DIPLOMARBEIT

Measurement of the $B^0 \to D^{*-}\ \ell^+\ \nu_\ell$ Branching Fraction using the Belle II data

ausgeführt an der
Fakultät für Physik der Technischen Universität Wien

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

zur Erlangung des akademischen Grades eines

Diplom-Ingeneurs

im Rahmen des Studiums

066 461 Technische Physik

unter der Anleitung von

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

eingereicht von

Daniel Dorner, BSc
Matr.Nr.: 01128534

Wien, am 16. Oktober 2020 _____   _____
Unterschrift Verfasser   Unterschrift Betreuer

Zusammenfassung

Das Ziel dieser Masterarbeit ist es eine Messung des Verzweigungsverhältnisses des B Mesonen Zerfalls $B \to D^*\ell\nu_\ell$ ($\ell = e, \mu$) durchzuführen anhand einer Analyse der Daten, die in den frühen Phasen des Belle II Experiments aufgezeichnet worden sind. Belle II wird am asymmetrischen Elektron-Positron Collider SuperKEKB von der Japanischen Hochenergie Beschleuniger Forschungsorganisation (KEK) durchgeführt. Der Collider befindet sich in Tsukuba, Japan und das Experiment hat mit der Aufzeichnung von Daten Anfang 2018 begonnen.

Semileptonische Zerfälle wie $B \to D^*\ell\nu_\ell$ können dazu verwendet werden um das Matrixelement $|V_{cb}|$ der 3×3 Cabibbo-Kobayashi-Maskawa (CKM) Matrix mit hoher Präzession zu messen. Die CKM Matrix beschreibt die Mischung von Quarks, die bei der schwachen Wechselwirkung beobachtet werden kann. Zum Beispiel beschreibt das Matrixelement $|V_{cb}|$ den Übergang von Bottom zu Charm Quarks.

Für die Datenanalyse wurden B Mesonen des Zerfalls $B \to D^*\ell\nu_\ell$ unter Verwendung der zwei Unterzerfälle $D^* \to D\pi$ und $D \to K\pi$ rekonstruiert. Der Rekonstruktionsprozess wurde mit dem Belle II Software Framework basf2 durchgeführt. Es wurde ein Datensatz von 34.6 fb^{-1}, der bei Belle II aufgezeichnet worden ist, verwendet und mit Monte Carlo (MC) generierten Simulationen mit einer integrierten Luminosität von 200 fb^{-1} verglichen.

Nach der Rekonstruktion der B Meson Kandidaten wurde die Signalauswahl basierend auf Signal Peaks in verschiedenen Verteilungen und der Figure of Merit optimiert. Nach der Optimierung des Signals wurde die Übereinstimmung zwischen MC und Daten überprüft und Diskrepanzen durch die Implementierung von Korrekturgewichten korrigiert.

Das Signal wurde von den rekonstruierten Ereignissen des Datensatzes mit Hilfe eines Maximum Likelihood Fits basierend auf den MC Simulationen extrahiert. Unter Verwendung der vom Fit geschätzten Anzahl von Signal Ereignissen, der Effizienz der Rekonstruktionskette, der Anzahl von B Mesonen, die im Datensatz enthalten sind, und des weltweiten Durchschnittswertes der Verzweigungsverhältnisse von den Unterzerfällen $D^* \to D\pi$ und $D \to K\pi$, die von der Particle Data Group (PDG) bereitgestellt werden, wurde das Verzweigungsverhältnis von $B \to D^*\ell\nu_\ell$ bestimmt. Nach der Berechnung des Verzweigungsverhältnisses wurde die totale, relative systematische Unsicherheit der Messung bestimmt.

Die Analyse führte zu einem Verzweigungsverhältnis von $(4.669 \pm 0.078(stat) \pm 0.495(sys))\%$ und $(4.576 \pm 0.067(stat) \pm 0.494(sys))\%$ für die Elektronen beziehungsweise Myonen Mode. Ebenso wurde das kombinierte Verzweigungsverhältnis beider Moden mit Hilfe eines nach Unsicherheiten gewichteten Mittelwerts berechnet, der zu einem Ergebnis von $(4.616 \pm 0.051(stat) \pm 0.499(sys))\%$ führte. Der weltweite Durchschnittswert liegt bei $(5.05 \pm 0.14)\%$ laut PDG und ist somit innerhalb der gesamten Messunsicherheit der Messungen. Um die Leptonen-Universalität der schwachen Wechselwirkung zu testen wurde das Verhältnis des Verzweigungsverhältnisses der Elektronen Mode zur Myonen Mode berechnet, was zu einem Ergebnis von $1.020 \pm 0.023(stat) \pm 0.106(sys)$ führte. Ein Verhältnis von 1 wird von dem Standard Model vorausgesagt, was innerhalb des statistischen Fehlers liegt.

Eine Gegenprobe der Ergebnisse wurde mit einer unabhängigen Analyse

desselben Zerfalls des PhD Studenten Chaoyi Lyu der Universität von Bonn durchgeführt um die Vorgangsweise bei der Analyse dieser Masterarbeit zu verifizieren. Diese Gegenprobe hat gezeigt, dass obwohl verschiedene Techniken verwendet wurden die Ergebnisse der beiden unabhängigen Analysen gut übereinstimmen und die Ergebnisse innerhalb der statischen Messunsicherheit liegen.

Abstract

The goal of this master thesis is to measure the branching fraction of the B meson decay $B \to D^*\ell\nu_\ell$ ($\ell = e, \mu$) by analysing the early data obtained at the ongoing Belle II experiment. Belle II is being conducted at the asymmetric electron positron collider SuperKEKB by the Japanese High Energy Accelerator Research Organisation (KEK). This collider is located in Tsukuba, Japan and the experiment started taking data in early 2018.

Semileptonic decays like $B \to D^*\ell\nu_\ell$ can be used to make high precision measurements of the matrix element $|V_{cb}|$ of the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) Matrix. The CKM-Matrix describes the quark mixing occurring in weak interactions. E.g. the matrix element $|V_{cb}|$ describes the transition between bottom and charm quarks.

For this data analysis the B meson of the $B \to D^*\ell\nu_\ell$ decay was reconstructed using the two sub-decays $D^* \to D\pi$ and $D \to K\pi$. The reconstruction was done using the Belle II software framework basf2. It was performed on the early data set of 34.6 fb^{-1} obtained at Belle II and compared to Monte Carlo (MC) generated samples with an integrated luminosity of 200 fb^{-1}.

After the reconstruction of the B meson candidates the signal selection was optimized based on signal peaks in various distributions and the figure of merit. Following the signal selection the Data-MC agreement was checked and discrepancies corrected by implementing correction weights.

The signal was then extracted from the reconstructed data events using a maximum likelihood fit based on the MC sample. Using the number of signal events estimated by the fit, the efficiency of the reconstruction chain, the number of B mesons contained in the data sample and the world average branching fractions of the sub-decays $D^* \to D\pi$ and $D \to K\pi$ provided by the Particle Data Group (PDG) the branching fraction of $B \to D^*\ell\nu_\ell$ was determined. After calculating the branching fraction the total relative systematic uncertainty of the measurements were determined.

This analysis yielded branching fractions of $(4.669 \pm 0.078(stat) \pm 0.495(sys))\%$ and $(4.576 \pm 0.067(stat) \pm 0.494(sys))\%$, for the electron and muon mode, respectively. Additionally the combined branching fraction was determined by calculating the uncertainty weighted mean of both modes, which resulted in $(4.616 \pm 0.051(stat) \pm 0.499(sys))\%$. The world average of $(5.05 \pm 0.14)\%$ provided by the PDG lies within the range of the total uncertainty of these measurements. To test the lepton universality of the weak interaction the ratio between the electron and muon mode branching fraction was calculated, which yielded $1.020 \pm 0.023(stat) \pm 0.106(sys)$. Unity is predicted by the Standard Model for this ratio and lies within the range of the statistical uncertainty.

A cross check of the final results was done with an indepedent analysis of the same decay mode by the PhD student Chaoyi Lyu from Bonn university to verify the analysis process of this thesis. This cross check showed that while using different techniques the results of both independent analyses agree within the statistical uncertainties.

Contents

1 Introduction

Modern particle physics are described by the Standard Model of particle physics (SM). The SM describes the elemental particles, i.e. spin=1/2 fermions (quarks. leptons) and spin=1 gauge bosons and their electromagnetic, weak and strong interactions among each other.

Thus far the predictions of the SM were proven to be correct by experimental measurements. One example of these predictions would be the Higgs boson, which was only recently discovered in 2012 at the LHC particle collider at CERN. However the current SM is not complete since it can not explain phenomena like the matter and antimatter asymmetry, gravity or the origins of dark matter.[1]

One way to search for new physics or verify the SM are precise measurements of SM predictions of parameters. One set of parameters that is studied are the matrix elements of the Cabibbo-Kobayashi-Maskawa Matrix (CKM-Matrix) formulated by Makoto Kobayashi and Toshihide Maskawa The CKM matrix explained the mixing of quarks and predicted the CP (charge conjugation + parity transformation) symmetry violations. The hypothesis of Kobayashi and Maskawa was verified by the Belle I experiment (1999-2010) by measuring CP violations in the decays of B mesons. This verification lead to Kobayashi and Maskawa being awarded the Nobel Prize in physics in 2008. [[1], [2]]

The CKM matrix elements $|V_{cb}|$ and $|V_{ub}|$ can be measured using B^0 and $B^\pm$ meson decays that including either matrix element. Based on a measured branching fraction and by applying a theoretical model for QCD(quantum chromodynamic) contributions to the weak decay processes the matrix elements $|V_{cb}|$ and $|V_{ub}|$ can be determined. For these measurements the branching fractions of semileptonic B meson decays are used, i.e. decays leading to a lepton and its respective neutrino as well as at least one hadron. [3]

These measurements of $|V_{cb}|$ and $|V_{ub}|$ can be done either in an inclusive or exclusive analysis. Both measurements should lead to the same results, however at the moment this is not the case. The inclusive and exclusive measurements of $|V_{cb}|$ and $|V_{ub}|$ show a discrepancy of 3 σ and 4 σ, respectively.[4] Since inclusive and exclusive measurements use different theoretical descriptions of QCD contributions this could hint at a flaw in these theories. Another example could be that flaws in detectors during the data taking caused the discrepancies. The differences could also hint to new physics. Due to these discrepancies bigger data samples taken by newer detectors will be very valuable for understanding this issue. [3]

One experiment that will provide these bigger and newer data samples is the Belle II experiment, which is currently running at the SuperKEKB electron positron collider in Tsukuba, Japan at the Japanese High Energy Research Organisation (KEK) and started taking data in 2018. [5]

The goal of this thesis is to make an exlusive, untagged measurement of the branching fraction of the semileptonic decay $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ ($\ell = e, \mu$) using a data sample with an integrated luminosity of 34.6 fb^{-1} obtained at Belle II. The analysis can be continued and the measurement of the branching fraction lays the ground work for a determination of the CKM matrix element $|V_{cb}|$. [6]

The thesis begins with an overview of the relevant theoretical background in Chapter 2. This is followed up by a discussion of the SuperKEKB collider and

the Belle II detector in Chapter 3. Chapter 4 introduces general principles used during the data analysis of Belle II data. Chapter 5 discusses the reconstruction process of $B^0 \to D^{*-}\ell\nu_\ell$ and the signal selection as well as the agreement between data and MC simulated samples. The fitting procedure used for the signal extraction will be discussed in Chapter 6 and the results of the analysis are shown in Chapter 7 including cross checks and the estimation of systematic uncertainties. In Chapter 8 a conclusion to the results obtained during the analysis will be given.

2 Theoretical Overview

2.1 Natural Units

In this thesis all particle related quantities will be given using the natural units as the system of units. In this system the S.I. units kg, m and s are replaced by the reduced planck constant $\hbar = 1.055 \times 10^{-34}$ J s , the speed of light $c = 2.998 \times 10^8$ m s^{-1} and GeV with 1 GeV $= 1.602 \times 10^{-10}$ J. Furthermore it is chosen that $\hbar = c = 1$ and therefore all quantities can be expressed in powers of GeV. [1]

2.2 Standard Model of Particle Physics

The Standard Model of particle physics (SM) is a quantum field theory that describes the elemental particles and the fundamental interactions (forces) between them, with the exception of gravity. The elementary particles of the SM are fermions and the force intermediating gauge bosons. The forces described by the SM are the electromagnetic, weak and strong interactions. [[1], [7]]

The fermions have a spin of 1/2 and are divided into two types, i.e. leptons and quarks. The fermions can be split into three different generations, as depicted in Figure 2.1. Every lepton generation contains one particle with an electric charge of -1 and a neutrino, which is charge neutral. Every quark generation consists of a quark with an electric charge of $+2/3$ and a quark with an electric charge of -1/3. Apart from their mass, which increases with every generation, these three generations share the same physical properties. The quarks additionally carry a colour charge and are divided into the six different flavours up(u), down(d), charm(c), strange(s), top(t) and bottom(b). The colour charge they are carrying can either be blue, red or green. Furthermore there also exist charge conjugated anti particles for each of the fermions, which have charges(including colour charges) of opposite signs but the same mass.[[1], [7]]

The bosons have a spin of 1, with the exception of the Higgs boson with a spin of 0, and are the force carriers of the SM. An overview of the gauge bosons and their corresponding interaction is given in Table 2.1. Each of the forces in the SM is described by their respective Quantum Field Theory (QTF) and is part of a gauge symmetry group. To be affected by a certain interaction a particle has to carry the charge corresponding to the given force, e.g. only electrically charged particles can partake in electromagnetic interactions. An overview of the charges of each fermion type is given in Table 2.2.[[1], [7]]

Boson	Mass (GeV)	Interaction
Gluon	0	Strong
Photon	0	Electromagnetic
W$^\pm$	80.395 ± 0.015	Weak
Z^0	91.1876 ± 0.0021	Weak

Table 2.1: Gauge bosons. [7]

Standard Model of Elementary Particles
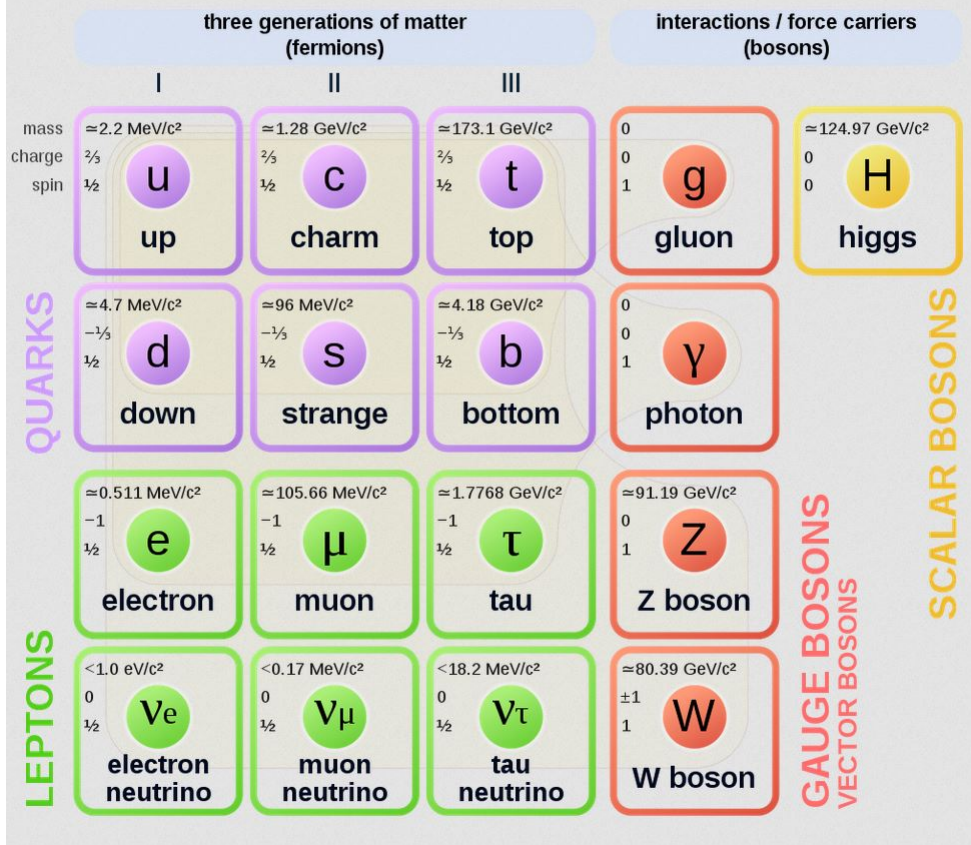
Figure 2.1: Elementary particles in the SM. [8]

Type	strong	electromagnetic	weak
Quarks	×	×	×
Charged leptons		×	×
Neutrinos			×

Table 2.2: Interaction summary of fermions. [1]

The strong interaction is described by Quantum Chromodynamics (QCD) and is part of the SU(3) symmetry group. The charge corresponding to the strong interaction is the colour charge, therefore only quarks can partake in strong interactions. The SU(3) symmetry of the QCD gives rise to eight different colour charge carrying, massless gluons as force mediators. Each gluon carries a colour and anti-colour charge. Because of the properties of QCD quarks can never be observed as isolated, free particles in experiments and are only seen as colour neutral bound states called hadrons. This phenomenon is called colour confinement. Because of the colour confinement there are only two types of hadrons, i.e. mesons, which consist of one quark and one anti-quark (e.g. π^+: $u\bar{d}$) and baryons, which consist of three quarks or anti-quarks (e.g. protons: uud). The strong interaction between two quarks gets stronger if the distance between them increases, which results in new quark pairs being formed if a certain energy threshold is reached. This process is called hadronization. [[1], [7]]

Quantum Electrodynamics (QED) describe the electromagnetic interaction between electrically charged particles. QED has a U(1) symmetry, which results in a massless and chargeless photon γ as the force mediator that gets exchanged during the interaction. Quarks and charged leptons can be affected by the electromagnetic interaction. [[1], [7]]

The weak interaction has a SU(2) symmetry and its gauge bosons are the massive, electrically neutral Z^0 and electrically charged $W^\pm$ bosons with a mass of 91.2 GeV and 80.4 GeV, respectively. At high energies of the order of 246 GeV the electromagnetic and weak interaction can be merged into the electroweak force with a SU(2) × U(1) symmetry. [[1], [7]]

2.3 Weak Interaction

The weak interaction has several unique traits:

- It is the only fundamental interaction that has massive gauge bosons, i.e. the electrically charged $W^\pm$ and charge neutral Z^0 bosons. [1]

- The weak interaction is the only interaction that can change the flavour of quarks. [1]

- The parity is not conserved due to its V-A structure and CP (charge conjugation + parity transformation) violation occurs due to quark mixing. [1]

Two examples of weak decays are the decays $B^+ \to \overline{D}^0 \pi^+$ and $B^+ \to \overline{D}^0 K^+$. The dominant Feynman diagrams of these decays are pictured in Figure 2.2.

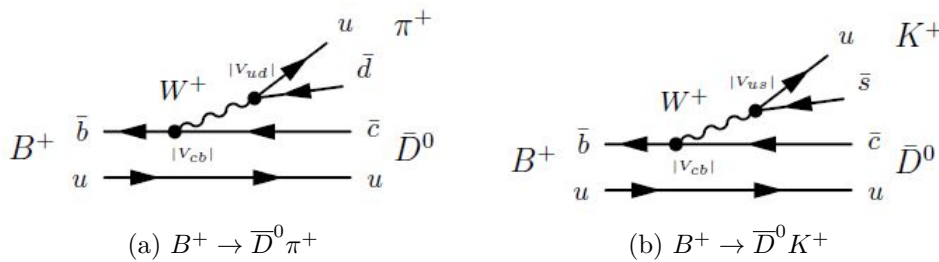(a) $B^+ \to \overline{D}^0\pi^+$ (b) $B^+ \to \overline{D}^0 K^+$

Figure 2.2: Dominant Feynman diagrams of the B meson decays $B^+ \to \overline{D}^0 \pi^+$ (left) and $B^+ \to \overline{D}^0 K^+$ (right). [3]

The modern description of quark mixing was preceded by the Cabibbo mechanism. The Cabibbo mechanism originated from the Cabibbo hypothesis. When the decay rates of the decays $K^-(s\overline{u}) \to \mu^-\overline{\nu}_\mu$ and $\pi^-(d\overline{u}) \to \mu^-\overline{\nu}_\mu$ depicted in Figure 2.3 were measured under the assumption of a universal weak coupling to quarks the decay rate of the K^- decay was smaller than what was expected by a factor 20. To save the universality of the weak coupling strength Cabibbo made the hypothesis that while the weak coupling strength is universal the weak eigenstates of quarks are different from their mass eigenstates. These eigenstates are connected by the relation given in Equation 2.1, where θ_c is the

Cabibbo angle and ′ denotes the weak eigenstates. At the time the Cabibbo hypothesis was proposed only the u, d and s quarks were known.[1]

$$
\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}
\tag{2.1}
$$

The Cabibbo matrix, which relates the two eigenstates solves the issue by introducing the factors $\cos\theta_c$ and $\sin\theta_c$ at the interaction vertices in Figure 2.3. [1]
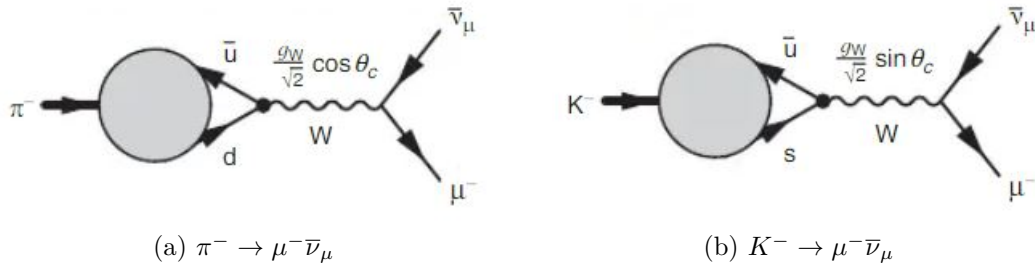

(a) $\pi^- \to \mu^- \overline{\nu}_\mu$

(b) $K^- \to \mu^- \overline{\nu}_\mu$

Figure 2.3: Expected dominant decay modes for $\pi^- \to \mu^- \overline{\nu}_\mu$ (left) and $K^-(s\overline{u}) \to \mu^- \overline{\nu}_\mu$ (right). [1]

The Cabibbo matrix was extended to three generation of quarks leading to the mixing of quarks in the SM to be described by the unitary, 3×3 Cabibbo-Kobayashi-Maskawa (CKM) Matrix. Like the Cabibbo matrix the CKM matrix therefore relates the weak eigenstates of the quarks to their mass eigenstates. This relation is described by Equation 2.2.[1]

$$
\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad V_{CKM}V_{CKM}^\dagger = 1, \quad V_{CKM}^\dagger V_{CKM} = 1
\tag{2.2}
$$

With $(d', s', b')^T$ being the weak eigenstates and $(d, s, b)^T$ the mass (flavour) eigenstates. The probability of a particular quark flavour change is proportional to the squared matrix element $|V_{ij}|^2$. E.g. the probability of a bottom quark changing its flavour to a charmed quark is proportional to $|V_{cb}|^2$. The CKM matrix can be parameterised using three mixing angles θ_{12}, θ_{23}, θ_{13} and the complex, CP-violating phase δ, with $s_{ij} = sin\theta_{ij}$, $c_{ij} = cos\theta_{ij}$ and $i, j = 1,2,3$, which results in Equation 2.3.[[1], [3], [5], [7]]

$$
V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}
\tag{2.3}
$$

The diagonal elements of the CKM matrix are in the magnitude order of unity and have the highest magnitude. They correspond to flavour changes

within the same generation of quarks ud, cs and tb. Therefore transitions within each generation are more likely to occur than transitions between different generations. The off diagonal elements are relatively small, therefore the transitions between different generations is suppressed. To illustrate this property the approximate values of the CKM matrix elements $|V_{ij}|$ are shown in Equation 2.4.[[1], [3]]

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix} \tag{2.4}$$

This hierarchy can be shown by using the Wolfenstein parameterisation in Equation 2.6, which uses the four real parameters λ, A, ρ and η defined in Equation 2.5. With the parameter $\lambda \approx |V_{cd}| = |V_{us}|$ the order of magnitude of a given matrix element is indicated. [[3], [5], [7]]

$$\lambda = s_{12}, \quad A\lambda^2 = s_{23}, \quad A\lambda^3(\rho - i\eta) = s_{13}e^{-i\delta} \tag{2.5}$$

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{2.6}$$

The conditions $V_{CKM}V_{CKM}^{\dagger} = 1$ and $V_{CKM}^{\dagger}V_{CKM} = 1$ of the unitary CKM matrix lead to different complex, independent relations that constraint the possible values of different elements of the matrix. Six of these constraints can be expressed as unitary triangles in the $\rho - \eta$ plane. E.g. the relation in Equation 2.7 leads to the B meson triangle. [[3], [5], [7], [9]]

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.7}$$

The unitary triangle is formed by dividing each side of the equation by $V_{cd}V_{cb}^*$ and defining the new parameters in Equation 2.8 using $\overline{\rho}$ and $\overline{\eta}$ related to ρ and η by Equation 2.9. The new parameters are equal to the Wolfenstein parameters up to terms of $\mathcal{O}(\lambda^4)$. The resulting triangle is depicted in Figure 2.4a. The unitary triangle contains three finite angles ϕ_1, ϕ_2, ϕ_3 and all unitary triangles share the same non zero area.[[5], [7], [9]]

$$\overline{\rho} + i\overline{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \quad 1 - (\overline{\rho} + i\overline{\eta}) = -\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \tag{2.8}$$

$$\rho + i\eta = \frac{\sqrt{1 - A^2\lambda^4}(\overline{\rho} + i\overline{\eta})}{\sqrt{1 - \lambda^2}[1 - A^2\lambda^4(\overline{\rho} + i\overline{\eta})]} \tag{2.9}$$

The area of the unitary triangles can be written as $J_{CP}/2$ using the Jarlskog parameter $J_{CP} \propto \sin(\delta)$. The sides and angles of the unitary triangle are

constrained by processes (e.g. B meson decays), which are used to measure the CKM matrix elements, as depicted in Figure 2.4b. The different constraints need to be consistent with each other and overlap at some point to lead to a closed unitary triangle. If the triangle closes the unitary condition is fulfilled and the prediction of the CP violation of the CKM matrix and the SM is correct.[[5], [7], [9]]
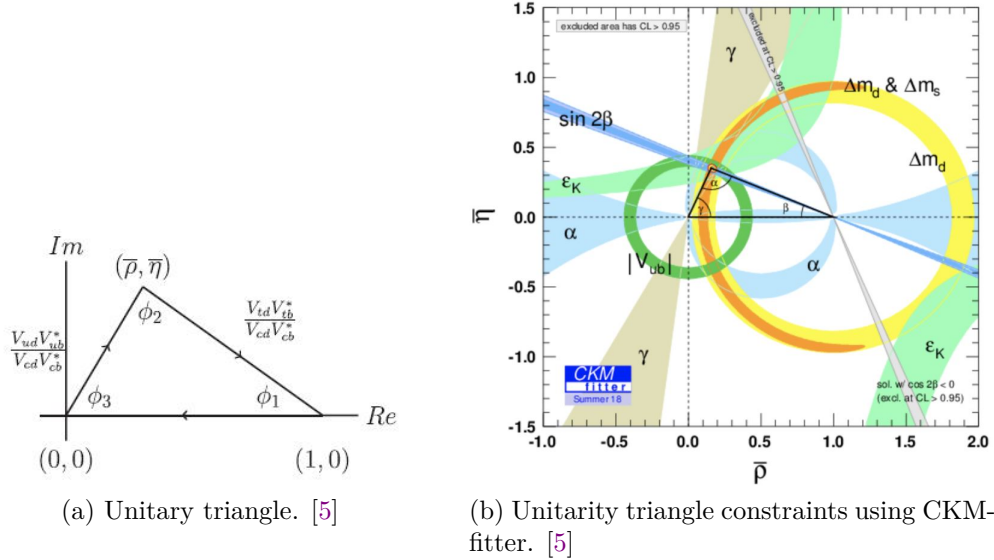
(a) Unitary triangle. [5]

(b) Unitarity triangle constraints using CKM-fitter. [5]

Figure 2.4: Unitarity triangle of the B meson without the constraint indications (left) and with constraint indications (right).

2.4 B Mesons

B mesons come in a variety of excited states, but the four most common B mesons are the ground states: B^0, B^+, B_s^0 and B_c^+. B mesons can be produced using Υ mesons, which are constituted of $b\bar{b}$ quarks and have a mass of ~ 9.46 GeV in their ground state. The B mesons are relatively massive mesons with a mass of ~ 5.3 GeV for B^0, B^+, B_s^0 and ~ 6.3 GeV for B_c^+. Each of the B mesons contains a $\bar{b}$ antiquark that is combined with either a c, u, d or s quark. The most common B meson quark combinations and their respective masses are listed in Table 2.3. [[1], [9], [10]]

Type	Quarks	Mass (GeV)
B^0	$d\bar{b}$	5.279
B$^+$	$u\bar{b}$	5.279
B$_s^0$	$s\bar{b}$	5.366
B$_c^+$	$c\bar{b}$	6.275

Table 2.3: B meson quark combinations and masses. [[1], [10], [11]]

For B^0 mesons $B^0(d\bar{b}) \longleftrightarrow \overline{B}^0(\bar{d}b)$ mixing occurs. It can be seen in the two dominant box diagrams of this mixing process in Figure 2.5 that this oscillation

is sensitive to the CKM matrix element $|V_{td}|$ and therefore the mixing frequency of the $B^0(d\bar{b}) \longleftrightarrow \overline{B}^0(\bar{d}b)$ oscillations can be used to determine the element. [1]
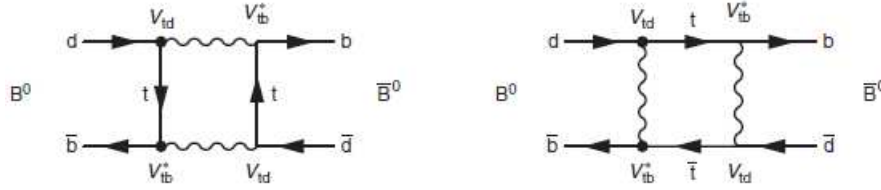

Figure 2.5: Dominant B^0 mixing box diagrams. [1]

2.5 Semileptonic Decays

Semileptonic decays of B^+ and B^0 mesons with charged, low mass leptons $\ell = e^\pm, \mu^\pm$ are expected to be free of non-Standard Model contributions, because they occur due to first-order weak interactions. The semileptonic decays involve either the electroweak transition $b \to c\ell\nu_\ell$ or $b \to u\ell\nu_\ell$ with the W boson decaying into the $\ell\nu_\ell$ pair (illustrated in Figure 2.7). [[3], [5]]

Theoretically one could also use purely leptonic decays like $B \to \ell\nu_\ell$ to measure the CKM matrix elements V_{cb} and V_{ub}. However compared to semileptonic decays the branching fractions of purely leptonic decays are very small, e.g. the branching ratio of $B \to \overline{D}\ell\nu_\ell$ is measured to be $(2.35 \pm 0.09)\%$ while for $B \to \mu\nu_\mu$ a measurement estimated $(6.46 \pm 2.22) \times 10^{-7}$ with a significance of only $2.4\,\sigma$ using the whole data set of Belle I. Therefore much more data would be required for purely leptonic measurements. [[11], [12]]

Furthermore hadronic decays, i.e. decays involving only hadrons like $B \to D\pi$ could be used, but while these decays have sufficiently high branching fractions the interaction of the two hadronic currents leads to very high uncertainties. Therefore with branching fractions being high enough and uncertainties of only one hadronic current being under control semileptonic decays are the best option for decays to be used for the determination of the CKM matrix elements V_{cb} and V_{ub}. [[3], [7]]

Another feature of semileptonic decays like $B \to D^*\ell\nu_\ell$ is the missing energy during its detection due to the decays having atleast one neutrino as a final state particle, which can not be detected. The Feynman Diagram of $B \to D^*\ell\nu_\ell$ is depicted in Figure 2.6. While in semileptonic decays via low mass leptons the new physics are suppressed the decays involving τ leptons are sensitive to new physics contributions like charged Higgs bosons. [[3], [5]]
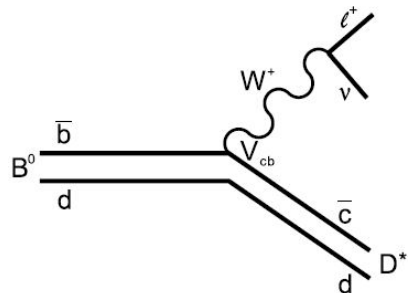

Figure 2.6: Feynman Diagram of the decay channel $B^0 \to D^{*-}\ell^+\nu_\ell$. [13]

The matrix elements can be determined by either studying the decay rates of exclusive decays, i.e. decays like $B \to D^* \ell \nu_\ell$, where a specific final state hadron like D* is looked for or inclusive decays, i.e. $B \to X \ell \nu_\ell$, with X being all possible final state hadrons containing either a c or u quark, depending on the given matrix element (see Figure 2.7). [[3], [5]]

The hadronic currents of the semileptonic decays are described using form factors. The decay rates are dependent on the momentum transfer during the decay, which is characterized by the form factor. To calculate the CKM matrix elements from these decay rates different theoretical models are used to determine the form factors for inclusive and exclusive measurements. However these models include parameters that can not be calculated and therefore need to be measured. [[3], [5]]

In theory both types of measurements should lead to the same values for $|V_{cb}|$ and $|V_{ub}|$. However at the moment there is a discrepancy of about 3 σ and 4 σ for $|V_{cb}|$ and $|V_{ub}|$, respectively between the inclusive measurement and the exclusive measurement. Current measurements for $|V_{cb}|$ and $|V_{ub}|$ are listed in Table 2.4. The observed discrepancy could for example hint to new physics processes, flaws in old detectors or flaws in theoretical descriptions of the QCD contributions. New measurements using more and newer data will therefore be valuable for understanding this discrepancy. [[3], [5]]
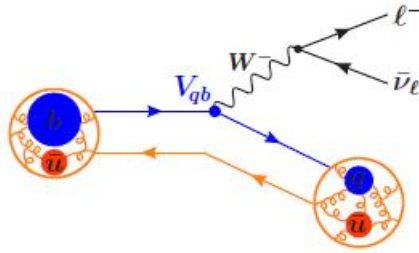

Figure 2.7: Feynman Diagram inclusive semileptonic $B \to X \ell^- \overline{\nu}_\ell$ decays. [3]

	$|V_{cb}|$ (10^{-3})	$|V_{ub}|$ (10^{-3})
exclusive	$38.76 \pm 0.42 \pm 0.55$	3.67 ± 0.15
inclusive	42.19 ± 0.78	$4.32 \pm 0.12 \pm 0.13$

Table 2.4: Comparison of exclusive and inclusive $|V_{cb}|$ and $|V_{ub}|$ measurements. [4]

3 Belle II and SuperKEKB

3.1 SuperKEKB

Like its predecessor KEKB, which was in operation from 1999 to 2010 for the Belle I experiment, the particle accelerator SuperKEKB used for the Belle II experiment is an asymmetric energy electron positron double-ring collider located in Tsukuba, Japan at the Japanese High Energy Accelerator Research Organisation (KEK). The Belle II experiment started taking data in 2018. [[5], [14]]

A schematic of SuperKEKB is depicted in Figure 3.1. The electron positron collider consists of the following main components: [14]

- The 7 GeV electron ring, i.e. the high-energy ring (HER).

- The 4 GeV positron ring, i.e. the low-energy ring (LER).

- An injector linear accelerator (linac)

- A positron damping ring (DR) to reduce the positron beam emittance
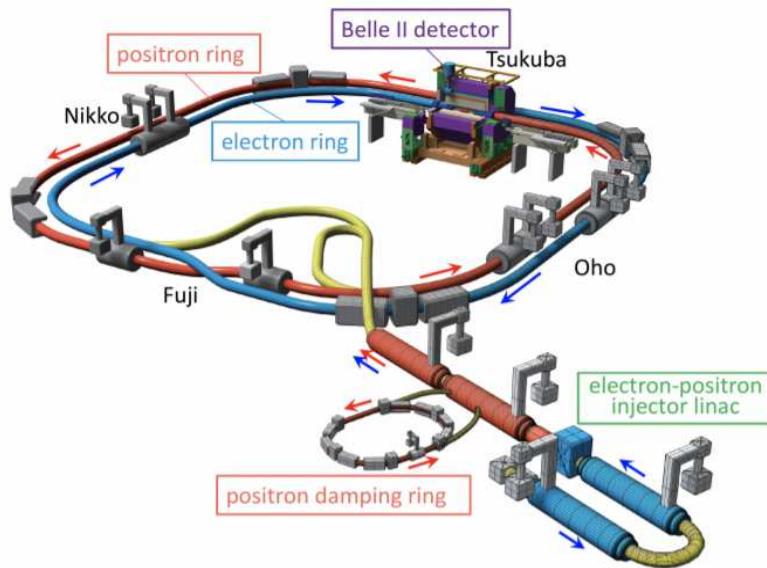

Figure 3.1: SuperKEKB schematic. [14]

An important parameter to characterise the performance of particle accelerators is their instantaneous luminosity $\mathcal{L}$, which is defined by Equation 3.1, where σ (in units of cm^{-2}) is the production cross section of a given interaction (event) and the number of events per second dN/dt. [15]

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

Leading to units of cm^{-2}s^{-1} for the luminosity. Furthermore the luminosity can be written as the relation in Equation 3.2 for particle collisions. Where N_1

and N_2 are the numbers of particles in a bunch for each beam, N_b the number of bunches, f the collision frequency and σ_x and σ_y are the standard deviations of the particle density, which is assumed to be gaussian. Therefore the luminosity is proportional to the beam currents and the beam's cross section area. The cross section area can be written using β functions, where β_y^* is the vertical β function at the interaction point (IP). The high luminosity of Belle II can be achieved due to higher currents and by squeezing the beam.[[15], [16]]

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi \sigma_x \sigma_y} \Rightarrow \mathcal{L} \propto \frac{I_{e^-} I_{e^+}}{\beta_y^*} \qquad (3.2)$$

The luminosity can be integrated over a given time period, which results in the integrated luminosity $\mathcal{L}_{int}$ in Equation 3.3. [15]

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t')dt' \qquad (3.3)$$

The integrated luminosity is often given in units of barn (b), where 1 b = 10^{-24} cm^2, and is directly related to the number of events N that occurred during a given time period. The number of events of various physics processes, which occur at Belle II can be approximated using Equation 3.4 and the cross sections in Table 3.1. The cross sections give a measure of probability that a given event occurs in a collision of two particles. [15]

$$\mathcal{L}_{int} \times \sigma = N \qquad (3.4)$$

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

Table 3.1: Cross section estimations of events occurring in e^-e^+ collisions at $\sqrt{s} = 10.58$ GeV. [5]

SuperKEKB is designed to collect data at the $\Upsilon(4S)$ resonance at a center of mass energy $\sqrt{s}$ of 10.58 GeV with a design luminosity of 8×10^{35}cm^{-2}s^{-1},

which is 40 times higher than the record luminosity achieved by KEKB. The high luminosity will be achieved by using a nano-beam scheme. The goal of SuperKEKB is to achieve an integrated luminosity of 50 ab^{-1} until ∼2030 as indicated in the roadmap plot in Figure 3.2. Its predecessor KEKB achieved an integrated luminosity of ∼710 fb^{-1} in its eleven years of operation.[[5], [11], [14]]
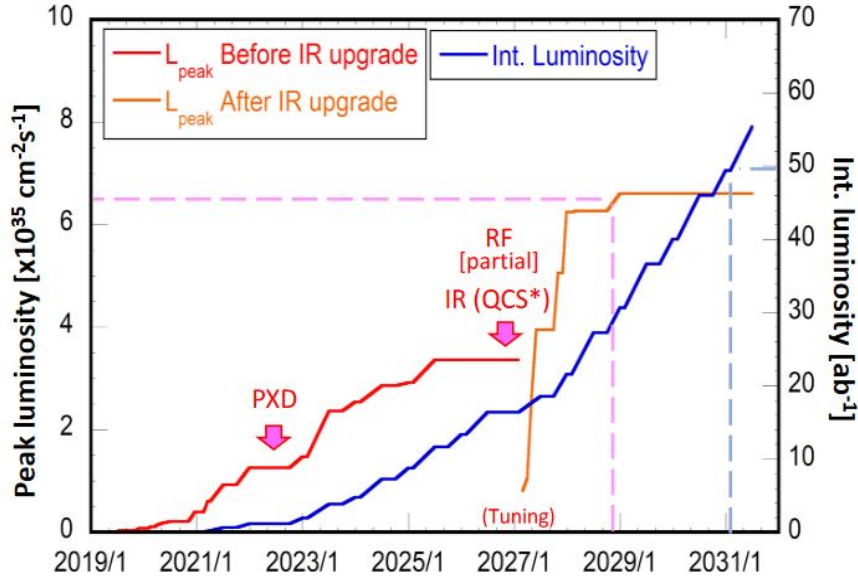

Figure 3.2: SuperKEKB roadmap. [6]

The $\Upsilon(4S)$ meson has a very high branching ratio of over 96% for the decay $\Upsilon(4S) \rightarrow B\overline{B}$, which classifies SuperKEKB as a B factory. The production process of the $B\overline{B}$ pair is shown in the Feynman diagram in Figure 3.3. The predominant production of $\Upsilon(4S)$ mesons is achieved by colliding electrons and positrons with asymmetric energies of 7 GeV and 4 GeV, respectively to achieve the targeted $\Upsilon(4S)$ resonance at a center of mass energy of 10.58 GeV. The energy of the resonance is just above the threshold for $B\overline{B}$ pair production, therefore the resulting boost from the energy asymmetry is needed to be able to differentiate between the decay products of the produced $B\overline{B}$ pair and to be able to measure time dependent CP violation since the B mesons would otherwise be produced at rest. [[5], [11], [14]]
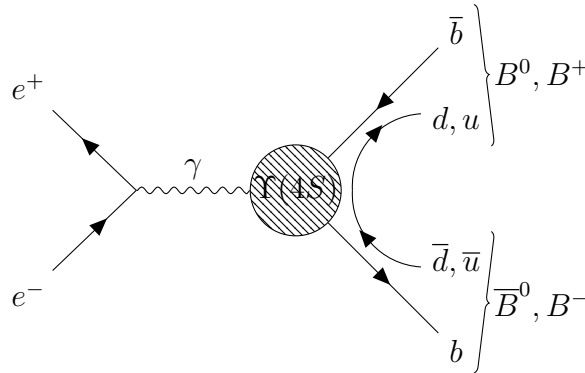

Figure 3.3: $\Upsilon(4S) \rightarrow B\overline{B}$ Feynman diagram. Adapted from [17].

3.2 Beam-induced Background

There are five main sources of beam-induced background at SuperKEKB: [[5], [16]]

- Touschek scattering

- Beam-Gas scattering

- Synchrotron radiation

- Radiative Bhabha scattering

- Two photon processes

Touschek scattering

Touschek scattering is an intra bunch scattering process that occurs inside the storage ring. This effect occurs due to Coulomb scattering of particles inside the same beam bunch, therefore the energy and momenta of the particles that participate in the scattering process is deviated from the nominal energy and momentum. This results in a particle with higher energy and a particle with lower energy. This leads to particles leaving the intended trajectory and hitting the inner wall of the beam pipe, which produces shower particles, which will subsequently be detected. [[5], [16], [18]]

The loss rate occurring due to the Touschek effect can be described using Bruck's formula described in [18]. The loss rate is proportional to the (beam size)$^{-1}$ and the (beam energy)$^{-3}$. Since SuperKEKB uses a nano-beam scheme the background resulting from the Touschek effect will be increased and is expected to be a major background source. The relation of the loss rate to the energy decreases the amount of Touschek scattering for the HER, but increases the Touschek effect for the LER. Therefore the amount of Touschek background is expected to be much higher for positrons than electrons. [[5], [16], [18]]

Beam-Gas scattering

Beam-Gas scattering can occur due to residual gas molecules in the beam pipe, which can happen either via Coulomb scattering or bremsstrahlung scattering. Similar to the Touschek scattering, these scattering processes change the energy and momentum of the scattered particles leading to the production of particle showers, which can be detected by the detector. [[5], [16], [18]]

Synchrotron radiation

Synchrotron radiation is electromagnetic radiation, which is emitted by a charged particle when it gets accelerated due to a magnetic field. This process can occur inside the detector due to the magnets of the final-focusing system. The beam emits photons due to this effect, which in turn can hit the detector leading to background. [[5], [16], [18]]

Radiative Bhabha scattering

Bhabha scattering is a $e^-e^+ \to e^-e^+$ scattering process. During this electron positron scattering process a photon can be produced, which is the so called radiative Bhabha scattering $e^-e^+ \to e^-e^+\gamma$. The Feynman diagram of this process is shown in Figure 3.4b. The rate of radiative Bhabha scattering processes is proportional to the luminosity of the accelerator and these events have two consequences: [[5], [16]]

- Photons get emitted along the direction of the beam axis and interact with the iron of the magnets. During these interactions a large amount of neutrons gets produced due to the photo-nuclear resonance mechanism ([19]). The neutrons produced in this interaction propagate to the K_L and muon detector.[[5], [16]]

- Due to the emitted photon the energy of the electron and positron decreases. Therefore they leave the intended trajectory and the particles create showers by hitting the beam pipe walls.[[5], [16]]

Two photon processes

During an electron positron collision the processes $e^+e^- \to e^+e^-\gamma^*\gamma^* \to e^+e^-e^+e^-$ and $e^+e^- \to e^+e^-\gamma^*\gamma^* \to e^+e^-\mu^+\mu^-$ can occur. During this processes two virtual (here denoted by *) photons are produced in the collision that in turn produce either a low momentum electron positron pair or a low momentum muon pair as shown in the Feynman diagram in Figure 3.4a. These lepton pairs can spiral in the magnetic field on the inside of the detector leading to multiple false hits in the inner detector. [[5], [16]]
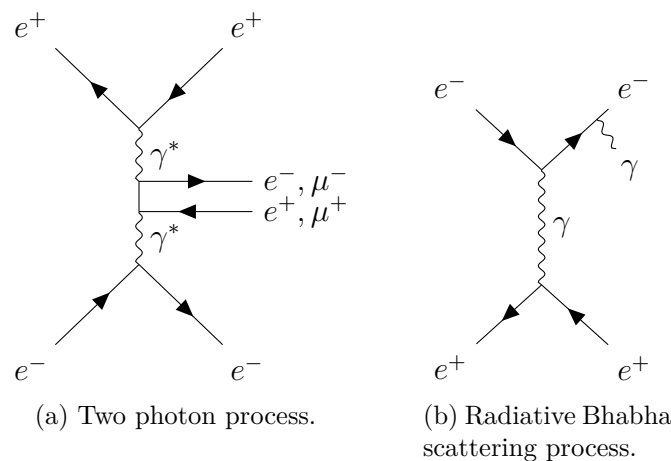

(a) Two photon process. (b) Radiative Bhabha scattering process.

Figure 3.4: Radiative Bhabha scattering(right) and two photon process(left) Feynman diagrams. Adapted from [20].

3.3 Belle II Detector

The various positions of the processes occurring in the detector are conventionally described either by a cartesian coordinate system or by polar coordinates.

The origin of the cartesian coordinate system is set to the interaction point (IP) with the z axis pointing in the direction of flight of the electron beam, the y axis pointing to the top and the x axis pointing outwards. The polar coordinates are set up with r being the distance from the IP, the azimuthal angle ϕ in the x-y plane and the polar angle θ. The detector covers a θ region of 17°-150°. The long lived charged particles, which can be detected in particle detectors are: $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, $p^{\pm}$ and the neutral particles: γ, K_L^0, n, d. A 3D model of the Belle II detector is depicted in Figure 3.5. The detector consists of five main components: [[5], [16]]

- Vertex Detector (VXD) consisting of a Silicon Pixel Detector (PXD) and a Silicon Vertex Detector (SVD)

- Central Drift Chamber (CDC)

- Particle Identification Chamber consisting of Time of propagation (TOP) counters and an Aerogel ring imaging detector (ARICH)

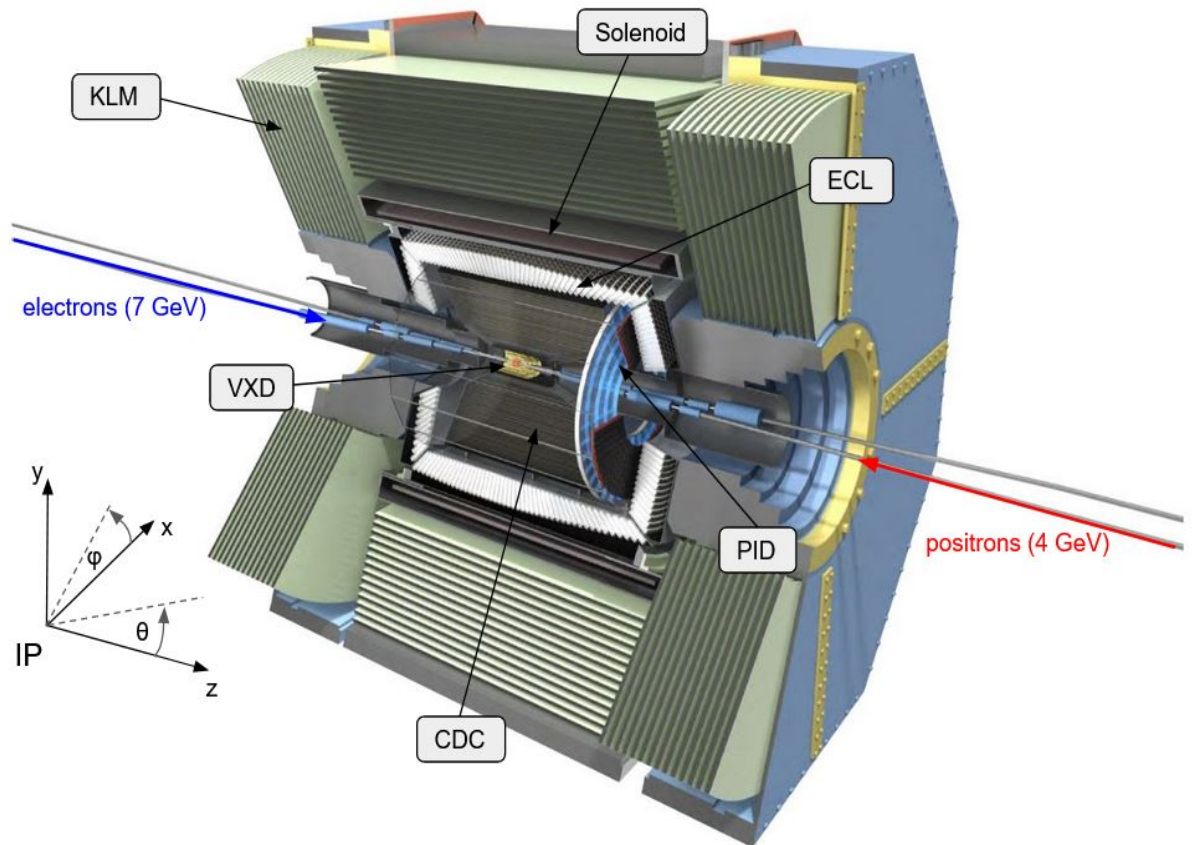- Electromagnetic Calorimeter (ECL)

- K_L and Muon Detector (KLM)

Figure 3.5: 3D model of the entire Belle II detector and coordinate system. Adapted from [[21], [22]].

3.3.1 Vertex Detector (VXD)

The vertex detector is the inner part of the tracking system. The purpose of the VXD is to measure the vertices of the particles with a short life time by extrapolating the tracks of the final state particles resulting from the decays to the interaction point (IP) and also to measure the momenta of the final state particles. The transverse momenta can be measured by using the trajectory of the helix movement of particles resulting from the Lorentz force due to the magnetic field inside the detector given by the relation in Equation 3.5, where R is the radius of the curvature, p_T the transverse momentum of the particle and B the magnetic flux density. [[1], [5], [16]]

$$p_T = BR \tag{3.5}$$

The vertex detector consists of two components: [[5], [16]]

- Silicon Pixel Detector (PXD)

- Silicon Vertex Detector (SVD)

The two sub-detectors consist of overall six layers around the 10mm radius Be beam pipe, depicted in the schematics in Figure 3.6. The PXD occupies the two innermost layers at the radii of $r = 14$ mm and $r = 22$ mm. The SVD consists of the four outer layers at the radii of $r = 39$ mm, $r = 80$ mm, $r = 104$ mm and $r = 135$ mm. [[5], [16]]
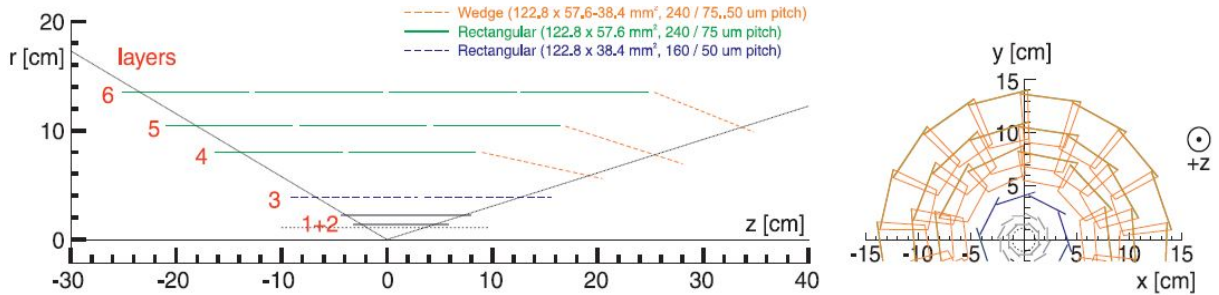

Figure 3.6: Schematic front view of the Belle II VXD consisting of six layers, two PXD layers and four SVD layers. [5]

Silicon Pixel Detector (PXD)

The PXD shown in Figure 3.8 is the innermost detector component of Belle II. It is used to measure the vertices of the particles with a short life time and the momentum of the final state particles. The PXD is a silicon detector, which is based on the Depleted P-channel Field Effect Transistors (DEPFET). [[16], [23]]

Whenever a charged particle passes through the sensor electron-hole pairs are created in the depleted silicon bulk. The electrons drift into the internal gate, while the holes drift to a p-doped backside, as illustrated in Figure 3.7. The charge in the internal gate then modulates the channel current and can be

measured. [[16], [24]]

The PXD consists of two layers at the radii of $r = 14$ mm and $r = 22$ mm currently containing 8 and 2 (the goal is 12) ladders, respectively. In the inner layer (L1) the ladders have an active area of 44.8 mm × 12.5 mm and in the outer layer (L2) they have an active area of 61.4 mm × 12.5 mm. Each ladder contains 768 × 250 pixels with a size of approximately 50 μm × (55 - 85) μm, with the pixels in the outer layer being slightly longer. [[16], [24]]
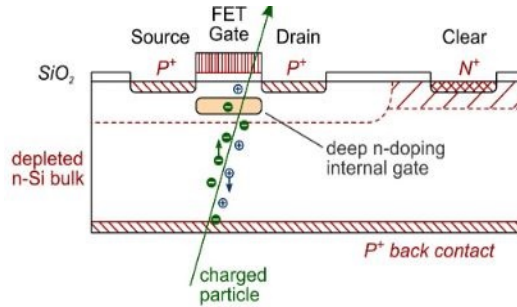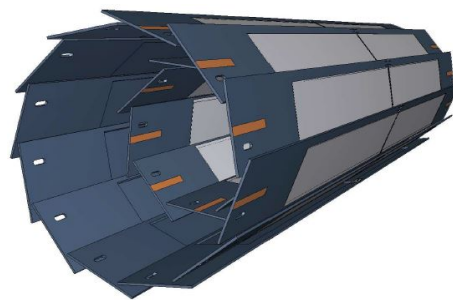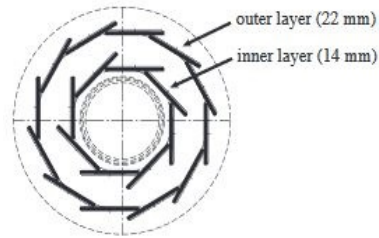

Figure 3.7: Production and collection of charges in a DEPFET silicon sensor. [24]

(a) Schematic of the PXD. [16]

(b) Frontal view of the ladder arrangements. [23]

Figure 3.8: 3D render (left) and front-view illustration (right) of the Belle II PXD consisting of two layers containing 8 in the inner layer and 12 ladders outer layer.

Silicon Vertex Detector (SVD)

The SVD is the outer sub-detector of the VXD and serves the same purpose as the PXD: to measure vertices of short lived particles and momenta of final state particles. The SVD consists of four layers of double sided silicon strip detectors (DSSD) of which the radii of their location, number of ladders and sensors are summarized in Table 3.2. Figure 3.9a and the left Figure in 3.6 show the cross section of the layout. Each line in the left Figure of 3.6 corresponds to a sensor of a given ladder in the corresponding layer. [[5], [16], [25]]

Layer	Radius [mm]	Number of Ladders	Number of Sensors per Ladder
3	39	7	2
4	80	10	3
5	104	12	4
6	135	16	5

Table 3.2: SVD layout summary. [25]

(a) Ladder layout of the SVD
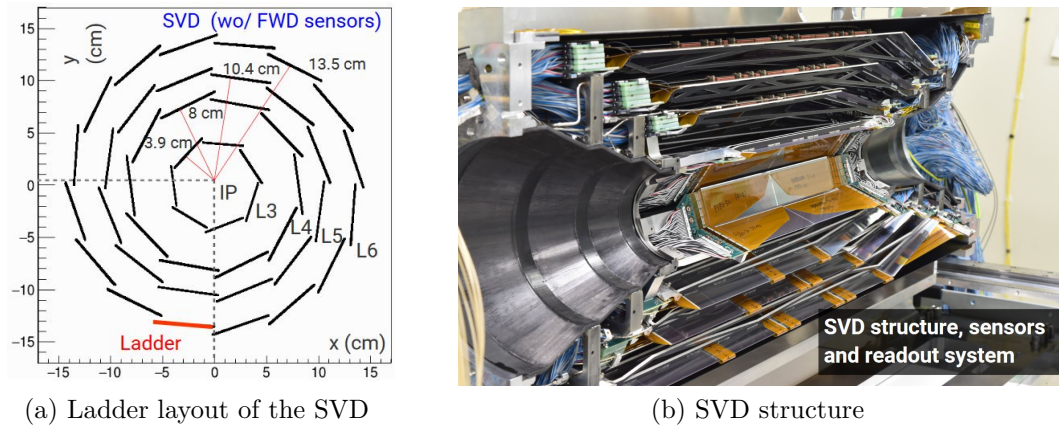
(b) SVD structure

Figure 3.9: Illustrations of the Belle II SVD showing its layout and a real picture of the SVD showing its structure. [25]

The sensors of the SVD depicted in Figure 3.10a are made out of a silicon bulk with p-doted strips on one side and n-doted strips on the opposite side. As a charged particle traverses through the silicon electron hole pairs are created, which drift to the n and p strips, respectively, thus creating measurable currents. The sensor plates are tilted (as seen in Figure 3.10b) to compensate that the deflections of the electrons due to Lorentz force are about three times larger than the deflections of the holes. [[5], [16], [25]]
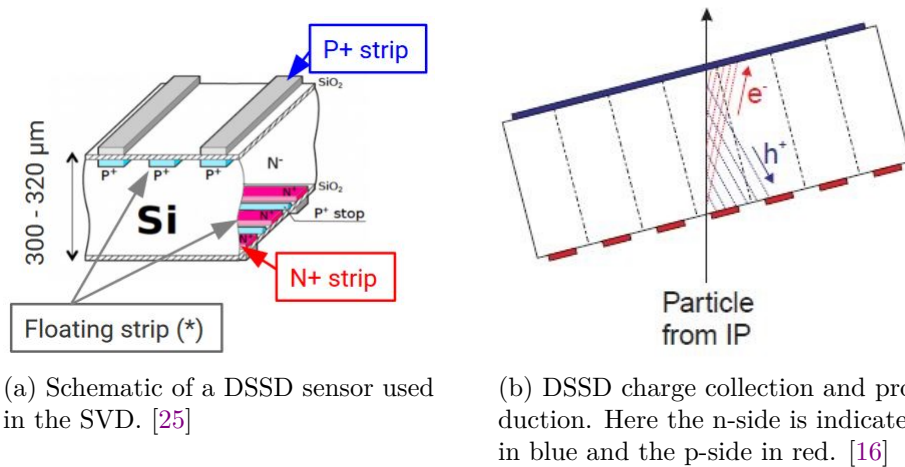

(a) Schematic of a DSSD sensor used in the SVD. [25]

(b) DSSD charge collection and production. Here the n-side is indicated in blue and the p-side in red. [16]

Figure 3.10: Illustrations of the DSSD sensors used in the Belle II SVD.

3.3.2 Central Drift Chamber (CDC)

The CDC is a wire chamber detector with a cylindrical shape, where the inner cylinder has a radius of 160 mm and the outer cylinder a radius of 1130 mm. It covers a polar angle range of $17° < \theta < 150°$ and the layout is divided into 9 super layers, which are further divided into 56 layers containing overall 14336 sense wires (160 - 384 wires per layer). The layout is illustrated in Figure 3.11. The chamber of the CDC is filled with a gas mixture of 50% He and 50% C_2H_6. The CDC's main purpose is to act as the main tracking detector of Belle II, but overall serves three purposes: [[5], [16], [26]]

- Reconstruction of the tracks of charged particles and measuring their momenta

- Obtaining particle identification (PID) information by measuring the energy loss dE/dx of the particle within the gas volume of the CDC

- Providing trigger signals for charged particles. Trigger signals determine whether data is kept or not. More information on the trigger system is given in Section 3.4.
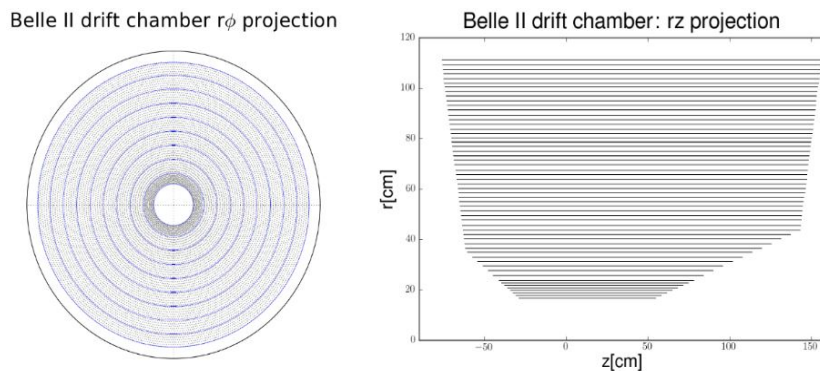

Figure 3.11: Illustrations of the layout of the Belle II CDC showing a rϕ projection (left) and a rz projection(right) [26]

The working principle of the CDC is the following: When a charged particle traverses through the gas inside the chamber the gas molecules get ionised. The electrons freed during the ionisation process ionise more gas molecules in the vicinity of the sense wire, which creates an electron avalanche.[[5], [16], [26], [22]]

To shape the electric field in an uniform manner each of the 14336 sense wires is surrounded by 8 field wires (overall 42240 field wires) resulting in the cells depicted in Figure 3.12a. A high voltage is applied between the sense wire and its 8 field wires, which creates an electric field and when the electron avalanche reaches the sense wire a hit is registered. The timing of the hits is proportional to the distance of the track of the particle, because the drift velocity of the electrons is almost constant due to the chosen wire configuration and gas.[[5], [16], [26], [22]]

The super layers of the CDC alternate between the axial wire super layers

and the stereo wire layers shown in Figure 3.12b. Information about the transverse plane coordinates is provided by the axial wires parallel to the z axis. The longitudinal coordinates are provided by the stereo wires. The currents measured in each layer can be summed up and the distances between layers approximated using straight lines. This is used to measure the overall strength of the signal, which can used to determine the energy loss dE/dx of the traversing particle for PID purposes. [[5], [16], [26], [22]]
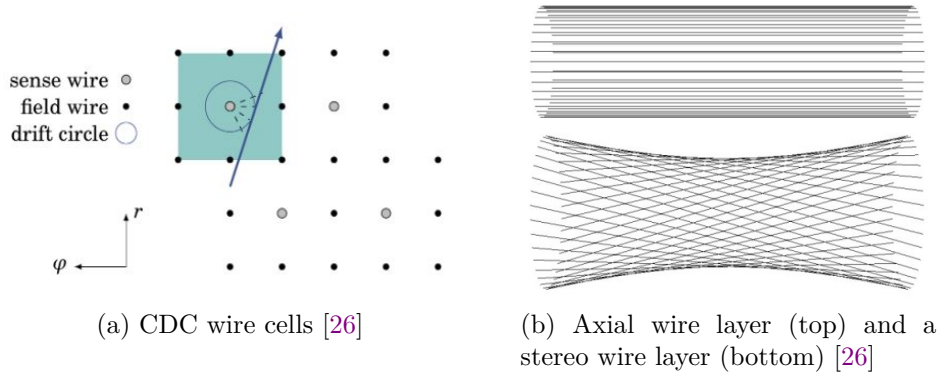
(a) CDC wire cells [26]

(b) Axial wire layer (top) and a stereo wire layer (bottom) [26]

Figure 3.12: CDC wire cells (left) and super layer wire layout types (right).

3.3.3 Particle Identification System (TOP and ARICH)

The Belle II particle identification system's main task is to separate protons (mass of 938 MeV), charged kaons (mass of 494 MeV) and charged pions (mass of 140 MeV) with a momentum of up to 4 GeV and contains two independent Cherenkov detectors, which cover different polar angel regions in the Belle II detector: [[5], [16], [27], [28]]

- 16 Time of propagation (TOP) counters covering the barrel region, i.e. $32° < \theta < 120°$

- An aerogel ring-imaging Cherenkov (ARICH) detector covering the forward end cap region, i.e. $17° < \theta < 35°$

Charged particles travelling through a dielectric medium will result in the molecules of the medium radiating photons (Cherenkov radiation) along its path if the particle travels with a velocity $v > c_n$, where v is the velocity of the particle and $c_n = c/n$ the speed of light in the given medium with a refractive index n. The photons are emitted at a fixed angle determined by the velocity of the particle and the refractive index of the medium. The relation between the emission angle θ_c of the photons, v and n is given in Equation 3.6, which results in the emission geometry in Figure 3.13. Detecting the Cherenkov radiation sets limits to the mass of the traversing charged particle, since the radiation is only emitted when the condition $v > c/n$ fulfilled. The angle of the photons can therefore be used to determine the velocity and mass of a particle. [[1], [28],

[29]]

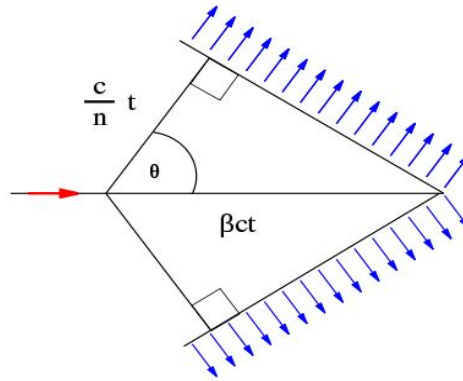$$\cos \theta_c = \frac{c'}{v} = \frac{1}{n\beta} \tag{3.6}$$

Figure 3.13: Cherenkov radiation geometry. [30]

TOP

Each of the 16 TOP counters is a quartz Cherenkov radiator bar. These counters surround the CDC and cover an angle of $32° < \theta < 120°$. Their working principle is depicted in Figure 3.14. The counters measure the time of propagation and the emission angles of the Cherenkov photons. The time needed for the photons to hit the photon detector at the end is proportional to the cherenkov emission angle θ_c enabling a measurement of θ_c. The photons that get emitted to the backwards region of the bar are reflected by a mirror at the back-end. The relation in Equation 3.6 can then be used to determine the velocity of the charged particle that traversed the crystal. This results in the possibility to differentiate kaons from pions due to their mass difference of $\sim$350 MeV, which results in a velocity and therefore timing difference. [[5], [16], [22], [27], [29]]
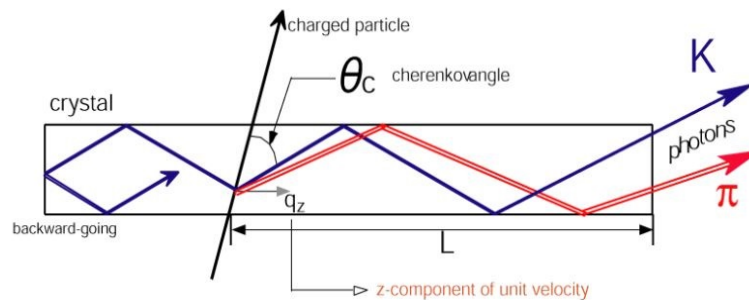

Figure 3.14: Side view of a TOP crystal. [28]

ARICH

ARICH covers the forward end cap region for a θ range of 17° to 35° and its purpose is to separate kaons from pions within a momentum range of 0.5 GeV to 4 GeV and additionaly to separate low momentum pions from muons within

a momentum range of 0.5 GeV to 1 GeV. ARICH consists of two stacked 2 cm aerogel layers with a slightly different refractive index n ($n_1 = 1.045$ and $n_2 = 1.055$) so that the radiated photons overlap to double the photon yield and the resolution, a 20 cm thick expansion volume and a photon detector consisting of 420 Hybrid Avalanche Photo diode sensors split into 7 rings. As a charged particle with sufficient velocity traverses through aerogel layers cherenkov radiation is produced, which forms a cone and the resulting photon ring is projected onto the photon detector. The radius of the resulting ring is proportional to the cherenkov angle and therefore the velocity of the charged particle. The detection process is illustrated in 3.15. [[5], [16], [27], [28], [31]]
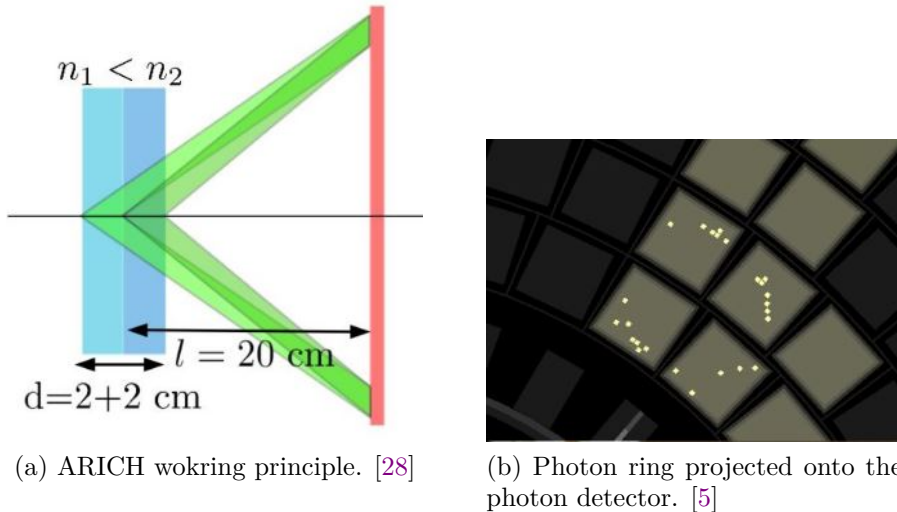

(a) ARICH wokring principle. [28]

(b) Photon ring projected onto the photon detector. [5]

Figure 3.15: Illustration of the working principle of ARICH (left) and a cherenkov photon ring hit on the photon detector (right).

3.3.4 Electromagnetic Calorimeter (ECL)

The first main purpose of the electromagnetic calorimeter is to detect photons with a wide energy range of 20 MeV to 4 GeV to identify neutral particles, which make up approximately one third of B decay products. The photons are the results of neutral particle decays like $\pi^0 \to \gamma\gamma$. The second main purpose of the ECL is to separate electrons from charged pions. Electrons are completely stopped in the ECL, while charged pions traverse further. The ECL contains 8736 crystals overall and is split into two sections: 6624 thallium-doped caesium iodide CsI(Tl) crystals around the barrel region and 2112 pure CsI crystals at the end caps. The calorimeter covers a polar angle range of $12.4° < \theta < 155.1°$. [[5] [16], [32], [33]]

When a photon reaches a crystal it produces an electron positron pair, which then emits bremsstrahlung due to their interaction with the medium producing further photons. This process continues and a shower of photons, electrons and positrons is created inside the crystal as depicted in Figure 3.16 until the energy of the particles reaches a threshold. The electromagnetic showers can be characterized by the radiation length X_0. After traversing the length X_0 inside a material the number of particles approximately doubles and the energy of an electron is reduced by $1/e$ due to bremsstrahlung emission inside a given

material. X_0 is proportional to Z^{-2}, therefore the radiation length is shorter for materials with a high atomic number.[[1], [5] [16], [22], [32]]

The location of the creation of the shower in the ECL can be compared to the tracks from previous detectors to determine whether the particle was neutral or charged, since only charged particles leave tracks in the CDC and VXD. Electrons deposit more energy in the ECL than charged pions, which only deposit a small amount of energy inside the ECL due to ionisation processes. The crystals of the ECL are scintillation crystals therefore the molecules of the crystal emit light after charged particles traversed the medium, which can be detected using photodiodes that are glued to the back of the crystal's surfaces. [[1], [5], [16], [32]]
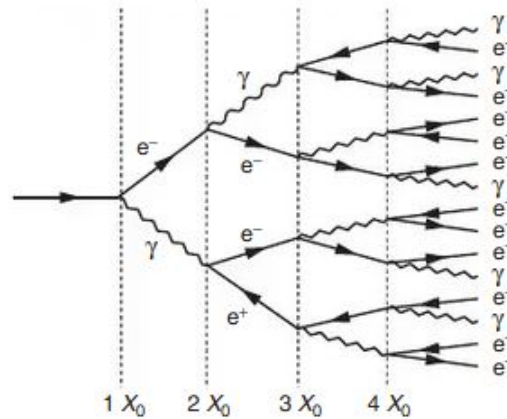

Figure 3.16: Electromagnetic shower development. [1]

3.3.5 K_L^0 Muon Detector (KLM)

The K_L^0 and muon Detector (KLM) detects K_L^0 mesons and muons (with a momentum of at least 0.6 GeV). The detector is located outside of the superconducting magnet and alternates between layers of 4.7 cm thick iron plates and active detector elements. The iron plates serve as return yokes for the magnetic field and as a material for the K_L^0 mesons to produce hadronic showers. The KLM consists of 15 detector layers and 14 iron plates in the barrel region and 14 iron plates and 14 detector layers in each of the end caps covering an overall angle range of $20° < \theta < 155°$. [[5], [16], [22], [34]]

While K_L^0 mesons create hadronic showers via interactions with the iron plates and muons with a momentum above 0.6 GeV traverse the KLM due to their high penetration power and leave hits. Muons additionally leave hits in the VXD and CDC detectors, while K_L^0 mesons do not. The hits in the KLM resulting from the muons can therefore be extrapolated to tracks from previous detector layers. If clusters are produced the location of the clusters left by K_L^0 mesons is connected to the IP. The K_L^0 mesons are then classified into two different classes depending whether an ECL hit on the way to the IP was found or not within 15° of the connection line: [[5], [16], [22], [34]]

- KLM only candidates, which require hits in at least two different detector layers

- KLM + ECL candidate, which require a hit in the ECL and at least one hit in the KLM

3.4   Trigger System

The purpose of the trigger system is to decide which data is kept during data taking. Due to the huge amount of data that gets produced by particle collisions it is impossible to keep all the data. This effect gets even more enhanced due to the high goal luminosity of $8 \times 10^{35} \text{cm}^{-2}\text{s}^{-1}$ and increased backgrounds at Belle II. The trigger system is required to have a high efficiency for $\Upsilon(4S) \rightarrow B\overline{B}$ and continuum events, high trigger rates of 30 kHz, a timing precision of less than 10 ns, a minimum two event separation of 200 ns and a total latency of the trigger system of ~ 5 μs to determine fast enough which data from detector signals occurs due to the relevant physics events and is kept and which data can be discarded. The latency is limited to ~ 5 μs due to the read out time of the SVD. The system consists of two trigger levels: [[5], [16], [35]]

- Hardware based low level trigger (L1) deciding, which data is kept during data taking

- Software based high level trigger (HLT) to further reduce data for specific analyses after the data was taken and saved

The L1 trigger consists of four sub triggers systems that send their sub trigger information to a Global Decision Logic (GDL), which checks if the event fulfills all conditions to be recorded. The L1 trigger system is depicted in Figure 3.17 and the sub trigger system consists of the following triggers: [[5], [16], [35]]

- CDC Trigger: The CDC trigger characterizes particles based on the tracks left by charged particles inside the CDC. The information given by the charged tracks include momentum, charge, position and multiplicity. [[16], [35]]

- ECL Trigger: The ECL trigger uses the energy deposit and energy cluster information obtained from the ECL providing a trigger for both charged and neutral particles. The deposit and cluster information are split up into a total energy trigger and into a cluster counting trigger. The ECL trigger can also be used to identify Bhabha and $\gamma\gamma$ events to monitor the luminosity. [[16], [35], [22]]

- BPID Trigger: The BPID (Barrel Particle Identification) trigger will use the information obtained from the TOP counters. The TOP counters have a good timing resolution of under a nanosecond and are therefore suitable to be used as a timing trigger for event timing. This trigger is not implemented yet. [[16], [35]]

- KLM Trigger: The KLM gives information about muon events and can be used to recognise $\mu^+\mu^-$ events, which are useful for detector calibrations. [[16], [35]]
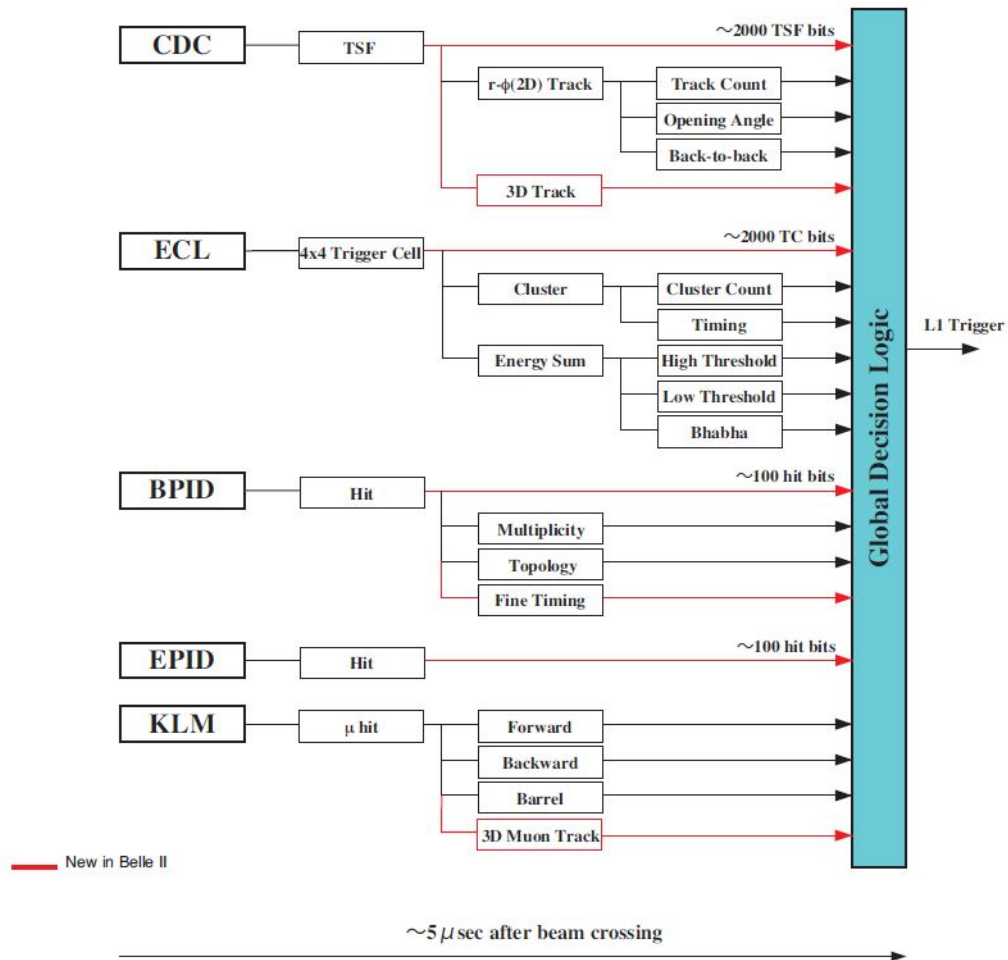
Figure 3.17: L1 trigger overview showing which information is used by the GDL to make its L1 trigger decision. The lines colored in red indicate, which trigger information was newly added in Belle II. [16]

3.5 Reconstruction

3.5.1 Track Reconstruction

The information obtained from the CDC and VXD is used to reconstruct the tracks of charged particles in two stages using a tracking software:[[5], [16]]

- Track finding: Building a track candidate from the detected hits

- Track fitting: Performing a fit to the hit positions of a candidate to obtain a trajectory

Track Finding

During track finding different software algorithms are used for each detector to build the tracks from hits. [5]

The CDC track finding is achieved by using two algorithms. First a global

26

track finder algorithm looks for tracks that originate from the IP. The wire hits of the CDC are combined to the Track Segments depicted in Figure 3.18 by a Track Segment Finder (TSF). By means of the information obtained from the axial layers and a Hough Transformation a two dimensional track is created from these segments. The found track can then be expanded to a three dimensional track by including the information of the stereo layers. [[5], [16], [22], [36]]

The second algorithm for the CDC is a local track finder algorithm aiming to detect tracks, which are displaced from the IP using a cellular automaton (CA). The information obtained by both algorithms is then combined. [[5], [16], [22], [36]]
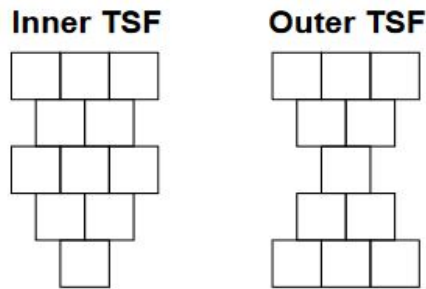

Figure 3.18: TSF segment shapes of the innermost super layer(left) and all other super layers(right) with the IP located downwards. Each wire cell of the CDC corresponds to a square in the segment with the signal wire in the middle of the square.[16]

The VXD track reconstruction is important for detecting charged, slow pions. They are called "slow" or "soft" pions due to their low momentum, i.e. a transverse momentum lower than 100 MeV. The charged, slow pions are forced to spiral inside the innermost detector layers due to the magnetic field and their low momentum. Therefore these pions don't reach the outer detector layers and leave multiple hits in the SVD sub detector of the VXD.[[5], [16], [36]]

Due to the SVD being located very near to the IP a lot of beam-induced background noise has to be filtered out. First the algorithm uses space points obtained by pre processing SVD measurements to create graphs of related space points. The second step is to create a sector map to obtain information how the different space points can be related to tracks. For this purpose each sensor is divided into sub sections as depicted in Figure 3.19. A set of paths is then produced by means of a CA of which the best path is chosen. [[5], [16], [36]]
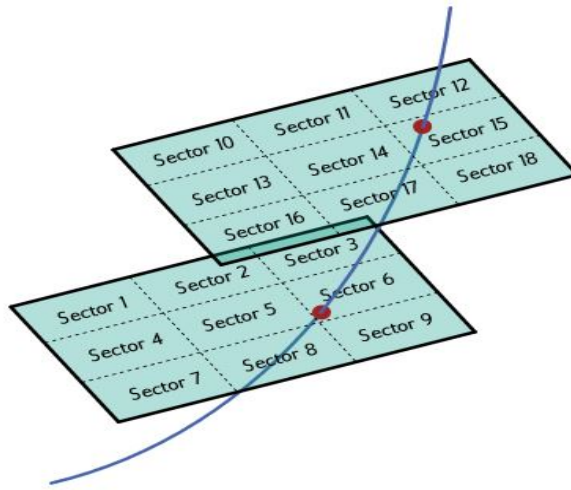
Figure 3.19: Illustration of SVD sectormaps of two sensors. [36]

Track Fitting

For track fitting the deterministic annealing filter (DAF) is used by implementing the GENFIT package [37]. The DAF is based on a Kalman filter, which is a least square method that takes the interactions with the materials into consideration. The five parameters used to describe a charged track are:[[5], [36], [37]]

- d_0: signed distance to the point of closest approach from the IP in the $r - \phi$ plane

- z_0: the signed distance of the point of closest approach from the IP on the z-axis

- ϕ_0: angle of the transverse momentum at the point of closest approach in the $r - \phi$ plane

- $\tan \lambda$: tangens of the angle between the momentum in the point of closest approach and the $r - \phi$ plane

- ω: the signed curvature of the track

3.5.2 Cluster Reconstruction

The energy deposits and their position obtained from the ECL are used to reconstruct clusters. Using software algorithms first the clusters need to be found and then clusters can be matched to tracks in previous detector layers. Based on the match it can be deduced if the clusters originiate from neutral or charged particles, due to the charged particles leaving tracks in the VXD and CDC. [[5], [16], [22]]

The clusters are reconstructed using cells of 4×4 crystals. Neighbouring cells are used to construct clusters containing 3×3 cells around the cell with the highest energy. The energy of the cluster is then determined by simply adding up the energies of all cells, while the timing and position of the clusters

is calculated using a energy weighted average over all cells. [[5], [16], [22]]

Tracks and clusters are then matched by extrapolating the tracks to the location of the ECL. If a cluster within a certain distance from the extrapolated track point is found the cluster and the track are combined. [[5], [16], [22]]

3.5.3 Particle Identification

Belle II needs to be able to identify five different particles with a long lifetime: π, K, μ, e and p. The particles are identified by combining the measurements obtained by different detectors. [5]

The ionization energy loss measurements dE/dx of the CDC and the SVD provides particularly good particle separation below a momentum of 1 GeV as depicted in the Figures 3.20 and 3.21. This separation is achieved by taking advantage of the Bethe-Bloch formula in Equation 3.7, which states that the measurement is proportional to the particle velocity β. Where z is the charge of the traversing particle, N the electron density of the material, m_e the mass of an electron, I the mean excitation energy, $\beta = v/c$, γ the Lorentz factor and e the electron charge. [[5], [26]]

$$\frac{dE}{dx} = \frac{4\pi N e^4}{m_e c^2 \beta^2} z^2 \left(\ln \frac{2 m_e c^2 \beta^2 \gamma^2}{I} - \beta^2 + corr \right) \tag{3.7}$$
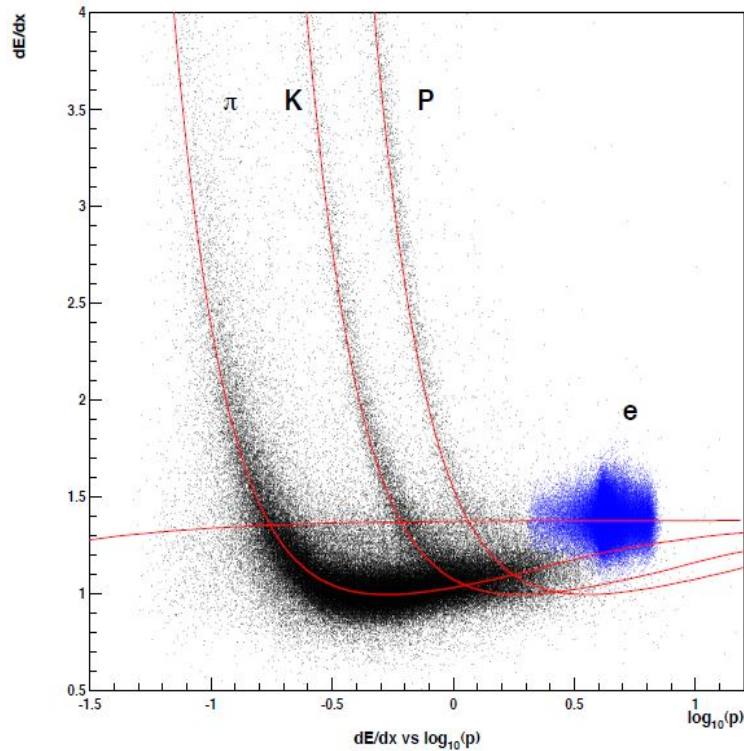

Figure 3.20: Plot showing the distinct particle separation by using dE/dx measurements from the CDC obtained from Belle I data. [7]
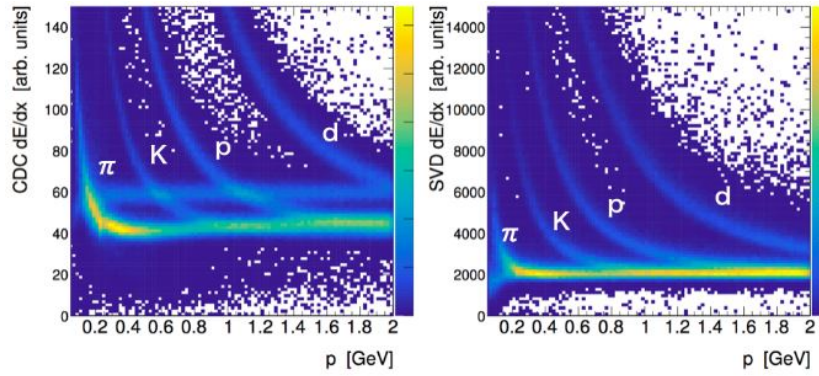
Figure 3.21: Particle separation using the energy loss measurements of the CDC (left) and SVD (right). [38]

ARICH provides additional information for particles that traverse through the forward end cap region using Cherenkov radiation to separate pions with a momentum of 0.5 GeV to 4 GeV from kaons or protons. [5]

The TOP counters are used mainly to separate pions from kaons in the barrel region. This is achieved by measuring the velocity of the particles using the cherenkov emission angle θ_c from cherenkov radiation. [5]

Electrons are mainly identified by the ECL using the shower shapes, the amount of energy deposited in the ECL and by combining the ECL clusters with charged tracks from previous layers. While hadrons continue to propagate the electrons are stopped in the ECL resulting in a larger energy deposit. Furthermore photons can be differentiated from electrons due to the photons leaving no tracks in the previous layers. [5]

Muons are identified by exploiting their penetration depth. Muons with a momentum of at least 0.6 GeV can mainly be identified by using the hits and amount of layers penetrated in the KLM detector. [5]

An overall particle likelihood for each particle type is then calculated by applying a particle hypothesis for each particle on the measurement of each detector independently and then combining all likelihoods to a final particle likelihood result. [5]

The ID of a particle can then be calculated using Equation 3.8, where $\mathcal{L}_i$ is the likelihood of the particle one is looking for and j runs over the likelihoods for the particles e, μ, π, K, p and d. [5]

$$ID = \frac{\mathcal{L}_i}{\sum_j \mathcal{L}_j} \tag{3.8}$$

3.6 Software Framework

The software framework used for online and offline data handling as well as reconstructing events and creating simulated MC samples is the Belle II analysis software framework (basf2). Basf2 is in active development and new software updates are released continuously. Basf2 contains a variety of packages, which contain modules. For the event reconstruction a python script (steering file) is written that loads these modules. The output of the steering file is written into

an ntuple and saved as a ROOT file [39]. [[5], [16]]

3.6.1 MC Samples

Before reconstructing events using real data samples obtained from the Belle II detector the steering file is tested using MC simulations. MC simulated samples are a very important tool for data analysis in general since they can be analysed exactly the same way as real data but the the analysed particles and all their properties are known. The used MC samples should be equivalent to at least three times of the real data luminosity.

The creation of a MC sample in basf2 starts with the generation of physical events by using generators. Depending on what kind of MC sample is produced different generators are used. The generated events are based on a decay file that contains the branching fractions of various decays. For the event generation the following generators are used: EvtGen for decays leading to exclusive final states, PYTHIA for the production of quark pairs and inclusive final states and τ pairs are produced using a combination of KKMC and TAUOLA. [5]
After the event generation a detector simulation using the Geant4 software [40] is done to simulate the physics processes of the particle interactions with the different detector layers and materials. [5]

4 Belle II Data Analysis

4.1 B Meson Reconstruction

Of the B mesons produced using the process $\Upsilon(4S) \to B\overline{B}$ are about 49% B^0 (and $\overline{B}^0$) mesons and 51% $B^\pm$ mesons. The B mesons travel approximately 130 μm before decaying due to the boost resulting from the energy asymmetry of the electron and positron beams. While this distance is enough to separate the two produced B mesons it is not enough for them to reach the detector, where the PXD is the first layer at 14 mm. Furthermore D mesons in their ground and excited states don't have a life time that would be long enough to reach the detector. The only final state particles that are sufficiently long lived to get detected and can be used to reconstruct the B mesons are: [[4], [5]]

- Charged particles: electrons, muons, charged kaons, protons and charged pions

- Neutral particles: neutrons, neutral kaons, photons, ...

During the reconstruction of a particle random combinations of particle "candidates" are tried. To reduce the amount of combinations and particle candidates constraints on various quantities are applied.

To describe the reconstruction process the decay $B^+ \to D^-\pi^+\pi^+$ with the sub-decay $D^- \to K^+\pi^-\pi^-$ will serve as an example. During the reconstruction the charge conjugated decay chain will also be considered.

To reconstruct the D^- meson combinations of a K^+ with two oppositely charged π^- are tried. The invariant mass in Equation 4.1 of the combined particles can then be calculated by using the 4-momenta p_{π_1}, p_{π_2} and p_K of the candidates. Since the nominal mass of D^- is 1.869 GeV the invariant mass of the combined candidates is required to be similar. [11]

$$m_{K\pi\pi} = |p_K + p_{\pi_1} + p_{\pi_2}| = \sqrt{(E_K + E_{\pi_1} + E_{\pi_2})^2 - (\vec{p}_K + \vec{p}_{\pi_1} + \vec{p}_{\pi_2})^2} \quad (4.1)$$

The B^+ meson is then reconstructed by trying combinations of the previously reconstructed D^- with two oppositely charged π^+. The 4-momentum of the B^+ candidate is then calculated adding up the 4-momenta of the combined particles. The reconstructed energy and momentum of the B^+ and half of the known center of mass energy of the beam E^*_{Beam} lead to two powerful variables for extracting the signal from the selected candidates, i.e. the beam constrained mass m_{BC} and the energy deviation ΔE.

The beam constrained mass m_{BC} is defined in Equation 4.2, where $\vec{p}^*_B$ is the reconstructed 3-dimensional momentum of the B meson in the center of mass frame. Due to the characteristics of the collider E^*_{Beam} is known to be equal to $\sqrt{s}/2$=5.29 GeV and m_{BC} peaks approximately at 5.279 GeV for signal events

as seen in Figure 4.1a. [3]

$$m_{BC} = \sqrt{E_{Beam}^{*2} - \vec{p}_B^{*2}} \tag{4.2}$$

The formula of the energy deviation ΔE is given in Equation 4.3, where E_B^* is the reconstructed energy of the B^+ in the center of mass frame. It can be seen in Figure 4.1b that signal events have a peak at zero for ΔE, due to the reconstructed B meson has to carrying approximately half of the center of mass energy.

$$\Delta E = E_B^* - E_{Beam}^* \tag{4.3}$$

(a) Beam constrained mass

(b) Energy deviation

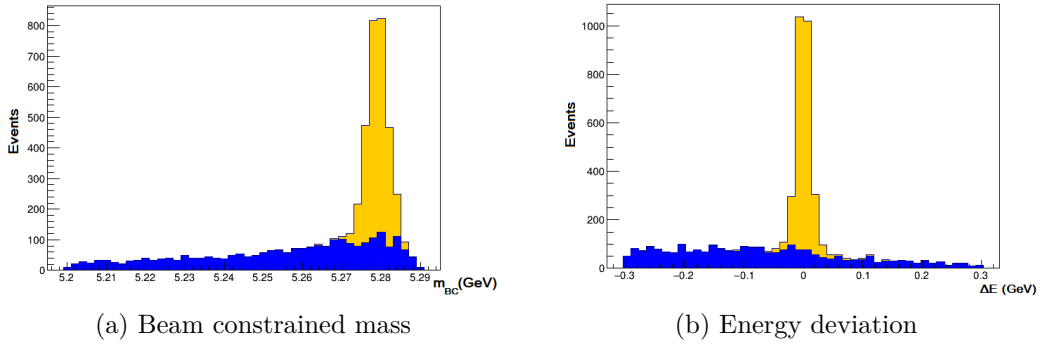Figure 4.1: Beam constrained mass and energy deviation histograms of the hadronic decay $B^+ \to D^- \pi^+ \pi^+$. The signal peaks (yellow) are stacked on top of the flat background (blue).

For semileptonic decays these two variables can not be utilized for signal extraction due to the missing momentum of the neutrino, which can not be measured. A powerful variable that serves as a discriminator between signal and background events in semileptonic decays is $\cos\theta_{B,Y}$, where $Y = D^*\ell$ in case of a $B^0 \to D^{*-}\ell^+\nu$ reconstruction. While the flight direction of the reconstructed B meson is not directly known its direction lies on a cone around the direction of the $D^*\ell$ system as depicted in Figure 4.3a. The cosine of the angle between the B and the $D^*\ell$ in the center of mass frame is given by Equation 4.4. The values of $E_{D^*\ell}^*$, $m_{D^*\ell}$ and $|\vec{p}_{D^*\ell}^*|$ are obtained from the reconstructed D^* and ℓ, $|\vec{p}_B^*|$ is $\sqrt{E_B^{*2} - m_B^2}$ and E_B^* is half of the center off mass energy. A property of $\cos\theta_{B,Y}$ is that the signal events predominantly lie in a range of [-1,1] while the background events are not restricted to that range (see Figure 4.3b). [[3], [13]]

$$cos\theta_{B,D^*\ell} = \frac{2E_{D^*\ell}^* E_B^* - m_B^2 - m_{D^*\ell}^2}{2|\vec{p}_B^*||\vec{p}_{D^*\ell}^*|} \tag{4.4}$$

Another variable that can be used for semileptonic decays as a discriminator between signal and background events for the signal extraction is the invariant mass squared m_{miss}^2 of the missing neutrino momentum defined in Equation 4.5 for a $B^0 \to D^{*-}\ell^+\nu_\ell$ decay. Since neutrinos have a very small mass < 1.1 eV

33

their m^2_{miss} peaks at zero for signal events as shown in an example distribution of a semileptonic B^0_s decay in Figure 4.2. [11]

$$m^2_{miss} = \left(\begin{pmatrix} \sqrt{s}/2 \\ 0 \\ 0 \\ 0 \end{pmatrix} - p^*_\ell - p^*_{D^*} \right)^2 \tag{4.5}$$
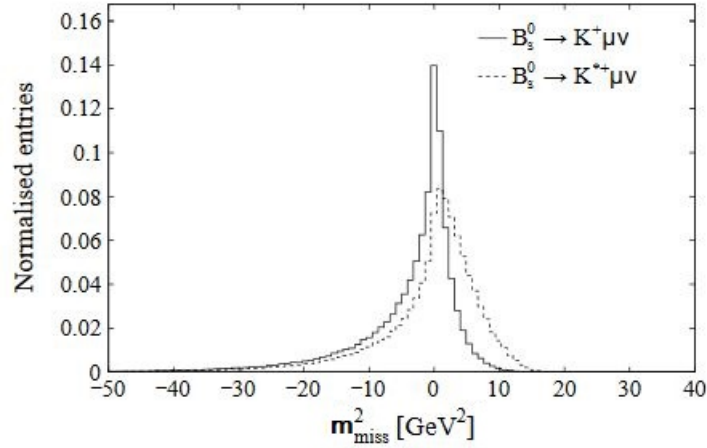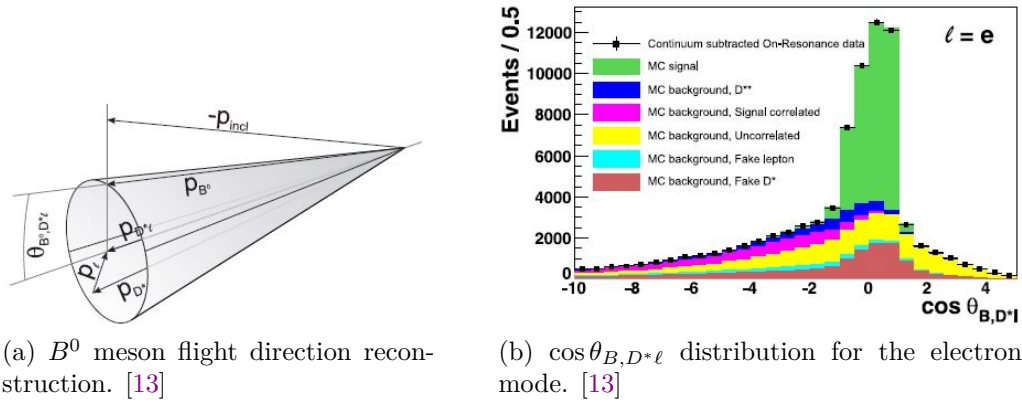

Figure 4.2: Example of a semileptonic B^0_s decay m^2_{miss} distribution. [41]

(a) B^0 meson flight direction reconstruction. [13]

(b) $\cos\theta_{B,D^*\ell}$ distribution for the electron mode. [13]

Figure 4.3: B^0 meson flight direction reconstruction (left) and $\cos\theta_{B,D^*\ell}$ histogram for the electron mode (right).

4.1.1 Tagged and Untagged

An analysis of a decay chain can be performed either tagged or untagged.

In a tagged analysis (see Figure 4.4) the first step of the event reconstruction is the reconstruction of one of two B mesons using a reconstruction algorithm. The B meson reconstructed by the algorithm is called B_{tag}. For the reconstruction of B_{tag} not one but a variety of final states that lead to a $B^\pm$ or B^0 are used, depending on whether the analysis is looking for charged or neutral B meson decays. [[7], [42]]

Due to the reconstruction of B_{tag} one can determine the 4-momentum of the

neutrino using Equation 4.6 (e.g. for $B \to D^* \ell \nu_\ell$) without the neutrino being seen in the detector, because all the kinematics of the event are known after reconstructing the detectable particles from the signal B meson B_{sig}. This can be important for the reduction of combinatorial background. [[7], [42]]

$$p_\nu = p_{HER} + p_{LER} - p_{B_{tag}} - p_{D^*} - p_\ell \tag{4.6}$$

In an untagged analysis (depicted in Figure 4.5) the second B meson is not reconstructed and its kinematics unknown. At first this seems like a disadvantage due to a missing constraint to control backgrounds, however at least at Belle I the algorithm for the B_{tag} reconstruction only had a tagging efficiency of $\mathcal{O}(0.1\%)$. Therefore a lot of events are lost due to the low efficiency and larger data samples are needed. [7]
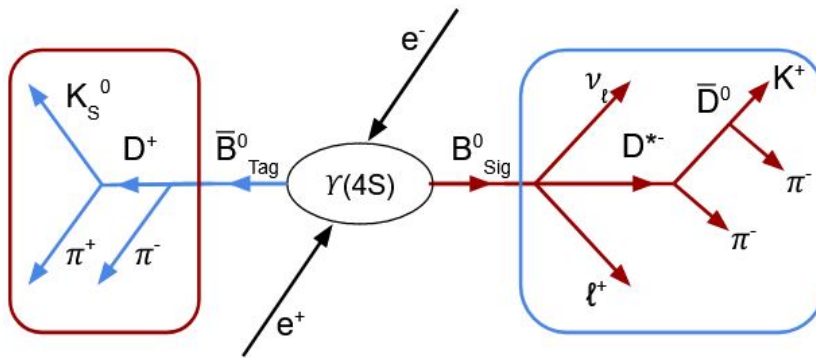
Figure 4.4: Example of a tagged exclusive reconstruction of a $B \to D^* \ell \nu_{ell}$ decay. Adapted from [43].

(a) Exclusive   (b) Inclusive
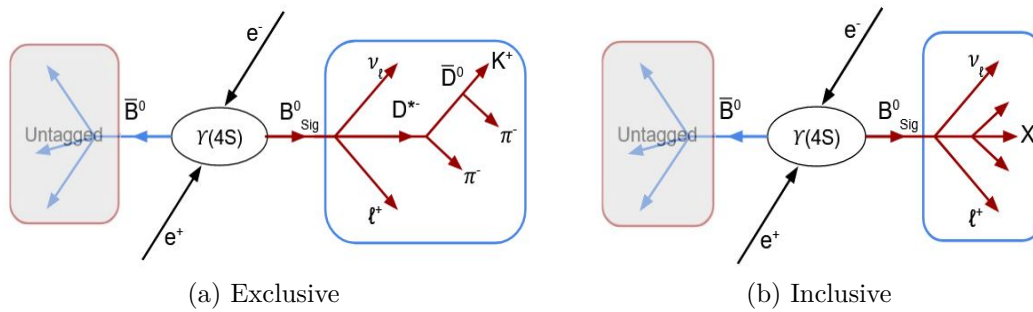
Figure 4.5: Illustration of the principle of untagged inclusive (right) and exclusive (left) reconstruction techniques. Adapted from [43].

4.2 Continuum Suppression

The background can be split up into three main categories:

- Fake leptons: Any background events, where the particle used as charged lepton candidate during the reconstruction is a misidentified hadron.

- Continuum background: Background resulting from $e^-e^+ \to \tau^-\tau^+$ events and from $e^-e^+ \to q\bar{q}$ events, where q = u,d,s,c.

- $B\overline{B}$ background: Any combinatorial background from $B\overline{B}$ events where random particles have a similar kinematics to the final state particles of the signal decay and therefore pass the selection criteria leading to wrong combinations. Another source of this background are $B\overline{B}$ events with similar final state particles.

To get a picture of the probability of different background events that can occur due to e^-e^+ collisions at the given center of mass energy of 10.58 GeV one can take a look at the cross sections of the different physical processes that are listed in Table 3.1. The fake leptons and the $B\overline{B}$ backgrounds are mainly suppressed by optimizing the selection criteria during or after the reconstruction. If the amount of background from similar decays is an issue it can be suppressed by implementing a veto during the reconstruction, i.e. rejecting decays that have kinematics that are too similar to a given background decay.

For continuum events the main suppression stems from making use of the different event shapes resulting from $e^-e^+ \to q\bar{q}$ and $e^-e^+ \to \tau^-\tau^+$ in comparison to $B\overline{B}$ events, as illustrated in Figure 4.7. The produced quark and τ pairs create jets that hit the detector. The particles from these jets may look like the kinematic signature of signal events. This leads to false combinations during the reconstruction process and results in an additional combinatorial background. The shape difference occurs due to the mass and velocity differences between $B\overline{B}$ and continuum events. To take advantage of these different shapes one uses the moments H_l introduced by Fox and Wolfram defined by Equation 4.7. [[7], [44]]

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

In Equation 4.7 the following quantities are used: i,j run over all the particles produced in the event, ϕ_{ij}^* is the opening angle between the particles i and j and P_l the Legendre polynomial of lth order. For the continuum suppression the Fox Wolfram moment ratio R_2 is used, i.e. the ratio of the 2nd and 0th order Fox Wolfram moments in Equation 4.8. The value of R_2 is close to 0 for spherical event shapes, as depicted in Figure 4.6. [[7], [44]]
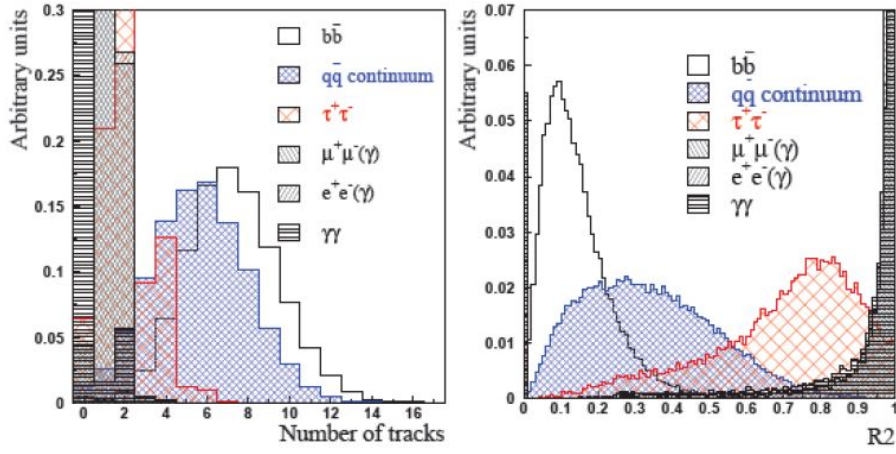
$$R_2 = \frac{H_2}{H_0} \tag{4.8}$$

Figure 4.6: Track multiplicity (left) and R_2 (right) distributions of various events that can occur at a center off mass energy of 10.58 GeV. [7]
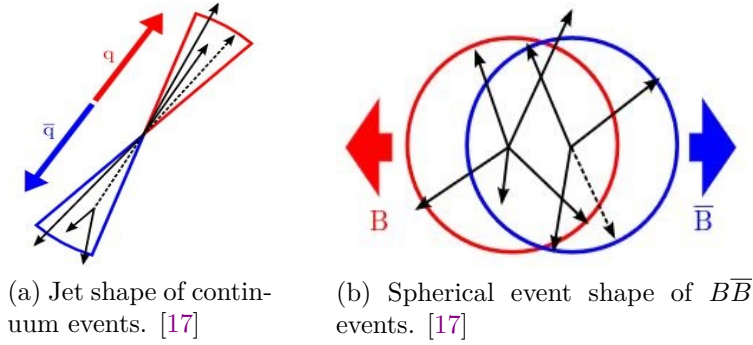

(a) Jet shape of continuum events. [17]

(b) Spherical event shape of $B\overline{B}$ events. [17]

Figure 4.7: Continuum and $B\overline{B}$ event shapes.

4.3 Statistics

4.3.1 Maximum Likelihood Fits (Barlow and Beeston)

In general a maximum likelihood fit is a fit that is determined by calculating the maximum of a likelihood, however in this section the explanation of a specific maximum likelihood algorithm invented by Roger Barlow and Christine Beeston from [45] will be reiterated.

From the event reconstruction one obtains data points. Each data point has a set of variables related to it and the goal is to determine the fractions of the sources of the data. The origin of the data obtained from MC samples is known preemptively and could be for example one of the following categories: Signal, $B\overline{B}$ background, continuum background, fake leptons. This can be used to estimate the amount of each category a sample of real data is composited of by means of a maximum likelihood fit.[[3], [45]]

The data is split up into n bins. The number of events in each bin can then be predicted by Equation 4.9 using the number of MC events a_{ji} for each bin i from a source j, the number of total events from a source in the MC sample N_j, the total number of data events N_D and a fraction P_j that sums up to one.[[3],

[45]]

$$f_i = N_D \sum_j P_i \frac{a_{ji}}{N_j} \tag{4.9}$$

The normalization can be combined to factors p_j resulting in:[[3], [45]]

$$f_i = \sum_j p_j a_{ji} \tag{4.10}$$

A Poisson distribution can be used to calculate probability to find d_i events in a bin i. The fractions p_j of the source components can then be estimated by maximizing the logarithm of the total likelihood in Equation 4.11. [[3], [45]]

$$\ln \mathcal{L} = \sum_i d_i \ln f_i - f_i \tag{4.11}$$

However one has to take a statistical fluctuation of a_{ji} into consideration, which arises from the MC samples having a finite size. Therefore the correct prediction of the number of data events would be Equation 4.12. Where A_{ji} is an unknown, expected number of events that generates a_{ji} using a Poisson distribution. [[3], [45]]

$$f_i = \sum_{j=1} p_j A_{ji} \tag{4.12}$$

This leads to a new logarithm of the total likelihood to be maximized by combining the probabilities for d_{ji} and a_{ji}: [[3], [45]]

$$\ln \mathcal{L} = \sum_i d_i \ln f_i - f_i + \sum_i \sum_j a_{ji} \ln A_{ji} - A_{ji} \tag{4.13}$$

4.3.2 Pull Distributions

Pull distributions are a means of validating maximum likelihood fits. If one creates random variables x based on a Gaussian distribution with a mean μ and width σ then the distribution of its pull defined in Equation 4.14 will be a Gaussian distribution with a mean of zero and a width of one. [[3], [46]]

$$g = \frac{x - \mu}{\sigma} \tag{4.14}$$

When fitting a data sample simulated using an MC simulation then the real values of the amount of signal and the different background components are

38

known. This feature of MC samples can be exploited to validate the fit by creating a distribution of the pull defined by Equation 4.15. Where n_{fit} is the number of events for a given component estimated by the fit, n_{true} is the known true value of events of the component and σ_{fit} is the statistical uncertainty estimated by the fit. By doing this multiple times one obtains a distribution of the pull g, which can be fitted using a Gaussian distribution. [3]

$$g = \frac{n_{fit} - n_{true}}{\sigma_{fit}} \tag{4.15}$$

As mentioned before this should lead to a Gaussian distribution with a mean of 0 and a width of 1 if the fit is unbiased and the statistical uncertainty well estimated. However if the fit is biased then the mean will deviate from 0 in either direction depending on whether the component tends to be over or underestimated. If the width σ of the distribution deviates from 1 then the fit overestimates the uncertainty if $\sigma < 1$ or underestimates the uncertainty if $\sigma > 1$. [3]

4.3.3 Figure of Merit

The figure of merit (fom) is the expected significance of a measurement and can be used to optimize the constraints that are applied to select particle candidates. [3]

The fom for the optimization of the measurement of a known process is defined in Equation 4.16, where N_{sig} is the number of signal events and N_{bg} the number of background events. This means the fom is the ratio of the signal events to an estimated Poisson error. The maximization of the fom can be used as an indicator whether the tightening (or loosening) of a requirement that particle candidates have to fulfill increases or decreases the signal purity. [3]

$$fom = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bg}}} \tag{4.16}$$

5   Signal Selection

5.1 $B \to D^* \ell \nu_\ell$ Reconstruction

The decay $B^0 \to D^{*-} \ell^+ \nu_\ell$ ($\ell = e, \mu$) was reconstructed using the two sub-decays $D^{*-} \to \overline{D}^0 \pi_s^-$ (s denotes the slow pion) and $\overline{D}^0 \to K^+ \pi^-$. During the reconstruction of the decay chain B^0 and $\overline{B}^0$ are both considered. The requirements applied for the particles candidates during the reconstruction itself are looser and will later on be tightened to optimize the signal purity before the fitting procedure. For the reconstruction process the basf2 release 04-02-08 was used.

The events of a data sample are preselected by a trigger (HLT skim). Only events that are classified to originate from a hadron are kept for B meson reconstruction. To be able to trigger this classification the events are required to have at least 3 good tracks (nTracksLE) and must not fulfill the trigger conditions for a Bhabha event with 2 or more tracks (Bhabha2Trk).

Before combining particle candidates to $\overline{D}^0$, D^{*-}, and B^0 an additional event selection is made. All events are required to have at least 3 tracks (nTracks) and the visible energy of the event in the center of mass frame has to be above 4 GeV (E_{vis}^*) for an event to be considered for the reconstruction procedure.

The next step is a track selection. All final state particle candidates for $B^0 \to D^{*-} \ell^+ \nu$, i.e. K^+, π^-, slow pions π_s^- and ℓ^+ are required to fulfill the following track selection criteria:

- The polar angle θ of the candidate has to lie inside the CDC acceptance range of $17° < \theta < 150°$

- The absolute value $|d_0|$ of the signed distance d_0 of the point of closest approach in the $r - \phi$ plane to the IP has to be below 0.5 cm

- The absolute value $|z_0|$ of the signed distance z_0 of the point of closest approach on the z-axis to the IP has to be below 2 cm.

To reconstruct $\overline{D}^0$ a π^- is combined with an oppositely charged K^+. For K^+ and π^- candidates no additional constraints are applied. The mass of the $\overline{D}^0$ resulting from the combination of the K^+ and π^- candidates is required to be within a range of ± 0.015 GeV of its nominal mass of 1.865 GeV [11].

The slow pion π_s^- candidate used for the D^{*-} reconstruction is required to have a low momentum due to the D^{*-} decay kinematics. During the reconstruction π_s^- is required to have a momentum p_{π_s} below 0.5 GeV in the laboratory frame. The π_s^- candidate is then combined with the previously reconstructed $\overline{D}^0$. A mass constraint is applied on the newly formed D^{*-}. Since the nominal mass of $\overline{D}^0$ is 1.865 GeV and for D^{*-} 2.010 GeV [11] the mass difference Δm_D between the D^{*-} and $\overline{D}^0$ candidate is required to be below 0.155 GeV. Additionally the momentum in the center of mass frame $p_{D^*}^*$ of the D^{*-} is required to be below 2.5 GeV.

The D^{*-} candidate is then combined with an oppositely charged lepton to

form a B^0 candidate. The neutrino can not be used during the reconstruction, since it can not be measured by the detector. In the laboratory frame the momentum of the electron has to be higher than 0.3 GeV, while for the muon it has to be higher than 0.6 GeV and their momentum in the center of mass frame has to be below 2.4 GeV. Additionally both leptons require a particle identification probability (ID, see Section 3.5.3) greater than 0.5.

The reconstructed B^0 itself has no additional requirements to fulfill. All of the constraints applied during the reconstruction itself are summarized in Table 5.1 and yield the $\cos\theta$ distributions shown in Figure 5.1.
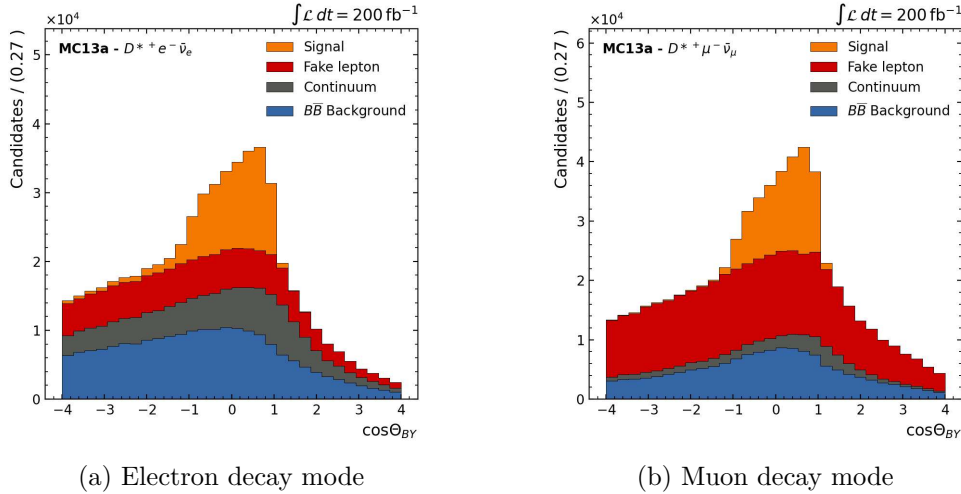

(a) Electron decay mode

(b) Muon decay mode

Figure 5.1: $B \to D^*\ell\nu_\ell$ $\cos\theta$ distributions acquired from the reconstruction with loose constraints for both decay modes.

Preselection	$nTracksLE \geq 3$
	$Bhabha2Trk = 0$
Event	$nTracks \geq 3$
	$E^*_{vis} > 4$ GeV
Tracking	$|d_0| < 0.5$ cm
	$|z_0| < 2$ cm
	θ in CDC acceptance
Leptons	$\ell\text{ID} > 0.5$
	$p_e > 0.3$ GeV
	$p_\mu > 0.6$ GeV
	$p^*_\ell < 2.4$ GeV
Slow pion	$p < 0.5$ GeV
D mesons	1.85 GeV $< m_D < 1.88$ GeV
	$\Delta m_D < 0.155$ GeV
	$p^*_{D^*} < 2.5$ GeV

Table 5.1: Constraints applied on the particle candidates during the $B \to D^*\ell\nu_\ell$ reconstruction.

5.2 Signal Selection Optimization

The signal selection is optimized by using MC simulated samples. In the MC samples the relations between the particles and their sources (continuum, fake lepton, $B\overline{B}$ background, signal) are known. Additionally, every particle candidate has a MC PDG ID assigned to it, which identifies every particle type and can be looked up in a table. Furthermore every particle candidate has a signal flag assigned to it, i.e. a value of 1 or 0, which enables one to identify whether a particle candidate was chosen correctly or not.

Using these properties of the MC samples one can split up the candidates into signal and background categories and can visualize the distributions of different variables to determine which regions are predominantly populated by either a background component or signal. This information in combination with the figure of merit, which is a measure for signal purity, can be used to optimize the signal selection.

The flow chart in Table 5.2 shows the changes in the amount of signal N_{sig}, the amount of background N_{bg}, the ratio of N_{sig} to the amount of overall candidates N_{all} and the fom and indicates which additional requirements are the most impactful. The flow chart starts with the constraints applied during the reconstruction and shows the changes after each additional applied selection. The most impactful constraints are the tightening of the lepton momentum in the center of mass frame and the mass difference Δm_D between D and D^*. The additional constraint on the Fox Wolfram moment ratio R_2 and the tightening of the lepton ID show a minor improvement.

$$B^0 \rightarrow D^{*-}e^+\nu_e$$

	N_{sig}	N_{bg}	N_{sig}/N_{all}	fom
Reconstruction	104816	808926	0.115	109.652
$R_2 < 0.3$	94420	622839	0.132	111.487
$0.144\ \text{GeV} < \Delta m_D < 0.148\ \text{GeV}$	85855	341787	0.201	131.288
$1.2\ \text{GeV} < p_\ell^* < 2.4\ \text{GeV}$	56340	47550	0.542	174.795
$eID > 0.9$	54511	42679	0.561	174.853

$$B^0 \rightarrow D^{*-}\mu^+\nu_\mu$$

	N_{sig}	N_{bg}	N_{sig}/N_{all}	fom
Reconstruction	101865	774810	0.116	108.794
$R_2 < 0.3$	91642	597765	0.133	110.371
$0.144\ \text{GeV} < \Delta m_D < 0.148\ \text{GeV}$	83232	335099	0.199	128.686
$1.2\ \text{GeV} < p_\ell^* < 2.4\ \text{GeV}$	62681	60981	0.507	178.245
$\mu ID > 0.9$	61181	53892	0.532	180.356

Table 5.2: Flow chart showing the changes of the amount of signal and background, as well as the signal ratio and figure of merit when applying tighter constraints.

The following plots of the distributions of the variables are shown after applying the final constraints on all the other variables. The vertical red lines in

the distributions indicate the final selection requirement on the shown variable and the transparent region indicates the part that did not pass the selection criterion.

A majority of the reconstructed events that include a lepton candidate with low momentum in the center of mass frame p_ℓ^* below 1.2 GeV are $B\overline{B}$ and fake leptons background events as indicated by the plots in Figure 5.2. The fake leptons are mainly misidentified pions. By applying the tighter constraint on the momentum in the center of mass frame 1.2 GeV $< p_\ell^* <$ 2.4 GeV of the selected lepton candidates a majority of the fake leptons and $B\overline{B}$ background for both the electron and muon mode gets rejected. Since there is no peak for the lepton momenta for signal events the restriction of 1.2 GeV for the minimum of the momentum was deduced by scanning the fom (see 4.3.3) in the region of 1.1 GeV to 1.25 GeV, where 1.2 GeV yielded the best result.

It can be seen in Figure 5.3 that the shapes of the Fox Wolfram moment ratio R_2 (see 4.8) of the continuum background and $B\overline{B}$ background differ in their shapes, which results in the continuum events mainly occupying regions with a higher moment. Therefore a large amount of continuum background gets reduced by applying a requirement of $R_2 < 0.3$.

The distributions of the mass difference between the D^* and D mesons can be used to reject large amounts of $B\overline{B}$ background events. In Figure 5.4 it can be seen that the signal predominantly populates the region of 0.144 GeV $< \Delta m_D <$ 0.148 GeV, while the region outside that range is mainly occupied by $B\overline{B}$ background, therefore the Δm_D requirement was tightened to this region.

An additional requirement for the lepton IDs of eID and μID to be above 0.9 is in comparison only a minor improvement since most fake leptons already get rejected by the tighter momentum cut Only a small amount of events is left in the region below 0.9 in the ID distributions, as seen in Figure 5.5. The thresholds of the lepton IDs can be set to either 0.5, 0.9 or 0.95 and 0.9 yielded the best fom.

A summary of the final signal selection requirements and a comparison to the constraints applied on the particle candidates during the reconstruction process is given in Table 5.3. The $\cos\theta$ distributions in Figure 5.6 help to visualize how much the ratios of the components has changed and the signal purity was increased by tightening the constrains on the discussed quantities in comparison to 5.1.
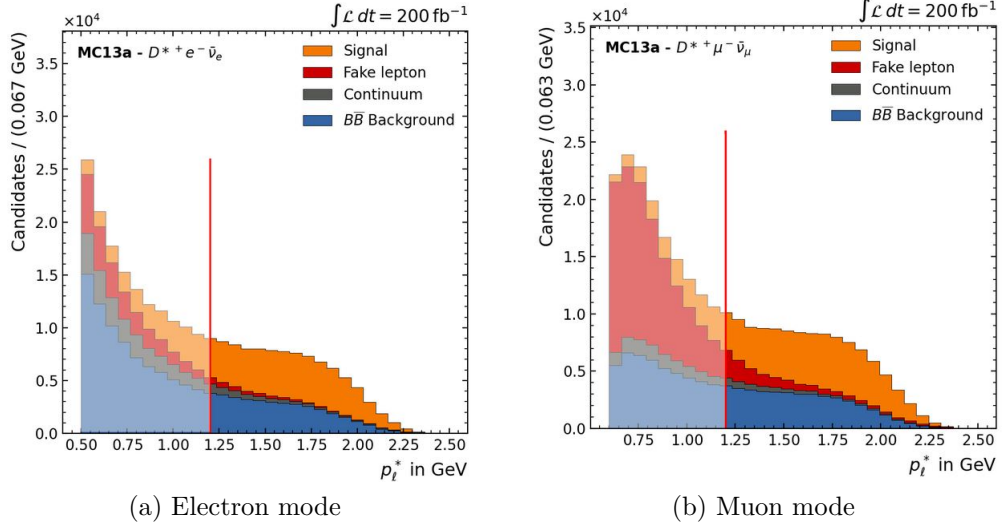
(a) Electron mode

(b) Muon mode

Figure 5.2: Distribution of the lepton momentum in the center of mass frame p_ℓ^* showing large amounts of fake leptons and $B\overline{B}$ background in the region below 1.2 GeV.
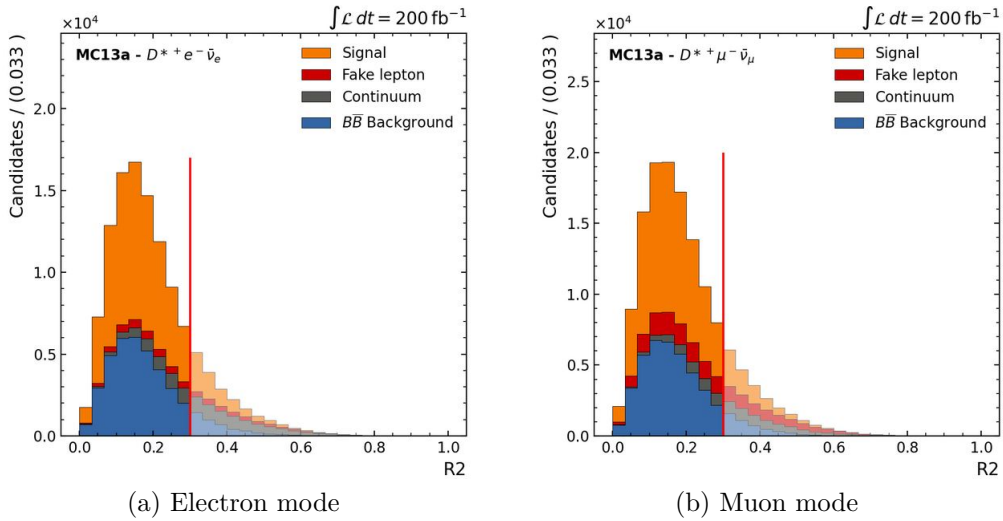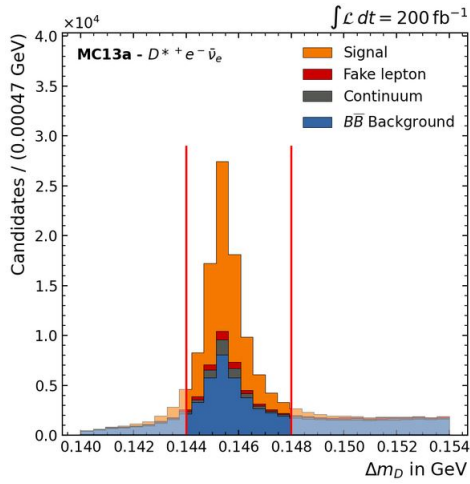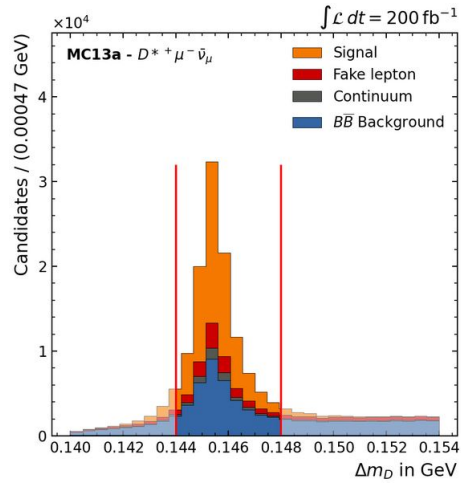

(a) Electron mode

(b) Muon mode

Figure 5.3: Fox Wolfram moment ratio R_2 distributions.
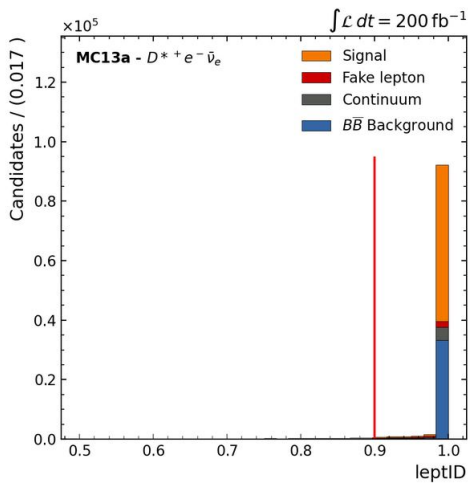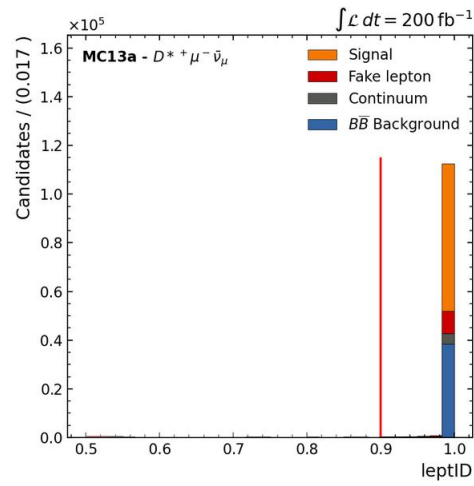
(a) Electron mode

(b) Muon mode

Figure 5.4: Distributions of the mass difference between D^* and D for the electron and muon mode.

(a) Electron mode

(b) Muon mode

Figure 5.5: LeptonID distributions showing that only very small amounts of background are left below 0.9.
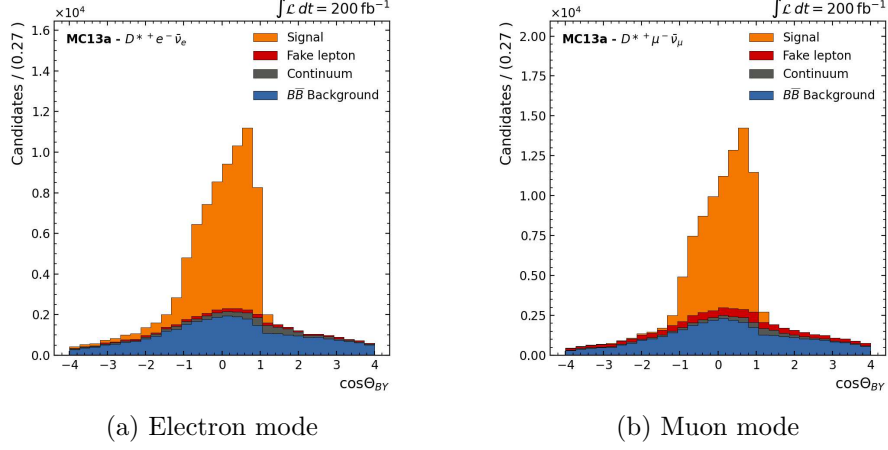
(a) Electron mode   (b) Muon mode

Figure 5.6: Components of the $\cos\theta$ distributions after the final signal selection.

	Reconstruction	**Final Selection**
Preselection	$nTracksLE \geq 3$	$nTracksLE \geq 3$
	$Bhabha2Trk = 0$	$Bhabha2Trk = 0$
Event	$nTracks \geq 3$	$nTracks \geq 3$
	$E_{vis}^* > 4$ GeV	$E_{vis}^* > 4$ GeV
		$R_2 < 0.3$
Tracking	$|d_0| < 0.5$ cm	$|d_0| < 0.5$ cm
	$|z_0| < 2$ cm	$|z_0| < 2$ cm
	θ in CDC acceptance	θ in CDC acceptance
Leptons	$\ell\mathrm{ID} > 0.5$	$\ell\mathrm{ID} > 0.9$
	$p_e > 0.3$ GeV	$p_\ell^* > 1.2$ GeV
	$p_\mu > 0.6$ GeV	
	$p_\ell^* < 2.4$ GeV	$p_\ell^* < 2.4$ GeV
Slow pion	$p_{\pi_s} < 0.5$ GeV	$p_{\pi_s}^* < 0.4$ GeV
D mesons	$1.85 < m_D < 1.88$ GeV	$1.85 < m_D < 1.88$ GeV
	$\Delta m_D < 0.155$ GeV	$0.144 < \Delta m_D < 0.148$ GeV
	$p_{D^*}^* < 2.5$ GeV	$p_{D^*}^* < 2.5$ GeV

$B^0 \rightarrow D^{*-}e^+\nu_e$		
	Reconstruction	**Final Selection**
N_{sig}	104816	54511
N_{bg}	808926	42679
N_{sig}/N_{all}	0.115	0.561
fom	109.652	174.853

$B^0 \rightarrow D^{*-}\mu^+\nu_\mu$		
	Reconstruction	**Final Selection**
N_{sig}	101865	61181
N_{bg}	774810	53892
N_{sig}/N_{all}	0.116	0.532
fom	108.794	180.356

Table 5.3: Comparison summary between the reconstruction cuts and the final signal selection constraints and N_{sig}, N_{bg}, the signal ratio and fom.

5.3   Prefit Data - MC Agreement

Using the requirements for particle candidates discussed in the previous chapter and summarized in Table 5.3 the distributions obtained from MC samples and real data samples are compared. The MC simulated samples were used from the production MC13a and amount to a luminosity of 200 fb^{-1}, while the luminosity of the real data used in this analysis, was processed at Belle II in 2019 and 2020 and amounts to 34.6 fb^{-1}. The real data sample includes the following data sets: proc11 (8.764 fb^{-1}), bucket9 (2.768 fb^{-1}), bucket 10 (10.361 fb^{-1}), bucket 11 (12.687 fb^{-1}). Both samples were reconstructed using release 04-02-08 of basf2.

To be able to compare data and MC the distributions obtained from MC samples were scaled down by a factor $\mathcal{L}_{data}/\mathcal{L}_{MC} = 0.173$. The distribution of each variable is shown without applying the final constraint on it and the $\cos\theta$ distribution is restricted to a range of $-4 < \cos\theta < 4$.

The Δm_D distributions (Figure 5.7) show that the electron mode of the MC simulated measurement is in good agreement with the real data while the muon mode shows MC overestimations in the signal region.

The distributions of the lepton momenta in the center of mass frame in Figure 5.8 show a bit of a disagreement between data and MC samples in the lower momentum region especially below $p_\ell^* < 1.2$ GeV, which gets removed due to constraints during the signal selection. However again the discrepancy for the muons is bigger than for the electrons.

In the distributions in Figure 5.9 of the missing squared mass of the neutrino momentum m_{miss}^2 defined in Equation 4.5 a disagreement in the region above 1 GeV^2 can be seen for both modes. The muon mode additionally also shows overestimations in the signal region.

The $\cos\theta$ distributions in Figure 5.10 also show a good agreement for the electron mode, but some disagreements for the muon mode in the signal region.

The shown disagreements in the distributions between the MC and the data samples can be explained due to mismodelling in the lepton identification efficiency and lepton mis-identification rate estimations of the lepton IDs implemented in the MC simulations. These issues can be corrected by applying correction weights in addition to the scaling.
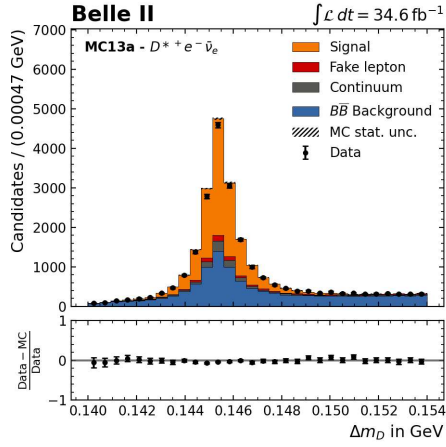
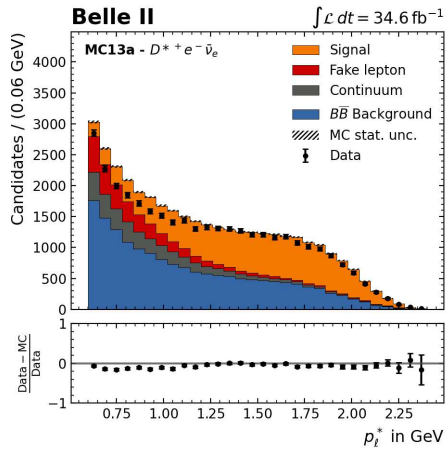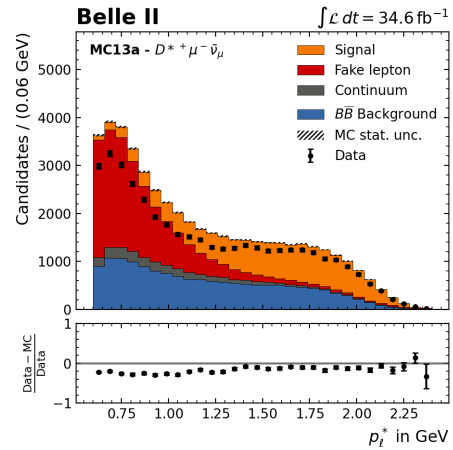(a) Electron mode
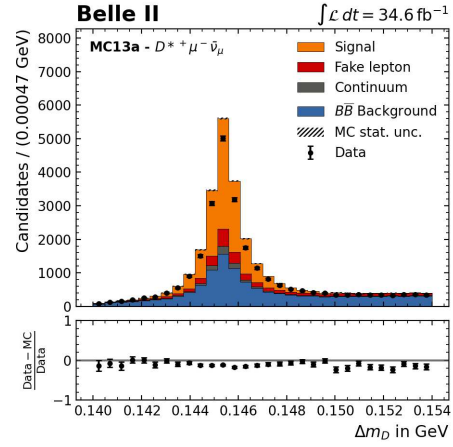
(b) Muon mode

Figure 5.7: MC - Data agreement of the Δm_D distributions showing some disagreement in the signal region of the muon mode.
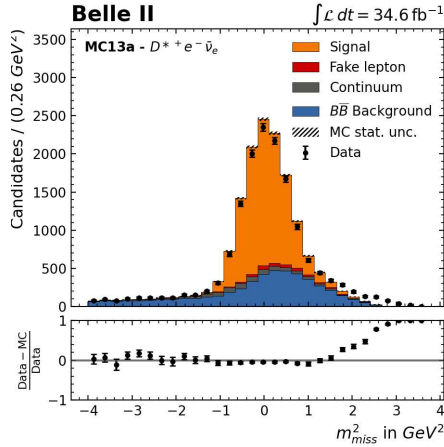

(a) Electron mode

(b) Muon mode

Figure 5.8: Agreement of the distributions of the momentum of the lepton in the center of mass frame. Some discrepancies in both modes can be seen, especially in the low momentum region below 1.2 GeV.
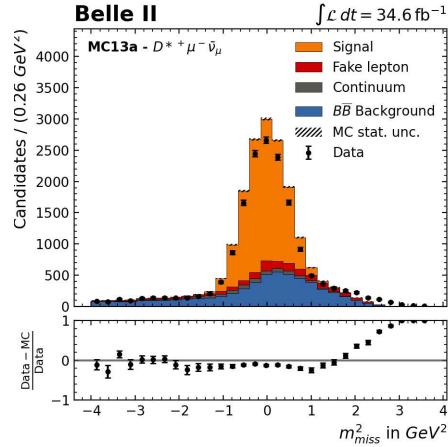
(a) Electron mode

(b) Muon mode

Figure 5.9: The comparison of the MC and data samples of the m_{miss}^2 shows a deviation for both modes in the region above 1 GeV^2 and additionally in the signal region of the muon mode.
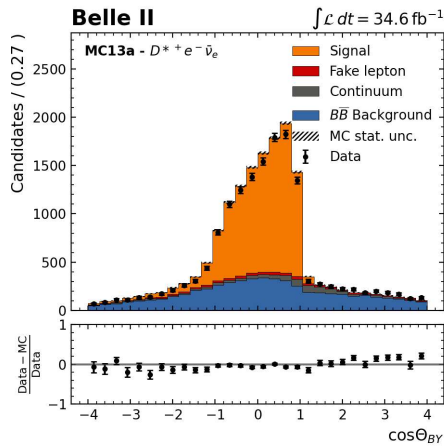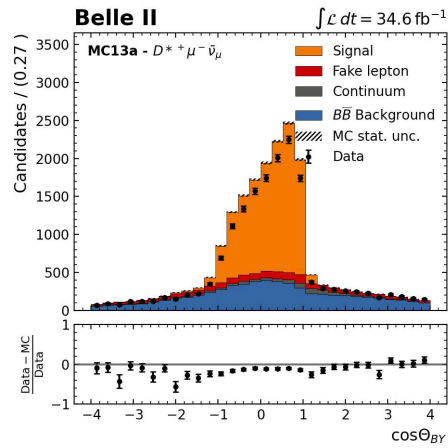
(a) Electron mode

(b) Muon mode

Figure 5.10: The data and MC samples of the $\cos\theta$ distributions agree well for the electron mode, while the muon mode shows some discrepancies in the signal region.

5.4 Lepton ID Corrections

To correct the issues in the implementation of the lepton identification efficiency and lepton mis-identification probability (fake rate) in the MC simulated samples weights are applied in addition to the scaling on the entries of the MC sample.

The correction tables are provided by the Lepton ID Performance Group. The tables are split up into different categories, i.e. efficiency corrections for true leptons, fake rate corrections for misidentified pions and fake rate corrections for misidentified kaons. The corrections are provided as weights with

statistical and systematic uncertainties. These weights are split into different (p, θ) bins, based on the lepton momentum (in the laboratory frame) and its polar angle.

The corrections are applied to the B meson candidates by iterating through the table entries and checking whether the momentum and theta angle of the lepton originating from the B meson fits into a correction bin. If the (p, θ) values of a lepton agree with a correction bin then the respective correction weight will be added to the B meson candidate. If no entry is found then the weight of the B meson will be the value 1. Since all of the leptons used to reconstruct the B mesons from the MC sample have a PDG ID (Particle Data Group identification code) attached to them one can identify them and knows which table has to be applied.

A summary of the coverage of the provided correction tables for their respective particle type after applying all final selection constraints is given in Table 5.4. For the coverage table it has to be noted that the numbers given is the percentage of the particles that are not covered by their respective correction table. It can be seen that the true leptons and misidentified pions are very well covered: for the electrons (muons) only 0.004% (0.962%) of true leptons and 0.083% (1.405%) of the misidentified prions are not covered. Furthermore 13.514%(12.147%) of the misidentified kaons are covered for the electron (muon) mode. At a first glance it seems that a lot of B meson candidates could slip through the fake rate corrections due to the worse kaon coverage but only approximately 0.08%(0.6%) of the B meson candidates in the electron (muon) mode contain a lepton that is a mis-identified kaon. Other mis-identified hadrons like protons are not corrected by these correction tables. In Figures 5.11, 5.12 and 5.13 two-dimensional (p, θ) heat map plots of the lepton momenta and polar angles are shown after the application of the final signal selection. In these plots the bins of the correction tables are indicated as the red squares.

	$real$ [%]	$fake\,\pi$ [%]	$fake\,K$ [%]
$B^0 \rightarrow D^{*-}e^+\nu_e$	0.004	0.083	13.514
$B^0 \rightarrow D^{*-}\mu^+\nu_\mu$	0.962	1.405	12.147

Table 5.4: Summary of the coverage for each correction table in both decay modes after applying the final selection constraints. The table shows the amount of leptons that are not covered.

(a) Electron mode

(b) Muon mode
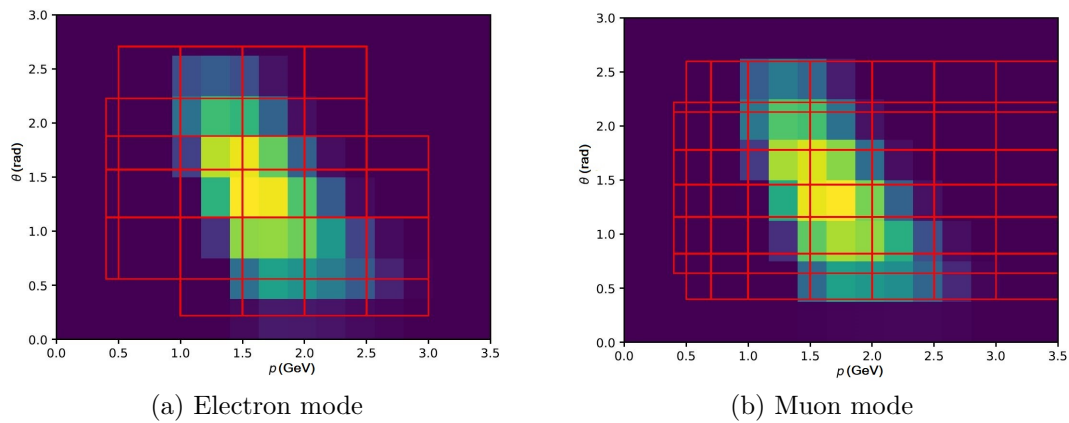
Figure 5.11: Heat map showing the two-dimensional (p, θ) distributions of the true leptons and the bins of the correction tables indicated by the red squares.

(a) Electron mode

(b) Muon mode
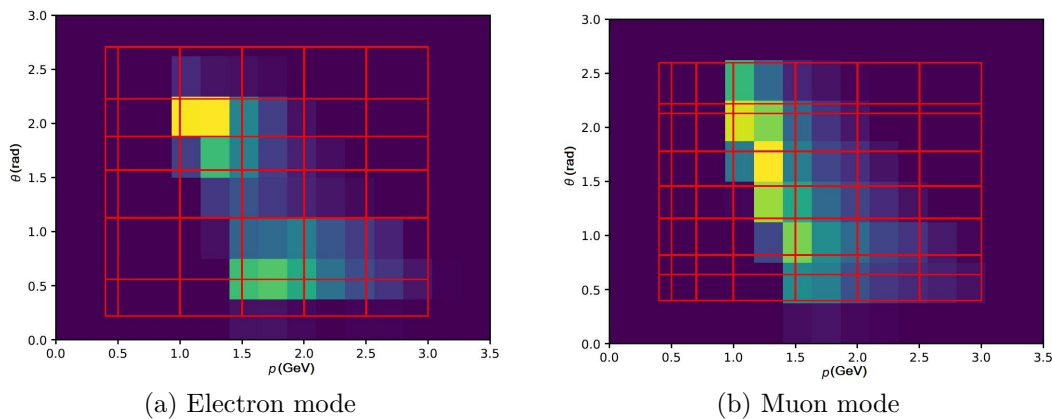
Figure 5.12: Two-dimensional heat map of the (p, θ) distributions of misidentified pions showing the coverage of the correction tables.

(a) Electron mode

(b) Muon mode
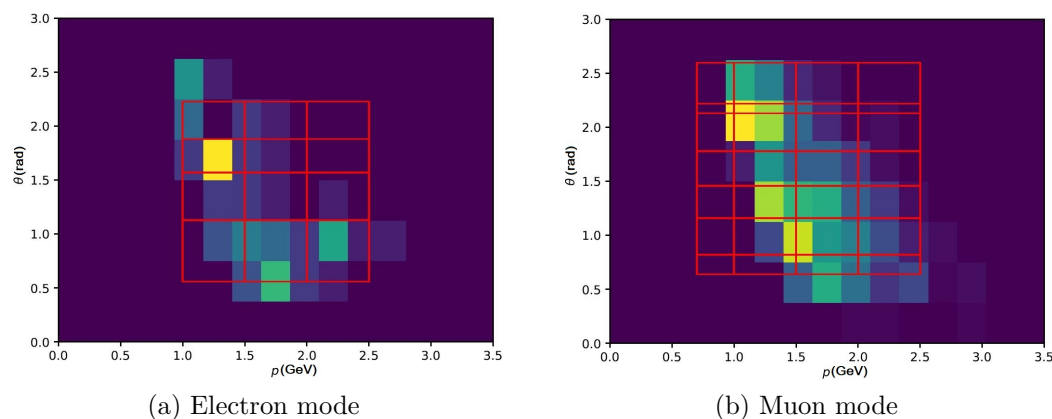
Figure 5.13: Coverage of the (p, θ) bins of the correction tables for misidentified kaons.

5.5 Corrected Prefit Data - MC Agreement

After applying the correction tables in the previously described manner the weights were applied to the MC distributions by multiplying the scaling factor with the correction weight attached for each B meson candidate. The agreement between the MC and data samples improved for all relevant regions used after the final signal selection criteria.

In the m_{miss}^2 distributions in Figure 5.14 one can see a major improvement in the agreement for both lepton modes above 1 GeV^2 as well as for the signal region in the muon mode and a minor improvement for the electron mode.

The same trend can be seen in the distributions of $\cos\theta$ (Figure 5.15) and Δm_D (Figure 5.16). While the electron modes already had a good agreement between MC and data samples there are still some minor improvements, while the discrepancies in the muon mode seen previously got corrected very well and a good agreement can be seen now as a result.

For the p_l^* distributions in Figure 5.17 an overall improvement in the agreement can be seen for the muon mode. However in the electron mode there is a slight improvement in the bins above 1.2 GeV, while in the lower momentum region the disagreement between the MC and data samples got worse. This could happen due to the low momentum regions of electrons not being described very well at the moment in the correction tables, but since the low momentum region below 1.2 GeV of the leptons gets rejected due to the very large amounts of various backgrounds it is not concerning for the current analysis.

(a) Electron mode   (b) Muon mode
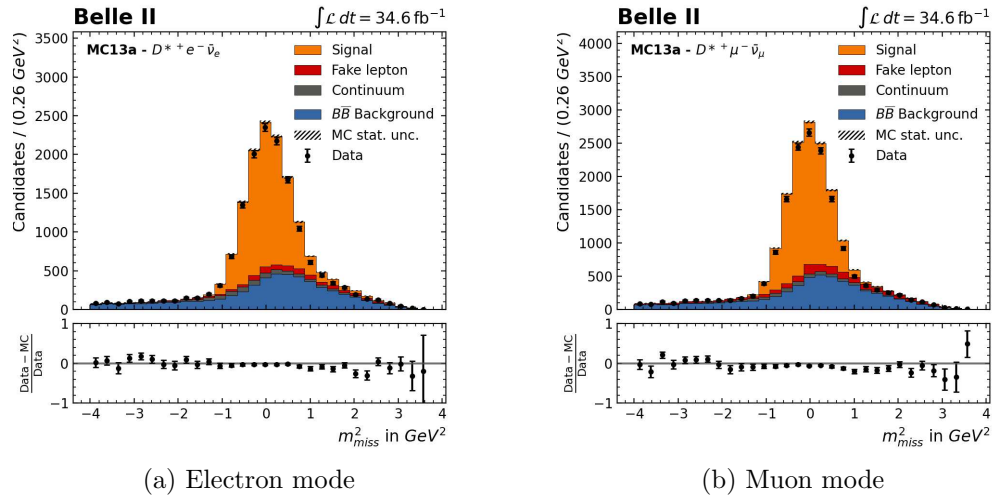
Figure 5.14: For the m_{miss}^2 distribution the correction weights solved the discrepancies in the region above 1 GeV for both modes and in the signal region of the muons.
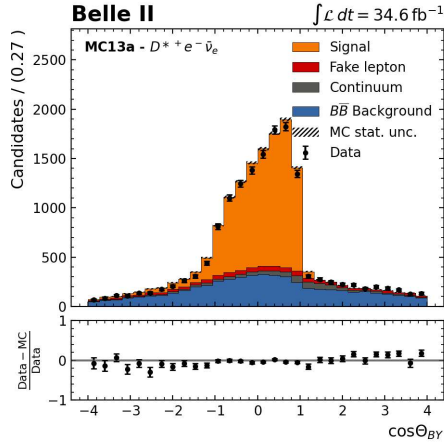
(a) Electron mode
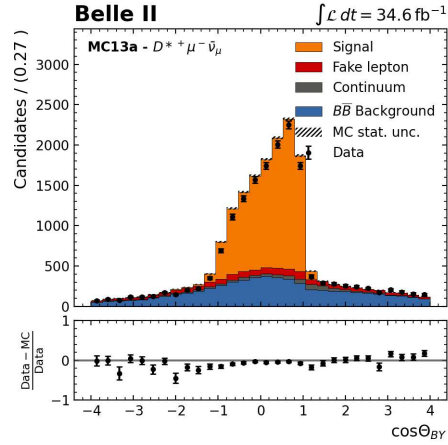
(b) Muon mode

Figure 5.15: The $\cos\theta$ distribution shows very good agreement for both modes after the application of the correction tables.
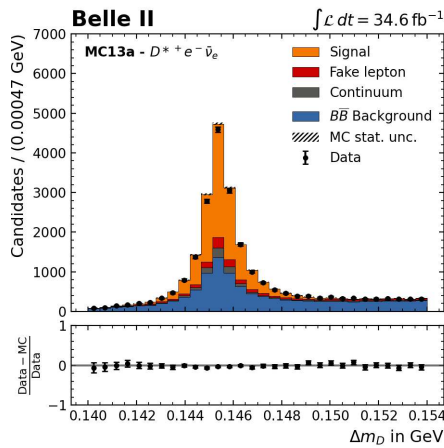
(a) Electron mode
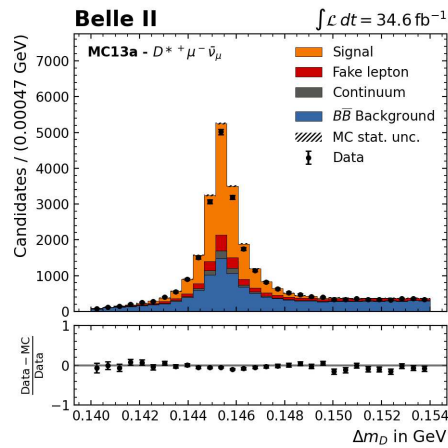
(b) Muon mode

Figure 5.16: After the application of the lepton corrections both modes show very good data - MC agreement.

(a) Electron mode   (b) Muon mode
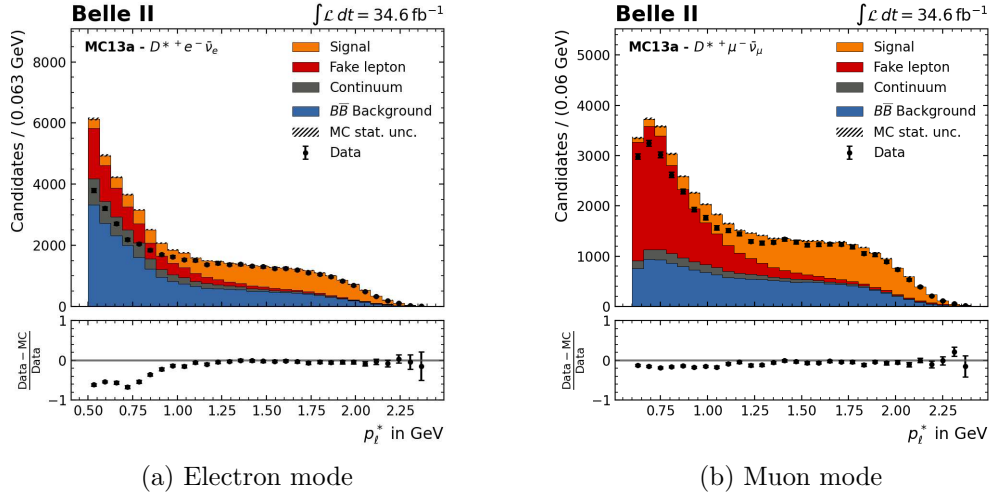
Figure 5.17: The application of the correction weights resulted in larger discrepancies in the low momentum region below 1 GeV for the electron mode, but for the relevant region above 1.2 GeV the agreement improved for both modes.

6 Fitting Procedure

6.1 Fitter Setup

The method used to create a fit for the data sample is the TFractionFitter, which is implemented in ROOT. The TFractionFitter creates a likelihood fit (see Section 4.3.1) based on the shapes of MC templates. These templates can be defined as different components of the MC and fit to the data histogram returning a fraction for each component that overall sum up to 1. [[45], [47]]

An additional feature of this algorithm is that it allows the implementation of weight histograms that can be handed over to the fitter for each template. These weight histograms will then be applied bin by bin for a given template when fitting the MC fractions on data.

However, the fitting algorithm has issues when facing templates of different components that have a very similar shape, therefore the original three background categories, i.e. continuum, fake leptons and $B\overline{B}$ background had to be reduced to two components. The background templates had very similar shapes when split up into three components leading to only continuum and $B\overline{B}$ background templates in addition to the signal template. The shapes used to create the fit are depicted in Figure 6.1 and 6.2 for the electron mode.

The weight histogram feature of the TFractionFitter is used to implement the lepton ID corrections during the fit. By creating an additional histogram for each template, where the weights of the correction gets applied one can create a weight histogram by dividing the bin values of the weighted histogram by the value of the non weighted entries. E.g. for the signal component: $W_{sig} = H_{sigW}/H_{sig}$, where W_{sig} is the resulting weight histogram, H_{sigW} the weighted histogram of the template and H_{sig} the non weighted histrogram.


(a) $B\overline{B}$ background shape
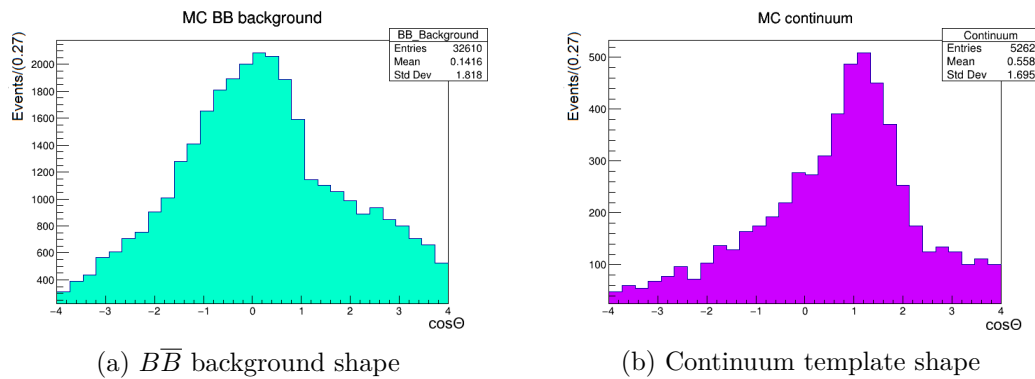
(b) Continuum template shape

Figure 6.1: The $\cos\theta$ distribution shapes of the MC templates of the $B\overline{B}$ background and continuum components. The additional flag for fake leptons was dropped during the fit due to it causing problems with the algorithm because its shape was to similar to the other components and the fake leptons are now added to the other components.
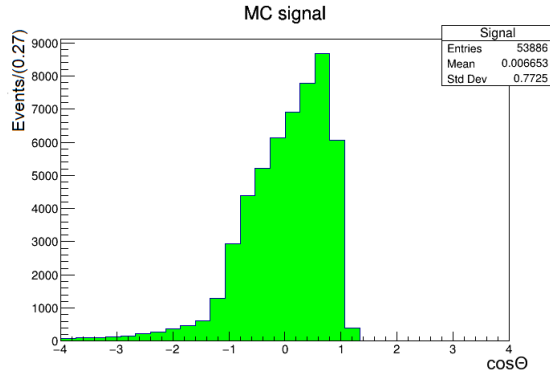
Figure 6.2: Shape of the MC signal template of the $\cos\theta$ distribution.

6.2 Fit Validation

Before applying the fit on real data the fitting procedure was first validated using MC samples and pull distributions. The following two methods were used to create the pull distributions (see Section 4.3.2) to test the fit:

- **GSIM pull test:** The entries of the ntuple obtained from the 200 fb^{-1} MC sample get shuffled to randomize the order of the entries and then split up into 11 equal parts. During this test 1/11 of the entries will be used as a simulated data sample while the other 10/11 parts will be used as the templates. This procedure will be done 11 times. Each time after the pulls were calculated the simulated data sample will be added to the template sample and another 1/11 part is chosen as the next test sample. Since the entries of the ntuple are randomized the 11 parts of the MC are not correlated to each other. The pull for each test sample is calculated for each template component using the known true value of the weighted fractions of the test sample and the estimated value by the fit as well as its estimated statistical uncertainty. The resulting pulls are then plotted and their distribution is fitted by an unbinned gaussian fit to obtain the mean and width σ of the pull distribution.

- **Toy MC pull test:** In the second method templates from the whole 200 fb^{-1} sample are used and not split up. These templates are applied on toy MCs to create a fit. The toy MCs are created by drawing poisson distributed random numbers based on the weighted MC histogram entries for each bin and filling the number of entries into the toy MC histogram bins. The pull of each component is calculated using the estimated fraction of the fit, its statistical uncertainty and the known weighted fractions of the of the components in the templates under the assumptions that on average the poisson distributed numbers of the entries of each bin in the toy MC will overall add up to the fractions of the template. This procedure is done 1500 times. The resulting pull distribution is again fitted with an unbinned gaussian fit to obtain the mean and width.

The results of the unbinned gaussian fit that was applied on the pull distribution are summarized in Table 6.1 and Table 6.2 for the GSIM and toy MC

method, respectively. The gaussian fits and the kernel density estimation of the pull distributions of the signal component for the electron mode are shown in Figure 6.3 for the GSIM test and in Figure 6.4 for the toy MC test.

Both methods result in a bias in the order of (sub)percent for every component in both decay modes, i.e. a bias of 0.5% (0.1 %) for the electron (muon) mode for the GSIM test and a bias of 1.5% (2.5 %) for the electron (muon) mode of the toy MC test for the signal component. Therefore it can be concluded that the fit shows no significant bias in either direction. The widths of the gaussian fits deviate a bit from 1 for the signal component of both implemented methods. The deviation from a width of 1 is bigger for the GSIM test. It can be assumed that the cause of the larger deviation for the GSIM test is due to the pull distribution being based on only 11 points in comparison to the 1500 data points of the toy MCs. For the signal component with a width σ below 1 for both methods this means that the fit overestimates the statistical uncertainty of the fit. However, this way the uncertainty is determined conservatively, which is not a concern.

$$B^0 \rightarrow D^{*-} e^+ \nu_e$$

	mean	σ
Signal	0.005	0.745
$B\overline{B}$ Background	0.005	1.278
Continuum Background	-0.041	1.344

$$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$$

	mean	σ
Signal	0.001	0.574
$B\overline{B}$ Background	-0.018	0.741
Continuum Background	0.007	1.027

Table 6.1: GSIM pull test results for all components.
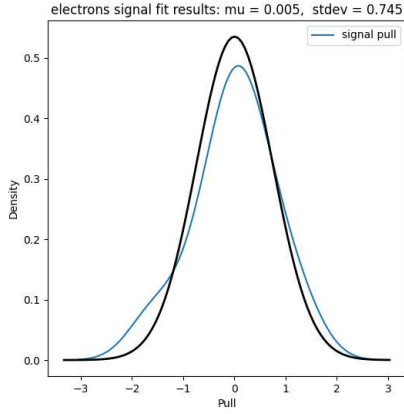
(a) Electron mode
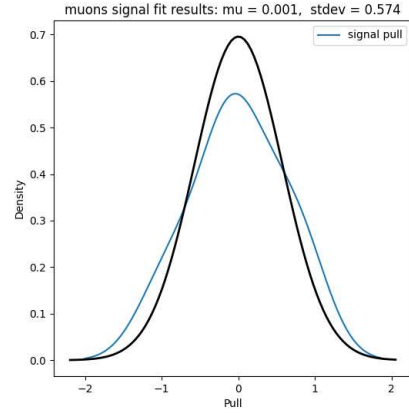
(b) Muon mode

Figure 6.3: Resulting kernel density estimation of the pull distributions (blue curve) from the GSIM pull test for the signal component for the electron and muon mode and their respective mean and width of the gaussian fit(black curve).

$$B^0 \to D^{*-}e^+\nu_e$$
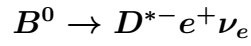
	mean	σ
Signal	0.015	0.813
$B\overline{B}$ *Background*	0.01	0.959
Continuum Background	-0.037	0.944

$$B^0 \to D^{*-}\mu^+\nu_\mu$$

	mean	σ
Signal	-0.025	0.787
$B\overline{B}$ *Background*	0.048	0.98
Continuum Background	-0.06	1.009

Table 6.2: Toy MC pull test results for each component.

(a) Electron mode

(b) Muon mode
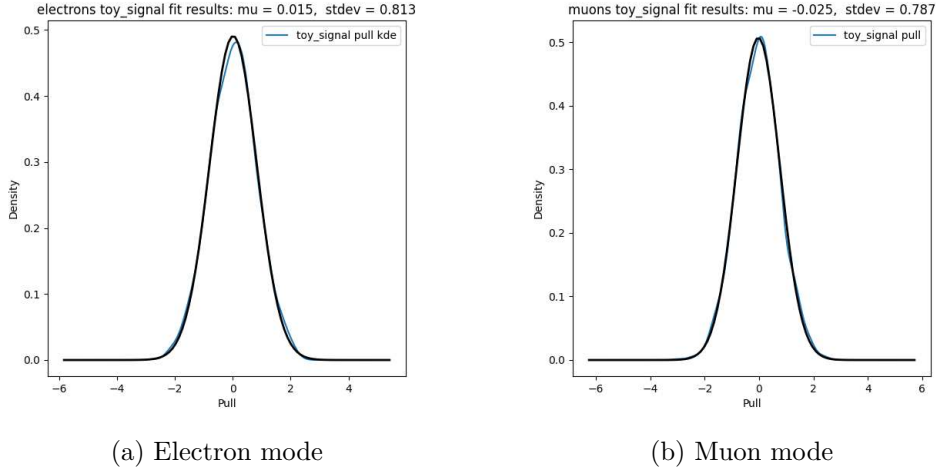
Figure 6.4: Unbinned gaussian fit (black curve) applied on a kernel density estimation (blue curve) of the distribution of 1500 toy MC pulls and the resulting mean and width of the fit for both decay modes.

6.3 Fit Results

After the application of the final selection criteria and restricting the region for the $\cos\theta$ distribution to $-4 < \cos\theta < 4$ the 34.6 fb^{-1} sample of real data yielded 15341 (16680) B meson candidates for the electron (muon) mode. E.g. for the signal components the fit estimated a fraction of 56.526% for the electron mode with an absolute (relative) statistical uncertainty of 0.936 % (1.656%) and for the muon mode it estimated a signal fraction of 56.624% with an absolute (relative) statistical uncertainty of 0.832% (1.469%). The resulting number of events for all component estimations, i.e. signal N_{sig}, continuum background N_{cont} and $B\overline{B}$ background $N_{B\overline{B}}$ after applying the fit is listed below:

- $B^0 \rightarrow D^{*-}e^+\nu_e$ yields: $N_{sig} = 8671 \pm 144$, $N_{B\overline{B}} = 5656 \pm 261$, $N_{cont} = 931 \pm 192$

- $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ yields: $N_{sig} = 9444 \pm 139$, $N_{B\overline{B}} = 5927 \pm 264$, $N_{cont} = 1308 \pm 225$

In Table 6.3 the estimated fraction of all three fitted components are summarized. In addition the expectation of the fractions based on the known source contributions of the MC sample is shown, as well as the absolute and relative statistical uncertainties estimated by the fitting algorithm.

The fits on the data sample resulting from applying the MC templates explained in Section 6.1 are depicted in Figure 6.5 for both decay modes. In the plots 6.5a and 6.5b the crosses in the plots indicate the amount of data events in the bins and the solid line indicates the shape resulting from the fit. The plots show that the fit shape and the data distribution are in good agreement. The postfit data MC agreement scaled with the weighted number of events instead of the luminosity is shown in 6.5c and 6.5d for the electron and muon mode, respectively.

$$B^0 \rightarrow D^{*-}e^+\nu_e$$

	MC expectation [%]	Fraction [%]	σ abs.[%]	σ rel.[%]
Signal	57.413	56.526	0.936	1.656
$B\overline{B}$ *Background*	36.233	36.871	1.705	4.624
Continuum	6.353	6.607	1.365	20.659

$$B^0 \rightarrow D^{*-}\mu^+\nu_\mu$$

	MC expectation [%]	Fraction [%]	σ abs.[%]	σ rel.[%]
Signal	56.452	56.624	0.832	1,469
$B\overline{B}$ *Background*	37.341	35.536	1.583	4.455
Continuum	6.206	7.842	1.347	17.177

Table 6.3: Fraction and uncertainty estimations resulting from the fit for each component and their respective expected fraction deduced from the MC simulated sample.

(a) Electron mode

(b) Muon mode

(c) Electron mode
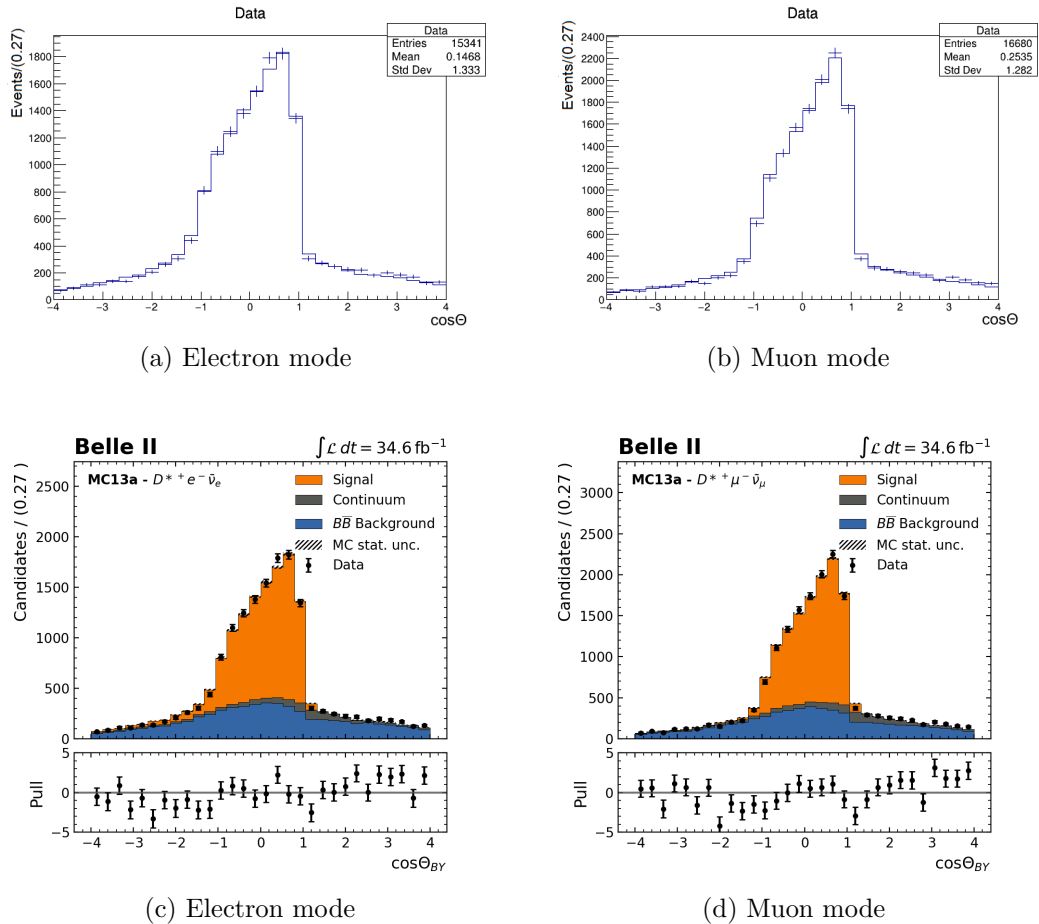
(d) Muon mode

Figure 6.5: In the upper figures the $\cos\theta$ distribution fit results of the electron and muon mode after applying the MC templates on the real data sample are shown. The data points are indicated by the crosses and the solid lines show the estimation of the fit. The lower plots show the postfit results. The pull in each bin is calculated using the number of data events, number of MC events and the square root of the number of data events as the uncertainty.

60

7 Results

7.1 Branching Fraction

To determine the branching fraction of the decay $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ Equation 7.1 is used. N_{sig} is the number of signal events estimated by the fit, N_{B^0} the number of B^0 mesons produced in the data sample, ϵ the efficiency of the whole reconstruction chain and the branching fractions of the subdecays $D^{*-} \rightarrow \overline{D}^0\pi^-$ and $\overline{D}^0 \rightarrow K^+\pi^-$.[4]

$$Br(B^0 \rightarrow \overline{D}^-\ell^+\nu_\ell) = \frac{N_{sig}}{N_{B^0} \times \epsilon \times Br(D^{*-} \rightarrow \overline{D}^0\pi^-) \times Br(\overline{D}^0 \rightarrow K^+\pi^-)} \quad (7.1)$$

The number of signal events was determined in Section 6.3 and yielded $N_{sig\_e} = 8671 \pm 144$ and $N_{sig\_\mu} = 9444 \pm 139$ for the electron and muon mode, respectively.

For the branching fractions of the subdecays $D^{*-} \rightarrow \overline{D}^0\pi^-$ and $\overline{D}^0 \rightarrow K^+\pi^-$ the world average values given by the particle data group in Equation 7.2 and 7.3 are used.[11]

$$Br(D^{*-} \rightarrow \overline{D}^0\pi^-) = (67.7 \pm 0.5)\% \quad (7.2)$$

$$Br(\overline{D}^0 \rightarrow K^+\pi^-) = (3.950 \pm 0.031)\% \quad (7.3)$$

The efficiency of the reconstruction chain can be determined by Equation 7.4, where N_{sigMC} is the number of signal events of the 200 fb^{-1} MC sample after the final signal selection requirements in a $\cos\theta$ region of [-4,4] with applied lepton ID correction weights and a flag for a successful HLT trigger. N_{count} is the number of signal events calculated from the branching fractions inside the decay file used to simulate the MC sample.

$$\epsilon = \frac{N_{sigMC}}{N_{count}} \quad (7.4)$$

The number of $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ decays in the MC simulated sample is estimated by Equation 7.5, where $N_{event} = 200000$ is the number of B^0 events in one MC file, $N_{file} = 510$ the number of files in the used 200 fb^{-1} MC batch that contain B^0 events and the branching fractions of the decay and $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ its sub decays given by the decay file.

$$N_{count} = N_{event} \times N_{file} \times \omega_B \times Br(B^0 \rightarrow D^{*-}\ell^+\nu_\ell) \times$$
$$\times \omega_{D^*} \times Br(D^{*-} \rightarrow \overline{D}^0\pi^-) \times \omega_D \times Br(\overline{D}^0 \rightarrow K^+\pi^-) \quad (7.5)$$

The value of N_{count} has to be renormalized, due to the branching fractions of the B^0, $\overline{B}^0$ and $\overline{D}^0$ not adding up to a total of 1. If the branching fractions don't add up to 1 the event generator creating the MC sample will normalize them automatically, which leads to a deviation of the number of events counted with the branching fractions of the decay file from the real amount. This requires the additional weights ω attached to the branching fractions in Equation 7.5.

The number of counted $B^0 \to D^{*-}\ell^+\nu_\ell$ decays inside the 200 fb^{-1} MC batch is determined to be $N_{count} = 272549$. The number of reconstructed signal events is $N_{sigMC} = 52301(57524)$ for the electron (muon) mode and therefore the calculated efficiencies of the reconstruction chain are:

$$\epsilon_e = 19.189\%, \;\; \epsilon_\mu = 21.106\% \tag{7.6}$$

The estimated number of signal events N_{count} in the MC batch was cross checked using a second method, i.e. the reconstructMCDecay() function from the basf2 framework. Using only direct decays the function reconstructed 272102 signal events in the electron mode and 272380 events in the muon mode. This leads to efficiencies of 19.221% (21.119%) for the electron (muon) mode, which is a relative discrepancy of only 0.167% (0.062%). Therefore both approaches are valid and show only negligible differences, but for this thesis the results from the previously discussed decay file method were chosen.

The number of B^0 and $\overline{B}^0$ mesons in the data samples was estimated using Equation 7.7. $N_{Proc11} = (9.56 \pm 0.02(stat) \pm 0.15(syst)) \times 10^6$ is the number of B mesons produced in the procc11 sample and $N_{Prompt} = (28.06 \pm 0.05(stat) \pm 0.45(syst)) \times 10^6$ is the number of B mesons produced in buckets $9 + 10 + 11$. N_{Proc11} and N_{Prompt} are both provided by the B meson counting group. f^{00} defined in 7.8 is the ratio of B^0 mesons produced from $\Upsilon(4\mathcal{S})$ decays. The factor 2 is multiplied to account for the charge conjugated $\overline{B}^0$ mesons. This leads to 36674532 produced B^0 and $\overline{B}^0$ mesons.[[4], [48]]

$$N_{B^0} = (N_{Proc11} + N_{Prompt}) \times 2 \times f^{00} = 36566640 \tag{7.7}$$

$$f^{00} = \frac{\Gamma(\Upsilon(4\mathcal{S}) \to B^0\overline{B}^0)}{\Gamma_{tot}(\Upsilon(4\mathcal{S}))} = 0.486 \pm 0.006 \tag{7.8}$$

All obtained quantities are summarized in Table 7.1. The branching fraction of the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ is determined for each lepton decay mode separately using Equation 7.1. This leads to the branching fractions in Equation 7.9 and 7.10. These branching fractions include only the statistical uncertainties, which result from the fitting procedure. The systematic uncertainties will be discussed in Section 7.2.

$$Br(B^0 \to \overline{D}^- e^+\nu_e) = (4.669 \pm 0.078(stat))\% \tag{7.9}$$

$$Br(B^0 \to \overline{D}^- \mu^+\nu_\mu) = (4.576 \pm 0.067(stat))\% \tag{7.10}$$

	$B^0 \rightarrow D^{*-}e^+\nu_e$	$B^0 \rightarrow D^{*-}\mu^+\nu_\mu$
N_{sig}	8671 ± 144	9444 ± 139
N_{B^0}	36566640	36566640
$Br(D^{*-} \rightarrow \overline{D}^0\pi^-)$	$(67.7 \pm 0.5)\%$	$(67.7 \pm 0.5)\%$
$Br(\overline{D}^0 \rightarrow K^+\pi^-)$	$(3.95 \pm 0.031)\%$	$(3.95 \pm 0.031)\%$
ϵ	19.189%	21.106%

Table 7.1: Summary of the quantities used to determine the branching fractions.

7.2 Systematic Uncertainties

In this section all systematic uncertainties that were considered for the branching fraction result are discussed. A summary of all relative systematics for the electron and muon mode is given in Table 7.3.

Tracking

Using $e^-e^+ \rightarrow \tau^-\tau^+$ decays a study at Belle II measured the track reconstruction efficiency and recommended an uncertainty of 0.8 % per track for leptons, kaons and pions. For the slow pions of the D^* an additional independent study was done. Since this uncertainty is multiplicative it leads to an additional relative, systematic uncertainty of **2.4%** for three tracks.[49]

Slow Pion Tracking Efficiency

A tracking efficiency uncertainty of the slow pion of **9.87%** per track is recommended by the physics performance group. This recommendation is based on a study of $B^0 \rightarrow D^*\pi$ and $B^0 \rightarrow D^*\rho$ decays.[50]

MC statistics of efficiency

The efficiencies ϵ were estimated to be 19.189% and 21.106 % for the electron and muon mode. The uncertainties of these estimations were calculated by means of an binominal error leading to relative uncertainties of **0.075%(0.078%)** for the electron (muon) mode.

Charm Branching Fractions

The branching fractions of the decays $Br(D^{*-} \rightarrow \overline{D}^0\pi^-) = (67.7 \pm 0.5)\%$ and $Br(\overline{D}^0 \rightarrow K^+\pi^-) = (3.95 \pm 0.031)\%$ provided by the particle data group have relative uncertainties of 0.74% and 0.78%, respectively. The sum of the squared uncertainties is **1.1%**.[11]

Number of B mesons

All quantities used to determine the number of B^0 mesons produced in the sample of real data come with statistical and systematic uncertainties. The quantities used to calculate N_{B^0} are: $N_{Proc11} = (9.56 \pm 0.02(stat.) \pm$

$0.15(syst.)) \times 10^6$, $N_{Prompt} = (28.06 \pm 0.05(stat.) \pm 0.45(syst.)) \times 10^6$ and $f^{00} = 0.486 \pm 0.006$. The relative systematic uncertainty of the N_{B^0} was estimated by adding all statistical and systematic, relative uncertainties of the mentioned quantities in quadrature. This leads to an relative systematic uncertainty of **2.58%**.

Lepton ID

The systematic uncertainty resulting from the application of the lepton ID corrections is estimated by generating 500 variations of the lepton ID weights for each (p,θ) bin. The systematics of the bins are assumed to be 100% correlated, while the statistical uncertainties are assumed to be independent from each other. The variation of the original weight from the table in a given (p,θ) bin is created based on the original weight, a random drawn Gauss number based on σ_{sys} of the bin entry in the table and a random drawn Gauss number based on σ_{stat} of the bin. This leads to 500 different distributions of N_{sigMC}. To estimate the relative uncertainty the weighted number of entries of the 500 N_{sigMC} distributions are summed up and their standard deviation σ_{var} is calculated. Dividing $\sigma_{var}s$ by the weighted number of events of the original N_{sigMC} distribution one yields the relative uncertainties of **0.42%** and **1.94%** for the electron and muon mode, respectively.

Form Factors

The form factors (see 2.5) used by the event generator to create $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ events were studied by Chaoyi Lyu from Bonn university, who also analyses the decay $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$. The form factor uncertainty originates from the uncertainties of the parameters measured to calculate the form factor. The uncertainty is estimated by varying the form factor parameters by the given uncertainties. He estimated an relative uncertainty of **1.1%** for both decay modes. [51]

Binning

The relative systematic uncertainty that arises from the binning of the events during the fit was estimated by varying the number of bins by ± 10 and repeating the fit. The relative deviation from the chosen number of 30 bins are listed in Table 7.2. The relative deviations were then added in quadrature to estimate the final binning systematic.

	20 bins	40 bins	σ_{bins}
$B^0 \rightarrow D^{*-}e^+\nu_e$	0.24%	0.02%	0.24%
$B^0 \rightarrow D^{*-}\mu^+\nu_\mu$	0.41 %	0.19%	0.45%

Table 7.2: Binning systematic result.

Branching Fractions

The resulting branching fractions for both decay modes with all systematic

64

uncertainties from Table 7.3 included are:

$$Br(B^0 \to \overline{D}^- e^+ \nu_e) = (4.669 \pm 0.078(stat) \pm 0.495(sys))\% \qquad (7.11)$$

$$Br(B^0 \to \overline{D}^- \mu^+ \nu_\mu) = (4.576 \pm 0.067(stat) \pm 0.494(sys))\% \qquad (7.12)$$

Additionally the combined branching fractions of both decay modes was determined by calculating the uncertainty weighted average of the electron and muon mode. The systematic uncertainty of the combined result was estimated by using the higher uncertainty of the muon decay mode in Table 7.3:

$$Br(B^0 \to \overline{D}^- \ell^+ \nu_\ell) = (4.616 \pm 0.051(stat) \pm 0.499(sys))\% \qquad (7.13)$$

The current world average branching fraction of the decay $B^0 \to D^{*-}\ell\nu_\ell = (5.05 \pm 0.14)\%$ lies within $1\ \sigma$ of both measurements after the inclusion of systematics. [11]

Furthermore the lepton universality of the weak interaction was probed by calculating the ratio of the electron and muon branching fractions:

$$R_{e\mu} = \frac{Br(B^0 \to \overline{D}^- e^+ \nu_e)}{Br(B^0 \to \overline{D}^- \mu^+ \nu_\mu)} = 1.020 \pm 0.023(stat) \pm 0.106(sys) \qquad (7.14)$$

The obtained ratio in Equation 7.14 is predicted to be unity by the Standard Model of particle physics, which lies within $1\ \sigma_{stat}$ of the result of this analysis. The systematic uncertainty was conservatively estimated by taking the values used for the combined uncertainty and removing only the N_{B^0} and the charm branching fraction uncertainties. It yielded a relative systematic uncertainty of 10.4%.

Source	$B^0 \to D^{*-}e^+\nu_e$	$B^0 \to D^{*-}\mu^+\nu_\mu$	$B^0 \to D^{*-}\ell^+\nu_\ell$
K, π, ℓ *Tracking*	2.4%	2.4%	2.4%
π_s *Tracking*	9.87%	9.87%	9.87%
MC Statistics	0.075%	0.078%	0.078%
Charm Br	1.1%	1.1%	1.1%
N_{B^0}	2.58%	2.58%	2.58%
Lepton ID	0.42%	1.94%	1.94%
Form Factor	1.1%	1.1%	1.1%
Binning	0.24%	0.45%	0.45%
Total	10.6 %	10.8%	10.8%

Table 7.3: Summary of the relative systematics and the resulting total relative systematic uncertainty for both decay modes.

7.3 Cross Checks

To validate the analysis procedure the final branching fraction result of both $Br(B^0 \to D^{*-}\ell^+\nu_\ell)$ decay modes as well as the determined values of ϵ, N_{B^0} and N_{sig} were cross checked with an independent, untagged exclusive analysis done by Chaoyi Lyu from Bonn University, who analyses the same decay channel. [51]

The compared quantities are summarized in Table 7.4 and additionally all comparisons will be discussed below:

- **N_{sig}**: The amount of signal in the real data sample for the electron mode estimated by the fit is $N_{sig.e} = 9583 \pm 134$ for Chaoyi's analysis and $N_{sig.e} = 8671 \pm 144$ for this thesis. The difference of the amount of estimated signal events arises from a few sources, i.e. both analyses use different final selection requirements, Chaoyi's analysis used a different fitting algorithm (namely the Asimov fit, while in my analysis the TFractionFitter was used) and additionally Chaoyi used the basf2 release light-2002-ichep and its Bremsstrahlung recovery function for the electrons, while I did not use any kind of Bremsstrahlung recovery. The difference in the number of signal events for the muon mode, i.e. $N_{sig.\mu} = 9860 \pm 132$ (Chaoyi) and $N_{sig.\mu} = 9444 \pm 139$ (me), arises from the fitting algorithm and the different signal selection requirements. Although there are a few discrepancies in the amount of signal for both modes there is a good agreement between the branching fraction from both analyses.

- **ϵ**: The efficiencies of the reconstruction chain were determined using two different techniques. In Chaoyi's analysis ϵ was determined by using MC samples containing only $B \to D^*\ell\nu_\ell$ events of which the number is known. Using this method he estimated an efficiency of 21.3% and 21.8% for the electron and muon mode, respectively. The decay file method used to estimate the efficiency for this thesis yielded 19.189% for the electron mode and 21.106% for the muon mode. The source of the discrepancies are the different methods used to determine ϵ, the different signal selection and the Bremsstrahlung recovery for the electrons.

- **N_{B^0}**: While both analyses estimated the amount of B^0 mesons based on the estimate performed by the B meson counting group, slightly different formulas were used to calculate the number of B^0 mesons. I.e. $N_{B^0} = 2 \times N_{B\overline{B}} \times (1 + f^{+0})^{-1}$ in Chaoyi's analysis, where f^{+0} is the ratio of B^+ to B^0 mesons in comparison to Equation 7.7 of this thesis. The two methods yielded $N_{B^0} = 36648202$ for Chaoyi and $N_{B^0} = 36566640$ for this thesis.

- **_Branching fractions_**: Despite the discrepancies for the individual quantities it can be seen in Table 7.4 that the branching fractions of both individual analyses converge well for both decay modes. The branching fractions estimated by Chaoyi's analysis, the determined ratio $R_{e\mu}$ and the combined branching fraction of both modes $Br(B^0 \to \overline{D}^-\ell^+\nu_\ell)$ are within the range of the statistical uncertainty of this thesis. Therefore it can be concluded that in neither of the independent analyses any major mishaps happened.

$$B^0 \to D^{*-}e^+\nu_e$$

	Chaoyi	Daniel
N_{sig}	9583 ± 134	8671 ± 144
ϵ	21.3 %	19.189%
N_{B^0}	36648202	36566640
Br_e	$(4.59 \pm 0.06(stat) \pm 0.48(sys))\%$	$(4.669 \pm 0.078(stat) \pm 0.495(sys))$ %

$$B^0 \to D^{*-}\mu^+\nu_\mu$$

	Chaoyi	Daniel
N_{sig}	9860 ± 132	9444 ± 139
ϵ	21.8 %	21.106%
N_{B^0}	36648202	36566640
Br_μ	$(4.62 \pm 0.06(stat) \pm 0.49(sys))\%$	$(4.576 \pm 0.067(stat) \pm 0.494(sys))\%$

$$B^0 \to D^{*-}\ell^+\nu_\ell$$

	Chaoyi	Daniel
Br_ℓ	$(4.60 \pm 0.05(stat) \pm 0.17(sys) \pm 0.45_{\pi_s})\%$	$(4.616 \pm 0.051(stat) \pm 0.499(sys))\%$

$$Ratio\ Br_e/Br_\mu$$

	Chaoyi	Daniel
$R_{e\mu}$	(0.99 ± 0.03)	$(1.020 \pm 0.023(stat) \pm 0.106(sys))$

Table 7.4: Summary of the quantities used to determine the branching fractions. Chaoyi's combined branching fraction singles out the systematic uncertainty originating from the slow pion tracking efficiency from the other systematics. [51]

8 Conclusion

An untagged, exclusive analysis of the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ ($\ell = e, \mu$) was done. For this analysis a data sample with an integrated luminosity of 34.6 fb^{-1} obtained during the first years of data taking at the Belle II experiment was used.

The decay $B^0 \to D^{*-}\ell^+\nu_\ell$ ($\ell = e, \mu$) was reconstructed via the sub-decays $D^{*-} \to \overline{D}^0\pi^-$ and $\overline{D}^0 \to K^+\pi^-$. The signal selection from the reconstructed B meson candidates was optimized using MC simulated samples. During this optimization the constraints applied on the selected B meson candidates were tightened either based on the variable distributions or by optimizing the figure of merit. if variable distributions could not clearly separate signal and background. The variables used for the signal optimization included: The Fox-Wolfram moment ratio R_2, the lepton momentum in the center of mass frame p_ℓ^*, the lepton identification probability and the mass difference Δm_D between the $\overline{D}^0$ and D^{*-} candidate. Of these quantities the tightening of the constraints on Δm_D and p_ℓ^* were the most impactful.

To ensure that a fit based on the MC simulated samples describes the data set well the agreement between a MC simulated sample with an integrated luminosity of 200 fb^{-1} and the data was checked. For this agreement check the size of the MC sample was scaled down to the size of the data sample based on their luminosity. The comparison of various distributions showed that there were minor disagreements in the electron mode and bigger disagreements in the muon mode. These deviations were corrected by applying corrections to the slightly mismodelled lepton identification efficiency and misidentification probabilities of the MC. After applying these corrections as an additional weight to the scaling the distributions of the MC and data samples showed good agreement for both decay modes.

To extract the signal from the data a maximum likelihood fit by R. Barlow and C. Beeston [45] was used. The fit was created based on the shapes of three MC templates: signal, $B\overline{B}$ background and continuum background. The fitting procedure was validated by means of a GSIM and toy MC pulltest. The testing of the fit showed no bias in the estimated number events of the components or underestimation of the uncertainty estimated by fit.

After the extraction of the signal events N_{sig} from the data the remaining two variables needed for the determination of the branching fraction were calculated, i.e. the efficiency of the reconstruction chain ϵ and the number of B^0 and $\overline{B}^0$ events N_{B^0}. For the branching fractions of the sub-decays $D^{*-} \to \overline{D}^0\pi^-$ and $\overline{D}^0 \to K^+\pi^-$ needed for the determination of the branching fraction of $B^0 \to D^{*-}\ell^+\nu_\ell$ the world average values provided by the Particle Data Group were taken. After calculating the branching fraction the total relative systematic uncertainty of the analysis chain was determined. The analysis resulted in the following branching fractions for the electron and muon mode:

$$Br(B^0 \to \overline{D}^- e^+\nu_e) = (4.669 \pm 0.078(stat) \pm 0.495(sys))\% \qquad (8.1)$$

$$Br(B^0 \to \overline{D}^- \mu^+\nu_\mu) = (4.576 \pm 0.067(stat) \pm 0.494(sys))\% \qquad (8.2)$$

Additionally the ratio of the branching fraction of the electron decay mode to the muon decay mode was calculated to probe the lepton universality of the weak interaction. Furthermore the combined branching fraction of both decay modes was determined by calculating the uncertainty weighted mean. The following results for the ratio and combined branching fraction were obtained:

$$Br(B^0 \to \overline{D}^- \ell^+ \nu_\ell) = (4.616 \pm 0.051(stat) \pm 0.499(sys))\% \qquad (8.3)$$

$$R_{e\mu} = \frac{Br(B^0 \to \overline{D}^- e^+ \nu_e)}{Br(B^0 \to \overline{D}^- \mu^+ \nu_\mu)} = 1.020 \pm 0.023(stat) \pm 0.106(sys) \qquad (8.4)$$

For a validation the results were compared to the current world average branching fraction of the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ provided by the Particle Data Group. It is $(5.05\pm0.14)\%$, which lies within the range of $1\,\sigma_{tot}$ of the branching fraction measurements of the analysis done for this thesis. Unity lies within 1 σ_{stat} of the ratio $R_{e\mu}$ as predicted by the Standard Model of particle physics.

A second validation of the analysis process was done by comparing the branching fraction measurements, the ratio and the quantities N_{sig}, N_{B^0} and ϵ with the independent, untagged exclusive $B^0 \to D^{*-}\ell^+\nu_\ell$ measurement of the same Belle II data sample by Chaoyi Lyu from Bonn University. While both analyses used different techniques and therefore measurements of single quantities deviated from each other the branching fractions agree within 1 σ_{stat} of the branching fraction measurements of this thesis. Thus the cross check further validated the analysis of this thesis.

While the measurements of this analysis are not yet competitive due to the small sample size of 34.6 fb^{-1} obtained since the operation started in 2018, it shows that the experimental setup and the software framework implemented at Belle II are already in good working condition. However, hardware and software improvements are continuously being implemented.

With the instantaneous luminosity of Belle II steadily increasing it is estimated that in three years an integrated luminosity of approximately 1 ab^{-1} will be reached. This will lead Belle II to the world first measurements of the CKM matrix elements $|V_{cb}|$ and $|V_{ub}|$ with a data sample of an integrated luminosity of 1 ab^{-1}. At that point this analysis can be repeated and extended to also include the $|V_{cb}|$ measurement.

References

[1] Mark Thomson. *Modern particle physics.* Cambridge University Press, New York, 4th edition, 2019.

[2] T.Browder, A.Sekiya. Belle results lead to Nobel Prize for Kobayashi and Maskawa . `https://belle.kek.jp/belle/km_nobel/index.html`. Accessed: 2020-09-18.

[3] A. J. Bevan, B. Golob, Th. Mannel, S. Prell, B. D. Yabsley, H. Aihara, F. Anulli, N. Arnaud, T. Aushev, M. Beneke, and et al. The Physics of the B Factories. *The European Physical Journal C*, 74(11), Nov 2014.

[4] Heavy Flavor Averaging Group and Y. Amhis et.al. Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018, 2019.

[5] Emi Kou, Phillip Urquijo, B2TiP Theory community, and Belle II Collaboration. The Belle II Physics Book. Aug 2018.

[6] Y. Suetsugu. 2020c run plan and MEXT SuperKEKBRoadmap2020. `https://indico.belle2.org/event/2419/contributions/11671/attachments/6003/9374/2020c_run_plan_MEXT_Roadmap2020_suetsugu_13.pdf`, Jun 2020. Accessed: 2020-08-10.

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

[8] Public Domain Image. Standard model of elementary particles. `https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg`, 2020. Accessed: 2020-07-20.

[9] Fernando Abudinén, Christian Kiesling, and Thomas Kuhr. *Development of a B^0 flavor tagger and performance study of a novel time-dependent CP analysis of the decay $B^0 -> \pi^0\pi^0$ at Belle II.* PhD thesis, Munich, Ludwig-Maximilians-Universität, Munich, 2018. Presented on 19 10 2018.

[10] Stephen Godfrey and Kenneth Moats. Spectroscopic assignments of the excited B-mesons. *The European Physical Journal A*, 55(5), May 2019.

[11] P.A. Zyla et al. (Particle Data Group). 2020 Review of Particle Physics. *to be published in Prog. Theor. Exp. Phys.*, 2020. 083C01 (2020).

[12] Alexei Sibidanov. B to $\mu\nu$ at Belle/Belle II. Sep 2018. 20 min.

[13] W. Dungel, C. Schwanda, I. Adachi, H. Aihara, T. Aushev, T. Aziz, A. M. Bakich, V. Balagura, E. Barberio, M. Bischofberger, and et al. Measurement of the form factors of the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ and determination of the CKM matrix element $|Vcb|$. *Physical Review D*, 82(11), Dec 2010.

[14] Akai, Kazunori and Furukawa, Kazuro and Koiso, Haruyo. SuperKEKB collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 907:188–199, Nov 2018.

[15] D. Brandt, editor. *Intermediate accelerator physics. Proceedings, CERN Accelerator School, Zeuthen, Germany, September 15-26, 2003*, CERN Yellow Reports: School Proceedings, 4 2006.

[16] Z. Dolezal and S. Uno. Belle II Technical Design Report. Oct 2010.

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

[18] Yukiyoshi Ohnishi et al. Accelerator design at SuperKEKB. *PTEP*, 2013:03A011, 2013.

[19] Claude Amsler et al. Review of Particle Physics. *Phys. Lett. B*, 667:1–1340, 2008.

[20] Hiroyuki Nakayama (KEK). SuperKEKB background simulations. `https://slideplayer.com/slide/7267437/`. Accessed: 2020-08-21.

[21] High Energy Accelerator Research Organization, Institute of Particle and Nuclear Studie. Belle II Detector 3D model. `https://www.belle2.org/archives/`. Accessed: 2020-08-11.

[22] Sara Pohl and Christian Kiesling. *Track Reconstruction at the First Level Trigger of the Belle II Experiment.* PhD thesis, Munich, Ludwig-Maximilians-Universität, Munich, 2018. Presented on 11 04 2018.

[23] F. Simon and K. Ackermann and L. Andricek and C. Heller and C. Kiesling and A. Moll and H-G. Moser and J. Ninkovic and K. Prothmann and B. Reisert and R. Richter and M. Ritter and S. Rummel and A Wassatsch. The Belle-II Pixel Vertex Detector at the SuperKEKB Flavor Factory, 2010.

[24] Christian Wessel. THE BELLE II PIXEL DETECTOR (PXD). `https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5651/B2SKW_PXD.pdf`, Jan 2020. Accessed: 2020-08-10.

[25] Luigi Corona. The Belle II Silicon Vertex Detector (SVD). `https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5651/B2SKW_PXD.pdf`, Jan 2020. Accessed: 2020-08-10.

[26] Henrikas Svidras. The Central Drift Chamber of Belle II. `https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5652/CDC_B2SK.pdf`, Jan 2020. Accessed: 2020-08-10.

[27] M. Akatsu, M. Aoki, K. Fujimoto, Y. Higashino, M. Hirose, K. Inami, A. Ishikawa, T. Matsumoto, K. Misono, I. Nagai, and et al. Time-of-propagation Cherenkov counter for particle identification. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 440(1):124–135, Jan 2000.

[28] Oskar Hartbrich. Particle ID in Belle II ARICH and TOP. `https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5668/2020_02_01_B2SKW_PID.pdf`, Jan 2020. Accessed: 2020-08-10.

[29] Rok Pestotnik. Particle identification devices at the Belle II experiment. Oct 2015.

[30] Arpad Horvath. The geometry of the Cherenkov radiation. `https://commons.wikimedia.org/wiki/File:Cherenkov.svg`. Accessed: 2020-08-11.

[31] Rok Pestotnik. Calibration of the Belle II Aerogel Ring Imaging Detector. Oct 2018. submitted to NIMA.

[32] Marco Milesi. Particle identification at Belle II with the Electromagnetic Calorimeter (ECL). `https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5670/mmilesi_ECLPid_B2SKW_2020.pdf`, Jan 2020. Accessed: 2020-08-10.

[33] V. Aulchenko et al. Electromagnetic calorimeter for Belle II. *Journal of Physics: Conference Series*, 587, 02 2015.

[34] Alberto Martini. KLM overview and μ identification algorithm. `https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5670/mmilesi_ECLPid_B2SKW_2020.pdf`, Jan 2020. Accessed: 2020-08-10.

[35] Iwasaki, Yoshihito and Cheon, ByungGu and Won, Eunil and Gao, Xin and Macchiarulo, Luca and Nishimura, Kurtis and Varner, Gary. Level 1 trigger system for the Belle II experiment. *IEEE Trans. Nucl. Sci.*, 58:1807–1815, 2011.

[36] Valerio Bertacchi et.al. Track Finding at Belle II. Mar 2020.

[37] Johannes Rauch, Tobias Schlüter. GENFIT and its Application for the Belle II Experiment. `https://indico.cern.ch/event/459865/contributions/1961746/attachments/1199055/1744107/GenfitTrackRecoForum.pdf`. Accessed: 2020-08-13.

[38] Hikaru Hirata. *Evaluation of PID performance and mass resolution of Belle II detector for hadron spectroscopy (in Japanese).* PhD thesis, Nagoya, Nagoya university, Nagoya, 2019. Presented on 14 02 2019.

[39] ROOT Team 2020. ROOT Cern. `https://root.cern/`. Accessed: 2020-08-20.

[40] S. Agostinelli et.al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

[41] G. Ciezarek, A. Lupato, M. Rotondo, and M. Vesterinen. Reconstruction of semileptonically decaying beauty hadrons produced in high energy pp collisions. *Journal of High Energy Physics*, 2017(2), Feb 2017.

[42] 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 $|V_{cb}|$. *Phys. Rev. D*, 93(3):032006, 2016.

[43] Racha Cheaib. First results on V_{ub} and V_{cb} with Belle II. https://indico.cern.ch/event/868940/contributions/3815659/attachments/2081582/3497706/ICHEPBelleII_VubVcb_Cheaib_FInal.pdf. Accessed: 2020-08-21.

[44] Fox, Geoffrey C. and Wolfram, Stephen. Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes. *Phys. Rev. Lett.*, 41:1581–1585, Dec 1978.

[45] Roger J. Barlow and Christine Beeston. Fitting using finite Monte Carlo samples. *Comput. Phys. Commun.*, 77:219–228, 1993.

[46] Luc Demortier1 and Louis Lyons2. Everything you always wanted to know about pulls. http://physics.rockefeller.edu/luc/technical_reports/cdf5776_pulls.pdf. Accessed: 2020-08-21.

[47] ROOT Team 2020. TFractionFitter. https://root.cern/doc/master/classTFractionFitter.html. Accessed: 2020-09-08.

[48] The Belle II Collaboration. Approved B counting analysis results with ICHEP2020 dataset. Jul 2020.

[49] Alexander Glazov and Petar Rados and Ami Rostomyan and Eugenio Paoloni and Laura Zani. Measurement of the tracking efficiency in Phase 3 data using tau-pair events. Feb 2020.

[50] S. Maity; and S. Patra; and S. Bahinipati; and V. Bhardwaj; and G. B. Mohanty. Study of slow-pion tracking efficiency using $B^0 -> D^*\pi$ and $B^0 -> D^*\rho$ decays. Jun 2020.

[51] Florian Bernlochner and Lu Cao and Daniel Dorner and Philipp Horak and Chaoyi Lyu and Minakshi Nayak and Christoph Schwanda and Abi Soffer. Studies of the semileptonic $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ and $B^- \to D^0\ell^-\bar{\nu}_\ell$ decay processes with 34.6 fb^{-1} of Belle II data. Aug 2020.