	0.95	0.96	0.96	0.95	0.93	0.9	0.86	0.83	0.79	0.75
6.0	0.87	0.87	0.89	0.9	0.91	0.93	0.94	0.95	0.95	0.94	0.93	0.9	0.87	0.84	0.81
6.5	0.8	0.81	0.82	0.83	0.85	0.87	0.9	0.92	0.93	0.93	0.94	0.92	0.9	0.88	0.86
7.0	0.72	0.73	0.74	0.76	0.78	0.81	0.84	0.87	0.89	0.9	0.92	0.92	0.91	0.9	0.88
7.5	0.64	0.64	0.66	0.67	0.7	0.73	0.77	0.81	0.83	0.86	0.88	0.89	0.91	0.9	0.89
8.0	0.56	0.56	0.58	0.59	0.62	0.66	0.7	0.74	0.77	0.8	0.83	0.85	0.87	0.89	0.89
8.5	0.48	0.49	0.51	0.52	0.55	0.59	0.63	0.67	0.71	0.74	0.78	0.8	0.83	0.85	0.87

Figure C.5: Full experimental correlations between $\langle q^2 \rangle$ and $\langle q^{2n} \rangle$ for $n = 2$–4.



$\langle q^4\rangle$ q^2_{th} [GeV^2]	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
1.5	1.0	0.99	0.97	0.95	0.94	0.92	0.9	0.86	0.82	0.76	0.69	0.61	0.53	0.45	0.38
2.0	0.99	1.0	0.99	0.98	0.97	0.96	0.93	0.89	0.85	0.79	0.71	0.63	0.55	0.47	0.4
2.5	0.97	0.99	1.0	1.0	0.99	0.98	0.95	0.92	0.87	0.81	0.74	0.66	0.58	0.5	0.43
3.0	0.95	0.98	1.0	1.0	0.99	0.99	0.96	0.93	0.89	0.83	0.76	0.68	0.6	0.52	0.46
3.5	0.94	0.97	0.99	0.99	1.0	1.0	0.98	0.95	0.91	0.86	0.79	0.71	0.64	0.56	0.49
4.0	0.92	0.96	0.98	0.99	1.0	1.0	0.99	0.97	0.94	0.9	0.83	0.76	0.69	0.62	0.55
4.5	0.9	0.93	0.95	0.96	0.98	0.99	1.0	0.99	0.97	0.94	0.89	0.83	0.76	0.69	0.63
5.0	0.86	0.89	0.92	0.93	0.95	0.97	0.99	1.0	0.99	0.97	0.93	0.88	0.82	0.76	0.7
5.5	0.82	0.85	0.87	0.89	0.91	0.94	0.97	0.99	1.0	0.99	0.96	0.93	0.88	0.82	0.77
6.0	0.76	0.79	0.81	0.83	0.86	0.9	0.94	0.97	0.99	1.0	0.99	0.96	0.92	0.88	0.83
6.5	0.69	0.71	0.74	0.76	0.79	0.83	0.89	0.93	0.96	0.99	1.0	0.99	0.96	0.93	0.89
7.0	0.61	0.63	0.66	0.68	0.71	0.76	0.83	0.88	0.93	0.96	0.99	1.0	0.99	0.97	0.93
7.5	0.53	0.55	0.58	0.6	0.64	0.69	0.76	0.82	0.88	0.92	0.96	0.99	1.0	0.99	0.97
8.0	0.45	0.47	0.5	0.52	0.56	0.62	0.69	0.76	0.82	0.88	0.93	0.97	0.99	1.0	0.99
8.5	0.38	0.4	0.43	0.46	0.49	0.55	0.63	0.7	0.77	0.83	0.89	0.93	0.97	0.99	1.0

q^2_{th} [GeV^2]
$\langle q^4\rangle$

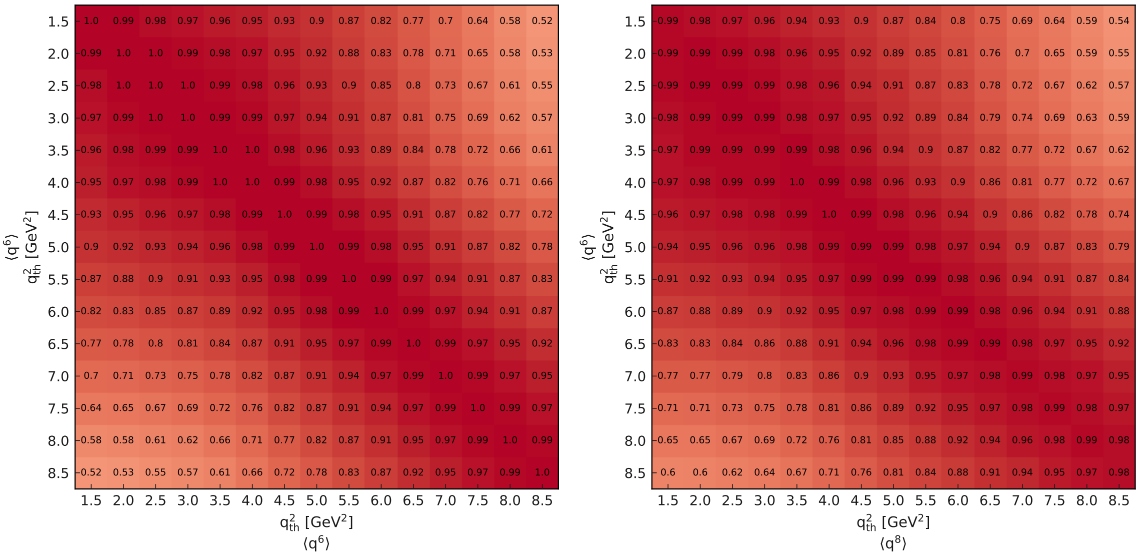
$\langle q^4\rangle$ q^2_{th} [GeV^2]	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
1.5	0.99	0.97	0.95	0.94	0.92	0.9	0.87	0.84	0.79	0.74	0.68	0.61	0.55	0.48	0.42
2.0	0.99	0.99	0.98	0.97	0.96	0.94	0.91	0.87	0.82	0.77	0.7	0.63	0.57	0.5	0.44
2.5	0.98	0.99	0.99	0.99	0.98	0.96	0.93	0.89	0.85	0.8	0.73	0.66	0.59	0.53	0.47
3.0	0.97	0.99	0.99	0.99	0.98	0.97	0.94	0.91	0.86	0.81	0.75	0.68	0.62	0.55	0.49
3.5	0.96	0.98	0.99	0.99	0.99	0.98	0.96	0.93	0.89	0.84	0.78	0.72	0.65	0.59	0.53
4.0	0.96	0.98	0.99	0.99	1.0	0.99	0.98	0.95	0.92	0.88	0.83	0.76	0.71	0.64	0.59
4.5	0.94	0.96	0.97	0.98	0.99	0.99	0.99	0.98	0.96	0.93	0.88	0.83	0.77	0.72	0.67
5.0	0.92	0.93	0.95	0.96	0.97	0.98	0.99	0.99	0.98	0.96	0.92	0.88	0.84	0.78	0.74
5.5	0.88	0.89	0.91	0.92	0.94	0.96	0.98	0.99	0.99	0.98	0.96	0.93	0.89	0.84	0.8
6.0	0.83	0.84	0.86	0.88	0.9	0.93	0.96	0.98	0.99	0.99	0.98	0.96	0.93	0.9	0.86
6.5	0.77	0.78	0.8	0.81	0.84	0.87	0.91	0.94	0.97	0.98	0.99	0.98	0.96	0.94	0.91
7.0	0.69	0.7	0.72	0.74	0.77	0.81	0.86	0.9	0.93	0.96	0.98	0.99	0.98	0.97	0.95
7.5	0.62	0.63	0.65	0.67	0.7	0.74	0.8	0.85	0.89	0.92	0.95	0.97	0.99	0.98	0.97
8.0	0.54	0.55	0.57	0.59	0.62	0.67	0.73	0.78	0.83	0.88	0.92	0.95	0.97	0.99	0.98
8.5	0.47	0.48	0.51	0.53	0.56	0.61	0.67	0.73	0.78	0.83	0.88	0.91	0.94	0.97	0.98

q^2_{th} [GeV^2]
$\langle q^6\rangle$

$\langle q^4\rangle$ q^2_{th} [GeV^2]	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
1.5	0.96	0.95	0.93	0.91	0.89	0.87	0.84	0.8	0.76	0.71	0.66	0.6	0.55	0.49	0.43
2.0	0.97	0.97	0.96	0.95	0.93	0.91	0.87	0.83	0.79	0.74	0.68	0.62	0.56	0.5	0.45
2.5	0.97	0.98	0.97	0.97	0.95	0.93	0.89	0.86	0.81	0.77	0.71	0.65	0.59	0.53	0.49
3.0	0.97	0.98	0.98	0.97	0.96	0.94	0.91	0.87	0.83	0.78	0.73	0.67	0.61	0.55	0.51
3.5	0.97	0.98	0.98	0.98	0.97	0.96	0.93	0.9	0.86	0.81	0.76	0.7	0.65	0.59	0.54
4.0	0.97	0.98	0.98	0.98	0.98	0.98	0.95	0.92	0.89	0.85	0.8	0.75	0.7	0.65	0.6
4.5	0.96	0.97	0.98	0.98	0.99	0.99	0.98	0.96	0.93	0.9	0.86	0.81	0.77	0.72	0.68
5.0	0.94	0.95	0.96	0.97	0.98	0.98	0.98	0.98	0.96	0.94	0.9	0.86	0.83	0.78	0.74
5.5	0.92	0.92	0.93	0.94	0.95	0.97	0.98	0.98	0.98	0.96	0.94	0.91	0.87	0.84	0.8
6.0	0.88	0.88	0.89	0.91	0.92	0.94	0.96	0.97	0.98	0.97	0.96	0.94	0.91	0.88	0.85
6.5	0.82	0.83	0.84	0.85	0.87	0.9	0.92	0.95	0.96	0.97	0.97	0.96	0.94	0.92	0.9
7.0	0.75	0.76	0.77	0.79	0.81	0.84	0.88	0.91	0.93	0.95	0.96	0.96	0.95	0.94	0.92
7.5	0.68	0.69	0.7	0.72	0.75	0.78	0.82	0.86	0.89	0.91	0.93	0.95	0.96	0.95	0.94
8.0	0.61	0.62	0.63	0.65	0.68	0.72	0.76	0.8	0.84	0.87	0.9	0.92	0.93	0.95	0.94
8.5	0.55	0.55	0.57	0.59	0.62	0.66	0.71	0.75	0.78	0.82	0.85	0.88	0.9	0.92	0.94

q^2_{th} [GeV^2]
$\langle q^8\rangle$

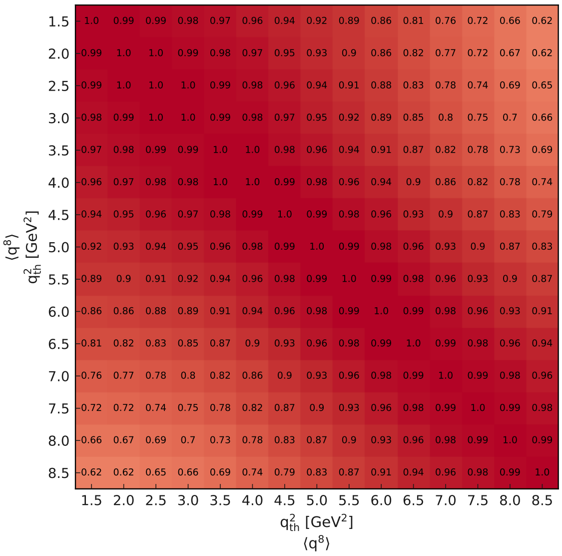
Figure C.6: Full experimental correlations between $\langle q^4\rangle$ and $\langle q^{2n}\rangle$ for n = 2–4.



	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
1.5	1.0	0.99	0.98	0.97	0.96	0.95	0.93	0.9	0.87	0.82	0.77	0.7	0.64	0.58	0.52
2.0	0.99	1.0	1.0	0.99	0.98	0.97	0.95	0.92	0.88	0.83	0.78	0.71	0.65	0.58	0.53
2.5	0.98	1.0	1.0	1.0	0.99	0.98	0.96	0.93	0.9	0.85	0.8	0.73	0.67	0.61	0.55
3.0	0.97	0.99	1.0	1.0	0.99	0.99	0.97	0.94	0.91	0.87	0.81	0.75	0.69	0.62	0.57
3.5	0.96	0.98	0.99	0.99	1.0	1.0	0.98	0.96	0.93	0.89	0.84	0.78	0.72	0.66	0.61
4.0	0.95	0.97	0.98	0.99	1.0	1.0	0.99	0.98	0.95	0.92	0.87	0.82	0.76	0.71	0.66
4.5	0.93	0.95	0.96	0.97	0.98	0.99	1.0	0.99	0.98	0.95	0.91	0.87	0.82	0.77	0.72
5.0	0.9	0.92	0.93	0.94	0.96	0.98	0.99	1.0	0.99	0.98	0.95	0.91	0.87	0.82	0.78
5.5	0.87	0.88	0.9	0.91	0.93	0.95	0.98	0.99	1.0	0.99	0.97	0.94	0.91	0.87	0.83
6.0	0.82	0.83	0.85	0.87	0.89	0.92	0.95	0.98	0.99	1.0	0.99	0.97	0.94	0.91	0.87
6.5	0.77	0.78	0.8	0.81	0.84	0.87	0.91	0.95	0.97	0.99	1.0	0.99	0.97	0.95	0.92
7.0	0.7	0.71	0.73	0.75	0.78	0.82	0.87	0.91	0.94	0.97	0.99	1.0	0.99	0.97	0.95
7.5	0.64	0.65	0.67	0.69	0.72	0.76	0.82	0.87	0.91	0.94	0.97	0.99	1.0	0.99	0.97
8.0	0.58	0.58	0.61	0.62	0.66	0.71	0.77	0.82	0.87	0.91	0.95	0.97	0.99	1.0	0.99
8.5	0.52	0.53	0.55	0.57	0.61	0.66	0.72	0.78	0.83	0.87	0.92	0.95	0.97	0.99	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
1.5	0.99	0.98	0.97	0.96	0.94	0.93	0.9	0.87	0.84	0.8	0.75	0.69	0.64	0.59	0.54
2.0	0.99	0.99	0.99	0.98	0.96	0.95	0.92	0.89	0.85	0.81	0.76	0.7	0.65	0.59	0.55
2.5	0.99	0.99	0.99	0.99	0.98	0.96	0.94	0.91	0.87	0.83	0.78	0.72	0.67	0.62	0.57
3.0	0.98	0.99	0.99	0.99	0.98	0.97	0.95	0.92	0.89	0.84	0.79	0.74	0.69	0.63	0.59
3.5	0.97	0.99	0.99	0.99	0.99	0.98	0.96	0.94	0.9	0.87	0.82	0.77	0.72	0.67	0.62
4.0	0.97	0.98	0.99	0.99	1.0	0.99	0.98	0.96	0.93	0.9	0.86	0.81	0.77	0.72	0.67
4.5	0.96	0.97	0.98	0.98	0.99	1.0	0.99	0.98	0.96	0.94	0.9	0.86	0.82	0.78	0.74
5.0	0.94	0.95	0.96	0.96	0.98	0.99	0.99	0.99	0.98	0.97	0.94	0.9	0.87	0.83	0.79
5.5	0.91	0.92	0.93	0.94	0.95	0.97	0.99	0.99	0.99	0.98	0.96	0.94	0.91	0.87	0.84
6.0	0.87	0.88	0.89	0.9	0.92	0.95	0.97	0.98	0.99	0.99	0.98	0.96	0.94	0.91	0.88
6.5	0.83	0.83	0.84	0.86	0.88	0.91	0.94	0.96	0.98	0.99	0.99	0.98	0.97	0.95	0.92
7.0	0.77	0.77	0.79	0.8	0.83	0.86	0.9	0.93	0.95	0.97	0.98	0.99	0.98	0.97	0.95
7.5	0.71	0.71	0.73	0.75	0.78	0.81	0.86	0.89	0.92	0.95	0.97	0.98	0.99	0.98	0.97
8.0	0.65	0.65	0.67	0.69	0.72	0.76	0.81	0.85	0.88	0.92	0.94	0.96	0.98	0.99	0.98
8.5	0.6	0.6	0.62	0.64	0.67	0.71	0.76	0.81	0.84	0.88	0.91	0.94	0.95	0.97	0.98

Figure C.7: Full experimental correlations between $\langle q^6 \rangle$ and $\langle q^{2n} \rangle$ for $n = 3, 4$.

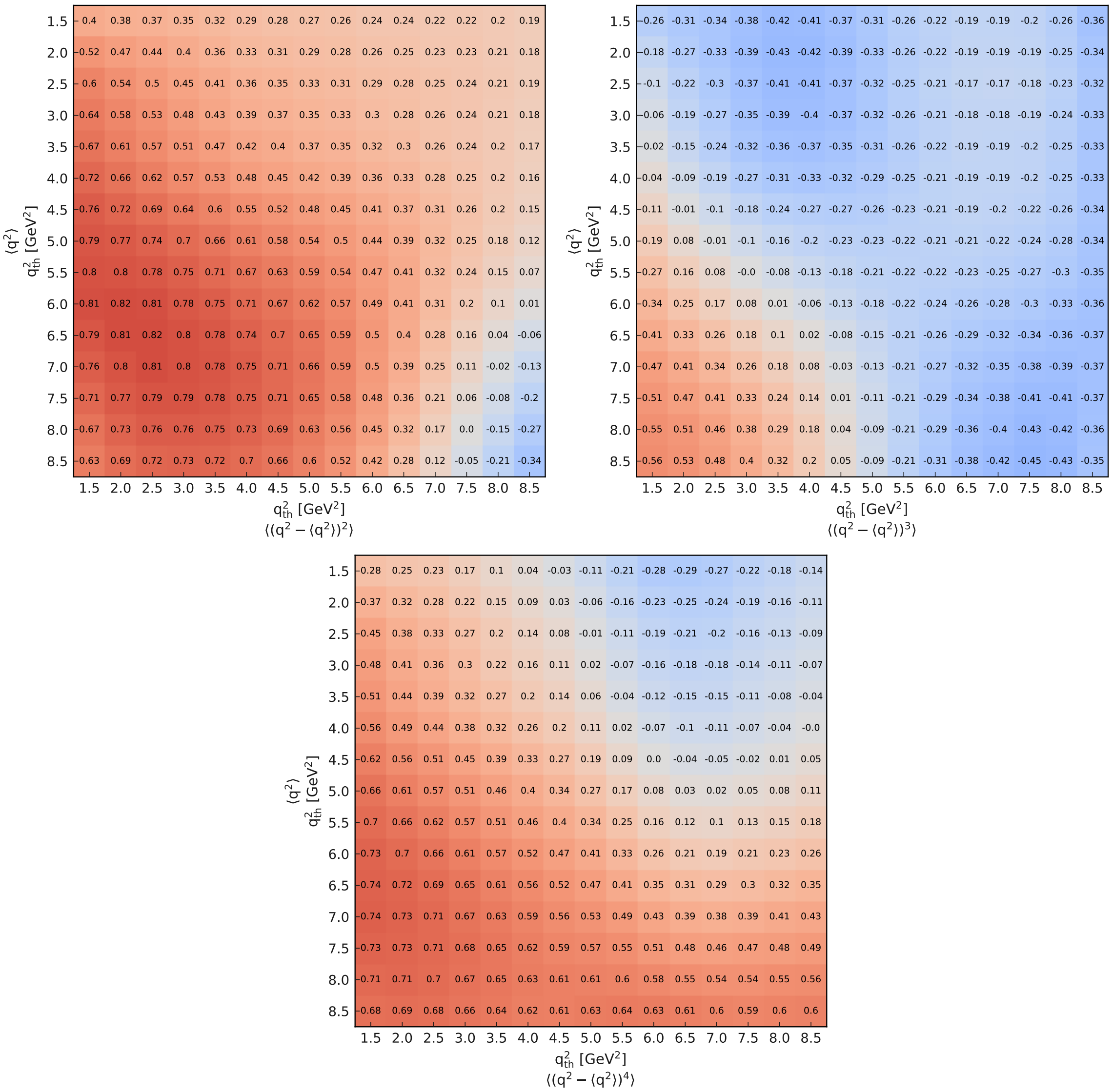


	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
1.5	1.0	0.99	0.99	0.98	0.97	0.96	0.94	0.92	0.89	0.86	0.81	0.76	0.72	0.66	0.62
2.0	0.99	1.0	1.0	0.99	0.98	0.97	0.95	0.93	0.9	0.86	0.82	0.77	0.72	0.67	0.62
2.5	0.99	1.0	1.0	1.0	0.99	0.98	0.96	0.94	0.91	0.88	0.83	0.78	0.74	0.69	0.65
3.0	0.98	0.99	1.0	1.0	0.99	0.98	0.97	0.95	0.92	0.89	0.85	0.8	0.75	0.7	0.66
3.5	0.97	0.98	0.99	0.99	1.0	1.0	0.98	0.96	0.94	0.91	0.87	0.82	0.78	0.73	0.69
4.0	0.96	0.97	0.98	0.98	1.0	1.0	0.99	0.98	0.96	0.94	0.9	0.86	0.82	0.78	0.74
4.5	0.94	0.95	0.96	0.97	0.98	0.99	1.0	0.99	0.98	0.96	0.93	0.9	0.87	0.83	0.79
5.0	0.92	0.93	0.94	0.95	0.96	0.98	0.99	1.0	0.99	0.98	0.96	0.93	0.9	0.87	0.83
5.5	0.89	0.9	0.91	0.92	0.94	0.96	0.98	0.99	1.0	0.99	0.98	0.96	0.93	0.9	0.87
6.0	0.86	0.86	0.88	0.89	0.91	0.94	0.96	0.98	0.99	1.0	0.99	0.98	0.96	0.93	0.91
6.5	0.81	0.82	0.83	0.85	0.87	0.9	0.93	0.96	0.98	0.99	1.0	0.99	0.98	0.96	0.94
7.0	0.76	0.77	0.78	0.8	0.82	0.86	0.9	0.93	0.96	0.98	0.99	1.0	0.99	0.98	0.96
7.5	0.72	0.72	0.74	0.75	0.78	0.82	0.87	0.9	0.93	0.96	0.98	0.99	1.0	0.99	0.98
8.0	0.66	0.67	0.69	0.7	0.73	0.78	0.83	0.87	0.9	0.93	0.96	0.98	0.99	1.0	0.99
8.5	0.62	0.62	0.65	0.66	0.69	0.74	0.79	0.83	0.87	0.91	0.94	0.96	0.98	0.99	1.0

Figure C.8: Full experimental correlations between $\langle q^8 \rangle$ and $\langle q^8 \rangle$.

Central Moments (Full Correlation)

Figures C.9 to C.12 summarize the numerical values for the full experimental (statistical and systematic) correlation between the central q^2 moments of different order and lower q^2 threshold. The full correlation structure is shown in Figure 5.27.



$\langle (q^2 - \langle q^2 \rangle)^2 \rangle$

	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
1.5	0.4	0.38	0.37	0.35	0.32	0.29	0.28	0.27	0.26	0.24	0.24	0.22	0.22	0.2	0.19
2.0	0.52	0.47	0.44	0.4	0.36	0.33	0.31	0.29	0.28	0.26	0.25	0.23	0.23	0.21	0.18
2.5	0.6	0.54	0.5	0.45	0.41	0.36	0.35	0.33	0.31	0.29	0.28	0.25	0.24	0.21	0.19
3.0	0.64	0.58	0.53	0.48	0.43	0.39	0.37	0.35	0.33	0.3	0.28	0.26	0.24	0.21	0.18
3.5	0.67	0.61	0.57	0.51	0.47	0.42	0.4	0.37	0.35	0.32	0.3	0.26	0.24	0.2	0.17
4.0	0.72	0.66	0.62	0.57	0.53	0.48	0.45	0.42	0.39	0.36	0.33	0.28	0.25	0.2	0.16
4.5	0.76	0.72	0.69	0.64	0.6	0.55	0.52	0.48	0.45	0.41	0.37	0.31	0.26	0.2	0.15
5.0	0.79	0.77	0.74	0.7	0.66	0.61	0.58	0.54	0.5	0.44	0.39	0.32	0.25	0.18	0.12
5.5	0.8	0.8	0.78	0.75	0.71	0.67	0.63	0.59	0.54	0.47	0.41	0.32	0.24	0.15	0.07
6.0	0.81	0.82	0.81	0.78	0.75	0.71	0.67	0.62	0.57	0.49	0.41	0.31	0.2	0.1	0.01
6.5	0.79	0.81	0.82	0.8	0.78	0.74	0.7	0.65	0.59	0.5	0.4	0.28	0.16	0.04	-0.06
7.0	0.76	0.8	0.81	0.8	0.78	0.75	0.71	0.66	0.59	0.5	0.39	0.25	0.11	-0.02	-0.13
7.5	0.71	0.77	0.79	0.79	0.78	0.75	0.71	0.65	0.58	0.48	0.36	0.21	0.06	-0.08	-0.2
8.0	0.67	0.73	0.76	0.76	0.75	0.73	0.69	0.63	0.56	0.45	0.32	0.17	0.0	-0.15	-0.27
8.5	0.63	0.69	0.72	0.73	0.72	0.7	0.66	0.6	0.52	0.42	0.28	0.12	-0.05	-0.21	-0.34

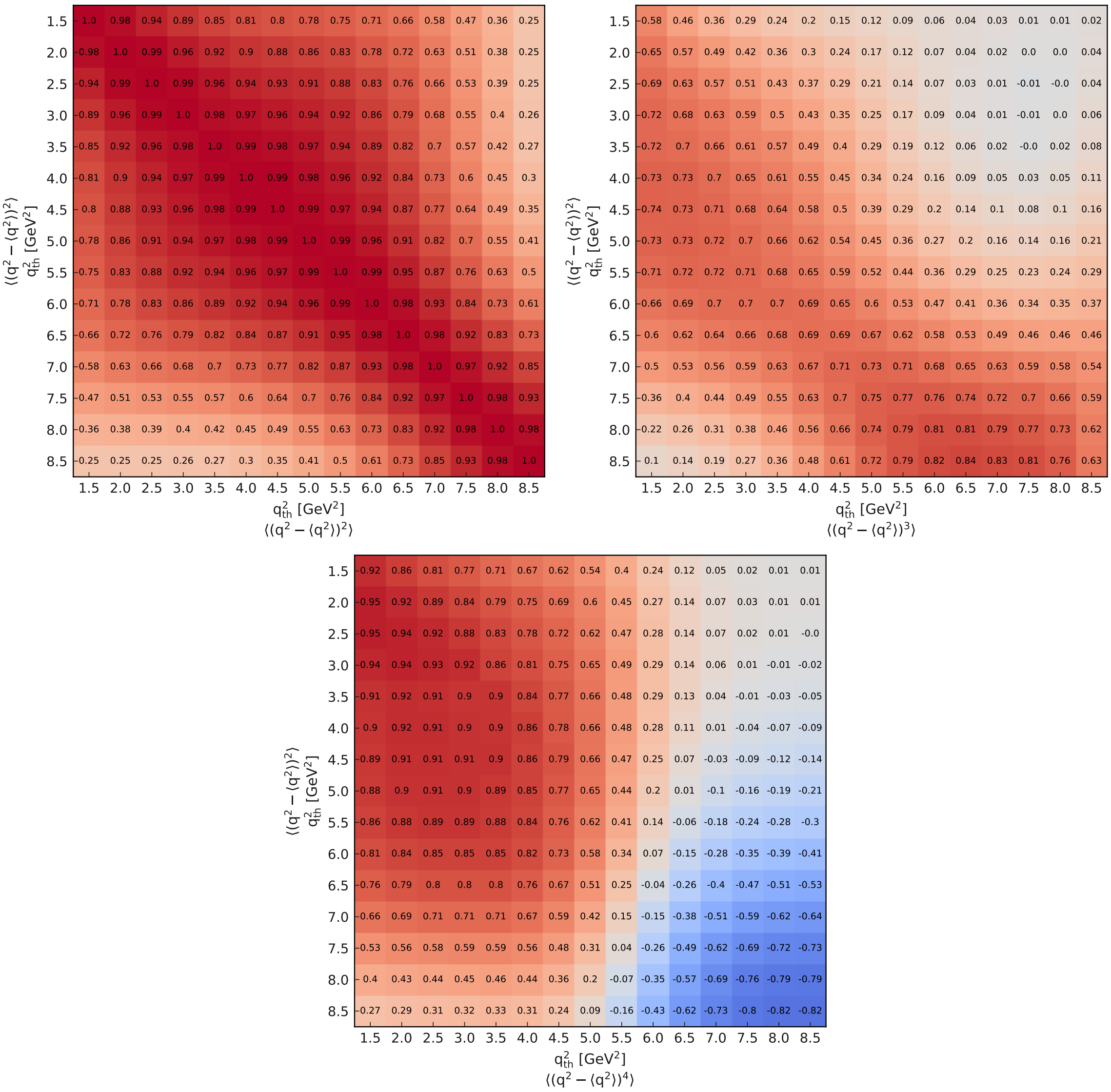
$\langle (q^2 - \langle q^2 \rangle)^3 \rangle$

	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
1.5	-0.26	-0.31	-0.34	-0.38	-0.42	-0.41	-0.37	-0.31	-0.26	-0.22	-0.19	-0.19	-0.2	-0.26	-0.36
2.0	-0.18	-0.27	-0.33	-0.39	-0.43	-0.42	-0.39	-0.33	-0.26	-0.22	-0.19	-0.19	-0.19	-0.25	-0.34
2.5	-0.1	-0.22	-0.3	-0.37	-0.41	-0.41	-0.37	-0.32	-0.25	-0.21	-0.17	-0.17	-0.18	-0.23	-0.32
3.0	-0.06	-0.19	-0.27	-0.35	-0.39	-0.4	-0.37	-0.32	-0.26	-0.21	-0.18	-0.18	-0.19	-0.24	-0.33
3.5	-0.02	-0.15	-0.24	-0.32	-0.36	-0.37	-0.35	-0.31	-0.26	-0.22	-0.19	-0.19	-0.2	-0.25	-0.33
4.0	0.04	-0.09	-0.19	-0.27	-0.31	-0.33	-0.32	-0.29	-0.25	-0.21	-0.19	-0.19	-0.2	-0.25	-0.33
4.5	0.11	-0.01	-0.1	-0.18	-0.24	-0.27	-0.27	-0.26	-0.23	-0.21	-0.19	-0.2	-0.22	-0.26	-0.34
5.0	0.19	0.08	-0.01	-0.1	-0.16	-0.2	-0.23	-0.23	-0.22	-0.21	-0.21	-0.22	-0.24	-0.28	-0.34
5.5	0.27	0.16	0.08	-0.0	-0.08	-0.13	-0.18	-0.21	-0.22	-0.22	-0.23	-0.25	-0.27	-0.3	-0.35
6.0	0.34	0.25	0.17	0.08	0.01	-0.06	-0.13	-0.18	-0.22	-0.24	-0.26	-0.28	-0.3	-0.33	-0.36
6.5	0.41	0.33	0.26	0.18	0.1	0.02	-0.08	-0.15	-0.21	-0.26	-0.29	-0.32	-0.34	-0.36	-0.37
7.0	0.47	0.41	0.34	0.26	0.18	0.08	-0.03	-0.13	-0.21	-0.27	-0.32	-0.35	-0.38	-0.39	-0.37
7.5	0.51	0.47	0.41	0.33	0.24	0.14	0.01	-0.11	-0.21	-0.29	-0.34	-0.38	-0.41	-0.41	-0.37
8.0	0.55	0.51	0.46	0.38	0.29	0.18	0.04	-0.09	-0.21	-0.29	-0.36	-0.4	-0.43	-0.42	-0.36
8.5	0.56	0.53	0.48	0.4	0.32	0.2	0.05	-0.09	-0.21	-0.31	-0.38	-0.42	-0.45	-0.43	-0.35

$\langle (q^2 - \langle q^2 \rangle)^4 \rangle$

	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
1.5	0.28	0.25	0.23	0.17	0.1	0.04	-0.03	-0.11	-0.21	-0.28	-0.29	-0.27	-0.22	-0.18	-0.14
2.0	0.37	0.32	0.28	0.22	0.15	0.09	0.03	-0.06	-0.16	-0.23	-0.25	-0.24	-0.19	-0.16	-0.11
2.5	0.45	0.38	0.33	0.27	0.2	0.14	0.08	-0.01	-0.11	-0.19	-0.21	-0.2	-0.16	-0.13	-0.09
3.0	0.48	0.41	0.36	0.3	0.22	0.16	0.11	0.02	-0.07	-0.16	-0.18	-0.18	-0.14	-0.11	-0.07
3.5	0.51	0.44	0.39	0.32	0.27	0.2	0.14	0.06	-0.04	-0.12	-0.15	-0.15	-0.11	-0.08	-0.04
4.0	0.56	0.49	0.44	0.38	0.32	0.26	0.2	0.11	0.02	-0.07	-0.1	-0.11	-0.07	-0.04	-0.0
4.5	0.62	0.56	0.51	0.45	0.39	0.33	0.27	0.19	0.09	0.0	-0.04	-0.05	-0.02	0.01	0.05
5.0	0.66	0.61	0.57	0.51	0.46	0.4	0.34	0.27	0.17	0.08	0.03	0.02	0.05	0.08	0.11
5.5	0.7	0.66	0.62	0.57	0.51	0.46	0.4	0.34	0.25	0.16	0.12	0.1	0.13	0.15	0.18
6.0	0.73	0.7	0.66	0.61	0.57	0.52	0.47	0.41	0.33	0.26	0.21	0.19	0.21	0.23	0.26
6.5	0.74	0.72	0.69	0.65	0.61	0.56	0.52	0.47	0.41	0.35	0.31	0.29	0.3	0.32	0.35
7.0	0.74	0.73	0.71	0.67	0.63	0.59	0.56	0.53	0.49	0.43	0.39	0.38	0.39	0.41	0.43
7.5	0.73	0.73	0.71	0.68	0.65	0.62	0.59	0.57	0.55	0.51	0.48	0.46	0.47	0.48	0.49
8.0	0.71	0.71	0.7	0.67	0.65	0.63	0.61	0.61	0.6	0.58	0.55	0.54	0.54	0.55	0.56
8.5	0.68	0.69	0.68	0.66	0.64	0.62	0.61	0.63	0.64	0.63	0.61	0.6	0.59	0.6	0.6

Figure C.9: Full experimental correlations between $\langle q^2 \rangle$ and $\langle (q^2 - \langle q^2 \rangle)^n \rangle$ for $n = 2$–4.

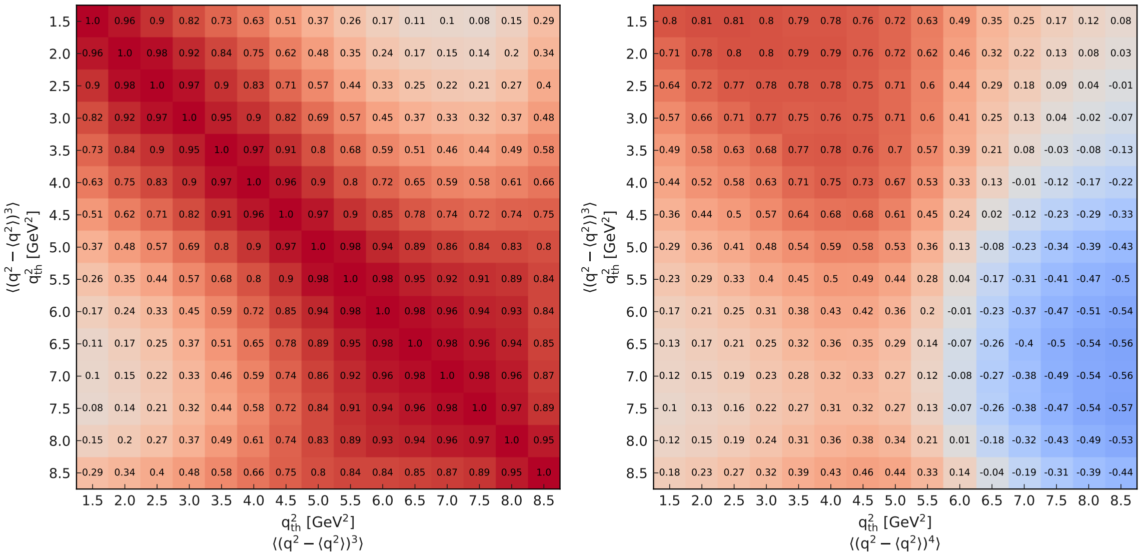


	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
1.5	1.0	0.98	0.94	0.89	0.85	0.81	0.8	0.78	0.75	0.71	0.66	0.58	0.47	0.36	0.25
2.0	0.98	1.0	0.99	0.96	0.92	0.9	0.88	0.86	0.83	0.78	0.72	0.63	0.51	0.38	0.25
2.5	0.94	0.99	1.0	0.99	0.96	0.94	0.93	0.91	0.88	0.83	0.76	0.66	0.53	0.39	0.25
3.0	0.89	0.96	0.99	1.0	0.98	0.97	0.96	0.94	0.92	0.86	0.79	0.68	0.55	0.4	0.26
3.5	0.85	0.92	0.96	0.98	1.0	0.99	0.98	0.97	0.94	0.89	0.82	0.7	0.57	0.42	0.27
4.0	0.81	0.9	0.94	0.97	0.99	1.0	0.99	0.98	0.96	0.92	0.84	0.73	0.6	0.45	0.3
4.5	0.8	0.88	0.93	0.96	0.98	0.99	1.0	0.99	0.97	0.94	0.87	0.77	0.64	0.49	0.35
5.0	0.78	0.86	0.91	0.94	0.97	0.98	0.99	1.0	0.99	0.96	0.91	0.82	0.7	0.55	0.41
5.5	0.75	0.83	0.88	0.92	0.94	0.96	0.97	0.99	1.0	0.99	0.95	0.87	0.76	0.63	0.5
6.0	0.71	0.78	0.83	0.86	0.89	0.92	0.94	0.96	0.99	1.0	0.98	0.93	0.84	0.73	0.61
6.5	0.66	0.72	0.76	0.79	0.82	0.84	0.87	0.91	0.95	0.98	1.0	0.98	0.92	0.83	0.73
7.0	0.58	0.63	0.66	0.68	0.7	0.73	0.77	0.82	0.87	0.93	0.98	1.0	0.97	0.92	0.85
7.5	0.47	0.51	0.53	0.55	0.57	0.6	0.64	0.7	0.76	0.84	0.92	0.97	1.0	0.98	0.93
8.0	0.36	0.38	0.39	0.4	0.42	0.45	0.49	0.55	0.63	0.73	0.83	0.92	0.98	1.0	0.98
8.5	0.25	0.25	0.25	0.26	0.27	0.3	0.35	0.41	0.5	0.61	0.73	0.85	0.93	0.98	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
1.5	0.58	0.46	0.36	0.29	0.24	0.2	0.15	0.12	0.09	0.06	0.04	0.03	0.01	0.01	0.02
2.0	0.65	0.57	0.49	0.42	0.36	0.3	0.24	0.17	0.12	0.07	0.04	0.02	0.0	0.0	0.04
2.5	0.69	0.63	0.57	0.51	0.43	0.37	0.29	0.21	0.14	0.07	0.03	0.01	-0.01	-0.0	0.04
3.0	0.72	0.68	0.63	0.59	0.5	0.43	0.35	0.25	0.17	0.09	0.04	0.01	-0.01	0.0	0.06
3.5	0.72	0.7	0.66	0.61	0.57	0.49	0.4	0.29	0.19	0.12	0.06	0.02	-0.0	0.02	0.08
4.0	0.73	0.73	0.7	0.65	0.61	0.55	0.45	0.34	0.24	0.16	0.09	0.05	0.03	0.05	0.11
4.5	0.74	0.73	0.71	0.68	0.64	0.58	0.5	0.39	0.29	0.2	0.14	0.1	0.08	0.1	0.16
5.0	0.73	0.73	0.72	0.7	0.66	0.62	0.54	0.45	0.36	0.27	0.2	0.16	0.14	0.16	0.21
5.5	0.71	0.72	0.72	0.71	0.68	0.65	0.59	0.52	0.44	0.36	0.29	0.25	0.23	0.24	0.29
6.0	0.66	0.69	0.7	0.7	0.7	0.69	0.65	0.6	0.53	0.47	0.41	0.36	0.34	0.35	0.37
6.5	0.6	0.62	0.64	0.66	0.68	0.69	0.69	0.67	0.62	0.58	0.53	0.49	0.46	0.46	0.46
7.0	0.5	0.53	0.56	0.59	0.63	0.67	0.71	0.73	0.71	0.68	0.65	0.63	0.59	0.58	0.54
7.5	0.36	0.4	0.44	0.49	0.55	0.63	0.7	0.75	0.77	0.76	0.74	0.72	0.7	0.66	0.59
8.0	0.22	0.26	0.31	0.38	0.46	0.56	0.66	0.74	0.79	0.81	0.81	0.79	0.77	0.73	0.62
8.5	0.1	0.14	0.19	0.27	0.36	0.48	0.61	0.72	0.79	0.82	0.84	0.83	0.81	0.76	0.63

	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
1.5	0.92	0.86	0.81	0.77	0.71	0.67	0.62	0.54	0.4	0.24	0.12	0.05	0.02	0.01	0.01
2.0	0.95	0.92	0.89	0.84	0.79	0.75	0.69	0.6	0.45	0.27	0.14	0.07	0.03	0.01	0.01
2.5	0.95	0.94	0.92	0.88	0.83	0.78	0.72	0.62	0.47	0.28	0.14	0.07	0.02	0.01	-0.0
3.0	0.94	0.94	0.93	0.92	0.86	0.81	0.75	0.65	0.49	0.29	0.14	0.06	0.01	-0.01	-0.02
3.5	0.91	0.92	0.91	0.9	0.9	0.84	0.77	0.66	0.48	0.29	0.13	0.04	-0.01	-0.03	-0.05
4.0	0.9	0.92	0.91	0.9	0.9	0.86	0.78	0.66	0.48	0.28	0.11	0.01	-0.04	-0.07	-0.09
4.5	0.89	0.91	0.91	0.91	0.9	0.86	0.79	0.66	0.47	0.25	0.07	-0.03	-0.09	-0.12	-0.14
5.0	0.88	0.9	0.91	0.9	0.89	0.85	0.77	0.65	0.44	0.2	0.01	-0.1	-0.16	-0.19	-0.21
5.5	0.86	0.88	0.89	0.89	0.88	0.84	0.76	0.62	0.41	0.14	-0.06	-0.18	-0.24	-0.28	-0.3
6.0	0.81	0.84	0.85	0.85	0.85	0.82	0.73	0.58	0.34	0.07	-0.15	-0.28	-0.35	-0.39	-0.41
6.5	0.76	0.79	0.8	0.8	0.8	0.76	0.67	0.51	0.25	-0.04	-0.26	-0.4	-0.47	-0.51	-0.53
7.0	0.66	0.69	0.71	0.71	0.71	0.67	0.59	0.42	0.15	-0.15	-0.38	-0.51	-0.59	-0.62	-0.64
7.5	0.53	0.56	0.58	0.59	0.59	0.56	0.48	0.31	0.04	-0.26	-0.49	-0.62	-0.69	-0.72	-0.73
8.0	0.4	0.43	0.44	0.45	0.46	0.44	0.36	0.2	-0.07	-0.35	-0.57	-0.69	-0.76	-0.79	-0.79
8.5	0.27	0.29	0.31	0.32	0.33	0.31	0.24	0.09	-0.16	-0.43	-0.62	-0.73	-0.8	-0.82	-0.82

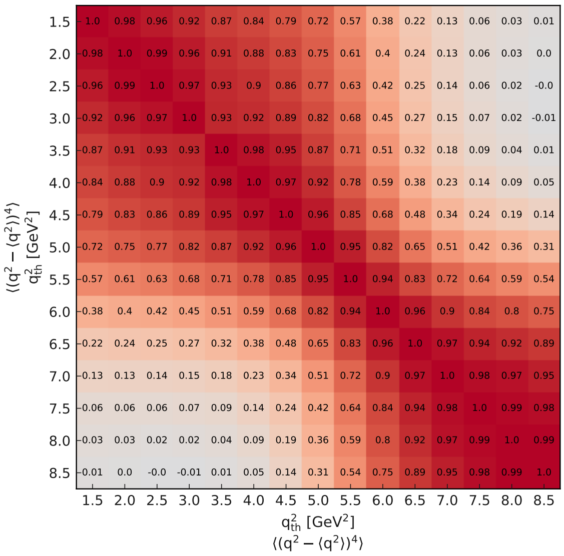
Figure C.10: Full experimental correlations between $\langle(q^2 - \langle q^2\rangle)^2\rangle$ and $\langle(q^2 - \langle q^2\rangle)^n\rangle$ for n = 2–4.



	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
1.5	1.0	0.96	0.9	0.82	0.73	0.63	0.51	0.37	0.26	0.17	0.11	0.1	0.08	0.15	0.29
2.0	0.96	1.0	0.98	0.92	0.84	0.75	0.62	0.48	0.35	0.24	0.17	0.15	0.14	0.2	0.34
2.5	0.9	0.98	1.0	0.97	0.9	0.83	0.71	0.57	0.44	0.33	0.25	0.22	0.21	0.27	0.4
3.0	0.82	0.92	0.97	1.0	0.95	0.9	0.82	0.69	0.57	0.45	0.37	0.33	0.32	0.37	0.48
3.5	0.73	0.84	0.9	0.95	1.0	0.97	0.91	0.8	0.68	0.59	0.51	0.46	0.44	0.49	0.58
4.0	0.63	0.75	0.83	0.9	0.97	1.0	0.96	0.9	0.8	0.72	0.65	0.59	0.58	0.61	0.66
4.5	0.51	0.62	0.71	0.82	0.91	0.96	1.0	0.97	0.9	0.85	0.78	0.74	0.72	0.74	0.75
5.0	0.37	0.48	0.57	0.69	0.8	0.9	0.97	1.0	0.98	0.94	0.89	0.86	0.84	0.83	0.8
5.5	0.26	0.35	0.44	0.57	0.68	0.8	0.9	0.98	1.0	0.98	0.95	0.92	0.91	0.89	0.84
6.0	0.17	0.24	0.33	0.45	0.59	0.72	0.85	0.94	0.98	1.0	0.98	0.96	0.94	0.93	0.84
6.5	0.11	0.17	0.25	0.37	0.51	0.65	0.78	0.89	0.95	0.98	1.0	0.98	0.96	0.94	0.85
7.0	0.1	0.15	0.22	0.33	0.46	0.59	0.74	0.86	0.92	0.96	0.98	1.0	0.98	0.96	0.87
7.5	0.08	0.14	0.21	0.32	0.44	0.58	0.72	0.84	0.91	0.94	0.96	0.98	1.0	0.97	0.89
8.0	0.15	0.2	0.27	0.37	0.49	0.61	0.74	0.83	0.89	0.93	0.94	0.96	0.97	1.0	0.95
8.5	0.29	0.34	0.4	0.48	0.58	0.66	0.75	0.8	0.84	0.84	0.85	0.87	0.89	0.95	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
1.5	0.8	0.81	0.81	0.8	0.79	0.78	0.76	0.72	0.63	0.49	0.35	0.25	0.17	0.12	0.08
2.0	0.71	0.78	0.8	0.8	0.79	0.79	0.76	0.72	0.62	0.46	0.32	0.22	0.13	0.08	0.03
2.5	0.64	0.72	0.77	0.78	0.78	0.78	0.75	0.71	0.6	0.44	0.29	0.18	0.09	0.04	-0.01
3.0	0.57	0.66	0.71	0.77	0.75	0.76	0.75	0.71	0.6	0.41	0.25	0.13	0.04	-0.02	-0.07
3.5	0.49	0.58	0.63	0.68	0.77	0.78	0.76	0.7	0.57	0.39	0.21	0.08	-0.03	-0.08	-0.13
4.0	0.44	0.52	0.58	0.63	0.71	0.75	0.73	0.67	0.53	0.33	0.13	-0.01	-0.12	-0.17	-0.22
4.5	0.36	0.44	0.5	0.57	0.64	0.68	0.68	0.61	0.45	0.24	0.02	-0.12	-0.23	-0.29	-0.33
5.0	0.29	0.36	0.41	0.48	0.54	0.59	0.58	0.53	0.36	0.13	-0.08	-0.23	-0.34	-0.39	-0.43
5.5	0.23	0.29	0.33	0.4	0.45	0.5	0.49	0.44	0.28	0.04	-0.17	-0.31	-0.41	-0.47	-0.5
6.0	0.17	0.21	0.25	0.31	0.38	0.43	0.42	0.36	0.2	-0.01	-0.23	-0.37	-0.47	-0.51	-0.54
6.5	0.13	0.17	0.21	0.25	0.32	0.36	0.35	0.29	0.14	-0.07	-0.26	-0.4	-0.5	-0.54	-0.56
7.0	0.12	0.15	0.19	0.23	0.28	0.32	0.33	0.27	0.12	-0.08	-0.27	-0.38	-0.49	-0.54	-0.56
7.5	0.1	0.13	0.16	0.22	0.27	0.31	0.32	0.27	0.13	-0.07	-0.26	-0.38	-0.47	-0.54	-0.57
8.0	0.12	0.15	0.19	0.24	0.31	0.36	0.38	0.34	0.21	0.01	-0.18	-0.32	-0.43	-0.49	-0.53
8.5	0.18	0.23	0.27	0.32	0.39	0.43	0.46	0.44	0.33	0.14	-0.04	-0.19	-0.31	-0.39	-0.44

Figure C.11: Full experimental correlations between $\langle (q^2 - \langle q^2 \rangle)^3 \rangle$ and $\langle (q^2 - \langle q^2 \rangle)^n \rangle$ for $n = 3, 4$.



	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
1.5	1.0	0.98	0.96	0.92	0.87	0.84	0.79	0.72	0.57	0.38	0.22	0.13	0.06	0.03	0.01
2.0	0.98	1.0	0.99	0.96	0.91	0.88	0.83	0.75	0.61	0.4	0.24	0.13	0.06	0.03	0.0
2.5	0.96	0.99	1.0	0.97	0.93	0.9	0.86	0.77	0.63	0.42	0.25	0.14	0.06	0.02	-0.0
3.0	0.92	0.96	0.97	1.0	0.93	0.92	0.89	0.82	0.68	0.45	0.27	0.15	0.07	0.02	-0.01
3.5	0.87	0.91	0.93	0.93	1.0	0.98	0.95	0.87	0.71	0.51	0.32	0.18	0.09	0.04	0.01
4.0	0.84	0.88	0.9	0.92	0.98	1.0	0.97	0.92	0.78	0.59	0.38	0.23	0.14	0.09	0.05
4.5	0.79	0.83	0.86	0.89	0.95	0.97	1.0	0.96	0.85	0.68	0.48	0.34	0.24	0.19	0.14
5.0	0.72	0.75	0.77	0.82	0.87	0.92	0.96	1.0	0.95	0.82	0.65	0.51	0.42	0.36	0.31
5.5	0.57	0.61	0.63	0.68	0.71	0.78	0.85	0.95	1.0	0.94	0.83	0.72	0.64	0.59	0.54
6.0	0.38	0.4	0.42	0.45	0.51	0.59	0.68	0.82	0.94	1.0	0.96	0.9	0.84	0.8	0.75
6.5	0.22	0.24	0.25	0.27	0.32	0.38	0.48	0.65	0.83	0.96	1.0	0.97	0.94	0.92	0.89
7.0	0.13	0.13	0.14	0.15	0.18	0.23	0.34	0.51	0.72	0.9	0.97	1.0	0.98	0.97	0.95
7.5	0.06	0.06	0.06	0.07	0.09	0.14	0.24	0.42	0.64	0.84	0.94	0.98	1.0	0.99	0.98
8.0	0.03	0.03	0.02	0.02	0.04	0.09	0.19	0.36	0.59	0.8	0.92	0.97	0.99	1.0	0.99
8.5	0.01	0.0	-0.0	-0.01	0.01	0.05	0.14	0.31	0.54	0.75	0.89	0.95	0.98	0.99	1.0

Figure C.12: Full experimental correlations between $\langle (q^2 - \langle q^2 \rangle)^4 \rangle$ and $\langle (q^2 - \langle q^2 \rangle)^4 \rangle$.

APPENDIX D

Additional Material Stability Checks

Measurement on Independent Data Samples

Figures D.1 and D.2 show the first second to fourth q^2 moments as functions of the lower q^2 threshold measured on sub-samples split by lepton flavor and B charge, respectively. Here, the considered uncertainties only include the statistical uncertainty on the moments and the uncertainties on the calibration coefficients and bias corrections factors. The latter are also considered to be independent since different MC sub-samples are used in their determination.

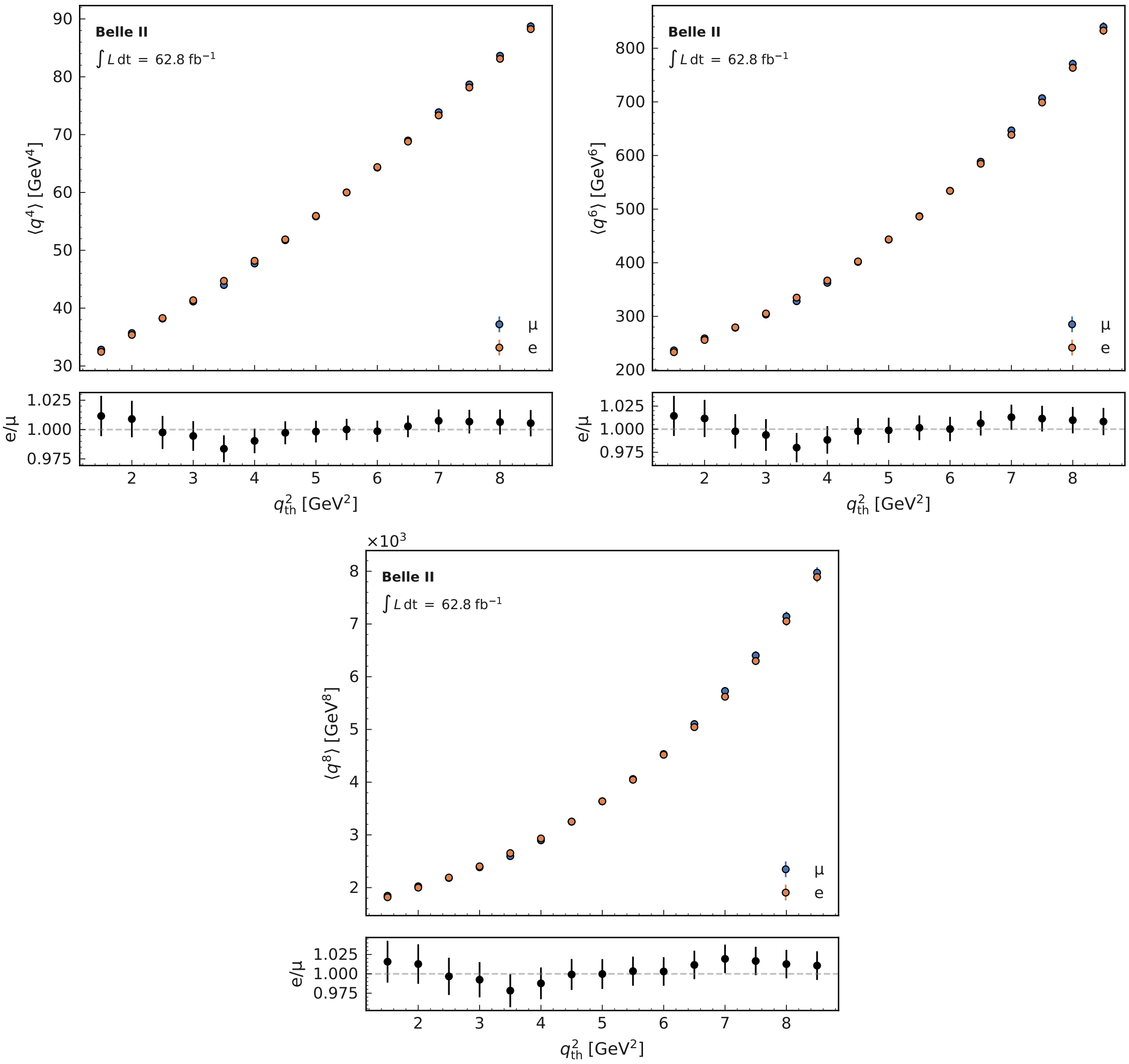


Figure D.1: Comparison of the second to fourth raw q^2 moments as functions of the lower q^2 threshold measured on statistically independent muon and electron sub-samples.

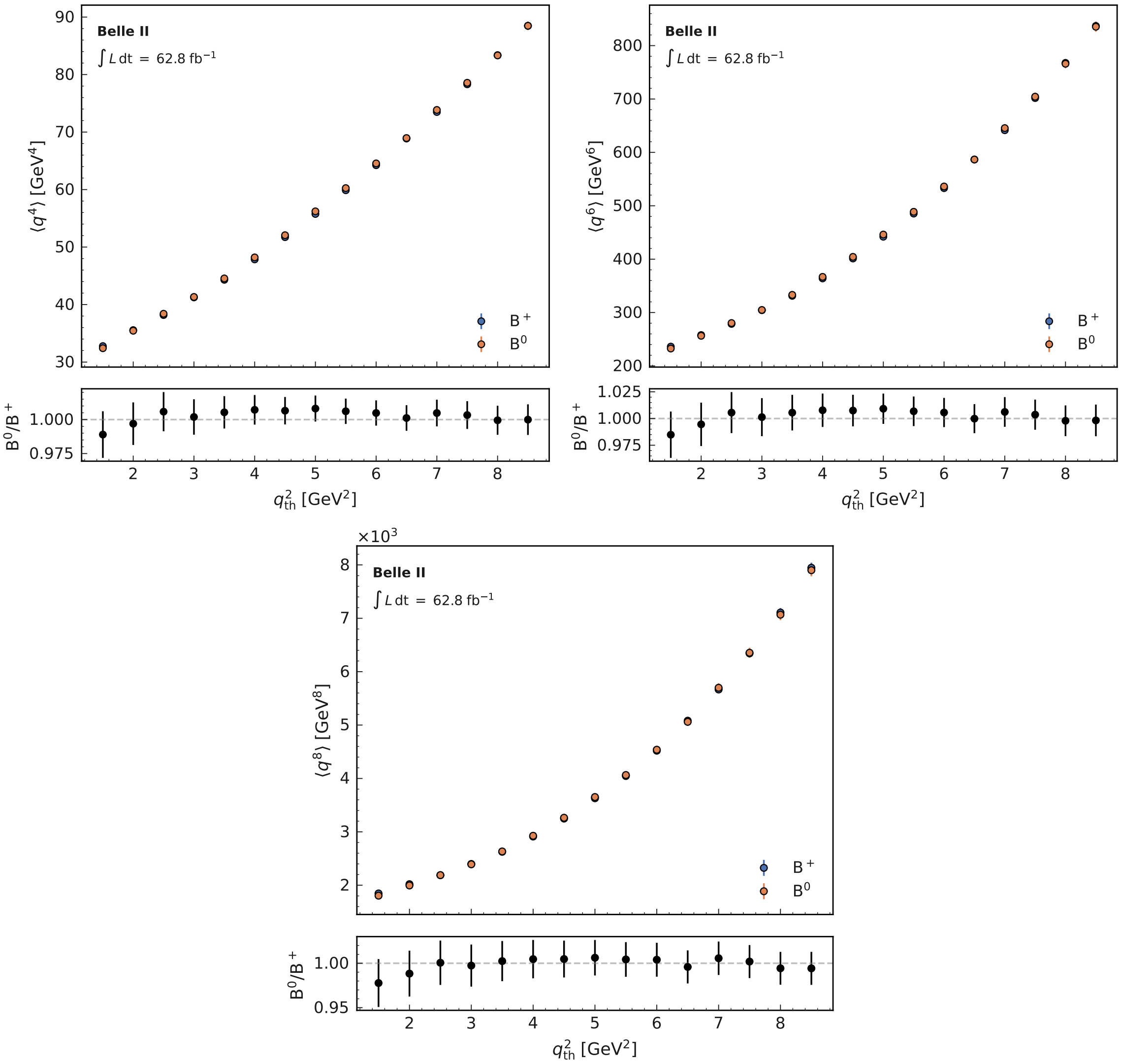


Figure D.2: Comparison of the second to fourth raw q^2 moments as functions of the lower q^2 threshold measured on statistically independent B^0 and B^+ sub-samples.

Robustness Against Altered q^2 Shapes

Figures D.3 to D.5 show the relative bias of the second to fourth raw q^2 moments extracted from the modified mock data sets compared using the nominal analysis procedure. The nominal the X_c modeling and background subtraction uncertainty budget of the nominal measurement is also given as reference. The observed remaining bias between the unfolded and generator level moments are covered by the uncertainties assumed in the measurement.

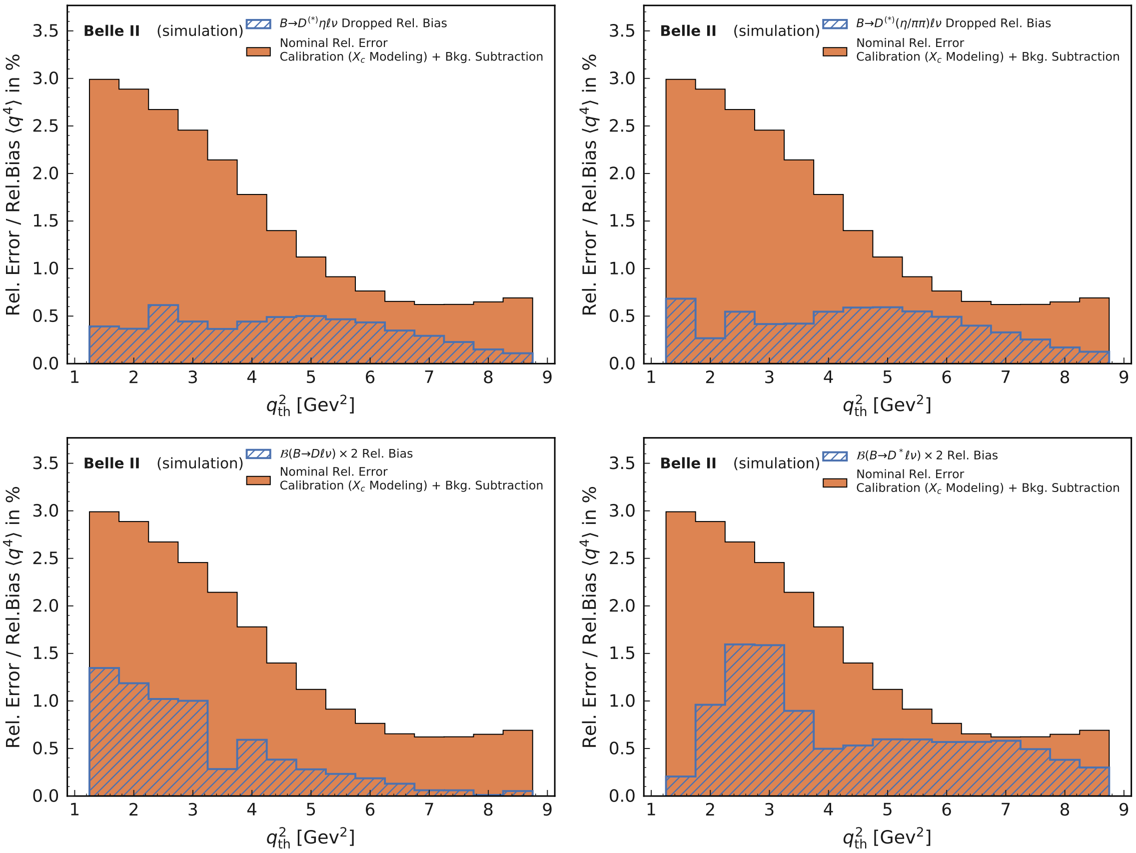


Figure D.3: Relative bias of the extracted second q^2 moments from the different mock data set with a modified q^2 shape compared to the X_c modeling and background subtraction uncertainty budget of the nominal result.

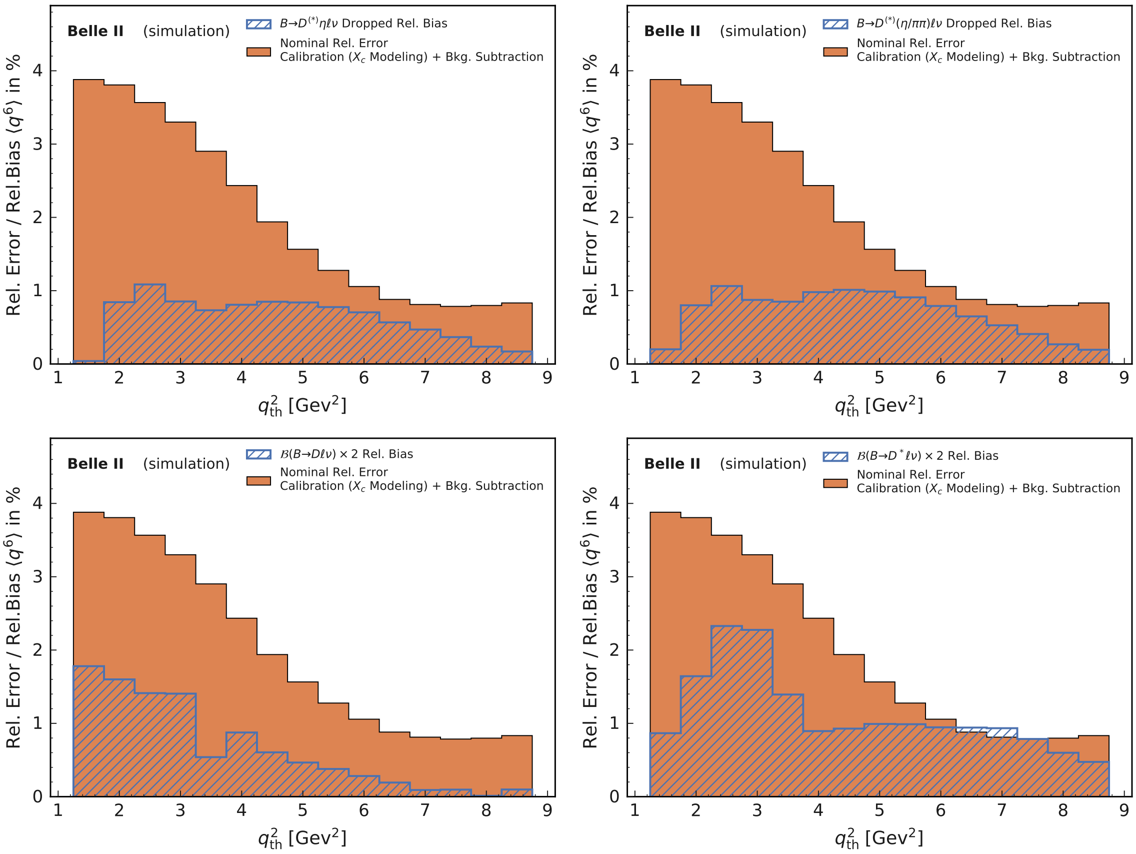


Figure D.4: Relative bias of the extracted third q^2 moments from the different mock data set with a modified q^2 shape compared to the X_c modeling and background subtraction uncertainty budget of the nominal result.

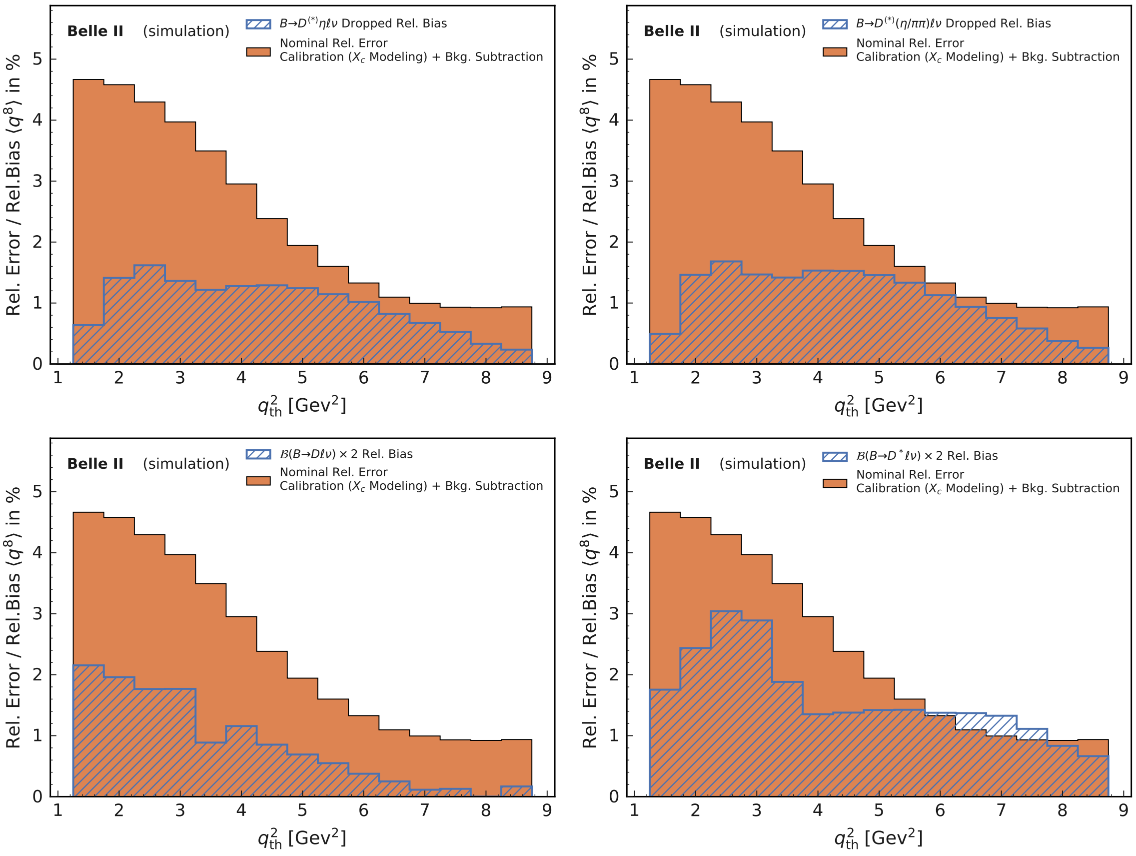


Figure D.5: Relative bias of the extracted fourth q^2 moments from the different mock data set with a modified q^2 shape compared to the X_c modeling and background subtraction uncertainty budget of the nominal result.

APPENDIX E

Additional Material First Extraction of Inclusive $|V_{cb}|$ from q^2 Moments

Toy Fits

This appendix presents the pull distribution for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 obtained by fits to pseudo-experiments generated with separated contributions to the total covariance matrix.

Figures E.1 and E.2 show the pulls obtained from toy moments generated with the experimental and theory covariance matrices, respectively. Here, all distributions follow a Gaussian distribution.

The effects of variations of the input branching fraction within the experimental uncertainty and the theory uncertainty on the prediction for the total rate are illustrated in Figures E.3 and E.4, respectively. These variations only contribute to the uncertainty on $|V_{cb}|$.

Lastly, the pull distributions for variations of the mean values for the constraints of m_b^{kin}, $\overline{m}_c$, μ_G^2, and μ_π^2 are shown in Figure E.5. Here, the distributions do not follow a Gaussian distribution.

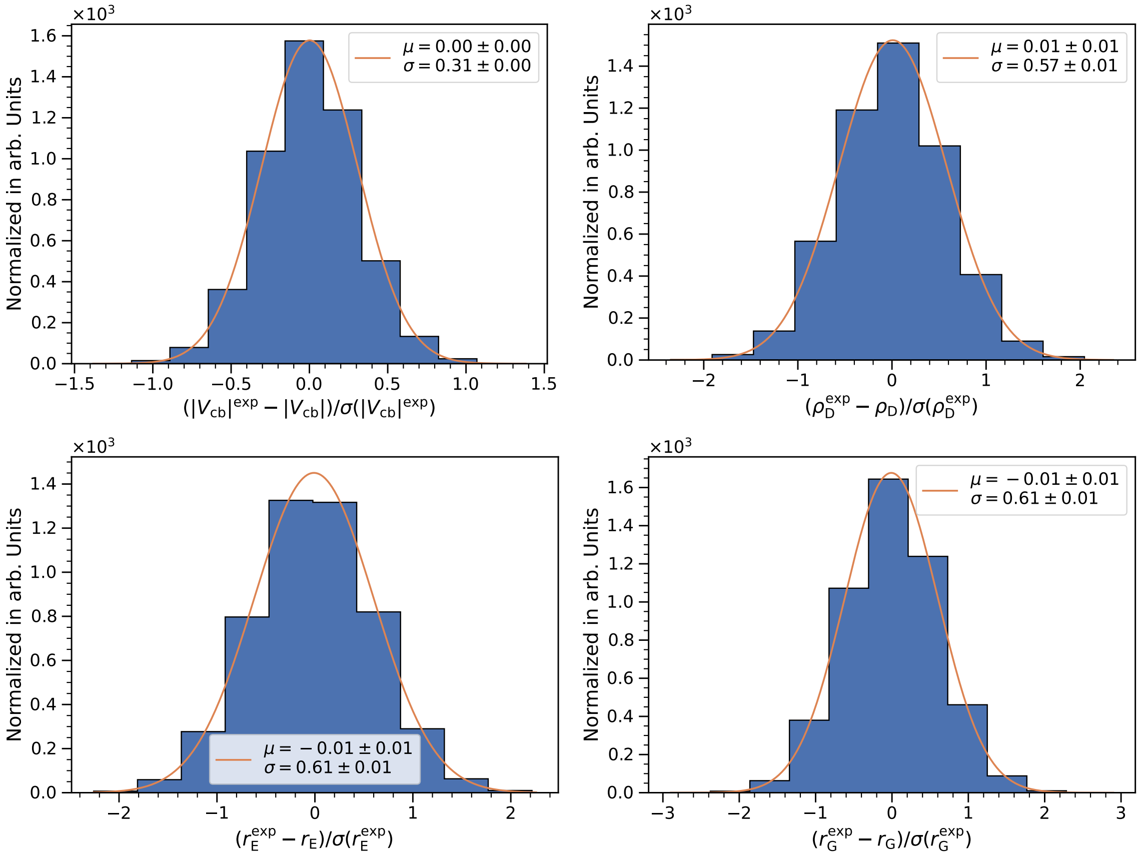


Figure E.1: Pull distributions for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 calculated from fits to toy data sets sampled using the experimental q^2 moment covariance matrix.

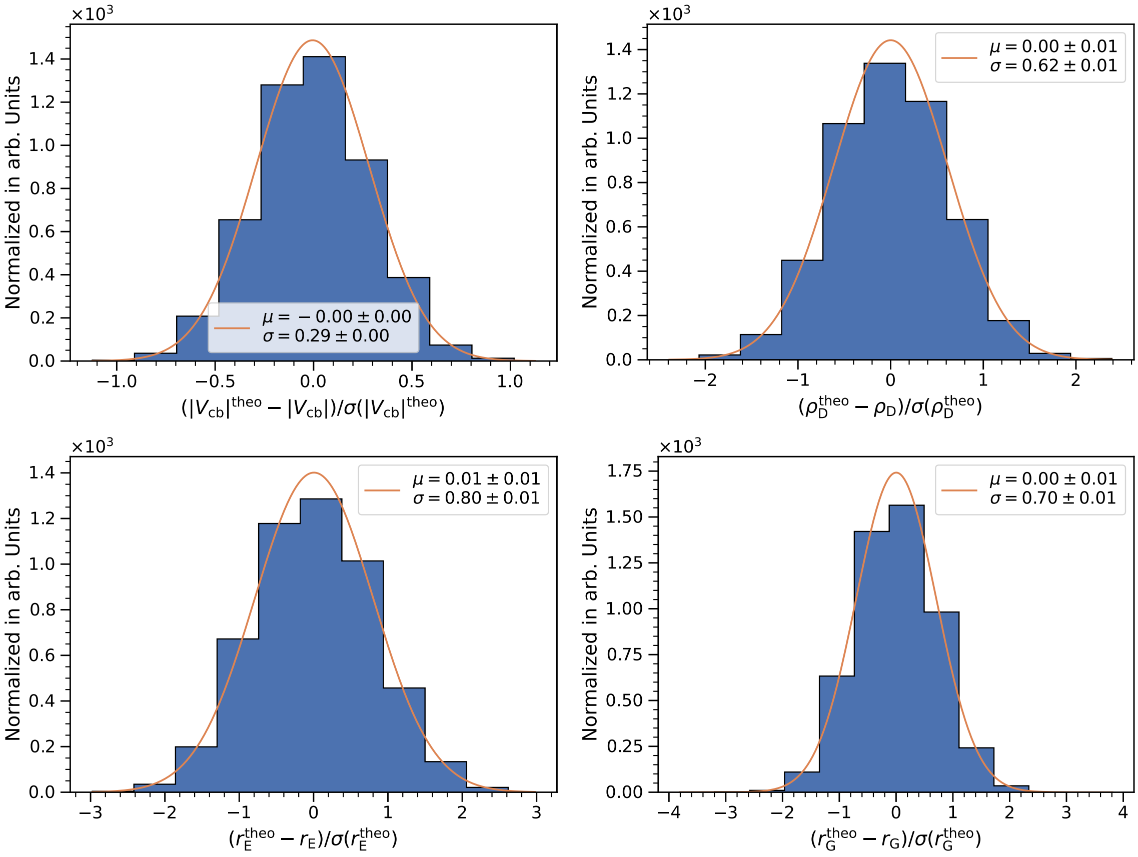


Figure E.2: Pull distributions for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 calculated from fits to toy data sets sampled using the theory q^2 moment covariance matrix.

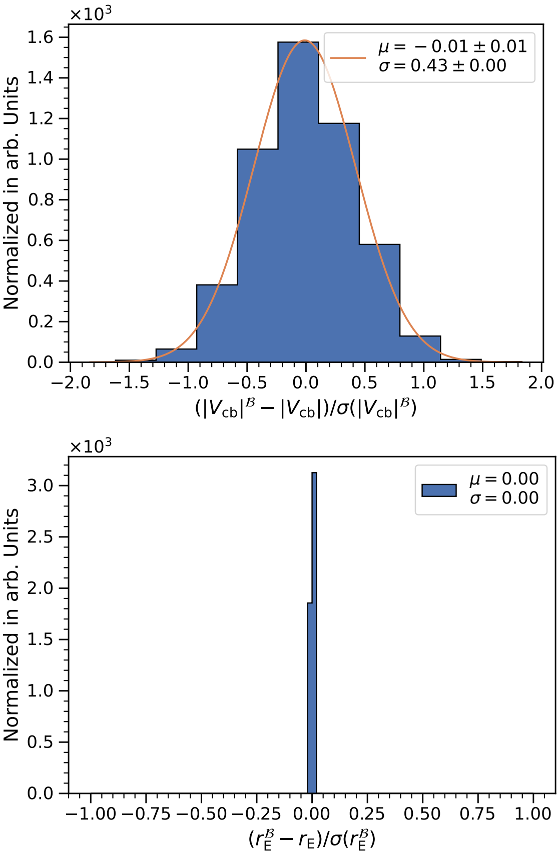


Figure E.3: Pull distributions for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 calculated from fits to toy data sets sampled using variations of the input $B \to X_c \ell \nu_\ell$ branching fractios.

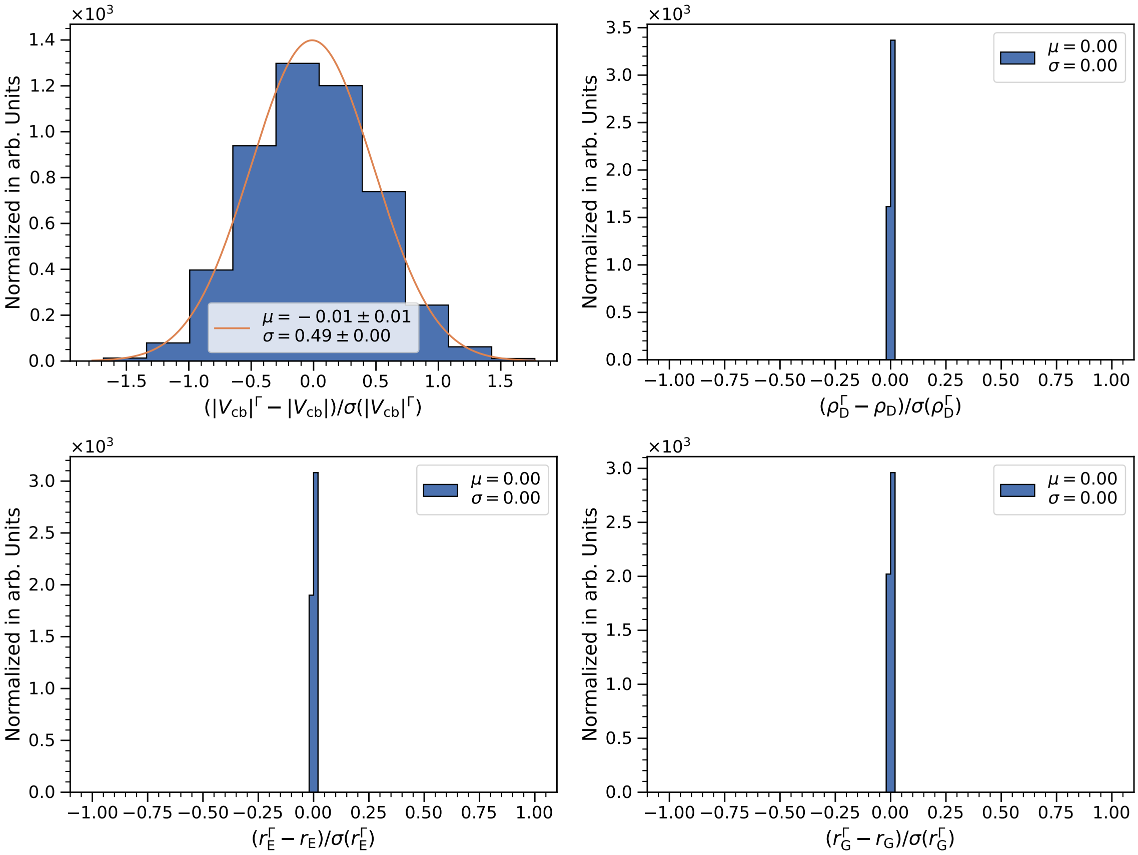


Figure E.4: Pull distributions for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 calculated from fits to toy data sets sampled using variations of the theory prediction for the total rate.

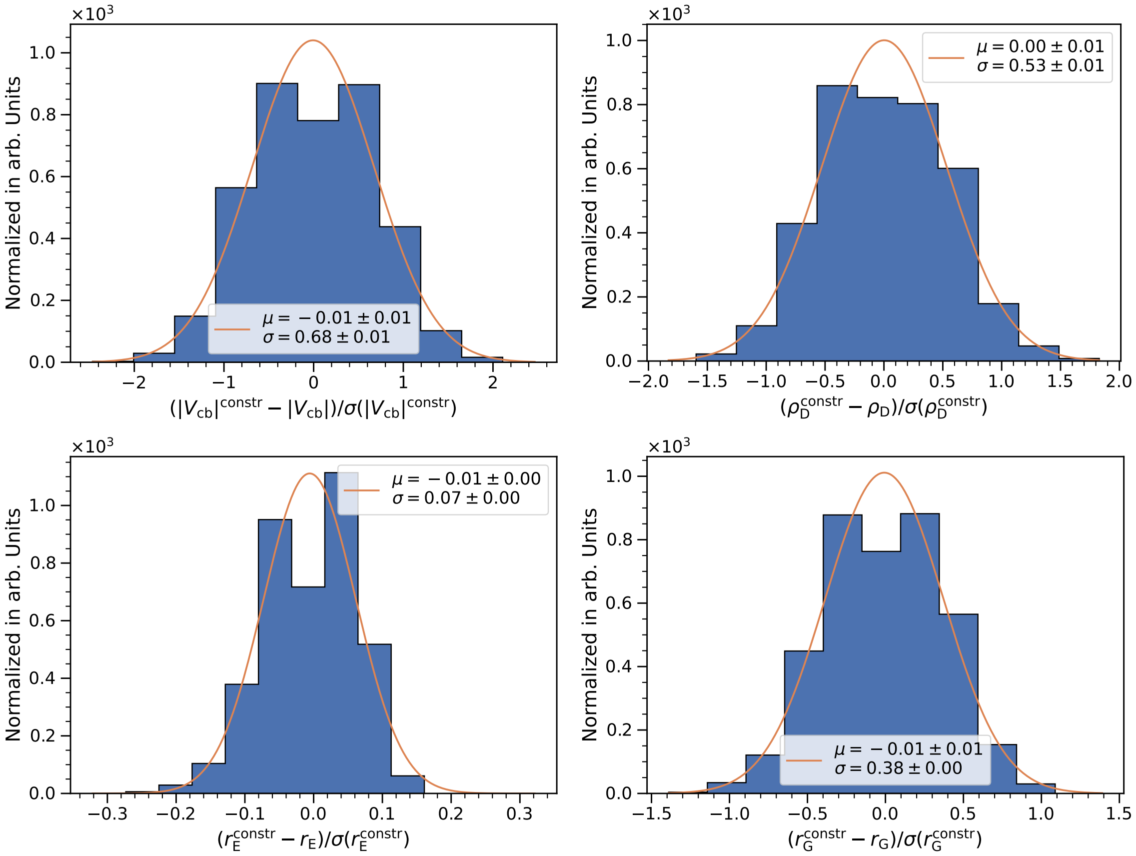


Figure E.5: Pull distributions for $|V_{cb}|$, ρ_D^3, r_E^4, and r_G^4 calculated from fits to toy data sets sampled using variations of the constraint mean values.

Belle Only Results

In this appendix, we present the results using Belle results only. The fit result is listed in Table E.1. Figure E.6 show the post-fit projections of the fitted theory predictions for the moments. The correlation matrix for the fitted parameters is illustrated in Figure E.7.

Table E.1: Results of our default fit using only Belle data for $|V_{cb}|$, $m_b^{\mathrm{kin}}(1\,\mathrm{GeV})$, $\overline{m}_c(2\,\mathrm{GeV})$, the HQE parameters, and the correlation parameters ρ_{cut} and ρ_{mom}. All parameters are expressed in GeV at the appropriate power.

	$\lvert V_{cb}\rvert \times 10^3$	m_b^{kin}	$\overline{m}_c$	μ_G^2	μ_π^2	ρ_D^3	r_G^4	$r_E^4 \times 10$	ρ_{cut}	ρ_{mom}
Value	41.54	4.56	1.09	0.36	0.43	0.12	-0.27	0.01	0.62	0.02
Uncertainty	0.66	0.02	0.01	0.07	0.24	0.28	0.99	0.48	-	-

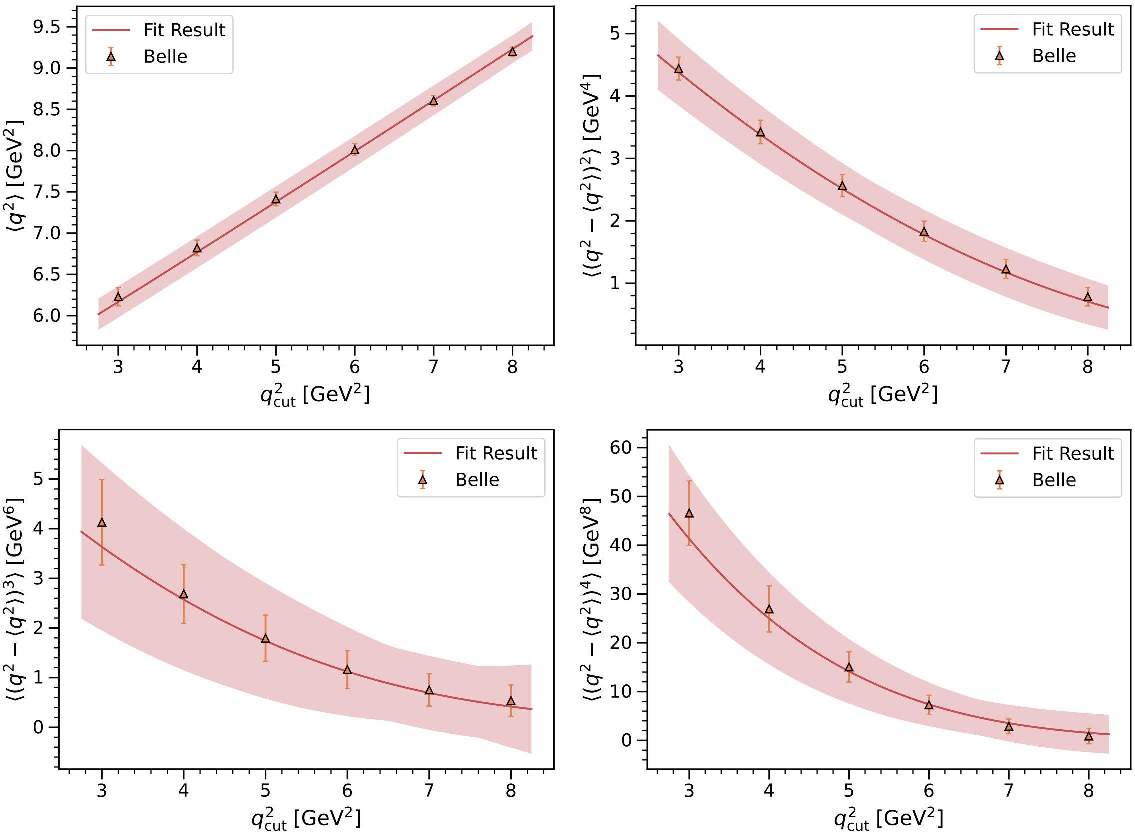
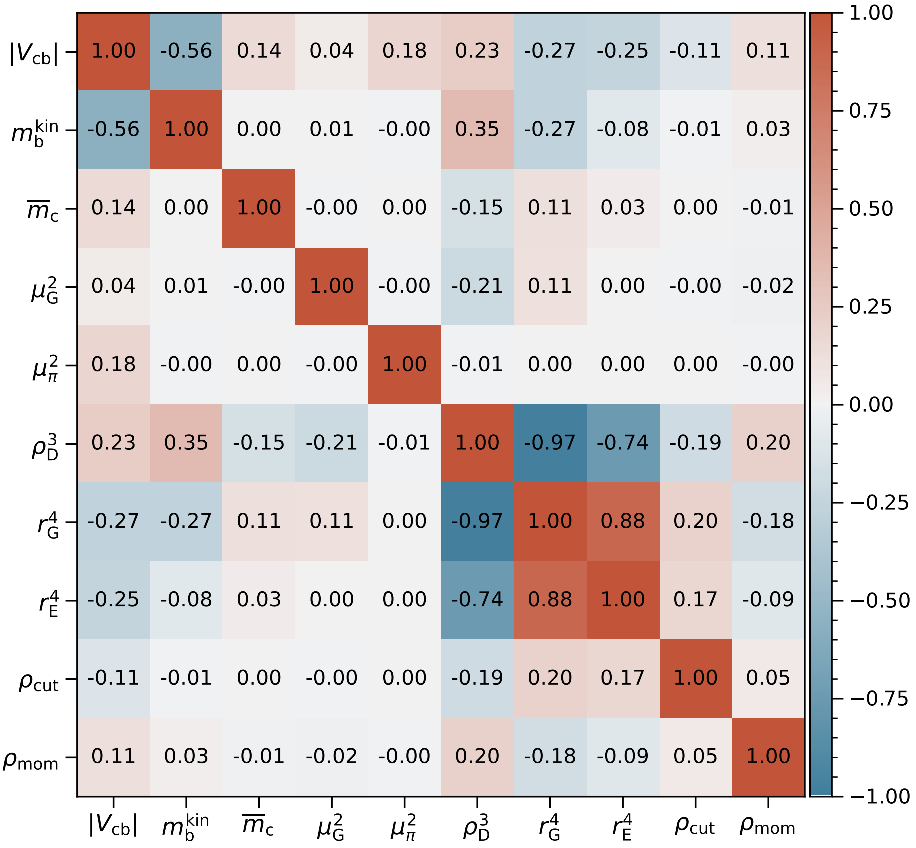


Figure E.6: Fit projections for q^2 moments using Belle data only.



	$\vert V_{\mathrm{cb}}\vert$	$m_{\mathrm{b}}^{\mathrm{kin}}$	$\overline{m}_{\mathrm{c}}$	μ_{G}^2	μ_{π}^2	ρ_{D}^3	r_{G}^4	r_{E}^4	ρ_{cut}	ρ_{mom}
$\vert V_{\mathrm{cb}}\vert$	1.00	-0.56	0.14	0.04	0.18	0.23	-0.27	-0.25	-0.11	0.11
$m_{\mathrm{b}}^{\mathrm{kin}}$	-0.56	1.00	0.00	0.01	-0.00	0.35	-0.27	-0.08	-0.01	0.03
$\overline{m}_{\mathrm{c}}$	0.14	0.00	1.00	-0.00	0.00	-0.15	0.11	0.03	0.00	-0.01
μ_{G}^2	0.04	0.01	-0.00	1.00	-0.00	-0.21	0.11	0.00	-0.00	-0.02
μ_{π}^2	0.18	-0.00	0.00	-0.00	1.00	-0.01	0.00	0.00	0.00	-0.00
ρ_{D}^3	0.23	0.35	-0.15	-0.21	-0.01	1.00	-0.97	-0.74	-0.19	0.20
r_{G}^4	-0.27	-0.27	0.11	0.11	0.00	-0.97	1.00	0.88	0.20	-0.18
r_{E}^4	-0.25	-0.08	0.03	0.00	0.00	-0.74	0.88	1.00	0.17	-0.09
ρ_{cut}	-0.11	-0.01	0.00	-0.00	0.00	-0.19	0.20	0.17	1.00	0.05
ρ_{mom}	0.11	0.03	-0.01	-0.02	-0.00	0.20	-0.18	-0.09	0.05	1.00

Figure E.7: Correlation matrix for $|V_{cb}|$, the HQE parameters, and the correlation parameters ρ_{mom} and ρ_{cut} using Belle data only.

Belle II Only Results

In this appendix, we present the results using Belle II results only. The fit result is listed in Table E.2. Figure E.8 show the post-fit projections of the fitted theory predictions for the moments. The correlation matrix for the fitted parameters is illustrated in Figure E.9.

Table E.2: Results of our default fit using only Belle II data for $|V_{cb}|$, $m_b^{\rm kin}(1\,{\rm GeV})$, $\overline{m}_c(2\,{\rm GeV})$, the HQE parameters, and the correlation parameters $\rho_{\rm cut}$ and $\rho_{\rm mom}$. All parameters are expressed in GeV at the appropriate power.

	$\lvert V_{cb}\rvert \times 10^3$	$m_b^{\rm kin}$	$\overline{m}_c$	μ_G^2	μ_π^2	ρ_D^3	r_G^4	$r_E^4 \times 10$	$\rho_{\rm cut}$	$\rho_{\rm mom}$
Value	41.70	4.56	1.09	0.36	0.43	0.12	-0.17	0.02	0.47	-0.38
Uncertainty	0.65	0.02	0.01	0.07	0.24	0.30	1.08	0.53	-	-

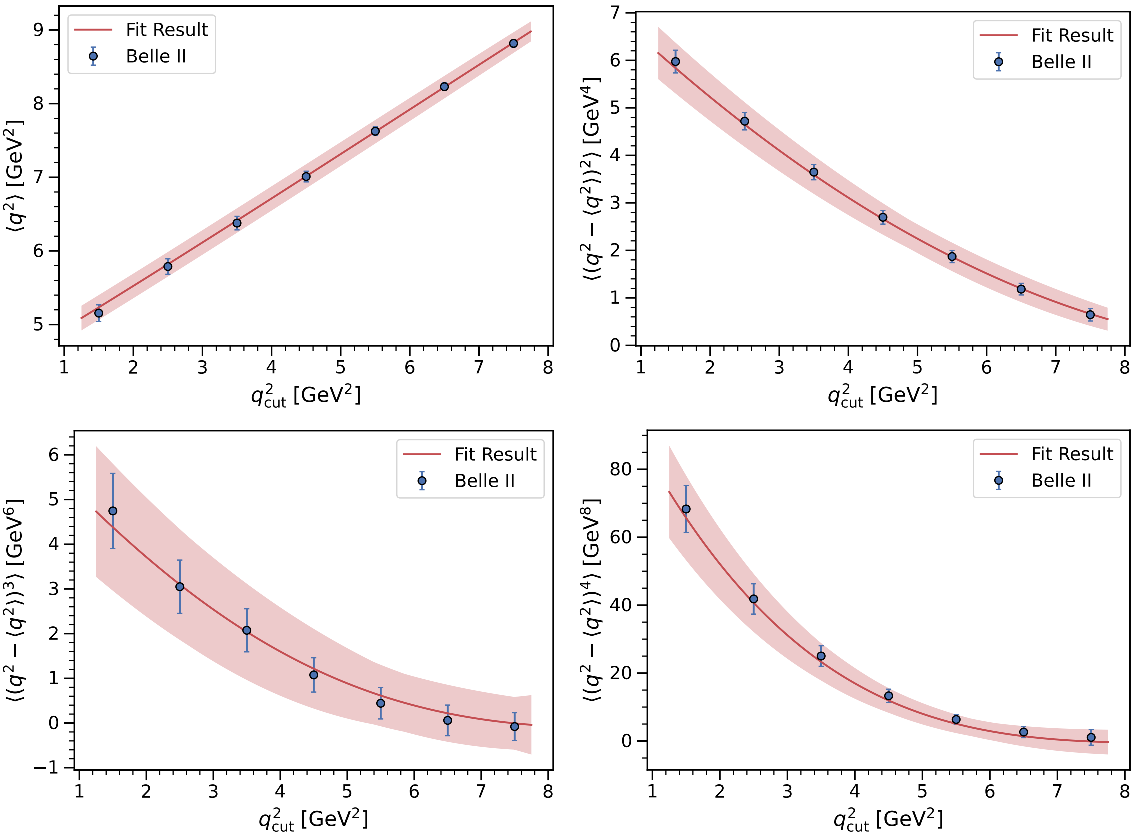
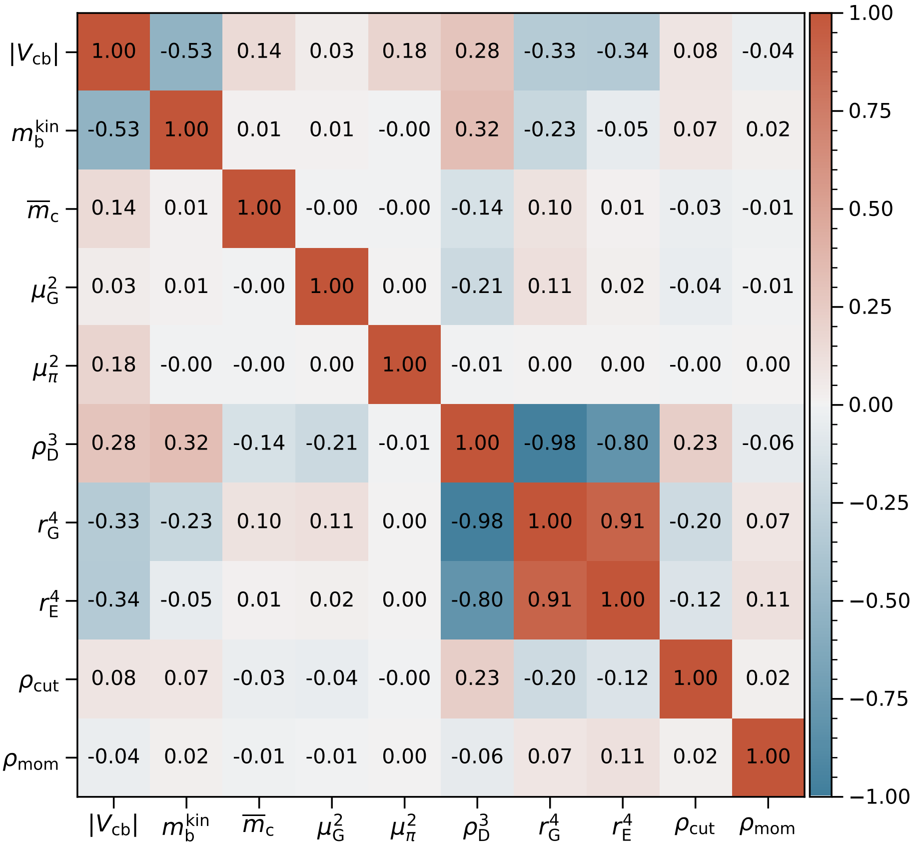


Figure E.8: Fit projections for q^2 moments using Belle II data only.



	$\lvert V_{\mathrm{cb}}\rvert$	$m_{\mathrm{b}}^{\mathrm{kin}}$	$\overline{m}_{\mathrm{c}}$	μ_{G}^2	μ_π^2	ρ_{D}^3	r_{G}^4	r_{E}^4	ρ_{cut}	ρ_{mom}
$\lvert V_{\mathrm{cb}}\rvert$	1.00	-0.53	0.14	0.03	0.18	0.28	-0.33	-0.34	0.08	-0.04
$m_{\mathrm{b}}^{\mathrm{kin}}$	-0.53	1.00	0.01	0.01	-0.00	0.32	-0.23	-0.05	0.07	0.02
$\overline{m}_{\mathrm{c}}$	0.14	0.01	1.00	-0.00	-0.00	-0.14	0.10	0.01	-0.03	-0.01
μ_{G}^2	0.03	0.01	-0.00	1.00	0.00	-0.21	0.11	0.02	-0.04	-0.01
μ_π^2	0.18	-0.00	-0.00	0.00	1.00	-0.01	0.00	0.00	-0.00	0.00
ρ_{D}^3	0.28	0.32	-0.14	-0.21	-0.01	1.00	-0.98	-0.80	0.23	-0.06
r_{G}^4	-0.33	-0.23	0.10	0.11	0.00	-0.98	1.00	0.91	-0.20	0.07
r_{E}^4	-0.34	-0.05	0.01	0.02	0.00	-0.80	0.91	1.00	-0.12	0.11
ρ_{cut}	0.08	0.07	-0.03	-0.04	-0.00	0.23	-0.20	-0.12	1.00	0.02
ρ_{mom}	-0.04	0.02	-0.01	-0.01	0.00	-0.06	0.07	0.11	0.02	1.00

Figure E.9: Correlation matrix for $|V_{cb}|$, the HQE parameters, and the correlation parameters ρ_{mom} and ρ_{cut} using Belle II data only.

Scan over the correlation parameters

In this appendix, we perform our default fit but for fixed combinations of ρ_{cut} and ρ_{mom}. The scans are performed with fits using Belle and Belle II data separately, and combined. The fit setup is identical to the default fit scenario describe in Section 6.2. The obtained $|V_{cb}|$ values are shown in Figure E.10. The scans for the HQE parameters ρ_D^3, r_E^4, and r_G^4 are illustrated in Figures E.11 to E.13, respectively.

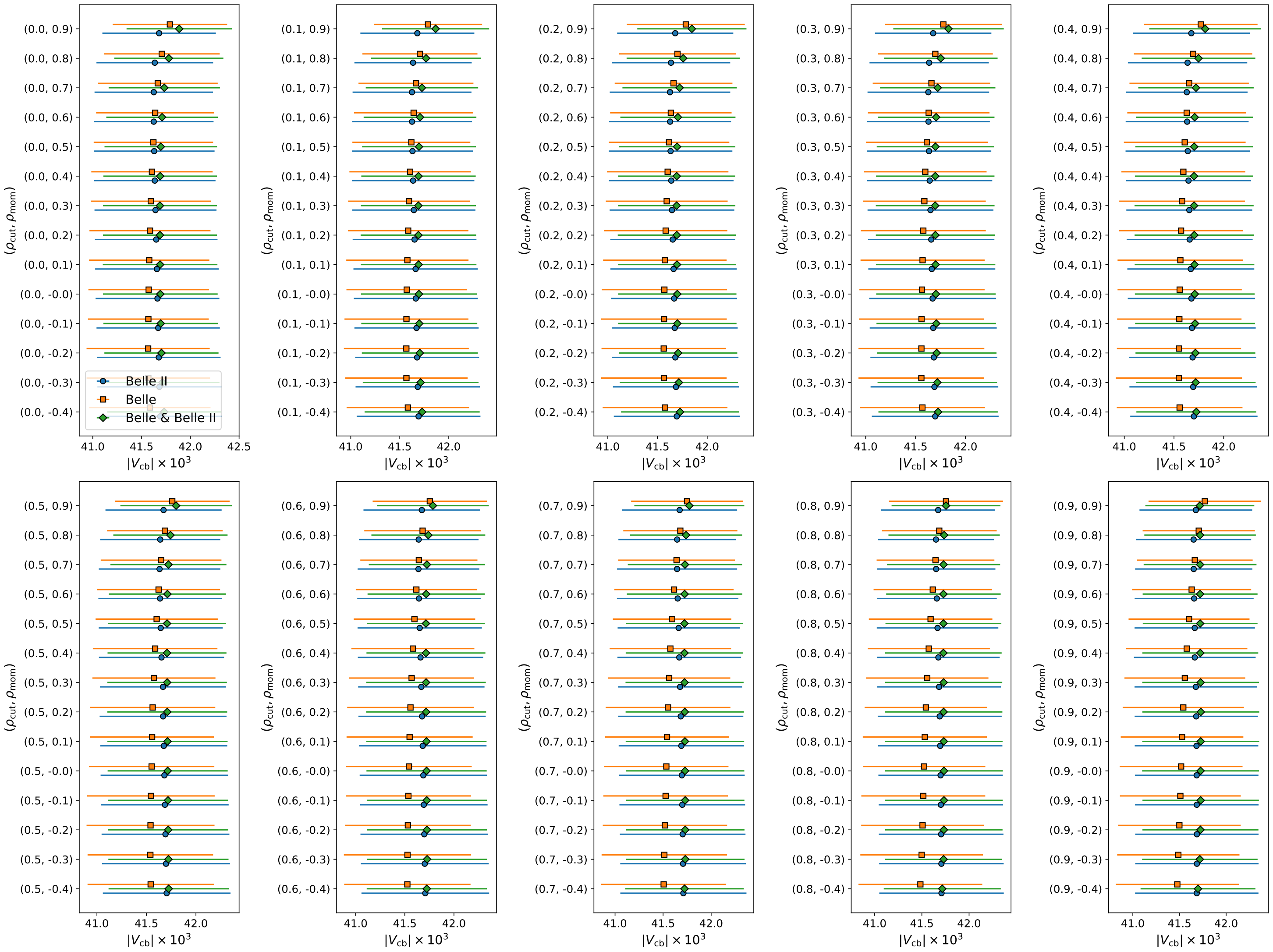


Figure E.10: Fit results for $|V_{cb}| \times 10^3$ for different combinations of correlation parameter values.

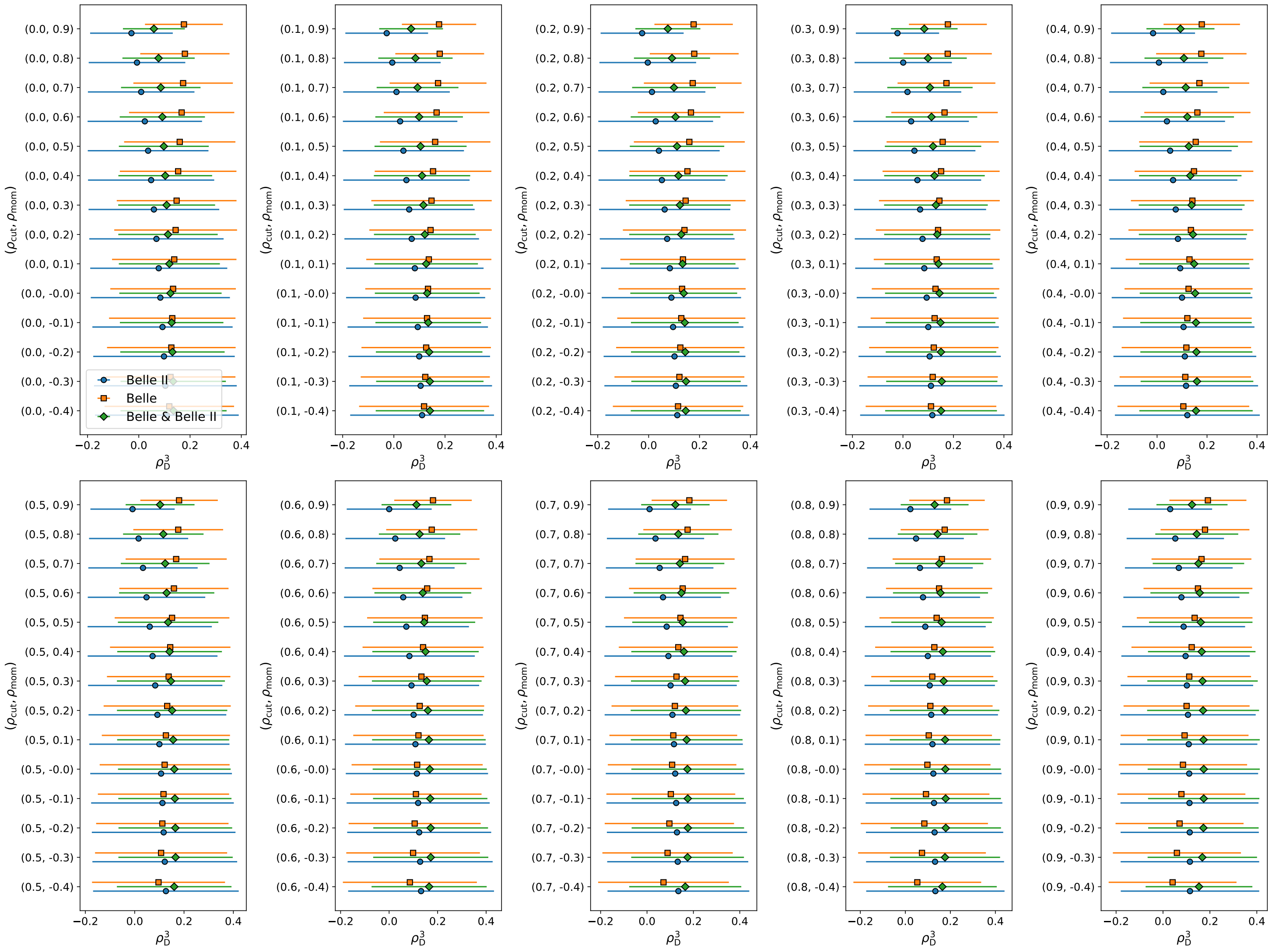


Figure E.11: Fit results for ρ_D^3 for different combinations of correlation parameter values.

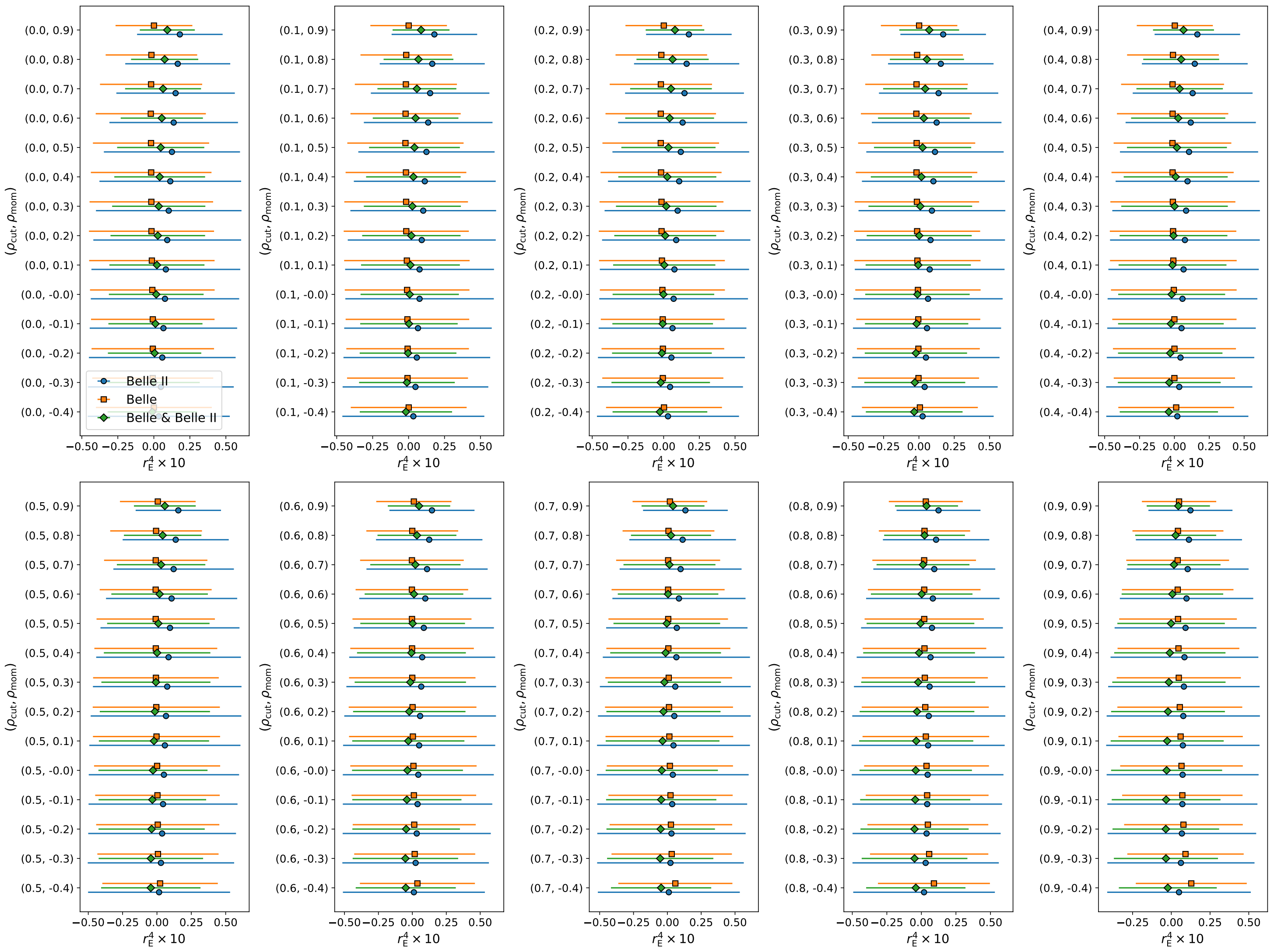


Figure E.12: Fit results for $r_E^4 \times 10$ for different combinations of correlation parameter values.

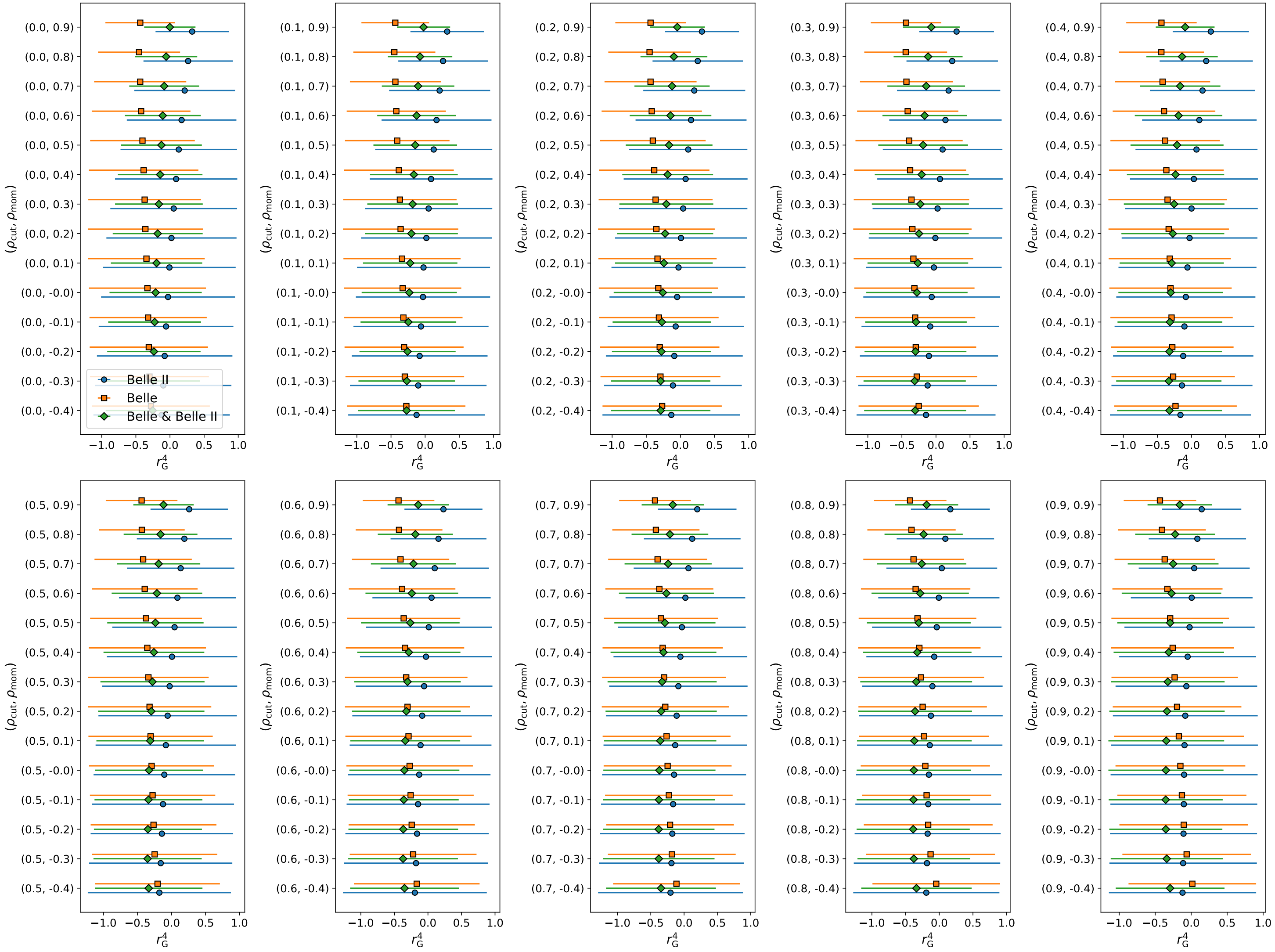


Figure E.13: Fit results for r_G^4 for different combinations of correlation parameter values.

Bibliography

---

[1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: `1207.7214 [hep-ex]`.

[2] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: `1207.7235 [hep-ex]`.

[3] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[4] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[5] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[6] M. Jezabek and J. H. Kuhn, *QCD Corrections to Semileptonic Decays of Heavy Quarks*, Nucl. Phys. B **314** (1989) 1.

[7] V. Aquila et al., *Perturbative corrections to semileptonic b decay distributions*, Nucl. Phys. B **719** (2005) 77, arXiv: `hep-ph/0503083 [hep-ph]`.

[8] T. Becher et al., *Kinetic corrections to* $B \to X_c \ell \overline{\nu}$ *at one loop*, JHEP **12** (2007) 062, arXiv: `0708.0855 [hep-ph]`.

[9] A. Pak and A. Czarnecki, *Heavy-to-heavy quark decays at NNLO*, Phys. Rev. D **78** (2008) 114015, arXiv: `0808.3509 [hep-ph]`.

[10] K. Melnikov, $\mathcal{O}(\alpha_s^2)$ *corrections to semileptonic decay* $b \to c\ell\overline{\nu}_\ell$, Phys. Lett. B **666** (2008) 336, arXiv: `0803.0951 [hep-ph]`.

[11] T. Mannel et al., *Higher Order Power Corrections in Inclusive B Decays*, JHEP **11** (2010) 109, arXiv: `1009.4622 [hep-ph]`.

[12] T. Mannel et al., *Inclusive semileptonic B decays from QCD with NLO accuracy for power suppressed terms*, Phys. Lett. B **741** (2015) 290, arXiv: `1405.5072 [hep-ph]`.

[13] A. Alberti et al., *Perturbative corrections to power suppressed effects in semileptonic B decays*, JHEP **01** (2014) 147, arXiv: `1311.7381 [hep-ph]`.

[14] P. Gambino and C. Schwanda, *Inclusive semileptonic fits, heavy quark masses, and* V_{cb}, Phys. Rev. D **89** (2014) 014022, arXiv: `1307.4551 [hep-ph]`.

[15] A. Alberti et al., *Precision Determination of the Cabibbo-Kobayashi-Maskawa Element V_{cb}*, Phys. Rev. Lett. **114** (2015) 061802, arXiv: `1411.6560 [hep-ph]`.

[16] CDF Collaboration, *Measurement of the moments of the hadronic invariant mass distribution in semileptonic B decays*, Phys. Rev. D **71** (2005) 051103, arXiv: `hep-ex/0502003 [hep-ex]`.

[17] CLEO Collaboration, *Moments of the B meson inclusive semileptonic decay rate using neutrino reconstruction*, Phys. Rev. D **70** (2004) 032002, arXiv: `hep-ex/0403052 [hep-ex]`.

[18] DELPHI Collaboration, *Determination of heavy quark non-perturbative parameters from spectral moments in semileptonic B decays*, Eur. Phys. J. C **45** (2006) 35, arXiv: `hep-ex/0510024 [hep-ex]`.

[19] BaBar Collaboration, *Measurement of the electron energy spectrum and its moments in inclusive $B \to Xe\nu$ decays*, Phys. Rev. D **69** (2004) 111104, arXiv: `hep-ex/0403030 [hep-ex]`.

[20] BaBar Collaboration, *Measurement and interpretation of moments in inclusive semileptonic decays $\overline{B} \to X_c \ell \overline{\nu}_\ell$*, Phys. Rev. D **81** (2010) 032003, arXiv: `0908.0415 [hep-ex]`.

[21] Belle Collaboration, *Moments of the electron energy spectrum and partial branching fraction of $B \to X_c e\nu$ decays at Belle*, Phys. Rev. D **75** (2007) 032001, arXiv: `hep-ex/0610012 [hep-ex]`.

[22] Belle Collaboration, *Moments of the Hadronic Invariant Mass Spectrum in $B \to X_c \ell\nu$ Decays at Belle*, Phys. Rev. D **75** (2007) 032005, arXiv: `hep-ex/0611044 [hep-ex]`.

[23] Particle Data Group, *Review of Particle Physics*, PTEP **2020** (2020) 083C01.

[24] D. Bigi and P. Gambino, *Revisiting $B \to D\ell\nu$*, Phys. Rev. D **94** (2016) 094008, arXiv: `1606.08030 [hep-ph]`.

[25] D. Bigi et al., *A fresh look at the determination of $|V_{cb}|$ from $B \to D^*\ell\nu$*, Phys. Lett. B **769** (2017) 441, arXiv: `1703.06124 [hep-ph]`.

[26] F. U. Bernlochner et al., *Tensions and correlations in $|V_{cb}|$ determinations*, Phys. Rev. D **96** (2017) 091503, arXiv: `1708.07134 [hep-ph]`.

[27] D. Ferlewicz et al., *Revisiting fits to $B^0 \to D^{*-}\ell^+\nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at nonzero recoil*, Phys. Rev. D **103** (7 2021) 073005, arXiv: `2008.09341 [hep-ph]`.

[28] D. Benson et al., *Imprecated, yet impeccable: On the theoretical evaluation of $\Gamma(B \to X_c\ell\nu)$*, Nucl. Phys. B **665** (2003) 367, arXiv: `hep-ph/0302262 [hep-ph]`.

[29] P. Gambino, *B semileptonic moments at NNLO*, JHEP **09** (2011) 055, arXiv: `1107.3100 [hep-ph]`.

[30] M. Fael et al., *V_{cb} determination from inclusive $b \to c$ decays: an alternative method*, JHEP **02** (2019) 177, arXiv: `1812.07472 [hep-ph]`.

[31] Belle Collaboration,
Measurements of q^2 Moments of Inclusive $B \to X_c \ell^+ \nu_\ell$ Decays with Hadronic Tagging,
Phys. Rev. D **104** (2021) 112011, arXiv: `2109.01685 [hep-ex]`.

[32] B. Povh et al., *Teilchen und Kerne. Eine Einführung in die physikalischen Konzepte*, Sixth,
Springer-Verlag GmbH, 2004, ISBN: 3540210652.

[33] M. Thomson, *Modern particle physics*, New York: Cambridge University Press, 2013,
ISBN: 978110734266.

[34] Particle Data Group, *CKM quark-mixing matrix in Review of Particle Physics*,
PTEP **2020** (2020) 083C01.

[35] A. Manohar et al., *Heavy Quark Physics*, Cambridge University Press, 2000,
ISBN: 9780521642415.

[36] *Standard Model of Elementary Particles*,
URL: `https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg` (visited on 06/07/2022).

[37] L.-L. Chau and W.-Y. Keung,
Comments on the Parametrization of the Kobayashi-Maskawa Matrix,
Phys. Rev. Lett. **53** (19 1984) 1802.

[38] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*,
Phys. Rev. Lett. **51** (21 1983) 1945.

[39] T. Mannel and K. K. Vos,
Reparametrization Invariance and Partial Re-Summations of the Heavy Quark Expansion,
JHEP **06** (2018) 115, arXiv: `1802.09409 [hep-ph]`.

[40] HFLAV Group, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*,
Eur. Phys. J. C **81** (2021) 226, arXiv: `1909.12524 [hep-ex]`.

[41] CKMFitter Group, *Recent CKMfitter updates on global fits of the CKM matrix*,
Workshop: 11th International Workshop on the CKM Unitarity Triangle, 2021,
URL: `https://indico.cern.ch/event/891123/contributions/4601722/attachments/2351890/4013122/CKMfitter2021.pdf`.

[42] A. Sirlin, *Current Algebra Formulation of Radiative Corrections in Gauge Theories and the Universality of the Weak Interactions*,
Rev. Mod. Phys. **50** (1978) 573, [Erratum: Rev.Mod.Phys. 50, 905 (1978)].

[43] I. Caprini et al., *Dispersive bounds on the shape of $\overline{B} \to D^{(*)} \ell \overline{\nu}$ form-factors*,
Nucl. Phys. B **530** (1998) 153, arXiv: `hep-ph/9712417 [hep-ph]`.

[44] Flavour Lattice Averaging Group,
FLAG Review 2019: Flavour Lattice Averaging Group (FLAG), Eur. Phys. J. C **80** (2020) 113,
arXiv: `1902.08191 [hep-lat]`.

[45] MILC Collaboration,
Update of $|V_{cb}|$ from the $\overline{B} \to D^* \ell \overline{\nu}$ form factor at zero recoil with three-flavor lattice QCD,
Phys. Rev. D **89** (2014) 114504, arXiv: `1403.0635 [hep-lat]`.

[46] J. Harrison et al., *Lattice QCD calculation of the* $B_{(s)} \to D^*_{(s)} \ell \nu$ *form factors at zero recoil and implications for* $|V_{cb}|$, Phys. Rev. D **97** (2018) 054502, arXiv: `1711.11013 [hep-lat]`.

[47] B. Grinstein and A. Kobach, *Model-Independent Extraction of* $|V_{cb}|$ *from* $\overline{B} \to D^* \ell \overline{\nu}$, Phys. Lett. B **771** (2017) 359, arXiv: `1703.08170 [hep-ph]`.

[48] Belle Collaboration,
Measurement of the CKM matrix element $|V_{cb}|$ *from* $B^0 \to D^{*-} \ell^+ \nu_\ell$ *at Belle*,
Phys. Rev. D **100** (2019) 052007, arXiv: `1809.03290 [hep-ex]`.

[49] BaBar collaboration,
Extraction of form Factors from a Four-Dimensional Angular Analysis of $\overline{B} \to D^* \ell^- \overline{\nu}_\ell$,
Phys. Rev. Lett. **123** (2019) 091801, arXiv: `1903.10002 [hep-ex]`.

[50] MILC Collaboration,
$B \to D \ell \nu$ *form factors at nonzero recoil and* $|V_{cb}|$ *from 2+1-flavor lattice QCD*,
Phys. Rev. D **92** (2015) 034506, arXiv: `1503.07237 [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* $|V_{cb}|$,
Phys. Rev. D **93** (2016) 032006, arXiv: `1510.03657 [hep-ex]`.

[52] BaBar Collaboration, *Measurement of |V(cb)| and the Form-Factor Slope in* $\overline{B} \to D \ell^- \overline{\nu}$ *Decays in Events Tagged by a Fully Reconstructed B Meson*,
Phys. Rev. Lett. **104** (2010) 011802, arXiv: `0904.4063 [hep-ex]`.

[53] LHCb Collaboration, *Measurement of* $|V_{cb}|$ *with* $B^0_s \to D_s^{(*)-} \mu^+ \nu_\mu$ *decays*,
Phys. Rev. D **101** (2020) 072004, arXiv: `2001.03225 [hep-ex]`.

[54] P. Gambino and N. Uraltsev,
Moments of semileptonic B decay distributions in the $1/m_b$ *expansion*,
Eur. Phys. J. C **34** (2004) 181, arXiv: `hep-ph/0401063 [hep-ph]`.

[55] C. W. Bauer et al., *Global analysis of inclusive B decays*, Phys. Rev. D **70** (2004) 094017, arXiv: `hep-ph/0408002 [hep-ph]`.

[56] Belle Collaboration, *Measurement of the Moments of the Photon Energy Spectrum in* $B \to X_s \gamma$ *Decays and Determination of* $|V_{cb}|$ *and* m_b *at Belle*,
Phys. Rev. D **78** (2008) 032016, arXiv: `0803.2158 [hep-ex]`.

[57] P. Gambino et al., *Taming the higher power corrections in semileptonic B decays*,
Phys. Lett. B **763** (2016) 60, arXiv: `1606.06174 [hep-ph]`.

[58] M. Bordone et al., *Three loop calculations and inclusive* V_{cb},
Phys. Lett. B **822** (2021) 136679, arXiv: `2107.00604 [hep-ph]`.

[59] M. Fael et al., *Third order corrections to the semileptonic* $b \to c$ *and the muon decays*,
Phys. Rev. D **104** (2021) 016003, arXiv: `2011.13654 [hep-ph]`.

[60] M. Fael et al., *Kinetic Heavy Quark Mass to Three Loops*, Phys. Rev. Lett. **125** (2020) 052003, arXiv: `2005.06487 [hep-ph]`.

[61] M. Fael et al., *Relation between the* $\overline{\text{MS}}$ *and the kinetic mass of heavy quarks*,
Phys. Rev. D **103** (2021) 014005, arXiv: `2011.11655 [hep-ph]`.

[62] P. Gambino et al., *The V_{cb} puzzle: An update*, Phys. Lett. B **795** (2019) 386, arXiv: `1905.08209 [hep-ph]`.

[63] Belle II Collaboration, *The Belle II Physics Book*, PTEP **2019** (2019) 123C01, [Erratum: PTEP 2020, 029201 (2020)], arXiv: `1808.10567 [hep-ex]`.

[64] SuperKEKB Collaboration, *SuperKEKB Collider*, Nucl. Instrum. Meth. A **907** (2018) 188, arXiv: `1809.01958 [physics.acc-ph]`.

[65] Belle II Collaboration, *Belle II Technical Design Report*, (2010), arXiv: `1011.0352 [physics.ins-det]`.

[66] *KEK B-Factory Design Report*, KEK-REPORT-95-7 (1995).

[67] SuperB Collaboration, *SuperB Technical Design Report*, (2013), arXiv: `1306.5655 [physics.ins-det]`.

[68] *SuperKEKB Raises the Bar*, URL: `https://cerncourier.com/a/superkekb-raises-the-bar/` (visited on 05/25/2022).

[69] Belle Collaboration, *The Belle Detector*, Nucl. Instrum. Meth. A **479** (2002) 117.

[70] *Model of the Belle II Detector*, URL: `https://www.mpp.mpg.de/en/research/structure-of-matter/belle-ii-detector-on-the-track-of-the-antimatter-puzzle` (visited on 03/30/2022).

[71] H. Kolanoski and N. Wermes, *Particle Detectors: Fundamentals and Applications*, Oxford: Oxford University Press, 2020 944, ISBN: 9780198858362.

[72] J. Kemmer and G. Lutz, *New Detector Concepts*, Nucl. Instrum. Meth. A **253** (1987) 365.

[73] Belle II Tracking Group, *Track finding at Belle II*, Comput. Phys. Commun. **259** (2021) 107610, arXiv: `2003.12466 [physics.ins-det]`.

[74] T. Alexopoulos et al., *Implementation of the Legendre Transform for track segment reconstruction in drift tube chambers*, Nucl. Instrum. Meth. A **592** (2008) 456.

[75] R. Frühwirth et al., *Track finding in silicon trackers with a small number of layers*, Nucl. Instrum. Meth. A **732** (2013) 95.

[76] T. Bilka et al., *Implementation of GENFIT2 as an experiment independent track-fitting framework*, (2019), arXiv: `1902.04405 [physics.data-an]`.

[77] J. Lettenbicher, *Real-time Pattern Recognition in the Central Tracking Detector of the Belle II Experiment*, PhD Thesis: Technische Universität Wien, 2016.

[78] P. Billoir, *Progressive track recognition with a Kalman-like fitting procedure*, Computer Physics Communications **57** (1989) 390, URL: `https://www.sciencedirect.com/science/article/pii/001046558990249X`.

[79] Belle II Tracking Group, *Tracking Performance in Early Belle II Data*, Unpublished manuscript, URL: `https://docs.belle2.org/record/2963`.

[80] Belle II Neutrals Performance Group, *Photon and neutral pion performance in early Belle II data*, Unpublished manuscript, URL: https://docs.belle2.org/record/2938.

[81] Belle II Lepton ID Group, *Muon and electron identification efficiencies and hadron-lepton mis-identification probabilities with Belle II Phase III data*, Unpublished manuscript, URL: https://docs.belle2.org/record/2340.

[82] Belle II Collaboration, *Measurement of the integrated luminosity of the Phase 2 data of the Belle II experiment*, Chinese Phys. C **44** (2020) 021001, URL: https://doi.org/10.1088/1674-1137/44/2/021001.

[83] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instr. Meth. A **462** (2001) 152, URL: http://www.sciencedirect.com/science/article/pii/S0168900201000894.

[84] S. Jadach et al., *The Precision Monte Carlo event generator $\mathcal{KK}$ for two fermion final states in e^+e^- collisions*, Comput. Phys. Commun. **130** (2000) 260, arXiv: hep-ph/9912214 [hep-ph].

[85] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[86] E. Barberio et al., *PHOTOS: A Universal Monte Carlo for QED radiative corrections in decays*, Comput. Phys. Commun. **66** (1991) 115.

[87] S. Agostinelli et al., *GEANT4: A Simulation Toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[88] Belle II Framework Software Group, *The Belle II Core Software*, Comput. Softw. Big Sci. **3** (2019) 1, arXiv: 1809.04299 [physics.comp-ph].

[89] C. G. Boyd et al., *Constraints on form-factors for exclusive semileptonic heavy to light meson decays*, Phys. Rev. Lett. **74** (1995) 4603, arXiv: hep-ph/9412324 [hep-ph].

[90] F. U. Bernlochner and Z. Ligeti, *Semileptonic $B_{(s)}$ decays to excited charmed mesons with e, μ, τ and searching for new physics with $R(D^{**})$*, Phys. Rev. D **95** (2017) 014022, arXiv: 1606.09300 [hep-ph].

[91] BaBar Collaboration, *Observation of $\overline{B} \to D^{(*)}\pi^+\pi^-\ell^-\overline{\nu}$ decays in e^+e^- collisions at the $\Upsilon(4S)$ resonance*, Phys. Rev. Lett. **116** (2016) 041801, arXiv: 1507.08303 [hep-ex].

[92] *Correction of $\mathcal{B}(B \to X_c\ell\nu)$ in generic MC*, Belle II Internal Acces Only, URL: https://confluence.desy.de/pages/viewpage.action?pageId=202394372 (visited on 05/29/2022).

[93] M. Prim, *b2-hive/eFFORT*, version v0.1.0, 2020, URL: https://doi.org/10.5281/zenodo.3965699.

[94] Belle II Collaboration, *Muon and electron identification efficiencies and hadron-lepton mis-identification rates at Belle II for Moriond 2021*, Internal Note - BELLE2-CONF-PH-2021-002 (2021), URL: `https://docs.belle2.org/record/2318/`.

[95] *Recommendations for LeptonID - Moriond 2021*, Belle II Internal Acces Only, URL: `https://confluence.desy.de/display/BI/Recommendations+for+LeptonID+-+Moriond+2021` (visited on 05/30/2022).

[96] T. Keck et al., *The Full Event Interpretation: An Exclusive Tagging Algorithm for the Belle II Experiment*, Comput. Softw. Big Sci. **3** (2019) 6, arXiv: `1807.08680 [hep-ex]`.

[97] T. Keck, *FastBDT: A Speed-Optimized Multivariate Classification Algorithm for the Belle II Experiment*, Comput. Softw. Big Sci. **1** (2017) 2, arXiv: `1609.06119 [cs.LG]`.

[98] M. Feindt et al., *A Hierarchical NeuroBayes-based Algorithm for Full Reconstruction of B Mesons at B Factories*, Nucl. Instrum. Meth. A **654** (2011) 432, arXiv: `1102.3876 [hep-ex]`.

[99] Belle II Collaboration, *A Calibration of the Belle II Hadronic Tag-side Reconstruction Algorithm with $B \to X\ell\nu$ Decays*, (2020), arXiv: `2008.06096 [hep-ex]`.

[100] G. Fox and S. Wolfram, *Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes*, Phys. Rev. Lett. **41** (1978) 1581.

[101] P. Ecker, *Study of Bremsstrahlung Finding at the Belle II Experiment*, Bachelor Thesis: Karlsruher Institut für Technologie, 2018.

[102] Belle Collaboration, *Measurement of $|V_{ub}|$ from Inclusive Charmless Semileptonic B Decays*, Phys. Rev. Lett. **104** (2 2010) 021801, arXiv: `0907.0379 [hep-ex]`.

[103] F. Zernike, *Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode*, Physica **1** (1934) 689.

[104] BaBar and Belle Collaborations, *The Physics of the B Factories*, Eur. Phys. J. C **74** (2014) 3026, arXiv: `1406.6311 [hep-ex]`.

[105] D. Kraft, *A software package for sequential quadratic programming*, Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt Köln: Forschungsbericht (1988).

[106] P. Virtanen et al., *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nat. Methods **17** (2020) 261.

[107] J. E. Sundermann, *Improved Measurements of Hadronic Mass Moments in $B \to X_c\ell\nu$ Decays*, PhD Thesis: Technische Universität Dresden, 2006.

[108] V. Klose, *Measurement and Interpretation of Moments of the combined Hadronic Mass and Energy Spectrum in Inclusive semileptoni B-Meson Decays*, PhD Thesis: Technische Universität Dresden, 2007.

[109] BaBar Collaboration, *Measurement and Interpretation of Moments in Inclusive Semileptonic Decays $\overline{B} \to X_c\ell^-\overline{\nu}$*, (2009), arXiv: `0908.0415`.

[110] F. James and M. Roos, *Minuit - a system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun. **10** (1975) 343.

[111] H. Dembinski et al., *scikit-hep/iminuit*, (2020), URL: `https://doi.org/10.5281/zenodo.3949207`.

[112] Belle II Collaboration,
Measurement of the tracking efficiency in Phase 3 data using tau-pair events,
Internal Note - BELLE2-NOTE-PH-2020-006 (2020),
URL: `https://docs.belle2.org/record/1867`.

[113] G. Cowan, *Statistical Data Analysis*, Oxford science publications, Clarendon Press, 1998, ISBN: 9780198501565.

[114] R. Simard and P. L'Ecuyer, *Computing the Two-Sided Kolmogorov-Smirnov Distribution*, J. Stat. Softw. **39** (2011) 1.

[115] C. De Boor, *A practical guide to splines*, New York: Springer, 1978, ISBN: 9780387953663.

[116] E. Prilepin, *CSAPS - Cubic Spline Approximation (Smoothing)*, 2021, URL: `https://github.com/espdev/csaps`.

[117] P. T. Boggs and J. E. Rogers, *Orthogonal distance regression*, Contemporary Mathematics **112** (1990) 183.

[118] K. G. Hayes et al.,
Application of the Bootstrap Statistical Method to the Tau Decay Mode Problem,
Phys. Rev. D **39** (1989) 274.

[119] P. Brodtkorb and J. D'Errico, *numdifftools 0.9.39*, (2019), URL: `https://github.com/pbrod/numdifftools`.

[120] F. Bernlochner et al., *First extraction of inclusive V_{cb} from q^2 moments*, (2022), arXiv: `2205.10274 [hep-ph]`.

[121] K. Olschewsky et al., *InclusiveVcb*, `https://github.com/kolschew/InclusiveVcb` (2022).

[122] M. J. Dugan et al., *On the Hilbert space of the heavy quark effective theory*, Physics Letters B **282** (1992) 142.

[123] Y.-Q. Chen, *On the reparameterization invariance in heavy quark effective theory*, Physics Letters B **317** (1993) 421.

[124] J. Heinonen et al., *Lorentz invariance in heavy particle effective theories*, Phys. Rev. D **86** (2012) 094020, arXiv: `1208.0601 [hep-ph]`.

[125] M. E. Luke and A. V. Manohar,
Reparametrization invariance constraints on heavy particle effective field theories,
Phys. Lett. B **286** (1992) 348, arXiv: `hep-ph/9205228 [hep-ph]`.

[126] A. V. Manohar,
Reparametrization Invariance Constraints on Inclusive Decay Spectra and Masses,
Phys. Rev. D **82** (2010) 014009, arXiv: `1005.1952 [hep-ph]`.

[127] K. G. Chetyrkin et al., *Charm and Bottom Quark Masses: An Update*, Phys. Rev. D **80** (2009) 074010, arXiv: `0907.2110 [hep-ph]`.

[128] K. G. Chetyrkin et al., *Addendum to "Charm and bottom quark masses: An update"*, (2017), [Addendum: Phys.Rev.D 96, 116007 (2017)], arXiv: `1710.04249 [hep-ph]`.

[129] K. Chetyrkin et al., *RunDec: a Mathematica package for running and decoupling of the strong coupling and quark masses*, Computer Physics Communications **133** (2000) 43.

[130] F. Herren and M. Steinhauser, *Version 3 of RunDec and CRunDec*, Comp. Phys. Commun. **224** (2018) 333.

[131] P. Gambino et al., *$B \to D^*$ Zero-Recoil Formfactor and the Heavy Quark Expansion in QCD: A Systematic Study*, JHEP **10** (2012) 169, arXiv: `1206.2296 [hep-ph]`.

[132] Belle Collaboration, *Measurements of Partial Branching Fractions of Inclusive $B \to X_u \ell^+ \nu_\ell$ Decays with Hadronic Tagging*, Phys. Rev. D **104** (2021) 012008, arXiv: `2102.00020 [hep-ex]`.

[133] CLEO Collaboration, *Measurement of the B-meson inclusive semileptonic branching fraction and electron energy moments*, Phys. Rev. D **70** (2004) 032003, arXiv: `hep-ex/0403053 [hep-ex]`.

[134] BaBar Collaboration, *Measurement of the inclusive electron spectrum from B meson decays and determination of $|V_{ub}|$*, Phys. Rev. D **95** (2017) 072001, arXiv: `1611.05624 [hep-ex]`.

[135] BaBar Collaboration, *Measurement of the ratio $\mathcal{B}(B^+ \to Xe\nu)/\mathcal{B}(B^0 \to Xe\nu)$*, Phys. Rev. D **74** (2006) 091105, arXiv: `hep-ex/0607111 [hep-ex]`.

[136] S. Turczyk, *Additional Information on Heavy Quark Parameters from Charged Lepton Forward-Backward Asymmetry*, JHEP **04** (2016) 131, arXiv: `1602.02678 [hep-ph]`.

[137] F. Herren, *The forward-backward asymmetry and differences of partial moments in inclusive semileptonic B decays*, (2022), arXiv: `2205.03427 [hep-ph]`.

List of Figures

List of Tables