TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Measurement of the CKM matrix element $|V_{cb}|$ using the decay $B \to D^* \ell \nu_\ell$ at Belle II

zur Erlangung des akademischen Grades

Doktor der Technischen Wissenschaften

eingereicht von

Dipl.-Ing. Daniel Dorner
Matrikelnummer: 01128534

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

Wien, 6. August 2024

Daniel Dorner

Christoph Schwanda

Kurzfassung

Im Standardmodell der Teilchenphysik wird der experimentell beobachtete Übergang zwischen Quarks durch die Cabibbo-Kobayashi-Maskawa Matrix beschrieben. Dabei ist die Übergangsrate von Bottom- zu Charm-Quarks proportional zur Größe des Matrixelements $|V_{cb}|$. Darüber hinaus ist $|V_{cb}|$ einer der fundamentalen Parameter des Standardmodells, d.h. die Größe kann nicht allein durch Theorie bestimmt werden. Diese Arbeit zielt darauf ab, $|V_{cb}|$ aus Daten zu extrahieren, die durch das Belle II Experiment am SuperKEKB-Beschleuniger in Tsukuba, Japan, gewonnen wurden. Dieser Beschleuniger kollidiert Elektronen und Positronen mit einer Schwerpunktsenergie von 10,58 GeV und begann 2019 mit der Datennahme. Der verwendete Datensatz für diese Studie entspricht einer integrierten Luminosität von 364 fb^{-1}. Zum Zeitpunkt des Schreibens ist die Bestimmung von $|V_{cb}|$ durch semileptonische $B \to X_c \ell \nu_\ell$ Zerfälle der bevorzugte Ansatz. Das B steht für ein B-Meson, X_c für ein hadronisches System, das ein Charm-Quark enthält und ℓ ist ein geladenes, leichtes Lepton, d.h. entweder ein Elektron oder ein Myon mit dem zugehörigen Neutrino ν_ℓ. Diese Messungen können mit zwei verschiedenen Ansätzen durchgeführt werden: dem inklusiven und dem exklusiven Ansatz. Bei inklusiven Messungen werden alle $B \to X_c \ell \nu_\ell$ Endzustände rekonstruiert, während bei exklusiven Messungen explizit eine Zerfallskette untersucht wird, z.B. $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$. Obwohl man eine Übereinstimmung zwischen der Ergebnisse der beiden Methoden erwartet, gibt es eine lang anhaltende Spannung von 3,3σ zwischen inklusiven und exklusiven Messungen, was die Motivation hinter dieser Arbeit ist. Die Extraktion von $|V_{cb}|$ wird anhand der Rekonstruierung von $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$ mit den Unterzerfällen $D^{*+} \to D^0 \pi^+$ und $D^0 \to K^- \pi^+$ untersucht. Durch die Bestimmung der Anzahl der Signal-Kandidaten aus dem Datensatz können die partiellen Zerfallsraten $\Delta\Gamma / \Delta w \Delta \cos\theta_\ell$ bestimmt werden, wobei es sich bei w um den hadronischen Rückstoß handelt und bei $\cos\theta_\ell$ um eine kinematische Winkelvariable. Diese partiellen Zerfallsraten sind mit $|V_{cb}|$ verknüpft. Die Zusammenhänge können mittels zwei verschiedener Parametrisierungen beschrieben werden, der Boyd-Grinstein-Lebed und der Caprini-Lellouch-Neubert Parametrisierung. Das Matrixelement $|V_{cb}|$ und die Formfaktoren können für beide Parametrisierungen durch einen numerischen Fit an die partiellen Zerfallsraten gemessen werden. Mit dieser Methode messen wir $|V_{cb}|_{\mathrm{BGL}} = (38.100 \pm 0.194 \pm 0.754 \pm 0.674) \times 10^{-3}$ und $|V_{cb}|_{\mathrm{CLN}} = (38.582 \pm 0.247 \pm 0.844 \pm 0.560) \times 10^{-3}$ für die Boyd-Grinstein-Lebed und die Caprini-Lellouch-Neubert Parametrisierung anhand von Simulationen mit wettbewerbsfähiger Präzision. Die angegebenen Unsicherheiten in der Reihenfolge ihres Auftretens sind statistisch, systematisch und theoretisch.

Abstract

In the standard model of particle physics the experimentally observed quark mixing is described by Cabibbo-Kobayashi-Maskawa matrix. The transition rate from bottom to charm quarks is proportional to the matrix element magnitude $|V_{cb}|$. Additionally $|V_{cb}|$ is one of the fundamental parameters of the standard model, i.e. it can not be determined by theory alone. This thesis aims to extract $|V_{cb}|$ from data obtained by the Belle II experiment at the SuperKEKB collider in Tsukuba, Japan, which collides electrons and positrons at a center of mass energy of 10.58 GeV and started taking data in 2019. The data sample size used for this study is equivalent to an integrated luminosity of 364 fb^{-1}. At the time of writing determining $|V_{cb}|$ through semileptonic $B \to X_c \ell \nu_\ell$ decays is the favored approach. The B corresponds to a B meson, X_c a hadronic system containing a charm quark and ℓ is a charged, light lepton, i.e. either an electron or muon with its associated neutrino ν_ℓ. These measurements can be done using two different approaches, namely inclusive and exclusive. For an inclusive measurement all $B \to X_c \ell \nu_\ell$ final states are reconstructed, while exclusive measurements investigate a single decay chain, e.g. $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$. While both approaches use different theoretical inputs their results should be the same. However, there is a long standing 3.3σ tension between inclusive and exclusive measurements, which is the main motivation behind this thesis. In this thesis the extraction of $|V_{cb}|$ is studied by reconstructing $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$ decays through the subsequent decays $D^{*+} \to D^0 \pi^+$ and $D^0 \to K^- \pi^+$. By extracting the amount of signal candidates from the sample the partial decay rates $\Delta\Gamma / \Delta w \Delta \cos\theta_\ell$ in bins of the hadronic recoil w and the angular kinematic variable $\cos\theta_\ell$ can be determined. These partial decay rates are related to $|V_{cb}|$ and these relations can be described via two different parameterizations, i.e. the Boyd-Grinstein-Lebed and Caprini-Lellouch-Neubert parameterizations. The matrix element $|V_{cb}|$ and the form factors can be measured for both parameterizations via a numerical fit to the decay rates. Through this analysis strategy we measure $|V_{cb}|_{\mathrm{BGL}} = (38.100 \pm 0.194 \pm 0.754 \pm 0.674) \times 10^{-3}$ and $|V_{cb}|_{\mathrm{CLN}} = (38.582 \pm 0.247 \pm 0.844 \pm 0.560) \times 10^{-3}$, for the Boyd-Grinstein-Lebed and Caprini-Lellouch-Neubert parameterizations, respectively from simulations with a competitive precision. The provided uncertainties in order of appearance are statistical, systematic and theoretical.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Christoph, for his guidance, support and encouragement throughout my PhD journey. I sincerely appreciate the time he dedicated to discussions about various tasks during the course of my PhD, as well as the insights he provided and his openness to new ideas and approaches. His door was always open and he would provide his counsel online while on travels, whenever I needed help.

I want to thank all my colleagues and friends from the HEPHY Belle II group: Philipp, Paul, Nadia, Huw, Géraldine, Michel, Manca, Rajesh, Petar and Abdul. The friendly environment fostered by our group made my time at the institute a truly enjoyable and unforgettable experience. I appreciate the time we spent together, not only during lunch brakes, but also after work and during travels, especially when we explored Japan.

I would also like to express my gratitude to the semileptonic group and the tracking group of Belle II for their contribution to this work. A special note of appreciation goes to Florian and Chaoyi from the University of Bonn for sharing their knowledge. It was a pleasure to work with you at your institute during my research visit.

I am grateful for to the Belle II collaboration, which provided the data and the necessary computational resources to carry this analysis out and allowed me to experience the world of particle physics. I would like to acknowledge HEPHY for providing the necessary resources for this study. I would also like to extend my gratitude to the FWF for their financial support under the project number P34529-N, which made my doctoral research possible.

Finally, I want to sincerely thank my parents and siblings for always having my back. Without their unwavering support, encouragement and patience none of this would have been possible.

Contents

CHAPTER 1

Introduction

The fundamental particles and the interactions between them are described by the standard model of particle physics. The standard model has been developed in the second half of the 20th century. The model might not be perfect, since it doesn't account for the gravitational force and dark matter among others and can't fully describe the matter-antimatter asymmetry in the universe. Nevertheless through the last decades numerous predictions by the standard model have been confirmed by experimental data, such as the Higgs boson in 2012.[1] To this day, the standard model is being tested rigorously by a variety of particle collider experiments.

In the standard model the transition rate between quark flavors is described by magnitudes of the elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix.[2] The CKM matrix is based on the Cabibbo Matrix introduced by Nicola Cabibbo in 1963, which includes two generations of quarks. In 1973 Makoto Kobayashi and Toshihide Maskawa expanded the Cabibbo Matrix matrix to a third generation of quarks to explain observed symmetry violations in weak interactions.

Several so-called fundamental parameters are present in the standard model. These parameters are require experimental measurements and can't be calculated theoretically. One of these fundamental parameters is the CKM matrix element magnitude $|V_{cb}|$, which describes the $b \to c$ quark transition rate, where b is a bottom and c a charm quark. At the time of writing the favored approach of measuring $|V_{cb}|$ is studying semileptonic B meson decays. The magnitude $|V_{cb}|$ can be determined through two different approaches: inclusively and exclusively. For an inclusive approach one analyzes the sum of all possible $B \to X_c \ell \nu_\ell$ decays, where X_c can be any hadronic system containing a c quark and ℓ is a charged, light lepton, i.e. either an electron or muon with their corresponding neutrino ν_ℓ. On the other hand, an exclusive analysis reconstructs a specific decay chain, e.g. $B \to D^* \ell \nu_l$. While the theoretical inputs are different for both approaches, one would assume an agreement between both methods. However $|V_{cb}|$ exhibits a long lasting 3.3σ discrepancy between inclusive and exclusive measurements. This yet unexplained tension proves that our understanding of $|V_{cb}|$ and as a consequence also the standard model is still incomplete.

The discrepancy between inclusive and exclusive measurements of $|V_{cb}|$ motivates this doctoral thesis. We determine $|V_{cb}|$ via an exclusive approach through semileptonic $B \to D^* \ell \nu_l$ decays using the data recorded at the Belle II experiment. The Belle II experiment is situated at the circular electron-positron collider SuperKEKB in Tsukuba, Japan and started taking data in 2019. SuperKEKB collides electrons and positrons at a center of mass energy of 10.58 GeV to produce vast amounts of B mesons, thus SuperKEKB is also referred to as a B factory.

In Ch.2 of this thesis we will introduce the standard model of particle physics and the theoretical framework surrounding $|V_{cb}|$ and its extraction from $B \to D^* \ell \nu_l$ decays. This is followed by a description of the SuperKEKB accelerator and the components of the Belle II detector in Ch.3. We then provide an overview of a typical workflow and concepts for data analysis at Belle II in Ch.4, which includes the used data and simulations, as well as statistical concepts. In Ch.5 we will determine and discuss the calibration of the slow pion efficiency; a leading systematic uncertainty for $B \to D^* \ell \nu_l$ studies. The measurement of $|V_{cb}|$ is presented in Ch.6, followed by the conclusion of our study in Ch.7.

CHAPTER 2

Theoretical Overview

2.1 Natural Units

The system of natural units replaces the S.I. units of kg, s and m in favor of the reduced Planck constant $\hbar = 1.055 \cdot 10^{-34}$Js, the speed of light $c = 2.998 \cdot 10^{8}$ ms^{-1} and GeV, where 1 GeV $= 1.602 \cdot 10^{-10}$J. Additionally it is chosen that $\hbar = c = 1$. [3] This thesis adopts the natural units for the following particle related quantities: mass, momentum and energies to express them in powers of GeV.

2.2 Standard Model of Particle Physics

The Standard Model of particle physics (SM) describes the elementary particles and three out of four fundamental interactions between them, i.e. strong, electromagnetic and weak interactions, excluding gravity. The description of the SM is based on the concept of the quantum gauge field theory. Each of the three interaction types is associated with a gauge group: electromagnetic U(1), weak SU(2) and strong SU(3). All forces between particles are described by an exchange of force carriers referred to as gauge bosons. An illustrative summary of all elementary particles of the SM is provided in Fig.2.1.[3]

Matter consists of elementary particles called fermions, which are half spin $s = 1/2$ particles. Fermions are divided into two types: quarks and leptons. The categorization of the fermions is based on the charges they carry with each category consisting of six particles. These six particles are split into three generations with an increase in mass between every generation and two particles per generation. Additionally a corresponding anti-particle with equal mass but opposite charge exists for each fermion. In general the second and third generations have a very short mean life time, while the first generations are stable, excluding neutrinos: they are part of the lepton family and for them all three generations are stable. The two sub-classes of the fermions can be summarized as follows:[3]

- Leptons carry up to two out of three charges. The leptons, that carry both an electric and weak isospin charge are electrons (e), muons (μ) and tauons (τ), which correspond to the first, second and third generation, respectively. In addition an associated neutrino, that only carries the weak isospin charge, exists for each of these three particles.[3]

- Quarks carry electric, weak and color charges. The color charge is the charge of the strong interaction, which is described by quantum chromodynamics (QCD). Overall six different types of quarks exist: up (u), down (d), charm (c), strange (s), top (t) and bottom (or beauty, b), with up and down being the first generation. Quarks can only be found as color neutral objects called hadrons, which are mainly divided into two categories: the so-called mesons, that are quark and antiquark pairs ($q\bar{q}$) and systems consisting of three quarks (qqq), the baryons. Additionally, a variety of very rare exotic hadrons also exist, such as pentaquarks ($qqqq\bar{q}$).[3]

Interactions between particles are described by an exchange of gauge bosons, which are full integer spin $s = 1$ particles, excluding the Higgs boson with $s = 0$. There are three types of interactions: strong, electromagnetic and weak, where the electromagnetic and weak forces can be merged into the electroweak interaction of the $U(1) \times SU(2)$ group at energies of the order of ~ 250 GeV. A particle can only partake in an interaction, if it carries the corresponding charge. Consequently only quarks can be a part of each interaction type, charged leptons can take part in weak and electromagnetic interactions and neutrinos can only be a part of weak interactions. Specific bosons are assigned for each interaction type in the SM: [3]

- Electromagnetic: The electromagnetic interaction between electrically charged particles is described by quantum electrodynamics (QED). The photon originates from the $U(1)$ structure of QED as a chargeless and massless force carrier.[3]
- Weak: The weak interaction is part of the $SU(2)$ symmetry group and can be described via the exchange of an electrically charged $W^{\pm}$ boson with a mass of ~ 80 GeV or a charge neutral Z boson with a mass of ~ 91 GeV. Only the bosons of the weak interaction have a mass and couple particles that differ with respect to their electrical charge.[3]
- Strong: The $SU(3)$ symmetry of QCD generates 8 different massless gluons, which all carry a color charge and anti charge. The unique property of the gluons themselves carrying the charge of QCD interactions leads to the phenomenon called color confinement. This results in quarks being only observed in color neutral, bound hadronic states.[3]

2.3 Global Symmetries

In general symmetries describe, that the physical properties of a system don't change when applying a given transformation. There are two different kinds of symmetries, i.e. continuous and discrete. E.g. with respect to rotation a square has a discrete symmetry, while a circle has a continuous one.[3]

All fundamental interactions of the SM are described using quantum field theories (QTF), which are formulated using special relativity. Consequently, it is implied, that the SM obeys the Poincaré symmetry, which leads to three continuous symmetries: translation, rotation

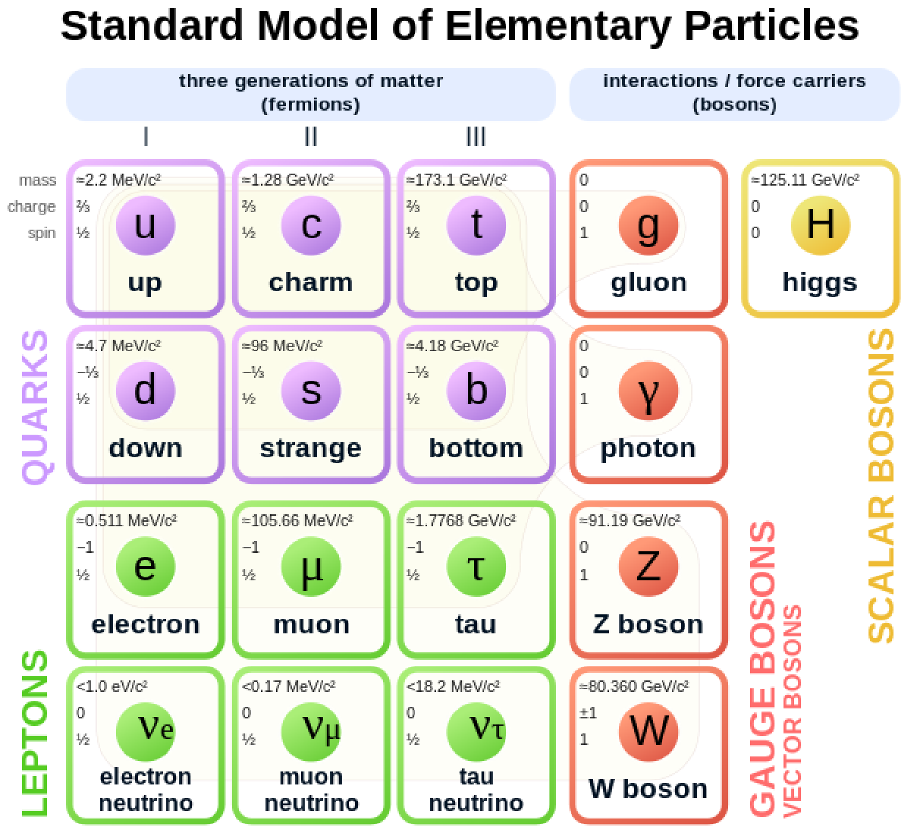


Figure 2.1: Classification of all elementary particles according to the SM.[4]

and boosts. Furthermore, Noether's theorem states, that every symmetry is linked to a conservation law.[5] The three discrete symmetries are the following:[3]

- Charge conjugation (C): Inverts all internal, discrete quantum numbers of a particle, basically creating the antiparticle. This affects electric charges, color charges, lepton numbers, baryon numbers etc., but does not affect mass or momenta.
- Parity transformation (P): Mirrors all spatial coordinates through the point of origin, i.e. $\vec{x} \to -\vec{x}$, where $\vec{x}$ is a vector of coordinates.
- Time revearsal (T): Reverses the time components $t \to -t$.

These discrete symmetry operations can be either applied individually or combined. If a symmetry operation changes the physical observables of a system, one speaks of a symmetry violation. It can be shown, that all QTF theories uphold a symmetry for combined CPT transformation. However, if all fundamental interactions between particles would be symmetric under all three individual transformation there would be an equal amount of matter and antimatter in the universe, which would lead to an annihilation of all matter. Therefore it can be concluded, that some form of symmetry violation must exist for at least one interaction. It is known, that QED and QCD interactions conserve C and P symmetries

Figure 2.2: Two exemplary Feynman diagrams of QCD interactions: Fig(a) shows a $rb \rightarrow br$ process, while (b) shows one of the lowest order Feynman diagrams of the $gg \rightarrow gg$ process.[3]

separately and therefore also the CP symmetry, which leaves only the weak interaction to contribute to the matter/antimatter asymmetry. The weak interaction in fact maximally violates C and P transformations individually and additionally, violates the CP symmetry to some extend.[3]

2.4 Strong Interaction

The charge related to the strong interaction is the so-called color charge and its associated quantum numbers are: r, g, b for particle or $\bar{r}$, $\bar{g}$, $\bar{b}$ for the antiparticle charges. Due to the charge naming convention, the QTF of the strong interaction is referred to as quantum chromodynamics (QCD). The only particles carrying color charge are quarks and the gauge bosons mediating the force of QCD, i.e. the massless gluons. Exemplary Feynman diagrams of QCD processes are given in Fig.2.2. Feynman diagrams are graphical representations of the mathematical description behind a particle physics processes and follow a given set of rules. Quarks can not be observed as individual free particles, but only in color neutral or "white" composite systems referred to as hadrons. The most common hadrons are quark antiquark pairs ($q\bar{q}$) and three quark systems (qqq) referred to as mesons and baryons, respectively. Besides mesons and baryons, there exists a variety of exotic hadrons, e.g. the recently discovered pentaquarks ($qqqq\bar{q}$).[3, 6, 7]

The phenomenon of the absence of individual propagating quarks is accredited to color confinement, which is a consequence of the color charge of gluons. Because the gluons carry color charge a self interaction between gluons is caused. The energy of the gluon field between the quarks increases proportional to the quark distance, making it impossible for them to separate from each other and propagate individually. Due to the color confinement $q\bar{q}$ events produced in high energy collisions are observed in the form of hadronisation jets. When the $q\bar{q}$ pair is produced during the collision, the q and $\bar{q}$ travel back to back into opposite directions: with increasing distance between the quark and antiquark, the energy stored in the color field between them also increases and produces new $q\bar{q}$ pairs, which leads to an unique topology for $q\bar{q}$ events.[3]

Another property of QCD is its large coupling constant $\alpha_s \sim \mathcal{O}(1)$ at lower energies, resulting in the inability to calculate QCD processes using perturbation theories. However, α_s is a running coupling strength that shows a dependence on the energy scale of the interaction. This property makes it possible to apply perturbation theories at higher energy scales, a phenomenon referred to as asymptotic freedom. The decrease of the coupling strength shows the dependence:[3, 8, 9]

$$\alpha_s(|q^2|) \propto \frac{1}{\beta_0 ln(|q^2|/\Lambda^2)}, \tag{2.1}$$

where $\beta_0 = 11 - \frac{2}{3}n_f$ is the first term of the beta series, with n_f being the number of flavors active at the scale of $|q^2|$ and Λ is the QCD scale parameter. This property separates α_s into a non perturbative low energy regime and a perturbative high energy regime.[3, 8, 9]

A common technique to calculate QCD processes at low energies without perturbation theory is the lattice QCD (LQCD) approach, where the strong interaction processes are calculated on four dimensional Euclidean lattice points in space time. The points of the lattice are inhabited by the quark fields and these sites are in turn connected by the gauge fields. The QCD need to be discretized for LQCD calculations. These discretized approximations are formulated in such a way, that a reduction of the spacing between lattice site to zero leads back to continuous QCD. By discretizing the gauge field theories of QCD to the lattice LQCD allows for numerical determinations of QCD effects in the low energy scale. The numerical LQCD calculations are typically done using Monte Carlo (MC) simulations and come at extremely high computational costs, which increase as the lattice spacing decreases. [10]

2.5 Weak Interaction

The weak interactions differs from various properties that the electromagnetic and strong interactions have in common. First of all, the force carriers of the weak interaction are $W^{\pm}$ and Z bosons, which have a mass of 80.377 ± 0.012 GeV and 91.188 ± 0.002 GeV, respectively.[10] The gauge bosons of QED and QCD are mass less. Additionly the $W^{\pm}$ bosons are the only gauge bosons to carry an electric charge, therefore the weak interaction is the only possible interaction between particles with a different electrical charge. Lastly, the weak interaction maximally violates C and P symmetry individually and also to some extend the CP symmetry. This is not the case for QED and QCD and therefore it is the only source for the matter antimatter asymmetry between these three fundamental interactions.[3]

The P symmetry violation of the weak interaction was first observed in 1957 at the Wu experiment, which studied the emission of electrons from β-decays of polarized ^{60}Co nuclei.[11] In general physical quantities can be categorized into four classes, based on their dimension and sign change under parity transformations:[3]

- Scalars: Single valued quantities, e.g. mass or time, that don't change their sign under parity transformations.
- Vectors: Multi dimensional quantities like momenta, which change their sign under parity transformation.
- Axialvectors: Are the cross products of two vectors and therefore don't change their sign, e.g. angular momenta. They are also referred to as pseudovectors.
- Pseudoscalars: The product of a vector and axialvector, which results in single valued quantities, e.g. helicity, that change their sign under parity transformations.

While QED and QCD currents are vectors, the necessary Lorentz transformation properties and observations from experiments lead to a (V-A) structure for the weak charged currents of $W^{\pm}$ boson exchanges.[3]

Another important concept, needed for the P symmetry violation of the weak interaction is chirality: particles can either have left or right handed chiral eigenstates. Chirality is connected to the helicity of a particle, i.e. a normalized spin component in the direction of the particle's momentum, that can also be either right handed or left handed. The right handed helicity spinor $u_{\uparrow}$ and left handed helicity spinor $u_{\downarrow}$ can be expressed in terms of left handed and right handed chirality components:[3]

$$u_{\uparrow} = \frac{1}{2}(1+\frac{p}{E+m})u_R + \frac{1}{2}(1-\frac{p}{E+m})u_L, \tag{2.2}$$

$$u_{\downarrow} = \frac{1}{2}(1+\frac{p}{E+m})u_L + \frac{1}{2}(1-\frac{p}{E+m})u_R, \tag{2.3}$$

where u_R and u_L are the right handed and left handed chirality components, respectively. From Eq.2.2 and Eq.2.3 one can see, that in high energy environments, where $E >> m$ the chirality eigenstates are equal to the helicity eigenstates. It can be shown, that, with respect to the chirality, only right handed particles and left handed antiparticles are able to take part in the weak interaction, which is the origin of the P symmetry violation. By applying a parity transformation the sign of the momenta of all particles gets inverted, turning a particle with left handed helicity into a right handed particle and a right handed antiparticle into a left handed anti particle. Since the weak interaction only couples to left handed particles and right handed anti particles, the previous interaction is not possible anymore after the application of the transformation, leading to a maximal violation of P symmetry. At lower energies or more massive particles the coupling with a particle with right handed helicity is possible, but very unlikely in comparison to the original process, still violating P symmetry.[3]

2.5.1 CP Violation

The main motivations behind the Belle experiment were to further investigate the CP violation and the quark mixing described by the Cabbibo Kobayashi Maskawa matrix. The CP violation of the weak interaction was first observed at an experiment measuring neutral kaon decay rates. The neutral kaons K^0 ($d\bar{s}$) and $\bar{K}^0$ ($\bar{d}s$) are the lightest mesons containing an s quark and differ from each other to conserve strangeness in strong interactions. However, K^0 and $\bar{K}^0$ are known to mix via the weak interaction, which leads to each K^0 developing a $\bar{K}^0$ component and vice versa. In the weak interaction of neutral kaons, the physical states are eigenstates of a K^0-$\bar{K}^0$ Hamiltonian, which are referred to as K^0 long (K^0_L) and K^0 short (K^0_S), due to their different lifetimes of $\tau_{K_L} = (5.116 \pm 0.021) \times 10^{-8}$ s and $\tau_{K_S} = (8.954 \pm 0.004) \times 10^{-1}$ s.[10] The dominant decays of $K_S \to \pi\pi$ and $K_L \to \pi\pi\pi$ differ in parity, namely CP($\pi\pi$) =+1 and CP($\pi\pi\pi$)=-1. The conservation of CP symmetry would imply a correspondence of life time to CP eigenstates, where $K^0_S = K_1$ and $K^0_L = K_2$, however $K_L \to \pi\pi$ decays could be observed by an experiment of Christenson et al. in 1964, proving the existence of CP violation.[3, 12]

In 1967, three years after the measurement of the CP violation in K^0 decays, Sakharov published a paper on baryogenesis, listing three conditions needed for the present matter antimatter asymmetry in our universe.[3, 13]

- Baryon number violation: The number of baryons is not constant.
- C and CP violation: Without C and CP violation there would be an equivalent amount of antiparticle decays, leading to a net zero.
- Departure from thermal equilibrium: Thermal equilibrium would imply, that all baryon number violating processes would be canceled out by an inverse process.

2.5.2 Cabibbo Kobayashi Maskawa Matrix

In the 1960s the observed decay rates of $K^-(\bar{u}s) \to \mu^-\bar{\nu}_\mu$ and $\pi^-(\bar{u}d) \to \mu^-\bar{\nu}_\mu$ deviated from the value expected from a universal coupling strength of the weak interaction to quarks and lepton flavor universality (LFU). This discrepancy was explained using a hypothesis proposed by Cabibbo in 1963, which states that the flavor eigenstates of quarks are not equal to their weak eigenstates involved in the weak interaction, but that the two eigenstates are related to each other by a rotation described by the cabibbo angle θ_c. [14] According to Cabibbo's hypothesis the two decay rates are linked to each other by a factor $\tan^2\theta_c$, which solves the issue by taking an angle of $\theta_c \approx 14°$ into account for the suppression of the K^- decay, therefore saving the LFU prediction for the weak interaction.[3]

When the Cabibbo hypothesis was first proposed only three quarks were known, i.e. u, d and s and it introduced ud and us couplings. However the measured branching fraction of $\mathcal{B}(K_L^0 \to \mu + \mu^-) = (6.84 \pm 0.11) \times 10^{-9}$ of the leptonic decay $K_L^0 \to \mu + \mu^-$ was lower than expected by theory.[3] By introducing a yet undiscovered, fourth quark the GIM mechanism proposed in 1970 would solve this issue by explaining the highly suppressed decay rate with a coupling between c and s quarks.[15] One would then need to sum up the decay amplitudes from the couplings, resulting in a suppressed branching fraction close to 0. The c quark was discovered in 1974 shortly after the proposition of the GIM mechanism. The resulting Feynman diagrams of the $K_L^0 \to \mu + \mu^-$ decay are shown in Fig.2.3. By combining the cabibbo hypothesis and the GIM mechanism the relation between the flavor eigenstates and the weak eigenstates of the four quarks can be written as a unitary 2×2 matrix, referred to as Cabibbo matrix:[3]

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} cos\theta_c & sin\theta_c \\ -sin\theta_c & cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{2.4}$$

where d', s' are the weak eigenstates and d, s the flavor eigenstates.

The Cabibbo matrix still has one major shortcoming: it can't explain the CP violation mandatory for the matter over antimatter asymmetry in the universe. To incoporate the CP vioaltion into the SM Kobayashi and Maskawa predicted a third generation of quarks in 1973. By introducing b and t quarks, the 2x2 Cabibbo matrix naturally expands to the unitary 3x3 Cabibbo-Kobayashi-Maskawa (CKM) matrix: [2, 3]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.5}$$

The magnitude of the CKM matrix elements $|V_{ij}|$ correspond to the transition rate between two quark flavors, e.g. $|V_{cb}|$ gives the rate for $b \to c$ transitions. Therefore the CKM matrix also allows for transitions between different generations. The CKM matrix can be described

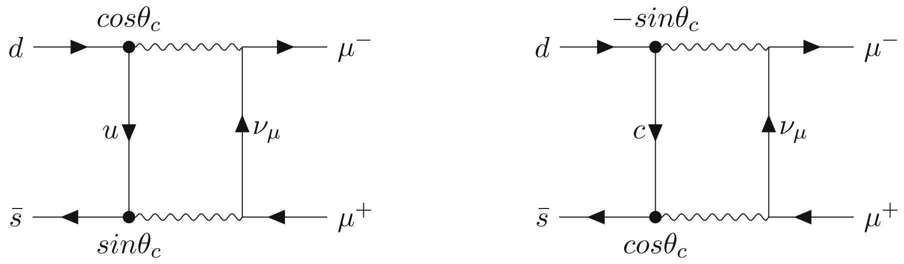


Figure 2.3: The two leading order Feynman diagrams of the $K_L^0 \rightarrow \mu + \mu^-$ decay. For the decay rate the sum of both processes has to be considered, leading to the suppression of the decay due to the cabibbo angle. [3]

by three Euler angles ϕ_{12}, ϕ_{13}, ϕ_{23} and the introduction of a complex phase δ, which is the source for CP violation: [16]

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.6}$$

where s_{ij} and c_{ij} correspond to sinϕ_{ij} and cosϕ_{ij}, respectively. While transitions between different generations are possible, they are suppressed in comparison to transitions within the same generation. This behavior is illustrated by the values of CKM matrix magnitudes $|V_{ij}|$, where the diagonal elements are close to 1: [10]

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.042 \\ 0.009 & 0.041 & 0.999 \end{pmatrix}. \tag{2.7}$$

The CKM matrix can be parameterized using the Wolfenstein parameters A, λ, ρ, η, by defining:[16, 17]

$$s_{12} \equiv \lambda, \;\; s_{23} \equiv A\lambda^2, \;\; s_{13}e^{-i\delta} = A\lambda^3(\rho - i\eta) \tag{2.8}$$

where $\lambda \approx 0.22$. Resulting in: [16]

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{2.9}$$

The unitarity conditions of the CKM matrix $V_{CKM}V_{CKM}^\dagger = 1$ and $V_{CKM}^\dagger V_{CKM} = 1$, where $\dagger$ denotes the conjugate transpose. The equations emerging from the unitarity conditions can be used to construct unitarity triangles (UT) in a complex plane. These equations have the property, that proving one equation leads to all equations being correct. The following equation is commonly used for B meson physics: [10, 16]

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.10}$$

To create the UT used for B physics Eq.2.10 is divided by $V_{cd}V_{cb}^*$, which yields:

$$1 + \frac{V_{ud}^*V_{ub}}{V_{cd}V_{cb}^*} + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0 \tag{2.11}$$

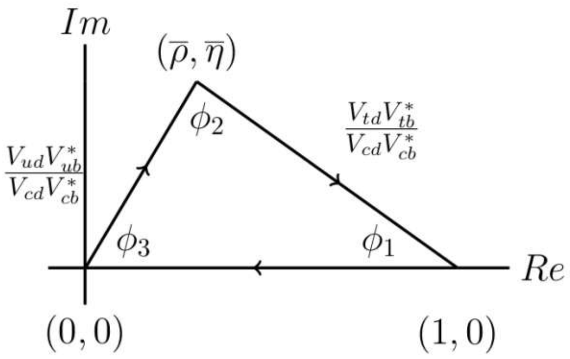


Figure 2.4: Schematic of the unitarity triangle. [16]

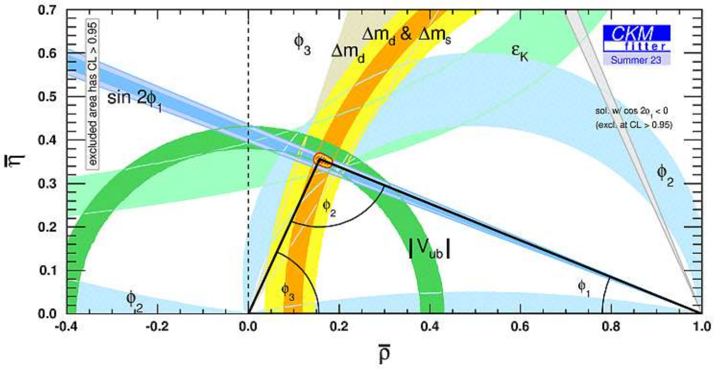


Figure 2.5: Current state of the unitarity triangle. [20]

Additionally, the the new parameters $\bar{\eta}$ and $\bar{\rho}$ are introduced, which are defined as: [18]

$$\bar{\rho} + i\bar{\eta} \equiv -\frac{V_{ud}^* V_{ub}}{V_{cd}V_{cb}^*}, \quad 1 - (\bar{\rho} + i\bar{\eta}) \equiv -\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \tag{2.12}$$

and are linked to the Wolfenstein parameters. By utilizing Eq.2.11 and the new parameters from 2.12, the UT in Fig.2.4 can be constructed in the complex $\bar{\rho} - \bar{\eta}$ plane.

According to the SM predictions and the unitarity of the CKM matrix, the UT has to be closed. This property provides a crucial opportunity to probe SM predictions, since precise measurements of angles and sides could result in an opening of the triangle, which would indicate new physics beyond the SM. [16]

The measurements of the B meson triangle complement each other. Measurements of time dependent CP asymmetries of B meson oscillations are used to determine the angle ϕ_1. Additionally one uses $b \to c$ and $b \to u$ transitions to measure $|V_{cb}|$ and $|V_{ub}|$, respectively: $|V_{cb}|$ normalizes the UT and the ratio of $|V_{ub}|/|V_{cb}|$ determines the length of the side opposite of ϕ_1. By increasing the precision of the measurements of the CKM matrix element magnitudes and other observables restricting the triangle, one can tighten the constraints of the corners. The current state of the UT is illustrated in Fig.2.5. [10, 16, 19]

Particle	Quarks	Spin and Parity (J^P)	Rest Mass [GeV]	Mean Life Time [s]
B^0	$d\bar{b}$	0^-	$5.279 \pm 0.120 \cdot 10^{-3}$	$(1.519 \pm 0.004) \cdot 10^{-12}$
B^+	$u\bar{b}$	0^-	$5.279 \pm 0.120 \cdot 10^{-3}$	$(1.638 \pm 0.004) \cdot 10^{-12}$
B_s^0	$s\bar{b}$	0^-	$5.367 \pm 0.100 \cdot 10^{-3}$	$(1.527 \pm 0.011) \cdot 10^{-12}$
B_c^+	$c\bar{b}$	0^-	$6.274 \pm 0.320 \cdot 10^{-3}$	$(0.510 \pm 0.009) \cdot 10^{-12}$

Table 2.1: A summary of the basic properties of B meson ground states. [10]

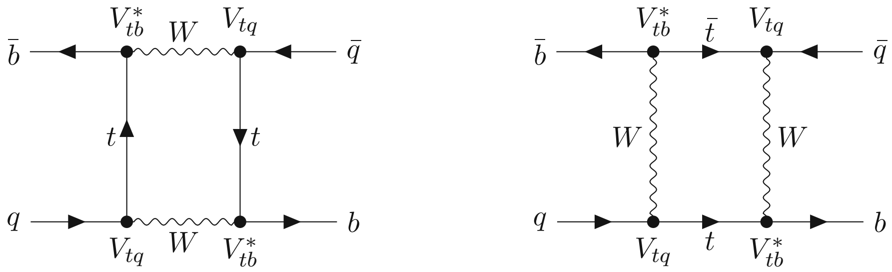


Figure 2.6: Leading order Feynman diagrams of neutral B^0 meson oscillations. [10].

2.5.3 B Mesons

A particle that provides an excellent opportunity to probe the CKM sector of the SM is the B meson. B mesons contain a third generation $b(\bar{b})$ quark with a mass of $\sim$ 4.18 GeV and an electric charge of $-\frac{1}{3}(+\frac{1}{3})$. There exists a variety of excited B meson states, however the most commonly observed ground states are B^0, B^+, B_s^0 and B_c^+. Due to their relatively high mass above 5 GeV, all B mesons decay rapidly with a mean life time in the order of $\mathcal{O}(10^{-12})$ s. The basic properties of the four ground states are summarized in Tab.2.1. At colliders B mesons can be produced via the $\Upsilon(4S)$ meson resonance with a mass of $\sim$ 10.58 GeV and a quark content of $b\bar{b}$, whose decay rate is heavily dominated by decays into $B^0\bar{B}^0$ or B^+B^- meson pairs with a rate above 96%. [10, 16, 19]

For the neutral B mesons $B_q^0 - \bar{B}_q^0$ $(q = d, s)$ mixing occurs due to the weak interaction, similar to the mixing of neutral kaons mentioned in Sec.2.5.1. At the lowest order, the $B_q^0 - \bar{B}_q^0$ oscillations occur by exchanging two W bosons and two t quarks, leading to the two dominant box diagrams in Fig.2.6. [10]

2.6 Semileptonic Decays

Measuring fundamental parameters, with high precision is crucial to probing the SM predictions. Two of the fundamental parameters of the SM are the CKM matrix element magnitudes $|V_{cb}|$ and $|V_{ub}|$, which can be measured via decays, that include $b \to c$ or $b \to u$ quark transitions. Another important feature of the magnitudes $|V_{cb}|$ and $|V_{ub}|$ is that they complement B meson CP asymmetry measurements, since $|V_{cb}|$ normalizes the unitarity triangle and the ratio of $\frac{|V_{ub}|}{|V_{cb}|}$ is proportional to the length of the side opposite of the ϕ_1 angle. Both $|V_{cb}|$ and $|V_{ub}|$ are equally important measurements, however this thesis will be focusing on decays involving $b \to c$ transitions and measuring $|V_{cb}|$. In principle there are three different types of decays one could analyse to determine CKM matrix element magnitudes: hadronic decays, leptonic decays and semileptonic decays. [19, 16]

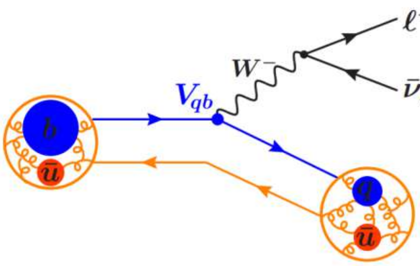


Figure 2.7: Feynman diagram of semileptonic B meson decays. [19]

In hadronic decays the B meson decays purely to hadronic daughters, e.g. $B \to D\pi$ decays. While hadronic decays are statistically not very limited and have rather high branching fractions, e.g. $(2.51 \pm 0.08) \cdot 10^{-3}$ for $B \to D\pi$, they suffer from large theoretical uncertainties from QCD calculations originating from the inability of factorize the two hadronic currents. [10, 19]

Leptonic decays, i.e. decays only involving lepton daughters e.g. $B \to e\nu_e$, don't suffer from the large uncertainty inflicted by the hadronic currents, however these decays are rare compared to hadronic decays and have very low branching fractions, e.g. $\mathcal{B}(B \to \mu\nu_\mu) < 8.6 \cdot 10^{-7}$.[10] For the current data samples sizes the very low branching fractions result in high statistical uncertainties severely limiting the precision of leptonic measurements making them not competitive for CKM measurements. Once more data hast been taken at Belle II the measurements of CKM magnitudes from leptonic decays will be competitive in terms of precision.[19]

With the contributions from one hadronic current being under control and high branching fractions, the semileptonic $B \to X_c\ell\nu_\ell$ decays provide the best opportunity for high precision measurements of $|V_{cb}|$. Here X_c refers to a hadronic system containing a c quark, ℓ a light charged lepton, i.e. either an electron e or muon μ and the neutrino ν_ℓ associated with the given lepton. The neutrinos do not interact with the matter of detectors and can't be detected directly. Additionally $B \to X_c\ell\nu_\ell$ decays involving light leptons decay via electroweak tree-level W transitions of $b \to c\ell\nu_\ell$, illustrated in the Feynman diagram in Fig.2.7, and are therefore expected to be free from non-SM contributions. Finally, the decays only involve one CKM matrix element, due to the leptonic decay of the W boson, leading to a critical role for probing the CKM predictions of the SM. Semileptonic decays are also used to test lepton flavor universality (LFU) of the electroweak interaction by measuring branching fraction ratios $\mathcal{B}(B \to X_c e\nu_e)/\mathcal{B}(B \to X_c\mu\nu_\mu)$. Semileptonic $B \to X_c\tau\nu_\tau$ decays on the other hand are susceptible to influences of physics beyond the SM like charged Higgs bosons, due to the large mass of the τ lepton. [16, 19]

The CKM matrix elements can be measured using two different approaches, i.e. inclusively and exclusively. In case of an inclusive determination, one measures the decay rates of $B \to X_c\ell\nu_\ell$ by summing over all possible decays involving a X_c, while exclusive decays reconstruct a specific decay, e.g. $B \to D^*\ell\nu_\ell$. [16, 19]

Both types use a different theoretical framework to determine the contribution of the hadronic current. In both cases one makes use of the large mass of the b quark $m_b \approx 4.18$ GeV. For inclusive measurements the QCD contribution is described using the operator product expansion (OPE) and the heavy quark expansion (HQE), expanding the QCD contribution in powers of $(\frac{1}{m_b})$. In exclusive measurements the hadronic contribution is parameterized in

	$\lvert V_{cb}\rvert$	$\lvert V_{ub}\rvert$
Exclusive	$(39.10 \pm 0.50) \cdot 10^{-3}$	$(3.51 \pm 0.12) \cdot 10^{-3}$
Inclusive	$(42.19 \pm 0.78) \cdot 10^{-3}$	$(4.19 \pm 0.17) \cdot 10^{-3}$

Table 2.2: Current world average values of the magnitudes $|V_{cb}|$ and $|V_{ub}|$, provided by the HFLAV collaboration. [21]

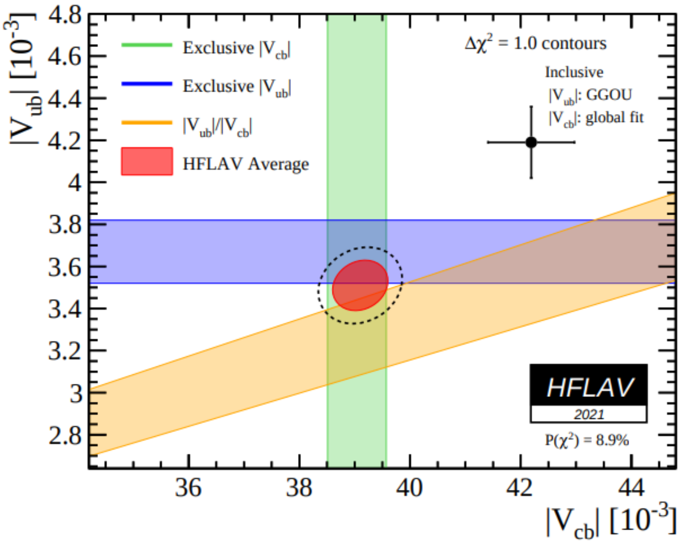


Figure 2.8: Graphical comparison between the world average of exclusive and inclusive measurements of $|V_{cb}|$ and $|V_{ub}|$, provided by HFLAV. [21]

terms of form factors, which are non perturbative functions of the momentum transfered to the leptons q^2. These form factors can be calculated using LQCD or light cone sum rules (LCSR), with LQCD generally being the favored approach, due to better quantifiable systematic uncertainties. In general $|V_{cb}|$ is determined by measuring the partial decay rates, which are calculated from the partial branching fractions and the mean life time of the B mesons. [10, 19]

In theory the values of $|V_{cb}|$ from exclusive and inclusive measurements should agree, however the current world averages of $|V_{cb}|$ and $|V_{ub}|$ listed in Tab.2.2 and illustrated in Fig.2.8 show a long standing 3.3σ tension between these measurements. This shows that our understanding of the CKM matrix is still incomplete. [16, 19]

2.6.1 Decay Rate of $B \to D^*\ell\nu_\ell$

One of the favored methods to measure the magnitude of the CKM matrix element $|V_{cb}|$ is by extracting the partial decay rates of semileptonic $B \to D^*\ell\nu$ decays, where the D^* meson $(c\bar{d})$ contributes the necessary c quark for the $b \to c$ transition and ℓ is a light charged lepton e, μ with its associated neutrino ν_ℓ. The $B \to D^*\ell\nu$ decay is described by the tree-level Feynman diagram in Fig.2.9. In general the differential decay rates of semileptonic B decays originate from the electroweak effective Hamiltonian

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} \sum_{q=u,c} V_{qb}(\bar{q}\gamma_\mu P_L b)(\ell\gamma^\mu P_L \nu_\ell), \tag{2.13}$$

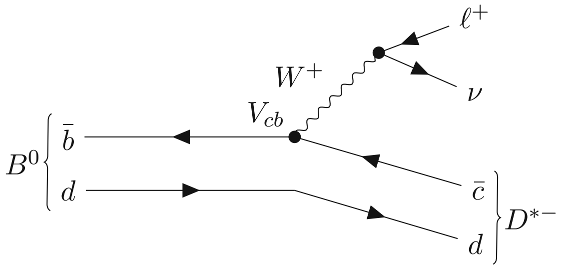


Figure 2.9: Tree-level Feynman diagram of the $B \to D^*\ell\nu$ decay process.

where γ^μ are the Dirac matrices, $P_L = (1-\gamma_5)/2$ and G_F is the Fermi constant. This results in the following dependence for differential decay rates with $b \to q$ $(q = c, u)$ transitions:

$$d\Gamma \propto G_F^2 |V_{qb}|^2 |L^\mu \langle X|\bar{q}\gamma_\mu P_L b|B\rangle|, \tag{2.14}$$

where L^μ is the leptonic current and X is the D^* meson in the case of $B \to D^*\ell\nu$ studies. For semileptonic B decays the QCD contributions to the decay amplitude are separated from the leptonic interactions and constrained in the term $\langle X|\bar{q}\gamma_\mu P_L b|B\rangle$. Any electromagnetic interactions between the leptons and quarks are taken into consideration by an electroweak correction factor η_{ew}. [22] For exclusive semileptonic decays the QCD contributions from the quark current is determined using form factor parameterizations.[19]

The kinematics and differential decay rate of the $B \to D^*\ell\nu$ decay can be fully characterized by four kinematic variables when following the formulation proposed by Ref.[23, 24]. These kinematic variables are the hadronic recoil w and three angular variables θ_ℓ, θ_v and χ. The hadronic recoil and the momentum transferred to the lepton system $q^2 = (P_\ell + P_\nu)^2$, where P_ℓ and P_ν refer to the four momenta of ℓ and ν, are related to each other and also kinematically restrict one another: [16, 19]

$$w = \frac{P_B P_{D^*}}{m_B m_{D^*}} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}}, \tag{2.15}$$

where P_B and P_{D^*} are the four momenta of the B and D^* mesons and m_B, m_{D^*} their respective masses of 5.279 GeV and 2.010 GeV in accordance with the current world average provided by the Particle Data Group (PDG). [10] Another property of w is its reduction to the Lortentz boost $\gamma_{D^*} = \frac{E_{D^*}}{m_{D^*}}$ in the B rest frame. Due to the relation between w and q^2 a minimum value of $q^2_{min} = 0$ results in a maximum value for w:

$$w_{max} = \frac{m_B^2 + m_{D^*}^2}{2 m_B m_{D^*}} \approx 1.504 \tag{2.16}$$

and consequently, zero recoil $w_{min} = 1$ corresponds to:

$$q^2_{max} = (m_B - m_{D^*})^2 \approx 10.69 \text{ GeV}. \tag{2.17}$$

The remaining three angular variables θ_ℓ, θ_v and χ, illustrated in Fig.2.10, correspond to:

- θ_ℓ: The angle between the lepton and the opposite direction of the B meson in the rest frame of the virtual W boson.

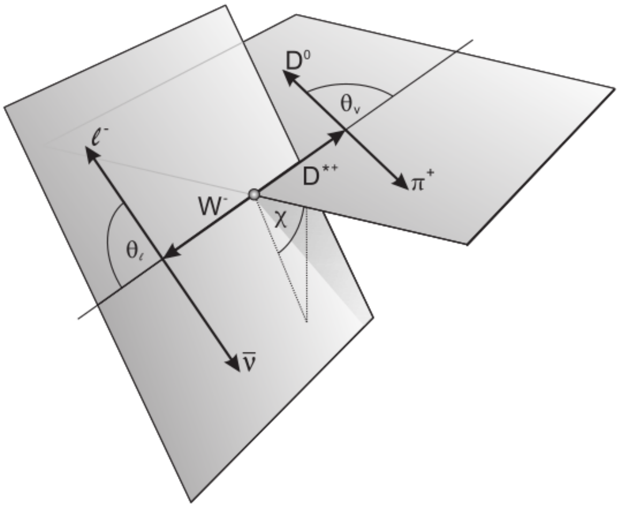


Figure 2.10: Schematic of the angular kinematic variables θ_ℓ, θ_v and χ. [19]

- θ_v: Angle between the opposite direction of the B meson in the D^* rest frame and the D^0.
- χ: Tilting angle between the D^* and W decay planes in the rest frame of the B meson.

For negligible lepton masses the differential decay rate of $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ reads as:

$$\begin{aligned}\frac{d^4\Gamma(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)}{dw\ dcos\theta_\ell\ dcos\theta_V\ d\chi} =& \frac{6m_Bm_{D*}^2}{8(4\pi)^4}\eta_{ew}^2G_F^2|V_{cb}|^2\sqrt{w^2-1}(1-2wr+r^2)\\ &[(1-cos\theta_\ell)^2sin^2\theta_vH_+^2(w)+(1+cos\theta_\ell)^2sin^2\theta_vH_-^2(w)]\\ &+4sin^2\theta_\ell cos^2\theta_vH_0^2(w)-2sin^2\theta_\ell sin^2\theta_v cos2\chi H_+(w)H_-(w)\\ &-4sin\theta_\ell(1-cos\theta_\ell)sin\theta_v cos\theta_v cos\chi H_+(w)H_0(w)\\ &+4sin\theta_\ell(1+cos\theta_\ell)sin\theta_v cos\theta_v cos\chi H_-(w)H_0(w)],\end{aligned} \tag{2.18}$$

where $r = m_{D^*}/m_B$ and H_+, H_- and H_0 are the three helicity amplitudes of the $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decay amplitude. Integrating over the two angular kinematic variables θ_v and χ yields:

$$\begin{aligned}\frac{d^2\Gamma(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)}{dw\ dcos\theta_\ell} =& \frac{6m_Bm_{D*}^2}{8(4\pi)^4}\eta_{ew}^2G_F^2|V_{cb}|^2\sqrt{w^2-1}(1-2wr+r^2)\\ &[(1-cos\theta_\ell)^2H_+^2(w)+(1+cos\theta_\ell)^2H_-^2(w)+\\ &+2(1-cos^2\theta_\ell)H_0^2(w)]\end{aligned} \tag{2.19}$$

In the heavy quark symmetry basis [25] the axial vector and vector currents of the hadronic contributions can be factorized using four independent form factors h_V and h_{A_i} $(i = 1, 2, 3)$, which are functions of w: [16, 19]

$$\frac{\langle D^*|\bar{c}\gamma^\mu b|B\rangle}{\sqrt{m_Bm_{D^*}}} = ih_V\epsilon^{\mu\nu\alpha\beta}, \tag{2.20}$$

$$\frac{\langle D^*|\bar{c}\gamma^\mu\gamma^5 b|B\rangle}{\sqrt{m_B m_{D^*}}} = h_{A_1}(w+1)\epsilon^{*\mu} - h_{A_2}(\epsilon^* \cdot v)v^\mu - h_{A_3}(\epsilon \cdot v)v'^\mu, \tag{2.21}$$

where the four velocities are $v = p_B/m_B$ and $v' = p_{D^*}/m_{D^*}$, ϵ^* is the polarization vector of the D^* and ϵ the Levi-Civita tensor. For negligible masses of the light leptons $\ell = e, \mu$ the hadronic matrix elements can be described by h_{A_1} and the two form factor ratios R_1 and R_2: [16, 19]

$$R_1(w) = \frac{h_V}{h_{A1}}, \quad R_2(w) = \frac{h_{A_3} + r h_{A_2}}{h_{A_1}}. \tag{2.22}$$

The form factors of the heavy quark symmetry basis and their ratios can be translated to the form factors $g, f, \mathcal{F}_1$ proposed by Boyd, Grinstein and Lebed (BGL)[26, 27]:

$$h_{A_1} = \frac{f}{m_B\sqrt{r}(w+1)} \tag{2.23}$$

$$h_V = g m_B \sqrt{r} \tag{2.24}$$

$$h_{A_1}(w - r - (w-1)R_2) = \frac{\mathcal{F}_1}{m_B^2\sqrt{r}(w+1)} \tag{2.25}$$

Following the approach of the BGL parameterization yields for the helicity amplitudes:

$$H_+ = f - m_B|\vec{p}_{D^*}|g, \tag{2.26}$$

$$H_- = f + m_B|\vec{p}_{D^*}|g, \tag{2.27}$$

$$H_0 = \frac{1}{m_{D^*}\sqrt{q^2}}\left[2m_B^2|\vec{p}_{D^*}|a_+ - \frac{1}{2}(q^2 - m_B^2 + m_{D^*}^2)f\right] \equiv \frac{\mathcal{F}_1}{\sqrt{q^2}}, \tag{2.28}$$

where $|\vec{p}_{D^*}|$ is the magnitude of the 3-momentum of the D^* in the B rest frame. In the BGL parameterization the form factors are expressed in terms of expansion coefficients a_n, b_n and c_n: [26, 27]

$$g(z) = \frac{1}{P_g(z)\phi_g(z)}\sum_{n=0}^{N} a_n z^n, \tag{2.29}$$

$$f(z) = \frac{1}{P_f(z)\phi_f(z)}\sum_{n=0}^{N} b_n z^n, \tag{2.30}$$

$$\mathcal{F}_1(z) = \frac{1}{P_{\mathcal{F}_1}(z)\phi_{\mathcal{F}_1}(z)}\sum_{n=0}^{N} c_n z^n, \tag{2.31}$$

where P are the Blaschke factors, ϕ are outer functions and z is a conformal variable defined as:

$$z \equiv \frac{\sqrt{w+1} - \sqrt{2a}}{\sqrt{w+1} + \sqrt{2a}}, \tag{2.32}$$

where it is chosen that $a = 0$ and therefore zero recoil corresponds to $z = 0$. The coefficients c_0 and b_0 are related by:

$$c_0 = \left(\frac{(m_B - m_{D^*}\phi_{\mathcal{F}_1}(0)}{\phi_f(0)}\right) b_0 \tag{2.33}$$

Another approach for the parameterization of the form factors is the CLN parameterization proposed by Caprini, Lellouch and Neubert. [28] The CLN parameterization expresses

h_{A_1} as an expansion of z with the slope parameter ρ^2 and the ratios $R_{1,2}$ in terms of the normalizations $R_{1,2}(1)$ and corrections up to the order of $(w-1)^2$:

$$h_{A_1}(w) = h_{A_1}(1)[1 - 8\rho^2 z + (53\rho^2 - 15)z^2 - (231\rho^2 - 91)z^3] \tag{2.34}$$

$$R_1(w) = R_1(1) - 0.12(w-1) + 0.05(w-1)^2 \tag{2.35}$$

$$R_2(w) = R_2(1) + 0.11(w-1) - 0.06(w-1)^2 \tag{2.36}$$

The parameters ρ^2 and $R_{1,2}(1)$ are free parameters, that have to be extracted from data. This makes the CLN parameterization model dependent, while the BGL parameterization is model independent.

CHAPTER 3

Belle II and SuperKEKB

3.1 SuperKEKB Accelerator

The data used for the $|V_{cb}|$ study of this thesis are obtained at the Belle II experiment, which measures particle momenta and energies from collision products of the particle collider SuperKEKB, located in Tsukuba, Japan, at the Japanese High Energy Accelerator Research Organisation (KEK). SuperKEKB is an asymmetric energy electron positron double ring collider with a circumference of ~ 3 km and the upgraded successor of the KEKB collider, which was in operation from 1991 to 2010 and produced collisions for the Belle experiment. The asymmetry of the beam energies provides a boost to the B mesons, which would otherwise be produced at rest. This boost makes it possible to separate the decay vertices of the B mesons from one another, providing an opportunity for CP violation measurements and probing of the CKM mechanism. One of the greatest achievements of Belle was to successfully provide experimental proof for the CKM mechanism of the quark mixing proposed by Kobayashi and Maskawa, which lead to them receiving the Nobel Prize in 2008. First test runs at SuperKEKB were commissioned in 2016 and the Belle II experiment started taking data in 2019. The accelerator complex consists of the following four main components, shown in Fig.3.1: [16, 29]

- LER: Low energy positron ring for e^+ with an energy of 4 GeV.
- HER: High energy electron ring for e^- with an energy of 7 GeV.
- Electron-Positron injector linac: An injector linear accelerator (linac), which accelerates the electron and positron beams to an energy of 7 GeV and 4 GeV, respectively and supplies them to the HER and LER.
- Positron damping ring: A ring placed at the end of the second sector of the linac to reduce the emittance of the positron beam.

In general the performance of particle colliders is characterized by their instantaneous luminosity $\mathcal{L}$ in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is related to the production rate of a certain type

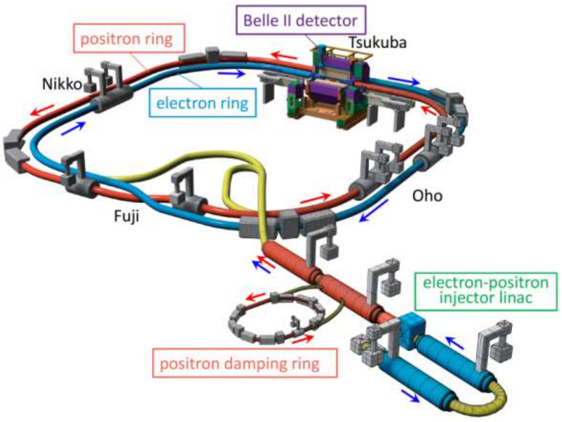


Figure 3.1: Illustration of the asymmetric energy SuperKEKB collider. [29]

of events per second dN/dt:

$$\mathcal{L} = \frac{dN}{dt}\frac{1}{\sigma}, \tag{3.1}$$

where σ is the cross section of the event of interest. The generic, instantaneous luminosity for the head-on collision of two beam bunches is determined by the collision frequency f and, assuming a Gaussian profile of the beams, the width of the horizontal and vertical beam sizes at the interaction point (IP) σ_x^* and σ_y^*: [10]

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x^*\sigma_y^*}, \tag{3.2}$$

where n_1 and n_2 are the particles contained in the colliding beam bunches and $\sigma_{x,y}^* = \sqrt{\epsilon_{x,y}\beta_{x,y}^*}$ is defined by the emittance ϵ and beta function β. By additionally introducing the beam current I, which is defined as the IP crossing rate of the beam, the luminosity of SuperKEKB can be written as: [29]

$$\mathcal{L} = \frac{\gamma_\pm}{2er_e}\left(1+\frac{\sigma_y^*}{\sigma_x^*}\right)\left(\frac{I_\pm\xi_{y\pm}}{\beta_y^*}\right)\left(\frac{R_\mathcal{L}}{R_{\xi_y}}\right), \tag{3.3}$$

where γ is the Lorentz factor, e the elementary particle charge and r_e the classical electron radius. Furthermore β_y^* is the vertical beta function at the IP, ξ_y is the vertical beam-beam parameter and $R_\mathcal{L}$ and R_{ξ_y} are reduction factors for the luminosity and beam-beam parameter, respectively. The expression of $\mathcal{L}$ in Eq.3.3 shows, that the luminosity is mainly determined by the beam current, the vertical beta function at the IP and vertical beam-beam parameter. [10, 16, 29, 30]

SuperKEKB was designed to achieve and instantaneous luminosity of $8 \times 10^{35}\text{cm}^{-2}\text{s}^{-1}$. For the purpose of comparison: the set goal is ~ 40 times the instantaneous luminosity achieved at KEKB, which peaked $\mathcal{L} = 2.11 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$. It is planned to achieve this increased luminosity goal by doubling the beam current and decreasing β_y^* by a factor of 20. While SuperKEKB didn't achieve its goal luminosity yet, it is the world record holder with $\mathcal{L} = 4.65 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$, at the time of writing.

In particle physics the size of the accumulated data samples is given in terms of the time integrated luminosity, $\mathcal{L}_{int} = \int_0^T \mathcal{L}dt$ in the units of barn, where $1\text{b} = 10^{-24}\text{cm}^2$. The number of events N of a given process is directly related to its probability of occurrence, the cross-section σ and the integrated luminosity: $N = \mathcal{L}_{int} \times \sigma$. Various estimated cross sections for the SuperKEKB setup are listed in Tab.3.1. The estimations were determined by Monte Carlo (MC) simulations.[16]

By colliding electrons and positrons with asymmetric energies of 7 GeV and 4 GeV, respectively SuperKEKB collides particles at $\Upsilon(4S)$ resonance with a center of mass energy of $\sqrt{s} = 10.58$ GeV. With a quark content of $b\bar{b}$ the $\Upsilon(4S)$ meson decays predominantly into charged or neutral B meson pairs with a branching fraction over 96%, which categorizes the SuperKEKB collider as a B factory. The production process of the $\Upsilon(4S)$ meson from e^-e^+ annihilation and its decay into B meson pairs is described by the Feynman diagram in Fig.3.2. The center of mass energy of 10.58 GeV is very close to the energy threshold for $B\bar{B}$ production, therefore the pair would be produced at rest, if the boost resulting from the asymmetric energy of the beams was missing. Additionally, SuperKEKB also collides electrons and positrons slightly below the $\Upsilon(4S)$ mass at $\sqrt{s} = 10.52$ GeV to produce samples exclusively containing $q\bar{q}$ events, which is referred to as off-resonance. The data taken at the $\Upsilon(4S)$ resonance is the so called on-resonance data. [16, 30]

Physics Process	**Cross Section [nb]**
$\Upsilon(4S)$	1.110 ± 0.008
$u\bar{u}(\gamma)$	1.61
$d\bar{d}(\gamma)$	0.40
$s\bar{s}(\gamma)$	0.38
$c\bar{c}(\gamma)$	1.30
$e^-e^+(\gamma)$	300 ± 3 (MC stat.)
$\gamma\gamma(\gamma)$	4.99 ± 0.05 (MC stat.)
$\mu^-\mu^+(\gamma)$	1.148
$\tau^-\tau^+(\gamma)$	0.919
$e^-e^+e^-e^+$	39.7 ± 0.1 (MC stat.)
$e^-e^+\mu^-\mu^+$	18.9 ± 0.1 (MC stat.)

Table 3.1: Cross-section estimations of events occurring in e^-e^+ collisions at $\sqrt{s} = 10.58$ GeV. [16]

Within its 19 years of operation the Belle experiment collected a data sample corresponding to 711 fb^{-1} at the $\Upsilon(4S)$ resonance and $\sim$ 1 ab^{-1}, when including the data with an energy offset. The initially set goal for the sample size of the Belle II experiment is equal to an integrated luminosity of 50 ab^{-1}. From the start of data taking in 2019 up until its long shutdown 1 (LS1), due to planned detector upgrades, in mid 2022 Belle II accumulated a sample size of 364 fb^{-1} at the $\Upsilon(4S)$ resonance and 43 fb^{-1} of off-resonance data. The development of the instantaneous and integrated luminosity at Belle II is summarized in Fig.3.3. [16, 30]

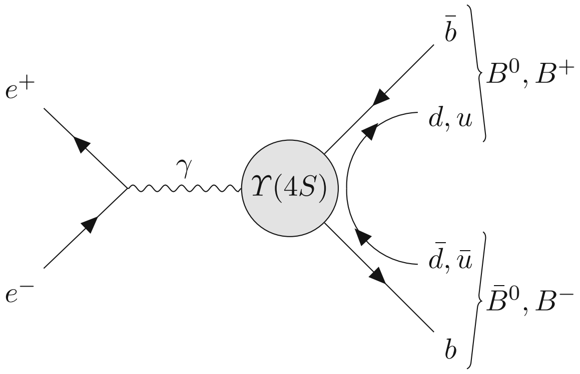


Figure 3.2: Feynman diagram of the $\Upsilon(4S)$ production through the annihilation of an e^-e^+ pair and its $\Upsilon(4S) \rightarrow B\bar{B}$ decay.

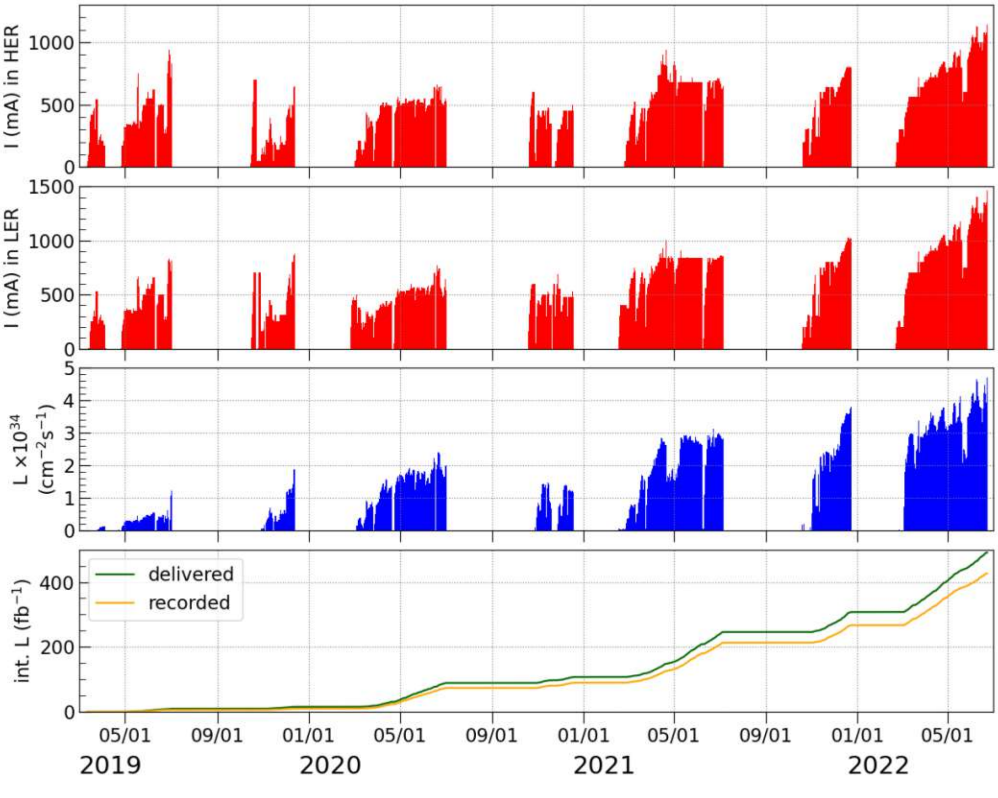


Figure 3.3: Summary of the beam currents of the HER, LER, the instantaneous luminosity and the integrated luminosity of SuperKEKB from its start of operation in 2019 up to the long shut down in 2022.

3.1.1 Beam Induced Backgrounds

As a consequence of higher beam currents and the lower vertical cross section at the IP the Belle II experiment will suffer from higher background rates than its predecessor. The expected five main background sources induced by the beam are Touschek scattering, beam-gas scattering, synchrotron radiation, radiative Bhabha processes and two photon processes.[16, 30]

Toushek Scattering

Toushek scattering is an intra-bunch scattering process, where coulomb scattering of particles inside the same bunch occurs. This leads to a deviation from the beam's nominal values for energies and momenta for scattered particles. As a consequence the scattered particles might hit vacuum pipe walls or magnet walls, which in turn creates showers, that might hit the detector if the scattering process happens near IP. The local loss rate from Toushek scattering is described by Bruck's formula.[31] Bruck's formula shows that the scattering probability is proportional to the bunch current density and second power of the beam current, as well as inversely proportional to the beam size and the third power of the beam energy. Due to the exhibited proportionality of the scattering rate, the effect will be larger for the LER beam, due to lower energies and higher beam currents. The backgrounds from Toushek scattering are expected to be 20-30 times higher than for KEKB.[16, 30]

Beam-Gas Scattering

Beam-gas scattering is a process that occurs, when beam particles collide with residual gas molecules left in the beam pipe. The scattering can occur either via Coulomb or Bremsstrahlung scattering. Coulomb scattering changes the direction of the particles, while Bremsstrahlung scattering changes their energy. The consequences are similar to Toushek scattering, i.e. due to the change of momentum the particles get lost in the beam pipe. If walls inside the detector get hit the created showers may be detected, creating background events. The probability of the Beam-gas scattering occurrence is proportional to the vacuum pressure and the beam currents. While the vacuum pressure is the same as for KEKB for the HER and LER rings it is increased near the IP. Additionally beam currents are increased by a factor of 2, which leads to an expected $\sim$ 100 times increase of this background component in comparison to its predecessor. However, the quality of the vacuum varies during a data taking run, i.e. a specific time period of operation.[16, 30]

Synchrotron Radiation

When relativistic, charged particles get accelerated by an magnetic field in a perpendicular direction, the particles emit electromagnetic radiation called synchrotron radiation. The emission rate of the photons is proportional to energy of the charged particle, therefore its main source is the HER beam. In addition to losing particles in the beam pipe the magnetic fields of the final focusing system near the IP can cause the emission of photons inside the detector, which leads to detected background showers from photons.[16, 30]

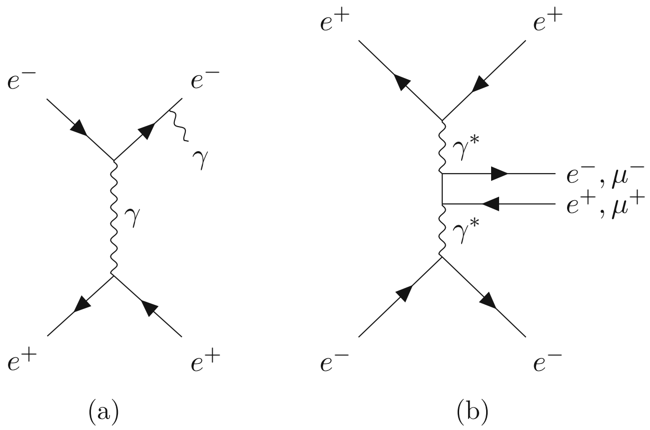


Figure 3.4: Radiative Bhabha scattering(left) and two photon process(right) Feynman diagrams.

Radiative Bhabha Process

Bhabha scattering is a $e^-e^+ \rightarrow e^-e^+$ scattering process, that can result in the production of a photon $e^-e^+ \rightarrow e^-e^+\gamma$, which is called radiative Bhabha scattering. The Feynman diagram of this process is shown in Fig.3.4a. Photons produced in this interaction travel in the direction of the beam axis and interact with the iron of the magnets. This interaction creates large amounts of neutrons due to the photo-nuclear resonance mechanism, which is the main background contribution for the K_L and muon detector placed in the return yoke of the Belle II detector. The neutron production rate from radiative Bhabha events is proportional to the detectors luminosity, that is designed to be ~ 40 times higher than its predecessors. In addition to the produced neutron from the photon emission, the trajectory of the electrons and positrons gets altered from their nominal value resulting in hits to the walls that create showers in the detector region or curl in the inner detector due to the magnetic field, leading to multiple false hits.[16, 30]

Two Photon Process

Electron and positron collisions have a chance to lead to the process $e^+e^- \rightarrow e^+e^-\gamma^*\gamma^* \rightarrow e^+e^-e^+e^-$ and $e^+e^- \rightarrow e^+e^-\gamma^*\gamma^* \rightarrow e^+e^-\mu^+\mu^-$, where * denotes that the photons are virtual particles. This interaction proceeds via the process illustrated in the Feynman diagram in Fig.3.4b. The produced electrons and positrons have a very low momentum and therefore they curl in the inner detectors and produce background hits.[16, 30]

3.2 Belle II Detector

The measurement and recording of the momenta and energies of final state particles produced from the collision at SuperKEKB and their identification is done by the Belle II detector. Heavy particles, e.g. B and D, decay too quickly to reach the detector, however their long lived daughters can be detected. The following charged and neutral particles can be detected: $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, $p^{\pm}$, γ, K_L^0, n and d.[16, 30]

The Belle II detector started data taking in 2019 and is an upgraded version of the hermetic, magnetic spectrometer Belle, which was in operation from 1999 to 2010. A key characteristic of the Belle II detector is its large solid angle coverage of $\theta \in [17°, 150°]$ due to the hermetic structure of the detector This results in the ability to reconstruct events with missing energy. Other features include its high electron and muon identification efficiency of 86% and 88%, respectively, as well as a high photon detection efficiency. [32, 33]

The coordinates of particle trajectories and various other spatial processes are provided either in Cartesian or polar coordinates. The center of origin of both coordinate systems is defined by the interaction point (IP) of the detector, where the electron and positron beams collide. For the Cartesian coordinates the z-axis is equivalent to the electron beam axis, the y axis points towards the top and the x axis points outwards. The polar coordinates are defined by the distance r from the IP, the polar angle θ and the azimuth angle ϕ in the x-y plane.[16, 30]

The spectrometer consists of a superconducting solenoid, that provides a magnetic field of 1.5 T, an iron return yoke for the magnetic flux and five sub-detector systems, as illustrated in the three dimensional model in Fig.3.5. These sub-detectors are from inner to outer layer:[16, 30]

- Vertex Detector (VXD) that consists of a Silicon Vertex Detector (SVD) and a silicon Pixel Detector (PXD).
- Central drift chamber (CDC)
- Particle Identification (PID) system consisting of a time of propagation counters (TOP) and aerogel ring imaging detectors (ARICH)
- Electromagnetic calorimeter (ECL)
- K_L^0 and muon detector (KLM)

3.2.1 Vertex Detector (VXD)

The Vertex Detector (VXD) is the inner most sub-detector system, situated near the IP around the Be beam pipe with a radius of 10 mm. The VXD measures the positions of the tracks of charged particles near the IP. These track positions are called hits and are used to reconstruct decay vertices.

The measurement of hits is based on the helix movement of charged particle trajectories caused by the Lorentz force of the magnetic field inside the detector. From the knowledge of the magnetic flux density and the curvature of the tracks one can then determine the momentum of the particles. By extrapolating the tracks from the measured hits to the IP one can reconstruct the position of the B meson decay vertices. The reconstruction of decay vertex positions is a mandatory requirement for CP violation measurements. Additionally the precision of the VXD allows for the measurement of secondary decay vertices from short lived B meson daughters, e.g. D mesons.[16, 30]

The VXD is also the only detector that can measure the tracks of very low momentum particles, which already curl inside the detector layers close to the IP and can't or barely reach the outer layers due to the Lorentz force. An example of such particles would be slow charged pions with momenta below 300 MeV. Finally, the VXD also measures the energy loss dE/dx for particle identification purposes. For information on the topic of PID see

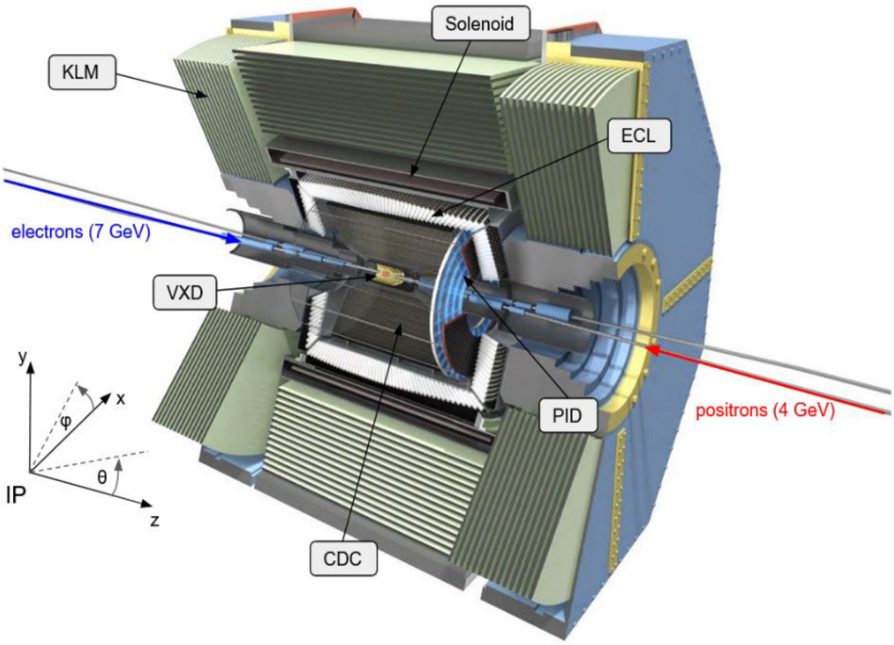


Figure 3.5: A three dimensional model of the cross-section of the Belle II detector. The coordinate systems used for spatial quantities and processes such as momenta and positions is illustrated in the lower left corner. Adapted from [34].

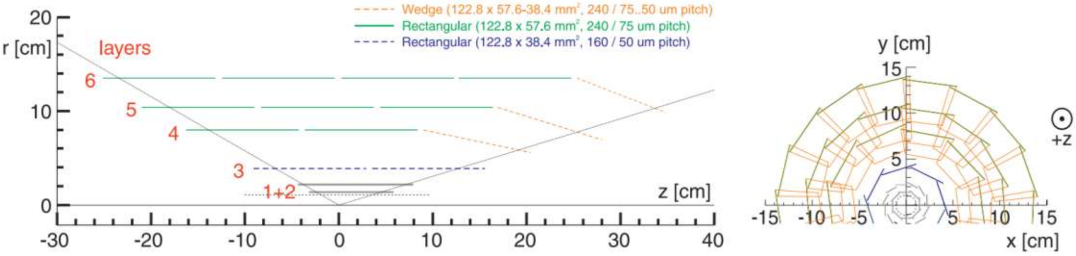


Figure 3.6: A schematic cross-section showing the shape and position of the first two PXD layers in layer one and two, followed by the SVD situated in the layers three to six from a side and front-view in the left and right figure, respectively. [16]

Sec3.3.[16, 30]

The VXD consists of six layers and two sub-detectors, namely the silicon Pixel Detector (PXD) and the Silicon Vertex Detector (SVD). Its layout is shown in the schematic in Fig.3.6. The asymmetric structure of the VXD is a consequence of the Lorentz boost from the asymmetric beam energies. The PXD consists of pixelated DEPFET sensors and occupies the first two layers at the radii of 14 mm and 20 mm. These layers are followed by the SVD, which contains four layers of double-sided silicon-strip sensors at the following radii: 39 mm, 80 mm, 104 mm, and 135 mm. Finally, the VXD covers a polar angle within the range of $\theta \in [17°, 150°]$. The hit resolution of the VXD system is approximately 20 µm.[16, 30]

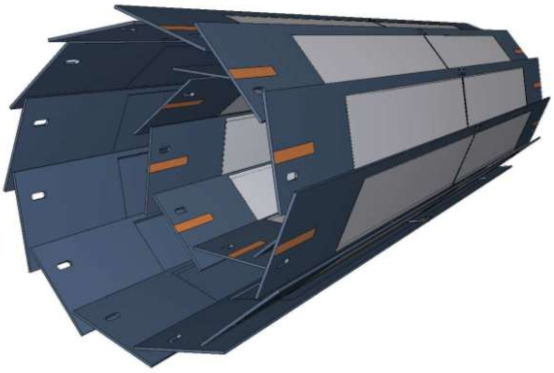


Figure 3.7: A three dimensional render of a side-view of the two PXD layers, showing its wind-mill like layout.[30]

Pixel Detector (PXD)

To improve the vertex reconstruction capabilities in comparison to its predecessor, it is necessary to situate detectors closer to the IP. The PXD consists of pixels based on silicon Depleted P-channel Field Effect Transistor (DEPFET) sensors. However, the close proximity to the IP comes at the cost of increased amounts of backgrounds. Among other things the increase occurs, due to the narrower beam pipe radius at the IP. This increase in backgrounds, especially at close proximity to the IP renders strip detectors useless at close range due to the large occupancy from background events. The occupancy is the fraction of occupied channels, i.e. strips or pixels, compared to unoccupied channels. The SVD layers consist of overall ~ 224 thousand strips, while the PXD contains ~ 8 million pixels, reaching a maximum occupancy of $\sim 3\%$.[16, 30]

The PXD consists of two layers, at the distance of 14 mm and 20 mm from the IP. Each layer consists of ladders, i.e. a sensor with a width of 15 mm and a length of 90 mm or 123 mm for the first and second layer, respectively. These ladders are orientated in a wind-mill like shape as illustrated in the three dimensional render of a side-view in Fig.3.7. The first layer consists of 8 ladders and the second layer currently consists of 2 ladders with a planned goal of 12 ladders. Unfortunately the integrated read out time of 19.2 μs of the PXD is slower than other sub-detectors making is unusable for trigger purposes. General information about Belle II's trigger system can be found in Sec.3.4.[16, 30]

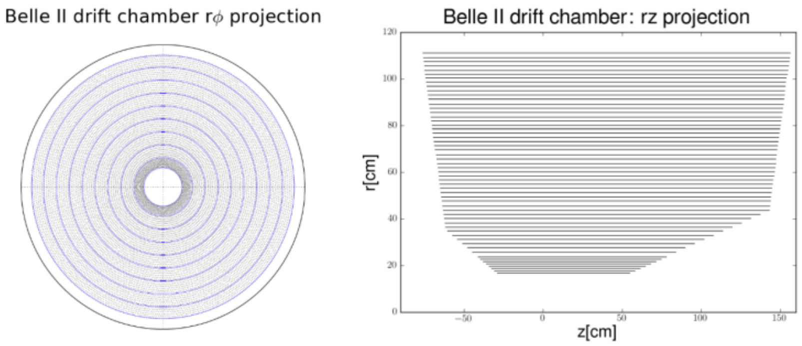
Silicon Vertex Detector (SVD)

Suffering from less backgrounds and radiation than the PXD, the structure of the sensors can be changed from pixels for strips to reduce the number of channels to read out. The SVD consists of four layers of sensors, which are labeled by the numbers 3, 4, 5 and 6, with increasing radius. The sensors of these layers consist of double-sided silicon strip detectors (DSSDs) based on six inch wafers with the number of ladders and sensors per ladder varying depending on the layer. These layers continue the windmill pattern of the PXD. A summary of the SVD's layout in numbers of ladders and sensors per layer is given in Tab.3.2.[16, 30]

The SVD is required to cover the full angular acceptance of Belle II, i.e. $\theta \in [17°, 150°]$. It serves as the detector for track reconstruction of low momentum tracks down to a

Layer	Radius [mm]	Ladders	Sensors per Ladder	Sensors
3	39	8	2	16
4	80	10	3	30
5	104	14	4	53
6	135	17	5	85

Table 3.2: Summary of the sensors and ladder layout of the SVD.[30]


Belle II drift chamber rϕ projection

Belle II drift chamber: rz projection

Figure 3.8: Schematic cross-sections of the CDC's layout from a front and side-view in the left and right figure, respectively.[35]

momentum of a few tens of MeV, as well as for the reconstruction of tracks from K_s decays outside of the PXD. The last point is required to be able to reconstruct decays like $B \to K_s K_s K_s$, where only the charged pion daughters of the K_s leave tracks. Lastly, the SVD also provides information for the particle identification by measuring the ionization energy loss dE/dx.[16, 30]

3.2.2 Central Drift Chamber (CDC)

The Central Drift Chamber (CDC) is the sub-detector following the SVD and is the main tracking detector of Belle II. It precisely determines the trajectories and consequently the momenta of charged particles. Additionally the CDC serves the purpose of obtaining information regarding the particle's identity by measuring the ionization energy loss dE/dx of the particles within its gas volume. Finally, the CDC is part of the trigger system, i.e. a system that determines whether the measured data will be recorded or not, and provides trigger signals for charged particles.[16, 30]

The CDC is a cylindrical chamber filled with a 50-50 gas mixture of He and C_2H_6. It occupies the detector region from $r = 160$ mm to $r = 1130$ mm and covers a polar angle region of $\theta \in [17°, 150°]$, equal to the whole angular acceptance region of Belle II. The CDC is organised in 9 super layers which consist of overall 59 layers with 160-384 wires per layer amounting to overall 14336 wires. A schematic of the CDC's layout is provided in Fig.3.8.[16, 30]

The working principle of the CDC is based on ionization, i.e. charged particle passing through the gas create electron-ion pairs due to electromagnetic interactions with the gas molecules. As the charged particles travel through the chamber they leave a track of ionized gas molecules, the resulting electrons of the ionization process then get accelerated by a

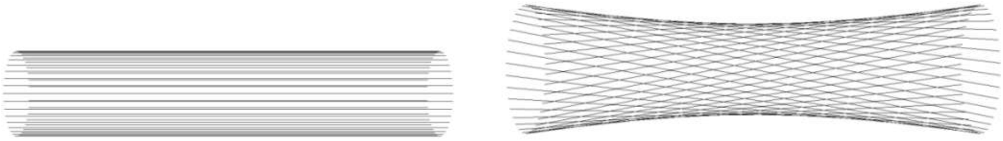


Figure 3.9: Axial and exaggerated stereo wire layers of the CDC on the left and right, respectively.[36]

homogeneous electrical field in the direction of so-called sense wires, ionizing further gas molecules and creating an avalanche. The electrical field is created by surrounding each sense wire with 8 field wires and applying a high voltage between the field wires and the sense wire. Due to the electrical field being homogeneous, the electrons drift to the sense wires with an approximately constant drift velocity. The constant drift velocity in turn enables the measurement of the drift time from the ionization of the first molecule to the arrival of the avalanche.[16, 30]

To reconstruct a three dimensional trajectory the layers within the super layers alternate between layers containing axial wires and stereo wires, both illustrated in Fig.3.9. The wires of the axial layers are parallel to the beam line providing the $r - \phi$ coordinates, while the stereo wires are slightly tilted resulting in a shift dependent on the polar angle θ, making it possible to estimate the z coordinate.[16, 30]

The overall strength of a signal can be summed up from all layers and used to determine the lost energy of the particle by estimating the ionization energy loss dE/dx for particle identification. The CDC is able to achieve an approximate hit precision of 100 µm and 2 mm for the $r - \phi$ plane and z coordinate, respectively.[16, 30]

3.2.3 Particle Identification System

The particle identification (PID) systems main purpose is to identify the charged final state hadrons of a decay with a momentum up to 4 GeV by providing an estimate of their rest mass. These hadrons include pions, kaons and protons with a rest mass of approximately 140 MeV, 494 MeV and 938 MeV, respectively.[16, 30]

Belle II's PID system consists of two independent cherenkov detectors, that cover different polar angle regions. The barrel region of the CDC, i.e. the polar angle region of $\theta \in [32°, 130°]$, is surrounded by 16 time of propagation (TOP) counter cells, whereas the forward end cap region of $\theta \in [17°, 35°]$ is covered by an aerogel ring-imaging cherenkov (ARICH) detector.[16, 30]

The momentum of charged particles is already known from the measurements of the CDC. Therefore to estimate a charged hadron's rest mass one needs to measure the velocity of the particles. This can be done by making use of cherenkov radiation. When a charged particle travels through a dielectric medium with a velocity greater than the propagation speed of light inside the medium photons are emitted at an angle determined by the particle's velocity, the speed of light and the refractive index of the medium. The relation between the emission angle θ_c of the photons, the particle's velocity v, the refractive n and the speed of light c is as follows: [3]

$$cos\theta_c = \frac{c'}{v} = \frac{1}{n\beta}. \tag{3.4}$$

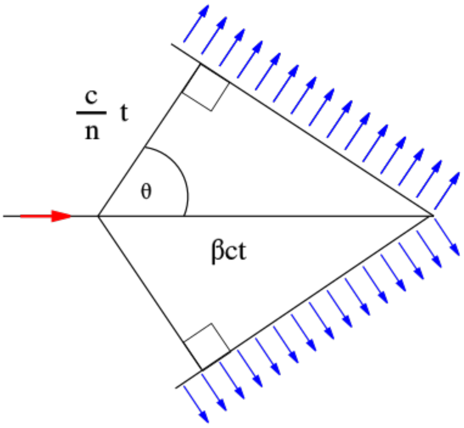


Figure 3.10: The characteristic geometry of cherenkov radiation emission. The direction of the charged particle is indicated by the red arrow.[37]

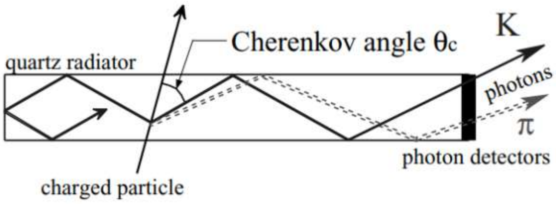


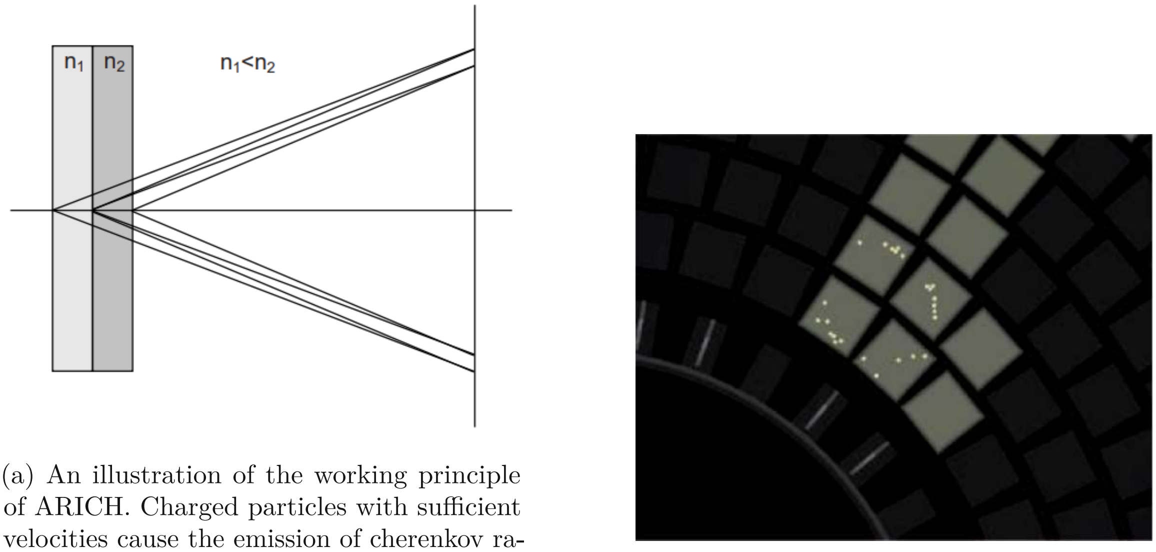
Figure 3.11: Side-view of TOP quartz bars illustrating their working principle, which shows the link between the emission angle of the cherenkov photons and a charged particle's velocity.[30]

The properties of the cherenkov radiation lead to its distinct emission geometry in Fig.3.10, thus the measurement of the emission angle allows to estimate a particles velocity and determine its mass.

TOP

The TOP counter surrounds the CDC barrel with 16 identical quartz modules, covering the polar angle region of $\theta \in [32°, 130°]$. Its working principle is based on the total photon reflection within the bar itself, a schematic side view of a quartz module is provided in Fig.3.11.[16, 30]

The time of propagation of the photons is based on the angle of the cherenkov photon emission, which in turn is proportional to the velocity and therefore mass of the charged particle. While the forward going photons are directly detected by front-end photon detectors, the backwards going photons first get reflected by a mirror at the back-end of the bar. By measuring the propagation time one can determine the angle of the emitted photons. From the angle and known refractive index the particle's velocity and thus its mass can be determined.[16, 30]



(a) An illustration of the working principle of ARICH. Charged particles with sufficient velocities cause the emission of cherenkov radiation from two aerogel layers, which overlap and form a cone on the photon detector.

(b) Photon ring originating from a cosmic muon passing through ARICH.

Figure 3.12: The layout of ARICH is depicted in Fig.a, while Fig.b shows the detected cone from a cosmic muon.[16]

ARICH

ARICH is an aerogel ring imaging cherenkov counter that covers the forward end-cap of the Belle II detector within the polar angle region of $\theta \in [17°, 35°]$. While the main purpose of ARICH is to provide information that separates charged kaons from pions within a momentum range of 0.5 GeV to 4 GeV it additionally allows for muon and pion separation in a momentum range up to 1 GeV. The ARICH consists of three components: two stacked 2 cm aerogel layers, a 20 cm expansion volume and a photon detector consisting of 7 rings, containing overall 420 Hybrid Avalanche Photon diode sensors. A schematic of the detector's layout is provided in Fig.3.12a. The two aerogel layers have a slightly different refractive index, namely $n_1 = 1.046$ and $n_2 = 1.056$ for the first and second layer, respectively.[16, 30]

As a particle with sufficient velocity traverses through the aerogel layers, photons will be emitted due to cherenkov radiation. Because the aerogel layers have different refractive indices the angle of the photon emissions from the first and second aerogel layer deviate slightly. This causes them to overlap when reaching the photon detector, effectively doubling the photon yield and resolution. The radiated photons form a cone and produce a ring on the photon detector rings, similar to the example of a detected cone originating from a cosmic muon in Fig.3.12b. The radius of the measured cone is proportional to the cherenkov angle of the radiated photons and therefore the mass of the traversing charged particle.[16, 30]

3.2.4 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECL) consists of three detectors, one sub-detector covering the barrel region and one for each end cap and follows up the PID system. Overall the calorimeter consists of 8736 scintillator crystals, i.e. 6624 CsI(Tl) crystals in the barrel region and 2112 CsI crystals in the end cap regions. The detector system covers a polar

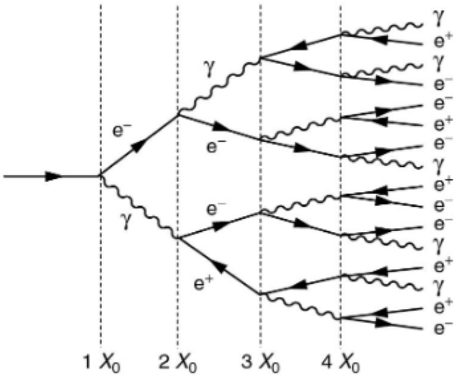


Figure 3.13: Development of electromagnetic showers in the ECL.[3]

angle region of $\theta \in [12°, 155°]$, with the exception of two $\sim 1°$ gaps between barrel and end cap sections. To detect the photons emitted by the scintillators crystals a photodiode is glued to the back of each crystal. [16, 30]

The main purpose of the ECL is the detection of photons in an energy range of 20 MeV to 4 GeV with a high efficiency by determining their energy and angular coordinates to be able to identify neutral particles. Approximately one third of the B mesons decay into neutral hadrons, e.g. π^0 mesons, which further decay to photons, e.g. $\pi^0 \to \gamma\gamma$. Secondly, the ECL contributes to particle identification by allowing to separate charged pions from electrons, because the latter deposit their whole energy in the calorimeter and get stopped, while the pions pass through, due to their mass difference of ~ 139 MeV. Finally, the ECL provides trigger signals and helps to identify K_L^0 mesons by combining its track information with the measurements from the surrounding K_L^0 and muon detector.[16, 30]

As electrons or photons interact with a medium they will produce e^-e^+ pairs that will in turn trigger an avalanche of photons, electrons and positrons due to Bremsstrahlung electron positron pair production. This will result in a shower, as illustrated in Fig.3.13, that develops until the energy of the particles falls below a certain threshold. Charged hadrons on the other hand lose their energy continuously due to ionization processes, thus not creating showers, but leaving measurable hits in the ECL. Typically, the performance of an ECL is provided by stating the relative energy resolution compared to the energy of the photons σ_E/E. In case of Belle II the simulations show an relative energy resolution of 7.7% and 2.3% for photons with an energy of 0.1 GeV and 1 GeV, respectively.[3, 16, 30, 38]

By combining the locations of the showers from energy deposits in the ECL with the tracks from previous detector layers, one can deduce if a given particle is neutral or charged. Charged particles leave tracks in the CDC and VXD, while neutral particles don't. The spatial resolution of the ECL is also dependent on the energy of the photons: $\sigma_{pos} = 0.5cm/\sqrt{E}$, where E is given in GeV.[16, 30, 38]

3.2.5 K_L^0 and Muon Detector

The K_L^0 and muon detector (KLM) serves not only the purpose of detecting K_L^0 mesons and muons with a momentum of at least 0.6 GeV, but also provides information to differentiate between them. The KLM is the only detector system situated outside of Belle II's superconducting solenoid and consists of alternating layers of 4.7 cm thick iron plates, additionally serving as return yokes for the magnetic flux, and active detector elements. Both the barrel and end cap regions are covered by the KLM, resulting in a polar angle coverage of $\theta \in [20°, 155°]$. While the barrel region consists of 15 detector elements and 14 iron plates, the end caps consist of 14 detector and 14 iron layers.[16, 30]

The neutral K_L^0 mesons leave hadronization showers due to interaction with the iron plates and can be separated into two types of candidates, depending on their polar angle with respect to the IP: K_L^0 that only leave hits in the KLM detector and candidates, which leave a hit in both the KLM and the ECL. Muons on the other hand don't create hadronic showers, but only leave hits inside the KLM. Additionally, K_L^0 and muons can be further separated by extrapolating the tracks from the KLM detector to the previous CDC and VXD layers, since only charged particles will leave hits in the first two sub detectors.[16, 30]

3.3 Particle Identification

Effective particle identification (PID) capabilities for leptons, as well as charged and neutral hadrons, are crucial for the Belle II experiment. The PID is a fundamental tool for background reduction and the identification of long lived final state particles.

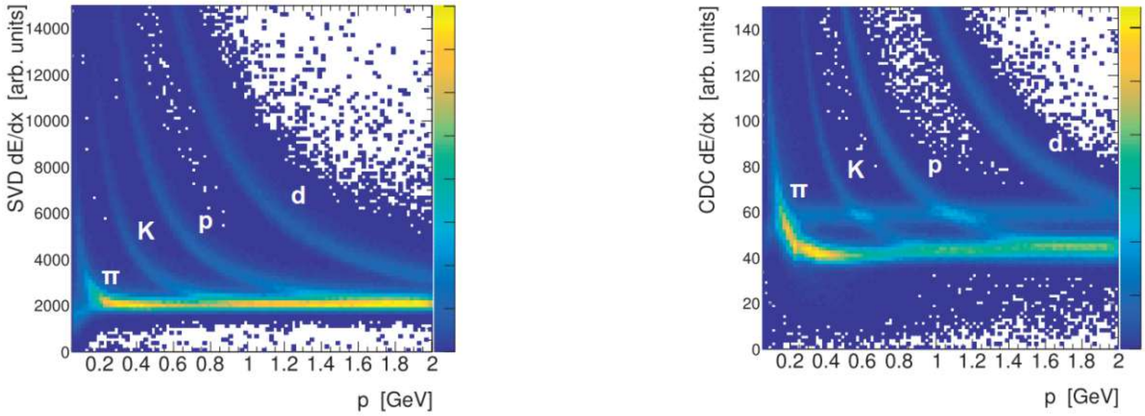
Belle II combines the complementary information obtained from each of the five sub-detector system to properly identify particles. For the separation between charged hadrons the SVD and CDC provide the main information via dE/dx ionization loss measurements, based on the Bethe-Bloch formula: [35]

$$\frac{dE}{dx} = \frac{4\pi N e^4}{m_e c^2 \beta^2} z^2 \left(\ln \frac{2 m_e c^2 \beta^2 \gamma^2}{I} - \beta^2 + corr \right), \tag{3.5}$$

where β is the velocity of the traversing particle and z its charge, γ is the Lorentz factor, m_e is the mass and e the charge of an electron, respectively. The material properties contributing to the Bethe-Bloch formula are its electron density N and the mean excitation energy I. In Fig.3.14a and Fig.3.14b the separation power for charged hadrons is shown for the SVD and CDC, respectively. The energy loss measurement is expected to be only dependent on a particle's momentum and provide better separation power below 1 GeV. Information obtained by the SVD and CDC can then be combined with the measurements from TOP and ARICH to further discriminate between charged kaons and pions as well as pions and muons, in the case of ARICH.[16, 30]

Leptons on the other hand are mainly identified by measurements from the ECL in case of electrons or the KLM for muons. The electrons have the distinct property that they deposit their entire energy in the ECL due to the mass dependency of the Bremsstrahlung emission. Combining the measured energy deposit E in the ECL with the track momentum p measurements from the CDC and SVD leads to a distinction between leptons and other charged particles using its distinct E/p distribution illustrated in Fig.3.15. While the ratio E/p provides good separation power for the momentum range of 1 to 3 GeV, this is not the case for lower momentum particles. The E/p ratio loses its distinct shape in the lower momentum range and in addition electrons might not even reach the ECL with low momenta making it especially difficult to separate electrons from pions.[16, 30]

Muon identification mainly makes use of the muons penetration power. The main idea is to extrapolate the tracks from the KLM measurements to the hits recorded in the CDC and VXD with the muon hypothesis to extract the muon likelihood.



(a) Ionization energy loss curves measured by the SVD.

(b) Ionization energy loss curves measured by the CDC.

Figure 3.14: The characteristic ionization energy loss dE/dx curves of charged hadrons as a function of the magnitude of the momentum. Adapted from [16]

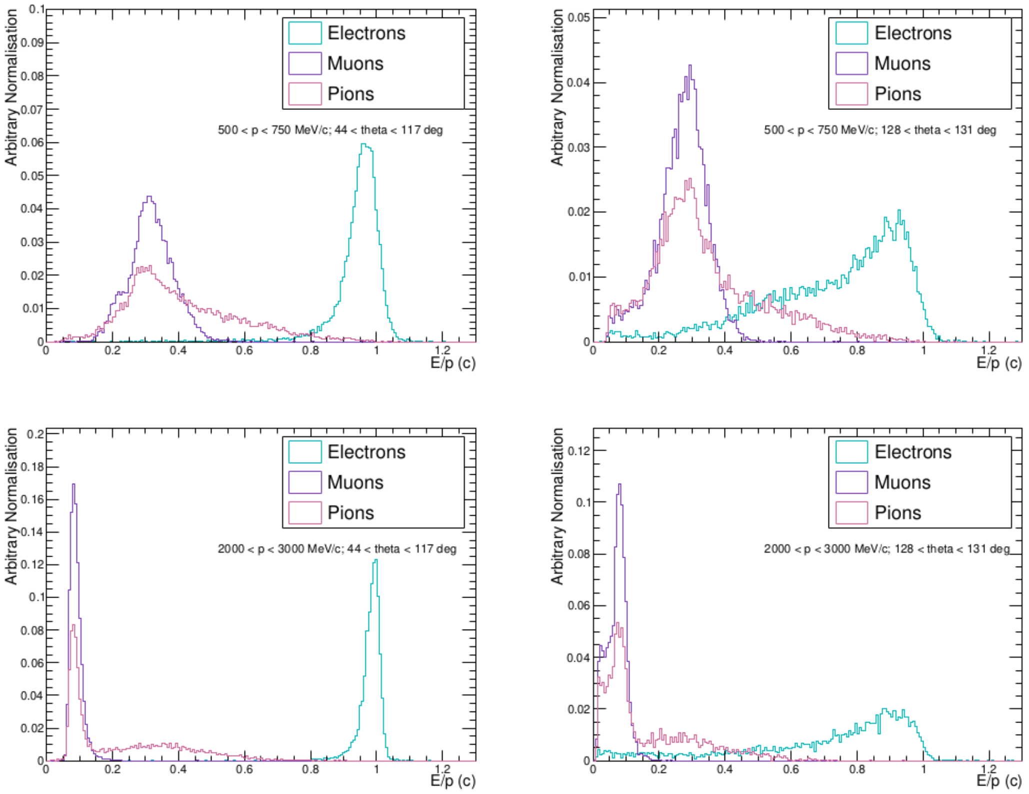


Figure 3.15: Shapes of the E/p distribution of electrons, muons and pions for various polar angle and momentum regions. This illustrates the loss of the distinct shape in lower momentum regions. The figures on the left side show the distributions for the polar angle range of $\theta \in [44°, 117°]$ and the momentum ranges of $500 < p < 750$ MeV and $2 < p < 3$ GeV on the top and bottom, respectively. The angular range of $\theta \in [128°, 131°]$ is depicted in a similar fashion on the right side.[16]

Photon and therefore π^0 identification relies on the shape of ECL clusters and the lack of associated tracks. While the shape of incident photon clusters is symmetrical, the main background showers caused by neutral and charged hadron showers are asymmetric and result in several clusters. Neutral pions decay via $\pi^0 \to \gamma\gamma$, which leads to specific ECL cluster shapes, depending on the energy of the decaying pion: below 1 GeV the the photons typically create two separate clusters, within [1, 2.5] GeV the clusters overlap but can still be separated, above 2.5 GeV the photon showers can not be separated anymore.[16, 30]

The K_L^0 meson identification relies mainly on information obtained by the KLM and ECL as they can leave hadronic clusters in the ECL and KLM, depending on their angle. In the KLM the K_L^0 are mainly identified by the number of penetrated layers, the distance between clusters and tracks, since neutral particles don't have any associated tracks from the inner detectors and the timing of the clusters. Information obtained from the ECL is complementary to the one obtained from the KLM. These include among others, the total energy deposited in the ECL, since K_L^0 only deposit very low amounts of energy below 50 MeV in the ECL.[16, 30]

The information obtained from the detector subsystems is then combined to obtain likelihoods for each particle hypothesis. In this way a likelihood for a given particle type can be obtained by comparing the PID likelihoods for all final state particles and construct the PID as follows:

$$\mathcal{L}(\alpha : \beta) = \frac{1}{1 + e^{\ln \mathcal{L}_\alpha - \ln \mathcal{L}_\beta}} = \frac{\prod_{\mathrm{det}} \mathcal{L}_\alpha}{\prod_{\mathrm{det}} \mathcal{L}_\alpha + \prod_{\mathrm{det}} \mathcal{L}_\beta}, \tag{3.6}$$

where α and β are the two compared particle types and the product is over the detector information of active detectors taken into account. This way the likelihoods of two or more particle types can be compared. Alternatively, one can also use particle identification likelihoods based on boosted decision trees. The concept of boosted decision trees is discussed in Sec.4.7.5. [16, 30]

For the slow pion calibration study featured in Ch.5, the binary kaon identification variable was used without of SVD information. This decision was based on studies that, at the time of writing, indicated a superior separation power between kaons and pions. The binary kaon identification likelihood ratio $\mathcal{L}_{K/\pi}$ is defined as follows:

$$\mathcal{L}_{K/\pi} = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi}, \tag{3.7}$$

where $\mathcal{L}_K$ and $\mathcal{L}_\pi$ are the likelihoods for a given particle to be a kaon or pion, respectively. The Belle II PID system exhibits an kaon identification efficiency of approximately 90% and pion fake rate of 5% for particles with a momentum magnitude of 2 GeV. A fake rate refers to the probability of a non kaon being falsely identified as a kaon, while the efficiency states the probability of a kaon being correctly identified.[16, 30]

The determination of the matrix element magnitude $|V_{cb}|$ through $B \to D^* \ell \nu_l$ (ℓ =e,µ) made use of the electron and muon identification systems. For the muon channel of $B \to D^* \ell \nu_l$ the global muon identification likelihood, i.e. the likelihood for all particle types was taken into consideration, without SVD information was used. With respect to muon identification, studies indicate an estimated efficiency of 92%-98% with a pion fake rate of 2%-6% for the momentum region above 1 GeV. The electron channel utilized an electron identification likelihood based on the multivariate classifier of boosted decision trees. The PID system of the Belle II detector shows a similar performance for electron

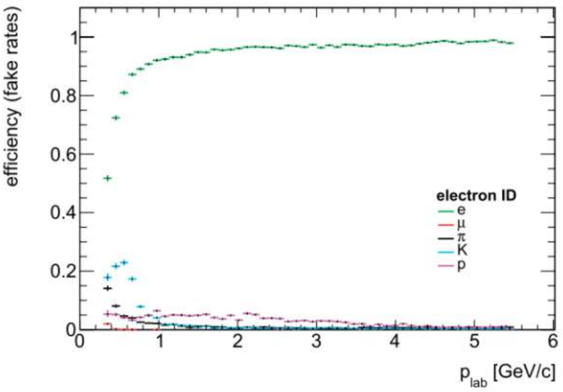


Figure 3.16: The electron efficiency and fake rate performance as a function of the momentum. A similar efficiency to the muon identification of approximately 92%-98% can be seen for the momentum region above 1 GeV and a pion fake rate below 5%. The lower efficiencies at low momenta can be explained, due to the electron lacking the momentum to reach the ECL.[16]

identification, which is illustrated in Fig.3.16. This choice of lepton identification variables follows recommendations of Belle II's performance group.[16, 30, 38]

3.4 Trigger System

Due to the copious amounts of collision data produced at the Belle II experiment, it is impossible to keep and record all data. The trigger system fulfills this task and decides whether data is recorded or discarded. As already mentioned in Sec.3.1.1, due to the design of SuperKEKB one has to take a variety of beam induced background events into account that are increased because of their proportionality to luminosity and beam related properties. Thus the number of background events is increased and the efficiency requirements of the trigger system are raised. While Belle II's physics program covers more than just B and D flavor physics, e.g. τ physics or dark sector searches, these flavor physics processes are luckily easily identifiable due to the rather high track multiplicity of at least three tracks in the CDC and the energy deposits in the ECL.[16, 30]

The trigger system consists of two sub-systems: first a hardware based trigger referred to as L1 containing four sub-trigger systems and a global decision logic (GDL). Sub-trigger systems collect the information obtained by detectors that are able to provide measurements fast enough to fulfill the trigger requirements of a latency of 5 μs and a maximum output rate of 30 kHz and send the collected information to the GDL that makes a decision based on the provided information.

The detectors included for the L1 trigger are: CDC, ECL, TOP, ARICH and the KLM. A schematic summary of the sub-trigger systems is provided in Fig.3.17. The information obtained by the sub-triggers is complementary. The track reconstruction from charged particles by the CDC characterizes properties of a particle that include its charge, opening angle and momentum as well as providing information about the charged track multiplicity of the event. The ECL sub-trigger consists of two triggers: a cluster trigger providing information about cluster shapes and their timing and an energy trigger based on the energy

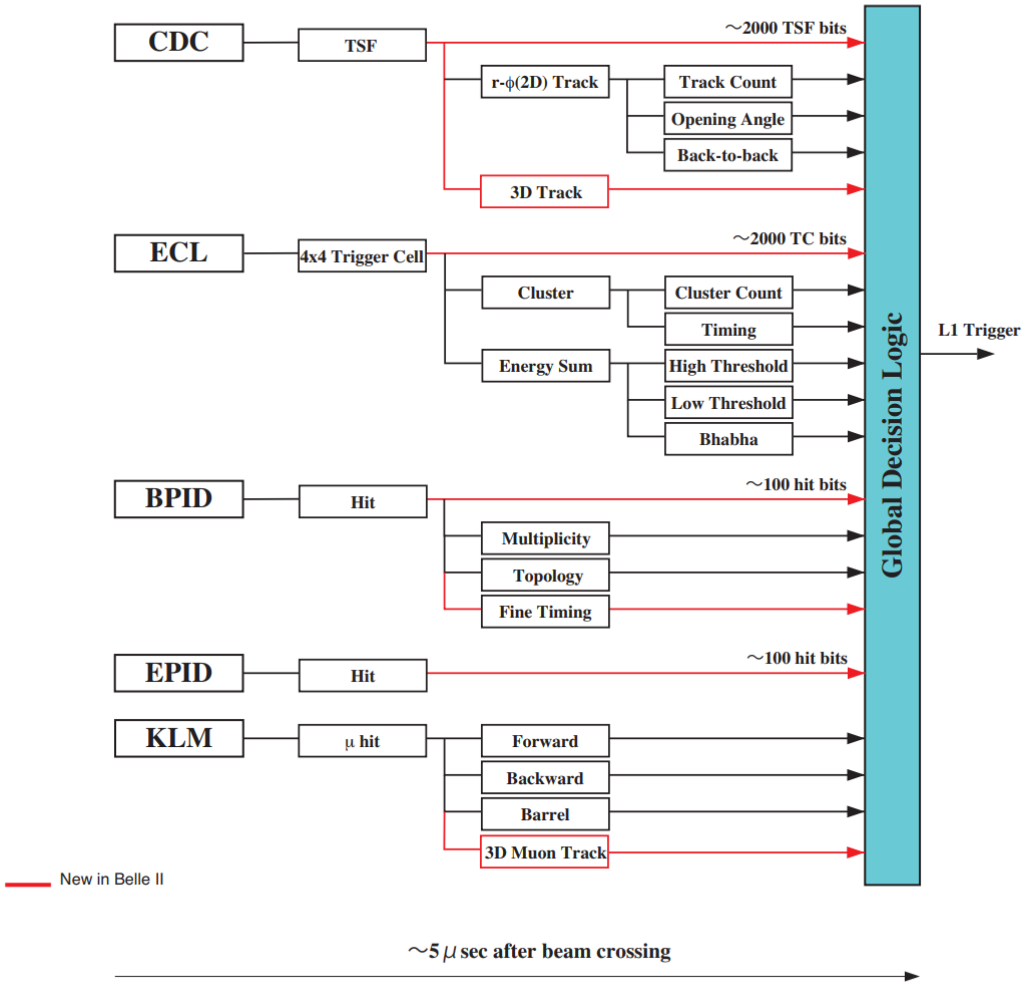


Figure 3.17: Schematic of the L1 trigger and its 5 sub-triggers. Newly added triggers for Belle II are marked in red.[16]

deposits. This allows the ECL trigger to monitor the luminosity by identifying Bhabha and $\gamma\gamma$ processes. The Barrel and end-cap PID systems provide timing information for charged particles. Finally, the KLM provides information about muon tracks. In addition to the events relevant for the physics program of Belle II, the L1 trigger also keeps $\sim 1\%$ of Bhabha events for luminosity measurements and detector calibration purposes.[16, 30]

After passing the L1 trigger system, the data gets further processed by the software based high level trigger (HLT). The HLT further reduces the recorded number of events by undertaking a full reconstruction of the events and filtering out events based on quantities like their track multiplicity, vertex position and energy deposits. This process is estimated to reduce the sample size by approximately 50%. After the initial filtering process the HLT also roughly categorizes the events based on various criteria, a process referred to as skimming. Some exemplary skim categories are: hadronic, Bhabha and τ events. Skimming is a very important process, that makes the sample sizes easier to handle for analysts, which gets more crucial as sample sizes increase.[16, 30]

For both the $|V_{cb}|$ and the slow pion study, the `hlt_hadron` skim of the whole pre long-shutdown 1 (LS1) data sample of the Belle II experiment was used. A hadron skim refers to a sample containing only events, that were categorized by the HLT to likely contain hadrons as decay products, based on a variety of criteria, e.g. track multiplicity. To be considered hadronic, the events must have 3 good tracks, i.e. tracks that can be reliably reconstructed, and additionally must pass a veto based on the trigger conditions for Bhabha events.

CHAPTER 4

Data Analysis at Belle II

4.1 Analysis Overview

Many processes of the analysis workflow are handled using the Belle II Analysis Software Framework (basf2) such as: the generation of Monte Carlo (MC) simulated samples by the data production group, the unpacking of real data and the reconstruction of particles. [39]

The basf2 framework is in active development and gets continuously updated and improved. In general basf2 is organized in modules, paths and packages. A module contains functions written in C++ for a specific part of the data processing. The modules are assigned to a path, i.e. an arrangement of modules in a linear order in which they get executed, where each module can in turn steer the direction of the process towards different paths depending on a condition. Finally, a package is a collection that contains multiple modules for an analysis or data processing task. For analysis purposes events are reconstructed using a steering file, i.e. a python code that initiates a path, adds modules to it and processes the data provided in analysis files (mDST) accordingly. After successfully processing an mDST file, the steering file defines an ntuple, which is saved using the ROOT file format. [40]

The analysis itself at Belle II follows roughly the steps illustrated in Fig.4.1. While the basics of these steps can be generalized to some degree, the implemented strategies can vary from analysis to analysis. The first step is the event reconstruction, where multiple long lived, so called final state particles, are combined to the short lived particle of interest also referred to as candidates, e.g. a B meson; a more detailed explanation of the reconstruction process is provided in Sec.4.3. This is done by processing mDST files with a self written steering file.

This reconstruction process is followed by an optimization of selection criteria for candidates based on a figure of merit obtained from MC simulations, the underlying concepts are discussed in Sec.4.4. Next, the reconstructed simulation and data samples will

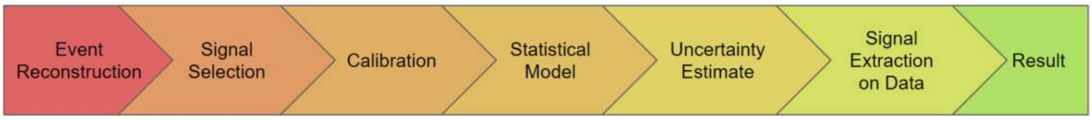


Figure 4.1: Simplified schematic of the typical analysis workflow.

be calibrated to account for potential mismodeling or biases during the data reconstruction process (see Sec.4.6).

The fourth step is dedicated to building a statistical model for a binned maximum likelihood fit based on the MC simulation and testing the model as well as the fit approach through bootstrapping. The method of bootstrapping is also employed for the estimation of systematic uncertainties of the measurement. In Sec.4.7, the statistical concepts utilized in these two steps are discussed.

Finally, after rigorous testing of the statistical model and fitting procedure as well as the statistical and systematic uncertainty estimation, the model will be used to extract the signal candidates from the data. Signal candidates refer to the particles, being sought in a given analysis.

4.2 Data and MC Samples

4.2.1 Data

The HLT hadron skim of the pre long shutdown 1 (LS1) data set was used for the studies presented in this thesis. As discussed in Sec.3.4, skims have the goal in mind to decrease the sample sizes by only including events that are potentially interesting for the analyst. The LS1 data sample includes all data recorded at Belle II from its start of operation in 2019 up to the temporary shut-down in summer 2022 for detector upgrade purposes, mainly the implementation of the missing PXD ladders of its second layer.

The sample size of the LS1 data corresponds to an integrated luminosity of $\sim 364\ \mathrm{fb}^{-1}$ of on-resonance data, taken at the $\Upsilon(4S)$ resonance with a center of mass energy of 10.58 GeV, containing $(387 \pm 6) \times 10^6$ B meson pairs. In general the data sets at Belle II are divided into experiments (exp), which are further divided into run periods. A run is a specific time frame in which the experiment operates and collects data. The runs within a period may have somewhat different beam parameters, e.g. slightly varying beam energies or vacuum conditions. Some runs are also dedicated to estimate e.g. beam-induced backgrounds or monitor the detector performance and calibration with cosmic muons.

For the $|V_{cb}|$ analysis, the off-resonance data, taken ∼5 MeV below the $\Upsilon(4S)$ resonance, was studied to monitor the agreement between data and MC simulated $q\bar{q}$ events. The overall data set size in fb^{-1} for the on- and off-resonance pre-LS1 data is presented in Tab.4.1. This summary is split into individual experiments and summed up, including their respective uncertainties. All luminosity values and their uncertainties are provided by the dedicated Belle II data processing group.

4.2.2 MC

MC simulated samples are produced by generating events using a event type specific generator followed by a simulation of the interaction of the particles with the matter of various detector layers and the magnetic field. In general the number of generated events is based on decay files, i.e. a file containing the branching fractions of a given particle. The branching fractions inside the decay file can be either set to only contain a given decay or be generic, i.e. a decay file containing the world average branching fractions provided by PDG. [10] The generator used to create an event is based on the event type: `PYTHIA`[41] is used for inclusive B and D meson decays, `EvtGen`[42] is used for exclusive B and D decays.

On-Resonance		Off-Resonance	
Exp.	$\int \mathcal{L}\, dt$ [fb^{-1}]	Exp.	$\int \mathcal{L}\, dt$ [fb^{-1}]
26	$54.795 \pm 0.009 \pm 0.351$	25	$24.574 \pm 0.006 \pm 0.158$
24	$85.642 \pm 0.011 \pm 0.549$	18	$8.482 \pm 0.004 \pm 0.055$
22	$32.060 \pm 0.007 \pm 0.206$	12	$8.691 \pm 0.002 \pm 0.056$
20	$3.788 \pm 0.003 \pm 0.025$	8	$0.808 \pm 0.001 \pm 0.006$
18	$89.900 \pm 0.011 \pm 0.576$		
17	$10.715 \pm 0.004 \pm 0.069$		
16	$10.294 \pm 0.004 \pm 0.066$		
14	$16.500 \pm 0.006 \pm 0.106$		
12	$54.573 \pm 0.005 \pm 0.350$		
10	$3.655 \pm 0.002 \pm 0.024$		
8	$1.663 \pm 0.003 \pm 0.011$		
7	$0.506 \pm 0.002 \pm 0.004$		
Total	$364.093 \pm 0.021 \pm 2.331$	Total	$42.556 \pm 0.007 \pm 0.273$

Table 4.1: Integrated Luminosity of the pre-LS1 on- and off-resonance data productions in fb^{-1} split into experiments. The first uncertainty is statistical and the second systematic.

The $q\bar{q}$ pairs as well as τ pairs are generated by `KKMC`[43] and `TAUOLA`[44]. Bhabha and two photon processes are generated by `BABAYGA.NLO`[45] and `AAFH`[46], respectively. The interaction of the particles with the detector layers is simulated using `GEANT4`. [16, 30]

Simulated MC samples are utilized to optimize the signal selection and also to extract the signal yield by building and testing statistical models based on the simulation. For all generated events in the MC sample the relations between particles and their origin, i.e. for example $B\bar{B}$ events or $q\bar{q}$ pairs, is known as well as the particle's properties like momenta, energy or polar angle. Furthermore, each particle has an associated ID and a MC PDG code, i.e. a numbering scheme provided by PDG, assigned to it. [10] The PDG code explicitly identifies the type of a particle. With all the information provided for the generated MC particles, one can assign signal flags, i.e. a boolean value of 1 or 0, to determine whether a decay chain has been correctly reconstructed.[16, 30]

The MC samples used for our studies are generic samples, i.e. samples where particles decay according to the branching fractions of their world averages provided by the PDG. The MC samples are run dependent, meaning they take the fluctuations of the beam energies of a given run into account during the simulation process, instead of assuming a constant center of mass beam energy of 10.58 GeV. The samples are split into the following categories, based on their event type: $B^0\overline{B}^0$, B^+B^-, $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$, and τ pairs. Each of the MC productions corresponds to a sample size of 1443.999 fb^{-1}, which is equivalent to four times the integrated data luminosity. Similarly to the data luminosity, a summary table for the luminosity of each MC sample is provided in Tab.4.2.

4.3 Event Reconstruction

At SuperKEKB B meson pairs get produced via the decay $\Upsilon(4S) \rightarrow B\bar{B}$ at a center of mass energy of 10.58 GeV. The center of mass energy SuperKEKB operates at is very close to the rest mass of the B mesons, therefore the $B\bar{B}$ pairs get practically produced at rest and

Process	exp7-18 $\int \mathcal{L}\,dt$ [fb^{-1}]	exp20-26 $\int \mathcal{L}\,dt$ [fb^{-1}]
$\Upsilon(4S) \to B^0\overline{B}^0$	744.764	699.235
$\Upsilon(4S) \to B^-B^+$	744.764	699.235
$e^-e^+ \to c\overline{c}$	744.764	699.235
$e^-e^+ \to d\overline{d}$	744.764	699.235
$e^-e^+ \to s\overline{s}$	744.764	699.235
$e^-e^+ \to u\overline{u}$	744.764	699.235
$e^-e^+ \to \tau^-\tau^+$	744.764	699.235

Table 4.2: Luminosity of the MC15rd simulated samples in fb^{-1} for each MC sample category.

wouldn't be separable, if it wasn't for the boost from the asymmetric beam energies. While the boost from asymmetric beam energies is sufficient to separate the B mesons, it is not large enough for them to reach the first detector layer of the PXD at 14 mm with a traveling distance of $\sim$ 130 µm before decaying.[16] Other short lived particles like D mesons and their excited states are also unable to reach the detector.

These short lived particles can be reconstructed by combining so-called final state particles, that have sufficiently long life times to be detected. During the reconstruction the final state particles get randomly combined, resulting in candidates. These candidates get selected or discarded based on requirements, that are enforced for various quantities, e.g. their invariant mass. The final state particles include: electrons, muons, photons, charged and neutral pions and kaons. Information about the momenta, energy and particle identification likelihoods of final state particles are provided by the detector.

It is not often the case, that all three types of information are directly available for a particle. The momentum of charged particles can be reconstructed by their tracks and they also leave hits inside of the ECL, depositing some energy in the ECL. Electrons however get completely stopped inside the ECL and deposit their entire energy. For charged particles with insufficient energy deposits in the ECL, the energy is estimated using a mass hypothesis based on the PID system.

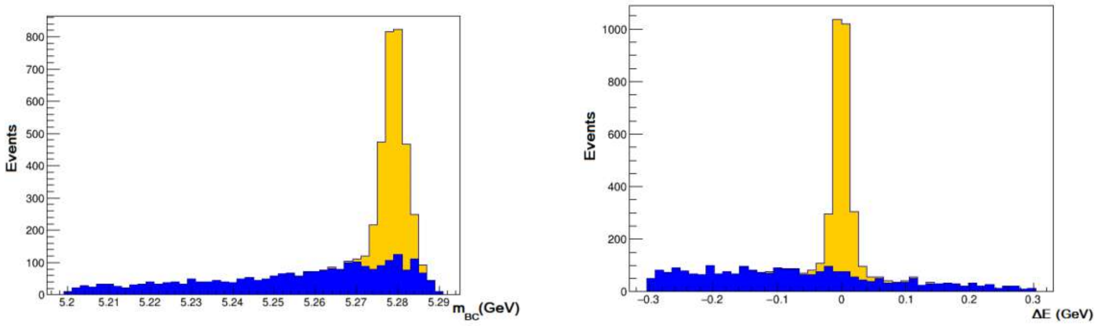
Neutral particles on the other hand don't leave any tracks in the CDC or PXD and their four momentum is mainly reconstructed based on their energy deposits in the ECL. Therefore their tracks have to be extrapolated to the IP based on a mass hypothesis, which is obtained from the PID system.

The combination of final state particles is best explained using an example. For this purpose we will use the decay $B^+ \to D^-\pi^+\pi^+$ with the subsequent decay $D^- \to K^+\pi^-\pi^-$. From the measured four momenta of the final state particles one can reconstruct short lived particles by taking the sum of the four momenta of a given number of particles. The four momenta of a kaon and two pions of opposite charge are combined to form a D candidate:

$$p_D = p_K + p_{\pi_1} + p_{\pi_2}. \tag{4.1}$$

During this combination process the reconstruction algorithm also takes charge conjugate decays into account, i.e. $D^+ \to K^-\pi^+\pi^+$. The newly combined D meson is then paired with two potential pions to reconstruct the B meson.

These particle combinations are random and of course not all of these combinations are correct, which leads to two basic categories of candidates, namely signal and combinatorial



(a) Beam constrained mass

(b) Energy deviation

Figure 4.2: The spectra of the m_{BC} and ΔE obtained from the reconstruction of the hadronic decay $B^+ \to D^- \pi^+ \pi^+$. In these illustrations the signal candidates, indicated in yellow, are stacked on top of the background candidates, indicated by the blue component.

background. Signal candidates are B mesons resulting from completely correctly reconstructed decay chains of the decay sought by the analyst. Background candidates on the other hand are falsely reconstructed candidates with similar kinematics.

To separate between signal and background events we can make use of the properties of the SuperKEKB collider. In case of hadronic decays two very important discriminators between background and signal events are the beam constrained mass m_{BC} and the energy difference between the reconstructed B meson and half the beam energy in the center of mass frame. The beam constrained mass is defined as:

$$m_{BC} = \sqrt{E^{*2}_{Beam} - \vec{p}^{*2}_{B}}, \tag{4.2}$$

where E^{*}_{Beam}[1] is the beam energy, which is known to be 5.29 GeV, due to the characteristics of the collider setup and $\vec{p}^{*2}_{B}$ is the reconstructed B meson momentum. This results in an expected peak at 5.28 GeV for the beam constrained mass for signal events, but a rather flat distribution for background events. This behaviour is illustrated by the distribution in Fig.4.2a.

A second powerful discriminator between signal and background events is the energy deviation ΔE of the reconstructed B meson candidate from the nominal beam energy:

$$\Delta E = E^{*}_{B} - E^{*}_{Beam}, \tag{4.3}$$

where E^{*}_{B} is the reconstructed B meson energy. For correctly reconstructed candidates a peak at zero can be observed in the ΔE distribution, which is depicted in Fig.4.2b. Background events exhibit a flat distribution.

Semileptonic decays always include a neutrino, which can't be detected due to their lack of any interaction with the detector material. Therefore the previously two mentioned variables m_{BC} and ΔE are only viable for hadronic decays, because the B meson energy and momentum can't be completely reconstructed.

One common approach of identifying signal events is the utilization of the low mass of less than 1.1 eV of the neutrinos. The invariant missing mass m^2_{miss} of the missing neutrino

---

[1] An asterisk indicates the center of mass frame.

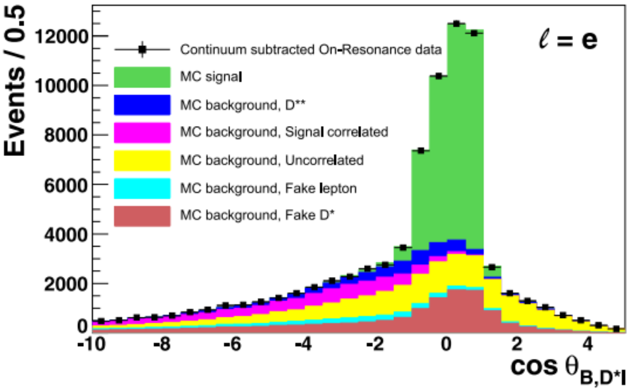


Figure 4.3: The binned distribution of $\cos\theta_{BY}$ for the electron channel of $B \to D^*\ell\nu_l$. The signal events only populate the region of [-1,1], while the background events are spread out over the range of [-10,5]. [47]

of a $B \to D^*\ell\nu_l$ decay can be determined as follows:

$$m^2_{miss} = \left(\begin{pmatrix} \sqrt{s}/2 \\ 0 \\ 0 \\ 0 \end{pmatrix} - p^*_\ell - p^*_{D^*} \right)^2 , \tag{4.4}$$

where p^*_ℓ and $p^*_{D^*}$ are the reconstructed four momenta of the lepton and D^* meson, respectively. Due to the very low neutrino mass a distribution with a peaking structure at zero can be expected for correctly reconstructed candidates.

For $B \to D^*\ell\nu_l$ and other semileptonic decays another powerful discriminator between signal and background events is the cosine of the opening angle θ_{BY}. It can be shown, that in the center of mass frame the direction of the B meson has to lie on a cone with the opening angle θ_{BY} around the direction of the incompletely reconstructed B meson system $Y(D^*\ell)$. The cosine of θ_{BY} is defined as:

$$\cos\theta_{BY} = \frac{2E^*_B E^*_Y - m^2_B - m^2_Y}{2|\vec{p}^*_B||\vec{p}^*_Y|}, \tag{4.5}$$

where E^*_Y and $|\vec{p}^*_Y|$ are the reconstructed energy and magnitude of the momentum in the center of mass frame of the Y system, respectively. The reconstructed mass of the Y system is denoted as m_Y, while m_B is the world average B meson mass of 5.28 GeV. Finally, $|\vec{p}^*_B|$ is the magnitude of the momentum of the B meson, calculated by $\sqrt{E^{*2}_B - m^2_B}$. The variable $\cos\theta_{BY}$ has the unique trait, that it only acts like a real cosine with a limited interval [-1,1] for signal events. This behavior is illustrated in an exemplary $\cos\theta_{BY}$ distribution for the electron channel of the $B \to D^*\ell\nu_l$ decay in Fig.4.3.

4.3.1 Full Event Interpretation

All semileptonic decays as well as the leptonic decays, $B \to e\nu$, $B \to \mu\nu$ and $B \to \tau\nu$, involve neutrinos. Because of the presence of the neutrinos in these decays, they suffer from missing kinematic information. This is a result of the absence of any interactions between

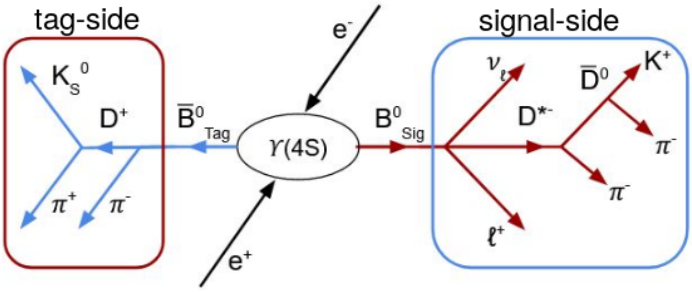


Figure 4.4: Schematic overview of the $B\bar{B}$-pair reconstruction. The left side shows the tag-side reconstruction of the hadronic decay $B \to D\pi$, while the signal-side on the right reconstructs $B \to D^*\ell\nu_l$.

the neutrinos and the detector material. However, thanks to the hermetic structure of the detector and its resulting high solid angle coverage, Belle II has the capability to measure the four momenta of the final state particles of both B mesons of a $\Upsilon(4S) \to B\bar{B}$ decay. Therefore it is possible to reconstruct the $\Upsilon(4S)$ mother, which provides the possibility to infer the missing momentum and energy of the neutrino with the knowledge of SuperKEKB's parameters. [48]

The concept of using the second B meson of an event to set constraints for the kinematics of the first B meson is referred to as tagging. For the tagging approach each $B\bar{B}$ pair is split into two parts, which is illustrated in Fig.4.4. On the one hand, there's the signal side, which contains the final state particles that show a distribution of tracks and energies hinting to the decay sought for a given analysis, e.g. $B \to D^*\ell\nu_l$. This B meson is the so-called B_{sig}. On the other hand is the tag-side with B_{tag}, which contains the remaining tracks and clusters of a random decay chain. [48]

The algorithm responsible for the tagging, implemented in basf2 is exclusive to Belle II and is called Full Event Interpretation (FEI) algorithm. The algorithm collects all tracks and clusters, that are not associated with B_{sig} and reconstructs a variety of $\mathcal{O}(10000)$ decay chains that are compatible with the detector information. Using boosted decision trees (see Sec.4.7.5), a probability is calculated for each decay chain to determine its likelihood of being a correct reconstruction. The FEI's tagging algorithm can operate in three different modes: [48]

- **Hadronic**: In this mode only hadronic decays are reconstructed for the tag-side. Therefore only decay chains, where all final state particles are detectable are reconstructed. This allows to determine the momentum of the neutrino via the relation $p_\nu = p_{e^+e^-} - p_Y - p_{B_{tag}}$, where Y corresponds to incompletely reconstructed B meson, e.g. $D^*\ell$ for $B \to D^*\ell\nu_l$. Additionally this approach can suppress background events, since all visible tracks must be used during the $\Upsilon(4S)$ reconstruction. If tracks are left over after the reconstruction this implies a background event. This approach only has an efficiency of $\mathcal{O}(0.1\%)$ for correctly reconstructing the tag-side resulting in a very high loss of potentially correct signal candidates. [48]

- **Semileptonic**: Using semileptonic decays for the tag-side reconstruction results in an efficiency of $\mathcal{O}(1\%)$, due to the higher branching fractions. Due to the missing tracks and clusters from the neutrino involved in semileptonic decays, one loses some information about B_{tag}. This results in the inability to calculate the neutrino momentum from its kinematic constraints and lessens the power of its background suppression. [48]

- **Inclusive**: This method is also referred to as untagged. For this approach all tracks and clusters not belonging to the final state particles of B_{sig} are collected and combined to the so-called rest of the event. This approach has an efficiency of $\mathcal{O}(100\%)$, because no decay chain reconstruction of the tag-side is attempted. As a trade off one can't place any constraints on the kinematics of the neutrino and no information about the B_{tag} decay chain is obtained. [48]

4.4 Candidate Selection Methods

4.4.1 Rectangular Cuts

During the reconstruction process and subsequent background suppression, certain selection criteria are required to be fulfilled by B meson candidates. These selections involve setting thresholds on a variety of variables, effectively defining a multi dimensional rectangular space which reconstructed candidates must lie within. Hence, these criteria are referred to as rectangular cuts. These cuts are necessary to increase the significance of the signal and reduce the sample size. An example for an rectangular cut would be a threshold on the reconstructed mass of $1.8 < m_D < 1.9$ GeV for a D meson. [16, 19]

While the simplicity of this concept makes cuts rather easy to implement and interpret, there are also drawbacks. When applying rectangular cuts the correlations between variables is not taken into account. Assuming the reconstruction of $D \to K\pi$: The D meson is reconstructed by combining the four momenta of the pion and the kaon. When applying a cut on the magnitude of the momentum of the π this will affect the mass distribution of the D mass, which we place another constraint on. However, the correlation between these two variables is ignored when placing rectangular cuts. Any candidates that might still fit into the required D mass window, but do not fit into the selected pion momentum region are still discarded. [16, 19]

Another issue is the cut optimization. Some variables like the beam constraint mass or the energy difference between the reconstructed B meson and half the center of mass energy show a distinct peak for signal candidates and rather flat distribution for background noise, which allows for an estimation of the selection criteria by eye by simply basing the selection criteria on the signal peak. Other variables like the Fox Wolfram moment ratio R_2 (see Sec.4.5) might have regions that are predominantly populated by signal or background events, however these regions overlap. This overlap makes it hard to determine the best value for the threshold, which it typically determined by optimizing a figure of merit.[16, 19]

4.4.2 Figure of Merit

In general the figure of merit (fom) summarizes the performance of the chosen constraints put on the variables of the reconstructed candidates and final state particles within a single quantity. This quantity can be optimized using the MC simulation, where the origin of all reconstructed candidates is known. The definition of the figure of merit can vary and depends on the context of the analysis. For precision measurements such as a $|V_{cb}|$ analysis based on $B \to D^*\ell\nu_l$ decays, the following quantity is maximized: [19]

$$fom = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bg}}}, \tag{4.6}$$

where N_{sig} and N_{bg} are the numbers of signal and background candidates, respectively. The denominator $\sqrt{N_{sig} + N_{bg}}$ is equivalent to the poisson error, i.e. the statistical fluctuation of the number of events. The poisson error in the figure of merits originates from the fact, that a number of events in a given time frame is described by the means of the Poisson distribution.[19]

4.5 Continuum Suppression

Events where $q\bar{q}$ pairs are produced by the e^+e^- collisions, rather than $\Upsilon(4S)$ mesons, are referred to as continuum events, where q can be one of the following quarks: u, d, c, s. To get an idea of the probability of $q\bar{q}$ production, one can take a look at the cross sections in Tab.3.1.

While most leptonic background events get rejected by the trigger, it is less effective at identifying and discarding the hadronic continuum events. Another feature of continuum events are their generally flat and continuous distributions with respect to the invariant mass of reconstructed B meson candidates and their intermediate daughter particles, such as D or D^* mesons. As a consequence, it is harder to separate continuum from signal candidates, due to missing background peaks. [16]

Nevertheless, there is a very useful characteristic to distinguish between $q\bar{q}$ and $B\bar{B}$ pair events, namely their event shape, illustrated in Fig.4.5. These distinct event shapes originate from the difference in the available momentum for $q\bar{q}$ and $B\bar{B}$ events. Due to the B meson pair mass being barely below the threshold of the $\Upsilon(4S)$ resonance of 10.58 GeV the B mesons are practically produced at rest and their decay vertices only get separated by the boost from the asymmetric e^+e^- energies. This results in an uniform event shape for $B\bar{B}$. The quark pairs on the other hand, are produced with higher momentum leading to narrow jets.

The distribution of particles in an event can be expressed by the Fox Wolfram moments H_l introduced by Fox and Wolfram: [50]

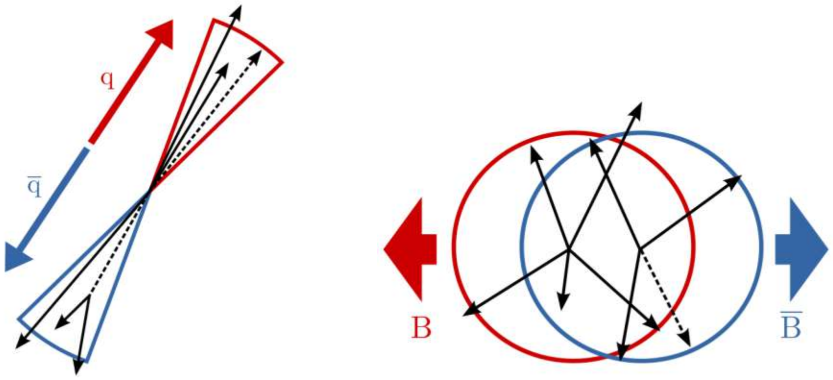
$$H_l = \sum_{i,j} \frac{|\vec{p}_i^*||\vec{p}_j^*|}{s} P_l(\cos\phi_{ij}^*), \tag{4.7}$$

where $\vec{p}_i^*$[2] and $\vec{p}_j^*$ are the momenta with i and j running over all final states particles of the event and ϕ_{ij}^* is defined as the opening angle between the two particles i and j. Finally, s is the energy in the center of mass frame and P_l the Legendre polynomial of the order l. The ratio of the second to the zeroth Fox Wolfram $R_2 = H_2/H_0$ serves as a powerful discriminator between continuum and $B\bar{B}$ events, owing to its peak near zero in the R_2 distribution. This feature is depicted in Fig.4.6 and allows for a selection that has a minimal impact on potential signal candidates, while rejecting a majority of $q\bar{q}$ events.

Based on the Fox Wolfram moments the Belle collaboration developed modified Fox Wolfram moments. The modified moments are separated into two categories: moments

---

[2] The asterisk denotes the center of mass frame.



(a) Tightly accumulated jets originating from $q\bar{q}$ events.

(b) Uniform $B\bar{B}$ event shape.

Figure 4.5: Comparison between the continuum and $B\bar{B}$ event shapes. [49]

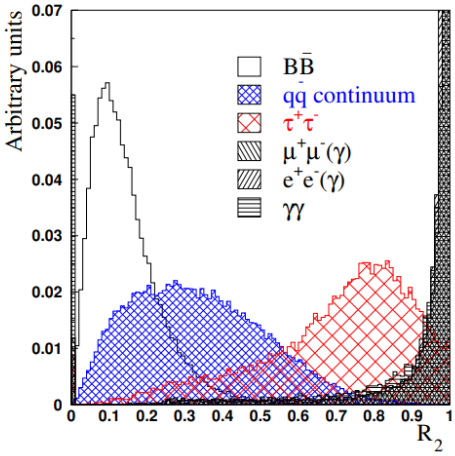


Figure 4.6: An example for the distinction power between signal and continuum events of the Fox Wolfram moments ratio R_2. [19]

that involve particles from both the rest of the event and the reconstructed candidate H^{so}_{xl} and moments that only include rest of event particles H^{oo}_{l}. Furthermore both modified moments H^{so}_{xl} and H^{oo}_{l} are divided into sub categories based on the order l of the Legendre polynomial, where $l \in [0, 4]$. Lastly, the H^{so}_{xl} moments are classified with respect to the charge of the particles involved in the calculation, i.e. $x = c, n, m$ for charged, neutral and missing particles, respectively. For even orders of l the formula for H^{so}_{xl} reads as: [19, 16]

$$H^{so}_{xl} = \sum_{i} \sum_{jx} |p_{jx}| P_l(\cos\theta_{i,jx}), \tag{4.8}$$

where i refers to the final state particles of the B candidate, jx to the final state particles of the rest of the event from a given class x and $\theta_{i,jx}$ is the angle between two particles. In case of odd orders of l follows:

$$H^{so}_{cl} = \sum_{i} \sum_{jx} Q_i Q_{jx} |p_{jx}| P_l(\cos\theta_{i,jx}),\ H^{so}_{nl} = H^{so}_{ml} = 0, \tag{4.9}$$

where Q_i, Q_{jx} refer to the charges of a given particle. The modified Fox Wolfram moments for even l run over the rest of event particles j and k are defined as:

$$H_l^{oo} = \sum_j \sum_k |p_j||p_k| P_l(\cos\theta_{j,k}), \tag{4.10}$$

while the moments for odd l yield:

$$H_l^{oo} = \sum_j \sum_k Q_i Q_k |p_j||p_k| P_l(\cos\theta_{j,k}). \tag{4.11}$$

Furthermore, the H_{xl}^{so} and H_l^{oo} moments are normalized with respect to H_0^{max} and $(H_0^{max})^2$, respectively, where $H_0^{max} = 2(E_{beam}^* - \Delta E)$. [19, 16]

Another set of effective variables for continuum suppression is based on the thrust vector of the final state particles used to reconstruct the B candidates. The thrust vector is the direction that maximizes the projection of all final state particle momenta associated with a B candidate or the rest of the event. The cosine of the angle between the thrust vectors of two B mesons will exhibit a flat spectrum, due to its uniform shape, whereas the $q\bar{q}$ pairs will peak toward one, due to their jet shape. [19, 16]

4.6 Data and MC Calibration

Some discrepancies between the data and MC samples may occur, due to mismodeling in the simulation for various sub-detectors during the MC production process. Additionally some biases might be observed in candidates that were reconstructed from data. For example the application of a biased magnetic field map during the data reconstruction process will affect the reconstructed momenta and result in invariant mass peak shifts.

These observed data reconstruction biases and mismodeled properties in the MC production are corrected by scaling certain variables of the final state particles in the samples during the reconstruction process or by multiplying correction weights after the reconstruction. The calibration process is done using python packages and values provided by the performance group of Belle II, which also offers guidance with respect to the application of the packages. Not only are these calibrations necessary for correctly interpreting the distribution of signal and background candidates in the data, since the estimated distributions are based on the simulation, but also for the estimation of systematic uncertainties.

While the calibration of the data and MC is an integral part of every physics analysis at Belle II, the amount and type of corrections vary depending on the studied decay: e.g. a hadronic decay like $B \to D^*\pi$ does need a calibration for mismodeling in the MC with respect to the particle identification efficiency for hadronic final state particles while a semileptonic $B \to D^*\ell\nu_l$ decay needs a correction for the lepton identification efficiency in addition to a Bremsstrahlung correction for the momentum of electron candidates. Due to the analysis specific nature of the calibrations, a complete run-down of all used corrections will be provided in dedicated sections Sec.5.3, Sec.5.4 and Sec.6.3 for the slow pion and $|V_{cb}|$ study, respectively.

4.7 Statistical Concepts

4.7.1 Poisson Distribution

In particle physics one studies a given number of events, that occur randomly and independently from each other, over a given time span at an average constant rate. In the context of particle physics, an event refers to a decay event. Due to the nature of the events, the optimal approach is to model them by utilizing the discrete Poisson distribution. The Poisson distribution is defined by the following probability density function: [10, 19]

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}, \tag{4.12}$$

where k is the number of events, λ is the expected event rate in a given time interval and e is the Euler's number. The Poisson distribution is discrete and therefore requires k to be a positive integer. Additionally the variance and mean of the distribution are both equal to λ, which results in statistical uncertainties being estimated by $\sqrt{\lambda}$. Another characteristic of the Poisson distribution is its convergence to the Gaussian distribution for large event rates λ, in accordance with the Central Limit theorem. [10, 19]

4.7.2 Maximum Likelihood

In general, the model for a given distribution of an observable is based on a probability density function (pdf). Due to the nature of the data in particle physics, we assume a model $f(\vec{x};\vec{\theta})$ with the parameters $\vec{\theta}=(\theta_1,...,\theta_k)$ based on a Poisson distribution (see Sec.4.7.1) for the measured data values $\vec{x}=(x_1,...,x_k)$ of an observable. The model also often includes nuisance parameters $\vec{q}$ that can be used to quantify the impact of various detector resolution effects on the relation between $\vec{\theta}$ and $\vec{x}$, which leads to the model $f(\vec{x};\vec{\theta},\vec{q})$. For simplicity, this section will stick to the notation $f(\vec{x};\vec{\theta})$ without nuisance parameters.[10, 19]

The probability of an observed data values $\vec{x}_0$ occurring is referred to as the likelihood $L(\vec{\theta})=f(\vec{x}_0;\vec{\theta})$. In case of multiple measurements, the overall likelihood can be written as a product of all individual measurements:[10, 19]

$$L(\vec{\theta})=\prod_{i=0}^{N} f(\vec{x}_i;\vec{\theta}) \tag{4.13}$$

One then seeks the optimal values of the parameters θ, under which the observed data is most probable. This is achieved by looking for a set of parameters, that maximize the logarithm of the likelihood $L(\vec{\theta})$ for each individual parameter θ_i:[10, 19]

$$\frac{\partial\ \log L(\vec{\theta})}{\partial\theta_i}=0,\ \ i=1,...,k. \tag{4.14}$$

Since the solution has to be obtained numerically, it is easier to calculate the minimum of the negative logarithm of the likelihood:

$$-\log L(\vec{\theta})=-\sum_{i=0}^{N}\log f(\vec{x}_i;\vec{\theta}). \tag{4.15}$$

Due to the large number of events in particle physics, $\mathcal{O}(10^8)$ B mesons in case of the current Belle II data set, the observed data values $\vec{x_i}$ are typically binned in N bins. Dividing the data into multiple bins converts the individual sets of $\vec{x_i}$ to vectors $\vec{n}(n_1, ..., n_N)$, where each entry n_i represents the number of events in a bin $i = 1, ..., N$. [10, 19]

The number of events in each bin is associated with a Poisson distribution with the mean $\mu_i(\vec{\theta})$ and a probability $f(n_i; \mu_i)$. The distribution over all bins is Poisson distributed with a mean of $\mu = \sum_i \mu_i$. This relationship allows to maximize the likelihood based on the binned event counts, where the content n_i of each bin is treated as independent. These properties yield the following likelihood $L(\vec{\theta})$ as a product of Poisson probabilities to be maximized:[10, 19]

$$L(\vec{\theta}) = \prod_{i=1}^{N} f(n_i; \vec{\theta}). \tag{4.16}$$

Alternatively, one can use the likelihood of an saturated model as a reference and maximize the likelihood ratio $\lambda(\vec{\theta}) = f(\vec{n}; \vec{\theta})/f_{model}(\vec{n}; \hat{\vec{\mu}})$. The likelihood $f_{model}(\vec{n}; \hat{\vec{\mu}})$ corresponds to a model with a set of parameters $\hat{\vec{\mu}}$ with one parameter for each bin i, that is equal to the expected bin counts, therefore $\hat{\vec{\mu}} = (n_1, ..., n_N)$. As a result, the saturated model perfectly reproduces the observed data and can be used as a benchmark. [10, 19]

Often the double of the negative logarithm of the likelihood ratio is minimized instead of the maximization:

$$-2 \ln \lambda(\vec{\theta}) = 2 \sum_{i=1}^{N} \left[\mu_i(\vec{\theta}) - n_i + n_i \ln \frac{n_i}{\mu_i(\vec{\theta})} \right], \tag{4.17}$$

where $-2 \ln \lambda(\vec{\theta})$ decreases as the agreement between $f(\vec{n}; \vec{\theta})$ and the saturated model improves. A key property of the minimum is that for large means μ it follows a χ^2 distribution according to Wilk's theorem. When comparing the event count between a binned data and MC simulation, sub-divided in multiple categories, the χ^2 distribution has i-1 degrees of freedom, where i is the number of categories. E.g. if a data sample is divided into four sub-sets based on the momentum of a particle, each sub-set would represent a category. [10, 19]

4.7.3 χ^2-Test

A goodness of fit test evaluates how well the observed data agrees with a model. One such test is the χ^2-test. While a goodness of fit test uses a p-value to quantify the result, similar to a significance test, it does not involve a comparison with an alternative hypothesis. Checks used to validate a model, like the χ^2-test, are also known as closure tests.[10]

As already mentioned in Sec.4.7.2, the maximum of the logarithm of a likelihood $L(\vec{\theta})$ follows a χ^2 distribution for a large number of events as they are often encountered in particle physics. Additionally it is also known, that the number of events in each bin is Poisson distributed, which can be approximate by a Gaussian distribution for high event counts. In this case, the logarithm of the likelihood function contains a sum of squares and converges with the method of least square and also follows the χ^2 distribution. Therefore a given set of parameters $\vec{\theta}$ optimizes both the minimum of χ^2 and maximum of the likelihood. In case of a completely uncorrelated, observed event counts y_i, this reads as: [10]

$$\chi^2(\theta) = -2 \ln L(\vec{\theta}) + \text{constant} = \sum_{i=1}^{N} \frac{(y_i - \mu(x_i; \vec{\theta}))^2}{\sigma_i^2}, \tag{4.18}$$

where $\mu(x_i; \vec{\theta})$ is the mean of the expected Gaussian of the data or equivalently the theoretically expected number of events in a bin, and σ_i the variance of the Gaussian distribution.

In case of correlated event counts y_i the χ^2 function gets slightly altered by the necessary implementation of the covariance matrix $V_{ij} = cov[y_i, y_j]$:[10]

$$\chi^2(\theta) = (\mathbf{y} - \mu(\vec{\theta}))^T \mathbf{V}^{-1} (\mathbf{y} - \mu(\vec{\theta})). \tag{4.19}$$

By integrating over the probability function $f(t; n_{dof})$ of χ^2, with the number of degrees of freedom n_{dof}, one obtains the p-value:

$$p = \int_{x^2_m in}^{\infty} f(t; n_{dof}) dt, \tag{4.20}$$

where t is equivalent to 4.19. The p-value quantifies the probability of the model correctly describing the observed data.[10]

4.7.4 Bootstrapping and Error Estimations

Bootstrapping is a valuable statistical technique, which can be used for validating fit models and estimating systematic uncertainties. The core concept of bootstrapping is based on creating multiple samples via resampling with replacement of the original data. These samples are referred to as toy simulations. In case of a binned distribution, this approach creates new samples by effectively varying the event count in each of the n bins. The bootstrapping process is done as follows: [10]

- Creating a toy samples from the original data set by resampling with replacements. Typically in the order of $\mathcal{O}(10^3)$ toys are created to have significant statistics for the distribution estimate.
- Determine the variable of interest for each resampled data set.
- From the calculated values a Gaussian distribution is obtained with a mean μ and standard deviation σ. The parameters of this distribution can be used to estimate systematic uncertainties or test the robustness of an underlying fit model.

While Bootstrapping is versatile and conceptually a rather simple approach, its results are heavily reliant on the size of the original data sample and can be computationally intensive for large samples.

Using the bootstrapping approach, the systematic uncertainties of a measurement can be determined. As mentioned in Sec.4.6, the data and the MC samples must both be calibrated for various mismodeled properties during detector simulations or biases during the reconstruction of final state particles. These calibrations are done by weighting each reconstructed candidates. The applied weights are based on performance studies done by the Belle II performance group. Each of these studies in itself is reliant on multiple parameters and has its own uncertainties attached to their result, these uncertainties can be both statistical and systematic. Therefore in each physics analysis, one has to estimate the impact of the uncertainties of a multitude of calibration factors. To estimate the systematic uncertainty, the bootstrapping method can be used to resample the original data set by varying all correction weights in accordance with their assigned uncertainty. This results in a variation of event counts in each bin and therefore the fit result and the measured variable of interest. An estimation of this style is done for each possible source of systematic uncertainties, which results in a Gaussian distribution with mean μ and standard deviation σ. By comparing the standard deviation with the mean of the distribution, one can estimate the relative impact of each systematic uncertainty on the measurement.[10]

Another use case for bootstrapping is the validation of the robustness of the fit model on simulated data before applying it to reality. In this case, the varied data samples are based on the model used for the maximum likelihood fit. Each of these samples is then fit using the original model, which results in k fitted signal strengths and uncertainties of said strengths. For each fit result a pull can be calculated: [19]

$$pull = \frac{\mu - \hat{\mu}}{\sigma}, \tag{4.21}$$

where μ is the estimated signal strength, $\hat{\mu}$ the known, true signal strength from the model and σ is the estimated uncertainty of μ. The resulting pull distribution can then be fitted by an unbinned Gaussian to obtain the mean and standard deviation of the distribution. If our model fits the data unbiased, then $\hat{\mu}$ is correctly estimated on average and the mean of the pull distribution should be zero. Otherwise, a bias will be observable in either direction. Similarly the estimated uncertainty of the model can be judged: if the uncertainty is correctly estimated the standard deviation of the Gaussian will be 1, while a width above 1 hints towards an underestimation of the uncertainty, while a width below 1 hints towards an overestimation of the uncertainty by the model. [19]

4.7.5 Boosted Decision Trees

In particle physics decision trees are commonly used to reduce background noise. To determine, whether candidates are kept or discarded, a signal probability is calculated based on a variety of n variables $\vec{x} = (x_1, ..., x_n)$, which can be used to place a cut. Decision trees are the basis for the multivariate tagging algorithm of the FEI at Belle II, discussed in Sec.4.3.1, and can also be used to determine the PID likelihood of individual final state particles. The concept of PID likelihoods is described in Sec.3.3. [19, 51]

A decision tree is a supervised learning algorithm consisting of multiple layers, with each layer containing multiple nodes. At each node, the data is split based on a threshold for a given variable, e.g. polar angle or momentum. Every node leads a candidate to another node and this process is continued through all subsequent layers until reaching a leaf, also referred to as terminal nodes. Each terminal node represents a signal probability. The conceptual design of the hierarchical splitting of a decision tree is illustrated in Fig.4.7. Supervised machine learning involves two steps: a fitting-phase, where training data with known signal and background labels is used, i.e. an MC simulated sample, and an application-phase, where the decision tree is applied.[19, 51]

Many shallow decision trees with few variables, so-called weak classifiers, can be combined to one stronger classification model, i.e. the boosted decision tree. This technique is known as boosting. During the fitting-phase, weak classifiers are created iteratively and the result of each tree is assigned a weight, depending on their accuracy. The output of all trees is then summed up to a weighted average. Additionally, each candidate in the training sample is assigned a weight, which is adjusted after each iteration, depending on how frequently it is misclassified, to correct the errors made by the previous trees.[19, 51]

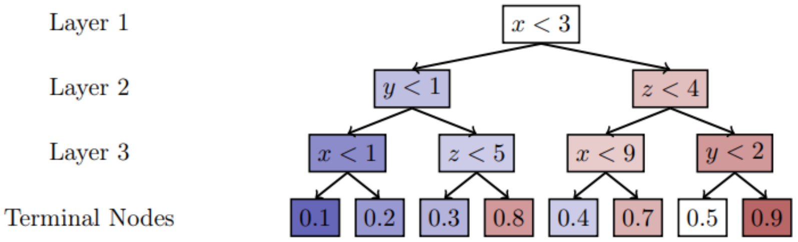
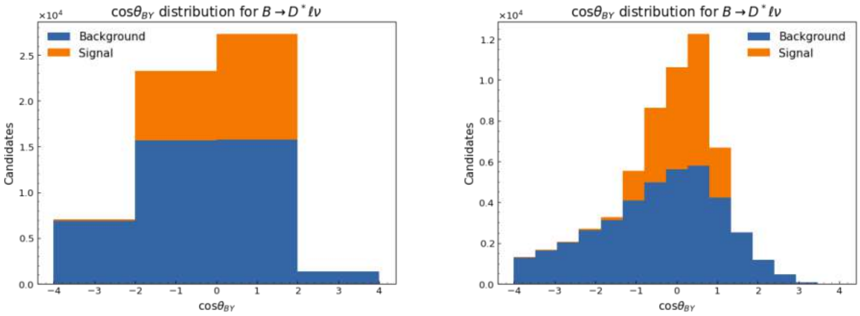


Figure 4.7: Schematic of a three layer decision tree. Each candidate has to traverse the tree until a terminal node is reached. A cut based decision is made at each node of the path for every candidate, here represented by the thresholds for three arbitrary variables x, y, z. This process ends, when a candidate finally reaches a terminal node that represents the probability of it being a signal candidate. The probability is based on the percentage of signal labeled candidates from the training-data that ended up on the node. [51]



(a) Too low bin count, leading to a loss of information about the shape.

(b) A reasonable bin count makes it possible to identify the shape of the signal region.

Figure 4.8: An example based on the $cos\theta_{BY}$ distribution of reconstructed $B \to D^*\ell\nu$ candidates showing the impact the choice of binning has to the underlying distribution. From Fig.a on the left with a number of 4 equal bins it can only be estimated, that our signal candidates mainly populate the region of [-2,2]. By increasing the number of bins to 15 the resolution of the distribution changes, revealing the structure of the signal component in the region of [-2,2], showing us that the main region populated by the signal is actually the region [-1,1].

4.7.6 Unfolding

In Sec.4.7.2 it was mentioned, that we look at our data in bins. Each bin has a well defined width within the distribution of the variable of interest, e.g. $\cos\theta_{BY}$. In general, binning simplifies the visualization of the distributions, making it easier to recognize patterns in larger data samples. However, it is essential to carefully choose an appropriate number and width of the bins, which should be based on the resolution of the observable. If the number of bins is chosen too low one loses information with respect to the shape. This behaviour is illustrated in Fig.4.8.

If the widths of the bins are too narrow, it can lead to measured events falling into wrong bins due to the finite detector resolution and limited acceptance. The procedure of transforming the measured distributions distorted by detector effects to the true distribution is called unfolding. There are a variety of different unfolding approaches, among others: Bayesian unfolding, matrix inversion, singular value decomposition unfolding and bin-by-bin unfolding. However, we will focus solely on bin-by-bin unfolding, which is a rather intuitive unfolding approach, that treats each bin independently and doesn't take possible correlations between bins into account. [10, 52]

Every unfolding method starts by constructing a migration or response matrix. The elements of the response matrix correspond to the probability, that a candidate with a simulated value in bin j has a reconstructed value in bin i:

$$\mathcal{M}_{ij} = \mathcal{P}(\text{measured value in bin i}|\text{ true value in bin j}). \tag{4.22}$$

Since the bin-by-bin unfolding approach doesn't take correlations between bins into account, only the diagonal elements are relevant. Simply by comparing the number of simulated events $N_{true,i}$ in bin i to the number of measured events $N_{rec,i}$ in the same bin by calculating an efficiency with the ratio $N_{true,i}/N_{rec,i}$ a correction factor can be obtained. By multiplying these correction factors to the number of measured events in each bin, one can obtain the corrected number of events.[10, 19]

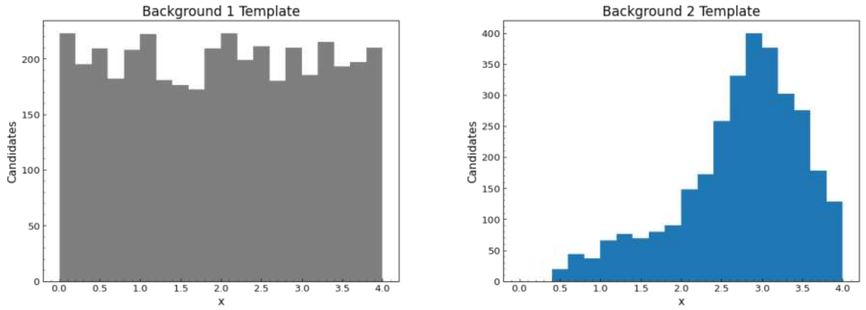
4.8 Template Fits

Measurements in particle physics are based on the prediction of the underlying physical processes in a data sample. This task is quite complex, due to background noise from misreconstructed candidates which show similar kinematics. This background can originate from a variety of different sources, e.g. $q\bar{q}$ events, falsely identified leptons or intermediate daughters like D mesons that are misreconstructed due to wrong final state particle combinations. In order to determine the observable of interest, one needs to reliably separate the signal candidates from the complex background noise. A popular approach to predict the number of signal candidates is by utilizing a template fit.

The concept of template fits is based on the obtained distributions of the signal and background candidates for a given observable (e.g. ΔE, cosθ_{BY}) from MC simulated samples. The shapes, i.e. the number of events in each bin of a binned distribution, predicted from the simulation of the various components within the spectrum are referred to as templates. The signal component and various processes that contributes to the background have their own dedicated template. An illustrative example containing two background templates and one signal template, is given in Fig.4.9.[10, 19]

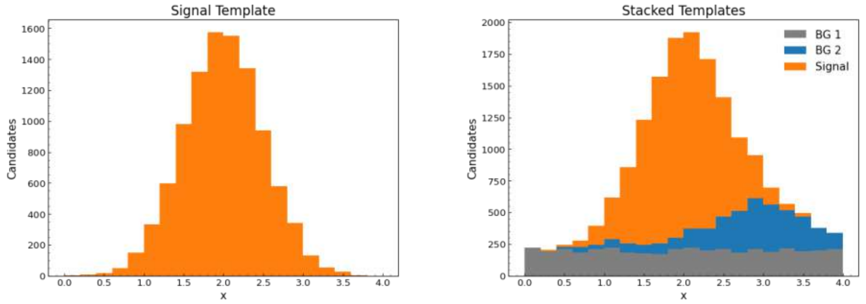
During the fitting process each template shape is in principle fixed, but nuisance parameters can be added to the model that allow some wiggle room for the shapes to estimate systematic uncertainties. However they get normalized in accordance with the number of events in the data sample by determining a factor of N_{rec}/N_{MC}, where N_{rec} and N_{MC} is the the number of events of the data and the simulated sample, respectively. The data from the experiment contains an unknown mixture of events.[10, 19]

A model is then built by adding together all templates. The sum of all templates is then varied by adjusting the scaling factor for each individual template. This is done by the means of an maximum likelihood estimator. We elaborated further on the concept of maximum likelihood fits in Sec.4.7.2. The result of the fit is then effectively the normalization factor for each template that most likely fits the shape of the data. By multiplying the number of signal events in the MC template with the normalization factor obtained by the fit, one then obtains the number of signal events that are predicted in the data.[10, 19]



(a) Uniform background shape.

(b) Peaking background shape.



(c) Signal shape.

(d) Sum of all templates.

Figure 4.9: A model in a template fit is based on the shapes of various components in a MC sample. In the pictured example the templates include a flat and peaking background in the upper row and a signal component in the lower left. Flat backgrounds typically originate from $q\bar{q}$ background events, while peaking backgrounds often originate from misreconstructed B meson pairs. The sum of all three templates is depicted in the lower right. All shapes are shown for an arbitrary observable x.

4.8.1 Implementation in pyhf

The python histfactory (pyhf) is a software tool that is designed for creating statistical models $f(x|\theta)$ for a binned analysis, which predicts the probability of the observed data x for a set of parameters θ of a given hypothesis. These models are then fitted onto the data based on a binned maximum likelihood fit for the likelihood $L(\theta)$ of the model describing the data. A detailed description of maximum likelihood fits is given in Sec.4.7.2. [53]

The pyhf library builds its statistical models based on the definition of Histfactory. [54] Conceptually the statistical models of the Histfactory are based on the approach of template fits. The likelihood specification is based on the following model: [53]

$$f(n, a \mid \eta, \chi) = \prod_{c\in\text{channels}} \prod_{b\in\text{bins}_c} \text{Pois}(n_{cb} \mid \nu_{cb}(\eta, \chi)) \prod_{X\in\mathcal{X}} c_X(a_X \mid \chi), \tag{4.23}$$

where the expected event rate ν_{cb} is defined as:

$$\nu_{cb}(\phi) = \sum_{s\in\text{samples}} \nu_{scb}(\eta, \chi) = \sum_{s\in\text{samples}} \left(\prod_{k\in\mathcal{K}} \kappa_{scb}(\eta, \chi)\right) \left(\nu^0_{scb}(\eta, \chi) + \sum_{\Delta\in\mathcal{A}} \Delta_{scb}(\eta, \chi)\right). \tag{4.24}$$

The binned distribution of the model $f(x|\theta)$ is represented by a channel. A model may include one or multiple independent channels and each of them is defined as the binned sum

n of events of the templates (signal, $q\bar{q}$ events, etc.) incorporated within the distribution, where each template is represented as a sample. Additionally the observed data may also include auxiliary data a, i.e. additional data obtained from simulations to constraint nuisance parameters, leading to $x = (n, a)$. Each set of parameters θ is split into free and constraint parameters, which are denoted as η and χ, respectively. In pyhf all models are defined using a plain-text JSON format for the ease of sharing and reproducibility. [53]

The expected event rate ν_{cb} is defined by its nominal rate ν^0_{scb} per sample, and its multiplicative κ and additive modifiers Δ. One of these modifiers is an unconstrained normalisation, which scales the model according to the data sample size and is used to derive the amount of signal events. Other modifiers include the statistical MC uncertainties and uncorrelated as well as correlated shape uncertainties. Further information about all available modifiers can be found in the official documentation in Ref.[53].

CHAPTER 5

Calibration of the Slow Pion Tracking Efficiency at Belle II

5.1 Analysis Overview

This chapter is dedicated to describing the calibration of the reconstruction efficiency of charged, slow pions for the LS1 data set and its run dependent MC simulation. A description of both samples is provided in Sec.4.2. Slow pions (π_s) refer to pions with a momentum below 300 MeV. Detecting these low momentum pions is challenging since they can't escape the inner detector layers due to the helix movement caused by magnetic field inside the detector.

The calibration is performed by measuring the relative tracking efficiency between data and MC. This is achieved by comparing the correctly reconstructed number of slow pions inferred from $B^0 \to D^{*-}\pi^+$ decays with the subsequent decay $D^{*-} \to \overline{D}^0\pi_s^-$. The subsequent decay $D^{*-} \to \overline{D}^0\pi_s^-$ is chosen due to small mass difference of ∼145 MeV between D and D^* mesons. Since the pion mass is already ∼140 MeV, only pions with a low momentum can be produced in this decay. This slow pion efficiency study is crucial, because the discrepancy between the slow pion tracking efficiency in data and MC is the source of a leading systematic uncertainty of $|V_{cb}|$ measurements based on $B \to D^*\ell\nu_l$ decays. Additionally this study can also be used to monitor charge dependent inefficiencies for the slow pion track finding in data and MC.

This chapter is structured as follows: In Sec.5.2 the reconstruction process of the $B^0 \to D^{*-}\pi^+$ decay will be discussed, followed by the application of various MC and data calibration weights in Sec.5.3 and Sec.5.4, respectively. This will be followed up by discussing the agreement between the data and MC samples in Sec.5.5 and the implemented fit procedure to extract the signal yields and its closure tests in Sec.5.6. The results of this calibration study will be discussed in Sec.5.7. Finally Sec.5.8 addresses the observed charge dependent inefficiencies.

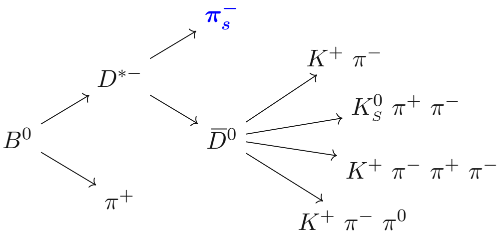


Figure 5.1: An illustration of the decay chain used to reconstruct B^0 candidates through $B^0 \to D^{*-}\pi^+$ with the subsequent decay $D^{*-} \to \bar{D}^0\pi_s^-$ and the four decay modes used for the $\bar{D}^0$ reconstruction.

5.2 Reconstruction of $B \to D^*\pi$

5.2.1 Event Reconstruction

We study the slow pion tracking efficiency at Belle II by reconstructing B^0 candidates with the untagged approach (see Sec.4.3.1) from the hadronic decay $B^0 \to D^{*-}\pi^+$, followed by the decay $D^{*-} \to \bar{D}^0\pi_s^-$[1], which includes the slow pion sought by this analysis. Due to the small mass difference of ~145 MeV between D^* and D mesons, this decay predominantly involves pions with a momentum below 300 MeV. The $\bar{D}^0$ mesons are reconstructed through the following four decay channels: $K^+\pi^-$, $K^+\pi^-\pi^+\pi^-$, $K_S^0\pi^+\pi^-$ and $K^+\pi^-\pi^0$. An illustration of the whole decay chain, including all four $\bar{D}^0$ decay modes, is given in Fig.5.1.

The K_S^0 of the $K_S^0\pi^+\pi^-$ channel is reconstructed by combining two charged pions, whose combined invariant mass must be within the range of $m \in [0.45, 0.55]$ GeV. Additionally the K_S^0 must successfully pass a vertex fit based on the `Treefitter` algorithm [55], which combines all particles in the decay chain to determine the best decay vertex position.

The π^0 of the $K^+\pi^-\pi^0$ decay channel is reconstructed by combining two photons. We based the selection of the photons on the recommendations provided by the Belle II performance group. The polar angle of the centroid position of the ECL cluster must lie within $0.2967 < \theta < 2.6180$ rad. Furthermore, photons must meet specific energy thresholds based on the ECL region, where the cluster is detected: in the forward ECL region ($12.4° < \theta < 31.4°$), the photon energy must exceed 0.08 GeV; in the barrel region ($32.2° < \theta < 128.7°$), the energy must be above 0.03 GeV; and in the backward ECL region ($130.7° < \theta < 155.1°$), an energy above 0.06 GeV is required. The π^0 must have an invariant mass within [0.120, 0.145] GeV.

Reliably reconstructed tracks of charged pions and kaons within the mid to high momentum region, excluding π_s, are selected by imposing constraints on the distance of the point of closest approach to the interaction point in the $r - \phi$ plane dr and along the z-axis dz: $dr < 2$ cm and $|dz| < 4$ cm. Additionally their polar angle is required to fall within the region of CDC acceptance, i.e. $17° < \theta < 150°$.

The reconstructed $\bar{D}^0$ mass is only allowed to deviate from the world average PDG value by 0.05 GeV. If this criterion is fulfilled, they are combined with a π_s^- to form a D^{*-} meson. The D^{*-} candidates are restricted to a momentum magnitude p^*[2] below 3 GeV to suppress the contribution from $q\bar{q}$ events and the mass difference between the D^{*-}

---

[1] The reconstruction of charge conjugated decays is also implied

[2] An asterisk denotes the center of mass frame.

and its $\overline{D}^0$ daughter must be below 0.158 GeV. Finally, we combine the D^{*-} with a π^+ to reconstruct the B^0 candidate. The B^0 meson must exhibit a beam constraint mass m_{bc} above 5.25 GeV and an energy deviation from the beam energy below 0.3 GeV.

5.2.2 MC Classification

After reconstructing the B^0 candidates with rather loose selection criteria we optimize the requirements by studying the figure of merit (see Sec.4.4.2). For this purpose the MC events are split into the following four categories:

- Signal: Correctly reconstructed candidates, based on a boolean flag from the truth-matching algorithm of basf2. The truthmatching compares the mothers and daughters of the whole reconstructed decay chain to the information provided by the generator.
- Continuum: Misreconstructed candidates originating from $q\overline{q}$ events.
- True D^*: Combinatorial background from B meson pairs, where the B is misreconstructed, however the reconstructed D^* candidate is a real D^*. This decision is based on the MC PDG code[3] provided by the generator.
- Fake D^*: Background events originating from B meson pairs without a correct B reconstruction and the D^* candidate is not true.

5.2.3 Signal Selection

To suppress the number of background events, we optimized the selection criteria by scanning the figure of merit (fom) defined in Sec.4.4.2. The change of the fom after each tightened selection criterion is shown in Fig.5.2 for each of the four $\overline{D}^0$ decay channels. Most cuts are optimized for the combination of all four channels, due to very minor deviations and for the sake of simplicity. However, there are some channel specific criteria: the binary kaonID without SVD information and the deviation of the reconstructed K_S^0 mass from its PDG value $|m_{K_S^0} - m_{K_S^0}^{PDG}|$[4]. The binary kaonID (see Sec.3.3) was only used for the $\overline{D}^0$ reconstruction via $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$, due to an lack of improvement of the fom for the other two channels $K^+\pi^-$ and $K_S^0\pi^+\pi^-$. We chose the binary instead of the global ID because it exhibited superior separation power between signal and background candidates.

The distributions of all discussed selections can be found in Fig.5.3. In these figures the threshold value of the cuts are indicated by a dashed red line and the colors of the rejected region are desaturated. For each distribution all cuts except the one for the shown variable are applied.

For K_S^0 candidates the absolute value of the mass difference $|m_{K_S^0}$ - $m_{K_S^0}^{PDG}|$ must be below 0.01 GeV. The contribution from continuum events is suppressed by requiring the ratio of the 2nd to the 0th Fox Wolfram moment R_2 to be below 0.5. While this cut could have been chosen tighter, it was deliberately kept rather loose to prevent deterioration of the signal efficiency, due to the main uncertainty of this analysis being statistical. Reconstructed $\overline{D}^0$ mesons are required to have a mass deviation from their world average PDG mass below 0.02 GeV. The mass difference Δm between the $\overline{D}^0$ and the D^{*-} candidates must fall

---

[3] Based on the MC numbering scheme of PDG. [10]

[4] PDG denotes, that the variable is the current world average provided by the particle data group

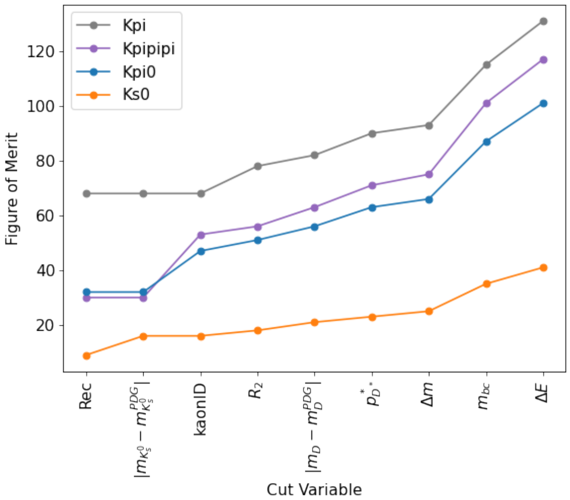


Figure 5.2: Cutflow of the fom for each decay channel used for the $\overline{D}^0$ reconstruction. Starting from the value of the fom directly after reconstruction up to the last cut. The x-axis shows the variable on which the cut was applied, while the y-axis shows the fom after its application.

within [0.141, 0.156] GeV. Furthermore the momentum in the center of mass frame of D^{*-} mesons is constrained to [2.0, 2.5] GeV to further suppress $q\overline{q}$ events. The kaons of the $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$ decay channels are required to surpass a binary PID likelihood ratio $\mathcal{L}_{K/\pi}$ of 0.1, with exclusion of the SVD information. For the B^0 candidates the energy difference $\Delta E = E_{B^0} - E_{beam}$ must be below 0.2 GeV and the beam constrained mass m_{bc} is constrained to the region of [5.270, 5.287] GeV.

5.3 MC Corrections

5.3.1 Gamma Efficiency

To account for observed discrepancies in the photon detection efficiency between data and MC, we remove photons from the MC samples. This is done by randomly dropping photons from the sample before initiating the reconstruction process. The dropout probability reflects the provided relative efficiency values, that take the photon energy dependence of the efficiency into consideration. These efficiency values are based on a study, that reconstructs $e^-e^+ \to \mu^-\mu^+(\gamma_{ISR})$, where γ_{ISR} is a high energy photon that gets radiated in the initial state e^-e^+. In this study, the photon is identified by predicting its momentum from ECL clusters based on the missing momentum of the muon system. [56]

5.3.2 HadronID

The binary kaon ID employed for the $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$ decay channel has been shown in Fig.5.3 to be a powerful tool for background suppression. However studies show, that there are discrepancies between data and MC when comparing the distributions of particle identification likelihoods. These likelihoods are among others based on momenta, energy clusters and polar angles, as discussed in the Sec.3.3. To correct for the differences in

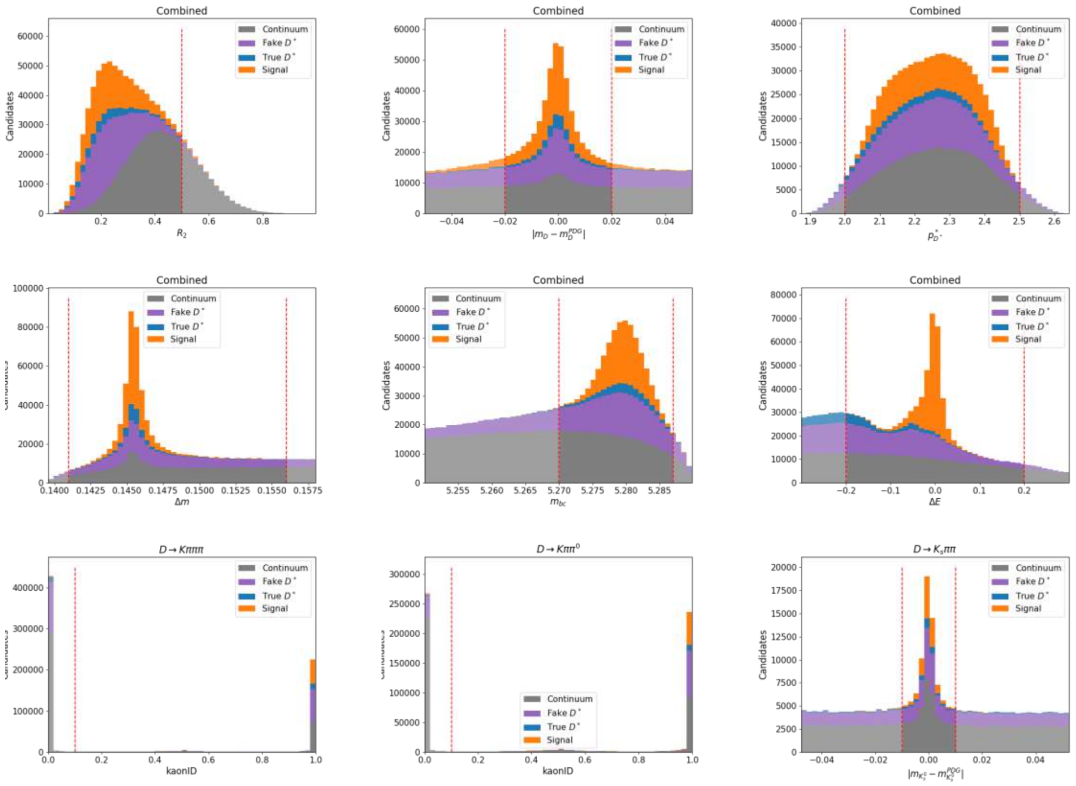
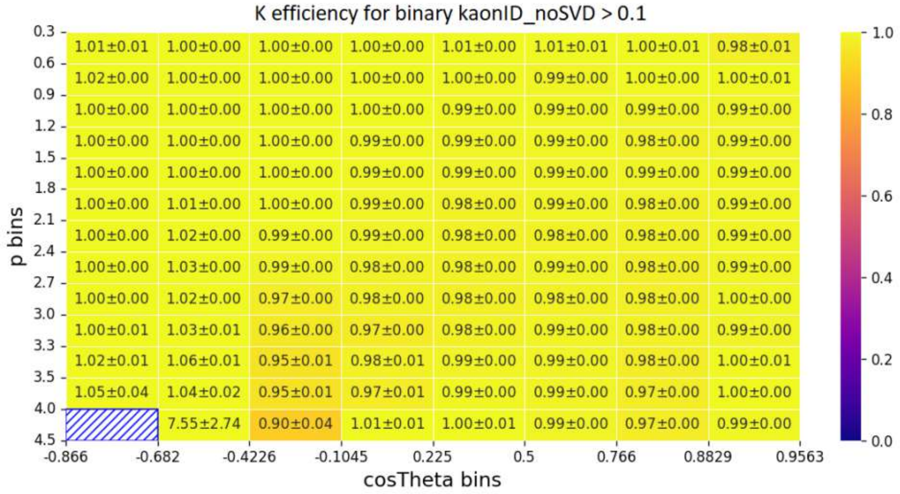


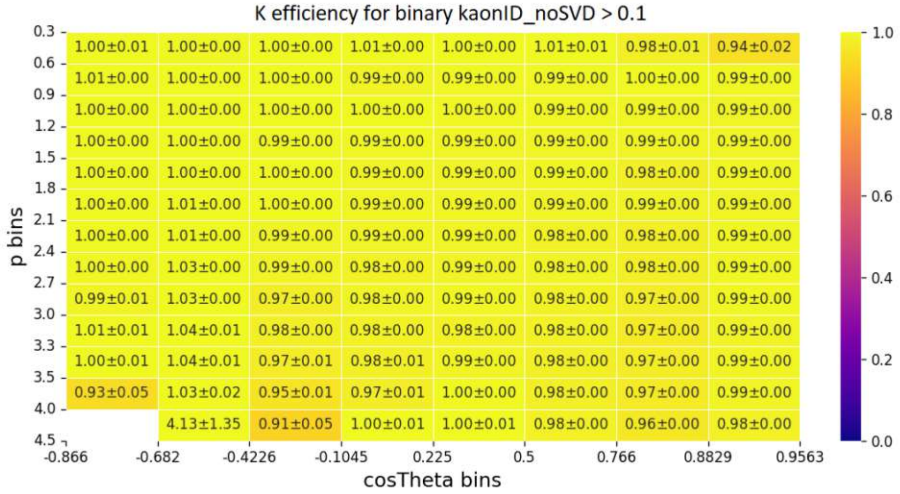
Figure 5.3: Here we show the distributions of all variables used for the tighter selections. In these figures all tight selection criteria with exclusion of the cut on the shown variable are applied. We indicate the rejected region by marking the selection threshold with a red dashed line and desaturating the rejected region. The variables all channels have in common are shown for the combined distributions in the first two rows, while the left and middle figure in last row show the kaonID selection for the $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$ channel, respectively. In the last row on the right we show the $|m_{K_S^0} - m_{K_S^0}^{PDG}|$ distribution of the $K_S^0\pi^+\pi^-$ channel.

the likelihood one compares the data/MC efficiency ratio of their ability to correctly assign an identification in bins of the momentum and polar angle for a selected PID threshold, i.e. kaonID > 0.1 in the case of our study. These correction ratios are determined by reconstructing a control channel, that involves the particle of the sought likelihood and where the tracks and polar angles can be reliably reconstructed. In case of the kaon ID the control channel would be $D^{*+} \to D^0[\to K^-\pi^+]\pi^+$. For the calculation of the calibration weights the performance group provides MC and data ntuples of said control channel and the exact same selections and PID cuts are applied on both samples. The ntuples are loaded and the weights calculated using the Belle II Systematic Framework [57]. The produced correction weights in bins of the kaon momentum and polar angle are provided in Fig.5.4 and Fig.5.5, for the efficiency and fake rate, respectively. While the efficiency compares the ability to correctly assign an identification, the fake rate compares the probabilities that a non-kaon is identified as a kaon. During the application of the weights with the `PIDVar`



1.01±0.01	1.00±0.00	1.00±0.00	1.00±0.00	1.01±0.00	1.01±0.01	1.00±0.01	0.98±0.01
1.02±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	0.99±0.00	1.00±0.00	1.00±0.01
1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00
1.00±0.00	1.00±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.98±0.00	0.99±0.00
1.00±0.00	1.00±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00
1.00±0.00	1.01±0.00	1.00±0.00	0.99±0.00	0.98±0.00	0.99±0.00	0.98±0.00	0.99±0.00
1.00±0.00	1.02±0.00	0.99±0.00	0.99±0.00	0.98±0.00	0.98±0.00	0.98±0.00	0.99±0.00
1.00±0.00	1.03±0.00	0.99±0.00	0.98±0.00	0.98±0.00	0.99±0.00	0.98±0.00	0.99±0.00
1.00±0.00	1.02±0.00	0.97±0.00	0.98±0.00	0.98±0.00	0.98±0.00	0.98±0.00	1.00±0.00
1.00±0.01	1.03±0.01	0.96±0.00	0.97±0.00	0.98±0.00	0.99±0.00	0.98±0.00	0.99±0.00
1.02±0.01	1.06±0.01	0.95±0.01	0.98±0.01	0.99±0.00	0.99±0.00	0.98±0.00	1.00±0.01
1.05±0.04	1.04±0.02	0.95±0.01	0.97±0.01	0.99±0.00	0.99±0.00	0.97±0.00	1.00±0.00
▨	7.55±2.74	0.90±0.04	1.01±0.01	1.00±0.01	0.99±0.00	0.97±0.00	0.99±0.00

(a) Correction tables for K candidates with a negative charge.



1.00±0.01	1.00±0.00	1.00±0.00	1.01±0.00	1.00±0.00	1.01±0.01	0.98±0.01	0.94±0.02
1.01±0.00	1.00±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00	1.00±0.00	0.99±0.00
1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00
1.00±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00
1.00±0.00	1.00±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.98±0.00	0.99±0.00
1.00±0.00	1.01±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00
1.00±0.00	1.01±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.98±0.00	0.98±0.00	0.99±0.00
1.00±0.00	1.03±0.00	0.99±0.00	0.98±0.00	0.99±0.00	0.98±0.00	0.98±0.00	0.99±0.00
0.99±0.01	1.03±0.00	0.97±0.00	0.98±0.00	0.99±0.00	0.98±0.00	0.97±0.00	0.99±0.00
1.01±0.01	1.04±0.01	0.98±0.00	0.98±0.00	0.98±0.00	0.98±0.00	0.97±0.00	0.99±0.00
1.00±0.01	1.04±0.01	0.97±0.01	0.98±0.01	0.99±0.00	0.98±0.00	0.97±0.00	0.99±0.00
0.93±0.05	1.03±0.02	0.95±0.01	0.97±0.01	1.00±0.00	0.98±0.00	0.97±0.00	0.99±0.00
	4.13±1.35	0.91±0.05	1.00±0.01	1.00±0.01	0.98±0.00	0.96±0.00	0.98±0.00

(b) Correction tables for K candidates with a positive charge.

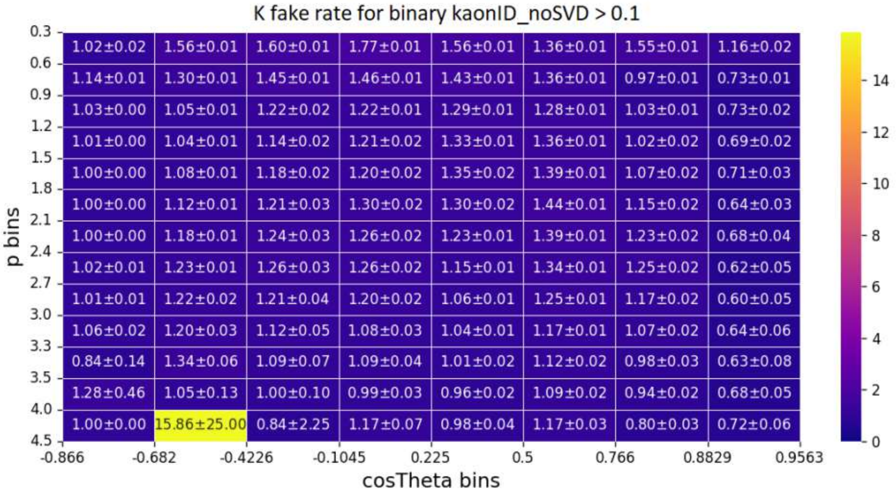
Figure 5.4: Efficiency weight tables for the binary kaonID, that excludes SVD information for a kaonID threshold > 0.1. The figures are binned in two dimensions, namely p (y-axis) and $cos\theta$ (x-axis). In Fig.5.4a we show the table for the candidates with a negative charge and Fig.5.4b shows positively charged candidates.

framework [58] we also check the coverage in bins of the momentum and polar angle of the kaon for both the efficiency and fake rate. The coverage shown for the $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$ in Fig.5.6 and Fig.5.7, respectively indicates a full coverage of our sample. [57]

5.4 Data Corrections

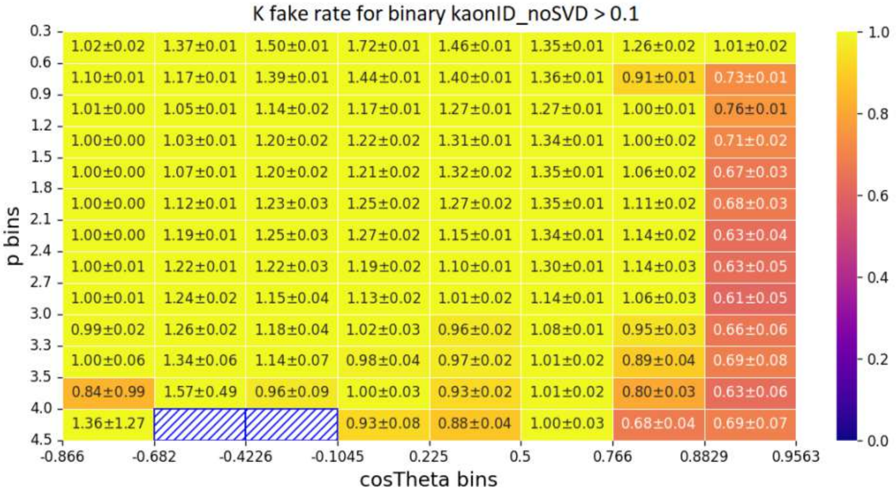
5.4.1 Track Momentum Scaling

During studies of $D^{+,0}$, J/ψ, K_S and Λ_c^+ mesons a peak shift of approximately -0.015 MeV was observed in the reconstructed invariant mass in comparison to their world average



1.02±0.02	1.56±0.01	1.60±0.01	1.77±0.01	1.56±0.01	1.36±0.01	1.55±0.01	1.16±0.02
1.14±0.01	1.30±0.01	1.45±0.01	1.46±0.01	1.43±0.01	1.36±0.01	0.97±0.01	0.73±0.01
1.03±0.00	1.05±0.01	1.22±0.02	1.22±0.01	1.29±0.01	1.28±0.01	1.03±0.01	0.73±0.02
1.01±0.00	1.04±0.01	1.14±0.02	1.21±0.02	1.33±0.01	1.36±0.01	1.02±0.02	0.69±0.02
1.00±0.00	1.08±0.01	1.18±0.02	1.20±0.02	1.35±0.02	1.39±0.01	1.07±0.02	0.71±0.03
1.00±0.00	1.12±0.01	1.21±0.03	1.30±0.02	1.30±0.02	1.44±0.01	1.15±0.02	0.64±0.03
1.00±0.00	1.18±0.01	1.24±0.03	1.26±0.02	1.23±0.01	1.39±0.01	1.23±0.02	0.68±0.04
1.02±0.01	1.23±0.01	1.26±0.03	1.26±0.02	1.15±0.01	1.34±0.01	1.25±0.02	0.62±0.05
1.01±0.01	1.22±0.02	1.21±0.04	1.20±0.02	1.06±0.01	1.25±0.01	1.17±0.02	0.60±0.05
1.06±0.02	1.20±0.03	1.12±0.05	1.08±0.03	1.04±0.01	1.17±0.01	1.07±0.02	0.64±0.06
0.84±0.14	1.34±0.06	1.09±0.07	1.09±0.04	1.01±0.02	1.12±0.02	0.98±0.03	0.63±0.08
1.28±0.46	1.05±0.13	1.00±0.10	0.99±0.03	0.96±0.02	1.09±0.02	0.94±0.02	0.68±0.05
1.00±0.00	15.86±25.00	0.84±2.25	1.17±0.07	0.98±0.04	1.17±0.03	0.80±0.03	0.72±0.06

(a) Correction tables for K candidates with a negative charge.



1.02±0.02	1.37±0.01	1.50±0.01	1.72±0.01	1.46±0.01	1.35±0.01	1.26±0.02	1.01±0.02
1.10±0.01	1.17±0.01	1.39±0.01	1.44±0.01	1.40±0.01	1.36±0.01	0.91±0.01	0.73±0.01
1.01±0.00	1.05±0.01	1.14±0.02	1.17±0.01	1.27±0.01	1.27±0.01	1.00±0.01	0.76±0.01
1.00±0.00	1.03±0.01	1.20±0.02	1.22±0.02	1.31±0.01	1.34±0.01	1.00±0.02	0.71±0.02
1.00±0.00	1.07±0.01	1.20±0.02	1.21±0.02	1.32±0.02	1.35±0.01	1.06±0.02	0.67±0.03
1.00±0.00	1.12±0.01	1.23±0.03	1.25±0.02	1.27±0.02	1.35±0.01	1.11±0.02	0.68±0.03
1.00±0.00	1.19±0.01	1.25±0.03	1.27±0.02	1.15±0.01	1.34±0.01	1.14±0.02	0.63±0.04
1.00±0.01	1.22±0.01	1.22±0.03	1.19±0.02	1.10±0.01	1.30±0.01	1.14±0.03	0.63±0.05
1.00±0.01	1.24±0.02	1.15±0.04	1.13±0.02	1.01±0.02	1.14±0.01	1.06±0.03	0.61±0.05
0.99±0.02	1.26±0.02	1.18±0.04	1.02±0.03	0.96±0.02	1.08±0.01	0.95±0.03	0.66±0.06
1.00±0.06	1.34±0.06	1.14±0.07	0.98±0.04	0.97±0.02	1.01±0.02	0.89±0.04	0.69±0.08
0.84±0.99	1.57±0.49	0.96±0.09	1.00±0.03	0.93±0.02	1.01±0.02	0.80±0.03	0.63±0.06
1.36±1.27	▨	▨	0.93±0.08	0.88±0.04	1.00±0.03	0.68±0.04	0.69±0.07

(b) Correction tables for K candidates with a positive charge.

Figure 5.5: Fake rate correction tables for the binary kaonID, that excludes SVD information for a kaonID threshold > 0.1. The figures are binned in two dimensions, namely p (y-axis) and $cos\theta$ (x-axis). In Fig.5.5a we show the table for the candidates with a negative charge and Fig.5.5b shows positively charged candidates.

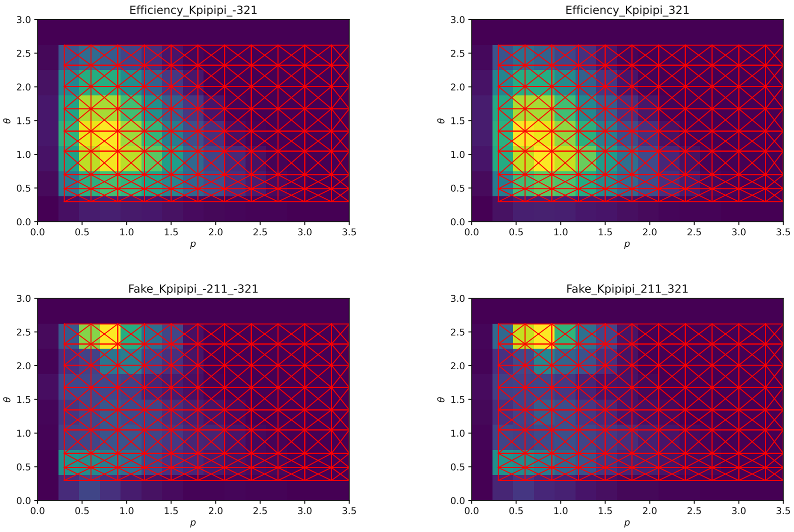


Figure 5.6: Charge separated coverage of the efficiency weight tables (upper) and fake rate tables (lower) for the $K^+\pi^-\pi^+\pi^-$ channel. The figures are binned in two dimensions, namely in p (x-axis) and θ (y-axis). The coverage of candidates with negative charge are shown on the left and with a positive charge on the right.

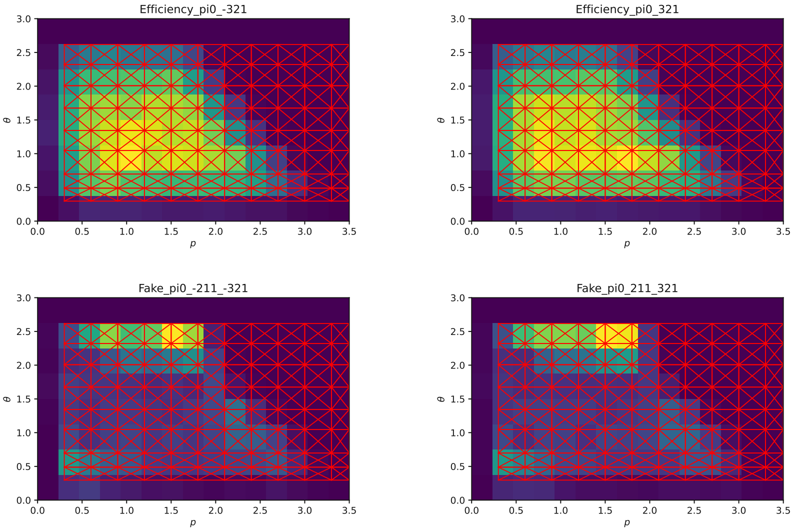


Figure 5.7: Charge separated coverage of the efficiency weight tables (upper) and fake rate tables (lower) for the $K^+\pi^-\pi^0$ channel. The figures are binned in two dimensions, namely in p (x-axis) and θ (y-axis). The coverage of candidates with negative charge are shown on the left and with a positive charge on the right.

PDG value. This shift originates from imperfections of the magnetic field map that was applied during the reconstruction process of the final state particles in the data. To counteract this discrepancy a global scaling factor is applied to all charged tracks before initiating the combination process of the particles. The scaling is determined by studying $D^{*+} \to D^0[\to K^-\pi^+]\pi^+$ and applying a fit on the invariant mass of the D meson. From the deviation of the fit result for the reconstructed invariant mass from its PDG value one can then derive a scaling factor.

5.4.2 Photon Energy Bias

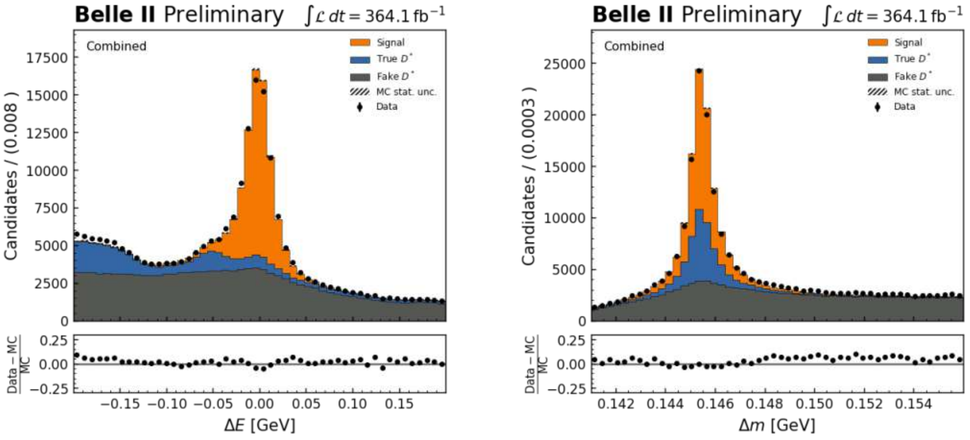
This calibration originates from observed energy shifts in the ΔE peak when comparing data to MC for decays that involve neutral final state particles. For example for the decay $B^+ \to K^+\pi^0$ a -21 ± 12 MeV shift of the data with respect to the simulation was determined. The correction for the observed photon energy bias in data was derived by reconstructing $B^+ \to \bar{D}^0[\to K^+\pi^-\pi^0]\pi^+$. We therefore correct the energy of photons by the calibration factor provided by performance group before combining particles.

5.5 Pre-Fit Data-MC Agreement

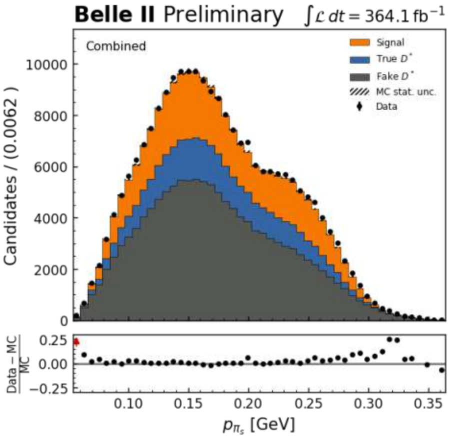
After the calibration of the data sample and the MC, we study the agreement between them before creating a model for the signal extraction to determine whether the calibration was sufficient. With respect to the classification of the MC components a minor change was made for this comparison, due to the chosen variables that will be used during the signal extraction: the energy deviation ΔE and the mass difference between the D^{*-} and $\bar{D}^0$ meson Δm. To achieve a better separation power for Δm, which is heavily dependent on the reconstructed D^{*-} being true, we drop the separate distinction of the $q\bar{q}$ events and add them to either the true or fake D^{*-} component in accordance with the MC PDG code of the D^{*-} candidate.

In Fig.5.8 and Fig.5.9 the distributions of ΔE, Δm and the slow pion momentum in the laboratory frame p_{π_s} are shown for the combination of all four decay channels and individual channels, respectively. We show specifically the distributions of these three variables because they will be used during the signal extraction: a two dimensional fit in ΔE and Δm will be performed for the signal extraction for four different p_{π_s} momentum regions. In these figures a comparison between the distribution of the data and the luminosity scaled (364/1443) MC samples are made. It has to be noted, that these plots only include the statistical uncertainties of the bin counts. By comparing the MC to the data we are able to determine how well the MC describes our data and if any additional corrections or nuisance parameters will be necessary during the signal extraction.

When examining the agreement for the combination of all channels, the data is generally well modeled by the MC. The combined distributions will be used during the fitting procedure to extract the signal. In Fig.5.8 we observe very good agreement for p_{π_s} between the data and MC across the whole momentum range. For the ΔE spectrum, the signal region, i.e. $\Delta E \in [-0.05, 0.05]$ GeV, is very well modeled, as are the lower and upper sideband regions of [-0.20,-0.05] GeV and [0.05, 0.20] GeV, respectively. However, there is some minor disagreement in the lower end of the lower sideband region of $\Delta E \in [-0.20, -0.15]$ GeV. In the prefit comparison of the Δm distribution similar properties are observed, with a very well modeled distribution in the signal region of $\Delta m \in [0.141, 0.148]$ GeV and a slight



(a) Combined ΔE (b) Combined Δm



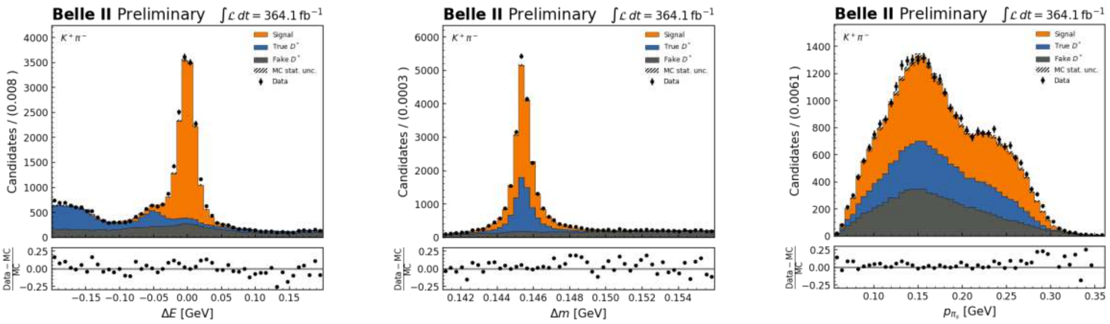
(c) Combined p_{π_s}

Figure 5.8: Shape comparison of pre-fit distribution of all variables relevant to the signal extraction for the combination of all decay channel. The signal, true D^{*-} and fake D^{*-} components of the MC are depicted by the orange, blue and gray components, respectively. Additionally the dots indicate the data. These figures only include statistical uncertainties.

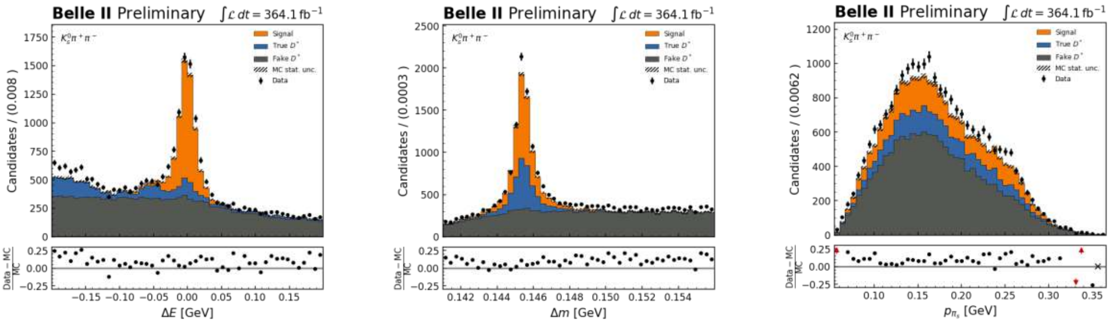
undershoot of the MC in the upper sideband region above 0.148 GeV.

The entire range of the distributions for ΔE, Δm and p_{π_s} of both $K^+\pi^-$ and $K^+\pi^-\pi^0$ show a very good agreement between data and MC in Fig.5.9. In general the slow pion momentum is well modeled for all individual channels. The only exception is the $K^0_S\pi^+\pi^-$ channel, where the expectations from the MC underestimate the data. However, the $K^0_S\pi^+\pi^-$ channel has by far the lowest event count and therefore rather high statistical uncertainties. This tendency of the slight underestimation of events by the MC is also observed for the ΔE and Δm distributions. The $K^+\pi^-\pi^+\pi^-$ channel shows good agreement in the signal region of ΔE and Δm, but an undershoot of the MC in the sideband regions causing minor discrepancies.

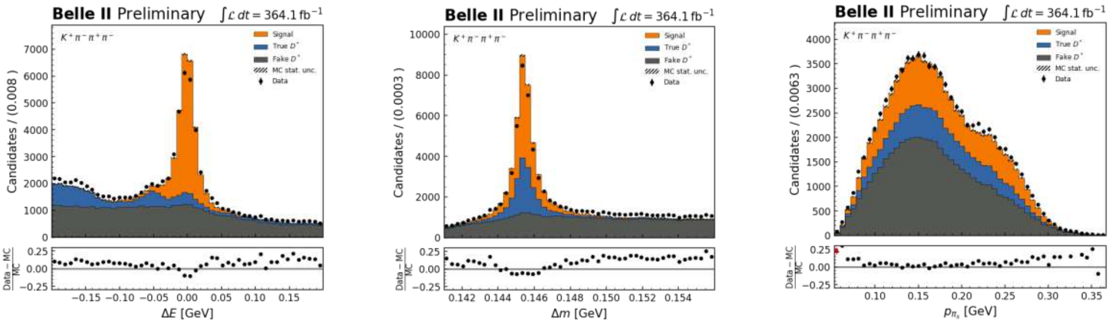
In addition to the agreement between data and MC, we investigated the signal yield, background yield and significance of each individual decay channel to quantitatively estimate their contributions to the overall combined distribution used for the signal extraction. For



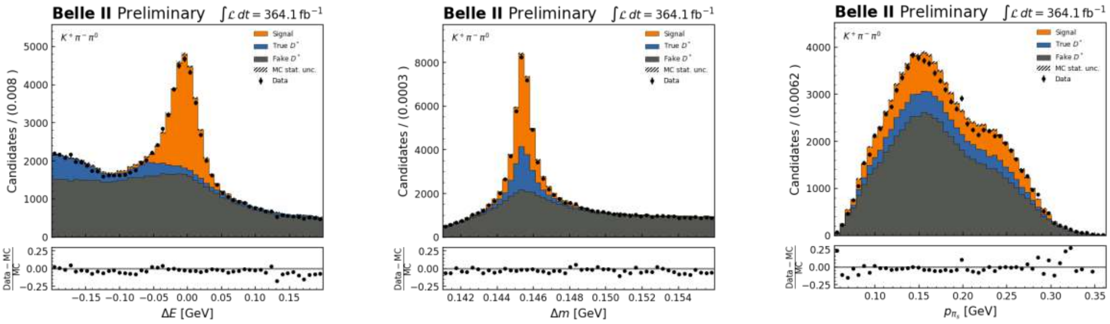
(a) $K^+\pi^-$: ΔE (b) $K^+\pi^-$: Δm (c) $K^+\pi^-$: p_{π_s}



(d) $K_S^0\pi^+\pi^-$: ΔE (e) $K_S^0\pi^+\pi^-$: Δm (f) $K_S^0\pi^+\pi^-$: p_{π_s}



(g) $K^+\pi^-\pi^+\pi^-$: ΔE (h) $K^+\pi^-\pi^+\pi^-$: Δm (i) $K^+\pi^-\pi^+\pi^-$: p_{π_s}



(j) $K^+\pi^-\pi^0$: ΔE (k) $K^+\pi^-\pi^0$: Δm (l) $K^+\pi^-\pi^0$: p_{π_s}

Figure 5.9: Shape comparison of pre-fit distribution of all variables relevant to the signal extraction for each individual channel. The signal, true D^* and fake D^* components of the MC are depicted by the orange, blue and gray components, respectively. Furthermore, the dots indicate the data and these illustrations only include statistical uncertainties.

	$K^+\pi^-$	$K^+\pi^-\pi^+\pi^-$	$K^0_S\pi^+\pi^-$	$K^+\pi^-\pi^0$
N_{sig}	13135 ± 57	20836 ± 72	4000 ± 31	17529 ± 66
N_{bg}	4374 ± 33	17337 ± 65	4948 ± 35	20970 ± 72
$\frac{N_{sig}}{\sqrt{N_{sig}+N_{bg}}}$	99.3 ± 0.4	106.6 ± 0.4	42.3 ± 0.3	89.3 ± 0.3

Table 5.1: Summary of the luminosity scaled and corrected signal yield, background yield and significance of each reconstructed decay channel with an additional restriction to the signal region of $|\Delta E| < 0.05$ GeV.

	Combined	$K^0_S\pi^+\pi^-$	$K^+\pi^-$	$K^+\pi^-\pi^+\pi^-$	$K^+\pi^-\pi^0$
Candidates	875107	83802	116621	353350	321334
Unique Candidates	870348	83801	116619	353338	321325
Unique Candidates in %	99.5%	99.9%	99.9%	99.9%	99.9%

Table 5.2: Summary of the candidates multiplicity without luminosity scaling applied to the number of candidates.

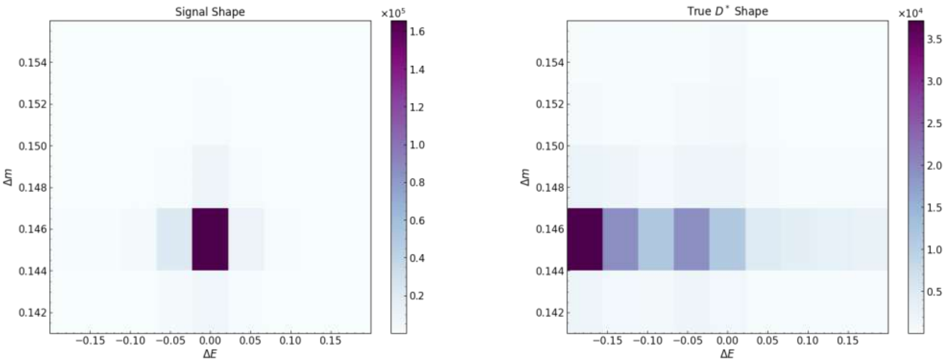
this check we compared the luminosity scaled and corrected number of events and restricted the region of ΔE to the signal region of [-0.05, 0.05]. The results summarized in Tab.5.1 show, that by far the $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$ channel contribute the highest number of signal candidates, closely followed by the $K^+\pi^-$ mode. However, it also shows that $K^+\pi^-$ and $K^+\pi^-\pi^+\pi^-$ have the best fom. While the $K^0_S\pi^+\pi^-$ has the least amount of signal candidates and the worst significance, it still contributes $\sim 7\%$ of the overall signal candidates and is therefore not negligible.

The results of study the candidate multiplicity of each individual channel and the combination is provided in Tab.5.2. This examination showed that due to $\sim 99.9\%$ unique candidates for individual channels and an candidate overlap of only $\sim 0.5\%$ there is no need for a dedicated best candidate selection.

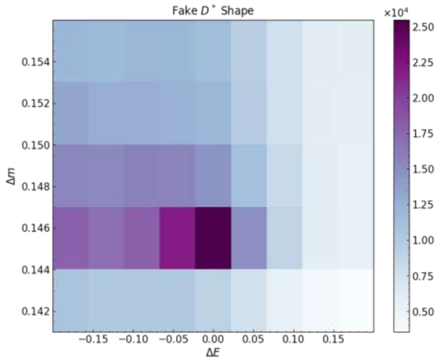
5.6 Signal Extraction

To determine the relative tracking efficiency of the slow pions, it is necessary to compare the amount of signal candidates in data and MC. We extract the signal candidates from the data sample using a two dimensional template fit (see Sec.4.8). This involves building a likelihood model based on the three event categories, as classified in Sec.5.5: Signal, true D^* and fake D^*. Each template corresponds to the distribution of events for a given event category. The background classes were chosen to enhance the separation power between each component. The fake D^* component will be spread out over the whole Δm range, while the true D^* component will only populate the signal region of $\Delta m \in [0.144, 0.148]$ GeV. In Fig.5.10 we present the two dimensional templates of all three event classes to highlight their distinct shapes.

The slow pion efficiency will be independently determined in four different p_{π_s} regions via a two dimensional model with 9 even bins in $\Delta E \in [-0.2, 0.2]$ GeV and 5 even bins in $\Delta m \in [0.141, 0.156]$ GeV. We split the samples in the following p_{π_s} regions: [0.05, 0.12] GeV, [0.12, 0.16] GeV, [0.16, 0.20] GeV and [0.20, 0.35] GeV. The python package `pyhf` [53] was used to create the model and implement this binned maximum likelihood fit.



(a) Signal Shape
(b) True D^* shape



(c) Fake D^* shape

Figure 5.10: The two dimensional templates in ΔE (x-axis) and ΔM (y-axis) for the three MC components. In Fig.5.10a and Fig.5.10b the distributions of the Signal and true D^* component are shown, respectively. Finally, Fig.5.10c shows the distribution of the fake D^* component. The shown spectra include the whole p_{π_s} range.

In our binned maximum likelihood model, both the statistical MC uncertainty and the systematic uncertainty arising from the uncertainties of the kaonID correction weights applied on the candidates of the $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$ channels were implemented as nuisance parameters in the form of shape variations. To account for mismodeling due to poorly known branching fractions for hadronic decay like $D^*\rho$, another nuisance parameter was added to reflect this incomplete knowledge as a systematic uncertainty. For that purpose we implement the nuisance parameter as an uncorrelated shape variation of up to 5% to the shape of ΔE for true D^{*-} background component. Additionally a correlated shape nuisance parameter of up to 10% was introduced to address a p_{π_s} momentum dependent slight peak of the fake D^{*-} component in signal region of $\Delta m \in [0.144, 0.147]$ GeV, as shown in Fig.5.11. This variation takes potential mismodeling of the ratio of the flat to the peaking component of the fake D^{*-} template into consideration as a systematic uncertainty. The impact of the 10% shape variation on the fake D^{*-} shape in the Δm signal region is illustrated in Fig.5.12

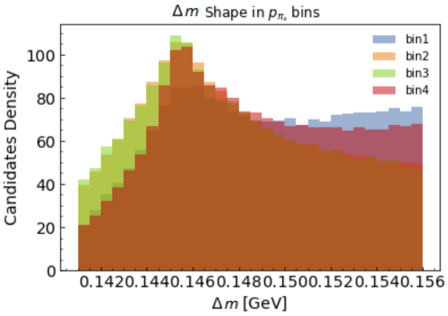


Figure 5.11: Slow pion momentum dependency of the shape of the fake D^* background. Due to the exhibited dependency on p_{π_s}, it was chosen to account for the variation of the slight peak in the Δm signal region via a shape nuisance parameter.

Figure 5.12: Impact of the correlated 10% up and down variations on the Δm shapes of the fake D^* component in the individual kinematic bins of p_{π_s}.

5.6.1 Closure Tests

To ensure a reliable result for the signal extraction two closure tests were made for the fitting procedure, namely an asimov test and a pull study.

An asimov test is done to evaluate if maximum likelihood is working as intended. This is done by using the combination of all MC templates without any classification as a pseudo data sample. We do this test for the exact setup as it will be applied on the data, i.e. a two dimensional binned maximum likelihood fit in 9 bins of ΔE and 5 bins in Δm, with the sample being split into four p_{π_s} momentum region. If the fit is working as intended, the resulting normalization factors of all MC templates should be equal to 1. Therefore estimating the amount of signal events in the pseudo data being equivalent to the signal yield in the simulation templates. The result of the asimov test for each momentum bin is provided in Fig.5.13a and shows perfect agreement between the predicted signal yield and the signal yield in the MC.

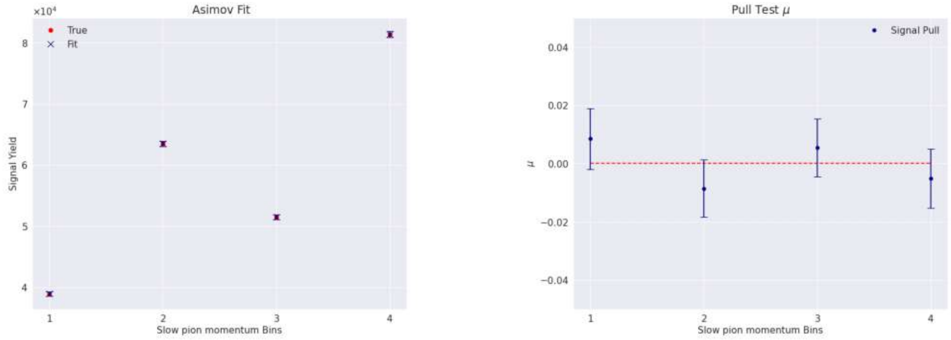
The second closure test was a pull test based on bootstrapping to ensure, that our fit is stable. In Sec.4.7.4 an overview about bootstrapping and pull tests is provided. For the pull test we calculate the pull of 10000 Poisson resamples of the original pseudo data used for the asimov test and also Poisson variations of the MC template. The distributions of the pull are then fitted with an unbinned Gaussian. In Fig.5.13b the mean of the Gaussian for each momentum bin is shown, where it can be seen that the mean of each pull is centered towards zero. Therefore no biases are expected for the resulting signal yields. Additionally no significant over or underestimation of the error on the fit can be observed from the width of the Gaussians shown in Fig.5.13c, with all widths being within a 1 to 1.5 σ agreement with 1 within their respective uncertainty.

5.7 Results

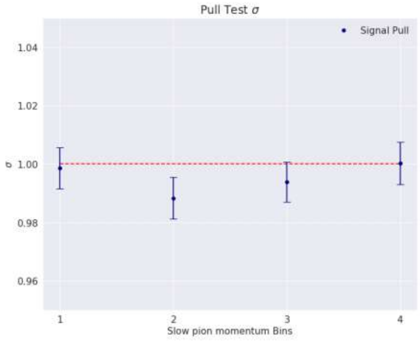
To determine the relative tracking efficiency calibration factor between data and MC we extract the signal yields based on the fitting procedure described in Sec.5.6. During the signal extraction we perform independent, two dimensional fits on ΔE and Δm for four different p_{π_s} momentum regions: [0.05, 0.12] GeV, [0.12, 0.16] GeV, [0.16, 0.20] GeV and [0.20, 0.35] GeV. In Fig.5.14 and Fig.5.15 we display the data and MC agreement before performing the binned maximum likelihood fit in 9 bins of $\Delta E \in [-0.2, 0.2]$ GeV and 5 bins of $\Delta m \in [0.141, 0.156]$ GeV, respectively.

For the ΔE distributions the data and MC comparisons exhibit the same features for three out of four momentum regions: overall good agreement with some minor disagreement in the very first ΔE bin below -0.15 GeV. The slow pion momentum bin of $[0.16, 0.20]$ GeV shows a very good agreement for the whole ΔE range. For Δm similar features can be observed, the $p_{\pi_s} \in [0.16, 0.20]$ GeV bin shows no diasgreement, while the other three bins show some minor disagreement for sideband region of $\Delta m > 0.146$ GeV.

The agreement between data and MC after scaling the MC templates with the resulting normalization factors from the fit is shown for the ΔE and Δm projection in Fig.5.16 and Fig.5.17, respectively for all four p_{π_s} bins. These figures demonstrate that the agreement for ΔE has significantly improved in comparison to the pre-fit distribution, resolving the minor disagreements in the lowest ΔE bin. For the Δm distribution we can also observe some improvements: the disagreement in the sideband region is corrected. However, a minor disagreement in the signal region remains between data and MC for three out of four



(a) Asimov (b) Bias



(c) Error Estimation

Figure 5.13: The closure test results for the two dimensional fit in ΔE and ΔM using the three MC templates: Signal, true D^* and fake D^*. In Fig.5.13a we show the resulting yields of the asimov test for each p_{π_s} bin, while Fig.5.13b and Fig.5.13c show the μ and σ pull results, respectively.

momentum region remains, excluding $p_{\pi_s} \in [0.16, 0.20]$ GeV.

To calculate the relative tracking efficiency we need to compare the signal yield N_{fit} of each kinematic slow pion bin in data with the MC prediction of the signal in the corresponding p_{π_s} bin. The MC prediction is defined as the luminosity scaled (364/1443) number of signal events in the signal template N_{sig}^{MC}. This yields the following formula for the relative tracking efficiency:

$$r_i = \frac{N_{fit}}{MC\ Prediction},\ MC\ Prediction = N_{sig}^{MC} \cdot \frac{luminosity\ data}{luminosity\ MC} \tag{5.1}$$

where r_i is the relative yield for a given momentum bin i. This relative yield is used as a measure of the discrepancy between the modeled tracking efficiency and the reconstructed tracking efficiency from data for the slow pions, with a ratio of 1 meaning perfect agreement.

These relative yields are not directly used as the calibration factors, because deviations in the ratios can be caused by factors other than track inefficiencies, such as mismodeling of branching fractions. Taking these branching fraction uncertainties into account for each decay would lead to significant systematic uncertainties for the yield ratio; e.g. the uncertainty of the $K^+\pi^-\pi^0$ alone would result in an additional systematic uncertainty of

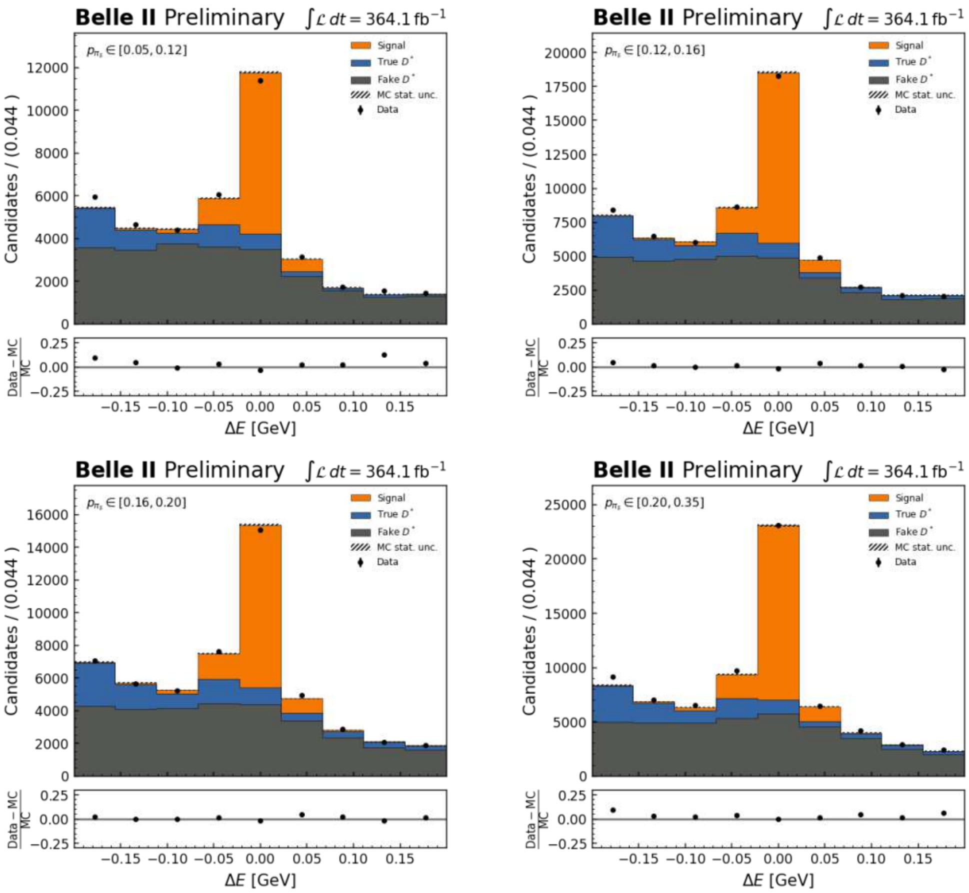


Figure 5.14: Pre-fit agreement for the distribution of ΔE in the four momentum bins of p_{π_s}. In these figures, the distributions of all decay channels are combined.

$\sim 5\%$. To cancel these global systematic uncertainties, which would affect all p_{π_s} bins equally, we normalize the individual yield ratios r_i to the yield of the highest kinematic bin, i.e. $\rho_i = r_i/r_4$. By normalizing to the highest momentum bin we assume that its tracking efficiency behaves like the tracking for the mid to high momentum regions and is therefore affected by their tracking efficiency systematic provided by the performance group. This approach cancels out all systematic uncertainties, except for the previously mentioned momentum-dependent shape variations from the fitting procedure.

The yield of the fourth bin to which we normalize is also extracted using a binned maximum likelihood fit with its implemented nuisance parameters. Therefore we need to propagate the uncertainty of the highest momentum bin to all three lower momentum bins as a correlated uncertainty. Additionally each lower momentum also has its own uncorrelated uncertainty from the individual fits. Through linear error propagation the following two uncertainties can be obtained for the results of each p_{π_s} bin:

$$\sigma_{\rho_i}^2 = \left(\frac{1}{r_4}\right)^2 \sigma_{r_i}^2 + \left(\frac{r_i}{r_4^2}\right)^2 \sigma_{r_4}^2 = \sigma_{uncorr_i}^2 + \sigma_{corr_i}^2, \forall i \in [1, 2, 3], \tag{5.2}$$

where σ_{r_i} is the statistical uncertainty of the fit for a given bin $i = 1, .., 3$. On top of the uncertainties of the signal extraction, the relative uncertainty of the mid to high momentum

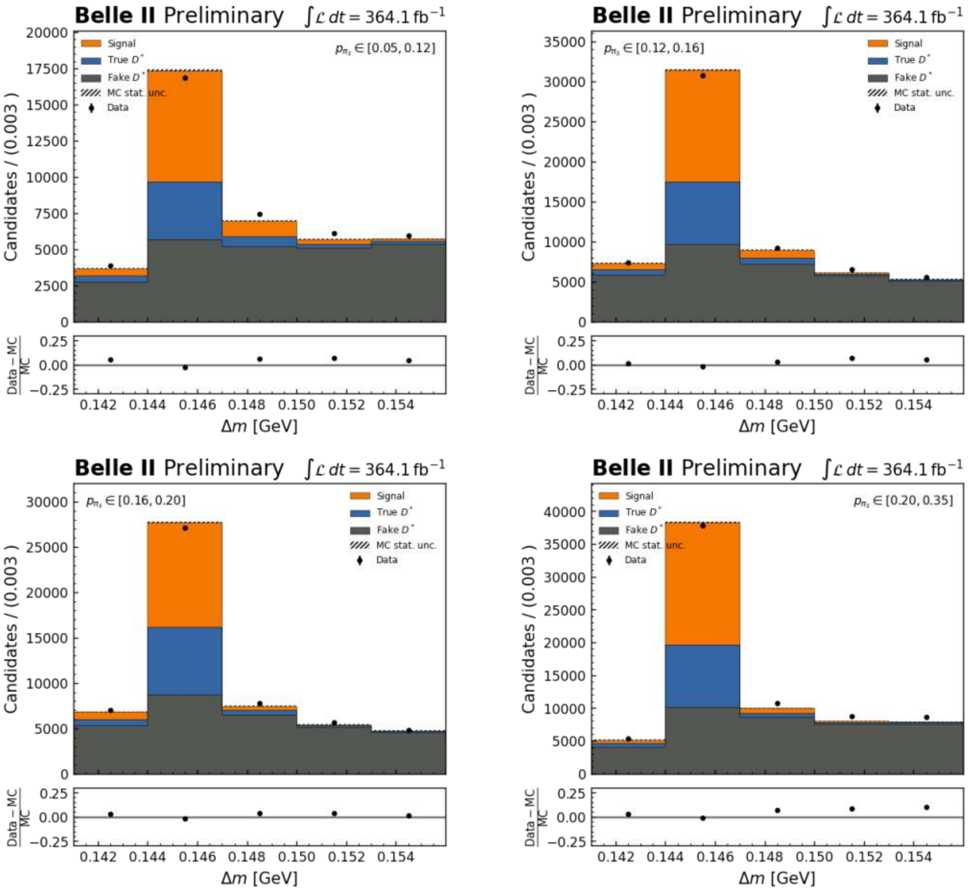
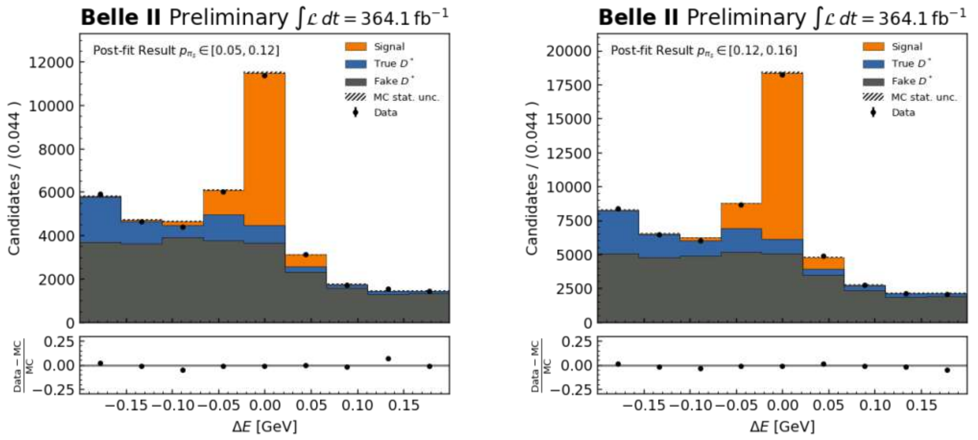


Figure 5.15: Pre-fit agreement for the distribution of Δm in the four momentum bins of p_{π_s}. In these figures, the distributions of all decay channels are combined.

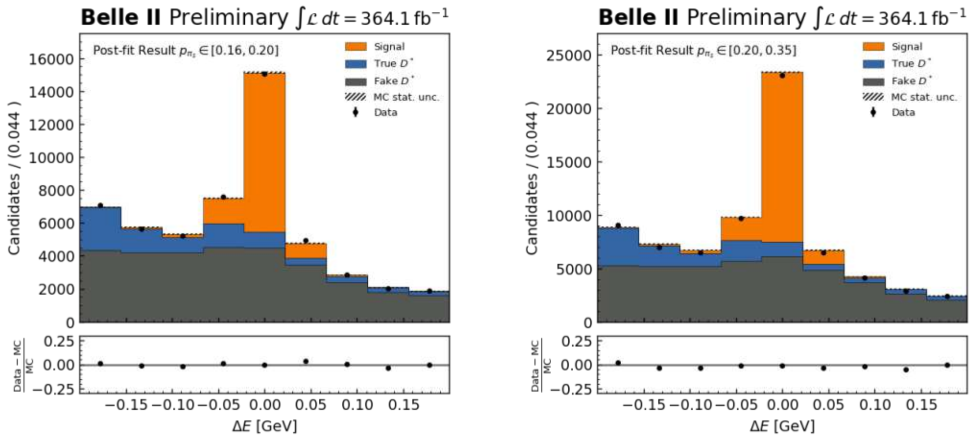
tracking $\sigma_{track} = 0.27\%$ is added in quadrature.

A summary of the yield ratios, normalized ratio and uncertainties for all slow pion momentum bins can be found in Tab.5.3. Additionally we visually compare the normalized ratios ρ_i in Fig.5.18. From Fig.5.18, we can observe that the tracking efficiency discrepancy for the second and third low momentum bins of the pions are rather low and agree with unity within their uncertainties, hinting towards a well modeled momentum region. For the lowest momentum bin however, i.e. $p_{\pi_s} \in [0.05, 0.12]$ GeV, a discrepancy of $\sim 5\%$ between data and MC can be observed. By comparing the discrepancies between the first bin and the other bins, one can assume, that even slow pions, which only barely reach the CDC are better modelled, than pions in the lowest momentum which almost completely have to rely on reconstruction via the SVD due to helix movement induced by the magnetic field in the innermost detector region. To get a measure of how the relative slow pion tracking efficiency discrepancy between data and MC will affect studies, that involve a slow pion in their decay chain we calculate an uncertainty weighted average based on the results provided in Tab.5.3:

$$(97.441 \pm 1.199_{stat} \pm 0.817_{sys})\%, \ \sigma_{\pi_s} = 1.46\% \tag{5.3}$$



(a) Asimov (b) Mu



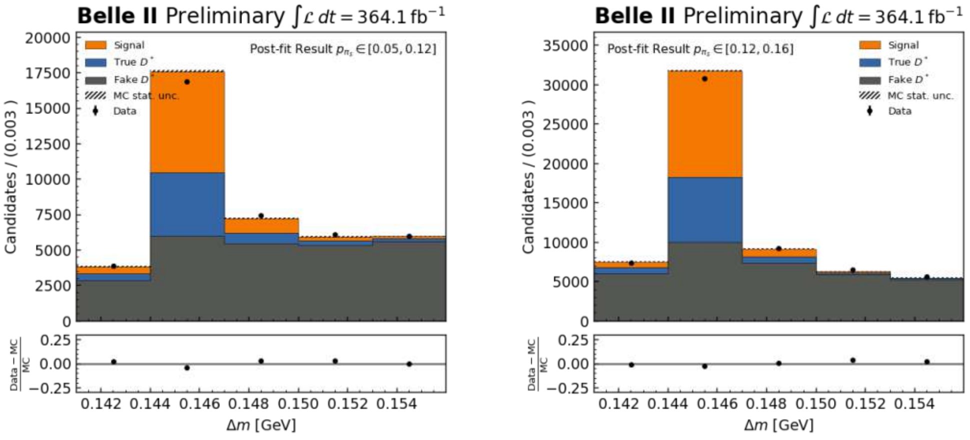
(c) Sigma (d) Sigma

Figure 5.16: Post-fit agreement for the distribution of ΔE in the four momentum bins of p_{π_s}. In these figures, the distributions of all decay channels are combined.

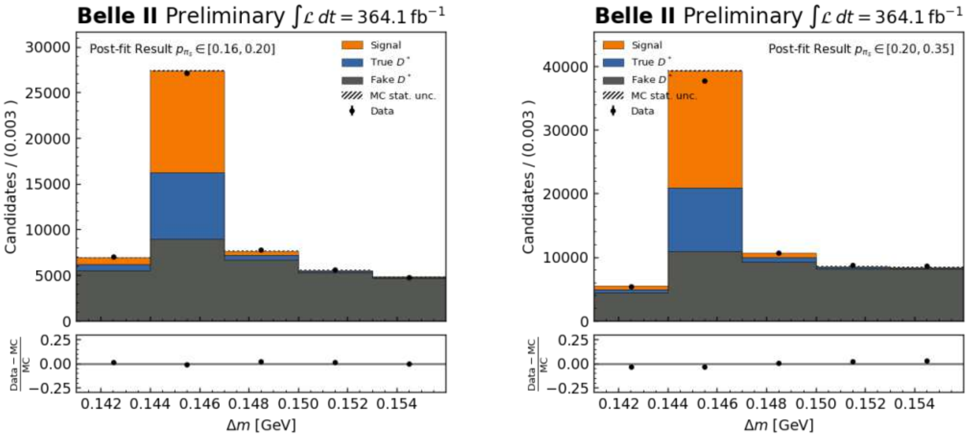
Calibration results for π_s

$\mathbf{p}_{\pi_s}$ **range[GeV]**	**0.05-0.12**	**0.12-0.16**	**0.16-0.20**	**0.20-0.35**
MC prediction	9733 ± 49	15832 ± 63	12863 ± 57	20259 ± 71
Fitted yield	9066 ± 144	15336 ± 188	12432 ± 176	19920 ± 220
r_i	0.932 ± 0.016	0.969 ± 0.013	0.967 ± 0.014	0.9833 ± 0.011
ρ_i	0.947	0.985	0.983	
$\sigma_{uncorr} \pm \sigma_{corr}$	0.016 ± 0.011	0.013 ± 0.011	0.015 ± 0.011	
σ_{track}	± 0.0027	± 0.0028	± 0.0028	

Table 5.3: Relative tracking efficiency of π_s for the LS1 data sample and the run dependent MC.



(a) Asimov (b) Mu



(c) Sigma (d) Sigma

Figure 5.17: Post-fit agreement for the distribution of Δm in the four momentum bins of p_{π_s}. In these figures, the distributions of all decay channels are combined.

5.8 Charge Asymmetry

The reconstruction chain used for the determination of the slow pion tracking efficiency comparison between data and MC can also be used to monitor charge dependent track finding inefficiencies for π_s. It has to be mentioned that this monitoring is entirely illustrative and not a precise study. For the investigation of the charge dependency of the track finding, we place an additional constraint to select the signal enriched region of $|\Delta E| < 0.05$ GeV. The charge of the pions is identified through the sign of their PDG code. This study is meant to show, that the data and MC samples show a similar behaviour and not provide to any calibration factors or systematic uncertainty estimates.

First we compare the charge dependent distributions of the transverse momentum p_T and the cosine of the polar angle cosθ in Fig.5.19. For this comparison the number of events in the MC is scaled down to the integrated luminosity of the data sample via the ratio 364/1443. The shown uncertainties are only statistical. From these distributions, we can conclude, that charge dependent inefficiencies are overall well modeled. However, we can

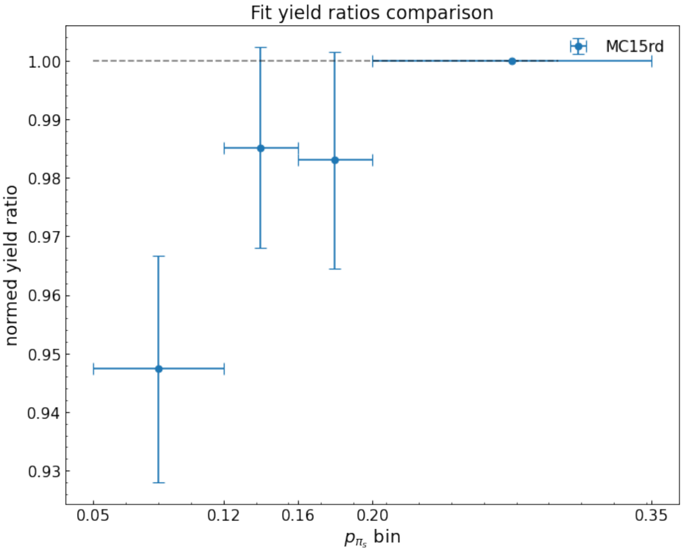


Figure 5.18: Visual comparison of the normalized yield ratios ρ_i for all kinematic p_{π_s} bins.

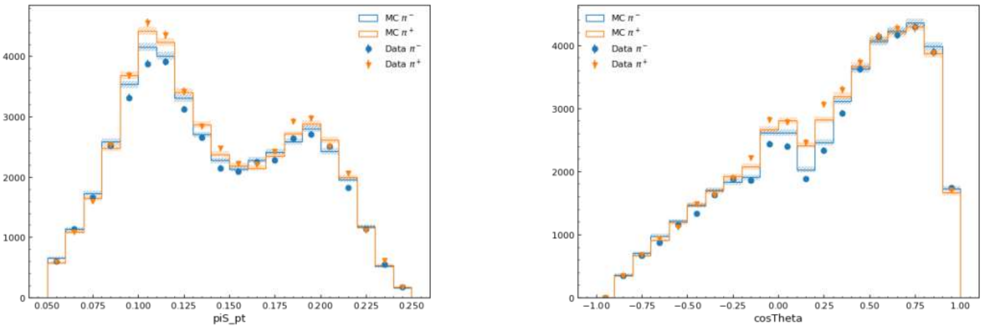


Figure 5.19: One dimensional comparison of the data and MC charge dependent efficiencies in bins of the transverse momentum of π_s p_T (left) and the cosine of the polar angle $cos\theta$ (right).

observe slight shape discrepancies for the region of $p_T \in [0.10, 0.15]$ GeV and cos$\theta \in [0, 0.25]$. In these regions the π_s has a higher reconstruction efficiency with a positive charge for both samples, however some minor shape discrepancies can be seen between data and MC.

In addition to the shape comparison, two dimensional asymmetry factors are determined, based on the number of reconstructed events:

$$asymmetry\ factor = \frac{N^{\pi^+} - N^{\pi^-}}{N^{\pi^+} + N^{\pi^-}}, \tag{5.4}$$

where N^{π^+} and N^{π^-} correspond to the number of reconstructed pions with a positive and negative charge, respectively. These asymmetry factors are calculated two dimensional in bins of p_T and $cos\theta$, separately for both the MC and data. The resulting asymmetry factors can be found for both samples in Fig.5.20. These asymmetry factors only have the statistical

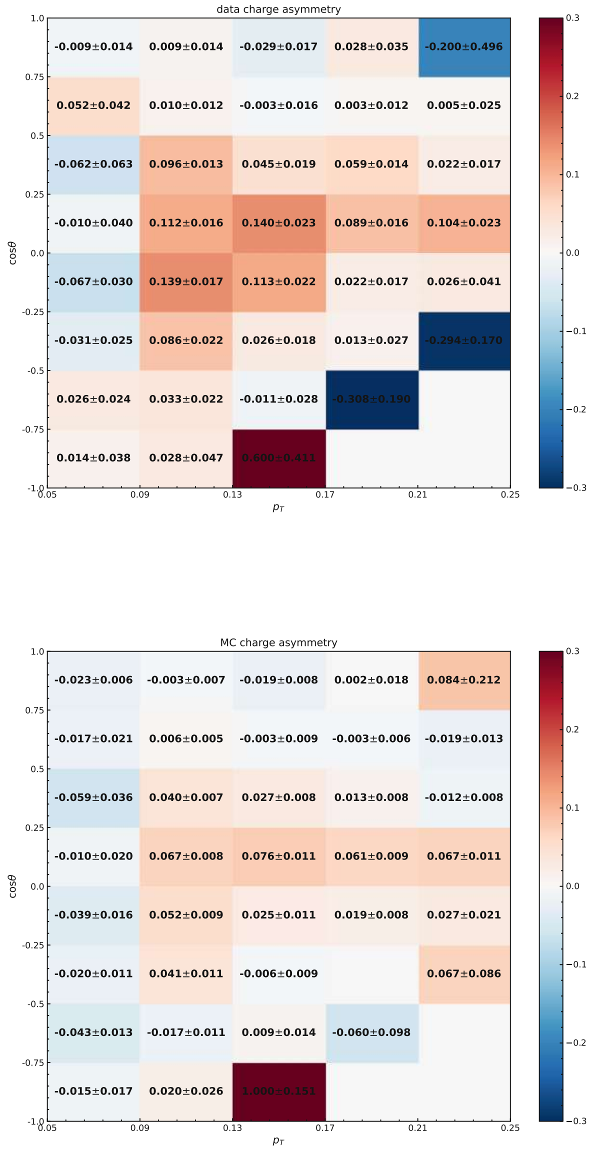


Figure 5.20: Resulting two dimensional charge asymmetry in bins of $cos\theta$ (y-axis) and p_T (x-axis) of the LS1 data in the upper plot and MC15 in the lower plot.

uncertainty attached, which were determined via a bootstrapping approach. Overall both the MC and data show the same features in the two dimensional distribution.

CHAPTER 6

Determination of $|V_{cb}|$ from $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays

6.1 Analysis Overview

In the standard model of particle physics, quark mixing is described by the CKM matrix, first introduced in 1973. [2, 14] The transition rate from b to c quarks is proportional to the squared CKM matrix element magnitude $|V_{cb}|$. At the time of writing, the favored approach at Belle II to measure $|V_{cb}|$ is through semileptonic $B \to X_c\ell\nu_\ell$ decays due to their high branching fractions and the factorization of the hadronic contributions (see Sec.2.6). In this decay the B is either a charged or neutral B meson. The term X_c refers to a hadronic system containing a c quark, ℓ is a light, charged lepton, i.e. either an electron or a muon and ν_ℓ is their respective neutrino.

In general the measurement of $|V_{cb}|$ can be done through two different approaches, i.e. either by reconstructing the decay inclusively or exclusively. During an inclusive reconstruction one reconstructs all $B \to X_c\ell\nu_\ell$ final states, while for exclusive reconstructions a single decay chain is explicitly reconstructed, e.g. $B \to D^*\ell\nu_\ell$. These two approaches serve as crucial cross-checks for each other. While their input from theory and the uncertainties differ between these two approaches, their results are still expected to agree with one another. However, this is not the case, since the two approaches exhibit a long lasting 3.3σ discrepancy. [21] This shows that our understanding of $|V_{cb}|$ is still incomplete.

We measure $|V_{cb}|$ through the exclusive reconstruction of $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays with the subsequent decay $D^{*+} \to D^0\pi^+$ followed by $D^0 \to K^-\pi^+$. These decays are reconstructed from the LS1 data sample of Belle II, which corresponds to an integrated luminosity of ~ 364 fb^{-1} of data taken at the $\Upsilon(4S)$ resonance of 10.58 GeV and a MC sample corresponding to an integrated luminosity of ~ 1443 fb^{-1}. Additionally, the off resonance data taken at ~ 5 MeV below the resonance will also come to use during a validation step. A detailed description of the data and MC samples can be found in Sec.4.2.

By reconstructing $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays it is possible to determine their decay rate and use the dependencies between the partial decay rates, $|V_{cb}|$, the form factors and the four kinematic variables, i.e. the hadronic recoil w and three angular variables. We extract $|V_{cb}|$ through the partial decay rates in bins of the kinematic variables w and $\cos\theta_\ell$. A

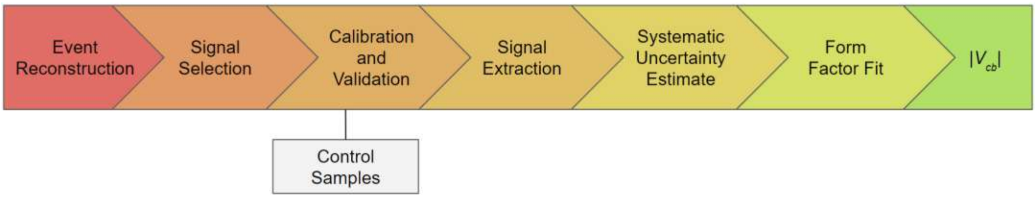


Figure 6.1: Simplified work flow schematic of the $|V_{cb}|$ analysis.

discussion of the theoretical framework of the $|V_{cb}|$ extraction from $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays and the kinematic variables is provided in Sec.2.6.1. In addition to the $|V_{cb}|$ measurement, we will also determine the branching fraction $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$ and probe lepton flavor universality by calculating the ratio $\mathcal{B}(\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$.

This chapter will be structured similar to the analysis workflow provided in Fig.6.1. First we present a discussion of the reconstruction process of the events, the signal selection and the reconstruction of kinematic variables in Sec.6.2. This will be followed up by discussing the necessary calibrations for data and MC in Sec.6.3 and a validation of the agreement between data and MC distributions in Sec.6.4. In Sec.6.5 we will elaborate on the signal extraction procedure and its closure tests. Furthermore the estimation process of the systematic uncertainties will be provided in Sec.6.6 and the $|V_{cb}|$ extraction method will be discussed in Sec.6.7. We will then present our results in Sec.6.8.

6.2 Reconstruction of $B \to D^*\ell\nu$

6.2.1 Event Reconstruction

Potential $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ $(\ell = e, \mu)$[1] candidates are reconstructed through the subsequent decays $D^{*+} \to D^0\pi^+$ and $D^0 \to K^-\pi^+$, as illustrated by the decay chain in Fig.6.2. The reconstruction process is done using an untagged approach, i.e. a reconstruction without attempting an explicit reconstruction of the tag-side B_{tag} in any manner and instead only gathering all left over particles after the B_{sig} to a so-called rest of event, see Sec.4.3.1 for a detailed discussion. This untagged approach comes at a cost of losing the ability to place kinematic constraints on the ν involved in the semileptonic decay and the additional background suppression provided by the FEI. The inability to reconstruct the neutrino momentum results in a worse resolution of the kinematic variables, because they have to be estimated through other strategies, e.g. by utilizing the information provided by the rest of the event. However using the untagged approach also has benefits: The event count is vastly improved due to an efficiency of $\mathcal{O}(100\%)$ compared to $\mathcal{O}(0.1\%)$ of the hadronic tagging. A higher event count of course also increases the background noise, however due to the low mass difference between the D^* and D mesons of ~ 145 MeV for the $D^{*+} \to D^0\pi^+$ decay the kinematics of the π is heavily restricted, resulting in $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ being a rather clean decay mode.

The first step of the reconstruction chain is to combine a K^- and π^+ candidate to a D^0 meson. To improve the chance of selecting accurately reconstructed, charged tracks of the final state particles K^-, π^+ and ℓ^-, we require the following distances of the point of closest approach (POCA) with respect to the interaction point (IP): the POCA in the $r - \phi$ plane

---

[1] The inclusion of charge conjugate decays is implied.

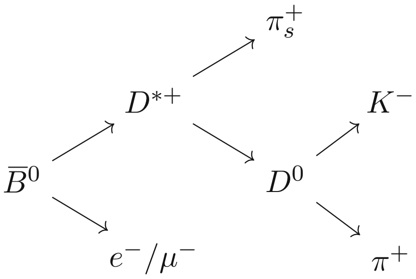


Figure 6.2: Illustration of the reconstructed $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decay chain with the subsequent decays $D^{*+} \to D^0\pi^+$ and $D^0 \to K^-\pi^+$.

dr is required to be below 1 cm and the distance to the POCA on the z-axis $|dz|$ is required to be below 3 cm. Furthermore, we require these particles to have a polar angle θ within the CDC acceptance range of $\theta \in [17°, 150°]$. The invariant mass of the newly combined π^+ and K^- candidates is required to only deviate by up to 25 MeV from the world average PDG value of the D^0 mass of ~ 1.864 GeV.

We then combine the reconstructed D^0 meson with a charged slow pion to a D^{*+} candidate. The π_s^+ candidates don't need to fulfill the previously mentioned tracking criteria due to their low momentum but are required to have momentum below 500 MeV. To suppress the contribution from $q\overline{q}$ events the D^{*+} candidates are required to have a momentum $p^*_{D^*}$[2] below 3 GeV. Additionally, the mass difference Δm between the D^{*+} and D^0 mesons is constrained to the region of [0.141, 0.156] GeV.

Finally, the $\overline{B}^0$ is reconstructed by combining the D^{*+} with a charged, light lepton ℓ^-, i.e. either an electron or a muon. Before initiating the last reconstruction step the energy and momenta of all electrons are corrected to account for Bremsstrahlung effects. The concept of Bremsstrahlung corrections is discussed in the following Sec.6.2.2. Both types of leptons are required to have a momentum within the region of $p^*_\ell \in [0.4, 2.6]$ GeV. Furthermore muon candidates are required to have a global muon identification probability above 0.9, excluding the SVD information. Electron candidates are required to have an electron ID likelihood above 0.9, which is based on a BDT score. These thresholds are based on recommendations of the Belle II performance group. The particle identification and BDTs are further discussed in Sec.3.3 and Sec.4.7.5, respectively.

The rest of the event (ROE) is reconstructed by collecting the remaining pions and photons of the event. To select only pions with reliably reconstructed tracks we require them to have a transverse momentum $p_t > 0.1$ GeV, a polar angle within CDC acceptance $\theta \in [17°, 150°]$ and IP parameters of $dr < 0.5$ cm and $|dz| < 3.0$ cm. All photons in the ROE must pass the same ECL selection criteria as photons for the Bremsstrahlung selection in Sec.6.2.2. Furthermore all particles are required to have a momentum in the center of mass frame below 3.2 GeV.

6.2.2 Bremsstrahlung

Due to Bremsstrahlung effects any electrons or positrons have to be corrected before involving them in the B meson reconstruction by recovering the lost energy and momentum. This process is done for both the data and MC events. It can be shown, that the Bremsstrahlung

---

[2] The asterisk denotes the center of mass frame for momenta and energies.

photons are mostly emitted in a cone around the direction of the momentum of the electrons[3]. Therefore we search for emitted photons on a cone around the initial momentum of the electron. This procedure is done via a Bremscorrection module implemented in basf2, that corrects the four momentum of the lepton by adding the energy and momenta of photons that fall within the predicted emission cone. Additionally the photons have to fulfill a variety of criteria based on their ECL clusters and their emission angle ϕ between the electron and the photon.

The ECL based baseline criteria each photon has to pass to be considered for Bremsstrahlung recovery are the following: The polar angle of the centroid position of the ECL cluster must lie within $0.2967 < \theta < 2.6180$ rad. Additionally the absolute value of the ECL cluster timing has to be below 200 ns. The cluster timing is defined as the difference between the photon timing and the event time. If the cluster originates from a photon that was emitted from the interaction point, the cluster timing would be 0. Finally all photons must have over 1.5 ECL cluster hits, i.e. a sum of weights w_i ($w_i < 1$) of all crystals within an ECL cluster.

The next set of selection criteria is based on the ECL region, where their cluster was registered. The ECL is split into three regions, i.e. the forward region within a particle's polar angle range of $12.4°\theta < 31.4°$, a barrel region within $32.2°\theta < 128.7°$ and lastly the backward region of $130.7°\theta < 155.1°$. For each region the ECL cluster energy E_{ECL} of the photons have to pass different thresholds:

- Forward: $E_{ECL} \in [0.04, 0.09]$ GeV
- Barrel: $E_{ECL} \in [0.09, 1.20]$ GeV
- Backward: $E_{ECL} \in [0.055, 0.900]$ GeV

Additionally the photons have to pass selection criteria, which are based on the momentum of the electron and the angle ψ between the electron and the photon, as well as the region where the ECL cluster was registered:

- $p_e < 0.6$ GeV: Forward region and $\psi < 0.1368$ rad
- $p_e \in [0.6, 1.0]$ GeV: Barrel region and $\psi < 0.0737$ rad
- $p_e \geq 1.0$ GeV: Backward region and $\psi < 0.0632$ rad

6.2.3 Event Classification

After the reconstruction process, the candidate selection will be further optimized by utilizing the MC simulation. For the reconstructed MC events we obtain information about the relations between all particles of each event. The knowledge of the origin of the events and the relations between the particles allows us to separate the events into a signal class and various background categories to improve our selection criteria based on a figure of merit (fom), see Sec.4.4.2. The events are split into the following four classes:

---

[3] The same behaviour and correction procedures for positrons is implied.

- **Signal**: The $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decay chain is identified as correctly reconstructed. The correct reconstruction of the B candidate is checked by utilizing the output of the truth matching algorithm of basf2 that checks the relations between the combined particles in the whole decay chain. We set the truth matching algorithm to take into consideration, that a correctly reconstructed decay chain still has a missing neutrino. The truth matching will provide each B candidate with an associated boolean signal flag, which is equal to 1 for correctly reconstructed events. This means, the combined D^* and ℓ are correctly reconstructed, true daughters of the B meson and the D daughter of the D^* is also correctly reconstructed and a true granddaughter of the B candidate.
- **Continuum**: Candidates that originate from $q\overline{q}$ events instead of $B\bar{B}$ events. These events are identified through an event based boolean continuum flag provided by the MC information.
- **True D^***: Candidates originating from $\Upsilon(4S) \to B\bar{B}$ events, which is identified based on a false continuum flag. Furthermore this event type is required to have a B meson that is not flagged as a signal candidate. However, the D^* meson in the decay chain is required to be a true D^* meson. This identification is made based on its associated MC PDG code, that explicitly defines the particle type.
- **Fake D^***: This background category has to flagged as a $B\bar{B}$ event and must have a false signal flag. Additionally the D^* candidate is not allowed to have a correctly matched MC PDG code.

6.2.4 Candidate Selection

After the reconstruction process, the candidate selection is further optimized based on the MC simulation. During this cut optimization the MC will be split into one signal and three background categories, mentioned in Sec.6.2.3: Signal, Continuum, True D^*, Fake D^*. This process will be done by scanning the fom. In Fig.6.3, the cut flow of the fom is shown for the electron and muon channel separately. This process leads to the following selection criteria:

- The visible energy of the event E^*_{vis} is required to be within [4.0, 10.5] GeV.
- As a main measure for the suppression of $q\overline{q}$ events a requirement on the 2nd to 0th ratio of the Fox Wolfram moments R_2 is placed: $R_2 < 0.3$. Additionally the Kakuno-Super-Fox-Wolfram (KSFW) variable H^{so}_{02} is restricted to be below 0.25 to further reduce continuum contributions. An discussion of both the Fox Wolfram moments and KSFW variables is given in Sec.4.5.
- The mass deviation of D meson candidates from the world average PDG mass $|m_{D^0} - m^{PDG}_{D^0}|$ must be below 0.015 GeV. As a last measure for $q\overline{q}$ suppression, the momentum $p^*_{D^*}$ of the D^* candidates is required to be below 2.5 GeV.
- Slow pions are restricted to a momentum range of $p_{\pi_s} \in [0.05, 0.30]$ GeV, which is the region covered by the slow pion efficiency study.
- The leptons are required to have a momentum within $p^*_e \in [0.5, 2.4]$ GeV and $p^*_\mu \in [0.9, 2.4]$ GeV in case of an electron or a muon, respectively. The difference between

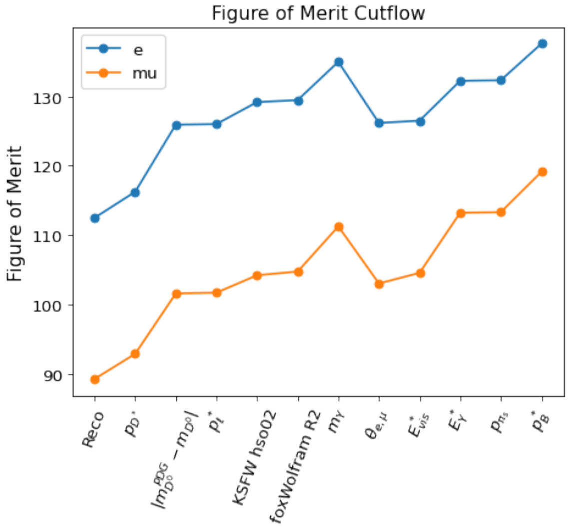


Figure 6.3: Development of the figure of merit with each additional selection criterion.

the selection criteria originates from the worse agreement between the data and MC distributions for muons in the lower momentum regions. To reduce the potential impact of systematic uncertainties of the lepton ID, the lepton candidates are restricted to the polar angle region of $\theta \in [0.56, 2.23]$ rad of the barrel.

- Lastly the Y[4] system, must exhibit a momentum of $p_Y^* < 2.2$ GeV, a mass $m_Y > 2.5$ GeV and an energy within $E_Y^* \in [3.2, 5.1]$ GeV.

6.2.5 Control Samples

To calibrate the MC modeling of the background distributions, two additional control samples were reconstructed: an off-resonance sample and a same sign sample. The control samples were reconstructed through the same procedure and by applying the same selection criteria as for the general sample, discussed in Sec.6.2.1 and Sec.6.2.4, respectively.

The off-resonance sample contains only data taken with a center of mass energy that is ∼ 5 MeV below the $\Upsilon(4S)$ resonance of 10.58 GeV. This sample consists only of $q\overline{q}$ events. Therefore it can be used to validate the continuum model of the MC and correct its continuum distributions. For the LS1 data set, this sample has a size that corresponds to an integrated luminosity of ~ 42.56 fb^{-1}.

A second control sample will consist of reconstructed $\bar{B}^0 \to D^{*-}\ell^-\nu_\ell$ candidates. This sample will be referred to as same sign sample. The reconstruction of the same charged candidates is carried out on the same samples as the original reconstruction, i.e. the on-resonance LS1 data and the general MC. We will use this wrong charge sample to calibrate the MC modeling of fake and true D^* background events.

---

[4] Y refers to the $D^*\ell$ system, see Sec.4.3.

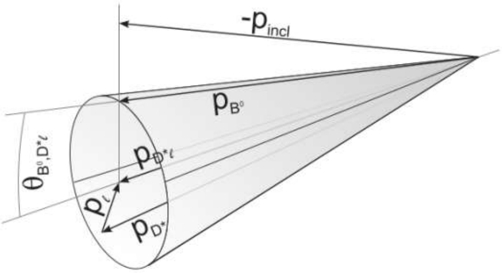


Figure 6.4: The illustration shows the cone with an opening angle of θ_{BY} and the estimated flight direction of the B meson based on $-\vec{p}_{incl}$ of the ROE. [47]

6.2.6 Kinematic Variables

In order to measure the magnitude $|V_{cb}|$ we need to determine the partial decay rates $\Gamma(\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l)$ as a function of the kinematic variables, i.e. the hadronic recoil w and the three angular variables $\cos\theta_\ell, \cos\theta_v$ and χ. An introduction of the kinematic variables can be found in Sec.2.6.1. To calculate the kinematic variable knowledge about the direction and rest frame of the signal B meson is needed. However, since we chose to reconstruct the events using an untagged approach to increase the number of events, the neutrino momentum can't be reconstructed by determining the missing momentum based on a reconstructed B_{tag}. Therefore we need to estimate the necessary information of the B candidates through other means.

The starting point of the B momentum direction estimate is the opening angle θ_{BY}. It can be shown that the B direction has to lie on a cone with the opening angle of 2 θ_{BY} around the momentum $p^*_{D^*\ell}$ of the Y system, as shown in Fig.6.4. The approach that makes use of this property and combines it with information available from the rest of the event is called the Rest of Event (ROE) frame. In the ROE frame approach, the momentum $\vec{p}_{incl}$ of the second B meson in the laboratory frame is estimated by combining the momenta of left over particles not used during the reconstruction of B_{sig}:

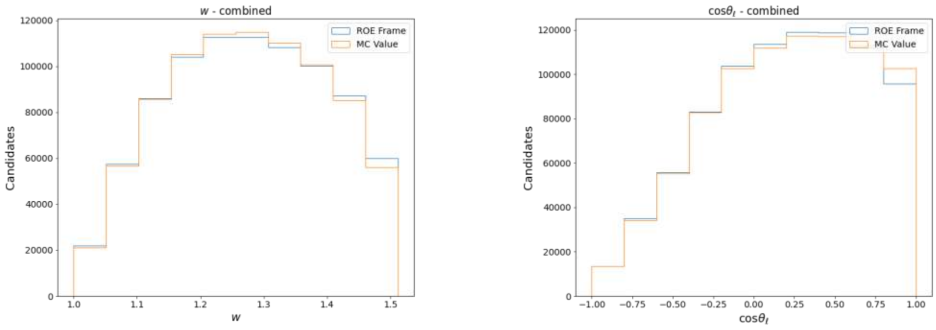
$$\vec{p}_{incl} = \sum_i \vec{p}_i. \tag{6.1}$$

All particles used for the reconstruction of $\vec{p}_{incl}$ must pass the selection criteria for ROE particles mentioned in Sec.6.2. The estimated momentum is then transformed to the center of mass frame and the energy of the partner B meson is estimated to be half the center of mass energy $E^*_{incl} = E^*_{beam} = \sqrt{s}/2$ to estimate the four momentum. We then choose the B direction to lie on the cone direction with minimal distance to $-\vec{p}^*_{incl}$.

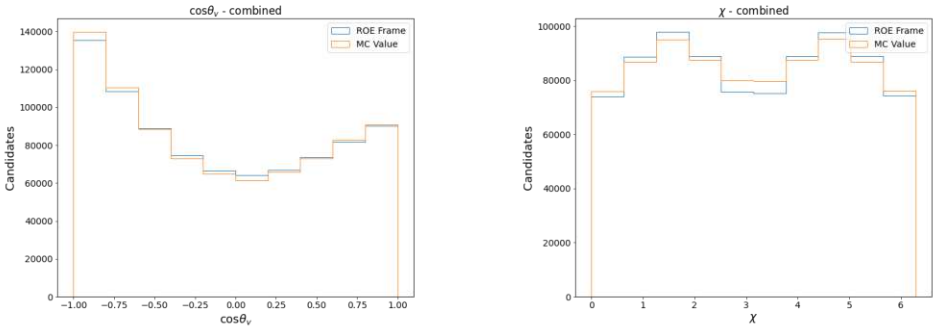
Since this ROE approach gives an estimate that is only based on rest of event particles instead of a B_{tag} reconstruction its resolution is limited. Therefore we estimate the resolution for the kinematic variables to know how to choose our binning for the decay rates and whether an unfolding process (see Sec.4.7.6) will be necessary to take bin by bin migrations into consideration for the signal extraction. The resolution of the kinematic variables is determined by calculating the difference between the reconstructed value kin_{ROE} and the true MC value kin_{MC}, which is calculated using the simulated neutrino. A summary of the calculated resolution values for each kinematic variable is given in Tab.6.1. These values are determined based on percentiles and the $68-95-99.7$ rule. Additionally we also checked if

Variables	Median	Percentiles	Resolution
w	0.00094	[-0.01891, 0.02221]	0.02056
$\cos\theta_\ell$	-0.00348	[-0.05382, 0.03492]	0.04437
$\cos\theta_V$	0.00371	[-0.05938, 0.06972]	0.06455
χ	-0.00026	[-0.2796, 0.28027]	0.27994

Table 6.1: Summary of the estimated resolution for each kinematic variables using the 68-95-99.7 rule of statistics. This table provides the median, the lower and upper 0.68 percentiles and the resulting resolution estimate.



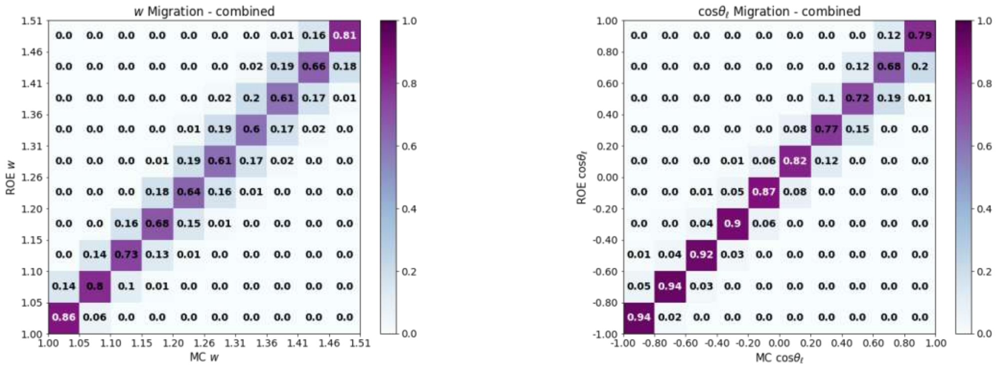
(a) w (b) $\cos\theta_\ell$



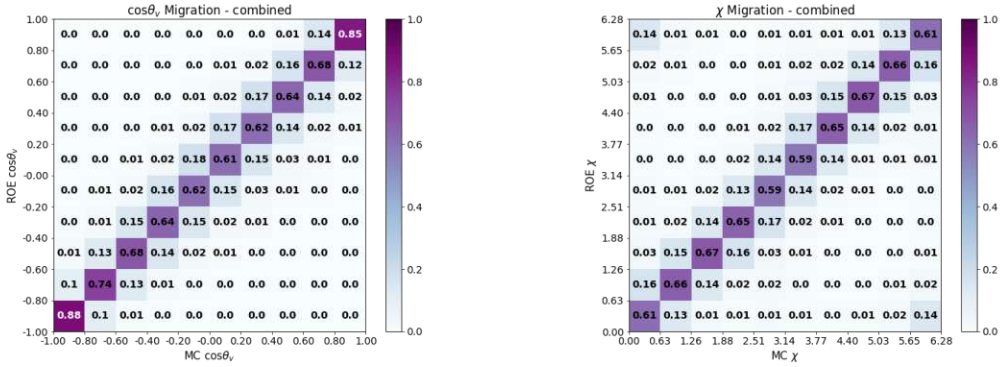
(c) $\cos\theta_V$ (d) χ

Figure 6.5: The validation of the ROE reconstruction process of the kinematic variables through a comparison of the shape between the ROE value kin_{ROE} (blue) and the true value $kin_{MC}(orange)$. In all figures we show the combined agreement of both lepton channels.

the ROE algorithm is working as intended by comparing the shapes of reconstructed and true values of the same sample in Fig.6.5, which show some minor disagreement between the two shapes. However overall they agree well and this comparison is made without showing any uncertainties and some disagreement between the estimate and the true value is expected. Lastly, we also checked the bin migration matrix (see Sec.4.7.6) of the kinematic variables for a binning scheme of 10 even bins for any unusual behavior of the reconstruction. The migration matrices are presented in Fig.6.6 and show the highest values in the diagonal with some migration to mainly adjacent bins, i.e. the behaviour one would expect.



(a) w (b) $\cos\theta_\ell$



(c) $\cos\theta_V$ (d) χ

Figure 6.6: Bin migration of the kinematic variables for the combination of both lepton channels. These two dimensional heat maps show the probability, that a kinematic variable with a value kin_{MC} falls into a reconstructed bin kin_{ROE}. The x-axis corresponds to the true value of the variables obtained from MC, while the y-axis represents to the ROE value.

6.3 Data and MC Calibration

The B meson candidates reconstructed from both the data and the MC samples need to be calibrates to counteract various biases, e.g. during the track reconstruction, or mismodeling in the MC, e.g. outdated branching fractions for the generation. In general the correction factors and guidance on their application are provided by the Belle I performance group.

For the data calibration, we apply the track momentum scaling for any charged final state particle tracks and photon energy bias corrections for any photons involved. The momentum scaling and the energy bias correction were already discussed for the slow pion study in Sec.5.4.1 and Sec.5.4.2, respectively and therefore won't be discussed further since it follows the same principles. Additionally the energy of the electrons of the data and MC sample are corrected to account for Bremsstrahlung, as described in Sec.6.2.2.

For the MC samples a variety of new corrections will be introduced: First a leptonID correction based on the same principles as the hadronID for the slow pions. Also the slow pion efficiency correction weights determined in Sec.5.7 will be applied. The other corrections include: an $X_c\ell\nu$ branching fraction calibration, BGL form factor corrections and lastly, $\cos\theta_{BY}$ background shape corrections.

6.3.1 Slow Pion Efficiency

Since we measure $|V_{cb}|$ through the decay $B \to D^*\ell\nu_l$ with the subsequent decay $D^* \to D\pi$, the tracking efficiency difference between data and MC has to be taken into account. The whole process of determining the relative tracking efficiency of π_s from $B \to D^*\pi$ decays is described in Ch.5. We calibrate our MC event count by applying correction weights that are dependent on the momentum p_{π_s} of the slow pion used in the reconstruction chain, according to summary in Tab.5.3. If the slow pion in a reconstruction chain is identified as a true pion and its momentum falls within the p_{π_s} region of the first bin, then the associated normalized yield ratio will be used as a correction weight for the candidate. This process is done for all reconstructed events.

6.3.2 LeptonID

The leptonID correction is generally done using the same procedures as the hadronID correction described in Sec.5.3.2. While the muonID fake rate tables are created via the systematics frame work, correction tables for the muonID efficiency and the BDT based electronID are provided by the performance group. For both the electron and muon channel we observed full coverage of our candidates by the correction tables. The two dimensional coverage plots for both decay channels can be found in App.A.1.

6.3.3 $B \to X_c\ell\nu$ branching fractions

The branching fractions of semileptonic $B \to X_c\ell\nu$ decays used in the decay file to generate the MC sample are updated to iso-spin averaged branching fractions based on most recent measurements. These averages are provided by the collaboration and listed in Tab.6.2, where a comparison between the branching fractions of the decay file and the updated values is provided. When comparing the measured inclusive $B \to X_c\ell\nu$ branching fraction to the sum of all exclusive decays a discrepancy can be observed, i.e. the so-called semileptonic gap. This gap is bridged by adding equal amounts of $B \to D\eta\ell\nu_\ell$ and $B \to D^*\eta\ell\nu_\ell$ decays. To be able to apply this correction, we first identify the decay chains of the B mesons using the information about the daughters and their particle type through their relations and MC PDG codes provided by the generator. Depending on the determined decay chain we then associate each candidate with a weight equal to the ratio of $\mathcal{B}_{update}/\mathcal{B}_{file}$, where $\mathcal{B}_{update}$ and $\mathcal{B}_{file}$ are the updated and original branching fraction, respectively.

6.3.4 Form Factors

Similarly to the $B \to X_c\ell\nu$ branching fractions, we also update the form factors used to model $B \to D^*\ell\nu_l$ decays in the decay file. We update the $B \to D^*\ell\nu_l$ form factors to the recent BGL$_{121}$ results from Ref.[59]. The events are reweighed using the following ratio of normalized decay rates to account for shape differences of the form factors:

$$w_{FF} = \frac{\Gamma_{\text{decay file}}}{\Gamma_{\text{final version}}} \frac{\frac{d^4\Gamma_{\text{final version}}}{dw\,d\cos\theta_\ell\,d\cos\theta_v\,d\chi}}{\frac{d^4\Gamma_{\text{decay file}}}{dw\,d\cos\theta_\ell\,d\cos\theta_v\,d\chi}}, \tag{6.2}$$

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 6.2: Overview of the updated iso-spin averaged branching fractions and their original values.

where Γ is the total decay rate and $\frac{d^4\Gamma}{dw\,d\cos\theta_\ell d\cos\theta_V d\chi}$ the differential decay rate calculated for a given set of kinematic variables. The calibration weight w_{FF} is determined for each signal candidate and the form factors for the decay rates are estimated using the MC values of all four kinematic variables $w, \cos\theta_\ell, \cos\theta_v$ and χ.

6.3.5 Background Shape of cosθ_{BY}

The cosine of the opening angle $\cos\theta_{BY}$ will be one of two observables used during signal extraction. We observed some mismodeling in the $\cos\theta_{BY}$ side-band region of $\cos\theta_{BY} > 2$. This mismodeled region is mainly dominated by contributions of fake D^* and true D^* events, which is the reason behind the reconstruction the same sign sample mentioned in Sec.6.2.5. This wrong charge control sample is used to derive weights that correct the mismodeled shape of these two background contributions.

To determine the correction weights for the background shape we first apply all previously mentioned calibration factors and scale down the MC event count using the integrated luminosity ratio of 364/1443. For this calibration process we split our samples into $B \to D^*e\nu_e$ and $B \to D^*\mu\nu_\mu$ samples. We then bin the $\cos\theta_{BY}$ distribution into 30 bins within the region of $\cos\theta_{BY} \in [-4, 4]$. The choice of 30 bins was made as a compromise between computation time and impact of the corrections, i.e. we saw no significant improvement by further increasing the bin counts and only made the process more time consuming. Since it can be observed that the continuum is well modeled by the MC, we subtract the continuum events from the simulation and also remove the same amount of data event in each $\cos\theta_{BY}$ bin. Furthermore signal events are excluded by design, i.e. reconstructing only candidates with a mismatched charge, therefore this correction will only affect true D^* and fake D^* events. Next we calculate the data/MC candidate ratio for each of 30 bins and add the weight to all true and fake D^* events depending on their respective $\cos\theta_{BY}$ bin.

6.4 Data-MC Agreement

The purpose of investigating the following data-MC agreement plots is to validate that the MC shows no major mismodeling of the data and no further calibration is needed. The MC samples shown in the figures presented in this section are scaled down to the integrated luminosity of the data by applying a weight equal to the data/MC luminosity ratio of 364/1443. For both samples the selection criteria listed in Sec.6.2.4 are applied. All distributions are shown only for the $B \to D^* e\nu_e$ channel, due to the similarities between channels unless otherwise mentioned. The spectra of the $B \to D^* \mu\nu_\mu$ channel can be found in App.A.2. Furthermore both the MC and data include all correction weights mentioned in Sec.6.3. The only exception to the listed properties is the n-1 cutflow, on which we elaborate further in Sec.6.4.3.

Each figure also features a subplot that shows the 1 and 3 σ band of the following pull:

$$pull = \frac{N_{data} - N_{MC}}{\sqrt{\sigma^2_{data} + \sigma^2_{MC}}}, \tag{6.3}$$

where N_{data} and N_{MC} are the number of signal and background events in a given bin, respectively and σ_{data} and σ_{MC} their associated uncertainties. In the pulls and for the uncertainties shown in the distributions we include not only the statistical uncertainties, but also the systematic uncertainties arising from: leptonID, slow pion efficiency, form factor reweighing and the $X_c\ell\nu$ branching fraction update.

6.4.1 Off-Resonance

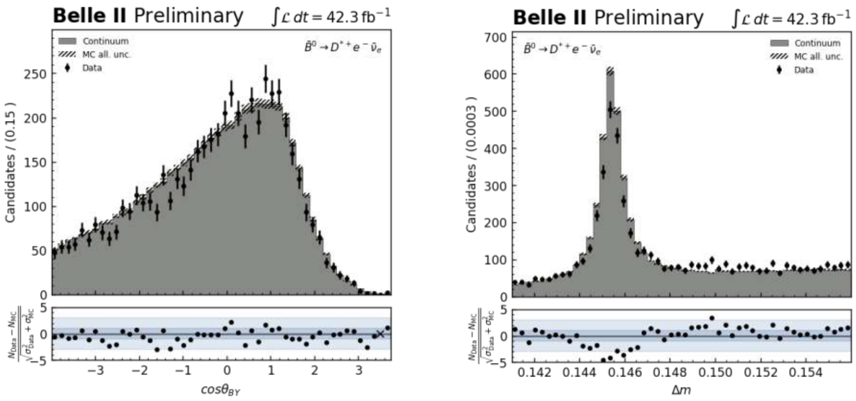
In this section we investigate how well the continuum MC models $q\overline{q}$ events by comparing it to the distribution obtained from reconstructing B mesons on the off-resonance data corresponding to an integrated luminosity of 42.3 fb^{-1}. Further details about the off-resonance data can be found in Sec.4.2. We show the agreement between the distributions of data and MC for all parameters used during the signal extraction for $B \to D^* e\nu_e$ in Fig.6.7. For these comparisons, all correction weights are applied, the number of MC candidates is scaled down and they are selected based on the continuum classification in Sec.6.2.5.

Overall very good agreement between the off-resonance data and the MC can be observed for both lepton channels, with barely any deviations above 3σ. Therefore we conclude that no additional calibration with respect to the continuum events is necessary.

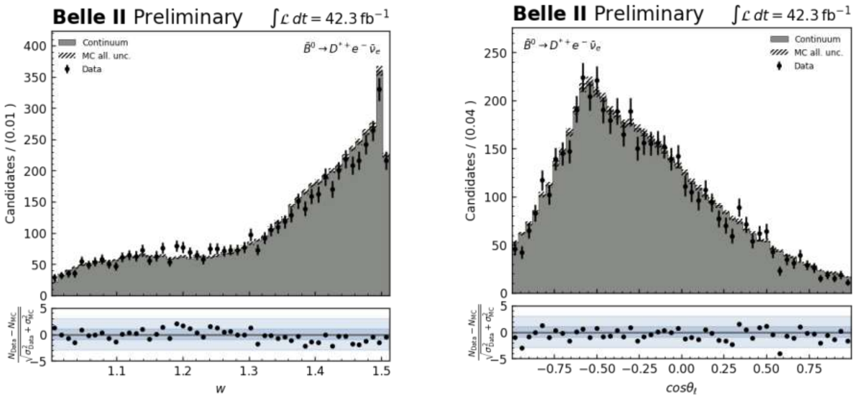
6.4.2 Same Sign

Due to observed discrepancies in the data-MC agreement of the upper side-band region of $\cos\theta_{BY} \in [2, 4]$ it was decided to reconstruct another control sample for calibration purposes, i.e. the same sign sample consisting of reconstructed $\bar{B}^0 \to D^{*-}\ell^-\nu_\ell$ decays discussed in Sec.6.2.5. The reasoning behind choosing this sample is that the upper side band region is only populated by fake D^*, true D^* and continuum events. However due to the lack of other viable control channels it was chosen to use a reconstruction with a same sign for both D^{*-} and ℓ^-, which exhibits similar kinematic properties. Additionally the $\cos\theta_{BY}$ distribution of the continuum has shown good agreement and is therefore proven not to be the source of this discrepancy.

The distributions before and after the application of the calibration factor is shown in



(a) $\cos\theta_{BY}$ (b) Δm



(c) w (d) $\cos\theta_\ell$

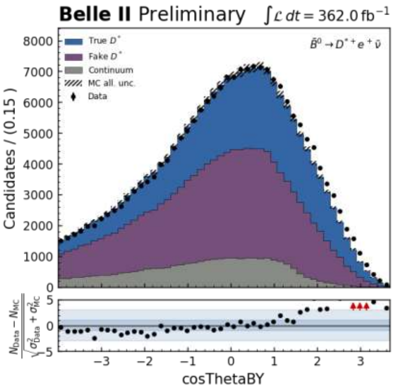
Figure 6.7: The MC and off-resonance data are shown to agree well for all parameters of interest for $B \to D^* e \nu_e$. The agreement for the $B \to D^* \mu \nu_\mu$ shows a similar behaviour and can be found in App.A.2.

Fig.6.8 for the electron channel. From these figures one can observe the big discrepancy caused in the side-bands by the fake and true D^* background components, which exceeds even 5σ for some bins. One can also observe that while in the upper side band the MC underestimates the data, it overestimates the data in the lower side-band. This behavior in the lower side-band is also corrected by the calibration.

6.4.3 N-1 Cutflow

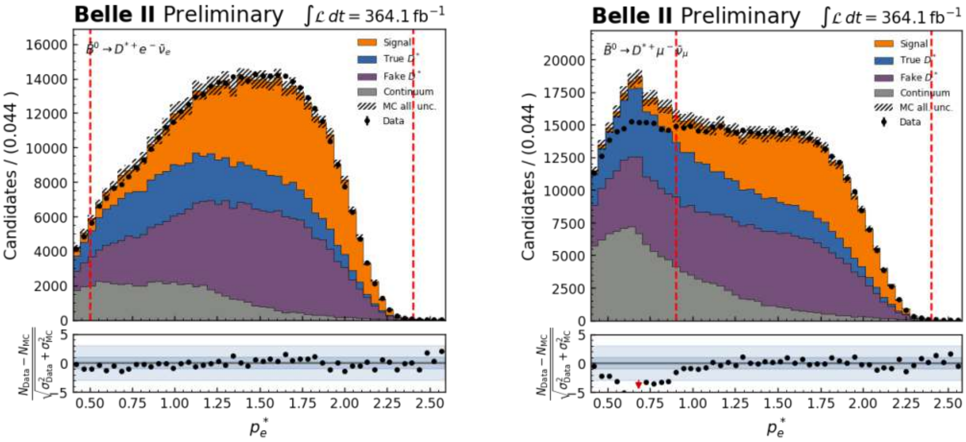
In the so-called n-1 cutflow plots, we show each variable from the tight signal selection section in Sec.6.2.4 with all cuts applied, except for the cut on the shown variable. Furthermore these cutflow plots don't show the whole MC sample of 1443 fb^{-1}, but only 1/10 of the whole MC sample, due to computational restrictions. These 143.3 fb^{-1} of the MC are created by randomly picking 1/10 of the events of each MC production listed in Sec.4.2 and are used to investigate whether a variable used during the selection is mismodeled and to visualize their impact on background suppression.

In these plots the red line indicates the selection threshold. The distributions of all parameters of interest are shown in App.A.2 for both the $B \to D^* e \nu_e$ and $B \to D^* \mu \nu_\mu$



(a) Before correction. (b) After correction.

Figure 6.8: The comparison between the $\cos\theta_{BY}$ distribution before and after the calibration is shown. It can be observed, that after the application of the newly calculated correction factors to the fake and true D^* component only very minor discrepancies from the continuum component are left over. This observation further solidifies the assumption, that the original disagreement is not caused by continuum events.



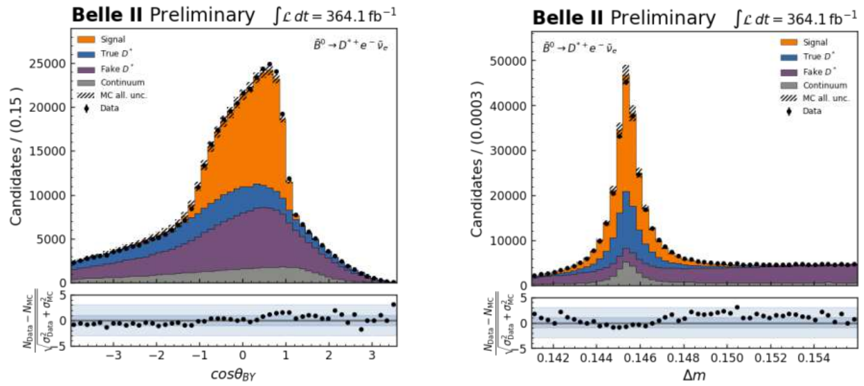
(a) $B \to D^* e\nu_e$ (b) $B \to D^*\mu\nu_\mu$

Figure 6.9: Comparison of the lepton momentum p^*_ℓ between the electron and muon channel.

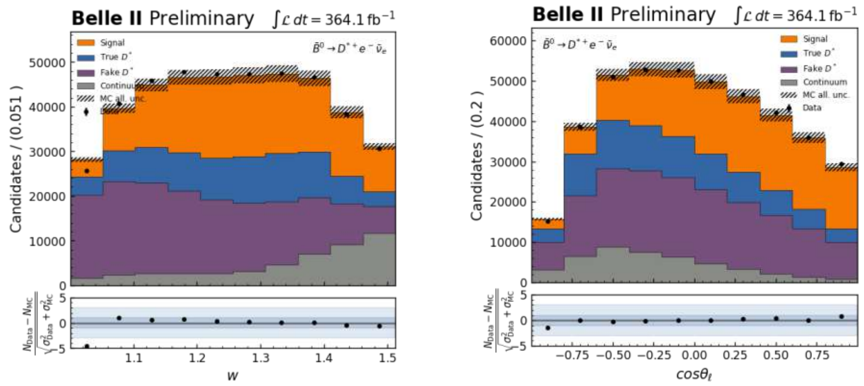
channel. Overall we can observe a good agreement between shapes. However the lepton momentum plots in Fig.6.9 highlight the necessity of the tighter muon momentum cut due to its heavily mismodeled lower momentum region.

6.4.4 $D^*\ell\nu$ Reconstruction

In this section the agreement between the reconstructed $B \to D^*\ell\nu_l$ candidates of the on-resonance sample and the MC is shown in Fig.6.10. This discussion focuses on the variables that will be directly involved in the signal extraction, i.e. $\cos\theta_{BY}$, Δm, w, $\cos\theta_\ell$. We will show the $B \to D^* e\nu_e$ figures in this section, while the $B \to D^*\mu\nu_\mu$ figures can be found in App.A.2.



(a) $\cos\theta_{BY}$ (b) Δm



(c) w (d) $\cos\theta_\ell$

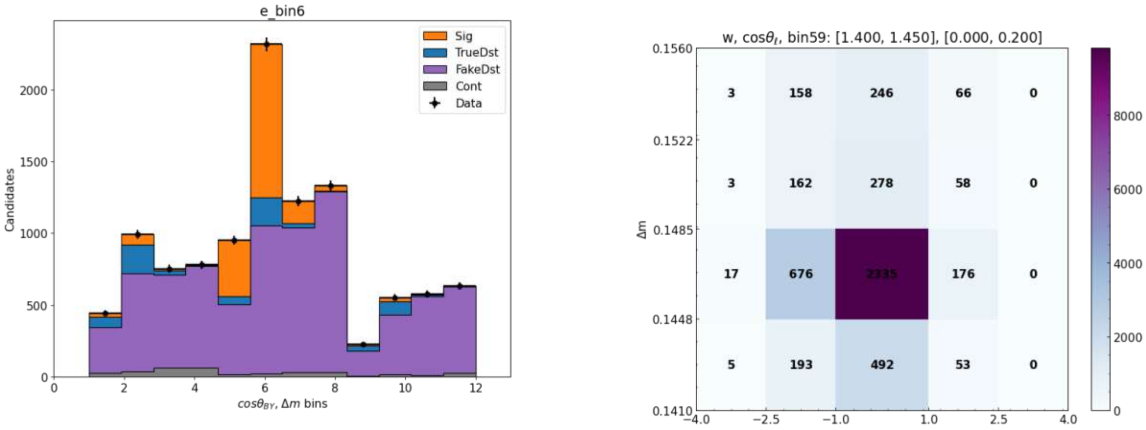
Figure 6.10: Agreement between data and MC of the distributions of the parameters used during the signal extraction for $B \to D^* e\nu_e$.

After the application of the correction weights we observe no major discrepancies between the data and the simulation. For both the electron and muon channels, the shapes of the distributions agree well and no deviations above 3σ can be observed, with exception of the lowest hadronic recoil bin. The first bin of the w distribution shows discrepancy of 3-5σ. The source of this disagreement is expected to be the high contribution of fake D^* events. However in general we don't expect a very well modeled lowest w, due to its inherent theoretical limitations of the modeling. Additionally the fake D^* component will be constrained during the signal extraction due to inclusion of the Δm distribution.

6.5 Signal Extraction

6.5.1 Procedure

To extract the number of signal events from data a binned maximum likelihood fit is performed using the `pyhf` package, which relies on template shapes. [53] More information about the concept of maximum likelihood fits and the `pyhf` library can be found in Sec.4.7.2 and Sec.4.8.1, respectively. The signal extraction will be done separately for the $B \to D^* e\nu_e$ and the $B \to D^* \mu\nu_\mu$ channel.



(a) Exemplary flattened distribution of $\cos\theta_{BY}$ and Δm to illustrate the low continuum contribution for $w \in [1.00, 1.05]$ and $\cos\theta_\ell \in [0.60, 0.80]$.

(b) A two-dimensional distribution of $\cos\theta_{BY}$ and Δm depicting the low event counts in the region $\cos\theta_{BY} < -2.5$ and $\cos\theta_{BY} > 2.5$.

Figure 6.11: Examples for the reasoning behind the choice of a fixed continuum component and the $\cos\theta_{BY}$ binning scheme.

During the signal extraction the templates are determined by the event classifications described in Sec.6.2.3: Signal, True D^*, Fake D^*, Continuum. The contribution of each event class within the data distribution is provided by the fit using a normalization factor, that scales the number of events of a template by a free parameter to best fit the data. This normalization factor is floating without constraints for the Signal, True D^* and Fake D^* component. The continuum template was chosen to be fixed using the integrated luminosity ratio of 364/1443, due to its very minor contribution. Especially in the lower w region next to no continuum events can be found, as illustrated in Fig.6.11a. The statistical uncertainty resulting from the finite size of the MC is considered by adding a nuisance parameter to the model.

In order to measure $|V_{cb}|$ we extract the partial decay rates $\Delta\Gamma/\Delta w \Delta \cos\theta_\ell$ with a simultaneous fit in bins of w and $\cos\theta_\ell$. This is done by extracting the number of signal candidates for a given $\cos\theta_\ell$ region within a w bin. For this purpose we split the events into 10 w bins, where each w bin is further subdivided into 7 $\cos\theta_\ell$ bins, resulting in overall 70 kinematic bins. A brief overview of the chosen binning scheme in w and $\cos\theta_\ell$ is provided in Tab.6.3. While an even binning scheme for w is employed, it was necessary to implement a variable binning for the lower region of $\cos\theta_\ell$, i.e. [1.00, 0.25, 0.00]. This variable binning in the region below 0 for $\cos\theta_\ell$ was needed due to the very limited number of events, as observed in Fig.6.10. Especially in the $B \rightarrow D^*\mu\nu_\mu$ channel a lower event count for $\cos\theta_\ell < 0$ can be observed, because of its tighter lepton momentum cut and its correlation to $\cos\theta_\ell$. The highest bin of w, i.e. $w \in [1.450, 1.504]$, is cut off at 1.504, which is the maximum value obtainable from B meson reconstruction with a simulated neutrino. Since the ROE frame reconstructed value w_{ROE} is only an estimate of w some deviations between w_{ROE} and w_{MC} are expected. The lost number of signal events due to the cut-off was observed to be insignificant with only 877 (0.15%) and 349 (0.07%) lost events for the $B \rightarrow D^*e\nu_e$ and $B \rightarrow D^*\mu\nu_\mu$ channel, respectively. With respect to the naming convention of the kinematic bins moving forward, $w \in [1.00, 1.05], \cos\theta_\ell \in [-1.00, -0.25]$ will be referred to as bin1,

w Bin Edges	$cos\theta_\ell$ Bin Edges
$[1.000, 1.050]$	[-1.00, -0.25, 0.00, 0.20, 0.40, 0.60, 0.80, 1.00]
$[1.050, 1.100]$	[-1.00, -0.25, 0.00, 0.20, 0.40, 0.60, 0.80, 1.00]
$\vdots$	$\vdots$
$[1.450, 1.504]$	[-1.00, -0.25, 0.00, 0.20, 0.40, 0.60, 0.80, 1.00]

Table 6.3: Overview of the kinematic binning theme of the partial decay rates.

Variable	Bin Edges	n_{bins}
w	[1.000, 1.050, 1.100, 1.150, 1.200, 1.250, 1.300, 1.350, 1.400, 1.450, 1.504]	10
$cos\theta_\ell$	[-1.00, -0.25, 0.00, 0.20, 0.40, 0.60, 0.80, 1.00]	7
$cos\theta_{BY}$	[-4, -1, 1, 4]	3
Δm	[0.141, 0.145, 0.149, 0.152, 0.156]	4

Table 6.4: Bin edges of each variable used during the signal extraction.

$w \in [1.00, 1.05], \cos\theta_\ell \in [-0.25, -0.00]$ as bin2 and so on. Splitting each w bin into its 7 $\cos\theta_\ell$ bins and then moving on to the next sub-divided w bin.

To determine the partial decay rate in each of these 70 kinematic bins the signal extracted through a two dimensional binned maximum likelihood fit in 3 variable bins of $\cos\theta_{BY}$ within [-4,4] and 4 even bins in $\Delta m \in [0.141, 0.156]$ GeV. We therefore extract the number of signal candidates in overall $10 \times 7 \times 3 \times 4 = 840$ bins. A summary of the binning scheme for each parameter is given in Tab.6.4. The binning of $\cos\theta_{BY}$ had to be chosen to be variable, i.e. [-4, -1, 1, 4], because of the very low number of events in the side-bands of [-4,-2.5] and [2.5,4]. This behavior is exhibited especially in the lower region of $\cos\theta_\ell \in [-1, 0]$ of high w bins. An example of this behavior is provided in the heat map of Fig.6.11b.

6.5.2 Asimov Test

The first closure test used to validate our signal extraction procedure is the asimov test. For this test we fit the statistical model derived from the MC onto the same MC simulated sample as pseudo-data without any event categorization. Therefore the expected result of the normalization factors of the fit equals 1 and the signal yield predicted by the fit and the known amount of signal events in the MC should be equal. This test serves the purpose to determine whether the fit procedure works as intended. Both the $B \to D^* e\nu_e$ and the $B \to D^* \mu\nu_\mu$ channel show perfect agreement between the true signal yield of the MC and the estimation by the fit. The results are shown in Fig.6.12 for the electron channel, while the muon channel result can be found in App.A.3.

6.5.3 Pull Test

The second closure test is a pull test based on the principles discussed in Sec.4.7.4. We calculated a pull distribution for 5000 toy MC samples based on Poisson resamples of the pseudo data and the MC templates for each kinematic bin. These pull distributions are fitted using an unbinned Gaussian, using its mean μ and width σ as indicators for potential biases or faulty error estimations of the fit. Due to the rather limited sample size for each kinematic bin, it was also investigated how the mean of the pull would translate to the

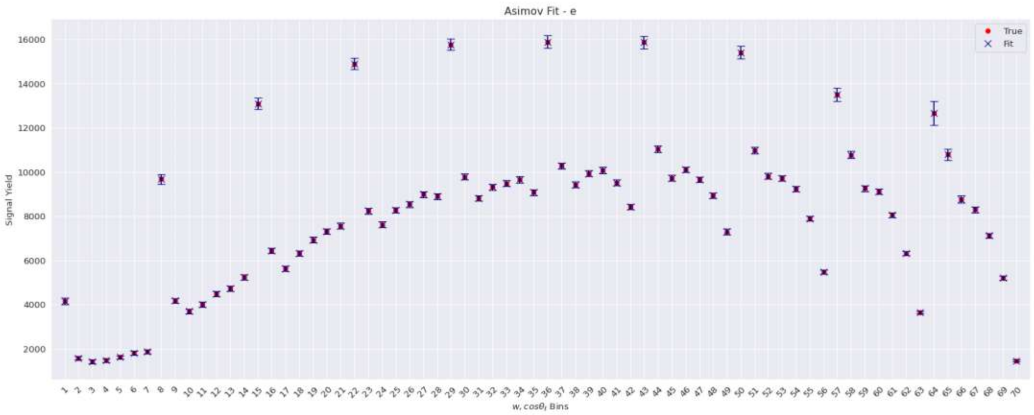


Figure 6.12: Asimov result of the $B \to D^* e\nu_e$ channel, showing perfect agreement between the known signal yield and the yield estimated by the fit. The red dots indicate the true amount of signal events and the blue crosses with their respective uncertainties show the estimation.

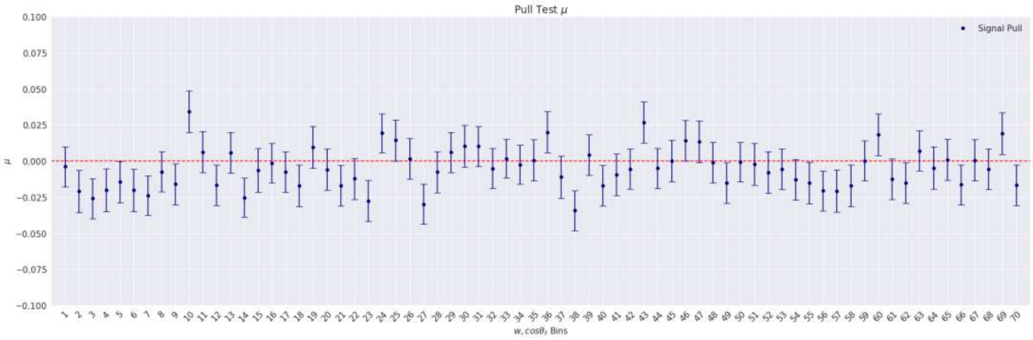
number of signal events. This was done by taking the mean of $\mu_{fit} - \mu_{true}$ over the 5000 MCs per bin. Similar to the asimov test, this closure test was also done for the $B \to D^* e\nu_e$ and $B \to D^* \mu\nu_\mu$ channel separately.

The results of this study for the mean and width of the pull, as well as the number of candidate estimation are shown for the $B \to D^* e\nu_e$ and $B \to D^* \mu\nu_\mu$ channel in Fig.6.13 and App.A.3, respectively. For neither decay channel a significant deviation from the expected value of $\mu = 0$ could be observed, with both channels mostly agreeing with 0 within 1 σ of the uncertainty of the Gaussian estimate. The study of the estimation of the mean in number of signal events also exhibits no biases, with the deviations being well below 5 events per bin and heavily centered around 0. For the error estimations the unbinned Gaussian also does not indicate any major over or underestimation of the errors by the fit, with all widths agreeing with unity within 1 - 1.5σ of their respective uncertainties.

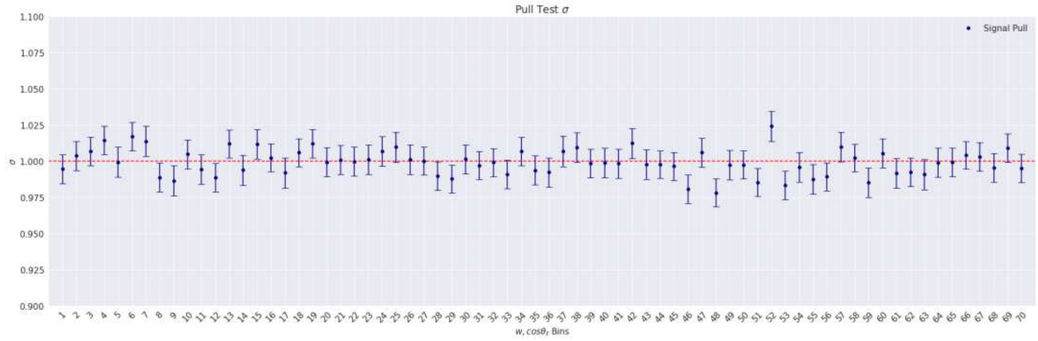
6.6 Systematic Uncertainties

All applied corrections mentioned in Sec.4.6 are determined by various analyses and have their own associated uncertainties. Furthermore various external input parameters used during the branching fraction and decay rate calculations are also determined from experiment with a given uncertainty. Thus all of the uncertainties from corrections and external input parameters will be needed to be taken into consideration as systematic uncertainties for our measurements. Additionally it is also necessary to determine the overall correlation between the systematic uncertainties of all 70 kinematic bins. For example, while the tracking efficiency of mid to high momentum tracks will affect all bins equally, the same variation of a form factor might have a different impact on bin 1 and bin 20. These correlations are also needed to average the results from $B \to D^* e\nu_e$ and $B \to D^* \mu\nu_\mu$.

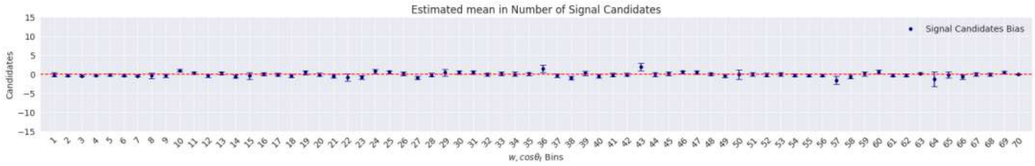
To estimate the impact on the systematic uncertainty on our measurements and also the correlation between the bins, we use a bootstrapping approach (see Sec.4.7.4). For each lepton channel we create 1000 toy MCs for every source of a systematic uncertainty and for a combination of all uncertainties. Each event has an associated correction weight for



(a) Summary of the bias estimation.



(b) Summary of the error estimation



(c) Estimated mean in number of signal candidates.

Figure 6.13: Pull result summary of $B \to D^{*} e \nu_e$ for all kinematic bins.

every calibration, when multiplying all weights, one obtains the overall correction factor. We create the toy MCs for this study not by directly varying the amount of events in each bin, but by varying the corrections weights according to their associated uncertainty. The fitting procedure is then done for each variation, ultimately obtaining 1000x70 signal yields for each source. In order to obtain the systematic correlation from this study, all variations are applied in the same way for all bins. This means, if a variation of +5% is pulled for a certain weight, then all 70 kinematic bins for both the $B \to D^{*} e \nu_e$ and $B \to D^{*} \mu \nu_\mu$ channel will be varied up by +5%.

We estimate the impact of each source on the branching fraction $\mathcal{B}(\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l)$ and partial decay rate$\Delta\Gamma/\Delta w \Delta \cos\theta_\ell$ by calculating them for each toy. The branching fraction $\mathcal{B}(\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l)$ is defined as follows:

$$\mathcal{B}(\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l) = \frac{N_{sig}}{N_{B^0} \mathcal{B}(D^{*+} \to D^0 \pi^+) \mathcal{B}(D^0 \to K^- \pi^+) \epsilon}, \tag{6.4}$$

where N_{sig} is the signal yield estimated by the fit, N_{B^0} is the number of neutral B mesons in the sample and $\mathcal{B}(D^{*+} \to D^0 \pi^+)$ and $\mathcal{B}(D^0 \to K^- \pi^+)$ are the world average branching fractions of the sub decays taken from PDG.[10] Lastly, the equation also includes the efficiency ϵ. The branching $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$ is first calculated for each kinematic bin individually and then summed up while taking systematic correlation into account to obtain the overall branching fraction. The standard deviations and mean of the branching fractions of the 1000 variations for each kinematic bin are used to determine the relative uncertainties of each source.

Through the efficiency we take any signal candidates that might have been lost during the analysis process into account, whether due to signal selection criteria or detector effects. During the calculation of the efficiency we also unfold (see Sec.4.7.6) our signal yields by taking bin by bin migrations into account through bin by bin unfolding. This is done by comparing the reconstructed amount of events within $w_{ROE}, \cos\theta_{\ell_{ROE}}$ bins with the event count provided by the generator in bins of $w_{MC}, \cos\theta_{\ell_{MC}}$ with equal bin edges:

$$\epsilon = \frac{N_{rec}}{N_{MC}}, \tag{6.5}$$

where N_{rec} is the number of reconstructed events in a ROE bin e.g. $w_{ROE} \in [1.00, 1.05], \cos\theta_{\ell_{ROE}} \in [-1.00, -0.25]$ and N_{MC} is the number of generated events within the bin with the true values $w_{MC} \in [1.00, 1.05], \cos\theta_{\ell_{MC}} \in [-1.00, -0.25]$.

The partial decay rates exhibit the following relation to $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$:

$$\frac{\Delta\Gamma_i}{\Delta w_i \Delta cos\theta_{\ell_i}} = \frac{\Delta\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)_i}{\tau_{B^0}\Delta w_i \Delta cos\theta_{\ell_i}}, i \in [1, 70], \tag{6.6}$$

where τ_{B^0} is the B^0 life time and Δw_i and $\Delta cos\theta_{\ell_i}$ the width of a given kinematic bin i. The determination of the relative error of the various sources of systematic uncertainties are determined in the same way as for the branching fractions.

6.6.1 Tracking

An additive per track uncertainty of the data/MC efficiency of mid to high momentum tracks is provided by the tracking and vertexing group. The recommendation corresponds to a relative uncertainty 0.27% per charged track. Thus we have to consider an relative tracking uncertainty of 0.81% due to three charged final state particle being part of the $B \to D^*\ell\nu_l$ decay chain, namely a pion, a kaon and a lepton. The slow pion involved in the $D^* \to D\pi$ decays has its own dedicated systematic. We estimate the tracking uncertainty by adding an additional weight to our events and drawing a random number from a Gaussian distribution with a mean of 1 and width of 0.0081.

6.6.2 $\cos\theta_{BY}$ Background Shapes

Another uncertainty arises from the reweighing of the fake and true D^* background shapes of $\cos\theta_{BY}$. To determine the background shape uncertainty we vary the correction weight $w_{\cos\theta_{BY}}$ within the region of $[-w_{\cos\theta_{BY}}, w_{\cos\theta_{BY}}]$. This is achieved by calculating the varied weight using $w_{\cos\theta_{BY}} + x_{rand}(w_{\cos\theta_{BY}} - 1)$, where x_{rand} is random number drawn from a Gaussian distribution with mean 0 and width 1. All weights for the 30 $\cos\theta_{BY}$ shape weights are varied in this manner.

6.6.3 Continuum Normalization

The continuum shape is fixed to the integrated luminosity ratio between data and MC during the fit. However the integrated luminosity of the data is estimated through Bhabha events and not an exact knowledge without uncertainties. Therefore we vary the luminosity scaling within the provided uncertainty in Sec.4.2.

6.6.4 LeptonID

The systematic uncertainty arising from the leptonID weights is based on the statistical and systematic errors provided with the tables for each p, θ bin of the leptons. We generate 200 event wise variations:

$$w_i' = w_i + \mathcal{G}(0, \sigma_{stat}^T \rho_{stat} \sigma_{stat}) + \mathcal{G}(0, \sigma_{sys}^T \rho_{sys} \sigma_{sys}), \tag{6.7}$$

where w_i' is the variation of a weight w_i and $\mathcal{G}(0, \sigma^T \rho \sigma)$ are random numbers based on a multivariate Gaussian with the correlation ρ and uncertainties σ. For the generated weights we assume full correlation over all bins for the systematic uncertainties and no correlation for the statistical uncertainties.

6.6.5 Slow Pion Efficiency

To estimate the uncertainty from the slow pion tracking efficiency we use a similar approach to the leptonID. We generate 200 variations of the slow pion weights through the following approach:

$$w_i' = w_i + \mathcal{G}(0, \sigma_{uncorr}) + \mathcal{G}(0, \sigma_{corr}) + \mathcal{G}(0, \sigma_{track}), \tag{6.8}$$

where $\mathcal{G}(0, \sigma)$ are random numbers drawn from a Gaussian distribution. The correlation between bins is taken into consideration by drawing the random numbers $\mathcal{G}(0, \sigma_{corr})$ and $\mathcal{G}(0, \sigma_{track})$ from the same random seed for each momentum bin. Random numbers from uncorrelated uncertainties are drawn with a different seed.

6.6.6 Form Factors

The uncertainties arising from updating the form factors of the decay file are estimated by generating 200 BGL models based on the updated form factor values. We create the 200 models through eigenvariations of the original form factors to take correlations between form factors into account. These eigenvariations are based on the uncertainties and correlations provided with the results of the measurement from Ref.[59]. For each of these variations new event-wise form factor weights will be calculated and a random number between 1 and 200 will be drawn to determine which set of variations will be used for a toy MC.

6.6.7 $X_c \ell \nu$ Branching Fractions

We estimate the uncertainties of the updated $X_c \ell \nu$ branching fractions by varying the updated branching fractions using a random Gaussian number based on their respective uncertainties provided in Tab.6.2.

6.6.8 Charm Branching Fractions

The impact of the uncertainties of the world averaged values of $\mathcal{B}(D^{*+} \to D^0\pi^+)$ and $\mathcal{B}(D^0 \to K^-\pi^+)$ used to determine the branching fraction $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$ have to be taken into consideration. The uncertainty arising from these input parameters is estimated through a Gaussian variation of $\mathcal{B}(D^{*+} \to D^0\pi^+)$ and $\mathcal{B}(D^0 \to K^-\pi^+)$ within their respective uncertainties and calculating a set of 1000 $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$ results using the varied charm branching fractions.

6.6.9 $B\bar{B}$ Counting

In order to determine $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$ the knowledge of the number of neutral B meson pairs is necessary. The amount of overall $B\bar{B}$ pairs $n_{B\bar{B}} = (387 \pm 6) \cdot 10^6$ is provided by the collaboration and determined via Bhabha decays. However the counted B meson pairs $n_{B\bar{B}}$ include both charged and neutral pairs. We determine the amount of B^0 mesons in the sample from the overall $B\bar{B}$ pairs as follows:

$$N_{B^0} = \frac{2 \cdot n_{B\bar{B}}}{1 + R^{+-/00}}, \tag{6.9}$$

where $R^{+-/00}$ the following ratio of $\Upsilon(4S)$ decay rates, measured by Belle [60]:

$$R^{+-/00} = \frac{f^{+-}}{f^{00}} = \frac{\Gamma(\Upsilon(4S) \to B^+B^-)}{\Gamma(\Upsilon(4S) \to B^0\bar{B}^0)} = 1.065 \pm 0.052. \tag{6.10}$$

Similar to the charm branching fraction uncertainty estimate, we calculate 1000 branching fractions with a Gaussian variation of the number of B^0 mesons.

6.6.10 B Lifetime

To determine the decay rates from $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$ knowledge of the B^0 meson life time τ_{B^0} is necessary. This external parameter is provided by the PDG [10] and varied within its associated uncertainty. Currently the world average $\tau_{B^0} = 1.519 \pm 0.004$ ps has a relative uncertainty of $\sim 0.3\%$.

6.6.11 Unfolding

The bin by bin migration of events due to the limited resolution of the reconstructed kinematic variables is taken into consideration by implementing bin by bin unfolding through the efficiency. The uncertainty arising from the unfolding procedure is partially taken into account by the form factor uncertainty. When altering the form factor we also change the number of events per kinematic bin, due to the weights being dependent on the kinematic variables. Furthermore we also have to take statistical fluctuation of the MC events into account due to its finite sample size. We achieve this by varying the amount of MC signal candidates in the true kinematic bins within their statistical uncertainty and calculating varied efficiencies.

6.6.12 Correlation Matrix

In order to sum up the branching fractions of $\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e$ and $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ of the kinematic bins and average the branching fractions and decay rates over both channels it

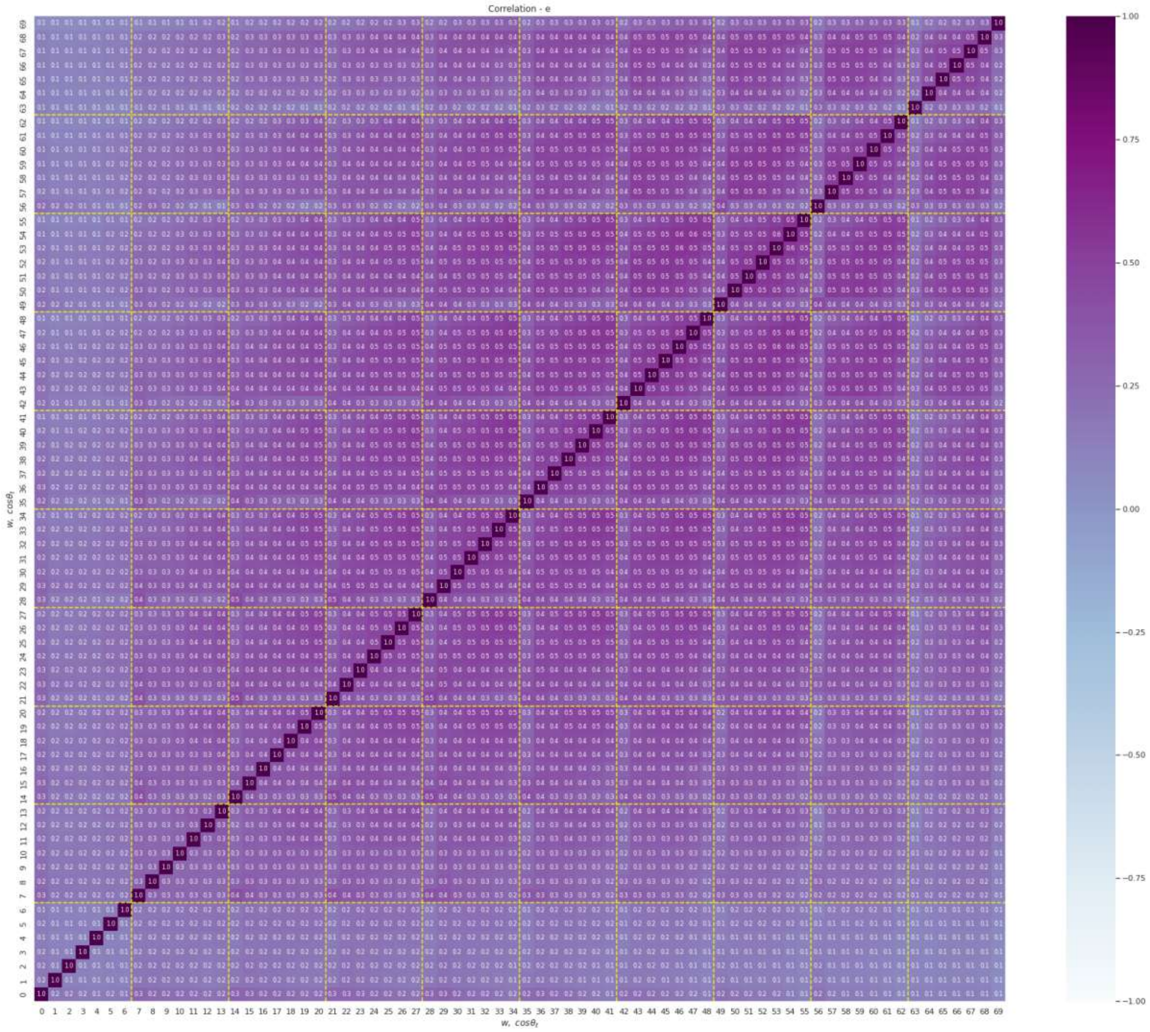


Figure 6.14: Statistical and systematic correlation matrix of $B \to D^* e \nu_e$.

is necessary to determine the systematic correlation between the kinematic bins of each channel and between both channels. The statistical correlation between the bins and the channels is equal to a diagonal matrix due to the signal extraction being independent for each kinematic bin. The full 70x70 correlation matrices including systematic and statistical correlations for the partial decay rates is provided in Fig.6.14 and App.A.4 for $B \to D^* e \nu_e$ and $B \to D^* \mu \nu_\mu$, respectively.

6.6.13 Estimation Results

The resulting estimated systematic and statistical uncertainties of each source for the branching fractions are provided in Tab.6.5. We provide a summary in the form of branching fraction instead of decay rates, due to the unwieldy nature of presenting three times 70 partial decay rates. By simply showing off the uncertainties on the determination of $\mathcal{B}(\overline{B}^0 \to D^{*+} e^- \overline{\nu}_e)$ and $\mathcal{B}(\overline{B}^0 \to D^{*+} \mu^- \overline{\nu}_\mu)$ separately as well as their average $\mathcal{B}(\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l)$, it is simpler to provide an overview of magnitude of each systematic source and their sum.

As we can observe from the summary table, the leading systematic by far is the external

Source	$\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e$ [%]	$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ [%]	$\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ [%]
Tracking	0.80	0.80	0.80
$cos\theta_{BY}$ BG Shape	0.42	1.20	0.81
LepID Efficiency	0.31	0.15	0.17
LepID Fake Rate	0.01	0.03	0.01
Slow Pions	1.35	1.38	1.37
$X_c\ell\nu$	0.13	0.19	0.11
Form Factors	0.15	0.13	0.14
$R^{+-/00}$	2.60	2.60	2.60
n_{BB}	1.41	1.41	1.40
Continuum Norm.	0.02	0.13	0.08
Unfolding	0.07	0.07	0.05
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	0.74	0.74	0.73
$\mathcal{B}(D^0 \to K^-\pi^+)$	0.73	0.73	0.74
MC stat.	0.27	0.44	0.26
Data stat.	0.54	0.87	0.51
All Sys	3.66	3.89	3.73
All Stat	0.61	0.98	0.57
Total	3.71	4.01	3.77

Table 6.5: Summary of the systematic and statistical uncertainties of each source and their sum. The results are provided for each lepton channel individually and their average $B \to D^*\ell\nu_l$

Channel	Fit Result	Generator Value
$\mathcal{B}(\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e)$	$(5.099 \pm 0.031_{stat} \pm 0.187_{sys})\%$	5.110%
$\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$	$(5.091 \pm 0.050_{stat} \pm 0.198_{sys})\%$	5.110%
$\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$	$(5.095 \pm 0.029_{stat} \pm 0.190_{sys})\%$	5.110%

Table 6.6: Comparison between the asimov branching fraction results and the generator value.

input of the $\Upsilon(4S)$ decay rate ratio $R^{+-/00}$ with 2.6%. This is followed by the slow pion efficiency and the $n_{B\bar{B}}$ counting, with 1.37% and 1.40%, respectively. The difference of the statistical uncertainties between $B \to D^*e\nu_e$ and $B \to D^*\mu\nu_\mu$ are expected, due to the higher cut on the lepton momentum for the muon channel. From this study we can already conclude, that the uncertainty of our measurement is heavily dominated by the systematic uncertainties and not statistically limited. Especially a more precise determination of the ratio $R^{+-/00}$ will be needed to improve the precision of future measurements. Additionally we provide a comparison between our asimov result of the branching fractions with the value of the MC generator in Tab.6.6, which easily agree with one another within 1σ.

6.7 $|V_{cb}|$ Extraction

The following relation between the decay rate and the CKM magnitude $|V_{cb}|$ is used to determine $|V_{cb}|$ and the form factors:

$$\begin{aligned}\frac{d^2\Gamma(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)}{dw\ dcos\theta_\ell} =& \frac{6m_B m_{D*}^2}{8(4\pi)^4}\eta_{ew}^2 G_F^2 |V_{cb}|^2 \sqrt{w^2-1}(1-2wr+r^2) \\ & [(1-cos\theta_\ell)^2 H_+^2(w) + (1+cos\theta_\ell)^2 H_-^2(w) + \\ & + 2(1-cos^2\theta_\ell)H_0^2(w)]\end{aligned} \tag{6.11}$$

Further details on the origins of this relation, the variables and the parameterizations of the form factors can be found in Sec.2.6.1. We fit $|V_{cb}|$ and the form factors for both the CLN [28] and the BGL [26, 27] parameterization onto the partial decay rates obtained from the simulation. Then we estimate the form factors and $|V_{cb}|$ by minimizing the following χ^2 function:

$$\chi^2 = (\vec{N}^{\,\text{obs}} - \vec{N}^{\,\text{exp}})C^{-1}(\vec{N}^{\,\text{obs}} - \vec{N}^{\,\text{exp}}), \tag{6.12}$$

where $\vec{N}^{\,\text{obs}}$ are the decay rates obtained from the data, $\vec{N}^{\,\text{exp}}$ their expected values based on a set of CLN or BGL parameters and the full covariance matrix C. Additionally we implement the recent beyond zero recoil LQCD results published by the FNAL/MILC collaborations in Ref.[61] for both parameterizations. To take the LQCD input into consideration a modification of the χ^2 function is necessary:

$$\chi^2 = (\vec{N}^{\,\text{obs}} - \vec{N}^{\,\text{exp}})C^{-1}(\vec{N}^{\,\text{obs}} - \vec{N}^{\,\text{exp}}) + (\vec{T}^{\,\text{obs}} - \vec{T}^{\,\text{exp}})U^{-1}(\vec{T}^{\,\text{obs}} - \vec{T}^{\,\text{exp}}), \tag{6.13}$$

where $\vec{T}^{\,\text{exp}}$ are the form factor values provided by the LQCD input at a given set of w values, $\vec{T}^{\,\text{obs}}$ the form factor values based on the observed decay rates and the correlation U^{-1} between the form factors from the LQCD measurements.

For both CLN and BGL we executed the same closure tests for the fit, that were mentioned during the signal extraction in Sec.6.5, i.e. an asimov and a pull test.

6.7.1 BGL Parameterization

To determine $|V_{cb}|$ and the form factors through the BGL parameterization it is necessary to truncate the BGL expansion. It was chosen to truncate the expansions in Eq.2.29-Eq.2.31 at n=2, since it was proven to be sufficient in Ref.[62]. This results in 3 parameters for each expansion, thus this approach is referred to as BGL$_{333}$. The parameter c_0 is not fitted directly, but determined by the following relation to b_0:

$$c_0 = \frac{(1-r)\cdot b_0}{\sqrt{2}\cdot(1+\sqrt{r})^2}, \tag{6.14}$$

where r is the ratio of the mass of the D^* over the B meson

The results of the asimov fit for all fitted parameters can be found in Tab.6.7. In Fig.6.15 we present the agreement between the measured partial decay rates in bins of $w, \cos\theta_\ell$ and the curve obtained from the fitted parameters. Figures for the individual w bins can be found in App.A.5. Overall the shape of the curve agrees well with the measured decay rates with all data points agreeing with the curve within 1σ. That being said, we can observe some minor disagreement between the measured decay rates and the curve. This

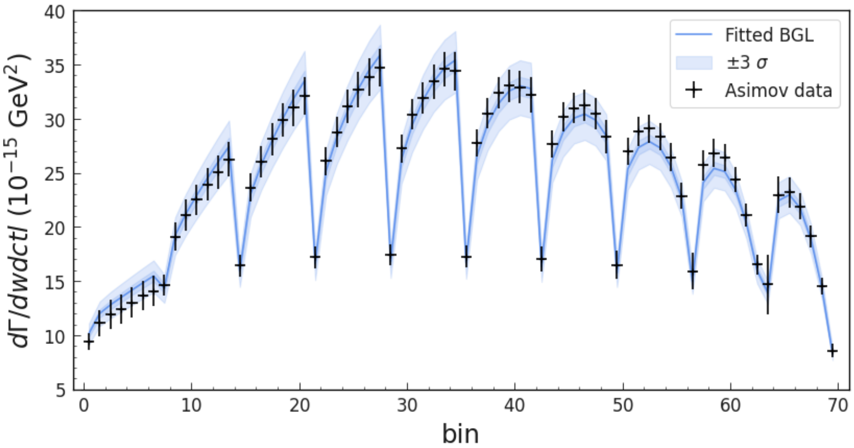


Figure 6.15: Comparison between the partial decay rates obtained from the MC with the shape of the curve determined by the χ^2 fit using the BGL parameterization.

$\lvert V_{cb}\rvert$	$(38.100 \pm 1.000) \times 10^{-3}$
a_0	0.032 ± 0.001
a_1	-0.084 ± 0.064
a_2	-1.687 ± 1.847
b_0	0.012 ± 0.001
b_1	0.031 ± 0.011
b_2	-0.015 ± 0.385
c_1	0.005 ± 0.002
c_2	-0.054 ± 0.032
χ^2/n_{dof}	50.16 / 70

Table 6.7: Results obtained from the asimov data through a BGL fit.

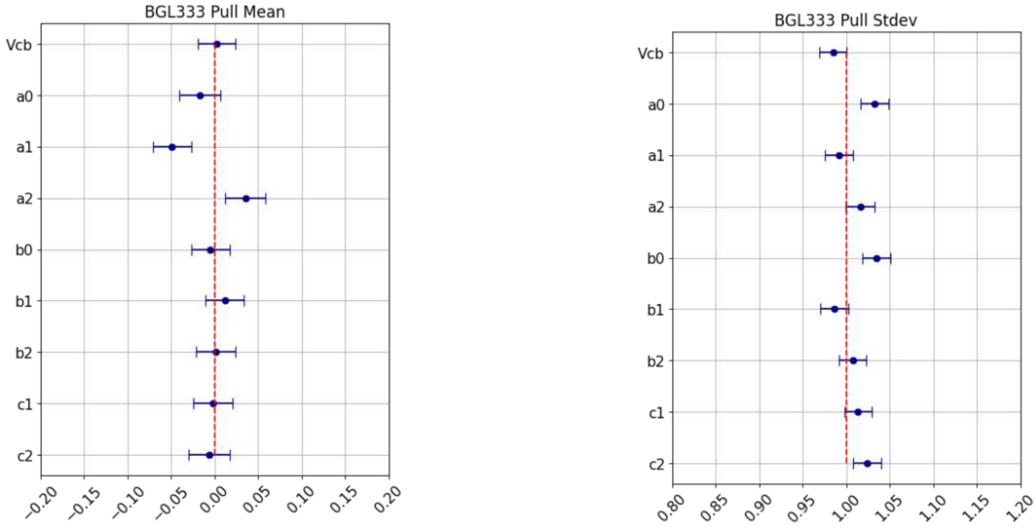
behaviour originates from the different BGL$_{121}$ truncation in the generator and additionally the generator doesn't use any inputs from LQCD results beyond zero recoil.

The pull test was done as described in Sec.4.7.4 with the results from the asimov fit acting as the expected value $\hat{\mu}$. To obtain the fitted values μ and σ we resample the averaged the decay rates through a multivariate Gaussian based on the full covariance matrix and additionally repeat the same procedure for the theory input. Through this approach we obtain the mean and width illustrated in Fig.6.16a and Fig.6.16b, respectively. From the obtained pull results, we can conclude that our BGL fit is unbiased and the errors for all parameters well estimated.

6.7.2 CLN Parameterization

When using the CLN parameterization, the fit doesn't directly provide $|V_{cb}|$, but instead $\mathcal{F}(1)|V_{cb}|$, where $\mathcal{F}(1)$ corresponds to the form factor normalization at zero recoil $\mathcal{F}(1) = 0.910 \pm 0.013$, provided by Ref.[63], which is estimated theoretically.

For the CLN form factors and $|V_{cb}|$ we obtain the asimov results listed in Tab.6.8. When comparing the MC result of $|V_{cb}|$ from the CLN fit with the BGL fit, we find that both results have an agreement within 1σ of each other. The comparison between the shape



(a) Summary of the bias estimation. (b) Summary of the error estimation

Figure 6.16: Results obtained from the pull test of the $|V_{cb}|$ and form factor fit using the BGL$_{333}$ parameterization.

$\mathcal{F}(1)\lvert V_{cb}\rvert$	$(35.110 \pm 0.780) \times 10^{-3}$
ρ^2	1.200 ± 0.070
$R_1(1)$	1.214 ± 0.088
$R_2(1)$	0.797 ± 0.067
χ^2/n_{dof}	69.95 / 69

Table 6.8: Results obtained from the asimov data through a CLN fit.

obtained from the fitted parameters and measured partial decay rates from the MC is provided in Fig.6.17. For the CLN fit we also don't expect a perfect agreement between the decay rates and the shape obtained from the fit, due to the different parameterization of the MC generator. Nevertheless, the decay rates of all bins agree with the fitted shape within 1σ. The shapes of the individual w bins are provided in App.A.6.

We studied the CLN fit for biases using a pull test with the same approach mentioned for the BGL parameterization. The resulting mean and width of the pull can be found in Fig.6.18a and Fig.6.18b, respectively. From the results of the pull, we don't expect any biases by the CLN fit. However, the uncertainties on the form factors seem to be underestimated by up to 20%, thus we will scale up these uncertainties.

6.7.3 Sensitivity

To estimate the achievable precision of the form factor and $|V_{cb}|$ measurements for our analysis setup we use a bootstrapping approach similar to the systematic uncertainty estimate of the branching fractions. We create a random variation of the averaged decay rates by drawing new decay rates based on a multivariate Gaussian using the covariance matrix obtained from the previous study on the impact of the individual systematic uncertainty sources on the partial decay rates and branching fractions.

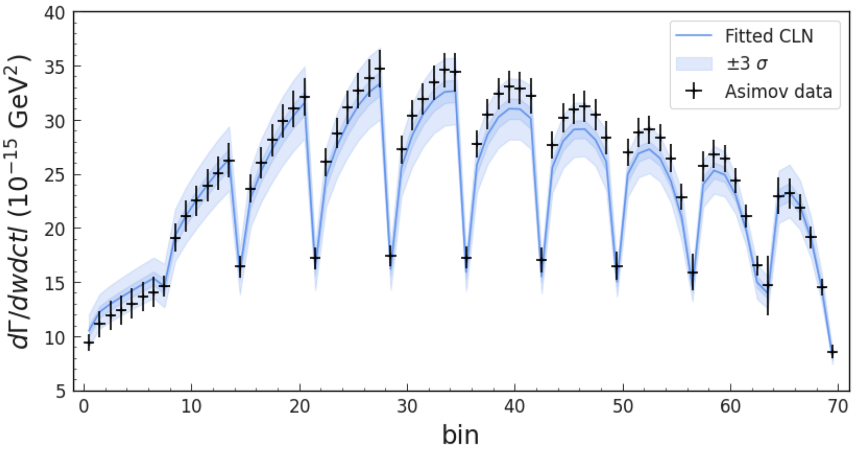
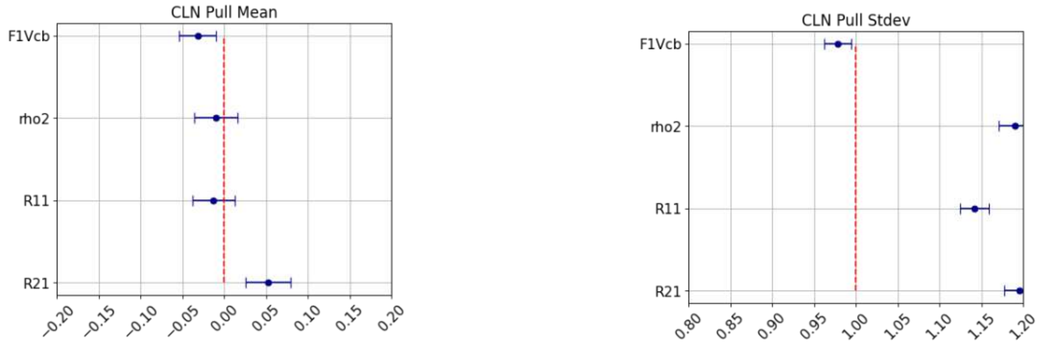


Figure 6.17: Comparison between the partial decay rates obtained from the MC with the shape of the curve determined by the χ^2 fit using the CLN parameterization



(a) Summary of the bias estimation.

(b) Summary of the error estimation

Figure 6.18: Results obtained from the pull test of the $\mathcal{F}(1)|V_{cb}|$ and form factor fit using the CLN parameterization.

Using this approach we obtain the uncertainty estimates in Tab.6.9 for the BGL parameterization. We obtain a split of the uncertainties for BGL of $|V_{cb}|_{\mathrm{BGL}} = (38.100 \pm 0.194 \pm 0.754 \pm 0.674) \times 10^{-3}$ from the simulation with an overall uncertainty of 2.62%. The uncertainties in order of appearance are: statistical, systematic and theoretical.

The results of the CLN study are provided in Tab.6.10. However this summary does not yet include the theoretical uncertainty arising from the form factor normalization at zero recoil $\mathcal{F}(1) = 0.910 \pm 0.013$ provided by Ref.[63]. For the sensitivity study of the CLN parameterization we can observe the same tendency shown in the pull: when using the complete covariance matrix for the fit we obtain an uncertainty $\sim 20\%$ lower than the sum of the individual sources for the form factors. Thus we opt for choosing the sum of the individual sources as our total uncertainty instead of the error estimated by the fit from the total covariance matrix. This way we obtain $\mathcal{F}(1)|V_{cb}|_{\mathrm{CLN}} = (35.110 \pm 0.225 \pm 0.769 \pm 0.095) \times 10^{-3}$ from the simulation. By inserting the value of $\mathcal{F}(1)$ from Ref.[63] and taking its uncertainty into consideration we obtain: $|V_{cb}|_{\mathrm{CLN}} = (38.582 \pm 0.247 \pm 0.844 \pm 0.560) \times 10^{-3}$. The uncertainty from $\mathcal{F}(1)$ contributes another relative uncertainty of 1.43% to the theoretical part, thus we obtain a precision of 2.70% for the CLN parameterization.

Source	$\lvert V_{cb}\rvert$	a_0	a_1	a_2	b_0	b_1	b_2	c_1	c_2
Tracking	0.36	0.03	0.26	1.24	0.01	0.83	34.22	1.63	2.16
$cos\theta_{BY}$ BG Shape	0.15	0.53	5.18	24.32	0.02	6.09	847.70	9.75	14.02
LepID Efficiency	0.07	0.08	2.15	9.00	0.02	1.15	163.52	2.47	4.18
LepID Fake Rate	0.01	0.02	0.88	2.45	0.01	0.33	39.71	0.65	1.09
Slow Pions	0.83	0.08	5.15	3.67	0.07	4.02	201.95	5.99	9.37
$X_c\ell\nu$	0.10	0.18	6.50	13.80	0.05	2.86	272.79	4.72	8.48
Form Factors	0.54	1.20	19.14	21.65	0.33	4.67	955.29	6.22	14.96
$R^{+-/00}$	1.24	0.10	0.84	4.04	0.03	2.88	137.56	5.59	7.33
n_{BB}	0.70	0.05	0.47	2.28	0.01	1.60	67.15	3.12	4.09
Continuum Norm.	0.01	0.02	0.68	3.02	0.01	0.65	70.78	1.16	2.17
Unfolding	0.07	0.09	5.17	10.55	0.04	2.08	155.57	3.60	6.19
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	0.35	0.03	0.22	1.13	0.01	0.81	34.01	1.58	2.07
$\mathcal{B}(D^0 \to K^-\pi^+)$	0.34	0.03	0.22	1.11	0.01	0.78	32.13	1.53	2.02
τ_{B^0}	0.12	0.01	0.08	0.41	0.01	0.28	11.87	0.56	0.73
MC stat.	0.23	0.33	14.06	30.05	0.12	6.42	445.89	10.35	17.18
Data stat.	0.48	0.65	29.91	60.47	0.26	13.34	914.56	21.35	35.76
All Sys	1.98	1.28	21.23	36.72	0.34	9.92	740.11	14.70	24.56
All Stat	0.51	0.74	33.50	68.42	0.29	14.47	1975.36	23.14	37.84
Theory	1.77	3.28	64.62	80.74	1.80	28.42	1350.55	27.99	36.93
Total	2.62	3.12	76.19	109.48	8.33	35.48	2566.66	40.00	59.25

Table 6.9: Summary of the systematic, statistical and theoretical uncertainties of each source and their sum for the BGL parameters of $B \to D^*\ell\nu_l$. The uncertainties are relative and given in %.

6.8 Results

The study of the sensitivity of an $|V_{cb}|$ extraction from partial decay rates $\Delta\Gamma/\Delta w\Delta\cos\theta_\ell$ yielded:

$$\begin{aligned} |V_{cb}|_{\mathrm{BGL}} &= (38.100 \pm 0.194 \pm 0.754 \pm 0.674) \times 10^{-3}, \\ |V_{cb}|_{\mathrm{CLN}} &= (38.582 \pm 0.247 \pm 0.844 \pm 0.560) \times 10^{-3}. \end{aligned} \tag{6.15}$$

The first uncertainty is statistical, the second is systematic and the third uncertainty is arising from the theoretical beyond zero recoil LQCD input. Additionally the uncertainty from the theoretical zero recoil normalization $\mathcal{F}(1)$ from Ref.[63] is also incorporated in the theoretical uncertainty of $|V_{cb}|_{\mathrm{CLN}}$. The results of the form factors obtained by the fits are provided in Tab.6.7 and Tab.6.8 for BGL and CLN, respectively. For both parameterization we used beyond zero recoil LQCD results provided by Ref.[61]. From the sensitivity studies we can assume, that the novel approach of measuring $|V_{cb}|$ from partial decay rates binned in both w and $\cos\theta_\ell$ simultaneously can achieve a precision comparable to recent $|V_{cb}|$ studies at Belle II. In Ref.[64] a Belle II sample corresponding to 189 fb^{-1} was studied for the same $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decay chain. The Measurement of $|V_{cb}|$ from Ref.[64] yielded a relative uncertainty of 2.85% for $|V_{cb}|$ for BGL and 2.81% for CLN in comparison to our estimated sensitivity of 2.62% and 2.70%.

From the simulation the following branching fractions were obtained for the individual

Source	$\mathcal{F}(1)\lvert V_{cb}\rvert$	ρ^2	$R_1(1)$	$R_2(1)$
Tracking	0.38	0.09	0.06	0.11
$cos\theta_{BY}$ BG Shape	0.23	1.08	2.62	2.00
LepID Efficiency	0.07	0.06	0.37	0.10
LepID Fake Rate	0.01	0.05	0.08	0.07
Slow Pions	0.93	0.87	0.29	0.59
$X_c\ell\nu$	0.14	0.67	1.07	0.98
Form Factors	0.85	4.39	6.53	6.98
$R^{+-/00}$	1.28	0.31	0.20	0.37
n_{BB}	0.74	0.18	0.11	0.21
Continuum Norm.	0.01	0.01	0.07	0.01
Unfolding	0.08	0.53	0.50	0.80
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	0.37	0.09	0.06	0.11
$\mathcal{B}(D^0 \to K^-\pi^+)$	0.35	0.08	0.05	0.10
τ_{B^0}	0.12	0.03	0.02	0.37
MC stat.	0.29	1.54	1.87	2.27
Data stat.	0.58	3.23	3.61	4.67
All Sys	2.19	4.42	7.02	7.03
All Stat	0.64	3.43	3.95	5.03
Theory	0.27	5.11	3.34	6.08
Total	2.22	7.58	8.72	10.57

Table 6.10: Summary of the systematic, statistical and theoretical uncertainties of each source and their sum for the CLN parameters of $B \to D^*\ell\nu_l$. The uncertainties are relative and given in %. This table does not include uncertainty arising from $\mathcal{F}(1)$, which contributes another 1.43% to the theoretical uncertainty of $|V_{cb}|$.

lepton channels:

$$\begin{aligned} \mathcal{B}(\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e) &= (5.099 \pm 0.031_{stat} \pm 0.187_{sys})\%, \\ \mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu) &= (5.091 \pm 0.050_{stat} \pm 0.198_{sys})\%. \end{aligned} \tag{6.16}$$

Which results in the following branching fraction ratio:

$$R = \frac{\mathcal{B}(\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e)}{\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)} = 1.002 \pm 0.012_{stat} \pm 0.036_{sys}, \tag{6.17}$$

in agreement with the SM prediction. The combination of both channels yielded:

$$\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l) = (5.095 \pm 0.029_{stat} \pm 0.190_{sys})\%. \tag{6.18}$$

The whole analysis chain was done blinded, i.e. the data was not used while establishing the signal and $|V_{cb}|$ extraction strategies. By withholding the data and using the MC as pseudo-data instead we eliminate potential biases for the outcome of our study. Due to the policies enforced by the Belle II collaboration it is not allowed to apply the extraction strategy on the data until a rigorous peer review of the analysis chain and the results from the simulation was done. At the time of writing this study was not peer reviewed by the collaboration and therefore only the sensitivity tests can be provided as a result.

CHAPTER 7

Conclusion

We studied the achievable precision for a novel approach to extract $|V_{cb}|$ and its form factors from the Belle II data with a sample size corresponding to 364 fb^{-1}. In this approach we extract $|V_{cb}|$ through the partial decay rates $\Delta\Gamma/\Delta w \Delta\cos\theta_\ell$ of semileptonic $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays. These decays are reconstructed through the subsequent decays $D^{*+} \to D^0\pi^+$ and $D^0 \to K^-\pi^+$. All conducted sensitivity studies were done blinded, i.e. the data was withheld to eliminate possible biases, and are based on MC simulations. The extraction of $|V_{cb}|$ was done for both the BGL parameterization formulated by Boyd, Grinstein and Lebed [26, 27] and the CLN parameterization proposed by Caprini, Lellouch and Neubert [28]. For both parameterizations we incorporated the beyond zero recoil LQCD results provided by the FNAL/MILC collaborations in Ref.[61] into our fit.

Using the proposed approach we measure the following values for $|V_{cb}|$ from the Belle II MC simulations for the BGL and CLN parameterization, respectively:

$$\begin{aligned} |V_{cb}|_{\text{BGL}} &= (38.100 \pm 0.194 \pm 0.754 \pm 0.674) \times 10^{-3}, \\ |V_{cb}|_{\text{CLN}} &= (38.582 \pm 0.247 \pm 0.844 \pm 0.560) \times 10^{-3}, \end{aligned} \tag{7.1}$$

where the uncertainties are statistical, systematic and theoretical. For both approaches we find that the systematic uncertainty is leading. However, the uncertainty arising from theory is only ~ 11% smaller than the systematic. Our studies show that the main three sources contributing to the systematic uncertainty of ~ 2% that need improvement to achieve a better precision in the future in descending order are:

- The $\Upsilon(4S)$ decay rate ratio to charged and neutral B mesons $R^{+-/00}$ provided by Belle in Ref.[60] contributes 1.24%.
- The calibration of the slow pion tracking efficiency derived from our study in Ch.5 through $B^0 \to D^{*-}\pi^+$ decays yields an uncertainty of 0.83%. At the time of writing this uncertainty is mainly driven by the statistical limitation of $B^0 \to D^{*-}\pi^+$ study.
- Finally, the counting of $n_{B\bar{B}}$ events provided by the Belle II collaboration contributes 0.70% to the systematic uncertainty.

To determine whether the precision of our approach to measure $|V_{cb}|$ is competitive, we compare the results of our sensitivity study to the recently published $|V_{cb}|$ measurement of Belle II in Ref.[64], which uses a data sample with a size corresponding to 189 fb^{-1}. The sensitivity test for our analysis yielded 2.62% and 2.70% for BGL and CLN, respectively. Compared to 2.85% and 2.81% from Ref.[64], which makes our achievable precision competitive. The main difference of the precision stems from the statistically dominated uncertainty of the slow pion tracking efficiency, whose impact on the uncertainty of $|V_{cb}|$ decreased from 1.50% to 0.83%, mainly due to the increase of the data used to derive the tracking efficiency.

Since the measurement of the partial decay rates requires the determination of the branching fraction $\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l)$ along the way, we also provide a precision estimate on the branching fraction for the combination of both lepton channels:

$$\mathcal{B}(\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l) = (5.095 \pm 0.029_{stat} \pm 0.190_{sys})\%, \tag{7.2}$$

which is in good agreement with the MC generator value of 5.110%. Additionally we also provide an estimate on the ratio between the branching fraction for the electron and muon channel, which offers a test for lepton flavor universality predictions of the SM:

$$R = \frac{\mathcal{B}(\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e)}{\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)} = 1.002 \pm 0.012_{stat} \pm 0.036_{sys}, \tag{7.3}$$

The sensitivity studies were done blinded on MC simulations due to internal Belle II policies allowing the application of the analysis chain on data only after a rigorous peer review of the studies on MC by the collaboration, which was not done by the time of writing. Nevertheless we can conclude, that we managed to proof that the concept of a $|V_{cb}|$ extraction through the partial decay rates $\Delta\Gamma/\Delta w \Delta\cos\theta_\ell$ of semileptonic $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays can be done with a potentially competitive precision.

APPENDIX A

Additional Figures for $|V_{cb}|$ from $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays

A.1 LeptonID Coverage

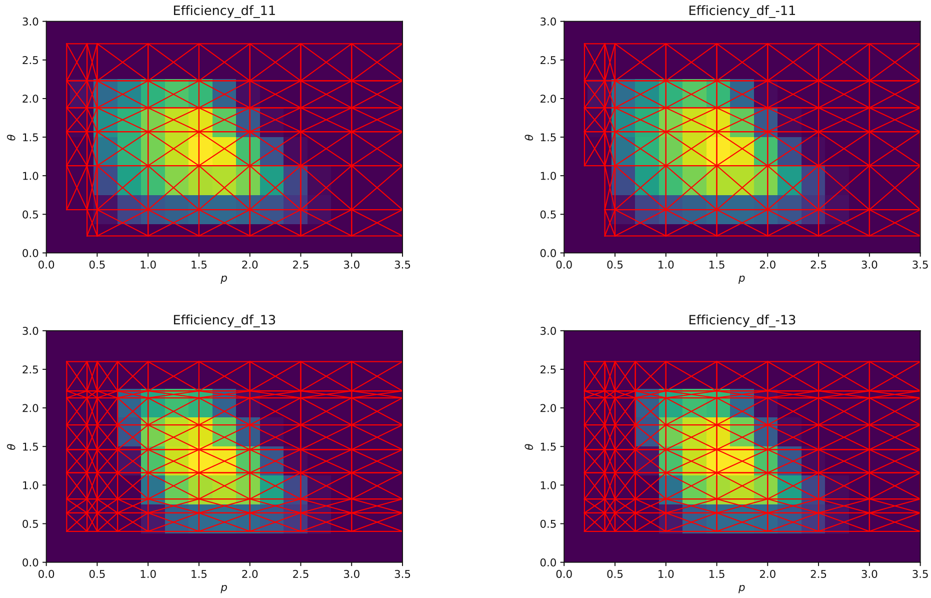


Figure A.1: Coverage of the leptonID efficiency correction tables for the $B \to D^* e \nu_e$ in the top row and $B \to D^* \mu \nu_\mu$ in the lower row. The left column is dedicated to leptons with a negative charge and the right column to leptons with a positive charge.

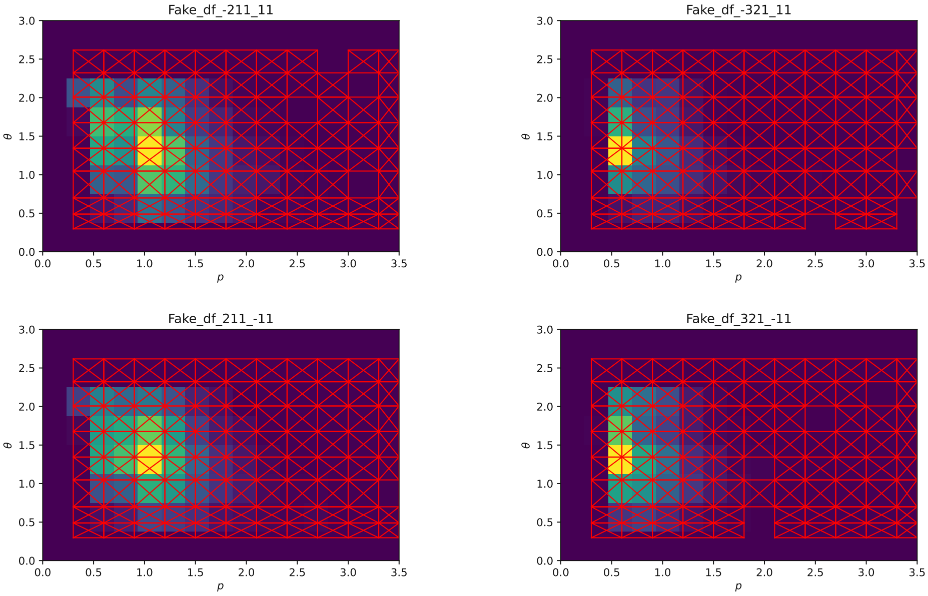


Figure A.2: Coverage of the leptonID fake rate correction tables for the $B \to D^* e \nu_e$. On the left we show the coverage for misidentified pions and on the right for kaons.

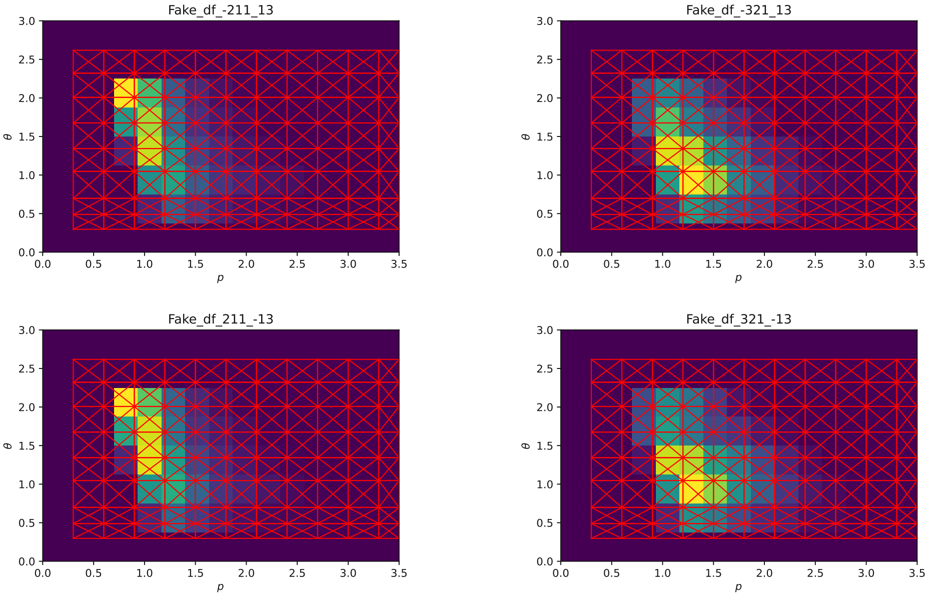


Figure A.3: Coverage of the leptonID fake rate correction tables for the $B \to D^* \mu \nu_\mu$. On the left we show the coverage for misidentified pions and on the right for kaons.

A.2 Data/MC Agreement and Cutflow

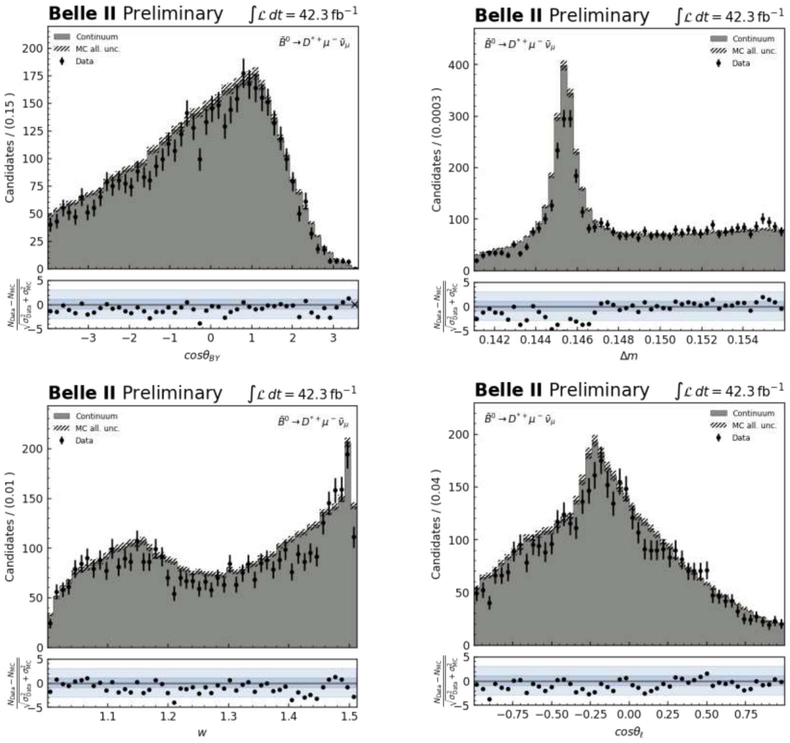


Figure A.4: Agreement of the continuum shapes for the distributions of the parameters of interest for the $B \to D^*\mu\nu_\mu$ channel that are used during the signal extraction.

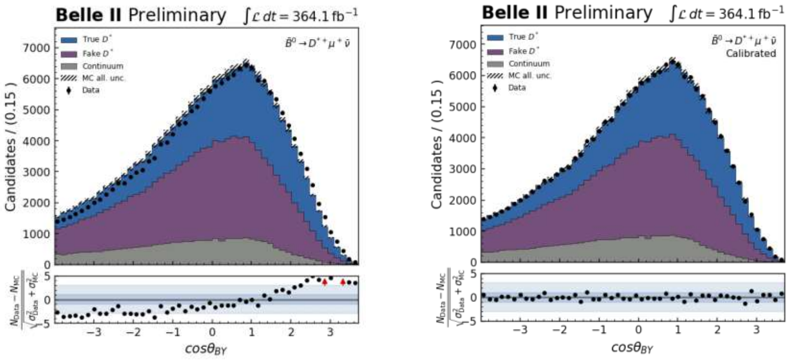


Figure A.5: Comparison of the $\cos\theta_{BY}$ distribution of the $B \to D^*\mu\nu_\mu$ channel before (left) and after (right) the application of the background shape correction factor.

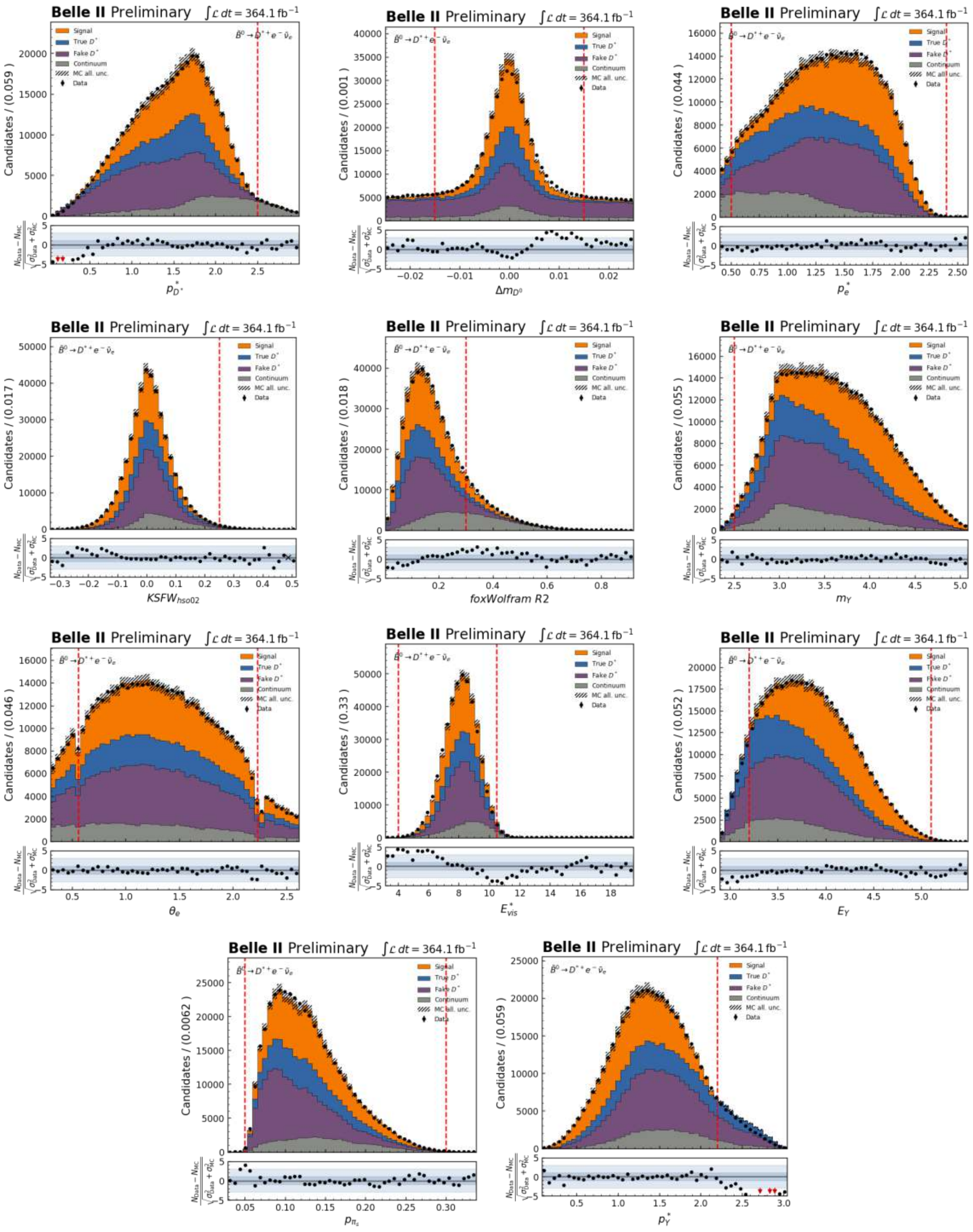


Figure A.6: On-resonance data and MC pre-fit agreement for the n-1 cutflow of $B \to D^* e \nu_e$.

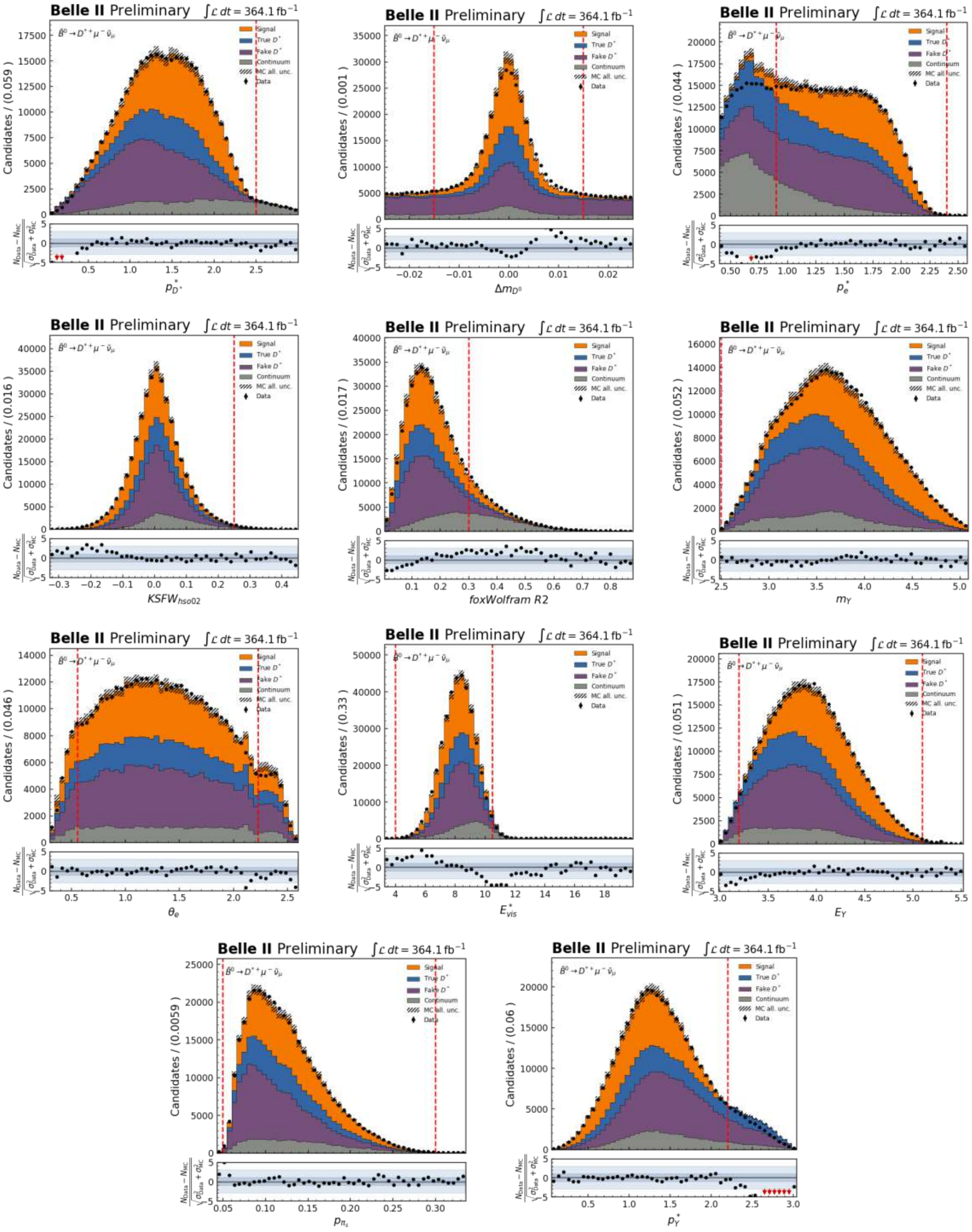
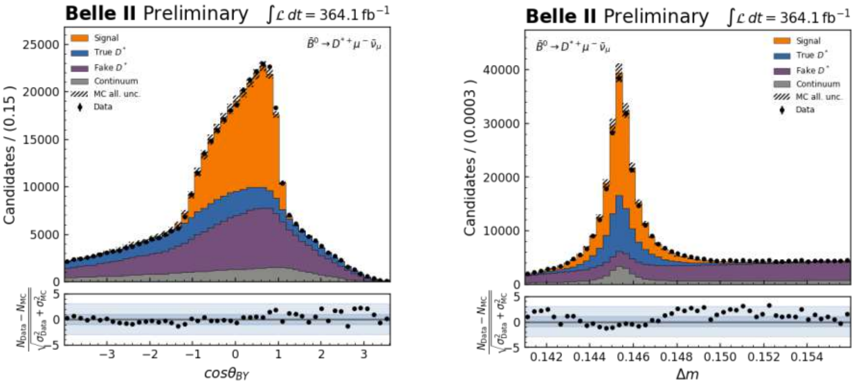
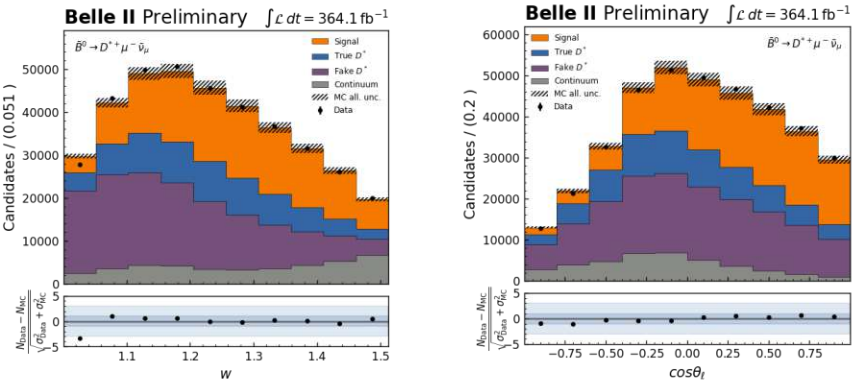


Figure A.7: On-resonance data and MC pre-fit agreement for the n-1 cutflow of $B \to D^*\mu\nu_\mu$.



(a) $\cos\theta_{BY}$

(b) Δm

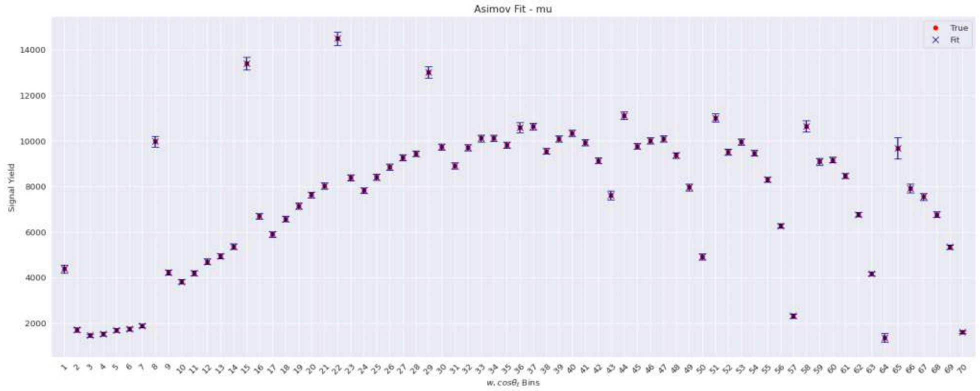


(c) w

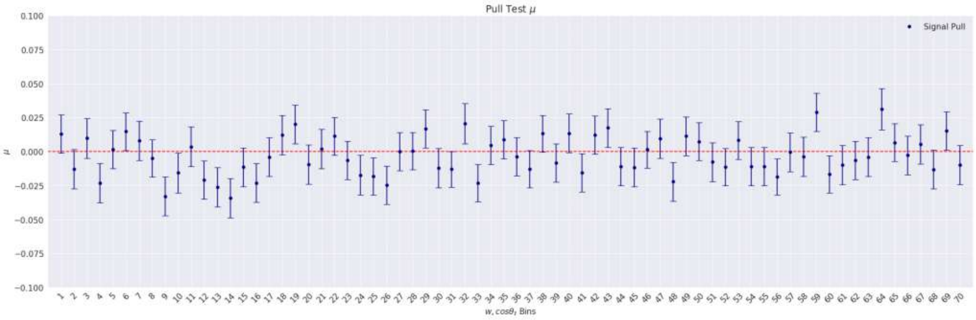
(d) $\cos\theta_\ell$

Figure A.8: Agreement between data and MC of the distributions of the parameters used during the signal extraction for $B \to D^*\mu\nu_\mu$.

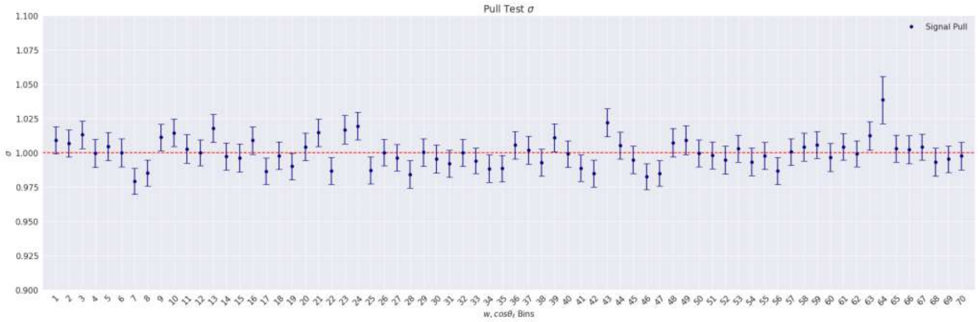
A.3 Signal Extraction Closure Tests



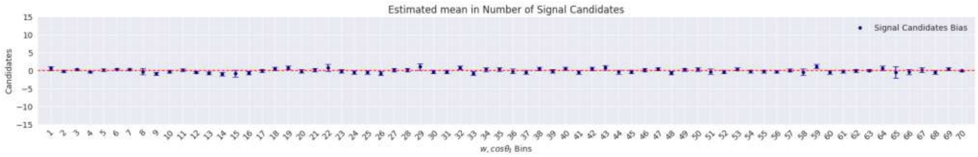
(a) Results of the asimov closure test for $B \to D^* \mu\nu_\mu$.



(b) Mean



(c) Width



(d) Estimated mean in number of candidates.

Figure A.9: Closure test results summary for the $B \to D^* \mu\nu_\mu$ decay channel.

A.4 Correlation Matrix

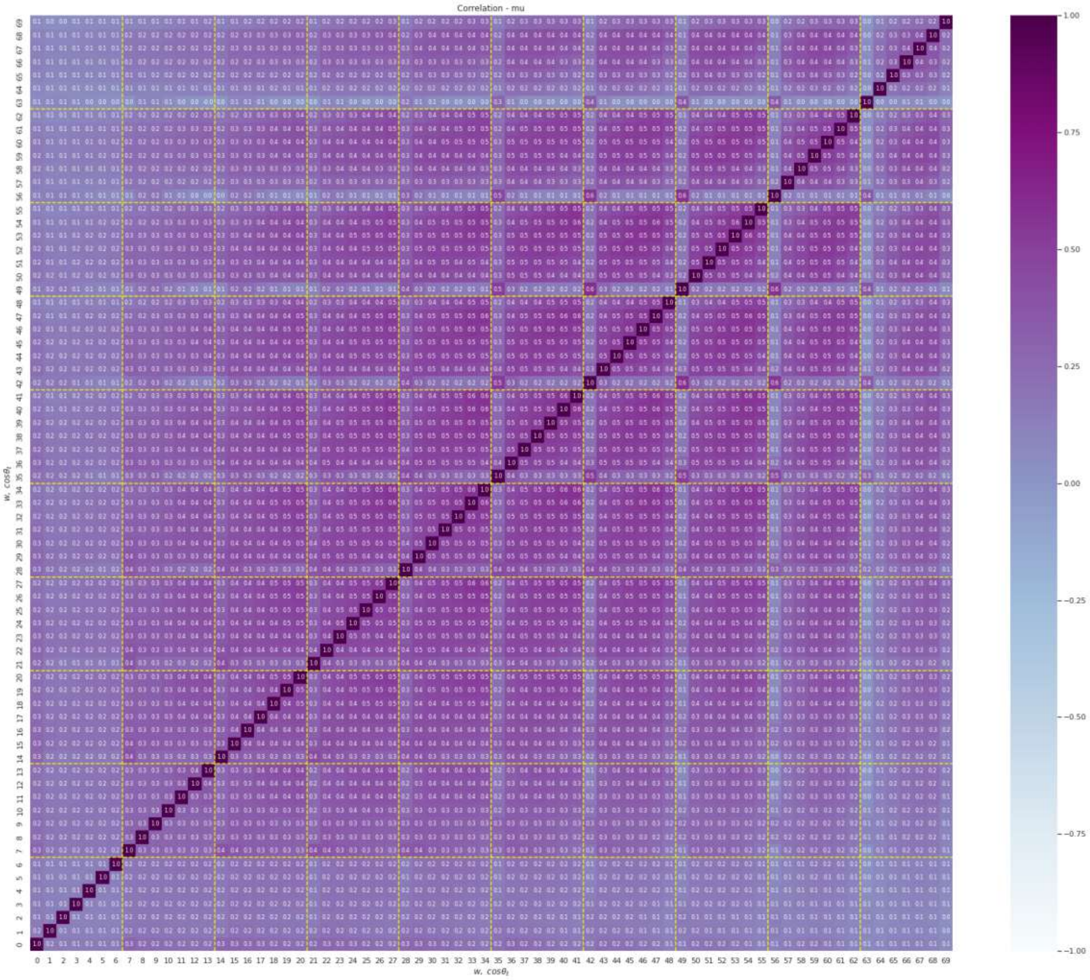


Figure A.10: Systematic and statistical correlation matrix of partial $B \to D^* \mu\nu_\mu$ decay rates.

A.5 BGL Parameterization

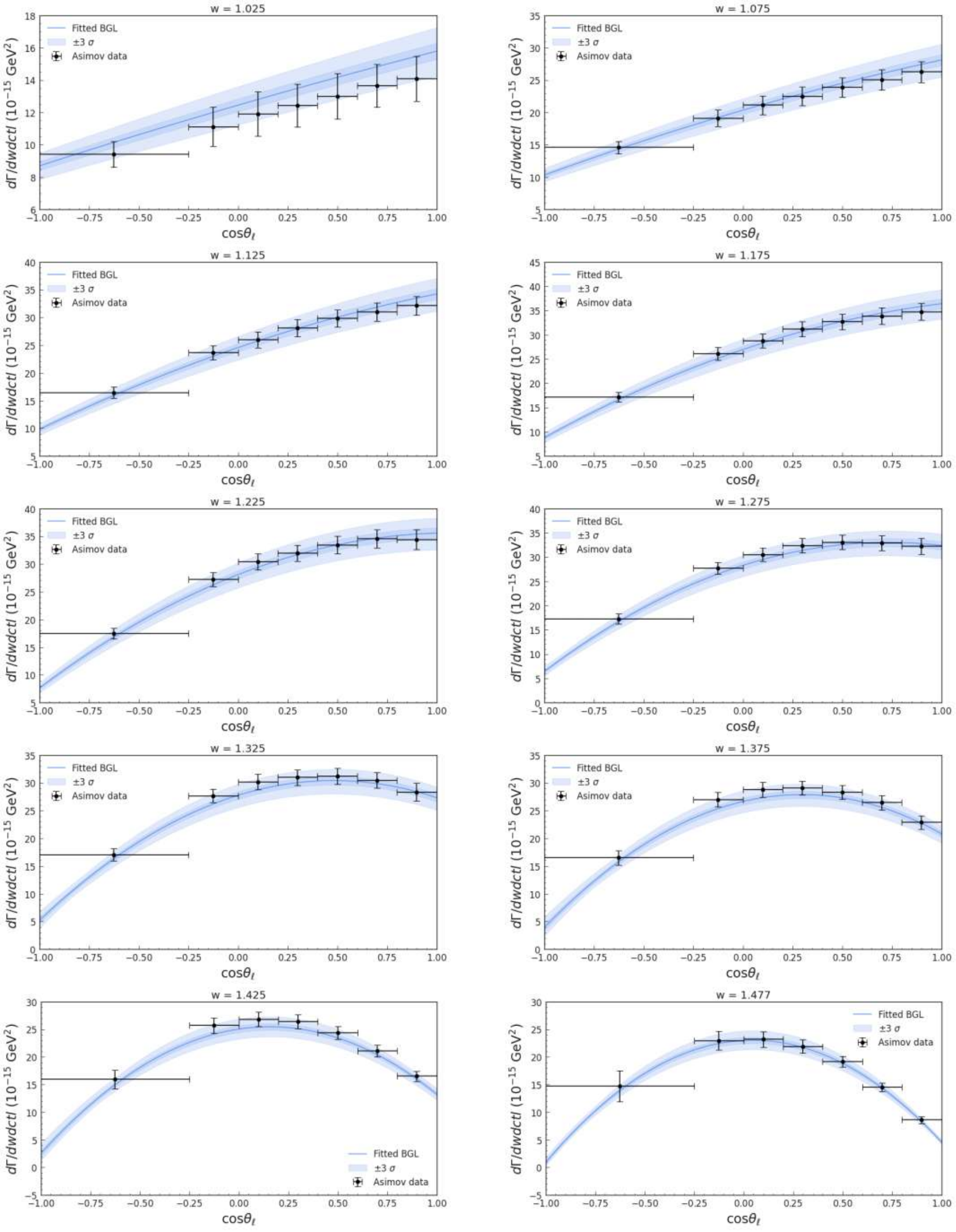


Figure A.11: Agreement between the partial decay rates obtained from the simulation and the shape of the predicted curve from the BGL$_{333}$ χ^2-fit. Each figure shows the partial decay rates of the $\cos\theta_\ell$ for the average value of a w bin.

A.6 CLN Parameterization

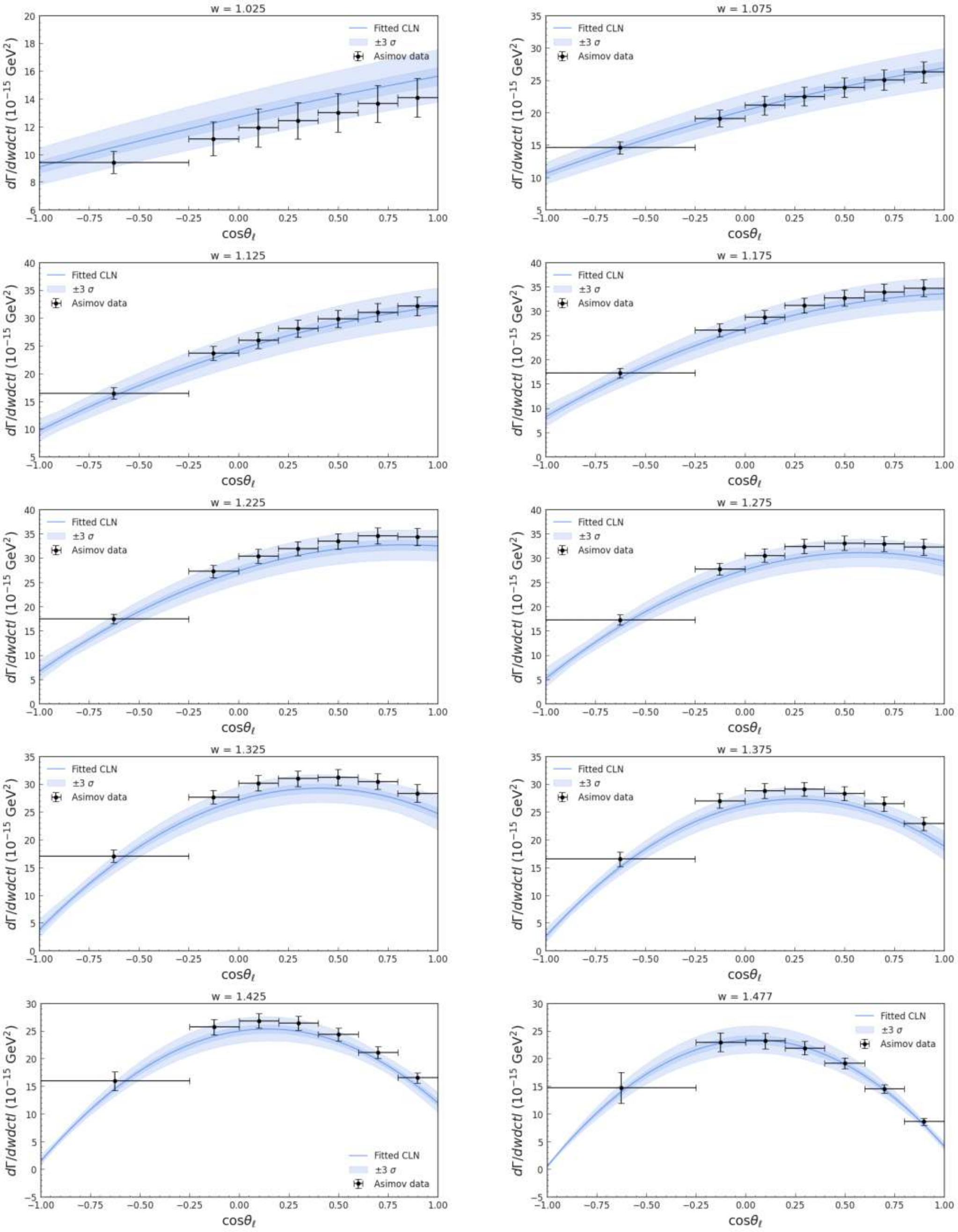


Figure A.12: Agreement between the partial decay rates obtained from the simulation and the shape of the predicted curve from the CLN χ^2-fit. Each figure shows the partial decay rates of the $\cos\theta_\ell$ for the average value of a w bin.

Bibliography

[1] G. Aad, T. Abajyan, et al., ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B 716 (Sept., 2012) 129. `http://dx.doi.org/10.1016/j.physletb.2012.08.020`.

[2] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics 49 (02, 1973) 652–657, https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. `https://doi.org/10.1143/PTP.49.652`.

[3] M. Thomson, *Modern particle physics.* Cambridge University Press, New York, 4^{th} ed., 2019.

[4] Public, *Standard model of elementary particles*, `https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg`, 2020. Accessed: 2024-02-26.

[5] E. Noether, *Invariant variation problems*, Transport Theory and Statistical Physics 1 (Jan., 1971) 186207.

[6] R. Aaij et al., LHCb Collaboration, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. 115 (Aug, 2015) 072001.

[7] R. Aaij, LHCb Collaboration, *Observation of a Narrow Pentaquark State, $P_c(4312)^+$, and of the Two-Peak Structure of the $P_c(4450)^+$*, Phys. Rev. Lett. 122 (Jun, 2019) 222001.

[8] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. 30 (Jun, 1973) 1343–1346. `https://link.aps.org/doi/10.1103/PhysRevLett.30.1343`.

[9] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. 30 (Jun, 1973) 1346–1349. `https://link.aps.org/doi/10.1103/PhysRevLett.30.1346`.

[10] R. L. Workman et al., Particle Data Group, *Review of Particle Physics*, PTEP 2022 (2022) 083C01.

[11] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. 105 (Feb, 1957) 1413–1415.

[12] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. 13 (Jul, 1964) 138–140.

[13] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. 5 (1967) 32–35.

[14] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. 10 (Jun, 1963) 531–533.

[15] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D 2 (Oct, 1970) 1285–1292.

[16] E. Kou, *The Belle II Physics Book*, Progress of Theoretical and Experimental Physics 2019 (Dec., 2019) . `http://dx.doi.org/10.1093/ptep/ptz106`.

[17] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. 51 (Nov, 1983) 1945–1947.
`https://link.aps.org/doi/10.1103/PhysRevLett.51.1945`.

[18] J. Charles, A. Höcker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malclés, J. Ocariz, M. Pivk, and L. Roos, *CP violation and the CKM matrix: assessing the impact of the asymmetric B factories*, The European Physical Journal C 41 (May, 2005) 1131.
`http://dx.doi.org/10.1140/epjc/s2005-02169-1`.

[19] A. Bevan, *The Physics of the B Factories*, The European Physical Journal C 74 (Nov., 2014) . `http://dx.doi.org/10.1140/epjc/s10052-014-3026-9`.

[20] CKMfitter Group, J. Charles et al., *Standard model of elementary particles*, `http://ckmfitter.in2p3.fr`, 2023. The European Physical Journal C (2023) .

[21] Y. Amhis et al., *Averages of b-hadron, c-hadron, and* τ*-lepton properties as of 2021*, Phys. Rev. D 107 (2023) 052008, arXiv:2206.07501 [hep-ex].

[22] A. Sirlin, *Large* m_W*,* m_Z *behaviour of the* $\mathcal{O}(\alpha)$ *corrections to semileptonic processes mediated by* W, Nuclear Physics B 196 (1982) no. 1, 83–92.

[23] M. Neubert, *Heavy-quark symmetry*, Physics Reports 245 (Sep, 1994) 259395.

[24] J. D. Richman and P. R. Burchat, *Leptonic and semileptonic decays of charm and bottom hadrons*, Reviews of Modern Physics 67 (Oct, 1995) 893976.

[25] A. V. Manohar and M. B. Wise, *Heavy quark physics*, vol. 10. 2000.

[26] C. G. Boyd, B. Grinstein, and R. F. Lebed, *Model-Independent Determinations of B -> D l nu , D* l nu Form Factors*, 1995. arXiv:hep-ph/9508211 [hep-ph].

[27] C. G. Boyd, B. Grinstein, and R. F. Lebed, *Precision corrections to dispersive bounds on form factors*, Physical Review D 56 (dec, 1997) 68956911.

[28] I. Caprini, L. Lellouch, and M. Neubert, *Dispersive bounds on the shape of form factors*, Nuclear Physics B 530 (Oct., 1998) 153181.
`http://dx.doi.org/10.1016/S0550-3213(98)00350-2`.

[29] K. Akai, K. Furukawa, and H. Koiso, *SuperKEKB collider*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 907 (Nov., 2018) 188199.
`http://dx.doi.org/10.1016/j.nima.2018.08.017`.

[30] T. Abe et al., *Belle II Technical Design Report*, 2010. arXiv:1011.0352 [physics.ins-det].

[31] Y. Ohnishi et al., *Accelerator design at SuperKEKB*, PTEP 2013 (2013) 03A011.

[32] The Belle II Collaboration, *Muon and electron identification efficiencies and hadron-lepton mis-identification rates at Belle II for Moriond 2022*, .
`https://docs.belle2.org/record/2895`.

[33] H. Svidras et al., The Belle II Collaboration, *Measurement of the data to MC ratio of photon reconstruction efficiency of the Belle II calorimeter using radiative muon pair events*, .

[34] The Belle II Collaboration, KEK, *Belle II archives*, 2024.
`https://www.belle2.org/archives/`. Accessed: 2024-02-14.

[35] The Belle II Collaboration, H. Svidras, *The Central Drift Chamber of Belle 2*, 2020.
`https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5652/CDC_B2SK.pdf`. Accessed: 2024-02-14.

[36] V. Bertacchi et al., The Belle II Collaboration, *Track finding at Belle II*, Computer Physics Communications 259 (Feb., 2021) 107610.

[37] A. Horvath, *The geometry of the Cherenkov radiation*, 2006.
`https://commons.wikimedia.org/wiki/File:Cherenkov.svg`. Accessed: 2024-02-14.

[38] I. Adachi, T. Browder, P. Krian, S. Tanaka, and Y. Ushiroda, *Detectors for extreme luminosity: Belle II*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 907 (2018) 46–59.
`https://www.sciencedirect.com/science/article/pii/S0168900218304200`.

[39] T. Kuhr, C. Pulvermacher, M. Ritter, et al., *The Belle II Core Software*, Computing and Software for Big Science 3 (Feb., 2019) 107610.

[40] R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Computing and Software for Big Science 3 (Feb., 2019) 107610.

[41] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, Computer Physics Communications 191 (June, 2015) 159177.
`http://dx.doi.org/10.1016/j.cpc.2015.01.024`.

[42] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A 462 (2001) 152–155.

[43] S. Jadach, B. Ward, and Z. Was, *The precision Monte Carlo event generator for two-fermion final states in collisions*, Computer Physics Communications 130 (Aug., 2000) 260325. `http://dx.doi.org/10.1016/S0010-4655(00)00048-5`.

[44] N. Davidson, G. Nanava, T. Przedziski, E. Richter-Was, and Z. Was, *Universal interface of TAUOLA: Technical and physics documentation*, Computer Physics Communications 183 (Mar., 2012) 821843. `http://dx.doi.org/10.1016/j.cpc.2011.12.009`.

[45] C. Carloni Calame, C. Lunardini, G. Montagna, O. Nicrosini, and F. Piccinini, *Large-angle Bhabha scattering and luminosity at flavour factories*, Nuclear Physics B 584 (Sept., 2000) 459479. `http://dx.doi.org/10.1016/S0550-3213(00)00356-4`.

[46] F. A. Berends, P. H. Daverveldt, and R. Kleiss, *Monte Carlo Simulation of Two Photon Processes: II: Complete Lowest Order Calculations for Four Lepton Production Processes in electron Positron Collisions*, Comput. Phys. Commun. 40 (1986) 285–307.

[47] W. Dungel, C. Schwanda, I. Adachi, et al., The Belle Collaboration, *Measurement of the form factors of the decay*, Physical Review D 82 (Dec., 2010) . `http://dx.doi.org/10.1103/PhysRevD.82.112007`.

[48] T. Keck, F. Abudinén, F. U. Bernlochner, et al., *The Full Event Interpretation: An Exclusive Tagging Algorithm for the Belle II Experiment*, Computing and Software for Big Science 3 (Feb., 2019) . `http://dx.doi.org/10.1007/s41781-019-0021-8`.

[49] M. Röhrken, *Time-Dependent CP Violation Measurements in Neutral B Meson to Double-Charm Decays at the Japanese Belle Experiment.* PhD thesis, Karlsruhe U., 2012.

[50] G. C. Fox and S. Wolfram, *Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes*, Phys. Rev. Lett. 41 (Dec, 1978) 1581–1585. `https://link.aps.org/doi/10.1103/PhysRevLett.41.1581`.

[51] T. Keck, *FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*, 2016. arXiv:1609.06119.

[52] V. Blobel, *Unfolding Methods in Particle Physics*, 2011. `https://cds.cern.ch/record/2203257`.

[53] L. Heinrich, M. Feickert, G. Stark, and K. Cranmer, *pyhf: pure-Python implementation of HistFactory statistical models*, Journal of Open Source Software 6 (2021) no. 58, 2823. `https://doi.org/10.21105/joss.02823`.

[54] ROOT, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., New York U., New York, 2012. `https://cds.cern.ch/record/1456844`.

[55] J.-F. Krohn, F. Tenchini, P. Urquijo, et al., *Global decay chain vertex fitting at Belle II*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators,

Spectrometers, Detectors and Associated Equipment 976 (2020) 164269. `https://www.sciencedirect.com/science/article/pii/S0168900220306653`.

[56] The Belle II Collaboration, H. Svidras, M. Röhrken, S. Wehle, and K. Tackmann, *Measurement of the data to MC ratio of photon reconstruction efficiency of the Belle II calorimeter using radiative muon pair events*, Sep, 2021.

[57] The Belle II Collaboration, *Systematic Corrections Framework*, 2024. `https://syscorrfw.readthedocs.io/en/latest/index.html#`. Accessed: 2024-04-12.

[58] The Belle II Collaboration, W. Sutcliffe, M. Eliachevitch, H. Junkerkalefeld, et al., *Particle ID efficiency and fake rate corrections for Belle II*, 2023. `https://gitlab.desy.de/meliache/pidvar`. Accessed: 2024-04-12.

[59] M. Prim, F. Bernlochner, et al., *Measurement of differential distributions of* $B \to D^* \ell \nu$, Physical Review D 108 (July, 2023) . `http://dx.doi.org/10.1103/PhysRevD.108.012002`.

[60] S. Choudhury et al., Belle Collaboration, *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*, Phys. Rev. D 107 (Feb, 2023) L031102. `https://link.aps.org/doi/10.1103/PhysRevD.107.L031102`.

[61] A. Bazavov, C. E. DeTar, et al., Fermilab Lattice and MILC Collaborations, *Semileptonic form factors for* $B \to D^* \ell \nu$ *at nonzero recoil from* $2+1$*-flavor lattice QCD*, The European Physical Journal C 82 (Dec., 2022) . `http://dx.doi.org/10.1140/epjc/s10052-022-10984-9`.

[62] D. Bigi, P. Gambino, and S. Schacht, *A fresh look at the determination of* $|V_{cb}|$ *from* $B \to D^* \ell \nu$, Physics Letters B 769 (June, 2017) 441445. `http://dx.doi.org/10.1016/j.physletb.2017.04.022`.

[63] Y. Aoki and T. o. Blum, *FLAG Review 2021*, The European Physical Journal C 82 (Oct., 2022) . `http://dx.doi.org/10.1140/epjc/s10052-022-10536-1`.

[64] I. Adachi et al., Belle II Collaboration, *Determination of* $|V_{cb}|$ *using* $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ *decays with Belle II*, Phys. Rev. D 108 (Nov, 2023) 092013. `https://link.aps.org/doi/10.1103/PhysRevD.108.092013`.