Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

TECHNISCHE
UNIVERSITÄT
WIEN

DIPLOMARBEIT

Measurement of the $B \to D^0 \ell^- \overline{\nu}_\ell$ Branching Fraction using Belle II Data

Ausgeführt am Atominstitut der Technischen Universität Wien

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

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Technische Physik

unter der Anleitung von

Privatdoz. Dipl.-Ing. Dr.techn. Christoph SCHWANDA

eingereicht an der Technischen Universität Wien von

Philipp HORAK, BSc

Unterschrift des Verfassers

Unterschrift des Betreuers

Wien, am 16. Oktober 2020

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Kurzfassung

Acht Jahre nach Beendigung des Belle-Experiments im Jahr 2010 wurde der verbesserte Belle II Detektor 2018 in Betrieb genommen. Belle II ist ein Teilchenbeschleunigerexperiment und befindet sich am KEK Forschungszentrum für Hochenergiephysik in Tsukuba, Japan. Durch Kollisionen von Elektronen und Positronen bei einer Schwerpunktsenergie von 10.58 GeV, welche ungefähr der doppelten B-Meson Masse entspricht, können große Mengen an B-Mesonen produziert und analysiert werden.

Ziel dieser Arbeit ist es, das Verzweigungsverhältnis der Zerfälle $B^- \to D^0 e^- \overline{\nu}_e$ und $B^- \to D^0 \mu^- \overline{\nu}_\mu$ mittels der Belle II Daten von 2019 und 2020 zu messen. Die Endprodukte des Zerfalls sind im Detektor sichtbar, mit Ausnahme des Neutrinos, und werden kombiniert um den Zerfall zu rekonstruieren. Der dominante $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$ $(\ell = e, \mu)$ Hintergrund wird mit Hilfe von maschinellem Lernen und anderen Methoden der multivariaten Statistik unterdrückt.
Ein Datensatz von 34.6 fb^{-1}, gleichzusetzen mit schätzungsweise 39 Millionen $B^\pm$ Mesonen, wird verwendet um ein $B^- \to D^0 \ell^- \overline{\nu}_l$ $(\ell = e, \mu)$ Verzweigungsverhältnis von $(2.173 \pm 0.060(\text{stat.}) \pm 0.163(\text{sys.}))\,\%$ zu messen. Das Messergebnis liegt in Übereinstimmung mit Messungen des Zerfalls bei vorhergehenden Experimenten wie Belle und BaBar und des gesammelten weltweiten Durchschnittswerts $(2.35 \pm 0.03 \pm 0.09)\%$.
Das Verhältnis von Verzweigungsverhältnissen des Zerfalls mittels e und μ $R\,(e/\mu) = \frac{\text{Br}(B^- \to D^0 e^- \overline{\nu}_e)}{\text{Br}(B^- \to D^0 \mu^- \overline{\nu}_\mu)} = 1.026 \pm 0.055(\text{stat.}) \pm 0.074(\text{sys.})$ wird gemessen. Das Ergebnis liegt in Übereinstimmung mit der Erwartung für Leptonuniversalität $R\,(e/\mu) = 1$

Die Arbeit bietet einen ersten Blick auf Messung des $B^- \to D^0 \ell^- \overline{\nu}_l$ Verzweigungsverhältnisses mittels Daten des Belle II Experiments und liefert einen Grundstein für umfassendere Analysen dieser Zerfälle.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

Eight years after the conclusion of the Belle experiment, the upgraded Belle II detector started operation in 2018. Belle II is a particle collider experiment located at the KEK Laboratory in Tsukuba, Japan. By colliding electrons and positrons at a center-of-mass energy of 10.58 GeV, equivalent to approximately two times the B meson rest mass, large amounts of B mesons are produced.

The goal of this analysis is to measure the branching ratio of the semileptonic decays $B^- \to D^0 e^- \overline{\nu}_e$ and $B^- \to D^0 \mu^- \overline{\nu}_\mu$, using data of the 2019 and 2020 physics runs. The decay is reconstructed by combining final state particles originating in the decay from detector data. The predominant $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$ ($\ell = e, \mu$) background is suppressed by employing a deep neural network and other multivariate analysis techniques.
A dataset of 34.6 fb^{-1}, corresponding to approximately 39 million $B^\pm$ mesons is analyzed to yield a $B^- \to D^0 \ell^- \overline{\nu}_l$ ($\ell = e, \mu$) branching fraction of $(2.173 \pm 0.060(\text{stat.}) \pm 0.163(\text{sys.}))\%$. This is in agreement with measurements of previous experiments such as Belle and BaBar as well as with the world average $(2.35 \pm 0.03 \pm 0.09)\%$.
The measured ratio of electron to muon branching ratios $R\,(e/\mu) = \frac{\text{Br}(B^- \to D^0 e^- \overline{\nu}_e)}{\text{Br}(B^- \to D^0 \mu^- \overline{\nu}_\mu)} = 1.026 \pm 0.055(\text{stat.}) \pm 0.074(\text{sys.})$ is also in agreement with the lepton universality expectation $R\,(e/\mu) = 1$.

This thesis provides a first look at measuring the $B^- \to D^0 \ell^- \overline{\nu}_l$ branching fraction at Belle II, laying grounds for a comprehensive study of these modes with the further goal of extracting the CKM parameter $|V_{cb}|$ via this decay mode.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Chapter 1

Introduction

Since its inception in the 1970s, the Standard Model of particle physics is the best known description of elementary particles and the forces that govern them. It succesfully incorporates the electromagnetic, the weak and the strong interaction, to accurately predict particle behavior, decays and collision rates. The Standard Model is able to calculate decay rates to remarkable precision and succesfully predicted particles before being seen experimentally, such as the bottom quark discovered in 1977 and most recently the Higgs boson discovered in 2012.

However, despite its tremendous success, the Standard Model has shortcomings. Because it fails to incorporate the fourth fundamental interaction, gravity, and cannot account for dark energy and dark matter, the Standard Model is not seen as the final theory of everything. Attempts to unify the Standard Model with gravity, such as string theory, could not be confirmed experimentally. Another problem is the inability to explain why the amount of observed matter significantly outweighs the amount of observed antimatter, known as *matter-antimatter asymmetry problem*.

For matter-antimatter asymmetry to occur, processes that violate a special symmetry, the *charge-parity symmetry* (CP symmetry), are required. A mechanism to describe violation of CP symmetry in the Standard Model, called the quark mixing mechanism, was introduced by Cabibbo in 1963 and later generalized by Kobayashi and Maskawa in 1973. In order for CP violation to occur in quark mixing, three quark generations are necessary, while only quarks of two generations were known at the time. The discovery of the bottom quark four years later gave validation to the quark mixing mechanism and allowed CP violation in the Standard Model to be measured succesfully, although at an order of magnitude that is far too low to explain the observed matter-antimatter asymmetry. The quark mixing mechanism is described by the nine elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The parameters have been studied and measured extensively at experiments such as Belle, BaBar and LHCb at CERN. For two of the parameters, $|V_{ub}|$ and $|V_{cb}|$, there are long standing discrepancies between two different methods of measuring the parameters. For $|V_{cb}|$ specifically a 3σ anomaly is observed between measuring through the explicit $B^- \to D^0\ell^-\overline{\nu}_l$ or $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decay branching fractions and measuring inclusively through $B \to X_c\ell\nu$, where X_c can be any particle containing a charm quark.

This thesis aims at measuring the branching fraction of the decay $B^- \to D^0\ell^-\overline{\nu}_l$,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

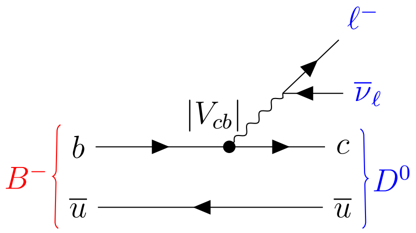


FIGURE 1.1: A Feynman diagram illustrating the focal decay of this thesis, $B^- \to D^0\ell^-\overline{\nu}_l$. The initial state particle on the left, a B^- meson composed of a b quark and a $\overline{u}$ quark, is produced in electron-positron collisions at the Belle II experiment. It decays to the final state particles pictured on the right side, a D^0 meson composed of c and $\overline{u}$ quarks, a lepton ℓ, and a corresponding neutrino $\overline{\nu}_\ell$. The b to c quark transition introduces the CKM parameter $|V_{cb}|$, making the decay rate dependant on it and thus allowing $|V_{cb}|$ measurement through measuring the $B^- \to D^0\ell^-\overline{\nu}_l$ branching fraction.

shown in Figure 1.1, using data from the Belle II experiment, preparing exclusive measurement of $|V_{cb}|$ through this mode. The Belle II experiment takes particle collision data by producing B mesons in large amounts. As successor to the Belle experiment, improvements have been done to the accelerator, detector and other parts involved, with plans to collect up to $\sim$ 70 times more data at higher measurement sensitivity. Belle II began to collect physics data in 2019. Collision data from spring 2019 to summer 2020 is used in this thesis to reconstruct $B^- \to D^0\ell^-\overline{\nu}_l$ and measure the branching fraction. The thesis is outlined as follows:

An overview of the theoretical physics background and the Standard Model is given in Chapter 2. Chapter 3 gives an introduction to the SuperKEKB accelerator and the Belle II detector including all its subdetectors and software framework. Data analysis techniques and statistical principles used in the reconstruction are described in Chapter 4. The $B^- \to D^0\ell^-\overline{\nu}_l$ reconstruction with selection criteria and background selection, as well as a comparison between physics data and simulated data is given in Chapter 5. A fitting procedure used to extract a number of $B^- \to D^0\ell^-\overline{\nu}_l$ signal events from the dataset, necessary to calculate the branching fraction is introduced in Chapter 6. A resulting branching fraction with statistical and systematic uncertainties is given in Chapter 7. Finally, Chapter 8 gives a conclusion and an outlook to the analysis performed.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Chapter 2

Theoretical Background

This chapter aims to establish the theoretical foundation required for the following analysis. The physics background starts with a brief introduction of the Standard Model and its fundamental particles and forces, with special focus on the weak interaction. A short digression into cosmology explains the requirement of charge-parity (CP) violation to explain the matter-antimatter asymmetry. The formulation of CP violation in the standard model is presented via the quark mixing mechanism and the CKM matrix, to demonstrate the significance of measuring $B \to D$ decays. To conclude the physics introduction, the current experimental status is given.

2.1 The Standard Model

The Standard Model (SM) of particle physics [2][3][4] is a theory describing the elementary particles and the three fundamental forces that govern them. Particles with integer intrinsic angular momentum (spin) are called bosons, particles with half-integer spin are called fermions. All Standard Model particles are shown in Figure 2.1.

Fermions are the particles that make up matter. They can be split up further into 6 leptons and 6 quarks and a respective antiparticle for each. The charged electron e^-, muon μ^- and tau τ^- particles as well as the chargeless electron neutrino ν_e, muon neutrino ν_μ and tau neutrino ν_τ make up the leptons. The quarks consist of 3 up-type quarks with charge $\frac{2}{3}$, the up quark u, the charm quark c and the top quark t and 3 down-type quarks with charge $-\frac{1}{3}$, the down quark d, the strange quark s and the bottom[*] quark b. Each particle's antiparticle has opposite electrical charge but otherwise shares the same properties with the original particle. Antiparticles are either noted with opposite charge or with an overline, *i.e.* the positron e^+ and the anti-up quark $\overline{u}$. The quarks and leptons can both be grouped into 3 generations with increasing mass. Heavier particles can then always decay into the lighter equivalents.

Bosons are the particles that mediate fundamental forces. The three fundamental forces of the SM are the electromagnetic force between electrically charged particles, the strong force between color charged particles, *i.e.* quarks and gluons, and the weak force affecting all fundamental particles. The fourth fundamental force, gravity, is not included in the SM. A summary of the properties of the forces can be found in Table 2.1.

---

[*] In literature this is also known as the beauty quark

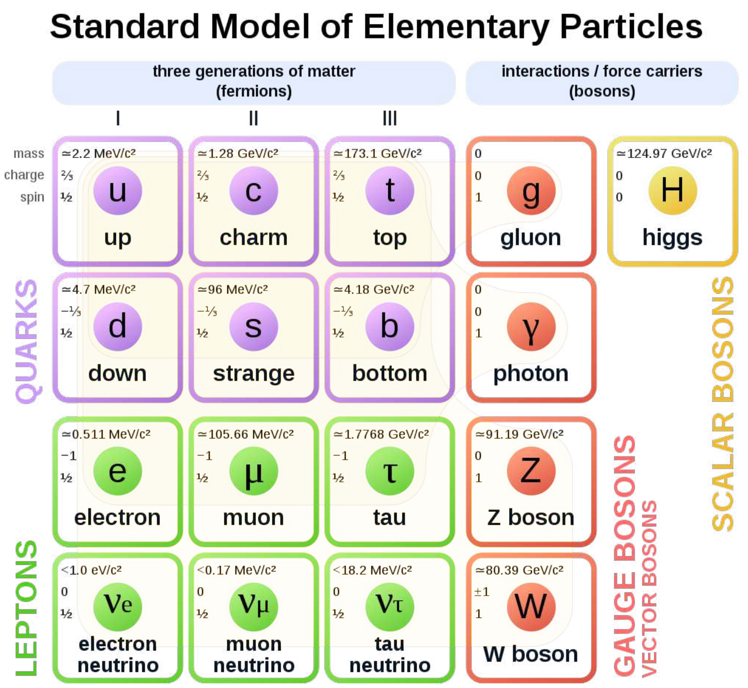


FIGURE 2.1: The particles of the Standard Model [1].

The massless and chargeless photon γ carries the electromagnetic force. The electromagnetic force between charged particles has infinite range, decreasing in strength with $\frac{1}{r^2}$. Its strength is characterized by the fine-structure constant $\alpha_f \approx \frac{1}{137}$.

In contrast, the strong force is short-ranged. It is mediated by massless gluons g and acts on all particles with non-zero color charge, *i.e.* quarks and gluons. The strong force has a larger coupling strength than both the electromagnetic and the weak interaction, and is thus responsible for keeping the quark constituents in atomic nuclei together, despite them containing same-charge quarks that repel electromagnetically.

An important property of Quantum Chromodynamics (QCD), the quantum theory describing the strong interaction, is *color confinement*. Color confinement is the phenomenon that quarks can not be observerved as free particles, but instead clump together to form bound states. It is described by the axiom that only *color singlet* compound states can be observed in nature, meaning combinations such as $q\bar{q}$, qqq and $\overline{qqq}$ are allowed while combinations such as q or $qq\bar{q}$ are not. Composite particles consisting of quarks bound by the strong force are called hadrons. Quark - antiquark pairs are called mesons, combinations of 3 (anti-)quarks are called (anti-)baryons.

The weak force has a coupling several orders of magnitude weaker than the strong force. It is mediated by massive, charged $W^{\pm}$ and neutral Z bosons. All SM particles are affected by the weak force and it is unique in the fact that particles interacting with W bosons can change flavor.

Standard model processes involving particles and forces can be graphically represented by Feynman diagrams. Examples involving all the forces can be seen

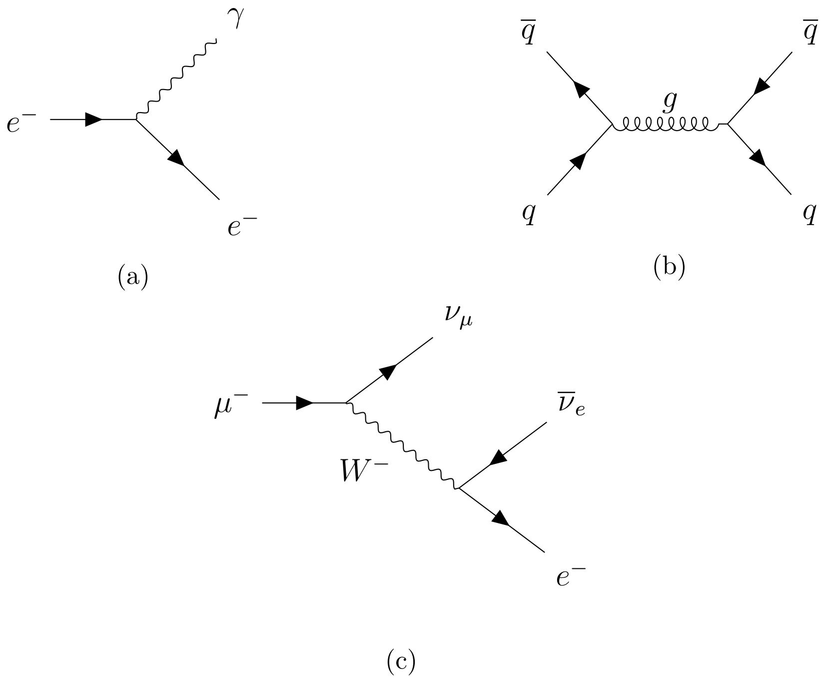
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

in Figure 2.2.

Interaction	Mediator	Relative strength (at ~ 1 GeV*)	Range
Electromagnetic	γ	$\alpha_f \approx \frac{1}{137}$	∞
Strong Force	g	$\alpha_s \approx 1$	~ 1 fm
Weak Force	W and Z	$\alpha_W \approx 10^{-6}$	$\sim 10^{-3}$ fm

TABLE 2.1: Properties of the three Standard Model interactions. Strength and range should only be viewed relative to each other.



(a) (b) (c)

FIGURE 2.2: Feynman diagrams for each of the fundamental forces. The time axis is from left to right. (a) An electron emitting a photon in an electromagnetic process. (Bremsstrahlung) (b) Strong process with a quark and an antiquark annihilating into a gluon followed by a new pair production. (c) A muon decaying into a lighter electron due to the weak force, giving off a neutrino and an anti-neutrino. ($\mu^- \to e^- \overline{\nu}_e \nu_\mu$)

2.2 Charge-parity violation

The Standard Model is a CPT-symmetric theory, meaning that application of simultaneous discrete transformations of charge conjugation (C), parity transformation (P) and time reversal (T) has no effects on observable physics. Violations of the symmetry of a single component, such as T, can be counteracted by a violation in the other two components while preserving overall CPT-symmetry [4]. A brief discourse into cosmology follows to motivate the importance of measuring CP violation in the standard model.

---

[*] The fundamental force coupling constants, and with it their relative strength, are energy-dependent. For the Belle II experiment, values at the scale of $1 - 10$ GeV are relevant.

If CP were to be a conserved quantity our universe would have equal amounts of matter and antimatter. A parameter to measure potential matter - antimatter discrepancies is the baryon asymmetry parameter

$$\eta_B = \frac{n_B - n_{\overline{B}}}{n_\gamma}\,, \tag{2.1}$$

where n_B, $n_{\overline{B}}$ and n_γ are the number of baryons, anti-baryons and photons in our universe.
In a universe without CP-symmetry violation, baryon to photon density ratios can be calculated from thermodynamic considerations [5][6] as

$$\frac{n_B}{n_\gamma} = \frac{n_{\overline{B}}}{n_\gamma} \sim 10^{-18}. \tag{2.2}$$

Current cosmological measurements from relic light element compositions [7] and the cosmic microwave background [8] give

$$\eta_B \approx 6 \cdot 10^{-10} \tag{2.3}$$

This 8 orders of magnitude discrepancy is known as the matter-antimatter asymmetry problem.
To give rise to an asymmetry between matter and antimatter, Sakharov formulated three conditions [9] that have to be fulfilled:

1. Baryon number symmetry violation such that $n_B - n_{\overline{B}}$ is not constant
2. Violation of the discrete symmetries C and CP
3. Departure from thermal equilibrium

The first condition states that there must be processes that produce an influx of baryons or anti-baryons, else they would always exist in equal amounts. Then, C and CP violation is necessary. If CP was a conserved quantity, the charge conjugated process (c.c.) of any process producing surplus baryons would produce surplus antibaryons with the exact same rate, and would thus cancel out. With CP violation, the decay rate of a process and its c.c. can differ. Finally, CPT symmetry would always balance processes causing surplus baryons in thermal equilibrium. Thus departure from thermal equilibrium is the third necessary condition.

2.3 Quark mixing

In the quark sector of the Standard Model, CP violation can be introduced naturally [10]. This originates from how the weak force interacts with quarks. Coupling strengths of bosons to charged leptons are equal for all flavors, this property is called lepton universality.* In contrast, the decay rates for the weak

---

*Hints of deviations of lepton universality have been found in recent Belle, BaBar and LHCb experiments when comparing τ to μ and e processes, but so far the stastical significance is too low to claim loss of lepton universality [11].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

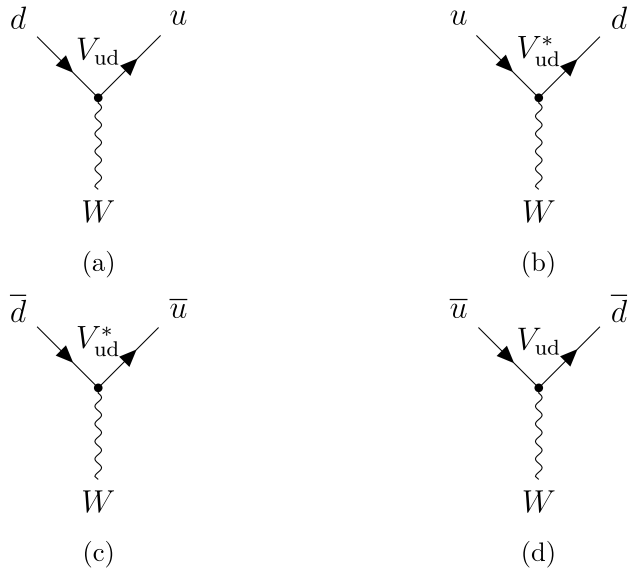


FIGURE 2.3: A Feynman diagram involving 2 quarks introduces an additional $V_{q_1 q_2}$ factor, due to the quark mixing mechanism differentiating mass from weak eigenstates. For a process with an incoming down-type and outgoing up-type quark, such as (a), this factor is the respective CKM matrix parameter. For the case of an incoming up-type quark and outgoing down-type quark, such as (b), the complex conjugate CKM parameter is used. As seen in the bottom row (c) and (d), applying CP transformation changes quarks into their anti-quarks while picking up another complex conjugation on the matrix element.

decays $K^-(u\overline{s}) \to \mu^-\overline{\nu}_\mu$ and $\pi^-(u\overline{d}) \to \mu^-\overline{\nu}_\mu$ differ by a factor of approximately 20, which implies that there is no such thing as quark universality [12]. These observations were originally explained by the Cabibbo hypothesis. The Cabibbo hypothesis explains the different u and d couplings by making a distinction between the free particle, mass eigenstate of the quarks and the weak eigenstate that participates in the weak interaction. The weak eigenstates labelled d' and s' are then given as

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

with the Cabibbo angle θ_c. This mechanism succesfully explains difference in u and d couplings by introducing additional terms into the decay rate calculations, dependant on participating quark flavors. From observed decays, the Cabibbo angle can be measured as $\theta_c \simeq 13.02°$.

Kobayashi and Maskawa extended the quark mixing mechanism to a third quark generation

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix} \tag{2.5}$$

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

with the Cabibbo-Kobayashi-Maskawa (CKM) matrix [10]

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix}. \tag{2.6}$$

The weak quark eigenstates are then given as a mixture of mass eigenstates

$$|d'\rangle = V_{\mathrm{ud}}\,|d\rangle + V_{\mathrm{us}}\,|s\rangle + V_{\mathrm{ub}}\,|b\rangle\,. \tag{2.7}$$

A d quark coupling to a u quark, such as in Figure 2.3, introduces an additional factor V_{ud}, contributing to decay rate calculations.

To preserve the overall quark current normalization, the CKM matrix is required to be unitary. In general, the parameters of the CKM matrix are complex numbers, with the unitarity constraint reducing the number of free parameters to 9. Since decays such as $d \to u$ and its complex conjugate $\overline{d} \to \overline{u}$ pick up factors V_{ud} and V^*_{ud} respectively as shown in Figure 2.3, an imaginary part in any of the CKM matrix elements leads to differing decay rates and immediately implies CP violation in the quark sector.

By measuring quark flavor changing decays, the magnitudes of the CKM elements can be measured. The Particle Data Group [12] lists the world averages of the CKM matrix element magnitudes as

$$\begin{pmatrix} |V_{\mathrm{ud}}| & |V_{\mathrm{us}}| & |V_{\mathrm{ub}}| \\ |V_{\mathrm{cd}}| & |V_{\mathrm{cs}}| & |V_{\mathrm{cb}}| \\ |V_{\mathrm{td}}| & |V_{\mathrm{ts}}| & |V_{\mathrm{tb}}| \end{pmatrix} = \begin{pmatrix} 0.97446 & 0.22452 & 0.00365 \\ 0.22438 & 0.97359 & 0.04214 \\ 0.00896 & 0.04133 & 0.999105 \end{pmatrix}. \tag{2.8}$$

The diagonal couplings between quarks of the same generation dominate the off-diagonal flavor changing terms. Because of the dominating diagonal in the CKM matrix a common parameterisation used is the Wolfenstein paramterisation [13], an expansion around the parameter [12]

$$\lambda = \frac{|V_{\mathrm{us}}|}{\sqrt{|V_{\mathrm{ud}}|^2 + |V_{\mathrm{us}}|^2}} \approx 0.225. \tag{2.9}$$

The CKM matrix can be parametrised with the 4 real parameters λ, A, ρ and η and up to $\mathcal{O}(\lambda^4) \sim 10^{-3}$ written as:

$$\begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(1-i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.10}$$

In this parametrisation, the hierachy between matrix elements with dominant diagonal terms becomes visible. The complex part is limited to V_{ub} and V_{td}, entirely parametrized by η (the next contributing terms in V_{cd} and V_{ts} are $\mathcal{O}(\lambda^5)$). Experimental measurements [12] yield

$$\eta = 0.355^{+0.012}_{-0.011}, \tag{2.11}$$

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

confirming CP violation in the Standard Model, although at an order of magnitude that is insufficient to explain the level of baryon asymmetry [14] outlined in Section 2.2 [15]. Therefore it is crucial to find additional sources of CP violation beyond the Standard Model.

2.3.1 Unitarity triangle

The unitarity condition on the CKM matrix (2.6) can be expressed as the constraints $V_{\text{CKM}}^{\dagger} V_{\text{CKM}} = \mathbb{1}$ and $V_{\text{CKM}} V_{\text{CKM}}^{\dagger} = \mathbb{1}$. Writing the unitarity equations in terms of the matrix elements gives [16]

$$\sum_{\alpha \in (u,c,t)} V_{\alpha i} V_{\alpha j}^{*} = \delta_{ij} \tag{2.12}$$

from $V_{\text{CKM}}^{\dagger} V_{\text{CKM}} = \mathbb{1}$ and

$$\sum_{i \in (d,s,b)} V_{\alpha i} V_{\beta i}^{*} = \delta_{\alpha\beta} \tag{2.13}$$

from $V_{\text{CKM}} V_{\text{CKM}}^{\dagger} = \mathbb{1}$ with the Kronecker delta

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{2.14}$$

These constraints provide a strong test of the quark mixing mechanism and the SM in general.
The 6 equations where $i \neq j$ and $\alpha \neq \beta$ are of special interest. For three generations of quarks the equations can be depicted geometrically in complex space as triangles. The 6 triangles that can be formed by this are called unitarity triangles [17]. The most relevant triangle for B factory experiments, shown in Figure 2.4a, fulfills the equation

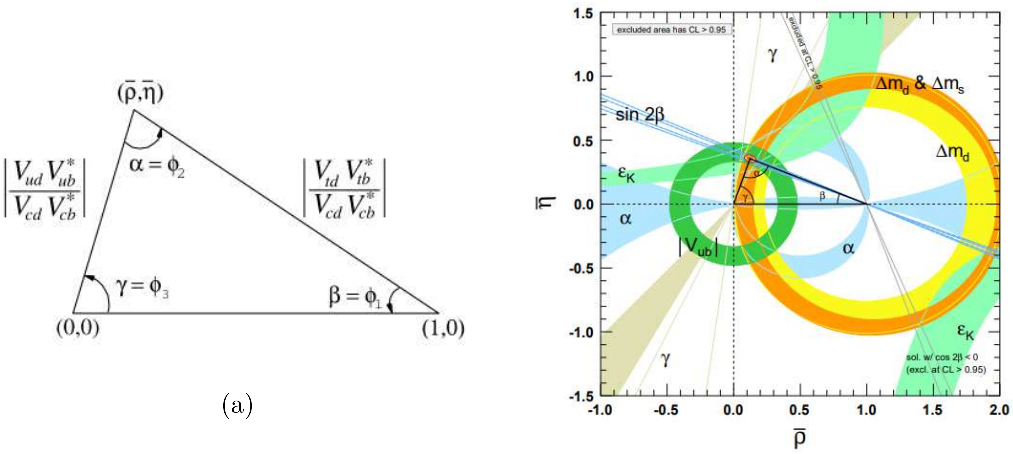
$$V_{ud} V_{ub}^{*} + V_{cd} V_{cb}^{*} + V_{td} V_{tb}^{*} = 0. \tag{2.15}$$

The lengths of the triangle sides are proportional to CKM element ratios, while the angles are determined by the complex phases inbetween them. In order for the theory to be validated, the triangle has to close exactly. Thus any deviations from a closed triangle imply flaws in the SM. Combining constraints from experiments such as Belle, BaBar, LHCb among others results in the current experimental status seen in Figure 2.4. Measurements of CKM matrix elements give constraints on side lengths, shown in colored circles. Measurements of the CKM angles are graphed with lines. The independent measurements of angles and sides are consistent and give strong validation to the quark mixing mechanism and the Standard Model.

2.4 Semileptonic B decays

Particle decays can generally be divided into three categories based on their decay products:

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a)

(b)

Figure 2.4: (a) The B meson unitarity triangle with side lengths and angles determined by CKM matrix elements [12].(b) The shaded 95% confidence interval experimental constraints yield exceptional agreement between theory and experiment [12].

1. Hadronic decays: the decay products are purely hadronic. Only states made up of quarks, most commonly mesons, are formed.
2. Leptonic decays: a particle decays into a combination of charged leptons and neutrinos.
3. Semileptonic decays: the decay products are a combination of leptonic and hadronic particles.

In principle, CKM matrix elements such as V_{ub} and V_{cb} can be measured from any of these types of decays.
Purely hadronic decays, such as $B^0 \to D^-\pi^+$ shown in Figure 2.5a, have an experimental advantage over decays containing leptons, because there are no invisible neutrinos that make measurement more difficult. However, in theoretical calculations of the matrix elements via hadronic decay branching fractions, hadronic currents between final state particles have to be taken into account. This leads to large correction terms, complicating the theoretical calculations.
Hadronic currents can be completely eliminated by measuring the matrix elements from purely leptonic decays, such as $B^-/B_c^- \to \mu^-\overline{\nu}_\mu$ shown in Figure 2.5b. The disadvantage of this method are the small branching fractions of leptonic decays. Searches for $B^- \to \mu^-\overline{\nu}_\mu$ in the Belle experiment were inconclusive, finding the decay at an excess of 2.4 standard deviations above background levels [18]. A much larger data sample is needed to measure the branching fraction precisely.
Finally, semileptonic decays such as $B^- \to D^0\ell^-\overline{\nu}_l$ or $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ offer a compromise between theoretical and experimental limitations. Because the hadronic final state products are contained, the currents are reduced in comparison to purely hadronic decays. The correction factors are reduced significantly, allowing for more precise calculation.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

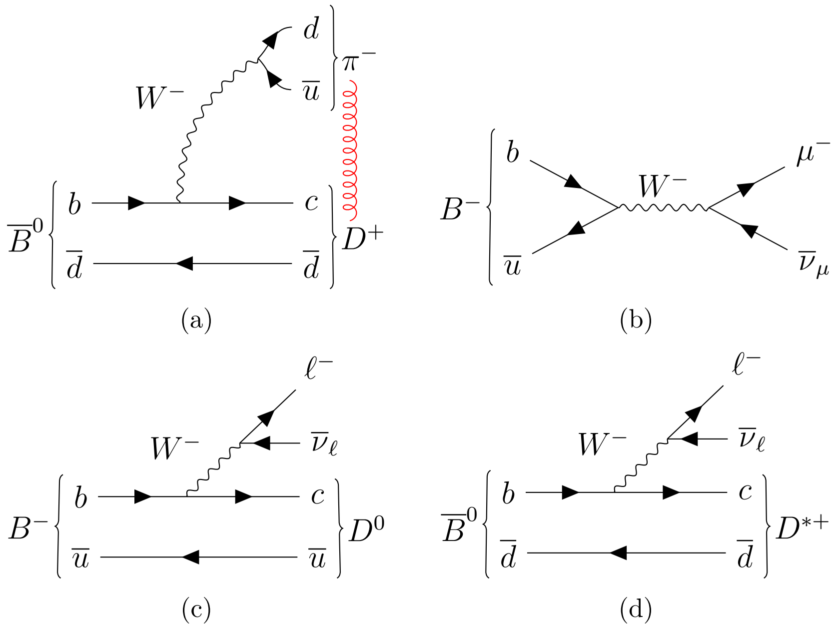


FIGURE 2.5: (a) Hadronic currents between final state particles makes CKM matrix element calculations via purely hadronic decays such as the pictured $\overline{B} \to \pi^- D^+$ difficult. (b) The purely leptonic decay $B^- \to \mu^- \overline{\nu}_\mu$ offers clean calculation of V_{ub}, but the branching fraction is too low to be measured precisely. (c) (d) Feynman diagrams of the semileptonic "golden modes" $B^- \to D^0 \ell^- \overline{\nu}_l$ and $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$ for measuring the CKM element V_{cb}. The coupling vertex between b and c quarks makes these decays sensitive to V_{cb}.

The focus of this thesis is the semileptonic decay $B^- \to D^0 \ell^- \overline{\nu}_l$ pictured in Figure 2.5c. In this decay, a charged B meson ($b\overline{u}$ or $\overline{b}u$) decays to a neutral D meson ($c\overline{u}$ or $\overline{c}u$), giving off a charged lepton and its corresponding antineutrino in the process. Because of the $b \to c$ transition, this decay depends on the V_{cb} matrix element. Thus by measuring the $B^- \to D^0 \ell^- \overline{\nu}_l$ branching fraction, the CKM matrix element can be determined.

Because of the rather large branching fractions of $B^- \to D^0 \ell^- \overline{\nu}_l$ (2.35%) and the closely related decay $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$ (5.05%), measuring these decays is a primary semileptonic method for determining V_{cb} and testing the unitarity triangle.

The resulting D^0 mesons have a lifetime of $\sim 10^{-13}$ s after which they can decay in a variety of different modes. In this thesis the branching fraction is measured for the case of D^0 decaying into a kaon K^- ($\overline{u}s$) and a pion π^+ ($u\overline{d}$) [12]. The full decay chain can then be written as $B^- \to [D^0 \to K^- \pi^+]\, \ell^- \overline{\nu}_l$. Throughout this thesis, any decay written includes the charge conjugated decay.

2.4.1 Inclusive and exclusive measurements

Another motivation for measuring V_{cb} via the $B^- \to D^0 \ell^- \overline{\nu}_l$ branching fraction is an anomaly in recent measurements. There are two ways to calculate CKM elements via semileptonic decays:

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

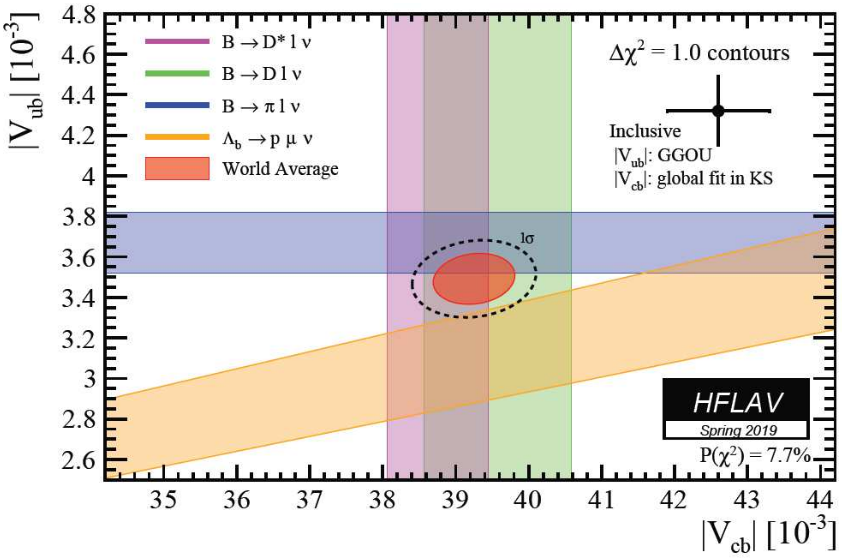


FIGURE 2.6: Exclusive V_{cb} and V_{ub} measurements are depicted as colored bands, with the red ellipse showing the combined average of exclusive measurements. The combined inclusive measurement is shown as the black point with error bars in the upper right quadrant. The anomaly is at a level of 3.3σ for $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ and 2.0σ for $B^- \to D^0\ell^-\overline{\nu}_l$ [19].

- **Exclusive decays.** In an exclusive reconstruction a specific decay mode such as $B^- \to D^0\ell^-\overline{\nu}_l$ or $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ is chosen. Its branching fraction is calculated by reconstructing the decay mode and the matrix element is calculated from lattice QCD.

- **Inclusive decays.** Inclusive reconstruction includes all possible decays $B \to X\ell\nu_\ell$ where X can be any resulting hadronic state containing a u quark for measuring V_{ub} or containing a c quark for measuring V_{cb}. The matrix elements are calculated in heavy quark effective theory.

The resulting CKM elements V_{ub} and V_{cb} from both methods show significant discrepancies at a combined order of 3σ [19]. The discrepancy is illustrated in Figure 2.6. The exclusive $B^- \to D^0\ell^-\overline{\nu}_l$ contribution is pictured as the green band. Belle II is looking to narrow the band by significantly reducing the measurement uncertainty in exclusive $Br(B^- \to D^0\ell^-\overline{\nu}_l)$ measurements through larger data samples and improved sensitivity.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Chapter 3

Experimental setup

This chapter gives an introduction to the features and setup of the SuperKEKB accelerator and the Belle II detector. After briefly explaining the concept and advantages of B factories as well as the various beam backgrounds that can be encountered, the individual sub-detectors are presented. An outline of the present and future data taking plan is given. Finally, the Belle II software framework and generators for simulated Monte-Carlo events are introduced briefly.

3.1 Particle colliders

Particle colliders allow probing of the standard model and measure its parameters. By accelerating particles to high kinetic energies, a collision frees up energy to produce particles that otherwise rarely occur in nature, such as antimatter or baryons containing second or third generation quarks.

Only particles with a rest mass lower than the collider center-of-mass energy $\sqrt{s} = E_1 + E_2$, with E_1 and E_2 being the colliding particle energies, can be produced.

An important measure of particle colliders performance besides the center-of-mass energy is the luminosity $\mathcal{L}$. Considering a particle collider based on an interaction with cross section σ_{col} and an event rate dN/dt, the instantaneous luminosity is defined as

$$\mathcal{L} = \frac{1}{\sigma_{\text{col}}} \frac{dN}{dt}. \tag{3.1}$$

This gives an expectation of how many events of a given decay can be expected. The integrated luminosity

$$\mathcal{L}_{\text{int}} = \int \mathcal{L}\, dt \tag{3.2}$$

gives a measure of the total experimental data taken over a given time frame. With this, the expected amount of events for an arbitrary decay with cross section σ in a dataset of luminosity $\mathcal{L}_{\text{int}}$ is simply $\langle N \rangle = \mathcal{L}_{\text{int}} \times \sigma$.

3.2 B factories

The most commonly used technique to produce a large amount of B mesons is by colliding electrons and positrons at a center of mass energy of $\sqrt{s} = 10.58$ GeV. The center-of-mass energy is chosen to correspond with the rest mass

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

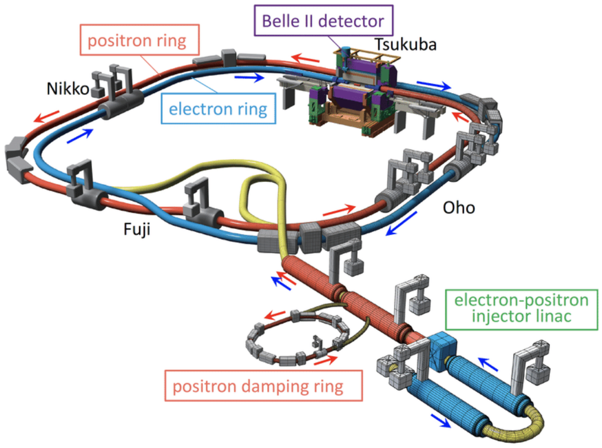


Figure 3.1: Schematic overview of the SuperKEKB collider located in Tsukuba, Japan [20].

of the $\Upsilon(4S)$ particle to achieve *on-threshold production*. The $\Upsilon(4S) \rightarrow B\overline{B}$ branching fraction is measured at over 96%, making this collider setup ideal for studying flavor physics via B meson decays. Because of this high B meson purity, B factories enable high precision measurements of rare decays. Another advantage of producing B mesons this way is the precisely known center of mass energy. Because of that, the missing mass in the event can be calculated easily, enabling measurements of decays with "invisible" particles such as neutrinos. or hypothesized dark matter candidates.

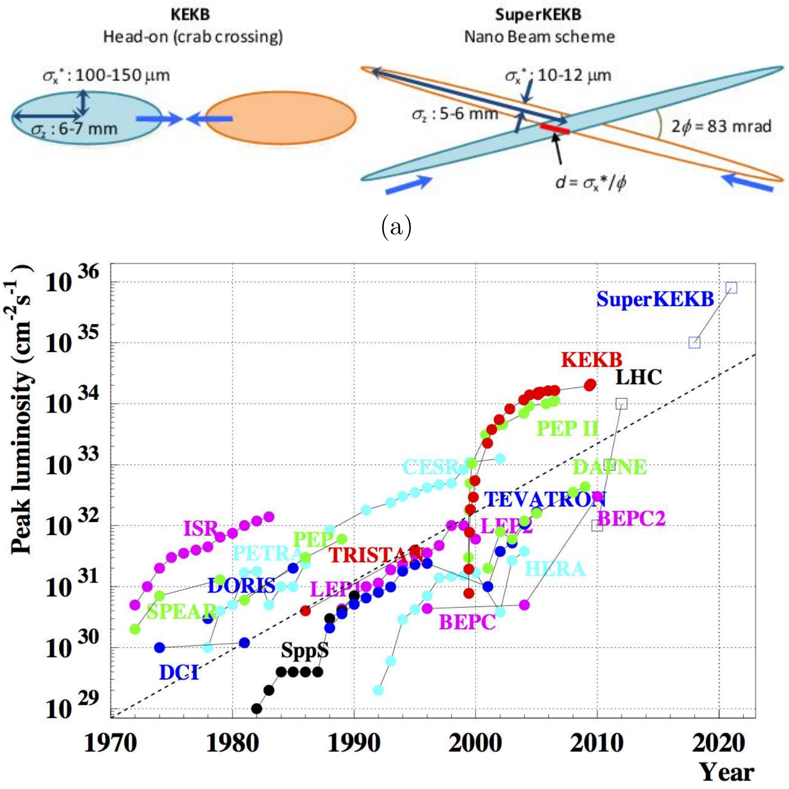
Because the $\Upsilon(4S)$ rest mass is very close to two times the B meson rest mass $2\ m_B = 10.56$ GeV, a resting $Y(4S)$ produces B mesons almost at rest. Due to their short lifetime $\tau_B \sim 1.5 \times 10^{-12}$ [12] the B mesons would only travel a short distance before decaying with little spatial seperation. The decay products of the two mesons would be hard to seperate. To avoid this problem, B factories are operated as asymmetric $e^+ e^-$ colliders, with a higher energy e^- beam and a lower energy e^+ beam. This way the $\Upsilon(4S)$ is boosted along the beam axis and as a result of this, the B mesons build enough spatial seperation before decaying, to accurately seperate which decay products originate from which B meson.

3.3 SuperKEKB

The SuperKEKB collider [21][22] located in Tsukuba, Japan is the latest generation B factory. Succeeding the KEKB collider used in the Belle experiment, SuperKEKB operates in a similar way with upgrades in many areas. A schematic overview of the collider is shown in Figure 3.1.

SuperKEKB collisions take place in the same tunnel as KEKB. A photo-cathode high-current radio-frequency gun generates electron bunches with charges of up to 4 nC. To generate equal charge positron bunches, a thermionic gun emits 10 nC electron bunches aimed at a 14 mm thick tungsten target. The positrons produced in the process are captured and focused by pulsed magnetic coils and

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a)

(b)

FIGURE 3.2: (a) Illustration of the nano-beam scheme geometry used in SuperKEKB [23]. (b) The design peak luminosity of SuperKEKB 8×10^{35} cm^{-2} s^{-1} is increased by a factor of 40 over previous record luminosities recorded by KEKB and LHC [24].

apperture accelerating structures, producing 4 nC positron bunches [21].
The resulting positron and electron bunches are accelerated via a linear accelerator and injected into seperate rings. The electrons are accelerated to 7 GeV before injection into the high energy storage ring (HER). The positron target is located within the linear accelerator, causing the accelerated electrons to produce positrons. The positrons are then accelerated to 4 GeV before being injected into the low energy storage ring (HER). This energy asymmetry corresponds to a Lorentz boost of $\beta\gamma = 0.28$ which leads to a B meson seperation of order $\Delta z \sim c \times \beta\gamma \times \tau_B \sim 130$ µm. The boost is reduced compared to KEKs $\beta\gamma = 0.42$, leading to a lower seperation between vertices. This is compensated by detector upgrades giving improved spatial and temporal resolution.
The electron positron beams collide at the interaction point (IP) located at the Belle II detector.

3.3.1 Nano-beam scheme

SuperKEKB is the first accelerator to employ the *nano-beam scheme* [22]. Through squeezing of the beam in the vertical direction and increase of the

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

crossing angle, the instantaneous luminosity can be improved significantly. By using doublets of vertical and horizontal focus quadrupole magnets, the beta function at the interaction point β_y^*, which measures the transverse beam size, is reduced from about 5.9 mm to 0.3 mm [25]. This factor 20 improvement, in addition with a factor 2 improvement of beam currents, leads to a total luminosity improvement of factor 40 compared to the Belle experiment. This way the instantaneous luminosity reaches 8×10^{35} $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at maximum beam squeezing. The nano beam-scheme with the resulting instantenous luminosity compared to other particle colliders is illustrated in Figure 3.2.

3.3.2 Beam backgrounds

Beam backgrounds are recorded particles that originate from interactions other than the desired electron-positron collision. With the drastically improved luminosity, the amount of beam background increases significantly. To ensure good performance and safety of the detector, a good assessment of possible beam backgrounds and their magnitudes is required. There are 5 main types of beam backgrounds [25][26]:

- **Touschek scattering.** The Touschek effect describes particle and energy loss in the beam due to Coulomb scattering of two particles within the same bunch. The momentum transfer of a intra-beam collision leads to one particle increasing energy and one decreasing energy, causing deviation from the nominal beam energy in both. After further propagation through the ring, Touschek scattered particles can exceed the longitudinal or transversal acceptance and consequently be lost.
  The Touschek scattering rate is inversely proportional to the beam size $\propto {\beta_y^*}^{-1}$ and thus increased by a factor of 20 compared to KEKB. To reduce Touschek background, collimators and metal shields are placed at different positions around the ring, stopping particles that leave the beam and preventing them from reaching the Belle II detector.

- **Beam-gas scattering.** Residual gas molecules in the beam pipe can interact with beam particles in two main ways: Coulomb scattering leads to transversal and longitudinal momentum transfer and changes direction of the beam particles. Bremsstrahlung scattering leads to energy loss of the beam particles. Due to the beam squeezing, beam-gas Coulomb scattering effects are expected to be amplified by a factor of 100 compared to KEKB, while the effects of Beam-gas Bremsstrahlung losses are negligible compared to Touschek scattering.
  The horizontal and especially vertical collimators used to suppress Touschek backgrounds are also effective at reducing Beam-gas backgrounds.

- **Synchrotron radiation.** Beam particles in the pipes get accelerated radially by bending and focusing magnets, leading to emission of synchrotron radiation. When unhindered, the emitted photons can reach the Belle II detector and, while not energetic enough to produce e^- e^+ pairs, can create hits in the detector. To reduce this background the inner wall

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

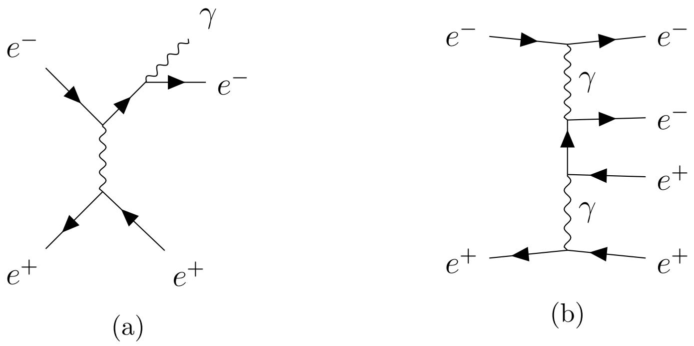


FIGURE 3.3: Scattering between beams at the interaction point leads to additional beam background. (a) Radiative Bhabha scattering emitting a low angle photon. (b) Two photon process producing additional e^- e^+ pairs.

of the interaction point beam pipe is covered with a gold layer, absorbing the synchrotron photons.

- **Radiative Bhabha scattering.** At the interaction point, electrons and positrons can scatter and produce additional photons as illustrated in Figure 3.3a. These photons can produce electromagnetic showers in the detector, diluting the measurements. Due to the typically small scattering angles, only a small amount of photons reach the detector. However, Bhabha photons scattered at a low angle can interact with the magnet iron, producing a large amount of neutrons via the photo-nuclear resonance mechanism. A dedicated shielding system in the accelerator tunnel mitigates the amount of neutrons reaching the outermost detector, where they make up the main background source.

- **Two photon process.** Similar to Bhabha scattering, electrons and positrons at the IP can produce e^- e^+ pairs in e^- e^+ $\rightarrow$ e^- $e^+ e^-$ e^+ processes. This process can be seen in Figure 3.3b The resulting e have low momentum and can spiral in the inner region of the detector, leaving excess hits. Bhabha scattering and two photon processes increase with increasing luminosity, which is why they are sometimes grouped as *luminosity backgrounds*.

In the first operating phase of the Belle II experiment in 2016, a seperate detector called BEAST 2 (Beam Exorcism for A Stable Belle II Experiment) has been used to measure beam background levels and ensure that the Belle II detector can be installed safely [27]. All main backgrounds were able to be observed and the Belle II installation was judged to be safe.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

3.4 Belle II detector

Belle II is the detector located at the interaction point of the SuperKEKB collider. Optimized for $\Upsilon(4S) \rightarrow B\overline{B}$ events, it has the capability to reconstruct charged tracks of produced particles. A superconducting magnet provides a 1.5T magnetic field, curving charged particle paths and thus allowing momentum measurements. Energies of neutral and charge particles can be measured from their interaction with detection material. Particle types can be identified via a sophisticated particle identification system (PID). Vertex reconstruction allows precise time-dependent measurements. As a hermetic detector, all possible event decay products can be recorded and because the initial center-of-mass energy is known this allows for elaborate missing mass analysis. A trigger system rapidly decides what events are kept to cope with the large amounts of beam backgrounds.

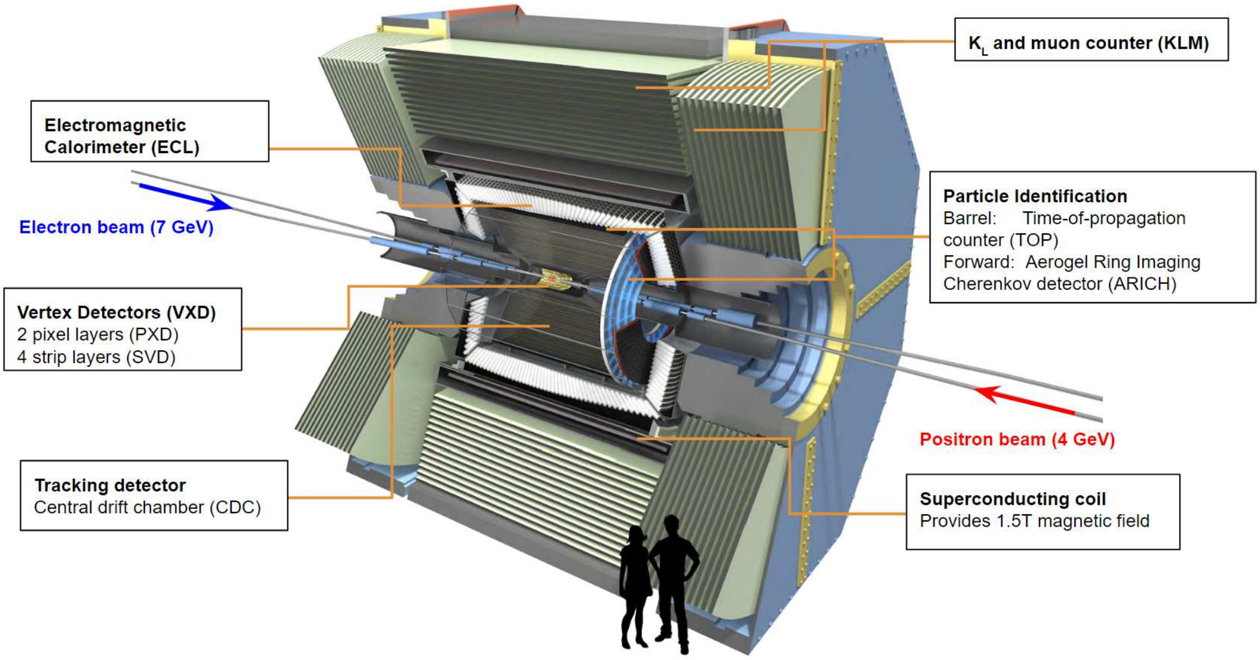


FIGURE 3.4: Cross section of the Belle II detector.

3.4.1 Detector overview

Belle II is made up of 7 subdetectors arranged in layers radially around the interaction point. The structure is shown in Figure 3.4.

The Belle II polar coordinate system origin is located at the IP. The z-axis is aligned along the beam pipe in the electron direction, the y-axis points to the top of the detector and the x-axis points outward of the detector ring. The azimuthal angle ϕ has its origin ϕ at $(x, y, z) = (1, 0, 0)$ and is in the $x-y$ plane. The zenith angle θ is in the $y - z$ plane and has its origin $\theta = 0$ at $(0, 0, 1)$. The three regions $17° < \theta < 30°$, $30° < \theta < 125°$ and $125° < \theta < 155°$ are called forward, barrel and backward regions respectively and cover the angular

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

acceptance of the detector. The asymmetry of the beam energies is reflected in the larger backward region.
The subdetectors from inner-most to outermost are:

- **Pixel detector (PXD).** 50µm thin pixels based on the Depleted field effect transistor technology (DEPFET) are aranged in two layers directly around the interaction point. Readout electronics record hits of charged particles passing through.
- **Silicon vertex detector (SVD).** The SVD consists of 4 layers of silicon strip detectors. Together with the PXD they form the Vertex Detector (VXD) covering the first 135mm around the IP.
- **Central drift chamber (CDC).** Wires spanned across a gas-filled chamber allow further tracking of charged particle paths through ionization of the gas.
- **Time-of-propagation counter (TOP).** A quartz radiatior propagates emitted Cherenkov photons to a photon detector. The time-of-propagation is proportional to the Cherenkov emission angle, which again is proportional to the particle mass. This allows for particle type identification.
- **Aerogel ring imaging Cherenkov counter (ARICH).** In the forward end-cap an aerogel layer produces Cherenkov photons and the emission angle is measured. ARICH and TOP combine to form the particle identification system (PID).
- **Electromagnetic calorimeter (ECL).** Over 8000 CsI crystals cause incoming particles to produce electromagnetic showers. Measuring the resulting photon energies allows for particle energy measurement.
- **K_L and muon detector (KLM).** Muons and long-lived kaons are not absorbed completely in the ECL. Alternated layers of iron plates and active material detectors allow for measurement of these particles.

Additionally, a superconducting magnet placed between ECL and KLM provides a homogenous magnetic field of 1.5T. This is essential for curving particle tracks. Particle momenta can be calculated from their track curvature.
A multi-part trigger system selects events interesting for physics analyses while discarding all other events. This is necessary to reduce the incoming data to levels that can be handled by the data acquisition system.
The following sections on the sub-detectors follow the Belle II Technical Design Report [25] and the Belle II Physics Book [26].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

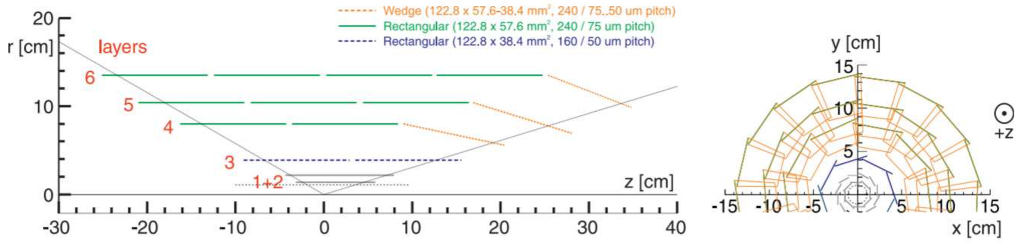


FIGURE 3.5: The configuration of PXD and SVD sensors [26]. PXD consists of layers 1 and 2 while layers 3 to 6 are made up of the SVDs silicon strip detectors. The z asymmetry compensates the electron and positron beam energy difference.

3.4.2 Pixel detector

Because of the large beam backgrounds at small radius, as listed in Section 3.3.2, the first layers require fine segmentation due to the extremely high hit rates. Thus 50µm thin pixel detectors are employed instead of silicon strips.
The DEPFET pixel sensors are semiconductor detectors that combine detection and amplification. Charged particles passing through generate electrons in the silicium layer of the pixel. A measured conductivity increase then gives evidence of particles passing through. Around the 10mm radius beam pipe, around 8 million DEPFET sensors are placed in two layers at 14mm and 22mm.* Due to the Lorentz boosted collisions, events in the forward region are favored over events in the backward region. Thus the angular coverage of the PXD spreads an asymmetric interval from 17° in the forward region to 155° in the backward region.
The readout electronics located outside of the acceptance region are capable of reading all pixels within a time frame of 20µs.

3.4.3 Silicon vertex detector

At a range unreachable by most of the low momentum beam background, pixel sensors can be replaced by double-sided silicon strip detectors (DSSDs). The SVD is made up of 4 layers of DSSDs, located at distances of 39 mm, 80 mm, 104 mm and 135 mm to the IP. Depending on the region, SVD uses sensors of different geometrical shape. As opposed to the pixel sensors, strip detectors measure two coordinates where a particle passed through the strip, rather than binary information on whether a particle passed. The edges of the sensors overlap by $5-18\%$, as can be seen in Figure 3.5, to provide information for alignment procedures.

3.4.4 Central drift chamber

The central drift chamber is a large-volume gas chamber, filled with He(50%) : C_2H_6(50%) gas mixture. 14436 tungsten sense wires are arranged throughout

---

*Only roughly half of the planned PXD sensors have been installed for the 2019 and 2020 runs after problems occured in the installation.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

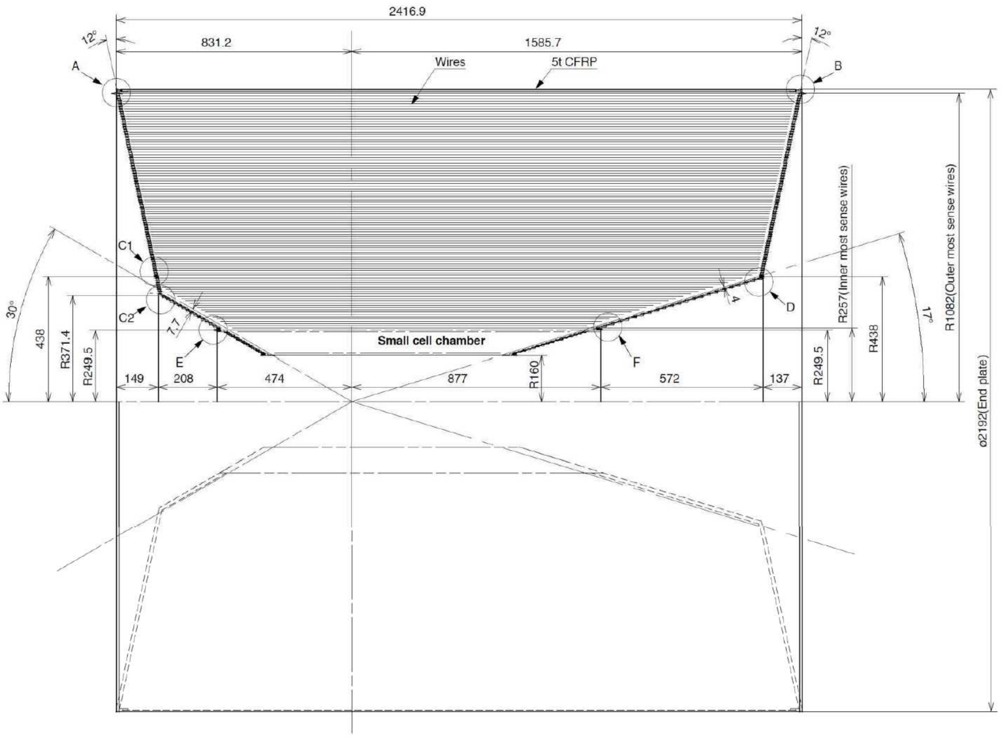


FIGURE 3.6: Schematic view of the central drift chamber [25]. All measures are in mm.

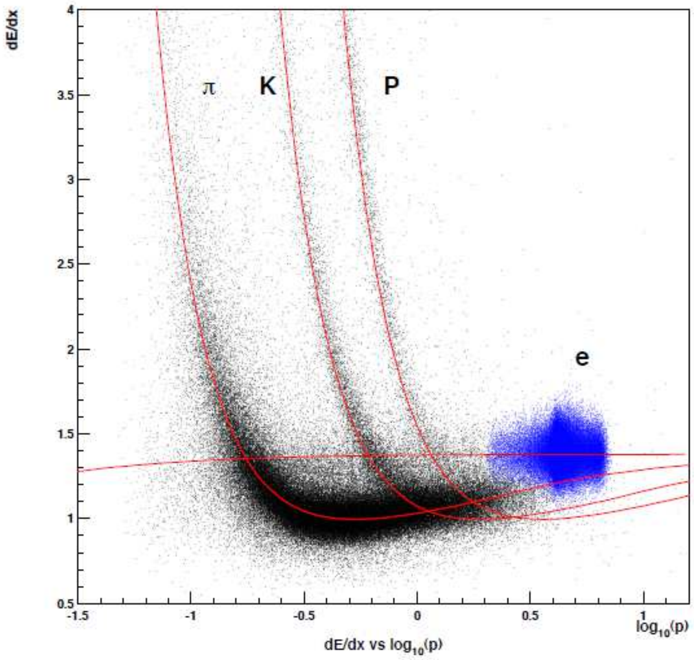


FIGURE 3.7: Different types of particles show different shapes of energy loss $\frac{dE}{dx}$ in the CDC observed in the Belle experiment [28].

the chamber, either aligned parallel to the beam axis or slightly tilted. Additionally, aluminium field wires generate an electric field. A schematic view of the CDC is shown in Figure 3.6. The CDC fulfills three major functions:

- **Momentum measurement.** Together with the PXD and the SVD, the CDC tracks charged particle trajectories to compute the momenta. 3D resolution is possible due to ionization of the gas as reaction to charged particles passing through. When a charged particle passes through the CDC it produces electrons. A uniform electric field shaped by field wires accelerates the electrons and produces a charge avalanche, inducing signal on the sense wires. Through drift time measurements in the sense wires the particles position can be reconstructed. A charged particles trajectory curvature in a magnetic field is proportional to its momentum, allowing for momentum measurement.

- **Particle identification.** Due to the gas ionization, charged particles traversing through the CDC continuously lose energy. This characteristic energy loss $\frac{dE}{dx}$ gives valuable information about the particle type. The different spectra observed in Belle are pictured in Figure 3.7.

- **Trigger.** Signals in the CDC contribute to the trigger system, essential for reducing the amount of background events. An overview of the trigger system is given in Section 3.4.8.

The measured spatial resolution of the CDC is about 100µm, while the energy loss $\frac{dE}{dx}$ can be measured to about 12% relative precision.

3.4.5 Particle identification

While the energy loss in VXD and CDC contributes to particle identification, the time-of-propagation counter and the Aerogel Ring Imaging Cherenkov counter are specifically dedicated to identifying charged particles.

Their fundamental principle is based on Cherenkov radiation. A charged particle traversing through a dielectric medium of refractive index n excites the molecules to higher energy levels. Photons are emitted when the molecules return to their ground states. When the particle moves faster than the speed of light in the medium $c_n = \frac{c}{n}$, the emitted photons interact constructively, leading to coherent radiation at a fixed Cherenkov angle θ_C proportional to the particles velocity $\cos(\theta_C) = \frac{vc}{n}$. This mechanism is shown in Figure 3.8. After measuring the particles momentum in the CDC, information about the Cherenkov radiation angle can be used to compute a particles rest mass, and thus identify the particle type.

Time-of-propagation counter

The Time-of-propagation counter consists of 16 45cm wide and 2cm thick quartz bars with sensors at one end. The main principle of the TOP is to calculate the Cherenkov angle θ_C by measuring the time of arrival of Cherenkov photons at the sensors. Charged particles flying through the quark bars emit photons

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

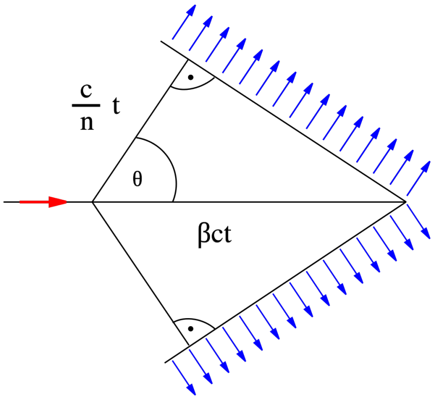
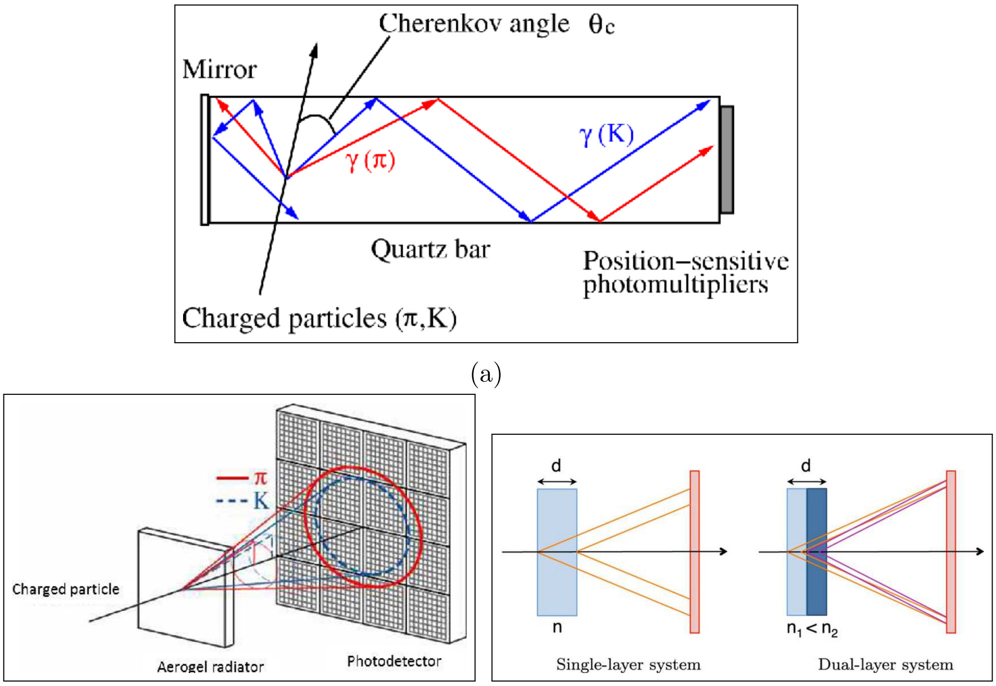


Figure 3.8: Emitted ionization photons by a particle moving faster than the speed of light in a medium add up coherently, leading to a "wavefront" of photons at a measurable angle θ.



(a)

(b)

(c)

Figure 3.9: (a) Quartz bars propagate Cherenkov photons to a position-sensitive photomultiplier in the TOP [29]. (b)(c) In the ARICH Cherenkov photons propagate freely to enlarge into rings. Belle II uses two aerogel layers of increasing refractive index to improve ring resolution.

at an angle dependent on their mass. The emitted photons propagate through the quartz bar by bouncing off the edges, as illustrated in Figure 3.9a. Photons get reflected more often at smaller θ_C, leading to longer time of flights. At the sensor end, an array of photomultipliers allows a position-sensitive measurement of the Cherenkov photons.
By combining information of charged particle arrival time at the TOP, photon position and time of flight, the Cherenkov angle θ_C is extracted. The TOP is most sensitive to differentiating K from π.

Aerogel ring imaging Cherenkov counter

The ARICH is the second sub-detector devoted to particle identification. Two 2 cm thick aerogel radiator layers oriented perpendicular to the beam are located in the forward end-cap. Their refractive indices differ slightly, $n_1 = 1.045$ and $n_2 = 1.055$. Photon detectors are located at a distance of 20 cm away from the aerogel radiators. Charged particles passing through emit Cherenkov photons in both aerogel layers. The photons spread apart while propagating over 20 cm, enlarging into rings, shown in Figure 3.9b.
The unique Belle II feature of using two aerogel layers with different refractive indices, shown in Figure 3.9c, increases the number of detected photons without degrading the ring resolution.

3.4.6 Electromagnetic calorimeter

The electromagnetic calorimeter consists of 6624 CsI scintillation crystals in the barrel-region and 2112 CsI crystals in the end-cap. It serves multiple functions:

- **Detection and energy measurements of photons.** Photons entering the ECL interact with the scintillator material, producing an e^+e^- pair. The leptons radiate additional photons via bremsstrahlung, which in turn continue to produce further e^+e^- pairs. This process repeats until the photon energy is below the threshold necessary to produce further pairs. This cascade of photons, electrons and positrons, shown in Figure 3.10, is called electromagnetic shower. Photons present in the event can thus be detected by observing showers and angular coordinates can be extracted. The scintillation crystals translate the deposited energy into photons, allowing measurement of the incoming particle energy.

- **Electron identification.** Electrons also produce electromagnetic showers depositing their entire energy in the ECL. Minimally ionizing particles, such as the heavier muons, pass through the ECL without fully depositing their energies. This way shower shapes can be used to identify and distinguish electrons from other particles.

- **Detection and energy measurements of neutral hadrons.** Neutral particles, that make up for over a third of B-decay products, decay in large parts to photons. Detection, energy measurement and identification of π^0 from the photon clusters from its primary decay mode $\pi^0 \to \gamma\gamma$ is especially important.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

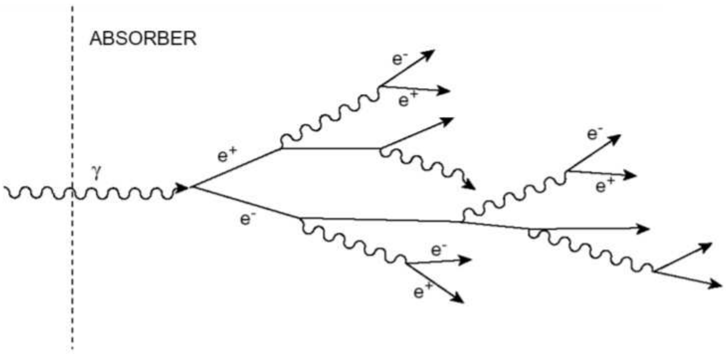


FIGURE 3.10: Schematic diagram of the electromagnetic showers created by photons entering the ECL crystals.

- **Trigger.** Similar to the CDC, clusters in the ECL contribute to the hardware trigger. The trigger system is covered in Section 3.4.8.

3.4.7 K_L and muon detector

Long-lived neutral Kaons K_L^0 and muons above momentum of 0.6 GeV pass through the ECL. The KLM is the outer-most detector, located outside of the magnet and has the purpose of detecting K_L^0 and $\mu^{\pm}$. It extends into the endcaps of the detector, to reach angular acceptance of $20° < \theta < 155°$.

The KLM consists of alternating sensor layers and iron plates. The 4.7 cm thick iron plates serve two purposes: they return the magnetic flux produced by the superconducting coil and they allow K_L to shower hadronically. Resistive plate chambers (RPC) in the barrel region of the KLM detect the hadronic showers and allow for identification of K_L, since muons do not shower hadronically. Due to the high background rates in the endcaps, scintillator strips are used for shower detection instead.

Muons can be identified effectively by extrapolating charged tracks in the CDC to the KLM. Charged tracks that continue through the ECL into the KLM and do not shower hadronically can be identified as muons. A likelihood can be computed based on the goodness of the track fit, among other factors.

3.4.8 Trigger

Beam bunches cross roughly every 8 ns. Although the physics event rate of 15 kHz is lower than the maximum data taking frequency of 30 kHz, the significantly increased beam backgrounds require a stringent trigger system, to decide which events are recorded. Because $B\overline{B}$ events have a high charge track multiplicity compared to beam backgrounds like Bhabha scattering, the number of tracks in the event gives a strong trigger identifier. However, to achieve a high $B\overline{B}$ efficiency, low multiplicity events such as τ decays or potential dark matter decays have to be recorded, requiring additional conditions other than the number of tracks.

The Belle II trigger consists of a hardware-based Level 1 trigger (L1) and a software-based High Level Trigger (HLT).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

- **L1 trigger.** The L1 trigger merges signals from various sub-detectors to make a trigger decision. The CDC gives charged track information, ECL gives the number and property of clusters, the KLM gives information on muon hits and the seperate particle identification in barrel (BPID) trigger module provides precise timing information. The subdetector information is sent to the global reconstruction logic (GRL), which provides an output to the global decision logic (GDL).
  The Belle II L1 trigger reduces the incoming events to an event rate of maximum 30 kHz with a latency of about 5 µs. Over 99.9 % of $B\overline{B}$ and $e^+e^- \to q\overline{q}$ $(q = u, d, c, s)$ (continuum) events are retained.

- **High Level Trigger.** For offline data storage, the HLT has to reduce incoming event rates of 30 kHz to roughly 10 kHz. For the pre-selected L1 output events, the HLT reconstructs the event fully and computes various event variables. Different physics modules make a decision on recording the event.

3.5 Data taking plans

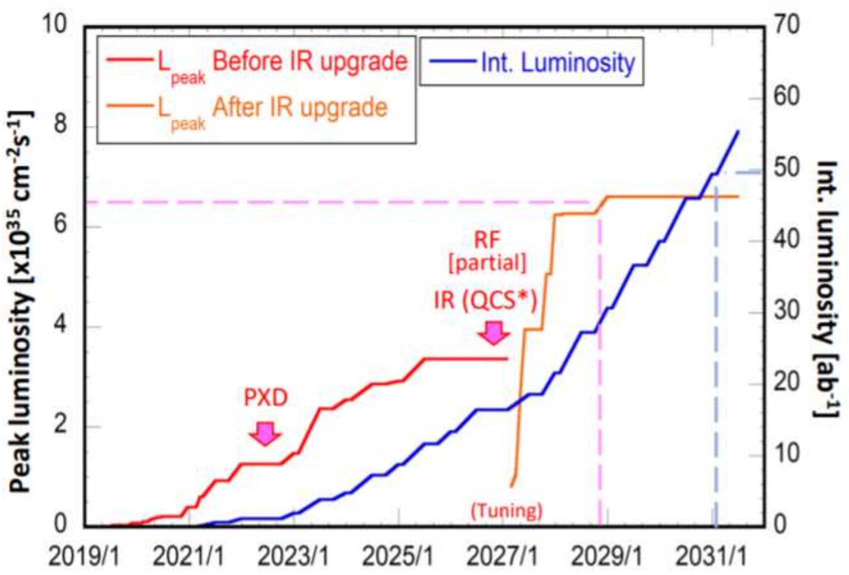


FIGURE 3.11: Current long-time luminosity projections for the Belle II experiment. Shutdowns to implement the full pixel detector and the IR (magnet and beam pipes) are planned before reaching the peak luminosity. The target integrated luminosity of 50 ab^{-1} is planned to be reached by 2031.

With the 40 times improved instantenous luminosity, the Belle II experiment is planning to collect 50 ab of data, corresponding to over 5×10^{10} $B\overline{B}$ pairs, compared to Belle's $\sim$ 0.7 ab. After two commissioning runs in 2016 (*Phase I*) and 2018 (*Phase II*) to probe the beam backgrounds, calibrate the detector and establish the nano-beam scheme, the physics runs (*Phase III*) began in 2019. The long-time luminosity projection, covering the next decade, is shown

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

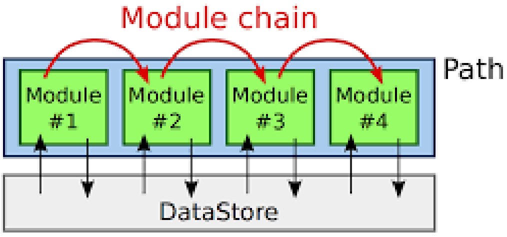


FIGURE 3.12: Individual modules arranged linearly form a BASF2 path. Each module can interact with the DataStore to read and write event information.

in Figure 3.11.
On June 15th 2020, SuperKEKB achieved an instantenous luminosity of 2.22×10^{34} $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ surpassing peak luminosities of KEK and LHC to set a new world record [30].

3.6 Belle II Analysis Software Framework

After event data is stored offline, the Belle II Analysis Software Framework (BASF2) is used for analysis [31]. BASF2 is built by individual modules arranged linearly in a *path*, shown schematically in Figure 3.12. The individiual modules interact with the data storage to perform self-contained tasks, such as reading and writing of information to files, simulating particle decays and detections and reconstructing certain decay modes.

The desired modules are written in a Python streering file, creating a path and executing it linearly. Information such as particle tracks, momenta and cluster information are stored in the common storage DataStore. The DataStore is built from an input file, either from simulated events or real Belle II collision data. The DataStore contains lists of tracks, track fitting results, ParticleID likelihoods, cluster information, trigger information, etc. In the case of a simulated input file, information about the true particles involved are available.
The details on reconstrucing a chosen decay mode from the DataStore objects are elaborated on in Section 4.1.

3.7 Event simulation (Monte Carlo)

Simulating collisions via the Monte Carlo method (MC) is an important part of high energy particle physics analyses. Based on known decay probabilities a large amount of collisions is generated. Opposite to real collisions where particle types can only be identified up to a certain likelihood, the particle types in simulated events are known. By generating decay chains according to world average branching fractions and simulating detector efficiencies, the sensitivity of physics analyses can be probed. Additionally, distributions of

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

physical variables can be seen.

The Belle II MC simulations use `EvtGen` originally developed for the BaBar and Belle experiments to generate B decays [32]. Branching ratios are provided via a decay table listing all possible decays. Additional algorithms such as `PHOTOS`, `PYTHIA` and `BABAYAGA.NLO` are used to simulate further details of decays such as Bremsstrahlung photons and beam backgrounds [26].

The generated events are then supplied to the BASF2 simulation package to simulate detector effects. The simulation package contains modules based on the Geant4 software [33]. Geant4 simulates transporting the particles generated by `EvtGen` through the detector, translating the events to list of tracks and clusters stored in the DataStore. The output is equivalent to stored data from a real physics collision, with additional knowledge of precisely what decay chains are involved in the B decays.

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Chapter 4

Belle II Data Analysis

This chapter elaborates how the data of the Belle II experiment is used to reconstruct B meson decays. Basic decay reconstruction is explained and the necessary variables are introduced. In the second half of this chapter, the statistical methods and principles necessary for the analysis of experimental data are introduced, including data selection, maximum likelihood fitting and error optimization.

4.1 Event reconstruction

As elaborated in Chapter 3, the main goal of the Belle II experiment is to probe the standard model via B meson decays. The SuperKEKB collider produces $\Upsilon(4S)$ mesons which decay into pairs of B mesons. A charged pair of B mesons B^+B^- is produced with a fraction of [19]

$$f^{+-} = \frac{\Gamma\left(\Upsilon(4S) \to B^+B^-\right)}{\Gamma_{\text{tot}}\left(\Upsilon(4S)\right)} = 0.514 \pm 0.006, \tag{4.1}$$

where Γ and Γ_{tot} refer to the partial and total decay widths[*], while neutral pairs of B mesons $B^0\overline{B^0}$ are produced with slightly disadvantaged fraction

$$f^{00} = \frac{\Gamma\left(\Upsilon(4S) \to B^0\overline{B^0}\right)}{\Gamma_{\text{tot}}\left(\Upsilon(4S)\right)} = 0.486 \pm 0.006. \tag{4.2}$$

The produced B mesons travel approximately 130µm before decaying into a cascade of further particles. An example is pictured in Figure 4.1. Hadronic B decays result in an average of 10 charged tracks and 10 neutral particles. Energy and momentum measurements are provided by the ECL, VXD and CDC. Because the vertex detector resolution is not precise enough to see the decay topology, reconstructing which decay mode ocurred has to be inferred from final state particle energies and momenta alone.
To reconstruct a decay such as $B^+ \to J/\psi\ \pi^+$, with the J/ψ meson further

---

[*] A decay width is defined as the probability per unit time that a given particle will decay, either in a specific decay mode or in total. The inverse of the decay width is the lifetime $\tau = \frac{1}{\Gamma}$

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

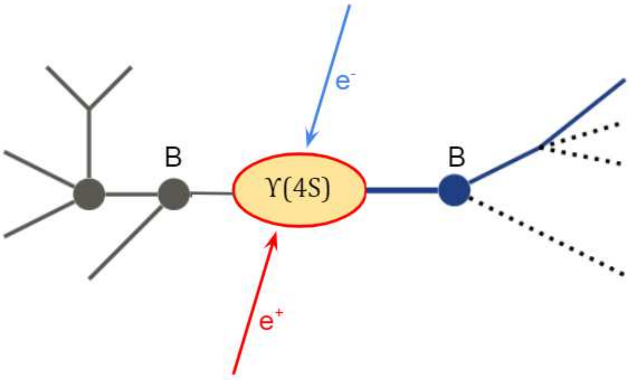


FIGURE 4.1: $\Upsilon(4S)$ mesons produced by $e^- e^+$ collisions decay into a pair of B mesons over 96% of the time. Pictured is an example process of 2 B mesons decaying further into various combinations of leptonic and hadronic final state particles.

decaying into a pair of electrons $J/\psi \to e^+ \ e^-$, first the J/ψ has to be reconstructed. All combinations of electrons and positrons* observed in the detector are pieced together to build a J/ψ candidate. The invariant mass of the electron positron combination

$$m_{e^-e^+} = |p_{e^-} + p_{e^+}| = \sqrt{(E_{e^-} + E_{e^+})^2 - (\vec{p}_{e^-} + \vec{p}_{e^-})^2} \tag{4.3}$$

has to be close to the nominal J/ψ mass $m_{J/\psi} = 3.097$ GeV. Only $e^- e^+$ combinations within a certain mass range of $m_{J/\psi}$ are used for further analysis. Repeating the same process to form a B meson, all π^+ candidates are combined with the J/Ψ candidates to compute a B meson candidate with mass $m_B^{\text{cand}} = |p_{e^-e^+} + p_{\pi^+}|$. Once again restricting m_B^{cand} to be within a mass window of the nominal B meson mass $m_B = 5.279$ GeV gives a set of events where the decay $B^+ \to (J/\psi \to e^+ \ e^-) \ \pi^+$ likely ocurred.

There are various other requirements that can be set on events to further exclude event candidates from the resulting set. A non-exhaustive list of methods is given in the following subsections.

4.1.1 Particle identification

Many Belle II sub-detectors contribute information about the type of particle encountered. Combining all information, each track is associated with six likelihoods, one for each of the six most common charged particles (electrons, muons, pions, kaons, protons, deuterons) encountered:

$$\frac{\mathcal{L}_i}{(\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_d)} \tag{4.4}$$

---

*The particle identification system is not unambiguous, but rather gives a likelihood that a certain track is an electron. In practice, all charged tracks above a certain electron likelihood are treated as potential electrons and used to form a J/ψ candidate.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

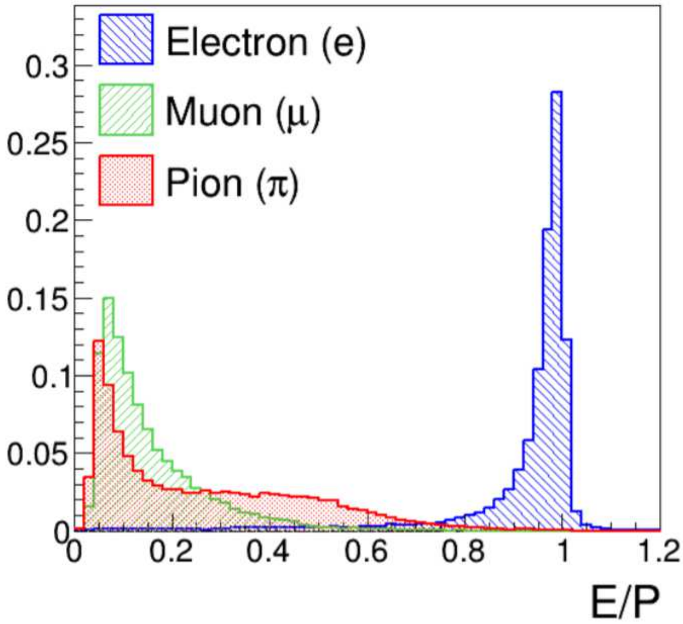


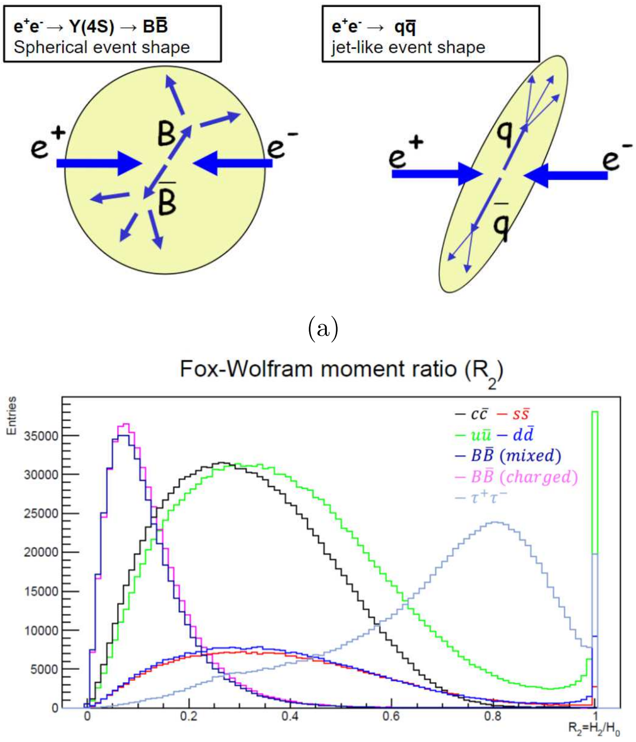
FIGURE 4.2: $\frac{E}{p}$ distributions measured in τ decays. The low $\frac{E}{p}$ is dominated by μ and π while the electron dominates the $\frac{E}{p} > 0.8$ region [34].

with $i = (e, \mu, \pi, K, p, d)$.
Particle identification can be divided into three groups. Hadron identification information, primarily distinguishing K and π mesons, is mainly gathered in the TOP for the barrel region and in the ARICH for the end-cap region. Electrons can be identified from ECL shower structures, due to them being fully contained within the ECL. Lastly, the most important information for muon identification is from the KLM. μ can be distinguished from K_L by their lack of hadronic showering.
The characteristic energy loss $\frac{dE}{dx}$ measured in the CDC and ECL gives an additional likelihood contribution. The differences in CDC $\frac{dE}{dx}$ distributions between particles are shown in Figure 3.7.
For a charged particle with a measured energy E based on the energy deposited in the ECL and a measured momentum p based on the charged track curvature, the ratio $\frac{E}{p}$ also contributes to particle identification. The electron and pion $\frac{E}{p}$ spectra differ significantly as shown in Figure 4.2, helping to distinguish them.
By requiring a minimum particle likelihood (particleID) for particle candidates, the reconstruction subset can be reduced further.

4.1.2 Continuum suppression

The cross section of $e^+e^- \to q\bar{q}$ (continuum) is approximately three times higher than the B meson production $e^+e^- \to \Upsilon(4S)$ and thus provide a large background called continuum background. Particle candidates originating from continuum events can pass kinematic selection criteria and be included in the resulting set of events.
To reduce the amount of candidates originating from continuum events, the

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a)

(b)

FIGURE 4.3: (a) $B\overline{B}$ events have a more spherical event shape than continuum events. (b) Distributions of $B\overline{B}$ events compared to different channels of continuum background. By only selecting events with low R_2 the majority of continuum events can be rejected [36].

differing event shapes shown in Figure 4.3, can be exploited. Because the B mesons are produced almost at rest, the event shape is much more spherical than in the case of lighter quark pairs $q\overline{q}$, which are produced with a significant boost.

The Fox-Wolfram moments [35]

$$H_l = \sum_{i,j} \frac{|\vec{p_i}||\vec{p_j}|}{s} P_l(\cos\,\phi^*_{ij}) \tag{4.5}$$

summed over all particles i, j in the event, with the center-of-mass energy s, the angle between particles i and j ϕ^*_{ij} and the 1^{th} Legendre polynomial P_l, and especially the ratio between 2^{nd} and 0^{th} Fox-Wolfram moment

$$R_2 = \frac{H_2}{H_0} \tag{4.6}$$

gives a measure of event sphericity. For spherically uniform distributed events R_2 is close to 0 while for jet-like events R_2 is close to 1. By only selecting events with low R_2, continuum events can be surpressed effectively.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

4.1.3 Veto

In some reconstruction cases, a specific decay gives a substantial background. *I.e.* in the decay $B^0 \rightarrow \rho^0\gamma$ the photon candidate might originate from a $\pi^0 \rightarrow \gamma\gamma$ decay.

As similarly done in Section 4.1, by seperately combining photon candidates to form a π^0, $m_{\gamma\gamma}$ can be calculated. A veto can be introduced by discarding all events where the photon candidate can be used to form a π_0 with a mass within an acceptance window of the nominal π^0 mass. Other requirements such as γ shower shape can be required additionally. This way, it is possible to significantly reduce the background coming from $\pi^0 \rightarrow \gamma\gamma$ photons.

4.1.4 Full Event Interpretation (tagged and untagged)

In general B meson decays from $\Upsilon(4S) \rightarrow B\overline{B}$ can be reconstructed through two methods called *tagged* and *untagged* reconstruction.

In untagged reconstruction a decay chain is reconstructed and daughter particles are combined to form a signal candidate B_{sig}. All other tracks in the event are ignored and treated as products of the second decaying companion B_{comp} meson. Adding up all the unused tracks and clusters allows measurement of the B_{comp} momentum, and with it the missing momentum $p_{\text{miss}} = p_{\text{beam}} - p_{B_{\text{sig}}} - p_{B_{\text{comp}}}$ can be estimated. Due to potential invisible neutrinos originating from B_{comp} the missing momentum resolution is limited to about $\mathcal{O}(10\ \%)$ and thus untagged analysis are better suited for smaller data samples.

In tagged analyses both B mesons are fully reconstructed using the Full Event Interpretation algorithm (FEI). In addition to reconstructing B_{sig}, the second B meson, B_{tag} is reconstructed via a number of exclusive decay modes. By reconstructing B_{tag} explicitely, the entire $\Upsilon(4S)$ decay chain is known, allowing for precise determination of the neutrino momentum in semi-leptonic decays. The downside of tagged analyses is the low efficiency. For semi-leptonic decays the FEI is able to fully reconstruct $\mathcal{O}(1\%)$ of events, significantly reducing the data sample size.

The analysis in this thesis uses the first 34.6 fb^{-1} of the Belle II experiment, equivalent to roughly 0.1% of the planned final data set. Thus, an untagged analysis is chosen.

4.2 Multivariate analysis

It is not immediately obvious what selection of events is ideal for further analysis. To get the best possible measurement, the statistical and systematic uncertainties should be reduced as much as possible. To estimate and optimize uncertainties multivariate analysis (MVA) offers a wide array of tools. When reconstructing a particle decay from experimental data, the data has to be categorized into “signal”, events where the sought-after decay ocurred, and “background”, made up of all other events. A measurement is always associated with a statistical uncertainty and a systematic uncertainty. MVA methods enable

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

reduction of background events and the statistical uncertainty.

4.2.1 Figures of merit

To select a dataset with minimal statistical uncertainty, a trade-off has to be made between the purity

$$\rho = \frac{\text{\# of signal events in the dataset}}{\text{\# of events in the dataset}}, \tag{4.7}$$

i.e. the ratio of true positives and true plus false positives, and the efficiency

$$\epsilon = \frac{\text{\# of signal events in the data subset}}{\text{\# of signal events in the total dataset}}, \tag{4.8}$$

i.e. the ratio of true positives and true positives plus true negatives[*].
To determine the data subset that leads to the smallest statistical uncertainty a common approach is to define a figure of merit (FOM) that quantifies the precision of a given dataset, and then maximize it [37].
Discrete counting processes with a fixed rate are Poisson distributed [4]. In a Poisson distribution with the expected value λ, the standard deviation is proportional to $\sqrt{\lambda}$. In a dataset of experimental data with S signal events and B background events, the standard deviation will be $\sqrt{N} = \sqrt{S+B}$. Optimizing the precision of the measurement can then be formulated as maximizing the figure of merit

$$FOM = \frac{S}{\sqrt{S+B}}. \tag{4.9}$$

Depending on the analysis there are other figure of merits that can be used instead. The approximation

$$\frac{S}{\sqrt{B}} \tag{4.10}$$

has been used in studies with $B \gg S$ in experiments such as the Belle $\tau \to \ell h h'$ analysis [38].
The Punzi figure of merit [39]

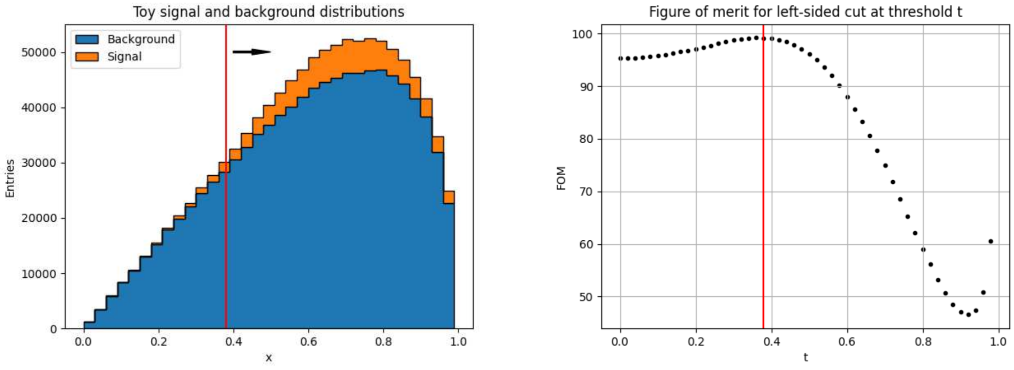
$$\frac{S}{n_\sigma/2 + \sqrt{B}}, \tag{4.11}$$

with the desired level of significance n_σ, has been used in searches for new physics such as the $B^+ \to \ell^+ \nu_\ell$ search at BaBar [40].
For rediscoveries of known processes, such as the analysis in this thesis, it is sufficient to optimize the FOM defined in (4.9).

---

[*] To give a number example, it is not intuitively obvious whether a dataset with loose restrictions containing 100 signal events and 1000 background events or a dataset with tighter restrictions containing 10 signal events and 10 background events leads to the measurement with higher statistical significance.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) (b)

Figure 4.4: (a) A generated toy data set with a gaussian distributed signal component and an argus distributed background component. The red horizontal line marks the ideal rectangular selection. (b) The resulting figures of merit $\frac{S}{\sqrt{S+B}}$ after removing events with $x < t$. The FOM is optimal for a cut at threshold $t = 0.38$

4.2.2 Rectangular selections

A simple method of increasing the figure of merit is a rectangular selection [37]. For a given variable x and a threshold t, cutting away regions below the threshold, such that every event in the new subset fulfills $x > t$ is referred to as a rectangular selection. By cutting away regions that contain more background than signal the FOM can be improved. This concept is illustrated in Figure 4.4. Rectangular selections can be applied simultaneously on different variables.

4.2.3 Neural nets

If correlations between variables exist rectangular selections may no longer be ideal. A selection dependent on multiple variables, for example the ellipsoid selections pictured in Figure 4.5, can lead to a higher FOM than the simple rectangular selection, by cutting more closely around the signal. In the general case of many variables, it is not trivial to find the optimal multivariate selection. A MVA method that can be used is the neural network [37]. Aiming to recognize patterns through a model of simulated neurons, a neural net is built off multiple layers. With the goal of seperating two different categories of events, such as signal and background, a set of selection variables s_i is used as the input layer. Each consequent layer takes inputs from the previous layer and applies a non-linear transformation. After an arbitrary number if intermediate “hidden” layers, the output layer can be expressed as a single number.

For the case of a binary classifier, either the event is signal or background, the neural net output can be seen as a signal likelihood. By applying a rectangular cut on this classifier, effectively a simultaneous selection on all of the selection variables s_i is made.

The neural nets layer parameters are optimized to minimize the classification error by inputting a training dataset with known categories. Then, the neural

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

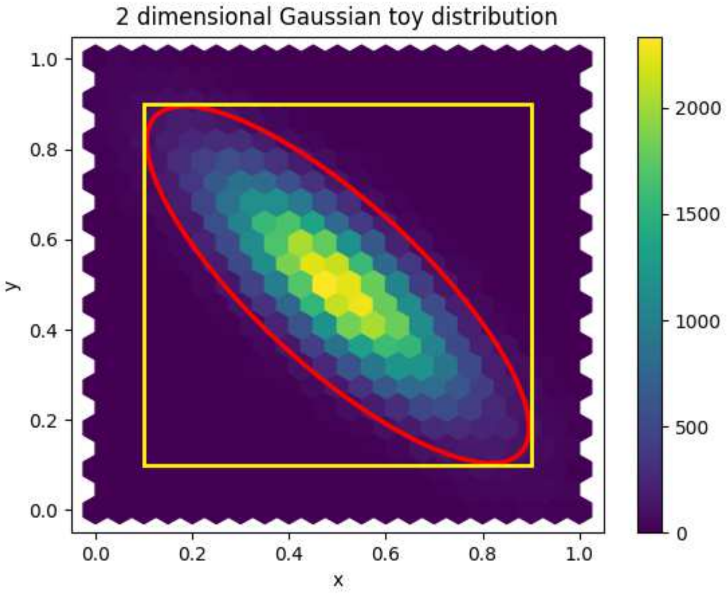


FIGURE 4.5: A generated multivariate Gaussian with correlations between variables is no longer optimally selected by rectangular selections. In the case of a uniform background, the ellipsoid cut pictured in red will lead to a higher figure of merit than the rectangular cut in yellow.

net can be applied to a testing dataset with unknown categories, to yield a classifier probability.

4.2.4 Decision trees

Another method of finding optimized multivariate selections is the binary decision tree [37]. A binary decision tree is built of consecutive split decisions based on the input variables, with the goal of seperating signal and background with maximum FOM.

After sorting all input events of the training data set, the splitting value with best seperation for each value is computed. The variable and splitting value with best seperation is chosen as the first split, yielding two seperate datasets (nodes). The process is then repeated for each node recursively. The algorithm stops when the FOM can not be improved through additional seperations. This decision tree can then be applied to a testing data set.

This algorithm can be improved by “boosting”. Feeding the decision tree output into another decision tree with increased weight of the misidentified events improves the classifier. This method is then called boosted decision tree (BDT) [41].

4.3 Signal extraction

Even after suppressing backgrounds, the resulting set of events will not be composed exclusively of the sought after decay. To compute a branching fraction,

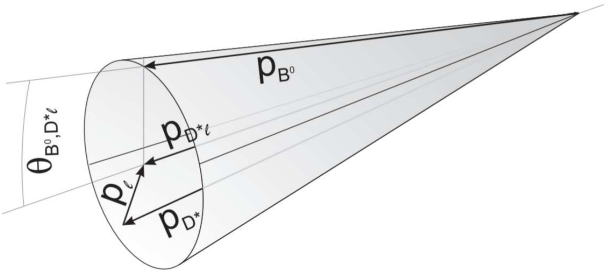


FIGURE 4.6: Reconstructed lepton and $D^{(*)}$ momenta are added up to obtain $p_Y = p_{D^{(*)}\ell}$. The angle between p_Y and the nominal B direction is defined as the angle θ_{BY} [42].

it is necessary to estimate the amount of signal events in the selected dataset. By finding a variable with different distributions for signal and background components, the components can be seperated. For decays with only visible final state particles, such as $B^- \to [D^0 \to K^- \pi^+] \pi^-$ commonly used fitting variables are the beam constrained mass

$$m_{\rm bc} = \sqrt{E^{*2}_{\rm beam} - \vec{p}^{*2}_B} \tag{4.12}$$

and the energy difference

$$\Delta E = E^*_B - E^*_{\rm beam}. \tag{4.13}$$

Belle II collisions occur at $\sqrt{s} = 10.58$ GeV, producing two B mesons with mass $m_B = 5.279$ GeV $\approx \sqrt{s}/2$. Thus, the center-of-mass collision energy $E^*_{\rm beam}$ is known without measuring it and $m_{\rm bc}$ and ΔE can be computed with high resolution. Signal events peak at $m_{\rm bc} = 5.279$GeV and $\Delta E = 0$GeV. Combinatorial background can be seperated due to its flat distribution as shown in Figure 4.7a.

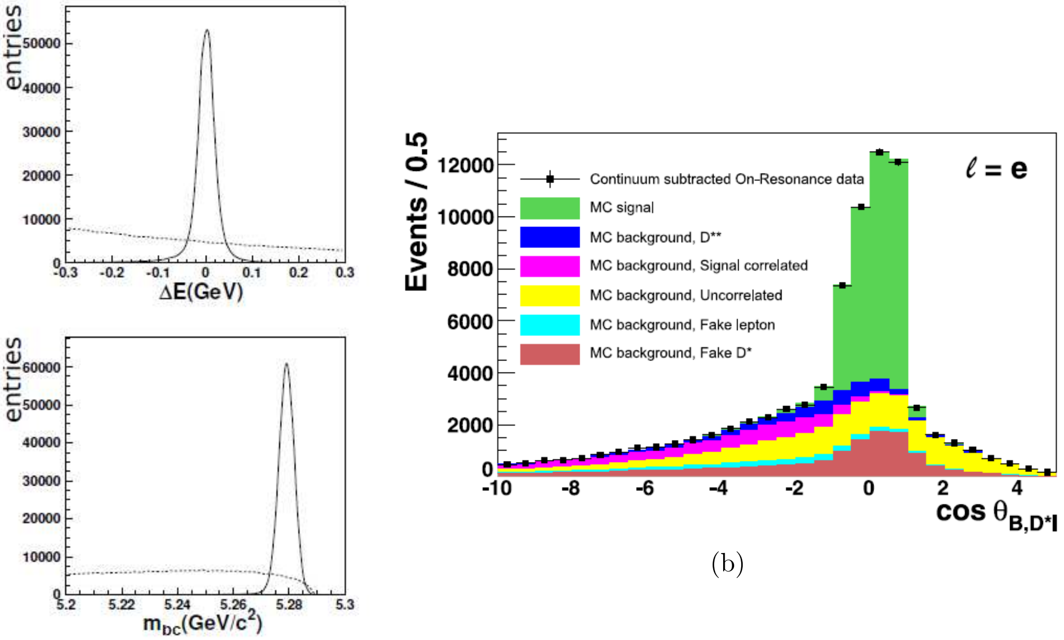
For semi-leptonic decays with neutrinos such as $B^- \to D^0 \ell^- \overline{\nu}_l$ $\vec{p}_B$ and E_B can not be measured precisely due to the difficulty of measuring invisible neutrinos kinematics. Thus, $m_{\rm bc}$ and ΔE are not well suited to extract signal. An alternative variable used in previous $\overline{B}^0 \to D^{*+} \ell^- \overline{\nu}_l$ and $B^- \to D^0 \ell^- \overline{\nu}_l$ analyses [42][43] is

$$\cos\theta_{BY} = \frac{2E^*_{\rm beam}E^*_Y - m_B^2 - m_Y^2}{2|\vec{p}^*_{\rm beam}||\vec{p}^*_Y|} \tag{4.14}$$

where Y is the combined $D^0\ell^-$ system and thus E^*_Y and $\vec{p}^*_Y$ are calculated from the reconstructed $D^0\ell^-$ system [42]. $E^*_{\rm beam}$ and $\vec{p}^*_{\rm beam}$ are known machine parameters. An illustration of the angle θ_{BY} is shown in Figure 4.6.

Because the derivation of equation (4.14) only holds under assumption that the neutrino is the only missing particle, $\cos\theta_{BY}$ is strictly contained in the interval $(-1, 1)$ for signal events only, making it a powerful discriminator between signal and background. A $\cos\theta_{BY}$ distribution from a previous $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ analysis is shown in Figure 4.7b [42].

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a)

(b)

FIGURE 4.7: (a) m_{bc} and ΔE distributions in $B^+ \to K_S^0 \pi^+$ show clear seperation between background (dashed lines) and signal (solid lines) [37]. (b) cos θ_{BY} is confined between -1 and 1 only for signal in semi-leptonic decays and thus gives good seperation power in analyses such as the pictured $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ [42].

4.3.1 Maximum likelihood fit

When the probability density functions (PDF) of fitting variable distributions are known, the analytical functions can be fit to the graphs. For example, the m_{bc} distribution shown in Figure 4.7a shows a signal component distributed as a Gaussian peak and a ARGUS distribution-like background [44]. A maximum likelihood fit is done by defining a likelihood estimator based on the PDFs evaluated at measured data points. The set of PDF parameters that maximizes the likelihood estimator is then the best model for the dataset. By integrating the maximum likelhood fit result Gaussian PDF describing the signal component, the number of signal events in the dataset can be calculated. For a more detailed description of likelihood estimators see *e.g.* [45].
In the case of a cos θ_{BY} analysis, such as the one pictured in Figure 4.7b, the analytic expressions of the underlying shapes are not intuitively known. Instead of fitting PDFs, a maximum likelihood fit can be performed by fitting the binned histogram shapes of the categories to data. The details of the special case of binned maximum likelihood fits commonly used in high energy physics are described by Barlow and Beeston [46]. Chosen MC fractions such as signal, $B\overline{B}$ background or continuum are fitted to the data distribution. While allowing for Poissonian $\sqrt{N_{\text{bin}}}$ variations in both MC fractions and data, the MC fractions are scaled to best describe the data distribution. The resulting fraction scalings are used to calculate the number of signal events in the dataset. Additionally, the binned likelihood fit estimates the statistical error on MC fractions [46].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

4.3.2 Fit tests

In the limit of infinite statistics, maximum likelihood estimators satisfy three conditions making them an *ideal* estimator [37]:

- **Consistency.** An estimator is consistent if, as the number of data points converges to infinity, the parameter that is estimated converges to the true value.

- **Unbiased.** A bias of an estimator is the difference between the average estimation on finite sample and the true value. If the bias is 0 the estimator is unbiased.

- **Efficient.** The spread around an estimator, *i.e.* the standard deviation σ, is a measure of estimator efficiency. An estimator with lower spread is more efficient. Specifically, an estimator with spread equal to the lowest possible theoretical value, given by the Cramér–Rao lower bound [47], is called an efficient estimator.

Because data and MC samples in physics analyses are finite, the maximum likelihood fit does not necessarily satisfy these conditions. To ensure that the fit on the available samples does not show a significant bias and estimates the uncertainty accurately, fit tests can be performed.
For a fitted category in a Monte Carlo sample with a true amount of events n_{true}, a fitted amount of events n_{fitted} and a fitted statistical error $\widetilde{\sigma}$, the pull [48]

$$g = \frac{n_{\text{fitted}} - n_{\text{true}}}{\widetilde{\sigma}} \tag{4.15}$$

is a useful quantity for testing. By definition, the distribution of pulls fitted with an ideal estimator evaluated on seperate samples is Gaussian distributed with mean $\mu = 0$ and standard deviation $\sigma = 1$. A biased estimator would correspond to a pull mean $\mu \neq 0$, while an inefficient estimator would correspond to a pull standard deviation $\sigma > 1$, implying that the estimated statistical error is too low. By fitting multiple samples and calculating the pull for each, the resulting pull distribution allows for testing of the fit behavior. Two common methods to generate multiple samples are used in high energy particle physics:

- **GSIM test.** The full MC sample is split into n equal-sized streams. Evaluating the fit on each stream and calculating g gives n data points for the pull distribution.

- **Toy MC test (Bootstrapping).** Using the bootstrapping technique [49], new MC samples are generated by random Poisson variations on the original MC sample. For a sample containing N bins a random weight $w_i, \; i \in (1, N)$ is assigned to every bin, drawn from a Poisson distribution $w_i \sim \text{Pois}\,(\lambda = 1)$. This way, an arbitrary amount of new samples can be generated to get a pull distribution.

The resulting pull distributions are evaluated to ensure an unbiased fit with a well estimated statistical error.

Chapter 5

$B^- \rightarrow D^0 \ell^- \overline{\nu}_l$ reconstruction

After introducing the physics background, the experimental setup and the data analysis techniques involved, this chapter gives an overview of the $B^- \rightarrow D^0 \ell^- \overline{\nu}_l$ reconstruction analysis performed. This section begins by giving an overview of the data and MC samples used in the analysis. Next, the selection criteria are presented, including kinematic selections and particle identification. The types of backgrounds encountered are discussed and selections to suppress the background are given. Vetoes and multivariate methods are used to suppress the dominant $\overline{B}^0 \rightarrow D^{*+} \ell^- \overline{\nu}_l$ background. An overview of selection steps with their statistical significance improvements is given.
After optimizing the signal selection on simulated Monte-Carlo events, the selection criteria are applied to available Belle II physics data to show agreement between MC and data. Known discrepancies between data and MC are corrected with correction factors. Distributions of kinematic variables are shown for MC and data overlaid with different scalings. The deep neural network classifier is compared on the two samples.

5.1 Data and MC samples

Belle II data taking is grouped into *experiments*, each consisting of runs. This thesis covers data from experiments 7, 8, 10 and 12, containing collisions from spring 2019 to summer 2020. 2019 experiments 7, 8 and 10 are grouped into `proc11`, covering 8.7 fb^{-1} of data available with finalized calibrations. In 2020 62.5 fb^{-1} of data were taken in experiment 12, grouped into 7 buckets. The first three experiment 12 buckets are used in this thesis, giving a total dataset of 34.6 fb^{-1}. A summary is given in Table 5.1.
The thirteenth official Belle II Monte-Carlo campaign MC13 is used as the simulated event sample. 200 fb^{-1} of run-independent MC13 is used, consisting of $\sim$ 200 million charged and neutral $B\overline{B}$ events and $\sim$ 900 million $u\overline{u}$, $d\overline{d}$, $c\overline{c}$ and $s\overline{s}$ continuum events.
The analysis is performed using the BASF2 `release-04-02-08`, available since June 11, 2020.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Experiment	Bucket	Luminosity
7		0.43 fb^{-1}
8		4.60 fb^{-1}
10		3.74 fb^{-1}
12	bucket9	2.7 fb^{-1}
	bucket10	10.4 fb^{-1}
	bucket11	13.1 fb^{-1}
Total		34.6 fb^{-1}

TABLE 5.1: Data samples available for analysis.

5.2 Selection criteria

To reconstruct $B^- \to D^0 \ell^- \overline{\nu}_l$ efficiently, a variety of selection criteria is used. The selection criteria introduced in this section are chosen to optimize the statistical significance. Using simulated $\Upsilon(4S)$, continuum and background events, the selections were chosen to maximize the figure of merit given in equation (4.9). To optimize the figure of merit, a dataset with loose preselections was generated using 100 fb^{-1} of Monte-Carlo simulation. The loose selections, referred to as *minimal selection*, were then tightened to maximize the figure of merit in the data sample. The minimal and final selection criteria, as well as the the figure of merit and purity improvements through final analysis selections are summarized at the end of the section in Table 5.2 and Table 5.3.
Unless otherwise specified, plots in this section are shown with all final selection criteria applied except for the variable that is being plotted.

5.2.1 Hadronic event selection

Only collision events categorized by the High Level Trigger as hadronic decays are considered for analysis. To be considered a B hadron event by the HLT, events are required to have at least 3 *good tracks*, defined as tracks with transverse momentum $p_T > 0.2$ GeV, distance between point-of-closest-approach (POCA) and interaction point $|d_0| < 2$ cm and a POCA z coordinate $|z_0| < 4$ cm, as well as fail the requirements to be considered a Bhabha scattering event[*]. To reduce continuum events, a requirement on the second Fox-Wolfram moment, introduced in Section 4.1.2, $R_2 < 0.4$ is set in the preliminary minimal selection. This selection is further tightened in the final analysis selection (Section 5.2.6).

5.2.2 Tracking requirements

Charged particle tracks from $\Upsilon(4S)$ events originate from the interaction point. To reduce the amount of beam background, all charged tracks used in reconstruction are required to fulfill $|d_0| < 0.5$ cm and $|z_0| < 2.0$ cm.

---

[*] The requirements for an event to be considered as Bhabha scattering by the HLT are beyond the scope of this thesis. In broad terms, two tracks identified as electrons are required to be highly energetic with a large angle inbetween.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

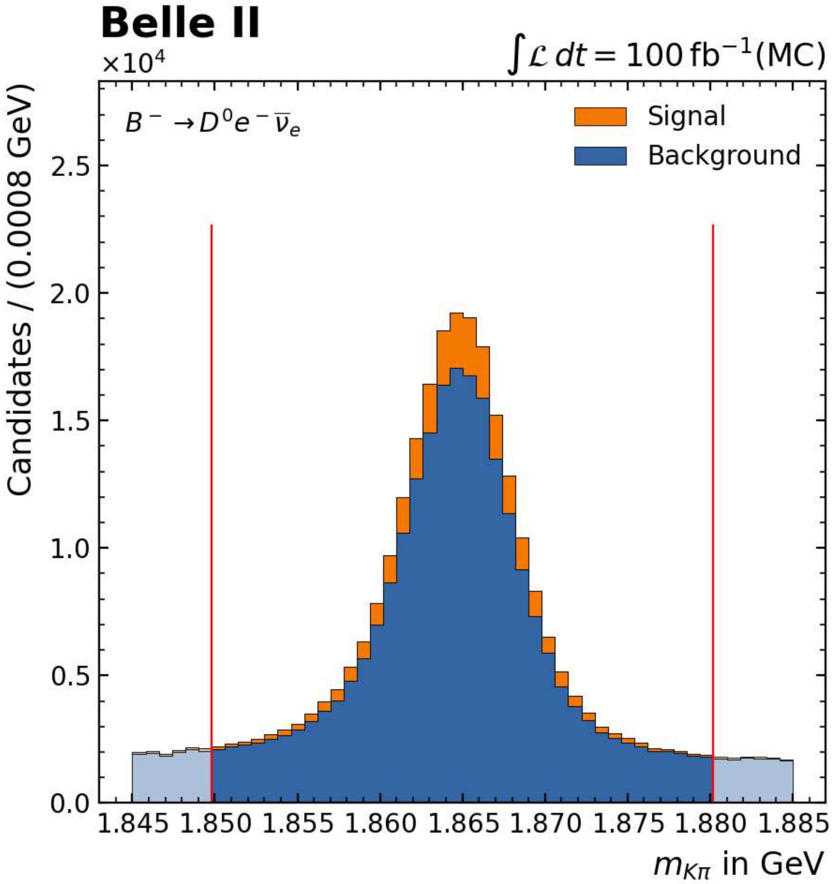
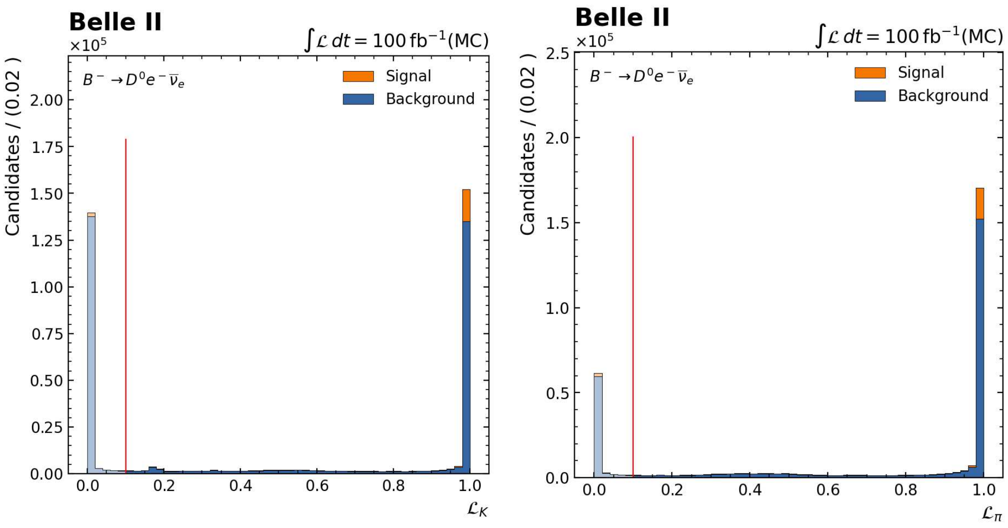


FIGURE 5.1: The D mass window selection for the electron mode. The figure of merit is improved by requiring 1.85 GeV $< m_{K\pi} <$ 1.88 GeV.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a)

(b)

FIGURE 5.2: Loose constraints $\mathcal{L}_K > 0.1$ and $\mathcal{L}_\pi > 0.1$ are placed on the hadronic particle identification likelihoods. Distributions are shown for the electron mode

To ensure that tracks are within acceptance of the CDC the polar angle θ is required to be $17° < \theta < 150°$.

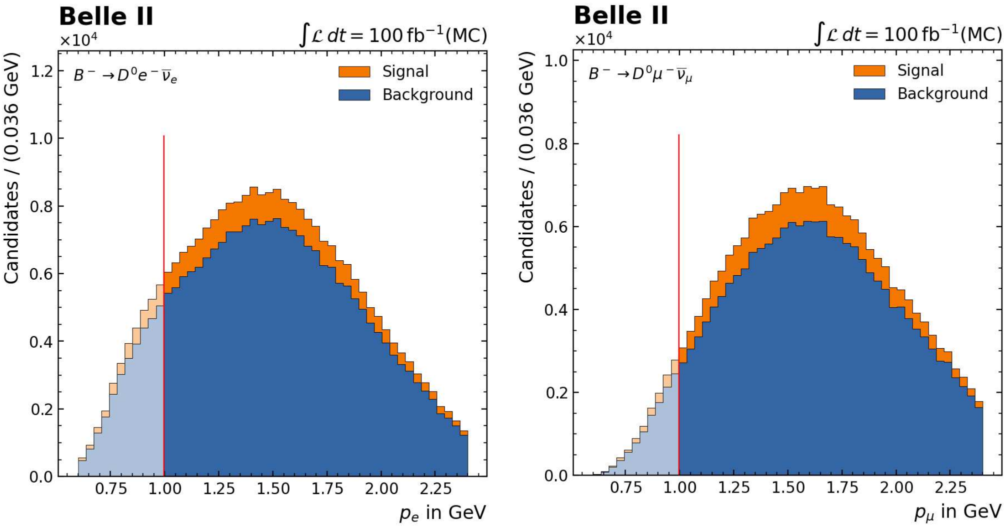
5.2.3 D meson reconstruction

The signal decay $B^- \to D^0\ell^-\overline{\nu}_l$ is reconstructed for the case where the D meson decays to a kaon and a pion. $D^0 \to K^-\pi^+$ has a branching fraction of $\sim 3.95\%$ and can be reconstructed by using two charged tracks. A pion candidate and a kaon candidate, required to be opposite charge, are kinematically combined into a D candidate. Their combined mass $m_{K\pi}$, calculated as seen in equation (4.3), is required to be within 1.85 GeV and 1.88 GeV, corresponding to a 15 MeV mass window around the nominal D^0 mass $m_{D^0} = 1864 \pm 0.17$ MeV. The D meson mass window selection for the electron mode is shown in Figure 5.1. Loose particle identification constraints are placed on the hadron daughters, requiring $\mathcal{L}_\pi > 0.1$ and $\mathcal{L}_K > 0.1$ for pion and kaon candidates respectively.

To reject D mesons originating from continuum background, the center-of-mass momentum p^*_D is required to be smaller than 2.5 GeV.

5.2.4 Lepton selection

For selecting the lepton, the selections differ between electron and muon candidates. To reject low-momentum background, a lower limit is placed on the lepton momentum. From the figure of merit optimization, electron center-of-mass momentum $p^*_e > 0.8$ GeV and muon center-of-mass momentum $p^*_\mu > 0.95$ GeV are optimal selection criteria. The Belle II performance group provides correction factors for lepton identification that are applied at a later stage of the analysis. Due to large systematic uncertainties in these correction factors for



(a) (b)

FIGURE 5.3: Momentum distributions for the (a) electron mode and (b) muon mode.

$p < 1$ GeV leptons the lower momentum limit is placed at $p_\ell > 1$ GeV instead of the statistically favorable p^* selections.
Due to the lepton momentum end-point for $B^- \to D^0\ell^-\overline{\nu}_l$ p_e^* and p_μ^* are required to be smaller than 2.4 GeV.
To seperate electrons from hadrons a stringent selection is placed on the particle identification. `electronID` $\mathcal{L}_e$ and `muonID` $\mathcal{L}_\mu$ are required to be greater than 0.9.

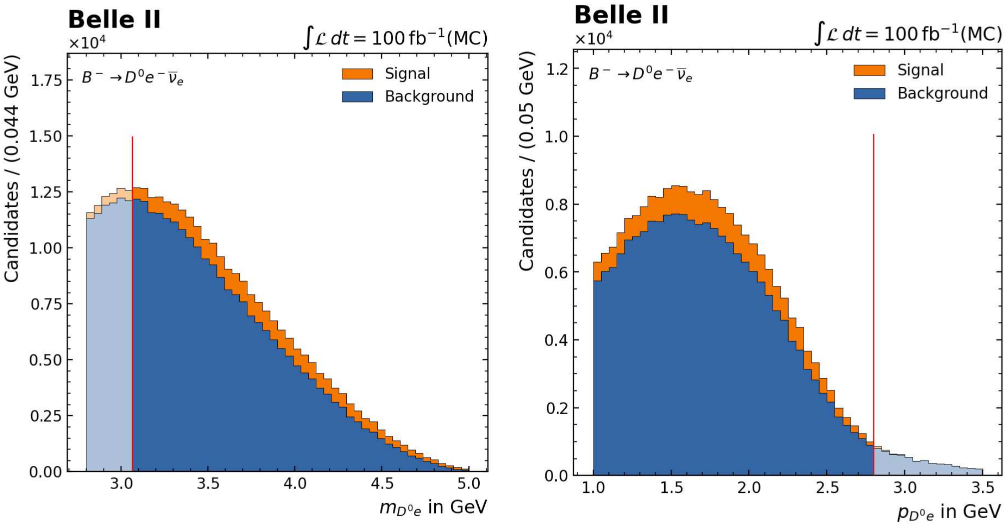
5.2.5 B meson reconstruction

Due to the invisible neutrino the B meson can not be reconstructed fully. Instead, selections are placed on the combined lepton and D meson. Following the convention introcuced in (4.14) the combined $D^0 - \ell^-$ system is referred to as Y.
Due to the difference in electron and muon rest mass, Y selections differ between the two modes. For the electron mode the combined mass m_{De} is required to be larger than 3.15 GeV while the combined laboratory frame momentum p_{De} is required to be smaller than 3 GeV. For the muon mode the respective selections are $m_{D\mu} < 3.35$ GeV and $p_{D\mu} < 2.8$ GeV.

5.2.6 Continuum and beam background suppression

To suppress $e^+e^- \to q\overline{q}$ continuum events, as well as beam background events where no $\Upsilon(4S)$ is present, multiple additional constraints are placed.
All events are required to record at least three charged tracks and contain an overall visible energy E_{vis}, added up from all clusters and tracks, above 4 GeV.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



(a) (b)

FIGURE 5.4: A lower limit $m_{De} > 3.05$ GeV is placed on the mass and an upper limit of $p_{De} < 2.8$ GeV is placed on the momentum of the reconstructed D^0e^- system.

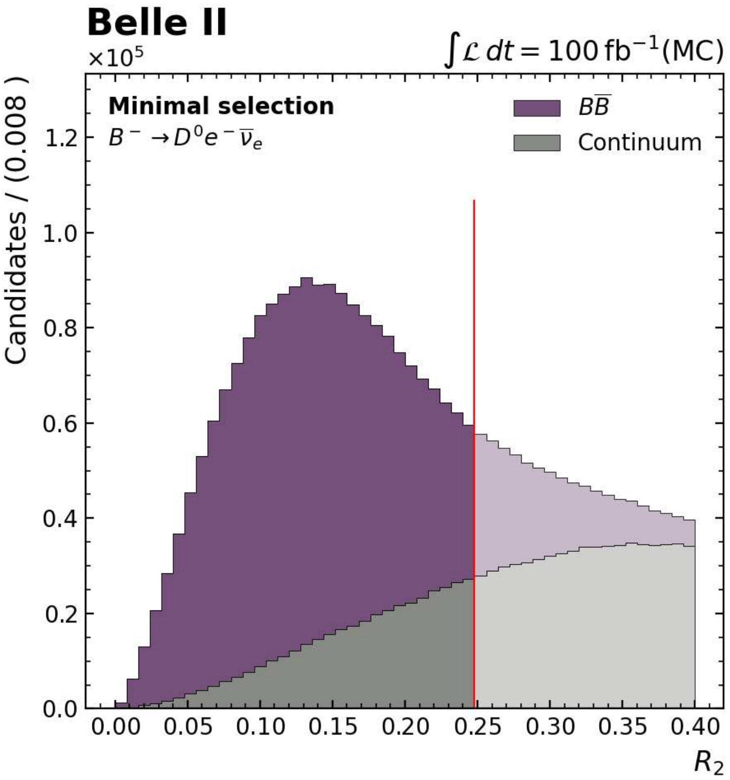


FIGURE 5.5: The Fox-Wolfram moment R_2 distributions differ in shape between continuum and $B\overline{B}$ events. A dataset selected with minimal selection criteria shows that by requiring $R_2 < 0.25$, the majority of continuum background is suppressed.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

To suppress continuum events, the second Fox-Wolfram moment R_2 is required to be smaller than 0.25. This selection is shown in Figure 5.5.

	Minimal selection	Final analysis selection
Tracks	$d_0 < 2$ cm	$d_0 < 2$ cm
	$\|z_0\| < 2$ cm	$\|z_0\| < 2$ cm
	$17° < \theta < 155°$	$17° < \theta < 155°$
Event	Pass hadronic HLT criteria	Pass hadronic HLT criteria
	$E_{\text{vis}} > 4$ GeV	$E_{\text{vis}} > 4$ GeV
	nTracks > 2	nTracks > 2
	$R_2 < 0.4$	$R_2 < 0.25$
D meson	-	$\mathcal{L}_\pi > 0.1$
	-	$\mathcal{L}_K > 0.1$
	1.83 GeV $< m_{K\pi} <$ 1.9 GeV	1.85 GeV $< m_{K\pi} <$ 1.88 GeV
Lepton	0.6 GeV $< p^*_e <$ 2.4 GeV	0.6 GeV $< p^*_e <$ 2.4 GeV
	0.6 GeV $< p^*_\mu <$ 2.4 GeV	0.6 GeV $< p^*_\mu <$ 2.4 GeV
	-	$p_e > 1$ GeV
	-	$p_\mu > 1$ GeV
	$\mathcal{L}_e > 0.7$	$\mathcal{L}_e > 0.9$
	$\mathcal{L}_\mu > 0.7$	$\mathcal{L}_\mu > 0.9$
Combined $D^0 - \ell^-$	$m_{D^0e} > 2.8$ GeV	$m_{D^0e} > 3.05$ GeV
	$m_{D\mu} > 2.8$ GeV	$m_{D\mu} > 3.35$ GeV
	$p_{De} < 3.5$ GeV	$p_{De} < 2.8$ GeV
	$p_{D\mu} < 3.5$ GeV	$p_{D\mu} < 3$ GeV

TABLE 5.2: A comparison of the minimal selection criteria used for pre-selection and the final analysis selections.

5.3 D^* suppression

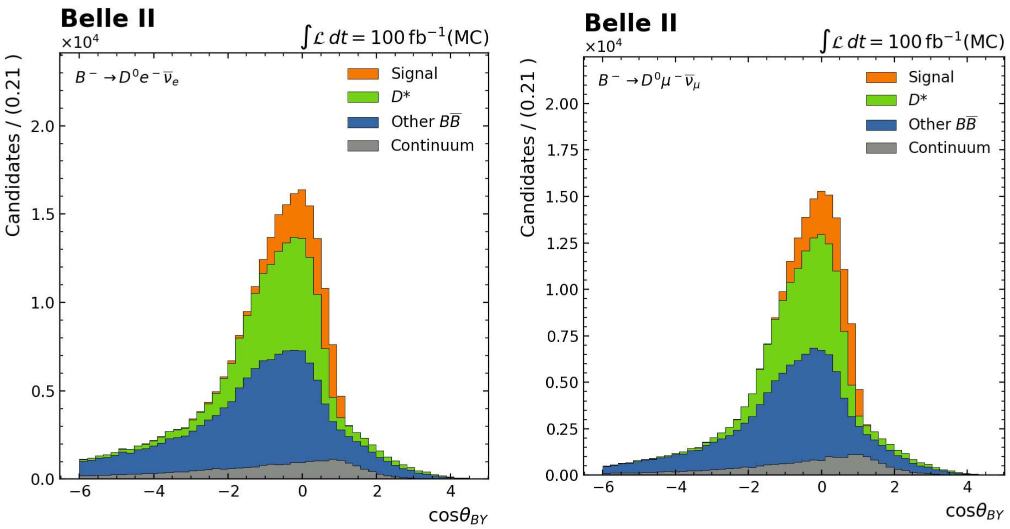
After applying the final analysis selection constraints, the $\cos\theta_{BY}$ is shown in Figure 5.6. Background events have been grouped into three distinct categories:

- D^* **events.** A major background contribution originates from events containing either a charged $D^{*\pm}$ or a neutral $D^{*0}/\overline{D}^{*0}$. The component marked in green in Figure 5.6 is made up of all events with B decays $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ or $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$. Because excited D^* mesons always decay to a D, emitting either a photon or a low-momentum pion in the process, these decays are difficult to seperate from $B^- \to D^0\ell^-\overline{\nu}_l$, and thus contribute significantly to the background.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

	n_{sig}	n_{bkg}	ρ	FOM $\frac{n_{\text{sig}}}{\sqrt{n_{\text{sig}}+n_{\text{bkg}}}}$
$B^- \to D^0 e^- \overline{\nu}_e$				
Minimal selection	51008	2829423	1.8%	30.05
Final analysis selection	27186	235546	10.3%	53.04
$B^- \to D^0 \mu^- \overline{\nu}_\mu$				
Minimal selection	52692	4491858	1.2%	24.72
Final analysis selection	23521	183670	11.4%	51.67

TABLE 5.3: Figure of merit and purity ρ (equation (4.7)) improvements through rectangular selections.



(a) (b)

FIGURE 5.6: $\cos\theta_{BY}$ distributions for electron and muon mode after applying the final selection constraints seen in Table 5.2. The signal component is constrained, up to detector resolution, within -1 and 1. A major background component are $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

- **Other $B\overline{B}$ events.** All $\Upsilon(4S) \to B\overline{B}$ events except for signal decays and D^* events are combined into a $B\overline{B}$ background component. This category consists of a multitude of sub-categories, including combinatorial background, where tracks pass the selection criteria by chance, B to D^{**} events and others.

- **Continuum background.** As mentioned in Section 4.1.2, $e^+e^- \to q\overline{q}$ events provide a large background source. Through constraints on the Fox-Wolfram moment R_2 and the number of charged tracks in the event, the continuum component has been reduced siginificantly.

Around $\sim 30\%$ of events in the dataset can be attributed to the specific D^* decays. Hence, an important part of $B^- \to D^0\ell^-\overline{\nu}_l$ analysis is to suppress D^* events. In this thesis, two techniques are used to suppress D^* background.

5.3.1 D^* veto

In $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decays, the excited D^{*+} further decays like $D^{*+} \to D^0\pi^+$ approximately $\frac{2}{3}$ of the time [12]. The difference between D^{*+} and D^0 rest masses $\Delta m = m_{D^{*+}} - m_{D^{*+}} = 145.4$ MeV is close to the π^+ rest mass $m_{\pi^\pm} = 139.6$ MeV, causing the produced pion to be almost at rest. In literature this pion is referred to as *slow pion*. Due to the additional hadrons, $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ can be reconstructed cleaner than $B^- \to D^0\ell^-\overline{\nu}_l$. This property can be used to implement a veto, described in Section 4.1.3, discarding events where a D^* can be reconstructed.
A summary of the veto conditions is listed in Table 5.4. Specifically three vetoes are used in this analysis:

- **$D^{*+} \to D^0\pi^+$ veto.** For the case described above, the D^{*+} decays to a D^0, emitting a slow pion π_s^+. The reconstructed D meson with mass $m_{K\pi}$ is combined with any pion with momentum $p < 0.35$ GeV. If the difference between combined D^* mass and D^0 mass $\Delta m = m_{K\pi\pi_s} - m_{K\pi}$ fulfills 144 MeV $< \Delta m <$ 148 MeV, a D^{*+} meson is built and the event is discarded. Any $B^- \to D^0\ell^-\overline{\nu}_l$ candidates from this event are not included in the dataset, due to the high likelihood of being a $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$ decay. For the slow pion, no $|d_0|$, $|z_0|$ and θ constraints are required due to their low momenta hindering them from reaching the CDC.

- **$D^{*0} \to D^0\gamma$ veto.** In addition to the charged D^{*+} veto, two neutral D^{*0} vetoes are implemented. For the case of $D^{*0} \to D^0\gamma$, a D^{*0} can be built of the reconstructed D^0 and a photon. Because the photon is restricted to low momentum, selection criteria can be used to veto the decay. Recommendations for soft photon selection critera are given in [50].
  The photon cluster is required to be within angular acceptance of the ECL. Its measured momentum is restricted within 32 and 105 MeV. Additionally, two constraints are placed on ECL shower shape variables. The ratio of energies deposited in the inner ECL crystal and 3×3 matrix of crystals around the central crystal E_1E_9 is required to be larger than 0.35. The lateral energy distribution of the cluster shower S is required to be

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

$D^{*+} \to D^0\pi^+$ veto	
Pion	$p < 0.35$ GeV
D^*	144 MeV $< \Delta m <$ 148 MeV

$D^{*0} \to D^0\gamma$ veto	
Photon	$17° < \theta < 155°$ 32 MeV $< p <$ 105 MeV $E_1E_9 > 0.35$ $S > 0.007$
D^*	141 MeV $< \Delta m <$ 145 MeV

$D^{*0} \to D^0[\pi^0 \to \gamma\gamma]$ veto	
Photons	Same selections as in $D^{*0} \to D^0\gamma$
Pion	$\alpha < 2$ rad $\lvert\Delta\phi\rvert > 1.8$ rad 115 MeV $< m_{\gamma\gamma} <$ 145 MeV
D^*	141 MeV $< \Delta m <$ 145 MeV

TABLE 5.4: Contraints required for D^* vetoes. Photon and π^0 selections based on recommendations in internal Belle II note [50].

larger than 0.007.
To form a D^{*0} and to discard the event, Δm has to be within 141 and 145 MeV

- $D^{*0} \to D^0[\pi^0 \to \gamma\gamma]$ **veto.** Lastly, a veto is implemented for the case of a D^{*0} decaying to a D^0 and a π^0. π^0 decay almost instantly into a pair of photons at a branching fraction > 98.8%. Thus, the π^0 has to be reconstructed via two ECL clusters. The selection criteria are again based on recommendations in [50].
  γ ECL clusters have to fulfill the same conditions as in the $D^{*0} \to D^0\gamma$ veto. To form a π^0, the 3D angle inbetween two photon clusters α is required to be smaller than 2 rad. The difference in azimuthal angles $|\Delta\phi| = |\phi_{\gamma_1} - \phi_{\gamma_1}|$ is required to be larger than 1.8 rad. The combined mass of two photons $\mathrm{m}_{\pi\pi}$ has to be within 115 and 145 MeV.
  Again, the mass difference between reconstructed D^{*0} and reconstructed D^0 has to be within 141 and 145 MeV for the event to be discarded.

Due to the prevalence of photon clusters, γ and neutral pions are difficult to reconstruct correctly compared to the charged D^* veto. Hence, the charged veto is more effective in the analysis. Numbers on veto effectiveness are shown in Table 5.6 at the end of the section.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

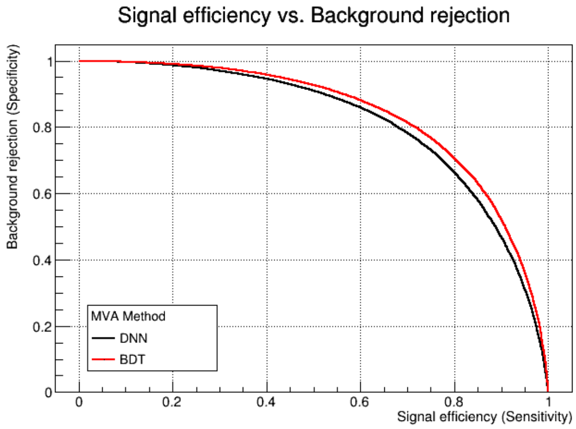


FIGURE 5.7: Signal efficiency against background rejection of the two trained models evaluated on the testing dataset.

5.3.2 Multivariate selections

While the vetoes suppress 75.5% and 73.6% of D^{*+} events in electron and muon mode respectively, the neutral D^{*0} events are only reduced by 27.7% and 20.1%. To further suppress D^* and especially D^{*0} events, multivariate selections as outlined in Section 4.2 are applied.
A deep neural network and a boosted decision tree are trained using the `ROOT` Toolkit for Multivariate Analysis (TMVA) [51]. The models are trained on 50% of the total MC sample to distinguish signal events originating from $B^- \to D^0\ell^-\overline{\nu}_l$ from background events specfically originating from $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$ events.

Input variables

A multitude of selection variables is used to help the models seperate the decays:

- **Kinematic variables.** Kinematic variables from reconstructed particles D, ℓ and Y are used. (Center-of-mass) momentum $p^{(*)}$ and the polar angle θ of the D meson, $p^{(*)}$ of the lepton and p, mass m and transverse momentum component p_T of the combined $D^0\ell^-$ are a part of the selection variables.

- **Event variables.** In addition to the kinematic variables, the number of photon-like clusters n_γ and the total energy of these clusters $E_{\gamma,\mathrm{ECL}}$ are given to the networks.

- **Individual cluster variables.** In $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$ decays, the generated photons and pions are emitted with little angular seperation to the D^0. Thus, shape and energy of clusters close to the D^0 direction can give additional seperation power. For every $B^- \to D^0\ell^-\overline{\nu}_l$ candidate, the closest clusters are selected by calculating the angle $\cos\alpha = \frac{\vec{p}_D\vec{p}_i}{|\vec{p}_D|\ |\vec{p}_i|}$ between the D meson and cluster i. For the three clusters with maximum $\cos\alpha$, the kinematic variables E, p_T, θ, ϕ and $\cos\alpha$, as well as the shape variables E_1E_9, E_9E_{21}, lateral energy distribution S and a BASF2

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Classifier selection	DNN		BDT	
	Training	Testing	Training	Testing
> 0.01	0.159	0.164	0.223	0.203
> 0.1	0.517	0.516	0.567	0.553
> 0.3	0.777	0.777	0.805	0.800

TABLE 5.5: A comparison of signal efficiencies in testing and training datasets. A classifier with significantly higher efficiency in the training dataset than in the testing dataset points to overtraining of the model.

computed BDT classifier ZernikeMVA are added to the set of selection variables.

Performance

Both models are trained on one half of the MC sample and evaluated on the other half. An overtraining test via TMVA is performed for different selections on the classifier and shown in Table 5.5. No overtraining is visible for the deep neural network, while overtraining in the order of a few percent, decreasing with higher cutoffs, is seen for the boosted decision tree.

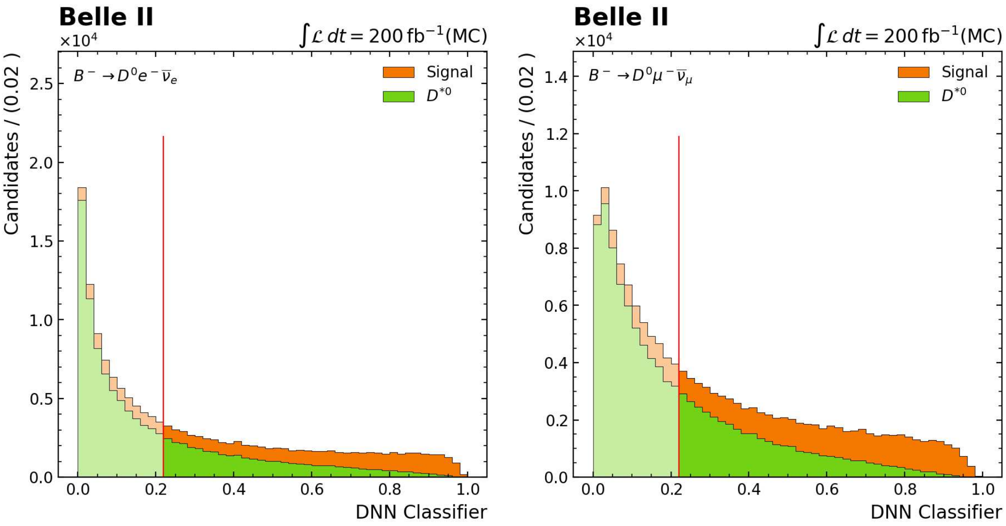
A common measure for classifier performance is the receiver operating characteristic (ROC) shown in Figure 5.7. It shows signal efficiency against background rejection for different cutoffs on the trained classifier. While the BDT performs slightly better on the testing dataset, the FOM improvement between DNN and BDT is similar due to the observed overtraining. In this analysis, the DNN classifier is used for a multivariate selection.

DNN classifier distributions for signal and D^{*0} samples obtained by applying the trained model to a 200 fb^{-1} MC sample are shown in Figure 5.8. The figure of merit is optimized for a rectangular selection at DNN > 0.22, rejecting 66.8% and 63.3% of D^{*0} events for electron and muon channels respectively, while retaining 80.0% and 82.5% of signal events.

While especially effective for suppressing neutral D^* events, the multivariate selection is effective against all types of backgrounds as shown in Figure 5.9. In total, the electron and muon datasets improve in purity from 13.3% and 14.1% to 26.7% and 26.2%.

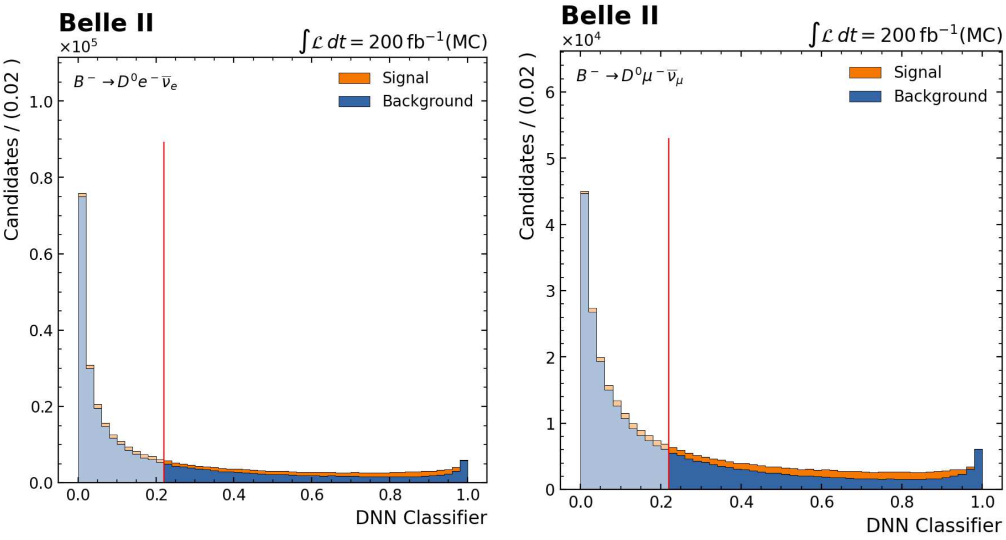
Full event compositions before and after each analysis step are shown in Table 5.6. The $\cos\theta_{BY}$ distributions after multivariate selections is shown in Figure 5.10.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) (b)

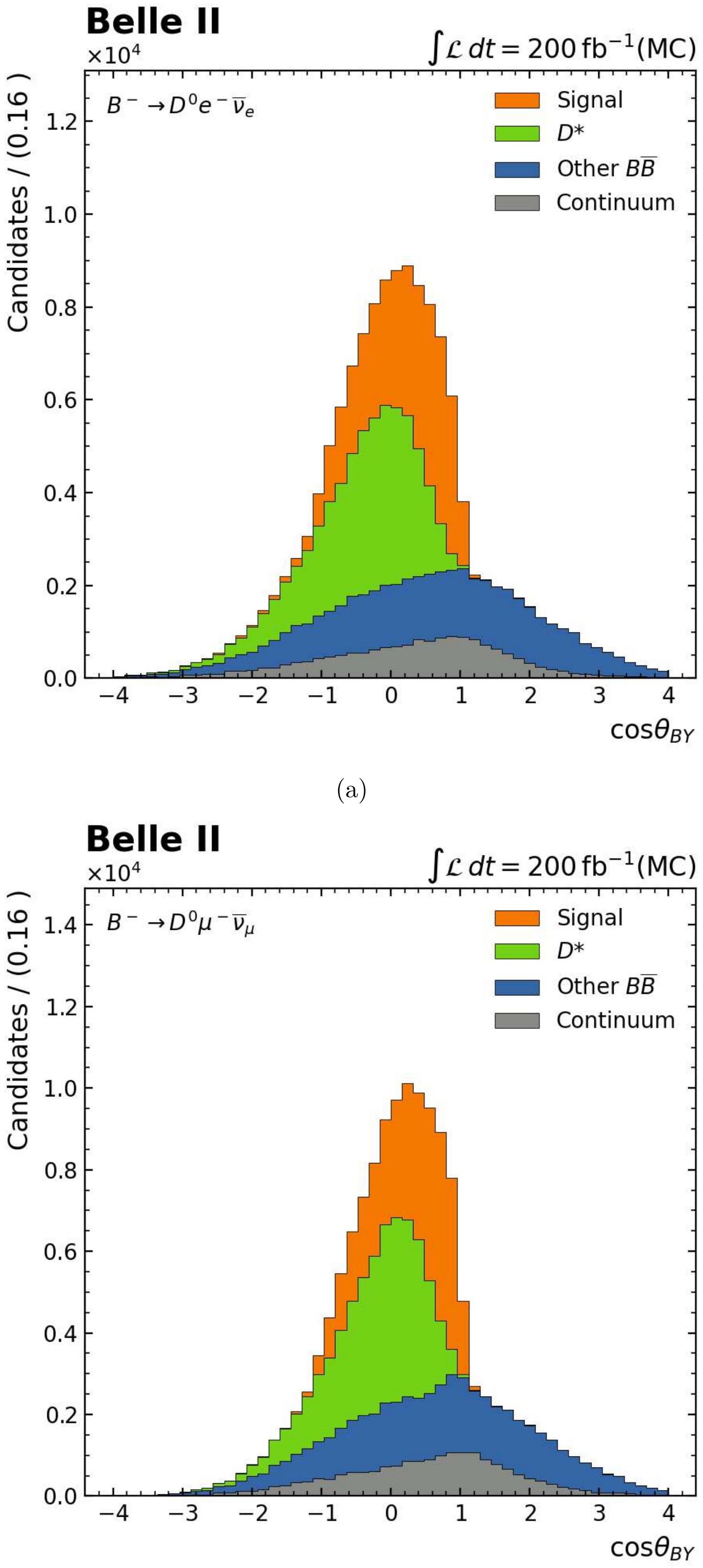
FIGURE 5.8: DNN classifier distributions for signal and D^{*0} events in a 200 fb^{-1} Monte-Carlo sample. A selection at DNN > 0.22 rejects the majority of neutral D^{*0} events.



(a) (b)

FIGURE 5.9: The DNN classifier is effective at separating all types of backgrounds from signal. The chosen selection DNN > 0.22 improves the figure of merit from 77.6 to 98.3 in the electron mode and from 78.84 to 97.73 in the muon mode.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a)

(b)

FIGURE 5.10: cos θ_{BY} distributions after multivariate selections.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

	Number of events					Significance measures		
	Signal	D^{*+}	D^{*0}	Cont.	Other BG	ρ	ϵ	FOM
$B^- \to D^0 e^- \overline{\nu}_e$								
MC sample	190563	-	-	-	-	-	100%	-
Minimal sel.*†	102016	-	-	-	5658846	1.8%	53.5%	42.5
Final sel.*	48728	97080	147614	42326	137768	10.3%	25.6%	70.81
Post-Veto	45361	23826	106662	31509	134272	13.3%	23.8%	77.61
Multivar. sel.	36268	6619	35390	17127	40674	26.7%	19.0%	98.32

$B^- \to D^0 \mu^- \overline{\nu}_\mu$								
MC sample	190629	-	-	-	-	-	100%	-
Min. sel.*†	105384	-	-	-	8983716	1.8%	55.3%	34.96
Fin. sel.*	45334	79664	125720	38368	109842	11.4%	23.8%	71.78
Veto	44122	21014	100424	31640	116003	14.1%	23.1%	78.84
Multi-var.	36385	6424	36902	18394	40514	26.2%	19.1%	97.73

TABLE 5.6: Full analysis flow for electron and muon modes. For every step the purity ρ, the efficiency ϵ and the figure of merit are given as defined in equations (4.7), (4.8) and (4.9).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

---

[*] The dataset for evaluating vetoes and the multivariate selection is 200 fb^{-1}, twice as large as the dataset used to for rectangular selection optimization. To account for that, the number of events in selection and final analysis selections have been multiplied by a factor of 2.

[†] Due to file size constraints the background types have not been further classified in the minimal selection.

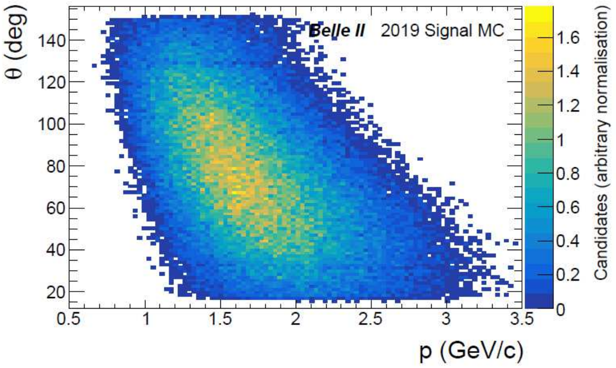


FIGURE 5.11: The two-dimensional J/ψ e^+e^- lepton distribution, used to evaluate leptonID efficiency corrections, has low event sample size at smaller momenta [52].

5.4 Data - Monte-Carlo agreement

After optimizing selection criteria, the reconstruction is applied to data of the 2019 and 2020 Belle II physics runs specified in Section 5.1. Before comparing the two data samples, known discrepancies are corrected.

5.4.1 Lepton identification corrections

A known discrepancy between data and MC, evaluated by the Belle II performance group arises in lepton particle identification. Tables are provided giving correction factors for electrons and muons, binned in their momentum p and polar angle θ. Corrections are provided for lepton ID selections $\mathcal{L}_\ell > 0.5$, $\mathcal{L}_\ell > 0.9$ and $\mathcal{L}_\ell > 0.95$. The correction weights are based on the ratio of data to MC events observed in channels that can be reconstructed cleanly. Applying weights to the MC sample counterbalances the discrepancy.
Each correction factor has an associated systematic and statistical uncertainty. These uncertainties enter the branching fraction calculation as a systematic error. The order of magnitude of this systematic is evaluated in Chapter 7.
Two types of leptonID corrections are available:

- **Efficiency corrections.** Electron and muon efficiency corrections are applied on lepton candidates that were correctly identified in reconstruction. The discrepancies determining the correction factors were measured mainly by reconstructing $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$ [52]. Because the $J/\psi \to \ell\ell$ spectrum is sparsely populated in low momentum regions, particularly $p_\ell < 1$ GeV, these regions have large uncertanties associated and motivate the $p_\ell > 1$ GeV rectangular selection introduced in Section 5.2.4.

- **Mis-identification corrections.** Mis-identification corrections are applied on lepton candidates that were falsely identified as electrons or muons. The mis-identification probability is refered to as *fake rates*. Fake rates are available for pions and kaons that are falsely identified.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

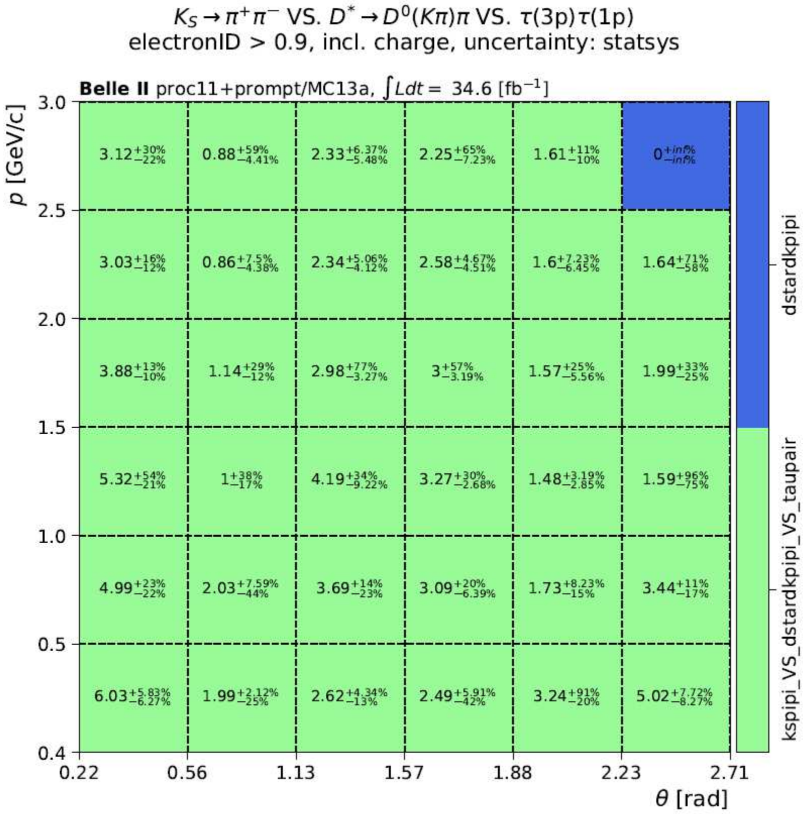
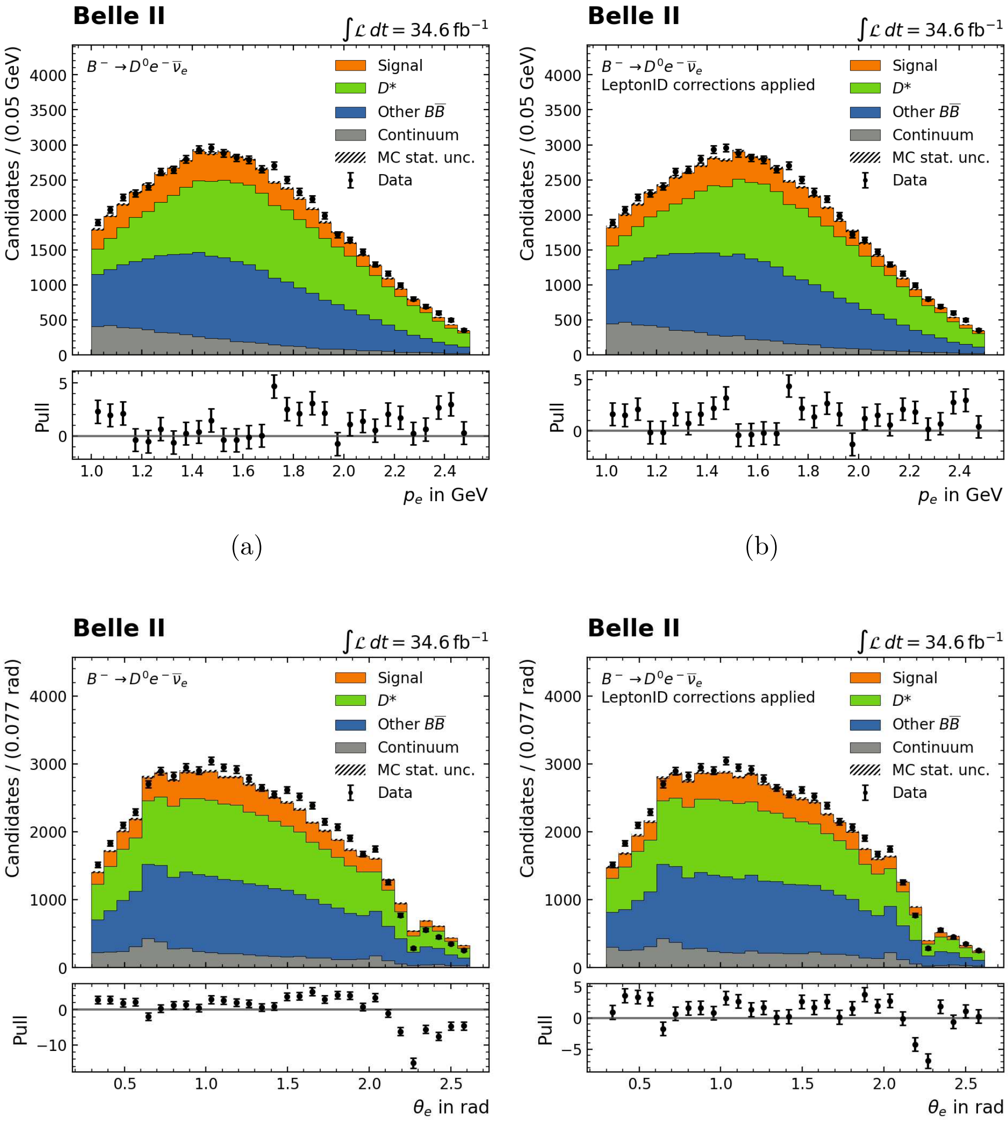


FIGURE 5.12: An example for a lepton identification correction table. A pion that is wrongly identified as an electron is assigned the correction factor associated to its p and θ bin.

Pion and kaon fake rates are mainly measured from $K_S^0 \to \pi^+\pi^-$ and $D^{*+} \to [D^0 \to K^-\pi^+]\pi^+$ reconstruction respectively [52][53]. An example for correction factors applied to pions incorrectly identified as electrons with a selection $\mathcal{L}_e > 0.9$ is shown in Figure 5.12.

Data-MC agreement for lepton momentum and polar angle spectra before and after applying lepton correction factors is shown in Figure 5.13 and Figure 5.14. Pull distributions in the lower half of the plot, defined in equation (4.15), show data-MC difference in units of standard deviations. In the shown dataset the final selection criteria and vetoes have been applied but no multivariate cut is performed. The correction factors improve the agreement in both lepton channels. In the electron channel, the improvement is especially visible in $\theta >$ 1.5 rad region. In the muon channel data-MC agreement is improved as MC is scaled down in all θ regions.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

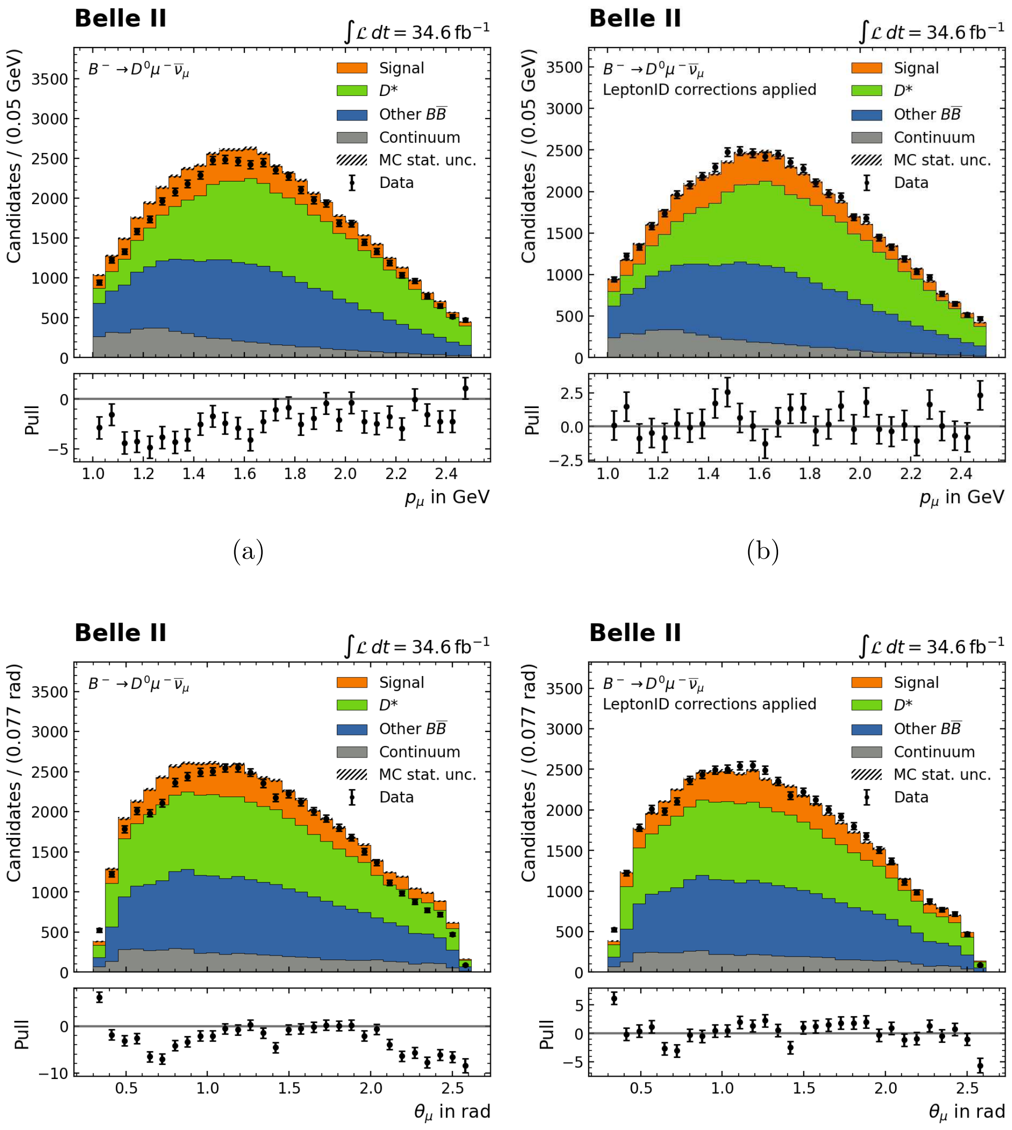


(a) (b)

(c) (d)

FIGURE 5.13: Electron mode p and θ distributions before and after applying lepton correction factors. An improvement is especially visible in the high θ backward region.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
Bibliothek, Your knowledge hub
TU WIEN

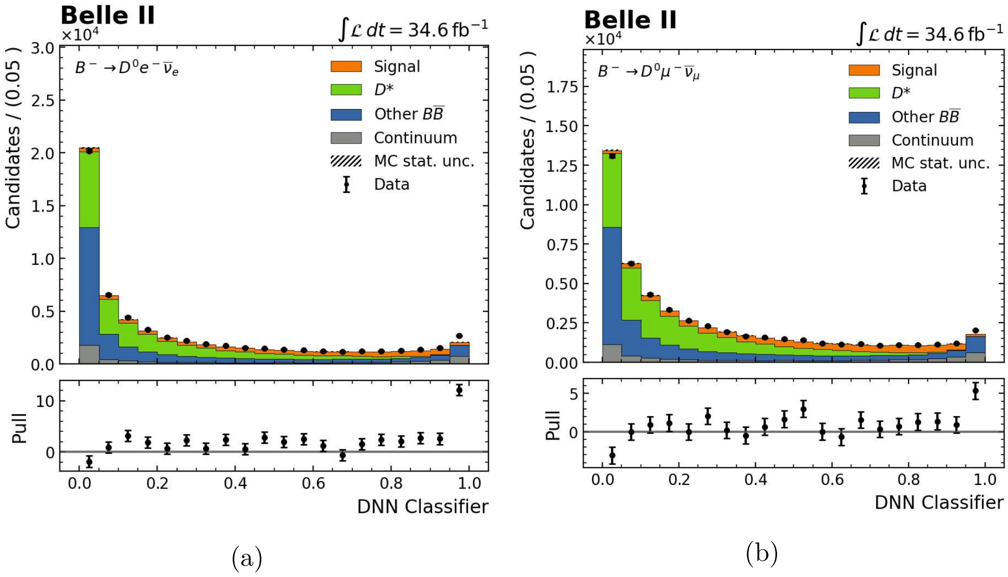


(a) (b)

(c) (d)

FIGURE 5.14: Muon mode p and θ distributions before and after applying lepton correction factors. The overall normalization is improved after applying corrections.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) (b)

FIGURE 5.15: Deep neural network classifier distributions for electrons and muons, with MC and data overlaid.

5.4.2 Pre-fit agreement

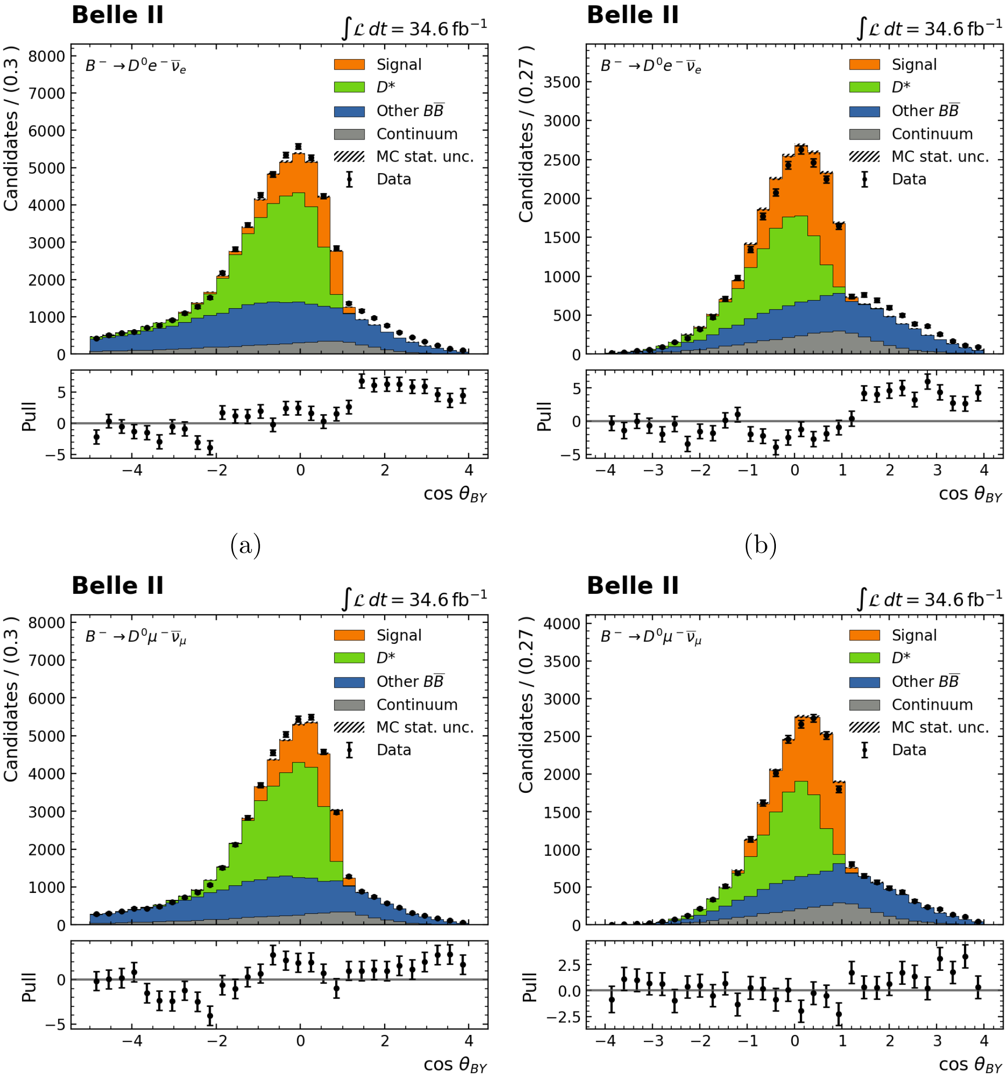
By applying the classifier trained on MC events following Section 5.3.2 to data, each event is assigned a signal likelihood. The distributions of deep neural network classifiers compared on data and MC are plotted in Figure 5.15. In the plots, electron and muon identification correction factors have been applied, and MC events are additionally weighted by the luminosity ratios

$$\frac{\mathcal{L}_{data}}{\mathcal{L}_{MC}} = \frac{34.6\ \mathrm{fb}^{-1}}{200\ \mathrm{fb}^{-1}} = 0.173. \tag{5.1}$$

While data and MC show good shape agreement in the middle region, some disagreement is observed in the first and last bins of both modes. Muons show good normalization agreement, while the electron mode shows a slight overall excess in data compared to MC. The cos θ_{BY} distributions for data and MC scaled by luminosities before and after applying a multivariate selection are shown in Figure 5.16. While the muon mode shows good shape agreement between data and MC, there is noticable disagreement in the cos $\theta_{BY} > 1$ sideband of the electron mode. This disagreement is visible both before and after the cut on the multivariate classifier.

Comparing the normalization between data and Monte-Carlo samples, a 2.98% excess of events in data compared to MC is observed in the electron mode, while a 0.99% excess is observed in the muon mode. A possible explanation to the larger excess in the electron mode could be differently modelled radiative Bremsstrahlung photons emitted from the lepton in the MC compared to real data.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) (b)

(c) (d)

FIGURE 5.16: Weighted $\cos\theta_{BY}$ data-MC agreement for electrons and muons before and after applying a multivariate selection. MC events are scaled by the luminosity ratio (5.1). In the plots (b) and (d) a rectangular selection on the multivariate classifierDNN > 0.22 is applied.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Chapter 6

Fit

This chapter covers the extraction of signal from the optimized dataset. With the ROOT software framework, MC fractions are fitted to data to obtain a number of signal events and an associated uncertainty. After explaining the fit setup including the categories used, fit tests are shown to evaluate the fit. Finally, results of the fit to data are shown and signal yields for both modes are given.

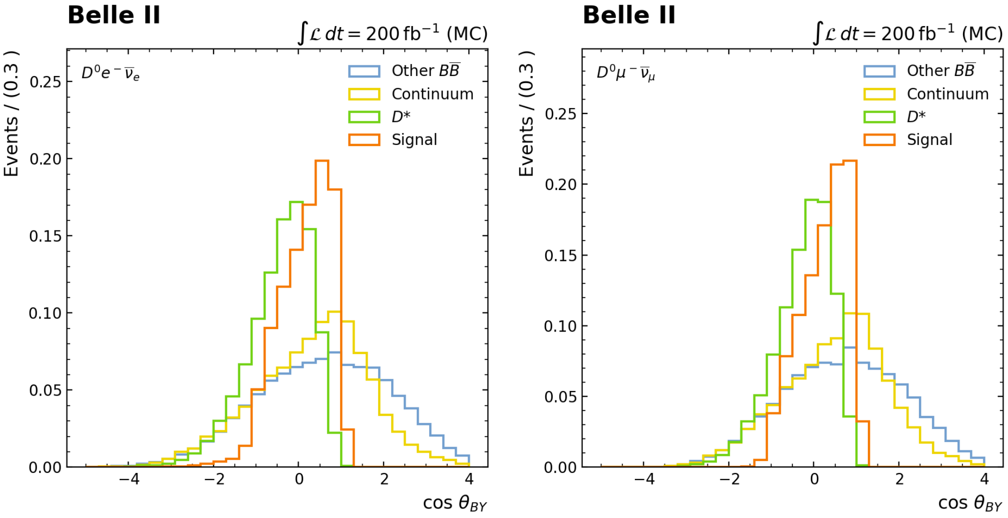
6.1 Fit setup

To calculate the branching fraction from the reconstructed dataset, the number of signal events in data has to be extracted. This is done by performing a binned maximum likelihood fit introduced in Section 4.3.1. Components of the Monte-Carlo dataset are used as template distributions to fit the data. Specifically, four components are used: signal, D^*, other $B\overline{B}$ background and continuum background. The components are required to differ in shape in the fitting variable, to be distinguishable in the fitting procedure. Shapes for the chosen fitting variable $\cos\theta_{BY}$ are shown with normalized scaling in Figure 6.1. Signal and D^* components are observed as peaks at different values of $\cos\theta_{BY}$, while continuum and other $B\overline{B}$ form differently shaped background. All four components are allowed to float, meaning that there are no constraints placed on minimum or maximum size of the component.
The range $-4 < \cos\theta_{BY} < 4$ is chosen as the fit range, divided into 30 bins.
The high energy physics software framework ROOT offers the binned maximum likelihood fit method TFractionFitter, chosen for extracting the signal in this analysis. The fitter maximizes the likelihood to return fitted component fractions and evaluates the statistical uncertainty for each component [46].

6.2 Fit tests

To ensure that the fit is unbiased and estimates the uncertainty accurately, a variety of tests can be used. In this analysis, the GSIM test and the toy-MC test introcuced in Section 4.3.2 are used to evaluate the fit before fitting to the data.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) (b)

FIGURE 6.1: The category shapes of cos θ_{BY} distributions are distinguishable for both electron and muon modes.

6.2.1 GSIM test

To perform the GSIM test, the 200 fb^{-1} Monte-Carlo dataset is split into 11 equal sized streams. Each stream is used as pseudo-data once, and fitted against the other 10 streams. The fitted fractions and statistical uncertainties are used to calculate the pull (4.15). The pull distributions are then made up of 11 data points for every fitting component.
Fit results are shown in Table 6.1. No significant bias is observed, and the statistical uncertainty is well estimated for the signal component.

6.2.2 Toy MC test

The second test performed is the toy MC test. Using the bootstrapping method, new resampled MC distributions are generated by randomly assigning Poisson distributed weights $w_i = X \sim \mathrm{Pois}(\lambda = 1)$ to each bin and then downscaling the resulting sample to the data size 34.6 fb^{-1}. The new toy distribution generated this way is then fitted to the real MC distribution and the pull is calculated.
In this analysis, 1000 toy distributions are generated to calculate pull distributions. The pull distributions of electron and muon mode for the signal component are shown in Figure 6.2. Means and standard deviations for all components are listed in Table 6.1. Confirming observations of the pull test, no biases are observed and the pull standard deviations are tendentially smaller than the fitted standard deviations. Observing a smaller standard deviation in testing than given by the fit implies that the statistical error is overestimated and can be used as is for a conservative error estimate.

It can be concluded that the fitting procedure shows no significant bias and the estimated statistical error can be used as measurement uncertainty.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

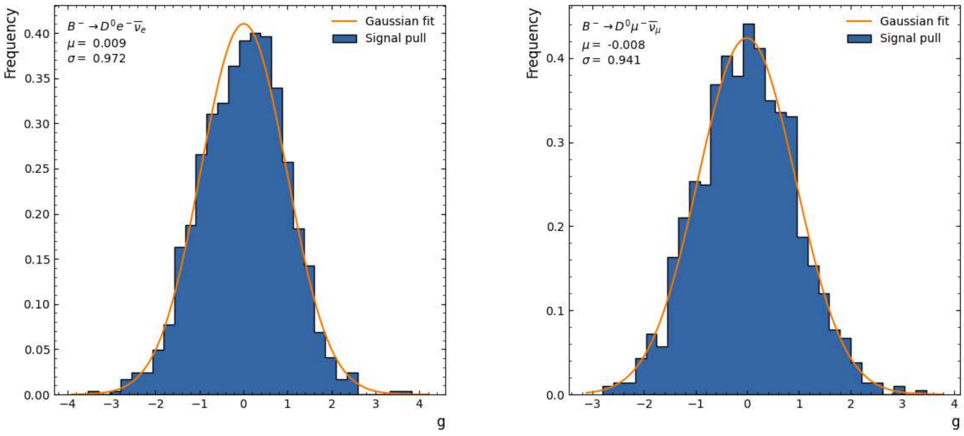
	GSIM test		Toy MC test	
	μ	σ	μ	σ
$B^- \to D^0 e^- \overline{\nu}_e$				
Signal	0.017	0.982	0.009	0.972
D^*	-0.002	1.355	0.001	0.94
Continuum	0.026	1.033	-0.028	1.011
Other BB	-0.031	0.939	0.006	1.038
$B^- \to D^0 \mu^- \overline{\nu}_\mu$				
Signal	0.018	0.896	-0.008	0.941
D^*	-0.016	1.134	-0.007	0.943
Continuum	-0.01	1.072	0.022	1.002
Other BB	-0.015	1.02	-0.020	1.002

TABLE 6.1: Fit validation results for the GSIM and Toy MC test applied on electron and muon mode.

6.3 Fit to data

After validation of the fitting procedure on simulated pseudo-data it is applied on the 34.6 fb^{-1} Belle II data sample. With the fit parameters outlined in Section 6.1 the four components are fitted in $\cos\theta_{BY}$. Fraction and yield results are shown in Table 6.2. In general, the fit is in good agreement with MC expectations, with a relative statistical uncertainty of approximately 4% in the signal component. The fitted continuum component is smaller than the MC expectation, with a large associated error, especially in the electron mode. This can partly be explained by the sideband disagreement, causing the $B\overline{B}$ component to take away from the continuum component to compensate for the mismatch. The post-fit $\cos\theta_{BY}$ distributions with the components weighted by their fitted fractions are shown in Figure 6.3.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



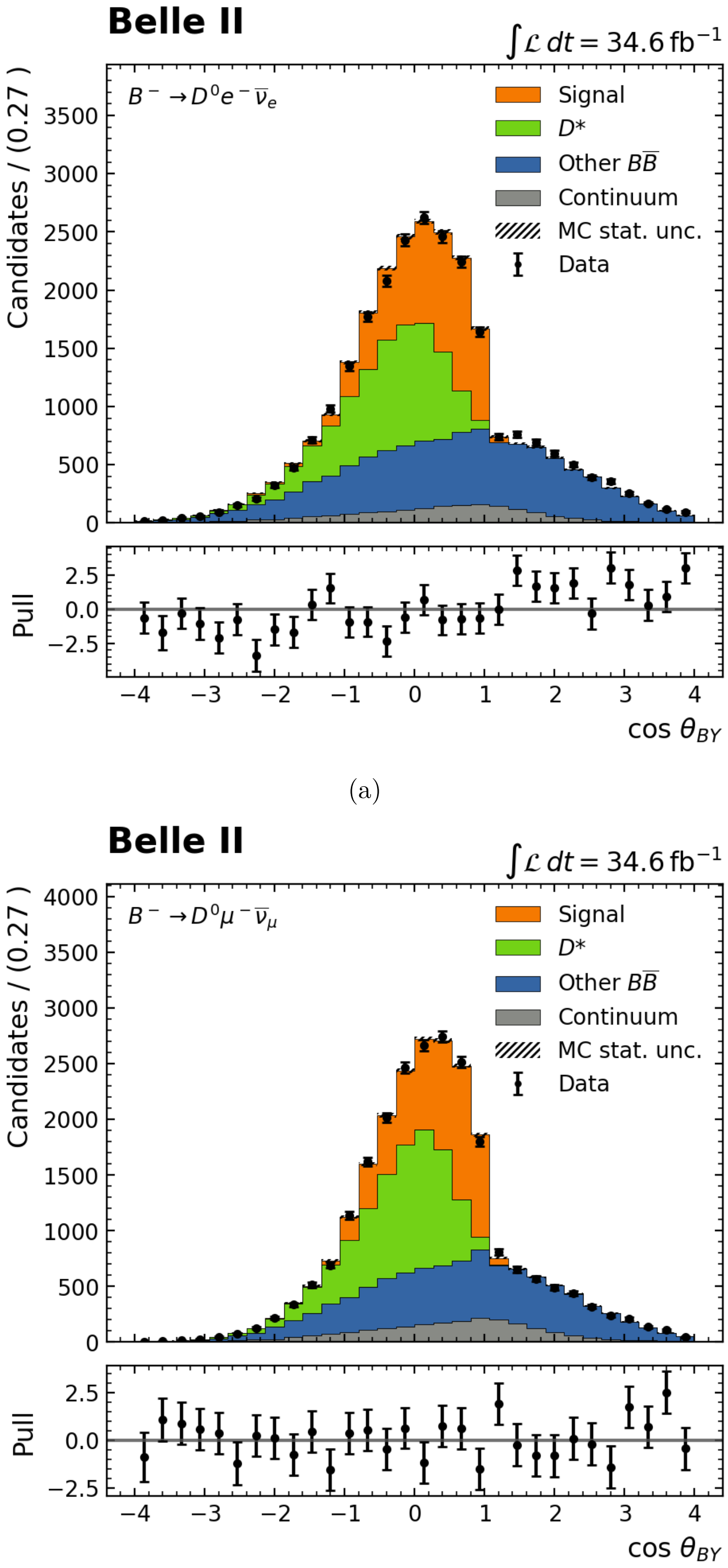
(a) (b)

FIGURE 6.2: Signal component pull distributions for 1000 generated toy samples for the (a) electron and (b) muon channel.

	MC expectation		Fitted results		
	Fraction	Yield	Fraction	Rel. stat. error	Yield
$B^- \to D^0 e^- \nu_e$					
Signal	26.0%	34993.7	$(25.40 \pm 0.96)\%$	3.78%	6186.1 ± 234.2
D^*	30.1%	40418.1	$(27.84 \pm 0.92)\%$	3.30%	6780 ± 224
Continuum	13.7%	18453.0	$(7.27 \pm 1.50)\%$	24.43%	1770 ± 432.5
Other BB	30.2%	40588.8	$(39.49 \pm 1.78)\%$	3.79%	9616.8 ± 364.6
$B^- \to D^0 \mu^- \nu_\mu$					
Signal	26.5%	34234.5	$(25.19 \pm 1.00)\%$	3.98%	5800.2 ± 231
D^*	31.5%	40761.0	$(30.98 \pm 0.99)\%$	3.20%	7132.4 ± 228.4
Continuum	13.0%	16712.1	$(9.48 \pm 1.93)\%$	20.33%	2181.5 ± 443.5
Other BB	28.9%	37340.1	$(34.35 \pm 1.76)\%$	5.11%	7907.9 ± 404.1

TABLE 6.2: A comparison between fractions in the MC and fitted data fractions. The dataset is restricted to $-4 < \cos\,\theta_{BY} < 4$. MC yields are not integer due to applied lepton identification correction factors.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a)

(b)

FIGURE 6.3: cos θ_{BY} distributions after fitting to data. Individual components are weighed by the fitted fractions resulting from the maximum likelihood fit.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Chapter 7

Results and systematic uncertainties

In this final analysis chapter, the signal yield in data evaluated from the fit is used to derive the $B^- \to D^0 \ell^- \overline{\nu}_l$ branching fraction with associated statistical error. Systematic error sources are evaluated and gaussian error propagation is used to calculate a total systematic error. The branching fraction is quoted with both statistical and systematic uncertainties.

7.1 Branching fraction

The fitting procedure resulted in a signal yield of 6186.1 ± 234.2 electron channel signal events and 5800.2 ± 231 muon channel signal events in a 34.6 fb^{-1} data sample. To calculate the resulting $B^- \to D^0 \ell^- \overline{\nu}_l$ branching fraction the efficiency and the number of total B mesons in the data sample is required. For selection criteria with n_{sig} signal events, efficiency ϵ and number of charged B mesons $n_{B^\pm}$ the branching ratio can be calculated as

$$\mathrm{Br}\left(B^- \to \left[D^0 \to K^-\pi^+\right] \ell^- \overline{\nu}_l\right) = \frac{n_{\mathrm{sig}}}{\epsilon\, n_{B^\pm}} \tag{7.1}$$

or with $\mathrm{Br}\left(B^- \to [D^0 \to K^-\pi^+]\, \ell^- \overline{\nu}_l\right) = \mathrm{Br}\left(B^- \to D^0 \ell^- \overline{\nu}_l\right) \times \mathrm{Br}\left(D^0 \to K^-\pi^+\right)$

$$\mathrm{Br}\left(B^- \to D^0 \ell^- \overline{\nu}_l\right) = \frac{n_{\mathrm{sig}}}{\epsilon\, n_{B^\pm}\ \mathrm{Br}\left(D^0 \to K^-\pi^+\right)}. \tag{7.2}$$

The branching ratio for $D^0 \to K^-\pi^+$ is measured at $\mathrm{Br}\left(D^0 \to K^-\pi^+\right) = (3.95 \pm 0.031)\%$ [12].

A dedicated study has been performed to measure the amount of B mesons in the 34.6 fb^{-1} 2019 and 2020 Belle II data sample [54]. The quoted value for the total number of $B\overline{B}$ pairs $N_{B\overline{B}} = 3.771 \cdot 10^7$ is multiplied by the ratio of charged over total B mesons f^{+-} (4.1) and by a factor of two to yield

$$n_{B^\pm} = 2 \cdot f^{+-} \cdot N_{B\overline{B}} = 3.877 \cdot 10^7. \tag{7.3}$$

The selection criteria efficiency ϵ is calculated as the ratio of lepton identification corrected MC signal events, listed in Table 6.2, over the total amount of signal events in the MC sample, listed in Table 5.6. The resulting efficiencies for electron and muon modes are $\epsilon_e = 18.36\%$ and $\epsilon_\mu = 17.66\%$.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Inserting the values for n_{sig}, ϵ, $n_{B^\pm}$ and $\text{Br}\left(D^0 \to K^-\pi^+\right)$ into the branching fraction (7.2) yields

$$\text{Br}\left(B^- \to D^0 e^- \overline{\nu}_e\right) = (2.200 \pm 0.083(\text{stat.}))\,\% \tag{7.4}$$

and

$$\text{Br}\left(B^- \to D^0 \mu^- \overline{\nu}_\mu\right) = (2.145 \pm 0.085(\text{stat.}))\,\%. \tag{7.5}$$

The ratio of the lepton mode branching ratios is measured at

$$R\left(e/\mu\right) = \frac{\text{Br}\left(B^- \to D^0 e^- \overline{\nu}_e\right)}{\text{Br}\left(B^- \to D^0 \mu^- \overline{\nu}_\mu\right)} = 1.026 \pm 0.055(\text{stat.}), \tag{7.6}$$

in agreement with lepton universality.

Electron and muon mode branching ratios can be combined into an overall $B^- \to D^0 \ell^- \overline{\nu}_l$ branching ratio. To combine n measured values x_i with variances σ_i a weighted average [12]

$$\hat{\mu} = \frac{1}{w} \sum_{i=1}^{n} w_i x_i, \tag{7.7}$$

with $w_i = \frac{1}{\sigma_i^2}$ and $w = \sum_i w_i$, is an unbiased estimator for the combined average. The standard deviation of the estimator is then given by $\frac{1}{\sqrt{w}}$.

Applying this to equations (7.4) and (7.5) yields

$$\text{Br}\left(B^- \to D^0 \ell^- \overline{\nu}_l\right) = (2.173 \pm 0.060(\text{stat.}))\,\%. \tag{7.8}$$

7.2 Systematic uncertainties

The uncertainties listed in the branching ratios (7.4) and (7.5) are purely statistical and originate from the data sample size. All other uncertainties encountered are included as additional systematic errors.

After evaluating sources of systematic errors and computing relative uncertainties $\delta a, \delta b, \delta c, \ldots$, the total systematic error on the branching fraction is computed using Gaussian error propagation

$$\sigma_{\text{tot}} = \sqrt{(\delta a)^2 + (\delta b)^2 + (\delta c)^2 + \ldots}. \tag{7.9}$$

7.2.1 Number of B mesons

The number of B mesons in the data sample is evaluated seperately for the 2019 and 2020 parts with associated statistical and systematic uncertainties [54]

$$N_{B\overline{B}}\ (2019) = (9.65 \pm 0.02(\text{stat.}) \pm 0.15(\text{sys.})) \times 10^6 \tag{7.10}$$

$$N_{B\overline{B}}\ (2020) = (28.06 \pm 0.05(\text{stat.}) \pm 0.45(\text{sys.})) \times 10^6. \tag{7.11}$$

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Additionally, the ratio of charged B mesons over total B mesons f^{+-} is measured with an associated uncertainty [19]

$$f^{+-} = 0.514 \pm 0.006. \tag{7.12}$$

The total resulting systematic on $n_{B^\pm}$ is evaluated by bootstrapping. The given errors on $N_{B\overline{B}}$ and f^{+-} are randomly drawn from a Gaussian distribution 10^6 times. Systematic errors on $N_{B\overline{B}}$ (2019) and $N_{B\overline{B}}$ (2019) are treated as fully correlated. The overall systematic error is then given as the width of the resulting $n_{B^\pm}$ distribution.
This computation yields

$$n_{B^\pm} = (3.877 \pm 0.067) \times 10^7 \tag{7.13}$$

corresponding to a relative error of 1.72%.

7.2.2 Charm branching fraction

The branching fraction of $D^0 \to K^-\pi^+$ is used directly to calculate the $B^- \to D^0\ell^-\overline{\nu}_l$ branching fraction. It is measured with an associated uncertainty [12]

$$\mathrm{Br}\left(D^0 \to K^-\pi^+\right) = (3.950 \pm 0.031)\%, \tag{7.14}$$

equivalent to a relative uncertainty of 0.78%.

7.2.3 Tracking

Uncertainties in momenta obtained from track finding contribute to an additional systematic error. A $e^+e^- \to \tau^+\tau^-$ performance study measures discrepancies between data and MC to evaluate Belle II tracking performance [55]. The study covers tracks from $p = 0.2$ to $p = 3.5$ GeV.
The resulting discrepancy in momentum between data and MC

$$\delta^* = (0.28 \pm 0.15 \pm 0.73)\,\% \tag{7.15}$$

is recommended to be taken into account by associating a relative systematic uncertainty of 0.8% per charged track. Thus, the three charged $B^- \to D^0\ell^-\overline{\nu}_l$ final state particles ℓ, K and π introduce a systematic error of 2.4%.

7.2.4 Efficiency statistics

Estimating the selection efficiency ϵ from a finite MC sample is only accurate up to a binomial statistical uncertainty. The binomial standard error for estimating a likelihood p in a sample of n entries is given by

$$\sigma_{\mathrm{binom}} = \sqrt{\frac{p\,(1-p)}{n}}. \tag{7.16}$$

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

The MC sample used contains 190563 electron and 190629 muon signal events, with calculated efficiencies of $\epsilon_e = 18.36\%$ and $\epsilon_\mu = 17.66\%$. This results in a 0.09% relative error for both modes.

7.2.5 Lepton identification

The lepton identification correction factors introduced in Section 5.4.1 are given with associated statistical and systematic uncertainties for each bin. To evaluate the total resulting uncertainty originating from lepton identification, the bootstrapping technique is used.

Starting from the correction table used, a new correction table is generated by pulling Gaussian distributed errors for each bin. The systematic errors of each bin are treated as fully correlated. Using the new table, the number of signal events in the data sample is calculated. By generating 500 tables and computing n_{sig} for each, a distribution is obtained. The width of the distribution divided by the mean yields an additional systematic error.

The resulting relative systematic uncertainty is 0.62% for the electron mode and 2.12% for the muon mode.

7.2.6 Hadron identification

For hadron identification, no performance study equivalent to the lepton correction studies has been performed. To estimate the effect of hadron identification selections on the data sample the branching fraction is calculated for varying hadron ID selections.

Requiring $\mathcal{L}_{K,\pi} > 0.2$ instead of $\mathcal{L}_{K,\pi} > 0.1$ results in a 1.36% reduced branching fraction for the electron mode, and a 2.26% reduced branching fraction for the muon mode. Requiring $\mathcal{L}_{K,\pi} > 0.5$ yields 0.33% and 3.02% reduced branching fractions for the electron mode and muon modes respectively. When omitting hadron ID requirements, the fitted branching ratio increases by 4.4% and 5.5% for electron and muon modes.

A relative systematic uncertainty of 4.5% for electrons and 5.5% for muons is included in the analysis to incorporate the observed inconsistency.

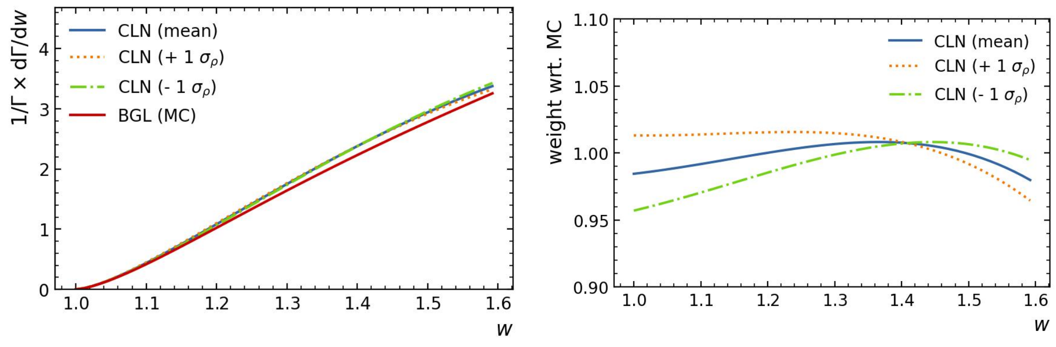
7.2.7 Binning uncertainty

The effect of chosen $\cos\theta_{BY}$ binning is estimated by varying the number of bins and fitting data. Choosing 20 and 40 bins instead of 30, the branching fraction varies by up to 0.98% in the electron mode and up to 1.35% in the muon mode. A conservative value of 1.5% is used as binning systematic for both lepton modes.

7.2.8 Deep neural network uncertainty

The DNN classifier sensitivity of the branching ratio is tested by varying the classifier rectangular selection in both directions. Requiring DNN > 0.5 instead of DNN > 0.22 results in a 2.93% higher branching ratio for the electron mode and a 2.73% lower branching fraction for the muon mode. Omitting the DNN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a) (b)

FIGURE 7.1: (a) The BGL form factor parametrization used in the MC compared to the commonly used CLN parametrization, pictured for the ρ measured mean value and $\rho = 1.131 \pm \sigma_\rho$. (b) By dividing differential decay rates of CLN by those of BGL, a weight dependent on w can be assigned to each MC signal candidate to rescale the form factor.

requirement results in 2.17% and 2.19% lower branching fractions for electrons and muons respectively.
A 3% systematic uncertainty is added to account for DNN sensitivity.

7.2.9 Form factor uncertainty

The $B^- \to D^0\ell^-\overline{\nu}_l$ decay rate is dependant on the momentum transfer from B to D meson $q^2 = (p_B - p_D)^2$. q^2 is commonly rewritten as

$$w = \frac{m_B^2 + m_D^2 - q^2}{2m_B m_D} = \frac{p_B \cdot p_D}{m_B m_D} \tag{7.17}$$

to give a limited kinematically allowed region from the zero recoil point $w = 1$ to the minimum momentum transfer point $w = \frac{m_B^2 + m_D^2}{2m_B m_D} \approx 1.6$ corresponding to $q^2 = 0$.
The differential decay rate for $B^- \to D^0\ell^-\overline{\nu}_l$

$$\frac{d\Gamma}{dw} \propto |V_{cb}|^2 |\mathcal{G}(w)|^2 \tag{7.18}$$

with the CKM matrix element V_{cb} and the form factor $\mathcal{G}(w)$. There are multiple parametrization models for $\mathcal{G}(w)$. The most commonly used parametrization is the CLN model [56]

$$\mathcal{G}(z) = \mathcal{G}(1)(1 - 8\rho^2 z + (51\rho^2 - 10)z^2 - (252\rho^2 - 84)z^3) \tag{7.19}$$

with a normalization $\mathcal{G}(1)$ and a slope ρ measured at $1.131 \pm 0.024 \pm 0.023$ [19]. The MC event generator uses a different parametrization, the BGL model [57]. To estimate the measured form factor sensitivity of the branching ratio, weights are applied to signal MC candidates to rescale the MC BGL form factor to the CLN model. This is done three times, for $\rho = 1.131$ and for $\rho = 1.131 \pm \sigma_\rho$.

Form Factor	e mode	Deviation	μ mode	Deviation
BGL	2.20%		2.15%	
CLN (mean)	2.20%	+0.06%	2.15%	+0.06%
CLN ($+1\sigma_\rho$)	2.19%	-0.24%	2.14%	-0.24%
CLN ($-1\sigma_\rho$)	2.21%	+0.34%	2.16%	+0.35%

TABLE 7.1: The fitted branching ratios and the deviation from the BGL model after reweighting the form factor to the CLN model.

	Relative uncertainty [%]	
Source	Electron mode	Muon mode
$n_{B^\pm}$	1.72	1.72
$\mathrm{Br}\left(D^0 \to K^-\pi^+\right)$	0.78	0.78
Tracking	2.4	2.4
Efficiency statistics	0.09	0.09
Lepton identification	0.62	2.12
Hadron identification	4.5	5.5
Form factor	0.35	0.35
Deep neural network uncertainty	3	3
Binning	1.5	1.5
Total	6.49	7.50

TABLE 7.2: Summary of the systematic errors for the $B^- \to D^0\ell^-\overline{\nu}_l$ branching fraction measurement.

After applying weights, the distributions are fit to data. The resulting branching fractions are shown in Table 7.1.
A systematic uncertainty of 0.35% is applied to both modes to account for the chosen form factor.

7.3 Branching fraction result

A final measurent result is obtained by combining statistical and systematic uncertainties:

$$\mathrm{Br}\left(B^- \to D^0 e^- \overline{\nu}_e\right) = (2.200 \pm 0.083(\text{stat.}) \pm 0.143(\text{sys.}))\,\% \tag{7.20}$$

$$\mathrm{Br}\left(B^- \to D^0 \mu^- \overline{\nu}_\mu\right) = (2.145 \pm 0.085(\text{stat.}) \pm 0.161(\text{sys.}))\,\%. \tag{7.21}$$

For the ratio of electron and muon mode branching fractions $R\,(e/\mu)$, the systematic errors on $n_{B^\pm}$ and $\mathrm{Br}\,(D^0 \to K^-\pi^+)$ cancel out and can be omitted, resulting in a total relative systematic uncertainty of 7.2%. This yields a ratio

$$R\,(e/\mu) = 1.026 \pm 0.055(\text{stat.}) \pm 0.074(\text{sys.}) \tag{7.22}$$

in agreement with lepton universality.
For the combined $B^- \to D^0\ell^-\overline{\nu}_l$ branching fraction, the systematic error is

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Experiment	$\mathrm{Br}\left(B^- \to D^0\ell^-\overline{\nu}_l\right)$ [%]
CLEO [46]	$2.19 \pm 0.13 \pm 0.17$
BaBar [58]	$2.19 \pm 0.08 \pm 0.13$
Belle [43]	$2.53 \pm 0.04 \pm 0.12$
HFLAV average [19]	$2.35 \pm 0.03(\text{stat.}) \pm 0.09(\text{sys.})$

TABLE 7.3: Previous $B^- \to D^0\ell^-\overline{\nu}_l$ measurements and the combined world average.

obtained by taking the larger value for every systematic uncertainty listed in Table 7.2. In this case, the muon mode uncertainty is always greater than or equal to the electron channel uncertainty, resulting in a 7.50% relative systematic uncertainty. This yields

$$\mathrm{Br}\left(B^- \to D^0\ell^-\overline{\nu}_l\right) = (2.173 \pm 0.060(\text{stat.}) \pm 0.163(\text{sys.}))\,\%. \tag{7.23}$$

Previous $\mathrm{Br}\left(B^- \to D^0\ell^-\overline{\nu}_l\right)$ measurements and the combined world average calculated by the Heavy Flavor Averaging Group (HFLAV) [19] are shown in Table 7.3.

The obtained branching fraction is in very good agreement with CLEO and BaBar measurements and within one combined σ of the world average. The Belle measurement is 1.6 combined σ removed of the Belle II measurement in this analysis.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Chapter 8

Conclusions and outlook

8.1 Conclusions

In this thesis, an analysis of the decay $B^- \to D^0\ell^-\overline{\nu}_l$ using 2019 and 2020 Belle II data was performed. The decay was reconstructed for $\ell = e, \mu$ and via $D^0 \to K^-\pi^+$. Before applying the reconstruction to data, signal selection was optimized on simulated Monte-Carlo collisions using techniques from multivariate data analysis. Rectangular selections were applied on variables of interest in order to optimize the statistical significance of the resulting dataset, quantized by the figure of merit. An important type of background observed in the analysis is downfeed from $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_l$. By employing a deep neural network, a classifier was trained to seperate signal from D^* background and suppress the amount of background in the final dataset.
After optimizing the selection criteria, distributions of kinematic variables were compared between Belle II physics runs and simulated Monte Carlo events. Discrepancies in data and MC were corrected by applying weights to Monte-Carlo events.
By fitting the MC distributions to the data, the decay composition in data was determined. A binned maximum likelihood fit was validated by applying it to varying sets of pseudo-data, to confirm an unbiased and efficient fitting procedure. Applying the validated fitting procedure to data yields a number of signal events in the dataset, that was used to calculate the $B^- \to D^0\ell^-\overline{\nu}_l$ branching fraction.
Taking into account statistical errors originating from the limited sample size, and ocurring systematic uncertainties from multiple sources, the final branching ratio

$$\mathrm{Br}\left(B^- \to D^0\ell^-\overline{\nu}_l\right) = (2.173 \pm 0.060(\mathrm{stat.}) \pm 0.163(\mathrm{sys.}))\,\%.$$

was obtained.
The resulting branching ratio is in good agreement with previous measurements at other experiments Belle, BaBar and CLEO and with the combined world average $(2.35 \pm 0.03 \pm 0.09)\%$.
A ratio of electron and muon mode branching fractions

$$R\,(e/\mu) = \frac{\mathrm{Br}\,(B^- \to D^0 e^-\overline{\nu}_e)}{\mathrm{Br}\,(B^- \to D^0\mu^-\overline{\nu}_\mu)} = 1.026 \pm 0.055(\mathrm{stat.}) \pm 0.074(\mathrm{sys.}),$$

was calculated, showing experimental agreement with lepton universality.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

8.2 Outlook

The analyzed dataset contains 0.0346 ab^{-1} of data with a total of 50 ab^{-1} planned over the full Belle II runtime. With a larger dataset and additional performance studies on detector sensitivity, statistical and systematic uncertainties can be further decreased. With a dataset of $\sim$ 1 ab^{-1} the branching fraction measurement significance is expected to exceed all previous measurements.

Following up on the branching fraction measurement, the next step is to obtain the CKM matrix element $|V_{cb}|$. By binning the branching fraction in momentum transferred to the D meson q^2 and extrapolating the distribution to the zero recoil point q^2_{max}, $|V_{cb}|$ can be calculated. The analysis in this thesis shows no major technical roadblocks for achieving a precise $|V_{cb}|$ measurement with $B^- \to D^0\ell^-\overline{\nu}_l$ reconstruction at Belle II.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Bibliography

[1] Public domain image. URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

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

[3] F. Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Jan. 1984. ISBN: 978-0-471-88741-6.

[4] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 2013. DOI: 10.1017/CBO9781139525367.

[5] Thomas Hambye. "CP violation and the matterantimatter asymmetry of the Universe". In: *Comptes Rendus Physique* 13.2 (2012). Flavour physics and CP violation / Physique de la saveur et violation de CP, pp. 193 –203. ISSN: 1631-0705. DOI: https://doi.org/10.1016/j.crhy.2011.09.007. URL: http://www.sciencedirect.com/science/article/pii/S1631070511001873.

[6] Edward W. Kolb and Michael S. Turner. *The early universe*. Vol. 69. 1990.

[7] Gary Steigman. *Primordial Nucleosynthesis: The Predicted and Observed Abundances and Their Consequences*. 2010. arXiv: 1008.4765 [astro-ph.CO].

[8] E. Komatsu et al. "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation". In: *Astrophys. J. Suppl.* 192 (2011), p. 18. DOI: 10.1088/0067-0049/192/2/18. arXiv: 1001.4538 [astro-ph.CO].

[9] A.D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Sov. Phys. Usp.* 34.5 (1991), pp. 392–393. DOI: 10.1070/PU1991v034n05ABEH002497.

[10] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[11] Gregory Ciezarek et al. "A challenge to lepton universality in B-meson decays". In: *Nature* 546.7657 (June 2017), pp. 227233. ISSN: 1476-4687. DOI: 10.1038/nature22346. URL: http://dx.doi.org/10.1038/nature22346.

[12] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev.* D98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

[13] Lincoln Wolfenstein. “Parametrization of the Kobayashi-Maskawa Matrix”. In: *Phys. Rev. Lett.* 51 (21 Nov. 1983), pp. 1945–1947. DOI: 10.1103/PhysRevLett.51.1945. URL: https://link.aps.org/doi/10.1103/PhysRevLett.51.1945.

[14] Renata Jora. *Baryon asymmetry in the standard model revisited.* 2018. arXiv: 1806.00597 [hep-ph].

[15] A G Cohen, D B Kaplan, and A E Nelson. “Progress in Electroweak Baryogenesis”. In: *Annual Review of Nuclear and Particle Science* 43.1 (Dec. 1993), pp. 2770. ISSN: 1545-4134. DOI: 10.1146/annurev.ns.43.120193.000331. URL: http://dx.doi.org/10.1146/annurev.ns.43.120193.000331.

[16] Richard F. Lebed. “Relating CKM parametrizations and unitarity triangles”. In: *Phys. Rev. D* 55 (1997), pp. 348–352. DOI: 10.1103/PhysRevD.55.348. arXiv: hep-ph/9607305.

[17] Konrad Kleinknecht and Burkhard Renk. “Unitarity triangle from CP invariant quantities”. In: *Physics Letters B* 639.6 (Aug. 2006), pp. 612615. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2006.07.015. URL: http://dx.doi.org/10.1016/j.physletb.2006.07.015.

[18] A. Sibidanov et al. “Search for $B^- \to \mu^- \bar{\nu}_\mu$ Decays at the Belle Experiment”. In: *Phys. Rev. Lett.* 121.3 (2018), p. 031801. DOI: 10.1103/PhysRevLett.121.031801. arXiv: 1712.04123 [hep-ex].

[19] Heavy Flavor Averaging Group et al. *Averages of b-hadron, c-hadron, and -lepton properties as of 2018.* 2019. arXiv: 1909.12524 [hep-ex].

[20] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. “SuperKEKB collider”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (Aug. 2018). DOI: 10.1016/j.nima.2018.08.017.

[21] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. “SuperKEKB collider”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 907 (Nov. 2018), pp. 188199. ISSN: 0168-9002. DOI: 10.1016/j.nima.2018.08.017. URL: http://dx.doi.org/10.1016/j.nima.2018.08.017.

[22] Yukiyoshi Ohnishi et al. “Accelerator design at SuperKEKB”. In: *Progress of Theoretical and Experimental Physics* 2013.3 (Mar. 2013). 03A011. ISSN: 2050-3911. DOI: 10.1093/ptep/pts083. eprint: https://academic.oup.com/ptep/article-pdf/2013/3/03A011/4439973/pts083.pdf. URL: https://doi.org/10.1093/ptep/pts083.

[23] Kyo Shibata. “Status and schedule of SuperKEKB”. 2011.

[24] G. B. Mohanty. “Belle II status and prospects”. *Beauty 2019* conference talk. Oct. 2019.

[25] T. Abe et al. “Belle II Technical Design Report”. Nov. 2010.

[26] E Kou et al. “The Belle II Physics Book”. In: *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019). ISSN: 2050-3911. DOI: 10.1093/ptep/ptz106. URL: http://dx.doi.org/10.1093/ptep/ptz106.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

[27] P.M. Lewis et al. “First measurements of beam backgrounds at SuperKEKB”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 914 (Jan. 2019), pp. 69144. ISSN: 0168-9002. DOI: `10.1016/j.nima.2018.05.071`. URL: `http://dx.doi.org/10.1016/j.nima.2018.05.071`.

[28] Christoph Schwanda. “Determination of the Cabbibo-Kobayashi-Maskawa Matrix Elements $|V_{cb}|$ and $|V_{ub}|$ at the Belle Experiment”. Aug. 2012.

[29] M. Akatsu et al. “Time-of-propagation Cherenkov counter for particle identification”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 440.1 (2000), pp. 124 –135. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/S0168-9002(99)00819-0`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900299008190`.

[30] SuperKEKB collider achieves the world’s highest luminosity. June 2020. URL: `https://www.kek.jp/en/newsroom/2020/06/26/1400/`.

[31] T Kuhr et al. “The Belle II core software”. In: *Computing and Software for Big Science* 3.1 (2019), p. 1.

[32] David J Lange. “The EvtGen particle decay simulation package”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1-2 (2001), pp. 152–155.

[33] Sea Agostinelli et al. “GEANT4a simulation toolkit”. In: *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303.

[34] Joseph T. Nunziata, Atanu Pathak, and Swagato Banerjee. *Charged Particle Identification Using Calorimetry and Tracking at the Belle II Experiment.* Apr. 2020. URL: `https://ir.library.louisville.edu/uars/25`.

[35] Geoffrey C Fox and Stephen Wolfram. “Observables for the analysis of event shapes in e+ e- annihilation and other processes”. In: *Physical Review Letters* 41.23 (1978), p. 1581.

[36] B. Asavapibhop, N. Suwonjandee, and J. Ineead. *EWP Research Progress 2020 #2.* May 2020.

[37] A. J. Bevan et al. “The Physics of the B Factories”. In: *The European Physical Journal C* 74.11 (Nov. 2014). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-014-3026-9`. URL: `http://dx.doi.org/10.1140/epjc/s10052-014-3026-9`.

[38] Y. Miyazaki et al. “Search for Lepton-Flavor-Violating and Lepton-Number-Violating $\tau \to \ell h h'$ Decay Modes”. In: *Phys. Lett. B* 719 (2013), pp. 346–353. DOI: `10.1016/j.physletb.2013.01.032`. arXiv: `1206.5595 [hep-ex]`.

[39] Giovanni Punzi. *Sensitivity of searches for new signals and its optimization.* 2003. arXiv: `physics/0308063 [physics.data-an]`.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

[40] Bernard Aubert et al. “A Search for $B^+ \to \ell^+\nu_\ell$ Recoiling Against $B^- \to D^0\ell^-\bar{\nu}X$”. In: *Phys. Rev. D* 81 (2010), p. 051101. DOI: 10.1103/PhysRevD.81.051101. arXiv: 0912.2453 [hep-ex].

[41] Robert E. Schapire. “The Strength of Weak Learnability”. In: *Mach. Learn.* 5.2 (July 1990), pp. 197227. ISSN: 0885-6125. DOI: 10.1023/A:1022648800760. URL: https://doi.org/10.1023/A:1022648800760.

[42] W. Dungel et al. “Measurement of the form factors of the decay $B^0 - > D^{*-}\ell + \nu$ and determination of the CKM matrix element $|V_{cb}|$”. In: *Phys. Rev. D* 82 (2010), p. 112007. DOI: 10.1103/PhysRevD.82.112007. arXiv: 1010.5620 [hep-ex].

[43] R. Glattauer et al. “Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element$|V_{cb}|$”. In: *Physical Review D* 93.3 (Feb. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.032006. URL: http://dx.doi.org/10.1103/PhysRevD.93.032006.

[44] H. Albrecht et al. “Search for hadronic bu decays”. In: *Physics Letters B* 241.2 (1990), pp. 278 –282. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(90)91293-K. URL: http://www.sciencedirect.com/science/article/pii/037026939091293K.

[45] In Jae Myung. “Tutorial on maximum likelihood estimation”. In: *Journal of Mathematical Psychology* 47.1 (2003), pp. 90 –100. ISSN: 0022-2496. DOI: https://doi.org/10.1016/S0022-2496(02)00028-7. URL: http://www.sciencedirect.com/science/article/pii/S0022249602000287.

[46] Roger Barlow and Christine Beeston. “Fitting using finite Monte Carlo samples”. In: *Computer Physics Communications* 77.2 (1993), pp. 219 –228. ISSN: 0010-4655. DOI: https://doi.org/10.1016/0010-4655(93)90005-W. URL: http://www.sciencedirect.com/science/article/pii/001046559390005W.

[47] Harald Cramér. *Mathematical methods of statistics*. Vol. 43. Princeton university press, 1999.

[48] Luc Demortier and Louis Lyons. “Everything you always wanted to know about pulls”. URL: http://physics.rockefeller.edu/luc/technical_reports/cdf5776_pulls.pdf.

[49] A. Kulesa et al. “Sampling distributions and the bootstrap”. In: *Nat. Methods* 12.6 (June 2015), pp. 477–478.

[50] Andrea Selce. “Soft pi0 investigation on phase3 data”. Belle II Note. Apr. 2020. URL: https://docs.belle2.org/record/1953.

[51] Andreas Hocker et al. “TMVA - Toolkit for Multivariate Data Analysis”. Mar. 2007.

[52] Daniel Ferlewicz, Marco Milesi, and Phillip Urquijo. “Lepton identification performance using $J\psi \to \ell^+\ell^-$ and $K_S^0\pi\pi^-$ in Phase 3 data at Belle II”. Belle II Note. July 2019. URL: https://docs.belle2.org/record/1581.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[53] Rashmi Dhamija, Saurabh Sandilya, and Anjan Giri. "Study of Lepton ID Fake Rate using D* sample." Belle II Note. June 2020. URL: https://docs.belle2.org/record/1953.

[54] The Belle II Collaboration. "Approved B counting analysis results with ICHEP2020 dataset". Belle II Note. July 2020. URL: https://docs.belle2.org/record/1997.

[55] Alexander Glazov et al. "Measurement of the tracking efficiency in Phase 3 data using tau-pair events." Belle II Note. Feb. 2020. URL: https://docs.belle2.org/record/2035.

[56] Irinel Caprini, Laurent Lellouch, and Matthias Neubert. "Dispersive bounds on the shape of form factors". In: *Nuclear Physics B* 530.1-2 (1998), pp. 153181. ISSN: 0550-3213. DOI: 10.1016/s0550-3213(98)00350-2. URL: http://dx.doi.org/10.1016/S0550-3213(98)00350-2.

[57] C. Glenn Boyd, Benjamín Grinstein, and Richard F. Lebed. "Precision corrections to dispersive bounds on form factors". In: *Physical Review D* 56.11 (1997), pp. 68956911. ISSN: 1089-4918. DOI: 10.1103/physrevd.56.6895. URL: http://dx.doi.org/10.1103/PhysRevD.56.6895.

[58] B. Aubert et al. "Measurement of $|V_{cb}|$ and the Form-Factor Slope in $\overline{B} \to D\ell\overline{\nu}_\ell$ Decays in Events Tagged by a Fully Reconstructed B Meson". In: *Physical Review Letters* 104.1 (Jan. 2010). ISSN: 1079-7114. DOI: 10.1103/physrevlett.104.011802. URL: http://dx.doi.org/10.1103/PhysRevLett.104.011802.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.