

Study of $B \rightarrow \pi \ell \nu$ and $B \rightarrow \rho \ell \nu$ decays and extraction of $|V_{ub}|$ at Belle II

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

vorgelegt von

Svenja Katharina Granderath
aus
Dormagen

Bonn 2024

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

Gutachter/Betreuer:	Prof. Dr. Jochen Dingfelder
Gutachter:	Prof. Dr. Florian Bernlochner
Tag der Promotion:	29.11.2024
Erscheinungsjahr:	2025

Abstract

This thesis presents a measurement of the magnitude of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $|V_{ub}|$ from a simultaneous study of the semileptonic decays $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, where $\ell = e, \mu$. The analysis utilizes a data sample corresponding to 364 fb^{-1} of integrated luminosity, collected by the Belle II detector at the SuperKEKB electron-positron accelerator from 2019 to 2022. The SuperKEKB accelerator operates at a center-of-mass energy corresponding to the mass of the $\Upsilon(4S)$ resonance. This results in nearly all produced $\Upsilon(4S)$ mesons decaying into two B mesons, which then decay into lighter particles. An untagged measurement approach is employed in this thesis, where both decay modes are reconstructed without identifying the partner B mesons. After initial signal reconstruction and selection, machine learning methods, specifically boosted decision trees, are trained and applied to suppress background contributions. The differential branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ are measured simultaneously as functions of q^2 (momentum transfer squared). This is achieved by introducing a novel fit method, in which yields are inherently corrected for finite detector resolution. The total branching fractions are determined to be $\mathcal{B}(B^0 \rightarrow \pi^- \ell^+ \nu_\ell) = (1.516 \pm 0.042(\text{stat}) \pm 0.059(\text{syst})) \times 10^{-4}$ and $\mathcal{B}(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell) = (1.625 \pm 0.079(\text{stat}) \pm 0.180(\text{syst})) \times 10^{-4}$. Fitting the measured partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ with lattice QCD constraints yields $|V_{ub}| = (3.93 \pm 0.09(\text{stat}) \pm 0.13(\text{syst}) \pm 0.19(\text{theo})) \times 10^{-3}$. Including additional constraints from light-cone sum rules (LCSR), the result is $(3.73 \pm 0.07(\text{stat}) \pm 0.07(\text{syst}) \pm 0.16(\text{theo})) \times 10^{-3}$. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, applying LCSR constraints yields $|V_{ub}| = (3.19 \pm 0.12(\text{stat}) \pm 0.17(\text{syst}) \pm 0.26(\text{theo})) \times 10^{-3}$. While the results are limited by theoretical uncertainties, the leading systematic uncertainties arise from the sizes of the available off-resonance and simulated samples.

Contents

1	Introduction	1
2	Theoretical background and status of V_{ub} measurements	3
2.1	The Standard Model of particle physics	3
2.1.1	Strong interactions	5
2.1.2	Electroweak unification and symmetry-breaking	6
2.1.3	Neutral and charged current interactions	7
2.1.4	Mass generation and Yukawa couplings	8
2.1.5	The CKM matrix and CP violation	9
2.2	Semileptonic $B \rightarrow X_u \ell \nu_\ell$ decays	11
2.2.1	Inclusive description	13
2.2.2	Exclusive description	15
2.3	Status of $ V_{ub} $ measurements	18
3	SuperKEKB and Belle II	21
3.1	The SuperKEKB accelerator	21
3.2	Particle collisions and beam-induced backgrounds	23
3.3	The Belle II detector	24
3.4	Particle reconstruction and identification	30
3.4.1	Track and vertex reconstruction	30
3.4.2	Cluster reconstruction	31
3.4.3	Track-cluster matching	31
3.4.4	Particle identification	32
4	Collision and simulated datasets	35
4.1	Collision data	35
4.2	Simulated data	35
4.2.1	Branching fractions	36
4.2.2	Form-factor models	37
4.2.3	Lineshape of the ρ meson	38
4.2.4	Nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ spectrum	40
4.2.5	Hybrid model implementation	42
4.3	Corrections based on data-MC agreement	42
4.3.1	Particle identification corrections	43
4.3.2	Track momentum and photon energy bias	44
4.3.3	Continuum corrections	44

5	Event reconstruction and selection	45
5.1	Event reconstruction	45
5.2	Signal B meson reconstruction and selection	46
5.2.1	Reconstruction and preselection of signal events	46
5.2.2	Signal-specific selections	49
5.3	Missing momentum reconstruction	51
5.4	Event categorization	53
5.5	Signal extraction variables	54
6	Background suppression	61
6.1	Multivariate analysis	62
6.2	Background suppression using boosted decision trees	63
6.2.1	Classifier setup and data samples	63
6.2.2	Input variables	64
6.2.3	Classifier validation and performance	68
6.2.4	Classifier selection	69
6.3	Best candidate selection	71
6.4	Signal selection efficiencies	72
7	Branching fraction measurement	75
7.1	Parameter estimation through fits	76
7.2	Maximum likelihood fit method	76
7.3	Construction of the likelihood function	77
7.4	Fit templates, parameters and variables	78
7.4.1	Continuum template: Reweighting	80
7.4.2	Signal templates: Composition	82
7.5	Fit validation, results and stability	85
7.6	Branching fraction results	91
8	Systematic uncertainties	93
8.1	Uncertainty types	93
8.2	Uncertainty evaluation	94
8.3	Uncertainty sources	94
8.3.1	Detector effects	95
8.3.2	Beam-energy difference	97
8.3.3	BDT selection efficiency	97
8.3.4	Simulated sample size	98
8.3.5	Physics constraints	99
8.3.6	$B \rightarrow X_u \ell \nu_\ell$ background	100
8.3.7	$B \rightarrow X_c \ell \nu_\ell$ background	101
8.3.8	Continuum component	102
8.3.9	Signal model and ρ meson lineshape	102

9	Determination of V_{ub}	105
9.1	Extraction of $ V_{ub} $ from χ^2 fits	105
9.1.1	Setup for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays	105
9.1.2	Setup for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays	106
9.2	$ V_{ub} $ fit results	107
9.3	Uncertainties	109
9.4	Stability tests	111
10	Conclusion and outlook	113
A	Further studies	115
A.1	Bremsstrahlung parameter optimization	115
A.2	ROE track and cluster selection optimization	115
A.3	Control mode	117
A.3.1	Reconstruction	117
A.3.2	Validation of q^2 reconstruction	118
A.4	Continuum reweighting	119
B	Distributions of ΔE and M_{bc}	125
B.1	Prefit distributions	125
B.2	Postfit distributions	125
C	Correlation matrices	135
	Bibliography	139
	Acronyms	147
	List of Figures	149
	List of Tables	151

Introduction

The Standard Model (SM) of particle physics describes the known fundamental particles – quarks, leptons, and bosons – that constitute the universe and their interactions through the strong, weak, and electromagnetic forces. Since its formulation, the SM has successfully predicted the existence of particles, such as the third generation of quarks [1], confirmed by the discovery of the top quark [2, 3], and the Higgs boson [4], discovered at the Large Hadron Collider (LHC) in 2012 [5, 6].

Despite its many successes, the SM currently cannot account for all phenomena, such as the observed matter-antimatter asymmetry in the universe. Additionally, the SM relies on numerous free parameters with in principle arbitrary magnitudes. This is in conflict with the hierarchical structures observed, for example, in quark masses or the magnitudes of the elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix [7, 8]. The CKM matrix describes the rotation of quark-mass eigenstates to the eigenstates involved in the weak interaction. It is a unitary matrix, meaning the magnitudes and complex phases of its elements can be used to construct unitarity triangles. By precisely measuring the sides and angles, the unitarity triangles can be over-constrained, helping to fully constrain the SM.

The smallest and least well-known CKM matrix element is V_{ub} , and its magnitude $|V_{ub}|$ can be determined by comparing measured to predicted decay rates of semileptonic $B \rightarrow X_u \ell \nu_\ell$ decays, where X_u is a hadron containing an up quark, and ℓ is a light charged lepton (electron or muon). The predicted decay rate is derived from the matrix element of the $B \rightarrow X_u \ell \nu_\ell$ decay, factorized into hadronic and leptonic contributions due to its semileptonic nature. While the leptonic part can be explicitly calculated, the hadronic part involves non-perturbative hadronic interactions, making it difficult to calculate and dependent on the nature of the X_u hadron.

If X_u represents the inclusive sum of possible final states, the hadronic part takes the form of inclusive rates. Conversely, if X_u exclusively represents a specific final state, the hadronic contribution is given as a form factor, typically parameterized in terms of the momentum transfer squared q^2 between the B meson and the hadron. These complementary approaches yield independent $|V_{ub}|$ values, displaying a persistent tension of approximately 2.3 standard deviations [9]. The origin of this discrepancy remains unclear, possibly stemming from experimental or theoretical sources, but resolving it is crucial, as it limits the constraining power of the unitarity triangles.

One point of investigation is comparing the $|V_{ub}|$ results obtained from different exclusive modes. While measuring decay rates is experimentally feasible for most exclusive modes, the limiting factor is the presence of precise predictions of the form factors. In the past, reliable predictions were only available for the $B \rightarrow \pi \ell \nu_\ell$ decay, which is both experimentally and theoretically the cleanest exclusive

mode to extract $|V_{ub}|$ from B meson decays. Consequently, the exclusive world average is constructed solely from $|V_{ub}|$ measurements from $B \rightarrow \pi \ell \nu_\ell$ decays. Only recently have form-factor predictions for the $B \rightarrow \rho \ell \nu_\ell$ decay reached a precision that allows for a meaningful determination of $|V_{ub}|$ from this mode. Using these predictions and a combination of previous measurements of $B \rightarrow \rho \ell \nu_\ell$ decay rates, Ref. [10] obtained a $|V_{ub}|$ result lower than the exclusive world average, introducing a further tension that remains to be understood.

These tensions can be tested using data recorded by the Belle II detector, which is located at the interaction point of the asymmetric-energy electron-positron SuperKEKB accelerator in Tsukuba, Japan. The collision energy of the SuperKEKB accelerator is tuned to the resonance energy of the $\Upsilon(4S)$ meson. The $\Upsilon(4S)$ mesons decay almost exclusively into pairs of B mesons, which subsequently decay further, for example $B \rightarrow X_u \ell \nu$. To reconstruct these decays, either both B mesons or only the signal B meson can be reconstructed, referred to as tagged and untagged measurements, respectively. In this work, an untagged measurement from the $B \rightarrow \pi \ell \nu_\ell$ mode is performed, resulting in lower purity but higher signal efficiency. This choice is driven by the need for high signal efficiency due to the rarity of $B \rightarrow X_u \ell \nu_\ell$ decays, and the limited size of the 2019–2022 Belle II dataset, corresponding to an integrated luminosity of 364 fb^{-1} [11].

The reconstruction of $B \rightarrow \pi \ell \nu_\ell$ decays is often limited by significant cross-feed from $B \rightarrow \rho \ell \nu_\ell$ decays, where the ρ meson decays into two pions, one of which is selected as the signal pion. To reduce the dependence, the decay rates of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays are simultaneously reconstructed and measured, focusing on decays with entirely charged final states in order to avoid backgrounds associated with neutral particles. The simultaneous measurement constrains the cross-feed modeling to improve the precision on $|V_{ub}|$ from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays and it provides access to determine $|V_{ub}|$ from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays. Previous simultaneous measurements of $B \rightarrow \pi \ell \nu_\ell$ and $B \rightarrow \rho \ell \nu_\ell$ decays were often limited by the resolution of the reconstructed q^2 distribution. In this work, a novel method is introduced that naturally accounts for the q^2 resolution by extracting the signal yields and linking the cross-feed components as functions of true, instead of reconstructed, q^2 . Using this approach, $|V_{ub}|$ is determined from both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays by fitting the measured partial decay rates using form-factor parameterizations [12, 13] and constraints from theoretical calculations [13–18].

In the following, Chapter 2 provides theoretical background information and details on the current status of $|V_{ub}|$ measurements. Next, the experimental setup and the relevant data samples are discussed in Chapter 3 and Chapter 4, respectively. The selection of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events is explained in Chapter 5. Both modes suffer from large levels of backgrounds, which are suppressed using boosted decision trees, described in Chapter 6. The extraction of the partial branching fractions, the associated systematic uncertainties, and the determination of $|V_{ub}|$ are discussed in Chapter 7, 8 and 9, respectively. Finally, Chapter 10 provides a conclusion and outlook.

The results discussed in this thesis are obtained following a *hidden signal box* approach [19] to avoid experimenter’s bias. This means access to collision data is restricted until the analysis setup has been finalized using simulated data. Furthermore throughout this thesis, charge-conjugation is implied and natural units are used, with the reduced Planck constant and the speed of light set to unity ($\hbar = c = 1$).

Theoretical background and status of $|V_{ub}|$ measurements

This chapter provides an introduction to the Standard Model of particle physics. Section 2.1 includes descriptions of the fundamental particles and their interactions, as well as details about the mass generation mechanism and the Cabibbo-Kobayashi-Maskawa matrix. In addition, the mathematical description of semileptonic $B \rightarrow X_u \ell \nu_\ell$ decays is described in Section 2.2. The current experimental status of $|V_{ub}|$ measurements is discussed in Section 2.3.

2.1 The Standard Model of particle physics

The Standard Model (SM) of particle physics is a relativistic quantum field theory [20] that characterizes the fundamental particles based on their spin, charge, and mass, and describes the interactions between these particles. A graphic representation of the SM is shown in Figure 2.1.

The SM includes descriptions of integer-spin bosons and half-integer spin fermions. Massless particles are described as right- or left-handed, depending on whether their spin points in the same or opposite direction to their momentum, respectively. Massive particles have both right- and left-handed components. A fully left-handed particle transforms into a fully right-handed particle under the parity (P) operation. Each fermion has a corresponding antiparticle, which is an exact copy of the particle, except for possessing opposite charges. The antiparticle can be obtained by applying the charge-conjugation (C) operation to the particle. There is also a time-reversal (T) operation which reverses the direction of time. The product of the three operations CPT is conserved by all fundamental interactions [22].

The group of fermions consists of quarks and leptons, each of which has six flavors. Three of the quarks (the up, charm, and top quarks) carry electric charges of $+2/3$, while the other three quarks (the down, strange, and bottom quarks) carry electric charges of $-1/3$. Quarks also carry color charge, which comes in six types (red, green, blue, and their corresponding anticolors). Pairs of one negatively and one positively charged quark belong to one of three generations. The generations share all properties except for their masses, with quarks belonging to higher generations having larger masses.

The leptons can similarly be grouped into three generations. Three of the leptons (the electron, muon, and tau) carry electric charges of -1 . The electron is the lightest of the charged leptons, while

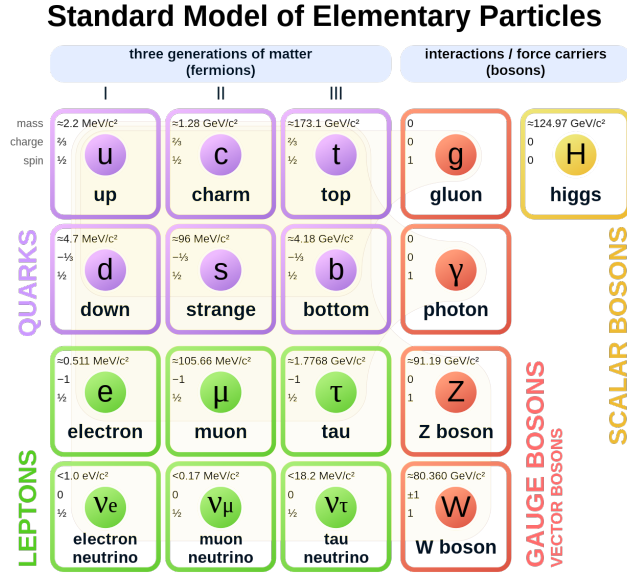


Figure 2.1: Graphic representation of the Standard Model of particle physics. The elementary particles (quarks, leptons, and bosons) are shown. The spin, charge, and mass of each particle are given. The fermions are grouped into three generations. Image from Ref. [21].

the tau is the heaviest. There are also three electrically-neutral leptons, called electron neutrino, muon neutrino, and tau neutrino. In the SM, neutrinos are considered massless, making them the only fermions that possess only a left-handed component. A pair of each charged lepton and its corresponding neutrino belongs to one of three generations.

The SM also describes the fundamental interactions between the elementary particles. These are the strong, electromagnetic, and weak interactions. Gravity is not included in the SM due to its negligible strength at the particle scale. The forces are mediated by vector (spin-1) gauge bosons. The exchange of electrically-neutral, massless gluons generates the strong interaction, described by Quantum Chromodynamics (QCD) [23, 24] and a local SU(3) symmetry. Gluons couple to particles carrying color charge, and they themselves carry color charge, giving rise to gluon self-interactions. The electromagnetic interaction is described by Quantum Electrodynamics (QED) [25] with a local U(1) symmetry and occurs between two particles through the exchange of an electrically-neutral, massless photon. The photon couples to all particles carrying electric charge.

Lastly, weak interactions are described by a local SU(2) symmetry and can be categorized into charged-current and neutral-current interactions. Weak charged-current interactions are mediated by electrically-charged W bosons, while weak neutral-current interactions are mediated by electrically-neutral Z bosons. In contrast to the photon, the W and Z bosons are massive. While the Z boson couples to all particles, the W bosons couple only to left-handed fermions and right-handed antifermions. The weak interaction, therefore, is the only interaction that is not invariant under the application of the combined CP operation and is said to violate CP symmetry. It is also the only interaction that can transform quarks from one flavor to another. The probabilities of such transitions are given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [7, 8], which incorporates a CP-violating factor.

In the underlying quantum field theory describing the SM, all particles are represented as excitations

of fields $\psi(x)$, where x is a vector containing space and time coordinates $x = \{\vec{x}, t\}$. The Lagrangian \mathcal{L} provides a description of the dynamics of the system caused by the excitations of the fields. It, therefore, describes the propagation of free particles and their interactions, depicted by Feynman diagrams. The SM Lagrangian is invariant under complex phase transformations under a local symmetry group

$$SU(3) \times SU(2) \times U(1),$$

which combines the QCD symmetry group and the electroweak symmetry group, resulting from the unification of the electromagnetic and weak interactions at high energies as described by the Glashow-Weinberg-Salam electroweak theory [26–28].

While the invariance allows for the existence of massless spin-1 bosons, it cannot explain the occurrence of the massive weak vector bosons. Instead, their presence is derived from spontaneous electroweak symmetry-breaking at low energies, caused by the introduction of a massive spin-0 field, the Higgs field. The electroweak symmetry-breaking creates the photon and the massive W and Z bosons, and concurs with the existence of the massive scalar (spin-0) Higgs boson, which corresponds to the excitation of the Higgs field. The Higgs boson couples to the massive vector bosons via the Higgs mechanism [29–31] and to the fermions using the Yukawa couplings [28].

The following subsections provide more details about specific aspects of the SM relevant for the work discussed in this thesis. Section 2.1.1 describes the properties of the strong interaction, while Section 2.1.2 discusses electroweak interactions, focusing on the mathematical framework describing electroweak symmetry-breaking. The emergence of neutral and charged current interactions within the electroweak framework, and the process of mass generation through Yukawa couplings are discussed in Section 2.1.3 and 2.1.4, respectively. In addition, the occurrence of CP violation, and the appearance of the CKM matrix elements is explained in Section 2.1.5.

2.1.1 Strong interactions

This subsection focuses on the description of the strong interaction, which arises due to the invariance of the SM Lagrangian under local gauge transformations defined by the $SU(3)_c$ symmetry, where c corresponds to the conserved color charge. The mathematical description of the strong force is provided by the theory of Quantum Chromodynamics (QCD) [23, 24], in which the quarks are described as color triplets $\psi = (\psi_1, \psi_2, \psi_3)^T$ with mass m_ψ . The interactions are mediated by gluons with corresponding gauge fields G_μ^a , which exist in eight color states ($a \in \{1, 8\}$), and belong to the gluon field strength tensor $G_{\mu\nu}^a$.

The QCD Lagrangian contains kinetic terms for the fermions and gluons, as well as interaction terms between quarks and gluons, and gluon self-interaction terms. It is given by:

$$\mathcal{L}_{\text{QCD}} = \sum_{\psi} \bar{\psi} (i\gamma^\mu (\partial_\mu - ig_s T^a G_\mu^a) - m_\psi) \psi - \frac{1}{4} G_{\mu\nu}^a G^{a,\mu\nu}.$$

Here, γ^μ are the Dirac matrices and T^a are the generators of the $SU(3)$ group. The strong coupling constant g_s can also be expressed as an effective coupling constant $\alpha_s = g_s^2/4\pi$.

It was observed that α_s depends on the energy scale of the process μ^2 , which is inversely proportional to the distance between the interacting particles. At short distances and high energies, where $\mu^2 \rightarrow \infty$, α_s decreases. In this regime, the interacting particles are quasi-free and experience asymptotic

freedom. The behavior can be described by perturbation theory. On the other hand, at low energies and larger distances, where $\mu^2 \rightarrow 0$, two competing effects contribute. One is a screening effect resulting from the polarization of the vacuum by $q\bar{q}$ pairs, which reduces α_s . The other effect comes from larger gluon self-interactions, resulting in an anti-screening and larger α_s . The anti-screening effect dominates, and thus α_s is larger in this regime.

This results in confinement, where quarks and gluons are never observed as free particles, but are always confined within color-neutral hadrons. Non-perturbative methods need to be employed to describe these processes. The functional dependence of α_s on the energy scale is described by

$$\alpha_s(\mu^2) = \frac{12\pi}{(33 - 2N_c) \ln(\mu^2/\Lambda_{\text{QCD}})},$$

where N_c is the number of colors and Λ_{QCD} characterizes the energy scale at which the transition from the perturbative to the non-perturbative regime occurs.

2.1.2 Electroweak unification and symmetry-breaking

This subsection discusses the properties of the electroweak interaction, and the process of electroweak symmetry-breaking.

The unification of the electromagnetic and weak interactions into a single electroweak interaction at high energies, described by the $SU(2)_L \times U(1)_Y$ group, is proposed by the Glashow-Weinberg-Salam theory [26–28]. It introduces the concept of left-handedness L and a quantity known as the weak hypercharge, denoted by $Y = 2(Q - I_3)$, where Q represents the electric charge and I_3 is the third component of the weak isospin. The weak-isospin quantum number I is $1/2$ for left-handed fermions and 0 for right-handed fermions.

The unified theory specifies the $U(1)_Y$ group's generator as the weak hypercharge acting on a gauge field B_μ , and the generators of the $SU(2)_L$ group as the Pauli spin matrices τ acting on the components of a triplet of gauge fields $W_\mu = (W_\mu^1, W_\mu^2, W_\mu^3)^T$. In contrast to the $SU(3)$ QCD group, the electroweak gauge fields of the $SU(2)_L \times U(1)_Y$ group, B_μ and W_μ , do not correspond directly to the physical gauge boson fields mediating the interactions.

The emergence of the difference between the electroweak and physical gauge bosons is explained by the Brout-Englert-Higgs (or Higgs) mechanism [29–31]. The mechanism introduces a complex scalar Higgs field ϕ with a non-trivial potential, the Higgs potential, which is not minimized at the origin. At low energies, the Higgs field acquires a non-zero vacuum expectation value v , leading to spontaneous electroweak symmetry-breaking and the emergence of the physical electromagnetic and weak boson fields. The fields A_μ and Z_μ , corresponding to the photon and the neutral Z boson fields, respectively, emerge from a rotation of B_μ and W_μ^3 by the weak mixing angle θ_W . The charged W^\pm boson fields are defined as linear combinations of W_μ^1 and W_μ^2 . The relation between the electroweak gauge and physical fields is given by the equations:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \quad \text{and} \quad W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2). \quad (2.1)$$

The electroweak Lagrangian post symmetry-breaking encompasses kinetic terms, Higgs self-interaction terms, terms describing the interactions between the Higgs field and the weak boson fields,

fermion-Higgs interaction terms (Yukawa couplings), self-interaction terms among the weak bosons, and terms describing neutral and charged currents.

For the work discussed in this thesis the neutral and charged current terms, and the Yukawa couplings are most relevant. For a comprehensive mathematical description of the electroweak interaction post symmetry-breaking, including a more detailed formulation of the physical boson fields, refer to standard electroweak theory textbooks such as Refs. [32, 33].

2.1.3 Neutral and charged current interactions

This subsection focuses on describing the neutral and charged current terms. The neutral and charged current terms are fundamental in describing weak interactions between fermions. The neutral current term includes interactions of the photon and the Z boson with fermions, while the charged current term involves the interaction of the W bosons with fermions.

First, three quark and three lepton weak-isospin doublets are defined, consisting of pairs of left-handed weak interaction eigenstates with opposite third components of weak isospin. These are represented by $\varphi_L = (u'_L, d'_L)^T$ for quarks and $(\nu', e'_L)^T$ for leptons (where the prime denotes weak eigenstates), and the right-handed components are described as singlets $\varphi_R = u'_R, d'_R, e'_R$. The projection operators $P_L = \frac{1}{2}(1 - \gamma_5)$ and $P_R = \frac{1}{2}(1 + \gamma_5)$ are used to separate the left- and right-handed components of the φ states. Here, $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$.

The charged current for the quark doublet is expressed as:

$$j_\mu^+ = \bar{u}'_L \gamma_\mu d'_L = \bar{u}' \gamma_\mu P_L d', \quad \text{and} \quad j_\mu^- = \bar{d}'_L \gamma_\mu u'_L = \bar{d}' \gamma_\mu P_L u'. \quad (2.2)$$

This formulation includes γ_μ and $\gamma_\mu \gamma_5$ terms, which reflect the $V - A$ structure of the weak interaction, where V stands for vector and A for axial-vector components. Considering the charged currents as interactions between the fermions and the physical W^\pm bosons, the currents are rewritten using the Pauli spin matrices τ , to generalize to the interaction between the fermions and the three gauge bosons W_i^μ :

$$j_\mu^i = \bar{\varphi}_L \gamma_\mu \frac{1}{2} \tau_i \varphi_L.$$

In this way, a triplet of weak currents following the $SU(2)_L$ symmetry is constructed, called the weak isospin current j_μ^i . Here, the third component describes neutral currents between the fermions and the W^3 boson. However, since this description only includes left-handed states, but right-handed states also exist, an additional term is needed to account for neutral currents including right-handed terms. Therefore, a separate weak isospin singlet for a generic fermion field ψ is constructed as:

$$j_\mu^Y = \bar{\psi} \gamma_\mu Y \psi.$$

This is called the weak hypercharge current, which follows the $U(1)_Y$ symmetry and acts on the B^μ gauge boson.

Combining the weak isospin and the weak hypercharge currents, the contributions of the charged and neutral current terms to the electroweak Lagrangian before symmetry-breaking can be described as:

$$\mathcal{L}_{\text{EW,pre}} = -ig j_\mu^i W_i^\mu - i \frac{g'}{2} j_\mu^Y B^\mu, \quad (2.3)$$

where g and g' are the $SU(2)_L$ and $U(1)_Y$ coupling constants, respectively. By combining Equations 2.1 with Equation 2.3, terms contributing to the electroweak Lagrangian after symmetry-breaking are defined, which describe the interactions with the physical gauge bosons. The resulting contributions from the charged current, electromagnetic neutral, and weak neutral current interactions are:

$$\mathcal{L}_{CC} = \frac{g}{\sqrt{2}} j_{\pm}^{\mu} W_{\mu}^{\mp}, \quad (2.4)$$

$$\mathcal{L}_{NC}^{\gamma} = e j_{\mu}^{\text{em}} A^{\mu}, \quad (2.5)$$

$$\mathcal{L}_{NC}^Z = \frac{g'}{\sin \theta_W} j_{\mu}^{\text{NC}} Z^{\mu}, \quad (2.6)$$

with the coupling strength of the W and Z bosons to fermions now defined as $g = e/\sin \theta_W$ and $g' = e/\cos \theta_W$, respectively. They relate the electromagnetic coupling strength e and the weak mixing angle θ_W . In the above, both the effective electromagnetic current j_{μ}^{em} and the effective weak neutral current j_{μ}^{NC} are constructed as combinations of the orthogonal currents j_{μ}^3 and j_{μ}^Y :

$$j_{\mu}^{\text{em}} = j_{\mu}^3 + \frac{1}{2} j_{\mu}^Y = \bar{\psi} \gamma_{\mu} Q \psi,$$

$$j_{\mu}^{\text{NC}} = j_{\mu}^3 - \sin^2 \theta_W j_{\mu}^{\text{em}},$$

where Q is the electric charge operator.

This formulation underlines the dual role of the electroweak theory in describing both the weak and electromagnetic neutral interactions through the exchange of the Z boson and the photon, respectively. The distinction and interaction between left- and right-handed states, as mediated by the projection operators P_L and P_R , are key, especially for the charged currents that change the flavor of quarks and leptons through the exchange of W^{\pm} bosons.

2.1.4 Mass generation and Yukawa couplings

This subsection describes the process of mass generation through Yukawa couplings within the Higgs mechanism. The Higgs mechanism, involving spontaneous symmetry-breaking, not only forms the basis for the existence of the physical photon, W , and Z bosons but also leads to mass generation when particles interact with the Higgs field, which has a non-zero vacuum expectation value v . The physical W and Z bosons acquire mass in this way, while the photon remains massless.

The interaction of fermions with the Higgs field, giving them mass, is called a Yukawa interaction and it is mediated by the exchange of Higgs bosons [28]. The Yukawa term includes the Yukawa coupling y_f to the fermion f in the Lagrangian:

$$\mathcal{L}_{\text{Yukawa}} = -y_f (\bar{\varphi}_L \phi \varphi_R + \bar{\varphi}_R \phi^{\dagger} \varphi_L). \quad (2.7)$$

The minimal Higgs model is selected to represent ϕ as an isospin doublet of two complex scalar fields $\phi = (\phi^+, \phi^0)^T$. Using the unitary gauge, the Higgs doublet is rewritten either as $\phi = \phi_1 = \frac{1}{\sqrt{2}}(0, v+h)^T$, or as $\phi = \phi_2 = \frac{1}{\sqrt{2}}(v+h, 0)^T$ where h is the remaining neutral Higgs field. By substituting ϕ_1 into Equation 2.7, the contribution to the Lagrangian for the leptons and down-type quarks is obtained.

Similarly, substituting ϕ_2 into Equation 2.7 results in the contribution for the up-type quarks. The Lagrangian terms include mass terms, as well as interaction terms between the fermions and the neutral Higgs field h . Focusing on the quark mass terms and generalizing for multiple quark generations $i \in \{1, 2, 3\}$:

$$\mathcal{L}_{\text{Yukawa}} = -M_d^{ij} \bar{d}_L'^i d_R'^j + M_u^{ij} \bar{u}_L'^i u_R'^j + \text{h.c.},$$

with mass matrices $M_d^{ij} = \nu y_d^{ij} / \sqrt{2}$ and $M_u^{ij} = \nu y_u^{ij} / \sqrt{2}$.

The mass matrices can be diagonalized by unitary transformations using the matrices A_u, A_d, B_u, B_d , which transform between mass and weak quark eigenstates. The right-handed quark isospin singlets transform as $\varphi_R = u_R' = B_u u_R$ and $\varphi_R = d_R' = B_d d_R$, while, using $V = A_u^\dagger A_d$, the left-handed doublets transform as

$$\varphi_L = \begin{pmatrix} u_L' \\ d_L' \end{pmatrix} = \begin{pmatrix} A_u u_L \\ A_d d_L \end{pmatrix} = A_u \begin{pmatrix} u_L \\ V d_L \end{pmatrix}.$$

The effect on the charged and neutral current interactions is obtained by analyzing the Lagrangian terms in Equations 2.4, 2.5 and 2.6, and replacing the weak eigenstates with the mass eigenstates. While the neutral current interactions are unaffected, the charged current interaction Lagrangian term takes the form:

$$\mathcal{L}_{\text{CC}} = \frac{g}{\sqrt{2}} W_\mu^+ \bar{u}_L^i \gamma_\mu V^{ij} d_L^j + \text{h.c.}, \quad (2.8)$$

where V^{ij} are the elements of the V matrix, and i and j represent the up- and down-type quarks, respectively. The occurrence of the matrix elements V^{ij} introduces flavor changes in charged current weak interactions, which are not allowed in other interactions. The matrix V is called the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix, and it will be discussed in more detail in the following section.

2.1.5 The CKM matrix and CP violation

This subsection describes the properties of the Cabibbo-Kobayashi-Maskawa (CKM) matrix and how it relates to CP violation.

The CKM matrix elements appear in the description of the charged current in Equation 2.8, describes the mixing between the generations of quarks when they interact via the weak force. It was initially introduced by Nicola Cabibbo for two quark generations [7] and extended later by Makoto Kobayashi and Toshihide Maskawa to three generations [8]. The CKM matrix is mathematically represented as a unitary 3×3 matrix:

$$\underbrace{\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}}_{\text{Weak Eigenstates}} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{\text{CKM Matrix, } V_{\text{CKM}}} \underbrace{\begin{pmatrix} d \\ s \\ b \end{pmatrix}}_{\text{Mass Eigenstates}}.$$

Each element is a complex number indicating the probability amplitude for the quark of one flavor to change into a quark of another flavor. The probability of a transition from a quark of type i to a quark of type j is proportional to $|V_{ij}|^2$. The unitarity of the matrix ensures that the total probability of a given up-type quark transforming into any of the down-type quarks is conserved. The elements of the CKM matrix are fundamental parameters of the SM and their magnitudes can only be determined experimentally.

The CKM matrix can be expressed with various parameterizations. One commonly used parameterization includes three mixing angles $\theta_{12}, \theta_{13}, \theta_{23}$ and one complex phase δ to write the CKM matrix as [34]:

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (2.9)$$

with $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$. Another parameterization called the Wolfenstein parameterization [35] focuses on the hierarchy of the CKM matrix elements revealed by experimental measurements. It describes the CKM matrix using four real parameters A, λ, η , and ρ , with $\lambda \simeq 0.22$:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \quad (2.10)$$

The transitions within the same generation have probabilities close to unity, the ones between neighboring generations are suppressed (of the order λ or λ^2), and the transitions skipping a generation are highly suppressed (of the order λ^3).

The complex nature of the matrix elements determined in Equation 2.9 by the phase $\delta \neq 0$ and in Equation 2.10 by $\eta \neq 0$ is key when investigating the properties of the charged current weak interactions under the application of the CP operator. Due to the appearance of the complex CKM matrix elements in Equation 2.8, with $V_{ij} \neq V_{ji}^*$, the possibility of CP violation in flavor-changing charged currents is granted. The complex phase δ and the parameter η therefore introduce and account for CP violation in weak interactions.

In order to conserve total probability, the unitarity condition $VV^\dagger = 1$ needs to be fulfilled. By imposing orthogonality constraints on elements within rows and columns of the CKM matrix $V_{ij}V_{ik}^* = \delta_{jk}$ and $V_{ij}V_{kj}^* = \delta_{ik}$, vanishing relationships related to non-neighboring rows and columns of the CKM matrix appear. The most-quoted relationship is:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (2.11)$$

Dividing Equation 2.11 by $V_{cd}V_{cb}^*$, which is the best known of the three terms, and introducing rescaled parameters $\bar{\rho} = \rho(1 - \frac{1}{2}\lambda^2)$ and $\bar{\eta} = \eta(1 - \frac{1}{2}\lambda^2)$, Equation 2.11 can be represented graphically as a unitarity triangle in the $(\bar{\eta}, \bar{\rho})$ plane shown in Figure 2.2. The base of the triangle has unit length and the vertices of the triangle are located at $(0,0)$, $(0,1)$, and $(\bar{\eta}, \bar{\rho})$. The area of the triangle is therefore directly related to the parameter η , and is an indicator of the size of the CP -violation. The non-unit sides of the triangle are given by $V_{ud}V_{ub}^*/V_{cd}V_{cb}^* = \bar{\rho} + i\bar{\eta}$ and $V_{td}V_{tb}^*/V_{cd}V_{cb}^* = 1 - \bar{\rho} - i\bar{\eta}$. The angles are defined as:

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \quad \text{and} \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right).$$

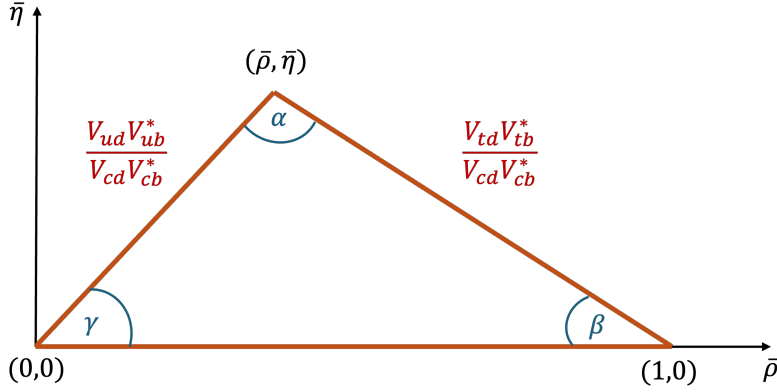


Figure 2.2: Visual representation of the CKM unitarity triangle described by Equation 2.11 in the $(\bar{\eta}, \bar{\rho})$ plane. The coordinates of the vertices and the non-unit lengths of the sides of the triangle are given.

2.2 Semileptonic $B \rightarrow X_u \ell \nu_\ell$ decays

Semileptonic decays of heavy mesons are mediated by the weak force and are characterized by the exchange of a W boson between hadronic and leptonic currents. This section provides details about the properties of these semileptonic decays and the various approaches to describe them mathematically.

One example of a semileptonic decay is the decay of a B meson into a lepton-neutrino pair and a hadron containing a charm or up quark ($B \rightarrow X_c \ell \nu_\ell$ and $B \rightarrow X_u \ell \nu_\ell$ decay, respectively). A Feynman diagram of this type of decay is shown in Figure 2.3. In this example, the \bar{b} quark of the B meson transforms into an \bar{u} or \bar{c} quark through the emission of a W boson. The W boson subsequently decays into a lepton-neutrino pair, and the spectator quark d combines with the \bar{u} or \bar{c} quark to form the final-state hadron, containing a charm (X_c) or an up (X_u) quark.

As described in the previous section, the probability of such a flavor-changing semileptonic weak decay is proportional to the magnitude of the relevant CKM matrix element squared. The origin of this dependence of the decay rate on the CKM matrix element is discussed in the following. For more detail, see Ref. [36].

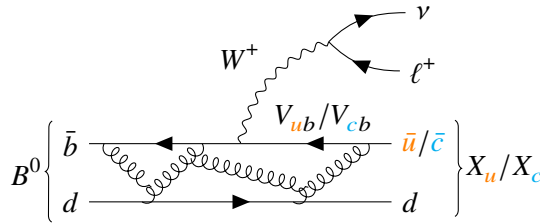


Figure 2.3: Feynman diagram for the flavor-changing semileptonic weak decay $B^0 \rightarrow X_u \ell \nu$ or $B^0 \rightarrow X_c \ell \nu$. The decay rates are proportional to $|V_{ub}|^2$ and $|V_{cb}|^2$, respectively. Examples of quark-gluon and gluon self-interactions are shown.

One of the two main ingredients for describing the decay rate is a phase-space factor, which accounts for the available kinematic configurations of the final-state particles. The kinematics of semileptonic decays into hadrons of a fixed mass m_X can be fully described using two variables. Typically, the

squared momentum transfer between the B meson and the hadron (q^2) and the lepton energy (E_ℓ) are chosen for this purpose. They are defined as

$$q^2 = (p_\ell + p_\nu)^2 = (p_B - p_X)^2 \quad \text{and} \quad E_\ell = \frac{p_B p_\ell}{m_B},$$

where p_B , p_ℓ and p_X are the momenta of the B meson, lepton, and hadron, respectively, and m_B is the mass of the B meson. Additionally, a recoil variable $w = (m_B^2 + m_X^2 - q^2)/(2m_B m_X)$ is defined, which is sometimes used in the kinematic description of the decay instead of q^2 . Even though the quantities E_ℓ and q^2 (or w) are fully sufficient to describe the kinematics, they are not independent. They are constrained by the momentum of the initial-state B meson, which defines the phase space for these decays.

The second main ingredient to the description of the decay rate is the squared magnitude of the matrix element. A distinctive feature of semileptonic decays is the factorization of the decay amplitude into separate hadronic and leptonic parts at leading electroweak order, in the regime where the W boson can be considered infinitely heavy. This is valid for all quark flavors, except for the top quark, which does not form hadrons, due to its short lifetime of 10^{-25} seconds [37]. For $B \rightarrow X_u \ell \nu_\ell$ decays, the hadronic and leptonic currents J_{had}^μ and $J_{\mu, \text{lep}}$, respectively, factorize, and the decay is described by an effective Hamiltonian

$$\mathcal{H}_{\text{eff}} = \frac{4G_F}{\sqrt{2}} J_{\text{had}}^\mu J_{\mu, \text{lep}}, \quad (2.12)$$

where G_F is the Fermi coupling constant. The effective Hamiltonian facilitates the calculation of the matrix element, which encompasses the dynamics of the decay.

In order to evaluate the decay rate, knowledge of the leptonic and hadronic currents is necessary. The leptonic current, corresponding to the decay of the W boson into a lepton and an antineutrino, is described using the definition of the charged-current interaction given in Equation 2.2:

$$J_{\text{lep}}^\mu = \bar{\nu}_\ell \gamma^\mu (1 - \gamma^5) \ell,$$

where ℓ stands for the lepton (electron or muon) and $\bar{\nu}_\ell$ for the corresponding antineutrino. The hadronic current involves the transition of a bottom (b) quark to either a charm or an up quark. The hadronic current is similarly related to the charged current given in Equation 2.2. Expressed in terms of the quark mass eigenstates it takes the form:

$$J_{\text{had}}^\mu = V_{\text{CKM}} \bar{q} \gamma^\mu (1 - \gamma^5) b,$$

where q represents the charm or up quark, and V_{CKM} denotes the element of the CKM matrix relevant for the transition. The CKM matrix element enters the hadronic current due to the nature of the weak quark eigenstates being different from the mass eigenstates.

Since the quarks are contained within mesons, the above expression is, however, not sufficient to describe the total hadronic current, and non-perturbative QCD effects and hadronization processes need to be considered. The prediction of these effects is the main challenge in calculating the total hadronic current, since they cannot be straightforwardly accounted for using perturbative techniques. The treatment of these non-perturbative effects depends on the nature of the hadron involved in the $B \rightarrow X \ell \nu_\ell$ decay. There are two methods that are commonly used. The first describes the hadronic current exclusively for specific final-state hadrons, such as π , D , ρ , or D^* . The other method takes X

as the inclusive sum of all individual final-state hadrons.

For both methods, non-perturbative effects in a hadronic matrix element are included as $\langle X_u | \bar{u} \gamma^\mu (1 - \gamma^5) b | B \rangle$. The matrix element of the local effective Hamiltonian given in Equation 2.12 for $B \rightarrow X_u \ell \nu_\ell$ decays is then written as

$$\mathcal{M} = -i \frac{G_F}{4\sqrt{2}} V_{ub} (\bar{\ell} \gamma^\mu (1 - \gamma^5) \nu_\ell) \langle X_u | \bar{u} \gamma^\mu (1 - \gamma^5) b | B \rangle, \quad (2.13)$$

where the leptonic and hadronic matrix elements factorize, as did the currents before. In the following subsections, the predictions of the hadronic matrix elements of charmless semileptonic decays are discussed. Section 2.2.1 and 2.2.2 describe the hadronic matrix elements for inclusive and exclusive decays, respectively.

2.2.1 Inclusive description

This subsection describes $B \rightarrow X_u \ell \nu_\ell$ decays, where X_u represents the inclusive sum of possible final states. To describe these decays mathematically, typically the framework of Heavy Quark Effective Theory (HQET) [38, 39] is employed. This framework addresses heavy quarks inside heavy mesons, such as the b quark within the B meson, a condition satisfied by $m_b \gg \Lambda_{\text{QCD}}$. In the HQET framework, the heavy quark is considered static, leading to a major finding: heavy-quark symmetries, where both the flavor and spin of the heavy quark are independent of the dynamics of the total system. Due to this decoupling, the hadronic matrix element can be expanded in powers of Λ_{QCD}/m_b .

Specifically, for inclusive decays, this expansion method is referred to as Heavy Quark Expansion (HQE) [40, 41]. Initially, the inclusive differential decay rate is expressed using the matrix element squared, as given in Equation 2.13. The phase-space factor described in the previous section enters into the integration of the differential decay rate:

$$\frac{d^3\Gamma}{dq^2 E_\ell E_{\nu_\ell}} = 2G_F^2 |V_{ub}|^2 W_{\mu\nu} L^{\mu\nu},$$

where the hadronic contribution to the decay rate is labeled as $W_{\mu\nu}$ and the leptonic contribution as $L_{\mu\nu}$. While $L_{\mu\nu}$ can be calculated analytically, the complex hadronic contribution contains the square of the hadronic matrix element, which includes non-perturbative effects. The key is to use the optical theorem [42] to relate the square of the hadronic matrix element to the forward matrix element of a scattering amplitude, thus connecting it to a time-ordered product of currents:

$$W_{\mu\nu} = -\frac{1}{\pi} \text{Im} \int d^4x e^{-iq_\alpha x^\alpha} \frac{\langle \bar{B} | T[J_{\text{had}}^{\dagger\mu}(x) J_{\text{had}}^\nu(0)] | B \rangle}{2m_B}.$$

The time-ordered product is then expressed as a series of i matrix elements of local operators $\mathcal{O}_{n+3,i}$ of increasing dimension n multiplied by Wilson coefficients $C_{n,i}$:

$$\int d^4x e^{-iq_\alpha x^\alpha} T[J_{\text{had}}^{\dagger\mu}(x) J_{\text{had}}^\nu(0)] = \sum_{n,i} \frac{1}{m_b^n} C_{n,i} \mathcal{O}_{n+3,i}.$$

While the Wilson coefficients are fully computed using perturbative QCD, the operators describe the

non-perturbative, long-distance effects. This method, known as operator-product expansion (OPE) [43], is valid in the limit $m_b \rightarrow \infty$ and effectively separates the perturbative from the non-perturbative effects.

The first nontrivial contributions in the expansion of the decay rate appear at order $1/m_b^2$ ($n = 2$), containing two matrix elements. The matrix elements are parameterized using the HQE parameters μ_π^2 and μ_G^2 , corresponding to the kinetic energy and the chromomagnetic moment of the b quark inside the B meson, respectively. At order $1/m_b^3$ ($n = 3$), two additional matrix elements contribute to the overall decay rate. They are parameterized using the HQE parameters ρ_D^3 and ρ_{LS}^3 , responsible for the Darwin and spin-orbit terms, respectively. At higher orders ($n > 3$), the number of contributing matrix elements becomes too large to introduce sensible parameterizations.

For the fully inclusive description, HQE using OPE works well, but the method reaches its limits for $B \rightarrow X_u \ell \nu_\ell$ decays, when it describes specific phase-space regions. Since the decay rate of $B \rightarrow X_u \ell \nu_\ell$ decays is suppressed by a factor of $|V_{ub}|^2/|V_{cb}|^2$, its measurement is only feasible in regions of phase space not dominated by $B \rightarrow X_c \ell \nu_\ell$ decays, such as the high lepton momentum region. In these regions, the fully inclusive description using OPE is no longer valid, and instead, a partial resummation of the local OPE is needed. In this case, the differential decay rate is no longer described as an explicit sum, but at leading order is written as:

$$\frac{d^3\Gamma}{dP^+ dP^- E_\ell} = \frac{G_F^2 |V_{ub}|^2}{192\pi^3} \int d\omega C(P^+, P^-, E_\ell, \omega) f(\omega) + \mathcal{O}\left(\frac{\lambda_{\text{QCD}}}{m_b}\right), \quad (2.14)$$

where $P^\pm = E_X \pm |\vec{p}_X|$ are the sum and difference of the energy and momentum of the hadron. The Wilson coefficients $C_{n,i}$ are now contained within a function $C(P^+, P^-, E_\ell, \omega)$ to be perturbatively calculated. The non-perturbative contribution described by the matrix elements of operators in OPE is now contained within the shape function $f(\omega)$. This function describes the residual b -quark momentum within the B meson, accounting for the deviation from the ideal HQE behavior, and the moments of $f(\omega)$ are given in terms of the HQE parameters. Information about the shape function may be extracted from $B \rightarrow X_u \ell \nu_\ell$, $B \rightarrow X_c \ell \nu_\ell$, and $B \rightarrow X_s \gamma$ decays, but its parameterization remains a source of uncertainty.

The free parameters in the description of the shape function are typically the HQE parameters, as well as the mass of the b quark, dictating the width of the shape function: $\bar{\lambda} = m_B - m_b$. Due to its confinement within the B meson, the mass of the b quark is affected by non-perturbative interactions with other quarks. Various schemes define m_b , each treating the perturbative and non-perturbative contributions differently. Three common mass schemes are the kinetic, the shape-function, and the Kagan-Neubert schemes. The kinetic scheme defines a scale-dependent b -quark mass based on the non-relativistic kinetic energy of the b quark. The shape-function scheme [44] specifically addresses the kinematic endpoint region by attributing some of the non-perturbative effects directly to m_b . The Kagan-Neubert scheme [45] defines m_b based on the mass of the Υ ($b\bar{b}$) bound state, correcting for binding energy and perturbative effects.

To fully describe the inclusive differential decay rates, various models have been developed, employing different mass schemes. One model is the Bosch, Lange, Neubert, and Paz (BLNP) model [46], which has two free parameters: the b -quark mass in the shape-function scheme m_b^{SF} and the kinetic-energy HQE parameter μ_π^2 . A competing model is the De Fazio and Neubert (DFN) model [47], which models the shape function using an ad-hoc exponential model with two free parameters. One is the mass of the b quark in the Kagan-Neubert scheme m_b^{KN} , defining the parameter

$\bar{\lambda} = m_B - m_b^{\text{KN}}$. The second parameter is $a^{\text{KN}} = -3\bar{\lambda}^2/\lambda_1 - 1$, with $\lambda_1 = -\mu_\pi^2$.

By selecting an appropriate model to predict the differential decay rate of $B \rightarrow X_u \ell \nu_\ell$ decays given in Equation 2.14 and comparing it to measured partial rates, values of $|V_{ub}|$ become accessible.

2.2.2 Exclusive description

This subsection describes exclusive $B \rightarrow X_u \ell \nu_\ell$ decays, where X_u corresponds to a specific final state. To describe these decays mathematically, the hadronic matrix element in Equation 2.13 is separated into contributions from vector and axial-vector currents. Lorentz-invariant functions, known as form factors, are constructed to parameterize these matrix elements, taking non-perturbative effects into account. Due to the $V-A$ structure of the currents, depending on the nature of the X_u meson, either only the vector current or both vector and axial-vector currents contribute to the total hadronic matrix element.

For decays to pseudoscalar mesons, $B \rightarrow P \ell \nu_\ell$ (such as $B \rightarrow \pi \ell \nu_\ell$), only the vector current contributes, and the corresponding matrix element is parameterized in terms of two form factors $f_+(q^2)$ and $f_0(q^2)$, which are typically given as functions of q^2 . The matrix element is defined as

$$\langle P(p_P) | \bar{u} \gamma^\mu b | B(p_B) \rangle = f_+(q^2) \left(p_B^\mu + p_P^\mu - \frac{m_B^2 - m_P^2}{q^2} q^\mu \right) + f_0(q^2) \frac{m_B^2 - m_P^2}{q^2} q^\mu,$$

where p_P and m_P are the momentum and mass of the hadron. Since $f_0(q^2)$ is negligible in the case $\ell = e, \mu$, the differential decay rate of $B \rightarrow P \ell \nu_\ell$ is expressed as a function of $f_+(q^2)$ and $|V_{ub}|$:

$$\frac{d\Gamma(B \rightarrow P \ell \nu)}{dq^2} = \frac{|\vec{p}_P|^3}{24\pi^3} G_F^2 |V_{ub}|^2 |f_+(q^2)|^2, \quad (2.15)$$

Often the decay rate is also expressed in terms of the Källén function $\lambda(q^2)$ with $\lambda(q^2) = (m_B^2 + m_P^2 - q^2)^2 - 4m_B^2 m_P^2$ and thus $p_P = \lambda^{1/2}(q^2)/2m_B$.

For decays to vector mesons, $B \rightarrow V \ell \nu_\ell$ (such as $B \rightarrow \rho \ell \nu_\ell$), both the vector and axial-vector currents contribute to the total hadronic matrix element. In addition, the polarization vector of the meson ϵ needs to be considered. The vector current contribution to the matrix element is parameterized by the form factor $V(q^2)$, while the axial-vector current contribution is parameterized by three form factors $A_i(q^2)$ with $i \in \{0, 1, 2\}$:

$$\begin{aligned} \langle V(p_V, \epsilon) | \bar{u} \gamma^\mu b | B(p_B) \rangle &= V(q^2) \epsilon^{\mu\sigma} \frac{2p_B^\nu p_V^\rho}{m_B + m_V}, \\ \langle V(p_V, \epsilon) | \bar{u} \gamma^\mu \gamma^5 b | B(p_B) \rangle &= i\epsilon_\nu^* \left[A_0(q^2) \frac{2m_V q^\mu q^\nu}{q^2} \right. \\ &\quad \left. + A_1(q^2) (m_B + m_V) \eta^{\mu\nu} - A_2(q^2) \frac{(p_B + p_V)_\sigma q^\sigma}{m_B + m_V} \eta^{\mu\sigma} \right], \end{aligned}$$

where p_V and m_V are the momentum and mass of the hadron, and $\eta^{\mu\nu} = g^{\mu\nu} - q^\mu q^\nu / q^2$ is the transverse part of the metric. Furthermore, a form factor $A_{12}(q^2)$ is defined in terms of $A_1(q^2)$ and

$A_2(q^2)$:

$$A_{12}(q^2) = A_2(q^2) \frac{|\vec{p}_V|^2 m_B}{4m_V(m_B + m_V)} - A_1(q^2) \frac{(m_B^2 - m_V^2 - q^2)(m_B + m_V)}{16m_B m_V}.$$

This modification allows for a simpler translation of the four form factors $A_0(q^2)$, $A_1(q^2)$, $A_{12}(q^2)$, and $V(q^2)$ to the helicity basis:

$$\begin{aligned} H_{\pm}(q^2) &= V(q^2) \frac{2m_B |\vec{p}_V|}{m_B + m_V} \pm A_1(q^2)(m_B + m_V), \\ H_0(q^2) &= A_{12}(q^2) \frac{8m_B m_V}{q}, \\ H_s(q^2) &= A_0(q^2) \frac{2m_B |\vec{p}_V|}{q}. \end{aligned}$$

For decays to light leptons $\ell = e, \mu$, only three form factors are relevant: H_{\pm} and $H_0(q^2)$ (corresponding to $A_1(q^2)$, $A_2(q^2)$, and $V(q^2)$). The differential decay rate of $B \rightarrow V\ell\nu_{\ell}$ as a function of H_{\pm} , $H_0(q^2)$, and $|V_{ub}|$ is then given by

$$\frac{d\Gamma(B \rightarrow V\ell\nu)}{dq^2} = \frac{|\vec{p}_V| q^2}{96\pi^3 m_B^2} G_F^2 |V_{ub}|^2 \left(|H_0(q^2)|^2 + |H_+(q^2)|^2 + |H_-(q^2)|^2 \right). \quad (2.16)$$

To predict the differential decay rates of such exclusive decays using Equations 2.15 and 2.16, the nature of the form factors must be understood. Since they are hadronic functions containing non-perturbative effects, they cannot be calculated analytically. Instead, they are predicted using QCD methods, such as QCD light-cone sum rule (LCSR) techniques or lattice QCD (LQCD) [36].

LCSR methods combine elements of QCD with methods from OPE, discussed in the previous section, and perturbation theory. These techniques are based on the concept of a light-cone, referring to the limit at which the relevant interactions occur nearly at the speed of light. This implies that LCSR predictions hold at low q^2 where the final-state hadron is highly relativistic. Towards higher q^2 , the LCSR predictions become less reliable, as the assumptions and truncations made in the sum-rule approach become less justified.

In contrast, LQCD predictions are most reliable in the high q^2 region. These predictions are obtained through numerical simulations of QCD on a discretized spacetime lattice. At high q^2 , the relevant energy scales are low, allowing the full QCD dynamics to be simulated. This becomes increasingly difficult towards lower q^2 .

The two methods therefore provide complementary predictions at opposite ends of the phase space. By combining these predictions with measured decay rates in the relevant phase-space regions, $|V_{ub}|$ values can be determined. To utilize available experimental information from the entire phase space, form-factor parameterizations that account for the entire q^2 range may be introduced. Possible form-factor parameterizations are described below.

By fitting these form-factor parameterizations to measured decay-rate spectra, free parameters of the form-factor parameterizations can be extracted. At this point, both the normalizations of the form factors and $|V_{ub}|$ are absorbed by these parameters. Once theoretical input from LCSR or LQCD calculations is added to provide information about the normalizations of the form factors, $|V_{ub}|$

becomes accessible.

Form-factor parameterizations

To account for the entire q^2 range, the q^2 dependence of the form factors is interpolated between the high and low q^2 regions using analyticity and unitarity arguments.

One approach uses simple phenomenologically motivated parameterizations for the form factors, such as the Becirevic-Kaidalov (BK) parameterization [48]. Another technique employs dispersion relations to expand in powers of the conformal variable $z(q^2, q_0^2)$ defined as

$$z(q^2, q_0^2) = \frac{\sqrt{m_+^2 - q^2} - \sqrt{m_+^2 - q_0^2}}{\sqrt{m_+^2 - q^2} + \sqrt{m_+^2 - q_0^2}}, \quad (2.17)$$

where $m_\pm = m_B \pm m_X$ is the sum (or difference) of the masses of the B meson and the hadron, corresponding to the threshold of production of hadron pairs from the vacuum. The expansions in powers of the parameter $z(q^2, q_0^2)$ correspond to a conformal mapping of q^2 to a small domain where the series can converge rapidly. The optimal value of the free parameter q_0^2 is chosen to ensure convergence of the series.

Examples of such expansions are the Boyd-Ginstein-Lebed (BGL) [49], the Bourrely-Caprini-Lellouch (BCL) [12], and the Bharucha-Straub-Zwicky (BSZ) [13] parameterizations. The form-factor parameterizations most commonly used to describe exclusive semileptonic $B \rightarrow X_u \ell \nu_\ell$ decays are described briefly below.

The BGL expansion: The Boyd-Ginstein-Lebed (BGL) parameterization [49] expands the form factors of decays to pseudoscalar mesons $f(q^2)$ ($f_0(q^2)$ and $f_+(q^2)$) in terms of the parameter z defined in Equation 2.17 using form-factor coefficients b_k up to the expansion order K :

$$f(q^2) = P(q^2) \phi^{-1}(q^2) \sum_{k=0}^{K-1} b_k z^k. \quad (2.18)$$

Here $\phi(q^2)$ is an outer function accounting for unitarity constraints. The exact form is chosen to ensure the series converges [49]. The inverse Blaschke factor $P(q^2) = z(q^2, m_R)$ accounts for poles below the pair production threshold. The mass of the resonance m_R depends on the allowed angular momentum and parity. The optimal value of the free parameter proposed by Ref. [50] is $q_0^2 = -0.65m_-^2$. For decays to vector mesons, the form factors (H_\pm and $H_0(q^2)$) are first rewritten as

$$\begin{aligned} \mathcal{F}_1(q^2) &= q H_0(q^2), \\ f(q^2) &= \left(H_-(q^2) + H_+(q^2) \right) / 2, \\ g(q^2) &= \left(H_-(q^2) - H_+(q^2) \right) / 2T, \end{aligned}$$

with $T = m_B m_X \sqrt{w^2 - 1}$. The newly defined form factors $\mathcal{F}_1(q^2)$, $f(q^2)$, and $g(q^2)$ are then expanded as described by Equation 2.18.

The BCL expansion: The Bourrely-Caprini-Lellouch (BCL) parameterization [12] is similar to the BGL parameterization. It expands the form factors for decays to pseudoscalar and vector mesons $f(q^2)$ in terms of the parameter z defined in Equation 2.17 up to expansion order K :

$$f(q^2) = (1 - q^2/m_R^2)^{-1} \sum_{k=0}^{K-1} b_k \left[z^k - (-1)^{k-K} \frac{k}{K} z^K \right]. \quad (2.19)$$

Here, instead of employing Blaschke factors, the pole with m_R is directly included in the expansion. The optimal free parameter for the BCL expansion is $q_0^2 = m_+ (\sqrt{m_B} - \sqrt{m_X})^2$ [12].

The BSZ expansion: A variation of the BCL expansion is the Bharucha-Straub-Zwicky (BSZ) parameterization [13]. The BSZ parameterization orders the power expansion differently, as a series expansion around $q^2 = 0$. The form factors $f(q^2)$ take the form

$$f(q^2) = (1 - q^2/m_R^2)^{-1} \sum_{k=0}^{K-1} b_k \left(z(q^2) - z(0) \right)^k. \quad (2.20)$$

The BK parameterization: The Becirevic-Kaidalov (BK) parameterization [48] differs from the above expansions as it relies on a more phenomenological approach to describe the form factors. The form factor $f_+(q^2)$ is parameterized by two free parameters $f_+(0)$ and α_{BK} , accounting for the normalization and shape of the form factor, respectively. The form factor $f_+(q^2)$ is defined as

$$f_+(q^2) = (1 - q^2/m_R^2)^{-1} \frac{f_+(0)}{(1 - \alpha_{BK} q^2/m_R^2)}. \quad (2.21)$$

2.3 Status of $|V_{ub}|$ measurements

Since the first measurements of $|V_{ub}|$ conducted by the CLEO collaboration in the early 1990s [51], numerous exclusive and inclusive measurements have been made by the BaBar, Belle, CLEO, and LHCb collaborations. Recently, the Belle collaboration performed a simultaneous exclusive and inclusive measurement of $|V_{ub}|$ [52]. This section describes the current status of $|V_{ub}|$ measurements.

Over time, a persistent tension has been observed between the $|V_{ub}|$ world averages from exclusive and inclusive methods, presenting a longstanding puzzle regarding the origin of this discrepancy. The current world averages determined by the Heavy Flavor Averaging Group (HFLAV) [9] are:

$$\begin{aligned} |V_{ub}|^{\text{excl}} &= (3.67 \pm 0.09 \pm 0.12) \times 10^{-3}, \\ |V_{ub}|^{\text{incl}} &= (4.19 \pm 0.12^{+0.11}_{-0.12}) \times 10^{-3}, \end{aligned}$$

where the first uncertainty is experimental and the second is theoretical. The results differ by approximately 2.3 standard deviations.

This tension is also evident in measurements of $|V_{cb}|$. Figure 2.4 shows the world averages of the exclusive and inclusive measurements of $|V_{ub}|$ and $|V_{cb}|$ in the two-dimensional plane of $|V_{ub}|$ versus $|V_{cb}|$. The result of a combined fit performed by the HFLAV group using the world average exclusive $|V_{ub}|$ and $|V_{cb}|$ measurements, as well as the world average $|V_{ub}|/|V_{cb}|$ measurement, is also shown. This result is compared to the inclusive values of $|V_{ub}|$ and $|V_{cb}|$, highlighting the visible tension.

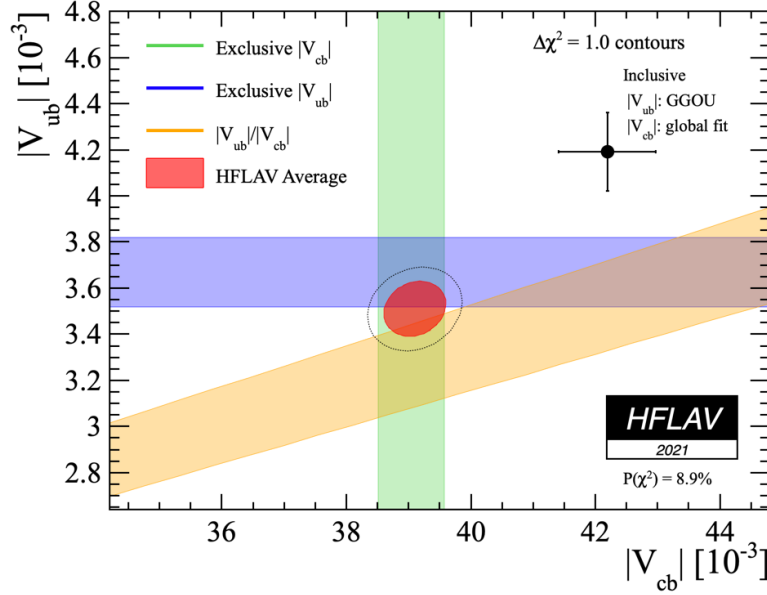


Figure 2.4: A two-dimensional plot of $|V_{ub}|$ versus $|V_{cb}|$. The blue and green bands show the world average exclusive values of $|V_{ub}|$ and $|V_{cb}|$, respectively. The orange band shows the world average of the $|V_{ub}|/|V_{cb}|$ measurements. The result of a combined two-dimensional fit to these values is shown in red, with the $\Delta\chi^2 = 1.0$ contour given. The χ^2 probability of the fit is 8.9%. A tension is observed between the fit result and the inclusive values of $|V_{ub}|$ and $|V_{cb}|$ (black data point), where the measurements were performed using the GGOU shape-function [53] and a global fit, respectively. Image from Ref. [9].

The discrepancy may originate from both experimental and theoretical sources. On the theoretical side, in exclusive measurements, theoretical predictions of the form factors from LQCD or LCSR calculations are needed to separate the form-factor normalization from the magnitude of the CKM matrix element. Uncertainties in these predictions directly affect the measured values. In inclusive measurements, predictions of the total rate are derived from HQE, OPE, and shape-function modeling, each with their own uncertainties contributing to the determination of the CKM matrix element. Experimentally, for both exclusive and inclusive measurements, factors such as background mismodeling can have systematic effects on the extracted results. A combined effort from theory and experiment is necessary to resolve the tension.

To investigate experimental effects in exclusive decays, testing whether the tension is present for different decay modes is valuable. For example, the measurements composing the exclusive world average of $|V_{ub}|$ come from $B \rightarrow \pi \ell \nu_\ell$ decays only. One could consider measuring $|V_{ub}|$ from other charmless decays, such as $B \rightarrow \rho \ell \nu_\ell$, $B \rightarrow \omega \ell \nu_\ell$, or $B \rightarrow \eta^{(\prime)} \ell \nu_\ell$. However, the determinations of $|V_{ub}|$ from these decays are limited by theoretical predictions. Since the ρ , ω , η , and η' mesons are unstable, LQCD cannot reliably predict the form factors.

While LCSR predictions for $B \rightarrow \rho \ell \nu_\ell$ and $B \rightarrow \omega \ell \nu_\ell$ are available, discussions about their validity are ongoing. Assuming they are reliable, these predictions have been used by Ref. [10] in combination with averages of previous measurements of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and $B \rightarrow \omega \ell \nu_\ell$ to obtain $|V_{ub}|$ results:

$$\begin{aligned} |V_{ub}|^{\rho\ell\nu} &= (2.96 \pm 0.29) \times 10^{-3}, \\ |V_{ub}|^{\omega\ell\nu} &= (2.99 \pm 0.35) \times 10^{-3}. \end{aligned}$$

These results are lower than the exclusive ($B \rightarrow \pi\ell\nu_\ell$) world average, approximately 2.2 standard deviations away. They also display a larger tension of approximately 3.7 standard deviations from the inclusive world average $|V_{ub}|$. Instead of contributing to resolving the tension, this difference introduces another feature yet to be understood.

As discussed in Section 2.1.5, the CKM matrix elements are crucial inputs in defining the unitarity triangles. The current status of the constraints on the sides and angles of the CKM unitarity triangle described by Equation 2.11 in the $(\bar{\eta}, \bar{\rho})$ plane is shown in Figure 2.5. The green band corresponds to the constraint imposed by $|V_{ub}|$.

To define the constraints, the CKMfitter group determined an average of the exclusive and inclusive measurements, $|V_{ub}| = (3.86 \pm 0.14) \times 10^{-3}$, and used it as the nominal input [54] to the global fit. The global fit is able to constrain $|V_{ub}|$, predicting a value of $|V_{ub}| = (3.67 \pm 0.11) \times 10^{-3}$, which is in agreement with the exclusive HFLAV average. Assuming the tension in $|V_{ub}|$ is reduced, the unitarity triangle could be further constrained. The tension in $|V_{ub}|$ currently limits the power of $|V_{ub}|$ in determining the vertex of the unitarity triangle, directly limiting the constraints on the Standard Model.

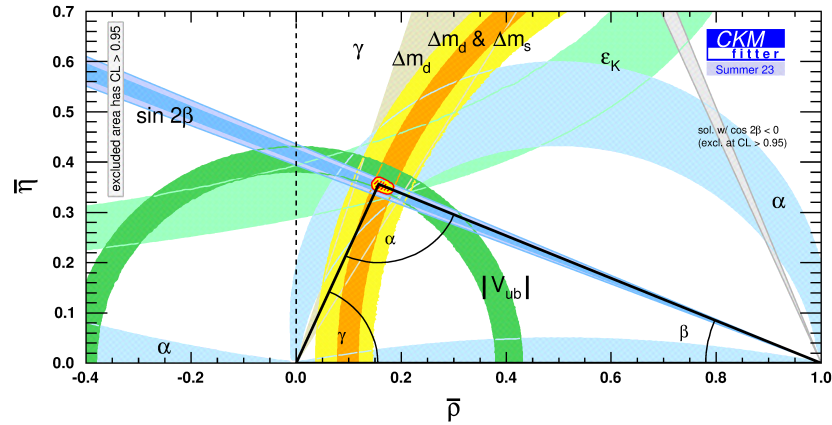


Figure 2.5: Constraints on the sides and angles of the CKM unitarity triangle in the $(\bar{\eta}, \bar{\rho})$ plane, including constraints from direct measurements of the angles α , β , and γ . In addition, constraints are placed using the kaon mixing parameter ϵ_K , and the mass-mixing parameters in the B_d and B_s meson systems, Δm_d and Δm_s . The constraint from $|V_{ub}|$ is given by the green band. The result of the global fit is shown as the red-hashed region. Areas with confidence levels greater than 0.95 were excluded. Image from Ref. [55].

SuperKEKB and Belle II

This chapter describes the experimental facilities used to collect the data used in this work. The main goal of the Belle II experiment is to achieve precision measurements, focusing primarily on B mesons and charmed mesons. This includes refining the understanding of the fundamental CKM parameters and exploring phenomena beyond the Standard Model. The Belle II experiment relies on electron-positron collisions produced by the SuperKEKB accelerator, described in Section 3.1.

The electron-positron collisions produce pairs of B mesons, along with other processes, discussed in Section 3.2. The Belle II detector (see Section 3.3), located at the interaction point of the accelerator, detects the subsequent particle interactions and decays. Detailed information about the SuperKEKB accelerator and the Belle II detector is available in Refs. [56, 57] and Ref. [58], respectively. Descriptions of the algorithms and methods used for particle reconstruction and identification are provided in Section 3.4.

3.1 The SuperKEKB accelerator

This section provides details about the key specifications and capabilities of the SuperKEKB accelerator.

The SuperKEKB accelerator, located at the High Energy Accelerator Research Organization (KEK) in Tsukuba, Japan, is an asymmetric-energy two-ring electron-positron collider. The primary objective of the SuperKEKB accelerator is to accelerate electrons and positrons, which are tuned to collide at the center of mass (CM) energy corresponding to the $\Upsilon(4S)$ resonance, at $\sqrt{s} = 10.58$ GeV. The $\Upsilon(4S)$ resonance is a radially excited bottomonium state, consisting of a bottom and an anti-bottom quark ($b\bar{b}$). The collision energy of the SuperKEKB accelerator is just above the threshold to produce a pair of B mesons, which are particles containing two quarks, one of which is a bottom quark. The CM energy is therefore specifically chosen to facilitate the production of $B\bar{B}$ pairs, which is why the SuperKEKB accelerator is considered a “B-factory”.

The SuperKEKB accelerator is the successor of the KEKB accelerator, which operated from 1999 to 2010 [59]. It has undergone several phases of operation. In 2016, Phase 1 marked the initiation of operations with single beams. By 2018, Phase 2 witnessed the assembly of most of the detector components, resulting in the first collisions. Subsequently, in 2019, Phase 3 saw the completion of the full detector setup. The accelerator experienced its first extended shutdown in 2022 to optimize its performance, and restarted collisions in 2024.

The acceleration of electrons begins with their production. A thermionic gun with a barium-impregnated tungsten cathode produces electrons, some of which are led to a tungsten target producing positrons. Subsequently, positrons undergo emittance reduction in a damping ring, while electrons follow a different path. The positrons and electrons are accelerated to energies of 4.0 GeV and 7.0 GeV, respectively, using a linear accelerator. Positrons then circulate in the low-energy ring (LER), and electrons in the high-energy ring (HER). A final-focus superconducting magnet system brings the beams to cross to produce electron-positron collisions inside the Belle II detector, described further in Section 3.3. A schematic layout of the accelerator is depicted in Figure 3.1.

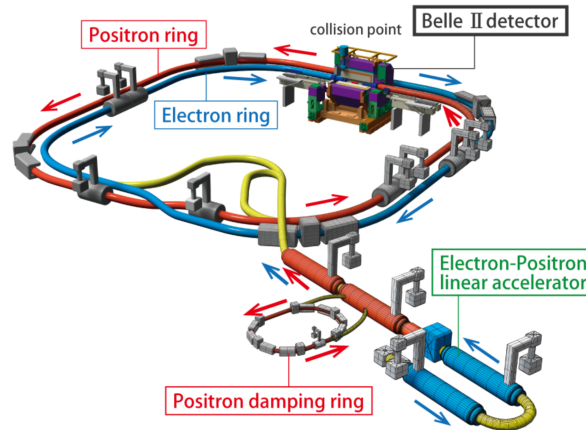


Figure 3.1: Sketch of the layout of the SuperKEKB electron-positron accelerator. Image from Ref. [56].

The intensity of particle collisions produced by accelerators is defined by the instantaneous luminosity. It quantifies the rate at which particles interact per unit area and time. The design luminosity of the SuperKEKB accelerator is $6.5 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ [60], which is 40 times higher than the luminosity of its predecessor the KEKB accelerator.

This ambitious goal is based on the implementation of the nanobeam scheme [56, 61, 62]. In addition to increasing the beam currents, contributing a factor two increase in luminosity, it aims to minimize the longitudinal beam overlap. The decrease in collision area is achieved by introducing a large crossing angle and keeping the horizontal beam size small. This results in an additional factor 20 increase in luminosity.

The SuperKEKB accelerator already holds the current world record for peak instantaneous luminosity with $3.1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ [60]. Some encountered challenges related to beam injection and collimation, resulting in intolerable levels of beam-induced backgrounds, have however also led to reduced performance compared to the designed performance.

In addition to the instantaneous luminosity, another key performance metric is the total integrated luminosity, which represents the cumulative measure of particle collisions over a given time interval. The total recorded integrated luminosity of the SuperKEKB accelerator from the beginning of Phase 3 in 2019 to the extended shutdown in 2022 is 424 fb^{-1} [11, 63].

3.2 Particle collisions and beam-induced backgrounds

During the collision of the electron and positron beams, in addition to the production of $B\bar{B}$ pairs, various other interaction processes occur. In the context of this thesis, these will collectively be referred to as backgrounds. The dominant processes can be categorized into continuum, other QED, and beam-induced backgrounds. This section will discuss the production of $B\bar{B}$ pairs and the types of backgrounds encountered and their sources.

$B\bar{B}$ production

The primary target of the SuperKEKB accelerator is the production of $B\bar{B}$ pairs, which occurs via the production of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ meson is the third radial excitation of the Υ ($J^{PC} = 1^{--}$) bottomonium ($b\bar{b}$) bound state. The mass of the $\Upsilon(4S)$ meson, 10.58 GeV, is just above the threshold to produce a $B\bar{B}$ pair, as twice the mass of a B meson is approximately 10.56 GeV. Due to the proximity to the energy threshold, the $B\bar{B}$ production leaves very little available energy to produce additional particles. This allows for tight constraints on the kinematics of the produced $B\bar{B}$ pair. The $\Upsilon(4S)$ meson decays into a $B\bar{B}$ pair more than 96% of the time [37]. The resonant production of an $\Upsilon(4S)$ meson with the subsequent production of a $B\bar{B}$ pair is shown in a Feynman-style diagram in Figure 3.2.

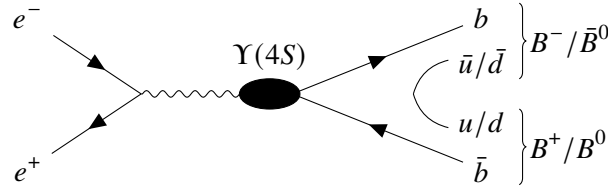


Figure 3.2: Sketch of an electron-positron interaction resulting in resonant production of an $\Upsilon(4S)$ meson with subsequent production of a $B\bar{B}$ pair.

Continuum production

Instead of yielding an $\Upsilon(4S)$ meson, the electron-positron collisions at the SuperKEKB accelerator can produce nonresonant quark pairs. These QED background events, collectively referred to as continuum events, are characterized by the production of quark-antiquark ($q\bar{q}$) pairs without the presence of a specific resonance peak. The term “continuum” is chosen because $q\bar{q}$ pair production is possible at any energy surpassing twice the respective quark mass threshold. The absence of a resonance peak holds true even when scanning across energy levels. At the CM energy of the SuperKEKB accelerator, the potential nonresonant $q\bar{q}$ pairs include up-quark pairs ($u\bar{u}$), down-quark pairs ($d\bar{d}$), charm-quark pairs ($c\bar{c}$), and strange-quark pairs ($s\bar{s}$), with $u\bar{u}$ and $c\bar{c}$ pair production being the most dominant.

Other QED background production

In addition to the production of continuum $q\bar{q}$ events, other QED production processes result from electron-positron collisions at the SuperKEKB accelerator. These processes include the pair production of leptons ($e^+e^- \rightarrow \ell^+\ell^-$), where $\ell \in \{e, \mu, \tau\}$, and the annihilation into two photons ($e^+e^- \rightarrow \gamma\gamma$).

Electron-positron collisions may also produce two virtual photons, which can then generate a pair of low-momentum electrons, muons, or tau leptons through pair production while retaining the original electron-positron pair. This is known as a two-photon process: $e^+e^- \rightarrow e^+e^-\ell^+\ell^-$, where $\ell \in \{e, \mu, \tau\}$.

The dominant QED process is Bhabha scattering ($e^+e^- \rightarrow e^+e^-$), with a cross section approximately three times larger than that of $\Upsilon(4S)$ meson production [64]. Its significance is further emphasized when a photon is generated during the scattering process, resulting in radiative Bhabha scattering: $e^+e^- \rightarrow e^+e^-\gamma$. The emitted photon may interact with the detector material, leading to pair production and the generation of false hits within the detector.

Most of these QED background processes are characterized by low-multiplicity track topologies, meaning four or fewer particles are in the final state. These are commonly referred to as “low-multiplicity backgrounds”. Exceptions include processes involving tau leptons, which decay further, often resulting in a larger number of final-state particles. For these backgrounds, the multiplicity depends on the decay mode of each tau lepton.

Beam-induced backgrounds

A significant background category not directly related to the electron-positron collisions, includes beam-induced backgrounds. These backgrounds arise from interactions among beam particles or between beam particles and residual gas particles in the beam pipe. Although these processes occur along the entire length of the beam pipe, they are particularly problematic near the interaction region. The particles produced in these interactions can mimic signal hits, making event reconstruction challenging. The three main sources of beam-induced backgrounds are Touschek scattering, beam-gas scattering, and synchrotron radiation, as discussed below. For more detailed information on the levels of these backgrounds at the SuperKEKB accelerator, refer to Ref. [65].

Touschek scattering backgrounds [66] result from collisions and Coulomb scattering between particles within the same bunch, causing particles to leave the bunch. Since the rate of this process is inversely proportional to the beam size, it is a dominant effect at the SuperKEKB accelerator. However, its impact is reduced slightly because the rate is also inversely proportional to the beam energy.

Beam-gas scattering occurs due to the imperfect vacuum inside the beam pipe. Residual gas molecules can interact with beam particles through Coulomb interactions and Bremsstrahlung. This process is similar to Touschek scattering, where beam particles leave the bunch and create showers in the detector.

Synchrotron radiation is a third type of beam-induced background, emitted as particles are accelerated radially outward. Within the detector, strong magnetic focusing amplifies this effect. As with the other two background processes, the emitted low-energy photons can interact with the detector, generating false hits.

3.3 The Belle II detector

This section provides a comprehensive description of the Belle II detector. The Belle II detector [58] is positioned around the collision point of the SuperKEKB accelerator, serving the purpose of detecting and recording the particle collisions. It succeeds the Belle detector [67], which operated at the collision point of the KEKB accelerator from 1999 to 2010. The Belle II detector commenced data collection

after its full commissioning in March 2019.

The detector is constructed of multiple subdetectors placed around a double-walled beryllium beam pipe with an inner radius of 10 mm. The subdetectors within the Belle II detector serve specific functions, each contributing to the overall functionality. The detector encompasses a tracking system, featuring the pixel detector (PXD), a silicon vertex detector (SVD), and the central drift chamber (CDC). Beyond the CDC, the time-of-propagation counter (TOP) and the aerogel ring-imaging Cherenkov detector (ARICH) cover the barrel and forward end-cap regions, respectively, contributing to charged particle identification. The electromagnetic calorimeter (ECL) is positioned outside the TOP and occupies the remaining volume within a superconducting solenoid magnet with a field strength of 1.5 T. The magnetic field is crucial for measuring the momentum and charge of detected particles. The K_L and muon detector (KLM) is situated outside the solenoid coil. The data recorded by the Belle II subdetectors passes through a trigger system before being passed on to the data acquisition system. A schematic representation of the detector layout, illustrating the arrangement of the subdetectors, can be found in Figure 3.3.

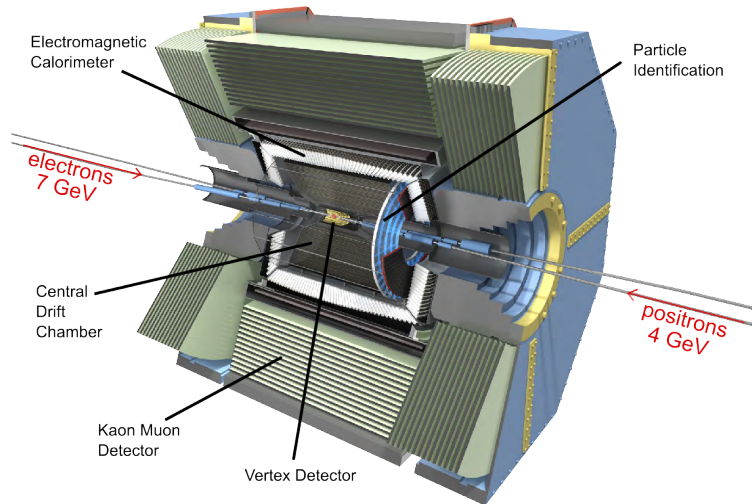


Figure 3.3: Sketch of the layout of the Belle II detector. The incoming positron and electron beams are shown, and the subdetectors labeled. The vertex detector consists of the pixel detector and the silicon vertex detector. Image adapted from Ref. [68].

The Belle II detector adheres to a right-handed Cartesian coordinate system, with the origin at the interaction point (IP). The z -axis in the laboratory frame aligns with the electron beam direction (HER), the y -axis points upwards, and the x -axis extends radially outward. Polar coordinates (r, θ, ϕ) are also utilized for precise descriptions. The detector's angular acceptance ranges from 17 to 150 degrees in θ , covering the entire 2π range in ϕ . The detector is divided into three regions in θ : forward, barrel, and back. The asymmetry in the beam energies results in a Lorentz boost of the CM frame with respect to the laboratory frame, and the detector layout intentionally reflects this asymmetry.

Below, more detail is provided about the purpose and the underlying physical processes resulting in particle detection for each subdetector. For more detailed information about interactions of particles with detectors see, for example, Ref. [69].

Pixel detector (PXD)

The Belle II pixel detector (PXD) is positioned closest to the interaction point (IP) as the innermost subdetector within the vertex detector (VXD), enabling precise vertexing for identifying short-lived particles within an angular acceptance of 17 to 150 degrees [58]. The PXD comprises two layers of DEpleted p-channel Field Effect Transistor (DEPFET) [70, 71] silicon detectors located at 14 mm and 22 mm from the IP, utilizing monolithic silicon sensors arranged in modules. Each module consists of 768×250 DEPFET sensors, totaling approximately 8 million pixels. Until the first extended shutdown, the first layer includes 16 modules, and the second layer comprises 8 modules. For further information on the current status and plans, see Ref [72].

The DEPFET operating principle involves a p-channel Metal Oxide Semiconductor Field-Effect Transistor (MOSFET) integrated onto a fully depleted silicon substrate, creating a potential minimum, the “Internal Gate,” for electron accumulation [70, 71]. Incident particles produce electron-hole pairs within the fully depleted bulk. The electrons accumulate in the Internal Gate, and upon switching on the transistor, the electrons modulate the channel current. The sensor simultaneously detects and internally amplifies signals, allowing for a non-destructive readout.

To address challenges such as high hit rates near the beam pipe due to beam-induced background, the PXD employs thin sensors ($75 \mu\text{m}$) with no additional support and cooling material in the active region, ensuring minimal multiple scattering. The detector has variable pixel sizes of $50 \mu\text{m} \times (55 - 85) \mu\text{m}$, and a high readout rate of 50 kHz, effectively limiting background occupancy to 1-2%. In addition to mitigating multiple scattering, minimizing radiation damage is a challenge. Both, DEPFET sensors and the associated readout chips, are designed to withstand up to 10 MRad without significant deterioration of performance [58]. The average spatial resolution after four years of data taking is approximately $15 \mu\text{m}$, with a hit efficiency of 98% [73].

Silicon vertex detector (SVD)

The silicon vertex detector (SVD) is the second component of the vertex detector (VXD) and surrounds the PXD [58, 74]. Its primary objectives include precise vertexing of short-lived particles and momentum measurements. It is particularly effective in tracking low-momentum particles and facilitates ionization-loss measurements for particle identification. The SVD is composed of four layers of double-sided silicon strip detectors (DSSDs) made of six-inch wafers, amounting to a total of 172 DSSDs. The layers are positioned at radii between 38 and 140 mm from the IP. To achieve the same angular acceptance as the PXD, the first layer is aligned entirely parallel to the beam pipe, while the forward sections of subsequent layers are slanted with respect to the beam axis.

The main structure of DSSDs consists of *p*-doped sensing strips implanted in the *n*-type silicon bulk, facing the beam axis (*z*), and *n*-doped strips implanted on the opposite side along *r*– ϕ . The *n*-doped strips are located on the sensor face towards the outside and are placed orthogonally to the *p*-doped sensors. Upon the incidence of a particle, an electron-hole pair is generated, and through the application of an electric field, electrons and holes move to opposite sides, inducing a charge on the sensors. The orthogonal arrangement of the sensors is important for providing three-dimensional trajectory information.

Using DSSDs with a thickness of approximately $300 \mu\text{m}$ results in a low material budget. The use of strips instead of pixels, which are used in the PXD, is feasible since the background levels for the SVD at larger radii than the PXD are more manageable. This choice results in lower costs and fewer

readout channels. The average sensor hit efficiency exceeds 99%, with measured resolutions ranging from 9 to 25 μm , depending on the layer and side [75–80].

Central Drift Chamber (CDC)

The Central Drift Chamber (CDC) provides important information for three-dimensional tracking, momentum measurements, and energy-loss measurements to identify low-momentum tracks [58]. The CDC's inner and outer radii are 16 and 113 cm from the IP, respectively, covering the full angular acceptance of 17 to 150 degrees. It spans a length of 2.3 m with a diameter of 2.2 m. The CDC is a multi-wire proportional drift chamber featuring 14,336 tungsten sense wires, each with a diameter of 30 μm , surrounded by eight aluminum field wires with a diameter of 126 μm . To mitigate multiple scattering effects, the CDC is filled with a gas mixture comprising 50% helium and 50% ethane (C_2H_6).

When a charged particle traverses the gas, it ionizes the medium. The resulting electrons are guided by an electric field along the field lines. Near the wire, a strong electric field causes secondary ionization, leading to avalanche multiplication and inducing a charge at the drift chamber wire. The currents in the wires are directly proportional to the deposited energy, enabling the measurement of dE/dx , which is a crucial input for particle identification.

The CDC wires are organized to provide both x - y and z -axis information. Wires in the axial layers are aligned parallel to the beam axis (z -axis), allowing for timing measurements to determine the position in the x - y plane. To further obtain z -axis information, some wires are skewed in stereo layers with an angular offset ranging from -74 to 70 mrad. The sense wires are organized into alternating superlayers of axial and stereo layers, with the first superlayer comprising eight layers and subsequent superlayers consisting of six layers, totaling nine superlayers. The position resolution varies from 50 to 120 μm , depending on the layer and the angle of the incident particle [81].

Time-of-propagation (TOP) counter

The time-of-propagation (TOP) counter is part of the particle identification (PID) system in the barrel region. The TOP counter uses Cherenkov radiation to discriminate primarily between charged kaons and pions, as well as protons and deuterons [58]. It surrounds the CDC at a radius of 1.2 m from the IP and covers a polar angle acceptance of 32.2 to 128.7 degrees. The TOP counter consists of 16 quartz radiators, evenly distributed in ϕ . Each radiator is made of two quartz crystals, each 20 mm thick, 45 cm wide, and 1.25 m long, with an array of photomultiplier tubes at the end of each radiator.

Cherenkov radiation is produced when particles travel through a medium at a speed greater than the local speed of light. The particles emit photons at the Cherenkov angle, given by $\cos \theta_c = 1/\beta n$, where β is the ratio of the particle's momentum to its energy, and n is the refractive index of the medium. Total internal reflection within the quartz radiator preserves the Cherenkov angle, and the photomultiplier tubes detect the photons, allowing measurement of the (x, y) coordinates and the arrival time. By reconstructing the Cherenkov angle and obtaining β from momentum information, the TOP counter can determine the particle mass and distinguish between different particle types. The TOP counter achieves an 85% kaon identification efficiency with a 10% pion misidentification rate [82, 83]. Muons also leave detectable signals in this counter.

Aerogel ring-imaging Cherenkov (ARICH) counter

The proximity-focusing aerogel ring-imaging Cherenkov (ARICH) counter constitutes the second component of the PID system [58, 84]. Positioned in the forward endcap with an acceptance angle of 14.8 to 33.7 degrees, the ARICH features inner and outer radii of 0.5 and 1.14 m, respectively, at a distance of 1.7 m along the z -axis from the IP. Operating on the Cherenkov principle, the ARICH employs a medium of silica aerogel, consisting of two layers with a thickness of 20 mm each. The aerogel material is organized into square tiles, each with sides measuring 17 cm, arranged in four concentric rings. The two layers exhibit different refractive indices, specifically 1.045 and 1.055 [64].

The ARICH counter's operating principle is similar to that of the TOP counter. Cherenkov radiation is generated within both aerogel layers, and the Cherenkov cones are projected onto arrays of hybrid avalanche photodetectors after traversing the expansion volume. The refractive indices are selected to focus the photons effectively, reducing spread by overlapping the rings from the two layers at the photon detectors. By measuring the radius of the projected circle, the Cherenkov angle and, consequently, β can be determined. Different mass hypotheses can be compared when momentum information is taken into account. The ARICH counter exhibits a kaon identification efficiency of 90% with a 10% pion misidentification rate [85].

Electromagnetic calorimeter (ECL)

The electromagnetic calorimeter (ECL) is designed to measure the energy and angular coordinates of photons [58]. Accurate energy measurements for minimum-ionizing particles, such as muons, are challenging because they do not deposit their entire energy inside the ECL.

The ECL in the barrel consists of 6624 thallium-doped scintillating caesium-iodide (CSI(Tl)) crystals, each with a pair of attached photodiodes to detect visible light. The endcaps use pure CSI crystals, 2112 in total, which have a shorter scintillation decay time to handle the high pile-up in endcap regions. These crystals have lower light output, especially in the ultraviolet spectrum. Each crystal is equipped with an individual photodetector with internal gain. On average, the crystals have cross-section sizes of $6 \times 6 \text{ cm}^2$ and lengths of 30 cm. The barrel section, extending 3 m in length with an inner radius of 1.25 m, covers an angular acceptance of 32.2 to 128.7 degrees. The endcaps, positioned at 2 m in the forward direction and 1 m in the backward direction, cover an acceptance range of 12 to 155 degrees, with a one-degree gap between the barrel and endcap regions.

As particles traverse the detector, they generate electromagnetic or hadronic showers through processes like Bremsstrahlung and pair-production, leading to the production of photons and subsequent showers. The energy deposited is measured as scintillation light, enabling the recovery of the shower's energy. The ECL can detect photon energies ranging from 20 MeV to 4 GeV.

K_L and muon (KLM) detector

The K_L and muon (KLM) detector is designed to identify K_L and muons with momenta ranging from 0.6 GeV to 1.5 GeV. These particles are not stopped by previously traversed detectors, reaching the KLM detector where they leave energy deposits. While muons leave tracks in the CDC, K_L particles do not. The energy deposits allow detection of the particle's flight direction, but not its energy.

The KLM surrounds the superconducting magnet within an acceptance angle of 20 to 155 degrees. The detector comprises alternating layers of 14 iron plates, each with a thickness of 4.7 cm, and 14 (15

in the barrel) layers of active detector material [64]. The iron plates decelerate particles and function as the return yoke for the magnet.

Except for the first two layers, the barrel detector layers consist of glass-electrode resistive-plate chambers (RPCs). The RPCs are composed of two parallel glass sheets separated by a 1.9 mm gap filled with gas. Stacked in pairs, two RPCs form a module, with detector readout strips placed in orthogonal planes on either side of the RPC pair. The module has a total thickness of 31.6 mm, a length of 2.2 m, and a width ranging from 1.5 to 2.7 m. Charged particles ionize the gas in the RPCs. By applying an electric field, the electrons drift, cause secondary ionizations, and induce a signal in the readout strips.

The two innermost layers in the barrel and all detector layers in the endcap consist of scintillator strips embedded with wavelength-shifting fibers. This design, totaling 16,800 scintillator strips in the endcap, addresses the high background levels. Each layer has two orthogonal planes of scintillator strips. These scintillator strips, made of doped polystyrene, generate photons when a charged particle excites an electron into a higher energy level. These photons are guided through the wavelength-shifting fibers, which are connected to multipixel silicon photodiodes for signal readout.

Trigger and data acquisition system

The electron-positron bunches at the SuperKEKB accelerator cross every 8 ns, generating vast amounts of data that must be managed efficiently. A trigger system is employed to reduce the data volume by selectively focusing on and storing events of interest. Typically, these events include $\Upsilon(4S)$ production, continuum events, two-photon interactions, Bhabha scattering, and $\mu^+\mu^-$ production. In addition to selecting specific event types, the trigger system helps mitigate beam-induced backgrounds, as discussed in Section 3.2.

The trigger system operates in multiple stages, with the first being the Level 1 (L1) trigger [86]. This hardware-based stage relies on Field Programmable Gate Array (FPGA) boards and collects input from individual subdetectors such as the CDC and ECL. The CDC provides track-related information like momentum, position, and charge, while the ECL offers details on energy deposits and energy clusters. The TOP counter contributes timing and hit-topology information, while the KLM detector provides data on muon tracks. A global reconstruction logic component receives information from these detectors, performs logical calculations and low-level event reconstruction, handles pre-scaling, and triggers the L1 decision within 2-4 μs of a bunch-crossing event. If an event passes the L1 trigger, it proceeds to the data acquisition (DAQ) system.

The next stage is the software-based High-Level Trigger (HLT), which utilizes data from all subdetectors except for the PXD to perform event reconstruction. This process mirrors the offline reconstruction described in the following section and determines the relevance of the event. The HLT operates as a farm with five units equipped with a total of 1600 cores. During the initial physics runs, the L1 trigger achieved a maximum rate of 3.5 kHz [87]. The entire trigger system is designed to reach a maximum trigger rate of 30 kHz eventually. With each event averaging around 100 kB in size, the HLT will eventually process an influx of 3 GB of data per second [58, 88]. Since the PXD alone generates larger information chunks of 1 MB per event, its data is incorporated only after the HLT decision has been made, at which point the information is stored to disk. The trigger system achieves a total data reduction factor of 8 [87].

3.4 Particle reconstruction and identification

The process of particle reconstruction and identification using data collected by the Belle II detector is described in this section. The reconstruction process is carried out within the Belle II Analysis Software Framework (basf2) [89, 90], which features a modular structure that enables flexible and efficient data processing and analysis.

After data acquisition and storage by the DAQ system, offline software is used to reconstruct particles from the collected data. The raw data, which includes signals, hits, and energy deposits, is processed to extract physical objects, such as track candidates in the tracking system and energy clusters in the calorimeter. For details, see Section 3.4.1 and 3.4.2, respectively. Subsequently, a track-cluster matching is performed, described in Section 3.4.3, to refine the particle candidates and provide information for particle identification, discussed in Section 3.4.4.

3.4.1 Track and vertex reconstruction

The reconstruction of charged particles at Belle II begins with the detection of hits in the subdetectors. The algorithms and methods used for track and vertex reconstruction using these hits are described in this subsection.

Central to the track reconstruction process are track finding and track fitting, which enable the reconstruction of trajectories (tracks) originating from primary and secondary vertices. Subsequently, decay vertices are reconstructed, and particle interaction points are determined, laying the foundation for reconstructing complex decay trees. The steps of track finding, track fitting, and vertex reconstruction are discussed in more detail below.

Despite the use of advanced algorithms for both track and vertex reconstruction, challenges arise due to detector inefficiencies and additional hits from beam-induced backgrounds. The efficiency of the track reconstruction algorithms, defined as the ratio of correctly reconstructed tracks to the total number of charged particles produced in the collisions, is a key performance metric. The tracking efficiency, the precision in vertex reconstruction and their uncertainties directly impact the accuracy of the physics analyses that rely on the reconstructed decays.

Track finding

Track finding involves a pattern recognition process to determine which hits correspond to the same particle trajectory. The specific approach adopted within Belle II depends on the characteristics of the relevant subdetector. For example, the CDC and VXD require different algorithms to accommodate differences in track multiplicity, background levels, and proximity to the IP.

For the VXD, a cellular automaton approach [91] is used, where track segments connecting hits in adjacent layers constitute a cell. Cells that share hits are assessed for neighboring relationships, with track candidates comprising sets of these neighboring cells. In the CDC, both a local track finder, which employs a mechanism similar to the VXD's, and a global track finder, which investigates all hits simultaneously via the Legendre algorithm [91], are utilized. Combinatorial Kalman filters (CKFs) are employed to extend track candidates from the CDC to the SVD, and finally to the PXD, to obtain final track candidates with associated hits and rough estimates of the track parameters.

Track fitting

Following the identification of track candidates, track fitting employs a chosen track model to refine the estimation of track parameters. At Belle II, a helix track model is predominantly used [92, 93]. This model is chosen due to the near-homogeneous magnetic field environment, although it is imperfect due to detector interactions such as multiple scattering and energy loss. Simulating these interactions accurately and estimating the associated uncertainties is challenging, especially without knowing the correct mass hypothesis.

An iterative process refines the estimation of the track parameters and updates the track hypothesis. This process uses a deterministic annealing filter based on a CKF. While the CKF approximates a least squares method and considers material interactions, it cannot handle outliers. The deterministic annealing filter mitigates this by assigning greater weight to hits with significant residuals. This method determines the momentum, charge, and impact parameters d_r and d_z , which denote the closest distance to the IP in radial and z -coordinates upon track extrapolation towards the IP.

Vertex reconstruction

Vertex reconstruction builds upon the results of the track reconstruction process by using the identified track candidates to locate vertices, which are the points where particles interact or decay. This process involves initial vertex finding through pattern recognition, followed by vertex fitting to precisely determine the positions of vertices and refine track parameters at these interaction points. Accurate reconstruction of the vertices is necessary for mapping out the decay chains of particles, such as those of B mesons.

3.4.2 Cluster reconstruction

Details about the reconstruction of clusters are provided in this subsection. Cluster reconstruction relies on information from two subdetectors, the ECL and the KLM detector. The ECL identifies photons and measures their energies by reconstructing clusters from electromagnetic showers, while the KLM detector identifies hadrons and muons by reconstructing clusters based on energy deposits generated by particles passing through its layers. The methods for cluster reconstruction in these detectors depend on their respective detection capabilities and the physical principles governing particle interactions with their materials.

Clusters within the KLM detector are formed by energy deposits within a 5° opening angle of each other. Cluster reconstruction in the ECL primarily utilizes timing and energy information. The process begins with the selection of a seed crystal with an energy above a certain threshold representing a local energy maximum among its neighboring crystals. Next, all neighboring crystals with energy greater than 0.5 MeV are added to form a cluster. In instances where clusters share crystals, the energies are divided based on the ratio of the energy of each cluster to the sum of energies of the overlapping crystals. This method ensures a representative distribution of energy among adjacent clusters, accurately determining the photon energies.

3.4.3 Track-cluster matching

To identify the origin of clusters, they must be checked for compatibility with track candidates and other clusters. This subsection describes this process of track-cluster matching.

For the ECL clusters, each track candidate from the tracking system is extrapolated to the entrance of the ECL. If a track candidate matches an ECL cluster, it is assigned a charged-particle hypothesis. If there is no match, a neutral-particle hypothesis is assumed.

A similar process occurs for the KLM clusters, which can either originate from K_L particles or muons. Each track candidate is extrapolated to the entrance of the KLM detector, and a straight line is drawn between the IP and the KLM detector entrance point. If a cluster's center aligns with the line within a 15° angle, the cluster is assigned a muon hypothesis. The process is repeated to check for alignment between KLM and ECL clusters. If a cluster in the KLM detector aligns with an ECL cluster within 15° , they are associated with each other and assigned a K_L hypothesis. The remaining ECL clusters that have not been matched to a track or a KLM cluster are hypothesized to originate from purely electromagnetic showers.

3.4.4 Particle identification

During particle identification, information from all subdetectors is combined, the results from track-cluster matching are applied, and various particle hypotheses are tested. The subsection discusses the approaches to charged and neutral particle identification with the Belle II experiment.

Charged particle identification

For a track matched with an associated cluster, charged-particle hypotheses $\alpha \in \{e, \mu, \pi, K, \text{proton}(p), \text{deuteron}(d)\}$ are applied. For each hypothesis α and each relevant subdetector k , a likelihood $\mathcal{L}_{\alpha,k}$ is constructed. The total likelihood for a single charged-particle hypothesis is given by the product of individual likelihoods from all relevant subdetectors:

$$\mathcal{L}_\alpha = \prod_k \mathcal{L}_{\alpha,k} \quad (3.1)$$

Global PID likelihood ratios are then defined for each charged-particle hypothesis α as the ratio of its likelihood to the sum of the likelihoods for all considered hypotheses:

$$\text{PID}_\alpha = \frac{\mathcal{L}_\alpha}{\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_d} \quad (3.2)$$

The contribution of each subdetector to the likelihood varies depending on the particle hypothesis. The dE/dx energy-loss information provided by the CDC gives insights into the energy deposition patterns. Additionally, the TOP and ARICH counters contribute valuable particle identification information by analyzing the timing and Cherenkov radiation emitted by particles as they traverse these detectors. Furthermore, inputs from the ECL, the SVD, and the KLM detector enhance the PID process by offering measurements on particle energy, vertex positioning, and muon identification, respectively.

For the PID ratios of the pion, kaon, proton, and deuteron, the CDC, TOP, and ARICH counters contribute the most information. The dE/dx from the CDC is most important below 0.7 GeV, while the TOP and ARICH counter information becomes more significant above 0.7 GeV. For the pion PID ratio, the SVD also provides important input. For muon identification, track candidates are extrapolated into the KLM detector using the muon mass hypothesis. Nearby KLM clusters are added,

with the difference between predicted and measured ranges, along with the goodness of fit, used as inputs for the likelihood ratio, along with information from the CDC, TOP, and ARICH counters. For electron identification, the E/p (energy-to-momentum) ratio from the ECL is primarily used, with further contributions from dE/dx measurements from the SVD and CDC, and information from the TOP and ARICH counters.

Neutral particle identification

Identification of neutral particles, such as photons, neutral pions, and neutral kaons, is primarily based on information from the ECL and the KLM detector. Photons are identified from clusters in the ECL resulting from electromagnetic showers. The absence of an associated track in the tracking system, combined with the shower's shape and energy distribution in the ECL, helps distinguish photons from other particles. Neutral pions (π^0) are primarily identified through their decay into two photons. After identifying the photon candidates, the identification process involves reconstructing the invariant masses of photon pairs. Neutral pion candidates are selected based on mass windows consistent with the π^0 mass. The spatial and energy resolution of the ECL plays a key role in accurately reconstructing the photons and, by extension, the π^0 candidates.

Long-lived neutral kaons (K_L) are identified by their interactions in the KLM detector, where they produce characteristic shower patterns, and by the absence of matching hits in the tracking detectors. Neutral K_S particles are typically identified through their decay into two charged pions, which are reconstructed back to a common vertex. After reconstructing and identifying the pions, the invariant mass of the decay products and the characteristic flight length aid in confirming the K_S identity.

Collision and simulated datasets

This chapter describes the collision and simulated datasets used in the work discussed in this thesis. The measurement is based on collision data, described in Section 4.1, collected by the Belle II detector. Additionally, simulated datasets, referred to as Monte Carlo (MC) data and introduced in Section 4.2, are employed to estimate background levels, efficiencies, and model distributions. Both collision and simulated datasets, described below, are processed using the basf2 software framework [89, 90]. Corrections to both the simulated and collision datasets are performed based on the agreement between simulated and collision data. They are described in Section 4.3.

4.1 Collision data

This section discusses the properties of the collision data collected by the Belle II detector.

Between 2019 and 2022, leading up to the first long shutdown, the Belle II detector collected a dataset totaling an integrated luminosity of 424 fb^{-1} [11]. The bulk of this dataset originates from electron-positron collisions at a CM energy of $\sqrt{s} = 10.584 \text{ GeV}$, targeting the $\Upsilon(4S)$ resonance. The integrated luminosity of this on-resonance sample corresponds to $364 \pm 2 \text{ fb}^{-1}$.

An off-resonance dataset, collected at a CM energy 60 MeV below the $\Upsilon(4S)$ resonance energy, comprises an integrated luminosity of $42.3 \pm 0.3 \text{ fb}^{-1}$. Its primary function is to characterize backgrounds, as discussed in Section 3.2. These predominantly consist of $q\bar{q}$ production events ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $c\bar{c}$), other QED processes such as $\tau^+\tau^-$ production, and two-photon processes resulting in final states like $e^+e^-\ell^+\ell^-$. The off-resonance dataset also aids in estimating the number of $B\bar{B}$ pairs present in the on-resonance data. A study by the Belle II performance group estimates $N_{B\bar{B}} = (387 \pm 6)$ million events of $\Upsilon(4S)$ production with subsequent $B\bar{B}$ pair production.

4.2 Simulated data

To identify background-discriminating processes and determine efficiencies, all relevant physics processes are simulated, producing what are known as Monte Carlo (MC) samples. The generation of the MC samples is a two-step process. Here, the simulated datasets used for comparison and analysis are described.

First, the physics processes are simulated using event generators, followed by the simulation of

the detector response. The EvtGen event generator software package [94] generates generic MC samples for both charged and neutral $B\bar{B}$ events, where both B mesons decay into semileptonic and hadronic B decays. EvtGen simulates cascade decays controlled by a decay table, listing the decays with respective branching ratios and decay models, with PYTHIA [95] used for inclusive final states. In contrast to the generic samples, signal MC samples simulate specific final states (such as $B \rightarrow X_c \ell \nu_\ell$ or $B \rightarrow X_u \ell \nu_\ell$) for one B meson, while the other B meson decays generically. For continuum processes, quark-antiquark ($q\bar{q}$) and tau-lepton pair ($\tau\tau$) samples are produced using PYTHIA and KKMC [96], respectively. Two-photon production events are simulated using the AAFH software package [97]. Hadronization and tau-lepton decays are simulated with PYTHIA and TAUOLA [98], respectively.

For all samples, final-state radiation of photons from stable charged particles is simulated using the PHOTOS [99] package. After simulating the physics processes, simulated beam-induced backgrounds are added to all generated events [65]. The Geant4 [100] software package is used to account for all interactions with the material upon propagation of the particles through the detector.

In this work, the generic $B\bar{B}$ samples are used to model all B meson decays, except for $B \rightarrow X_u \ell \nu_\ell$ decays. For $B \rightarrow X_u \ell \nu_\ell$ decays, separate signal MC samples, containing 50 million events each of resonant and nonresonant $B^0 \rightarrow X_u \ell \nu_\ell$, and $B^+ \rightarrow X_u \ell \nu_\ell$ events are simulated. Additionally, 10 million events each of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays are simulated separately. More detailed information about the models used in the simulation of $B \rightarrow X_c \ell \nu_\ell$ and $B \rightarrow X_u \ell \nu_\ell$ decays is provided in the following subsections. This includes descriptions of branching fractions (Section 4.2.1), form-factor models (Section 4.2.2), and for $B \rightarrow X_u \ell \nu_\ell$ decays more specifically, the lineshape of the ρ meson (Section 4.2.3), and the modeling of the nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ (Section 4.2.4) and total nonresonant (Section 4.2.5) components.

4.2.1 Branching fractions

This subsection provides an overview of the branching fractions used in the simulation. For resonant $B \rightarrow X_c \ell \nu_\ell$ decays, world average branching fractions are employed as reported by the Particle Data Group (PDG) [37], coupled with an assumption of isospin symmetry following the methodology in Ref. [101]. The branching fractions of these resonant $B \rightarrow X_c \ell \nu_\ell$ decays, with $X_c \in \{D, D^*, D_1, D_0^*, D_1', D_2^*\}$, are shown in Table 4.1.

For both charged and neutral B meson decays, the sums of the branching fractions of the resonant decays do not agree with the total inclusively measured $B \rightarrow X_c \ell \nu_\ell$ branching fractions, as shown in Table 4.1. Following the procedure in Ref. [102], the remaining difference is filled by nonresonant decays. Some of the nonresonant modes have been measured directly, such as $B \rightarrow D^{(*)} \pi \pi \ell \nu$ (or additionally $B \rightarrow D_s^{(*)} K \ell \nu$ for charged B mesons). After subtracting the resonant contributions to these modes from the measured branching fractions, they are assigned to the corresponding simulated nonresonant decays shown in Table 4.1. At this point, there is still a “gap” between the sum of the resonant and nonresonant decays and the total inclusively measured $B \rightarrow X_c \ell \nu_\ell$ branching fraction for both charged and neutral B meson decays, accounting for approximately 4% of the $B \rightarrow X_c \ell \nu_\ell$ decays. In this work, this gap is filled with $B \rightarrow D^{(*)} \eta \ell \nu_\ell$ decays. Since this is just a motivated guess following Ref. [102], an uncertainty of 100% is assigned to the branching fractions of these gap decays. The branching fractions assigned to the gap decays are presented in Table 4.1.

To model resonant $B \rightarrow X_u \ell \nu_\ell$ decays, where $X_u \in \{\pi, \rho, \omega, \eta, \eta'\}$ (or for neutral B decays $X_u \in \{\pi, \rho\}$), world average $B \rightarrow X_u \ell \nu_\ell$ branching fractions [37] are used. The sums of these resonant branching fractions approximately account for 20% of the total inclusively measured

Table 4.1: Branching fractions of the $B \rightarrow X_c \ell \nu_\ell$ decays used in simulation. Values are derived from Ref. [37] following the methodology in Ref. [101].

	Decay	Unit	B^\pm	B^0
Resonant	$B \rightarrow D \ell \nu$	10^{-2}	2.41 ± 0.07	2.24 ± 0.07
	$B \rightarrow D^* \ell \nu$	10^{-2}	5.50 ± 0.11	5.11 ± 0.11
	$B \rightarrow D_1 \ell \nu$	10^{-3}	6.63 ± 1.09	6.16 ± 1.01
	$B \rightarrow D_0^* \ell \nu$	10^{-3}	4.20 ± 0.75	3.90 ± 0.70
	$B \rightarrow D_1' \ell \nu$	10^{-3}	4.20 ± 0.90	3.90 ± 0.84
	$B \rightarrow D_2^* \ell \nu$	10^{-3}	2.93 ± 0.32	2.73 ± 0.30
Nonresonant	$B \rightarrow D \pi \pi \ell \nu$	10^{-3}	0.26 ± 0.89	0.58 ± 0.82
	$B \rightarrow D^* \pi \pi \ell \nu$	10^{-3}	2.16 ± 1.02	2.01 ± 0.95
	$B \rightarrow D_s K \ell \nu$	10^{-3}	0.30 ± 0.14	
	$B \rightarrow D_s^* K \ell \nu$	10^{-3}	0.29 ± 0.19	
Gap	$B \rightarrow D \eta \ell \nu$	10^{-3}	3.77 ± 3.77	4.09 ± 4.09
	$B \rightarrow D^* \eta \ell \nu$	10^{-3}	3.77 ± 3.77	4.09 ± 4.09
Total	$B \rightarrow X_c \ell \nu$	10^{-2}	10.8 ± 0.4	10.1 ± 0.4

$B \rightarrow X_u \ell \nu_\ell$ branching fractions for both charged and neutral B meson decays. The remaining difference is saturated by nonresonant $B \rightarrow X_u \ell \nu_\ell$ decays, in which the X_u state hadronizes into multiple hadrons. The branching fractions of the $B \rightarrow X_u \ell \nu_\ell$ decays used in the simulation are shown in Table 4.2.

Table 4.2: Branching fractions of the $B \rightarrow X_u \ell \nu_\ell$ decays used in simulation. Values are provided in Ref. [37].

	Decay	Unit	B^\pm	B^0
Resonant	$B \rightarrow \pi \ell \nu$	10^{-4}	0.78 ± 0.03	1.50 ± 0.06
	$B \rightarrow \rho \ell \nu$	10^{-4}	1.58 ± 0.11	2.94 ± 0.21
	$B \rightarrow \omega \ell \nu$	10^{-4}	1.19 ± 0.09	
	$B \rightarrow \eta \ell \nu$	10^{-4}	0.39 ± 0.05	
	$B \rightarrow \eta' \ell \nu$	10^{-4}	0.23 ± 0.08	
Nonresonant	$B \rightarrow X_u \ell \nu$	10^{-3}	1.34 ± 0.22	1.48 ± 0.24
Total	$B \rightarrow X_u \ell \nu$	10^{-3}	1.76 ± 0.22	1.92 ± 0.24

4.2.2 Form-factor models

To describe the decay rates of the semileptonic $B \rightarrow X \ell \nu_\ell$ decays as functions of q^2 , form-factor parameterizations introduced in Section 2.2.2 are used. This subsection discusses the form-factor models implemented in the simulation to represent particle interactions accurately.

The BGL [49] parameterization given in Equation 2.18 is chosen to describe the form factors of

$B \rightarrow D\ell\nu_\ell$ and $B \rightarrow D^*\ell\nu_\ell$ decays. The form-factor expansion coefficients used in the simulation are provided in Table 4.3. For $B \rightarrow D\ell\nu_\ell$ decays, the central values of the expansion coefficients to order $k = 3$ for $f_+(q^2)$ and $f_0(q^2)$ provided by Ref. [103] are implemented. Central values of the expansion coefficients to order $k = 1$ for the $\mathcal{F}_1(q^2)$, $f(q^2)$, and $g(q^2)$ form factors given in Ref. [104] are used in the simulation of $B \rightarrow D^*\ell\nu_\ell$ decays. Table 4.3 provides these values in a modified form: $\tilde{b}_k = \eta_{\text{EW}}|V_{cb}|b_k$, where $\eta_{\text{EW}} = 1.0066$ [105] is an electroweak correction.

Table 4.3: Central values and uncertainties of the form-factor expansion coefficients for $B \rightarrow D\ell\nu_\ell$, $B \rightarrow \pi\ell\nu_\ell$ and $B \rightarrow D^*\ell\nu_\ell$ decays used in simulation. The BGL parameterization is used to describe $B \rightarrow D\ell\nu_\ell$ and $B \rightarrow D^*\ell\nu_\ell$ decays, while the BCL parameterization is used in the description of $B \rightarrow \pi\ell\nu_\ell$ decays.

Coefficient		$B \rightarrow D\ell\nu_\ell$	$B \rightarrow \pi\ell\nu_\ell$
$f_+(q^2)$	b_0^+	0.0126 ± 0.0001	0.42 ± 0.02
	b_1^+	-0.094 ± 0.003	-0.51 ± 0.09
	b_2^+	0.34 ± 0.04	-0.75 ± 0.34
	b_3^+	-0.1 ± 0.6	
$f_0(q^2)$	b_0^0	0.0115 ± 0.0001	0.56 ± 0.02
	b_1^0	-0.057 ± 0.002	-1.42 ± 0.11
	b_2^0	0.12 ± 0.04	
	b_3^0	0.4 ± 0.7	

Coefficient $\times 10^3$		$B \rightarrow D^*\ell\nu_\ell$
$\mathcal{F}_1(q^2)$	$\tilde{b}_0^{\mathcal{F}_1}$	0.30 ± 0.10
	$\tilde{b}_1^{\mathcal{F}_1}$	-3.68 ± 1.74
$f(q^2)$	\tilde{b}_0^f	0.51 ± 0.01
	\tilde{b}_1^f	0.67 ± 0.34
$g(q^2)$	\tilde{b}_0^g	1.00 ± 0.03
	\tilde{b}_1^g	-2.35 ± 0.90

In this work, $B \rightarrow \pi\ell\nu_\ell$ decays are described using the BCL [12] parameterization given in Equation 2.19. In the simulation, central values for the expansion coefficients b_k from Ref. [14], consisting of results from Refs. [15–17], are implemented and shown in Table 4.3. The expansion orders are $k = 2$ and $k = 1$ for the $f_+(q^2)$ and $f_0(q^2)$ form factors, respectively. The form factors of $B \rightarrow \rho\ell\nu_\ell$ and $B \rightarrow \omega\ell\nu_\ell$ decays are described using the BSZ [13] parameterization given in Equation 2.20. The central values for b_k used in the simulation are provided by Ref. [10], which extracted the BSZ expansion coefficients from averaged $B \rightarrow \rho\ell\nu_\ell$ and $B \rightarrow \omega\ell\nu_\ell$ measurements. The expansion order is $k = 2$ for the $A_1(q^2)$, $A_{12}(q^2)$, and $V(q^2)$ form factors, and it is $k = 1$ for the $A_0(q^2)$ form factor. The implemented central values are shown in Table 4.4.

To describe the form factors of the $B \rightarrow \eta^{(\prime)}\ell\nu_\ell$ decays, the BK [48] parameterization given in Equation 2.21 is used. Due to the mixing of the η and η' states, additional nontrivial modifications would be necessary to implement a valid expansion-based parameterization. Results from LCSR calculations in Ref. [106] are used as central values of the free parameters in the simulation and are provided in Table 4.4.

4.2.3 Lineshape of the ρ meson

In contrast to the relatively long lifetime of 26 ns of the pion, the ρ meson decays rapidly, with a lifetime of only 4×10^{-15} ns [37]. This results in a significantly larger natural width of the ρ meson compared to the pion, whose width can be considered negligible in most cases. Since the lineshape of a particle is a defining property of the invariant mass spectrum of the particle's decay products, often used to reconstruct the particle, careful consideration is required in modeling the lineshape. This subsection provides a detailed description of the lineshape modeling of the ρ meson within the simulation, including any assumptions and parameters.

Table 4.4: Central values and uncertainties of the BSZ form-factor expansion coefficients for $B \rightarrow \rho \ell \nu_\ell$ and $B \rightarrow \omega \ell \nu_\ell$ decays, and of the free parameters in the BK parameterization for $B \rightarrow \eta^{(\prime)} \ell \nu_\ell$ decays, used in simulation. For $B \rightarrow \eta^{(\prime)} \ell \nu_\ell$ decays, the larger of the two uncertainties is used where Ref. [106] quotes asymmetric uncertainties.

Coefficient		$B \rightarrow \rho \ell \nu_\ell$	$B \rightarrow \omega \ell \nu_\ell$
$A_0(q^2)$	$b_0^{A_0}$	0.86 ± 0.19	0.94 ± 0.28
	$b_1^{A_0}$	1.43 ± 1.02	1.78 ± 1.20
$A_1(q^2)$	$b_0^{A_1}$	0.26 ± 0.03	0.24 ± 0.03
	$b_1^{A_1}$	0.38 ± 0.13	0.30 ± 0.22
	$b_2^{A_1}$	0.16 ± 0.41	0.00 ± 0.55
$A_{12}(q^2)$	$b_0^{A_{12}}$	0.29 ± 0.03	0.25 ± 0.04
	$b_1^{A_{12}}$	0.72 ± 0.17	0.54 ± 0.24
	$b_2^{A_{12}}$	0.37 ± 0.70	-0.03 ± 0.96
$V(q^2)$	b_0^V	0.33 ± 0.03	0.31 ± 0.04
	b_1^V	-0.87 ± 0.18	0.89 ± 0.27
	b_2^V	1.88 ± 0.94	1.81 ± 1.19

Parameter	$B \rightarrow \eta \ell \nu_\ell$	$B \rightarrow \eta' \ell \nu_\ell$
$f_+(0)$	0.168 ± 0.047	0.130 ± 0.036
α_{BK}	0.46 ± 0.06	0.45 ± 0.17

The dynamic amplitude A describing the lineshape $|A^2|$ of a particle can be parameterized as a function of the energy \sqrt{s} using a fixed-width relativistic Breit-Wigner of the form:

$$A(\sqrt{s}) = \frac{1}{M_R^2 - s + iM_R\Gamma_R}, \quad (4.1)$$

where M_R and Γ_R define the nominal mass and width of the particle. For the decay of a ρ meson into two pions, $\sqrt{s} = m_{\pi\pi}$, corresponding to the di-pion invariant mass. Often the nominal mass and width provided by the PDG [37], $M_R = 775.26 \pm 0.23$ MeV and $\Gamma_R = 149.1 \pm 0.8$ MeV, are used to describe the lineshape of the ρ meson. The $m_{\pi\pi}$ distribution of $\rho \rightarrow \pi\pi$ in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays resulting from substituting the PDG values into Equation 4.1 is given as the grey histogram in Figure 4.1.

It is, however, not always appropriate to substitute these values without further consideration, as they depend on the chosen parameterization of the ρ amplitude. Incorrectly implementing these values may alter the pole position of the lineshape, which is an unintended feature since the pole position is an intrinsic property of the resonance. To account for this effect, in this work, the pole-conserving values of M_R and Γ_R reported in Ref. [107] are implemented: $M_R = 0.7602 \pm 0.0017$ MeV and $\Gamma_R = 0.1471 \pm 0.002$ MeV. In Ref. [107], these values were determined from a conversion of the pole position $\sqrt{s_0} = (0.763 - 0.073i)$ reported in Ref. [108]. The $m_{\pi\pi}$ spectrum of $\rho \rightarrow \pi\pi$ in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays resulting from substituting these values into Equation 4.1 is shown as the light-blue histogram in Figure 4.1. The shift in the position of the peak between the two options is visible.

Other effects on the lineshape related to possible interference between particles also need to be considered. For example, in Ref. [107], the level of interference between the ρ and ω mesons in semileptonic decays is examined using the partial branching-fraction spectrum of $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ as a function of the $m_{\pi\pi}$ invariant mass measured by Ref. [109]. The measured spectrum within the range

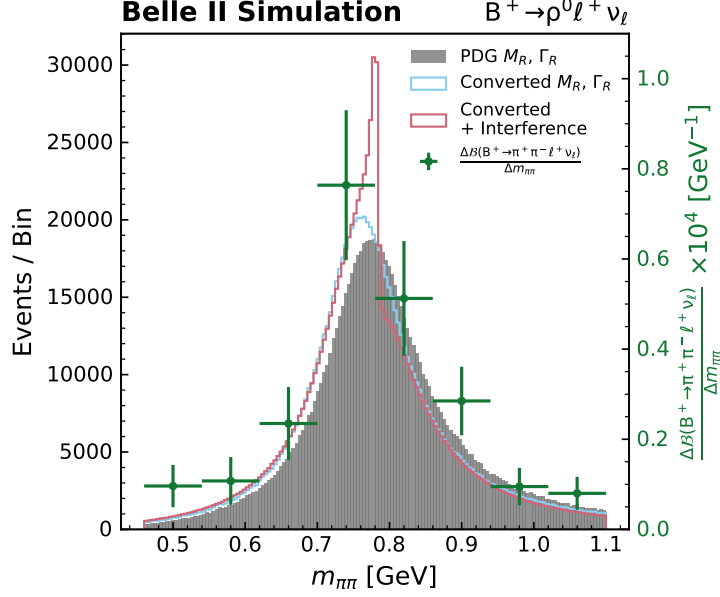


Figure 4.1: Invariant mass $m_{\pi\pi}$ spectrum of $\rho \rightarrow \pi\pi$ in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays using the mass and width of the ρ meson: 1) directly from the PDG, 2) converted from the pole position [108], or 3) using the lineshape of the ρ meson including an estimation of the interference between the ρ and ω mesons [107]. The partial branching-fraction spectrum of $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ as a function of $m_{\pi\pi}$ measured by Ref. [109] is shown on a second axis.

of 0.46 to 1.10 GeV is shown in green in Figure 4.1. Ref. [107] performed an amplitude fit incorporating an additional ρ - ω meson interference term to the measured $m_{\pi\pi}$ spectrum of $B \rightarrow \pi\pi \ell \nu_\ell$ decays. The $m_{\pi\pi}$ spectrum of $\rho \rightarrow \pi\pi$ in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays obtained from the fit results is shown as the red histogram in Figure 4.1. We choose to model the ρ meson lineshape neglecting the contributions from interference, but include potential interference effects as a systematic uncertainty discussed in Section 8.3.9.

4.2.4 Nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ spectrum

A significant component of the nonresonant $B \rightarrow X_u \ell \nu_\ell$ decays consists of events where the X_u system corresponds to a pair of charged pions. Accurately representing the abundance of these decays in simulation is crucial for overall background estimation, and a correct prediction of the shape of the $m_{\pi\pi}$ spectrum is necessary whenever a ρ meson is reconstructed from two charged pions. This subsection discusses the modeling of the nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ component, including the methodologies and results.

In this thesis, results from Ref. [109] are used to address both concerns in a data-driven way. In Ref. [109], the partial branching-fraction spectrum of $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ decays as a function of $m_{\pi\pi}$ has been measured. Since the measured spectrum includes contributions from resonant $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, it cannot be used directly to describe the nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ decays in the simulation. While a naive subtraction of the world average total branching fraction of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ from the

measured total branching fraction could provide a reasonable description of the total abundance of the nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ decays, it would not lead to an accurate representation of the $m_{\pi\pi}$ spectrum of these decays.

Instead, in this work, a region of the measured spectrum contaminated by $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays is defined. This corresponds to the region in $m_{\pi\pi}$ located at the ρ meson mass. We combine all measurement points from Ref. [109] reaching into the $[m_R - \Gamma_R, m_R + \Gamma_R]$ region, with m_R and Γ_R corresponding to the PDG mass and width of the ρ meson, respectively. This results in a combined bin with edges: $[0.46, 1.10]$ GeV. The individual measurements within this region are shown in Figure 4.1. In Figure 4.2, the combination of the measurements within this region is shown as the black data point within the grey band. The remaining bins shown in Figure 4.2 correspond to the ones in Ref. [109]. The bin edges are defined as: $m_{\pi\pi} = [0.28, 0.46, 1.10, 1.18, 1.26, 1.34, 2.00]$ GeV.

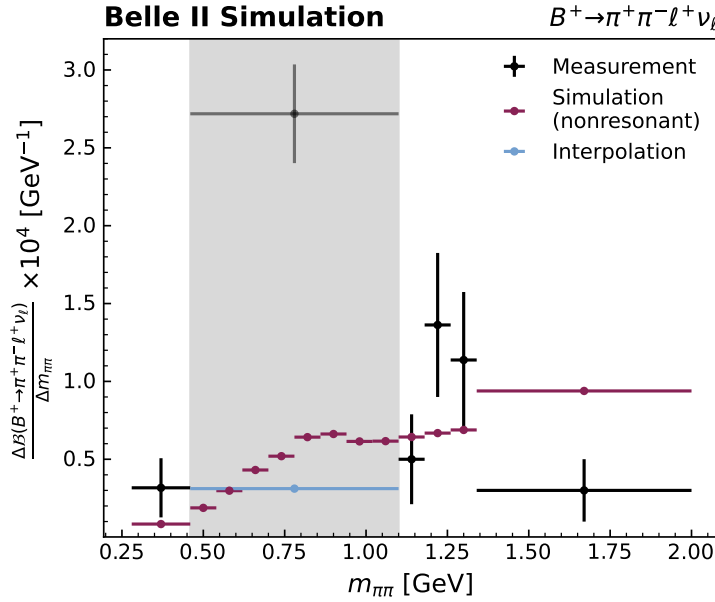


Figure 4.2: The partial branching-fraction spectrum of $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ measured as a function of the $m_{\pi\pi}$ invariant mass in Ref. [109]. The simulated nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ spectrum, and the result of the interpolation of the spectrum into the ρ meson mass window (shown in grey) is also shown.

The uncorrected $m_{\pi\pi}$ spectrum of nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ decays in simulation, shown in purple in Figure 4.2, and the measured spectrum do not agree well. One reason is that a relatively simple phase-space model is used to simulate the nonresonant decays with EvtGen [94], corresponding only to a physically-motivated estimate. Additionally, the simulation does not account for additional resonances, such as the $f_2(1270)$, whose mass peak is observed in the measured spectrum in the region surrounding 1.27 GeV.

To correct the simulated spectrum, first, the partial branching fractions measured by Ref. [109] are assigned to the simulated nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ decays in all bins except for the bin corresponding to the ρ meson mass window marked in grey in Figure 4.2. Due to the dominance of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events in this bin, the partial branching fraction in this bin is estimated by interpolating the measured partial branching fractions from the first and last bins.

However, the appropriate functional form of the interpolation depends on the behavior of the nonresonant decays as a function of $m_{\pi\pi}$. A simple model of a straight line with two free parameters is selected, under the assumption that nonresonant $B \rightarrow X_u \ell \nu_\ell$ events exhibit no resonant structure as a function of $m_{\pi\pi}$. This is in agreement with the findings of Ref. [107], showing that any of the proposed models can be represented by a sloped straight line as a function of $m_{\pi\pi}$ within the range of the ρ meson mass bin. The interpolated point is shown in light-blue in Figure 4.2. This interpolated partial branching fraction is assigned to all $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ events falling into this bin.

The $m_{\pi\pi}$ spectrum of the nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ events in simulation thus corresponds to the combination of the black and blue points shown in Figure 4.2. In this way, both the normalization and the shape in $m_{\pi\pi}$ are assigned in a data-driven way. The estimation of the uncertainties related to this procedure is described in Section 8.3.6.

4.2.5 Hybrid model implementation

This subsection provides details on the modeling of the total nonresonant $B \rightarrow X_u \ell \nu_\ell$ component in simulation. The total nonresonant $B \rightarrow X_u \ell \nu_\ell$ decays is described using the DFN [47] model, as detailed in Section 2.2.1. This model predicts the triple-differential rate in X_u particle mass m_X , the B meson rest-frame lepton energy E_ℓ^B , and q^2 , combined with a nonperturbative shape function using an exponential model. In the simulation, the central values for the two relevant parameters provided in Ref. [110] are used:

$$\bar{\lambda} = 0.621 \pm 0.041 \text{ GeV} \quad \text{and} \quad \lambda_1 = -0.497^{+0.072}_{-0.086} \text{ GeV}^2.$$

The distributions of m_X , E_ℓ^B , and q^2 for the nonresonant $B \rightarrow X_u \ell \nu_\ell$ decays are shown in Figure 4.3.

To accurately describe the total $B \rightarrow X_u \ell \nu_\ell$ sample composition as a function of m_X , E_ℓ^B , and q^2 , a hybrid model [111] approach is implemented, closely following the methodology in Ref. [112]. This approach combines the resonant and nonresonant decay rates in bins of m_X , E_ℓ^B , and q^2 to recover the inclusively measured rates. Weights w_i are assigned to the nonresonant events, such that the sample size in each bin i is $H_i = R_i + w_i I_i$, where R_i is the number of resonant events in the bin, and I_i is the number of nonresonant events in the bin. The binning scheme is as follows:

$$\begin{aligned} m_X &= [0, 1.4, 1.6, 1.8, 2, 2.5, 3, 3.5] \text{ GeV}, \\ E_\ell^B &= [0, 0.5, 1, 1.25, 1.5, 1.75, 2, 2.25, 3] \text{ GeV}, \\ q^2 &= [0, 2.5, 5, 7.5, 10, 12.5, 15, 25] \text{ GeV}^2, \end{aligned}$$

The resulting distributions of m_X , E_ℓ^B , and q^2 for the total $B^+ \rightarrow X_u \ell \nu_\ell$ samples, separated into resonant and nonresonant contributions, are shown in Figure 4.3. The effect of implementing the hybrid model on the distributions of the nonresonant samples is most prominently observed in the m_X distribution.

4.3 Corrections based on data-MC agreement

To accurately estimate background contributions and efficiencies from simulation, the simulated (MC) data needs to represent the collision data. Whenever deviations between the two are observed,

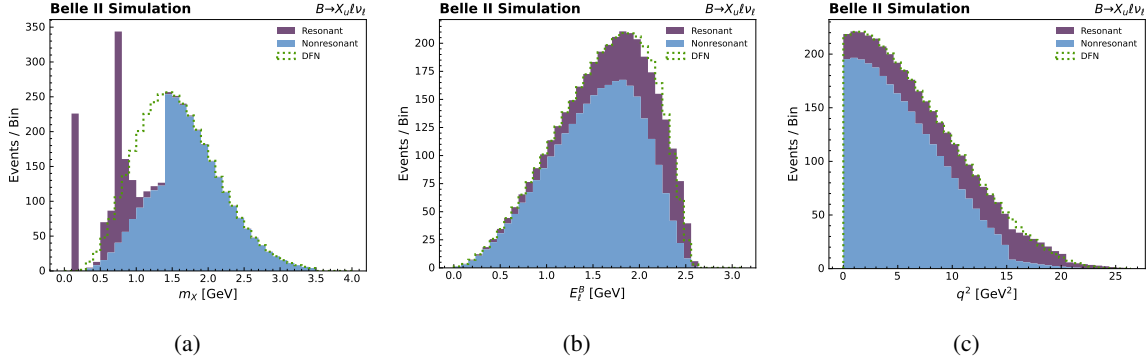


Figure 4.3: Distributions of (a) X_u particle mass m_X , (b) the B meson rest-frame lepton energy E_ℓ^B , and (c) q^2 in $B \rightarrow X_u \ell \nu_\ell$ decays. The nonresonant component is modeled using the DFN model, shown before and after implementing the hybrid model. The resonant component is also shown.

corrections are applied to efficiencies, fake rates, and overall event rates. The following subsections describe the corrections for particle identification (PID) performance (Section 4.3.1), track momentum and photon energy biases (Section 4.3.2), and differences in event rates between off- and on-resonance data (Section 4.3.3).

4.3.1 Particle identification corrections

Since the PID ratios, which were introduced in Section 3.4.4, are not perfectly modeled in the simulation, selections based on these ratios can introduce biases in the selection efficiencies of events between collision and simulated data. This subsection describes the dedicated corrections for PID discrepancies between collision data and simulation.

Identification efficiencies and misidentification probabilities for charged pions, kaons, electrons, and muons are determined separately for each particle and its charge in bins of momentum and polar angle using data samples from physics control modes. Correction tables for these efficiencies and misidentification rates are provided by the Belle II performance group in bins of particle momentum and polar angle for various selections on the PID ratio. These tables are then used to correct the simulated data.

For lepton PID ratios, a study of the agreement between the efficiencies in collision and simulated data has been conducted using $J/\psi \rightarrow \ell^+ \ell^-$, $e^+ e^- \rightarrow \ell^+ \ell^- (\gamma)$, and $e^+ e^- \rightarrow (e^+ e^-) \ell^+ \ell^-$ modes. Similarly, the agreement in lepton misidentification rates has been investigated in $K_s^0 \rightarrow \pi^+ \pi^-$ and $e^+ e^- \rightarrow \tau(1p)\tau(3p)$ modes. For pion and kaon PID ratios, studies of $D^* \rightarrow D^0 [K\pi]\pi$, $K_s^0 \rightarrow \pi^+ \pi^-$, and $\Lambda^0 \rightarrow p\pi^-$ have been performed to determine corrections for efficiency and misidentification rates.

The regions in momentum and polar angle not covered by the correction tables are negligible, accounting for less than 0.01% of the simulated events. Of the covered events, the average efficiency correction for the electron PID ratio is 0.96, with no corrections larger than 1.17. For the muon PID ratio, the average correction is 0.92, with maximum corrections of 1.26.

4.3.2 Track momentum and photon energy bias

This subsection describes the adjustments made to correct biases in track momentum and photon energy measurements.

In a study of $D^* \rightarrow D^0 [K^- \pi^+] \pi^+$ decays, it was observed that the D meson invariant mass distribution is shifted in collision data compared to the PDG value present in the simulation. This discrepancy originates from a minor inaccuracy in the magnetic field map used in reconstructing the collision data, leading to a bias in track momentum. This effect is corrected by scaling the momentum of each track using a scaling factor of 0.99987.

In studies of π^0 and η decays into two photons, it was also discovered that there is a finite bias in the energy and resolution of the low-energy photon spectrum in collision data. This effect is corrected by scaling the energy of each cluster not matched to a track.

4.3.3 Continuum corrections

Before comparing the simulated continuum to the off-resonance data, normalization differences in event rates due to different luminosities and CM energies are accounted for. These corrections are described in the following.

First, the off-resonance sample is scaled to match the luminosity of the on-resonance dataset. Additionally, a correction factor is needed to account for $1/s$ dependence of the continuum production cross-section. This means fewer continuum events occur at the on-resonance energy than predicted by the luminosity-scaled off-resonance sample. The total scaling factor for the off-resonance sample is given by:

$$\frac{\mathcal{L}_{\text{on}}}{\mathcal{L}_{\text{off}}} \times \frac{E_{\text{off}}^2}{E_{\text{on}}^2},$$

where \mathcal{L}_{on} and \mathcal{L}_{off} are the luminosities of the on- and off-resonance data samples, respectively, and the average CM energies of the on- and off-resonance data samples are given by E_{on} and E_{off} .

Event reconstruction and selection

In Chapter 3, the methods used to reconstruct particles at the Belle II experiment were described. This chapter provides details about the specifics of event reconstruction from particle candidates.

First, the tagged and untagged approach to event reconstruction are introduced in Section 5.1. Next, Section 5.2 described the reconstruction of the signal B meson and signal selection steps. For decays involving a final-state neutrino, such as $B \rightarrow \pi \ell \nu_\ell$ and $B \rightarrow \rho \ell \nu_\ell$, the reconstruction process includes estimating the so-called missing momentum attributed to the neutrino. This process is described in Section 5.3.

Event reconstruction involves selection steps designed to enhance signal purity while maintaining high efficiency. Two key metrics used to evaluate reconstruction and selection performance are purity and efficiency. The total reconstruction or selection efficiency is defined as the ratio of the number of events after reconstruction or selection to the initial number of events. Additionally, a step-wise efficiency is defined as the ratio of the number of events after a particular selection step to the number of events before that step. Purity is defined as the ratio of signal events to the total number of events.

During the reconstruction process, multiple event candidates may arise, each representing a different combination of particle candidates. The number of candidates per event is referred to as the multiplicity. In Section 5.4 and 5.5, the categorization of simulated events and the quantities used for signal extraction are described, respectively.

5.1 Event reconstruction

This section describes two approaches to event reconstruction at the Belle II experiment. In cases where a B meson decays hadronically, the B meson can be fully reconstructed by sequentially reconstructing its daughter particles from the lowest to highest levels. However, when the B meson decay includes a neutrino in the final state, as in semileptonic decays, or in inclusive B meson decays with multiple final states, kinematic constraints are only provided by the experimental setup. As discussed in Chapter 3, the CM energy of the SuperKEKB accelerator is slightly above the threshold needed to produce two B mesons. Knowing the energy of the $\Upsilon(4S)$, one B meson can be kinematically constrained, if information about the kinematics of the partner B meson is available. There are two main approaches for this, which are detailed below.

Tagged approach: In the tagged approach, first one B meson is reconstructed, known as the tag side, and then reconstruct the signal B meson using the remaining tracks and clusters in the event. The tag-side reconstruction involves many channels with various selection stages, and Belle II employs a dedicated algorithm called Full Event Interpretation [113]. The strong kinematic constraints from the reconstructed tag side result in high signal purity; however, this approach also reduces signal reconstruction efficiency due to the limited number of tag-side channels.

Untagged approach: Alternatively, in the untagged or inclusive-tag approach, the tag side is not explicitly reconstructed and the focus lies solely on reconstructing the signal B meson. The sum of all tracks and clusters not assigned to the signal B meson constitutes the rest of the event (ROE). While the untagged approach lacks direct constraints from the second B meson, kinematic constraints implied by the ROE can still be utilized through some assumptions. Although this approach results in lower signal purity compared to the tagged approach, it offers higher signal reconstruction efficiency.

The high signal reconstruction efficiency of the untagged approach is particularly desirable in the early stages of a particle physics experiment, where the tagged approach provides statistically limited results. The current size of the Belle II dataset is one of the reasons the untagged approach is chosen for analyzing $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays. In this approach, the signal B meson is reconstructed from a lepton and a hadron (π or ρ) candidate, with neutrino properties inferred from the missing momentum, which is determined from the difference between the known initial state and the observed final state.

5.2 Signal B meson reconstruction and selection

The reconstruction of signal B mesons begins by selecting possible track candidates for the lepton and pion. The preselection of signal event candidates is detailed in the Section 5.2.1. Subsequent selection steps aim to reject backgrounds and enhance purity while maintaining high signal selection efficiency. These are described in Section 5.2.2.

5.2.1 Reconstruction and preselection of signal events

This section discusses the reconstruction of signal events, including initial signal selection steps.

Initially, track candidates are required to have polar angles within the acceptance of the CDC and transverse momenta greater than 0.05 GeV. This requirement filters out low-quality tracks that are partially outside the full detector acceptance. At this stage, the fraction of generated signal events correctly reconstructed in simulation is 83% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 72% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. Detector and tracking inefficiencies, as well as acceptance losses, contribute to the reduction in signal reconstruction efficiency.

The sample purity is very low, at 3.1×10^{-7} and 0.7×10^{-7} signal events per background event. This is due to the large cross-sections of continuum and low-multiplicity processes. Additionally, event multiplicities are very high, with approximately 132 candidates for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 700 candidates for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ per simulated event. For continuum events, the numbers are about a factor of three lower, and low-multiplicity events yield around five candidates per event. To improve signal reconstruction, issues of purity and multiplicity need to be addressed. Applying stricter selection

criteria and advanced algorithms can enhance the identification of true signal events. The following subsections will detail these additional selection steps and their impact on improving the overall quality of the reconstructed signal.

Application of the `hlt_hadron` filter

Before focusing on further quality improvements, the `hlt_hadron` filter acting on collision data is applied to simulation. Only those simulated events that pass this filter are then selected for further reconstruction steps. The `hlt_hadron` filter is designed to reject events inconsistent with containing a hadron at the HLT level. It mainly includes selection steps to reject events consistent with characteristic Bhabha scattering and is effective at rejecting these low-multiplicity events. This is reflected in the reconstruction efficiencies, with greater than 96% efficiency for simulated $B\bar{B}$ and continuum events, and more than 99.95% of low-multiplicity events rejected.

Impact parameter selections

To ensure only tracks from the IP are considered, selections are imposed on the impact parameters recommended by the Belle II performance group. For lepton candidates, $d_r < 1.0$ cm and $|d_z| < 3.0$ cm are required, while for pion candidates, $d_r < 2.0$ cm and $|d_z| < 4.0$ cm are required. These requirements correspond to a cylindrical region centered on the IP with a length of 3.0 cm (4.0 cm for pions) along the z -axis and a radius of 1.0 cm (2.0 cm for pions) in the transverse plane. The chosen thresholds are loose enough to reject particles not coming from the IP without introducing bias due to impact parameter modeling. The looseness results in a selection efficiency greater than 97% for signal events but it ranges from 60% to 70% for background events.

CDC hits for pion candidates

In addition, a selection criterion is applied on the number of hits in the CDC recommended by the Belle II performance group to identify pion candidates, selecting only those with more than 20 hits. This ensures the selected pion candidates leave enough information in the CDC to reconstruct a high-quality track and are more likely to reach the PID detectors. Information from the PID detectors improves performance and helps differentiate pions from electrons. The signal selection efficiency is 93% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 90% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal events. Approximately 80% of $B\bar{B}$ background, 85% of continuum, and 65% of low-multiplicity events are retained.

Pion PID selection

PID likelihood ratios, as discussed in Chapter 3.4.4 and described by Equations 3.1 and 3.2, are used to select reconstructed electron, muon, and pion candidates compatible with the corresponding particle hypotheses. All pion candidates are required to have PID likelihood ratios greater than 0.1. The pion PID likelihood ratio includes information from the SVD, TOP counter, CDC, ECL, ARICH counter, and KLM detector. The average pion identification efficiency is 86%, with signal selection efficiencies of 91% and 81% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, respectively. The kaon and lepton misidentification rates are 7% and 0.4%, respectively, retaining 70% of the continuum and $B\bar{B}$ background events in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode and 50% in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, while only 30% of the remaining low-multiplicity backgrounds are retained. Normalized distributions of the pion PID

likelihood ratio for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ signal, continuum, and $B\bar{B}$ background events, and the selection thresholds, are shown in Figure 5.1(a). Similar distributions are observed for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode.

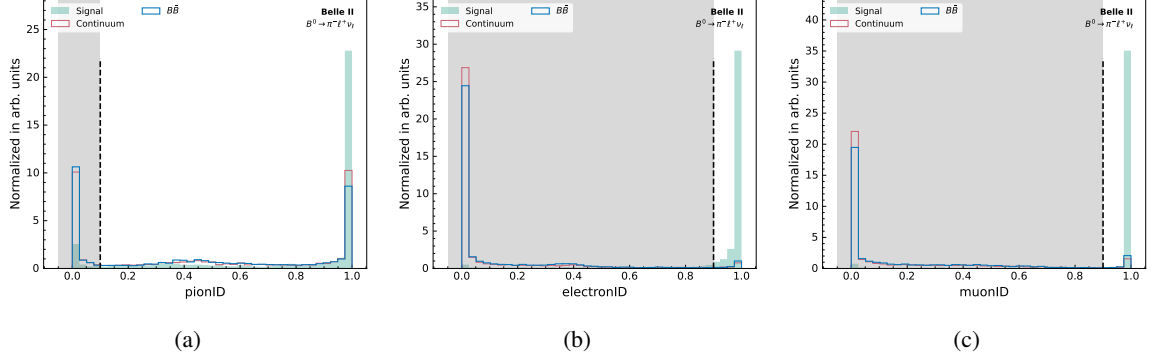


Figure 5.1: Normalized simulated distributions of the (a) pion, (b) electron, and (c) muon PID likelihood ratios for simulated signal, continuum, and $B\bar{B}$ background events in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode. The selection thresholds are indicated, and the rejected regions are shaded grey.

Lepton PID selection

Similarly, the electron PID likelihood ratio combines information from the CDC, ECL, ARICH counter, and KLM detector, while the muon PID likelihood ratio also includes data from the TOP counter. The PID likelihood ratios for electron and muon candidates are required to be greater than 0.9. The average lepton identification efficiency is 92%, resulting in a signal selection efficiency of 90%. The hadron misidentification rates are 0.2% for electron PID selection and 3.2% for muon PID selection. Lepton PID selections remove approximately 95% of $B\bar{B}$ and continuum background events, while about 20% of the remaining low-multiplicity events are retained. Normalized distributions of the electron and muon PID likelihood ratios for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode are shown in Figures 5.1(b) and 5.1(c). Similar distributions are observed for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode.

Polar angle selection

As discussed in Section 4.3.1, corrections for identification efficiencies between collision and simulated data are applied in polar angle and momentum bins of the particle. This setup is selected, since the identification performance varies with detector specifications, resulting in resolutions that depend on detector region and momentum. The modeling of polar angle and momentum distributions of lepton and pion candidates after applying corrections described in Section 4.3.1 is examined. Some residual disagreement is observed in the extreme regions of the lepton and pion-candidate polar angle distributions (θ_ℓ and θ_π). This likely results from mismodeling of the detector response in these regions, not fully captured by the physics control modes. To address this, the mismodeled regions are removed, retaining events with $\cos \theta_\ell > -0.85$ and $-0.85 < \cos \theta_\pi < 0.85$. The signal selection efficiency remains at 90%, and a similar fraction of background events is retained. During the validation of fit results in Section 7.5, stability tests are performed to ensure this selection does not bias the results.

5.2.2 Signal-specific selections

Following the preselection steps, the overall signal selection efficiency is 59% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events and 48% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. Purities have improved significantly, increasing by factors of 35 and 54 to 110×10^{-7} for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 38×10^{-7} for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. The selections are extended by performing vertex fits [114] to the hadron and lepton candidates and require that they converge. The fraction of candidates rejected by this vertex fit condition is very low, less than 0.01% for all event categories. Additional signal-specific selections are implemented, as discussed in the following subsections.

Lepton momentum

In semileptonic $B \rightarrow X \ell \nu_\ell$ decays, the lepton phase space depends on the mass of the X hadron. Hadrons containing a charm quark, such as X_c hadrons, are heavier than the pion or the ρ meson. Consequently, the lepton momentum spectrum for $B \rightarrow X_c \ell \nu_\ell$ decays peaks at lower momenta compared to $B \rightarrow X_u \ell \nu_\ell$ decays. Additionally, $B \rightarrow X_c \ell \nu_\ell$ decays have larger branching fractions due to the suppression of bottom quark transitions to up quarks compared to transitions to charm quarks. This is reflected in the relative magnitudes of $|V_{ub}|$ and $|V_{cb}|$, as discussed in Section 2.1.5.

The $B \rightarrow X_c \ell \nu_\ell$ decays thus represent a significant background in $B \rightarrow X_u \ell \nu_\ell$ analyses, and it is key to reduce their contribution. One way of targeting these decays is through the lepton momentum spectrum. First the four-momenta of electron candidates are corrected for bremsstrahlung by adding the four-momenta of photons with cluster energies below 1.0 GeV for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 0.5 GeV for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, found within a cone of 0.05 rad around the electron momentum vector. The maximum opening angle and cluster energies were optimized using simulated signal event samples. The optimization of the bremsstrahlung correction parameters is detailed in Appendix A.1.

Normalized distributions of the CM lepton momenta p_ℓ^* for simulated signal events, $B\bar{B}$, and continuum background events are shown in Figures 5.2(a) and 5.2(b) for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, respectively. In the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, leptons with CM momenta in the range [1.0, 2.85] GeV are selected to retain 92% of the signal events while rejecting approximately 80% of the background candidates. The lower threshold is particularly effective at targeting leptons from $B \rightarrow X_c \ell \nu_\ell$ events, while the upper threshold primarily filters out continuum events inconsistent with signal B meson decays.

Given the larger combinatorial backgrounds in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, a tighter lower threshold is applied to balance signal retention and background rejection. This is possible without a significant loss in signal selection efficiency, since the CM lepton momentum spectrum of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays peaks at larger values than for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays, due to the different spin structure of the decay to a vector ρ meson. $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events with CM lepton momenta in the range [1.4, 2.85] GeV are selected. This selection yields a signal efficiency of 83%, with 90% of the background events removed.

Di-pion invariant mass

In reconstructing the ρ meson candidate from two charged pion candidates in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, the di-pion invariant mass $m_{\pi\pi}$ is investigated. The normalized distributions of $m_{\pi\pi}$ for simulated signal, $B\bar{B}$, and continuum background events in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode are presented

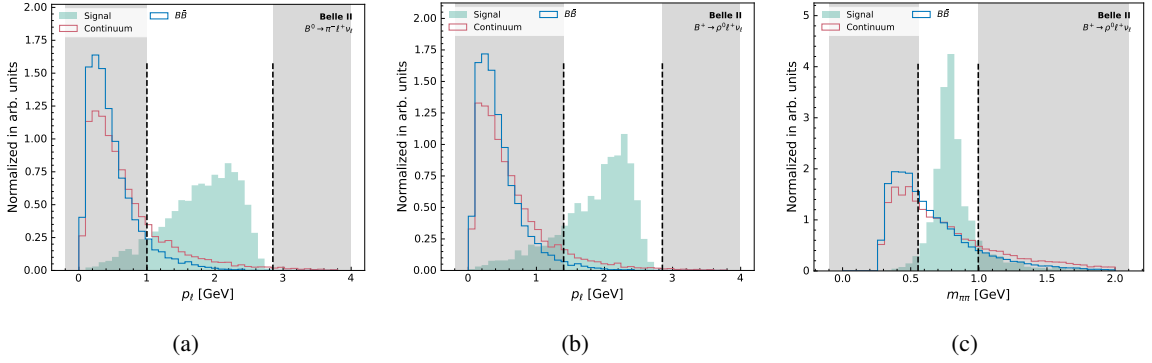


Figure 5.2: Normalized simulated distributions of the CM lepton momentum in the (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes. The normalized distribution of the invariant mass $m_{\pi\pi}$ of the $\rho \rightarrow \pi\pi$ decay in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays is shown in (c). The distributions are separately shown for simulated signal, continuum, and $B\bar{B}$ background events. The selection thresholds are indicated, and the rejected regions are shaded grey.

in Figure 5.2(c). These distributions incorporate specific models for the ρ meson lineshape and the branching fractions of nonresonant $B^+ \rightarrow \pi^+ \pi^- \ell^+ \nu_\ell$ events, as detailed in Section 4.2.3 and 4.2.4.

While correctly reconstructed ρ mesons, which do not originate from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal decays, contribute to a peak at the ρ meson mass, simulated $B\bar{B}$ and continuum background samples show a more pronounced combinatorial background with no resonant structure below the ρ meson peak. To effectively reduce this combinatorial background, a tight selection around the ρ meson mass would be ideal. However, to avoid introducing bias from potential mismodeling of the ρ meson mass, a stringent selection is avoided. Potential mismodeling could arise from ρ - ω meson interference, as discussed in Section 4.2.3. Thus, events within a window centered on the PDG mass of the ρ meson, with a total width of three times the PDG decay width: $0.554 < m_{\pi\pi} < 0.996$ GeV [37], are retained. This selection results in a signal efficiency of 85%, retaining 35% of $B\bar{B}$ and continuum background events, and only 0.3% of low-multiplicity events.

Hadron-lepton angle: $\cos \theta_{BY}$

To filter out candidates inconsistent with the kinematics of signal B meson decays, the hadron-lepton angle distribution is used. In semileptonic $B \rightarrow X \ell \nu_\ell$ decays, the combination of the signal hadron and lepton candidates is referred to as a Y candidate. Assuming only a single massless particle (the neutrino) has not been reconstructed, the recoil momentum can be used to define the cosine of the angle between the Y and the B meson as:

$$\cos \theta_{BY} = \frac{2E_B^* E_Y^* - m_B^2 - m_Y^2}{2|\vec{p}_B^*| |\vec{p}_Y^*|}, \quad (5.1)$$

where E_B^* and $|\vec{p}_B^*|$ are the energy and magnitude of the three-momentum of the B meson, respectively, calculated from beam properties, and m_B is the mass of the B meson [37]. Similarly, E_Y^* , $|\vec{p}_Y^*|$, and m_Y are the energy, magnitude of the three-momentum, and invariant mass of the Y candidate in the CM frame, respectively. For correctly reconstructed signal events with perfect resolution, $\cos \theta_{BY}$ should range from -1 to 1 .

To retain sufficient background events for training the classifiers discussed in Chapter 6, a loose selection of $|\cos \theta_{BY}| < 1.6$ is applied. This selection achieves approximately 96% signal efficiency, while retaining 4% and 19% of $B\bar{B}$ background events, and 15% and 46% of continuum background events in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, respectively. As a result, the purity increases by factors of 14 and 4 in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, respectively. The simulated distribution of $\cos \theta_{BY}$ for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, with the selection indicated by dashed lines, is shown in Figure 5.3(b). A similar distribution is observed for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode. The histogram event categories are described in Section 5.4.

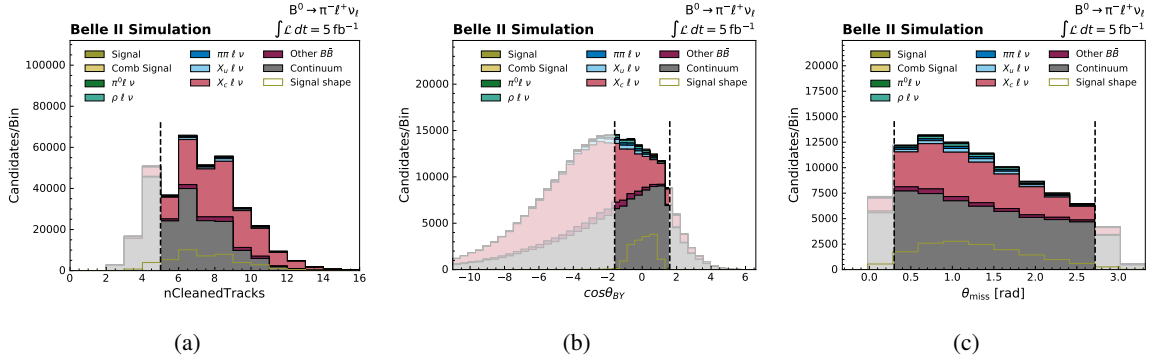


Figure 5.3: Simulated distributions of the (a) number of tracks, (b) $\cos \theta_{BY}$, and (c) missing momentum angle in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode. The events are categorized as described in Section 5.4. Each distribution is shown after the preceding selection has been applied. Selection thresholds are indicated, and the rejected regions are shaded grey.

5.3 Missing momentum reconstruction

In untagged semileptonic measurements at B-factories, although the signal neutrino is undetected, its properties can be inferred from the ROE. This section introduces the concept of missing momentum, which is attributed to the neutrino. The accuracy of predicting the missing momentum from ROE information depends on the quality of the tracks and clusters included, and thus specific requirements on these components are imposed. These requirements and additional event selection steps based on ROE information are also discussed below.

Concept and definition of missing momentum

At B-factories such as the SuperKEKB accelerator, the initial conditions are well-known, including that the sum of the momenta of the two B mesons corresponds to the momentum of the $\Upsilon(4S)$ meson, $\vec{p}_{\Upsilon(4S)}$. Assuming that the ROE includes all tracks and clusters associated with the non-signal B meson:

$$\vec{p}_{\Upsilon(4S)} = \vec{p}_{\text{ROE}} + \vec{p}_X + \vec{p}_\ell + \vec{p}_\nu,$$

where \vec{p}_{ROE} is the sum of the momenta of all tracks and clusters in the ROE, and \vec{p}_X , \vec{p}_ℓ , and \vec{p}_ν are the three-momenta of the hadron, lepton, and neutrino, respectively. In the CM frame of the $\Upsilon(4S)$

meson, the missing momentum is defined as:

$$\vec{p}_\nu^* = -(\vec{p}_X^* + \vec{p}_\ell^* + \vec{p}_{\text{ROE}}^*).$$

Since the neutrino is not detected, its momentum is considered the missing momentum, i.e., $\vec{p}_{\text{miss}}^* = \vec{p}_\nu^*$. Similarly, the missing energy can be defined as $E_{\text{miss}}^* = E_\nu^*$. The missing four-momentum in the CM frame is then given by:

$$(E_{\text{miss}}^*, \vec{p}_{\text{miss}}^*) = (E_{Y(4S)}^*, \vec{p}_{Y(4S)}^*) - \left(\sum_i E_i^*, \sum_i \vec{p}_i^* \right),$$

where E_i^* and \vec{p}_i^* are the CM energy and momentum of the i th track or cluster in the event, respectively.

An alternative definition of the missing energy is often used: $E_{\text{miss}}^* = E_\nu^* = |\vec{p}_\nu^*| = |\vec{p}_{\text{miss}}^*|$. This definition is preferred to avoid using information about the energy of the ROE. This approach is advantageous because the energy of each particle in the ROE depends on the mass hypothesis applied, making energy determination prone to errors from particle identification. Nevertheless, even using the momentum of the ROE to estimate energy introduces a residual error due to the mass hypothesis applied when transforming individual momenta from the lab frame to the CM frame.

ROE track and cluster selection

To obtain an accurate prediction of the neutrino momentum, the ROE must be as pure and complete as possible. To achieve this the quality of both tracks and clusters contributing to the ROE needs to be considered. As discussed earlier, applying the correct mass hypothesis to tracks is crucial for defining the energy and momentum of the ROE. Thus, each track in the ROE is assigned its “most-likely” particle hypothesis by selecting the mass hypothesis α that maximizes the PID likelihood ratio PID_α .

To minimize the impact of beam-induced background and acceptance losses, the following requirements are imposed on track candidates contributing to the ROE:

- Tracks must be within the CDC acceptance,
- Tracks must have transverse momenta greater than 0.05 GeV,
- Tracks must have impact parameters satisfying $dr < 1.0$ cm and $|dz| < 3.0$ cm.

For clusters included in the ROE, recommendations from the Belle II performance group re applied. Each cluster must be within the CDC acceptance and consist of more than one crystal. Additionally, only clusters detected within 200 ns of the collision time are considered, which corresponds to approximately five times the mean timing resolution of the calorimeter.

To ensure that the simulated ROE matches the ROE in collision data, the track and cluster distributions, such as transverse momentum p_t and cluster energy, are compared between simulation and collision data. The agreement for various p_t and cluster energy thresholds is examined in the forward, barrel, and backward detector regions. The procedure is detailed in Appendix A.2. Good agreement is observed in the p_t distributions with a threshold of $p_t > 0.05$ GeV, indicating no need for tighter track requirements.

For cluster energy spectra, the agreement between simulated and collision data is highly dependent on beam-induced background levels, which vary significantly by detector region. Given that beam-induced backgrounds primarily produce low-energy clusters, regions with significant discrepancies

are eliminated using energy thresholds. Therefore, region-dependent cluster energy thresholds are introduced. Clusters with energies greater than 0.060, 0.050, and 0.075 GeV in the forward, barrel, and backward directions, respectively, are allowed to contribute to the ROE.

Selection on missing momentum angle

After selecting the tracks and clusters contributing to the ROE, events are addressed where massive particles escape undetected, potentially biasing the missing momentum estimation. Accurately estimating the missing momentum is essential for achieving high resolution in the signal extraction variables (see Section 5.5). Events with the polar angle of the missing momentum θ_{miss} in the lab frame outside the CDC acceptance are therefore rejected, requiring $0.30 \text{ rad} < \theta_{\text{miss}} < 2.72 \text{ rad}$. The simulated distribution of θ_{miss} is shown in Figure 5.3(c). A similar distribution is observed for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode. This selection results in an efficiency of approximately 95% for all event categories in both modes.

Selection on track multiplicity

The average track multiplicity of B meson decays at the $\Upsilon(4S)$ energy is about 5.4 [115], so the number of tracks in the ROE is expected to peak around five. Including an additional two (three) tracks from the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ ($B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$) signal side, the total number of tracks for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ ($B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$) signal events should peak around seven (eight). The simulated distribution of track number for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode is shown in Figure 5.3(a). A similar distribution is observed for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode.

Track multiplicity is a critical feature for distinguishing between signal and low-multiplicity events. A significant fraction of Bhabha, two-photon, and other low-multiplicity events can be rejected without major losses in signal selection efficiency by setting a minimum track threshold. This threshold is determined by examining the background with the highest track count, which are primarily two-photon events like $ee \rightarrow eeee$ and $ee \rightarrow ee\mu\mu$, each resulting in four tracks. Thus, only events with more than four tracks are retained. The threshold is indicated by the dashed line in Figure 5.3(a). The signal selection efficiency is 88% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events and 99% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, due to the different number of tracks on the signal B meson side. This discrepancy arises from the additional track on the signal B meson side in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events compared to $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events. This selection retains only 0.07% of low-multiplicity events and eliminates an additional 20% of continuum backgrounds.

After these reconstruction and selection steps, the overall signal selection efficiency is 45% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 30% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. Purities have increased by factors of 58 and 77 compared to those mentioned in Section 5.2.2, reaching 6.4×10^{-4} ($B^0 \rightarrow \pi^- \ell^+ \nu_\ell$) and 3.0×10^{-4} ($B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$).

5.4 Event categorization

The simulated event categories depicted in Figure 5.3 are based on the origin of the signal lepton candidate. Events are divided into two primary categories: $B\bar{B}$ events and continuum backgrounds. The $B\bar{B}$ event category is further subdivided into signal, cross-feed, $B \rightarrow X_u \ell \nu_\ell$, $B \rightarrow X_c \ell \nu_\ell$, and

other $B\bar{B}$ background categories. These event categories are described in the following and they are revisited during the signal extraction discussion in Chapter 7.

Signal: Correctly reconstructed signal events are classified as *signal* or true signal events. Events where the signal lepton is correctly reconstructed but the signal hadron is not fall into the *combinatorial signal* category. Another signal category is *isospin-conjugate signal* events. In the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, these events correspond to cases where the signal lepton originates from a $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ decay. In the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, they refer to events where the signal lepton is produced in a $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ decay. In Figure 5.3, true, combinatorial, and isospin-conjugate signal events are shown as separate event categories.

Cross-feed: In the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, the *cross-feed* category includes events where the signal lepton originates from a $B \rightarrow \rho \ell \nu_\ell$ decay. In the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, a similar category is defined for $B \rightarrow \pi \ell \nu_\ell$ cross-feed events. Defining these cross-feed components helps in cross-linking the two reconstruction modes during signal extraction, as discussed in Section 7.4.

$B \rightarrow X_u \ell \nu_\ell$ and $B \rightarrow X_c \ell \nu_\ell$: Events where the signal lepton originates from other $B \rightarrow X_u \ell \nu_\ell$ decays are categorized under $B \rightarrow X_u \ell \nu_\ell$ background. This category includes events where X_u is a well-known resonant state (ω, η, η'). It also contains events where the signal lepton comes from nonresonant $B \rightarrow X_u \ell \nu_\ell$ decays, such as $B \rightarrow \pi \pi \ell \nu_\ell$ decays. Figure 5.3 shows these nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ background events as a separate event category. Similarly, the $B \rightarrow X_c \ell \nu_\ell$ background category includes all events where the signal lepton comes from a $B \rightarrow X_c \ell \nu_\ell$ decay.

Other $B\bar{B}$: The previously mentioned categories cover events where the signal lepton is produced directly in a semileptonic B meson decay (primary lepton) and has been correctly identified (true lepton). The *other $B\bar{B}$* background category encompasses remaining $B\bar{B}$ background events. This category includes events with incorrectly identified signal leptons (fake leptons) or cases where the signal lepton does not originate directly from a semileptonic decay (secondary lepton). Additionally, it includes a small number of events where leptons are produced in other B meson decays, such as $B \rightarrow X_s \ell \ell$ with X_s being a hadron containing a strange quark.

Continuum: The *continuum* background category consists of events where the electron-positron collision produces a $q\bar{q}$ or $\tau\tau$ pair, rather than an $\Upsilon(4S)$ meson and a $B\bar{B}$ pair. Low-multiplicity events remaining after the previous selections are also included in the continuum background category. This predominantly includes $ee \rightarrow ee\tau\tau$ events.

5.5 Signal extraction variables

In this section the variables used to extract signal events are described. This includes the beam-constrained mass M_{bc} and the energy difference ΔE , which are constructed using beam energy information. The method used to reconstruct the squared momentum transfer q^2 from the B meson to the hadron is also described.

Beam-constrained mass and energy difference

To test the kinematic consistency of an event with a signal B meson decay, ROE information is used to define the beam-constrained mass M_{bc} and the energy difference ΔE . The beam-constrained mass is the invariant mass of the B meson, with the single-beam energy E_{beam}^* in the $\Upsilon(4S)$ rest frame assigned as the reconstructed energy of the B meson, which is half of the CM energy \sqrt{s} . In the CM frame, M_{bc} is thus defined as

$$M_{bc} = \sqrt{E_{\text{beam}}^{*2} - |\vec{p}_B^*|^2} = \sqrt{\left(\frac{\sqrt{s}}{2}\right)^2 - |\vec{p}_B^*|^2}, \quad (5.2)$$

where \vec{p}_B^* is the reconstructed momentum of the B meson in the $\Upsilon(4S)$ rest frame. The energy difference ΔE is defined as the difference between the reconstructed and expected B meson energy. In the CM frame, it is given by

$$\Delta E = E_B^* - E_{\text{beam}}^* = E_B^* - \frac{\sqrt{s}}{2}, \quad (5.3)$$

where E_B^* is the reconstructed energy of the B meson, and the expected B meson energy is E_{beam}^* .

To calculate both ΔE and M_{bc} , the four-momentum of the reconstructed B meson in the CM frame p_B^* must be known. Using the definition of missing momentum from Section 5.3, $\vec{p}_B^* = \vec{p}_X^* + \vec{p}_\ell^* + \vec{p}_{\text{miss}}^*$ and $E_B^* = E_X^* + E_\ell^* + |\vec{p}_{\text{miss}}^*|$ are defined. The resolutions of ΔE and M_{bc} are limited by the accuracy of estimating the neutrino momentum using ROE information. Thus, the contents and purity of the ROE, described in Section 5.3, are crucial.

The distributions of ΔE and M_{bc} in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode (before applying selections on the number of tracks, missing momentum, and $\cos \theta_{BY}$) are shown in Figure 5.4(a) and 5.4(b), respectively. Similar distributions are observed for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode. As expected, the energy difference ΔE is consistent with 0 GeV for correctly reconstructed signal events. For these events, M_{bc} is consistent with the B meson mass, approximately 5.279 GeV [37].

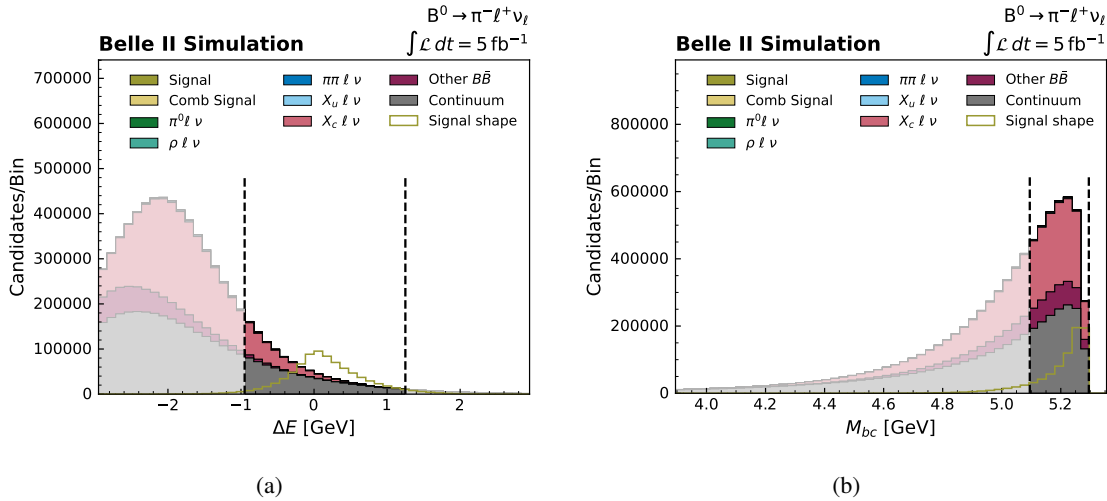


Figure 5.4: Simulated distributions of (a) ΔE and (b) M_{bc} in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode before applying cuts on the number of tracks, missing momentum, and $\cos \theta_{BY}$.

A two-dimensional fit region in ΔE and M_{bc} with $-0.95 < \Delta E < 1.25$ GeV and $5.095 < M_{bc} < 5.295$ GeV is defined. Visualizing the effect of the selection is helpful with two-dimensional distributions of ΔE and M_{bc} , shown for signal and background events in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode in Figure 5.5(a) and 5.5(b), respectively. Similar distributions are observed for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode. In Figure 5.4 the fit region selection is indicated by dashed lines, while in Figure 5.5, it is represented by the green box. The selected region is enriched in signal, while it retains sub-regions sufficiently populated by backgrounds for distinguishing events during signal extraction.

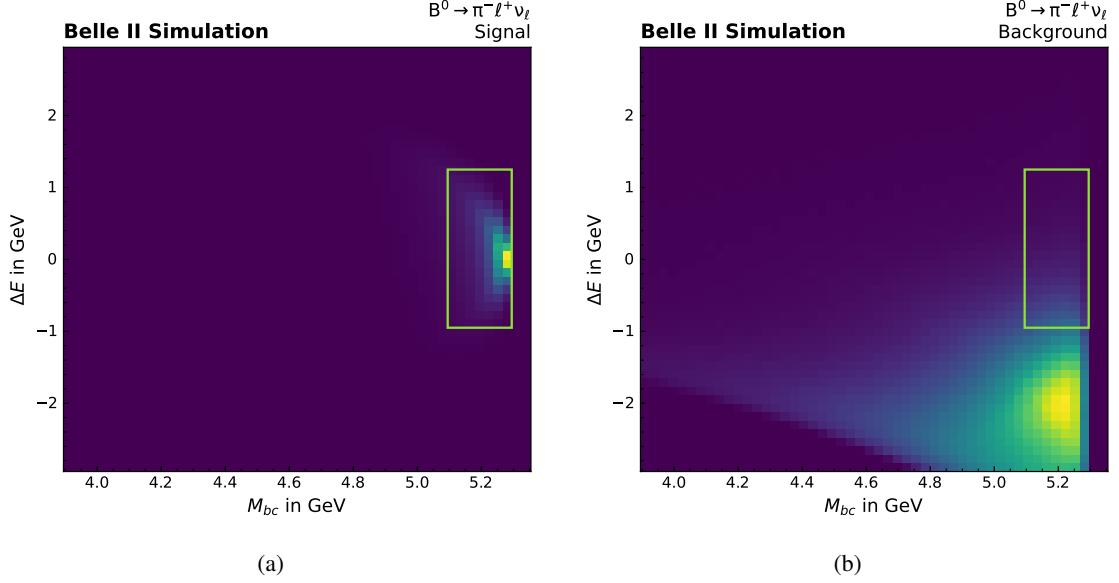


Figure 5.5: Two-dimensional distributions of ΔE and M_{bc} for simulated (a) signal and (b) background events in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode. The selected fit region is marked by the green rectangle.

The signal selection efficiency for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events is 82%, while 85% of the background events are rejected. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, the signal efficiency is similar, but the fraction of rejected background events is smaller at 70%. The reduced background rejection is due to the background in the ΔE distribution peaking closer to the selected fit region than in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode. This feature is due to the combination of the selection on the di-pion invariant mass of the $\rho \rightarrow \pi\pi$ decay, the tighter lepton momentum selection, and reconstruction losses in the ρ meson.

Momentum transfer squared

As discussed in Section 2.2, the momentum transfer squared q^2 describes the kinematics of semileptonic decays. For exclusive final states, q^2 is used to parameterize the form factor, which includes non-perturbative effects. By measuring the differential decay rate as a function of q^2 , the value of $|V_{ub}|$ can be determined, as detailed in Equations 2.15 and 2.16.

For semileptonic $B \rightarrow X \ell \nu_\ell$ decays, q^2 is defined by

$$q^2 = (p_B - p_X)^2 = (p_\ell + p_\nu)^2, \quad (5.4)$$

where p_B , p_X , p_ℓ , and p_ν represent the four-momenta of the B meson, hadron, lepton, and neutrino,

respectively. The upper endpoint of q^2 is determined by the masses of the B meson and the hadron. For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays, this endpoint is 26.4 GeV^2 , and for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, it is 20.3 GeV^2 . The lower endpoint is set by the mass of the lepton. For decays involving light leptons (e, μ), the lepton mass is negligible, resulting in a lower endpoint of 0 GeV^2 .

To perform a differential measurement, the q^2 range is divided into 13 bins for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 10 bins for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. The first 12 (9) bins have uniform widths, while the final bins extend to the respective kinematic endpoints. The bin labels and edges for q^2 are: $q1 : q^2 \in [0, 2]$, $q2 : [2, 4]$, $q3 : [4, 6]$, $q4 : [6, 8]$, $q5 : [8, 10]$, $q6 : [10, 12]$, $q7 : [12, 14]$, $q8 : [14, 16]$, $q9 : [16, 18]$, $q10 : [18, 20(20.3)]$, $q11 : [20, 22]$, $q12 : [22, 24]$, $q13 : [24, 26.4] \text{ GeV}^2$.

In simulation, q^2 bins can be defined both in true and reconstructed q^2 spaces. True values of q^2 are calculated from the generated B meson and hadron momenta using Equation 5.4. Ideally, reconstructed values of q^2 are obtained by reconstructing the momenta of the B meson and hadron. However, due to the missing neutrino component, this process is less straightforward as the B meson momentum must be estimated. Several approaches to estimating this momentum are discussed and compared below:

Rest frame: A straightforward approach assumes that the B mesons decay at rest. In this scenario, the B meson energy is half the CM energy, and its momentum is zero. In reality, the B meson is not at rest but carries a total momentum of 0.332 GeV due to the difference between the $\Upsilon(4S)$ mass and twice the B meson mass. Therefore, the resolution in q^2 using the rest frame method is limited by this assumption.

Diamond Frame: Alternative methods use the kinematics of the signal hadron and lepton candidates to predict the B meson momentum \vec{p}_B^* . The Diamond Frame technique [116] predicts the B meson momentum using the $\cos \theta_{BY}$ variable defined in Equation 5.1. This method constrains the B meson momentum to lie on the surface of a cone defined by the θ_{BY} angle and the Y candidate momentum. Figure 5.6(a) shows a sketch of the cone construction. The B meson momentum is determined as a weighted average over possible configurations on the cone surface, with a weight of $\sin^2 \theta_B$. This weight is based on the assumption that the electron-positron beams are effectively massless and polarized along the beam axis. The produced spin-1 $\Upsilon(4S)$ meson thus has its spin aligned with the axis, and it decays to two spinless B mesons with prior probability of the B meson flight direction relative to the beam axis following a $\sin^2 \theta_B$ distribution. The average over ten configurations of the B meson momentum, evenly distributed in azimuthal angle on the cone, is predicted as the B meson momentum. Figure 5.6(a) illustrates two such configurations.

ROE method: The ROE method combines kinematic information from the signal hadron and lepton candidates, represented by θ_{BY} , with the ROE momentum vector. The B meson momentum vector on the θ_{BY} cone that is most back-to-back with the ROE vector is selected. Figure 5.6(b) illustrates the ROE method [117]. The chosen B meson momentum vector is the one on the θ_{BY} cone that maximizes the projection onto the inverted ROE momentum vector, $-\vec{p}_{\text{ROE}}^*$.

Combined method: The combined method integrates the Diamond Frame and ROE methods. This approach resembles the Diamond Frame technique, but before averaging over the ten vectors on the θ_{BY} cone, each vector is weighted by $\frac{1}{2}(1 - \hat{p}_B^* \cdot \hat{p}_{\text{ROE}}^*)$. Here, \hat{p}_B^* and \hat{p}_{ROE}^* are the unit vectors of

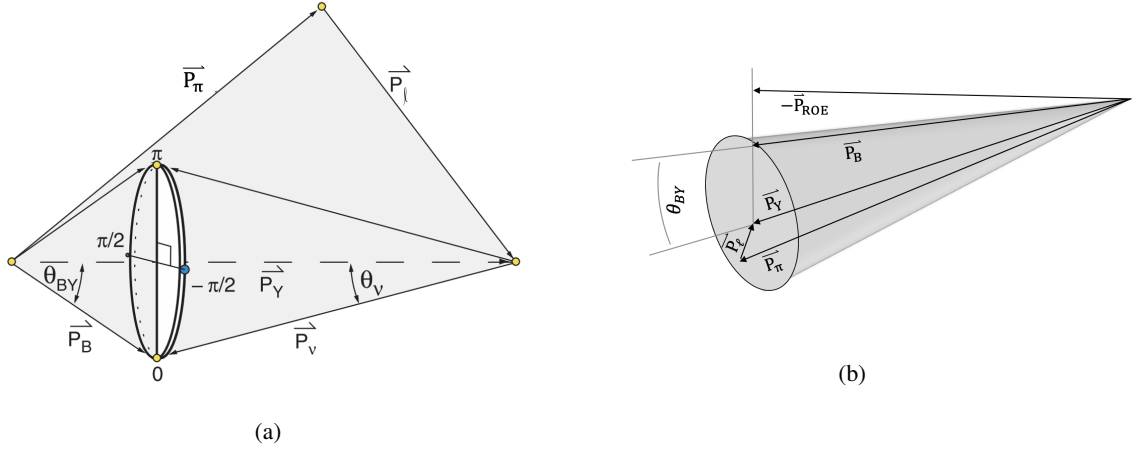


Figure 5.6: (a) Illustration of the cone defined by θ_{BY} and the Y candidate momentum used to predict the B meson momentum using the Diamond Frame technique. Image from Ref. [116]. (b) Illustration of the ROE method, which selects the B meson momentum vector on the θ_{BY} cone that is most back-to-back with the ROE momentum vector.

\vec{p}_B^* and \vec{p}_{ROE}^* , respectively. This factor adjusts for the projection onto the ROE momentum vector. Additionally, the $\sin^2 \theta_B$ weight is included in the combined method [118].

The resolutions in q^2 obtained using the rest frame, Diamond Frame, ROE, and combined methods for simulated $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ and $B^+ \rightarrow \rho^0 \mu^+ \nu_\mu$ are shown in Figure 5.7(a) and 5.7(b), respectively. The ROE method exhibits the most prominent peak in resolution but also displays long tails. The Diamond Frame technique has a less pronounced peak and does not exhibit long tails. The combined method effectively integrates the advantages of both the ROE and Diamond Frame techniques.

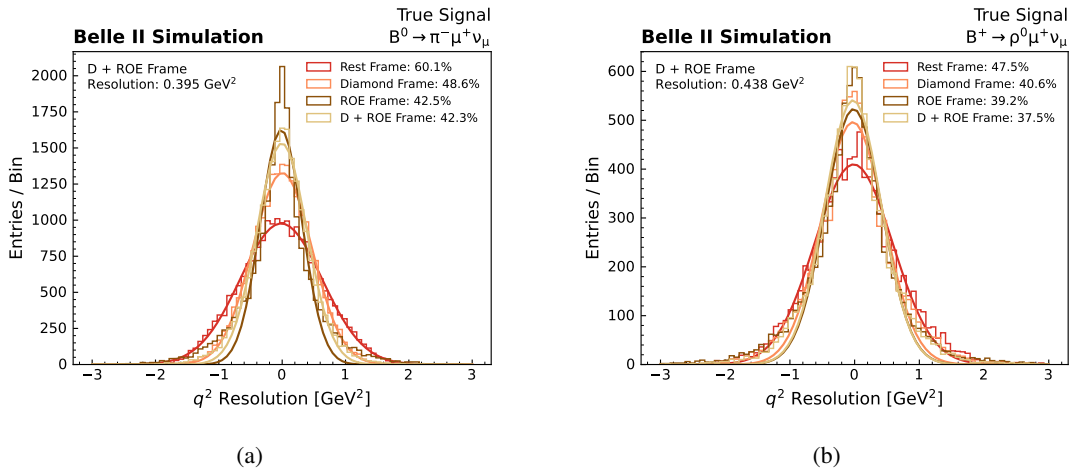


Figure 5.7: Resolution in q^2 achieved in simulation using the rest frame, Diamond Frame, ROE, and combined (labeled as $D + ROE$ Frame) methods for simulated (a) $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ and (b) $B^+ \rightarrow \rho^0 \mu^+ \nu_\mu$ signal events.

The resolutions are quantitatively compared by approximating them as Gaussian functions, as shown in Figure 5.7. For each method, the width of the corresponding Gaussian function is determined. The smallest width σ is achieved using the combined method for both modes. For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal events, the resolutions are $\sigma = 0.395 \text{ GeV}^2$ and $\sigma = 0.438 \text{ GeV}^2$, respectively. The poorer resolution in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode is mainly due to reconstruction losses associated with the ρ meson. To fully assess the most suitable method, it is necessary to evaluate not just the widths but also the tail behavior of the resolution functions. Therefore, the fraction of events falling outside the $[-\sigma, \sigma]$ range for each method, as indicated in the legends of Figure 5.7, is also determined. For both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, the combined method results in the smallest fraction of events outside this range. Thus, the combined method is used for estimating q^2 , as in simulation it better assigns events to the correct true and reconstructed q^2 bins compared to the other methods.

The resolutions in q^2 decrease with increasing q^2 , ranging from 0.09–0.60 GeV^2 for simulated $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events and 0.16–0.84 GeV^2 for simulated $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. These values are smaller than the defined momentum transfer bin widths. Since the reconstructed values of q^2 determined this way depend on $\cos \theta_{BY}$, which is calculated using the beam energy, the effects of potential beam-energy differences in collision and simulated data on the q^2 resolutions need to be evaluated. This is described in Section 8.3.2.

The performance of the combined method is further evaluated using bin migration matrices. The process of bin migration refers to the misclassification of events from one q^2 bin to another due to the resolution limits of the detector system or the statistical fluctuations inherent in the simulated datasets used to model the physical processes. This misclassification can distort the true shape of the q^2 distribution, leading to systematic uncertainties in the extracted parameters. The bin migration matrix entry with indices i, j represents the fraction of events reconstructed in q^2 bin i with true q^2 in bin j . Ideally, diagonal entries with $i = j$ would be unity, with no events reconstructed in incorrect q^2 bins. Figure 5.8 presents the bin migration matrices for simulated $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ and $B^+ \rightarrow \rho^0 \mu^+ \nu_\mu$ events. The $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode shows significantly larger bin migrations compared to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode.

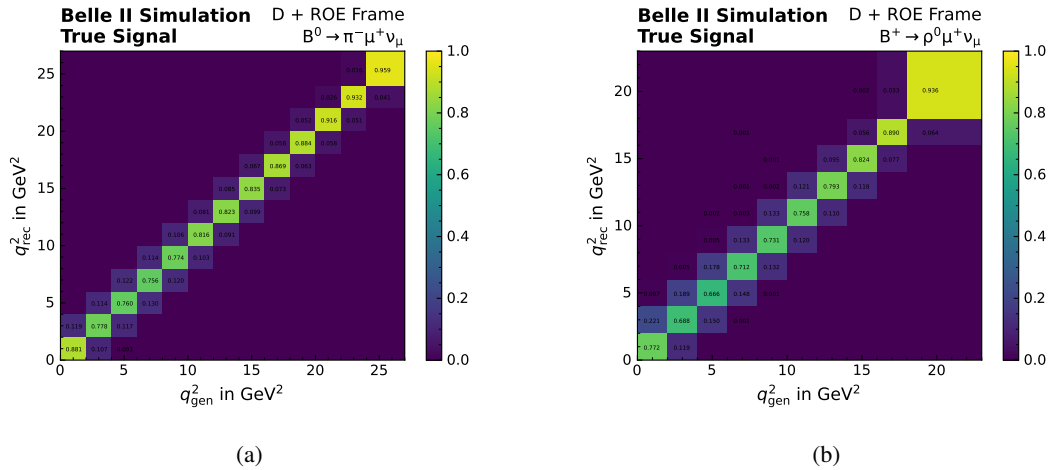


Figure 5.8: Migration matrices for q^2 bins from the combined method (labeled as $D + ROE \text{ Frame}$) for simulated (a) $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ and (b) $B^+ \rightarrow \rho^0 \mu^+ \nu_\mu$ signal events.

Background suppression

After reconstructing and selecting signal events, as discussed in Chapter 5, the dominant background contributions from continuum and $B\bar{B}$ events are addressed. Figure 6.1 shows the simulated distributions of ΔE , M_{bc} , and reconstructed q^2 for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes. The contributions from continuum and $B\bar{B}$ backgrounds stand out. This chapter describes the suppression of these backgrounds using multivariate analysis (MVA) to create classifiers that differentiate between signal and background events.

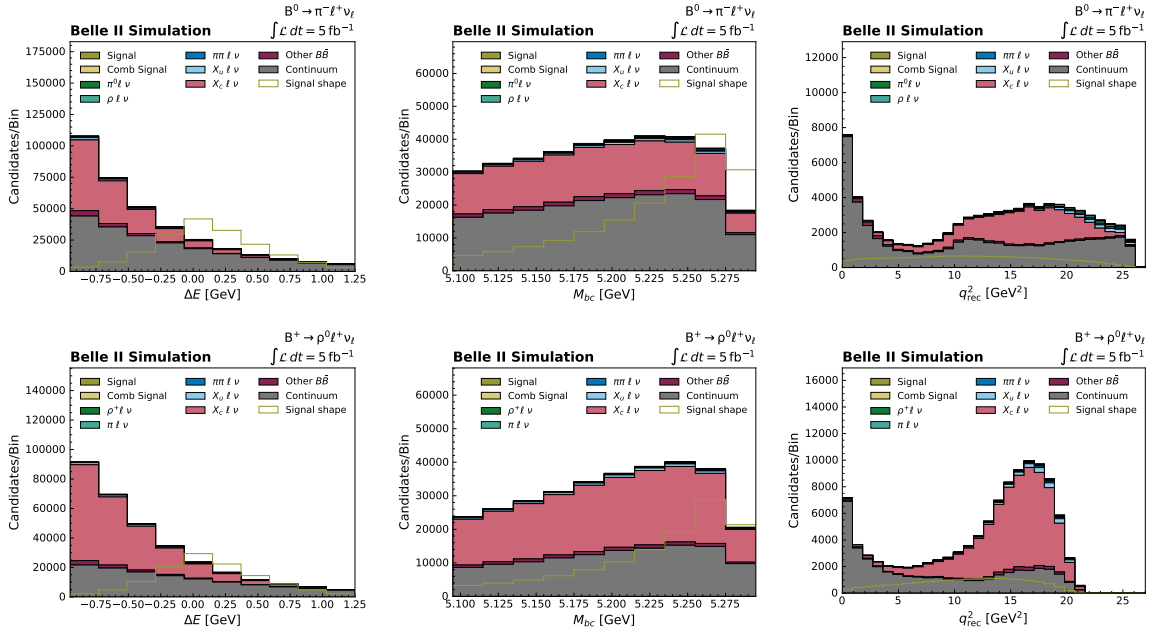


Figure 6.1: Simulated distributions of ΔE (left), M_{bc} (middle), and reconstructed q^2 (right) for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (top) and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (bottom) modes, following the selections described in Chapter 5.

First, Section 6.1 introduces the concept of MVA. In Section 6.2 its implementation and performance within the work discussed in this thesis are outlined. Subsequent to applying these background-suppression techniques, a best candidate selection is conducted, described in Section 6.3. The overall

effectiveness of the selection process is discussed in Section 6.4.

6.1 Multivariate analysis

This section introduces the concept of multivariate analysis (MVA), which involves techniques for extracting information from data using multiple variables. In particle physics, MVA methods are often employed to enhance signal-to-background ratios by classifying events. MVA algorithms may rely on linear algebra or integrate machine-learning (ML) techniques. ML algorithms exploit correlations in underlying statistical distributions that might not be immediately visible.

Common ML methods for classification in particle physics include neural networks and decision trees. These methods are trained using datasets with labeled data points categorized as *signal* or *background*. The output of the MVA classifier measures the ability to distinguish between signal and background categories and is, in some respects, proportional to a signal probability. Neural networks consist of layers of interconnected nodes (neurons), where connections are weights adjusted iteratively to maximize the signal probability. Decision trees use sequential cuts at nodes, with the final path determined by the highest signal probability. This discussion will focus on decision trees, although many concepts also apply to neural networks.

A key metric for evaluating MVA performance is the receiver operating characteristic (ROC) curve. The ROC curve plots background rejection versus signal efficiency, with the area under the ROC curve (AUC) obtained by integrating this function. Ideally, a perfect separation between signal and background categories results in an AUC of 1.0. If the MVA method performs no better than random guessing, the AUC is 0.5.

Further validation is required to ensure that the model generalizes beyond the training data to avoid overfitting (also known as overtraining). Overfitting occurs when the model memorizes the training data, capturing specific features or noise unique to that sample. A hypothesis test using both training and independent test samples with labeled data points can detect overfitting. The null hypothesis assumes that the test and training samples are statistically equivalent. A p-value is calculated based on the difference between test and training sample classifier distributions. The p-value quantifies the probability of observing the measured (or a more extreme) difference. Overfitting, indicated by a p-value below 0.05 [119], is common with single decision trees that have too many nodes, enabling them to target specific features.

An overly simple model can also be problematic, as it may fail to capture all key features of the sample. This results in disagreement between test and training sample classifiers, known as underfitting (or undertraining). A hypothesis test can detect underfitting with p-values greater than 0.95 [119]. Underfitting can often be mitigated by slightly increasing model complexity.

Reducing overfitting is typically more challenging, as simplifying the model reduces sensitivity to both noise and key features. To address overfitting, multiple trees (an ensemble) can be trained and combined using a boosting algorithm. Data points are weighted based on how well they are classified by each tree, with weights propagated during subsequent tree training, and each tree contributes to the final signal probability. This type of decision tree is known as a boosted decision tree (BDT). When the final classifier is constructed by combining the outputs of all trees rather than using only the final tree, it is called a gradient boosted decision tree. A stochastic gradient boosted decision tree (SGBDT) further reduces overfitting by training individual trees on subsets of the total sample.

To some extent, SGBDT performance can be tuned using hyperparameters, including:

- Number of trees: The total number of trees in the ensemble.
- Depth of the trees: Controls how deep each tree can grow.
- Number of cuts in a tree: Determines the total number of allowed cuts within a tree.
- Shrinkage: The rate at which trees are combined, scaling each tree's contribution.
- Sampling rate: The size of the sample subset used for training each tree within the SGBDT.

Tuning these hyperparameters can reduce overfitting and improve SGBDT classification performance.

6.2 Background suppression using boosted decision trees

Following the introduction of multivariate analysis (MVA), this section describes its application in the work discussed in this thesis. Given the dominance of $B\bar{B}$ and continuum backgrounds, as observed in Figure 6.1, the goal is to target these backgrounds specifically.

The setup of the BDT classifiers including the used data samples are discussed in Section 6.2.1, while the selected input variables are described in Section 6.2.2. Multiple SGBDTs are trained using the `FastBDT` [120] package integrated within the `basf2` framework. The validation of the classifiers and the rationale for selecting SGBDTs over simpler cut-based methods is discussed in Section 6.2.3. The performance of selecting on the output classifiers is described in Section 6.2.4.

6.2.1 Classifier setup and data samples

To address the dominant $B\bar{B}$ and continuum backgrounds in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, methods are implemented to suppress these events. The approaches, selected setup and data samples used to achieve this goal are discussed in this subsection.

One approach to suppress background events involves training a single BDT classifier for each mode to distinguish between signal and background events. An alternative approach trains separate classifiers to target continuum and $B\bar{B}$ backgrounds individually. This separation is motivated by the expectation that distinguishing signal from continuum events is easier compared to $B\bar{B}$ events, due to more distinct event shapes. The difference in event shape between continuum and $B\bar{B}$ events arises from the absence of B mesons and the hadronization of quark pairs into jets in continuum events. Further details are provided in Section 6.2.2.

Initially, continuum suppression classifiers are trained to separate signal from continuum events. In the second stage, the focus lies on $B\bar{B}$ events that are signal-like, identified by selecting regions with continuum suppression classifier outputs greater than 0.8. In these signal-enriched regions, $B\bar{B}$ background suppression classifiers are trained to target events that were not rejected by the continuum suppression classifiers.

Figure 6.1(c) shows that background composition varies in each q^2 bin, especially for $B\bar{B}$ background, where the other $B\bar{B}$ category dominates at low q^2 , the $B \rightarrow X_c \ell \nu_\ell$ events dominate at intermediate q^2 and $B \rightarrow X_u \ell \nu_\ell$ events dominate at high q^2 . To optimize background suppression in each q^2 bin, separate BDT classifiers are trained for each reconstructed q^2 bin. The $q10$ bin in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode is further split into two bins ($q10: q^2 \in [18, 19.5]$, $q11: q^2 \in [19.5, 20.3]$) to improve signal sensitivity in the highest q^2 bin. This results in a total of $2 \times (13 + 11) = 48$ BDTs.

The data samples used for training and validating the BDT classifiers are defined as follows. Each of the 48 BDT classifiers is trained and tested using simulated events with a 50:50 split between signal and

background events. Continuum background samples include simulated continuum, τ -pair, and two-photon $e\bar{e}\tau\tau$ events, with negligible contributions from other two-photon and low-multiplicity events. The continuum background samples consist of simulated continuum and taupair data corresponding to an integrated luminosity of 600 fb^{-1} to ensure the q^2 bin with the fewest events has at least 500,000 events. Additionally, the full sample of simulated two-photon and low-multiplicity events, scaled to 600 fb^{-1} , is used.

$B\bar{B}$ background samples are derived from 1 ab^{-1} of generic simulated data, including both neutral and charged B -meson events, excluding true signal events. A larger luminosity compared to the continuum sample is required because $B\bar{B}$ background suppression classifiers are trained in the signal-enriched regions of the continuum suppression classifiers, resulting in lower $B\bar{B}$ background retention. Signal samples consist of correctly-reconstructed signal events from the full simulated signal sample. In each q^2 bin, the number of signal events is scaled to match the number of background events. The samples for $B\bar{B}$ and continuum background suppression are split into training and test samples with a 70:30 ratio.

After selecting the final list of input variables (see Section 6.2.2), a grid search is conducted to determine the optimal hyperparameter settings. The number of trees has the most significant impact on the AUC. The focus lies thus on finding the value that maximizes it. For continuum suppression BDT classifiers, the best performance is achieved with 200 trees, while $B\bar{B}$ suppression BDT classifiers perform best with 150 trees. Other hyperparameters have a lesser impact on the AUC and are set to default values in the `basf2` framework, as follows:

- Number of trees = 200 / 150
- Depth of the trees = 3
- Number of cuts in a tree = 8
- Shrinkage = 0.2
- Sampling rate = 0.5

6.2.2 Input variables

Before training the BDT classifiers using the training data samples described above, potential input variables are selected. This process is described in this subsection.

First a preliminary list of input variables is produced and any poorly modeled variables are removed. From the remaining variables, those with the highest feature importance are selected to create the final list of input variables for each BDT classifier. In the `FastBDT` package, feature importance is calculated by comparing the AUC when including the variable in the training to the AUC when it is excluded [120]. This approach quantitatively measures the impact of an input variable on the final classifier.

For the 24 continuum suppression BDT classifiers, the ten input variables with the highest feature importance are used. For the 24 $B\bar{B}$ background suppression BDT classifiers, the top twelve variables are selected. The reason for including two additional variables in the $B\bar{B}$ background suppression BDT classifiers is that $B\bar{B}$ background events are more similar to signal events compared to continuum events based on event shape. As a result, the individual discriminating power of each input variable is lower. By allowing the $B\bar{B}$ background suppression BDT classifiers to use information from two additional input variables, a partial compensation for this effect occurs.

In producing the final list of input variables, none of the selected input variables are allowed to strongly correlate with ΔE , M_{bc} , or q^2 . Specifically, correlations between these variables and ΔE , M_{bc} , or q^2 are required to be lower than 0.7. Additionally, to remove redundancy in the provided information, correlations between input variables for signal events are required to be less than 0.95.

In the following, the input variables are described in detail, starting with the event shape variables and continuing with variables describing aspects of event kinematics and topology. Finally, the final lists of input variables is presented.

Event-shape variables

The variables describing the event shape are particularly relevant for suppressing continuum backgrounds. Signal and continuum events at B-factories exhibit distinct event shapes due to their production mechanisms. For continuum events, in the CM frame of the incoming electron-positron pair, the quark and antiquark are produced back-to-back. Due to their low masses compared to the collision energy, the produced quarks have high energies. The quarks hadronize due to color confinement, leading to jet-like behavior in the lab frame because of the boost from the CM frame. This effect is illustrated in Figure 6.2(a).

In contrast, for $B\bar{B}$ events, the $\Upsilon(4S)$ meson energy is just above the threshold for producing two B mesons. In the CM frame, the B mesons are nearly at rest. In the lab frame, this results in near-isotropic (spherical) angular distributions of the B meson decay products, as shown in Figure 6.2(b).

The distinct event shapes of $B\bar{B}$ and continuum events provide substantial discriminating power. However, this discriminating power is not limited to continuum events alone. Signal and $B\bar{B}$ background events also exhibit distinct event shapes, though the differences are more subtle compared to continuum events. Below, three categories of event-shape variables are described.

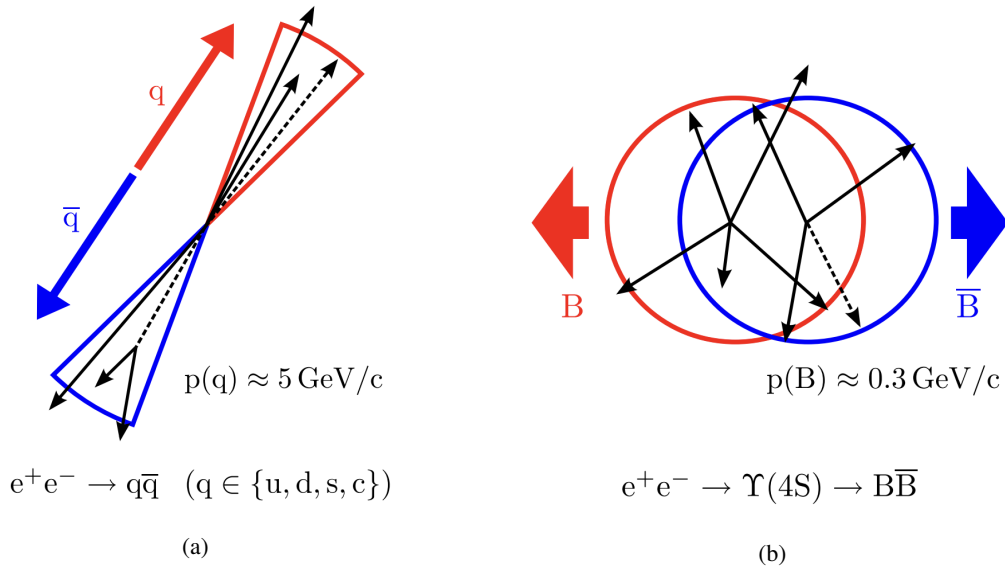


Figure 6.2: Illustration of (a) the jet-like continuum and (b) the spherical $B\bar{B}$ event shape in the lab frame, resulting from the high momenta of the produced quarks $p(q)$ and the very low momenta of the B mesons $p(B)$. Image from Ref. [121]

Thrust variables: The thrust axis (\vec{T}) is defined as the axis along which the sum of the projected momenta of the tracks (\vec{p}_i) in the event is maximal [122]:

$$\vec{T} = \max \left(\sum_i \vec{T} \cdot \vec{p}_i \right)$$

The thrust is the magnitude of this axis. For jet-like events, such as those from continuum processes, the thrust is expected to be close to 1. For isotropic events, the thrust value is distributed around 0.5. Distinct thrust axes are also defined for the tracks related to the signal B meson and those belonging to the ROE. The magnitudes of these axes are called $\text{Thrust}_{\text{sig}}$ and $\text{Thrust}_{\text{ROE}}$, respectively.

The cosine of the angle between the total thrust axis and the axis describing the sum of the momenta of all tracks in the event is referred to as $\cos \theta_{\text{Thrust}}$. This angle is expected to be close to 0 for jet-like events, such as continuum events, where the two axes should coincide. This results in a $\cos \theta_{\text{Thrust}}$ distribution peaking at 1. Conversely, for $B\bar{B}$ events, a uniform angular distribution is expected without structure in $\cos \theta_{\text{Thrust}}$. Additionally, $\cos \theta_z$ is defined as the cosine of the angle between the total thrust axis and the beam axis.

Fox-Wolfram moments: The Fox-Wolfram moments describe the momentum distribution in an N -particle system [123]. The ℓ -th order normalized moment is given by:

$$h_\ell^k = \frac{\sum_{i,j}^N |\vec{p}_i| |\vec{p}_j| P_\ell(\cos \theta_{ij})}{\sum_{i,j}^N |\vec{p}_i| |\vec{p}_j|}$$

where $|\vec{p}_i|$ is the momentum magnitude of the i -th particle, $P_\ell(\cos \theta_{ij})$ is the Legendre polynomial of degree ℓ , and θ_{ij} is the angle between the vectors \vec{p}_i and \vec{p}_j . In this work, the index $k = so$ indicates that particle i is related to the signal B meson, while particle j is part of the ROE. For $B\bar{B}$ events, the momentum vectors of the particles related to the signal B meson are largely uncorrelated with those of the ROE particles. In contrast, for continuum events, the momentum vectors of the reconstructed B meson and the ROE are correlated. Thus, the distributions of h_ℓ^k are expected to differ between continuum and $B\bar{B}$ events.

Cleo cones: The CLEO cones describe the momentum flow around the thrust axis of a particle system [124]. Nine cones (cc1–9) are defined centered around the total thrust axis, with each successive cone having an opening angle increasing in ten-degree steps. The momentum within each of these cones is measured. For $B\bar{B}$ events, the flight directions of the particles should be uncorrelated. In contrast, for jet-like continuum events, the flight directions of particles within the jet are correlated, resulting in different distributions for the CLEO cone variables.

Kinematic and topological variables

In addition to event-shape variables, there is another category of variables that focus on individual aspects of the reconstructed B meson candidate or the ROE. These variables describe the event topology or the kinematics of particles within the event. Some of these variables are effective in distinguishing signal events from both continuum and $B\bar{B}$ background events. Others are specifically designed to target differences between signal and the remaining $B\bar{B}$ events.

Some of these variables have already been introduced in previous reconstruction or selection steps. These include the χ^2 probability of the vertex fit to the signal pion and lepton candidates, $\cos \theta_{BY}$, θ_{miss} , the number of tracks, p_{ROE} , and $\cos \theta_\ell$. In addition, the cosine of the angle between the signal B meson momentum vector and the vector connecting its fitted vertex to the IP, both in the plane parallel and perpendicular to the beam axis, is included. These are defined as $\cos \theta_{\text{p,ip}}$ and $\cos \theta_{\text{p,ip}}^{\text{XY}}$, respectively. Another important discriminating variable is $\cos \theta_{W\ell}$, which is the cosine of the angle between the lepton's direction in the W boson frame and the W boson's direction in the B meson frame. In the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, the distribution of $\cos \theta_\nu$, defined as the cosine of the angle between the ρ meson direction in the B meson frame and the pion direction in the ρ meson rest frame, also provides a clear distinction between signal and background events.

Final list of input variables

After training the 48 BDT classifiers using the preliminary list of variables described above, the ten (twelve) variables with the highest feature importance for each continuum ($B\bar{B}$ background) suppression BDT classifier are selected.

The final input variables for each of the continuum suppression BDT classifiers for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes are listed in Figure 6.3(a) and 6.3(b). The relative importance of each input variable is shown. The “x” indicates that the variable was not selected for the relevant BDT classifier. The $\cos \theta_{\text{Thrust}}$ is among the most discriminating variables for suppressing continuum backgrounds across all q^2 bins, for both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes. For some variables, such as $\cos \theta_{W\ell}$ and h_{22}^{so} , there is a strong dependence of feature importance on q^2 . This observed behavior is one reason for training separate classifiers for different q^2 bins.

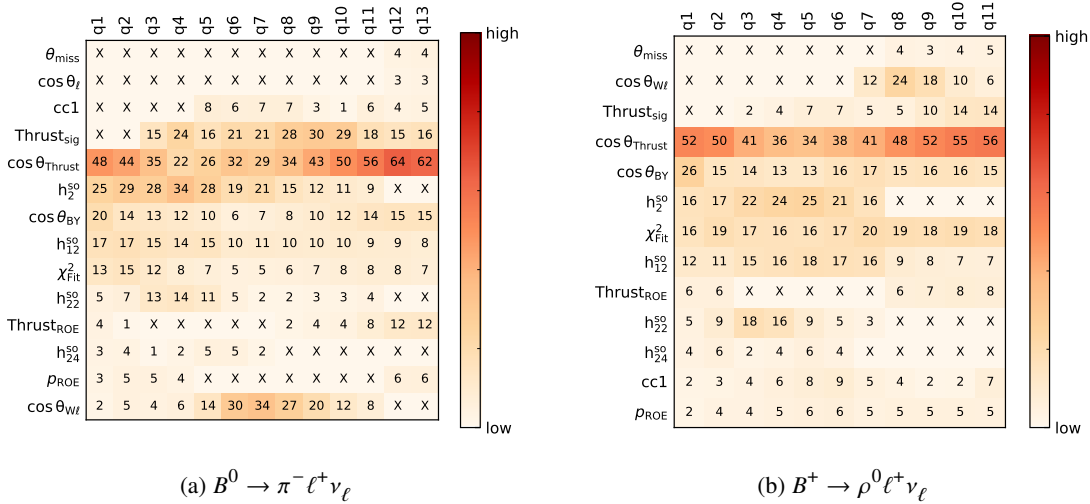


Figure 6.3: The input variables and their relative feature importance for the continuum suppression BDT classifier in each q^2 bin for the (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes.

Similarly, the final input variables for each of the $B\bar{B}$ background suppression BDT classifiers for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes are listed in Figure 6.4(a) and 6.4(b). Except for the lowest q^2 bins in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, where $\cos \theta_{BY}$ dominates, the greatest discriminating

power is provided by $\cos \theta_{BY}$, $\cos \theta_{W\ell}$, and the χ^2 vertex fit probability. In the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, the latter two variables exhibit the highest feature importance. The feature importance also shows a dependence on the q^2 bin for the $B\bar{B}$ background suppression classifiers. This reflects the difference in background composition across different q^2 bins. Variables that effectively distinguish signal from $B \rightarrow X_c \ell \nu_\ell$ events are important in intermediate q^2 bins, where $B \rightarrow X_c \ell \nu_\ell$ background is predominant. However, if these variables do not differentiate well between signal and $B \rightarrow X_u \ell \nu_\ell$ events, their importance decreases in high q^2 bins, where $B \rightarrow X_u \ell \nu_\ell$ events are more common.

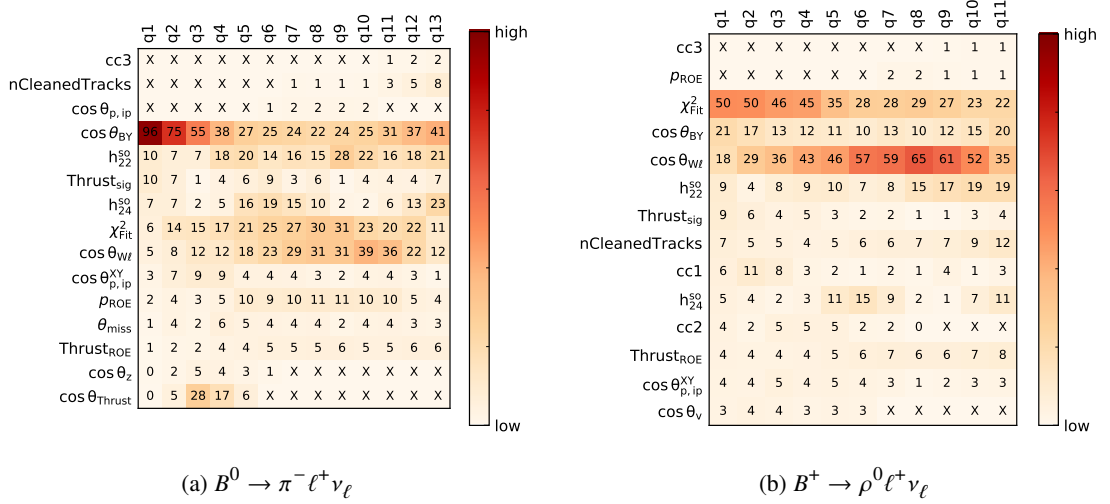


Figure 6.4: The input variables and their relative feature importance for the $B\bar{B}$ background suppression BDT classifier in each q^2 bin for the (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes.

6.2.3 Classifier validation and performance

After finalizing the list of input variables and training the classifiers, the absence of overfitting or underfitting needs to be verified. The validation of the classifiers and their performance is described in this subsection.

To validate the classifiers, the Kolmogorov-Smirnov test [125] is used to evaluate p-values, as described in Section 6.1. The p-values obtained separately for signal and background events are within the range of 0.05–0.95. Additionally, the ROC curves for the training and test samples are compared and good agreement is observed. Both tests indicate that no overfitting or underfitting is occurring.

Next, the performance of the BDT classifiers is evaluated by examining the ROC curves. The ROC curves for the continuum suppression BDT classifiers for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes are shown in Figure 6.5. The AUCs are all greater than 0.95, indicating that each classifier provides substantial discriminating power between signal and continuum events. The ROC curves for the $B\bar{B}$ background suppression BDT classifiers for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes are also shown in Figure 6.5. The AUCs are all greater than 0.77 but are lower than those of the continuum suppression BDT classifiers. The reduced discriminating power of the $B\bar{B}$ background suppression classifiers is expected due to the more similar event shapes of the backgrounds. This effect is further

enhanced by training these classifiers using a signal-enriched sample.

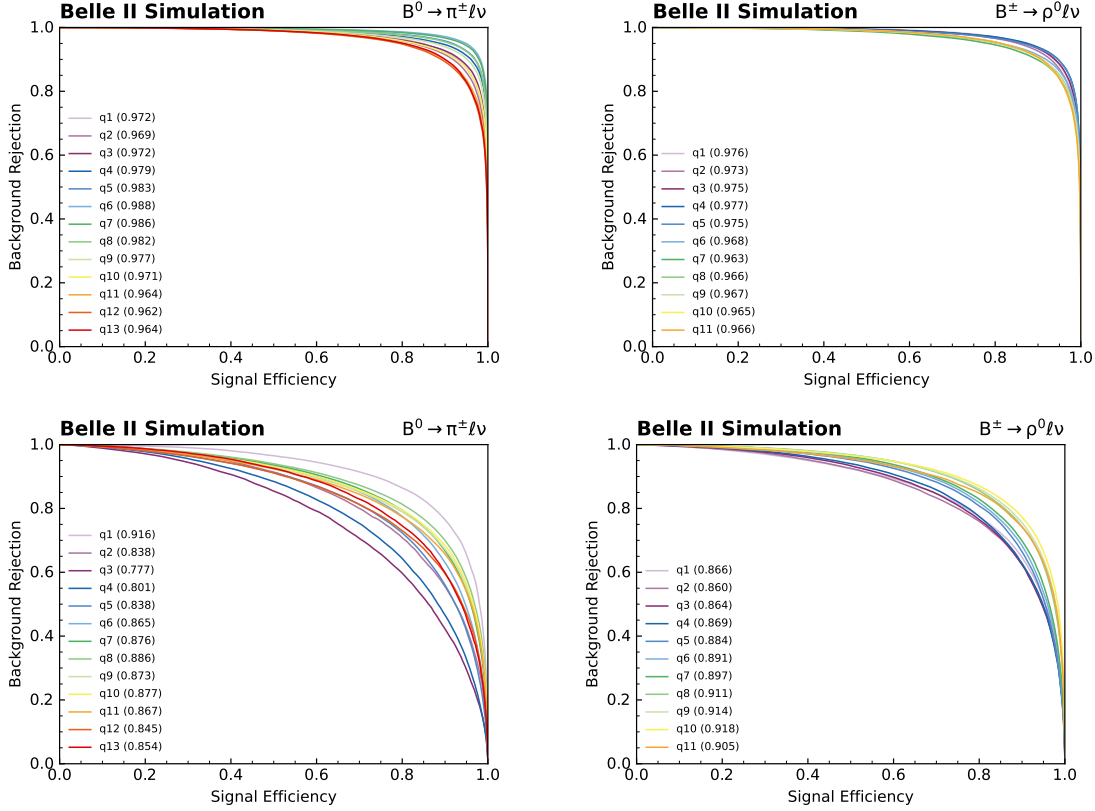


Figure 6.5: ROC curves of the continuum (top) and $B\bar{B}$ background (bottom) suppression BDT classifiers for the $B^0 \rightarrow \pi^+ \ell^- \nu$ (left) and $B^+ \rightarrow \rho^0 \ell^+ \nu$ (right) modes. The areas under the ROC curves are provided.

Further, the performance achieved by the classifiers is compared to that of a cut-based approach within each q^2 bin. To achieve the same level of background rejection with a cut-based approach, significant losses in signal selection efficiency are conceded. When using the three most discriminating variables— $\cos \theta_{BY}$, $\cos \theta_{Thrust}$, and the χ^2 vertex fit probability—the signal selection efficiencies are up to 30% lower. Thus, for the purpose of the work described in this thesis, the BDT classifiers outperform the cut-based approach.

6.2.4 Classifier selection

After applying the classifiers to the simulated and collision data samples, they are used to suppress backgrounds. The selections on the classifiers and their performance at suppressing backgrounds are discussed in this subsection.

The distributions of the continuum and $B\bar{B}$ background suppression classifiers for one example q^2 bin in simulated and collision data for the $B^0 \rightarrow \pi^+ \ell^- \nu$ mode are shown in Figure 6.6(a) and 6.6(b). The continuum suppression classifier distribution clearly peaks at 0 for continuum events and at 1 for signal events. In contrast, the $B\bar{B}$ background suppression classifier distribution shows more overlap between signal and $B\bar{B}$ events, with peaks towards 0 and 1 but less pronounced.

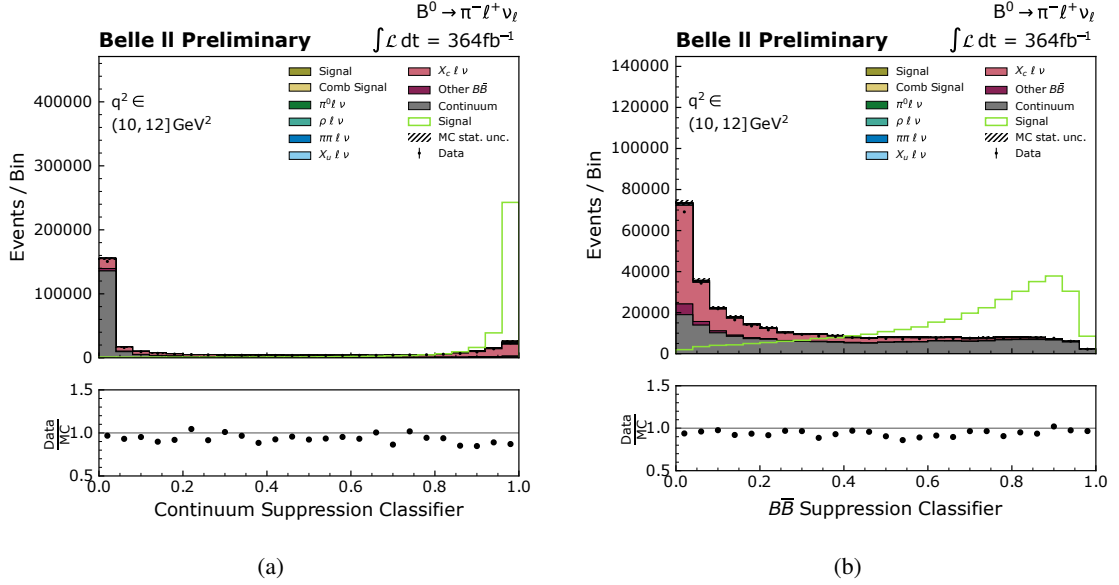


Figure 6.6: Distributions of the (a) continuum and (b) $B\bar{B}$ background suppression classifiers in the q_6 bin of the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode in simulated and collision data. The ratio between the entries in simulated and collision data is shown in the lower panel.

The backgrounds are suppressed by performing a two-dimensional selection on the continuum and $B\bar{B}$ background suppression BDT classifiers for each reconstructed q^2 bin. The reasons for using separate thresholds for each q^2 bin are similar to those for training separate classifiers. Firstly, it accounts for the varying contributions of continuum and $B\bar{B}$ background events to the total background. Secondly, selecting per reconstructed q^2 bin reduces the risk of introducing sharp features into the true q^2 spectrum, which could bias the measurement of $|V_{ub}|$ from form-factor fits.

Minimum thresholds on the classifiers ($C_{\text{out},1}$, $C_{\text{out},2}$) are selected by maximizing the figure of merit (FOM) predicted by simulation for each reconstructed q^2 bin. The FOM is a measure of statistical signal significance and is defined as:

$$\text{FOM}(C_{\text{out},1}, C_{\text{out},2}) = \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + N_{\text{bkg}}}},$$

where N_{sig} and N_{bkg} are the numbers of signal and background events in the q^2 bin when thresholds $C_{\text{out},1}$ and $C_{\text{out},2}$ are applied. Example plots of the FOM as a function of the $B\bar{B}$ background and continuum suppression classifiers for the q_1 and q_8 bin of the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode are shown in Figure 6.7. While the continuum classifier dominates at low q^2 where the continuum background is most dominant, the $B\bar{B}$ background classifier gains in importance in intermediate q^2 bins, where the contribution from the $B\bar{B}$ background grows.

$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events with classifiers above the thresholds listed in Table 6.1 are selected. In the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ ($B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$) mode, these selections retain 55% (48%) of the signal, but only 4% (1%) of the $B\bar{B}$ background, and 1% (0.6%) of the continuum events in simulation.

The agreement of the selection efficiency between simulated and collision data relies on correctly

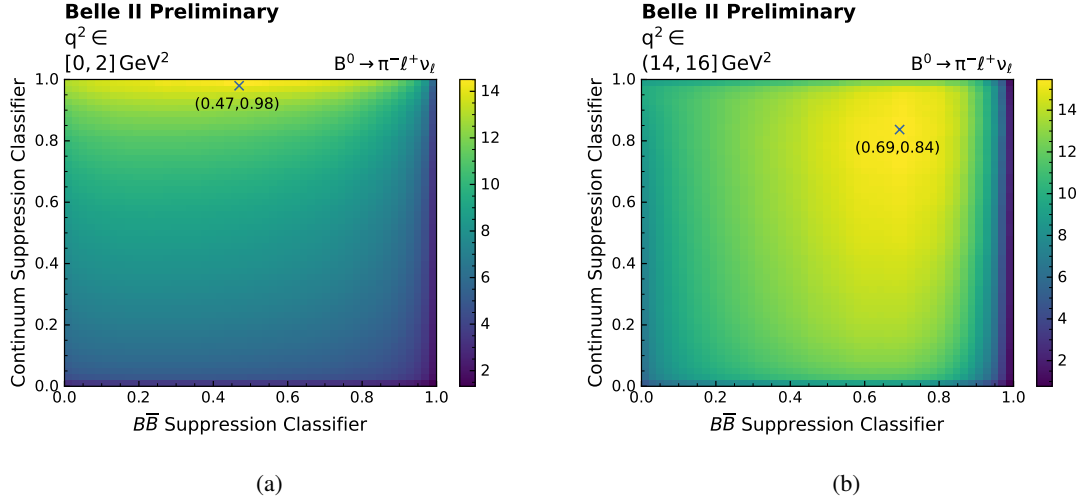


Figure 6.7: Plot of the FOM as a function of lower thresholds on the continuum and $B\bar{B}$ background suppression classifiers in the (a) $q1$ and (b) $q8$ bin of the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode from simulation. The thresholds corresponding to the maximum FOM values are marked with a cross.

modeled input variable distributions. The presence of any potentially concerning mismodeling is enhanced by selecting on the BDT classifiers. The agreement in the input variable distributions between simulated and collision data after applying the minimum thresholds listed in Table 6.1 is examined. Generally good agreement is observed, with p-values in the range 0.05–0.95. To account for residual uncertainties in the signal selection efficiencies the classifier distributions in simulated and collision data are compared (see Section 8.3.3).

6.3 Best candidate selection

After selecting on the BDT classifiers, different event multiplicities are accounted for using a best candidate selection discussed in this section. While the multiplicity for reconstructed $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events is 1.08, for reconstructed $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events it is 1.18. For $B\bar{B}$ background and continuum events, the multiplicities are larger. While in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, these are 1.13 and 1.29, respectively, in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, the multiplicities are 1.29 and 1.43.

In events with multiple candidates in either or both the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, one candidate is randomly selected and the others are discarded to avoid any bias. This best candidate selection ensures that a single event cannot contribute multiple times to either or both the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes.

By checking the agreement of the overall multiplicities between simulated and collision data, a correct modeling of the multiplicities is verified. This check is repeated for different data-taking periods and consistently good agreement is observed. This indicates a negligible uncertainty on the signal selection efficiency.

Table 6.1: Continuum and $B\bar{B}$ background classifier thresholds for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events across different reconstructed q^2 bins determined by maximizing the FOM.

q^2 bin	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$		$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	
	Continuum	$B\bar{B}$	Continuum	$B\bar{B}$
$q1$	0.975	0.467	0.983	0.647
$q2$	0.918	0.450	0.950	0.592
$q3$	0.902	0.432	0.945	0.647
$q4$	0.898	0.588	0.901	0.691
$q5$	0.912	0.612	0.901	0.758
$q6$	0.923	0.707	0.851	0.769
$q7$	0.898	0.667	0.829	0.824
$q8$	0.842	0.694	0.923	0.857
$q9$	0.822	0.633	0.934	0.857
$q10$	0.877	0.671	0.945	0.868
$q11$	0.898	0.593	0.950	0.846
$q12$	0.916	0.547		
$q13$	0.941	0.534		

6.4 Signal selection efficiencies

The best candidate selection completes the signal selection process. The performance of the selection, focusing on the total signal selection efficiencies, is described in this section.

In the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, the purity has increased to 1.4 signal events per 100 events. In the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, there are 1.6 signal events per 100 events. These results correspond to increases by factors of 21 and 54 compared to the samples before background suppression using the BDT classifiers. Overall, the purities have increased by factors of 10^4 in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode and 10^5 in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode. The total signal selection efficiencies are 15% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events and 7% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. The reduced efficiency in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, compared to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, is necessary to achieve similar levels of purity in both modes.

To ensure no sharp features were introduced into the q^2 spectrum, the signal efficiencies are examined in bins of true q^2 . The signal selection efficiency in a given true q^2 bin is defined as the number of signal events in that bin after all selections divided by the number of initial signal events in the same bin. For simulated signal events divided into true q^2 bins based on the boundaries given in Section 5.5, the signal selection efficiencies ϵ range from 9% to 19% in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode and from 3% to 9% in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode. The signal efficiencies in each q^2 bin, including statistical and systematic uncertainties, are presented in Table 6.2. These efficiencies are also shown as functions of q^2 in Figure 6.8.

The efficiencies presented in Table 6.2 include contributions from both true and combinatorial signal events. These are labeled as “total signal” in Figure 6.8. Efficiencies related to true signal events only are also shown. Overall, nearly flat efficiency distributions are observed, confirming that the q^2 spectra remain largely undistorted by the selections. Combinatorial signal events contribute mainly at high q^2 , as indicated by the increase in total signal selection efficiencies compared to true signal selection efficiencies in those regions.

Table 6.2: Signal efficiencies in % for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events in each true q^2 bin from simulation after all selections have been performed. The signal samples include true and combinatorial signal events. The first uncertainty is statistical and the second is systematic.

q^2 bin	Signal efficiency [%]	
	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
$q1$	$9.90 \pm 0.09 \pm 0.06$	$3.812 \pm 0.058 \pm 0.017$
$q2$	$13.15 \pm 0.11 \pm 0.08$	$5.821 \pm 0.067 \pm 0.038$
$q3$	$16.03 \pm 0.12 \pm 0.10$	$6.989 \pm 0.069 \pm 0.049$
$q4$	$16.67 \pm 0.12 \pm 0.10$	$7.951 \pm 0.072 \pm 0.060$
$q5$	$16.76 \pm 0.13 \pm 0.10$	$8.259 \pm 0.071 \pm 0.069$
$q6$	$16.96 \pm 0.13 \pm 0.09$	$8.246 \pm 0.070 \pm 0.068$
$q7$	$17.39 \pm 0.13 \pm 0.08$	$7.465 \pm 0.067 \pm 0.066$
$q8$	$17.48 \pm 0.14 \pm 0.08$	$6.353 \pm 0.064 \pm 0.060$
$q9$	$18.35 \pm 0.15 \pm 0.09$	$6.442 \pm 0.071 \pm 0.065$
$q10$	$18.34 \pm 0.16 \pm 0.12$	$6.517 \pm 0.088 \pm 0.089$
$q11$	$18.03 \pm 0.17 \pm 0.16$	
$q12$	$16.70 \pm 0.19 \pm 0.25$	
$q13$	$14.27 \pm 0.27 \pm 0.28$	

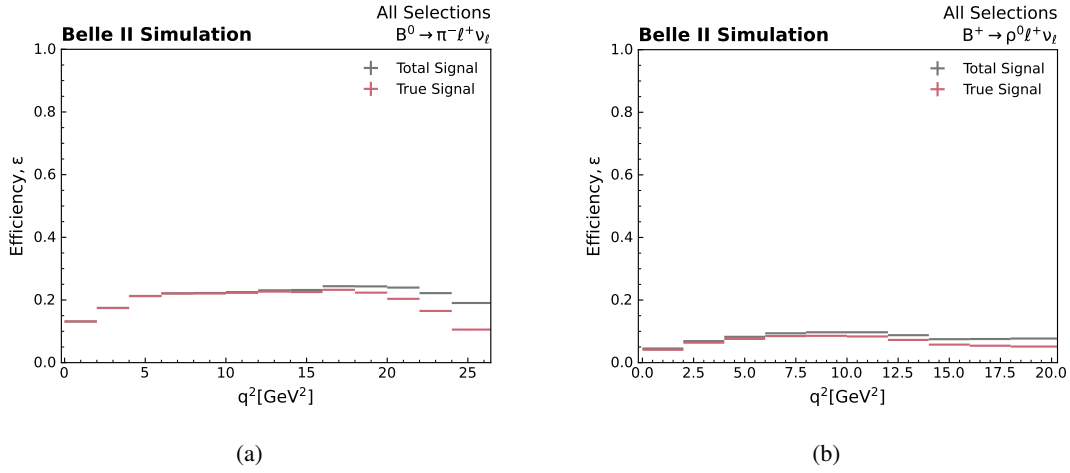


Figure 6.8: Signal efficiencies as a function of true q^2 for (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal events. Efficiencies are shown for the sum of true and combinatorial signal events (labeled as total signal) and separately for true signal events.

Statistical uncertainties on the efficiencies are calculated assuming the number of selected events follows Poisson distributions. Systematic uncertainties include those due to lepton and pion PID performance. For the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, uncertainties also arise from potential mismodeling of the ρ meson lineshape. Additionally, systematic uncertainties include contributions from mismodeling of the BDT classifiers, as discussed in Section 6.2.4. Other uncertainties account for bin migrations due to limited simulated sample sizes or residual form-factor dependence. The origin and evaluation of

these uncertainties are described in Chapter [8.3.3](#).

Branching fraction measurement

After signal selection and background suppression, the distributions of ΔE , M_{bc} , and q^2 for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes in simulated and collision data are shown in Figure 7.1. The lower panels provide a comparison of the data samples, illustrating the difference between collision and simulated data divided by the combined uncertainty. Distributions of ΔE and M_{bc} within each reconstructed q^2 bin are provided in Appendix B.1.

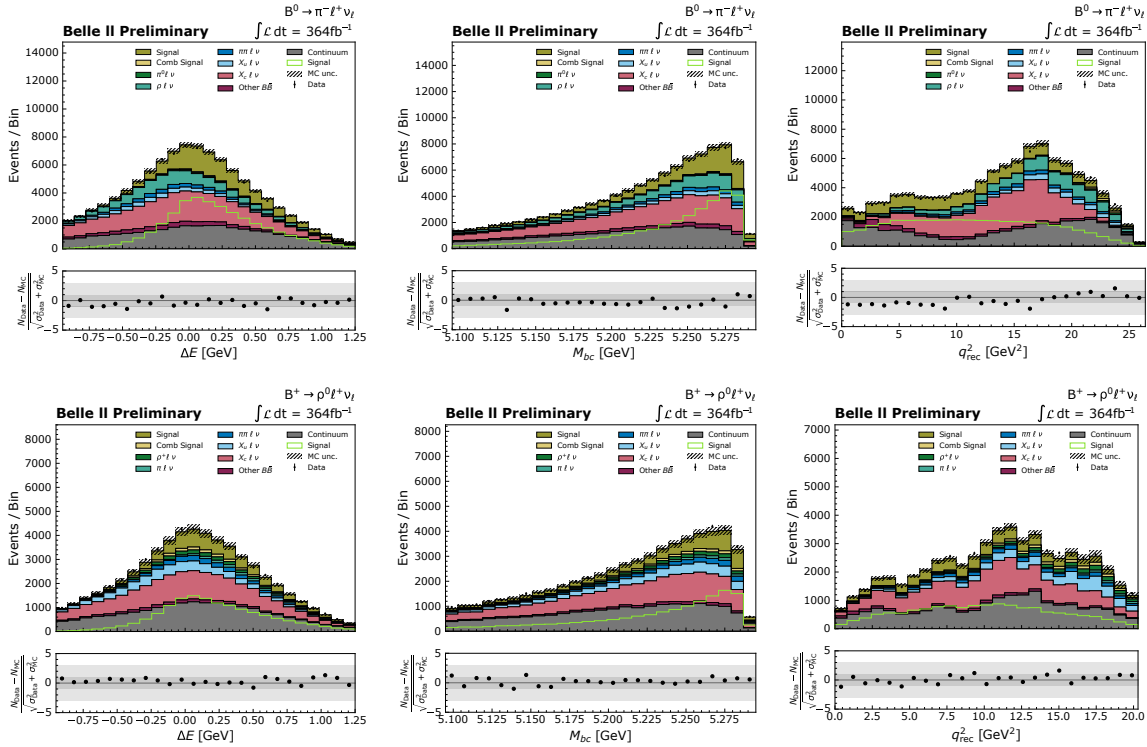


Figure 7.1: Distributions of ΔE (left), M_{bc} (middle), and q^2 (right) in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (top) and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (bottom) modes for simulated and collision data after signal selection and background suppression. The hatched areas represent statistical and systematic uncertainties on the simulated distributions, detailed in Chapter 8. The lower panels show the difference between collision and simulated data divided by the combined uncertainty.

In this chapter the development of a signal extraction method to simultaneously measure $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal events in q^2 bins is described. In addition, the application of the developed method in the determination of the partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, which are used to extract $|V_{ub}|$, is discussed. First, the methods for parameter estimation through fits are described in Section 7.1. The selected implementation as a simultaneous extended binned maximum likelihood fit is explained in Section 7.2. This further includes a discussion of the likelihood function in Section 7.3. Next, the fit components (templates), parameters and variables are discussed in Section 7.4, highlighting the treatment of the continuum and signal templates.

In Section 7.5, the results obtained from fitting the three-dimensional distribution of ΔE , M_{bc} , and reconstructed q^2 are presented. Finally, the determination of the partial and total branching fractions is described in Section 7.6.

7.1 Parameter estimation through fits

This section introduces methods for parameter estimation, which aims to determine the values of parameters within a theoretical model that best explain the observed data. Two primary methods for parameter estimation through fits are chi-square and likelihood-based fits, each with distinct advantages and suitable applications.

Chi-square fits: The chi-square statistic is calculated as the sum of the squared differences between observed and expected values, normalized by the variance of the observed values. The best-fit parameters are those that minimize the chi-square statistic, indicating the smallest relative deviation between the model predictions and the observed data. Chi-square fits are most appropriate when the uncertainties in the data are well-understood and can be approximated by a Gaussian distribution.

Likelihood-based fits: Likelihood-based fits are more flexible, as they can accommodate data with non-Gaussian error distributions. They are especially useful for complex models or datasets with varying sizes. Maximum likelihood estimation techniques construct likelihood functions that quantify how well the model's predicted distributions match the observed data.

7.2 Maximum likelihood fit method

Given the flexibility of likelihood-based fits in handling varying dataset sizes and uncertainty estimations, a maximum likelihood fit is developed for the work discussed in this thesis to extract the number of signal events. Different approaches exist for implementing maximum likelihood fits, and in this section the selected setup is described.

First the choice between binned and unbinned maximum likelihood estimation is considered. Unbinned maximum likelihood estimation utilizes each data point individually without grouping the data points into bins. While this approach makes full use of the available data, it can be computationally intensive for large datasets. In contrast, binned maximum likelihood estimation groups the data into bins, which reduces computational complexity and is making it particularly advantageous for large datasets. Due to these reasons, a binned approach is selected, though the choice of bin size and range may introduce potential biases.

The predicted distributions used in the maximum likelihood estimation are derived from probability density functions associated with each model component, known as templates. Each template is characterized by a yield, representing the expected number of observations or entries. The fit templates are discussed in more detail in Section 7.4. During the fitting process, the yields of specific components are either fixed at predetermined values or they can be allowed to vary within a defined range. This flexibility enables the fit to incorporate prior knowledge or constraints on certain components while exploring the parameter space for others.

The outcome of a maximum likelihood fit is a set of yields, one for each component. These yields correspond to the point of maximum likelihood, reflecting the best agreement between the model and the observed data under the constraint that the total yield (the sum of all component yields) remains constant. In this work, the total number of events is not known a priori and must be determined as part of the fitting process. An extended binned maximum likelihood fit is implemented, which allows the total yield to vary within constraints derived from Poisson statistics, accounting for statistical fluctuations in the counts of observed events.

Finally, to constrain the cross-feed components between the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, a simultaneous fit is employed. With this approach common parameters across both datasets can be extracted, while also accommodating dataset-specific parameters.

7.3 Construction of the likelihood function

Next, a likelihood function is constructed to be maximized during signal extraction. this process and the resulting likelihood function is described in this section.

For a binned maximum likelihood fit, the likelihood function is based on the counts (frequencies) of data points within each bin. Therefore, to define the likelihood function for the fit, first the expected number of events for component p in each bin l , denoted as $v_{l,p}^{\text{exp}}$, is defined. Ideally, $v_{l,p}^{\text{exp}}$ matches the number of events for component p in bin l from simulation, $v_{l,p}^{\text{sim}}$. However, by introducing per-component scale factors a_p the fit may adapt the model based on observed data:

$$v_{l,p}^{\text{exp}} = a_p \times v_{l,p}^{\text{sim}}$$

These scale factors describe the ratio of the measured to simulated yield of each component. The total number of events in bin l is given by:

$$v_l^{\text{exp}} = \sum_p a_p v_{l,p}^{\text{sim}}.$$

The likelihood function to be maximized is defined as:

$$-2 \ln \mathcal{L}(\vec{a}) = -2 \ln \prod_l \text{Poisson}(N_l, v_l^{\text{exp}}), \quad (7.1)$$

where N_l is the observed number of events in bin l . The fit parameters are represented by the vector \vec{a} , which contains the scale factors a_p for each fit component. The Poisson distribution term in the

likelihood function is given by:

$$\text{Poisson}(N_l, v_l^{\text{exp}}) = \frac{(v_l^{\text{exp}})^{N_l} e^{-v_l^{\text{exp}}}}{N_l!}.$$

For a simultaneous fit to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, the likelihood function in Equation 7.1 is extended to include individual likelihoods for each mode: $\mathcal{L}_{\text{comb}} = \mathcal{L}_{\pi\ell\nu} \times \mathcal{L}_{\rho\ell\nu}$. The total likelihood function is therefore:

$$-2 \ln \mathcal{L}_{\text{comb}}(\vec{a}_{\text{comb}}) = -2 \ln \left[\prod_l \text{Poisson}(N_l^{\pi\ell\nu}, v_l^{\text{exp}, \pi\ell\nu}) \times \prod_m \text{Poisson}(N_m^{\rho\ell\nu}, v_m^{\text{exp}, \rho\ell\nu}) \right], \quad (7.2)$$

where l and m denote the bins in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, respectively. The vector \vec{a}_{comb} contains scale factors for components in both modes, including those shared between them.

In the work described in this thesis, the likelihood function in Equation 7.2 is numerically maximized using the `iminuit` package [126].

7.4 Fit templates, parameters and variables

To extract the signal yields of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, a simultaneous extended binned maximum likelihood fit is performed. In this section, the templates and parameters used in the fit, as well as the fit variables and binning choices, are described. A special focus is laid on the treatment of the continuum template in Section 7.4.1 and the composition of the signal templates in Section 7.4.2.

Fit templates

First, the model components associated with background event categories are defined. The background in each signal mode is considered separately and categorized into four components: $B \rightarrow X_c \ell \nu_\ell$, $B \rightarrow X_u \ell \nu_\ell$, *other* $B\bar{B}$, and continuum events, as detailed in Section 5.4. The fit templates are generated from simulated samples corresponding to 3 ab^{-1} of $B\bar{B}$ events and 1 ab^{-1} of continuum events. Due to the limited size of these samples, there are uncertainties in the template shapes that must be considered in the fit.

The simulated samples are modeled and corrected as outlined in Chapter 4. Continuum templates are derived from simulated continuum samples, which are then reweighted to correct for discrepancies between simulated and off-resonance data. The process of reweighting continuum templates is described in Section 7.4.1.

In addition to the eight background templates, there is one independent total signal template for each true q^2 bin in each mode. By dividing the signal samples into true- q^2 templates, branching fractions can be derived within true q^2 bins, which are essential for extracting $|V_{ub}|$ from differential measurements. This binning approach in true instead of reconstructed q^2 also reduces the impact of uncertainties in q^2 resolution modeling.

The $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ signal templates include contributions from true, combinatorial, and isospin-conjugate signal events, and cross-feed events ($B \rightarrow \pi \ell \nu_\ell$ events reconstructed in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$

mode). Similarly, the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal templates consist of true, combinatorial, and isospin-conjugate signal events, as well as cross-feed events ($B \rightarrow \rho \ell \nu_\ell$ events reconstructed in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode). Due to these cross-feed events, the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes are interconnected through shared signal templates. The construction and composition of these signal templates are discussed in more detail in Section 7.4.2.

Fit parameters

Next, the fit parameters are defined, which are estimated during the maximization of the likelihood function given by Equation 7.2. These parameters are elements of the vector \vec{a}_{comb} . Since distinct background templates are constructed for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, the fit parameters include separate scale factors $a_p^{(\pi/\rho)}$ for each mode $((\pi/\rho)\ell\nu)$ and each background component p .

Additionally, for each true q^2 bin i and each mode, there are signal scale factors a_i^π and a_i^ρ . Because the cross-feed events in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode are included in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal template, and vice versa, the signal scale factors are shared between the two modes.

The scale factors are generally unconstrained. However, for the two continuum background templates, Gaussian penalty factors are introduced to constrain the continuum yields to match the entries of the scaled off-resonance dataset. Table 7.1 summarizes the 31 templates and corresponding scale factors.

Table 7.1: Summary of templates and scale factors used in the fit for background and signal components. For each true q^2 bin i and each signal mode, there is one signal scale factor: a_i^π for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and a_i^ρ for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. All fit parameters are free, except for a_{cont}^π and a_{cont}^ρ , which are constrained by off-resonance data.

Component		Reconstructed mode	
		$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
Signal:	True signal	a_i^π	a_i^ρ
	Combinatorial signal	a_i^π	a_i^ρ
	Isospin-conjugate signal	a_i^π	a_i^ρ
	Cross-feed	a_i^ρ	a_i^π
Background:	$B \rightarrow X_u \ell \nu_\ell$	$a_{X_u \ell \nu}^\pi$	$a_{X_u \ell \nu}^\rho$
	$B \rightarrow X_c \ell \nu_\ell$	$a_{X_c \ell \nu}^\pi$	$a_{X_c \ell \nu}^\rho$
	Other $B\bar{B}$	$a_{B\bar{B}}^\pi$	$a_{B\bar{B}}^\rho$
	Continuum	a_{cont}^π	a_{cont}^ρ

Fit variables

Lastly, ΔE and M_{bc} are selected as fit variables, due to their significant discriminating power, as outlined in Section 5.5. These variables effectively separate signal from background and differentiate between various background components. While ΔE and M_{bc} distributions vary slightly between different true q^2 signal templates, they do not sufficiently differentiate between these templates.

To enhance the fit's ability to distinguish between different true q^2 signal templates, the reconstructed q^2 is reconstructed as an additional fit variable. As demonstrated by the bin migration matrices in Figure 5.8, there is a strong correlation between true and reconstructed q^2 , which helps the fit to discriminate between signal templates corresponding to distinct true q^2 bins. Moreover, the reconstructed q^2 distribution provides additional discrimination among the backgrounds, as seen in Figure 7.1.

This approach of utilizing true q^2 templates to fit the reconstructed q^2 distribution is novel compared to the traditional method, which employs reconstructed q^2 templates along with unfolding to true q^2 bins. A key advantage of this novel method is that it inherently accounts for resolution effects that may cause events generated in one q^2 bin to be reconstructed in another. This approach includes q^2 resolution uncertainties directly in the fit uncertainty, avoiding the need for additional uncertainty propagation associated with standard unfolding procedures.

Thus, a fit to the binned three-dimensional distributions of ΔE , M_{bc} , and reconstructed q^2 is performed. The reconstructed q^2 distribution is binned according to the same q^2 bins used for the true q^2 templates, as detailed in Section 5.5. The two-dimensional fit region in ΔE and M_{bc} , defined in Section 5.5, is divided into 20 bins with the following boundaries:

- Five bins in ΔE : $[-0.95, -0.35, -0.15, 0.15, 0.40, 1.25]$
- Four bins in M_{bc} : $[5.095, 5.175, 5.250, 5.295]$

The binning is initially chosen by dividing the ΔE and M_{bc} ranges into equidistant bins. A finer binning is then applied in signal-enriched regions to create bins with high purity for signal definition. Similarly, bin boundaries are adjusted to achieve very low purity in bins for effective background constraint. Although bin boundaries are tuned manually, the impact of binning optimization is evaluated. The significance of the extracted signal templates (defined as the fitted value of the related scale factor divided by its uncertainty) does not strongly depend on the chosen binning. To avoid potential biases and sensitivity to detector or resolution mismodeling, the binning is not further optimized. Thus, there are $13 (q^2) \times 5 (\Delta E) \times 4 (M_{bc}) = 260$ bins in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, and $10 (q^2) \times 5 (\Delta E) \times 4 (M_{bc}) = 200$ bins in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, resulting in a total of 460 bins.

7.4.1 Continuum template: Reweighting

Before performing the fit to extract the signal scale factors as outlined above, the continuum templates need to be constructed. This process is described in this subsection.

Ideally, the off-resonance data would be used directly in the fit, after correcting it as described in Section 4.3.3, since it should closely represent the continuum background component in on-resonance data. However, due to the small size of the off-resonance sample, particularly after suppressing continuum events, using it directly to construct fit templates leads to significant statistical fluctuations in individual bins.

Instead, the simulated continuum events are reweighted using off-resonance data to create a reweighted simulated sample to use as the continuum template during signal extraction. Reweighting reduces the dimensionality of the fit binning described above by using fewer bins for the reweighting process. This method incorporates shape constraints from off-resonance data without the large statistical fluctuations encountered when using the data directly in the fit.

To produce the weights, first the reconstructed q^2 shapes of simulated continuum events are compared to those of off-resonance data. Similar normalization differences are present in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and

$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, with a ratio of 1.2 between the number of events in off-resonance and simulated data. After correcting the simulated continuum data for total normalization, the agreement between simulated continuum and off-resonance data in the three signal extraction variables: ΔE , M_{bc} , and reconstructed q^2 , is evaluated. The distributions for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes are shown in Figure 7.2 using the fit binning described above.

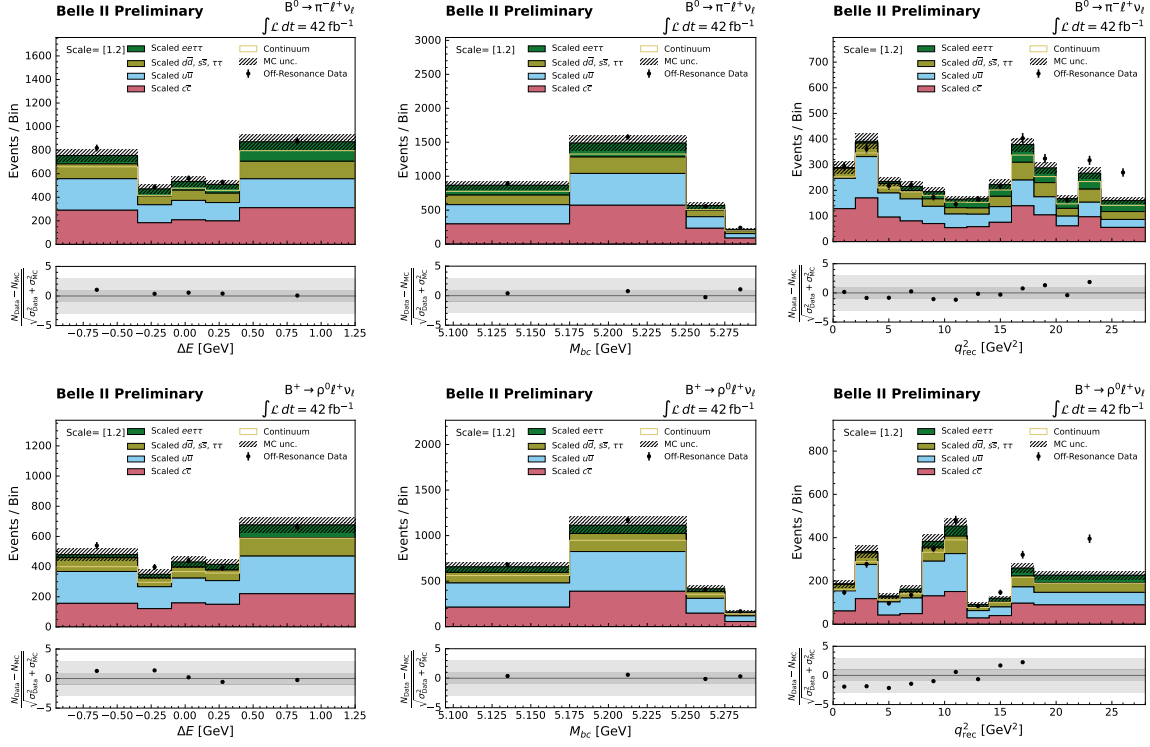


Figure 7.2: Distributions of ΔE (left), M_{bc} (middle), and q^2 (right) in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (top) and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (bottom) modes, comparing simulated continuum and off-resonance data after scaling the simulated continuum samples to match the number of events in off-resonance data (scale factor of 1.2). The hatched areas represent statistical and systematic uncertainties on the simulated distributions, as described in Chapter 8. The lower panels show the difference between collision and simulated data divided by the combined uncertainty. The bright green line represents the simulated distribution prior to reweighting.

While the shape agreement in M_{bc} appears reasonable, there is a potential shift in ΔE in the simulated continuum data relative to off-resonance data. In addition, the reconstructed q^2 distributions in simulated continuum data do not closely resemble those in off-resonance data, particularly at high reconstructed q^2 . The excess of events in off-resonance data at high reconstructed q^2 is investigated by examining the characteristics of these events. First, the samples are split by lepton flavor and charge, where it is found that the excess is predominantly in the electron sample. The excess is also observed at low missing momentum angles and in the forward and backward endcaps. Additionally, event charge and cluster distributions align with those of two-photon processes with high multiplicity.

The excess may be due to two-photon processes simulated with insufficient cross sections or mismodeled low-multiplicity backgrounds with more tracks in data than in simulation. Given the relatively well-known cross sections for two-photon processes and the agreement in the number of

tracks between simulated and collision data, it is more likely that the excess arises from two-photon processes not included in the simulation. These include nonresonant processes, such as $ee \rightarrow eeX$ where X is a hadronic system with multiple final-state particles. They can mimic signal events and contribute to observed excesses at high q^2 when high-momentum electrons and low-momentum particles within X are reconstructed as signal lepton and hadron candidates.

These residual differences are corrected by adjusting the simulated two-dimensional reconstructed q^2 and ΔE spectra using bin-by-bin event weights, defined as the ratio of off-resonance events to simulated continuum events in each bin. The reconstructed q^2 distribution is binned using the fit binning, while ΔE is split into low and high ΔE regions based on the fit binning ($[-0.95, -0.1]$ and $[-0.1, 1.25]$). The choice of two ΔE bins allows for potential shifts while maintaining a sufficient number of off-resonance events for reweighting. The reweighted distributions of ΔE and M_{bc} in each reconstructed q^2 bin for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes are provided in the Appendix A.4. Any remaining discrepancies in ΔE shapes are accounted for by the uncertainty assigned to the continuum template, as described in Section 8.3.8.

This method of reweighting the distributions of reconstructed q^2 and ΔE without considering M_{bc} assumes that the difference between off-resonance and simulated continuum data is independent of M_{bc} . This assumption is validated by examining p-values from Kolmogorov-Smirnov tests [125], comparing M_{bc} distributions in the reweighted simulated continuum and off-resonance data. The obtained p-values agree with consistent distributions, justifying the use of the reweighted continuum samples as the continuum templates for signal extraction.

7.4.2 Signal templates: Composition

In addition to describing the continuum templates in detail, the composition of the signal templates is discussed in this subsection. The simulated signal samples are categorized into true q^2 bins, treating each bin as an independent signal template. These templates include true signal, combinatorial signal, isospin-conjugate signal, and cross-feed events. First, the combination of true and combinatorial signal components is addressed, before incorporating isospin-conjugate and cross-feed components.

Combinatorial signal events occur at the same rate as true signal events due to the same underlying branching fraction. Thus, the number of combinatorial signal events scales proportionally with true signal events and can be described with a single scale factor. Figure 7.3(a) and 7.3(b) show the simulated relative ratio of true to combinatorial signal events within each true q^2 bin for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, respectively. These ratios align with observations from Section 6.4, where an increased signal efficiency at high q^2 was observed, due to combinatorial signal events. The fitted yields of these combined true and combinatorial signal events are used to determine the partial branching fractions for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ in Section 7.6.

In addition to true and combinatorial signal components, the total signal templates also include isospin-conjugate and cross-feed components. The isospin-conjugate signal component can be included to constrain these events during signal extraction under the assumption of isospin symmetry:

$$\mathcal{B}(B^0 \rightarrow \pi^- \ell^+ \nu_\ell) = 2\tau_0/\tau_+ \times \mathcal{B}(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell), \quad (7.3a)$$

$$\mathcal{B}(B^0 \rightarrow \rho^- \ell^+ \nu_\ell) = 2\tau_0/\tau_+ \times \mathcal{B}(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell), \quad (7.3b)$$

where $\tau_+/\tau_0 = 1.076 \pm 0.004$ is the B meson lifetime ratio [37]. The decay rates of charged and neutral $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (and $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$) are thus proportional, justifying the introduction of a single scale

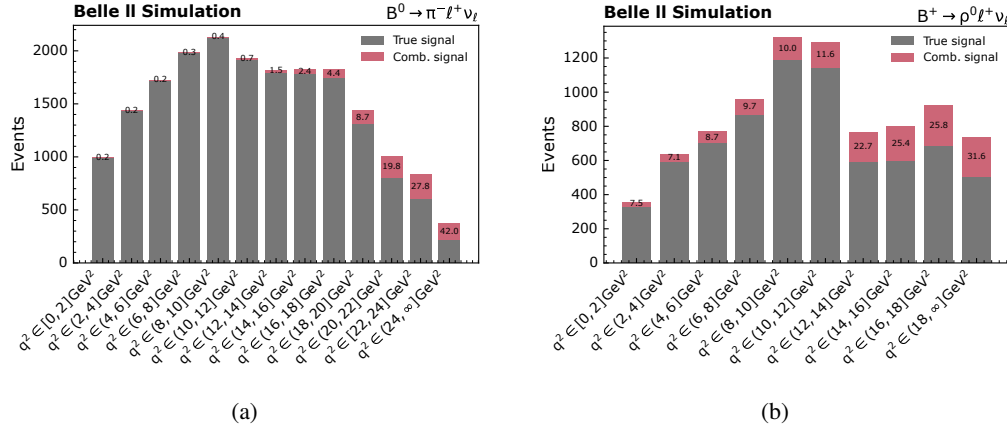


Figure 7.3: The composition of the total signal component in each true q^2 bin for (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal events. The percentage represents the contribution of the combinatorial signal to the total signal events in that bin.

factor for these events. The uncertainties due to the assumption of isospin symmetry in the creation of the signal templates are discussed in Section 8.3.5.

Similar considerations apply to integrating cross-feed events, which consist of two components, into the total signal template. The first component comprises $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events reconstructed as $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events (termed “true cross-feed events”), and vice versa for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal template. The number of these events scales with true signal events, similar to combinatorial signal events. The second component includes “cross-isospin events,” where $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ ($B^0 \rightarrow \rho^- \ell^+ \nu_\ell$) events are reconstructed as $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ ($B^0 \rightarrow \pi^- \ell^+ \nu_\ell$) events. The same isospin symmetry arguments are applied to assign the true-signal scale factor to these events. Including these samples in the signal templates allows the fit to constrain these backgrounds.

Figure 7.4(a) and 7.4(b) show the relative contributions of each component to the total number of events within each true q^2 bin for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal templates, respectively. For both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, the relative contributions of cross-feed, cross-isospin, and isospin-conjugate signal events increase with higher true q^2 values. The contribution from cross-feed and cross-isospin events is notably larger for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ than for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$. For most true q^2 bins, more than 50% of events in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal templates are reconstructed as $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events, indicating a significant total cross-feed component. These events often involve missed pions during reconstruction. The cross-isospin component is approximately twice the size of the true cross-feed component, consistent with the relations in Equations 7.3. The substantial cross-feed component underscores the importance of constraining cross-feed backgrounds by simultaneously reconstructing $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events.

Since the partial branching fractions are extracted from the combined samples of true and combinatorial signal events, attention is owed to the relative contributions of the combined samples to each total signal template. These contributions are referred to as signal strengths. Table 7.2 provides the simulated signal strengths for each true q^2 bin for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, including statistical and systematic uncertainties. For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$, the signal strength ranges from 69% to 99%, while for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, it ranges from 23% to 57%.

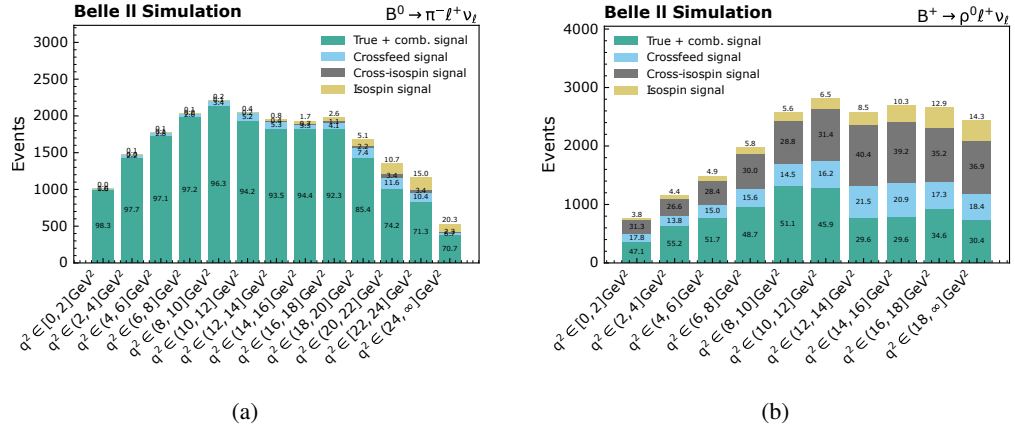


Figure 7.4: The composition of events scaling with the number of true signal events in each true q^2 bin for (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal events. The percentage indicates the contribution of each component to the total number of events scaling with true signal in that bin.

Statistical uncertainties on the signal strengths reflect the underlying Poisson distributions. Systematic uncertainties encompass lepton and pion PID performance, interference effects on the ρ meson lineshape for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, tracking efficiency, signal form factors, shape uncertainties due to limited simulated sample sizes leading to bin migrations, and isospin symmetry assumptions. The contribution of each uncertainty source to the total uncertainty budget is discussed in the following chapter.

Table 7.2: Signal strengths in % for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events in each true q^2 bin from simulation. The signal samples include true and combinatorial signal events. Statistical uncertainties are listed first, followed by systematic uncertainties.

q^2 bin	Signal strength [%]	
	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
$q1$	$98.3 \pm 0.43 \pm 0.01$	$49.8 \pm 0.33 \pm 0.09$
$q2$	$97.1 \pm 0.37 \pm 0.01$	$56.7 \pm 0.29 \pm 0.08$
$q3$	$96.8 \pm 0.34 \pm 0.01$	$56.5 \pm 0.25 \pm 0.07$
$q4$	$95.9 \pm 0.33 \pm 0.01$	$54.6 \pm 0.21 \pm 0.05$
$q5$	$94.6 \pm 0.32 \pm 0.02$	$49.4 \pm 0.18 \pm 0.04$
$q6$	$93.9 \pm 0.32 \pm 0.02$	$45.2 \pm 0.16 \pm 0.04$
$q7$	$92.5 \pm 0.31 \pm 0.02$	$38.3 \pm 0.13 \pm 0.03$
$q8$	$91.8 \pm 0.31 \pm 0.02$	$31.2 \pm 0.12 \pm 0.03$
$q9$	$90.1 \pm 0.31 \pm 0.02$	$28.1 \pm 0.11 \pm 0.03$
$q10$	$86.0 \pm 0.31 \pm 0.02$	$23.7 \pm 0.11 \pm 0.03$
$q11$	$79.0 \pm 0.31 \pm 0.03$	
$q12$	$72.8 \pm 0.34 \pm 0.04$	
$q13$	$69.7 \pm 0.51 \pm 0.06$	

7.5 Fit validation, results and stability

This section describes the validation of the fit, the obtained results and the methods used to test the fit stability.

Fit validation

Before applying the fit to collision data, potential biases need to be identified and the estimation of uncertainties validated. In the work discussed in this thesis, a “toy MC method” is used. This simulation-based approach generates a large number of pseudo-experiments, or toys, from a specified model or distribution. To validate the fit, three-dimensional toy distributions are created for each template by sampling event counts from the nominal templates, assuming these counts follow Poisson distributions. Next, multiple total toy distributions are created by summing the toy distributions of each template.

The fit is then repeatedly performed to each of these toy distributions using the nominal templates and the resulting distributions of each fit parameter are recorded. For each fit parameter, a pull is defined as the difference between the expected and fitted result divided by the fit uncertainty. As outlined in Section 7.3, the scale factors expected from simulation are unity by construction. By examining the means and standard deviations of these pull distributions, bias can be detected and the estimation of the uncertainties validated. The means and standard deviations are consistent with zero and unity, respectively, indicating unbiased results with correctly estimated uncertainties.

Fit results

Now, the fit to collision data is performed and the fit results are described in the following. Figure 7.5 shows the two-dimensional fit projections of the ΔE and M_{bc} distributions in bins of reconstructed q^2 for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes. The goodness-of-fit is evaluated using a chi-square test, with the chi-square statistic detailed in Section 7.1. The number of degrees of freedom (DOF) is calculated as the number of fit bins minus the number of templates: $460 - 31 = 429$. The resulting $\chi^2/\text{DOF} = 468.5/429$, which corresponds to a p-value of 0.091.

The statistical correlation matrix of the component scale factors is shown in Figure 7.6, with the elements detailed in Table C.1 in Appendix C. The observed correlations are all below 0.75, with the largest correlations occurring between the $B \rightarrow X_c \ell \nu_\ell$ and *other* $B\bar{B}$ background scale factors in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode. Signal scale factors for high true q^2 templates become increasingly correlated with the $B \rightarrow X_u \ell \nu_\ell$ scale factor, for both $B \rightarrow \pi \ell \nu_\ell$ and $B \rightarrow \rho \ell \nu_\ell$ events.

First the fitted background scale factors are examined, before looking at the signal scale factors. This validates the understanding of the backgrounds and the associated uncertainties. Only with a well-understood background and properly estimated uncertainties the signal yields can be reliably extracted. The expected background yields from simulation and the fitted background scale factors, along with their statistical and systematic uncertainties, are given in Table 7.3. Systematic uncertainty sources are discussed in the next chapter. The fitted background scale factors are consistent with unity, supporting the assumption that the backgrounds are well-understood and uncertainties correctly assigned.

Next, the signal scale factors are examined. The expected total signal yields from simulation and the fitted signal scale factors for $B \rightarrow \pi \ell \nu_\ell$ and $B \rightarrow \rho \ell \nu_\ell$ events are provided in Table 7.4 and

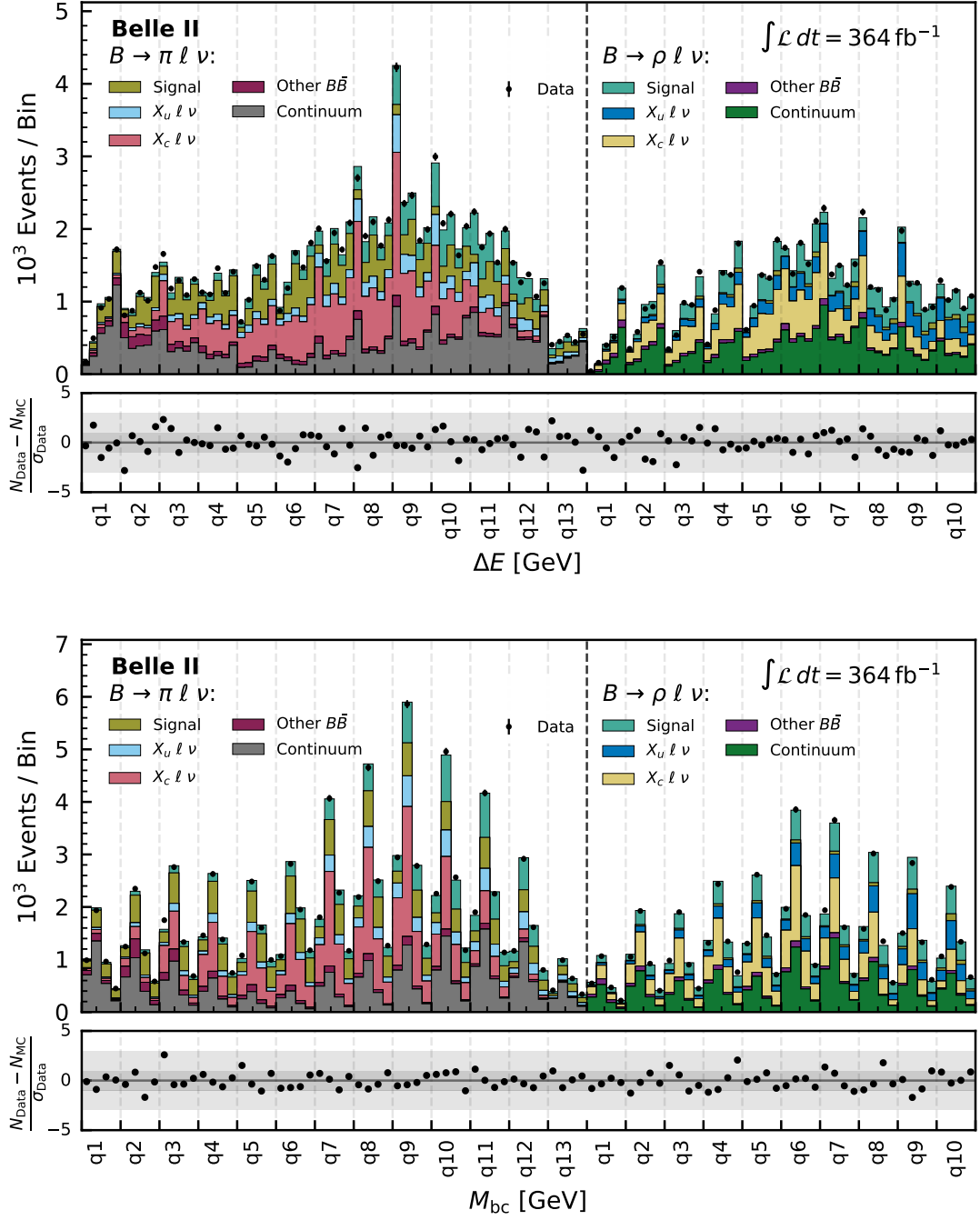


Figure 7.5: Distributions of ΔE (top) and M_{bc} (bottom) in the q^2 bins for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ candidates reconstructed in Belle II data with fit projections from simulation overlaid. The difference between collision and simulated data divided by the collision data uncertainty is shown in the panels below the histograms.

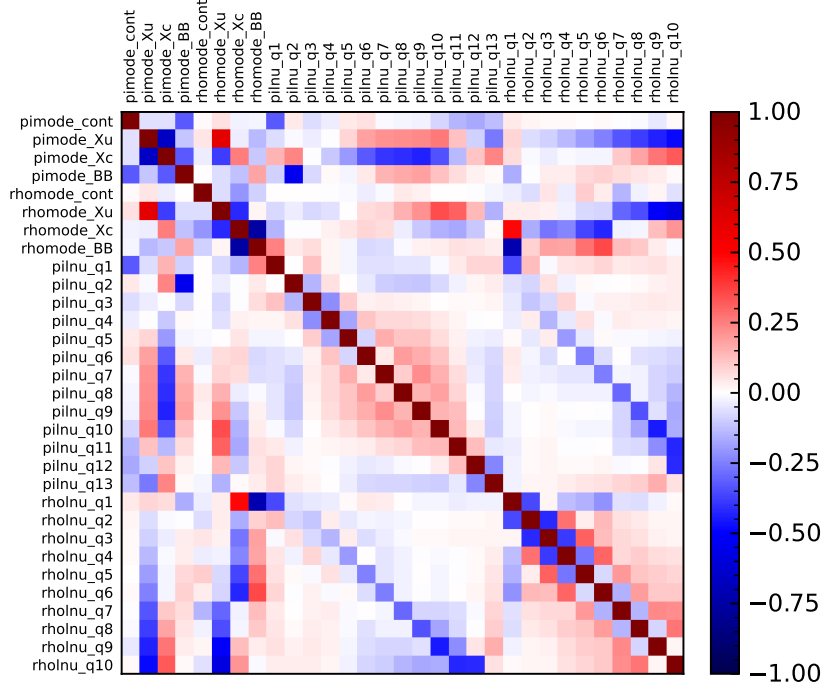


Figure 7.6: Illustration of the statistical correlation matrix of the fit component scale factors.

Table 7.3: Expected yields and fitted scale factors for the background templates. The first uncertainties are statistical and the second are systematic.

Component	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$		$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	
	Exp. yield	Scale factor	Exp. yield	Scale factor
$B \rightarrow X_c \ell \nu_\ell$	29847	$0.95 \pm 0.03 \pm 0.06$	16666	$1.04 \pm 0.06 \pm 0.10$
$B \rightarrow X_u \ell \nu_\ell$	7486	$1.11 \pm 0.14 \pm 0.37$	7918	$0.97 \pm 0.09 \pm 0.27$
Other $B\bar{B}$	4563	$0.90 \pm 0.11 \pm 0.17$	1796	$1.05 \pm 0.37 \pm 0.45$
Continuum	27410	$0.99 \pm 0.02 \pm 0.04$	18801	$0.97 \pm 0.02 \pm 0.04$

Table 7.5, respectively. Statistical and systematic uncertainties on the scale factors are included, covering all relevant systematic uncertainties described in the next chapter. Signal scale factors range from 0.85–1.27 for $B \rightarrow \pi \ell \nu_\ell$ and 0.86–1.39 for $B \rightarrow \rho \ell \nu_\ell$ events. Figure 7.7 and Figure 7.8 show the one-dimensional fit projections of ΔE and M_{bc} distributions in low, medium, and high reconstructed q^2 bins for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, respectively. The distributions in the remaining q^2 bins are provided in Appendix B.2.

As described in the previous section, the combined sample of true and combinatorial signal events is used to determine the partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays. The total signal template and expected yield also include contributions from isospin-conjugate signal and cross-feed events. Using the signal strengths shown in Table 7.2, the fitted scale factors, and

Table 7.4: Expected yields and fitted scale factors for each $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ signal template. The fitted yields of true and combinatorial signal events are provided. The first uncertainties are statistical and the second are systematic.

q^2 bin	Exp. yield	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	
		Scale factor	Yield
$q1$	1043	$0.85 \pm 0.09 \pm 0.14$	$869 \pm 95 \pm 139$
$q2$	1375	$1.05 \pm 0.08 \pm 0.13$	$1406 \pm 123 \pm 172$
$q3$	1694	$0.87 \pm 0.06 \pm 0.10$	$1426 \pm 112 \pm 124$
$q4$	1748	$1.02 \pm 0.06 \pm 0.08$	$1714 \pm 120 \pm 139$
$q5$	1794	$0.95 \pm 0.07 \pm 0.09$	$1617 \pm 120 \pm 113$
$q6$	1991	$1.16 \pm 0.08 \pm 0.13$	$2167 \pm 138 \pm 151$
$q7$	1826	$1.08 \pm 0.08 \pm 0.14$	$1817 \pm 143 \pm 172$
$q8$	1793	$1.17 \pm 0.08 \pm 0.17$	$1921 \pm 147 \pm 181$
$q9$	1765	$1.03 \pm 0.09 \pm 0.17$	$1640 \pm 149 \pm 174$
$q10$	1647	$0.94 \pm 0.09 \pm 0.18$	$1328 \pm 142 \pm 156$
$q11$	1471	$1.27 \pm 0.12 \pm 0.25$	$1472 \pm 140 \pm 239$
$q12$	1092	$1.03 \pm 0.14 \pm 0.27$	$819 \pm 120 \pm 211$
$q13$	534	$0.93 \pm 0.20 \pm 0.39$	$295 \pm 66 \pm 122$

Table 7.5: Expected yields and fitted scale factors for each $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal template. The fitted yields of true and combinatorial signal events are provided. The first uncertainties are statistical and the second are systematic.

q^2 bin	Exp. yield	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	
		Scale factor	Yield
$q1$	688	$0.97 \pm 0.28 \pm 0.34$	$332 \pm 100 \pm 118$
$q2$	1089	$1.06 \pm 0.18 \pm 0.29$	$651 \pm 114 \pm 178$
$q3$	1460	$0.76 \pm 0.16 \pm 0.24$	$630 \pm 131 \pm 131$
$q4$	1881	$1.00 \pm 0.14 \pm 0.23$	$1028 \pm 147 \pm 240$
$q5$	2300	$1.12 \pm 0.13 \pm 0.21$	$1273 \pm 158 \pm 236$
$q6$	2715	$0.98 \pm 0.13 \pm 0.20$	$1207 \pm 164 \pm 244$
$q7$	2915	$0.86 \pm 0.12 \pm 0.18$	$962 \pm 136 \pm 206$
$q8$	2892	$1.26 \pm 0.13 \pm 0.22$	$1141 \pm 118 \pm 218$
$q9$	2737	$1.22 \pm 0.14 \pm 0.24$	$936 \pm 114 \pm 186$
$q10$	2504	$1.39 \pm 0.15 \pm 0.37$	$821 \pm 96 \pm 220$

the expected total signal yields, the fitted signal yields in each true q^2 bin are obtained. The fitted signal yields for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, along with their statistical and systematic uncertainties, are provided in Table 7.4 and Table 7.5, respectively. The results include uncertainties related to the scale factors and those on the signal strengths, and are systematically limited,

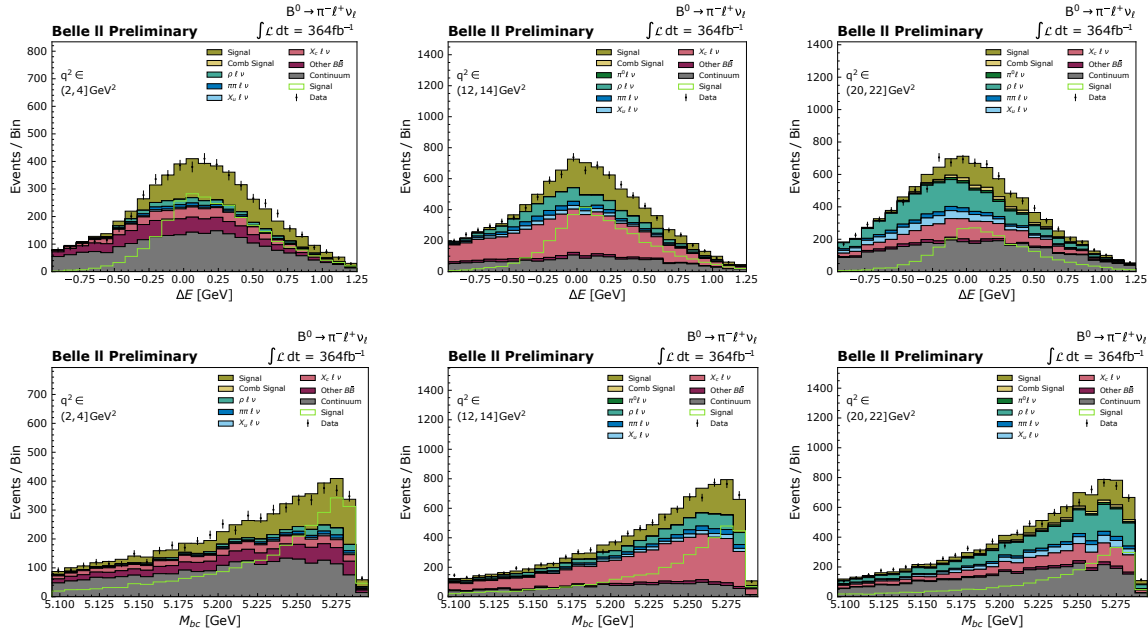


Figure 7.7: Distributions of ΔE (top) and M_{bc} (bottom) in the $q2$ (left), $q7$ (middle) and $q11$ (right) reconstructed q^2 bins in collision data for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ candidates with one-dimensional fit projections from simulation overlaid.

Fit stability

To assess the stability of the fit, additional tests are conducted. For instance, the dataset is divided by lepton flavor, lepton charge, or θ_{miss} region. The fit is then performed separately on each sub-sample and it is verified that the results are consistent within statistical uncertainties. As a specific example, the agreement of the background scale factors between the combined $B^0 \rightarrow \pi^- e^+ \nu_\mu$ and $B^+ \rightarrow \rho^0 e^+ \nu_\mu$ samples versus the combined $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ and $B^+ \rightarrow \rho^0 \mu^+ \nu_\mu$ samples is observed. Figure 7.9 shows a comparison of the background scale factors for the electron, muon, and total samples. The scale factors are consistent with each other and match the expected value of unity.

The fit stability is further verified by analyzing the reconstructed $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ samples separately. To achieve this, the cross-feed components are isolated from the total signal templates and they are treated as independent q^2 -integrated templates. The results for the signal scale factors are consistent with the nominal results. It is also observed that these scale factors exhibit larger relative uncertainties, which is expected as the cross-feed background is no longer constrained by the second mode.

Additionally, as discussed in Section 5.2.1, stability tests are performed to ensure that the fit results are not influenced by the mismodeled regions in lepton and pion polar angles. This is accomplished by successively tightening the thresholds on the polar angles and repeating the fit to these subsets of data. By demonstrating that the resulting scale factors agree within statistical uncertainties, after accounting for sample overlaps, this test confirms that no bias has been introduced.

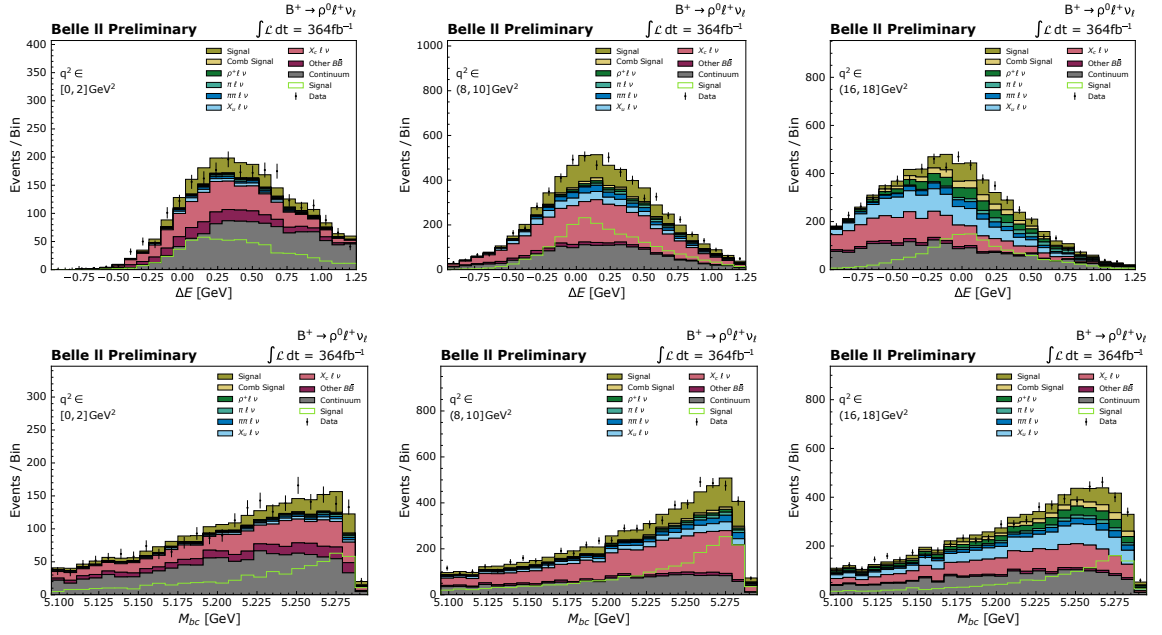


Figure 7.8: Distributions of ΔE (top) and M_{bc} (bottom) in the $q1$ (left), $q5$ (middle) and $q9$ (right) reconstructed q^2 bins in collision data for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ candidates with one-dimensional fit projections from simulation overlaid.

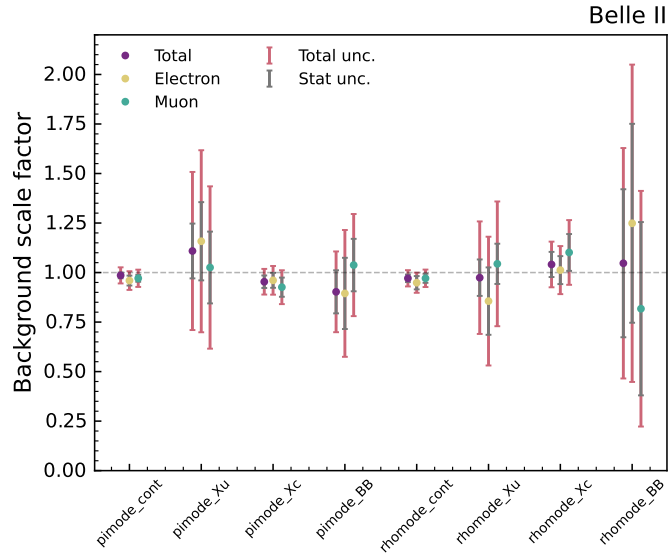


Figure 7.9: Fitted background scale factors for the nominal (total) fit with statistical and total uncertainties, compared to those obtained when the electron and muon samples are fit separately.

7.6 Branching fraction results

Next, the partial branching fractions are determined from the fitted yields. This process and the obtained results are presented in this section.

The partial branching fractions $\Delta\mathcal{B}_i$ for each true q^2 bin i are computed using the signal yields, N_i , provided in Table 7.4 and Table 7.5. Additionally, corrections are applied for the selection efficiencies using the signal efficiencies ϵ_i listed in Table 6.2. The partial branching fractions are given by the following equations:

$$\Delta\mathcal{B}_i(B^0 \rightarrow \pi^- \ell^+ \nu_\ell) = \frac{N_i(1 + f_{+0})}{4\epsilon_i \times N_{B\bar{B}}}, \quad (7.4a)$$

$$\Delta\mathcal{B}_i(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell) = \frac{N_i(1 + f_{+0})}{4\epsilon_i \times N_{B\bar{B}}} \times \frac{1}{f_{+0}}, \quad (7.4b)$$

where $f_{+0} = \mathcal{B}(\Upsilon(4S) \rightarrow B^+ B^-) / \mathcal{B}(\Upsilon(4S) \rightarrow B^0 \bar{B}^0) = 1.065 \pm 0.052$ [127] converts the number of $B\bar{B}$ pairs $N_{B\bar{B}}$ determined by the Belle II performance group into the number of neutral or charged B mesons in the collision data sample. For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$, the number of neutral B mesons is required, whereas for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, the number of charged B mesons is relevant. The factors of four in the denominators account for the two B mesons per $B\bar{B}$ pair and for the two lepton flavors (e, μ) that are combined in the measurement.

The partial branching fractions $\Delta\mathcal{B}_i$ for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays are presented in Table 7.6. The uncertainties quoted include both statistical and systematic components. These uncertainties reflect the variations in fitted yields and efficiencies, as well as those in $N_{B\bar{B}}$ and f_{+0} , which are discussed in the following chapter. Systematic uncertainties dominate the results. The full experimental correlation matrices for the partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ are provided in Tables C.2 and C.3 in Appendix C.

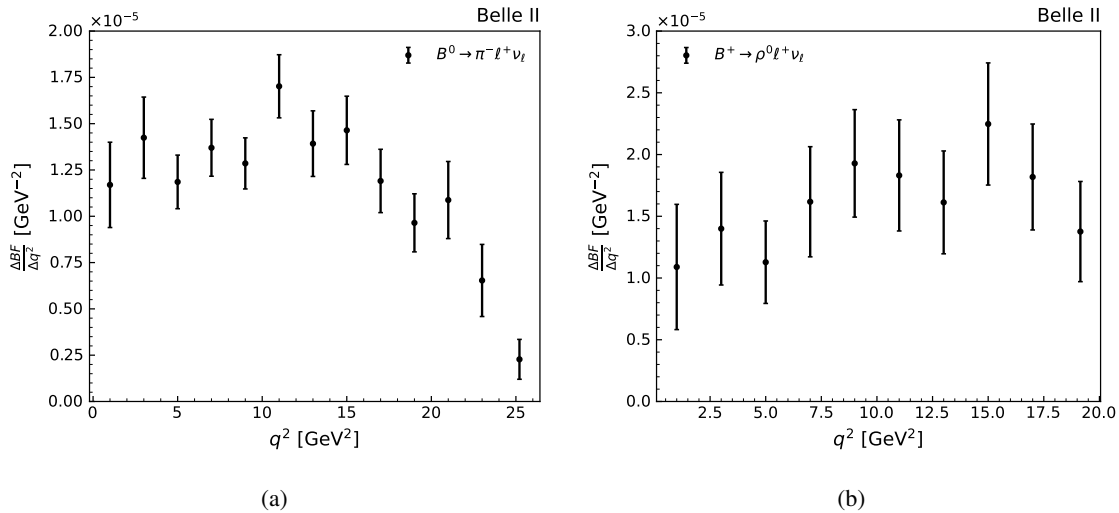


Figure 7.10: Measured differential rates of (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays as a function of true q^2 . The error bars display the total uncertainty.

Table 7.6: Partial branching fractions $\Delta\mathcal{B} (\times 10^4)$ of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events in each true q^2 bin. The first uncertainty is statistical and the second is systematic.

q^2 bin	$\Delta\mathcal{B} (\times 10^4)$	
	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
$q1$	$0.117 \pm 0.013 \pm 0.019$	$0.109 \pm 0.033 \pm 0.039$
$q2$	$0.142 \pm 0.013 \pm 0.018$	$0.140 \pm 0.025 \pm 0.038$
$q3$	$0.119 \pm 0.009 \pm 0.011$	$0.113 \pm 0.024 \pm 0.024$
$q4$	$0.137 \pm 0.010 \pm 0.012$	$0.162 \pm 0.023 \pm 0.038$
$q5$	$0.129 \pm 0.010 \pm 0.010$	$0.193 \pm 0.024 \pm 0.036$
$q6$	$0.170 \pm 0.011 \pm 0.013$	$0.183 \pm 0.025 \pm 0.037$
$q7$	$0.139 \pm 0.011 \pm 0.014$	$0.161 \pm 0.023 \pm 0.035$
$q8$	$0.146 \pm 0.011 \pm 0.015$	$0.225 \pm 0.023 \pm 0.044$
$q9$	$0.119 \pm 0.011 \pm 0.013$	$0.182 \pm 0.022 \pm 0.037$
$q10$	$0.096 \pm 0.010 \pm 0.012$	$0.158 \pm 0.019 \pm 0.043$
$q11$	$0.109 \pm 0.010 \pm 0.018$	
$q12$	$0.065 \pm 0.010 \pm 0.017$	
$q13$	$0.028 \pm 0.006 \pm 0.011$	

Before proceeding to measure $|V_{ub}|$ using the differential rates shown in Figure 7.10, the total branching fractions are calculated by summing the partial branching fractions:

$$\begin{aligned}\mathcal{B}(B^0 \rightarrow \pi^- \ell^+ \nu_\ell) &= (1.516 \pm 0.042 \pm 0.059) \times 10^{-4} \\ \mathcal{B}(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell) &= (1.625 \pm 0.079 \pm 0.180) \times 10^{-4}\end{aligned}$$

where the first uncertainties are statistical and the second are systematic. The full experimental correlation between these values is -0.16 . The results are consistent with the world averages [37] and are primarily limited by the systematic uncertainties discussed in the subsequent chapter.

Systematic uncertainties

Identifying and evaluating sources of uncertainty is essential for the accurate interpretation of results. This process is described in this chapter. Only after quantifying the uncertainties, the results, their limitations, and potential strategies for reducing uncertainty contributions in the future can be fully understood and developed.

First, the types of uncertainties and how they affect the results are described in Section 8.1. Next, the methods used to evaluate the uncertainties are introduced in Section 8.2. Finally, Section 8.3 describes the sources of uncertainties and their impact on the results.

8.1 Uncertainty types

Several types of uncertainty impact the measurement of partial branching fractions and thus influence the determination of $|V_{ub}|$. Beyond statistical uncertainties due to limited sample sizes, systematic uncertainties can arise at various stages of the measurement process. Different kinds of systematic uncertainty are discussed in this section.

Systematic uncertainties from external inputs directly affect the calculation of partial branching fractions, as outlined in Equations 7.4a and 7.4b. These include uncertainties related to $N_{B\bar{B}}$ and f_{+0} , which impact the overall normalization of the measurements without altering the shape of the partial branching fraction spectra.

Other sources of uncertainty influence the measurement of signal yields N_i in Equations 7.4a and 7.4b. Such uncertainties can introduce variations in the shape of the partial branching fraction spectra. Since signal yields are derived from fitted signal scale factors and signal strengths, uncertainties associated with these components must be accounted for. Additionally, uncertainties in signal efficiencies can also affect the shapes of the partial branching fraction spectra. The uncertainties related to these components encompass both statistical and systematic contributions, with systematic uncertainties arising from multiple sources, as discussed below.

Systematic uncertainties can affect the branching fraction results in different ways. When a source of uncertainty impacts the signal templates, uncertainties on all three components—fitted scale factors, signal strengths, and efficiencies—must be considered. In contrast, uncertainties that affect only the background templates generally influence only the fitted scale factors.

8.2 Uncertainty evaluation

The evaluation of the systematic uncertainties is described in this section. To assess systematic uncertainties, two methods are employed: a toy MC method and a bootstrapping technique. The toy MC method is introduced in Section 7.5. The bootstrapping technique involves generating new samples that replicate the statistical properties of the original data by resampling with replacement from the simulated or experimental dataset [128].

For the results presented in this thesis, systematic uncertainties are consistently evaluated using the toy MC method, except in a few cases where bootstrapping is more appropriate. For each identified source of uncertainty, 1000 variations of the templates are produced. Variations are generated by sampling from distributions of central values with associated uncertainties and accounting for correlations. 1000 variations are selected to balance computational feasibility with statistical robustness.

For instance, the procedure for evaluating the impact of uncertainties related to the $B \rightarrow \omega \ell \nu_\ell$ form factors, as described in Section 4.2.2, involves generating 1000 alternative $B \rightarrow X_u \ell \nu_\ell$ background distributions. These alternative distributions assume that the form-factor coefficients follow Gaussian distributions with standard deviations corresponding to the uncertainties on the coefficients. Subsequently, 1000 toy distributions are created by adding the alternative $B \rightarrow X_u \ell \nu_\ell$ templates to the unaffected templates. Fitting the nominal templates to these toy distributions and recording the fitted scale factors enables the derivation of a covariance matrix for the signal scale factors using Pearson correlation [129]. In this example, the covariance matrix is related to the $B \rightarrow \omega \ell \nu_\ell$ form factors.

The toy MC method provides a distinct covariance matrix for the signal scale factors associated with each source of systematic uncertainty. If a source impacts the signal templates, the toy MC evaluations extend to signal strengths and efficiencies. For each set of varied signal templates, the efficiencies and signal strengths are recorded. The systematic covariance matrices for these components are then obtained through Pearson correlation.

The systematic covariance matrix for the partial branching fractions, related to each uncertainty source, is determined by propagating the covariance matrices of the fitted signal scale factors, signal strengths, and efficiencies, accounting for all correlations. Similarly, the statistical covariance matrix for the partial branching fractions is derived from the statistical covariance matrices of the three components. The total systematic covariance matrix for the partial branching fractions is the sum of covariance matrices from each systematic uncertainty source, assuming these sources are uncorrelated. Ultimately, the full experimental covariance matrix for the partial branching fractions is the sum of the total systematic and statistical covariance matrices.

8.3 Uncertainty sources

This section describes the various known sources of systematic uncertainty affecting the partial branching fraction results in each true q^2 bin. This includes a discussion of effects related to detector inefficiencies (Section 8.3.1), beam-energy differences (Section 8.3.2), differences in the BDT selection efficiency (Section 8.3.3), uncertainties due to limited sizes of simulated samples (Section 8.3.4), and uncertainties on physics constraints (Section 8.3.5). In addition, Sections 8.3.6, 8.3.7 and 8.3.8 describe uncertainties related to the $B \rightarrow X_u \ell \nu_\ell$, $B \rightarrow X_c \ell \nu_\ell$ and continuum background templates, respectively. Lastly, Section 8.3.9 discusses uncertainties related to the signal model and the ρ meson

lineshape.

The absolute uncertainties on these results due to a specific source of uncertainty are the square roots of the diagonal elements of the relevant covariance matrix. Tables 8.1 and 8.2 show the fractional uncertainties on the partial branching fractions in each true q^2 bin, arising from known systematic uncertainty sources, for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. These tables also display the relative statistical, total systematic, and full experimental uncertainties. The elements of the full experimental covariance matrices for the partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays are presented in Tables C.2 and C.3 in Appendix C.

For the partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events, the relative total uncertainties range from 10.0% to 47.3%. The uncertainties are dominated by those associated with the continuum template, leading to large total uncertainties at high q^2 , but smaller uncertainties for intermediate q^2 bins. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, the relative total uncertainties are largest at low q^2 , reaching 46.6%, and are relatively constant in the remaining q^2 bins, ranging from 22.0% to 32.6%.

Table 8.1: Summary of fractional uncertainties in % by source on the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ partial branching fractions $\Delta\mathcal{B}$ in each true q^2 bin.

Source	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$												
	$q1$	$q2$	$q3$	$q4$	$q5$	$q6$	$q7$	$q8$	$q9$	$q10$	$q11$	$q12$	$q13$
Detector effects	2.0	0.9	1.1	1.0	1.0	1.1	1.1	1.0	0.9	1.2	2.3	4.1	5.8
Beam energy	0.6	0.8	0.7	0.8	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.6	0.7
Sim. sample size	4.7	3.8	3.3	3.2	3.2	2.9	3.8	3.7	4.0	4.5	5.9	8.0	13.6
BDT efficiency	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
Phys. constraints	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9
Signal model	0.1	0.1	0.2	0.1	0.0	0.2	0.2	0.4	0.3	0.8	0.9	0.2	4.9
ρ lineshape	0.1	0.1	0.3	0.3	0.2	0.1	0.3	0.1	0.3	0.1	0.2	0.2	0.6
N. $B \rightarrow \pi\pi\ell\nu_\ell$	0.5	0.6	0.4	0.4	0.5	1.0	1.2	1.0	0.8	1.8	1.2	2.3	14.3
DFN parameters	0.8	0.4	1.5	1.6	1.4	1.7	1.2	0.1	0.7	1.2	2.9	3.5	3.7
Oth. $B \rightarrow X_u\ell\nu_\ell$	0.2	0.4	0.3	0.4	0.2	0.9	1.1	1.2	1.0	1.3	1.6	0.7	8.7
$B \rightarrow X_c\ell\nu_\ell$	1.4	2.0	1.7	1.3	1.3	1.4	1.8	1.6	1.3	1.4	1.1	0.5	1.7
Continuum	15.1	11.3	7.6	7.1	5.8	5.7	8.1	8.3	9.6	10.4	14.5	23.8	34.4
Total systematic	16.4	12.6	9.3	8.7	7.7	7.7	10.0	9.9	11.1	12.2	16.6	26.0	41.6
Statistical	11.0	8.8	7.9	7.0	7.5	6.4	7.9	7.7	9.1	10.7	9.6	14.6	22.6
Total	19.7	15.4	12.2	11.2	10.7	10.0	12.7	12.6	14.4	16.3	19.1	29.8	47.3

8.3.1 Detector effects

Detector uncertainties arise from variations in tracking efficiency and uncertainties assigned to the corrections to lepton and pion PID efficiencies, which are discussed in this subsection. The uncertainties on the partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events range from 0.9–5.8%. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, the uncertainties range from 1.1–2.8%.

Tracking efficiency

Uncertainties in tracking efficiency, introduced in Section 3.4.1, affect all simulated fit templates. To quantify these uncertainties, the Belle II performance group compared tracking efficiency in data to

Table 8.2: Summary of fractional uncertainties in % by source on the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ partial branching fractions $\Delta\mathcal{B}$ in each true q^2 bin.

Source	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$									
	$q1$	$q2$	$q3$	$q4$	$q5$	$q6$	$q7$	$q8$	$q9$	$q10$
Detector effects	2.8	2.0	1.6	1.1	1.7	1.9	2.4	1.4	1.4	1.6
Beam energy	2.1	1.9	1.9	1.5	1.3	1.1	1.0	0.9	0.8	0.5
Sim. sample size	14.1	7.8	7.4	6.3	6.3	5.2	6.4	5.6	6.2	7.3
BDT efficiency	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6
Phys. constraints	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
Signal model	0.7	0.2	0.2	0.2	0.3	0.4	0.5	0.3	1.8	2.4
ρ lineshape	1.7	1.6	2.0	1.0	1.9	1.8	1.4	0.9	1.6	1.7
Nonres. $B \rightarrow \pi\pi\ell\nu_\ell$	5.6	6.3	6.7	8.6	9.3	10.7	10.1	7.0	7.8	11.8
DFN parameters	3.6	5.5	4.1	3.5	1.1	1.2	2.7	1.7	1.9	2.3
Oth. $B \rightarrow X_u\ell\nu_\ell$	1.7	3.0	3.8	5.0	5.8	6.1	6.3	1.9	7.2	12.4
$B \rightarrow X_c\ell\nu_\ell$	1.8	1.9	1.7	1.1	1.4	1.7	0.9	0.9	1.9	2.6
Continuum	31.5	24.3	17.0	19.6	13.2	14.8	16.0	16.6	15.2	18.7
Total systematic	35.6	27.5	21.0	23.5	18.8	20.5	21.6	19.4	20.2	27.0
Statistical	30.0	17.5	20.8	14.4	12.4	13.6	14.1	10.4	12.2	11.8
Total	46.6	32.6	29.6	27.6	22.6	24.6	25.8	22.0	23.6	29.5

that in simulation using an $e^+e^- \rightarrow \tau^+\tau^-$ control sample. The deviation from unity, measured as 0.24%, is designated as the per-track uncertainty.

The impact of this uncertainty depends on the number of tracks per event, denoted as N_{track} . Reconstructing $B^0 \rightarrow \pi^-\ell^+\nu_\ell$ and $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ events involves two and three tracks, respectively. Background samples consist solely of reconstructed $B^0 \rightarrow \pi^-\ell^+\nu_\ell$ and $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ events, without mixing. Thus, the uncertainty translates into a normalization shift of the templates. For signal templates, which include both two and three-track reconstructed events due to the cross-feed component, the uncertainty affects not only the normalization but also the shapes of the signal templates.

To model these variations, 1000 alternative templates are generated, assuming the tracking efficiency follows a Gaussian distribution with standard deviation $N_{\text{track}} \times 0.0024$. The distributions of scale factors, signal strengths, and efficiencies from these templates are used to approximate the covariance matrices for each. These covariance matrices are then propagated to the partial branching fractions. The resulting uncertainties on the partial branching fractions range from 0.4% to 1.5%.

Particle ID corrections

Corrections to particle ID efficiencies, described in Section 4.3.1, affect the shapes of all fit templates. Uncertainties in these correction factors lead to uncertainties in scale factors, signal strengths, and efficiencies. These effects are addressed using the statistical and systematic uncertainties provided by the Belle II performance group. Whenever asymmetric uncertainties are provided, they are symmetrized by adopting the larger value for both upward and downward directions. Statistical and systematic uncertainties σ_i^{stat} and σ_i^{syst} , respectively, are derived for each momentum and polar angle

bin i .

To incorporate these uncertainties, 1000 varied correction factors are generated per bin. The varied correction factor for variation n and bin i is denoted as $R_i^n = R_i^{\text{nom}} + \Delta\sigma_i^{\text{stat},n} + \Delta\sigma_i^{\text{syst},n}$, where R_i^{nom} is the nominal correction factor. Statistical and systematic uncertainties are represented by $\Delta\sigma_i^{\text{stat},n}$ and $\Delta\sigma_i^{\text{syst},n}$, respectively. Correlations across bins are accounted for in determining these components.

The statistical component is treated as uncorrelated between bins, sampled from a Gaussian distribution with mean zero and standard deviation σ_i^{stat} . The systematic component is fully correlated across bins, with a global scaling factor for each variation n extracted from a unit normal distribution. This systematic component is calculated as the product of the scaling factor and σ_i^{syst} .

These varied correction factors are used to create 1000 toy distributions. The relative systematic uncertainties on the partial branching fractions due to lepton efficiency and misidentification rate corrections range from 0.7% to 5.6%. Uncertainties from pion efficiency and misidentification rate corrections are smaller, ranging from 0.2% to 0.4%.

8.3.2 Beam-energy difference

As discussed in Section 5.5, effects of differences between the CM energy in simulated and collision data on the signal extraction variables are studied. The evaluation of these effects is described in this subsection.

The simulated CM energy is 10.5796 GeV, while the mean CM energy observed in collision data is 10.5771 GeV, resulting in a 2.8 MeV difference. This CM energy difference affects the modeling of the reconstructed q^2 distribution in simulation, leading to variations in the shapes of the signal templates.

To estimate this effect and account for the associated systematic uncertainty, a control mode is studied by fully reconstructing $B^+ \rightarrow J/\psi[\rightarrow \mu^+\mu^-]K^+$ events. Simultaneously, $B^0 \rightarrow \mu^- K^+$ events are reconstructed within the same fully reconstructed event set. In this way the missing neutrino in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal decays can be simulated using the undetected muon from the J/ψ decay, while the ROE of the fully reconstructed event mirrors the conditions of the signal decays. In this control mode, a 4% difference in q^2 resolution is observed between collision and simulated data. Further details are provided in Appendix A.3.

To address this resolution difference, the q^2 resolution in each true q^2 bin of the simulated data is adjusted using a Gaussian-smearing method. For each true q^2 bin, 1000 alternative reconstructed q^2 distributions are generated by drawing each reconstructed q^2 value from a Gaussian distribution centered at the event's true q^2 value, with a standard deviation corresponding to the updated resolution for that bin. This process produces 1000 alternative signal templates with varied reconstructed q^2 distributions. These toy distributions are then fit using the nominal templates. The resulting systematic uncertainties on the partial branching fractions for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events range from 0.5% to 0.8%, while for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, the uncertainties range from 0.5% to 2.1%.

8.3.3 BDT selection efficiency

The uncertainties on the signal efficiencies arising from potential discrepancies between the signal selection efficiencies in collision and simulated data due to the selection on the 48 BDT classifiers, are described in this subsection. For each BDT classifier selection discussed in Section 6.2.4, the $B^+ \rightarrow J/\psi[\rightarrow \mu^+\mu^-]K^+$ control mode, described above and in Appendix A.3, is used to determine

the ratio between the selection efficiencies in collision and simulated data. These ratios are found to be consistent with unity within their respective uncertainties, ranging from 0.963 ± 0.043 to 1.008 ± 0.009 .

To propagate these uncertainties to the efficiencies, the standard deviations of the ratios for each type of BDT (continuum or $B\bar{B}$ suppression) and mode ($B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ or $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$) are independently evaluated assuming Gaussian distributions. This results in four independent uncertainties: σ_{cont}^π and $\sigma_{B\bar{B}}^\pi$ for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ signal efficiencies across all true q^2 bins, and $\sigma_{\text{cont}}^\rho$ and $\sigma_{B\bar{B}}^\rho$ for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal efficiencies across all true q^2 bins.

This approach is appropriate because the efficiency ratios for continuum (or $B\bar{B}$) suppression classifier selections are correlated within each q^2 bin. This correlation arises because all continuum (or $B\bar{B}$) suppression classifiers rely on similar input variables, and the same data sample is used to determine each efficiency ratio. The uncertainties on the efficiencies from the continuum suppression BDT classifiers are 1.1% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 0.6% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. For the $B\bar{B}$ suppression BDT classifiers, the uncertainties are 0.7% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 1.5% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$.

8.3.4 Simulated sample size

The limited size of the simulated samples (MC statistics) used to generate the fit templates, as discussed in Section 7.4, introduces uncertainties by potentially altering the shapes of the templates. For the signal templates, it also induces uncertainties in scale factors, signal strengths and efficiencies, due to bin migrations. These uncertainties are discussed in this subsection. The combined relative uncertainties on the partial branching fractions from these effects range from 2.9% to 14.1%, making them one of the dominant sources of uncertainty. The uncertainties are most pronounced at the extremes of the q^2 spectrum, consistent with the lower number of simulated background events in these regions.

Background sample statistics

The shapes of the background templates are influenced by uncertainties arising from the limited sizes of the simulated samples used for their generation. To assess the impact, the counts are adjusted within the Poisson statistical errors in each bin of the fit variable 1000 times for all background templates. From the signal scale factors resulting from these 1000 fits to the corresponding toy distributions, the uncertainties on the partial branching fractions are derived.

Separating these uncertainties into contributions from the simulated continuum sample and other samples reveals that the uncertainties related to continuum sample sizes dominate at the extremes of the q^2 spectra, with relative uncertainties up to 11.4%. In contrast, the maximum uncertainty attributed to other simulated samples is 7.9%. This discrepancy aligns with the data volumes: 1 ab^{-1} for simulated continuum samples versus 5 ab^{-1} for the generic samples. Increasing the simulated sample sizes, especially for continuum backgrounds, would significantly reduce these uncertainties in future studies.

Signal sample statistics: Bin migrations

An additional source of systematic uncertainty from the limited size of the simulated samples impacting the signal templates is bin migration. Since the nominal templates are defined by true q^2 values, effects changing the distribution of true q^2 , such as those caused by limited sample sizes, need to be

considered. Due to this, uncertainties are assigned based on the migration of signal events between true q^2 bins, reflecting signal template migrations.

To evaluate these uncertainties, a bootstrapping technique is used instead of the usual toy MC method. By employing a bootstrapping technique instead, the effect of events migrating between templates can be simulated. 1000 alternative signal samples are generated by sampling with replacement from the true q^2 distributions. Each sample is then separated into true q^2 templates and fitted to the nominal dataset constructed from the nominal templates. The usual procedure is followed to determine the uncertainties on the scale factors, signal strengths, and efficiencies associated with bin migrations using Pearson correlation.

The advantage of bootstrapping is that it inherently accounts for additional effects present in the background templates. The uncertainties on the partial branching fractions due to bin migrations are smaller than those from limited background sample sizes. They range from 1.0% to 3.4% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and from 1.9% to 3.8% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal events.

8.3.5 Physics constraints

Systematic uncertainties on the partial branching fractions arising from physics constraints, described in this subsection, include uncertainties related to the number of $B\bar{B}$ pairs and the value of f_{+0} . These uncertainties primarily impact the overall normalization of the partial branching fractions without altering the shape of the q^2 spectra. In contrast, uncertainties associated with the assumption of isospin symmetry affect the normalization between charged and neutral B meson decays within the total signal templates, influencing the shapes of the branching fraction spectra.

$N_{B\bar{B}}$ and f_{+0}

As described in Section 7.6, the number of $B\bar{B}$ pairs, $N_{B\bar{B}}$, and $f_{+0} = \mathcal{B}(\Upsilon(4S) \rightarrow B^+ B^-) / \mathcal{B}(\Upsilon(4S) \rightarrow B^0 \bar{B}^0)$ are needed for evaluating the number of charged and neutral B mesons in the collision dataset. Systematic uncertainties on the partial branching fractions arise from uncertainties in these inputs. The Belle II performance group provides the value of $N_{B\bar{B}} = (387 \pm 6)$, while Ref. [127] determines $f_{+0} = 1.065 \pm 0.052$. Propagating these uncertainties to the partial branching fractions, relative normalization uncertainties of 2.9% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events and 2.8% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events are derived.

Isospin symmetry

An uncertainty is assigned to the assumption of isospin symmetry, which is applied during the creation of the total signal templates, as described in Section 7.4.2. This assumption involves using the same scale factor in the fit for both charged and neutral $B \rightarrow \pi \ell \nu_\ell$ and $B \rightarrow \rho \ell \nu_\ell$ decays, as specified in Equation 7.3, and fixing the B meson lifetime ratio τ_+/τ_0 to the constant value used in simulation.

To evaluate the impact of deviations from the simulation value on the fitted scale factors and signal strengths, the relative uncertainty of $\tau_+/\tau_0 = 1.076 \pm 0.004$ [37] is employed. The relative fractions of neutral and charged B meson events in the signal templates is adjusted accordingly. Gaussian variations of the fractions are sampled within the relative uncertainty of 0.37% to produce varied signal templates. The fits of the nominal templates to the toy distributions reveal that the contribution of the isospin symmetry assumption to the systematic uncertainties on the partial branching fractions

for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events is negligible. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, the contribution is also negligible at low q^2 , but introduces small relative uncertainties of 0.1% at high q^2 , where the isospin-conjugate signal component is significant.

8.3.6 $B \rightarrow X_u \ell \nu_\ell$ background

The uncertainties associated with the partial branching fractions due to $B \rightarrow X_u \ell \nu_\ell$ background events, described in this subsection, include uncertainties related to the modeling of the nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ component, the parameters of the DFN model describing the shape of the total nonresonant $B \rightarrow X_u \ell \nu_\ell$ component, and the form-factor coefficients and branching fractions. These uncertainties impact the shape of the $B \rightarrow X_u \ell \nu_\ell$ template. The contributions to uncertainties from the latter two sources are summarized in the “Oth. $B \rightarrow X_u \ell \nu_\ell$ ” category in Tables 8.1 and 8.2.

Nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ model

One significant uncertainty related to the $B \rightarrow X_u \ell \nu_\ell$ background events is associated with the modeling of the nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ component, as discussed in Section 4.2.4. Due to the kinematic similarity between $B \rightarrow \pi \pi \ell \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, deviations from the simulated behavior can substantially affect the measured partial branching fractions.

In the fit, both the shape in $m_{\pi\pi}$ and the absolute normalization of the $B \rightarrow \pi \pi \ell \nu_\ell$ background relative to the $B \rightarrow X_u \ell \nu_\ell$ template are fixed according to the measurements reported by Ref. [109]. To account for uncertainties in these measurements, 1000 varied $m_{\pi\pi}$ spectra are generated based on the covariance matrix provided by Ref. [109]. These varied spectra are used to assess both normalization and $m_{\pi\pi}$ shape differences. Since the nominal $m_{\pi\pi}$ spectrum around the ρ meson mass peak is obtained by interpolation from measurements on either side of the peak, a similar interpolation approach is employed to produce the 1000 varied spectra.

Next, 1000 alternative $B \rightarrow X_u \ell \nu_\ell$ distributions are created and the usual fitting procedure is performed to evaluate uncertainties on the fitted scale factors. The uncertainties on the partial branching fractions are significant in the high q^2 region for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events, reaching up to 14.3%. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, these uncertainties are among the most dominant across all true q^2 bins, ranging from 5.6% to 11.8%.

DFN parameters

Another source of uncertainty arises from the description of the remaining nonresonant $B \rightarrow X_u \ell \nu_\ell$ background events, as discussed in Section 4.2.5. This uncertainty is associated with the shape-function parameters that define the nonresonant $B \rightarrow X_u \ell \nu_\ell$ background shape within the DFN model. The uncertainties on the partial branching fractions related to these parameters are assessed by following the methodology outlined in Ref. [9]. This evaluation includes accounting for the correlation between the parameters and determining the uncertainties in the eigenvector basis of the relevant covariance matrix. By aligning variations with the principal axes of the parameter uncertainty space, this approach simplifies the analysis. Thus, the uncertainties due to both DFN parameters can be simultaneously evaluated while considering parameter correlations.

Initially, the $i = 2$ orthogonal uncertainties on the vector containing the nominal parameters \vec{u}_{nom} : $\Delta \vec{u}_i = \sqrt{\lambda_i} \vec{e}_i$ are generated, where λ_i and \vec{e}_i represent the i -th eigenvalue and eigenvector of the

covariance matrix, respectively. Using these uncertainties, varied DFN parameters are sampled from multivariate Gaussian distributions with standard deviations $\Delta\vec{u}_i$ to create 1000 varied $B \rightarrow X_u \ell \nu_\ell$ templates. The resulting uncertainties on the partial branching fractions range from 0.1% to 3.7% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events and from 1.1% to 5.5% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events.

Additionally, the results obtained using the DFN model are compared with those from an alternative nonresonant $B \rightarrow X_u \ell \nu_\ell$ model. This comparison is performed using the BLNP model [46], introduced in Section 2.2.1. Since the observed differences are smaller than the uncertainties assigned to the DFN model parameters, no additional uncertainties are introduced due to model choice.

$B \rightarrow X_u \ell \nu_\ell$ background: Form factors

Uncertainties on the partial branching fractions arising from the background $B \rightarrow X_u \ell \nu_\ell$ form factors, where $X_u \in \{\omega, \eta, \eta'\}$, are considered. The form-factor coefficients used in the simulation, listed in Table 4.4, are provided with covariance matrices for $B \rightarrow \omega \ell \nu_\ell$ events in Ref. [10], and for $B \rightarrow \eta \ell \nu_\ell$ and $B \rightarrow \eta' \ell \nu_\ell$ events in Ref. [106]. To assess the uncertainties due to correlated form-factor coefficients, a method similar to that used for the DFN parameters is used.

First, uncertainties in the eigenvector basis of the covariance matrices are generated. From these uncertainties, $\Delta\vec{u}_i = \sqrt{\lambda_i} \vec{e}_i$, varied coefficients are sampled from multivariate Gaussian distributions with standard deviations $\Delta\vec{u}_i$ to create 1000 varied $B \rightarrow X_u \ell \nu_\ell$ distributions. Fitting the 1000 resulting toy distributions, negligible uncertainties on the partial branching fractions for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events are determined in most true q^2 bins. Significant uncertainties appear only in the highest q^2 bins, reaching up to 4.6% and 1.4% for the $B \rightarrow \omega \ell \nu_\ell$ and $B \rightarrow \eta \ell \nu_\ell$ form-factor coefficients, respectively. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, the relative uncertainties on the partial branching fractions are larger across the q^2 spectrum.

$B \rightarrow X_u \ell \nu_\ell$ background: Branching fractions

Uncertainties associated with the branching fractions of $B \rightarrow X_u \ell \nu_\ell$ background events, both exclusively and inclusively measured, as listed in Table 4.2, are considered. 1000 alternative $B \rightarrow X_u \ell \nu_\ell$ templates are created for $B \rightarrow \omega \ell \nu_\ell$, $B \rightarrow \eta \ell \nu_\ell$, $B \rightarrow \eta' \ell \nu_\ell$, and the inclusively measured $B \rightarrow X_u \ell \nu_\ell$ branching fractions using Gaussian variations with standard deviations given in Table 4.2.

The uncertainties on the partial branching fractions from the $B \rightarrow \omega \ell \nu_\ell$ and $B \rightarrow \eta \ell \nu_\ell$ branching fractions are similar, ranging from below 0.1% to 0.8%. Uncertainties from the inclusively measured $B \rightarrow X_u \ell \nu_\ell$ branching fraction are significantly larger. For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events, these uncertainties range from 0.1% to 1.2%, with an exception in the highest q^2 bin, where the uncertainty reaches 7.1%. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, the uncertainties are more pronounced, varying from 1.4% to 6.3%, and up to 12.2% in the highest q^2 bin.

8.3.7 $B \rightarrow X_c \ell \nu_\ell$ background

The uncertainties in the form-factor coefficients and branching fractions of the $B \rightarrow X_c \ell \nu_\ell$ background events are discussed in this subsection. These uncertainties affect the shapes of the $B \rightarrow X_c \ell \nu_\ell$ templates. The methodologies used are similar to those applied for evaluating the $B \rightarrow X_u \ell \nu_\ell$ background form-factor and branching fraction uncertainties.

$B \rightarrow X_c \ell \nu_\ell$ background: Form factors

Uncertainties on the partial branching fractions arising from the form-factor coefficients of $B \rightarrow D \ell \nu_\ell$ and $B \rightarrow D^* \ell \nu_\ell$, as provided in Table 4.3, are described in the following. The covariance matrices for these coefficients are detailed in Ref. [103] and Ref. [104], respectively. Following the procedure used for $B \rightarrow X_u \ell \nu_\ell$ background form-factor uncertainties, variations are generated in the eigenvector basis of these covariance matrices. From these variations, 1000 varied $B \rightarrow X_c \ell \nu_\ell$ toy distributions are produced to assess the uncertainties on the partial branching fractions. The resulting uncertainties range from 0.1% to 1.7%.

$B \rightarrow X_c \ell \nu_\ell$ background: Branching fractions

In addition, uncertainties arising from both exclusively and inclusively measured $B \rightarrow X_c \ell \nu_\ell$ branching fractions are considered. The relevant branching fractions and their uncertainties are listed in Table 4.1. Varied $B \rightarrow X_c \ell \nu_\ell$ templates are created using Gaussian variations based on these uncertainties.

For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events, the uncertainties on the partial branching fraction range from 0.1% to 1.9%. Uncertainties from exclusively measured branching fractions are generally smaller compared to those from the unknown “gap” modes. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, uncertainties due to $B \rightarrow D \ell \nu_\ell$ and $B \rightarrow D^* \ell \nu_\ell$ branching fractions can reach up to 1.5% and at low q^2 are larger than the “gap” and $B \rightarrow D^{**} \ell \nu_\ell$ branching fraction uncertainties. At high q^2 , uncertainties from $B \rightarrow D^{**} \ell \nu_\ell$ and “gap” modes become more significant, surpassing those from $B \rightarrow D \ell \nu_\ell$ and $B \rightarrow D^* \ell \nu_\ell$ branching fraction uncertainties.

8.3.8 Continuum component

The continuum templates are derived from simulated continuum data reweighted to match off-resonance data, as described in Section 7.4.1. Uncertainties due to shape differences in the continuum background templates must thus be considered. These are described in this subsection.

Uncertainties related to the simulated sample size are already covered in Section 8.3.4. Additionally, Section 7.4.1 demonstrated that the reweighted distributions show negligible dependence on the M_{bc} distribution. Therefore, no additional uncertainty is assigned for this effect, as it is already included in the uncertainty from the limited off-resonance sample size.

The continuum weights, which are obtained during the reweighting procedure, depend on the number of off-resonance events in each bin. To account for uncertainties in these numbers, sets of 1000 varied continuum weights are generated for each bin. These weights are based on off-resonance event numbers, which are assumed to follow Poisson distributions. By fitting the nominal templates to the resulting toy distributions, the uncertainties on the partial branching fractions are determined. These uncertainties are significant for both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, being largest in the q^2 bins where the continuum background component is dominant. In these bins, uncertainties reach up to 34.4% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and 31.5% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. In the remaining bins, uncertainties range from 5.7% to 23.8% for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and from 13.2% to 24.3% for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events.

8.3.9 Signal model and ρ meson lineshape

The uncertainties related to shape differences in the signal templates are discussed in this subsection. This includes the residual dependence of the templates on signal form-factor uncertainties and

bin-migration effects, referred to as “signal model” effects. It also includes uncertainties associated with the modeling of the ρ meson lineshape.

Signal model

Although separate signal templates are created for different true q^2 bins, allowing the fit to adjust the q^2 spectrum, there is still a residual dependence of the true q^2 distribution within each template on the form-factor model. This includes effects from uncertainties in the form-factor coefficients listed in Tables 4.3 and 4.4 for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. To assess the uncertainties on the partial branching fractions, the approach described for $B \rightarrow X_u \ell \nu_\ell$ and $B \rightarrow X_c \ell \nu_\ell$ form-factor uncertainties is employed. As expected, the contribution to the total uncertainties is small for most q^2 bins but not negligible.

Additionally, uncertainties arise from bin migrations between true q^2 signal templates resulting from signal form-factor model variations. The same bootstrapping approach is employed as the one used for evaluating uncertainties on scale factors and signal strengths due to limited sample sizes, as detailed in Section 8.3.4. The resulting uncertainties on the partial branching fractions due to these combined effects range from 0.1% to 0.9% for all but the highest q^2 bin for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$, and the two highest q^2 bins for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. In these bins, uncertainties extend up to 4.9%.

Lineshape of the ρ meson

The model used to describe the ρ meson lineshape was discussed in Section 4.2.3. In the nominal model, interference between the ρ and ω mesons was neglected. Uncertainties related to potential interference effects, which could impact the fitted scale factors, signal strengths, and efficiencies, are considered in the following.

To evaluate these uncertainties, the fit results from Ref. [107], shown in Figure 4.1, are used. Uncertainties on the fit parameters from Ref. [107] are generated in the eigenvector basis of these parameters to cover the maximum possible deviation from the nominal model. Alternative signal templates are produced for each variation of the ρ meson lineshape. The nominal fitted scale factors are then compared with those obtained from fitting each varied toy distribution using the nominal templates.

Since the variations are produced in the eigenvector basis of the fit results (shown as the red histogram in Figure 4.1), rather than in the basis of the nominal no-interference model (shown as the blue histogram), a simultaneous variation of all fit parameter uncertainties would not be appropriate. Instead, the largest difference between the fitted scale factors is assigned as the uncertainty.

Similarly, the uncertainties on the signal efficiencies and signal strengths are evaluated. Uncertainties are assigned based on the largest differences between the nominal and varied efficiencies and signal strengths. For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ events, the resulting uncertainties are small, ranging from 0.1% to 0.6%. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events, uncertainties are larger, ranging from 0.9% to 2.0%.

Determination of $|V_{ub}|$

Finally, the measured partial branching fraction spectra obtained in Chapter 7 are used to extract values of $|V_{ub}|$ from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays. In this chapter, first the method used to derive $|V_{ub}|$ through χ^2 minimization are introduced in Section 9.1. This includes a discussion of the relevant theoretical constraints. Section 9.2 presents the results and Section 9.3 includes a discussion of the relevant uncertainties. In Section 9.4 stability checks are described.

9.1 Extraction of $|V_{ub}|$ from χ^2 fits

As described in Equations 2.15 and 2.16, the measured partial branching fractions $\Delta\mathcal{B}_i$, listed in Table 7.6 and illustrated in Figure 7.10, are used to determine $|V_{ub}|$ with the aid of theoretical form-factor predictions. The method used to extract $|V_{ub}|$ are described in this section.

Independent $|V_{ub}|$ values from form-factor fits to the partial branching fraction spectra of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays are extracted using the chi-square statistic, detailed in Section 7.1. The χ^2 to be minimized is defined as

$$\chi^2 = \sum_{i,j=1}^N (\Delta\mathcal{B}_i - \Delta\Gamma_i \tau) C_{ij}^{-1} (\Delta\mathcal{B}_j - \Delta\Gamma_j \tau) + \sum_m \chi_{\text{Theory},m}^2, \quad (9.1)$$

where N represents the number of true q^2 bins, τ denotes the B meson lifetime, and C_{ij}^{-1} is the inverse of the total covariance matrix for $\Delta\mathcal{B}_i$ in true q^2 bin i . The partial decay rate predictions for bin i , $\Delta\Gamma_i$, given in Equations 2.15 and 2.16, include the form-factor coefficients and $|V_{ub}|$. The term $\chi_{\text{Theory},m}^2$ associated with the theory calculation m imposes constraints on the shape and normalization of the form factors entering the partial decay rate, and is key in determining $|V_{ub}|$.

The following subsections will detail the fit setups used to extract $|V_{ub}|$ from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (Section 9.1.1) and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays (Section 9.1.1), including the forms of the relevant theoretical terms.

9.1.1 Setup for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays

To extract $|V_{ub}|$ from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays using the relationship described by Equation 2.15, two different sets of theoretical constraints are considered in this work. These are described in this

subsection.

First, a theory term based solely on LQCD constraints is included. In this case, LQCD predictions at high q^2 are extrapolated to low q^2 , which introduces significant uncertainties. The second option incorporates an additional theory term that includes constraints from LCSR methods. By using LCSR predictions at low q^2 , constraints can be placed on both ends of the q^2 spectrum.

Both approaches impose constraints on the BCL form-factor parameterizations. The LQCD-based term constrains the form-factor coefficients, while the LCSR term directly constrains the magnitudes of the form factors. Details of these theory terms are discussed separately below. Regardless of the theory term used, m_R is set to 5.325 GeV [12] when evaluating the inverse Blaschke factor in the form-factor expansion described by Equation 2.19.

LQCD constraints: For $|V_{ub}|$ extraction from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays using LQCD constraints, nuisance parameters are included for three $f_+(q^2)$ and two $f_0(q^2)$ BCL form-factor coefficients b_k^+ and b_k^0 , respectively, contained within the form-factor expansion in Equation 2.19. The χ_{LQCD}^2 term is defined as:

$$\chi_{\text{LQCD}}^2 = \sum_{k,l=1}^5 (b_k - b_k^{\text{LQCD}}) C_{\text{LQCD},kl}^{-1} (b_l - b_l^{\text{LQCD}}), \quad (9.2)$$

where b_k^{LQCD} and $C_{\text{LQCD},kl}^{-1}$ are the constraints and inverse covariance matrix for the BCL form-factor coefficients provided in the February 2023 version of the FLAG 21 review [14]. These constraints combine results from the FNAL/MILC [15], RBC/UKQCD [16], and JLQCD [17] collaborations.

LCSR constraints: Alternatively, $|V_{ub}|$ can be extracted from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays by including LCSR constraints on $f_+(q^2)$ and $f_0(q^2)$ at five discrete q^2 points. In this case, the additional χ_{LCSR}^2 term is given by:

$$\chi_{\text{LCSR}}^2 = \sum_{k,l=1}^{10} (f_k - f_k^{\text{LCSR}}) C_{\text{LCSR},kl}^{-1} (f_l - f_l^{\text{LCSR}}), \quad (9.3)$$

where f_k^{LCSR} are the LCSR-constrained form factors, and $C_{\text{LCSR},kl}^{-1}$ is the corresponding inverse covariance matrix provided by Ref. [18]. This approach directly constrains the BCL form factors f_k for $f_+(q^2)$ and $f_0(q^2)$, incorporating the relevant inverse covariance matrix.

9.1.2 Setup for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays

To extract $|V_{ub}|$ from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays using the relationship described by Equation 2.16, a single set of theory constraints is used. The setup is described in this subsection.

Since no LQCD predictions are available for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, only constraints on the BSZ form-factor parameterization derived from LCSR predictions are incorporated. Unlike the LCSR constraints used for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays, which constrain the form factors directly, the LCSR constraints for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays target the BSZ form-factor coefficients, similar to the LQCD constraints used for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays. The m_R factors within the inverse Blaschke factors in

Equation 2.20 are set to 5.724 GeV for the form-factor expansions of $A_1(q^2)$ and $A_2(q^2)$, and to 5.325 GeV for the form-factor expansion of $V(q^2)$ [13].

LCSR constraints: In extracting $|V_{ub}|$ from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, LCSR constraints are incorporated as nuisance parameters for the BSZ form-factor coefficients. Specifically, constraints on two form-factor coefficients b_k^i each for $A_1(q^2)$, $A_2(q^2)$, and $V(q^2)$ are included, resulting in a total of six BSZ form-factor coefficients. These coefficients are part of the form-factor expansion described by Equation 2.20. The χ_{LCSR}^2 term is given by:

$$\chi_{\text{LCSR}}^2 = \sum_{k,l=1}^6 (b_k - b_k^{\text{LCSR}}) C_{\text{LCSR},kl}^{-1} (b_l - b_l^{\text{LCSR}}), \quad (9.4)$$

where b_k^{LCSR} are the theory constraints on the form-factor coefficients from Ref. [13], and $C_{\text{LCSR},kl}^{-1}$ is the corresponding inverse covariance matrix.

9.2 $|V_{ub}|$ fit results

By performing the fits described in the previous section, three values for $|V_{ub}|$, presented in this section, are obtained. For all quoted $|V_{ub}|$ results in this section, the uncertainties are presented as follows: the first is statistical, the second is systematic, and the third is theoretical. Details on the composition and evaluation of these uncertainties are provided in the next section.

The first value of $|V_{ub}|$ from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays, using only LQCD constraints, is:

$$|V_{ub}|_{B \rightarrow \pi \ell \nu_\ell} = (3.93 \pm 0.09 \pm 0.13 \pm 0.19) \times 10^{-3}.$$

When including LCSR constraints, the result for $|V_{ub}|$ from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays becomes:

$$|V_{ub}|_{B \rightarrow \pi \ell \nu_\ell} = (3.73 \pm 0.07 \pm 0.07 \pm 0.16) \times 10^{-3}.$$

Figure 9.1 shows the partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays as a function of q^2 . The fitted differential rates are also displayed, along with the one, two, and three standard-deviation uncertainty bands. The χ^2 per degree of freedom (DOF) for both fits is 8.4/7, corresponding to a p-value of 0.40. The DOF is calculated as the number of q^2 bins minus the number of fit parameters, which is $13 - 6 = 7$. The measured values of $|V_{ub}|$ and the BCL form-factor coefficients from the fits to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ partial branching fraction spectrum are provided in Table 9.1. The elements of the full correlation matrices for these values are provided in Tables C.4 and C.5 in Appendix C.

The two $|V_{ub}|$ results obtained from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays are consistent, differing by only 0.7 standard deviations. Both results are in agreement with previous exclusive $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ measurements [9], described in Section 2.3, with the combined LQCD and LCSR $|V_{ub}|$ fit result agreeing exceptionally well, within 0.3 standard deviations. However, while the LQCD $|V_{ub}|$ fit result agrees equally well with both the exclusive and inclusive world averages within 0.9 standard deviations, the combined LQCD and LCSR $|V_{ub}|$ fit result is lower and deviates by 1.9 standard deviations from the inclusive world average. This indicates a similar tension as observed between the exclusive and inclusive world averages.

The third value for $|V_{ub}|$ is obtained from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays using LCSR constraints:

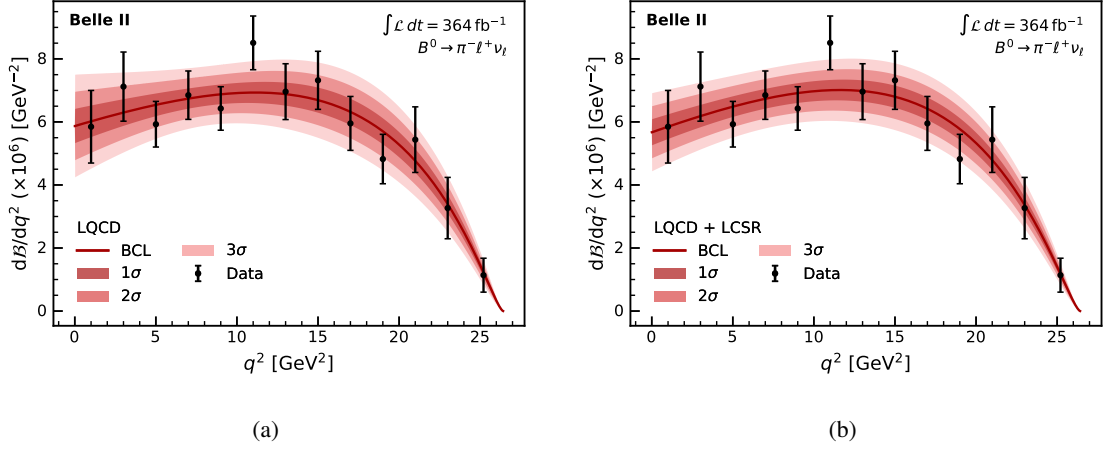


Figure 9.1: Measured partial branching fractions as a function of q^2 for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays. The fitted differential rates are shown, along with one, two, and three standard-deviation uncertainty bands for fits using constraints from (a) LQCD and (b) LQCD plus LCSR predictions.

Table 9.1: Measured values of $|V_{ub}|$ and the BCL form-factor coefficients with total uncertainties from fits to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ partial branching fraction spectrum. The χ^2/DOF is also provided.

$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$		
	LQCD	LQCD + LCSR
$ V_{ub} (10^{-3})$	3.93 ± 0.25	3.73 ± 0.19
$f_+(q^2)$	b_0^+	0.42 ± 0.02
	b_1^+	-0.52 ± 0.05
	b_2^+	-0.81 ± 0.21
$f_0(q^2)$	b_0^0	0.02 ± 0.25
	b_1^0	-1.43 ± 0.08
χ^2/DOF	8.39/7	8.36/7

$$|V_{ub}|_{B \rightarrow \rho \ell \nu_\ell} = (3.19 \pm 0.12 \pm 0.17 \pm 0.26) \times 10^{-3}.$$

Figure 9.2 shows the partial branching fractions of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays as a function of q^2 . The fitted differential rate is also displayed, with one, two, and three standard-deviation uncertainty bands. The χ^2 per DOF is 3.9/3, corresponding to a p-value of 0.27. The DOF is $10 - 7 = 3$, given that partial branching fractions were measured in ten q^2 bins and seven fit parameters are included. The measured values of $|V_{ub}|$ and the BSZ form-factor coefficients from the fit to the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ partial branching fraction spectrum are summarized in Table 9.2. The elements of the full correlation matrix for these values are provided in Table C.6 in Appendix C.

The $|V_{ub}|$ result obtained from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays is lower compared to the $|V_{ub}|$ results from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays. It agrees only within 1.8 and 1.4 standard deviations with the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ results obtained using LQCD and combined LQCD and LCSR fits, respectively. However, the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ $|V_{ub}|$ result is consistent with previous $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ measurements [10], as described

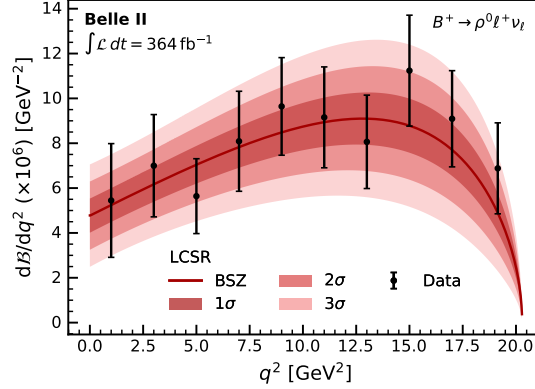


Figure 9.2: Measured partial branching fractions as a function of q^2 for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays. The fitted differential rate is shown along with one, two, and three standard-deviation uncertainty bands for a fit using LCSR constraints.

Table 9.2: Measured values of $|V_{ub}|$ and the BSZ form-factor coefficients with total uncertainties from a fit to the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ partial branching fraction spectrum. The χ^2/DOF is also provided.

$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$		
LCSR		
$ V_{ub} (10^{-3})$		3.19 ± 0.33
$A_1(q^2)$	$b_0^{A_1}$	0.27 ± 0.03
	$b_1^{A_1}$	0.34 ± 0.13
$A_2(q^2)$	$b_0^{A_2}$	0.29 ± 0.03
	$b_1^{A_2}$	0.66 ± 0.17
$V(q^2)$	b_0^V	0.33 ± 0.03
	b_1^V	-0.93 ± 0.17
χ^2/DOF		$3.85/3$

in Section 2.3, within 0.5 standard deviations. Since this result is higher than previous measurements, the tensions with the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ result and the inclusive and exclusive ($B^0 \rightarrow \pi^- \ell^+ \nu_\ell$) world averages are reduced to 2.7 and 1.3 standard deviations, respectively.

9.3 Uncertainties

The uncertainties on the $|V_{ub}|$ results quoted in the previous section consist of statistical, systematic, and theoretical components. These uncertainties are evaluated using a toy method and described in this section. To determine the statistical and systematic components, toy datasets are generated by varying the partial branching fractions according to their statistical or systematic covariance matrices. This process includes all sources of uncertainty discussed in Chapter 8.

200 variations are created and a χ^2 fit on each toy dataset is performed, using the nominal covariance matrix and theory constraints for the χ^2 calculation. The width of the distribution of $|V_{ub}|$ results from these fits gives the statistical or systematic error. Similarly, for the theoretical uncertainty, the theory constraints are varied using their respective covariance matrices, rather than varying the input partial branching fractions.

All three $|V_{ub}|$ results are limited by the theoretical uncertainties. For the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ results, the relative theoretical uncertainties are 4.8% for the LQCD fit and 4.3% for the combined LQCD and LCSR fit. For the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ result, the theoretical uncertainty is larger, at 8.2%.

Adding the LCSR constraint to the LQCD constraint for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays reduces the relative theoretical uncertainty. This reduction is because LCSR predictions provide low q^2 information, decreasing the dependence on low q^2 extrapolations from LQCD. This reduction is evident in the b_0^0 coefficient of the $f_0(q^2)$ form factor, as shown in Table 9.1. While improvements in the uncertainties of the $f_+(q^2)$ form-factor coefficients are minor with the inclusion of LCSR information, it significantly constrains the $f_0(q^2)$ form factor.

Including LCSR information also improves the statistical and systematic uncertainty contributions for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ fits. To understand this effect, the relative contribution of each uncertainty source discussed in Chapter 8, is assessed. Using the toy method, the input partial branching fractions are varied according to the systematic uncertainty matrix corresponding to each uncertainty source. The relative uncertainties on $|V_{ub}|$ from various systematic uncertainty sources are shown in Table 9.3.

Table 9.3: Summary of fractional uncertainties in % by source on the extracted $|V_{ub}|$ values.

	$B^0 \rightarrow \pi^- \ell^+ \nu_\ell$		$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
	LQCD	LQCD + LCSR	LCSR
Detector effects	0.64	0.24	0.44
Beam energy	0.05	0.03	0.09
Simulated sample size	1.51	0.78	1.41
BDT efficiency	0.31	0.21	0.28
Physics constraints	0.61	0.43	0.88
Signal model	0.38	0.13	0.41
ρ lineshape	0.26	0.21	0.13
Nonres. $B \rightarrow \pi\pi\ell\nu_\ell$	0.43	0.11	1.97
DFN parameters	0.64	0.32	0.88
$B \rightarrow X_u \ell \nu_\ell$ model	0.61	0.40	1.56
$B \rightarrow X_c \ell \nu_\ell$ model	0.51	0.43	0.50
Continuum	2.39	1.37	4.91
Total systematic	3.26	1.91	5.33
Statistical	2.31	1.82	3.76
Theory	4.83	4.29	8.15
Total	6.40	5.13	10.34

For both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ results, the largest contribution to the systematic uncertainty is from the limited size of the off-resonance data sample. For the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ LQCD

result and the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ result, this contribution alone exceeds the statistical uncertainty. Another significant contribution to the uncertainties comes from the limited simulated sample sizes. For the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ result, the uncertainties from modeling nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ decays and the $B \rightarrow X_u \ell \nu_\ell$ background are also significant.

To understand the reduction in relative statistical and systematic uncertainties on $|V_{ub}|$ from adding LCSR constraints, the uncertainties on the input partial branching fractions given in Table 8.1 are examined. The statistical uncertainty as a function of q^2 is analyzed to assess its importance. For systematic uncertainties, the behavior of dominant sources as a function of q^2 are examined. These include uncertainties from limited sizes of off-resonance and simulated samples.

The uncertainties are largest at high q^2 , but they also dominate at low q^2 . While LQCD constraints can constrain high q^2 uncertainties they cannot constrain those at low q^2 . Including LCSR constraints reduces this dependence, similarly to the reduction observed for theoretical uncertainties. This results in decreased relative statistical and systematic uncertainties on $|V_{ub}|$.

9.4 Stability tests

The stability of the $|V_{ub}|$ results is assessed by performing fits with various q^2 cut-off values, described in this section.

For the LQCD fit of the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ partial branching fraction spectrum, the lowest q^2 bin is sequentially removed to examine the extrapolation of the LQCD constraints to low q^2 . Figure 9.3(a) displays the extracted $|V_{ub}|$ values as a function of the q^2_{\min} cut-off values used in the χ^2 fit. These results are compared with the exclusive and inclusive world averages of $|V_{ub}|$ [9], with the left-most point representing the nominal $|V_{ub}|$ result.

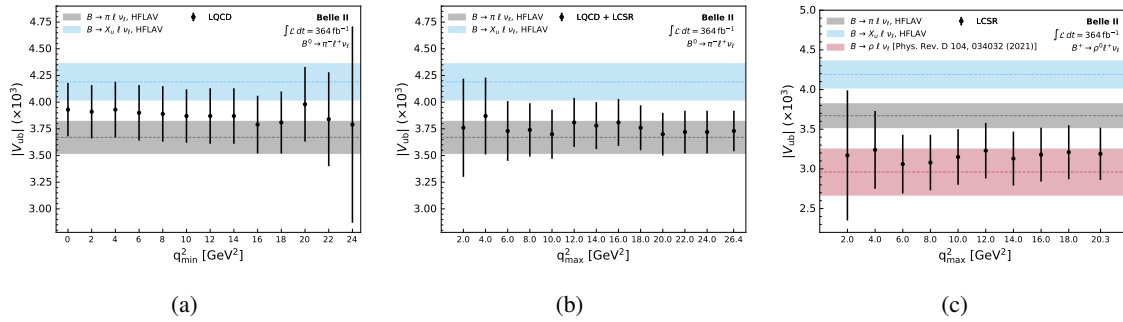


Figure 9.3: Measured $|V_{ub}|$ values for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (a, b) and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (c) decays with varying cut-off values of q^2 in the fit. In (a) results with LQCD constraints, where minimum cut-off values are used, are shown. In (b) and (c) results with LCSR constraints, where maximum cut-off values are investigated, are shown. The results are compared with the world average $|V_{ub}|$ values from $B \rightarrow \pi \ell \nu_\ell$ and inclusive $B \rightarrow X_u \ell \nu_\ell$ decays from Ref. [9]. In (c) a further comparison to the $|V_{ub}|$ result obtained in Ref. [10] from $B \rightarrow \rho \ell \nu_\ell$ measurements is provided.

For the combined LQCD and LCSR fit to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ partial branching fraction spectrum and for the fit to the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ partial branching fraction spectrum, the dependence on the high q^2 extrapolation is tested by sequentially removing the largest q^2 bins. The $|V_{ub}|$ results for different q^2_{\max} cut-off values are shown in Figure 9.3(b, c). The right-most points in these figures correspond to

the nominal $|V_{ub}|$ results. In addition to comparing these results with exclusive and inclusive world averages, the fit-stability $|V_{ub}|$ results for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays are compared with a previous $|V_{ub}|$ measurement [10]. For all three fits, the $|V_{ub}|$ results remain stable with varying q^2 cut-off values.

Additionally, the self-consistency of the form-factor coefficients is verified by repeating the signal extraction described in Chapter 7 using modified signal templates with q^2 shapes that match the measured form-factor coefficients. After updating the branching fraction spectra with these modified templates, the $|V_{ub}|$ fits are re-performed and consistent results are obtained.

Conclusion and outlook

This thesis presented a simultaneous measurement of untagged $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays using data corresponding to an integrated luminosity of 364 fb^{-1} recorded by the Belle II detector. Signal events were selected, and the dominant continuum and $B\bar{B}$ backgrounds were suppressed using boosted decision trees. A novel approach was developed to simultaneously extract signal yields of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events in bins of true q^2 , effectively reducing the dependence on the modeling of the q^2 resolution. To achieve this, a simultaneous extended binned maximum likelihood fit to the distributions of ΔE , M_{bc} , and reconstructed q^2 was constructed.

The signal yields were used to determine the partial and total branching fractions. The branching fraction of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ was found to be $(1.516 \pm 0.042(\text{stat}) \pm 0.059(\text{syst})) \times 10^{-4}$, while that of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ is $(1.625 \pm 0.079(\text{stat}) \pm 0.180(\text{syst})) \times 10^{-4}$. These results are consistent with the world averages [37] and are primarily limited by the size of the off-resonance data sample and the description of the nonresonant $B \rightarrow X_u \ell \nu_\ell$ background.

Form-factor fits to the measured partial branching fraction spectra were performed to extract values of the CKM matrix element magnitude $|V_{ub}|$ from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays. For $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays, two sets of theoretical constraints were applied. Using LQCD constraints from the FLAG 21 review, updated in February 2023 [14], $|V_{ub}|$ was determined to be $(3.93 \pm 0.09(\text{stat}) \pm 0.13(\text{syst}) \pm 0.19(\text{theo})) \times 10^{-3}$. When additional LCSR constraints were included [18], the result became $(3.73 \pm 0.07(\text{stat}) \pm 0.07(\text{syst}) \pm 0.16(\text{theo})) \times 10^{-3}$. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, applying LCSR constraints from Ref. [13] resulted in a $|V_{ub}|$ value of $(3.19 \pm 0.12(\text{stat}) \pm 0.17(\text{syst}) \pm 0.26(\text{theo})) \times 10^{-3}$.

The $|V_{ub}|$ results are dominated by theoretical uncertainties, highlighting the need for further refinement of these models in the future. Experimentally, the $|V_{ub}|$ results are systematically limited by the size of the off-resonance data sample and the available simulated samples. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, additional significant uncertainties arise from the modeling of $B \rightarrow X_u \ell \nu_\ell$ decays, particularly from nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ decays.

Reducing uncertainties due to limited sample sizes could primarily be achieved by increasing the sample sizes. The dominant off-resonance sample uncertainty could also be mitigated by improving the simulation of continuum backgrounds, allowing the off-resonance sample to be used only for validation. To address uncertainties related to the modeling of $B \rightarrow X_u \ell \nu_\ell$ decays, exploring alternatives to the hybrid model may be beneficial. Developing such techniques will require more inclusive and exclusive measurements of $B \rightarrow X_u \ell \nu_\ell$ decays, particularly nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ decays, in the future.

One of the motivations for the work described by this thesis was to test the discrepancy between the

exclusive and inclusive values of $|V_{ub}|$ [9]. Both $|V_{ub}|$ results from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays are consistent with each other and with the exclusive world average. The $|V_{ub}|$ result constrained only by LQCD predictions also agrees with the inclusive world average, which is not true when LCSR constraints are included. The $|V_{ub}|$ result from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays is lower than those from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays, yet consistent with previous $B \rightarrow \rho \ell \nu_\ell$ measurements in Ref. [10].

The discrepancies observed in $|V_{ub}|$ results when using LCSR constraints to the inclusive world averages, potentially suggest issues within the theoretical predictions or their associated uncertainties. Although the LQCD-constrained result aligns well with both the exclusive and inclusive world averages, slightly reducing the overall discrepancy, it does not resolve it. Consequently, it is likely that using the same experimental techniques in the future may not lead to further insights. Similarly, the measurements from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays show the same discrepancies as observed in previous studies. To improve the theoretical predictions of form factors for $B \rightarrow \rho \ell \nu_\ell$ decays, a simultaneous analysis of $B \rightarrow \pi \pi \ell \nu_\ell$ and $B \rightarrow \rho \ell \nu_\ell$ decays could help in distinguishing between resonant and nonresonant contributions.

Further studies

This appendix presents additional studies that support the results discussed in this thesis.

Section A.1 describes the optimization of the bremsstrahlung correction parameters. Section A.2 focuses on the selection of optimal thresholds for track momentum and cluster energy of the ROE particles. Section A.3 introduces the reconstruction and application of the $B^+ \rightarrow J/\psi K^+$ control mode. Finally, Section A.4 presents additional distributions from the reweighting of simulated continuum events.

A.1 Bremsstrahlung parameter optimization

Lepton candidates are corrected for bremsstrahlung effects as described in Section 5.2.2. This section details the optimization of the bremsstrahlung correction parameters, which include the maximum opening angle and cluster energies.

Preliminary selections of 0.8 GeV and 1.2 GeV are applied to the lab-frame momentum of lepton candidates in the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, respectively. The resolution of the lepton momentum is evaluated for various angle and energy thresholds. Gaussian functions are fitted to each resolution peak. Given that the resolution exhibits a Gaussian core with long tails, the figure of merit (FOM), representing the fraction of total events lying outside the fitted peak width, is minimized to determine the optimal parameters.

Figure A.1 shows the FOM as a function of cluster energy and opening angle in a two-dimensional plot for both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events. The optimal thresholds are marked, and the projections of the FOM as functions of energy and angle at the optimal points are also presented. While the optimal opening angle is consistent for both $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, the optimal cluster energy differs due to the distinct lepton momentum thresholds in each mode.

A.2 ROE track and cluster selection optimization

Section 5.3 discusses the selection criteria for the tracks and clusters contributing to the ROE. This section focuses on the optimization of these selections.

Initially, the agreement between simulated and collision data is examined for the transverse momentum (p_t) distribution of all ROE tracks. Figure A.2 presents these distributions, where

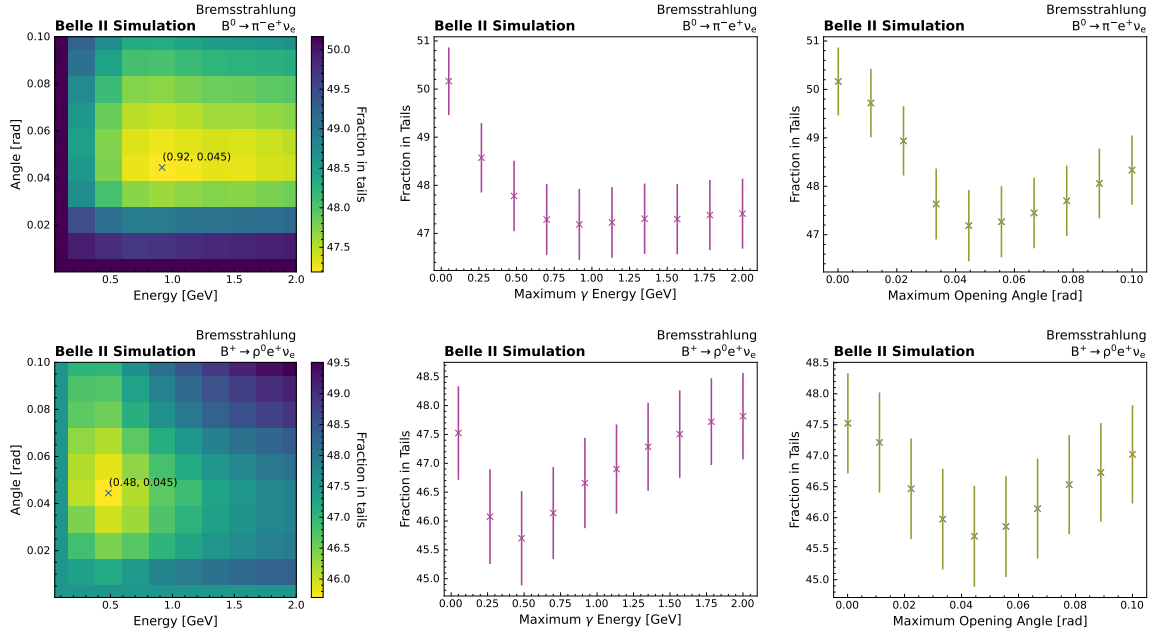


Figure A.1: Two-dimensional plots of the FOM as a function of cluster energy and opening angle (left) for $B^0 \rightarrow \pi^- e^+ \nu_e$ (top) and $B^+ \rightarrow \rho^0 e^+ \nu_e$ events (bottom) with optimal points marked. The projections of the FOM as functions of cluster energy (middle) and opening angle (right) at the optimal points are also shown.

off-resonance data (*e18*) represents the simulated continuum component, while the $B\bar{B}$ backgrounds are derived from simulation. The analysis shows that no tighter p_t track selection than $p_t > 0.05$ GeV is necessary, as the agreement between simulated and collision data is good.

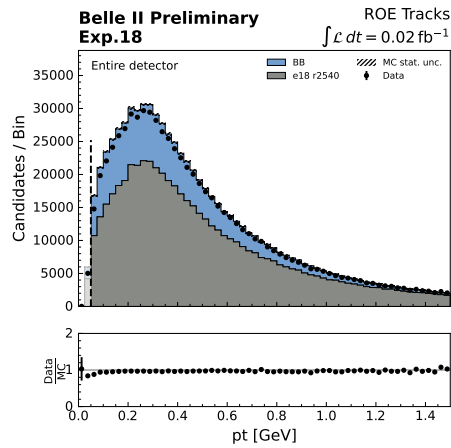


Figure A.2: Distributions of track p_t in simulated and collision data. Off-resonance data (*e18*) represents the simulated continuum component. The threshold is indicated.

Subsequently, the cluster energy spectra in the forward, barrel, and backward detector regions are compared between simulated and collision data. Figure A.3 shows these distributions, with

off-resonance data used to represent the simulated continuum component. The comparison reveals that the low-energy region does not match well between on-resonance and off-resonance data across all detector regions. This discrepancy is likely due to variations in beam-induced background levels depending on accelerator conditions. To address this, thresholds for cluster energy are applied based on the detector region. These thresholds are chosen to ensure that a maximum of 5% of the remaining photons are attributed to beam-induced backgrounds, as indicated by simulation.

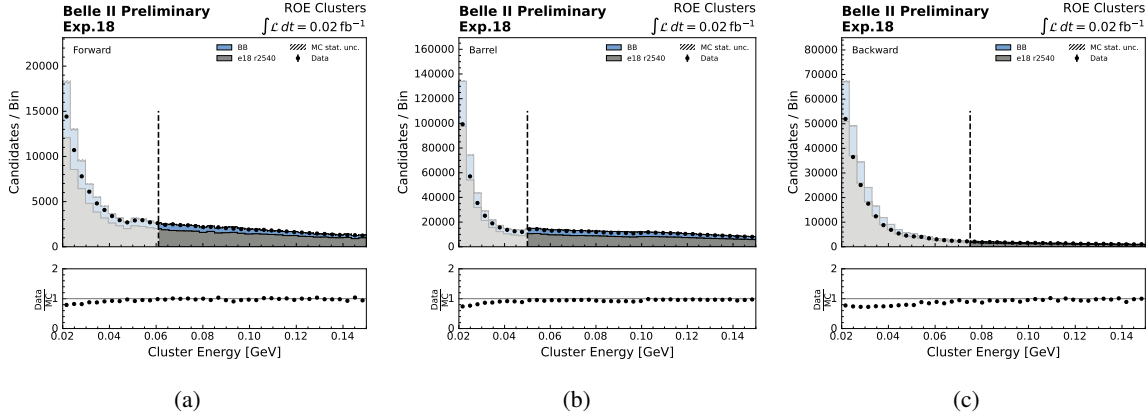


Figure A.3: Distributions of cluster energy in (a) forward, (b) barrel and (c) backward detector regions in simulated and collision data. Off-resonance data (*e18*) represents the simulated continuum component. The thresholds are indicated.

A.3 Control mode

To validate the reconstruction of q^2 and test the consistency of the signal efficiency of the BDT classifier selection between simulated and collision data, a control mode is introduced, which is described in this section. Section A.3.1 describes the reconstruction of the $B^+ \rightarrow J/\psi K^+$ control mode, while the validation study for the q^2 resolution is described in Section

A.3.1 Reconstruction

This subsection describes the reconstruction of the $B^+ \rightarrow J/\psi K^+$ control mode. First, $B^+ \rightarrow J/\psi K^+$ events, where the J/ψ decays to two muons, are fully reconstructed. Subsequently, $B^+ \rightarrow \mu K^+$ events are reconstructed, ensuring that these events form a subset of the fully reconstructed events. The selection criteria are outlined in Table A.1 and are based on ongoing dedicated studies at the Belle II experiment.

Muon and kaon candidates are required to have polar angles within the CDC acceptance and transverse momenta greater than 0.05 GeV. Additionally, their impact parameters must satisfy $dr < 1.0$ cm and $|dz| < 3.0$ cm. Thresholds are applied to the PID ratios for both muon and kaon candidates, and PID ratio corrections are made as detailed in Section 4.3.1. Moreover, the muon candidate's momentum must exceed 0.8 GeV.

J/ψ candidates reconstructed from two muons must have an invariant mass within the range: $m_{ll} \in [2.96, 3.19]$ GeV. Fully reconstructed B meson candidates are required to pass selections on M_{bc} and

Table A.1: Requirements for particle candidates and the ROE in the reconstruction of the $B^+ \rightarrow J/\psi K^+$ control mode.

Selection	
Muon	muon PID > 0.9 and $p > 0.8$ GeV
Kaon	kaon PID > 0.6
J/ψ	$m_{ll} \in [2.96, 3.19]$ GeV
B meson	$M_{bc} > 5.2$ GeV, $ \Delta E < 0.2$ GeV, $ \Delta M < 0.05$ GeV
ROE	see Section 5.3

ΔE , with an additional constraint on the difference between the reconstructed and PDG mass of the B meson, ΔM [37]. The ROE requirements for the fully reconstructed events align with those for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, as described in Section 5.3. Finally, a random selection of the best candidate is performed.

A.3.2 Validation of q^2 reconstruction

This subsection describes the validation study for the reconstruction of q^2 using the control mode. After reconstructing the control mode as outlined above, the q^2 resolutions in collision and simulated data are compared. While the root-mean-square (RMS) of the q^2 distribution in collision data is found to be 0.505 GeV^2 , in simulated data it is 0.526 GeV^2 , corresponding to a 4% difference. The two distributions are shown in Figure A.4(a). The *mixed* and *charged* components are the backgrounds from neutral and charged B meson events, respectively.

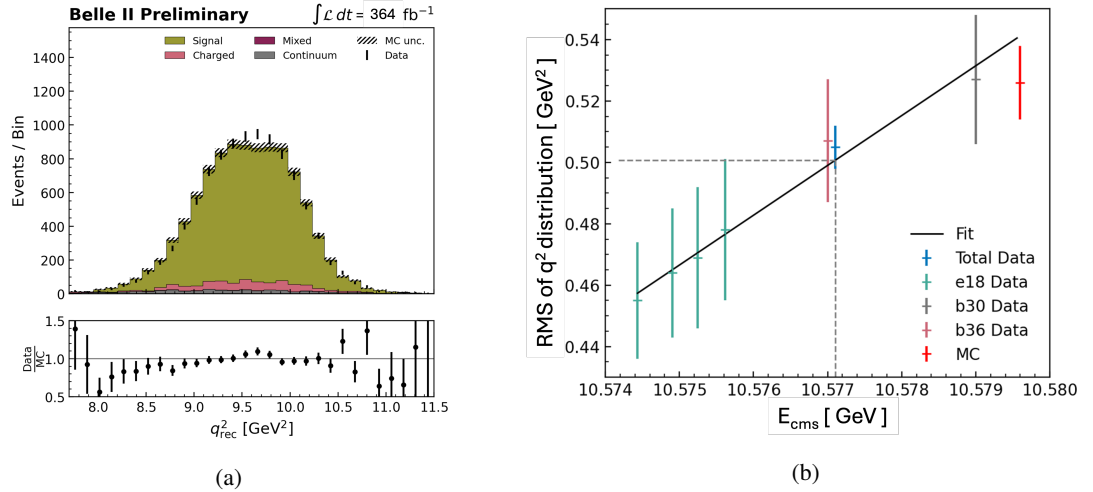


Figure A.4: (a) Distributions of reconstructed q^2 in the control mode in simulated and collision data, with the lower panel showing the ratio between collision and simulated data. (b) Plot of RMS values of the q^2 distributions as a function of E_{cms} for different collision datasets.

A dependence of the q^2 resolution on the CM energy E_{cms} is observed. The E_{cms} in simulated

data is 10.5796 GeV, whereas the mean E_{cms} in collision data is 10.5771 GeV. Three subsets of the collision data with different E_{cms} ranges are examined: e18 (low E_{cms}), b30 (approximately the mean data E_{cms}), and b36 (approximately the simulated E_{cms}). The RMS value of the q^2 distribution for each subset is determined, showing a linear trend with respect to E_{cms} . Figure A.4(b) displays the RMS values of the q^2 distribution as a function of E_{cms} .

Since mismodeling of the q^2 distribution could impact the signal template shapes during signal extraction, the effect of a narrower q^2 resolution in collision data must be investigated. To simulate a 4% difference in resolution, as detailed in Section 8.3.2, a 4% difference is assumed for all q^2 bins. The validity of this assumption needs to be investigated, since the control mode occurs at a single q^2 value (9.59 GeV²).

To validate this assumption, $\Upsilon(4S) \rightarrow B\bar{B}(\rightarrow \pi/\rho \ell \nu)$ events are generated using the TGenPhaseSpace class [130, 131]. Two sets of events are generated: one using the mean E_{cms} from collision data and the other using the E_{cms} from simulated data. The combined Diamond Frame and ROE method is then used to reconstruct q^2 . Comparing the q^2 resolutions, a smaller resolution is observed with the mean collision data E_{cms} than with the simulated data E_{cms} , consistent with the control mode findings.

To examine the effect in different q^2 bins, generated events are split into true q^2 bins. The fractional difference between the RMS values of q^2 distributions in samples generated with collision and simulated E_{cms} values is plotted against true q^2 , as shown for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events in Figures A.5(a) and A.5(b), respectively. The results are consistent with an overall 4% fractional difference in RMS values across all true q^2 bins.

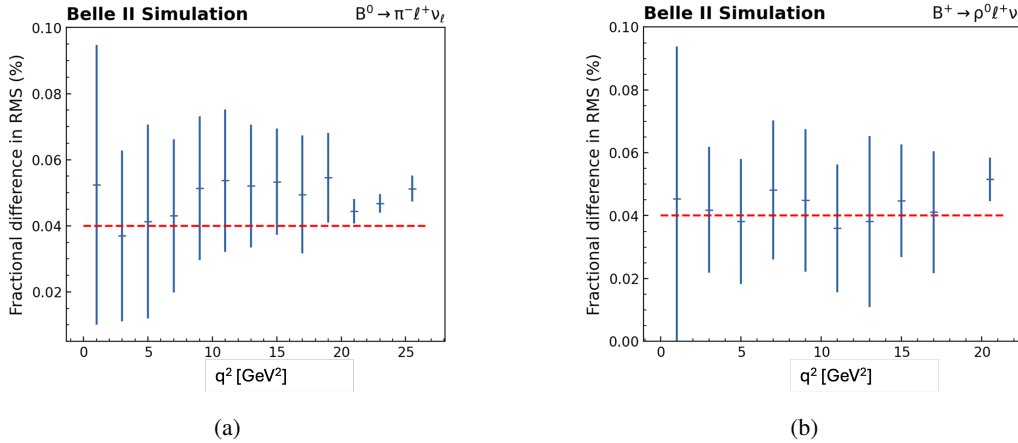


Figure A.5: Fractional differences between the RMS values of q^2 distributions in generated samples using collision and simulated E_{cms} values as a function of true q^2 for (a) $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and (b) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ events.

A.4 Continuum reweighting

This section presents the distributions of ΔE and M_{bc} in each reconstructed q^2 bin for both the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ modes, using off-resonance and reweighted simulated data. The reweighting procedure is detailed in Section 7.4.1.

Figures A.6 and A.7 display the distributions of ΔE and M_{bc} , respectively, for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode. Figures A.8 and A.9 show the distributions of ΔE and M_{bc} , respectively, for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode.

mode. The binning chosen corresponds to that used during signal extraction, and the scale factors provided refer to the scale factors applied to the lower and higher ΔE bins during reweighting.

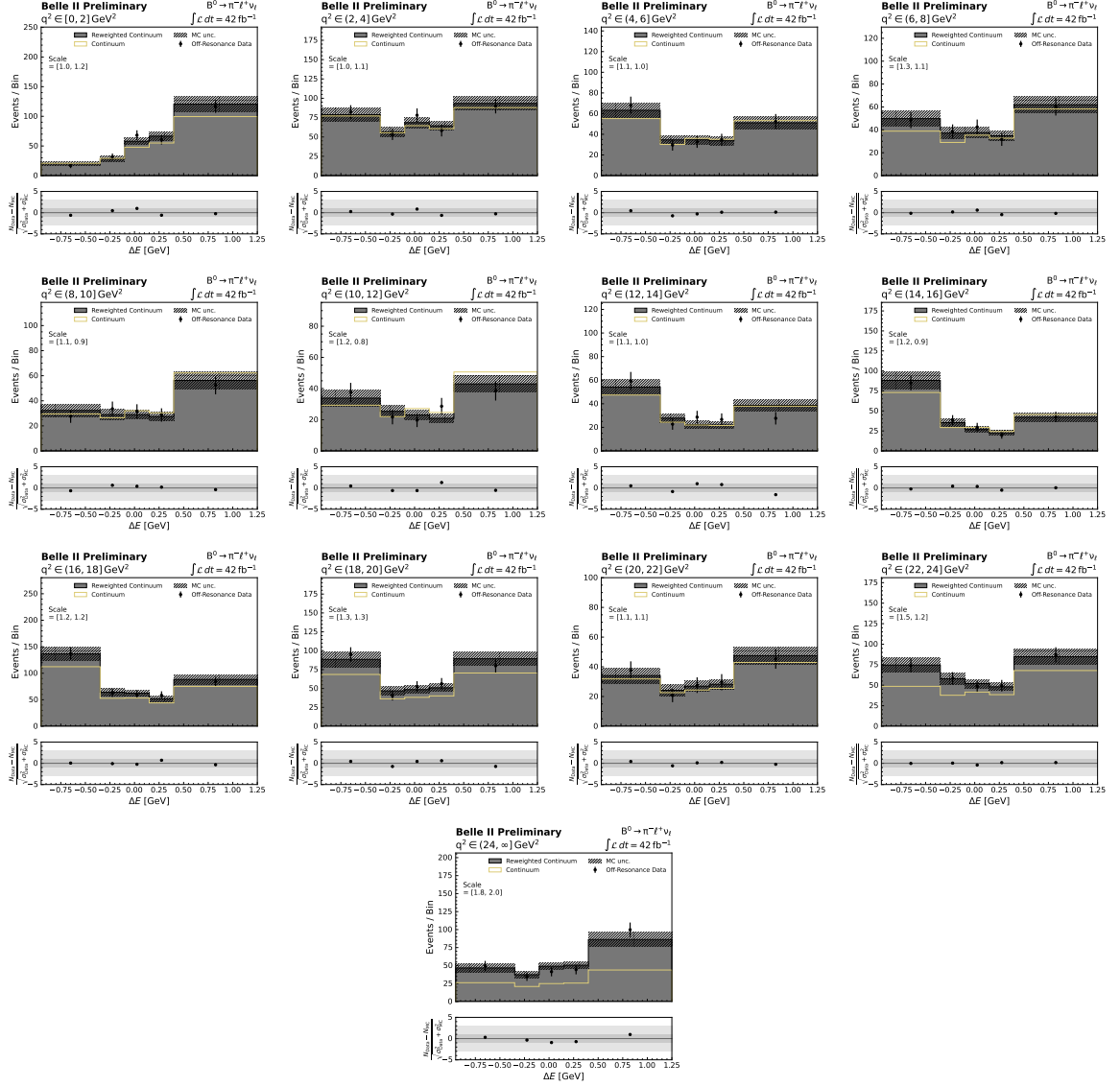


Figure A.6: Distributions of ΔE for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode in off-resonance and reweighted simulated data (grey) for each reconstructed q^2 bin. The scale factors for the lower and higher ΔE bins used in reweighting are indicated. The lower panels show the difference between collision and reweighted simulated data divided by the combined uncertainty. The olive line represents the simulated distribution before reweighting.

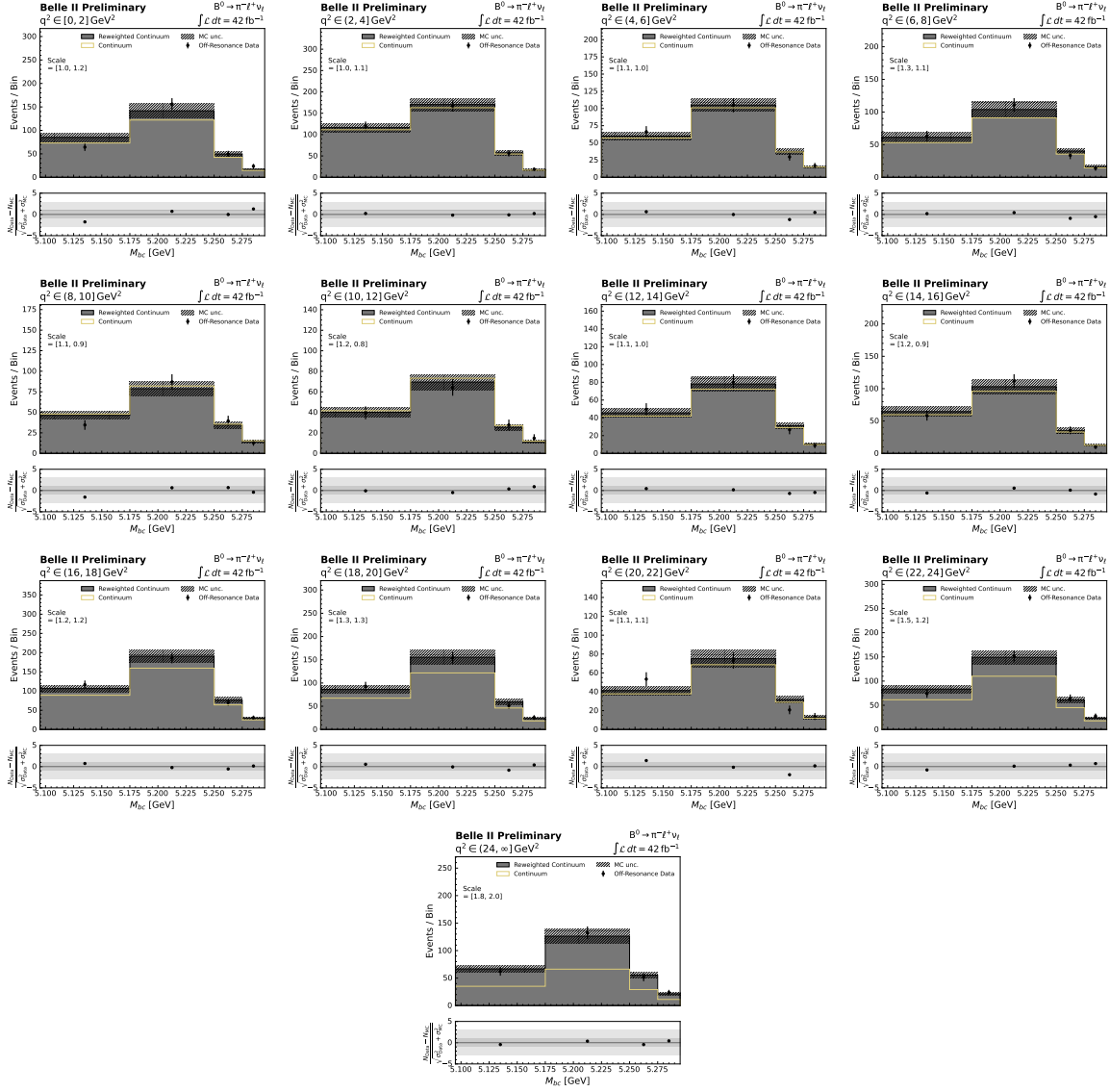


Figure A.7: Distributions of M_{bc} for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode in off-resonance and reweighted simulated data (grey) for each reconstructed q^2 bin. The scale factors for the lower and higher ΔE bins used in reweighting are indicated. The lower panels show the difference between collision and reweighted simulated data divided by the combined uncertainty. The olive line represents the simulated distribution before reweighting.

Appendix A Further studies

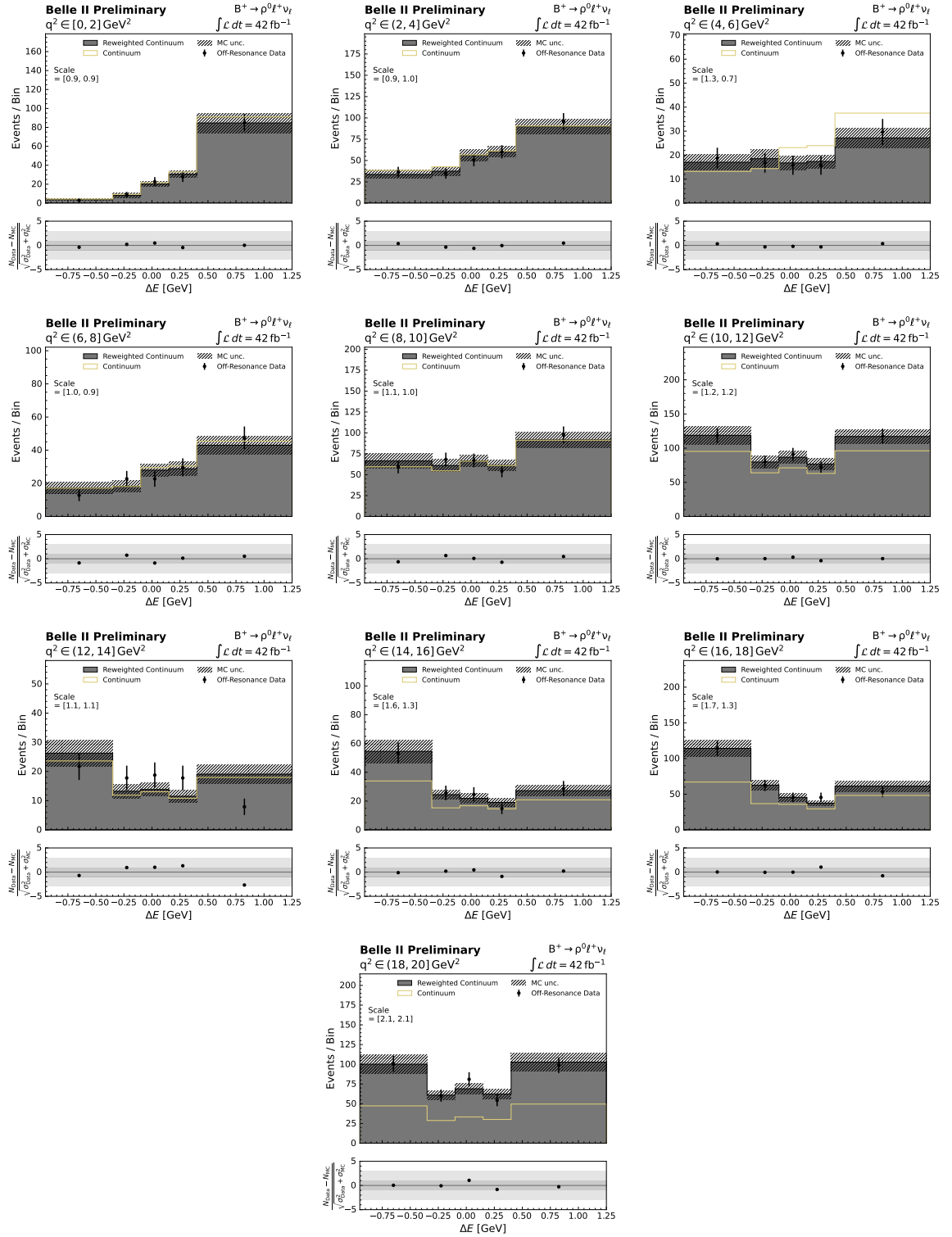


Figure A.8: Distributions of ΔE for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode in off-resonance and reweighted simulated data (grey) for each reconstructed q^2 bin. The scale factors for the lower and higher ΔE bins used in reweighting are indicated. The lower panels show the difference between collision and reweighted simulated data divided by the combined uncertainty. The olive line represents the simulated distribution before reweighting.

A.4 Continuum reweighting

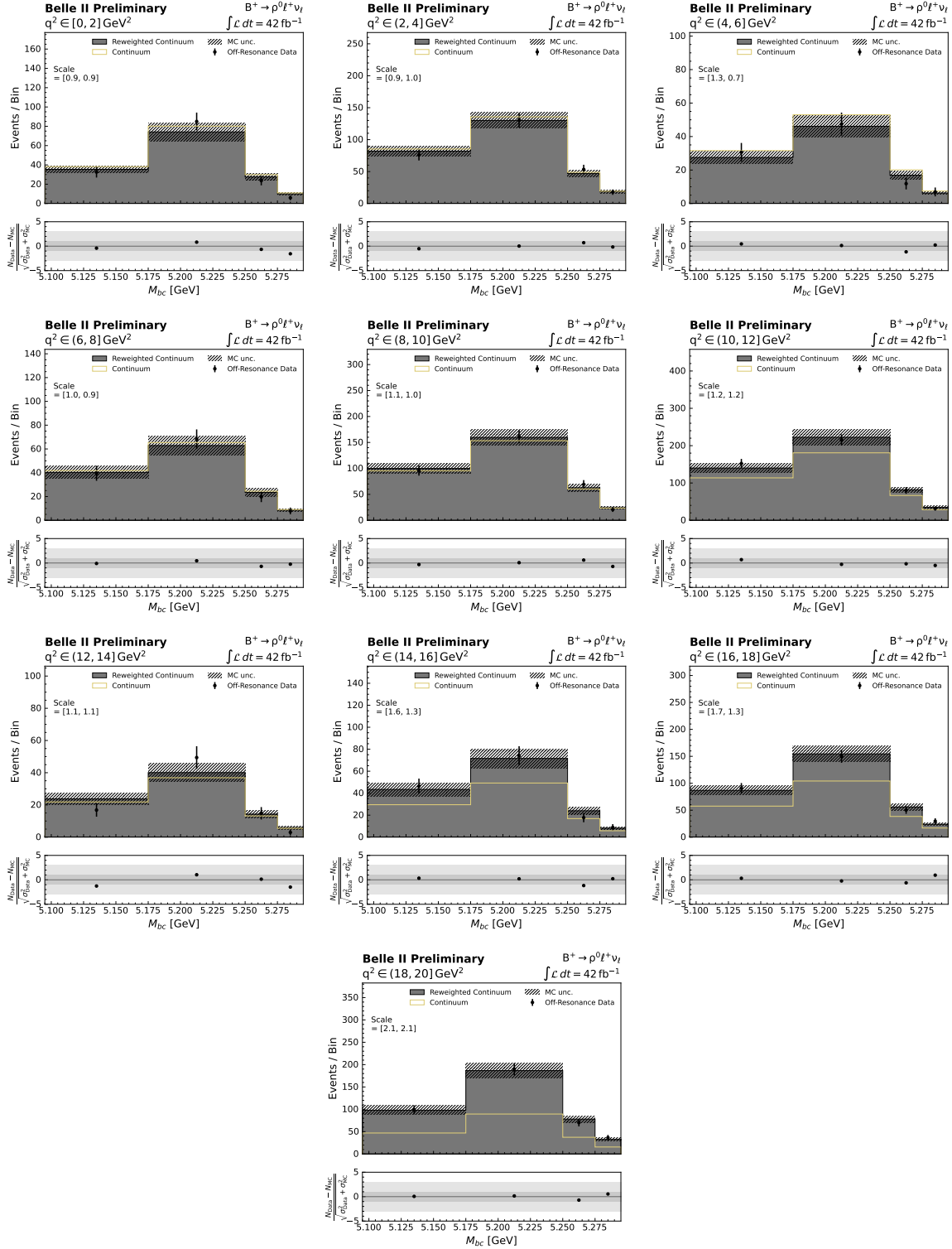


Figure A.9: Distributions of M_{bc} for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode in off-resonance and reweighted simulated data (grey) for each reconstructed q^2 bin. The scale factors for the lower and higher ΔE bins used in reweighting are indicated. The lower panels show the difference between collision and reweighted simulated data divided by the combined uncertainty. The olive line represents the simulated distribution before reweighting.

Distributions of ΔE and M_{bc}

This appendix presents the distributions of ΔE and M_{bc} for each reconstructed q^2 bin for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decay modes. The distributions following all signal selection and background suppression steps are detailed in Section B.1. One-dimensional fit projections of ΔE and M_{bc} are provided in Section B.2.

B.1 Prefit distributions

This section displays the distributions of ΔE and M_{bc} within each reconstructed q^2 bin after signal selection and background suppression.

For the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode, distributions are shown in Figure B.1, B.2, and B.3 for the q^2 bins 1–5, 6–10, and 11–13, respectively.

For the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode, distributions are provided in Figure B.4 and B.5 for the q^2 bins 1–5 and 6–10, respectively.

B.2 Postfit distributions

This section provides the distributions of ΔE and M_{bc} within each reconstructed q^2 bin for collision data, including one-dimensional fit projections from simulation.

The one-dimensional fit projections of ΔE and M_{bc} for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ candidates are shown in Figure B.6 and B.7, respectively.

The corresponding distributions for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ candidates are presented in Figure B.8 and B.9, respectively.

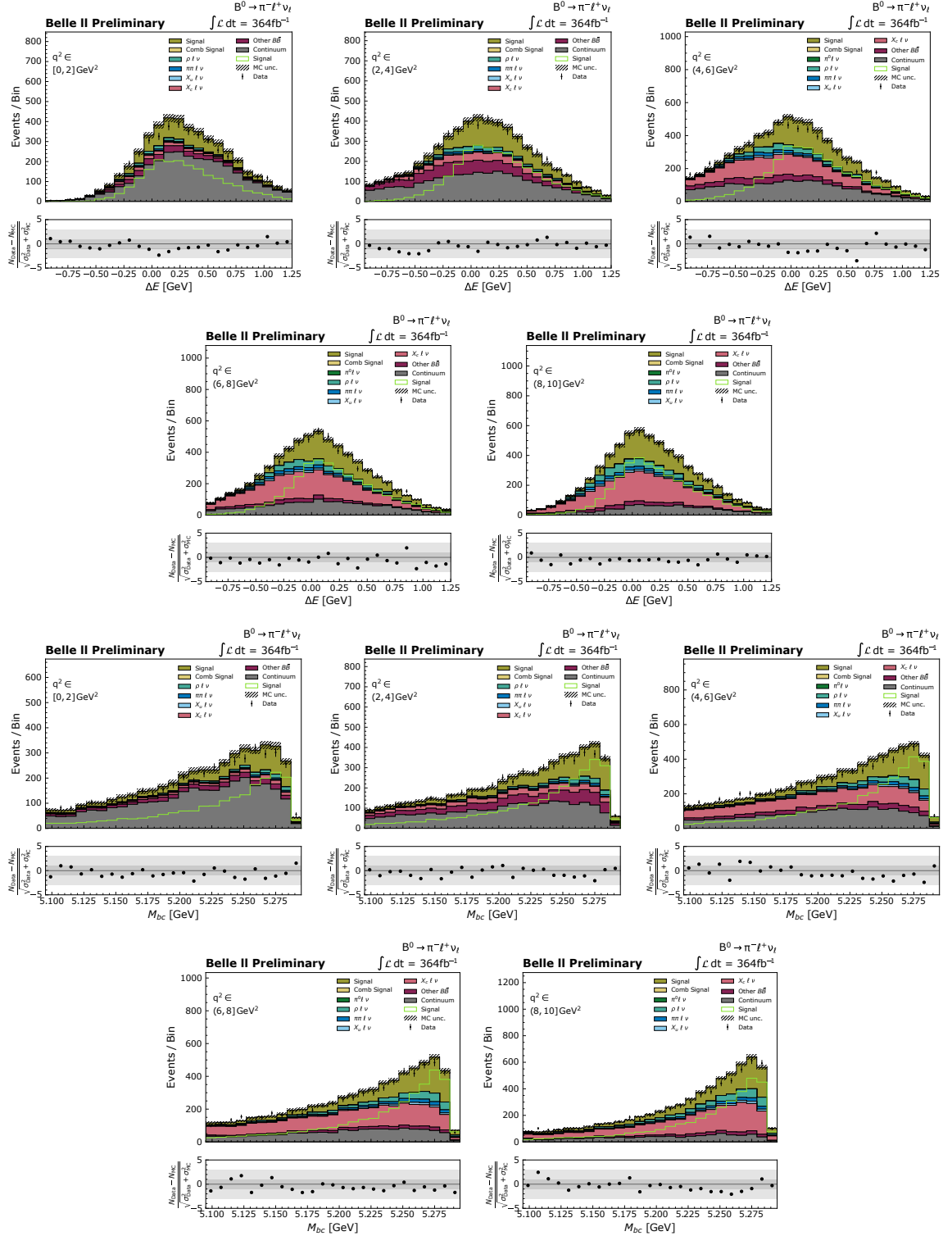


Figure B.1: Distributions of ΔE (top) and M_{bc} (bottom) in the first five reconstructed q^2 bins for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode for simulated and collision data after signal selection and background suppression. The hatched areas represent statistical and systematic uncertainties on the simulated distributions, as detailed in Chapter 8. The lower panels show the difference between collision and simulated data divided by the combined uncertainty.

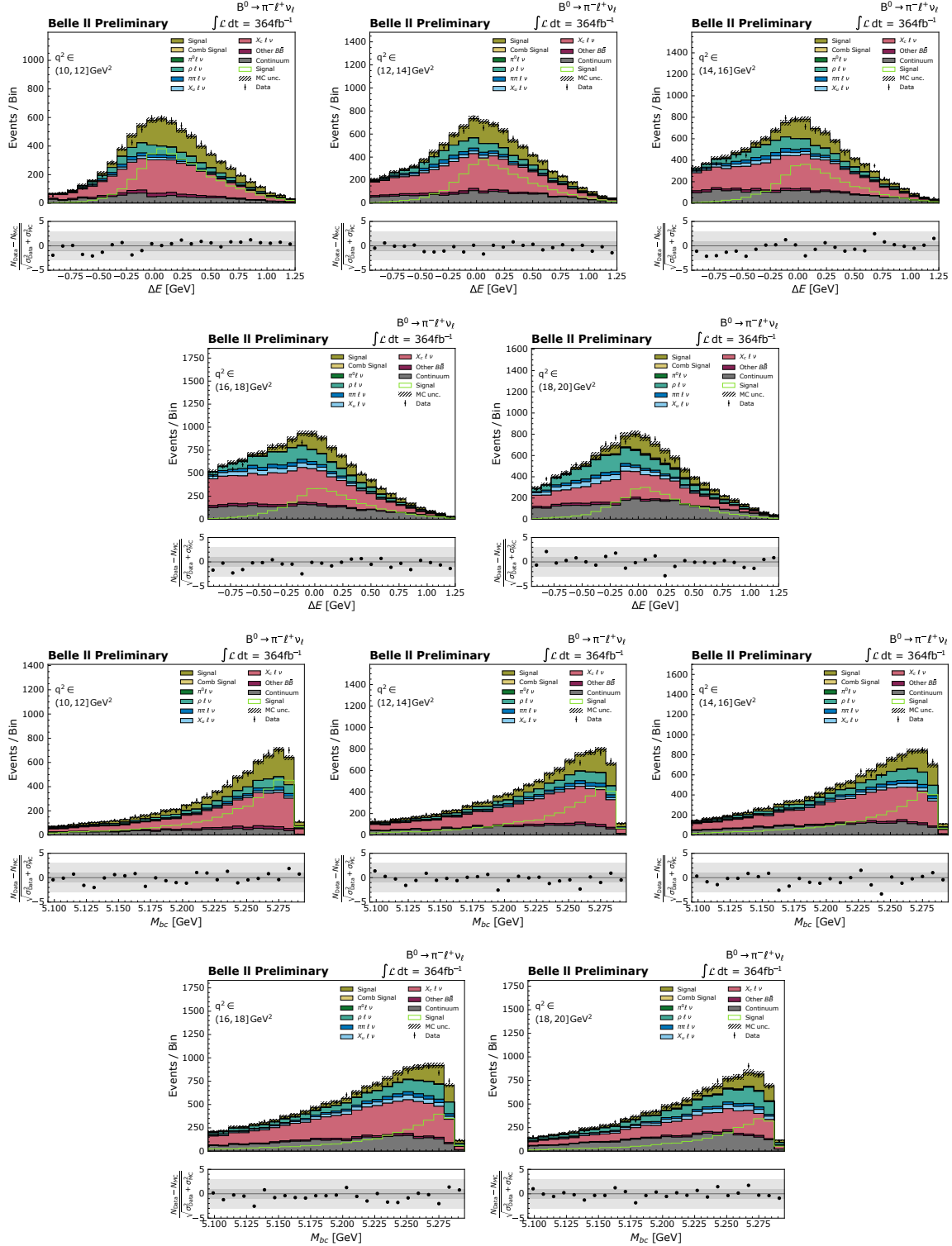


Figure B.2: Distributions of ΔE (top) and M_{bc} (bottom) in the sixth to the tenth reconstructed q^2 bins for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode for simulated and collision data after signal selection and background suppression. The hatched areas represent statistical and systematic uncertainties on the simulated distributions, as detailed in Chapter 8. The lower panels show the difference between collision and simulated data divided by the combined uncertainty.

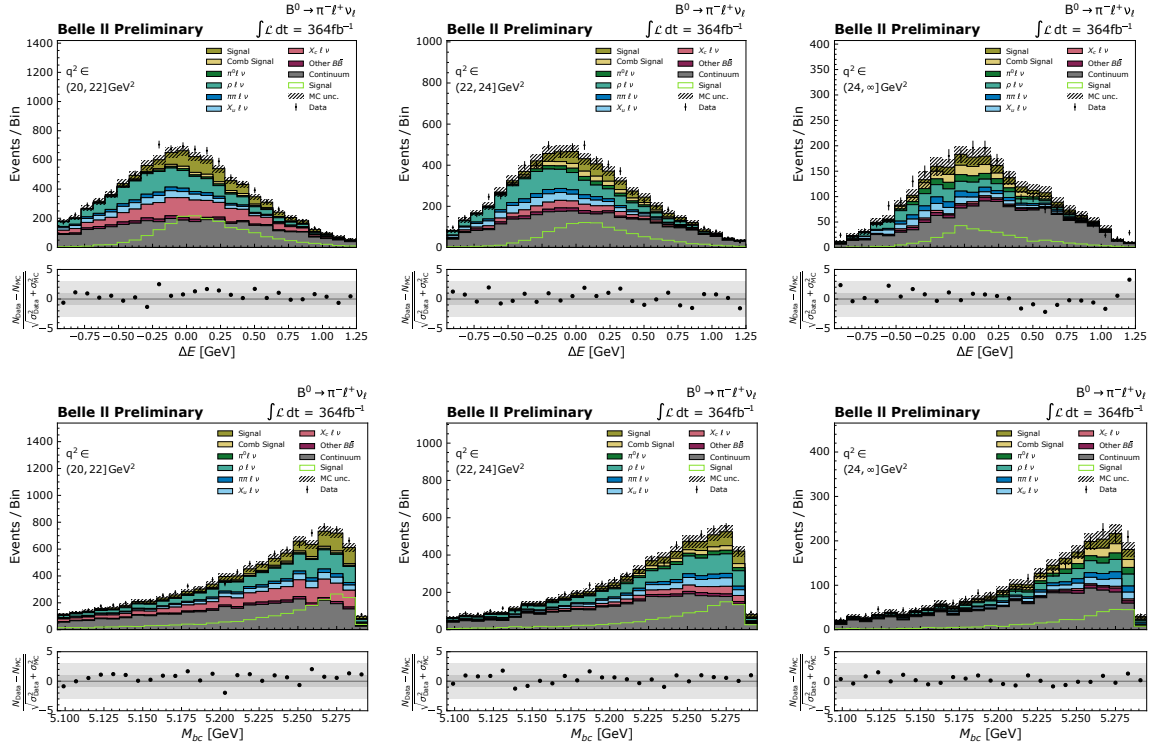


Figure B.3: Distributions of ΔE (top) and M_{bc} (bottom) in the last three reconstructed q^2 bins for the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ mode for simulated and collision data after signal selection and background suppression. The hatched areas represent statistical and systematic uncertainties on the simulated distributions, as detailed in Chapter 8. The lower panels show the difference between collision and simulated data divided by the combined uncertainty.

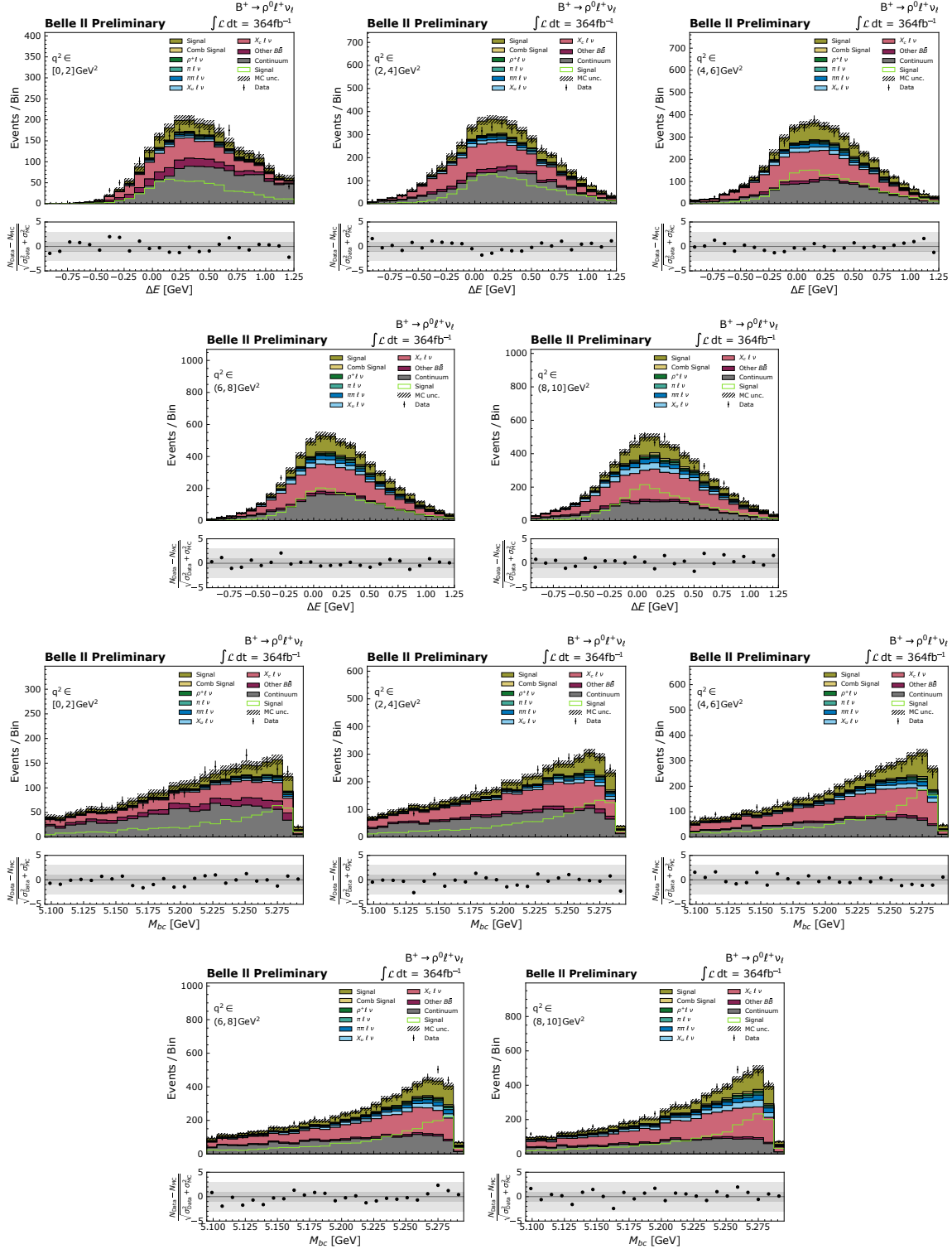


Figure B.4: Distributions of ΔE (top) and M_{bc} (bottom) in the first five reconstructed q^2 bins for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode for simulated and collision data after signal selection and background suppression. The hatched areas represent statistical and systematic uncertainties on the simulated distributions, as detailed in Chapter 8. The lower panels show the difference between collision and simulated data divided by the combined uncertainty.

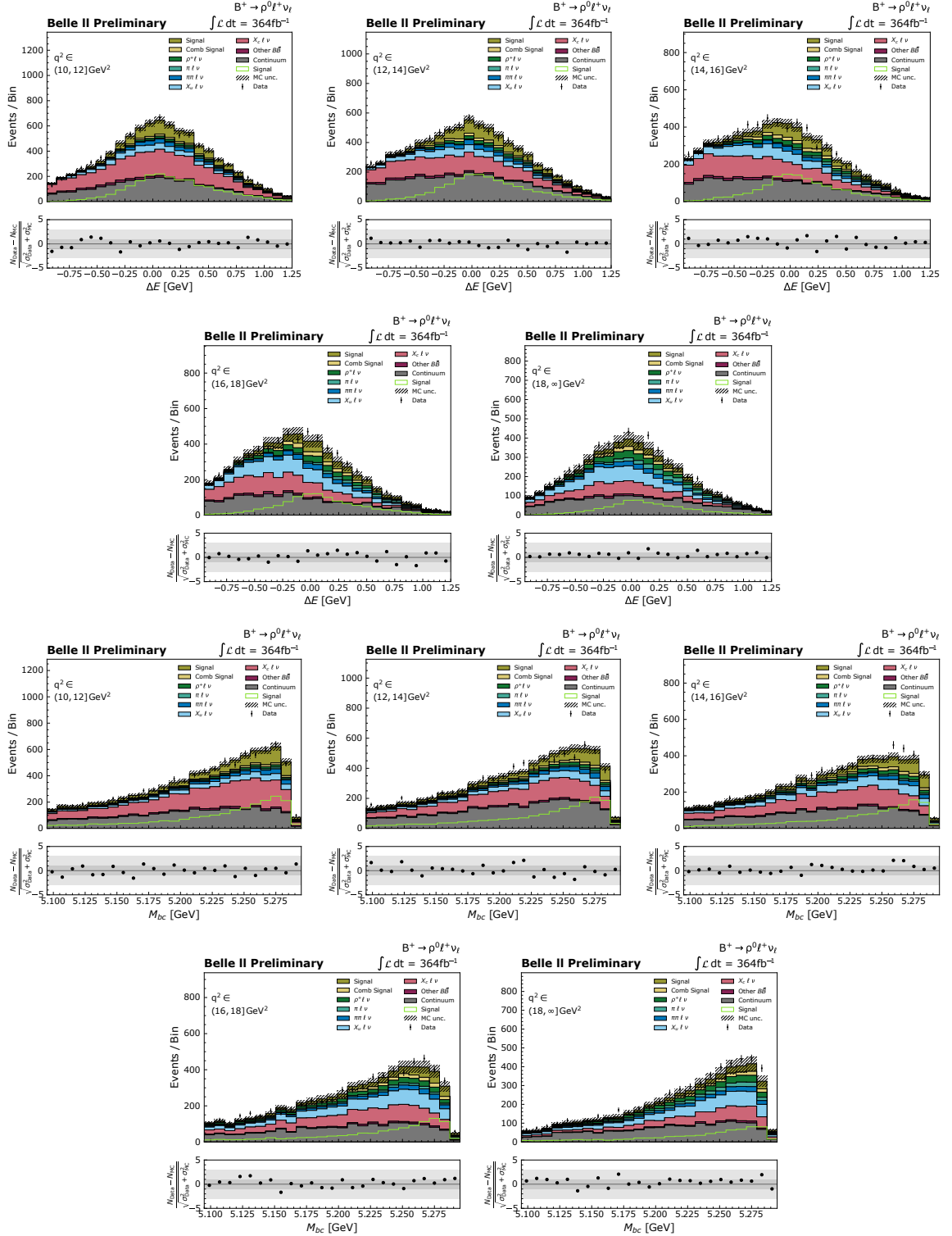


Figure B.5: Distributions of ΔE (top) and M_{bc} (bottom) in the last five reconstructed q^2 bins for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ mode for simulated and collision data after signal selection and background suppression. The hatched areas represent statistical and systematic uncertainties on the simulated distributions, as detailed in Chapter 8. The lower panels show the difference between collision and simulated data divided by the combined uncertainty.

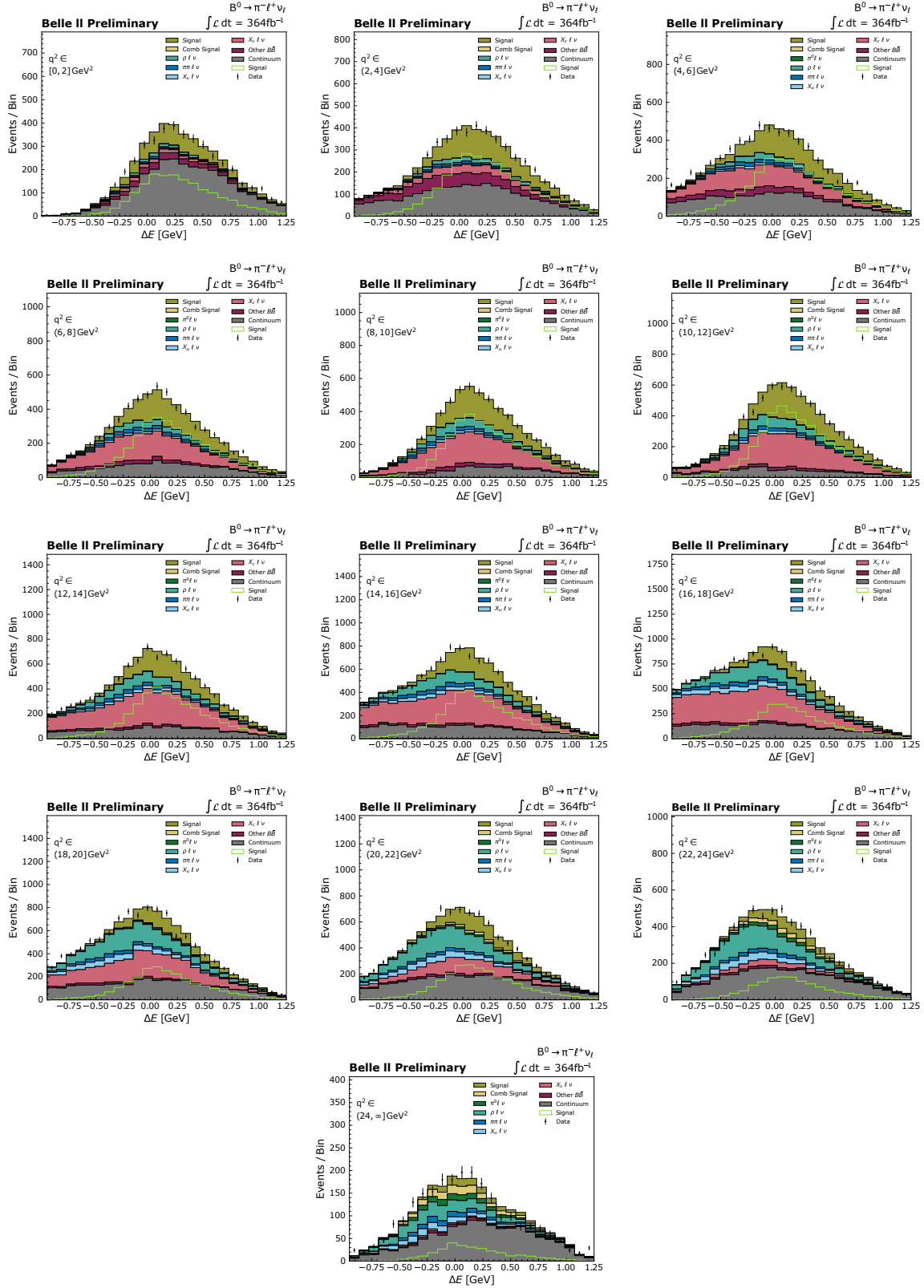


Figure B.6: Distributions of ΔE in reconstructed q^2 bins in collision data for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ candidates with one-dimensional fit projections from simulation overlaid.

Appendix B Distributions of ΔE and M_{bc}

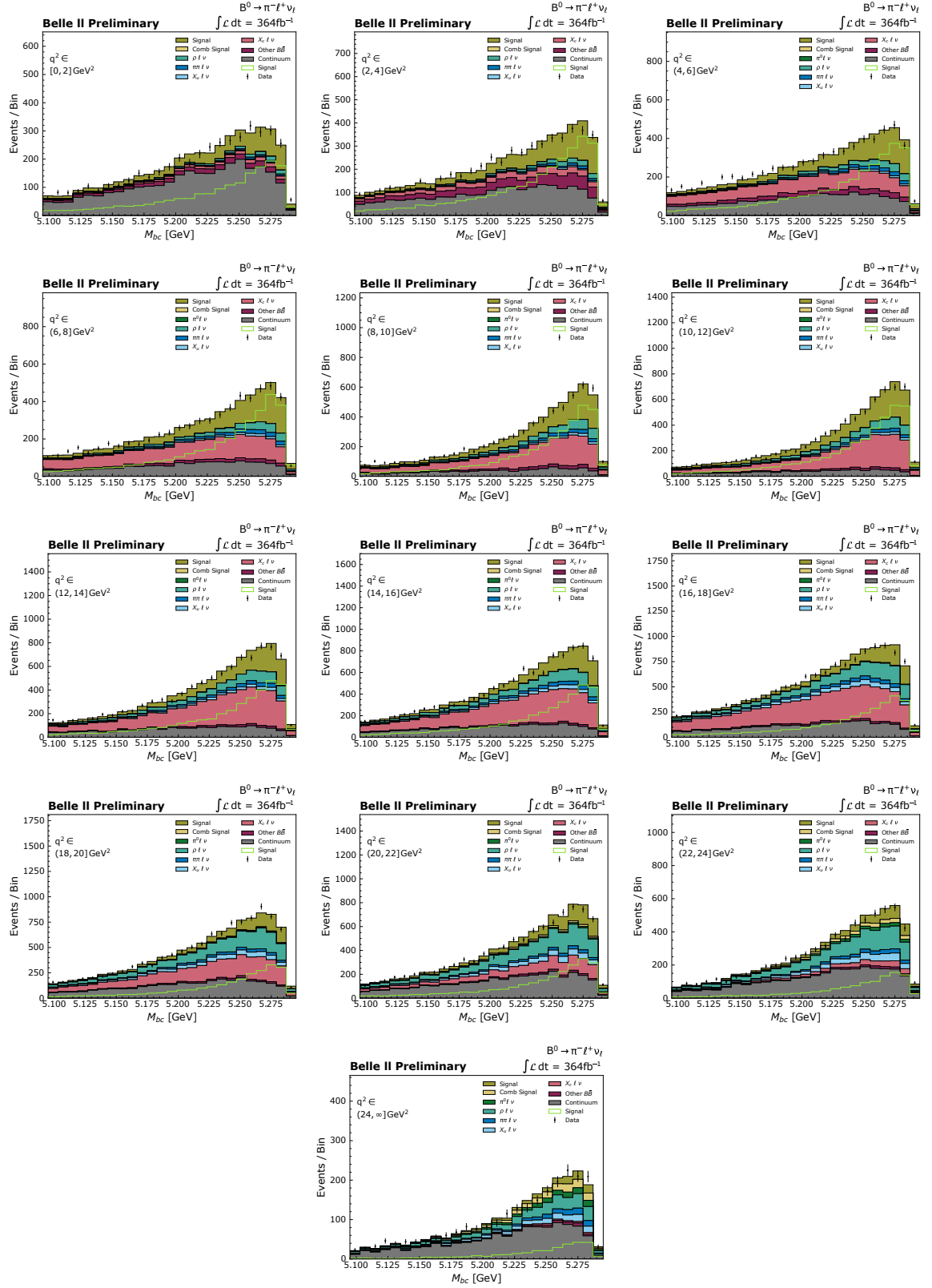


Figure B.7: Distributions of M_{bc} in reconstructed q^2 bins in collision data for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ candidates with one-dimensional fit projections from simulation overlaid.

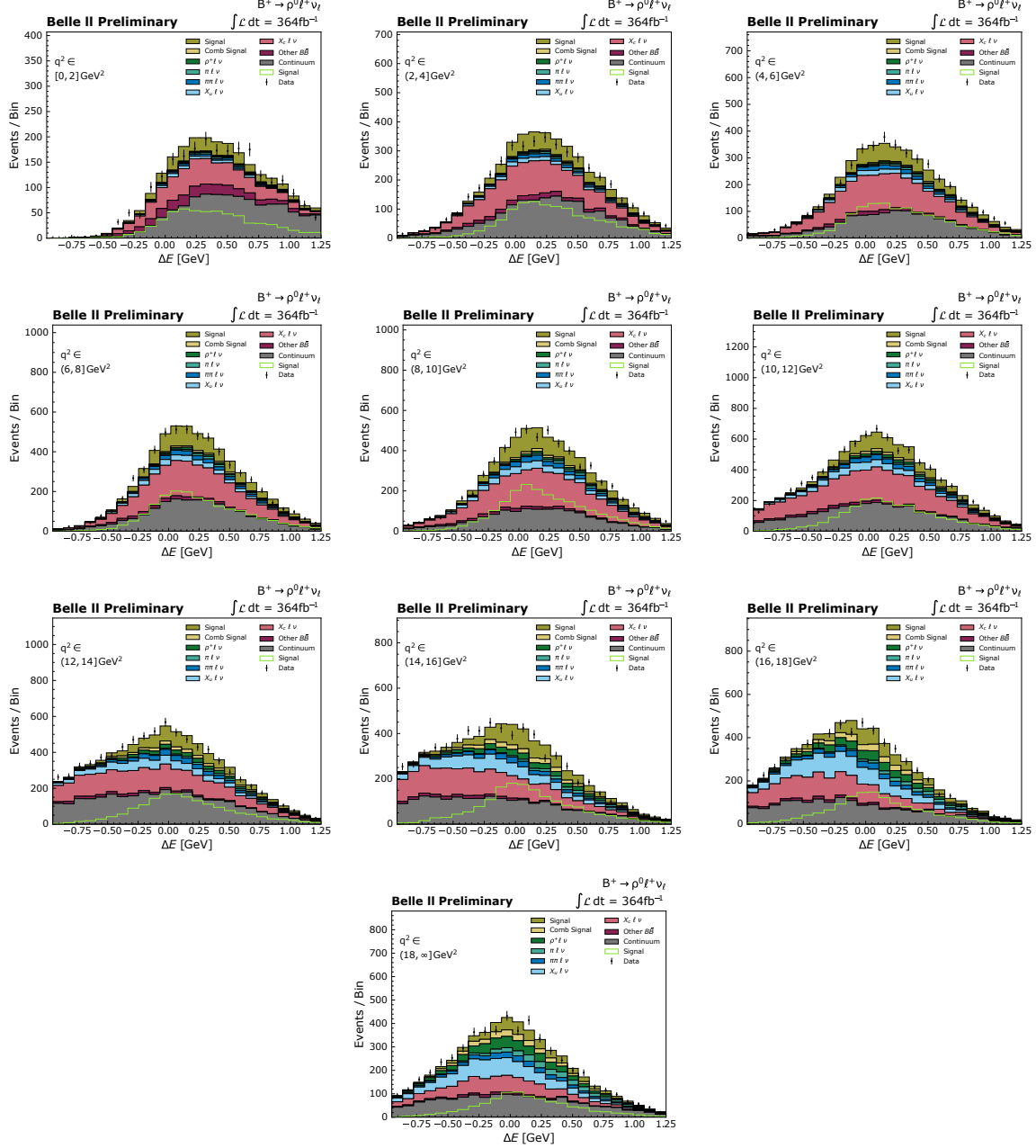


Figure B.8: Distributions of ΔE in reconstructed q^2 bins in collision data for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ candidates with one-dimensional fit projections from simulation overlaid.

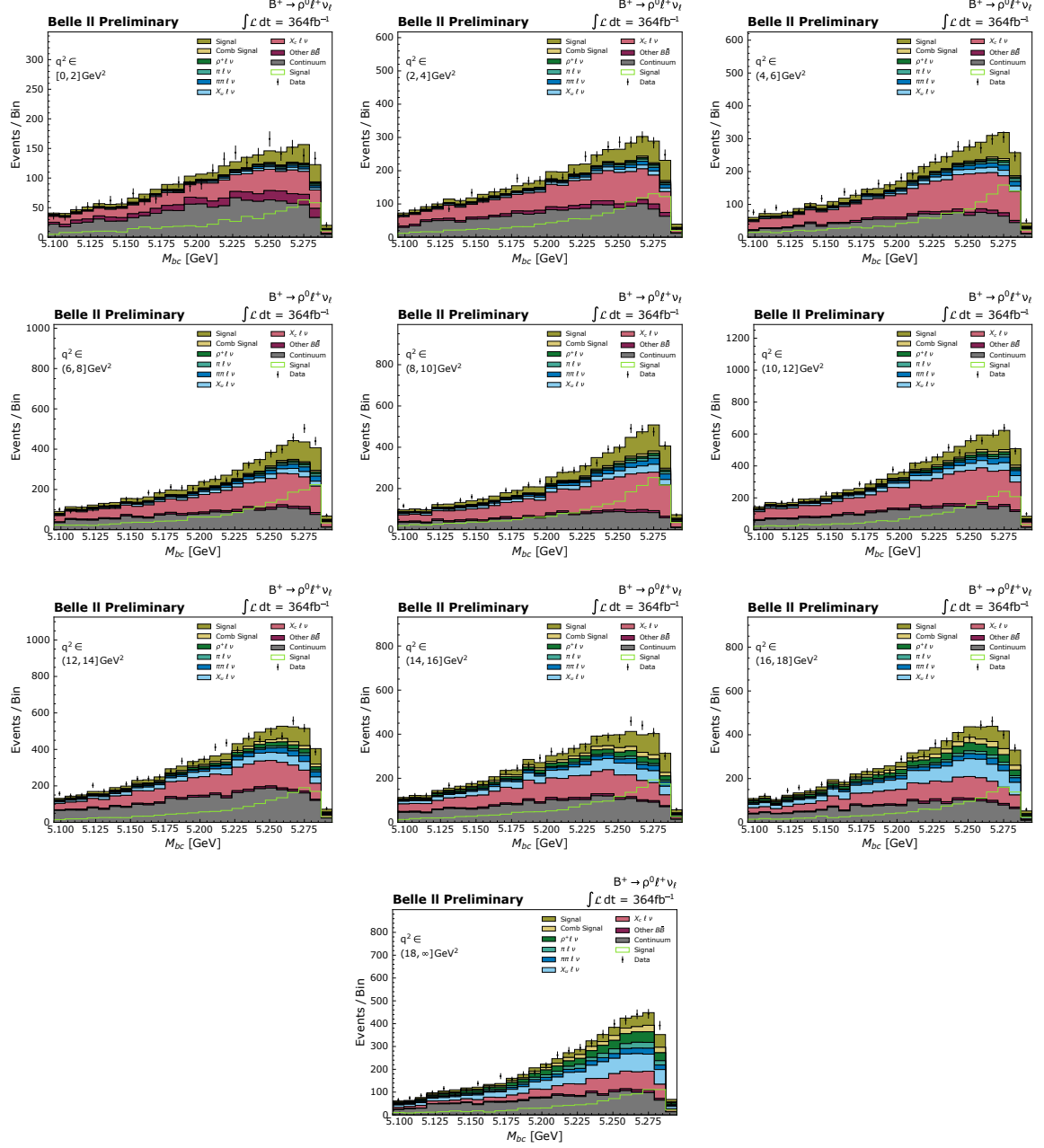


Figure B.9: Distributions of M_{bc} in reconstructed q^2 bins in collision data for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ candidates with one-dimensional fit projections from simulation overlaid.

Correlation matrices

This appendix provides the correlation matrices for the results discussed throughout this thesis.

Table C.1 displays the elements of the statistical correlation matrix for the component scale factors used in the fit, as discussed in Section 7.5.

Tables C.2 and C.3 present the elements of the full experimental covariance matrices for the partial branching fractions of the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays, detailed in Section 7.6, respectively.

Tables C.4 and C.5 include the elements of the full correlation matrices for $|V_{ub}|$ and the BCL form-factor coefficients. These matrices result from the fit to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ spectrum with LQCD and LQCD plus LCSR constraints, as discussed in Section 9.2.

Similarly, Table C.6 shows the elements of the full correlation matrix for $|V_{ub}|$ and the BSZ form-factor coefficients from the fit to the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ spectrum with LCSR constraints.

Table C.1: Statistical correlation matrix of the 31 scale factors corresponding to the fit templates.

	pimode_cont	pimode_Xu	pimode_Xc	pimode_BB	rhomode_cont	rhomode_Xu	rhomode_Xc	rhomode_BB	p1nu.q1	p1nu.q2	p1nu.q3	p1nu.q4	p1nu.q5	p1nu.q6	p1nu.q7	p1nu.q8	p1nu.q9	p1nu.q10	p1nu.q11	p1nu.q12	p1nu.q13	rholnu.q1	rholnu.q2	rholnu.q3	rholnu.q4	rholnu.q5	rholnu.q6	rholnu.q7	rholnu.q8	rholnu.q9	rholnu.q10
1	1.00																														
2	-0.06	1.00																													
3	-0.06	-0.66	1.00																												
4	-0.32	-0.11	-0.33	1.00																											
5	0.01	0.05	-0.04	0.01	1.00																										
6	0.06	0.60	-0.38	-0.06	-0.07	1.00																									
7	-0.03	-0.04	0.25	-0.12	-0.21	-0.42	1.00																								
8	-0.02	-0.14	-0.10	0.17	-0.09	0.02	-0.75	1.00																							
9	-0.33	-0.07	0.14	-0.09	0.01	-0.07	-0.14	0.24	1.00																						
10	0.05	-0.02	0.24	-0.55	0.01	-0.04	-0.01	0.04	0.01	1.00																					
11	-0.07	-0.04	-0.01	-0.07	0.01	-0.08	-0.01	0.06	0.12	-0.14	1.00																				
12	-0.03	-0.01	-0.11	0.01	-0.01	-0.06	0.03	0.02	0.02	0.06	-0.22	1.00																			
13	0.04	0.08	-0.20	-0.02	-0.01	0.01	0.05	-0.02	-0.02	0.01	0.10	-0.18	1.00																		
14	0.06	0.19	-0.32	0.04	-0.03	0.07	0.08	-0.07	-0.06	-0.05	0.02	0.12	-0.08	1.00																	
15	-0.01	0.21	-0.39	0.14	-0.01	0.08	0.07	-0.07	-0.06	-0.10	0.03	0.08	0.16	0.04	1.00																
16	-0.02	0.22	-0.41	0.16	0.04	0.15	-0.04	-0.01	-0.05	-0.10	0.02	0.07	0.11	0.19	0.10	1.00															
17	-0.03	0.23	-0.44	0.19	0.03	0.22	-0.11	0.03	-0.05	-0.12	0.02	0.07	0.11	0.16	0.22	0.15	1.00														
18	-0.08	0.26	-0.34	0.12	0.01	0.35	-0.15	0.03	-0.02	-0.08	0.01	0.04	0.07	0.12	0.15	0.19	0.15	1.00													
19	-0.15	0.12	-0.13	0.07	0.01	0.31	-0.17	0.06	0.05	-0.02	0.01	0.02	0.02	0.04	0.06	0.09	0.13	0.14	1.00												
20	-0.17	-0.09	0.12	0.02	-0.01	0.14	-0.10	0.05	0.08	0.02	0.01	-0.01	-0.02	-0.04	-0.03	-0.02	-0.01	0.04	0.13	1.00											
21	-0.13	-0.26	0.24	0.01	-0.02	-0.15	0.01	0.04	0.08	0.04	0.02	0.01	-0.04	-0.08	-0.08	-0.09	-0.09	-0.04	-0.04	-0.24	1.00										
22	0.04	0.08	0.07	-0.16	-0.04	0.03	0.49	-0.71	-0.36	-0.06	-0.04	-0.04	0.01	0.04	0.04	0.01	-0.02	-0.02	-0.03	-0.03	-0.02	1.00									
23	0.02	-0.07	-0.02	-0.01	-0.07	0.04	-0.16	0.09	0.13	-0.08	-0.12	0.04	-0.04	-0.02	-0.03	-0.02	0.01	0.01	0.01	0.01	0.02	-0.35	1.00								
24	0.01	-0.09	-0.04	0.04	0.05	0.02	-0.27	0.18	-0.01	0.06	-0.07	-0.14	0.04	-0.07	-0.03	-0.01	0.01	0.01	0.02	0.02	0.02	0.02	-0.42	1.00							
25	0.01	-0.14	-0.01	0.04	-0.03	-0.03	-0.24	0.17	0.05	-0.03	0.08	-0.04	-0.20	0.01	-0.05	-0.02	-0.01	-0.02	-0.01	0.01	0.03	-0.13	0.27	-0.39	1.00						
26	0.01	-0.19	-0.02	0.07	0.10	-0.08	-0.36	0.28	0.06	0.01	-0.01	0.06	-0.04	-0.25	-0.05	-0.03	0.01	-0.02	-0.01	0.01	0.05	-0.16	0.03	0.31	-0.27	1.00					
27	0.01	-0.24	-0.02	0.09	0.03	-0.08	-0.43	0.36	0.08	-0.01	0.03	-0.01	0.01	-0.07	-0.25	-0.03	-0.01	-0.02	0.01	0.02	0.06	-0.21	0.14	0.13	0.30	-0.08	1.00				
28	-0.01	-0.33	0.11	0.07	-0.15	-0.30	-0.03	0.13	0.04	0.01	0.03	0.03	-0.01	0.01	-0.02	-0.29	-0.08	-0.08	-0.07	-0.01	0.08	-0.06	0.06	0.08	0.08	0.20	-0.15	1.00			
29	-0.01	-0.38	0.17	0.05	-0.03	-0.35	-0.02	0.11	0.05	0.01	0.03	0.02	-0.01	-0.06	-0.03	-0.03	-0.34	-0.18	-0.07	-0.01	0.10	-0.06	0.05	0.06	0.09	0.11	0.19	-0.14	1.00		
30	-0.04	-0.44	0.27	0.04	0.02	-0.51	0.13	0.04	0.06	0.03	0.04	0.03	-0.02	-0.07	-0.08	-0.08	-0.06	-0.45	-0.23	0.05	0.16	-0.03	0.02	0.03	0.07	0.10	0.11	0.23	-0.08	1.00	
31	0.01	-0.48	0.32	0.01	-0.06	-0.58	0.21	-0.01	0.03	0.03	0.03	0.02	-0.03	-0.08	-0.10	-0.15	-0.17	-0.16	-0.42	-0.42	0.06	0.01	0.02	0.02	0.06	0.09	0.11	0.22	0.27	0.02	1.00

Table C.2: Full experimental correlation matrix of the partial branching fractions $\Delta\mathcal{B}_i$ for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays.

q^2 bin	$q1$	$q2$	$q3$	$q4$	$q5$	$q6$	$q7$	$q8$	$q9$	$q10$	$q11$	$q12$	$q13$
$q1$	1.000												
$q2$	0.021	1.000											
$q3$	0.105	-0.193	1.000										
$q4$	-0.018	0.019	-0.139	1.000									
$q5$	-0.031	-0.052	0.202	-0.053	1.000								
$q6$	0.065	-0.058	0.034	0.097	0.004	1.000							
$q7$	-0.097	-0.160	0.069	0.226	0.223	0.090	1.000						
$q8$	-0.067	-0.097	0.026	0.026	0.194	0.255	0.213	1.000					
$q9$	0.088	0.035	-0.019	-0.027	0.053	0.170	0.108	0.110	1.000				
$q10$	0.007	-0.007	0.001	-0.053	0.067	0.100	0.050	0.058	0.196	1.000			
$q11$	0.075	0.001	0.059	-0.005	0.021	0.056	0.028	-0.035	0.148	0.236	1.000		
$q12$	0.050	0.080	0.014	0.004	-0.035	-0.044	-0.038	-0.101	0.074	0.187	0.297	1.000	
$q13$	0.030	-0.053	0.115	0.024	0.041	-0.048	-0.011	-0.078	-0.092	-0.129	-0.212	-0.355	1.000

Table C.3: Full experimental correlation matrix of the partial branching fractions $\Delta\mathcal{B}_i$ for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays.

q^2 bin	$q1$	$q2$	$q3$	$q4$	$q5$	$q6$	$q7$	$q8$	$q9$	$q10$
$q1$	1.000									
$q2$	-0.340	1.000								
$q3$	0.146	-0.322	1.000							
$q4$	0.023	0.241	-0.241	1.000						
$q5$	-0.052	0.131	0.275	-0.060	1.000					
$q6$	0.017	0.139	0.183	0.464	0.148	1.000				
$q7$	-0.021	0.197	0.068	0.184	0.428	0.030	1.000			
$q8$	0.149	0.018	0.054	0.216	0.205	0.311	-0.063	1.000		
$q9$	0.095	0.101	0.050	0.115	0.136	0.156	0.235	-0.005	1.000	
$q10$	0.004	0.187	-0.083	0.153	0.151	0.133	0.188	0.341	0.241	1.000

Table C.4: Full correlation matrix of $|V_{ub}|$ and the BCL form-factor coefficients from the fit to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ spectrum with LQCD constraints.

	$ V_{ub} $	b_0^+	b_1^+	b_2^+	b_0^0	b_1^0
$ V_{ub} $	1.000					
b_0^+	-0.806	1.000				
b_1^+	-0.053	-0.273	1.000			
b_2^+	0.062	-0.319	-0.338	1.000		
b_0^0	-0.315	0.409	-0.073	-0.204	1.000	
b_1^0	-0.142	-0.048	0.150	0.258	-0.775	1.000

Table C.5: Full correlation matrix of $|V_{ub}|$ and the BCL form-factor coefficients from the fit to the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ spectrum with LQCD and LCSR constraints.

	$ V_{ub} $	b_0^+	b_1^+	b_2^+	b_0^0	b_1^0
$ V_{ub} $	1.000					
b_0^+	-0.791	1.000				
b_1^+	0.007	-0.339	1.000			
b_2^+	0.243	-0.375	-0.448	1.000		
b_0^0	-0.376	0.430	-0.065	-0.190	1.000	
b_1^0	0.003	-0.164	0.127	0.244	-0.830	1.000

 Table C.6: Full correlation matrix of $|V_{ub}|$ and the BSZ form-factor coefficients from the fit to the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ spectrum with LCSR constraints.

	$ V_{ub} $	$b_0^{A_1}$	$b_1^{A_1}$	$b_0^{A_2}$	$b_1^{A_2}$	b_0^V	b_1^V
$ V_{ub} $	1.000						
$b_0^{A_1}$	-0.464	1.000					
$b_1^{A_1}$	0.035	0.542	1.000				
$b_0^{A_2}$	-0.735	0.241	-0.117	1.000			
$b_1^{A_2}$	-0.126	-0.007	0.023	0.472	1.000		
b_0^V	-0.473	0.894	0.493	0.255	-0.056	1.000	
b_1^V	0.064	0.538	0.946	-0.144	0.127	0.558	1.000

Bibliography

- [1] H. Harari, *A new quark model for hadrons*, *Physics Letters B* **57** (1975) 265.
- [2] S. Abachi et al., *Observation of the Top Quark*, *Phys. Rev. Lett.* **74** (1995) 2632.
- [3] F. Abe et al., *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, *Phys. Rev. Lett.* **74** (1995) 2626.
- [4] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [5] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Physics Letters B* **716** (2012) 1.
- [6] S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Physics Letters B* **716** (2012) 30.
- [7] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531.
- [8] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, *Prog. Theor. Phys.* **49** (1973) 652.
- [9] Y. Amhis et al., *Averages of b -hadron, c -hadron, and τ -lepton properties as of 2021*, *Phys. Rev. D* **107** (2023) 052008, updated results and plots available at <https://hflav.web.cern.ch/>.
- [10] F. U. Bernlochner, M. T. Prim and D. J. Robinson, *$B \rightarrow \rho \ell \nu$ and $\omega \ell \nu$ in and beyond the Standard Model: Improved predictions and $|V_{ub}|$* , *Phys. Rev. D* **104** (2021) 034032.
- [11] I. Adachi et al., *Measurement of the integrated luminosity of data samples collected during 2019-2022 by the Belle II experiment*, (2024), arXiv: [2407.00965](https://arxiv.org/abs/2407.00965) [hep-ex].
- [12] C. Bourrely, L. Lellouch and I. Caprini, *Model-independent description of $B \rightarrow \pi \ell \nu$ decays and a determination of $|V_{ub}|$* , *Phys. Rev. D* **79** (2009) 013008.
- [13] A. Bharucha, D. M. Straub and R. Zwicky, *$B \rightarrow V \ell \ell$ in the Standard Model from light-cone sum rules*, *J. High Energ. Phys.* **8** (2016) 98.
- [14] Y. Aoki et al., *FLAG Review 2021*, *Eur. Phys. J. C* **82** (2023) 869.
- [15] J. A. Bailey et al., *$|V_{ub}|$ from $B \rightarrow \pi \ell \nu$ decays and $(2+1)$ -flavor lattice QCD*, *Phys. Rev. D* **92** (2015) 014024.
- [16] J. M. Flynn et al., *$B \rightarrow \pi \ell \nu$ and $B_s \rightarrow K \ell \nu$ form factors and $|V_{ub}|$ from $2+1$ -flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks*, *Phys. Rev. D* **91** (2015) 074510.
- [17] B. Colquhoun et al., *Form factors of $B \rightarrow \pi \ell \nu$ and a determination of $|V_{ub}|$ with Möbius domain-wall fermions*, *Phys. Rev. D* **106** (2022) 054502.

- [18] D. Lejcek, B. Melić and D. van Dyk, *The $\bar{B} \rightarrow \pi$ form factors from QCD and their impact on $|V_{ub}|$* , *J. High Energ. Phys.* **07** (2021) 36.
- [19] A. Roodman, *Blind Analysis in Particle Physics*, 2003, arXiv: [physics/0312102](#) [[physics.data-an](#)].
- [20] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, *Nuclear Physics B* **44** (1972) 189.
- [21] Wikipedia, *Standard Model*, <http://en.wikipedia.org/w/index.php?title=Standard%20Model&oldid=1147861546>, [Online; accessed 02-May-2023], 2023.
- [22] G. Lüders, *Proof of the TCP theorem*, *Annals of Physics* **2** (1957) 1.
- [23] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346.
- [24] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343.
- [25] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, *Phys. Rev.* **80** (1950) 440.
- [26] S. L. Glashow, *The renormalizability of vector meson interactions*, *Nuclear Physics* **10** (1959) 107.
- [27] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, *Il Nuovo Cimento* (1955-1965) **11** (1959) 568.
- [28] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.
- [29] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [30] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [31] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.
- [32] F. Halzen and A. Martin, *Quarks and Leptons: An introductory course in modern particle physics*, New York, USA: John Wiley and Sons, 1984.
- [33] D. Griffiths, *Introduction to Elementary Particles*, New York, USA: John Wiley and Sons, 1987.
- [34] L.-L. Chau and W. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **53** (1984) 1802.
- [35] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **51** (1983) 1945.
- [36] J. Dingfelder and T. Mannel, *Leptonic and semileptonic decays of B mesons*, *Rev. Mod. Phys.* **88** (2016) 035008.
- [37] R. L. Workman et al., *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [38] A. V. Manohar and M. B. Wise, *Heavy quark physics*, vol. 10, 2000 1.

-
- [39] T. Mannel, *Effective Field Theories in Flavor Physics*, Springer Tracts Mod. Phys. **203** (2004) 1.
 - [40] J. Chay, H. Georgi and B. Grinstein, *Lepton energy distributions in heavy meson decays from QCD*, Physics Letters B **247** (1990) 399.
 - [41] I. I. Bigi, M. Shifman, N. G. Uraltsev and A. Vainshtein, *QCD predictions for lepton spectra in inclusive heavy flavor decays*, Phys. Rev. Lett. **71** (1993) 496.
 - [42] U. Fano, *Effects of Configuration Interaction on Intensities and Phase Shifts*, Phys. Rev. **124** (1961) 1866.
 - [43] T. Mannel, *Operator product expansion for inclusive semileptonic decays in heavy quark effective field theory*, Nuclear Physics B **413** (1994) 396.
 - [44] M. Beneke and T. Feldmann, *Factorization of heavy-to-light form factors in soft-collinear effective theory*, Nuclear Physics B **685** (2004) 249.
 - [45] A. L. Kagan and M. Neubert, *QCD anatomy of $B \rightarrow X_s \gamma$ decays*, Eur. Phys. J. C **7** (1999) 5.
 - [46] B. O. Lange et al., *Theory of charmless inclusive B decays and the extraction of $|V_{ub}|$* , Phys. Rev. D **72** (2005) 073006.
 - [47] F. D. Fazio and M. Neubert, *$B \rightarrow X_u \ell \nu$ decay distributions to order α_s* , J. High Energ. Phys. **1999** (1999) 017.
 - [48] D. Becirevic and A. B. Kaidalov, *Comment on the heavy to light form factors*, Physics Letters B **478** (2000) 417.
 - [49] C. G. Boyd, B. Grinstein and R. F. Lebed, *Constraints on Form Factors for Exclusive Semileptonic Heavy to Light Meson Decays*, Phys. Rev. Lett. **74** (1995) 4603.
 - [50] J. A. Bailey et al., *$B \rightarrow \pi l \nu$ semileptonic form factor from three-flavor lattice QCD: A model-independent determination of $|V_{ub}|$* , Phys. Rev. D **79** (2009) 054507.
 - [51] J. Bartelt et al., *Measurement of charmless semileptonic decays of B mesons*, Phys. Rev. Lett. **71** (1993) 4111.
 - [52] L. Cao, F. Bernlochner, K. Tackmann et al., *First Simultaneous Determination of Inclusive and Exclusive $|V_{ub}|$* , Phys. Rev. Lett. **131** (2023) 211801.
 - [53] P. Gambino, P. Giordano, G. Ossola and N. Uraltsev, *Inclusive semileptonic B decays and the determination of $|V_{ub}|$* , J. High Energ. Phys. **2007** (2007) 058.
 - [54] L. V. Silva, *2023 update of the extraction of the CKM matrix elements*, 2024, arXiv: 2405.08046.
 - [55] J. Charles et al., *CKM Fitter*, Eur. Phys. J C **41** (2005) 1.
 - [56] K. Akai et al., *SuperKEKB collider*, Nucl. Instrum. Meth. A **907** (2018) 188.
 - [57] Y. Ohnishi et al., *Accelerator design at SuperKEKB*, PTEP **2013** (2013).
 - [58] T. Abe et al., *Belle II Technical Design Report*, 2010, arXiv: 1011.0352 [physics.ins-det].
 - [59] *KEKB B-Factory Design Report*, KEK Report **95-7** (1995).
 - [60] CERN Courier, *SuperKEKB raises the bar*, <https://cerncourier.com/a/superkekb-raises-the-bar/>, [Online; accessed 21-Nov-2023], 2021.

- [61] P. Raimondi, D. N. Shatilov and M. Zobov, *Beam-Beam Issues for Colliding Schemes with Large Piwinski Angle and Crabbed Waist*, 2007, arXiv: [physics/0702033](#) [[physics.acc-ph](#)].
- [62] M. Baszczyk et al., *SuperB Technical Design Report*, 2013, arXiv: [1306.5655](#) [[physics.ins-det](#)].
- [63] Belle II collaboration, *Belle II Luminosity*, <https://confluence.desy.de/display/BI/Belle+II+Luminosity>, [Online; accessed 21-Nov-2023], 2023.
- [64] E. Kou et al., *The Belle II Physics*, [PTEP 2019 \(2019\) 123C01](#).
- [65] A. Natochii et al., *Beam background expectations for Belle II at SuperKEKB*, (2023), eprint: [2203.05731](#) (hep-ex).
- [66] A. Piwinski, *The Touschek Effect in Strong Focusing Storage Rings*, (1999), arXiv: [physics/9903034](#) [[physics.acc-ph](#)].
- [67] A. Abashian et al., *The Belle detector*, [Nucl. Instrum. Meth. A 479 \(2002\) 117](#).
- [68] Belle II Collaboration, *Belle II Detector 3D model*, <https://www.belle2.org/archives/>, [Online; accessed 27-June-2024], 2024.
- [69] H. Kolanoski and N. Wermes, *Teilchendetektoren: Grundlagen und Anwendungen*, 1st ed., Springer, 2016.
- [70] J. Kemmer and G. Lutz, *New detector concepts*, [Nucl. Instrum. Meth. A 253 \(1987\) 365](#).
- [71] C. Marinas, *The Belle II pixel detector: High precision with low material*, [Nucl. Instrum. Meth. A 731 \(2013\) 31](#).
- [72] G. Giakoustidis et al., *Status of the BELLE II Pixel Detector*, [PoS Pixel2022 \(2023\) 005](#).
- [73] B. Wang et al., *Operational experience of the Belle II pixel detector*, [Nucl. Instrum. Meth. A 1032 \(2022\) 166631](#).
- [74] K. Adamczyk et al., *The design, construction, operation and performance of the Belle II silicon vertex detector*, [JINST 17 \(2022\) P11042](#).
- [75] G. Dujany et al., *The silicon vertex detector of the Belle II experiment*, [JINST 17 \(2022\) C08014](#).
- [76] R. Leboucher et al., *Measurement of the cluster position resolution of the Belle II Silicon Vertex Detector*, [Nucl. Instrum. Meth. A 1033 \(2022\) 166746](#).
- [77] C. Irmmler et al., *The silicon vertex detector of the Belle II experiment*, [Nucl. Instrum. Meth. A 1045 \(2023\) 167578](#).
- [78] Y. Uematsu et al., *The Silicon Vertex Detector of the Belle II experiment*, [Nucl. Instrum. Meth. A 1033 \(2022\) 166688](#).
- [79] K. Adamczyk et al., *The design, construction, operation and performance of the Belle II silicon vertex detector*, [JINST 17 \(2022\) P11042](#).
- [80] L. Zani et al., *The Silicon Vertex Detector of the Belle II experiment*, [Nucl. Instrum. Meth. A 1038 \(2022\) 166952](#).
- [81] T. Dong et al., *Calibration and alignment of the Belle II central drift chamber*, [Nucl. Instrum. Meth. A 930 \(2019\) 132](#).
- [82] K. Kojima, *The operation and performance of the TOP detector at the Belle II experiment*, [PoS EPS-HEP2021 \(2022\) 803](#).

-
- [83] S. Sandilya et al., *Charged particle identification performance of the TOP counters in Belle II*, *J. Phys.: Conf. Ser.* **2374** (2022) 012107.
- [84] S. Korpar et al., *Proximity focusing RICH with TOF capabilities*, *Nucl. Instrum. Meth. A* **572** (2007) 432.
- [85] L. Santelj et al., *Recent developments in data reconstruction for aerogel RICH at Belle II*, *Nucl. Instrum. Meth. A* **1055** (2023) 168502.
- [86] Y. Iwasaki et al., *Level 1 trigger system for the Belle II experiment*, *2010 17th IEEE-NPSS Real Time Conference* (2010) 1.
- [87] R. Itoh et al., *The Performance of Belle II High Level Trigger in the First Physics Run*, *EPJ Web Conf.* **245** (2020) 01040.
- [88] S. Lee et al., *J. Phys.: Conf. Ser.* **396** (2012) 012029.
- [89] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth and N. Braun, *The Belle II Core Software*, *Comput. Softw. Big Sci.* **3** (2019) 1.
- [90] *Belle II Analysis Software Framework (basf2)*, <https://doi.org/10.5281/zenodo.5574115>.
- [91] V. Bertacchi et al., *Track finding at Belle II*, *Comput. Phys. Comm.* **259** (2021) 107610.
- [92] S. Spataro, *Track Fitting for the Belle II Experiment*, *EPJ Web Conf.* **214** (2019) 02039.
- [93] T. Bilka et al., *Implementation of GENFIT2 as an experiment independent track-fitting framework*, (2019), arXiv: [1902.04405](https://arxiv.org/abs/1902.04405) [physics.data-an].
- [94] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **462** (2001) 152.
- [95] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159.
- [96] S. Jadach, B. F. L. Ward and Z. Was, *The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions*, *Comput. Phys. Commun.* **130** (2000) 260.
- [97] F. A. Berends and R. van Gulik, *GaGaRes: A Monte Carlo generator for resonance production in two-photon physics*, *Computer Physics Communications* **144** (2002) 82.
- [98] S. Jadach, J. H. Kuhn and Z. Was, *TAUOLA: A library of Monte Carlo programs to simulate decays of polarized tau leptons*, *Comput. Phys. Commun.* **64** (1990) 275.
- [99] E. Barberio, B. van Eijk and Z. Was, *PHOTOS: A universal Monte Carlo for QED radiative corrections in decays*, *Comput. Phys. Commun.* **66** (1991) 115.
- [100] S. Agostinelli et al., *GEANT4: A simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [101] 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.
- [102] F. Abudinén et al., *Measurement of Lepton Mass Squared Moments in $B \rightarrow X_c \ell \bar{\nu}_\ell$ Decays with the Belle II Experiment*, *Phys. Rev. D* **107** (2023) 072002.
- [103] R. Glattauer et al., *Measurement of the decay $B \rightarrow D \ell \nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$* , *Phys. Rev. D* **93** (2016) 032006.

- [104] D. Ferlewicz et al., *Revisiting fits to $B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at nonzero recoil*, *Phys. Rev. D* **103** (2021) 112001.
- [105] J. A. Bailey et al., *Update of $|V_{cb}|$ from the $\bar{B} \rightarrow D^* \ell \bar{\nu}$ form factor at zero recoil with three-flavor lattice QCD*, *Phys. Rev. D* **89** (2014) 114504.
- [106] G. Duplancic and B. Melic, *Form factors of $B, B_s \rightarrow \eta^{(\prime)}$ and $D, D_s \rightarrow \eta^{(\prime)}$ transitions from QCD light-cone sum rules*, *J. High Energy Phys.* **11** (2015) 138.
- [107] F. U. Bernlochner and S. Wallner, *Role of ρ - ω interference in semileptonic $B \rightarrow \pi^+ \pi^- \ell \bar{\nu}_\ell$ decays*, *Phys. Rev. D* **109** (2024) 074040.
- [108] R. Garcia-Martin, R. Kaminski, J. R. Pelaez and J. R. de Elvira, *Precise Determination of the $f_0(600)$ and $f_0(980)$ Pole Parameters from a Dispersive Data Analysis*, *Phys. Rev. Lett.* **107** (2011) 072001.
- [109] C. Beleño et al., *Measurement of the branching fraction of the decay $B \rightarrow \pi \pi \ell \nu$ in fully reconstructed events at Belle*, *Phys. Rev. D* **103** (2021) 112001.
- [110] O. L. Buchmüller and H. U. Flächer, *Fit to moments of inclusive $B \rightarrow X_c \ell \bar{\nu}$ and $B \rightarrow X_s \gamma$ decay distributions using heavy quark expansions in the kinetic scheme*, *Phys. Rev. D* **73** (2006) 073008.
- [111] C. Ramirez et al., *Semileptonic $b \rightarrow u$ decays*, *Phys. Rev. D* **41** (1990) 1496.
- [112] M. T. Prim et al., *Search for $B^+ \rightarrow \mu^+ \nu_\mu$ and $B^+ \rightarrow \mu^+ N$ with inclusive tagging*, *Phys. Rev. D* **101** (2020) 032007.
- [113] T. Keck et al., *The Full Event Interpretation*, *Comput. Softw. Big Sci.* **3** (2019) 6.
- [114] J.-F. Krohn et al., *Global decay chain vertex fitting at Belle II*, *Nucl. Instrum. Meth. A* **976** (2020) 164269.
- [115] G. Brandenburg et al., *Charged track multiplicity in B meson decay*, *Phys. Rev. D* **61** (2000) 072002.
- [116] B. Aubert et al., *Measurements of the $B \rightarrow D^*$ form factors using the decay $\bar{B}^0 \rightarrow D^{*+} e^- \nu_e$* , *Phys. Rev. D* **74** (2006) 092004.
- [117] E. Waheed et al., *Measurement of the CKM Matrix Element $|V_{cb}|$ from $B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$ at Belle*, *Phys. Rev. D* **100** (2019) 052007.
- [118] I. Adachi et al., *Determination of $|V_{cb}|$ using $\bar{B}^0 \rightarrow D^{*+} \ell^- \bar{\nu}_\ell$ decays with Belle II*, *Phys. Rev. D* **108** (2023) 092013.
- [119] A. G. Patriota, *A classical measure of evidence for general null hypotheses*, *Fuzzy Sets and Systems* **233** (2013) 74.
- [120] T. Keck, *FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*, *Comput. Softw. Big Sci.* **1** (2017) 2.
- [121] M. Rohrken, *Time-Dependent CP Violation Measurements in Neutral B Meson to Double-Charm Decays at the Japanese Belle Experiment*, IEKP-KA-2012-13, PhD Thesis: Karlsruhe Institute of Technology, 2012.

-
- [122] E. A. J. Bevan, B. Golob, T. Mannel, S. Prell, and B. D. Yabsley, *The physics of the B factories*, *Eur. Phys. J. C* **74** (2014) 3026.
- [123] G. C. Fox and S. Wolfram, *Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes*, *Phys. Rev. Lett.* **41** (1978) 1581.
- [124] D. M. Asner et al., *Search for exclusive charmless hadronic B decays*, *Phys. Rev. D* **53** (1996) 1039.
- [125] F. J. Massey, *The Kolmogorov-Smirnov Test for Goodness of Fit*, *Journal of the American Statistical Association* **46** (1951) 68.
- [126] 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.
- [127] S. Choudhury et al., *Measurement of the B^+/B^0 production ratio in e^+e^- collisions at the $\Upsilon(4S)$ resonance using $B \rightarrow J/\psi(\ell\ell)K$ decays at Belle*, *Phys. Rev. D* **107** (2023) L031102.
- [128] B. Efron, *Bootstrap Methods: Another Look at the Jackknife*, *The Annals of Statistics* **7** (1979) 1.
- [129] J. Benesty, J. Chen, Y. Huang and I. Cohen, *Noise Reduction in Speech Processing*, vol. 2, 2009 1.
- [130] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.
- [131] Brun, R. and Rademakers, F., *ROOT - An Object Oriented Data Analysis Framework*, <https://doi.org/10.5281/zenodo.3895860>.

Acronyms

ARICH	Aerogel R ing- I maging C herenkov
AUC	Area u nder the ROC C urve
basf2	B elle II (2) analysis software f ramework
BCL	B ourrely- C aprini- L ellouch
BDT	B oosted D ecision T ree
BGL	B oyd- G rinstein- L ebed
BK	B ecirevic- K aidalov
BLNP	B osch- L ange- N eubert- P az
BSZ	B harucha- S traub- Z wicky
CDC	C entral D rift C hamber
CKF	C ombinatorial K alman F ilter
CKM	C abibbo- K obayashi- M askawa
CM	C enter of M ass
DAQ	D ata A cquisition
DEPFET	D epleted F ield E ffect T ransistor
DFN	D e F azio- N eubert
DOF	D egrees of F reedom
DSSD	D ouble- S ided S ilicon S trip D etector
ECL	E lectromagnetic C alorimeter
FOM	F igure of M erit
FPGA	F ield P rogrammable G ate A rray
GGOU	G ambino- G iordano- O ssola- U raltsev
h.c.	H ermitian conjugate
HER	H igh- E nergy R ing
HFLAV	H eavy F lavor A veraging
HLT	H igh- L evel T rigger
HQE	H eavy Q uark E xpansion
HQET	H eavy Q uark E ffective T heory
IP	I nteraction P oint
KEK	H igh E nergy A ccelerator R esearch O rganization
KLM	K_L and M uon D etector
L1	L evel 1
LCSR	Q CD L ight- C one S um R ules
LER	L ow- E nergy R ing
LHC	L arge H adron C ollider

Acronyms

LQCD	L attice Q CD
MC	M onte C arlo
ML	M achine L earning
MOSFET	M etal O xide F ield E ffect T ransistor
MVA	M ultivariate A nalysis
OPE	O perator P roduct E xpansion
PDG	P article D ata G roup
PID	P article I dentification
PXD	P ixel D etector
QCD	Q uantum C hromod Y namics
QED	Q uantum E lectro D ynamics
Ref.	Reference
RMS	R oot- M ean- S quare
ROC	R eciever O perating C haracteristic
ROE	R est of E vent
RPC	R esistive- P late C hamber
SGBDT	S tochastic G radient B oosted D ecision T ree
SM	S tandard M odel of particle physics
stat	statistical
SVD	S ilicon V ertex D etector
syst	systematic
theo	theoretical
TOP	T ime- o f- P ropagation
VXD	V ertex D etector

List of Figures

2.1	Graphic representation of the Standard Model of particle physics	4
2.2	Visual representation of the CKM unitarity triangle	11
2.3	Feynman diagram for the semileptonic decay $B \rightarrow X\ell\nu$	11
2.4	A two-dimensional plot of $ V_{ub} $ versus $ V_{cb} $	19
2.5	Constraints on the sides and angles of the CKM unitarity triangle in the $(\bar{\eta}, \bar{\rho})$ plane	20
3.1	Sketch of the layout of the SuperKEKB electron-positron accelerator	22
3.2	Sketch of $\Upsilon(4S)$ production with subsequent production of a $B\bar{B}$ pair	23
3.3	Sketch of the layout of the Belle II detector	25
4.1	Invariant mass $m_{\pi\pi}$ spectrum of $\rho \rightarrow \pi\pi$ in $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ decays	40
4.2	Partial branching-fraction spectrum of $B^+ \rightarrow \pi^+\pi^-\ell^+\nu_\ell$ as a function of $m_{\pi\pi}$	41
4.3	True distributions of m_X , E_ℓ^B and q^2 in $B \rightarrow X_u\ell\nu_\ell$ decays	43
5.1	Normalized simulated distributions of the pion, electron, and muon PID likelihood ratios	48
5.2	Normalized simulated distributions of the CM lepton momentum	50
5.3	Simulated distributions of the number of tracks, $\cos\theta_{B\gamma}$, and θ_{miss}	51
5.4	Simulated distributions of ΔE and M_{bc} before selections	55
5.5	Two-dimensional distributions of ΔE and M_{bc} for simulated signal and background events	56
5.6	Illustration of the cones used in the Diamond Frame and ROE method	58
5.7	Resolution in q^2 achieved in simulation	58
5.8	Migration matrices from the combined method for simulated signal events	59
6.1	Simulated distributions of ΔE and M_{bc} after selections	61
6.2	Illustration of the continuum and $B\bar{B}$ event shapes	65
6.3	Input variables and relative feature importance for the continuum suppression BDT classifier	67
6.4	Input variables and relative feature importance for the $B\bar{B}$ background suppression BDT classifier	68
6.5	ROC curves of the continuum and $B\bar{B}$ background suppression BDT classifiers	69
6.6	Example distributions of the continuum and $B\bar{B}$ background suppression classifiers in simulated and collision data	70
6.7	FOM as a function of the thresholds on the continuum and $B\bar{B}$ background suppression classifiers	71
6.8	Signal efficiencies as a function of true q^2	73

7.1	Prefit distributions of ΔE , M_{bc} , and q^2 for simulated and collision data	75
7.2	Distributions of ΔE , M_{bc} , and q^2 in simulated continuum and off-resonance data . . .	81
7.3	Composition of the total signal component in each true q^2 bin	83
7.4	Composition of events scaling with the number of true signal events in each true q^2 bin	84
7.5	Distributions of ΔE and M_{bc} in the q^2 bins in collision data with fit projections overlaid	86
7.6	Illustration of the statistical correlation matrix of the fit component scale factors . . .	87
7.7	Distributions of ΔE and M_{bc} in three reconstructed q^2 bins in collision data for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ candidates with one-dimensional fit projections overlaid	89
7.8	Distributions of ΔE and M_{bc} in three reconstructed q^2 bins in collision data for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ candidates with one-dimensional fit projections overlaid	90
7.9	Fitted background scale factors from the nominal, electron and muon fits	90
7.10	Measured differential rates as a function of true q^2	91
9.1	Measured partial branching fractions and fitted differential rates: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$. . .	108
9.2	Measured partial branching fractions and fitted differential rates: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. . .	109
9.3	Measured $ V_{ub} $ values with varying cut-off values of q^2 in the fit.	111
A.1	Plots of the FOM for bremsstrahlung parameter optimization	116
A.2	Distributions of track p_t for ROE optimization	116
A.3	Distributions of cluster energy for ROE optimization	117
A.4	Distributions of q^2 in the control mode and resolution as a function of E_{cms}	118
A.5	Fractional differences between RMS values of q^2 distributions in generated samples .	119
A.6	Distributions of ΔE in off-resonance and reweighted simulated data for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	120
A.7	Distributions of M_{bc} in off-resonance and reweighted simulated data for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	121
A.8	Distributions of ΔE in off-resonance and reweighted simulated data for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	122
A.9	Distributions of M_{bc} in off-resonance and reweighted simulated data for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	123
B.1	Prefit distributions of ΔE and M_{bc} in reconstructed q^2 bins: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (1)	126
B.2	Prefit distributions of ΔE and M_{bc} in reconstructed q^2 bins: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (2)	127
B.3	Prefit distributions of ΔE and M_{bc} in reconstructed q^2 bins: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (3)	128
B.4	Prefit distributions of ΔE and M_{bc} in reconstructed q^2 bins: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (1)	129
B.5	Prefit distributions of ΔE and M_{bc} in reconstructed q^2 bins: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (2)	130
B.6	Postfit distributions of ΔE in reconstructed q^2 bins: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	131
B.7	Postfit distributions of M_{bc} in reconstructed q^2 bins: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	132
B.8	Postfit distributions of ΔE in reconstructed q^2 bins: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	133
B.9	Postfit distributions of M_{bc} in reconstructed q^2 bins: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	134

List of Tables

4.1	Branching fractions of the $B \rightarrow X_c \ell \nu_\ell$ decays used in simulation	37
4.2	Branching fractions of the $B \rightarrow X_u \ell \nu_\ell$ decays used in simulation	37
4.3	Form-factor coefficients for $B \rightarrow D \ell \nu_\ell$, $B \rightarrow \pi \ell \nu_\ell$ and $B \rightarrow D^* \ell \nu_\ell$ decays used in the simulation	38
4.4	Form-factor coefficients for $B \rightarrow \rho \ell \nu_\ell$ and $B \rightarrow \omega \ell \nu_\ell$ decays, and free parameters for $B \rightarrow \eta^{(\prime)} \ell \nu_\ell$ decays used in simulation	39
6.1	Continuum and $B\bar{B}$ background classifier thresholds	72
6.2	Simulated signal efficiencies in each true q^2 bin after all selections	73
7.1	Summary of templates and scale factors for all fit components	79
7.2	Signal strengths in each true q^2 bin from simulation	84
7.3	Expected yields and fitted scale factors for the background templates	87
7.4	Expected yields and fitted scale factors for each $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ signal template	88
7.5	Expected yields and fitted scale factors for each $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal template	88
7.6	Partial branching fractions $\Delta\mathcal{B}$ in each true q^2 bin	92
8.1	Fractional uncertainties by source on the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ partial branching fractions	95
8.2	Fractional uncertainties by source on the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ partial branching fractions	96
9.1	Measured values of $ V_{ub} $ and the BCL form-factor coefficients: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	108
9.2	Measured values of $ V_{ub} $ and the BSZ form-factor coefficients: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	109
9.3	Fractional uncertainties by source on the extracted $ V_{ub} $ values	110
A.1	Requirements for the reconstruction of the $B^+ \rightarrow J/\psi K^+$ control mode	118
C.1	Statistical correlation matrix of the fit scale factors	136
C.2	Full experimental correlation matrix of the partial branching fractions: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$	137
C.3	Full experimental correlation matrix of the partial branching fractions: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	137
C.4	Full correlation matrix of $ V_{ub} $ and the BCL form-factor coefficients: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$, LQCD	137
C.5	Full correlation matrix of $ V_{ub} $ and the BCL form-factor coefficients: $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$, LQCD + LCSR	138
C.6	Full correlation matrix of $ V_{ub} $ and the BSZ form-factor coefficients: $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	138