TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Determination of $|V_{cb}|$ using $B \to D\ell\nu_\ell$ decays in Belle II data

zur Erlangung des akademischen Grades

Doktor der Technischen Wissenschaften

eingereicht von

Philipp Horak, M.Sc.
Matrikelnummer 01225685

an der Fakultät für Physik der Technischen Universität Wien

in Zusammenarbeit mit dem Institut für Hochenergiephysik
der Österreichischen Akademie der Wissenschaften

Betreuung: **Privatdoz. Dipl.-Ing. Dr.techn. Christoph Schwanda**
E141 - Atominstitut, TU Wien

Philipp Horak

Christoph Schwanda

Wien, 7. August 2024

Kurzfassung

Diese Dissertation behandelt die langjährig beobachtete Abweichung zwischen exklusiven und inklusiven Bestimmungen des CKM-Matrixelements $|V_{cb}|$, indem der exklusive Ansatz mittels $B^+ \to \bar{D}^0 \ell^+ \nu_\ell$ und $B^0 \to D^- \ell^+ \nu_\ell$ Zerfällen untersucht wird. Unter Verwendung von Daten des Belle II-Experiments, welches seinen Betrieb in 2019 aufnahm, wird eine umfangreiche Datenmenge von B-Meson-Zerfällen analysiert. Die Dissertation präsentiert ein vorläufiges Ergebnis basierend auf Daten, die zwischen 2019 und 2021 gesammelt wurden, sowie eine aktualisierte Messung mit etwa der doppelten Datenmenge.
Um die differentiellen Zerfallsraten zu bestimmen werden die Signalraten der verschiedenen Endzustände als Funktion des hadronischen Rückstoßes w extrahiert und in differentielle Zerfallsraten umgewandelt. Der theoretische Formfaktor beschreibt die Abhängigkeit der differentiellen Zerfallsraten von w. Ein Fit des Formfaktors an die gemessenen Zerfallsraten, unter Verwendung von zusätzlichen theoretischen Daten, ermöglicht die Messung von $|V_{cb}|$. Rekonstruktion von $B \to D\ell\nu_\ell$ Zerfällen in einer Datenmenge von 189 fb^{-1} und Verwendung des Formfaktors in der Parametrisierung von Boyd, Grinstein und Lebed [1, 2], ergibt $|V_{cb}| = (38.28 \pm 1.16) \times 10^{-3}$, wobei der Fehler aus Unsicherheiten aus statistischen und systematischen Quellen sowie theoretischen Eingaben besteht. Das Ergebnis liegt innerhalb einer Standardabweichung des weltweiten Durchschnitts aller exklusiven Bestimmungen $|V_{cb}| = (39.10 \pm 0.50) \times 10^{-3}$ [3]. Die gesamte relative Unsicherheit auf $|V_{cb}|$ entspricht etwa 3%, vergleichbar mit der Messung von $B \to D\ell\nu_\ell$ Zerfällen bei Belle, die mit einem etwa viermal größeren Datensatz eine Gesamtunsicherheit von etwa 2.8% aufweist [4].
Die Sensitivitätsschätzung für die aktualisierte Messung prognostiziert eine Reduktion der relativen Unsicherheit auf etwa 2.1%, was einer weltweit führenden Präzision für die Messung von $|V_{cb}|$ mit $B \to D\ell\nu_\ell$ Zerfällen entspricht. Diese Verbesserung ist unter anderem auf eine größere Datenmenge, verbesserte Signalextraktion und detaillertere Behandlung von systematischen Unsicherheiten zurückzuführen. Darüber hinaus kann eine Messung der Lepton-Flavour-Universalität durchgeführt werden, indem Verzweigungsverhältnisse mit Elektronen im Endzustand mit denen mit Myonen verglichen werden. Die geschätzte gesamte Unsicherheit auf $R_{e/\mu}$ beträgt 2%.

Abstract

This thesis addresses the long-standing tension between exclusive and inclusive determinations of the CKM matrix element $|V_{cb}|$ by focusing on the exclusive approach using $B^+ \to \bar{D}^0 \ell^+ \nu_\ell$ and $B^0 \to D^- \ell^+ \nu_\ell$ decays. Leveraging data from the Belle II experiment, which began operations in 2019, a substantial data sample of B meson decays is analyzed. The thesis presents a preliminary result based on data collected between 2019 and 2021, and an updated measurement using a data sample approximately twice that size.

A binned template fit is employed to extract the signal yields for the different modes in bins of the hadronic recoil w. The signal yields are unfolded and converted into differential decay rates. The theoretical form factor describes the shape of the differential decay rates as a function of w. A fit of the form factor to decay rates measured in experiment, with data from theory as additional constraining factor, allows measuring $|V_{cb}|$.

Fitting the form factor in the parameterization by Boyd, Grinstein and Lebed [1, 2] to a spectrum of differential decay rates obtained from reconstructing $B \to D\ell\nu_\ell$ decays in a data sample corresponding to 189 fb^{-1} yields $|V_{cb}| = (38.28 \pm 1.16) \times 10^{-3}$, with the error comprising uncertainties from statistical and systematic sources, as well as contributions from theoretical inputs. The result lies within one standard deviation of the world average across all exclusive determinations $|V_{cb}| = (39.10 \pm 0.50) \times 10^{-3}$ [3]. The measurement is sensitive to $|V_{cb}|$ at a total relative uncertainty of $\sim$3%, comparable to the Belle measurement of $B \to D\ell\nu_\ell$ decays, resulting in a total uncertainty of $\sim$2.8% [4] using a data sample approximately four times larger.

The sensitivity estimate for the updated measurement projects a relative uncertainty reduction to approximately 2.1%, corresponding to world-leading precision for measuring $|V_{cb}|$ with the $D\ell\nu$ final state. This improvement is attributed to, among other factors, an increased data sample, enhanced signal extraction, refined unfolding methods, and improved form factor fitting. In addition, light-lepton flavour universality can be measured by comparing branching ratios with electrons in the final state to those with muons, with an estimated total sensitivity on $R_{e/\mu}$ of 2%.

Show me a completely smooth operation
and I'll show you someone who's covering mistakes.
Real boats rock.

FRANK HERBERT

Acknowledgements

First, I would like to express my appreciation to my supervisor Christoph for always providing an open ear for me during my time as a graduate student. I am thankful for all the extensive conversations, not only about physics but also about the broader picture. Your always-open office door made a real difference, and I really appreciate it.

I am grateful to the Belle II collaboration for providing an excellent working environment. The collaboration meetings in Japan served as a valuable reminder that I am working with not just email addresses, but rather engaging and talented individuals. I would like to thank the semileptonic working group, and Markus in particular, whose insightful feedback helped me navigate the most challenging moments. I would also like to thank Kenta-san for his consistently kind and supportive attitude when helping us solve problems in the LeptonID group.

I feel extremely lucky to have had the privilege to work alongside my colleagues at HEPHY, in particular Daniel, Géraldine, Huw, Michel, Nadia, Paul, and Rajesh. Our dynamics truly felt like a family and I will miss the time we spent together a lot. Thanks to you guys, I would not trade these three years for anything. Maybe we can go for just one more beer sometime.

I would like to extend my gratitude to HEPHY Vienna for providing the essential infrastructure. Special thanks to the FWF project P 34529 for its crucial funding support, and the EU grant JENNIFER2 for providing additional travel support, which allowed me to work productively in Japan.

Finally, I owe a huge *Dankeschön* to my family for their constant and unwavering support. Your encouragement has been an important foundation for me, and I couldnt have done any of this without you. Thank you.

Contents

Kurzfassung iii

Abstract v

Contents xi

1 Introduction 1
- 1.1 Motivation . . . 1
- 1.2 Conventions and operational practices . . . 2

2 $b \to c$ transitions in the Standard Model of Particle Physics 5
- 2.1 The Standard Model . . . 5
- 2.2 Quantum field theoretical formulation of interactions . . . 7
  - 2.2.1 Quantum electrodynamics . . . 8
  - 2.2.2 Electroweak unification and the Higgs mechanism . . . 9
  - 2.2.3 Quantum chromodynamics . . . 16
- 2.3 Semileptonic $B \to D\ell\nu_\ell$ decays . . . 18
  - 2.3.1 $B \to D\ell\nu_\ell$ form factor parameterizations . . . 19
- 2.4 Current experimental status of $|V_{cb}|$ measurements . . . 21

3 The Belle II experiment at the SuperKEKB collider 23
- 3.1 B factories . . . 23
- 3.2 The SuperKEKB collider . . . 24
- 3.3 The Belle II detector . . . 26
  - 3.3.1 Coordinate system . . . 29
  - 3.3.2 Pixel and vertex detectors . . . 29
  - 3.3.3 Central drift chamber . . . 30
  - 3.3.4 Particle identification . . . 31
  - 3.3.5 Electromagnetic calorimeter . . . 32
  - 3.3.6 Solenoid magnet . . . 32
  - 3.3.7 K_L and muon detector . . . 32
- 3.4 Trigger . . . 32
- 3.5 The Belle II Analysis Software Framework . . . 33
- 3.6 Monte Carlo simulation . . . 33

3.7 Event reconstruction . . . 34
3.7.1 Charged particle identification . . . 35
3.7.2 Full Event Interpretation . . . 36

4 Statistical data analysis in high-energy particle physics 39
4.1 Multivariate analysis . . . 39
4.1.1 Figures of merit . . . 39
4.1.2 Multivariate selections . . . 41
4.2 Binned maximum likelihood fits . . . 44
4.3 Systematic uncertainties . . . 48
4.3.1 Toy resampling . . . 48
4.3.2 Eigenvalue decomposition . . . 48
4.4 Blind analysis . . . 49

5 Calibrating electron identification using Bhabha scattering events 51
5.1 Bhabha scattering . . . 51
5.2 Tag-and-probe method . . . 52
5.3 Data samples . . . 54
5.4 Candidate selection . . . 54
5.4.1 L1 Trigger lines . . . 56
5.4.2 HLT Trigger lines . . . 56
5.5 Binning . . . 57
5.6 Systematic uncertainties . . . 59
5.6.1 Trigger line selection . . . 59
5.6.2 Background estimation . . . 59
5.7 Results . . . 59

6 Determination of $|V_{cb}|$ from exclusive $B \to D\ell\nu_\ell$ decays 63
6.1 Analysis overview . . . 63
6.2 Input data samples . . . 66
6.3 Variable definitions . . . 66
6.3.1 Track impact parameters . . . 66
6.3.2 Event properties . . . 67
6.3.3 Kinematic variables . . . 69
6.3.4 Rest-of-event variables . . . 70
6.3.5 Particle identification . . . 70
6.4 Reconstruction of $B \to D\ell\nu_\ell$ and event selection . . . 71
6.4.1 Candidate reconstruction . . . 71
6.4.2 Candidate classification . . . 73
6.4.3 $B \to D^*\ell\nu_\ell$ vetoes . . . 73
6.4.4 Background suppression . . . 75
6.4.5 Sample composition and signal efficiency . . . 76
6.4.6 Kinematic variable reconstruction . . . 78
6.5 Reconstruction of control samples . . . 82

6.5.1 Background control samples 82
6.5.2 Signal control samples 85
6.6 External corrections 87
6.7 Control sample derived corrections and validation 90
6.8 Comparison between Data and MC 93
6.9 Signal extraction 96
6.9.1 Template fit setup 96
6.9.2 Systematic uncertainties 98
6.9.3 Recovering the MC expectation 104
6.10 Form factor fit 105
6.11 Fit validation 107
6.12 Results 109
6.12.1 Sensitivity estimate 109
6.12.2 Fit to data 115

7 Summary and outlook **119**

A Electron identification correction factors **121**

B Cutflow diagrams **127**

C Selection comparison in Data and MC **133**

List of Figures **139**

List of Tables **145**

Bibliography **147**

CHAPTER 1

Introduction

1.1 Motivation

To explain phenomena observed in weak transitions, particularly the suppression of strange-particle decays, Nicola Cabibbo postulated the concept of quark mixing in 1963 [5], introducing a systematic difference in how strongly the flavour-changing weak interaction acts on free quark states. In 1973, Makoto Kobayashi and Toshihide Maskawa proposed to expand the quark mixing mechanism to a third generation of quarks [6] to explain observations of charge-parity symmetry (CP) violating phenomena in neutral kaon systems. The third generation of quarks was subsequently discovered [7–11], confirming their proposed theory of a 3x3 quark mixing matrix, now known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The nine elements of the matrix, characterizing the strength of the weak interaction between the different quark types, are fundamental parameters of the Standard Model, *i.e.* they have to be determined through experiment.
Of the nine different coupling strengths, the up-to-date least well-known elements are the couplings between bottom and up quarks, and bottom and charm quarks, V_{ub} and V_{cb}. Studying decays of B mesons offers access to measuring these parameters. B factory experiments, such as Belle II, operate at collision energies fine-tuned to produce a large amount of B mesons, with a detector around the interaction point observing the decay products and measuring their properties.
Decays of B mesons into hadrons are ideal for reconstruction at B factories because all final state particles are detectable, allowing for precise measurements. However, interactions between the hadronic final state systems complicate theoretical calculations of the amplitudes needed to determine $|V_{cb}|$ and $|V_{ub}|$ from measurements. On the other hand, purely leptonic decays of B and B_c mesons provide the cleanest environment for theoretical predictions, but their small branching ratios necessitate larger data samples than are currently available to achieve competitive measurements. Semileptonic B decays, such as $B \to \pi\ell\nu_\ell$ and $B \to D\ell\nu_\ell$, offer a balanced compromise between theory and experiment. The factorizable leptonic and hadronic currents enable theoretical computations of high precision, while the

relatively high branching ratios allow statistically significant measurements, despite the challenge posed by the invisible neutrino in the reconstruction process.
In general, there are two possible approaches of measuring $|V_{cb}|$ and $|V_{ub}|$ via semileptonic decays: *exclusive* determination through reconstructing a specific final state – such as $B \to D\ell\nu_\ell$ – or *inclusive* determination of all $B \to X_{c/u}\ell\nu_\ell$ decays, where $X_{c/u}$ refers to any hadronic system containing an up or charm quark. Due to the independent inputs from theorists in the two methods they are complementary, and it is crucial to study both approaches. Unfortunately, a long-standing tension between exclusive and inclusive determinations has been observed, making the measurements irreconcilable. The combined tension between all exclusive and inclusive determinations of $|V_{cb}|$ and $|V_{ub}|$ reaches more than three standard deviations. While individual measurements are performed with high precision, the observed tension cloudens the knowledge of these fundamental parameters.
The B factory experiment Belle II, which has been operating since 2018 in Tsukuba, Japan, is the successor to the original Belle experiment. By upgrading the accelerator and all detector parts, Belle II is designed to accumulate data containing an unprecedented amount of B decays. From the start of collision data collection in 2019 to the first long shutdown in 2022, Belle II collected a data sample corresponding to 387 million pairs of B mesons, roughly equal to half of the total Belle data sample collected over 11 years. With future upgrades planned, the rate of data taking at Belle II is still scheduled to increase by an order of magnitude over its operating lifetime.

This work aims to determine $|V_{cb}|$ through an exclusive measurement of $B \to D\ell\nu_\ell$ decays in the Belle II data collected from 2019 to 2022. A 2016 measurement of this decay mode at Belle reached a precision of up to 2.8% on $|V_{cb}|$ [4]. The Belle II measurement is expected to be competitive with roughly half the size of the available data sample thanks to improved analysis techniques and upgraded detector performance.
The thesis is organized as follows: Chapter 2 outlines the required theoretical concepts, focusing on semileptonic B decays. In Chapter 3, the SuperKEKB accelerator, the Belle II detector with all its subdetectors and the Belle II software framework are introduced. Following this, statistical concepts and data analysis techniques relevant to the analysis are introduced in Chapter 4. A performance study of electron identification with $e^+e^- \to e^+e^-$ decays is outlined in Chapter 5, resulting in an essential input for the $|V_{cb}|$ measurement. The $B \to D\ell\nu$ measurement is described in Chapter 6, detailing the analysis procedure step-by-step. Finally, the analysis results and a future outlook are given in Chapter 7.

1.2 Conventions and operational practices

Several important conventions and practices are highlighted here and used throughout the thesis.

Natural units

This work employs the system of natural units, where fundamental constants such as the speed of light c, reduced Planck constant $\hbar$ and elementary charge are set to 1. Different

quantities can be expressed in the same units in this convention, simplifying computations. In this work, quantities such as mass, momenta, and energies are expressed in electron volts (eV) – and their multiples keV, MeV, and GeV – unless explicitly stated otherwise.

Charge conjugation

When discussing decay processes, it is standard practice to imply the inclusion of both particle decay and its charge-conjugate process. For instance, if the decay $B^0 \to D^- e^+ \nu_e$ is mentioned, it is understood that the charge-conjugate decay $\overline{B}^0 \to D^+ e^- \bar{\nu}_e$ is also considered. Throughout this thesis, unless explicitly specified, any mention of a particular decay implicitly includes its charge-conjugate counterpart.

Collaborative work

While performing the analysis of $B \to D\ell\nu_\ell$ decays as the principal analyst alongside my supervisor and the electron identification study in collaboration with Paul Feichtinger, the work represents a group effort by the Belle II and SuperKEKB collaborations. It is essential to acknowledge that the contributions of the collaborations are integral to every aspect of this research, including data collection, simulation, performance inputs, and analysis methodologies.

Belle II review process

Measurements in large high-energy physics collaborations undergo an extensive internal review process. Every qualifying member has the option to be author, and multiple steps involving different reviewing bodies ensure that the measurement meets the experiment's standards. Often, this process spans longer than a year until the final results are unblinded, *i.e.*, a measurement is obtained.
In the context of this thesis, a preliminary measurement of $|V_{cb}|$, performed by my supervisor and me as principal analysts, has been approved internally and published as preprint [12]. This thesis focuses on the update to the preliminary measurement using a larger data sample and improved analysis techniques, and provides details and sensitivity estimates using this approach. At the time of writing, the analysis is in the final stages of the internal review process; therefore, no updated result is presented in the conclusion. Instead, the preliminary result is cited while noting that differences in the analysis strategy are present at several levels.

CHAPTER 2

$b \to c$ transitions in the Standard Model of Particle Physics

This chapter provides the particle physics background necessary to contextualize the main studies of the thesis. Section 2.1 introduces the basic building blocks of the fundamental particles and the three forces that make up the Standard Model of particle physics. The quantum field theoretical formulation of the fundamental interactions is then outlined in Section 2.2, adhering to the treatment presented in Ref. [13]. Section 2.3 introduces semileptonic decays and, in particular, the state-of-the-art description of $b \to c$ transitions before the experimental status of these measurements is summarized in Section 2.4.
This chapter does not aim to provide a complete treatise of the Standard Model but rather summarizes the crucial aspects concerning the physics analyses described in the thesis. See, *e.g.*, Refs. [13–16] for a more comprehensive description.

2.1 The Standard Model

The Standard Model (SM) [13, 17] is a theory describing the fundamental particles and their interactions. Fundamental interactions carried by integer-spin *bosons* act on the half-integer spin *fermions* that constitute matter. Three of the four fundamental forces, excluding gravity, are successfully described in this common framework. The three forces of the SM, with their corresponding mediator bosons, are:

- **Electromagnetic force:** Electromagnetism is mediated by the massless photon (γ). The photon has neutral electric charge and carries the electromagnetic interaction between any electrically charged fermion.
- **Weak force:** The mediators of the weak force are the massive bosons $W^{\pm}$ and Z. The weak force is unique among fundamental interactions in the fact that it affects

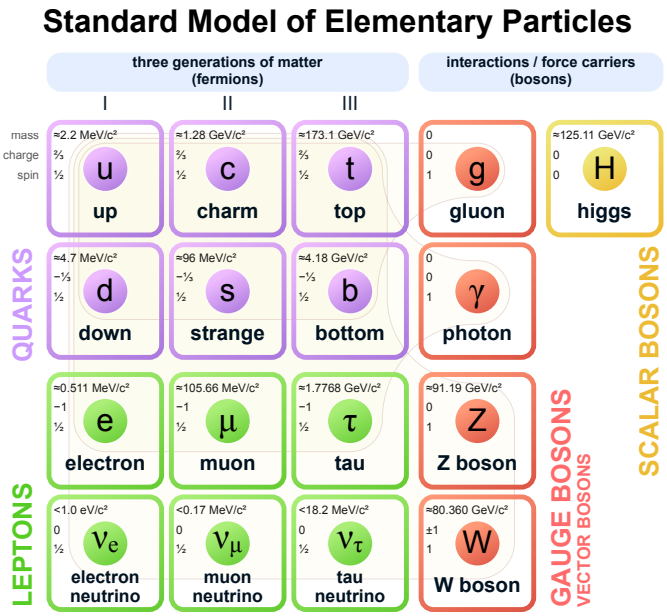


Figure 2.1: A summary table of all fundamental particles. Each particle is shown with its respective mass, charge, and spin in the upper left corner [18].

all fermions and can change one type of particle into another, such as causing a muon to decay into an electron.

- **Strong force:** The strong force is mediated by the eight types of massless gluons (g). The strong force affects any fermion carrying a color charge.

The fermions can be categorized into two groups:

- **Leptons:** Three charged, massive leptons are grouped with their respective three chargeless neutrinos

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}.$$

The charged leptons interact electromagnetically, while all six leptons interact with the weak force. The coupling strengths of each lepton are the same magnitude, a property called *lepton universality*. This suggests that each type of lepton interacts with fundamental forces in a similar manner. If their masses were identical, the probability of interaction with these forces would be the same for each lepton type. Neutrinos are almost massless, but observation of a phenomenon called neutrino oscillation [19] implies that at least two of the three neutrinos have nonzero mass. Because the SM in its current state does not incorporate a mechanism of giving mass to neutrinos, neutrinos are described as massless in the SM.

- **Quarks:** Six different types of quarks constitute the second group of fermions. They can be divided in the *up*-type quarks (u, c, t) possessing an electric charge of $+\frac{2}{3}$, and the *down*-type quarks (d, s, b) with a charge of $-\frac{1}{3}$. A similarity to leptons is that they can be grouped into three generations:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix},$$

  each consisting of an up-type and a down-type quark, with masses of similar order of magnitude. Quarks are distinct from leptons in that they carry a color charge and are, therefore, involved in the strong force. Consequently, they exhibit a property called *quark confinement* [20, 21]. They can not be isolated and thus cannot be directly observed as free particles. Quarks only occur in bound states, with the two main states being *mesons* and *baryons*. A meson state consists of a quark and its anti-quark $q\bar{q}$, and a baryon state consists of three quarks qqq.

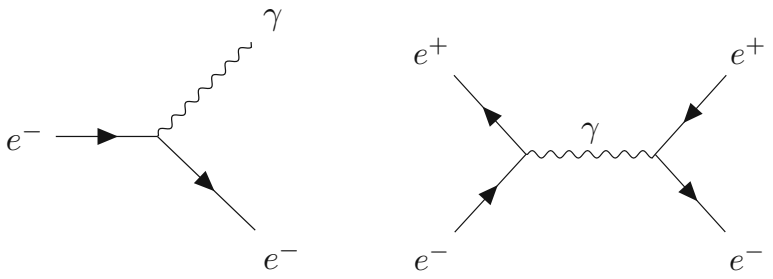
A summary of all fundamental particles and their properties is shown in Figure 2.1. In addition to the particles listed, each fermion f has its corresponding antiparticle, written $\bar{\text{f}}$. Each fundamental force exhibits distinct observable phenomena; however, they are all governed by a shared underlying principle in the SM. This principle is known as *local gauge symmetry*. This symmetry is mathematically represented by Lie groups [22], each corresponding to a set of gauge transformations under which physical laws remain invariant. Specifically, the Lie groups U(1), SU(2), and SU(3) govern electromagnetism, the weak force, and the strong force, respectively. Gauge fields are added to the formalism to guarantee invariance under these transformations. When such fields are quantized, their excitations can be identified as gauge bosons, *i.e.* the mediators carrying each fundamental interaction.

2.2 Quantum field theoretical formulation of interactions

The SM is formulated as a quantum field theory, describing free particles as excited states of the underlying fields. The dynamics of each interaction are encoded into the Lagrangian density, abbreviated Lagrangian,

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{Gauge}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \tag{2.1}$$

with the electroweak (EW) interaction providing a unified description of the electromagnetic and weak forces, and quantum chromodynamics (QCD) describing the strong force.
The time evolution of any system of fundamental particles can, in principle, be derived from the Lagrangian equations of motion of $\mathcal{L}_{\text{SM}}$. $\mathcal{L}_{\text{EW}}$ and $\mathcal{L}_{\text{QCD}}$ contain interactions of fermions with the gauge bosons γ, $W^{\pm}$, Z and g. $\mathcal{L}_{\text{Gauge}}$ contains self-interactions of gauge bosons. Masses of fundamental particles arise through interactions with the spin-0 (*scalar*) Higgs boson, as described by the Higgs mechanism [23–25]. Interactions of the



(a) (b)

Figure 2.2: Feynman diagrams for electromagnetic interactions. Subfigure (a) depicts an electron that emits a photon in a process known as *bremsstrahlung*. Subfigure (b) shows $e^+e^- \to e^+e^-$ scattering, also known as *Bhabha scattering*.

Higgs boson with gauge bosons are contained in $\mathcal{L}_{\text{Higgs}}$, giving mass to W and Z bosons. Finally, $\mathcal{L}_{\text{Yukawa}}$ consists of interaction terms between fermions and the Higgs boson, giving mass to charged leptons and quarks. The SM Lagrangian does not account for massive neutrinos, requiring further extensions to explain the observations. In the subsequent sections, the individual terms of the Standard Model Lagrangian will be described in more detail.

2.2.1 Quantum electrodynamics

The Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{2.2}$$

is a relativistic version of the quantum mechanical Schrödinger equation. It describes the time evolution of a free spinor field ψ representing a fermion with spin $\frac{1}{2}$. In its Lagrangian form, it reads

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi. \tag{2.3}$$

To arrive at the Lagrangian of quantum electrodynamics (QED), one has to impose invariance under a local U(1) gauge transformation

$$\psi(x) \to \psi'(x) = e^{ie\lambda(x)}\psi(x), \tag{2.4}$$

with a coupling constant e that can be identified as the electric charge, and an arbitrary local phase $\lambda(x)$. Equation (2.2) is by itself not invariant under the U(1) transformation, but can be made invariant by replacing the derivative ∂_μ with the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu, \tag{2.5}$$

introducing the new field A_μ that follows a transformation law

$$A_\mu \to A'_\mu + \partial_\mu \lambda. \tag{2.6}$$

This yields the QED Lagrangian

$$\mathcal{L}_{\text{QED}} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\overline{\psi}\gamma^\mu\psi A_\mu, \tag{2.7}$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The new field A_μ can be identified as the photon. The equations of motion of the QED Lagrangian recover the covariant form of Maxwell's equations, thereby encompassing classical electrodynamics.
The first and second terms in the QED Lagrangian are called the *propagators* of the fermion and the U(1) gauge boson, respectively. The third term is an interaction term of fermions with photons. *Feynman diagrams* [26] are used to illustrate allowed processes in a quantum field theory and are built out of propagators and interaction terms in the Lagrangian. A solid line represents each propagator, and a vertex represents each interaction. Example processes for QED are shown in Figure 2.2

2.2.2 Electroweak unification and the Higgs mechanism

The underlying symmetry of QED is the local gauge symmetry U(1). Repeating the procedure in the previous section for the SU(2) symmetry group by imposing invariance under the transformation

$$\varphi(x) \to \varphi'(x) = e^{ig_W\boldsymbol{\alpha}(x)\boldsymbol{\tau}}\varphi(x). \tag{2.8}$$

results in the Lagrangian

$$\mathcal{L}_{\text{SU(2)}} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \frac{1}{4}\mathbf{G}_{\mu\nu}\mathbf{G}^{\mu\nu} - (g\overline{\psi}\gamma^\mu\boldsymbol{\tau}\psi)\mathbf{B}'_\mu, \tag{2.9}$$

with a field strength tensor $\boldsymbol{F}^{\mu\nu} \equiv \partial^\mu\boldsymbol{A}^\nu - \partial^\nu\boldsymbol{A}^\mu - 2g(\boldsymbol{A}^\mu \times \boldsymbol{A}^\nu)$, the Pauli matrices $\boldsymbol{\tau} = (\sigma_1, \sigma_2, \sigma_3)$ and a new set of three gauge fields $\mathbf{B}'_\mu$. While the number of gauge fields obtained in this way corresponds to the number of Gauge bosons mediating the weak force (W^+, W^- and Z), the combined Lagrangian $\mathcal{L}_{\text{U(1)}} + \mathcal{L}_{\text{SU(2)}}$ can not be identified directly with the Lagrangians of QED and the weak force in the SM. There are two inconsistencies with the obtained Lagrangian:

- Fermion mass terms of the form $e\overline{\psi}\psi$ appearing in the Lagrangian are not symmetric under SU(2) symmetry and would lead to the imposed gauge symmetry being explicitly broken in the Lagrangian.
- Observations of massive gauge bosons are irreconcilable with the gauge symmetries in the obtained Lagrangian without introducing additional mechanisms.

By introducing an additional scalar field as part of the Higgs mechanism and unifying electromagnetic and weak force into a combined U(1) × SU(2) electroweak (EW) theory, explicit gauge symmetry in the Lagrangian can be recovered.

In the unified EW theory [27], the QED local gauge symmetry U(1) is replaced by $U(1)_Y$ with symmetry transformation

$$\varphi(x) \to \varphi'(x) = e^{ig'\frac{Y}{2}\zeta(x)}\varphi(x), \tag{2.10}$$

where introducing the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + ig'\frac{Y}{2}B_\mu \tag{2.11}$$

gives rise to a new gauge field B_μ that couples to the weak hypercharge Y with the coupling strength g'.
For SU(2) in the electroweak unification, the generators are represented by the 2×2 Pauli matrices, and the charge associated with this symmetry is known as the weak isospin, denoted by I_W. Chirality is a key feature in the weak interaction, distinguishing between left-handed and right-handed components of fermions. Left-handed fermions with $I_W = \frac{1}{2}$ are grouped into weak isospin doublets of the form

$$\varphi(x) = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \tag{2.12}$$

where the components of the doublet differ by one unit of electric charge, allowing them to be converted into each other via interactions with charged W boson. In contrast, right-handed fermions are weak isospin singlets with $I_W = 0$, meaning they do not participate in the SU(2) interactions and thus do not couple to the W bosons. Because only left-handed particles are affected, the SU(2) of the EW theory is referred to as $SU(2)_L$.
The $SU(2)_L$ Lagrangian is obtained by imposing gauge invariance analogous to Equation (2.9). The Lagrangian then contains interaction terms of the form

$$(g_W\overline{\psi}\gamma^\mu\boldsymbol{\tau}\psi)\mathbf{W}_\mu, \tag{2.13}$$

giving rise to three currents

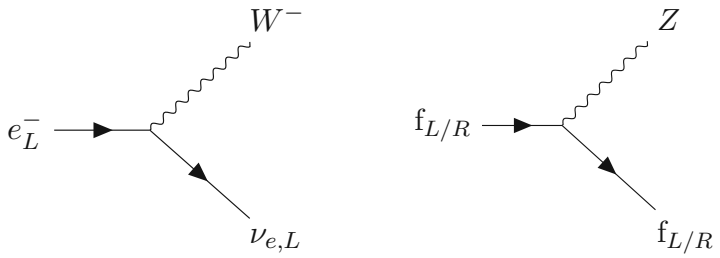
$$j_i^\mu = \frac{g_W}{2}\bar{\varphi}_L\gamma^\mu\sigma_i\varphi_L \tag{2.14}$$

for each of the three Pauli matrices.
The physical $W^\pm$ bosons act on the lower particle in an isospin doublet, changing it into the upper partner or vice versa. Such a process is shown in Figure 2.3. The charged W bosons can therefore be identified with the Pauli matrix raising and lowering operators $(\sigma_1 \pm i\sigma_2)$ and written as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^{(1)} \mp iW_\mu^{(2)}). \tag{2.15}$$

The physical photon and Z boson are obtained through mixing of the third W_μ component and the field of $U(1)_Y$, B_μ:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \tag{2.16}$$



(a) (b)

Figure 2.3: Feynman diagram of (a) charged-current and (b) neutral-current weak interactions. The letters L and R denote left- and right-handed chiral states.

with the weak mixing angle θ_W. The weak mixing angle is a free parameter of the Standard Model; therefore, there is no predicted value. It has to be measured, with the current world average of measurements determining it to be $\sin^2\theta_W = 0.23122 \pm 0.00004$ [28]. The weak mixing angle connects the $\mathrm{U}(1)_Y$ and $\mathrm{SU}(2)_L$ coupling constants to the electromagnetic coupling e through the relation

$$e = g_W \sin\theta_W = g' \cos\theta_W. \tag{2.17}$$

Higgs mechanism

To resolve the problem of gauge symmetry breaking fermion and gauge boson mass terms, it is necessary to introduce the Higgs mechanism [23–25]. After the introduction of the Higgs mechanism, rather than working with an explicit gauge broken Lagrangian, the electroweak Lagrangian is manifestly gauge invariant under $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_L$ transformations.
The minimal Higgs model consists of two complex scalar fields in a weak isospin doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.18}$$

described by the Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi),\ V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \tag{2.19}$$

where $V(\phi)$ is the so-called Higgs potential. Because the photon is required to remain massless, the minimum of the potential should have a nonzero vacuum expectation value for only ϕ^0, which can be chosen as $v/\sqrt{2}$. In the unitary gauge obtained by selecting the gauge that forces ϕ to be entirely real, the Higgs doublet can be written as an excitation $h(x)$ around its expectation value

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.20}$$

The Lagrangian resulting from the Higgs mechanism is then obtained by writing Equation (2.19) as a function of $h(x)$ and replacing derivatives with the appropriate covariant

derivatives

$$\partial_\mu \to D_\mu = \partial_\mu + ig_{\mathrm{W}}\mathbf{T}\cdot\mathbf{W}_\mu + ig'\frac{Y}{2}B_\mu. \tag{2.21}$$

The resulting form of the kinetic term is

$$\begin{aligned}(D_\mu\phi)^\dagger(D^\mu\phi) = &\frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{8}g_{\mathrm{W}}^2(W_\mu^{(1)} + iW_\mu^{(2)})(W^{(1)\mu} - iW^{(2)\mu})(v+h)^2 \\ &+\frac{1}{8}(g_{\mathrm{W}}W_\mu^{(3)} - g'B_\mu)(g_{\mathrm{W}}W^{(3)\mu} - g'B^\mu)(v+h)^2.\end{aligned} \tag{2.22}$$

By introducing the scalar Higgs field, the previously massless gauge fields acquired mass terms of the form $mW_\mu W^\mu$. The mass of the W bosons can be identified from the quadratic terms in $W_\mu^{(1)/(2)}$ as

$$m_W = \frac{1}{2}g_W v \tag{2.23}$$

and is therefore determined from the Higgs vacuum expectation value and the $\mathrm{SU}(2)_L$ coupling constant.
Substituting the physical form of photon and Z boson fields (2.16), as well as the relation between couplings (2.17), the term containing W_μ^3 and B_μ can be rewritten as

$$\frac{1}{8}v^2\begin{pmatrix} A_\mu & Z_\mu \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & g_{\mathrm{W}}^2 + g'^2 \end{pmatrix}\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix}. \tag{2.24}$$

The gauge boson mass terms can then be identified as

$$m_\gamma = 0,\ m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2}. \tag{2.25}$$

Thus, by implementing a complex scalar field ϕ with a non-vanishing vacuum expectation value, the gauge bosons $W^\pm$ and Z obtained mass terms in a fully gauge-invariant way while the photon remained massless.
Simultaneously, the Higgs mechanism also solves the problem of gauge symmetry violating fermion mass terms. Terms such as

$$-m\overline{\psi}\psi = -m(\overline{\psi}_R\psi_L + \overline{\psi}_L\psi_R), \tag{2.26}$$

decomposed into left and right-handed chiral states, do not adhere to the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry. However, combining left-handed $\mathrm{SU}(2)_L$ doublets L with the complex scalar field, the term $\overline{L}\phi = L^\dagger\gamma^0\phi$ becomes gauge invariant due to the complementary transformation properties

$$\phi \to \phi' = (I + ig_{\mathrm{W}}\boldsymbol{\epsilon}(x)\cdot\mathbf{T})\phi \tag{2.27}$$

and

$$\overline{L} \to \overline{L}' = \overline{L}(I - ig_{\mathrm{W}}\boldsymbol{\epsilon}(x)\cdot\mathbf{T}). \tag{2.28}$$

By combining it with a right-handed chiral state, the terms $\overline{L}\phi R$ and $\overline{R}\phi^\dagger L$ are invariant under both $\mathrm{U}(1)_Y$ and $\mathrm{SU}(2)_L$. By writing ϕ as excitation around v, the gauge-invariant

terms for the fermion state $\begin{pmatrix} \nu_e^- \\ e^- \end{pmatrix}$ becomes

$$\mathcal{L}_{\mathrm{e}} = -\frac{g_e}{\sqrt{2}}\nu(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{g_e}{\sqrt{2}}h(\bar{e}_L e_R + \bar{e}_R e_L) = -m_e\bar{e}e - \frac{m_e}{v}\bar{e}eh. \tag{2.29}$$

The first term can be identified as a fermion mass term with

$$m_e = \frac{g_e v}{\sqrt{2}}, \tag{2.30}$$

following directly from introducing a complex scalar field with nonzero expectation value. The term g_e is called the Yukawa coupling, resulting from the fermion coupling to the Higgs field. The second term in Equation (2.29) is interpreted as an interaction term of the fermion with the Higgs boson.
Since the complex field ϕ only has a nonzero expectation value in the neutral component, mass terms are obtained for charged leptons and down-type quarks. The procedure has to be repeated with the conjugate doublet $\phi_c = \begin{pmatrix} -\phi^{0*} \\ \phi^- \end{pmatrix}$ to arrive at mass terms for up-type quarks [1].
By generalizing to all three generations of quarks, the Lagrangian can be written as [30]

$$\mathcal{L}_{\mathrm{Yukawa}} = -\overline{Q}_L(Y^u\phi_c u_R + Y^c\phi d_R) + \mathrm{h.c.}. \tag{2.31}$$

Q_L contains the three left-handed $\mathrm{SU}(2)_L$ doublets of the form $\begin{pmatrix} u' \\ d' \end{pmatrix}_L$, u_R and d_R are the six right-handed singlets, and Y^c and Y^u are 3x3 matrices dependent on the 18 bare Yukawa couplings. The prime indicates that while these quark states are the eigenstates of the electroweak interaction, they are not the physical eigenstates of the mass operator known as flavour eigenstates. To obtain the Lagrangian as a function of flavour eigenstates, Y matrices are diagonalized as $M^f_{diag} = v/\sqrt{2}V^f_L Y^f V^{f\dagger}_R$ with $f = u, d$. Equation (2.31) can then be written as

$$\mathcal{L}_{\mathrm{Yukawa}} = -\frac{g}{\sqrt{2}}(\overline{u}_L, \overline{c}_L, \overline{t}_L)\gamma^\mu W^+_\mu V^u_L V^{d\dagger}_L \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \mathrm{h.c.}. \tag{2.32}$$

Quark mixing

Equation (2.32) contains the unitary 3×3 matrix

$$V_{\mathrm{CKM}} \equiv V^u_L V^{d\dagger}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.33}$$

[1] Treatment of neutrino masses does not follow this formalism and extensions to the Standard Model are necessary to introduce them, see e.g. [29].

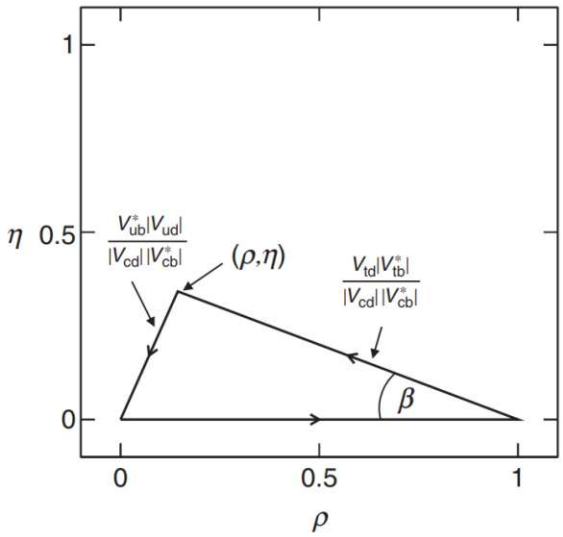


Figure 2.4: The unitarity triangle arising from the equation $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ [13].

called the Cabibbo-Kobayashi-Maskawa (CKM) matrix [5, 6], or quark mixing matrix. The CKM matrix describes how electroweak eigenstates of quarks relate to the physical flavour states. The absolute magnitude squared of the elements $|V_{qq'}|^2$ describe the relative probability of a $q \to q'$ quark transition.
Due to the unitarity of V_{CKM}it can be parameterized by three mixing angles $(\theta_{12}, \theta_{13}, \theta_{23})$ and a charge-parity symmetry violating phase δ. The common convention in literature is

$$\begin{aligned} V_{\mathrm{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \end{aligned} \tag{2.34}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$.
The nine complex elements of the CKM matrix are free parameters of the standard model that must be measured. The most recent measurements [28] determine the magnitude of the elements to

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.97435 \pm 0.00016 & 0.22500 \pm 0.00067 & 0.00369 \pm 0.00011 \\ 0.22486 \pm 0.00067 & 0.97349 \pm 0.00016 & 0.04182^{+0.00085}_{-0.00074} \\ 0.00857^{+0.00020}_{-0.00018} & 0.04110^{+0.00083}_{-0.00072} & 0.999118^{+0.000031}_{-0.000036} \end{pmatrix}. \tag{2.35}$$

The observed near-diagonal structure of the CKM matrix, with small off-diagonal contributions, allows the so-called Wolfenstein parameterization [32], an expansion around the

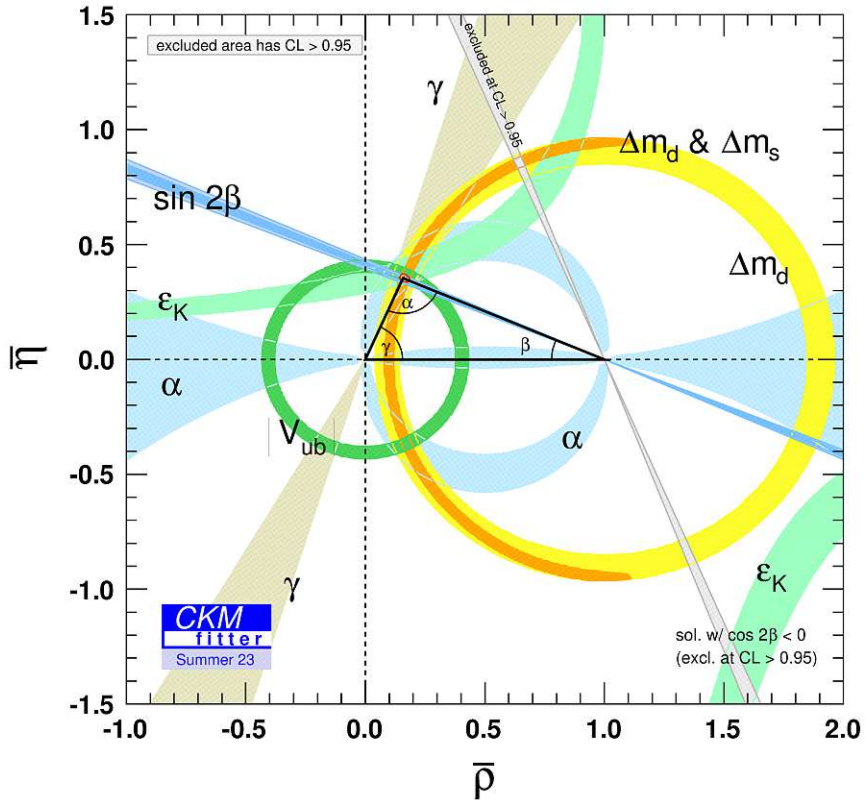


Figure 2.5: Experimental constraints on the unitarity triangle in the $\rho - \eta$ plane [31]. The red hashed region around the top vertex of the triangle indicates the 68% confidence interval of the global combination.

parameter $\lambda = \sin\theta_{12} \simeq 0.225$. In this parameterization, V_{CKM} can be written as

$$\begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{2.36}$$

with the definitions

$$A = \frac{s_{23}}{s_{12}^2}, \ \rho = \mathfrak{Re}\left(\frac{s_{13}e^{-i\delta}}{s_{12}s_{23}}\right), \quad \eta = -\mathfrak{Im}\left(\frac{s_{13}e^{-i\delta}}{s_{12}s_{23}}\right). \tag{2.37}$$

The approximation up to $O(\lambda^4)$ is accurate to ∼0.3%.
The unitarity of the CKM matrix allows for a powerful test of the standard model through the determination of its elements. The unitarity conditions can be stated as $\sum_i V_{ij}V_{ik}^* = \delta_{ij}$ and $\sum_j V_{ij}V_{kj}^* = \delta_{ik}$. These constraints can be expressed as the so-called unitarity triangles.

The most commonly used of these triangles is obtained by dividing the equation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.38}$$

by $V_{cd}V_{cb}^*$. Using the Wolfenstein parameterization, the resulting triangle in the $\rho - \eta$ plane is shown in Figure 2.4.
By performing a global fit to all measurements, the closure of the unitarity triangle can be tested. The CKMfitter [31] and UTfit [33] groups perform these global fits independently, with agreement between results. The current experimental picture of constraints, obtained by the CKMfitter group, imposed on the unitarity triangle is shown in Figure 2.5.

2.2.3 Quantum chromodynamics

The final term of the SM Lagrangian Equation (2.1) yet to be discussed is the contribution from the strong interaction. The underlying symmetry group of the strong interaction is the SU(3) gauge group. In an analogous approach as employed in the electroweak unification, the QCD Lagrangian is obtained by imposing upon the Dirac equation invariance under the transformation

$$\psi(x) \to \psi'(x) = e^{ig_S\alpha_k(x)\hat{T}^k}\psi(x) \tag{2.39}$$

and introducing the covariant derivative and the new gauge fields

$$(D_\mu)_{ij} = \partial_\mu\delta_{ij} - ig(\hat{T}_a)_{ij}\mathcal{A}_\mu^a \tag{2.40}$$

with the transformation property

$$\mathcal{A}_\mu^k \to \mathcal{A}_\mu^{k\prime} = \mathcal{A}_\mu^k - \partial_\mu\alpha_k - g_S f_{ijk}\alpha_i\mathcal{A}_\mu^j. \tag{2.41}$$

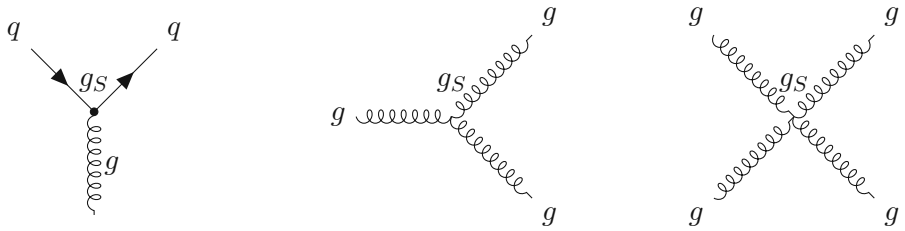
Here, $\hat{T}^k$ are the 8 Gell-Mann matrices and f_{ijk} are the structure constants that follow from their commutation relations $[\lambda_a, \lambda_b] = 2i\sum_c f^{abc}\lambda_c$,. The Gell-Man matrices act as the generators of the symmetry group and give rise to the eight gauge fields $\mathcal{A}^k$. They can be identified as the eight gluons that couple to quark fields with the coupling strength g_S. The resulting Lagrangian is

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}_i\left(i\gamma^\mu(D_\mu)_{ij} - m\,\delta_{ij}\right)\psi_j) - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}, \tag{2.42}$$

with the gluon field strength tensor

$$G_{\mu\nu}^a = \partial_\mu\mathcal{A}_\nu^a - \partial_\nu\mathcal{A}_\mu^a + gf^{abc}\mathcal{A}_\mu^b\mathcal{A}_\nu^c. \tag{2.43}$$

In contrast to the electroweak theory, which features weak isospin and hypercharge as its fundamental charges, QCD has three conserved charges known as color charges. Only particles that carry color charge, *i.e.* quarks and gluons, partake in the strong interaction. In contrast to the electroweak interaction, the non-vanishing commutation relations of the SU(3) generators lead to terms such as $\mathcal{A}_\mu^b\mathcal{A}_\nu^c$ that give rise to self-interactions between gluons, as shown in their Feynman diagrams in Figure 2.6. The gluon self-interactions are a potential contributor to the observed phenomenon of color confinement [20]. Ample



(a) (b) (c)

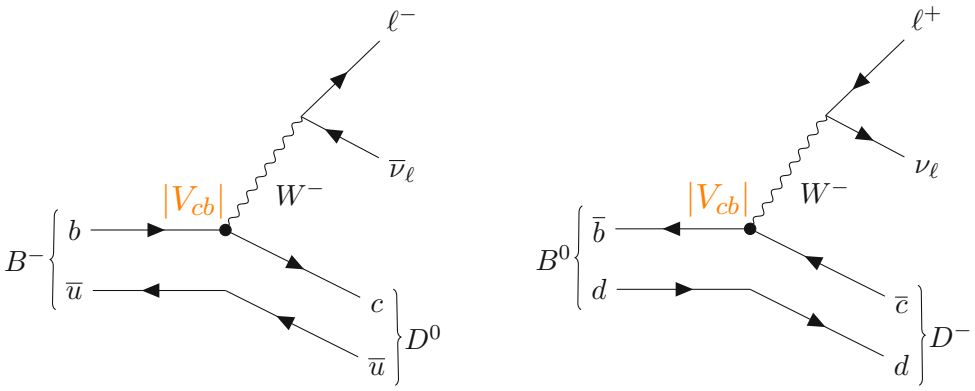
Figure 2.6: Feynman diagrams of (a) quark-gluon interactions, (b) three-point gluon self-interactions, and (c) four-point gluon self-interactions.

experimental evidence validates the quark-gluon interaction model of QCD. However, there have been no observations of free quarks. The color confinement hypothesis states that all observed quark systems must be in color singlet states, *i.e.* states that carry neutral total color charge. This is the case for *e.g.* $q\bar{q}$ mesons and qqq baryons, but not for isolated quarks q. All confirmed observed hadronic states so far have been in color singlet states, including mesons and baryons but also exotic $qqqq\bar{q}$ pentaquark states [34, 35].

Coupling constants of quantum field theories are not constant but depend on the energy scale of the considered process. A distinctive feature of the strong interaction is that, in contrast to the electromagnetic and weak interactions, is its large coupling strength $\alpha_s \sim O(1)$. Consequently, QCD can not be treated as a perturbative theory where tree-level processes dominate rate calculations. Nonperturbative techniques must be considered to evaluate the contributions of strong interactions.

A prominent nonperturbative approach is lattice QCD (LQCD) [36]. In lattice QCD, the continuous spacetime is discretized with a grid spacing a. Quark fields are defined at grid points or lattice sites, while gluons act as connecting edges between sites. In the limit $a \to 0$, this approximation converges to QCD in continuous spacetime. Computations are expensive, often taking multiple months to obtain a single data point, and cost increases with smaller grid spacing a. A common approach is to perform calculations for ensembles with varying grid spacing and then extrapolate to $a \to 0$. This technique has successfully reproduced experimental observations, such as determining the proton mass to precisions of less than 2% [37].

Results of lattice QCD calculations are an essential contribution to the measurement of $|V_{cb}|$ in this thesis, particularly results of the MILC [38] and HPQCD groups [39] on $B \to D\ell\nu_\ell$ form factors. Additionally, the FLAG group provides averages of the two groups' results, considering correlations between computations [40].



(a) (b)

Figure 2.7: Feynman diagrams representing $B^- \to D^0 \ell^- \bar{\nu}_\ell$ and $B^0 \to D^- \ell^+ \nu_\ell$ decays.

2.3 Semileptonic $B \to D\ell\nu_\ell$ decays

Semileptonic B decays $B \to X_c \ell \nu_\ell$ are mediated through the electroweak interaction and include $b \to c$ transitions, as seen in the Feynman diagrams shown in Figure 2.7. Therefore, their decay rates are directly proportional to the CKM matrix element $|V_{cb}|$ and can be used to provide direct measurements. An important advantage of semileptonic decays over hadronic decays is that QCD corrections play a reduced role. The outgoing leptonic and hadronic currents factorize, and only the hadronic current is sensitive to QCD effects up to a small electroweak correction term η_{EW} that takes into account higher-order effects. The differential decay rates of the B meson thus take the form

$$d\Gamma \propto G_F^2 |V_{cb}|^2 |L^\mu H_\mu|^2, \tag{2.44}$$

with factorizing leptonic and hadronic currents L^μ and H_μ and the Fermi constant

$$G_F = \frac{\sqrt{2} g_W^2}{8 m_W^2}. \tag{2.45}$$

The leptonic current contains the small electroweak correction η_{EW} and is otherwise independent of QCD effects. The hadronic current is determined from the matrix element

$$\langle X | \bar{c} \gamma_\mu P_L b | B \rangle \,, \tag{2.46}$$

with the left-handed chiral projection operator $P_L = (1 - \gamma_5)/2$, and contains all other QCD dependencies. In $|V_{cb}|$ measurements, the vector meson D^* and the pseudoscalar meson D are considered for X. In these cases, the nonperturbative hadronic matrix element can be parameterized by form factors dependent on the lepton momentum transfer $q^2 = (p_\ell + p_\nu)^2 = (p_B - p_X)^2$. Information about these form factors can be computed numerically from first principles with a lattice QCD approach.
The main topic of this thesis is the study of B decays to the pseudoscalar D meson. For

$B \to D\ell\nu_\ell$ decays, the matrix element (2.46) can be written in terms of the form factors f_+ and f_0 as

$$\langle D|\bar{c}\gamma_\mu P_L b|B\rangle = f_+(q^2)\left(p_B^\mu + p_D^\mu - \frac{m_B^2 - m_D^2}{q^2}q^\mu\right) + f_0(q^2)\frac{m_B^2 - m_D^2}{q^2}q^\mu. \tag{2.47}$$

An alternative kinematic variable to q^2 is the velocity transfer $w = v_B \cdot v_D$, which can be obtained from

$$w = \frac{m_B^2 + m_D^2 - q^2}{2m_B m_D}\ . \tag{2.48}$$

w can also be viewed as a measure of the hadronic recoil, ranging from zero recoil of the hadronic final state $w = 1$ (equivalent to $q^2 \to \infty$) to maximal hadronic recoil $w_{\max} = \frac{m_B^2+m_D^2}{2m_B m_D} \approx 1.59$ ($q^2 = 0$). In terms of the hadronic recoil, the differential decay rate of $B \to D\ell\nu_\ell$ can be obtained from (2.44), the appropriate prefactors, and the hadronic matrix element (2.47) as

$$\frac{d\Gamma}{dw}(B \to D\ell\nu_\ell) = \frac{G_F^2 m_D^3}{48\pi^3}(m_B + m_D)^2(w^2-1)^{3/2}\eta_{\mathrm{EW}}^2|V_{cb}|^2\mathcal{G}(w)^2. \tag{2.49}$$

Here, $\mathcal{G}(w)$ contains the dependency on the form factors $f_+(w)$ and $f_0(w)$. Lattice QCD results provide the form factor at low values of w. Using these theoretical predictions as input and measuring the differential decay rate of $B \to D\ell\nu_\ell$ decays, $|V_{cb}|$ can be directly measured. Due to term $(w^2-1)^{3/2}$ the differential rate rapidly approaches 0 as w approaches 1. This makes the kinematic region where lattice QCD can provide the most accurate predictions hard to probe experimentally. Therefore, it is crucial to measure $\Delta\Gamma/\Delta w$ across the full range of w, which allows performing a fit of the form factor to the distribution.
There are multiple available parameterizations of the form factor $\mathcal{G}(w)$, which are introduced in the following sections.

2.3.1 $B \to D\ell\nu_\ell$ form factor parameterizations

The form factor $\mathcal{G}(w)$ can generally be expressed in terms of f_+ and f_0, with

$$f_0(q^2) = f_+(q^2) + \frac{q^2}{m_B^2 - m_D^2}f_-(q^2), \tag{2.50}$$

from which the identity

$$f_0(q^2 = 0) = f_+(q^2 = 0) \tag{2.51}$$

follows. This is referred to as the *kinematic constraint*. In the limit of vanishing lepton mass, which is an adequate approximation for the modes decaying into e and μ where the lepton mass is negligible in comparison to the involved hadronic mass scales, the form factor depends only on $f_+(w)$ and can be written as

$$\mathcal{G}(w) = \frac{4r}{(1+r)^2}f_+^2(w), \tag{2.52}$$

where $r = m_D/m_B$.

BGL parameterization

The parameterization introduced by Boyd, Grinstein and Lebed (BGL) is a model-independent formulation of form factors relying only on dispersion relations obtained from QCD [1, 2]. In BGL, the conformal transformation

$$z(w,\mathcal{N}) = \frac{\sqrt{w+1}-\sqrt{2\mathcal{N}}}{\sqrt{w+1}+\sqrt{2\mathcal{N}}} \tag{2.53}$$

maps w onto the disk $|z(w,\mathcal{N})| < 1$ in the z complex plane. Here, $\mathcal{N}$ is a free parameter determining the value of w that is mapped onto the origin of the z plane [41]. By choosing $\mathcal{N} = 1$, corresponding to the zero-recoil point $w = 1$ at $z = 0$, z covers a range of $0 \leq z \leq \frac{(1-\sqrt{r})^2}{(1+\sqrt{r})^2} \simeq 0.06$. The form factors are then parameterized by

$$\begin{aligned} f_+(z) &= \frac{1}{P_+(z)\phi_+(z,\mathcal{N})} \sum_{n=0}^{\infty} a_n z^n \\ f_0(z) &= \frac{1}{P_0(z)\phi_0(z,\mathcal{N})} \sum_{n=0}^{\infty} b_n z^n, \end{aligned} \tag{2.54}$$

where $P_{+,0}(z)$ are *Blaschke factors*, containing explicit pole terms (*e.g.* B_c or B_c^* poles) and $\phi_{+,0}$ are *outer functions*, arbitrary functions without additional poles or branch cuts. An important feature of the BGL parameterization is that the sum of squares of its coefficients is constrained. The standard choice of outer functions sets this constraint to the so-called *unitarity bound*

$$\sum_{n=0}^{\infty} a_n^2 < 1, \quad \sum_{n=0}^{\infty} b_n^2 < 1. \tag{2.55}$$

This feature allows truncation of the BGL expansion (2.54) at a given order N, limiting the maximum error introduced through this truncation. In practice, the series is usually truncated at $N = 2$, $N = 3$ or $N = 4$.

BCL parameterization

Because of unknown pole terms, the standard choice of outer functions in the parameterization of BGL distorts the behaviour at large q^2 and introduces a singularity at the $B\pi$ threshold. A more recent choice of outer functions has been proposed by Bourrely, Caprini and Lellouch (BCL) and leads to the parameterization [42, 43]

$$f_+(q^2) = \frac{1}{1-q^2/M_+^2} \sum_{n=0}^{N} a_n^+(t_0) z(q^2,t_0)^n. \tag{2.56}$$

where M_+^2 is the mass of closest B_c resonance. Due to imprecise knowledge of the pole terms, the prefactor can be set to 1, implicitly absorbing these factors into the coefficients of the BCL expansion.

After truncating at order N, one can also impose the asymptotic behavior $\mathfrak{Im} f_+(q^2) \simeq (q^2 - t_+)^{3/2}$ near $t_+ = (m_B + m_D)^2$ to constrain the highest order f_+ coefficient, yielding

$$f_+(q^2) = \frac{1}{1 - q^2/M_+^2} \sum_{k=0}^{N-1} a_k \left[z^k - (-1)^{k-N} \frac{k}{N} z^N \right] \tag{2.57}$$

and

$$f_0\left(q^2\right) = \frac{1}{1 - q^2/M_0^2} \sum_{k=0}^{N-1} b_k z^k. \tag{2.58}$$

CLN parameterization

By introducing additional model-dependent constraints, the parameterization by Caprini, Lellouch and Neubert (CLN) reduces the form factor to two free parameters [44]. Due to its simplicity, it has been widely used in past measurements. The model dependency assumptions have an accuracy within 2%, and as recent measurements become more sensitive, the importance of BGL and BCL parameterization has increased.
The CLN parameterization exploits heavy quark effective theory (HQET) [45] to yield further dispersive constraints, allowing higher order coefficients to be expressed in terms of the slope ρ. The form factor f_+ can then be expressed using only two parameters as

$$f_+(z) \simeq f_+(0) \left(1 - 8\rho^2 z + (51\rho^2 - 10)z^2 - (252\rho^2 - 84)z^3\right). \tag{2.59}$$

To compare with past measurements, it is of interest to perform a fit to the CLN parameterized form factor.

2.4 Current experimental status of $|V_{cb}|$ measurements

As mentioned in the previous section, semileptonic decays have the advantage of factorizing hadronic and leptonic currents, simplifying computations of QCD effects. The cleanest theoretical modes are purely leptonic modes without QCD contributions, however due to their low branching ratios they are difficult to measure experimentally. Semileptonic decays, therefore, offer a compromise between theory and experiment for measuring CKM elements.
There are two distinct methods of measuring CKM elements, *exclusive* and *inclusive* approaches. Exclusive approaches rely on measuring differential decay rates in a specific decay mode and relating the measured rates to CKM elements using theory inputs from lattice QCD. This method was described for measuring $|V_{cb}|$ using $B \to D\ell\nu_\ell$ decays in Section 2.3. The decay modes measured most commonly for determining CKM elements in an exclusive approach are $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ for $|V_{cb}|$, and $B \to \pi\ell\nu_\ell$ for $|V_{ub}|$.
Inclusive measurements determine CKM elements by measuring the generic decay $B \to X_{c/u}\ell\nu_\ell$, where $X_{c/u}$ is a hadronic system that contains a charm or up quark to measure $|V_{cb}|$ or $|V_{ub}|$, respectively. In inclusive measurements, kinematic distributions are related to an operator product expansion (OPE) in terms of the heavy quark mass. In B decays

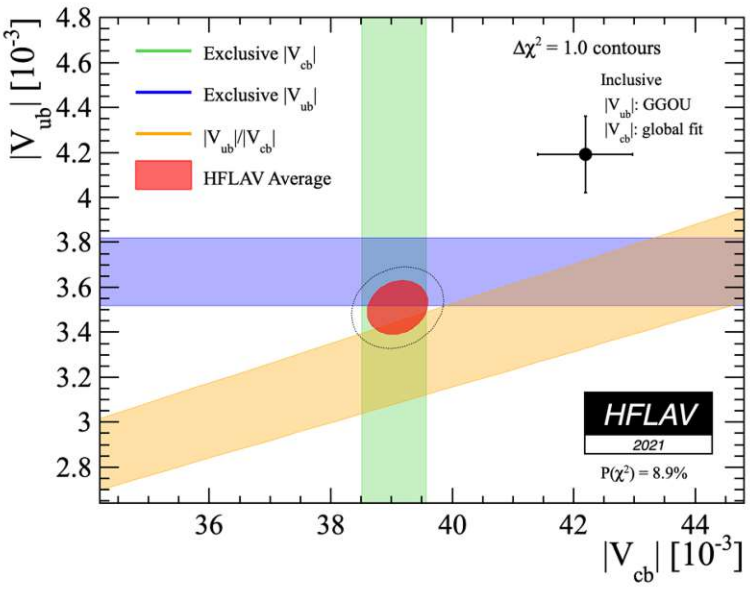


Figure 2.8: Experimental status of the observed discrepancy between exclusive and inclusive measurements of the CKM parameters $|V_{cb}|$ and $|V_{ub}|$ obtained from a global fit by the HFLAV group [3]. The experimental measurements and their global average are shown as colored bands and red ellipses, respectively. The inclusive measurement is shown as a black marker with error bars.

where the quark mass is high compared to the QCD scale, the approximation $m_q \to \infty$ is called the heavy-quark effective theory. These calculations are, in general, orthogonal to the theory inputs in exclusive measurements, and therefore, comparing results of exclusive and inclusive measurements is a viable test of the Standard Model.

A long-standing tension is observed when comparing exclusive and inclusive measurements of both $|V_{cb}|$ and $|V_{ub}|$. When performing a global fit to measurements, such as is done by the HFLAV group [3], the combined tension reaches the level of a combined $\sim 3\sigma$, shown in Figure 2.8. This tension is often called the *inclusive vs. exclusive puzzle.*

While the elements are measured to precisions of $O(2\%)$ in either exclusive or inclusive approaches [28], this observed discrepancy obstructs precise knowledge of the fundamental SM parameters. Improving the precision of measurements to determine the cause of the tension would be a crucial contribution to the understanding of particle physics.

CHAPTER 3

The Belle II experiment at the SuperKEKB collider

This chapter introduces the SuperKEKB particle collider and the Belle II experiment. After a short primer on B factories in Section 3.1, the operating principles of the SuperKEKB accelerator are summarized in Section 3.2. The Belle II detector, with all its subdetectors and the trigger system, is described in Sections 3.3 and 3.4. Finally, the software framework, the simulation of data samples, and the principles of reconstructing decays in collision data are outlined in Sections 3.5 to 3.7.

3.1 B factories

B factories are particle collider experiments operating at the optimal working point to produce large amounts of B mesons. Electrons and positrons collide at a center-of-mass energy of 10.58 GeV, equivalent to the rest mass of the $b\bar{b}$ resonance $\Upsilon(4S)$. Over 96% of the $\Upsilon(4S)$ produced in these e^+e^- collisions decay into either $B^0\overline{B}^0$ or B^+B^- meson pairs [28]. The large amount of B mesons can be analyzed to measure SM parameters precisely. Decays are mediated through the weak force, allowing measurements of, among many other studies, the elements of the CKM matrix.
With respect to higher energy hadron colliders such as the LHC, B factories have advantages in several aspects. The lower multiplicity of charged particles per event allows for cleaner reconstruction of specific decays. With the additional advantage of precisely known initial conditions of the e^+e^- collision, missing energies and momenta from, *e.g.*, invisible neutrinos can be inferred. This is particularly important for analyzing semileptonic decays.
The first generation of B factories, the *BABAR* [46] and Belle [47] experiments, were commissioned and built in the 1990s and began data-taking in 1999. The *BABAR*experiment was operated at the PEP-II accelerator at SLAC in California, USA, and remained in operation until 2008, collecting a data sample corresponding to 467 million B meson pairs.

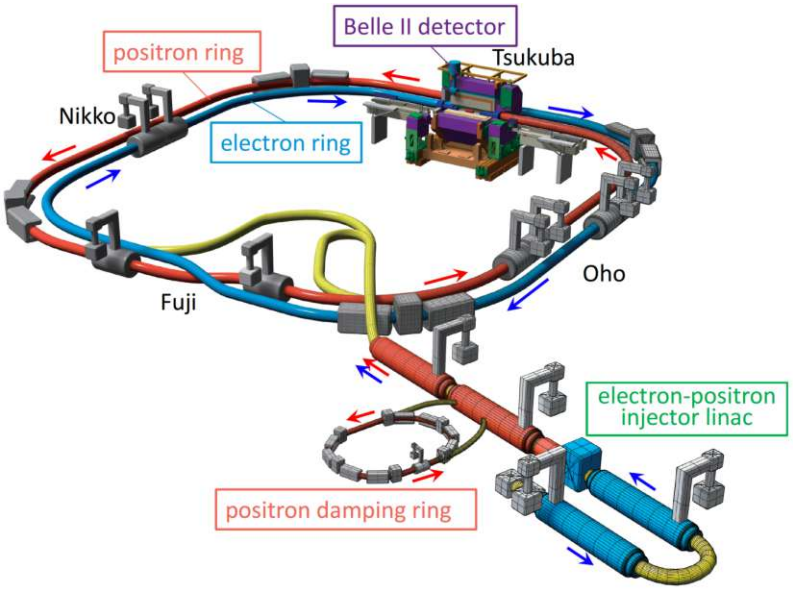


Figure 3.1: Schematic view of the SuperKEKB collider. Electron and positron beams collide at the interaction point located at the Belle II collider [48].

The Belle experiment was located at the KEKB collider at KEK in Tsukuba, Japan, and ran until 2010, recording data of around 770 million B meson pairs.
The Belle II experiment is the successor to the Belle experiment and is considered the only B factory of the second generation. It is operating at the upgraded SuperKEKB collider. Details on operation and improvements with regard to the previous generation will be elaborated on in upcoming sections.

3.2 The SuperKEKB collider

SuperKEKB operates as a double-ring e^+e^- collider, with a low-energy ring (LER) containing 4 GeV positrons and a high-energy ring (HER) with 7 GeV electrons. A schematic view is shown in Figure 3.1. The resulting collisions then take place at a center-of-mass (c.m.) energy[1]

$$\sqrt{s} = \sqrt{\left(E^*_{e^-} + E^*_{e^+}\right)^2 - \left(\vec{p}^*_{e^-} + \vec{p}^*_{e^+}\right)^2} \simeq 10.58\text{GeV/c}^2, \tag{3.1}$$

with the c.m. energies of the electrons and positrons E^*_e and the c.m. 3-momenta $\vec{p}_e$.
Note that the reason for choosing asymmetric beam energies is to obtain the resulting Lorentz boost

$$\beta\gamma = \frac{\vec{p}_{e^-} - \vec{p}_{e^+}}{\sqrt{s}} \simeq 0.28. \tag{3.2}$$

[1]Here and throughout the thesis a raised asterisk in quantites such as E^* and p^* refers to the quantities being given in the c.m. frame.

Process	σ [nb]
$\Upsilon(4S) \to B^+B^-$	0.54
$\Upsilon(4S) \to B^0\bar{B}^0$	0.51
$ee \to c\bar{c}$	1.33
$ee \to d\bar{d}$	0.40
$ee \to s\bar{s}$	0.38
$ee \to u\bar{u}$	1.61
$ee \to e\bar{e}$	300.0
$ee \to \mu\bar{\mu}$	1.15
$ee \to \tau\bar{\tau}$	0.92

Table 3.1: Cross-sections of processes occurring at the Belle II design beam c.m. energy.

The Lorentz-boosted collisions increase the B meson lifetime in the laboratory frame and extend the average flight distance to 130 µm, allowing measurement of the location of B meson decay vertices.

The cross-sections of processes occurring at the collision energy of $\sqrt{s} = 10.58\text{GeV}$ are shown in Table 3.1.

An important performance measure of particle colliders is the instantaneous luminosity. Considering a particle collider based on an interaction with cross-section σ and an event rate dN/dt, the instantaneous luminosity is given as

$$\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt}. \tag{3.3}$$

As a measure of the total collision data taken over a given time frame, the integrated luminosity can be calculated as

$$\mathcal{L}_{\text{int}} = \int \mathcal{L}\, dt. \tag{3.4}$$

To estimate the expected number of events for a decay with cross-section σ in a dataset, one can then multiply by the integrated luminosity $\langle N \rangle = \mathcal{L}_{\text{int}} \times \sigma$. High luminosities are necessary to obtain large samples of decays and reduce uncertainties of measurements.

SuperKEKB is the first accelerator to employ the so-called *nanobeam scheme*. By squeezing the beams to as low as 50 nm along the y-axis and 5 µm along the x-axis, in addition to increasing the crossing angle, the event rate can be increased significantly. This scheme is predicted to allow for an improvement in instantaneous luminosity by a factor of ∼20 with respect to the peak instantaneous luminosity achieved at Belle. An additional doubling of the beam currents yields a total of factor 40 improvement with respect to Belle, corresponding to a target luminosity of $8 \times 10^{35}\text{cm}^{-2}\text{s}^{-1}$. After encountering more beam backgrounds than expected in the first periods of data-taking, the target has been lowered to $6 \times 10^{35}\text{cm}^{-2}\text{s}^{-1}$. In June 2022, SuperKEKB and Belle II reached an instantaneous luminosity of $4.7 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$, setting a world record for particle colliders [49].

The Belle II experiment started data-taking in 2018 and is scheduled to operate until the

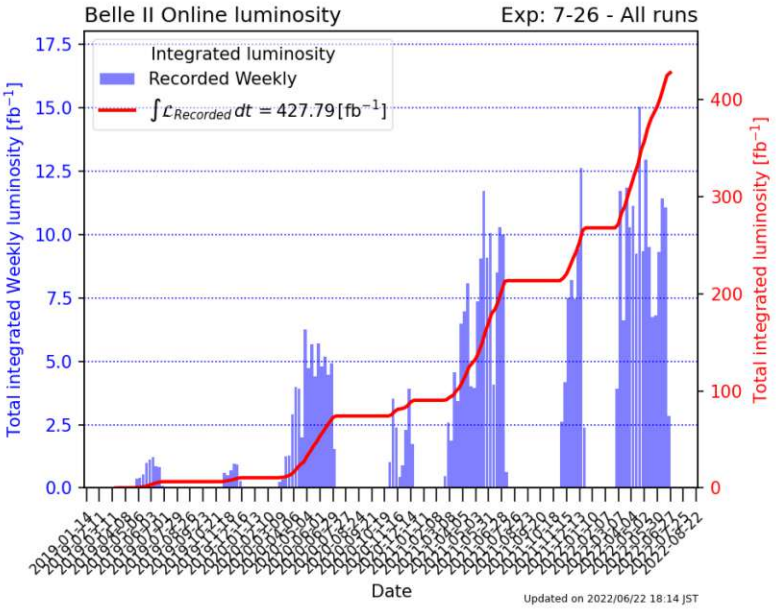


Figure 3.2: Total recorded luminosity at the SuperKEKB collider during run 1 data-taking [50].

2030s. In 2022, the experiment went into the first long shutdown to implement detector upgrades, ending the data-taking period called *run 1*. In run 1, Belle II collected a data sample corresponding to an integrated luminosity of 364 fb^{-1} at the $\Upsilon(4S)$ resonance, approximately equal to half of the total Belle data sample. Additional samples are collected at a c.m. energy of slightly below the $\Upsilon(4S)$ resonance and a scan sample where the beam energy is varied. The combined total luminosity of all samples corresponds to 426 fb^{-1}. Figure 3.2 shows the data-taking progress throughout run 1.

3.3 The Belle II detector

The Belle II detector is a large solid-angle cylindrical detector comprising seven subdetectors arranged in layers radially around the interaction point. The structure is shown in Figure 3.3.
The subdetectors from innermost to outermost are:

- Pixel Detector (PXD): Comprising 50mm thin silicon pixel detectors based on Depleted Field Effect Transistor (DEPFET) technology, the PXD is arranged in two layers around the interaction point to record hits of charged particles.
- Silicon vertex detector (SVD): The SVD consists of four layers of silicon strip detectors. Together with the PXD, they form the Vertex Detector (VXD) covering the first 135mm around the IP.

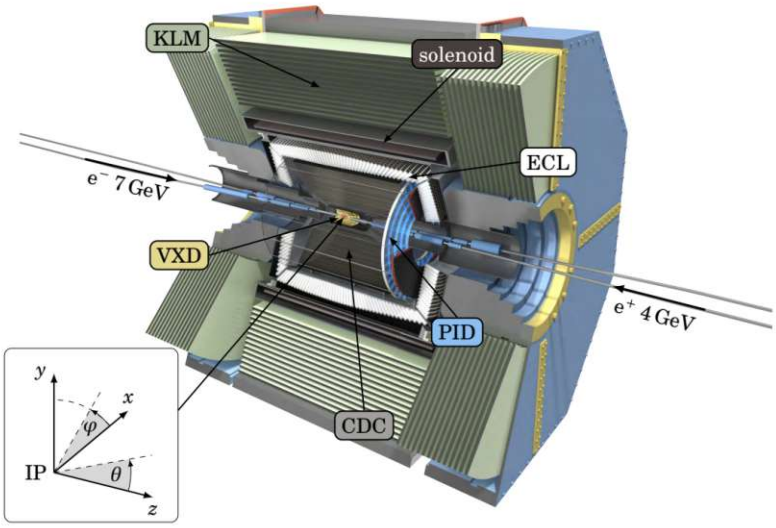


Figure 3.3: Schematic view of the Belle II detector, with its subdetectors indicated with arrows [51, 52]. The coordinate frame is superimposed in the bottom left corner.

- Central drift chamber (CDC): Wires spanned across a gas-filled chamber allow the tracking of charged particle paths through gas ionization and identification of particle type via its energy-loss curve.

- Particle identification (PID): The time-of-propagation counter (TOP) and the Aerogel ring imaging Cherenkov counter (ARICH) are dedicated subdetectors responsible for identifying particles. They offer particularly good distinguishing power between pions and kaons.

- Electromagnetic calorimeter (ECL): CsI crystals outside of the CDC cause incoming particles to produce electromagnetic showers. Particle energies can be determined from the measured energy deposits.

- Superconducting solenoid magnet: Provides a 1.5T magnetic field causing charged tracks to curve, allowing for momentum determination.

- K_L and muon detector (KLM): Alternated layers of iron plates and active material detectors allow for measurement of long-lived kaons and muons that do not get stopped in the ECL.

A schematic overview is shown in Figure 3.3, and a more detailed comparison to the Belle detector is shown in Figure 3.4. The following subsections describe each sub-detector in more detail, following the comprehensive descriptions in the Belle II Technical Design Report [54] and the Belle II Physics Book [52].

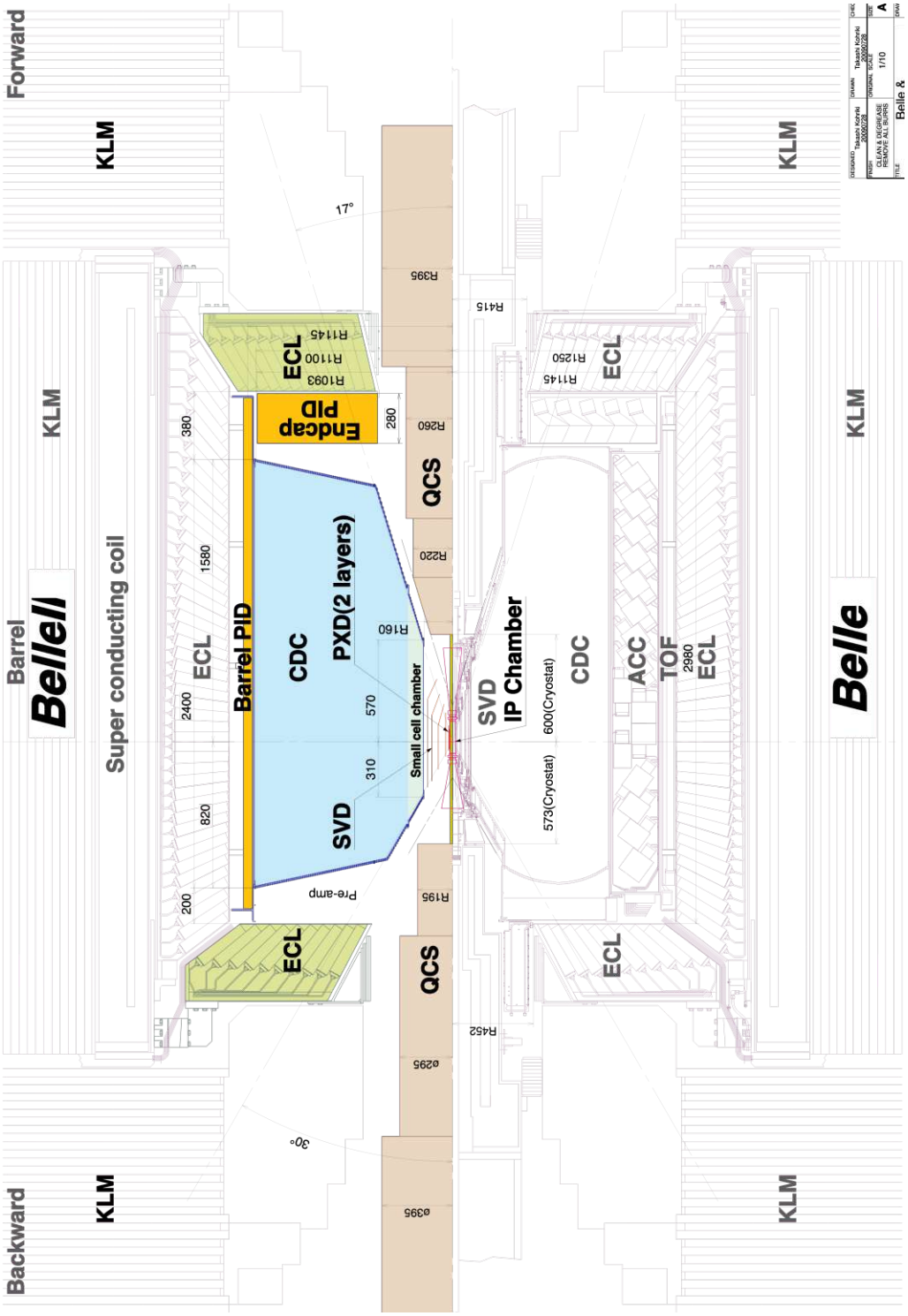
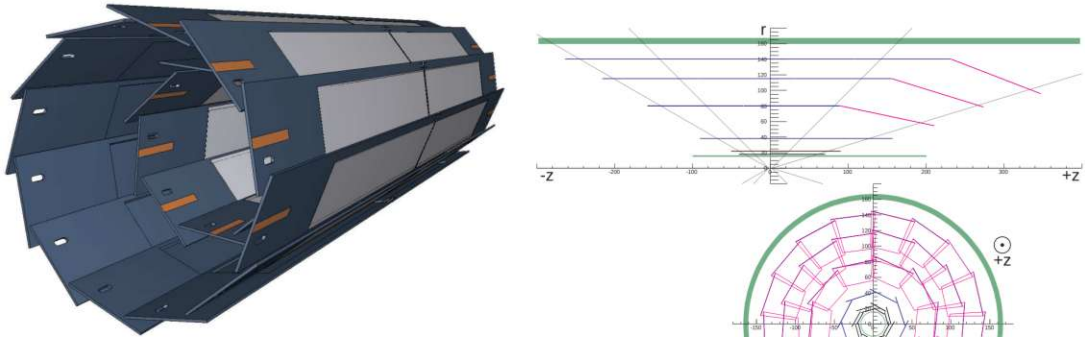


Figure 3.4: Schematic view of the Belle II detector (left) compared to the Belle detector (right) [53].



(a) (b)

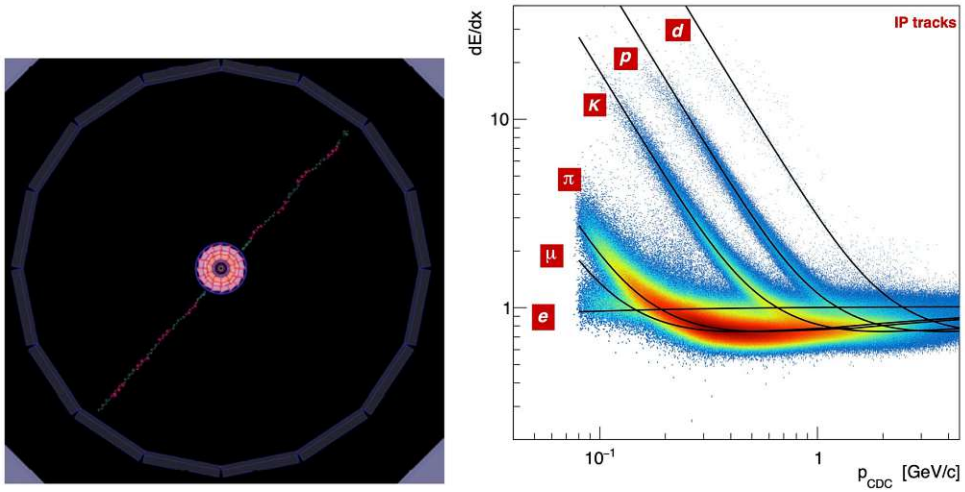
Figure 3.5: (a) Schematic overview of the PXD sensors and their geometric alignment. The DEPFET sensors are shown in light-grey areas [54]. (b) Geometrical alignment of the silicon strips in the SVD. The slanted layout in the forward region reduces the amount of required sensors [54].

3.3.1 Coordinate system

Belle II's right-handed cartesian coordinate system defines its z-axis parallel to the beam, its y-axis vertically upwards, and its x-axis perpendicular to x and y. The origin is set at the IP. Two important angles are the azimuthal angle ϕ, measured counter-clockwise from the x-axis to the projection into the $x-y$ axis of the radial line r, and the polar angle θ measured between the z-axis and the radial line. ϕ and θ range from 0 to 2π and $-\pi$ to π respectively. *Longitudinal* refers to the direction parallel to the beam, while *transverse* indicates the direction perpendicular to it. The axis and angles are shown overlaid in Figure 3.3

3.3.2 Pixel and vertex detectors

The innermost section of the Belle II detector is composed of two inner pixel layers forming the PXD, along with four silicon vertex detector layers constituting the SVD [55]. Together, these two subdetectors comprise the Belle II vertex detector (VXD). The VXD is crucial for reconstructing charged particle trajectories (tracks) and is especially important for locating the decay vertices of B and D mesons. Vertex reconstruction enables time-dependent analyses and helps suppress combinatorial backgrounds.
The PXD consists of two layers of pixelated DEPFET sensors. The high rates of beam backgrounds encountered at the high operating luminosities necessitate thin sensors in the innermost part of the detector, for which sensors utilizing the DEPFET structure are optimal. They combine detection and amplification and are stacked in two layers of 8 and 12 modules, located at radii of 14mm and 22mm around the IP. The dimensions of the sensors are chosen to match the polar angle acceptance of $17° < \theta < 155°$. The readout time for an entire frame is approximately 20 µs. Due to complications in the initial installation of the PXD, the complete second layer was only installed during the long shutdown in 2023.



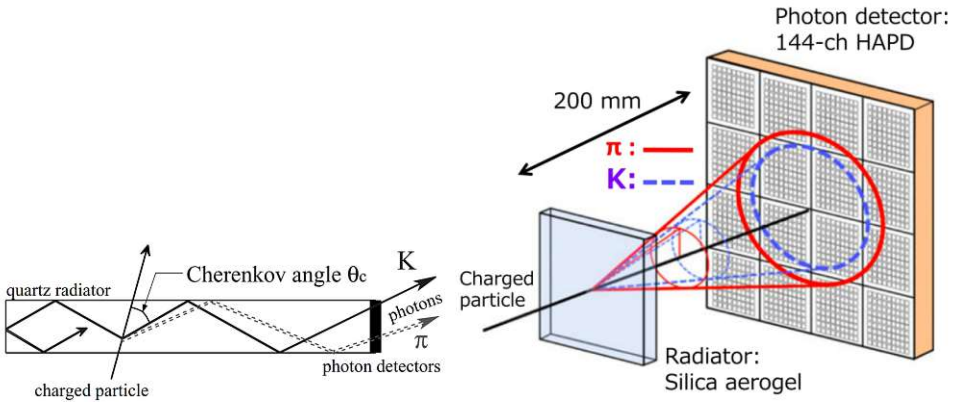
(a) (b)

Figure 3.6: (a) A cosmic muon being tracked in the CDC [52]. (b) Characteristic dE/dx energy loss curves in the CDC [57].

The silicon strips of the SVD and their geometry are designed to satisfy requirements on angular acceptance, hit occupancy, and momentum resolution. The hit occupancy – *i.e.* the average percentage of strips with signal above threshold on any trigger – should be below 10% to allow for association of VXD hits of a track with their hits in the CDC. The momentum resolution should be of the order of tens of MeV, essential for reconstructing slow pions in $D^* \to D\pi$ decays. To fulfill those requirements, four layers are installed at radii of 39, 80, 104, and 135 mm, each consisting of double-sided silicon strip detectors fabricated from six-inch wafers. Each layer consists of 7 to 16 ladders with 2 to 5 sensors each, totalling 35 ladders and 172 sensors. The SVD recorded hit occupancies of 0.3% and efficiencies of 99% in early data collection, showing promising results for the vertex detection system [56].

3.3.3 Central drift chamber

The CDC [58] combines with the VXD to complete the Belle II tracking system. The large chamber is filled with a He–C_2H_6 gas mixture that is ionized by charged particles passing through. Tungsten sense wires are arranged in layers, aligned parallel to the beam axis or skewed by angles of 45.4 and 74 mrad. A uniform electric field is generated by aluminum wires, causing charged particles to induce a signal into the wires when passing through. The particle position can be reconstructed through drift time measurements, and by combining the measurements in each wire the trajectory is measured throughout the CDC. Due to the present magnetic field, the tracks have curvature proportional to their momentum, allowing measurement. In addition to tracking, the CDC also provides information useful for particle identification. Due to gas ionization, charged particles lose energy while traversing the CDC. Different particle types have characteristic energy-loss



(a) (b)

Figure 3.7: (a) Functioning principle of the TOP detector, with photons radiated at the Cherenkov angle θ_c and subsequent propagation through the Quartz bar [54]. (b) Principle of the ARICH detector. The radius of the observed Cherenkov ring depends on the charged particles' mass [59].

curves, crucial for particle identification. The energy loss curves are shown in Figure 3.6. Finally, the CDC contributes to the trigger system by sending a signal based on occupancy. The spatial resolution in the CDC is about 2 mm in z and 100 µm in r, with a relative precision in the energy loss of approximately 12% for tracks at incident angle 90° [54].

3.3.4 Particle identification

The two subdetectors dedicated to particle identification are the ARICH [60] and the TOP [61]. Both detectors are complementary in their angular regions and are used in particular to separate pions and kaons. The functioning principle of both detectors is sketched in Figure 3.7.

The TOP operates in the barrel region and consists of modules made out of two quartz bars, a mirror, and a prism coupling to photomultiplier tubes. The quartz bars act as Cherenkov radiators, causing traversing particles to emit photons. The photons are then trapped through reflection and travel through the bars until the Cherenkov image is reconstructed at the prism. The photon detection time distribution is distinct for each particle type, and therefore, the observed photon distribution provides important information for particle identification.

ARICH is located in the region towards the forward end cap. It is composed of an aerogel Cherenkov radiator that produces Cherenkov photons when charged particles pass through, a photosensor positioned 20 cm from the radiator to allow photons to form Cherenkov rings, and a readout system. The radius of the measured Cherenkov rings allows for separation between kaons and pions.

3.3.5 Electromagnetic calorimeter

The ECL [62] measures particle energy by inducing electromagnetic showers in crystals from incoming leptons, hadrons, and photons. Approximately one third of the decay products from B decays are photons or neutral particles like π^0 that promptly decay into pairs of photons. Since photons are not seen in the tracking detectors, it is crucial to observe them with high efficiency in the electromagnetic calorimeter and precisely determine their energy and momentum. The Belle II ECL is split into three parts: the barrel, the forward and the backward end cap. The detector consists of arrays of thallium-doped caesium iodide CsI(Tl) crystals. The barrel ECL covers the $32° < \theta < 129°$ polar angle range, which extends to $12° < \theta < 155°$ when including the end caps. Between barrel and end caps, as well as at 90°, there are $\sim$1° gaps in the ECL. Upgraded readout electronics mitigate effects from higher background levels occupying crystals at a higher percentage.

3.3.6 Solenoid magnet

The magnetic field in the Belle II detector is provided by a superconducting solenoid filling a cylindrical volume with diameter and length of 3.4 m and 4.4 m respectively. The coil is cooled with a liquid helium cryogenic system, producing a magnetic field of 1.5 T. The magnetic field bends charged particles, allowing momentum measurement through the track curvature. The iron structure (*yoke*), weighing 608 and 524 tonnes for barrel and end caps, respectively, provides a return path for magnetic flux.

3.3.7 K_L and muon detector

The KLM [63] is located in the outermost region of the Belle II detector, outside the ECL and the solenoid. Nonshowering charged hadrons, such as long-lived kaons K_L and muons above $\sim$0.6 GeV, do not deposit their entire energy into the calorimeter but instead travel further. The main objective of the KLM is, therefore, to measure these particles. It is made up of alternating detector elements and iron yokes that serve as the magnetic flux return, as well as allow K_L to shower hadronically. To mitigate the effects of increased beam backgrounds, the resistive plate chamber detector elements used in Belle have been replaced by layers of scintillator strips, which are read out by silicon photomultipliers.
In addition to allowing the measurement of energy and momentum of the long-lived particles, the KLM significantly contributes to particle identification for muons.

3.4 Trigger

The trigger system is responsible for selecting events of interest over events considered as beam backgrounds. The lower cross-sections in ee colliders allow a higher budget of events deemed interesting with respect to pp colliders, making a well-designed trigger system crucial for enabling world-leading precision analyses in fields such as dark searches or τ physics. An important detail is that the Belle II trigger scheme is inclusive for hadronic events with more than three charged tracks, *i.e.* no events have to be discarded.

The Belle II trigger consists of a hardware-based low-level trigger (L1) and a software-based high-level trigger (HLT) [64, 65]. The L1 reduces the output to a rate of 30 kHz, equivalent to the maximum read-in rate of the data acquisition system (DAQ). The reduction is primarily achieved by significantly decreasing the share of $ee \to ee(\gamma)$ events, referred to as *Bhabha scattering*. The cross-section of these events is large compared to other observed events, as seen in Table 3.1. Bhabha scattering is characterized by two high-momentum tracks carrying the majority of the event's available energy, accompanied by a maximum of two clusters in the ECL. By identifying these events and applying a *prescale*, *i.e.* randomly discarding the event with a certain probability, the rate of events passed on to the DAQ can be reduced.
The software-based HLT reduces event rates further to 10 kHz to match the speed of offline storage processing. The HLT performs full, real-time reconstruction using information from all detectors, excluding the PXD, to characterize physics events. The event rate is then reduced by excluding mostly residual beam background events that the L1 trigger has not identified.
After physics events pass the trigger system, the raw information containing track momenta, energy and additional subdetector information is saved as *mdst* data objects and made available by the collaboration for physics analysis.

3.5 The Belle II Analysis Software Framework

The Belle II Analysis Software Framework (`basf2`) [66] is responsible for processing both offline and online data. `mdst` data objects containing particle information are read into the *DataStore*. Specific conditions of the run period, such as the most recent magnetic field map, are saved into the *DBStore* as objects called *payloads*. A collection of multiple payloads is called a *global tag*. Independent analysis modules are chained together in a linear path, each performing a dedicated task required for physics analysis. Each module has access to the DataStore and the DBStore. The `Python` script containing the chain of modules is commonly referred to as the *steering file*, where the specific decay chains and selections chosen by the analysts are specified, and variables of interest are computed from raw data. The resulting subset of events in the input `mdst` with variables of interest saved as columns is then saved as a ROOT data object, called *ntuple*.

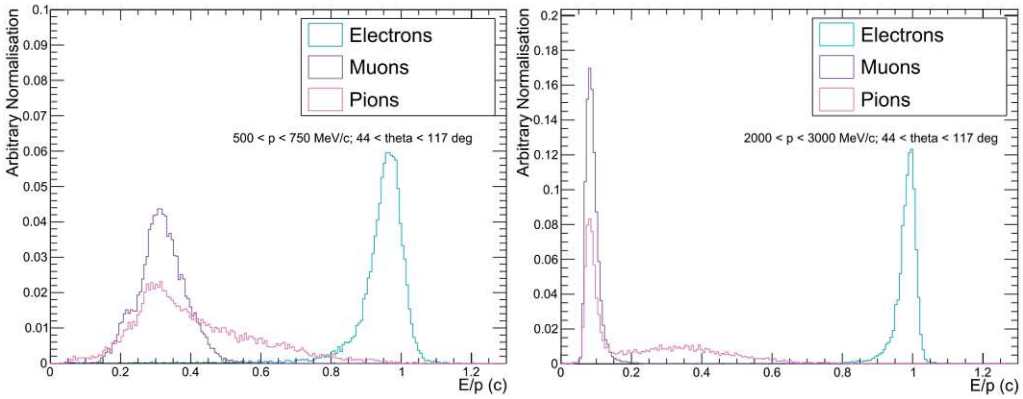
3.6 Monte Carlo simulation

To develop analysis procedures and understand the effects of systematic uncertainties, it is vital to have simulated data available in addition to the collision data collected at the experiment. This is commonly achieved by using Monte Carlo (MC) based event generators, simulating decays of B mesons and other processes occurring in ee collisions; see Table 3.1; by drawing random numbers according to their measured decay rates. The sample of events generated this way is then propagated into detailed detector simulation software, simulating propagation of tracks throughout the detector to reproduce resolution effects, inhomogeneities in the magnetic field, etc.

In Belle II, multiple event generators are used to simulate different types of physics processes. B and D meson decays into exclusive final states are modelled using `EvtGen` [67], allowing implementation of kinematic form factors such as those described in Section 2.3. Implementation of a decay table (`DECAY.DEC`) supplies currently known cross-sections and branching ratios, allowing detailed simulation of cascade decays involving multiple decay vertices. For so-called *continuum* events, constituted of $ee \to q\overline{q}$ events with $q \in [u, d, s, c]$, `Pythia` is used [68]. The QED processes $ee \to ee(\gamma)$ and $ee \to \tau\overline{\tau}$ are simulated in `BABAYAGA.NLO` [69] and a combination of `TAUOLA` [70] and `KKMC` [71] respectively. The event generators account for effects such as the asymmetric beam energies, the location of the interaction point vertex, and the angle between the beams. `BABAYAGA` additionally accounts for smearing in the initial beam energy.
The detector simulation software used is `GEANT4`. `GEANT4` [72] propagates each particle produced in an event subsequently through the individual subdetectors, producing secondary particles where applicable. Hit information is sent to the `basf2` DataStore, allowing the analysis modules to access it.
Belle II MC samples are produced in two categories, *run-independent* and *run-dependent*. Run-independent MC has fixed detector conditions corresponding to the average over a given run period, whereas run-dependent MC accounts for effects such as the beam energy spread and mirror conditions in each run period individually.

3.7 Event reconstruction

To analyze specific decays in the collected data sample, a reconstruction algorithm is used to select events where the decay has potentially occurred. The reconstruction algorithm specified in the steering file combines tracks into composite *candidates*. For instance, the momenta and energies of a kaon and pion with opposite charges are combined to form a composite D candidate from $D^0 \to K^-\pi^+$ decays. Criteria can then be applied to these composite D meson candidates, such as requiring their reconstructed invariant mass (m) to fall within a specified mass range around its expected mass. Such selection criteria are referred to as *rectangular selections* and will be expanded on in Chapter 4. By placing appropriate rectangular selections, the full collected data sample can be scanned, and only events that fulfill every condition and are, therefore, likely to originate from the specified decay are added to the ntuple. Loose selections during reconstruction can be tightened in further offline steps by saving the composite particle variables of interest into the ntuples. Vertex fit algorithms such as the `basf2` implemented *TreeFitter* [73] can be used to analyze the vertex information of each track in the decay chain. Fitting the tracks involved results in a χ^2 distributed probability. Placing a selection on this probability helps suppress combinatorial backgrounds, *i.e.* candidates with tracks from separate processes that fulfill the requirements by chance.
For events with missing energy and momentum, such as semileptonic decays with neutrinos involved, it is useful to construct *rest-of-event* (ROE) quantities. After reconstructing a specific decay, such as $B \to D\ell\nu_\ell$, from one B meson, all additional tracks and clusters not assigned to the signal reconstruction are assigned to the ROE. Variables such as the



(a) (b)

Figure 3.8: The histograms show the E/p distributions in the polar angle range $44° < \theta < 117°$ in the two momentum ranges (a) $500 < p < 750$ MeV and (b) $2000 < p < 3000$ MeV. E/p offers good discriminating power between electrons and pions/muons. [52]

invariant mass of the ROE or the ROE momentum can be used to separate signal from backgrounds or estimate the precise direction of flight of the signal B meson.
Reconstructed candidates in simulated data can be *truth matched* by using ground-truth information from the event generator before detector simulation, attaching flags that describe whether any constituents of the candidate are misreconstructed. In this way, candidates can be separated into signal, *i.e.* correct reconstruction, or backgrounds, where the reconstruction does not match the ground truth. Backgrounds can come, among others, from similar B decays, continuum events, or combinatorial backgrounds.

3.7.1 Charged particle identification

One important aspect of reconstructing physics processes is correctly identifying particle types of final state particles. The information from all sub-detectors enters into a likelihood or a *boosted decision tree* (BDT), introduced further in Chapter 4, to obtain a classifier called particleID (PID). The likelihood or BDT output is scaled from 0 to 1 and can be viewed as pseudo-probability of the particle being of the specific type. For instance, a track with electronID of 1 has maximal indications of being an electron. Classifiers can be further categorized into global particleID, *e.g.* the likelihood of a particle being a kaon over all other particle type hypotheses, or a binary particleID, *e.g.* the likelihood of a particle being a kaon over a pion.
For electronID the main discriminator is the E/p ratio obtained from the ECL and CDC, shown in Figure 3.8. BDT-based electronID additionally considers the shape of ECL showers by considering variables such as the ratio of energy deposits of the inner 9 and outer 25 ECL crystals E_9E_{25}. Muon identification mainly uses KLM information, where muons and long-lived kaons are responsible for most hits. Tracks are extrapolated from

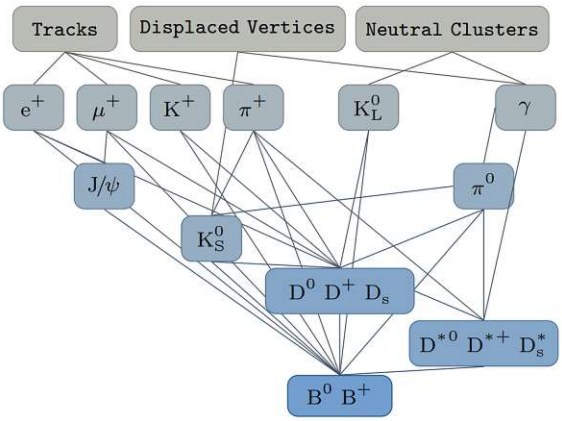


Figure 3.9: Schematic overview of the individual layers that each gather information and feed forward into other layers to reconstruct the tag-side B meson [74].

the inner tracking layers to the KLM, and a muon likelihood is assigned depending on where and how the track ended [52].
Charged hadrons K and π are identified using a combination of CDC, TOP and ARICH information. The Cherenkov-radiation-based ParticleID subdetectors contribute significantly to separating kaons from pions.

3.7.2 Full Event Interpretation

Determining the momentum of the neutrino in semileptonic decays is crucial because of the q^2 dependence of the differential decay rate discussed in Section 2.3. Due to the non-interacting nature of the neutrino, its momentum cannot be directly measured and must be inferred from other information. When reconstructing a semileptonic B decay mode, one approach to constraining the neutrino is to reconstruct the other B meson in the event. Belle II employs an algorithm called full event interpretation (FEI) [74] to reconstruct the decay chain of the so-called *tag-side* B meson B_{tag} in a large number of possible decay modes. With a reconstructed B_{tag}, the additional information can be used to compute the neutrino momentum on the *signal-side* B meson B_{sig}. An analysis employing FEI is therefore referred to as a *tagged* analysis.
The FEI is constructed in layers, starting with final state particles including leptons, pions, kaons and photons, for each of which a boosted decision tree is trained on simulated data. The outputs of the BDTs are fed into the next layer consisting of BDTs representing intermediate composite particles, such as π^0, D or J/ψ mesons, which are subsequently fed into a layer representing B mesons. A schematic view is shown in Figure 3.9. In this way, the FEI employs over 200 BDTs to reconstruct $\mathcal{O}(10\,000)$ distinct hadronic decay chains. A final output classifies the signal probability of the B_{tag}, i.e. a classifier output scaled between 0 and 1 representing the likelihood that the B_{tag} is correctly reconstructed in its assigned decay mode. The FEI can be used either purely hadronically or by including semileptonic decays, which increases the efficiency due to the high semileptonic branching

ratio but reduces the information about B_{tag}. The Belle II hadronic FEI has an efficiency of ∼0.5% for B^+ and ∼0.3% for B^0 at ∼15% purity, which is equivalent to an improvement of ∼50% over Belle [74].
Although the FEI approach has the advantage of increased q^2 resolution, the efficiency of the tag-side reconstruction limits the sample size and therefore increases the statistical uncertainty. The more luminosity is collected, the more advantageous tagged analyses become. In *untagged* analyses that forgo the use of the FEI and rely solely on signal-side reconstruction, such as the $B \to D\ell\nu_\ell$ analysis described in this thesis, alternative approaches must be explored to determine the neutrino momentum precisely. These approaches are described in Chapter 6.

CHAPTER 4

Statistical data analysis in high-energy particle physics

The following chapter emphasizes fundamental principles and techniques from statistics and data analysis that are essential for achieving precise measurements. Section 4.1 motivates performance metrics such as a figure of merit obtained from assuming Poisson distributed data and introduces multivariate methods of optimizing such a figure of merit. Section 4.2 highlights the binned maximum likelihood fit method used throughout this thesis to extract signal yields from collision data. The treatment of systematic uncertainties in a binned fit framework is covered in Section 4.3. Finally, the concept of a blind analysis is explained in Section 4.4.

4.1 Multivariate analysis

Analyzing high-energy data sets obtained in experiments and measuring parameters is an inherently statistical process – the finite amount of data available associates uncertainties to any measurement. The goal of any analysis is to minimize the overall uncertainty of the measured parameter of interest, and this is done by employing principles and techniques from statistical data analysis. A primary use is binary classification, *i.e.* separating signal from background candidates. Concepts from multivariate analysis used to do binary classification are introduced in the following subsections.

4.1.1 Figures of merit

Before discussing methods to maximize statistical sensitivity, it is instructive first to introduce the metrics that characterize the level of sensitivity. In high-energy particle physics, one commonly used metric is the *figure of merit* [28]. The distribution underlying a counting experiment with independent events is the Poisson distribution, expressing the

probability of an event occurring k times in a fixed time interval, given a constant mean rate. Its probability mass function

$$f(k;\mu) = \Pr(\mathrm{X} = \mathrm{k}) = \frac{\mu^k e^{-\mu}}{k!}, \tag{4.1}$$

provides the probability that a random variable X occurs k times in a fixed time interval, where the rate μ represents the average number of occurrences per that interval. The Poisson distribution is described by the single parameter μ, characterizing its expected value and variance

$$\mathrm{E}(X) = \mathrm{Var}(X) = \mu. \tag{4.2}$$

To estimate the rate governing the random process from experimental data, such as a decay rate from a particle collider experiment, one can define a likelihood function $L(\boldsymbol{\theta}) = P(\boldsymbol{x}|\boldsymbol{\theta})$, where one or more parameters $\boldsymbol{\theta}$ characterize the probability distribution of the data. In general, the likelihood function is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i;\boldsymbol{\theta}), \tag{4.3}$$

where $f(x_i;\boldsymbol{\theta})$ is the probability mass function of the distribution for the i-th observation x_i. The parameters can then be estimated using the *maximum likelihood* (ML) estimator, which identifies the values of $\boldsymbol{\theta}$ that maximize the likelihood, or equivalently and often simpler the logarithm of the likelihood, by solving

$$\frac{\partial \ln L}{\partial \theta_i} = 0, \quad i = 1, \ldots, N. \tag{4.4}$$

As the sample size approaches infinity, the maximum likelihood estimator is unbiased, *i.e.* its expectation is equal to the true value and efficient, *i.e.* the variance of the estimator is equal to its theoretical minimum [75]. For the special case of the Poisson distribution, the likelihood is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \frac{\mu^{x_i}}{x_i!} e^{-\mu}, \tag{4.5}$$

where x_i constitutes a set of Poisson-distributed data. Consider a counting experiment with an observed number of events n, distributed as $\mu s + b$, where s and b correspond to the number of expected signal and background events, and the signal strength μ is the parameter to be measured[1]. The likelihood function for this scenario becomes

$$L(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)}, \tag{4.6}$$

and its maximum likelihood estimator is $\hat{\mu} = (n - b)/s$. Under the assumption that the observed signal is equal to the expectation, $\mu = 1$, the variance of the MLE is then

$$V[\hat{\mu}] = V\left[\frac{n-b}{s}\right] = \frac{1}{s^2} V[n] = \frac{s+b}{s^2}. \tag{4.7}$$

[1] In practice, the expected values of signal and background are obtained from MC samples. Measuring $\mu = 1$ is then equivalent to observing exactly the amount of signal events predicted in the simulated data samples.

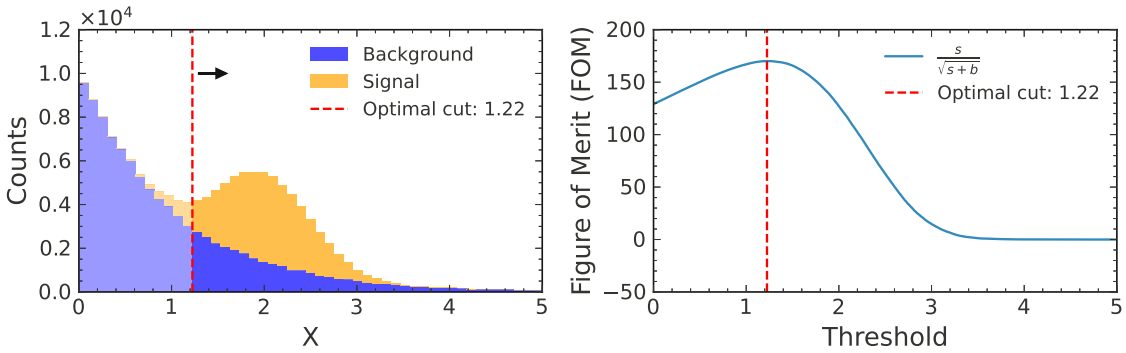


Figure 4.1: Stacked example histograms of an exponentially decaying background component and a Gaussian distributed signal component over a feature X. For each potential rectangular cut threshold $X > \lambda$, the right-hand side plot shows the FOM of the resulting sample. The optimal cut is shown as a red dashed line.

The most precise measurement of the parameter of interest is then obtained by minimizing the variance, or as commonly used, maximizing the figure of merit

$$\text{FOM} = \sqrt{\frac{1}{V[\hat{\mu}]}} = \frac{s}{\sqrt{s+b}}. \tag{4.8}$$

To optimize an analysis for minimal statistical uncertainty, one must select the subset of events in data with the highest figure of merit.
Related concepts describing the distribution of signal and background event in a given subset are the efficiency and the purity:

- Efficiency:
$$\epsilon = \frac{\#\text{ of signal events in the data subset}}{\#\text{ of signal events in the total data set}} \tag{4.9}$$
i.e., the ratio of true positives (signal events correctly identified) to the total number of true positives (all signal events in the dataset).

- Purity:
$$\rho = \frac{\#\text{ of signal events in the data set}}{\#\text{ of events in the data set}} \tag{4.10}$$
i.e. the ratio of true positives over true positives plus true negatives.

Equations (4.8) to (4.10) are used throughout the analysis described in the thesis to define selections that result in minimal statistical uncertainties.

4.1.2 Multivariate selections

For multivariate data samples, such as the ones obtained from reconstructing decays in collider data, there are multiple ways to improve the figure of merit. The simplest way

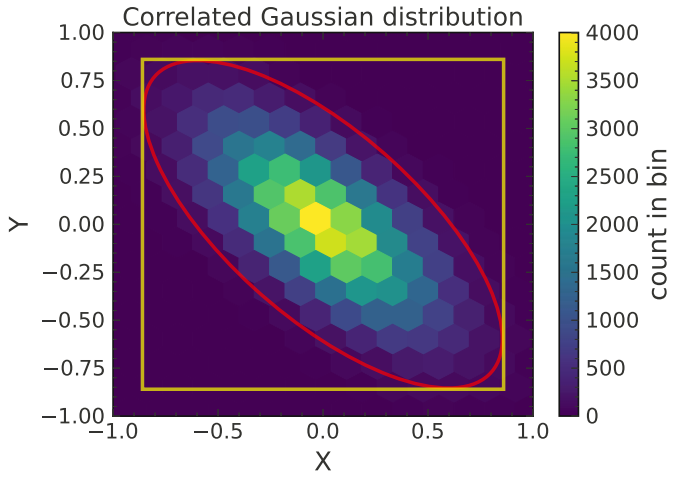


Figure 4.2: The plot shows a heatmap of an example signal distribution in two correlated variables X and Y. If overlaid on a uniform background distribution, the multivariate selection in the red ellipsis leads to a higher FOM than the optimized rectangular cuts in yellow.

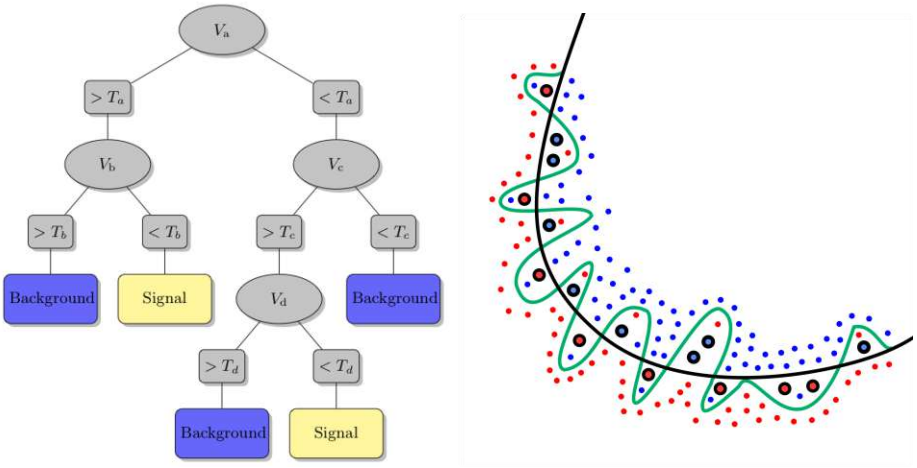
is to perform a *rectangular selection* or *rectangular cut*, *i.e.* require every candidate to pass a requirement $X > \lambda$, where X is a feature of the multivariate data sample, and λ is a chosen cut-off value that ideally maximizes the figure of merit $\frac{s}{\sqrt{s+b}}$. An example of optimizing such a selection is shown in Figure 4.1. The process of consecutively optimizing multiple rectangular cuts is generally not commutative, *i.e.* the order of optimization can affect the obtained set of selections, and the FOM will, in almost all cases, not be the true optimum. For example, background candidates already removed in the first selection could potentially be removed more efficiently in a subsequent selection. One method of obtaining an optimal set of rectangular selections is by using an algorithm such as simulated annealing to optimize selections simultaneously.
Simulated annealing [76] is an iterative probabilistic technique for approximating the global minimum of a given function $f(x)$. The algorithm starts with an initial solution x_0 and an initial temperature T_0. Each iteration generates a new candidate solution x_{new} in the neighbourhood of the current solution x_k. The objective function value

$$\Delta E = f(x_{\text{new}}) - f(x_k) \tag{4.11}$$

decides the acceptance of the new solution. The new solution is always accepted in the case of $\Delta E < 0$. To avoid getting trapped in a local minimum of the function contour, solutions with positive ΔE are accepted according to a probability

$$P(\Delta E) = \exp\left(-\frac{\Delta E}{T_k}\right), \tag{4.12}$$

dependent on the current temperature T_k. In each iteration step, the temperature is reduced by a factor α, set in the algorithm's initialization.



(a) (b)

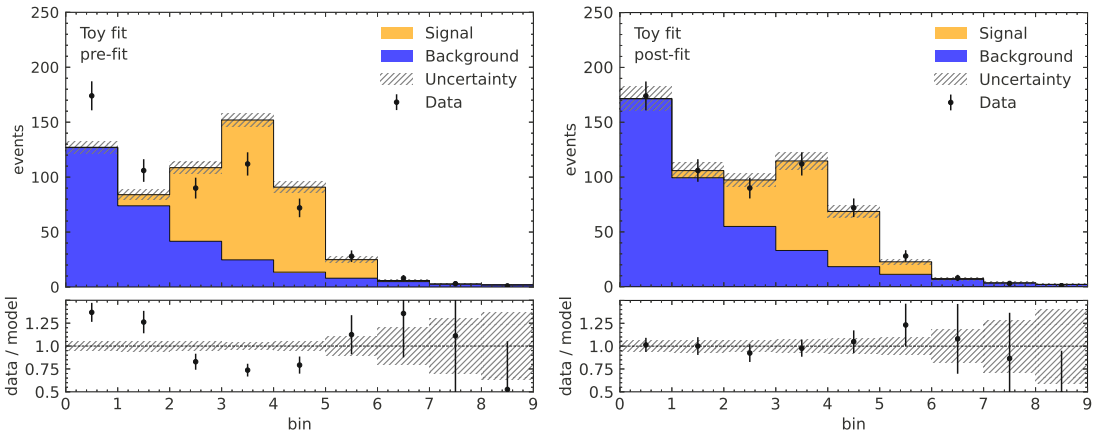
Figure 4.3: (a) Schematic view of a decision tree. Each node represents a test in a feature V_i, with thresholds T_i. After traversing each test, every instance in the training set is classified into signal or background [78]. (b) The green line schematically represents an overfitted model distinguishing blue and red categories. Individual decision trees are prone to overfitting; boosting can mitigate the effect by combining multiple decision trees into a single classifier output [79].

In the case of applying simulated annealing to optimizing rectangular selections, the initial solution corresponds to a set of selections $\alpha_i < X_i < \beta_i$ for each feature $X_i, (i = 1, \ldots, N)$. The function $f(x)$ corresponds to the figure of merit of a subset of the data sample passing the current set of selections. The following solution in the iteration is generated by randomly altering one α_i or β_i in the selections. The Python library `selanneal` [77] implements the feature optimization in Python.

When correlations between features are present, it is advantageous to define selections dependent on multiple features rather than simple rectangular selections. This concept is shown schematically in Figure 4.2. Multivariate selections include requirements such as $X < \lambda Y$, where X and Y are different features of the sample or more advanced concepts based on machine learning. A standard machine learning tool for multivariate classification used in high-energy particle physics is the boosted decision tree.

Boosted decision trees

A boosted decision tree (BDT) is a supervised learning method combining weak classifiers into a single strong classifier. It is effective when the underlying relationships between input features are complex and non-linear, making it suitable for high-energy physics analysis. A decision tree is a flowchart-like structure where each node represents a test on a feature, and each branch represents its outcome. A schematic illustration is shown in Figure 4.3. The decision tree is built to classify each candidate into a category by training



(a) (b)

Figure 4.4: A fit to a toy data sample with a signal and background template. The histograms are shown before the fit on the left and after the fit on the right. Data is shown as black dots with error bars, and the expected templates from simulation as colored, stacked histograms. The overall normalization of the templates is left unconstrained, while shape variations are allowed in the templates according to their Poisson statistics. In this example, the signal yield in the toy data is significantly overestimated in simulation with respect to observed data.

on labelled input data. To mitigate the risks of overfitting and reduce the variance of the decision tree outcome, *boosting* can be used. Boosting involves combining multiple decision trees sequentially. The classifier performance can be increased by focusing on instances misclassified in the previous iteration of decision trees. The most common algorithm to train a BDT is called *AdaBoost* (Adaptive Boosting) [80]. Each instance in the training set is initialized with an equal weight. Individual decision trees, called weak classifiers, are trained on the input data. In the next iteration, often misidentified instances are assigned higher weights to increase the focus on them. After multiple iterations, the final classifier output is obtained by taking a weighted sum of the weak classifiers where the weak classifier accuracy determines the weight. The classifier output is generally scaled to a range of $(0, 1)$, representing a signal likelihood. By then performing a selection on the classifier output, all input features are used simultaneously to reject backgrounds.

4.2 Binned maximum likelihood fits

After adequately optimizing the data sample, the resulting signal yield has to be extracted. This is often done by performing a maximum likelihood fit to the data in a given feature X. The fit can be done in two ways: by fitting parameters that describe a probability density function (pdf) to the experimental data without grouping the data into discrete intervals (*bins*) or by fitting binned histograms of different categories obtained from a simulation (*templates*). When the shape of the signal and background distributions is

well known, *e.g.* a Gaussian signal peak, an unbinned fit has the advantage of being less reliant on simulation. In the case of more complex distributions, binned template fits are advantageous as long as discrepancies between data and simulation are well understood. In the context of this thesis, binned template fits are employed and are therefore introduced in more detail.
Binned template fits involve dividing the data into bins and comparing the number of events observed to the expected events. The expected events are obtained from simulation and are generally split into multiple categories. The simple case of two templates, signal and background, fitted to data is shown in Figure 4.4. In practice, different types of physics processes are split up to constitute the background. In the fit, the shape and normalization of the templates can vary, either with free parameters η, which are often used to model the overall normalization (*i.e.* signal yield), or constrained parameters χ, which are used to model the effect of systematic uncertainties on the prediction.

In a binned maximum likelihood fit, the probability of model parameters ϕ given the data x is represented by the statistical model $f(x|\phi)$. The likelihood of a given set of parameters ϕ describing the data is then denoted by $\mathcal{L}(\phi)$. This document outlines the specification of f and $\mathcal{L}$ within the `HistFactory` framework used by `ROOT` [81].
The parameters describing the model can be split into free and constrained parameters η and χ, while the parameters describing the data can be divided into *actual data* a, referring to the per-bin yields of the data histogram, and the *auxiliary data*, which inform the constraints on the constrained model parameters. In practice, the auxiliary parameters are *e.g.* correction factors or shape uncertainties obtained from orthogonal calibration samples. Conventionally, it is also useful to split the model parameters into *parameters of interest*, *i.e.* signal yields, and label the remaining parameters as *nuisance parameters*. The symbols and notations are summarised in Table 4.1.

The general form of the statistical model $f(x|\phi)$ can then be written as

$$f(\boldsymbol{n}, \boldsymbol{a} \mid \boldsymbol{\eta}, \boldsymbol{\chi}) = \prod_{b \in \text{bins}} \text{Pois}\left(n_b \mid \nu_b(\boldsymbol{\eta}, \boldsymbol{\chi})\right) \quad \prod_{\chi \in \boldsymbol{\chi}} c_\chi\left(a_\chi \mid \chi\right) \tag{4.13}$$

with the event rates $\nu_b(\boldsymbol{\eta}, \boldsymbol{\chi})$, and constraint terms $c_\chi\left(a_\chi \mid \chi\right)$. The first part parametrizes the probability of the model with parameters η and χ describing the Poisson distributed bin counts of the observed data, while the second part describes the compatibility of the constraint terms with the auxiliary data. The event rates are defined as

$$\nu_b(\boldsymbol{\phi}) = \sum_{s \in \text{templates}} \nu_{sb}(\boldsymbol{\eta}, \boldsymbol{\chi}) = \sum_{s \in \text{templates}} \left(\prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{sb}(\boldsymbol{\eta}, \boldsymbol{\chi})\right) \left(\nu_{sb}^0 + \sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{sb}(\boldsymbol{\eta}, \boldsymbol{\chi})\right) \tag{4.14}$$

with a nominal rate ν_{sb}^0 and rate modifying terms κ_{sb} and Δ_{sb}. The result of the maximum likelihood fit is the set of parameters $\boldsymbol{\eta}$ and $\boldsymbol{\chi}$ that minimizes Equation (4.13).
Throughout the thesis, the Python framework `pyhf` [82] is employed to perform binned likelihood fits, and its conventions for rate modifying terms are adapted. The two modifiers most commonly used in the analyses described in this thesis are:

Symbol	Name
$f(\boldsymbol{x} \mid \boldsymbol{\phi})$	Model
$\boldsymbol{n}$	Channel data (or event counts)
$\boldsymbol{a}$	Auxiliary data
ν^0	Nominal event rates
$\nu(\boldsymbol{\phi})$	Calculated event rates
$\boldsymbol{\phi} = \{\boldsymbol{\eta}, \boldsymbol{\chi}\}$	All parameters
$\boldsymbol{\eta}$	Free parameters
$\boldsymbol{\chi}$	Constrained parameters
$\boldsymbol{\kappa}(\boldsymbol{\phi})$	Multiplicative rate modifier
$\boldsymbol{\Delta}(\boldsymbol{\phi})$	Additive rate modifier
$c_\chi\,(a_\chi \mid \chi)$	Constraint term for constrained parameter χ

Table 4.1: Symbol notation in the `pyhf` notation [82].

- *normsys* modifier: The *normsys* modifier adjusts the overall normalization of a template. It is governed by auxiliary data, which consists of a lower and an upper multiplicative factor. A single nuisance parameter α, which is subject to a Gaussian constraint, interpolates the scale of the histogram. For instance, a $\alpha = 1$ corresponds to scaling the nominal histogram by the upper multiplicative factor, while $\alpha = -1$ corresponds to the lower multiplicative factor. For other values of, the scaling is linearly interpolated between these two factors. The Gaussian penalty term in the fit function discourages the nuisance parameter α from deviating significantly from 0.

- *shapesys* modifier: The *shapesys* modifier adjusts the shape of a histogram based on a single nuisance parameter, which is also subject to a Gaussian constraint. This parameter linearly interpolates between a nominal histogram and two variation histograms provided as auxiliary data – one representing a lower variation and the other an upper variation. If the nuisance parameter is fit to, for example, 0.5, the resulting histogram will be interpolated halfway between the nominal histogram and the upper variation histogram for each bin. This approach allows for shape variations that are fully correlated between bins.

All available modifiers and their constraint terms are summarised in Table 4.2.

Description	Modification	Constraint Term c_χ	Input
Uncorrelated Shape (shapesys)	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Pois}\left(r_b = \sigma_b^{-2} \mid \rho_b = \sigma_b^{-2}\gamma_b\right)$	σ_b
Correlated Shape (histosys)	$\Delta_{scb}(\alpha) = f_p\left(\alpha \mid \Delta_{scb,\alpha=-1}, \Delta_{scb,\alpha=1}\right)$	$\text{Gaus}(a = 0 \mid \alpha, \sigma = 1)$	$\Delta_{scb,\alpha=\pm1}$
Normalisation Unc. (normsys)	$\kappa_{scb}(\alpha) = g_p\left(\alpha \mid \kappa_{scb,\alpha=-1}, \kappa_{scb,\alpha=1}\right)$	$\text{Gaus}\left(a = 0 \mid \alpha, \sigma = 1\right)$	$\kappa_{scb,\alpha=\pm1}$
MC Stat. Uncertainty (staterror)	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Gaus}\left(a_{\gamma_b} = 1 \mid \gamma_b, \delta_b\right)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
Luminosity (lumi)	$\kappa_{scb}(\lambda) = \lambda$	$\text{Gaus}\left(l = \lambda_0 \mid \lambda, \sigma_\lambda\right)$	$\lambda_0, \sigma_\lambda$
Normalisation (normfactor)	$\kappa_{scb}(\mu_b) = \mu_b$		

Table 4.2: Relevant modifiers and constraints in the `pyhf` formulation [82]. The additional index c compared to the likelihood specification Equation (4.13) indicates a *channel*. Multiple channels, *e.g.*, decay modes or multiple dimensions, can be fit simultaneously.

4.3 Systematic uncertainties

In high-energy physics analysis, systematic uncertainties from several sources can affect the shape and normalization of templates associated with the fit. Using the `pyhf` statistical model and its modifiers defined in Table 4.2, the uncertainties can be directly included and modelled as nuisance parameters in the likelihood fit. The constraint terms can then allow flexibility in the shape to account for effects such as a difference in performance of particle identification in data and MC. Such effects then enter the likelihood fit and are mapped directly into the resulting uncertainties on the fitted signal strengths and the correlation matrix between parameters of interest.

4.3.1 Toy resampling

One approach to evaluating systematic uncertainties is by using a so-called *toy approach*. Toy approaches employ a resampling strategy to generate alternative input data sets, allowing studies of the impact of specific sources of uncertainties. One resampling approach is *bootstrapping* [83]. Bootstrapping is a statistical method of performing a test by resampling with replacement. From a set of N observed data points, a toy set is produced by drawing a random data point X N times, allowing for duplication of the original data points.
An alternative way of generating toy samples is by drawing Gaussian-distributed random numbers. Often, a particular correction is provided in the form of binned correction factors, with bin-wise statistical and systematic uncertainties. To estimate the impact of this correction, a toy set of corrections can be produced by generating new correction factors

$$w_{k,\text{ toy}} = w_{k,\text{ nominal}} + \mathcal{G}\left(0, \sigma^T_{\text{stat.}}\ \rho_{\text{stat.}}\sigma_{\text{stat.}}\right) + \mathcal{G}\left(0, \sigma^T_{\text{sys.}}\ \rho_{\text{sys.}}\sigma_{\text{sys.}}\right). \tag{4.15}$$

Here, $w_{k,\text{ nominal}}$ represents the nominal values of the correction factors and $\mathcal{G}\left(0, \sigma^T\rho\sigma\right)$ is a randomly drawn number out of a multivariate Gaussian distribution with means 0 and covariance $\sigma^T\rho\sigma$. Uncorrelated statistic correlations correspond to a diagonal form of $\rho_{\text{stat.}}$. For systematic uncertainties, the exact nature of the correlation matrix is often unknown, and different assumptions for correlations should be studied.

By generating a large number of toy sets, the distribution of a particular parameter of interest across all toy sets provides a test for potential biases and efficiencies of estimators. For instance, it is instructive to introduce the *pull*

$$g = \frac{x - \mu}{\sigma} \tag{4.16}$$

with a given parameter of interest x and corresponding mean and standard deviation μ and σ. Since the mean and standard deviation of the distribution over g should be 0 and 1, respectively, for the estimator to be unbiased and efficient, computing the pull serves as an effective test for bias.

4.3.2 Eigenvalue decomposition

After generating toy sets of correction factors for a given systematic uncertainty, the effect on the shape and the normalization of the fit templates should be included in the statistical

model. By calculating the bin counts S for each set of toy correction factors $n \in N$, the correlation matrix can be obtained as

$$\rho_{ij} = \frac{\langle (S_i - \langle S_i \rangle)\,(S_j - \langle S_j \rangle) \rangle}{\sqrt{\left\langle (S_i - \langle S_i \rangle)^2 \right\rangle}\sqrt{\left\langle (S_j - \langle S_j \rangle)^2 \right\rangle}}. \tag{4.17}$$

Here, angular brackets denote averaging over all N set of toy factors, and i and j sum over each bin of every fit template, setting the matrix dimension to $(n_{bin} \times n_{templates}, n_{bin} \times n_{templates})$. The covariance matrix is then obtained as $C_{ij} = \rho_{ij}\sigma_i\sigma_j$ with the bin-wise standard deviations σ.
To describe a covariance matrix in terms of independent shape variations, one can use the eigenvalue decomposition

$$C^{sys} = VUV^T = (V\sqrt{U})(V\sqrt{U})^T \tag{4.18}$$

with eigenvectors V and eigenvalues U in addition to the fact that a covariance matrix can be decomposed into individual systematic shape variations Γ

$$C^{sys}_{ik} = \sum \Gamma_{il}\Gamma_{lk} \tag{4.19}$$

to identify individual systematic shape variations as

$$\Gamma_i = v_i\sqrt{\lambda_i} \tag{4.20}$$

with the i-th eigenvalue and eigenvector λ_i and v_i. If the eigenvalues fall off quickly, the often high-dimensional covariance matrix can be effectively modelled by including only the leading eigenvectors.

4.4 Blind analysis

In high-energy particle physics, a *blind analysis* [84] refers to a specific measurement where the result is blinded until the last step of setting up the measurement. This is done to avoid or reduce bias induced by the experimenter to unintentionally steer the analysis choices in a way of obtaining a result in a particular direction. This includes changing cut values of selections, binning choices in a maximum likelihood fit or choosing a particular set of boundaries for the signal window.
To avoid this, the analysis workflow is calibrated and performed as far as possible by only looking at simulated data. When input from collider data is necessary, *e.g.* for calibration of Data/MC discrepancies, only data in non-sensitive regions or from related physics channels, so-called *control channels*, is used to perform calibrations. After the analysis is sufficiently optimized, the unblinding follows in several steps. In the blind analysis described in Chapter 6, the first step of the unblinding procedure is to look at Data/MC agreements in the appropriate signal windows. If the agreement is appropriate, meaning that the corrections calibrated on control samples map well on the signal reconstruction, the next step is to perform the fit with a blinded result of the signal yields. By then looking

at the fitted values of the nuisance parameters, it can be checked if a particular source of uncertainty is modelled significantly differently in data compared to MC. If everything is under control, the next step is to unblind the fitted result with a limited data sample, followed by unblinding the fitted result with the full data sample.

CHAPTER 5

Calibrating electron identification using Bhabha scattering events

This chapter summarizes a study of the electron identification algorithm using $ee \to ee(\gamma)$ events. By comparing the identification efficiency in data and MC, correction factors are obtained and provided to analysts. Section 5.1 introduces the physical properties of radiative electron scattering events and motivates their use for performance studies. The study's methodology is described in Section 5.2. Sections 5.3 to 5.5 cover details of the analysis choices, including the data samples used, the selection of candidates and the binning of the final correction factors are provided in. The treatment of systematic uncertainties is described in Section 5.6. Finally, the resulting correction factors and uncertainties are presented in Section 5.7.
I performed the study described in this section in collaboration with Paul Feichtinger.

5.1 Bhabha scattering

Electron-positron scattering $ee \to ee$ or $ee \to ee\gamma$, also termed Bhabha and radiative Bhabha scattering, respectively, is a major source of backgrounds in B factories. Due to its high cross-section, approximately ∼300 times larger than $b\bar{b}$ production at the Belle II

Figure 5.1: Feynman diagrams for $e^+e^- \to e^+e^-$ in the s (left) and t (right) channels.

design beam c.m. energy, vetoes are applied to reduce the load faced by the data acquisition system, as described in Section 3.4. Due to its clean signature of exactly two final state tracks – mostly back-to-back in the c.m. frame (with some deviations due to radiated photons) – vetoes with high efficiency can be applied. While Bhabha events are primarily seen as backgrounds and, therefore, vetoed in most physics analyses, the events can also be used for calibration studies. In most e^+e^- collision experiments, including *BABAR*, Belle, and Belle II, Bhabha events are used to estimate the integrated luminosity collected over a certain run period, *e.g.* Ref. [85].
Furthermore, since the events are easily identifiable with high certainty that the two tracks observed are electrons, Bhabha events are also an excellent environment to study and compare the performance of electron identification in simulated data against collision data. Electrons from Bhabha scattering cover a wide range of laboratory frame momentum, from as low as 0.2 GeV to 7 GeV, as well as the full laboratory frame polar angle coverage of the detector. This makes corrections derived from Bhabha events applicable to most analyses involving electrons. Kinematic distributions in laboratory momentum and polar angle are shown in Figure 5.2 for electrons and positrons. The asymmetric beams in the collision cause asymmetry in the distributions.
As introduced in Section 3.7.1, particle identification variables are computed using different algorithms in the Belle II software framework. Throughout the thesis, the notation of particleID variables is chosen as $\mathcal{P}^a_{p,d}$, where a refers to the underlying algorithm of BDT- or likelihood-based particleID, p refers to the particle species being identified, and d refers to whether the variable is trained to discriminate against all particle types (global) or one particle type (binary)[1].

5.2 Tag-and-probe method

The goal of the study is to evaluate the efficiency of a given $\mathcal{P}_e$ selection

$$\epsilon = \frac{N_{\text{sig, after cut}}}{N_{\text{sig, before cut}}} \tag{5.1}$$

in both data and MC and compute the ratio. The ratio is then provided to analysts as a correction factor when applying the same $\mathcal{P}_e$ selection. In most calibration channels, the number of signal events can be obtained from performing a fit to an invariant mass distribution, such as in $J/\psi \to e^+e^-$ events. In Bhabha events, no clear mass peak exists to perform a fit to. However, since the sample can be reconstructed with very high purities $\mathcal{O}(99.99\%)$, as seen in the logarithmically scaled distribution plots in Figure 5.2, the impact of the background rate is small for the resulting ratios and can simply be estimated from simulation. This estimation is accounted for in a systematic uncertainty.
This analysis method of counting the events is referred to as *tag-and-probe*. Here, the two final state electrons are labelled in two categories:

- Tag: The tag electron is always required to pass a fixed electronID selection; in the case of this study, it is required to fulfill $\mathcal{P}_e^{\text{BDT}} > 0.5$.

[1] d is omitted for global discriminators and $d = b$ for binary discriminators

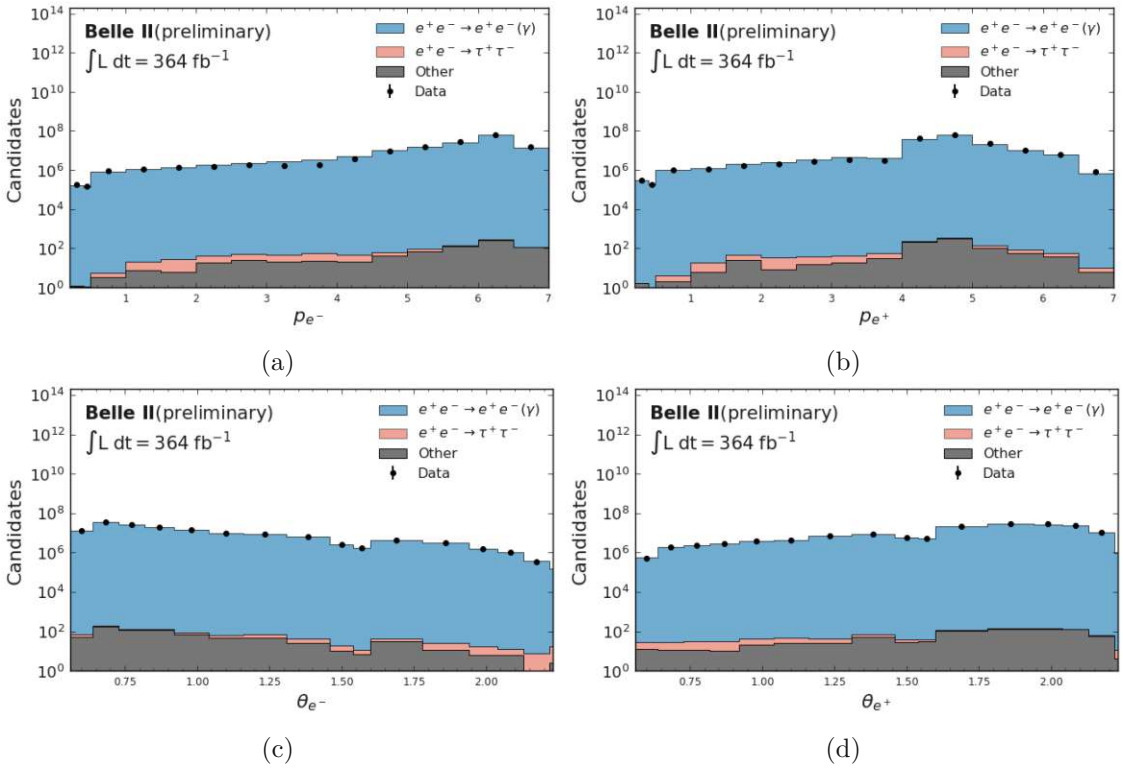


Figure 5.2: Laboratory frame momentum and polar angle distributions for the probe in reconstructed Bhabha events, with a loose electronID selection on the tag. The y axis is scaled logarithmically. The left plots shows electrons, and the right plots positrons.

- Probe: The other electron in the event is then an unbiased set of candidates, with minimal requirements placed upon it. The efficiency of a given $\mathcal{P}_e$ requirement can then be evaluated as the fraction of probe candidates passing the event over all probe candidates.

In simulated data, the efficiency is evaluated using only signal events and is expressed as the ratio

$$\epsilon_{\text{MC}} = \frac{N_{\text{tag+probe}}}{N_{\text{tag}}}, \tag{5.2}$$

where $N_{\text{tag+probe}}$ counts the events meeting the requirements for both electrons, while N_{tag} counts the events that satisfy only the preselection for the tag. In data, the background must first be subtracted by estimating the fake rate β from MC for both tag and tag+probe selections. The resulting data efficiency is then

$$\epsilon_{\text{Data}} = \frac{(1-\beta_{\text{tag+probe}})N_{\text{tag+probe}}}{(1-\beta_{\text{tag}})N_{\text{tag}}}. \tag{5.3}$$

Due to the high purity in the sample, the fake rate correction is of order $\mathcal{O}(10^{-3}) - \mathcal{O}(10^{-4})$ depending on the p and θ region.

Process	Run-dependent MC N_{evt} [$\times 10^9$]	Run-dependent MC $\mathcal{L}_{\text{int}}$ [fb^{-1}]	Run-independent MC N_{evt} [$\times 10^9$]	Run-independent MC $\mathcal{L}_{\text{int}}$ [fb^{-1}]
$e^+e^- \to e^+e^-$	2.7	36	7.4	100
$e^+e^- \to e^+e^-e^+e^-$	14.1	361	1.18	200
$e^+e^- \to e^+e^-\mu^+\mu^-$	6.8	361	1.1	200
$e^+e^- \to \mu^+\mu^-$	1.7	1444	1.1	1000
$e^+e^- \to \tau^+\tau^-$	1.3	1444	0.9	1000

Table 5.1: Simulated data samples available for the Bhabha study. The ratio of N_{evt} and $\mathcal{L}_{\text{int}}$ differs due to run-dependent prescales.

Finally, the correction factor is computed as

$$r = \frac{\epsilon_{\text{Data}}}{\epsilon_{\text{MC}}}. \tag{5.4}$$

The statistical uncertainty on efficiencies is obtained from the binomial standard deviation

$$\sigma = \sqrt{np(1-p)}, \tag{5.5}$$

with the probability p of the probe passing the selection and the sample size n.
While the analysis principle is straightforward, special attention has to be paid to keep the selection unbiased with respect to other analyses, to estimate occurring systematic uncertainties, and to account for the effect of prescales at various levels of the trigger involved.

5.3 Data samples

The study is performed on the full run 1 data sample corresponding to a luminosity of 364 fb^{-1}. Correction factors are applied for the most recent campaign of Monte Carlo production, including run-independent and run-dependent MC. Simulated samples include signal Bhabha samples, as well as several different background samples. These include di-photon $e^+e^- \to e^+e^-\ell^+\ell^-$ events, $e^+e^- \to \mu^+\mu^-$ events, and τ pair $e^+e^- \to \tau^+\tau^-$ events. While still small, the largest contribution of the backgrounds stems from $\tau^+\tau^-$ events, where both τ can decay into electrons. An overview of the luminosities of the simulated samples is shown in Table 5.1.

5.4 Candidate selection

As discussed in Section 3.4, the read-in rate of the data acquisition system is limited, and to maximize the amount of $B\overline{B}$ events recorded, certain events must be rejected by the trigger in real-time. Different Bhabha recognition trigger lines are defined in both the L1 and HLT triggers and come with prescales that only allow a given percentage to be recorded by the DAQ. The selection must rely on trigger lines to properly compare data to

MC. Otherwise, prescales that can differ by run period and event characteristics can not properly be taken into account. Before detailing the characteristics of the relevant trigger lines, there are some basic requirements to loosely suppress backgrounds:

- The number of charged tracks in the event is required to be exactly two.
- Since there are no neutrinos in the event, the laboratory frame energy of the combined electron and photon system is constrained to $10.8 < E_{\gamma,e} < 11.2$ GeV.
- Some events with one spurious or misreconstructed track in data are not simulated. Such events are rejected by requiring the total energy of e^+ and e^- to be larger than 5.5 GeV.
- The tag cut of the tag-and-probe method requires that at least one of the electrons satisfies the condition of its BDT-based electronID to be greater than 0.5. This electron is then the tag-side electron.

The remaining selection criteria rely on L1 and HLT trigger lines. Events are distinguished between two types of Bhabha events, and independent trigger requirements are used to select them:

- $e^+e^- \to e^+e^-$: Bhabha scattering without a radiated photon present in the final state, referred to as *flat* Bhabha events, are characterized by being back-to-back in the c.m. frame, with both tracks carrying the majority of the available energy and momentum. Due to the boosted collision, the events can still be reconstructed in different areas of polar angle and laboratory frame momentum. However, the two kinematic variables are highly correlated in this case.
- $e^+e^- \to e^+e^-\gamma$: Radiative Bhabha events cover a wider range of momenta and angles in the c.m. frame due to the radiated off photon. The radiative Bhabha trigger lines typically require a highly energetic photon cluster in addition to the two charged tracks. In radiative Bhabha, the correlation between laboratory frame momentum and polar angle is less correlated than for flat Bhabha.

In addition to separating into flat and radiative Bhabhas, the study employs several control samples using different selection criteria of trigger lines to evaluate systematic uncertainties. Differences between the samples in the electron identification performance are accounted for when providing the correction factors and uncertainties. It is important to note that, while this systematic uncertainty is based on different trigger line selections, the origin of the discrepancies does not necessarily lie directly in problems with the trigger, but that differences in the selected physics processes can lead to different electronID performance.

5.4.1 L1 Trigger lines

The following L1 trigger lines are used in separating samples:

- 3D Bhabha: The 3D Bhabha trigger, commonly used as veto in physics analysis, selects Bhabha events by the energy of the electrons, and the angular difference in the c.m. frame to identify back-to-back tracks. Because of a limited read-out rate budget in the L1 triggers, energy requirements use only information from the ECL. The energy requirements are that both clusters deposit at least 3 GeV in the ECL and one cluster at least 4.5 GeV. The back-to-back requirements are $165° < \theta_1^* + \theta_2^* < 190°$ and $160° < |\phi_1^* - \phi_2^*| < 200°$.
- High Energy (*hie*): The hie trigger requires energy deposited in the ECL to be at least 1 GeV, and that the 3D Bhabha trigger did not fire. Because of the low energy requirements and the complementary nature with 3D Bhabha, this is a suitable L1 line for selecting radiative Bhabha decays.
- Bhapur: Looser version of the 3D Bhabha trigger bit, relaxing the angular requirements to $160° < \theta_1^* + \theta_2^* < 200°$ and $140° < |\phi_1^* - \phi_2^*| < 220°$. The energy requirements are also loosened to at least 2.5 GeV for both clusters and at least 4 GeV for one of the two.
- Low Multiplicity (*lml1*): The lml1 requires at least one ECL cluster located in the ECL barrel θ region with more than 2 GeV. Due to its comparably low prescale compared to other lines that do not veto Bhabha events, it is suitable to select flat Bhabha candidates. To avoid biasing the set of probe electrons when lml1 is required, it is ensured that the tag fired the trigger by additionally requiring $E_{tag}^{ECL} > 2$ GeV and $0.61 < \theta_{tag} < 2.15$.

5.4.2 HLT Trigger lines

While in early data taking the HLT functioned primarily to classify events, data taken in later run periods requires at least one HLT line to be fired. Subsequently, to correctly account for prescales, at least one HLT line has to be required in the selection. The following lines are considered:

- n2GeVPhotonBarrel: Requires a photon cluster in the ECL barrel with at least 2 GeV.
- radBhabha: Selects radiative Bhabha events without information from the photons directly, *i.e.* at least 1 GeV of missing c.m. frame momentum with corresponding polar angle between 10° and 170°. Additionally, the two tracks are required to have momenta of at least 1 GeV in the c.m. frame; their azimuthal angle difference is required to be smaller than 170°; and one of the tracks is required to have $\frac{E_{\mathrm{ECL}}}{p} > 0.8$. When applying this trigger line, the E/p requirement has to be placed on the tag electron to avoid biasing the selection on the probe.

		$e^+e^- \to e^+e^-$			$e^+e^- \to e^+e^-\gamma$		
		eeflat + lml1	selectee + bhapur	selectee + bhapur + lml1	eegamma + hie	eegamma + lml1	eegamma + lml1 + allow b2b
HLT	ee_flat_xx_yy	✓					
	selectee		✓	✓			
	n2GeVPhoton or radiativeBhabha				✓	✓	✓
L1	lml1	✓		✓		✓	✓
	bhapur		✓	✓			
	bhapur (unprescaled, via HLT)	✓					
	hie				✓		
Other	b2b veto (not L1 bh3d)				✓	✓	

Table 5.2: The different sample selections according to their trigger requirements. In ee_flat, xx and yy refer to lower and upper limits of θ_e. They are separated into individual lines to allow theta-dependent prescale.

- selectee: This trigger for flat $ee \to ee$ events requires the invariant mass of the combined electrons to be larger than 5.29 GeV and at least one of the tracks to deposit at least 1.5 GeV into the ECL.
- ee_flat: Another flat Bhabha trigger with identical conditions to the selectee line, with the main difference of applying polar angle dependent prescales rather than one combined prescale across the full angular range. Additionally, ee_flat requires that the Bhapur L1 trigger bit was fired before applying the prescale.

To fully address the differences in electronID performance between different physics processes, the samples listed with their trigger line selections in Table 5.2 are all used to compute correction factors independently.

For candidates reconstructed in each category listed in Table 5.1, several prescales are applied at different stages, differing between data and MC and between run periods. To properly account for prescales, candidates are weighed by the cumulative prescale obtained by multiplying relevant prescales. Certain HLT lines implicitly require L1 lines that must be additionally accounted for.

5.5 Binning

The electronID efficiency ratio, given in Equation (5.4), is computed in bins of the laboratory frame momentum and the polar angle theta due to potential detector geometry effects influencing the performance differently. For instance, electrons in the barrel are identified

Bin	θ [rad]	θ [deg]	Region
1	(0.22, 0.4)	(12.6, 22.9)	Forward endcap
2	(0.4, 0.56)	(22.9, 32.1)	Forward endcap
3	(0.56, 0.64)	(32.1, 36.7)	Barrel edge
4	(0.64, 0.73)	(36.7, 41.8)	Barrel edge
5	(0.73, 0.82)	(41.8, 47.0)	Barrel edge
6	(0.82, 0.92)	(47.0, 52.7)	Barrel
7	(0.92, 1.04)	(52.7, 59.6)	Barrel
8	(1.04, 1.16)	(59.6, 66.5)	Barrel
9	(1.16, 1.31)	(66.5, 75.1)	Barrel
10	(1.31, 1.46)	(75.1, 83.7)	Barrel
11	(1.46, 1.54)	(83.7, 88.2)	Barrel
12	(1.54, 1.6)	(88.2, 91.7)	ECL gap
13	(1.6, 1.78)	(91.7, 102.0)	Barrel
14	(1.78, 1.94)	(102.0, 111.2)	Barrel
15	(1.94, 2.04)	(111.2, 116.9)	Barrel
16	(2.04, 2.13)	(116.9, 122.0)	Barrel
17	(2.13, 2.22)	(122.0, 127.2)	Barrel edge
18	(2.22, 2.4)	(127.2, 137.5)	Backward endcap
19	(2.4, 2.6)	(137.5, 149.0)	Backward endcap
20	(2.6, 2.71)	(149.0, 155.3)	Backward endcap

Table 5.3: Bin edges of provided electronID correction factors, in radians and degrees, and characterization of the respective detector region.

more efficiently than in the endcaps. Additionally, correction factors are split by charge into electrons and positrons since the opposite curvature of the tracks can affect the performance. Finally, correction factors are provided for multiple versions of the electronID variable, including $\mathcal{P}_e^{\text{BDT}}$, $\mathcal{P}_e^{\text{LH}}$, $\mathcal{P}_{e,b}^{\text{BDT}}$ and $\mathcal{P}_{e,b}^{\text{LH}}$.

The Bhabha sample reaches as low as 200 MeV, and the first bin edges in laboratory frame momentum are chosen in tight intervals at $[0.2, 0.4, 0.5]$ GeV. The remaining bin edges are chosen with bin widths of 500 MeV, matching the choices in other particleID control samples.

When choosing bins in the polar angle theta, detector geometry effects have to be taken into account. Compared to previous performance studies, a tighter binning in theta was chosen, particularly for bins at the edge of the barrel region. Due to the non-uniform angular distributions of electrons in the Bhabha sample, shown in Figure 5.2, wide bins with high angular dependency can lead to biased correction factors. This was of particular importance for a Belle II measurement of light-lepton flavour universality, which is dominated by uncertainties from LeptonID performance [86]. The resulting binning isolates key areas, such as the gap in the ECL, and tightens bins around the edges of the barrel, where the performance declines rapidly. The final binning is shown in Table 5.3.

5.6 Systematic uncertainties

The computed binned correction factors are supplied alongside their uncertainties. In addition to the statistical uncertainty in Equation (5.5), systematic uncertainties from two sources are considered and evaluated:

5.6.1 Trigger line selection

As described above, a trend of discrepancies in efficiencies between trigger line selection is observed, which indicates a dependency on the type of Bhabha event reconstructed. To estimate the uncertainty associated with an observed dependency of correction factors depending on the trigger selections, the control samples defined in Table 5.2 are used to compute corrections independently. The samples are generally split between selections focusing on either radiative or flat Bhabha. Discrepancies primarily occur between all radiative samples and all flat samples, but some discrepancies are also observed within each category.
A systematic uncertainty is assigned if the computed correction factors between two samples are outside of their combined statistical uncertainties. The difference is determined by taking the largest difference between any sample and the nominal correction factor value.

5.6.2 Background estimation

In the efficiency computation for data, the background rate β is estimated from MC. In a tag-and-probe approach, the difference in β between the data and MC is not accounted for. To assess the impact of this uncertainty, an additional control sample is defined with an assumption of $\beta = 0$, corresponding to 100% purity. The sample is included along with the other trigger line control samples, and the systematic uncertainty is calculated according to the description in the previous section. In this way, the full impact of the background estimation correction is included in the provided uncertainty.

The discrepancies between different samples for both sources of systematics are shown in Figure 5.3 in a central barrel bin $\theta \in [1.31, 1.46]$ rad for positrons at multiple working points of $\mathcal{P}_e^{\mathrm{BDT}}$. While samples still agree at low electronID thresholds, higher thresholds introduce increasing discrepancies, leading to higher systematic uncertainties at higher working points.

5.7 Results

Radiative Bhabha events dominate the lower momenta region between 200 MeV and 4 GeV, while flat Bhabha events primarily occur above 4 GeV. Most analyses are sensitive to momenta below 3 GeV. Due to the strong correlation between p and θ in flat Bhabha events, these events offer limited sensitivity in the relevant kinematic region. Therefore, radiative Bhabha selected with the ”eegamma+hie” selection listed in Table 5.2 are used as nominal values. The ratios from flat Bhabha events enter the correction factors through its systematic uncertainty. The resulting correction factors for run-dependent MC are shown

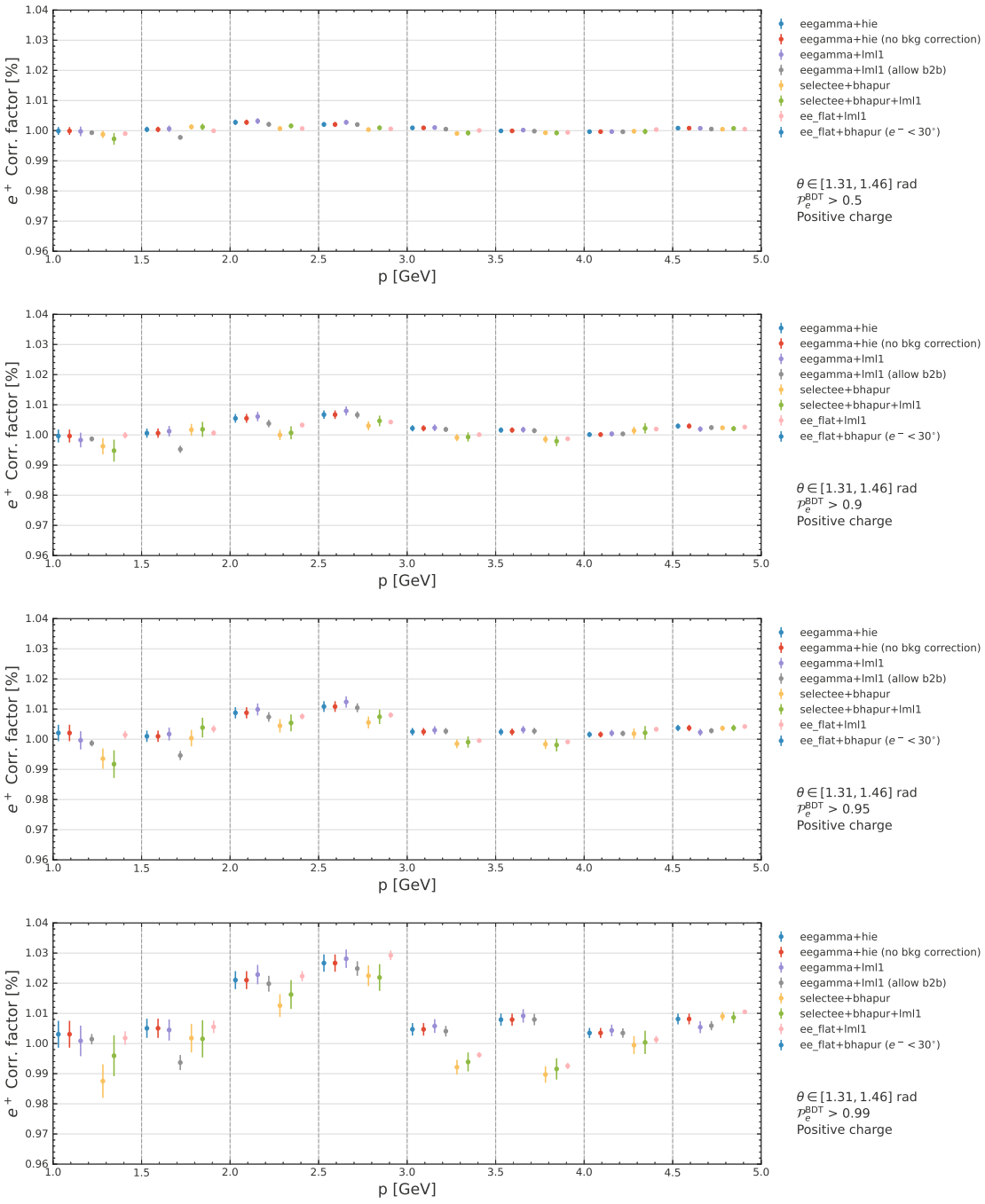


Figure 5.3: Differences in correction factors between control samples for multiple increasing thresholds from top to bottom. The correction factors shown are for positron momenta between 1 and 5 GeV in the barrel bin $\theta \in [1.31, 1.46]$ rad and different $\mathcal{P}_e^{\mathrm{BDT}}$ selections.

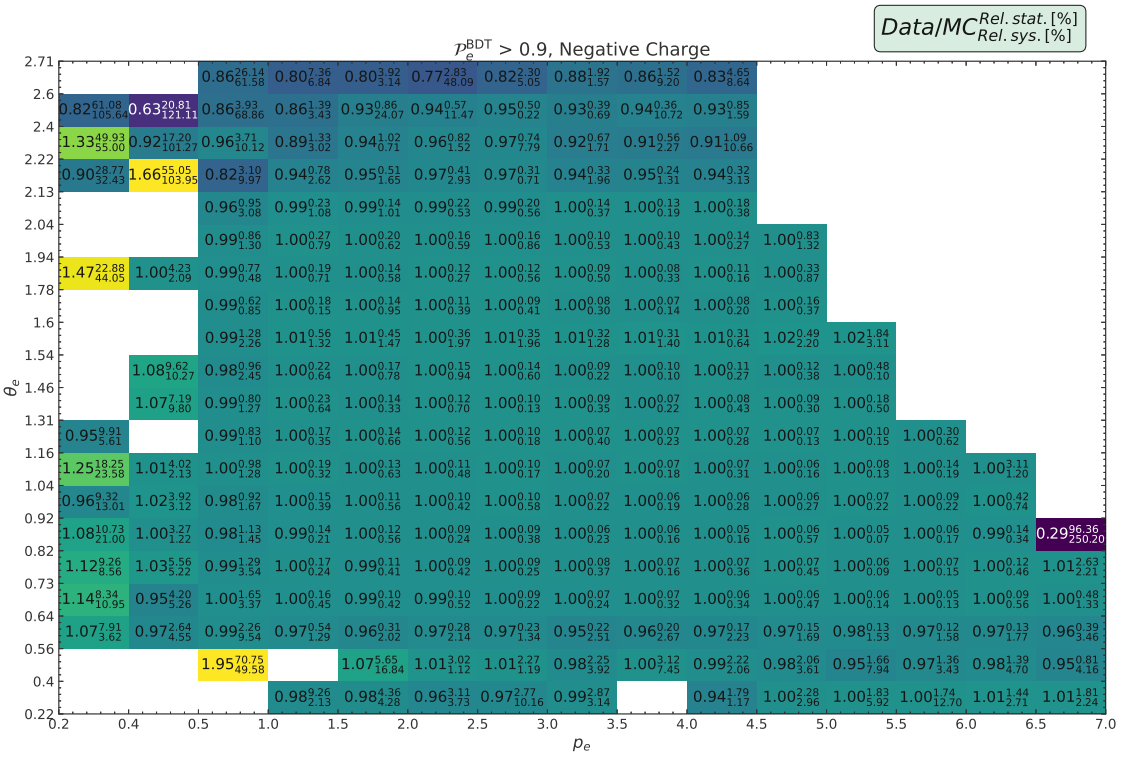


Figure 5.4: Results of correction factors with statistical and systematic uncertainty, for $\mathcal{P}_e^{\mathrm{BDT}} > 0.9$ and electrons.

for $\mathcal{P}_e^{\mathrm{BDT}} > 0.9$ in Figures 5.4 and 5.5. Uncertainties are dominated by systematics in most observed bins. In the central kinematic region, observed correction factors are close to 1, indicating good modelling of electronID in simulation when compared to data. Additional correction factors are shown in Appendix A.

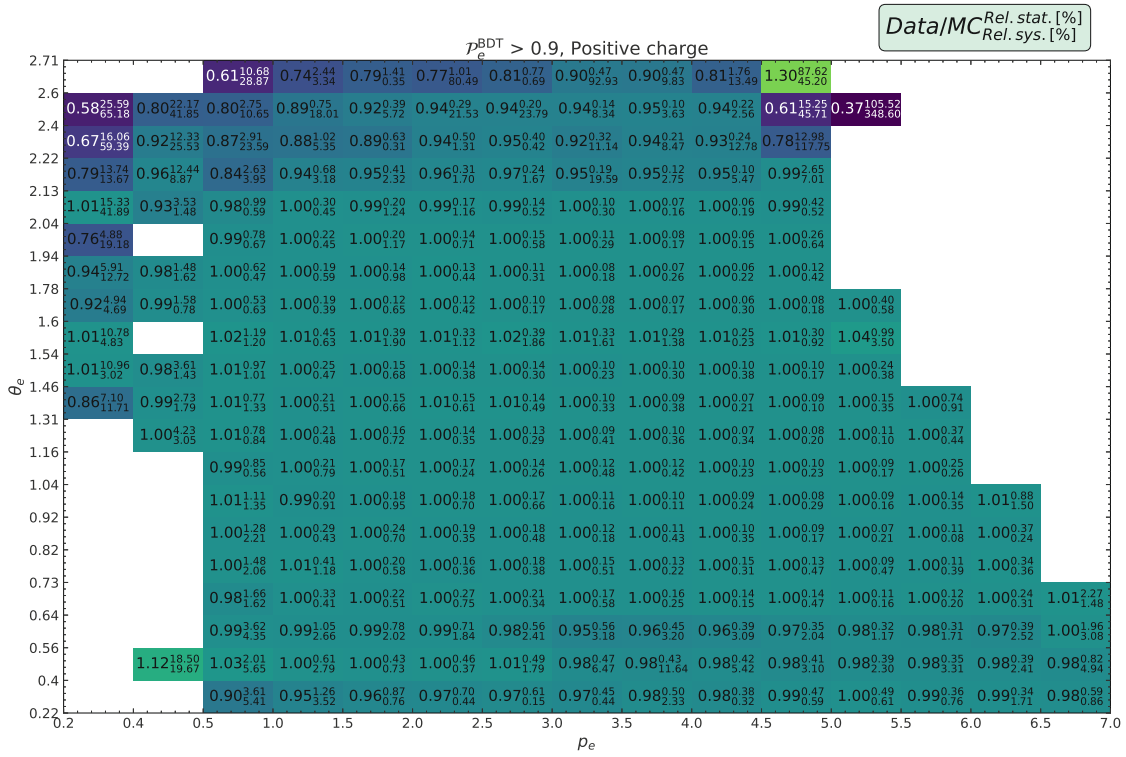


Figure 5.5: Results of correction factors with statistical and systematic uncertainty, for $\mathcal{P}_e^{\mathrm{BDT}} > 0.9$ and positrons.

CHAPTER 6

Determination of $|V_{cb}|$ from exclusive $B \to D\ell\nu_\ell$ decays

This chapter describes the $|V_{cb}|$ measurement in detail. Following an outline of the key stages discussed in Section 6.1, the utilized data samples are presented in Section 6.2. Definitions for important variables used throughout the analysis description are given in Section 6.3. In Section 6.4, the reconstruction of $B \to D\ell\nu_\ell$ in four different final states, with the respective selection criteria to suppress backgrounds and the reconstruction approach for kinematic variables is detailed. Control samples used throughout the analysis are defined in Section 6.5. The procedures applied to correct for mismodellings are described in Sections 6.6 and 6.7, including corrections from external inputs and control samples. A simulated and collision data comparison is presented in Section 6.8. The setup of the signal extraction template fit and the form factor fit is explained in Sections 6.9 and 6.10. Validation of the fit procedures is presented in Section 6.11, followed by the results, including a sensitivity estimate and a preliminary result on a smaller data sample, in Section 6.12.
As principal analysts, my supervisor and I conducted the analysis described as part of the Belle II collaboration.

6.1 Analysis overview

This section provides an overview of the experimental approach to obtain a measurement of $|V_{cb}|$ using $B \to D\ell\nu_\ell$ decays, introducing the essential ingredients of each stage. A schematic summary is shown in Figure 6.1.
Using reconstruction techniques introduced in Section 3.7, final state particles observed in the Belle II detector are combined into composite candidates. Particle identification algorithms and vertex fitting are employed to reduce candidates from backgrounds. The signal channel is reconstructed in four final states, charged $B^+ \to \bar{D}^0\ell^+\nu_\ell$ and neutral

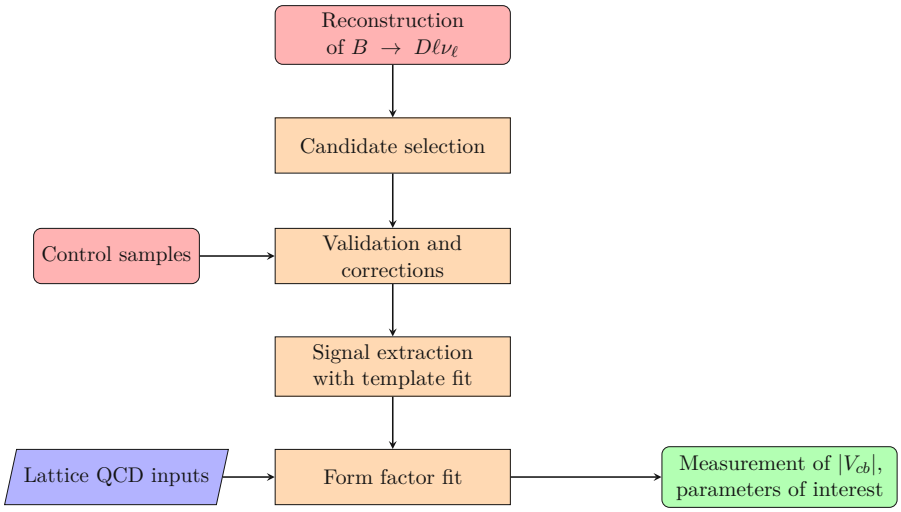


Figure 6.1: A schematic overview of the steps involved in the measurement. After reconstructing $B \to D\ell\nu_\ell$ events in different final states in collision data, selection criteria are defined for optimal sensitivity. The modelling of signal and backgrounds is validated using control samples. Corrections are derived from control samples or external inputs to account for known and observed effects. Signal yields are extracted using a template fit, and the resulting differential decay rates are fit separately to the form factor. Using inputs from Lattice QCD in the form factor, $|V_{cb}|$ and additional parameters of interest are accessible.

$B^0 \to D^-\ell^+\nu_\ell$ channels, each with $\ell = (e, \mu)$.
Following the guiding statistical principles described in Chapter 4, backgrounds are suppressed with rectangular and multivariate selection criteria designed to optimize the statistical sensitivity of the resulting subsample. The dominant background component originates from $B \to D^*\ell\nu_\ell$ decays, in which the excited D^* state decays to a D meson. The small difference in invariant mass between D^* and D mesons results in the emission of a slow pion or a soft photon in the D^* decay, which can easily be missed. Due to the larger branching fraction of $B \to D^*\ell\nu_\ell$, these decays contribute significantly to the observed backgrounds. Explicit vetoes are employed to reduce the impact of these contributions.
After the selection criteria are optimized and before comparing the resulting simulation data set with collision data, distributions are first studied in control samples enriched in certain categories of background processes. Discrepancies observed between data and simulation are corrected by calibrating corrections in the control samples or external performance inputs, such as the electron identification correction factors introduced and computed in Chapter 5. The signal distribution is validated through the use of additional control samples. Related signal decays that mirror $B \to D\ell\nu_\ell$ decays in important aspects but can be reconstructed with higher purity are studied. Once all corrections are in place, the $B \to D\ell\nu_\ell$ reconstruction sample on MC is compared to Belle II data.
The signal yields of the reconstructed data sample are then extracted using a binned maximum likelihood fit, as introduced in Section 4.2. Signal and backgrounds are categorized

Sample	Experiment	Buckets	$\int \mathcal{L} dt$ [fb^{-1}]
Chunk 1	7	-	0.5
	8	-	1.7
	10	-	3.7
Chunk 2	12	-	54.6
Chunk 3	14	16, 16b	16.5
Chunk 4	16	17	10.3
	17	18	10.7
Chunk 5	18	19a, 20, 21, 22, 23, 24, 25	89.9
Prompt	20	26	3.8
	22	28, 29	32.1
	24	30, 31, 32, 33	85.6
	26	35, 36	54.8
Total	7 - 26		364.1 ± 2.3
	10 - 26		361.9 ± 2.3

Table 6.1: Data luminosities separated by experiment.

into templates, and the statistical model, as specified in the `pyhf` formulation of maximum likelihood fits, is optimized to determine the best-fit values for the parameters controlling the template shapes. The signal yields are obtained simultaneously in 10 bins of the kinematic variable w, introduced in Section 2.3. Systematic uncertainties are included in the fit as nuisance parameters and map directly onto the fit result. The resulting covariance matrix then contains a combination of uncertainties from both statistical and systematic sources. The fit result consists of 40 signal yields and the total covariance matrix.
The differential decay rates can be computed from the fit result using the determined amount of B meson decays in the sample, the B lifetime, and the involved charm branching ratios. The differential decay rate, formulated in Section 2.3, depends on the kinematic variable w and the form factor. By performing a second fit of the form factor in various parameterizations to the measured differential decay rates and including data from Lattice QCD as additional nuisance parameters, a corresponding value of $|V_{cb}|$ is obtained.

While the measurement is described primarily as a $|V_{cb}|$ measurement, additional parameters of interest can be extracted from the analysis. The differential decay rate is measured in 10 bins of w for all modes, with the corresponding covariance matrix. The total branching ratio of the involved decays can be computed from the measured signal yields by summing over bins of w. Finally, since the measurement is done for both e and μ modes, a measurement

of light-lepton flavour universality

$$R_{e/\mu} = \frac{\mathcal{B}(B \to D\,\mu\,\overline{\nu})}{\mathcal{B}(B \to D\,e\,\overline{\nu})} \tag{6.1}$$

can be obtained from the fit result.

6.2 Input data samples

The analysis is based on the data collected at the Belle II experiment during run 1 data taking, lasting from 2019 to the summer of 2022. The total integrated luminosity of the data sample taken at the $\Upsilon(4S)$ resonance is 364.1 fb^{-1}, with a detailed breakup into run periods given in Table 6.1. The Belle II performance group measures the amount of B mesons collected in the collision data sample with a data-driven method [87]. The amount of B meson pairs obtained in the run 1 data set is

$$N_{B\overline{B}} = (387.1 \pm 6) \times 10^6, \tag{6.2}$$

where the uncertainty includes statistical and systematic uncertainties.
In addition, Belle II collects data at energies 60 MeV below the $\Upsilon(4S)$ resonance where the energy is not sufficient to produce B mesons and continuum processes dominate. This analysis uses a sample of 42 fb^{-1} of so-called *off-resonance* data for calibrating and validating continuum backgrounds.
Simulated data produced through means described in Section 3.6 are used in the analysis. The MC samples used are run-dependent, *i.e.* variations such as a run period dependence of the beam energy and beam backgrounds are included in the simulation. A breakup of samples included in the analysis, with their corresponding cross sections and luminosities, is shown in Table 6.2. The integrated luminosity corresponds to four times the collected data luminosity.
Run-dependent beam backgrounds are overlaid on simulated data. Due to missing beam background information in early data-taking periods, not all experiments have corresponding run-dependent MC produced. Subsequently, those experiments are excluded from the analysis, reducing the overall integrated luminosity to 361.9 fb^{-1}.

6.3 Variable definitions

When reconstructing candidates, as described in Section 3.7, it is useful to compute quantities such as kinematic variables of composites or variables describing the event shape to characterize a specific candidate and potentially suppress it with a selection. In this section, definitions of variables used throughout the analysis are given.

6.3.1 Track impact parameters

To suppress charged tracks not originating from the IP, curved particle tracks reconstructed in the event are characterized by impact parameters describing the helix trajectory:

Process	σ [nb]	$\int \mathcal{L} dt$ [fb^{-1}]	N [10^6]
$\Upsilon(4S) \to B^+B^-$	0.54	1444	785.7
$\Upsilon(4S) \to B^0\bar{B}^0$	0.51	1444	742.1
$ee \to c\bar{c}$	1.329	1444	1933.8
$ee \to d\bar{d}$	0.401	1444	583.5
$ee \to s\bar{s}$	0.383	1444	557.3
$ee \to u\bar{u}$	1.605	1444	2335.4
$ee \to \tau\overline{\tau}$	0.919	1444	1337.2

Table 6.2: The different physics processes simulated, listed with their corresponding cross section, integrated luminosity, and total number of events.

- dz and dr parameterize the point-of-closest approach (POCA) of the trajectory with respect to the interaction point. dz corresponds to the signed distance between POCA and IP in the z plane, while dr refers to the magnitude of the distance in the transversal plane.

- nCDCHits is equivalent to the number of individual hits a track registers in the CDC. A minimum requirement on nCDCHits suppresses low momentum particles that curve strong enough not to reach the CDC.

- ϕ and θ characterize the angular distribution of the track. The polar angle θ is used to divide the detector into three regions: the $12.01° < \theta < 31.36°$ forward endcap, the $32.2° < \theta < 128.7°$ barrel and the $131.5° < \theta < 155.03°$ backward endcap. θ can also be used to require a track to be within acceptance of the CDC ($17° < \theta < 150°$) or the ECL ($12.4° < \theta < 155.1°$). In contrast to the polar angle, the detector is mostly symmetric in its azimuthal angle ϕ, defined from $-\pi$ to π.

6.3.2 Event properties

Event variables are computed from every observed cluster and track in the event and are therefore independent of which decay is reconstructed. *Event shape* variables are a subclass of event variables specifically used to characterize the geometric shape of events. In B factories, event shape variables are commonly used to distinguish $\Upsilon(4S) \to b\bar{b}$ events, which are in general more spherically distributed, from continuum events, which are more jet-like distributed due to the lower invariant mass of the produced quarks. This is illustrated in Figure 6.2. Commonly used event shape variables include:

- The Fox-Wolfram moments [89]

$$H_l = \sum_{i,j} \frac{|\vec{p_i}||\vec{p_j}|}{s} P_l(\cos\,\theta_{ij}) \tag{6.3}$$

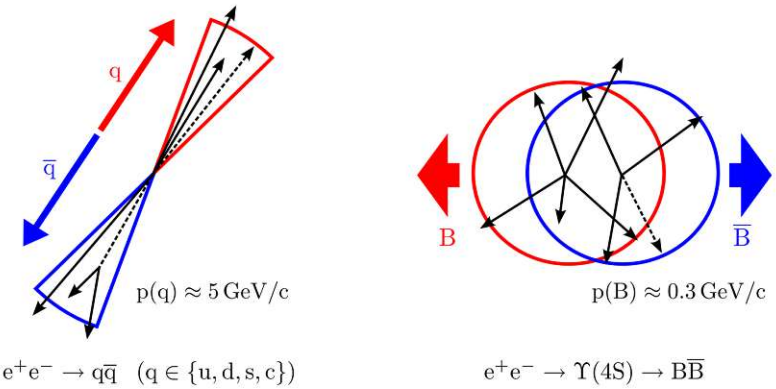


Figure 6.2: A schematic illustration of the geometrical event shape difference between continuum and $B\overline{B}$ events [88].

provide a measure of event sphericity. Here, indices i and j run over all particles in the event, θ_{ij} corresponds to the angle between particles i and j and P_l refers to the order l Legendre polynomial. Particularly the ratio between the second and zeroth Fox-Wolfram moment

$$R_2 = \frac{H_2}{H_0} \tag{6.4}$$

is used to suppress backgrounds from continuum events. R_2 ranges from 0 to 1, with typical values closer to 0 for spherically uniformly distributed events and closer to 1 for jet-like events.

- To further improve suppression of continuum events, modified Fox-Wolfram moments were developed by the Belle collaboration [90, 91]. In the modification, particles are treated differently depending on whether they are involved in the signal reconstruction. The Fox-Wolfram moments are then further divided into three categories of charged, neutral and missing particles. The modified Fox-Wolfram moments used in this analysis are defined as

$$H_{\mathrm{xl}}^{so} = \sum_i \sum_{\mathrm{jx}} \begin{cases} Q_i Q_{\mathrm{jx}} \left|p_{\mathrm{jx}}\right| P_l\left(\cos\theta_{i,jx}\right), & \text{for } l \in [1,3], \\ \left|p_{\mathrm{jx}}\right| P_l\left(\cos\theta_{i,jx}\right) & \text{for } l \in [0,2,4]. \end{cases} \tag{6.5}$$

Here, i runs over particles involved in the signal reconstruction; jx indexes the other particles in categories $x = (\text{charged}, \text{neutral}, \text{missing})$; Q_i and Q_{jx} are the charges of the particles i and jx; p_{jx} is the momentum of particle jx; and $P_l\left(\cos\theta_{i,jx}\right)$ is the l-th Legendre polynomial of the cosine of the angle between particles i and jx.

Further event variables that are not specifically related to the geometric shape of the event can also be used to distinguish signal from background in $B \to D\ell\nu_\ell$ analysis:

- Quantities related to missing energy and momentum such as p^*_{miss} and m_{miss} are computed by subtracting the momentum and energy of all visible tracks and clusters

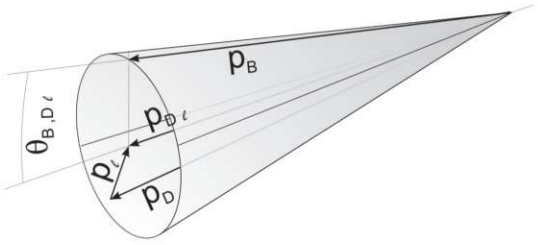


Figure 6.3: Construction of the angle $\cos\theta_{BY}$.

from the beam parameter expectation. Events involving *e.g.* neutrinos can then be preselected by requiring a minimum amount of missing momentum.

- Complementary, visible energy variables such as E^*_{vis} are computed by summing up all seen particles and tracks. Typical $B\overline{B}$ events have visible energies between 6 and 11 GeV.

6.3.3 Kinematic variables

Besides standard kinematic quantities, such as p, E and m, several additional kinematic variables are defined and used throughout this analysis:

- The composite of reconstructed D and ℓ is referred to as Y. Its invariant mass and momentum p_Y and m_Y are useful in distinguishing $B \to D\ell\nu_\ell$ signal from $B \to D^*\ell\nu_\ell$ backgrounds.

- A useful metric is the cosine of the angle between nominal B mesons and the reconstructed Y system

$$\cos\theta_{BY} = \frac{2\,E^*_B E^*_Y - m^2_B - m^2_Y}{2|p^*_B||p^*_Y|}. \tag{6.6}$$

  The construction of θ_{BY} is illustrated in Figure 6.3. Due to the mass assumptions made for the Y system, θ_{BY} is only a physical angle for signal decays; in this case, its cosine is restricted to the region of -1 to 1. For background processes, $\cos\theta_{BY}$ can have values outside of that interval, making it a useful variable in background suppression. In the context of this analysis, $\cos\theta_{BY}$ is used as a fit variable in the signal extraction.

- Owed to the fact that the D meson is a spin 0 odd-parity pseudo-scalar, the differential decay rate is dependent kinematically only on the hadronic recoil w, as defined in Section 2.3. The vector D^* meson has additional dependencies on three angles that can be defined as shown in Figure 6.4. The angles θ_V and χ are defined through the $D^* \to D\pi$ decay plane, and therefore there is no equivalent in $B \to D\ell\nu_\ell$ decays. The angle θ_ℓ is defined as the angle between the virtual W boson and the lepton and can be reconstructed for both $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ decays. Because the

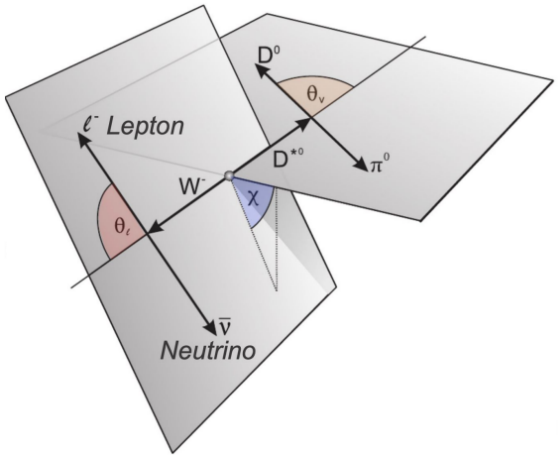


Figure 6.4: The differential decay rate of $B \to D^*\ell\nu_\ell$ decays depends on the hadronic recoil w and the three angles θ_ℓ, θ_V, and χ, defined as shown in the figure.

distribution of this angle differs depending on the final state, it is a useful variable to suppress backgrounds. The angle is labelled $\theta_{\ell,W}$ to distinguish it from the laboratory frame polar angle of the reconstructed lepton track, θ_ℓ.

- The angle between reconstructed D and ℓ, $\theta_{D,\ell}$, offers additional descriptive power in separating $B \to D\ell\nu_\ell$ from semileptonic backgrounds.

6.3.4 Rest-of-event variables

After building composite particles from reconstructed tracks and clusters, the rest-of-event (ROE) is computed by summing up all unused tracks and clusters into a composite object. A so-called *mask* defines minimum conditions for unused tracks to contribute to the ROE to avoid contributions from spurious tracks or clusters.
The invariant mass m_{ROE}, the momentum p_{ROE}, the c..m. energy E^*_{ROE} and the polar angle of the ROE momentum direction θ^*_{ROE} can all be used in further classification of candidates.

6.3.5 Particle identification

Identification algorithms described in Section 3.7.1 and Chapter 5 assign classifiers to each final state track. Global particle identification classifiers based on likelihoods are referred to as $\mathcal{P}_e^{\text{LH}}$, $\mathcal{P}_K^{\text{LH}}$, etc. BDT-based classifiers are labelled $\mathcal{P}_e^{\text{BDT}}$. Binary classification is labelled $\mathcal{P}_{K,b}$, and in this thesis, specifically represents the likelihood of kaons over pions.

6.4 Reconstruction of $B \to D\ell\nu_\ell$ and event selection

6.4.1 Candidate reconstruction

Four $D\ell\nu$ final states are reconstructed independently: $B^+ \to \bar{D}^0\ell^+\nu_\ell$ and $B^0 \to D^-\ell^+\nu_\ell$, with ℓ being either e or μ. The D mesons are reconstructed from the decays $D^0 \to K^-\pi^+$ for neutral B meson modes and $D^+ \to K^-\pi^+\pi^+$ for charged B meson modes. The reconstruction strategy is *untagged*, as described in Section 3.7.2, *i.e.* no explicit decay is reconstructed on the opposite side B meson. Nevertheless, there is some dependency on the second B meson through means such as ROE variable selections, sometimes referred to as *inclusive tag*. Throughout the chapter, untagged and inclusive tags are used interchangeably. The data samples considered for reconstruction are presented in Section 6.2. The MC samples are run-dependent. To reduce computation time and resources, the samples are first skimmed with a loose hadron High Level Trigger condition that consists of a Bhabha veto and a requirement of at least three good[1] tracks. The HLT skim reduces the file sizes by about 80%.

To reconstruct candidates, charged tracks with distances of closest approach $\mathrm{d}r < 1cm$ and $|\mathrm{d}z| < 3$ cm are considered. Tracks are required to lie within the acceptance range of the CDC, and the transverse momentum p_T is required to be greater than 50 MeV.

Reconstructing photons is necessary for event shape variables and ROE. To reject spurious clusters being identified as photons, clusters must have at least 1.5 hits in the ECL. The non-integer requirement comes from cases where energy is split among neighbouring clusters, causing each crystal to record a shared hit. Additionally, the cluster should be within the CDC's accepted polar angle range range. Finally, the timing of the cluster in the ECL must align with the collision timing, allowing for a window of ±200 ns due to limited timing resolution.

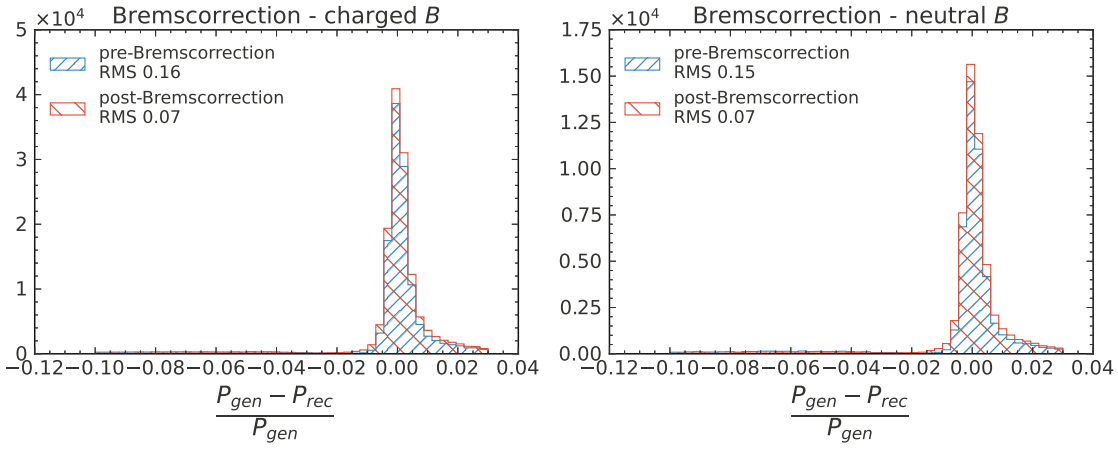
Leptons are selected by requiring a center-of-mass momentum between 0.6 and 2.4 GeV for electrons and between 1.0 and 2.4 GeV for muons. For electrons, $\mathcal{P}_e^{\mathrm{BDT}}$ is required to be larger than 0.9. For muons, $\mathcal{P}_\mu^{\mathrm{LH}}$ is required to be at least 0.9 [2].

To allow for the correction of radiated-off photons from electrons, photons within an angular acceptance cone around the electron direction and pass certain criteria are added to the electron candidate by correcting its momentum. In addition to the base requirements for ROE photons, Bremsstrahlung photons have additional energy requirements based on their polar angle:

- $\theta_{\mathrm{cluster}} < 0.56$: $E_{\mathrm{cluster}} > 40$ MeV
- $0.56 < \theta_{\mathrm{cluster}} < 2.23$: $E_{\mathrm{cluster}} > 55$ MeV
- $\theta_{\mathrm{cluster}} > 2.23$: $E_{\mathrm{cluster}} > 90$ MeV

[1] Good tracks are defined as tracks that satisfy basic impact parameter requirements, *i.e.* $p_T > 0.2$ GeV, $\mathrm{d}r < 2$ cm and $|\mathrm{d}z| < 4$ cm.

[2] Due to differences in Data/MC in the SVD likelihood contribution, observed by the Belle II performance group, the SVD likelihood is omitted from the muonID computation.



(a) (b)

Figure 6.5: The resolution of electron momentum for truth-matched electrons in both the charged and neutral B modes. The blue curve shows the resolution before applying bremscorrection, while the red curve shows the improved resolution when allowing to correct electrons with radiated photons. The legends denote the root mean square deviations of the candidates.

Electrons below 1 GeV are corrected with photons up to 900 MeV, and those above 1 GeV with photons up to 1.2 GeV. The momentum resolution is defined as the deviation of the reconstructed momentum from the generated momentum in MC samples over the generated momentum:

$$\delta p = \frac{p_{\text{gen}} - p_{\text{reco}}}{p_{\text{gen}}}. \tag{6.7}$$

The momentum resolution before and after allowing for Bremsstrahlung correction is shown in Figure 6.5.

Final state hadrons must have momenta greater than 500 MeV and nCDCHits greater than 20 to ensure they register in the CDC. This requirement avoids a large systematic uncertainty usually associated with slow pions. Final state kaons are identified using a selction of $\mathcal{P}^{\text{LH}}_{K,b} > 0.1$.

Appropriately charged K and π candidates are combined into composite D mesons in the decay modes mentioned above. A selection on the invariant mass equivalent to three times the mass resolution is placed on the composite D meson to reduce combinatorial backgrounds. This is equivalent to $1.85 < m_{K\pi} < 1.88$ GeV for the charged B modes and $1.86 < m_{K\pi\pi} < 1.88$ GeV for the neutral B modes. To reduce backgrounds from $c\bar{c}$ continuum events, the c.m. momentum of the D meson candidate is required to be smaller than 2.4 GeV.

Appropriately charged D composites are combined with lepton candidates to form Y candidates. A vertex fit is performed on the full Y decay chain using the *TreeFitter* module, introduced in Section 3.7. The initial decay vertex is set to the average B meson origin. The χ^2 distributed probability associated with the fit result is required to be at least 0.05.

6.4.2 Candidate classification

The selection criteria are designed for $B \to D\ell\nu_\ell$ decays. Nevertheless, not all candidates passing every criterion are signal candidates, and various backgrounds enter the obtained sample. Backgrounds are classified into categories encompassing different groups of physics processes:

- Continuum: Candidates from events where no $\Upsilon(4S)$ meson was produced in the collision, but rather e^-e^+ resulted in $q\overline{q}$ ($q \in (u, d, c, s)$) or $\tau^+\tau^-$
- $B \to D^*\ell\nu$: A prominent background in the analysis is down-feed from $B \to D^*\ell\nu$ decays. Candidates where the mother B meson of the lepton decays to $B \to D^*\ell\nu$ fall in this category.
- True D backgrounds: The remaining $B\overline{B}$ backgrounds are split into two categories, classified by whether the reconstructed D candidate is truth-matched as a D meson. If a candidate does not fall into signal or other background categories, and the D meson candidate is truth matched as a D meson with proper charge, the candidate is classified as true D background;
- False D backgrounds: If a candidate does not meet the requirements for the signal or other backgrounds, and the D meson candidate is not truth-matched, it is classified as a false D background. Most candidates in this category are combinatorial backgrounds resulting from random combinations.

Candidates where every particle is reconstructed correctly, with appropriate mother and daughter particles, except for the invisible neutrino that is not reconstructed, are classified as signal events.

6.4.3 $B \to D^*\ell\nu_\ell$ vetoes

Explicit vetoes of several $D^*\ell\nu$ final states are studied and implemented to reduce the amount $B \to D^*\ell\nu_\ell$ backgrounds. The approach involves pairing the D meson candidate with either a slow pion or a soft photon and then excluding it from the data set if the combined invariant mass aligns with a $D^* \to D\pi$ or $D^* \to D\gamma$ decay.
To optimize the veto criteria, a sample of $B \to D^*\ell\nu_\ell$ decays is reconstructed in multiple final states and selections are calibrated to optimize the $FOM = \frac{S}{\sqrt{S+B}}$, where S corresponds to the number of $B \to D^*\ell\nu$ events and B to the number of $B \to D\ell\nu$ events. The following modes are reconstructed and studied:

- $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$ with
  - $D^{*0} \to D^0\pi^0$
  - $D^{*0} \to D^0\gamma$
- $B^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ with

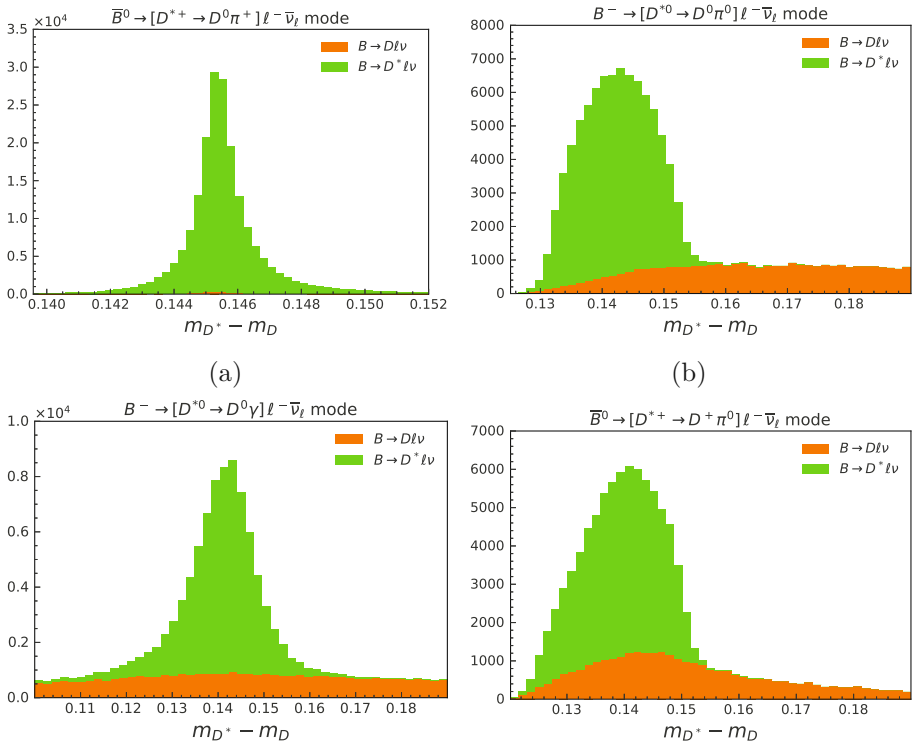


Figure 6.6: Distributions for the mass difference in each explicitly reconstructed $B \to D^*\ell\nu_\ell$ mode after optimization of selections. The modes in (a) and (b) are used to veto candidates in $B^+ \to \bar{D}^0\ell^+\nu_\ell$ reconstruction, while a veto of the mode in (d) is used for $B^0 \to D^-\ell^+\nu_\ell$.

- $D^{*+} \to D^0\pi^+$
- $D^{*+} \to D^+\pi^0$,

with $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ in all cases. Following the FOM optimization, slow π^+ are selected by requiring a momentum between 50 and 350 MeV. In the $D^* \to D^0\pi^+$ case, the angular difference between the D candidate and the slow pion must be smaller than 51.5°. In the soft photon decay mode, photons are required to have between 150 and 500 MeV; register at least 1.5 weighted ECL cluster hits; have cluster timing within a ±200 ns window; and have cluster polar angles within the CDC acceptance. For the π^0 modes, neutral pions immediately decay to a pair of photons. The photons are reconstructed with the same selections as for the soft photon; the invariant mass of the combined π^0 is required to be within 125 and 145 MeV; and the angle between D and π^0 mesons must be less than 74.5°.

The difference in invariance mass between reconstructed D meson and reconstructed D^* meson $\Delta m = m_{D^*} - m_D$ is a strong discriminator for $B \to D^*\ell\nu_\ell$ decays. The Δm distributions for each of the reconstructed $D^*\ell\nu$ modes after FOM optimization are shown in Figure 6.6.

Variable	Description
H_{20}^{so}	Modified Fox-Wolfram moment.
E_Y^*	Center-of-mass energy of the combined $D-\ell$ system.
p_{miss}^*	Missing momentum after subtracting all visible particles in the event.
m_{ROE}	Invariant mass in the rest-of-event, consisting of all particles and clusters not involved in signal reconstruction.
p_{ROE}	Magnitude of the rest-of-event momentum.
$\cos\theta_{\ell,W}$	Cosine of the angle between the direction of the lepton in the virtual W rest frame and the direction of the W boson in the B rest frame.
$\theta_{D,\ell}$	Angle between D meson and lepton in the lab frame.

Table 6.3: Final set of simulated annealing input variables deemed optimal in the feature preselection.

Using the optimized selections for slow pions and soft photons, the D^* candidate reconstruction is performed as part of the nominal $B \to D\ell\nu_\ell$ reconstruction. The reconstructed D meson is combined into a composite D^* candidate. The candidate is vetoed if Δm lies within a mass window around the peak. In the charged B modes, an improvement in figure of merit of the $D\ell\nu$ sample is seen when applying $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$ vetoes with the mass windows $\Delta m \in [140\text{ MeV}, 150\text{ MeV}]$ and $\Delta m \in [135\text{ MeV}, 150\text{ MeV}]$ respectively. In the neutral B mode, the $D^{*+} \to D^+\pi^0$ veto with the mass window $\Delta m \in [135\text{ MeV}, 145\text{ MeV}]$ is advantageous. The veto with soft photon reconstruction does not improve the figure of merit and is therefore not applied in either B mode.

6.4.4 Background suppression

The selections outlined in Section 6.4.1 and Section 6.4.3 are rectangular selections, used either to loosely preselect $B \to D\ell\nu_\ell$ candidates, exclude kinematic regions that are badly modelled or are disadvantageous in terms of systematic uncertainties. Various multivariate tools are available to further optimize the statistical sensitivity of the obtained sample, such as training a neural network or BDT based on input features. In the case of this analysis, the high efficiencies of the untagged approach lead to the measurement being dominated by systematic rather than statistical uncertainties. For this reason, the method of choice to further optimize the sample selection is a simultaneous optimization of rectangular selections using the simulated annealing algorithm introduced Section 4.1.2. Due to the systematic limitations, there is merit in keeping distributions of backgrounds under control rather than a pure statistical optimization. By relying on rectangular cuts, the optimized selection offers a high degree of interpretability and ensures that the selection can be studied in-depth using control samples.
The Python library `selanneal` [77] is employed to do a grid search of rectangular selections and maximize a given figure of merit by placing requirements on input features.

Variable	$B^+ \to \bar{D}^0\ell^+\nu_\ell$	$B^0 \to D^-\ell^+\nu_\ell$
H_{20}^{so}	[-0.371,0.485]	[-0.361,0.457]
E_Y^* [GeV]	[3.247,5.169]	[3.340, 5.174]
p_{miss}^* [GeV][3]	[0.698,4.345]	[0.745, 4.389]
m_{ROE} [GeV]	[0.510, 7.967]	[0.756,6.771]
$\cos\theta_{\ell,W}$	[-0.821, 0.553]	[-0.828, 0.866]
p_{ROE} [GeV]	[0.069, 2.858]	[0.091, 2.738]
$\theta_{D,\ell}$	[0.120, 3.137]	[0.339, 3.136]

Table 6.4: Rectangular selection optimized using simulated annealing.

To maximize the statistical sensitivity of the $|V_{cb}|$ measurement, the quantity

$$\sqrt[n]{\prod_{i=1}^{n} \frac{S_i}{\sqrt{(S_i + B_i)}}} \tag{6.8}$$

is used as a figure of merit, where i indicates the bin in the kinematic variable w. This metric represents a geometric mean over the classic figure of merits in each bin of w. It is chosen to prevent the optimization from excluding particular regions of phase space that naturally show lower purity.
The `selanneal` algorithm can handle a maximum of seven input features. A separate simulated annealing optimization is employed to select the best-performing subset from a larger pool of features to determine the optimal set of these seven features.

The selection is optimized independently for charged and neutral B modes. The input variables and the result of the simulated annealing algorithm are shown in Tables 6.3 and 6.4.
Distributions of all selections of MC with data overlaid are shown in Appendix C.

6.4.5 Sample composition and signal efficiency

After applying all selections, the sample is binned the sample into ten bins of w_{MC} with bin edges

$$[1., 1.06, 1.12, 1.18, 1.24, 1.3, 1.36, 1.42, 1.48, 1.54, w_{\max}], \tag{6.9}$$

where the value corresponding to maximal hadronic recoil $w_{\max} = \frac{m_B^2+m_D^2}{2m_B m_D}$ differs between charged and neutral B modes because of the invariant mass differences. The bin-wise event counts obtained from the run-dependent MC sample are shown in Table 6.5 for each category.

In Chapter 4, multiple performance metrics for sample composition were introduced, including efficiency, purity and figure of merit. To compute the efficiency, the total number of signal events in the MC samples is obtained by separately reconstructing them on a generator-truth level before detector simulation.
The different metrics are tracked throughout all individual selections in so-called *cut-flow*

	w bin	$B^+ \to \bar{D}^0 e^+ \nu_e$	$B^+ \to \bar{D}^0 \mu^+ \nu_\mu$	$B^0 \to D^- e^+ \nu_e$	$B^0 \to D^- \mu^+ \nu_\mu$
Signal1	$[1.0, 1.06]$	476.9	530.6	108.9	109.0
Signal2	$[1.06, 1.12]$	2072.4	2214.8	599.2	628.8
Signal3	$[1.12, 1.18]$	3986.6	4072.6	1421.3	1506.4
Signal4	$[1.18, 1.24]$	5933.9	5906.0	2563.3	2622.5
Signal5	$[1.24, 1.30]$	7814.0	7284.3	4057.8	3943.1
Signal6	$[1.30, 1.36]$	9678.8	8294.2	5658.6	5099.4
Signal7	$[1.36, 1.42]$	10542.2	8443.0	7049.4	5875.4
Signal8	$[1.42, 1.48]$	10833.6	7919.4	7802.7	6145.1
Signal9	$[1.48, 1.54]$	10115.5	7093.0	8027.3	5878.3
Signal10	$[1.54, w_{max}]$	7629.2	5121.5	5948.7	4187.9
$B \to D^* \ell\nu$	$[1.0, w_{max}]$	154609.1	141865.5	21315.8	19298.8
True D	$[1.0, w_{max}]$	61766.4	57856.9	17375.4	16240.1
Fake D	$[1.0, w_{max}]$	57722.4	54287.4	49434.0	47342.6
Continuum	$[1.0, w_{max}]$	38063.7	41305.8	16100.9	16803.4
Total		381244.8	352195.0	147463.6	135680.8

Table 6.5: Resulting yields for each component after applying all selections outlined in this section.

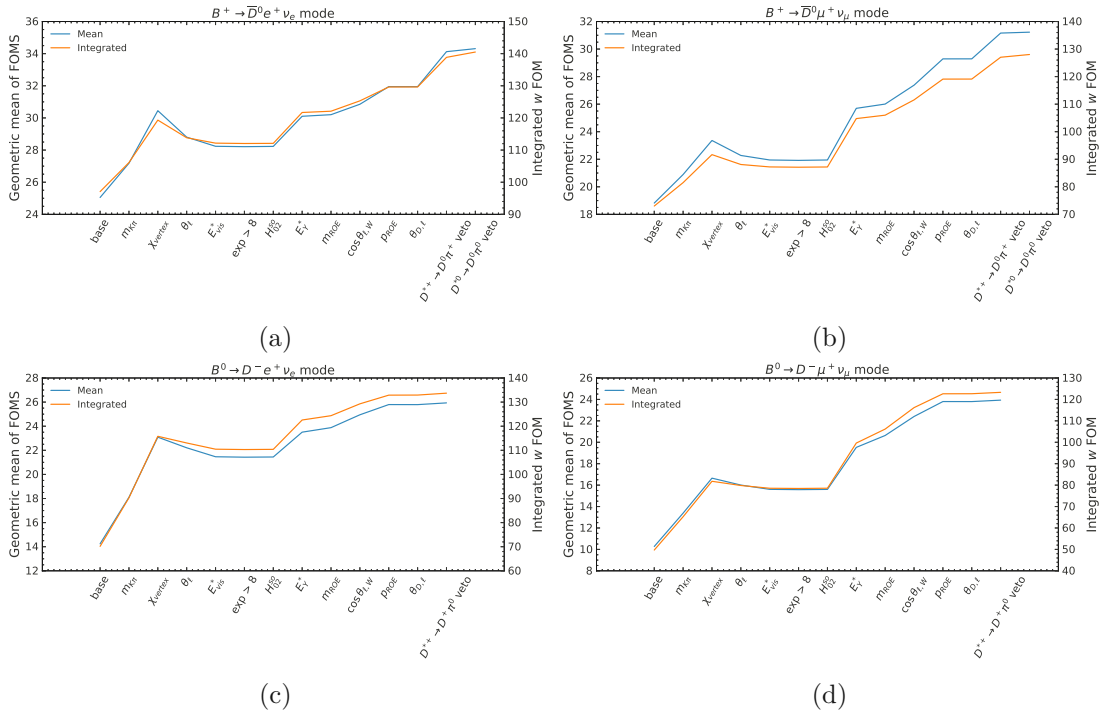


Figure 6.7: The figure of merit after each cut for all four modes, integrated over the full w range, is shown. The overall figure of merit for the entire sample and the geometric mean of the FOMs in each w bin, which serves as the target in the selected annealing, are displayed. Cuts not shown on the x axis but described in the text have already been applied in the steering file.

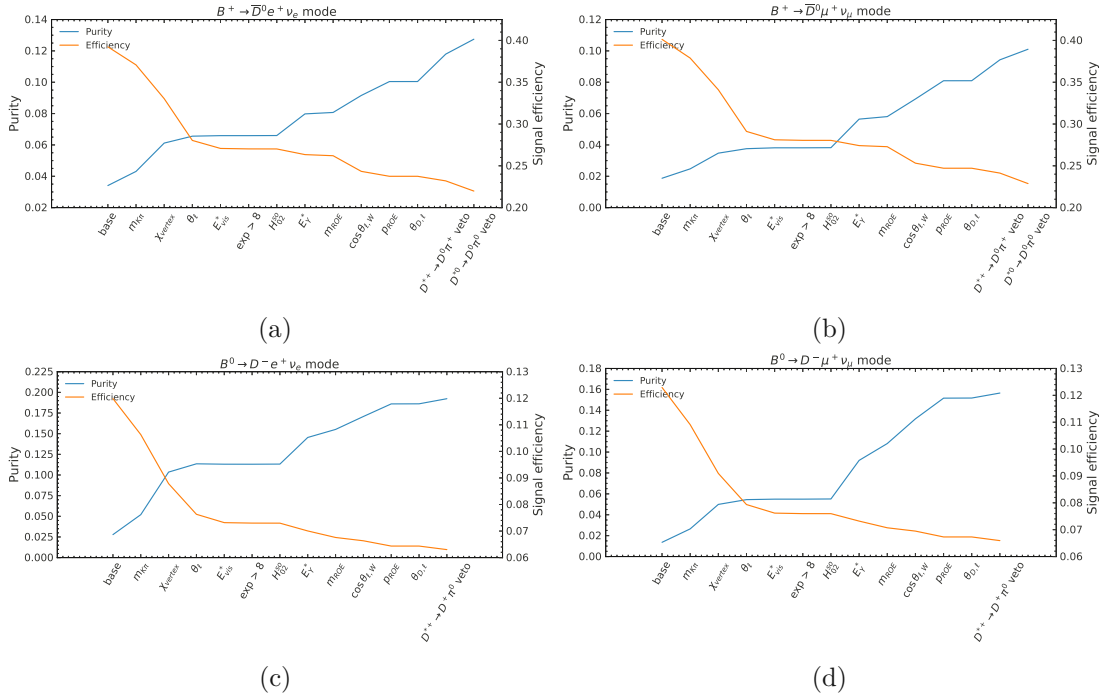


Figure 6.8: Purities and efficiencies after each cut for all four modes, integrated over the full w range.

diagrams in Figure 6.7 and Figure 6.8. The integrated FOM is computed over all bins of w, while the mean FOM gives the geometric mean over bins of w. The starting base is equivalent to the loose preselections already applied in the steering file. Note that not every selection is solely aimed to improve the FOM. For instance, early run periods were excluded due to missing beam background overlay files, and the lepton polar angle cut was applied to minimize the systematic uncertainty related to lepton identification..
The signal efficiency after all selections in each bin of w for all modes is given in Table 6.6. Cut-flow diagrams for each bin of w are shown in Appendix B.

6.4.6 Kinematic variable reconstruction

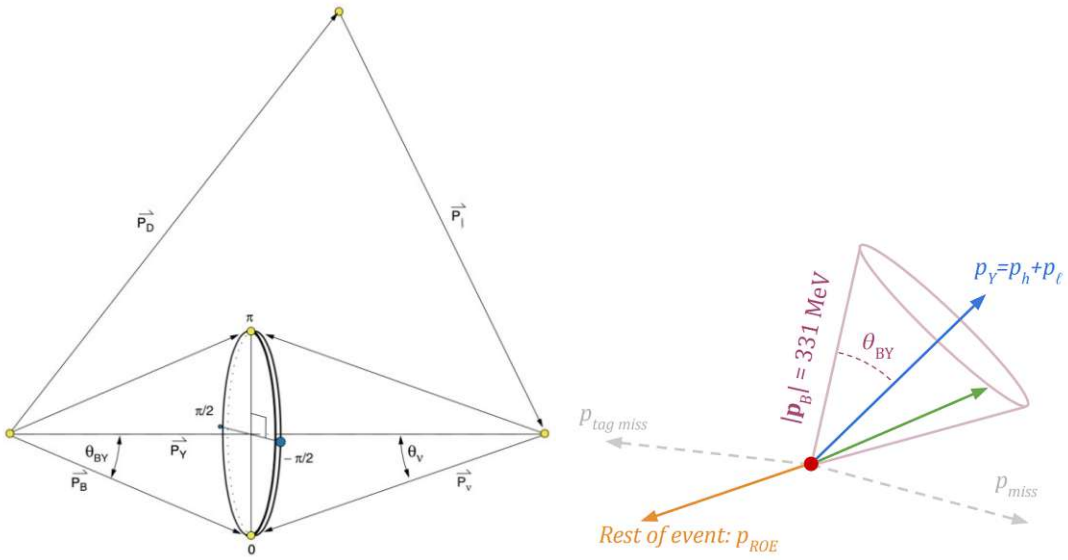
The differential decay rate of $B \to D\ell\nu_\ell$, introduced in Equation (2.44), depends, both explicitly and as a dependence of the form factor $\mathcal{G}(w)$, on the kinematic variable

$$w = \frac{m_B^2 + m_D^2 - q^2}{2m_B m_D} \tag{6.10}$$

which parametrizes the recoil momentum of the D meson in the B rest frame. The precision of determining the hadronic recoil relies on knowing the B meson direction. However, this direction cannot be measured directly because of the presence of the invisible neutrino. Since the analysis is untagged and the second B meson is not explicitly reconstructed, the neutrino momentum also cannot be directly inferred from the missing momentum. Several

	$B^+ \to \bar{D}^0 e^+ \nu_e$	$B^+ \to \bar{D}^0 \mu^+ \nu_\mu$	$B^0 \to D^- e^+ \nu_e$	$B^0 \to D^- \mu^+ \nu_\mu$
$w \in [1.00, 1.06]$	34.0%	29.9%	3.2%	3.0%
$w \in [1.06, 1.12]$	30.0%	28.9%	4.2%	4.0%
$w \in [1.12, 1.18]$	28.6%	27.1%	4.8%	4.9%
$w \in [1.18, 1.24]$	26.8%	25.8%	5.7%	5.6%
$w \in [1.24, 1.30]$	25.9%	23.6%	6.6%	6.3%
$w \in [1.30, 1.36]$	25.6%	21.8%	7.4%	6.7%
$w \in [1.36, 1.42]$	24.1%	18.8%	7.9%	6.6%
$w \in [1.42, 1.48]$	20.4%	15.3%	7.5%	6.2%
$w \in [1.48, 1.54]$	17.0%	12.1%	7.0%	5.3%
$w \in [1.54, w_{\max}]$	12.6%	8.9%	5.2%	3.9%

Table 6.6: Differential signal efficiencies with all selections in place.



(a) (b)

Figure 6.9: (a) Schematic of the diamond frame method, averaging over four directions. (b) Schematic of the ROE method, placing the B meson direction back-to-back with the p_{ROE}.

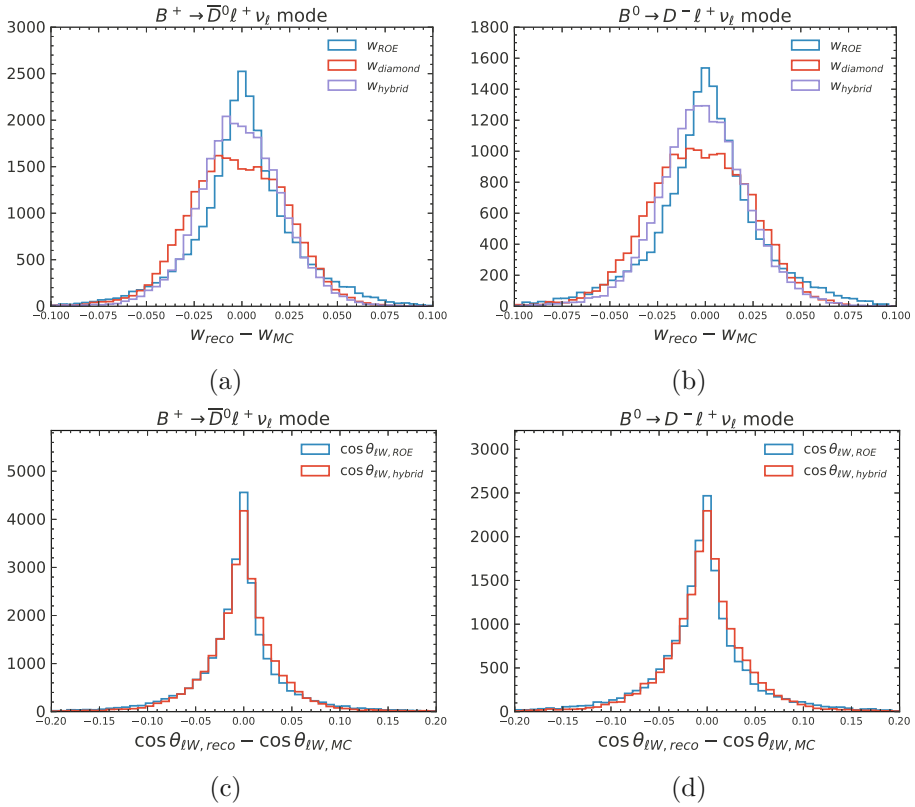


Figure 6.10: Resolutions for the kinematic variables w and $\cos\theta_{\ell,W}$. We compare approaches of ROE information only, diamond frame, and a combination of both. The combination results in the lowest width of the distributions. e and μ channels are combined for each B mode.

approaches are possible for reconstructing the B meson direction in untagged analyses: One technique is the *diamond frame* approach, developed by *BABAR* [92] in the context of analyzing $\overline{B}^0 \to D^{*+} e^- \overline{\nu}_e$ decays. The angle between Y and a nominal B meson θ_{BY}, defined in Equation (6.6), can be computed from reconstructed particles. However, while this angle determines the opening angle of a cone around Y, the flight direction of the signal B meson on the cone is unknown. The diamond frame technique assumes an arbitrary flight direction by choosing an azimuthal angle ϕ on the cone and computing the kinematic variables – *i.e.* w or q^2 – under that assumption. By repeating the procedure for multiple directions on the cone and assigning a weight for each, the parameters are obtained in a weighted average over the directions. Here, the weight is a $\sin^2\theta^*_B$ factor, obtained from the angular distribution of a spin-1 particle decaying to two spinless B meson daughters, as is the case in $\Upsilon(4S) \to B\overline{B}$ decays. In Ref. [92], *BABAR*chooses four uniformly spaced directions on the θ_{BY} spanned cone, illustrated in Figure 6.9a.
Another approach is to use information from the ROE to constrain the B meson direction, as was done at Belle [93]. After summing up unused tracks and clusters to obtain $\vec{p}^*_{ROE}$,

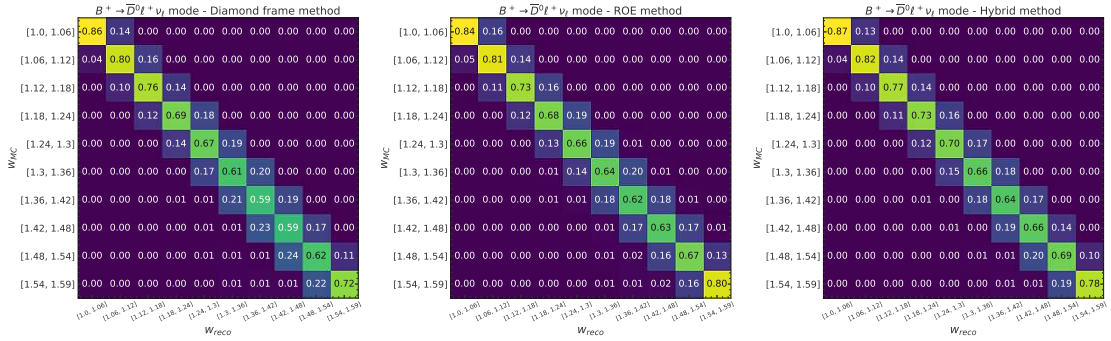
	w			$\cos\theta_{\ell,W}$	
	Diamond	ROE	Hybrid	ROE	Hybrid
$B^+ \to \bar{D}^0\ell^+\nu_\ell$	0.034	0.034	0.030	0.079	0.074
$B^0 \to D^-\ell^+\nu_\ell$	0.032	0.032	0.028	0.080	0.073

Table 6.7: The resolution of the reconstructed kinematic variables with different methods and for different B modes. The new hybrid method results in the best overall resolutions.

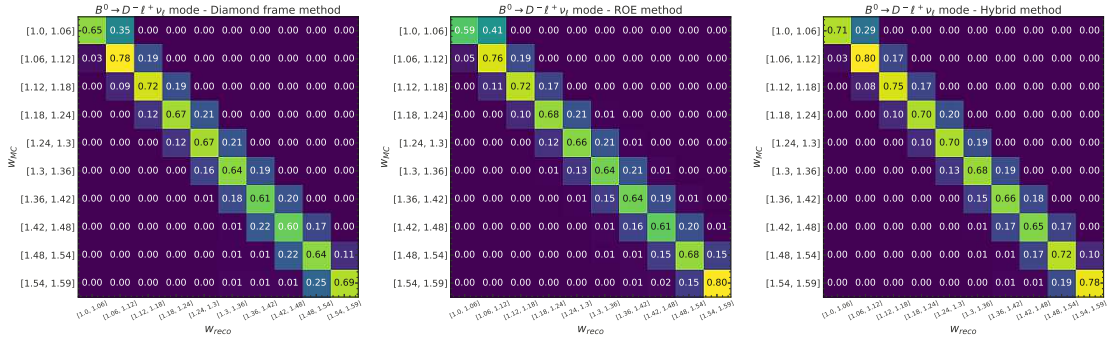
momentum conservation can be leveraged to place the B meson on the direction on the cone that is most back-to-back with $\vec{p}^*_{ROE}$, shown in Figure 6.9b.
In this analysis, a new method that combines the diamond frame approach with the ROE approach has been co-developed in collaboration with another Belle II analysis [94]. Like in the diamond frame approach, a weighted average is computed over arbitrarily chosen cone directions. In this approach, the weight

$$\alpha = \frac{1}{2}(1 - \vec{p}^*_B \cdot \vec{p}^*_{ROE}) \sin^2\theta^*_B, \tag{6.11}$$

takes into account how back-to-back the B meson lies with the ROE and the angular distribution in $\Upsilon(4S)$ decays. The same approach is used to compute the angle between the direction of the virtual W boson in the B rest frame and the lepton in the W rest frame $\cos\theta_{\ell,W}$. The method is implemented by taking a weighted average over ten randomly chosen, uniformly spread directions on the cone.
Figure Figure 6.10 shows histograms of the deviations to the generator value in w and $\cos\theta_{\ell,W}$ for the different approaches to computing kinematic variables using charged and neutral B meson samples. Electron and muon modes are combined in the diagrams. The hybrid technique, combining the diamond frame and information from the ROE, results in the best resolution for both w and $\cos\theta_{\ell,W}$. The method draws advantages from the ROE method in the more distinct peak while also having narrower tails similar to the diamond frame method. The resolution is defined as the Gaussian width of the distributions shown in Figure 6.10. A tabulated form of the resolutions is shown in Table 6.7. The novel hybrid method improves the resolution of w and $\cos\theta_{\ell,W}$ by about 12% and 6% with respect to preceding methods.
An additional important performance metric for different reconstruction approaches are bin-to-bin migrations. Since there is a difference between true and reconstructed w, a candidate reconstructed in a particular w bin can belong to a different bin according to its true value of w. This can be represented graphically by plotting the migration matrix, *i.e.* the fraction of events from true bin i ending up in reconstructed bin j. The hybrid method shows the highest diagonal terms on average, indicating less bin-by-bin migration than in other reconstruction techniques. To address the effects of migration and correct for the discrepancies between true and reconstructed values, *unfolding* techniques are applied. Details of the unfolding process are discussed in Section 6.9.1.



(a)



(b)

Figure 6.11: Migration matrices with each reconstruction approach for the charged and neutral B modes. The bin entries list the fraction of events reconstructed in each bin relative to their true origin.

6.5 Reconstruction of control samples

Defining control samples is crucial to validate the modelling of signal and background distributions. For different categories of backgrounds, finding *sidebands*, *i.e.* regions that are usually cut away in signal selection, or alternative reconstruction approaches that lead to a sample enriched in a particular type of background allows comparing collision data with simulation as a test of whether the behaviour is well-understood. Alternative decay channels that closely resemble the signal decay in one or more aspects but can be reconstructed with greater purity are useful to validate signal distributions. These channels help verify reconstruction elements such as the vertex fit.
Particularly in a blinded analysis, control samples allow for validation before unblinding the collision data in sensitive regions for the measurement. This section defines and motivates the various control samples used for backgrounds and signals.

6.5.1 Background control samples

The backgrounds are categorized into four categories introduced in Section 6.4.2. For each of false D, true D, $B \to D^*\ell\nu_\ell$, and continuum backgrounds, a control sample is defined

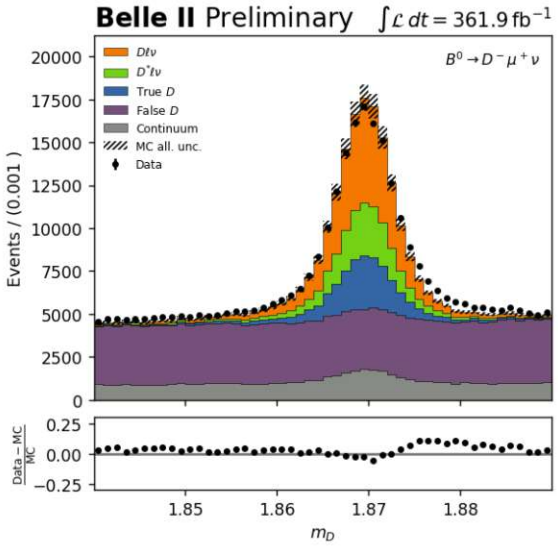


Figure 6.12: The m_D distribution peaks at the nominal D meson mass, with some width due to detector resolution. Combinatorial false D backgrounds dominate the sidebands outside of the signal region and can, therefore, be used as a control sample.

that is dominated by the respective background type.

D mass sideband

The false D component is dominated mainly by backgrounds where reconstructed final state hadrons fulfill the mass requirements by chance. Since random particle combinations are uniformly distributed within a specific mass window, while other backgrounds typically form Gaussian peaks around the nominal D meson mass, examining the sidebands far from this mass yields a sample composed predominantly of combinatorial backgrounds. The distributions of the components around the m_D peak are shown in Figure 6.12. Therefore, the requirements for this control sample are for m_D to be below 1.84 GeV or above 1.89 GeV. To further isolate the false D component from continuum backgrounds in the control sample, R_2 must be smaller than 0.4. This results in a sample with a false D component share of 73% and 88% for charged and neutral B modes, respectively.

Off-resonance data

A natural control sample for continuum backgrounds arises from off-resonance data. Data taken below the $\Upsilon(4S)$ resonance with c.m. energies below the $B\overline{B}$ production threshold consists entirely of continuum backgrounds. Compared to other control samples, the statistical significance is diminished due to the limited amount of integrated luminosity available for off-resonance data. During run 1 data taking, Belle II collected a data sample taken at 60 MeV below the $\Upsilon(4S)$ resonance corresponding to 42 fb^{-1}.

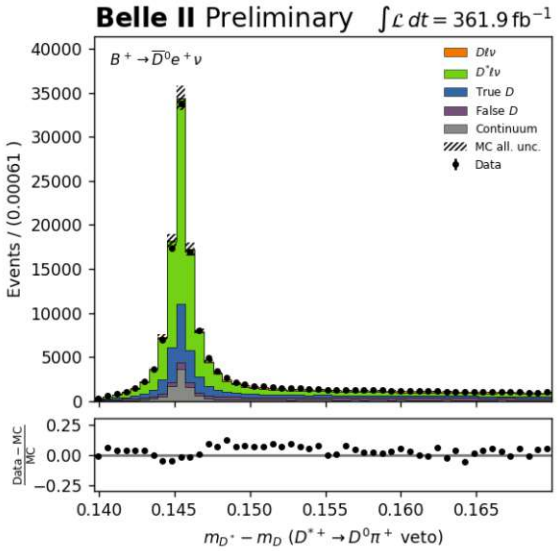


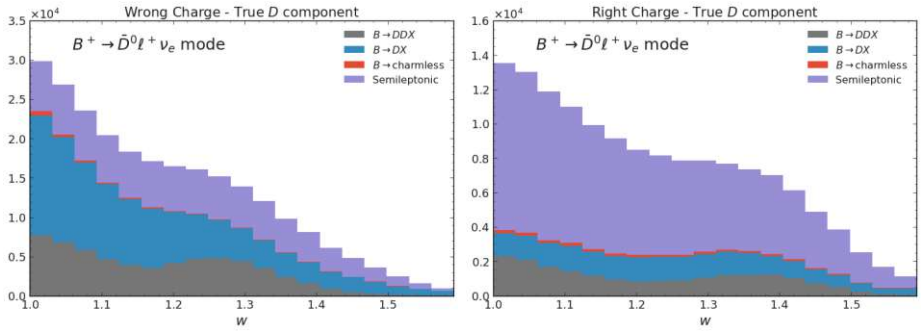
Figure 6.13: The plot shows the stacked Δm distributions for candidates where a slow π^+ fulfilling the requirements stated in Section 6.4.3 is found. A $B \to D^*\ell\nu_\ell$ control sample can be defined by selecting candidates close to the peak.

Inverted $B \to D^*\ell\nu_\ell$ veto

The explicit $B \to D^*\ell\nu_\ell$ vetoes designed to reduce down-feed can be used to design control samples. Samples enriched in this background can be obtained by inverting the selection and selecting only candidates within the mass windows consistent with $D^* \to D$ decays. The relevant distributions for the $B^+ \to \bar{D}^0 e^+ \nu_e$ sample are shown in Figure 6.13. In the charged B modes, this corresponds to candidates passing either the $D^{*0} \to D^0\pi^+$ or the $D^{*0} \to D^0\pi^0$ veto. In the neutral B mode, only the $D^{*+} \to D^+\pi^0$ veto is applied and is therefore inverted. This way, a sample consisting of 63% and 42% of $B \to D^*$ decays in charged and neutral B modes, respectively, is obtained.

Wrong charge reconstruction

Backgrounds where the D meson is reconstructed correctly, but the reconstructed decay is neither $B \to D\ell\nu_\ell$ nor $B \to D^*\ell\nu_\ell$, can contain various different physics processes. A significant contribution is from hadronic decays, where the lepton stems from the other side B meson. Such a lepton is termed *secondary*. Additionally, other semileptonic B decays, such as decays into higher excitation D^{**} states, fall into this category. Because of the different contributions, finding a control sample that mirrors the composition of this background type is challenging.
A sample enriched in secondary lepton contributions is obtained by reconstructing $B \to D$ decays in charge conservation violating combinations, such as $B^0 \to D^- e^- \overline{\nu}$. False D contributions are reduced by additionally requiring a tighter mass window of $1.86 < m_D < 1.87$ GeV. While this sample contains a true D share of 75% and 38% for charged



(a) (b)

Figure 6.14: The compositions of the true D component when reconstructing while requiring the (a) wrong charge and (b) right charge combinations. A bigger share of semileptonic decays is observed in the right charge reconstruction. Among hadronic decays, there are proportionally more $B \to DX$ decays in the wrong charge sample, particularly towards low w.

Selections	$B \to D\pi$	$B \to D\pi$ from $B \to [D^* \to D\pi_{\text{slow}}]\pi$	$B \to K^*\ell$ from $B \to [J/\psi \to \ell\ell][K^* \to K\pi]$
Full reconstruction	$1.85 < m_D < 1.88$ GeV $m_{bc} > 5.27$ GeV $\lvert\Delta E\rvert < 50$ MeV $R_2 < 0.4$	$1.85 < m_D < 1.88$ GeV $m_{bc} > 5.27$ GeV $\lvert\Delta E\rvert < 50$ MeV $140 < m_{D^*} - m_D < 150$ MeV $p_{\pi_s} < 400$ MeV $p_{D^*} < 2.5$ GeV	$\lvert\Delta m_{K*}\rvert < 75$ MeV $m_{bc} > 5.27$ GeV $\lvert\Delta E\rvert < 100$ MeV $\lvert\Delta m_B\rvert < 50$ MeV $2.96 < m_{J/\psi} < 3.19$ GeV
Partial reconstruction		π daughter of $B_{fullreco}$ D grand-daughter of $B_{fullreco}$	K^* daughter of $B_{fullreco}$ ℓ grand-daughter of $B_{fullreco}$

Table 6.8: Selection criteria for the control samples used to study signal distributions.

and neutral B modes, it is important to remember that the composition is not entirely equivalent to the true D component in the nominal reconstruction. A comparison of compositions of the true D component in wrong charge and nominal reconstruction samples is shown in Figure 6.14. The most significant difference is a reduced contribution from other semileptonic decays.

6.5.2 Signal control samples

If the signal distributions differ in Data and MC, a particular selection could lead to a bias in selection efficiency, which would then propagate further into signal yields. It is essential to reconstruct signal control samples to study distributions and verify that selections do not bias signal efficiency through Data/MC disagreements. Three control samples with similarities to $B \to D\ell\nu_\ell$ are reconstructed, with an overview of selections given in Table 6.8.

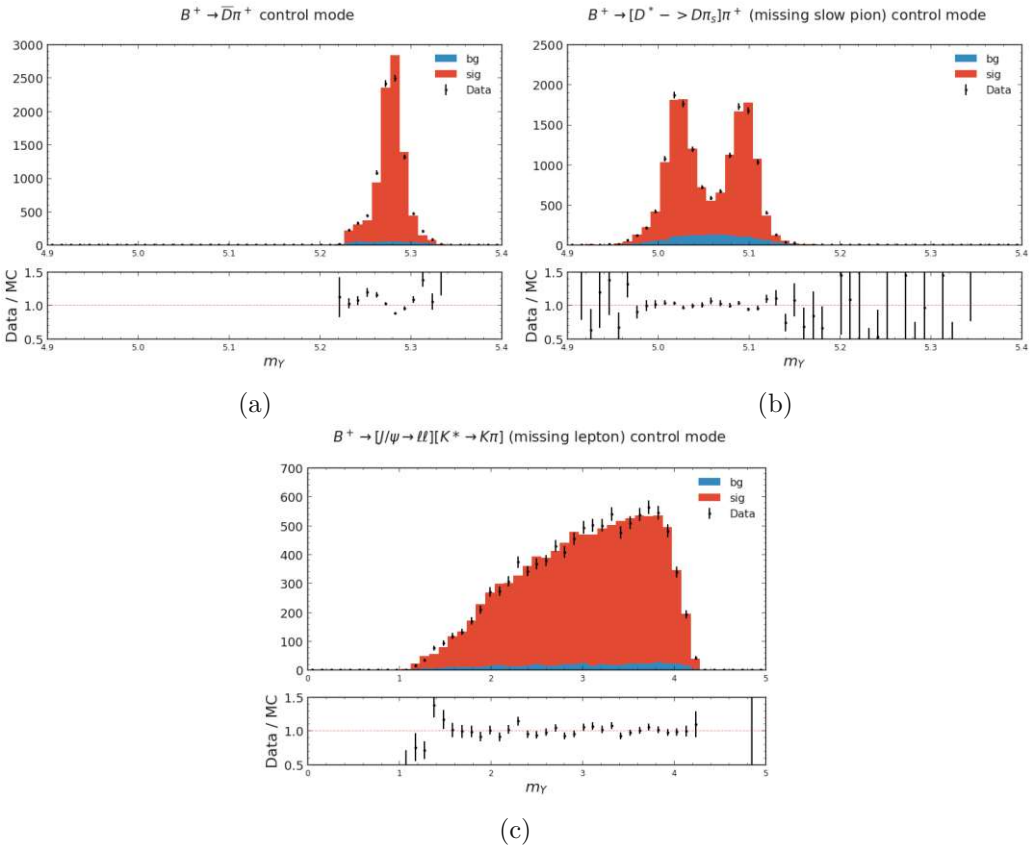


Figure 6.15: Distribution of the hadron-lepton system invariant mass in each control sample. The J/ψ mode has the highest amount of missing momentum and energy due to the missing second lepton in the reconstruction.

Hadronic $B \to D\pi$ decays

In fully hadronic $B \to D\pi$ decays, the signal side B meson can be reconstructed fully. Backgrounds can then be removed efficiently by placing selections on the *beam-constrained mass*

$$m_{bc} = \sqrt{E^{*2}_{\text{beam}} - \vec{p}^{*2}_B} \tag{6.12}$$

and the *energy difference*

$$\Delta E = E^*_B - E^*_{\text{beam}}, \tag{6.13}$$

where quantities indexed with B refer to the reconstructed signal B meson, and $E^*_{\text{beam}} = 5.29$ GeV is half of $\sqrt{s}$.
By treating the reconstructed π as the lepton in the event, the decay mirrors the track multiplicity observed in $B \to D\ell\nu_\ell$ and is therefore suitable for studying variables associated with tracking. On the other hand, there is no missing energy or momentum that mirrors the neutrino in $B \to D\ell\nu_\ell$.

$B \to D\pi$ from $B \to [D^* \to D\pi]\pi$ decays

In addition to the fully reconstructed $B \to D\pi$ sample, the other two control samples are based on a partial reconstruction approach. The final state is first fully reconstructed, and ΔE and m_{bc} selections are used to obtain a high purity. Subsequently, certain daughters of the full reconstruction are used to partially reconstruct the B meson, omitting one or more daughters in the reconstruction. With this approach, the partially reconstructed B meson has missing momentum and energy contributions similar to the $B \to D\ell\nu_\ell$ neutrino.
In the first partial reconstruction mode, $B \to [D^* \to D\pi_{\text{slow}}]\pi$ with a slow pion from the $D^* \to D$ decay are first fully reconstructed. Subsequently, the D meson and the pion from the B decay are selected to form a quasi-Y system. As shown in Figure 6.15 the missing slow pion causes the invariant mass of Y to have more spread than in the fully reconstructed $B \to D\pi$ sample, where any spread is only due to detector resolution.

$B \to K^*\ell$ from $B \to [K^* \to K\pi][J/\psi \to \ell\ell]$ decays

The second partial reconstruction mode is based on $B \to K^* J/\psi$ decays. In this case, the J/ψ is reconstructed from a pair of electrons or muons, and a mass window selection suppresses background contributions. The K^* meson is effectively treated as the D meson and is reconstructed in a $K^-\pi^+$ mode. Selections on the composite B meson include m_{bc}, ΔE and m_B cuts, and yield a sample with 95% purity.
In the partial reconstruction method, the second lepton is omitted, and the Y is directly reconstructed from one lepton and the K^* candidate. This leads to a significant amount of missing energy and momentum, as shown in Figure 6.15c, making this a suitable mode to validate quantities such as p_{miss}, or the hybrid reconstruction approach of $\cos\theta_{\ell,W}$. In contrast to the other modes, the sample contains a true lepton, allowing validation of *e.g.* the θ_ℓ distribution after applying a lepton identification selection.

6.6 External corrections

Before comparing the obtained simulated samples with collision data, known mismodellings must be corrected. These adjustments include calibration measurements by the Belle II performance groups, updated external inputs like branching ratios and form factors, and corrections calibrated specifically for this analysis using control samples.

Track momentum scaling

Precise knowledge of the magnetic field inside the detector is required to extract momentum measurements from charged particle trajectories. Imperfections in the magnetic field map then lead to a shifted momentum scale that has to be corrected with calibrated correction factors. The Belle II tracking performance group studies the overall scale of charged particle track momentum by comparing the mass peak in $D^0 \to K^-\pi^+$ decays [95, 96]. The obtained scale factors are computed independently for different run periods. The effect is taken into account by applying the scale to reconstructed momentum in data.

	Leptons	Kaons
Efficiency	$J/\psi \to \ell^+\ell^-$ $e^+e^- \to \ell^+\ell^-(\gamma)$ $e^+e^- \to e^+e^-\ell^+\ell^-$	$D^{*+} \to [D^0 \to K^-\pi^+]\pi^+$ $K^0_S \to \pi^+\pi^-$
Fake rates	$e^+e^- \to (\tau^+ \to h^\pm X)(\tau^- \to h^\pm h^\pm h^\pm X)$	$D^{*+} \to [D^0 \to K^-\pi^+]\pi^+$ $K^0_S \to \pi^+\pi^-$

Table 6.9: Calibration samples used to study deviations in particle identification performance between data and MC.

ParticleID corrections

The Belle II ParticleID group conducts studies, including the $ee \to ee(\gamma)$ study described in Chapter 5, to measure deviations in particle identification efficiencies between data and simulation. These deviations are corrected by applying a correction factor to the MC candidates. The scale factors are measured in bins of p and θ and are computed separately for positively and negatively charged particles. In this analysis, selection criteria are placed on electron, muon and kaon identification, each of which has to be corrected by applying the appropriate scale factors as a weight to each candidate based on its momentum, polar angle and charge.
The studies are divided into two types: measuring correction factors for efficiency, which involves correctly identified particles, and measuring fake rates, which involve tracks misidentified as different particle species. The calibration samples for leptons and kaons are listed in Table 6.9.

Photon energy and efficiency

Photons are not used directly in the $B \to D\ell\nu_\ell$ reconstruction but exert an effect through event shape and ROE variables. The measured photon energy from its ECL cluster may differ from the actual photon energy. A study conducted by the Belle II neutrals performance group on $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ decays measures the photon energy bias and provides correction factors binned by photon energy, ϕ, and θ. These correction factors are then applied to reconstructed photons in data.
In a separate neutrals performance study, the Belle II group measures differences in the photon reconstruction efficiency between data and MC using high-energy initial-state radiation photons in $e^+e^- \to \mu^+\mu^-\gamma_{ISR}$ decays. Calibration ratios are provided in bins of momentum, θ, and ϕ, and are used as a probability for randomly removing photons from MC events.

$B \to X_c\ell\nu$ branching ratios

In addition to corrections measured by performance groups, external parameters used in MC generation are updated to the most recent values. Branching ratios of semileptonic

Decay	$\mathcal{B}(B^+)$		$\mathcal{B}(B^0)$	
	Decay file	Updated	Decay file	Updated
$B \to D\ell\nu_\ell$	$2.31 \cdot 10^{-2}$	$(2.41 \pm 0.07) \cdot 10^{-2}$	$2.14 \cdot 10^{-2}$	$(2.24 \pm 0.07) \cdot 10^{-2}$
$B \to D^*\ell\nu_\ell$	$5.49 \cdot 10^{-2}$	$(5.50 \pm 0.12) \cdot 10^{-2}$	$5.11 \cdot 10^{-2}$	$(5.11 \pm 0.11) \cdot 10^{-2}$
$B \to D_1\ell\nu_\ell$	$7.57 \cdot 10^{-3}$	$(6.63 \pm 1.09) \cdot 10^{-3}$	$7.04 \cdot 10^{-3}$	$(6.16 \pm 1.01) \cdot 10^{-3}$
$B \to D_0^*\ell\nu_\ell$	$3.89 \cdot 10^{-3}$	$(4.20 \pm 0.75) \cdot 10^{-3}$	$3.62 \cdot 10^{-3}$	$(3.90 \pm 0.70) \cdot 10^{-3}$
$B \to D_1'\ell\nu_\ell$	$4.31 \cdot 10^{-3}$	$(4.20 \pm 0.90) \cdot 10^{-3}$	$4.01 \cdot 10^{-3}$	$(3.90 \pm 0.84) \cdot 10^{-3}$
$B \to D_2^*\ell\nu_\ell$	$3.73 \cdot 10^{-3}$	$(2.93 \pm 0.33) \cdot 10^{-3}$	$3.47 \cdot 10^{-3}$	$(2.73 \pm 0.30) \cdot 10^{-3}$
$B \to D\pi\pi\ell\nu_\ell$	$0.53 \cdot 10^{-3}$	$(0.62 \pm 0.89) \cdot 10^{-3}$	$0.49 \cdot 10^{-3}$	$(0.58 \pm 0.82) \cdot 10^{-3}$
$B \to D^*\pi\pi\ell\nu_\ell$	$2.63 \cdot 10^{-3}$	$(2.16 \pm 1.03) \cdot 10^{-3}$	$2.45 \cdot 10^{-3}$	$(2.00 \pm 0.95) \cdot 10^{-3}$
$B \to D_sK\ell\nu_\ell$	$0.3 \cdot 10^{-3}$	$(0.30 \pm 0.14) \cdot 10^{-3}$	0	0
$B \to D_s^*K\ell\nu_\ell$	$0.3 \cdot 10^{-3}$	$(0.29 \pm 0.19) \cdot 10^{-3}$	0	0
$B \to D\pi\ell\nu_\ell$	$1.50 \cdot 10^{-3}$	0	$1.38 \cdot 10^{-3}$	0
$B \to D^*\pi\ell\nu_\ell$	$1.50 \cdot 10^{-3}$	0	$1.38 \cdot 10^{-3}$	0
$B \to D\eta\ell\nu_\ell$	$2.01 \cdot 10^{-3}$	$(3.77 \pm 3.77) \cdot 10^{-3}$	$2.17 \cdot 10^{-3}$	$(4.09 \pm 4.09) \cdot 10^{-3}$
$B \to D^*\eta\ell\nu_\ell$	$2.01 \cdot 10^{-3}$	$(3.77 \pm 3.77) \cdot 10^{-3}$	$2.17 \cdot 10^{-3}$	$(4.09 \pm 4.09) \cdot 10^{-3}$

Table 6.10: Branching ratios of $X_c\ell\nu$ decays in the decay file and the reweighed values.

$b \to c$ decays are updated to the recent world averages [28] after computing values under the assumption of iso-spin average.
When summing up measured exclusive semileptonic $b \to c$ branching fractions, the obtained total branching fraction falls short of the measured inclusive branching ratio $\mathcal{B}(B \to X_c\ell\nu_\ell)$. This observed discrepancy is known as the *semileptonic gap* [97]. In this analysis, the gap is modelled with equal amounts of hitherto unmeasured $B \to D\eta\ell\nu$ and $B \to D^*\eta\ell\nu$ decays. The updated branching fractions are listed in Table 6.10.

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

A parameterization of $b \to c$ form factors and their parameters are assumed in MC generation to model the decay rate dependency on kinematic variables. The recent MC campaign uses the BGL parameterized form factor. After reconstruction, a reweighing of candidates may be applied to update the assumed parameters. The recent measurements of $B \to D^{(*)}\ell\nu$ form factors used for reweighing are:

- $D^*\ell\nu$: BGL parameterized form factor truncated at order $(n_a, n_b, n_c) = (1, 2, 1)$ from the Belle measurement in Ref. [98].
- $D\ell\nu$: BGL parameterized form factor truncated at $N = 3$ from Ref. [4].

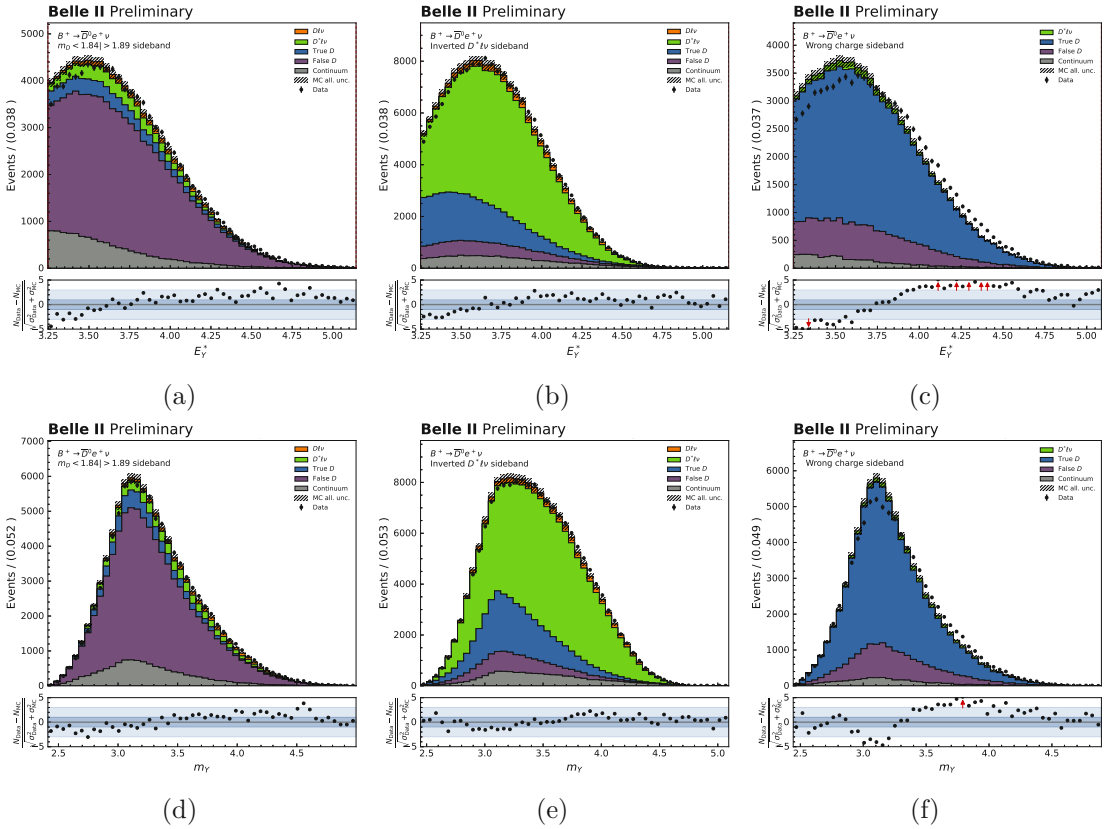


Figure 6.16: Data-MC agreements for the variables E_Y^* and m_Y in the control samples for the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode.

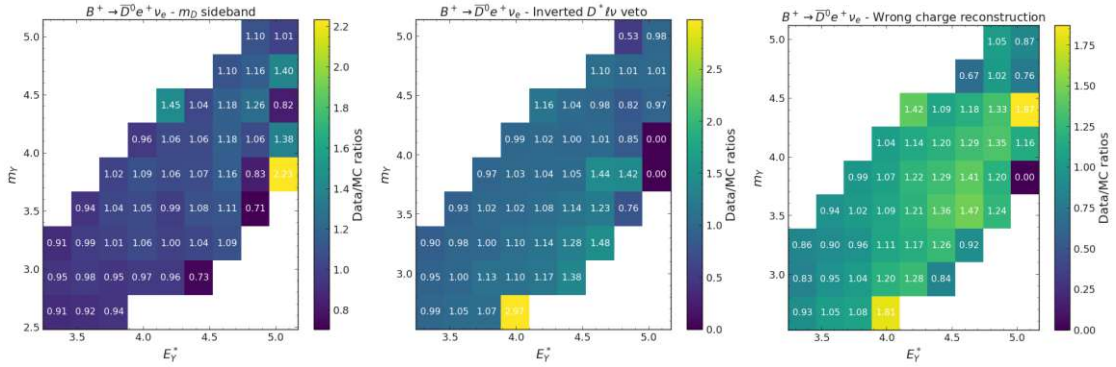
Correction factors depending on the generator-truth level w_{mc} of each candidate

$$\mathcal{G}^{nom}(w_{mc})/\mathcal{G}^{MC}(w_{mc}) \tag{6.14}$$

are applied, where $\mathcal{G}^{MC}$ and $\mathcal{G}^{nom}$ are the generator and updated form factors.

6.7 Control sample derived corrections and validation

After external corrections are implemented, distributions of key variables in control samples are compared between data and MC as a validation test. This comparison also helps determine if additional corrections are needed.
When comparing data with MC, the standard visualization involves colored stacked histograms for the simulations and black data points for comparison. Corrections are applied as weights to MC candidates, and the MC is rescaled according to the ratio of luminosities in the data samples. Uncertainties are represented by black hatched lines around the histograms for MC and error bars for data. The uncertainty for the MC



(a) (b) (c)

Figure 6.17: Correction factors derived from control samples in the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode.

includes systematic uncertainties from particle identification, branching ratio corrections, and luminosity. The lower part shows a distribution of pulls

$$\frac{N_{\text{data}} - N_{\text{MC}}}{\sqrt{\sigma_{\text{data}} + \sigma_{\text{MC}}}} \tag{6.15}$$

where the MC uncertainty includes statistical and systematic uncertainties, and the data uncertainty is only statistical.

Background $\cos\theta_{BY}$ shape

A noticeable disagreement between simulation and data is observed in the E_Y^*, m_Y and therefore also $\cos\theta_{BY}$, which is used as a fitting variable. Distributions are shown for control samples in the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode in the style of the above description in Figure 6.16. The disagreement is strongest in the wrong charge reconstruction sample enriched in true D background.
The mismodelling is corrected by individually calibrating correction factors in the $E_Y^* - m_Y$ plane for each of the three control samples shown in Figure 6.16. The obtained correction factors are displayed in Figure 6.17.
Applying the correction to the nominal $B \to D\ell\nu_\ell$ reconstruction significantly improves Data/MC agreement in m_Y and E_Y^*, but also in $\cos\theta_{BY}$. The distributions are shown before and after the correction in Figure 6.18.

Selection criteria in signal control samples

The signal control samples are compared to data to verify that the optimized selection criteria do not bias the selection efficiency due to differing distributions in Data and MC. Selected distributions in the $J/\psi \to e(e)$ sample, which most closely resembles $B \to D\ell\nu_\ell$ kinematically, are shown in Figure 6.19. The collision data aligns with expectations within the statistical uncertainties for most selection variables. The p_{ROE} and $\cos\theta_{\ell,W}$ diagrams

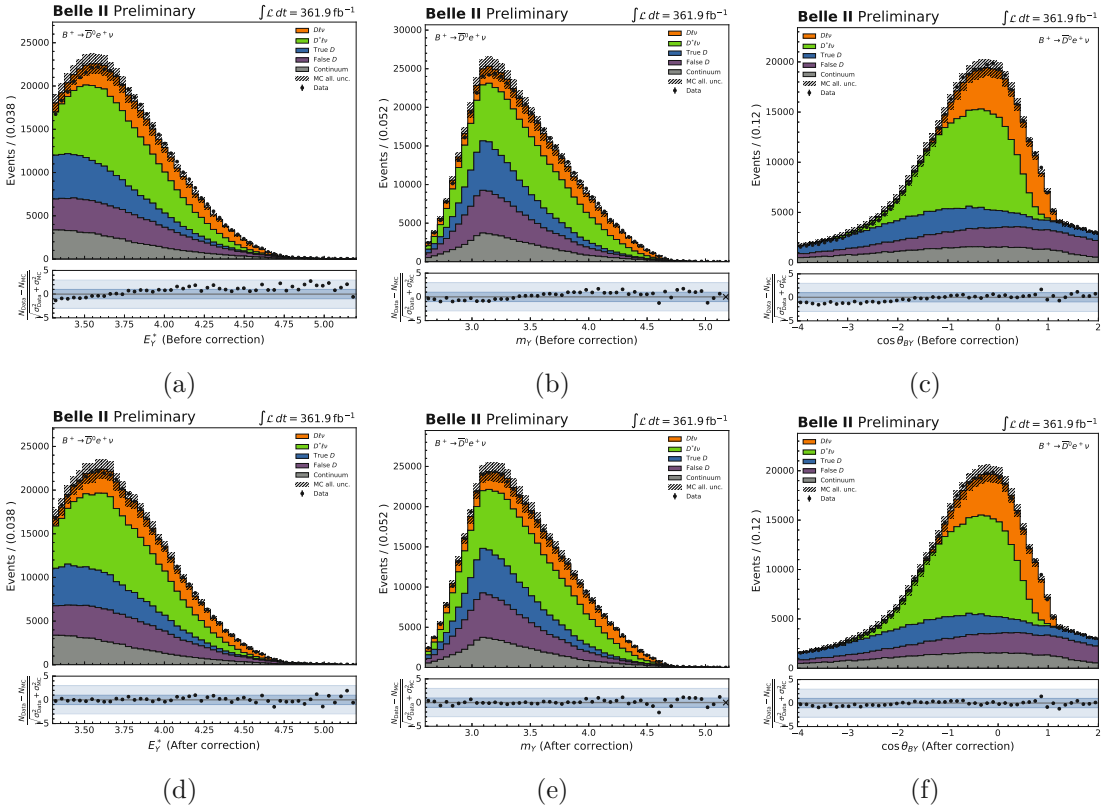


Figure 6.18: Data-MC agreements for the variables E_Y^*, m_Y and $\cos\theta_{BY}$ in the nominal $B^+ \to \bar{D}^0 e^+ \nu_e$ sample before (top) and after (bottom) applying the reweight correction.

	$B \to [J/\psi \to e(e)][K^* \to K\pi]$	$B \to [J/\psi \to \mu(\mu)][K^* \to K\pi]$
χ_{vertex}	1.000 ± 0.005	
θ_ℓ	1.004 ± 0.006	1.007 ± 0.006
E^*_{vis}	0.998 ± 0.002	0.998 ± 0.002
H^{so}_{02}	1.000 ± 0.000	1.001 ± 0.000
E^*_Y	0.997 ± 0.010	1.002 ± 0.009
p^*_{miss}	0.983 ± 0.015	0.978 ± 0.014
m_{ROE}	0.996 ± 0.004	1.002 ± 0.004
$\cos\theta_{\ell,W}$	1.006 ± 0.008	1.000 ± 0.008
p_{ROE}	0.998 ± 0.013	0.997 ± 0.012
$\theta_{D,\ell}$	1.000 ± 0.000	1.000 ± 0.000
p^*_ℓ		1.000 ± 0.000

Table 6.11: Data/MC of selection efficiencies in the J/ψ control sample. We drop the selection on p^*_{miss} due to different efficiencies observed in Data and MC.

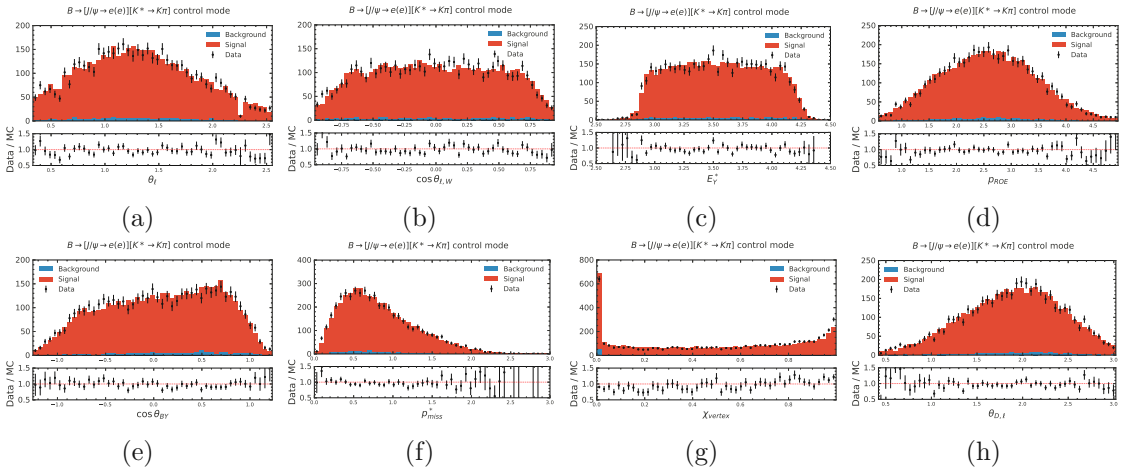
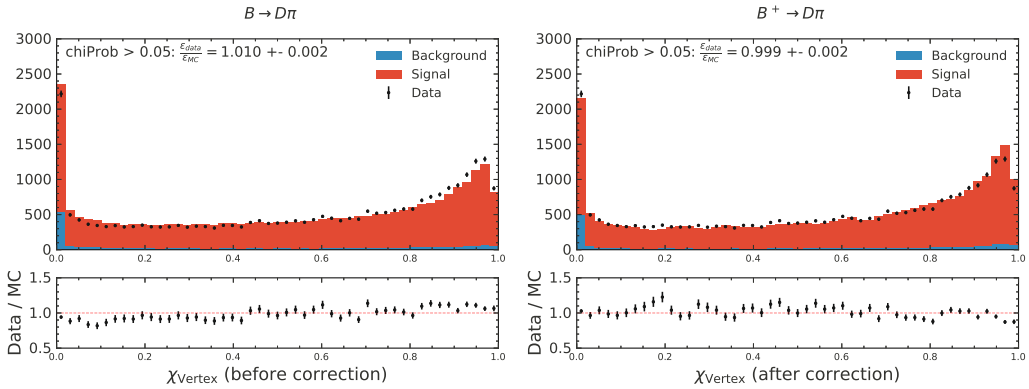


Figure 6.19: Distributions of selection variables in the $B \to J/\psi K^*$ control sample with electrons in data and MC. Uncertainties are only statistical.

confirm that the construction of the ROE and the hybrid reconstruction approach are valid and do not introduce significant biases.
Two exceptions are the missing momentum of the event p_{miss} and the vertex fit χ^2 probability distribution. The selection on p_{miss} has a minor impact on the overall statistical significance and is therefore dropped following this observation. The selection on the vertex fit probability is crucial for suppressing combinatorial backgrounds, particularly in the neutral B mode, and is therefore corrected for.
The observed trend is consistent between all three signal control samples described in Section 6.5.2. To correct it, scale factors are computed in bins of vertex fit probability based on the $B \to J/\psi K^*$ control sample. By applying the correction factors to the $B \to D\pi$ control sample and successfully correcting the selection efficiency as shown in Figure 6.20, it is verified that the correction tracks among different decays. Finally, the correction is applied to the signal distribution in the nominal sample.
After applying corrections for the vertex fit mismodelling and removing the selection on missing momentum, the control sample efficiency of the remaining selections is compatible with the expectations within uncertainties. An overview of selection efficiency ratios $\frac{\varepsilon_{\text{Data}}}{\varepsilon_{\text{MC}}}$ is given in Table 6.11

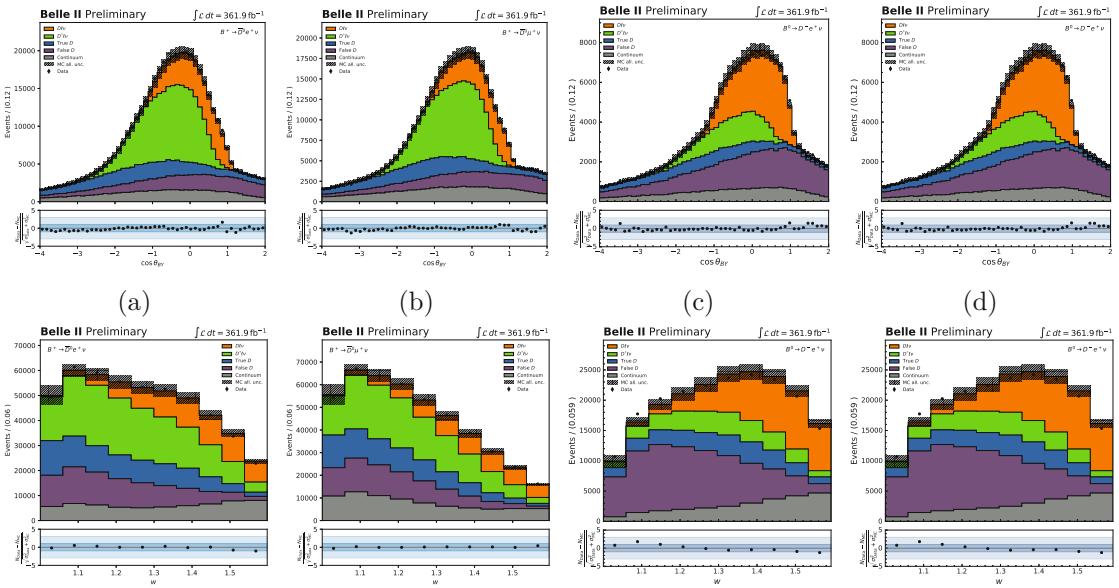
6.8 Comparison between Data and MC

After applying all corrections, the nominal reconstruction samples can be compared to collision data. The signal extraction is performed by fitting the $\cos\theta_{BY}$ distribution. A comparison of data and MC distributions, using the visualization style described in Section 6.7, of $\cos\theta_{BY}$, w, lepton momentum, polar angle, and $\cos\theta_{\ell,W}$ is shown in Figures 6.21 and 6.22. The diagrams are shown with all selections applied and before performing a fit to the data (*pre-fit*). Additional comparison diagrams are shown in



(a) (b)

Figure 6.20: Vertex fit probability distribution in the $B \to D\pi$ control sample before and after applying the correction calibrated on $B \to J/\psi K^*$ decays. The selection efficiency ratio improves to move within the statistical uncertainty of unity.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 6.21: $\cos\theta_{BY}$ and w distributions in data and MC for all four modes.

Appendix C.

In the distributions of w, additional background normalization uncertainties are assigned based on control samples. The procedure to derive these uncertainties is described in the later section Section 6.9.2.

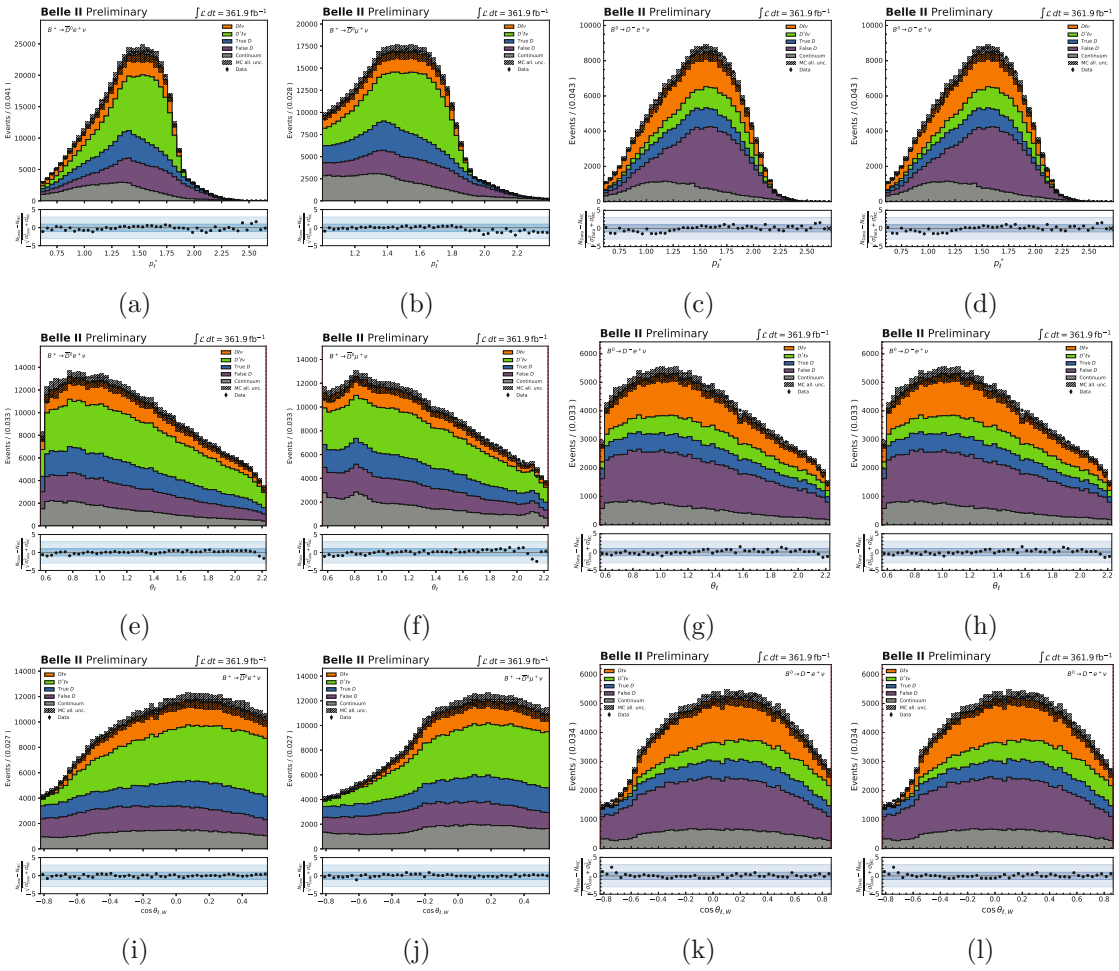


Figure 6.22: p^*_ℓ, θ_ℓ and $\cos\theta_{\ell,W}$ distributions in data and MC for all four modes.

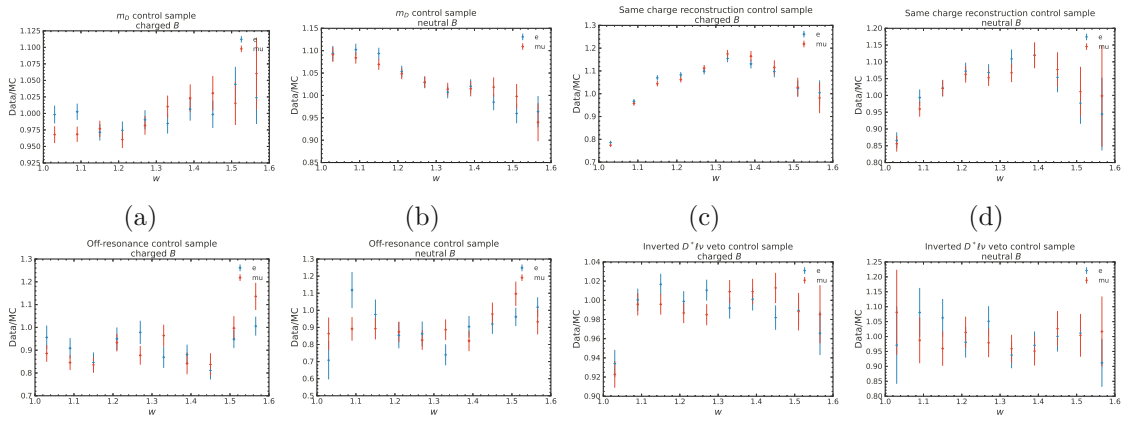


Figure 6.23: Data-MC ratios in bins of w in the background control samples.

6.9 Signal extraction

The determination of $|V_{cb}|$ is derived from the differential decay rates $\Delta\Gamma_i/\Delta w$. These differential decay rates are computed from the partial branching ratios

$$\frac{\Delta\Gamma_i}{\Delta w} = \frac{\Delta\mathcal{B}_i\left(B \to D\ell\nu_\ell\right)}{\tau_B \Delta w} \ , \quad i = 0, \ldots, 9 \ , \tag{6.16}$$

which in turn depend on the signal yields in bins of w

$$\Delta\mathcal{B}_i\left(B \to D\ell\nu_\ell\right) = \frac{N_{\mathrm{sig},i}}{2N_{B\overline{B}}\,\epsilon_i\,\mathcal{B}\left(D^{0(+)} \to K^-\pi^+(\pi^+)\right)}. \tag{6.17}$$

Here, τ_B are the lifetimes of charged and neutral B mesons; $N_{B\overline{B}}$ is the number of B meson pairs in the data sample; $N_{\mathrm{sig},i}$ denotes the signal yields in each w bin i; ϵ_i are the bin-wise selection efficiencies; and $\mathcal{B}\left(D^0 \to K^-\pi^+(\pi^+)\right)$ are the branching ratios of the involved D decays.

High-precision determination of each signal yield directly corresponds to a precise measurement of $|V_{cb}|$. The signal extraction strategy in this analysis employs a binned maximum likelihood fit using templates derived from simulation. Details of the statistical model used in the fit are provided in Section 4.2. Systematic uncertainties are incorporated directly into the fit by using nuisance parameters that adjust the shape and normalization of the templates, which are then reflected in the fitted signal yields. The fit is set up and executed using the `pyhf` library [82], a Python adaptation of `HistFactory` [81].

6.9.1 Template fit setup

A one-dimensional fit to the $\cos\theta_{BY}$ distribution is performed across ten bins of w and in all four reconstructed modes, yielding a total of $10 \times 4 = 40$ signal windows. To improve constraints on background components and to obtain systematic correlations between

w bin	$\cos\theta_{BY}$ bin edges	n_{Bins}
(1.0, 1.06)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.06, 1.12)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.12, 1.18)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.18, 1.24)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.24, 1.3)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.3, 1.36)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.36, 1.42)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.42, 1.48)	[-4.0, -3.6, -3.2, -2.8, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	15
(1.48, 1.54)	[-4.0, -2.4, -2.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	12
(1.54, $w_{\max}$)	[-4.0, -1.6, -1.2, -0.8, -0.4, 0.0, 0.4, 0.8, 1.2, 1.6, 2.0]	10

Table 6.12: $\cos\theta_{BY}$ binning is determined by merging the lowest bins in the presence of empty bins.

signal yields, the fit is conducted simultaneously across all 40 windows. The categories introduced in Section 6.4.2 are used as separate templates, each with parameters that control and vary the normalization and shape. The normalization of signal templates is left free-floating in order to obtain a measurement. For background templates, a single unconstrained normalization parameter applies across all w bins, while per-bin Gaussian nuisance parameters manage deviations from simulation in the w distribution. The width of the Gaussian distributed nuisance parameters is derived from control samples, as outlined in Section 6.9.2. omparison of bin-by-bin normalizations in background control samples indicates good agreement between e and μ modes, as shown in Figure 6.23, which justifies the use of joint bin-by-bin parameters.

Binning

In the fit, each channel consists of ten $\cos\theta_{BY}$ distributions with 15 bins each for a total of 150 bins per mode and 600 bins total.
The $\cos\theta_{BY}$ fit range is from -4 to 2, with the signal being constrained from -1 to 1, up to some smearing due to detector resolution. In higher w bins, the background distributions shift towards higher values of $\cos\theta_{BY}$. If any bins are empty, the lowest two bins are merged, and the process is iteratively repeated until every bin contains at least 50 events. After this merging process, the total number of bins is reduced from 600 to 568. The final bin edges are listed in Table 6.12.

Unfolding

Due to detector resolution effects, reconstructed values of w deviate from the true values, subjecting the measured spectra to smearing across bins. The process of converting the measured to the underlying spectra is called *unfolding*. Using the migration matrix, as introduced in Section 6.4.6, allows unfolding the fitted signal yields by accounting for candidates migrating from one bin to another. Various methods for unfolding exist, including bin-by-bin unfolding [99], SVD unfolding [100], matrix inversion and iterative

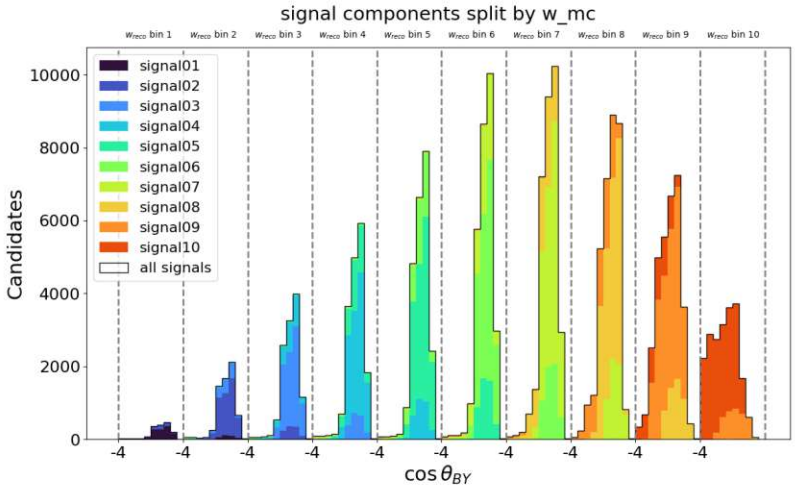


Figure 6.24: The signal templates $\cos\theta_{BY}$ distributions in each bin of w flattened onto one combined axis. By categorizing the signal templates using w_{MC} rather than w_{reco}, migration effects between w bins are directly mapped into the fit.

Bayesian unfolding [101]. For instance, matrix inversion derives the true spectrum by directly inverting the migration matrix. In SVD unfolding, singular value decomposition decomposes the migration matrix in its singular values and vectors, and regularization smoothes out statistical fluctuations.
This thesis introduces a novel unfolding method that incorporates migration effects directly into the template fit. While the reconstructed value w_{reco} is used to define the overall bin counts of the signal (and background) histograms, the bin-wise signal templates are split using w_{MC} at the generator level. In this way, each signal template has contributions primarily in its nominal w bin, with smaller migration components in neighbouring bins. Nuisance parameters that vary the shape of signal templates consequently affect the magnitude of bin-by-bin migrations. As a result, the fit yields are obtained in terms of values that are already unfolded. A schematic of the signal template definitions in this *direct* unfolding approach is shown in Figure 6.24.

6.9.2 Systematic uncertainties

Systematic uncertainties are incorporated into the templates through nuisance parameters. An overview of the available modifiers in `pyhf` and their corresponding constraint terms is shown in Table 4.2. *normsys* modifiers allow normalization variations and *histosys* for correlated shape variations. Additionally, *staterror* is used for the statistical uncertainty from the finite MC sample size and *luminosity* for the luminosity uncertainty.
Before describing the implementation of each source of systematic uncertainty in more detail, an overview of implemented nuisance parameters is given in Table 6.13.

Nuisance parameter	Implementation	Type	# of parameters
LeptonID	Generate toys, eigendecompose covariance matrix	normsys	5
KaonID	Generate toys, eigendecompose covariance matrix	normsys	5
$B \to X_c \ell\nu$ branching ratios	24 decay modes as individiual parameters correlated between modes	histosys	24
$B \to D^{(*)}\ell\nu$ form factors	Variations in eigendirections 8 nuisance paramters for $D\ell\nu$ 5 for $D^*\ell\nu$	histosys	13
Background $\cos\theta_{BY}$ shapes	Correlated shape uncertainties	histosys	12
Background w shapes	Normalization uncertainties	normsys	20
Tracking uncertainty	0.24% per track Correlated between categories and modes	normsys	1
Signal efficiency	From vertex χ correction	normsys	1
D branching ratios	Separate for B modes	normsys	2
B lifetime	Correlated between categories and modes	normsys	2
N_{bb}	Correlated between categories and modes	luminosity	1
f_{00}/f_{+-}	correlated between categories Anti-correlated between B modes	normsys	1
MC Stat. Error	Uncorrelated between all bins	staterror	568

Table 6.13: An overview of the implemented nuisance parameters that assess the systematic uncertainties in the template fit.

Form factors

The form factor used in MC generation is reweighed as described in Section 6.6. The systematic uncertainty arising from the effect of the form factor on the signal template shape is assessed by utilizing the covariance matrix from the recent reweighted measurements. An analogous approach is applied to $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$.
The covariance matrix between parameters parameterizing the BGL form factor is decomposed into its eigendirections. Varying the input parameters in each eigendirection by $\pm 1\sqrt{\lambda_i}\vec{v_i}$, with eigenvalues λ and eigenvectors $\vec{v_i}$ gives rise to new sets of BGL parameters. Reweighing the $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ component histograms according to the form factor constructed with these parameters results in shape variations in the templates. These variations correspond to shifts along the eigendirections, with positive and negative shifts representing the up- and down-variations, respectively. One nuisance parameter for each eigendirection then controls the correlated shape modifier. The form factors from external inputs are truncated at $N = 3$ for $B \to D\ell\nu_\ell$ and $(n_a, n_b, n_c) = (1, 2, 1)$ for $B \to D^*\ell\nu_\ell$, yielding 8 and 5 eigenvariations respectively, for a total of 13 nuisance parameters.
The signal templates in the fit directly model the migration effects by classifying them using w_{MC}. Therefore, systematic uncertainties affecting the w distribution directly impact the

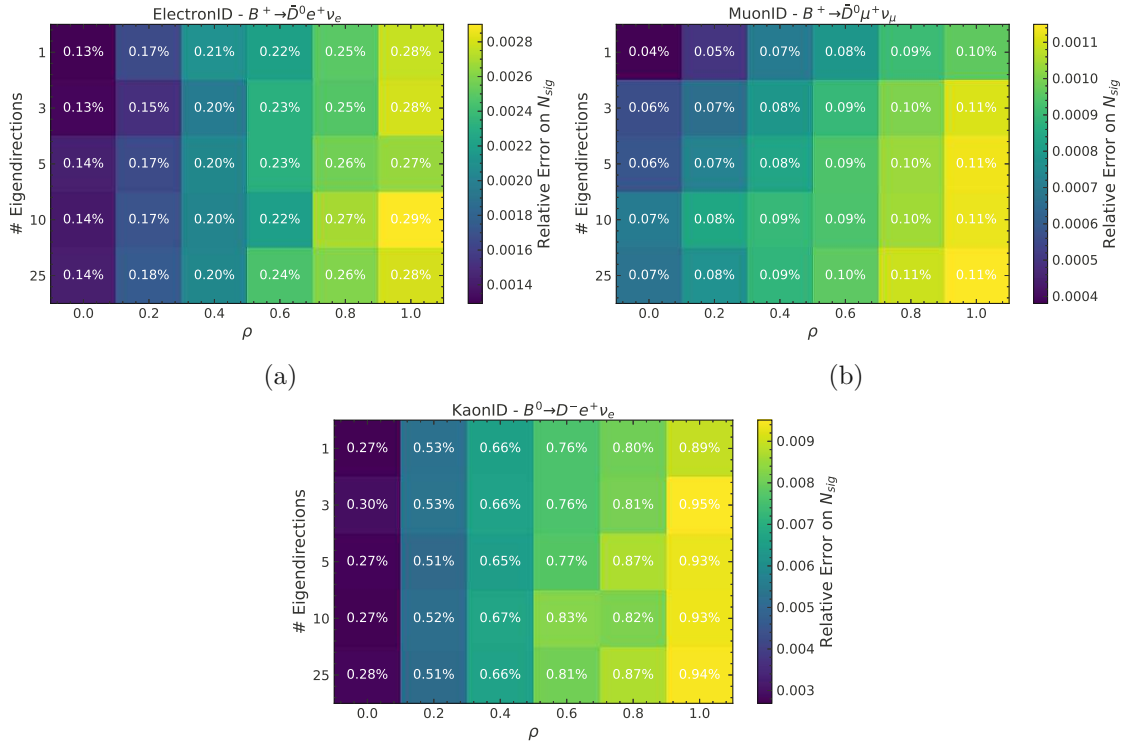


Figure 6.25: The uncertainty from particleID on the total signal yield N_{sig} as a function of assumed correlation coefficient ρ between systematic uncertainties bins in the correction tables, and a truncation order of the eigendecomposition that models the uncertainty. Full correlation $\rho = 1$ shows the largest uncertainty for all particles.

size of the migration effects. Consequently, the form factor uncertainty for $B \to D\ell\nu_\ell$ has a more significant impact on the differential decay rates than a more traditional unfolding approach, where the effect from the unfolding method is assessed independently.

$B \to X_c\ell\nu$ branching ratios

After reweighing each $B \to X_c\ell\nu_\ell$ branching ratio to its updated value according to the values in Table 6.10, each decay mode can also vary within its uncertainty. This is implemented in the fit as a *histosys* modifier for each of the 24 decays, applied to the $B\overline{B}$ background templates. The nuisance parameters controlling these modifiers are correlated across different decay modes and w bins, resulting in a total of 24 independent nuisance parameters.

Particle identification

The correction factors for particleID provided by the Belle II performance groups come with associated statistical and systematic uncertainties for each bin of p and θ. The correlation

of the systematic uncertainty between these bins is unknown, requiring assumptions to model these uncertainties.
To address this, toy sets of correction factors are generated by resampling from the nominal values with an assumed systematic correlation matrix ρ_{sys} between bins. Statistical uncertainties are assumed to be uncorrelated. For candidates not covered by the provided tables, such as protons misidentified as kaons, a weight of 1 with 100% uncertainty is assigned.
The toy correction factors are applied to the $B \to D\ell\nu_\ell$ candidates, and the appropriate weights are propagated into $\cos\theta_{BY}$ histograms. The covariance matrix between bins of $\cos\theta_{BY}$ for each template and decay mode is decomposed into eigenvalues and eigenvectors. Each eigenvector with associated eigenvalue is then modelled as a *histosys* modifier controlled by one nuisance parameter, as described in Section 4.3.2. Including all such modifiers is computationally expensive, so approximations are necessary. Given that eigenvalues decrease significantly, it is generally sufficient to truncate after N eigenvalues.
To determine the most conservative approach, the sensitivity on the total signal yield is examined with varying correlation coefficients ρ_{sys} and truncation orders N. The resulting uncertainties are shown in Figure 6.25. The most conservative assumption for the systematic correlation coefficient is found to be $\rho_{\text{sys}} = 1$. The uncertainty does not increase significantly after adding more than two variations. The first five variations are included for LeptonID and KaonID to ensure a conservative estimate.

Signal efficiency

Differences between data and MC distributions in cut variables can bias the selection efficiency. The data/MC discrepancy in the vertex fit probability was described and corrected in Section 6.7. To account for statistical fluctuations in the control sample-derived correction, a systematic uncertainty is assessed by resampling the applied correction factors within their uncertainties. The impact on the signal efficiency from resampled correction factors is a relative uncertainty of 0.5%. The uncertainty is modelled with a *normsys* modifier controlled by one nuisance parameter fully correlated between w bins and decay modes. Remaining selections other than the vertex fit probability do not show significant biases of the efficiency, as shown in Table 6.11.

Background w shapes

The per-bin normalizations in w for the different background components can be constrained using the background control samples. As described in Section 6.5.1, the sample composition of the true D component differs in the nominal and the wrong charge reconstruction. Therefore, the normalization of this component is left unconstrained, as the overall normalization can not be accurately obtained. The normalization of the other backgrounds is obtained by performing individual binned likelihood fits to the w distribution for the control samples enriched in continuum, $B \to D^*\ell\nu_\ell$ and false D backgrounds.
An unconstrained shape uncertainty is added to the background component of interest if the fit result has a p value lower than 0.5 to account for shape deviations between data and MC. This uncertainty is increased iteratively until the fit p value exceeds 0.5. Pre-

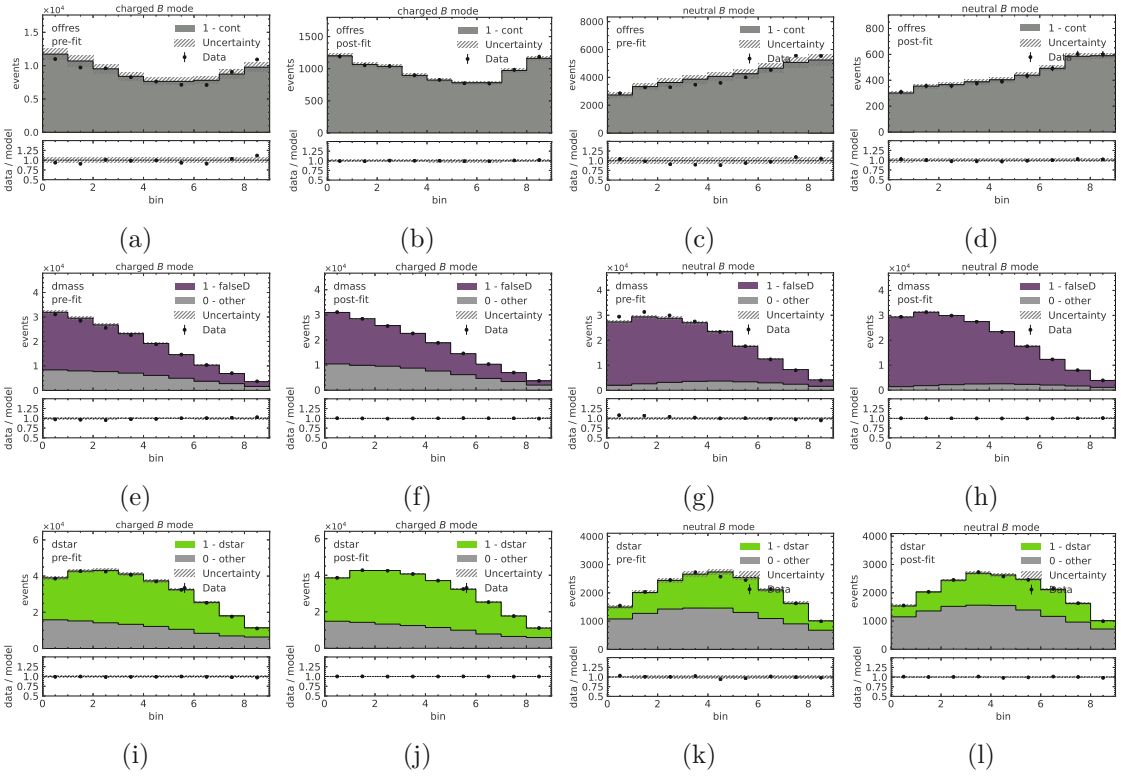


Figure 6.26: Pre- and post-fit plots for the fits to the background components in the control samples. The fit is performed separately for each control sample and for neutral and charged B modes.

		True D	Fake D	$B \to D^*\ell\nu$	Continuum
$B^+ \to \bar{D}^0\ell^+\nu_\ell$	Unconstrained	3.2%	2.4%	2.6%
$B^0 \to D^-\ell^+\nu_\ell$	Unconstrained	2.0%	10.6%	2.9%

Table 6.14: The resulting uncertainties on background components obtained from a fit to control samples.

and post-fit plots of the fits to control samples for all control sample fits are shown in Figure 6.26. The resulting Gaussian uncertainties are shown in Table 6.14. The resulting uncertainties are then used as Gaussian widths on the per-bin normalization modifiers for the respective background components. Larger discrepancies are observed for the lowest bin of w, and an uncertainty of 30% is assigned to cover all deviations.

Background $\cos\theta_{BY}$ shapes

The correction of background shapes by using a reweighing in E_Y^* and m_Y is described in Section 6.7. The calibrated weight factors are assigned in full as up- and down-variations on

the reweighed templates to assess the systematic uncertainty associated with this correction. This is implemented as individual *histosys* modifiers for each mode and background template, controlled by one nuisance parameter each.

Tracking efficiency

The Belle II tracking group studies the Data/MC ratios of track efficiencies for medium-to-high momentum charged tracks. A flat uncertainty of 0.24% per charged final state track is determined, resulting in a total uncertainty of 0.72% for the charged B modes and 0.96% for the neutral B modes. The tracking uncertainty is fully correlated over w bins and templates, and included as a single nuisance parameter.

Luminosity and f_{00}/f_{+-}

The amount of B meson pairs recorded in the full run 1 data sample is evaluated by the Belle II performance group in a study using Bhabha decays. Because experiments 7 and 8 are excluded, as motivated in Section 6.2, $N_{B\overline{B}}$ is scaled down using run period-wise integrated luminosities, while the uncertainty is kept in full. The uncertainty is included with a single, fully correlated nuisance parameter.
The share of charged over neutral B meson pair production in $\Upsilon(4S)$ is described by the parameter

$$R^{+-/00} = \frac{f^{+-}}{f^{00}} = \frac{\Gamma(\Upsilon(4S) \to B^+B^-)}{\Gamma(\Upsilon(4S) \to B^0B^0)} = 1.065 \pm 0.052, \tag{6.18}$$

measured at Belle [102]. The signal yields are sensitive to $f^{+-}/(f^{00}+f^{+-})$ and $f^{00}/(f^{00}+f^{+-})$, rather than the ratio directly. The uncertainty on these factors is obtained using Gaussian error propagation. With $\frac{f^{+-}}{f^{00}+f^{+-}} = \frac{R^{+-/00}}{1+R^{+-/00}}$ the uncertainty becomes

$$\sigma_{f^{+-}} = \frac{\partial f}{\partial R^{+-/00}}\sigma_{R^{+-/00}} = \frac{1}{(1+R^{+-/00})^2}\sigma_{R^{+-/00}}, \tag{6.19}$$

resulting in an uncertainty of 0.0121 on f^{00} and f^{+-}. Under the assumption of $f^{00}+f^{+-} = 1$, the uncertainty on f^{00} and f^{+-} is modelled with one nuisance parameter controlling a *normsys* modifier, fully anti-correlated between charged and neutral B modes. This leads to partial cancellation of this uncertainty when averaging across B modes.

$D^{0(+)} \to K^-\pi^+(\pi^+)$ branching ratios

D mesons are reconstructed in a single mode for each of charged and neutral B mesons, and its branching ratio has to be divided out when converting the signal yields to the $B \to D\ell\nu_\ell$ branching ratios, see Equation (6.17). The current precision with which the D branching ratios are known thus affects the results for the branching ratios and decay rates. The uncertainties on the current world averages [28]

$$\mathcal{B}\left(D^0 \to K^-\pi^+\right) = (3.95 \pm 0.03)\%, \tag{6.20}$$

and

$$\mathcal{B}\left(D^+ \to K^-\pi^+\pi^+\right) = (9.38 \pm 0.16)\%, \tag{6.21}$$

enter the result as relative uncertainties of 0.8% and 1.7% for charged and neutral B modes respectively. The uncertainties are included independently for the types of B modes but fully correlated across w bins and lepton modes for a total of two nuisance parameters controlling *normsys* modifiers.

B lifetime

The lifetimes of charged and neutral B mesons enter the computation of $\Delta\Gamma_i/\Delta w$ via Equation (6.16). The current world averages [3], dominated by measurements from LHCb [103] and CMS [104], are

$$\tau_{B^0} = 1.519 \pm 0.004 \text{ ps} \tag{6.22}$$

and

$$\tau_{B^\pm} = 1.638 \pm 0.004 \text{ ps}. \tag{6.23}$$

The relative uncertainties on τ_{B^0} and $\tau_{B^\pm}$ of 0.3% and 0.2% respectively are applied as normalization uncertainty on signal templates, correlated between e and μ modes.

MC statistical uncertainty

The simulated MC sample is produced at integrated luminosities four times the data luminosity. Under the assumption of Poisson distributions, where uncertainties scale with the square root of the number of events $\sqrt{N}$, the size of statistical fluctuations of simulated samples is approximately half that of the data. The *staterror* modifier implemented in `pyhf` is used to model these statistical fluctuations of the templates. It assigns one uncorrelated nuisance parameter per bin, allowing the total bin count to vary. A Gaussian width is computed for each bin based on the individual uncertainties of all fit templates.

6.9.3 Recovering the MC expectation

Signal yields in simulation can be recovered by fitting the MC data onto itself, by using *pseudo-data*, which involves treating the total number of candidates from the MC as if it were real data. Fit bias can be tested by checking whether all signal strengths are fitted to the underlying ground truth in this so-called *Asimov* fit. The resulting signal strengths are listed in Table 6.5, where a value of one corresponds to exactly the MC expectation. The test is conclusive, as all the signal strengths are consistent with unity.
Another test involves summing the obtained signal yields over bins of w while accounting for correlations to recover the branching ratios used in the MC generation. The recovered branching ratios with the generator values are presented in Table 6.16.
Furthermore, the resulting 40×40 correlation matrix between fitted signal strengths from the Asimov fit is shown in Figure 6.27. Correlations are stronger within the same charge B meson modes due to the coupled background normalizations and the asymmetrical

	w bin	$B^+ \to \bar{D}^0 e^+ \nu_e$	$B^+ \to \bar{D}^0 \mu^+ \nu_\mu$	$B^0 \to D^- e^+ \nu_e$	$B^0 \to D^- \mu^+ \nu_\mu$
Signal1	$[1.0, 1.06]$	1.000 ± 0.770	1.000 ± 0.728	1.000 ± 1.399	1.000 ± 1.362
Signal2	$[1.06, 1.12]$	1.000 ± 0.177	1.000 ± 0.164	1.000 ± 0.288	1.000 ± 0.280
Signal3	$[1.12, 1.18]$	1.000 ± 0.091	1.000 ± 0.092	1.000 ± 0.141	1.000 ± 0.136
Signal4	$[1.18, 1.24]$	1.000 ± 0.065	1.000 ± 0.066	1.000 ± 0.091	1.000 ± 0.091
Signal5	$[1.24, 1.30]$	1.000 ± 0.054	1.000 ± 0.055	1.000 ± 0.068	1.000 ± 0.071
Signal6	$[1.30, 1.36]$	1.000 ± 0.048	1.000 ± 0.050	1.000 ± 0.058	1.000 ± 0.061
Signal7	$[1.36, 1.42]$	1.000 ± 0.044	1.000 ± 0.047	1.000 ± 0.053	1.000 ± 0.056
Signal8	$[1.42, 1.48]$	1.000 ± 0.042	1.000 ± 0.046	1.000 ± 0.049	1.000 ± 0.053
Signal9	$[1.48, 1.54]$	1.000 ± 0.045	1.000 ± 0.048	1.000 ± 0.049	1.000 ± 0.053
Signal10	$[1.54, w_{max}]$	1.000 ± 0.055	1.000 ± 0.056	1.000 ± 0.057	1.000 ± 0.058

Table 6.15: Fitted signal strength and total statistic plus systematic uncertainty when fitting the MC templates to itself. All fitted strengths are equal to one, equivalent to the MC expectation.

	Fitted value	Generator value
$B^+ \to \bar{D}^0 e^+ \nu_e$	$2.30\% \pm 0.08\%$	2.31%
$B^+ \to \bar{D}^0 \mu^+ \nu_\mu$	$2.31\% \pm 0.09\%$	2.31%
$B^0 \to D^- e^+ \nu_e$	$2.14\% \pm 0.09\%$	2.14%
$B^0 \to D^- \mu^+ \nu_\mu$	$2.14\% \pm 0.10\%$	2.14%

Table 6.16: The resulting total branching ratios for each of the four modes, obtained by summing the signal yields while accounting for the correlation matrix, are consistent with the generator-level branching ratios.

uncertainty on f^{+-}/f^{00}. Elements one off the diagonal show lower correlations than their near neighbours due to the included migration effects in the templates.

6.10 Form factor fit

The fitted signal yields can be converted to differential decay rates using Equations (6.16) and (6.17). Before fitting the form factor to the obtained rates, these rates are averaged across decay modes using a resampling approach incorporating the full covariance matrix from the fit. Specifically, the signal yields are resampled 10,000 times using the 40×40 covariance matrix of all signal yields. By calculating the decay rates for each resampled set and then taking the mean and standard deviation across all samples, the correlations from the fit are fully propagated into the averaged $\Delta\Gamma_i/\Delta w$ spectrum.
To determine the CKM matrix element $|V_{cb}|$ from the averaged rates, a fit to the BCL form factor introduced in Section 2.3 is performed. The fitted form factor is truncated at order $N = 3$ to align with the theoretical inputs from the FLAG collaboration, which provides lattice QCD averages [40]. The averages incorporate results from the MILC [38] and HPQCD groups [39]. After truncating the BCL expansion at $N = 3$, FLAG provides values

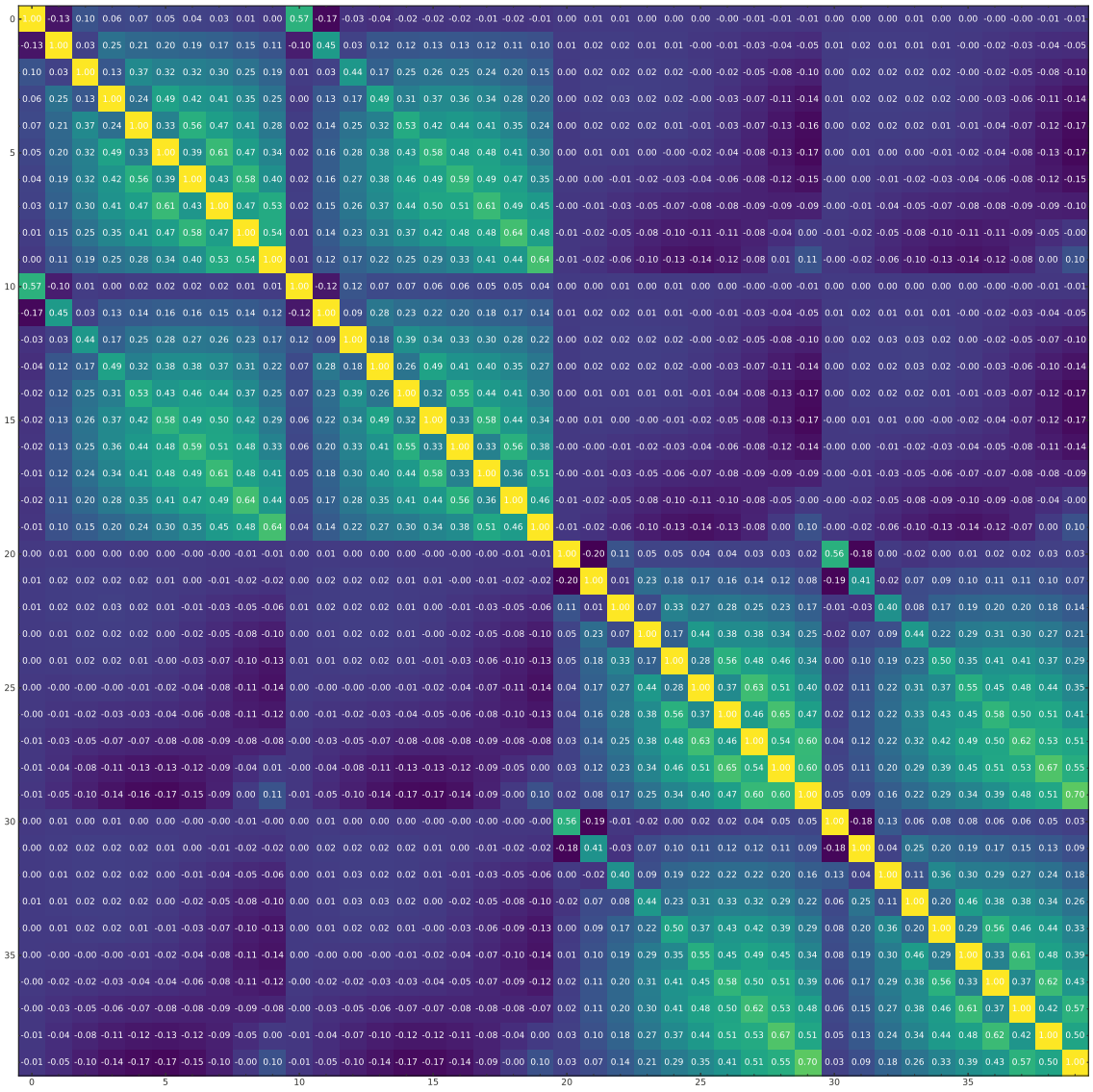


Figure 6.27: The combined statistical and systematic correlation matrix between the 40 signal strengths. The indices correspond to 0-9 $B^+ \to \bar{D}^0 e^+ \nu_e$, 10-19 $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$, 20-29 $B^0 \to D^- e^+ \nu_e$ and 30-39 $B^0 \to D^- \mu^+ \nu_\mu$,

of the five remaining BCL parameters $a^{\mathrm{FLAG}} = (a_0^+, a_1^+, a_2^+, a_0^0, a_1^0)$ with their correlations. The form factor fit is performed by minimizing

$$\begin{aligned}\chi^2 &= \sum_{i,j}^{10}\left(\frac{\Delta\Gamma_i}{\Delta w} - \frac{\Delta\Gamma_{i,\mathrm{BCL}}}{\Delta w}\right)\mathbf{C}_{ij}^{-1}\left(\frac{\Delta\Gamma_j}{\Delta w} - \frac{\Delta\Gamma_{j,\mathrm{BCL}}}{\Delta w}\right) \\ &+ \sum_{k,l}^{5}\left(a_k^{\mathrm{FLAG}} - a_k^{\mathrm{BCL}}\right)\mathbf{D}_{kl}^{-1}\left(a_l^{\mathrm{FLAG}} - a_l^{\mathrm{BCL}}\right)\end{aligned} \tag{6.24}$$

with the differential decay rates determined from the template fit $\Delta\Gamma_i/\Delta w$ and the partial widths calculated from the fit parameters $\frac{\Delta\Gamma_{i,\mathrm{BCL}}}{\Delta w}$ by using Equations (2.44), (2.51), (2.52), (2.57) and (2.58). The covariance matrix $\mathbf{C}$ includes the statistical and systematic uncertainties in the measurements of $\Delta\Gamma_i/\Delta w$. In the second term, a_k^{BCL} are the five BCL parameters also used to calculate $\frac{\Delta\Gamma_{j,\mathrm{BCL}}}{\Delta w}$ and $\mathbf{D}$ is the covariance matrix of these parameters.

The $\Delta\Gamma_i/\Delta w$ distribution depends only on the vector form factor f_+. However, because the kinematic constraint given in Equation (2.51) links f_+ and f_0 at the hadronic zero-recoil point, the fit retains sensitivity to the parameters of the scalar form factor a_i^0. Following Ref. [40], the kinematic constraint is expressed by writing a_2^0 in terms of the other coefficients.

In addition to the BCL fit, a fit to the BGL parameterization is implemented to check for closure with the MC generation input BGL form factor. Rather than constraining the parameters of the fit directly with LQCD, synthetic values of the form factors f_+ and f_0 at $w = 1$, $w = 1.08$ and $w = 1.16$ computed by the LQCD groups mentioned above are used as constraints. The theoretical constraint term in χ^2 is then replaced by

$$\sum_{k,l}\left(f_{+,0}^{\mathrm{LQCD}}(w_k) - f_{+,0}^{\mathrm{BGL}}(w_k)\right)\mathbf{D}_{kl}^{-1}\left((f_{+,0}^{\mathrm{LQCD}}(w_l) - f_{+,0}^{\mathrm{BGL}}(w_l)\right) \tag{6.25}$$

6.11 Fit validation

Closure test

Similar to the closure test of the template fit discussed in Section 6.9.3, the form factor fit can be validated by fitting the data sample obtained from simulation. After converting the signal strengths from the Asimov fit into differential decay rates and averaging across the four modes, the form factor fit is performed on the resulting spectrum to compare with the input value of $|V_{cb}|$. Although the decay file does not explicitly specify a value for $|V_{cb}|$, it can be inferred from

$$|V_{cb,\mathrm{generator}}| = \sqrt{\frac{\mathcal{B}(B \to D\ell\nu_\ell)}{\tau_B \Gamma_{\mathrm{theory}}}}, \tag{6.26}$$

where $\mathcal{B}(B \to D\ell\nu_\ell)$ is the relevant branching ratio in the decay file, and Γ_{theory} is obtained by integrating the BGL form factor with the generator-level form factor parameters as input. The value for $|V_{cb,\mathrm{generator}}|$ averaged across charged and neutral B modes is found

BGL	Fitted value	Generator value
V_{cb} [x1]	39.55 ± 0.81	39.76
a_{+0} [x100]	1.27 ± 0.01	1.26
a_{+1} [x100]	-9.48 ± 0.31	-9.40
a_{+2} [x10]	3.96 ± 1.52	3.40
a_{+3} [x100]	-33.73 ± 224.61	-10.00
a_{01} [x100]	-5.84 ± 0.26	-5.70
a_{02} [x10]	2.28 ± 1.37	1.20
a_{03} [x100]	-44.91 ± 219.73	40.00

Table 6.17: The form factor parameters are obtained by minimizing the χ^2 function for the BGL fit applied to the signal yields from the simulation pseudo-data. The resulting values are consistent with the input parameters within the uncertainties.

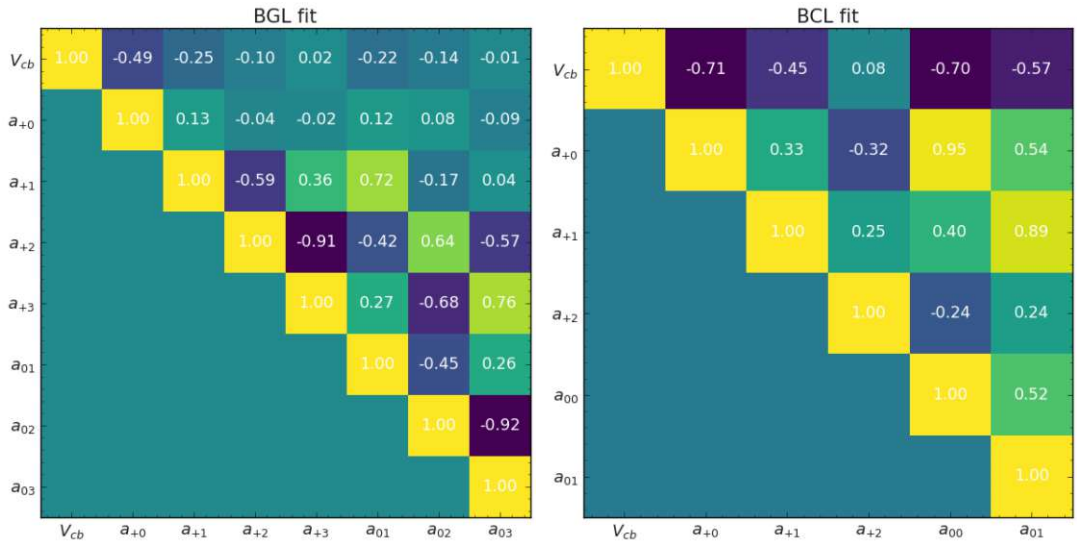
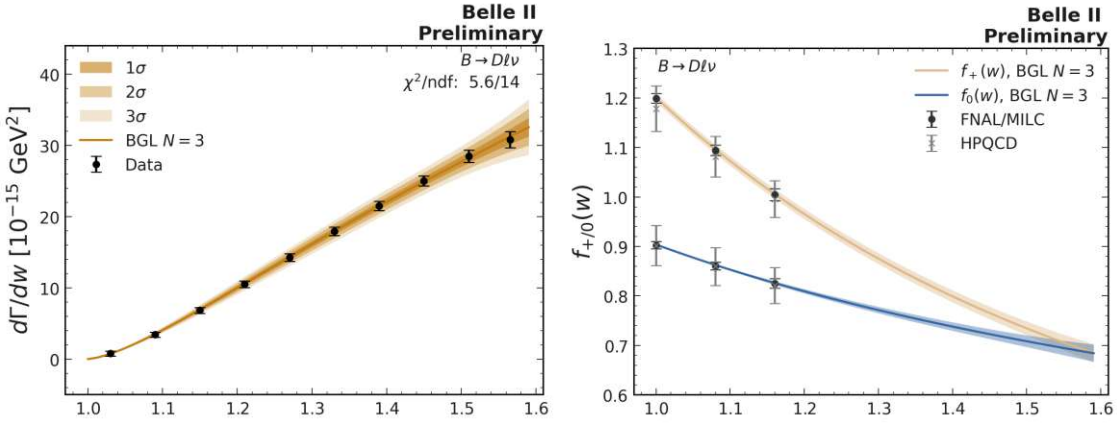


Figure 6.28: Correlation between BGL/BCL form factor parameters in the fit to MC.

to be $|V_{cb,generator}| = 39.76$. The values obtained from fitting Asimov data are shown in Table 6.17 and agree within uncertainties with the expected values. For the form factor fit to remain sensitive to $|V_{cb}|$, inputs from Lattice QCD can not be omitted in the χ^2, leading to expected deviations from the pure input values.

The correlation matrix between the BGL form factor parameters, obtained from the fit to the differential decay rates, is shown in Figure 6.28. By performing a similar Asimov fit with the BCL parameterization, the corresponding correlation matrix is shown in the same figure.
Figure 6.29 shows the fitted BGL form factor with the $\Delta\Gamma_i/\Delta w$ spectrum.



(a) (b)

Figure 6.29: The fitted BGL form factor with the extracted differential decay rates when fitting MC to itself.

Pull test

Further validation of the fit is conducted through a pull test, which evaluates the distribution of pulls as introduced in Section 4.3.1. In this test, toy samples of pseudo-data are generated by resampling bin-wise counts from the MC expectation, taking into account their uncertainties. Initially, a signal extraction fit is performed on these toy data sets. Subsequently, the BGL and BCL parameterized form factor fits are applied to the results of the signal extraction. For each toy sample, key parameters, including $|V_{cb}|$, the branching ratios of $B^+ \to \bar{D}^0 \ell^+ \nu_\ell$ and $B^+ \to \bar{D}^0 e^+ \nu_e$, and the ratio of light-lepton universality $R_{e/\mu}$ are computed.
For an unbiased fit that accurately estimates uncertainties, the pull distribution is expected to have a mean of 0 and a standard deviation of 1. The resulting pull distributions for the parameters of interest, based on 1000 toy samples, are shown in Figure 6.30 and are consistent with these expectations.

6.12 Results

6.12.1 Sensitivity estimate

A toy simulation approach is employed to assess the individual impacts of systematic and statistical sources on the measured parameters of interest. For each systematic uncertainty, the associated auxiliary data for the relevant `pyhf` modifiers are resampled to generate a set of pseudo-data. For instance, to evaluate the systematic effect on the $B \to D\ell\nu_\ell$ form factor, variations of each eigendirection are sampled, and the resulting histograms are used as pseudo-data. The full signal extraction fit is then applied to these pseudo-data sets.
This process is repeated 300 times for each source of uncertainty to create a distribution of signal yields, with its width reflecting the impact of the systematic uncertainty. The

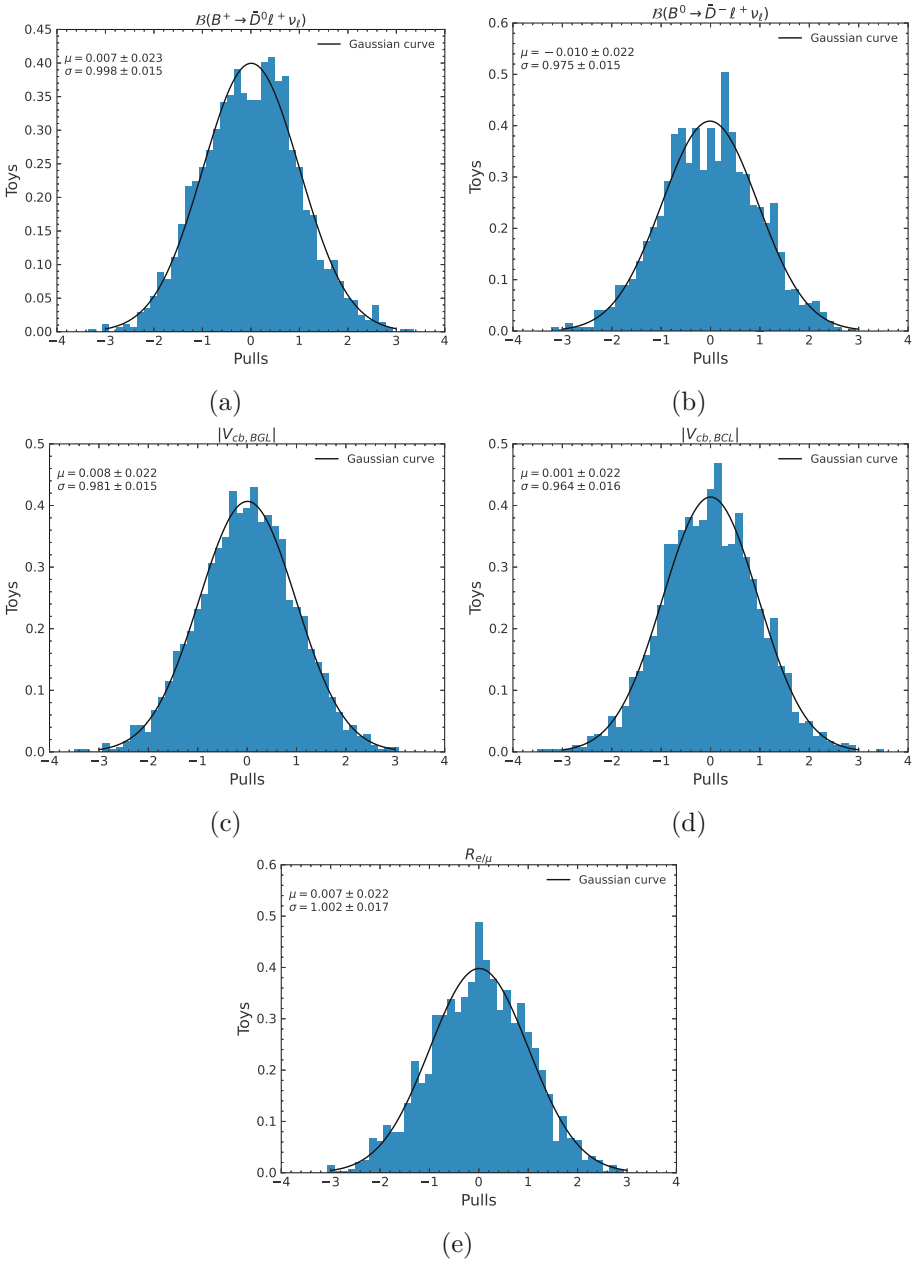


Figure 6.30: Pull distributions for the toy fit results for the average $B^+ \to \bar{D}^0\ell^+\nu_\ell$ and $B^0 \to D^-\ell^+\nu_\ell$ branching ratios, $|V_{cb}|$ with both BGL and BCL fits, and $R_{e/\mu}$.

$B \to D\ell\nu_\ell$	$\Delta\Gamma_i/\Delta w$ uncertainty [%]									
	0	1	2	3	4	5	6	7	8	9
Stat. Error	42.5	9.6	4.8	2.8	2.4	1.8	1.7	1.5	1.5	1.7
MC Stat. Error	21.3	4.8	2.4	1.4	1.2	0.9	0.9	0.7	0.7	0.9
N_{bb}	1.6	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
f_{00}/f_{+-}	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
$\mathcal{B}(D \to K\pi(\pi))$	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
Selection	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
$\mathcal{B}(B \to X_c\ell\nu)$	8.7	1.1	0.5	0.4	0.3	0.3	0.3	0.3	0.3	0.3
LeptonID	1.3	0.2	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.2
KaonID	10.8	2.1	1.4	1.1	0.9	0.8	0.8	0.8	0.8	0.8
Tracking efficiency	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$B \to D\ell\nu_\ell$ form factor	2.2	2.0	1.8	1.6	1.3	1.1	0.8	0.5	1.0	2.3
$B \to D^*\ell\nu_\ell$ form factor	6.7	0.9	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1
$\cos\theta_{BY}$ background modelling	2.8	0.8	0.5	0.3	0.3	0.2	0.2	0.2	0.2	0.3
w background modelling	55.2	10.1	4.4	2.5	1.7	1.3	1.0	0.9	1.1	1.4
$\tau_{B^{0/\pm}}$	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Total systematic	61.3	11.9	6.0	4.0	3.4	3.0	2.7	2.6	2.8	3.6
Total	74.6	15.3	7.6	4.9	4.2	3.5	3.2	3.0	3.2	4.0

Table 6.18: Split up systematic uncertainty for the differential decay rates averaged over the four modes.

uncertainties are propagated into sets of differential decay rates while accounting for correlations. By performing the form factor fit to the obtained differential decay rates, distributions for parameters such as $|V_{cb}|$ are obtained, enabling a detailed breakdown of uncertainties.

Differential decay rates

The resulting signal yields are converted into decay rates and averaged across all four modes using the covariance matrix. The split up of uncertainties obtained from the toy procedure is shown in Table 6.18.

Branching ratios

To obtain the branching ratios, the yields are summed up across bins of w, taking into account the correlations. The toy resampling method is used to average electron and muon modes. The resulting table of uncertainties is shown in Table 6.19.

	$\mathcal{B}(B \to D\ell\nu_\ell)$ [%]	
	$B^+ \to \bar{D}^0\ell^+\nu_\ell$	$B^0 \to D^-\ell^+\nu_\ell$
Stat. Error	1.2	1.5
MC Stat. Error	0.6	0.8
N_{bb}	1.5	1.5
f_{00}/f_{+-}	2.6	2.6
$\mathcal{B}(D \to K\pi(\pi))$	0.7	1.7
Selection	0.5	0.5
$\mathcal{B}(B \to X_c\ell\nu)$	0.3	0.3
LeptonID	0.2	0.2
KaonID	0.6	1.3
Tracking efficiency	0.9	1.1
$B \to D\ell\nu_\ell$ form factor	< 0.1	< 0.1
$B \to D^*\ell\nu_\ell$ form factor	0.1	< 0.1
$\cos\theta_{BY}$ background modelling	0.2	0.4
w background modelling	0.6	0.7
Total systematic	3.4	4.1
Total	3.6	4.3

Table 6.19: Split up systematic uncertainties on the overall branching ratios of the $B^+ \to \bar{D}^0\ell^+\nu_\ell$ and $B^0 \to D^-\ell^+\nu_\ell$ decays.

Lepton universality

When propagating the full covariance matrix to a ratio of e over μ $R_{e/\mu}$, most systematics cancel out in large parts. LeptonID and background modelling effects are the most significant remaining systematics, as seen in Table 6.20. A sensitivity of approximately 2% is expected on $R_{e/\mu}$.

$|V_{cb}|$ and form factor parameters

By performing the BCL/BGL fits on each toy data set, a systematics breakdown on $|V_{cb}|$ is obtained and shown in tables Table 6.21. Here, we obtain an estimate of the theoretical uncertainty by also drawing toys for the LQCD input. We use the published covariance matrices to draw these toys.
Systematic uncertainties dominate the overall sensitivity on $|V_{cb}|$. The dominant sources include the knowledge of the amount of B mesons in the sample and the effect of varying the signal template shape through the form factor. The next leading systematics at the order of 0.5% include background shape systematics, KaonID and tracking efficiency. A total uncertainty of around 2.1% on V_{cb} is expected, consisting of 1.7% experimental uncertainty and 1.2% theoretical uncertainty. A smaller overall uncertainty is obtained with the more constrained BCL fit.

	$R_{e/\mu}$ [%]		
	$B^+ \to \bar{D}^0 \ell^+ \nu_\ell$	$B^0 \to D^- \ell^+ \nu_\ell$	$B \to D \ell \nu_\ell$
Stat. Error	2.1	2.7	1.7
MC Stat. Error	1.0	1.4	0.8
N_{bb}	< 0.1	< 0.1	< 0.1
f_{00}/f_{+-}	< 0.1	< 0.1	< 0.1
$\mathcal{B}(D \to K\pi(\pi))$	< 0.1	< 0.1	< 0.1
Selection	< 0.1	< 0.1	< 0.1
$\mathcal{B}(B \to X_c \ell \nu)$	0.1	0.1	0.1
LeptonID	0.4	0.4	0.4
KaonID	< 0.1	0.1	< 0.1
Tracking efficiency	< 0.1	< 0.1	< 0.1
$B \to D \ell \nu_\ell$ form factor	< 0.1	< 0.1	< 0.1
$B \to D^* \ell \nu_\ell$ form factor	0.1	< 0.1	0.1
$\cos\theta_{BY}$ background modelling	0.5	0.7	0.4
w background modelling	0.3	0.3	0.3
Total systematic	1.2	1.6	1.0
Total	2.4	3.2	2.0

Table 6.20: Breakdown of systematic uncertainties when measuring the ratio of the electron over muon decay rates for $B^+ \to \bar{D}^0 \ell^+ \nu_\ell$, $B^0 \to D^- \ell^+ \nu_\ell$, and the average over all modes.

	$\lvert V_{cb,BGL}\rvert$ [%]	$\lvert V_{cb,BCL}\rvert$ [%]
Stat. Error	0.7	0.6
MC Stat. Error	0.4	0.3
N_{bb}	0.8	0.8
f_{00}/f_{+-}	< 0.1	< 0.1
$\mathcal{B}(D \to K\pi(\pi))$	0.4	0.4
Selection	0.2	0.2
$\mathcal{B}(B \to X_c\ell\nu)$	0.2	0.1
LeptonID	0.1	0.1
KaonID	0.4	0.4
Tracking efficiency	0.5	0.5
$B \to D\ell\nu_\ell$ form factor	0.8	0.4
$B \to D^*\ell\nu_\ell$ form factor	0.1	0.1
$\cos\theta_{BY}$ background modelling	0.1	0.1
w background modelling	0.5	0.4
$\tau_{B^{0/\pm}}$	0.1	0.1
Total systematic	1.5	1.4
Theory	1.3	1.2
Total	2.1	1.9

Table 6.21: Breakdown of systematic uncertainties on V_{cb} when fitting both the BCL and the BGL parameterized form factor.

6.12.2 Fit to data

As mentioned in Chapter 1, the blinded analysis described in this chapter is under internal review as of the time of writing the thesis. The fit to data with the final measurement of $|V_{cb}|$ has not been unblinded and is therefore not presented here.
A preliminary result based on a data sample with an integrated luminosity of 189.2 fb^{-1} [12] passed internal review and is presented here. Compared to the updated measurement described in this chapter, the preliminary result differs in several ways:

- The number of B meson pairs in the data sample is approximately half that of the updated measurement, resulting in increased statistical uncertainty.
- The simulated data samples are run-independent and do not reflect dependencies on run periods.
- The template fit is conducted individually for each mode and each signal window rather than using a combined global fit. Although correlations of systematic uncertainties are considered using a toy approach, the backgrounds are constrained to a lesser extent, leading to larger systematic uncertainties associated with background modelling.
- Rather than directly mapping unfolding effects into the fit, as described in Section 6.9.1, the preliminary result uses bin-by-bin unfolding [99]. In this approach, w-dependent efficiencies are computed to map the reconstructed yields into unfolded yields. Compared to the direct unfolding approach in the updated measurement, different systematic uncertainties affecting the magnitude of the migration effects do not propagate into the unfolding method.
- The central value of $|V_{cb}|$ is derived by averaging the $|V_{cb}|$ values obtained from each mode individually, assuming full correlation of systematic uncertainties. In the updated measurement, the full uncertainty matrix from the global fit is propagated into the differential decay rates, which allows for the cancellation of systematic uncertainties. The most significant cancellation arises from $f^{\pm}/f^{00}$.
- The theoretical inputs from lattice QCD are used in the form of synthetic measurements of f_+ and f_0 at $w = 1$, 1.08, and 1.16, obtained from the FNAL/MILC [38] and HPQCD [39] collaborations. In this approach, uncertainties from lattice calculations are assumed to be fully correlated between the individual groups. In contrast, the updated measurement employs an official average of these collaborations reported by FLAG [40].

The resulting values of $|V_{cb}|$ obtained from the fit to the BGL parameterized form factor are shown in Table 6.22. The result of the form factor fit overlaid over the differential decay rates in each mode is shown in Figure 6.31.
Averaging over the four final states yields

$$\eta_{\text{EW}}|V_{cb}| = (38.53 \pm 1.15) \times 10^{-3}. \tag{6.27}$$

	$B^+ \to \bar{D}^0 e^+ \nu_e$	$B^+ \to \bar{D}^0 \mu^+ \nu_\mu$	$B^0 \to D^- e^+ \nu_e$	$B^0 \to D^- \mu^+ \nu_\mu$
$a_{+,0} \times 10^2$	1.27 ± 0.01	1.27 ± 0.01	1.26 ± 0.01	1.26 ± 0.01
$a_{+,1} \times 10$	-0.96 ± 0.03	-0.95 ± 0.03	-0.95 ± 0.03	-0.95 ± 0.03
$a_{+,2}$	0.44 ± 0.17	0.33 ± 0.17	0.36 ± 0.17	0.41 ± 0.17
$a_{+,3}$	-3.06 ± 2.76	0.73 ± 3.09	0.10 ± 2.79	-1.41 ± 2.93
$a_{0,1} \times 10$	-0.59 ± 0.03	-0.58 ± 0.03	-0.58 ± 0.03	-0.59 ± 0.03
$a_{0,2}$	0.29 ± 0.15	0.18 ± 0.15	0.21 ± 0.15	0.25 ± 0.15
$a_{0,3}$	-3.59 ± 2.73	0.35 ± 3.09	-0.13 ± 2.75	-1.76 ± 2.89
$\eta_{EW}\|V_{cb}\| \times 10^3$	38.72 ± 1.09	37.91 ± 1.27	38.47 ± 1.10	38.89 ± 1.17
χ^2/ndf	18.2/14	12.3/14	11.0/14	15.1/14

Table 6.22: Result of the BGL fit to the partial widths and two lattice QCD calculations of the form factors f_+ and f_0 [38, 39]. The BGL series is truncated at $N = 3$.

The quoted error includes experimental and theoretical uncertainties. Assuming $\eta_{\text{EW}} = 1.0066 \pm 0.0050$ [105] for the electroweak correction introduced in Equation (2.49) results in

$$|V_{cb}| = (38.28 \pm 1.16) \times 10^{-3}, \tag{6.28}$$

which is consistent within one standard deviation with the current world average from exclusive determinations [3]

$$|V_{cb}| = (39.10 \pm 0.50) \times 10^{3}. \tag{6.29}$$

The total uncertainty on $|V_{cb}|$ in the preliminary result is ∼3%. In comparison, the updated measurement aims for a sensitivity slightly above the 2% mark, which would represent the most precise measurement of $|V_{cb}|$ using the $D\ell\nu$ final state to date.

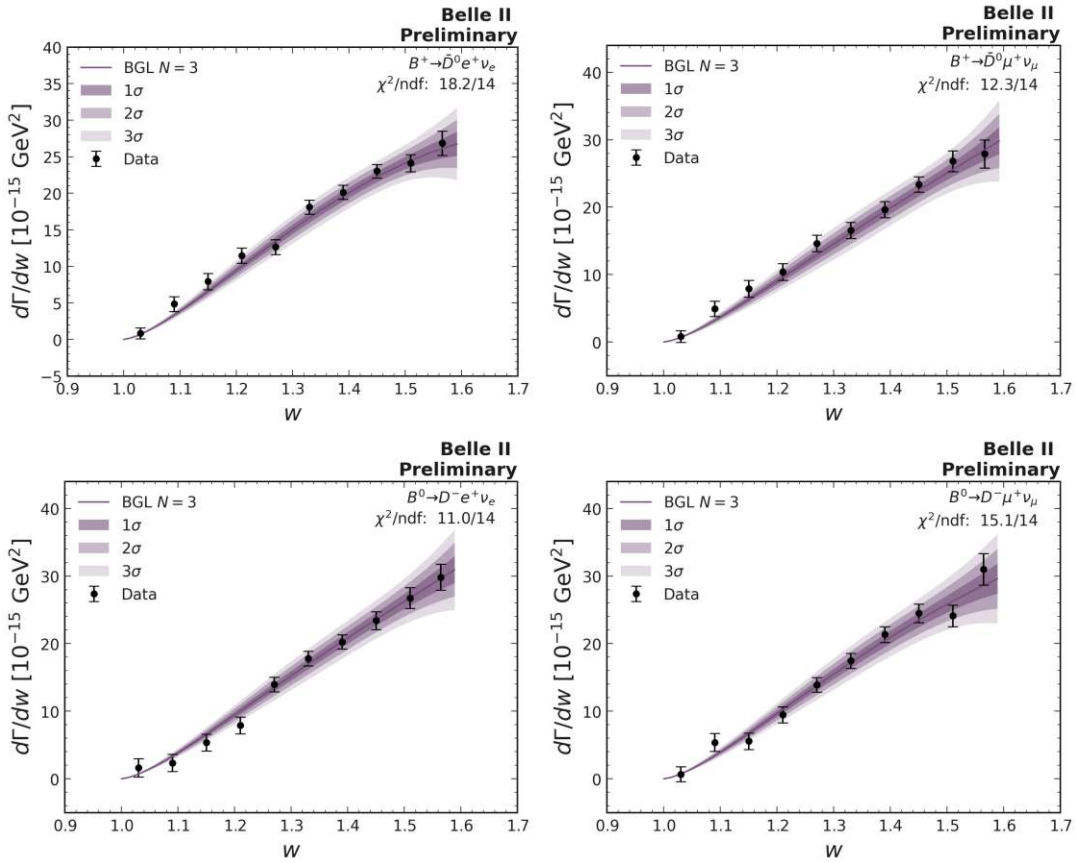


Figure 6.31: $B \to D\ell\nu_\ell$ differential decay rates and the result of the BGL fit for all four final states.

CHAPTER 7

Summary and outlook

A long-standing tension between exclusive and inclusive determinations of the CKM matrix elements $|V_{cb}|$ and $|V_{ub}|$ motivates further measurements of these fundamental SM parameters. The second generation B factory Belle II started operating in 2019, with the goal of collecting a data sample containing an unprecedented amount of B mesons with improved performance of every subdetector. As of 2022, the data sample amounts to $N_{B\bar{B}} = (387.1 \pm 6) \times 10^6$, equivalent to approximately half of the B meson decays collected throughout the full Belle run time.
By making use of this data sample to analyze the $B^+ \to \bar{D}^0 \ell^+ \nu_\ell$ and $B^0 \to D^- \ell^+ \nu_\ell$ decays, $|V_{cb}|$ can be measured in an exclusive approach. Signal yields are extracted in differential bins of w, and after unfolding, the differential decay rates are obtained. By using external inputs from lattice QCD as constraints, performing a fit of the form factor to the measured differential decay rates gives access to measuring $|V_{cb}|$. In a preliminary result using a data sample of 189 fb^{-1}, a fit to the BGL parameterized form factor yields [12]

$$|V_{cb}| = (38.28 \pm 1.16) \times 10^{-3}, \tag{7.1}$$

where the uncertainty represents the sum of statistical, systematic and theoretical sources. The result is in good agreement with the world average across all exclusive determinations [3]

$$|V_{cb}| = (39.10 \pm 0.50) \times 10^{3}. \tag{7.2}$$

In the preliminary result, the relative uncertainty from all sources combined totals ∼3%. The corresponding Belle analysis of the $B \to D\ell\nu_\ell$ decays [4] results in a measurement of $|V_{cb}|$ with a precision of ∼2.7% using a significantly larger data sample that corresponds to a luminosity of 711 fb^{-1}.
An update to the preliminary measurement improves the sensitivity through multiple aspects: the size of the data sample is increased to 362 fb^{-1}; simulated data takes into account run period dependent conditions; the signal extraction is performed simultaneously across all signal windows; the full resulting covariance matrix can be used to cancel

uncertainties from systematic sources partially; direct unfolding allows mapping effects of systematic uncertainties directly into the fit result; and the form factor fit includes updated inputs from theory. With the new approach, the estimated sensitivity improves to a total relative uncertainty of ∼2.1% on $|V_{cb}|$, corresponding to world-leading precision with the $B \to D\ell\nu_\ell$ final state. Additionally, the lepton flavour universality between e and μ can be tested to a precision of 2%.

The Belle II experiment is still in its starting phase, and a significantly larger data sample is expected to be collected over its lifetime. Complementary measurements to the one described in this thesis can help shed light on the exclusive-inclusive puzzle. A measurement of $|V_{cb}|$ via the $B \to D^*\ell\nu_\ell$ decay [94] results in comparable precision in the data sample corresponding to 189 fb^{-1}. While being able to reconstruct the mode with higher purity, reliance on systematics such as the tracking of slow pions represents a trade-off with respect to a $B \to D\ell\nu_\ell$ analysis. Looking ahead, a simultaneous determination of $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ in one combined analysis allows leveraging of advantages of each approach, leading to an improved measurement.
On the inclusive front, Belle II can contribute by analyzing $B \to X_c\ell\nu_\ell$ decays. Measuring spectral moments of the momentum transfer q^2 for different minimum q^2 thresholds allows extraction of $|V_{cb}|$ [106, 107]. Additionally, a key input to the inclusive determination is the semileptonic branching ratio $\mathcal{B}(B \to X_c\ell\nu_\ell)$, where Belle II has the unique opportunity of providing world-leading measurements [108].

APPENDIX A

Electron identification correction factors

This appendix presents the results of the electronID performance study in terms of the obtained correction factors. The correction factors shown with their relative statistical and systematic uncertainties for different variables, thresholds and charges.
The resulting correction factors are close to 1 for low thresholds, indicating good agreement between simulation and data especially in the medium momentum range and barrel region. With increasing threshold, deviations from 1 increase and in turn systematic uncertainties increase. The increase in systematic uncertainties is driven from the comparison between samples selected with different trigger line requirements.

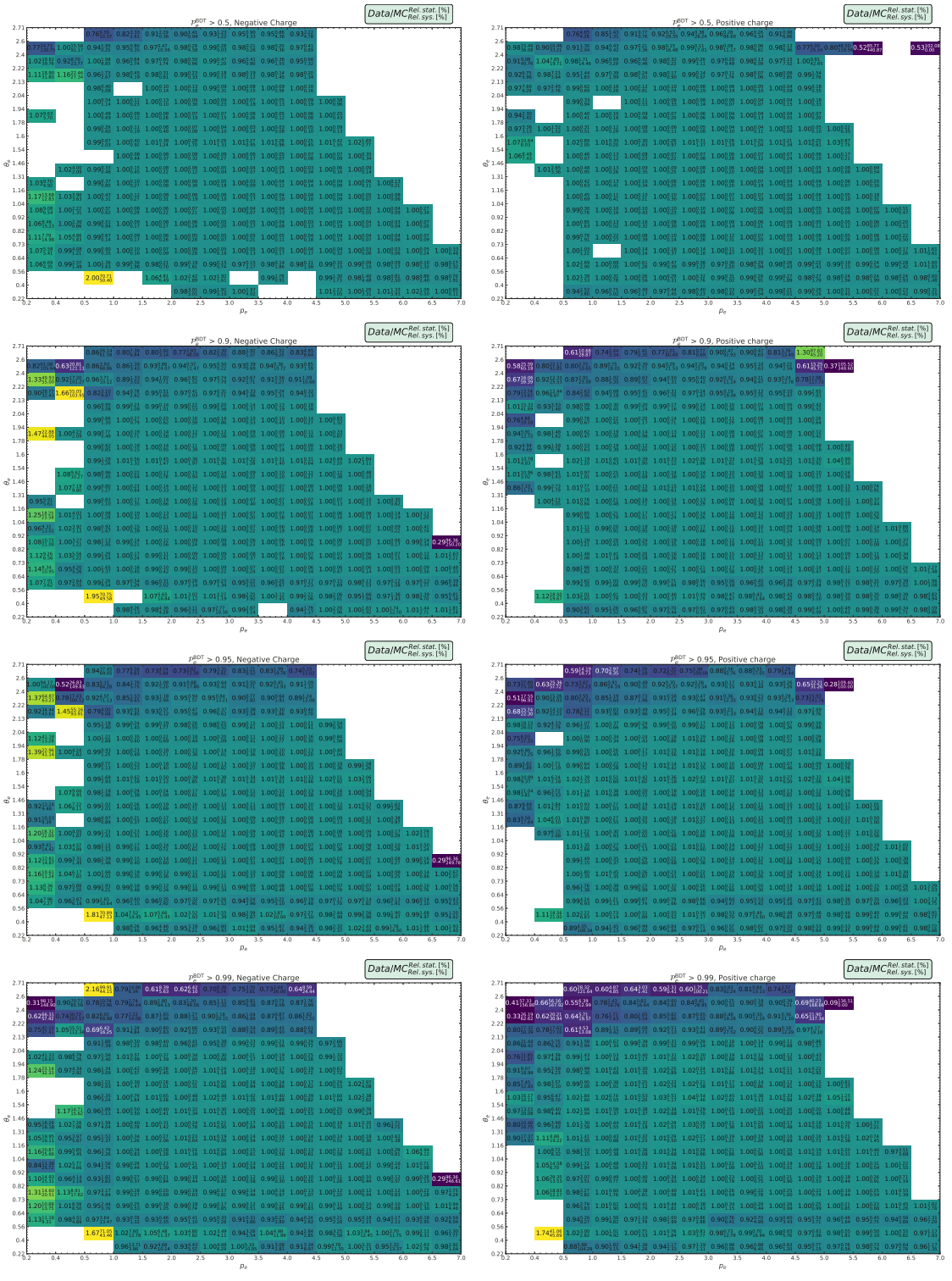


Figure A.1: Correction factors and uncertainties for BDT-based global electronID with different thresholds and charges.

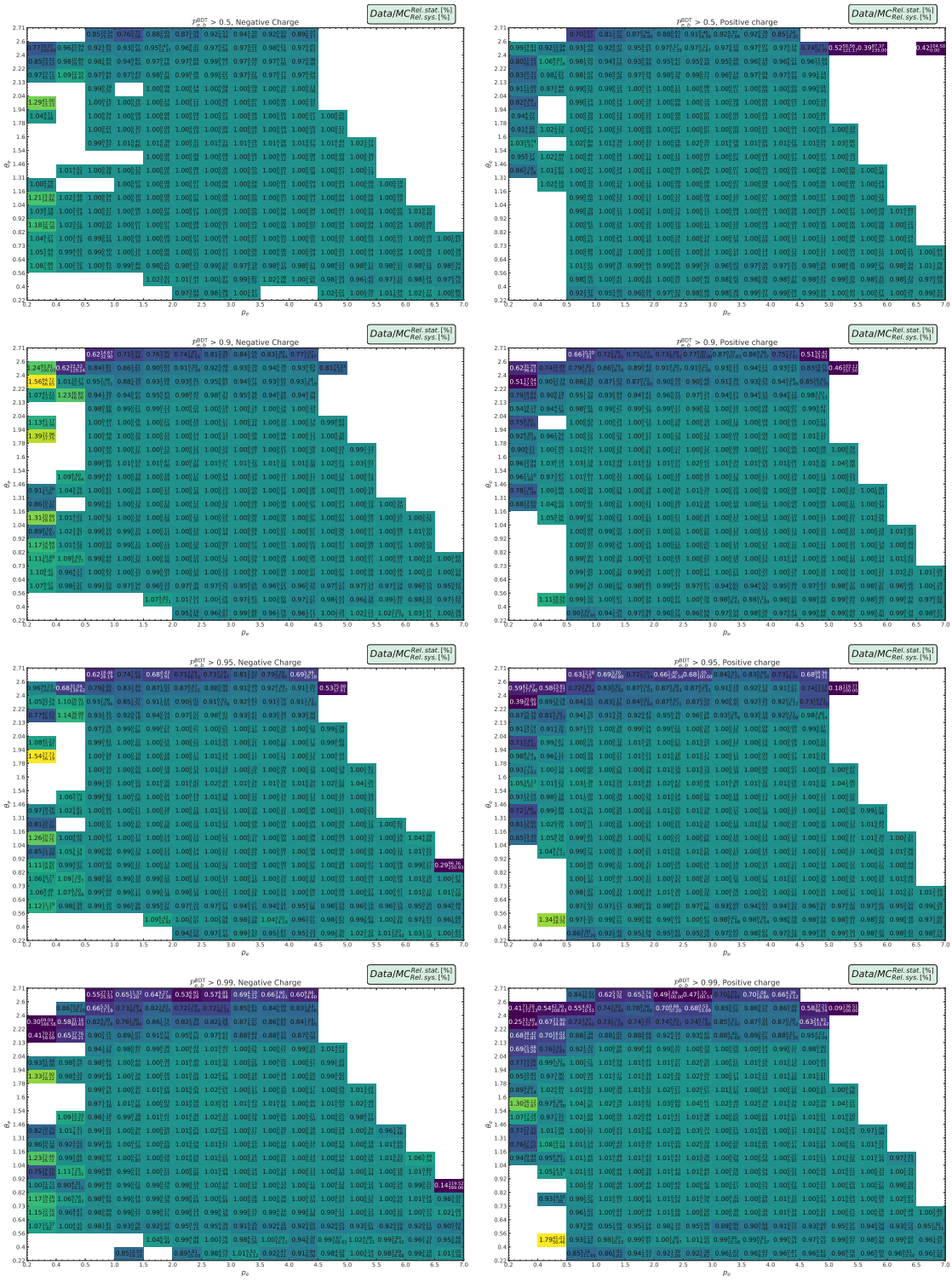


Figure A.2: Correction factors and uncertainties for BDT-based binary electronID with different thresholds and charges.

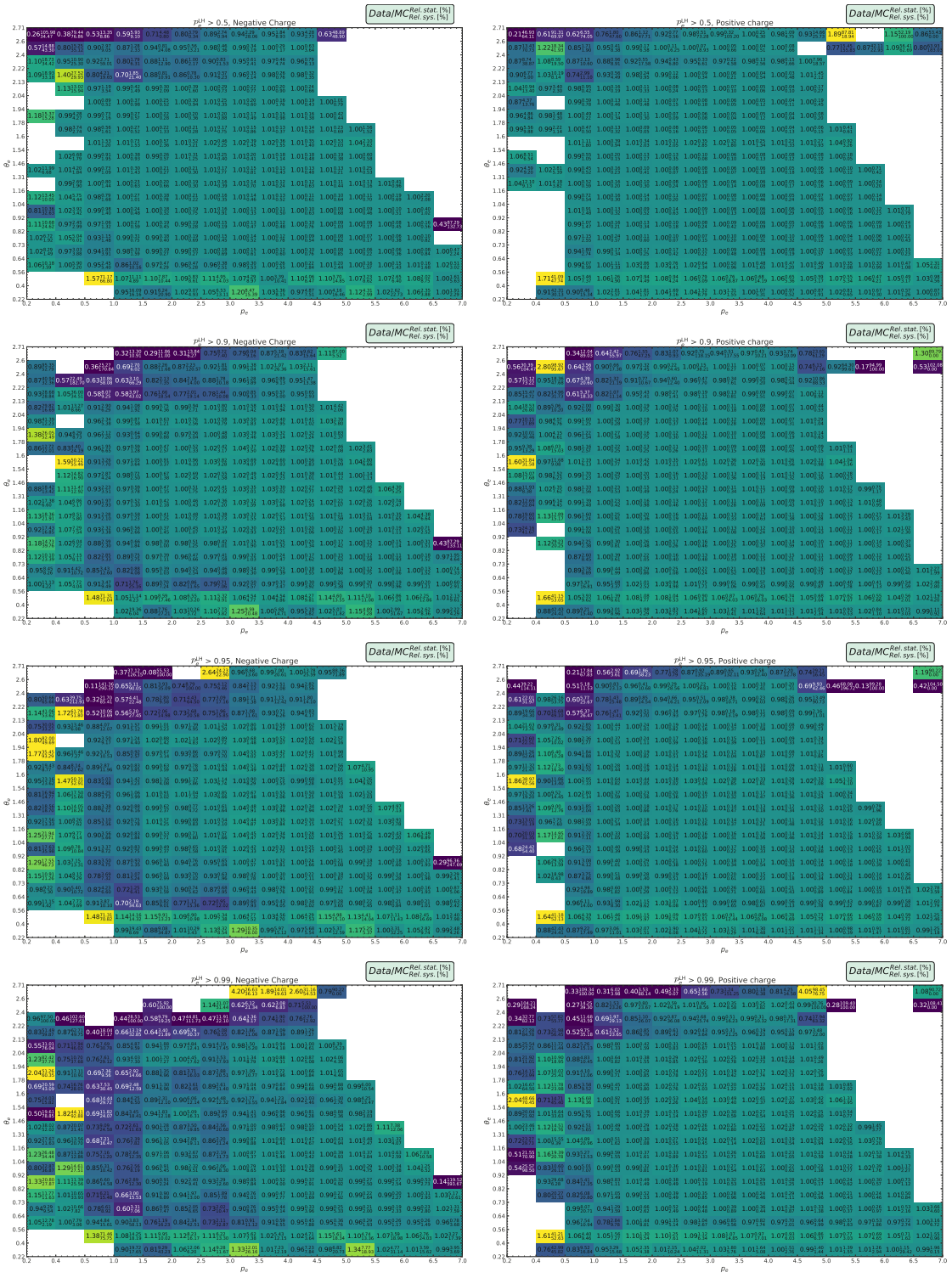


Figure A.3: Correction factors and uncertainties for likelihood-based global electronID with different thresholds and charges.

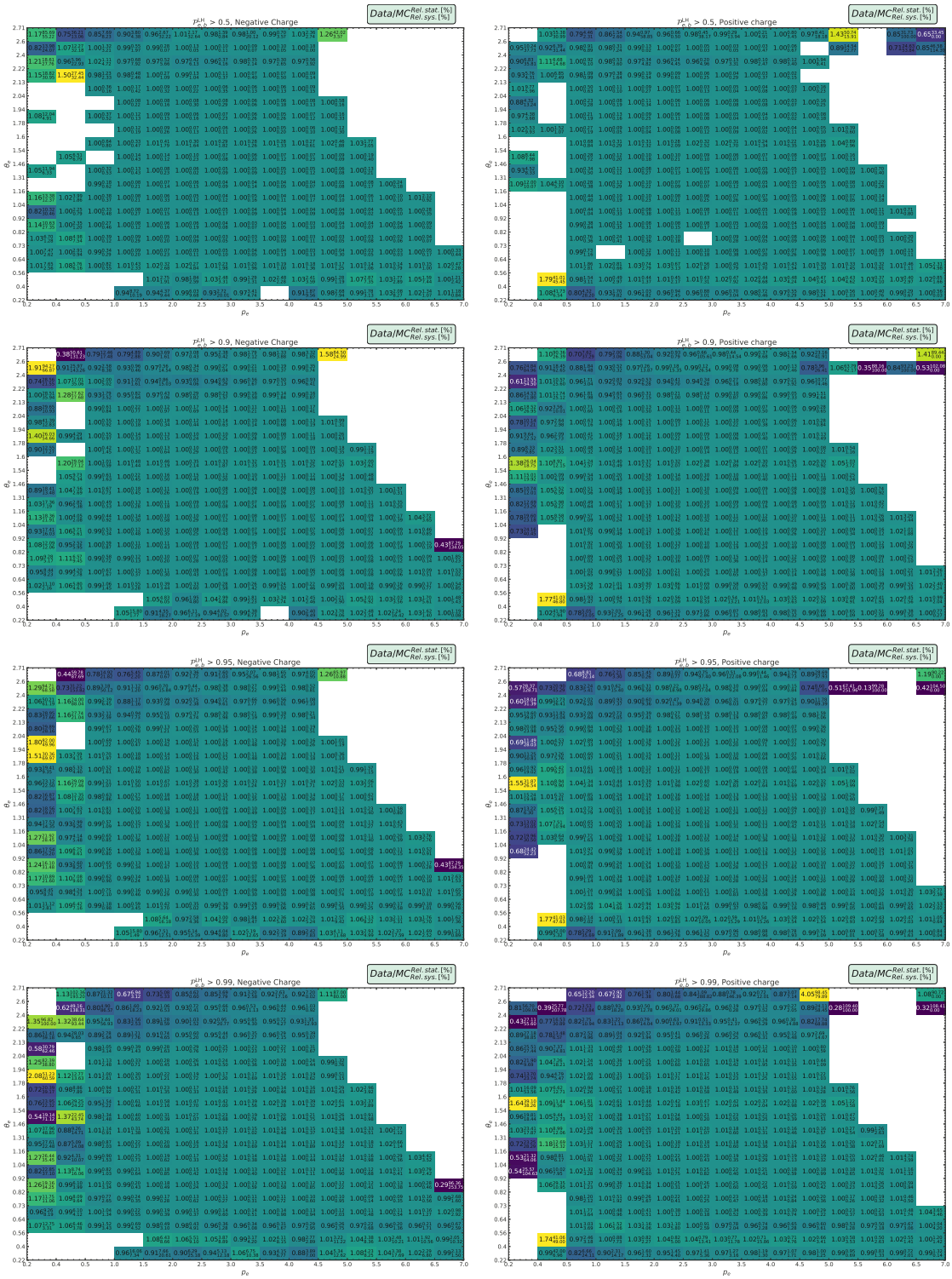


Figure A.4: Correction factors and uncertainties for likelihood-based binary electronID with different thresholds and charges.

APPENDIX B

Cutflow diagrams

To further illustrate the effects of the optimized selections, the plots in this appendix show the flow of purity and efficiency after each selection. One diagram is shown for each bin of w, in each of the 4 $B \to D\ell\nu_\ell$ modes.

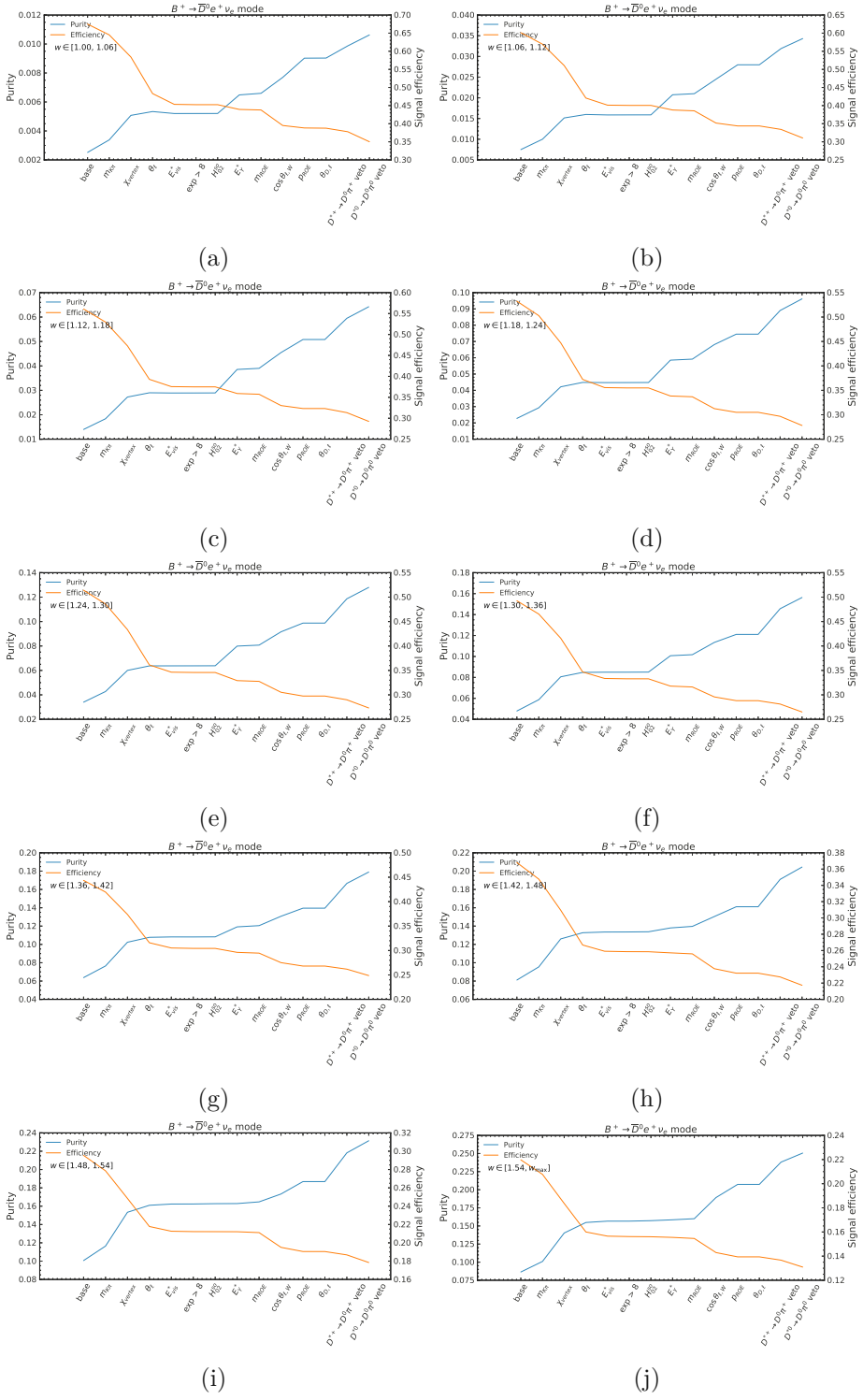


Figure B.1: Purity and signal efficiency in bins of w for the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode.

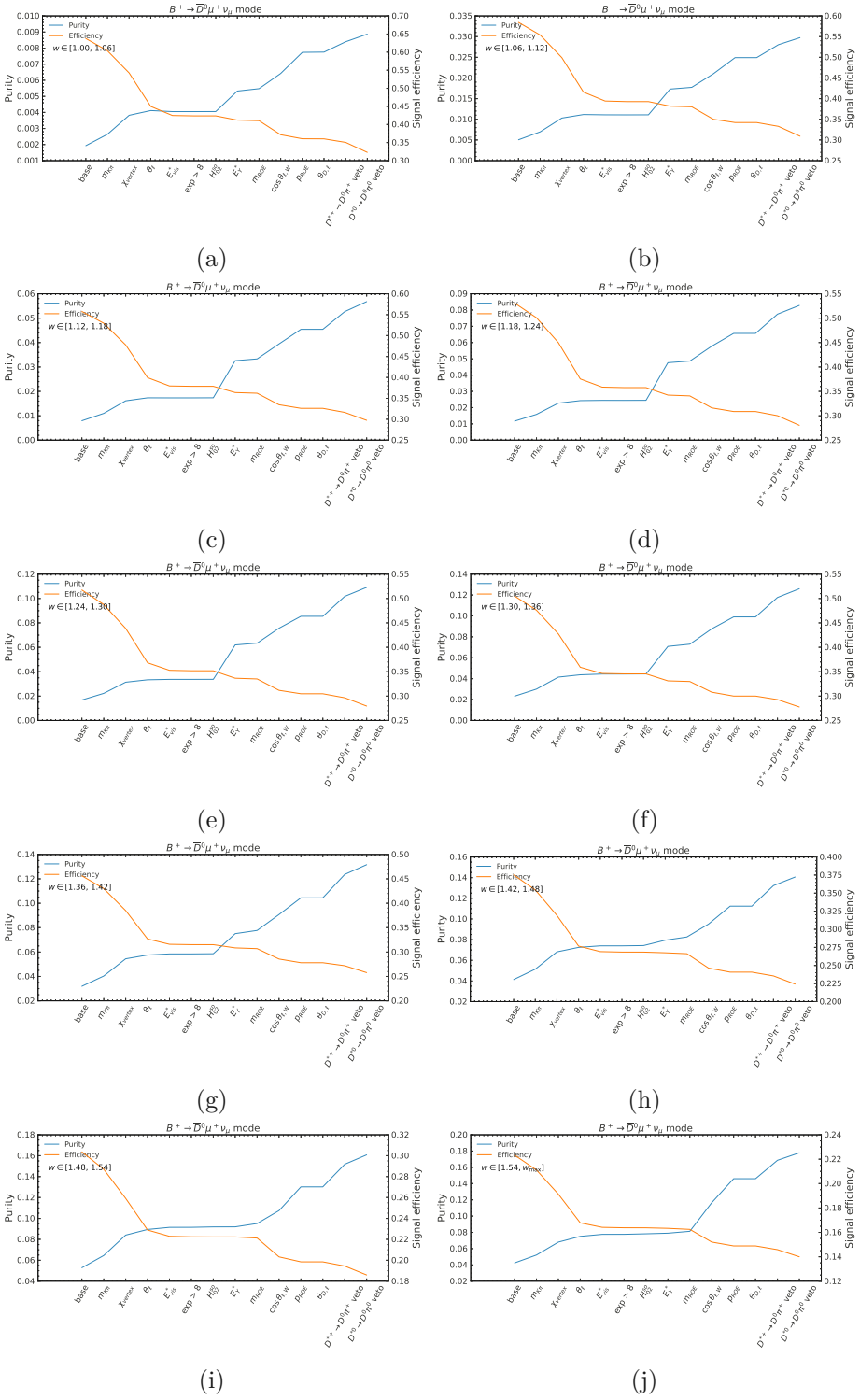


Figure B.2: Purity and signal efficiency in bins of w for the $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$ mode.

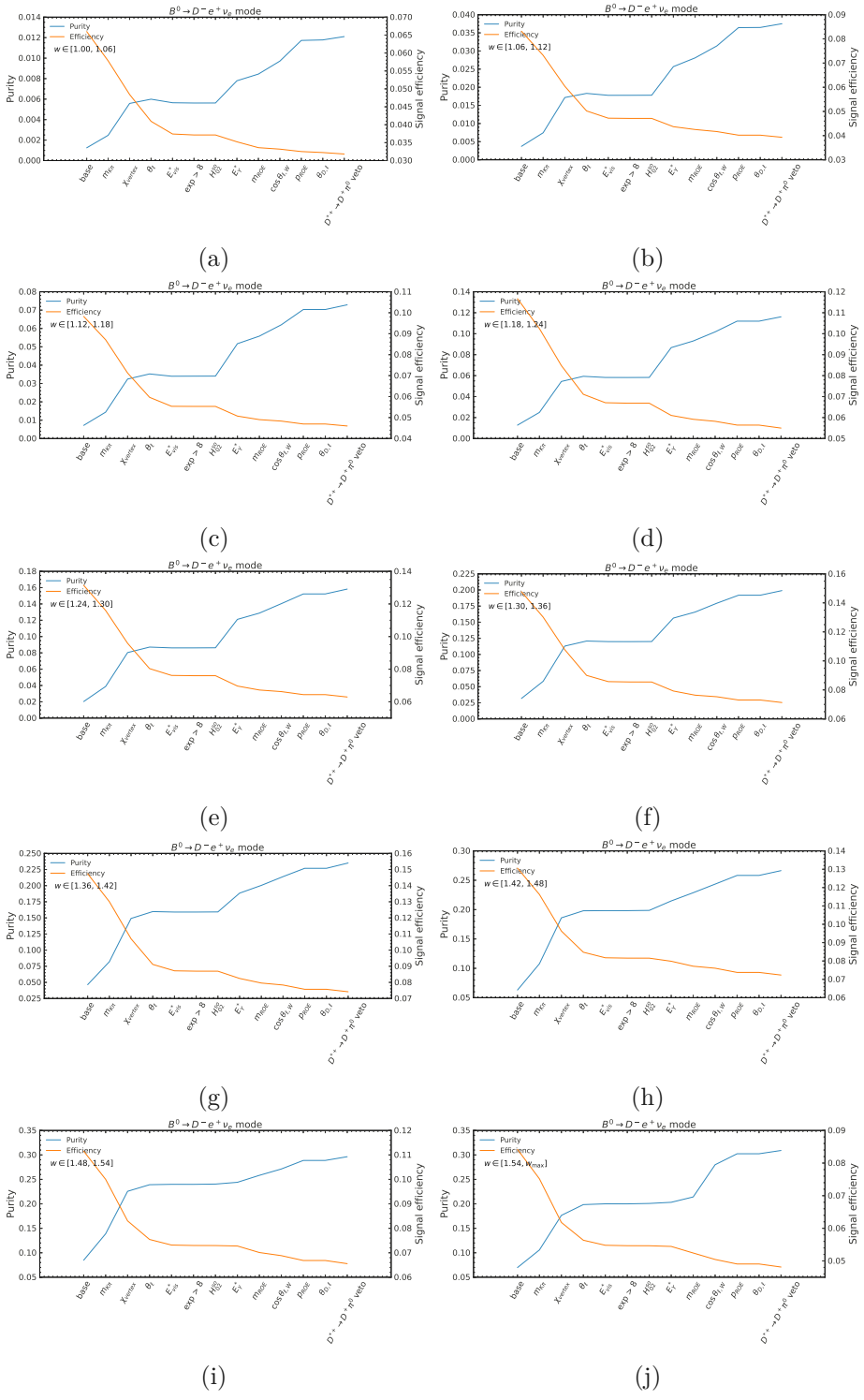


Figure B.3: Purity and signal efficiency in bins of w for the $B^0 \to D^- e^+ \nu_e$ mode.

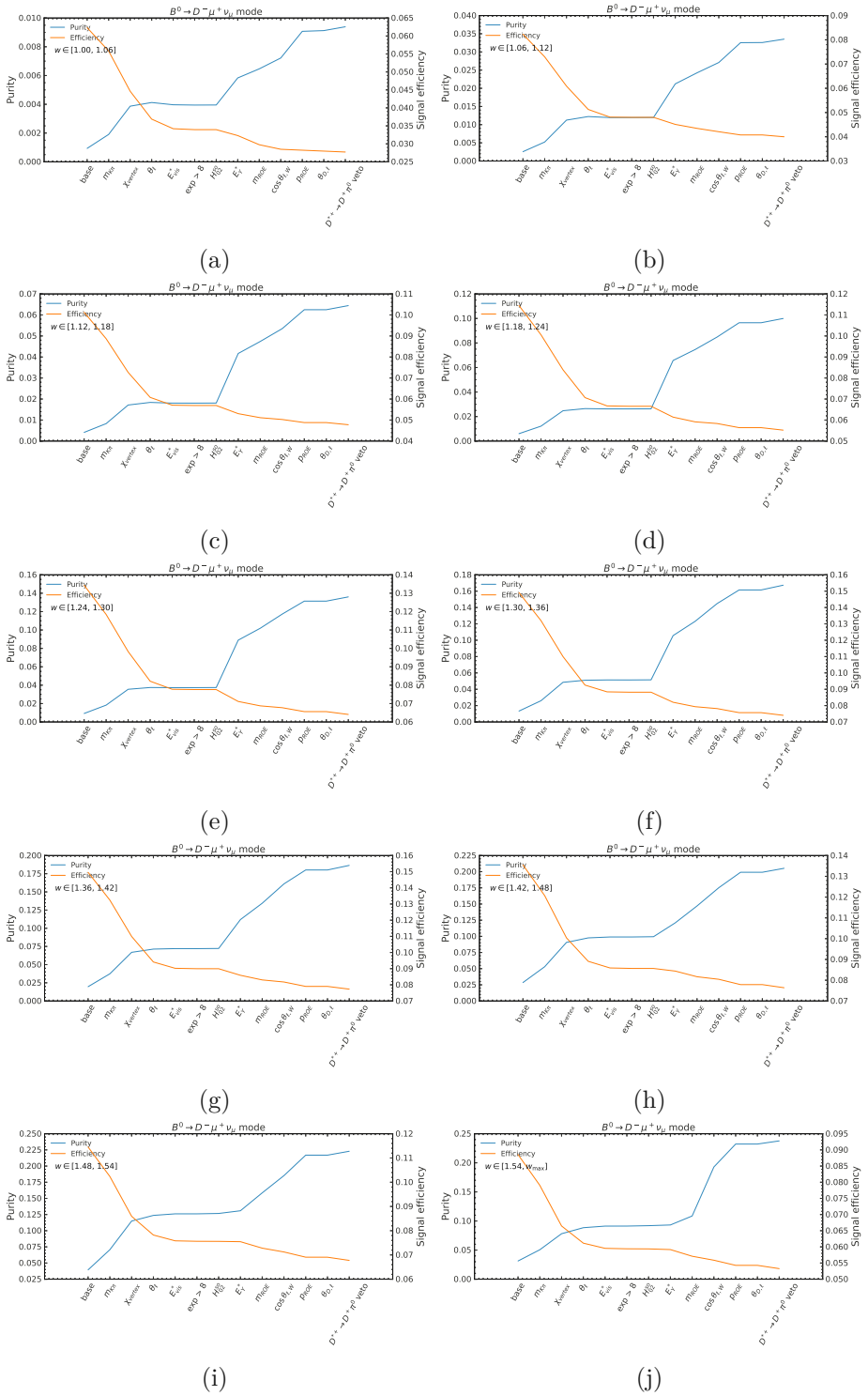


Figure B.4: Purity and signal efficiency in bins of w for the $B^0 \to D^- \mu^+ \nu_\mu$ mode.

APPENDIX C

Selection comparison in Data and MC

This appendix shows a comparison of selection variables between data and simulation. The selection requirements are indicated in red dashed lines. In each diagram, every other selection except for the one depicted is applied. Some choices of loose preselections have to be made in the initial phase of reconstructing candidates and are therefore not shown here. The selections shown include data-driven selections, selecions optimized from simulated annealing, and explicit $B \to D^* \ell \nu_\ell$ vetoes. For vetoes, the diagrams only show candidates in which an appropriate slow pion is found.
The uncertainties pictured as hashed bands on the MC histograms include systematic uncertainties from $B \to X_c \ell \nu_\ell$ corrections, particleID, and background shape corrections. The uncertainty on data pictured as error bars around the black dots are purely statistical. After applying all corrections, while most variables are seen to be within expectation of the systematic uncertainties, some residual disagreement is observed. To account for the effect this has on the reconstruction efficiency, a check of control samples is performed, described in Section 6.5.2. For the vertex fit probability, a correction to the signal efficiency and a corresponding systematic uncertainty is applied.

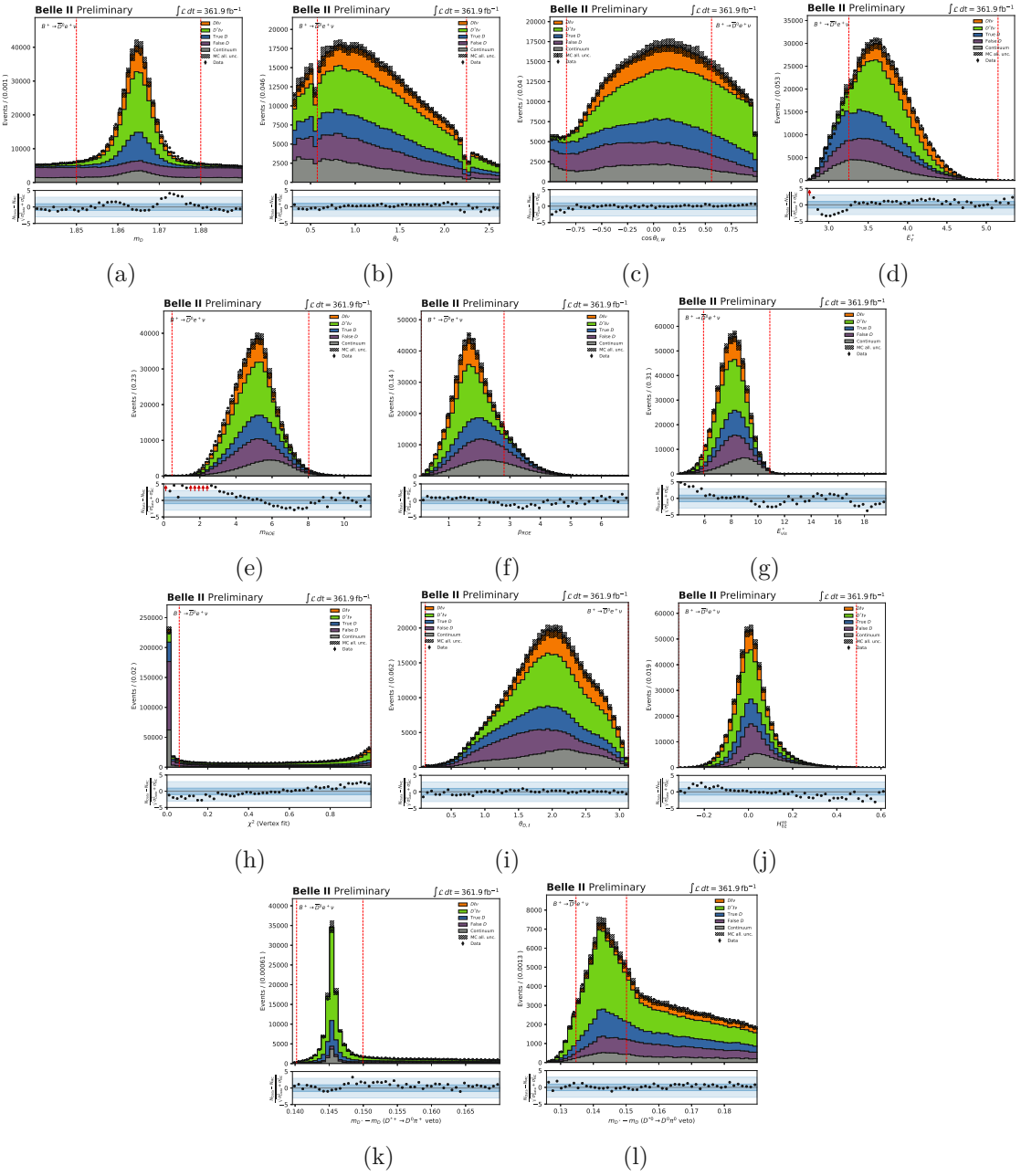


Figure C.1: All selections applied except the one shown in the plot for the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode.

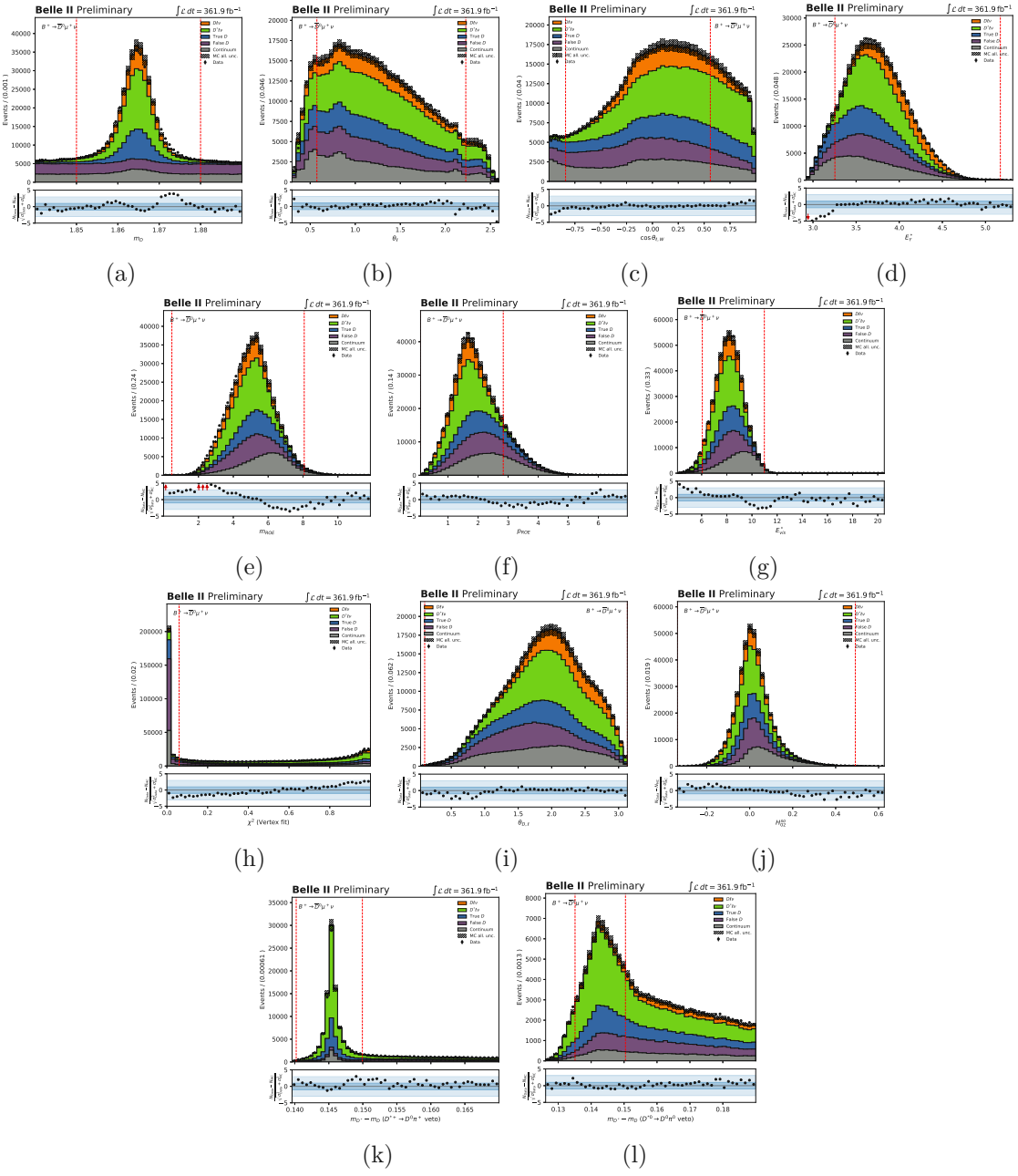


Figure C.2: All selections applied except the one shown in the plot for the $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$ mode.

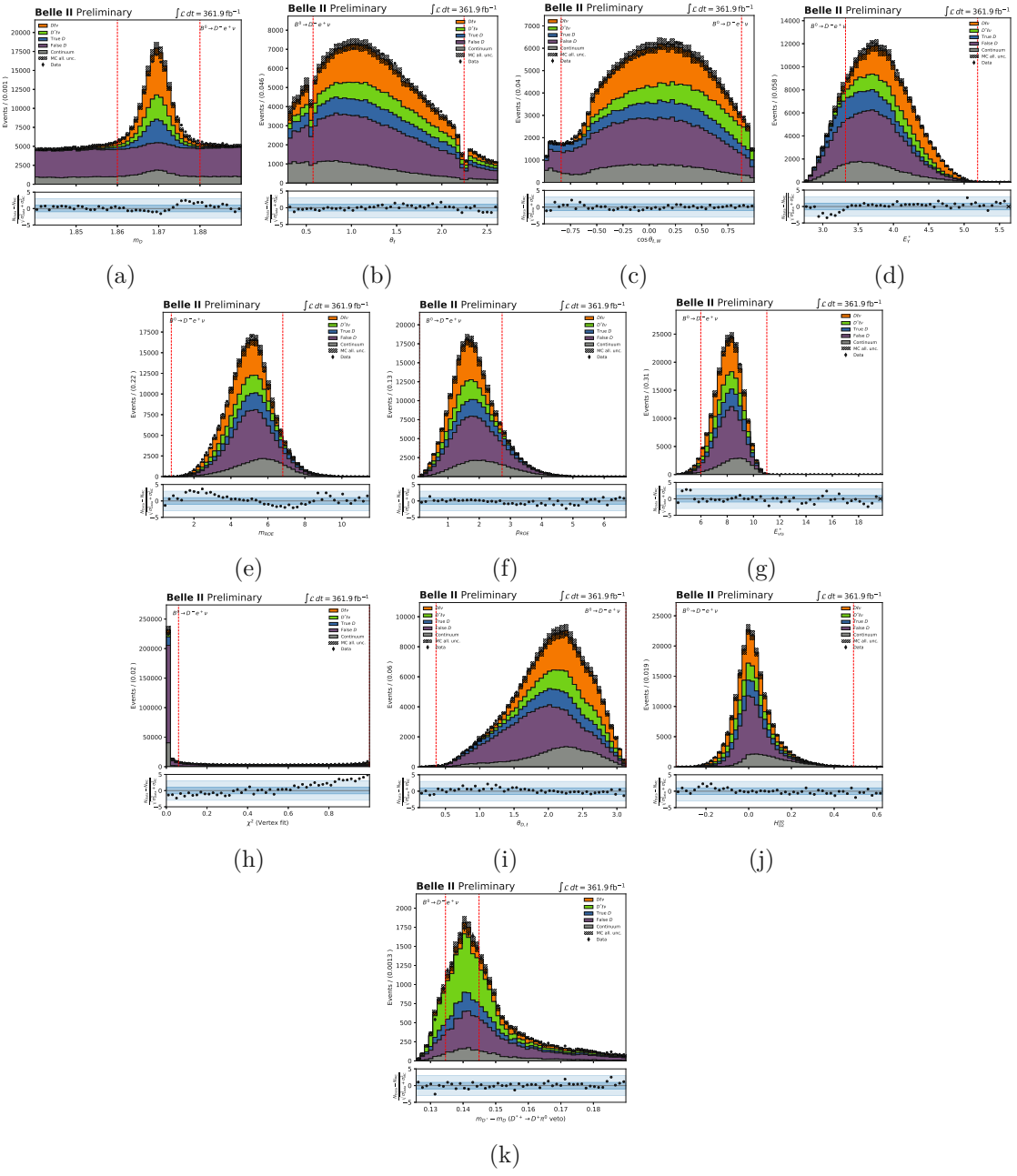


Figure C.3: All selections applied except the one shown in the plot for the $B^0 \to D^- e^+ \nu_e$ mode.

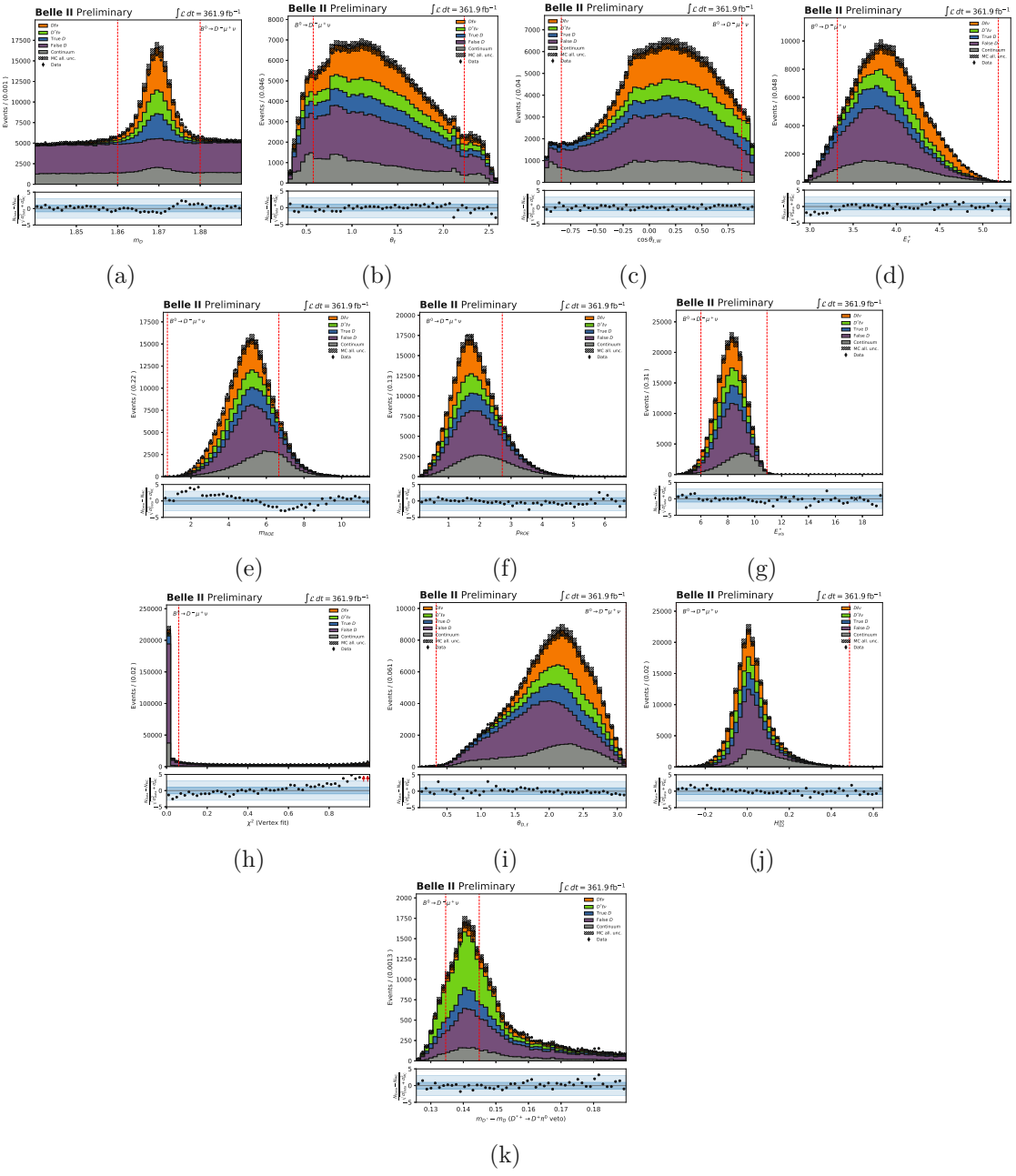


Figure C.4: All selections applied except the one shown in the plot for the $B^0 \to D^- \mu^+ \nu_\mu$ mode.

List of Figures

2.1 A summary table of all fundamental particles. Each particle is shown with its respective mass, charge, and spin in the upper left corner [18]. 6
2.2 Feynman diagrams for electromagnetic interactions. Subfigure (a) depicts an electron that emits a photon in a process known as *bremsstrahlung*. Subfigure (b) shows $e^+e^- \to e^+e^-$ scattering, also known as *Bhabha scattering*. . . . 8
2.3 Feynman diagram of (a) charged-current and (b) neutral-current weak interactions. The letters L and R denote left- and right-handed chiral states. . . . 11
2.4 The unitarity triangle arising from the equation $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ [13]. 14
2.5 Experimental constraints on the unitarity triangle in the $\rho - \eta$ plane [31]. The red hashed region around the top vertex of the triangle indicates the 68% confidence interval of the global combination. 15
2.6 Feynman diagrams of (a) quark-gluon interactions, (b) three-point gluon self-interactions, and (c) four-point gluon self-interactions. 17
2.7 Feynman diagrams representing $B^- \to D^0\ell^-\bar{\nu}_\ell$ and $B^0 \to D^-\ell^+\nu_\ell$ decays. 18
2.8 Experimental status of the observed discrepancy between exclusive and inclusive measurements of the CKM parameters $|V_{cb}|$ and $|V_{ub}|$ obtained from a global fit by the HFLAV group [3]. The experimental measurements and their global average are shown as colored bands and red ellipses, respectively. The inclusive measurement is shown as a black marker with error bars. 22

3.1 Schematic view of the SuperKEKB collider. Electron and positron beams collide at the interaction point located at the Belle II collider [48]. 24
3.2 Total recorded luminosity at the SuperKEKB collider during run 1 data-taking [50]. 26
3.3 Schematic view of the Belle II detector, with its subdetectors indicated with arrows [51, 52]. The coordinate frame is superimposed in the bottom left corner. 27
3.4 Schematic view of the Belle II detector (left) compared to the Belle detector (right) [53]. 28
3.5 (a) Schematic overview of the PXD sensors and their geometric alignment. The DEPFET sensors are shown in light-grey areas [54]. (b) Geometrical alignment of the silicon strips in the SVD. The slanted layout in the forward region reduces the amount of required sensors [54]. 29

3.6 (a) A cosmic muon being tracked in the CDC [52]. (b) Characteristic dE/dx energy loss curves in the CDC [57]. 30

3.7 (a) Functioning principle of the TOP detector, with photons radiated at the Cherenkov angle θ_c and subsequent propagation through the Quartz bar [54]. (b) Principle of the ARICH detector. The radius of the observed Cherenkov ring depends on the charged particles' mass [59]. 31

3.8 The histograms show the E/p distributions in the polar angle range $44° < \theta < 117°$ in the two momentum ranges (a) $500 < p < 750$ MeV and (b) $2000 < p < 3000$ MeV. E/p offers good discriminating power between electrons and pions/muons. [52] 35

3.9 Schematic overview of the individual layers that each gather information and feed forward into other layers to reconstruct the tag-side B meson [74]. . . 36

4.1 Stacked example histograms of an exponentially decaying background component and a Gaussian distributed signal component over a feature X. For each potential rectangular cut threshold $X > \lambda$, the right-hand side plot shows the FOM of the resulting sample. The optimal cut is shown as a red dashed line. 41

4.2 The plot shows a heatmap of an example signal distribution in two correlated variables X and Y. If overlaid on a uniform background distribution, the multivariate selection in the red ellipsis leads to a higher FOM than the optimized rectangular cuts in yellow. 42

4.3 (a) Schematic view of a decision tree. Each node represents a test in a feature V_i, with thresholds T_i. After traversing each test, every instance in the training set is classified into signal or background [78]. (b) The green line schematically represents an overfitted model distinguishing blue and red categories. Individual decision trees are prone to overfitting; boosting can mitigate the effect by combining multiple decision trees into a single classifier output [79]. 43

4.4 A fit to a toy data sample with a signal and background template. The histograms are shown before the fit on the left and after the fit on the right. Data is shown as black dots with error bars, and the expected templates from simulation as colored, stacked histograms. The overall normalization of the templates is left unconstrained, while shape variations are allowed in the templates according to their Poisson statistics. In this example, the signal yield in the toy data is significantly overestimated in simulation with respect to observed data. 44

5.1 Feynman diagrams for $e^+e^- \to e^+e^-$ in the s (left) and t (right) channels. . 51

5.2 Laboratory frame momentum and polar angle distributions for the probe in reconstructed Bhabha events, with a loose electronID selection on the tag. The y axis is scaled logarithmically. The left plots shows electrons, and the right plots positrons. 53

5.3 Differences in correction factors between control samples for multiple increasing thresholds from top to bottom. The correction factors shown are for positron momenta between 1 and 5 GeV in the barrel bin $\theta \in [1.31, 1.46]$ rad and different $\mathcal{P}_e^{\text{BDT}}$ selections. 60

5.4 Results of correction factors with statistical and systematic uncertainty, for $\mathcal{P}_e^{\text{BDT}} > 0.9$ and electrons. 61

5.5 Results of correction factors with statistical and systematic uncertainty, for $\mathcal{P}_e^{\text{BDT}} > 0.9$ and positrons. 62

6.1 A schematic overview of the steps involved in the measurement. After reconstructing $B \to D\ell\nu_\ell$ events in different final states in collision data, selection criteria are defined for optimal sensitivity. The modelling of signal and backgrounds is validated using control samples. Corrections are derived from control samples or external inputs to account for known and observed effects. Signal yields are extracted using a template fit, and the resulting differential decay rates are fit separately to the form factor. Using inputs from Lattice QCD in the form factor, $|V_{cb}|$ and additional parameters of interest are accessible. . 64

6.2 A schematic illustration of the geometrical event shape difference between continuum and $B\overline{B}$ events [88]. 68

6.3 Construction of the angle $\cos\theta_{BY}$. 69

6.4 The differential decay rate of $B \to D^*\ell\nu_\ell$ decays depends on the hadronic recoil w and the three angles θ_ℓ, θ_V, and χ, defined as shown in the figure. 70

6.5 The resolution of electron momentum for truth-matched electrons in both the charged and neutral B modes. The blue curve shows the resolution before applying bremscorrection, while the red curve shows the improved resolution when allowing to correct electrons with radiated photons. The legends denote the root mean square deviations of the candidates. 72

6.6 Distributions for the mass difference in each explicitly reconstructed $B \to D^*\ell\nu_\ell$ mode after optimization of selections. The modes in (a) and (b) are used to veto candidates in $B^+ \to \bar{D}^0\ell^+\nu_\ell$ reconstruction, while a veto of the mode in (d) is used for $B^0 \to D^-\ell^+\nu_\ell$. 74

6.7 The figure of merit after each cut for all four modes, integrated over the full w range, is shown. The overall figure of merit for the entire sample and the geometric mean of the FOMs in each w bin, which serves as the target in the selected annealing, are displayed. Cuts not shown on the x axis but described in the text have already been applied in the steering file. 77

6.8 Purities and efficiencies after each cut for all four modes, integrated over the full w range. 78

6.9 (a) Schematic of the diamond frame method, averaging over four directions. (b) Schematic of the ROE method, placing the B meson direction back-to-back with the p_{ROE}. 79

6.10 Resolutions for the kinematic variables w and $\cos\theta_{\ell,W}$. We compare approaches of ROE information only, diamond frame, and a combination of both. The combination results in the lowest width of the distributions. e and μ channels are combined for each B mode. 80
6.11 Migration matrices with each reconstruction approach for the charged and neutral B modes. The bin entries list the fraction of events reconstructed in each bin relative to their true origin. 82
6.12 The m_D distribution peaks at the nominal D meson mass, with some width due to detector resolution. Combinatorial false D backgrounds dominate the sidebands outside of the signal region and can, therefore, be used as a control sample. 83
6.13 The plot shows the stacked Δm distributions for candidates where a slow π^+ fulfilling the requirements stated in Section 6.4.3 is found. A $B \to D^*\ell\nu_\ell$ control sample can be defined by selecting candidates close to the peak. 84
6.14 The compositions of the true D component when reconstructing while requiring the (a) wrong charge and (b) right charge combinations. A bigger share of semileptonic decays is observed in the right charge reconstruction. Among hadronic decays, there are proportionally more $B \to DX$ decays in the wrong charge sample, particularly towards low w. 85
6.15 Distribution of the hadron-lepton system invariant mass in each control sample. The J/ψ mode has the highest amount of missing momentum and energy due to the missing second lepton in the reconstruction. 86
6.16 Data-MC agreements for the variables E_Y^* and m_Y in the control samples for the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode. 90
6.17 Correction factors derived from control samples in the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode. 91
6.18 Data-MC agreements for the variables E_Y^*, m_Y and $\cos\theta_{BY}$ in the nominal $B^+ \to \bar{D}^0 e^+ \nu_e$ sample before (top) and after (bottom) applying the reweight correction. 92
6.19 Distributions of selection variables in the $B \to J/\psi K^*$ control sample with electrons in data and MC. Uncertainties are only statistical. 93
6.20 Vertex fit probability distribution in the $B \to D\pi$ control sample before and after applying the correction calibrated on $B \to J/\psi K^*$ decays. The selection efficiency ratio improves to move within the statistical uncertainty of unity. 94
6.21 $\cos\theta_{BY}$ and w distributions in data and MC for all four modes. 94
6.22 p_ℓ^*, θ_ℓ and $\cos\theta_{\ell,W}$ distributions in data and MC for all four modes. 95
6.23 Data-MC ratios in bins of w in the background control samples. 96
6.24 The signal templates $\cos\theta_{BY}$ distributions in each bin of w flattened onto one combined axis. By categorizing the signal templates using w_{MC} rather than w_{reco}, migration effects between w bins are directly mapped into the fit. . 98
6.25 The uncertainty from particleID on the total signal yield N_{sig} as a function of assumed correlation coefficient ρ between systematic uncertainties bins in the correction tables, and a truncation order of the eigendecomposition that models the uncertainty. Full correlation $\rho = 1$ shows the largest uncertainty for all particles. 100

6.26 Pre- and post-fit plots for the fits to the background components in the control samples. The fit is performed separately for each control sample and for neutral and charged B modes. 102
6.27 The combined statistical and systematic correlation matrix between the 40 signal strengths. The indices correspond to 0-9 $B^+ \to \bar{D}^0 e^+ \nu_e$, 10-19 $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$, 20-29 $B^0 \to D^- e^+ \nu_e$ and 30-39 $B^0 \to D^- \mu^+ \nu_\mu$, ... 106
6.28 Correlation between BGL/BCL form factor parameters in the fit to MC. . . 108
6.29 The fitted BGL form factor with the extracted differential decay rates when fitting MC to itself. 109
6.30 Pull distributions for the toy fit results for the average $B^+ \to \bar{D}^0 \ell^+ \nu_\ell$ and $B^0 \to D^- \ell^+ \nu_\ell$ branching ratios, $|V_{cb}|$ with both BGL and BCL fits, and $R_{e/\mu}$. 110
6.31 $B \to D \ell \nu_\ell$ differential decay rates and the result of the BGL fit for all four final states. 117

A.1 Correction factors and uncertainties for BDT-based global electronID with different thresholds and charges. 122
A.2 Correction factors and uncertainties for BDT-based binary electronID with different thresholds and charges. 123
A.3 Correction factors and uncertainties for likelihood-based global electronID with different thresholds and charges. 124
A.4 Correction factors and uncertainties for likelihood-based binary electronID with different thresholds and charges. 125

B.1 Purity and signal efficiency in bins of w for the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode. 128
B.2 Purity and signal efficiency in bins of w for the $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$ mode. . . . 129
B.3 Purity and signal efficiency in bins of w for the $B^0 \to D^- e^+ \nu_e$ mode. 130
B.4 Purity and signal efficiency in bins of w for the $B^0 \to D^- \mu^+ \nu_\mu$ mode. . . . 131

C.1 All selections applied except the one shown in the plot for the $B^+ \to \bar{D}^0 e^+ \nu_e$ mode. 134
C.2 All selections applied except the one shown in the plot for the $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$ mode. 135
C.3 All selections applied except the one shown in the plot for the $B^0 \to D^- e^+ \nu_e$ mode. 136
C.4 All selections applied except the one shown in the plot for the $B^0 \to D^- \mu^+ \nu_\mu$ mode. 137

List of Tables

3.1 Cross-sections of processes occurring at the Belle II design beam c.m. energy. 25

4.1 Symbol notation in the `pyhf` notation [82]. 46
4.2 Relevant modifiers and constraints in the `pyhf` formulation [82]. The additional index c compared to the likelihood specification Equation (4.13) indicates a *channel*. Multiple channels, *e.g.*, decay modes or multiple dimensions, can be fit simultaneously. 47

5.1 Simulated data samples available for the Bhabha study. The ratio of N_{evt} and $\mathcal{L}_{\text{int}}$ differs due to run-dependent prescales. 54
5.2 The different sample selections according to their trigger requirements. In ee_flat, xx and yy refer to lower and upper limits of θ_e. They are separated into individual lines to allow theta-dependent prescale. 57
5.3 Bin edges of provided electronID correction factors, in radians and degrees, and characterization of the respective detector region. 58

6.1 Data luminosities separated by experiment. 65
6.2 The different physics processes simulated, listed with their corresponding cross section, integrated luminosity, and total number of events. 67
6.3 Final set of simulated annealing input variables deemed optimal in the feature preselection. 75
6.4 Rectangular selection optimized using simulated annealing. 76
6.5 Resulting yields for each component after applying all selections outlined in this section. 77
6.6 Differential signal efficiencies with all selections in place. 79
6.7 The resolution of the reconstructed kinematic variables with different methods and for different B modes. The new hybrid method results in the best overall resolutions. 81
6.8 Selection criteria for the control samples used to study signal distributions. 85
6.9 Calibration samples used to study deviations in particle identification performance between data and MC. 88
6.10 Branching ratios of $X_c\ell\nu$ decays in the decay file and the reweighed values. 89
6.11 Data/MC of selection efficiencies in the J/ψ control sample. We drop the selection on p^*_{miss} due to different efficiencies observed in Data and MC. . . 92

6.12 $\cos\theta_{BY}$ binning is determined by merging the lowest bins in the presence of empty bins. 97
6.13 An overview of the implemented nuisance parameters that assess the systematic uncertainties in the template fit. 99
6.14 The resulting uncertainties on background components obtained from a fit to control samples. 102
6.15 Fitted signal strength and total statistic plus systematic uncertainty when fitting the MC templates to itself. All fitted strengths are equal to one, equivalent to the MC expectation. 105
6.16 The resulting total branching ratios for each of the four modes, obtained by summing the signal yields while accounting for the correlation matrix, are consistent with the generator-level branching ratios. 105
6.17 The form factor parameters are obtained by minimizing the χ^2 function for the BGL fit applied to the signal yields from the simulation pseudo-data. The resulting values are consistent with the input parameters within the uncertainties. 108
6.18 Split up systematic uncertainty for the differential decay rates averaged over the four modes. 111
6.19 Split up systematic uncertainties on the overall branching ratios of the $B^+ \to \bar{D}^0\ell^+\nu_\ell$ and $B^0 \to D^-\ell^+\nu_\ell$ decays. 112
6.20 Breakdown of systematic uncertainties when measuring the ratio of the electron over muon decay rates for $B^+ \to \bar{D}^0\ell^+\nu_\ell$, $B^0 \to D^-\ell^+\nu_\ell$, and the average over all modes. 113
6.21 Breakdown of systematic uncertainties on V_{cb} when fitting both the BCL and the BGL parameterized form factor. 114
6.22 Result of the BGL fit to the partial widths and two lattice QCD calculations of the form factors f_+ and f_0 [38, 39]. The BGL series is truncated at $N = 3$. 116

Bibliography

[1] C. Glenn Boyd, Benjamín Grinstein, and Richard F. Lebed. „Model-independent determinations of $B \to D\ell\nu$, $D\ell\nu$ form factors“. In: *Nucl. Phys. B* 461.3 (1996), pp. 493–511. DOI: https://doi.org/10.1016/0550-3213(95)00653-2.

[2] C. Glenn Boyd, Benjamín Grinstein, and Richard F. Lebed. „Precision corrections to dispersive bounds on form factors“. In: *Phys. Rev. D* 56 (1997), pp. 6895–6911. DOI: 10.1103/PhysRevD.56.6895.

[3] Y. Amhis et al. „Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021“. In: *Phys. Rev. D* 107 (2023), p. 052008. DOI: 10.1103/PhysRevD.107.052008. arXiv: 2206.07501 [hep-ex].

[4] Robin Glattauer et al. „Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$“. In: *Phys. Rev. D* 93 (2016), p. 032006. DOI: 10.1103/PhysRevD.93.032006.

[5] Nicola Cabibbo. „Unitary Symmetry and Leptonic Decays“. In: *Phys. Rev. Lett.* 10.12 (1963), pp. 531533. DOI: 10.1103/physrevlett.10.531.

[6] Makoto Kobayashi and Toshihide Maskawa. „CP Violation in the Renormalizable Theory of Weak Interaction“. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[7] J. E. Augustin et al. „Discovery of a Narrow Resonance in e^+e^- Annihilation“. In: *Phys. Rev. Lett.* 33 (1974), pp. 1406–1408. DOI: 10.1103/PhysRevLett.33.1406.

[8] J. J. Aubert et al. „Experimental Observation of a Heavy Particle J“. In: *Phys. Rev. Lett.* 33 (1974), pp. 1404–1406. DOI: 10.1103/PhysRevLett.33.1404.

[9] S. W. Herb et al. „Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions“. In: *Phys. Rev. Lett.* 39 (1977), pp. 252–255. DOI: 10.1103/PhysRevLett.39.252.

[10] S. Abachi et al. „Observation of the Top Quark“. In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632.

[11] F. Abe et al. „Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab“. In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626.

[12] Belle II collaboration, F. Abudinén, et al. *Determination of $|V_{cb}|$ from $B \to D\ell\nu$ decays using 2019-2021 Belle II data*. 2022. arXiv: 2210.13143 [hep-ex].

[13] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03426-6. DOI: 10.1017/CBO9781139525367.

[14] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5, 978-0-429-50355-9, 978-0-429-49417-8. DOI: 10.1201/9780429503559.

[15] David Griffiths. *Introduction to elementary particles*. 2008. ISBN: 978-3-527-40601-2.

[16] Haim Harari. „Quarks and Leptons“. In: *Phys. Rept.* 42 (1978), pp. 235–309. DOI: 10.1016/0370-1573(78)90206-5.

[17] M. J. Herrero. *The Standard Model*. 1998. arXiv: hep-ph/9812242 [hep-ph].

[18] Wikimedia Commons. *Standard Model of Elementary Particles*. [Online; accessed March 19, 2024]. 2019. URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[19] S. Bilenky. „Neutrino oscillations: From a historical perspective to the present status“. In: *Nucl. Phys. B* 908 (2016). Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015, pp. 2–13. DOI: https://doi.org/10.1016/j.nuclphysb.2016.01.025.

[20] Kenneth G. Wilson. „Confinement of quarks“. In: *Phys. Rev. D* 10 (1974), pp. 2445–2459. DOI: 10.1103/PhysRevD.10.2445.

[21] Vladimir Gribov. „The theory of quark confinement“. In: *Eur. Phys. J. C* 10.1 (1999), pp. 91105. DOI: 10.1007/s100529900052.

[22] Daniel Bump. *Lie Groups*. Springer New York, 2013. ISBN: 9781461480242. DOI: 10.1007/978-1-4614-8024-2.

[23] F. Englert and R. Brout. „Broken Symmetry and the Mass of Gauge Vector Mesons“. In: *Phys. Rev. Lett.* 13.9 (1964), pp. 321323. DOI: 10.1103/physrevlett.13.321.

[24] Peter W. Higgs. „Broken Symmetries and the Masses of Gauge Bosons“. In: *Phys. Rev. Lett.* 13.16 (1964), pp. 508509. DOI: 10.1103/physrevlett.13.508.

[25] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. „Global Conservation Laws and Massless Particles“. In: *Phys. Rev. Lett.* 13.20 (1964), pp. 585587. DOI: 10.1103/physrevlett.13.585.

[26] R. P. Feynman. „The Theory of Positrons“. In: *Phys. Rev.* 76 (1949), pp. 749–759. DOI: 10.1103/PhysRev.76.749.

[27] Sheldon L. Glashow. „Partial-symmetries of weak interactions“. In: *Nuclear Physics* 22.4 (1961), pp. 579588. DOI: 10.1016/0029-5582(61)90469-2.

[28] R. L. Workman et al. „Review of Particle Physics“. In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[29] S. M. Bilenky. *Neutrino in Standard Model and beyond*. 2015. arXiv: 1501.00232 [hep-ph].

[30] W J Marciano. *Fermion masses and CKM mixing*. 1993. URL: https://www.osti.gov/biblio/10191720.

[31] J. Charles et al. „CP violation and the CKM matrix: assessing the impact of the asymmetric B factories". In: *Eur. Phys. J. C* 41.1 (2005). [and updates at: http://ckmfitter.in2p3.fr/.], pp. 1131. DOI: 10.1140/epjc/s2005-02169-1.

[32] Lincoln Wolfenstein. „Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett.* 51.21 (1983), pp. 19451947. DOI: 10.1103/physrevlett.51.1945.

[33] M. Bona et al. „The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD". In: *JHEP* 10 (2006). [and updates at: http://www.utfit.org/UTfit/WebHome.], p. 081. DOI: 10.1088/1126-6708/2006/10/081. arXiv: hep-ph/0606167.

[34] R. Aaij et al. „Observation of $J/\psi\, p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001.

[35] R. Aaij et al. „Observation of a Narrow Pentaquark State, $P_c(4312)^+$, and of the Two-Peak Structure of the $P_c(4450)^+$". In: *Phys. Rev. Lett.* 122 (2019), p. 222001. DOI: 10.1103/PhysRevLett.122.222001.

[36] C. T. H. Davies et al. „High-Precision Lattice QCD Confronts Experiment". In: *Phys. Rev. Lett.* 92.2 (2004). DOI: 10.1103/physrevlett.92.022001.

[37] S. Dürr et al. „Ab Initio Determination of Light Hadron Masses". In: *Science* 322.5905 (2008), pp. 1224–1227. DOI: 10.1126/science.1163233.

[38] Jon A. Bailey et al. „$B \to D\ell\nu$ form factors at nonzero recoil and $|V_{cb}|$ from $2+1$-flavor lattice QCD". In: *Phys. Rev. D* 92.3 (2015), p. 034506. DOI: 10.1103/PhysRevD.92.034506. arXiv: 1503.07237 [hep-lat].

[39] Heechang Na et al. „$B \to Dl\nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$". In: *Phys. Rev. D* 92.5 (2015). [Erratum: Phys. Rev. D **93**, 119906 (2016)], p. 054510. DOI: 10.1103/PhysRevD.93.119906. arXiv: 1505.03925 [hep-lat].

[40] Y. Aoki et al. http://flag.unibe.ch/2021/.

[41] Dante Bigi and Paolo Gambino. „Revisiting $B \to D\ell\nu$". In: *Phys. Rev. D* 94 (2016), p. 094008. DOI: 10.1103/PhysRevD.94.094008.

[42] Claude Bourrely, Laurent Lellouch, and Irinel Caprini. „Model-independent description of $B \to \pi l\nu$ decays and a determination of $|V_{ub}|$". In: *Phys. Rev. D* 79 (2009), p. 013008. DOI: 10.1103/PhysRevD.79.013008.

[43] Sinya Aoki et al. *Review of lattice results concerning low energy particle physics*. 2014. arXiv: 1310.8555 [hep-lat]. URL: https://arxiv.org/abs/1310.8555.

[44] Irinel Caprini, Laurent Lellouch, and Matthias Neubert. „Dispersive bounds on the shape of $B \to D^{(*)}\ell\nu$ form factors". In: *Nucl. Phys. B* 530.1 (1998), pp. 153–181. DOI: https://doi.org/10.1016/S0550-3213(98)00350-2.

[45] Matthias Neubert. „Heavy quark effective theory“. In: *Subnucl. Ser.* 34 (1997). Ed. by A. Zichichi, pp. 98–165. arXiv: hep-ph/9610266.

[46] D. Boutigny et al. *BABARtechnical design report*. 1995.

[47] A. Abashian et al. „The Belle Detector“. In: *Nucl. Instrum. Meth. A* 479 (2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[48] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. „SuperKEKB collider“. In: *Nucl. Instrum. Meth. A* 907 (2018), pp. 188199. DOI: 10.1016/j.nima.2018.08.017.

[49] Yoshihiro Funakoshi et al. „The SuperKEKB Has Broken the World Record of the Luminosity“. In: *JACoW* IPAC2022 (2022), pp. 1–5. DOI: 10.18429/JACoW-IPAC2022-MOPLXGD1.

[50] *Belle II Luminosity webpage*. https://www.belle2.org/research/luminosity/. Accessed: 2024-04-06.

[51] Michel Bertemes. „Search for dark sector physics and performance of the Belle II detector in final state events with muons and large missing energy“. en. PhD thesis. 2021. DOI: 10.34726/HSS.2021.85804.

[52] W. Altmannshofer et al. „The Belle II Physics Book“. In: *PTEP* 2019.12 (2019). Ed. by E. Kou and P. Urquijo. [Erratum: PTEP 2020, 029201 (2020)], p. 123C01. DOI: 10.1093/ptep/ptz106. arXiv: 1808.10567 [hep-ex].

[53] Belle-ECL et al. „Electromagnetic calorimeter for Belle II“. In: *Journal of Physics: Conference Series* 587 (2015), p. 012045. DOI: 10.1088/1742-6596/587/1/012045.

[54] T. Abe et al. „Belle II Technical Design Report“. In: (2010). arXiv: 1011.0352 [physics.ins-det].

[55] K. Adamczyk et al. „The design, construction, operation and performance of the Belle II silicon vertex detector“. In: *Journal of Instrumentation* 17.11 (2022), P11042. DOI: 10.1088/1748-0221/17/11/p11042.

[56] Christoph Schwanda, K. Adamczyk, et al. „Performance of the Belle II Silicon Vertex Detector“. In: *PoS* Vertex2019 (2020), p. 014. DOI: 10.22323/1.373.0014.

[57] Tom Browder. „Belle II Overview“. US Belle II Operations Review. 2020. URL: https://indico.bnl.gov/event/9843/.

[58] N. Taniguchi. „Central Drift Chamber for Belle-II“. In: *Journal of Instrumentation* 12.06 (2017), pp. C06014C06014. DOI: 10.1088/1748-0221/12/06/c06014.

[59] S. Iwata et al. „Particle identification performance of the prototype aerogel RICH counter for the Belle II experiment“. In: *Prog. Theor. Exp. Phys* 2016.3 (2016), 033H01. DOI: 10.1093/ptep/ptw005.

[60] L. Burmistrov et al. „Belle II Aerogel RICH detector“. In: *Nucl. Instrum. Meth. A* 958 (2020), p. 162232. DOI: 10.1016/j.nima.2019.05.073.

[61] Umberto Tamponi. „The TOP counter of Belle II: Status and first results“. In: *Nucl. Instrum. Methods Phys. Res.* 952 (2020), p. 162208. DOI: 10.1016/j.nima.2019.05.049.

[62] A. Kuzmin. „Electromagnetic calorimeter of Belle II“. In: *Nucl. Instrum. Methods Phys. Res.* 958 (2020), p. 162235. DOI: 10.1016/j.nima.2019.05.076.

[63] T. Aushev et al. „A scintillator based endcap K and muon detector for the Belle II experiment“. In: *Nucl. Instrum. Methods Phys. Res.* 789 (2015), pp. 134142. DOI: 10.1016/j.nima.2015.03.060.

[64] Yoshihito Iwasaki et al. „Level 1 trigger system for the Belle II experiment“. In: *IEEE Trans. Nucl. Sci.* 58 (2011), pp. 1807–1815. DOI: 10.1109/TNS.2011.2119329.

[65] S. Lee et al. „Belle-II high level trigger at SuperKEKB“. In: *J. Phys. Conf. Ser.* 396 (2012), p. 012029. DOI: 10.1088/1742-6596/396/1/012029.

[66] T. Kuhr et al. „The Belle II Core Software: Belle II Framework Software Group“. In: *Computing and Software for Big Science* 3.1 (2018). DOI: 10.1007/s41781-018-0017-9.

[67] David J. Lange. „The EvtGen particle decay simulation package“. In: *Nucl. Instrum. Meth. A* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. DOI: https://doi.org/10.1016/S0168-9002(01)00089-4.

[68] Torbjörn Sjöstrand et al. „An introduction to PYTHIA 8.2“. In: *Comput. Phys. Commun* 191 (2015), pp. 159–177. DOI: https://doi.org/10.1016/j.cpc.2015.01.024.

[69] C.M. Carloni Calame et al. „The BABAYAGA event generator“. In: *Nucl. Phys. B - Proceedings Supplements* 131 (2004). SIGHADO3, pp. 48–55. DOI: https://doi.org/10.1016/j.nuclphysbps.2004.02.008.

[70] S. Jadach et al. „The τ decay library TAUOLA, version 2.4“. In: *Comput. Phys. Commun* 76.3 (1993), pp. 361–380. DOI: https://doi.org/10.1016/0010-4655(93)90061-G.

[71] A. Arbuzov et al. „The Monte Carlo Program KKMC, for the Lepton or Quark Pair Production at LEP/SLC EnergiesUpdates of electroweak calculations“. In: *Comput. Phys. Commun* 260 (2021), p. 107734. DOI: https://doi.org/10.1016/j.cpc.2020.107734.

[72] S. Agostinelli et al. „Geant4 - a simulation toolkit“. In: *Nucl. Instrum. Meth. A* 506.3 (2003), pp. 250–303. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8.

[73] Wouter D. Hulsbergen. „Decay chain fitting with a Kalman filter“. In: *Nucl. Instrum. Meth. A* 552.3 (2005), pp. 566–575. DOI: https://doi.org/10.1016/j.nima.2005.06.078.

[74] T. Keck et al. „The Full Event Interpretation: An Exclusive Tagging Algorithm for the Belle II Experiment“. In: *Computing and Software for Big Science* 3.1 (2019). DOI: 10.1007/s41781-019-0021-8.

[75] In Jae Myung. „Tutorial on maximum likelihood estimation“. In: *Journal of Mathematical Psychology* 47.1 (2003), pp. 90100. DOI: 10.1016/s0022-2496(02)00028-7.

[76] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. „Optimization by Simulated Annealing“. In: *Science* 220.4598 (1983), pp. 671–680. DOI: 10.1126/science.220.4598.671.

[77] Paul Feichtinger. *selanneal*. https://github.com/feichtip/selanneal. [Accessed 28-05-2024].

[78] Thomas S. Adams et al. „Gravitational-wave detection using multivariate analysis“. In: *Phys. Rev. D* 88.6 (2013). DOI: 10.1103/physrevd.88.062006.

[79] Wikimedia Commons. *File:Overfitting.svg — Wikimedia Commons, the free media repository*. [Online; accessed 29-May-2024]. 2023. URL: https://commons.wikimedia.org/w/index.php?title=File:Overfitting.svg&oldid=793934668.

[80] Yoav Freund, Robert Schapire, and Naoki Abe. „A short introduction to boosting“. In: *Journal-Japanese Society For Artificial Intelligence* 14.771-780 (1999), p. 1612.

[81] Kyle Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., 2012. DOI: 10.17181/CERN-OPEN-2012-016.

[82] Lukas Heinrich, Matthew Feickert, and Giordon Stark. *scikit-hep/pyhf: v0.7.6*. 2024. DOI: 10.5281/ZENODO.10470033.

[83] B. Efron. „Bootstrap Methods: Another Look at the Jackknife“. In: *The Annals of Statistics* 7.1 (1979), pp. 1 –26. DOI: 10.1214/aos/1176344552.

[84] Aaron Roodman. *Blind Analysis in Particle Physics*. 2003. arXiv: physics/0312102 [physics.data-an].

[85] 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 [hep-ex].

[86] Ichiro Adachi et al. *Test of light-lepton universality in τ decays with the Belle II experiment*. 2024. arXiv: 2405.14625 [hep-ex].

[87] Claudia Cecchi et al. „B counting measurement of the Run 1 Belle II data sample“. In: (2023). URL: https://docs.belle2.org/record/3468?ln=en.

[88] Markus Röhrken. „Time-Dependent CP Violation Measurements in Neutral B Meson to Double-Charm Decays at the Japanese Belle Experiment“. en. In: (2012). DOI: 10.5445/IR/1000028856.

[89] Geoffrey C. Fox and Stephen Wolfram. „Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes“. In: *Phys. Rev. Lett.* 41 (1978), pp. 1581–1585. DOI: 10.1103/PhysRevLett.41.1581.

[90] A. J. Bevan et al. „The Physics of the B Factories“. In: *Eur. Phys. J. C* 74.11 (2014). DOI: 10.1140/epjc/s10052-014-3026-9.

[91] S. H. Lee et al. „Evidence for $B^0 \to \pi^0\pi^0$“. In: *Phys. Rev. Lett.* 91 (2003), p. 261801. DOI: 10.1103/PhysRevLett.91.261801.

[92] B. Aubert, R. Barate, et al. „Measurements of the $B \to D^*$ form factors using the decay $\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e$“. In: *Phys. Rev. D* 74.9 (2006). DOI: 10.1103/physrevd.74.092004.

[93] W. Dungel, C. Schwanda, et al. „Measurement of the form factors of the decay $B^0 \to D^{*-}\ell^+\nu$ and determination of the CKM matrix element $|V_{cb}|$“. In: *Phys. Rev. D* 82.11 (2010). DOI: 10.1103/physrevd.82.112007.

[94] I. Adachi et al. „Determination of $|V_{cb}|$ using $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays with Belle II“. In: *Phys. Rev. D* 108 (2023), p. 092013. DOI: 10.1103/PhysRevD.108.092013.

[95] Qi-Dong Zhou et al. „Correction for tracking momentum bias“. In: (2022). URL: https://docs.belle2.org/record/2843?ln=en.

[96] Raimundo Hoppe and Daniel Pitzl. „Correction for tracking momentum and energy bias based on D0 invariant mass peak for proc13 and prompt processing of buckets 26-36“. In: (2023). URL: https://docs.belle2.org/record/4027?ln=en.

[97] Matthew Rudolph. „An experimentalists guide to the semileptonic bottom to charm branching fractions“. In: *Int.J.Mod.Phys. A* 33.32 (2018), p. 1850176. DOI: 10.1142/S0217751X18501762.

[98] M. T. Prim et al. „Measurement of differential distributions of $B \to D^*\ell\nu$ and implications on $|V_{cb}|$“. In: *Phys. Rev. D* 108.1 (2023). DOI: 10.1103/physrevd.108.012002.

[99] Schmitt, Stefan. „Data Unfolding Methods in High Energy Physics“. In: *EPJ Web Conf.* 137 (2017), p. 11008. DOI: 10.1051/epjconf/201713711008.

[100] Andreas Höcker and Vakhtang Kartvelishvili. „SVD approach to data unfolding“. In: *Nucl. Instrum. Meth. A* 372.3 (1996), pp. 469–481. DOI: https://doi.org/10.1016/0168-9002(95)01478-0.

[101] G. D'Agostini. „A multidimensional unfolding method based on Bayes' theorem“. In: *Nucl. Instrum. Meth. A* 362.2 (1995), pp. 487–498. DOI: https://doi.org/10.1016/0168-9002(95)00274-X.

[102] S. et al. Choudhury. „Measurement of the B^+/B^0 production ratio in e^+e^- collisions at the $\Upsilon(4S)$ resonance using $B \to J/\psi(\ell\ell)K$ decays at Belle“. In: *Phys. Rev. D* 107 (2023), p. L031102. DOI: 10.1103/PhysRevD.107.L031102.

[103] Roel Aaij et al. „Measurements of the B^+, B^0, B_s^0 meson and Λ_b^0 baryon lifetimes“. In: *JHEP* 04 (2014), p. 114. DOI: 10.1007/JHEP04(2014)114. arXiv: 1402.2554 [hep-ex].

[104] Albert M Sirunyan et al. „Measurement of b hadron lifetimes in pp collisions at $\sqrt{s} = 8$ TeV“. In: *Eur. Phys. J. C* 78.6 (2018). [Erratum: Eur.Phys.J.C 78, 561 (2018)], p. 457. DOI: 10.1140/epjc/s10052-018-5929-3. arXiv: 1710.08949 [hep-ex].

[105] A. Sirlin. „Large $m(W)$, $m(Z)$ Behavior of the $\mathcal{O}(\alpha)$ Corrections to Semileptonic Processes Mediated by W“. In: *Nucl. Phys. B* 196 (1982), pp. 83–92. DOI: 10.1016/0550-3213(82)90303-0.

[106] Belle II Collaboration. *Measurement of Lepton Mass Squared Moments in $B \to X_c \ell \bar{\nu}_\ell$ Decays with the Belle II Experiment*. 2022. arXiv: 2205.06372 [hep-ex]. URL: https://arxiv.org/abs/2205.06372.

[107] Florian Bernlochner et al. *First extraction of inclusive V_{cb} from q^2 moments*. 2022. arXiv: 2205.10274 [hep-ph].

[108] Belle II Collaboration. *Measurement of the inclusive semileptonic B meson branching fraction in 62.8 fb^{-1} of Belle II data*. 2021. arXiv: 2111.09405 [hep-ex].